AIP | The Journal of Chemical Physics

VR-SCOSMO: A smooth conductor-like screening model with charge-dependent radii for modeling chemical reactions

Erich R. Kuechler, Timothy J. Giese, and Darrin M. York

Citation: The Journal of Chemical Physics **144**, 164115 (2016); doi: 10.1063/1.4946779
View online: http://dx.doi.org/10.1063/1.4946779
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/144/16?ver=pdfcov
Published by the AIP Publishing

VR-SCOSMO: A smooth conductor-like screening model with charge-dependent radii for modeling chemical reactions

Erich R. Kuechler,[1,2] Timothy J. Giese,[1] and Darrin M. York[1]
[1]*Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854-8087, USA*
[2]*Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, USA*

(Received 18 December 2015; accepted 1 April 2016; published online 28 April 2016)

To better represent the solvation effects observed along reaction pathways, and of ionic species in general, a charge-dependent variable-radii smooth conductor-like screening model (VR-SCOSMO) is developed. This model is implemented and parameterized with a third order density-functional tight binding quantum model, DFTB3/3OB-OPhyd, a quantum method which was developed for organic and biological compounds, utilizing a specific parameterization for phosphate hydrolysis reactions. Unlike most other applications with the DFTB3/3OB model, an auxiliary set of atomic multipoles is constructed from the underlying DFTB3 density matrix which is used to interact the solute with the solvent response surface. The resulting method is variational, produces smooth energies, and has analytic gradients. As a baseline, a conventional SCOSMO model with fixed radii is also parameterized. The SCOSMO and VR-SCOSMO models shown have comparable accuracy in reproducing neutral-molecule absolute solvation free energies; however, the VR-SCOSMO model is shown to reduce the mean unsigned errors (MUEs) of ionic compounds by half (about 2-3 kcal/mol). The VR-SCOSMO model presents similar accuracy as a charge-dependent Poisson-Boltzmann model introduced by Hou *et al.* [J. Chem. Theory Comput. **6**, 2303 (2010)]. VR-SCOSMO is then used to examine the hydrolysis of trimethylphosphate and seven other phosphoryl transesterification reactions with different leaving groups. Two-dimensional energy landscapes are constructed for these reactions and calculated barriers are compared to those obtained from *ab initio* polarizable continuum calculations and experiment. Results of the VR-SCOSMO model are in good agreement in both cases, capturing the rate-limiting reaction barrier and the nature of the transition state. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4946779]

I. INTRODUCTION

Computational modeling of biological reactions is of great interest to those seeking molecular-level insight into reaction mechanisms. Such studies can help interpret experimental results which are pursuing a deeper understanding of the fundamental principles governing chemical activity and encompass a wide range of archetypes including, but not limited to, reactions occurring within protein environments,[1–3] model phosphoryl transfer reactions,[4–8] and studies of RNA catalysis.[9–11] The end goal of these studies is frequently to aid in the design of new drugs and therapeutic treatments.[12–14] Multiscale modeling approaches including long timescale molecular dynamics simulations to explore fluctuations and conformational changes of biomolecules, combined quantum mechanical/molecular mechanical (QM/MM) simulations to examine deeply embedded reactive chemical events, and implicit solvent calculations of small model reactions are used for this task.

Biological reactions almost exclusively take place in solution, which has drastic ramifications on the pathways and barriers of the reaction mechanism.[4,6,15–22] These phenomena are especially apparent when studying highly charged systems such as RNA[11] as shown in Figure 1. To model these effects, simulations can be performed that include a vast buffer of explicit solvent molecules; however, such approaches are often extremely computationally demanding. Frequently, this cost leads the simulation into becoming impractical, especially when using expensive quantum mechanical models. Implicit solvation techniques forego the need to perform extensive sampling of explicit solvent degrees of freedom by replacing the solvent response of an effective, empirical model.[23–26] In addition to enabling stand-alone applications to chemical reactivity, implicit solvation models can also be used to guide sampling in more expensive QM/MM free energy simulations.[27–30] Implicit solvent models have come in a variety of forms, including Poisson-Boltzmann (PB),[31] the Minnesota solvation (SMx),[32–34] the conductor-like screening (COSMO),[35–40] generalized Born (GB),[41,42] and polarizable continuum (PCM) models,[43–45] among many others.

One particularly useful class of solvation models is based on boundary element methods. The original boundary element solvation models suffered from discontinuities in the solvation energy as atoms became buried or solvent exposed as a function of conformation (thus exposing boundary elements discontinuously), making them difficult, if not impossible to use effectively in geometry optimizations or explorations of reaction pathways. York and Karplus introduced the first model to fully address this issue[46] within the COSMO[35]

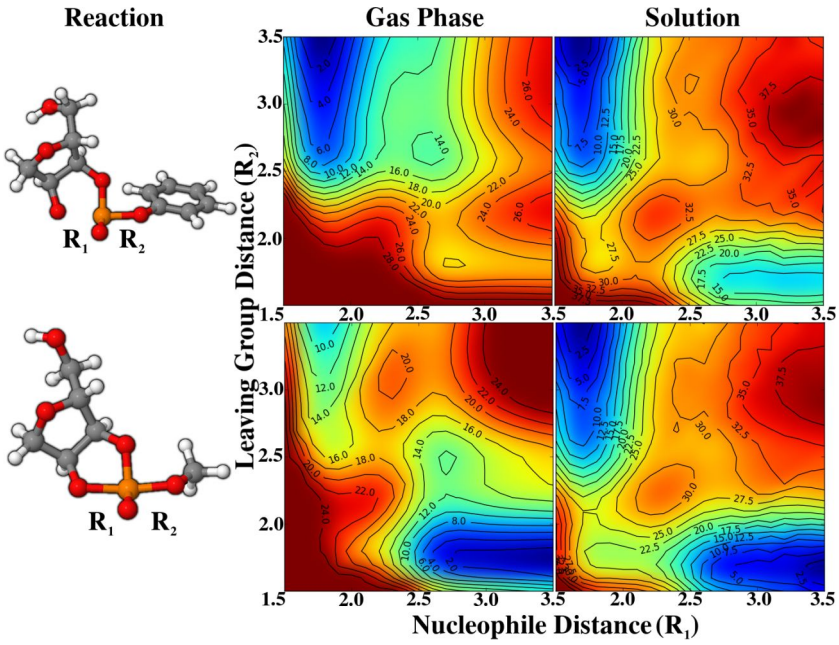


FIG. 1. Comparison between gas phase and solution reaction energy profiles for two model RNA-like reactions. Special note should be made of the difference in color scale between the gas and solution phase surfaces. All distances and energies are reported in Å and kcal/mol respectively.

framework. Other similar smooth boundary element models have followed this formalism and have proven useful to the larger quantum chemistry community.[45,47]

However, another limitation to traditional boundary element solvation methods exists which involves the use of fixed radii to define the solute cavity. In some instances, to obtain more accurate solvation free energies, different radii are assigned to the same element depending on its chemical environment. This procedure becomes ill-defined when studying a chemical reaction where the chemical environment can change along the reaction coordinate. A mechanism whereby the radii themselves can change smoothly in this process is needed. Furthermore, while this fixed-radii strategy often works reasonably well for neutral solutes near local minima, it often failed when considering charged solutes, or chemical reactions where localized charge varied along the reaction coordinate.

The present work introduces the variable-radii smooth conductor-like screening model (VR-SCOSMO) whereby the solvation radii themselves vary smoothly as a function of the underlying electronic structure, as represented by atomic charge. Influenced by the pioneering work of Hou *et al.*[20] and others,[48–50] the present work applies a similar strategy of introducing solvation radii that have a charge dependence into the SCOSMO model.[46] In this work, the DFTB3/3OB semiempirical method is used,[51–53] and parameterization of the SCOSMO and VR-SCOSMO models is created to reproduce absolute solvation energies. The parameterized SCOSMO and VR-SCOSMO models are compared to each other and to the SCC-DFTBPR model.[20] The VR-SCOSMO model is then used to construct eight two-dimensional relaxed potential energy surfaces of biologically motivated reactions, which are then compared to available experimental results. One of these surfaces is for the trimethylphosphate hydrolysis reaction. The other surfaces are transesterification phosphoryl transfer reactions involving seven different leaving groups.

II. COMPUTATIONAL METHODS

This manuscript uses the semiempirical DFTB3/3OB Hamiltonian[52] supplemented with a specific reaction parameterization, OPhyd,[53] which modulates the O and P parameters for improving performance on phosphate hydrolysis reactions. The DFTB3/3OB-OPhyd parameters nor the functional form for the solute-solute self-interaction energy has been altered in this manuscript. Therefore, for further discussion of technical details of the DFTB3 model, the reader is referred to Refs. 51–53. It can be stated that standard gas-phase DFTB3 calculations depend on the nuclear positions and the atomic orbital density matrix

$$E_{\text{gas}} \equiv E_{\text{DFTB3}}(\mathbf{R}, \mathbf{P}), \tag{1}$$

where $\mathbf{R}$ is a vector of atomic positions,

$$\mathbf{P} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}, \tag{2}$$

$$P^{\sigma}_{\mu\nu} = \sum_i n^{\sigma}_i C^{\sigma}_{\mu i} C^{\sigma}_{\nu i}, \tag{3}$$

where $\mathbf{n}^{\sigma}$ and $\mathbf{C}^{\sigma}$ are spin-resolved orbital occupation numbers and molecular orbital coefficients, respectively, in the atomic orbital basis of μ and ν. For all results presented, DFTB3/3OB and VR-SCOSMO calculations were performed using an in-house program which will be integrated into the Sander module AMBER and into the CHARMM molecular simulation software suites and made available in future full releases.

A. Solute charge density interaction correction

The DFTB3/3OB energy happens to use Mulliken charges to interact the solute atoms with the other solute atoms;[52] however, previous work has found that the interaction of DFTB3 molecules with an "external environment" can be improved by choosing a second charge representation.[54] This

approach has been adopted in the present work. In essence, a second set of atomic charges and higher-order atomic multipoles are chosen to interact with the implicit solvent response, while leaving the DFTB3 solute-solute interactions unchanged from their original form. The solute charge density can be approximated by a sum of atomic point multipoles

$$q(\mathbf{r}) = \sum_a Z_a \delta(\mathbf{r} - \mathbf{R}_a) - \sum_{\mu\nu} P_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}) \approx \sum_{a,lm\in a} q_{a,lm} \frac{C_{lm}(\nabla_a)}{(2l-1)!!} \delta(\mathbf{r} - \mathbf{R}_a), \quad (4)$$

where Z_a is the nuclear core charge of atom a, $q_{a,lm}$ is a multipole moment, $\mathbf{R}_a$ is an atomic position, ∇_a is the Cartesian gradient operator acting on the coordinates of a, and $C_{lm}(\nabla)$ is a spherical tensor gradient operator which is constructed by replacing the Cartesian coordinate arguments of the real-valued regular solid harmonic with their corresponding derivative operators. In this notation,

$$\chi_\mu(\mathbf{r}) \equiv \chi_\mu(r) Y_{l_\mu m_\mu}(\Omega), \quad (5)$$

where $Y_{lm}(\Omega)$ is a real-valued spherical harmonic. Following Refs. 54 and 55, the atomic charges are a biased Mulliken partitioning of the density matrix, and the higher-order multipole moments are constructed from the one-center blocks of the density matrix,

$$q_{a,lm} = \begin{cases} Z_a - \frac{b_{aa}}{2} - \sum_{b\neq a} f_{ab}(b_{ab}) b_{ab} & \text{if } l = 0 \\ \sum_{\mu\nu\in a} P_{\mu\nu} M^{(l)}_{\mu\nu} \sqrt{\frac{4\pi}{2l+1}} & \text{if } l > 0 \\ \times \int Y_{lm}(\Omega) Y_{l_\mu m_\mu}(\Omega) Y_{l_\nu m_\nu}(\Omega) d\Omega & \end{cases}, \quad (6)$$

where b_{ab} is a Mulliken bond order

$$b_{ab} = 2 \sum_{\substack{\mu\in a \\ \nu\in b}} P_{\mu\nu} S_{\mu\nu}, \quad (7)$$

where $\mathbf{S}$ is the atomic orbital overlap matrix

$$S_{\mu\nu} = \int \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}) d^3r \quad (8)$$

and f_{ab} is a "charge bias." This bias is constructed such that at $f_{ab} = 1/2$, a Mulliken charge decomposition is produced, but it is generalized to switch from f^{s}_{ab} to f^{d}_{ab} as the Mulliken bond order changes from single bond character, b^{s}_{ab}, to double bond character, b^{d}_{ab},

$$f_{ab}(b_{ab}) = f^{\mathrm{s}}_{ab} + S_{\mathrm{on}}(b_{ab}, b^{\mathrm{s}}_{ab}, b^{\mathrm{d}}_{ab})(f^{\mathrm{d}}_{ab} - f^{\mathrm{s}}_{ab}), \quad (9)$$

where $S_{\mathrm{on}}(x, x_{\mathrm{lo}}, x_{\mathrm{hi}})$ is a smooth switching function

$$S_{\mathrm{on}}(x, x_{\mathrm{lo}}, x_{\mathrm{hi}}) = \begin{cases} 0 & \text{if } x < x_{\mathrm{lo}} \\ 1 & \text{if } x > x_{\mathrm{hi}} \\ 1 - 10u^3 - 15u^4 + 6u^5 & \text{otherwise} \\ u \equiv (x_{\mathrm{hi}} - x)/(x_{\mathrm{hi}} - x_{\mathrm{lo}}) & \end{cases} \quad (10)$$

and $M^{(l)}_{\mu\nu}$ are the one-center radial integrals

$$M^{(l)}_{\mu\nu} = \int_0^\infty \chi_\mu(r) \chi_\nu(r) r^{2+l} dr \quad (11)$$

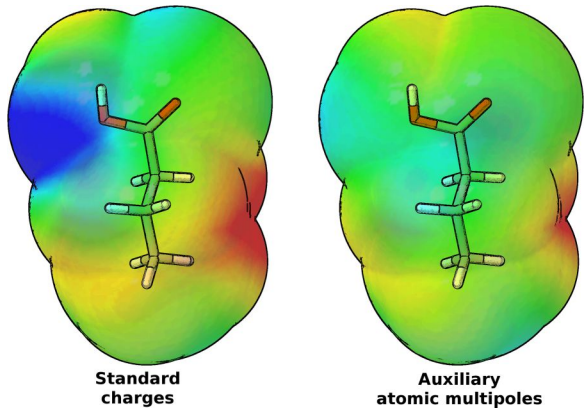


FIG. 2. The electrostatic potential difference between B3LYP/6-311++G** and DFTB3/3OB (left) and DFTB3/3OB using the auxiliary atomic multipoles (right) evaluated on the solvent accessible surface of butanoic acid. Blue and red indicate that the semiempirical model requires more negative charge and more positive charge relative to B3LYP reference, respectively. Colors are bounded with a range ±0.003 a.u.

which are treated as parameters to empirically improve molecular electrostatic potentials. The parameters of this model, f^{s}_{ab}, f^{d}_{ab}, b^{s}_{ab}, b^{d}_{ab}, and $M^{(l)}_{\mu\nu}$, are listed in the supplementary material of Ref. 55 and are not modified in this work. The effects of the inclusion of this correction can be seen in Figure 2 for butanoic acid, showing how the use of multipoles can improve the description of sp^3 oxygen lone pairs. Previous work has noted improvements to the description of sp^2 carbon bonds, and sp^3 sulfur and sp^2 nitrogen lone pairs.

B. The SCOSMO implicit solvation model

The interface between the SCOSMO model[56,57] and the DFTB3 semiempirical method closely follows the formulation provided by Khandogin *et al.*[58] Briefly, the SCOSMO model treats the solvent environment as if it were a block of conducting metal whose dielectric constant is chosen to mimic the solvent on a macroscopic scale. Then, a cavity with a dielectric constant of unity, $\epsilon_0 = 1$, is carved out of the metal and the solute is placed within it. Given the discontinuous nature between the cavity and the bulk solvent dielectric constants at the cavity's boundary, the electrostatic response of the metal occurs solely at the cavity surface. The cavity is constructed from a union of solute spheres whose radii, $R_{\mathrm{rad},a}$, are parameters of the model. The cavity surface is discretized by Gaussian functions,

$$g_t(\mathbf{r} - \mathbf{R}_t; \zeta_t) = (\zeta_t/\pi)^{3/2} e^{-\zeta_t |\mathbf{r} - \mathbf{R}_t|^2} \quad (12)$$

such that the solvent's charge density response at the surface is

$$\sigma(\mathbf{r}) \approx \sum_t s_t g_t(\mathbf{r} - \mathbf{R}_t; \zeta_t), \quad (13)$$

where $\mathbf{s}$ is the vector of solvent response coefficients, s_t, that remain to be determined. In our notation, t indexes a discretized point on the surface and

$$\mathbf{R}_t = \mathbf{R}_a + R_{\mathrm{rad},a} \hat{\mathbf{R}}_{\mathrm{q},t} \quad (14)$$

is the location of that point. In other words, there are $N_{q,a}$ discretized points, each point t is tethered to one atom $t \in a$ separated by $R_{rad,a}$, in the direction $\hat{\mathbf{R}}_{q,t}$, where $\hat{\mathbf{R}}_{q,t}$ is a Lebedev quadrature point that is associated with corresponding quadrature weight $w_{q,t}$.[59,60]

The solvated energy of a compound can be modeled as an adiabatic energy of that compound in the presence of SCOSMO implicit

$$E_{aq} \equiv E_{DFTB3}(\mathbf{R},\mathbf{P}) + E_s(\mathbf{s},\mathbf{q}), \tag{15}$$

where

$$E_s(\mathbf{s},\mathbf{q}) = \frac{1}{2f(\epsilon)}\mathbf{s}^T \cdot \mathbf{A} \cdot \mathbf{s} + \sum_{a,t\in a} \gamma w_{q,t} p_t R_{rad,a}^2 + \mathbf{s}^T \cdot \mathbf{B} \cdot \mathbf{q} \tag{16}$$

and

$$f(\epsilon) = \frac{\epsilon - 1}{\epsilon} \tag{17}$$

is the scaled deviation of the dielectric medium of the implicit solvent from an ideal conductor. Throughout this work, the dielectric $\epsilon = 78.4$ (unitless) is used for water. The scaled self-interaction of the polarized surface charges of residing upon the cavitation barrier, $\mathbf{A}$, takes the form

$$A_{tt'} = \int\int \frac{g_t(\mathbf{r}-\mathbf{R}_t;\zeta_t) g_{t'}(\mathbf{r}'-\mathbf{R}_{t'};\zeta_{t'})}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r' \times \begin{cases} p_t^{-1}, & \text{if } t = t' \\ 1, & \text{otherwise} \end{cases}. \tag{18}$$

The switching function p_t smoothly transitions between 0 and 1, preventing surface elements from interacting with other surface elements which have been partially or fully removed by the intersection of the cavitation sphere from another particle,

$$p_{t\in a} = \prod_{b\neq a} S_{on}(R_{bt}^2, R_{in,b}^2, R_{out,b}^2). \tag{19}$$

When $p_t = 0$ dimensionality of the arrays can be decreased safely without causing discontinuous changes in the energy or gradients. In this case, $R_{in,a}$ and $R_{out,a}$ are the limits of the switching region, whose explicit definitions are

$$R_{in,a} = (1 - oh)R_{rad,a} \tag{20}$$

and

$$R_{out,a} = (1 + (1-o)h)R_{rad,a}, \tag{21}$$

where

$$o = \frac{1}{2} + \frac{1}{h} - \sqrt{\frac{1}{h^2} - \frac{1}{28}} \tag{22}$$

such that h is some constant, which was set to $h = 0.15$ (unitless) for this work. Continuing, γ is the "surface tension" of the cavity which accounts for the free energy cost associated with the formation of the cavity. This term is treated as a parameter for optimization in the current protocols. $\mathbf{B}$ is the interaction between the previously defined surface charges and the solute charge distribution, $\mathbf{q}$,

$$B_{ta} = \frac{C_{l_a m_a}(\nabla_a)}{(2l_a-1)!!}\int\int \frac{g_t(\mathbf{r}-\mathbf{R}_t;\zeta_t)\delta(\mathbf{r}'-\mathbf{R}_a)}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r' = \frac{C_{l_a m_a}(\mathbf{R}_{at})}{(2l_a-1)!!}\left(\frac{d}{dR_{at}^2}\right)^{l_a} \frac{\mathrm{erf}(\sqrt{\zeta_t} R_{at})}{R_{at}}, \tag{23}$$

where

$$\zeta_{t\in a} = \frac{1}{w_{q,t}}\left(\frac{\zeta(N_{q,a})}{R_{rad,a}}\right)^2. \tag{24}$$

Values of $\zeta(N_{q,a})$ constants are tabulated in Ref. 61 and are completely determined upon defining the discretization level of each atom, $N_{q,a}$. Special note should be taken that in the present work the interaction of the electrostatic potential with the surface response elements has been modified. The standard monopole representation is replaced with an auxiliary set of atomic multipole moments, constructed from the underlying DFTB3 density matrix.

The response coefficients, s_t, are determined by minimizing

$$\delta\left\{E_s(\mathbf{s},\mathbf{q}) - \lambda\left(\sum_t s_t + f(\epsilon)Q\right)\right\} = 0, \tag{25}$$

where

$$Q = \sum_a q_{a,00} \tag{26}$$

is the solute charge, q_a is the atomic charge of atom a, and the Lagrange multiplier, λ, is used to enforce the constraint

$$\int \sigma(\mathbf{r}) d^3r = -f(\epsilon)Q. \tag{27}$$

Following this constraint, one finds

$$\mathbf{s} = f(\epsilon)\left(\lambda \mathbf{A}^{-1}\cdot\mathbf{v} - \mathbf{A}^{-1}\cdot\mathbf{B}\cdot\mathbf{q}\right), \tag{28}$$

$$\lambda = \frac{\mathbf{v}^T\cdot\mathbf{A}^{-1}\cdot\mathbf{B}\cdot\mathbf{q} - Q}{\mathbf{v}^T\cdot\mathbf{A}^{-1}\cdot\mathbf{v}}, \tag{29}$$

where $v_t = 1\,\forall t$. The quantities treated as adjustable parameters in the present work are the surface tension term, γ, and the cavitation radii $R_{rad,a}$.

The SCOSMO model enters the solution for the DFTB3 molecular orbital coefficients variationally through the Fock matrix

$$F_{\mu\nu}^{\sigma} = \left.\frac{\partial E_{aq}}{\partial P_{\mu\nu}^{\sigma}}\right|_{\mathbf{R}} = \left.\frac{\partial E_{DFTB3}}{\partial P_{\mu\nu}^{\sigma}}\right|_{\mathbf{R}} + \sum_{a,lm\in a}\left.\frac{\partial E_s}{\partial q_{a,lm}}\right|_{\mathbf{R},\mathbf{s}}\left.\frac{\partial q_{a,lm}}{\partial P_{\mu\nu}}\right|_{\mathbf{R}}. \tag{30}$$

Upon reaching self-consistent convergence, the molecular orbitals satisfy the generalized eigenvalue equation

$$\mathbf{F}^{\sigma}\cdot\mathbf{C}^{\sigma} = \mathbf{S}\cdot\mathbf{C}^{\sigma}\cdot\mathbf{E}^{\sigma}, \tag{31}$$

where $\mathbf{S}$ is the atomic orbital overlap matrix

$$S_{\mu\nu} = \int \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) d^3r \tag{32}$$

and the Cartesian gradient X_a of atom a becomes

$$\frac{dE_{\mathrm{aq}}}{dX_a} = \left.\frac{\partial E_{\mathrm{DFTB3}}}{\partial X_a}\right|_{\mathbf{P}} - \sum_{\mu\nu} Q_{\mu\nu}\frac{dS_{\mu\nu}}{dX_a} + \left.\frac{\partial E_s}{\partial X_a}\right|_{\mathbf{s},\mathbf{q}} + \sum_b \left.\frac{\partial E_s}{\partial q_{b,lm}}\right|_{\mathbf{R},\mathbf{s}} \left.\frac{\partial q_{b,lm}}{\partial X_a}\right|_{\mathbf{P}}, \quad (33)$$

where

$$Q_{\mu\nu} = \sum_{\sigma\in(\alpha,\beta)} \sum_i n_i^\sigma E_{ii}^\sigma C_{\mu i}^\sigma C_{\nu i}^\sigma \quad (34)$$

and E_{ii}^σ is a spin-resolved molecular orbital eigenvalue.

C. The VR-SCOSMO implicit solvation model

The VR-SCOSMO model modifies the SCOSMO framework by incorporating charge-dependent behavior into the definition of the solvation radii

$$R_{\mathrm{rad},a}(q_{a,00}) = R_{\mathrm{rad},a} + \alpha(q_{a,00} - q_a^{(0)}) + \frac{1}{2}\beta(q_{a,00} - q_a^{(0)})^2, \quad (35)$$

where $R_{\mathrm{rad},a}$, α, β, and $q_a^{(0)}$ are parameters. With this change, the surface response Gaussian positions $\mathbf{R}_t$ and exponents ζ_t [Eq. (24)] also become charge-dependent. Upon replacing $R_{\mathrm{rad},a}$ with $R_{\mathrm{rad},a}(q_{a,00})$, the multipole moment-derivatives appearing in Eqs. (30)-(33) require additional chain-rules; that is,

$$\left.\frac{\partial E_s}{\partial q_{a,lm}}\right|_{\mathbf{R},\mathbf{s}} = \left.\frac{\partial E_s}{\partial q_{a,lm}}\right|_{\mathbf{R},\mathbf{s},\zeta,\mathbf{R_t}} + \delta_{l0}\delta_{m0}\sum_{t\in a}\left.\frac{\partial E_s}{\partial \zeta_t}\right|_{\mathbf{R},\mathbf{s},\mathbf{R_t}} \frac{d\zeta_t}{dR_{\mathrm{rad},a}}\frac{dR_{\mathrm{rad},a}}{\partial q_{a,00}} + \delta_{l0}\delta_{m0}\sum_{t\in a}\left.\frac{\partial E_s}{\partial X_t}\right|_{\mathbf{R},\mathbf{s},\zeta} \frac{dX_t}{dR_{\mathrm{rad},a}}\frac{dR_{\mathrm{rad},a}}{\partial q_{a,00}}. \quad (36)$$

D. Parameterization of the SCOSMO and VR-SCOSMO models

Parameters for the SCOSMO and VR-SCOSMO models, both utilizing the multipole correction, were optimized for elements H, C, N, O, and P to reproduce the absolute solvation free energy of a subset of the SM6 database[32] consisting of 169 molecules (139 neutral, 16 anionic, 14 cationic). A table listing these molecules can be found in the supplementary material, Table S1.[81] The absolute solvation free energy is approximated by the difference in the adiabatic electronic energy of the molecule in the implicit solvent and gas phase environments upon geometry optimization in its respective state,

$$\Delta G_{\mathrm{solv}} \equiv E_{\mathrm{aq}}(\mathbf{R}_{\mathrm{aq}}, \mathbf{P}_{\mathrm{aq}}) - E_{\mathrm{gas}}(\mathbf{R}_{\mathrm{gas}}, \mathbf{P}_{\mathrm{gas}}). \quad (37)$$

The parameterizations of the SCOSMO and VR-SCOSMO models were treated independently. Parameters obtained for both models are not constrained to have the same values. Parameter optimization was performed by direction set chi-squared minimization and geometry optimization was performed using the DL-Find software package.[62] Initial parameterization steps were performed with the optimized gas phase geometry, as simultaneous optimization of solvation model parameters and molecular geometry can be computationally prohibitive. However, after initial parameterization, geometries and parameters were allowed to relax interchangeably until convergence was reached. In the final step of optimization, both geometries and parameters were relaxed simultaneously in the neighborhood of the previously obtained values. The optimized parameters are listed in Table I. The training set mean signed error (MSE), mean unsigned error (MUE), and root mean squared error (RMSE) for the SCOSMO and VR-SCOSMO models are listed in Table II and model performance, as compared to experiment, is shown in Figure 3.

After parameter optimization, the SCOSMO and VR-SCOSMO models were evaluated on a separate test set of molecules, consisting of a total of 65 molecules (34 neutral, 17 anionic, 14 cationic) taken from Ref. 20. A table of these molecules can be found in the supplementary material, Table S2.[81] These molecules were used in the parameterization of the SCC-DFTBPR implicit solvation model developed by Hou *et al.*, which utilized a charge-dependent non-linear PB model implemented with an earlier variant of the DFTB3 semiempirical method.[63,64] Comparisons of SCOSMO, VR-SCOSMO, and this reference model's performance are tabulated in Table III. Similarly, performance of these models as compared to experimental values can be found in Figure 4.

E. Phosphate hydrolysis and phosphoryl transfer reactions

Using the parameters obtained from absolute solvation free energy optimization, the VR-SCOSMO model was used to examine various biologically motivated reactions. A series of

TABLE I. Optimized COSMO and VRSCOSMO parameters. All values are in atomic units. The surface tension parameters γ for SCOSMO and VRSCOSMO are 5×10^{-6} a.u. and 7.5×10^{-6} a.u., respectively.

Element	SCOSMO	VR-SCOSMO			
	R_{rad}	R_{rad}	$q^{(0)}$	α	β
C	4.35	3.75	−0.0339	−0.2493	0.0000
H	1.90	2.20	0.0954	−0.0502	−0.1794
O	2.85	3.19	−0.7056	0.0179	−0.4350
N	2.65	3.17	−0.2450	0.0128	−0.0000
P	5.30	5.30	−0.2000	0.0000	0.0000

TABLE II. Absolute solvation energy training set error statistics (kcal/mol).

Type	VR-SCOSMO			SCOSMO		
	MSE	MUE	RMSE	MSE	MUE	RMSE
Neutral	−0.0	1.4	1.9	0.4	1.2	1.7
Anion	0.8	2.5	3.0	3.9	5.6	6.3
Cation	1.4	2.8	4.3	0.7	4.6	5.7
Cation[a]	−0.1	1.5	1.9	−1.0	3.5	3.9
Total	0.2	1.6	2.4	0.8	2.0	3.0

[a]12 molecules. Excludes the compounds where the sum of bond orders around any oxygen is ≥3. Not included in the "total" statistics.

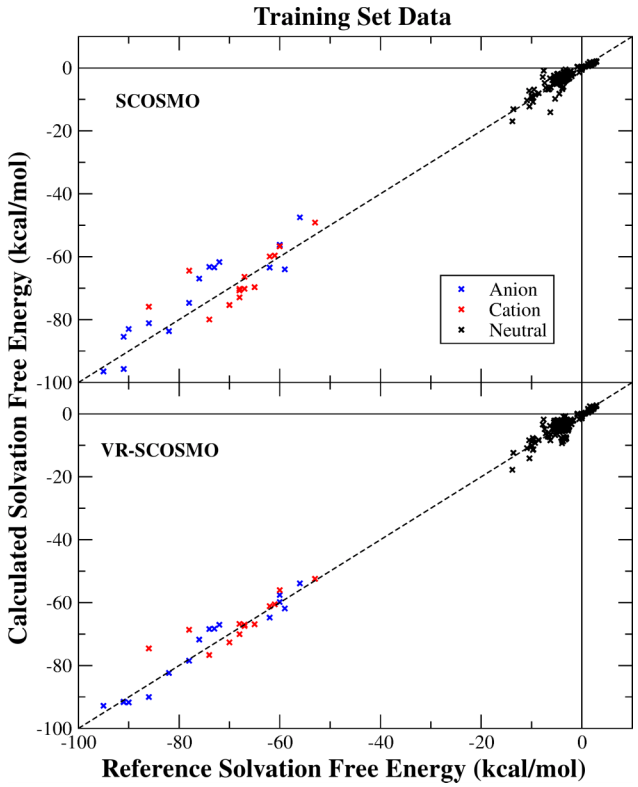


FIG. 3. Comparison between the SCOSMO and VR-SCOSMO absolute solvation free energy values with the reference values used in the training set.

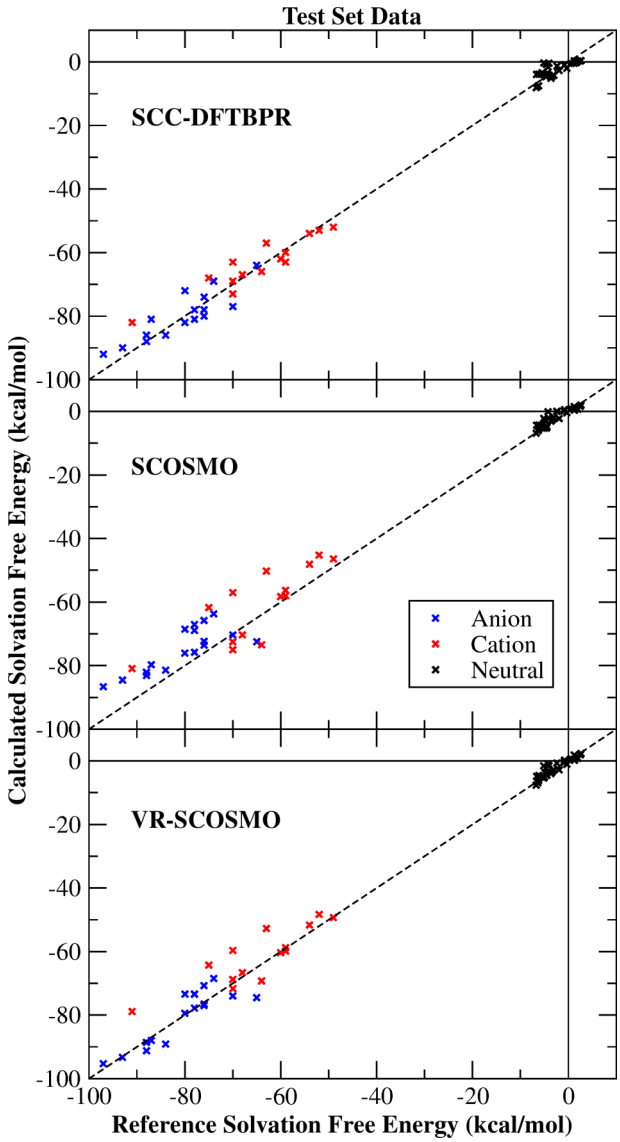


FIG. 4. Comparison between SCOSMO and VR-SCOSMO modeled absolute solvation free energies with the reference values used in the test set. SCC-DFTBPR is taken from Ref. 20.

phosphoryl transfer reactions were performed which serve as informative models for reactive RNA systems. The cleavage of the phosphodiester bond in these hydrolysis/transfer reactions can be seen in many catalytic motifs: in the protein RNase A[65,66] and in the hairpin,[67,68] hammerhead,[69–71] *glm*S,[72,73] hepatitis delta virus[74,75] ribozymes. An understanding of the fundamental chemical pathway in these reactions can glean further insight into these complex biochemical processes.

Two-dimensional reaction Potential Energy Surfaces (PESs) were generated, with reaction coordinates defined as the distance between the attacking and leaving group oxygen to phosphorus center, respectively. Distances were constrained to the value of each reaction coordinate and the system was allowed to relax using the DL-Find geometry optimization software suite.[62] After which, energy landscapes were constructed using two-dimensional spline interpolation. Stationary point calculations were performed with no restraints to obtain reaction barriers and product state calculations were completed at infinite separation. No additional, higher-level corrections were made to the resulting energy surfaces.

TABLE III. Absolute solvation energy test set error statistics (kcal/mol).

	VR-SCOSMO		SCOSMO		SCC-DFTBPR[a]	
Type	MSE	MUE	MSE	MUE	MSE	MUE
Neutral	0.4	0.9	0.6	1.0	0.2	1.6
Anion	0.2	3.2	5.6	6.6	0.7	3.1
Cation	3.5	4.6	3.2	6.1	1.1	3.4
Cation[b]	0.0	1.6	−0.2	3.9	−1.4	1.8
Total	1.1	2.3	2.4	3.6	0.5	2.4

[a]Reference 20.
[b]10 molecules. Excludes the compounds where the sum of bond orders around any oxygen is ≥3. This row is not included in the "total" statistics.

Two different types of phosphoryl reactions were modeled: the phosphate hydrolysis reaction of trimethylphosphate (TMP) and a series of phosphoryl transesterification transfer reactions corresponding to the self-attack of the O2' oxygen on the phosphate center in methyloxyribose (MOR) with seven different leaving groups [HO^-, CH_3O^-, $CH_3CH_2O^-$, $CH_3CH_2CH_2O^-$, $(CH_3)_2CHO^-$, CH_3COO^-, and PhO^-]. The TMP PES is shown in Figure 5 and the MOR PESs are shown in Figure 6. Table IV compares the modeled free energy barriers of the MOR reactions to previously calculated and experimental values and Figure 7 displays geometric information for the MOR transition states. For these reactions, a steady-state

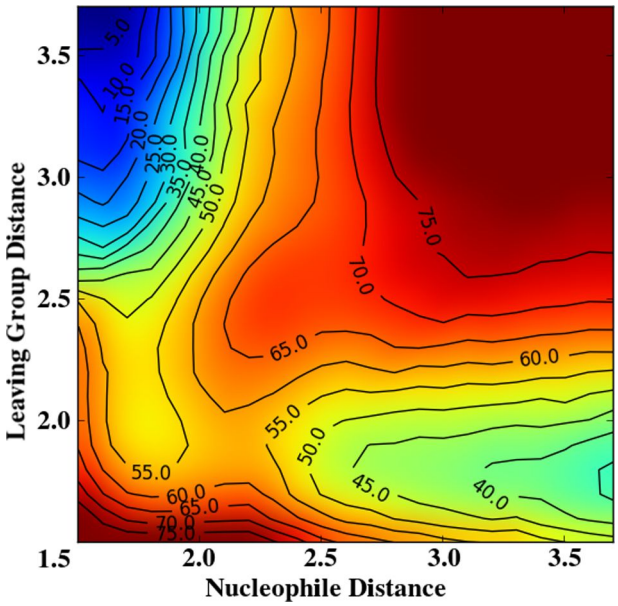


FIG. 5. DFTB3+VR-SCOSMO modeled potential energy surface of the trimethylphosphate hydrolysis reaction. Energies and distances are in kcal/mol and Å, respectively.

approximation was made in which the O2' oxygen of MOR is assumed to be deprotonated, and that the reaction is following a pseudo-first-order relate law with a transmission coefficient equal to one. According to the reaction scheme

$$\mathrm{MOR-OH} \rightleftharpoons \mathrm{MOR-O^-} \Rightarrow \mathrm{products} \qquad (38)$$

it is assumed that, as these reactions were carried out under basic conditions, modeling the reaction starting from the deprotonated state is sufficient.

III. RESULTS AND DISCUSSION

A. VR-SCOSMO model parameterization

Initial tests performed during the preparation of this manuscript found that the standard monopole representation of the DFTB3/3OB electrostatic potential encountered difficulties in distinguishing different functional groups. Large, systematic errors were observed in the absolute solvation free energy between molecules containing oxygen atoms with different numbers of lone pairs, such as when comparing alcohol and acid functional groups. Additionally, conjugated rings were frequently found to produce an electrostatic potential which, at the ring center, was too positive when compared to a B3LYP reference. The implicit solvent model response is caused by the electrostatic potential of the solute; therefore, an auxiliary set of atomic multipole moments from the DFTB3 density matrix was constructed using the prescription developed in Ref. 54. The inclusion of these multipole moments leads to the alleviation or complete elimination of these systematic errors, thus informing the decision to abandon parameterization efforts using the conventional DFTB3 monopole approximation.

Overall, the VR-SCOSMO model performs well for a vast majority of the molecules in the training set, shown in Table II and Figure 3. The SCOSMO non-charge-dependent

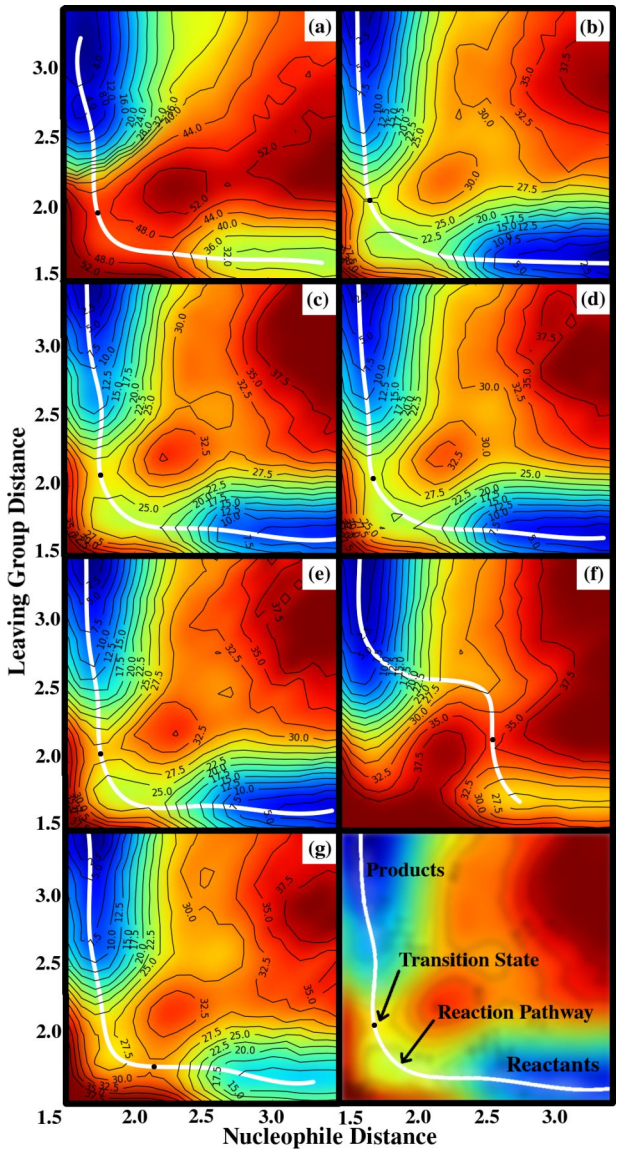


FIG. 6. DFTB3+VR-SCOSMO modeled potential energy surfaces for the phosphoryl transesterification reactions with a series of leaving groups. Energies and distances are in kcal/mol and Å, respectively. The leaving groups are (a) HO^-, (b) CH_3O^-, (c) $CH_3CH_2O^-$, (d) $CH_3CH_2CH_2O^-$, (e) $(CH_3)_2CHO^-$, (f) CH_3COO^-, and (g) PhO^-. Note that the color scale used in (a) is different from the other plots.

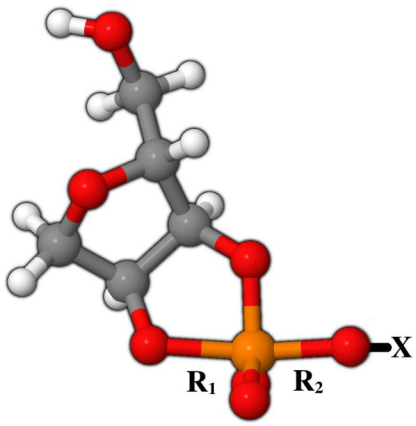
and VR-SCOSMO charge-dependent parameterizations show similar performance for neutral molecules, having MUEs of approximately 1 kcal/mol. However, the VR-SCOSMO model presents MUE 2-3 kcal/mol lower for the charged species. For both models, the largest errors occur in cationic molecules containing oxygen atoms whose sum of formal charge is +1, that is, molecules containing the "additionally protonated" alcohol compounds or molecules containing a protonated oxygen which also contains a double bond. These compounds are systematically undersolvated by approximately 10 kcal/mol, which is likely due to the over-simplification of proton chemical bonding in the absence of explicit solvent. Better estimations of solvation free energies could likely be obtained if explicit water molecules were included as part of the solute. Due to this logic, these types of compounds were assigned a smaller weight during chi-squared

TABLE IV. Reaction barriers for the self-attack of MOR with different leaving groups. All values are kcal/mol. The experimental barriers for the leaving groups with 2-hydroxypropyl phosphate (Expt. I) and uridine 3′-phosphate (Expt. II) are shown due to the lack of available experimental data with MOR. Calculations performed with DFTB3+VR-SCOSMO with no additional, higher-level corrections.

	Calculated		Experiment	
Model/Reference	VR-SCOSMO	PCM[a]	Expt. I[b]	Expt. II[c]
HO^-	24.90	…	…	…
CH_3O^-	28.10	24.43	27.09	…
$CH_3CH_2O^-$	26.60	25.84	28.55	23.37
$CH_3CH_2CH_2O^-$	27.38	24.26	…	…
$(CH_3)_2CHO^-$	28.87	25.96	30.76	26.02
CH_3COO^-	8.19	11.91	…	…
PhO^-	22.72	14.76	22.62	…

[a]M06-2X/6-311++G(3df,2p)//6-311++G(d,p) with PCM implicit solvation. Ref. 19.
[b]2-hydroxypropyl phosphate. Ref. 76.
[c]Uridine 3′-phosphate. Ref. 77.

minimization as to not bias the parameterization into over-correcting these errors. For comparison purposes, additional solvation free energy data for the more stable cations, bereft of these compounds, have been added to Tables II and III labeled with a and b, respectively. Additional issues were present in the nitrogen containing compounds which, as a whole, tended to contribute the largest difference from experiment in the solvation free energies tested. Detailed information of individual molecule absolute solvation free energy values can be found in the supplementary material.[81]



	TS 1			TS 2			Exp.
Leaving Group	R_1	R_2	$\Delta G^‡$	R_1	R_2	$\Delta G^‡$	pK_a
HO^-	2.05	1.72	23.6	1.75	2.00	24.9	15.7
CH_3O^-	2.02	1.73	26.0	1.75	2.07	28.1	15.5
$CH_3CH_2O^-$	2.02	1.72	23.8	1.74	2.08	26.6	16.0
$CH_3CH_2CH_2O^-$	2.00	1.73	24.9	1.75	2.06	27.4	16.1
$(CH_3)_2CHO^-$	2.01	1.72	25.3	1.75	2.08	28.9	17.1
CH_3COO^-	2.60	2.16	8.2	–	–	–	4.46
PhO^-	2.12	1.77	22.7	–	–	–	9.95

FIG. 7. The model phosphoryl transfer reaction with different leaving groups. Shown is the methyloxyribose ring, which undergoes an internal nucleophilic attack to model the backbone cleavage in RNA. Also shown are the attacking and leaving oxygen bond lengths, R_1 and R_2, respectively, computed with DFTB3+VR-SCOSMO for each leaving group, X–O$^-$. TS 1 and TS 2 are the early and late transition states. Experimental pK_a values are taken from Ref. 78.

After parameter optimization the SCOSMO and VR-SCOSMO models were evaluated against a separate test set of molecules. Absolute solvation data from these calculations can be seen in Table III and Figure 4. These results are also compared to those generated by Hou *et al.*,[20] who developed a similar charge-dependent nonlinear PB model, SCC-DFTBPR. The molecules within this test set were used as part of the SCC-DFTBPR training set and do not appear within this manuscript's training set. Again, the largest errors occur for the cationic compounds containing unusual oxygen protonation states. As expected, the SCC-DFTBPR model performs better for these unusual cases, as a significant number of these compounds appeared in the training set for this model. However, if these compounds where removed, the MSE and MUE of the test set for VR-SCOSMO would be 0.3 kcal/mol and 1.6 kcal/mol, respectively, while the SCC-DFTBPR model would report MSEs and MUEs of 0.1 kcal/mol and 2.1 kcal/mol, respectively. Both VR-SCOSMO and SCC-DFTBPR perform significantly better than the SCOSMO parameterization.

The improved accuracy of the VR-SCOSMO model does come at an additional computational cost. Energy calculations with VR-SCOSMO require reconstruction of the surface element mesh at each iteration of the SCF cycle as well as re-evaluation of the electrostatic potential at the surface elements and solution of the solvation polarization response density on the surface. Unlike a fixed surface mesh that can be inverted once and formulated as a Green's function term added directly to the Fock matrix, the VR-SCOSMO method requires inversion of the surface element interaction matrix each time the surface elements change during the SCF. Alternatively, the solution of the solvation polarization response density can be achieved by linear algebraic methods (e.g., conjugate gradient minimization), which is more practical for large systems, in which case the overhead of the VR-SCOSMO method is much less. Relative timings for the small systems studied here varied greatly, but the VR-SCOSMO/SCOSMO timing ratio was generally around 4:1. It is observed that the number of SCF cycle iterations generally does not change when using either the VR-SCOSMO or SCOSMO models, supporting the thought that the increase in computational cost is believed to largely emerge from the necessary increase in calculation complexity.

B. Phosphoryl transfer reactions

Phosphoryl transfer reactions are ubiquitous in biology[7,79] and have been studied extensively with quantum chemistry calculations and implicit solvation models.[4,6,15–20] Two different types of phosphoryl reactions were used to test the phosphate hydrolysis applicability of the DFTB3+VR-SCOSMO framework to RNA-like systems, a phosphate hydrolysis reaction, and a series of phosphoryl transesterification transfer reactions with different leaving groups. These reactions have a large amount of local charge transfer, totaling a net $2e$ charge, which makes them ideal candidates for use with a charge-dependent model with an aim of accurately

modeling biological reactions. These reactions in particular serve as analogs for a key step in the self-cleavage of the RNA backbone seen in many ribozymes.

The two-dimensional reaction PESs for hydrolysis of TMP can be seen in Figure 5. The energy barrier is found to be 23.19 kcal/mol which is in good agreement with the experimentally known value of 24.6 kcal/mol.[80,81] The model predicts an "early" transition state; that is, it forms the attacking bond P–O in advance of the leaving group bond cleavage. This observation is consistent with previous computational studies of this system.[20]

Similarly, DFTB3+VR-SCOSMO reaction surfaces for the phosphoryl transesterification reactions of MOR with different leaving groups can be found in Figure 6. A summary of the reaction barriers as compared to previously calculated *ab initio* barriers using the PCM with UFF radii and experimentally known barriers can be found in Table IV. Additionally, Figure 7 details various transition state geometries for the modeled reactions as well as listing non-rate-limiting reaction barriers for those reactions where two separate transition states were observed. VR-SCOSMO attack barriers compare favorably to those calculated with the M06-2X/6-311++G(3df,2p)//M06-2X/6-311++G(d,p) PCM model[19] as well as to the known experiments.[76,77] It should be noted, however, that the experiments presented here examine reactions with sightly different chemical structures than those which have been computed in this manuscript. Specifically, Expt. I[76] performed experiments with 2-hydroxypropyl phosphate while Expt. II[77] examined various esters of uridine 3′-phosphate. However, despite these structural differences, it is thought that the reaction barriers should not be significantly different and general trends of the data should remain fairly consistent.

IV. CONCLUSION

A charge-dependent, variable radii variant of the SCOSMO implicit solvation model is presented. The model, VR-SCOSMO, is parameterized using the DFTB3/3OB-OPhyd semiempirical Hamiltonian, while using an auxiliary set of atomic multipoles to interact the DFTB3 solute with the implicit solvent model. As a baseline for comparison, a traditional SCOSMO solvation is similarly parameterized. The two methods produce similar neutral molecule solvation energies; however, VR-SCOSMO is shown to provide a significant 2-3 kcal/mol improvement in MUE for charged species. Further comparisons were made to a charge-dependent nonlinear PB implicit solvation model,[20] with VR-SCOSMO showing similar accuracy. Example DFTB3+VR-SCOSMO applications are performed on phosphoryl hydrolysis and phosphoryl transesterification reactions. The reaction barriers were shown to agree well with available experiment and previous *ab initio* calculations. The VR-SCOSMO implicit solvation model serves as a efficient technique to probe biochemical reaction landscapes to understand their mechanistic pathways. Its advantages are in producing smooth potential energy surfaces while being able to dynamically adjust itself to changes in the local charge state of atoms over the course of a reaction.

ACKNOWLEDGMENTS

The authors are grateful for financial support provided by the National Institutes of Health (No. GM107485 to DY). Computational resources were provided by the Minnesota Supercomputing Institute for Advanced Computational Research (MSI) and Rutgers, the State University of New Jersey.

[1]S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, J. Phys. Chem. B **113**, 1253 (2009).
[2]V. H. Teixeira, C. M. Soares, and A. M. Baptista, Proteins **70**, 1010 (2008).
[3]A. Pabis, I. Geronimo, D. M. York, and P. Paneth, J. Chem. Theory Comput. **10**, 2246 (2014).
[4]Y. Liu, B. A. Gregersen, A. Hengge, and D. M. York, Biochemistry **45**, 10043 (2006).
[5]K. Range, M. J. McGrath, X. Lopez, and D. M. York, J. Am. Chem. Soc. **126**, 1654 (2004).
[6]H. Chen, T. J. Giese, M. Huang, K.-Y. Wong, M. E. Harris, and D. M. York, Chem. - Eur. J. **20**, 14336 (2014).
[7]S. C. L. Kamerlin, P. K. Sharma, R. B. Prasad, and A. Warshel, Q. Rev. Biophys. **46**, 1 (2013).
[8]X. Zhang, D. H. T. Harrison, and Q. Cui, J. Am. Chem. Soc. **124**, 14871 (2002).
[9]T.-S. Lee, K.-Y. Wong, G. M. Giambasu, and D. M. York, "Bridging the gap between theory and experiment to derive a detailed understanding of hammerhead ribozyme catalysis," Prog. Mol. Biol. Transl. Sci. **120**, 25–91 (2013).
[10]B. K. Radak, T.-S. Lee, M. E. Harris, and D. M. York, RNA **21**, 1566 (2015).
[11]M. T. Panteva, T. Dissanayake, H. Chen, B. K. Radak, E. R. Kuechler, G. M. Giambaşu, T.-S. Lee, and D. M. York, in *Multiscale Methods for Computational RNA Enzymology* (Elsevier, 2015), Chap. 14.
[12]P. V. Klimovich, M. R. Shirts, and D. L. Mobley, J. Comput.-Aided Mol. Des. **29**(5), 397–411 (2015).
[13]J. Mortier, C. Rakers, M. Bermudez, M. S. Murgueitio, S. Riniker, and G. Wolber, Drug Discovery Today **20**(6), 686–702 (2015).
[14]L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood *et al.*, J. Am. Chem. Soc. **137**, 2695 (2015).
[15]J. Florián and A. Warshel, J. Phys. Chem. B **102**, 719 (1998).
[16]J. Florián, J. Åqvist, and A. Warshel, J. Am. Chem. Soc. **120**, 11524 (1998).
[17]X. Lopez, D. M. York, A. Dejaegere, and M. Karplus, Int. J. Quantum Chem. **86**, 10 (2002).
[18]Y. Liu, B. A. Gregersen, X. Lopez, and D. M. York, J. Phys. Chem. B **109**, 19987 (2005).
[19]M. Huang and D. M. York, Phys. Chem. Chem. Phys. **16**, 15846 (2014).
[20]G. Hou, X. Zhu, and Q. Cui, J. Chem. Theory Comput. **6**, 2303 (2010).
[21]M. W. van der Kamp and A. J. Mulholland, "Computational enzymology: Insight into biological catalysts from modelling," Nat. Prod. Rep. **25**(6), 1001–1014 (2008).
[22]E. R. Kuechler and D. M. York, J. Chem. Phys. **140**, 054109 (2014).
[23]C. J. Cramer and D. G. Truhlar, Chem. Rev. **99**, 2161 (1999).
[24]B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).
[25]T. N. Truong, Int. Rev. Phys. Chem. **17**, 525 (1998).
[26]W. Im, J. Chen, and C. L. Brooks III, Adv. Protein Chem. **72**, 173 (2006).
[27]H. M. Senn and W. Thiel, Angew. Chem., Int. Ed. **48**, 1198 (2009).
[28]H. M. Senn and W. Thiel, Top. Curr. Chem. **268**, 173 (2007).
[29]J. Gao and D. G. Truhlar, Annu. Rev. Phys. Chem. **53**, 467 (2002).
[30]M. Kläln, E. Rosta, and A. Warshel, J. Am. Chem. Soc. **128**, 15310 (2006).
[31]K. Sharp and B. Honig, J. Phys. Chem. **94**, 7684 (1990).
[32]C. P. Kelly, C. J. Cramer, and D. G. Truhlar, J. Chem. Theory Comput. **1**, 1177 (2005).
[33]D. J. Geisen, G. D. Hawkins, D. A. Liotard, C. J. Cramer, and D. G. Truhlar, Theor. Chem. Acc. **98**, 85 (1997).
[34]C. J. Cramer and D. G. Truhlar, Science **256**, 213 (1992).
[35]A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2 **2**, 799 (1993).
[36]A. Klamt, J. Phys. Chem. **99**, 2224 (1995).
[37]A. Klamt, F. Eckert, M. Diedenhofen, and M. E. Beck, J. Phys. Chem. A **107**, 9380 (2003).
[38]F. Eckert and A. Klamt, J. Comput. Chem. **27**, 11 (2006).
[39]F. Eckert and A. Klamt, AIChE J. **48**, 369 (2002).
[40]A. Schäfer, J. C. W. L. Andreas Klamt, D. Sattel, and F. Eckert, Phys. Chem. Chem. Phys. **2**, 2187 (2000).

[41]D. A. Liotard, G. D. Hawkins, G. C. Lynch, C. J. Cramer, and D. G. Truhlar, J. Comput. Chem. **16**, 422 (1995).
[42]D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still, J. Phys. Chem. A **101**, 3005 (1997).
[43]R. A. Pierotti, Chem. Rev. **76**, 717 (1976).
[44]S. Miertuš, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).
[45]A. W. Lange and J. M. Herbert, J. Chem. Phys. **133**, 244111 (2010).
[46]D. M. York and M. Karplus, J. Phys. Chem. A **103**, 11060 (1999).
[47]G. Scalmani and M. J. Frisch, J. Chem. Phys. **132**, 114110 (2010).
[48]B. Ginovska, D. M. Camaioni, M. Dupuis, C. A. Schwerdtfeger, and Q. Gil, J. Phys. Chem. A **112**, 10604 (2008).
[49]M. Aguilar, M. Martin, S. Tolosa, and F. Olivares Del Valle, J. Mol. Struct.: THEOCHEM **166**, 313 (1988).
[50]M. Aguilar and F. Olivares Del Valle, Chem. Phys. **129**, 439 (1989).
[51]M. Gaus, Q. Cui, and M. Elstner, J. Chem. Theory Comput. **7**, 931 (2011).
[52]M. Gaus, A. Goez, and M. Elstner, J. Chem. Theory Comput. **9**, 338 (2013).
[53]M. Gaus, X. Lu, M. Elstner, and Q. Cui, J. Chem. Theory Comput. **10**, 1518 (2014).
[54]T. J. Giese, H. Chen, M. Huang, and D. M. York, J. Chem. Theory Comput. **10**, 1086 (2014).
[55]T. J. Giese, M. T. Panteva, H. Chen, and D. M. York, J. Chem. Theory Comput. **11**, 451 (2015).
[56]D. M. York, T.-S. Lee, and W. Yang, Chem. Phys. Lett. **263**, 297 (1996).
[57]B. A. Gregersen, J. Khandogin, W. Thiel, and D. M. York, J. Phys. Chem. B **109**, 9810 (2005).
[58]J. Khandogin, B. A. Gregersen, W. Thiel, and D. M. York, J. Phys. Chem. B **109**, 9799 (2005).
[59]V. I. Lebedev, Ž. Vyčisl. Mat. i Mat. Fiz. **16**, 293 (1976).
[60]V. I. Lebedev, Sibirsk. Mat. Ž. **18**, 99 (1977).
[61]B. A. Gregersen and D. M. York, J. Chem. Phys. **122**, 194110 (2005).
[62]J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, and P. Sherwood, J. Phys. Chem. A **113**, 11856 (2009).
[63]M. Elstner, J. Phys. Chem. A **111**, 5614 (2007).
[64]Y. Yang, H. Yu, D. M. York, Q. Cui, and M. Elstner, J. Phys. Chem. A **111**, 10861 (2007).
[65]E. Formoso, J. M. Matxain, X. Lopez, and D. M. York, J. Phys. Chem. B **114**, 7371 (2010).
[66]H. Gu, S. Zhang, K.-Y. Wong, B. K. Radak, T. Dissanayake, D. L. Kellerman, Q. Dai, M. Miyagi, V. E. Anderson, D. M. York *et al.*, Proc. Natl. Acad. Sci. U. S. A. **110**, 13002 (2013).
[67]G. M. Giambaşu, D. M. York, and D. A. Case, RNA **21**, 963 (2015).
[68]H. Heldenbrand, P. A. Janowski, G. Giambaşu, T. J. Giese, J. E. Wedekind, and D. M. York, J. Am. Chem. Soc. **136**, 7789 (2014).
[69]W. G. Scott, *What Can the New Hammerhead Ribozyme Structures Teach us About Design?* (Springer-Verlag, Berlin Heidelberg, 2010), pp. 305–323.
[70]W. G. Scott, Biol. Chem. **388**, 727 (2007).
[71]W. G. Scott, J. B. Murray, J. R. P. Arnold, B. L. Stoddard, and A. Klug, Science **274**, 2065 (1996).
[72]S. Zhang, A. Ganguly, P. Goyal, J. L. Bingaman, P. C. Bevilacqua, and S. Hammes-Schiffer, J. Am. Chem. Soc. **137**, 784 (2015).
[73]J. Viladoms and M. J. Fedor, J. Am. Chem. Soc. **134**, 19043 (2012).
[74]P. Thaplyal, A. Ganguly, S. Hammes-Schiffer, and P. C. Bevilacqua, Biochemistry **54**, 2160 (2015).
[75]A. R. Ferré-D'Amaré and W. G. Scott, Cold Spring Harbor Perspect. Biol. **2**, a003574 (2010).
[76]D. M. Brown and D. A. Usher, J. Chem. Soc. **87**, 6558 (1965).
[77]M. Kosonen, E. Youseti-Salakdeh, R. Strömberg, and H. Lönnberg, J. Chem. Soc., Perkin Trans. 2 **1997**, 2661–2666.
[78]E. P. Serjeant and B. Dempsey, *Ionisation Constants of Organic Acids in Aqueous Solution* (Pergamon Press, New York, 1979).
[79]J. K. Lassila, J. G. Zalatan, and D. Herschlag, Annu. Rev. Biochem. **80**, 669 (2011).
[80]P. W. Barnard, C. A. Bunton, D. R. Llewellyn, C. A. Vernon, and V. A. Welch, J. Chem. Soc. **1961**, 2670–2676.
[81]See supplementary material at http://dx.doi.org/10.1063/1.4946779 for all materials stated to be in the supplementary material.