COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient

A. Klamt* and G. Schüürmann†
Bayer AG, Q18, D-5090 Leverkusen-Bayerwerk, Germany

Starting from the screening in conductors, an algorithm for the accurate calculation of dielectric screening effects in solvents is presented, which leads to rather simple explicit expressions for the screening energy and its analytic gradient with respect to the solute coordinates. Thus geometry optimization of a solute within a realistic dielectric continuum model becomes practicable for the first time. The algorithm is suited for molecular mechanics as well as for any molecular orbital algorithm. The implementation into MOPAC and some example applications are reported.

Owing to the outstanding importance of solvation effects in chemistry the proper handling of solvation in molecular mechanics as well as in molecular orbital (MO) calculations is a field of active research. Even a brief discussion of the several very different approaches to this problem would be beyond the scope of this paper and we therefore refer the interested reader to some reviews[1,2,3] and literature cited therein.

The new approach presented in this article belongs to the class of dielectric continuum models.[4–20] In these models the solute molecule is embedded in a dielectric continuum of permittivity ε. Thus the solute forms a cavity within the dielectric. The cavity surface, *i.e.* the interface between the cavity and the dielectric is usually called the 'solvent accessible surface' (SAS). It is well known from basic electrostatics[21] that the response of a homogeneous dielectric continuum to any charge distribution of the solute consists of a surface charge distribution on the interface, *i.e.* the SAS, arising from the polarization of the dielectric medium. The only, but non-trivial, problem consists in the calculation of the screening charge densities $\sigma(\boldsymbol{r})$ which are implicitly given by eqn. (1) where $\boldsymbol{n}(\boldsymbol{r})$ is the surface normal vector at a point $\boldsymbol{r}$ and $\boldsymbol{E}^-(\boldsymbol{r})$ denotes

$$4\pi\varepsilon\,\sigma(\boldsymbol{r}) = (\varepsilon - 1)\,\boldsymbol{n}(\boldsymbol{r})\,\boldsymbol{E}^-(\boldsymbol{r}) \qquad (1)$$

the total electric field at the inner side of the surface at this point. The latter consists of contributions arising from the solute charge distribution and from the screening charges. For an arbitrarily shaped surface eqn. (1) cannot be solved by analytical means and different numerical approaches for its solution have been proposed. All of them require the segmentation of the SAS into small segments, and usually a constant charge density σ_s is assumed for each segment S_s. Thus, with m being the number of segments on the SAS, the screening charge distribution can be represented by a m-dimensional vector $\boldsymbol{\sigma}$.

One approach to the continuum model has been elaborated by Miertus, Tomasi and co-workers.[4–8] For a given charge distribution $\rho(\boldsymbol{r})$ within the cavity they calculate the screening charges $\boldsymbol{\sigma}$ iteratively from eqn. (1). This has the advantage of being quite straightforward, but it requires a separate iterative calculation for each charge distribution $\rho(\boldsymbol{r})$, even if the SAS is kept fixed. Thus the calculation of the self-consistent solution, often called self-consistent reaction field (SCRF), within a MO-calculation results in a double iterative procedure within a single SCF-cycle. Obviously this is computationally quite expensive, but owing to the enormous increase in computer power it has become practicable during recent years. Thus implementations in different *ab initio* and semi-empirical codes have been reported. Even more serious is the lack of an analytical expression for the gradient of the screening energy with respect to the solute geometry. The problem of geometry optimization within this approach has been considered by Bonaccorsi, Cammi and Tomasi, recently.[7] Numerical gradients are prohibitively expensive owing to the iterative procedure. In addition they are unstable owing to fluctuations arising from the segmentation of the surface. Therefore geometry optimization has to be done using strategies which do not use gradients, but these can hardly be expected to work properly for systems with many degrees of freedom.

Hoshi and co-workers[13] have developed a Green function solution of eqn. (1) (although not recognized as such) which allows one to express the screening charge distribution σ as a linear function of the charge distribution $\rho(\boldsymbol{r})$. As a consequence, the solvation energy becomes a quadratic expression with respect to $\rho(\boldsymbol{r})$, very much like the Coulomb energy. Thus it can be included in the Hamiltonian of the solute in analogy to the Coulomb interactions, as soon as the corresponding dielectric operator $\mathbf{D}$, *i.e.* the Green function of the dielectric cavity, is evaluated. The evaluation of $\mathbf{D}$ is quite expensive and includes the inversion of a $m \times m$-matrix, where m is the number of segments, but since $\mathbf{D}$ does only depend on geometry, it has to be evaluated only once for a fixed geometry. As soon as $\mathbf{D}$ is known and included in the Hamiltonian, the calculation of the fully self-consistent wavefunction in the presence of the dielectric is a single SCF-cycle. Thus this non-iterative approach should be more efficient than the iterative ones, as long as the evaluation of $\mathbf{D}$ is not much more expensive than a single iterative solution of eqn. (1), but this question has not been discussed in the literature. The approach of Hoshi *et al.* has not been taken up by the scientific community so far. This may be due to the quite complicated formalism, which arises partly from the general treatment of anisotropic dielectric media. In addition, they did not realize, that, at least in principle, the Green function approach opens the way for the calculation of analytical gradients, although within their formalism this task would be quite hard.

For spherical and ellipsoidal cavities eqn. (1) can be solved analytically. This was done first by Onsager.[22] His solution, which is an expansion with respect to the electric multipole moments, is the basis for another class of dielectric continuum approaches,[14–18] which work with spherical or ellipsoidal cavities. In these approaches the first and most crucial step is the fitting of a suitable sphere or ellipsoid for the solute under consideration, while the subsequent calculation of the screening energy or the integration of the corresponding operator into the

† *Present address:* Umweltforschungszentrum Leipzig-Halle, O-7050 Leipzig, Germany.

Hamiltonian is straightforward. But it is necessary to take into account several orders of the multipole expansion, since the convergence in general is very poor.[17] This has not been done in the implementation of the spherical approach into semi-empirical MO-codes[23,24] by Katritzky, Zerner and co-workers,[18-20] where only dipoles are taken into account. Their approach can easily be shown to fail in several cases, for example for symmetric molecules with two dipoles, for which the total dipole moment is zero. Nevertheless, carefully applied, the ellipsoidal approach especially has been shown to give reliable results for a wide range of molecules, as long as their shape is roughly ellipsoidal. Owing to the simplicity of the expression for the screening energy it is no problem to calculate an analytical gradient with respect to the solute coordinates for a fixed cavity. This becomes much more complicated if the cavity is simultaneously adjusted. Recently such an algorithm, which allows for the fast geometry optimization of solutes within the ellipsoidal approximation, has been developed and implemented into different *ab initio* and semi-empirical codes by the group of Rivail.[16] Nevertheless, apart from being very fast, it is questionable whether such a geometry optimization, which is obviously restricted to ellipsoid-like conformations of a molecule, is useful anyway.

In summary, within their range of validity, the different realizations of the dielectric continuum model give comparable results. The values of the screening energy are very sensitive to the size of the cavity used. Therefore differences in the results between the approaches are mainly due to differences in the definition of the cavity. The use of van der Waals surfaces with slightly increased radii (up to 20%) turned out to yield surprisingly good estimates for free energies of solvation.[7,8,17] Therefore the dielectric continuum model is a valuable tool for the handling of solvation effects.

In this article we present an approximate, but very accurate, non-iterative approach for the solution of eqn. (1) for arbitrarily shaped cavities. It is a Green function solution and thus, although being developed independently, the formalism is to some extent similar to the approach of Hoshi *et al.*,[13] but it is less complicated with the consequence that the calculation of analytical gradients, and hence efficient geometry optimization without shape constraints, becomes practicable. Since our approach is based on the screening in conductors, we call it '**C**onductor-like **S**creening **M**odel' or COSMO.

Theory

Dielectric Screening Energy.—It is well known from basic electrostatics[21] that dielectric screening energies for a given geometry scale as $(\varepsilon - 1)/(\varepsilon + x)$ with the dielectric permittivity ε of the screening medium, where x is in the range 0–2 (see Appendix A). Hence screening effects in strong dielectric media such as water, which has a relative dielectric permittivity of ε *ca.* 80, are well approximated by the corresponding screening effects in an infinitely strong dielectric ($\varepsilon = \infty$), *i.e.* in a conductor. Screening in conductors can be handled much more easily than in dielectric media. As an example we first consider the problem of N point charges Q_i at positions $\boldsymbol{r}_i$ within a sphere of radius R. For a conductor this problem can be solved in closed form using the image charge method,[21,25] and the total screening energy of the system reads

$$\Delta E = -\tfrac{1}{2}\boldsymbol{Q}\mathbf{D}\boldsymbol{Q} \tag{2}$$

with

$$D_{ij} = R/(R^4 - 2R^2\,\boldsymbol{r}_i\boldsymbol{r}_j + \boldsymbol{r}_i^2\boldsymbol{r}_j^2)^{\frac{1}{2}} \tag{3}$$

Here, as well as throughout the article, we use electrostatic energy units for the sake of shortness, *i.e.* we have dropped the factor $(4\pi\varepsilon_0)^{-1}$, which would appear in all energy expressions otherwise. Furthermore, we have adopted the vector notation $\boldsymbol{Q}$ for the N charges Q_i. ΔE includes the energy gain of the source charges $\boldsymbol{Q}$ due to the screening as well as the positive energy due to the interaction of the screening charges on the sphere. Thus, using the dielectric operator $\mathbf{D}$, *i.e.* the Green function of the sphere, the total screening energy is a quadratic expression with respect to the charge vector $\boldsymbol{Q}$ in close analogy to the Coulomb interaction of the charges itself, which in an analogous notation reads

$$E_{\text{coulomb}} = -\tfrac{1}{2}\boldsymbol{Q}\mathbf{C}\boldsymbol{Q} \tag{4}$$

with

$$C_{ij} = \|\boldsymbol{r}_i - \boldsymbol{r}_j\|^{-1} \text{ and } C_{ii} = 0 \tag{5}$$

Eqn. (2) is the exact solution of the screening problem only for conductors, but with a weakly ε-dependent correction factor discussed in Appendix A it is also very accurate for most dielectric media of interest. Hence this simple quadratic expression provides an accurate and much easier to handle solution for the dielectric screening energy of a molecule within a spherical cavity than Onsager's multipole expansion,[22] and it is by far preferable to the dipole approach of Katritzky *et al.*[18,19]

Since, in our opinion, the assumption of a spherical cavity is a too severe restriction, we decided to develop a similar Green function formulation for realistic cavity shapes, starting again from the screening in conductors. If the charges Q_i are enclosed by an arbitrary closed surface S the screening charge distribution on S and the screening energy can be found by dividing S into a large number M of small segments S_μ centred at $\boldsymbol{t}_\mu$ with constant surface charge density σ_μ on each segment. Let $|S_\mu|$ denote the area and $q_\mu = |S_\mu|\,\sigma_\mu$ the charge of segment μ. Then

$$b_{i\mu} = \frac{1}{|S_\mu|}\int_{S_\mu} \|\boldsymbol{r} - \boldsymbol{r}_i\|^{-1}\,\mathrm{d}^2\boldsymbol{r} \tag{6}$$

$$\approx \|\boldsymbol{t}_\mu - \boldsymbol{r}_i\|^{-1} \tag{6a}$$

is the electrostatic interaction of a unit charge at $\boldsymbol{r}_j$ with a unit charge on S_μ and

$$a_{\mu\nu} = \frac{1}{|S_\mu|\,|S_\nu|}\int_{S_\mu}\int_{S_\nu} \|\boldsymbol{r} - \boldsymbol{r}'\|^{-1}\,\mathrm{d}^2\boldsymbol{r}'\,\mathrm{d}^2\boldsymbol{r} \tag{7}$$

is the electrostatic interaction of unit charges on S_μ and S_ν. For $\mu \neq \nu$, $a_{\mu\nu}$ can be approximated by

$$a_{\mu\nu} \approx \|\boldsymbol{t}_\mu - \boldsymbol{t}_\nu\|^{-1} \tag{7a}$$

while for a diagonal element $a_{\mu\mu}$, which represents twice the self energy of a unit charge on segment μ, a detailed analysis (see Appendix B) yields

$$a_{\mu\mu} \approx 3.8\,|S_\mu|^{-\frac{1}{2}} \tag{7b}$$

as a good approximation. Using vector notations $\boldsymbol{Q}$ for the N source charges Q_i and $\boldsymbol{q}$ for the M surface charges q_j, the total energy of the system may be written as

$$E(\boldsymbol{q}) = \tfrac{1}{2}\boldsymbol{Q}\mathbf{C}\boldsymbol{Q} + \boldsymbol{Q}\mathbf{B}\boldsymbol{q} + \tfrac{1}{2}\boldsymbol{q}\mathbf{A}\boldsymbol{q} \tag{8}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the matrices formed by $a_{\mu\nu}$ and $b_{i\mu}$, respectively, and $\mathbf{C}$ is the Coulomb matrix [eqn. (5)]. The actual

screening charge distribution q^* minimises this total energy. Hence we have

$$\nabla_q E(q)|_{q^*} = \mathbf{B}Q + \mathbf{A}q^* = 0 \quad (9)$$

and

$$q^* = -\mathbf{A}^{-1}\mathbf{B}Q \quad (10)$$

and the total energy of the screened system becomes

$$E(\sigma^*) = \tfrac{1}{2}Q(\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B})Q \quad (11)$$

and

$$\Delta E = -\tfrac{1}{2}Q\mathbf{B}\mathbf{A}^{-1}\mathbf{B}Q = -\tfrac{1}{2}Q\mathbf{D}Q \quad (12)$$

Thus the expression for the screening energy in an arbitrarily shaped cavity now is analogous to the solution for the sphere, with the Green function or dielectric operator now being $\mathbf{B}\mathbf{A}^{-1}\mathbf{B}$. Eqn. (9) and thereby also eqns. (10), (11) and (12) can also be derived from the boundary condition of vanishing potential on the surface of a conductor. This can be written as $\Phi = 0$, where Φ is the vector of potentials Φ_μ on the segments. Realising that $\Phi = \mathbf{B}Q + \mathbf{A}q$, where $\mathbf{B}Q$ is the potential arising on the segments from the source charges Q, and $\mathbf{A}q$ is the potential arising from the surface charges q, the condition $\Phi = 0$ is just equivalent to eqn. (9).

Up to now we have only considered point charges. A generalization to charge distributions arising from basis functions, which, in general, is required for the use of our method in molecular orbital calculations, is straightforward and given in Appendix C.

Gradients.—The relative simplicity of the explicit results for the total dielectric screening energy ΔE allows for the calculation of the analytic gradient of the dielectric energy with respect to each atomic position R_α:

$$\begin{aligned}\nabla_{R_\alpha}\Delta E &= -\tfrac{1}{2}\nabla_{R_\alpha}(Q^t\mathbf{B}^t\mathbf{A}^{-1}\mathbf{B}Q)\\ &= -Q^t\mathbf{B}^t\mathbf{A}^{-1}(\nabla_{R_\alpha}\mathbf{B})Q + \tfrac{1}{2}Q^t\mathbf{B}^t\mathbf{A}^{-1}(\nabla_{R_\alpha}\mathbf{A})\mathbf{A}^{-1}\mathbf{B}Q \quad (13)\\ &= -q^*(\nabla_{R_\alpha}\mathbf{B})Q + \tfrac{1}{2}q^*(\nabla_{R_\alpha}\mathbf{A})q^*\end{aligned}$$

Physically the first part of this gradient represents the change of Coulomb interaction between the source charges Q and the screening charges q^* due to a change of R_α, while the second part represents the corresponding change in the interaction of the screening charges. As long as each segment S_μ is fixed to a single atom, the gradient of the matrices $\mathbf{B}$ and $\mathbf{A}$ with respect to R_α can easily be calculated as

$$\begin{aligned}\nabla_{R_\alpha} b_{i\mu} &\approx \nabla_{R_\alpha}\|t_\mu - r_i\|^{-1}\\ &\approx \frac{t_\mu - r_i}{\|t_\mu - r_i\|^3}(\tilde{\delta}_{i\alpha} - \delta_{\mu\alpha}) \quad (14)\end{aligned}$$

$$\nabla_{R_\alpha} a_{\mu\mu} \approx 0 \quad (15)$$

and

$$\begin{aligned}\nabla_{R_\alpha} a_{\mu\nu} &\approx \nabla_{R_\alpha}\|t_\mu - t_\nu\|^{-1}\\ &\approx \frac{t_\mu - t_\nu}{\|t_\mu - t_\nu\|^3}(\delta_{\nu\alpha} - \delta_{\mu\alpha}) \quad (16)\end{aligned}$$

where $\tilde{\delta}_{i\alpha} = 1$ or 0 depending on whether the representation point r_i does or does not belong to atom α, respectively, and $\delta_{\mu\alpha} = 1$ if surface segment S_ν is associated with atom α and $\delta_{\mu\alpha} = 0$ otherwise. The generalization of eqn. (14) to the case of charge densities is straightforward and left out here for the sake of shortness. In eqns. (14)–(16) only the direct geometry dependence of the segment positions has been taken into account while the indirect geometry dependence *via* a possible change of the segment areas has been neglected. The handling of the latter contributions would be much harder, but not impossible. Fortunately they turn out to be of minor importance.

So far we have developed the theory for conductors. By the introduction of the ε-dependent correction factor $f(\varepsilon) = (\varepsilon - 1)/(\varepsilon + \frac{1}{2})$ into the expressions for the screening energy and its gradient the theory can be extended to finite values of the dielectric constant with a relative error of less than $\frac{1}{2}\varepsilon^{-1}$ (see Appendix A).

Results

Method of Implementation and Efficiency Aspects

Before going into the details of a special implementation, the general method of use and some efficiency aspects of the COSMO-approach presented in the previous section will be discussed.

The first step, *i.e.* the construction and segmentation of the SAS is common to all 'real'-shape continuum approaches. Usually this is not important for the total computation time.

The second step is the evaluation of all interaction coefficients, *i.e.* the construction of the matrices $\mathbf{A}$ and $\mathbf{B}$. In one way or another this has to be done in any continuum approach, but again it is less critical with respect to computation time.

The first really time-consuming step is the inversion of $M \times M$-matrix $\mathbf{A}$. This step scales as M^3, where M is the number of segments on the SAS. Depending on the desired accuracy, segment numbers M between 50 and 500 turn out to be reasonable for molecules up to 20 atoms, even if some effort is made to keep M as small as possible (see below). Therefore the inversion is the time-determining step in semi-empirical and force-field COSMO-calculations. It should be done with the most efficient inversion routines taking advantage of the positive definiteness of the matrix. Fortunately it has to be done only once for a given geometry.

After the inversion of $\mathbf{A}$ the Green function $\mathbf{D} = \mathbf{B}\mathbf{A}^{-1}\mathbf{B}$ has to be evaluated. Since $\mathbf{B}$ is a $N \times M$-matrix, and since for *ab initio* calculations N is of the order of $n^2/2$, where n is the number of basis functions, the evaluation of $\mathbf{D}$ is an $n^4 \times M^2$-step process, and therefore it becomes highly relevant for the calculation time. The total time effect of a COSMO-implementation into *ab initio* codes can thus be estimated by a comparison with the other n^4-step process in *ab initio* calculations, *i.e.* the calculation of the two-electron integrals. A preliminary and crude estimation yields that both steps should require comparable computation time. Hence we expect a time increase of about a factor 2 or 3 for COSMO *ab initio* calculations. This appears to be rather modest for an accurate representation of solvent effects. In semi-empirical calculations most of the overlap densities are neglected and as a result N is only proportional to the number of orbitals n and the evaluation of $\mathbf{D}$ is not time-relevant.

For force-field calculations the solvation energy can now be calculated directly from the point charges and the matrix $\mathbf{D}$. For MO calculations we have to include $\mathbf{D}$ in the Hamiltonian. Owing to the formal analogy to the Coulomb interactions this is rather straightforward. Having done this the self-consistent wave function and the corresponding fully self-consistent total energy of the system in the presence of the dielectric environment can be calculated in a standard SCF-cycle with the modified Hamiltonian.

At this point in any case the screening charges $q^* = \mathbf{A}^{-1}\mathbf{B}Q$ are available at negligible costs and thus the gradient with

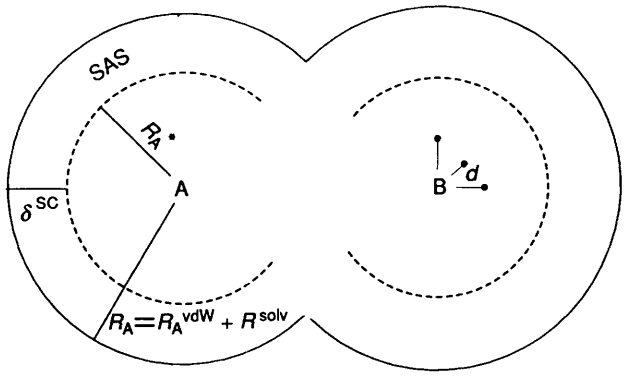

Fig. 1 Construction of the SAS (schematic); Solid circles indicate the surface accessible to the centres of solvent molecules, dashed lines the surface accessible to solvent charges. For atom B the three dipole representation points are indicated in addition.

respect to an atomic position $\boldsymbol{R}_\alpha$ can be calculated from eqn. (13). This does not involve any additional M^3- or n^4-step process. Thus the gradient is not only available, but even at no relevant additional costs, as soon as the screening energy is calculated. This is a unique feature within the non-ellipsoidal continuum approaches. We consider this to be the greatest advantage of COSMO.

Construction of a Suitable Solvent Accessible Surface

Because the number of segments on the surface strongly influences the computational costs, and because we have to ensure that all segment positions are connected to the atom positions in a simple way for an efficient calculation of gradients, we did not use a published segmentation strategy. Instead we have developed a more efficient algorithm for the construction of a suitable set of segments on the solvent accessible surface. We assume that the geometry of the solvent molecules may be described by an effective radius R^{solv}. Thus the centres of the solvent molecules are excluded from a sphere of radius $R_\alpha = R_\alpha^{vdW} + R^{solv}$ around each atom α of the solute with R_α^{vdW} being the van der Waals radius of atom α. The effective charges which are responsible for the dielectric screening will not be located at the centres of solvent molecules. They may be assumed to be a distance δ^{sc} outside the centre, so that their minimum distance to a solute atom α is $R_\alpha^* = R_\alpha - \delta^{sc}$. δ^{sc} will be in the range of 0.5 Å to R^{solv}. The most reliable value for each solvent has to be found empirically. The distances R_α^* define the SAS as indicated in Fig. 1.

In more detail the construction and segmentation of the SAS is done as follows: As an initial preparation a basic grid of points on the unit sphere is generated by iterative refinement of triangles starting from a regular icosahedron. Although developed independently our refinement procedure is to some degree similar to the GEPOL algorithm published recently.[26] Thus a large number of basic points per atom (NPPA) is generated. To a high accuracy, each basic point represents an area of 4π/NPPA on the unit sphere. The following steps are done atomwise, in order to ensure that each segment of the SAS is connected to a single solute atom. As previously mentioned this is essential for the calculation of gradients.

In the first step, the basic grid is projected to the R_α-sphere and those points lying within the R_β-sphere of any other atom β are excluded. Thus only allowed positions for the centres of solvent molecules are left.

In a second step the remainder of these basic points is contracted to the R_α^*-sphere. Now these points represent allowed positions for the screening charges and define the SAS.

In the third step the basic points on the SAS are gathered into segments. By the parameter NSPA (number of segments per atom) the size of the segments is controlled; the mean number of basic points per segment is set to NPPA/NSPA, corresponding to NSPA segments on a complete sphere. The basic points are put together in such a way as to achieve the most compact segments.

Finally, for each segment the segment area is calculated from the number of basic points associated with it, and the corresponding representation vector $\boldsymbol{t}_\nu$ is obtained as the centre of these basic points.

The described algorithm of defining a relatively small number of segments composed of sets of basic points enables us to calculate the segment interactions $a_{\mu\nu}$ by summation over the interactions of the corresponding basic points. Owing to the corresponding increase in accuracy we can achieve the continuum limit with a moderate total number of segments M, *i.e.* with 50 to 500 segments for molecules of up to 20 atoms.

Implementation into MOPAC

For a first implementation of COSMO we have chosen the semi-empirical MO package MOPAC.[23,27] Within the framework of MOPAC, the charge distribution of a molecule can be represented quite accurately by the sum of atomic point charges and atomic dipoles. We have adopted this notation and thus have avoided the slightly more complicated basis function notation. Introducing three additional dipole points $\boldsymbol{r}_{\alpha k}$ (see Fig. 1) for each non-hydrogen atom at $\boldsymbol{R}^\alpha + d\boldsymbol{e}_k$ ($k = x, y, z$), where $\boldsymbol{e}_k$ denotes the unit vectors in Cartesian space, the charge distribution can be represented by point charges $Q_{\alpha k}$ on the resulting set of representation points. The number N of these points is just equal to the number of atomic orbitals. The dipole length d is a parameter which has to be chosen as a compromise between artificial quadrupole moments associated with dipoles of finite length and numeric instability arising from too small distances of the point charges. We found $d = 0.01$ Å to be a reasonable value. The charges $Q_{\alpha k}$ can be expressed by the electron density matrix **P** and the core charges Q_α^{core} *via*

$$Q_{\alpha 0} = Q_\alpha^{core} - (P_\alpha^{ss} + P_\alpha^{xx} + P_\alpha^{yy} + P_\alpha^{zz}) - (Q_{\alpha x} + Q_{\alpha y} + Q_{\alpha z}) \quad (17)$$

and

$$Q_{\alpha k} = -\frac{\bar{d}_\alpha}{d} p_\alpha^{sk} \; (k = x, y, z) \quad (18)$$

where $\bar{d}_\alpha$ denotes the hybridization dipole length coded for the different elements in MOPAC. These charges can be expected to represent quite accurately the electrostatic interactions of the solute with its environment. To increase the accuracy up to atomic quadrupole contributions, an additional 15 points per atom would be required. Alternatively one could use the basis function formulation of the **B** matrix as given in eqn. (C2). In our opinion the latter way would be preferable, but it has not been implemented yet.

Since the $Q_{\alpha k}$ are linear with respect to the density matrix **P**, the total energy reduction ΔE to the solvent is a quadratic form with respect to **P**, in perfect analogy to the Coulomb energy. Thus the integration of the solvation energy into the Hamiltonian is straightforward: The terms not depending on **P**, *i.e.* the dielectric reduction of the core–core interactions, are added to the core–core interaction energy, the linear terms with respect to **P** correspond to a reduction of the core–electron interaction and are thus added to the core–electron interaction Hamiltonian. The quadratic terms correspond to a reduction of the electron–electron interaction and they are included in the two-electron operator. One should be aware that these dielectric contributions to the Hamiltonian are very large. For medium sized molecules they cause changes in the core–core interaction and electronic energies of the order of some 10^4 eV while the

Table 1 Calculated heats of formation and hydration energies/enthalpies (in kcal mol $^{-1}$)

Compound	Method	Heat of formation: Gas phase geom. in gas phase	Heat of formation: Gas phase geom. in solution	Heat of formation: Solution geom. in solution	Hydration: Energy (calc.)	Hydration: Enthalpy (exp.)
NH_4^+	AM1	150.6	59.6	59.5	91.1	88.0[17]
$N(CH_3)_4^+$	AM1	157.1	101.6	101.1	56.0	59.9[17]
$N(C_2H_5)_4^+$	AM1	132.1	84.2	84.2	47.9	57.0[17]
Formamide	AM1	−44.7	−60.3	−61.4	16.7	—
Glycine						
neutral	AM1	−101.6	−117.1	−117.3	15.7	—
zwitterion	AM1	−59.2	−118.3	−125.6	66.4	—
H_2O	PM3	−53.4	−61.0	−61.1	7.7	≈10[a]
$(H_2O)_2$	PM3	−110.4	−122.3	−123.8	13.4	—
$(H_2O)_3$	PM3	−170.4	−183.7	−185.8	15.4	—
HF	PM3	−62.7	−67.3	−67.3	4.6	—
$(HF)_2$	PM3	−131.2	−138.8	−139.3	8.1	—

[a] Enthalpy of vaporization.

real effect, *i.e.* the difference of these changes, typically is of the order of 0.1 eV. Fortunately, using double precision accuracy throughout the implementation, this does not cause any problems.

Using the resulting self-consistent density matrix **P** the associated charges Q_{ak} [eqns. (17) and (18)] can be evaluated and used for the calculation of analytical gradients. Since each segment position does depend on a single atom position, the implementation of the analytical gradient according to eqns. (13)–(16) is straightforward. Now, having the self-consistent total energy as well as the gradient, standard geometry optimization can be done. Owing to the finite segmentation of the SAS some random fluctuations in the total energy do unavoidably occur during the optimization procedure. These turn out to cause problems with the standard BFGS optimizer in MOPAC, but the eigenvector-following algorithm appears to be less sensitive to these fluctuations. Thus, using the MOPAC keywords EF and ANALYT we have achieved final gradient norms of about 0.1–0.4 kcal mol $^{-1}$ Å $^{-1}$ for the test molecules considered so far.* This clearly shows that the implemented analytical gradient for the screening energy is rather accurate and that the approximations used are less serious.

Some Preliminary Applications

In this section we present some preliminary applications of the COSMO implementation into MOPAC in order to demonstrate the performance of our approach. All calculations have been carried out with $\varepsilon = 78.4$ corresponding to water. We set the parameters NPPA to 1082 and NSPA to 60; this choice turned out to give a convergence for the dielectric screening energy of better than 1%.

Since the screening energy is very sensitive to the distance of the SAS from the atoms, the choice of proper parameters for the construction of the SAS is crucial. Tomasi and co-workers[6,7,8] have shown that distances in the order of van der Waals radii lead to reasonable values of the screening energies in continuum models. They usually work with van der Waals radii increased by 20%, but in some cases they use slightly smaller values. In the work of Negre *et al.*[9] van der Waals radii have been used. Since a detailed optimization of the surface parameters is beyond the scope of this paper, we also decided to use van der Waals distances, corresponding to a setting of $\delta_{sc} = R_{solv}$ in our model. The value of R_{solv} itself is less critical. In the following applications we have used $R_{solv} = 1$ Å.

Screening of Cations.—The series of cations NH_4^+, $N(CH_3)_4^+$ and $N(C_2H_5)_4^+$ has been studied using AM1.[28] For the three cations the unscreened and screened heats of formation in their gas phase geometries have been calculated and a geometry optimization in solution has been carried out. These results and the corresponding hydration energies are given in Table 1 together with experimental hydration enthalpies taken from ref. 17. In general there is a quite good agreement between calculated and experimental numbers. As one might have expected for such rigid molecules, geometry optimization in solution causes only slight contractions of the ions with marginal energy gains.

Screening of Formamide.—Formamide has been used as an example in the work of Rinaldi *et al.*[16] on geometry optimization in ellipsoidal cavities. Our AM1 results on formamide are given in Table 1. We find an energy gain of 1.1 kcal mol^{-1} due to geometry optimization in solution. This is accompanied by several changes in the bond lengths and bond angles, which all have the same tendency as those found by Rinaldi *et al.*, but they are larger by a factor of *ca.* 3. Unfortunately, at the moment it is impossible to decide whether this difference is due to the differences in the MO-methods, or due to limitations of the ellipsoidal cavity approach.

Screening of Glycine.—The screening effects on glycine have been frequently studied[8,29] in order to explain the experimental fact, that, in aqueous solution, the zwitterionic species of glycine is about 10 kcal mol^{-1} more stable than the neutral, while in the gas phase the neutral species is by far the more stable.

The different heats of formation from an AM1 analysis for the two species of glycine are given in Table 1. The energy difference between the zwitterion and the neutral species decreases from 42 to −1 kcal mol^{-1} owing to the dielectric screening for the gas phase geometry. This difference is further decreased by geometry optimization within the solvent. While for the neutral species this only yields 0.2 kcal mol^{-1} with no qualitative geometry changes, there arise strong effects for the zwitterion (see Fig. 2). The hydrogen bond existing between one of the NH_3^+-hydrogens and one of the oxygens in the gas phase is

* 1 cal = 4.184 J.

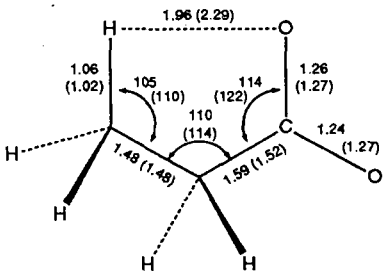

Fig. 2 Gas phase and solution optimized geometries of the glycine zwitterion: the hydrogen bond (dashed) is removed in solution

removed by an increase in the corresponding distance of 0.3 Å and most of the bond lengths and angles of the five-membered ring formed by the hydrogen bond are relaxed considerably. This relaxation goes along with an energy gain of 7.3 kcal mol^{-1}. Thus we finally obtain a total energy difference of 8.3 kcal mol^{-1} between the two species, in good agreement with the experimental finding. The same qualitative result, *i.e.* the release of the hydrogen bond in the zwitterion was obtained within the supermolecule approach,[29] but with orders of magnitude higher computational effort. In addition, it is hard to imagine, that such a ring-opening 'reaction' could be described within the ellipsoidal cavity approximation.

Screening of the Water Molecule and its Di- and Tri-mers.—The conformations and interaction energies of small clusters of water molecules have been the subject of several theoretical investigations.[30] For the dimer the gas phase conformation is generally accepted to have C_s-symmetry. The hydrogen bond between the two molecules is entirely straight, *i.e.* the donating OH-bond of the one molecule is oriented directly towards the accepting oxygen of the other molecule, while the remaining hydrogen of the donating molecule is pointing outside to minimize the total dipole of the dimer. While AM1 does not yield this geometry at all, it is reproduced correctly by the PM3-Hamiltonian. Hence we studied the dielectric screening of water and its di- and tri-mers within a dielectric continuum model of water using PM3 (see Table 1).

Dielectric screening with $\varepsilon = 78.4$ reduces the heat of formation of H_2O by 7.7 kcal mol^{-1}. Geometry optimization in solution only yields additional 0.02 kcal mol^{-1}, accompanied by a marginal increase in the bond lengths and a decrease in the bond angle of 1°. The hydration energy of 7.7 kcal mol^{-1} is in quite good agreement with the heat of vaporization of about 10 kcal mol^{-1}. The screening effect on the dipole moment is an increase from 1.74 to 2.06 Debye.

The heat of formation of the water dimer in the PM3 gas phase geometry is reduced by the dielectric environment by 13.4 kcal mol^{-1}. Geometry optimization in solution leads to a 180°-rotation of the donating water molecule around the hydrogen bond axis, while the other degrees of freedom are entirely unaffected (see Fig. 3). This rotation, which goes along with an increase of the dipole moment from 2.7 to 4.3 Debye, yields an additional 1.5 kcal mol^{-1}. Thus we find a stabilization of the screened dimer of 1.6 kcal mol^{-1} compared to two isolated screened water molecules. The corresponding interaction energy in the gas phase is 3.5 kcal mol^{-1}.

A similar rotation has been found for the triangular water trimer. In this case the rotation leads to full C_3-symmetry of the cluster, while in the gas phase this symmetry is prevented by a high dipole moment. Although the relevance of these conformational changes is not clear at the moment, they clearly demonstrate the benefit of geometry optimization within the dielectric environment since unexpected changes in the geometry might otherwise remain undetected.

Fig. 3 Gas phase and solution optimized geometries of the water dimer: the left molecule is rotated by 180° in solution (gas phase geometry dashed). Both conformations have C_s-symmetry.

Screening of Hydrogen Fluoride and its Dimer.—The HF dimer has been studied in the paper of Bonaccorsi *et al.*[7] on geometry optimization of molecular solutes. Even on this rather simple molecule they report considerable difficulties appearing in an automated geometry optimization within their continuum approach. These are due to the lack of analytical gradients. Therefore we have also considered HF and its dimer. Again we have used the PM3-Hamiltonian, owing to a better agreement of the dimer geometry with *ab initio* calculations. The results are reported in Table 1. Within COSMO the geometry optimization of the dimer in solution causes no problems. It yields an additional solvation energy of about 0.5 kcal mol^{-1}. Only small changes in the geometry are observed. As far as is comparable these are in agreement with the trends in the geometry optimisation reported by Bonaccorsi *et al.*

Discussion

The COSMO approach introduced in this article allows for the calculation of a comparatively simple, explicit expression for the screening energy of a molecule in a dielectric medium. For the first time this enables the calculation of analytic gradients within a real-shape cavity. Hence reliable geometry optimization within solvents becomes practicable and this should enable considerable progress in the computational treatment of solvation effects.

Obviously we could not resolve all the questions arising from the new approach, yet. The most critical problem in our opinion is the optimum choice of the solvent parameters R^{solv}, δ^{sc} and of a set of van der Waals radii, which in summary define the SAS. The values used throughout our examples should only be looked at as a reasonable first guess for water, but further work is needed on this point, especially if different solvents are considered. Nevertheless, similar problems arise from any continuum approach and are not specific for COSMO. More specific, but probably less important, is the question of the effects arising from the neglect of small parts of the SAS due to the atomwise segmentation which is needed for the calculation of gradients. These effects can be shown to be of the order of 1% of the total screening energy, or less.

Although our preliminary implementation of COSMO into MOPAC has proved to work properly, a lot of work remains to be done. First of all this implementation should be optimized, especially with respect to faster geometry optimization, for example by the use of updating algorithms for the A^{-1}-matrix. Furthermore it should be generalized to become compatible with all the different functionalities of MOPAC like vibrational analysis, eigenvector following, CI, *etc.* Furthermore a refinement of the charge representation by taking into account atomic quadrupoles would be desirable for the sake of consistency with the MOPAC calculation of Coulomb interactions, although we do not expect a large influence on the results. These improvements should be possible without major problems, but a good knowledge of the architecture of MOPAC is required. They are beyond the scope of our work at a chemical company. Therefore we have submitted our implementation to J. J. P. Stewart for further development and distribution as part of the MOPAC software.

Implementations of COSMO into *ab initio* as well as force-field codes should be practicable without larger problems, but

they too are beyond the scope of our work. Thus we would like to encourage people who are familiar with such codes to do this work and to contact us for details.

Acknowledgements

The authors would like to thank Henry S. Rzepa and James J. P. Stewart as well as their colleagues Thorsten Pötter and Felix Reichel for stimulating discussions and valuable hints. We thank A. Reed for critically reading the manuscript.

Appendix A

As previously mentioned, the screening energy of a charge distribution due to a dielectric medium scales as $(\varepsilon - 1)/(\varepsilon + x)$ with $0 \leqslant x \leqslant 2$. This can be seen from different standard problems of electrostatics. In Onsager's solution for the screening of a multipole of order l at the centre of a spherical cavity,[10] x takes the value $l/(l + 1)$. For a monopole, *i.e.* a point charge, in front of a planar surface of a dielectric medium the value of x is 1 while x becomes 2 for a monopole far away from a sphere with relative dielectric permittivity ε. The latter case seems to be very unrealistic for charges within the cavity formed by a molecule. Thus we can expect the best value of x to be between zero and one. We propose to use $x = \frac{1}{2}$ and hence an ε-dependent correction factor of $f(\varepsilon) = (\varepsilon - 1)/(\varepsilon + \frac{1}{2})$ for the total screening energies and gradients appearing in our approach. Thus the relative error should be less than $\frac{1}{2}\varepsilon^{-1}$. This is negligible, or at least tolerably small, for most of the relevant polar solvents in organic chemistry. For weak dielectrics with $\varepsilon \approx 1$, the relative error may become 50%, but in this case the solvation energy itself is very small and thus the absolute error remains irrelevant.

Appendix B

The total electrostatic energy of a sphere of radius R with a unit charge homogeneously distributed over its surface is $0.5/R$. On the other hand for a homogeneous segmentation of the sphere into M equivalent segments of area $|S_\mu| = 4\pi R^2/M$ the total energy can be calculated as the sum of the Coulomb interactions of the charges M^{-1} on different segments plus the sum of their self energies. Hence using the notation of the Theory section we have

$$\frac{1}{2R} = \frac{M}{2}\sum_{\nu=2}^{M}\frac{M^{-2}}{\|t_1 - t_\nu\|} + MM^{-2}a_{diag}/2 \qquad (B1)$$

and

$$a_{diag} = (4\pi|S_\mu|)^{\frac{1}{2}}M^{-\frac{1}{2}}\left(M - \sum_{\nu=2}\frac{R}{\|t_1 - t_\nu\|}\right) \qquad (B2)$$

For $M = 4, 6, 8$ and 12, *i.e.* for segmentations corresponding to the vertices of a tetrahedron, octahedron, cube and icosahedron, respectively, the sum of the inverse distances can be calculated easily. In these cases the last two terms of eqn. (B2) approximately take the value 1.07. This result even holds for homogeneous segmentations with values of M up to 3242, which have been generated by the algorithm described for construction of a solvent accessible surface. Thus we have got the desired expression for the diagonal elements of **A**.

Appendix C

During the text we only considered point charges Q_i as sources. To adjust the formalism to MO-calculations we have to consider charge density distributions $\rho_{\kappa\lambda}(r) = P_{\kappa\lambda}\varphi_\kappa(r)\varphi_\lambda(r)$ instead of, or, more precisely, in addition to, point charges Q_i at positions r_i. Here $\varphi_\kappa(r)$ and $\varphi_\lambda(r)$ denote basis functions describing the position of the resulting overlap charge density and hence taking the role of r_i, while the density matrix elements $P_{\kappa\lambda}$ describe the amount of charge corresponding to the overlap of these basis functions and hence they take the part of Q_i. In this analogy the index i is replaced by a double indexation $\kappa\lambda$ denoting the combinations of basis functions. Exploiting the analogy further we have

$$b_{\kappa\lambda,\mu} = \frac{1}{|S_\mu|}\int_{S_\mu}\int\frac{\varphi_\kappa(r)\varphi_\lambda(r)}{\|r' - r\|}\,d^3r\,d^2r' \qquad (C1)$$

$$\approx \int\frac{\varphi_\kappa(r)\varphi_\lambda(r)}{\|t_\mu - r\|}\,d^3r \qquad (C2)$$

The integral in eqn. (C2) is of the core–electron interaction type and thus readily available in any MO-code. If we define a generalized charge vector $\boldsymbol{Q}$ consisting of the point charges of the nuclei or atom cores, respectively, and of electron densities $P_{\kappa\lambda}$, eqn. (12) does hold even in the MO-formulation.

References

1 A. Warshel and S. Russel, *Quart. Rev. Biophys.*, 1984, **17**, 283.
2 A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley, New York, 1991.
3 W. F. van Gunsteren, *Angew. Chem.*, 1990, **102**, 1020.
4 B. Pullman, S. Miertus and C. Perahia, *Theor. Chem. Acta*, 1979, **50**, 317.
5 S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
6 F. M. Floris, J. Tomasi and J. L. Pascual-Ahuir, *J. Comput. Chem.*, 1991, **12**, 784.
7 R. Bonaccorsi, R. Cammi and J. Tomasi, *J. Comput. Chem.*, 1991, **12**, 301.
8 R. Bonaccorsi, P. Palla and J. Tomasi, *J. Am. Chem. Soc.*, 1984, **106**, 1945.
9 N. Negre, M. Orozco and F. J. Luque, *Chem. Phys. Lett.*, 1992, **196**, 26.
10 D. Morales-Lagos and J. S. Gomez-Jeria, *J. Phys. Chem.*, 1991, **12**, 784.
11 C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc.*, 1991, **113**, 8305.
12 O. Tapia, *J. Mol. Struct. (Theochem.)*, 1991, **226**, 59.
13 H. H. Hoshi, M. Sakurai, Y. Inoue and R. Chujo, *J. Chem. Phys.*, 1987, **87**, 1107.
14 D. Rinaldi and J. L. Rivail, *Theor. Chem. Acta*, 1973, **32**, 57.
15 D. Rinaldi, P. E. Hoggan and A. Cartier, GEOMOS, QCPE 584, *QCPE Bull.*, 1989, **9**, 128.
16 D. Rinaldi, J. L. Rivail and N. Rguini, *J. Comput. Chem.*, 1992, **13**, 675.
17 G. P. Ford and B. Wang, *J. Comput. Chem.*, 1992, **13**, 229.
18 M. M. Karelson, A. R. Katritzky, M. Szafran and M. C. Zerner, *J. Org. Chem.*, 1989, **54**, 6030.
19 M. M. Karelson, T. Tamm, A. R. Katritzky, S. J. Cato and M. C. Zerner, *Tetrahedron Comp. Meth.*, 1989, **2**, 295.
20 H. S. Rzepa, M. M. Yi, M. M. Karelson and M. C. Zerner, *J. Chem. Soc., Perkin Trans. 2*, 1991, 635.
21 See *e.g.*: J. D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1975.
22 L. J. Onsager, *J. Am. Chem. Soc.*, 1936, **58**, 1436.
23 J. J. P. Stewart, MOPAC program package, QCPE Nr. 455.
24 M. J. S. Dewar, AMPAC program package, QCPE Nr. 506.
25 H. L. Friedman, *Mol. Phys.*, 1975, **29**, 1533.
26 F. Silla, I. Tunon and J. L. Pascual-Ahuir, *J. Comput. Chem.*, 1991, **12**, 1077.
27 J. J. P. Stewart, *J. Comp.-Aided Mol. Design*, 1990, **4**, 1.
28 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.
29 H. S. Rzepa and M. M. Yi, *J. Chem. Soc., Perkin Trans. 2*, 1991, 531.
30 See *e.g.*: E. Honegger and S. Leutwyler, *J. Chem. Phys.*, 1988, **88**, 2582.

Paper 3/00240C
Received 14*th January* 1993
Accepted 9*th February* 1993