



Optimized atomic radii for quantum dielectric continuum solvation models

Eugene V. Stefanovich¹, Thanh N. Truong^{*}

Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA

Received 23 June 1995; in final form 26 July 1995

Abstract

We present an optimized set of atomic radii for H, C, N, O, F, P, S and Cl, which can be used in quantum dielectric continuum calculations. These atomic radii yield differences in hydration energies with experimental data of about 1 kcal/mol for neutral molecules, 2 kcal/mol for cations and 2–5 kcal/mol for anions at the Hartree–Fock, second order Møller–Plesset perturbation, and various non-local density functional theories within the framework of the generalized conductor-like screening model. Almost the same accuracy was also found when these radii were used with the polarizable continuum model (PCM). The accuracy of the recently proposed self-consistent isodensity PCM model is also discussed.

1. Introduction

Solvent effects play an important role in chemistry and biochemistry. From a theoretical point of view, dielectric continuum solvation theory provides a simple, accurate and efficient approach for incorporating the solvent effects into the quantum-mechanical (QM) calculations of the solute molecules. In this approach, the solvent reaction field can be represented by the charge density $\sigma(\mathbf{r})$ on the surface of the cavity in a dielectric medium. This surface charge density is determined self-consistently with the solute electron density distribution $\rho(\mathbf{r})$. An excellent review of different dielectric continuum solvation models was given by Tomasi and Persico [1].

It is clear that the accuracy of such a solvation theory critically depends on the size and shape of the cavity. There are two different approaches for defining the cavity. One is to use a spherical or ellipsoidal cavity with an effective size. Such simplistic approach allows more sophisticated treatments of solvent effects in quantum electronic structure theories and has provided useful insights [1]. The other approach is to define the cavity with a more realistic shape according to the molecular structure of the solute. Such a cavity can be formed from a set of interlocking spheres with centers on solute atoms. The boundary surface is then determined either by the exposed regions of these spheres (van der Waals (VDW) surface) [2], or the trace of the center of the solvent sphere rolling on the van der Waals surface (solvent accessible surface) [3], or the smooth surface created by rolling the solvent sphere on the VDW surface (molecular surface) [4]. The radii of

^{*} Corresponding author.

¹ On leave from the Institute of Chemical Physics, University of Latvia, 19 Rainis Blvd., Riga LV1586, Latvia.

the atomic spheres are empirically fitted parameters depending on the atom type. Several sets of atomic radii were proposed in the literature. According to recent comparative analysis [5], the most accurate hydration free energies can be obtained with the atomic radii suggested by Rashin and Namboodiri [6]. In the original version of the polarizable continuum model (PCM) [7] to be described below, it was suggested that atomic radii can be obtained by scaling the standard van der Waals radii of atoms. However, further studies showed that the use of an unique scaling factor is not accurate. Different scaling factors (1.20–1.25) for neutral molecules [8] and (1.10–1.15) for ions [9] were then suggested. However, such multiple scaling factors are not particularly suitable for studying charge transfer chemical reactions. To avoid this difficulty, a more complex computational scheme with variable atomic radii depending on the partial charges of atoms was proposed [10]. This approach improves the accuracy, however, at the expense of introducing additional basis set dependent parameters. A logical continuation on these ideas is to define the boundary of the cavity as an isodensity surface which depends only on one parameter – the isodensity value. This surface was recently introduced within the PCM formalism in the GAUSSIAN 94 computer program [11]. However, our recent study of the S_N2 $Cl^- + CH_3Cl$ chemical reaction in aqueous solvent [12] indicated that this interesting approach is not accurate for calculating reaction energy profiles of charge transfer reactions and requires further testing.

Once the cavity boundary is defined, the solvent reaction field can be determined within the boundary element framework by using either the PCM or generalized conductor-like screening model (GCOSMO) [12–15]. Although other *ab initio* approaches [1], including the recent coupled DFT/classical Poisson–Boltzmann method [16–18], that was not included in the review by Tomasi and Persico [1], can also be used for calculations of the solvent reaction field, our present work focuses on the PCM and GCOSMO models, thus we briefly describe them here.

The PCM model was first proposed by Miertus et al. [7] and has been further developed by numerous authors. In this model, the self-consistent surface charge density is calculated from the normal compo-

nent of the total electric field $E_n(\mathbf{r})$ produced on the cavity surface by both the solute and surface charges,

$$\sigma(\mathbf{r}) = \frac{1 - \epsilon}{4\pi\epsilon} E_n(\mathbf{r}), \quad (1)$$

where ϵ is the dielectric constant of the solvent. In most implementations of the PCM model, this charge density is additionally scaled to satisfy explicitly the Gauss theorem for the total surface charge Q_s ,

$$Q_s \equiv \int_S \sigma(\mathbf{r}) d^2r = \frac{1 - \epsilon}{\epsilon} Q, \quad (2)$$

where Q is the total charge of the solute [7]. Such a scaling corrects numerical errors in the classical PCM model whereas it also corresponds to the assumption that the electron density of the solute is localized completely inside the cavity in the QM PCM model. The self-consistent isodensity PCM (SCIPCM) model [11] is a recent variation based on the above PCM formalism where the isodensity surface was also converged along with the solute and surface charge densities.

The generalized conductor-like screening model (GCOSMO) was recently proposed by us for incorporating solvent effects into the classical, *ab initio* molecular orbital and density functional theory (DFT) frameworks [12–15]. It is based on the original COSMO model of Klamt and Schüürman [19] where the solute charge distribution within a semiempirical molecular orbital theory is represented by a set of point charges and dipoles. GCOSMO, however, has a general description for the solute charge distribution and also includes dispersion, repulsion and cavitation contributions. Calculation of both COSMO and GCOSMO surface charges consists of two steps. First, the surface charge density $\sigma_{sc}(\mathbf{r})$ is determined from the boundary condition for a cavity in a screening conductor ($\epsilon = \infty$),

$$\sum_i \frac{z_i}{|\mathbf{r} - \mathbf{R}_i|} - \int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' + \int_S \frac{\sigma_{sc}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^2r' = 0, \quad (3)$$

where z_i are nuclear charges, V and S are volume and surface of the cavity, respectively, and vector \mathbf{r} lies on S . The second step is to scale $\sigma_{sc}(\mathbf{r})$ by an appropriate constant factor depending on the dielectric constant of the solvent,

$$\sigma(\mathbf{r}) = f(\epsilon) \sigma_{sc}(\mathbf{r}). \quad (4)$$

In the GCOSMO method, the scaling factor of

$$f(\epsilon) = (\epsilon - 1)/\epsilon \quad (5)$$

is chosen to satisfy the Gauss theorem. Note, that unlike in the PCM model, this scaling is theoretically motivated and is not intended to correct numerical errors nor to assume the solute charge distribution being entirely inside the cavity.

We are interested presently only in the aqueous solvent ($\epsilon = 80$) for which the scaling factor (5) is quite accurate [13]. In fact, one can expect an error of the order of ϵ^{-1} in hydration energies when using this scaling factor. For most molecular species, this corresponds to less than 1 kcal/mol difference with the more accurate PCM results, providing that the cavity size and boundary elements are the same for both methods. Furthermore, previous studies [14,20–22] have shown that in most cases the solute electron correlation effects change hydration free energies by only 1–2 kcal/mol in the second-order Møller–Plesset perturbation (MP2) and density functional theories. These findings support the following hypothesis: there exists a unique set of optimized atomic radii that can be used with either the PCM or GCOSMO model for calculations of hydration energies within the frameworks of either Hartree–Fock (HF), MP2 or density functional theories. The goal of this study is to determine this set of atomic radii for the elements that are most important in organic and biochemistry: H, C, N, O, F, P, S and Cl. This is done by optimizing the atomic radii R_j to experimental hydration free energies, ΔG_{exp}^k , for a selected set of molecules with different functional groups and ions. In other words, we minimized the average unsigned error function

$$E\{R_j\} = \frac{1}{N} \sum_{k=1}^N |\Delta G_{\text{QM}}^k - \Delta G_{\text{exp}}^k|, \quad (6)$$

for a representative set of N molecular species labelled with the index k , where ΔG_{QM}^k are calculated hydration free energies.

In spite of previous attempts [8,9,23], complete optimization of the atomic radii by solving the minimization problem above has not yet been done. This is because such an optimization, if implemented straightforwardly, requires thousands of QM calculations; that is impractical. In the present Letter, we suggest a practical method for minimizing the func-

tion (6), which avoids heavy QM calculations. This method is based on the good agreement between QM and classical GCOSMO models observed in our previous works [12,14].

Due to the recent availability of the PCM and SCIPCM solvation models in the widely used GAUSSIAN 94 program [11], the number of applications of these methods will sharply increase. Thus, it is also important and timely in this study (a) to provide assessment on the accuracy of the atomic radii used as a default in the GAUSSIAN 94 program for the PCM model (results denoted as PCM/G94) by comparing with the PCM results calculated with our optimized radii (denoted as PCM/ST); (b) to check the accuracy of the SCIPCM method.

In Section 2, we outline the optimization procedure used here and other computational details. In Section 3, we present the resulting set of atomic radii and compare GCOSMO hydration energies calculated at different levels of theory and with those computed from the PCM/G94, PCM/ST and SCIPCM models. Conclusions are formulated in Section 4.

2. Method

We begin with the procedure for optimizing the atomic radii for H, C, N, O, P and S. Previous work [24] (and references therein) showed that for better agreement with experiment, the aromatic carbon center should have a different atomic radius than other types of carbon center. Generalizing on this idea, we treated carbon centers with sp^2 or sp hybridizations ($C_{\sigma\pi}$) different from those with sp^3 hybridization (C_{σ}). Furthermore, atomic radii of nitrogen in hydrogen-bonding (N_{hb}) and non-hydrogen-bonding (N_{nhb}) environments may differ substantially [6]. Thus, the error function in Eq. (6) depends on 8 variables. We selected a set of 17 neutral polar and non-polar molecules, 7 anions and 14 cations for calculating this error function. This representative set (marked with asterisk in Tables 2 and 3) includes molecules with different functional groups and ions that are important in organic and biochemistry. The geometries of these molecules were first optimized in the gas phase at the HF/6-31G(d) level. These opti-

mized geometries were used in all subsequent solvation calculations. The 6-31G(d) basis set was also used throughout except for the fitting of F and Cl atomic radii as described below.

For the QM calculations in solution, we have chosen the hybrid non-local DFT BH and HLYP method as implemented in the G92/DFT computer program [25]. From our own experience, this method yields reasonably accurate properties not only at equilibrium structures but also at transition states. The self-consistent electrostatic solvent reaction field was calculated using the GCOSMO solvation model. The cavity surfaces were defined as Molecular Surfaces generated by the GEPOL93 computer program [26] assuming the solvent radius of 1.4 Å. The cavity surface was divided into 60 surface elements per each full atomic sphere. This discretization level was shown to yield converged results for aqueous solvent [13].

The classical dispersion and repulsion contributions to the hydration free energy were calculated by using the Floris et al. method [27] with the pair-potential parameters from the OPLS force field [28]. For the cavitation term, we employed the modified scaled particle fluid theory proposed by Huron and Claverie [29]. For calculation of these three non-electrostatic contributions the radius of the spherical solvent water R_w is required. Since, the non-electrostatic terms are dominant in the non-polar hydrocarbon molecules, we first optimized R_w for 5 linear alkanes, C_nH_{2n+2} ($n = 1-5$) using atomic radii for C and H from Ref. [14]. This optimization yielded R_w equal to 1.29 Å that is somewhat smaller than value of 1.4 Å recommended in Ref. [27].

The automated optimization of the atomic radii was carried out using an iterative procedure consisting of three steps. First, we performed both BH and HLYP and classical GCOSMO calculations for the whole set of molecules using the initial guess of atomic radii from our previous work [14] and the

optimized $R_w = 1.29$ Å. Partial charges used in the classical GCOSMO calculations were determined from fitting to the solute gas-phase electrostatic potential by using the CHELPG method [30]. Second, for each molecule, we defined an effective hydration energy ΔG_{eff} by adding to the experimental value ΔG_{exp} the difference between the classical and quantum hydration free energies $\Delta G_{\text{class}} - \Delta G_{\text{QM}}$, i.e.

$$\Delta G_{\text{eff}} = \Delta G_{\text{exp}} + \Delta G_{\text{class}} - \Delta G_{\text{quant}} \quad (7)$$

The central idea is to optimize the atomic radii using the simple classical GCOSMO model to reproduce ΔG_{eff} , correspondingly, quantum GCOSMO hydration energies ΔG_{QM} are expected to be in good agreement with ΔG_{exp} . This is done in the third step. We repeated this three-step procedure several times until convergence.

To optimize atomic radii for F and Cl, we performed separate BH and HLYP-GCOSMO calculations for F^- and Cl^- anions using the 6-31 + G* basis set. Addition of the diffuse functions is necessary for representing the delocalized character of the valence electron density of these ions. For all quantum GCOSMO calculations we used our locally modified version of the GAUSSIAN 92/DFT program [25].

For the SCIPCM calculations we used the suggested isodensity value of 0.001 au [11]. The van der Waals cavities (the only choice currently available in the GAUSSIAN 94 program) were employed for the PCM/ST and PCM/G94 calculations.

3. Results and discussion

The optimized atomic radii listed in Table 1 resemble that from our previous set [14] with the largest difference of 0.24 Å for the hydrogen-bonding nitrogen. The calculated deviations $\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$ are listed in Tables 2–5 along with the experi-

Table 1
Atomic radii (in Å)

	H	C_σ	$C_{\sigma\pi}$	N_{hb}	N_{nhb}	O	F	P	S	Cl
Ref. [6]	1.16	2.46	2.46	1.5	2.3	1.5	1.423		1.97	1.937
Ref. [14]	1.16	2.3	1.7	1.5	2.2	1.4	1.423	2.35	1.97	1.937
this work	1.172	2.096	1.635	1.738	2.126	1.576	1.28	2.279	2.023	1.75

mental hydration free energies which were taken from Refs. [31,32]. Thus, in order to obtain the actual calculated hydration free energies one should add to these deviations the experimental values. To test the transferability of our fitted radii to other molecular species, we performed BH and HLYP-GCOSMO calculations for the additional molecules CH_3O^- , $\text{C}_6\text{H}_5\text{NH}_3^+$, HF, CH_3Cl , MeCONH_2 . We also calculated hydration energies of HPO_3 and $\text{H}_4\text{P}_2\text{O}_7$ molecules. Although no experimental data is available for these two molecules, their importance in biosynthetic reactions, such as the synthesis of polynucleotides and recent findings [33] on the failure of the semiempirical SMx solvation models in calculating their hydration free energies support the inclusion of such calculations here. In these two cases, we report the actual computed ΔG_{QM} in Ta-

bles 2 and 4 and do not include these results in the computation of the average errors.

To confirm our hypothesis on the transferability of the fitted atomic radii among different levels of theory, we performed GCOSMO calculations for the same set of molecules and ions at the HF, MP2 (in the PTE approximation from Ref. [34]), BLYP and B3LYP levels. We found that the average unsigned errors (see Eq. (6)) are of the order of 1 kcal/mol for neutral polar and non-polar molecules, 2 kcal/mol for cations and 2–5 kcal/mol for anions at all levels of theory considered here (see Tables 2 and 3). Note that the experimental uncertainty in the hydration energies of ions is of the order of 2 kcal/mol or more [32] whereas it is smaller for neutral molecules. Thus, our method is quite accurate for neutral molecules and cations but not as

Table 2
Experimental hydration energies and GCOSMO calculated $\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$ differences for neutral molecules (in kcal/mol) ^a

	ΔG_{exp}	$\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$				
		BH&HLYP	HF	MP2	BLYP	B3LYP
HF	-5.6	-1.78	-2.33	-1.89	-0.87	-1.29
* H ₂ O	-6.3	-2.95	-3.52	-3.32	-2.03	-2.49
* NH ₃	-4.3	-2.11	-2.38	-2.40	-1.60	-1.88
CH ₃ Cl	-0.6	-1.43	-1.69	-1.23	-1.05	-1.23
* CH ₃ OH	-5.1	0.17	-0.35	-0.01	1.03	0.62
* CH ₃ SH	-1.2	-0.53	-0.59	-0.61	-0.34	-0.45
* CH ₃ CN	-3.9	0.04	-0.54	0.62	0.69	0.40
* MeNH ₂	-4.6	0.89	0.63	0.65	1.39	1.13
* MeCOOH	-6.7	-0.84	-1.93	-0.15	0.49	-0.11
* MeOMe	-1.9	0.75	0.31	0.70	1.60	1.19
MeCONH ₂	-9.7	1.93	-1.66	0.46	1.17	0.44
* MeCOMe	-3.0	-0.55	-1.61	0.50	0.69	0.15
* MeCOOMe	-3.3	-0.40	-1.34	0.32	0.78	0.26
* C ₅ H ₅ N	-4.7	-0.65	-1.42	-0.39	0.50	-0.03
* C ₆ H ₅ OH	-6.6	-1.43	-2.23	-1.31	0.00	-0.69
* CH ₄	2.00	0.08	0.10	0.07	0.06	0.05
* C ₂ H ₆	1.83	0.65	0.66	0.62	0.65	0.62
* C ₃ H ₈	1.96	0.86	0.87	0.82	0.86	0.83
* n-C ₄ H ₁₀	2.08	0.98	0.99	0.94	0.99	0.96
* n-C ₅ H ₁₂	2.33	1.27	1.27	1.22	1.28	1.24
average unsigned error		1.01	1.32	0.91	0.90	1.25
HPO ₃ ^b	(-18.6) ^c	-12.08	-14.89	-10.33	-8.74	-10.16
H ₄ P ₂ O ₇ ^b	(-20.0) ^c	-32.52	-35.67	-31.93	-28.64	-30.39

^a Molecules marked with an asterisk were used for fitting of atomic radii in Table 1. 'Me' denotes the methyl group.

^b ΔG_{QM} is shown in columns 2–6 instead of $\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$.

^c Taken from Ref. [35].

accurate for anions. At least partially, this problem is related to the diffuse character of the wave function describing the negatively charged species that leads to the strong basis set dependence of the hydration energies. For instance, ΔG_{QM} for F^- and Cl^- anions computed with the 6-31 + G(d) basis sets are -102.33 and -72.70 kcal/mol, respectively, in good agreement with the experimental data. Without the diffuse functions (basis 6-31G(d)), the ΔG_{QM} is overestimated by about 15 and 7 kcal/mol for these two ions. This is because electrons are localized more inside the cavity with the 6-31G(d) basis set

than with the 6-31 + G(d) one. The largest difference (13.72 kcal/mol) between the calculated and experimental hydration free energies is in the case of the CH_3O^- molecule. The origin of such discrepancies is not clear at this point.

The above results indicate that the hydration free energies for most equilibrium molecular species are not very sensitive to correlation effects. The CH_3O^- , $\text{C}_6\text{H}_5\text{O}^-$, HPO_3 and $\text{H}_4\text{P}_2\text{O}_7$ molecules are exceptional: inclusion of the electron correlation changes the hydration energies of these molecules by about 3–8 kcal/mol. This magnitude of the correlation

Table 3
Experimental hydration energies and GCOSMO calculated $\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$ for ions (in kcal/mol)^a

	ΔG_{exp}	$\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$				
		BH&HLYP	HF	MP2	BLYP	B3LYP
anions						
F^-	-105	-14.16	-14.31	-13.88	-14.10	-14.83
Cl^-	-75	-4.34	-4.98	-3.99	-4.30	-6.54
* OH^-	-104	0.26	-0.50	-0.27	1.58	0.91
* SH^-	-74	-0.62	-0.27	-0.47	-0.57	-0.63
* NH_2^-	-93	0.35	-0.32	0.01	1.71	1.01
* PH_2^-	-65	-0.18	0.10	-0.05	-0.32	-0.29
CH_3O^-	-93	13.72	8.79	14.07	17.03	15.53
* CH_3S^-	-77	4.76	4.59	5.04	5.30	4.99
* MeCOO^-	-75	1.06	-0.86	2.31	3.74	2.54
* $\text{C}_6\text{H}_5\text{O}^-$	-70	4.65	2.56	6.02	8.19	6.55
average unsigned error ^b		3.20	2.25	3.53	4.81	4.06
cations						
* NH_4^+	-77	-0.01	-0.04	-0.04	0.09	0.04
* MeNH_3^+	-68	-1.54	-1.52	-1.67	-1.55	-1.55
* Me_2NH_2^+	-61	-2.27	-2.24	-2.43	-2.33	-2.31
* Me_3NH^+	-57	-1.25	-1.23	-1.41	-1.31	-1.29
$\text{C}_6\text{H}_5\text{NH}_3^+$	-66	0.53	-0.07	0.45	1.46	0.85
* PH_4^+	-71	0.28	0.28	0.16	0.27	0.27
* MePH_3^+	-64	0.08	-0.17	-0.15	0.35	0.23
* Me_2PH_2^+	-55	-3.61	-4.04	-3.87	-3.20	-3.36
* Me_3PH^+	-50	-4.23	-4.78	-4.49	-3.75	-3.94
* H_3O^+	-102	5.90	5.83	5.70	5.99	5.95
* MeOH_2^+	-83	4.95	5.23	4.66	4.64	4.76
* Me_2OH^+	-68	1.82	1.96	1.69	1.65	1.72
* H_3S^+	-85	0.97	0.66	0.66	1.18	1.07
* MeSH_2^+	-72	0.80	0.63	0.33	1.04	0.92
* Me_2SH^+	-59	-3.18	-3.46	-3.56	-2.88	-3.02
average unsigned error		2.09	2.14	2.08	2.11	2.09

^a Molecules marked with an asterisk were used for fitting of atomic radii in Table 1. 'Me' denotes the methyl group.

^b Not including results for F^- and Cl^- .

effect is similar to that found in our previous study for the hydration energy of the transition state for the S_N2 $Cl^- + CH_3Cl$ reaction [12].

PCM/ST results reported in Tables 4 and 5 prove the applicability of our optimized radii for the PCM solvation model. The electrostatic hydration energies have average unsigned errors of 2.05 kcal/mol for neutral molecules, 3.76 kcal/mol for cations, and 5.05 kcal/mol for anions, respectively. However, with the inclusion of the non-electrostatic contributions taken from our GCOSMO calculations the level of accuracy becomes similar to that of the GCOSMO model: 1.20, 2.99 and 3.83 kcal/mol for neutrals, cations and anions, respectively (see Tables 4 and 5). The remaining differences between PCM/ST and GCOSMO results can be explained by the following factors: (1) different surface charge scaling procedures; (2) different cavity surfaces: the van der Waals in the PCM/ST model and molecular surface in the

GCOSMO model; (3) different surface discretization procedures in the two models. A more systematic comparison on the performance of these two models within the HF, MP2 and DFT frameworks will be presented in a forthcoming paper.

The results of SCIPCM and PCM/G94 electrostatic calculations are also summarized in Tables 4 and 5. They are less accurate than GCOSMO and PCM/ST results, especially for ions. The inaccuracy of the PCM/G94 model is due to the inadequate atomic radii used and neglect of the non-electrostatic contributions to the ΔG_{OM} . It is interesting to note that in the SCIPCM method the ΔG_{OM} values for negative ions are systematically underestimated and those for positive ions overestimated. These systematic errors result from the main deficiency of the isodensity surface. In particular, the electron density of anions is very delocalized, thus the isodensity surface lies far from the atoms, this corresponds to

Table 4
Comparison of the GCOSMO and PCM hydration energy results for neutral molecules (in kcal/mol) obtained at the BH&HLYP level

	ΔG_{exp}	$\Delta G_{OM} - \Delta G_{exp}$			
		GCOSMO	PCM/ST	SCIPCM	PCM/G94
HF	-5.6	-1.78	1.96	-2.61	0.13
H ₂ O	-6.3	-2.95	-2.32	-2.02	-0.74
NH ₃	-4.3	-2.11	-2.06	-1.03	-2.27
CH ₃ Cl	-0.6	-1.43	1.69	-2.30	0.04
CH ₃ OH	-5.1	0.17	1.00	-0.60	-0.97
CH ₃ SH	-1.2	-0.53	0.27	0.30	0.10
CH ₃ CN	-3.9	0.04	0.74	-3.03	-3.32
MeNH ₂	-4.6	0.89	1.21	0.53	-0.76
MeCOOH	-6.7	-0.84	-0.15	1.70	-3.50
MeOMe	-1.9	0.75	0.18	-1.03	-3.14
MeCONH ₂	-9.7	1.93	2.69	1.86	-4.34
MeCOMe	-3.0	-0.55	0.57	-2.60	-3.73
MeCOOMe	-3.3	-0.40	-0.26	-2.55	-6.24
C ₅ H ₅ N	-4.7	-0.65	-0.43	0.28	-2.52
C ₆ H ₅ OH	-6.6	-1.43	3.36	-0.43	3.36
CH ₄	2.00	0.08	-0.13	-2.31	-2.02
C ₂ H ₆	1.83	0.65	0.95	-2.12	-1.83
C ₃ H ₈	1.96	0.86	1.19	-2.32	-1.96
n-C ₄ H ₁₀	2.08	0.98	1.21	-2.56	-2.09
n-C ₅ H ₁₂	2.33	1.27	1.64	-2.88	-2.35
average unsigned error		1.01	1.20	1.75	2.27
HPO ₃ ^a	(-18.6) ^b	-12.08	-8.97	-14.13	-11.99
H ₄ P ₂ O ₇ ^a	(-20.0) ^b	-32.52	-30.26	-33.01	-30.23

^a ΔG_{OM} is shown in columns 2–5 instead of $\Delta G_{OM} - \Delta G_{exp}$.

^b Taken from Ref. [35].

effectively larger atomic radii and lower (less negative) hydration energies. The electron density of cations is generally localized closer to the nuclei, this corresponds to effective shrinking of ‘atomic radii’ and higher hydration energies. As we pointed out in Ref. [12], this strong dependence of the cavity shape on the charges may lead not only to inaccurate computed ΔG_{QM} value but also to incorrect reaction profiles for chemical reactions involving charged species. It is important to also point out that the SCIPCM model is computationally comparable to the GCOSMO model whereas the PCM model is

about several times slower and the SCF step sometimes fail to converge, as in the case of $\text{C}_6\text{H}_5\text{NH}_3^+$.

Hydration energies for HPO_3 and $\text{H}_4\text{P}_2\text{O}_7$ molecules computed in this work with different ab initio dielectric continuum models are about -9 to -15 kcal/mol for HPO_3 , and -29 to -36 kcal/mol for $\text{H}_4\text{P}_2\text{O}_7$ (see Tables 2 and 4). Due to the consistency of these results we believe that they provide a better estimation for the hydration free energies of HPO_3 and $\text{H}_4\text{P}_2\text{O}_7$ than -18.6 and -20.0 kcal/mol, respectively, obtained from previous PCM calculations [33,35].

Table 5
Comparison of the GCOSMO and PCM hydration energy results for ions (in kcal/mol) obtained at the BH&HLYP level

	ΔG_{exp}	$\Delta G_{\text{QM}} - \Delta G_{\text{exp}}$			
		GCOSMO	PCM/ST	SCIPCM	PCM/G94
anions					
F^-	-105	-14.16	-17.80	8.21	-14.31
Cl^-	-75	-4.34	-10.03	9.89	-13.77
OH^-	-104	0.26	-2.15	17.32	-15.36
SH^-	-74	-0.62	-4.87	14.19	-13.24
NH_2^-	-93	0.35	-4.50	16.21	-19.47
PH_2^-	-65	-0.18	-4.02	10.68	-18.78
CH_3O^-	-93	13.72	8.40	24.92	-9.23
CH_3S^-	-77	4.76	1.15	20.33	-7.81
MeCOO^-	-75	1.06	-1.51	12.68	-13.92
$\text{C}_6\text{H}_5\text{O}^-$	-70	4.65	-4.00	19.69	-10.19
average unsigned error ^a		3.20	3.83	17.00	13.50
cations					
NH_4^+	-77	-0.01	-0.06	-15.02	-14.29
MeNH_3^+	-68	-1.54	0.59	-16.24	-11.25
Me_2NH_2^+	-61	-2.27	-0.91	-16.90	-14.26
Me_3NH^+	-57	-1.25	-0.49	-15.00	-15.70
$\text{C}_6\text{H}_5\text{NH}_3^+$	-66	0.53	12.17	-15.46	NC ^b
PH_4^+	-71	0.28	1.17	-12.31	-12.60
MePH_3^+	-64	0.08	1.07	-13.87	-12.41
Me_2PH_2^+	-55	-3.61	-2.35	-16.45	-15.04
Me_3PH^+	-50	-4.23	-2.43	-17.50	-16.54
H_3O^+	-102	5.90	6.71	-1.95	6.00
MeOH_2^+	-83	4.95	7.78	-6.03	1.21
Me_2OH^+	-68	1.82	3.33	-10.83	-6.42
H_3S^+	-85	0.97	2.55	-5.95	-2.79
MeSH_2^+	-72	0.8	2.18	-8.23	-6.39
Me_2SH^+	-59	-3.18	-1.05	-15.67	-13.08
average unsigned error		2.09	2.99	12.49	10.57

^a Not including results for F^- and Cl^- .

^b Calculation was not converged.

4. Conclusions

(1) By using an automated minimization procedure we found a set of atomic radii for H, C, N, O, F, P, S and Cl for generation of molecular-shape cavities in the dielectric continuum calculations for molecules in aqueous solution. These radii can be used with the PCM or GCOSMO model at different levels of theory, particularly, HF, MP2 and DFT methods, though other correlated molecular orbital theories can also be employed. When non-electrostatic, particularly dispersion, repulsion and cavitation solute–solvent interactions are included, these atomic radii yield hydration free energies with the accuracy of about 1 kcal/mol for neutral molecules, 2 kcal/mol for cations and 2–5 kcal/mol for anions.

The optimized atomic radii presented here allow studies of the hydration effects on structure and reactivity of biochemical systems, mechanisms of reactions in solutions by using either the PCM or GCOSMO quantum mechanical dielectric continuum model at an improved quantitative level of accuracy.

(2) The SCIPCM solvation model with a unique level of isodensity used for generation of the cavity is not accurate for calculations of hydration free energies for charged molecules. More testing and further development are required for this approach.

Acknowledgements

This work was supported by the University of Utah and by the National Science Foundation through a Young Investigator Award to TNT.

References

- [1] J. Tomasi and M. Persico, *Chem. Rev.* 94 (1994) 2027.
- [2] M.L. Connolly, *J. Appl. Cryst.* 16 (1983) 548.
- [3] B. Lee and F.M. Richards, *J. Mol. Biol.* 55 (1971) 379.
- [4] F.M. Richards, *Ann. Rev. Biophys. Bioeng.* 6 (1977) 151.
- [5] T. Furuki, A. Umeda, M. Sakurai, Y. Inoue, R. Chûjô and K. Harata, *J. Comp. Chem.* 15 (1994) 90.
- [6] A.A. Rashin and K. Nambodiri, *J. Phys. Chem.* 91 (1987) 6003.
- [7] S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.* 55 (1981) 117.
- [8] M. Bachs, F.J. Luque and M. Orozco, *J. Comput. Chem.* 15 (1994) 446.
- [9] M. Orozco and F.J. Luque, *Chem. Phys.* 182 (1994) 237.
- [10] M.A. Aguilar and F.J. Olivares del Valle, *Chem. Phys.* 129 (1989) 439.
- [11] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Petersson, J.A. Montgomery, B. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. DeFrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez and J.A. Pople, *GAUSSIAN 94, Version B.1* (Gaussian, Pittsburgh, 1995).
- [12] T.N. Truong and E.V. Stefanovich, *J. Phys. Chem.* (1995), in press.
- [13] E.V. Stefanovich and T.N. Truong, submitted for publication.
- [14] T.N. Truong and E.V. Stefanovich, *Chem. Phys. Letters* 240 (1995) 253.
- [15] T.N. Truong and E.V. Stefanovich, *J. Chem. Phys.* 103 (1995), in press.
- [16] D.J. Tannor, B. Marten, R. Murphy, R.A. Friesner, D. Sitkoff, A. Nicholls, M. Ringnalda, W.A. Goddard and B. Honig, *J. Am. Chem. Soc.* 116 (1994) 11875.
- [17] J.L. Chen, L. Noodleman, D.A. Case and D. Bashford, *J. Phys. Chem.* 98 (1994) 11059.
- [18] A.A. Rashin, M.A. Bukatin, J. Andzelm and A.T. Hagler, *Biophys. Chem.* 51 (1994) 375.
- [19] A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. II* (1993) 799.
- [20] F.J. Olivares del Valle, M.A. Aguilar and S. Tolosa, *J. Mol. Struct. THEOCHEM* 98 (1993) 223.
- [21] F.J. Olivares del Valle and M.A. Aguilar, *J. Mol. Struct. THEOCHEM* 99 (1993) 25.
- [22] F.J. Olivares del Valle and M.A. Aguilar, *J. Comput. Chem.* 13 (1992) 115.
- [23] F.J. Luque, M. Bachs and M. Orozco, *J. Comp. Chem.* 15 (1994) 847.
- [24] D. Sitkoff, K.A. Sharp and B. Honig, *J. Phys. Chem.* 98 (1994) 1978.
- [25] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.W. Wong, J.B. Foresman, M.A. Robb, M. Head-Gordon, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. DeFrees, J. Baker, J.J.P. Stewart and J.A. Pople, *GAUSSIAN 92/DFT, Version G.3* (Gaussian, Pittsburgh, 1993).
- [26] J.L. Pascual-Ahuir, E. Silla and I. Tuñon, *J. Comput. Chem.* 15 (1994) 1127.
- [27] F.M. Floris, J. Tomasi and J.L.P. Ahuir, *J. Comput. Chem.* 12 (1991) 784.
- [28] W.L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* 110 (1988) 1657.

- [29] M.J. Huron and P. Claverie, *J. Phys. Chem.* 76 (1972) 2123.
- [30] C.M. Breneman and K.B. Wiberg, *J. Comput. Chem.* 11 (1990) 361.
- [31] S. Cabani, P. Gianni, V. Mollica and L. Lepori, *J. Solution Chem.* 10 (1981) 563.
- [32] R.G. Pearson, *J. Am. Chem. Soc.* 108 (1986) 6109.
- [33] M.E. Colvin, E.M. Evleth and Y. Akacem, *J. Am. Chem. Soc.* 117 (1995) 4357.
- [34] F.J. Olivares del Valle and J. Tomasi, *Chem. Phys.* 150 (1991) 139.
- [35] E.M. Evleth, A.Y. and M. Colvin, *Chem. Phys. Letters* 227 (1994) 412.