# Description of hydration free energy density as a function of molecular physical properties

Kyoung Tai No[a,b,*], Seong Gon Kim[a], Kwang-Hwi Cho[b], Harold A. Scheraga[b,*]

[a]*Department of Chemistry and CAMD Research Center, Soong Sil University, Seoul 156-743, Korea*
[b]*Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853-1301, USA*

## Abstract

  A method to calculate the solvation free energy density (SFED) at any point in the cavity surface or solvent volume surrounding a solute is proposed. In the special case in which the solvent is water, the SFED is referred to as the hydration free energy density (HFED). The HFED is described as a function of some physical properties of the molecules. These properties are represented by simple basis functions. The hydration free energy of a solute was obtained by integrating the HFED over the solvent volume surrounding the solute, using a grid model. Of 34 basis functions that were introduced to describe the HFED, only six contribute significantly to the HFED. These functions are representations of the surface area and volume of the solute, of the polarization and dispersion of the solute, and of two types of electrostatic interactions between the solute and its environment. The HFED is described as a linear combination of these basis functions, evaluated by summing the interaction energy between each atom of the solute with a grid point in the solvent, where each grid point is a representation of a finite volume of the solvent. The linear combination coefficients were determined by minimizing the error between the calculated and experimental hydration free energies of 81 neutral organic molecules that have a variety of functional groups. The calculated hydration free energies agree well with the experimental results. The hydration free energy of any other solute molecule can then be calculated by summing the product of the linear combination coefficients and the basis functions for the solute.   © 1999 Elsevier Science B.V. All rights reserved.

*Keywords:*  Solvation free energy; Solvation free energy density; Hydration free energy; Hydration free energy density; Empirical solvation model; Implicit solvent model

---

[*]Corresponding authors.

# 1. Introduction

The development of a practical dielectric continuum model to calculate solvation free energies is crucial for an understanding of the physical and chemical properties of solutions [1–4], especially solutions of macromolecules. In continuum models [5–8], the solvent is treated as an ensemble-averaged continuum medium, i.e. the solvent is treated implicitly.

The original concept of a continuum model for a charged solute was proposed by Born [5] and Bell [6], and extended to the case of a dipolar solute by Onsager [7], and given a quantum mechanical description by Kirkwood [8]. In the dielectric medium, the charge or dipole of the solute induces an electric field, the reaction field (RF), and this RF in turn perturbs the solute, leading to a modification of the charge distribution of the solute. All proposed continuum solvent models, in both the classical and quantum mechanical approaches, were derived by making use of the concept of the reaction field.

In the classical continuum models, only first-order perturbation effects are taken into account. The solute has a fixed charge distribution and the continuum medium can be polarized (the reaction field can be modified) or, alternatively, the solute can be polarized and the reaction field is not modified.

In the classical approach, most works are concerned with describing the electrostatic contribution to the solvation in terms of the Poisson equation (PE) [9] or the Poisson–Boltzmann equation (PBE) [10]. Some workers have dealt with the Born model [11–16], where the solute is treated as a point charge whereas, recently, most works deal with a distribution of charge throughout the solute. Either the Boundary Element Method (BEM) [13–15,17–26] or the Finite Difference Method (FDM) [27–42] can provide numerical solutions for random charge distributions. In the BEM approach, the reaction potential is described by an apparent charge distribution, the apparent surface charge (ASC) spread over the boundary surface between the dielectric media, i.e. on the cavity surface of the solute. One of the popular volume integration methods

is the FDM. In the FDM, the dielectric property of the solvent is described in terms of a three-dimensional grid around the solute. In both methods, the results of the calculations depend critically on the choice of the position of the dielectric boundary and on the set of atomic charges. Since there are several parameters to be determined empirically, many workers have been concerned with finding the optimum parameters for continuum solvation models [11–42]. The parameters are determined so that the calculated hydration free energies reproduce the experimental values.

A simplification of the BEM and FDM methods is the Solvent Accessible Surface Area (SASA) and the Hydration Shell models [43–71]. The free energy of hydration can be described simply by multiplying the SASA or hydration shell volume of each atom in a solute by the average free energy density of each atom and summing up all the contributions. These simple expressions are computationally efficient and can reproduce the experimental data well. However, they cannot describe the polarization of the solvent in the cavity formed by the solute.

In quantum mechanical continuum models, an iterative procedure is used: (i) the wave function of the solute embedded in the solvent is evaluated under the influence of the fixed solvent reaction field, and then (ii) the solvent reaction field is modified with the calculated solute wave function, and this procedure is repeated variationally and self-consistently. The free energy of the charge distribution in a homogeneous dielectric medium was given by Kirkwood [8]. The formulation contains the Born [5] solvation energy term (for an ion located at the origin) and the Onsager [7] dipole moment contribution term (for a dipole located at the origin). The expression for the reaction field is the same as with the classical model mentioned above. Even in the quantum mechanical model, the parameters to define the shape of the solute must be determined empirically by fitting to experimental data.

The quantum mechanical continuum models [72–95], such as the self-consistent-reaction-field (SCRF) model, were developed on the basis of the Onsager reaction field model [7]. In this

model, the solute cavity is specified as a set of overlapping spheres, and both the wave function and the reaction field are repeatedly modified in each SCF cycle. In the Polarization Continuum Model (PCM), proposed by Tomasi and co-workers [77−82], the effective surface charges on the cavity, which mimic the effects of the solvent, are calculated by using the solute wave function and the dielectric constant of the solvent; then, the interaction energy between the surface charges and the charges of the solute is computed. Recently, a new infinite-order solvation scheme, the Isoelectrondensity Polarizable Continuum Model (IPCM), which defines the cavity of the solute in a solvent with an isoelectrondensity surface of the solute, was proposed by Wong et al. [83−87]. Quantum mechanical continuum dielectric models at the semi-empirical level have also been proposed [96−103].

For calculating the solvation free energy, the free energy required to form a cavity of the shape of the solute must also be evaluated [59−61], in addition to the solute−solvent electrostatic and non-electrostatic interactions,

The purpose of this work is to propose a method to calculate a solvation free energy density (SFED) which will be described as a function of molecular physical properties with some analytical functions. The parameters resulting from the calculation of the SFED of some organic molecules will be applied to calculate the hydration free energy density (HFED) and the hydration free energy of some organic molecules when the solvent is water.

## 2. Empirical model for solvation free energy density (SFED) and calculation of solvation free energy

### 2.1. Definitions of relevant quantities

The process of solvation of a solute can be broken up into two steps: (i) creating a cavity of the size of the solute in the solvent; and (ii) the interaction of the solute with its environment, the solvent. The free energy of solvation can be written in terms of these two contributions as

$$\Delta G_{\text{sol}} = \Delta G_{\text{cav}} + \Delta G_{\text{inter}} \tag{1}$$

where $\Delta G_{\text{sol}}$ is the difference in free energy between the solvated solute and the unsolvated (gas-phase) solute. The free energy of interaction, $\Delta G_{\text{inter}}$, is the difference between the free energy of the solution, $G_{\text{sys}}$, and the sum of the free energies of the pure solvent (e.g. water) and the solute in the gas phase.

$$\Delta G_{\text{inter}} = G_{\text{sys}} - (G_w + G_{\text{solute}}) \tag{2}$$

At temperature $T$, this equation can be written as follows:

$$\begin{aligned} \Delta G_{\text{inter}} = {} & -RT \, \ln\langle e^{-H_{\text{sys}}/RT} \rangle \\ & - [-RT \, \ln\langle e^{-H_w/RT} \rangle \\ & - RT \, \ln\langle e^{-H_{\text{solute}}/RT} \rangle] \end{aligned} \tag{3}$$

where $H_{\text{sys}}$, $H_w$ and $H_{\text{solute}}$ represent the Hamiltonian of the solution, the pure solvent, and the gas phase solute, respectively, and the symbols $\langle \rangle$ represent ensemble averages. It is possible that some part of the $\Delta G_{\text{inter}}$ may be included in $\Delta G_{\text{cav}}$; since the decomposition expression, Eq. (1), is not defined explicitly, the degree of double counting may depend on the definition of the state of the cavity and the reference state of $\Delta G_{\text{sys}}$.

If the enthalpy of the reference state of each phase is set equal to zero, i.e. the reference state of the force field is set equal to zero, and the averaging is taken over a finite number of configurations, Eq. (3) becomes

$$\begin{aligned} \Delta G_{\text{inter}} \approx {} & -RT \Bigg[ \sum_I w_I(-H_{I,\text{sys}}/RT) \\ & - \sum_J w_J(-H_{J,W}/RT) \\ & - \sum_K w_K(-H_{K,\text{solute}}/RT) \Bigg] \end{aligned} \tag{4}$$

$$= \sum_I w_I H_{I,\text{sys}} - \sum_J w_J H_{J,w} - \sum_K w_K H_{K,\text{solute}} \tag{5}$$

where $w_I$, $w_J$ and $w_K$ are the weight factors of the $I$th, $J$th, and $K$th configurations of the solution, pure solvent, and pure solute, respectively. As a result of this approximation, some of the perturbations of the solute−solvent energy may be small, and are neglected.

## 2.2. Components of the Hamiltonian

The Hamiltonian of the solution, $H_{I,\text{sys}}$, can be described as a sum of components.

$$H_{I,\text{sys}} = H_{I,\text{sys}}^{\text{solute}} + H_{I,\text{sys}}^{w} + H_{I,\text{sys}}^{s-w} \tag{6}$$

where $H_{I,\text{sys}}^{\text{solute}}$ and $H_{I,\text{sys}}^{w}$ represent the Hamiltonian of the solute and solvent, respectively, in the solution. $H_{I,\text{sys}}^{s-w}$ represents the solute−solvent interaction energy.

In Fig. 1, the solute and the solvent of the solution are described as assemblages of interacting compartments. The subscripts $i$ and $j$ are used to represent the interacting compartments of the solute (e.g. the atoms of the solute), and $k$ and $l$ represent those of the solvent. The model is composed of two domains, solute ($q$) and solvent ($Q$). In this continuum model, the components of the Hamiltonian $H_{I,\text{sys}}$ are described as follows.

$$
\begin{aligned}
H_{I,\text{sys}}^{\text{solute}} = \frac{1}{2} \int_q \int_q \Bigg[ & V_1^o(\mathbf{a};\mathbf{r}_i,\mathbf{r}_j) \\
& + \int_Q V_1'(\mathbf{a};\mathbf{r}_{ki},\mathbf{r}_{kj}) d\mathbf{r}_k \Bigg]_I
\end{aligned}
$$
$$ d\mathbf{r}_i d\mathbf{r}_j = E_I^{o,\text{solute}} + E_I^{\prime,\text{solute}} \tag{7}$$

where $\mathbf{a}$ represents the physical properties of the solute, and $V_1^o$ and $V_1'$ are the potential energy of the solute without perturbation by the solvent and the change of the interaction potential between interacting compartments $i$ and $j$ of the solute due to the perturbation by the solvent, respectively, i.e. $V_1'$ represents the perturbation of the interaction between the two compartments $i$ and $j$ of the solute by the solvent.
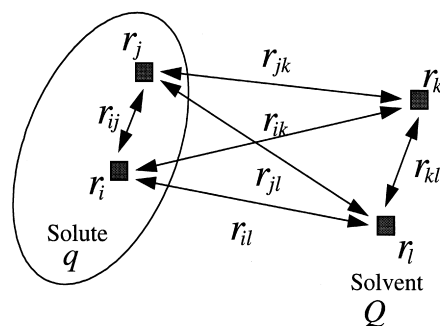


Fig. 1. The solute and solvent of a solution are described as an assemblage of interaction compartments.

$$
\begin{aligned}
H_{I,\text{sys}}^{w} = \frac{1}{2} \int_Q \int_Q \Bigg[ & V_2^o(\mathbf{r}_k,\mathbf{r}_l) \\
& + \int_q V_2'(\mathbf{a};\mathbf{r}_{ik},\mathbf{r}_{il}) d\mathbf{r}_i \Bigg]_I d\mathbf{r}_k d\mathbf{r}_l
\end{aligned}
\tag{8}
$$

where $V_2^o$ and $V_2'$ are the potential energy of the pure solvent without perturbation by the solute and the changes of the interaction potential between interacting compartments $k$ and $l$ of the solvent molecules due to the perturbation by the solute, respectively, i.e. $V_2'$ represents the perturbation of the interaction between the two compartments $k$ and $l$ of the solvent by the solute. Finally,

$$H_{I,\text{sys}}^{s-w} = \int_q \int_Q [V_3(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) d\mathbf{r}_i d\mathbf{r}_k]_I \tag{9}$$

where $V_3$ is the interaction potential between the solute and the solvent.

The Hamiltonians of the pure solvent and the gas phase solute are

$$H_{J,w} = \frac{1}{2} \int_Q \int_Q [V_2^o(\mathbf{r}_k,\mathbf{r}_l)]_J d\mathbf{r}_k d\mathbf{r}_l \tag{10}$$

and

$$H_{K,\text{solute}} = \frac{1}{2} \int_q \int_q [V_1^o(\mathbf{a};\mathbf{r}_i,\mathbf{r}_j)]_K d\mathbf{r}_i d\mathbf{r}_j = E_K^{o,\text{solute}} \tag{11}$$

respectively. By introducing Eqs. (6)–(11) into Eq. (5), $\Delta G_{\text{inter}}$ becomes

$$
\begin{aligned}
\Delta G_{\text{inter}} = \\
\Bigg[ \sum_I w_I E_I^{o,\text{solute}} - \sum_K w_K E_K^{o,\text{solute}} \Bigg] \\
+ \frac{1}{2} \sum_I w_I \int_{q'} \int_q \left\{ \int_Q V_1'(\mathbf{a};\mathbf{r}_{ki},\mathbf{r}_{kj}) d\mathbf{r}_k \right\} d\mathbf{r}_i d\mathbf{r}_j \\
+ \frac{1}{2} \sum_I w_I \int_Q \int_Q \left\{ \int_q V_2'(\mathbf{a};\mathbf{r}_{ik},\mathbf{r}_{il}) d\mathbf{r}_i \right\} d\mathbf{r}_k d\mathbf{r}_l \\
+ \sum_I w_I \int_q \int_Q V_3(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) d\mathbf{r}_i \mathbf{r}_k
\end{aligned}
\tag{12}
$$

### 2.3. Mean field approximation

By introducing the *mean field approximation* into the second term of Eq. (12), the change of the interaction potential between the interacting compartments $i$ and $j$ of the solute by the solvent is simplified as the additional interaction between the solute and solvent compartments, $i$ and $k$.

$$
\begin{aligned}
\int_q \int_q \left\{ \int_Q V_1'(\mathbf{a};\mathbf{r}_{ki},\mathbf{r}_{kj}) d\mathbf{r}_k \right\} d\mathbf{r}_i d\mathbf{r}_j \\
= \int_q \int_Q \left\{ \int_q V_1'(\mathbf{a};\mathbf{r}_{ki},\mathbf{r}_{kj}) d\mathbf{r}_j \right\} d\mathbf{r}_i d\mathbf{r}_k \\
= \int_Q \int_q V_1^*(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) d\mathbf{r}_i d\mathbf{r}_k
\end{aligned}
\tag{13}
$$

where $V_1^*$ is the result of the integration of the term in braces on the right-hand side of Eq. (13). The quantity $V_1^*$ plays the same role as the screening in the orbital approximation of the Hartree–Fock method.

By introducing the mean field approximation into the third term of Eq. (12), the change of the interaction potential between the interacting compartments $k$ and $l$ of the solvent by the solute is also simplified as the additional interaction between the solute and solvent compartments, $i$ and $k$; i.e. by use of the mean field

approximation, we avoid the necessity of evaluating integrals of the type $\int_Q \int_Q d\mathbf{r}_k d\mathbf{r}_l$.

$$
\begin{aligned}
\int_Q \int_Q \left\{ \int_q V_2'(\mathbf{a};\mathbf{r}_{ik},\mathbf{r}_{il}) d\mathbf{r}_i \right\} d\mathbf{r}_k d\mathbf{r}_l \\
= \int_Q \int_q \left\{ \int_Q V_2'(\mathbf{a};\mathbf{r}_{ik},\mathbf{r}_{il}) d\mathbf{r}_l \right\} d\mathbf{r}_i d\mathbf{r}_k \\
= \int_Q \int_q V_2^*(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) d\mathbf{r}_i d\mathbf{r}_k
\end{aligned}
\tag{14}
$$

A rigid solute has only one stable conformation; hence, $I = 1$ and $K = 1$. Therefore, the first term of Eq. (12) can be ignored, and hence only rigid or not very flexible molecules will be used as constraints for determining the empirical parameters of this model. Thus, $\Delta G_{\text{inter}}$ becomes,

$$
\begin{aligned}
\Delta G_{\text{inter}} = \int_Q \int_q [V_1^*(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) \\
+ V_2^*(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k) + V_3(\mathbf{a};\mathbf{r}_i,\mathbf{r}_k)] d\mathbf{r}_i, d\mathbf{r}_k
\end{aligned}
\tag{15}
$$

$\Delta G_{\text{inter}}$ corresponds to the interaction energy between the solute and water up to the first order perturbation.

If the solute molecule is described as a finite number of interacting fragments, denoted by $i$, and the solvent is assumed to be a continuum, Eq. (15) can be written in terms of two alternative formulas by replacing $\int_q$ by $\Sigma_i$.

$$
\begin{aligned}
\Delta G_{\text{inter}} = \int_Q \sum_i [V_1^*(\mathbf{a};\mathbf{r}_{ik}) \\
+ V_2^*(\mathbf{a};\mathbf{r}_{ik}) + V_3(\mathbf{a};\mathbf{r}_{ik})] d\mathbf{r}_k
\end{aligned}
\tag{16}
$$

or

$$
\begin{aligned}
\Delta G_{\text{inter}} = \sum_i \int_Q [V_1^*(\mathbf{a};\mathbf{r}_{ik}) \\
+ V_2^*(\mathbf{a};\mathbf{r}_{ik}) + V_3(\mathbf{a};\mathbf{r}_{ik})] d\mathbf{r}_k
\end{aligned}
\tag{17}
$$

Whether one uses Eq. (16) or (17), the interaction energy up to first order perturbation is

$$f(\mathbf{r}_{ik}) \equiv V_1^*(\mathbf{a};\mathbf{r}_{ik}) + V_2^*(\mathbf{a};\mathbf{r}_{ik}) + V_3(\mathbf{a};\mathbf{r}_{ik}) \quad (18)$$

In this work, we will use Eq. (16) because it is easy to visualize $\sum_i f(\mathbf{r}_{ik})$ in the Q space, the solvent domain around the solute. Hence,

$$\Delta G_{\text{inter}} = \int_Q \sum_i f(\mathbf{r}_{ik}) d\mathbf{r}_k \quad (19)$$

$\sum_i f(\mathbf{r}_{ik})$ represents the contribution of $\Delta G_{\text{inter}}$ to the free energy density of solvation at the point $\mathbf{r}_k$. Finally, from Eq. (1), the free energy of solvation becomes

$$\Delta G_{\text{sol}} = \Delta G_{\text{cav}} + \int_Q \sum_i f(\mathbf{r}_{ik}) d\mathbf{r}_k \quad (20)$$

The above relations are represented schematically in Fig. 2.

### 2.4. Evaluation of free energy

To evaluate Eq. (20), we must decided how to calculate $\Delta G_{\text{cav}}$, how to evaluate the integral $\int_Q$, how to carry out the summation, and how to determine the form of the function $f(\mathbf{r}_{ik})$.

For the evaluation of $\Delta G_{\text{cav}}$, several methods have been proposed [43–71]. Among them, the description in terms of the surface tension, $\gamma$, is
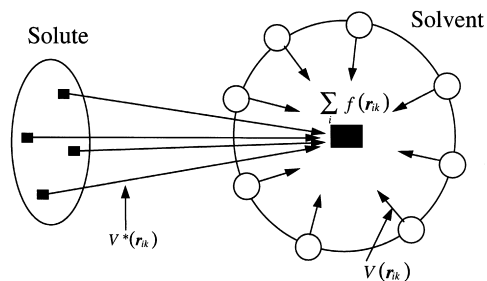


Fig. 2. The interaction between the solute and the water which are perturbed by the solute. The large circle designates the boundary of the grid points for the calculation of $V(\mathbf{r}_{ik})$. The small open circles indicate the grid points within the boundary. The sum of the arrows from the solute to the dark rectangle within the boundary of the solvent grid points corresponds to $\sum_i V^*(\mathbf{r}_{ik})$. The sum of the arrows from the small open circles to the dark rectangle corresponds to $\sum_i V(\mathbf{r}_{ik})$. The sum of *all* the arrows in the figure corresponds to $\sum_i f(\mathbf{r}_{ik})$, where $f(\mathbf{r}_{ik})$ is given by Eq. (18).

the most widely used method in empirical models. Eq. (20) can be modified for use with a solvent accessible surface area (SASA) model as

$$\Delta G_{\text{sol}} = \sum_i \gamma A_i + \sum_i \int_\sigma f(\mathbf{r}_{i\sigma}) d\sigma \quad (21)$$

where $A_i$ is the surface area of the $i$th atom, the volume integral $\int_Q$ is transformed into the surface integral $\int_\sigma$, and $f(\mathbf{r}_{i\sigma}) d\sigma$ represents the effective interaction of the $i$th interacting compartment of the solute with the surface element $d\sigma$ of the cavity surface. If the average value of $f(\mathbf{r}_{i\sigma})$ for each atomic species (or for each interacting segment) of the solute is taken as $\bar{f}_{\sigma i}$, then $\int_\sigma f(\mathbf{r}_{i\sigma}) d\sigma$ in the second term of Eq. (21) becomes $\bar{f}_{\sigma i} A_i$, and Eq. (21) becomes

$$\Delta G_{\text{sol}} = \sum_i \left(\gamma + \bar{f}_{\sigma i}\right) A_i = \sum_i g_i A_i \quad (22)$$

Eq. (20) can also be modified for use with a Hydration Shell model. If the integration $\int_Q$ of Eq. (20) is carried out from the solvent accessible surface (SAS) to the surface at a distance $\Delta r_c$ from the SAS, then Eq. (21) becomes,

$$\Delta G_{\text{sol}} = \sum_i \gamma A_i + \sum_i \int_{V_{io}}^{V_{io}+\Delta V_C} f(\mathbf{r}_{i,V}) d\mathbf{V}_i \quad (23)$$

or

$$\Delta G_{\text{sol}} = \sum_i \frac{\gamma}{\Delta r_C} A_i \Delta r_C + \sum_i \bar{f}_{V,i} \Delta V_i$$
$$= \sum_i \left\{\frac{\gamma}{\Delta r_i} + \bar{f}_{V,i}\right\} \Delta V_i = \sum_i \bar{g}_i \Delta V_i \quad (24)$$

where $V_{io}$ is the effective volume of the $i$th atomic species which is surrounded by the SAS, and $\Delta V_C$ is the effective solvent accessible volume (i.e. the volume of the layer between the SAS and the surface at the distance $\Delta r_c$ from the SAS). The first term of Eq. (23) is modified to the first term of Eq. (24) in order to describe $\Delta G_{\text{cav}}$ with the effective solvent accessible volume. The term $f(\mathbf{r}_{i,V}) d\mathbf{V}_i$ represents the interaction of the $i$th interacting point of the solute with the volume

element $d\mathbf{V}_i$ in the shell. The quantity $\bar{f}_{V,i}$ is the average value of $\int f(\mathbf{r}_{i,V})d\mathbf{V}_i$ for the $i$th atomic species. When either model is used, the calculation of the surface area or the volume of the hydration shell is the crucial step, and the empirical parameters that define the surface or the shell must be determined with experimental free energy of solvation data through a least-squares fitting procedure.

The concept of *solvation free energy density* is a very useful tool with which to investigate the water binding affinity distribution of the solute surface. If is evenly distributed over the volume or surface defined by $\int_Q$ or $\int_\sigma$, then $\Delta G_{\mathrm{sol}}$ of Eq. (20) can be written as

$$\Delta G_{\mathrm{sol}} = \int_Q \left[ \sum_i^{N_A} f(\mathbf{r}_{ik}) + g_{\mathrm{cav}} \right] d\mathbf{r}_k = \int_Q g(\mathbf{r}_k) d\mathbf{r}_k \tag{25}$$

where $N_A$ is the number of atoms in the solute, $g_{\mathrm{cav}}$ is $\Delta G_{\mathrm{cav}}$/*unit volume* or $\Delta G_{\mathrm{cav}}$/*unit area* depending on whether a volume or a surface integration, respectively, is used. The quantity $g(\mathbf{r}_k)$ is the SFED at point $\mathbf{r}_k$.

## 3. Computations

In this work, the SFED calculation method was applied to obtain the HFED and hydration free energy (writing $\Delta G_{\mathrm{hyd}}$ for $\Delta G_{\mathrm{sol}}$) of organic molecules. As an approximation, $\int_Q$ of Eq. (25) was replaced by a summation over a grid around the solute, $\sum_Q$. The solute was described as an assembly of atoms, and each atom was treated as a compartment of the solute. Then, Eq. (25) was approximated as

$$\Delta G_{\mathrm{hyd}} = a_l \sum_k^Q \sum_i^{N_A} f(\mathbf{r}_{ik}) + \Delta G_{\mathrm{cav}} \tag{26}$$

where $a_l$ is a proportionality constant which depends on the size of the grid interval, $\Delta_l$. The grid points were located between the SAS and some distance from the SAS, namely, $R_{\mathrm{shell}}$. The space between the SAS and the outer surface corre-
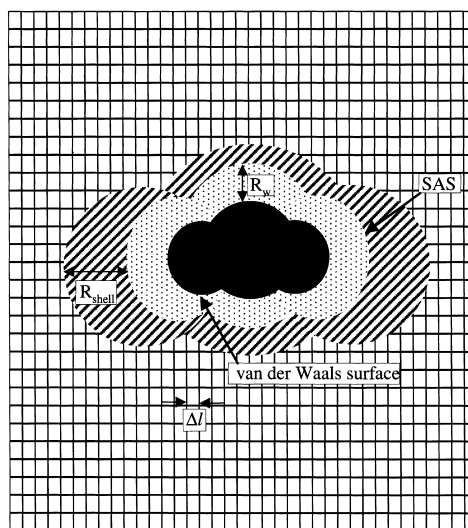


Fig. 3. Solvent Accessible Surface Area (SASA) of the solute, $R_W$, $R_{\mathrm{shell}}$, $\Delta l$, and the van der Waals surface of the solute.

sponds to the 'hydration shell'. The SAS of the solute is defined by the overlap of the atomic SASs. The atomic SAS is the spherical surface whose radius is the sum of its van der Waals radius, $R_{\mathrm{vdw}}$, and the effective solvent (water) shell thickness, $R_W$ (Fig. 3). Since the computation time depends linearly on the number of grid points, the choice of $\Delta l$ and $R_{\mathrm{shell}}$ is very important for economical computations. The optimum values of $\Delta l$, $R_W$, and $R_{\mathrm{shell}}$ were taken as a compromise by considering both the computing time and the accuracy of the calculation (see Section 4).

The quantity $f(\mathbf{r}_{ik})$ is described with a combination of $m$ basis functions, $h(\mathbf{r}_{ik})$.

$$f(\mathbf{r}_{ik}) = \sum_j^m C_j h_j(\mathbf{r}_{ik}) \tag{27}$$

Then, by inserting Eq. (27) into Eq. (26), $\Delta G_{\mathrm{hyd}}$ becomes

$$\Delta G_{\mathrm{hyd}} = a_l \sum_k^Q \sum_i^{N_A} \sum_j^m C_j h_j(\mathbf{r}_{ik}) + \Delta G_{\mathrm{cav}}$$

$$= a_l \sum_k^Q \sum_j^m \sum_i^{N_A} C_j h_j(\mathbf{r}_{ik}) + \Delta G_{\mathrm{cav}}$$

$$= a_l \sum_k^Q \sum_j^m C_j h_j^o(\mathbf{r}_k) + \Delta G_{\text{cav}}$$

$$= \sum_k^Q \sum_j^m C_j(\Delta l) h_j^o(\mathbf{r}_k) + \Delta G_{\text{cav}} \qquad (28)$$

where $h_j^o(\mathbf{r}_k)$ is the contribution of the $j$th basis functions of the atoms of the solute, and $C_j$ is a function of $\Delta l$.

To obtain the $h_j^o(\mathbf{r}_k)$ terms of Eq. (28), it is necessary to know their functional forms. As can be seen from Eq. (15), $\Delta G_{\text{inter}}$ has contributions from several terms, namely, the interaction between the solute and the solvent, and the perturbation of both the solute and the solvent by their environments. Since the solvent is described with a continuum model, one can not introduce well-defined analytical energy functions for those interactions. Although the functional form is not clear, it is believed that the dominant terms in molecular interactions in the gas phase also play an important role in solution. Hence, in the description of $h_j^o(\mathbf{r}_k)$, and therefore of $f(\mathbf{r}_{ik})$, gas-phase type intermolecular interaction functions were introduced. The appropriate functional form was determined through an iterative trial and error procedure in which the difference between the predicted and experimental hydration free energies was minimized (see below).

For the expression for $\Delta G_{\text{cav}}$, three models were used, namely, (i) the solute volume ($V_s$); (ii) the solvent accessible surface area ($S_s$); or (iii) a linear combination of $V_s$ and $S_s$, the latter being

$$\Delta G_{\text{cav}} = C_V V_s + C_s S_s + \text{const} \qquad (29)$$

where the *const* term is a residual contribution, even for a small cavity for which $V_s$ and $S_s$ approach zero. Hence, Eq. (28) becomes

$$\Delta G_{\text{hyd}} = \sum_k^Q \sum_j^m C_j(\Delta l) h_j^o(\mathbf{r}) + C_V V_s + C_s S_s + \text{const} \qquad (30)$$

For the expression for $\sum_j^m C_j(\Delta l) h_j^o(\mathbf{r}_k)$, a linear combination [given by the coefficients $C_j(\Delta_l)$ of

Eq. (30)] of the basis functions which are a subset of the following trial function set $\mathbf{X}$ was used:

$$\mathbf{X} = \left\{ \sum_i \frac{q_i}{r_{ik}^{n1}}, \left| \sum_i \frac{q_i}{r_{ik}^{n1}} \right|, \sum_i \frac{q_i^2}{r_{ik}^{n1}}, \sum_i \frac{q_i|q_i|}{r_{ik}^{n1}}, \sum_i \frac{\alpha_i}{r_{ik}^{n2}}, \right.$$

$$\left. \sum_i \frac{\alpha_i^2}{r_{ik}^{n2}}, \sum_i \frac{C_i}{r_{ik}^{n2}}, \sum_i \frac{q_i \alpha_i}{r_{ik}^{n2}}, \sum_i \frac{\alpha_i C_i}{r_{ik}^{n2}} \right\} \qquad (31)$$

where $q_i$, $\alpha_i$ and $C_i$ are the net atomic charge, the effective atomic polarizability, and the dispersion energy coefficient of the $i$th atom in the solute, respectively. $n1$ is an integer from 1 to 3, and $n2$ is an integer from 3 to 6. Considering the commonly used intermolecular interaction functional forms, we initially selected 32 basis functions in $\mathbf{X}$. After optimization, this number was reduced (see Section 4).

The coefficients, $C_j(\Delta l)$, $C_V$, $C_S$, and the const term in Eq. (28) and Eq. (29) were determined by minimizing the following function:

$$F = \sum_N W_N \left| \Delta G_{\text{hyd},N}^{\text{calc.}} - \Delta G_{\text{hyd},N}^{\text{exp}t.} \right| \qquad (32)$$

$W_N$ is the weight factor of the $N$th data value, and depends on the reliability of the data and the frequency of its appearance in a system of biological molecules. If the molecule is rigid, then a large value of $W_N$ was assigned because the geometry of the molecule does not change much upon solvation. For flexible molecules smaller weight factors were assigned.

For $q_i$, atom-centered net atomic charges were used. The charges were calculated with an empirical net atomic charge calculation method, the Modified Partial Equalization of Orbital Electronegativity (MPEOE) method [104,105]. The atomic polarizabilities, $\alpha_i$, were calculated from an empirical formula in which the effective atomic polarizability in a molecule is described as a function of the net atomic charge, the Charge Dependent Effective Atomic Polarizability (CDEAP) [106]. The dispersion coefficients, $C_i$, were calculated from the $\alpha_i$'s using the Slater−Kirkwood formula [107−109]. The van der Waals radii of the atoms of the solute molecules were taken from

our previous work on the force fields for proteins [107]. In this work, the van der Waals radii, summarized in Table 1, were not optimized.

The surface area is approximately linear in the number of grid points $n_Q$ in the solvation shell. The molecular volume is approximately linear in the sum of the distances $\Sigma_k^Q r_k$ between the center of mass and the grid points. The quantity $\Delta G_{cav}$ was taken as a linear combination of the contributions from both the surface area and the molecular volume. Eq. (29) then becomes

$$
\begin{aligned}
\Delta G_{cav} &= C_s A(\Delta l, R_{shell}) n_Q \\
&\quad + C_V B(\Delta l, R_{shell}) \sum_k^Q r_k + \text{const} \\
&= C_s'(\Delta l, R_{shell}) n_Q \\
&\quad + C_V'(\Delta l, R_{shell}) \sum_k^Q r_k + \text{const}
\end{aligned}
\tag{33}
$$

where $A(\Delta l, R_{shell})$ and $B(\Delta l, R_{shell})$ are proportionality constants and are functions of both $\Delta l$ and $R_{shell}$.

Since $f(\mathbf{r}_{ik})$ of Eq. (27) was described as a function of the physical properties of molecules in Eq. (28) and Eq. (31), the HFED is a function of both position and the physical properties. Hence, Eq. (25) becomes

$$
\Delta G_{hyd} = \sum_k^Q g(\mathbf{r}_k, \mathbf{X}, \Delta l)
\tag{34}
$$

To save computation time, the grid interval was determined as large as possible under the condition that the value of $F$ of Eq. (32) does not change significantly with changes in the grid interval. In choosing the molecules that were used as constraints for the determination of the coefficients and the functional form of $f(\mathbf{r}_{ik})$, two criteria were used: (i) the molecules should appear frequently as fragments of biomolecules; and (ii) the molecules should be rigid. If the molecule is flexible, its conformation may change upon hydration. Since the number of rigid molecules is not sufficiently large to establish reliable constraints, molecules with one or two dihedral angles were also introduced as constraints. For those flexible molecules, only the lowest-energy conformations were used, and smaller weight factors were assigned than those for rigid molecules. For rigid molecules, the weight factor was 1, and for

Table 1
The van der Waals radii (in Å) used in the HFED calculation

| Atom type | Description | van der Waals Radii[a] |
|---|---|---|
| H1 | Aliphatic H | 1.474 |
| H2 | H bonded to aromatic system | 1.579 |
| H3 | Hydroxyl H in alcohol | 1.175 |
| H4 | Hydroxyl H in carboxylic acid | 1.220 |
| H5 | H bonded to amide N | 1.165 |
| H6 | H bonded to amine N | 1.130 |
| C1 | Aliphatic C | 2.074 |
| C2 | Aromatic C | 1.994 |
| C3 | C in carboxylic acid | 1.725 |
| C4 | C in amide | 1.725 |
| O1 | O in carboxylic acid | 1.764 |
| O2 | O in carboxylic acid COH | 1.490 |
| O3 | O in alcohol | 1.489 |
| N1 | Aromatic N with three bonds | 1.589 |
| N2 | Aromatic N with two bonds | 1.349 |
| N3 | N in amine | 1.435 |
| N4 | N in amide | 1.689 |

[a] Calculated using Table 5 and Eq. (10c) of No et al. [107], except for H5, H6 and N4 whose values were determined during the optimization procedure.

flexible molecules the weight factor was taken between 0.5 and 0.1 depending on the flexibility of the molecule.

For the determination of the coefficients in Eq. (30), 81 neutral molecules that contain H, C, N, O atoms and encompass a variety of organic functional groups were introduced as constraints. The geometry of the molecule was obtained by energy minimization in the gas phase, using CVFF[1] for the energy calculations. All the geometry optimization was carried out with Cerius2 [110].

In summary, all the basis functions of Eqs. (31) and (33) were used in Eq. (28), with the coefficients $C_j$, $C_s'$, $C_V'$ and the const term of Eq. (33) optimized by minimizing $F$ of Eq. (32). Then, those basis functions that contributed negligibly to the calculated free energy were removed, and optimization was continued with the remaining basis functions that contribute significantly to $\Delta G_{\text{hyd}}$. By evaluating the sum $\Sigma_i^{N_A} h_j(\mathbf{r}_{ik}), h_j^o(\mathbf{r}_k)$ was calculated at each grid point $\mathbf{r}_k$ (the sum of the distances from each atom $i$ to the grid point $k$) in the solvent. Then, with the aid of Eqs. (28),(33), $\Delta G_{\text{hyd}}$ was calculated for the solute; this calculation involved the calculation of the sum of the interaction energies between each atom $i$ of the solute and each grid point $\mathbf{r}_k$ of the solvent.

## 4. Results and discussion

### 4.1. Parameterization without a hydrogen bond effect: non-polar molecules

Of 34 basis functions, two for $\Delta G_{\text{cav}}$ and 32 for $f(\mathbf{r}_{ik})$, only six basis functions were found to contribute significantly to $\Delta G_{\text{hyd}}$. $\{S_s, V_s\}$ are the functions for $\Delta G_{\text{cav}}$ and

$$\left\{ \left| \sum_i \frac{q_i}{r_{ik}} \right|, \sum_i \frac{q_i^2}{r_{ik}}, \sum_i \frac{\alpha_i}{r_{ik}^3}, \sum_i \frac{C_i}{r_{ik}^6} \right\}$$

are those for $f(\mathbf{r}_{ik})$. The second group of basis functions will be abbreviated as $\{f_{\text{inter}}\}$.

The optimum values for $\Delta l$, $R_{\text{shell}}$, and $R_W$
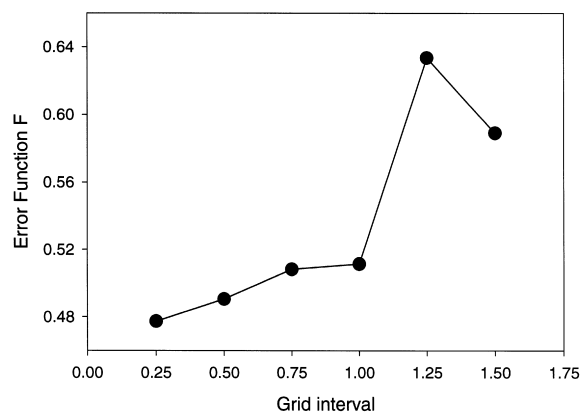
----

[1]CVFF: the force field in Cerius2 was used.



Fig. 4. The error $F$ (defined in Eq. (32)) in the calculated free energy of hydration (in kcal/mol) is plotted against the grid interval, $\Delta l$, in Å. In this illustration, the free energy was calculated with the HFED model.

were determined as 0.5, 3.0, and 1.4 Å, respectively. The grid interval $\Delta l$ was determined as large as possible without introducing much change in the error function F (Fig. 4). When $R_{\text{shell}}$ is larger than 3.0 Å, the calculated value of $\Delta G_{\text{hyd}}$ does not change appreciably compared to its value when $R_{\text{shell}}$ is 3 Å. To save computation time, 3.0 Å was used for $R_{\text{shell}}$. A value of 1.4 Å was used for the effective radius of water. This value was obtained from the experimental oxygen−oxygen pair correlation function of water [103].

In this work, three models were introduced, depending on the expression for the cavity energy, $\Delta G_{\text{cav}}$. In model I, $S_s$, const, and $\{f_{\text{inter}}\}$, in model II, $V_s$, const, and $\{f_{\text{inter}}\}$, and in model III, both $S_s$ and $V_s$, const, and $\{f_{\text{inter}}\}$ were used; only six parameters had to be determined for each of models I and II, and only seven for model III. In Table 2 the optimized coefficients of models I−III are summarized. The deviations described by Eq. (32) were obtained as 0.4906, 0.4489, and 0.4302 kcal/mol, for models I, II and III, respectively.

From the values of the deviations, the shell volume is a better descriptor than the surface area in describing the cavity free energy. When both shell volume and surface area are used, the cavity free energy can be described more accurately compared with the case in which only one descriptor is used. In model III, the coefficient of the surface area has a negative sign and the

Table 2
The linear coefficients for the basis functions of $\Delta G_{\text{hyd.}}$ of models I−III

| | $\left\lvert \sum_i \dfrac{q_i}{r_{ik}} \right\rvert$ $(\times 10^{-1})$ | $\sum_i \dfrac{q_i^2}{r_{ik}}$ $(\times 10^{-2})$ | $\sum_i \dfrac{\alpha_i}{r_{ik}^3}$ $(\times 10^{-4})$ | $\sum_i \dfrac{C_i}{r_{ik}^6}$ $(\times 10^{-3})$ | $S_s$ $(\times 10^{-4})$ | $V_s$ $(\times 10^{-4})$ | Const. |
|---|---|---|---|---|---|---|---|
| Model I | −0.3466 | −0.6359 | −0.4348 | −0.3042 | 0.8407 | | 1.21 |
| Model II | −0.3475 | −0.7686 | −0.2252 | −0.5299 | | 0.1869 | 1.23 |
| Model III | −0.3476 | −0.7686 | −0.2252 | −0.5299 | −0.4820 | 0.2886 | 1.23 |

The coefficients have units that enable the product of the basis function and the coefficient to be expressed in kcal/mol.

coefficient of the volume is larger than that in model II (see Table 2). The coefficients of $\Delta G_{\text{inter}}$ of models II and III are almost identical whereas the coefficients of $\sum_i \dfrac{\alpha_i}{r_{ik}^3}$ and $\sum_i \dfrac{C_i}{r_{ik}^6}$ of model I are very different from those of models II and III.

For alkanes, alkenes, alkylated benzenes, and alkylated naphthalenes, the calculated values of $\Delta G_{\text{hyd}}$ agreed reasonably well with the experimental data, especially when the molecules are rigid or have only one or two torsional degrees of freedom (see Table 3). Since the conformations of flexible molecules in aqueous solution are not known, some of the flexible molecules show large deviations between $\Delta G_{\text{hyd}}^{\text{calc}}$ and $\Delta G_{\text{hyd}}^{\text{exp}}$. It seems that, for non-polar molecules, the volume descriptor works better than the surface area in describing the cavity free energy.

### 4.2. Parameterization with a hydrogen bond effect: polar molecules

For polar molecules, especially those that can form a hydrogen bond with water, the estimated values of $\Delta G_{\text{hyd.}}^{\text{calc.}}$ are smaller than those for $\Delta G_{\text{hyd.}}^{\text{exp.}}$ except for amides (see Table 3). Since there is no extra basis function for describing the hydrogen bond in the expression for $\Delta G_{\text{hyd}}$, the contribution from the stabilization due to the hydrogen bond between solute and water (described partially by electrostatic interactions) is less well estimated. The contribution of the electrostatic interaction between the solute and water to the hydrogen bond energy is dominant compared with the contribution from the other interactions. For amides and alkylated acetates, the point charges on the atoms that participate in hydrogen bond

formation with water are large enough to take the hydrogen bond contribution into account. However, the point charges on the atoms in alcohols, carboxylic acids, and amines that form hydrogen bonds with water are too small to express the hydrogen bond effect quantitatively.

The hydrogen bond is a phenomenon that is traditionally studied outside of continuum solvation models because of its high specificity and directionality. The hydration free energies of the alcohol, carboxylic acid, and amine molecules that form a hydrogen bond with water are less well estimated compared with experimental data. Those molecules have X−H functional groups that form hydrogen bonds. It seems that the strength of the interaction between X−H and the continuum solvent is not large enough to describe the hydrogen bond between X−H and water. To overcome this shortcoming, the magnitude of the X−H dipole was increased until the error between the calculated and experimental values of $\Delta G_{\text{hyd}}$ of alcohols, carboxylic acids, and amines reached a minimum. During this optimization, the charges of all atoms *not* included in hydrogen bonding were fixed, while *all* the coefficients of $\Delta G_{\text{hyd}}$ and the charges of the atoms that participate in the hydrogen bonding with water were optimized. Since model I gives better results for the group of molecules that are fragments of biomolecules, model I was modified. For amines, the net atomic charge $\delta_{\text{H}}$ of H in N−H was increased by $\Delta \delta_{\text{H}}$ (13.8% of $\delta_{\text{H}}$), and for the net atomic charge $\delta_{\text{N}}$ of N, $\Delta \delta_{\text{H}}$ (which is equal to $-\Delta \delta_{\text{N}}$) was subtracted from $\delta_{\text{N}}$. For carboxylic acids, $\delta_{\text{H}}$ was increased by 12.7%, and for alcohols, $\delta_{\text{H}}$ was increased by 5.4%. The point charges on the

hydrogen bond acceptors were not optimized. This point charge modified model will be designated as an HFED-Effective Charge Model (HFED-ECM). The optimum values of $\Delta l$ and $R_{\text{shell}}$ were obtained as 0.75 and 3.0 Å, respectively, whereas 0.5 and 3.0 Å were obtained for the HFED model. In an analog of Fig. 4 (not shown here), the error $F$ described by Eq. (32) was plotted against $\Delta l$ for the HFED-ECM. In the calculation, $R_{\text{shell}}$, and $R_W$ were fixed at 3.0 and 1.4 Å, respectively. When $\Delta l$ is 0.75 Å, $F$ is smaller than those at $\Delta l$ equal to 0.25 and 0.5 Å. As $\Delta l$ increases above 0.75 Å, the value of $F$ increases rapidly. It seems that 0.75 Å is appropriate for $\Delta l$ both for accuracy and for saving computation time in the free energy calculation in the HFED-ECM.

The values of $\Delta G_{\text{hyd}}$ calculated with the HFED-ECM for the same set of 81 compounds are listed in Table 4. The values of $\Delta G_{\text{hyd}}^{\text{calc.}}$ for the molecules that form a hydrogen bond with water agreed well with the values of $\Delta G_{\text{hyd}}^{\text{exp.}}$ except for methanol and ethanol (see Table 4).

In Table 5, the components of $\Delta G_{\text{hyd}}$ calculated with the HFED-ECM are summarized for some molecules. Some terms from Eq. (31) were introduced to describe the contribution from electrostatic interactions, i.e. $\sum_i \frac{q_i}{r^n}, \left| \sum_i \frac{q_i}{r^n} \right|$, $\sum_i \frac{q_i^2}{r^n}$, and $\sum_i \frac{q_i|q_i|}{r^n}$, where $n$ takes values from 1 to 3. Among those 12 terms, only two terms, $\left| \sum_i \frac{q_i}{r} \right|$ and $\sum_i \frac{q_i^2}{r}$, contribute considerably to $\Delta G_{\text{hyd}}$ and the contributions from the other terms are negligible. Since $\left| \sum_i \frac{q_i}{r^n} \right|$ is the magnitude of the electric potential at the grid point k, this term is responsible for the electrostatic stabilization of the dipole of the solvent by the electric potential produced by the solute. $\sum_i \frac{q_i^2}{r^n}$ contributes more than $\left| \sum_i \frac{q_i}{r^n} \right|$ to $\Delta G_{\text{hyd}}$ (see Table 4). $\sum_i \frac{q_i^2}{r^n}$ may be responsible for the interaction between the point charge on the solute and its image charge induced in the water, especially in the hydration shell.

For describing the contributions from polarization and non-bonded interactions, the $\sum_i \frac{\alpha_i}{r_{ik}^n}$, $\sum_i \frac{\alpha_i^2}{r_{ik}^n}$, and $\sum_i \frac{C_i}{r_{ik}^n}$ terms were introduced as basis functions of $\Delta G_{\text{hyd}}$. Among them, only two functions, $\sum_i \frac{\alpha_i}{r_{ik}^3}$ and $\sum_i \frac{C_i}{r_{ik}^6}$, contribute considerably, especially the second function. These two functions may be responsible for the polarization of the solute by water and the dispersion interaction between the solute and water. The surface area and the optimized term 'const' of Eq. (33), 1.15, (for the HFED-ECM) contribute to the formation of the shape of the solute cavity. The cross-terms, i.e. $\sum_i \frac{q_i \alpha_i}{r_{ik}^n}$ and $\sum_i \frac{\alpha_i C_i}{r_{ik}^n}$, do not contribute appreciably to $\Delta G_{\text{hyd}}$. Thus, in the HFED-ECM, $\Delta G_{\text{hyd}}$ was obtained from the following expression:

$$\Delta G_{\text{hyd}} = \sum_k^Q \left[ -0.112878 \left| \sum_i \frac{q_i}{r_{ik}} \right| \right.$$
$$-0.021990 \sum_i \frac{q_i^2}{r_{ik}}$$
$$-0.301314 \times 10^{-3} \sum_i \frac{\alpha_i}{r_{ik}^3}$$
$$\left. -0.176349 \times 10^{-2} \sum_i \frac{C_i}{r_{ik}^6} \right]$$
$$+0.370383 \times 10^{-3} S_S + 1.15 \qquad (35)$$

These coefficients, optimized for the HFED-ECM, differ from those of Table 2, which were optimized for the HFED model. All the basis functions that contribute considerably to $\Delta G_{\text{hyd}}$ have physical meanings such as a coulomb potential, a van der Waals potential, etc.

*4.3. Final comment*

The advantage of the SFED method of calculation is that it is expressed in terms of analytical functions with only six or seven adjustable parameters. The SFED can be transformed easily

Table 3
Calculated and experimental free energies of solvation (kcal/mol)

| Molcules | $\Delta G_{hyd.}^{calc.}$, HFED model | | | $\Delta G_{hyd.}^{exp.a}$ |
|---|---|---|---|---|
| | Model I | Model II | Model III | |
| Methane | 1.79 | 1.89 | 1.87 | 2.00 |
| Ethane | 1.78 | 1.93 | 1.94 | 1.83 |
| Propane | 1.81 | 2.03 | 2.06 | 1.96 |
| Butane | 1.89 | 2.16 | 2.21 | 2.08 |
| Pentane | 1.95 | 2.30 | 2.41 | 2.33 |
| 2,2-Dimethylbutane | 1.79 | 2.14 | 2.26 | 2.59 |
| 2-Methylpropane | 1.90 | 2.16 | 2.22 | 2.32 |
| 2,2-Dimethylpropane | 1.84 | 2.15 | 2.23 | 2.50 |
| 3-Methylpentane | 1.81 | 2.18 | 2.30 | 2.51 |
| Ethene | 1.25 | 1.24 | 1.27 | 1.27 |
| 1-Propene | 1.28 | 1.34 | 1.39 | 1.27 |
| 1-Butene | 1.27 | 1.47 | 1.59 | 1.38 |
| 1-Pentene | 1.25 | 1.45 | 1.58 | 1.66 |
| Formaldehyde | −2.28 | −2.55 | −2.60 | −2.77 |
| 3 − Methyl-2-butanone | −3.44 | −3.54 | −3.48 | −3.26 |
| Cyclopentanone | −3.57 | −3.74 | −3.70 | −4.72 |
| Benzene | −0.77 | −0.98 | −0.97 | −0.87 |
| Toluene | −0.93 | −1.07 | −1.00 | −0.89 |
| 2-Ethyltoluene | −1.08 | −1.09 | −0.98 | −1.04 |
| 4-Ethyltoluene | −0.98 | −0.96 | −0.83 | −0.96 |
| 1,2,3-Trimethylbenzene | −1.21 | −1.22 | −1.16 | −1.22 |
| 1,2,4-Trimethylbenzene | −1.15 | −1.15 | −1.04 | −0.86 |
| 1,3,5-Trimethylbenzene | −1.09 | −1.06 | −0.98 | −0.90 |
| Ethylbenzene | −0.90 | −0.97 | −0.88 | −0.79 |
| Propylbenzene | −1.02 | −1.01 | −0.88 | −0.53 |
| Isopropylbenzene | −0.88 | −0.89 | −0.78 | −0.30 |
| o-Xylene | −1.10 | −1.18 | −1.09 | −0.90 |
| m-Xylene | −1.04 | −1.10 | −1.03 | −0.84 |
| p-Xylene | −1.00 | −1.05 | −1.01 | −0.81 |
| Naphthalene | −2.14 | −2.41 | −2.38 | −2.39 |
| 1-Methylnaphthalene | −2.31 | −2.52 | −2.46 | −2.45 |
| 1,3-Dimethylnaphthalene | −2.45 | −2.58 | −2.48 | −2.48 |
| 1,4-Dimethylnaphthalene | −2.40 | −2.53 | −2.46 | −2.83 |
| 2,3-Dimethylnaphthalene | −2.51 | −2.63 | −2.52 | −2.79 |
| 2,6-Dimethylnaphthalene | −2.47 | −2.58 | −2.44 | −2.64 |
| 1-Ethylnaphthalene | −2.35 | −2.48 | −2.37 | −2.41 |
| Methanol | −3.29 | −3.89 | −3.76 | −5.11 |
| Ethanol | −3.38 | −3.96 | −3.83 | −5.01 |
| 1-Propanol | −3.71 | −4.28 | −4.12 | −4.83 |
| 2-Propanol | −3.63 | −4.19 | −4.04 | −4.76 |
| 2-Methyl − 1-propanol | −3.86 | −4.42 | −4.25 | −4.52 |
| 2-Methyl-2-propanol | −3.74 | −4.28 | −4.09 | −4.52 |
| 1-Butanol | −3.90 | −4.43 | −4.24 | −4.74 |
| 2-Butanol | −3.71 | −4.25 | −4.07 | −4.58 |
| 3-Methyl-1-butanol | −4.01 | −4.53 | −4.33 | −4.44 |
| 2-Methyl-2-butanol | −3.89 | −4.41 | −4.19 | −4.45 |
| 1-Pentanol | −4.05 | −4.54 | −4.30 | −4.47 |
| 2-Pentanol | −4.06 | −4.57 | −4.34 | −4.39 |
| 3-Pentanol | −4.05 | −4.55 | −4.34 | −4.35 |

Table 3 (*Continued*)

| Molcules | $\Delta G_{\text{hyd.}}^{\text{calc.}}$, HFED model | | | $\Delta G_{\text{hyd.}}^{\text{exp.a}}$ |
|---|---|---|---|---|
| | Model I | Model II | Model III | |
| Acetic acid | −4.76 | −5.58 | −5.48 | −6.70 |
| Propanoic acid | −4.99 | −5.80 | −5.68 | −6.47 |
| Butanoic acid | −5.25 | −6.04 | −5.86 | −6.38 |
| Acetone | −3.36 | −3.55 | −3.54 | −3.80 |
| 2-Butanone | −3.42 | −3.57 | −3.53 | −3.64 |
| 2-Pentanone | −3.54 | −3.63 | −3.55 | −3.53 |
| 3-Pentanone | −3.44 | −3.55 | −3.49 | −3.41 |
| Acetophenone | −4.92 | −5.38 | −5.32 | −4.58 |
| Propanal | −2.93 | −3.12 | −3.12 | −3.44 |
| Butanal | −3.05 | −3.19 | −3.14 | −3.19 |
| Acetamide | −9.75 | −11.16 | −10.89 | −9.75 |
| Propionamide | −10.08 | −11.49 | −11.27 | −9.42 |
| $t-N$-Methylacetamide | −10.01 | −11.28 | −11.05 | −10.07 |
| Dimethylacetamide | −8.51 | −9.48 | −9.27 | −8.54 |
| Methylamine | −2.84 | −3.29 | −3.22 | −4.56 |
| Ethylamine | −2.98 | −3.41 | −3.30 | −4.50 |
| Propylamine | −3.06 | −3.45 | −3.33 | −4.39 |
| Butylamine | −3.26 | −3.60 | −3.42 | −4.26 |
| Pyridine | −5.01 | −5.76 | −5.62 | −4.70 |
| 3,5-Dimethyl pyridine | −5.75 | −6.40 | −6.23 | −4.84 |
| 3-Methyl pyridine | −5.38 | −6.08 | −5.89 | −4.77 |
| 4-Methyl pyridine | −5.46 | −6.16 | −5.98 | −4.96 |
| Dimethyl ether | −1.81 | −2.02 | −1.94 | −1.90 |
| Diethyl ether | −1.69 | −1.78 | −1.65 | −1.64 |
| Dipropyl ether | −1.99 | −1.94 | −1.72 | −1.15 |
| Methylisopropyl ether | −1.75 | −1.86 | −1.73 | −2.01 |
| Methylpropyl ether | −1.91 | −2.02 | −1.86 | −1.66 |
| Diisopropyl ether | −1.61 | −1.61 | −1.39 | −0.53 |
| Ethyl acetate | −3.21 | −3.59 | −3.47 | −3.10 |
| Propyl acetate | −3.37 | −3.70 | −3.55 | −2.86 |
| Isopropyl acetate | −3.32 | −3.65 | −3.50 | −2.65 |
| Methyl propionate | −3.29 | −3.69 | −3.55 | −2.94 |
| Av.[b] | 0.4906 | 0.4489 | 0.4302 | |

[a] Experimental data taken from Cabani et al. [111], Pearson [112] and Abraham et al. [113].
[b] Deviation defined by Eq. (32).

into other mathematical forms for some useful physical quantities, for example, the molecular descriptors for quantitative structure−activity relations (QSAR) and quantitative structure−property relations (QSPR). Since $\Delta G_{\text{hyd}}$ is expressed as a function of position, the HFED can be plotted around a molecule. In Fig. 5a,b, $\Delta G_{\text{inter}}$ of acetamide and acetic acid are plotted in two dimensions. Since the hydration free energy density is described in terms of the physical properties of molecules, especially the electrostatic fields produced by molecules, we believe that the free

energy density is describable with a general functional from, and this expression is transferable to other similar molecules.

## 5. Conclusion

A model has been proposed to calculate the SFED. In the model, the SFED was expressed as a function of the physical properties of the solute molecule with the parameters $(q_i, \alpha_i, C_i)$ pertaining to the atoms $i$ and $j$ of the solute. This model was used for the calculation of the free energy

Table 4
Free energy of hydration (in kcal/mol) calculated with the HFED-ECM

| Molecules | $\Delta G_{\text{hyd.}}^{\text{calc.}}$ | $\Delta G_{\text{hyd.}}^{\text{exp.}}$ | Molecules | $\Delta G_{\text{hyd.}}^{\text{calc.}}$ | $\Delta G_{\text{hyd.}}^{\text{exp.}}$ |
|---|---|---|---|---|---|
| Methane | 1.94 | 2.00 | 2-Methyl-2-propanol | −4.32 | −4.52 |
| Ethane | 1.94 | 1.83 | 1-Butanol | −4.45 | −4.74 |
| Propane | 1.98 | 1.96 | 2-Butanol | −4.30 | −4.58 |
| Butane | 2.06 | 2.08 | 3-Methyl-1-butanol | −4.60 | −4.44 |
| Pentane | 2.12 | 2.33 | 2-Methyl-2-butanol | −4.49 | −4.45 |
| 2,2-Dimethylbutane | 1.96 | 2.59 | 1 − Pentanol | −4.65 | −4.47 |
| 2-Methylpropane | 2.06 | 2.32 | 2-Pentanol | −4.61 | −4.39 |
| 2,2-Dimethylpropane | 2.00 | 2.50 | 3-Pentanol | −4.63 | −4.35 |
| 3-Methylpentane | 1.98 | 2.51 | Acetic acid | −6.71 | −6.70 |
| Ethene | 1.34 | 1.27 | Propanoic acid | −7.09 | −6.47 |
| 1-Propene | 1.37 | 1.27 | Butanoic acid | −7.39 | −6.38 |
| 1-Butene | 1.40 | 1.38 | Acetone | −3.13 | −3.80 |
| 1-Pentene | 1.38 | 1.66 | 2-Butanone | −3.23 | −3.64 |
| Formaldehyde | −2.12 | −2.77 | 2-Pentanone | −3.33 | −3.53 |
| 3-Methyl-2-butanone | −3.24 | −3.26 | 3-Pentanone | −3.24 | −3.41 |
| Cyclopentanone | −3.40 | −4.72 | Acetophenone | −4.96 | −4.58 |
| Benzene | −0.77 | −0.87 | Propanal | −2.75 | −3.44 |
| Toluene | −0.86 | −0.89 | Butanal | −2.85 | −3.19 |
| 2-Ethyltoluene | −1.04 | −1.04 | Acetamide | −9.77 | −9.75 |
| 4-Ethyltoluene | −0.89 | −0.96 | Propionamide | −10.04 | −9.42 |
| 1,2,3-Trimethylbenzene | −1.18 | −1.22 | t-N-methylacetamide | −10.03 | −10.07 |
| 1,2,4-Trimethylbenzene | −1.06 | −0.86 | Dimethylacetamide | −8.41 | −8.54 |
| 1,3,5-Trimethylbenzene | −1.09 | −0.90 | Methylamine | −4.23 | −4.56 |
| Ethylbenzene | −0.86 | −0.79 | Ethylamine | −4.43 | −4.50 |
| Propylbenzene | −0.97 | −0.53 | Propylamine | −4.58 | −4.39 |
| Isopropylbenzene | −0.81 | −0.30 | Butylamine | −4.87 | −4.26 |
| o-Xylene | −1.06 | −0.90 | Pyridine | −5.04 | −4.70 |
| m-Xylene | −1.01 | −0.84 | 3,5-Dimethyl pyridine | −5.76 | −4.84 |
| p-Xylene | −1.03 | −0.81 | 3-Methyl pyridine | −5.37 | −4.77 |
| Naphthalene | −2.23 | −2.39 | 4-Methyl pyridine | −5.44 | −4.96 |
| 1-Methylnaphthalene | −2.36 | −2.45 | Dimethyl ether | −1.65 | −1.90 |
| 1,3-Dimethylnaphthalene | −2.48 | −2.48 | Diethyl ether | −1.53 | −1.64 |
| 1,4-Dimethylnaphthalene | −2.47 | −2.83 | Dipropyl ether | −1.82 | −1.15 |
| 2,3-Dimethylnaphthalene | −2.54 | −2.79 | Methylisopropyl ether | −1.55 | −2.01 |
| 2,6-Dimethylnaphthalene | −2.47 | −2.64 | Methylpropyl ether | −1.75 | −1.66 |
| 1-Ethylnaphthalene | −2.41 | −2.41 | Diisoproply ether | −1.45 | −0.53 |
| Methanol | −3.66 | −5.11 | Ethyl acetate | −3.10 | −3.10 |
| Ethanol | −3.86 | −5.01 | Propyl acetate | −3.25 | −2.86 |
| 1-Propanol | −4.24 | −4.83 | Isopropyl acetate | −3.21 | −2.65 |
| 2-Propanol | −4.18 | −4.76 | Methyl propionate | −3.18 | −2.94 |
| 2 − Methyl-1-propanol | −4.44 | −4.52 | | | |
| Av[a] | | 0.430 | | | |

[a] Deviation defined by Eq. (32).

density and free energy of hydration of organic molecules. Thirty-four physical properties, described by simple basis functions, were introduced, and the HFED was expressed as a linear combination of these basis functions. The coefficients of the linear combination were de-termined in order to minimize the deviation between the calculated and observed free ener-gies of hydration. Among the 34 basis functions, only six functions contribute considerably to the HFED. All six basis functions are physically real-istic for the description of hydration. The hydra-

Table 5

Contribution of each term (in kcal/mol) to $\Delta G_{\text{hyd}}$ [The values of $\Delta G_{\text{hyd.}}$ were calculated with the HFED-ECM[a] (in kcal/mol)]

| | $\left\lvert\sum_i \dfrac{q_i}{r_{ik}}\right\rvert$ | $\sum_i \dfrac{q_i^2}{r_{ik}}$ | $\sum_i \dfrac{\alpha_i}{r_{ik}^3}$ | $S_s$ | $\sum_i \dfrac{C_i}{r_{ik}^6}$ | $\Delta G_{\text{hyd.}}^{\text{calc.}}$ | $\Delta G_{\text{hyd.}}^{\text{exp.}}$ |
|---|---|---|---|---|---|---|---|
| Ethane | −0.121 | −0.015 | −0.022 | 1.18 | −0.236 | 1.94 | 1.83 |
| Propane | −0.125 | −0.024 | −0.031 | 1.307 | −0.299 | 1.98 | 1.96 |
| Benzene | −0.87 | −1.703 | −0.05 | 1.463 | −0.762 | −0.77 | −0.87 |
| Toluene | −1.037 | −1.717 | −0.058 | 1.568 | −0.769 | −0.86 | −0.89 |
| Naphthalene | −1.507 | −2.4 | −0.079 | 1.717 | −1.113 | −2.23 | −2.39 |
| Pyridine | −2.553 | −4.317 | −0.045 | 1.423 | −0.694 | −5.04 | −4.70 |
| 1,2,3-Trimethylbenzene | −1.423 | −1.717 | −0.075 | 1.74 | −0.857 | −1.18 | −1.22 |
| 1,2,4-Trimethylbenzene | −1.332 | −1.723 | −0.075 | 1.757 | −0.839 | −1.06 | −0.86 |
| 1,3,5-Trimethylbenzene | −1.317 | −1.755 | −0.077 | 1.794 | −0.885 | −1.09 | −0.90 |
| Ethylbenzene | −1.032 | −1.776 | −0.067 | 1.685 | −0.818 | −0.86 | −0.79 |
| *o*-xylene | −1.261 | −1.724 | −0.067 | 1.662 | −0.824 | −1.06 | −0.90 |
| *m*-xylene | −1.211 | −1.74 | −0.068 | 1.684 | −0.83 | −1.01 | −0.84 |
| *p*-xylene | −1.195 | −1.74 | −0.068 | 1.68 | −0.853 | −1.03 | −0.81 |
| Methanol | −1.801 | −3.823 | −0.017 | 1.06 | −0.231 | −3.66 | −5.11 |
| Ethanol | −1.732 | −4.173 | −0.025 | 1.214 | −0.297 | −3.86 | −5.01 |
| Acetic acid | −1.811 | −6.872 | −0.026 | 1.219 | −0.369 | −6.71 | −6.70 |
| Acetamide | −4.118 | −7.594 | −0.03 | 1.243 | −0.425 | −9.77 | −9.75 |
| Propionamide | −4.1 | −7.953 | −0.038 | 1.367 | −0.467 | −10.04 | −9.42 |
| *t-N*-Methylacetamide | −4.631 | −7.493 | −0.038 | 1.422 | −0.441 | −10.03 | −10.07 |
| Dimethylacetamide | −4.32 | −6.247 | −0.044 | 1.507 | −0.46 | −8.41 | −8.54 |
| Methyl amine | −2.221 | −3.914 | −0.021 | 1.083 | −0.309 | −4.23 | −4.56 |
| Ethyl amine | −2.205 | −4.202 | −0.03 | 1.226 | −0.369 | −4.43 | −4.50 |

[a] The parameters for the HFED-ECM were re-optimized for the 81 compounds, including those (e.g. benzene) that do not involve hydrogen bonds to water.

tion free energy was obtained by integration (summation) of the HFED. For the integration (summation), a grid model was used. Optimum values of $\Delta l$, $R_{\text{W}}$, and $R_{\text{shell}}$ for the summation were obtained. The calculated hydration free energies agree well with experimental data except for some molecules that form hydrogen bonds with water. In order to introduce the hydrogen bonding effect into the HFED calculation method, an effective charge model (ECM) was added to the HFED method. Since the HFED model is expressed in terms of an analytical function, it can be visualized both in two- and three-dimensional space.

Finally, for conformational energy calculations for molecules (whether included or not in the parameterization), the total energy, $E_{\text{tot}}$, is

$$E_{\text{tot}} = E_{\text{anhyd}} + \Delta G_{\text{hyd}} \tag{36}$$

where $E_{\text{anhyd}}$ is the anhydrous energy (calculated by any of a number of current force fields), and $\Delta G_{\text{hyd}}$ is calculated by integration of the HFED around the solute molecule, using Eq. (28) with the coefficients of Table 2 for the HFED model, or those of Eq. (35) for the HFED-ECM. For consistency, and for proper balance between $E_{\text{anhyd}}$ and $\Delta G_{\text{hyd}}$, the van der Waals radii of Table 1 and the charges used here should be used for the calculation of $\Delta G_{\text{hyd}}$.

### Acknowledgements

(a)

**Acetamide**



(b)
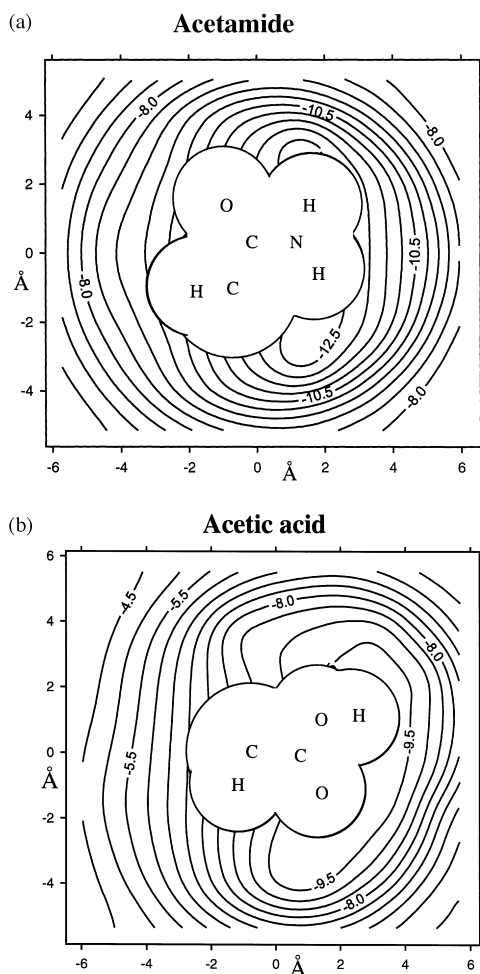
**Acetic acid**



Fig. 5. $\Delta G_{inter}$ is plotted as contours (in units of kcal/mole) for (a) acetamide and (b) acetic acid.

## References

[1] M.E. Davis, J.A. McCammon, Chem. Rev. 90 (1990) 509.

[2] J. Tomasi, M. Persico, Chem. Rev. 94 (1994) 2027.

[3] C.J. Cramer, D.G. Truhlar, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 6, VCH, New York, 1995, p. 1.

[4] J. Tomasi, B. Mennucci, R. Cammi, M. Cossi, in: G. Náray-Szabó, A. Warshel (Eds.), Computational Approaches to Biochemical Reactivity, vol. 19, Kluwer, Dordrecht, 1997, pp. 1–102.

[5] M. Born, Z. Phys. 1 (1920) 45.

[6] R.P. Bell, Trans. Farad. Soc. 27 (1931) 797.

[7] L. Onsager, J. Am. Chem. Soc. 58 (1936) 1486.

[8] J.G. Kirkwood, J. Chem. Phys. 2 (1934) 351.

[9] N.H. Frank, W. Tobocman, in: D.H. Menzel (Ed.), Fundamental Formulas of Physics, vol. 1, Dover, New York, 1960, p. 307.

[10] R. Fowler, E.A. Guggenheim, Statistical Thermodynamics, Cambridge University Press, London, 1956, p. 385.

[11] A.A. Rashin, B. Honig, J. Phys. Chem. 89 (1985) 5588.

[12] B. Jayaram, R. Fine, K. Sharp, B. Honig, J. Phys. Chem. 93 (1989) 4320.

[13] A.A. Rashin, K. Namboodiri, J. Phys. Chem. 91 (1987) 6003.

[14] A.A. Rashin, Int. J. Quantum Chem., Quantum Biol. Symp. 15 (1988) 103.

[15] A.A. Rashin, J. Phys. Chem. 94 (1990) 1725.

[16] F. Hirata, P. Redfern, R.M. Levy, Int. J. Quantum Chem., Quantum Biol. Symp. 15 (1988) 179.

[17] R.J. Zauhar, R.S. Morgan, J. Mol. Biol. 186 (1985) 815.

[18] R.J. Zauhar, R.S. Morgan, J. Comput. Chem. 9 (1988) 171.

[19] T. Kozaki, K. Morihashi, O. Kikuchi, J. Am. Chem. Soc. 111 (1989) 1547.

[20] B.J. Yoon, A.M. Lenhoff, J. Comput. Chem. 11 (1990) 1080.

[21] A. Jean-Charles, A. Nicholls, K. Sharp, et al., J. Am. Chem. Soc. 113 (1991) 1454.

[22] Y.N. Vorobjev, J.A. Grant, H.A. Scheraga, J. Am. Chem. Soc. 114 (1992) 3189.

[23] Y.N. Vorobjev, H.A. Scheraga, J. Phys. Chem. 97 (1993) 4855.

[24] Y.N. Vorobjev, H.A. Scheraga, B. Hitz, B. Honig, J. Phys. Chem. 98 (1994) 10940.

[25] Y.N. Vorobjev, H.A. Scheraga, B. Honig, J. Phys. Chem. 99 (1995) 7180.

[26] Y.N. Vorobjev, H.A. Scheraga, J. Comput. Chem. 18 (1997) 569.

[27] N.K. Rogers, M.J.E. Sternberg, J. Mol. Biol. 174 (1984) 527.

[28] D.T. Edmonds, N.K. Rogers, M.J.E. Sternberg, Mol. Phys. 52 (1984) 1487.

[29] N.K. Rogers, G.R. Moore, M.J.E. Sternberg, J. Mol. Biol. 182 (1985) 613.

[30] N.K. Rogers, Prog. Biophys. Mol. Biol. 48 (1986) 37.

[31] M.E. Davis, J.A. McCammon, J. Comput. Chem. 10 (1989) 386.

[32] B.A. Luty, M.E. Davis, J.A. McCammon, J. Comput. Chem. 13 (1992) 768.

[33] V. Mohan, M.E. Davis, J.A. McCammon, B.M. Pettitt, J. Phys. Chem. 96 (1992) 6428.

[34] M.K. Gilson, B. Honig, Proteins 4 (1988) 7.

[35] M.K. Gilson, K.A. Sharp, B.H. Honig, J. Comput. Chem. 9 (1988) 327.

[36] A. Nicholls, B. Honig, J. Comput. Chem. 12 (1991) 435.

[37] B. Honig, K. Sharp, A.S. Yang, J. Phys. Chem. 97 (1993) 1101.

[38] B. Honig, A. Nicholls, Science 268 (1995) 1144.

[39] K. Sharp, A. Jean-Charles, B. Honig, J. Phys. Chem. 96 (1992) 3822.

[40] J. Langlet, P. Claverie, J. Caillet, A. Pullman, J. Phys. Chem. 92 (1988) 1617.

[41] H. Nakamura, J. Phys. Soc. Jpn 57 (1988) 3702.

[42] D. Bashford, M. Karplus, G.W. Canters, J. Mol. Biol. 203 (1988) 507.

[43] G. Némethy, H.A. Scheraga, J. Chem. Phys. 36 (1962) 3401.

[44] K.D. Gibson, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. 58 (1967) 420.

[45] Z.I. Hodes, G. Némethy, H.A. Scheraga, Biopolymers 18 (1979) 1565.

[46] G. Némethy, W.J. Peer, H.A. Scheraga, Annu. Rev. Biophys. Bioeng. 10 (1981) 459.

[47] T. Ooi, M. Oobatake, G. Némethy, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. 84 (1987) 3086.

[48] K.D. Gibson, H.A. Scheraga, Mol. Phys. 62 (1987) 1247.

[49] K.D. Gibson, H.A. Scheraga, Mol. Phys. 64 (1988) 641.

[50] Y.K. Kang, G. Némethy, H.A. Scheraga, J. Phys. Chem. 91 (1987) 4105.

[51] Y.K. Kang, G. Némethy, H.A. Scheraga, J. Phys. Chem. 91 (1987) 4109.

[52] Y.K. Kang, G. Némethy, H.A. Scheraga, J. Phys. Chem. 91 (1987) 4118.

[53] Y.K. Kang, K.D. Gibson, G. Némethy, H.A. Scheraga, J. Phys. Chem. 92 (1988) 4739.

[54] J. Vila, R.L. Williams, M. Vasquez, H.A. Scheraga, Proteins: Struct. Funct. Gen. 10 (1991) 199.

[55] G. Perrot, B. Cheng, K.D. Gibson, et al., J. Comput. Chem. 13 (1992) 1.

[56] A.J. Hopfinger, Macromolecular 4 (1971) 731.

[57] A.J. Hopfinger, Macromolecular 5 (1972) 334.

[58] A. Shrake, J.A. Rupley, J. Mol. Biol. 79 (1973) 351.

[59] B. Lee, F.M. Richards, J. Mol. Biol. 55 (1971) 379.

[60] T.J. Richmond, F.M. Richards, J. Mol. Biol. 119 (1978) 537.

[61] T.J. Richmond, J. Mol. Biol. 178 (1984) 63.

[62] M.L. Connolly, J. Appl. Crystallogr. 16 (1983) 548.

[63] M.L. Connolly, Science 221 (1983) 709.

[64] M.L. Connolly, J. Am. Chem. Soc. 107 (1985) 1118.

[65] D. Eisenberg, A.D. McLachlan, Nature 319 (1986) 199.

[66] L. Wesson, D. Eisenberg, Protein Sci. 1 (1992) 227.

[67] A.Y. Meyer, J. Chem. Soc., Perkin Trans. 2 2 (1985) 1161.

[68] A.Y. Meyer, J. Comput. Chem. 7 (1986) 144.

[69] G. Perrot, B. Maigret, J. Mol. Graph. 8 (1990) 141.

[70] C.A. Schiffer, J.W. Caldwell, P.A. Kollman, R.M. Stroud, Mol. Simul. 10 (1993) 121.

[71] S.M. Le Grand, K.M. Merz, Jr., J. Comput. Chem. 14 (1993) 349.

[72] M.-J. Huron, P. Claverie, J. Phys. Chem. 76 (1972) 2123.

[73] M.-J. Huron, P. Claverie, J. Phys. Chem. 78 (1974) 1853.

[74] M.-J. Huron, P. Claverie, J. Phys. Chem. 78 (1974) 1862.

[75] O. Tapia, O. Goscinski, Mol. Phys. 29 (1975) 1653.

[76] O. Tapia, G. Johannin, J. Chem. Phys. 75 (1981) 3624.

[77] S. Miertus, E. Scrocoo, J. Tomasi, Chem. Phys. 55 (1981) 117.

[78] S. Miertus, J. Tomasi, Chem. Phys. 65 (1985) 239.

[79] J.L. Pascual-Ahuir, E. Silla, J. Tomasi, R. Bonaccorsi, J. Comput. Chem. 8 (1987) 778.

[80] R. Cammi, J. Tomasi, J. Chem. Phys. 100 (1994) 7495.

[81] R. Cammi, J. Tomasi, J. Chem. Phys. 101 (1994) 3888.

[82] C.S. Pomelli, J. Tomasi, Theor. Chim. Acta. 96 (1997) 39.

[83] M.W. Wong, K.B. Wiberg, M. Frisch, J. Chem. Phys. 95 (1991) 8991.

[84] M.W. Wong, M.J. Frisch, K.B. Wiberg, J. Am. Chem. Soc. 113 (1991) 4776.

[85] M.W. Wong, K.B. Wiberg, M.J. Frisch, J. Am. Chem. Soc. 114 (1992) 523.

[86] M.W. Wong, K.B. Wiberg, M.J. Frisch, J. Am. Chem. Soc. 114 (1992) 1645.

[87] J.B. Foresman, T.A. Keith, K.B. Wiberg, J. Snoonian, M.J. Frisch, J. Phys. Chem. 100 (1996) 16098.

[88] J.A. Grant, R.L. Williams, H.A. Scheraga, Biopolymers 30 (1990) 929.

[89] I. Tuñón, E. Silla, J.L. Pascual-Ahuir, Chem. Phys. Lett. 203 (1993) 289.

[90] D. Rinaldi, M.F. Ruiz-Lopez, J.-L. Rivail, J. Chem. Phys. 78 (1983) 834.

[91] C. Chipot, D. Rinaldi, J.-L. Rivail, Chem. Phys. Lett. 91 (1992) 287.

[92] D. Rinaldi, J.-L. Rivail, N. Rguini, J. Comput. Chem. 13 (1992) 675.

[93] V. Dillet, D. Rinaldi, J.-L. Rivail, J. Phys. Chem. 98 (1994) 5034.

[94] K.K. Stavrev, T. Tamm, M.C. Zerner, Int. J. Quantum Chem: Quantum Chem. Symp. 30 (1996) 373.

[95] C. Amovilli, B. Mennucci, J. Phys. Chem. B 101 (1997) 1051.

[96] A. Warshel, S.T. Russell, Q. Rev. Biophys. 17 (1984) 283.

[97] S.T. Russell, A. Warshel, J. Mol. Biol. 185 (1985) 389.

[98] V. Luzhkov, A. Warshel, J. Comput. Chem. 13 (1992) 199.

[99] C.J. Cramer, D.G. Truhlar, J. Am. Chem. Soc. 113 (1991) 8305.

[100] C.J. Cramer, D.G. Truhlar, Science 256 (1992) 213.

[101] C.J. Cramer, D.G. Truhlar, J. Comput. Chem. 13 (1992) 1089.

[102] C.J. Cramer, D.G. Truhlar, J. Comp. Aided Mol. Des. 6 (1992) 629.

[103] C.C. Chambers, G.D. Hawkins, C.J. Cramer, D.G. Truhlar, J. Phys. Chem. 100 (1996) 16385.

[104] K.T. No, J.A. Grant, M.S. Jhon, H.A. Scheraga, J. Phys. Chem. 94 (1990) 4740.

[105] K.T. No, J.A. Grant, H.A. Scheraga, J. Phys. Chem. 94 (1990) 4732.

[106] K.T. No, K.H. Cho, M.S. Jhon, H.A. Scheraga, J. Am. Chem. Soc. 115 (1993) 2005.

[107] K.T. No, O.Y. Kwon, S.Y. Kim, et al., J. Phys. Chem. 99 (1995) 13019.

[108] K.T. No, O.Y. Kwon, S.Y. Kim, M.S. Jhon, H.A. Scheraga, J. Phys. Chem. 99 (1995) 3478.

[109] J.C. Slater, J.G. Kirkwood, Phys. Rev. 37 (1931) 682.

[110] Cerius2, MSI, San Diego (1997).

[111] S. Cabani, P. Gianni, V. Mollica, L. Lepori, J. Solution Chem. 10 (1981) 563.

[112] R.G. Pearson, J. Am. Chem. Soc. 108 (1986) 6109.

[113] M.H. Abraham, J. Andonian-Haftvan, G.S. Whiting, A. Leo, R.S. Taft, J. Chem. Soc. Perkin Trans. 2 (1994) 1777.