Outlying Charge, Stability, Efficiency, and Algorithmic Enhancements in the Quantum-Mechanical Solvation Method, COSab-GAMESS

by **Laura N. Gregerson** and **Kim K. Baldridge***[1])

San Diego Supercomputer Center, La Jolla, CA 92093-0505 and
Department of Chemistry, University of California-San Diego, La Jolla, CA 92093-0358, USA

This and the companion paper are dedicated to Professor *Duilio Arigoni* on the occasion of his 75th birthday. His steadfast work focused on the structure and biosynthesis of natural products, and mechanisms and stereochemistry of related biochemical processes, foreshadows important insights in our own investigations presented here.

In this work, we present algorithmic modifications and extensions to our quantum-mechanical approach for the inclusion of solvent effects by means of molecule-shaped cavities. The theory of conductor-like screening, modified and extended for quantum-mechanical techniques, serves as the basis for our solvation methodology. The modified method is being referred to as COSab-GAMESS and is available within the GAMESS package. Our previous work has emphasized the implementation of this model by way of a distributed multipole approach for handling the effects of outlying charge. The method has been enabled within the framework of open- and closed-shell RHF and MP2. In the present work, we present *a*) a second method to handle outlying charge effects, *b*) algorithmic extensions to open- and closed-shell density-functional theory, second-derivative analysis, and reaction-path following, and *c*) enhancements to improve performance, convergence, and predictability. The method is now surtable for large molecular systems. New features of the enhanced continuum model are highlighted by means of a set of neutral and charged species. Computations on a series of structures with roughly the same molecular shape and volume provides an evaluation of cavitation effects.

Introduction. – In the development and refinement of continuum-solvation models (CSMs), highly accurate continuum-dielectric models are a desirable endpoint [1–6]. Noncoupled iterative solute-specific shell approaches, including those of higher-order quantum-mechanical methods, such as coupled cluster [7][8], *Møller–Plesset* [9–11], and multiconfigurational theories [12–14], are now applied in this pursuit. Nonetheless, it is instructive to focus on some fine details of continuum-model implementations relevant for accurate predictions of chemical and biochemical systems, as well as in the efficiency and extensibility of CSMs. Particularly pertinent to our 'COSab' model, are details that include methods to incorporate outlying-charge effects (which are mostly lost in the evaluation of E_{diel}) [15], nonelectrostatic components of the total solvation energy (*e.g.*, cavitation), and self-consistency and convergence properties. As well, we have now implemented the ability to carry out open- and closed-shell density-functional-theory (DFT) computations, second-derivative analysis, and reaction-path-following studies, which provide substantial upgrades to our model capabilities.

[1]) Current address: Institute of Organic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich.

While, in most CSMs, a molecular-shaped cavity is constructed around the solute of interest and the volume outside of that cavity is defined as a dielectric continuum that represents the solvent, it is important to keep in mind that, when used in conjunction with wave-function-based *ab initio* methods, the cavity, constructed from solute coordinates only, is typically outside the loop of the quantum-mechanical self-consistent-field-iteration procedure. The consequence of this is that, depending on the solute of interest and the level of the electronic-structure theory considered, some electron density can reside outside of the cavity (*Fig. 1*). This phenomenon, which we denote 'outlying-charge error' (OCE) [15], has been addressed in the literature, and several strategies have been proposed to handle this loss of energy [10][16–21]. We as well as others [22–27] have shown specific cases where OCE can be quite large, particularly for solute systems with rather diffuse electron density (*e.g.*, for anions, OCE can be up to 25% of the solvation energy). Indeed, coupling of CSMs with quantum-chemical calculations reveals that proper treatment of OCE in the computation of electrostatic solvation can provide higher accuracy and better understanding of the reaction-field approach.

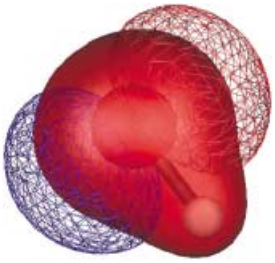

Fig. 1. *Exaggerated view illustrating wave-function 'leakage' outside a molecular-shaped cavity, which is shown here as abruptly 'cutting off' the molecular orbital*

Although original approaches to account for OCE involved correction of the screening charges through the use of scaling factors, it has been pointed out that such strategies are not rigorous: even though they may, in fact, correct the global error in the screening charges, they leave local errors as well as error in the potential [15]. In the context of the present model, two prominent methodologies, the distributed-multipole (DM) approach [10][16], and the double-cavity (DC) approach [15], have been proposed to account for local correction of the screening charges and of the potential. With the DM method, OCE is included as part of the description of the solute through construction of an atom-centered distributed multipolar representation of the electron density (up to hexadecapoles). The method has been shown to be readily implemented and highly accurate [10][11][16].

The DC scheme [15] involves construction of a second molecule-shaped cavity after a normal COSab calculation, one that lies an appropriate distance further out from the original one to enclose entirely the full electron density of the solute. In the context of this method, there are two possibilities for implementation to evaluate OCE. The scheme that is currently implemented creates the secondary cavity as a final stage in the analysis to obtain 'corrected' charges, which include the perturbation from the part of the electronic wave function that extends beyond the inner cavity. These charges are

then projected back onto the primary cavity, and the corresponding 'corrected' potential is calculated. In principal, one could forego the projection, and, instead, simply correct the potential with the charges in place on the secondary cavity, which also leads to an OCE correction. We feel that the former method is the more-efficient and the more-stable of the two, and enables one to adequately account for outlying charge.

Both DM and DC schemes provide very accurate predictions of solvation phenomenon. A systematic comparison in terms of practicality, accuracy, and extensibility illustrates the relative merits of the two methods, both in the present work as well as the companion application paper. On the basis of a similar comparison, several algorithmic enhancements to our combined methodology show clear benefits for the prediction of general solvation phenomenon in reaction processes.

In the accurate treatment of solvation phenomenon for large-molecule systems, issues of efficiency, nonelectrostatic effects, and extensibility to other models are key to methodological advancement. We have introduced several developments to address such issues and enable a clear path forward in our development approach.

Theory. – Our approach to solvent modeling is based on the original concepts of screening in conductors, COSMO, presented by *Klamt* [28]. We have developed this basic approach more expansively to solvation modeling for quantum chemistry in several key areas [16], and have established that it *a*) is highly accurate, *b*) has the potential for extension to important large-scale molecular-modeling techniques, and *c*) provides the sound basis needed to carry out fundamental investigations of full reaction processes (*e.g.*, see companion paper). We have presented extensions of the method to include dynamic correlation *via Møller–Plesset* theory [10] as well as DFT [29] and extensions to include second-derivative analysis for computation of vibrational analysis in solution [29], both of which are necessary for many chemical and biochemical investigations, an example of which is provided in the companion paper. As the basic methodology has been outlined in detail elsewhere, the following entails only those details pertinent to the present discussion.

Important concepts of solvation phenomena lie in the details of the more-subtle features of CSM approaches, providing a better understanding of how to not only maximize accuracy and predictability for complex molecular environments, but also to better understand how to further enhance the method and even create novel hybrid techniques for application to particularly large molecular systems. Key to any of these lies in the description of charge distribution and the construction of the molecular cavity, which acts as the boundary interface between the solute and the solvent. In our approach, which is an approximate but very accurate noniterative approach formulated in terms of a *Greens* function [28], one is allowed to express the screening-charge distribution σ as a linear function of the electronic-charge distribution $\rho(r)$. Thus, the solvation energy, represented as a quadratic expression with respect to $\rho(r)$, can be included in the Hamiltonian of the solute, analogous to that of the Coulombic interactions, as soon as the corresponding dielectric operator (*i.e.*, the *Greens* function of the dielectric cavity) is evaluated. As a consequence, this method not only eliminates the iterative procedure, but also allows calculation of analytical gradients (hence, geometry optimization) without shape constraints.

The solute molecule is enclosed by a cavity consisting of m surface segments with corresponding potential, $\boldsymbol{\Phi} = (\Phi_1,\ldots,\Phi_m)$. The potential, $\boldsymbol{\Phi}$, is directly related to the charge distribution, Q, of the solute, *i.e.*, of the positions and charges of the nuclei, and the electron density. This potential evokes screening charges, $\boldsymbol{q} = (q_1,\ldots,q_m)$ that arise from the polarization of the continuum. The matrix **A** is defined to be the Coulombic-interaction matrix of these screening charges. In principle, $\boldsymbol{\Phi}$ can be evaluated directly and exactly from the charge distribution by direct integration within the quantum-mechanical procedure. Such direct integration of the potential may cause some problems with a small part of the electron density that extends outside the cavity [15], since the cavity is constructed based on the coordinates of the solute alone and not based on the electrostatics. As mentioned above, when the OCE is not corrected, serious discrepancies result in predictions of the solvent-environment effects OCE. In this work, we specifically consider two effective methods for inclusion of OCE into our solvation model, each of which not only presents interesting perspectives into continuum models, but, when compared and contrasted, enable a deeper understanding of the continuum-model approach.

DM Method. In our initial continuum-model developments, we found it advantageous to represent Q by a set of k multipoles $\boldsymbol{M}(Q)$, in addition to evaluating the charge distribution by direct integration. Thus, $\boldsymbol{\Phi}$ is approximated by the potential, $\boldsymbol{\Phi}' = \mathbf{B}\boldsymbol{M}(Q)$, which arises on the m segments from the k multipoles, where the $(k \times m)$ matrix **B** corresponds to the Coulombic-interaction matrix of the multipoles with the segments. In the present implementation, a distributed multipole analysis is used [30–32] that takes into account spherical multipoles up to hexadecapoles for each atom. As such, although the direct integration produces a solvation energy that is subject to some OCE, the screening charges themselves are calculated from the multipole potential, which is insensitive to OCE. Note that this total energy includes the back-polarization of the solute by the solvent. In the end, the self-consistent-field (SCF) energy includes the cost of back-polarization, and the corresponding energy gain, which is about twice as large as the back-polarization cost, is subsumed in the dielectric energy. The resulting method is very accurate and cost efficient, and has the advantage of having limited empirical constraints.

DC Method. A second approach for including effects of outlying charge, which has been the focus of our current efforts, is the DC method originally proposed by *Klamt* [15]. The strategy behind this method is to effectively encompass the charge that lies outside the primary molecular shaped cavity with a secondary molecular shaped cavity. In general, the primary cavity has been shown to be optimally constructed as a *Van der Waals* surface plus 20% [33]. Ideally, one would like to find the optimal expanse of a secondary cavity such that its surface is at a radius that encloses essentially all of the solute electron density. Since this is obviously a function of the molecular system, the initial stages involve creating an outer cavity that represents a *ca.* 85% expansion of the inner cavity, (*i.e.*, $1.85 \times$ primary cavity radii), which has been asserted to be the optimal value [15].

The segmentation of the outer cavity corresponds directly to that of the inner cavity, such that there is a 1:1 correspondence between outer- and inner-cavity charge points, q. The outer cavity-segment charges, $\boldsymbol{q}'$, are calculated *via* $\boldsymbol{A}'\boldsymbol{q}' = -\Phi'$, where $\boldsymbol{A}'$ is the interaction matrix involving the segment points on the outer cavity, and Φ' represents

the potential for interaction of the secondary cavity with the solute and the screening charges from the inner cavity. Since the segmentation points on both surfaces correspond, the final screening charge is represented as a correction, $\boldsymbol{q}'' = \boldsymbol{q} + \boldsymbol{q}'$, with associated potential $\boldsymbol{\Phi}'' = -\boldsymbol{A}\boldsymbol{q}''$. The final corrected energy is, as before, half of the interaction energy needed for the creation of the screening charges, or $\mathbf{E}'' = \mathbf{0.5}\boldsymbol{q}''\boldsymbol{\Phi}''$.

Theoretical Methods. – All calculations have been carried out with the GAMESS software package [34], with a variety of levels of theory employed for comparative purposes. The wave-function-based methods considered include *Hartree–Fock* (HF) [35][36] and second-order *Møller–Plesset* perturbation theory (MP2) [37][38]. For an illustration of calculations involving density-functional theory (DFT) [39–41], see our companion paper. For discussion of the effect of basis set on solvated structure and energetic properties, a variety of basis sets have been employed, including, 3-21G(d) [42][43], 6-31G(nd,mp) [44–46], 6-311(nd,mp) [47], DZV(2d,p) [48][49], TZV(2d,p) [50][51], cc-pvdz [52], and cc-pvtz [53].

The solvation calculations were performed with our new version of a continuum model for *ab initio* methods, adapted to the GAMESS source, termed the COSab solvation method. The method has been implemented at the restricted and unrestricted HF (RHF and UHF), MP2, and DFT levels of theory. The details of the method and its implementation have been described previously for HF [16] and MP2 [10][11] and will not be repeated here. In this study, HF-COSab calculations on gas-phase and COSab-optimized geometries, as well as MP2-COSab calculations on gas-phase geometries, by both the DM and DC approaches are performed.

A dielectric permittivity value ε of 78.4 was employed for H_2O at room temperature for all calculations. The parameters of the cavity construction are 1082 points for the basic grid, 92 segments on a complete sphere, and a solvent radius of 1.3 Å. Atomic radii were taken from the previous parameterization [33][54].

Extensions of methodology from details reported previously include finalizations for treatment of open-shell molecules at the correlated level, a more-accurate surface-construction routine [55], algorithmic modifications for better efficiency and stability, and the new DC scheme to account for OCE. The COSab calculations shown here do not account for any nonelectrostatic interactions, as it is our feeling that it is more important to address the accuracy of the core method before one includes effects that tend to be more empirically derived.

Results and Discussion. – A flowchart for COSab implementation within the GAMESS software is illustrated in *Fig. 2*. We highlight several alternatives of strategy within this diagram that provide reference points for comparison and general discussion of the *ab initio* based solvation theories presented here. The first of these strategy choices involves the method for handling the OCE within the SCF procedure.

There are two OCE components that should be considered to capture its full effect in any rigorous methodology. The first component involves correction of the screening charges on the cavity, and the second component involves correction of the electrostatic potential Φ arising from the solute-solvent interactions. The two strategies that we have formulated within COSab address these two components in quite different but equally satisfying ways. Reference to *Fig. 2* helps to illustrate these differences.

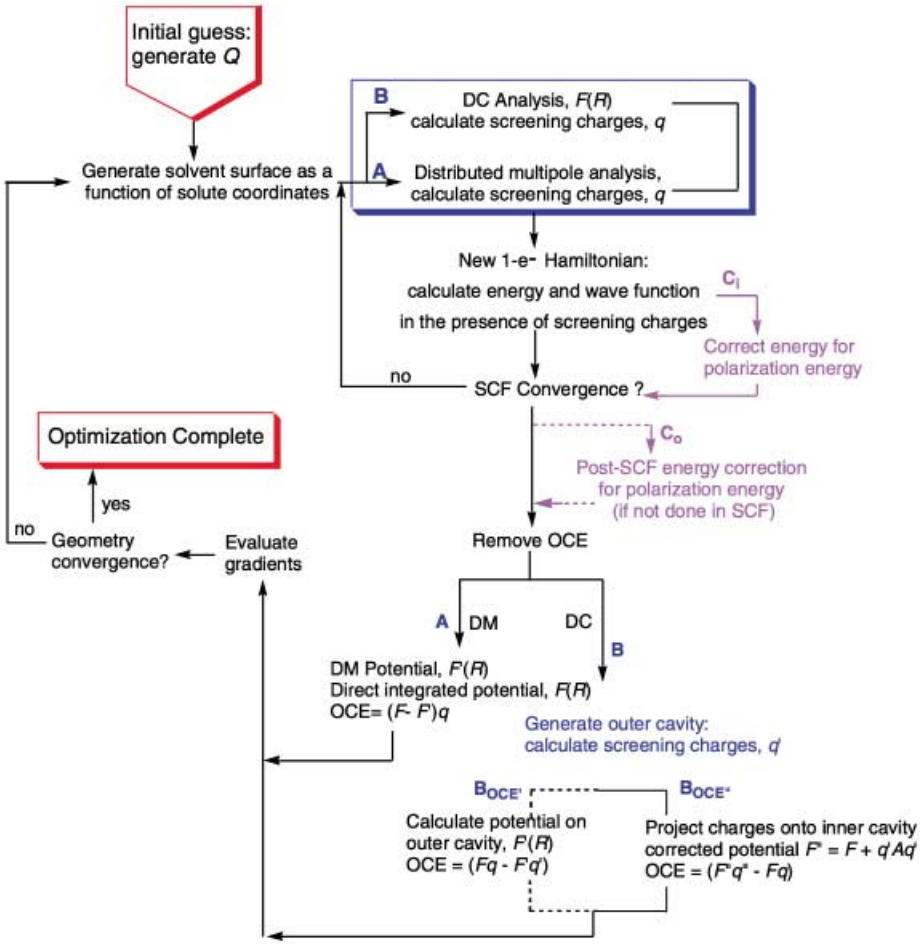

Fig. 2. *Flow diagram for COSab method within the* Hartree–Fock-*SCF procedure*

DM Method. In the case of the DM strategy, path A, screening charges are calculated from the distributed multipole representation of the solute charges and are, therefore, the 'corrected' charges, as they include the full extent of the wave function, albeit from a point-charge formulation of that wave function. The potential is handled in two ways. The 'corrected' potential is approximated by multiplication of the multipolar representation of the solute with the coulombic interaction matrix of the multipoles with the cavity segments. This represents the corrected potential since it includes the full extent of the wave function. The second potential calculated is the result of the direct integration *via* the SCF solution of the HF equations, which now includes not only the solute coordinates, but also the 'corrected' screening charges. In the end, the difference between the directly integrated potential and the 'corrected' potential represents the error associated with charge lying outside the molecular cavity.

DC Method. In the DC approach, path B, both corrected screening charges as well as corrected potential, are handled much later in the calculation procedure – post-SCF, in fact. Initially, the screening charges are calculated from the directly integrated potential of the SCF procedure, and, thereby, are missing the component of charge that resides outside of the cavity surface. Once the SCF has converged, a second cavity is constructed, and the corresponding charges are calculated, as described above.

As discussed above, of the possible strategies for calculation of OCEs specific to the DC method ($B_{OCE'}$ and $B_{OCE''}$, in *Fig. 2*), we have chosen to implement the projection method. As depicted in *Fig. 2*, the final corrected charges $B_{OCE''}$ are calculated *via* projection of secondary 'perturbation' charges onto the primary cavity, and the corresponding corrected potential then calculated. The second cavity correction is performed only at the final stages of the evaluation of the solvation effects. In principle, one could enable an iterative procedure to optimize the OCE correction, or implement a multiple-cavity approach that not only enables one to optimize the ability to capture any outlying effects, but also offers the opportunity to include such things as variable permittivity effects.

In this work, we focus on three aspects of methodology comparison for these strategies for *ab initio* solvation phenomenon: *a*) the general behavior of methodology for prediction of accurate solvation effects, *b*) stability, computational cost, and efficiency issues, and *c*) potential for algorithmic extensions.

General Behavior of COSab-DC and COSab-DM. As laid out in the discussions above, the 'corrected' potential by either the DM and DC methods is approximate, but different, depending on the method used. In the DM method, the 'corrected' potential is approximated by means of a distributed multipole representation of the solute, which is not a fully 'exact' representation of the wave function, for obvious reasons. In the DC method, the positioning of the second cavity beyond the initial cavity is essentially a 'parameter', and so the resulting 'corrected' potential can only be an approximation. In either case, however, the 'corrected' potential is more accurate for the solution-phase computation than is the exactly integrated potential, since it includes the influence of any charge lying outside the cavity. We have illustrated, in fact, in previous studies [10][15][16] as well as in this work, that the use of DM or DC methodology completely eliminates the OCE, is basis-set insensitive, and accurate to within experimental error for the electrostatic component of the solvation energy.

By construction, CSM methods retain some nature of approximation, and, therefore, instabilities can arise for a number of purely algorithmic reasons. Examples of numerical instabilities may be as dramatic as the inability to converge either a single-point solvation energy for a specific geometry or to an optimized solution-phase geometry and corresponding energy, to a more-minor but disconcerting fluctuation in any of these energies (*e.g.*, fluctuations in the SCF procedure reflected in the fourth or fifth decimal place). Typically, such effects will be more pronounced for some molecules than others. Thus, although a computation may successfully converge towards the correct *ab initio* solvation energy, it becomes important to carefully inspect the result, as such instabilities can be indicative of more-fundamental problems, as we will discuss.

In general, numerical instability is inherent to any model that is not completely variationally embedded within the SCF procedure. In other words, in most continuum

electrostatic methods, there is a branching back and forth between specialized solvation routines and the SCF iterative procedure. For example, with each new geometry, one must reconstruct the molecular cavity and associated cavity charges for the next SCF cycle. As such, the overall behavior of the model is influenced by such things as the specifics of the cavity construction on each update and the representation of charge on the cavity surface, and how the energetic cost for the creation of cavity segment charges is handled within the algorithm. Below, we address the representation of charge together with the cavity construction and update procedure and the influence on the SCF procedure, followed by a discussion of energy-cost issues.

Cavity Construction and Surface-Charge Generation. Cavity construction, in general, is based on the generation of a basic grid from an iterative refinement of triangles beginning with a regular icosahedron (*Fig. 3*). This is followed by a series of projections onto the final solvent surface, such that the cavity size is essentially *Van der Waals* radii +20% [33]. The basic points on the surface are then gathered into segments. The detailed nature of the cavity-construction algorithm is an inherent part of the quality of the solvation model. In particular, for a variety of interesting molecular constructs, different regions of the molecular cavity can have different sensitivity to surface discretization, requiring special considerations of segment partitioning and size, specific inclusions of effects of segment–segment charge-contact areas embedded in crevices of molecular surfaces, and/or properly treating cavities having very large segments near very small segments. A specific example involves the ability to include effects of major charged regions, which would have remain 'hidden' or embedded in crevices by standard routines, that relies only on *Van der Waals* or solvent-accessible surface construction (*Figs. 3* and *4*,*a*). Specialized segmentation procedures become necessary at the junctions of any steep surfaces in the molecular configuration, so that their interaction energies are realistically assessed.

Fig. 3. *Schematic of molecular-shaped cavity-surface representation*

Additionally, a notoriously difficult situation for many CSMs is the handling of T-shape molecules (*Fig. 4*,*b*), where the construction of any solvent surface is inhibited the slight contact of interacting spheres. By some surface routines, the problem is solved by constructing two separate molecular surfaces, whose interaction with the solvent has to be synchronized at each step of the iterative procedure. This can cause numerical instability and inaccurate assessment of solvation phenomena.

Similar situations can be found when dealing with transition states and other structures that occur along a reaction path, where one potentially encounters situations of unusually long bonds and unusual angles. Since the ability to accommodate change in the cavity that occurs in response to the molecular changes through the progression along the reaction path is crucial to the realistic assessment of solvation phenomenon for the reaction process, as in our companion paper, the enhanced cavity-construction algorithm is very important for realistic solvation modeling in general.

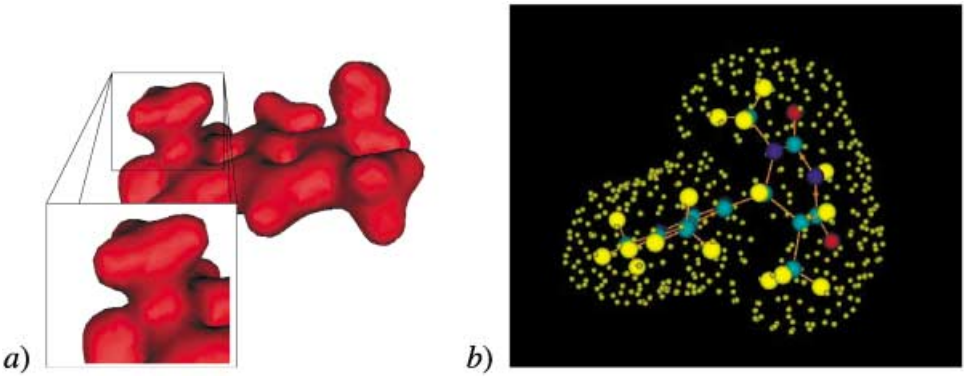

Fig. 4. *Depictions of solvent surfaces:* a) *simple molecular surface of irofulven* (HMAF; see companion paper) *and* b) *ball-and-stick with translucent surface of a T-shaped molecule* ((6*R*)-1-methyl-6-(1,5-dimethylpyrimidone)-5,6-dihydrothymine)

Many of these issues can specifically be addressed through careful cavity-construction algorithms. In fact, the current cavity-construction algorithm within our COSab model is attributed to one of our collaborators *Klamt* [55], who has continually made enhancements to the routine, targeting problems as they have arisen during studies of various molecular constructs. Algorithmically, such problems arise partly from calculation of an interaction matrix of elements for the interaction of very small segments associated with the crowded areas (*e.g.*, crevices) with the more normally sized segments, by means of the center approximation for the latter segments. If a very small segment is very close to a much larger sphere segment, this sometimes causes a numerical instability in the *Cholesky* factorization, which is associated with the charge calculations. Additionally, when one tries to construct the segmentation for molecular features that have perpendicular junctions, care has to be taken to avoid spheres that slightly intersect at the tight junction. If not properly accounted for, then the resulting charges that are calculated and fed back into the SCF procedure do not properly represent these molecular features, and, therefore, the resulting solvation energies will be missing these key interactions.

A basic assumption in the representation of charge on any particular segment of the cavity surface is that there is a homogeneous distribution on that segment – in other words, that a single charge can represent that segment. If the segment size is taken too large, this assumption naturally breaks down. Since the set of surface charges for the entire molecular cavity is generated only once per geometry optimization, when the assumption of homogenous electron density on a segment is not fully realized, the energies in the SCF cycle can show small observable fluctuations.

Table 1 shows the variation of electrostatic solvation energy with finer and finer cavity discretization for several small molecular constructs, for both the DM and DC implementations. The value of discretization represents the number of segments per atom and controls the size of the segments on the cavity surface. The mean number of basic points per segment is set to a value that corresponds to the number of segments (NSPA) on a complete sphere. The basic points are constructed to achieve the most-compact segments. The algorithm described for defining a relatively small number of

Table 1. *Effects of Discretization on RHF/6-31G(2d,p) Solvation Energies* [kcal/mol] [a]) *for CH_4, H_2O, and HF*

Discretization[b])	CH_4		HF		H_2O		C_2H_2	
	ΔG_{DC}	ΔG_{DM}	ΔG_{DC}	ΔG_{DM}	ΔG_{DC}	ΔG_{DM}	ΔG_{DC}	ΔG_{DM}
12	−0.37	−0.39	−5.61 (−6.3)	−7.00 (−8.0)	−7.72	−9.61	–	–
32	−0.41	−0.39	−5.91 (−6.6)	−7.14 (−8.1)	−8.09	−9.44	−4.3	−5.2
42	−0.40	−0.38	−5.94 (−6.7)	−7.15 (−8.2)	−8.12	−9.43	−4.4	−5.2
92	−0.39	−0.35	−6.10 (−6.8)	−6.60 (−7.5)	−8.13	−8.81	−4.5	−4.9
162	−0.38	−0.35	−6.08 (−6.8)	−6.44 (−7.3)	−8.14	−8.64	−4.5	−4.8
362	−0.36	−0.33	−6.10 (−6.8)	−6.21 (−7.0)	−8.10	−8.31	−4.5	−4.6
482	−0.36	−0.33	−6.10 (−6.9)	−6.16 (−7.0)	−8.10	−8.22	−4.5	−4.5

[a]) Numbers in parenthesis refer to DZV + (2d,p) computations performed additionally for HF. [b]) Number of segments per atom. Calculations are geometry optimizations performed at RHF/6-31G(2d,p).

segments composed of sets of basic points enables the calculation of the segment interactions by summation over the interactions of the corresponding basic points. The resulting increase in accuracy allows achievement of the continuum limit with a moderate total number of segments.

In these rather simple molecular systems, we see the importance of using an adequate discretization to obtain the desired chemical accuracy with such computations. While in a relatively nonpolar molecular system such as CH_4, there is very little effect of discretization on predicted solvation energy, for the more-polarized systems, the effect can be more dramatic, with errors over 1 kcal/mol. The effect is slightly more dramatic when more-sensitive basis-set descriptions are used, as shown for HF. For more-complicated molecular systems, such as the structures in the companion paper, the effect can be much larger. For example, in those compounds, at very low values of discretization (*e.g.*, NSPA = 12), one can find up to 50% errors in energetics (as compared to NSPA = 92), 10–15 kcal/mol of which is due to OCE. Additionally, often it is the case that the calculation will not converge at all for low values of discretization (*e.g.*, acetylene in *Table 1*). Nonetheless, we see a quick leveling off such that the highest levels of discretization are often unnecessary.

Energetic Cost for the Creation of Cavity-Segment Charges. For both DC and DM methods, the cost for creating the surface charges, $1/2*\Phi'*q_{DM}$ and $1/2*\Phi*q_{DC}$, has been, until now, included outside the SCF (*Fig. 2*, path C_o). In the DM implementation, the correction is evaluated from the 'corrected' potential and not the directly integrated potential generated within the SCF cycle. In the DC method, the correction is evaluated from the directly integrated potential, but does not yet include the effects of any outlying charge. As a result, it is prudent to investigate the behavior of the SCF convergence so ensure that iterative procedure exhibits variable behavior. Even though the evaluation of one and two electron integrals and subsequent formation of FOCK matrix does include the effect of the solvent, it is not completely balanced with the cost for the creation of the charges immediately within the SCF, but done outside the SCF, as is the correction due to outlying charge.

We have evaluated the behavior of the SCF with the cost for creating surface charges performed both inside and outside the SCF cycle (*Fig. 2*, path C_i). We have carried this out for both the DC and DM methods and have found that there can be

slight fluctuation in the SCF energetics depending on the placement of the correction factor, thereby interfering with our ultimate goal of chemical accuracy in CSM methods. We illustrate the extent of this fluctuation for two molecular systems, OH^-, which has been shown to have large OCE effects, and irofulvene (HMAF) an example from the companion paper.

In a small diffuse system such as the OH^- ion, the magnitude of the fluctuation even for particularly coarse segment size is well below 0.005 kcal/mol regardless of the placement of the cavity charge cost as illustrated in *Fig. 5*. The fluctuation is completely eliminated with proper segmentation size and basis set functionality, and placement of the cost function inside the SCF.

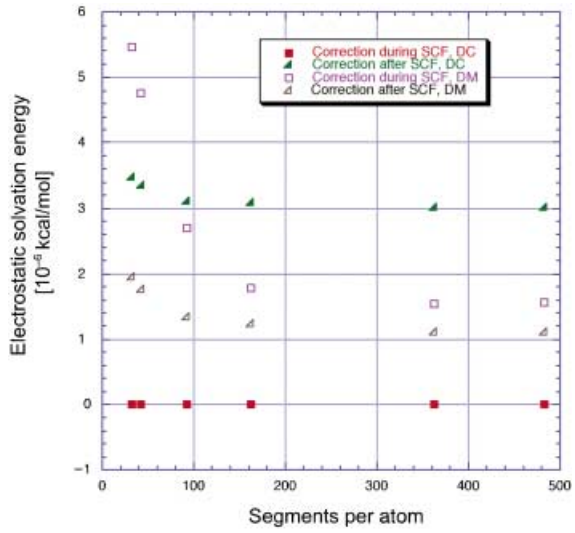

Fig. 5. *Variation in SCF fluctuation with DC and DM methods as a function of segmentation size for OH^-*

Perhaps a more-relevant example has been taken from one of the structures (HMAF) in the companion paper. For this illustration, we performed a solvent calculation with a lower-level basis set (RHF/6-31G(d)), and a typical segmentation size (92), with the cavity-charge-creation cost both inside and outside the SCF. In this case, due to the more-complicated nature of the cavity structure, and the limited flexibility of the basis set, one finds much more-dramatic SCF fluctuations depending on where the correction is placed. With the correction placed outside the SCF, the fluctuation in energetics is observed to vary over 100 kcal/mol before settling down to the final prediction. When the correction is placed inside the SCF iterative procedure, the fluctuation is completely eliminated by implementation of the DC method and nearly completely eliminated (*e.g.*, $<$sub 0.005 kcal/mol) with the DM method.

In the DM method, an additional characteristic of the model that causes this tiny remaining fluctuation noted in the above analysis is the result of the charges in the DM model being generated from the 'corrected' potential, which was computed by means of the distributed multipole representation of the quantum-mechanical field. The

'corrected' charges generated that, therefore, include the effects of the charge lying outside the cavity are then not only used to calculate $1/2*\Phi*q_{DC}$, but are also fed back into the quantum-mechanical SCF procedure for generation of the field (which inherently does not include the outlying charge). Depending on how large a difference there is between Φ and Φ', fluctuations within the SCF cycle could actually occur. Such a phenomenon is a direct result of underestimating of the solvent back-polarization in the direct integration, and, in fact, the more outlying the charge associated with a molecule, the more noticeable such a phenomenon would be. Thus, proper segmentation is an important component, again relating back to the discussions above. However, again with careful consideration of computational set-up, these effects are minimized.

While such arguments may seem somewhat esoteric since, with proper implementation and care in the selection of such factors as cavity segmentation and level of theory, the methods as presented provide very accurate predictions of electrostatic solvation, our goal is to investigate all sources of error before considering non-electrostatic effects. It is our experience that several of these errors can be extensive in some continuum solvation methods and, in certain cases, computations are not feasible. Additionally, there is a concern that the parameterization schemes used to include nonelectrostatic effects in some model implementations become contaminated by contributions from errors made due to factors such as OCE, inconsistencies in method, and other sources. Our preference is to understand nonelectrostatic effects more from a quantitative view when possible. We begin to address such effects below.

Computational Costs and Efficiency. We have analyzed the two OCE methods in terms of computational expense both with respect to the corresponding gas-phase computations, as well as with respect to each other. Basically, the DM and DC strategies have nearly identical timing, with the exception of one component of the former method that we have found can be rather expensive in the current implementation. This component has to do with the distributed multipole computations that are performed for each geometry cycle. This is a portion of the code that has yet to be optimized either sequentially or with parallel strategies. We find that the additional analysis within the DM routine can take from 10–50% longer in CPU time compared to gas-phase calculations depending on molecular construction. *Fig. 6* illustrates the timing data for a complete calculation by the DM method in a case where the additional computations are on the higher end of the cost range, along with the specific timing for only the distributed multipole component of the computation. One can see that the majority of the computation is in the evaluation of the distributed multipoles. We have recently discovered a strategy for parallelization of this portion of the code that will bring this cost down considerably.

Given that the DC and DM behave almost identically with the exception noted above, we show two additional comparisons only with the former method. *Fig. 6* shows timing results for a larger set of molecules of variant functionality. Over the set of 20 molecules, we find an average increase in CPU time of *ca.* 32% for the solution-phase computations, with values as low as 5%. In general, as we perform such computations for research studies of real molecules where the number of atoms per molecule is increased over our test set, that we see a lower and lower overhead cost for the solution computation, depending on the complexity.

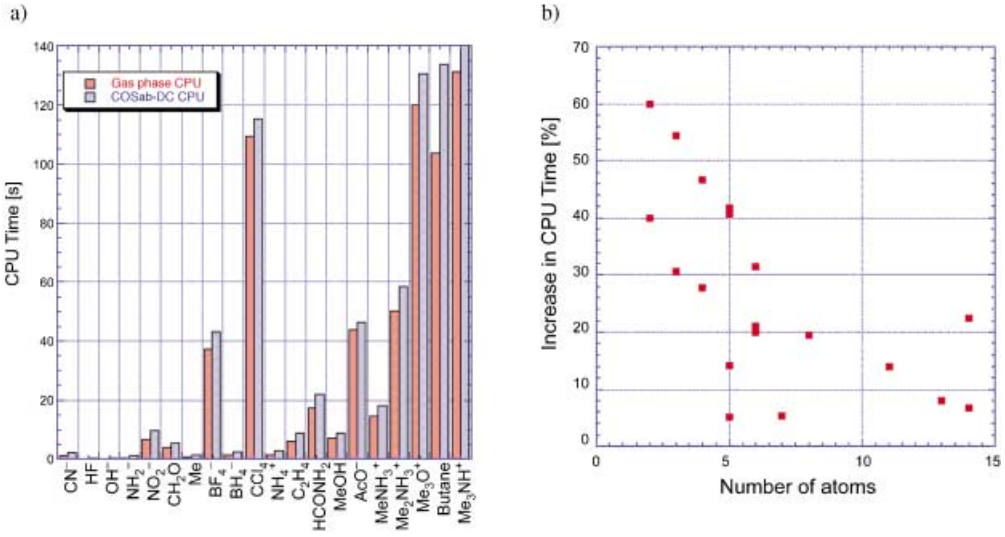

Fig. 6. *Gas-phase* vs. *solution-phase timing results for a set of 20 molecules. a*) CPU Time [s] for gas- and solution-phase computations. *b*) Percent increase in CPU time for solution-phase compared to gas-phase computation *vs.* number of atoms.

This latter point is demonstrated in *Fig. 6*, *b* for the set of 20 molecules, where we see a rather dramatic decline in percentage CPU change from gas phase to solution phase in the progression from atoms and diatomics to larger molecules. In particular, a specific example from our companion paper on the chemical-reaction processes for the illudin-based molecules, required, on average, at most 6% more time for the solution-phase computations, and, in several systems, the computations were even somewhat (*e.g.*, 10%) faster, since the density actually converged faster (*i.e.*, after fewer iterative cycles) in those cases. All of these computations were performed on a *Pentium4* 2.6 GHz processor. In general, due to the localizing electric field created by the surface charges, the convergence of the solution-phase wave function should actually take no more, and, as observed, can take less, time than the respective gas-phase calculation, depending on the molecular system and the structure from which the calculation is initiated.

Finally, we note some additional timings from the new DFT- and MP2 implementations of COSab in GAMESS in *Table 2*. The direct overhead for solvation can be gleaned from the first two entries showing gas-phase and DC results, and as already pointed out, this is fairly small and a larger component of the time than are larger molecular systems. Beyond these small percentages, the remaining increase in timing in the next three entries give one a feel for the costs associated with DFT- and MP2-type correlative treatments, with MP2 being the most costly, as is well known, and our grid-free DFT implementation being more costly than the grid-based DFT.

Nonelectrostatic Considerations. In the forgoing solvation methodology analysis, we have considered only the electrostatic component of the solvation energy. The total solute–solvent system energy (E_{tot} in *Eqn. 1*) is expressed as a sum of a component due

Table 2. *Timing Data for HDFT/TZV(2d,p)//RHF/TZV(2d,p) and MP2/TZV(2d,p)//RHF/TZV(2d,p) Solvation-Energy Calculations* (running on *Pentium4* 2.6GHz *Dell*)

	H_2O		AcOH		Octanol	
	ΔG_{elect} [kcal/mol]	Total CPU time [s]	ΔG_{elect} [kcal/mol]	Total CPU time [min]	ΔG_{elect} [kcal/mol]	Total CPU time [min]
Gas phase	–	1.4	–	2.4	–	51.0
DBLCAV[a])	– 8.4	2.3	– 8.9	2.7	– 6.5	53.6
DFT-GF[a])[b])	– 7.8	16.5	– 7.3	22.7	– 5.6	426.6
DFT-G[a])[c])	– 7.7	8.8	– 7.3	5.2	– 5.5	94.5
MP2[a])	– 6.7	3.7	– 8.2	20.7	– 6.3	833.0

[a]) DBLCAV OCE Correction method used. [b]) Grid-free DFT in GAMESS. [c]) Grid-based DFT in GAMESS.

to electrostatics, which includes back-polarization of the solvent onto the solute, as well as a component that includes dispersive and cavitation energies, which are typically addressed together as the nonelectrostatic energy [56].

$$E_{tot} = E_{gp} + E_{sol} \tag{1}$$

$$E_{sol} = (E_{diel} + E_{back\text{-}pol}) + E_{nelect} \tag{2}$$

The creation of the cavity in a medium costs energy, *i.e.*, is destabilizing, while the dispersion interactions between the solute and solution contribute to stabilization. Such short-range effects, which include costs associated with, *e.g.*, H-bonding and preferential orientation of a solvent near an ion, are characteristically concentrated in the first solvation shell. These effects are typically not captured in any continuum model approach to solvation. Cavitation necessitates the inclusion of entropy factors and a specific accounting for the loss of solvent–solute *Van der Waals* interactions. Dispersion effects require the accounting of the new *Van der Waals* interactions established between the solute and solvent.

Many continuum-model approaches account for cavitation and dispersion by way of parameteric fits to experimental data [2][57][58]. Both factors are proportional to surface area and can be fit *via Eqn. 3:*

$$\Lambda G_{cavity} + \Lambda G_{expt} = \sum_{atoms,i} \xi_i S_i \tag{3}$$

Until now, we have chosen to concentrate on first understanding many of the details of the basic CSM method associated with calculation of the electrostatic component, such as those presented in this work, before moving on to the inclusion of nonelectrostatic effects so as not to mask inherent model errors. Additionally, it turns out that through careful consideration of the fine points of the model itself, one can derive appropriate ways to include features such as nonelectrostatic effects, the appropriate solvent radii, and predictions made for nonaqueous media. A case in point is the careful consideration of solvent surfaces made by *Klamt et al.* in the COSMO-RS model

[33][59]. Given this, it is important to establish some perspective on this issue with respect to our current model. As such, we have put together an array of structures (*Fig. 7*) that fall into roughly the same molecular shape and volume categories, providing a consistent set from which to evaluate cavitation effects, under the constraints of the current model implementation.

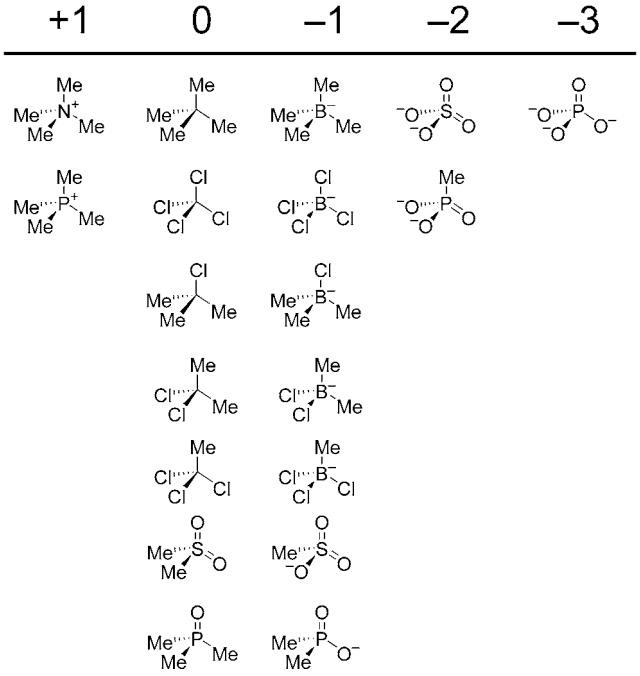

Fig. 7. *Molecule systems that have similar shapes and volume properties over a range of ionic states*

We have optimized all structures in the gas phase and solution phase at the RHF/6-31G(2d,p) level of theory, and performed subsequent single point MP2/6-31G(2d,p) computations to check for effects of dynamic correlation. Through preliminary investigation we find that the use of more-substantial basis sets does not dramatically change the results. *Table 3* shows the results for all molecules in *Fig. 7*, including the computed single-point volume, RHF-COSab/6-31G(2d,p)//RHF/6-31G(2d,p), optimized volume, RHF-COSab/6-31G(2d,p)//RHF-COSab/6-31G(2d,p), and solvation energetics at the MP2-COSab/6-31G(2d,op)//RHF-COSab/6-31G(2d,p). The MP2 single-point energies were performed with our newly implemented MP2-COSab methodology, which is now available for both open- and closed-shell molecules.

At the outset, it is useful to look only at the neutral species in *Table 3*. In this set, we can eliminate the effect of cavitation based on surface area (*e.g.*, *Van der Waals* interactions) and volume effects since these are essentially constant, and the effect of cavitation based on any ionic contribution. Of these, then, when one compares the two zero-dipole compounds, $C(Me)_4$ and CCl_4, the difference in solvation energy is very small, -0.7 kcal/mol in the former to -1.4 kcal/mol in the latter, *i.e.*, only 0.7 kcal/mol.

Table 3. *RHF/6-31G(2d,p)//RHF/6-31G(2d,p) Molecular-Solvent-Volume Extents, Electrostatic Energies* [kcal/mol], *and Molecular Dipoles* [D] *for Molecules in* Fig. 7

Molecule	Charge	Single-point volume	Optimized volume	Gas-phase dipole	Dipole in H_2O	MP2 Dipole in H_2O	Solvation energy	MP2[a]) Solvation energy
CMe_4	0	102.9	102.7	0	0	0	−0.8	−0.7
CCl_4	0	118.2	118.2	0	0	0	−1.4	−1.4
$CMeCl_3$	0	114.4	114.4	2.0	2.6	2.5	−3.5	−3.6
CMe_2Cl_2	0	110.5	110.9	2.6	3.5	3.2	−4.9	−4.8
CMe_3Cl	0	107.6	108.0	2.6	3.8	3.4	−5.0	−4.6
PMe_3O	0	103.7	103.7	4.4	6.1	5.6	−16.1	−14.0
SMe_2O_2	0	91.3	91.5	4.7	6.6	5.8	−17.4	−14.3
PMe_4	1	115.2	115.1	0	0	0	−53.2	−52.7
NMe_4	1	101.3	100.9	0	0	0	−55.2	−54.5
BCl_4	−1	123.7	123.2	0	0	0	−53.5	−53.3
BMe_4	−1	108.1	108.1	0	0	0	−53.8	−53.1
$BMeCl_3$	−1	119.8	119.9	1.3	2.1	1.9	−56.6	−56.4
BMe_3Cl	−1	112.5	112.8	0.41	1.6	10.1	−57.4	−75.6
BMe_2Cl_2	−1	116.7	117.0	1.7	3.0	2.8	−59.3	−59.0
$MESO_3$	−1	81.1	80.8	3.5	5.6	5.0	−77.3	−74.6
PMe_2O_2	−1	93.1	93.0	4.1	6.7	6.1	−81.8	−79.2
SO_4	−2	70.1	70.0	0	0	0	−252.4	−250.6
$PMeO_3$	−2	82.9	82.7	2.7	5.5	5.2	−261.1	−259.0
PO_4	−3	73.8	72.6	0	0	0	−561.9	−563.8

[a]) Single-point solvation energies also computed at the MP2/6-31G(2d,p)//RHF/6-31G(2d,p) level.

This is an indication of the polarizability effect of the additional Cl environment, and of the ability of the model to make such an assessment. Taking a look at the full set of data for neutral compounds in *Fig. 8,a*, where we now have included also the much larger dipole contribution, the plot of dipole *vs.* electrostatic solvation is linear with a $r^2 = 0.88$, which shows the ability of the model to respond to the multipolar (dipole) features of the molecules.

One might expect, in a comparison of tetramethylborate with tetrachloroborate, that there would be a similar if not larger electrostatic component from the octapole-moment change than was observed for the neutral case. However, the data actually shows that the two compounds have almost identical solvation energies, −53.1 and −53.3 kcal/mol respectively. A key difference here is that tetrachloroborate has a larger volume than tetramethylborate, because the B−Cl bond is so much longer (1.87 Å) than the respective B−C bond (1.66 Å). The overall larger ionic radius implies a lower solvation-energy contribution from the ion. In this case, the volume-change penalty nearly perfectly balances the benefit from the higher multipole contributions.

Fig. 8,b shows that a plot of dipole moment *vs.* solvation energy of the B−Cl compounds is linear with a high correlation factor ($r^2 = 0.99$). This would indicate that the dominant factor causing variance in the *modeled* solvation of these compounds is the multipole-moment character. This may appear to be in contrast to the volume arguments we just made, but the volume changes here are simply too small and are compensated by yet higher polar moments.

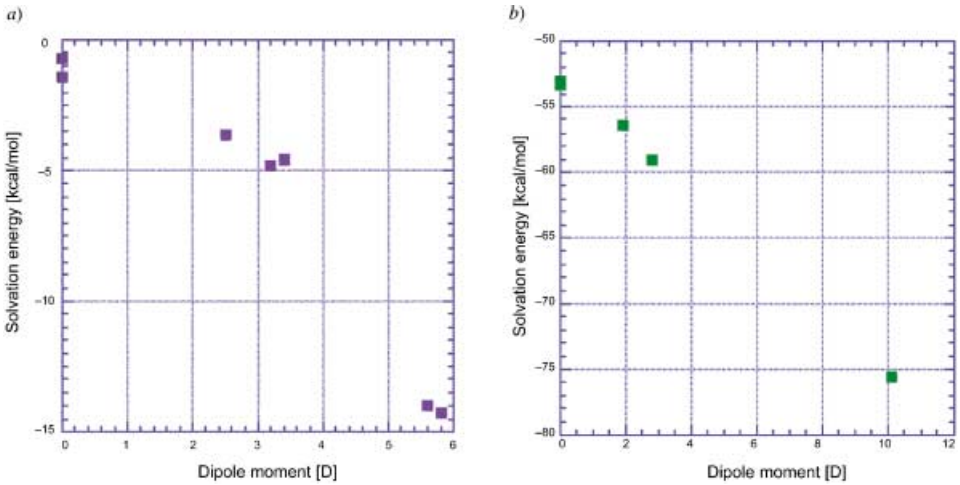

Fig. 8. *Plots of electrostatic solvation energy* vs. *dipole moment for* a) *neutral molecules and* b) *boron compounds* (from *Fig. 7* and *Table 3*)

In fact, all ions of unit charge should actually exhibit constant solvation energy per volume, when the condition that *only* charge interactions are important is met. This can be tested across the four compounds, NMe_4, PMe_4, and BMe_4, and, additionally, the species NH_4, BH_4, BF_4, and K to establish a more-adequate regression analysis. In this extracted set of data, although symmetry restricts the magnitude of the higher multipole moments in these compounds (octapole/hexadecapole moments only), volumes across this series changes considerably, over the range of 30–115 Å^3. A plot of volume *vs.* electrostatic solvation energy for these chemically unrelated compounds is monotonic and quasi-linear, as shown in *Fig. 9*. In our original analysis, we included BCl_4, which one might expect to fit in this linear relationship, but does not quite (*e.g.*, with such a small test set the change in r^2 to 0.96 is relatively dramatic with the addition). This will be further addressed below.

From this regression analysis (*Fig. 9,a*), a hypothetical value for a unit-charged ion of volume 70 Å^3 (the volume of SO_4^{2-}) can be calculated as −66.6 kcal/mol. As the solvation energy responds to the square of the charge, one would then predict a value of −266.6 kcal/mol, for the solvation energy of the −2 ion based *solely* on a regression of the unit-charged ions. Indeed, this value agrees well in this context with the independently calculated solvation energy of SO_4^{2-} (252.4 *vs.* 266.6, 5% overestimation), and further highlights the dominance of electrostatic interactions in our solvent model.

That BCl_4 is an apparent outlier in the above analysis (*Fig. 9,b*) would imply that something besides volume affects the linear relationship. Such nonvolume-related effects could be geometric in nature, or due to differences in the multipolar natures of the compounds. A look at the geometries across the set reveals no significant differences. If one analyzes the distributed multipole data from the compounds NMe_4, PMe_4, BMe_4, and BCl_4, however, one finds some interesting features.

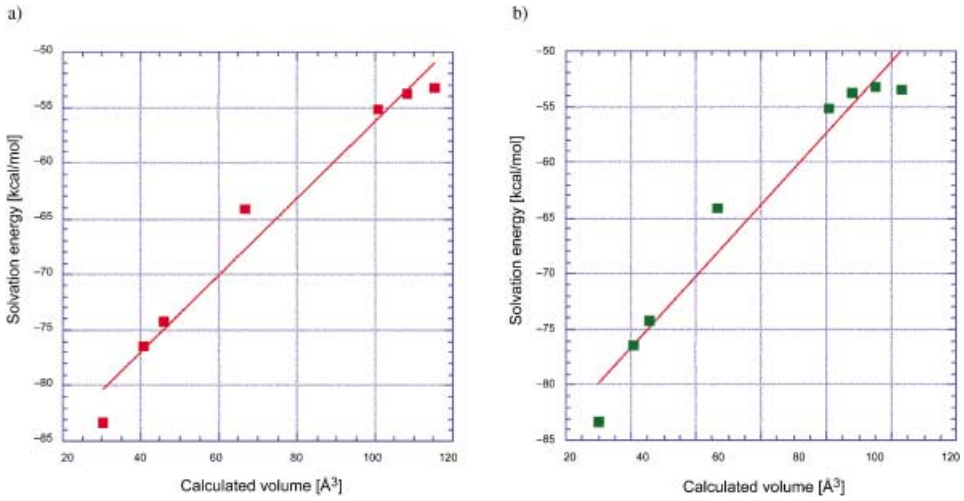

Fig. 9. *Regression analysis of volume* vs. *electrostatic solvation energy for a subset of the species in* Fig. 7. *a*) BCl_4 excluded; *b*) BCl_4 included.

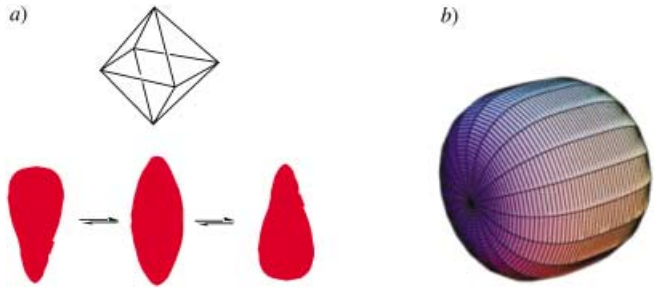

Fig. 10. *Schematic representations of an octapole and a hexadecapole*

As expected for tetrahedral systems, the first multipole found in these tetrahedral compounds is an octapole. Compared to a quadrupole, an octapole has more degrees of freedom and is, therefore, more difficult to model with theory (*Fig. 10*).

According to the definitions of spherical harmonic multipoles, there are six components of the octapole. In the tetrahedral compounds analyzed, only one of these components (along the z axis, $Q_{32s} = f^*xyz$) is nonzero. For the N- and P-analogues, this component is very, very small (for NMe_4, 2.0 and PMe_4, 1.2). With the addition of the B, as in BMe_4, the component is much larger in magnitude, 47.0. This would imply a much larger electrostatic distribution along the lines of *Fig. 10*. When the ligands are changed to Cl, in BCl_4, the magnitude of this component again diminishes, but only to 22.7.

A further look at the hexadecapoles completes the analysis. The hexadecapole moment is even more difficult to visualize, but can be related to not only the z extent, but also a more-cylindrical extent, with nonzero components related to $Q_{40} = f(z)$ and $Q_{44c} = f(x,y)$ as illustrated in *Fig. 10, b*. While the hexadecapole components of NMe_4,

PMe_4, and BMe_4 are all nearly identical, and large (60–90), the analogous components of BCl_4 are *ca.* fourfold smaller, which would indicate a much tighter organization of charge, likely due to the difference between B and Cl in electronegativity. Also, the small size of the B-atom enables the Cl-atoms at the vertices to concentrate most of the charge in a smaller space. As such, it would appear that BCl_4 is, in fact, a bit different in its multipolar makeup, enough to create the offset from linearity observed in the above analysis.

Including Nonelectrostatic Effects in the Model. In reality, electric fields on the molecular surfaces of polar solutes are typically so strong that the majority of the solvent polarizability (reorientation of static dipoles) is almost at saturation. As such, there is no longer the linear behavior associated with the macroscopic limit. A novel extension to our present solvation model, as first presented by *Klamt* [59], would actually avoid the questionable dielectric approach, providing a modification for nonlinear behavior. In such an extension, deviations from ideal screening that occur within any solvent would be described as pairwise misfit interactions of the ideal screening charges on contact points of the molecules in the fluid. Such a description represents an overall generalization of CSMs, as it no longer depends on experimental data or any parameterization of the solvent. On the basis of screening energies, surface areas, and screening surface-charge densities, the intermolecular interactions in a liquid system can, therefore, be described, providing a much more-efficient means than the traditional description *via* electrostatic and *Van der Waals* interactions, since the explicit position of all molecules in space has not to be taken into account, and fewer parameters are required for the model.

Conclusions and Future Work. – Detailed chemical treatments of molecular and electronic structure with the inclusion of the effects of the environment now enable one to predict molecular geometry, follow the reaction paths of chemical transformations, predict electrostatic effects in a variety of environments, estimate pK_a shifts, redox potentials, and provide interpretations of spectroscopic probes of molecular environments with an accuracy not offered by other more-simplistic modeling techniques.

In the methodology presented, solvent effects are included directly in the *Fock* matrix such that self-consistency is maintained with respect to the solvent charges and the interaction potential. In the present extensions and evaluations, we show comparison of and implications for capturing the outlying charge effects with two different strategies, both of which prove to be solid methods for capturing this error, but have different advantages and disadvantages for providing a generalized methodology. Additionally, we present a variety of significant modifications for overall improvement in performance and stability in these models, from improved numerical procedures (elimination of computations involving inverse matrices, positioning of mathematical manipulations within the SCF, and clever strategies for dynamic memory allocation and parallel computing efficiency), to more-advanced cavity construction and surface-charge-evaluation techniques, a majority of which are very important for our applications that involve transition states and reaction paths, dynamics applications, and weak complexes. We feel that such a detailed analysis has opened avenues to extend the methodology to larger molecular constructs by means of rigorous hybrid methodologies, variable permittivity strategies, and, most importantly, to a fully

variable methodology, the latter of which should, in principle, eliminate many of the problems associated with small variations and parameterizations associated with the incorporation of nonelectrostatic components.

Investigations of biochemical systems by such treatments have proven to offer an auspicious interface between experiment and theory, an example of which is shown in the companion paper with the interpretation of the mechanistic information that is important to the understanding of the selective toxicity of such compounds towards certain forms of malignant tumors, and, thereby, significantly enhancing drug design studies. The ability to study the potential-energy surface (PES) around specific reaction sites in more detail is critical. In the context of the continuum model already presented, we have incorporated the modifications necessary to follow the solute geometry through the progression along the reaction path as it is accompanied by a change in the shape of the cavity, which contributes to the energy of the system. We have recently also incorporated the ability to carry out vibrational analysis in solution, which enables us to fully establish the character of the stationary points on such a reaction profile. Future work will involve the extensions towards prediction of rate constants in solution.

This work was mostly funded by the *NIH*-sponsored resource, NBCR-RR08605. *L. G.* has also been supported by the *GAAN Foundation* for the promotion of science graduate students for research (from 2000–2003). Additional support for computation has been provided by the resource allocation program of SDSC, and the SDSC-ROCKS cluster team. *K. K. B.* would also like to thank Prof. *J. S. Siegel*, and Dr. *A. Klamt* for helpful discussions pertaining to some of the related work herein.

REFERENCES

[1] C. J. Cramer, D. G. Truhlar, *J. Am. Chem. Soc.* **1991**, *113*, 8552.
[2] J. Tomasi, M. Persico, *Chem. Rev.* **1994**, *94*, 2027.
[3] C. J. Cramer, D. G. Truhlar, 'Continuum Solvation Models: Classical and Quantum Mechanical Implementation', Vol. 6, VCH Publishers, New York, 1995.
[4] C. J. Cramer, D. G. Truhlar, *Chem. Rev.* **1999**, *99*, 2161.
[5] M. G. Cossi, G. Scalmani, N. Rega, V. Barone, *J. Chem. Phys.* **2002**, *117*, 43.
[6] D. M. Chipman, M. Dupuis, *Theor. Chem. Acta* **2002**, *107*, 90.
[7] O. Christiansen, K. V. Mikkelsen, *J. Chem. Phys.* **1999**, *110*, 8348.
[8] O. Christiansen, K. V. Mikkelsen, *J. Chem. Phys.* **1999**, *110*, 1365.
[9] C. B. Nielsen, K. V. Mikkelsen, S. P. A. Sauer, *J. Chem. Phys.* **2001**, *114*, 7753.
[10] K. K. Baldridge, V. Jonas, *J. Chem. Phys.* **2000**, *113*, 7511.
[11] K. K. Baldridge, V. Jonas, A. Bain, *J. Chem. Phys.* **2000**, *113*, 7519.
[12] R. Cammi, L. Frediani, B. Mennucci, J. Tomasi, K. Ruud, K. V. Mikkelsen, *J. Chem. Phys.* **2002**, *117*, 13.
[13] T. D. Poulsen, P. R. Ogilby, K. V. Mikkelsen, *J. Chem. Phys.* **2001**, *115*, 7843.
[14] T. D. Poulsen, P. R. Ogilby, K. V. Mikkelsen, *J. Chem. Phys.* **2002**, *116*, 3730.
[15] A. Klamt, V. Jonas, *J. Chem. Phys.* **1996**, *105*, 9972.
[16] K. K. Baldridge, A. Klamt, *J. Chem. Phys.* **1997**, *106*, 6622.
[17] E. Cances, B. Mennucci, *J. Chem. Phys.* **2001**, *115*, 6130.
[18] M. Cossi, B. Mennucci, J. Pitarch, J. Tomasi, *J. Comp. Chem.* **1998**, *19*, 833.
[19] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Chem. Phys.* **2001**, *114*, 5691.
[20] C. C. Pye, T. Ziegler, *Theor. Chem. Acc.* **1999**, *101*, 396.
[21] D. M. Chipman, *J. Chem. Phys.* **2000**, *112*, 5558.
[22] C.-G. Zhan, J. Bentley, D. M. Chipman, *J. Chem. Phys.* **1998**, *108*, 177.
[23] C.-G. Zhan, D. M. Chipman, *J. Chem. Phys.* **1998**, *109*, 10543.
[24] C.-G. Zhan, D. M. Chipman, *J. Chem. Phys.* **1999**, *110*, 1611.
[25] D. M. Chipman, *J. Chem. Phys.* **1996**, *104*, 3276.
[26] D. M. Chipman, *J. Chem. Phys.* **1997**, *106*, 10194.

[27] B. Mennucci, E. Cances, J. Tomasi, *J. Phys. Chem. B.* **1997**, *101*, 10506.
[28] A. Klamt, G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.
[29] L. N. Gregerson, K. K. Baldridge, unpublished results.
[30] C. W. Murray, D. J. Tozer, C. van Heusden, CADPAC, Issue 5.0 Cambridge, 1995.
[31] A. J. Stone, *Chem. Phys. Lett.* **1981**, *83*, 233.
[32] A. J. Stone, M. Alderton, *Mol. Phys.* **1985**, *56*, 1047.
[33] A. Klamt, V. Jonas, T. Bürger, C. W. Lohrenz, *J. Phys. Chem.* **1997**, *102*, 5074.
[34] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, S. T. Elbert, *J. Comp. Chem.* **1993**, *14*, 1347.
[35] J. A. Pople, R. K. Nesbet, *J. Chem. Phys.* **1954**, *22*, 571.
[36] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69.
[37] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.
[38] J. A. Pople, J. S. Binkley, R. Seeger, *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1.
[39] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864.
[40] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.
[41] R. G. Parr, W. Yang, 'Density Functional Theory of Atoms and Molecules', Oxford University Press, New York, 1989.
[42] J. S. Binkley, J. A. Pople, W. J. Hehre, *J. Am. Chem. Soc.* **1980**, *102*, 939.
[43] M. S. Gordon, J. S. Binkley, J. A. Pople, W. Pietro, W. J. Hehre, *J. Am. Chem. Soc.* **1982**, *104*, 2797.
[44] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.
[45] W. J. Hehre, W. A. Lathan, *J. Chem. Phys.* **1972**, *56*, 5255.
[46] M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265.
[47] R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 650.
[48] T. H. Dunning Jr., P. J. Hay, 'Modern Theoretical Chemistry', Plenum Press, New York, 1976.
[49] T. H. Dunning Jr., P. J. Hay, 'Methods of Electronic Structure Theory', Plenum Press, New York, 1977.
[50] A. D. McClean, G. S. Chandler, *J. Chem. Phys.* **1980**, *72*, 5639.
[51] T. H. Dunning Jr., *J. Chem. Phys.* **1971**, *55*, 716.
[52] D. E. Woon, T. H. Dunning Jr., *J. Chem. Phys.* **1993**, *98*, 1358.
[53] R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.
[54] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441.
[55] A. Klamt, *COSMOlogie* GmbH & Co., Leverkusen, personal communication.
[56] C. Reichardt, 'Solvents and Solvent Effects in Organic Chemistry', VCH, Weinheim, 1988.
[57] C. Colominas, F. J. Luque, J. Teixido, M. Orozco, *Chem. Phys.* **1999**, *240*, 253.
[58] M. Zacharias, *J. Phys. Chem.* **2003**, *107*, 3000.
[59] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224.

Received August 31, 2003