Free Energies of Transfer of Trp Analogs from Chloroform to Water: Comparison of Theory and Experiment and the Importance of Adequate Treatment of Electrostatic and Internal Interactions

Xavier Daura,[†] Philippe H. Hünenberger,[‡] Alan E. Mark,[‡] Enrique Querol,[†] Francesc X. Avilés,[†] and Wilfred F. van Gunsteren*[,‡]

Contribution from the Institut de Biologia Fonamental and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain, and Laboratorium für Physikalische Chemie, ETH Zentrum, CH-8092 Zürich, Switzerland

Received November 13, 1995. Revised Manuscript Received April 22, 1996[⊗]

Abstract: Experimentally determined water/chloroform partition coefficients for three indole derivatives (3-methylindole, *N*-acetyltryptamine, and 3-(3′-indolyl)propionic acid *N*-methylamide), models of the amino acid tryptophan, are compared to free-energy differences calculated using molecular dynamics simulations. The effect of the choice of force field, the choice of pathway along which the free-energy change is calculated, the inclusion of free-energy contributions from constraints, and the treatment of long-range interactions on the agreement with the experimental data for this system are investigated with an eye to understanding the importance of the different aspects of the molecular model and computational procedure. It is demonstrated that, although the compounds are neutral and do not differ much in dipole moment or charge distribution, the incorporation of a reaction field to treat long-range electrostatic interactions is necessary to reproduce the experimentally observed trends. The implications of these findings for free-energy calculations in general and for the estimation of partition coefficients in particular are discussed.

Introduction

The extent to which a compound partitions between two environments, for example, between an aqueous phase and an organic phase, is central to a wide variety of chemical and biochemical processes. These include separation,[1−3] ligand binding,[4,5] and protein folding,[6−8] to name a few. Considerable effort has, therefore, been directed toward the tabulation of experimentally determined partition coefficients and the development of theoretical approaches to predict partition coefficients. A variety of methods to predict partition coefficients have been proposed,[9−14] primarily based on empirically derived group contributions. Such approaches are highly accurate when used for homogeneous systems. They fail, however, in cases where the influence of different substituent groups is correlated, or in cases where the environment cannot be approximated as a homogeneous continuum. This is often the case in biological systems. The effects of amino acid substitution on protein stability, for example, frequently correlate poorly with experimentally determined partition coefficients between water and a low dielectric solvent such as octanol for related compounds.

In cases where the system cannot be treated as a homogeneous continuum or where the neighboring groups are not independent, methods that directly incorporate specific intramolecular interactions and structural information are required. Partition coefficients are directly related to the difference in free energy of the compound in the two environments. In principle, such free-energy differences can, on the basis of statistical mechanics, be determined using molecular simulation techniques. Free-energy calculations are potentially highly accurate. A number of previous studies have shown that such calculations can be used to reproduce experimental partition coefficients for simple compounds.[15,16] The generality of this approach in practical terms is still an open question, however, as is the relative importance of different factors that affect the reliability of free-energy calculations when applied to estimating partition coefficients. The aim of the current study, therefore, is to build on

* To whom correspondence should be addressed at the Laboratorium für Physikalische Chemie, ETH Zentrum, Universtätstrasse 6, CH-8092 Zürich, Switzerland. Phone: 41-1-632 5501. Fax: 41-1-632 1039. E-mail: wfvgn@igc.phys.chem.ethz.ch.

[†] Universitat Autònoma de Barcelona.

[‡] ETH Zürich.

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1996.

(1) Yang, Z.; Robb, D. A. *Biotechnol. Bioeng.* **1994**, *43*, 365−370.

(2) Yang, W.-Y.; Lin, C.-D.; Chu, I.-M.; Lee, C.-J. *Biotechnol. Bioeng.* **1994**, *43*, 439−445.

(3) Ariga, O.; Miyakawa, I.; Aota, T.; Sano, Y. *J. Ferment. Bioeng.* **1994**, *77*, 71−74.

(4) Richieri, G. V.; Ogata, R. T.; Kleinfeld, A. M. *J. Biol. Chem.* **1995**, *270*, 15076−15084.

(5) Elling, L.; Kula, M. R.; Hadas, E.; Katchalski-Katzir, E. *Anal. Biochem.* **1991**, *192*, 74−77.

(6) Pérez-Gil, J.; Cruz, A.; Casals, C. *Biochem. Biophys. Acta* **1993**, *1168*, 261−270.

(7) Nakano, T.; Fink, A. L. *J. Biol. Chem.* **1990**, *265*, 12356−12362.

(8) Arnold, F. H. *TIBTECH* **1990**, *8*, 244−249.

(9) Bodor, N. S.; Gabanyi, Z.; Wong, C.-K. *J. Am. Chem. Soc.* **1989**, *111*, 3783−3786.

(10) Warne, M. J.; Connell, D. W.; Hawker, D. W.; Schueuermann, G. *Chemosphere* **1990**, *21*, 877−888.

(11) Klopman, G.; Wang, S. *J. Comput. Chem.* **1991**, *12*, 1025−1032.

(12) Moriguchi, I.; Hirono, S.; Lui, Q.; Nakagome, I.; Matsushita, Y. *Chem. Pharm. Bull.* **1992**, *40*, 127−130.

(13) Niemi, G. J.; Basak, S. C.; Veith, G. D.; Grunwald, G. *Environ. Toxicol. Chem.* **1992**, *11*, 893−900.

(14) Eiteman, M. A.; Hassinen, C.; Veide, A. *Biotechnol. Prog.* **1994**, *10*, 513−519.

(15) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683−1686.

(16) Essex, J. W.; Reynolds, C. A.; Richards, W. G. *J. Chem. Soc., Chem. Commun.* **1989**, *16*, 1152−1154.

S0002-7863(95)03794-2 CCC: $12.00

Table 1. Experimental Partition Coefficients between Chloroform and Water for the Three Studied Molecules

molecule[a]	$\log K_D(CHCl_3 \rightarrow H_2O)$[b]	$\Delta G(CHCl_3 \rightarrow H_2O)$[c] (kJ/mol)
Trp	−2.24	12.6
Ac-Trp	−1.04	5.8
Trp-NMe	−0.97	5.4

[a] The abbreviations correspond to 3-methylindole (Trp), *N*-acetyltryptamine (Ac-Trp), and 3-(3′-indolyl)propionic acid *N*-methylamide (Trp-NMe). [b] Experimental data from Radzicka and Wolfenden.[20] [c] Free energies of transfer calculated as $\Delta G = -RT \ln K_D$, with $T = 293$ K.

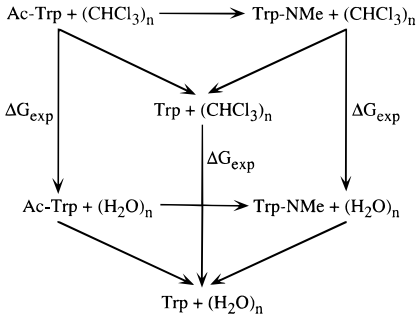

Figure 1. Thermodynamic cycles under study. The abbreviations correspond to *N*-acetyltryptamine (Ac-Trp), 3-(3′-indolyl)propionic acid *N*-methylamide (Trp-NMe), and 3-methylindole (Trp). The change in free energy related to each of the vertical arrows is known as derived from experimental partition coefficients of the three compounds between chloroform and water.

previous work by investigating the ability of free-energy calculations to predict partition coefficients for a case not readily treated by empirical methods.

The water accessibility of tryptophan is commonly used to follow the process of protein folding or unfolding.[7,17] Despite this, tryptophan may be found exposed to solvent in folded proteins.[18,19] In a number of respects, tryptophan displays an unusual partitioning behavior. Despite its aromatic character, the side chain of tryptophan is appreciably soluble in water. Its position in hydrophobicity scales based on water/organic solvent partition coefficients depends strongly on the reference organic solvent.[20] This ambiguous character makes tryptophan and its derivatives a challenging test case for theoretical approaches designed to predict partition coefficients.

In this paper we compare experimentally determined water/chloroform partition coefficients for three indole derivative models of the amino acid tryptophan to free-energy differences calculated using molecular dynamics simulations. The three derivatives studied were 3-methylindole (Trp), *N*-acetyltryptamine (Ac-Trp) and 3-(3′-indolyl)propionic acid *N*-methylamide (Trp-NMe). The experimentally determined partition coefficients as well as the corresponding differences in free energy between chloroform and water are given in Table 1. The thermodynamic cycles used to relate the calculated free-energy differences to the experimental partition coefficients are shown in Figure 1. To investigate the system in detail, the effect of the choice of force field, the choice of pathway along which the free-energy change is calculated, the inclusion of free-energy contributions from constraints, and the treatment of long-range interactions on the ability of the calculations to reproduce the experimental results in this system are each considered. It is demonstrated that, although the compounds are neutral and do not differ much in dipole moment or charge distribution, the incorporation of a reaction field to treat long-range electrostatic interactions is necessary to reproduce the experimentally observed trends. The implications of these findings for free-energy calculations in general and for the estimation of partition coefficients in particular are discussed.

Methods

Theory. (1) Free-energy Differences by Thermodynamic Integration. The difference in free energy between two states, A and B, of a molecular system, with interaction potential energy functions denoted by $V_A(\mathbf{r})$ and $V_B(\mathbf{r})$, may be expressed as[21]

$$\Delta G_{BA} = G_B - G_A = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial V(\mathbf{r};\lambda)}{\partial \lambda} \right\rangle_{\lambda'} d\lambda' \qquad (1)$$

where $\mathbf{r}$ is the ($3N$-dimensional) position vector of the N atoms of the system, and $V(\mathbf{r};\lambda)$ is a potential energy function parametrized by the coupling variable λ, continuous in λ, and satisfying

$$V(\mathbf{r};\lambda_A) = V_A(\mathbf{r}) \quad \text{and} \quad V(\mathbf{r};\lambda_B) = V_B(\mathbf{r}) \qquad (2)$$

The angular brackets, $\langle ... \rangle_{\lambda'}$, denote averaging over an equilibrium ensemble generated with the potential energy function $V(\mathbf{r};\lambda')$.

If molecular dynamics simulations are used to generate the ensembles of molecular configurations, the integral can be evaluated in different ways. In the slow-growth method, the coupling parameter λ is made a function of time and slowly changed from λ_A to λ_B during the course of a simulation. The ensemble is then reduced to a single conformation per λ' point, and the integral in (1) is approximated by a sum.[21] Alternatively, the ensemble average in (1) may be evaluated at a number of discrete λ' points by performing a separate simulation at each chosen λ' point. The integral is then determined numerically.[21] In this work, unless stated otherwise, the term thermodynamic integration will refer to this latter approach.

(2) Free-Energy Differences by Thermodynamic Perturbation. An alternative to the thermodynamic-integration formula (1) is to compute the difference in free energy between two states of a system directly, using the so-called free-energy perturbation formula.[22] It can be shown that

$$\Delta G_{BA} = -k_B T \ln \langle \exp\{-[V(\mathbf{r};\lambda_B) - V(\mathbf{r};\lambda_A)]/k_B T\} \rangle_{\lambda_A} \qquad (3)$$

Although formally exact, the ensemble average in (3) converges slowly unless the difference between the states described by $V(\mathbf{r};\lambda_A)$ and $V(\mathbf{r};\lambda_B)$ is small.

(3) λ Dependence of the Potential Energy Function. The dependence on λ of the bond, bond angle, improper dihedral angle, and torsional dihedral angle terms is essentially linear and has been described in detail elsewhere.[23] When molecular dynamics simulations are used to calculate the change in free energy involved in the creation or annihilation of van der Waals or electrostatic interaction sites, special care has to be taken to choose the λ dependence of the potential energy function for nonbonded atom-pair interactions. The presence of a singularity in this function at interparticle distances r_{ij} close to 0, which become accessible when λ approaches 0 (creation) or when λ approaches 1 (annihilation), can lead to numerical instabilities in the integration of the equations of motion. This problem can be overcome by the use of a so-called soft-core λ-dependent potential energy function

(17) Khorasanizadeh, S.; Peters, I. D.; Butt, T. R.; Roder, H. *Biochemistry* **1993**, *32*, 7054−7063.
(18) Czurylo, E. A.; Emelyanenko, V. I.; Permyakov, E. A.; Dabrowska, R. *Biophys. Chem.* **1991**, *40*, 181−188.
(19) Chabbert, M.; Hillen, W.; Hansen, D.; Takahashi, M.; Bousquet, J. A. *Biochemistry* **1992**, *31*, 1951−1960.
(20) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664−1670.
(21) van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. In *Computer simulation of biomolecular systems, theoretical and experimental applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1993; Vol. 2, pp 315−348.
(22) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(23) van Gunsteren, W. F. *Protein Eng.* **1988**, *2*, 5−13.

for the nonbonded interactions.[24,25] In the soft-core function used in this work the van der Waals and electrostatic interactions of atoms i and j at a distance r_{ij} are given by[24]

$$V_{ij}^{\mathrm{nb}}(r_{ij};\lambda) = (1-\lambda)^n\left[\frac{C12_{ij}^{\mathrm{A}}}{(\alpha_{ij}^{\mathrm{LJ}}(\sigma_{ij}^{\mathrm{A}})^6\lambda^2 + \mathrm{r}_{ij}^{\ 6})^2} - \frac{C6_{ij}^{\mathrm{A}}}{\alpha_{ij}^{\mathrm{LJ}}(\sigma_{ij}^{\mathrm{A}})^6\lambda^2 + r_{ij}^{\ 6}} + \frac{q_i^{\mathrm{A}}q_j^{\mathrm{A}}}{4\pi\epsilon_0\epsilon_1(\alpha_{ij}^{\mathrm{C}}\lambda^2 + r_{ij}^{\ 2})^{1/2}}\right] + \lambda^n\left[\frac{C12_{ij}^{\mathrm{B}}}{(\alpha_{ij}^{\mathrm{LJ}}(\sigma_{ij}^{\mathrm{B}})^6(1-\lambda)^2 + r_{ij}^{\ 6})^2} - \frac{C6_{ij}^{\mathrm{B}}}{\alpha_{ij}^{\mathrm{LJ}}(\sigma_{ij}^{\mathrm{B}})^6(1-\lambda)^2 + r_{ij}^{\ 6}} + \frac{q_i^{\mathrm{B}}q_j^{\mathrm{B}}}{4\pi\epsilon_0\epsilon_1(\alpha_{ij}^{\mathrm{C}}(1-\lambda)^2 + r_{ij}^{\ 2})^{1/2}}\right] \quad (4)$$

where $C6_{ij}$ and $C12_{ij}$ are the van der Waals parameters for the dispersion and repulsion energy terms of the interaction, respectively, q_i and q_j are the atomic charges, and $\sigma_{ij}^{\ 6}$, n, $\alpha_{ij}^{\mathrm{LJ}}$, and α_{ij}^{C} are set, in this work, to

$$\sigma_{ij}^{\ 6} = \frac{C12_{ij}}{C6_{ij}} \quad (C6_{ij} \neq 0 \text{ and } C12_{ij} \neq 0) \quad (5)$$

$$n = 1$$

$$\alpha_{ij}^{\mathrm{LJ}} = \begin{cases} 0.5 & \text{if } i \text{ or } j \text{ perturbed} \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{ij}^{\mathrm{C}} = \begin{cases} 0.5\ \mathrm{nm}^2 & \text{if } i \text{ or } j \text{ perturbed} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

For the interaction between pairs of unperturbed atoms ($\alpha_{ij}^{\mathrm{LJ}} = 0$ and $\alpha_{ij}^{\mathrm{C}} = 0$) the standard form of the nonbonded term of the interaction is recovered. For interactions involving one or two perturbed atoms the standard form of the interaction is recovered only at $\lambda = 0$ and $\lambda = 1$.

(4) Reaction-Field Contribution to the Electrostatic Interaction. When periodic boundary conditions are applied in molecular dynamics simulations, the electrostatic interactions are often truncated at a cutoff distance R_c. Interactions beyond this cutoff are normally neglected. This is equivalent to a vacuum existing outside the cutoff sphere. The error introduced by this approximation in the calculations will be dependent on the dielectric permittivity of the system and the property of interest. A better approximation is to treat the medium outside the cutoff sphere as a dielectric continuum generating a reaction-field potential.[26−29] To study the sensitivity of free-energy calculations to the use of a cutoff radius for the electrostatic interactions, we have performed simulations with and without a reaction-field correction. If there is no ionic strength in the system, the electrostatic potential energy including the reaction-field correction for the interaction between atoms i and j is given by

$$V_{ij}^{\mathrm{el}}(r_{ij}) = \frac{q_iq_j}{4\pi\epsilon_0\epsilon_1}\left(\frac{1}{r_{ij}} + \frac{r_{ij}^{\ 2}}{R_c^{\ 3}}\frac{\epsilon_2 - \epsilon_1}{2\epsilon_2 + \epsilon_1}\right) \quad (7)$$

where R_c is the cutoff radius, ϵ_0 is the permittivity of vacuum, ϵ_1 is the relative permittivity within the sphere of radius R_c centred in atom i (i.e., $\epsilon_1 = 1$), and ϵ_2 is the permittivity of the continuum beyond R_c. The reaction-field approach assumes that interactions beyond the cutoff distance are equivalent to those of the bulk solvent. Therefore, the cutoff should be large enough that entropic effects observed beyond R_c are negligible, and the permittivity of the continuum ϵ_2 should correspond to that of the bulk solvent model that is used. The soft-core λ-dependent electrostatic potential energy term including the reaction-field correction has the form

$$V_{ij}^{\mathrm{el}}(r_{ij};\lambda) = (1-\lambda)^n\frac{q_i^{\mathrm{A}}q_j^{\mathrm{A}}}{4\pi\epsilon_0\epsilon_1}\left[\frac{1}{(\alpha_{ij}^{\mathrm{C}}\lambda^2 + r_{ij}^{\ 2})^{1/2}} + \frac{r_{ij}^{\ 2}}{(\alpha_{ij}^{\mathrm{RF}}\lambda^2 + R_c^{\ 2})^{3/2}}\frac{\epsilon_2 - \epsilon_1}{2\epsilon_2 + \epsilon_1}\right] + \lambda^n\frac{q_i^{\mathrm{B}}q_j^{\mathrm{B}}}{4\pi\epsilon_0\epsilon_1}\left[\frac{1}{(\alpha_{ij}^{\mathrm{C}}(1-\lambda)^2 + r_{ij}^{\ 2})^{1/2}} + \frac{r_{ij}^{\ 2}}{(\alpha_{ij}^{\mathrm{RF}}(1-\lambda)^2 + R_c^{\ 2})^{3/2}}\frac{\epsilon_2 - \epsilon_1}{2\epsilon_2 + \epsilon_1}\right] \quad (8)$$

where n and $\alpha_{ij}^{\mathrm{RF}}$ are set, in this work, to $n = 1$ and $\alpha_{ij}^{\mathrm{RF}} = 0$ nm^2. Again, (8) reduces to (7) when $\alpha_{ij}^{\mathrm{C}} = 0$ and $\alpha_{ij}^{\mathrm{RF}} = 0$, when $\lambda = 0$, or when $\lambda = 1$.

(5) Constraint Contribution to the Change in Free Energy. When the SHAKE algorithm[30] is applied to constrain bond lengths in an MD simulation, after each unconstrained dynamics time step the bonded atoms are displaced in an iterative way until all constraints are satisfied. This process involves the displacement of the point of application of the forces exerted by the environment on the bonded atoms and will contribute to the change in the free energy of the system if the bond length changes with λ.[31] This contribution can be easily computed, as shown by Straatsma *et al.*[32] and van Gunsteren *et al.*[21]

(6) Estimation of Errors. To estimate the error in $\langle\partial V/\partial\lambda\rangle_{\lambda'}$, we have used its standard deviation $\sigma(\langle\partial V/\partial\lambda\rangle_{\lambda'})$, as defined by Fincham *et al.*[33,34]

To estimate the error in ΔG due to the trapezoidal integration, $\langle\partial V/\partial\lambda\rangle_{\lambda'}$ is assumed to be Gaussian distributed with a maximum probability equal to $\langle\partial V/\partial\lambda\rangle_{\lambda'}$ and a width of $\sigma(\langle\partial V/\partial\lambda\rangle_{\lambda'})$. The distributions are assumed to be uncorrelated among different λ' points. The standard deviation on ΔG is then calculated as[35]

$$\sigma(\Delta G) = \left[\sum_{n=1}^{N_\lambda} w(\lambda_n)\,\sigma^2(\langle\partial V/\partial\lambda\rangle_{\lambda_n})\right]^{1/2} \quad (9)$$

where N_λ is the number of λ' points at which the ensemble average has been calculated, $w(\lambda_n)$ is the weight factor due to the trapezoidal integration, and $\sigma(\langle\partial V/\partial\lambda\rangle_{\lambda_n})$ is the error in the ensemble averages.

Simulation Setup. All simulations were performed using a modified version of the GROMOS87 package of programs[36] in conjunction with a modified version of the GROMOS87 force field.[36−39] In this force field, aliphatic hydrogen atoms are treated as united atoms together with the carbon atom to which they are attached. Polar hydrogens are treated explicitly and experience electrostatic but no van der Waals interactions. Aromatic hydrogens are also treated explicitly and experience both electrostatic and van der Waals interactions. The force-field parameters for the nonbonded interactions that have been used are given in Table 2. The water parameters are taken from the SPC

(24) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(25) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025−9031.
(26) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486−1493.
(27) Barker, J. A.; Watts, R. O. *Mol. Phys.* **1973**, *26*, 789−792.
(28) Neumann, M. *Mol. Phys.* **1983**, *50*, 841−858.
(29) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(30) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(31) Wang, L.; Hermans, J. *J. Chem. Phys.* **1994**, *100*, 9129−9139.
(32) Straatsma, T. P.; Zacharias, M.; McCammon, J. A. *Chem. Phys. Lett.* **1992**, *196*, 297−302.
(33) Fincham, D.; Quirke, N.; Tildesley, D. J. *J. Chem. Phys.* **1986**, *84*, 4535−4546.
(34) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: New York, 1987.
(35) Spiegel, M. R. *Theory and Problems of Statistics*; McGraw-Hill: New York, 1972.
(36) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen molecular simulation (GROMOS) library manual*; Biomos, Nijenborgh 4, 9747 AG Groningen, The Netherlands, 1987.
(37) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513−1518.
(38) Mark, A. E.; van Helden, S. P.; Smith, P. E.; Janssen, L. H. M.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 6293−6302.
(39) Smith, L. J.; Mark, A. E.; Dobson, C. M.; van Gunsteren, W. F. *Biochemistry* **1995**, *34*, 10918−10931.

Table 2. Nonbonded Interaction Parameters for the Atom Types of the Force Field

molecule[a]	atom	partial charge (e)	$C6^{1/2}_{i,i}$ [(kcal mol^{-1} Å ^{6})$^{1/2}$][b,d,e]	$C12^{1/2}_{i,i}$ [(kcal mol^{-1} Å^{12})$^{1/2}$][b−e]
Trp	CH_3	0.000	46.06	2500.0
	$C\gamma$	−0.140	23.65	898.0
	$C\delta_1$, $C\epsilon_3$, $C\zeta_2$, $C\zeta_3$, $C\eta_2$	−0.100	36.30	1901.0
	$H\delta_1$, $H\epsilon_3$, $H\zeta_2$, $H\zeta_3$, $H\eta_2$	0.100	4.50	60.1
	$C\delta_2$, $C\epsilon_2$	0.000	23.65	898.0
	$N\epsilon_1$	−0.050	24.13	636.0/900.0
	$H\epsilon_1$	0.190	0.00	0.0
Ac-Trp	CH_3	0.000	46.06	2500.0
	C	0.380	23.65	898.0
	O	−0.380	23.25	421.0/550.0
	N	−0.280	24.13	636.0/950.0
	H	0.280	0.00	0.0
	C_α, C_β	0.000	46.63	2906.0
H_2O	O_w	−0.820	25.01	793.3/793.3
	H_w	0.410	0.00	0.0
$CHCl_3$	C_c	0.179	25.08	985.6
	H_c	0.082	3.00	32.1
	Cl_c	−0.087	44.56	1813.8

[a] The water parameters are taken from the SPC model.[40] The chloroform parameters are taken from the Dietz-Heinzinger model.[41−43] [b] The Lennard-Jones parameters $C6_{i,j}$ and $C12_{i,j}$ were obtained using the following combination rules: $C6_{i,j} = C6^{1/2}_{i,i}\, C6^{1/2}_{j,j}$ and $C12_{i,j} = C12^{1/2}_{i,i}\, C12^{1/2}_{j,j}$. [c] To properly model hydrogen-bonding interactions, an alternative set of $C12$ parameters, corresponding to an increased short-range repulsion, is used for interactions between hydrogen-bond donor−acceptor pairs.[36] This is the second value in the column. [d] $C6^{1/2}_{i,i}$ and $C12^{1/2}_{i,i}$ parameters for aromatic hydrogens were taken from Hermans *et al*.[37] [e] Interactions between different types of chloroform atoms have specific parameters: $C6_{C_c,Cl_c} = 1117.4$, $C6_{C_c,H_c} = 86.6$, $C6_{Cl_c,H_c} = 155.2$ (kcal mol^{-1} Å^6), and $C12_{C_c,Cl_c} = 1\,788\,074$, $C12_{C_c,H_c} = 41\,706$, $C12_{Cl_c,H_c} = 78\,059$ (kcal mol^{-1} Å^{12}).

model,[40] and the chloroform parameters are taken from the Dietz−Heinzinger model.[41−43]

Ac-Trp was placed at the center of each of two truncated octahedrons with box lengths of 3.07 and 4.99 nm for the water and chloroform systems, respectively. Trp-NMe was also placed in the center of each of two truncated octahedrons with the same box lengths. The box sizes corresponded to a closest distance of a solute atom to the box wall of about 1.0 nm for the water systems and 2.0 nm for the chloroform systems, or approximately four solvation shells. The boxes were filled with an equilibrium distribution of the appropriate solvent. Solvent molecules that lie outside the box or which overlapped with a non-hydrogen solute atom (distance solute−water < 0.23 nm and distance solute−chloroform < 0.30 nm) were removed. In each of the four cases, the final systems included 473 solvent molecules. Each system was first energy minimized to obtain the starting structure for the dynamics calculations. All calculations were performed using periodic boundary conditions at a temperature of 300 K and a pressure of 1 atm. The temperature and the pressure were maintained by weak coupling to an external bath,[44] using a relaxation time of 0.1 and 0.5 ps, respectively, for the simulations in water, and 0.1 and 0.2 ps, respectively, for the simulations in chloroform. Bond lengths were constrained using the SHAKE algorithm[30] with a geometric tolerance of 10^{-4}. The integration time step was 2 fs. A cutoff radius of 1.4 nm was used for all nonbonded interactions. The nonbonded interactions were calculated at each time step using a charge-group pair list that was updated every 10 time steps. Each system was equilibrated for 50 ps before starting the free-energy calculations. The relative permittivities of bulk water and chloroform were taken as those calculated for the solvent models used, that is, $\epsilon_2 = 54.0$ for SPC water,[45] and $\epsilon_2 = 2.4$ for Dietz−Heinzinger chloroform.[43]

Perturbations. To estimate the differences between the free energies of transfer from chloroform to water of Trp-NMe, Ac-Trp, and Trp, mutations between the four states shown in Figure 2 have been performed in both solvents. The calculated free energies can be related

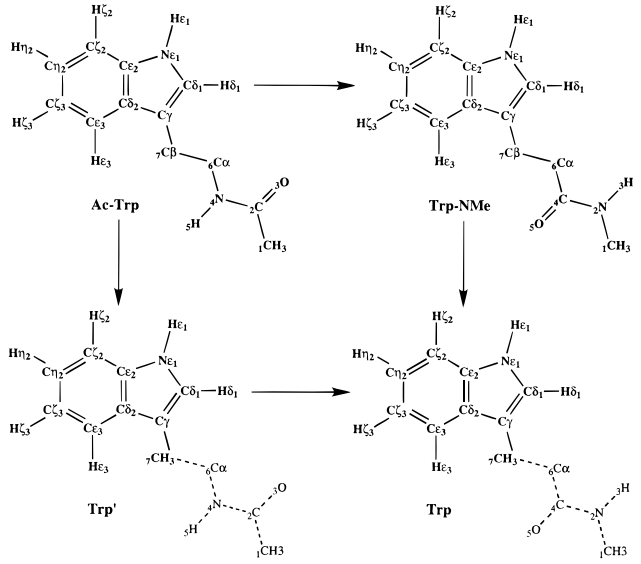

Figure 2. Molecular models that have been used to define the different states in the perturbation cycle. The nonaromatic atoms are numbered from 1 to 7. Atoms 1−6 in Trp and Trp′ preserve their original mass and parameters for bond, bond angle, improper dihedral angle, and torsional dihedral angle interactions, but experience no van der Waals or electrostatic interactions.

to the experimental partition coefficients using the thermodynamic cycles shown in Figure 1. The mutations Trp-NMe to Trp and Ac-Trp to Trp′ involve only changes in nonbonded parameters. The interactions of atoms 1−6 with the rest of the system are removed, and atom 7 is changed from a $-CH_2$ to a $-CH_3$. The mutation Ac-Trp to Trp-NMe involves changes in bonded and nonbonded parameters. It should be noted that although atoms 1−6 from Trp and Trp′ do not have nonbonded interactions, they are still coupled to the system via bonded terms. For this reason, the change in the free energy for the mutation Trp′ to Trp must be determined to obtain closure of the thermodynamic cycle.[38]

For each mutation shown in Figure 2 slow-growth calculations in water and chloroform (Table 3) were used to generate the initial structures for the simulations at different λ' values, and the shape of the resulting free-energy curves was used to guide the choice of λ' points at which ensemble averages were calculated for thermodynamic integration. At each λ' value, the first 10 ps of the simulation was

(40) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331−342.
(41) Dietz, W.; Heinzinger, K. *Ber. Bunsen-Ges. Phys. Chem.* **1984**, *88*, 543−546.
(42) Dietz, W.; Heinzinger, K. *Ber. Bunsen-Ges. Phys. Chem.* **1985**, *89*, 968−977.
(43) Tironi, I. G.; van Gunsteren, W. F. *Mol. Phys.* **1994**, *83*, 381−403.
(44) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(45) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169−3174.

Table 3. Summary of Slow-Growth Calculations[a]

mutation	solvent	$V^{nb}(\lambda)$[b]	ΔG (kJ/mol)	mutation	solvent	$V^{nb}(\lambda)$[b]	ΔG (kJ/mol)
Trp-NMe → Trp	H_2O	soft-core	70.0	Trp-NMe → Trp	$CHCl_3$	soft-core	82.9
Ac-Trp → Trp′	H_2O	soft-core	75.1	Ac-Trp → Trp′	$CHCl_3$	soft-core	77.2
Trp′ → Trp	H_2O	standard	0.3	Trp′ → Trp	$CHCl_3$	standard	0.2
Ac-Trp → Trp-NMe	H_2O	standard	7.1	Ac-Trp → Trp-NMe	$CHCl_3$	standard	1.2
Ac-Trp → Trp-NMe	H_2O	soft-core	2.3	Ac-Trp → Trp-NMe	$CHCl_3$	soft-core	3.3

[a] The molecular models represented by the abbreviations Trp-NMe, Ac-Trp, Trp, and Trp′ are shown in Figure 2. [b] λ-dependent potential energy function that is used for the nonbonded interactions. The nonbonded interaction parameters are identical to those of set 1 in Table 4. The slow-growth MD simulations are 100 ps long, with a λ step of 2×10^{-5}.

Table 4. Summary of Thermodynamic Integration Calculations[a]

set	mutation	solvent	V^{el} [b]	$C12^{1/2}_{O_w,O_w}$ [c] [(kcal mol^{-1} Å^{12})$^{1/2}$]	ΔG[d] (kJ/mol)
1	Trp-NMe → Trp	H_2O	standard	793.3	70.3 ± 2.1
1	Ac-Trp → Trp′	H_2O	standard	793.3	69.8 ± 2.3
1	Ac-Trp → Trp-NMe	H_2O	standard	793.3	−0.1 ± 1.9
1	Trp-NMe → Trp	$CHCl_3$	standard		73.2 ± 2.1
1	Ac-Trp → Trp′	$CHCl_3$	standard		73.3 ± 2.2
1	Ac-Trp → Trp-NMe	$CHCl_3$	standard		2.5 ± 1.9
2	Ac-Trp → Trp′	H_2O	standard	421.0	76.3 ± 2.3
3	Ac-Trp → Trp′	H_2O	standard	690.0	71.5 ± 2.5
4	Ac-Trp → Trp′	H_2O	RF correction	793.3	74.1 ± 2.1
5	Trp-NMe → Trp	H_2O	RF correction	690.0	74.0 ± 2.2
5	Ac-Trp → Trp′	H_2O	RF correction	690.0	75.9 ± 2.1
5	Ac-Trp → Trp-NMe	H_2O	RF correction	690.0	2.3 ± 1.8
5	Trp-NMe → Trp	$CHCl_3$	RF correction		72.6 ± 2.1
5	Ac-Trp → Trp′	$CHCl_3$	RF correction		73.6 ± 2.0
5	Ac-Trp → Trp-NMe	$CHCl_3$	RF correction		2.5 ± 2.0
	Trp′ → Trp	vacuum	standard		−0.2 ± 0.1

[a] The perturbations are grouped in sets according to the parameters given in the table. The molecular models represented by the abbreviations Trp-NMe, Ac-Trp, Trp, and Trp′ are shown in Figure 2. The thermodynamic integration method has been used in all cases, with ensemble averages $\langle \partial V/\partial\lambda \rangle_{\lambda'}$, at 13 λ' points between 0 and 1 (21 λ' points in the vacuum case). The simulation time at each λ' point ranges from 50 to 200 ps depending on the convergence properties of $\langle \partial V/\partial\lambda \rangle_{\lambda'}$. The soft-core λ-dependent potential energy function has been used in all but the vacuum case. [b] Potential energy function that is used for the electrostatic interactions; RF refers to reaction field. [c] $C12^{1/2}_{O_w,O_w}$ van der Waals parameter that is used for the interactions of the water oxygen with nonpolar atoms (see Table 2). [d] ΔG is calculated by trapezoidal integration.

discarded (equilibration). The sampling time at each λ' value was chosen according to the convergence properties of $\langle \partial V/\partial\lambda \rangle_{\lambda'}$, such that the error in the ensemble average was roughly uniform over all λ' points. To study the influence of different treatments of the nonbonded interactions on the calculated free-energy differences, various sets of perturbations were performed with the conditions summarized in Table 4. For the first set of perturbations (set 1), the initial coordinates and velocities for the simulations at each λ' value were taken from the corresponding λ' points in the slow-growth calculations. For the perturbation sets 2−5, the initial coordinates were taken from the corresponding λ' points in set 1 after a further 50 ps equilibration.

Results

Table 3 shows the results from the slow-growth calculations. These comprise the mutations described in Figure 2, performed both in water and in chloroform. The perturbation Ac-Trp to Trp-NMe was performed using the standard λ dependence of the GROMOS87 potential energy function[23,36] (linear for nonbonded interactions) and using the soft-core λ-dependent potential energy function (4). To compare the results from the use of these two coupling schemes, plots of $\partial V/\partial\lambda$ and ΔG versus λ are shown for the mutation in water and in chloroform in Figure 3.

Table 4 shows the results from the thermodynamic-integration calculations. Five different variants of the force field have been used to perform the calculations (sets 1−5). All included a soft-core λ-dependent potential energy function for the evaluation of nonbonded interactions. In set 1, the parameter $C12^{1/2}_{O_w,O_w}$ = 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ has been used for interactions between water oxygens and nonpolar atoms (Table 2). This value has been recently shown to be more appropriate than the value 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ of the original force-field parameter set.[38,39,46−49] In set 2, the value 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ has been used to evaluate the effect of this parameter on the change in free energy for the mutation Ac-Trp to Trp′ in water. Due to the sensitivity of the calculated change in free energy to the choice of the $C12^{1/2}_{O_w,O_w}$ parameter for water−nonpolar carbon interactions this term was reparametrized to correctly reproduce the free energy of hydration of CH_4 (Figure 4). The value obtained was 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$, and was validated by determining the free energies of hydration of a series of aliphatic compounds (Table 5). In set 3 the value 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$ has been used for the mutation Ac-Trp to Trp′ in water.

To determine the sensitivity of the calculated free energies to the truncation of the electrostatic interactions at a given cutoff distance, the long-range electrostatic interactions were evaluated using a reaction-field potential. The effect of the reaction-field correction was determined for the mutation Ac-Trp to Trp′ in water (set 4) and in chloroform (set 5). The $C12^{1/2}_{O_w,O_w}$ parameter in set 4 was set to 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ to enable comparison with set 1. In set 5, the complete cycles in water and in chloroform were performed. The parameter $C12^{1/2}_{O_w,O_w}$ = 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$ was used for the simulations in water, and the reaction-field correction to the electrostatic potential was applied in all cases. The change in the free energy for the mutation Trp′ to Trp was determined from vacuum

(46) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385−391.
(47) van Buuren, A. R.; Marrink, S.-J.; Berendsen, H. J. C. *J. Phys. Chem.* **1993**, *97*, 9206−9212.
(48) van Buuren, A. R.; Berendsen, H. J. C. *Biopolymers* **1993**, *33*, 1159−1166.
(49) Daura, X.; Oliva, B.; Querol, E.; Avilés, F. X.; Tapia, O. *Proteins: Struct., Funct. Genet.* **1996**, *25*, 89−103.

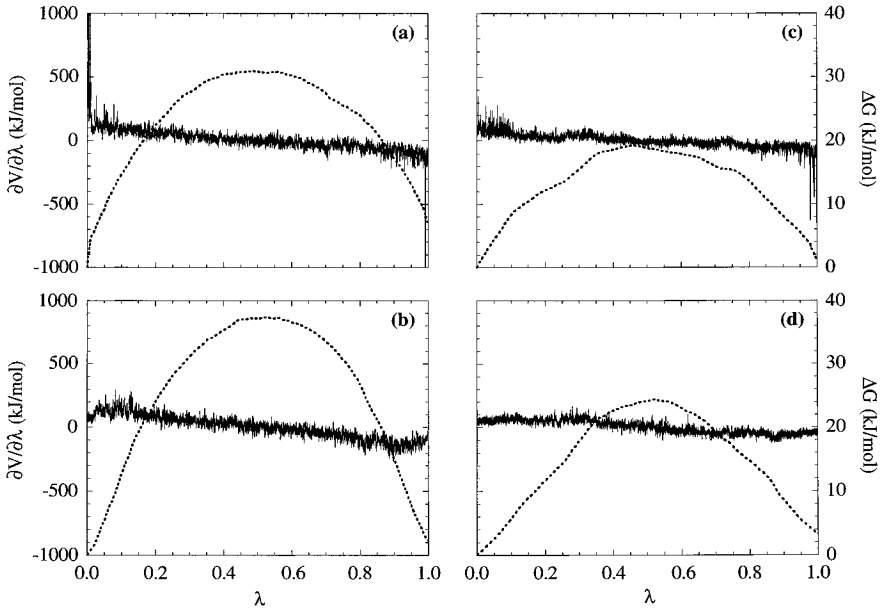

Figure 3. $\partial V/\partial\lambda$ and ΔG as functions of the coupling parameter λ from different slow-growth simulations of the mutation Ac-Trp to Trp-NMe. λ changes from 0 to 1 in 100 ps. The solid line corresponds to $\partial V/\partial\lambda$. The dotted line corresponds to ΔG. (a) Mutation in water, using the standard λ-dependent potential energy function. (b) Mutation in water, using the soft-core λ-dependent potential energy function. (c) Mutation in chloroform, using the standard λ-dependent potential energy function. (d) Mutation in chloroform, using the soft-core λ-dependent potential energy function. The truncated peaks of $\partial V/\partial\lambda$ in plot a go up to 2000 kJ/mol at λ' values close to 0 and down to −1500 kJ/mol at λ' values close to 1.

Table 5. Comparison between Experimental and Calculated Free Energies of Hydration for a Series of Aliphatic Compounds

compound[a]	ΔG_{hyd} (exptl) (kJ/mol)[b]	ΔG_{hyd} (calcd) (kJ/mol)[c]	compound[a]	ΔG_{hyd} (exptl) (kJ/mol)[b]	ΔG_{hyd} (calcd) (kJ/mol)[c]
methane	8.37	8.30	butane	8.70	6.81
ethane	7.66	7.38	isobutane	9.70	5.06
propane	8.18	7.44			

[a] Aliphatic hydrogen atoms are treated as united atoms together with the carbon atom to which they are attached.[54] The interactions of the united atoms with water are purely van der Waals. [b] Values obtained from Cabani *et al.*[55] [c] $C12^{1/2}_{O_w,O_w} = 690.0$ (kcal mol^{-1} Å^{12})$^{1/2}$ has been used for van der Waals interactions between water and aliphatic united atoms. The compound is initially placed, with no interactions, in a box of 512 SPC waters equilibrated at 1 atm and 300 K. Then the compound is grown using the soft-core λ-dependent potential energy function, with a cutoff radius for van der Waals interactions of 0.9 nm. $\langle\partial V/\partial\lambda\rangle_{\lambda'}$, is calculated at 23 λ' points (21 for methane), with 50 ps sampling. ΔG_{hyd} is calculated by trapezoidal integration.

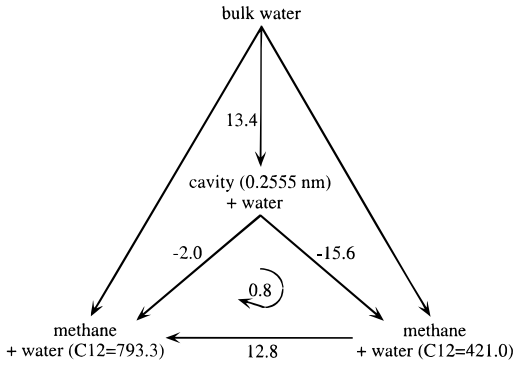

Figure 4. Extrapolation of the parameter $C12^{1/2}_{O_w,O_w}$ for water−methane interactions. A cavity in water of diameter 0.2555 nm has been used as a reference state. The free energy of creation of this cavity in bulk water has been shown to be 13.4 kJ/mol.[56] Methane is modeled as a united atom,[54] with van der Waals but no electrostatic interactions with water. The calculated free energy of hydration of methane is −2.2 kJ/mol when the value 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ is used, and 11.4 kJ/mol when the value 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ is used. The experimental free energy of hydration is 8.37 kJ/mol.[55] If we assume that $\Delta G_{hyd}(C12^{1/2}_{O_w,O_w})$ is a linear function, the experimental $\Delta G_{hyd} = 8.37$ kJ/mol corresponds approximately to $C12^{1/2}_{O_w,O_w} = 690.0$ (kcal mol^{-1} Å^{12})$^{1/2}$.

simulations. The mutated atoms (2−5 in Figure 2) do not interact with the solvent and are, thus, unaffected by it.

Table 6 shows, for each of the perturbations in set 5 and for the mutation Trp′ to Trp, $\langle\partial V/\partial\lambda\rangle_{\lambda'}$ calculated at 13 λ' points. ΔG was calculated by trapezoidal integration. To examine the dependence of ΔG on the integration method, the integrals were also determined using a cubic-spline function to interpolate between the (λ', $\langle\partial V/\partial\lambda\rangle_{\lambda'}$) points. The difference in ΔG obtained with the two integration methods was in all cases less than 0.3 kJ/mol, much smaller than the estimated errors. The (λ', $\langle\partial V/\partial\lambda\rangle_{\lambda'}$) points and corresponding error bars are plotted in Figure 5.

Table 7 shows a comparison between the experimental and calculated $\Delta\Delta G$ values. $\Delta\Delta G_1$ and $\Delta\Delta G_2$ correspond to the closed thermodynamic cycles for the mutations in water and in chloroform, respectively. $\Delta\Delta G_3$, $\Delta\Delta G_4$, and $\Delta\Delta G_5$ are the differences between the change in free energy in water and in chloroform for the mutations Trp-NMe to Trp, Ac-Trp to Trp′, and Ac-Trp to Trp-NMe, respectively. These should equal the corresponding differences between the free energies of transfer from chloroform to water of each pair of molecules involved in the mutations (see the cycles in Figure 1). $\Delta\Delta G_1$ and $\Delta\Delta G_2$ are indicative of the precision of the calculations. $\Delta\Delta G_3$, $\Delta\Delta G_4$, and $\Delta\Delta G_5$ are indicative of the ability of the molecular model and force field to reproduce the properties of the real system.

Discussion

(1) λ Dependence of the Potential Energy Function: Standard versus Soft-Core. The use of a soft-core interaction has been shown to improve the efficiency of free-energy calculations in cases which involve the creation or annihilation of exposed atoms.[24,25] By removing the singularity in the

Table 6. Results from Set 5 Calculations

	water						chloroform						vacuum	
	Trp-NMe → Trp		Ac-Trp → Trp′		Ac-Trp → Trp-NMe		Trp-NMe → Trp		Ac-Trp → Trp′		Ac-Trp → Trp-NMe		Trp′ → Trp[c]	
λ	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)	simulation time (ps)	$\langle\partial V/\partial\lambda\rangle^a$ (kJ/mol)
0.00	50	128.5 ± 1.9	50	131.7 ± 1.1	50	74.6 ± 1.1	50	115.9 ± 1.1	50	115.8 ± 1.0	50	46.2 ± 0.9	100	−0.14 ± 0.05
0.05	100	195.3 ± 2.3	100	188.2 ± 2.3	100	138.2 ± 2.2	100	124.6 ± 1.0	100	136.0 ± 1.2	50	58.9 ± 2.0	100	0.05 ± 0.07
0.10	100	194.4 ± 2.6	100	196.7 ± 2.8	125	148.9 ± 2.5	100	145.0 ± 1.4	100	137.6 ± 1.2	50	69.9 ± 1.6	100	−0.07 ± 0.06
0.20	100	147.2 ± 1.8	100	148.8 ± 1.5	100	100.7 ± 1.7	100	127.2 ± 1.7	100	138.5 ± 1.4	50	67.8 ± 3.1	100	−0.05 ± 0.04
0.30	50	106.8 ± 1.5	50	112.4 ± 1.1	50	62.3 ± 1.5	50	110.3 ± 1.4	50	112.5 ± 1.3	50	51.5 ± 2.2	100	0.02 ± 0.05
0.40	50	79.4 ± 1.5	50	87.9 ± 1.3	50	34.2 ± 1.1	50	85.5 ± 2.4	50	92.5 ± 2.0	50	30.9 ± 1.9	100	−0.15 ± 0.03
0.50	50	58.5 ± 1.4	50	67.4 ± 1.9	50	3.6 ± 1.4	50	71.5 ± 1.4	50	73.7 ± 2.3	50	2.5 ± 1.8	100	−0.11 ± 0.04
0.60	50	45.3 ± 2.0	50	48.1 ± 2.1	50	−24.0 ± 1.2	50	57.1 ± 1.6	50	59.5 ± 2.1	50	−15.8 ± 1.8	100	−0.16 ± 0.04
0.70	100	20.9 ± 2.4	100	27.3 ± 2.0	50	−58.6 ± 2.4	100	41.7 ± 2.1	100	42.5 ± 1.7	50	−46.2 ± 1.8	100	−0.17 ± 0.04
0.80	200	−1.5 ± 3.4	200	−10.5 ± 3.7	100	−96.6 ± 1.5	150	17.5 ± 3.2	150	9.1 ± 3.1	50	−64.0 ± 2.0	100	−0.49 ± 0.11
0.90	100	−10.3 ± 2.7	100	−12.5 ± 2.2	125	−142.1 ± 2.3	150	−1.9 ± 3.1	150	−6.4 ± 2.9	50	−68.6 ± 1.8	100	−0.55 ± 0.15
0.95	100	14.4 ± 1.6	100	9.4 ± 1.7	125	−145.2 ± 1.7	125	13.5 ± 2.3	125	9.3 ± 2.1	50	−64.0 ± 1.3	100	−0.43 ± 0.14
1.00	50	34.3 ± 1.2	50	29.1 ± 1.8	50	−77.7 ± 1.2	50	38.2 ± 2.0	50	32.5 ± 2.2	50	−45.6 ± 0.8	100	−1.02 ± 0.21
ΔG (kJ/mol)[b]		74.0 ± 2.2		75.9 ± 2.1		2.3 ± 1.8		72.6 ± 2.1		73.6 ± 2.0		2.5 ± 2.0		−0.21 ± 0.09

[a] The first 10 ps of simulation at each λ' point is not considered for the calculation of the average. [b] ΔG is calculated by trapezoidal integration. [c] This column shows 13 of the 21 λ' points used for the calculation of the integral.

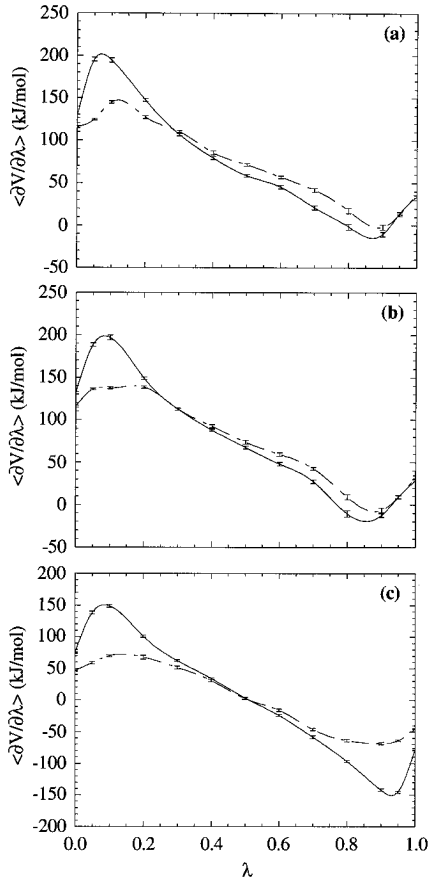

Figure 5. $\langle\partial V/\partial\lambda\rangle$ plotted against λ for the perturbations from set 5 (see Table 6). Solid lines: cubic spline interpolation of the (λ', $\langle\partial V/\partial\lambda\rangle_{\lambda'}$) points for the perturbations in water. Broken lines: cubic spline interpolation of the (λ', $\langle\partial V/\partial\lambda\rangle_{\lambda'}$) points for the perturbations in chloroform. (a) Trp-NMe to Trp. (b) Ac-Trp to Trp′. (c) Ac-Trp to Trp-NMe.

Table 7. Comparison between Expected and Calculated $\Delta\Delta G$ Values[a]

	$\Delta\Delta G$ (kJ/mol)		
	expected	set 1	set 5
$\Delta\Delta G_1 = \Delta G_{H_2O}$(Ac-Trp → Trp-NMe) + ΔG_{H_2O}(Trp-NMe → Trp) − ΔG_{H_2O}(Ac-Trp → Trp′) − ΔG_{H_2O}(Trp′ → Trp)	0.0	0.6	0.6
$\Delta\Delta G_2 = \Delta G_{CHCl_3}$(Ac-Trp → Trp-NMe) + ΔG_{CHCl_3}(Trp-NMe → Trp) − ΔG_{CHCl_3}(Ac-Trp → Trp′) − ΔG_{CHCl_3}(Trp′ → Trp)	0.0	2.6	1.7
$\Delta\Delta G_3 = \Delta G_{H_2O}$(Trp-NMe → Trp) − ΔG_{CHCl_3}(Trp-NMe → Trp)	7.2	−2.9	1.4
$\Delta\Delta G_4 = \Delta G_{H_2O}$(Ac-Trp → Trp′) − ΔG_{CHCl_3}(Ac-Trp → Trp′)	6.8	−3.5	2.3
$\Delta\Delta G_5 = \Delta G_{H_2O}$(Ac-Trp → Trp-NMe) − ΔG_{CHCl_3}(Ac-Trp → Trp-NMe)	−0.4	−2.6	−0.2

[a] The first two expected values, $\Delta\Delta G_1$ and $\Delta\Delta G_2$, are purely theoretical. They are based on closure of the cycles in Figure 1. $\Delta\Delta G_3$, $\Delta\Delta G_4$, and $\Delta\Delta G_5$ are derived from the experimental partition coefficients in Table 1.

potential energy function at short interaction distances, it is possible to create or annihilate atoms without first ensuring that they lie within the van der Waals radius of another atom. Of the four mutations shown in Figure 2, Ac-Trp to Trp′ and Trp-NMe to Trp are clear candidates for the use of such a potential energy function. In contrast, the mutation Ac-Trp to Trp-NMe does not involve annihilation or creation of exposed atoms. Two groups, −NH and −CO, are interconverted (see Figure 2).

Although the −NH hydrogen has no van der Waals interactions, this atom always lies well within the van der Waals radius of the atom to which it is attached. To investigate whether in such a case the use of a soft-core λ-dependent interaction improves the efficiency of the simulation, yielding stable calculations at λ' values close to 0 and 1, slow-growth calculations were performed for this mutation using both the standard and the soft-core potential energy functions. In Figure 3 it is seen that when the calculations were performed in water and the standard λ-dependent energy function was used (a) $\partial V/\partial\lambda$ diverges for λ' values close to 0 and 1. The peaks in the derivative of the potential are due to repulsive van der Waals interactions in the initial stages of the transformation of the hydrogen in position 5 into an oxygen and in the final stages of the inverse transformation in position 3. This problem is avoided when the soft-core λ-dependent energy function is used (b). Similar effects are observed in the calculations in chloroform, although smaller in scale.

The rate of mutation in the slow-growth calculations is too rapid to ensure that the mutation is performed along a sequence of representative equilibrium substates. Thus, the obtained ΔG values must be taken only as indicative. It is worth noting, however, that the results for the mutation Ac-Trp to Trp-NMe when the soft-core λ-dependent energy function is used (2.3 kJ/mol in water and 3.3 kJ/mol in chloroform) are much closer to the results from the equivalent thermodynamic-integration calculations (set 1) (−0.1 kJ/mol in water and 2.5 kJ/mol in chloroform) than when the standard (linear) λ-dependent energy function is used (7.1 kJ/mol in water and 1.2 kJ/mol in chloroform).

(2) Closure of Theoretical Cycles: Effect of Sampling, Integration, and Constraint Contributions. A necessary condition for correctness in all free-energy calculations is that the change in free energy for any closed cycle is zero. Closure of simulated cycles is perhaps the best means to ensuring the self-consistency of the calculations and estimating the intrinsic uncertainty of the calculated free-energy differences, independent of force-field considerations. The change in free energy for the closed cycle Ac-Trp to Trp-NMe to Trp to Trp′ to Ac-Trp in water ($\Delta\Delta G_1$) and in chloroform ($\Delta\Delta G_2$) (see Figure 2) is given in Table 7 for sets 1 and 5. The hysteresis in $\Delta\Delta G_1$ (water), 0.6 kJ/mol for both sets 1 and 5, is smaller than the hysteresis in $\Delta\Delta G_2$ (chloroform), 2.6 and 1.7 kJ/mol for sets 1 and 5, respectively. The hysteresis is a measure of the minimum uncertainty in the calculated free-energy differences. Failure to close a simulated cycle can result from (i) insufficient sampling, (ii) errors in the integration over λ, or (iii) failure to include all theoretical contributions to the change in free energy.

(a) Sampling. The simulations were continued at each λ' value until the estimated error in $\langle\partial V/\partial\lambda\rangle_{\lambda'}$ was less than ±3.0 kJ/mol or a maximum simulation time of 200 ps was reached (see Table 6). It is important to note that, because the correlation time of the derivative of the potential energy cannot be definitively determined, it is not possible to precisely estimate the true error in either $\langle\partial V/\partial\lambda\rangle_{\lambda'}$ or ΔG. The errors quoted should only be taken as estimates and used primarily as a tool to compare the uncertainty in these quantities between λ' points.

(b) Integration. Plots of $\langle\partial V/\partial\lambda\rangle$ versus λ for the different perturbations (see Figure 5) indicate that there is little scope for improving the estimates of ΔG by the inclusion of more points in the integration. Adding additional points in regions of high curvature, $\lambda = 0.15$ and $\lambda = 0.85$, resulted in an overall change in ΔG (set 5) which was in all cases less than 0.8 kJ/mol with respect to the values shown in Table 4, without significantly affecting the $\Delta\Delta G$ values presented in Table 7.

(c) Constraint Contributions. If the length of a bond is perturbed in a given mutation, there will be a contribution from the bond to the total change in free energy.[21,32,50,51] This is true irrespective of whether the bonds are treated harmonically or held fixed by constraints. Nevertheless, there has been continued debate in the literature over whether the term is significant or will always cancel within a thermodynamic cycle.[31] In the present case, the change in free energy due to the constraint forces is about −1.4 kJ/mol for the mutation Ac-Trp to Trp-NMe in either water or chloroform and for either set 1 or 5, and about −0.2 kJ/mol for the mutation Trp′ to Trp in vacuum. If these contributions were not considered, $\Delta\Delta G_1$ and $\Delta\Delta G_2$ would be about 1.8 and 3.8 kJ/mol, respectively, for set 1, and 1.8 and 2.9 kJ/mol, respectively, for set 5. $\Delta\Delta G_5$ would not be affected, due to cancellation of errors.

(3) Agreement with Experimental Data: $\Delta\Delta G_3$, $\Delta\Delta G_4$, and $\Delta\Delta G_5$. From Table 7 it can be seen that, for set 1, the difference between the change in free energy calculated in water and the change in free energy calculated in chloroform for the mutations Trp-NMe to Trp ($\Delta\Delta G_3$) and Ac-Trp to Trp′ ($\Delta\Delta G_4$) are opposite in sign to the expected values, derived from the experimental partition coefficients (see Table 1). The difference between the experimental and calculated free-energy differences is also substantially larger than the intrinsic uncertainty in the calculated values, indicated by $\Delta\Delta G_1$ and $\Delta\Delta G_2$. Contrary to what is expected, the calculations suggest that the removal of the nonbonded interactions of atoms 2−5 (see Figure 2) with the rest of the system is more favorable in water than in chloroform. Following the same trend, the mutation Ac-Trp to Trp-NMe ($\Delta\Delta G_5$), although having the correct sign, shows that the energy necessary for the inversion of the dipole in the mutation Ac-Trp to Trp-NMe is either underestimated in water or overestimated in chloroform (or both). Together, the results for $\Delta\Delta G_3$, $\Delta\Delta G_4$, and $\Delta\Delta G_5$ indicate that the force field does not correctly describe the relative differences in the nature of the solute−solvent interactions in the water and chloroform systems. To sort possible reasons for this disagreement between the model and the real system interaction properties, we have studied the influence of individual corrections to the van der Waals and electrostatic potential energy functions.

(4) Parametrization of the Water−Nonpolar Carbon van der Waals Repulsion Energy. Recently, a number of researchers have questioned the parametrization of the $C12^{1/2}_{O_w,O_w}$ term used for water−nonpolar carbon interactions in the GROMOS87 force field.[38,39,46−49] The value of 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ used for set 1 has been shown to be more appropriate than the original value of 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ in a number of simulations. This term has not, however, been directly parametrized. In set 2 the mutation Ac-Trp to Trp′ was performed using the original GROMOS87 value of 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$. Use of 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ will increase the energy of interaction between SPC water and the solute, making $\Delta\Delta G_4$ less negative, in better agreement with experiment. Comparison of the results from sets 1 and 2 for the mutation Ac-Trp to Trp′ in water (Table 4) shows that the change in the free energy is 6.5 kJ/mol more positive when 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ is used than when 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ is used. By using the value 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$, we have delimited the maximum effect on the calculations from changing the $C12^{1/2}_{O_w,O_w}$ parameter. Although use of this parameter significantly improves agreement with experiment, it has been clearly demonstrated in other studies that 421.0 (kcal

(50) Edberg, R.; Evans, D. J.; Morris, G. P. *J. Chem. Phys.* **1986**, *84*, 6933−6939.
(51) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4545.

mol^{-1} Å^{12})$^{1/2}$ overestimates the interaction energy between water and aliphatic or aromatic groups. Nevertheless, these results and others[49] suggest that 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$, although a good first-order correction, underestimates the interaction energy between water and aliphatic or aromatic groups. The $C12^{1/2}_{O_w,O_w}$ term used for interactions of water with nonpolar carbons was, therefore, independently reparametrized on the basis of the free energy of hydration of CH_4 (Figure 4), and the obtained value of 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$ was further tested against a series of hydrocarbons of small molecular weight (Table 5). Although the calculated free energies of hydration shown in Table 5 are not equally accurate for the whole series of alkanes studied, our results are as good as those obtained by Kaminski *et al.*[52] for the same series when the united-atom OPLS force field is used. Comparison of the results from sets 1 and 3 for the mutation Ac-Trp to Trp′ in water shows that the change in free energy is 1.7 kJ/mol more positive when the reparametrized value of 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$ is used than when 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$ is used. These results show a linear dependence of the change in free energy for the mutation Ac-Trp to Trp′ on the $C12^{1/2}_{O_w,O_w}$ value that is used.

(5) Reaction-Field Correction to the Electrostatic Potential Energy. A possible source of systematic error in the calculations in water and chloroform is the use of a cutoff distance for the electrostatic interactions. The exclusion of long-range dipole−dipole interactions will have different effects in these two systems, due to the difference in dielectric permittivity.

Comparison of the results from sets 1 and 4 for the mutation Ac-Trp to Trp′ in water (see Table 4) shows that the change in free energy is 4.3 kJ/mol more positive when the reaction-field correction is used. Not surprisingly, the use of the reaction-field correction does not significantly modify the result (0.3 kJ/mol) for the same mutation in chloroform (set 5). It should be noted that the direct contribution to the change in free energy due to the reaction field, obtained by integrating $\langle \partial V^{RF}/\partial\lambda \rangle_{\lambda'}$ over λ, V^{RF} being the reaction-field correction to the λ-dependent electrostatic potential energy (see (8)), is only 0.4 kJ/mol for the mutation in water (set 4). The indirect effects of the reaction field, however, result in an increase of 2.2 kJ/mol in the short-range electrostatic contribution, and an increase of 1.3 kJ/mol in the van der Waals contribution. This underlines the necessity of using the reaction field in the simulations. When the perturbation is performed in chloroform (set 5), the integration of $\langle \partial V^{RF}/\partial\lambda \rangle_{\lambda'}$ over λ gives a contribution to the total change in free energy of 0.1 kJ/mol.

In set 5 both corrections to the force field have been applied: the change of the $C12^{1/2}_{O_w,O_w}$ van der Waals parameter for water−nonpolar carbon interactions to 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$, and the use of the reaction-field correction to the electrostatic interactions. With respect to the results from set 1, ΔG_{H_2O}(Trp-NMe → Trp) is 3.7 kJ/mol more positive, ΔG_{H_2O}(Ac-Trp → Trp′) is 6.1 kJ/mol more positive, ΔG_{H_2O}(Ac-Trp → Trp-NMe) is 2.4 kJ/mol more positive, ΔG_{CHCl_3}(Trp-NMe → Trp) is 0.6 kJ/mol more negative, ΔG_{CHCl_3}(Ac-Trp → Trp′) is 0.3 kJ/mol more positive, and ΔG_{CHCl_3}(Ac-Trp → Trp-NMe) is unchanged. Comparing the results from sets 1 and 3−5 in Table 4, it can be seen that for the mutation Ac-Trp to Trp′ in water the two corrections have additive effects. It is interesting to note that the effect of the reaction field is not equivalent for the mutations Trp-NMe to Trp and Ac-Trp to Trp′ in water, affecting the latter significantly more. In set 1 the species Ac-Trp and Trp-NMe were energetically indistinguishable in water. The mutations Trp-NMe to Trp and Ac-Trp to Trp (Ac-Trp to Trp′ + Trp′ to Trp) gave almost the same free-energy change, and the mutation Ac-Trp to Trp-NMe gave a zero free-energy change. In contrast, in set 5, where the reaction-field correction is applied, the calculated free energies are sensitive to the small difference in the overall dipole of the molecules. The difference between the free-energy changes calculated for the mutations Trp-NMe to Trp and Ac-Trp to Trp (1.7 kJ/mol) is very close to the free-energy change calculated for the mutation Ac-Trp to Trp-NMe (2.3 kJ/mol). This is a good example of how the closure of theoretical cycles ($\Delta\Delta G_1$ in sets 1 and 5) is independent of force-field considerations. In chloroform, the reaction field has almost no effect on the free-energy changes, and the differences between the results from sets 1 and 5 are small.

Table 8. Free-Energy Differences from Charge Scaling of Atoms C, O, N, and H (2−5 in Figure 2) of Ac-Trp in Water and in Chloroform[a]

scaling factor	$\Delta G(H_2O)^b - \Delta G(CHCl_3)^c$ (kJ/mol)	scaling factor	$\Delta G(H_2O)^b - \Delta G(CHCl_3)^c$ (kJ/mol)
1.000	0.00	1.100	−3.84
1.025	−0.89	1.125	−4.91
1.050	−1.83	1.150	−6.02
1.075	−2.81		

[a] The initial charges (scaling factor 1) are (e) 0.38 for C, −0.38 for O, −0.28 for N, and 0.28 for H. [b] Free energy of charge scaling in water. [c] Free energy of charge scaling in chloroform.

Improved treatment of long-range electrostatic interactions by inclusion of the reaction field and reparametrization of the water−nonpolar carbon interaction to reproduce hydration free energies of alkanes together significantly improve the agreement between the calculated and experimental free energies. $\Delta\Delta G_3$ and $\Delta\Delta G_4$ have the correct sign, and there is close agreement between the expected and calculated values of $\Delta\Delta G_5$ (Table 7).

(6) Estimation of Charge Correction. Despite the improvement in the agreement between the calculated and experimental free-energy differences $\Delta\Delta G_3$ and $\Delta\Delta G_4$ on going from set 1 to set 5, the calculated values are still lower than the experimental values by 5.8 and 4.5 kJ/mol, respectively. This remaining discrepancy could result either from the parametrization of the solvent models or from the choice of partial charges on the Trp derivatives. To determine the change in partial charge on the atoms, C, O, N, and H (2−5 in Figure 2) of Ac-Trp required to reproduce the experimental free-energy difference between the mutation Ac-Trp to Trp′ in water and in chloroform ($\Delta\Delta G_4$), the following procedure was used.

The charges were scaled linearly, and the change in free energy was estimated by application of the perturbation formula (3). One hundred fifty picosecond trajectories of Ac-Trp in water and in chloroform were used (with previous 50 ps equilibration). Reaction-field effects were included (R_c = 1.4 nm), and configurations were stored every 0.02 ps. The change in free energy due to charge scaling was determined for different charge sets by rescanning the trajectories using the approach described by Liu *et al.*[53] The results show that scaling the partial charges on the atoms C, O, N, and H by 1.125 (Table 8) would be sufficient to reproduce the experimental free-energy difference. This is equivalent to increasing the charges (e) on the C and O atoms from ±0.38 to ±0.43, and on the N and H atoms from ±0.28 to ±0.31. Two points should be noted with respect

(52) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077−13082.
(53) Liu, H.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.*, in press.
(54) Dunfield, L. G.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1978**, *82*, 2609−2616.
(55) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563−593.
(56) Beutler, T. C.; Béguelin, D. R.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 3787−3793.

to this result. First, the partial charges that will best reproduce the experimental free-energy differences will depend on the choice of solvent model, van der Waals parameters, and treatment of long-range interactions. Second, the change in partial charge is small compared to differences commonly found for equivalent functional groups between comparable empirically derived force fields. Thus, although these results suggest that a slight increase of these partial charges in the GROMOS87 force field might be appropriate, further more direct evidence would be required before such a change is recommended.

Conclusions

Typically, the reliability of free-energy calculations is judged by comparison of the results with experimental data. Although agreement with experiment is clearly a necessary condition, it is not a sufficient condition to validate the molecular model and the force field used. A minimum set of requirements include that the ensemble average at each λ' point must be properly converged, and that the number and distribution of λ' points must be such that adding extra points will not significantly change the results. In addition, the change in free energy for any closed cycle (including nonphysical cycles) should be zero. In these respects we have demonstrated that the calculated ΔG values listed in Tables 4−7 truly reflect the changes in free energy corresponding to the underlying model. Despite the mutations involving the deletion of up to six atoms, these ΔG values were determined with a precision of k_BT or better. This has in turn allowed us to refine our model, on the basis of comparison to the experimental results. We have also shown that the computation of the constraint contribution to the change in free energy is necessary if the constraints are dependent on λ and the environment in which they are applied is different for different perturbations in a cycle, and thus should always be considered.

In MD simulations the precision in the numerical integration of the equations of motion, for a given time-step length, decreases as the second derivative of the potential energy function with respect to the coordinates, the curvature of the potential energy function, increases. In free-energy calculations, when atoms are created or annihilated, small interparticle distances become energetically accessible, and can be sampled as λ approaches 0 (creation) or 1 (annihilation). If a standard 6−12 Lennard-Jones and Coulombic nonbonded interaction function is used, the first and second derivatives of the potential energy function with respect to the coordinates will diverge for these small interparticle distances. To overcome this problem, we have used the soft-core λ-dependent potential of Beutler *et al.*,[24] and clearly shown that this form of potential energy function improves the efficiency of free-energy calculations.

Ultimately, the ability to reproduce the experimental partition coefficients will depend on the parametrization of the force field that is used. The GROMOS force field is a so-called empirical or effective force field. It has been parametrized to reproduce thermodynamic properties of simple liquids and to reproduce structural and dynamic properties of (primarily) peptides in aqueous solution or in crystals. The parameters in effective force fields are highly correlated. Underestimation of one parameter may be compensated by altering another. Although this may result in an appropriate parametrization for a specific environment, such compensation will result in systematic errors when the same processes in different environments are compared.

For the mutations investigated there are two main sources of systematic errors: (i) the choice of van der Waals and electrostatic solute−solvent interaction parameters and (ii) the use of a cutoff radius for the nonbonded interactions. The sensitivity of the free-energy calculations to different values of the $C12^{1/2}_{O_w,O_w}$ van der Waals parameter used for interactions between water and nonpolar atoms has been evaluated. We have shown that, due to compensation of errors, the original value of 421.0 (kcal mol^{-1} Å^{12})$^{1/2}$ results in better agreement with the experimental partition coefficients than the recently proposed value of 793.3 (kcal mol^{-1} Å^{12})$^{1/2}$. However, neither value correctly reproduces the hydration free energy of simple alkanes. The $C12^{1/2}_{O_w,O_w}$ van der Waals parameter used for interactions between water and nonpolar atoms was reparametrized, and it was shown that a value of approximately 690.0 (kcal mol^{-1} Å^{12})$^{1/2}$ is more appropriate. This underlines the importance of directly parametrizing cross interactions where possible rather than relying solely on combination rules.

The importance of the adequate treatment of long-range electrostatic interactions when relative free-energy differences are compared for a process that takes place in solvents with different dielectric properties was also evaluated. The use of a reaction-field correction has been shown to significantly affect (up to $2k_BT$) the results, improving the qualitative agreement with experiment, despite the fact that the molecular systems under study have no net charge, and that the cutoff radius that is used for the electrostatic interactions is large (1.4 nm). Larger reaction-field effects are expected for mutations involving changes in net charge or when using a smaller cutoff for electrostatic interactions. The predominant contribution to the change in free energy does not come from the reaction-field term itself. Instead, the effect of the reaction field is indirect, altering the conformations sampled during the trajectory. Thus, it is not possible to properly correct the change in free energy for reaction-field effects unless the simulation is performed with the reaction field included.

Neither the correction of the $C12^{1/2}_{O_w,O_w}$ parameter nor the inclusion of the reaction-field correction was sufficient to obtain quantitative agreement with the experimentally derived $\Delta\Delta G$ values. Agreement with experiment could be achieved by increasing the partial charges on the atoms that were mutated. Slight scaling of the charges (by a factor of 1.125) was sufficient. Compared to similar effective force fields, the partial charges in the GROMOS force field are small. Of course, larger charges would discriminate better between a high and a low dielectric environment, and it would be possible to parametrize a model which reproduces the experimental partition coefficients without the need to include long-range interactions by overemphasizing short-range interactions. If transferable force fields are to be developed, it is important that this approach is not taken. Furthermore, partial charges cannot be considered independently of van der Waals interactions, and cross interactions between different atom types cannot always be considered by blind application of combination rules. To this end, we have demonstrated that the calculation of relative partition coefficients can be effectively used to identify and rectify inadequacies and systematic errors in a force field, which might otherwise go undetected.

Acknowledgment. The authors wish to thank T. C. Beutler, I. G. Tironi, and H. Liu for helpful comments and discussion on the methods used in this work. X.D. is a recipient of a fellowship from the Programa de Química Fina from the CIRIT (Generalitat de Catalunya). F.X.A. and E.Q. acknowledge financial support from Grants BIO92-0458, BIO95-0848, and IN94-0347 from the CICYT (Ministerio de Educación y Ciencia, Spain), and from the Centre de Referència de R+D en Biotecnologia (Generalitat de Catalunya).

JA9537944