Calculations of Relative Hydration Free Energies: A Comparative Study Using Thermodynamic Integration and an Extrapolation Method Based on a Single Reference State

Tiziana Z. Mordasini*,† and J. Andrew McCammon‡

Department of Chemistry and Biochemistry, and Department of Pharmacology, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0365

Received: September 1, 1999; In Final Form: November 4, 1999

The relative hydration free energies of a series of organic molecules were calculated from molecular dynamics (MD) simulations using an extrapolation method in combination with soft-core potential scaling. This technique consists in first generating one single long trajectory using an unphysical reference state and in using afterward this trajectory for the estimation of the free energy difference between several molecules. First, we investigated the accuracy of the method for the deletion of small functional groups. A trajectory from an MD simulation of pyrogallol (1,2,3-trihydroxybenzene) was used to calculate by extrapolation the free energy changes for the mutations of pyrogallol, catechol, and phenol to benzene in water and in vacuo. The results were compared to those obtained by thermodynamic integration, to experimental data, and to values calculated using a semiempirical method. In a second step, increasingly larger mutations were studied in order to investigate the limitations of the method. The influence of various simulation parameters (choice of the unphysical reference state, simulation length, soft-core scaling parameter α) on the final free energy values was examined. Benzene derivatives with hydroxyalkyl and/or bulky functional groups of increasing sizes were mutated into benzene. The results for simulations both in water and in vacuo were compared to the free energy results obtained by thermodynamic integration and to the experimental values of similar molecules. The results showed that for small mutations (deletion of functional groups with up to three atoms) the extrapolation method is reliable. However, the free energies calculated for the deletion of larger functional groups showed different accuracy levels depending on the chosen simulation parameters. For the largest mutations the thermodynamic integration method also showed convergence problems. This study therefore demonstrated the usefulness of the extrapolation method for molecules of similar size, but showed the difficulties of obtaining reliable results for molecules that substantially differ from each other.

Introduction

Lots of efforts have been dedicated to devise methods for the calculation of free energies from computer simulations. Many methods with several variations have been developed allowing the calculation of the free energy difference between two states A and B of a molecular system. Most of them define a path for the mutation of an initial state A to a final state B by means of a coupling parameter λ. The free energy difference between states A and B is then calculated as the work performed on the system when reversibly going from A to B along the defined path. Alternatively, the free energy difference can be obtained from the difference in relative probability of both states. A detailed overview on free energy methods can be found in a number of recent reviews.[1−5]

Most of the methods for the calculation of free energy differences are based either on the thermodynamic integration (TI) method or on the perturbation formula. Both methods lead to accurate results within a reasonable amount of time if the states A and B do not strongly differ. However, they are quite expensive from a computational point of view if the free energy difference between several states, i.e., several molecules, have to be investigated. There is therefore an increasing demand for methods for the rapid estimation of free energy differences. Such methods would in particular be extremely valuable in biochemical applications, such as structure-based drug design, where a fast screening of several compounds based on the binding free energy to an enzyme is required.

For this reason, several extrapolation methods[6−11] were developed recently, allowing the estimation of the free energy for the mutations between several molecules based on a single simulation of a reference state. One of them, described by Liu et al.,[7] utilizes the perturbation formula in combination with soft-core potential scaling for atoms that are created or deleted. This definition of certain atoms as *soft* allows the sampling of a larger configurational space, which becomes representative for both the initial and final states in perturbations where atoms are deleted or created.

In this study, we focus on this extrapolation method and we apply it for calculating the differences in hydration free energies between benzene, phenol, catechol, and pyrogallol (1,2,3-trihydroxybenzene, see Figure 1). A single trajectory, generated for pyrogallol with soft-core interaction sites for all three hydroxyl groups, was used for the estimation of the free energy difference between the four molecules shown in Figure 1. In contrast to the previous work of Liu et al.,[7] intramolecular

* To whom correspondence should be addressed. E-mail: tmordasi@ucsd.edu.
† Department of Chemistry and Biochemistry.
‡ Department of Chemistry and Biochemistry and Department of Pharmacology. E-mail: jmccammon@ucsd.edu.

10.1021/jp993102o CCC: $19.00 © 2000 American Chemical Society
Published on Web 12/22/1999

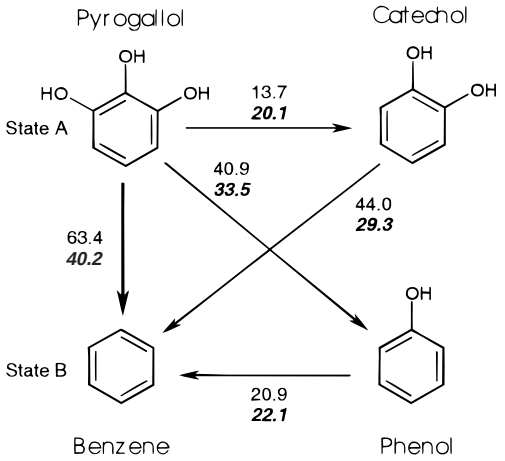

Figure 1. Structures of the hydroxylated benzene derivatives investigated in the present study. The arrows indicate the mutations applied on these systems. The final free energies of hydration (kJ/mol) calculated using the TI method (upper numbers) and extrapolation method (lower numbers) are shown for each of the investigated mutations. The reference trajectory used for the extrapolations was generated at $\lambda = 0.5$ for the mutation of pyrogallol to benzene.

degrees of freedom were sampled for all molecules, therefore resulting in flexible and more realistic systems. Moreover, partial charges were used for the different atoms, requiring a soft-core scaling of both the electrostatic and the Lennard-Jones interactions for the deleted groups. The accuracy of the extrapolation results was checked against the values obtained using the TI method. In order to compare the calculated results with experimental hydration free energies, MD simulations in water and in vacuo were carried out. Moreover, semiempirical calculations were carried out for estimating the free energy of hydration of pyrogallol, which was not available experimentally.

The method was further tested for the deletion of larger partially polar groups, such as hydroxylated alkyl chains with and without bulky substituents attached to a benzene molecule. The free energy differences were also calculated using the TI method and an estimation of the experimental values that were not available was obtained from experimental data for similar molecules.[12] The extrapolations were performed using different reference trajectories and different values for the soft-core scaling parameter α to appraise the importance of the choice of the reference state.

Theory

It is common practice in free energy calculations to use a coupling parameter λ for defining the mutation of an initial state A, with Hamiltonian H_A, to a final state B, with Hamiltonian H_B. During the perturbation the Hamiltonian becomes a function of λ, such that at $\lambda = 0$, $H(\lambda) = H_A$ and at $\lambda = 1$, $H(\lambda) = H_B$. The simplest coupling is linear scaling

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B \tag{1}$$

but since the free energy F is a state function, any kind of scaling can be used. The free energy required for the transformation of state A into state B can be calculated as[13]

$$\Delta F_{BA} = F(B) - F(A) = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{2}$$

where $\langle ... \rangle_\lambda$ corresponds to the ensemble average obtained using the Hamiltonian $H(\lambda)$. In practice, the ensemble of configurations can be obtained, e.g., by MD simulations. Equation 2 represents the so-called thermodynamic integration formula.[14] The exact calculation of ΔF_{BA} would require an infinite number of ensemble averages for λ ranging from 0 to 1. Therefore, the integral in eq 2 needs to be approximated, e.g., by a summation over a discrete number of points λ_i, leading to

$$\Delta F = \sum_i \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta\lambda_i \tag{3}$$

A finite number of λ_i values between 0 and 1 is chosen and for each of them a complete MD simulation is carried out resulting in an ensemble of configurations generated with $H(\lambda_i)$. The ensemble average of the derivative of the Hamiltonian with respect to λ is then calculated for each $\lambda = \lambda_i$.

An alternative way for the calculation of ΔF is the application of the perturbation formula[14]

$$\Delta F_{BA} = -k_B T \ln\langle \exp\{-[H(B) - H(A)]/k_B T\}\rangle_A \tag{4}$$

where k_B denotes the Boltzmann constant and T the temperature.

The extrapolation method used in this work is based on eq 4. An ensemble average generated for a certain reference state A is used for estimating the free energy difference to an alternate state B. The results are accurate as long as the ensemble of configurations generated for the reference state A is also representative of the perturbed state B. When a linear scaling (1) is used, this condition is not satisfied when atoms are deleted (or created). In solution, the presence of a functional group creates a cavity, where the solvent molecules cannot diffuse. The singularity of the electrostatic and Lennard-Jones potentials at the atomic distance $r_{ij} = 0$ prevents the solvent from sampling the entire configurational space. The ensemble of configurations generated in this way is not suitable for representing the low-energy conformations of a system where the functional group is deleted, requiring solvent molecules diffusing in the empty space left by the deleted functional group. In such cases, the use of a soft-core potential scaling[15−17] has been shown[7] to solve the problem. In this work, we used the soft-core potential scaling implemented in the GROMOS96 package[18,19]

$$V(\lambda) = (1 - \lambda)\left\{\frac{C_{12}^A(i,j)}{[\alpha_{LJ}\lambda^2(C_{12}^A(i,j)/C_6^A(i,j)) + r^6]^2} - \frac{C_6^A(i,j)}{[\alpha_{LJ}\lambda^2(C_{12}^A(i,j)/C_6^A(i,j)) + r^6]}\right\} + (1 - \lambda)\frac{q_i^A q_j^A}{4\pi\epsilon_0\epsilon_1} \times \left\{\frac{1}{[\alpha_C\lambda^2 + r^2]^{1/2}}\right\} + \lambda\left\{\frac{C_{12}^B(i,j)}{[\alpha_{LJ}(1 - \lambda)^2(C_{12}^B(i,j)/C_6^B(i,j)) + r^6]^2} - \frac{C_6^B(i,j)}{[\alpha_{LJ}(1 - \lambda)^2(C_{12}^B(i,j)/C_6^B(i,j)) + r^6]}\right\} + \lambda\frac{q_i^B q_j^B}{4\pi\epsilon_0\epsilon_1} \times \left\{\frac{1}{[\alpha_C(1 - \lambda)^2 + r^2]^{1/2}}\right\} \tag{5}$$

where $C_{12}(i,j)$ and $C_6(i,j)$ are the Lennard-Jones parameters for the van der Waals interactions between atoms i and j. α_{LJ} and α_C represent the soft-core parameters for the Lennard-Jones and Coulomb interactions, respectively, and determine together with λ the magnitude of the repulsion between two atoms i and j at $r_{ij} = 0$. Usually, the same value for α_{LJ} and α_C is used (unitless for α_{LJ} and in nm^2 for α_C), and will be referred to in the text as α. The partial charges on atoms i and j are represented by q_i

and q_j, respectively, and $\epsilon_0\epsilon_1$ represents the absolute dielectric permittivity of the medium. The parameters labeled with A and B correspond to the initial and final states A and B.

The regular Lennard-Jones and Coulomb interactions have a singularity at $r_{ij} = 0$ and this infinite repulsion prevents the solvent molecules (or other solute atoms) from diffusing over the region occupied by other atoms. The utilization of the soft-core potential scaling allows to scale down to a finite value the infinite repulsion occurring when two atoms share the same coordinates. The core of the atoms becomes therefore *soft*, allowing, e.g., solvent molecules to pass through solute atoms.

This characteristic is particularly useful in free energy calculations where the mutation involves the deletion (or creation) of atoms, since it eliminates the well-known singularity problem.[20−22] The extrapolation method used by Liu et al.[7] and in this work consists of generating a long trajectory with the soft-core potential scaling for the deleted (or created) functional groups, so that configurations relevant for both initial and final states are sampled. These configurations are afterward used in the perturbation formula (4) for estimating the free energy difference.

The advantage of the method is that the same reference trajectory can be used for evaluating free energy differences between several related states. When applying the perturbation formula, only the interactions between the perturbed atoms and the rest of the system need to be evaluated based on the previously stored trajectory. This method has been successfully used by Liu et al. and by Mark et al. to estimate free energy differences for a series of para-substituted phenols[7] and methyl-substituted benzenes,[23] respectively. Since this extrapolation method is much faster than performing a new series of simulations for each mutation, it becomes suitable for, e.g., structure-based drug design applications,[23] where many compounds need to be scored based on the binding free energy to a certain protein.

The accuracy of the extrapolation with the perturbation formula depends on the choice of the reference state used for generating the reference trajectory, which has to be representative for all the states for which the free energy has to be calculated. Liu et al.[7] showed that the best results are obtained, when an unphysical reference state A is used which "contains" all the final states B_1, B_2, ..., B_N. In our approach we define the reference unphysical state as being halfway between the two states A and B ($\lambda = 0.5$) and with soft-core scaling for the electrostatic and Lennard-Jones interactions of the atoms which are deleted. The initial state A is chosen such that all final states B_1, B_2, ..., B_N can be obtained by deletion of atoms, whereas the final state B corresponds to a molecule where all functional groups are deleted (see Figures 1 and 2). In this way, two extrapolations are needed for evaluating the entire free energy difference between A and each B_i (from $\lambda = 0.5$ to $\lambda = 0.0$ and from $\lambda = 0.5$ to $\lambda = 1.0$).

Computational Details

The molecular dynamics simulations and extrapolations were performed using the GROMOS96 computer simulation package[18] together with the GROMOS force field and the SPC water model. Simulations in vacuo were performed with addition of stochastic and frictional forces to improve the sampling efficiency.[18] Bond lengths were kept constrained using the SHAKE algorithm[24] with a relative tolerance of 10^{-4}. A time step of 2 fs was used. Reference trajectories were collected by storing configurations every 10 fs.

Temperature and pressure were kept constant by coupling the system to a Berendsen thermostat[25] and manostat at 300 K and 1 atm using relaxation times of 0.1 and 0.5, respectively. Periodic boundary conditions for a rectangular box were applied. The number of solvent molecules was chosen to ensure a minimum distance of 0.9 nm between the solute and any wall of the rectangular box. Nonbonded interactions were evaluated using a cutoff radius of 1.2 nm.

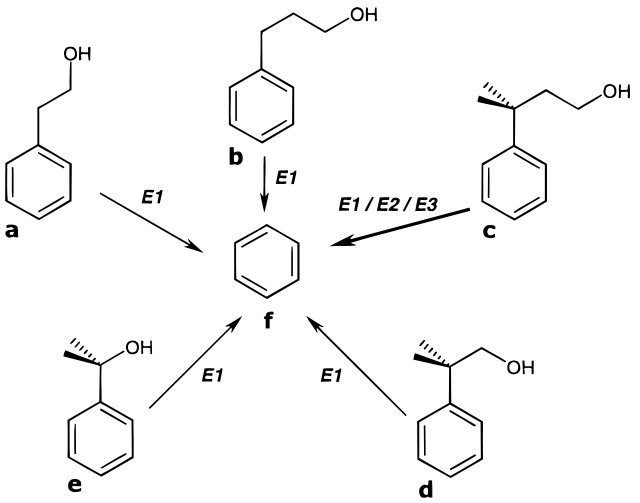

Figure 2. Structures of the five investigated benzene derivatives. The arrows indicate the mutations applied on these systems. The reference trajectories (E1, E2, and E3, as indicated in Table 2) used for the extrapolations were generated at $\lambda = 0.5$ for the mutations shown by the arrows.

The free energy values calculated using the TI method were obtained by first fitting to a cubic spline the ensemble averages of the derivative of the Hamiltonian at the different λ_i points and a subsequent numerical integration of the fitted curve.

Hydroxylated Benzenes. The reference state was defined as a pyrogallol molecule (1,2,3-trihydroxybenzene) solvated by 442 water molecules with the three hydroxyl groups treated as soft sites. The reference trajectory was generated at $\lambda = 0.5$ for a mutation of pyrogallol to benzene, corresponding to the deletion of all three hydroxyl groups (see Figure 1). The reference trajectories in water and in vacuo spanned 0.9 ns each, after 5 and 10 ps of equilibration, respectively.

An α value of 0.1 and 0.1 nm^2 was used in the soft-core interaction scaling for both the Lennard-Jones and electrostatic interactions, respectively. Two extrapolations ($\lambda = 0.5$ to $\lambda = 0.0$ and $\lambda = 0.5$ to $\lambda = 1.0$) were carried out for each of the following mutations: pyrogallol to benzene, pyrogallol to phenol, and pyrogallol to catechol, corresponding respectively to the deletion of three, two, and one hydroxyl groups. For the last two mutations, the three and respectively two possible combinations for deleting two and one hydroxyl groups were evaluated and the results averaged for obtaining the final value.

The free energy differences for the same three mutations were also calculated using the TI method with soft-core potential scaling for the hydroxyl groups. A total of 12−14 λ-points was used, each of them obtained from 20 to 40 ps simulations in vacuo and in water, preceded by an equilibration period of 10 ps.

To check the consistency of the results, extrapolations were carried out for mutations from catechol to benzene and from phenol to benzene, closing two thermodynamic cycles (see Figure 1).

No experimental hydration free energies were available for pyrogallol. We therefore performed semiempirical calculations at the AM1 level[26] for all four molecules shown in Figure 1.

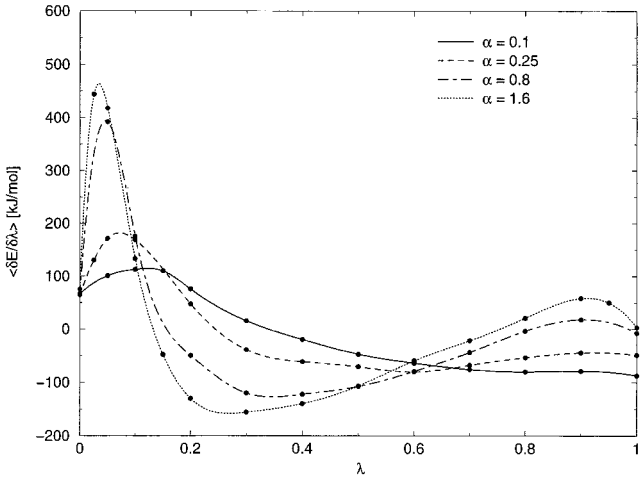

Figure 3. Dependence of the free energy derivative and of free energy profile on the soft-core parameter α for the deletion of the three hydroxyl groups of pyrogallol in water.

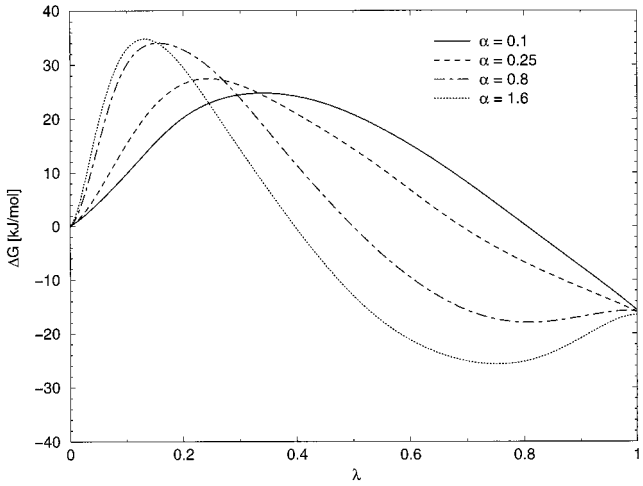

Figure 4. Dependence of the free energy derivative and of free energy profile on the soft-core parameter α for the deletion of the three hydroxyl groups of pyrogallol in water.

These semiempirical calculations were carried out with the MNDO97 program.[27] All molecules were optimized both in vacuo and with the continuum solvation model COSMO[28] with additional cavitation and dispersion terms[29] always leading to planar minimum conformations. The COSMO calculations were performed using a dielectric constant of 78.4, corresponding to water solution. The number of segments per atom for the construction of the solvent accessible surface was set to 162, and the van der Waals radii of Bondi[30] were used. For pyrogallol and catechol several planar conformers are possible. In order to obtain a better estimate of the free energy, all low-energy conformers with the hydroxyl groups lying in the same plane as the phenyl ring were calculated. The obtained heats of formation in vacuo and in solution were averaged with Boltzmann weighting.

Alkyl Alcohol Derivatives of Benzene. The free energy differences between the five substituted benzene molecules shown in Figure 2 and benzene were calculated using the extrapolation method and TI. All the atoms of the perturbed functional groups were in each case deleted using a soft-core potential scaling with $\alpha = 0.6$, corresponding to the value used by Liu et al.[7] For the mutations of molecules **c**, **d**, and **e** to benzene using the TI method, additional calculation were carried out using $\alpha = 10$.

The TI calculations were performed in vacuo and in water, where the five molecules **a**, **b**, **c**, **d**, and **e** were solvated with 590, 695, 509, 488, and 470 water molecules, respectively. A total of 11−14 λ-points for each mutation of the molecules **a**−**e** to benzene were used. For each λ-point, the average value of the derivative of the potential energy with respect to λ was obtained from a 100 ps (vacuum) or 50 ps (water) simulation, after an equilibration period of 20 ps (vacuum) or 5 ps (water).

The extrapolation of the free energy differences for the five mutations was carried out using different reference trajectories:

E1. A series of five trajectories was generated, each with 0.3 ns in water and 0.7 ns in vacuo, with $\alpha = 0.6$ and at $\lambda = 0.5$ between each one of the five molecules studied (**a**−**e**) and benzene.

E2. Two reference trajectories (1 ns in water and 1.4 ns in vacuo) were generated with $\alpha = 0.6$ at $\lambda = 0.5$ between the largest molecule **c** and benzene, and were subsequently used for estimating the free energy differences between all five molecules **a**−**e** and benzene.

E3. Using the same procedure as used in E2, two other trajectories (0.9 ns in water and 1.4 ns in vacuo) were generated using $\alpha = 1.45$.

E4. As a check for consistency, the mutation of molecule **d** to **a** was studied, consisting of the deletion of two methyl groups. The reference trajectory was generated with the same procedure as used in E1.

Results and Discussion

Hydroxylated Benzenes. The choice of the α-parameter, determining the amount of repulsion between two atoms at short distances, is arbitrary. In principle, the choice of α should not affect the results for the free energy change corresponding to a certain mutation.[15] However, it influences the shape of the λ-dependent free energy profile. The optimal α value should be chosen so as to minimize the perturbation introduced into the system and therefore to produce a smooth transition from the initial to the final state.

For the deletion of the three hydroxyl groups of pyrogallol in water, we investigated this effect by comparing the results obtained with four different α values (0.1, 0.25, 0.8, 1.6).

As can be seen from the profiles displayed in Figures 3 and 4, the smoother transition is obtained with $\alpha = 0.1$, a value which is smaller than the values of 0.6−0.8 used in previous studies.[7,23] This may be due to the fact that the deleted hydroxyl groups lie close to the carbon atoms of the phenyl ring. The van der Waals radius of these carbon atoms is large enough to "protect" the core of the oxygen atoms from solvent molecules. Even when the hydroxyl groups are deleted, water molecules cannot approach too closely the center of the oxygen atom, due to the repulsion caused by the carbon atoms of the phenyl ring. For this reason, in this case there is no need for a high α-parameter. Consequently, a rather *hard* scaling, that is a small α value, is most suitable for the deletion of such small functional groups. We therefore used for all the investigated mutations involving the deletion of hydroxyl groups an α value of 0.1.

The results obtained for the mutations between pyrogallol, catechol, phenol, and benzene are shown in Table 1.

The overall agreement between extrapolation and TI is reasonably good, especially in vacuo. As expected, the largest deviation from the TI results is observed for the largest mutation of pyrogallol to benzene, where the three hydroxyl groups are deleted. In order to check if the discrepancies between extrapolation and TI methods were due to inefficient sampling, all

TABLE 1: Calculated Free Energies (kJ/mol) for the Mutations between Pyrogallol, Catechol, Phenol, and Benzene Using $\alpha = 0.1$[a]

	ΔG in vacuo		ΔG in water		$\Delta\Delta G_{hydr}$		$\Delta\Delta H_f$	
	TI	extrap	TI	extrap	TI	extrap	AM1	expt[12]
pyrogallol → benzene	−79.0	−75.2	−15.6	−35.0	63.4	40.2	46.7	
pyrogallol → phenol	−114.7	−114.8	−73.8	−81.3	40.9	33.5	26.0	
pyrogallol → catechol	−40.7	−44.1	−27.0	−24.0	13.7	20.1	10.0	
catechol → benzene	−34.4	−32.7	9.6	−3.4	44.0[b] (49.7[c])	29.3[b] (20.1[c])	36.8	35.6
phenol → benzene	38.8	34.8	59.7	56.9	20.9[b] (22.5[c])	22.1[b] (6.7[c])	20.8	23.9
catechol → phenol					23.1[d]/27.2[c]	7.2[d]/13.4[c]	16.0	12.0

[a] $\Delta\Delta H_f$ corresponds to the difference in free energy of formation between the molecule optimized with a continuum solvation model for water and in vacuo. Extrapolation results are obtained from reference trajectories of 0.9 ns in water and vacuo, respectively. The hydration free energy ($\Delta\Delta G_{hydr}$) corresponds to the difference between the calculated free energy change in water and in vacuo. [b] Results for the direct mutation. [c] Results obtained from the summation over a thermodynamic cycle using pyrogallol as intermediate molecule. [d] Results obtained from the summation over a thermodynamic cycle using benzene as intermediate molecule.

extrapolations in vacuo were repeated using a longer (2 ns) trajectory which ensured sampling of all the low-energy conformations with respect to the hydroxyl orientations. No substantial changes in the final extrapolated values could be observed, pointing out that some errors are introduced by the extrapolation method itself and by the choice of the α value. By building thermodynamic cycles, it is possible to calculate the difference in free energy of hydration between the three molecules.

The calculated free energy differences, from both TI and extrapolation, qualitatively reproduce the trends observed experimentally[12] and in the semiempirical results (see Table 1). Both TI and extrapolation methods very well reproduce the free energy difference between benzene and phenol. However, the TI calculations predict an almost identical free energy contribution for the introduction of the second hydroxyl group, overestimating the free energy difference between phenol and catechol. For the same reason, also the change in free energy between benzene and catechol is overestimated. On the other hand, the extrapolation method slightly underestimates the experimental results, but yields an overall good agreement for the differences in hydration free energies between catechol, phenol, and benzene (last three rows of Table 1). The discrepancies between TI and extrapolation data are probably due to sampling insufficiencies in solution and by approximations implied in the extrapolation method, since all other parameters were kept identical. The largest discrepancies occur in the water simulations, where accurate sampling is more difficult. For quantitative results, very long reference trajectories may be required. However, the goal of the extrapolation method is to yield an estimated value and therefore only a qualitative scoring of free energies for several molecules and not to replace the more accurate but also more computationally expensive TI method. Given the performed mutations (see Figure 1), it is possible to check the consistency of the results by building thermodynamic cycles that should theoretically give a zero total free energy difference. The cycle pyrogallol → catechol → benzene → pyrogallol has a residual free energy of 5.7 kJ/mol using TI and 9.2 kJ/mol using the extrapolation method. Similarly, residual free energies of 1.6 kJ/mol for TI and 15.4 kJ/mol for the extrapolation method are obtained for the cycle pyrogallol → phenol → benzene → pyrogallol. The largest errors are therefore clearly present in the results obtained with the extrapolation method. However, the major discrepancy seems to be concentrated only in the largest mutation, i.e., from pyrogallol to benzene. The data (see Table 1) show that the results calculated using the cycle leg pyrogallol → benzene (those shown in parentheses in the extrapolation row) are off by up to 15 kJ/mol, whereas those calculated either directly or using other cycle legs reproduce much better the experimental data. Because of the nature of the extrapolation method, where the same single trajectory is used for several mutations, it is possible that the sampling achieved during a certain period of simulation time is accurate enough to ensure good results with respect to the smallest mutations, but at the same time not complete enough for the larger ones.

For the mutations of pyrogallol to phenol and catechol, respectively three and two combinations for the deletion of the hydroxyl groups are possible. The free energy values obtained for the same mutation using the deletion of different hydroxyl groups were very similar (with a maximal deviation of 5 kJ/mol; data not shown), showing that the three hydroxyl groups of pyrogallol were equally sampled in the reference trajectory.

Alkyl Alcohol Derivatives of Benzene. Table 2 displays the free energy differences calculated for the mutations of the molecules shown in Figure 2 to benzene and for the mutation of **d** to **a**. Three values for the extrapolation method are reported, depending on the reference trajectory used (see Computational Details section).

The experimental values for ΔG_{hydr} can be estimated from the data for similar molecules without the phenyl ring,[12] leading to expected values in the range of 18−21 kJ/mol for all five substituted benzenes shown in Figure 2.

The mutations investigated are quite large and when branched bulky groups are deleted, even the results obtained with TI show inconsistencies. Surprisingly, the simulations in vacuo represented a substantial problem. When using $\alpha = 0.6$ for the soft-core potential scaling, TI overestimated the free energy differences for the one-step deletion of the branched functional groups (**c**−**e**) in vacuo. In order to quantify this overestimation, the mutation of **d** into benzene, which was first carried out in one step (value in parentheses in Table 2), was split into the mutation of **d** to **a** and **a** to benzene (**f**). The splitting of such a large mutation using an intermediate state resulted in more accurate results. The difference in free energy between **d** and benzene in vacuo obtained using TI and $\alpha = 0.6$ with either the direct or split mutations is important (25 kJ/mol). In order to further investigate this large overestimation of ΔG in vacuo, TI calculations for the mutation of **d** to benzene were repeated using different α values.

The results, shown in Table 3, and generated from MD simulations of 20 ps for every λ-point, systematically depend on the α value used. For $\alpha = 0$, the expression 5 is reduced to a linear scaling and all the known problems connected to such a scaling[15−17,20−22] when deleting atoms become evident from the completely unreliable result. By increasing α, ΔG increases reaching convergence only when very high values of the α parameter are used. The free energy value obtained with $\alpha =$

TABLE 2: Calculated Free Energy Differences (kJ/mol) for the Five Mutations Shown in Figure 2[a]

	ΔG in vacuo				ΔG in water				$\Delta\Delta G_{hydr}$			
	TI	E1[b]	E2[c]	E3[d]	TI	E1[b]	E2[c]	E3[d]	TI	E1[b]	E2[c]	E3[d]
a to **f**	−4.4[f]	−1.0	1.0	−1.7	22.4[f]	18.5	32.7	21.1	26.8[f]	19.5	31.8	22.8
b to **f**	−5.3[f]	−0.2	6.1	−6.2	20.0[f]	34.5	37.0	16.2	25.3[f]	34.6	30.9	22.4
c to **f**	−45.1[e]	−38.5	−38.5	−50.1	−13.8[e]	−20.5	−20.5	−59.5	31.3[e]	18.0	18.0	−9.3
	(−66.3[f])				(−12.8[f])				(53.4[f])			
d to **f**	−36.4[e]	−40.8	−24.6	−29.5	−4.1[e]	−12.3	31.8	−6.0	32.3[e]	27.8	56.4	23.5
	(−59.1[f])				(−8.6[f])				(50.0[f])			
e to **f**	−32.9[e]	−34.9	−12.3	−27.4	−6.5[e]	−6.4	19.2	−19.0	26.4[e]	28.5	31.5	8.4
	(−50.5[f])				(−5.6[f])				(44.9[f])			
d to **a**	−29.7[f]	−30.2			−28.5[f]	−27.0			1.1[f]	3.2		

[a] The hydration free energy ($\Delta\Delta G_{hydr}$) corresponds to the difference between the calculated free energy change in water and in vacuo. [b] Data obtained with extrapolation from reference trajectories (0.3 ns in water and 0.7 ns in vacuo) generated at $\lambda = 0.5$ and $\alpha = 0.6$ between each of the five molecules **a**−**e** and benzene. [c] Data obtained with extrapolation from one unique reference trajectory (1 ns in water and 1.4 ns in vacuo) generated at $\lambda = 0.5$ between **c** and benzene using $\alpha = 0.6$. [d] Data obtained with extrapolation from one unique reference trajectory (0.9 ns in water and 1.4 ns in vacuo) generated at $\lambda = 0.5$ between **c** and benzene using $\alpha = 1.45$. [e] Values obtained using $\alpha = 10$. [f] Values obtained using $\alpha = 0.6$.

TABLE 3: Free Energy Differences (kJ/mol) Calculated in Vacuo Using the TI Method for the Mutation of d to Benzene Using Increasing α Values

α	0.0	0.1	0.4	0.6	1.0	10	30
ΔG_{vac}	∼−2500	−79.1	−63.5	−58.8	−52.3	−36.4	−35.7

10 (−36 kJ/mol) is in agreement with the value obtained by splitting the mutation of **d** to benzene in two steps (−34 kJ/mol). Such an α-dependency was surprising, since the calculated ΔG should not depend[15] on the choice of α as also shown in Figure 4. This effect is therefore probably due to sampling problems, caused by the combination of a large mutation with a too hard α parameter. It is possible to rationalize these results by considering the large size of the deleted functional groups and the molecule on which this mutation is performed. The larger the mutation, the softer the scaling should be in order to avoid a large repulsion between the atoms that are deleted and the atoms of the phenyl ring. In order to prove that high α values deliver more accurate free energies for the branched functional groups, TI calculations in vacuo and in water were repeated for the mutations of **c**, **d**, and **e** to benzene using $\alpha = 10$. The new results, reported in Table 2, are in much better agreement with the expected experimental values. Using equal simulation lengths, a higher α value causes in this case a more rapid convergence of the results.

Interestingly, the values from the extrapolations using a reference trajectory generated with $\alpha = 0.6$ for the five systems (column E1 in Table 2) deliver more accurate results than TI with the same α value, when considering the $\Delta\Delta G_{hydr}$. The reference trajectories used for the extrapolations are longer than one single MD simulation for every λ-point and should therefore ensure better sampling. The errors due to the approximation of the free energy using an extrapolation cancel out between vacuum and water simulations, resulting in more realistic $\Delta\Delta G_{hydr}$ for all molecules irrespective of whether their functional groups are branched or linear. These results were, however, obtained with extrapolation from reference trajectories that were specifically generated with the topology of the molecule that had to be investigated. The sampling was therefore tailored for the desired mutation.

The results change considerably if one single reference trajectory, based on the largest molecule **c**, is used for calculating all five molecules **a**−**e** using $\alpha = 0.6$ (column E2 in Table 2). As expected, the free energy values converge rather slowly. Even in vacuo, after 1.4 ns simulation the results from the extrapolation using one single trajectory often significantly differ from those obtained with the first extrapolation (E1). As already shown by Liu et al.,[7] the cumulative values of the free energy obtained by extrapolation display a stepwise evolution when low-energy conformations are sampled. Because of the exponential function used for the ensemble average in the perturbation formula (4), configurations with large negative interaction energies provide the most significant contributions to the ensemble average and therefore strongly determine the extrapolation results. If a larger functional group is deleted, more degrees of freedom have to be sampled, because all molecules were treated as flexible, requiring a longer reference trajectory to obtain meaningful extrapolation results. It is, however, difficult to know in advance how long the trajectory should be and to check the convergence of the extrapolated values. As shown in Figure 5 for the mutation of **c** to **d** in vacuo, although the ΔG values seemed to have reached convergence after 1 ns, a continuation of the simulation caused a sudden drop, when a new low-energy configuration was sampled.

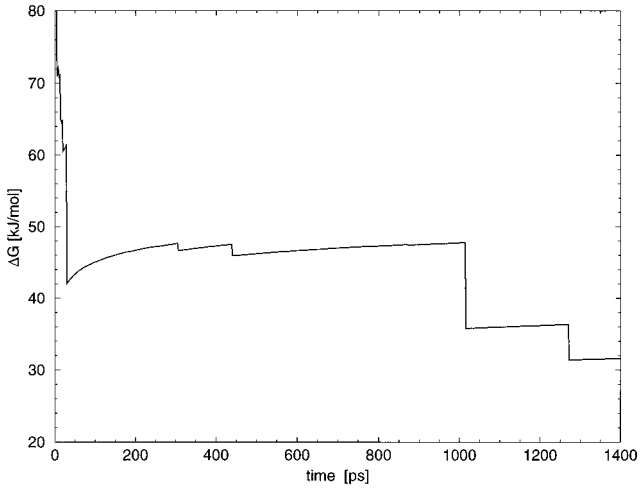

Figure 5. Cumulative value of the extrapolated ΔG as a function of the simulation time for the mutation of **c** to **d** in vacuo.

From Table 2, we can also observe that for three mutations (**a** to **f**, **b** to **f**, and **e** to **f**) the extrapolated values using one single trajectory (column E2) approximate the results from the first extrapolations reasonably well, consistently predicting the same free energy difference for all three molecules, as expected from the estimated experimental values. However, the free energy values for the other two molecules (**c** and **d**) are not yet converged.

The third series of extrapolation results (column E3 in Table 2) has been obtained from a single reference trajectory generated

at $\lambda = 0.5$ for the mutation of **c** to benzene using a higher α value of 1.45. The influence of the α parameter was investigated by using a softer scaling of the functional group. Already, with the TI method, it appeared that larger α values yielded more accurate results for the deletion of large functional groups. On increasing the α parameter, the repulsion between two atoms at $r_{ij} = 0$ is decreased and a better sampling for the final state (benzene) is expected. Similar to the results obtained with $\alpha = 0.6$, using $\alpha = 1.45$ three mutations (**a** to **f**, **b** to **f**, and **d** to **f**) lead to very similar results (ca. 23 kJ/mol) and closer to the estimated experimental data (18−21 kJ/mol). On the other hand, the free energies associated with the mutation of **c** and **e** into benzene are significantly underestimated. This indicates that the sampling obtained during the 0.9 ns simulation is not sufficient even for a qualitative estimation of the largest mutation considered. The increase of the α parameter from 0.6 to 1.45 improves the quality of the results only for three of the five mutations considered. It is not surprising that the same reference trajectory can lead to converged free energy differences for only some of the molecules considered (such as **a**, **b**, and **d**). The same effect can be observed from the results using the reference trajectory E2. In order to deliver qualitatively good results for all five evaluated molecules, the reference trajectory should contain representative low-energy conformations for each molecule **a**−**f**. In the case of the reference trajectory E3, low-energy conformations for **c** and **e** are clearly not yet sampled, whereas the extrapolated results for **a**, **b**, and **d** are in very good agreement with the experimental data ($\Delta\Delta G < 3$ kJ/mol).

The overall agreement of the extrapolation results with TI and experimental data has to be judged also considering the complexity of the studied systems. The mutations studied involved the deletion of large, flexible and partially polar functional groups. More sampling problems are caused by taking into account the flexibility of the molecules. Moreover, both the large size of the deleted groups and the presence of the polar hydroxyl groups make it more difficult to sample solvent configurations that are representative for both the hydrophobic benzene and the partially polar molecules **a**−**e**, since polar functional groups require a more structured solvent distribution than the hydrophobic benzene.

In order to better understand the extrapolation method, other kinds of mutation were investigated. Given a certain reference trajectory, the extrapolation method allows to use the same ensemble of configurations for several different mutations. The data shown in column E2 and E3 of Table 2 were calculated from a trajectory generated on the basis of the largest molecule **c** mutated to benzene. Using this trajectory, the free energy values for the mutation of different molecules, such as **a**, **b**, **d**, and **e**, to benzene were evaluated. However, other types of mutations are also possible, such as direct mutations between the molecules **a**−**e** without going through benzene as an intermediate. Mutations such as **c** to **d** or **c** to **b** are relatively small, corresponding to a shortening of the alkyl chain by one unit or deleting two methyl groups, and could therefore give fewer problems in terms of convergence. However, the evaluation of those mutations using the same reference trajectories generated with $\alpha = 0.6$ and $\alpha = 1.45$ led to poor results (data not shown). This is nevertheless an interesting result, pointing out that the choice of the reference trajectory should match the type of the desired mutation. In this case, the reference trajectory was generated for a certain type of mutation, i.e., the deletion of the *entire* functional group attached to the phenyl ring (mutation of **c** to benzene), and could therefore be used for similar mutations, such as **a**−**e** to benzene. However, the configurations sampled were not representative for other kinds of mutations, even if smaller, such as a *partial* deletion of the functional group (mutations of **c** to **d** or **c** to **b**).

Conclusions

In this study, the efficiency of the extrapolation method using a single unphysical reference state was investigated and compared to the results obtained with TI and to experimental data. Functional groups of increasing size were deleted. The influence of the soft-core parameter α, of the length of the reference trajectory, and of the choice of the unphysical reference state on the extrapolation results was analyzed.

We could confirm that the extrapolation method in combination with soft-core scaling developed by Liu et al.[7] is a powerful tool for a qualitative estimation of free energy differences between similar molecular systems. For mutations involving the deletion of functional groups with up to three atoms, one single reference trajectory can be used for a qualitative scoring of several compounds.

For mutations involving the deletion of large functional groups with several degrees of freedom, the accurate sampling of the entire configurational space becomes problematic. In such cases, the results obtained with the extrapolation method strongly depend on the reference trajectory used, and as expected, the differences become more evident for shorter trajectories. Moreover, in contrast to the work of Liu et al.,[7] all the molecules were treated as flexible with partial charges on the atoms. This choice made it necessary to carry out simulations both in vacuo and in solution, in order to obtain the hydration free energy, and to use a soft-core potential scaling for both the Lennard-Jones and electrostatic interactions. Flexible partially polar molecules represent more realistic systems; however, the increased number of degrees of freedom also implies more sampling problems. The presence of the polar hydroxyl groups also makes it more difficult to sample solvent configurations representative for both the polar functional group and the hydrophobic benzene.

The choice of the unphysical reference state, the length and number of configurations of the reference trajectory, and the α parameter become crucial for large mutations, since they determine the efficiency of the sampling. If the reference trajectory is generated using the topologies of the desired initial and final states and using the same kind of mutation for which the free energy has to be calculated, reliable results can be obtained from relatively short simulations. However, if one single trajectory is used for extrapolations to differing molecular states, the convergence of the calculated free energy values is not assured. The production of predictive values becomes difficult, since it is impossible to establish the degree of convergence of the results. Because of the presence of an exponential weighting in the perturbation formula for the calculation of the ensemble average, the sampling of a new configuration with low interaction energy causes a sudden drop of the extrapolated free energy (Figure 5). A constant running average of the ΔG value over a trajectory of several hundred picoseconds does not necessarily mean that the value has converged. Moreover, the same reference trajectory may yield converged results for some mutations and unreliable results for other, depending on the presence or absence of representative low-energy conformations for the evaluated molecule. Therefore, the length of the reference trajectory should increase with the number of degrees of freedom of the deleted (or created) functional group. The generation of very long trajectories (at least in the nanosecond range) containing hundreds of thousands

of configurations may be required for more accurate results. This can eventually cause practical storage problems, especially for simulations in explicit solvent. However, even if the reference trajectory has to cover a long simulation time, the extrapolation method is still more convenient than TI when several molecules have to be qualitatively scored.

The choice of the α parameter, determining the softness of the potential scaling, turns out to be also very important. Even if the total free energy difference between two states does not theoretically depend on α, the parameter strongly influences the shape of the free energy profile (Figures 3 and 4), and therefore determines how fast low-energy conformations are sampled. Because of the finite length of the simulations, this automatically implies that different α-parameters can lead to more or less converged free energy values. The results of this work suggest that high α values are required for larger mutations, whereas a relatively hard scaling (small α values) gives better results for smaller mutations, at least in the case where the deleted groups are attached to a hard molecule. It is, however, not possible, based on these results, to give a general recipe for the choice of the α value which should still be optimized if other systems are analyzed. Moreover, also the environment influences the choice of the parameter α, e.g. in the case that another solvent is used or for molecules bound to an enzyme.

In general, the overall agreement between the different kinds of extrapolations, TI, and experimental data for the deletion of large functional groups was relatively satisfactory, given the large mutations investigated on flexible molecules with polar functional groups. However, for more reliable qualitative results in the case of large mutations, the use of a longer reference trajectory (>1 ns) or the splitting of the mutation using intermediate states is recommended.

Acknowledgment. This work was supported in part by a grant from NSF. T.Z.M. is grateful to the Swiss National Science Foundation for a postdoctoral fellowship (no. 8220-053397) and to Dr. Philippe Hünenberger and Dr. Wolfgang Weber for a critical reading of this manuscript and helpful discussions. We also thank Prof. Wilfred van Gunsteren for providing a copy of the GROMOS96 computer simulation package and Prof. Walter Thiel for allowing us to perform the semiempirical calculations with his program MNDO97.

References and Notes

(1) King, P. M. In *Computer Simulations of Biomolecular Systems*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden,The Netherlands, 1993; Vol. 2, pp 267−314.
(2) McCammon, J. A. *Curr. Opin. Struct. Biol.* **1991**, *1*, 196.
(3) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.
(4) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
(5) Jorgensen, W. L.; Tirado-Rives, J. *Perspect. Drug Discov.* **1995**, *3*, 123.
(6) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577.
(7) Liu, H.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 9485.
(8) Levy, R. M.; Berhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.
(9) King, G.; Barford, R. A. *J. Phys. Chem.* **1993**, *97*, 8798.
(10) Cieplak, P.; Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1994**, *101*, 627.
(11) Radmer, R. J.; Kollman, P. A. *J. Comput. Chem.* **1997**, *18*, 902.
(12) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292.
(13) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(14) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(15) Beutler, T. C.; Mark, A. E.; van Shaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529.
(16) Gerber, P. R.; Mark, A. E.; van Gunsteren, W. F. *J. Comput.-Aid. Mol. Des.* **1993**, *7*, 305.
(17) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025.
(18) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdF: Hochschulverlag AG and der ETH Zürich and BIOMOS b.v., Zürich, Groningen 1996; ISBN 3 7281 2422 2.
(19) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.
(20) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1.
(21) Mruzik, M. R.; Abraham, F. F.; Schreiber, D. E.; Pound, G. M. *J. Chem. Phys.* **1976**, *64*, 481.
(22) Cross, A. J. *Chem. Phys. Lett.* **1986**, *128*, 198.
(23) Mark, A. E.; Xu, Y.; Liu, H.; van Gunsteren, W. F. *Acta Biochim. Pol.* **1995**, *42*, 525.
(24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(25) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(26) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(27) Thiel, W. Program MNDO97, University of Zürich, Zürich, 1998.
(28) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.
(29) Gelessus, A. Ph.D. Dissertation, University of Zürich, 1997.
(30) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.