Computational and Theoretical Chemistry 1040–1041 (2014) 347–359

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

Modeling environment effects on pigment site energies: Frozen density embedding with fully quantum-chemical protein densities

Albrecht Goez [a], Christoph R. Jacob [b,*], Johannes Neugebauer [a,*]

[a] Theoretische Organische Chemie, Organisch-Chemisches Institut and Center for Multiscale Theory and Computation, Westfälische Wilhelms-Universität Münster, Corrensstraße 40, 48149 Münster, Germany
[b] Center for Functional Nanostructures and Institute of Physical Chemistry, Karlsruhe Institute of Technology (KIT), Wolfgang-Gaede-Str. 1a, 76131 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 19 January 2014
Accepted 10 February 2014
Available online 25 February 2014

Keywords:
Protein-pigment interactions
Protein densities
Subsystem density-functional theory
Embedding methods
Time-dependent density-functional theory

ABSTRACT

The recently developed 3-FDE method (Jacob and Visscher, 2008) combines ideas from the Molecular Fractionation with Conjugate Caps (MFCC) method (Zhang and Zhang, 2003) with Frozen Density Embedding (FDE) (Wesołowski and Warshel, 1993) and is thus able to produce fully quantum-chemical electron densities of entire proteins. In contrast to the original FDE method, 3-FDE facilitates a fragmentation into covalently bound subunits. We apply the method for the first time to the calculation of excitation energies, where we use the Fenna–Matthews–Olson (FMO) pigment–protein complex as a test case. Several technical and conceptual parameters for the preparation stages are tested and a robust protocol for this type of embedding is established. We present calculations of excitation energies of the individual pigments (site energies) under the influence of full protein densities obtained with different settings and compare them to the results from a simple point charge model. Our results indicate that 3-FDE is a well-suited method for the description of excitation energies within density-based embedding.

1. Introduction

Photosynthesis is one of the most important processes in nature and a prerequisite for almost all life on earth. A prototypical example of a pigment–protein complex occuring in photosynthesis is the Fenna–Matthews–Olson (FMO) protein, which is found in green sulfur bacteria. It acts as an antenna complex that transfers excitation energy from the peripheral regions to the photosynthetic reaction center, where the actual conversion of light to chemical energy takes place [1]. Since the experimental determination of site energies [2] and the discovery of long-lived coherences in this complex [3,4], many theoretical studies aimed at modeling the energy transfer dynamics on the basis of electronic structure calculations (see, e.g., the two recent reviews in Refs. [5,6]). In all these studies, the key quantities are the individual excitation energies of the bacteriochlorophyll (BChl) *a* pigments under the influence of the protein environment (site energies) and the excitonic couplings between them. Recently, an interesting comparison to earlier work by Schulten, Kleinekathöfer and co-workers [7] was published by Shim et al. [8], wherein it was demonstrated that the details of the electronic structure method, which is used to obtain the optical properties, determine the simulated dynamics to a large extent. It is therefore most desirable to be able to carry out reliable first-principles calculations of the pigments including as much of the environment as possible in the quantum-chemical description.

Unfortunately, the size of even the smallest light-harvesting complexes makes a direct calculation of the entire electronic structure practically unfeasible. Numerous studies concerning the FMO complex have therefore resorted to a QM/MM approach, where the environment of the pigments is represented by a point charge distribution (see, e.g., Refs. [8–11]) or continuum models such as the polarizable continuum model (PCM) [12,13] (often in its integral equation formalism [14–16]) or the conductor-like screening model (COSMO) [17–19], which can mimick at least the bulk effects of a protein environment on a solute. However, it is obvious that such approaches cannot capture effects of a quantum-mechanical nature, but solely coarse-grained electrostatic influences. A more accurate way to model the influence of the surroundings on the pigments is provided by density-based embedding approaches, in which also the environmental degrees of freedom are calculated from first principles. Examples in this field are the Fragment Molecular Orbital method by Kitaura et al. [20–22], which can be regarded as a truncated incremental scheme, and the Effective Fragment Potential method by Gordon and co-workers [23,24], where the total interaction energy of an

* Corresponding authors. Tel.: +49 721 608 48032 (C.R. Jacob), +49 251 83 33241 (J. Neugebauer).
E-mail addresses: christoph.jacob@kit.edu (C.R. Jacob), j.neugebauer@uni-muenster.de (J. Neugebauer).

http://dx.doi.org/10.1016/j.comptc.2014.02.009
2210-271X/

active subsystem is approximated as the sum of the interaction energies of individual environmental subsystems with the embedded one, which partly have to be parametrized. Both schemes have been applied to obtain excitation energies and optical properties of embedded subsystems [25–30]. The two methods have also been combined with each other [31].

An alternative to the aforementioned schemes is Subsystem Density Functional Theory [32] and the related Frozen Density Embedding (FDE) formalism [33] (for details, see the recent review in Ref. [34]). Strictly speaking, one has to distinguish FDE, where only one subsystem's density is optimized in a given, fixed background density, and subsystem DFT, where all subsystem densities are optimized in the environmental density given by all other subsystems. But for the reasons discussed in full detail in Ref. [34], we use the acronym FDE as a general term for both methods in the following. In contrast to physically motivated embedding methods, it is an approximation to DFT free of system-specific parameters and can be considered exact in the limit of exact exchange–correlation and non-additive kinetic energy functionals. FDE in combination with Time-Dependent Density Functional Theory (TDDFT) [35] is already well-established as a useful tool for calculations of excited-state properties of light-harvesting complexes [36,37]. However, one drawback of FDE within standard approximations is its failure to describe fragments which are connected through covalent bonds. This is a well-known problem and can be traced back to the absence of a suitable approximation for the non-additive kinetic energy functional (and thus, the corresponding component of the embedding potential). Possible solutions for this problem are offered by potential reconstruction schemes [38–41]; however, in general these methods also require a calculation of the supermolecular complex and are therefore not applicable for the present purpose. Past attempts to calculate the effects of a large, single covalent unit (e.g., a protein) on the optical properties of chromophores have therefore necessitated the use of fragmented model structures (see, e.g., Ref. [42]). In these schemes, some connecting groups are removed to split the protein into a few peptide chains. The remaining fragments are saturated to generate a number of environmental subsystems which are large, but still small enough to be treated in a conventional Kohn–Sham DFT calculation. Even if most of the protein is kept intact in this way, this procedure naturally forces one to alter the structure of the protein.

One way to avoid the calculation of subsystems with dangling bonds without resorting to such disconnected-environment models is the method of Molecular Fractionation with Conjugate Caps (MFCC) pioneered by Zhang and co-workers [43]. Although the method in its original form was intended for the calculation of energy terms only, it was soon extended to densities and related quantities [44]. In the MFCC scheme, the protein is fragmented by cutting covalent bonds, but every fragment is then saturated with capping groups in order to circumvent the problem of dangling bonds. To reconstruct the density of the supermolecule, combinations of the individual capping groups are also calculated and their density is subtracted from the sum of capped fragment densities. This is an approximation and leads to density errors, especially in the regions around the cut bonds [45,46]. Based on an earlier idea by Casida and Wesołowski [47], Jacob and Visscher combined the MFCC scheme with the FDE method to improve the quality of the reconstructed density. The resulting approach was termed the 3-FDE method [45,46]. Similar to regular FDE, an embedding potential is constructed from the densities of the individual subsystems (in this case the capped protein fragments). The electronic structure of each subsystem is then relaxed in the potential resulting from all the other fragments in so-called freeze-and-thaw cycles [48]. The 3-FDE method has only been used to calculate ground-state properties such as densities and dipole moments so far. In this work, we apply the MFCC and 3-FDE methods to the calculation of excited-state properties of the FMO protein and test a number of model parameters as well as technical factors with regard to performance, robustness and accuracy of the results.

The remainder of this article is organized as follows: In Section 2, the MFCC and 3-FDE methods are presented and their crucial points are briefly explained. In Section 3.1, the commonly used point-charge embedding is applied to generate reference data, which are then compared to results from a disconnected-environment model for a binding pocket in Section 3.2. In Section 3.3, different parameters are tested with regard to their performance in the preparation of the protein MFCC density. Finally, the resulting densities are used in 3-FDE(0) calculations and their effect on the excitation energies is assessed in Section 3.4. Section 3.5 briefly describes some of the difficulties in performing 3-FDE calculations for the whole protein. A summary and conclusions are given in Section 4.

2. Computational methods

In this section, a brief introduction to the MFCC and 3-FDE methods is given. For details and derivations, see Refs. [43,45,46].

2.1. MFCC

The basic idea of the MFCC scheme is as follows: A supersystem is partitioned into smaller entities that are connected through covalent bonds. To saturate the resulting dangling bonds, capping groups are introduced as illustrated in Fig. 1. A separate calculation is then carried out for each capped fragment. Subsequently, neighbouring (conjugated) capping groups are combined to form one cap molecule which is also subjected to a calculation. Finally, the density of the supermolecule is reconstructed by summing up the densities of all capped fragments A and subtracting the densities of the cap molecules B:

$$\rho_{\mathrm{tot}} = \sum_{A}^{K} \rho_{A}^{\mathrm{capped}} - \sum_{B}^{L} \rho_{B}^{\mathrm{cap}} \tag{1}$$

Here, K is the total number of fragments and L is the corresponding number of cap molecules. Of course it is not necessary for the supersystem to consist of only one covalent unit; oligomers of interacting proteins or non-covalently bound cofactors (like the pigments in this work) can be treated just as well. In the latter case, each

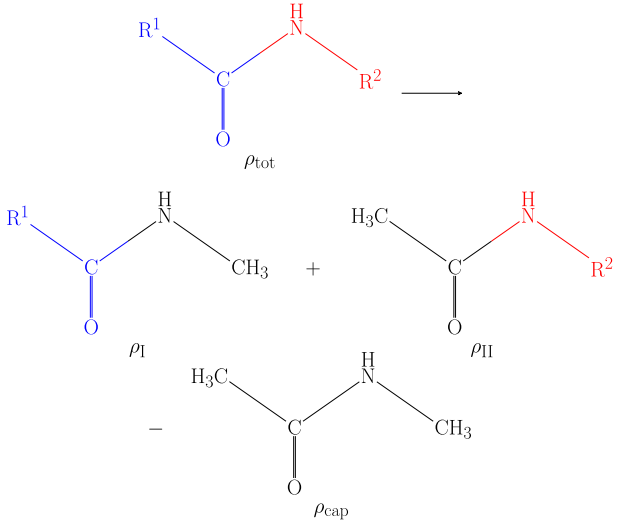

Fig. 1. Schematic representation of the partitioning and capping in the MFCC and 3-FDE methods.

cofactor can be considered as an individual subsystem without any capping groups.

2.2. 3-FDE

The 3-FDE method is based on the MFCC scheme and requires a complete calculation of all capped fragments and capping molecules beforehand. In the first step, an embedding potential for each fragment (with density ρ_I) is constructed from the MFCC densities of all the other fragments. This results in the Kohn–Sham-like equations [45]

$$\left[-\frac{\nabla^2}{2}+v_\mathrm{eff}^{\text{3-FDE}}[\rho_\mathrm{I},\rho_\mathrm{II}](\mathbf{r})\right]\phi_i^{(\mathrm{I})}(\mathbf{r})=\epsilon_i\phi_i^{(\mathrm{I})}(\mathbf{r}), \tag{2}$$

where ρ_I denotes the density of the currently active subsystem and ρ_II represents the frozen density containing contributions from all the other capped fragments (i.e., ρ_II is calculated from Eq. (1), with the sums over A and B excluding the active fragment). This equation is very similar to the standard FDE case [33]. However, one specialty of the 3-FDE method is the form of the effective embedding potential $v_\mathrm{eff}^{\text{3-FDE}}$. To guarantee that the total density is non-negative at all points in space when the cap density is subtracted, a constraint is introduced to ensure that the density at each grid point in the cap region of the capped fragment is equal to the one in the subtracted cap molecule. To this end, the regular embedding potential is replaced by a special cap potential within a certain region around the cap atoms (in the current implementation, this area is defined by all grid points within $3\,a_0$ of any cap atom that are not closer to a non-cap atom). The embedding potential thus is of the following form [45]:

$$v_\mathrm{eff}^{\text{3-FDE}}[\rho_\mathrm{I},\rho_\mathrm{II}](\mathbf{r})=\begin{cases} v_\mathrm{eff}^{\mathrm{KS}}[\rho_\mathrm{I}](\mathbf{r})+v_\mathrm{eff}^{\mathrm{emb}}[\rho_\mathrm{I},\rho_\mathrm{II}](\mathbf{r}) & \text{for } \mathbf{r}\notin V_\mathrm{I}^{\mathrm{cap}} \\ v_\mathrm{cap}(\mathbf{r}) & \text{for } \mathbf{r}\in V_\mathrm{I}^{\mathrm{cap}} \end{cases} \tag{3}$$

In the first line, $v_\mathrm{eff}^{\mathrm{KS}}[\rho_\mathrm{I}](\mathbf{r})$ is the regular Kohn–Sham potential of the active subsystem density and $v_\mathrm{eff}^{\mathrm{emb}}[\rho_\mathrm{I},\rho_\mathrm{II}](\mathbf{r})$ is the effective embedding potential due to the presence of the frozen capped fragment densities. However, it must be noted that all terms resulting from the frozen caps are excluded from the latter term.

The cap potential $v_\mathrm{cap}(\mathbf{r})$ is constructed in an iterative fashion. As a starting point, the potential of the corresponding atoms in the isolated cap molecule is used and the Self-Consistent Field (SCF) procedure starts. As soon as the capped fragment is converged under the current cap potential and the embedding potential caused by the other fragments, the cap potential is updated to change towards reproducing the desired density. Two mechanisms are available for the update, with the error in the number of electrons being the decisive factor as to which one is used. If the number of electrons is far from the correct value (more than 0.05 electrons in the current implementation), a direct shift is equally applied to the potential at all grid points. If the error falls below that threshold, the cap potential is updated according to the procedure by van Leeuwen and Baerends (LB step) [49]. As soon as a cap potential is converged, it is kept fixed. When all caps present in the fragment are converged, one last SCF cycle is passed and an updated density for the whole capped fragment is obtained. Details concerning the implementation can be found in Ref. [46].

Usually, each fragment density is updated once in the embedding potential due to the MFCC density resulting from all other fragments, which is denoted as 3-FDE(0) (see Section 3.4). When all capped fragments have been relaxed within this initial potential, freeze-and-thaw cycles are employed, i.e., the fragment densities are updated in the embedding potential of all the other fragments in turns. In this case, the embedding potential changes after every fragment calculation, not only after a complete cycle. This procedure is called 3-FDE(n) in the following.

2.3. Structure preparation

The structure of the Fenna–Matthews–Olson complex used in all calculations presented here was prepared by a Molecular Dynamics (MD) equilibration run. We used a setup closely related to the one by Olbrich et al. [50]. In compliance with that setup, the initial coordinates were those of one monomer of the seven-pigment X-ray structure by Tronrud et al. [51] taken from the RCSB Protein Data Bank (PDB ID: 3ENI). Hydrogen atoms were added and the complex was then immersed in water. After a short energy minimization, a 15 ns equilibration run was carried out and the resulting final structure was extracted from the trajectory. We used the program NAMD2 [52] with the CHARMM27 [53,54] parameters for protein and water and a specially developed set of parameters for the BChl *a* pigments [55,56]. Details concerning the setup of the MD simulation can be found in the Supporting Information. After the equilibration, the system was stripped of water and ions to yield a structure with 6505 atoms to be used for the following quantum-chemical calculations. The resulting structure is shown in Fig. 2.

It should be noted that this arbitrary snapshot merely serves as an example in this pilot study. The purpose of our present efforts is to establish a robust protocol for the calculation of excitation energies in pigment–protein complexes using 3-FDE. Strictly speaking, there is thus no need to produce an entire MD trajectory. However, since we are aiming at establishing the methodology presented here for subsequent studies on dynamical effects in light-harvesting proteins, we use a setup that in principle allows sampling over many snapshots.

It has to be kept in mind that the structure from the simulation was used directly without any post-optimization procedure. This can lead to two problems: First, there is always an uncertainty concerning how realistically the used parameters can model the molecule. For instance, it was found that in the present structure, the phytyl double bond in all pigments is significantly non-planar. It is known that the phytyl tail does not affect the excitations of the pigments to a large degree [57,37], but this serves as an example on how carefully MD parameters have to be checked. Second, if the structure in a quantum-chemical calculation differs much from the equilibrium structure that would be obtained with the first-principles method at hand, SCF convergence problems can arise. At the moment, however, parametrized MD simulations are still the only way to obtain enough conformations of a protein like the FMO complex to reliably sample properties like excitation energies.

2.4. Electronic structure calculations

All calculations were carried out with the Amsterdam Density Functional (ADF) program suite [58,59] and the PBE exchange–correlation functional [60,61]. If not stated otherwise, a DZP basis set with a 1*s* frozen core for C, N and O, and a [1*s*, 2*s*, 2*p*] frozen core for S and Mg was used. All MFCC and 3-FDE calculations were set up with the aid of the PYADF scripting framework [62]. We use the FDE implementation in ADF [45], including the extension to (uncoupled) excited states [35,63]. If not stated otherwise, all given excitation energies correspond to the excitation that is dominated by the transition from the highest π orbital to the lowest π^* orbital. This is expected to be the BChl Q_y excitation [1]. In some cases, an unambiguous assignment was not possible and a second value is given in brackets in the corresponding tables.

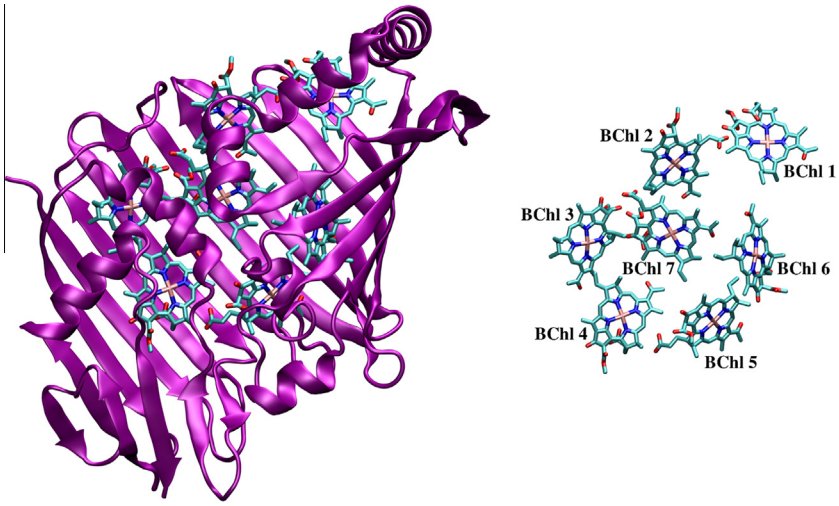

Fig. 2. Cartoon representation of the MD-equilibrated structure of a monomer from the Fenna–Matthews–Olson complex with its seven BChl pigments. Figure created with VMD [74,75].

3. Results

In this section, the results of the different methods used to evaluate the effect of the surroundings on the optical properties of the pigments are presented.

3.1. Classical embedding

As the most simple scenario, the seven BChl molecules were extracted from the generated snapshot and calculated individually *in vacuo*. The resulting excitation energies are presented in Table 1. Furthermore, the complete array of seven pigments was subjected to an uncoupled FDE (FDEu) calculation with one freeze-and-thaw cycle, in which every pigment was treated as an individual subsystem. In such an approach, only the polarization (no excitonic interaction) by the neighbouring subsystems is taken into account. In the present case, the polarization effect of the uncharged pigments is expected to be minor. However, an FDEu calculation is a prerequisite for a so-called coupled FDE (FDEc) calculation [64], which makes it possible to obtain excitonic coupling parameters. The latter might be the target of a subsequent study. For completeness, the FDEu results are also given in Table 1.

The computationally most efficient way to include polarization of the pigments by the surrounding protein is to model the latter as a point charge distribution. This is easily possible for the given scenario, as the force field charges from the MD simulation can be used directly. The number of point charges hardly affects the CPU demands of the calculations, so that all charges (protein, water, ions, other pigments) can easily be included. However, it has to be kept in mind that these charges are not parametrized to be used in QM calculations, but to reproduce correct interactions between molecular building blocks in an MD simulation. The calculated excitation energies are also presented in Table 1.

In all cases, the changes in the excitation energies are fairly small and reach maximum values of around 0.02 eV. For the comparison between isolated pigments and the seven-pigment array, this was expected (see above) and the shifts for the point charge distribution are of the same order of magnitude as in an earlier study that used the seven-pigment model and point charges from the CHARMM27 force field [65]. The results will be compared to those from 3-FDE calculations in Section 3.4.

Table 1
Excitation energies of the individual pigments in the FMO complex and the seven-pigment array with and without a background point charge (PC) distribution resembling the protein. All values in eV.

Pigment	Isolated	Array (FDEu)	Isolated (PC)	Array (FDEu, PC)
1	1.858	1.876	1.850	1.850
2	1.869	1.872	1.850	1.855
3	1.826	1.823	1.807	1.806
4	1.815	1.815	1.812	1.804
5	1.887	1.876	1.878	1.881
6	1.854	1.848	1.841	1.839
7	1.887	1.904	1.893	1.889
		(1.811)		

3.2. Disconnected-environment models

Following our previous work [42], a first quantum-chemical model of the pigment surroundings was constructed by setting up binding-pocket models for the individual pigments. Each binding pocket included the amino acid residues and water molecules with at least one atom closer than 8 Å to the pigment's central Mg atom. Contrary to Ref. [42], the resulting oligopeptides were not saturated with hydrogen atoms, but capped with neutral amino acid termini from the official CHARMM27 topology files [53,54] and no subsequent optimization was carried out here. Note that no other pigments were allowed in the calculations, as the purpose of the setup was to study the pigment–protein interactions. As an example, BChl 3 in its 8 Å binding pocket is shown in Fig. 3.

The binding pocket models obtained in this way were then calculated by four different strategies of increasing complexity:

(a) Isolated pigments
For reference, each pigment was extracted from the structure and calculated *in vacuo*. This setup is identical to the isolated calculations in Section 3.1.
(b) Point charges
In this case, only the pigment itself was treated quantum-mechanically, whereas the capped amino acid fragments were not modeled explicitly, but as classical point charges.
(c) FDE
The model of the pigment in its binding pocket was partitioned into individual subsystems, where the pigment as well as each covalently saturated protein fragment and each water molecule constituted one subsystem. The electron densities of

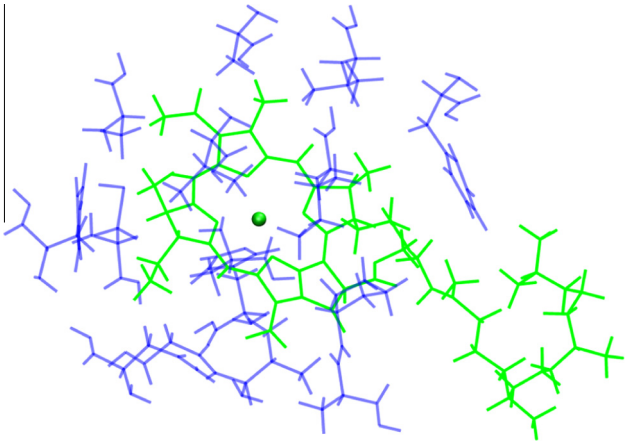

Fig. 3. 8 Å binding pocket around BChl 3 from the FMO protein. The pigment is colored green, while the protein surroundings are blue. Figure created with VMD [74]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the subsystems were then relaxed in two consecutive freeze-and-thaw cycles and the excitation energies were calculated only for the pigment subsystem.

(d) Supermolecular calculation

Owing to the relatively small size of the overall 8 Å binding pockets (about 370 atoms on average), a supermolecular reference calculation was still feasible and was carried out for each of them.

It must be noted that the minimal environment in this setup is not an accurate representation of the whole pigment–protein complex, since long-range electrostatic interactions between the pigments and remote parts of the protein are neglected. However, the purpose of this test is to benchmark the point charge approach and FDE against supermolecular calculations. Unfortunately, some of the supermolecular calculations for the artificially cut-out binding pockets could not be converged, which is probably due to the fact that the structure was obtained from a parametrized simulation (see Section 2.3). However, no significantly different behaviour is expected for these structures concerning their excitation energies.

Table 2
Excitation energies of the individual pigments in their binding pockets calculated with different approximations to the environmental interaction for two basis sets. A dash denotes that the respective calculation could not be converged. All values in eV.

Pigment	Isolated	Point charges	FDEu	Supermolecular
DZP				
1	1.858	1.854	1.835	1.794
2	1.869	1.851	1.835	1.799
3	1.826	1.837	1.815	1.768
4	1.815	1.796	1.793	1.748
5	1.887	1.859	1.871	1.842
6	1.854	1.848	1.824	1.795
7	1.887	1.899	1.880	–
TZP				
1	1.795	1.797	1.797	1.765
2	1.839	1.805	1.788	1.757
3	1.792	1.794	1.782	1.760 (1.716)
4	1.783	1.759	1.757	–
5	1.842	1.838	1.844	1.791
6	1.814	1.800	1.783	1.756
7	1.866 (1.778)	1.848	1.849	–

The results from the different models are presented for two basis sets (DZP and TZP) in Table 2.

When comparing the resulting DZP excitation energies, it can be easily seen that neither a point charge model, nor FDE are able to fully reproduce the shift from the supermolecular calculation. However, in line with these reference calculations, FDE predicts a redshift for all pigments, while the point charge model yields a blueshift for some of them. Naturally, both methods are not capable of describing the differential polarization caused by the excited chromophore, which is expected to be responsible for part of the residual shift. Furthermore, though we consider the supermolecular approach as a reference here, there are arguments for actually favouring a subsystem approach. One reason is that FDE as used here is inherently free from the Basis Set Superposition Error, since all basis functions are confined to their respective fragments. The supermolecular calculations of the present binding-pocket models most likely suffer from this problem, so that the resulting excitation energies are artificially low. This is supported by the fact that the shifts between supermolecular and model calculations for all pigments are much smaller in the TZP case. Furthermore, it was shown that the well-known self-interaction error in DFT is reduced in a subsystem formalism [66]. In conclusion, we note that in many cases FDE can improve the excitation energies compared to a point charge model, even though the differences may be small.

3.3. Convergence of individual fragments in the MFCC framework: role of embedding in preparation steps

The basis for each 3-FDE calculation is a complete MFCC run for the system, from which the initial frozen density ρ_{II} is constructed (see Section 2.2). We therefore decided to conduct a systematic test to analyze how different computational settings affect the efficiency of the calculation and the convergence behaviour for the individual capped fragments. From now on, we use the whole structure of the FMO complex that was obtained from the MD simulation (see Section 2.3) as our test system. Starting from default settings (see below), the following parameters were adjusted: Basis set, fragmentation pattern, environment model and miscellaneous SCF settings. In each of the following subsections, only one setting was varied while the others were kept constant.

As a starting point, a systematic fragmentation pattern with 10 amino acids per fragment was chosen, and the calculations were carried out with a DZP basis set and ADF default settings for the SCF. The environment was modeled to be pure water in the framework of COSMO [17–19] (the requirement of an environment model already in the preparation steps is detailed below). If not stated otherwise, these defaults were used in all the following tests.

- Computational settings
  A variety of SCF settings were tested to accelerate SCF convergence. These included the mixing parameter in a simple damping scheme, the number of expansion vectors for the "Direct inversion in the iterative subspace" (DIIS) method [67] as well as its starting criteria, and the usage of the Augmented Roothaan–Hall DIIS (ADIIS) scheme [68]. However, the optimal SCF parameters that emerged from this test resembled the ADF default settings (see Supporting Information) so closely that it was decided to use the latter for the sake of simplicity.
- Environment model
  Converged densities of the individual building blocks of the system are required as a prerequisite for the 3-FDE procedure. The more complicated these fragments are, e.g., in terms of charged side chains, the higher the chance that their isolated densities cannot be converged *in vacuo*. Thus it can be necessary to already introduce a model of the surroundings for the individual fragments in the preparation stage. In the following, we

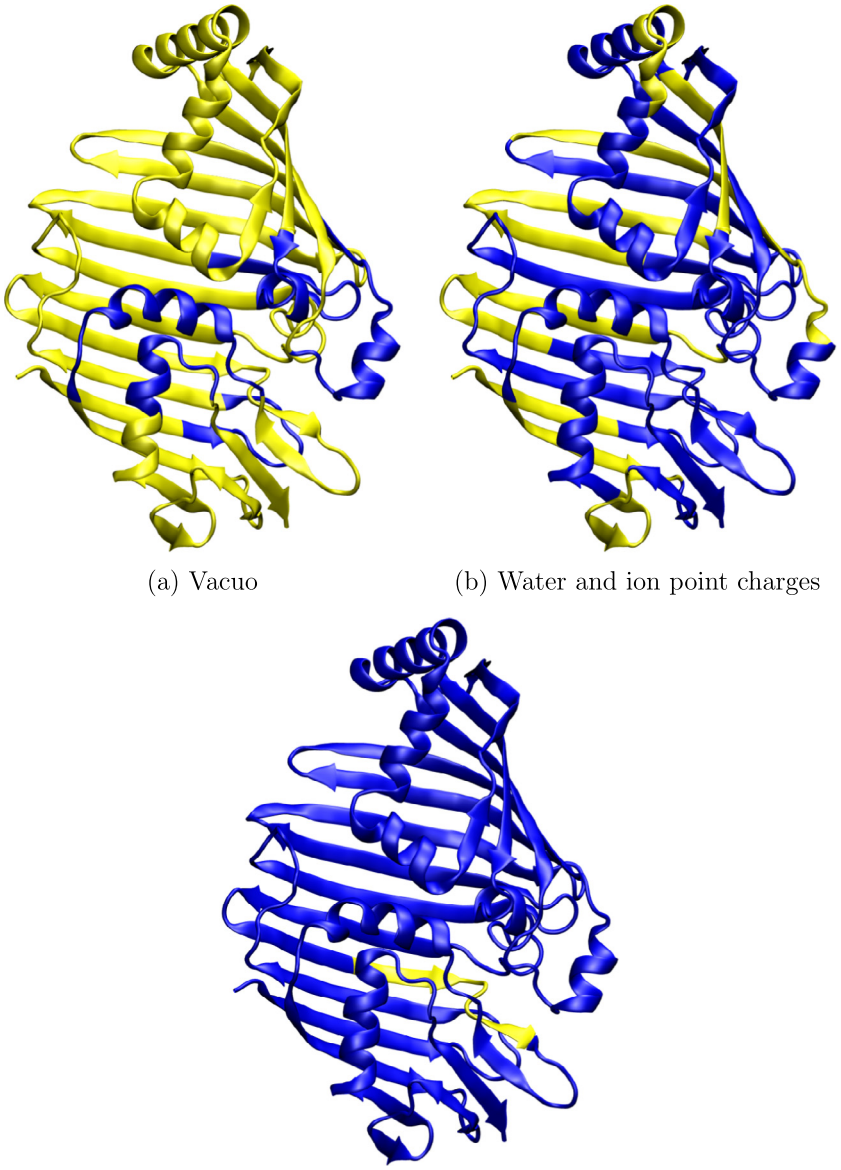

Fig. 4. Convergence of individual 10-amino-acid fragments in a cartoon representation of the FMO protein for different amounts of environment point charges. Yellow residues belong to non-converged fragments, blue residues are converged. Figure created with VMD [74,75]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

refer to this type of environment model as "preparational embedding" in order to distinguish it from any embedding in the subsequent 3-FDE excitation energy calculations, which will be referred to as "outer embedding".

Different environmental models for the individual fragments of the pigment–protein complex were tested. Starting from calculations *in vacuo*, the model was gradually expanded to include different sections of the background point charges from the MD simulation. As the largest point charge model, all water molecules, ions and protein residues (apart from the fragment in question) were taken into account. Furthermore, the usage of the polarizable continuum model COSMO was explored.

- *Vacuo*
  When the protein fragments are isolated from the snapshot and no environment whatsoever (besides the capping groups) is used, most fragment calculations suffer from severe convergence problems. In a test calculation with the default settings (10 residues per fragment), 28 of 36 fragments did not converge within 100 SCF cycles. This suggests that stabilizing interactions between specific amino acid fragments and their surroundings play an important role for the overall electronic structure of the protein. A graphical representation of the non-converged fragments is given in Fig. 4, panel (a).
- Water and ion point charges
  If all surrounding water molecules and solvated ions from the MD simulation are included as a background charge distribution, convergence is substantially improved compared to the isolated situation. However, 11 out of 36 fragments still have convergence problems, which points to the fact that not only the interaction with the solvent on the outside of the protein is important, but also electrostatic interactions between spatially close (although not necessarily neighbouring) amino acids stabilize the electronic structure. The convergence behaviour for this scenario is depicted in Fig. 4, panel (b).

– Water, ion and protein point charges
To include the electrostatic effect of water, ions and neighbouring protein residues, all atoms from the MD-equilibrated structure were considered in the fragment calculations. However, the calculated fragment itself as well as the neighbouring amino acid on each side of the fragment were deleted from the distribution to avoid overlap of the capping groups with some of the point charges. When using this approach with the aforementioned parameters, all fragments but one converge easily. Upon inspecting the structure of the non-converged fragment, one deprotonated ASP residue turned out to feature a strong interaction with a protonated LYS residue through a H_3N^+–COO^- ion bridge (see Fig. 5). However, the LYS residue was not included in the point charge distribution, as it is one of the directly neighbouring amino acids that partly overlapped with the capping group of the fragment in question.
A calculation in which only the LYS backbone charges overlapping with cap atoms of the neighbouring fragment were deleted and the ones representing the LYS side chain were kept, could be converged without any problems. This example stresses the importance of individual interactions between the protein fragments and the necessity to model them as realistically as possible.
Unfortunately, with this kind of environment model, convergence is hampered again when going to large fragment sizes (beyond 30). For both our models with larger fragment size (40 and 50 amino acids per fragment), about half of the fragments could not be converged.
– COSMO
When applying COSMO with parameters corresponding to water ($\epsilon = 78$) to the individual calculations of the fragments in the MFCC framework, all fragments converge without problems, even for fragmentation sizes of 40 or 50 amino acids per fragment. However, a continuum model for water is certainly not a very realistic depiction of the individual fragments' environment, since several of them are buried deeply in the protein and have thus very little or no contact to actual water. When setting the relative permittivity to $\epsilon = 4$, which has been suggested as a suitable value to mimick a protein environment [69], no convergence problems are encountered either. Continuum models are therefore very helpful in constructing initial densities, which can then be further refined using the 3-FDE scheme. Still, it should be kept in mind that a continuum model cannot incorporate specific interactions between amino acids in different fragments.

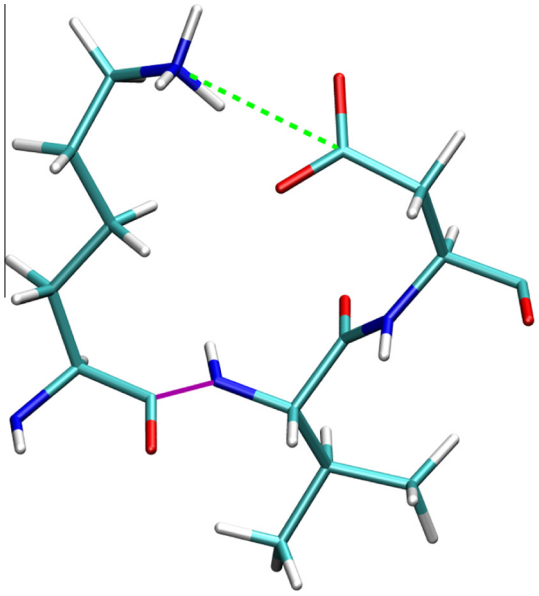

Fig. 5. Stabilizing interaction between a protonated amino group and a deprotonated carboxyl group of a second-nearest neighbour amino acid side chain. The solid purple line denotes the connection of one fragment to the next one. The dashed green line represents the ion bridge. Figure created with VMD [74]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In conclusion, we find that both continuum models and the inclusion of background charges from the MD simulation are promising ways to obtain initial densities for the isolated fragments. If possible, we suggest to use a point charge distribution of as much environment as possible in this preparation step to mimick individual stabilizing interactions most closely. An embedding potential constructed from the resulting MFCC densities is then already pre-polarized and will probably change much less in the subsequent 3-FDE calculation than one constructed from non-polarized fragments. This should reduce the number of required freeze-and-thaw cycles. However, if very large protein fragments are used or if no suitable parameters for the point charges are available, it is advisable to use a continuum model for the preparational embedding. It has to be kept in mind, though, that apart from the resulting computational overhead this does not represent a very realistic environment for the individual fragments, as some of them will experience an aqueous environment, while others are surrounded by neighbouring amino acids.

- Basis set effects
Several standard basis sets from the ADF library were compared with regard to their applicability to the problem at hand. The use of a minimal SZ basis was explored first in order to simplify the construction of the protein density as much as possible. With this choice, however, 16 out of 36 fragments failed to converge. Various additional adjustments of the SCF parameters could not change this outcome. Apparently, this minimal basis does not offer sufficient flexibility for the electronic structure of the present protein environment.
When going beyond a minimal basis, convergence presents no problems if at least a double-zeta basis with one set of polarization functions (DZP) is used. It should be mentioned that the construction of a DZP-quality protein density with the MFCC scheme took only around 10 h on a 12-core Intel Xeon machine (2.67 GHz) in the present case. Enlarging the basis even more to TZP or TZ2P increases the total wall times to 16 or 22 h, respectively. No convergence problems were encountered here either. The influence of the basis set on the calculated excitation energies will be discussed in Section 3.4.
- Fragmentation patterns
The choice of fragmentation pattern is critical for the MFCC/3-FDE procedure, as it represents a trade-off between accuracy and computational cost. The smaller the fragments are chosen, the lower the overall computational demands will be. With small fragments, however, many bonds have to be cut, which leads to larger errors in the density and the derived quantities. Different systematic fragmentation patterns have been tested to analyze the dependence of the convergence on the fragment size *N*. Interestingly, calculations with different values of *N* between 1 and 50 exhibit no convergence problems at all, as long as a sufficiently large basis set (DZP or larger) is used. When using a minimal SZ basis, very small fragments can be converged, but already fragments with a size of 10 amino acids each present serious convergence problems. We did not try to increase the fragmentation size further than 50, since the calculations naturally become more and more time-consuming (see below). Furthermore, common DFT weaknesses such as

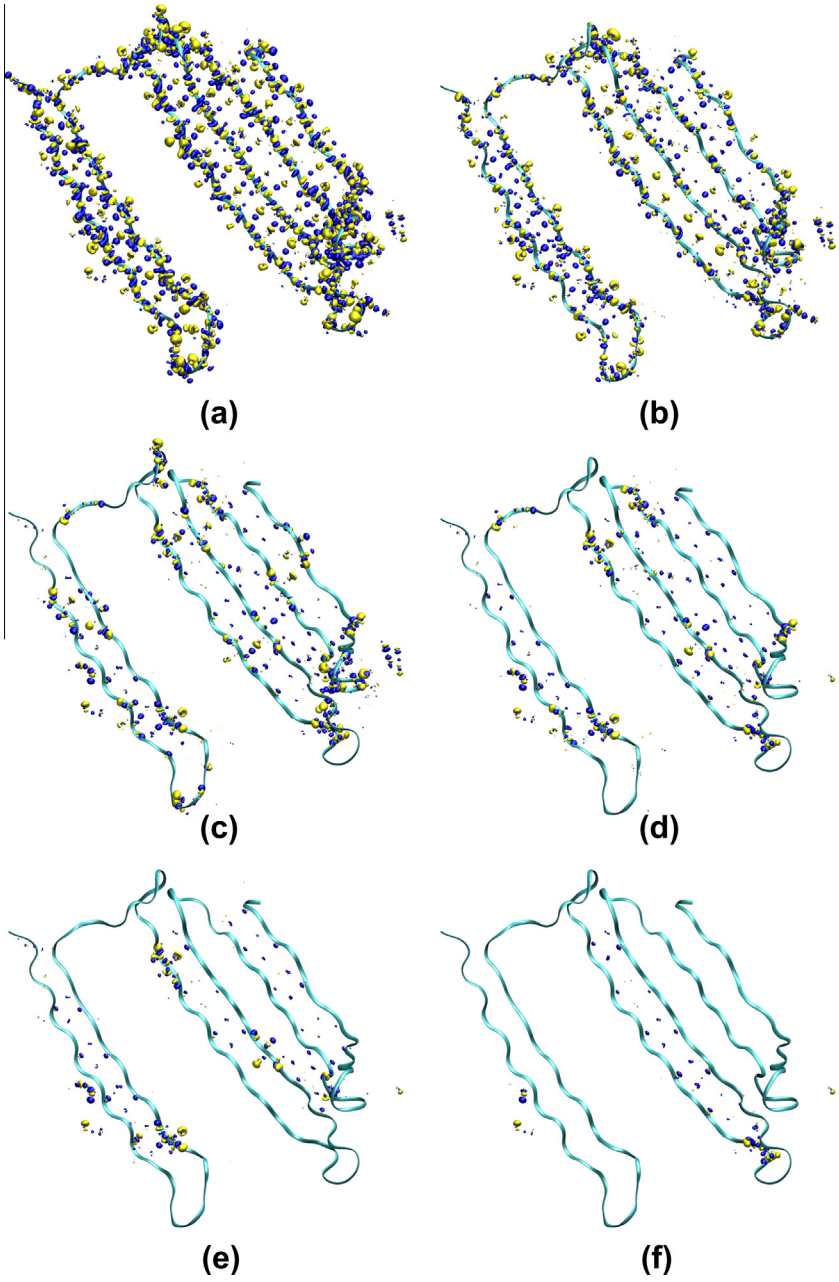

Fig. 6. Difference density plots for the first 100 amino acid residues from the FMO complex with different systematic fragmentation patterns of (a) 1, (b) 2, (c) 5, (d) 10, (e) 20, and (f) 50 residues per fragment (Isovalue: 0.0025). Figure created with VMD [74,75].

over-delocalization, especially concerning multiple charged amino acids in one fragment, can partially be avoided by restricting the electrons to certain fragments in the calculation [46,66]. Such advantages would be forfeit when increasing the fragmentation size to large numbers.

To visualize the density errors brought about by the partitioning, we created some difference density plots. Since a supermolecular calculation of the whole FMO protein is not feasible, we extracted the first 100 amino acid residues from the total structure. This subunit of 1510 atoms could be converged in a reasonable amount of time with a DZP basis set. We then applied systematic fragmentation patterns for $N = 1$ (single residues) up to $N = 50$ (only two fragments in total) to the structure and computed the density differences. The resulting plots are shown in Fig. 6.

As expected, very large density errors arise with the smallest fragmentation size. Already with $N = 5$, these are much smaller and concentrate in the regions around the cuts. With the largest fragmentation size, a very good match to the supermolecular density is obtained. Apart from the single fragmentation site, some more errors can still be observed relatively far away from it, mainly due to hydrogen bonds of the β-sheet structure. However, these are very small.

As mentioned above, the computer time spent for constructing an MFCC density with our default fragmentation size of 10 residues per fragment and a DZP basis set was approximately 120 CPU hours. For the smallest tested N of 1 the calculations took only 36 CPU hours, while for $N = 5$ this time increased only a little to 50 CPU hours. Calculations using large fragments with N of 30 and 50 required 330 and 580 CPU hours, respectively. Although such calculations are still feasible for single structures, the effort spent is too large to be useful for sampling along a trajectory.

Apart from systematically fragmenting the protein, an adaptive pattern was tested in which the number of charged residues

was limited to one or two per fragment. For the former, this resulted in 83 fragments of between one and seventeen amino acids with a mean size of 4.3 residues per fragment, for the latter it yielded 42 fragments of between two and 24 residues with an average size of 8.5 residues. Both setups did not present any convergence problems either. The effect of the fragmentation pattern on the excitation energies is discussed in the next section.

3.4. The influence of a QM environment on excitation energies

Our primary motivation for the usage of the MFCC/3-FDE method is the construction of an approximate protein density to study the protein's effect on the electronic properties of pigments in pigment–protein complexes. Depending on the level of self-consistency to which the environment density was constructed, different cases can be distinguished:

- MFCC: Since the pigments in our setup are neither cut nor capped, their densities (and thus, their electronic properties) are identical to those from calculations of the isolated pigments. Therefore, the MFCC scheme by itself cannot improve the description of the excitation energies.
- 3-FDE(0): The first step towards modeling the effect of the protein on the chromophores consists of relaxing the individual subsystem densities in the embedding potential constructed by all the other subsystem densities (i.e., the sum of MFCC densities of all fragments as in Eq. (1), but with all contributions of the active subsystem removed). However, the relaxed densities are not used to update the embedding potential until the first cycle is complete. In compliance with Refs. [45,46], this is denoted as 3-FDE(0).
  This scheme offers a first glance at the expected changes in the optical properties of the chromophores and is computationally not very demanding, since it is not even necessary to walk through a single complete 3-FDE cycle. As long as the protein density is not of direct interest, only the changes in the pigment densities need to be evaluated and used as a basis for excitation energy calculations. In fact, the use of unrelaxed environment densities often (but not always) gives surprisingly good results, partially due to error cancellation effects. For a more detailed discussion, see, e.g., Refs. [70–72].
- 3-FDE(n): Further improvement requires to complete the 3-FDE(0) cycle and successively update the individual fragment densities using freeze-and-thaw cycles. Here, n denotes the number of complete cycles in which every fragment was relaxed in the new embedding potential constructed from the other fragment densities. The outcome of a 3-FDE(n) calculation generally depends on the order of going through the subsystems. However, as soon as self-consistency is reached, this dependence should be lifted. Our first attempts to obtain also 3-FDE(n) excitation energies for the FMO protein revealed several unforeseen challenges, which are discussed in Section 3.5. The remainder of this section is dedicated to describing the effects of different computational setups on the 3-FDE(0) excitation energies.

The following test calculations were carried out analogously to the tests in Section 3.3, i.e., with the default settings listed there, if not stated otherwise. Apart from the embedding potential, which only models the protein itself, an outer embedding model to describe the solvation shell around the protein should be applied. As explained in more detail in Section 3.5, the current COSMO implementation in ADF is not optimized for 3-FDE calculations and led to technical problems in our applications. All calculations were therefore carried out with a background point charge distribution resembling water and ions from the MD simulation. This choice should not have a large effect on the calculated excitation energies, since the protein (which is modeled by the embedding potential) wraps the pigments almost completely.

3.4.1. Basis set effects

We have evaluated three different basis sets (DZP, TZP, TZ2P) with regard to their description of the excitation energies. In each case, a frozen core approximation was applied to the 1*s* electrons of C, N and O and to the 1*s*, 2*s* and 2*p* electrons of S and Mg. It should be noted that the frozen-core approximation was only applied in order to simplify the construction of the protein density. In principle, one can use a different basis set for the embedded pigments. However, it is desirable to use a consistent description for all fragments if freeze-and-thaw cycles are to be employed in a 3-FDE(n) calculation. The results for the different basis sets are presented in Table 3.

It is observed that the excitation energies with a TZP and TZ2P basis set hardly differ at all (maximum deviation: 0.002 eV), with the exception of BChl 4. In that case, however, the assignment of the excitations is not trivial, as an orbital from the twisted phytyl side chain (see Section 2.3) mixes with the $\pi \rightarrow \pi^*$ transitions in the TZ2P calculation. Additional basis functions are thus not expected to improve the energies very much. In contrast, the DZP results are all blue-shifted by approximately 0.04 eV. However, the relative energies among the individual pigments are completely conserved, with the exception of BChl 1, where the shift is about 0.01 eV larger. Thus, it is already possible to assess the qualitative features of the pigment network with a medium-sized basis set, although the excitation energies are certainly more reliable when a triple-zeta basis set is used.

To investigate the consequences of the frozen core approximation, additional all-electron calculations as well as calculations with an intermediate degree of frozen cores (ADF: "Frozen Core Small") were carried out for the case of the DZP basis set. The corresponding excitation energies are listed in Table 4.

As expected, the usage of a frozen core does not change the resulting excitation energies dramatically. The mean absolute deviation for a "Large" frozen core from the all-electron calculation is merely 0.004 eV and it is thus concluded that a frozen core approximation for all but the valence electrons is justified for the research question at hand.

3.4.2. Fragmentation pattern

3-FDE(0) calculations were carried out for different systematic fragmentation patterns between 1 and 50 amino acids and two special partitioning schemes where the number of charged residues per fragment was limited to one and two, respectively. The resulting excitation energies are plotted in Fig. 7.

For small fragmentation sizes, the largest changes in excitation energy are observed, although these are already fairly small. Interestingly, with fragment sizes of only around 5, all pigment excitation energies are relatively well converged. From a size of 20 on, hardly any changes can be spotted. This weak dependence of the

Table 3
Comparison of the 3-FDE(0) excitation energies of the individual BChl Q_y transitions in the FMO protein with different basis sets. All values in eV.

Pigment	DZP	TZP	TZ2P
1	1.852	1.794	1.794
2	1.856	1.809	1.810
3	1.795	1.758	1.759
4	1.805	1.764	1.754
			(1.772)
5	1.883	1.843	1.845
6	1.844	1.803	1.803
7	1.891	1.861	1.859

Table 4
Comparison of the 3-FDE(0) excitation energies of the individual BChl Q_y transitions in the FMO protein with different frozen core approximations. All values in eV.

Pigment	Large	Small	None
1	1.852	1.854	1.855 (1.812)
2	1.856	1.857	1.854
3	1.795	1.796	1.793
4	1.805	1.806	1.802
5	1.883	1.883	1.882
6	1.844	1.844	1.841
7	1.891	1.892	1.891

Table 5
Comparison of the 3-FDE (0) excitation energies of the individual BChl Q_y transitions in the FMO protein for different solvent models in the MFCC calculations. All values in eV.

Pigment	COSMO	PC
1	1.846	1.852
2	1.856	1.856
3	1.797	1.795
4	1.792	1.805
5	1.875	1.883
6	1.838	1.844
7	1.896	1.891

pigment excitations on the fragmentation size can be interpreted as a quality assessment of the MFCC scheme for the preparation of a frozen density. As it is not expected that the excitation energies will change any further when going to even larger fragmentation sizes up to a hypothetical supermolecular calculation, it can be concluded that at least for the given purpose, the MFCC density is equivalent to a non-fragmented protein density. The two models created with the adaptive partitioning scheme (in order to minimize the number of charged amino acids per residue, see Section 3.3) did not exhibit any special behaviour. Their excitation energies were very similar to those observed with fragmentation patterns of similar average fragment size. This indicates that the fragment size is the decisive parameter and the exact location of the cuts does not play a major role.

3.4.3. Preparational embedding

In this section, the effect of the preparational embedding for the individual fragments is analyzed. The preparation of the MFCC fragments *in vacuo* was not considered any further due to the massive convergence problems. As the most promising approaches, the calculation of the capped fragments with COSMO (water) and in a background point charge distribution resembling the complete protein in water were compared. The results are presented in Table 5. In both cases, the outer embedding model for the 3-FDE(0) calculations was constructed by using the water and ion point charges from the MD simulation.

Apparently, the preparational environment has a small effect on the excitation energies in a 3-FDE(0) step, although the embedding potential seems to be dominated by the intrinsic properties of the protein fragments and not by the way they were prepared. The largest effect occurs for BChl 4, where a shift of 0.013 eV is observed. This is comparable, in the present case, to the shifts between isolated pigment calculations and calculations using a point charge model (see Section 3.1). The pre-polarization due to the preparational embedding with point charges may thus not be considered negligible in general. Our initial preference for a preparational embedding of atomistic nature is therefore reassured,

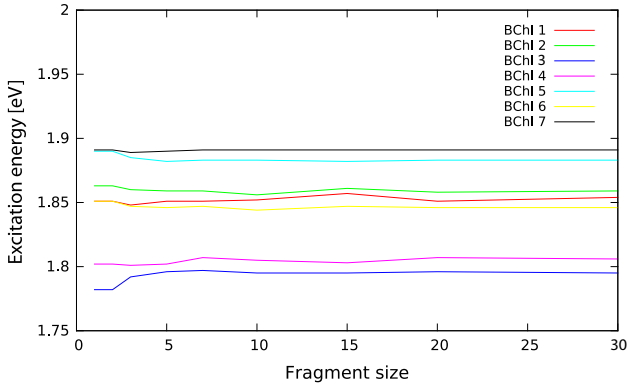

Fig. 7. 3-FDE (0) excitation energy of the Q_y transition for the BChl pigments in the FMO complex for different fragment sizes.

as we expect that less freeze-and-thaw cycles will be required to obtain fully self-consistent excitation energies. Furthermore, in cases where not the properties of non-covalently bound cofactors, but those of the protein are of interest, the effects of the pre-polarization might be much larger.

3.5. Outlook: Relaxation of the obtained MFCC fragment densities using the 3-FDE method: current challenges

As shown in the previous chapter, 3-FDE(0) already represents a very good compromise between the accuracy of the embedding density and the efficiency of the calculation. The next step would be the improvement of the fragment densities within the 3-FDE framework. For this purpose, however, several technical challenges have to be overcome.

3.5.1. Cap convergence problems

Converging the cap potentials that assure a positive density everywhere in space proved to be very difficult for the present system. As explained in Section 2.2, in a 3-FDE fragment calculation these potentials have to be converged first, and only when they are fixed, the final relaxation of the fragment density can happen. Since, strictly speaking, the cap convergence procedure adds another layer of self-consistency to the overall procedure, we increased the maximum number of SCF iterations to 300, but for 29 of 36 fragments, one or both caps could still not be converged in a representative 3-FDE calculation. This could not be changed by including COSMO as an outer embedding model. Furthermore, we find that the current COSMO implementation in ADF [73] is not optimized for calculations of this type (see next paragraph). Choosing smaller fragments could not remedy the convergence problems either. We suspect the reason to be the currently used procedure for converging the cap potentials. When using the LB step, an extra damping factor is applied, which is identical for all grid points. However, the deviation from the desired potential at points near the nuclei is naturally very different from the one at grid points far away from the latter. This can lead to oscillations in the potential updates if too large steps are used. On the other hand, the smaller the steps have to be, the more iterations are needed to converge each potential. Since every potential update requires electronic self-consistency for the whole capped fragment, this can dominate the overall procedure very easily and make the calculation very lengthy. The step size can be adjusted in the current implementation (see also Ref. [46]), but different choices seem to be necessary for different fragments, which makes the whole process very cumbersome. In the future, we will therefore try to find improved schemes to converge the cap potentials.

3.5.2. Usage of COSMO with 3-FDE

As mentioned in the last paragraph, embedding calculations for full-size proteins with COSMO lead to some difficulties. The primary problem is that in contrast to the preparation-stage MFCC

calculations, where a cavity is constructed around each small fragment, in the 3-FDE case a vast cavity has to be created around the entire system. This leads to an enormous number of point charges and the disk I/O spent on storing their interaction matrix makes the calculation unfeasible. To circumvent this problem, we prepared a locally modified version of the ADF code in order to re-calculate the matrix directly whenever it is needed instead of writing it to disk. With this scheme the calculation actually becomes possible. However the CPU time spent on calculating the interaction matrix completely dominates the calculation (for details, see the Supporting Information).

The update of all COSMO charges in every SCF step turns out to be rather inefficient in the context of subsystem approaches, because only the charges near the active fragment will differ significantly from their values in the previous step. To make the COSMO scheme more efficient for FDE and 3-FDE calculations for large systems, it could be advantageous to introduce a distance criterion, based upon which it is decided whether an update is required.

Both of the issues discussed in this section will be adressed in our future work in order to facilitate also fully polarized 3-FDE(n) calculations for pigment-protein complexes.

4. Summary and conclusions

In this paper, we have established a robust protocol for density-based embedding calculations of chromophores in pigment–protein complexes using the 3-FDE(0) method. First, we tested different technical and conceptual parameters with regard to the individual MFCC calculations, which represent the preparation stage for the embedding calculations. We found that at least a medium-sized basis set (e.g., DZP) has to be used to be able to converge the individual protein fragments. The choice of an environment model already for the individual capped fragments (preparational embedding) proved to be crucial. Only with the aid of either COSMO or a background distribution of point charges from the MD simulation, we could converge all the fragments. If possible, we suggest to use the latter approach, in which the fragments will already be pre-polarized in the subsequent 3-FDE calculations. It is thus expected that less iterations of the embedding potential will be required, or that even the initial embedding potential constructed from the frozen MFCC density will model the real environment with fairly good accuracy. For very large fragments we found that only the continuum solvation model was able to reliably provide converged electron densities, even for fragment sizes of up to 50 amino acid residues per subsystem. This may, however, partly be due to the non-relaxed, force-field derived structure employed in this pilot study.

We also compared several of the mentioned parameters with respect to the corresponding 3-FDE(0) excitation energies. Concerning the fragmentation pattern, we found that the excitation energies hardly depend on this decision as long as at least five amino acid residues are treated within one fragment, which is easily possible on modern computers. This is an encouraging result, as it shows that the approximation of representing the protein density through the sum of its fragments is justified, at least with regard to the calculation of cofactor excitation energies. We also compared different basis sets and it was shown that already with DZP, the relative excitation energies among the pigments are almost identical to those with a triple-zeta basis set, although a common redshift is observed. The effects of the preparational embedding were very small, but not negligible in the present case.

When comparing the excitation energies for our default settings (fragmentation size of ten residues per fragment, preparational embedding with point charges and a DZP basis set) to the values initially obtained with a simple point charge model, we find that they are very similar. On the one hand, this is encouraging as it shows that the embedding procedure does not produce any artifacts connected with the partitioning of the protein. It seems that at least for the present arbitrary snapshot, the conceptually more simple point charge approach already gives results which are similar to density-based embedding. In this sense, 3-FDE may also serve as a benchmark reference for computationally even more efficient point-charge models in future applications.

Our future plans include using the methodology tested and established in this paper to sample along an MD trajectory in order to produce site energy distributions of the pigments. When combined with the FDEc approach, also excitonic coupling parameters can be produced with our current setup. In conjunction with the site energies, these can be used to model excitation energy transfer dynamics. Furthermore, we want to probe the effects of a relaxation of the initial frozen MFCC density in freeze-and-thaw cycles (3-FDE(n) calculations). To this end, we plan to improve the existing cap convergence procedure so that routine calculations of excitation energies with fully polarized protein densities become possible.

Acknowledgement

We would like to thank Prof. Dr. Ulrich Kleinekathöfer for providing MD topology and parameter files for BChl *a*. JN is supported by a VIDI grant (700.59.422) of the Netherlands Organisation for Scientific Research (NWO). CRJ acknowledges funding from the DFG Center for Functional Nanostructures at KIT. Support by COST action CODECS is acknowledged.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.comptc.2014.02.009.

References

[1] Robert E. Blankenship, Molecular Mechanisms of Photosynthesis, Blackwell, Science Ltd, Oxford, 2002.
[2] Markus Wendling, Milosz Przyjalgowski, Demet Gülen, Simone Vulto, Thijs J. Aartsma, Rienk van Grondelle, Herbert van Amerongen, The quantitative relationship between structure and polarized spectroscopy in the FMO complex of prosthecochloris aestuarii: refining experiments and simulations, Photosynth. Res. 71 (2002) 99–123.
[3] Gregory S. Engel, Tessa R. Calhoun, Elizabeth L. Read, Tae-Kyu Ahn, Tomáš Mančal, Yuan-Chung Cheng, Robert E. Blankenship, Graham R. Fleming, Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems, Nature 446 (2007) 782–786.
[4] Gitt Panitchayangkoon, Dugan Hayes, Kelly A. Fransted, Justin R. Caram, Elad Harel, Jianzhong Wen, Robert E. Blankenship, Gregory S. Engel, Long-lived quantum coherence in photosynthetic complexes at physiological temperature, Proc. Natl. Acad. Sci. USA 107 (29) (2010) 12766–12770.
[5] Leonardo A. Pachón, Paul Brumer, Computational methodologies and physical insights into electronic energy transfer in photosynthetic light-harvesting complexes, Phys. Chem. Chem. Phys. 14 (2012) 10094–10108.
[6] Carolin König, Johannes Neugebauer, Quantum chemical description of absorption properties and excited-state processes in photosynthetic systems, ChemPhysChem 13 (2) (2012) 386–425.
[7] Carsten Olbrich, Thomas L.C. Jansen, Jörg Liebers, Mortaza Aghtar, Johan Strümpfer, Klaus Schulten, Jasper Knoester, Ulrich Kleinekathöfer, From atomistic modeling to excitation transfer and two-dimensional spectra of the FMO light-harvesting complex, J. Phys. Chem. B 115 (2011) 8609–8621.
[8] Sangwoo Shim, Patrick Rebentrost, Stéphanie Valleau, Alán Aspuru-Guzik, Atomistic study of the long-lived quantum coherence in the Fenna–Matthews–Olson complex, Biophys. J. 102 (2012) 649–660.
[9] Carsten Olbrich, Johan Strümpfer, Klaus Schulten, Ulrich Kleinekathöfer, Theory and simulation of the environmental effects on FMO electronic transitions, J. Phys. Chem. Lett 2 (2011) 1771–1776.
[10] Thomas Renger, Alexander Klinger, Florian Steinecker, Marcel Schmidt am Busch, Jorge Numata, Frank Müh, Normal mode analysis of the spectral density of the Fenna–Matthews–Olson light-harvesting protein: how the protein dissipates the excess energy of excitons, J. Phys. Chem. B 116 (50) (2012) 14565–14580.

[11] Junkuo Gao, Wu-Jun Shi, Jun Ye, Xiaoqing Wang, Hajime Hirao, Yang Zhao, QM/MM modeling of environmental effects on electronic transitions of the FMO complex, J. Phys. Chem. B 117 (13) (2013) 3488–3495.
[12] S. Miertuš, E. Scrocco, J. Tomasi, Electrostatic interaction of a solute with a continuum: a direct utilization of ab initio molecular potentials for the prevision of solvent effects, Chem. Phys. 55 (1) (1981) 117–129.
[13] R. Cammi, J. Tomasi, Remarks on the use of the apparent surface charges (ASC) methods in solvation problems: iterative versus matrix-inversion procedures and the renormalization of the apparent charges, J. Comput. Chem. 16 (12) (1995) 1449–1458.
[14] E. Cancès, B. Mennucci, J. Tomasi, A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics, J. Chem. Phys. 107 (8) (1997) 3032–3041.
[15] Eric Cancès, Benedetta Mennucci, New applications of integral equations methods for solvation continuum models: ionic solutions and liquid crystals, J. Math. Chem. 23 (3–4) (1998) 309–326.
[16] B. Mennucci, E. Cancès, J. Tomasi, Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications, J. Phys. Chem. B 101 (49) (1997) 10506–10517.
[17] A. Klamt, G. Schuurmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc. Perkin Trans. 2 (1993) 799–805.
[18] Andreas Klamt, Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena, J. Phys. Chem. 99 (7) (1995) 2224–2235.
[19] Andreas Klamt, Volker Jonas, Treatment of the outlying charge in continuum solvation models, J. Chem. Phys. 105 (22) (1996) 9972–9981.
[20] Kazuo Kitaura, Eiji Ikeo, Toshio Asada, Tatsuya Nakano, Masami Uebayasi, Fragment molecular orbital method: an approximate computational method for large molecules, Chem. Phys. Lett. 313 (3–4) (1999) 701–706.
[21] Tatsuya Nakano, Tsuguchika Kaminuma, Toshiyuki Sato, Yutaka Akiyama, Masami Uebayasi, Kazuo Kitaura, Fragment molecular orbital method: application to polypeptides, Chem. Phys. Lett. 318 (6) (2000) 614–618.
[22] Dmitri G. Fedorov, Kazuo Kitaura, Extending the power of quantum chemistry to large systems with the fragment molecular orbital method, J. Phys. Chem. A 111 (30) (2007) 6904–6914.
[23] Mark S. Gordon, Mark A. Freitag, Pradipta Bandyopadhyay, Jan H. Jensen, Visvaldas Kairys, Walter J. Stevens, The effective fragment potential method: a QM-based MM approach to modeling environmental effects in chemistry, J. Phys. Chem. A 105 (2) (2001) 293–307.
[24] Soohaeng Yoo, Federico Zahariev, Sarom Sok, Mark S. Gordon, Solvent effects on optical properties of molecules: a combined time-dependent density functional theory/effective fragment potential approach, J. Chem. Phys. 129 (14) (2008) 144112.
[25] Yuji Mochizuki, Tatsuya Nakano, Shinji Amari, Takeshi Ishikawa, Kiyoshi Tanaka, Minoru Sakurai, Shigenori Tanaka, Fragment molecular orbital calculations on red fluorescent protein (DsRed), Chem. Phys. Lett. 433 (4–6) (2007) 360–367.
[26] Tsutomu Ikegami, Toyokazu Ishida, Dmitri G. Fedorov, Kazuo Kitaura, Yuichi Inadomi, Hiroaki Umeda, Mitsuo Yokokawa, Satochi Sekiguchi, Fragment molecular orbital study of the electronic excitations in the photosynthetic reaction center of blastochloris viridis, J. Comput. Chem. 31 (2010) 447–454.
[27] Naoki Taguchi, Yuji Mochizuki, Tatsuya Nakano, Fragment molecular orbital calculations for excitation energies of blue- and yellow-fluorescent proteins, Chem. Phys. Lett. 504 (13) (2011) 76–82.
[28] Noriyuki Minezawa, Nuwan De Silva, Federico Zahariev, Mark S. Gordon, Implementation of the analytic energy gradient for the combined time-dependent density functional theory/effective fragment potential method: application to excited-state molecular dynamics simulations, J. Chem. Phys. 134 (5) (2011) 054111.
[29] Georgios Fradelos, Jesse J. Lutz, Tomasz A. Wesołowski, Piotr Piecuch, Marta Włoch, Embedding vs. supermolecular strategies in evaluating the hydrogen-bonding-induced shifts of excitation energies, J. Chem. Theory Comput. 7 (6) (2011) 1647–1666.
[30] Nuwan De Silva, Soohaeng Y. Willow, Mark S. Gordon, Solvent induced shifts in the UV spectrum of amides, J. Phys. Chem. A 117 (46) (2013) 11847–11855.
[31] Takeshi Nagata, Dmitri G. Fedorov, Kazuo Kitaura, Mark S. Gordon, A combined effective fragment potential–fragment molecular orbital method. I: The energy expression and initial applications, J. Chem. Phys. 131 (2) (2009) 024101.
[32] Pietro Cortona, Self-consistently determined properties of solids without band-structure calculations, Phys. Rev. B 44 (16) (1991) 8454–8458.
[33] Tomasz Adam Wesołowski, Arieh Warshel, Frozen density functional approach for ab initio calculations of solvated molecules, J. Phys. Chem. 97 (1993) 8050–8053.
[34] Christoph R. Jacob, Johannes Neugebauer, Subsystem density-functional theory, WIREs Comput. Mol. Sci. (2013), http://dx.doi.org/10.1002/wcms.1175.
[35] Tomasz A. Wesołowski, Hydrogen-bonding-induced shifts of the excitation energies in nucleic acid bases: an interplay between electrostatic and electron density overlap effects, J. Am. Chem. Soc. 126 (2004) 11444–11445.
[36] Johannes Neugebauer, Photophysical properties of natural light-harvesting complexes studied by subsystem density functional theory, J. Phys. Chem. B 112 (2008) 2207–2217.
[37] Carolin König, Johannes Neugebauer, First-principles calculation of electronic spectra of light-harvesting complex II, Phys. Chem. Chem. Phys. 13 (2011) 10475–10490.
[38] O. Roncero, M.P. de Lara-Castells, P. Villarreal, F. Flores, J. Ortega, M. Paniagua, A. Aguado, An inversion technique for the calculation of embedding potentials, J. Chem. Phys. 129 (2008) 184104.
[39] Samuel Fux, Christoph R. Jacob, Johannes Neugebauer, Lucas Visscher, Markus Reiher, Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds, J. Chem. Phys. 132 (2010) 164101.
[40] J.D. Goodpaster, N. Ananth, F.R. Manby, T.F. Miller III, Exact nonadditive kinetic potentials for embedded density functional theory, J. Chem. Phys. 133 (2010) 084103.
[41] Chen Huang, Emily A. Carter, Potential-functional embedding theory for molecules and materials, J. Chem. Phys. 135 (2011) 194104.
[42] Carolin König, Johannes Neugebauer, Protein effects on the optical spectrum of the Fenna–Matthews–Olson complex from fully quantum chemical calculations, J. Chem. Theory Comput. 9 (3) (2013) 1808–1820.
[43] Da W. Zhang, J.Z.H. Zhang, Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy, J. Chem. Phys. 119 (7) (2003) 3599–3605.
[44] Ai M. Gao, Da W. Zhang, John Z.H. Zhang, Yingkai Zhang, An efficient linear scaling method for ab initio calculation of electron density of proteins, Chem. Phys. Lett. 394 (4–6) (2004) 293–297.
[45] Christoph R. Jacob, Lucas Visscher, A subsystem density-functional theory approach for the quantum chemical treatment of proteins, J. Chem. Phys. 128 (15) (2008) 155102.
[46] Karin Kiewisch, Christoph R. Jacob, Lucas Visscher, Quantum-chemical electron densities of proteins and of selected protein sites from subsystem density functional theory, J. Chem. Theory Comput. 9 (5) (2013) 2425–2440.
[47] Mark E. Casida, Tomasz A. Wesołowski, Generalization of Kohn–Sham equations with constrained electron density formalism and its time-dependent response theory formulation, Int. J. Quantum Chem. 96 (2004) 577–588.
[48] T.A. Wesołowski, J. Weber, Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules, Chem. Phys. Lett. 248 (1996) 71–76.
[49] R. van Leeuwen, E.J. Baerends, Exchange-correlation potential with correct asymptotic behavior, Phys. Rev. A 49 (1994) 2421–2431.
[50] Carsten Olbrich, Johan Strümpfer, Klaus Schulten, Ulrich Kleinekathöfer, Quest for spatially correlated fluctuations in the FMO light-harvesting complex, J. Phys. Chem. B 115 (4) (2011) 758–764.
[51] Dale E. Tronrud, Jianzhong Wen, Leslie Gay, Robert E. Blankenship, The structural basis for the difference in absorbance spectra for the FMO antenna protein from various green sulfur bacteria, Photosynth. Res. 100 (2009) 79–87.
[52] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, Klaus Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.
[53] A.D. MacKerell Jr., D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.
[54] Alexander D. Mackerell, Michael Feig, Charles L. Brooks, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25 (11) (2004) 1400–1415.
[55] Nicolas Foloppe, Michel Ferrand, Jacques Breton, Jeremy C. Smith, Structural model of the photosynthetic reaction center of rhodobacter capsulatus, Proteins: Struct. Funct. Genet. 22 (1995) 226–244.
[56] Ana Damjanović, Ioan Kosztin, Ulrich Kleinekathöfer, Klaus Schulten, Excitons in a photosynthetic light-harvesting system: a combined molecular dynamics, quantum chemistry, and polaron model study, Phys. Rev. E 65 (2002) 031919.
[57] Dage Sundholm, Comparison of the electronic excitation spectra of chlorophyll a and pheophytin a calculated on density functional theory level, Chem. Phys. Lett. 317 (2000) 545–552.
[58] Amsterdam density functional program. Theoretical Chemistry, Vrije Universiteit, Amsterdam. <http://www.scm.com>.
[59] G. Te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, Chemistry with ADF, J. Comput. Chem. 22 (9) (2001) 931–967.
[60] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868.
[61] John P. Perdew, Kieron Burke, Matthias Ernzerhof, Generalized gradient approximation made simple [Phys. Rev. Lett. 77, 3865 (1996)], Phys. Rev. Lett. 78 (1997) 1396.
[62] Christoph R. Jacob, S. Maya Beyhan, Rosa E. Bulo, André Severo Pereira Gomes, Andreas W. Götz, Karin Kiewisch, Jetze Sikkema, Lucas Visscher, PyADF – a scripting framework for multiscale quantum chemistry, J. Comput. Chem. 32 (10) (2011) 2328–2338.
[63] Johannes Neugebauer, Christoph R. Jacob, Tomasz A. Wesołowski, Evert Jan Baerends, An explicit quantum chemical method for modeling large solvation shells applied to aminocoumarin C151, J. Phys. Chem. A 109 (2005) 7805–7814.

[64] Johannes Neugebauer, Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory, J. Chem. Phys. 126 (2007) 134116.
[65] Julia Adolphs, Thomas Renger, How proteins trigger excitation energy transfer in the FMO complex of green sulfur bacteria, Biophys. J. 91 (2006) 2778–2797.
[66] Alisa Solovyeva, Michele Pavanello, Johannes Neugebauer, Spin densities from subsystem density-functional theory: assessment and application to a photosynthetic reaction center complex model, J. Chem. Phys. 136 (19) (2012) 194104.
[67] Peter Pulay, Convergence acceleration of iterative sequences. The case of SCF iteration, Chem. Phys. Lett. 73 (1980) 393–398.
[68] Xiangqian Hu, Weitao Yang, Accelerating self-consistent field convergence with the augmented Roothaan–Hall energy function, J. Chem. Phys. 132 (2010) 054109.
[69] Margareta R.A. Blomberg, Per E.M. Siegbahn, Gerald T. Babcock, Modeling electron transfer in biochemistry: a quantum chemical study of charge separation in rhodobacter sphaeroides and photosystem II, J. Am. Chem. Soc. 120 (34) (1998) 8812–8824.
[70] Johannes Neugebauer, Manuel J. Louwerse, Evert Jan Baerends, Tomasz A. Wesołowski, The merits of the frozen-density embedding scheme to model solvatochromic shifts, J. Chem. Phys. 122 (2005) 094115.
[71] Arseny Kovyrshin, Johannes Neugebauer, Potential-energy surfaces of local excited states from subsystem- and selective Kohn–Sham-TDDFT, Chem. Phys. 391 (1) (2011) 147–156.
[72] Marie Humbert-Droz, Xiuwen Zhou, Sapana V. Shedge, Tomasz A. Wesolowski, How to choose the frozen density in frozen-density embedding theory-based numerical simulations of local excitations, Theor. Chem. Acc. 132 (2013) 1405.
[73] Cory C. Pye, Tom Ziegler, An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package, Theor. Chem. Acc. 101 (6) (1999) 396–408.
[74] W. Humphrey, A. Dalke, K. Schulten, VMD — visual molecular dynamics, J. Molec. Graph. 14 (1996) 33–38.
[75] D. Frishman, P. Argos, Knowledge-based secondary structure assignment, Proteins: Struct. Funct. Genet. 23 (1995) 566–579.