Quantum mechanical/molecular mechanical/continuum style solvation model: Time-dependent density functional theory

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · AUGUST 2013

Impact Factor: 2.95 · DOI: 10.1063/1.4819139 · Source: PubMed

3 AUTHORS, INCLUDING:

Nandun M Thellamurege
University of Nebraska at Lincoln
9 PUBLICATIONS **38** CITATIONS
SEE PROFILE

Fengchao Cui
Changchun Institute of Applied Chemistry Ch…
9 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Quantum mechanical/molecular mechanical/continuum style solvation model: Time-dependent density functional theory

Nandun M. Thellamurege, Fengchao Cui, and Hui Li

Citation: The Journal of Chemical Physics **139**, 084106 (2013); doi: 10.1063/1.4819139
View online: http://dx.doi.org/10.1063/1.4819139
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/139/8?ver=pdfcov
Published by the AIP Publishing

Quantum mechanical/molecular mechanical/continuum style solvation model: Time-dependent density functional theory

Nandun M. Thellamurege, Fengchao Cui, and Hui Li[a)]
Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA

(Received 4 April 2013; accepted 9 August 2013; published online 27 August 2013)

A combined quantum mechanical/molecular mechanical/continuum (QM/MMpol/C) style method is developed for time-dependent density functional theory (TDDFT, including long-range corrected TDDFT) method, induced dipole polarizable force field, and induced surface charge continuum model. Induced dipoles and induced charges are included in the TDDFT equations to solve for the transition energies, relaxed density, and transition density. Analytic gradient is derived and implemented for geometry optimization and molecular dynamics simulation. QM/MMpol/C style DFT and TDDFT methods are used to study the hydrogen bonding of the photoactive yellow protein chromopore in ground state and excited state. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4819139]

I. INTRODUCTION

Combined quantum mechanical/molecular mechanical/continuum (QM/MM/C) models can be used to study solvent effects.[1–10] In such combined models, some solvent molecules around the QM solute molecules are described explicitly with MM force field models, and the bulk solvent is simplified as a continuum model. Using a polarizable force field in QM/MM calculations is advantageous because the electronic polarization of the MM region can be described, especially when electronic excitation is involved in the QM region. A conceptually simple and efficient method to incorporate electronic polarization is to use induced point dipoles.[10–15] In this work, we focus on QM/MMpol/C style methods that incorporate induced point dipoles into the MM force field to describe the electronic polarization of the MM region and induced surface charges to describe the dielectric polarization of the bulk solvent.

Variational Hartree-Fock (HF) and density functional theory (DFT) methods are of fundamental importance in quantum chemistry. For example, analytic gradients and some other properties can be evaluated efficiently, and post-HF and post-DFT methods can be formulated. It is important to incorporate solvation models into HF and DFT methods in a variational manner. A rigorous variational treatment of induced dipoles and induced surface charges for QM/MMpol/C style HF and DFT methods were described in a previous work.[9]

Time-dependent density functional theory (TDDFT) methods are efficient and relatively accurate QM methods for the study of valence and singly excited states.[16,17] Analytical gradients for QM/MMpol style and QM/continuum style TDDFT methods have been derived and implemented by the authors group.[18,19] Recently, Steindal *et al.*[8] developed a QM/MMpol/C style TDDFT method (but with no analytic gradient). In this work, the analytic gradient is derived and implemented for a QM/MMpol/C style TDDFT method.

a)Author to whom correspondence should be addressed. Electronic mail: hli4@unl.edu

Photoactive yellow protein (PYP) is a 125-residue photoreceptor found in the bacterium *Ectothiprhodospira* (or *halorhodospira*) *halophila*.[20] After blue light absorption (maximum 446 nm or 2.78 eV[20]), the chromophore in PYP undergoes an ultrafast *trans*-to-*cis* isomerization in the time scale of a few ps.[21] Due to its relatively small size and thermal stability, PYP has been a model system for extensive experimental and theoretical studies. Many theoretical studies used complete active space self-consistent field (CASSCF) or complete active space second order perturbation theory (CASPT2) method,[22–26] or coupled cluster method.[27] TDDFT has also been used to study the PYP chromophore and its analogs, for example, by Sergi *et al.*,[28] Groenhof *et al.*,[29] Thompson *et al.*,[30] Levine *et al.*,[31] Isborn *et al.*,[32] González *et al.*,[33] and Rocha-Rinza *et al.*[34]

In this work, we use QM/MMpol/C style DFT and TDDFT methods to study the hydrogen bonding of the PYP chromophore in the ground (S_0) and the singlet exited states (S_1, but could be S_2). According to experimental and computational studies,[35–44] in the ground state the PYP chromophore is a deprotonated *p*-coumaric acid (*p*CA) in hydrogen bonding with Tyr42 and Glu46. It is believed that after photoexcitation, *p*CA remains anionic, finishes the isomerization and deexcitation in the anionic state, and evolves into an anionic ground state intermediate denoted as I_0 in 1–3 ps. The I_0 intermediate then evolves into an intermediate denoted as I_1 in 1-3 ns.[45–50] The anionic chromophore of I_1 is protonated by Glu46[22,35,50–53] or the solvent[50,54,55] and evolves into an intermediate denoted as I_2 in a time scale of ∼200 μs.[56,57] The I_2 intermediate returns back to the dark state in a time scale of 1-3 s, completing the PYP photocycle.

We first describe the theory and implementation of the QM/MMpol/C style TDDFT gradient, then the computational methods used for the calculation of PYP. The hydrogen bonding distances and their strengths at the PYP active site are analyzed and discussed, and compared to theoretical and experimental results in the literature. Finally, we summarize the results and conclude.

II. THEORY

A. QM/MMpol/C style DFT gradient

The variational theory for QM/MMpol/C style density functional theory and Hartree-Fock methods has been established earlier.[9] Here, the main equations are summarized for the purpose of establishing the notations for the current paper.

The total energy of a molecular system described with a QM/MMpol/C style DFT method (including Hartree-Fock as a special case) can be written as

$$E_{DFT/MMpol/C} = E_{DFT} + E_{MM} + E_{rep} + E_{disp} + E_{mul} + E_{polsol}. \quad (1)$$

E_{DFT} in Eq. (1) is the DFT electronic (including nuclear repulsion) energy of the QM region,

$$E_{DFT} = \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} h_{\mu\nu} + \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} (\mu\nu|\lambda\kappa) - c_x \delta_{\sigma\sigma'} \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} (\mu\lambda|\nu\kappa) + \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} f^{XC}_{\mu\nu\sigma,\lambda\kappa\sigma'} + E_{N,N}. \quad (2)$$

In Eq. (2), μ, ν, λ, and κ denote basis functions (Mulliken notation is used for two-electron integrals); σ and σ' (and later σ'') are spin indices; $\mathbf{P}$ is the one-particle density matrix; $\mathbf{h}$ is the one-electron integral; c_x is the mixing coefficient of the Hartree-Fock exchange energy; $f^{XC}_{\mu\nu\sigma,\lambda\kappa\sigma'}$ is the DFT exchange-correlation kernel; $E_{N,N}$ is the nuclear repulsion energy.

E_{MM} in Eq. (1) is the force field energy of the MM region.

E_{rep} and E_{disp} in Eq. (1) are the repulsion and dispersion interaction energies, respectively, between QM electrons and MM repulsion and dispersion points; E_{mul} is the electrostatic interaction energy between QM electrons/nuclei and MM electric multipole points. Usually they are modeled by Lennard-Jones (LJ) potentials and atomic point charges

$$E_{rep} + E_{disp} + E_{mul} = E_{LJ} + \sum_{\mu\nu} P_{\mu\nu} V_{chg,\mu\nu} + E_{N,chg}. \quad (3)$$

In Eq. (3), $V_{chg,\mu\nu}$ is the one-electron integrals of the MM point charge potentials, $E_{N,chg}$ is the electrostatic energy between QM nuclei and MM force field atomic charges.

In the Effective Fragment Potential (EFP)[9,10,58,59] method, repulsion and dispersion are described with effective Gaussian potentials that directly act on the QM electrons, similar to the electric multipole interaction,

$$E_{rep} + E_{disp} + E_{mul} = \sum_{\mu\nu} P_{\mu\nu} V_{rep,\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} V_{disp,\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} V_{mul,\mu\nu} + E_{N,mul}. \quad (4)$$

Here, $V_{rep,\mu\nu}$, $V_{disp,\mu\nu}$, and $V_{mul,\mu\nu}$ are the one-electron integrals of the EFP repulsion, dispersion, and multipole potentials, $E_{N,mul}$ is the electrostatic energy between QM nuclei and EFP electric multipole points.

E_{polsol} in Eq. (1) is the polarization-solvation energy of the entire QM/MMpol/C system, including the induced dipole and induced charge contributions

$$\begin{aligned} E_{polsol} &= -\tfrac{1}{2}\mathbf{U}^T\mathbf{w} \\ &= -\tfrac{1}{2}(\mathbf{U}_N + \mathbf{U}_{mul} + \mathbf{U}_e)^T(\mathbf{w}_N + \mathbf{w}_{mul} + \mathbf{w}_e) \\ &= -\tfrac{1}{2}(\mathbf{U}_N + \mathbf{U}_{mul})^T(\mathbf{w}_N + \mathbf{w}_{mul}) \\ &\quad -\tfrac{1}{2}\sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} (\mathbf{U}_{\mu\nu})^T \mathbf{w}_{\lambda\kappa} \\ &\quad -\tfrac{1}{2}\sum_{\mu\nu\sigma} P_{\mu\nu\sigma} [(\mathbf{U}_N + \mathbf{U}_{mul})^T \mathbf{w}_{\mu\nu} \\ &\quad + (\mathbf{U}_{\mu\nu})^T(\mathbf{w}_N + \mathbf{w}_{mul})]. \end{aligned} \quad (5)$$

Here, $\mathbf{U}$ is the electrostatic fields/potentials at the dipole polarizability points/surface tesserae created by QM nuclei ($\mathbf{U}_N$), DFT density ($\mathbf{U}_e$), and MM electric multipoles ($\mathbf{U}_{mul}$); the superscript T denotes transpose; $\mathbf{w}$ is a set of induced dipoles/charges, which satisfies the following linear equation:

$$\mathbf{B}\cdot\mathbf{w} = \mathbf{U}. \quad (6)$$

Here, $\mathbf{B}$ is the supermatrix that describes the interactions among induced dipoles/charges.[9] Due to the linearity of Eq. (6), $\mathbf{w}$ can be determined separately for the nuclei ($\mathbf{w}_N$,) and DFT density ($\mathbf{w}_e$) of the QM region and the electric multipoles ($\mathbf{w}_{mul}$) of the MM region. In addition, the induced dipoles/charges can be determined for the electrostatic field/potential due to a product of two basis functions such as $\mu\nu$, for example, $\mathbf{w}_{\mu\nu}$ and $\mathbf{U}_{\mu\nu}$ in Eq. (5).

Once $E_{DFT/MMpol/C}$ is minimized via a self-consistent field procedure, Pulay's method[60] can be used to determine the first derivative (gradient) with respect to a coordinate x:

$$\begin{aligned} E^x_{DFT/MMpol/C} &= E^x_{DFT} + E^x_{MM} + E^x_{rep} + E^x_{disp} + E^x_{mul} + E^x_{polsol} \\ &= -\sum_{\mu\nu\sigma} W_{\mu\nu\sigma} S^x_{\mu\nu} + \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} h^x_{\mu\nu} \\ &\quad + \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} (\mu\nu|\lambda\kappa)^x \\ &\quad - c_x \delta_{\sigma\sigma'} \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} (\mu\lambda|\nu\kappa)^x \\ &\quad + \sum_{\mu\nu\sigma\lambda\kappa\sigma'} P_{\mu\nu\sigma} P_{\lambda\kappa\sigma'} f^{XC,x}_{\mu\nu\sigma,\lambda\kappa\sigma'} + E^x_{N,N} \\ &\quad + E^x_{MM} + \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} V^x_{rep,\mu\nu} + \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} V^x_{disp,\mu\nu} \\ &\quad + \sum_{\mu\nu\sigma} P_{\mu\nu\sigma} V^x_{mul,\mu\nu} + E^x_{N,mul} \\ &\quad - \frac{1}{2}[(\mathbf{U}_N + \mathbf{U}_{mul})^T(\mathbf{w}_N + \mathbf{w}_{mul})]^x \\ &\quad - \frac{1}{2}\sum_{\mu\nu\rho\sigma} P_{\mu\nu} P_{\rho\sigma} [(\mathbf{U}_{\mu\nu})^T \mathbf{w}_{\rho\sigma}]^x \end{aligned}$$

$$-\frac{1}{2}\sum_{\mu\nu}P_{\mu\nu}[(\mathbf{U}_N+\mathbf{U}_{mul})^T\mathbf{w}_{\mu\nu} + (\mathbf{U}_{\mu\nu})^T(\mathbf{w}_N+\mathbf{w}_{mul})]^x. \quad (7)$$

Here, **P** and **W** are the DFT density matrix and energy-weighted density matrix, respectively.

There is a general expression for the derivative of the dot product of an electric field/potential vector $\mathbf{U}_K$ and an induced dipole/charge vector $\mathbf{w}_L$,

$$\left[(\mathbf{U}_K)^T\mathbf{w}_L\right]^x = \left(\mathbf{U}_K^x\right)^T\mathbf{w}_L + (\tilde{\mathbf{w}}_K)^T\mathbf{B}^x\mathbf{w}_L + (\tilde{\mathbf{w}}_K)^T\mathbf{U}_L^x \quad (8)$$

with

$$\tilde{\mathbf{w}}_K = (\mathbf{B}^{-1})^T\mathbf{U}_K. \quad (9)$$

So the last three terms in Eq. (7) can be written as

$$\begin{aligned}&-\frac{1}{2}[(\mathbf{U}_N+\mathbf{U}_{mul})^T(\mathbf{w}_N+\mathbf{w}_{mul})]^x\\ &-\frac{1}{2}\sum_{\mu\nu\rho\sigma}P_{\mu\nu}P_{\rho\sigma}[(\mathbf{U}_{\mu\nu})^T\mathbf{w}_{\rho\sigma}]^x\\ &-\frac{1}{2}\sum_{\mu\nu}P_{\mu\nu}[(\mathbf{U}_N+\mathbf{U}_{mul})^T\mathbf{w}_{\mu\nu}+(\mathbf{U}_{\mu\nu})^T(\mathbf{w}_N+\mathbf{w}_{mul})]^x\\ &=-\frac{1}{2}[(\mathbf{U}^x)^T\mathbf{w}+(\tilde{\mathbf{w}})^T\mathbf{B}^x\mathbf{w}+(\tilde{\mathbf{w}})^T\mathbf{U}^x]. \quad (10)\end{aligned}$$

The first and third terms in the bracket on the right-hand side of Eq. (10) represent the forces and torques between the induced dipoles/charges and the QM nuclei, QM density, and the MM electric multipoles; the second term represents the forces and torques among the induced dipoles/charges. In the EFP method, the **B** matrix is not symmetric. If the **B** matrix is symmetric, we have $(\mathbf{B}^{-1})^T = \mathbf{B}^{-1}$ and $\tilde{\mathbf{w}}_K = \mathbf{w}_K$.

B. QM/MMpol/C style TDDFT gradient

In a QM/MMpol/C style TDDFT method, the MM region should affect the TDDFT excitation energy in two ways. First, the MM region (including induced dipoles) and the continuum solvent (induced charges) affect the ground state DFT spin orbitals and energies. Second, the MM induced dipoles (and only the induced dipoles) and the continuum induced surface charges directly affect the TDDFT excitation energies.

The central equation in combined QM/MMpol/C style TDDFT calculation of the excitation energy is the same as that in regular TDDFT calculations,

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} = \omega\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix}. \quad (11)$$

Here, **X** and **Y** are the linear response amplitudes in the excitation and deexcitation processes; ω is the excitation energy, **A** and **B** matrices are the orbital rotation Hessian,

$$\begin{aligned}A_{ia\sigma,jb\sigma'} &= \delta_{ij}\delta_{ab}\delta_{\sigma\sigma'}(\varepsilon_{a\sigma}-\varepsilon_{i\sigma}) + (ia\sigma|bj\sigma')\\ &\quad - c_x\delta_{\sigma\sigma'}(ib\sigma|aj\sigma')\\ &\quad + f^{XC}_{ia\sigma bj\sigma'} - \frac{1}{2}(\mathbf{U}_{ia\sigma})^T(\tilde{\mathbf{w}}_{jb\sigma'}+\mathbf{w}_{jb\sigma'}), \quad (12)\end{aligned}$$

$$\begin{aligned}B_{ia\sigma,jb\sigma'} &= (ia\sigma|bj\sigma') - c_x\delta_{\sigma\sigma'}(ij\sigma|ab\sigma') + f^{XC}_{ia\sigma bj\sigma'}\\ &\quad - \frac{1}{2}(\mathbf{U}_{ia\sigma})^T(\tilde{\mathbf{w}}_{jb\sigma'}+\mathbf{w}_{jb\sigma'}). \quad (13)\end{aligned}$$

i and j represent occupied orbitals, a and b represent virtual orbitals, ε represents orbital energies. Here, both $\tilde{\mathbf{w}}_{jb\sigma'}$ and $\mathbf{w}_{jb\sigma'}$ are used for general asymmetric polarization.

The first derivative of the TDDFT excitation energy ω with respect to a coordinate x is

$$\begin{aligned}\omega^x &= \sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}h^x_{\mu\nu} - \sum_{\mu\nu\sigma}W^{\Delta}_{\mu\nu\sigma}S^x_{\mu\nu} + \sum_{\mu\nu\sigma\lambda\kappa\sigma'}\Gamma^{\Delta}_{\mu\nu\sigma\lambda\kappa\sigma'}(\mu\nu|\lambda\kappa)^x\\ &\quad + \sum_{\mu\nu\sigma\lambda\kappa\sigma'}P^{\Delta}_{\mu\nu\sigma}P_{\lambda\kappa\sigma'}\left(f^{XC}_{\mu\nu\sigma,\lambda\kappa\sigma'}\right)^x\\ &\quad + \sum_{\mu\nu\sigma\lambda\kappa\sigma'}(X+Y)_{\mu\nu\sigma}(X+Y)_{\lambda\kappa\sigma'}\left(f^{XC}_{\mu\nu\sigma,\lambda\kappa\sigma'}\right)^x\\ &\quad + \sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}V^x_{rep,\mu\nu} + \sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}V^x_{disp,\mu\nu} + \sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}V^x_{mul,\mu\nu}\\ &\quad - \frac{1}{2}\sum_{\mu\nu\sigma\lambda\kappa\sigma'}\left(P^{\Delta}_{\mu\nu\sigma}P_{\lambda\kappa\sigma'} + P_{\mu\nu\sigma}P^{\Delta}_{\lambda\kappa\sigma'}\right)\\ &\quad \times\left[\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_{\lambda\kappa}+\mathbf{w}_{\lambda\kappa})\right]^x\\ &\quad - \frac{1}{2}\sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}\left[\frac{1}{2}(\mathbf{U}_N+\mathbf{U}_{mul})^T(\tilde{\mathbf{w}}_{\mu\nu}+\mathbf{w}_{\mu\nu})\right.\\ &\quad \left.+\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_N+\tilde{\mathbf{w}}_{mul}+\mathbf{w}_N+\mathbf{w}_{mul})\right]^x\\ &\quad - \sum_{\mu\nu\sigma\lambda\kappa\sigma'}(X+Y)_{\mu\nu\sigma}(X+Y)_{\lambda\kappa\sigma'}\\ &\quad \times\left[\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_{\lambda\kappa}+\mathbf{w}_{\lambda\kappa})\right]^x. \quad (14)\end{aligned}$$

Here, $\mathbf{P}^{\Delta}$ is the relaxed one-particle difference density matrix; **P** is the ground state one-particle density matrix; $\mathbf{W}^{\Delta}$ is the energy-weighted difference density matrix; $\mathbf{\Gamma}^{\Delta}$ is the two-particle difference density matrix, including the Hartree-Fock exchange contribution in hybrid DFT methods.

The first five terms in Eq. (14) have exactly the same forms as those in the regular TDDFT excitation energy gradient formula derived by Furche and Ahlrichs,[61,62] but the $\mathbf{P}^{\Delta}$, $\mathbf{W}^{\Delta}$, and $\mathbf{\Gamma}^{\Delta}$ must be determined from QM/MMpol/C style calculations.

$\mathbf{P}^{\Delta}$ is the sum of the unrelaxed difference density matrix $\mathbf{T}^{\Delta}$ and the **Z** density matrix,

$$\mathbf{P}^{\Delta} = \mathbf{T}^{\Delta} + \mathbf{Z}. \quad (15)$$

The unrelaxed difference density matrix $\mathbf{T}^{\Delta}$,

$$T^{\Delta}_{ab\sigma} = \frac{1}{2}\sum_i[(X+Y)_{ia\sigma}(X+Y)_{ib\sigma} + (X-Y)_{ia\sigma}(X-Y)_{ib\sigma}], \quad (16)$$

$$T^{\Delta}_{ij\sigma} = -\frac{1}{2}\sum_{a}[(X+Y)_{ia\sigma}(X+Y)_{ja\sigma} + (X-Y)_{ia\sigma}(X-Y)_{ja\sigma}], \tag{17}$$

$$T^{\Delta}_{ia\sigma} = T^{\Delta}_{ai\sigma} = 0, \tag{18}$$

is defined with the **X** and **Y** from the QM/MMpol/C style TDDFT calculation as shown by Eq. (11). The **Z** matrix, when written as a vector (Lagrange multiplier), can be determined using the Z-vector equation[63]

$$\sum_{jb\sigma'}(A+B)_{ia\sigma jb\sigma'}Z_{jb\sigma'} = -R_{ia\sigma}. \tag{19}$$

Here, **A** and **B** matrices are the orbital rotation Hessian, which includes induced dipoles, as shown by Eqs. (12) and (13); the Lagrangian **R**,

$$\begin{aligned} R_{ia\sigma} = &\sum_{b}[(X+Y)_{ib\sigma}H^{+}_{ab\sigma}[X+Y] + (X-Y)_{ib\sigma}H^{-}_{ab\sigma}[X-Y]] \\ &-\sum_{j}[(X+Y)_{ja\sigma}H^{+}_{ji\sigma}[X+Y] + (X-Y)_{ja\sigma}H^{-}_{ji\sigma}[X-Y]] \\ &+H^{+}_{ia\sigma}[T^{\Delta}] + 2\sum_{jb\sigma'kc\sigma''}g^{XC}_{ia\sigma jb\sigma' kc\sigma''}(X+Y)_{jb\sigma'}(X+Y)_{kc\sigma''}, \end{aligned} \tag{20}$$

is also defined with the **X** and **Y** from the QM/MMpol/C style TDDFT calculation as shown by Eq. (11), and with $g^{XC}_{ia\sigma jb\sigma' kc\sigma''}$ being a matrix element of the third order derivative of the density functional, which is the same as in regular TDDFT cases. The transformed vectors $\mathbf{H}^{+}$ and $\mathbf{H}^{-}$ originally defined by Furche and Ahlrichs[61,62] for arbitrary vector **V** can be written as

$$H^{+}_{pq\sigma}[V] = \sum_{rs\sigma'}\left\{\begin{array}{l} 2(pq\sigma|rs\sigma') + 2f^{XC}_{pq\sigma rs\sigma'} - (\mathbf{U}_{pq})^{T}(\tilde{\mathbf{w}}_{rs} + \mathbf{w}_{rs}) \\ -c_x\delta_{\sigma\sigma'}[(ps\sigma|rq\sigma) + (pr\sigma|sq\sigma)] \end{array}\right\}V_{rs\sigma}, \tag{21}$$

$$H^{-}_{pq\sigma}[V] = \sum_{rs\sigma'}\{c_x\delta_{\sigma\sigma'}[(ps\sigma|rq\sigma) - (pr\sigma|sq\sigma)]\}V_{rs\sigma}. \tag{22}$$

Here, r and s run over all occupied and virtual molecular orbitals, while p and q run over select occupied and virtual molecular orbitals. Compared to regular TDDFT cases, $\mathbf{H}^{+}[\mathbf{V}]$ now contains induced dipoles and induced charges.

Once **Z** and $\mathbf{P}^{\Delta}$ are available, $\mathbf{W}^{\Delta}$ and $\mathbf{\Gamma}^{\Delta}$ can be evaluated using Furche and Ahlrichs' formulas,[61,62]

$$W^{\Delta}_{ij\sigma} = \left(1-\frac{1}{2}\delta_{ij}\right)\left\{\begin{array}{l} +\sum_{a}\omega[(X+Y)_{ia\sigma}(X-Y)_{ja\sigma} + (X-Y)_{ia\sigma}(X+Y)_{ja\sigma}] \\ -\sum_{a}\varepsilon_{a\sigma}[(X+Y)_{ia\sigma}(X+Y)_{ja\sigma} + (X-Y)_{ia\sigma}(X-Y)_{ja\sigma}] \\ +H^{+}_{ij\sigma}[P^{\Delta}] \\ +2\sum_{kc\sigma' ld\sigma''}g^{xc}_{ij\sigma kc\sigma' ld\sigma''}(X+Y)_{kc\sigma'}(X+Y)_{ld\sigma''} \end{array}\right\}, \tag{23}$$

$$W^{\Delta}_{ab\sigma} = \left(1-\frac{1}{2}\delta_{ab}\right)\left\{\begin{array}{l} +\sum_{i}\omega[(X+Y)_{ia\sigma}(X-Y)_{ib\sigma} + (X-Y)_{ia\sigma}(X+Y)_{ib\sigma}] \\ +\sum_{i}\varepsilon_{i\sigma}[(X+Y)_{ia\sigma}(X+Y)_{ib\sigma} + (X-Y)_{ia\sigma}(X-Y)_{ib\sigma}] \end{array}\right\}, \tag{24}$$

$$W^{\Delta}_{ia\sigma} = \sum_{j}\{(X+Y)_{ja\sigma}H^{+}_{ji\sigma}[X+Y] + (X-Y)_{ia\sigma}H^{-}_{ji\sigma}[X-Y]\} + \varepsilon_{i\sigma}Z_{ia\sigma}, \tag{25}$$

$$\begin{aligned} \Gamma^{\Delta}_{\mu\nu\sigma\kappa\lambda\sigma'} = &\frac{1}{2}\left[2P^{\Delta}_{\mu\nu\sigma}P_{\kappa\lambda\sigma'} + 2(X+Y)_{\mu\nu\sigma}(X+Y)_{\kappa\lambda\sigma'}\right] \\ &-c_x\delta_{\sigma\sigma'}\left[\begin{array}{l} +P^{\Delta}_{\mu\lambda\sigma}P_{\kappa\nu\sigma'} + P^{\Delta}_{\mu\kappa\sigma}P_{\lambda\nu\sigma'} \\ +(X+Y)_{\mu\lambda\sigma}(X+Y)_{\kappa\nu\sigma'} + (X+Y)_{\mu\kappa\sigma}(X+Y)_{\lambda\nu\sigma'} \\ -(X-Y)_{\mu\lambda\sigma}(X+Y)_{\kappa\nu\sigma'} + (X-Y)_{\mu\kappa\sigma}(X-Y)_{\lambda\nu\sigma'} \end{array}\right], \end{aligned} \tag{26}$$

with the **X** and **Y** from the QM/MMpol/C style TDDFT calculation as shown by Eq. (11). We note again that the transformed vector $\mathbf{H}^{+}$ in Eqs. (23) and (25) contains induced dipoles and induced charges [see Eq. (21)].

The sixth, seventh, and eighth terms in Eq. (14) represent the forces between the TDDFT difference density and MM repulsion/dispersion/multipole points. As shown in Eq. (4), repulsion, dispersion, and multipole points are used in the EFP method. For typical force field methods, the repulsion and dispersion are described by LJ potentials so they are simply zero in Eq. (14). These three forces are analogs of the force between the TDDFT difference density and QM nuclei, and can be evaluated using similar techniques.

Using Eq. (8), the ninth and tenth terms in Eq. (14) can be combined and simplified

$$
\begin{aligned}
&-\frac{1}{2}\sum_{\mu\nu\sigma\lambda\kappa\sigma'}\left(P^{\Delta}_{\mu\nu\sigma}P_{\lambda\kappa\sigma'}+P_{\mu\nu\sigma}P^{\Delta}_{\lambda\kappa\sigma'}\right)\\
&\times\left[\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_{\lambda\kappa}+\mathbf{w}_{\lambda\kappa})\right]^x\\
&-\frac{1}{2}\sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}\left[\frac{1}{2}(\mathbf{U}_N+\mathbf{U}_{mul})^T(\tilde{\mathbf{w}}_{\mu\nu}+\mathbf{w}_{\mu\nu})\right.\\
&\left.+\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_N+\tilde{\mathbf{w}}_{mul}+\mathbf{w}_N+\mathbf{w}_{mul})\right]^x\\
&=-\frac{1}{2}\left[(\mathbf{U}^{\Delta,x})^T\mathbf{w}_e+(\tilde{\mathbf{w}}^{\Delta})^T\mathbf{B}^x\mathbf{w}_e+(\tilde{\mathbf{w}}^{\Delta})^T\mathbf{U}^x_e\right]\\
&-\frac{1}{2}\left[\left(\mathbf{U}^x_e\right)^T\mathbf{w}^{\Delta}+(\tilde{\mathbf{w}}_e)^T\mathbf{B}^x\mathbf{w}^{\Delta}+(\tilde{\mathbf{w}}_e)^T\mathbf{U}^{\Delta,x}\right]\\
&-\frac{1}{2}\left[\left(\mathbf{U}^x_N+\mathbf{U}^x_{mul}\right)^T\mathbf{w}^{\Delta}+(\tilde{\mathbf{w}}_N+\tilde{\mathbf{w}}_{mul})^T\mathbf{B}^x\mathbf{w}^{\Delta}\right.\\
&\left.+(\tilde{\mathbf{w}}_N+\tilde{\mathbf{w}}_{mul})^T\mathbf{U}^{\Delta,x}\right]\\
&-\frac{1}{2}\left[(\mathbf{U}^{\Delta,x})^T(\mathbf{w}_N+\mathbf{w}_{mul})+(\tilde{\mathbf{w}}^{\Delta})^T\mathbf{B}^x(\mathbf{w}_N+\mathbf{w}_{mul})\right.\\
&\left.+(\tilde{\mathbf{w}}^{\Delta})^T\left(\mathbf{U}^x_N+\mathbf{U}^x_{mul}\right)\right]\\
&=-\frac{1}{2}[(\mathbf{U}^x)^T(\mathbf{w}^{\Delta}+\tilde{\mathbf{w}}^{\Delta})+(\mathbf{U}^{\Delta,x})^T(\mathbf{w}+\tilde{\mathbf{w}})\\
&+(\tilde{\mathbf{w}})^T\mathbf{B}^x\mathbf{w}^{\Delta}+(\tilde{\mathbf{w}}^{\Delta})^T\mathbf{B}^x\mathbf{w}].
\end{aligned}
\tag{27}
$$

Here, $\mathbf{U}^{\Delta,x}$ represents the derivatives of the electrostatic field and potential due to the TDDFT excitation difference density, and $\mathbf{w}^{\Delta}$ and $\tilde{\mathbf{w}}^{\Delta}$ represent induced dipoles and induced charges due to the TDDFT excitation difference density,

$$
\mathbf{U}^{\Delta,x}=\sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}\mathbf{U}^x_{\mu\nu},
\tag{28}
$$

$$
\mathbf{w}^{\Delta}=-\sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}\mathbf{B}^{-1}\mathbf{U}_{\mu\nu},
\tag{29}
$$

$$
\tilde{\mathbf{w}}^{\Delta}=-\sum_{\mu\nu\sigma}P^{\Delta}_{\mu\nu\sigma}(\mathbf{B}^{-1})^T\mathbf{U}_{\mu\nu}.
\tag{30}
$$

The first term in the final expression of Eq. (27) represents the electrostatic force and torque imposed on $\mathbf{w}^{\Delta}$ (and $\tilde{\mathbf{w}}^{\Delta}$) by QM nuclei, electrons, and MM multipoles; the second term represents the electrostatic force and torque imposed on $\mathbf{w}^{\Delta}$ (and $\tilde{\mathbf{w}}^{\Delta}$) by the TDDFT difference electron density; the third and fourth terms represent the electrostatic force and torque among $\mathbf{w}^{\Delta}$ (and $\tilde{\mathbf{w}}^{\Delta}$). These forces and torques can be evaluated efficiently using standard electrostatic formulas as if the induced dipoles and induced charges were permanent dipoles and charges.[9,10,59]

The last term in Eq. (14) can be written as

$$
\begin{aligned}
&\sum_{\mu\nu\sigma\lambda\kappa\sigma'}(X+Y)_{\mu\nu\sigma}(X+Y)_{\lambda\kappa\sigma'}\left[\frac{1}{2}(\mathbf{U}_{\mu\nu})^T(\tilde{\mathbf{w}}_{\lambda\kappa}+\mathbf{w}_{\lambda\kappa})\right]^x\\
&=(\mathbf{U}^x_{XY})^T(\mathbf{w}_{XY}+\tilde{\mathbf{w}}_{XY})+(\tilde{\mathbf{w}}_{XY})^T\mathbf{B}^x\mathbf{w}_{XY}.
\end{aligned}
\tag{31}
$$

Here, $\mathbf{U}^x_{XY}$ represents the derivatives of electrostatic field and potential due to the TDDFT transition state density $(X+Y)_{\mu\nu\sigma}$, and $\mathbf{w}_{XY}$ represents the induced dipoles and induced charges due to the TDDFT transition state density $(X+Y)_{\mu\nu\sigma}$,

$$
\mathbf{U}^x_{XY}=\sum_{\mu\nu\sigma}(X+Y)_{\mu\nu\sigma}\mathbf{U}^x_{\mu\nu},
\tag{32}
$$

$$
\mathbf{w}_{XY}=-\sum_{\mu\nu\sigma}(X+Y)_{\mu\nu\sigma}\mathbf{B}^{-1}\mathbf{U}_{\mu\nu},
\tag{33}
$$

$$
\tilde{\mathbf{w}}_{XY}=-\sum_{\mu\nu\sigma}(X+Y)_{\mu\nu\sigma}(\mathbf{B}^{-1})^T\mathbf{U}_{\mu\nu}.
\tag{34}
$$

The first term on the right side of Eq. (31) represents the force and torque imposed on the transition state density induced dipoles $\mathbf{w}_{XY}$ (and $\tilde{\mathbf{w}}_{XY}$) by the electrostatic field gradient due to the transition state density $(X+Y)_{\mu\nu\sigma}$; the second term represents the force and torque between $\mathbf{w}_{XY}$ and $\tilde{\mathbf{w}}_{XY}$. Again, these forces and torques can be evaluated efficiently using standard electrostatic formulas as if the transition-state-density-induced dipoles/charges were permanent dipoles/charges.[9,10,59]

III. IMPLEMENTATION

We implemented the QM/MMpol/C style TDDFT methods in the General Atomic and Molecular Electronic Structure System (GAMESS)[64,65] based on the TDDFT and long-range correction TDDFT (denoted as TD-LC-DFT in this paper) program implemented by Chiba *et al.*[66,67] This TDDFT program can also perform configuration interaction singles (CIS)[68] calculations. Two different versions were implemented. One is for the conductorlike polarizable continuum model (CPCM,[69,70] a variant of COSMO[71]) and the EFP[9,10,58,59] method (code already released for public use in Aug 2011), the other is for the FixSol[72] solvation model (similar to COSMO[71] and CPCM[69,70] but with modified short-range surface charge interactions) and general induced dipole polarizable force field methods in the quantum chemistry polarizable force field (QuanPol)[73] program, which is designed for general MM and QM/MM calculations. In QuanPol calculations, the QM methods can be HF, generalized valence bond theory (GVB),[74] multiconfiguration self-consistent-field (MCSCF,[75–78] including state-average MCSCF, or SA-MCSCF[79–81]), DFT,[82] TDDFT,[16,17] and second order Møller-Plesset perturbation theory (MP2)[83] methods. The QuanPol program is integrated in and distributed with

the GAMESS package. A new version of QuanPol will be released soon. Using GAMESS default settings, the accuracy of the analytic QM/MMpol/C style TDDFT gradient is typically 10^{-6} hartree/bohr. In the GAMESS source code file *grd2a.src*, the analytic gradient is projected out when DFT or TDDFT is used. The projection will alter the genuine Cartesian analytic gradient but can greatly improve geometry optimization convergence. To reproduce the genuine gradients, the projection in *grd2a.src* should be turned off.

IV. COMPUTATIONAL METHODS

QM/MMpol/C style calculations were performed for PYP. The X-ray structure file 2PHY[39] was obtained from the Protein Data Bank (PDB).[84] All water molecules in the PDB file 2PHY were deleted. Hydrogen atoms were added to the X-ray structure using the WHAT IF web interface.[85] The total number of atoms is 1929, including the chromophore *p*CA linked to Cys69. The total charge of the protein and chromophore is −6 *e*. All amino acids (1908 atoms without the chromophore) were described with the AMBER induced dipole polarizable force field (topology and parameter files "all_amino02.in," "all_aminoct02.in," "all_aminont02.in," and "parm99.dat").[86–88] The force field parameters of the chromophore were obtained from the general AMBER force field (GAFF)[89] by using the AmberTools12.[90] The GAFF atomic charges were obtained by using the RED-III.52-Tools[91] and GAMESS.[64,65] The atomic charges and LJ parameters are listed in Table SI in the supplementary material;[92] the molecule used to generate the GAFF parameters for the chromophore is shown in Figure S1 in the supplementary material.[92] The covalent terms were not used in the QM/MM calculation because the chromophore was described by QM methods. So the covalent parameters generated by GAFF for the chromophore are not discussed in this paper. Charge-charge and LJ interactions between atoms separated by three covalent bonds were scaled by 1/1.2 and 0.50, respectively, as in standard AMBER force field. The interactions between induced dipoles and the interactions between charges and induced dipoles are excluded for 1–2 and 1–3 atom pairs, and are included with no scaling for 1–4 atom pairs and beyond. There is no guarantee that the AMBER dipole polarizabilities are optimum for QM/MM calculations. However, using another set of polarizabilities may destroy the integrity of the AMBER protein force field.

In the QM/MMpol/C calculation, the QM region had 46 atoms: the side chain of Tyr42, the partial side chain of Glu46, and the chromophore (Figure 1). Two QM capping H atoms for Tyr42 and Cys69 were placed at the positions of their alpha carbon atoms, while the QM capping H atom for Glu46 was placed at the position of its beta carbon atom. In all QM calculations, the 6-31++G(d,p)[93] basis set was used. We used the standard method implemented in QuanPol to treat the covalent bonds between QM and MM atoms.[73] The details of the method are given in the supplementary material.[92] In the QM/MM calculation, the force field atomic charges of the QM atoms were not used because they interact with the MM atoms using their electrons and nuclear charges; the QM atoms use their LJ potentials (LJ parameters from GAFF) to interact with the MM atoms (LJ parameters from AMBER). We emphasize here that the AMBER and GAFF charge-charge and LJ interactions are parameterized together for pure MM methods. There is no guarantee that the LJ parameters work similarly well for QM/MM methods because the QM atoms are represented by nuclear charges and electrons instead of atomic point charges. In general, we found that the AMBER and GAFF LJ potential can produce reasonable hydrogen bond distances and intermolecular distances in QM/MM calculations.

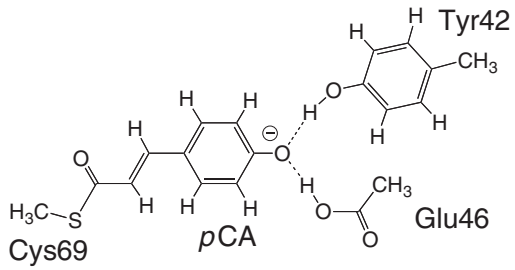

FIG. 1. The 46 QM atoms in QM/AMBER(pol)/FixSol calculations of the ground and excited states of photoactive yellow protein (PYP).

The aqueous solvent was described with the FixSol[72] model and 60 initial tesserae per sphere. Periodic boundary condition and shifting function were not used. In the FixSol calculations, atomic radii of 0.00, 2.10, 2.00, 1.90, and 2.40 Å were used for H, C, N, O, and S atoms, respectively, to define the molecular cavity (using zero for H means that they do not contribute to form the surface). For all of the 1929 QM and MM atoms in the calculation, FixSol used ∼16 500 surface tesserae and ∼8200 Å^2 total solvent accessible surface area. A dielectric constant of 78.39 was used to represent aqueous solution at 298.15 K. The induced surface charges were determined by an iterative procedure with no charge renormalization.[6,94] It took ∼56 iterations to solve ∼16 500 induced surface charges to an accuracy of averagely 10^{-10} *e* per tessera. The FixSol total induced surface charge was 5.923 ± 0.001 *e*, in excellent agreement with the theoretical value 5.9235 *e* (note the total charge of the protein is −6 *e*). In both the ground and excited states, the FixSol electrostatic solvation free energy was around −1800 kcal/mol.

The QM/MMpol/C style TDDFT calculations require more computing time than regular gas phase TDDFT calculations. Using 24 computer cores (3 nodes, each with 2 Intel Xeon E5430 CPUs and 8 cores) running at 2.66 GHz, it took 2.1 h (1.2 h for energy and 0.9 h for gradient) to finish one geometry optimization step with the TD-LC-BLYP/AMBER(pol)/FixSol method (46 atoms in the QM region, total 1929 atoms). The master core CPU efficiency is ∼93%. Using the same computer, it took 1.3 h (0.6 h for energy and 0.7 h for gradient) to finish one geometry optimization step with the gas phase TD-LC-BLYP method for the 46-atom QM region. The master core CPU efficiency is 72%.

V. RESULTS AND DISCUSSION

A. Ground state hydrogen bonding

We performed B3LYP[95]/AMBER(pol)/FixSol geometry optimization from the PDB structure in that Tyr42 and Glu46

were neutral, *p*CA was anionic. All of the 1929 QM and MM atoms were optimized. For typical small molecules, we usually converge the maximum and root mean square gradients to 1.0×10^{-4} and 3.3×10^{-5} hartree/bohr, respectively. It is difficult to optimize the geometry of a protein with a QM/MM/C method to such a tight convergence criterion. We submitted 14 consecutive jobs and performed a total of 500 steps of optimization. We used a standard geometry optimization method in GAMESS, which is usually good for small molecules, but not good for large molecules like proteins. If a better optimization method is available, the optimization can be faster and easier to converge. After 500 steps, the overall geometry of the protein is very similar to the PDB geometry. The maximum gradient and root mean square gradient were lowered to 5.9×10^{-4} and 7.0×10^{-5} hartree/bohr, respectively. The energy was still slowly lowering at a rate of ∼0.015 kcal/mol per step in the last 40 steps of the 500. This is due to the slow conformational changes of protein surface residues (in the MM region), a common behavior when a continuum solvation model is used for proteins. However, we found that after roughly 30 steps, the molecular geometry of the 46-atom QM region was already very close to that optimized for 500 steps. This is not surprising because the QM region is buried in a hydrophobic core with a relatively low flexibility. Therefore, a few tens of steps of optimization should already be sufficient for analyzing the QM region geometry and interaction.

After 500 steps of optimization, Tyr42 remained neutral, with the H_{Tyr}–O_{Tyr} distance optimized to 1.02 Å, the O_{pCA}–H_{Tyr} distance optimized to 1.53 Å, and the O_{pCA}–O_{Tyr} distance optimized to 2.54 Å. Glu46 remained neutral, with the H_{Glu}–O_{Glu} distance optimized to 1.02 Å, the O_{pCA}–H_{Glu} distance optimized to 1.60 Å, and the O_{pCA}–O_{Glu} distance optimized to 2.62 Å (Table I). Due to the negative charge of the phenolate oxygen, the H_{Tyr}–O_{Tyr} and H_{Glu}–O_{Glu} distances are slightly elongated. The O_{pCA}–O_{Tyr} distance is shorter than the O_{pCA}–O_{Glu} distance by 0.08 Å. One may assume that the *p*CA–Tyr42 hydrogen bonding is stronger than the *p*CA–Glu46 one. However, as shown later in Subsection V C, despite its shorter distance, the hydrogen bonding between *p*CA and Tyr42 is actually weaker than that between *p*CA and Glu46 by ∼5 kcal/mol.

TABLE I. Comparison of hydrogen bond distances (Å) in the S_0 ground and S_1 excited states of the photoactive yellow protein (PYP).

	X-ray[a] S_0	Neutron[b] S_0	QM/MM[c] S_0	B3LYP[d] S_0	TD-LC-BLYP[d] S_1
O_{pCA}–O_{Tyr}	2.71	2.52	2.50	2.54	2.55
H_{Tyr}–O_{Tyr}	NA	0.96	1.01	1.02	1.02
O_{pCA}–H_{Tyr}	NA	1.65	1.50	1.53	1.54
O_{pCA}–O_{Glu}	2.69	2.57	2.57	2.62	2.60
H_{Glu}–O_{Glu}	NA	1.21	1.00	1.02	1.02
O_{pCA}–H_{Glu}	NA	1.37	1.58	1.60	1.59

[a]X-ray structure 2PHY, Ref. 39.
[b]Yamaguchi *et al.*, Ref. 96.
[c]Saito and Ishikita, Ref. 97.
[d]QM/AMBER(pol)/FixSol geometry optimization from the X-ray structure 2PHY.

Based on neutron diffraction study of S_0 ground state PYP, Yamaguchi *et al.*[96] assigned a H_{Glu}–O_{Glu} distance of 1.21 Å and a O_{pCA}–H_{Glu} distance of 1.37 Å for the *p*CA–Glu46 hydrogen bond, and a H_{Tyr}–O_{Tyr} distance of 0.96 Å and a O_{pCA}–H_{Tyr} distance of 1.65 Å for the *p*CA–Tyr42 hydrogen bond (Table I). They proposed that the ground state chromophore phenolate forms a low-barrier hydrogen bond (LBHB) with Glu46 and a short ionic hydrogen bond (SIHB) with Tyr42. Saito and Ishikita[97] performed QM/MM geometry optimization of PYP, and found a H_{Glu}–O_{Glu} distance of 1.00 Å and a O_{pCA}–H_{Glu} distance of 1.58 Å for the *p*CA–Glu46 hydrogen bond, and a H_{Tyr}–O_{Tyr} distance of 1.01 Å and a O_{pCA}–H_{Tyr} distance of 1.50 Å for the *p*CA–Tyr42 hydrogen bond (Table I). Saito and Ishikita[97] therefore suggested that the *p*CA–Glu46 hydrogen bond in the ground state should be better classified as a short hydrogen bond (SHB) rather than a LBHB. The QM/MM/C optimized distances in our present study are very similar to Saito and Ishikita's[97] results (Table I).

B. Vertical excitation: Charge transfer from Tyr42 to *p*CA

Using the gas phase B3LYP optimized ground state geometry of the 22-atom anionic chromophore in Figure 1, we performed gas phase single point energy calculations with TD-B3LYP, TD-LC-BLYP,[66,67,98] and CIS methods. The excitation is dominated by the transition from HOMO (the 51st orbital) to HOMO + *n*. We visualized the two dominant molecular orbitals in each case using the GAMESS graphic software MacMolPlt,[99] and found that there is a noticeable electron density translocation from the phenolate ring to the ethylene chain (Figure 2). The excitation energies calculated with TD-B3LYP, TD-LC-BLYP, and CIS are 3.214, 3.235, and 3.907 eV, respectively (Table II). According to the experimental measurement performed by Nielsen *et al.*,[100] the anionic chromophore should have an absorption maximum of 2.70 eV in the gas phase. Molina and Merchan[23] performed multistate-CASPT2 calculation and obtained 2.58 eV. Gromov *et al.*[27] performed CC2 and EOM-CCSD calculations and obtained 2.89 eV and 3.18 eV, respectively. Ma *et al.*[101] performed many-body Green's function theory (MBGFT) calculation and obtained 2.80 eV. Compared to more accurate methods, TD-B3LYP and TD-LC-BLYP tend to overestimate the excitation energy of *p*CA by ∼0.5 eV.

Based on the B3LYP/AMBER(pol)/FixSol optimized PYP geometry, we performed vertical excitation energy calculations using TD-B3LYP/AMBER(pol)/FixSol, TD-LC-BLYP/AMBER(pol)/FixSol, and CIS/AMBER(pol)/FixSol methods. The results are listed in Table II and the orbitals are displayed in Figure 3. The main findings are that in PYP there is a similar charge translocation from the *p*CA phenolate ring to the ethylene chain (Figures 2 and 3), and that there is some charge transfer from Tyr42 to *p*CA (Figure 3).

TD-B3LYP calculation gives an excitation energy of 2.868 eV with an oscillator strength of 1.306. The excited state is dominated by the excitation (amplitude 0.984) from the highest occupied molecular orbital (HOMO, the 96th

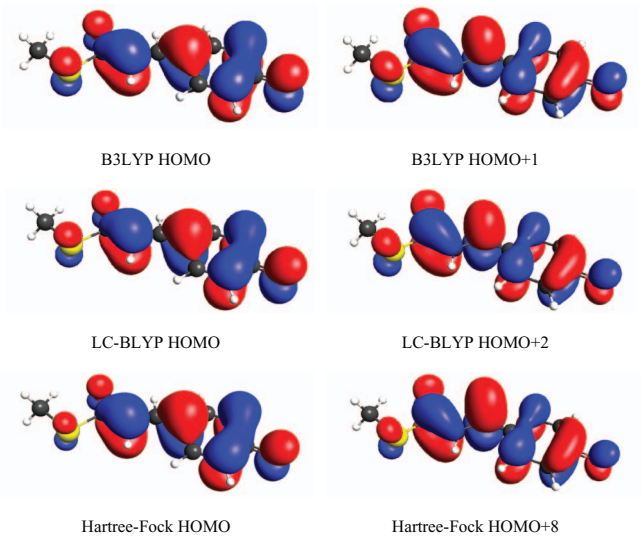

FIG. 2. The two main molecular orbitals involved in the photoexcitation of gas phase anionic chromophore (22-atom model) calculated using different methods. There is some electron density translocation from the *p*CA phenolate ring to the ethylene chain in the excited state.

TABLE II. Computed excitation energies (ω) and oscillator strength (f) for photoactive yellow protein (PYP) chromophore in gas phase and protein.

Molecule	Method	ω (eV)	f	Charge transfer	Protein effect (eV)
Gas phase[a]	TD-B3LYP	3.214	1.057	NA	NA
	TD-LC-BLYP	3.235	1.140	NA	NA
	CIS	3.907	1.493	NA	NA
Protein[b]	TD-B3LYP	2.868	1.306	Significant	−0.346
	TD-LC-BLYP	3.203	1.413	Little	−0.032
	CIS	3.860	1.455	Little	−0.047

[a]22-atom model, B3LYP optimized.
[b]B3LYP/AMBER(pol)/FixSol optimized.

orbital) to HOMO+1 (the 97th orbital). The amplitude of the excitation from the HOMO-1 (the 95th orbital) to HOMO+1 (the 97th orbital) is 0.136, also significant. The HOMO-1 is mainly localized in Tyr42 but significantly delocalized to *p*CA, the HOMO is mainly localized in *p*CA but significantly delocalized to Tyr42, and the HOMO+1 is localized in *p*CA (Figure 3). Because the electron in the HOMO is significantly delocalized into Tyr42, and is almost completely localized in *p*CA after excitation, it causes a significant (exact value is unknown) charge transfer from Tyr42 to *p*CA. In addition, the excitation from the HOMO-1 (localized in Tyr42) to HOMO+1 (localized in *p*CA) also contributes roughly 0.018 electron (estimated as 0.136 × 0.136) to the charge transfer from Tyr42 to *p*CA. It is well known that TDDFT (including TD-B3LYP) methods tend to overestimate intermolecular long-range charge transfer.[102–104] Therefore, we suspect that the charge transfer observed in the TD-B3LYP/AMBER(pol)/FixSol calculation is erroneously overestimated.

TD-LC-BLYP/AMBER(pol)/FixSol gives 3.203 eV with an oscillator strength of 1.413. The excited state is dominated by the excitation (amplitude −0.922) from HOMO (the 96th orbital) to HOMO+1 (the 97th orbital). The amplitude of the excitation from HOMO-1 (the 95th orbital) to HOMO+1 (the 97th orbital) is 0.247, also significant. The HOMO-1 is localized in Tyr42, the HOMO is localized in *p*CA, and the HOMO+1 is localized in *p*CA (Figure 3). This means that the excitation of PYP involves little (roughly 0.06 electron as estimated from 0.247 × 0.247) charge transfer from Tyr42 to *p*CA. Compared to the TD-B3LYP/AMBER(pol)/FixSol case discussed in the above paragraph, the long-range charge transfer in TD-LC-BLYP/AMBER(pol)/FixSol calculation is largely suppressed and should not be problematic.

CIS/AMBER(pol)/FixSol gives 3.860 eV with an oscillator strength of 1.455. The excited state is dominated by the

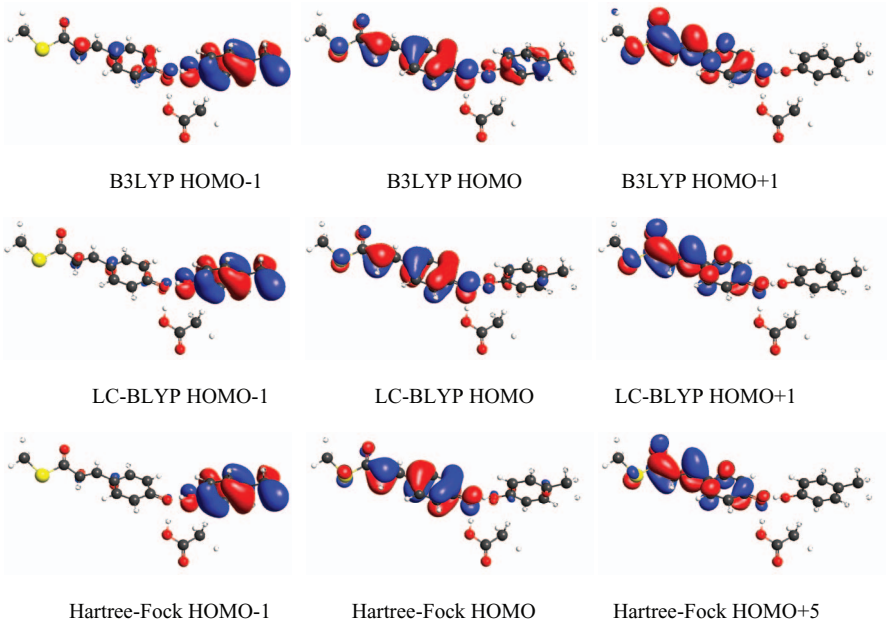

FIG. 3. The main molecular orbitals involved in the photoexcitation of anionic chromophore (46-atom QM model) calculated using different QM/AMBER(pol)/FixSol methods.

excitation from HOMO (the 96th orbital) to HOMO+5 (the 101st orbital, amplitude −0.845). The amplitude of the excitation from HOMO-1 (the 95th orbital) to HOMO+5 (the 101st orbital) is 0.190. We visualized the HOMO-1, HOMO, and HOMO+5 orbitals in the reference Hartree-Fock calculation. The HOMO-1 is localized in Tyr42, the HOMO is localized in *p*CA, and HOMO+5 is localized in *p*CA (Figure 3). This means that the excitation of PYP involves little (roughly 0.04 electron as estimated from 0.190 × 0.190) charge transfer from Tyr42 to *p*CA. Although CIS method does not have the electron correlation interaction TDDFT methods have, it is free of erroneous long-range charge transfer. Therefore, we estimate that the charge transfer from Tyr42 to *p*CA is about ∼0.04 electron.

TD-B3LYP/AMBER(pol)/FixSol gives an excitation energy in protein that is lower by 0.346 eV than that in the gas phase (2.868 versus 3.214 eV, Table II). This result is problematic because of the erroneous charge transfer in the TD-B3LYP method. TD-LC-BLYP and CIS methods are more reliable than TD-B3LYP method because they give better description of intermolecular long-range charge transfer. TD-LC-BLYP/AMBER(pol)/FixSol and CIS/AMBER(pol)/FixSol give excitation energies in protein similar (within −0.03 to −0.04 eV, Table II) to those calculated for the gas phase, suggesting that the protein and solvent interactions do not significantly change the vertical excitation energy. This result is in accordance with the experimental value 2.70 eV for the gas phase chromophore and 2.78 eV for the chromophore in PYP. Previously, González *et al.*[33] obtained 3.32 eV from a gas phase B3LYP calculation for the chromphore, and 3.21 eV from a QM/MM style TD-B3LYP calculation for the chromphore in PYP. Chiba *et al.*[105,106] obtained a similarly small change using the fragment molecular orbital (FMO) based TDDFT and CPCM/TDDFT methods.

Recently, Rocha-Rinza *et al.*[34] found that the red shift due to the chromophore geometric distortion in the protein and the blue shift due to protein interaction are both in the magnitude of roughly 0.1 eV, and largely cancel each other. Following their procedure, here we present a similar analysis. Using the same geometry optimized for 500 steps with the B3LYP/AMBER(pol)/FixSol method, we performed gas phase TD-LC-BLYP single point energy calculations for the 46-atom QM region (no MM, no FixSol, only *p*CA, Tyr42, and Glu46) and found that the excitation energy is 3.436 eV. Clearly, the protein interactions (other than Tyr42 and Glu46) can cause a red shift in the excitation energy by 0.233 eV. We also performed similar calculations for the *p*CA–Tyr42 and *p*CA–Glu46 pairs and *p*CA itself, and obtained excitation energies of 3.266, 3.289, and 3.128 eV, respectively. These results show that the geometric distortion of the *p*CA can cause a red shift of 0.107 eV (from 3.235 to 3.128 eV), the hydrogen bonding of Tyr42 can cause a blue shift of 0.138 eV (from 3.128 to 3.266 eV), the hydrogen bonding of Glu46 can cause a blue shift of 0.161 eV (from 3.128 to 3.289 eV). It seems that the hydrogen bonding effects from Tyr42 and Glu46 are roughly additive: 0.138 + 0.161 = 0.299 eV, and the total blue shift due to Tyr42 and Glu46 is 0.308 eV (from 3.128 to 3.436 eV). The blue shift due to the hydrogen bonding of Tyr42 and Glu46 can be explained by the electron density translocation that weakens the hydrogen bonding upon excitation. In other words, the excitation of *p*CA requires more energy when the negative charge on the phenolate oxygen is stabilized by hydrogen bonds.

C. Excited state hydrogen bonding

We performed TD-B3LYP/AMBER(pol)/FixSol geometry optimization starting from the B3LYP/AMBER(pol)/FixSol optimized structure, in which Tyr42 and Glu46 were neutral, *p*CA was anionic. After 40 steps of optimization, the proton transferred from Tyr42 to *p*CA, and Glu46 remained neutral. We suspect that the proton transfer is due to the Tyr42-to-*p*CA charge transfer: the proton tends to follow the electron. Because the charge transfer is erroneously overestimated in the TD-B3LYP calculation, the proton transfer predicted by the TD-B3LYP method is most likely a false prediction.

We performed similar TD-LC-BLYP/AMBER(pol)/FixSol excited state geometry optimizations for 40 steps and found that the Tyr42 proton did not transfer to *p*CA. As discussed in Subsection V A, the QM region optimizes much faster than the whole protein, and a few tens of steps are sufficient. Here, after 40 steps of optimization, the geometry of the QM region was very close to being fully optimized, and the hydrogen bonding distances were changing at a level like 0.0005 Å. The maximum gradient and root mean square gradient of the 46 QM atoms were lowered to 4.4×10^{-4} and 1.0×10^{-4} hartree/bohr, respectively, while the maximum gradient and root mean square gradient of all 1929 QM/MM atoms were 2.8×10^{-3} and 1.0×10^{-4} hartree/bohr. Clearly, further optimization would not lead to any significant differences in the QM geometry, thus not worthy of trying due to the high computational cost.

The excited state hydrogen bond distances from the TD-LC-BLYP/AMBER(pol)/FixSol calculation are very similar to the ground state ones from the B3LYP//AMBER(pol)/FixSol calculation (Table I). Although there is no significant change in the *p*CA hydrogen bond distances after excitation, the intermolecular interaction (mainly hydrogen bonding) strengths for the *p*CA–Tyr42 and *p*CA–Glu46 pairs show a general weakening (Table III). Using the B3LYP/AMBER(pol)/FixSol optimized S_0 state geometry of the 46-atom QM region, we performed gas phase LC-BLYP ground state and/or TD-LC-BLYP excited state single point energy calculations for the Tyr42 molecule, the Glu46 molecule, the *p*CA molecule, the *p*CA–Tyr42 pair, and the *p*CA–Glu46 pair (Figure 1, pure QM with no MM and no FixSol). The intermolecular interaction energies were obtained by taking the energy differences (e.g., $E_{int} = E_{pCA\text{-}Tyr42} - E_{pCA} - E_{Tyr42}$) with no basis set superposition correction, which is usually small for DFT methods with a moderate polarizable basis set. These interaction energies are two-body values for the two involved groups. When other groups are involved, the total interaction energy tends to be smaller in magnitude than the sum of the two-body values, due to many-body effects. Nevertheless, the two-body interaction energies

TABLE III. Intermolecular interaction energy (kcal/mol) of the *p*CA–Tyr42 pair and the *p*CA–Glu46 pair in S_0 and S_1 states.

*p*CA state	Method	*p*CA–Tyr42	*p*CA–Glu46
S_0	LC-BLYP//B3LYP/AMBER(pol)/FixSol	−28.22	−33.94
S_1	(LC-BLYP and TD-LC-BLYP)//B3LYP/AMBER(pol)/FixSol	−25.05	−30.22
S_1	(LC-BLYP and TD-LC-BLYP)//TD-LC-BLYP/AMBER(pol)/FixSol	−24.51	−30.32

are still very meaningful for comparing the relative interaction strength between the *p*CA–Tyr42 pair and the *p*CA–Glu46 pair. These interactions contain mainly hydrogen bonding, and are general indicators of the hydrogen bonding strengths. In general, DFT methods cannot predict accurate intermolecular interactions because they cannot describe long-range dispersion interactions, but may overestimate short-range electronic correlation energies. The absolute interaction energies obtained here using LC-BLYP and TD-LC-BLYP methods may significantly overestimate or underestimate the true interactions, and should not be accepted as very accurate results. However, their relative values should still be meaningful for comparison. The interaction energies calculated with the same exchange and correlation functionals (e.g., LC-BLYP) should also be meaningful for comparing DFT and TDDFT results.

In the S_0 state geometry, S_0 state *p*CA shows a 5.72 kcal/mol stronger interaction with S_0 state Glu46 (−33.94 kcal/mol) than with S_0 state Tyr42 (−28.22 kcal/mol). Upon vertical excitation (no geometric change), the S_1 state *p*CA shows weaker interactions with S_0 state Tyr42 (weaker by 3.17 kcal/mol) and S_0 state Glu46 (weaker by 3.72 kcal/mol) as compared to the S_0 state *p*CA. Since both are weakened, the Glu46 interaction is still 5.17 kcal/mol stronger than Tyr42. The weakening of these interactions is caused by the electron density translocation from the *p*CA phenolate ring to the ethylene chain, which results in less negative charge on the phenolate oxygen atom.

We also obtained the interaction energies using the TD-LC-BLYP/AMBER(pol)/FixSol optimized S_1 state geometry. Clearly, after relaxation on the S_1 state potential energy surface, the S_1 state *p*CA shows almost unchanged hydrogen bond strengths: *p*CA–Tyr42 is weakened by 0.54 kcal/mol and *p*CA–Glu46 is strengthened by 0.10 kcal/mol. The Glu46 interaction is still 5.81 kcal/mol stronger than the Tyr42 one (Table III).

D. Comparison to experiments

Based on the analysis of ultrafast (instrument response time is ∼200 fs) infrared spectra of PYP, van Wilderen *et al.*[107] noted that *p*CA shows an instantaneous change in infrared spectrum (mainly the phenolate-ring bands at 1160, 1443, 1485, and 1550 cm^{-1}) upon photoexcitation, and the change is similar to the difference spectrum between anionic *p*CA and neutral *p*CA. They said that this implies that the *p*CA in PYP phenolate ring is less negatively charged after excitation. Our results in the present work clearly indicate that upon excitation there is an electron density translocation from the *p*CA phenolate ring to the ethylene chain (Figure 3). This electron density translocation changes the single and double bond lengths and strengths of the conjugated *p*CA phenolate ring and ethylene chain, causing differences in infrared absorption band position and intensity.

Based on the analyses of ultraviolet resonance Raman (UVRR) spectra of PYP, Mizuno *et al.*[108] concluded that, after photoexcitation, the *p*CA–Tyr42 hydrogen bond is strengthened, while the *p*CA–Glu46 hydrogen bond is weakened. According to their description, the *p*CA–Glu46 hydrogen bond strength in the excited state PYP is similar to that of *p*CA–Gln46 in the excited state Glu46Gln–PYP. The key assumption made by Mizuno *et al.*[108] is that the strengthening of the *p*CA–Tyr42 hydrogen bond can cause blue shift of the Tyr42 absorption profile, resulting in a depletion of the Tyr42 absorption band (especially the Y8a mode of Tyr42 at 1620 cm^{-1}). However, the interaction energies (Table III) obtained in the current work suggest that the Tyr42 and Glu46 hydrogen bonds are weakened by 3.2 and 3.7 kcal/mol, respectively, immediately upon (vertical) excitation. It is easy to explain that the Tyr42 and Glu46 hydrogen bonds are both weakened because there is a loss of electron density on the phenolate oxygen atom, but difficult to explain why the Tyr42 one would be preferentially or asymmetrically strengthened. In addition, we found that the initial geometry relaxation on the S_1 potential energy surface does not appear to alter the Tyr42 and Glu46 hydrogen bond strengths (Table III). Therefore, our current result does not support the main conclusion (preferential strengthening of the *p*CA–Tyr42 hydrogen bond upon excitation) in the paper of Mizuno *et al.*[108]

Based on the analyses of femtosecond stimulated Raman spectroscopy spectra of PYP, Nakamura *et al.*[109] found that the 1555 cm^{-1} mode of the wild type PYP has a reduced intensity that is similar to the same mode of Glu46Gln-PYP, implying that the pCA–Glu46 hydrogen bond is weakened upon excitation. They observed that the weakening process is accomplished within ∼150 fs after the excitation. Other authors, for example, Mizuno *et al.*[108] also concluded that there is a weakening in the *p*CA–Glu46 hydrogen bond upon excitation. Indeed, our calculation of the interaction energies (Table III) suggest that the *p*CA–Tyr42 and *p*CA–Glu46 hydrogen bonds are weakened by ∼3.2 and ∼3.7 kcal/mol, respectively. The hydrogen bond weakening is a direct consequence of the charge translocation from *p*CA phenolate ring to the ethylene chain, and takes only a few fm to finish. In addition, we found that the initial geometry relaxation on the S_1 potential energy surface does not alter the Tyr42 and Glu46 hydrogen bond strengths, implying that the weakening of these hydrogen bonds can be observed immediately after the excitation and probably for a short period of time. This is consistent with Nakamura *et al.*[109] observation that the *p*CA

hydrogen bond rearrangements occurs within ∼150 fs upon excitation.

We note that the term "hydrogen bond rearrangement" is vague because it does not give any specific description. Based on our results, we can give a more clear description: the hydrogen bonding geometry is almost unchanged upon excitation, the only significant change is the electron density translocation from the *p*CA phenolate oxygen to the ethylene chain, which causes the immediate weakening of both the Tyr42 and Glu46 hydrogen bonds.

VI. CONCLUSION

The analytic gradient of the QM/MMpol/C style TDDFT method (including TD-LC-DFT and CIS) is derived and implemented for the EFP method and general induced dipole polarizable force field methods in the QuanPol program. Using the analytic gradient, geometry optimization and molecular dynamics simulation can be performed. Using QM to describe the photoactive site (the chromophore, Tyr42, and Glu46), AMBER induced dipole polarizable force field to describe the rest of the protein, and the FixSol solvation model to describe the aqueous solvent, geometry optimizations were performed for the S_0 ground state and the S_1 excited state of PYP. The main findings are summarized below:

(1) Ground state QM/MM(pol)/C style B3LYP geometry optimization suggests that in PYP the chromophore *p*CA is anionic, forming two normal hydrogen bonds with neutral Tyr42 and neutral Glu46. In the ground state, the *p*CA–Glu46 hydrogen bond is stronger than the *p*CA–Tyr42 one by ∼5 kcal/mol (Table III).

(2) Excited state QM/MM(pol)/C style TD-LC-BLYP and CIS calculations suggest that there is a significant electron density translocation from the *p*CA phenolate ring to the ethylene chain, and some charge transfer from Tyr42 to *p*CA in the photoexcitation (Figures 2 and 3). These calculations also show that the protein and solvent interactions have very little effects (roughly −0.03 eV, Table II) on the excitation energy of the chromophore *p*CA.

(3) Excited state QM/MM(pol)/C style TD-LC-BLYP geometry optimization of the excited state show that there are very small changes in the ground and excited state *p*CA–Tyr42 and *p*CA–Glu46 hydrogen bond distances (Table I). However, after excitation, both the *p*CA–Tyr42 and *p*CA–Glu46 hydrogen bonds show a reduction in strength by 3-4 kcal/mol due to the less negative charge on the phenolate oxygen atom (Table III). In both the ground and excited states, the *p*CA–Glu46 hydrogen bond is stronger than the *p*CA–Tyr42 one by ∼5 kcal/mol (Table III).

ACKNOWLEDGMENTS

This work is supported by the U.S. National Science Foundation (NSF) (Award No. 1010674). The authors are grateful to the reviewers of the paper for their critical comments that lead to improvement of the paper.

[1] A. H. Devries, P. T. van Duijnen, A. H. Juffer, J. A. C. Rullmann, J. P. Dijkman, H. Merenga, and B. T. Thole, J. Comput. Chem. **16**(1), 37–55 (1995).
[2] P. Bandyopadhyay and M. S. Gordon, J. Chem. Phys. **113**(3), 1104–1109 (2000).
[3] A. Ohrn and G. Karlstrom, Mol. Phys. **104**(19), 3087–3099 (2006).
[4] Q. Cui, J. Chem. Phys. **117**(10), 4720–4728 (2002).
[5] P. Bandyopadhyay, M. S. Gordon, B. Mennucci, and J. Tomasi, J. Chem. Phys. **116**(12), 5023–5032 (2002).
[6] H. Li, C. S. Pomelli, and J. H. Jensen, Theor. Chem. Acc. **109**(2), 71–84 (2003).
[7] S. Chalmet, D. Rinaldi, and M. F. Ruiz-López, Int. J. Quantum Chem. **84**(5), 559–564 (2001).
[8] A. H. Steindal, K. Ruud, L. Frediani, K. Aidas, and J. Kongsted, J. Phys. Chem. B **115**(12), 3027–3037 (2011).
[9] H. Li, J. Chem. Phys. **131**(18), 184103 (2009).
[10] H. Li and M. S. Gordon, J. Chem. Phys. **126**(12), 124112 (2007).
[11] F. J. Vesely, J. Comput. Phys. **24**(4), 361–371 (1977).
[12] M. Neumann, F. J. Vesely, O. Steinhauser, and P. Schuster, Mol. Phys. **35**(3), 841–855 (1978).
[13] F. H. Stillinger and C. W. David, J. Chem. Phys. **69**(4), 1473–1484 (1978).
[14] P. Barnes, J. L. Finney, J. D. Nicholas, and J. E. Quinn, Nature (London) **282**(5738), 459–464 (1979).
[15] A. Warshel, J. Phys. Chem. **83**(12), 1640–1652 (1979).
[16] M. E. Casida, in *Recent Advances in Density Functional Methods*, edited by D. P. Chong (World Scientific, Singapore, 1995), p. 155.
[17] M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, J. Chem. Phys. **108**(11), 4439–4449 (1998).
[18] Y. Wang and H. Li, J. Chem. Phys. **133**(3), 034108 (2010).
[19] D. Si and H. Li, J. Chem. Phys. **133**(14), 144112 (2010).
[20] T. E. Meyer, Biochim. Biophys. Acta **806**(1), 175–183 (1985).
[21] U. K. Genick, S. M. Soltis, P. Kuhn, I. L. Canestrelli, and E. D. Getzoff, Nature (London) **392**(6672), 206–209 (1998).
[22] G. Groenhof, M. Bouxin-Cademartory, B. Hess, S. P. de Visser, H. J. C. Berendsen, M. Olivucci, A. E. Mark, and M. A. Robb, J. Am. Chem. Soc. **126**(13), 4228–4233 (2004).
[23] V. Molina and M. Merchan, Proc. Natl. Acad. Sci. U.S.A. **98**(8), 4299–4304 (2001).
[24] C. Ko, B. Levine, A. Toniolo, L. Manohar, S. Olsen, H.-J. Werner, and T. J. Martinez, J. Am. Chem. Soc. **125**(42), 12710–12711 (2003).
[25] P. B. Coto, D. Roca-Sanjuan, L. Serrano-Andres, A. Martin-Pendas, S. Marti, and J. Andres, J. Chem. Theory Comput. **5**(11), 3032–3038 (2009).
[26] M. Boggio-Pasqua, M. A. Robb, and G. Groenhof, J. Am. Chem. Soc. **131**(38), 13580–13581 (2009).
[27] E. V. Gromov, I. Burghardt, H. Koppel, and L. S. Cederbaum, J. Am. Chem. Soc. **129**(21), 6798–6806 (2007).
[28] A. Sergi, M. Gruning, M. Ferrario, and F. Buda, J. Phys. Chem. B **105**(19), 4386–4391 (2001).
[29] G. Groenhof, M. F. Lensink, H. J. C. Berendsen, J. G. Snijders, and A. E. Mark, Proteins: Struct., Funct., Genet. **48**(2), 202–211 (2002).
[30] M. J. Thompson, D. Bashford, L. Noodleman, and E. D. Getzoff, J. Am. Chem. Soc. **125**(27), 8186–8194 (2003).
[31] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martinez, Mol. Phys. **104**(5–7), 1039–1051 (2006).
[32] C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker, and T. J. Martínez, J. Chem. Theory Comput. **8**(12), 5092–5106 (2012).
[33] E. M. González, L. Guidoni, and C. Molteni, Phys. Chem. Chem. Phys. **11**(22), 4556–4563 (2009).
[34] T. Rocha-Rinza, K. Sneskov, O. Christiansen, U. Ryde, and J. Kongsted, Phys. Chem. Chem. Phys. **13**(4), 1585–1589 (2011).
[35] A. Xie, W. D. Hoff, A. R. Kroon, and K. J. Hellingwerf, Biochemistry **35**(47), 14671–14678 (1996).
[36] W. D. Hoff, A. Xie, I. H. M. Van Stokkum, X.-J. Tang, J. Gural, A. R. Kroon, and K. J. Hellingwerf, Biochemistry **38**(3), 1009–1017 (1999).
[37] E. D. Getzoff, K. N. Gutwin, and U. K. Genick, Nat. Struct. Mol. Biol. **5**, 568–570 (1998).
[38] M. Baca, G. E. O. Borgstahl, M. Boissinot, P. M. Burke, D. R. Williams, K. A. Slater, and E. D. Getzoff, Biochemistry **33**(48), 14369–14377 (1994).
[39] G. E. O. Borgstahl, D. R. Williams, and E. D. Getzoff, Biochemistry **34**(19), 6278–6287 (1995).
[40] M. Kim, R. A. Mathies, W. D. Hoff, and K. J. Hellingwerf, Biochemistry **34**(39), 12669–12672 (1995).
[41] E. Demchuk, U. K. Genick, T. T. Woo, E. D. Getzoff, and D. Bashford, Biochemistry **39**(5), 1100–1113 (2000).

[42]M. Yoda, Y. Inoue, and M. Sakurai, J. Phys. Chem. B **107**(51), 14569–14575 (2003).
[43]A. R. Kroon, W. D. Hoff, H. P. M. Fennema, J. Gijzen, G.-J. Koomen, J. W. Verhoeven, W. Crielaard, and K. J. Hellingwerf, J. Biol. Chem. **271**(50), 31949–31956 (1996).
[44]J. J. Vanbeeumen, B. V. Devreese, S. M. Vanbun, W. D. Hoff, K. J. Hellingwerf, T. E. Meyer, D. E. McRee, and M. A. Cusanovich, Protein Sci. **2**(7), 1114–1125 (1993).
[45]A. Baltuška, I. H. M. van Stokkum, A. Kroon, R. Monshouwer, K. J. Hellingwerf, and R. van Grondelle, Chem. Phys. Lett. **270**(3–4), 263–266 (1997).
[46]L. Ujj, S. Devanathan, T. E. Meyer, M. A. Cusanovich, G. Tollin, and G. H. Atkinson, Biophys. J. **75**(1), 406–412 (1998).
[47]S. Devanathan, A. Pacheco, L. Ujj, M. Cusanovich, G. Tollin, S. Lin, and N. Woodbury, Biophys. J. **77**(2), 1017–1023 (1999).
[48]Y. Imamoto, M. Kataoka, F. Tokunaga, T. Asahi, and H. Masuhara, Biochemistry **40**(20), 6047–6052 (2001).
[49]T. Gensch, C. C. Gradinaru, I. H. M. van Stokkum, J. Hendriks, K. J. Hellingwerf, and R. van Grondelle, Chem. Phys. Lett. **356**(3–4), 347–354 (2002).
[50]R. Brudler, R. Rammelsberg, T. T. Woo, E. D. Getzoff, and K. Gerwert, Nat. Struct. Biol. **8**, 265–270 (2001).
[51]U. K. Genick, S. Devanathan, T. E. Meyer, I. L. Canestrelli, E. Williams, M. A. Cusanovich, G. Tollin, and E. D. Getzoff, Biochemistry **36**(1), 8–14 (1997).
[52]Y. Imamoto, K. I. Mihara, O. Hisatomi, M. Kataoka, F. Tokunaga, N. Bojkova, and K. Yoshihara, J. Biol. Chem. **272**(20), 12905–12908 (1997).
[53]G. Groenhof, M. F. Lensink, H. J. C. Berendsen, and A. E. Mark, Proteins: Struct., Funct., Bioinf. **48**(2), 212–219 (2002).
[54]T. E. Meyer, M. A. Cusanovich, and G. Tollin, Arch. Biochem. Biophys. **306**(2), 515–517 (1993).
[55]B. Borucki, S. Devanathan, H. Otto, M. A. Cusanovich, G. Tollin, and M. P. Heyn, Biochemistry **41**(31), 10026–10037 (2002).
[56]T. E. Meyer, G. Tollin, J. H. Hazzard, and M. A. Cusanovich, Biophys. J. **56**(3), 559–564 (1989).
[57]U. K. Genick, G. E. O. Borgstahl, K. Ng, Z. Ren, C. Pradervand, P. M. Burke, V. Šrajer, T.-Y. Teng, W. Schildkamp, D. E. McRee, K. Moffat, and E. D. Getzoff, Science **275**(5305), 1471–1475 (1997).
[58]P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, J. Chem. Phys. **105**(5), 1968–1986 (1996).
[59]H. Li, H. M. Netzloff, and M. S. Gordon, J. Chem. Phys. **125**(19), 194103 (2006).
[60]P. Pulay, Mol. Phys. **17**(2), 197–204 (1969).
[61]F. Furche and R. Ahlrichs, J. Chem. Phys. **117**(16), 7433–7447 (2002).
[62]F. Furche and R. Ahlrichs, J. Chem. Phys. **121**(24), 12772–12773 (2004).
[63]N. C. Handy and H. F. Schaefer III, J. Chem. Phys. **81**(11), 5031–5033 (1984).
[64]M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**(11), 1347–1363 (1993).
[65]M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, 2005).
[66]M. Chiba, T. Tsuneda, and K. Hirao, J. Chem. Phys. **124**(14), 144106 (2006).
[67]M. Chiba, T. Tsuneda, and K. Hirao, Chem. Phys. Lett. **420**(4–6), 391–396 (2006).
[68]J. B. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, J. Phys. Chem. **96**(1), 135–149 (1992).
[69]H. Li and J. H. Jensen, J. Comput. Chem. **25**(12), 1449–1462 (2004).
[70]V. Barone and M. Cossi, J. Phys. Chem. A **102**(11), 1995–2001 (1998).
[71]A. Klamt and G. Schuurmann, J. Chem. Soc., Perkin Trans. 2 **1993**(5), 799–805.
[72]N. M. Thellamurege and H. Li, J. Chem. Phys. **137**(24), 246101 (2012).
[73]N. M. Thellamurege, D. J. Si, F. C. Cui, H. B. Zhu, R. Lai, and H. Li, "QuanPol: A full spectrum and seamless QM/MM program," J. Comput. Chem. (in press).
[74]W. A. Goddard, T. H. Dunning, W. J. Hunt, and P. J. Hay, Acc. Chem. Res. **6**(11), 368–376 (1973).
[75]P. E. M. Siegbahn, J. Almlof, A. Heiberg, and B. O. Roos, J. Chem. Phys. **74**(4), 2384–2396 (1981).
[76]J. Ivanic and K. Ruedenberg, J. Comput. Chem. **24**(10), 1250–1262 (2003).
[77]J. Ivanic and K. Ruedenberg, Theor. Chem. Acc. **106**(5), 339–351 (2001).
[78]J. M. Rintelman, M. S. Gordon, G. D. Fletcher, and J. Ivanic, J. Chem. Phys. **124**(3), 034303 (2006).
[79]K. K. Docken and J. Hinze, J. Chem. Phys. **57**(11), 4928–4936 (1972).
[80]K. Ruedenberg, L. M. Cheung, and S. T. Elbert, Int. J. Quantum Chem. **16**(5), 1069–1101 (1979).
[81]H.-J. Werner and W. Meyer, J. Chem. Phys. **74**(10), 5794–5801 (1981).
[82]W. Kohn and L. J. Sham, Phys. Rev. **140**(4A), A1133–A1138 (1965).
[83]C. Møller and M. S. Plesset, Phys. Rev. **46**(7), 618–622 (1934).
[84]H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, Nucleic Acids Res. **28**(1), 235–242 (2000).
[85]R. Rodriguez, G. Chinea, N. Lopez, T. Pons, and G. Vriend, Bioinformatics **14**(6), 523–528 (1998).
[86]P. Cieplak, J. Caldwell, and P. Kollman, J. Comput. Chem. **22**(10), 1048–1057 (2001).
[87]D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**(16), 1668–1688 (2005).
[88]Z. X. Wang, W. Zhang, C. Wu, H. X. Lei, P. Cieplak, and Y. Duan, J. Comput. Chem. **27**(6), 781–790 (2006).
[89]J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25**(9), 1157–1174 (2004).
[90]D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman (2012), *AMBER 12*, University of California, San Francisco.
[91]F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, and P. Cieplak, Phys. Chem. Chem. Phys. **12**(28), 7821–7839 (2010).
[92]See supplementary material at http://dx.doi.org/10.1063/1.4819139 for QuanPol method for treating QM-MM interaction, Table SI, and Figure S1.
[93]M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees, and J. A. Pople, J. Chem. Phys. **77**(7), 3654–3665 (1982).
[94]C. S. Pomelli, J. Tomasi, and V. Barone, Theor. Chem. Acc. **105**(6), 446–451 (2001).
[95]R. H. Hertwig and W. Koch, Chem. Phys. Lett. **268**(5–6), 345–351 (1997).
[96]S. Yamaguchi, H. Kamikubo, K. Kurihara, R. Kuroki, N. Niimura, N. Shimizu, Y. Yamazaki, and M. Kataoka, Proc. Natl. Acad. Sci. U.S.A. **106**, 440–444 (2009).
[97]K. Saito and H. Ishikita, Proc. Natl. Acad. Sci. U.S.A. **109**(1), 167–172 (2011).
[98]Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, J. Chem. Phys. **120**(18), 8425–8433 (2004).
[99]B. M. Bode and M. S. Gordon, J. Mol. Graphics Modell. **16**(3), 133–138 (1998).
[100]I. B. Nielsen, S. Boye-Peronne, M. O. A. El Ghazaly, M. B. Kristensen, S. Brondsted Nielsen, and L. H. Andersen, Biophys. J. **89**(4), 2597–2604 (2005).
[101]Y. Ma, M. Rohlfing, and C. Molteni, J. Chem. Theory Comput. **6**(1), 257–265 (2010).
[102]D. J. Tozer and N. C. Handy, J. Chem. Phys. **109**(23), 10180–10189 (1998).
[103]D. J. Tozer, J. Chem. Phys. **112**(8), 3507–3515 (2000).
[104]A. Dreuw, J. L. Weisman, and M. Head-Gordon, J. Chem. Phys. **119**(6), 2943–2946 (2003).
[105]M. Chiba, D. G. Fedorov, and K. Kitaura, Chem. Phys. Lett. **444**(4–6), 346–350 (2007).
[106]M. Chiba, D. G. Fedorov, and K. Kitaura, J. Comput. Chem. **29**(16), 2667–2676 (2008).
[107]L. J. G. W. van Wilderen, M. A. van der Horst, I. H. M. van Stokkum, K. J. Hellingwerf, R. van Grondelle, and M. L. Groot, Proc. Natl. Acad. Sci. U.S.A. **103**(41), 15050–15055 (2006).
[108]M. Mizuno, H. Kamikubo, M. Kataoka, and Y. Mizutani, J. Phys. Chem. B **115**(29), 9306–9310 (2011).
[109]R. Nakamura, N. Hamada, K. Abe, and M. Yoshizawa, J. Phys. Chem. B **116**(51), 14768–14775 (2012).