Theor Chem Acc (2004) 112: 240–246
DOI 10.1007/s00214-004-0583-2

Theoretical
Chemistry Accounts

Regular article

The costly process of creating a cavity in *n*-octanol

Siegfried Höfinger[1], **Francesco Zerbetto**[2]

[1]Novartis Forschungsinstitut, Brunnerstrasse 59, 1235 Vienna, Austria
[2]Dipartimento di Chimica "G. Ciamician", Via F. Selmi 2, Università di Bologna, 40126 Bologna, Italy

Received: 17 July 2003 / Accepted: 18 November 2003 / Published online: 29 April 2004

Abstract. Free-energy perturbation calculations are used to evaluate the free energy of cavity formation in *n*-octanol. A detailed theoretical analysis of the procedure is given and some limiting value phenomena are discussed. The data become subject to a three-parameter fit and a revised formulation of the popular approach due to Pierotti of calculating cavitation free energies is given. Pierotti's approach is based on the equation derived from scaled particle theory (SPT) by Reiss et al. [(2000) *J. Chem. Phys.* 31:369–380]. The revision of Pierotti's approach has the important advantage of being completely independent of the solvent hard-sphere radius, an empirical parameter in the standard procedure, which is hard to define in a uniformly valid way.

Keywords: Cavitation free energy – Pierotti's approach – Free energy – Perturbation – log *P* – *n*-octanol

1 Introduction

Embedding a molecule in a solvent may be formally considered a two-step process. The initial step creates an empty space inside the solvent, which then shall be occupied by the molecule in question – the solute. The second step consists of switching on all the intermolecular interactions that occur at the solute/solvent boundary. While basic physical principles may be applied in solving for the second part [1], the initial process of creating a void of appropriate shape and size to accommodate the solute still poses a major challenge to the theoretical as well as to the practicing physical chemist of today. In this context the enormous advantage coming along with bare theoretical approaches must be highlighted again, since only these have – at least in principle – the power to describe such a rather artificial scenario properly. Any experimental approach would most likely fail, because this molecular volume – the cavity – really needs to be perfectly empty, but experimentally observed cavities, for example, from surface tension data, are not. Indeed, it has long been realized that, for example, the cavitation free energy in water could successfully be obtained from employing one such barely theoretical technique, namely free-energy perturbation (FEP) [2] in combination with molecular dynamics (MD) simulation which nowadays appears to be a standard methodology [3]. Recently the authors of the present article have re-evaluated the MD/FEP approach for the calculation of free energies required to create perfectly empty cavities in various solvents [4, 5]. The main findings of these studies may be summarized as follows:

- Little or practically no dependence on the force field parameters.
- Discouraging disagreement with the analytical model of Pierotti [6], especially for the standard set of solvent hard-sphere radii used, for example, in a standard quantum chemistry program [7]. However, rather close similarity between the MD/FEP results with surface tension data [8] was observed.
- A transformation property was identified when converting any arbitrary molecular volume to a perfectly spherical one and employing the MD/FEP data for the latter. Thus, with the help of this transformation property the rather theoretical process, which was originally reserved for the creation of an empty sphere inside a solvent only, immediately becomes suitable for real-world solvation problems too, which usually deal with molecular shapes that are far from perfectly spherical.
- The MD/FEP data may be interpolated with a polynomial expansion of up to second-order terms and thus gives rise to a "revision of Pierotti's approach" (rPA) which reads

$$\Delta G_{\mathrm{cav}} = k_0 + (k_1 B) + (k_2 B^2), \tag{1}$$

Proceedings of the 11th International Congress of Quantum Chemistry satellite meeting in honor of Jean-Louis Rivail

Correspondence to: S. Höfinger
e-mail: siegfried.hoefinger@pharma.novartis.com

with B being the effective radius of the volume-equivalent representative sphere of the molecule and k_0, k_1 and k_2 data tabulated in Ref. [5].

Previous work on the important issue of hydrophobic solvation has been carried out by Pohorille and Pratt [9, 10], Hummer et al. [11], Ashbaugh and Paulaitis [12] and Floris et al. [13], to mention just a few names.

One of the simplest ways to calculate ΔG_{cav} energies has been devised by Pohorille and Pratt [9, 10], where the focus was on evaluating the statistical occurrence of transient cavities – of limited maximum size – that are obtained during plain MD or Monte Carlo (MC) simulations. The simulation data formed the basis for a probability distribution of solvent molecules to enter a certain volume. Thus, the probability of finding exactly zero solvent molecules in a certain volume (transient cavity) could be related to the chemical potential of a hard-sphere solute dissolved in the solvent. While limited to rather small cavities, it must be noted that this approach is heavily dependent on the particular choice of model parameters used for the description of the solvent (e.g., compare the two water models employed in Ref. [9]; Figs. 3, 6).

An extension towards larger cavities was pioneered by the observation that the probability distribution of having exactly n solvent molecules in a confined cavity volume forms a quasi-parabolic function, which, surprisingly, may be further parameterized from nonphysical information theory models with a clear preference for a very simple two-moment flat-model [11]. The extrapolation to the zero-solvent-molecule probability then again connects the simulation data with actual cavitation free energies. However, as already mentioned, the strong intrinsic dependence on the absolute microscopic validity of model parameters adjusted to reproduce solvent properties in the macroscopic domain remains the critical issue here.

Another interesting distinction concerning dewetting and rewetting in the first hydration shell of solutes has been described by Ashbaugh and Paulaitits [12]. For extension towards nonaqueous cavitation free-energy calculations and alternative semiexperimental approaches see Refs. [14, 15, 16] and references therein.

Free-energy calculations have already been recommended by Floris et al. [13] to provide one of the most accurate ways of estimating the cavitation energy. In that study MC smulations were used and cavities up to a hard-sphere radius of 5 Å were analyzed. Despite the finding of a convincing correlation between scaled particle theory (SPT) and FEP results, the major critical dependences were identified to be the choice of an appropriate hard-sphere radius for the solvent and the definition of the relationship between soft-sphere radii and the corresponding hard-sphere values.

In the present article, we restrict ourselves to the report of how to accurately compute the cavitation free energy in n-octanol, which forms a major contribution to the net solvation free energy. In the next section, we briefly recall basic theoretical concepts. Then a detailed description of computational parameters is given. We present the results in the following section and discuss them in detail before summarizing the main findings in the conclusion.

2 Theory

In order to account for the effect of creating a perfectly empty cavity inside an "equilibrated structure of n-octanol" a repulsive potential (V_{rep}) of the type

$$V_{\mathrm{rep}} = \lambda \left(\frac{B^*}{r} \right)^{12} \tag{2}$$

was introduced at the origin of the simulation box that contained the n-octanol molecules. The parameter B^* in Eq. (2) defines the maximum radius of the final repulsive sphere and r is the distance between any of the atoms of all the molecules of n-octanol and the origin of the repulsive cavity. It is important to note that B^* in this form is given in units of $(\mathrm{kcal/mol})^{\frac{1}{12}}$Å and thus facilitates the connection between a repulsive radius and the energy penalty due to repulsion. In analogy to Postma et.al. [2], however, we prefer to switch to a description of cavity dimensions that does not involve any units of energy and thus allows us to envision the cavity as a pure spatial construct. Therefore for the remainder of the paper we shall make use of the reduced quantity "thermal cavity radius", which we shall refer to as B, given in reduced units of angstroms, where at a certain cavity dimension the corresponding "thermal radius" gives a repulsive action of exactly 1.0 kcal/mol. In a way the reduced quantities B may be considered as the cavity's boundaries at which an intruding particle would actually start to "feel" the acting repulsive potential and owing to the steepness of the potential this effect would grow rather rapidly the further the particle enters the cavity's space. Furthermore the variable λ in Eq. (2) is a dimensionless scalar parameter in the range $0 \leq \lambda \leq 1$ characteristic to all free energy calculations. It smoothly introduces the perturbation as it is modified from $\lambda = 0$ (unperturbed state) to $\lambda = 1$ (fully perturbed state). At some particular intermediate value of λ_i, a corresponding free-energy change due to a small perturbation $\lambda_i + \delta\lambda$ is given by Zwanzig's formula[17]

$$\Delta G(\lambda_i) = -k_{\mathrm{B}} T \ln \left\langle e^{-\frac{1}{k_{\mathrm{B}}T}[\mathscr{H}(\lambda_i+\delta\lambda)-\mathscr{H}(\lambda_i)]} \right\rangle_{\lambda_i}, \tag{3}$$

where $< >$ depicts a thermodynamic average, for example, as obtained from some MD trajectory sampled including a λ_i-fold contribution of the potential given by Eq. (2), k_{B} is the Boltzmann constant, T is the absolute temperature, $\mathscr{H}$ means the total energy of the system, and the average is built at λ_i with small perturbative variations $\delta\lambda$ around λ_i. It is important to note that the perturbations should not exceed an upper-limit value which is on the order of $k_{\mathrm{B}}T$ (at least within the FEP framework). If this criterion is violated, one needs to split the perturbation again into smaller increments of λ_i. Finally, the net change in free energy is given by the sum over all incremental changes;

$$\Delta G_{\lambda=0\rightarrow 1} = \sum_i \Delta G(\lambda_i). \tag{4}$$

In this context it is of interest to note that we split the whole approach into two parts ($\Delta G_{\mathrm{cav},1} + \Delta G_{\mathrm{cav},2}$), where the first one considered the early onset of a repulsive cavity ($\lambda B = 0 \rightarrow 1$ Å), while the second part was concerned with the extension of the already established primary cavity ($\lambda B > 1$ Å). The reason for this distinction comes from the need to employ a modified repulsive potential in the small cavity regime owing to the apparent singularity in Eq. (2) when B^* approaches 0 [18,19], in particular

$$V_{\mathrm{rep}}^{\mathrm{mod}} = \frac{\lambda^n}{\left[0.3(1-\lambda)^2 + \left(\frac{r}{B^*=1.0}\right)^6\right]^2}, \qquad (5)$$

with n having been set to 12. A graphical representation of the modified potential of Eq. (5) is given in Fig. 1.

Referring to the pioneering work of Postma et al. [2] a further generalization of the free energy expression of Eq. (3) must be recalled. Within the so-called overlapping spheres technique (OST) [2] the emphasis in defining the repulsive cavities is shifted from working with various λ-scaled radii of one single maximum value of B^* to a whole set of precisely fixed and individual values of repulsive radii B_i^*. In this sense both these worlds are entirely complementary because each repulsive radius B_i^* has one corresponding λ_i that generates itself by scaling the maximum radius B^* appropriately – $B_i^* = \lambda_i B^*$. Hence a similar expression to Eq. (3) within OST when using a V_{rep} of the type of Eq. (2) therefore reads

$$\Delta G_i(j) = -k_{\mathrm{B}}T \ln \left\langle \mathrm{e}^{-\frac{1}{k_B T}\sum_k^{\mathrm{natoms}} \frac{(B_i^*+j\Delta B)^{12}-(B_i^*)^{12}}{(r_k)^{12}}} \right\rangle_{B_i^*}$$
$$j \in \{-25,-24,-23,\ldots,0,\ldots.,23,24,25\}, \qquad (6)$$

where $\langle\,\rangle_{B_i^*}$ again means an average over a trajectory recorded with the inclusion of an extra potential term of the type of Eq. (2) at some constant value of B_i^*. The perturbations now come into play. At each of the time steps of the MD simulation one computes deviations, with respect to B_i^*, that consider the induced energy change if B_i^* had suddenly become $B_i^* + (j\Delta B)$, with a very small incremental value of ΔB and j adopting negative as well as positive integral numbers, thus enabling an increase as well as a decrease of the repulsive cavity set as B_i^*. Our choice of having used exactly 25 positive and negative values for j in Eq. (6) is arbitrary. The important fact to realize is that when another trajectory is recorded, for example, employing a repulsive cavity radius of $B_{i+1}^* = B_i^* + (25\Delta B)$, then one obtains 25 different possibilities from the FEP results to estimate the free energy required to increase the repulsive cavity radius from B_i^* to B_{i+1}^*. These additional 25 possible ways of calculating the same value for $\Delta G_{\mathrm{cav},i\rightarrow i+1}$ naturally combine corresponding j-pairs within the overlapping region, for example, $\Delta G_i(1) + \Delta G_{i+1}(-24)$ or $\Delta G_i(2) + \Delta G_{i+1}(-23)$, etc. In principle, all of them should result in the same value. In reality, they do, however, differ slightly and these variations provide a perfect basis to estimate the numerical error of the method.

A similar expression to Eq. (6) but now for the early onset of a cavity (0 Å$\leq B_i \leq$ 1 Å) where the repulsive potential is of the type of Eq. (5) is defined by

$$\Delta G_i(j) = -k_{\mathrm{B}}T \ln \left\langle \mathrm{e}^{-\frac{1}{k_B T}\sum_k^{\mathrm{natoms}} \frac{\alpha_i}{\left[\beta_i+\left(\frac{r_k}{1.0+(j\Delta B_i)}\right)^6\right]^2} - \frac{\alpha_i}{\left[\beta_i+\left(\frac{r_k}{1.0}\right)^6\right]^2}} \right\rangle_{B_i^*}$$
$$j \in \{-25,-24,-23\ldots,0,\ldots 23,24,25\}, \qquad (7)$$

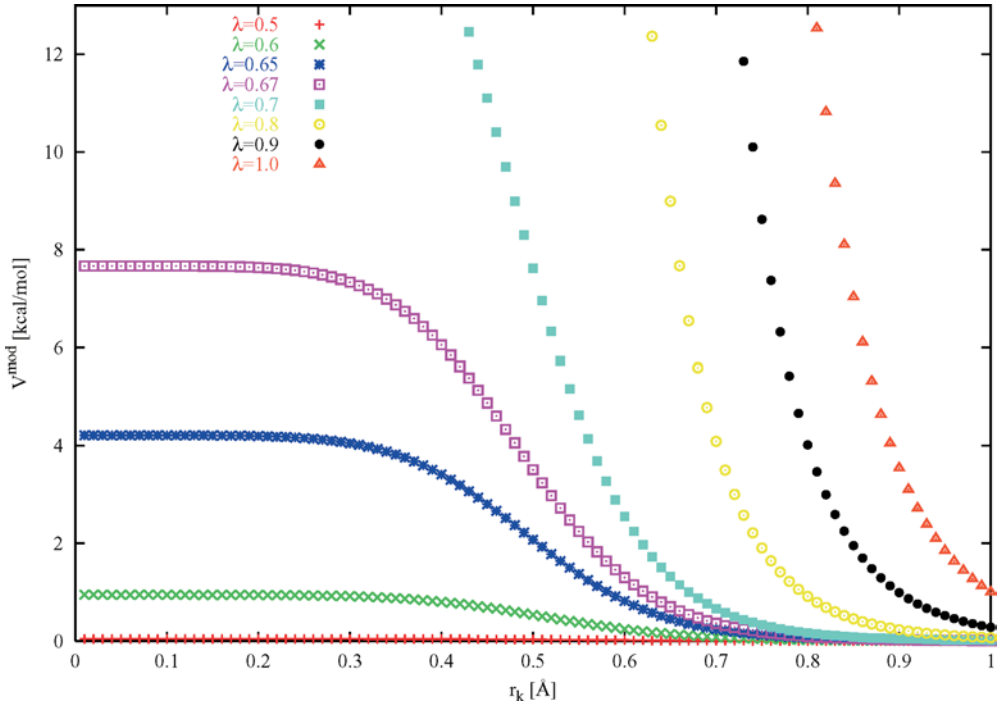

Fig. 1. Modified repulsive van der Waals potential [18] employed for small cavities of radii $B_i \leq 1.0$ Å (in the text this is referred to as the early onset domain, $\Delta G_{\mathrm{cav},1}$). The intent is to avoid the apparent singularity arising from the use of the standard repulsive van der Waals potential of type $(\frac{B_i}{r_k})^{12}$ when B_i tends to 0.0 Å and thus r_k – the distance between an atom and the cavity's center – might become 0.0 Å too. The plot shows the shape of all the eight different potentials that were actually used in each of the initial eight simulations starting with $B_i = 0.5$ Å (e.g., $\lambda = 0.5$) and reaching up to $B_i = 1.0$ Å (e.g., $\lambda = 1.0$), where finally the connection to standard van der Waals is re-established

where ΔB now is varied, because the overlapping regions for adjacent B_i^* have become asymmetric and one needs to carefully combine corresponding j values between subsequent B_i^* and B_{i+1}^* trajectories. Furthermore, several subterms in Eq. (5) are adsorbed into parameters α_i and β_i, which are all represented in Table 1.

3 Computational

All MD trajectories were computed with the TINKER program version 3.9[20, 21] using MM3-2000 parameters[22]. The ***NPT*** type of thermodynamic ensembles was modeled at 1.0 atm pressure with the pressure-coupling method due to Berendsen et al. [23], which accounts for box-size changes that naturally occur when the volume of the repulsive cavity is introduced. Initially a cubic box composed of $6 \times 6 \times 6$ grid cells was set up, where the volume of one of these grid elements was determined from the number density of *n*-octanol. A single optimized structure of *n*-octanol centered at the origin was periodically translated and copied to each of the centers in these sub-cells. The initial simulation box contained a total number of 216 *n*-octanol molecules and was roughly dimensioned to reproduce the experimental density at $T = 300\text{K}$ ($\rho = 0.8262\,\text{g/cm}^3$). Periodic boundary conditions were applied and a cutoff distance of 14.0 Å was used for non-bonded-type interaction potentials. The system was minimized and subsequently subjected to a simulated-annealing procedure, where the following protocol was used: 2000 steps of 1.0 fs each, starting with random velocities leading to a system temperature of 1000 K and using a linear cooling algorithm to finally end up at a system temperature of 300 K. The change in volume during this initial simulated annealing run was less than 1% compared with the box dimensions upon start-up.

Individual trajectories were recorded for 100 ps each, the time step was 1.0 fs, the initial 25 ps was always used for equilibration, bond lengths involving hydrogen were held constant via the RATTLE algorithm[24] and the FEP terms shown in Eqs. (6) and (7) were computed in each of the time steps, thus every femtosecond.

An initial set of eight trajectories was produced using the potential of Eq. (5) at the following values of λ 0.5, 0.6, 0.65, 0.67, 0.7, 0.8, 0.9, and 1.0, from which it becomes clear that in this initial run ($\Delta G_{\text{cav},1}$), where the maximum value of B was set to 1.0 Å, all the individual λ values can be directly taken as explicit B_i values. The choice for this particular set of λs had its reasons in the effect that they actually smooth out the singularity at $B_i = 0$ Å (Fig. 1). All parameter details of the modified potential are summarized in Table 1.

Table 1. Modified potential parameters – Eqs. (5) and (7)

$\lambda = B_i^*$	Formula (α_i, β_i)	ΔB_i	+/− Perturbative terms
0.5	$\dfrac{0.000244}{\left[0.075 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.004	$\dfrac{0.000244}{\left[0.075 + \left(\frac{r}{1.0+(j\times 0.004)}\right)^6\right]^2}$
0.6	$\dfrac{0.002177}{\left[0.048 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.0034	$\dfrac{0.002177}{\left[0.048 + \left(\frac{r}{1.0+(j\times 0.0034)}\right)^6\right]^2}$
0.65	$\dfrac{0.005688}{\left[0.037 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.0012	$\dfrac{0.005688}{\left[0.037 + \left(\frac{r}{1.0+(j\times 0.0012)}\right)^6\right]^2}$
0.67	$\dfrac{0.008183}{\left[0.033 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.0018	$\dfrac{0.008183}{\left[0.033 + \left(\frac{r}{1.0+(j\times 0.0018)}\right)^6\right]^2}$
0.7	$\dfrac{0.013841}{\left[0.027 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.0057	$\dfrac{0.013841}{\left[0.027 + \left(\frac{r}{1.0+(j\times 0.0057)}\right)^6\right]^2}$
0.8	$\dfrac{0.068720}{\left[0.012 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.005	$\dfrac{0.068720}{\left[0.012 + \left(\frac{r}{1.0+(j\times 0.005)}\right)^6\right]^2}$
0.9	$\dfrac{0.282430}{\left[0.003 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.0045	$\dfrac{0.282430}{\left[0.003 + \left(\frac{r}{1.0+(j\times 0.0045)}\right)^6\right]^2}$
1.0	$\dfrac{1.000000}{\left[0.000 + \left(\frac{r}{1.0}\right)^6\right]^2}$	0.004	$\dfrac{1.000000}{\left[0.000 + \left(\frac{r}{1.0+(j\times 0.004)}\right)^6\right]^2}$

In addition, 16 more trajectories were recorded with a potential of the type of Eq. (2) at explicit values of B_i set to 1.08, 1.16, 1.24, 1.32, 1.40, 1.48, 1.56, 1.64, 1.72, 1.80, 1.88, 1.96, 2.04, 2.12, 2.20, and 2.28 Å. For this latter set of calculations ($\Delta G_{\text{cav},2}$) a constant value of $\Delta B = 0.0032$ Å was used to compute the ± 25 perturbations according to Eq. (6).

A maximum allowed perturbation of 2 $k_{\text{B}}T$ was used throughout and any results exceeding that threshold were discarded. Final averages for a certain j perturbation away from the reference cavity radius B_i had to exhibit a failure rate of less than 1% in order to be taken into account.

Following philosophy of the Postma et al. [2], all the actually employed repulsive radii were scaled by a multiplicative factor of 1.117548 that accounts for transformations to thermal radii followed by conversion to hard-sphere radii.

4 Results and discussion

MD/FEP data for the cavitation free energy in *n*-octanol are shown in Table 2 and Fig. 2. As outlined earlier the procedure allows a rigorous estimation of the numerical error and thus error bars and lower/upper bounds are included in Table 2 and Fig. 2. Because the approach is inherently additive, the error grows with increasing cavity size. The data were interpolated with a second-order polynomial (see Eq. 1, $k_0^{\text{octanol}} = 0.331$, $k_1^{\text{octanol}} = -0.962$, $k_2^{\text{octanol}} = 0.708$) with a root-mean square of residuals of 0.011 using the Marquardt–Levenberg method [25, 26] as implemented in GNUPLOT version 3.7 [27]. Extension beyond the simulated B domain to larger cavity sizes using this polynomial is also shown in Fig. 2.

In order to estimate the uncertainty of the approach with respect to the "solvent structure" the data were analyzed a second time taking into account an initial equilibration interval of 40 ps instead of 25 ps. It is assumed that the solvent structure between two adjacent cavity simulations has then become even more disperse, because the solvent molecules have had even more time to relax and move around their cavity to find some agreeable configuration to accommodate them. In so doing we obtain another fit of $k_0^{\text{check}} = 0.362$, $k_1^{\text{check}} = -1.008$, $k_2^{\text{check}} = 0.721$, which differs from the original by only 8.4%, 4.6%, and 1.8% with notably increasing stability for the more significant k_2 coefficient. From this we conclude that the procedure is numerically stable and does not depend on arbitrarily chosen starting geometries.

Restriction of the power-series expansion to maximal quadratic terms gives rise to a rPA, which does not depend on any particular choice of a hard-sphere radius for *n*-octanol, which has been reported to be the main problematic aspect of the standard Pierotti approach (PA) [28, 29, 30, 31].

In order to justify the generalization of the second-order polynomial fit into the medium-to-large cavity domain, several more free-energy calculations were

Table 2. Cavitation free energies for growing spheres of radii B in n-octanol as obtained from free-energy perturbation calculations including the confidence interval

B (Å)	ΔG_{cav} $\left(\frac{kcal}{mol}\right)$	Lower ΔG_{cav} bound $\left(\frac{kcal}{mol}\right)$	Upper ΔG_{cav} bound $\left(\frac{kcal}{mol}\right)$
0.558774	0.000000	0.000000	0.000000
0.670529	0.002000	0.001600	0.002600
0.726406	0.004900	0.003300	0.008700
0.748757	0.008700	0.006700	0.014600
0.782284	0.015400	0.013400	0.026700
0.894039	0.049400	0.044100	0.073000
1.005793	0.091000	0.078100	0.124000
1.117548	0.148200	0.129700	0.181600
1.206952	0.211600	0.189100	0.246000
1.296356	0.281800	0.246100	0.326400
1.385760	0.348400	0.306700	0.394800
1.475164	0.439600	0.386700	0.491100
1.564568	0.550700	0.489300	0.605400
1.653971	0.659000	0.580200	0.738800
1.743375	0.785500	0.683500	0.878200
1.832779	0.949800	0.841800	1.042800
1.922183	1.109100	0.985600	1.216100
2.011587	1.276800	1.150400	1.385900
2.100991	1.435800	1.289100	1.565000
2.190395	1.623500	1.446600	1.774800
2.279799	1.831400	1.638900	1.992300
2.369202	2.031700	1.828600	2.200000
2.458606	2.248300	2.031500	2.418900
2.548010	2.464700	2.230100	2.644000

carried out for a different case using the TIP3P [32] potential of water and AMBER-6 [33]. These results will be the subject of a future communication. These calculations considered a pseudo-particle of arbitrarily scaled size that vanishes in the course of the free-energy calculation. The cavity size range was $B = 2.4 — 10.0$ Å. Interestingly enough the data obtained in this manner could be very well reproduced by the second-order polynomial fit of the MD/FEP data reported in Ref. [4] which is an identical approach to the one presented here except that the solvent considered there was water. As a consequence it seems to be sufficient to employ a rigorous MD/FEP study for small cavities and to derive a good interpolation from those data in order to accurately describe cavitation free energies in general. In so doing one may go far beyond the cavity range covered by MD/FEP in the first place.

An attempt towards verification and realization of the standard PA approach is represented in Fig. 3. Here the calculated MD/FEP data were taken and the most appropriate values for the hard-sphere radius of *n*-octanol were derived at all points from backward substitution into the equation used by Pierotti, which historically dates back to the work of Reiss et al. [34]. Furthermore, the B range was extended towards larger cavities using the second-order polynomial outside the range directly covered by MD/FEP. As becomes clear from Fig. 3, the critical property of the solvent hard-sphere radius comes close to the standard value of 3.425 Å [35], but it never really converges to a constant value throughout the entire region. Thus, ideally one would have to make the hard-sphere radius dependent on the cavity size in order to use the standard PA right away. On the other hand, instead of always deriving a most appropriate hard-sphere radius for any given problem, one could equally well stick to directly using Eq. (1).

In contrast to the standard PA, in the MD/FEP-derived rPA, the cavitation free energy is only a function of the cavity radius B, rather than a function of the ratio between the solute hard-sphere radius and the solvent

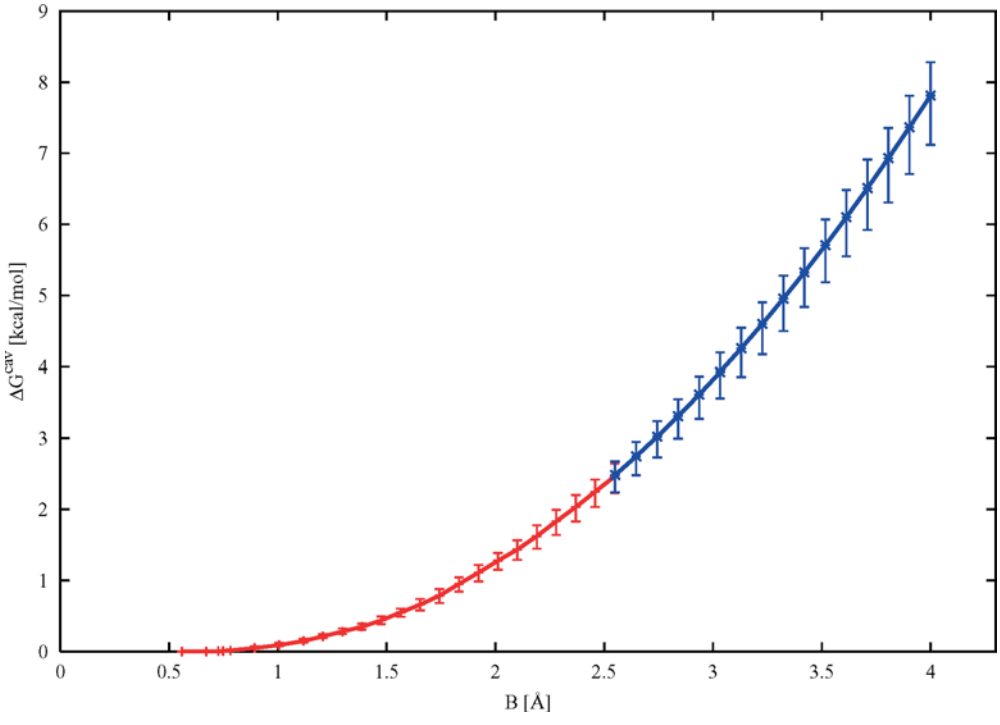

Fig. 2. Free-energy perturbation results for cavity formation in *n*-octanol (*red*) and approximate extension via a second-order polynomial fit (*blue*) as a function of the cavity size

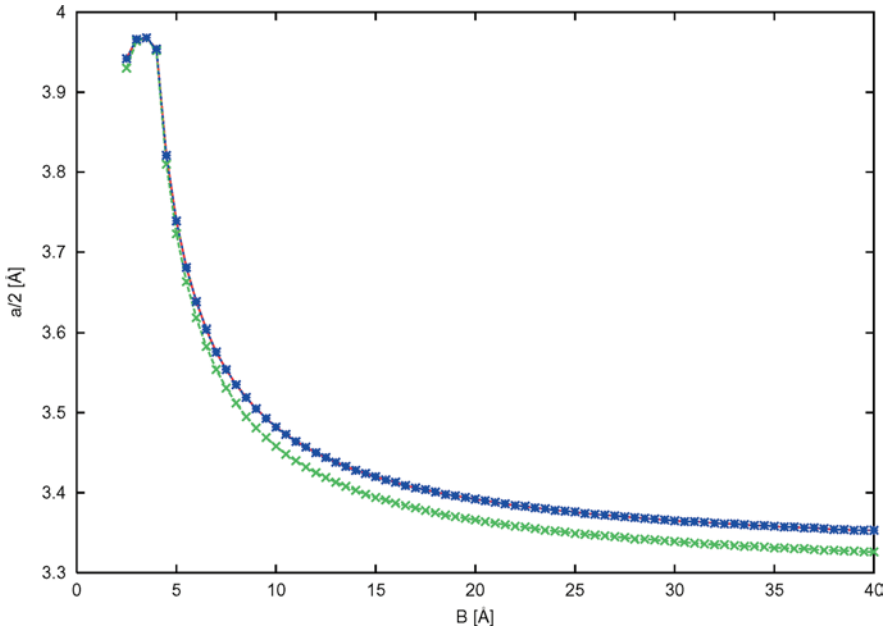

Fig. 3. Molecular dynamics (MD)/Free-energy perturbation (FEP)-derived solvent hard-sphere radius for *n*-octanol. The data obtained from MD/FEP (together with the corresponding second-order polynomial in the B_i range beyond the directly covered space) were taken as a basis and solvent hard-sphere radii were deduced via backward substitution into the equation used by Pierotti and originally derived from scaled particle theory by Reiss et al. The plot indicates that one would have to go to fairly large cavity sizes (*x*-axes) to assume a constant value of the solvent hard-sphere radius (*y*-axes) as is implicit in Pierotti's approach. Upper and lower bounds to the MD/FEP data led to the curves shown

hard-sphere radius. In addition, B exhibits an important transformation property via the solvent excluded volume, $V^{\mathrm{slv.xcl}}$ [4]. In detail, this means that for any regularly shaped molecular cavity, which usually is non-spherical, one may always deduce an effective value of B suitable for direct application of Eq. (1) by simply taking into account the solvent excluded volume, such as $B_{\mathrm{eff}} = \left(\frac{3}{4\pi} V^{\mathrm{slv.xcl}}\right)^{\frac{1}{3}}$. The great advantage is that this is always the same – at least approximately – no matter which particular value has been assigned to the probe sphere radius that formally acts as the analogue to the solvent hard-sphere radius in most of the commonly employed molecular surface generation algorithms [36, 37, 38].

Another interesting hint of the validity of using a power series truncated after the quadratic term is the fact that it adopts Neumann's rule [39] of converging to a constant ratio of $\left(\frac{B_1}{B_2}\right)^2$ for growing cavity pairs B_1 and B_2 in all kind of solvents.

A further important consequence of using a second-order polynomial fit for the ΔG_{cav} data obtained from MD/FEP calculations is that at large cavity sizes ΔG_{cav} scales with the cavity surface; hence, surface tension data might prove useful in providing an alternative estimate to the net cavitation free energy for large systems [40].

Finally, working with a second-order polynomial fit for the ΔG_{cav} data obtained from MD/FEP calculations

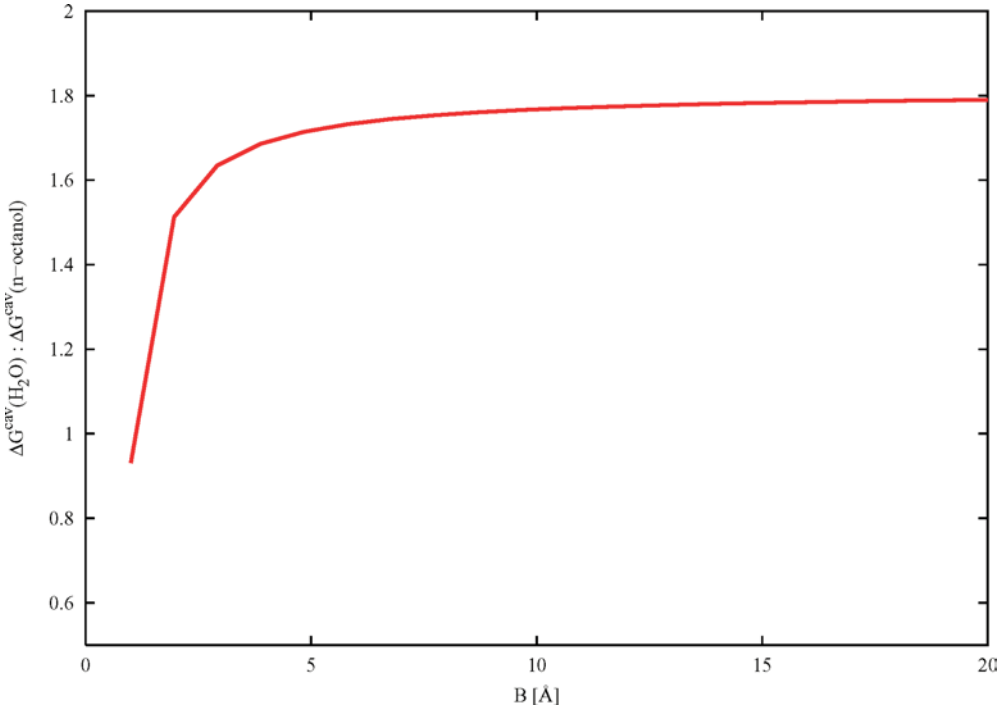

Fig. 4. Ratio between cavitation free energies in water and *n*-octanol – $\Delta G_{\mathrm{cav}}^{H_2O} : \Delta G_{\mathrm{cav}}^{n-\mathrm{Octanol}}$ – for growing cavity sizes

instead of the equation used by Pierotti [6] and originally derived from SPT by Reiss et al. [34] also implies that one has omitted the final term appearing in this equation. This term deals with the pressure dependence, $p\mathrm{d}V$, and thus would not pose any problems with small cavities since the change in volume must be considered small there. Its neglect for larger cavities is not obvious and was the subject of independent investigations [40].

As far as the log P value is concerned it might be of interest to show the trend for the ratio between ΔG_{cav}^{water} and $\Delta G_{cav}^{octanol}$. This is certainly just a subset of the overall effect, but still can give important insight into the complex relationship of the relative contributions to solvation free energies. The ratio between ΔG_{cav}^{water} and $\Delta G_{cav}^{octanol}$ is plotted in Fig. 4 (rPA coefficients for water taken from Ref. [5], $k_0^{water} = 0.823$, $k_1^{water} = -2.034$, $k_2^{water} = 1.283$). One can clearly see how even for medium-sized cavities the water term becomes a constant multiple of the octanol cavitation. In the limit of large B values, the ratio is mainly given by the relationship between the k_2 coefficients, which is $1.283/0.708 = 1.81$. Thus for molecules with an effective B radius of 10.0 Å and above, the water cavitation seems to be 1.8 times as large as the octanol cavitation. However, the cavitation effect (positive ΔG) is usually counterbalanced by the polarization (and dispersion) effect (negative ΔG), which is much larger in water, because of its largely increased dielectric constant: $\varepsilon^{20} = 80$ in water versus $\varepsilon^{20} = 10$ in octanol. In any event, the proper computation of the net solvation free energy requires as accurate as possible estimates of the individual partial terms independently and therefore the present data can serve to improve the quality of solvation free-energy calculations in *n*-octanol significantly.

5 Conclusion

FEP calculations obtained from MD simulations have been used to approximate the cavitation free energy in *n*-octanol. The data were interpolated by a power-series expansion truncated at the quadratic term, which gave a rPA. This revised version of the equation used by Pierotti and originally derived from SPT by Reiss et al. is not sensitive to any particular choice of the solvent hard-sphere radius in *n*-octanol and may be directly applied to any arbitrary molecular system via the solvent excluded volume. The present results are not only of interest to theoretical approaches that describe the solvation effect in general, but also may have a significant impact on the estimation of $\log P$ values, an important property to characterize the bio-availability of potential drug substances in different tissues.

Acknowledgments. S.H. gratefully acknowledges supercomputing support from the University of Linz, GUP, Professor Volkert and Dr. Kranzlmüller, and the University of Salzburg, RIST^{++}, Professor Zinterhof.

References

1. Landau LD, Lifshitz EM, Pitaevskii LP (1998) Electrodynamics of continuous media, 2nd edn Butterworth-Heinemann Oxford
2. Postma PM, Berendsen HJC, Haak JR (1982) Faradray Symp Chem Soc 17: 55–67
3. Simonson T, Archontis G, Karplus M (2002) Acc Chem Res 35: 430–437
4. Höfinger S, Zerbetto F (2003) Chem Eur J 9: 566–569
5. Höfinger S, Zerbetto F (2003) J Phys Chem A 107: 1253–11257
6. Pierotti RA (1976) Chem Rev 76: 717
7. Frisch MJ et al (2001) Gaussian 98 Gaussian, Pittsburgh, PA
8. Jasper JJ et al (1972) J Phys Chem 1: 841
9. Pohorille A, Pratt LR (1990) J Am Chem Soc 112: 5066–5074
10. (a) Pratt LR, Pohorille A (1992) Proc Natl Acad Sci USA 89: 2995–2999 (b) Pratt LR (2002) Ann Rev Phys Chem 53: 409–436
11. Hummer G, Garde S, Garcia AE, Pohorille A, Pratt LR (1996) Proc Natl Acad Sci USA 93: 8951–8955
12. Ashbaugh HS, Paulaitis ME (2001) J Am Chem Soc 123: 10721–10728
13. Floris FM, Selmi M, Tani A, Tomasi J (1997) J Chem Phys 107: 6353–6365
14. Matyushov DV, Schmid R (1996) J Chem Phys 105: 4729–4741
15. Boulougouris GC, Votsas EC, Economou IG, Theodorou DN, Tassios DP (2001) J Phys Chem B 105: 7792–7798
16. Prévost M, Oliveira IT, Kocher J-P, Wodak SJ (1996) J Phys Chem 100: 2738–2743
17. Zwanzig RW (1954) J Chem Phys 22: 1420–1426
18. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Chem Phys Lett 222: 529–539
19. Simonson T (1993) Mol Phys 80: 441–447
20. Kundrot C, Ponder JW, Richards F (1991) J Comput Chem 12: 402–409 (1991)
21. Dudek MJ, Ponder JW (1995) J Comput Chem 16: 791–816
22. Lii J-H, Allinger NL (1998) J Comput Chem 19: 1001–1016
23. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) J Chem Phys 81: 3684–3690
24. Andersen HC (1983) J Comput Phys 52: 24–34
25. Marquardt D (1963) SIAM J Appl Math 11: 431–441
26. Levenberg K (1944) Q Appl Math 2: 164–168
27. Williams T, Kelley C et al (1999) GNUPLOT, Linux version 3.7, patchlevel 1 http://www.ucc.ie/gnuplot/
28. Morel-Desrosiers N, Morel J-P (1981) Can J Chem 59: 1–7
29. De Voe H (1976) J Am Chem Soc 98: 1724–1729
30. Tang KES, Bloomfield VA (2000) Biophys J 79: 2222–2234
31. Gogonea V, Bãleanu-Gogonea C, Osawa E (1998) J Mol Struct 432: 177–189
32. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79: 926–935
33. Case DA et al (1999) AMBER 6 University of California, San Francisco
34. Reiss H, Frisch HL, Lebowitz JL (1959) J Chem Phys 31: 369–380
35. Marcus Y (1998) The properties of solvents. Wiley series in solution chemistry 4, Wiley, New York, chapt 3
36. Connolly ML (1985) J Am Chem Soc 107: 1118
37. Vorobjev YN, Hermans J (1997) Biophys J 73: 722–732
38. Sanner MF, Olson AJ, Spehner J-C (1995) Fast and robust computation of molecular surfaces. In: Proceedings of the 11th ACM Symposium Computational Geometry, C6–C7
39. Neumann HM (1977) J Solution Chem 6: 33–38
40. Hawkins GD, Cramer CJ, Truhlar DG (1997) J Phys Chem B 101: 7147–7157
41. Shimizu S, Ikeguchi M, Nakamura S, Shimizu K (1999) J Chem Phys 110: 2971–2982