



Projected Hybrid Orbitals: A General QM/MM Method

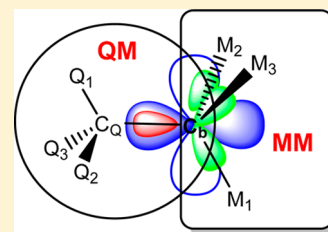
Yingjie Wang[‡] and Jiali Gao^{*,†,‡}

[†]Theoretical Chemistry Institute, State Key Laboratory of Theoretical and Computational Chemistry, Jilin University, Changchun, Jilin Province 130028, People's Republic of China

[‡]Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States

S Supporting Information

ABSTRACT: A projected hybrid orbital (PHO) method was described to model the covalent boundary in a hybrid quantum mechanical and molecular mechanical (QM/MM) system. The PHO approach can be used in *ab initio* wave function theory and in density functional theory with any basis set without introducing system-dependent parameters. In this method, a secondary basis set on the boundary atom is introduced to formulate a set of hybrid atomic orbitals. The primary basis set on the boundary atom used for the QM subsystem is projected onto the secondary basis to yield a representation that provides a good approximation to the electron-withdrawing power of the primary basis set to balance electronic interactions between QM and MM subsystems. The PHO method has been tested on a range of molecules and properties. Comparison with results obtained from QM calculations on the entire system shows that the present PHO method is a robust and balanced QM/MM scheme that preserves the structural and electronic properties of the QM region.



1. INTRODUCTION

Computer simulation of solvent effects on chemical reactions was pioneered by Jorgensen and co-workers in their seminal study of the S_N2 reaction between chloride ion and chloromethane in aqueous solution.^{1,2} In that work, quantum mechanical (QM) calculations were used to parametrize a reactive force field, or molecular mechanics (MM), for the chemical process along with its interactions with the solvent system. This potential was subsequently used in statistical mechanical Monte Carlo simulations to obtain the solvent effects, an approach that may be called QM+MM. Importantly, Jorgensen's studies stimulated the development and application of a range of computational methods to investigate the mechanism and dynamics of condensed-phase reactions, including the re-emergence of combined QM/MM methods.^{3–6}

Combined QM/MM methods provide a convenient and practical procedure to study chemical processes in condensed-phases and biological systems by modeling the chemical transformation directly with an electronic structure method.^{3–14} In application to macromolecular systems, critical to success is the treatment of the covalent boundary that separates the QM region from the remaining MM region.^{15–18} One may wonder whether this question has been well-resolved since combined QM/MM methods are widely used in a diverse range of fields for a long time,^{3,4,8,10,11,19} and indeed, satisfactory approaches are available in special cases, especially when semiempirical quantum mechanical models are used.^{20–23} However, there is still a lack of general methods for *ab initio* molecular orbital and density functional theory¹⁸ with the use of any arbitrary basis sets without introducing system-dependent parameters. Consequently, the treatment of QM and MM boundary remains an active subject of current research.^{3,4,18–51} In this article, we describe a systematic approach that employs a projected hybrid

orbital (PHO) technique for treating the covalent boundary between QM and MM molecular fragments. The PHO approach is system-independent and can be applied to any basis sets in *ab initio* wave function theory and density functional theory.

There are three main criteria to validate methods for treating QM–MM boundaries.^{15,20} First, the electron-withdrawing power between the QM region and the MM region is properly balanced such that the electronegativity of the boundary atom of the MM region closely mimics that of the full QM system treated by the same quantum mechanical model. This would allow a smooth transition from the QM region into the MM region without severely altering the reactivity of the central part of interest (i.e., the QM region). Second, a well-defined QM/MM system should preserve the integrity of the system without introducing or eliminating any degrees of freedom in the system.^{18,20} This also includes the need that electrostatic interactions from atoms in close proximity of the QM region are preserved. Finally, the boundary method is capable of yielding consistent molecular geometry and relative energy in comparison with that when the system is either fully treated by the QM method or by the corresponding MM approximation.

Numerous methods have been developed for treating the QM–MM boundary.^{3,4,18–51} Generally, these techniques may be grouped into two main categories (Figure 1). The first category includes methods that introduce additional degrees of freedom into the system or alter the local electrostatic environment of the original system.^{4,19,28,29,35–39,42–45,48} The boundary that sepa-

Special Issue: William L. Jorgensen Festschrift

Received: August 7, 2014

Revised: September 26, 2014

Published: October 15, 2014

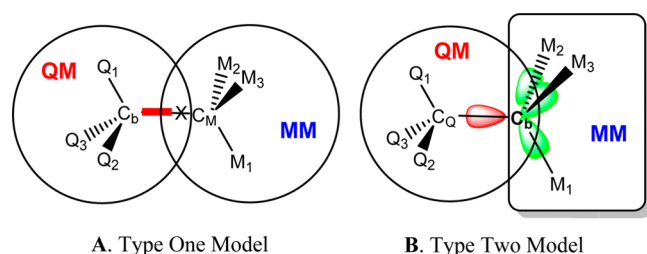


Figure 1. Classification of methods for treating the boundary between QM and MM fragments in a combined QM/MM approach. (A) Category one models are depicted, in which a QM fragment is separated from an MM subsystem across a covalent bond, typically, but not required, between two sp^3 carbons. The valency of the QM fragment is satisfied by placing a link-atom, X, which can be a hydrogen or a pseudohalogen atom to mimic the electronic properties of the C_M atom. X can be parametrized to mimic the C_b-C_M bond, namely a pseudobond model. Constraints are needed to enforce the link-atom aligned along the C_b-C_M bond; thus, it belongs both to the QM and MM region as indicated by inclusion in the two circles. (B) In type two models, the QM and MM partition is made over a boundary atom, C_b , typically an sp^3 carbon atom, which is both a “QM” atom and an “MM” atom. Typically hybrid orbitals are used. Unlike category one models, no degrees of freedom are added (Cartesian coordinates of the link-atom) or eliminated (neighboring atomic charges) here.

rates QM and MM regions is defined by a chemical bond, typically between two sp^3 carbon atoms; these two atoms are called frontier atoms. The so-called link-atom approach is a prototypical example of this category, in which the valency of the molecular fragment treated quantum mechanically is saturated by a hydrogen atom, i.e., link-atom (Figure 1a).^{4,19,38} The hydrogen link-atom is typically placed along the chemical bond between the two frontier atoms connecting the QM and MM region, and a standard bond length is adopted.^{4,19,38} Aside from hydrogen, halogen-like atoms parametrized to mimic the covalent bonds of the carbon atom in the MM region have also been used.^{35,43,44} The addition of the link-atom into the system has a number of consequences,¹⁵ including force redistribution, kinetic energy and temperature adjustment in molecular dynamics simulations (or simply ignored), and removal of electrostatic overpolarization from near contacts.^{12,38} The latter issue is a major shortcoming of the link-atom type approach since the atom that the link-atom replaces is too close to the QM fragment, and its atomic partial charges from the MM force field must be removed and redistributed.^{16,38,49} Furthermore, the atomic charges on atoms directly connected to the MM frontier atom also need to be adjusted to maintain charge neutrality of the MM fragment.^{43,44} Clearly, the local alteration of partial atomic charges would affect the local electrostatic environment and polarization of the QM region, although the significance of this effect in real systems remains to be carefully examined.¹⁵

Methods belonging to the second category do not introduce nor eliminate any degrees of freedom of the system, and they do not change the partial charges on atoms assigned to the MM region.^{3,18,20,22,30,32,46,47} Thus, these methods can best maintain the electrostatic interactions in the original system. Here, the boundary between QM and classical regions is anchored on an atomic site, often, but not restricted to, an sp^3 hybridized carbon atom (Figure 1B). Thus, literally, such a boundary atom is both a “QM” atom, being included in the QM self-consistent field (SCF) optimization, and an “MM” atom, keeping the partial charge assigned to this atom in the force field. Hybrid orbitals

were used in ref 3 as the basis set. The local self-consistent field (LSCF) method developed by Rivail and co-workers,^{18,22–26} in which three hybrid orbitals on the boundary atom are included in the SCF optimization with the remaining hybrid orbital fixed at a parametrized density, is a good example in this category. The method has been extended to a broad range of situations by Friesner and co-workers.^{32–34} An alternative is the generalized hybrid orbital (GHO) approach, in which the hybrid orbitals are obtained consistently depending on the local, instantaneous geometry during molecular dynamics simulation, and the parameters are no longer system-dependent.^{20,21,30,31,46} Ten-no and co-workers further extended the GHO method with a restricted hybridization (RH) scheme,⁴⁷ and used it to study excited states,⁵¹ NMR magnetic shielding tensors,⁵² circular dichroism spectra,⁵³ and reaction paths.⁵⁴ A main difference between the LSCF and GHO method, however, is the way the boundary atom is recognized in SCF optimizations; three bonds from the boundary atoms are optimized in the LSCF method, whereas just one covalent bond is part of the SCF procedure. Thus, the error caused by the boundary approximation on the QM region may be minimized in the GHO method.

Parameterization of the GHO method is straightforward using semiempirical QM models, requiring only two (U_{ss} and U_{pp}) of the 14 parameters to be adjusted in the standard AM1 and PM3 set for carbon.^{20,21} The extension of the GHO approach to ab initio methods becomes more cumbersome if the same basis set is used in the QM fragment and the boundary atom since it is difficult to define hybrid orbitals beyond the minimum basis. In an earlier study, the valence-only STO-3G basis was used both in ab initio Hartree–Fock theory and in density functional theory.^{30,31} Because different basis sets were employed in the QM region and on the boundary atom, electronic integrals had to be adjusted to maintain a similar electronegativity of the boundary atom as that in a full QM system.^{30,31} On the other hand, Ten-no and co-workers used the electronic integrals involving mixed basis sets directly.⁴⁶ In this Article, we present a new strategy to circumvent the need for parametrization of the electronic integrals of the boundary atom in the QM region. The main idea is to represent the core and valence electrons with a secondary, minimal basis set. Then, the original (primary) basis set used in the QM system⁵⁵ is projected onto the valence orbitals and transformed into a set of hybrid orbitals, defined in exactly the same way as that used in the GHO method.^{20,21} The hybrid orbital pointing toward the QM fragment from the boundary atom is included in the standard SCF optimization. Since the minimal basis set used in the present PHOs is a close representation of the original basis set, it retains the essential properties to balance interactions between the QM and MM subsystems. This differs from previous studies that employed the minimum basis set directly in QM/MM boundary treatment.^{30,31,46} It was found that it is possible to obtain good results without system-dependent parameters in the present approach.

In the following, we first describe the PHO method, in which we employ two optimization procedures depending on the way the total Fock matrix is partitioned. Then, we present test cases to validate the present PHO method as a general approach to model QM and MM covalent bond separations.

2. METHOD

We present an approach to treat the QM and MM covalent boundary with any basis set without introducing *system-dependent* parameters in combined QM/MM calculations using ab initio

wave function theory and density functional theory. The GHO method^{20,21} introduced previously provides a good starting point, but two major issues need to be addressed in ab initio calculations.^{18,25,30–32} First, although intuitive,^{18,20,22} the hybrid orbitals are not conveniently defined using a large basis set that includes split valence and diffuse functions.^{30,31} Previously, a valence-only minimum basis was used on the boundary atom, different from that for the QM fragment, resulting in a mixed basis description of the QM–MM bond that required the scaling of electronic integrals to balance electronegativity.^{30,31,46} Here, we employ a coarse-grain-like approach by introducing a secondary, minimal basis on the boundary atom to represent the original (primary), typically larger, basis set in the QM region. This is accomplished by projecting the large basis set of the boundary atom onto the minimal basis set, which defines a representation most closely resembling the original basis set, and thus, best preserving the electronegativity of the boundary atom in the original basis.

Second, the auxiliary orbitals need to be orthogonalized with respect to the active orbitals in the QM region. This has been carefully studied in the development of the LSCF method and the GHO model.^{18,25,30–32} Here, we employ two procedures to enforce the orthogonality constraints.

In this section, we briefly review the definition of hybrid orbitals on a QM–MM boundary atom that separates the two regions. For clarity, we use only one boundary atom in the discussion, while generalization to any number of boundary atoms between the two regions is straightforward.^{30,46} Then, we describe a strategy for a minimal basis representation in the sense of least-squares resemblance of the original, larger basis set on the boundary atom. Next, the energy formulation is presented, along with the SCF procedure and the associated density and Fock matrix formation. Finally, we present the expression for the first analytic energy derivatives of the present PHO method.

2.1. Generalized Hybrid Orbitals. We consider a system partitioned into a QM fragment and an MM subsystem across a boundary atom C_B , which is assumed to be an sp^3 carbon (Figure 1B). The C_B atom is bonded with three other MM atoms, denoted by the symbols M_1 , M_2 , and M_3 , respectively (Figure 1B). The atoms in the QM region other than C_B are defined as “full QM atoms”, as compared to the boundary atom C_B whose representation is both quantum mechanical and classical. In the original GHO method, both at semiempirical^{20,21} and at ab initio levels,^{30,31,46} there are four atomic valence orbitals on the C_B atom. Here, we also include the 1s core orbital (χ_c) (the exact nature of the five atomic orbitals on the boundary atom will be discussed in the next section).

The atomic orbitals, $\{\chi_c, 2s, p_x, p_y, p_z\}$, on the boundary atom are transformed into a set of core χ_c and valence hybrid orbitals η :

$$\begin{pmatrix} \chi_c \\ \eta_B \\ \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = T_b^\dagger \begin{pmatrix} \chi_c \\ 2s \\ p_x \\ p_y \\ p_z \end{pmatrix} \quad (1)$$

where T_b^\dagger is the basis transformation matrix, which depends on the local geometry about the C_B atom and has been explicitly defined previously,²¹ with the addition of a unity in the diagonal element corresponding to the core orbital. The hybrid orbitals

are split into two categories: two active orbitals, denoted by χ_c and η_B , with the latter pointing toward the QM frontier atom C_Q (Figure 1B), and three auxiliary orbitals that are not optimized in the SCF of the QM system, pointing roughly to the three MM neighbors (M_1, M_2, M_3).

The five core and valence hybrid orbitals plus the N atomic basis functions $\{\chi_\mu\}$ on the full QM atoms form an $(N+5)$ -dimensional hybrid-atomic-orbital (HO) space. The transformation T_H that relates the AO space with the HO space is written as

$$T_H = \begin{pmatrix} I & 0 \\ 0 & T_b \end{pmatrix} \quad (2)$$

where the transformation on the basis functions of full QM atoms is just an identity matrix of dimension N .

2.2. Projected Hybrid Orbitals. When we use an arbitrarily large atomic basis set to represent the QM region, the number of AOs on the boundary atom C_B is generally greater than that of the minimal basis functions needed to define the GHO orbitals above. When polarization and diffuse functions are included, it is even more difficult to divide these atomic orbitals directly into active and auxiliary bases. Previously, we explored an approximate approach, called GHO-AIHF,³⁰ in which the boundary atom was represented by the valence-only minimal basis, STO-3G(v), different from the basis set used for the rest of the system.^{30,31,46} The imbalance due to mixing different basis sets on the QM fragment and on the boundary atom is compensated by scaling the electronic integrals involving the STO-3G(v) basis.³⁰ To eliminate the need to scale electronic integrals in such a mixed basis approach, we present an orbital projection technique to construct hybrid orbitals from the same basis on the boundary atom as that for the full QM atoms for QM–MM boundaries. Note that in the work of Ten-no and co-workers, electronic integrals from mixed basis sets were directly used.⁴⁶

Let $\{\chi\}$ be the primary basis set used for the atoms in the QM region, including the boundary atom C_B that is treated equally as the rest of the “full QM atoms”, and $\{\zeta^b\}$ be the secondary minimal basis functions located only on the boundary atom (indicated by the superscript b). The essential step is to use basis projection to represent the primary basis set of the boundary carbon by the secondary, minimal basis. One straightforward strategy is to use Mulliken modified atomic orbital (MAO) projection scheme,^{56,57} which has been applied to population analysis on a minimal basis set.⁵⁵ In particular, the projection from the primary basis set to the minimal basis on the boundary atom is given by

$$P_b = (S^b)^{-1/2} \bar{S} (\bar{S}^b)^{-1} \bar{S})^{-1/2} \quad (3)$$

where S^b is the atomic orbital overlap of the primary basis on the boundary atom $\{\chi^b\}$, and $\bar{S}_{\mu\nu} = \langle \chi_\mu^b | \chi_\nu^b \rangle$ is the rectangular overlap matrix between the primary basis $\{\chi^b\}$ and the secondary basis $\{\zeta^b\}$. In the present study, we report results obtained with the STO-3G basis as the secondary orbitals on the boundary atom, although any STO-nG could be used. Löwdin symmetric orthogonalization $(S^b)^{-(1/2)}$ is required to preserve the maximum resemblance between the original basis and the projected basis.⁵⁸ Further, normalization is applied to retain the unitary property of the projection transformation. To this end, the N_b atomic orbitals in the primary basis is reduced to five STO-3G orbitals on the boundary atom C_B , and the $(N+N_b)$ total primary atomic orbitals of the QM fragment is reduced to $(N+5)$ mixed (primary and

secondary) atomic orbitals. The overall transformation is a $(N + N_b) \times (N + 5)$ rectangular matrix that consists of an identity transformation $(N \times N)$ on the full QM atoms and the minimal basis projection of the boundary atom B:

$$T_P = \begin{pmatrix} I & 0 \\ 0 & P_b \end{pmatrix} \quad (4)$$

The five STO-3G orbitals on the boundary carbon atom include one 1s core orbital, one 2s orbital and three 2p orbitals. Hybridization transformation defined in eq 1 is then performed on the 2s and three 2p orbitals. Thus, the overall basis set projection and hybridization transformation is given by

$$T_{PH} = T_P T_H \quad (5)$$

Furthermore, the 1s core orbital can be assigned either as an active orbital along with the active hybrid orbital to participate in SCF optimization or as a frozen orbital that contributes to the external potential from the three auxiliary hybrid orbitals. In the present study, the 1s core orbital, the one active hybrid orbital η_b , and the N basis functions on the fully QM atoms are grouped together to form an $(N+2)$ -dimensional active HO space for the SCF iteration.

2.3. Orthogonality Constraint. The overlap matrix in the $(N+5)$ -dimensional HO space, S_{N+5}^{HO} , is related to the overlap matrix of the primary atomic orbital basis, S , by the successive projection and hybridization transformations:

$$S_{N+5}^{HO} = T_H^\dagger T_P^\dagger S T_P T_H \quad (6)$$

The active orbitals are generally not orthogonal to the auxiliary orbitals that do not participate in the SCF optimization:

$$\langle \phi_a^H | \eta_b \rangle \neq 0 \quad (a = 1, \dots, N + 2; \quad b = 1, 2, 3) \quad (7)$$

where ϕ_a^H and η_b are the active and the three nonoptimized, hybrid auxiliary MOs, respectively. The superscript H is to emphasize that the molecular orbitals in the QM region are linear combinations of the primary atomic orbitals on the full QM atoms and the core and active hybrid orbitals on the boundary atom.

Orthogonality constraint between auxiliary and active MOs is a general condition for both GHO^{30,46} and LSCF-type methods,^{25,32} which must be imposed in the optimization of the active MOs. Given the orthogonality among the four hybrid orbitals on the boundary atom, the constraints of the PHO method are

$$S_{ub} = \langle \chi_u | \eta_b \rangle = 0 \quad (u = 1, 2, \dots, N + 1; \quad b = 1, 2, 3) \quad (8)$$

where χ_u includes the atomic basis on the full QM atoms and the projected core orbital (χ_c) on the boundary atom C_B , and η_b are the three auxiliary hybrid orbitals.

Several orthogonalization techniques have been described by Pu et al. in the GHO approach,³⁰ and two of them are adopted in the present PHO method: (a) the symmetric global Löwdin orthogonalization (GLO) method in which all the orbitals in HO space are diagonalized, and (b) the projected basis method in which the auxiliary basis is projected out of the active space and the Fock matrix is diagonalized in this projected active basis.

(a). *Global Löwdin Orthogonalization (GLO).* Global Löwdin transformation produces a set of orthogonal hybrid orbitals (OHO) of the entire mixed QM and boundary subsystems:

$$\phi^{OHO} = T_{LO} \phi^H \quad (9)$$

where ϕ^H specifies the column vector of the nonorthogonal active and auxiliary hybrid basis, and T_{LO} is the symmetric Löwdin orthogonalization matrix,

$$T_{LO} = (S_{N+5}^{HO})^{-1/2} \quad (10)$$

The transformation makes all the OHOs resemble the original HO in a least-squares sense. The OHOs form an orthonormal set; thereby, the orthonormality between the three auxiliary orbitals with other active basis functions is satisfied in OHOs.

The total transformation matrix T that relates the original AO basis to the OHO basis is the matrix product of consecutive transformations that include projection, hybridization, and orthogonalization:

$$T_{GLO} = T_P T_H T_{LO} \quad (11)$$

The transformation relates the Fock matrix in the OHO basis to the Fock matrix expressed in the primary AO basis by

$$F_{N+5}^{OHO} = T_{GLO}^\dagger F^{AO} T_{GLO} \quad (12)$$

Due to orthogonalization tails in the Löwdin transformation, the elements in the rows and columns corresponding to the auxiliary orbitals in F_{N+5}^{OHO} are generally not zero. In principle, a further projection step is required to uncouple the two block matrices; however, for simplicity, these elements are dropped as an approximation for simplicity to reduce the Fock matrix to $(N + 2)$ -dimension in the active OHO space for subsequent diagonalization.

(b). *Projected Basis Method.* The second approach to enforce the orthogonality constraint is the projected basis method, in which successive Schmidt orthogonalization is performed on the $(N+2)$ active basis $|\chi_u\rangle$ to remove its linear dependency with the three auxiliary hybrid orbitals. This results in a new set of projected hybrid basis, denoted by $|\tilde{\chi}_u\rangle$:³⁰

$$|\tilde{\chi}_u\rangle = [1 - \sum_{b=1}^3 (S_{ub}^{HO})^2]^{-1/2} (|\chi_u\rangle - \sum_{b=1}^3 |\eta_b\rangle \langle \eta_b | \chi_u \rangle) \quad (13)$$

$$u = 1, \dots, N + 1$$

where u is the index of active orbitals on the fully QM atoms and the 1s core orbital on the boundary atom C_B . S_{ub}^{HO} is the overlap matrix between mixed active atomic and hybrid orbitals χ_u and auxiliary orbital η_b . The transformation matrix, T_M , of projected basis was explicitly given in ref 30 and the total transformation matrix of the projected basis method that relates the original AOs to the projected active hybrid basis is

$$T_{PHO} = T_P T_H T_M \quad (14)$$

The Fock matrix defined in the projected hybrid space is obtained from the Fock matrix in primary atomic orbital basis via the transformation:

$$F_{N+5}^{PH} = T_{PHO}^\dagger F^{AO} T_{PHO} \quad (15)$$

By projecting the auxiliary orbitals out of the active space, all the elements in the rows and columns of auxiliary orbitals in the $(N+5)$ -dimensional Fock matrix F_{N+5}^{PH} are zero, resulting in the reduced Fock matrix F_{N+2}^{PH} in the projected active hybrid basis fully uncoupled to the auxiliary block. In contrast to the global Löwdin orthogonalization scheme where the $(N+2)$ -dimensional Fock matrix is constructed approximately (without the

additional projection step), the reduced Fock matrix in the projected basis method is variationally determined and analytic gradients can be computed in a more straightforward fashion.

2.4. Self-Consistent Field Procedure for the PHO Energy. To summarize the successive transformations and subsequent SCF steps, the procedure used in the PHO calculation is given below.

(a) Construct the minimal basis projection matrix T_p and the hybridization matrix T_H according to eqs 4 and 2, respectively.

(b) Construct the orthogonalization matrix, either using the global Löwdin transformation (eq 10), or the projected basis scheme. The total transformation matrix T_t is the matrix product of the consecutive operations that include projection, hybridization, and orthogonalization. Specifically, $T_t = T_{\text{GLO}}$ in the global Löwdin transformation (eq 11), and $T_t = T_{\text{PHO}}$ in the projected basis procedure (eq 14).

(c) Form the $(N+N_b)$ -dimensional Fock matrix in the primary AO basis as in standard electronic structure calculations, and transform it into the $(N+5)$ -dimensional Fock matrix in the HO basis in the PHO method using either eq 12 in the global Löwdin orthogonalization scheme or eq 15 in the projected basis method.

$$F_{N+5}^{\text{HO}} = T_t^\dagger F^{\text{AO}} T_t \quad (16)$$

(d) Diagonalize the $(N+2)$ -dimensional block Fock matrix of the active orbitals in the QM region by dropping the columns and rows that correspond to the auxiliary hybrid orbitals from F_{N+5}^{HO} , and solve the Roothaan equation in the reduced active HO space to obtain the orbital coefficients C_{N+2}^{HO} :

$$F_{N+2}^{\text{HO}} C_{N+2}^{\text{HO}} = \epsilon S_{N+2}^{\text{HO}} C_{N+2}^{\text{HO}} \quad (17)$$

Note that in the GLO procedure, S_{N+2}^{HO} is the identity matrix already.

(e) Form the new density matrix P_{N+2}^{HO} using C_{N+2}^{HO} , and append the electron densities of auxiliary orbitals to the diagonal terms to form the total density matrix in the $(N+5)$ -dimensional HO space.

$$P_{N+5}^{\text{HO}} = \begin{pmatrix} P_{N+2}^{\text{HO}} & 0 & 0 & 0 \\ 0 & 1 - q_B/3 & 0 & 0 \\ 0 & 0 & 1 - q_B/3 & 0 \\ 0 & 0 & 0 & 1 - q_B/3 \end{pmatrix} \quad (18)$$

where q_B is the partial atomic charge on the boundary atom in the MM force field.

(f) Transform the density matrix in the HO space back to the density matrix in the primary AO space:

$$P^{\text{AO}} = T_t P_{N+5}^{\text{HO}} T_t^\dagger \quad (19)$$

(g) Check the convergence in density and total electronic energy of the combined QM/MM system. If convergence is not yet achieved, the procedure returns to step (c) and the SCF iteration is repeated by another increment.

Note that the density of the partial (MM) atomic charge of the boundary atom, q_B , was equally distributed to the three auxiliary orbitals in step (e), resulting in a value of $1 - (q_B/3)$. Implicitly, this assumed that both the QM and MM fragments have no net fractional charges. This is typically satisfied using the CHARMM force field where the neutral group convention is used. In the case where there is a net nonzero fractional charge (δq_{QM}) in the QM

fragment, it is added to the boundary atom and the charge density of the auxiliary orbitals is modified as $1 - (q_B + \delta q_{\text{QM}})/3$. It is possible to further distinguish different bond polarities by using a weighted distribution ($q_{\text{Mi}}/\sum_j q_{\text{Mj}}$) based on MM charges if the three MM atoms attached to the boundary atom are different. Alternatively, Jung et al. introduced a fractional occupation scheme based on relative electronegativities.⁴⁶

The total energy of the QM/MM system using the PHO method is determined by

$$E = \sum_{\mu\nu}^{N+N_b} P_{\mu\nu}^{\text{AO}} H_{\mu\nu}^{\text{eff}} + \frac{1}{2} \sum_{\mu\nu}^{N+N_b} \sum_{\lambda\sigma}^{N+N_b} P_{\mu\nu}^{\text{AO}} P_{\lambda\sigma}^{\text{AO}} \times \left[(\mu\nu, \sigma\lambda) - \frac{1}{2} (\mu\sigma, \lambda\nu) \right] + E_{\text{nuc}}^{\text{QM}} + E_{\text{nuc}}^{\text{QM/MM}} + E^{\text{MM}} + E_{\text{B}}^{\text{tet}} \quad (20)$$

where the first two summations account for the QM electronic energy, in which all matrix elements are expressed in terms of the primary atomic orbitals, including the boundary atom, $E_{\text{nuc}}^{\text{QM}}$ and $E_{\text{nuc}}^{\text{QM/MM}}$ are the nuclear repulsion energy in the QM region and between QM and MM atoms, respectively, and E^{MM} is the energy of the MM region. In eq 20, the term $E_{\text{B}}^{\text{tet}}$ is a tetrahedral restraining potential of the boundary atom, which is introduced to correct electronic (Pauli) repulsion error among the auxiliary orbitals, which only have the electron density from the boundary atom rather than fully localized “bond” orbitals for the MM fragment. $E_{\text{B}}^{\text{tet}}$ is expressed by standard bond stretch terms:

$$E_{\text{B}}^{\text{tet}} = \sum_b^6 K_b^{\text{tet}} (\theta_b - \theta_{\text{tet}}^0)^2 \quad (21)$$

where b specifies the 6 possible bond angles about the boundary atom, $\theta_{\text{tet}}^0 = 109.47^\circ$, and $K_b^{\text{tet}} = 300 \text{ kcal mol}^{-1} \text{ rad}^{-2}$. K_b^{tet} may be considered as the only parameter of the PHO method, which is basis set- and system-independent.

Jung and Ten-no proposed an RH scheme to account the Pauli exchange repulsion errors among the auxiliary orbitals, in which the hybridization matrix for the boundary atom is kept as a perfect sp^3 center,⁴⁷ although it should be noted that the optimized geometry is not necessarily a perfect tetrahedral. The RH matrix itself serves as a restraining potential. We have also tested the RH approach.

To accelerate the SCF convergence, we used an energy-based DIIS procedure.⁵⁹ In the PHO method, the density matrices are optimized in the active HO space, whereas the Fock matrix constructed in the AO space has an implicit dependence both on the optimized density and on the frozen density of the auxiliary orbitals. Because of the contribution from frozen densities, the popular DIIS scheme based on the commutator of the density and Fock matrices is not applicable to the PHO procedure.⁶⁰

2.5. Analytic First Gradient of the PHO Energy. The reduced Fock matrix in the projected active hybrid orbital basis is variationally determined, resulting in a straightforward derivation of the analytic first gradient. The same expression may be used as an approximation to the first derivatives of the energy computed using the GLO method. The gradient of the PHO energy (eq 20) with respect to the nuclear coordinates $\{R_a\}$ is defined by

$$\frac{\partial E}{\partial R_a} = \frac{\partial E^{\text{QM}}}{\partial R_a} + \frac{\partial E_{\text{nuc}}^{\text{QM}}}{\partial R_a} + \frac{\partial E_{\text{nuc}}^{\text{QM/MM}}}{\partial R_a} + \frac{\partial E^{\text{MM}}}{\partial R_a} + \frac{\partial E_{\text{B}}^{\text{tet}}}{\partial R_a} \quad (22)$$

The gradient includes contribution from the QM electronic energy, nuclear Coulombic energy in the QM fragment and

between QM and MM atoms, and the MM energy. All terms but the first are trivial to derive.

We use Hartree–Fock theory to illustrate the gradient calculation, arising from the QM electronic energy term:

$$\begin{aligned} \frac{\partial E_{\text{QM}}}{\partial R_a} &= 2 \sum_{\mu\nu} P_{\mu\nu}^{\text{AO}} \frac{\partial H_{\mu\nu}}{\partial R_a} + \sum_{\mu\nu\rho\sigma} (2P_{\mu\nu}^{\text{AO}} P_{\rho\sigma}^{\text{AO}} - P_{\mu\rho}^{\text{AO}} P_{\nu\sigma}^{\text{AO}}) \\ &\quad \times \frac{\partial G_{\mu\nu}}{\partial R_a} + 2 \sum_{\mu\nu} \frac{\partial P_{\mu\nu}^{\text{AO}}}{\partial R_a} (H_{\mu\nu} + G_{\mu\nu}) \end{aligned} \quad (23)$$

where the first two summations are the derivatives on the one-electron and two-electron integrals, respectively, which are readily obtained from standard HF gradient calculations. The main difference between the PHO gradient and the standard HF gradient is the “density force” terms, denoted by $(\partial P_{\mu\nu}^{\text{AO}}/\partial R_a)$. Note that in standard HF, the “density force” term is transformed to energy weighted density matrix, but in PHO, $(\partial P_{\mu\nu}^{\text{AO}}/\partial R_a)$ has explicit dependency on the derivatives of the transformation matrix as follows:

$$\begin{aligned} \frac{\partial P_{\mu\nu}^{\text{AO}}}{\partial R_a} &= \frac{\partial (T_{\text{PHO}} P_{N+5}^{\text{H}} T_{\text{PHO}}^\dagger)}{\partial R_a} \\ &= \frac{\partial T_{\text{PHO}}}{\partial R_a} P_{N+5}^{\text{H}} T_{\text{PHO}}^\dagger + T_{\text{PHO}} P_{N+5}^{\text{H}} \frac{\partial T_{\text{PHO}}^\dagger}{\partial R_a} + T_{\text{PHO}} \frac{\partial P_{N+5}^{\text{H}}}{\partial R_a} T_{\text{PHO}}^\dagger \end{aligned} \quad (24)$$

where the derivatives of the transformation matrix $(\partial T_{\text{PHO}}/\partial R_a)$ is computed by

$$\frac{\partial T_{\text{PHO}}}{\partial R_a} = \frac{\partial (T_{\text{P}} T_{\text{H}} T_{\text{M}})}{\partial R_a} = \frac{\partial T_{\text{P}}}{\partial R_a} T_{\text{H}} T_{\text{M}} + T_{\text{P}} \frac{\partial T_{\text{H}}}{\partial R_a} T_{\text{M}} + T_{\text{P}} T_{\text{H}} \frac{\partial T_{\text{M}}}{\partial R_a} \quad (25)$$

Note that $(\partial T_{\text{P}}/\partial R_a) = 0$, because the projection operation is on the boundary atom only and is invariant to coordinate changes. The explicit expressions on the specific terms $(\partial T_{\text{H}}/\partial q)$ and $(\partial T_{\text{M}}/\partial q)$ have been detailed previously in the GHO method.^{21,30}

The last term in eq 24 is written as

$$\frac{\partial P_{N+5}^{\text{H}}}{\partial R_a} = \frac{\partial P_{N+2}^{\text{H}}}{\partial R_a} + \frac{\partial P_{\text{bb}}^{\text{H}}}{\partial R_a} = \frac{\partial P_{N+2}^{\text{H}}}{\partial R_a} \quad (26)$$

where the fixed density terms P_{bb}^{H} do not contribute to the gradient, and for $(\partial P_{N+2}^{\text{H}}/\partial R_a)$, energy weighted density matrix is constructed in the $(N+2)$ -dimensional projected active hybrid basis and is transformed back to AO representation.

3. COMPUTATIONAL DETAILS

We have implemented the PHO method into the Gaussian program, developmental version (gdv-H35).⁶¹ Single-point energy calculations using PHO are available both for the global Löwdin orthogonalization and the projected basis method, and the exact first analytic gradient of PHO is available for the projected basis method. In PHO calculations, the QM part can be represented either by WFT or by DFT, whereas the MM region adopts the AMBER force field⁶² except in cases that are specifically noted. We follow the same strategy in the original GHO method^{15,20} to determine the MM energy terms: the general rule is that MM energy terms containing any atom in the MM region will be retained. Therefore, for bonded interactions between QM and MM atoms, i.e., bond stretching, angle bending and torsional terms, we discard covalent terms among all QM atoms, but preserve MM energy terms that contain at least one

MM atom. Nonbonded van der Waals interactions between QM and MM atoms are fully counted for in the MM energy expression, whereas electrostatic interactions between QM and MM atoms are determined quantum mechanically by including the partial charges of all MM atoms in the QM/MM interaction Hamiltonian.

The performance of the PHO method is examined in various aspects as follows.

- Ethane.** The prototypical test case, ethane, was optimized using PHO-QM/MM with different QM models and basis sets; they include, respectively, HF, B3LYP, and M06-2X methods, and STO-3G, 6-31G(d), 6-311+G(d,p), and aug-cc-pVTZ basis sets. This simple test is, in fact, the most crucial example to validate the balance of electronegativity between QM and MM regions by placing one methyl group in the QM region and the other in the MM region.^{15,20} In addition, the force constant in the tetrahedral restraining term (eq 21) was optimized to yield the best agreement in molecular geometry in comparison with the corresponding QM results for the entire molecule.
- Equilibrium Geometry.** To validate the performance of the PHO method and the generality of the tetrahedral restraining term, geometry optimizations were carried out on a test set of molecules that contain different functional groups. The optimized geometries from PHO calculations using HF with different basis sets are compared with the corresponding standard HF results for the full system.
- Torsion Energy Profile.** The potential energy profile about rotation of the central C₂–C₃ bond of *n*-butane was determined using PHO-QM/MM with the boundary atom placed at different positions along the carbon chain both at the HF/3-21G level and at the HF/6-31G(d) level. The PHO results are compared with standard QM profile and MM results. Since the purpose here is to compare the difference between PHO and full ab initio results, the use of relatively small basis sets is reasonable.
- Energetics.** Proton affinities for a set of organic molecules were determined from QM/MM calculations employing the PHO boundary treatment using Hartree–Fock and the B3LYP and M06-2X density functionals. In each case, the organic acid and its conjugated base were optimized using M06-2X/6-311+G(d,p). Both orthogonalization schemes, the global Löwdin orthogonalization and projected basis method, were examined.

4. RESULT AND DISCUSSION

The performance of the PHO method was examined by comparing the results obtained from QM/MM calculations and from full QM study. These include geometry optimization and charge polarization between QM and MM regions in ethane. Then, the performance of the PHO boundary method was further tested on molecular geometries, deprotonation energies, and torsional potential profiles.

A. Ethane. As in the original development of the GHO method,²⁰ ethane was selected as a prototype for testing the PHO algorithm. Although ethane might be considered to be too small for this purpose, in fact, it provides the most direct and strict test on the balance of electron-withdrawing power between the QM region and the MM subsystem across their boundary. When one methyl group is partitioned into the QM region and

Table 1. MUEs in Bond Lengths (Å) and Bond Angles (deg) of Ethane between Results from the Hybrid QM/MM Optimization with (PHO) and without the Tetrahedral Restraining Correction (NC), and the Corresponding Full QM Calculations^a

	STO-3G		6-31G(d)		6-311+G(d,p)		aug-cc-pVTZ	
	NC	PHO	NC	PHO	NC	PHO	NC	PHO
C–H	0.002	0.001	0.002	0.003	0.004	0.004	0.003	0.002
C–C _B	0.054	0.072	0.140	0.155	0.101	0.118	0.057	0.070
C _B –H _B	0.012	0.013	0.020	0.021	0.025	0.025	0.027	0.027
H–C–H	1.3	1.3	1.3	1.0	1.1	0.9	1.8	1.6
H–C–C _B	1.3	1.2	1.2	1.0	1.1	0.8	1.7	1.5
C _A –C _B –H _B	12.8	1.9	8.6	0.4	12.6	1.4	11.3	1.1
H _B –C _B –H _B	15.8	2.0	10.3	0.4	15.6	1.5	13.8	1.1

^aFor each basis set listed, the MUE is averaged over HF, B3LYP and M062X calculations, which have similar deviations. The subscript B specifies the atom is the boundary atom.

Table 2. MUEs in Bond Lengths (Å) and Bond Angles (deg) of Ethane Using PHO with the Tetrahedral Restraining Potential and with the RH Scheme

	STO-3G		6-31G(d)		6-311+G(d,p)		aug-cc-pVTZ	
	RH	PHO	RH	PHO	RH	PHO	RH	PHO
C–H	0.001	0.001	0.002	0.003	0.004	0.004	0.002	0.002
C–C _B	0.080	0.072	0.159	0.155	0.127	0.118	0.076	0.070
C _B –H _B	0.011	0.013	0.019	0.021	0.023	0.025	0.025	0.027
H–C–H	1.2	1.3	1.0	1.0	0.8	0.9	1.6	1.6
H–C–C _B	1.1	1.2	0.9	1.0	0.7	0.8	1.5	1.5
C _A –C _B –H _B	1.0	1.9	1.6	0.4	1.5	1.4	1.4	1.1
H _B –C _B –H _B	1.0	2.0	1.6	0.4	1.5	1.5	1.5	1.1

the other in the MM region along with the PHO boundary atom (C_B), this can be tested simply by inspecting the net partial charge of each methyl group since a perfectly balanced QM and MM division would have no charge transfer between the two methyl groups. We used Hartree–Fock (HF) theory and two popular hybrid density functionals, B3LYP and M06-2X, combined with four representative basis sets, including STO-3G, 6-31G(d), 6-311+G(d,p), and aug-cc-pVTZ to illustrate the performance of the PHO method. In each QM model and basis set combination, the optimized structures and electronic properties obtained from the hybrid QM/MM calculations were compared with those from the corresponding full QM results.

Table 1 lists the mean unsigned errors (MUEs) both in bond length and in bond angle of ethane between the hybrid QM/MM method and the full QM calculations over HF, B3LYP and M06-2X optimizations for a given basis set. The range of the four basis functions selected are deemed to be sufficient to illustrate the generality of the PHO method here, although a broader range of other basis sets have been tested. Both results obtained with and without the E_B^{tet} correction term (eq 21) are shown. First, the optimized bond lengths in the QM region and in the MM region are in good accord with the full QM results, except the C–C_B bond across the two regions, which exhibits greater deviations, ranging from 0.054 to 0.155 Å. We attribute the larger error in the C–C_B bond length to the restriction on the projected orbitals, which are not relaxed in the SCF optimization. The errors in Table 1 do not show systematic dependence on the size of the basis set, although it is important to note that the C–H bonds report the perturbation to the QM region due to the boundary atom and that results on the C_B–H_B bonds are mainly dictated by the MM force field. The error trends in Table 1 have been found previously in the semiempirical^{20,21} and ab initio GHO-AIHF models,³⁰ and we consider that they are acceptable in a combined QM/MM method.

Without introducing any correction terms (Table 1, NC column), we found that the bond angles about the boundary atom C_B have significant errors in QM/MM geometry optimizations. The origin of this discrepancy is due to the use of system-independent auxiliary hybrid orbitals that are not optimized (or localized) for the specific bonding environment from model calculations.^{18,25,32–34} In PHO, as well as the original GHO,^{20,21,30} the three auxiliary orbitals are represented by atomic orbitals with an effective charge density of $(1 - q_B/3)$, where q_B is the partial atomic charge on the boundary atom defined in the MM force field. The electronic Pauli repulsions do not account for the full electron pair interactions among the three auxiliary orbitals, and they are weaker than interactions with the C–C_B bond orbital in the full QM system. This results in a locally distorted tetrahedral geometry. To correct this local structure distortion, a tetrahedral restraining potential, E_B^{tet} , is introduced (eq 21). As it turns out, a single parameter (the force constant) is sufficient for all basis sets examined. With the inclusion of this term, which is system-independent, the mean unsigned errors in the optimized bond angles are of similar size as that in all other bond types (Table 1, PHO column).

Alternatively, the Pauli repulsion error among the auxiliary orbitals can be removed by using the RH scheme examined by Jung and Ten-no. We have implemented this scheme on a different application to partition a system into QM-QM fragments, and the errors on QM/MM representation of ethane are compared with those of the PHO optimizations in Table 2. Clearly, the RH scheme and the tetrahedral restrained PHO method yield essentially the same results for all basis sets tested. The RH approach has the elegance without introducing empirical parameters other than fixing the transformation matrix, whereas the original definition of the hybridization matrix in the GHO method is directly related to the instantaneous geometry of the system.

The electronic polarization between the two methyl groups of ethane that are partitioned into the QM and MM subsystems was characterized by comparison of the partial atomic charges obtained from the hybrid QM/MM method and from the corresponding full QM calculations. Because of symmetry, the net atomic charge of each methyl fragment must be zero if the entire system is treated by the same method. In a QM/MM partition, and the difference between the two methyl groups provides a simple, but most critical, test of the relative electron-withdrawing power between the QM and MM subsystems. Two charge models, namely Mulliken population analysis (MPA)⁶³ and Charge Model 5 (CM5),⁶⁴ were used; the former is known to be a poor model for large basis sets (although it is convenient to use), whereas the latter is a parametrized density-mapping approach that shows remarkable stability across different basis sets and theoretical models.⁶⁴

Figures 2 and 3 illustrate deviations in partial charge between the PHO model and the full QM results, $\Delta q = q(\text{PHO}) -$

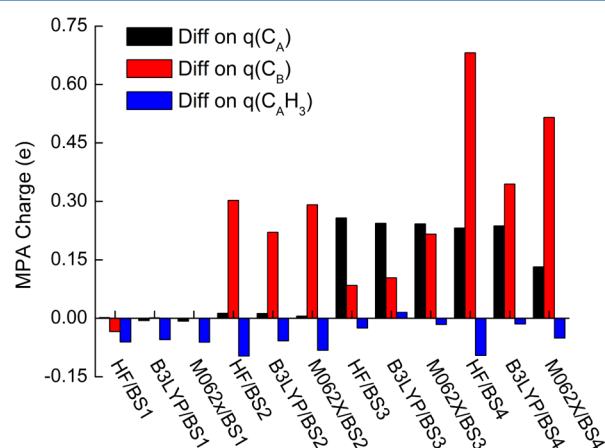


Figure 2. Difference in Mulliken population charges on the QM carbon (C_A) and the boundary carbon atom (C_B), and on the “QM” methyl group between PHO-QM/MM and full QM calculations. The basis set abbreviations BS1 through BS4 denote STO-3G, 6-31G(d), 6-311+G(d,p), and aug-cc-pVTZ.

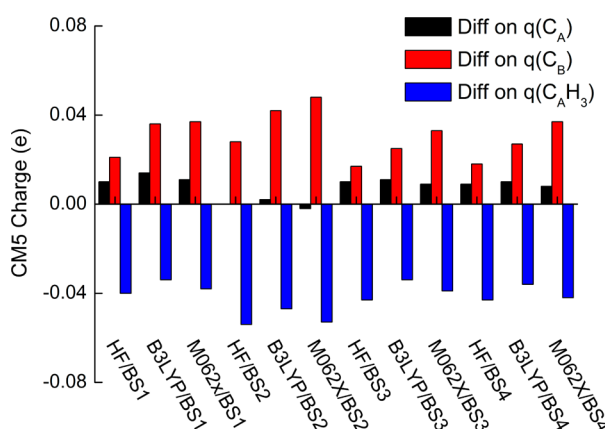


Figure 3. Difference in CM5 charges on the QM carbon (C_A) and the boundary carbon (C_B), and on the “QM” methyl group between PHO-QM/MM and full QM calculations. The basis set abbreviations are the same as those in Figure 2.

$q(\text{QM})$. Not surprisingly, the partial atomic charges on the individual atoms from MPA are very sensitive to the basis set used and have large errors between the PHO and full QM results

(Figure 2), but the total net charges of the entire methyl group are quite stable. Importantly, the total charge on the $C_B H_3$ group in the hybrid method exhibits only relatively small departure from the ideal value, indicating that charge transfer between QM and MM fragments is reasonably small. The largest errors came from HF/6-31G(d) and HF/aug-cc-pVTZ, which showed a net charge accumulation of about -0.07 e. On the other hand, the partial charges determined using the CM5 method have smaller differences between the PHO and the full QM methods on the individual atoms as well as the methyl fragments (Figure 3). For all methods examined, the CM5 charge deviations in PHO-QM/MM calculations are less than 0.05 e. The small deviation may also be attributed to the unrelaxed nature of the projected secondary orbitals. This restriction can be lifted,⁶⁵ which will be addressed in a forthcoming study, but the results displayed in Figures 2 and 3 indicate that the small charge deviations in the PHO method are acceptable for QM/MM applications.

B. Molecular Geometries and Further Validation of the PHO Method. To validate the performance of the PHO method, we constructed a set of 20 small compounds that include alkanes and molecules with different functional groups (Supporting Information). The optimized geometries from the hybrid PHO-QM/MM method were compared against results from HF treatment of the full system with three basis sets, STO-3G, 6-31G(d), and 6-311+G(d,p). In addition, we included results obtained by using the RH scheme to define the hybrid orbitals. We focus on bond lengths and bond angles associated with the frontier atom C_A in the QM fragment and the boundary atom C_B of the MM fragment (Table 3). The MUEs in bond

Table 3. MUEs in Bond Length (Å) and Bond Angle (deg) from Hybrid PHO-QM/MM and Restricted Hybridization (RH) Optimizations Relative to Full Hartree–Fock Optimizations for a Set of 20 Small Compounds^a

	STO-3G		6-31G(d)		6-311+G(d,p)	
	GHO	RH	GHO	RH	GHO	RH
C_A-Q	0.005	0.003	0.004	0.003	0.005	0.003
C_A-C_B	0.068	0.076	0.158	0.163	0.108	0.116
C_B-M	0.022	0.022	0.026	0.026	0.029	0.029
$Q-C_A-Q$	1.3	1.4	1.3	1.4	1.5	1.5
$Q-C_A-C_B$	1.3	1.5	1.3	1.4	1.4	1.4
C_A-C_B-M	2.6	1.4	1.8	1.8	2.6	1.7
$M-C_B-M$	2.4	1.1	1.2	1.5	2.2	1.5

^aQ indicates an atom present in the QM region, M is an atom treated classically by an MM model, and C_A and C_B are the frontier atoms in the QM and MM regions.

length are about 0.005 Å for covalent bonds between C_A and atoms in the QM region (C_A-Q), 0.03 Å between C_B and atoms in the MM region (C_B-M), and 0.068 to 0.158 Å in the QM/MM frontier bond (C_A-C_B). For bond angles in the QM region, the MUEs are less than 1.5° , whereas the errors are slightly larger for bond angles about C_B . Similar trends are obtained employing the RH transformation matrix. Overall, the geometrical results in Table 3 have mean unsigned errors of similar magnitude as that in ethane used to define the PHO method.

To illustrate the variation of the frontier bond between QM and MM regions as it is influenced by the neighboring functional group, the optimized bond lengths, $C-C_B$, for some key functional groups are compared for different basis sets (Table 4).

Table 4. Comparison of the Bond Length C_A-C_B for Selected Molecules between PHO and Standard HF Optimizations Using Different Basis Sets^a

	STO-3G		6-31G(d)		6-311+G(d,p)	
	PHO	HF	PHO	HF	PHO	HF
$\text{HOOC}_A-\text{C}_B\text{H}_3$	1.589	1.537	1.619	1.502	1.569	1.500
$\text{HCOCC}_A\text{H}_2-\text{C}_B\text{H}_3$	1.616	1.542	1.629	1.504	1.578	1.502
$\text{C}_6\text{H}_5-\text{C}_B\text{H}_3$	1.596	1.527	1.659	1.512	1.613	1.511
$\text{HOC}_A\text{H}_2-\text{C}_B\text{H}_3$	1.613	1.547	1.675	1.522	1.622	1.520
$\text{HOCC}_A\text{H}_2-\text{C}_B\text{H}_3$	1.611	1.539	1.681	1.524	1.631	1.523
$\text{HOCH}_2\text{C}_A\text{H}_2-\text{C}_B\text{H}_3$	1.613	1.540	1.692	1.528	1.642	1.528
$\text{H}_2\text{NC}_A\text{H}_2-\text{C}_B\text{H}_3$	1.615	1.546	1.690	1.529	1.639	1.528
$\text{H}_2\text{NCH}_2\text{C}_A\text{H}_2-\text{C}_B\text{H}_3$	1.614	1.540	1.693	1.528	1.643	1.528
$\text{H}(\text{O}=\text{C})_A-\text{C}_B\text{H}_3$	1.594	1.536	1.696	1.533	1.642	1.532

^a C_B is the boundary atom, and atoms to the right of C_B belong to the MM region.

Although the optimized frontier bonds from the PHO method are uniformly longer than the corresponding HF results, the trends due to functional group substitutions are reasonably reproduced.

C. Torsional Potential Energy Profile of *n*-Butane. The potential energy profile about the central C_2-C_3 single bond torsion as well as the equilibrium geometries of *n*-butane were determined using full QM and combined QM/MM with the PHO boundary atom, C_B , placed at C_2 , C_3 , and C_4 positions at the HF/6-31G(d) level (Figure 4). For comparison, the torsional

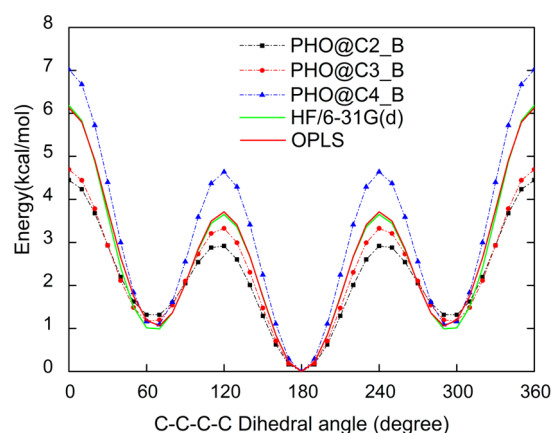


Figure 4. Torsional barriers about the C_2-C_3 bond in *n*-butane. Complete energy minimization was performed at each constrained value of the dihedral angle. PHO calculations are done at HF/6-31G(d) level, with the boundary atom placed at the C_2 , C_3 and C_4 carbons, denoted by PHO@ C_2_B , PHO@ C_3_B and PHO@ C_4_B , respectively.

potential energy profile from the OPLS force field,⁶⁶ which was also used in the QM/MM calculations of *n*-butane, was included.

The torsion energies from the OPLS potential and the HF/6-31G(d) optimization are very similar, whereas the QM/MM results tend to underestimate the full eclipse and partial eclipse conformational energy by about 1.5 and 0.5 kcal/mol, respectively, when the boundary atom is placed at the C_2 and C_3 positions. In the latter cases, the dominant contributions to the torsional energy in QM/MM calculations are due to the MM terms. Thus, the local QM/MM electrostatic interactions, which are fully included in the PHO method, lowers the energies of the full and partial eclipsed conformers. When the boundary atom is placed at the C_4 position, the potential energy profile is largely determined by the QM method, and the higher energies for the full and partial eclipsed conformers are due to the inclusion of the unmodified torsional terms in the MM force field. In all situations, the relative energy between the trans and gauche conformers are in good agreement with the corresponding QM and MM results. Overall, the PHO method yields qualitatively correct results of the internal rotations about the C_2-C_3 bond in *n*-butane, although slight modifications of the torsion terms may result in better agreement for high-energy conformations.

D. Deprotonation Energy. To evaluate the performance of the PHO method in combined QM/MM calculations, the deprotonation energies of selected organic acids are computed. Here, the purpose is not on the accuracy of the electronic structure method on the computed deprotonation energy relative to the experimental data, but, rather, the main goal is to evaluate the errors introduced in the present QM/MM approach employing the PHO boundary method with respect to the results from the corresponding full QM calculations. The deprotonation energy (DPE) is defined as the zero-point-exclusive energy difference between the protonated and deprotonated species, which differs from the experimental proton affinity, corresponding to the enthalpy difference. For simplicity, we used a single set of molecular geometries for all

Table 5. Computed Deprotonation Energies (kcal/mol) Using HF/6-311+G(d,p) and Signed Errors from the PHO-QM/MM Calculation and the ONIOM Method for the All-Trans Conformation of *n*-Octanol, 1-Hydroxy-($x+1$)-octanone, and x -Methoxy-1-octanol, Where x Denotes the PHO Boundary Atom Placed at the x th Carbon along the Main Chain

C_x	<i>n</i> -octanol		1-HO-($x+1$)-octanone			x -MeO-1-octanol		
	HF	PHO	HF	PHO	ONIOM	HF	PHO	ONIOM
C_2	396.3	5.8	388.4	7.5	8.6	390.6	7.8	8.1
C_3	396.3	3.5	390.1	4.9	6.9	393.3	5.9	8.4
C_4	396.3	2.0	394.5	3.1	4.7	395.1	4.2	5.8
C_5	396.3	0.7	395.0	1.6	3.4	395.6	2.8	3.8
C_6	396.3	0.2	395.3	0.8	2.0	395.9	1.6	2.2

compounds (Supporting Information), optimized using M06-2X/6-311+G(d,p).

We first examined the effect of the boundary atom placed at different distances away from a functional group. The all-trans conformation of *n*-octanol (and its conjugated base) was chosen for this test. In addition, we have placed two functional groups next to the boundary atom: a carbonyl group in the MM region along the carbon chain, and a methoxy group as a substituent. Specifically, if the boundary atom C_x is located on the x th carbon along the carbon chain, the carbonyl compound is called 1-hydroxy-($x+1$)-octanone, and the methoxy-substituted compound is called x -methoxy-1-octanol. For comparison, we have also computed the DPEs using the electrostatically embedded ONIOM model, where the C_x atom is the frontier atom in the MM region. For these compounds, HF/6-311+G(d,p) was used in full QM and hybrid QM/MM single-point energy calculations.

Table 5 shows that significant errors exist when the boundary atom is placed too close to the functional group. When the boundary atom is directly connected to the alcohol group, an error of 5.8 kcal/mol is found in the present PHO-QM/MM calculation for *n*-octanol. The errors are increased by about 2 kcal/mol when additional functional groups are included in the MM region. We see a gradual decrease in computation error as the QM-MM boundary is moved further away from the alcoholic functional group. Chemical accuracy with errors less than 1 kcal/mol can be obtained when the boundary atom is placed at least three covalent bonds away at the C_5 position in *n*-octanol, whereas at least another covalent bond away is required for compounds with neighboring functional groups. For the two functionalized alcohol series tested, the PHO boundary approach appears to have a better agreement with the full HF results than the standard ONIOM method.

The deprotonation energies further tested and compared with the corresponding full QM results for 22 selected organic acids, including alcohols, carboxylate acids, and alkylammonium ions. Single-point energy calculations were performed to yield the reference DPEs at HF/6-31G(d), HF/6-311+G(d,p), M062X/6-31G(d), and B3LYP/6-31G(d) levels, and the PHO-QM/MM results employing both the global Löwdin orthogonalization (GLO) and projected basis (PB) schemes. The computed protonation energies are provided as Supporting Information. The mean unsigned errors for the computed DPEs using different methods are summarized in Table 6. In this collection of

for the methods tested, and they do not seem to be systematically dependent on the details of the orthogonalization scheme used. This error range is consistent with the test case illustrated for *n*-octanol in Table 5, and the strong message from these tests is that the MM region should be defined at least three covalent bonds (or four carbon atoms) away from the functional group where chemical transformation takes place.

The average error on the computed proton affinities using the present PHO method may be comparable to that from other QM/MM boundary treatments. Amara et al. used the hydrogen link-atom model and found an average error of about 3 kcal/mol.³⁸ Zhang et al. used a pseudobond approach with various basis sets to yield MUEs of 2.9 to 7.7 kcal/mol in computed proton affinity.³⁵ In the GHO-AIHF method with parametrized STO-3G(v) basis set on the boundary carbon, an average error of 2.6 kcal/mol was reported using the MIDI! basis set for the QM region.³⁰

5. CONCLUSIONS

A PHO method for treating covalent boundary between QM and MM subsystems in combined QM/MM calculations has been developed. The PHO approach can be used in ab initio wave function theory (WFT) and density functional theory (DFT) with any basis set without introducing system-dependent parameters. The present method is an extension of the original GHO technique by introducing a secondary, minimum basis set on the boundary atom that separates the QM and MM region in order to formulate a set of hybrid orbitals. The primary basis set on the boundary atom, which is the same as that used in the QM region, is projected onto the secondary basis by means of Mulliken modified atomic orbital projection to yield a best representation. Thus, the PHOs in the secondary basis set provide a good approximation to the electron-withdrawing power of the primary basis set to balance electronic interactions between QM and MM regions.

The PHO method has been tested on a range of molecules and properties using WFT and DFT along with several representative basis sets. Comparison with results obtained from QM calculations on the entire system shows that the present PHO method is a robust and balanced QM/MM scheme that preserves the structural and electronic properties of the QM region. The computed torsional potential energy profiles about the central C–C bond of butane are qualitatively in agreement with the corresponding QM and MM results, respectively, with the energies of the full and partial eclipsed conformers underestimated by 0.5 to 1.5 kcal/mol when the boundary atom is placed at the C_2 and C_3 positions, but they are overestimated by about 1 kcal/mol when the boundary atom is located on C_4 . The agreement between full QM and PHO-QM/MM results on protonation energies is found to be dependent on the distance of the boundary atom from the functional group. When the boundary atoms are anchored directly to the functional group or placed just one covalent bond away, the mean unsigned errors in protonation energy are 2.6 to 4.8 kcal/mol. A mean deviation of less than 1 kcal/mol from the full QM results can be obtained if the boundary atom is placed at least 3 covalent bonds away from the functional group. Thus, it is important to separate the MM region sufficiently far away from the active site when chemical transformations take place in combined QM/MM calculations.

Table 6. MUEs (kcal/mol) on Computed Protonation Energies between PHO-QM/MM and Full QM Calculations^a

QM method	PB	GLO
HF/6-31G(d)	3.8	3.6
HF/6-311+G(d,p)	2.6	4.8
M062X/6-31G(d)	3.8	4.0
B3LYP/6-31G(d)	4.0	3.6

^aThe QM method used in combined QM/MM calculations is indicated, and the AMBER force field is employed to model the MM region. PB denotes the projected basis method and GLO specifies the global Löwdin orthogonalization scheme.

molecules, whose sizes are relatively small, the boundary atoms are located at the C_2 and C_3 carbons (i.e., directly connected to, or one bond away from the functional group), except one case in which it is situated at the C_4 position. Thus, the errors shown in Table 6 represent the worst case scenarios in QM/MM applications. The MUEs are in the range of 2.6 to 4.8 kcal/mol

■ ASSOCIATED CONTENT

■ Supporting Information

Optimized geometries, and computed deprotonation energies for a set of 22 organic acids using HF, B3LYP and M06-2X methods are provided. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: gao@jialigao.org.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the National Institutes of Health (GM46736) and National Science Foundation (CHE09-57162) for partial support of this research.

■ REFERENCES

- (1) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. SN2 reaction profiles in the gas phase and aqueous solution. *J. Am. Chem. Soc.* **1984**, *106* (10), 3049–3050.
- (2) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. Theoretical examination of the SN2 reaction involving chloride ion and methyl chloride in the gas phase and aqueous solution. *J. Am. Chem. Soc.* **1985**, *107* (1), 154–163.
- (3) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249.
- (4) Singh, U. C.; Kollman, P. A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH₃Cl + Cl[−] exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7* (6), 718–730.
- (5) Bash, P. A.; Field, M. J.; Karplus, M. Free energy perturbation method for chemical reactions in the condensed phase: A dynamic approach based on a combined quantum and molecular mechanics potential. *J. Am. Chem. Soc.* **1987**, *109* (26), 8092–8094.
- (6) Gao, J.; Xia, X. A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* **1992**, *258* (5082), 631–635.
- (7) Gao, J. Methods and applications of combined quantum mechanical and molecular mechanical potentials. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; Wiley VCH: New York, 1995; Vol. 7, pp 119–185.
- (8) Gao, J. Hybrid quantum and molecular mechanical simulations: An alternative avenue to solvent effects in organic chemistry. *Acc. Chem. Res.* **1996**, *29* (6), 298–305.
- (9) Field, M. J. Simulating enzyme reactions: Challenges and perspectives. *J. Comput. Chem.* **2002**, *23* (1), 48–58.
- (10) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. Mechanisms and free energies of enzymatic reactions. *Chem. Rev.* **2006**, *106* (8), 3188–3209.
- (11) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48* (7), 1198–1229.
- (12) Eurenio, K. P.; Chatfield, D. C.; Brooks, B. R.; Hodoscek, M. Enzyme mechanisms with hybrid quantum and molecular mechanical potentials. I. Theoretical considerations. *Int. J. Quantum Chem.* **1996**, *60* (6), 1189–1200.
- (13) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.* **1999**, *110* (21), 10452–10467.
- (14) Lyne, P. D.; Hodoscek, M.; Karplus, M. A hybrid QM-MM potential employing Hartree–Fock or density functional methods in the quantum region. *J. Phys. Chem. A* **1999**, *103* (18), 3462–3471.
- (15) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. A* **2000**, *104* (8), 1720–1735.
- (16) Ferré, N.; Olivucci, M. The amide bond: Pitfalls and drawbacks of the link atom scheme. *J. Mol. Struct.: THEOCHEM* **2003**, *632* (1–3), 71–82.
- (17) Zhang, Y. Pseudobond ab initio QM/MM approach and its applications to enzyme reactions. *Theor. Chem. Acc.* **2006**, *116* (1–3), 43–50.
- (18) Monari, A.; Rivail, J.-L.; Assfeld, X. Theoretical modeling of large molecular systems. Advances in the local self consistent field method for mixed quantum mechanics/molecular mechanics calculations. *Acc. Chem. Res.* **2012**, *46* (2), 596–603.
- (19) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11* (6), 700–733.
- (20) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A* **1998**, *102* (24), 4714–4721.
- (21) Amara, P.; Field, M. J.; Alhambra, C.; Gao, J. The generalized hybrid orbital method for combined quantum mechanical/molecular mechanical calculations: formulation and tests of the analytical derivatives. *Theor. Chem. Acc.* **2000**, *104* (5), 336–343.
- (22) Ferenczy, G. G.; Rivail, J.-L.; Surján, P. R.; Náray-Szabó, G. NDDO fragment self-consistent field approximation for large electronic systems. *J. Comput. Chem.* **1992**, *13* (7), 830–837.
- (23) Théry, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. G. Quantum mechanical computations on very large molecular systems: The local self-consistent field method. *J. Comput. Chem.* **1994**, *15* (3), 269–282.
- (24) Monard, G.; Loos, M.; Théry, V.; Baka, K.; Rivail, J.-L. Hybrid classical quantum force field for modeling very large molecules. *Int. J. Quantum Chem.* **1996**, *58* (2), 153–159.
- (25) Assfeld, X.; Rivail, J.-L. Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem. Phys. Lett.* **1996**, *263* (1–2), 100–106.
- (26) Antonczak, S.; Monard, G.; Ruiz-López, M. F.; Rivail, J.-L. Modeling of peptide hydrolysis by thermolysin. A semiempirical and QM/MM study. *J. Am. Chem. Soc.* **1998**, *120* (34), 8825–8833.
- (27) Ferré, N.; Assfeld, X. Application of the local self-consistent-field method to core-ionized and core-excited molecules, polymers, and proteins: True orthogonality between ground and excited states. *J. Chem. Phys.* **2002**, *117* (9), 4119–4125.
- (28) Maseras, F.; Morokuma, K. IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* **1995**, *16* (9), 1170–1179.
- (29) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels–Alder reactions and Pt(P(*t*-Bu)₃)₂ + H₂ oxidative addition. *J. Phys. Chem.* **1996**, *100* (50), 19357–19363.
- (30) Pu, J.; Gao, J.; Truhlar, D. G. Generalized hybrid orbital (GHO) method for combining ab initio Hartree–Fock wave functions with molecular mechanics. *J. Phys. Chem. A* **2004**, *108* (4), 632–650.
- (31) Pu, J.; Gao, J.; Truhlar, D. G. Generalized Hybrid-Orbital method for combining density functional theory with molecular mechanics. *ChemPhysChem* **2005**, *6* (9), 1853–1865.
- (32) Philipp, D. M.; Friesner, R. A. Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *J. Comput. Chem.* **1999**, *20* (14), 1468–1494.
- (33) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. Frozen orbital QM/MM methods for density functional theory. *Chem. Phys. Lett.* **2000**, *321* (1–2), 113–120.
- (34) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *J. Comput. Chem.* **2000**, *21* (16), 1442–1457.

- (35) Zhang, Y.; Lee, T.-S.; Yang, W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.* **1999**, *110* (1), 46–54.
- (36) Antes, I.; Thiel, W. Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. *J. Phys. Chem. A* **1999**, *103* (46), 9290–9295.
- (37) Swart, M. AddRemove: A new link model for use in QM/MM studies. *Int. J. Quantum Chem.* **2003**, *91* (2), 177–183.
- (38) Amara, P.; Field, M. J. Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor. Chem. Acc.* **2003**, *109* (1), 43–52.
- (39) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem.* **2003**, *24* (6), 760–769.
- (40) DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. Simple one-electron quantum capping potentials for use in hybrid QM/MM studies of biological molecules. *J. Chem. Phys.* **2002**, *116* (22), 9578–9584.
- (41) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. Variational optimization of effective atom centered potentials for molecular properties. *J. Chem. Phys.* **2005**, *122* (1), 14113.
- (42) Lin, H.; Truhlar, D. G. Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations. *J. Phys. Chem. A* **2005**, *109* (17), 3991–4004.
- (43) Wang, B.; Truhlar, D. G. Combined quantum mechanical and molecular mechanical methods for calculating potential energy surfaces: Tuned and balanced redistributed-charge algorithm. *J. Chem. Theory Comput.* **2010**, *6* (2), 359–369.
- (44) Wang, B.; Truhlar, D. G. Tuned and balanced redistributed charge scheme for combined quantum mechanical and molecular mechanical (QM/MM) methods and fragment methods: tuning based on the CM5 charge model. *J. Chem. Theory Comput.* **2013**, *9* (2), 1036–1042.
- (45) Xiao, C.; Zhang, Y. Design-atom approach for the quantum mechanical/molecular mechanical covalent boundary: A design-carbon atom with five valence electrons. *J. Chem. Phys.* **2007**, *127* (12), 124102.
- (46) Jung, J.; Choi, C. H.; Sugita, Y.; Ten-no, S. New implementation of a combined quantum mechanical and molecular mechanical method using modified generalized hybrid orbitals. *J. Chem. Phys.* **2007**, *127* (20), 204102.
- (47) Jung, J.; Ten-no, S. Simple avoidance of Pauli repulsion errors in the generalized hybrid orbital (GHO) method. *Chem. Phys. Lett.* **2010**, *484* (4–6), 344–348.
- (48) Schiffmann, C.; Sebastiani, D. Artificial Bee Colony Optimization of Capping Potentials for Hybrid Quantum Mechanical/Molecular Mechanical Calculations. *J. Chem. Theory Comput.* **2011**, *7* (5), 1307–1315.
- (49) Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7* (3), 761–777.
- (50) Das, D.; Eurenus, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *J. Chem. Phys.* **2002**, *117* (23), 10534–10547.
- (51) Kawashima, Y.; Nakano, H.; Jung, J.; Ten-no, S. A combined quantum mechanical and molecular mechanical method using modified generalized hybrid orbitals: Implementation for electronic excited states. *Phys. Chem. Chem. Phys.* **2011**, *13* (24), 11731–11738.
- (52) Akinaga, Y.; Jung, J.; Ten-no, S. QM/MM calculation of protein magnetic shielding tensors with generalized hybrid-orbital method: A GIAO approach. *Phys. Chem. Chem. Phys.* **2011**, *13* (32), 14490–14499.
- (53) Kitagawa, Y.; Akinaga, Y.; Kawashima, Y.; Jung, J.; Ten-no, S. A QM/MM-MD study on protein electronic properties: Circular dichroism spectra of oxytocin and insulin. *Chem. Phys.* **2012**, *401* (0), 95–102.
- (54) Jung, J.; Re, S.; Sugita, Y.; Ten-no, S. Improved constrained optimization method for reaction-path determination in the generalized hybrid orbital quantum mechanical/molecular mechanical calculations. *J. Chem. Phys.* **2013**, *138* (4), 044106.
- (55) Montgomery, J., Jr.; Frisch, M.; Ochterski, J.; Petersson, G. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **2000**, *112* (15), 6532–6542.
- (56) Mulliken, R. S. Criteria for the construction of good self-consistent-field molecular orbital wave functions, and the significance of LCAO-MO population analysis. *J. Chem. Phys.* **1962**, *36* (12), 3428–3439.
- (57) Davidson, E. R. Electronic population analysis of molecular wavefunctions. *J. Chem. Phys.* **1967**, *46* (9), 3320–3324.
- (58) Löwdin, P. O. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **1950**, *18* (3), 365–375.
- (59) Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980**, *73* (2), 393–398.
- (60) Pulay, P. Improved SCF convergence acceleration. *J. Comput. Chem.* **1982**, *3* (4), 556–560.
- (61) Frisch, M.; Trucks, G.; Schlegel, H. B.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G., *Gaussian Development Version*; Gaussian, Inc.: Wallingford, CT, 2013.
- (62) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197.
- (63) Mulliken, R. S. Electronic population analysis on LCAO–MO molecular wave functions. I. *J. Chem. Phys.* **1955**, *23* (10), 1833–1840.
- (64) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge model 5: An extension of Hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.
- (65) Lee, M. S.; Head-Gordon, M. Polarized atomic orbitals for self-consistent field electronic structure calculations. *J. Chem. Phys.* **1997**, *107* (21), 9085–9095.
- (66) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657–1666.