

Testing a Heterogeneous Polarizable Continuum Model against Exact Poisson Boundary Conditions

Paige E. Bowling,[§] Montgomery Gray,[§] Suranjan K. Paul, and John M. Herbert*

 Cite This: *J. Chem. Theory Comput.* 2025, 21, 1722–1738

 Read Online

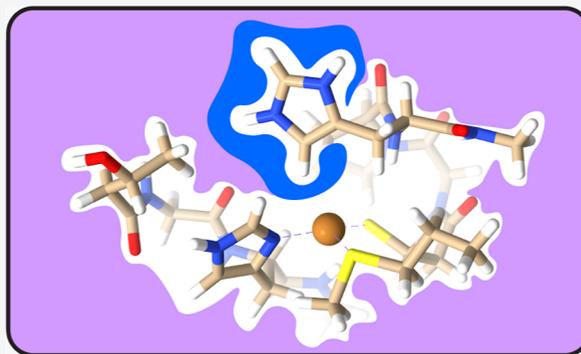
ACCESS |

 Metrics & More

 Article Recommendations

 Supporting Information

ABSTRACT: The polarizable continuum model (PCM) is a computationally efficient way to incorporate dielectric boundary conditions into electronic structure calculations, via a boundary-element reformulation of Poisson's equation. This transformation is only rigorously valid for an isotropic dielectric medium. To simulate anisotropic solvation, as encountered at an interface or when parts of a system are solvent-exposed while other parts are in a nonpolar environment, *ad hoc* modifications to the PCM formalism have been suggested, in which a dielectric constant is assigned separately to each atomic sphere that contributes to the solute cavity. The accuracy of this “heterogeneous” PCM (HetPCM) method is tested here for the first time, by comparison to results from a generalized Poisson equation solver. The latter is a more expensive and cumbersome approach to incorporate arbitrary dielectric boundary conditions, but one that corresponds to a well-defined scalar permittivity function, $\epsilon(\mathbf{r})$. We examine simple model systems for which a function $\epsilon(\mathbf{r})$ can be constructed in a manner that maps reasonably well onto a dielectric constant for each atomic sphere, using a solvent-exposed dielectric constant $\epsilon_{\text{solv}} = 78$ and a range of smaller values to represent hydrophobic environments. For nonpolar dielectric constants $\epsilon_{\text{nonp}} \leq 2$, differences between the HetPCM and Poisson solvation energies are large compared to the effect of anisotropy on the solvation energy. For $\epsilon_{\text{nonp}} = 4$ and $\epsilon_{\text{nonp}} = 10$, however, HetPCM and anisotropic Poisson solvation energies agree to within 2 kcal/mol in most cases. As a realistic use case, we apply the HetPCM method to predict solvation energies and pK_a values for blue copper proteins. The HetPCM method affords pK_a values that are more in line with experimental results as compared to either gas-phase calculations or homogeneous (isotropic) PCM results.



1. INTRODUCTION

The polarizable continuum model^{1–3} (PCM) is a popular approach for incorporating solvation into electronic structure calculations, which describes electrostatic and polarization interactions between an atomistic solute and its continuum environment. Although accurate modeling of solvation energies requires that this model be augmented by non-electrostatic contributions,^{3–5} what the PCM does correctly is to furnish boundary conditions for an electronic structure calculation that are superior to gas-phase boundary conditions. For typical solutes that are amenable to quantum mechanics (QM) calculations, a PCM contributes only modest computational overhead. Furthermore, linear-scaling PCM algorithms have been developed for much larger solutes,^{2,3,6–11} and PCM calculations on full proteins have been reported.^{10–14} Elimination of explicit solvent is beneficial when “cluster-QM” models are used to investigate enzymatic reaction mechanisms,^{15–17} because truncated models can be susceptible to the multiple-minimum problem¹⁸ when molecular dynamics is eschewed in favor of geometry optimization. Use of continuum boundary conditions might provide opportunities for additional truncation, further limiting the appearance of multiple minima. For large protein models, however, it is not

always clear that an *isotropic* dielectric medium is the appropriate choice of boundary conditions. The heterogeneous PCM (HetPCM) approach described herein is an attractive alternative for this purpose.

The PCM approach is based on a reformulation of Poisson's equation for a sharp dielectric interface between an atomistic solute region (characterized by relative electric permittivity $\epsilon = 1$) and a continuum solvent that is characterized by permittivity ϵ_s .³ This corresponds to an electric permittivity function

$$\epsilon(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in \Omega \\ \epsilon_s, & \mathbf{r} \notin \Omega \end{cases} \quad (1)$$

Received: December 6, 2024

Revised: January 23, 2025

Accepted: January 31, 2025

Published: February 17, 2025



where Ω represents a molecule-sized cavity. Whereas solution of Poisson's equation requires discretization of the solute's charge density and electrostatic potential throughout three-dimensional space,^{19–28} the PCM only requires the electrostatic potential to be evaluated on the surface of the cavity that defines Ω , the interface with the continuum. The reformulation of Poisson's equation is exact if the entirety of the solute's charge density is contained within Ω ,³ but is highly accurate even when the tails of the charge distribution penetrate into the dielectric medium,^{29–31} as they do in any QM calculation with a realistic, molecule-sized cavity.

Derivation of the PCM starting from Poisson's equation is predicated on the assumption that the dielectric medium is isotropic. To describe *anisotropic* solvation, such as an air/water interface or a water/biomolecule interface, one can always resort to solution of a generalized Poisson equation with a permittivity function $\epsilon(\mathbf{r})$ that is defined pointwise throughout three-dimensional space.^{24–27} This approach is more common in plane-wave electronic structure codes,³² where solution of Poisson's equation is already a part of the standard computational machinery, but in localized Gaussian orbital codes it sacrifices the efficiency advantages of the PCM and the computational overhead to solve Poisson's equation is not negligible.

There have been various efforts to extend PCMs to anisotropic solvation environments without sacrificing simplicity and computational expedience. Some approaches with good formal properties, but which lack the generality that we are seeking, include methods that generalize the dielectric “constant” to a 3×3 tensor,^{33–36} as appropriate for liquid crystals, and other methods that modify the Green's function for the Coulomb potential appearing in Poisson's equation, for the case of a two-dimensional interface.^{37–44} However, the most convenient (and potentially general) extension is the heterogeneous model that we call HetPCM, in which each atomic sphere that is used to construct the solute cavity is assigned its own dielectric constant.^{45–47} In this way, one might hope to provide appropriate boundary conditions for a protein (for example), in which certain residues are exposed to the aqueous solvent ($\epsilon = 78$) while others are buried in the protein's hydrophobic interior. For the latter environment, values $\epsilon \approx 4$ are often used in classical biomolecular electrostatics calculations,^{48–53} e.g., to compute pK_a values.^{52–62} In some cases, larger values have been used for the nonpolar dielectric constant, up to $\epsilon = 10–20$.^{53–55,63–67}

The HetPCM approach has an appealing simplicity and would be easy to combine with fragment-based methods that can be used to extend the reach of quantum chemistry to proteins.^{12–14,68} To the best of our knowledge, however, this model is not derivable starting from a well-defined permittivity model $\epsilon(\mathbf{r})$, in contrast to the original (homogeneous or isotropic) PCM. Thus, HetPCM has been introduced as an *ad hoc* modification of the original model, which has not been rigorously tested against exact continuum electrostatics theory. We do so in the present work, using a generalized Poisson equation solver^{24–26} (PEqS) to provide a benchmark result for the solute–continuum polarization energy associated with any model permittivity function. This function, $\epsilon(\mathbf{r})$, is defined pointwise in three-dimensional space, allowing different spatial domains to have different permittivities. Crucial to this testing regiment is the construction of model systems for which the function $\epsilon(\mathbf{r})$ unambiguously places each continuum-exposed atomic sphere into a region where the value of ϵ is

approximately constant, such that the model function $\epsilon(\mathbf{r})$ can be used to assign a permittivity to each atomic sphere. We accomplish this using model functions $\epsilon(\mathbf{r})$ based on Voronoi cells, which can be mapped onto atomic dielectric constants.

2. THEORY

2.1. PCM Formalism. Theoretical underpinnings of the PCM approach have been reviewed recently.³ The underlying physical model is based upon a generalized form of Poisson's equation,

$$\hat{\nabla} \cdot [\epsilon(\mathbf{r}) \hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})] = -4\pi\rho_{\text{sol}}(\mathbf{r}) \quad (2)$$

in which $\rho_{\text{sol}}(\mathbf{r})$ is the charge density of the atomistic solute, obtained herein from a quantum chemistry calculation, and $\varphi_{\text{tot}}(\mathbf{r})$ is the total electrostatic potential. The latter includes the potential generated by $\rho_{\text{sol}}(\mathbf{r})$ but also a contribution from polarizing the continuum. Gaussian electrostatic units are used in eq 2, such that $4\pi\epsilon_0 = 1$.

The original PCM solves the model problem defined by eq 2 and the sharp dielectric interface in eq 1, where $\mathbf{r} \in \Omega$ in eq 1 indicates the interior of the molecular cavity (see Figure 1a)

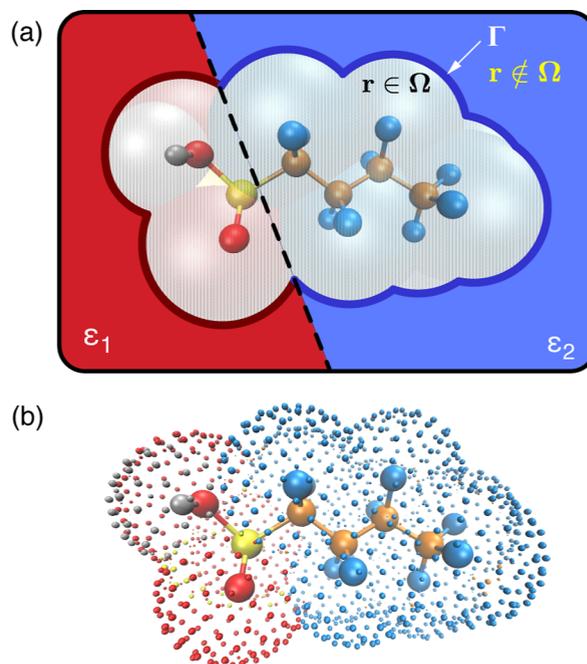


Figure 1. (a) Schematic depiction of the $\text{CF}_3(\text{CF}_2)_3\text{SO}_3\text{H}$ molecule in a heterogeneous solvation environment, such that the acidic SO_3H group and the perfluorocarbon tail are embedding in media with different dielectric constants, ϵ_1 and ϵ_2 . A solute cavity consisting of atom-centered spheres is indicated, whose boundary is denoted by Γ . Note that $\epsilon = 1$ within the solute cavity ($\mathbf{r} \in \Omega$). (b) Depiction of the surface quadrature grid points used for PCM calculations. The size of each discretization point s_i is an indication of its contribution a_i to the cavity surface area, as defined in eq 10.

and ϵ_s is the (static) dielectric constant of the solvent. In the special case where $\rho_{\text{sol}}(\mathbf{r})$ vanishes for $\mathbf{r} \notin \Omega$, the model defined by eqs 1 and 2 can be mapped onto an equivalent boundary-element or *apparent surface charge* (ASC) problem, defined at the cavity surface Γ that is indicated in Figure 1a.^{3,69} That remapping defines the PCM.³

The most fundamental version of this remapping has been called the *integral equation formulation* (IEF).^{34,69,70} For

classical solutes where there is no charge leakage into the continuum, IEF-PCM is an exact reformulation of isotropic Poisson boundary conditions, as can be demonstrated numerically.^{10,71} For QM charge densities, the tails of $\rho_{\text{sol}}(\mathbf{r})$ penetrate into the medium, to the tune of ~ 0.1 – 0.2 electrons for small molecules.⁷² However, an alternative derivation of the IEF-PCM equation demonstrates that this approach implicitly (albeit approximately) accounts for the volume polarization due to this escaped charge.³¹ Therefore, IEF-PCM is an accurate reformulation of Poisson boundary conditions even in the case of a QM solute.^{29–31}

In matrix form, the IEF-PCM equation is³

$$\mathbf{K}_\epsilon \mathbf{q} = \mathbf{Y}_\epsilon \mathbf{v} \quad (3)$$

The input is a vector \mathbf{v} containing the solute's electrostatic potential $\varphi_{\text{sol}}(\mathbf{s}_i)$, evaluated at a set of discretization points \mathbf{s}_i on the cavity surface (Figure 1b), while the output is a vector of surface charges \mathbf{q} at the same points. The charges $\{q_i\}$ describe the polarizing effect of the medium. Several other ASC-PCM methods can be cast in the form of eq 3,^{3,10,71} and IEF-PCM is defined by a particular choice of the matrices \mathbf{K}_ϵ and \mathbf{Y}_ϵ :

$$\mathbf{K}_\epsilon^{\text{IEF-PCM}} = \mathbf{S} - \frac{f_\epsilon}{2\pi} \mathbf{DAS} \quad (4)$$

and

$$\mathbf{Y}_\epsilon^{\text{IEF-PCM}} = -f_\epsilon \left(\mathbf{1} - \frac{1}{2\pi} \mathbf{DA} \right) \quad (5)$$

Here, \mathbf{A} is a diagonal matrix containing surface areas of individual discretization elements and

$$f_\epsilon = \frac{\epsilon_s - 1}{\epsilon_s + 1} \quad (6)$$

The matrices \mathbf{S} and \mathbf{D} are discretized forms of so-called single- and double-layer operators, \hat{S} and \hat{D} .³ The former generates the surface electrostatic potential, and its matrix representation \mathbf{S} consists of the Coulomb interaction between surface elements. The operator \hat{D}^\dagger generates the normal electric field at the cavity surface.³ The IEF-PCM version of Poisson's equation provides a theoretical basis for several other approaches to implicit solvation including Generalized Born models^{73,74} and Debye–Hückel theory.⁷⁵

Replacing \mathbf{DAS} in eq 4 with $(\mathbf{DAS} + \mathbf{SAD}^\dagger)/2$ affords the *surface and simulation of volume polarization for electrostatics* [SS(V)PE] method,^{31,76} which is formally equivalent to IEF-PCM at the level of integral equations.^{3,77} However, the SS(V)PE form is more sensitive to the quality of the surface discretization and may exhibit numerical artifacts at “crevices” between atomic spheres.⁷¹ These artifacts generally disappear for larger atomic spheres, which afford smoother cavities,³ but for this reason we will use the IEF-PCM form exclusively.

The *conductor-like* C-PCM method⁷⁸ is another popular model that is derivable from IEF-PCM in the limit $\epsilon_s \rightarrow \infty$.⁷⁵ It can be placed in the same form as eq 3 but with alternative definitions for the matrices \mathbf{K}_ϵ and \mathbf{Y}_ϵ .^{3,10,71} For C-PCM, these are

$$\mathbf{K}_\epsilon^{\text{C-PCM}} = \mathbf{S} \quad (7)$$

and

$$\mathbf{Y}_\epsilon^{\text{C-PCM}} = -\tilde{f}_\epsilon(\zeta) \mathbf{1} \quad (8)$$

where

$$\tilde{f}_\epsilon(\zeta) = \frac{\epsilon_s - 1}{\epsilon_s + \zeta} \quad (9)$$

The conventional choice for C-PCM is $\zeta = 0$ in eq 9, corresponding to the conductor limit of IEF-PCM. Other choices (chiefly $\zeta = 1/2$) are sometimes encountered but will not be used here. For $\epsilon_s \gtrsim 30$, C-PCM is numerically indistinguishable from IEF-PCM.⁷¹ Even for smaller values of ϵ_s the differences are modest, perhaps 1–2 kcal/mol in the electrostatic solvation energies of small molecules.^{71,75}

2.2. HetPCM. The IEF-PCM and C-PCM methods are defined by eq 3 and the dependence on ϵ_s is contained wholly within the factors f_ϵ and $\tilde{f}_\epsilon(\zeta)$ in eqs 6 and 9, respectively. The HetPCM approach to be tested here, which was introduced in ref 45 for IEF-PCM and in ref 46 for C-PCM, consists in modifying this factor to use a different value of ϵ_s for each atomic sphere. To fully specify the model, however, it is necessary to consider how we discretize the solute cavity surface.

We employ the switching/Gaussian (“SwiG”) discretization algorithm,^{3,10,71,79–81} which uses atom-centered Lebedev quadrature grids for each atomic sphere.^{82,83} Each surface point \mathbf{s}_i is assigned a Lebedev quadrature weight w_i^{Leb} and a switching amplitude F_i^{sw} with $0 \leq F_i^{\text{sw}} \leq 1$, as shown in Figure 1b. The nature of the switching functions is detailed elsewhere.^{10,79,80}

The size of each discretization point \mathbf{s}_i in Figure 1b corresponds to the surface area a_i that is assigned to that point, which is the diagonal entry of the matrix \mathbf{A} that was introduced in Section 2.1. For a discretization point \mathbf{s}_i on the surface of atom B , whose atomic radius is $R_{\text{vdW},B}$, we set

$$a_i = w_i^{\text{Leb}} F_i^{\text{sw}} R_{\text{vdW},B}^2 \quad (10)$$

For a single spherical cavity, where every point \mathbf{s}_i lies on the exterior of the cavity (so that $F_i^{\text{sw}} = 1$), this ensures a surface area of $4\pi R_{\text{vdW},B}^2$ since

$$\sum_i w_i^{\text{Leb}} = 4\pi \quad (11)$$

For molecular cavities, the switching weights F_i^{sw} rapidly but smoothly attenuate a_i as this point passes into the interior of the cavity.^{79,80} The total solvent-accessible surface area for atom B in a molecular solute cavity is

$$SA_B = R_{\text{vdW},B}^2 \sum_{i \in B} w_i^{\text{Leb}} F_i^{\text{sw}} \quad (12)$$

HetPCM modifies the PCM equations of Section 2.1 by defining f_ϵ or $\tilde{f}_\epsilon(\zeta)$ pointwise across the cavity surface. For IEF-PCM, the factor f_ϵ in eq 6 is replaced by

$$f_i = \frac{\epsilon_i - 1}{\epsilon_i + 1} \quad (13)$$

where $\epsilon_i = \epsilon(\mathbf{s}_i)$ is a dielectric constant for the surface element \mathbf{s}_i . For C-PCM, the factor $\tilde{f}_\epsilon(\zeta)$ in eq 9 is modified analogously. These are trivial modifications, in any existing implementation of C-PCM or IEF-PCM. From a user perspective, this simply entails specifying a dielectric constant for each atom. In the examples considered herein, we use only

two such values: $\epsilon_{\text{sol}} = 78$ for solvent-exposed parts of the cavity surface, and something smaller (ϵ_{nonp}) for hydrophobic portions of the surface. The latter will be varied up to $\epsilon_{\text{nonp}} = 10$ for testing.

The same SwiG discretization that ensures continuity of the potential energy surface also means that the energy gradient of HetPCM is well-defined and continuous. For truncated cluster-QM models of biomolecules, for which the method is intended, it is already necessary to constrain certain atoms in order to prevent collapse of the protein structure when the geometry is relaxed.^{84,85} As such, these are local rather than global optimizations and we do not foresee problems with the use of heterogeneous boundary conditions.

2.3. Poisson Boundary Conditions. The HetPCM approach is a simple but *ad hoc* modification of the ASC-PCM formalism, which we intend to test against rigorous Poisson boundary conditions that can describe an anisotropic continuum environment in a general way, by specifying a permittivity function $\epsilon(\mathbf{r})$.^{24–27} This flexibility facilitates the use of a heterogeneous (anisotropic) dielectric environment. The function $\epsilon(\mathbf{r})$ represents a model that can be used, for example, to describe the air/water interface,^{24,25,86} or different regions of a protein^{87–89} or other complex system.⁹⁰ Given a model $\epsilon(\mathbf{r})$, the corresponding Poisson boundary conditions are implemented in an exact way, up to controllable discretization errors. Such methods have a long history in classical biomolecular electrostatics calculations,^{91–99} and a variety of numerical solvers have been developed,^{92–99} but implementations of Poisson boundary conditions for electronic structure calculations have also been reported.^{21–27} The finite-difference algorithms that are typically used to solve eq 2 require the permittivity function $\epsilon(\mathbf{r})$ to be smoothly varying,^{100–102} else the induced charge may vary wildly in space and adequate discretization becomes a challenge.

In the present work, eq 2 is solved for densities $\rho_{\text{sol}}(\mathbf{r})$ from Hartree–Fock (HF) or density-functional theory (DFT) calculations, using the PEqS algorithm described previously.²⁵ We partition the electrostatic potential φ_{tot} into a contribution arising directly from the solute's charge density (φ_{sol}) and a polarization potential (φ_{pol}) arising from the charge induced in the continuum:

$$\varphi_{\text{tot}}(\mathbf{r}) = \varphi_{\text{sol}}(\mathbf{r}) + \varphi_{\text{pol}}(\mathbf{r}) \quad (14)$$

As compared to using $\rho_{\text{sol}}(\mathbf{r})$ directly,²⁴ we find that the algorithm is more stable if φ_{sol} is computed from the one-electron density matrix and electrostatic potential integrals and used as the basic variable, from which the charge density can be obtained.²⁵ Formally, this means

$$\rho_{\text{sol}}(\mathbf{r}) = -\frac{1}{4\pi} \hat{\nabla}^2 \varphi_{\text{sol}}(\mathbf{r}) \quad (15)$$

The quantity $\varphi_{\text{sol}}(\mathbf{r})$ is computed on a real-space Cartesian grid and eq 2 is solved iteratively (because φ_{sol} polarizes the solute's charge density), using a multigrid algorithm.²⁵

Following work by others,^{22,23} implementation of Poisson boundary conditions for a smoothly varying (but otherwise arbitrary) permittivity function $\epsilon(\mathbf{r})$ is accomplished by reformulating eq 2 as a vacuum-like Poisson equation,²⁵

$$\hat{\nabla}^2 \varphi_{\text{tot}}(\mathbf{r}) = -4\pi\rho_{\text{tot}}(\mathbf{r}) \quad (16)$$

wherein the effects of anisotropic solvation manifest within the polarization density $\rho_{\text{pol}}(\mathbf{r})$, which is defined by a partition analogous to that in eq 14:

$$\rho_{\text{tot}}(\mathbf{r}) = \rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r}) \quad (17)$$

Spatial inhomogeneity in the permittivity $\epsilon(\mathbf{r})$ leads to inhomogeneity in $\rho_{\text{pol}}(\mathbf{r})$, represented as

$$\rho_{\text{pol}}(\mathbf{r}) = \left[\frac{1 - \epsilon(\mathbf{r})}{\epsilon(\mathbf{r})} \right] \rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{iter}}(\mathbf{r}) \quad (18)$$

The quantity $\rho_{\text{iter}}(\mathbf{r})$ is updated iteratively in solving eq 16 for a fixed solute charge density (ρ_{sol}), the latter of which is obtained via self-consistent field (SCF) iterations.²⁵ For the permittivity function in eq 1 that forms the basis of PCM theory, the factor $[1 - \epsilon(\mathbf{r})]/\epsilon(\mathbf{r})$ in eq 18 vanishes inside of the solute cavity.

Following iterative solution of eq 16, the polarization density is used to augment the one-electron contributions to the solute's Fock operator, to include the interaction between $\rho_{\text{sol}}(\mathbf{r})$ and $\rho_{\text{pol}}(\mathbf{r})$.²⁵ The total energy is

$$E = E_0 + \Delta G_{\text{elst}} \quad (19)$$

where E_0 is the ground-state SCF energy functional and

$$\Delta G_{\text{elst}} = \frac{1}{2} \int \varphi_{\text{sol}}(\mathbf{r}) \rho_{\text{pol}}(\mathbf{r}) \, d\mathbf{r} \quad (20)$$

represents the interaction of the SCF density (ρ_{sol}) with the polarized continuum. For a PCM, this polarized electrostatic interaction is the entirety of the solvation energy.³

The permittivity function in eq 1 is the one that defines the PCM but the equations outlined in this section can be solved for any smoothly varying function $\epsilon(\mathbf{r})$, subject only to numerical limitations. For example, it is not possible to use an absolutely sharp dielectric boundary (as in eq 1), and for direct comparison to PCM results this boundary must be smoothed somewhat. In practice this must be done carefully. Too much smoothing causes the model to deviate from the sharp dielectric boundary that defines the PCM and other reaction-field solvation models that are predicated on the use of a sharp spherical boundary for the continuum.^{3,103} Conversely, an absolutely sharp boundary will cause wild oscillations in $\rho_{\text{iter}}(\mathbf{r})$, leading to convergence failure, unless the discretization grid is made extremely dense (at greatly increased cost). These features make the PEqS approach cumbersome in practice, even if its generality is appealing. HetPCM, on the other hand, is just as computationally simple and robust as any other PCM, requiring relatively inexpensive two-dimensional discretization of the cavity surface rather than three-dimensional discretization of all space. This makes it an attractive alternative to PEqS, provided that it realistically models the same phenomenology.

3. METHODS

The HetPCM method was implemented in a locally modified copy of Q-Chem,¹⁰⁴ where it can be used alongside the PEqS algorithm²⁵ and the isotropic SwiG-PCM algorithms.^{79,80} A code to generate the Voronoi-PEqS grids (introduced in Section 3.1.2) is available.¹⁰⁵

3.1. Permittivity Models. PEqS calculations require specification of the permittivity function $\epsilon(\mathbf{r})$ in three-dimensional space. This procedure was originally introduced for use with a van der Waals (vdW) molecular cavity,^{24–27} i.e., a sharp dielectric interface (as in eq 1), defined by atomic

spheres, somewhat smoothed for the numerical reasons discussed in Section 2.3. We will call this the “vdW-PEqS” approach, for which the construction of $\epsilon(\mathbf{r})$ is described in Section 3.1.1. Alternatively, the “Voronoi-PEqS” method introduced in Section 3.1.2 uses atom-centered dielectric regions to define $\epsilon(\mathbf{r})$, as a model that can be compared to HetPCM.

3.1.1. van der Waals Construction of $\epsilon(\mathbf{r})$. In the original PEqS approach,²⁵ the permittivity function is constructed by smoothing the sharp interface used in PCM (eq 1), connecting the values $\epsilon = 1$ (inside the cavity) and $\epsilon = \epsilon_{\text{solv}}$ (outside) by means of a switching function. Mathematically, this can be expressed as

$$\epsilon(\mathbf{r}) = 1 + (\epsilon_{\text{solv}} - 1) \prod_A^{\text{atoms}} F_A^{\text{PEqS}}(\|\mathbf{r} - \mathbf{R}_A\|) \quad (21)$$

where

$$F_A^{\text{PEqS}}(r) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{r - R_{\text{vdW},A}}{\Delta} \right) \right] \quad (22)$$

is an atom-centered switching function. The width parameter in this function is set to $\Delta = 0.265 \text{ \AA}$,²⁵ which affords solvation energies that are converged to within about 1 kcal/mol.²³ For this and other methods, the vdW radii are taken to be

$$R_{\text{vdW},A} = 1.2R_{\text{Bondi},A} \quad (23)$$

where the values $R_{\text{Bondi},A}$ are taken from ref 106 and differ from Bondi’s original vdW radii¹⁰⁷ in that $R_{\text{Bondi},H}$ is reduced from 1.2 to 1.1 \AA .¹⁰⁶ The use of eq 23 is a standard cavity construction for PCM calculations.³

For a cavity consisting of a single sphere, however, we find that better agreement with the Born ion model is obtained using a switching function based on $\tanh(x)$ rather than $\operatorname{erf}(x)$. For a single sphere of radius R_{vdW} , centered at \mathbf{R}_0 , we take the permittivity function to be

$$\epsilon(\mathbf{r}) = \frac{1}{2} \left\{ (\epsilon_{\text{solv}} - 1) \tanh[\alpha(r_0 - r_{\text{mid}})] + \epsilon_{\text{solv}} + 1 \right\} \quad (24)$$

where

$$r_0 = \|\mathbf{r} - \mathbf{R}_0\| \quad (25)$$

and

$$r_{\text{mid}} = R_{\text{vdW}} + L/2 \quad (26)$$

is the distance (along the vector $\mathbf{r} - \mathbf{R}_0$) to the center point of the switching function. That point is situated at a distance $L/2$ outside of the sphere, as shown in Figure 2. We take $L = 0.5 \text{ \AA}$ as the length scale of the switching region and set $\alpha = 4/L$.

3.1.2. Voronoi Construction of $\epsilon(\mathbf{r})$. To test the HetPCM method against exact Poisson boundary conditions, we need a model function $\epsilon(\mathbf{r})$ for the latter. For fair comparison, this function must correspond reasonably well with a model in which a dielectric constant (either ϵ_{solv} or ϵ_{nonp}) is assigned to each atomic sphere, at the user’s discretion. To construct a corresponding function $\epsilon(\mathbf{r})$, we use Becke’s definition of “fuzzy” Voronoi cells¹⁰⁸ to define $\epsilon(\mathbf{r})$ in pointwise fashion over the real-space grid that is used in the PEqS calculations. This smoothing of the Voronoi cell boundaries is widely used as a numerical quadrature device in DFT codes, and the atom-based smoothing functions $w_A(\mathbf{r})$ are often called “Becke

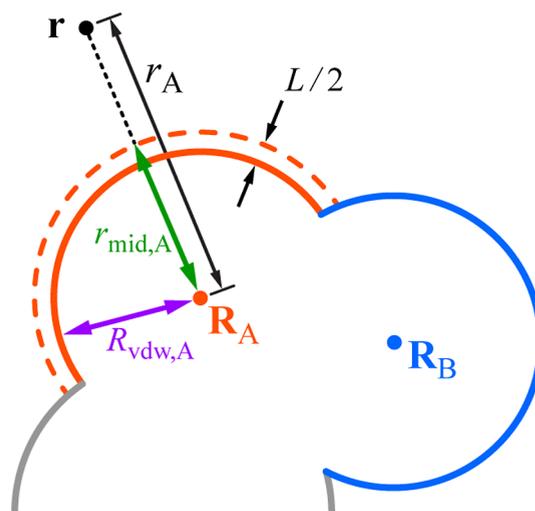


Figure 2. Illustration of the atomic spheres (A, B, ...) that comprise the vdW molecular cavity surface, along with distances used to define switching functions in eqs 24 and 29. For the indicated point \mathbf{r} , the value $\min_x \|\mathbf{r} - \mathbf{R}_x\|$ is obtained for the sphere centered at \mathbf{R}_A . As such, $\bar{A} = A$ in eq 32.

weights”. Details are provided in Section S1 of the Supporting Information and the weight function $w_A(\mathbf{r})$ is defined in eq S14.

In the present context, the function $w_A(\mathbf{r})$ is used to smooth the boundaries of the Voronoi cells defined by atom (nucleus) A , and then we impose that $\epsilon(\mathbf{r}) = 1$ for all points inside of the molecular cavity. The result is a “primitive” permittivity function,

$$\tilde{\epsilon}(\mathbf{r}) = \begin{cases} \sum_A^{\text{atoms}} \epsilon_A w_A(\mathbf{r}), & \mathbf{r} \notin \Omega \\ 1, & \mathbf{r} \in \Omega \end{cases} \quad (27)$$

where $\epsilon_A = \epsilon_{\text{solv}}$ if atom A is solvent exposed, or else $\epsilon_A = \epsilon_{\text{nonp}}$ for atoms embedded in hydrophobic regions. The choice of ϵ_A for each atom is made by the user.

In the definition of $\tilde{\epsilon}(\mathbf{r})$, the transition to $\epsilon = 1$ at the cavity surface remains abrupt. This is addressed by means of a switching function F_{cav} , and the permittivity function used in what we call the “Voronoi-PEqS” method is

$$\epsilon(\mathbf{r}) = \tilde{\epsilon}(\mathbf{r}) F_{\text{cav}}(\mathbf{r}) \quad (28)$$

The function $F_{\text{cav}}(\mathbf{r})$ interpolates between the interior value $\epsilon = 1$ and the value $\epsilon_{\bar{A}}$ that is just outside of the cavity surface, for a given point $\mathbf{r} \in \Omega$. We use a $\tanh(x)$ -based form similar to that in eq 24. Specifically

$$F_{\text{cav}}(\mathbf{r}) = \frac{1}{2} \{ (\epsilon_{\bar{A}} - 1) \tanh[\alpha(r_{\bar{A}} - r_{\text{mid},\bar{A}})] + \epsilon_{\bar{A}} + 1 \} \quad (29)$$

where

$$r_{\bar{A}} = \|\mathbf{r} - \mathbf{R}_{\bar{A}}\| \quad (30)$$

and

$$r_{\text{mid},\bar{A}} = R_{\text{vdW},\bar{A}} + L/2 \quad (31)$$

in analogy to eqs 25 and 26, respectively. The index \bar{A} in these equations denotes the sphere that is closest to the point \mathbf{r} , as shown in Figure 2. Formally,

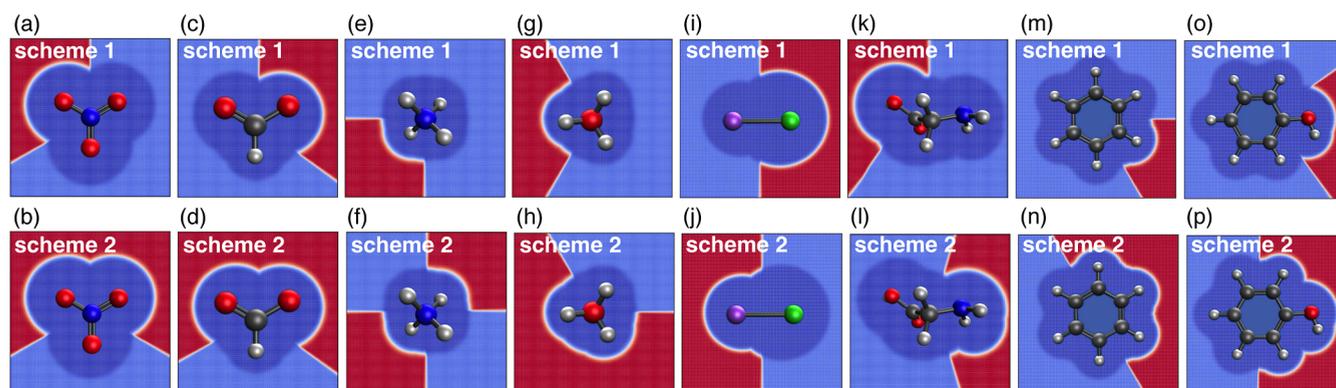


Figure 3. Permittivity functions $\epsilon(\mathbf{r})$, generated using the fuzzy Voronoi method of eq 28, for: (a,b) NO_3^- , (c,d) CHOO^- , (e,f) NH_4^+ , (g,h) H_3O^+ , (i,j) NaCl , (k,l) glycine, (m,n) benzene, and (o,p) phenol. Red regions represent $\epsilon_{\text{solv}} = 78.4$ whereas light blue represents ϵ_{nonp} (set to $\epsilon_{\text{nonp}} = 10$ in these particular images), and finally the dark blue region is the cavity interior where $\epsilon = 1$. The switching regions between these values are highlighted in white but that switching smoothly interpolates between the indicated values of ϵ . For each solute, two different constructions of $\epsilon(\mathbf{r})$ are tested: scheme 1 (along the top row) places less of the solute in contact with the polar region whereas scheme 2 (bottom row) puts more of it into that region.

$$\bar{A} = \operatorname{argmin}_{\mathbf{x}} (\min \|\mathbf{r} - \mathbf{R}_{\mathbf{x}}\|) \quad (32)$$

The effect of $F_{\text{cav}}(\mathbf{r})$ in eq 28 is to make sure that $\epsilon \approx 1$ for points near the vdW cavity surface, which is constructed using the atomic radii defined in eq 23. We take $L = 0.5 \text{ \AA}$ and $\alpha = 4/L$, as for the single-sphere cavity, so that the switching function is centered $L/2 = 0.25 \text{ \AA}$ outside of the vdW surface as shown in Figure 2.

As an example, Figure 3 presents two different versions of the function $\epsilon(\mathbf{r})$ for several different solutes. These two versions are labeled “scheme 1” and “scheme 2” and correspond to different portions of each solute cavity being exposed to solvent ($\epsilon = \epsilon_{\text{solv}}$) or else classified as hydrophobic ($\epsilon = \epsilon_{\text{nonp}}$). These two schemes are discussed in Section 4.3. For now, it suffices to note that the Voronoi-based permittivity functions that are plotted in Figure 3 can easily be mapped onto atomic choices for ϵ_A in a HetPCM calculation.

3.2. Computational Details. This section provides computational details for the electronic structure calculations and, separately, the PCM and PEQs boundary conditions that are applied to them. All calculations were performed using a locally modified copy of Q-Chem v. 5.4.¹⁰⁴

3.2.1. Electronic Structure Calculations. Calculations used to benchmark the model were performed using HF theory for the solute because solvation energies computed with continuum models are no more accurate when DFT is used instead.^{109,110} (A matching set of calculations performed with the $\omega\text{B97M-V}$ functional¹¹¹ can be found in Tables S19 and S20 and will be referenced briefly in the discussion that follows.) The B3LYP+D3(BJ) functional is used for applications to protein pK_a prediction in Section 4.5, as a representative use case. Here, D3(BJ) indicates Grimme’s D3 dispersion correction based on Becke–Johnson damping.¹¹²

For all calculations, the integral screening threshold and shell-pair drop tolerance were set to $\tau_{\text{ints}} = 10^{-12}$ a.u., as appropriate for medium-size systems with diffuse basis functions.¹¹³ The SCF convergence threshold was set to $\tau_{\text{SCF}} = 10^{-8} E_h$.

3.2.2. PCM Boundary Conditions. All PCM calculations use the C-PCM model and the HetPCM adaptation of it, meaning eqs 7–9 with $\zeta = 0$. The cavity surface is constructed using scaled vdW radii, as in eq 23, then discretized using atom-centered Lebedev grids. Except where noted, we use 110

Lebedev points per atomic sphere for hydrogen and 194 points per sphere for other atoms.

3.2.3. Poisson Boundary Conditions. Solution of Poisson’s equation requires discretization of three-dimensional space, to a distance sufficiently far from the molecule such that $\varphi_{\text{tot}}(\mathbf{r}) \approx 0$. In PEQs,^{24–27} this is accomplished using a dense Cartesian grid so that the requisite Laplacian, $\hat{\nabla}^2 \varphi_{\text{tot}}(\mathbf{r})$, can be evaluated using a finite-difference procedure and a Cartesian multigrid algorithm.²⁵ For some of the results presented in Sections 4.1 and 4.2, where the solute cavities are spherical, we use a relatively coarse grid spacing $\Delta x = \Delta y = \Delta z = 0.2877 \text{ \AA}$, which is comparable to the spacing used in previous work on excitation and ionization energies.^{24–26,86,90} These quantities are relatively insensitive to the grid spacing as compared to ΔG_{elst} which is quite sensitive to how quickly $\epsilon(\mathbf{r})$ changes at the boundaries of the molecular cavity. As such, when the cavity is constructed from atomic spheres (as in Section 4.3), we use a very fine grid spacing of $\Delta x = \Delta y = \Delta z = 0.05 \text{ \AA}$. See Table S1 for tests of how ΔG_{elst} converges as a function of the grid spacing.

The coordinate origin is placed at the centroid of the atomic coordinates and the grid extends to a total length of 8 \AA in each direction. In most cases, this means that all nuclei are at least 2 \AA away from the edges of the grid. The multigrid PEQs algorithm described in ref 25 is used to solve Poisson’s equation at each SCF iteration.

3.3. Protein Systems. Previously, Su and Li¹¹⁴ used a heterogeneous PCM⁴⁶ to investigate the active sites of five type-1 copper proteins (1EYS, 1ID2, 1KDI, 1PZA, and 2CAK), and we will explore the same systems in this work. Relaxed structures were not provided in ref 114 so we started from crystal structures. Following previous work,^{12–14} we first protonated these structures using the H++ web server (pH = 7.0, salinity of 0.15 M, $\epsilon_{\text{in}} = 10$, and $\epsilon_{\text{out}} = 80$),^{115,116} then relaxed the geometries using GFN2-xTB^{117,118} in conjunction with a generalized Born/surface area implicit solvent model for water.¹¹⁹

The relaxed structures were trimmed prior to further relaxation at the DFT level, following the protocol in ref 114. A significant number of atoms were fixed (i.e., not relaxed); see Figure 1 of ref 114. These fixed-atom models were then relaxed at the B3LYP+D3(BJ)/6-31+G* level, in

Table 1. Solvation Energies for Monatomic Ions, in kcal/mol^a

solute	Born model	PEqS (vdW cavity)				PEqS (Voronoi)				PCM (spherical)			
		ΔG_{elst}		error		ΔG_{elst}		error		ΔG_{elst}		error	
				abs.	(%)			abs.	(%)			abs.	(%)
F ⁻	-25.02	-25.21	0.19	(0.7%)	-24.76	0.26	(1.1%)	-25.02	0.00	(0.0%)			
Cl ⁻	-29.80	-30.08	0.28	(1.1%)	-29.44	0.36	(1.4%)	-29.80	0.00	(0.0%)			
Li ⁺	-24.18	-24.35	0.17	(0.7%)	-23.93	0.25	(1.0%)	-24.18	0.00	(0.0%)			
Na ⁺	-19.28	-19.39	0.11	(0.4%)	-19.12	0.16	(0.7%)	-19.28	0.00	(0.0%)			

^aHF/6-31+G* for anions and HF/6-31G* for cations, with $\epsilon_{\text{sol}} = 78.4$, using radii $R = 5.55 \text{ \AA}$ (F⁻), 6.55 \AA (Cl⁻), 6.78 \AA (Li⁺), and 8.50 \AA (Na⁺).

Table 2. Solvation Energies for Small Molecules and Ions, in kcal/mol^a

solute	Kirkwood model ^b	PEqS (vdW cavity)				PEqS (Voronoi)				PCM (spherical)			
		ΔG_{elst}		error		ΔG_{elst}		error		ΔG_{elst}		error	
				abs.	(%)			abs.	(%)			abs.	(%)
ClH	-0.08	-0.07	0.01	(8.5%)	-0.12	0.04	(56.8%)	-0.07	0.01	(8.9%)			
FH	-0.22	-0.22	0.00	(0.8%)	-0.12	0.10	(45.9%)	-0.21	0.01	(3.5%)			
NO ₃ ⁻	-29.89	-30.18	0.29	(1.0%)	-29.52	0.37	(1.2%)	-29.89	0.00	(0.0%)			

^aHF/6-31+G* level with $\epsilon_{\text{sol}} = 78.4$, using spherical cavities with $R = 5.50 \text{ \AA}$ (FH), 6.55 \AA (ClH), and 5.50 \AA (NO₃⁻). ^bMultipole expansion up to $l = 20$ to represent $\rho_{\text{sol}}(\mathbf{r})$.

isotropic PCM with $\epsilon_s = 78.4$. (Note that we use the VWN5 version of B3LYP,¹²⁰ which is slightly different from the version used in ref 114.) Relaxed structures are provided in the Supporting Information and are similar to those reported in ref 114.

In 1E5Y, 1KDI, 1PZA, and 2CAK the copper ion is in its reduced form (Cu⁺), whereas in 1ID2 it is in the oxidized form (Cu²⁺). These charge states were specified as part of the SCF initial guess. For HetPCM calculations, a dielectric constant $\epsilon_{\text{sol}} = 78.4$ was assigned to the solvent-exposed histidine side chains, whereas the remaining atoms were assigned a value of either $\epsilon_{\text{nonp}} = 4$ or 10. The solute cavity was constructed and discretized as described in Section 3.2.

4. RESULTS AND DISCUSSION

For reproducibility purposes, PEqS and HetPCM solvation energies for various molecules and ions can be found in Section S3 of the Supporting Information, computed in a variety of basis sets and using either HF or ω B97M-V for the solute. The detailed discussion of solvation energies that follows is focused on calculations at the HF/def2-TZVPD level, as this separates the comparison of PEqS and HetPCM from any issues related to DFT, such as delocalization error for the ionic solutes.

4.1. Evaluating Discretization Errors. As a test of the PEqS setup, including the use of smooth Voronoi cells to construct $\epsilon(\mathbf{r})$, we first consider several examples using spherical cavities. These are not intended as realistic models of solvation but they do furnish model problems for which exact analytic results can be obtained, under the assumption that the cavity is large enough so that there is no charge penetration into the continuum. Then, the only source of error in an isotropic PCM calculation is the discretization procedure, while errors in numerical PEqS calculations arise both from discretization and from the fact that the sharp dielectric interface (eq 1) that is assumed in the analytic solution must be smoothed in practice for numerical implementation. Comparison against analytic results allows us to quantify these numerical errors, before moving to realistic molecular cavities.

The Born ion model affords an analytic result for solvation of a monatomic ion in a spherical cavity,³ and Table 1 compares exact values for ΔG_{elst} to PEqS and PCM results for several such ions. PCM calculations reproduce the analytic results essentially exactly (errors <0.01 kcal/mol), whereas PEqS errors average about 1% of ΔG_{elst} for either of two versions of PEqS boundaries that are examined. These include a standard vdW cavity construction, as in the original PEqS approach,^{24–26} which simply means smoothing the sharp dielectric boundary according to eq 21. (For the present calculations, that boundary surface consists of a single sphere.) In addition, we consider pointwise construction of $\epsilon(\mathbf{r})$ according to the Voronoi scheme in eq 28. These calculations use large atomic radii ($R = 3.12R_{\text{vdW}}$) to ensure that there is no escaped charge, since the analytic Born model is only exact under that assumption. A $21 \text{ \AA} \times 21 \text{ \AA} \times 21 \text{ \AA}$ grid is employed, with $\Delta x = \Delta y = \Delta z = 0.2877 \text{ \AA}$, which is typical of earlier PEqS calculations.²⁵

For the PEqS calculations using the standard vdW cavity construction, the discretized Poisson model solvates the ions slightly more strongly than the exact analytic solution to Poisson's equation, although the discretization errors are smaller than 0.3 kcal/mol or about 1% of ΔG_{elst} . For the Voronoi construction of $\epsilon(\mathbf{r})$, the errors are in the other direction but also about 1% of ΔG_{elst} .

The Born model can be generalized to multipoles of arbitrary order (Kirkwood model),³ centered in a spherical cavity surrounded by an isotropic medium with a sharp dielectric interface. To use this model, $\rho_{\text{sol}}(\mathbf{r})$ is represented by a multipole expansion (up to $l = 20$ for these examples) and the analytic result is applied for each spherical harmonic function.¹⁰³ In Table 2, we present results for two polar diatomic molecules and one polyatomic ion, using spherical cavities with $R = 5.5\text{--}6.5 \text{ \AA}$. These are large cavities ($R \sim 3R_{\text{vdW}}$), to ensure that there is no escaped charge so that the Kirkwood multipolar result is exact. Consequently, the solvation energies are correspondingly small. (For comparison, C-PCM calculations for the same species afford $\Delta G_{\text{elst}} = -5$ kcal/mol for ClH, $\Delta G_{\text{elst}} = -8$ kcal/mol for FH, and $\Delta G_{\text{elst}} = -64$ kcal/mol for NO₃⁻.) Due to the smallness of the solvation

energies for the two neutral diatomic molecules, the PEqS grid spacing is set to $\Delta x = \Delta y = \Delta z = 0.05 \text{ \AA}$ in order to provide well-converged results.

For the standard vdW cavity construction, PEqS calculations reproduce analytic results to $<0.01 \text{ kcal/mol}$ accuracy for the molecules FH and ClH, although the solvation energies themselves are $|\Delta G_{\text{elst}}| = 0.1\text{--}0.2 \text{ kcal/mol}$. For the Voronoi construction, the error for FH is large in relative terms but is only 0.1 kcal/mol in absolute terms, while the error for ClH is negligibly small. The NO_3^- ion has a much larger solvation energy and larger absolute errors for PEqS as compared to the analytic Kirkwood result, on the order of $0.3\text{--}0.4 \text{ kcal/mol}$ and thus similar to discretization errors for the monatomic ions. As in that case, this error is only about 1% of $|\Delta G_{\text{elst}}|$. This level of agreement suggests that we can use the PEqS code, in conjunction with the Voronoi construction of $\epsilon(\mathbf{r})$, to generate near-exact benchmarks that can be used as reference values for other methods. As such, only the Voronoi construction of $\epsilon(\mathbf{r})$ will be used in subsequent PEqS calculations.

4.2. Comparison to Isotropic PCM. Before testing HetPCM, we first confirm that the isotropic PCM agrees with the Voronoi-PEqS method, in the case of a homogeneous dielectric environment. Table 3 compares solvation energies for aqueous ions and small molecules using these two methods. A grid spacing of 0.05 \AA is used for the Voronoi-PEqS calculations.

Table 3. Solvation Energies for Small Molecules and Ions^a

solute	ΔG_{elst} (kcal/mol)		difference	
	PCM	PEqS ^b	Abs.	(%)
NO_3^-	-63.7	-62.5	1.2	(2.0%)
NH_4^+	-81.3	-79.1	2.2	(2.8%)
NaCl	-23.3	-22.1	1.3	(5.7%)
CHOO^-	-75.2	-73.1	2.1	(2.9%)
H_3O^+	-89.9	-86.9	2.9	(3.4%)
glycine	-12.2	-11.5	0.7	(6.0%)
benzene	-3.7	-3.5	0.2	(5.5%)
phenol	-8.0	-6.6	1.4	(20.9%)

^aHF/def2-TZVPD with $\epsilon_{\text{solv}} = 78.4$ and atomic radii from eq 23.

^bVoronoi construction of $\epsilon(\mathbf{r})$, eq 28.

The Voronoi-PEqS solvation energies are systematically smaller than PCM values, with absolute differences ranging up to 2.9 kcal/mol . However, these differences amount to no more than 6% of ΔG_{elst} , except in the case of phenol where the error is 21% although the solvation energy is also smaller than that of most other solutes in Table 3. Errors for ions are $\lesssim 3\%$.

4.3. Quantitative Evaluation of HetPCM. In order to justify the use of HetPCM, we next compute solvation energies using different dielectric setups and compare HetPCM results to Voronoi-PEqS values, as the Voronoi construction of $\epsilon(\mathbf{r})$ best maps onto the concept of “different dielectric constants for different atoms”. We use $\epsilon_{\text{solv}} = 78.4$ for all of these calculations but vary ϵ_{nonp} . Results are provided in Table 4 for $\epsilon_{\text{nonp}} = 1$ and 2, and in Table 5 for $\epsilon_{\text{nonp}} = 4$ and 10. These data use HF/def2-TZVPD calculations for the solute but analogous calculations using $\omega\text{B97M-V/def2-TZVPD}$ can be found in Tables S19 and S20.

For each combination of dielectric constants, we examine two different Voronoi-based constructions of $\epsilon(\mathbf{r})$ that are called “scheme 1” and “scheme 2” in Tables 4 and 5. This is a reference to the numbering scheme introduced in Figure 3,

which shows how three-dimensional space is partitioned into polar and nonpolar regions characterized by ϵ_{solv} and ϵ_{nonp} , respectively. Taking NO_3^- as an example, the difference between schemes 1 and 2 lies in whether a single N–O moiety is exposed to the high-dielectric region (scheme 1, Figure 3a) or else two N–O moieties are immersed in ϵ_{solv} (scheme 2, Figure 3b). One may object that this is not a physically realistic setup for a strongly polarizing species such as NO_3^- , although it is a well-defined model problem. Considering glycine as another example, the difference between schemes 1 and 2 lies in whether the carboxylate moiety is solvent-exposed (scheme 1, Figure 3k) or whether the amino group is instead solvent-exposed (scheme 2, Figure 3l). For other solutes, the meaning of scheme 1 versus scheme 2 can be gleaned from Figure 3, and in each case amounts to exposing different parts of the vdW cavity surface to $\epsilon_{\text{solv}} = 78.4$ versus a smaller dielectric value, ϵ_{nonp} . In the Voronoi-PEqS calculations this is done in a smooth way—and Figure 3 is actually a color map of the permittivity function $\epsilon(\mathbf{r})$ that is constructed—whereas for HetPCM, ϵ_{solv} and ϵ_{nonp} are assigned at the atomic level, in a manner that should be obvious from Figure 3.

The most extreme example of heterogeneity is to set $\epsilon_{\text{nonp}} = 1$, for which results are provided in Table 4. Here, the mean absolute difference between the HetPCM and Voronoi-PEqS values of ΔG_{elst} is 11.2 kcal/mol (or 64% of ΔG_{elst}) with a maximum deviation of 35.2 kcal/mol . These very large deviations reflect the *ad hoc* nature of the HetPCM construction. It is notable, however, that the largest errors are incurred for the least realistic solvation environments. For example, in the case of NO_3^- there is a 56% difference between HetPCM and Voronoi-PEqS values of ΔG_{elst} when one N–O moiety is in contact with the solvent (represented by $\epsilon_{\text{solv}} = 78.4$), which is reduced to 19% error when the high-dielectric environment surrounds two N–O moieties. Similar remarks can be made for the other ions, for which scheme 1 is a less realistic solvation model because it puts less of the ion’s vdW surface in contact with ϵ_{solv} . In contrast, scheme 2 surrounds more of the vdW surface with ϵ_{solv} , which is closer to homogeneous solvation.

For the neutral species (NaCl, glycine, benzene, and phenol), absolute differences between HetPCM and Voronoi-PEqS solvation energies are $\lesssim 1 \text{ kcal/mol}$ even for the case where $\epsilon_{\text{nonp}} = 1$ and $\epsilon_{\text{solv}} = 78.4$. These solvation energies are much smaller as compared to those for ions, and in several of the charge-neutral cases the percentage errors in ΔG_{elst} are rather large. For example, phenol exhibits the largest absolute differences in ΔG_{elst} among the neutral solutes (at 2 kcal/mol), which translates into relative errors as large as 118% since the Voronoi-PEqS benchmarks lie in the range $|\Delta G_{\text{elst}}| = 4\text{--}6 \text{ kcal/mol}$.

However, these deviations are substantially reduced simply by increasing ϵ_{nonp} from 1 to 2 (Table 4). For the combination of $\epsilon_{\text{nonp}} = 2$ and $\epsilon_{\text{solv}} = 78.4$, the maximum deviation between the HetPCM and Voronoi-PEqS methods is no greater than 1.7 kcal/mol (41% of ΔG_{elst}) for the neutral solutes. Including the ions, the average deviation is 4.8 kcal/mol . (The average deviation is 5.2 kcal/mol at the $\omega\text{B97M-V/def2-TZVPD}$ level; see Table S19.) Although the change from $\epsilon_{\text{nonp}} = 1$ to 2 seems quite modest, there is precedent for such a change inducing relatively large changes in reaction energies in QM/PCM calculations on cluster models of enzymes. Those changes are

Table 4. Solvation Energies (in kcal/mol) Computed Using $\epsilon_{\text{nonp}} = 1$ or 2^a

solute	scheme ^b	$\epsilon_{\text{nonp}} = 1, \epsilon_{\text{solv}} = 78.4$				$\epsilon_{\text{nonp}} = 2, \epsilon_{\text{solv}} = 78.4$			
		ΔG_{elst}		difference		ΔG_{elst}		difference	
		PEqS ^c	HetPCM	Abs.	(%)	PEqS ^c	HetPCM	Abs.	(%)
NO ₃ ⁻	1	-39.8	-25.5	14.4	(56%)	-49.1	-43.8	5.3	(12%)
NO ₃ ⁻	2	-55.5	-46.6	8.9	(19%)	-58.4	-54.3	4.1	(8%)
CHOO ⁻	1	-58.2	-45.9	12.3	(27%)	-64.3	-59.6	4.7	(8%)
CHOO ⁻	2	-73.2	-71.2	2.0	(3%)	-73.8	-73.2	0.6	(1%)
NH ₄ ⁺	1	-45.3	-15.2	30.1	(199%)	-58.8	-48.3	10.5	(22%)
NH ₄ ⁺	2	-64.3	-29.4	34.8	(118%)	-69.9	-55.2	14.6	(26%)
H ₃ O ⁺	1	-54.6	-21.4	33.2	(156%)	-67.3	-55.6	11.7	(21%)
H ₃ O ⁺	2	-76.8	-41.6	35.2	(85%)	-80.9	-65.7	15.3	(23%)
NaCl	1	-7.7	-9.4	1.6	(18%)	-13.2	-15.9	2.7	(17%)
NaCl	2	-11.5	-11.9	0.4	(3%)	-15.4	-17.6	2.2	(12%)
glycine	1	-3.1	-3.1	0.0	(1%)	-5.3	-6.8	1.5	(23%)
glycine	2	-4.6	-3.6	1.0	(29%)	-6.4	-7.2	0.8	(11%)
benzene	1	-1.0	-0.5	0.4	(85%)	-1.6	-1.8	0.3	(15%)
benzene	2	-2.2	-1.6	0.5	(33%)	-2.5	-2.4	0.0	(1%)
phenol	1	-4.0	-1.8	2.1	(118%)	-4.7	-3.4	1.3	(37%)
phenol	2	-5.7	-3.4	2.3	(68%)	-6.0	-4.3	1.7	(41%)

^aCalculations at the HF/def2-TZVPD level. ^bPermittivity construction; see Figure 3. ^cVoronoi construction of $\epsilon(r)$ using $L = 0.5 \text{ \AA}$, $k = 10$, and a 0.05 \AA grid spacing.

Table 5. Solvation Energies (in kcal/mol) Computed Using $\epsilon_{\text{nonp}} = 4$ or 10^a

solute	scheme ^b	$\epsilon_{\text{nonp}} = 4, \epsilon_{\text{solv}} = 78.4$				$\epsilon_{\text{nonp}} = 10, \epsilon_{\text{solv}} = 78.4$			
		ΔG_{elst}		difference		ΔG_{elst}		difference	
		PEqS ^c	HetPCM	Abs.	(%)	PEqS ^c	HetPCM	Abs.	(%)
NO ₃ ⁻	1	-54.8	-53.7	1.1	(2%)	-59.2	-60.0	0.8	(1%)
NO ₃ ⁻	2	-60.1	-58.8	1.3	(2%)	-61.6	-61.8	0.3	(0%)
CHOO ⁻	1	-67.9	-67.3	0.7	(1%)	-71.0	-72.2	1.2	(2%)
CHOO ⁻	2	-73.9	-74.2	0.4	(1%)	-74.0	-74.8	0.9	(1%)
NO ₄ ⁺	1	-67.2	-65.1	2.1	(3%)	-73.8	-75.3	1.5	(2%)
NH ₄ ⁺	2	-73.4	-68.4	5.0	(7%)	-76.6	-76.6	0.0	(0%)
H ₃ O ⁺	1	-75.3	-73.1	2.2	(3%)	-81.8	-83.6	1.9	(2%)
H ₃ O ⁺	2	-83.4	-78.0	5.4	(7%)	-85.6	-85.5	0.2	(0%)
NaCl	1	-16.8	-19.6	2.7	(14%)	-19.6	-21.9	2.3	(10%)
NaCl	2	-17.9	-20.5	2.6	(13%)	-20.0	-22.3	2.3	(10%)
glycine	1	-7.1	-9.1	1.9	(21%)	-9.1	-11.1	2.1	(18%)
glycine	2	-8.0	-9.4	1.4	(15%)	-9.5	-11.3	1.8	(16%)
benzene	1	-2.1	-2.6	0.5	(19%)	-2.7	-3.3	0.6	(18%)
benzene	2	-2.8	-2.9	0.1	(4%)	-3.1	-3.4	0.4	(10%)
phenol	1	-5.4	-4.4	1.0	(23%)	-6.1	-5.0	1.1	(21%)
phenol	2	-6.3	-4.8	1.5	(32%)	-6.6	-5.2	1.4	(27%)

^aCalculations at the HF/def2-TZVPD level. ^bPermittivity construction; see Figure 3. ^cVoronoi construction of $\epsilon(r)$ using $L = 0.5 \text{ \AA}$, $k = 10$, and a 0.05 \AA grid spacing.

often large upon passing from vacuum dielectric boundaries to $\epsilon = 2$, with the effect saturating by $\epsilon = 4$.^{12,84,85,121–124}

The agreement between HetPCM and Voronoi-PEqS improves even further for $\epsilon_{\text{nonp}} = 4$ or 10 , as demonstrated in Table 5. For $\epsilon_{\text{nonp}} = 4$, which is often used to represent the hydrophobic interiors of proteins,^{48–53,91} the absolute differences for charge-neutral solutes are all <2 kcal/mol, which is no more than a 32% error, while the errors for ions are all <10%. For $\epsilon_{\text{nonp}} = 10$, this difference drops to $\lesssim 2\%$ for both neutral and ionic solutes. Nevertheless, there remain substantial differences between HetPCM solvation, with $\epsilon_{\text{nonp}} = 10$ and $\epsilon_{\text{solv}} = 78$ (Table 5), and isotropic PCM solvation with $\epsilon_s = 78$ (Table 3). It is not the case that the former simply converges to the latter by $\epsilon_{\text{nonp}} = 10$.

For $\omega\text{B97M-V/def2-TZVPD}$ calculations with $\epsilon_{\text{nonp}} = 10$, the average deviation between HetPCM and Voronoi-PEqS is 1.1 kcal/mol for the same data set (Table S20). For context, we note that absolute errors in C-PCM solvation energies average 1.6 kcal/mol for small, charge-neutral solutes in water, if one compares calculated values of ΔG_{elst} to the experimental free energies of solvation (ΔG°), with no corrections for nonelectrostatic interactions.^{3,125} (The experimental data have estimated uncertainties of only ± 0.2 kcal/mol.^{126,127}) As such, discrepancies of 1–2 kcal/mol between HetPCM and Voronoi-PEqS are of the same order of magnitude as the intrinsic accuracy of continuum solvation models themselves. For ions, the accuracy of C-PCM drops to 7–8 kcal/mol for

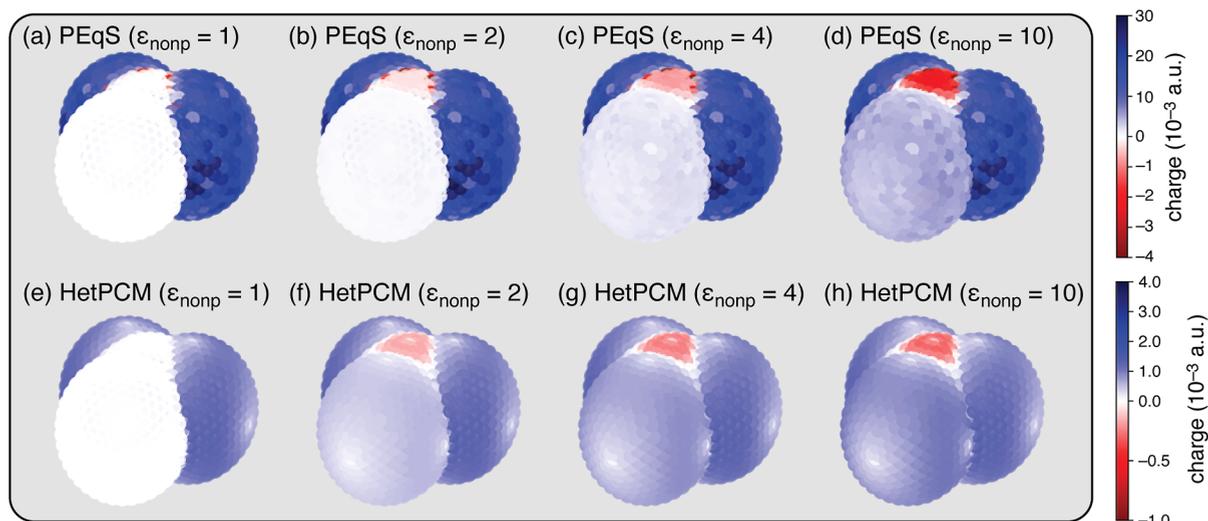


Figure 4. Comparison of (a–d) the PEqS polarization charge interpolated onto the vdW cavity surface, versus (e–h) the corresponding HetPCM surface charge distribution, for NO_3^- in dielectric scheme 2 (see Figure 3b). The two N–O moieties farthest from the viewer are immersed in $\epsilon_{\text{solv}} = 78.4$ while that nearest the observer resides in a region described by ϵ_{nonp} with the values indicated.

aqueous solvation,^{3,125} although the uncertainties associated with the experimental data are ± 3 kcal/mol for ions.¹²⁷

In our judgment, HetPCM behaves reasonably well as assessed by comparison to the Voronoi-PEqS method and in comparison to the inherent accuracy of continuum solvation models, provided that $\epsilon_{\text{nonp}} \geq 4$. The explanation for this is likely the rapidity with which the continuum solvation energy converges to the $\epsilon \rightarrow \infty$ limit,⁷¹ which is at least part of the reason why values $\epsilon > 4$ have so little effect in enzyme modeling.^{12,84,85,121–124} In contrast, the combination of $\epsilon_{\text{nonp}} = 1$ with $\epsilon_{\text{solv}} = 78.4$ affords much larger discrepancies between HetPCM and Poisson electrostatic solvation energies.

4.4. Surface Potentials. Solvation energies are used in the analysis above because they provide a single, interpretable number to compare HetPCM and PEqS results. However, the HetPCM method is not actually intended for calculation of solvation energies, as these are not meaningful for the truncated protein models that we have in mind to investigate with this approach. In this section, we consider some qualitative comparisons of surface quantities computed using the HetPCM and PEqS methods.

4.4.1. Anisotropic Solvation of NO_3^- . We begin with a detailed examination of an intentionally unrealistic model, namely, NO_3^- in permittivity scheme 2 (Figure 3b), where one oxygen atom is exposed to a low-dielectric region. This setup provides an incisive probe of how differences between HetPCM and PEqS arise. Figure 4 provides a side-by-side comparison of the PEqS polarization charge density, ρ_{pol} (eq 18), interpolated onto the vdW cavity surface for NO_3^- , along with the HetPCM surface charges $\{q_i\}$. Both quantities are evaluated at the same set of surface discretization points $\{s_i\}$ that are used in the HetPCM calculation.

As ϵ_{nonp} is increased from 1 to 10 (from left to right in Figure 4), both $\rho_{\text{pol}}(s)$ and the PCM surface charge gradually increase from a uniform value of zero, over the entire oxygen sphere that is exposed to $\epsilon_{\text{nonp}} = 1$, to a larger, positive value as ϵ_{nonp} increases. This change is more gradual for the PEqS method, however, and for HetPCM the surface charge is much closer to isotropic (across the three oxygen atoms) by $\epsilon_{\text{nonp}} = 10$ than it is in the PEqS case. We speculate that this behavior

arises due to the increased locality of the HetPCM method, where the value of the solute's electrostatic potential $\varphi_{\text{sol}}(s_i)$ induces a charge in proportion to ϵ_i at the discretization point s_i . In contrast, the Poisson model is controlled by a global electrostatic potential φ_{tot} that is obligated to be continuous and smooth across the boundaries between atomic regions with different dielectric constants. This implies that the neighboring high-dielectric regions have a somewhat greater influence on the polarization charge in the low-dielectric region, in the PEqS calculation, as compared to the HetPCM calculation.

4.4.2. Protein Surface Potentials. In contrast to the highly construed example of NO_3^- with different environments for different oxygen atoms, a more compelling rationale for development of anisotropic solvation models is to provide dielectric boundary conditions that can be used with sizable biomolecular models. In such cases, it may not make sense to use the same dielectric constant for all residues because some of them may be solvent-exposed while others are buried within the hydrophobic interior of the protein. What we desire is a qualitative means to provide compensating charge when an ionic residue is solvent-exposed, while not overpolarizing the hydrophobic regions with a large dielectric constant.

Figure 5 compares the induced surface charge for a protonated His(+)-Ile-His-Ile tetrapeptide, computed using either a conventional PCM with $\epsilon_s = 78.4$, or else HetPCM with $\epsilon_{\text{solv}} = 78.4$ and $\epsilon_{\text{nonp}} = 4$. The first histidine in the sequence is protonated and positively charged, so we select ϵ_{solv} for the atoms of this cationic histidine and ϵ_{nonp} for the other three residues.

For the protonated histidine residue that we select to be solvent-exposed, there is essentially no difference in the surface charge predicted by the isotropic PCM and HetPCM methods (Figure 5b). However, subtle differences emerge on the other three residues, which are described with nonpolar boundaries in HetPCM but are solvent-exposed in the conventional PCM calculation. In the latter case, there is buildup of charge associated with the nitrogen and oxygen atoms that is not observed with HetPCM, where the residues in question are exposed to a low-dielectric environment that apparently cannot

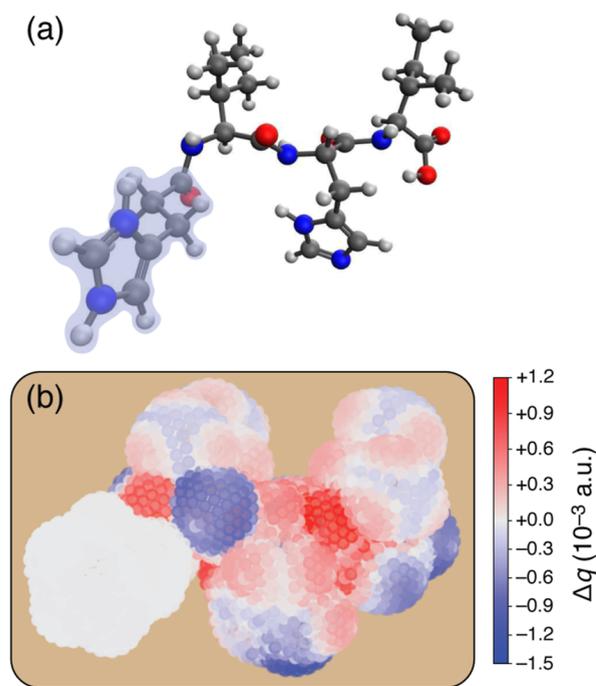


Figure 5. (a) Protonated His(+)-Ile-His-Ile tetrapeptide structure, with the solvent-exposed atoms enveloped in blue. (b) Surface map of the difference Δq_i in the induced surface charges, comparing a HetPCM calculation ($\epsilon_{\text{nonp}} = 4$, $\epsilon_{\text{solv}} = 78.4$) to a homogeneous PCM calculation ($\epsilon_s = 78.4$), at the B3LYP+D3(BJ)/6-31+G* level.

support such a charge buildup. We expect this to be important in biomolecular modeling, in order to prevent anomalous buildup of charge on what are supposed to be hydrophobic parts of the surface. In contrast, and not unexpectedly, the surface charge at the alkyl groups is more similar between these two models.

4.5. pK_a Calculations for Type-1 Cu Proteins. As an illustrative application, we consider pK_a calculations in protein models containing ca. 140 atoms, comparing homogeneous PCM and HetPCM boundary conditions. For biomolecular model systems of this size, Voronoi-PEqS calculations are intractable due to the requisite dense three-dimensional grid.

Whereas tests against exact Poisson boundaries in Section 4.3 suggest that HetPCM should probably *not* be used to predict absolute solvation energies, results below demonstrate that the use of an appropriate reference state does allow calculation of a relative solvation energy, and thus a pK_a . This methodology builds upon work by others to compute relative ΔG° and pK_a values using a heterogeneous PCM.¹¹⁴ We also provide an assessment of how the HetPCM results change with respect to the value of ϵ_{nonp} that is used for the surface boundary of the hydrophobic portions of the protein model.

The type-1 Cu centers considered here are critical to electron transfer in biological systems. The Cu center is coordinated by two histidines and one cysteine residue in a trigonal planar structure, and mutations or modifications of this site are known to cause disorders related to copper homeostasis.¹²⁸ Two protonated forms of these centers exist in solution, which are depicted for the enzyme fern plastocyanin (PDB code 1KDI) in Figure 6. In the HetPCM models, a singular histidine residue is exposed to a higher dielectric, as indicated.

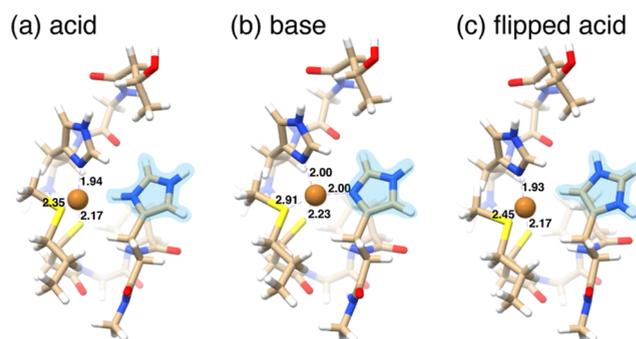


Figure 6. Relaxed structure of fern plastocyanin (1KDI) in its (a) acid, (b) base, and (c) flipped acid forms. Coordination distances to the Cu ion are labeled, in Å. Atoms enveloped in blue are described using $\epsilon_{\text{solv}} = 78.4$ in HetPCM models, whereas the remaining atoms are described using $\epsilon_{\text{nonp}} = 4$ or 10.

Following ref 114, 1KDI is used as a reference for relative pK_a calculations in other enzymes. The free energy of each protein (Pro), relative to that of plastocyanin (Plc), is estimated using the reaction



where PlcH is the protonated form of fern plastocyanin. The standard-state free energy change ΔG° will be approximated using the electronic energies for each system, neglecting entropic effects. From this, the pK_a of each protein is calculated (at $T = 298$ K) using 1KDI as a reference ($pK_a = 4.4 \pm 0.1$),¹²⁹ so the pK_a for Pro in eq 33 is

$$pK_a = 4.4 + \frac{\Delta G^\circ}{RT \ln(10)} \quad (34)$$

Relative energies of two different acid forms, analogous to the 1KDI structures in Figure 6a,c, are listed in Table 6 with a

Table 6. Energy Difference between the Acid and Imidazolium-Flipped Forms of Proteins under Different Boundary Conditions^a

protein	ΔE (kcal/mol) ^b			
	gas phase	PCM ($\epsilon_s = 78.4$)	HetPCM ($\epsilon_{\text{solv}} = 78.4$)	
			$\epsilon_{\text{nonp}} = 4$	$\epsilon_{\text{nonp}} = 10$
1ESY	-1.92	-0.32	-0.07	-0.24
1ID2	-4.89	-1.75	-2.46	-2.03
1KDI	4.01	1.35	-0.58	0.60
1PZA	0.04	0.29	-0.44	0.01
2CAK	-0.12	-0.19	-4.24	-1.71

^aB3LYP+D3(BJ)/6-31+G*. ^b $\Delta E < 0$ indicates that the imidazolium-flipped structure is more stable.

sign convention such that the imidazolium-flipped form is more stable when $\Delta E < 0$. These calculations, performed at the B3LYP+D3(BJ)/6-31+G* level, suggest that the flipped form is preferred for the enzymes 1ESY, 1ID2, and 2CAK, as it emerges lower in energy regardless of the boundary conditions. For the other two examples (1KDI and 1PZA), it is unclear which form is preferred because different boundary conditions afford different relative energies.

Experiments on 1ID2 suggest that the imidazolium-flipped acid form is more stable than the unflipped form by about 1 kcal/mol, in solution at $T = 297$ K.¹³⁰ Focusing on 1ID2, the

Table 7. Solvation Energies for Protein Models Computed with PCM Boundary Conditions^a

protein	ΔG_{elst} (kcal/mol)								
	PCM ($\epsilon_s = 78.4$)			HetPCM ($\epsilon_{\text{solv}} = 78.4$)					
	base	acid	flipped	$\epsilon_{\text{nonp}} = 4$			$\epsilon_{\text{nonp}} = 10$		
	base	acid	flipped	base	acid	flipped	base	acid	flipped
1E5Y	-68.5	-99.0	-97.4	-50.3	-77.1	-75.3	-61.6	-90.7	-89.0
1ID2	-73.3	-108.0	-104.9	-53.3	-84.6	-82.2	-65.6	-99.1	-96.3
1KDI	-66.1	-99.2	-101.9	-48.7	-77.6	-82.1	-59.5	-91.0	-94.4
1PZA	-75.8	-99.5	-99.2	-55.7	-77.3	-77.8	-68.1	-91.1	-91.1
2CAK	-60.1	-100.6	-100.7	-44.8	-76.0	-80.2	-54.3	-91.3	-92.9

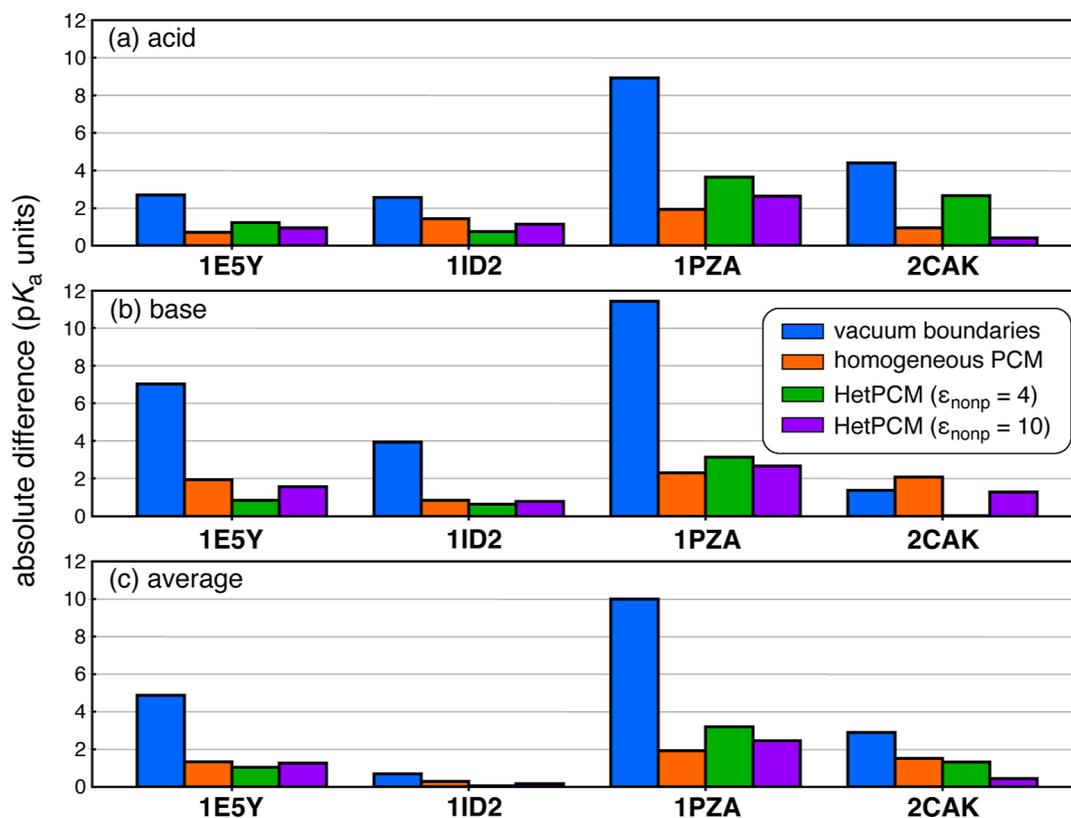
^aB3LYP+D3(BJ)/6-31+G*.

Figure 7. Absolute differences in pK_a values, as compared to experiment, for (a) acid forms of Cu proteins, (b) imidazolium-flipped forms of the same proteins, and (c) the average pK_a for these two forms.

conventional PCM with $\epsilon_s = 78.4$ is closest to that experimental value, predicting that the flipped form is 1.75 kcal/mol more stable, although the two HetPCM methods predict that the flipped form is 2.0–2.5 kcal/mol more stable, which is much closer to experiment than the gas-phase value (4.9 kcal/mol more stable).

Relative energies reported in Table 6 are not consistent in sign with the B3LYP-based heterogeneous PCM calculations reported previously by Su and Li.¹¹⁴ We have not attempted to mimic their computational paradigm exactly; for example, calculations in ref 114 omit diffuse functions and any dispersion correction. Experiments suggest that π - π interactions near the active sites of Cu-containing metalloenzymes have a significant impact on pK_a values,¹³¹ suggesting that dispersion effects may be important. Another difference, relative to the calculations reported in ref 114, is that the latter employ a united-atom cavity construction with no explicit atomic spheres for the hydrogen atoms. Although this

seems like an odd choice for pK_a calculations, we have performed calculations at the B3LYP+D3(BJ)/6-31+G* level using the same cavity construction, for comparison. Results in Table S21 demonstrate that the relative energies are rather erratic in sign and magnitude.

The polarized electrostatic interaction ΔG_{elst} is a key factor in determining how amino acid residues interact with their surroundings, which is essential for understanding protein stability in biological environments. Table 7 compares ΔG_{elst} for the three forms of each protein (acid, base, and imidazolium-flipped acid), computed using different PCM-based boundary conditions. We now understand that absolute HetPCM solvation energies do not always agree with Poisson benchmarks, and there is significant variation in ΔG_{elst} across different solvation models in Table 7. The isotropic PCM affords the largest values of $|\Delta G_{\text{elst}}|$, because it allows the most polarization charge to accumulate at the solute cavity surface.

One might anticipate that relative solvation energies (comparing different isomers of the same protein) are likely to be more accurate than absolute solvation energies, and indeed we find that differences in ΔG_{elst} are more consistent, comparing HetPCM with $\epsilon_{\text{nonp}} = 4$ versus $\epsilon_{\text{nonp}} = 10$ (Table S22). These results are consistent in trend (but not necessarily in magnitude) with those presented by Su and Li.¹¹⁴ In general, our calculations afford larger values of $|\Delta G_{\text{elst}}|$ as compared to those reported in ref 114, consistent with the use of smaller atomic radii in the present work. Results obtained with united-atom radii (Table S23) are more consistent with those in ref 114.

Finally, we compute $\text{p}K_{\text{a}}$ values for these Cu-containing proteins. Prediction of $\text{p}K_{\text{a}}$ s is a notoriously difficult computational problem,⁵³ but one where continuum solvation has historically played a role.^{132–135} Figure 7 reports $\text{p}K_{\text{a}}$ values for each system in its flipped and unflipped forms using various PCM-type boundary conditions. (Numerical data are provided in Table S24.). Experimental values, used here and in ref 114 to assess errors, are $\text{p}K_{\text{a}} < 2$ for 1ESY and 2CAK,^{136–140} $\text{p}K_{\text{a}} = 7.2$ for 1ID2,¹⁴¹ $\text{p}K_{\text{a}} = 4.4$ for 1KDI,¹²⁹ and $\text{p}K_{\text{a}} = 4.8$ for 1PZA.¹⁴²

For each system except 2CAK, we see a considerable decrease in the predicted $\text{p}K_{\text{a}}$ upon application of any solvent model. Reduction in $\text{p}K_{\text{a}}$ upon solvation is expected based on stabilization of the charged species, and the one exception is also the only system to exhibit a negative $\text{p}K_{\text{a}}$. This indicates that 2CAK is a much stronger acid than fern plastocyanin. Looking at the different solvation models, the best-performing one (as compared to experiment) is HetPCM with $\epsilon_{\text{nonp}} = 10$. However, all of the solvation models reduce the error as compared to the gas-phase result. Excluding fern plastocyanin (1KDI), which is the reference value, the average absolute difference from experiment, expressed in $\text{p}K_{\text{a}}$ units, is 4.6 (gas phase), 1.3 (isotropic PCM), 1.4 (HetPCM with $\epsilon_{\text{nonp}} = 4$) and 1.1 (HetPCM with $\epsilon_{\text{nonp}} = 10$). Expressed in energy units (ΔG°), both versions of HetPCM predict these values within 1.3 kcal/mol of experiment.

Results obtained with gas-phase (vacuum) boundary conditions are effectively useless, with average $\text{p}K_{\text{a}}$ errors that exceed 2 in most cases and exceed 6 in several cases. The largest differences with respect to experiment are for the enzyme 1PZA, which was also true in ref 114. (In the present work, the relaxed structure for 1PZA exhibits coordination distances to the Cu center that differ by only about 0.1 Å from those reported in ref 114.) Moreover, gas-phase calculations cannot rationalize differences in $\text{p}K_{\text{a}}$ upon imidazolium flipping. The difference between the two acid forms ($\Delta \text{p}K_{\text{a}}$), averaged across the data set in Figure 7, is $\Delta \text{p}K_{\text{a}} > 1$ for gas-phase calculations but $\Delta \text{p}K_{\text{a}} < 1$ for solvated calculations.

On average, HetPCM with $\epsilon_{\text{nonp}} = 4$ shows a smaller difference from the experimentally measured $\text{p}K_{\text{a}}$ value for the flipped acid form (1.16 $\text{p}K_{\text{a}}$ units) and the HetPCM model using $\epsilon_{\text{nonp}} = 10$ has a smaller difference for the acid form (1.28 $\text{p}K_{\text{a}}$ units). For the unflipped acid, the homogeneous PCM has a very slightly smaller average difference (1.26 $\text{p}K_{\text{a}}$ units) as compared to the HetPCM methods. To obtain better agreement with experiment, one might construct a larger QM model of the active site, or else include a larger number of low-energy configurations in the averaging.

Overall, HetPCM with $\epsilon_{\text{nonp}} = 10$ is in better agreement with the homogeneous PCM representation with $\epsilon_{\text{s}} = 78.4$ across all three metrics investigated, as compared to the HetPCM with

$\epsilon_{\text{nonp}} = 4$. This behavior highlights the diminishing sensitivity of the model to differences in the nonpolar dielectric constant. At higher values of ϵ_{nonp} , the effects of using dielectric boundary conditions become saturated and the differences between internal and external dielectric environments is less pronounced. This convergence suggests that the choice of dielectric constant (within a reasonable range) may not significantly impact the predictions for solvation free energies or $\text{p}K_{\text{a}}$ values. These findings emphasize the importance of using a PCM to improve the accuracy of cluster-QM enzyme models, while showcasing a range of ϵ_{nonp} values within which the results are not too sensitive to the specific choice.

5. CONCLUSION

The Voronoi-PEqS method has been shown to replicate exact (Kirkwood multipolar) solvation results for spherical cavities, and near-exact PCM results for molecular cavities. This justifies our smooth Voronoi construction of the permittivity function $\epsilon(\mathbf{r})$, as a way to define atom-specific dielectric constants that are consistent with a physical model, namely, Poisson's equation. Numerical solution of the latter can then be used as a test of the *ad hoc* HetPCM approach. As compared to this Voronoi-PEqS methods, the much simpler HetPCM procedure affords solvation energies (ΔG_{elst}) within 2 kcal/mol when the nonpolar dielectric constant is set to $\epsilon_{\text{nonp}} = 4$ or 10. Agreement deteriorates considerably for $\epsilon_{\text{nonp}} \leq 2$.

As an exemplary application, we tested HetPCM (with $\epsilon_{\text{nonp}} = 4$ and 10) for DFT prediction of $\text{p}K_{\text{a}}$ values in a set of Cu proteins, using ϵ_{nonp} to represent hydrophobic parts of a protein model and $\epsilon_{\text{sol}} = 78$ for solvent-exposed parts. Reasonable agreement with experiment is obtained, in contrast to results computed using vacuum boundary conditions. HetPCM may thus represent a simple means to implement heterogeneous dielectric boundary conditions for cluster-QM enzymatic modeling with electronic structure theory.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.4c01665>.

Additional data and technical details (PDF)

Protein structures (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

John M. Herbert – *Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, United States; Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0000-0002-1663-2278; Email: herbert@chemistry.ohio-state.edu*

Authors

Paige E. Bowling – *Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, United States; Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; Present Address: Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109 USA*
Montgomery Gray – *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Suranjan K. Paul – Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; Present Address: Department of Chemistry, Purdue University, West Lafayette, Indiana 47907 USA.

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jctc.4c01665>

Author Contributions

[§]P.E.B. and M.G. contributed equally to this work.

Notes

The authors declare the following competing financial interest(s): J.M.H. is part owner of Q-Chem Inc. and serves on its board of directors.

ACKNOWLEDGMENTS

Initial work to develop the PEQs method (by S.K.P.) was supported by the National Science Foundation under grant no. CHE-1955282. Development and testing of the HetPCM model was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award no. 1R43GM148095-01A1. M.G. further acknowledges a Presidential Fellowship from The Ohio State University. Calculations were performed at the Ohio Supercomputer Center.¹⁴³ M.G. thanks Caleb Lehman for helpful discussions regarding Voronoi polyhedra.

REFERENCES

- (1) Mennucci, B. Polarizable continuum model. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 386–404.
- (2) Lipparini, F.; Mennucci, B. Perspective: Polarizable continuum models for quantum-mechanical descriptions. *J. Chem. Phys.* **2016**, *144*, 160901.
- (3) Herbert, J. M. Dielectric continuum methods for quantum chemistry. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, No. e1519.
- (4) Cramer, C. J.; Truhlar, D. G. A universal approach to solvation modeling. *Acc. Chem. Res.* **2008**, *41*, 760–768.
- (5) Klamt, A.; Mennucci, B.; Tomasi, J.; Barone, V.; Curutchet, C.; Orozco, M.; Luque, F. J. On the performance of continuum solvation methods. A comment on “Universal approaches to solvation modeling”. *Acc. Chem. Res.* **2009**, *42*, 489–492.
- (6) Scalmani, G.; Barone, V.; Kudin, K. N.; Pomelli, C. S.; Scuseria, G. E.; Frisch, M. J. Achieving linear-scaling computational cost for the polarizable continuum model of solvation. *Theor. Chem. Acc.* **2004**, *111*, 90–100.
- (7) Lipparini, F.; Lagardère, L.; Scalmani, G.; Stamm, B.; Cancès, E.; Maday, Y.; Piquemal, J.-P.; Frisch, M. J.; Mennucci, B. Quantum calculations in solution for large to very large molecules: A new linear scaling QM/continuum approach. *J. Phys. Chem. Lett.* **2014**, *5*, 953–958.
- (8) Lipparini, F.; Scalmani, G.; Lagardère, L.; Stamm, B.; Cancès, E.; Maday, Y.; Piquemal, J.-P.; Frisch, M. J.; Mennucci, B. Quantum, classical, and hybrid QM/MM calculations in solution: General implementation of the ddCOSMO linear scaling strategy. *J. Chem. Phys.* **2014**, *141*, 184108.
- (9) Stamm, B.; Lagardère, L.; Scalmani, G.; Gatto, P.; Cancès, E.; Piquemal, J.-P.; Maday, Y.; Mennucci, B.; Lipparini, F. How to make continuum solvation incredibly fast in a few simple steps: A practical guide to the domain decomposition paradigm for the conductor-like screening model. *Int. J. Quantum Chem.* **2019**, *119*, No. e25669.
- (10) Herbert, J. M.; Lange, A. W. Polarizable continuum models for (bio)molecular electrostatics: Basic theory and recent developments for macromolecules and simulations. In *Many-Body Effects and Electrostatics in Biomolecules*; Cui, Q.; Ren, P.; Meuwly, M., Eds.; CRC Press: Boca Raton, 2016; Chapter 11, pages 363–416.
- (11) Nottoli, M.; Herbst, M. F.; Mikhalev, A.; Jha, A.; Lipparini, F.; Stamm, B. ddX: Polarizable continuum solvation from small molecules to proteins. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2024**, *14*, No. e1726.
- (12) Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Fragment-based calculations of enzymatic thermochemistry require dielectric boundary conditions. *J. Phys. Chem. Lett.* **2023**, *14*, 3826–3834.
- (13) Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Convergent protocols for protein–ligand interaction energies using fragment-based quantum chemistry. *J. Chem. Theory Comput.* **2025**, *21*, 951–966.
- (14) Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Quick-and-easy validation of protein–ligand binding models using fragment-based semiempirical quantum chemistry. *J. Chem. Inf. Model.* **2025**, *65*, 937–949.
- (15) Siegbahn, P. E. M.; Himo, F. The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 323–336.
- (16) Himo, F. Recent trends in quantum chemical modeling of enzymatic reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780–6786.
- (17) Wappett, D. A.; DeYonker, N. J. Accessible and predictable QM-cluster model building for enzymes with the Residue Interaction Network Residue Selector. *Annu. Rep. Comput. Chem.* **2024**, *20*, 131–155.
- (18) Scheraga, H. A. The multiple-minima problem in protein folding. *AIP Conf. Proc.* **1991**, *239*, 97–107.
- (19) Zhan, C.-G.; Bentley, J.; Chipman, D. M. Volume polarization in reaction field theory. *J. Chem. Phys.* **1998**, *108*, 177–192.
- (20) Chipman, D. M. New formulation and implementation for volume polarization in dielectric continuum theory. *J. Chem. Phys.* **2006**, *124*, 224111.
- (21) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. Incorporating solvation effects into density functional electronic structure calculations. *J. Phys. Chem.* **1994**, *98*, 11059–11068.
- (22) Andreussi, O.; Dabo, I.; Marzari, N. Revised self-consistent continuum solvation in electronic-structure calculations. *J. Chem. Phys.* **2012**, *136*, 064102.
- (23) Fiscaro, G.; Genovese, L.; Andreussi, O.; Marzari, N.; Goedecker, S. A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments. *J. Chem. Phys.* **2016**, *144*, 014103.
- (24) Coons, M. P.; You, Z.-Q.; Herbert, J. M. The hydrated electron at the surface of neat liquid water appears to be indistinguishable from the bulk species. *J. Am. Chem. Soc.* **2016**, *138*, 10879–10886.
- (25) Coons, M. P.; Herbert, J. M. Quantum chemistry in arbitrary dielectric environments: Theory and implementation of nonequilibrium Poisson boundary conditions and application to compute vertical ionization energies at the air/water interface. *J. Chem. Phys.* **2018**, *148*, 222834.
- (26) Paul, S. K.; Coons, M. P.; Herbert, J. M. Erratum: “Quantum chemistry in arbitrary dielectric environments: Theory and implementation of nonequilibrium Poisson boundary conditions and application to compute vertical ionization energies at the air/water interface”. *J. Chem. Phys.* **2019**, *151*, 189901.
- (27) Stein, C. J.; Herbert, J. M.; Head-Gordon, M. The Poisson–Boltzmann model for implicit solvation of electrolyte solutions: Quantum chemical implementation and assessment via Sechenov coefficients. *J. Chem. Phys.* **2019**, *151*, 224111.
- (28) Womack, J. C.; Anton, L.; Dziedzic, J.; Hasnip, P. J.; Probert, M. I. J.; Skylaris, C.-K. DL_MG: A parallel multigrid Poisson and Poisson–Boltzmann solver for electronic structure calculations in vacuum and solution. *J. Chem. Theory Comput.* **2018**, *14*, 1412–1432.
- (29) Chipman, D. M. Charge penetration in dielectric models of solvation. *J. Chem. Phys.* **1997**, *106*, 10194–10206.
- (30) Chipman, D. M. Simulation of volume polarization in reaction field theory. *J. Chem. Phys.* **1999**, *110*, 8012–8018.
- (31) Chipman, D. M. Reaction field treatment of charge penetration. *J. Chem. Phys.* **2000**, *112*, 5558–5565.

- (32) Andreussi, O.; Fiscaro, G. Continuum embeddings in condensed-matter simulations. *Int. J. Quantum Chem.* **2019**, *119*, No. e25725.
- (33) Mennucci, B.; Cossi, M.; Tomasi, J. A theoretical model of solvation in continuum anisotropic dielectrics. *J. Chem. Phys.* **1995**, *102*, 6837–6845.
- (34) Cancès, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
- (35) Cancès, E.; Mennucci, B. New applications of integral equations methods for solvation continuum models: Ionic solutions and liquid crystals. *J. Math. Chem.* **1998**, *23*, 309–326.
- (36) Mennucci, B.; Cammi, R. Ab initio model to predict NMR shielding tensors for solutes in liquid crystals. *Int. J. Quantum Chem.* **2003**, *93*, 121–130.
- (37) Frediani, L.; Cammi, R.; Corni, S.; Tomasi, J. A polarizable continuum model for molecules at diffuse interfaces. *J. Chem. Phys.* **2004**, *120*, 3893–3907.
- (38) Frediani, L.; Mennucci, B.; Cammi, R. Quantum-mechanical continuum solvation study of the polarizability of halides at the water/air interface. *J. Phys. Chem. B* **2004**, *108*, 13796–13806.
- (39) Bondesson, L.; Frediani, L.; Ågren, H.; Mennucci, B. Solvation of N_3 at the water surface: The polarizable continuum model approach. *J. Phys. Chem. B* **2006**, *110*, 11361–11368.
- (40) Mozgawa, K.; Mennucci, B.; Frediani, L. Solvation at surfaces and interfaces: A quantum-mechanical/continuum approach including nonelectrostatic contributions. *J. Phys. Chem. C* **2014**, *118*, 4715–4725.
- (41) Mozgawa, K.; Frediani, L. Electronic structure of small surfactants: A continuum solvation study. *J. Phys. Chem. C* **2016**, *120*, 17501–17513.
- (42) Ghosh, S.; Horvath, S.; Soudackov, A. V.; Hammes-Schiffer, S. Electrochemical solvent reorganization energies in the framework of the polarizable continuum model. *J. Chem. Theory Comput.* **2014**, *10*, 2091–2102.
- (43) Ghosh, S.; Hammes-Schiffer, S. Calculation of electrochemical reorganization energies for redox molecules at self-assembled monolayer modified electrodes. *J. Phys. Chem. Lett.* **2015**, *6*, 1–5.
- (44) Ghosh, S.; Soudackov, A. V.; Hammes-Schiffer, S. Electrochemical electron transfer and proton-coupled electron transfer: Effects of double layer and ionic environment on solvent reorganization energies. *J. Chem. Theory Comput.* **2016**, *12*, 2917–2925.
- (45) Iozzi, M. F.; Cossi, M.; Improta, R.; Rega, N.; Barone, V. A polarizable continuum approach for the study of heterogeneous dielectric environments. *J. Chem. Phys.* **2006**, *124*, 184103.
- (46) Si, D.; Li, H. Heterogeneous conductor like solvation model. *J. Chem. Phys.* **2009**, *131*, 044123.
- (47) Wang, J.-B.; Ma, J.-Y.; Li, X.-Y. Polarizable continuum model associated with the self-consistent-reaction field for molecular adsorbates at the interface. *Phys. Chem. Chem. Phys.* **2010**, *12*, 207–214.
- (48) Gilson, M. K.; Honig, B. H. The dielectric constant of a folded protein. *Biopolymers* **1986**, *25*, 2097–2119.
- (49) Gilson, M. K.; Honig, B. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins* **1988**, *4*, 7–18.
- (50) Rodgers, K. K.; Silgar, S. G. Surface electrostatics, reduction potentials, and internal dielectric constant of proteins. *J. Am. Chem. Soc.* **1991**, *113*, 9419–9421.
- (51) Nakamura, H. Roles of electrostatic interaction in proteins. *Q. Rev. Biophys.* **1996**, *29*, 1–90.
- (52) Kuhn, B.; Kollman, P. A.; Stahl, M. Prediction of pK_a shifts in proteins using a combination of molecular mechanical and continuum solvent calculations. *J. Comput. Chem.* **2004**, *25*, 1865–1872.
- (53) Alexov, E.; Mehler, E. L.; Baker, N.; M Baptista, A.; Huang, Y.; Milletti, F.; Erik Nielsen, J.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; Shen, J. K.; Warwicker, J.; Williams, S.; Word, J. M. Progress in the prediction of pK_a values in proteins. *Proteins* **2011**, *79*, 3260–3275.
- (54) Demchuk, E.; Wade, R. C. Improving the continuum dielectric approach to calculating pK_a s of ionizable groups in proteins. *J. Phys. Chem.* **1996**, *100*, 17373–17387.
- (55) Fitch, C. A.; Karp, D. A.; Lee, K. K.; Stites, W. E.; Lattman, E. E.; Garcia-Moreno, E. B. Experimental pK_a values of buried residues: Analysis with continuum methods and role of water penetration. *Biophys. J.* **2002**, *82*, 3289–3304.
- (56) Yang, A. S.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. On the calculation of pK_a s in proteins. *Proteins* **1993**, *15*, 252–265.
- (57) Lim, C.; Bashford, D.; Karplus, M. Absolute pK_a calculations with continuum dielectric methods. *J. Phys. Chem.* **1991**, *95*, 5610–5620.
- (58) Barone, V.; Improta, R.; Rega, N. Computation of protein pK_a s values by an integrated density functional theory/polarizable continuum model approach. *Theor. Chem. Acc.* **2004**, *111*, 237–245.
- (59) Burger, S. K.; Ayers, P. W. A parameterized, continuum electrostatic model for predicting protein pK_a values. *Proteins* **2011**, *79*, 2044–2052.
- (60) Casanovas, R.; Ortega-Castro, J.; Frau, J.; Donoso, J.; Muñoz, F. Theoretical pK_a calculations with continuum model solvents, alternative protocols to thermodynamic cycles. *Int. J. Quantum Chem.* **2014**, *114*, 1350–1363.
- (61) Gunner, M. R.; Baker, N. A. Continuum electrostatics approaches to calculating pK_a s and E_m s in proteins. *Methods Enzymol.* **2016**, *578*, 1–20.
- (62) Vorobjev, Y. N.; Scheraga, H. A.; Vila, J. A. Coupled molecular dynamics and continuum electrostatic method to compute the ionization pK_a s of proteins as a function of pH. Test on a large set of proteins. *J. Biomol. Struct. Dyn.* **2018**, *36*, 561–574.
- (63) King, G.; Lee, F. S.; Warshel, A. Microscopic simulations of macroscopic dielectric constants of solvated proteins. *J. Chem. Phys.* **1991**, *95*, 4366–4377.
- (64) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **1994**, *238*, 415–436.
- (65) Grycuk, T. Revision of the model system concept for the prediction of pK_a s in proteins. *J. Phys. Chem. B* **2002**, *106*, 1434–1445.
- (66) Truchon, J.-F.; Nicholls, A.; Roux, B.; Iftimie, R. I.; Bayly, C. I. Integrated continuum dielectric approaches to treat molecular polarizability and the condensed phase: Refractive index and implicit solvation. *J. Chem. Theory Comput.* **2009**, *5*, 1785–1802.
- (67) Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the dielectric “constant” of proteins: Smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136.
- (68) Herbert, J. M. Fantasy versus reality in fragment-based quantum chemistry. *J. Chem. Phys.* **2019**, *151*, 170901.
- (69) Cancès, E. Integral equation approaches for continuum models. In *Continuum Solvation Models in Chemical Physics*; Mennucci, B., Cammi, R., Eds.; Wiley: Chichester, 2007; pp 29–48.
- (70) Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct.* **1999**, *464*, 211–226.
- (71) Lange, A. W.; Herbert, J. M. Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models. *Chem. Phys. Lett.* **2011**, *509*, 77–87.
- (72) Chipman, D. M. Energy correction to simulation of volume polarization in reaction field theory. *J. Chem. Phys.* **2002**, *116*, 10129–10138.
- (73) Lange, A. W.; Herbert, J. M. Improving generalized Born models by exploiting connections to polarizable continuum models. I. An improved effective Coulomb operator. *J. Chem. Theory Comput.* **2012**, *8*, 1999–2011.
- (74) Lange, A. W.; Herbert, J. M. Improving generalized Born models by exploiting connections to polarizable continuum models.

- II. Corrections for salt effects. *J. Chem. Theory Comput.* **2012**, *8*, 4381–4392.
- (75) Lange, A. W.; Herbert, J. M. A simple polarizable continuum solvation model for electrolyte solutions. *J. Chem. Phys.* **2011**, *134*, 204110.
- (76) Chipman, D. M. Comparison of solvent reaction field representations. *Theor. Chem. Acc.* **2002**, *107*, 80–89.
- (77) Cancès, E.; Mennucci, B. Comment on “Reaction field treatment of charge penetration”. *J. Chem. Phys.* **2001**, *114*, 4744–4745.
- (78) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681.
- (79) Lange, A. W.; Herbert, J. M. Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces. *J. Phys. Chem. Lett.* **2010**, *1*, 556–561.
- (80) Lange, A. W.; Herbert, J. M. A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.* **2010**, *133*, 244111.
- (81) Lange, A. W.; Herbert, J. M.; Albrecht, B. J.; You, Z.-Q. Intrinsically smooth discretisation of Connolly’s solvent-excluded molecular surface. *Mol. Phys.* **2020**, *118*, No. e1644384.
- (82) York, D. M.; Karplus, M. Smooth solvation potential based on the conductor-like screening model. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.
- (83) Gregersen, B. A.; York, D. M. High-order discretization schemes for biochemical applications of boundary element solvation and variational electrostatic projection methods. *J. Chem. Phys.* **2005**, *122*, 194110.
- (84) Dasgupta, S.; Herbert, J. M. Using atomic confining potentials for geometry optimization and vibrational frequency calculations in quantum-chemical models of enzyme active sites. *J. Phys. Chem. B* **2020**, *124*, 1137–1147.
- (85) Bowling, P. E.; Dasgupta, S.; Herbert, J. M. Eliminating imaginary frequencies in quantum-chemical cluster models of enzymatic active sites. *J. Chem. Inf. Model.* **2024**, *64*, 3912–3922.
- (86) Paul, S. K.; Herbert, J. M. Probing interfacial effects on ionization energies: The surprising banality of anion–water hydrogen bonding at the air/water interface. *J. Am. Chem. Soc.* **2021**, *143*, 10189–10202.
- (87) Mouesca, J.-M.; Chen, J. L.; Noodleman, L.; Bashford, D.; Case, D. A. Density functional/Poisson–Boltzmann calculations of redox potentials for iron–sulfur clusters. *J. Am. Chem. Soc.* **1994**, *116*, 11898–11914.
- (88) Li, J.; Nelson, M. R.; Peng, C. Y.; Bashford, D.; Noodleman, L. Incorporating protein environments in density functional theory: A self-consistent reaction field calculation of redox potentials of [2Fe2S] clusters in ferredoxin and phthalate dioxygenase reductase. *J. Phys. Chem. A* **1998**, *102*, 6311–6324.
- (89) Rocchia, W.; Alexov, E.; Honig, B. Extending the applicability of the nonlinear Poisson–Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B* **2001**, *105*, 6507–6514.
- (90) Aksu, H.; Paul, S. K.; Herbert, J. M.; Dunietz, B. D. How well does a solvated octa-acid capsule shield the embedded chromophore? A computational analysis based on an anisotropic dielectric continuum model. *J. Phys. Chem. B* **2020**, *124*, 6998–7004.
- (91) Grochowski, P.; Trylska, J. Continuum molecular electrostatics, salt effects, and counterion binding—A review of the Poisson–Boltzmann theory and its modifications. *Biopolymers* **2008**, *89*, 93–113.
- (92) Lamm, G. The Poisson–Boltzmann equation. In *Reviews in Computational Chemistry*, Vol. 19; Lipkowitz, K. B.; Larter, R.; Cundari, T. R.; Boyd, D. B., Eds.; Wiley VCH: New York, 2003; Chapter 4, pages 147–365.
- (93) Baker, N. A. Poisson–Boltzmann methods for biomolecular electrostatics. *Methods Enzymol.* **2004**, *383*, 94–118.
- (94) Baker, N. A. Biomolecular applications of Poisson–Boltzmann methods. In *Reviews in Computational Chemistry*, Vol. 21; Lipkowitz, K.; Larter, R.; Cundari, T. R., Eds.; John Wiley & Sons: Hoboken, 2005; Chapter 5, pages 349–379.
- (95) Lu, B. Z.; Zhou, Y. C.; Holst, M. J.; McCammon, J. A. Recent progress in numerical methods for the Poisson–Boltzmann equation in biophysical applications. *Commun. Comput. Phys.* **2008**, *3*, 973–1009.
- (96) Wang, J.; Luo, R. Assessment of linear finite-difference Poisson–Boltzmann solvers. *J. Comput. Chem.* **2010**, *31*, 1689–1698.
- (97) Cai, Q.; Wang, J.; Hsieh, M.-J.; Ye, X.; Luo, R. Poisson–Boltzmann implicit solvation models. *Annu. Rep. Comput. Chem.* **2012**, *8*, 149–162.
- (98) Li, C.; Li, L.; Petukh, M.; Alexov, E. Progress in developing Poisson–Boltzmann equation solvers. *Mol. Based Math. Biol.* **2013**, *1*, 42–62.
- (99) Wang, C.; Wang, J.; Cai, Q.; Li, Z.; Zhao, H.-K.; Luo, R. Exploring accurate Poisson–Boltzmann methods for biomolecular simulations. *Comput. Theor. Chem.* **2013**, *1024*, 34–44.
- (100) Grant, J. A.; Pickup, B. T.; Nicholls, A. A smooth permittivity function for Poisson–Boltzmann solvation methods. *J. Comput. Chem.* **2001**, *22*, 608–640.
- (101) Davis, M. E.; McCammon, J. A. Dielectric boundary smoothing in finite difference solutions of the Poisson equation: An approach to improve accuracy and convergence. *J. Comput. Chem.* **1991**, *12*, 909–912.
- (102) Wang, J.; Cai, Q.; Xiang, Y.; Luo, R. Reducing grid dependence in finite-difference Poisson–Boltzmann calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2741–2751.
- (103) Mikkelsen, K. V.; Ågren, H.; Jensen, H. J. A.; Helgaker, T. A multiconfigurational self-consistent reaction-field method. *J. Chem. Phys.* **1988**, *89*, 3086–3095.
- (104) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.
- (105) Voronoi dielectric grid generator, <https://gitlab.com/john-herbert-group/voronoi> (accessed Dec 6, 2024).
- (106) Rowland, R. S.; Taylor, R. Intermolecular nonbonded contact distances in organic crystal structures: Comparison with distances expected from van der Waals radii. *J. Phys. Chem.* **1996**, *100*, 7384–7391.
- (107) Bondi, A. Van der Waals volumes and radii. *J. Phys. Chem.* **1964**, *68*, 441–451.
- (108) Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547–2553.
- (109) Soteras, I.; Curutchet, C.; Bidon-Chanal, A.; Orozco, M.; Luque, F. J. Extension of the MST model to the IEF formalism: HF and B3LYP parameterizations. *J. Mol. Struct.* **2005**, *727*, 29–40.
- (110) Liu, J.; Kelly, C. P.; Goren, A. C.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Zhan, C.-G. Free energies of solvation with surface, volume, and local electrostatic effects and atomic surface tensions to represent the first solvation shell. *J. Chem. Theory Comput.* **2010**, *6*, 1109–1117.
- (111) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (112) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (113) Gray, M.; Bowling, P. E.; Herbert, J. M. Comment on “Benchmarking basis sets for density functional theory thermochemistry calculations: Why unpolarized basis sets and the polarized 6-311G family should be avoided”. *J. Phys. Chem. A* **2024**, *128*, 7739–7745.
- (114) Su, P.; Li, H. Protonation of type-I Cu bound histidines: A quantum chemical study. *Inorg. Chem.* **2010**, *49*, 435–444.
- (115) H++ web server, <http://newbiophysics.cs.vt.edu/H++> (accessed Dec 3, 2024).

- (116) Anandkrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation. *Nucleic Acids Res.* **2012**, *40*, 537–541.
- (117) Semiempirical Extended Tight-Binding Program Package, <https://github.com/grimme-lab/xtb> (accessed Dec 3, 2024).
- (118) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—an accurate and broadly parameterized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- (119) Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and efficient implicit solvation model for fast semiempirical methods. *J. Chem. Theory Comput.* **2021**, *17*, 4250–4261.
- (120) Hertwig, R. H.; Koch, W. On the parameterization of the local correlation functional. What is Becke-3-LYP? *Chem. Phys. Lett.* **1997**, *268*, 345–351.
- (121) Himo, F. Quantum chemical modeling of enzyme active sites and reaction mechanisms. *Theor. Chem. Acc.* **2006**, *116*, 232–240.
- (122) Sevastik, R.; Himo, F. Quantum chemical modeling on enzymatic reactions: The case of 4-oxalocrotonate tautomerase. *Bioorg. Chem.* **2007**, *35*, 444–457.
- (123) Liao, R.-Z.; Yu, J.-G.; Himo, F. Mechanism of tungsten-dependent acetylene hydratase from quantum chemical calculations. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22523–22527.
- (124) Georgieva, P.; Himo, F. Quantum chemical modeling of enzymatic reactions: The case of histone lysine methyltransferase. *J. Comput. Chem.* **2010**, *31*, 1707–1714.
- (125) Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011–2033.
- (126) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. New universal solvation model and comparison of the accuracy of the SM5.42R, SM5.43R, C-PCM, D-PCM, and IEF-PCM continuum solvation models for aqueous and organic solvation free energies and for vapor pressures. *J. Phys. Chem. A* **2004**, *108*, 6532–6542.
- (127) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute–water clusters. *J. Chem. Theory Comput.* **2005**, *1*, 1133–1152.
- (128) National Research Council *Copper in Drinking Water*; The National Academies Press: Washington DC, 2000.
- (129) Hulsker, R.; Mery, A.; Thomassen, E. A.; Ranieri, A.; Sola, M.; Verbeet, M. P.; Kohzuma, T.; Ubbink, M. Protonation of a histidine copper ligand in fern plastocyanin. *J. Am. Chem. Soc.* **2007**, *129*, 4423–4429.
- (130) Canters, G. W.; Lommen, A.; Kamp, M.; Hoitink, C. W. G. Structure and reactivity of type-I copper sites. *BioMetals* **1990**, *3*, 67–72.
- (131) Yanagisawa, S.; Crowley, P. B.; Firbank, S. J.; Lawler, A. T.; Hunter, D. M.; McFarlane, W.; Li, C.; Kohzuma, T.; Banfield, M. J.; Dennison, C. π -interaction tuning of the active site properties of metalloproteins. *J. Am. Chem. Soc.* **2008**, *130*, 15420–15428.
- (132) Ho, J. Predicting pK_a in implicit solvents: Current status and future directions. *Aust. J. Chem.* **2014**, *67*, 1441–1460.
- (133) Thapa, B.; Schlegel, H. B. Calculations of pK_a's and redox potentials of nucleobases with explicit waters and polarizable continuum solvation. *J. Phys. Chem. A* **2015**, *119*, 5134–5144.
- (134) Ho, J.; Ertem, M. Z. Calculating free energy changes in continuum solvation models. *J. Phys. Chem. B* **2016**, *120*, 1319–1329.
- (135) Thapa, B.; Schlegel, H. B. Density functional theory calculation of pK_a's of thiols in aqueous solution using explicit water molecules and the polarizable continuum model. *J. Phys. Chem. A* **2016**, *120*, 5726–5735.
- (136) Groeneveld, C. M.; Feiter, M.; Hasnain, S. S.; van Rijn, J.; Reedijk, J.; Canters, G. W. The pH and redox-state dependence of the copper site in azurin from *Pseudomonas aeruginosa* as studied by EXAFS. *Biochim. Biophys. Acta* **1986**, *873*, 214–227.
- (137) McGinnis, J.; Ingledew, W. J.; Sykes, A. G. Kinetic studies of 1:1 electron-transfer reactions of blue copper proteins. 13. reactions of rusticyanin from *Thiobacillus ferrooxidans* with inorganic redox partners. *Inorg. Chem.* **1986**, *25*, 3730–3733.
- (138) Hunt, A. H.; Toypalmer, A.; Assamunt, N.; Cavanagh, J.; Blake, R. C.; Dyson, H. J. Nuclear magnetic resonance ¹⁵N and ¹H resonance assignments and global fold of rusticyanin: Insights into the ligation and acid stability of the blue copper site. *J. Mol. Biol.* **1994**, *244*, 370–384.
- (139) Jeuken, L. J. C.; van Vliet, P.; Verbeet, M. P.; Camba, R.; McEvoy, J. P.; Armstrong, F. K.; Canters, G. W. Role of the surface-exposed and copper-coordinating histidine in blue copper proteins: The electron-transfer and redox-coupled ligand binding properties of His117Gly azurin. *J. Am. Chem. Soc.* **2000**, *122*, 12186–12194.
- (140) Jeuken, L. J. C.; Ubbink, M.; Bitter, J. H.; van Vliet, P.; Meyer-Klaucke, W.; Canters, G. W. The structural role of the copper-coordinating and surface-exposed histidine residue in the blue copper protein azurin. *J. Mol. Biol.* **2000**, *299*, 737–755.
- (141) Lommen, A.; Canters, G. W. pH-dependent redox activity and fluxionality of the copper site in amicyanin from *Thiobacillus yersutus* as studied by 300- and 600-MHz ¹H NMR. *J. Biol. Chem.* **1990**, *265*, 2768–2774.
- (142) Impagliazzo, A. *Transient Protein Interactions; The Case of Pseudoazurin and Nitrite Reductase*; Universiteit Leiden, 2005; . Ph.D. Thesis.
- (143) Ohio Supercomputer Center, <https://osc.edu/ark:/19495/f5s1ph73> (accessed Jan 23, 2025).