



HAL
open science

Efficient analytic second derivative of electrostatic embedding QM/MM energy: normal mode analysis of plant cryptochrome

Karno Schwinn, Nicolas Ferré, Miquel Huix-Rotllant

► **To cite this version:**

Karno Schwinn, Nicolas Ferré, Miquel Huix-Rotllant. Efficient analytic second derivative of electrostatic embedding QM/MM energy: normal mode analysis of plant cryptochrome. *Journal of Chemical Theory and Computation*, 2020, 16 (6), pp.3816-3824. 10.1021/acs.jctc.9b01145 . hal-02395105

HAL Id: hal-02395105

<https://amu.hal.science/hal-02395105>

Submitted on 6 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficient analytic second derivative of electrostatic embedding QM/MM energy: normal mode analysis of plant cryptochrome

Karno Schwinn, Nicolas Ferré, and Miquel Huix-Rotllant*

Aix-Marseille Univ, CNRS, ICR, Marseille, France

E-mail: miquel.huixrotllant@univ-amu.fr

Abstract

Analytic second derivatives of electrostatic embedding (EE) quantum mechanics/molecular mechanics (QM/MM) energy are important for performing vibrational analysis and simulating vibrational spectra of quantum systems interacting with an environment represented as a classical electrostatic potential. The main bottleneck of EE-QM/MM second derivatives is the solution of coupled perturbed equations for each MM atom perturbation. Here, we exploit the Q-vector method [J. Chem. Phys., 151, 041102 (2019)] to work around this bottleneck. We derive the full analytic second derivative of the EE-QM/MM energy, which allows to compute QM, MM and QM-MM Hessian blocks in an efficient and easy to implement manner. To show the capabilities of our method, we compute the normal modes for the full *arabidopsis thaliana* plant cryptochrome. We show that the flavin adenine dinucleotide vibrations (QM subsystem) strongly mix with protein modes. We compute approximate vibronic couplings for the lowest bright transition, from which we extract spectral densities and the homogeneous broadening of FAD absorption spectrum in protein using vibrationally resolved electronic spectrum simulations.

1 Introduction

Analytic derivatives of the energy revolutionized quantum chemistry,¹ giving access to gradients and Hessians (respectively first and second derivative with respect to nuclear positions) to explore potential energy surfaces,² in addition to derivatives with respect to electromagnetic fields, from which properties are nowadays routinely extracted.

The development of mixed quantum mechanics (QM) and molecular mechanics (MM) models by Warshel, Levitt and Karplus, allowed a more realistic simulation of quantum systems embedded in complex environments.³⁻⁶ Analytic derivatives of QM/MM energy have essentially focused on analytic gradient, which have been implemented for the main self-consistent

field (SCF) and post-SCF QM methods.⁷⁻¹⁴ Analytic second derivatives of QM/MM energy are less commonly available. As a matter of fact, most of the QM/MM Hessian constructions are based on the so-called Partial Hessian approach (see 15 and references therein) in which only the QM degrees of freedom are considered. A first derivation of an analytic Hessian was introduced by Morokuma and coworkers in the framework of the subtractive ONIOM QM/MM scheme.¹⁶ One year later, Cui and Karplus proposed an analytic Hessian evaluation in their additive QM/MM approach.^{17,18} The latter method is expensive, mainly due to the fact that the Hessian calculation involves solving a coupled perturbed (CP) SCF equation not only for the QM atoms but also for the MM atoms. Thus, the method

of Cui and Karplus resorted to approximations of the CPSCF equations. More recently, Ghysels, Woodcock III and coworkers developed a coarse-graining method for the MM environment to reduce the number of CPSCF equations to solve.¹⁹ A Hessian with fragment molecular orbital method has been derived by Nakata and coworkers.²⁰ Other approximate second derivative methods have been proposed by Sakai and Morita,²¹ and Gadre and coworkers.²² The most up-to-date formulation of the Hessian is certainly the one by Capelli and coworkers, especially derived for their very promising polarizable QM/MM method which can be used to calculate infra-red spectra of very large systems in the so-called Partial Hessian Vibrational Approach, hence ignoring the major part of the Hessian that only depends on the MM atoms.²³

Herein, we develop a new method for calculating the complete second derivatives of the QM/MM energy with additive QM-MM interaction terms. We base our derivation on single-determinant SCF wavefunctions characterizing a QM charge density interacting with a classical point-like atomic environment via the Electrostatic Potential Fitting (ESPF) interaction Hamiltonian developed by one of us.^{9,10} For that purpose, we apply the CPSCF Q-vector method for MM atoms,²⁴ avoiding one of the main bottlenecks of second derivatives of QM-MM energies. The method works even when the QM and MM regions are covalently linked via the use of the link atom method of Morokuma and coworkers.¹⁶ For testing our implementation, we apply this method to perform a normal mode analysis of *arabidopsis thaliana* cryptochrome,²⁵ and we compute the vibronic couplings, which allows us to efficiently construct spectral densities and account of homogeneous broadening of the spectrum via vibrationally resolved spectrum simulations.

2 Analytic Derivatives of the ESPF QM/MM energy

We use the notation $\mathbf{R} = (x, y, z)$ to symbolize the coordinates of QM atom, and $\tilde{\mathbf{R}} = (\tilde{x}, \tilde{y}, \tilde{z})$ to symbolize the coordinates of a MM atom. The partial derivatives with respect to a coordinate will be written in the short-hand notation

$$E^x = \frac{\partial E}{\partial x} \quad (1)$$

The indexes i, j, \dots will refer to occupied molecular orbitals (MO), a, b, \dots will refer to virtual MOs, p, q, \dots to generic MOs (both occupied and virtual), and μ, ν, \dots to atomic orbital (AO) index. Capital roman letters will be used for QM atom indexes.

2.1 ESPF QM/MM energy

In the ESPF method, the QM/MM energy can be written as

$$E = E_{QM} + E_{MM} \quad (2)$$

in which E_{QM} is the internal energy of the QM subsystem whose electron density is polarized by the MM subsystem (thus containing the energy contribution due to the electrostatic interaction with the MM subsystem), while E_{MM} is the internal energy of the MM part complemented with classical interaction terms between QM and MM atoms. The electrostatic interaction energy in the ESPF method is evaluated as $\Delta E = \text{Tr}[\mathbf{P}\mathbf{h}]$, in which \mathbf{P} is the atomic density and \mathbf{h} is the QM/MM interaction hamiltonian.^{9,10} This interaction energy is also expressed as the classical electrostatic interaction between quantum-derived atomic multipoles q_A, μ_A, \dots and the MM electrostatic potential ϕ_A , the electric field vector $\nabla\phi_A, \dots$

$$\begin{aligned} \Delta E &= \sum_A (q_A \phi_A - \boldsymbol{\mu}_A \cdot \nabla \phi_A + \dots) \\ &= \Delta E^0 + \Delta E^1 + \dots \end{aligned} \quad (3)$$

in which the index A runs over the number of QM atoms N_{QM} . For the sake of simplicity, we will restrict the multipolar expansion (Eq. 3) to the order 0 in the following, even though our method can be easily generalized to higher orders.²⁴ The expressions for the derivatives of the QM-only and MM-only energy terms will be omitted here, since they have been lengthy discussed in the literature.^{9,26,27} Here, we will focus on the electrostatic interaction term and its derivatives.

At the order 0 of the QM density of charges, each atomic charge q_A is expressed as the sum of the nuclear charge Z_A and the mean value of an atomic one-electron operator, whose matrix elements are denoted $Q_{A,\mu\nu}$. The interaction energy in atomic orbital basis reads,

$$\begin{aligned}\Delta E &= \sum_{A,\mu\nu} (-P_{\mu\nu}Q_{A,\mu\nu} + Z_A) \phi_A \\ &= \sum_A q_A \phi_A = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \sum_A Z_A \phi_A\end{aligned}\quad (4)$$

in which we have introduced a new one-electron operator whose elements are;

$$h_{\mu\nu} = - \sum_A Q_{A,\mu\nu} \phi_A \quad (5)$$

In a nutshell, the ESPF method consists in fitting the $Q_{A,\mu\nu}$ matrix elements to reproduce the purely electronic QM electrostatic integrals calculated on a grid surrounding the QM subsystem.⁹

2.2 First derivative

The analytic first derivatives of the ESPF interaction term were first derived in Ref. 9. For completeness, we reexpress them with the current notation. The first derivative of Eq. 4 can be expressed as a vector gradient,

$$\nabla[\Delta E] = \begin{bmatrix} \Delta E^x \\ \Delta E^{\tilde{x}} \end{bmatrix} \quad (6)$$

in which the first block corresponds to the derivative with respect to QM atom positions and the second block to MM atoms. The expression of the energy gradient with respect to

QM coordinates is given by

$$\Delta E^x = \sum_{\mu\nu} P_{\mu\nu}^x h_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu}^x + \sum_A Z_A \phi_A^x \quad (7)$$

As usual, the first term is not calculated thanks to the introduction of the so-called energy-weighted density matrix $-\mathbf{PhPS}^x$.^{9,10,28} The derivative of the ESPF operator element $h_{\mu\nu}$ is given in Appendix 5.1. Regarding the derivative with respect to the MM coordinates, the energy gradient can be expressed as

$$\Delta E^{\tilde{x}} = \sum_A q_A \phi_A^{\tilde{x}}, \quad (8)$$

noting that $\mathbf{S}^{\tilde{x}} = 0$ for MM perturbations.

2.3 Second derivative

The analytic second derivative of Eq. 4 can be expressed as a Hessian matrix

$$\nabla^2[\Delta E] = \begin{bmatrix} \Delta E^{xy} & \Delta E^{x\tilde{y}} \\ \Delta E^{\tilde{x}y} & \Delta E^{\tilde{x}\tilde{y}} \end{bmatrix} \quad (9)$$

in which the ΔE^{xy} is the block of second derivatives with respect to QM atom coordinates, $\Delta E^{\tilde{x}\tilde{y}}$ is the second derivative block with respect to MM atom coordinates, and $\Delta E^{x\tilde{y}} = \Delta E^{\tilde{x}y}$ is the second derivative block with respect to one MM and one QM atom coordinates.

The analytic formula for the ΔE^{xy} second derivative is given by

$$\Delta E^{xy} = \sum_A Z_A \phi_A^{xy} + \sum_{\mu\nu} (P_{\mu\nu} h_{\mu\nu}^{xy} + P_{\mu\nu}^y h_{\mu\nu}^x) \quad (10)$$

The second derivative of the Fock operator requires the second derivative of the ESPF operator, which is given in 5.1. This second derivative block also requires the density matrix derivative \mathbf{P}^y with respect to QM atom coordinates, which is obtained by solving a CPSCF equation for each QM coordinate.²⁹

The $\Delta E^{x\tilde{y}}$ block over a QM and an MM atom

coordinates is given by

$$\Delta E^{x\tilde{y}} = \sum_A Z_A \phi_A^{x\tilde{y}} + \sum_{\mu\nu} (P_{\mu\nu} h_{\mu\nu}^{x\tilde{y}} + P_{\mu\nu}^{\tilde{y}} h_{\mu\nu}^x) \quad (11)$$

The explicit form of the second derivative of the ESPF operator is given in Appendix 5.1. This energy derivative requires the explicit construction of $\mathbf{P}^{\tilde{x}}$. These would require the solution of a CPSCF equation for all MM atom coordinates, which is prohibitive. A more convenient form to implement, is to compute the equivalent $\Delta E^{\tilde{x}y}$ derivative block, which is given by

$$\Delta E^{\tilde{x}y} = \sum_A \left(q_A^y \phi_A^{\tilde{x}} + q_A \phi_A^{\tilde{x}y} \right) \quad (12)$$

in which the charge derivative is given by

$$q_A^y = - \sum_{\mu\nu} (P_{\mu\nu}^y Q_{A,\mu\nu} + P_{\mu\nu} Q_{A,\mu\nu}^y) \quad (13)$$

In this case, only derivatives of the density matrix with respect to QM coordinates are required, which is already constructed for the QM energy second derivative.²⁶

Finally, the $\Delta E^{\tilde{x}y}$ derivatives over MM atom coordinates is given by

$$\Delta E^{\tilde{x}y} = \sum_A \left(q_A^{\tilde{y}} \phi_A^{\tilde{x}} + q_A \phi_A^{\tilde{x}y} \right) \quad (14)$$

requires the derivatives of the ESPF atomic charges, which would in principle require the density derivatives

$$q_A^{\tilde{y}} = - \sum_{\mu\nu} P_{\mu\nu}^{\tilde{y}} Q_{A,\mu\nu} \quad (15)$$

The construction of $\mathbf{P}^{\tilde{y}}$ would imply solving a coupled-perturbed type equation for each MM perturbation. Usually, the number of MM atoms is hundreds to thousands, becoming rapidly intractable. In Ref. 24, we showed how to avoid this by defining an auxiliary set of coupled-perturbed equations (Q-vector method) that only scales with the number of QM atoms, and from which we can efficiently

reconstruct the charge derivatives,

$$q_A^{\tilde{y}} = - \sum_{B,ia} Q_{A,ia} \phi_B^{\tilde{y}} \tilde{Q}_{B,ia} \quad (16)$$

where $\tilde{Q}_{B,ia}$ is the solution of the Q-vector coupled-perturbed equation,²⁴ which allows to rewrite the MM contribution to second derivative as

$$\Delta E^{\tilde{x}y} = \sum_A \left(q_A \phi_A^{\tilde{x}y} + \sum_{B,ia} Q_{A,ia} \tilde{Q}_{B,ia} \phi_B^{\tilde{y}} \phi_A^{\tilde{x}} \right) \quad (17)$$

This equation can be interpreted in terms of classical atomistic mechanics,

$$\Delta E^{\tilde{x}y} = \sum_A \left(q_A \phi_A^{\tilde{x}y} + \sum_B Q_A Q'_{AB} \phi_B^{\tilde{y}} \phi_A^{\tilde{x}} \right) \quad (18)$$

in which $Q_A = -Tr[\mathbf{P}Q_A]$, and

$$Q'_{AB} = \frac{\sum_{ia} \tilde{Q}_{B,ia} Q_{A,ia}}{Q_A} \quad (19)$$

Then, Q'_{AB} can be interpreted as a tensor that, given an electric charge Q_A on QM atom A interacting with a gradient of the external potential, will give the “effective” charge on QM center B that interacts with another gradient of the external field.

2.4 Link atoms

For molecular systems in which there exist covalent bonds at the interface between the MM and QM subsystems, unphysical free valences are created in the QM side. Morokuma and coworkers proposed to use a fictitious link atom in order to saturate these dangling bonds, most often using a hydrogen atom. The link atom is constrained to remain positioned between the frontier MM and QM atoms,¹⁶

$$\mathbf{R}_{link} = \mathbf{R}_{QM} + g \left(\tilde{\mathbf{R}}_{MM} - \mathbf{R}_{QM} \right) \quad (20)$$

in which g is a parameter that fixes the coordinates of the link atom between the covalently bonded QM and MM atoms.

From the QM perspective, the “auxiliary” link

atom is considered as a QM atom, and therefore, N_{QM} includes all real QM atoms plus the link atoms. Consequently, there will be an extra contribution for the QM and MM frontier atoms that comes from the link atom. In the case of first energy derivatives, this is given by

$$\begin{aligned} E^x|_{link} &= E^x + (1-g)\frac{\partial E}{\partial x_{link}} \\ E^{\tilde{x}}|_{link} &= E^{\tilde{x}} + g\frac{\partial E}{\partial \tilde{x}_{link}} \end{aligned} \quad (21)$$

In other words, the link atom gradient is projected onto the frontier QM and MM atoms. Similarly, the contribution of the link atom to the MM and QM second derivatives can be expressed as

$$\begin{aligned} E^{xy}|_{link} &= E^{xy} + (1-g)^2\frac{\partial^2 E}{\partial x_{link}\partial y_{link}} \\ &+ (1-g)\left(\frac{\partial^2 E}{\partial x\partial y_{link}} + \frac{\partial^2 E}{\partial x_{link}\partial y}\right) \\ E^{x\tilde{y}}|_{link} &= E^{x\tilde{y}} + g(1-g)\frac{\partial^2 E}{\partial x_{link}\partial y_{link}} \\ &+ (1-g)\frac{\partial^2 E}{\partial x_{link}\partial \tilde{y}} + g\frac{\partial^2 E}{\partial x\partial y_{link}} \\ E^{\tilde{x}y}|_{link} &= E^{\tilde{x}y} + g(1-g)\frac{\partial^2 E}{\partial x_{link}\partial y_{link}} \\ &+ g\frac{\partial^2 E}{\partial x_{link}\partial y} + (1-g)\frac{\partial^2 E}{\partial \tilde{x}\partial y_{link}} \\ E^{\tilde{x}\tilde{y}}|_{link} &= E^{\tilde{x}\tilde{y}} + g^2\frac{\partial^2 E}{\partial x_{link}\partial y_{link}} \\ &+ g\left(\frac{\partial^2 E}{\partial \tilde{x}\partial y_{link}} + \frac{\partial^2 E}{\partial x_{link}\partial \tilde{y}}\right) \end{aligned} \quad (22)$$

3 Computational Details

All electronic structure computations have been performed with a modified version of GAUSSIAN16 including the Q-vector method and the ESPF energy, gradient and Hessian. The MM information is obtained via an interface with a modified version of TINKER 6.3.3.^{9,30,31}

The vibrationally resolved spectrum for FAD has been obtained by Fourier transforming the time-dependent auto-correlation function for the multi-dimensional harmonic oscillator, as implemented in GAUSSIAN16.^{30,32} The Franck-

Condon approximation of the auto-correlation function has been integrated on 2^{20} time steps, over an interval of about 10.5 ps. A Gaussian broadening of 235 cm^{-1} has been applied to the vibrational peaks.

All QM calculations have been performed at the DFT and full TDDFT levels,³³ using the B3LYP functional³⁴⁻³⁶ and the TZVP/TZVPfit basis set.³⁷⁻⁴⁰ For the MM calculations, we have employed the AMBER99 force field.⁴¹ Van der Waals parameters for FAD have been taken from Ref. 42.

4 Results and Discussion

As a first example of our method, we have computed the frequencies of flavin adenine dinucleotide at the QM/MM level. In the QM region, we have placed the isoalloxazine ring while the rest of the molecule is considered MM. This selection requires the introduction of a link atom to separate the QM and MM regions, which has been placed 3 bonds far away from the isoalloxazine ring to avoid overpolarization, as shown in Fig. 1. In this figure, we show the frequencies computed at the QM/MM level compared to the QM frequencies, which we take as reference. For the sake of comparison, we also show the pure MM frequencies. Overall, the quality of the QM/MM and MM frequencies is in very good accordance with the QM frequencies. This is expected since the MM parameters have been obtained from purely *ab initio* data. Some appreciable errors are observed for the highest frequency modes, as well as the mid-range. The QM/MM frequencies follow a similar trend than the MM frequencies, indicating that these modes have simply a small mixing and remain rather purely MM or purely QM. Indeed, the most delocalized frequencies are found only in the low frequency range (see further below). This results indicate that the choice of force field is indeed critical for obtaining a good spectrum of frequencies that is comparable to the *ab initio* spectrum.

To further illustrate the potential of our method, we have performed a normal mode analysis of plant cryptochrome, including vi-

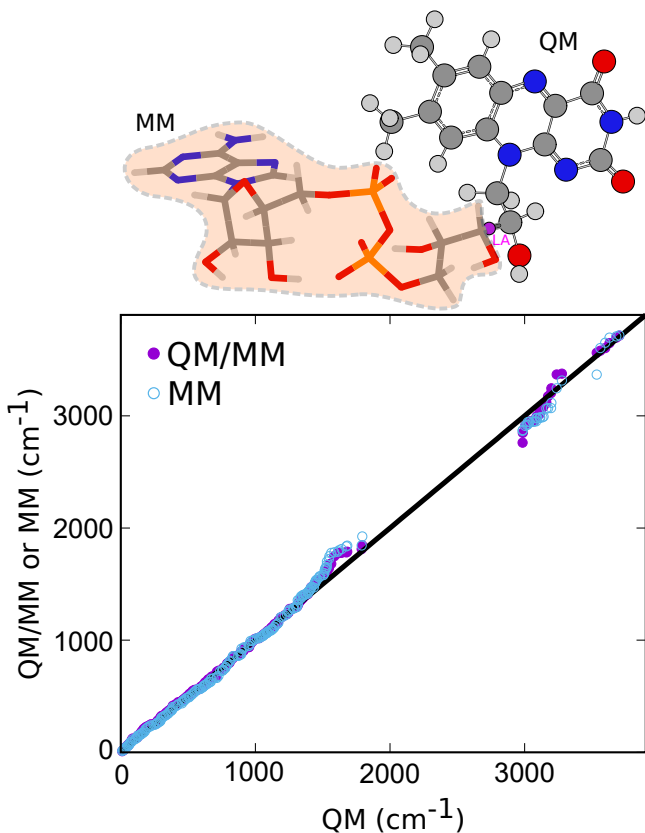


Figure 1: Comparison of the QM/MM and MM frequencies (y-axis) with respect to full QM frequencies (x-axis) of flavin adenine dinucleotide. For the computation of frequencies, a geometry optimization has been performed at each level of theory (QM: B3LYP/6-31G*, MM: Amber99; QM/MM: B3LYP/6-31G*/Amber99). For the QM/MM calculation, a link atom has been placed between the isoalloxazine ring (treated at the QM level) and the rest of the molecule as indicated in the graph.

bronic coupling and homogeneous broadening of absorption spectrum using vibrationally resolved spectra simulations. For this, we have selected chain A of *arabidopsis thaliana* cryptochrome (PDB code 2J4D, see Fig. 2). Crystallographic water molecules have been kept in the model. Hydrogens have been added for standard aminoacid protonation states at physiological pH. The entire FAD coenzyme was defined as the QM subsystem. FAD is not covalently linked to the rest of the protein, thus not requiring any special care of the QM/MM boundary. The initial PDB structure was pre-optimized at the Amber99 level using microiterations. Subsequent geometry optimization, hessian calculation and normal mode analysis were carried out at B3LYP/TZVP/TZVPfit/Amber99 level for the entire protein model.

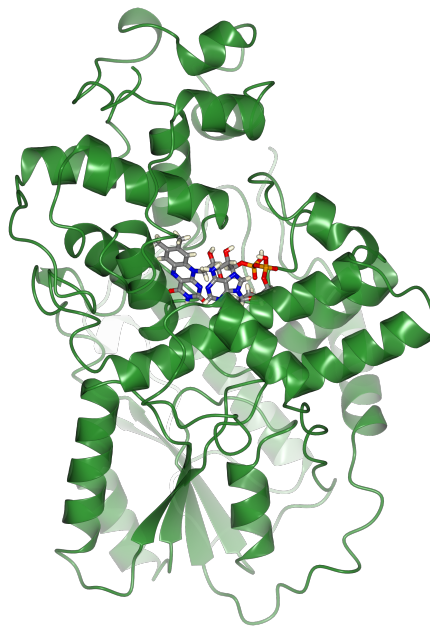


Figure 2: Representation of the *arabidopsis thaliana* cryptochrome (PDBID: 2J4D) used in the QM/MM simulations. The FAD fragment (ball and stick) constitutes the QM subsystem, while the protein (ribbon) is treated at the MM level. Crystallographic water molecules have been excluded from the plot, but included in the simulations.

One of the main advantages of the method presented here is that the CPSCF equations for the MM perturbations are not scaling with

the MM subsystem size. As we showed, we can decouple this dependence by solving the Q-vector equations for N_{QM} atoms and reconstruct the atomic charge derivatives with respect to MM perturbations by contracting with the external field derivatives. This induces an obvious saving in computational time. For the present case (plant cryptochrome, 9432 atoms at the B3LYP/TZVP/TZVPFit/Amber99 level using 32 cores of 2.6 GHz), the solution of conventional CPSCF equations for the full system would take around 86.5 hours, while the Q-vector method (84 atom) just takes 17 minutes. This allows the construction of the full Hessian including the charge polarization terms at the cost of a mechanical embedding QM/MM full Hessian.

4.1 Normal mode analysis

We performed the normal mode analysis of the plant cryptochrome shown in Fig. 2 at the minimum energy structure of the QM/MM internal energy. This geometry was obtained from a quadratically convergent optimization of the QM subsystem, with a steepest descent optimization of the MM subsystem at every step of optimization (microiterations). At the minimum structure, the FAD remains in a U-folded conformation due to the electrostatic interactions with the protein environment. The protein model has 8121 atoms, from which 84 atoms are in the QM subsystem. This leads to a total of 24357 normal modes. In Fig. 3, we analyse the degree of delocalization as well as the QM character of these modes. First, we computed the participation ratio in order to measure the degree of delocalization. For this, we defined the participation ratio as

$$PR(\omega_i) = \left(3N \cdot \sum_j^{3N} ({}^0L_{i,j}^\dagger \cdot {}^0L_{j,i})^2 \right)^{-1} \quad (23)$$

in which 0L_i is the mass-weighted normal mode displacement of the ground state. For a fully delocalized mode, all atoms contribute equally to the mass-weighted displacement, and the participation ratio is equal to one, whereas if

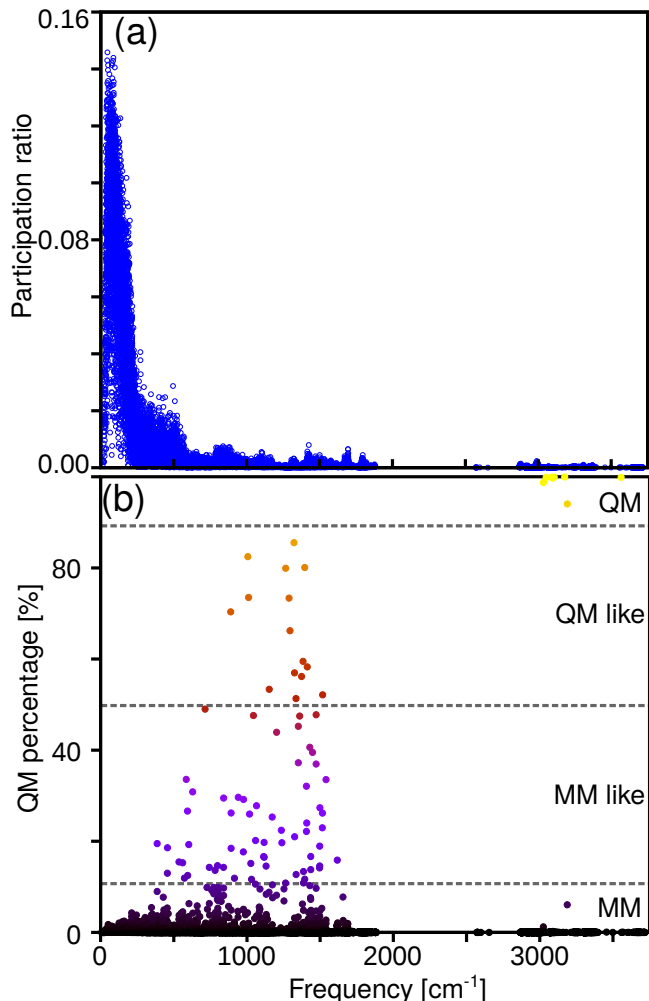


Figure 3: Normal mode analysis of plant cryptochrome. (a) participation ratio showing the degree of delocalization of normal modes, and (b) analysis of the character QM character. Modes with less than 10% QM atom displacement are classified as MM (black) while modes with more than 90% are classified as QM (yellow).

only few atoms contribute to the displacement, the participation ratio has the value of $(3N)^{-1}$ (which is close to 0 in the case of plant cryptochrome.) Following this analysis, we can observe that the low frequency modes (<200 cm^{-1}) show on average the largest degree of delocalization, involving thus a large number of atoms in each normal mode frequency. These modes are usually involved in describing protein phase transitions. The rest of the normal modes involve only a few atoms.

The participation ratio analysis is not sufficient to determine the degree of mixing between QM and MM atom motions. If one would perform a QM subsystem normal mode analysis, one would obtain 246 full QM normal modes. For the full protein diagonalization, these modes are strongly mixed. In Fig. 3, we report the QM character of each normal mode, defined as

$$QM(\omega_i) = \sum_{j \in QM}^{3N} \left({}^0\mathbf{L}_{i,j}^\dagger \right)^2. \quad (24)$$

As we can observe, the low frequency modes below 200 cm^{-1} have a negligible QM character. This implies a strong mixing between the QM and MM atoms in this region. Since the number of MM atoms is so large, the QM character is “diluted” within the MM motion. Only two normal modes between 1200 - 1400 cm^{-1} keep a clear QM character ($>90\%$), and another 23 normal modes have a QM character larger than 60% , comprised in the range of 400 - 1600 cm^{-1} . Therefore, only 10% of the FAD normal modes keep in essence its original character. For the rest, QM and MM atom motions are essentially mixed, even if the modes are not fully delocalized over the whole protein. Importantly, this mixing is present in the whole range of frequencies, and it is not only appearing for the fully delocalized low-frequency modes.

4.2 Vibronic couplings, Huang-Rhys factors and spectral densities

The Huang-Rhys couplings are a fundamental measure for the electron-nuclear (vibronic) coupling. The HR factor is directly related to the reorganization energy of one mode in the excited state. They can be defined as

$$\text{HR}_{i,I} = \frac{1}{2} \omega_i \Delta q_{i,I}^2, \quad (25)$$

in which ω_i is the normal mode frequency and $\Delta q_{i,I} = {}^0\mathbf{L}_i \mathbf{M}^{1/2} (\mathbf{R}_I - \mathbf{R}_0)$ is the projection in the ground state normal mode direction ${}^0\mathbf{L}_i$ of the Cartesian geometry difference between excited (\mathbf{R}_I) and ground state (\mathbf{R}_0) minima, mass weighted by the mass metric $\mathbf{M} = \delta_{ij} m_i$, in which m_i is the atomic mass of atom i .

The Δq_i^I can be approximately obtained by the so-called vertical gradient approximation, in which the mass-weighted energy gradient of the excited state I (Δq_i^I) is projected to the normal mode direction, leading to

$$\Delta q_{i,I} = - \frac{(\nabla \mathbf{E}_I | {}^0\mathbf{L}_i)}{\omega_i^2} \quad (26)$$

In order to define the effect of the bath modes, the notion of spectral density $J(\omega)$ is useful. It can be defined as

$$J_I(\omega) = \pi \sum_i \text{HR}_{i,I} \omega_i^2 \delta(\omega - \omega_i) \quad (27)$$

Indeed, the spectral density contains information about each normal mode frequency, and its coupling to the system, which suffices to characterize the internal dissipation of an excited state due to nuclear motion.

Using the information of the QM/MM analytic second derivatives together with the excited-state gradient at the ground-state minimum, we can construct the spectral densities characterizing the same cryptochrome model. In Figure 4, we have plotted such spectral density for the lowest energy bright excited state. Two spectral densities are shown, one corresponding to the normal modes and gradient of the full protein, and the other corresponding

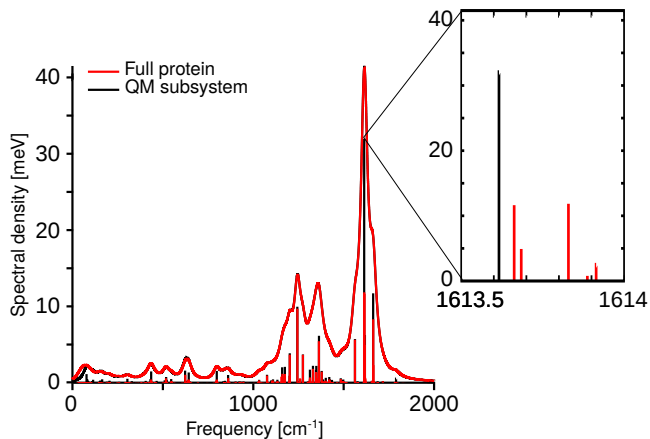


Figure 4: Stick spectrum corresponding to the spectral density for the first bright excited state. The continuous spectral density has been obtained from a Lorentzian broadening of 500 cm^{-1} . The red line correspond to the normal modes of the full protein, while the black line to the QM subsystem normal modes. The zoom-in of the most intense peak is shown in the right-hand inset.

to the normal modes and gradient of the QM subsystem.

The overall contribution of MM atoms to the broadened spectral density is only minor. This is expected because the transition density is mainly located on the chromophore, which is entirely located in the QM subsystem, and thus only modes with substantial QM motion should be contributing to the spectral density. Still, an indirect contribution of the MM atoms appear in the spectral density due to the mode mixing, as shown for example in the most intense peak around $1620\text{-}1625 \text{ cm}^{-1}$. While the QM spectral density has only one peak of 40 meV , the full protein spectral density shows a spectral density contribution of 4 modes. This is due to the mixing of QM and MM atom motions, alike to the intensity borrowing mechanism of electronic spectra, in which dark states couple to bright states borrowing dipole intensity and thus becoming observable. This can be observed for almost all peaks in the spectral density, which reflects once again the strong normal mode mixed character that we showed in Fig. 3. It is important to notice though that even though there is a strong mixing, the spec-

tral densities of the QM and the full protein are essentially equivalent. This is because the normal mode mixing occurs only between QM and MM modes at almost the same frequency. Eventually, the effect of mixing has only a minor impact on the frequency of the modes (less than 1 cm^{-1} for the most intense peak). This results justify the extraction of vibronic couplings from the normal mode analysis of the QM subsystem.

4.3 Vibrationally resolved absorption spectrum

Flavoproteins like cryptochrome act as blue-light sensors in biological systems.⁴³ These proteins have a well-resolved vibrational progression of the absorption spectrum, characteristic of the isoalloxazine chromophore. Unlike other types of proteins in which inhomogeneous broadening dominates, the homogeneous broadening has to be explicitly taken into account to reproduce the color of absorption of the flavin chromophore.^{44,45} The approximate vibronic couplings shown in Eq. 26 can be used to compute the absorption spectrum including the vibrational structure due to homogeneous broadening.

Here, we compute the vibrational resolved spectrum of FAD in plant cryptochrome, by using the QM subsystem normal modes in conjunction with the vibrationally-resolved time-dependent approach of Bloino and coworkers.³² In Fig. 5, we compare the experimental and theoretical absorption spectra for *arabidopsis thaliana* plant cryptochrome with FAD in its fully oxidized form.²⁵ The experimental absorption spectrum is characterized by two main absorption peaks of $\pi \rightarrow \pi^*$ character. The lowest bright transition has a peak maximum around 450 nm and the second bright state at around 360 nm .²⁵ These peaks have a marked vibrational structure, especially due to in-plane C-N stretchings.⁴⁶ Indeed, the first electronic transition is distributed in resolved peaks, approximately found at $470, 450$ and 420 nm . Our simulated spectrum perfectly matches the relative position and intensities of the experimental peaks. From our simulations, we observe that

these peaks are in fact formed from two electronic transitions, a HOMO→LUMO transition which give the basic vibrational structure, and a HOMO-1→LUMO transition which is centered at 420 nm. For the second absorption peak, we have a reasonable agreement between simulated and experimental spectra. Indeed, the relative peak position and the relative intensities between the two vibronic peaks are well reproduced by the theory. However, our theoretical simulations underestimate by around 40% the relative intensity with respect to the lowest energy absorption peak. We attribute this either to the approximations of the level of theory or some missing contributions of low-intensity high-energy peak.

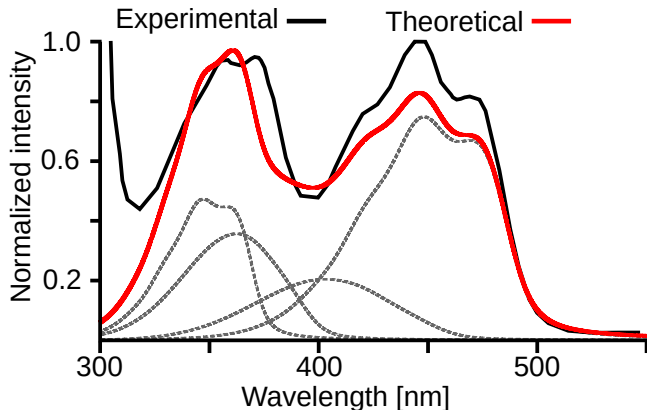


Figure 5: Comparison of the simulated vibrationally resolved absorption spectrum (red) with the experimental spectrum (black) of *Arabidopsis thaliana*. The grey dotted lines indicate the individual contributions from each electronic state. For the sake of comparison, both spectra have been normalized for the highest intensity peak between the range of 350 nm and 500 nm. The simulated spectrum has been red-shifted by 12 nm to match the experimental maximum. Experimental spectrum has been taken from Ref. 25

5 Conclusion

We have derived a fully analytic expression for the second derivative of the electrostatic embedding ESPF QM/MM energy. In this approach, the atomic charges representing the

QM charge distribution are not only polarizable by the usual QM/MM electrostatic interactions, but also by the motion of the MM atoms. QM/MM electrostatic energy second derivative contains MM-block contributions too expensive to compute, and usually simply ignored or approximated. Using the ESPF method, we have shown that these terms can be equivalently computed with a set of auxiliary coupled-perturbed self-consistent field equations that only scales with the QM subsystem size,²⁴ reducing thus the total computational cost of the full electrostatic embedding QM/MM hessian to a similar cost than mechanical embedding QM/MM hessian.

To show the capabilities of our method, we have computed the full hessian of *Arabidopsis thaliana* plant cryptochrome. We have evidenced that most of the FAD chromophore (QM) vibrational modes are strongly mixed with the modes of the protein, and only a few high-frequency modes around 1200-1500 cm^{-1} keep a full QM character. In addition, we have computed the vibronic couplings and spectral densities for the bright excited state, together with the vibrationally resolved spectra of FAD in the protein. These applications open up the use of QM/MM Hessians construct vibronic models for photoactive proteins and QM systems in complex environments.

Several limitations are currently imposed in our current model for a practical use. First, one of the main limitations of our QM/MM method is the lack of periodic boundary conditions to include solvent effects. In the future, we plan to implement our method in an extension of the ESPF QM/MM method to include Ewald summations.⁴⁷ Meantime, the only strategy is to freeze a state of the solvent, and relax the protein within a solvation shell. Since the QM equations do not scale with the MM subsystem size, the calculations considering a large number of solvent molecules should be tractable at a reasonable computational cost. Second, the current simulation of vibrationally resolved absorption spectrum requires a positive semi-definite Hessian matrix. This implies that any initial geometry must be optimized to a minimum energy structure so

that all frequencies are positive. The geometry optimization step is time-consuming, hence severely limiting the usual efficiency associated with the inhomogenous broadening as it is frequently reported in the literature, i.e., by selection of snapshots from a MD simulation and performing the QM/MM analysis at each of them. This limitation could be overcome by using for example instantaneous normal modes,⁴⁸ although further investigation in this direction is required.

Finally, we plan to extend our model to address computed vibrational spectroscopies. Such simulations face frequently the problem of comparing with experimental spectra, which could ultimately be attributed to the lack of anharmonic effects, both from the QM and the MM frequency spectrum. To better match the QM frequencies to experimental data, it is common to scale the Hessian via a scaling factor that depends on the method and the basis set. Morokuma and coworkers applied a similar strategy to their ONIOM QM/MM Hessian, in which three different scaling factors for the QM, MM and QM-MM Hessian blocks are applied.¹⁶ Such scheme would be readily applied to our analytic QM/MM Hessian too. In addition, some recent developments at the level of force fields for better describing infra-red spectra can also be trivially integrated in our Hessian scheme.^{49,50}

5.1 Appendix: Derivatives of the ESPF operator

The ESPF atomic charge operator elements $Q_{A,\mu\nu}$ are obtained by fitting the corresponding electrostatic potential integrals $V_{k,\mu\nu}$,

$$V_{k,\mu\nu} = \int d\mathbf{r} \chi_\mu(\mathbf{r}) |\mathbf{r} - \mathbf{R}_k|^{-1} \chi_\nu(\mathbf{r}) \quad (28)$$

calculated over a grid of N_{grid} points k , at position \mathbf{R}_k , around the QM atoms:

$$\sum_A T_{A,k} Q_{A,\mu\nu} = V_{k,\mu\nu} \quad (29)$$

The kernel \mathbf{T} is a rectangular matrix of dimensions $N_{grid} \cdot N_{QM}$. The kernel definition depends

on the multipolar expansion order. For example, for atomic charges it contains the inverse distance between QM atom and the grid point $T_{A,k} = |\mathbf{R}_A - \mathbf{R}_k|^{-1}$. Thus, we can give an expression for the $Q_{A,\mu\nu} = \langle \chi_\mu | \hat{Q}_A | \chi_\nu \rangle$ atomic charge operator elements by solving the previous set of linear equations, leading to

$$Q_{A,\mu\nu} = \sum_k^{N_{grid}} \left[(\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \right]_{A,k} V_{k,\mu\nu} \quad (30)$$

The derivative of the ESPF hamiltonian involves the derivative of the atomic charge operator. A recursive formula for the n-derivative of the charge operator can be obtained by directly deriving n-times Eq. 29,

$$Q_{A,\mu\nu}^n = \sum_k^{N_{grid}} \left[(\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \right]_{A,k} F_{k,\mu\nu}^n \quad (31)$$

in which

$$F_{k,\mu\nu}^n = V_{k,\mu\nu}^n - \sum_{i=0}^{n-1} \sum_B^{N_{QM}} T_{k,B}^{n-i} Q_{B,\mu\nu}^i \quad (32)$$

Here we give the explicit first and second derivatives of the \mathbf{h} operator defined in Eq. 5. The first derivative of the operator is given by

$$h_{\mu\nu}^x = \sum_A (Q_{A,\mu\nu}^x \phi_A + Q_{A,\mu\nu} \phi_A^x) \quad (33)$$

The second term involves the derivative of the classical external field, which have simple expression given the form of the force field. The derivative of the atomic charge operator can be in term expanded (in matrix notation) as

$$\mathbf{Q}^x = (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger [\mathbf{V}^x - \mathbf{T}^x \mathbf{Q}] \quad (34)$$

we can easily see that the computation of first-order derivatives of \mathbf{h} operators requires the computation of ϕ_A^x , $V_{k,\mu\nu}^x$ and \mathbf{T}^x in addition to the zero-order field and charges.

The second derivatives of the \mathbf{h} operator is

given by

$$h_{\mu\nu}^{xy} = \sum_A (Q_{A,\mu\nu}^{xy} \phi_A + Q_{A,\mu\nu} \phi_A^{xy}) \quad (35)$$

$$+ \sum_A (Q_{A,\mu\nu}^x \phi_A^y + Q_{A,\mu\nu}^y \phi_A^x)$$

The second term involves the second derivative of the classical external field, which again has a simple expression given the form of the force field, while the third and fourth term are constructed with elements of the first derivative of the operator. Hence, one needs only to develop the second derivative of the atomic charge operator. This is given by

$$\mathbf{Q}^{xy} = (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger [\mathbf{V}^{xy} - \mathbf{T}^{xy} \mathbf{Q}]$$

$$- (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger [\mathbf{T}^x \mathbf{Q}^y + \mathbf{T}^y \mathbf{Q}^x] \quad (36)$$

Therefore, in order to obtain the second derivative of the ESPF operator, we need to compute ϕ_A^{xy} , $V_{k,\mu\nu}^{xy}$ and \mathbf{T}^{xy} in addition to the first derivatives and the zero-order quantities.

Acknowledgement The authors acknowledge financial support by the ‘‘Agence Nationale pour la Recherche’’ through the project BIOMAGNET (ANR-16CE29-0008-01). ‘‘Centre de Calcul Intensif d’Aix-Marseille’’ is acknowledged for granting access to its high performance computing resources.

References

- (1) Pulay, P. Analytical derivatives, forces, force constants, molecular geometries, and related response properties in electronic structure theory. *WIREs: Comput. Molec. Sci.* **2014**, *4*, 169–181.
- (2) Pulay, P. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. *Mol. Phys.* **1969**, *17*, 197–204.
- (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. Molec. Biol.* **1976**, *103*, 227–249.
- (4) 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*, 700–733.
- (5) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.
- (6) Nobelprize.org, The Nobel Prize in Chemistry 2013 - Advanced Information. www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/advanced.html.
- (7) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Mol. Phys.* **2002**, *100*, 1813–1828.
- (8) Sinicropi, A.; Basosi, R.; Olivucci, M. Recent applications of a QM/MM scheme at the CASPT2//CASSCF/AMBER (or CHARMM) level of theory in photochemistry and photobiology. *J. Phys. Conf. Ser.* **2008**, *101*, 012001.
- (9) Ferré, N.; Ángyán, J. G. Approximate electrostatic interaction operator for QM/MM calculations. *Chem. Phys. Lett.* **2002**, *356*, 331–339.
- (10) Melaccio, F.; Olivucci, M.; Lindh, R.; Ferré, N. Unique QM/MM potential energy surface exploration using microiterations. *Int. J. Quant. Chem.* **2011**, *111*, 3339–3346.
- (11) Zhang, R.; Lev, B.; Cuervo, J. E.; Noskov, S. Y.; Salahub, D. R. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Brändas, E., Eds.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 353 – 400.

- (12) Tu, Y.; Laaksonen, A. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Brändas, E., Eds.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 1 – 15.
- (13) Barone, V.; Biczysko, M.; Brancato, G. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Brändas, E., Eds.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 17 – 57.
- (14) Carnimeo, I.; Cappelli, C.; Barone, V. Analytical gradients for MP2, double hybrid functionals, and TD-DFT with polarizable embedding described by fluctuating charges. *J. Comput. Chem.* **2015**, *36*, 2271–2290.
- (15) Mroginiski, M. A. *Encyclopedia of Biophysics*; Springer Berlin Heidelberg, 2013; pp 2149–2154.
- (16) Dapprich, S.; Komáromi, I.; Byun, K.; Morokuma, K.; J, F. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Molec. Struct.: THEOCHEM* **1999**, *461-462*, 1–21.
- (17) Cui, Q.; Karplus, M. Molecular properties from combined QM/MM methods. I. Analytical second derivative and vibrational calculations. *J. Chem. Phys.* **2000**, *112*, 1133–1149.
- (18) Cui, Q.; Karplus, M. Molecular Properties from Combined QM/MM Methods. 2. Chemical Shifts in Large Molecules. *J. Phys. Chem. B* **2000**, *104*, 3721–3743.
- (19) Ghysels, A.; Woodcock, H. L.; Larkin, J. D.; Miller, B. T.; Shao, Y.; Kong, J.; Neck, D. V.; Speybroeck, V. V.; Waroquier, M.; Brooks, B. R. Efficient Calculation of QM/MM Frequencies with the Mobile Block Hessian. *J. Chem. Theor. Comput.* **2011**, *7*, 496–514.
- (20) Nakata, H.; Nagata, T.; Fedorov, D. G.; Yokojima, S.; Kitaura, K.; Nakamura, S. Analytic second derivatives of the energy in the fragment molecular orbital method. *J. Chem. Phys.* **2013**, *138*, 164103.
- (21) Sakai, S.; Morita, S. Ab Initio Integrated Multi-Center Molecular Orbitals Method for Large Cluster Systems: Total Energy and Normal Vibration. *J. Phys. Chem. A* **2005**, *109*, 8424–8429.
- (22) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. Enabling ab initio Hessian and frequency calculations of large molecules. *J. Chem. Phys.* **2008**, *129*, 234101.
- (23) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of IR Spectra with a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495–5507.
- (24) Schwinn, K.; Ferré, N.; Huix-Rotllant, M. Analytic QM/MM atomic charge derivatives avoiding the scaling of coupled perturbed equations with the MM subsystem size. *J. Chem. Phys.* **2019**, *151*, 041102.
- (25) Müller, P.; Bouly, J.-P.; Hitomi, K.; Balland, V.; Getzoff, E. D.; Ritz, T.; Brettel, K. ATP Binding Turns Plant Cryptochrome Into an Efficient Natural Photoswitch. *Sci. Rep.* **2014**, *4*, 5175.
- (26) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. Derivative studies in Hartree-Fock and Møller-Plesset theories. *Int. J. Quant. Chem.* **1979**, *16*, 225–241.
- (27) Kaledin, A. L.; Kaledin, M.; Bowman, J. M. All-Atom Calculation of the Normal Modes of Bacteriorhodopsin Using a Sliding Block Iterative Diagonalization Method. *J. Chem. Theor. Comput.* **2006**, *2*, 166–174.

- (28) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. Derivative studies in Hartree-Fock and Moller-Plesset theories. *Int. J. Quant. Chem.: Quant. Chem. Symp.* **1979**, *13*, 225–241.
- (29) McWeeny, R. Some Recent Advances in Density Matrix Theory. *Rev. Mod. Phys.* **1960**, *32*, 335–369.
- (30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Menucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Revision A.01. 2016; Gaussian Inc. Wallingford CT.
- (31) Ponder, J. W. TINKER: Software tools for molecular design, version 6.3. 2010; <http://dasher.wustl.edu/tinker/>.
- (32) Baiardi, A.; Bloino, J.; Barone, V. General Time Dependent Approach to Vibronic Spectroscopy Including Franck-Condon, Herzberg-Teller, and Duschinsky Effects. *J. Chem. Theor. Comput.* **2013**, *9*, 4097–4115.
- (33) Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Ann. Rev. Phys. Chem.* **2012**, *63*, 287–323.
- (34) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (35) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (36) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (37) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (38) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
- (39) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283 – 290.
- (40) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
- (41) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S.

- Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2018), AMBER 2018, University of California, San Francisco.
- (42) Antony, J.; Medvedev, D.; Stuchebukhovich, A. Theoretical study of electron transfer between the photolyase catalytic cofactor FADH(-) and DNA thymine dimer. *J. Am. Chem. Soc.* **2000**, *122*, 1057.
- (43) Conrad, K. S.; Manahan, C. C.; Crane, B. R. Photochemistry of flavoprotein light sensors. *Nat. Chem. Biol.* **2014**, *10*, 801–809.
- (44) Kabir, M. P.; Orozco-Gonzalez, Y.; Gozem, S. Electronic spectra of flavin in different redox and protonation states: a computational perspective on the effect of the electrostatic environment. *Phys. Chem. Chem. Phys.* **2019**, *21*, 16526–16537.
- (45) Mondal, P.; Schwinn, K.; Huix-Rotllant, M. Impact of the redox state of flavin chromophores on the UV-Vis spectra, redox and acidity constants and electron affinities. *J. Photochem. Photobiol. A: Chem.* **2020**, *387*, 112164.
- (46) Klaumünzer, B.; Kröner, D.; Saalfrank, P. (TD-)DFT Calculation of Vibrational and Vibronic Spectra of Riboflavin in Solution. *J. Phys. Chem. B* **2010**, *114*, 10826–10834.
- (47) Holden, Z. C.; Rana, B.; Herbert, J. M. Analytic gradient for the QM/MM-Ewald method using charges derived from the electrostatic potential: Theory, implementation, and application to ab initio molecular dynamics simulation of the aqueous electron. *J. Chem. Phys.* **2019**, *150*, 144115.
- (48) Kalstein, A.; Fernández-Alberti, S.; Bastida, A.; Soler, M. A.; Farag, M. H.; Zúñiga, J.; Requena, A. Vibrational dynamics of polyatomic molecules in solution: assignment, time evolution and mixing of instantaneous normal modes. *Theor. Chem. Acc.* **2011**, *128*, 769–782.
- (49) Nutt, D. R.; Meuwly, M. Theoretical Investigation of Infrared Spectra and Pocket Dynamics of Photodissociated Carbonmonoxy Myoglobin. *Biophys. J.* **2003**, *85*, 3612–3623.
- (50) Mackie, C. J.; Candian, A.; Huang, X.; Maltseva, E.; Petrigiani, A.; Oomens, J.; Buma, W. J.; Lee, T. J.; Tielens, A. G. G. M. The anharmonic quartic force field infrared spectra of three polycyclic aromatic hydrocarbons: Naphthalene, anthracene, and tetracene. *J. Chem. Phys.* **2015**, *143*, 224314.

Graphical TOC Entry

