JCTC Journal of Chemical Theory and Computation

Evaluation of the SCF Combination of KS-DFT and 3D-RISM-KH; Solvation Effect on Conformational Equilibria, Tautomerization Energies, and Activation Barriers

David Casanova,[†,||] Sergey Gusarov,[‡] Andriy Kovalenko,*[,‡,§] and Tom Ziegler[||]

Departament de Química Inorgànica and Centre de Recerca en Química Teòrica, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain, National Institute for Nanotechnology, National Research Council of Canada, 421 Saskatchewan Drive, Edmonton, Alberta, T6G 2M9, Canada, Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, T6G 2G8, Canada, and Department of Chemistry, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

Received May 20, 2006

Abstract: The effect of solvation on conformational equilibria, tautomerization energies, and activation barriers in simple S_N2 reactions is reproduced by using the self-consistent field coupling of the Kohn−Sham density functional theory (KS-DFT) for electronic structure and the three-dimensional reference interaction site model with the closure approximation of Kovalenko and Hirata (3D-RISM-KH) for molecular solvation structure. These examples are used in order to validate the implementation of the 3D-RISM-KH method in the Amsterdam Density Functional package. The computations of the free energy difference in the trans/gauche conformational equilibrium for 1,2-dichloroethane in different solvents; the relative tautomerization free energy for cytosine, isocytosine, and guanine; and the free energy activation barrier for a CH_3X-type ($X = F, Cl, Br$) S_N2 reaction exhibit agreement with the experimental data. The method is also applied to the electronic and hydration structure of carbon single-wall nanotubes of different diameters, including the effect of water located in the inner space of the nanotubes. A comparison with continuum models of solvation (including COSMO) as well as with other more precise and computationally more expensive calculations is made to demonstrate the accuracy and predictive capability of the new KS-DFT/3D-RISM-KH method.

1. Introduction

One of the challenges of contemporary quantum chemistry has been to give an accurate description of solute−solvent interactions. A number of methods have been used to account for solvent effects on molecular properties such as structure and reactivity. The methods currently in use can be divided into three categories: dielectric continuum schemes, procedures based on a combination of quantum mechanics with molecular mechanics (QM/MM) or with molecular dynamics (QM/MD and Car−Parrinello), and methods utilizing the integral equations derived in the statistical-mechanical theory of molecular liquids.

In this work, we compare the results obtained by the approaches falling into two of these categories. One is the well-known conductor-like screening model (COSMO),[6] already implemented some time ago in the Amsterdam Density Functional (ADF)[4] package. The other is the three-dimensional reference interaction site model with the closure proposed by Kovalenko and Hirata (3D-RISM-KH),[1,2] which has recently been implemented in a self-consistent field form

* Corresponding author e-mail: andriy.kovalenko@nrc-cnrc.gc.ca.
† Universitat de Barcelona.
‡ National Research Council of Canada.
§ University of Alberta.
|| University of Calgary.

10.1021/ct6001785 CCC: $37.00
Published on Web 02/22/2007

in ADF.[6] Our objective in this paper is to assess how well the 3D-RISM-KH method, more complex computationally but significantly more advanced theoretically, works in practical calculations in comparison with the empirical COSMO approach.

2. Theory

Here, we present a brief background theory for the two solvation methods that are used in the present study.

2.1. Dielectric Continuum Model of Solvation. The conductor-like screening model (COSMO) by Klamt and Schüürmann[5] belongs to the category of dielectric continuum methods. The solute molecule is considered as being situated in a cavity inside a structureless medium/continuum, with the cavity shapes and medium parameters empirically fitted to reproduce solvation thermochemistry data for a reference set of pure solvents. In the QM/COSMO approach, the original Hamiltonian is modified by adding a solvent potential which is constructed on the basis of the electrostatic potential of the charge distribution $\rho_s(\mathbf{r}_S)$ induced on the cavity surface S by the nuclear charges Z_A and the electronic density distribution $n(\mathbf{r})$ of the solute molecule.[6] The energy E^{es} due to the electrostatic interaction between the solute and cavity surface charges as well the self-interaction of the surface charges constitute the total electrostatic solvation energy given by

$$E^{es} = \int_S \sum_A \frac{Z_A\, \rho_s(\mathbf{r}_S)}{|\mathbf{R}_A - \mathbf{r}_S|}\, d\mathbf{r}_S + \int_S \int_S \frac{\rho_s(\mathbf{r}_S)\, \rho_s(\mathbf{r}_S)\, (\mathbf{r}'_S)}{|\mathbf{r}_S - \mathbf{r}'_S|}\, d\mathbf{r}_S\, d\mathbf{r}'_S + \int_V \int_S \frac{n(\mathbf{r})\, \rho_s(\mathbf{r}_S)}{|\mathbf{r} - \mathbf{r}_S|}\, d\mathbf{r}\, d\mathbf{r}_S \quad (1)$$

In addition to the electrostatic contribution, there is also a nonelectrostatic term $E^{non-elst}$ arising from cavitation, dispersion, and repulsion. These contributions are usually modeled as functions of the area of the cavity surface.[6]

The COSMO approach readily yields the solvation chemical potential of a given solute in the pure solvents that have been parametrized at certain (typically ambient) conditions, provided the COSMO molecular surface can be constructed for the given solute molecule. (The latter is meaningless, for instance, inside a nanotube channel.[7]) However, COSMO does not provide the complete physical picture of the solvation structure. Furthermore, the whole approach is not transferable to a given solvent at given thermodynamic conditions, not present in the database. Nor does it work for solvent mixtures or electrolyte solutions.[8]

The above limitations are partially obviated in the COSMO for real solvents (COSMO-RS) method developed by Klamt which gives quite accurate thermochemical information on liquids and mixtures.[8,9] It is based on the statistics of solvation shells coarse-grained into walls (or double layers) between COSMO cavities encapsulating molecules in the liquid. The solvation chemical potential is obtained as the statistical average of the interaction of wall charges over the distribution in the space of screening charge density values (σ profile). The interaction energy functional is approximated by the self-energy of the total charge density on the two adjacent sides of a wall segment (electrostatic misfit energy) plus the hydrogen-bonding energy, whereas the interaction between different segments is neglected. The ensemble of interacting molecules is thus reduced to an ensemble of independently interacting surface segments, with each molecule represented by a histogram of cavity surface area with respect to screening charge density. This enormously simplifies the problem and yields the solvation chemical potential in a simple calculation. The adjustment parameters for each type of chemical element constituting molecules are determined by using full QM/COSMO calculations for about 800 small molecules to fit experimental physicochemical data for their pure compounds and for the partition of solvent mixtures (typically, water with octanol, benzene, hexane, CCl_4, ether, and ethylacetate) at temperatures around $T = 260-340$ K and ambient pressure.

As seen from the above, COSMO-RS is a post-COSMO method interpolating the thermochemistry of a molecular liquid or a mixture between the QM/COSMO fits for a number of simple, pure compounds at thermodynamic conditions around ambient values. The interpolation is based, in fact, on very simplified and yet rather reasonable approximations: (i) coarse-graining three-dimensional solvation shells into two-dimensional molecular surfaces and (ii) representing the solvent-mediated effective potential between two molecules by local interaction dependent just on the screening charge densities at the contact of their molecular surfaces, with the partition coefficient resolved in the space of screening charge density values. That is, the two "labeled" molecules of the liquid effectively interact through a local contact of their molecular surfaces. This essentially contracts the orientational and, in part, translational degrees of freedom of a "third" molecule in the overlapping solvation shells mediating the effective interaction (a field vertex in the diagrammatic language). It is hence clear where this approach works and where it does not. COSMO-RS gives good results for liquid−liquid, vapor−liquid, and solid−liquid equilibrium properties and the solubility of many compounds and mixtures. Currently, the parametrization is available for molecules that contain only hydrogen, carbon, nitrogen, oxygen, fluorine, phosphorus, sulfur, chlorine, bromine, and iodine.[8,10] However, COSMO and COSMO-RS produce poor results for the phase equilibria of a number of substances, in particular, for polymers and highly polar compounds of complex shape like amines.[10] This is apparently caused by a more ordered solvation structure of such compounds, which is mistreated by the COSMO-RS coarse-graining. Furthermore, COSMO-related approaches are not able to deal with volumetric properties, as the COSMO representation of the solvation structure by a molecular surface totally discards the information about the structure in the dimension across the solvation shells. It reduces the degrees of freedom across the shells to average values on the molecular surface. We emphasize that, although such a mean field can always be introduced in the equations, it cannot be derived in a transferable form within COSMO for a number of cases when the 3D solvation structure is important. It has been shown that the second solvation shell contributes equally to

or even more than the first one into the partial molar volume and compressibility of the solute, which are related to excluded volume and electrostriction.[11] The physics of these effects is beyond COSMO-RS. Moreover, COSMO-related approaches are not applicable at all to systems with confined geometry for which the notion of a molecular surface just doesn't make sense, such as a solvent inside and in pockets of an organic rosette nanotube[12] and inside a carbon nanotube (see section 4.3.3 below).

Of crucial importance in computational chemistry is the description of chemical reactions in solution. Concerning this task, COSMO-RS has significant difficulty with handling geometry optimization, as it requires rerunning the COSMO calculation for the solute molecule at each step of geometry modification.[13] Meanwhile, a single calculation of a screening charge distribution for large molecules with hundreds of atoms becomes so time-consuming[3] that even a self-consistent field QM/COSMO calculation becomes unfeasible and is usually replaced with one COSMO iteration, thus treating the solvent effect just as a perturbation. However, a principal limitation of COSMO-related approaches in treating chemical reactions lies in their inability to accurately yield transition states. The analytical gradients of the solvation free energy are available in COSMO only for the electrostatic term and not for the nonelectrostatic contributions arising from cavitation, dispersion, and repulsion, which are empirically constructed.[6] COSMO also requires additional approximations to avoid discontinuities and singularities in the derivatives arising from the polyhedral representation of the molecular surface. This is done in some practical realizations like the smooth solvation model,[14] which can reproduce the potential energy surface for most of the cases, with some exceptions.[14] However, transition states with the elements having the electron density strongly different from that in their stable states fall out of the COSMO-RS parametrization database. They are practically not amenable to the COSMO-RS parametrization procedure because of the necessity to possess reference solvation thermodynamics data from simulations (or perhaps experiments) for all of the COSMO-RS large training set compounds but with their elements in transition states with a range of different partial charges and sizes.

To address all of the above principal drawbacks of COSMO-related empirical approaches to solvation, one needs to resort to the integral equation theory of molecular liquids, or molecular theory of solvation, which stands on the first principles of statistical mechanics and thus reproduces the essential physics of solvation.

2.2. Interaction Site Model. The reference interaction site model (RISM) is one of the commonly used methods of the integral equation theory of liquids.[15] The RISM integral equation, or site−site Ornstein−Zernike integral equation, was first introduced in 1972 by Chandler and Andersen[16] to study the classical solvation structure of molecular liquids. It is an extension of the Ornstein−Zernike equation for atomic liquids[15] to molecular liquid with atoms strongly bound by intramolecular correlations representing chemical bonds. The RISM integral equation yields the site−site radial distribution functions between interaction sites of molecules of liquid, completely averaged over orientations of the molecules.

As distinct, the 3D-RISM[1,2,17,18] method treats a solute molecule fully in a three-dimensional manner, whereas orientational averaging around interaction sites is applied to solvent molecules. This yields the three-dimensional classical density distributions of interaction sites of solvent molecules around a solute particle of arbitrary shape. The 3D-RISM integral equation can be derived from the molecular Ornstein−Zernike equation[15] for the six-dimensional solute−solvent correlation functions by reducing the orientations of solvent molecules.[1,2] Alternatively, the 3D-RISM integral equation can be obtained from the 3D-RISM in the hypernetted chain approximation (3D-RISM-HNC theory)[17] constructed on the basis of the density functional theory of nonuniform polyatomic liquids.[19] The 3D-RISM integral equation gives the 3D density correlations representing the response of a molecular solvent to the external field of the solute:

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int \mathrm{d}\mathbf{r}'\, c_\alpha(\mathbf{r}-\mathbf{r}')\, \chi_{\alpha\gamma}(r') \qquad (2)$$

where $h_\gamma(\mathbf{r})$ and $c_\gamma(\mathbf{r})$ are respectively the 3D total and direct correlation functions of solvent site γ around the solute and $\chi_{\alpha\gamma}(r)$ is the site−site susceptibility of the pure solvent. The 3D total correlation functions $h_\gamma(\mathbf{r})$ are related to the 3D distribution function $g_\gamma(\mathbf{r}) = h_\gamma(\mathbf{r}) + 1$, which gives the normalized probability of finding solvent site γ at position $\mathbf{r}$ with respect to the solute molecule. Outside the molecular core, the 3D direct correlation function $c_\gamma(\mathbf{r})$ has the asymptotics of the interaction potential $u_\gamma(\mathbf{r})$ between solvent site γ and the solute, $c_\gamma(\mathbf{r}) \sim -u_\gamma(\mathbf{r})/(k_\mathrm{B}T)$, where $k_\mathrm{B}T$ is the Boltzmann constant times the solvent temperature. The solvent susceptibility breaks up into the intra- and intermolecular parts

$$\chi_{\alpha\gamma}(\mathbf{r}) = \omega_{\alpha\gamma}(r) + \rho_\alpha\, h_{\alpha\gamma}(r) \qquad (3)$$

where $\omega_{\alpha\gamma}(r)$ is the intramolecular pair correlation matrix function defining the geometry of the solvent molecule, $h_{\alpha\gamma}(r)$ is the radial total correlation function between sites α and γ in the pure solvent, and ρ_α is the mean number density of solvent site α. The intramolecular matrix of a rigid molecule has the form $\omega_{\alpha\gamma}(r) = \delta(r - z_{\alpha\gamma})/(4\pi r_{\alpha\gamma}^2)$, or $\omega_{\alpha\gamma}(k) = \sin(kz_{\alpha\gamma})/(kz_{\alpha\gamma})$ in the reciprocal space, where $z_{\alpha\gamma}$ are the Z matrix elements. Equations 2 and 3 split the effective interaction between two molecules of liquid up into the direct intermolecular interaction and the medium induced correlation propagating through other molecules of liquid.

In liquid state theory, the integral equation for the total and direct correlation functions must be complemented with an additional relation between them, called "closure". The exact closure is nonlocal and extremely cumbersome; therefore, it is usually replaced with one of the known approximations, for instance, the hypernetted chain (HNC) closure and the mean spherical approximation (MSA).[15] Effects predicted by an integral equation theory of liquid stem from singularities and asymptotics inherent in a closure approximation involved. Kovalenko and Hirata[1,2] proposed a closure (KH approximation) which nontrivially combines HNC and MSA, appropriately applying them to regions of

density depletion and enrichment. This properly accounts for density distribution peaks and long-range tails brought about by chemical specificities of molecules, as well as for density depletion within their repulsive cores.[2,20,21] In this study, we complement the 3D-RISM integral eq 2 for the 3D total and direct correlations functions with the 3D-KH closure which reads[1,2]

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp(d_\gamma(\mathbf{r})) & \text{for } d_\gamma(\mathbf{r}) \le 0 \\ 1 + d_\gamma(\mathbf{r}) & \text{for } d_\gamma(\mathbf{r}) > 0 \end{cases} \quad (4)$$

where

$$d_\gamma(\mathbf{r}) = -\frac{u_\gamma(\mathbf{r})}{k_B T} + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$$

To properly treat electrostatic forces in electrolyte solution with a polar solvent when numerically solving the 3D-RISM-KH eqs 2 and 4, the Coulomb singularities of all the correlation functions $h_\gamma(\mathbf{r})$, $c_\gamma(\mathbf{r})$, and $\chi_{\alpha\gamma}(\mathbf{r})$ are separated out in both the direct and reciprocal spaces and accounted for analytically.[2,22] This includes correction for the supercell periodicity artifacts arising from the 3D fast Fourier transform employed to calculate the convolution in eq 2.

To obtain the radial total correlation functions of solvent $h_{\alpha\gamma}(\mathbf{r})$ constituting the intermolecular part of the solvent susceptibility in eq 3, we use the dielectrically consistent RISM theory (DRISM) developed by Perkyns and Pettitt,[23,24] which consistently describes the dielectric properties of an electrolyte solution comprising ions and a polar molecular solvent. We solve the DRISM integral equation complemented with the KH closure, with the long-range electrostatic asymptotics of the total and direct correlation functions being separated out and treated analytically.[2,22] The computational cost of solving the DRISM integral equations is orders of magnitude smaller than that for 3D-RISM, and the solution can be stored for further use.

Similarly to the RISM-HNC[25] and 3D-RISM-HNC[1,2] theories, the 3D-RISM-KH eqs 2 and 4 lead to the closed analytical expression for the excess chemical potential of solvation[1,2]

$$\Delta\mu^{\mathrm{KH}} = k_B T \sum_\gamma \rho_\gamma \int \mathrm{d}\mathbf{r} \left\{ \frac{1}{2} (h_\gamma(\mathbf{r}))^2 \Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r})\, c_\gamma(\mathbf{r}) \right\} \quad (5)$$

In what follows, we denote the changes due to solvation by Δ for all quantities.

Unlike the COSMO approach constructing the electrostatic potential of the solvent cavity,[6] our method combined with Kohn−Sham density functional theory (KS-DFT/3D-RISM-KH) yields the self-consistent field of solvent from the first principles by introducing the Helmholtz free energy functional defined as

$$A = E^{\mathrm{el}} + \Delta\mu[n(\mathbf{r}),\rho_\alpha(\mathbf{r})] \quad (6)$$

where E^{el} is the internal energy of the solute including translational, rotational, and vibrational degrees of freedom and $\Delta\mu_{\mathrm{solv}}$ is the part of the solvation free energy (excess chemical potential of solvation) due to the solute−solvent interaction and solvent reorganization. The solvent potential term $v(\mathbf{r})$ in the KS equation as well as the classical potential $u_\alpha(\mathbf{r})$ of the solute molecule acting on site α of the solvent molecules are obtained analytically by variational differentiation of the functional (eq 6) with respect to the 3D density of solute valence electrons $n(\mathbf{r})$ and the 3D site density distributions of solvent molecules $\rho_\alpha(\mathbf{r}) = \rho_\alpha g_\alpha(\mathbf{r})$:

$$v(\mathbf{r}) = \frac{\delta\Delta\mu}{\delta n(\mathbf{r})} \quad (7)$$

$$u_\alpha(\mathbf{r}) = \frac{\delta\Delta\mu}{\delta\rho_\alpha(\mathbf{r})} \quad (8)$$

With the term $\Delta\mu_{\mathrm{solv}}$ given by eq 5, the electronic and classical potentials (eqs 7 and 8) coupling the solute electronic and solvent classical subsystems are obtained in a closed analytical form in terms of pseudopotentials $v_\alpha^{(\mathrm{ps})}(\mathbf{r})$ representing the total electronic density of interaction sites of solvent molecules:[1,2]

$$v(\mathbf{r}) = \sum_\alpha \int \mathrm{d}\mathbf{r}'\, \rho_\alpha(\mathbf{r} - \mathbf{r}')\, v_\alpha^{(\mathrm{ps})}(\mathbf{r}') \quad (7')$$

$$u_\alpha(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\, n(\mathbf{r} - \mathbf{r}')\, v_\alpha^{(\mathrm{ps})}(\mathbf{r}') \quad (8')$$

In the simplest way, the solvent site pseudopotentials $v_\alpha^{(\mathrm{ps})}(\mathbf{r})$ can be represented as a sum of the core repulsion, dispersion, and electrostatic terms.[1,2] This results in the subdivision of the classical interaction potentials $u_\alpha(\mathbf{r})$ between the whole solute and solvent interaction sites into the Lennard-Jones and Coulomb parts.

Notice that, unlike COSMO, the KS-DFT/3D-RISM-KH theory is based on the proper physical concept following from the first principles and statistical physics. The self-consistency is ensured for both the solute electronic structure and the solvent distribution determining the shape of the excluded volume and solvation shell. We emphasize that the latter replaces the concepts of a solvent cavity and molecular surface in continuum solvation models.

Another important advantage is that the 3D-RISM-KH theory allows one to analytically calculate the derivatives of the free energy functional.[3] This makes the method readily applicable to molecular geometry optimization and the study of chemical reactions in solutions. The analytical gradients from the RISM theory in combination with KS-DFT for molecules in solution were first derived by Sato et al.[26] It should be noted that our analytical gradients following from the 3D-RISM-KH theory are consistent since they are based on the solvation potential $v(\mathbf{r})$ obtained directly from the definition (eq 7). The latter is not possible for 1D-RISM, which gives $\Delta\mu_{\mathrm{solv}}$ in terms of the radial correlations dependent on partial site charges. It thus requires contraction of the electron density by using the population operator, which, in addition, can be defined in different ways. Another approach to the analytical gradients within the 3D-RISM method is based on separation of the space around the solute into long- and short-range regions.[27] The resulting estimate

expressions for the analytical gradients are more complicated and depend on a definition of the regions.

2.3. Thermodynamic Calculations. For comparison with experimental results, we compute the Gibbs free energies. In the gas phase, the Gibbs free energy of the solute is related to its enthalpy by

$$G_{gas} = H_{gas}(T) - TS_{gas}^{trans} - TS_{gas}^{rot} - TS_{gas}^{vib} \quad (9)$$

The part $H_{gas}(T)$ contains the internal potential energy of the ideal gas and the standard contributions due to the zero-point temperature corrections and the finite temperature corrections to the enthalpy. The entropic terms for translational, rotational, and vibrational motion are obtained from the standard expressions.[28] The free energy of a particle in solution G_{sol} can be split up into that in the gas phase and the solvation free energy:

$$G_{sol} = G_{gas} + \Delta G \quad (10)$$

In COSMO, the solvation free energy is broken up as

$$\Delta G = \Delta G^{solute} + E^{es} + E^{non-es} \quad (11)$$

where the electrostatic energy E^{es} of the solute−solvent interaction is given by eq 1 and the nonelectrostatic contribution E^{non-es} is modeled empirically as a function of the cavity surface area. Notice that eq 11 does not include the change in the translational entropy of solvent ΔS induced by the solute. The term ΔG^{solute} represents the changes in the solute and is composed in COSMO as follows:

$$\Delta G^{solute} = \Delta E^{el} + \Delta H^{rot}(T) + \Delta H^{vib}(T) - T\Delta S^{rot} - T\Delta S^{vib} \quad (12)$$

where ΔE^{el} is the energy of solute electronic reorganization upon transfer from the gas phase to solution, including the response to the change in the solute geometry. The other terms $\Delta H^{rot,vib}(T)$ and $T\Delta S^{rot,vib}$ are the changes induced by the solvent in the enthalpy and entropy of rotational and vibrational degrees of freedom of the solute. The energy and enthalpy terms in eq 12 are replaced altogether with $\Delta H(T)$, obtained from the same functional form $H_{gas}(T)$ but with the inertia moments and vibrational frequencies modified by solvation. Furthermore, one has to construct the rotational and vibrational entropic terms $T\Delta S^{rot,vib}$, as well as the translational term $T\Delta S$ missing in eq 11. They are usually estimated on the basis of the entropies in the gas phase by using the procedure of Wertz.[29] For a water solvent, this procedure gives the following empirical expression:[30]

$$S_{sol} = 0.258 + 0.54\, S_g^{\circ} \quad (13)$$

In sections 4.1 and 4.2, where we deal with the relative free energies between molecules with similar modes, ΔE^{el} is the major contribution and the other terms in eq 12 are neglected.[31]

With the 3D-RISM methodology, the solvation free energy ΔG is split up into the solute electronic reorganization energy ΔE^{el}, the change in the rotational and vibrational free energies of the solute upon transfer from the gas phase to solution ΔG^{rot} and ΔG^{vib}, and the excess chemical potential of solvation $\Delta\mu$ due to the solute−solvent interaction and solvent reorganization:

$$\Delta G = \Delta E^{el} + \Delta G^{rot} + \Delta G^{vib} + \Delta\mu \quad (14)$$

The excess chemical potential of solvation $\Delta\mu$ obtained by expression 5 contains the solute−solvent (uv) interaction energy E^{uv}, the energy of solvent reorganization (vv) around the solute ΔE^{vv}, and the solvent translational entropy change ΔS induced by the solute:

$$\Delta\mu = E^{uv} + \Delta E^{vv} - T\Delta S \quad (15)$$

The change in the vibrational terms is small ($\Delta G^{vib} < 1$ kcal/mol) and can be neglected, especially in the calculation of relative free energies as in sections 4.1 and 4.2. We can readily obtain the solvent effect on the vibrational free energy ΔG^{vib} by the standard frequency calculation. However, it is small and is not included in the results presented below.

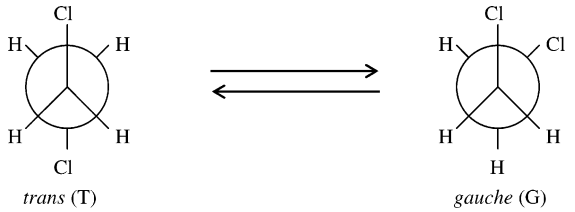

Figure 1. Trans/gauche equilibrium for 1,2-dichloroethane.

3. Computational Details

All calculations were carried out using the ADF program package.[4] We used both the local-density approximation (LDA)[32] and the BP86 scheme in which LDA is augmented with gradient corrections for exchange[33] and correlation.[34] The majority of the results presented in this work were computed using the BP86 scheme. The calculations carried out with LDA are indicated explicitly. The basis set was of triple-ζ quality with one polarization function added (TZP). In the KS-DFT/3D-RISM-KH calculations, the 3D-FFT grid size was 64 × 64 × 64 in the supercell of size 32 × 32 × 32 Å^3. Further refinement of the grid did not result in any considerable changes. In one particular case of the reaction $Br^- + CH_3Cl \rightarrow CH_3Br + Cl^-$, the geometry of the species and the shape of the solvation potential required a doubling of the grid resolution.

The COSMO implementation is that introduced into ADF by Pye and Ziegler,[6] whereas the 3D-RISM-KH scheme has been recently implemented in ADF by Gusarov et al.[3]

4. Results and Discussion

We shall now evaluate the self-consistent field combination of KS-DFT with the 3D-RISM-KH method in a number of applications studying the effect of solvation on several molecular properties and compare the results with experimental data as well as other solvation approaches, including COSMO.

4.1. Solvent Effect on Conformational Equilibria. The free energy difference between the two conformers (trans and gauche) of the 1,2-dichloroethane molecule (Figure 1)

$$\Delta G^{T\rightarrow G} = G^G - G^T \quad (16)$$

Table 1. Free Energy Difference $\Delta G^{T\to G}$ (kcal/mol) between the trans and gauche Conformers of 1,2-Dichloroethane in the Gas Phase (Upper Part) and in Solvents (Lower Part)[a]

	this work	Capelli [b]	experiment [c]		
gas phase	1.83	1.5	1.35		
solvent	dielectric constant[c]	COSMO	PCM[d]	3D-RISM-KH	experiment[c]
n-hexane	2.02	1.34	0.95	1.57	1.15
1,4-dioxane	2.20	1.28	0.82	0.04	0.54
carbon tetrachloride	2.23	1.27	0.81	0.32	1.03
tetrachloroethylene	2.50	1.18	0.79	1.20	0.95
diethyl ether	4.34	0.84	0.61	0.54	0.77
ethyl acetate	6.00	0.70	0.48	0.63	0.50
THF	7.52	0.62	0.38	−0.11	0.40
acetone	20.70	0.4	0.13	0.16	0.18

[a] Predictions of the KS-DFT coupled with COSMO, PCM, and 3D-RISM-KH methods, versus experimental results. [b] B3LYP and 6-31G**, ref 36. [c] From refs 35 and 36. [d] From ref 36.

Table 2. Solvent Effect on the 1,2-Dichloroethane trans/gauche Conformational Free Energy Difference, eq 17, $\Delta\Delta G^{T\to G}_{gas\to sol}$ (kcal/mol)

solvent	dielectric constant[a]	COSMO	PCM[b]	3D-RISM-KH[c]	experiment[a]
n-hexane	2.02	0.49	0.55	0.13	0.20
1,4-dioxane	2.20	0.55	0.68	1.66	0.81
carbon tetrachloride	2.23	0.56	0.69	1.38	0.32
tetrachloroethylene	2.50	0.65	0.71	0.50	0.40
diethyl ether	4.34	0.99	0.89	1.16	0.58
ethyl acetate	6.00	1.13	1.02	1.07	0.85
THF	7.52	1.21	1.12	1.81	0.95
acetone	20.70	1.43	1.37	1.54	1.17

[a] From refs 35 and 36. [b] From ref 36. [c] Without the vibrational correction.

has been computed in different solvents using the COSMO and 3D-RISM-KH methods, as well as in the gas phase. The values are compared with the results of previous calculations[35−37] and experimental data[35,36] for several solvents in Table 1. This is a good test for theoretical predictions; the conformational equilibrium of the 1,2-dichloroethane molecule in solution is quite sensitive to the solute reorganization energy and the solvation shell structure, as was shown for instance in the case of 1,2-dichloroethane in water.[38]

The experimental value for $\Delta G^{T\to G}_{gas}$ in the gas phase is positive. Thus, the trans conformer is the most stable in the gas phase, as one would expect on the basis of steric grounds. The value of $\Delta G^{T\to G}_{gas}$ calculated in this work is 1.83 kcal/mol, which is about 0.5 kcal/mol larger than the experimental estimate[35,36] (Table 1). Fortunately, this error cancels out when we turn to the point of interest for our study, namely, the solvent effect on the trans/gauche equilibrium. We expect that the errors in the electronic structure calculation for $\Delta G^{T\to G}_{gas}$ in the gas phase and $\Delta G^{T\to G}_{sol}$ in solution are similar and almost cancel out in their difference, giving the solvent effect on the conformational stability.

In solution, the free energy difference between the two conformers $\Delta G^{T\to G}_{sol}$ becomes smaller (Table 1). This is understood, as only the gauche conformer has a permanent dipole, and is likely to be stabilized more than the trans conformer by solvation. Compared to the experiment, all the utilized theoretical methods seem to reproduce the overall trend in $\Delta G^{T\to G}_{sol}$ with respect to different solvents.

In order to highlight the solvent effects, we also computed the change in the conformational energy difference $\Delta G^{T\to G}$ upon transfer from the gas phase to solution:

$$\Delta\Delta G^{T\to G}_{gas\to sol} = \Delta G^{T\to G}_{gas} - \Delta G^{T\to G}_{sol} \qquad (17)$$

Table 2 makes a comparison of the theoretical and experimental results for the same solvents as in Table 1. The values of $\Delta\Delta G^{T\to G}_{gas-sol}$ obtained using the polarizable continuum model (PCM) by Capelli et al.,[36] COSMO, and 3D-RISM-KH are close to the experimental ones. In most cases, they all slightly overestimate $\Delta\Delta G^{T\to G}_{gas\to sol}$. The results of COSMO and PCM are close to each other. The COSMO values are within a 0.4 kcal/mol range from the experimental data. However, the COSMO and PCM curves run monotonically with the solvent dielectric constant ϵ and do not reproduce some features. In Figure 2, we visualize the data from Tables 1 and 2, making the comparison between the theoretical models and experimental results for $\Delta\Delta G^{T\to G}_{gas\to sol}$ as well as $\Delta G^{T\to G}_{sol}$ as a function of the dielectric constant ϵ of the solvent. The curves for the continuum models show the expected increase of $\Delta\Delta G^{T\to G}_{gas\to sol}$ with the value of ϵ : $f(\epsilon) = (\epsilon - 1)/(\epsilon + 0.5)$. However, solvation effects in molecular solvents with a low ϵ can significantly deviate from the monotonic behavior predicted by continuum models. Indeed, the experimental data for both the conformation energy difference $\Delta G^{T\to G}$ and its change $\Delta\Delta G^{T\to G}_{gas\to sol}$ in a 1,4-dioxane solvent clearly stand out of the smooth dependence

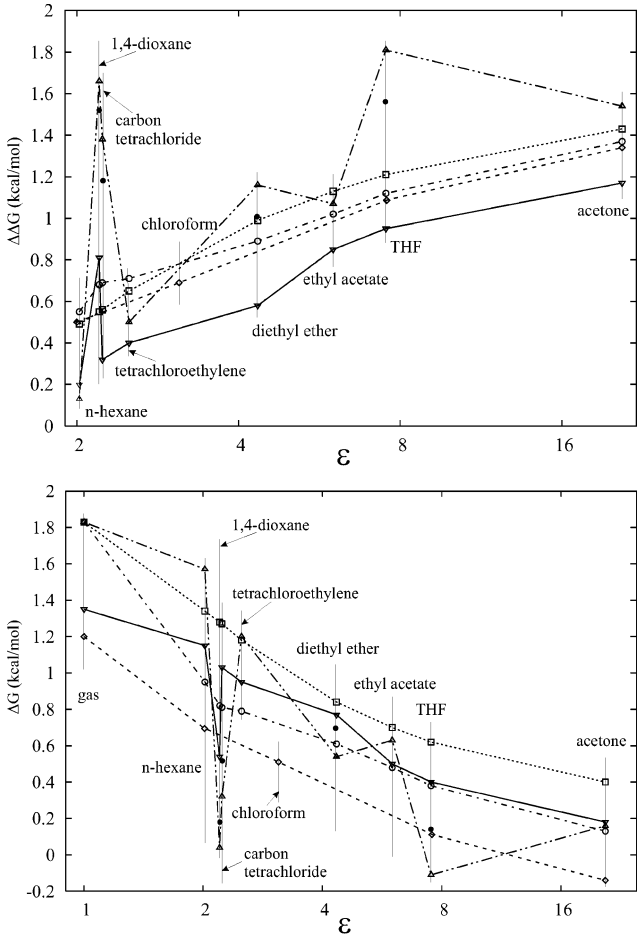

Figure 2. Conformational free energy difference between the trans and gauche conformers of 1,2-dichloroethane in solution $\Delta G^{T\rightarrow G}$ (lower part) as well as its change $\Delta\Delta G^{T\rightarrow G}$ upon transfer from the gas phase to solution (upper part) given by eq 17. Predictions of the KS-DFT coupled with continuum models (COSMO = squares and dotted lines, PCM = circles and dash-dotted lines; the COSMO results are without the translational, vibrational, and rotational thermodynamic corrections to the free energy, given by eq 12) and with the 3D-RISM-KH theory[b] (united-atom solvent models = triangles and dash-double-dotted lines, all-atom models = filled circles; the 3D-RISM values are without the vibrational free energy corrections). Experimental data (references from Capelli[35,36] = diamonds, from Colominas[54] = upside down triangles, and solid lines). The connecting lines are added for the eye.

for the other solvents with similar low ϵ values. This effect cannot be explained simply as dielectric screening by a cavity in a structureless solvent and, obviously, involves short-range interactions of the solute and solvent molecules in the solvation shell, which requires resolution of the three-dimensional solvation structure and a proper account of both energetic and entropic contributions into the solvation free energy. The quality of the electronic basis set and the three-dimensional detail of the solvation model are particularly important for the conformational equilibrium of solvated 1,2-dichlorethane.[38] The 3D-RISM-KH method adequately represents the 3D structure of a molecular solvation shell and, therefore, readily reproduces, in agreement with experimental results, the off-trend values of $\Delta G^{T\rightarrow G}$ and $\Delta\Delta G^{T\rightarrow G}_{gas\rightarrow sol}$ in the 1,4-dioxane solvent. The qualitative agreement with experimental results improves when the united-atom model utilized for the solvent throughout this work is replaced with the more precise all-atom one (filled pentagons in Figure 2). The remaining discrepancy is largely due to the quality of the electronic part, and we were able to further refine the agreement by optimizing the basis set and functional used (data not presented here). We stress that, for this solvent with a low ϵ, the 3D-RISM-KH method readily reproduces the deviation from the monotonic behavior peculiar to dielectric screening, which is missed by continuum solvation models but is indeed observed in experimental results.[35,36]

4.2. Solvent Effect on Tautomerization Energies and Structures. Since the work by Watson and Crick,[39] the problem of tautomeric equilibria has been studied widely. Also, tautomerism has been used in several models to explain the spontaneous mutation of DNA.[40,41] For this reason, a great amount of experimental[42−51] and theoretical[52−57] work based on the study of the tautomerism of nucleic bases has been published. In spite of these efforts, there are still unresolved details. The main theoretical problem has been obtaining accurate results from quantum mechanical calculations in the aqueous phase.

In the analysis of the solvent effect on tautomerization energies and structures of cytosine and isocytosine, we compare our method with the available experimental data and other theoretical studies. As there are no experimental data related to the stability of guanine, protonated cytosine, and protonated guanine tautomers, we compare our calculations just with the theoretical studies available in the literature.

4.2.1. Cytosine. We begin with studying the relative stability of the seven possible tautomers of the cytosine molecule (Figure 3). Experimental IR spectra[42−51,58] and theoretical[52,53,55−57] studies in the gas phase have shown that cytosine exists in three tautomeric forms. The two most stable forms are the aminooxo (canonic) and iminohydroxo forms (CYT2a and CYT2b in Figure 3), while the third most stable form is the aminooxo species, which is observed in small amounts. The aminohydroxo form has no biological significance due to the fact that in DNA the proton at the N1 position is substituted by a sugar moiety.

The energies obtained in this work (Table 3) predict relative stabilities that agree quite well with previous experimental[41−45] and theoretical studies:[52−57]

$$\text{CYT1} \geq \text{CYT4a} \geq \text{CYT2a} \geq \text{CYT4b} \approx \text{CYT2b} \gg \text{CYT3} \gg \text{CYT5}$$

Our calculations predict that in the gas phase the cytosine molecule exists at equilibrium conditions as a mixture of the following tautomers: aminooxo (CYT1), iminooxo (CYT4a and CYT4b), and aminohydroxo (CYT2a and CYT2b). The energy difference between the iminooxo and the aminooxo forms found in the present work is quite small (1.6 kcal/mol), which is in agreement with the previous results (Table 3). Also, we found, in line with other studies,[54] that the enol-imino form (CYT5) is less stable than other tautomers by 15−20 kcal/mol. In a polar solvent like water, there is a change in the relative stability of the cytosine

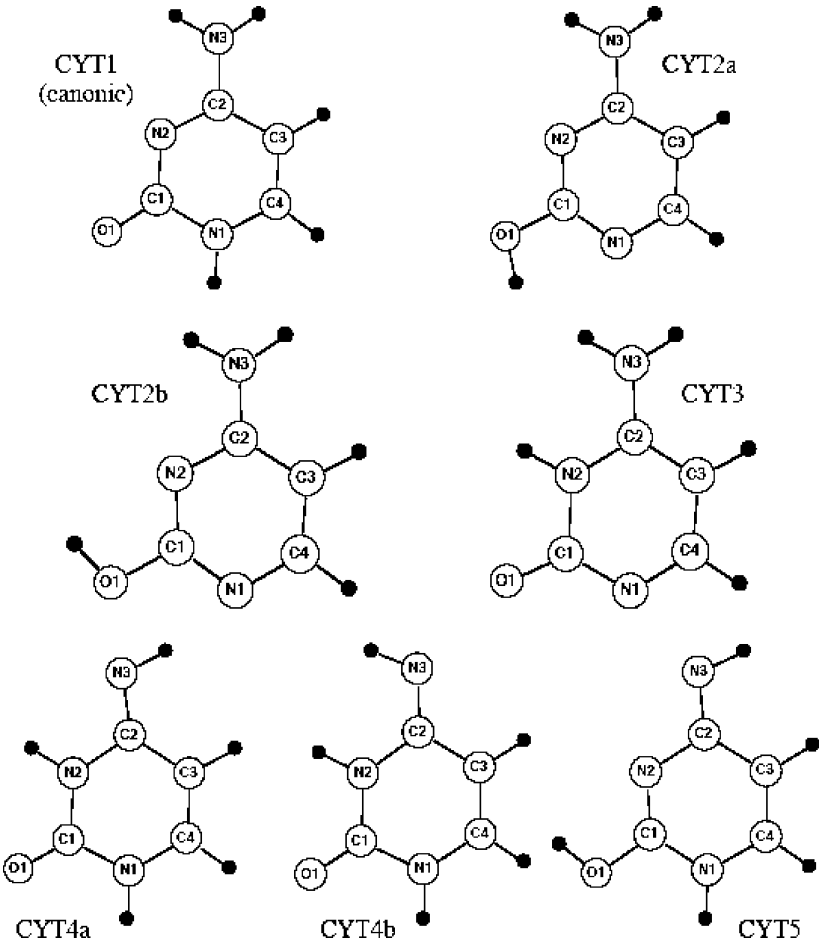

Figure 3. Cytosine tautomers.

Table 3. Free Energy (kcal/mol) of the Cytosine Tautomers in the Gas Phase with Respect to the CYT1 Tautomer

tautomer	this work	Gorb[a]	Colominas[b]	Kobayashi[c]	experiment[d]
CYT 1	0.00	0.0	0.0/0.0	0.0/0.0	
CYT 2a	2.62	0.5	−0.8/1.3	0.330/1.545	0.965
CYT 2b	3.30		0.0/2.0	0.660/0.727	
CYT 3	6.41	7.3	7.0/7.1		
CYT 4a	1.80	1.4	3.0/3.7	−1.45/0.335	
CYT 4b	3.10		1.6/2.2	3.274/2.919	
CYT 5	21.68	20.7			

[a] From ref 56. [b] MP4/6-311++G(d,p)//MP2/6-31G(d) and B3LYP(6-311++G(d,p))//MP2/6-31G(d), respectively.[54] [c] Electronic energies calculated at the CCSD(T)//MP2 and DFT//DFT (cc-pvtz(-f) basis) levels of theory, respectively.[52] [d] From ref 43.

Table 4. Free Energy (kcal/mol) of the Cytosine Tautomers in Aqueous Solution with Respect to the CYT1 Tautomer

tautomer	COSMO	3D-RISM-KH	Gorb[a]	Gorb[b]	Colominas[c]
CYT 1	0.00	0.00	0.0	0.0	0.0
CYT 2a	10.86	8.17	5.6	5.0/3.2	6.8
CYT 2b	10.34	7.31			7.1
CYT 3	2.63	3.75	2.9	7.8	5.6
CYT 4a	6.63	7.59	5.4	4.5/7.0	6.1
CYT 4b	7.11	8.81			6.1
CYT 5	23.35	22.24	22.7	20.5	

[a] SCRF model.[56] [b] Tautomer plus one water molecule.[56] [c] SCRF model.[54]

tautomers; only the canonic forms have been observed experimentally.[46−51]

Here, we assess the accuracy of the 3D-RISM-KH method by comparing its predictions for the relative stabilities with experimental data and other theories,[54,56] including the COSMO scheme. Our calculations for the solvent effects by the 3D-RISM-KH method and continuum model (COSMO) predict that the aminooxo (CYT1) is the only tautomer observable in solution (Table 4), in agreement with experimental results for IR spectra in a water solution.[46−51] The consonance of the 3D-RISM-KH and COSMO results with those of previous theoretical works[54,56] is quite remarkable. All the methods find that the canonical form CYT1 is much more stable in solution than the other tautomers. Most of the calculations seem to identify the tautomer CYT3 as the second most stable in solution. The most important stabilization in solution takes place in the CYT3 tautomer, which has the largest dipole moment (Table 5).

The free energy of solvation for the canonic form is much larger than that for the other stable tautomers in the gas phase (CYT2a, CYT2b, CYT4a, and CYT4b). This explains why the iminooxo and aminohydroxo forms are not detected in solution. In general, as we can expect, there is a correlation between the solvation free energy and the dipole moment of the tautomers. The computed dipole moments in solution are always by several kilocalories per mole larger than those in the gas phase because of the polarity of the water solvent. Notice that the dipole moment values obtained by using the

Table 5. Solvent Effect on the Difference of the Free Energy of the Cytosine Tautomers with Respect to the CYT1 Tautomer, $\Delta\Delta G_{gas\rightarrow sol}^{CYT1\rightarrow t}$, and Dipole Moments μ of the Cytosine Tautomers in the Gas Phase and in Aqueous Solution

	$\Delta\Delta G_{gas\rightarrow sol}^{CYT1\rightarrow t}$ (kcal/mol)			μ (Debye)		
tautomer *t*	COSMO	3D-RISM-KH	Colominas[a]	gas phase	COSMO	3D-RISM-KH
CYT 1	0.00	0.00	0.0	6.50	10.96	10.93
CYT 2a	8.23	5.55	7.6	3.50	5.69	5.57
CYT 2b	7.04	4.02	7.1	4.86	7.74	7.61
CYT 3	−3.78	−2.67	−1.4	8.22	14.01	13.83
CYT 4a	4.83	5.79	4.5	4.69	7.42	7.38
CYT 4b	4.01	5.71	4.5	2.41	4.28	4.15
CYT 5	1.67	0.56		5.52	9.79	9.87

[a] From ref 54.

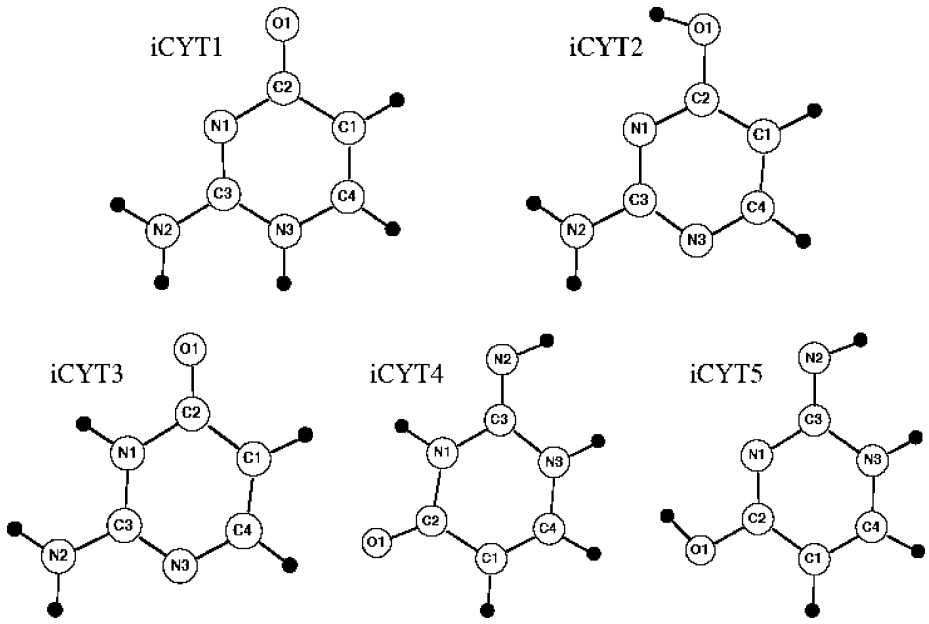

Figure 4. Isocytosine tautomers.

3D-RISM-KH method and the COSMO scheme are very similar and show the same trend (Table 5).

4.2.2. Isocytosine. Isocytosine, in contrast to cytosine, has not been detected in natural DNA or RNA. On the other hand, the C nucleoside of isocytosine is known as an antileukemia drug.[59] Moreover, isocytosine has been incorporated enzymatically into both DNA and RNA.[60−67]

We have studied the different stabilities of the five tautomers of the isocytosine molecule (Figure 4) in the gas phase and in solution. The experimental data in the gas phase[68−72] show that the isocytosine exists in a mixture of two forms, iCYT2 and iCYT3, where the aminohydroxo (iCYT2) seems to be the most stable one. Our calculations in the gas phase (Table 6) reproduce this trend and predict iCYT3 and iCYT2 as the most stable tautomers, although the former is 0.5 kcal/mol more stable than the latter. We determined all the other tautomers to be considerably less stable, in agreement with the fact that iCYT2 and iCYT3 are the only detectable tautomers in the gas phase. In general, our calculations of the relative free energy follow the same trend as that in previous work by Gorb et al.[56] (Table 6). Much as for the tautomers of cytosine, the enol-imino form (iCYT5) is less stable by 10 kcal/mol.

In aqueous solution, the relative stabilities are quite different. Experimentally, only the aminooxo forms (iCYT1 and iCYT3) have been found.[41,46−51,58,73,74] Previous self-consistent reaction field (SCRF)[75] calculations on these tautomers did not reproduce the relative stability order observed experimentally. The KS-DFT/3D-RISM-KH method predicts the existence of three tautomers in solution: iCYT1, iCYT2, and iCYT3. The results of our COSMO calculations give a substantial difference between the relative stabilities of iCYT1, iCYT2, and iCYT3. These values are not quite in agreement with the experimental observation of the three species. Meanwhile, 3D-RISM-KH predicts the aminooxo forms as the most stable in solution. The difference in stability between the two aminooxo forms iCYT1 and iCYT3 is quite small (1.3 kcal/mol), and their free energy in solution is lower than iCYT2 by about 4 and 5 kcal/mol, respectively

Table 6. Free Energy (kcal/mol) of the Isocytosine Tautomers in Gas Phase with Respect to the iCYT2 Tautomer

tautomer	this work	Gorb[a]
iCYT 1	8.98	9.7
iCYT 2	0.00	0.0
iCYT 3	−0.50	1.2
iCYT 4	5.55	7.0
iCYT 5	18.39	22.4

[a] From ref 56.

Table 7. Free Energy (kcal/mol) of the Isocytosine Tautomers in Aqueous Solution with Respect to the iCYT3 Tautomer

tautomer	COSMO	3D-RISM-KH	Gorb[a]	Gorb[b]
iCYT 1	−4.70	1.26	1.9	8.4
iCYT 2	8.93	5.37	1.3	0.9
iCYT 3	0.00	0.00	0.0	0.0
iCYT 4	5.60	6.44	5.7	8.2/9.4
iCYT 5	12.38	16.50	23.7	20.7

[a] SCRF model.[56] [b] Tautomer plus one water molecule.[56]

(Table 7). Thus, the 3D-RISM-KH method seems to give a more consistent account of the experimental fact that only the two aminooxo forms are detectable.

The calculations by Gorb et al.[56] used only a single water molecule to represent the solvent. They predict the iCYT1 and iCYT2 tautomers to be the most stable species, in disagreement with experimental data. Although they differ by about 5−10 kcal/mol from the KS-DFT calculations with the COSMO and 3D-RISM-KH treatments of solvation, all of these calculations agree that iCYT5 is the least stable tautomer in solution.

The relative free energy of solvation for the five tautomers of isocytosine (Table 8) calculated with 3D-RISM-KH and COSMO shows that the iCYT1 tautomer is the most stable in aqueous solution. The hydration thus changes the relative order of stability among the isocytosine tautomers in solution, compared to that found in the gas phase.

An important stabilizing factor for the tautomer iCYT1 in water is its dipole moment interacting with the solvent. On the other hand, the iCYT2 tautomer has the smallest solvation free energy, in agreement with its modest dipole moment. Much as for the cytosine tautomers, the calculated dipole moments increase in aqueous solution because of polarization by water.

4.2.3. Guanine. We shall next turn to a study of the relative stability for some of the tautomers of the guanine molecule. We have studied five tautomers of guanine (Figure 5): two keto-amino (G19 and G17) and three enol-amino (G96c, G96t, and G76c) species. The enol-imino forms are not included in this work, as it is well-known that they are very unstable with respect to the keto-amino and enol-amino tautomers.

The relative free energies computed in the gas phase show that the two keto-amino forms G19 and G17 are the most stable (Table 9). In our calculations, G17 appears to be the most stable and G19 slightly differs in energy by ∼0.5 kcal/mol. These values are very close to the DFT calculations combined with self-consistent reaction field and Monte Carlo simulations carried out by Colominas et al.[54]

The enol-amino forms G96c and G96t are ∼2.5 and ∼3.0 kcal/mol less stable than the G17 tautomer. Although these values differ by about 1 kcal/mol from the estimates of Colominas et al., the stability order predicted is the same. Finally, the tautomer G76c is the least stable in the gas phase; it has a free energy more than 5 kcal/mol higher than that of the most stable tautomer G17.

In aqueous solution, the order of relative stability remains nearly the same, but the difference in stability between some tautomers increases (Table 10). The results obtained with 3D-RISM-KH and COSMO agree with the order of tautomers stability computed previously by Colominas et al.[54] using optimized versions of the continuum model developed by Miertus et al.[76] Much as in the gas phase, the G17 and G19 tautomers are the most stable in aqueous solution. Their relative stability differs by just about 1 kcal/mol, with G17 being the most stable in the gas phase and G19 favored in solution. The most stable tautomer G19 has the largest dipole moment (Table 11).

The lower stability of the enol-amino species G96c, G96t, and G76c relative to the most stable keto-amino forms G17 and G19 is emphasized in aqueous solution, the difference considerably increasing by 7 kcal/mol (Table 11). The G76c tautomer is the least stable in water.

4.2.4. Protonated Cytosine. We have studied just two protonated cytosine molecules (Figure 6) derived from the most stable neutral cytosine tautomers (keto-amino and enol-amino forms). A previous theoretical study of the other protonated enol forms[54] revealed that they are more than 9 kcal/mol less stable. Further, the protonated imino species were shown to be very unstable by some 29 kcal/mol compared to the keto-amino derivatives.

The energy difference between the two studied protonated species pCYT13 and pCYT12c is very small in the gas phase (Table 12). The order of stability we obtained agrees with the previous DFT study by Colominas et al.;[54] however, the difference is larger by 1.5 kcal/mol.

The difference between the solvation free energy of pCYT13 and pCYT12c (Table 13) increases the stability gap between the two tautomers. Both 3D-RISM-KH and COSMO computations show that in aqueous solution the keto-amino form pCYT13 is much more stable than the enol-amino form pCYT12c. The difference between the solvation free energy of these protonated forms is likely related to the negative charge on the keto-oxygen (Table 14).

4.2.5. Protonated Guanine. Five protonated tautomers of guanine (Figure 7) have been studied in the gas phase (Table 15) and in aqueous solution. Colominas et al.[54] determined that all tautomers are within 6 kcal/mol. We found a slightly larger difference (about 6.6 kcal/mol) between the most and least stable tautomers pG179 and pG376c, respectively. The second most stable tautomer is pG376c, while pG137 and pG796c have quite similar stabilities.

In aqueous solution, the difference in the tautomers' stabilities increases (Table 16). Much as in the gas phase, pG179 and pG196t are the most and least stable tautomers in water. The pG137 tautomer has the larger hydration energy (Table 17) and, therefore, becomes the second most stable tautomer (very close to pG179). Again, the stabilization of the tautomers correlates quite well with the negative charge on the oxygen atom (Table 17).

4.3. Solvent Effect on Activation of S_N2 Reactions with CH_3X (X = F, Cl, Br). In organic chemistry, the bimolecular nucleophilic S_N2 substitution reaction

$$Y^- + RX \rightarrow YR + X^- \qquad (18)$$

has been studied extensively by both theoretical[77−88] and experimental[88−94] means. Experimental studies have shown that the rate of S_N2 reactions involving CH_3X species with

Table 8. Solvent Effect on the Difference of the Free Energy of the Isocytosine Tautomers with Respect to the iCYT1 Tautomer, $\Delta\Delta G_{gas\rightarrow sol}^{iCYT1\rightarrow t}$, and Dipole Moments μ of the Isocytosine Tautomers in the Gas Phase and in Aqueous Solution

	$\Delta\Delta G_{gas\rightarrow sol}^{iCYT1\rightarrow t}$ (kcal/mol)		μ (Debye)		
tautomer *t*	COSMO	3D-RISM-KH	gas phase	COSMO	3D-RISM-KH
iCYT 1	0.00	0.00	8.56	14.90	14.68
iCYT 2	17.02	13.10	1.41	1.98	1.87
iCYT 3	8.58	8.23	4.50	7.42	7.20
iCYT 4	8.14	8.62	5.69	9.15	8.97
iCYT 5	7.67	5.83	4.85	8.19	8.27

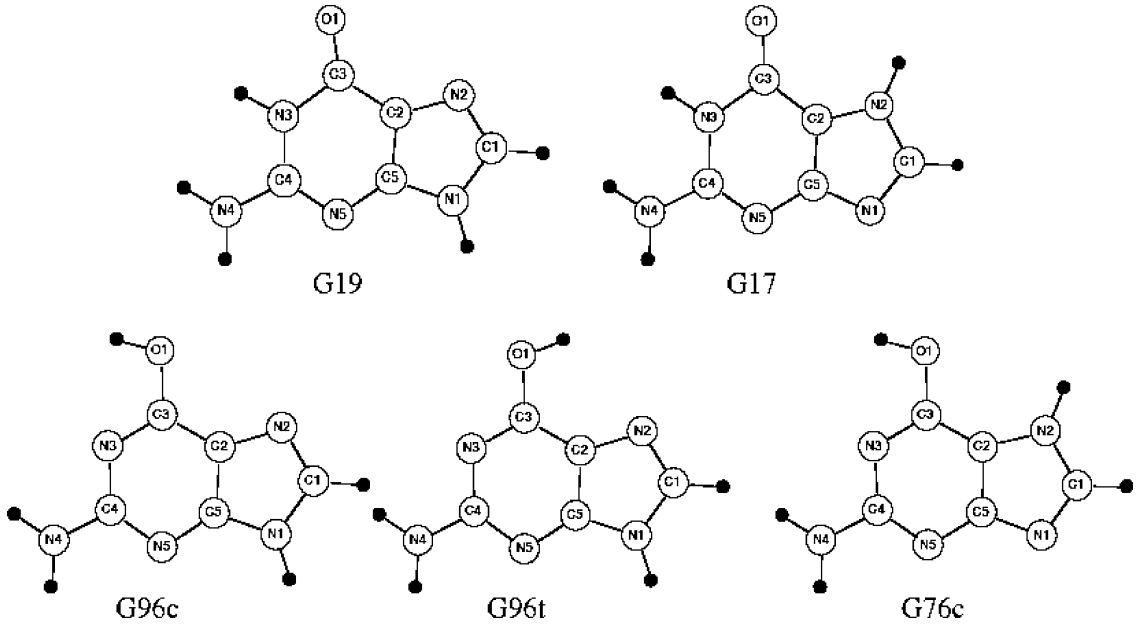

Figure 5. Guanine tautomers.

Table 9. Free Energy (kcal/mol) of the Guanine Tautomers in the Gas Phase with Respect to the G19 Tautomer

tautomer	this work	Colominas[a]	Colominas[b]
G19	0.00	0.0	0.0
G17	−0.57	0.2	−0.4
G96c	2.09	1.1	1.1
G96t	2.53	1.8	1.8
G76c	4.49	4.4	3.7

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54] [b] B3LYP(6-311++G(d,p))//MP2/6-31G(d).[54]

Table 10. Free Energy (kcal/mol) of the Guanine Tautomers in Aqueous Solution with Respect to the G19 Tautomer

tautomer	COSMO	3D-RISM-KH	Colominas[a]
G19	0.00	0.00	0.00
G17	0.40	1.10	1.00
G96c	10.11	7.33	7.20
G96t	10.57	7.18	8.00
G76c	10.94	9.32	8.80

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54]

halogen X = F, Cl, and Br are quite different in the gas phase and in solution. Our aim here is to demonstrate how well the 3D-RISM-KH method reproduces this difference, compared to the COSMO solvation model.

4.3.1. Identity S_N2 Reactions. The S_N2 reaction both in the gas phase and in solution is characterized by a typical double-well potential energy profile, although the depth of the potential well in solution is marginal (Figure 8). We shall first study the case in which the leaving group X is identical to the incoming nucleophile Y in eq 18. It is convenient to introduce the difference RC = R(X−C) − R(Y−C) as a reaction coordinate of the S_N2 reaction, where X and Y are the leaving and incoming halogens, respectively. Figure 8 displays the potential energy profile against the reaction coordinate RC.

The two minima of the double-well potential for the gas-phase identity S_N2 reaction correspond to the ion−dipole complexes. The minima are separated by a central energy barrier representing the transition state of the whole reaction. In the transition state at RC = 0, the carbon atom is in the center of a bipyramid, with the two halogens in the axial positions equally distant from the carbon (Figure 9). Typical free energy parameters that define an S_N2 reaction in the gas phase are the complexation free energy ΔG^C, the overall barrier ΔG^B relative to the reactants, and the intrinsic barrier ΔG^{IB}, as shown in Figure 8.

In solution, the potential energy profile is quite different. In this case, there are almost no ion−dipole minima and the reaction can be defined in terms of the overall barrier between the reactant and transition states. Solvation makes

Table 11. Solvent Effect on the Difference of the Free Energy of the Guanine Tautomers with Respect to the G19 Tautomer, $\Delta\Delta G_{gas\rightarrow sol}^{G19\rightarrow t}$, and Dipole Moments μ of the Guanine Tautomers in the Gas Phase and in Aqueous Solution

	$\Delta\Delta G_{gas\rightarrow sol}^{G19\rightarrow t}$ (kcal/mol)			μ (Debye)		
tautomer *t*	COSMO	3D-RISM-KH	Colominas[a]	gas phase	COSMO	3D-RISM-KH
G19	0.00	0.00	0.0	6.73	11.15	10.84
G17	0.97	1.67	0.8	2.17	3.72	3.85
G96c	8.02	5.25	6.1	3.30	4.96	4.92
G96t	8.04	4.66	6.2	3.50	5.56	5.16
G76c	6.46	4.84	4.4	3.62	6.39	6.55

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54]

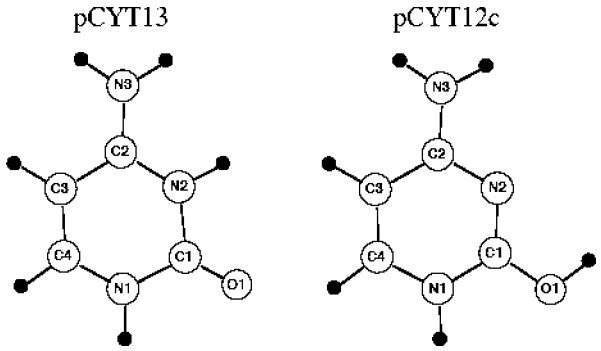

Figure 6. Protonated cytosine tautomers.

Table 12. Free Energy (kcal/mol) of the Protonated Forms of Cytosine in the Gas Phase with Respect to the pCYT13 Tautomer

tautomer	this work	Colominas
pCYT 13	0.00	0.0
pCYT 12c	1.96	−0.4[a]/0.3[b]

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54] [b] B3LYP(6-311++G(d,p))//MP2/6-31G(d).[54]

Table 13. Free Energy (kcal/mol) of the Protonated Forms of Cytosine in Aqueous Solution with Respect to the pCYT13 Tautomer

tautomers	COSMO	3D-RISM-KH	Colominas[a]
pCYT 13	0.00	0.00	0.0
pCYT 12c	11.97	8.43	10.7

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54]

a substantial contribution to the activation energy and can even change the relative order of stability between reactants and the transition state.

In this study, we consider three different cases of the methyl-transfer S_N2 identity reaction with X = Y = F, Cl, and Br. The angle Y−C−X assumes the values in the range from 160 to 180°, as one might expect from an S_N2-type reaction with inversion of the configuration. Moreover, the optimized geometries for the different ion−dipole complexes and the transition state are similar to those obtained in previous works.[85−87,95] The free energies in the gas phase and in aqueous solution were calculated by introducing the thermodynamic terms as detailed above. We computed the difference between the free energy barrier in aqueous solution and in the gas phase, which is shown in Figure 8:

$$\Delta\Delta G_{gas\rightarrow sol}^{B} = \Delta G_{sol}^{B} - \Delta G_{gas}^{B} \qquad (19)$$

For COSMO, the deviation in $\Delta\Delta G_{gas\rightarrow sol}^{B}$ from experimental results amounts to about 1 kcal/mol, whereas for 3D-RISM-KH, it is 3 kcal/mol (see Table 18). As is seen, the 3D-RISM-KH results considerably depend on the accuracy of the Lennard-Jones parameters used. Also presented in Table 18 are the results of the COSMO calculation without the translational, rotational, and vibrational thermodynamic corrections. In this case, the discrepancy with the experiment becomes much larger, which shows the importance of these terms.

4.3.2. Nonidentity S_N2 Reactions. We have also studied some methyl-transfer nonidentity S_N2 reactions. The transition structure is no longer a bond-symmetric arrangement at the midpoint of the reaction coordinate connecting reactants and products (RC ≠ 0). One of the most important distinctions between identity and nonidentity S_N2 reactions is the presence of a "thermodynamic driving force" in the latter due to a free energy difference between the reactants and products.

We obtained a typical potential energy profile and relevant structures both in the gas phase and in solution by using COSMO and 3D-RISM-KH (see Table 19). Notice that the COSMO calculation for $\Delta\Delta G_{gas\rightarrow sol}^{B}$ without the translational, rotational, and vibrational thermodynamic corrections by eq 12 is in poor agreement with experimental results. This shows that these corrections are important and should be accounted for in most of the cases by standard methods. The optimized structures do not differ appreciably from those obtained in previous calculations.[84,86] The prediction for the overall free energy difference $\Delta\Delta G_{gas\rightarrow sol}^{B}$ given by eq 19 for X = F and Y = Cl obtained with COSMO is very close to the experimental value (the deviation is about 0.5 kcal/mol), whereas it is 3.5 kcal/mol for 3D-RISM-KH. For the two other cases of X = F/Y = Br and X = Cl/Y = Br, the values of $\Delta\Delta G_{gas\rightarrow sol}^{B}$ obtained with COSMO are overestimated as 6 and 3 kcal/mol, respectively. The 3D-RISM-KH calculation for X = F and Y = Cl achieves an even better precision of 0.24 kcal/mol. For X = Cl and Y = Br, the value obtained with 3D-RISM-KH is at least 5−7 kcal/mol lower than the experimental values of$\Delta\Delta G_{gas\rightarrow sol}^{B}$. Unlike the case of X = F, 3D-RISM-KH considerably overemphasizes the experimentally observed decrease of $\Delta\Delta G_{gas\rightarrow sol}^{B}$ for X = Cl. This discrepancy can be eliminated upon further optimization of the Lennard-Jones parameters of the solute, as they substantially change in the reaction process. To illustrate this, we also show in parentheses the considerably improved values of $\Delta\Delta G_{gas\rightarrow sol}^{B}$ we obtained by switching in the transition state from the Lennard-Jones parameters of

Table 14. Water Solvent Effect on the Difference of the Free Energy of the Protonated Cytosine Tautomers with Respect to the pCYT13 Tautomer, $\Delta\Delta G_{gas\rightarrow sol}^{pCYT13\rightarrow t}$, and Partial Charge on the Oxygen Atom of the Protonated Cytosine

	$\Delta\Delta G_{gas\rightarrow sol}^{pCYT13\rightarrow t}$ (kcal/mol)			oxygen charge (atomic units)	
tautomer *t*	COSMO	3D-RISM-KH	Colominas[a]	COSMO[b]	3D-RISM-KH[b]
pCYT 13	0.00	0.00	0.0	−0.72	−0.65
pCYT 12c	10.01	6.70	11.1	−0.56	−0.53

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54] [b] Obtained from the Mulliken charge analysis.

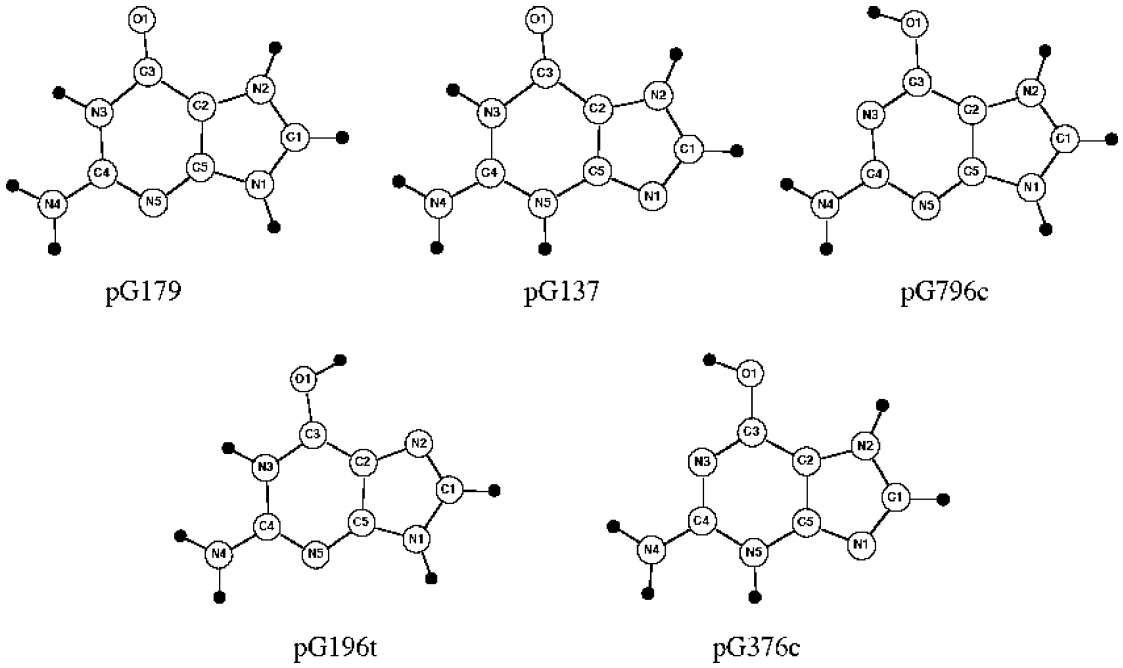

Figure 7. Protonated guanine tautomers.

Table 15. Free Energy (kcal/mol) of the Protonated Forms of Guanine in the Gas Phase with Respect to the pG179 Tautomer

tautomer	this work	Colominas
pG179	0.00	0.0
pG137	4.34	3.5[a]/4.8[b]
pG796c	4.84	3.6[a]/3.6[b]
pG196t	6.59	5.1[a]/5.9[b]
pG376c	3.35	1.8[a]/1.9[b]

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54] [b] B3LYP(6-311++G(d,p))//MP2/6-31G(d).[54]

Table 16. Free Energy (kcal/mol) of the Protonated Forms of Guanine in Aqueous Solution with Respect to the pG179 Tautomer

tautomer	COSMO	3D-RISM-KH	Colominas[a]
pG179	0.00	0.00	0.0
pG137	0.18	2.49	1.1
pG796c	11.95	8.98	10.7
pG196t	12.16	11.94	12.7
pG376c	10.02	9.25	9.8

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54]

the Br^- ion to those of the neutral Br atom. This outlines a possible way to further improve the description by interpolating the Lennard-Jones parameters in an intermediate state by projecting onto the initial and final states.

4.3.3. Water Distribution inside Carbon Nanotubes. Important effects can be revealed by exploring the solvent confined in inner spaces of nanomaterials.[96] In particular, introducing guest species into the inner phase of a carbon nanotube may give rise to altered composite system properties through spontaneous innerphase charge transfer and electrostatic interactions.[97] The energetics, mechanism, and dynamics of chemical reactions may be significantly altered inside carbon nanotubes because of their large electronic polarizabilities and because of the severely decreased reaction volume.[98] Aqueous solution confined in such nanosystems inside a molecular size volume shows a number of unique properties and processes which can be qualitatively different from the regular bulk environment. The prediction of these properties from the first principles is of significant interest since confined water can be found in many important systems, for example, in pockets of proteins, in channels of various biological and synthetic organic membranes, and inside various types of nanotubes.

Theoretical investigation of a solution confined in nanospaces is complicated by the fact that it cannot be treated by continuum solvation models because the concept of surface charge distribution in general loses meaning in this case.[7] So far, molecular dynamics and Monte Carlo simulations have been widely used for studying water in a nanospace environment.[99−101] These methods provide a realistic and accurate picture but are very expensive com-

Table 17. Water Solvent Effect on the Difference of the Free Energy of the Tautomers of Protonated Guanine with Respect to the G179 Tautomer, $\Delta\Delta G_{gas\rightarrow sol}^{pG179\rightarrow t}$, and Partial Charge on the Oxygen Atom of the Protonated Guanine

	$\Delta\Delta G_{gas\rightarrow sol}^{pG179\rightarrow t}$ (kcal/mol)			oxygen charge (atomic units)	
tautomer	COSMO	3D-RISM-KH	Colominas[a]	COSMO[b]	3D-RISM-KH[b]
pG179	0.00	0.00	0.0	−0.75	−0.68
pG137	−4.16	−2.39	−2.5	−0.70	−0.65
pG796c	7.11	4.56	7.5	−0.58	−0.54
pG196t	5.58	4.42	8.0	−0.55	−0.51
pG376c	6.66	5.93	7.1	−0.56	−0.52

[a] MP4/6-311++G(d,p)//MP2/6-31G(d).[54] [b] Obtained from the Mulliken charge analysis.

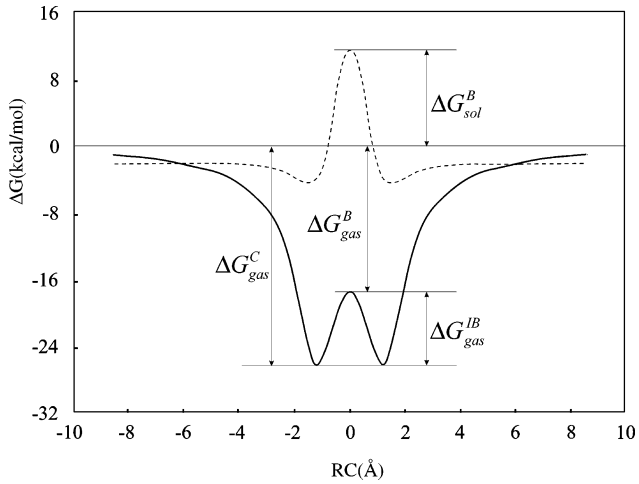

Figure 8. Free energy profiles (FEPs) for the identity S_N2 reaction $F^- + CH_3F$ in the gas phase (solid line) and in solution (dashed line). The energy parameters defining the FEP in the gas phase and in aqueous solution: the complexation free energy ΔG^C, the overall free energy barrier ΔG^B, and the intrinsic free energy barrier ΔG^{IB}.

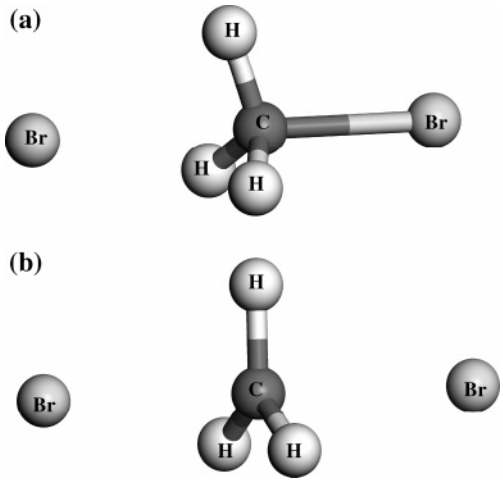

Figure 9. Ion-complex (a) and transition state (b) geometries for the $Br^- + BrCH_3 \rightarrow BrCH_3 + Br^-$ S_N2 reaction.

putationally. As an advantage, the KS-DFT/3D-RISM-KH theory readily yields a self-consistent field description with the relevant part of the nanosystem treated at the quantum chemical level and the properties of the confined solvent predicted from the first principles of statistical mechanics by the molecular theory of solvation.

Table 18. Difference $\Delta\Delta G_{gas\rightarrow sol}^{B}$ (kcal/mol) between the Free Energy of the Transition State of the Identity S_N2 Reactions in Aqueous Solution and in the Gas Phase, eq 19[a]

X = Y	COSMO[b]	COSMO[c]	3D-RISM-KH	experiment[c]
F	21.13	26.46	18.74	22.11
Cl	12.80	18.69	10.67	13.74
Br	13.14	16.61	16.18	14.97

[a] Theoretical predictions (calculated with the GGA functional) versus experimental results. The experimental values in the gas phase of the intrinsic barrier correspond to enthalpies. These are corrected with the calculated entropies to obtain free energies.[87−91] [b] With the translational, rotational, and vibrational thermodynamic corrections given by eq 12. [c] Without the translational, rotational, and vibrational thermodynamic corrections.

Table 19. Difference $\Delta\Delta G_{gas\rightarrow sol}^{B}$ (kcal/mol) Between the Free Energy of the Transition State of the Nonidentity S_N2 Reactions in Aqueous Solution and in the Gas Phase[a]

X, Y	COSMO[b]	COSMO[c]	3D-RISM-KH	experiment
F, Cl	31.24	36.54	35.27	31.71
F, Br	38.29	43.77	32.60	32.36
Cl, Br	26.86	32.66	7,6 (10.2)	15.95/23.80

[a] Theoretical predictions (calculated with the LDA functional) versus experimental results. The experimental values of the intrinsic barrier in the gas phase are derived from the measured enthalpy corrected by the calculated entropy to obtain the free energy.[87−91] [b] With the translational, rotational, and vibrational thermodynamic corrections given by eq 12. [c] Without the translational, rotational, and vibrational thermodynamic corrections.

To illustrate the capabilities of the KS-DFT/3D-RISM-KH theory for a solvent confined in inner spaces, we applied it to the simplest example of such a system: (6,6), (8,8), (12,12), and (20,20) carbon single-wall nanotubes (SWNTs) of different lengths immersed in water. So far, such a description level has been achieved only in the ONIOM[102] and QM/MD[102] calculations, which are significantly more expensive computationally. The whole nanotube is treated by KS-DFT with the LDA functional and the DZP basis set, which allows for the nanotube polarization. The computational load can be further decreased by applying the quantum mechanical treatment just to the most important part of the system (for instance, reactants inside the nanotube) and modeling the electronic response of the nanotube by the classical polarizability of its atoms.

The systems we studied included the SWNTs of five different diameters: (6,6), (8,8), (10,10), (12,12), and (20,-20). The nanotube length of ∼14 Å was chosen to be large enough to eliminate the end effects on the density distribu-

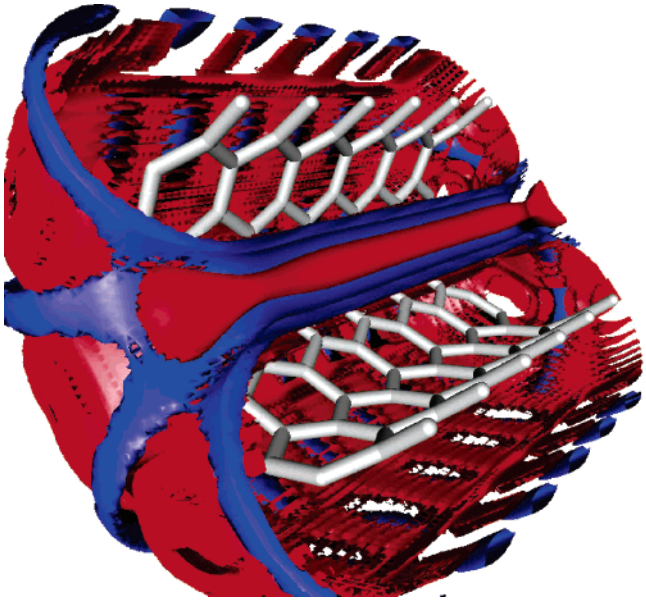

Figure 10. First hydration shell of the (6,6) single-wall carbon nanotube. Isosurfaces of the 3D distribution functions $g(\mathbf{r}) > 1.1$ of water oxygen (red) and hydrogen (blue) in the inner part as well as around the nanotube. Hidden is part of the nanotube and hydration shell.

tions in the central part the nanotube. Figure 10 presents the predictions of the KS-DFT/3D-RISM-KH theory for the 3D density distributions $g_\gamma(\mathbf{r})$ of water inside the (6,6) SWNT. The theory yields the hydration structure in 3D detail including the orientations of water molecules, both around the nanotube and in its inner part. Water molecules in contact with the outer surface of the nanotube are located mostly at the on-top and bridge-site positions of the surface lattice. Their hydrogens are oriented outward from the nanotube, with the HOH molecular plane at small angles to the nanotube axis. This alignment of water dipoles outward from the surface is caused by the strong polarization of the nanotube.

In Figure 11, we make a comparison between the predictions of the KS-DFT/3D-RISM-KH theory for the axial profiles of water sites (O = solid line, H = dashed line) inside the (6,6) SWNT and the MD simulations[99] (O = dash-dotted line, H = dotted line) available for this system. The 3D-RISM-KH results are in rather good agreement with the simulations and predict a number of features of the hydration structure. In the narrow channel of the (6,6) nanotube close to the water molecule size, water oxygens are almost centered at the nanotube axis, whereas water hydrogens are oriented mostly outward from the axis. As the OH intramolecular bond length is 1 Å, the hydrogen maximum at a distance of 0.6 Å from the channel axis suggests that water molecules are oriented with the dipole moment around normal to the axis and the HOH molecular plane at small angles to the axis. The relative position of the oxygen and hydrogen maxima is clearly seen also from the 1D axial distribution functions $g_\gamma(r)$ shown in Figure 11. In agreement with the MD simulations,[99] the theory predicts a low peak of the oxygen distribution at $r = 1.6$ Å for water molecules against the inner surface of the nanotube and a small peak of the hydrogen distribution at $r = 0$. The latter gives some

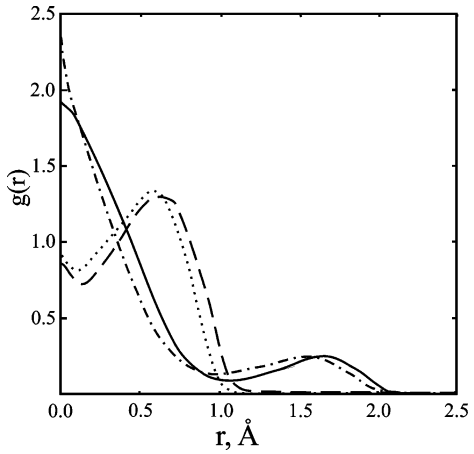

Figure 11. Water distribution profiles inside (6,6) SWNT as a function of the distance from the nanotube axis, predicted by the KS-DFT/3D-RISM-KH theory (O = solid line, H = dashed line) versus MD simulation[99] (O = dash-dotted line, H = dotted line).

probability of finding hydrogens around the channel axis, with a possibility to get in contact with oxygen cores of the adjacent water molecules and thus form hydrogen bonds. However, the integration $2\pi\rho_H \int_0^{r_1} r\, dr\, g_H(r)$ up to the distance $r_1 \approx 0.3$ Å estimating the region where hydrogens can be in contact with adjacent oxygens gives approximately just one hydrogen bond per 20 water molecules in the channel. We stress again that this level of description is comparable to molecular simulations and cannot be achieved by any continuum solvation model.

Further, the most probable orientations of a "labeled" water molecule for any given space region outside or inside the nanotube can be obtained explicitly in the form of the distribution function $g(\mathbf{r},\Omega)$ dependent on its position $\mathbf{r}$ and orientation Ω with respect to the nanotube by using the 3D-RISM approach to orientationally dependent potentials of mean force between solution molecules.[103]

Figure 12 presents the density profiles of water inside the nanotubes as a function of the distance from the nanotube center, obtained by azimuthal averaging of the 3D distribution functions $g_\gamma(\mathbf{r})$. In the (20,20) SWNT with the wide channel, the profile consists of four hydration shells. The height of the peaks is largest against the nanotube wall and decreases toward the nanotube center. All the peaks of water hydrogen are lower but wider than those of water oxygen, and the positions of the oxygen and hydrogen peaks are very close in the first and second hydration shells. This is typical for hydrophobic hydration with hydrogens aligned mainly along the hydrophobic surface but some assuming orientations toward and outward from it. This picture gets distorted for the narrower (12,12) SWNT, which holds about three hydration shells, the third one of the increasing height in the nanotube center. The even narrower (10,10) SWNT fits two hydration shells, the second one in the nanotube center elevated to the same height as that against the nanotube wall. Finally, the narrowest (8,8) SWNT accommodates just one hydration shell, with the complete void of water distribution in the nanotube center. The position of the hydrogen peak at $r = 1.7$ Å is shifted with respect to the oxygen one at r

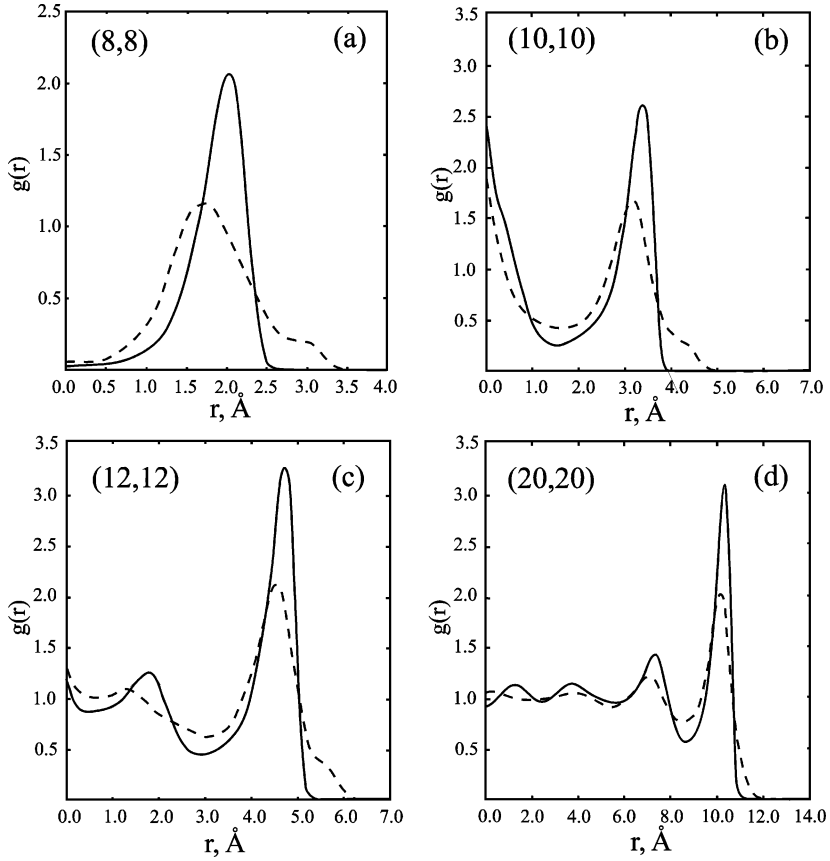

Figure 12. Water oxygen (solid line) and hydrogen (dashed line) distribution profiles inside SWNT as a function of the distance from the nanotube symmetry axis. Parts a−d: (8,8), (10,10), (12,12), and (20,20) SWNT, respectively.

= 2.0 Å, revealing the tilt of the OH intramolecular bonds by about 15−20° off the orientation along the nanotube wall toward the nanotube center. The considerable shoulder of the hydrogen profile at $r = 3.0$ Å corresponds apparently to the orientation of second hydrogens of such water molecules in the hydration shell in normal toward the nanotube wall.

Many theoretical investigations of water distributions inside carbon nanotubes are available,[99−102,104] and we shall not further discuss the electronic and classical effects in these systems presented in detail in the literature, such as the polarization of solvated nanotubes of different diameters. Appropriate for the goals of this article is to demonstrate that the KS-DFT/3D-RISM-KH theory reproduces the results of molecular simulations with a high level of accuracy, which is evident from Figure 11, making a comparison with MD simulation[99] for the hydration structure of (6,6) SWNT. This renders the 3D-RISM method self-consistently coupled with DFT or multireference methods as a valuable tool for the prediction of the properties of inner phases in nanomaterials, including chemical reactions in solution confined in inner spaces.

Conclusion

To validate the newly implemented KS-DFT/3D-RISM-KH theory by comparison with experimental results and other solvation approaches, we applied it to predict the effect of various solvents on the conformational equilibrium of 1,2-dichloroethane; the relative stability of some tautomers of cytosine, isocytosine, and guanine molecules in aqueous solution; and the solvation effect on the activation free energy of some S_N2 reactions. The results are quite similar to those obtained by high-level ab initio methods and continuum solvation approaches such as COSMO, when such calculations are available. Some differences have been detected in a few cases. We found that in general the KS-DFT/3D-RISM-KH method well-reproduces the free energy and predicts the solvation effects in agreement with experimental results.

We currently use the Lennard-Jones parametrization to represent the exchange and dispersive part of the classical potential between the atomic sites of solute and solvent molecules. Because these parameters considerably affect the results, it is crucial to take them from an appropriate force field optimized for liquids of the corresponding molecules rather than just for individual molecules in aqueous solvent. We took the Lennard-Jones parameters of both solute and solvent interaction sites, as well as the site charges of solvent molecules from the OPLS force field. Alternatively, these parameters can be determined from a post Hartree−Fock procedure[105] which is relatively inexpensive computationally and can be included in a self-consistent calculation. For chemical reactions, the Lennard-Jones parameters in an intermediate state can be interpolated between the initial and final states.

We found that the KS-DFT/3D-RISM-KH theory well-reproduces the conformational preferences of 1,2-dichloroethane in different solvents observed in experiments. It also predicts well the relative stabilities of several tautomers of

cytosine, isocytosine, and guanine molecules, determined experimentally. Moreover, KS-DFT/3D-RISM-KH has been the only method able to reproduce the experimental findings for the relative stability of the isocytosine tautomers in aqueous solution.

Most of the results obtained in the calculation of the solvent effects for the identity and nonidentity S_N2 reactions are in agreement with the latest experimental results. We obtained only one considerable discrepancy, which might be eliminated by parametrizing the Lennard-Jones parameters in intermediate states, as described above.

In general, the computational cost of the 3D-RISM-KH method is comparable with COSMO. While slower for small molecules, 3D-RISM-KH becomes more efficient in the case of macromolecules with hundreds of atoms.[3] The most computationally expensive part of 3D-RISM is the 3D fast Fourier transform, which scales as $N \log N$ with the number of 3D grid nodes N. The latter is roughly proportional to the number of atoms in the solute molecule and stays within reasonable size even for such large macromolecules as nanotubes. On the other hand, the scaling of COSMO is determined by the inversion of the matrix of size $M \times M$, where the number of nodes M covering the molecular surface with screening charge is proportional to the number of atoms in the solute. In the general case, the latter is the operation scaling as M^3. Therefore, for large macromolecules, COSMO gets increasingly inefficient and even unusable in converging the SCF loop.[106] Further linear-scaling modifications of COSMO are based on linear approximations in the calculation of the gradients.[106] However, a profound drawback of such methods is that the linear approximations shift the minimum of the potential energy surface in an uncontrolled way. This disadvantage of linear approximations added to conjugate gradient methods is well-known in numerical methods.[107] There is no guarantee in a general case that such linear-scaling modifications to COSMO reach the global minimum, and they should be used with considerable caution. As distinct, the 3D-RISM is precise in that respect and can be applied without reservation to *any* large system.

The 3D-RISM-KH theory presents principal advantages, compared to empirical continuum models such as COSMO, and is free from the principal limitations of post-COSMO thermodynamic methods like COSMO-RS. The 3D-RISM-KH theory has a predictive capability of a first-principle statistical-mechanical method. It can treat a given molecular solvent or mixture in a given thermodynamic state and is transferable. The theory gives the full physical view of molecular solvation: the 3D structure of the successive solvation shells, including association structures such as hydrogen bonds; the solvation free energy and all its thermodynamic derivatives such as entropy, partial molar volume, and so forth; the potentials of mean force between all species in solution, including the medium-induced effective interaction between a ligand and a solute supra- or macromolecule. The theory provides from the first principles the analytical gradients including all the solvation contributions, both electrostatic and nonelectrostatic (such as cavitation, dispersion, and repulsion). This allows one to accurately determine the potential energy surface and optimize molecular geometry in solution. The theory yields the solvent effect on the transition states, which enables prediction of the chemical reactions in solution. We emphasize that a crucial advantage of the KS-DFT/3D-RISM-KH method is a first-principle, physical view on electronic structure in solution. With reasonable computational efforts, it allows one to predict the electronic structure and geometry of solvated macromolecules, their thermochemistry, and chemical reactions in solution, including the effect of solvent in inner spaces of nanosystems such as channels of nanotubes. This renders it an attractive alternative to continuum solvation methods in theoretical chemistry.

Acknowledgment. This work was supported by the National Research Council (NRC) of Canada and by the Natural Science and Engineering Research Council (NSERC) of Canada. D.C. thanks the Generalitat de Catalunya for a scholarship. T.Z. thanks the Canadian Government for a Canada Research Chair.

References

(1) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **1999**, *110*, 10095.

(2) Kovalenko, A. Three-Dimensional RISM Theory for Molecular Liquids and Solid−Liquid Interfaces. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp 169−275.

(3) Gusarov, S.; Ziegler, T.; Kovalenko, A. *J. Phys. Chem. A* **2006**, *110*, 6083.

(4) Te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca, C. G.; Van Gisbergen, S. J.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931.

(5) Klamt, A.; Shüürmann, G. *J. Chem. Soc. Trans.* **1993**, *2*, 799.

(6) Pye, C. C.; Ziegler, T. *Theor. Chem. Acc.* **1999**, *101*, 396.

(7) Britz, D.; Khlobystov, A. *Chem. Soc. Rev.* **2006**, *35*, 637.

(8) Klamt, A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*; Elsevier: New York, 2005.

(9) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224.

(10) Mullins, E.; Oldland, R.; Liu, Y.; Wang, S.; Sandler, S.; Chen, C.; Zwolak, M.; Seavey, K. *Ind. Eng. Chem. Res.* **2006**, *45*, 4389.

(11) Imai, T.; Nomura, H.; Kinoshita, M.; Hirata, F. *J. Phys. Chem.* **2002**, *106*, 7308.

(12) Moralez, J. G.; Raez, J.; Yamazaki, T.; Motkuri, R. K.; Kovalenko, A.; Fenniri, H. *J. Am. Chem. Soc.* **2005**, *127*, 8307.

(13) Eckert, F.; Klamt, A. *J. Comput. Chem.* **2006**, *27*, 11.

(14) Gregersen, B.; Khandogin, J.; Thiel, W.; York, D. *J. Phys. Chem. B* **2005**, *109*, 9810.

(15) Hansen, J. P.; McDonald, I. R. *Theory of Simple Liquids*; Academic: London, 1986.

(16) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 193.

(17) Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 7821.

(18) Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **1998**, *290*, 237.

(19) Chandler, D.; McCoy, J.; Singer, J. *J. Chem. Phys.* **1996**, *85*, 5971.

(20) Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **2001**, *349*, 496.

(21) Kovalenko, A.; Hirata, F. *J. Theor. Comput. Chem.* **2002**, *1*, 381.

(22) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10391.

(23) Perkyns, J. S.; Pettitt, B. M. *J. Chem. Phys.* **1992**, *97*, 7656.

(24) Perkyns, J. S.; Pettitt, B. M. *Chem. Phys. Lett.* **1992**, *190*, 626.

(25) Singer, S. J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

(26) Sato, H.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1996**, *105*, 1546.

(27) Yoshida, N.; Hirata, F. *J. Comput. Chem.* **2006**, *27*, 453.

(28) McQuarrie, D. A. *Statistical Thermodynamics*; University Science Books: New York, 1973.

(29) Wertz, D. H. *J. Am. Chem. Soc.* **1980**, *102*, 5316.

(30) Cooper, J.; Ziegler, T. *Inorg. Chem.* **2002**, *41*, 6614.

(31) Yu, B.; Privalov, P.; Hodges, R. *Biophys. J.* **2001**, *81*, 1632.

(32) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(33) Perdew, J. P.; Wang, Y. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *46*, 6671.

(34) Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.

(35) Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M. *J. Phys. Chem.* **1995**, *99*, 9072.

(36) Capelli, C.; Corni, S.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 10807−10815.

(37) Christiansen, O.; Mikkelsen, K. V. *J. Chem. Phys.* **1999**, *110*, 1365.

(38) Lee, J.; Yoshida, N.; Hirata, F. *J. Phys. Chem. B* **2006**, *110*, 16018.

(39) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.

(40) Topal, M.; Fresco, J. R. *Nature* **1976**, *260*, 285.

(41) Kwiatkowski, J. S.; Pullman, B. *Adv. Heterocycl. Chem.* **1975**, *18*, 199.

(42) Gould, I. R.; Vincent, M. A.; Hiller, I. H.; Lapinski, L.; Nowak, M. J. *Spectrochim. Acta, Part A* **1992**, *48*, 811.

(43) Jaworski, A.; Szczepaniak, M.; Kubulat, K.; Person, W. B. *J. Mol. Struct.* **1990**, *63*, 223.

(44) Nowak, M. J.; Lapinski, L.; Fulara, J. *Spectrochim. Acta, Part A* **1989**, *45*, 229.

(45) Radchenko, E. D.; Sheina, G. G.; Smorygo, N.; Blagoi, Y. P. *J. Mol. Struct.* **1984**, *116*, 387.

(46) Brown, D. J.; Teitei, T. *Aust. J. Chem.* **1965**, *18*, 559.

(47) Helene, C.; Dozou, P. *CR Acad. Sci. (Paris)* **1964**, *259*, 4387.

(48) Kaito, A.; Hatano, M.; Ueda, T.; Shibuya, S. *Bull. Chem. Soc. Jpn.* **1980**, *53*, 3073.

(49) Lowe, P. R.; Schwalbe, C. H.; Williams, G. J. B. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1987**, *43*, 330.

(50) Morita, H.; Nagakura, S. *Theor. Chim. Acta* **1968**, *11*, 279.

(51) Sharma, B. D.; McConnel, J. F. *Acta Crystallogr.* **1965**, *19*, 797.

(52) Kobayashi, R. *J. Phys. Chem. A* **1998**, *102*, 10813.

(53) Les, A.; Adamowicz, L.; Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 4001.

(54) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811−6821.

(55) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1987**, *111*, 2308.

(56) Gorb, L.; Podolyan, Y.; Leszczynski, J. *THEOCHEM* **1999**, *487*, 47.

(57) Podolyan, Y.; Gorb, L.; Leszczunski, J. *Int. J. Mol. Sci.* **2003**, *4*, 410−421.

(58) Saenger, W. *Principles of Nucleic Acid Structure;* Springer-Verlag: New York, 1983.

(59) Burchenal, J. H.; Ciovacco, K.; Kalaher, K.; O'Toole, T.; Dowling, M. H.; Chu, C. K.; Watanabe, K. A.; Fox, J. J. *Cancer Res.* **1982**, *36*, 1520.

(60) Roberts, C.; Bandaru, R.; Switzer, C. Y. *Tetrahedron Lett.* **1995**, *36*, 3601.

(61) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322.

(62) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489.

(63) Bain, J. D.; Switzer, C. Y.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537.

(64) Strobel, S. A.; Cech, T. R.; Usman, N.; Beigelman, L. *Biochemistry* **1994**, *33*, 13824.

(65) Horn, T.; Chang, C. A.; Collins, M. L. *Tetrahedron Lett.* **1995**, *36*, 2033.

(66) Tor, Y.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 4461.

(67) Musier-Forsyth, K.; Shi, J.-P.; Henderson, B.; Bald, R.; Furste, J. P.; Erdmann, V. A.; Schimmel, P. *J. Am. Chem. Soc.* **1995**, *117*, 7253.

(68) Sheina, G. G.; Stepanian, S. G.; Radchenko, E. D.; Blagoi, Yu. P. *J. Mol. Struct.* **1987**, *158*, 275.

(69) Vranken, H.; Smets, J.; Maes, G.; Lapinski, L.; Nowak, M. J.; Adamowicz, L. *Spectrochim. Acta, Part A* **1994**, *50*, 875.

(70) Szczesniak, M. Ph.D. Thesis, University of Warsaw: Warsaw, Poland, 1985.

(71) Szczesniak, M.; Nowak, M. J.; Szczesniak, K. *J. Mol. Struct.* **1984**, *115*, 221.

(72) Stepanian, S. G.; Radchenko, E. D.; Sheina, G. G.; Blagoi, Yu. P. *J. Mol. Struct.* **1990**, *216*, 77.

(73) Kwiatkowski, J. S.; Leszczunski, J. *J. Phys. Chem.* **1996**, *100*, 941.

(74) Beak, P. *Acc. Chem. Res.* **1997**, *10*, 186.

(75) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.

(76) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

(77) Shi, Z.; Boyd, R. J. *J. Am. Chem. Soc.* **1990**, *112*, 6789.

(78) Tucker, S. C.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 8138.

(79) Vetter, R.; Zülicke, L. *J. Am. Chem. Soc.* **1990**, *112*, 5136.

(80) Cernusák, I.; Urban, M. *Collect. Czech. Chem. Commun.* **1988**, *53*, 2239.

(81) Dedieu, A.; Veillard, A. *J. Am. Chem. Soc.* **1972**, *94*, 6730.

(82) Dedieu, A.; Veillard, A. *Chem. Phys. Lett.* **1970**, *5*, 328.

(83) Sato, H.; Sakaki, S. *J. Phys. Chem. A* **2004**, *108*, 1629−1634.

(84) Yang, S.-Y.; Fleurat-Lessart, P.; Hristov, I.; Ziegler, T. *J. Phys. Chem. A* **2004**, *108*, 9461.

(85) Alemán, C.; Maseras, F.; Lledós, A.; Duran, M.; Bertrán, J. *J. Phys. Chem.* **1989**, *2*, 611.

(86) Gonzales, J. M.; Pak, C.; Cox, R. S.; Allen, W. D.; Shaefer, H. F., III; Császár, A. G.; Tarczay, G. *Chem.—Eur. J.* **2003**, *9*, 2173.

(87) Parthiban, S.; de Oliveira, G.; Martin, J. M. L. *J. Phys. Chem. A* **2001**, *105*, 895−904.

(88) Shaik, S. S.; Schlegel, H. B.; Wolfe, S. *Theoretical Aspects of Physical Organic Chemistry: The SN2 Mechanism*; John Wiley & Sons: New York, 1992.

(89) Bathgate, R. H.; Moelwyn-Huches, E. A. *J. Chem. Soc.* **1959**, 2642.

(90) McLennan, D. J. *Aust. J. Chem.* **1978**, *31*, 1897.

(91) Wladkowski, B. D.; Brauman, J. I. *J. Phys. Chem.* **1993**, *97*, 13158.

(92) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154.

(93) Pellerite, M. J.; Brauman, J. I. *J. Am. Chem. Soc.* **1983**, *105*, 2672.

(94) Wilbur, J. L.; Wladkowski, B. D.; Brauman, J. I. *J. Am. Chem. Soc.* **1993**, *115*, 10823.

(95) Glukhovtsev, M. N.; Pross, A.; Schlegel, H. B.; Bach, R. D.; Random, L. *J. Am. Chem. Soc.* **1996**, *118*, 11258.

(96) Lu, G. Q.; Zhao, X. S. *Nanoporous Materials: Science and Engineering*; Imperial College Press: London, 2005.

(97) Knox, J. E.; Halls, M. D.; Schlegel, H. B. *J. Comput. Nano. Sci.* **2006**, *3*, 398.

(98) Halls, M. D.; Schlegel, H. B. *J. Phys. Chem. B* **2002**, *106*, 1921.

(99) Gordillo, M. C.; Marti, J. *Chem. Phys. Lett.* **2000**, *329*, 341.

(100) Hanasaki, I.; Nakatani, A. *J. Chem. Phys.* **2006**, *124*, 174714.

(101) Moulin, F.; Devel, M.; Picaud, S. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *71*, 165401.

(102) Xu, S.; Irle, S.; Musaev, D. G.; Lin, M. C. *J. Phys. Chem. A* **2005**, *109*, 9563.

(103) Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **1999**, *103*, 7942.

(104) Grujicic, M. *Appl. Surf. Sci.* **2003**, *206*, 167.

(105) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.

(106) York, D. M.; Lee, T.-S.; Yang, W. *Chem. Phys. Lett.* **1996**, *263*, 297.

(107) Press, W. H.; Teukolsky, A. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes*; Cambridge University Press: New York, 1992.

CT6001785