PERSPECTIVE

Excited states in large molecular systems through polarizable embedding

Nanna Holmgaard List,*† Jógvan Magnus Haugaard Olsen and Jacob Kongsted*

In this perspective, we provide an overview of recent work within the polarizable embedding scheme to describe properties of molecules in realistic environments of increasing complexity. After an outline of the theoretical basis for the polarizable embedding model, we discuss the importance of using an accurate embedding potential, and how this may be used to significantly reduce the size of the part of the system treated using quantum mechanics without compromising the accuracy of the final results. Furthermore, we discuss the calculation of local electronic excited states based on response theory. We finally discuss aspects related to two recent extensions of the model (i) effective external field and (ii) polarizable density embedding emphasizing their importance for efficient yet accurate description of excited-state properties in complex environments.

1 Introduction

Photoresponsive systems are ubiquitous in Nature, and biological functions such as photosynthesis, phototropism, phototaxis and vision are all triggered and regulated by electronic excitation processes.[1] Excitation phenomena in molecular systems also find applications in many areas within physics and chemistry such as organic solar cells,[2] light-driven molecular rotors[3] and optical probes[4,5] and are thus critical to our understanding of physical, chemical and biological processes. To gain insight into such fundamental processes at the molecular level use of computational methods is highly beneficial. The great virtue of computational science is that simulations can provide information directly at the atomistic scale and classical simulation methods, like classical molecular dynamics (MD), have for a long time been established methods within the physical, chemical and biological sciences where they are used to complement theory and experiment. However, the main shortcoming of classical MD is its lack of flexibility in regards to processes that can be studied, and the fact that it relies on a predefined force field.[6]

Quantum chemistry and dynamics provide a powerful tool to elucidate key aspects of the fundamental mechanisms underlying light-triggered processes that are imperative to achieve rational molecular design. The large-scale nature of molecular materials and biomolecular systems, however, greatly challenges

the use of conventional quantum-chemical methods because of their high computational cost which rapidly increases with system size. A successful strategy to treat local electronic transitions in large molecular systems is combined quantum-classical models most notably combined quantum mechanics/molecular mechanics (QM/MM) models.[7–10] Such models rely on the assumption that the total system can be partitioned into two systems in which the quantum-mechanical description is restricted to only a smaller part of the system (the chromophore), whereas the remaining part (the environment) is treated classically. This effectively extends the range of applicability of quantum-mechanical methods to large molecular systems. Focusing on properties related to the part of the system treated using quantum mechanics, embedding methods may be introduced as models in which the environment is represented by an embedding potential. Such potentials may be of purely electrostatic nature or contain non-electrostatic components as well. One usually distinguishes between embedding methods that treat the environment as an unstructured medium or in atomic detail. In the continuum models, belonging to the first class of these methods, the environment is represented by a dielectric medium that surrounds a cavity hosting the quantum region.[11] The most advanced continuum models belong to the class of polarizable continuum models (PCM)[12,13] in which the environment polarization is represented by an apparent surface-charge distribution spread on the cavity surface. The great advantage of the continuum models is the implicit description of the environment dynamics meaning that there is no need for explicit configurational sampling. A physically more appealing strategy is atomistic embedding in which the environment is represented by its discrete charge distribution. The most common embedding strategy is electrostatic embedding, where

Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark. E-mail: kongsted@sdu.dk, jmo@sdu.dk
† Present address: Division of Theoretical Chemistry and Biology, School of Biotechnology, KTH Royal Institute of Technology, SE-106 91, Stockholm, Sweden. E-mail: nalist@kth.se

20234 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250

This journal is © the Owner Societies 2016

Fig. 1 Schematic representation of different classes of embedding models and some selected features. By sampling is meant configurational sampling and by 2 × SCF is meant double SCF, *i.e.*, a dual SCF solution for the quantum region and environment polarization.

the environment is represented by a static charge distribution (described by charges or more generally by multipoles). The impact of the environment then enters into the Hamiltonian for the quantum region as a one-electron operator mimicking the nucleus–electron interaction operator. While this scheme allows the quantum region to be polarized by the environment, the latter cannot respond to changes in the quantum region. This effect, however, can be very important and is crucial for a physically correct coupling, and mutual polarization between the quantum region and the environment—so-called polarized embedding—was in fact already considered in the pioneering work on QM/MM by Warshel and Levitt.[7] An overview of selected features of various embedding methods are summarized in Fig. 1. Within the context of embedding strategies designed for describing excited states, polarization is usually accounted for either on the basis of induced dipoles,[14–17] or by means of a fluctuating charge (FQ) model.[18] Fig. 1 also includes QM/QM embedding methods in which a quantum-mechanical description of the environment is retained. Examples are embedding fragment *ab initio* model potentials (FAIMP)[19,20] and frozen density embedding (FDE).[21,22] In such QM/QM-based embedding methods, polarization can be included through so-called freeze-and-thaw iterations making these models computationally more costly but provides on the other hand a more complete description of the electrostatics, and, in addition, they also include non-electrostatic interactions. A detailed comparison of various embedding methods can be found in ref. 23 and 24 together with general discussions of the relationship between QM/QM embedding approaches and fragmentation-based many-body computational schemes such as the fragment molecular orbital (FMO) method.[25]

With the aim of addressing general molecular properties and excited states of embedded systems, we have in recent years formulated and applied the polarizable embedding (PE) model[26,27]—as detailed in the following sections—to various molecular systems. The strength of the PE model lies in its formulation within quantum-mechanical response theory[28] which makes it possible to calculate a variety of molecular properties for complex systems

in an efficient yet accurate manner. The PE model is a focused computational scheme meaning that only a part of the total system is treated using quantum mechanics whereas the major part of the system is described through an embedding potential. This makes the PE model very efficient and thus ideal for combination with, *e.g.*, MD in order to include effects of configurational sampling – an issue which is very important for both homogeneous and heterogeneous molecular systems. In addition, we have recently introduced the focused polarizable density embedding (PDE)[29] model which, as detailed later in this perspective, captures the main physics of QM/QM embedding but at a reduced computational cost.

In this perspective, a description of the PE model will be presented. Particular emphasis will be given to the description of electronic transition properties associated with local electronic excitations and the physical nature of their coupling to the environment. The presentation will be organized in five parts: first, we define the model in a qualitative fashion highlighting the approximations made in going to an effective environment description followed by a more rigorous definition of the basic elements of the model. We also address the construction and accuracy of the central embedding potential. We proceed by extending the model to a quantum-mechanical response theory framework and discuss an important problem caused by the non-linear embedding operator. Then, we illustrate the capabilities of the model by its application to the calculation of absorption properties of two chromophore–protein complexes, including a discussion of the importance of the explicit polarization component. We further briefly outline the PDE model, which is a recent extension that addresses two issues of the PE model. Finally, we conclude with some reflections on future developments.

2 Fundamentals of the PE model

The PE model relies on a physical division of the system into subsystems: a quantum region and its environment, which are

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20234–20250 | 20235

treated at different levels of theory. We start with a brief outline of the series of approximations made in the transition from a full quantum-mechanical description to this differentiated treatment. To this end, consider a system that can be decomposed into subsystems A and B. In the end, the goal is to describe only subsystem A (defining the quantum region) explicitly, while incorporating the effect of subsystem B (the environment) through an effective embedding operator. Fig. 2 provides an overview of the approximations that lead to the embedding operator defining the PE model.

The key assumption in the PE model and other quantum-classical approaches is that the subsystems are strictly non-overlapping (step 1). This implies that exchange-repulsion

between subsystems vanishes and there is no need for inter-subsystem antisymmetrization.[30] Within this approximation, the exact wave function of the system can be written as a linear combination of the states in the direct-product space of the complete subsystem spaces.[31] The next step is to invoke a mean-field approximation of the inter-subsystem interactions by restricting the wave function of the composite system to a single generalized Hartree product of subsystem wave functions (step 2). The variational condition then leads to a set of coupled subsystem equations, each involving the electrostatic potential of the other subsystem.[32] At this point, no distinction is made between the description of subsystem A and B.

To achieve a differentiated treatment of the subsystems, a perturbation approach is taken for subsystem B (the environment), allowing it to be only linearly responsive. This corresponds to keeping terms in the electrostatic potential through second order in the interaction (step 3).[32,33] Hereby, we obtain an effective optimization equation for subsystem A (the quantum region), involving the electrostatic potential of the permanent and induced charge distributions of B. This potential depends on the density of subsystem A itself through the induced charge distribution of B and thus leads to a non-linear effective Hamiltonian. Finally, we make the approximation to represent the permanent and first-order induced electrostatic potentials of subsystem B by multipole expansions to eliminate the explicit reference to its wave function (step 4). The use of multipole expansions is computationally advantageous because the inter-subsystem two-electron Coulomb integrals involved in steps 2 and 3 are replaced by one-electron integrals.

In summary, this series of approximations leads to an effective Schrödinger equation for the quantum region in which the electrostatic effect of the environment enters through an embedding operator. In the PE model, the permanent and first-order induced charge distribution of the environment is described by distributed multipole expansions taking the isolated environment molecules as expansion points: each site in the environment is assigned a set of permanent multipole moments as well as an anisotropic dipole–dipole polarizability. The partitioning of the density into site contributions is not unique and several distribution schemes for multipole moments are available.[34–38] A comparison of different schemes can be found in ref. 39.

In the following, we detail the working equations for the PE model. While $|0\rangle$ will denote a generic variational wave function and $|\tilde{0}\rangle$ its time-dependent counterpart, we note that the PE model has been implemented at several quantum-mechanical levels, encompassing both variational and non-variational electronic structure methods. Starting with the original formulation and implementation of the PE model at the level of Hartree–Fock (HF) and Kohn–Sham density functional theory (KS-DFT) for linear and up to cubic response,[26,27,40] the PE model has been implemented within the framework of multiconfigurational self-consistent-field (MCSCF) wave function theory and linear response,[41,42] including the multiconfiguration short-range DFT (MC-srDFT) approach,[43] and the density matrix renormalization

full QM

$$\hat{H}_{AB}|0_{AB}\rangle = E_{AB}|0_{AB}\rangle$$

(1) non-overlapping subsystems

A   B

$$\{|0_A\rangle, |i_A\rangle\} \otimes \{|0_B\rangle, |j_B\rangle\}$$

(2) single Hartree-like product

$$(\hat{H}_A + \langle 0_B|\hat{V}_{AB}|0_B\rangle)|0_A\rangle = E_A|0_A\rangle$$
$$(\hat{H}_B + \langle 0_A|\hat{V}_{AB}|0_A\rangle)|0_B\rangle = E_B|0_B\rangle$$

(3) linearly responsive environment

$$(\hat{H}_A + \langle 0_B|\hat{V}_{AB}|0_B\rangle)^{(II)}|0_A\rangle = E_A|0_A\rangle$$

(4) multipole expansion

$q_1, \boldsymbol{\mu}_1, \boldsymbol{\Theta}_1, \ldots$
$\boldsymbol{\alpha}_1$
$q_2, \boldsymbol{\mu}_2, \boldsymbol{\Theta}_2, \ldots$
$\boldsymbol{\alpha}_2$
$q_3, \boldsymbol{\mu}_3, \boldsymbol{\Theta}_3, \ldots$
$\boldsymbol{\alpha}_3$

$$(\hat{H}_0 + \hat{v}_{PE})|0\rangle = E|0\rangle$$

the PE model
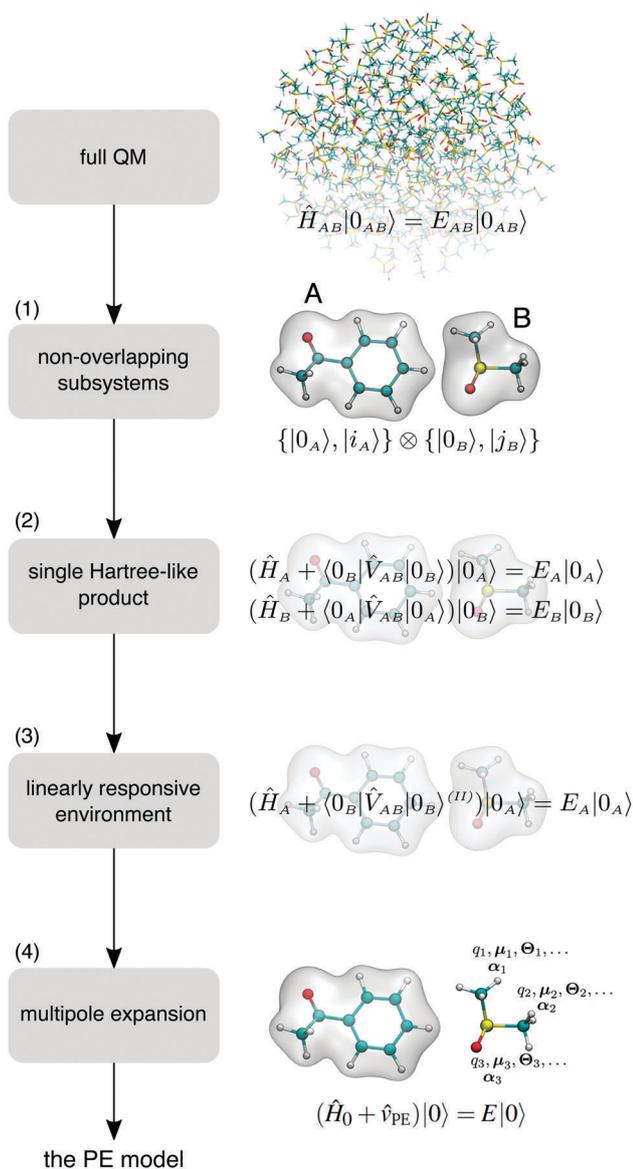
Fig. 2 Schematic representation of the approximations made in going from a full quantum mechanical description of a composite system, exemplified by an acetophenone–dimethylsulfoxide system, to the differentiated treatment defining the PE model, where the solute is the quantum region and the solvent is the environment. See description in the text.

20236 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250

This journal is © the Owner Societies 2016

group (DMRG) method,[44] coupled-cluster (CC) linear and quadratic response theory[45,46] including the very efficient RI-CC2 model,[47,48] as well as a linear response implementation within the second-order polarization propagator approximation (SOPPA) formalism.[49] Furthermore, the use of London atomic orbitals has made it possible to calculate gauge-origin independent magnetic properties.[50] We also note that the PE model has been combined with continuum solvation descriptions for inclusion of bulk solvation effects.[51,52]

2.1 The PE-QM energy functional

In the PE model, the energy functional for a composite system consisting of a quantum region embedded in a polarizable environment is partitioned into three contributions

$$\mathcal{E}_{\text{tot}} = \mathcal{E}_{\text{QM}} + \mathcal{E}_{\text{PE}} + \mathcal{E}_{\text{env}}, \tag{1}$$

where $\mathcal{E}_{\text{QM}}$ is the energy of the quantum region (including wave function polarization), and $\mathcal{E}_{\text{env}}$ is the internal energy of all fragments in the environment, including the electrostatic interaction between these fragments, but excluding the energy associated with the creation of the dipole polarization in the environment. The latter contributes to the variational determination of the wave function/density of the quantum region, and it is therefore expedient to include the polarization energy of the environment in $\mathcal{E}_{\text{PE}}$ together with the contributions arising from the direct interactions of the quantum region with the environment. The polarizable embedding energy can be decomposed as

$$\mathcal{E}_{\text{PE}} = \mathcal{E}_{\text{es}} + \mathcal{E}_{\text{ind}}, \tag{2}$$

where $\mathcal{E}_{\text{es}}$ and $\mathcal{E}_{\text{ind}}$ are the electrostatic contributions arising from interactions between the quantum region and the permanent and induced charge distributions of the environment, respectively. Terms modeling non-electrostatic interactions such as dispersion, non-electrostatic repulsion and charge transfer can in principle be added as energy corrections, but they are not part of the PE model because such energy corrections do not enter into the optimization of the wave function/density of the quantum region.

The electrostatic interaction energy can be written succinctly by introducing a multi-index notation.[53] A multi-index $k = (k_x, k_y, k_z)$ is a 3-tuple, i.e., an ordered list of non-negative integers, where each integer is associated to the indicated component of a Cartesian coordinate. The norm and factorial of a multi-index are defined as $|k| = k_x + k_y + k_z$ and $k! = k_x! \cdot k_y! \cdot k_z!$, respectively. Using this notation, the electrostatic interaction energy takes the form

$$\mathcal{E}_{\text{es}} = \langle 0|\hat{v}^{\text{es}}|0\rangle + \sum_{s=1}^{S} \sum_{|k|=0}^{K_s} \frac{(-1)^{|k|}}{k!} M_s^{(k)} \sum_{n=1}^{N} T_{sn}^{(k)} Z_n, \tag{3}$$

where the $|k|$ summation runs over the $(|k| + 1)(|k| + 2)/2$ multi-indices for a given $|k|$, e.g., for $|k| = 1$, the summation is over $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, which corresponds to x, y and z components. $M_s^{(k)}$ is the k-component of the $|k|$'th-order Cartesian multipole moment located at $\mathbf{R}_s$ in the environment, and Z_n is the nuclear charge of the n'th nucleus in the quantum region.

K_s is the truncation level of the multipole expansion assigned to site s. S is the number of sites in the environment.

Within a second-quantization formalism, the electrostatic $\hat{v}^{\text{es}}$ operator is defined as

$$\hat{v}^{\text{es}} = \sum_{s=1}^{S} \sum_{|k|=0}^{K_s} \frac{(-1)^{|k|}}{k!} M_s^{(k)} \sum_{pq} t_{pq}^{(k)}(\mathbf{R}_s) \hat{E}_{pq}, \tag{4}$$

where $\hat{E}_{pq}$ is a one-electron orbital excitation operator[54] and p and q label general molecular orbital indices belonging to the quantum region. The $t_{pq}^{(k)}(\mathbf{R}_s)$ integrals are given by

$$t_{pq}^{(k)}(\mathbf{R}_s) = -\int \phi_p^*(\mathbf{r}_i) T_{si}^{(k)} \phi_q(\mathbf{r}_i) \mathrm{d}\mathbf{r}_i \tag{5}$$

and involve the k'th component of the interaction tensor defined as a partial derivative of the potential between two sites i and j, i.e.,

$$T_{ij}^{(k)} = \frac{\partial^{|k|}}{\partial x_j^{k_x} \partial y_j^{k_y} \partial z_j^{k_z}} \left(\frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \right). \tag{6}$$

For a linearly responsive environment, the variationally optimized induction energy associated with the polarization of the environment both internally and by the electrons and nuclei in the quantum region amounts to half the interaction energy with the induced dipole moments[32,55,56]

$$\mathcal{E}_{\text{ind}} = -\frac{1}{2} \sum_{s=1}^{S} \bar{\boldsymbol{\mu}}_s(\mathbf{F})^{\text{T}} \mathbf{F}(\mathbf{R}_s), \tag{7}$$

where the vector $\bar{\boldsymbol{\mu}}_s$ is the induced dipole moment at site s in the environment. $\mathbf{F}(\mathbf{R}_s)$ is the electric field vector acting on site s that contains the fields from the nuclei and electrons in the quantum region as well as the fields from the permanent multipole moments in the environment

$$\mathbf{F}(\mathbf{R}_s) = \mathbf{F}^{\text{n}}(\mathbf{R}_s) + \langle 0|\hat{\mathbf{F}}^{\text{e}}(\mathbf{R}_s)|0\rangle + \mathbf{F}^{\text{m}}(\mathbf{R}_s). \tag{8}$$

The induced dipole moments are obtained from a variational optimization of the classical energy of a collection of induced dipoles interacting with an electric field. This can be cast into a set of classical linear response equations given as[57]

$$\bar{\boldsymbol{\mu}}(\mathbf{F}) = \mathbf{B}\mathbf{F}, \tag{9}$$

where the $3S$-dimensional vector $\bar{\boldsymbol{\mu}}$ contains all the induced dipole moments at the polarizable sites in the environment and $\mathbf{F}$ the corresponding electric fields. $\mathbf{B}$ is the classical linear response matrix (also called the relay matrix) of dimension $3S \times 3S$. It is defined as[57]

$$\mathbf{B} = \begin{pmatrix} \boldsymbol{\alpha}_1^{-1} & -\mathbf{T}_{12}^{(2)} & \cdots & & -\mathbf{T}_{1S}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \boldsymbol{\alpha}_2^{-1} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & -\mathbf{T}_{(S-1)S}^{(2)} \\ -\mathbf{T}_{S1}^{(2)} & \cdots & & -\mathbf{T}_{S(S-1)}^{(2)} & \boldsymbol{\alpha}_S^{-1} \end{pmatrix}^{-1}, \tag{10}$$

where the inverse of the electronic dipole–dipole polarizabilities are placed on the diagonal, while the dipole–dipole interaction tensors are found as the off-diagonal elements.

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, 18, 20234–20250 | 20237

The wave function/density describing the quantum region is determined by requiring $\mathcal{E}_{tot}$ to be stationary with respect to variations of the wave function/density of the quantum region. As discussed in the previous section, the only difference in the resulting wave function optimization equations compared to the vacuum counterparts is the presence of the embedding operator

$$\hat{v}_{PE} = \hat{v}^{es} + \hat{v}^{ind}, \tag{11}$$

where the electrostatic operator is defined in eqn (4) and the induction part is given by

$$\hat{v}^{ind} = -\sum_{s=1}^{S} \sum_{a=x,y,z} \bar{\mu}_{a,s}(\mathbf{F}) \hat{F}_a^e(\mathbf{R}_s). \tag{12}$$

A Cartesian component of the electric-field operator is given by

$$\hat{F}_a^e(\mathbf{R}_s) = \sum_{pq} t_{a,pq}(\mathbf{R}_s) \hat{E}_{pq}, \tag{13}$$

with electric-field integrals defined as

$$t_{a,pq}(\mathbf{R}_s) = -\int \phi_p^*(\mathbf{r}) \frac{R_{a,s} - r_a}{|\mathbf{R}_s - \mathbf{r}|^3} \phi_q(\mathbf{r}) d\mathbf{r}. \tag{14}$$

The induced dipole moments, and in turn the embedding operator, depend on the electronic wave function/density of the quantum region through the electric fields. In other words, the presence of the embedding operator leads to a non-linear effective Hamiltonian. In Section 3.1, we discuss a complication that arises from this non-linearity upon addressing excited electronic states.

As apparent from eqn (4), (10) and (12), an essential step in setting up a PE-QM calculation is the embedding potential, i.e., the distributed multipoles moments and polarizabilities in the environment representing its electrostatic effects on the quantum region. The next section will be concerned with the construction of the embedding potential.

2.2 The embedding potential

The embedding potential parameters (multipole moments and polarizabilities) that are used in the PE model are derived from quantum-mechanical calculations in order to achieve an accurate description of the environment. Special considerations are therefore required in the generation of the embedding potential of large environments involving covalently bonded units (e.g., proteins and DNA), where a full quantum-mechanical treatment cannot be afforded. To overcome this, we employ the fragment-based molecular fractionation with conjugate caps (MFCC) scheme, originally developed by Zhang and Zhang[58] for the calculation of interaction energies but later extended to distributed properties by Söderhjelm and Ryde.[59] Fig. 3 provides an overview of the steps involved in the generation of the embedding potential, as exemplified by a protein.

In the MFCC approach, the environment is divided into chemically-sound fragments (here, amino acids: A_1, A_2) that are capped with groups built from the neighboring fragments (step 1). The capping groups (c_1^*, c_2) serve to saturate the valencies of the broken bonds as well as to mimic the effects from the
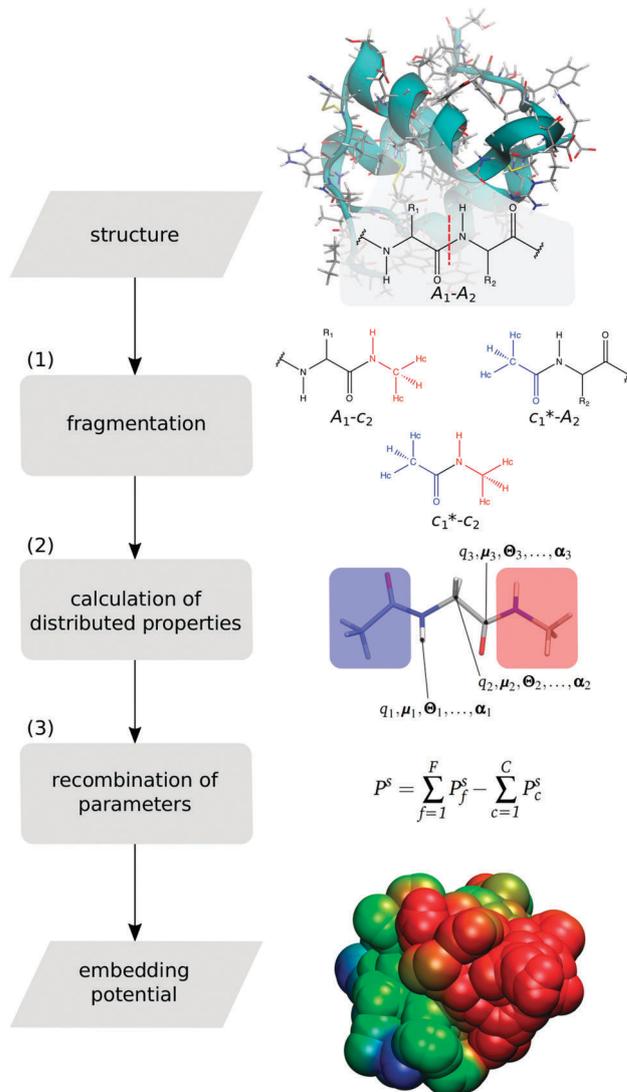
Fig. 3 Steps involved in the generation of the MFCC-based embedding potential: (1) decompose the environment into smaller capped fragments f and merge capping groups to form concaps c. (2) Compute distributed properties of all fragments generated in step 1. (3) Recombine the distributed property P^s for each site s in the environment according to the MFCC scheme to construct the final embedding potential. For environments comprised of individual molecules, e.g. solvents, the embedding parameters can be obtained directly from step 2.

covalently bonded fragments. Moreover, capping groups from neighboring fragments are merged to form a so-called concap to remove the double counting introduced by the capping. Independent distributed property calculations can now be performed on the individual capped fragments and concaps (step 2). The final set of embedding parameters $\{P^s\}$ for a given site s in the environment are finally obtained by subtracting the distributed properties associated with this site in concaps from those in the capped fragment (step 3). The MFCC method leaves the open question of an optimal choice of fragmentation site and size of capping groups. We recently investigated suitable choices for fragmentation of environments involving proteins,[60] nucleic acids[52] and phospholipids.[61]

20238 | Phys. Chem. Chem. Phys., 2016, 18, 20234–20250

This journal is © the Owner Societies 2016

An obvious drawback of this approach is the significant computational cost associated with the preparation of the embedding potential compared to standard QM/MM embedding using parametrized force fields. Indeed, when dealing with large environment fragments and relatively small quantum regions, this step can take more time than the molecular property calculation of the embedded quantum region itself, and this is worsened further by the need to consider dynamic effects due to finite temperature. We have shown, for a set of electric and optical molecular properties, that a combination of geometry-specific embedding parameters close to the quantum region and averaged isotropic embedding parameters further away constitutes a cost-effective embedding potential without compromising the accuracy significantly.[62] At present, such averaged isotropic embedding parameters have been derived and validated for a range of common solvents,[62] nucleic acids[52] and phospholipids.[61]

2.2.1   Performance of the embedding potential. The accuracy of the final PE-QM calculation relies to a high degree on the quality of the underlying structure, the quantum-chemical method, the embedding potential and the coupling between the quantum region and the environment. In this section, we will focus on the quality of the embedding potential, the importance of the explicit polarization component and how it affects molecular properties.

Within the limitations of the quantum-classical embedding strategy, the performance of the embedding potential is restricted by the introduction of a multipole representation of the electrostatic potential as well as by the use of the fragmentation procedure for environments involving covalently bonded units. The most direct way of evaluating the quality of the embedding potential is to compare the resulting electro-static potential (ESP) to a quantum-mechanical reference, since the ESP is the quantity entering the effective Hamiltonian [eqn (11)]. Here we will use the LoProp procedure[63] to obtain the distributed multipole moments and polarizabilities that define the embedding potential. Previous analyses, based on LoProp-derived parameters of solvents and amino acids, have shown that the embedding potential is essentially converged with respect to the order of multipole expansion at the quadrupole level.[60,64] The typical strategy, also followed in the examples in this work, is therefore to include permanent multipole moments up to quadrupoles and anisotropic dipole–dipole polarizabilities, giving the so-called M2P2 potential. All PE-QM calculations presented in this work were performed in a development version of the Dalton program. The PE contributions were handled by the PE library,[40,65] which has been interfaced with Dalton. The LoProp-based embedding potentials were computed using the Molcas program[66,67] or Dalton combined with the Loprop-for-Dalton Python script.[68] All steps in the generation of the potentials were handled using the PE Assistant Script.[40]

In Fig. 4, the ESP of insulin generated by the M2P2 potential (B3LYP[69–72]/cc-pVDZ[73,74]) derived using the MFCC scheme is benchmarked against a full-structure DEC-MP2[75]/cc-pVDZ reference ESP[76] (see Fig. 4b). The ESP difference is probed on
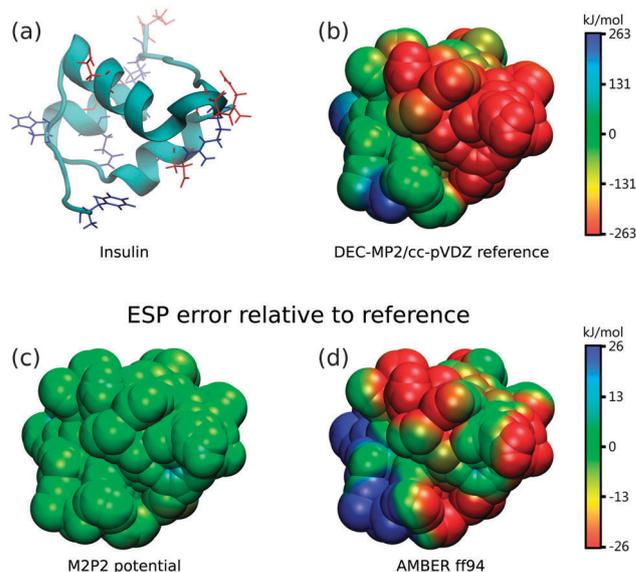
Fig. 4 (a) Structure of an insulin monomer (PDB ID: 1MSO[77]) with five cationic residues (blue) and six anionic residues (red). (b) The ESP of insulin computed at the DEC-MP2/cc-pVDZ level of theory. The difference between the ESP of insulin generated by (c) the M2P2 potential (B3LYP/cc-pVDZ) based on the fragmentation procedure described in Fig. 3 and (d) the embedding potential based on the electrostatic component of the AMBER ff94 force field, both relative to the MP2 reference. The surfaces are defined at twice the vdW atomic radii and the error is reported as interaction energies with an elementary charge. Results from ref. 60.
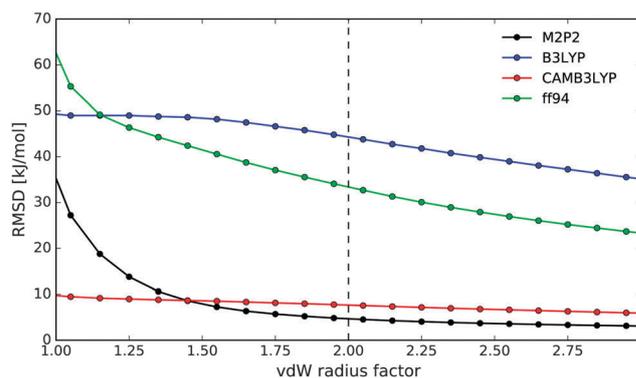
Fig. 5 RMSD of the ESP at different distances relative to the DEC-MP2/cc-pVDZ reference. The vertical dashed line indicates the vdW radius factor used for the ESPs in Fig. 4. The RMSDs are calculated at a molecular surface defined by spheres of the vdW atomic radii times a factor. Results from ref. 60.

a molecular surface defined by twice the van der Waals (vdW) atomic radii, and its distance dependence, given as root-mean-square-deviations (RMSDs), is shown in Fig. 5. From Fig. 4c, it is clearly seen that the fragment-based embedding potential reproduces the quantum-mechanical reference quite well (RMSDs of ~ 5 kJ mol^{-1}) except at short distances (as seen in Fig. 5) where the deviation expectedly increases as a result of charge penetration errors and divergence of the multipole expansions.[30] In fact, the M2P2 potential completely outperforms the corresponding full-structure B3LYP ESP, which has substantial errors (Fig. 5).

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20234–20250 | 20239

As discussed in ref. 76, the reason for these large errors is that B3LYP and other common exchange–correlation functionals suffer from charge-transfer overstabilization due to incomplete cancellation of self-interactions. This issue does not appear in the B3LYP-based embedding potential because the fragmentation procedure does not allow charge transfer between non-neighboring amino acids. The errors can be partly removed by using the long-range corrected CAM-B3LYP functional but the embedding potential still rivals the full-structure CAM-B3LYP ESP even relatively close to insulin. For comparison, we include in Fig. 4d and 5 the ESP derived from the electrostatic component of the AMBER ff94 force field[78] as a representative of an embedding potential that is often used in electrostatic embedding QM/MM calculations. The RMSD at twice the vdW radii increases by a factor of ~ 7 with respect to the M2P2 potential, illustrating that the implicit polarization included in ff94 through enhanced but fixed charges is not flexible enough to yield an accurate ESP. As shown in ref. 79, the inclusion of explicit polarization is particularly important for heterogeneous environments with many charged residues, which is indeed the case for insulin (see Fig. 4a).

We have seen that high-quality potentials can be obtained from fragment-based distributed property calculations and that explicit inclusion of polarization effects is a necessity for an accurate ESP. An important reason for using high-quality embedding potentials in actual calculations is that the size of the quantum region can be reduced without deteriorating the accuracy, as exemplified below. This is particularly important when considering higher-level quantum-chemical methods, such as CC theory, where larger quantum regions cannot be afforded.

To illustrate the importance of a good quality embedding potential on the molecular properties of the quantum region, we consider as an example the convergence of the excitation energy of the lowest $\pi \rightarrow \pi^\star$ transition in acrolein in aqueous solution with respect to the size of the quantum region (total system size defined by radius 12 Å). More specifically, we compare the behavior when the classically-treated water molecules are represented by a polarizable M2P2 potential (based on CAM-B3LYP/6-31+G*[80–84]) to that obtained using the non-polarizable TIP3P water potential.[85] Fig. 6 shows the convergence obtained using the M2P2 and TIP3P potentials and a center-of-mass-based distance criterion, R_{QM}, for selecting solvent molecules to be included in the quantum region.

The M2P2-based excitation energy shows a faster convergence with respect to the size of the quantum region than the TIP3P counterpart. Without any waters in the quantum region, the M2P2 result is already within 0.08 eV of the converged value, which is an almost 0.2 eV improvement over the TIP3P potential. As shown by the open markers in Fig. 6 at $R_{QM} = 2$ Å, the inclusion of two water molecules closest to the C=C bond midpoint, and thereby to the π-conjugated system involved in the transition, brings the excitation energy within 0.03 eV of the converged result using the M2P2 potential. This clearly shows that care should be taken when selecting solvent molecules to be included in the quantum region and if chosen wisely a faster convergence can indeed be obtained. It is, however, important
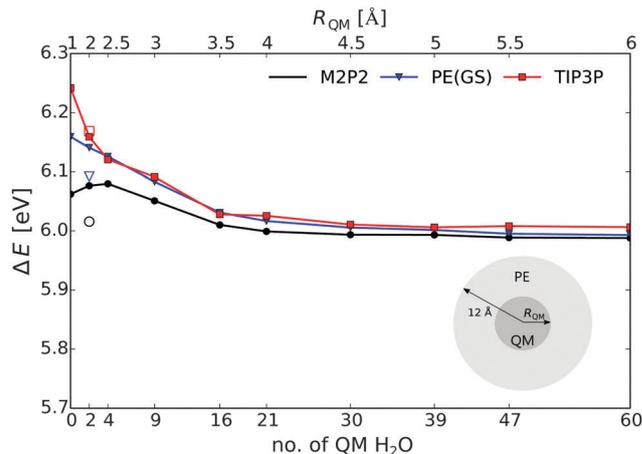
Fig. 6 Convergence of the excitation energy of the lowest π → π* transition in acrolein in aqueous solution with respect to the size of the quantum region (defined by R_{QM}) for different embedding potentials. Open markers at $R_{QM} = 2$ Å indicate results obtained by including the two water molecules closest to the C=C bond midpoint. The results are for an arbitrary MD configuration. Calculations were performed at the CAM-B3LYP/6-31+G* level of theory.

to note that this good agreement to some degree relies on fortuitous error cancellation that hides the effects of the inter-subsystem quantum interactions still neglected (as seen from the initial convergence behavior of the M2P2 results). We will return to this point in Section 5. Nevertheless, the different performances of the M2P2 and TIP3P potentials can also be attributed to the much more accurate ESP provided by the M2P2 potential.

For electronic transitions the explicit environment polarization not only affects the electronic ground state but also the excitation process, as will be discussed in more detail in Section 3. To distinguish these two polarization contributions, we have in Fig. 6 also included M2P2-based results obtained by neglecting the environmental response upon excitation (labeled PE(GS), see Table 1). In the present case, the account of polarization in the ground state provides a modest improvement for the smallest quantum regions, but it is similar to TIP3P in terms of convergence behavior with respect to the size of the quantum region. This illustrates that the response of the environment to the electronic excitation in the quantum region extends far in this environment (more than 6 Å), and that its inclusion is essential to obtain a realistic description of the excitation process.

In summary of this section, we have seen that classical embedding potentials constitute an efficient yet realistic way of describing the long-range electrostatic effects of an environment provided that the embedding potential is formulated accurately and includes explicit environmental polarization.

3 Molecular properties

Optical spectroscopy plays an important role in experimental characterization of molecular properties and the electronic structure of molecules. In such experiments the sample is

Table 1 Hierarchy of models for inclusion of polarization in response-theoretical calculations in the PE framework

Model		Description
PE(GS)	Static reaction field	The induced dipole moments are determined self-consistently in the optimization of the reference state but kept fixed in the subsequent response calculation. This corresponds to including only the $\mathbf{E}^{[2](0)}$ term [eqn (21)] of the electronic Hessian. Essentially, this is analogous to electrostatic embedding.
PE(−EEF)	Full (static and dynamic) reaction field	The environment is allowed to respond to the perturbation-induced changes in the quantum region such that both static and dynamic contributions to the reaction field are included. In a response picture, this effect is incorporated through $\mathbf{E}^{[2]d}$ in terms of transition electric fields [eqn (22)].
PE(+EEF)	Full reaction field and effective external field	All contributions to the local field [eqn (15)] are taken into account in this model. In particular, the environment polarization is modified directly by the external field resulting in an effective external field.

subjected to electromagnetic radiation and the electronic and/or geometric structure is probed by observing the molecular response to the applied perturbation. From a theoretical point of view, molecular properties are linked to response functions that can conveniently be calculated based on response theoretical methods.[28,86–88] In the standard formulation, the theory gives access—by analogy to exact-state theory—to the determination of transition properties associated with stationary states through pole and residue analyses of resonant-divergent molecular response functions. The alternative resonant-convergent formulation of response theory provides response functions that are physically sound across the entire frequency range.[89–91] This allows addressing resonance spectroscopies without having to resolve individual states, a feature that can be of particular importance in frequency regions with a high density of states.

In this section, we will briefly outline the formulation of PE within a quantum-mechanical response framework, which allows for the calculation of response and transition properties of molecules embedded in large environments. Two new concepts need to be introduced to describe the physical aspects of the interactions between the quantum region, the environment and the external perturbing fields: (i) the non-equilibrium regime due to the existence of different relaxation times for various degrees of freedom and (ii) the local field.

As mentioned in Section 2, the PE model has been developed and implemented for both linear and non-linear response functions at several quantum-mechanical levels, including variational and non-variational theories. Recently, we considered resonant-convergent response theory within the PE formalism. The extension to resonance-convergent linear response theory parallels the PE modifications in the standard formulation and has been presented in ref. 92 and 93 at the time-dependent DFT and CC levels, respectively. In the following, we restrict our attention to the coupling of PE to the conventional response theory approach and specifically consider the linear response to a uniform oscillating electric field. For simplicity of interpretation, we work within a configuration interaction (CI) framework, which is parametrized via the state-transfer operators: $\hat{R}_n^\dagger = |n\rangle\langle 0|$ and $\hat{R}_n = |0\rangle\langle n|$, where $\{|n\rangle\}$ is a set of orthonormalized states that spans the orthogonal complement of the reference state $|0\rangle$.

In the unperturbed ground state (no external field), the quantum region and the environment are fully equilibrated.

However, upon applying a time-dependent field, the nuclear frameworks cannot follow the dynamics of the field, and the molecular response to the external field therefore only involves the electronic degrees of freedom. This is characteristic for a vertical electronic (de-)excitation where the electronic degrees of freedom of the environment respond immediately to the electronic transition in the quantum region with no appreciable changes in the intra- or inter-subsystem nuclear configuration. For an isolated molecule, this is equivalent to the standard Franck–Condon approximation and corresponds, in the terminology of embedding models, to the so-called non-equilibrium regime.[94,95] The non-equilibrium response can be properly described with polarized embedding models. In the PE model, the electronic polarization of the environment due to the external perturbation is included through the induced dipole moments while the permanent multipole moments and induced dipole moments from the ground-state calculation represent the fixed orientational polarization of the environment.

The second important aspect to be taken into account in the definition of molecular properties of embedded systems is the distinction between the field actually acting on the quantum region—the local field—and the externally applied electric field, which would be experienced by the quantum region in isolation (see ref. 11, 96–100 and references therein). The local field at the embedded molecule differs from the externally applied field due to the interactions involving the quantum region, its environment and the external field: it is the superposition of the external field and the fields generated by the permanent and induced charge distributions of the environment. In the PE model, a Cartesian component of the local electric field probed at location $\mathbf{R}_m$ within the quantum region is given by[101]

$$F_a^{\mathrm{LF}}(\mathbf{R}_m) = F_a^t(\mathbf{R}_m) + F_a^m(\mathbf{R}_m) + \sum_{s=1}^{S} \sum_{b=x,y,z} T_{sm,ab}^{(2)} \bar{\mu}_{s,b}(\tilde{\mathbf{F}}), \quad (15)$$

where $\tilde{\mathbf{F}}$ is the time-dependent version of eqn (7) augmented with the external electric field:

$$\tilde{\mathbf{F}} = \mathbf{F}^n + \langle \tilde{0}(t)|\hat{\mathbf{F}}^e|\tilde{0}(t)\rangle + \mathbf{F}^m + \mathbf{F}^t. \quad (16)$$

The local field is the central quantity for deriving response and transition properties of embedded molecules that can be related to those measured in experiments, i.e., defined in terms

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20234–20250 | **20241**

of the applied external field. As will be discussed further below, the local field can be divided into two contributions: (i) the reaction field originating from the dipole polarization in eqn (15) induced by the first two terms of eqn (16), and (ii) the effective external field, which is the superposition of the external field and the field from the dipole polarization caused by the external field itself. An additional term to the local field arises in the PE model due to the permanent multipole moments of the environment [second term in eqn (15)]. In Section 4.1, the importance of these contributions in describing transition properties of embedded molecules will be illustrated.

The incorporation of the local-field effects in a response framework can be accomplished using the quasi-energy derivative formulation,[86] where the ground-state energy is replaced by the time-averaged quasi-energy. Within this formalism, the molecular response functions are determined by expanding the time-averaged quasi-energy in orders of the periodic perturbation, $\hat{V}(t) = \sum_{\pm\omega} \sum_{a=x,y,z} \hat{V}_a^\omega F_a^\omega e^{-i\omega t}$, and by imposing that the variational condition is satisfied at the various orders. Response functions can then be identified from the time-averaged Hellmann–Feynman theorem as derivatives of the time-averaged quasi-energy with respect to suitable external field strengths.[86]

The time-dependent quasi-energy in the PE framework is obtained by augmenting the quasi-energy of the isolated quantum region with the time-dependent analogue of $\mathcal{E}_{PE}$ and $\mathcal{E}_{env}$ in eqn (1). The resulting expression becomes

$$
\begin{aligned}
Q_{tot}(t) = & \langle \tilde{0}(t)| \hat{H}_0 + \hat{V}(t) - i\frac{\partial}{\partial t}|\tilde{0}(t)\rangle \\
& + \langle \tilde{0}(t)|\hat{v}^{es}|\tilde{0}(t)\rangle - \frac{1}{2}\sum_{s,s'=1}^{S} \tilde{\mathbf{F}}(\mathbf{R}_s)^T \mathbf{B}_{ss'} \tilde{\mathbf{F}}(\mathbf{R}_{s'}) + E'_{env},
\end{aligned}
\tag{17}
$$

where $\hat{H}_0$ denotes the vacuum Hamiltonian, and an apostrophe signifies inclusion of interaction terms with the external field. $|\tilde{0}(t)\rangle$ is the phase-isolated wave function built from the unperturbed reference state. The response functions are then defined as derivatives of the time average of eqn (17), evaluated at zero field strength. Applying the above procedure, the linear response function for the composite system becomes

$$
\frac{d^2\{Q_{tot}(t)\}_T}{dF_a^{-\omega} dF_b^\omega}\bigg|_{\mathbf{F}^\omega=0} = \alpha_{ab}^{env}(-\omega;\omega) + \langle\langle \hat{\bar{V}}_a^{-\omega}, \hat{\bar{V}}_b^\omega \rangle\rangle,
\tag{18}
$$

where

$$
\langle\langle \hat{\bar{V}}_a^{-\omega}, \hat{\bar{V}}_b^\omega \rangle\rangle = -\bar{\mathbf{V}}_a^{\omega\dagger}\left(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]}\right)^{-1}\bar{\mathbf{V}}_b^\omega.
\tag{19}
$$

As seen from eqn (18), the quasi-energy formulation results in a symmetric form of the response function involving two terms: $\boldsymbol{\alpha}^{env}$ is an explicit contribution from the isolated environment to the linear polarizability of the composite system, while the second term in eqn (18) encompasses the interactions between the embedded quantum region and the environment. This expression for the linear response function is suited for determining transition

properties of the embedded system. Note that an alternative formulation of response theory follows from the Ehrenfest approach in which properties are defined from a perturbation expansion of the expectation value of a one-electron operator.[28] Applied in the PE context, the latter approach leads to a single asymmetric term

$$
\langle\langle \hat{V}_a^{-\omega}, \hat{V}_b^\omega \rangle\rangle = -\mathbf{V}_a^{\omega\dagger}\left(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]}\right)^{-1}\bar{\mathbf{V}}_b^\omega,
\tag{20}
$$

which corresponds to the so-called effective linear response property of the embedded quantum region, i.e., defined with respect to the external field.[96] The choice of symmetric [eqn (19)] or asymmetric [eqn (20)] response functions corresponds to different divisions of the property of the total system, where the former is appropriate for transition strengths and the latter if focus is on properties of the quantum region.

Within a variational framework, the overall form of the linear response function in eqn (19) [or eqn (20)] is preserved as in vacuum but with environmental contributions entering through the electronic Hessian $\mathbf{E}^{[2]}$ and property gradient $\bar{\mathbf{V}}_a^\omega$. The situation is not this simple in a non-variational framework, where additional purely environment-related terms appear (see e.g. ref. 102 for a comparison of PE-QM linear response theory in the two frameworks). The metric matrix $\mathbf{S}^{[2]}$ is unchanged and its definition can be found, e.g., in ref. 86. The electronic Hessian can be partitioned into static and dynamic contributions that read as

$$
\mathbf{E}^{[2](0)} = \langle 0|[\hat{\mathbf{R}}, [\hat{H}_0 + \hat{v}_{PE}, \hat{\mathbf{R}}^\dagger]]|0\rangle.
\tag{21}
$$

$$
\mathbf{E}^{[2]d} = -\sum_{s,s'=1}^{S} \langle 0|[\hat{\mathbf{R}}, \hat{\mathbf{F}}_s^e]|0\rangle^T \mathbf{B}_{ss'} \langle 0|[\hat{\mathbf{R}}^\dagger, \hat{\mathbf{F}}_{s'}^e]|0\rangle.
\tag{22}
$$

The PE contributions to the electronic Hessian model the effects of the reaction-field contribution to the local field, i.e., the impact of the direct coupling between the quantum region and the environment. The static term reflects the situation in which the environment is not allowed to respond to the changes in the electronic density of the quantum region induced by the applied perturbation. This picture can be viewed as a special case of electrostatic embedding, where the environment is frozen at the polarization induced by the ground state of the quantum region. In standard electrostatic embedding schemes only the term from $\hat{v}^{es}$ of the embedding operator in eqn (11) is retained in $\mathbf{E}^{[2](0)}$. The inclusion of explicit environment polarization introduces the additional $\mathbf{E}^{[2]d}$ contribution. This term accounts for the indirect linear response of the environment to the applied field. In other words, it describes the environmental response to the oscillating density of the quantum region (at frequency ω) induced by the time-dependent external field, and thus it gives rise to a dynamic contribution to the reaction field. It should be noted that the polarizabilities entering the classical response matrix $\mathbf{B}$ in eqn (22) are usually approximated with their static limit irrespective of the frequency ω of the external field. This is a reasonable approximation for frequencies in the UV/Vis region that are far from resonances in the environment.[103,104]

The remaining part of the local field describes the implications of the direct environmental response to the applied

20242 | Phys. Chem. Chem. Phys., 2016, 18, 20234–20250

This journal is © the Owner Societies 2016

external field. The additional environment polarization gives rise to an effective external field in the void of the quantum region that differs from the applied field:[101]

$$F_a^{\text{EEF},\omega}(\mathbf{R}_m) = F_a^\omega + \sum_{s=1}^{S} \sum_{b=x,y,z} T_{sm,ab}^{(2)} \bar{\mu}_{s,b}(\mathbf{F}^\omega). \quad (23)$$

Such difference introduces an explicit environment contribution to the property gradient

$$\bar{\mathbf{V}}_a^\omega = \langle 0|[\hat{\mathbf{R}}, \hat{V}_a^\omega]|0\rangle - \sum_{s,s'=1}^{S} \mathbf{e}_{a,s}^{\text{T}} \mathbf{B}_{ss'} \langle 0|[\hat{\mathbf{R}}, \hat{F}^\text{e}(\mathbf{R}_{s'})]|0\rangle, \quad (24)$$

where the first term is the usual unmodified property gradient $\mathbf{V}_a^\omega$, and $\mathbf{e}_a$ is a column vector of length $3S$ holding the a'th Cartesian unit vector for each site in the environment. This effect is similar in origin to the so-called cavity field effect in polarizable continuum models,[97,100] although the latter is relative to the macroscopic (Maxwell) field inside the dielectric rather than the external field. The modification of the applied field due to its interactions with the environment can conveniently be expressed in terms of effective external field tensors that are defined as the linear response of the effective external field with respect to the externally applied field:[101,105]

$$L_{ab}^\omega(\mathbf{R}_m) = \delta_{ab} + \sum_{s=1}^{S} \sum_{c=x,y,z} T_{sm,ac}^{(2)} \frac{\partial \bar{\mu}_{s,c}(\mathbf{F}^\omega)}{\partial F_b^\omega}. \quad (25)$$

By recasting the transition electric field in the second term of eqn (24) to a dipolar form, the direct influence of the environment can be related to the effective external field tensors as

$$\bar{\mathbf{V}}_a^\omega \stackrel{(I)}{=} \sum_{b=x,y,z} \mathbf{V}_b^\omega L_{ba}^\omega(\mathbf{R}_m), \quad (26)$$

where $\stackrel{(I)}{=}$ signifies that the expression on the right-hand side is correct through first order. Insertion of this into eqn (19) gives an (approximate) expression for the response function in terms of the unmodified property gradient and effective external field tensors. The latter gives the connection between the measurable molecular properties, describing the molecular response to the external field, and those in terms of the effective external field (unmodified property gradients).

To distinguish the significance of the different local-field contributions, we introduce a hierarchy of polarization models based on the extent to which environment polarization is incorporated in the response properties of an embedded molecule. These models are summarized in Table 1. In Section 4.1, we will consider the importance of the different contributions as given by this hierarchy of polarization models. Having discussed the key equations in PE-QM linear response theory, we will continue to consider the transition properties resulting from these expressions and how they compare to those obtained from the alternative state-specific formalism.

3.1 Electronic excitations: energies and transition strengths

Within the response framework, excitation energies of electronic transitions in the quantum region are identified as the poles of

the linear response function in eqn (19). This involves the solution of a generalized eigenvalue equation for the electronic Hessian with the metric $\mathbf{S}^{[2]}$:

$$\left(\mathbf{E}^{[2]} - \omega_n \mathbf{S}^{[2]}\right)\mathbf{x}_n = \mathbf{0}, \quad (27)$$

where $\mathbf{x}_n$ represents the n'th excited state with excitation energy $\hbar\omega_n = \Delta E_{n0}$. State-specific approaches provide an alternative strategy for determining excitation energies involving explicit optimization of the initial and final states of interest and simple subtraction of the associated energies. For an isolated molecule the two formalisms give the same results in the limit of exact theory or for linearly parameterized approximate (CI-like) wave functions. Turning to the PE model, and polarized embedding models in general, this equivalence is, however, destroyed as a consequence of the non-linear nature of the effective embedding operator.[106–108]

The difference between the linear-response and state-specific excitation energies in the framework of PE can be illustrated within a first-order perturbation analysis using the eigenvectors $\{|0^{(0)}\rangle, |n^{(0)}\rangle\}$ and -energies $\{E_0^{(0)}, E_n^{(0)}\}$ of $\mathbf{E}^{[2](0)}$ in eqn (21) as zeroth-order basis.[108] The linear response and state-specific transition frequencies for an excitation from the initial state $|0\rangle$ to the final state $|n\rangle$ can through first order be written as

$$\omega_n^{\text{LR}} \stackrel{(I)}{=} \omega_n^{(0)} - \sum_{s,s'=1}^{S} \langle n^{(0)}|\hat{F}^\text{e}(\mathbf{R}_s)|0^{(0)}\rangle^\text{T} \mathbf{B}_{ss'} \langle 0^{(0)}|\hat{F}^\text{e}(\mathbf{R}_{s'})|n^{(0)}\rangle, \quad (28)$$

$$\omega_n^{\text{SS}} \stackrel{(I)}{=} \omega_n^{(0)} - \frac{1}{2}\sum_{s,s'=1}^{S} (\langle n^{(0)}|\hat{F}^\text{e}(\mathbf{R}_s)|n^{(0)}\rangle - \langle 0^{(0)}|\hat{F}^\text{e}(\mathbf{R}_s)|0^{(0)}\rangle)^\text{T} \quad (29)$$
$$\times \mathbf{B}_{ss'}(\langle n^{(0)}|\hat{F}^\text{e}(\mathbf{R}_{s'})|n^{(0)}\rangle - \langle 0^{(0)}|\hat{F}^\text{e}(\mathbf{R}_{s'})|0^{(0)}\rangle).$$

The first (zeroth-order) term is equivalent in the two formulations and can be interpreted as the excitation energy associated with an artificial transition with the environment frozen at the polarization of the ground state of the quantum region. Accordingly, the difference between the two approaches concerns only the second term describing the electronic response of the environment to the excitation process, i.e., the dynamic reaction field contribution. In a linear-response formulation, the environmental response reflects the transition electric field of the quantum region rather than the change in the electric field upon excitation as in state-specific models. Recasting these terms into dipolar forms translates into the environment response being induced either by the transition dipole moment of the quantum region or the difference dipole upon the electronic excitation. The behavior of these two terms will therefore depend crucially on the nature of the specific excitation under consideration. In particular, the dynamical contribution in linear response theory is only significant for bright singlet excitations that have non-vanishing transition dipole moments. The difference between the linear-response and state-specific approaches common to polarized embedding models has been analyzed in ref. 108 on the basis of an exact four-state-model

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, 18, 20234–20250 | 20243

(within the zero-overlap assumption) for two interacting molecules. Within this picture, the dynamical PE contribution retrieved in the linear-response framework has been interpreted as inter-subsystem correlation effects beyond the pure Casimir–Polder dispersion term. On the other hand, the state-specific term is purely electrostatic in nature, taking into account the differential environment response upon excitation and is therefore consistent with the classical description of the environment taken in the PE model.[108]

From the residues of the linear response function in eqn (19), we can identify the one-photon transition strengths associated with the electronic transitions in the quantum region. We obtain[101]

$$\lim_{\omega \to \omega_n} (\omega - \omega_n) \langle\langle \hat{\bar{V}}_a^{-\omega} ; \hat{\bar{V}}_b^{\omega} \rangle\rangle = \bar{\mathbf{V}}_a^{\omega_n \dagger} \mathbf{x}_n \mathbf{x}_n^\dagger \bar{\mathbf{V}}_b^{\omega_n}. \qquad (30)$$

From this it is clear that the environment affects the transition strengths in two ways: (i) indirectly through modified eigenvectors, as discussed above, and (ii) directly through the property gradients in terms of the effective external field. As seen in eqn (26), the impact of the latter can be approximated by means of effective external field tensors. As a consequence, the influence of the effective external field will depend on the orientation of the transition dipole moment for a given excitation, and it may therefore change the shape of the spectrum. The effect becomes even more pronounced when going to higher-order transition properties, which involve longer strings of external effective field tensor components. We will consider an example in Section 4.1.

The strength of combining the PE model with a response-theoretical framework is a direct generalization to higher-order (transition) properties. In the alternative state-specific approach, one-photon transition strengths can be computed in analogy to eqn (30), but using individually optimized electronic states. However, the need to perform separate calculations of all (relevant) electronic states and account for the fact that their mutual orthogonality is, in general, no longer ensured means that the generalization of the state-specific model to higher-order transition properties is not straightforward. Finally, we recall that the inequivalence between the response and state-specific approaches is not restricted to polarized embedding approaches, but is an issue that already enters for molecules *in vacuo* when using non-linearly parameterized wave functions.

4 Example calculations

4.1 Importance of local electric field contributions

Having described how local-field effects enter into the definitions of the transition properties of an embedded molecule in the PE-QM response framework, we turn to examine the importance of the various field contributions by comparing to a full quantum-mechanical reference calculation, keeping in mind inherent differences due to, *e.g.*, basis-set incompleteness. As an example, we consider the one- and two-photon absorption (1PA and 2PA, respectively) properties of the lowest $\pi \to \pi^*$ transition in a 242-atoms cluster model of the red fluorescent
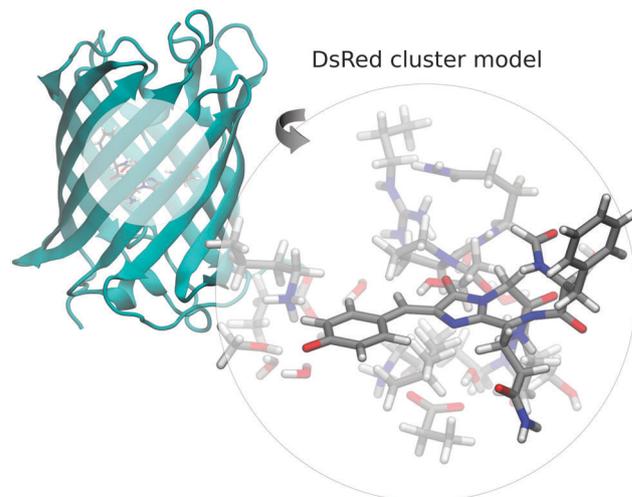
Fig. 7 Cutout of a monomer of the DsRed protein (PDB ID: 1ZGO[109]) defining the cluster model used in the calculations. The chromophore defining the quantum region in the PE-TDDFT calculations is shown in unfaded colors.

protein DsRed (Fig. 7) obtained using the hierarchy of polarization models introduced in Table 1.

The CAM-B3LYP functional was used for the quantum region, together with the 6-31+G* basis set, and the embedding potential was generated at the B3LYP/6-31+G* level of theory. For more details on the preparation of the system and the computational setup, we refer to the literature.[110] All calculations are based on a response treatment.

The transition energy (ΔE), associated oscillator strength (f) and 2PA cross section (σ^{2PA}) are listed in Table 2 for the different polarization models alongside with the quantum-mechanical reference results. In addition, we include the results for the isolated DsRed chromophore adopting the in-protein geometry. The static reaction field, represented by the incremental contribution when going from vacuum (vac) to a static reaction field (PE(GS)), gives rise to a substantial blue shift of the excitation energy that overshoots the reference shift by 0.15 eV. This is largely remedied by allowing the environment to respond to the excitation process in terms of the dynamic reaction field (PE(−EEF)), which brings the PE-TDDFT excitation energy within 0.04 eV of the reference (full QM). The 1PA and 2PA intensities, however, reflect the effective external field at the location of the chromophore rather than the applied external field and are thus not comparable to the full quantum-mechanical reference. Incorporating the effective external field effects (PE(+EEF)), *i.e.*, defining the properties with respect to the external field, leads to absolute intensities that are in good agreement with the full quantum-mechanical counterparts.

The results obtained with the hierarchy of models clearly demonstrate the efficacy of the PE model to simulate environmental effects in systems where the inter-subsystem interactions are dominated by long-range electrostatics. In particular, it provides physical insight into the relative importance of the various contributions to the local electric field, which is not easily accessible from full quantum-mechanical calculations.

20244 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250

This journal is © the Owner Societies 2016

Table 2 Excitation energies ΔE (eV), oscillator strengths f, and 2PA cross section σ^{2PA} (GM) of the lowest $\pi \rightarrow \pi^*$ transition in a DsRed cluster model computed using the various polarization models defined in Table 1. Results from ref. 101

Model	ΔE (eV)	f	σ^{2PA} (GM)
vac[a]	2.53	1.16	47
PE(GS)	3.07	1.08	58
PE($-$EEF)	2.96	1.11	77
PE($+$EEF)	2.96	0.80	42
full QM	2.92	0.84	38

[a] Vacuum results are taken from ref. 110 and correspond to the conformation adopted by the chromophore inside the protein.

4.2 *In silico* mutations in channelrhodopsin

There is great potential in the use of computational models and tools for the rationalization and optimization of the properties of biomolecular systems. The PE model can provide valuable information that can be used to investigate the molecular mechanisms underlying experimental measurements. Furthermore, since we have an atomistic description of the environment, this allows us to manipulate the composition of the environment and to estimate the effects it has on the properties of the quantum region.

Here we present a proof-of-principle example to demonstrate the usefulness of the PE model for *in silico* mutation studies. Specifically, we investigate the 2PA activity of a channelrhodopsin, which belongs to a subfamily of 7-transmembrane retinylidene proteins that act as light-gated ion channels. The channel-opening is triggered by a light-induced isomerization in the all-*trans*-retinylidene chromophore, a vitamin A derivative, which is bound covalently to a lysine residue through a protonated Schiff base (see Fig. 8). Channelrhodopsins are often used as actuators in optogenetics, which is a biological technique that allows *in vitro* or *in vivo* optical control of cellular processes.[111] Their use in living tissue is complicated, however, because of the relatively high absorption and scattering coefficients in biological tissue. The issues can potentially be overcome by using red-shifted variants, or, alternatively, activating the channel using 2PA techniques.

In Fig. 9, we present 2PA results for a Channelrhodopsin-1 (ChR1) and Channelrhodopsin-2 (ChR2) chimera called
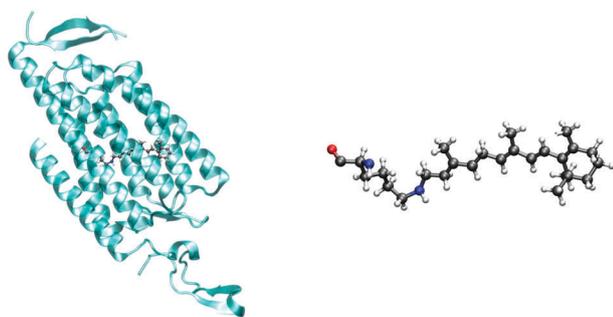
Fig. 8 Cartoon structure of a channelrhodopsin (left) and its retinylidene protonated Schiff base chromophore (right).
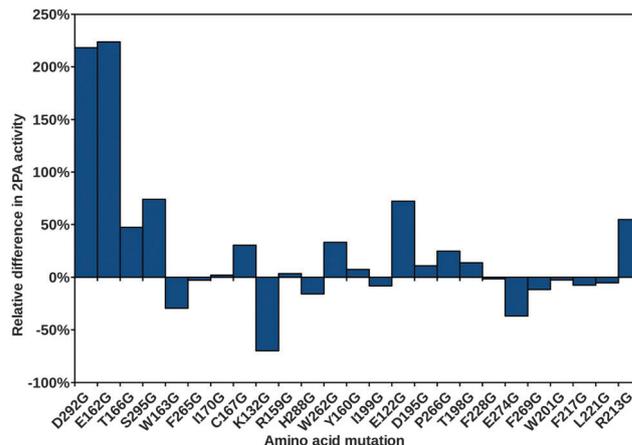
Fig. 9 The gain or loss of 2PA in the C1C2 chimera as a result of point mutations to glycine. Calculated as the relative difference between 2PA transition probabilities. The distance to the chromophore is increasing from left to right.

C1C2[112] (see Fig. 8). The 2PA transition probabilities were calculated at PE-CAM-B3LYP/6-31+G* level using an embedding potential derived at the B3LYP/6-31+G* level. The structures were taken from ref. 113. By mutating residues to glycine, the effect of each residue on the 2PA properties can be mapped, and this gives valuable information that can be used in a rational design process. For example, the largest effect comes from the two Schiff base counter-ions, D292 and E162, that cause a substantial decrease in the 2PA. It has been shown that a D292A mutant does not produce a photocurrent whereas a E162A mutation only has a moderate effect.[112] The E162 residue would therefore be a good candidate for further more elaborate computational mutation studies.

5 The next generation: polarizable density embedding

We have so far demonstrated the reliability and efficiency of the PE model to study localized properties in large molecular systems. The reliability, however, is only guaranteed if the interactions between the quantum region and its environment are dominated by long-range electrostatics. The next generation of our embedding model, polarizable density embedding (PDE),[29] addresses two potential issues that can arise when the distance between the quantum region and the molecules in its immediate environment is short.

As seen in Fig. 5, the error in the embedding potential increases rapidly at short distances. This is mainly because of an incomplete description of the short-range electrostatics partly due to the use of point multipoles, *i.e.*, charge-penetration effects are missing. Therefore, to improve the short-range electrostatics in the PDE model, we use the explicit electron density of the fragments in the environment, instead of a multipole representation of the permanent charge distribution. This increases the computational cost compared to the PE model because the interfragment two-electron Coulomb integrals between the

quantum region and the permanent charge distribution of each fragment in the environment are reintroduced. Compared to other QM/QM-embedding methods, PDE is much more efficient because the classical description of environment polarization from the PE model is retained, thus avoiding the costly iterations that require multiple calculations of the electron density of each fragment in the environment. PDE is designed as a focused model in contrast to *e.g.* FMO which aims for a description of the total molecular system at a consistent level of quantum mechanics through the use of many-body expansions.[25] Since the distributed multipoles in the PE model describe the long-range electrostatics quite well, a three-layer model may be introduced where the environment is split into two regions: an inner region containing fragment densities, and an outer region consisting of distributed multipoles. Correspondingly, we introduce a splitting of the electrostatic part of the embedding operator

$$\hat{v}^{\text{es}} = \hat{v}^{\text{mul}} + \hat{v}^{\text{fd}}, \tag{31}$$

where $\hat{v}^{\text{mul}}$ is defined as in eqn (4) but only includes fragments that are outside a defined threshold distance from the quantum region. The $\hat{v}^{\text{fd}}$ operator describes the electrostatic interactions between the quantum region and the fragment densities in the inner region, and is thus defined as

$$\hat{v}^{\text{fd}} = \sum_{f=1}^{N_{\text{fd}}} \sum_{pq \in \text{QM}} \left(\sum_{rs \in f} D_{rs} v_{pq,rs} - \sum_{m \in f}^{M_f} Z_m v_{pq}(\mathbf{R}_m) \right) \hat{E}_{pq}. \tag{32}$$

Here, the summation runs over N_{fd} fragments, D_{rs} is an element of the density matrix belonging to the f'th fragment, and Z_m is the charge of the m'th nucleus in the f'th fragment. The one-electron integrals in eqn (32) are given by

$$v_{pq}(\mathbf{R}) = \int \phi_p^*(\mathbf{r}) \frac{1}{|\mathbf{R}-\mathbf{r}|} \phi_q(\mathbf{r}) d\mathbf{r}, \tag{33}$$

and the interfragment two-electron integrals by

$$v_{pq,rs} = \iint \phi_p^*(\mathbf{r}_1) \phi_q^*(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} \phi_r(\mathbf{r}_2) \phi_s(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \tag{34}$$

where molecular orbital indices r and s belong to fragments in the inner region as indicated in eqn (32).

The other potential issue that can arise in the PE model is artificial stabilization of the electron density at classical sites in the environment – the so-called electron-spill effect. It is related to the lack of non-electrostatic repulsion, *i.e.*, exchange-repulsion or Pauli repulsion, which introduces substantial errors at short distances when there is significant wave-function overlap.[114,115] In the PDE model, this issue is remedied by introducing a projection operator, based on the work by Huzinaga and Cantu,[116] that models wave-function orthogonality between the quantum region and the fragments in its environment. This non-electrostatic repulsion operator is defined as

$$\hat{v}^{\text{rep}} = -\sum_{f=1}^{N_{\text{fd}}} \sum_{pq \in \text{QM}} \sum_{i \in f}^{N_{\text{occ}}} \varepsilon_i S_{pi} S_{iq} \hat{E}_{pq}, \tag{35}$$

where ε_i is the energy of the i'th occupied orbital of the a'th fragment in the environment, and S_{pi} and S_{iq} are overlap integrals between molecular orbitals in the quantum region and molecular orbitals belonging to fragments in the environment. As seen from eqn (35), the operator associates an energy penalty that scales with the square of the overlap and is weighted by the energy of the occupied molecular orbitals belonging to the fragments in the environment. This effectively prevents the electrons in the quantum region from penetrating too far into the environment. However, since we use the orbitals calculated for each fragment in isolation, the repulsive effect can be too strong, as shown in ref. 29. It may therefore be necessary to introduce a scaling of the non-electrostatic repulsion. In the preliminary analysis in ref. 29, we found that a simple linear scaling gave good results. Further investigations along this line are currently in progress.

6 Concluding remarks and perspectives

In this perspective, we have given an overview of the status as well as recent developments of the polarizable embedding method.

The high-quality embedding potential is a key ingredient for an accurate description. The examples in this perspective illustrate that the electrostatic component of the embedding potential is well-described by the combination of a fragmentation-based approach with quantum-mechanical distributed property calculations as used in the PE and PDE models. However, there remain some directions for further improvements. First of all the environment is assumed to be linearly responsive. Yet, strong electric fields are not unusual in *e.g.* proteins,[117–123] which can be expected to require a treatment of non-linear environmental polarization. The generalization of the LoProp scheme to first hyperpolarizabilities was recently presented in the context of a second-order Applequist model.[124] However, it has not been explored in the context of embedding calculations and the second hyperpolarizability is naturally expected to be of greater importance in isotropic systems. Furthermore, although the use of distributed properties to some extent account for anisotropies in the induced charge distribution, improvements can be made by including higher-order terms in the multipole expansion.[125,126] New computational methodologies for the calculation of distributed properties are currently being explored and initial investigations of distributed multipoles are promising.[39] Current efforts are focused on the extension to distributed any-order multipole polarizabilities and hyperpolarizabilities, as well as the inclusion of these in the PE model.

At present, the PE and PDE models do not include dispersion in the embedding operator but this component might become very important especially when studying apolar species. The simplest approach would be to add the leading-order Casimir–Polder like term as an energy-correction as done in *e.g.* the EFP model.[127] However, such an approach does not lead to an effective dispersion operator entering into the expression for the embedding potential. Ongoing work addresses the construction of such effective methods for inclusion of dispersion directly into the embedding potential for a discrete environment.

20246 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250

This journal is © the Owner Societies 2016

Several new features are in progress to broaden the use of the PE and PDE models. One of the strengths of the PE model is its formulation within a quantum-mechanical response framework. The PE model has recently been combined with an open-ended response theory framework[128,129] at the HF and KS-DFT levels. The details of the implementation together with pilot calculations will be published shortly. This will enable the calculation of electric response properties to any order including also multiphoton transition strengths *via* the recent single residue functionality.[130] In addition, efforts are presently focused on an extension of the capabilities of our previous molecular gradient implementation[131] to any-order geometric and mixed electric-geometric derivatives. This will open the door to the calculation of vibrational spectroscopies and vibrational contributions to molecular properties as well as the optimization of excited electronic states. Based on such novel developments, we expect that the PE and PDE models will play an important role in future advanced spectroscopies and materials design.

The main concern of this work has been on molecular systems consisting of a single chromophore embedded in a complex environment. However, in many cases molecular systems contain a number of chromophores and it is thereby necessary to consider collective excitation and energy-transfer processes, *i.e.*, the excitations become non-local. In such cases it is important to include environment effects directly in the description of the excitation energy and transfer processes. The environment not only tunes the excitation energies of the specific chromophores, it also contributes both directly and indirectly to the calculation of the electronic couplings.[15,132] The PE model has recently been extended to multichromophoric systems paying special attention to the role played by the environment.[133] In the future, the PE model as well as extensions hereof especially in the direction of energy transfer processes for multiconfigurational descriptions of molecular systems will be used to predict and rationalize energy-transfer mechanisms in complex biological systems.

Finally, an important point that has not been touched upon in this work is that of configurational sampling and finite temperature effects, which is crucial when relying on an atomistic representation of the environment. Typically, we have addressed such sampling based on a sequential approach in which the sampling and embedding calculations are treated in an uncoupled manner.[134] Usually the sampling is based on classical MD simulations and the quality of both the sampling and the produced molecular structures thus relies on the quality of the underlying force field. In the future, an interesting approach will be to sample the structures at the same level as the embedding, *i.e.*, use the embedding potential directly in the classical MD, or as an intermediate alternative to use the previously mentioned PE-QM molecular gradient implementation to optimize the geometries of snapshots extracted from a classical MD run using a conventional force field. The latter approach will, however, remove finite temperature effects from the central part and we are therefore investigating the possibility to use *ab initio* MD to fully address the issue of configurational sampling and finite temperature effects.

References

1 M. Mohseni, Y. Omar, G. S. Engel and M. B. Plenio, *Quantum effects in biology*, Cambridge University Press, 2014.
2 S. Günes, H. Neugebauer and N. S. Sariciftci, *Chem. Rev.*, 2007, **107**, 1324–1338.
3 N. Koumura, R. W. J. Zijlstra, R. A. van Delden, N. Harada and B. L. Feringa, *Nature*, 1999, **401**, 152–155.
4 M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes and A. Rebane, *Nat. Methods*, 2011, **8**, 393–399.
5 M. Pawlicki, H. A. Collins, R. G. Denning and H. L. Anderson, *Angew. Chem., Int. Ed.*, 2009, **48**, 3244–3266.
6 F. Jensen, *Introduction to Computational Chemistry*, Wiley, 2nd edn, 2006.
7 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
8 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
9 U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1986, **7**, 718–730.
10 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700–733.
11 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.
12 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.
13 E. Cances, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3032–3041.
14 S. Yoo, F. Zahariev, A. Sok and M. S. Gordon, *J. Chem. Phys.*, 2008, **129**, 144112.
15 C. Curutchet, A. Munoz-Losa, S. Monti, J. Kongsted, G. D. Scholes and B. Mennucci, *J. Chem. Theory Comput.*, 2009, **5**, 1838–1848.
16 L. Jensen, P. T. van Duijnen and J. G. Snijders, *J. Chem. Phys.*, 2003, **118**, 514–521.
17 P. Söderhjelm, C. Husberg, A. Strambi, M. Olivucci and U. Ryde, *J. Chem. Theory Comput.*, 2009, **5**, 649–658.
18 F. Lipparini, C. Cappelli and V. Barone, *J. Chem. Theory Comput.*, 2012, **8**, 4153–4165.
19 Z. Barandiaran and L. Seijo, *J. Chem. Phys.*, 1988, **89**, 5739.
20 B. Swerts, L. F. Chibotaru, R. Lindh, L. Seijo, Z. Barandiaran, S. Clima, K. Pierloot and M. F. A. Hendrickx, *J. Chem. Theory Comput.*, 2008, **4**, 586–594.
21 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050–8053.

22 C. R. Jacob and J. Neugebauer, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 325–362.

23 A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2012, **108**, 222–277.

24 M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko, *Chem. Rev.*, 2012, **112**, 632–672.

25 K. Kitaura, E. Ikeo, T. Asada, T. Nakano and M. Uebayasi, *Chem. Phys. Lett.*, 1999, **313**, 701–706.

26 J. M. H. Olsen, K. Aidas and J. Kongsted, *J. Chem. Theory Comput.*, 2010, **6**, 3721–3734.

27 J. M. H. Olsen and J. Kongsted, *Adv. Quantum Chem.*, 2011, **61**, 107–143.

28 J. Olsen and P. Jørgensen, *J. Chem. Phys.*, 1985, **82**, 3235–3264.

29 J. M. H. Olsen, C. Steinmann, K. Ruud and J. Kongsted, *J. Phys. Chem. A*, 2015, **119**, 5344–5355.

30 A. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, 2002.

31 R. McWeeny and B. T. Sutcliffe, *Methods of Molecular Quantum Mechanics*, Academic press London, 1989, vol. 2.

32 J. G. Ángyán, *J. Math. Chem.*, 1992, **10**, 93–137.

33 P. R. Surján and J. Ángyán, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1983, **28**, 45.

34 R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833–1840.

35 A. J. Stone, *Chem. Phys. Lett.*, 1981, **83**, 233–239.

36 A. J. Stone, *J. Chem. Theory Comput.*, 2005, **1**, 1128–1132.

37 L. Gagliardi, R. Lindh and G. Karlström, *J. Chem. Phys.*, 2004, **121**, 4494–4500.

38 G. Karlström, Proceedings of the 5th Seminar on Computational Methods in Quantum Chemistry, 1981, p. 353.

39 M. S. Nørby, J. M. H. Olsen, J. Kongsted and H. J. Aa. Jensen, *J. Comput. Chem.*, 2016, **37**, 1887–1896.

40 J. M. H. Olsen, PhD thesis, University of Southern Denmark, Odense, Denmark, 2012.

41 E. D. Hedegård, N. H. List, H. J. Aa. Jensen and J. Kongsted, *J. Chem. Phys.*, 2013, **139**, 044101.

42 N. H. List, H. J. Aa. Jensen, J. Kongsted and E. D. Hedegård, *Adv. Quantum Chem.*, 2013, **66**, 195.

43 E. D. Hedegård, J. M. H. Olsen, S. Knecht, J. Kongsted and H. J. Aa. Jensen, *J. Chem. Phys.*, 2015, **142**, 114113.

44 E. D. Hedegård, M. Reiher, arXiv:1605.03232 [physics.chem-ph], submitted, 2016.

45 K. Sneskov, E. Matito, J. Kongsted and O. Christiansen, *J. Chem. Theory Comput.*, 2010, **6**, 839–850.

46 K. Sneskov, T. Schwabe, J. Kongsted and O. Christiansen, *J. Chem. Phys.*, 2011, **134**, 104108.

47 T. Schwabe, K. Sneskov, J. M. H. Olsen, J. Kongsted, O. Christiansen and C. Hättig, *J. Chem. Theory Comput.*, 2012, **8**, 3274–3283.

48 D. Hršak, A. Marefat Khah, O. Christiansen and C. Hättig, *J. Chem. Theory Comput.*, 2015, **11**, 3669–3678.

49 J. J. Eriksen, S. Sauer, K. V. Mikkelsen, H. J. Aa. Jensen and J. Kongsted, *J. Comput. Chem.*, 2012, **33**, 2012–2022.

50 C. Steinmann, J. M. H. Olsen and J. Kongsted, *J. Chem. Theory Comput.*, 2014, **10**, 981–988.

51 A. H. Steindal, K. Ruud, L. Frediani, K. Aidas and J. Kongsted, *J. Phys. Chem. B*, 2011, **115**, 3027–3037.

52 M. S. Nørby, C. Steinmann, J. M. H. Olsen, H. Li and J. Kongsted, *J. Chem. Theory Comput.*, 2016, submitted.

53 X. Saint Raymond, *Elementary Introduction to the Theory of Pseudodifferential Operators*, CRC Press, 1991, vol. 3.

54 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, 2000.

55 J. G. Ángyán, *Int. J. Quantum Chem.*, 1993, **47**, 469–483.

56 M. Orozco and F. J. Luque, *Chem. Phys. Lett.*, 1997, **265**, 473–480.

57 J. Applequist, J. R. Carl and K.-K. Fung, *J. Am. Chem. Soc.*, 1972, **94**, 2952–2960.

58 D. W. Zhang and J. Z. H. Zhang, *J. Chem. Phys.*, 2003, **119**, 3599–3605.

59 P. Söderhjelm and U. Ryde, *J. Phys. Chem. A*, 2009, **113**, 617–627.

60 J. M. H. Olsen, N. H. List, K. Kristensen and J. Kongsted, *J. Chem. Theory Comput.*, 2015, **11**, 1832–1842.

61 S. Witzke, N. H. List, J. M. H. Olsen, C. Steinmann, M. T. P. Beerepoot, M. Petersen and J. Kongsted, 2016, submitted.

62 M. T. P. Beerepoot, A. H. Steindal, N. H. List, J. Kongsted and J. M. H. Olsen, *J. Chem. Theory Comput.*, 2016, **12**, 1684–1695.

63 L. Gagliardi, R. Lindh and G. Karlström, *J. Chem. Phys.*, 2004, **121**, 4494–4500.

64 T. Schwabe, J. M. H. Olsen, K. Sneskov, J. Kongsted and O. Christiansen, *J. Chem. Theory Comput.*, 2011, **7**, 2209–2217.

65 The PE library developers, *PElib: The Polarizable Embedding library (develop version)*, 2015, http://gitlab.com/pe-software/pelib-public.

66 G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady and L. Seijo, *Comput. Mater. Sci.*, 2003, **28**, 222–239.

67 F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov and R. Lindh, *J. Comput. Chem.*, 2010, **31**, 224–247.

68 O. Vahtras, LoProp for Dalton, 2014, http://dx.doi.org/10.5281/zenodo.13276.

69 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

70 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

71 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.

72 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

73 T. H. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007–1023.

74 D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.*, 1993, **98**, 1358–1371.

75 K. Kristensen, I.-M. Høyvik, B. Jansik, P. Jørgensen, T. Kjærgaard, S. Reine and J. Jakowski, *Phys. Chem. Chem. Phys.*, 2012, **14**, 15706–15714.

20248 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250

This journal is © the Owner Societies 2016

76 S. Jakobsen, K. Kristensen and F. Jensen, *J. Chem. Theory Comput.*, 2013, **9**, 3978–3985.

77 G. D. Smith, W. A. Pangborn and R. H. Blessing, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 474–482.

78 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.

79 M. T. P. Beerepoot, A. H. Steindal, K. Ruud, J. M. H. Olsen and J. Kongsted, *Comput. Theor. Chem.*, 2014, **1040–1041**, 304–311.

80 P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.

81 M. M. Francl, W. J. Pietro, W. J. Hehre, S. J. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654–3665.

82 T. Clark, J. Chandrasekhar, G. W. Spitznagel and P. V. R. Schleyer, *J. Comput. Chem.*, 1983, **4**, 294–301.

83 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650–654.

84 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.

85 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

86 O. Christiansen, P. Jørgensen and C. Hättig, *Int. J. Quantum Chem.*, 1998, **68**, 1–52.

87 T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen and K. Ruud, *Chem. Rev.*, 2012, **112**, 543–631.

88 P. Norman, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20519–20535.

89 P. Norman, D. M. Bishop, H. Jo and J. Oddershede, *J. Chem. Phys.*, 2001, **115**, 10323–10334.

90 P. Norman, D. M. Bishop, H. J. Aa. Jensen and J. Oddershede, *J. Chem. Phys.*, 2005, **123**, 194103.

91 K. Kristensen, J. Kauczor, T. Kjærgaard and P. Jørgensen, *J. Chem. Phys.*, 2009, **131**, 044112.

92 M. N. Pedersen, E. D. Hedegård, J. M. H. Olsen, J. Kauczor, P. Norman and J. Kongsted, *J. Chem. Theory Comput.*, 2014, **10**, 1164–1171.

93 N. H. List, S. Coriani, J. Kongsted and O. Christiansen, *J. Chem. Phys.*, 2014, **141**, 244107.

94 H. J. Kim and J. T. Hynes, *Int. J. Quantum Chem.*, 1990, **38**, 821–833.

95 M. A. Aguilar, F. J. O. Del Valle and J. Tomasi, *J. Chem. Phys.*, 1993, **98**, 7375–7384.

96 R. Wortmann and D. M. Bishop, *J. Chem. Phys.*, 1998, **108**, 1001–1007.

97 R. Cammi, B. Mennucci and J. Tomasi, *J. Phys. Chem. A*, 1998, **102**, 870–875.

98 L. Jensen, M. Swart and P. T. van Duijnen, *J. Chem. Phys.*, 2005, **122**, 034103.

99 S. Pipolo, S. Corni and R. Cammi, *Comput. Theor. Chem.*, 2014, **1040**, 112–119.

100 S. Pipolo, S. Corni and R. Cammi, *J. Chem. Phys.*, 2014, **140**, 164114.

101 N. H. List, H. J. Aa. Jensen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10070–10080.

102 N. H. List, S. Coriani, O. Christiansen and J. Kongsted, *J. Chem. Phys.*, 2014, **140**, 224103.

103 I. Harczuk, O. Vahtras and H. Ågren, *Phys. Chem. Chem. Phys.*, 2015, **17**, 7800–7812.

104 M. S. Nørby, O. Vahtras, P. Norman and J. Kongsted, *Mol. Phys.*, 2016, DOI: 10.1080/00268976.2016.1177667.

105 N. Davari, S. Haghdani, P.-O. Åstrand and G. C. Schatz, *RSC Adv.*, 2015, **5**, 31594–31605.

106 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.*, 2002, **100**, 1813–1828.

107 R. Cammi, S. Corni, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 2005, **122**, 104513.

108 S. Corni, R. Cammi, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 2005, **123**, 134512.

109 J. L. Tubbs, J. A. Tainer and E. D. Getzoff, *Biochemistry*, 2005, **44**, 9833–9840.

110 N. H. List, J. M. H. Olsen, H. J. Aa. Jensen, A. H. Steindal and J. Kongsted, *J. Phys. Chem. Lett.*, 2012, **3**, 3513–3521.

111 L. Fenno, O. Yizhar and K. Deisseroth, *Annu. Rev. Neurosci.*, 2011, **34**, 389–412.

112 H. E. Kato, F. Zhang, O. Yizhar, C. Ramakrishnan, T. Nishizawa, K. Hirata, J. Ito, Y. Aita, T. Tsukazaki, S. Hayashi, P. Hegemann, A. D. Maturana, R. Ishitani, K. Deisseroth and O. Nureki, *Nature*, 2012, **482**, 369–374.

113 K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7567–7576.

114 P. Surján and J. G. Ángyán, *Chem. Phys. Lett.*, 1994, **225**, 258–264.

115 G. Fradelos and T. A. Wesołowski, *J. Phys. Chem. A*, 2011, **115**, 10018–10026.

116 S. Huzinaga and A. A. Cantu, *J. Chem. Phys.*, 1971, **55**, 5543–5549.

117 S. D. Fried, S. Bagchi and S. G. Boxer, *Science*, 2014, **346**, 1510–1514.

118 A. T. Fafarman, P. A. Sigala, D. Herschlag and S. G. Boxer, *J. Am. Chem. Soc.*, 2010, **132**, 12811–12813.

119 S. D. Fried, S. Bagchi and S. G. Boxer, *J. Am. Chem. Soc.*, 2013, **135**, 11181–11192.

120 S. D. Fried, L.-P. Wang, S. G. Boxer, P. Ren and V. S. Pande, *J. Phys. Chem. B*, 2013, **117**, 16236–16248.

121 S. D. Fried and S. G. Boxer, *Acc. Chem. Res.*, 2015, **48**, 998–1006.

122 M. Drobizhev, S. Tillo, N. S. Makarov, T. E. Hughes and A. Rebane, *J. Phys. Chem. B*, 2009, **113**, 12860–12864.

123 M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes and A. Rebane, *Nat. Methods*, 2011, **8**, 393–399.

124 I. Harczuk, O. Vahtras and H. Ågren, *Phys. Chem. Chem. Phys.*, 2016, **18**, 8710–8722.

125 A. Holt and G. Karlström, *J. Comput. Chem.*, 2008, **29**, 2033–2038.

126 A. Holt, J. Boström, G. Karlström and R. Lindh, *J. Comput. Chem.*, 2010, **31**, 1583–1591.

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20234–20250 | 20249

127 M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys and W. J. Stevens, *J. Phys. Chem. A*, 2001, **105**, 293–307.

128 A. J. Thorvaldsen, K. Ruud, K. Kristensen, P. Jørgensen and S. Coriani, *J. Chem. Phys.*, 2008, **129**, 214108.

129 M. Ringholm, D. Jonsson and K. Ruud, *J. Comput. Chem.*, 2014, **35**, 622–633.

130 D. H. Friese, M. T. P. Beerepoot, M. Ringholm and K. Ruud, *J. Chem. Theory Comput.*, 2015, **11**, 1129–1144.

131 N. H. List, M. T. P. Beerepoot, J. M. H. Olsen, B. Gao, K. Ruud, H. J. Aa. Jensen and J. Kongsted, *J. Chem. Phys.*, 2015, **142**, 034119.

132 M. F. Iozzi, B. Mennucci, J. Tomasi and R. Cammi, *J. Chem. Phys.*, 2004, **120**, 7029–7040.

133 C. Steinmann and J. Kongsted, *J. Chem. Theory Comput.*, 2015, **11**, 4283–4293.

134 K. Coutinho and S. Canuto, *Adv. Quantum Chem.*, 1997, **28**, 89–105.