Definition and benchmarking of ab initio fragment methods for accurate excimer potential energy surfaces

Bónis Barcza,[†,¶] Ádám B. Szirmai,[†,¶] Attila Tajti,[†] John F. Stanton,[‡] and Péter G. Szalay[*,†]

†*Laboratory of Theoretical Chemistry, Institute of Chemistry, Eötvös University, P. O. Box 32, H-1518, Budapest 112, Hungary*

‡*Department of Chemistry, University of Florida, Gainesville, Florida 32611, United States*

¶*György Hevesy Doctoral School, ELTE Eötvös Loránd University, Institute of Chemistry*

E-mail: szalay@chem.elte.hu

Abstract

While Coupled-Cluster methods have been proven to provide an accurate description of excited electronic states, the scaling of the computational costs with the system size limits their applicability. In this study a fragment-based approach is presented which can be applied to non-covalently bound molecules with interacting chromophores localized on the fragments (so called *Frankel pairs*), such as π-stacked nucleobases.

The interaction of the fragments is considered at two distinct points. First, the states localized on the fragments are described in the presence of the other fragment(s), for which we test two approaches. A QM/MM type method that only includes the electrostatic interaction between the fragments in the electronic structure calculations,

while Pauli repulsion and dispersion effects are added separately. The other model, a Projection-based Embedding (PbE) using the Huzinaga equation includes both electrostatic and Pauli repulsion and only needs to be augmented by dispersion interactions. In both schemes the extended Effective Fragment Potential (EFP2) method of Gordon et al. was found to provide an accurate correction for the missing terms.

In the second step, the interaction of the localized chromophores is modeled for a proper description of the excitonic coupling. Here the inclusion of purely electrostatic contributions appears to be sufficient: it is found that the Coulomb part of the coupling, as evaluated using the Transition Density Cube method, provides accurate splitting of the energy of interacting chromophores above 4 Å intermolecular separations.

Keywords: excited states, exciton coupling, intermolecular interactions, Effective Fragment Potential, Pauli repulsion, embedding, QM/MM, dispersion ■

1 Introduction

The last decades brought about a huge development of quantum chemical methodology: larger and larger molecules can be treated with increasing accuracy, and, at the same time, the quest for calculations to support experimental observations became more and more relevant. For the molecular ground state well established methods are available and quantum chemistry, mostly by density functional theory (DFT), is able to study structures and even reactions of large molecules, as big as polipeptides. Recent experiments, however, discovered the importance of new processes in these large molecules: electron transfer between distant regions, or excitations affecting not just a local domain are responsible for many interesting phenomena in biology and material science. Since ground state methods cannot describe such processes, there is a challenge to develop new tools which can treat excited states as well.

There are two possible routes towards this goal. One option is to develop new approximate methods, but maintaining the accuracy and reliability is not a trivial task. Alternatively, one can aim at defining multiscale approaches where only the important part of the system is handled on the high level, while the rest is approximated at lower level. Different types of embedding methods, like quantum mechanics/molecular mechanics (QM/MM),[1] "our own N-layered integrated molecular orbital and molecular mechanics" ONIOM,[2] projector based embedding (PbE),[3] frozen density embedding (FDE)[4,5] or local correlation methods[6–12] are available for describing local events in large systems. However, non-local phenomena often require too large active partitions in these calculations, rendering the realization impossible.

Such collective events are indeed very important. For example, we have found earlier[13] that excitations can be delocalized to at least four nucleobases in oligonucleotides and this study left the question open whether even more units play role in excited states of RNA and DNA chains.

For such situations fragment methods[14] could be the preferred approach, where several "active" centers can be handled at a high level of theory and the property of the entire system is calculated from that of the individual fragments, considering proper coupling terms between them.

Such fragment methods are especially suited for non-covalently bound systems, since the fragmentation is natural. From the point of view of excited states, DNA is one of such systems since the sugars and the phosphate backbone play a spectator role as the excited states of the nucleobases, which are non-covalently bound with each other, do not delocalize to these parts.[15]

On the route towards developing such a fragment based method, in our previous work[16] we have investigated the possibilities of accurately obtaining the ground state potential curves of non-covalent dimers using high-level Coupled Cluster (CC) theory to describe the electronic structure of the fragments. Several schemes have been tested and we have found that both QM/MM and PbE methods, with a proper account for dispersion and Pauli repulsion ("exchange interaction") terms from the extended Effective Fragment Potential (EFP2)[17–19] model are capable of reproducing full dimer calculations at the same level of theory to a high accuracy. In this work we take another step forward by extending this methodology to (singly) excited states, thereby considering multi-chromophore systems.

The general Hamiltonian of two interacting fragments is given as

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2) = \hat{H}_1(\mathbf{r}_1) + \hat{H}_2(\mathbf{r}_2) + \hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2), \tag{1}$$

with $\hat{H}_i$ being the Hamiltonian of the non-interacting fragments and $\hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2)$ their interaction. With this Hamiltonian a perturbative treatment is possible, taking the product of the fragments' wave functions as the zeroth order. When considering the ground state and one excited state on each fragment, these product wave functions take the form

$$\begin{aligned} \Psi_0(\mathbf{r}_1, \mathbf{r}_2) &= \Phi_{1,0}(\mathbf{r}_1)\Phi_{2,0}(\mathbf{r}_2) \\ \Psi_1(\mathbf{r}_1, \mathbf{r}_2) &= \Phi_{1,1}(\mathbf{r}_1)\Phi_{2,0}(\mathbf{r}_2) \\ \Psi_2(\mathbf{r}_1, \mathbf{r}_2) &= \Phi_{1,0}(\mathbf{r}_1)\Phi_{2,1}(\mathbf{r}_2), \end{aligned} \tag{2}$$

where the indices 1 and 2 refer to the two subsystems. $\Phi_{i,k}$ is the kth eigenfunction of the

Hamiltonian of fragment i:

$$\hat{H}_i(\mathbf{r}_i)\Phi_{i,k}(\mathbf{r}_i) \;=\; E_{i,k}\Phi_{i,k}(\mathbf{r}_i). \tag{3}$$

The interacting states of the entire system can be obtained by the so-called *exciton model* by diagonalizing the Hamiltonian defined in the three-dimensional space spanned by the functions in Eqn. (2), as first suggested by Frenkel,[20] later by Davydov[21] and used in many applications (see e.g. Refs. 22–24 and references therein) ever since.

The quality of such a perturbational approach depends on the magnitude of the coupling terms. A fraction of their effects can be captured by repartitioning the Hamiltonian in Eqn. (1):

$$\hat{H}(\mathbf{r}_1,\mathbf{r}_2) \;=\; \hat{H}_1^{eff}(\mathbf{r}_1;\mathbf{r}_2) + \hat{H}_2(\mathbf{r}_2) + \Delta V_1(\mathbf{r}_1,\mathbf{r}_2) \tag{4}$$

where $\hat{H}_1^{eff} = \hat{H}_1 + \hat{V}_1^{eff}$ and $\Delta V_1(\mathbf{r}_1,\mathbf{r}_2) \approx \hat{V}_{1,2}(\mathbf{r}_1,\mathbf{r}_2) - \hat{V}_1^{eff}$. $\hat{H}_1^{eff}$ is essentially an embedded Hamiltonian which can be used to describe the ground state and also local excited states in the presence of the other fragment. The accuracy of this approach will primarily depend on the choice of $\hat{V}_1^{eff}$, and $\Delta V_1(\mathbf{r}_1,\mathbf{r}_2)$ and several choices have been tested in our previous study.[16] This formalism can also be applied to the excited state of the complex, provided the excitation is localized on the first fragment (see e.g. Ref. 25).

In this study we extend this scheme to interacting chromophores by considering the exciton model: after obtaining the fragment energies and wave functions corresponding to the effective Hamiltonians $\hat{H}_i^{eff}$, the final excited states are obtained by diagonalizing the Hamiltonian matrix defined in the space of the two locally excited product functions.

The scheme suggested here has thus two main ingredients. First, a proper definition of the effective Hamiltonian $\hat{H}_i^{eff}$ is needed, which includes the intermolecular interactions compatible with the electronic structure method in use. Eventually, a corresponding ΔV_i is necessary to include the missing interaction terms. Based on our experience for the ground state,[16] QM/MM and PbE approaches used with Coupled Cluster (CC) type electronic structure methods, augmented by EFP2 dispersion and Pauli repulsion[17–19] are able to provide proper interaction of the fragments, worth of generalizing to excited states. Second, an

appropriate approximation for the excitonic or Frenkel coupling between the excited states needs also to be established.

Similar methods have been suggested among others by Morrison et al.,[22] Sisto et al.,[23] Amadei et al.,[24] and Head-Gordon et al.[26] The novelty of the schemes presented here is the use of high-level CC treatment of the electronic structure of the fragments and the careful selection of all components of the non-covalent interaction energy.

Obtaining proper reference data for validating the models turns out to be a non-trivial task. Problems arise from, among others, converging the calculations to the proper states (in particular when diffuse basis sets are used), from the presence of Charge Transfer (CT) type electronic states in the low-energy spectrum at small intermolecular distances, as well as from the choice of counterpoise correction (CP) in excited states which is not well discussed in the literature. We also address these issues in this study.

The paper is organized as follows. Section 2 gives a short summary of the available methodologies, with emphasis on the differences in theoretical formulations and the possibility of incorporating them into *ab initio* calculations. Section 4 describes the computational details, while the results and their discussion are presented in Section 5.

2 Methodologies of modeling intermolecular interactions in excited states

Let us introduce the proposed methodology on the simplest example of a supersystem consisting of two fragments, with just one excited state considered on both. To describe the Frenkel coupling of the two excitations localized on the fragments, we select two product functions describing the two local excited states:

$$\begin{aligned} \tilde{\Psi}_1(\mathbf{r}_1, \mathbf{r}_2) &= \tilde{\Phi}_{1,1}(\mathbf{r}_1)\Phi_{2,0}(\mathbf{r}_2) \\ \tilde{\Psi}_2(\mathbf{r}_1, \mathbf{r}_2) &= \Phi_{1,0}(\mathbf{r}_1)\tilde{\Phi}_{2,1}(\mathbf{r}_2), \end{aligned} \tag{5}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ represent the coordinates of fragment 1 and 2, respectively. The first function describes the excited state localized on the first fragment, while the second function describes

the one localized on the second fragment, and they belong to two different partitioning of the Hamiltonian:

$$\begin{aligned} \hat{H}(\mathbf{r}_1, \mathbf{r}_2) &= \hat{H}_1^{eff}(\mathbf{r}_1; \mathbf{r}_2) + \hat{H}_2(\mathbf{r}_2) + \Delta V_1(\mathbf{r}_1, \mathbf{r}_2) \\ \hat{H}(\mathbf{r}_1, \mathbf{r}_2) &= \hat{H}_2^{eff}(\mathbf{r}_1; \mathbf{r}_2) + \hat{H}_1(\mathbf{r}_2) + \Delta V_2(\mathbf{r}_1, \mathbf{r}_2), \end{aligned} \tag{6}$$

where $\hat{H}_i^{eff} = \hat{H}_i + \hat{V}_i^{eff}$ includes the effect of the environment through an "embedding potential" $\hat{V}_i^{eff}$. The fragment wave functions with tilde ($\tilde{\Phi}_{1,1}(\mathbf{r}_1), \tilde{\Phi}_{2,1}(\mathbf{r}_2)$) are eigenfunctions of the appropriate effective Hamiltonians

$$\hat{H}_i^{eff} \tilde{\Phi}_{i,1} = \tilde{E}_{i,1} \tilde{\Phi}_{i,1}, \tag{7}$$

thus include the effect of the other fragment's ground state through $\hat{V}_i^{eff}$, while the bare ones ($\Phi_{1,1}(\mathbf{r}_1), \Phi_{2,1}(\mathbf{r}_2)$) are eigenfunctions of the unperturbed subsystems. The inclusion of the approximate "remainder potential" $\Delta V_i(\mathbf{r}_1, \mathbf{r}_2) \approx \hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2) - \hat{V}_i^{eff}$ is necessary to correct for such interactions which are not/cannot be accounted for when the eigenproblem of $\hat{H}_i^{eff}$ is solved (see later).

$\tilde{\Psi}_1(\mathbf{r}_1, \mathbf{r}_2)$ is a good approximation for local excited states with energy

$$E_1 = \tilde{E}_{1,1} + E_{2,0} + \Delta V_1 \tag{8}$$

but does not include any polarization between the two fragments. This expression is often used in QM/MM and other embedding schemes, we also have tested it in Ref. 25. The quality of this approximation depends on the electronic structure method, as well as on the choice of $\hat{V}_i^{eff}$ and ΔV_i and will be discussed in Subsection 2.1 in detail. Note that, due to the missing anti-symmetrization of the wave functions in Eqn. (5), the Pauli repulsion contribution is missing from the interaction potential, but can be conveniently included in ΔV_i.

To consider Frenkel[20] or exciton coupling of the two chromophores on the fragments, the matrix of the Hamiltonian on the basis of the functions in Eqn. (5) should be diagonalized:

$$\mathbf{H} = \begin{bmatrix} \tilde{E}_{1,1} + E_{2,0} + \Delta V_1 & \tilde{V}(1,2) \\ \tilde{V}(2,1) & E_{1,0} + \tilde{E}_{2,1} + \Delta V_2 \end{bmatrix} \quad (9)$$

with

$$\begin{aligned} \tilde{V}(1,2) &= \langle \tilde{\Psi}_1(\mathbf{r}_1, \mathbf{r}_2) | \hat{H}(\mathbf{r}_1, \mathbf{r}_2) | \tilde{\Psi}_2(\mathbf{r}_1, \mathbf{r}_2) \rangle \\ &= \langle \tilde{\Phi}_{1,1}(\mathbf{r}_1) \Phi_{2,0}(\mathbf{r}_2) | \hat{H}(\mathbf{r}_1, \mathbf{r}_2) | \Phi_{1,0}(\mathbf{r}_1) \tilde{\Phi}_{2,1}(\mathbf{r}_2) \rangle. \end{aligned} \quad (10)$$

To evaluate this coupling, appropriate approximations need to be introduced which are discussed below in Subsection 2.3.

First, however, the multilevel approaches that can be applied for solving the eigensystem of $\hat{H}_i^{eff}$, along with the appropriate choice of the "remainder potential" are reviewed in Subsections 2.1 and 2.2.

2.1 Definition of the effective Hamiltonian

As in Ref. 16, in this study we test two approaches to define the effective Hamiltonian that can be used with CC methods.

2.1.1 QM/MM

The multilevel scheme termed as QM/MM (Quantum Mechanics/Molecular Mechanics)[1] revolves around the representation of the environment with point charges that usually reside on atomic sites, allowing the treatment of large complexes in a very cheap manner. Within this scheme the effective Hamiltonian H_1^{eff} introduced in Eqn. (6) can be given as

$$\hat{H}_1^{eff,\ QM/MM}(\mathbf{r}_1, \mathbf{r}_2) = \hat{H}_1(\mathbf{r}_1) + \hat{V}_1^{electrostatic}(\mathbf{r}_1; \mathbf{r}_2), \quad (11)$$

where $\hat{V}_1^{electrostatic}$ represents the Coulomb interaction of fragments 1 and 2, the latter represented *via* the partial charges. The main advantage of this scheme is its straightforward application in any quantum chemical method and code, and therefore it is widely applied.

There are multiple ways to define the point charges, several sets are available e.g. in traditional molecular force fields like AMBER[27,28] or CHARMM,[29] and a more flexible parametrization can be achieved using atomic multipoles[30,31] as e.g. within the Effective Fragment Potential (EFP) framework.[32–34] In our previous paper on the ground state of non-covalently bound complexes[16] the CHELPG (CHarges from ELectrostatic Potentials using a Grid-based) algorithm[35] was found to work well with CC methods. In the context of CHELPG the partial charges at the atomic sites are fitted to reproduce the electrostatic potential and some lower order electrostatic moments of the molecule calculated at any level of theory, hence we can apply the one used for the active fragment (e.g. CCSD). This way, an accurate set of charges can be obtained from first principle considerations, which fit to the *ab initio* spirit of the investigated schemes, unlike methodologies used in most conventional force fields.

The choice of the effective Hamiltonian in Eqn. (11) does not include important interaction terms, such as dispersion and Pauli exchange between the fragments, thus for accurate interaction energies also at short distances, these need to be added *a posteriori* via $\Delta V_1(\mathbf{r}_1;\mathbf{r}_2)$ (see Subsection 2.2).

2.1.2 Huzinaga embedding scheme

A more sophisticated consideration of the environment and its interactions is possible using multilevel quantum chemical embedding methods. Bottom-up, frozen density embedding[4,5] and top-down, projector-based embedding (PbE)[3] strategies are both available. The former is computationally less demanding as subsystems are treated separately, while the latter avoids the issues that arise from the lack of orthogonality between subsystems. PbE was first introduced by Manby and Miller,[3] who included an arbitrary level shift parameter in the Fockian of the embedded subsystem, raising the energy of environment orbitals to near infinity. Orthogonality between subsystems can also be achieved by adapting the Huzinaga-equation[36] to the embedding problem[37] or via the projector scheme of Hoffmann and Khait.[38] PbE has been applied to the calculation of excited states with various goals in mind. Bennie et al.[39] improved the original PbE method by including the most important occupied orbitals of the environment in the excited state calculation. Parravicini and Jagau[40] studied

ionization, electron attachment and electronic resonances; also applying the concentric virtual orbital localization scheme.[41] A benchmark study by Hégely et al.[42] compared various multilevel approaches including PbE for the calculation of excitation energies.

In this study, the embedding scheme of Hégely and co-workers,[37] which is based on the Huzinaga-equation, was chosen as the second approach for the description of the active fragment. Embedding a fragment treated at the CC level of theory into the environment of the other fragment described using Density Funcional Theory (DFT) requires the *a priori* definition of the active subsystem and the environment. Following a DFT calculation on the entire system, the occupied molecular orbitals (MOs) of the fragments are localized, which results in the splitting of the density matrix into two parts $\mathbf{D}_1$ and $\mathbf{D}_2$. The localization can be extended to the virtual orbital space, reducing the computational cost of the CC calculations, as well as allowing a clearer treatment of localized excited states by avoiding charge transfer contributions (see later). The SPADE (Subsystem Projected AO DEcomposition)[43] method offers a black-box way of localizing and partitioning the orbital space based on the difference of singular values in a singular value decomposition scheme, and can be applied to both the occupied and the virtual orbitals.

The Fockian of the active subsystem is

$$\tilde{\mathbf{F}}_1 \;=\; \mathbf{h} + \mathbf{G}^{\mathrm{HF}}[\tilde{\mathbf{D}}_1] + (\mathbf{G}^{\mathrm{DFT}}[\mathbf{D}_{1,2}] - \mathbf{G}^{\mathrm{DFT}}[\mathbf{D}_1]), \tag{12}$$

where $\mathbf{h}$ and $\mathbf{G}^{\mathrm{HF}}[\tilde{\mathbf{D}}_1]$ are the core Hamiltonian and the two-electron part of the Fockian of subsystem 1, respectively, while the remaining terms in Eqn. (12) give the embedding potential. $\tilde{\mathbf{D}}_1$ is the density of the embedded active subsystem reoptimized in a second Hartree-Fock calculation during which the orthogonality of the subsystems is ensured by solving the Huzinaga-equation, while the orbitals of the environment are kept frozen:

$$(\tilde{\mathbf{F}}_1 - \mathbf{S}\mathbf{P}_2\tilde{\mathbf{F}}_1 - \tilde{\mathbf{F}}_1\mathbf{P}_2\mathbf{S} + 2\mathbf{S}\mathbf{P}_2\tilde{\mathbf{F}}_1\mathbf{P}_2\mathbf{S})\tilde{\mathbf{C}}_1 \;=\; \mathbf{S}\tilde{\mathbf{C}}_1\tilde{\boldsymbol{\mathcal{E}}}_1, \tag{13}$$

where $\mathbf{S}$ is the overlap matrix of the atomic orbitals, $\mathbf{P}_2$ is the projector of the environemnt orbitals and $\tilde{\mathbf{C}}_1$ is the MO coefficient matrix of the embedded orbitals, and their energies are in the diagonal matrix $\tilde{\boldsymbol{\mathcal{E}}}_1$. The orbitals obtained this way can be used in correlated

wavefunction (WF) calculations both in ground and excited state.

The energy ansatz of this method for WF-in-DFT embedding is:

$$E_1^{\mathrm{WF-in-DFT}}[\Psi_1;\tilde{\mathbf{D}}_1,\mathbf{D}_1,\mathbf{D}_2] = E_{1,2}^{\mathrm{DFT}}[\mathbf{D}_{1,2}] - E_1^{\mathrm{DFT}}[\mathbf{D}_1] + E_1^{\mathrm{WFT}}[\Psi_1;\tilde{\mathbf{D}}_1,\mathbf{D}_1,\mathbf{D}_2], \quad (14)$$

where $E_{1,2}^{\mathrm{DFT}}$ is the DFT energy of the dimer system, E_1^{DFT} is the DFT energy of the active fragment calculated using the localised dimer orbitals and E_1^{WFT} is the WF energy of the embedded fragment. This energy ansatz is general, meaning that the ground and excited state energies differ only in the WF term that is calculated using the same embedding potential in both cases. This is the result of the frozen environment orbitals as they remain unrelaxed to changes in the active orbital space. Note that the Fockian here is constructed from similar principles as the effective Hamiltonian introduced in Eqn. (6), i.e. it includes the electronic embedding by the enviroment's density. It is also apparent from the energy expression (Eqn. (14)) that inter-subsystem interactions are retained from the original ground state supersystem calculation at the DFT level of theory. This means that not only the electrostatic interaction, but the ground state Pauli repulsion is included in the final energy. Common DFT functionals are known to lack a proper description of dispersion interactions, therefore it has to included separately.

The first term on the right-hand side of Eqn. (14) can be formally divided into monomer energies calculated using the localized dimer MOs and the interaction energy term:

$$E_{1,2}^{\mathrm{DFT}}[\mathbf{D}_{1,2}] = E_1^{\mathrm{DFT}}[\mathbf{D}_1] + E_2^{\mathrm{DFT}}[\mathbf{D}_2] + E_{1,2,\mathrm{int}}^{\mathrm{DFT}}[\mathbf{D}_{1,2}]. \quad (15)$$

Therefore, the above original energy ansatz (Eqn. (14)) contains the fragment energies at two different levels of theory: that of the active fragment is from wavefunction calculation ($E_1^{\mathrm{WFT}}[\Psi_1;\tilde{\mathbf{D}}_1,\mathbf{D}_1,\mathbf{D}_2]$), while the fragment playing the role of environment at the DFT level ($E_2^{\mathrm{DFT}}[\mathbf{D}_2]$). This results in an incorrect asymptotic behavior of relative energy curves for non-homodimer systems. Therefore, we suggest the use of a modified energy expression:

$$\begin{aligned}\bar{E}_1^{\mathrm{WF-in-DFT}} &= E_1^{\mathrm{WF-in-DFT}}[\Psi_1;\tilde{\mathbf{D}}_1,\mathbf{D}_1,\mathbf{D}_2] + E_2^{\mathrm{WF-in-DFT}}[\Psi_2;\tilde{\mathbf{D}}_2,\mathbf{D}_1,\mathbf{D}_2] - E_{1,2}^{\mathrm{DFT}}[\mathbf{D}_{1,2}], \\ &= E_1^{\mathrm{WFT}}[\Psi_1;\tilde{\mathbf{D}}_1,\mathbf{D}_1,\mathbf{D}_2] + E_2^{\mathrm{WFT}}[\Psi_2;\tilde{\mathbf{D}}_2,\mathbf{D}_1,\mathbf{D}_2] + E_{1,2,\mathrm{int}}^{\mathrm{DFT}}[\mathbf{D}_{1,2}]. \quad (16)\end{aligned}$$

2.2 Choice for the remainder potential ΔV_i

The above outlined schemes lack some important intermolecular interactions: the *QM/MM* does not include any van der Waals type interactions, while dispersion is missing from the PbE approach. Based on our findings in Ref. 16, correcting for dispersion and Pauli repulsion via $\Delta V_1(\mathbf{r}_1, \mathbf{r}_2)$ term in the Hamiltonian is an appropriate approximation and the extended Effective Fragment Potential (EFP2) developed by Gordon and coworkers[17–19,44–46] is an excellent choice. See Ref. 16 for other possibilities.

The essence of the EFP2[17] is that an adequate parameter set of the potential is derived from first principle considerations. The algorithm is able predict every important contribution to the intermolecular interactions, but here we shortly discuss only the Pauli repulsion and dispersion terms, which are used in the calculations. For a more detailed description of the EFP2 formalism, the reader is referred to Refs. 18, 19, 46.

To obtain the dispersion contribution in EFP2,[47] a set of points are defined in the first step, located at the nuclear positions and the midpoints of covalent bonds. These serve as the center of localized molecular orbitals (LMO) obtained from a HF or DFT calculation on the subsystem. Then the dynamic dipole polarizabilities around the centroids of these LMOs are determined by applying the time-dependent extension of the chosen method, which are then used to calculate the dispersion between the fragments. Additional damping parameters, derived from the overlap of these orbitals, are also used to correct the behavior at short separations.

For the Pauli repulsion, in the EFP2 scheme[19] the necessary antisymmetrization is achieved by restricting the permutations to just two electrons and approximating the energetic effect with a power series with respect to the orbital overlaps S. The terms beyond $\mathcal{O}(S^3)$ are neglected, while additional simplifications are applied also to the lower-order terms.

According to the literature,[48,49] EFP2 is a good model for interactions of molecules in the ground state. However, to our knowledge, only one attempt has been made to treat excited states with EFP2: Rojas and Slipchenko[50] calculated the solvent effect on the excitation energies of nine chromophores. In this study, the Pauli repulsion term was included in the

effective Hamiltonian (so called QM/EFP2 scheme).[51,52] Important efforts have also been made by Hapka, Przybytek, and Pernal[53,54] to include the dispersion between the ground and excited states within SAPT (Symmetry-Adapted Perturbation Theory).[55,56] These methods are not yet available for routine applications.

Therefore, since the experience with the ground state clearly reveals that dispersion and Pauli repulsion are necessary to achieve the accuracy we aim for,[16] we decided to test whether the ground state corrections are also applicable to the excited states. The hope is that the error made this way is significantly smaller than by completely neglecting these terms. This idea is supported by the fact that in excited states that are dominated by a single substitution, only a small fraction of the electrons are affected by the excitation and the dominant part of the electron density is very similar in the ground and excited electronic states. Although the results below show that this approximation works quite well, we hope that the success of the whole scheme presented in this paper will inspire research that will enable us to remove this restriction in the future.

To summarize, two, conceptually different methods will be used in this study. One is a QM/MM approach based on CHELPG point charges, augmented with the (ground state) dispersion and Pauli repulsion from EFP2 (termed as *QM/MM + EFP2* hereafter), while the other is a Huzinaga projection based embedding model (PbE) augmented with an EFP2 (ground state) dispersion contribution (termed as *PbE + EFP2*).

2.3 Coupling schemes

The off-diagonal elements of the Hamiltonian matrix (Eqn. (10)) represent the coupling between the excitations of the chromophore groups. To consider possible approximations of $\tilde{V}(1,2)$, let us use the original partitioning of the Hamiltonian (Eqn. (1)) which treats both fragments equivalently. With this Hamiltonian it is more reasonable to use the unbiased basis (Eqn. (2)), to evaluate the matrix elements. This approximation is not severe since the considered non-covalent interactions presumably do not change the fragment wave functions

heavily. This means that

$$\begin{aligned}\tilde{V}(1,2) \approx V(1,2) &= \langle \Phi_{1,1}(\mathbf{r}_1)\Phi_{2,0}(\mathbf{r}_2)|\hat{H}_1(\mathbf{r}_1) + \hat{H}_2(\mathbf{r}_2) + \hat{V}_{1,2}(\mathbf{r}_1,\mathbf{r}_2)|\Phi_{1,0}(\mathbf{r}_1)\Phi_{2,1}(\mathbf{r}_2)\rangle \\ &= \langle \Phi_{1,1}(\mathbf{r}_1)\Phi_{2,0}(\mathbf{r}_2)|\hat{V}_{1,2}(\mathbf{r}_1,\mathbf{r}_2)|\Phi_{1,0}(\mathbf{r}_1)\Phi_{2,1}(\mathbf{r}_2)\rangle, \end{aligned} \tag{17}$$

i.e. the coupling is just the matrix element of the interaction potential $\hat{V}_{1,2}(\mathbf{r}_1,\mathbf{r}_2)$.

A simple way to calculate this term is to approximate the potential within the dipole approximation, first suggested by Förster[57]

$$\hat{V}_{1,2}(\mathbf{r}_1,\mathbf{r}_2) \approx \frac{\hat{\boldsymbol{\mu}}_1(\mathbf{r}_1)\cdot\hat{\boldsymbol{\mu}}_2(\mathbf{r}_2)}{R_{1,2}^3} - 3\frac{(\hat{\boldsymbol{\mu}}_1(\mathbf{r}_1)\cdot\mathbf{R}_{1,2})\,(\hat{\boldsymbol{\mu}}_2(\mathbf{r}_2)\cdot\mathbf{R}_{1,2})}{R_{1,2}^5}, \tag{18}$$

with $\mathbf{R}_{1,2}$ being the vector connecting the center-of-mass (COM) of the two choromophores, R_{12} being the distance between them, and $\hat{\boldsymbol{\mu}}_i(\mathbf{r}_i)$ is the dipole moment operator of fragment i.

With this approximation the coupling can be calculated from the *transition dipoles* of the fragments:

$$V(1,2) = \frac{\boldsymbol{\mu}_1^{10}\boldsymbol{\mu}_2^{01}}{R_{12}^3} - 3\frac{(\boldsymbol{\mu}_1^{10}\cdot\mathbf{R}_{12})(\boldsymbol{\mu}_2^{01}\cdot\mathbf{R}_{12})}{R_{12}^5}, \tag{19}$$

with

$$\boldsymbol{\mu}_i^{kl} = \langle \Phi_{i,k}(\mathbf{r}_i)|\hat{\boldsymbol{\mu}}_i(\mathbf{r}_i)|\Phi_{i,l}(\mathbf{r}_i)\rangle.$$

This approach, called the *Transition Dipole Approximation (TrDA)* hereafter, is a simple and cheap way to estimate the coupling. Note that the second term in Eq. 19 is needed only if the two transition dipole moments are not parallel, i.e. it does not contribute for the stacked systems treated in this paper.

The dipole approximation can be extended by considering transition monopoles e.g. at atomic sites[58] or by higher multipoles like in an extension of the PMM (Perturbed Matrix Method) appraoch (PMM-QQ).[59] However, the latter study reveals, that the extension in this form does not necessarily improve the accuracy of the coupling.

A more accurate approximation of $\hat{V}_{1,2}$ is to consider the exact Coulomb interaction between the electrons of the two fragments

$$\hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i(1)} \sum_{j(2)} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{20}$$

where the indices $i(1)$ and $j(2)$ run over the electrons of fragment 1 and 2, respectively. (The contribution of the nuclei to V vanishes in the commonly adopted Born-Oppenheimer and Condon approximations.[60]) With this the coupling becomes:

$$\begin{aligned} V_{1,2} &= \int d\mathbf{r}_1 \; d\mathbf{r}_2 \; \Phi_{1,1}(\mathbf{r}_1)\Phi_{1,0}(\mathbf{r}_1) \; \hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2) \; \Phi_{2,0}(\mathbf{r}_2)\Phi_{2,1}(\mathbf{r}_2) \\ &= \int d\mathbf{r}_1 \; d\mathbf{r}_2 \; \rho_1^{0\to 1}(\mathbf{r}_1) \; \hat{V}_{1,2}(\mathbf{r}_1, \mathbf{r}_2) \; \rho_2^{0\to 1}(\mathbf{r}_2) \end{aligned} \tag{21}$$

i.e. it can be calculated from the *transition densities* ($\rho_i^{0\to i}$) of the two fragments.[60,61]

Krueger and co-workers[61] suggested the evaluation of this expression on a grid and named the procedure the Transition Density Cube (TDC) method,[61,62] we keep this nomenclature in this paper.

We have implemented the grid representation of transition densities in the CFOUR[63,64] program system, which enables us to calculate within the numerical accuracy the electronic Coulomb interaction between the fragments at any electronic structure level for which the transition density is available.

There are also other attempts in the literature to include accurate couplings for Frenkel-type excitonic interaction with specific electronic structure method: CIS by Herbert and co-workers,[22] TDDFT by Martinez and co-workers[65] as well as by Head-Gordon and co-workers,[66] CC2 and related methods by Fückel et al.[60] These often also include the exchange type (Dexter) coupling,[67] which is relevent only at very short distances. Since the Dexter coupling is not available at the CCSD level, we can not include this term here. However, from the results presented we can obtain the distance where this term becomes important.

Note that the same coupling also appears in Electron Energy Transfer (EET) processes,[68] and in this regard many approximate schemes have been suggested. Overview of these

techniques is out of scope of the present paper, more detail can be found e.g. in Ref. 69.

3 Molecular systems

To find the best–suited approximation for the excited states of non-covalently interacting dimers of nucleobases, in the present paper we investigate bimolecular complexes of nitrogen–containing heterocycles: stacked pyrrole-pyrole dimer (denoted as *(Pyr)*$_2$ hereafter), and cytosine-uracil complex (*Cyt-Ura*) as in our previous paper on the ground state.[16] In addition, we include the homodimers of cytosine (*(Cyt)*$_2$) and uracil (*(Ura)*$_2$) to better understand trends, as well as that of formaldehyde (*(*CH_2O*)*$_2$) as a smaller complex. The test systems under study are illustrated in Fig. 1. In the discussion below, the distance of monomers refers to that between the centers of masses. For the *(Pyr)*$_2$ stacked conformer the fragments' dipoles point essentially in the opposite direction, but the symmetry was lowered by a 10 degree in-plane rotation away from the C_{2h} structure. The *Cyt-Ura* , *(Cyt)*$_2$, *(Ura)*$_2$, and *(*CH_2O*)*$_2$ complexes were investigated in a stacked (sandwich) setup, with oppositely oriented dipole moments. The equilibrium structures of the monomers, optimized at the MP2/6-31G* level, were taken from Ref. 70 and are documented in Tables S1-S4 of the *Supplementary material* along with the geometries of dimers at a representative distance in Tables S5-S9.

We envision application of the new methods not just in single point calculations, but rather in exploration of potential energy surfaces. Therefore, the tests presented in this paper include potential curves along the intermolecular separation. Although the proposed methodology can be generalized for practically any number of states, in this first investigation we stay with the interaction of just two states in order to better understand how the different approximations work in the presented scheme.

To investigate the interaction of just a pair of states, these need to be energetically well separated from the other states. It turned out that selecting such pairs is not an easy task, in particular in calculations with diffuse basis functions: Rydberg states present among the excited states of molecules with π electronic structure result in a high density of states with strong interactions between them. In addition, at short distances, also charge transfer (CT)

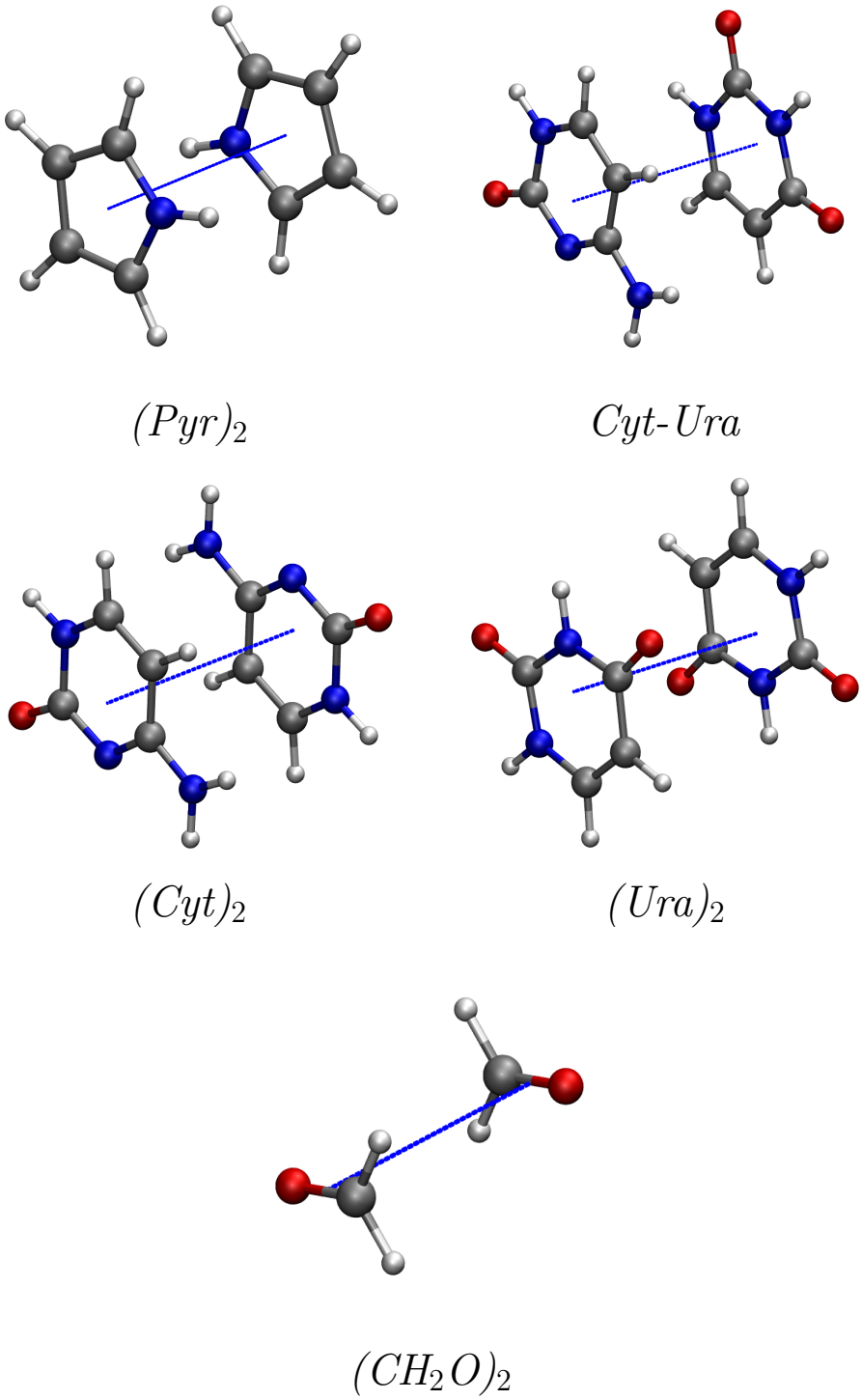

Figure 1: Orientation of the molecules in the test systems used in this study. The measure of the distance is represented by the blue dotted line connecting the centers of mass of the fragments.

states complicate the spectrum. Therefore, to find appropriate states for the tests, certain compromises were necessary, which will be discussed later. The monomer electronic states considered in this study are summarized in Table 1.

Table 1: Summary of the monomer electronic states investigated in this study.

	Basis set	State	ΔE [a] / eV	f [b]
CH_2O	cc-pVDZ	$1\,^1A_1$ (ground)	0.00	
		$1\,^1B_1$ ($\sigma - \pi^*$)	9.36	0.002
		$2\,^1A_1$ ($\pi - \pi^*$)	9.95	0.011
CH_2O	aug-cc-pVDZ	$1\,^1A_1$ (ground)	0.00	
		$2\,^1A_1$ ($n - R$)	8.07	0.058
		$1\,^1B_1$ ($\sigma - \pi^*$)	9.24	0.001
		$3\,^1A_1$ ($\pi - \pi^*$)	9.59	0.166
Pyrrole	cc-pVDZ	$1\,^1A_1$ (ground)	0.00	
		$1\,^1B_2$ ($\pi - \pi^*$)	7.03	0.146
Pyrrole	aug-cc-pVDZ	$1\,^1A_1$ (ground)	0.00	
		$1\,^1B_1$ ($n - R$)	5.14	0.000
		$1\,^1A_2$ ($n - R$)	5.87	0.021
Cytosine	cc-pVDZ	$1\,^1$A' (ground)	0.00	
		$3\,^1$A' ($\pi - \pi^*$)	6.07	0.157
Uracil	cc-pVDZ	$1\,^1$A' (ground)	0.00	
		$2\,^1$A' ($\pi - \pi^*$)	5.78	0.199

[a] EOM-CCSD vertical excitation energies, in electron volts; [b] Oscillator strength, evaluated at the EOM-CCSD level

4 Computational details

The primary goal with developing the fragment methods is to replace high level Coupled Cluster (CC) type electronic structure calculations on the complex with appropriate calculations on the fragments. The family of CC-type methods allows a systematic improvement of the accuracy, but approximate versions, like CC2[71] and ADC(2),[72] in particular with spin scaling,[73,74] enable us to increase the size of the fragments, as well. Still, to avoid any ambiguity caused by the approximate methods, the first tests presented in this paper will be performed at the Equation of Motion Coupled Cluster method with Singles and Doubles (EOM-CCSD) level,[75,76] which can be easily performed not only for middle sized fragments, but also for their complexes.

Double-ζ quality basis set with diffuse functions has been selected for the calculations,

since it gives a reasonable description of excited states and, at the same time, of intermolecular interaction,[77] while the reference EOM-CCSD calculations are also feasible on the complexes. However, to be able to investigate valence $\pi - \pi^*$ excitations, the primary interest for later applications, some calculations had to be performed without diffuse functions, even if the interaction energies obtained this way are clearly underestimated. Nevertheless, one has to proceed with great care while performing comparisons without diffuse functions, details are given below.

To summarize, all reference, QM/MM and high-level steps of the PbE calculations have been performed at the CCSD[78] and EOM-CCSD[75,76] level of theory with the core electrons excluded from the correlation treatment. Two basis sets were used in these calculations: cc-pVDZ[79] and aug-cc-pVDZ.[80] Only calculations using the same method and basis have been compared allowing the evaluation of the performance of the fragment methods. More details of the calculations are discussed in the next subsections.

4.1 Reference Ab initio calculations

To evaluate the accuracy of the new scheme, comparison to appropriate reference calculations is necessary. It turned out that these presented a major bottleneck of this study.

First, it was a major challenge to identify and converge valence excited states when diffuse functions were included in the basis, due to the presence of large number of Rydberg states and their strong mixing. This was not the case when the basis without diffuse functions was used. Clearly, this way just a limited portion of the interaction energy is captured in the model, but with a careful selection of the parameters of the fragment calculations (see below) meaningful comparisons can be made.

Second, the calculations on the complex need to be corrected for Basis Set Superposition Error (BSSE) since the fragment calculations are free from this artificial lowering of the interaction energy. In our previous paper[16] we used Counterpoise Correction (CP)[81] to correct the ground state energies of the complex. Although it is often claimed that CP corrections overestimate BSSE, in particular, in correlated calculations,[82] it is still the state of the art in most studies of non-covalent interactions.[83] Note that the SNOOP method by Jorgensen et al.[84] is a noteworthy attempt to calculate a CP correction for correlated ground

states, but requires special code which is at present only available in the LSDalton program suite.[85]

Not much literature exists about BSSE corrections for excited states . We have performed some test calculations, just applying the usual CP scheme for excited states. The results are shown in Figure 2 for $(CH_2O)_2$: while for some low-lying states the CP correction is close to the ground state value at any distance, some states show unreasonably large CP correction with strong growth with decreasing intermolecular distance. A quick look into this problem revealed that the CP correction removes the charge transfer component of the solution, which is an artifact: most excited states of interacting fragments (can) have indeed charge flow. Considering that CP correction is a basis set effect and influence mostly the orbitals,[82] we decided to use the CP correction obtained at the HF level for both the ground state and also for the excited states. Note that this way the CP correction will not change the excitation energy, but will indeed influence the shape of the excited state energy surfaces. It is certainly worth investigating the problem in more detail, and a project has been started already in our laboratory. Nevertheless, the results presented in this paper show that this approximation is reasonable.

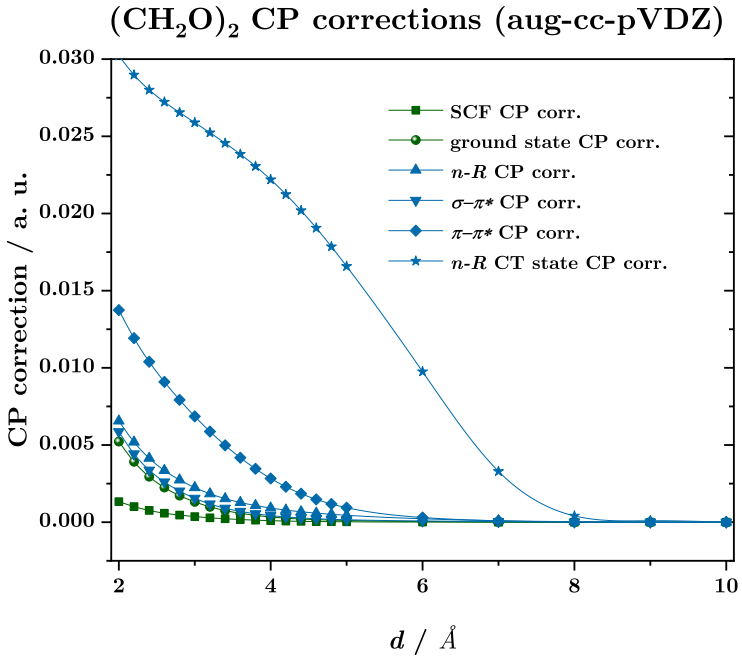

Figure 2: CP correction calculated for several states of $(CH_2O)_2$

All reference calculations have been performed using the CFOUR program package.[63,64]

4.2 QM/MM calculations

The QM/MM calculations were performed at the same levels as the reference calculations using the CFOUR[63,64] program package, with the CHELPG point charges incorporated into the one-electron Hamiltonian.

The CHELPG point charges were obtained at the CCSD level using GAMESS[86] and aug-cc-pvDZ basis. The grid with a step size of 0.8 Å and the most distant point from any atom of 3 Å was used to represent the electrostatic potential in the fit, with the additional requirement to reproduce the dipole and quadrupole moments of the molecule, except for the formaldehyde molecules, where the restriction was set to the dipole moment only, to keep the number of parameters correct. The CHELPG atomic charges are also found in Tables S1-S4 of the *Supplementary material*.

4.3 Embedding calculations

The PbE calculations were performed using the MRCC program suite[87] with the resolution of identity approximation used for both the SCF and correlation parts of the calculation. The active subsystem was treated at the same level of theory as the reference calculations and the PBE functional[88] was used as the low level method. The localisation and partitioning of the occupied and virtual orbital spaces was done using SPADE.[43] The energy expression suggested in Eqn. (16) was used as the diagonal term in the Frenkel matrix.

Calculations on all systems used localised virtual orbitals, otherwise excited states spectrum was spoiled by artificial CT states. However, the $(CH_2O)_2$ dimer was also investigated using the full virtual space (in ground state calculation), as well as an extended SPADE-localised virtual space containing two additional orbitals, to test the incorrect behavior of the method with diffuse basis functions (see later).

4.4 EFP2 calculations

The EFP2 dispersion and Pauli repulsion potentials were calculated on the restricted Hartree-Fock level using the built-in routine of GAMESS[86] with default settings. The basis sets of these calculations were adjusted to the electronic structure calculations: for the augmented basis tests these interaction terms were calculated on the 6-311++G(3df,2p) basis set as suggested by Slipchenko and Gordon.[89] On the other hand, to compare with the reference results obtained at the cc-pVDZ basis, which clearly includes only a smaller portion of the interaction energy, the EFP2 parameters were also obtained with this non-augmented basis set. Whether EFP2 scheme gives reasonable interaction energies in this case, will be checked below.

4.5 Calculation of interstate couplings

The transition moments and transition densities of the investigated excited states were determined from monomer calculations on the same level as the reference calculations using CFOUR.[64] In the TDC scheme the transition densities were represented on a COM centered grid of stepsize of 0.25 Å with maximum dimension of 16-by-20-by-24 Å for formaldehyde and 20-by-20-by-20 Å for all other monomers corresponding to the Cartesian coordinates given in Tables S1-S4 in the *Supplementary material*. The numerical accuracy of the TDC procedure was evaluated by integrating the transition density for the entire space, as well as by comparing the transition moment obtained on the grid with those from the EOM-CC calculation.

Due to the bi-orthogonality of the CC framework, the above schemes need further adjustment. There are (slightly) different left and right transition properties which would result in slightly different interaction energies. Following the procedure that the physically relevant oscillator strength is calculated from the product of left and right transition dipoles, we here suggest to use the geometric mean of the left and right transition properties to evaluate the coupling.

5 Results and discussion

In this section the different approximations defining our scheme are tested. The validity of approximations of the interaction energy are investigated in Section 5.1, the performance of different coupling schemes is presented in Section 5.3, while the final results, i.e. the total PESs of both states in question, are discussed in Section 5.1.2. In addition, in Subsection 5.2 the problem related to the virtual space localization of the PbE scheme is discussed.

5.1 Interaction energy

5.1.1 Ground states

First, we review the quality of the ground state interaction potentials including the EFP2 contributions for the setups used later in the excited state calculations. In Figure 3 relative potential energy curves of the ground states of $(CH_2O)_2$ and $(Pyr)_2$ dimers are shown, calculated with aug-cc-pVDZ basis. In the case of $(CH_2O)_2$, very good agreement can be observed for the *QM/MM + EFP2* curve, the difference to the reference getting clearly observable only at very short distances close to the equilibrium. The discrepancy seems to be larger for the $(Pyr)_2$ dimer, but one should recognize that the interaction energy is almost one order of magnitude smaller here than in the $(CH_2O)_2$ system.

On the other hand, the *PbE + EFP2* curves deviate considerably from the reference ones. This problem turns out to be a general flaw of the embedding approach and is discussed in detail in Subsection 5.2.

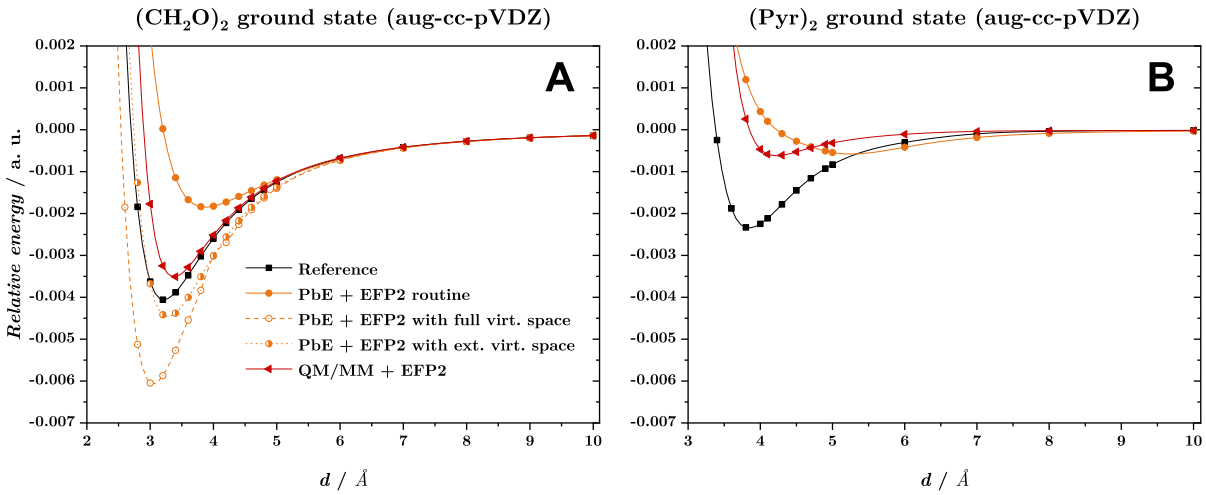

Figure 3: Relative potential energy curves of the ground states of the $(CH_2O)_2$ (Panel A) and $(Pyr)_2$ (Panel B) dimers, calculated with different models using the CCSD/aug-cc-pVDZ method as their wave function component.

In Figure 4 the relative potential energy curves in the ground states of the investigated dimers are shown, calculated with cc-pVDZ basis set.

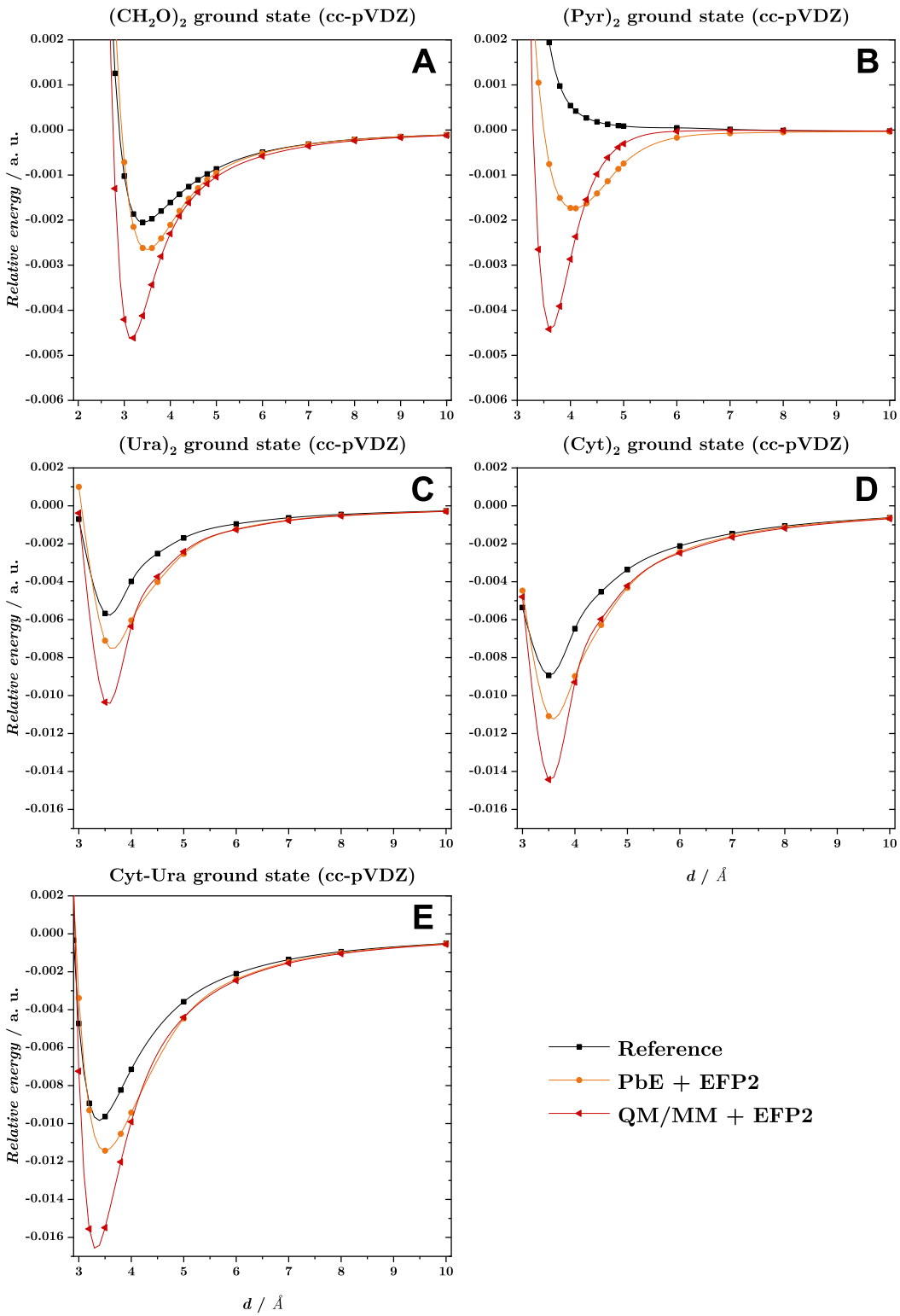

Figure 4: Relative potential energy curves of the ground states of the $(CH_2O)_2$ (Panel A), $(Pyr)_2$ (Panel B), $(Ura)_2$ (Panel C), $(Cyt)_2$ (Panel D), and *Cyt-Ura* (Panel E) complexes, calculated with the different models using CCSD/cc-pVDZ method as their wave function component.

Figure 4 shows that both *QM/MM + EFP2* and *PbE + EFP2* models overestimate the interaction energy for all systems, resulting in potential energy curves running below the reference ones. While at larger distances the two models give very similar results, at shorter distances the *PbE + EFP2* curve is less bound. This indicates that the dispersion is overestimated at the EFP2/cc-pVDZ level while the Pauli repulsion is better described within the *PbE + EFP2* model than with EFP2/cc-pVDZ. Note that the ground state of the $(Pyr)_2$ dimer is not bound at the CCSD/cc-pVDZ level, thus the comparison is less relevant in that case.

5.1.2 Excited states

For homodimers it is possible to separately investigate the quality of interaction potentials without the effect of the coupling, by comparing the EFP2 corrected QM/MM and PbE curves to the average of the two reference ones. We do this here focusing our attention on the most important, 4-10 Å range of separation. For states showing a minimum with respect to the intermolecular distance, this range covers the practically relevant parts of the attractive region.

In Figure 5 the distance dependence of various interaction potentials between two formaldehyde molecules, one in a valence excited state and the other in ground state, are plotted. The calculations are done at the CCSD/aug-cc-pVDZ level. A relatively good agreement of the *QM/MM + EFP2* curve and the averaged reference curve is observed for the $\sigma \rightarrow \pi^*$ state, suggesting that this model correctly describes the interaction and the ground state Pauli exchange and dispersion interactions are well suited for this excited state. In the $\pi - \pi^*$ state, similar agreement is observable only until 6 Å separation, below which the *QM/MM + EFP2* result deviates significantly from the reference. A closer analysis reveals a Charge Transfer (CT) state crossing the $\pi - \pi^*$ curve in this region, causing the fast descent of the reference curves (see Fugure S3 in *Supplementary material*). Since the effect of CT states is not included in the fragment model at the present stage, these curves obtained with the fragment models are only reasonable above 6 Å.

The curve obtained by PbE + EFP2 scheme is, on the other hand, repulsive, similarly to the ground state case seen above. The reason of this failure is discussed below in more

detail.

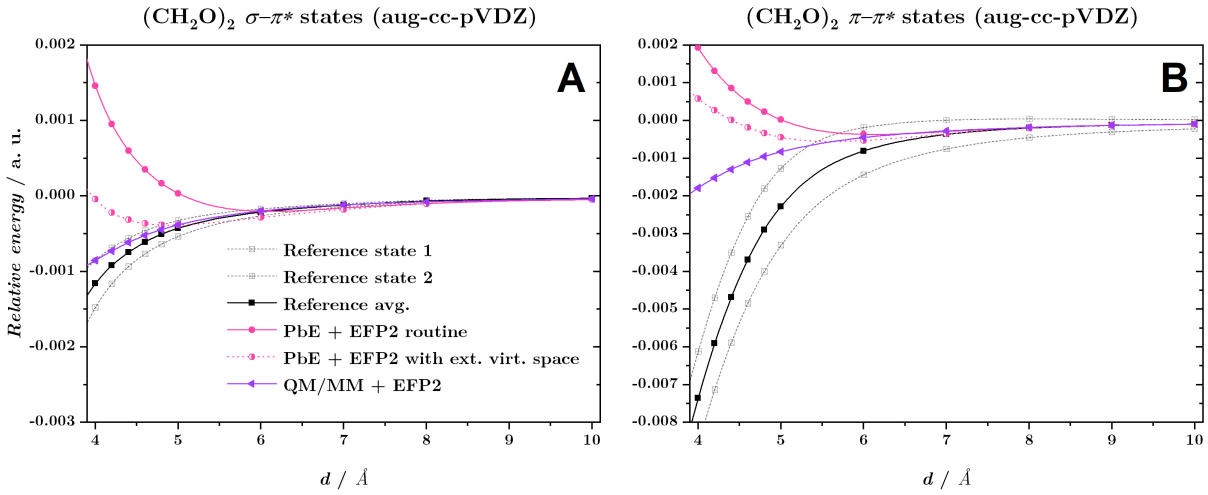

Figure 5: Distance dependence of various interaction potentials (not including interstate couplings) in the $\sigma - \pi^*$ (Panel A) and $\pi - \pi^*$ (Panel B) local valence excitations of the $(CH_2O)_2$ dimer, calculated with different models using CCSD/aug-cc-pVDZ method as their wave function component.

In Figure 6 the same comparisons are presented for valence excitations of the homodimers, obtained with the cc-pVDZ basis. Now the $\pi - \pi^*$ state of $(CH_2O)_2$ can also be considered, since no interaction with CT states are present in the reference. The *QM/MM + EFP2* curves overestimate the interaction energy in all cases, just like they did in the ground state. The *PbE + EFP2* results are remarkably close to the *QM/MM + EFP2* ones, except in the case of $(Pyr)_2$ where it produces an early minimum, deviating from the attractive *QM/MM + EFP2* and reference curves.

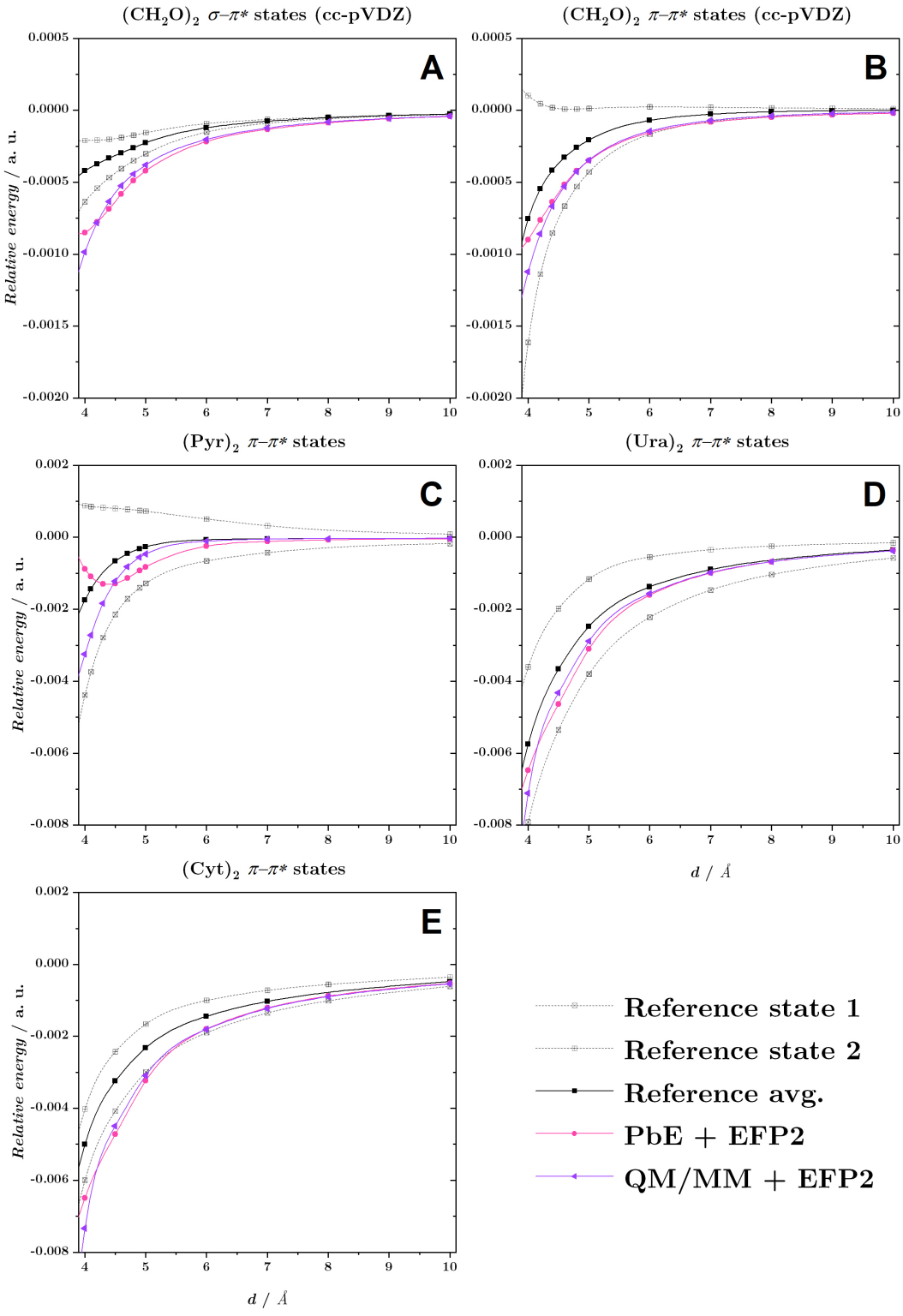

Figure 6: Distance dependence of various interaction potentials (not including interstate coupling) in valence excited states of the $(CH_2O)_2$ (Panels A and B), $(Pyr)_2$ (Panel C), $(Ura)_2$ (Panel D), and $(Cyt)_2$ (Panel E) homodimers calculated with different models using CCSD/cc-pVDZ method as their wave function component.

While the present results are encouraging, one has to keep in mind that the EFP2 corrections used here were calculated with the cc-pVDZ basis set. Obviously there is no evidence that EFP2 could reproduce the interaction energy of the reference dimer calculations in this rather artificial situation. This question is, however, of moderate importance here since in real applications the fragment methods should and will be used with an appropriate basis set. On the other hand, we can answer the question whether the vdW corrections obtained for the ground state represent a reasonable approximation for the case of an interacting ground state fragment and one in its excited state. To that end, we have calculated the 'exact' vdW contribution of the ground state as the energy difference of the reference ground state energy and that of the uncorrected QM/MM and PbE energies and used this quantity as an approximate vdW correction for the excited states. The corresponding curves are shown on Figure 7 for the $(Ura)_2$ and $(Cyt)_2$ systems. One can see a much better agreement with the averaged reference curve in both cases, the curves from our models running slightly above the reference ones. This means that the vdW correction from the ground state is slightly smaller than it should be in the excited state. The solid agreement of the QM/MM and PbE results indicates that it is mostly the dispersion part which is underestimated in this case.

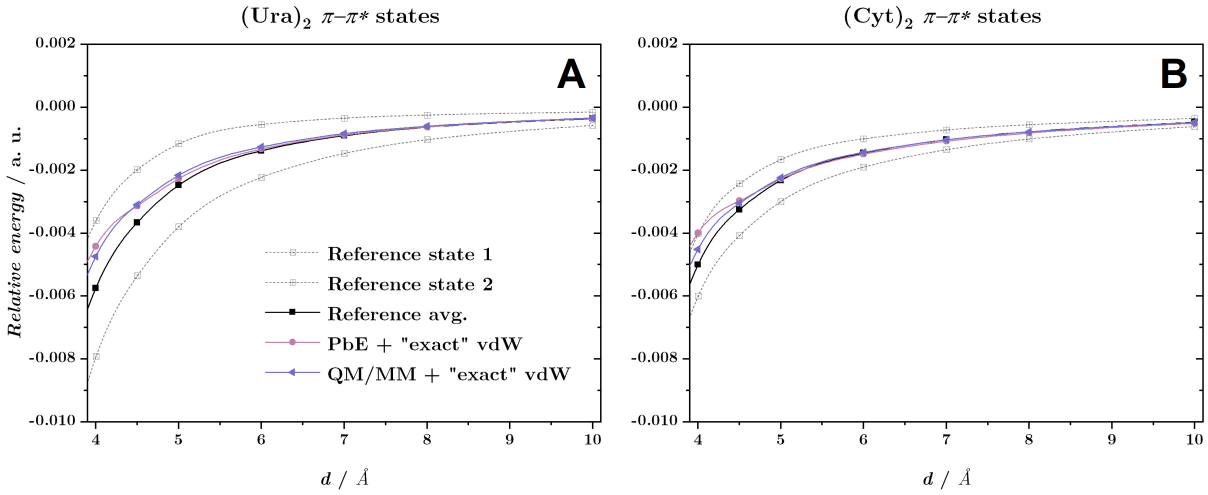

Figure 7: Distance dependence of various interaction potentials in the $\pi - \pi*$ excited states of the $(Ura)_2$ (Panel A) and $(Cyt)_2$ (Panel B) homodimers, with the vdW correction calculated from the ground state reference potential, at the CCSD/cc-pVDZ level of theory.

Continuing the investigation with the inspection of Rydberg states, in Figure 8 the distance dependence of various potentials for the Rydberg states of $(CH_2O)_2$ and $(Pyr)_2$ are

shown, calculated with the aug-cc-pVDZ basis. A striking feature one observes here is the clearly inappropriate, strongly repulsive curve produced by the by *PbE + EFP2* model (see discussion below). For *QM/MM + EFP2* the situation is much better, the curves are qualitatively correct, although the interaction energy is (except perhaps in the case of the 2nd Rydberg state pair of $(Pyr)_2$) is slightly too small.

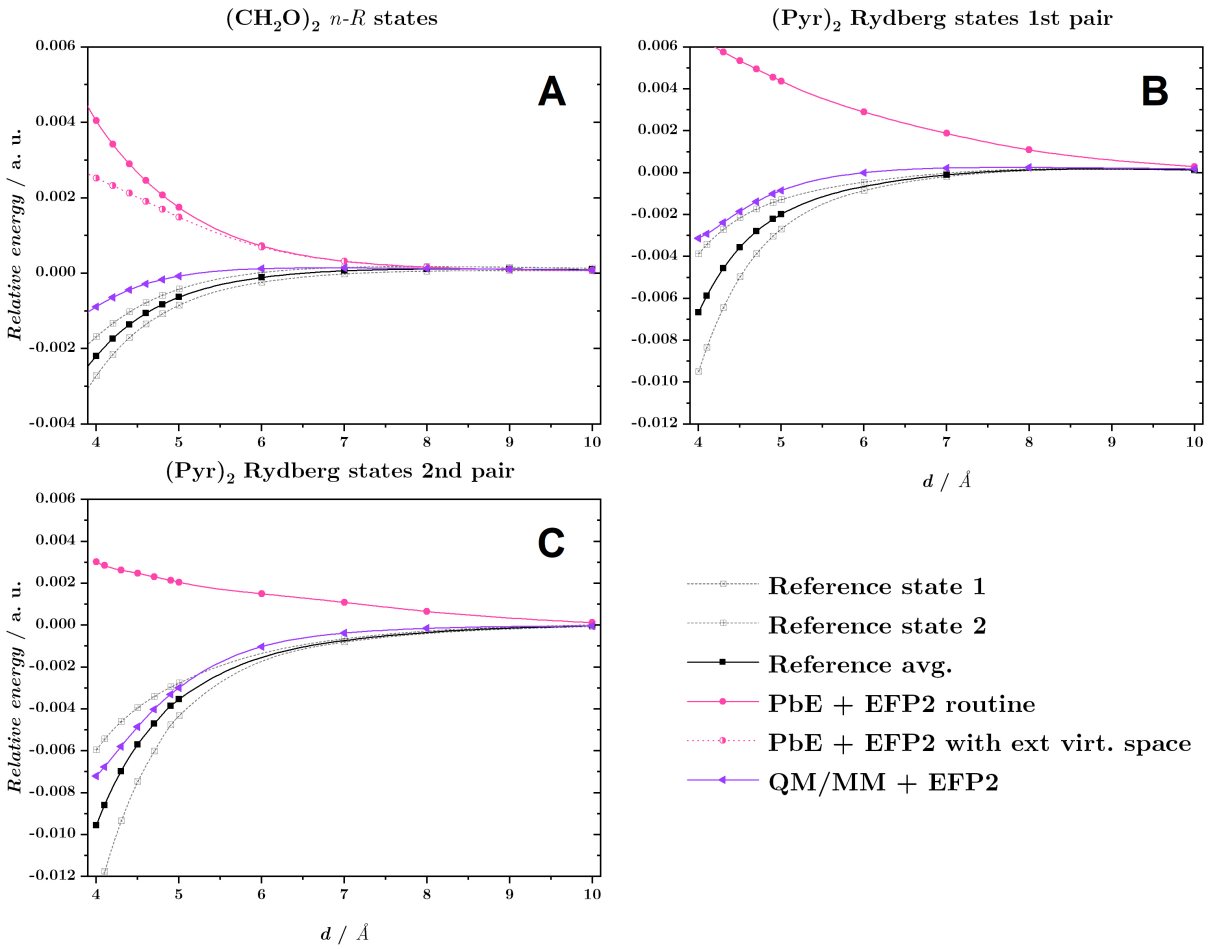

Figure 8: Distance dependence of various interaction potentials in Rydberg excited states of the $(CH_2O)_2$ (Panel A) and $(Pyr)_2$ (Panels B and C) homodimers, calculated with different models using CCSD/aug-cc-pVDZ method as their wave function component.

5.2 Failure of the Huzinaga embedding scheme with diffuse basis sets

In the above discussions we have observed that the *PbE + EFP2* scheme worked very well for valence excited states if used with the cc-pVDZ basis, but repulsive potential curves

have been obtained if the calculations are performed with the aug-cc-pVDZ basis set. The possibility that the failure is caused by the EFP2 dispersion term can be excluded, since the same dispersion is used in *QM/MM + EFP2*, which works excellently.

On the other hand, an important feature of the present PbE scheme is the restriction of the virtual space to the orbitals localized on the active fragment. As discussed earlier in Subsection 2.1.2, this step is necessary to avoid the appearance of artificially too many charge transfer type CC excited states and to keep the cost of the correlated calculation to that of a single fragment. This operation can be, however, a serious distortion of the virtual space if diffuse functions (Rydberg orbitals) are present whose localization to the individual fragments is less obvious.

To illustrate that this restriction causes the problem, we have performed additional calculations on the $(CH_2O)_2$ complex including a) all virtual orbitals of the complex (denoted as *PbE + EFP2 with full virtual space*) and b) two additional, normally excluded virtual orbitals that have the largest contribution on the active fragment (*PbE + EFP2 with extended virtual space*). The results are included in Figure 3, Figure 5 and Figure 8 for the ground state, for the valence excited states and for Rydberg states, respectively.

In case of the ground state (Figure 3), even with just two additional orbitals in the extended virtual space, the potential becomes attractive, in fact too attractive, especially if all virtual orbitals are included. Neither choice seems to be appropriate: with all orbitals included in the virtual space we see a special kind of BSSE since by decreasing the distance between the fragments, more and more virtual orbitals become available in the spatial region of the active fragment, an effect which is expected to be strong in particular with diffuse basis functions. On the other hand, restricting the virtual space to that of the active fragment causes a "reverse BSSE": by decreasing the distance, the diffuse functions of the active fragment will spatially overlap with those of the environment, causing the localization procedure to assign them to the latter. Although this effect is only visible with diffuse basis sets, it should eventually appear also with non-diffuse basis sets at very small distances.

In case of the valence excited states (see Figure 5) the *PbE + EFP2* curve is repulsive again, but now already with the extended virtual orbital space a minimum is observed. The $\sigma - \pi^*$ state is very close to the *QM/MM + EFP2* curve showing deviation only at shorter

distances. Also for the $\pi - \pi^*$ state the agreement is excellent until the the critical distance of 6Å (see above). Note that the identification of proper states with the full virtual space were not possible due to the proliferation of artificial CT states.

The concept of "reverse BSSE" is strongly supported by the fact that for Rydberg excited states the *PbE + EFP2* scheme always produces strongly repulsive potential curves (see Figure 8): here the effect is more severe since diffuse orbitals are needed to describe the dominant excitation but only orbitals spatially restricted to the active fragment are available. This effect is demonstrated on a video in the *Supplementary material*. Note that not even the above described extension of the virtual space does cure the problem here, the curve produced by embedding with extended virtual space is still repulsive.

In summary, the present PbE scheme does not work with diffuse basis sets, and a new version is warranted with proper definition of the virtual space that includes all the virtual orbitals of the active fragment, but none of the other.

5.3 Quality of the coupling schemes

The other factor that plays key role in reproducing the energy levels of interacting states is the accuracy of the coupling between local excitations, as defined in the models presented in Subsection 2.3. The models that include the coupling terms predict a certain splitting (in the case of homodimers) or a shift (for non-symmetric complexes) of the energy levels due to the interstate interaction. To evaluate the quality of the couplings, these energy changes have to be compared to those of the respective reference calculations, i.e., high-level calculations on the entire complex.

Figures 9 – 12 show the energy splittings of the investigated symmetric complexes as functions of the intermolecular separation. For the $(CH_2O)_2$ system with aug-cc-pVDZ basis (Figure 9), a large dependence of the results on the type of electronic state is apparent. In the $\sigma - \pi^*$ state the transition dipole approximation fails completely, predicting a practically zero splitting of the states. The TDC approach, on the other hand, gives very accurate results throughout the investigated range, indicating that in this state the interaction is appropriately described by the electrostatic model if the charge distribution is considered at a high resolution. For the $\pi - \pi*$ state the TDC curve is correct until 6Å and the

deviation from the reference is still moderate around 5Å . Below this point, as discussed above, the states in the reference calculations show strong CT character, thus our model is not appropriate. In the n-Rydberg state the coupling is overestimated by the transition dipole approximation, while the prediction based on TDC gives correct results well down to a 4.5Å intermolecular distance, where some deviation from the reference curve starts to show up, presumably the exchange and overlap contributions of the coupling also become significant that are ignored in the TDC model.

We note that the TDC results also have an obvious dependence on the electronic structure method through the accuracy of the transition density matrix predicted for the state in question. In this regard, both a close similarity as well as a significant discrepancy between the CC2 and CCSD levels could be observed, depending on the electronic state. Details are shown in Figure S2 of the *Supplementary material*.

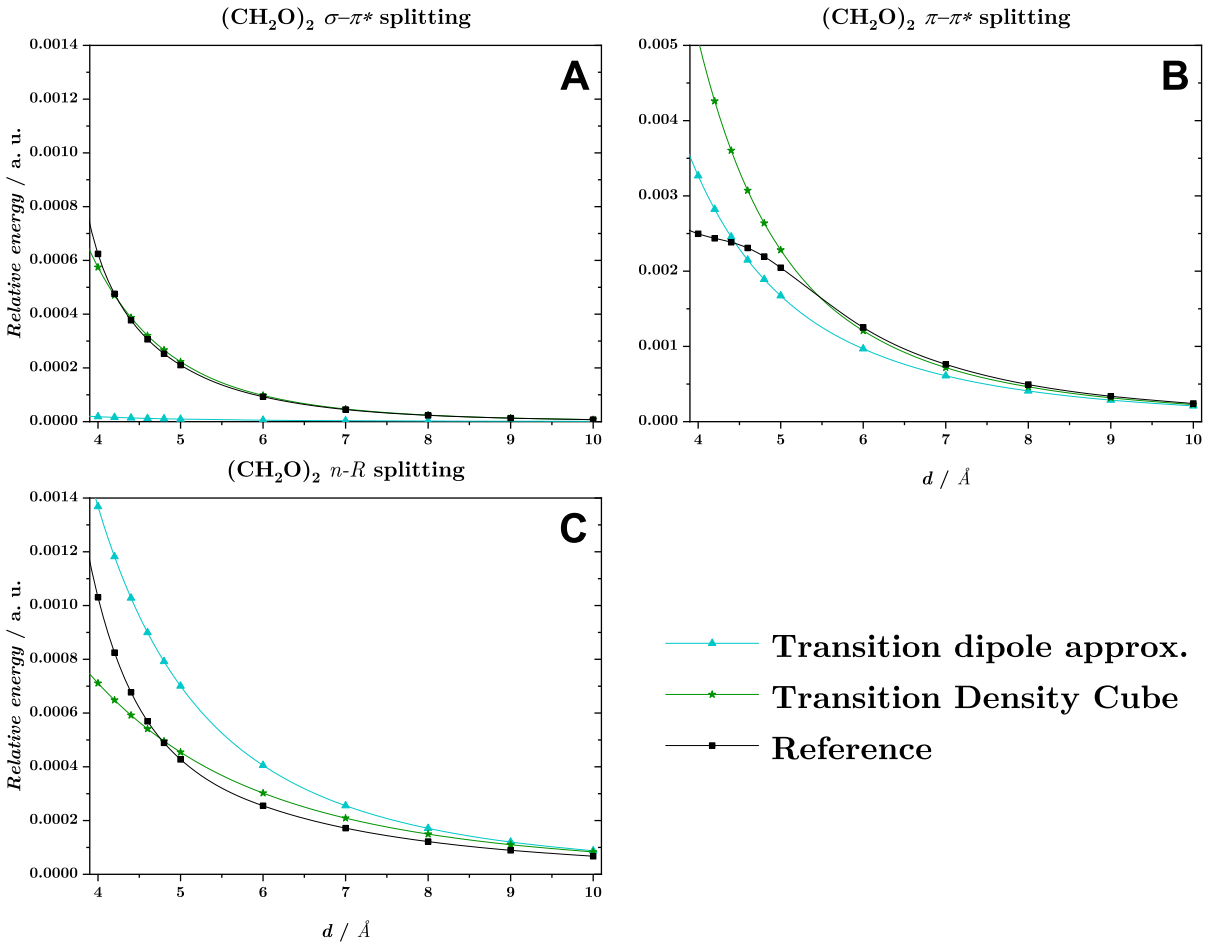

Figure 9: Dependence of the splitting (in atomic units) of the energy levels of the $(CH_2O)_2$ dimer on the intermolecular distance, as predicted by various coupling schemes.

Fig. 10 shows the couplings in the $\pi - \pi^*$ states of the $(Pyr)_2$, $(Cyt)_2$ and $(Ura)_2$ dimers. (Note that cc-pVDZ basis set was used for these calculations.) In these valence states, an inconsistent behavior of the transition dipole approximation is observed. While it underestimates the coupling for $(Pyr)_2$ below 4.5Å, a remarkable overestimation is seen for the $(Cyt)_2$ and $(Ura)_2$ states even at larger separations. TDC, however, closely follows the reference curve in all states in the range where the Coulomb contribution still dominates the coupling (about 5 Å for $(Pyr)_2$, 4 Å for $(Cyt)_2$ and 3.5 Å for $(Ura)_2$). Below these points, it starts to deviate from the reference, indicating the importance of the missing exchange (Dexter) contributions to the splitting. Since for the latter two systems this point falls comfortably close the ground state equilibrium separation, it can be inferred that the TDC model provides a reliable prediction of the splitting in the entire attractive side of the

interacting PES.

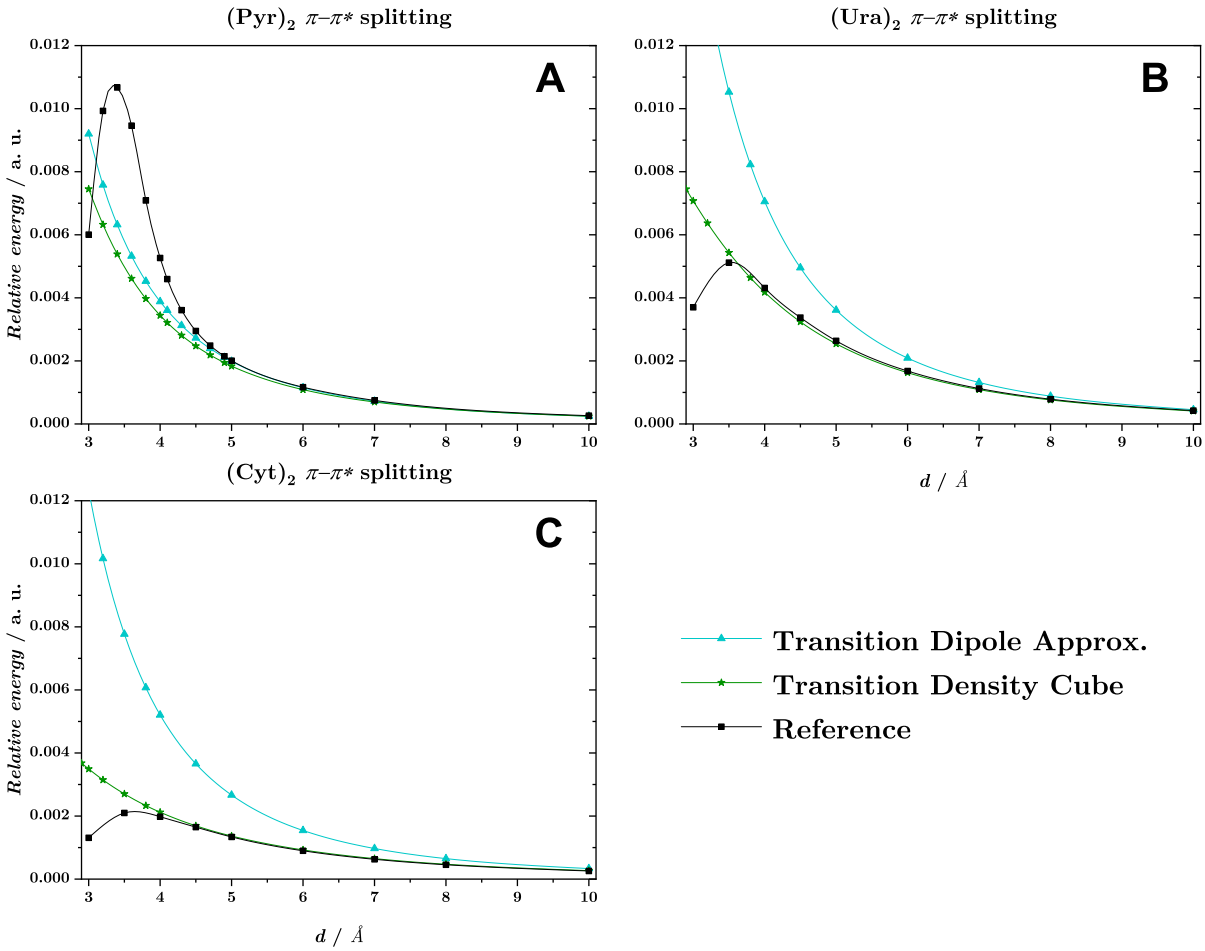

Figure 10: Dependence of the splitting (in atomic units) of the $\pi - \pi*$ energy levels of the *(Pyr)*$_2$ (panel A), *(Cyt)*$_2$ (panel B) and *(Ura)*$_2$ (panel C) dimers on the intermolecular distance, as predicted by various coupling schemes.

The energy splitting of the corresponding states in the asymmetric *Cyt-Ura* complex is shown on Fig. 11. Since in this case the energy gap is nonzero even at infinite separation, the splitting is defined here as

$$\Delta E_{split}(R) := \Delta E(R) - \Delta E(\infty), \tag{22}$$

for all methods.

The change of the gaps with the distance is much smaller in this complex than in either the *(Cyt)*$_2$ or the *(Ura)*$_2$ systems. The transition dipole approximation shows a behavior

similar to the above cases, severely overestimating the growth of the energy gap with the fragments approaching each other. The TDC model, on the other hand, shows a good agreement with the reference, with only a moderate underestimation of the splitting below 5Å.

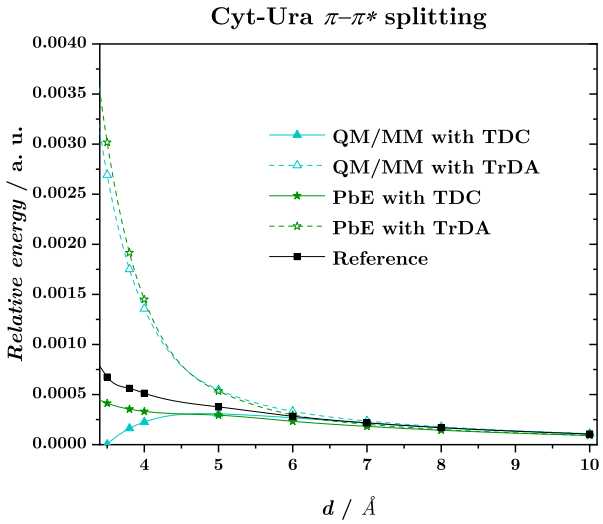

Figure 11: Dependence of the difference (in atomic units) of the $\pi - \pi^*$ energy levels in the *Cyt-Ura* complex on the intermolecular distance, as predicted by various coupling schemes. The reference curve is defined relative to the energy gap at infinite separation.

A very different picture is seen for the 1st and 2nd pairs Rydberg states of the $(Pyr)_2$ dimer on Fig. 12. (Note that, obviously, the aug-cc-pVDZ basis was used here.) The transition dipole approximation completely fails for the 1st pair, as it predicts a zero coupling for these dark excited states. The refined evaluation of the Coulombic coupling by the TDC model provides a reasonable result for the first pair of states until 6 Å, but severely underestimates the reference curve at smaller separations, indicating the necessity of the exchange (Dexter) terms for a reasonable prediction of the coupling. The monomer state of the 2nd Rydberg pair possesses a non-zero transition dipole moment, hence the splitting could in principle be approximated from this scheme. However, it is larger than its TDC counterpart, similarly to the n-R excitation of $(CH_2O)_2$ discussed above. This insinuates the conclusion that for these type of states the transition dipole approximation overestimates the Coulombic (Frenkel) coupling between the two monomers. The resulting splitting is, however, still much smaller than that of the reference curves, showing that this overestimation is far

from canceling the error from the absence of the Dexter terms in our methods. These findings make clear that both models investigated here are inappropriate for modeling couplings between Rydberg type excited states.

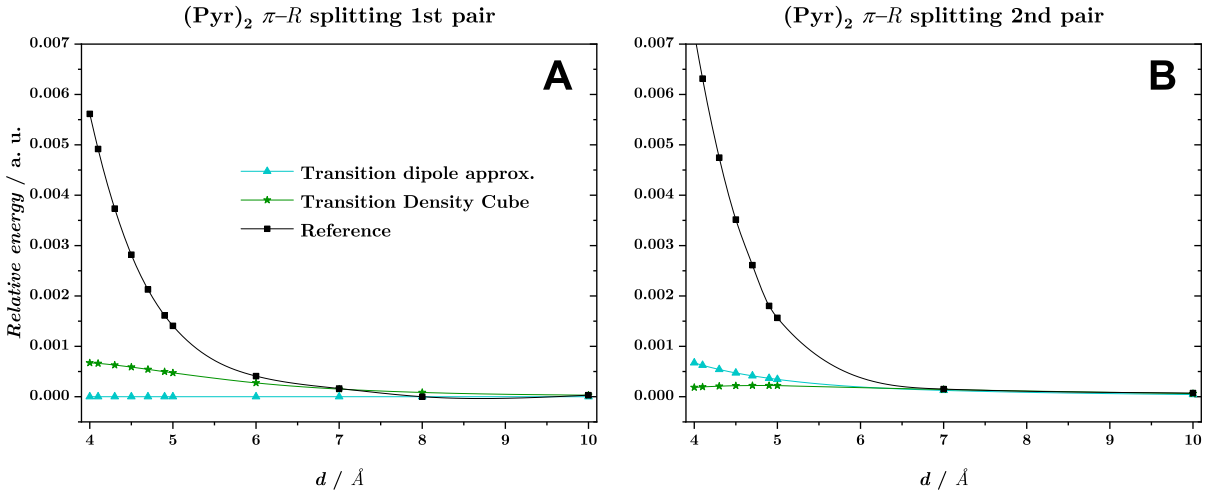

Figure 12: Distance dependence of the splitting (in atomic units) of the first (Panel A) and second (Panel B) Rydberg energy levels in the $(Pyr)_2$ complex, as predicted by various coupling schemes at the CCSD / aug-cc-pVDZ level of theory.

5.4 Potential energy curves

The different interaction energy and interstate coupling models result in a variety of final predictions for excited state potential energy curves. Since the evaluation of the coupling terms in Section 5.3 demonstrated the clear superiority of the TDC scheme over the one based solely on transition dipoles, surfaces resulting from the combination of TDC couplings (calculated at the CCSD level of theory) and various interaction energy potentials will be discussed in this chapter to assess the final results. As in Section 5.1, the 4 - 10 Å range of intermolecular separation is investigated.

5.4.1 Valence excited states

The potential energy surfaces, relative to the sum of the respective monomer electronic energies, are shown for the valence type electronic states on Figs. 13 and 14. For the $(CH_2O)_2$ system (Figure 13) results obtained with the aug-cc-pVDZ basis set are shown, while for the other complexes (Figure 14) the cc-pVDZ basis set was used.

In the $\sigma - \pi^*$ state of $(CH_2O)_2$ the *QM/MM + EFP2* model is generally well suited to describe the PES of both states in the whole investigated range, with the same accuracy. Severe discrepancies are seen for the *PbE + EFP2* scheme, however: both states turn strongly repulsive at as early as 6Å separation. This behavior stems from the incorrect virtual orbital selection as described earlier in Section 5.2.

The $\pi - \pi^*$ state of the $(CH_2O)_2$ system is, as discussed in Section 5.1.2 strongly influenced by a crossing CT state below 6 Å if diffuse functions are present in the basis set. Since the models do not include charge transfer effects, neither method can be expected to provide a satisfactory result at short separations. This is, as seen on Figure 13 indeed the case here, the surfaces obtained from the *PbE + EFP2* model being so overly repulsive that only the lower one features a shallow intermolecular minimum around 5Å, while the upper one exhibits a strong, obviously incorrect repulsive character. *QM/MM + EFP2* also underestimates the attraction in both states, as seen already on the coupling-free surfaces of Figure 5. On the other hand, anywhere above 6 Å, both models agree very well with the reference. From the two the *QM/MM + EFP2* results are to some degree more accurate, but the difference is insignificant. This suggests that without the interacting CT states complicating the wave functions, both approaches could produce meaningful potential energy surfaces even at shorter distances.

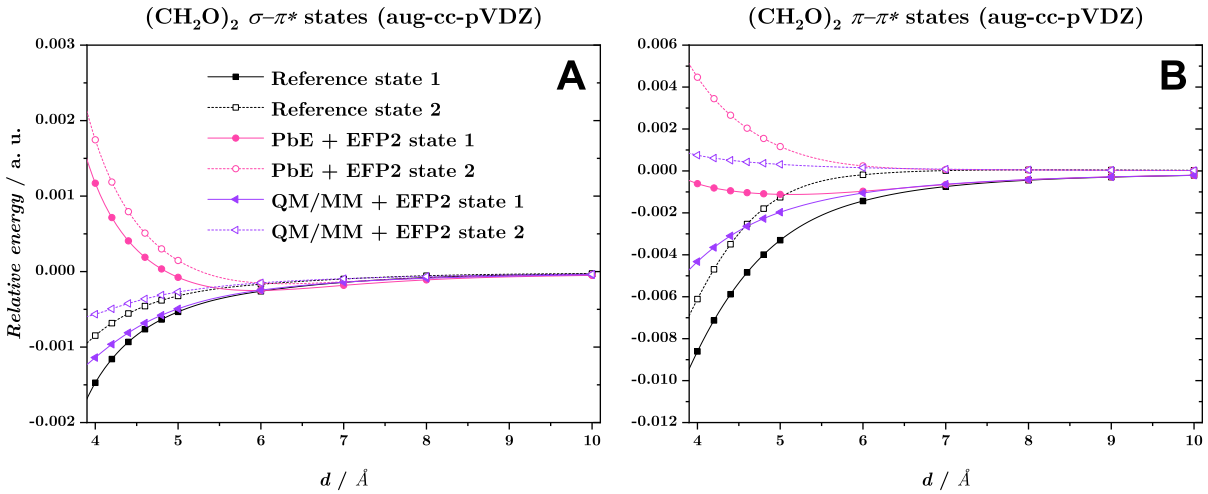

Figure 13: Potential energy curves (relative to the values at infinite separation, in atomic units) of the investigated $\sigma - \pi^*$ (Panel A) and $\pi - \pi^*$ (Panel B) valence excited states of the $(CH_2O)_2$ complex, calculated at with different models using CCSD/aug-cc-pVDZ method as their wave function component.

The effect of other states is mostly eliminated if no diffuse functions are present in the basis set, as in the results presented on Figure 14. In the investigated $\pi - \pi^*$ state pair of *(Pyr)*$_2$ (Panel A) both the *PbE + EFP2* and the *QM/MM + EFP2* models produce a nearly perfect agreement with the reference until 6 Å. Below this point, the potential energies predicted by *PbE + EFP2* start to become somewhat too low and show an unphysical minimum between 4 and 6 Å . This turn-back is likely caused by the truncation of the virtual orbital space in the PbE excitation energy calculations, artificially eliminating certain excitations of the dimer wave function. The *QM/MM + EFP2* results are, in principle, correct until 5 Å, where both curves start to underestimate the reference. While the error still remains moderate for the rapidly decreasing lower state, the upper one is significantly overstabilized and fails to reproduce the repulsive nature of the reference surface.

The approximate models are generally well suited for the *(Ura)*$_2$ and *(Cyt)*$_2$ $\pi - \pi^*$ dimer states (Panels B and C of Fig. 14). The shape of the potential energy curves follow those of the reference, with a modest underestimation by both methods. The *QM/MM + EFP2* approach turns out slightly more accurate until 4.5 Å, but the difference to the embedding results are negligible anywhere over 5 Å. The deviation from the reference is more significant in the *(Cyt)*$_2$ case, both in relative and nominal terms. Based on the findings discussed in Section 5.3 it is clear that the deviation is mostly caused by the error of the interaction energy modeling, as the interstate couplings are almost perfectly reproduced in the investigated region.

The potential energies of the same electronic states of the *Cyt-Ura* complex are shown on Panel D of Figure 14. (Note that the zero point of the energy scale is defined as the mean of the monomer excited state energies.) Here the *PbE + EFP2* potentials show very good agreement with the reference in the entire region. The *QM/MM + EFP2* result is also correct, almost indistinguishable from the *PbE + EFP2* one above 4.5 Å, while a somewhat stronger underestimation of the reference energies is produced below this point. Generally, the relative agreement with the reference is better in this non-symmetric complex than expected from the results seen for the *(Cyt)*$_2$ and *(Ura)*$_2$ homodimers. This accentuates the complexity of the excited state interactions in stacked π-complexes.

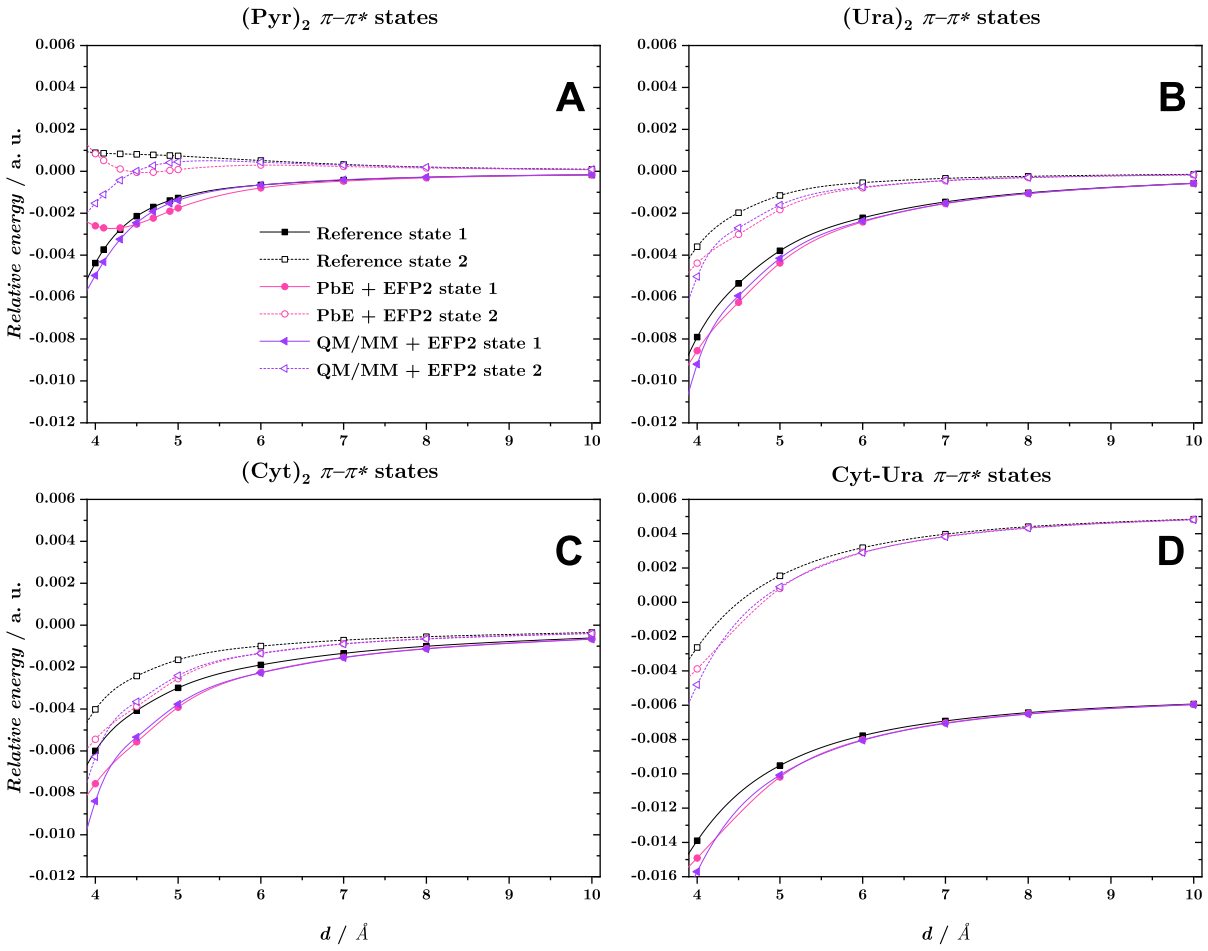

Figure 14: Potential energy curves (relative to the mean total energies at infinite separation, in atomic units) of the investigated $\pi - \pi^*$ excited states in the *(Pyr)*$_2$ (Panel A), *(Cyt)*$_2$ (Panel B), *(Ura)*$_2$ (Panel C) and *Cyt-Ura* (Panel D) complexes, calculated with different models using CCSD/cc-pVDZ method as their wave function component.

5.4.2 Rydberg states

Rydberg type electronic states, as seen above in Sections 5.1.2 and 5.3, are big challenge for the interaction modeling. The potential energy curves, shown on Fig. 15 for the *(Pyr)*$_2$ and the *(CH*$_2$*O)*$_2$ dimers, reflect many features already seen on the interaction energies. The *PbE + EFP2* model produces nonphysical, repulsive surfaces for both interacting states in all cases. This makes it clear that this approach in its present form is not suited for Rydberg type excited states. The *QM/MM + EFP2* energies, on the other hand, follow the reference with good parallelism in the whole range. However, due to the improper prediction of the coupling for Rydberg states (see Section 5.3), the splitting of the states is severely

underestimated. The resulting surfaces are thus also incorrect from the very point where the interaction becomes significant - as early as 6Å separation in these cases. This makes clear that for a qualitatively correct recovery of the Rydberg surfaces in the *QM/MM + EFP2* framework, the appropriate modeling of the coupling is also indispensable.

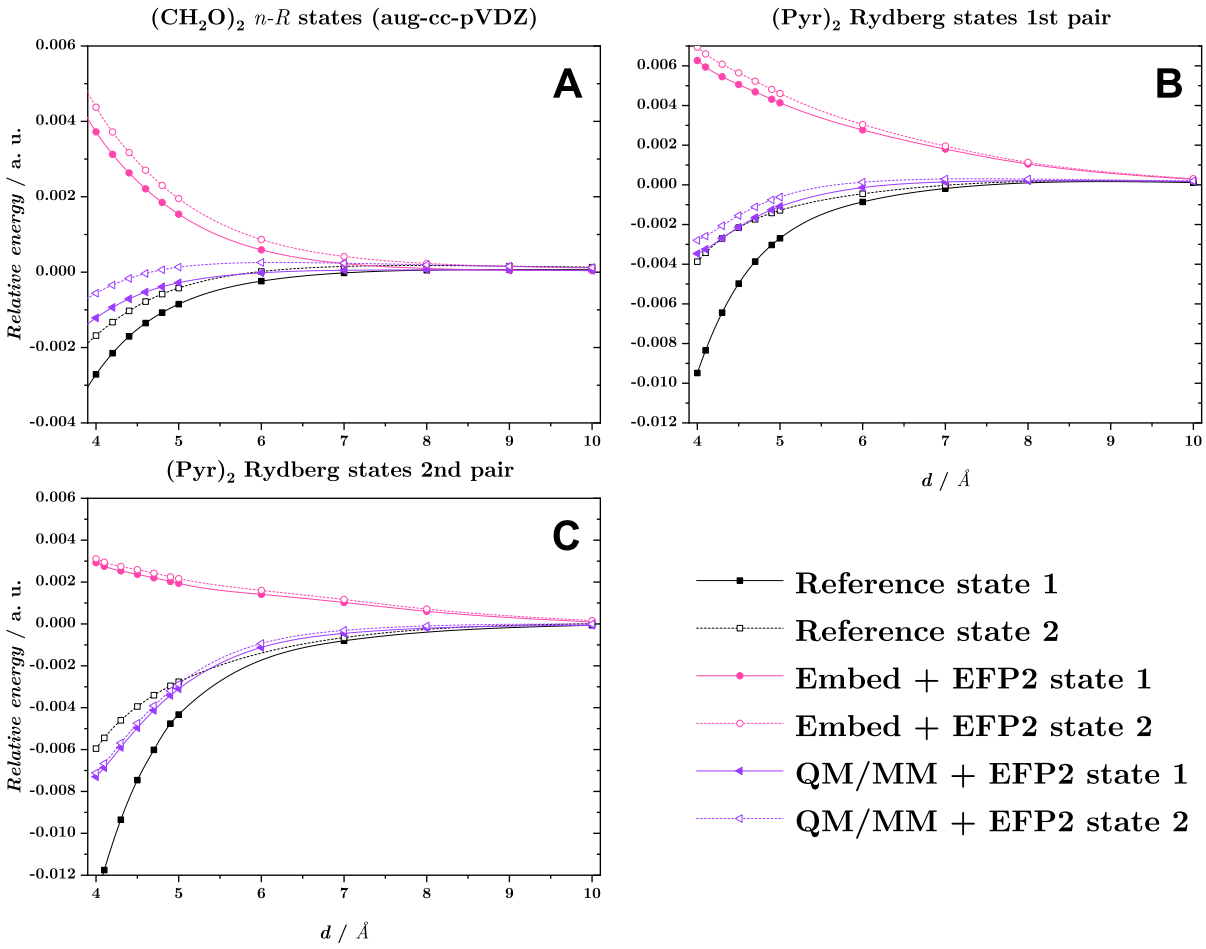

Figure 15: Potential energy curves (relative to the values at infinite separation, in atomic units) of the investigated Rydberg excited states in the $(CH_2O)_2$ (Panel A) and $(Pyr)_2$ (Panels B and D) dimers, evaluated with different models using CCSD/aug-cc-pVDZ method as their wave function component.

6 Conclusions

Various approximate methods for the description of non-covalent intermolecular interactions in excited state were defined and tested on stacked complexes of pyrrole, uracil, cytosine and formaldehyde. The conclusions can be summarized as:

- as of the electronic structure calculations, QM/MM with CHELPG point charges of the environment is a good choice, so is projector based embedding (PbE) if no diffuse functions are in the basis. However, PbE, due to the use of localized and truncated virtual space, fails with diffuse basis sets, giving incorrectly repulsive potential energy curves.
- correction potential to include van-der-Waals-type interactions (dispersion and Pauli repulsion) is necessary, and EFP2 provides a convenient and accurate way to do so.
- excitonic interaction can provide a qualitative description of the Frenkel splitting.
- excitonic coupling can be accurately calculated at TDC scheme.
- Dexter coupling becomes important below 4 Å.
- potential energy curves are reasonably accurate with the methodology presented here above 4 Å, i.e. from the minimum towards dissociation.

Acknowledgements

This work has been supported by the National Research, Innovation and Development Fund (NKFIA) of Hungary Grant No. 124018. We thank Prof. Mihály Kállay and his coworkers for assisting the embedding calculations and making adjustments in the MRCC code. JFS thanks for Fulbright fellowship.

Data availability statement

The data that supports the findings of this study are available in the supplementary material of this article.

Conflict of interest

The authors declare no conflict of interest.

Supporting Information

Structures of the monomers and complexes evaluated in this study, the impact of "exact" ground state van-der-Waals terms on the homodimer excited state surfaces, comparison of the CCSD and CC2 excited state splittings in the $(CH_2O)_2$ system and excited state total energy curves of $(CH_2O)_2$ are presented here.

References

1. Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew Chem Int Edit* **2009**, *48*, 1198–1229.

2. Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chemical Reviews* **2015**, *115*, 5678–5796.

3. Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. I. A Simple, Exact Density-Functional-Theory Embedding Scheme. *Journal of Chemical Theory and Computation* **2012**, *8*, 2564–2568, PMID: 22904692.

4. Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *The Journal of Physical Chemistry* **1993**, *97*, 8050–8053.

5. Wesołowski, T. A. Embedding a multideterminantal wave function in an orbital-free environment. *Physical Review A* **2008**, *77*, 012504.

6. Crawford, T. D.; King, R. A. Locally correlated equation-of-motion coupled cluster theory for the excited states of large molecules. *Chem. Phys. Lett.* **2002**, *366*, 611.

7. Korona, T.; Werner, H.-J. Local treatment of electron excitations in the EOM-CCSD method. *J. Chem. Phys.* **2003**, *118*, 3006.

8. Kats, D.; Korona, T.; Schütz, M. Local CC2 electronic excitation energies for large molecules with density fitting. *J. Chem. Phys.* **2006**, *125*, 104106.

9. Ma, Q.; Werner, H.-J. Explicitly correlated local coupled-cluster methods using pair natural orbitals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1371.

10. Dutta, A. K.; Saitow, M.; Riplinger, C.; Neese, F.; Izsák, R. A nearlinear scaling equation of motion coupled cluster method for ionized states. *J. Chem. Phys.* **2018**, *148*, 244101.

11. Mester, D.; Nagy, P. R.; Kállay, M. Reduced-cost second-order algebraic-diagrammatic construction method for excitation energies and transition moments. *J. Chem. Phys.* **2018**, *148*, 094111.

12. Mester, D.; Nagy, P. R.; Kállay, M. Reduced-scaling correlation methods for the excited states of large molecules: Implementation and benchmarks for the second-order algebraic-diagrammatic construction approach. *J. Chem. Theory Comput.* **2019**, *15*, 6111.

13. Benda, Z.; Szalay, P. G. Characterization of the Excited States of DNA Building Blocks: a Coupled Cluster Computational Study. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23596–23606.

14. Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2011**, *112*, 632–672.

15. Szalay, P. G.; Watson Jr., T. J.; Perera, A.; Lotrich, V. F.; Fogarasi, G.; Bartlett, R. J. Benchmark Studies on the Building Blocks of DNA: 2. Effect of Biological Environment on the Electronic Excitation Spectrum of Nucleobases. **2012**, *116*, 8851–8860.

16. Barcza, B.; Szirmai, Á. B.; Szántó, K. J.; Tajti, A.; Szalay, P. G. Comparison of approximate intermolecular potentials for ab initio fragment calculations on medium sized N-heterocycles. *Journal of Computational Chemistry* **2022**, *43*, 1079–1093.

17. Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. In *Chapter 10 The Effective Fragment Potential: A General Method for Predicting Intermolecular Interactions*; Spellmeyer, D., Wheeler, R., Eds.; Annual Reports in Computational Chemistry; Elsevier, 2007; Vol. 3; pp 177–193.

18. Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *Journal of Chemical Theory and Computation* **2005**, *1*, 1128–1132, PMID: 26631656.

19. Jensen, J. H.; Gordon, M. S. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application to the effective fragment potential method. *The Journal of Chemical Physics* **1998**, *108*, 4772–4782.

20. Frenkel, J. A. On the transformation of light into heat in solids. II. *Phys. Rev.* **1931**, *37*, 1276.

21. Davydov, A. S. *Soviet Phys.-Usp* **1964**, *530*, 145–180.

22. Morrison, A. F.; You, Z. Q.; Herbert, J. M. Ab initio implementation of the Frenkel–Davydov exciton model: a naturally parallelizable approach to computing collective excitations in crystals and aggregates. *Journal of Chemical Theory and Computation* **2014**, *10*, 5366 – 5376.

23. Sisto, A.; Stross, C.; van der Kamp, M. W.; O'Connor, M.; McIntosh-Smith, S.; Johnson, G. T.; Hohenstein, E. G.; Manby, F. R.; Glowacki, D. R.; Martinez, T. J. Atomistic non-adiabatic dynamics of the LH2 complex with a GPU-accelerated ab initio exciton model. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14924–14936.

24. Amadei, A.; D'Alessandro, M.; D'Abramo, M.; Aschi, M. Theoretical characterization of electronic states in interacting chemical systems. *The Journal of Chemical Physics* **2009**, *130*, 084109.

25. Hégely, B.; Szirmai, Á. B.; Mester, D.; Tajti, A.; Szalay, P. G.; Kállay, M. Performance of Multilevel Methods for Excited States. *The Journal of Physical Chemistry A* **2022**, *126*, 6548–6557, PMID: 36095318.

26. Ge, Q.; Head-Gordon, M. Energy decomposition analysis for excimers using absolutely localized molecular orbitals within time-dependent density functional theory and configuration interaction .... *Journal of Chemical Theory and Computation* **2018**, *14*, 5156–5168.

27. Ponder, J. W.; Case, D. A. *Protein Simulations*; Advances in Protein Chemistry; Academic Press: San Diego, CA, 2003; Vol. 66; pp 27–85.

28. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.

29. Brooks, B. R.; Brooks III, C. L.; Mackerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.;

Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry* **2009**, *30*, 1545–1614.

30. Stone, A. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chemical Physics Letters* **1981**, *83*, 233–239.

31. Stone, A. J. *The Theory of Intermolecular Forces*; OUP Oxford: Oxford, 2013.

32. Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *The Journal of Chemical Physics* **1996**, *105*, 1968–1986.

33. Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate First Principles Model Potentials for Intermolecular Interactions. *Annual Review of Physical Chemistry* **2013**, *64*, 553–578.

34. Sattasathuchana, T.; Xu, P.; Gordon, M. S. An Accurate Quantum-Based Approach to Explicit Solvent Effects: Interfacing the General Effective Fragment Potential Method with Ab Initio Electronic Structure Theory. *The Journal of Physical Chemistry A* **2019**, *123*, 8460–8475.

35. Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry* **1990**, *11*, 361–373.

36. Huzinaga, S.; Cantu, A. A. Theory of Separability of Many-Electron Systems. *The Journal of Chemical Physics* **1971**, *55*, 5543–5549.

37. Hégely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact density functional and wave function embedding schemes based on orbital localization. *The Journal of Chemical Physics* **2016**, *145*, 064107.

38. Khait, Y. G.; Hoffmann, M. R. In *Annual Reports in Computational Chemistry*; Wheeler, R. A., Ed.; Annual Reports in Computational Chemistry; Elsevier, 2012; Vol. 8; pp 53–70.

39. Bennie, S. J.; Curchod, B. F. E.; Manby, F. R.; Glowacki, D. R. Pushing the Limits of EOM-CCSD with Projector-Based Embedding for Excitation Energies. *The Journal of Physical Chemistry Letters* **2017**, *8*, 5559–5565, PMID: 29076727.

40. Parravicini, V.; Jagau, T.-C. Embedded equation-of-motion coupled-cluster theory for electronic excitation, ionisation, electron attachment, and electronic resonances. *Molecular Physics* **2021**, *119*, e1943029.

41. Claudino, D.; Mayhall, N. J. Simple and Efficient Truncation of Virtual Spaces in Embedded Wave Functions via Concentric Localization. *The Journal Chemical Theory and Computations* **2019**, *15*, 6085 – 6096.

42. Hégely, B.; Szirmai, Á. B.; Mester, D.; Tajti, A.; Szalay, P. G.; Kállay, M. Performance of Multilevel Methods for Excited States. *The Journal of Physical Chemistry A* **2022**, *126*, 6548–6557, PMID: 36095318.

43. Claudino, D.; Mayhall, N. J. Automatic Partition of Orbital Spaces Based on Singular Value Decomposition in the Context of Embedding Theories. *Journal of Chemical Theory and Computation* **2019**, *15*, 1053–1064.

44. Jensen, J. H.; Gordon, M. S. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. *Molecular Physics* **1996**, *89*, 1313–1325.

45. Jensen, J. H. Intermolecular exchange-induction and charge transfer: Derivation of approximate formulas using nonorthogonal localized molecular orbitals. *The Journal of Chemical Physics* **2001**, *114*, 8775–8783.

46. Adamovic, I.; *, M. S. G. Dynamic polarizability, dispersion coefficient C6 and dispersion energy in the effective fragment potential method. *Molecular Physics* **2005**, *103*, 379–387.

47. Smith, Q. A.; Ruedenberg, K.; Gordon, M. S.; Slipchenko, L. V. The dispersion interaction between quantum mechanics and effective fragment potential molecules. *The Journal of Chemical Physics* **2012**, *136*, 244107.

48. Smith, T.; Slipchenko, L. V.; Gordon*, M. S. Modeling $\pi - \pi$ Interactions with the Effective Fragment Potential Method: The Benzene Dimer and Substituents. *The Journal of Physical Chemistry A* **2008**, *112*, 5286–5294, PMID: 18476681.

49. Sattasathuchana, T.; Xu, P.; Gordon, M. S. An Accurate Quantum-Based Approach to Explicit Solvent Effects: Interfacing the General Effective Fragment Potential Method with Ab Initio Electronic Structure Theory. *The Journal of Physical Chemistry A* **2019**, *123*, 8460–8475, PMID: 31365250.

50. Rojas, C. I. V.; Slipchenko, L. V. Exchange Repulsion in Quantum Mechanical/Effective Fragment Potential Excitation Energies: Beyond Polarizable Embedding. *Journal of Physical Chemistry Letters* **2020**, *16*, 6408 – 6417.

51. Rojas, C. I. V.; Fine, J.; Slipchenko, L. V. Exchange-repulsion energy in QM/EFP. *Journal of Chemical Physics* **2018**, *149*, 094103.

52. Smith, Q. A.; Ruedenberg, K.; Gordon, M. S.; Slipchenko, L. V. The dispersion interaction between quantum mechanics and effective fragment potential molecules. *Journal of Chemical Physics* **2012**, *136*, 244107.

53. Hapka, M.; Przybytek, M.; Pernal, K. Symmetry-Adapted Perturbation Theory Based on Multiconfigurational Wave Function Description of Monomers. *Journal of Chemical Theory and Computation* **2021**, *17*, 5538–5555, PMID: 34517707.

54. Jangrouei, M. R.; Krzemińska, A.; Hapka, M.; Pastorczak, E.; Pernal, K. Dispersion Interactions in Exciton-Localized States. Theory and Applications to $\pi - \pi^*$ and $n - \pi^*$ Excited States. *Journal of Chemical Theory and Computation* **2022**, *18*, 3497–3511, PMID: 35587598.

55. Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chemical Reviews* **1994**, *94*, 1887–1930.

56. Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Computational Molecular Science* **2012**, *2*, 254–272.

57. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der physik* **1948**, *437*, 55 – 75.

58. Chang, J. C. Monopole effects on electronic excitation interactions between large molecules. I. Application to energy transfer in chlorophylls. *The Journal of Chemical Physics* **1977**, *67*, 3901–3909.

59. Zanetti-Polzi, L.; Galdo, S. D.; Daidone, I.; D'Abramo, M.; Barone, V.; Aschi, M.; Amadei, A. Extending the perturbed matrix method beyond the dipolar approximation: comparison of different levels of theory. *Physical Chemistry Chemical Physics* **2018**, *20*, 24369 – 24378.

60. Fückel, B.; Köhn, A.; Harding, M. E.; Diezemann, G.; Hinze, G.; Basché, T.; Gauss, J. Theoretical investigation of electronic excitation energy transfer in bichromophoric assemblies. *The Journal of Chemical Physics* **2008**, *128*, 074505.

61. Krueger, B. P.; Scholes, G. D.; Fleming, G. R. Calculation of Couplings and Energy-Transfer Pathways between the Pigments of LH2 by the ab Initio Transition Density Cube Method. *The Journal of Physical Chemistry B* **1998**, *102*, 5378–5386.

62. Krueger, B. P. The Transition Density Cube method. `http://www.chem.hope.edu/\string~krieg/TDC/TDC\_home.htm`, 1998; [Online; accessed Dec-2022].

63. Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-cluster techniques for computational chemistry: The CFOUR program package. *The Journal of Chemical Physics* **2020**, *152*, 214108.

64. Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package. With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, Y.J. Bomble, O. Christiansen, F. Engel, R. Faber, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, W.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

65. Li, X.; Parrish, R. M.; liu, F.; Schumacher, S. I. L. K.; Martinez, T. J. An Ab Initio Exciton Model Including Charge-Transfer Excited States. *Journal of Chemical Theory and Computation* **2017**, *13*, 3493 – 3504.

66. Hsu, C.-P.; Fleming, G. R.; Head-Gordon, M.; Head-Gordon, T. Excitation energy transfer in condensed media. *Journal of Chemical Physics* **2001**, *114*, 3065.

67. Dexter, D. L. A Theory of Sensitized Luminescence in Solids. *The Journal of Chemical Physics* **1953**, *21*, 836–850.

68. Förster, T. *Delocalized excitation and excitation transfer*; Modern Quantum Chemistry: Istanbul Lectures. Part III, Action of Light and Organic Crystals. Ed. O. Sinanoglu; Academic: New York, 1965.

69. Scholes, G. D. LONG-RANGE RESONANCE ENERGY TRANSFER IN MOLECULAR SYSTEMS. *Physical Chemistry* **2003**, *54*, 57–87.

70. Schreiber, M.; Silva, M. R. J.; Sauer, S. P. A.; Thiel, W. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *The Journal of Chemical Physics* **2008**, *128*.

71. Christiansen, O.; Koch, H.; Jørgensen, P. The Second-order Approximate Coupled-Cluster Singles and Doubles Model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

72. Schirmer, J.; Trofimov, A. B. Intermediate state representation approach to physical properties of electronically excited molecules. *J. Chem. Phys.* **2004**, *120*, 11449–11464.

73. Hellweg, A.; Grün, S. A.; Hättig, C. Benchmarking the performance of spin-component scaled CC2 in ground and electronically excited states. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4119–4127.

74. Winter, N. O. C.; Hättig, C. Scaled opposite-spin CC2 for ground and excited states with fourth order scaling computational costs. *J. Chem. Phys.* **2011**, *134*, 184101.

75. Stanton, J. F.; Bartlett, R. J. The Equation of Motion Coupled-Cluster Method - A Systematic Biorthogonal Approach to Molecular-Excitation Energies, Transition-Probabilities, and Excited-State Properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.

76. Comeau, D. C.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method - Applications to Open-Shell and Closed-Shell Reference States. *Chem. Phys. Lett.* **1993**, *207*, 414–423.

77. Sinnokrot, M. O.; Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: sandwich, T-shaped, and parallel-displaced configurations. *The Journal of Physical Chemistry A* **2004**, *108*, 10200 – 10207.

78. Bartlett, R. J.; Purvis III, G. D. Molecular Applications of Coupled Cluster and Many-Body Perturbation Methods. *Phys. Scr.* **1980**, *21*, 255–265.

79. Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

80. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron-Affinities of the 1st-Row Atoms Revisited - Systematic Basis-Sets and Wave-Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

81. Boys, S.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **1970**, *19*, 553–566.

82. Cook, D. B.; Sordo, J. A.; Sordo, T. L. Some comments on the counterpoise correction for the basis set superposition error at the correlated level. *International Journal of Quantum Chemistry* **1993**, *48*, 375–384.

83. Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *Journal of Chemical Theory and Computation* **2014**, *10*, 49–57.

84. Kristensen, K.; Ettenhuber, P.; Eriksen, J. J.; Jensen, F.; Jørgensen, P. The same number of optimized parameters scheme for determining intermolecular interaction energies. *The Journal of Chemical Physics* **2015**, *142*, 114116.

85. Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. Aa.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Sałek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269–284.

86. Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.;

Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S. S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruitt, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Galvez Vallejo, J. L.; Westheimer, B.; Wloch, M.; Xu, P.; Zahariev, F.; Gordon, M. S. Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics* **2020**, *152*, 154102.

87. Kállay, M.; Nagy, P. R.; Rolik, Z.; Mester, D.; Samu, G.; Csontos, J.; Csóka, J.; Szabó, B. P.; Gyevi-Nagy, L.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Petrov, K.; Farkas, M.; Mezei, P. D.; Hégely., B. Mrcc, a quantum chemical program suite. See also Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, and M. Kállay, J. Chem. Phys. 139, 094105 (2013), as well as: `www.mrcc.hu`.

88. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

89. Slipchenko, L. V.; Gordon, M. S. Electrostatic energy in the effective fragment potential method: Theory and application to benzene dimer. *Journal of Computational Chemistry* **2007**, *28*, 276–291.