JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Damped Response Theory in Combination with Polarizable Environments: The Polarizable Embedding Complex Polarization Propagator Method

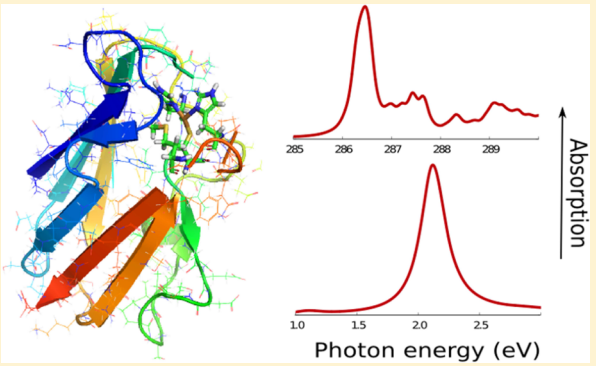
Morten N. Pedersen,*[,†,‡] Erik D. Hedegård,[†] Jógvan Magnus H. Olsen,[†] Joanna Kauczor,[‡] Patrick Norman,[‡] and Jacob Kongsted[†]

[†]Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark

[‡]Department of Physics, Chemistry and Biology, Linköping University, Linköping SE-58183, Sweden

Supporting Information

ABSTRACT: We present a combination of the polarizable embedding (PE) scheme with the complex polarization propagator (CPP) method with the aim of calculating response properties including relaxation for large and complex systems. This new approach, termed PE-CPP, will benefit from the highly advanced description of the environmental electrostatic potential and polarization in the PE method as well as the treatment of near-resonant effects in the CPP approach. The PE-CPP model has been implemented in a Kohn−Sham density functional theory approach, and we present pilot calculations exemplifying the implementation for the UV/vis and carbon *K*-edge X-ray absorption spectra of the protein plastocyanin. Furthermore, technical details associated with a PE-CPP calculation are discussed.



1. INTRODUCTION

Today's advancements of nanotechnologies such as photonics, electronics, light harvesting, light-emitting diodes, drug delivery, sensors, surface coating, catalysis, etc.[1−3] have encouraged the interplay between experiment and theory with the ultimate aim of manipulating and controlling molecular materials on an atomistic scale. On the experimental side, there exist a large number of spectroscopic techniques used for characterization such as, magnetic resonance spectroscopy, infrared absorption and Raman spectroscopies, visible/ultraviolet (UV/vis) absorption and fluorescence spectroscopies, circular dichroism, and near-edge X-ray absorption fine structure spectroscopy (NEXAFS).[4] All the above-mentioned spectroscopies can, from a theoretical point of view, be treated within the framework of response theory.[5] Traditionally, response theory formulations have been based on the Ehrenfest formulation,[6] but the more general quasi-energy formulation has recently been adopted.[7−9] Standard response theory is known to break down in the resonant or near-resonant part of the spectrum, which can be traced to the assumption that the excited states have an infinite lifetime. A solution can be to describe the near-resonant region of a molecular property in a rigorous way, such that the finite lifetime of the excited state is taken into account.[10] Response theory taking relaxation into account, also known as *damped* response theory, is thus well-defined in both the resonant and nonresonant parts of a spectrum.[11,12] The implications of damped response theory in spectroscopies within the high-energy regime are especially promising, since it allows for a direct evaluation of the relevant part of the spectrum without having to address the issues of resolving embedded states in high-density regions of the spectrum. The response formalism relies on a parametrization of the electronic wave function and is thus bound by the same restrictions upon size of the system as the underlying electronic structure theory, meaning that, in most applications, first principle calculations of large biological systems are beyond reach. One way to deal with this problem is to consider hybrid models combining a quantum mechanical (QM) method with a less accurate molecular mechanics (MM) description for the remaining part, known as a QM/MM approach.[13] In recent years, QM/MM methods have been used to study several spectroscopic details of solute−solvent systems as well as proteins.[14−19] However, the accuracy of standard force fields is usually not sufficient for accurate calculations of molecular properties, and effects of polarization of the MM region by the QM region is often completely neglected. To bring theory and experiment in even closer agreement, we present a model combining damped response theory with a hybrid QM/MM method. The great advantage of this combination is that we will be able to calculate theoretical spectra in a fraction of the time it will usually take to do the corresponding experiment. In detail, we will combine the complex polarization propagator (CPP) approach[11,12,20] with the polarizable embedding (PE)

Received: October 29, 2013
Published: January 22, 2014

   dx.doi.org/10.1021/ct400946k | *J. Chem. Theory Comput.* 2014, 10, 1164−1171

model[21,22] including polarization fully self-consistently and charge densities described by multipole expansions. We coin this new method the *polarizable embedding complex polarization propagator method* (PE-CPP). Earlier work along this line include the DIM/QM model by Morton and Jensen.[23] As an example, we will present the UV/vis spectrum of plastocyanin along with its carbon near-edge X-ray absorption fine structure spectrum. Plastocyanin is involved in the electron transfer process between photosystems I and II, and it has therefore been studied intensively with a range of spectroscopic and theoretical methods.[24,25] The intense signature band at 16700 cm^{-1} gives rise to the blue color and is known to have an origin in a $S_\pi \rightarrow d_{x^2-y^2}$ ligand to metal charge transfer (LMCT) excitation while a weaker band at 12800 cm^{-1} is also observed and assigned to a $d - d$ ligand field excitation.[26,27] Previous theoretical studies have dealt with both the UV/vis and NEXAFS spectra, also within a QM/MM framework.[28−30] This paper is structured as follows: In sections 2.1 and 2.2, we present the working equations for the CPP and the PE method, respectively, while in section 2.3 the equations coupling CPP and PE are presented. In sections 3−6, we discuss in detail the work flow and issues needed to be addressed in connection with PE-CPP calculations. Some of these issues are common for all QM/MM and embedding methods, such as the problem of defining the size of a region, which is treated with a quantum mechanical description. Finally, in section 7, we present the computed UV/vis and X-ray spectra of plastocyanin.

2. DAMPED RESPONSE THEORY IN A POLARIZABLE ENVIRONMENT

Throughout this paper, we will be working in a Kohn−Sham density functional theory formalism. We therefore adopt a single determinant reference state and parametrize it as[6]

$$|\psi(t)\rangle = e^{i\hat{\kappa}(t)}|0\rangle \tag{1}$$

where $|0\rangle$ is an SCF optimized reference state and the time-dependent Hermitian operator $\hat{\kappa}(t)$ is

$$\hat{\kappa}(t) = \sum_n [\kappa_n(t)\hat{q}_n^\dagger + \kappa_n^*(t)\hat{q}_n] \tag{2}$$

where the $\hat{q}_n^\dagger$ operator is defined in terms of creation and annihilation operators, $\hat{q}_n^\dagger = \hat{a}_s^\dagger \hat{a}_i$ and the sum runs over all virtual and occupied orbitals. In the following subsections, we briefly review the working equations of both the CPP approach and PE model. Finally, we move on to describe the coupling between damped response theory and PE.

2.1. Complex Polarization Propagator. The equation governing the time evolution of a quantum chemical system including relaxation of excited states can be written as[11]

$$\frac{\partial}{\partial t}\langle\psi(t)|\hat{t}_n^\dagger|\psi(t)\rangle - \left\langle \psi(t)\middle|\frac{\partial}{\partial t}\hat{t}_n^\dagger\middle|\psi(t)\right\rangle = \frac{1}{i\hbar}\langle\psi(t)|[\hat{t}_n^\dagger, \hat{H}]|\psi(t)\rangle - \gamma_n\langle\psi(t)|\hat{q}_n^\dagger|\psi(t)\rangle \tag{3}$$

under the assumption that there is no population of the electronic excited states at thermal equilibrium and defining $\hat{t}_n^\dagger = \exp(i\hat{\kappa})\hat{q}_n^\dagger\exp(-i\hat{\kappa})$. The Hamiltonian of eq 3 is the sum of the time-independent Hamiltonian $\hat{H}_0$ and a time-dependent perturbation $\hat{V}(t)$. In order to determine the time-evolution of the wave function, the time-dependent parameters from eq 2 are expanded in orders of the perturbation ($\kappa(t) = \kappa^{(1)}(t) + \cdots$) with Fourier components defined as

$$\kappa_n^{(1)}(t) = \int \kappa_n^\omega \exp(-i\omega t)\mathrm{d}\omega \tag{4}$$

Inserting $\kappa(t)$ into eq 3 along with the parametrized reference state of eq 1 and afterward expanding eq 3 using the Baker−Campbell−Hausdorff (BCH) expansion leads to a first order correction to the wave function parameters that may be determined from

$$\kappa^\omega = i[\mathbf{E}^{[2]} - \hbar\omega\mathbf{S}^{[2]} - i\hbar\mathbf{R}^{[2]}]^{-1}\mathbf{V}^{\omega[1]} \tag{5}$$

where $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ are the Hessian and the generalized overlap matrix,[6] respectively. $\mathbf{R}^{[2]}$ is the relaxation matrix,[11] and $\mathbf{V}^{\omega[1]} = \langle 0|[\hat{q},\hat{V}^\omega]|0\rangle$ is the property gradient. The time-dependent expectation value of a general operator $\hat{A}$ with an approximate reference state in eq 1 is given by

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle 0|e^{-i\hat{\kappa}(t)}\hat{A}e^{i\hat{\kappa}(t)}|0\rangle \tag{6}$$

The definition of the response functions now follows by using a BCH expansion in eq 6 and the linear response function can be identified as[12]

$$\langle\langle\hat{A};\hat{V}^\omega\rangle\rangle = -i\mathbf{A}^{[1]\dagger}\kappa^\omega \tag{7}$$

where the parameters κ^ω are determined from eq 5 and $\mathbf{A}^{[1]} = \langle 0|[\hat{q},\hat{A}]|0\rangle$ is a property gradient. Since the response vectors are generally complex (as a consequence of the relaxation matrix), this response function is complex. In order to evaluate eq 7, the linear response equation in eq 5 first needs to be solved. This can, in principle, be done by direct inversion, but due to the large dimensions of the Hessian, overlap, and relaxation matrices, it is usually solved using iterative methods.[20]

2.2. Polarizable Embedding Operators. In the polarizable embedding model,[21,22] the total energy is given as the sum of the energy from the part defining the quantum chemical system and the interactions with the environment. What separates the PE model from other QM/MM models, is the way the environment interacts with the quantum chemical system. The energy expression E_{PE} is given by

$$E_{PE} = E_{es} + E_{ind} \tag{8}$$

We will in general treat the permanent charge distribution of the environment by assigning distributed multipole expansions with expansion centers usually taken at the atomic nuclei that define the environment. The induced charge distribution in the environment is described by distributed polarizabilities that give rise to induced dipoles. Thus, the electrostatic interaction energy between an electronic charge distribution described by quantum mechanics (in terms of the molecular orbitals) and the permanent multipole moments defining the environment is given by

$$E_{es} = \sum_s \left(\sum_{|k|=0}^{K} \frac{(-1)^{|k|}}{k!} M_s^{(k)} \sum_n Z_n T_{sn}^{(k)} \right) + \langle 0|\hat{V}^{es}|0\rangle \tag{9}$$

Furthermore, the electronic interaction operator $\hat{V}^{es}$ is expressed as a sum over sites, s, in the environment

$$\hat{V}^{es} = \sum_s \sum_{pq} \nu_{s,pq}^{es}\hat{E}_{pq} \tag{10}$$

$$\nu_{s,pq}^{es} = \sum_{|k|=0}^{K} \frac{(-1)^{|k|}}{k!} M_s^{(k)} t_{s,pq}^{(k)} \tag{11}$$

dx.doi.org/10.1021/ct400946k | J. Chem. Theory Comput. 2014, 10, 1164−1171

The sum over $|k|$ in eqs 9 and 10 represents a compact notation[22,31] of a multipole expansion, where each site s has assigned multipole moments $M_s^{(k)}$ of order $|k|$. The $t_{s,pq}^{(k)}$ integrals in eq 11 are of the type

$$t_{s,pq}^{(k)} = -\langle \phi_p | T_s^{(k)}(\mathbf{r}) | \phi_q \rangle \tag{12}$$

$$T_s^{(k)}(\mathbf{r}) = \nabla_{\mathbf{r}}^k \frac{1}{|\mathbf{r} - \mathbf{r}_s|} \tag{13}$$

where the $T_s^{(k)}$ is a general interaction tensor describing the mutual interaction between sites in the environment[32] and $|\phi_p\rangle$ is the p-th molecular orbital. Next, the induction energy arising from the polarization of the environment, both internally and from the QM part of the system, is expressed through the field operator $\hat{\mathbf{F}}$

$$E_{\text{ind}} = -\frac{1}{2}\langle \hat{\mathbf{F}} \rangle^T \mathbf{A} \langle \hat{\mathbf{F}} \rangle \tag{14}$$

where we have used the short hand notation $\langle \hat{F} \rangle = \langle 0|\hat{F}|0\rangle$ and $\hat{F}$ is the combined field from (QM) nuclei and electrons, along with the field from the multipoles within the environment

$$\hat{\mathbf{F}} = \mathbf{F}^{\text{mul}} + \mathbf{F}^{\text{n}} + \hat{\mathbf{F}}^{\text{e}} \tag{15}$$

collected as a super vector with dimension $3N$, where N is the number of *polarizable* sites in the environment. Due to the method in which the site polarizabilities are constructed, polarizabilities within the same fragment are not allowed to polarize each other. The field due to the permanent charge distribution (represented by multipoles) of the environment is denoted by F^{mul} and the fields at site s' from QM nuclei and electrons are

$$\mathbf{F}_{s'}^{n} = \sum_M Z_M \mathbf{T}_{Ms'}^{(1)} \tag{16}$$

$$\hat{\mathbf{F}}_{s'}^{e} = -\sum_{pq} \mathbf{t}_{s',pq}^{(1)} \hat{E}_{pq} \tag{17}$$

Note that s' is used to denote a *polarizable* site. The symmetric classical response matrix $\boldsymbol{A}$ (also known as the *relay* matrix) in eq 14 is given as[33]

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{\alpha}_{11}^{-1} & -\mathbf{T}_{12}^{(2)} & \cdots & -\mathbf{T}_{1N}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \boldsymbol{\alpha}_{22}^{-1} & \cdots & -\mathbf{T}_{2N^{[2]}} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{N1}^{(2)} & -\mathbf{T}_{N2}^{(2)} & \cdots & \boldsymbol{\alpha}_{NN}^{-1} \end{pmatrix}^{-1} \tag{18}$$

where the diagonal contains the inverse polarizability tensors and the off-diagonal contains the dipole−dipole interaction tensor. Utilizing

$$\boldsymbol{\mu}^{\text{ind}} = \mathbf{A} \langle \hat{\mathbf{F}} \rangle \tag{19}$$

Equation 14 can be written directly in terms of the induced dipoles by

$$E_{\text{ind}} = -\frac{1}{2}(\boldsymbol{\mu}^{\text{ind}})^T \langle \hat{\mathbf{F}} \rangle \tag{20}$$

The PE model can be combined with DFT by means of an effective Kohn−Sham operator written as[21]

$$\hat{f}^{\text{KS}} = \hat{f}_{\text{vac}}^{\text{KS}} + \hat{v}_{\text{PE}} \tag{21}$$

which is used to build up the self-consistent Kohn−Sham equations. Here, $\hat{f}_{\text{vac}}^{\text{KS}}$ is the Kohn−Sham operator known from standard vacuum DFT, and $\hat{v}_{\text{PE}}$ is the PE potential operator. This operator can be derived as a variation of the total energy functional with respect to the density. Thereby, all terms not related to the electronic density vanish, and the only terms needed to be considered are the electronic parts of eqs 9 and 20. The PE potential operator thus reads

$$\begin{aligned} \hat{v}_{\text{PE}} &= \hat{V}^{\text{es}} + \hat{V}^{\text{ind}} \\ &= \sum_s \sum_{pq} v_{s,pq}^{\text{es}} \hat{E}_{pq} - \boldsymbol{\mu}^{\text{ind}} \hat{\mathbf{F}}^{\text{e}} \end{aligned} \tag{22}$$

This operator is used in the following to derive the damped response equations including a polarizable environment.

2.3. Polarizable Embedding in Linear Damped Response Theory. In the following section, we will give an overview of the equations associated with the coupling between the PE approach and damped response theory. Writing the electronic Hamiltonian as the sum of a time-independent part $\hat{H}_0$ and a time-dependent perturbation $\hat{V}(t)$ and inserting this into the Ehrenfest equation [eq 3], the following expression arises

$$\begin{aligned} &\frac{\partial}{\partial t}\langle \psi(t)|\hat{t}_n^\dagger|\psi(t)\rangle - \left\langle \psi(t)\middle|\frac{\partial}{\partial t}\hat{t}_n^\dagger\middle|\psi(t)\right\rangle \\ &= \frac{1}{i\hbar}[\langle \psi(t)|[\hat{t}_n^\dagger, \hat{H}_0 + \hat{V}(t)]|\psi(t)\rangle \\ &\quad + \langle \psi(t)|[\hat{t}_n^\dagger, \hat{v}_{PE}]|\psi(t)\rangle] - \gamma_n \langle \psi(t)|\hat{q}_n^\dagger|\psi(t)\rangle \end{aligned} \tag{23}$$

We note that the left-hand side and the first term on the right-hand side are equivalent to the vacuum equation.[34] The coupling between PE and standard response theory is obtained by also including the $\langle \psi(t)|[\hat{t}_n^\dagger, \hat{v}_{\text{PE}}]|\psi(t)\rangle$ in the $\mathbf{E}^{[2]}$ matrix and has been derived in the work by Olsen et al.[21] The extension to the CPP method is thus performed by straightforward replacement of $\mathbf{E}^{[2]}$ in eq 5 by

$$\mathbf{E}^{[2]} \rightarrow \mathbf{E}_{\text{tot}}^{[2]} = \mathbf{E}_{\text{vac}}^{[2]} + \mathbf{E}_{\text{PE}}^{[2]} \tag{24}$$

Note that, as a result of the density dependence of the second term in the PE potential operator in eq 22, the Hamiltonian of the system becomes nonlinear and an additional contribution to the Hessian matrix appears as

$$\mathbf{E}_{\text{PE}}^{[2]} = -\langle 0|[\hat{q}, \hat{Q}_1^{\omega} + \hat{Q}_2^{\omega}]|0\rangle \tag{25}$$

with

$$\begin{aligned} \hat{Q}_1^{\omega} &= \tilde{V}^{\text{es}}(\kappa^\omega) - \mathbf{A}\langle 0|\hat{F}|0\rangle \tilde{\mathbf{F}}^{\text{e}}(\kappa^\omega) \\ &= \tilde{V}^{\text{es}}(\kappa^\omega) - \boldsymbol{\mu}^{\text{ind}} \hat{\mathbf{F}}^{\text{e}}(\kappa^\omega) \end{aligned} \tag{26}$$

and

$$\begin{aligned} \hat{Q}_2^{\omega} &= -\mathbf{A}\langle 0|\tilde{\mathbf{F}}^{\text{e}}(\kappa^\omega)|0\rangle \hat{\mathbf{F}}^{\text{e}} \\ &= -\tilde{\boldsymbol{\mu}}^{\text{ind}}(\kappa^\omega) \hat{\mathbf{F}}^{\text{e}} \end{aligned} \tag{27}$$

where $\tilde{\mathbf{F}}^{\text{e}}$ is an one-index transformed electronic field operator.[35] The $\hat{Q}_1^{\omega}$ operator describes the PE contribution to the linear response function, corresponding to an environment where the excited states are described using the ground state polarization, while the $\hat{Q}_2^{\omega}$ operator describes the dynamic response of the environment due to the external perturbation.

dx.doi.org/10.1021/ct400946k | J. Chem. Theory Comput. 2014, 10, 1164−1171

The present implementation builds upon a newly developed PE library,[31] which handles the density-dependent operators by a density-driven formulation. Thus, the expectation value of one-index transformed field operator in eq 27 is for site s' described as

$$\langle 0|\tilde{\mathbf{F}}^{e}_{s'}(\kappa^{\omega})|0\rangle = \sum_{pq} D^{\omega}_{pq} t^{(1)}_{pq,s}, \tag{28}$$

in terms of the perturbed density matrix. With this, we have now identified the necessary equations to couple the PE method with the CPP approach for solving response equations and we move on to discussing the technical details associated with an actual calculation performed using the PE-CPP approach.

3. COMPUTATIONAL ISSUES

The work flow of a PE calculation can be split into four main steps:

1. Split the total system into a core region and an environmental region.
2. Determine distributed multipole moments and distributed polarizabilities for the region defining the environment.
3. Optimize the wave function (density) of the core region by including the polarizable embedding potential in the Fock or KS-operator.
4. Calculate the quantum mechanical properties of interest using response theory including explicit account of the environment.

The first step requires first an identification of the chemically interesting part of the system and a fragmentation of the surrounding medium. In the simple case, where the environment consists of only noncovalently bonded units, the fragmentation is usually straightforward: Each different solvent molecule is simply regarded as a fragment by itself. If, on the other hand, the environment consists of covalently bonded units the fragmentation is more elaborate. Here, we will only focus on the procedure used for polypeptides, comprising the backbone of protein systems. The fragmentation is for these systems carried out based on the molecular fractionation using conjugate caps (MFCC) approach,[36] which decomposes the system into amino acid fragments.

The second step takes as input all the generated fragments from step one and calculates distributed multipole moments and distributed polarizabilities. Currently this is processed using the LoProp method[37] as implemented in the MOLCAS program,[38] but other methods such as DMA[32] and MpProp[59] can be equally suitable. The potential is assembled and an exclusion list is constructed, making sure that polarizabilities within the same fragment do not interact, according to Söderhjelm and Ryde.[39]

The third step uses the distributed multipole moments and distributed polarizabilities to construct the polarizable embedding potential from eq 22, which is added to the vacuum KS operator, and the wave function is then optimized. Since the quantum mechanical and classical equations describing polarization are coupled, a double SCF procedure is performed at this point.

The fourth step uses the optimized wave function to calculate the properties of interest including an explicit account of the environment.

4. COMPUTATIONAL DETAILS

The PE and CPP models have both been implemented in a development version of the DALTON quantum chemistry program,[40] which was used to obtain all results presented in this work. All PE-CPP calculations utilize a standard long-range corrected CAMB3LYP functional[41] and a fully long-range corrected CAMB3LYP functional.[42] In all calculations, we used the Stuttgart effective core potential (SDD)[43] for the copper atom and the aug-cc-pVDZ[44] basis set for the remaining atoms in the QM region. The plastocyanin structure (1PLC) was obtained from the Protein Data Bank (PDB)[45] and can be seen in Figure 1. The active site (chosen as the core region for all PE-CPP calculations) comprises a cysteine, a methionine, and two histidine ligands and can be viewed in Figure 2. Each atom in the classical region was assigned multipole moments up to quadrupoles and anisotropic dipole−dipole polarizabilities using the procedure described in section 3. The induced dipoles are solved by using a parallelized iterative solver that is based on a combined Jacobi and Gauss−Seidel scheme. In the response calculations, the relaxation parameter in eq 23 was chosen as $\gamma = 1000\ \mathrm{cm}^{-1}$.

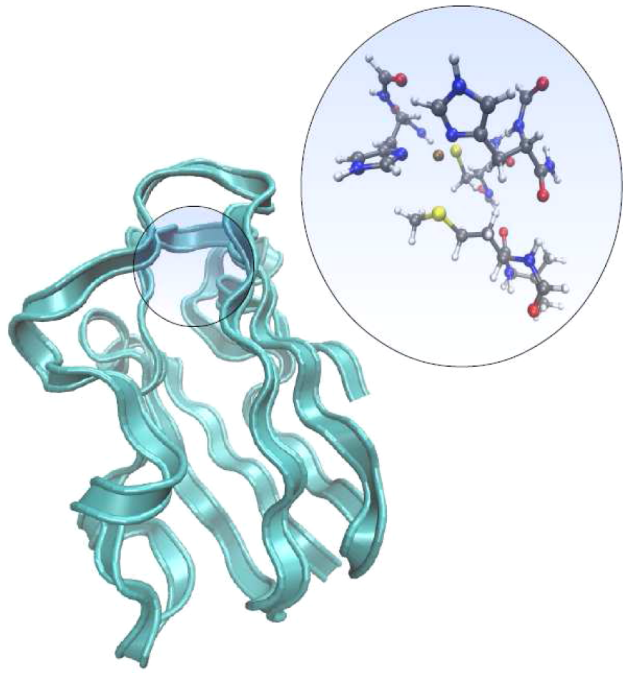


Figure 1. Plastocyanin and the active site used in the QM/MM geometry optimization.

5. PROTEIN GEOMETRY OPTIMIZATION

The plastocyanin protein used in this study contains in total 99 amino acids and the structure is known at an atomistic level (1.33 Å resolution) from X-ray diffraction analysis (PDB: 1PLC).[45] Hydrogen atoms where added to this structure using the protein preparation tool in the Maestro program which is a part of the Schrödinger software package.[46] All crystallographic water molecules were kept, and their orientations were sampled at neutral pH, using the exhaustive sampling option in the protein preparation wizard in Maestro.[46−49] Titratable residues were inspected and the following residues were found to have a negative charge (as assigned by the protein prepation wizard): Asp2, Asp8, Asp9, Glu18, Glu25, Asp42, Glu43, Asp44, Asp51,

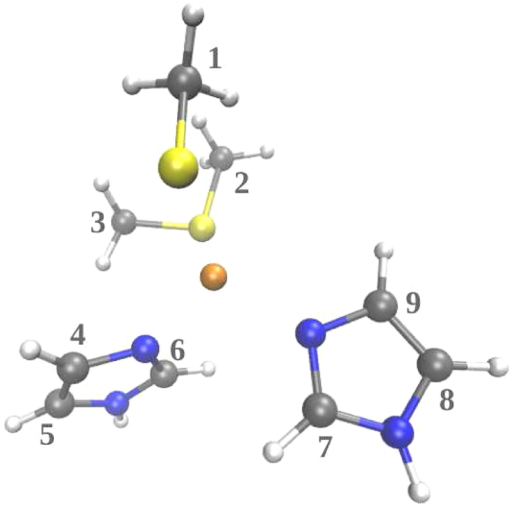


Figure 2. Active site of plastocyanin used in the QM/MM calculations of the UV/vis and NEXAFS spectra.

Glu59, Glu60, Asp61, Glu68, Glu71, and Glu79. A number of residues has been assigned charged $-NH_3^+$ groups; Ile1, Lys26, Lys30, Lys54, Lys66, Lys77, and Lys95 (as assigned by the preparation wizard). The above assignments were found consistent with the PROPKA procedure.[50] Thus, the total charge of the protein is −8, and all charged groups are at the protein surface, with the sole exception of Cys84, which coordinates to the Cu^{II} atom. For the QM/MM geometry optimizations the active site was chosen to consist of Cu^{II} and the four coordinating groups, His37, His87 (coordinates through N^{δ}) and Cys84, Met92, which coordinates through sulfur (see Figure 1). In all of the four coordinating groups, a part of the neighboring residues was also included in the QM system, which accordingly also contain the −CO group of Pro36, Pro86, Tyr83, and Gly91. Further, the QM system contains the −NH groups of Asn38, Gln88, and Ser85. The −CO and −NH groups were capped with hydrogen. For the residue Met92, the part included from the neighboring residue (Val93) in the QM region were slightly extended, such that this part contains $-NHCH(CH_3)_2$. This part was likewise capped with hydrogen. Explicitly, the QM/MM caps are between the atoms with entry 515−514, 544−545, 1279−1280, 1250−1249, 1208−1207, 1237−1238, 1338−1339, and 1315−1314 in the PDB file (the first numbers are QM atoms and the last are MM atoms). The geometry optimization on the final system comprising 98 atoms was performed at the B3LYP/6-31+G*[51−54] level of theory, using the QSite program.[55] The geometry of the environment was allowed to relax using the OPLS2005 force field[56] as implemented in the Schrödinger package. The optimized bond distances for the active site are given in Table 1. We observe that the calculated geometrical parameters deviate somewhat from the experimental values.

Table 1. Optimized Bond Distances for Plastocyanin[a]

bond	residue	calc.	exp.[b]
Cu−S	Met92	2.97	2.82
Cu−S	Cys84	2.19	2.07
Cu−N	His37	2.04	1.91
Cu−N	His87	2.00	2.06

[a]All distances are given in Å. [b]From ref 45.

However, we add here that the standard deviation associated with the Cu−ligand bond lengths in the experimental determined structure is estimated to be around 0.04 Å,[45] and another study has found that structural parameters similar to ours lead to better agreement between TD-DFT and experimental UV/vis spectra.[30] Furthermore, neglect of finite temperature effects will also contribute to discrepancies between the experimental and predicted spectroscopic parameters.

6. QM/MM BORDER

The distributed multipole moment scheme, used by some advanced force fields and also by the PE model, breaks down when the multipole moments or the polarizabilities are too close positioned to the QM region. This happens especially if the environment is covalently bonded to the QM region, as is often the case in proteins. In this section, we intend to test the sensitivity of the PE−CPP method for different ways of dealing with multipole moments and polarizabilities close to the QM region. We have considered two approaches to control the QM/MM border. The first is simply to remove all multipole moments and polarizabilities within a given distance of the QM region. The second approach redistributes all zero order multipole moments (charges) within a given distance of the QM region to the three nearest atom sites in the MM region and removes all higher order multipole moments and polarizabilities. As a reference for these tests, we used a full TD-DFT calculation of a cluster model consisting of the copper atom and coordinated amino acid residues capped with $-NHCH_3$ and $-COCH_3$. The remaining calculations utilizes the QM core seen in Figure 2, but with a smaller environmental part consisting of the remaining atoms from the reference calculation. The results of this analysis can be found in Table 2.

Table 2. LMCT Peak Position and Intensity Calculated for Different Distribution Schemes to Control the QM/MM Border Region

method	≈ LMCT peak position (cm^{-1})[a]	intensity
redist 1.2	19800	0.56
redist 1.4	24100	0.51
remove 1.2	19800	0.89
remove 1.4	22000	0.97
full QM	22000	0.66

[a]Peak positions based on a spline fit of six points (see Figure SI-1 in Supporting Information).

Here, we only consider the peak position and intensity of the LMCT transition. We consider two cutoff distances, 1.2 and 1.4 Å corresponding to removal/redistribution of 8 and 15 partial PE charges, respectively. The redistribution scheme shows a shift of approximately 4300 cm^{-1} when increasing the cutoff distance by 0.2 Å, while removing all multipole moments and polarizabilities shows a shift of approximately 2200 cm^{-1}. Both approaches, utilizing a cutoff distance of 1.2 Å, underestimate the full QM calculation by ca. 2200 cm^{-1}. While the shift going to a cutoff distance on 1.4 Å for the redistribution scheme is too large, the shift in the remove scheme reproduces the full QM calculation well. On the other hand, removing all multipole moments and polarizabilities produces in both cases a too intense peak, off by 0.31 and 0.23, respectively, while the redistribution scheme is much better, differing only by 0.15 and 0.10, respectively, from the full QM calculation. In conclusion,

we choose in the following to remove all parameters within a threshold of 1.4 Å.

7. RESULTS AND DISCUSSION

In this section, we demonstrate the use of the PE-CPP approach by calculating absorption spectra of plastocyanin. First, we present the results from the UV/vis spectrum, and second, we move on to the carbon *K*-edge NEXAFS spectrum. In both situations, it is possible to characterize the spectrum by a detailed analysis of the imaginary response vectors.[42,57] We will stress that the calculation of absorption spectra requires no explicit reference to the excited state and, thus, the calculation time does not increase with the number of excited states present in the user-defined frequency domain.

7.1. Protein UV/Vis Spectrum. Figure 3a shows the computed UV/vis spectrum of plastocyanin at the PE-TD-DFT level. Two transitions are observed, one at 19200 cm^{-1} and one at 8200 cm^{-1}. An analysis of the response vectors confirm that the two transitions are a $S(cys) \rightarrow d$ and a $d-d$ transition, respectively. The general features of the experimental spectra are thus well captured by the computational approach, and while our results are in good agreement with previous results obtained at the TD-DFT level by Robinson and Besley,[30] both peaks are perceptibly shifted compared to the experimental values. To demonstrate the environmental effects, we calculated the spectra of the active site in isolation, the active site surrounded by only the copper coordinated amino acid residues capped with $-NHCH_3$ and $-COCH_3$ and the active site in the full protein environment. The results are shown in Figure 3b, where we immediately observe a large shift of the intense LMCT transition while the less intense $d-d$ transition is only affected by including a minimum of the protein environment. Previous calculations of X-ray data using TD-DFT/CAM-B3LYP has shown that predicting X-ray data benefits from a parametrization of the CAMB3LYP functional with a correct asymptotic limit of the Coulomb hole−electron interaction (α = 0.19, β = 0.81, and μ = 0.33).[42] In Figure 3a, we compare the spectrum obtained with the modified CAMB3LYP functional versus the spectrum obtained using standard CAMB3LYP (α = 0.19, β = 0.46, and μ = 0.33). Contrary to X-ray spectra (as discussed in the next section), using the modified CAMB3LYP functional in the UV/vis region shifts the LMCT transition to higher energy and thus a slightly worse spectrum compared to experiment is obtained, but the general features of the spectrum still remains.

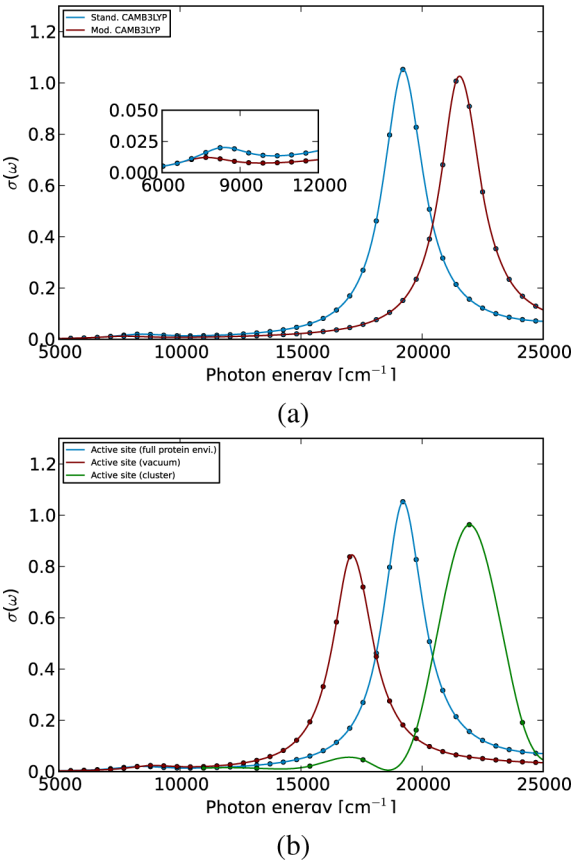


Figure 3. Computed UV/vis spectra of the plastocyanin complex; (a) using standard CAMB3LYP[41] and a fully long-range corrected CAMB3LYP functional,[42] (b) using different sizes of active site region and the standard CAMB3LYP functional.

7.2. NEXAFS Spectrum. In this section, we present the calculated carbon *K*-edge NEXAFS spectrum of plastocyanin, see Figure 4. The spectrum has been calculated using the CPP approach, both with and without PE, and utilizing a fully long-range corrected CAMB3LYP functional (α = 0.19, β = 0.81, and μ = 0.33). The spectrum was shifted 10 eV to account for the self-interaction error associated with the use of standard exchange-correlation functionals in Kohn−Sham DFT.[58] The characterization of the spectrum is in short summarized in Table 3. The spectrum comprises four peaks marked as A−D (see Figure 4). The energy region covered by peaks A−D spans some 2.6 and 3.2 eV, when using PE-CPP and CPP, respectively. When analyzing the complex response vectors, we find that the absorption in peak A is due to the histidine carbon atoms (C4−C6 and C7−C9, see Figure 2 for labeling). Peak B is assigned to the cystine and methionine carbon atoms, while peak C and D can be assigned to C1−C3, C4, C9 and C1−C3, C4−C5, C8−C9 respectively. Comparing the vacuum CPP with the PE-CPP calculations, we observe that the general characteristics of the spectrum are preserved, while the shift of the peak positions are in the order 0.2−0.6 eV. Considering the range of the vacuum spectrum of 3.2 eV a shift of 0.6 eV ($\approx$ 20%) is large enough to indicate that environmental effects are also important in the X-ray region of the spectrum and not only the in the UV/vis part.

8. CONCLUSIONS

We have derived and implemented a coupling between the polarizable embedding model and the complex polarization propagator approach. The coupling is straightforwardly implemented, as it only requires a modification to the Hessian matrix. To promote the capabilities of this new method, PE-CPP, we presented the UV/vis and carbon *K*-edge NEXAFS spectra of plastocyanin. The modeling of these spectra follow a specific work flow. First is the choice of the core region and the environmental region. When dealing with complexes where these two parts are covalently bonded, the issue of the border region arises. We have here presented a few possibilities of redistributing the multipole parameters present in the border region. Other schemes are currently under investigation in our group. The second step is the calculation of distributed multipole moments and distributed polarizabilities, where we have used the LoProp method. In the UV/vis region, we capture the general aspects of the spectrum, and as a result of the nature of standard KS-DFT, there is a need to apply an

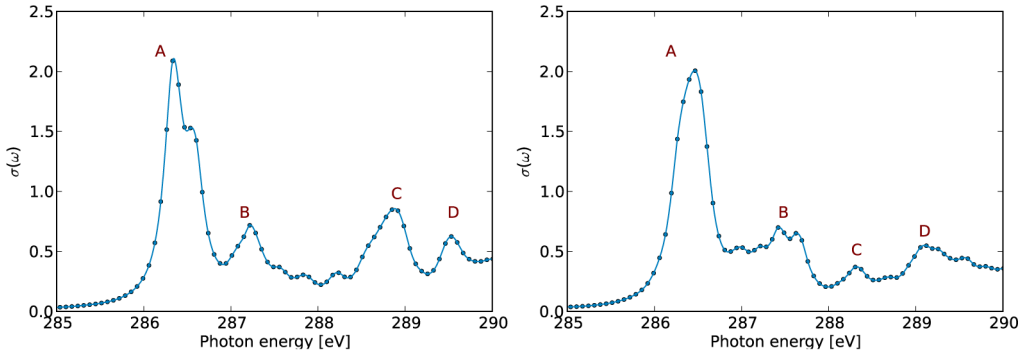


Figure 4. (a) NEXAFS spectrum of plastocyanin computed without PE, (b) NEXAFS spectrum of plastocyanin computed with PE-CPP. Both spectra have been shifted 10 eV.

Table 3. Assignment of Calculated Carbon *K*-edge NEXAFS Spectrum from Figure 4

peak		position (max abs.)/ [eV]	assignment[a]
A	PE-CPP	286.5	C4−C6, C7−C9
	vacuum (CPP)	286.3	
B	PE-CPP	287.4	C1−C3
	vacuum (CPP)	287.2	
C	PE-CPP	288.3	C1−C3, C4, C9
	vacuum (CPP)	288.9	
D	PE-CPP	289.1	C1−C3, C4−C5, C8−C9
	vacuum (CPP)	289.5	

[a]The numbering of carbon atoms refer to Figure 2.

overall spectral shift to account for the inherent self-interaction error.[58] Recently, the PE model has been formulated using a multi configurational self-consistent field (MCSCF) type of wave function,[60] and we intend to make the coupling to CPP at this level. We also showed the carbon *K*-edge NEXAFS spectrum of plastocyanin with and without the PE model to study the effects of the environment in this energy region. The main details of the spectrum are preserved even if the environmental part is not considered, but we do observe shifts in peak positions, indicating that environmental effects are needed for accurate modeling in the X-ray region. In conclusion, we presented a method that is able to generate good and reliable spectra in all regions of the energetic spectrum, while including environmental effects via permanent multipoles and distributed polarizabilities. We believe that this model can be used to complement experimental spectra and the characterization of these, especially when considering large biologically related complexes.

■ ASSOCIATED CONTENT

Supporting Information

More extensive figure for "QM/MM border" used for estimating the max absorption peak. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mortennp@sdu.dk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Danish Center for Scientific Computing. M.N.P. thanks the Augustinus foundation for financial support. J.K. thanks the Danish Councils for Independent Research (Sapere Aude), the Lundbeck Foundation, and the Villum foundation for financial support. P.N. acknowledges financial support from the Swedish Research Council (Grant No. 621-2010-5014).

■ REFERENCES

(1) *Handbook of Advanced Electronic and Photonic Materials and Devices*; Nalwa, H. S., Ed.; Academic Press: San Diego, 2001.
(2) *Molecular Nanoelectronics*; Reed, M. A. , Lee, T., Eds.; American Scientific Publishers: Valencia, CA, 2003.
(3) *Functional Nanomaterials*; Geckeler, K. E., Rosenberg, E., Eds.; American Scientific Publishers: Valencia, CA, 2005.
(4) *Encyclopedia of Spectroscopy and Spectrometry*; Lindon, J. C., Tranter, G. E., Holmes, J. L., Eds.; Academic Press: San Diego, 2010.
(5) Norman, P. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519−20535.
(6) Olsen, J.; Jørgensen, P. *J. Chem. Phys.* **1985**, *82*, 3235−3264.
(7) Christiansen, O.; Jørgensen, P.; Hättig, C. *Int. J. Quantum Chem.* **1998**, *68*, 1−52.
(8) Kristensen, K.; Kauczor, J.; Kjærgaard, T.; Jørgensen, P. *J. Chem. Phys.* **2009**, *131*, 044112.
(9) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. *Chem. Rev.* **2012**, *112*, 543−631.
(10) Boyd, R. W. *Nonlinear Optics*, 3rd ed.; Academic Press: San Diego, 2008.
(11) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. *J. Chem. Phys.* **2005**, *123*, 194193.
(12) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. *J. Chem. Phys.* **2001**, *115*, 10323.
(13) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(14) Curutchet, C.; Munoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. *J. Chem. Theory Comput.* **2009**, *5*, 1838−1848.
(15) Lipparini, F.; Cappelli, C.; Scalmani, S.; De Mitri, N.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4270−4278.
(16) Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4153−4166.
(17) Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Phys.* **2013**, *138*, 234108.
(18) Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2012**, *8*, 4527−4538.
(19) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(20) Kauczor, J.; Jørgensen, P.; Norman, P. *J. Chem. Theory Comput.* **2011**, *7*, 1610−1630.

(21) Olsen, J. M.; Aidas, K.; Kongsted, J. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.
(22) Olsen, J. M.; Kongsted, J. *Adv. Quantum Chem.* **2011**, *61*, 107−143.
(23) Morton, S. M.; Jensen, L. *J. Chem. Phys.* **2011**, *135*, 134103.
(24) Solomon, E. I.; Szilagyi, R. K.; DeBeer George, S.; Basumallick, L. *Chem. Rev.* **2004**, *104*, 419−458.
(25) Solomon, E. I.; Szilagyi, R. K.; DeBeer George, S.; Basumallick, L. *Coord. Chem. Rev.* **2011**, *255*, 774−789.
(26) Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1988**, *110*, 3811−3819.
(27) Pierloot, K.; Kerpel, J. O. A. D.; Ryde, U.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 218−226.
(28) Monari, A.; Very, T.; Rivail, J.-L.; Assfeld, X. *Comp. Theor. Chem.* **2012**, *990*, 119−125.
(29) Sinnecker, S.; Neese, F. *J. Comput. Chem.* **2006**, *27*, 1463−1475.
(30) Robinson, D.; Besley, N. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9667−9676.
(31) Olsen, J. M. H. Why should one use the polarizable embedding model for computational studies of photoactive proteins? Ph.D. Dissertation, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense M, Denmark, 2012; DOI: 10.6084/m9.figshare.156851
(32) Stone, A. J. *The theory of Intermolecular Forces*; Oxford University Press, Oxford, U.K., 1997.
(33) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952−2960.
(34) Salek, P.; Vahtras, O.; Helgaker, T.; Ågren, H. *J. Chem. Phys.* **2002**, *117*, 9630−9645.
(35) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory*; Wiley: New York, 2000.
(36) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599−3605.
(37) Gagliardi, L.; Lindh, R.; Karlström, G. *J. Chem. Phys.* **2004**, *121*, 4494−4500.
(38) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-A.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224−247.
(39) Söderhjelm, P.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 617−627.
(40) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. Aa.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Salek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs Comput Mol Sci.* **2013**, DOI: 10.1002/wcms.1172.
(41) Yanai, T.; Tew, D.; Handy, N. *Chem. Phys. Lett.* **2004**, *393*, 51−57.
(42) Ekström, U.; Norman, P. *Phys. Rev. A* **2006**, *74*, 042722.
(43) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866−872.
(44) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(45) Guss, J. M.; Bartunik, H. D.; Freeman, H. C. *Acta Cryst. B* **1992**, *48*, 790−811.
(46) *Maestro*, version 9.2; Schrödinger, LLC: New York, 2013.
(47) *Epik*, version 0.2; Schrödinger, LLC: New York, 2013.
(48) *Impact*, version 5.7; Schrödinger, LLC: New York, 2013.
(49) *Prime*, version 3.0; Schrödinger, LLC: New York, 2013.
(50) Olsson, M. H. M.; Søndergard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(51) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(52) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.
(53) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. *J. Comput. Chem.* **1983**, *4*, 294−301.
(54) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(55) *Qsite*, version 5.7; Schrödinger, LLC: New York, 2013.
(56) Banks, J. L.; et al. *J. Comput. Chem.* **2005**, *26*, 1752−1780.
(57) Ekström, U.; Norman, P.; Carravetta, V.; Ågren, H. *Phys. Rev. A* **2006**, *73*, 022506.
(58) Tu, G.; Rinkevicius, Z.; Vahtras, O.; Ågren, H.; Ekström, U.; Norman, P.; Carravetta, V. *Phys. Rev. A* **2007**, *76*, 022506.
(59) Söderhjelm, P.; Krogh, J. W.; Karlström, G.; Ryde, U.; Lindh, R. *J. Comput. Chem.* **2007**, *28*, 1083−1090.
(60) Hedegård, E. D.; List, N. H.; Jensen, H. J. Aa.; Kongsted, J. *J. Chem. Phys.* **2013**, *139*, 044101.