Smooth Solvation Method for d-Orbital Semiempirical Calculations of Biological Reactions. 2. Application to Transphosphorylation Thio Effects in Solution

Brent A. Gregersen,[†] Jana Khandogin,[†] Walter Thiel,[‡] and Darrin M. York*,[†]

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, and Max-Planck-Institut für Kohlenforschung, 45466 Mülheim an der Ruhr, Germany

Received: December 31, 2004; In Final Form: March 8, 2005

Density-functional and semiempirical quantum methods and continuum dielectric and explicit solvation models are applied to study the role of solvation on the stabilization of native and thio-substituted transphosphorylation reactions. Extensive comparison is made between results obtained from the different methods. For the semiempirical methods, explicit solvation was treated using a hybrid quantum mechanical/molecular mechanical (QM/MM) approach and the implicit solvation was treated using a recently developed smooth solvation model implemented into a d-orbital semiempirical framework (MNDO/d-SCOSMO) within CHARMM. The different quantum and solvation methods were applied to the transesterification of a 3′-ribose,5′-methyl phosphodiester that serves as a nonenzymatic model for the self-cleavage reaction catalyzed by the hammerhead and hairpin ribozymes. Thio effects were studied for a double sulfur substitution at the nonbridging phosphoryl oxygen positions. The reaction profiles of both the native and double sulfur-substituted reactions from the MNDO/d-SCOSMO calculations were similar to the QM/MM results and consistent with the experimentally observed trends. These results underscore the need for a d-orbital semiempirical representation for phosphorus and sulfur for the study of experimentally observed thio effects in enzymatic and nonenzymatic phosphoryl transfer reactions. One of the major advantages of the present approach is that it can be applied to model chemical reactions at a significantly lower computational cost than either the density-functional calculations with implicit solvation or the semiempirical QM/MM simulations with explicit solvent.

1. Introduction

There is growing interest in the development of new semiempirical and implicit solvation methods that can be used synchronously with molecular simulation and linear-scaling methods to address complicated biological problems. These integrated "multi-scale quantum models" are able to simultaneously span a broad range of spatial and temporal domains. Certain solvation methods, particularly those of the "boundary element" type,[1] are among the most widely applied in electronic structure calculations[2,3] and have been traditionally problematic in the rigorous development and implementation of smooth analytic gradients. Many reports of analytic gradients for these methods ignore terms associated with surface discretization which leads to severe problems in energy minima and transition state geometry optimization procedures and molecular dynamics simulations. Recently, a smooth COSMO method[4] has been introduced that circumvents these problems through the use of Gaussian functions for expansion of the reaction field surface charges and a differentiable switching layer that allows new surface elements to appear and disappear smoothly with respect to the energy. Incorporated into electronic structure or molecular mechanics methods, this method may provide a powerful tool to study solution-phase molecular properties, chemical reactions, and molecular dynamics.

Solvation is particularly important in phosphoryl transfer reactions that involve the association or dissociation of anionic species.[5−8] In addition to stabilizing the electronic structure of polyanions that are often unstable in the gas phase, solvation generally lowers the energy barrier of associating anions by reducing the effective Coulomb repulsion as they approach and also preferentially stabilizes transition state complexes. Despite a tremendous amount of experimental and theoretical study, the nonenzymatic and enzymatic chemical mechanism of phosphate hydrolysis reactions remains a topic of discussion and considerable debate.[9−13] Consequently, it is important to ascertain the reliability of new and existing quantum and solvation models in order to design improved multi-scale quantum models for these reactions.

The present work reports the application of a smooth solvation method to phosphoryl transfer reactions in solution. The method has been implemented with analytic gradients[14] within a d-orbital semiempirical framework[15,16] and integrated with CHARMM.[17] The focus of this paper is (1) to characterize the MNDO/d-smooth COSMO (MNDO/d-SCOSMO) solvation method in the description of potential energy surfaces (PESs) of phosphoryl transfer reactions and (2) to compare the results from the present MNDO/d-SCOSMO method with other quantum and solvation models in order to identify areas where improvements can be made. One of the key results that arise from this work is that the use of the present MNDO/d-SCOSMO model may afford an avenue for efficient parametrization of QM/MM van der Waals radii that are key to obtain reliable free energy profiles of reactions that involve ionic species.[18] The paper is organized as follows: the following (second) section outlines the computational methods, including the MNDO/d-SCOSMO, density-functional/implicit solvation and hybrid QM/MM models. The

* Corresponding author: Phone: (612) 624-8042. Fax: (612) 626-7541. E-mail: york@chem.umn.edu. Web site: http://riesling.chem.umn.edu.
[†] University of Minnesota.
[‡] Max-Planck-Institut für Kohlenforschung.

10.1021/jp044061l CCC: $30.25
Published on Web 04/20/2005

third section describes application to transphosphorylation thio effects in a 3′-ribose,5′-methyl phosphodiester model. The fourth section concludes with a summary of the main results and outline of future research directions.

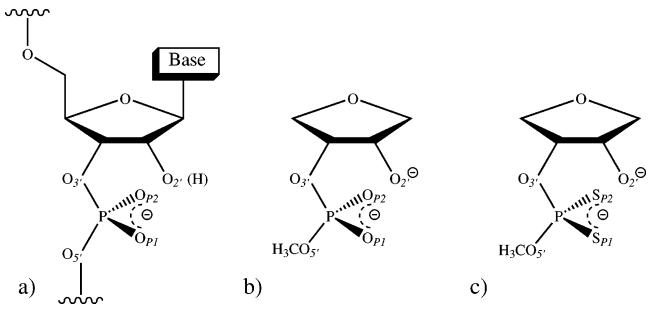
2. Methods

The focus of the present work is the application of the MNDO/d-SCOSMO method[14] to transphosphorylation thio effects of a 3′-ribose,5′-methyl phosphodiester model. Comparisons are made with results of density-functional theory (DFT) calculations with implicit solvation corrections, and hybrid QM/MM simulations with the same d-orbital semiempirical quantum model. The MNDO/d-SCOSMO method has been implemented into the MNDO code[19] and interfaced to the CHARMM molecular modeling package[17] and will be available in future MNDO and CHARMM releases. The use of d orbitals has been shown previously[16] to be important for pentavalent phosphorus and is particularly important for phosphorus and sulfur in the present reactions that proceed through dianionic phosphorane transition states and intermediate. Semiempirical methods without d orbitals such as AM1 and PM3 produce artificially stable dianionic phosphorane intermediates in the gas phase (compared with density-functional calculations) and unrealistic energy profiles (see Supporting Information Figure S1 and Table S1 for comparison).

2.1. Semiempirical Calculations. All semiempirical/implicit solvation calculations were performed using the combined MNDO/d Hamiltonian[16] and smooth COSMO method[4] implemented in MNDO97.[19] Details of the method, implementation, and testing have been presented elsewhere.[14] Convergence criteria for the SCF energy were 10^{-6} eV, and for geometry optimizations, they were 1.0 kcal/mol/ Å on the gradient norm. Solvation calculations for the 3′-ribose,5′-methyl phosphodiester model were performed using van der Waals radii from the CHARMM c27 force field for nucleic acids[20] with a 0.8 Å probe radius. The probe radius was optimized to obtain a best fit to the QM/MM reaction free energy profile[8] that used the same MNDO/d quantum model and CHARMM c27 van der Waals radii. Unless otherwise stated, the solvent accessible surface used a discretization level of 110 points per sphere, smooth COSMO switching parameter $\gamma_s = 1.0$ and shift parameter $\alpha \approx 0.5$ as described in previous work.[14] The nonpolar (nonelectrostatic) component of the solvation free energy (E_{npol} in eq 22 of ref 14), which has a very small effect on PESs of reactions that involve dianionic species, was neglected as in previous work.[14]

2.2. Hybrid QM/MM Calculations. Hybrid QM/MM calculations were performed on the transphosphorylation of a native and doubly sulfur substituted 3′-ribose,5′-methyl phosphodiester model system (Scheme 1) using the MNDO/d Hamiltonian.[16] Activated dynamics calculations were performed on the model system solvated in a 20 Å sphere of 1076 TIP3P water molecules[21] with a 4.0 Å stochastic buffer zone.[22,23] To prevent excessive drift of the solute, a weak harmonic restraint (1 kcal/mol/atom) on the center of mass of the solute was introduced. The reaction zone was propagated with Hamiltonian dynamics, and water molecules in the buffer zone were propagated with Langevin dynamics ($\tau = 62.0$ ps^{-1}) to serve as a constant-temperature bath (300 K), and a deformable boundary potential was used to reproduce the containment effect of bulk solvent outside the buffer zone.[23] Simulations used the SHAKE algorithm[24] to constrain the internal geometry of the water molecules and were performed with a 0.5 fs integration time step without cutoffs for electrostatic and van der Waals interactions. Umbrella sampling[25] was used to collect statistics

SCHEME 1: 3′-Ribose,5′-methyl phosphodiester used to model transphosphorylation thio effects in RNA: (a) an RNA nucleotide unit, (b) the native (un-substituted) 3′-ribose,5′-methyl phosphodiester model, and (c) the doubly thio-substituted) 3′-ribose, 5′-methyl phosphothioate model

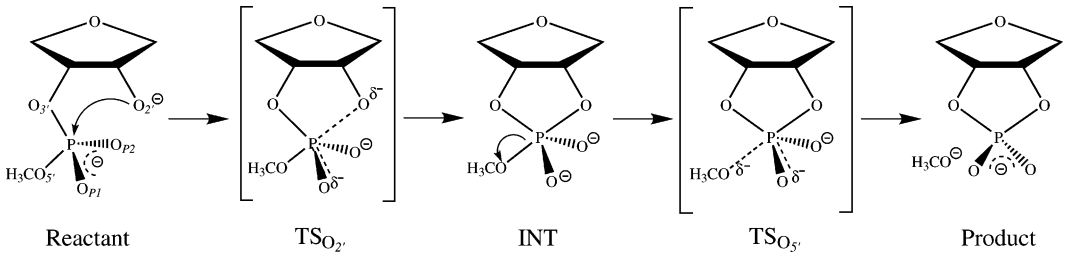


over the reaction coordinate range (-3.1 Å $\leq r_2 - r_1 \leq 3.7$ Å, where r_1 and r_2 are the P···$O_{2'}$ and P···$O_{5'}$ distances, respectively), and free energy profiles as a function of the $r_2 - r_1$ coordinate were calculated using the weighted histogram analysis method (WHAM)[26] with a 0.02 Å bin spacing. A detailed description of the calculations and discussion of the results can be found elsewhere.[8]

2.3. Density-Functional Calculations. Density-functional calculations were performed using the B3LYP exchange-correlation functional[27,28] with the 6-31++G(d,p) basis set for geometry and frequency calculations followed by single-point energy refinement with the 6-311++G(3df,2p) basis set in a manner analogous to recent studies of biological phosphates.[29−32] Energy minimum and transition state geometry optimizations were performed in redundant internal coordinates with default convergence criteria,[33] and stability conditions of the restricted closed shell Kohn−Sham determinant for each final structure were verified.[34,35] Frequency calculations were performed to establish the nature of all stationary points and to allow evaluation of thermodynamic quantities. The solvent was treated using the PCM solvation model,[2,36] with the CHARMM c27[20] van der Waals radii for consistency with the QM/MM and MNDO/d-SCOSMO calculations. Due to instabilities in the geometry optimization of the DFT calculations on the solvated potential energy surface, solvent corrections were taken into account by single-point calculations at the gas-phase-optimized B3LYP/6-31++G(d,p) geometries, as in previous work.[29−32] All density-functional calculations were performed with the Gaussian 03[37] suite of programs.

In the sections that follow, comparisons will be made between the MNDO/d-SCOSMO method, DFT with the PCM solvation method, and hybrid QM/MM calculations with MNDO/d and explicit solvation. As mentioned previously, the DFT calculations could not be performed on the solvated PES due to instabilities in the geometry optimization procedure, and hence solvation effects were included as single-point energy corrections at the gas-phase geometries. Indeed, this long-standing problem, prevalent in the vast majority of boundary element solvation methods, was a prime motivation for implementation of the smooth COSMO method into an electronic structure package. To examine the effect of solvent-induced structure relaxation, PESs and free energies in solution were derived from the MNDO/d-SCOSMO method in two ways: (1) with solvation treated as a single-point correction analogous to the DFT calculations with PCM solvation and (2) with optimization on the solvated PES. To make clear exactly what models have been used to derive structures and energies, the following notation has been introduced:

SCHEME 2: In-line dianionic mechanism for transphorylation of a 3′-ribose,5′-methyl phosphodiester as a model for RNA catalysis



DFT: denotes DFT energies and structures derived from the gas-phase PES.

DFT+PCM: denotes DFT energies with PCM solvation derived from single-point PCM solvation correction to the DFT gas phase-optimized structures.

MNDO/d: denotes MNDO/d energies and structures derived from the gas phase PES.

MNDO/d+SCOSMO: denotes MNDO/d energies with SCOSMO solvation derived from single-point SCOSMO solvation correction to the MNDO/d gas phase-optimized structures.

MNDO/d−SCOSMO: denotes MNDO/d energies with SCOSMO solvation derived from the solution phase-optimized structures.

QM/MM: denotes QM/MM free energies (potentials of mean force) and average structures dereived from hybrid QM/MM simulationa with explicit solvation.

3. Results and Discussion

In this section, a 3′-ribose,5′-methyl phosphodiester will be used as a model system to study transphosphorylation thio effects in RNA enzymes (Scheme 1). Reaction profiles for the native reaction and a chemically modified (doubly sulfur-substituted) reaction are compared using DFT+PCM, MNDO/d+SCOSMO, MNDO/d-SCOSMO, and hybrid QM/MM models (see the Methods section). A schematic of the in-line dianionic transphosphorylation mechanism is shown in Scheme 2. The two transition states corresponding to the formation of the endo-cyclic $P-O_{2'}$ bond and cleavage of the exo-cyclic $P-O_{5'}$ bond are denoted $TS_{O_{2'}}$ and $TS_{O_{5'}}$, respectively, whereas the phosphorane intermediate between the transition states is denoted INT.

Figure 1 illustrates the effect of solvation on the reaction energy profiles calculated with the MNDO/d-SCOSMO and MNDO/d for the native (unsubstituted) and the sulfur-substituted (on the nonbridging O_{P1} and O_{P2} positions) reactions of the dianionic 3′-ribose,5′-methyl phosphodiester. The MNDO/d gas-phase and MNDO/d-SCOSMO profiles were obtained by performing a constrained geometry optimization at successive reaction coordinate values. The reaction coordinate was defined as the difference between the phosphorus-leaving group (r_2 = $P-O_{5'}$) and phosphorus−nucleophile (r_1 = $P-O_{2'}$) distances. Figure 2 compares the reaction energy profiles in solution calculated with the MNDO/d-SCOSMO method and derived from hybrid QM/MM simulation with MNDO/d and explicit solvation.

Table 1 compares the key structural parameters for stationary points along the reaction PES in the gas-phase calculated with DFT and MNDO/d and for stationary points along the PES in solution calculated with MNDO/d-SCOSMO and QM/MM models. Table 2 compares the corresponding aqueous free energy values, along with the gas-phase and solvation free energy components. The ΔG_{gas} values contain zero-point

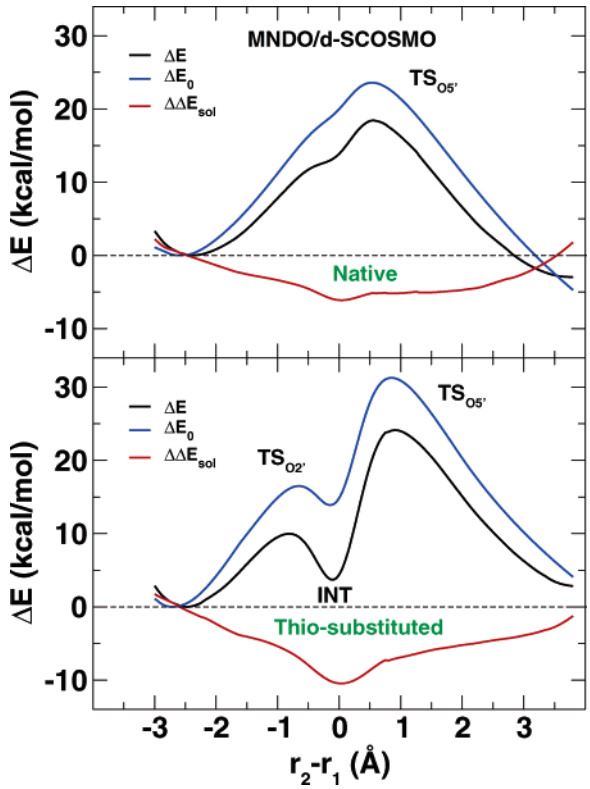


Figure 1. MNDO/d-SCOSMO and MNDO/d reaction profiles of the transphosphorylation of 3′-ribose,5′-methyl phosphodiester in the native (top) and doubly sulfur-substituted (bottom) forms (see Scheme 1). The reaction coordinate is defined to be $r_2 - r_1$ (see Scheme 2). Shown are the relaxed solution-phase energy profiles optimized with MNDO/d-SCOSMO (ΔE, black), the relaxed gas-phase energy profiles optimized with MNDO/d (ΔE_0, blue), and the solvation energy defined as the difference between the solution-phase and gas-phase curves ($\Delta\Delta E_{sol}$, red). All energies (kcal/mol) are with respect to the reactant.

energies and thermal corrections to the enthalpy and entropy, and the $\Delta\Delta G_{sol}$ values are defined as the difference between the aqueous-phase and gas-phase free energy values

$$\Delta\Delta G_{sol} = \Delta G_{aq} - \Delta G_{gas} \quad (1)$$

Figures 3−5 illustrate the structural differences of the rate-controlling transition states and reactive intermediates calculated with DFT and with MNDO/d-SCOSMO. In the case of the native reaction, no kinetically significant intermediate is present. The DFT structures were optimized in the gas phase, whereas the MNDO/d-SCOSMO structures include solvent-induced structural relaxation.

3.1. Effect of Solvation. This section discusses the effect of solvation by direct comparison of the PESs and free energy values calculated in the gas phase with MNDO/d and in solution

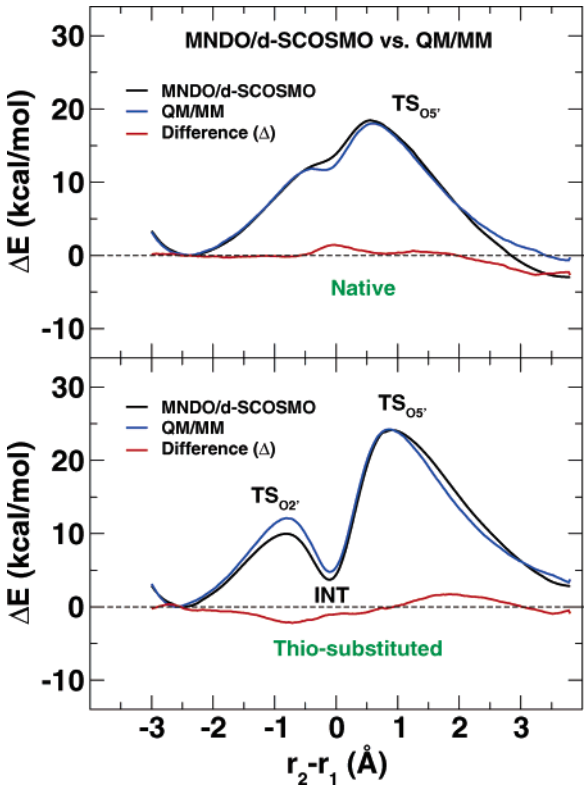


Figure 2. MNDO/d-SCOSMO and QM/MM reaction profiles of the transphosphorylation of 3′-ribose,5′-methyl phosphodiester in the native (top) and doubly sulfur-substituted (bottom) forms (see Scheme 1). The reaction coordinate is defined to be $r_2 - r_1$ (see Scheme 2). Shown are the relaxed solution-phase energy profiles optimized with MNDO/d-SCOSMO (black), the potential of mean force profile obtained from hybrid QM/MM simulation in explicit solvent (blue), and the difference between the two curves (red). All energies (kcal/mol) are with respect to the reactant.

calculated with MNDO/d-SCOSMO that includes solvent-induced structural relaxation.

3.1.1. Effect of Solvation on the Energetics of Reaction. Figure 1 shows the gas-phase PES and solution-phase free energy profiles for the native (top) and thio-substituted (bottom) transphosphorylation reactions. The gas-phase PES for the native reaction exhibits a single transition state ($TS_{O_5'}$) with a 23.6 kcal/mol activation energy barrier. If thermal corrections to the enthalpy and entropy are made at the transition state, the relative free energy in the gas phase is 24.6 kcal/mol (Table 2). Upon solvation, the activation free energy barrier calculated with the MNDO/d-SCOSMO method for the native reaction is reduced to 18.3 kcal/mol. Solvation has an even greater stabilizing effect on the transition states and reactive intermediate for the sulfur-substituted reaction (Figure 1, bottom). The solvation energy is most stabilizing (−13.0 kcal/mol) for the dithiophosphorane intermediate (INT), and is least stabilizing for the rate-controlling $TS_{O_5'}$ transition state (−7.5 kcal/mol).

3.1.2. Effect of Solvation on the Structure of Reactive Species. The effect of solvation on the structure of the stationary points along the reaction path is significant in several cases. The most prevalent solvent-induced structural relaxation occurs for the product minima where the $r_1 = P-O_{2'}$ distance in the native reaction contracts by 0.2 Å from 4.30 Å in the gas phase to 4.10 Å in solution, and the $O_{2'}-P-O_{5'}$ angle (θ_{inl}) increases by 7.2° to be more poised for in-line attack. This trend in solvent-induced structural relaxation is amplified for the thio-substituted reactant where the $r_1 = P-O_{2'}$ distance decreases upon solvation by 0.26 Å and the θ_{inl} value increases by 11.5°. Solvation has the effect of screening the Coulomb repulsions between the 2′ alkoxide and the phosphate and allows them to come into closer proximity than would be favorable in the gas phase. Consequently, the reactant minima for the native and thio-substituted reactions undergo significant stabilization and structural relaxation upon solvation.

The dianionic transition state and intermediate complexes are preferentially stabilized by solvent relative to the separated product monoanions. The average $P-O_{2'}/P-O_{5'}$ bond lengths in the rate-controlling $TS_{O_5'}$ transition state contract upon solvation by 0.050 and 0.025 Å for the native and thio-substituted reactions, respectively (Table 1). The corresponding reaction coordinate values for the native and thio-substituted reactions slightly decreases from $r_2 - r_1 = 0.55$ and 0.91 Å, respectively, in the gas phase to $r_2 - r_1 = 0.53$ and 0.86 Å, respectively, in solution. The structures of the rate-controlling transition states thus become slightly more associative upon solvation in both cases with character shifted toward that of the reactant 3′-ribose,5′-methyl phosphodiester. The increase in associative character is more pronounced for the thio-substituted reaction, although the overall effect of solvation on the structure of the rate-controlling transition state is not very large in either reaction.

In summary, solvation has the general effect of preferentially stabilizing the dianionic transition states and reactive intermediates relative to the reactants and products. Solvation increases the associative character of the reactive species by contraction of the axial P−O bond lengths, as observed in other work,[38,29] and increase of the axial O−P−O angle (θ_{inl}) to be more in-line and poised for nucleophilic attack or leaving-group departure.

3.2. Comparison of MNDO/d-SCOSMO with Hybrid QM/MM Simulations. *3.2.1. Comparison of Aqueous Free Energy Profiles.* The free energy profile for the native and thio-substituted transphosphorylation reactions calculated from MNDO/d-SCOSMO and QM/MM simulations with MNDO/d and explicit water are shown in Figure 2. The QM/MM profile (blue curve, top) for the native reaction proceeds through two transition states, $TS_{O_2'}$ and $TS_{O_5'}$, separated by a kinetically insignificant intermediate with a 0.2 kcal/mol barrier to collapse back to the reactants. The free energy profile for the native reaction obtained from the MNDO/d-SCOSMO method does not resolve the reactive intermediate observed in the QM/MM profile but exhibits a broad shoulder with an inflection point in the vicinity of the QM/MM intermediate. The activation free energy for the QM/MM profile (18.0 kcal/mol) is very close to that of the MNDO/d-SCOSMO value (18.3 kcal/mol). The QM/MM profile for the thio-substituted reaction features two distinct transition states (Figure 2: blue curve, bottom) and is overall quite similar to that of the MNDO/d-SCOSMO profile. The largest deviation occurs for the first $TS_{O_2'}$ activation barrier that has a QM/MM value (12.1 kcal/mol) that is 1.8 kcal/mol higher than the MNDO/d-SCOSMO value (10.3 kcal/mol). The QM/MM values for the INT and rate-controlling $TS_{O_5'}$ (4.8 and 24.2 kcal/mol, respectively) are within 1 kcal/mol of the corresponding MNDO/d-SCOSMO values (4.0 and 24.5 kcal/mol, respectively).

3.2.2. Comparison of Structures in Solution. The geometrical parameters for the transition states and intermediates are all in very close agreement between the QM/MM and MNDO/d-SCOSMO methods (Table 1). The largest deviation (0.06 Å) occurs for the $TS_{O_5'}$ structure in the thio-substituted reaction that has the longest leaving group bond length ($r_2 = 2.63$ and 2.69 Å for QM/MM and MNDO/d-SCOSMO, respectively). The

TABLE 1: Comparison of the Relevant Distances[a] and Angles[a] in the Reactant, Transition State, Intermediate, and Product along the Reaction Path PES for the Transphosphorylation of 3′-Ribose,5′-methyl Phosphodiester in the Gas Phase and in Solution

	DFT			MNDO/d			MNDO/d-SCOSMO			QM/MM[b]		
molecule	r_1	r_2	θ_{inl}	r_1	r_2	θ_{inl}	r_1	r_2	θ_{inl}	r_1	r_2	θ_{inl}
native												
reactant	4.60	1.71	108.6	4.30	1.76	125.3	4.10	1.72	132.5	4.07	1.69	130.3
$TS_{O_5'}$	1.85	2.51	162.2	1.91	2.44	162.7	1.85	2.40	165.4	1.80	2.40	164.1
product	1.71			1.76			1.74					
sulfur-substituted												
reactant	4.57	1.68	98.4	4.44	1.73	114.1	4.18	1.69	125.6	4.26	1.67	118.9
$TS_{O_2'}$	2.43	1.73	163.9	2.44	1.78	158.9	2.54	1.73	163.0	2.49	1.71	163.5
INT	1.90	1.77	165.3	1.95	1.82	160.5	1.89	1.79	164.8	1.88	1.77	165.0
$TS_{O_5'}$	1.77	2.84	158.9	1.83	2.69	156.6	1.78	2.69	159.5	1.76	2.63	160.4
product	1.70			1.75			1.73					

[a] Shown are the $P-O_{2'}$ and $P-O_{5'}$ distances (r_1 and r_2, respectively) and the in-line $O_{2'}$-$P-O_{5'}$ angle (θ_{inl}). [b] Geometric quantities were taken from QM/MM simulations[8] with explicit solvent, see text for additional details.

TABLE 2: Comparison of Relative Energies for Stationary Points along the Reaction Path PES for the Transphosphorylation of 3′-Ribose,5′-methyl Phosphodiester in the Gas Phase and in Solution[a]

	DFT+PCM[b]			MNDO/d+SCOSMO[b]			MNDO/d-SCOSMO[c]			QM/MM[d]
molecule	ΔG_{aq}	ΔG_{gas}	$\Delta\Delta G_{sol}$	ΔG_{aq}	ΔG_{gas}	$\Delta\Delta G_{sol}$	ΔG_{aq}	ΔG_{gas}	$\Delta\Delta G_{sol}$	ΔG_{aq}
native										
reactant	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$TS_{O_2'}$										11.8
INT										11.6
$TS_{O_5'}$	19.4	31.4	−12.0	18.2	24.6	−6.4	18.3	23.6	−5.2	18.0
product	−8.2	−61.6	53.4	−3.5	−64.4	60.9	−2.2	−67.8	65.6	
sulfur-substituted										
reactant	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$TS_{O_2'}$	10.5	20.6	−10.1	10.8	19.5	−8.7	10.3	15.9	−5.6	12.1
INT	5.5	19.0	−13.5	5.2	17.0	−11.9	4.0	14.2	−10.3	4.8
$TS_{O_5'}$	17.6	30.3	−12.7	24.0	32.0	−8.0	24.5	30.6	−6.1	24.2
product	−7.7	−56.0	48.3	2.6	−54.7	57.3	3.9	−58.5	62.4	

[a] Relative free energy values (kcal/mol) with respect to reactants for stationary points along the reaction path PES for the transphosphorylation of 3′-ribose,5′-methyl phosphodiester in solution. Shown are the relative free energy values in aqueous solution (ΔG_{aq}), and also listed for the DFT+PCM, MNDO/d+SCOSMO, and MNDO/d-SCOSMO methods are the separate gas-phase (ΔG_{gas}) and solvation ($\Delta\Delta G_{sol}$) free energy components. The product energy is defined as the sum of the total energy for the 3′,5′-cyclic ribose phosphate and methoxide anion infinitely separated in solution (Scheme 2). [b] Free energy values calculated with the DFT+PCM and MNDO/d+SCOSMO methods are based on the gas phase-optimized geometries. Energy values follow standard thermodynamic and statistical mechanical expressions within the harmonic oscillator/rigid rotor/ideal gas approximations, as described in previous work.[29] [c] Free energy values calculated based on solution phase-optimized geometries with MNDO/d-SCOSMO. [d] Relative Gibbs free energy values taken from QM/MM simulations[8] with explicit solvent, see text for additional details.

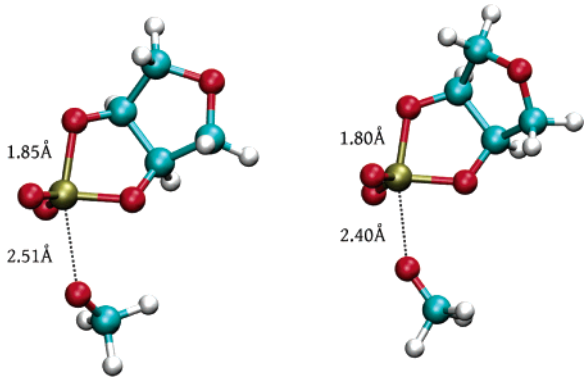


Figure 3. Rate-controlling transition state for the native transphosphorylation reaction: (a) DFT-optimized in the gas phase and (b) MNDO/d-optimized in solution.

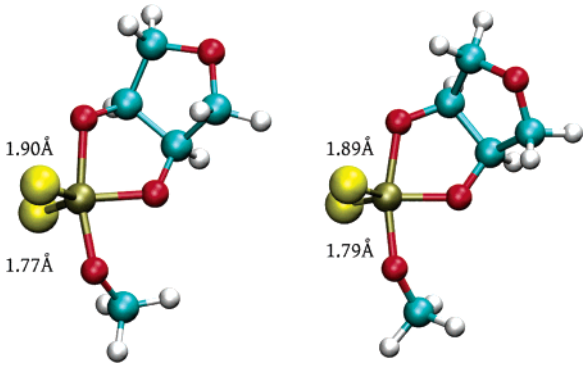


Figure 4. Stable gas-phase phosphorane intermediate for the doubly sulfur-substituted transphosphorylation reaction: (a) DFT-optimized in the gas phase and (b) MNDO/d-optimized in solution.

in-line $O_{2'}-P-O_{5'}$ angle (θ_{inl}) has the largest variation between the MNDO/d-SCOSMO and QM/MM methods for the thio-substitutes reactant complex (6.7°) where $r_1 = P-O_{2'}$ distance is the longest (around 4.2 Å). For the rate-limiting transition state of the native reaction, the θ_{inl} value for the MNDO/d-SCOSMO method is 1.3° larger than that of the QM/MM value (164.1°), and the MNDO/d-SCOSMO and QM/MM θ_{inl} values for the transition states, reactive intermediate, and product of the thio-substituted reaction agree to within 1°. These results suggest that the MNDO/d-SCOSMO is able to reliably reproduce the main structural and energetic features of more rigorous and computationally expensive QM/MM simulations of biological phosphate transesterification reactions in solution.

3.3. Comparison with Density-Functional Calculations. This section compares the structure and energetics for the native and thio-substituted dianionic 3′-ribose,5′-methyl phosphodiester

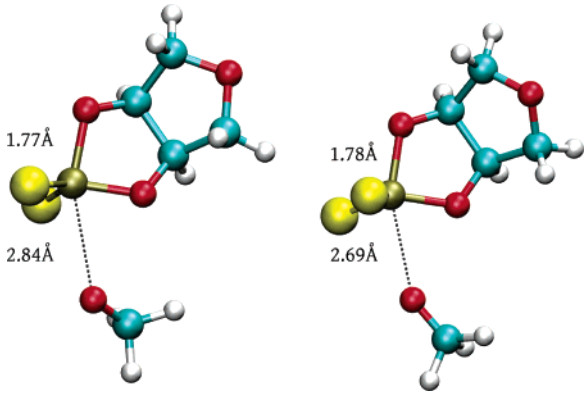


Figure 5. Rate-controlling transition state for the doubly sulfur-substituted transphosphorylation reaction: (a) DFT-optimized in the gas phase and (b) MNDO/d-optimized in solution.

transesterification reaction calculated with DFT+PCM, MNDO/d+SCOSMO, and MNDO/d-SCOSMO methods (Tables 1 and 2).

3.3.1. DFT and MNDO/d Gas-Phase Free Energy Values. The ΔG_{gas} values for the DFT and MNDO/d methods are overall in reasonable agreement. The biggest differences occur for the native reaction. This is in part due to the dianionic nature of the reaction that is particularly difficult to accommodate with a minimal valence basis set semiempirical method and particularly so with the harder phosphate oxygen ligands of the native reaction as compared with the softer thiophosphate sulfur ligands of the thio-substituted reaction. The forward $TS_{O_5'}$ transition state barrier in the gas phase for the native reaction (24.6 kcal/mol) is 6.8 kcal/mol lower than the corresponding DFT value (31.4 kcal/mol). The gas-phase product minimum, on the other hand, has a ΔG_{gas} value of −64.4 kcal/mol with MNDO/d, only 2.8 kcal/mol lower than the corresponding DFT value of −61.6 kcal/mol. The value of the reverse barrier (i.e., the $TS_{O_5'}$ barrier relative to the product) with MNDO/d (89.0 kcal/mol) is within 5% of the corresponding DFT value of 93.0 kcal/mol in the gas phase. The slightly reduced $TS_{O_5'}$ barrier for the gas-phase native reaction calculated with MNDO/d relative to the DFT value is correlated with constriction of the $r_2 = P-O_{5'}$ bond in $TS_{O_5'}$. The ΔG_{gas} values for the DFT and MNDO/d methods are in closer agreement for the thio-substituted reactions relative to the native reaction. The differences in the ΔG_{gas} values for the gas-phase thio-substituted reactions range from 1.1 to 2.0 kcal/mol, the largest difference occurs for the INT structure.

3.3.2. DFT and MNDO/d Gas-Phase Structures. In the native reactant minima, the $r_1 = P-O_{2'}$ distance calculated with DFT is longer by 0.3 Å than the MNDO/d distance, and the in-line angle (θ_{inl}) values are considerably different (108.6° and 125.3°, for DFT and MNDO/d, respectively) indicating slightly different reactant minimum structures in the gas phase. The most notable difference is that the leaving group P−O distance (r_2) is predicted by the gas-phase DFT calculation to be slightly elongated for the late $TS_{O_5'}$ transition state relative to MNDO/d. The difference between the DFT and MNDO/d values is likely due in part to the diffuse nature of the basis set used in the DFT transition state optimization that is able to partially help stabilize the separated anions. The $TS_{O_5'}$ r_2 values for the native and thio-substituted reactions calculated with the MNDO/d method (2.44 and 2.69 Å, respectively) are decreased by 0.07 and 0.15 Å, respectively, with respect to the DFT values (Table 1).

The gas-phase structural parameters for the DFT and MNDO/d methods are generally in closer agreement for the thio-substituted reactions than for the native reaction. The largest structural deviation occurs for the reactant minimum that, like the native reaction, shows a θ_{inl} value with DFT (98.4°) that is 15.7° less than the corresponding MNDO/d value (114.1°) in the gas phase. The other large difference occurs in the shortened gas-phase MNDO/d $r_2 = P-O_{5'}$ value of 2.69 Å relative to the DFT value of 2.84 Å. Overall, for the thio-substituted reactions, the DFT and MNDO/d methods are in remarkable agreement.

3.3.3. Comparison of Solvation Free Energy ($\Delta\Delta G_{sol}$) Values. The $\Delta\Delta G_{sol}$ values vary considerably between the DFT+PCM and the MNDO/d+SCOSMO and MNDO/d-SCOSMO methods. In general, the $\Delta\Delta G_{sol}$ values for MNDO/d+SCOSMO are in better agreement with the DFT+PCM values than are the MNDO/d-SCOSMO values. This is mainly due to the lack of solvent-induced structural relaxation for both the DFT+PCM and MNDO/d+SCOSMO methods. Solvent-induced structural relaxation tends to raise the $\Delta\Delta G_{sol}$ values (which are ΔG_{sol} values relative to the reactant minimum) in part because the reactants undergo the largest degree of structural relaxation upon solvation, as discussed in section 3.1.

3.3.4. Comparison of Aqueous Free Energy (ΔG_{aq}) Values. The differences in the DFT and MNDO/d gas-phase free energy and PCM and SCOSMO solvation energy largely cancel one another for the rate-controlling transition state that has an activation free energy barrier (ΔG_{aq}) value for DFT+PCM (19.4 kcal/mol) that is 1.1−1.2 kcal/mol higher than the MNDO/d+SCOSMO and MNDO/d-SCOSMO values (18.2 and 18.3 kcal/mol, respectively). The differences of the ΔG_{aq} values between the DFT+PCM and the MNDO/d+SCOSMO and MNDO/d-SCOSMO methods are less pronounced for the thio-substituted reaction. As with the $\Delta\Delta G_{sol}$ values, the MNDO/d+SCOSMO ΔG_{aq} values are in general closer to the DFT+PCM ΔG_{aq} values than are the those of MNDO/d-SCOSMO. Although the DFT and MNDO/d gas-phase free energies are closer to one another for the thio-substituted reaction than for the native reaction, the corresponding differences in solvation free energy ($\Delta\Delta G_{sol}$ values) show larger deviations. Unlike in the native reaction, the deviation in the $\Delta\Delta G_{sol}$ values do not have a large compensating effect on the ΔG_{gas} values for the rate-controlling activation barrier. The result is that the MNDO/d+SCOSMO and MNDO/d-SCOSMO methods predict a significantly elevated $TS_{O_5'}$ barrier (24.0 and 24.5 kcal/mol, respectively) by 6.4−6.9 kcal/mol relative to the DFT+PCM value (17.6 kcal/mol). These results underscore the delicate balance between electronic and solvation effects that govern biological phosphate reactivity, and in particular transphosphorylation thio effects.[8]

3.3.5. Comparison of DFT and MNDO/d-SCOSMO Structures. Despite the fairly significant differences in free energies for the DFTPCM and MNDO/d-SCOSMO methods, the corresponding structures are remarkably similar (Table 1). The structures optimized with MNDO/d-SCOSMO were overall more similar to the gas phase-optimized DFT structures than the gas phase-optimized MNDO/d structures. As noted previously, the P−O bond distances are observed to contract upon solvation. The bond-length contraction that occurs upon solvation with the MNDO/d-SCOSMO method is only significant for the native reaction (r_2 decreases from 2.44 Å in the gas phase to 2.40 Å in solution). These results imply that solvent-induced structural relaxation at the DFT level may be important, as solvation effects screen the Coulomb repulsion between the anionic charge centers and favor the association of like-charged ions due to the leading term in the Kirkwood expansion[39] that predicts the solvent energy to be proportional to the square of the total charge. Treatment of solvation as a post-correction to

gas-phase optimized structures with DFT methods is a fairly common procedure that, if properly calibrated, provides reasonable relative solvation energy results.[29−32] Optimizations at the DFT level with solvation, on the other hand, often fail due to the lacking smoothness of the solvation potential in available DFT software, a long-standing problem with many integrated electronic structure and solvation packages that has been overcome with the MNDO/d-SCOSMO method. The differences in structure of the native reaction optimized with MNDO/d in the gas phase and with MNDO/d-SCOSMO suggest that inclusion of solvation in optimizations of transition states and reactive intermediates of dianionic phosphoryl transfer reactions can be important. It therefore is of considerable interest to extend the solvation model of the present work to other quantum methods such as DFT.

3.4. Relation with Experiment. Substitution of the nonbridging oxygens by sulfurs modifies the observed energy profiles through several competing effects. The dianionic phosphorane becomes electronically stabilized due to the higher electronic polarizability of sulfur atoms (chemically softer). However, the sulfur atoms are larger and consequently are less well solvated. The sulfur atoms are also much weaker hydrogen bond acceptors than oxygen atoms, thus, solvent stabilization through solvent hydrogen-bond donation is expected to play a less significant role in the overall stabilization of the dianionic charge of the intermediates and transition states.

The native and thio-substituted reaction profiles obtained from all of the methods are in qualitative agreement with experimental observations that a stable dianionic intermediate for the native reaction is unlikely[9] and that double sulfur substitution at the nonbridging positions results in the accumulation of a kinetically significant intermediate.[10] The existence of a stable dianionic thiophosphorane intermediate (Figure 4) has important implications since, especially in the case of the dithiophosphorane, the lifetime may be sufficient to undergo protonation by solvent and subsequent pseudorotation to form migration products.[30]

In the absence of an enhanced leaving group, the 2′,3′-cyclic phosphate can be readily detected only under alkaline conditions owing to rapid acid-catalyzed hydrolysis to give 2′- or 3′-phosphates. If a metastable phosphorane intermediate were to exist, pseudorotation could occur followed by fast endo-cyclic cleavage to produce a mixture of 3′,5′- and 2′,5′-phosphates. The barriers to pseudorotation for native and thio-substituted phosphoranes have been predicted to be quite low[30] (less than 4 kcal/mol for monoanionic phosphoranes). Under alkaline conditions, where the phosphates are expected to exist in dianionic form, no such phosphate migration is observed in RNA.[11]

The results from the MNDO/d-SCOSMO and QM/MM methods suggest that alkaline transphosphorylation reactions show a considerable thio effect at the nonbridging phosphoryl positions. This is consistent with experimental observations that phosphorothioates with double sulfur substitution at the nonbridging positions are more resistant to hydrolysis.[10] However, the magnitude of the observed thio effects is in some contrast with experimental results for RNA analogues with nonbridging thio substitutions that exhibit only modest thio effects.[11,40,41] The DFT+PCM results, on the other hand, predict that the rate-controlling transition states of the native and thio-substituted reactions have more similar barriers, but suggest that there is a mild reverse thio effect, which is not supported by experiment. Clearly, a highly accurate quantum model and reliable representation of the solvent environment are both necessary for quantitative predictions of mechanisms and reaction rates.[42] These results underscore the need for the development of new, highly accurate semiempirical quantum methods[43−45] for biochemical reactions that can be used in combined QM/MM simulations for improved modeling of the macromolecular and/or solvent environment.

4. Conclusion

In the present work, a recently developed MNDO/d-SCOSMO method is applied to the study of phosphoryl transfer reactions in solution and the results are compared with those from other electronic structure and solvation methods. The MNDO/d-SCOSMO method is demonstrated to be stable in the calculation of PESs for these reactions and to provide a robust description of solvation effects on the reaction coordinate and activation barriers.

Results of the MNDO/d-SCOSMO method were compared to results from density-functional/implicit solvation methods and hybrid QM/MM simulations in explicit solvent using the same set of van der Waals radii. It was observed that optimization of the probe radius for the MNDO/d-SCOSMO method gave good agreement with the QM/MM results for the entire PES and excellent agreement for the reactant minimum and rate-limiting transition state. This observation provides impetus to further develop the smooth COSMO method into a tool to assist in the parametrization of QM/MM van der Waals radii.

The reaction energy profile of the native reaction reveals no intermediate in the gas phase calculated with DFT or MNDO/d or in solution calculated with MNDO/d-SCOSMO. The QM/MM simulation produces a kinetically insignificant late intermediate at 10 kcal/mol above the reactant. This difference is attributed partially to the absence of the stabilizing hydrogen bonding in the implicit model. The reaction profile of the double sulfur-substituted reaction shows both pentavalent early and late transition states in the gas-phase as well as in solution; the relative energies of both transition states from the MNDO/d-SCOSMO calculation are very similar to those from the QM/MM simulation. Generally, the reaction profiles from MNDO/d-SCOSMO and QM/MM closely resemble each other both for the native and double sulfur substituted reactions, and they are consistent with the experimentally observed thio effects and the associative nature of the mechanism. Furthermore, the solvent-induced structure relaxations at all stages of the native and sulfur-substituted reactions predicted by the MNDO/d-SCOSMO method are similar to those from the explicit solvent QM/MM simulation.

Overall the results of the present work demonstrate the utility of implicit solvation methods as an integral part of the development of multi-scale quantum models for complex biological reactions. It is hoped that further development of such models may provide new insight into the molecular mechanisms of phosphoryl transfer reactions and processes of protein phosphorylation, cell signaling, and RNA catalysis.

Acknowledgment. D.Y. is grateful for financial support provided by the National Institutes of Health (Grant GM62248) and the Army High Performance Computing Research Center (AHPCRC) under the auspices of the Department of the Army, Army Research Laboratory (ARL) under Cooperative Agreement Number DAAD19-01-2-0014. Computational resources were provided by the Minnesota Supercomputing Institute.

Supporting Information Available: Gas-phase reaction profiles and key geometric parameters of the rate-limiting transition state for common semiempirical methods are com-

pared with density-functional calculations This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Rashin, A. A. *J. Phys. Chem.* **1990**, *94*, 1725−1733.
(2) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.
(3) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(4) York, D. M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060−11079.
(5) Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **1993**, *115*, 5316−5317.
(6) Dejaegere, A.; Liang, X. L.; Karplus, M. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1763−1767.
(7) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719−734.
(8) Gregersen, B. A.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 7504−7513.
(9) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 432−450.
(10) Zhou, D.-M.; Taira, K. *Chem. Rev.* **1998**, *98*, 991−1026.
(11) Oivanen, M.; Kuusela, S.; Lönnberg, H. *Chem. Rev.* **1998**, *98*, 961−990.
(12) Hengge, A. C. *Acc. Chem. Res.* **2002**, *35*, 105−112.
(13) Takagi, Y.; Ikeda, Y.; Taira, K. *Top. Curr. Chem.* **2004**, *232*, 213−251.
(14) Khandogin, J.; Gregersen, B. A.; Thiel, W.; York, D. M. *J. Phys. Chem. B* **2005**, *109*, 9799−9809.
(15) Thiel, W.; Voityuk, A. A. *Theor. Chim. Acta* **1992**, *81*, 391−404.
(16) Thiel, W.; Voityuk, A. A. *J. Phys. Chem.* **1996**, *100*, 616−626.
(17) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(18) Riccardi, D.; Li, G.; Cui, Q. *J. Phys. Chem. B* **2004**, *108*, 6467−6478.
(19) Thiel, W. *MNDO97*, version 5.0; University of Zurich: Zurich, Switzerland, 1998.
(20) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104.
(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(22) Brooks, C. L., III.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(23) Brooks, C. L., III.; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843−865.
(24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(25) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578−581.
(26) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(28) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785−789.
(29) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 1654−1665.
(30) López, C. S.; Faza, O. N.; Gregersen, B. A.; Lopez, X.; de Lera, A. R.; York, D. M. *Chem. Phys. Chem.* **2004**, *5*, 1045−1049.
(31) Mayaan, E.; Range, K.; York, D. M. *J. Biol. Inorg. Chem.* **2004**, *9*, 807−817.
(32) López, C. S.; Faza, O. N.; de Lera, A. R.; York, D. M. *Chem. Eur. J.* **2005**, *11*, 2081−2093.
(33) Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996**, *17*, 49−56.
(34) Bauernschmitt, R.; Ahlrichs, R. *J. Chem. Phys.* **1996**, *104*, 9047−9052.
(35) Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1977**, *66*, 3045−3050.
(36) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43−54.
(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.01; Gaussian, Inc.: Pittsburgh, PA, 2003.
(38) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 5010−5018.
(39) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351−361.
(40) Almer, H.; Strömberg, R. *Tetrahedron Lett.* **1991**, *32*, 3723−3726.
(41) Ora, M.; Järvi, J.; Oivanen, M.; Lönnberg, H. *J. Org. Chem.* **2000**, *65*, 2651−2657.
(42) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.
(43) Thiel, W. Perspectives on semiempirical molecular orbital theory. In *Advanced Chemical Physics*, Vol. 93; Prigogine, I., Rice, S. A., Eds.; John Wiley and Sons: New York, 1996.
(44) Clark, T. *J. Mol. Struct. (THEOCHEM)* **2000**, *530*, 1−10.
(45) Thiel, W. Semiempirical Theories. In *Handbook of Molecular Physics and Quantum Chemistry*, Vol. 2; Wilson, S., Ed.; John Wiley and Sons: Chichester, U.K., 2003.