JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Letter

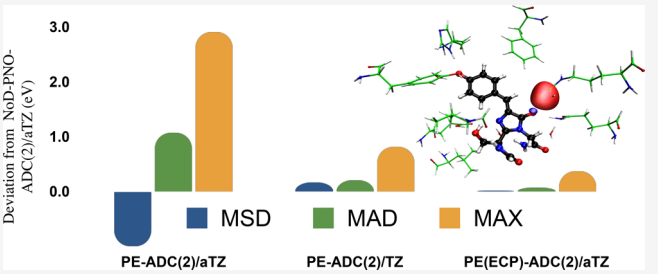
Avoiding Electron Spill-Out in QM/MM Calculations on Excited States with Simple Pseudopotentials

Alireza Marefat Khah,* Peter Reinholdt,* Jógvan Magnus Haugaard Olsen, Jacob Kongsted, and Christof Hättig

Cite This: https://dx.doi.org/10.1021/acs.jctc.9b01162 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: QM/MM calculations of electronic excitations with diffuse basis sets often have large errors due to spill-out of electrons from the quantum subsystem. The Pauli repulsion of the electrons by the environment has to be included to avoid this. We propose transferable atomic all-electron pseudopotentials that can readily be combined with most MM force fields to avoid electron spill-out. QM/MM excitation energies computed with time-dependent Hartree−Fock and the algebraic diagrammatic construction through second-order are benchmarked against supermolecular calculations to validate these new pseudopotentials. The QM/MM calculations with pseudopotentials give accurate results that are stable with augmentation of the basis set with diffuse functions. We show that the largest contribution to residual deviations from full QM calculations is caused by the missing London dispersion interaction.



Downloaded via BOCHUM LIBRARIES on February 28, 2020 at 08:11:58 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

A standard tool for multiscale simulations of molecular and spectroscopic properties in complex environments is QM/MM methods.[1−5] Many quantum chemistry packages are equipped with a QM/MM module or have been interfaced with molecular mechanics programs. For a fast convergence with the size of the quantum mechanical (QM) region, it is advantageous if the molecular mechanics (MM) part is polarizable.[6,7] Different polarizable QM/MM schemes have been developed that account for the polarization of the MM region through, e.g., point dipole−dipole polarizabilities,[8−15] fluctuating charges,[11,16−19] or Drude oscillators.[20−22] Up to today, polarizable QM/MM approaches have been combined with a wealth of *ab initio* methods, including (time-dependent) Hartree−Fock and density functional theory,[8,12,14,15,23] and some of the most widely used correlated wave function methods, e.g., coupled-cluster[24,25] (CC), the algebraic-diagrammatic construction[26,27] (ADC), multiconfigurational self-consistent field[28] (MCSCF), complete active space perturbation theory[29] (CASPT2), the second-order polarization propagator approach (SOPPA),[30] and the density matrix renormalization group[31] (DMRG) approach.

In additive QM/MM schemes,[2] the missing effects of Pauli repulsion and dispersion interactions due to electrons in the QM subsystem are one of the major deficiencies that require going beyond a purely classical electrostatic and mechanical description. In the following, we focus on additive polarizable QM/MM methods for which the energy can be written as the sum of the internal energies of the QM and MM subsystems and an interaction term:

$$E^{\text{full}} = E^{\text{QM}} + E^{\text{MM}} + E^{\text{QM/MM}} \quad (1)$$

The latter consists usually of a contribution from the permanent electrostatic moments of the particles in the MM region, $E_{\text{es}}^{\text{QM/MM}}$, a polarization contribution, $E_{\text{pol}}^{\text{QM/MM}}$, and contributions from nonclassical or van der Waals interactions, $E_{\text{vdW}}^{\text{QM/MM}}$:

$$E^{\text{QM/MM}} = E_{\text{es}}^{\text{QM/MM}} + E_{\text{pol}}^{\text{QM/MM}} + E_{\text{vdW}}^{\text{QM/MM}} \quad (2)$$

$E_{\text{vdW}}^{\text{QM/MM}}$ is usually approximated by purely mechanical potentials acting only on the nuclei, e.g., pairwise 6−12 Lennard-Jones potentials. The missing effect of the Pauli repulsion on the electronic wave function leads to the so-called electron spill-out (ESO) problem. This problem manifests itself as an unphysical leaking of the electron density into the MM environment. The ESO diminishes the accuracy of the calculations and in some cases causes completely unphysical results, in particular, when diffuse basis functions are required to capture the physical character of an excited state or to achieve the required accuracy. Several remedies have been proposed to avoid it. The modified Coulombic interaction approach[32] is often used in combination with standard MM force fields. It has also been suggested that replacing point

Received: November 22, 2019
Published: February 7, 2020

© XXXX American Chemical Society

multipole moments with a Gaussian-smeared charge density can cure the issue.[33] This can, in effect, lift the electrostatic singularities that occur at MM sites with point multipoles and provide a more accurate description of the continuous charge distribution in the short-range including charge penetration effects. However, these methods do not capture the effects of the Pauli exchange repulsion on the electronic wave function. In recent years, several extensions of the standard polarizable QM/MM approach have been proposed to substitute the purely mechanical description of $E_{\mathrm{vdW}}^{\mathrm{QM/MM}}$ by a density-dependent description of Pauli repulsion.[34−37] Some of these approaches, e.g., EFP2, are more accurate but also more involved and costly.[34,38] For some of them,[35] molecular gradients are not yet available, and some require preceding QM calculations for the MM subsystem or a reparametrization of the electrostatics,[36,37,39] which limit their applicability for large systems.

Therefore, there is a demand for techniques to avoid the ESO in QM/MM calculations, which satisfy the following criteria: **i**) low computational cost that does not scale steeply with the size of the MM region, **ii**) accurate enough to be used with small QM regions and high-level QM methods, **iii**) compatible with the calculation of molecular gradients and response properties, and **iv**) easy to set up and parametrize.

Using pseudopotential operators to model the effect of chemically inactive (core) electrons on the explicitly treated electrons is a well-known approach.[40] The method was first introduced by Hellmann[41,42] and Gombás[43] to model Pauli repulsion in metallic systems. Notably, for heavy elements, effective core potentials (ECPs) are frequently employed to reduce the computational costs and to include the most important relativistic effects on the valence electrons[44] in otherwise nonrelativistic calculations. ECPs are also widely used in embedded cluster calculations for defects in ionic crystals and on their surfaces to avoid ESO as for these systems, the MM subsystem contains only a few atom- or ion-types, and well-parametrized ECPs are available.[45] Parametrized ECPs have also been used in ground- and excited-state QM/MM calculations to model boundary atoms and functional groups at the QM/MM boundary.[46−49] In the current work, we investigate the possibility of using atomic pseudopotentials to avoid the electron spill-out problem in QM/MM calculations on electronically excited states.

For simplicity of the parametrization, we assume that all atoms in the MM region fulfill the octet rule, i.e., have no dangling bonds that might serve as electron acceptors, and that their contribution to the Pauli repulsion can be described by a superposition of transferable one-center pseudopotentials that mimic approximately the core and valence electrons. We use the same ansatz as for the usual ECPs[50] since for them the integrals are already available in most quantum chemistry packages:

$$\hat{V}_{\lambda}^{\mathrm{ECP}}(i) = \sum_{l=0}^{l_{\max}} \sum_{k} A_{l,k}^{\lambda} |\mathbf{r}_{\lambda} - \mathbf{r}_i|^{n_{l,k}^{\lambda}} \exp(-a_{l,k}^{\lambda}(\mathbf{r}_{\lambda} - \mathbf{r}_i)^2) \hat{P}_l^{\lambda}(i) \tag{3}$$

with

$$P_l^{\lambda}(i) = \sum_{m_l=-l}^{l} |\lambda l m_l(i)\rangle\langle\lambda l m_l(i)| \tag{4}$$

In eq 3, $P_l^{\lambda}(i)$ projects the orbital occupied by electron i on the spherical harmonic function with angular momentum l around the position $\mathbf{r}_{\lambda}$ of the MM site λ. In the current work, $l_{\max}$ is set to the highest angular momentum occupied in the ground state configuration of the atom plus one, the exponents $n_{l,k}^{\lambda}$ are set to zero, and only one Gaussian function ($k = 0$) is used per angular momentum. For simplicity, we assume that the ECPs only depend on the number of atomic orbitals that are (partially) filled after fulfillment of the octet rule, i.e., we use for H the same ECP as for He, for B−F the same as for Ne, and for Al−Cl the same as for Ar. For the current work, we thus only optimized ECP parameters for the three rare gas atoms He, Ne, and Ar. The coefficients $A_{l,0}^{\lambda}$ and exponents $a_{l,0}^{\lambda}$ for He, Ne, and Ar were determined by minimizing the objective function

$$L = \sum_{l=0}^{l_{\max}} \sum_{a=1}^{2} (\epsilon_{l,a}^{\mathrm{ECP}} - \epsilon_{l,a}^{\mathrm{all\text{-}el}})^2 \tag{5}$$

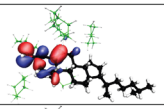
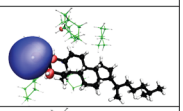
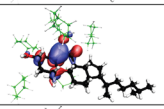
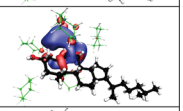
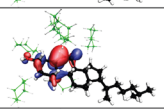
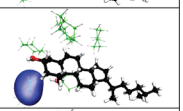
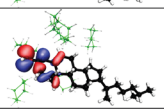
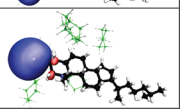
where $\epsilon_{l,a}^{\mathrm{ECP}}$ and $\epsilon_{l,a}^{\mathrm{all\text{-}el}}$ are the canonical Hartree−Fock orbital energies of the lowest unoccupied orbitals a with angular momentum l in, respectively, ECP and all-electron calculations on the atoms in a large basis set, see Tables S1−S3 in the Supporting Information. Adding such minimal ECPs to all MM sites increases the computational cost only moderately (see Table S4 in the Supporting Information), and since the ECP integrals are more short-ranged than the Coulombic charge and multipole interaction integrals, they will not affect the scaling of the overall cost with the sizes of the QM and MM regions. For the following, we focus on the polarizable embedding[10,12,25,26] (PE) variant of QM/MM, and, to distinguish between PE calculations with and without ECPs at the MM centers, we denote the ECP-augmented polarizable embedding as PE(ECP).

To assess the performance of PE(ECP), we compare below electronic excitation energies calculated within QM/MM setups with and without ECPs with full QM calculations at the time-dependent Hartree−Fock (TDHF) and the ADC(2) level. TDHF has been chosen in order to fully exclude correlation and thus dispersion effects. In this way, the deviations of PE-TDHF from full QM TDHF calculations can only come from the parametrization of the potential for the electrostatic interaction, including the polarization and (the lack of) Pauli repulsion. In the correlated wave function method ADC(2), the London dispersion interaction between the QM subsystem and the MM environment is, in addition to the Pauli repulsion, another source of discrepancy between the full QM and the QM/MM calculations. For the ADC(2) calculations we used the PNO-ADC(2) code[51,52] that is part of the development version[53] of TURBOMOLE. At this level of theory, we are able to separate the dispersion contribution from the supermolecular full QM calculations by decomposition of the correlation contributions to the excitation energies. The PNO-ADC(2) implementation is based on localized molecular orbitals (LMOs) for the occupied and pair natural orbitals for the virtual space which allows for splitting of the correlation contribution to the excitation energy originating from the double excitations, $T_2 = \sum_{aibj} t_{ij}^{ab} \tau_{ab}^{ij}$, according to the localization of the LMOs i and j into

- a major contribution where both LMOs are localized on the chromophore in the QM region,
- a smaller intermolecular contribution where one LMO is localized on the chromophore and one in the environment, which includes the dispersion interaction, and

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Full-QM NoD-PNO-ADC(2) (T_{PNO} = 9) Excitation Energy ΔE to the Second-Lowest Singlet Electronic Excited State S_2 of Cholesterol in Cyclohexane, Calculated with Mixed aTZ/DZ and TZ/DZ Basis Sets[a]

Level	S_2	ΔE (eV)	occ	vir
full-QM/aTZ	$n\pi \rightarrow$ O-Ryd.	6.40 (6.26)		
full-QM/TZ	$\pi \rightarrow$ C-Ryd.	7.54 (7.08)		
PE/aTZ	$\pi \rightarrow$ ESO	4.53		
PE(ECP)/aTZ	$n\pi \rightarrow$ O-Ryd.	6.41		

[a]The full-QM PNO-ADC(2) S_2 excitation energies are quoted in the parentheses. PE- and PE(ECP)-ADC(2)/aTZ excitation energies ΔE are calculated within the post-SCF linear response (LR) formalism. NTOs are visualized for isosurface values of ±0.0275 au. Carbon atoms of solvent molecules are shown in green.

- a minor contribution where both LMOs are localized in the environment.

By subtracting the second contributions from the reference full QM PNO-ADC(2) excitation energies (denoted below as NoD-PNO-ADC(2) results), we can isolate the electrostatic and Pauli repulsion contributions, which are the subject of the current study. In all the PNO-ADC(2) calculations, we used a tight PNO selection threshold ($T_{PNO} = 10^{-9}$) to ensure that the errors introduced by the local approximations are negligible.[51]

Here, we used basis sets with and without diffuse basis functions since the amount of ESO varies a great deal with the inclusion of diffuse functions.[26,54] For excited states, however, diffuse basis functions are usually needed to achieve reasonable accuracy. This requirement may be alleviated in full QM calculations for densely packed systems where the diffuseness can be described by basis functions of the surrounding atoms. However, this is not the case for QM/MM calculations where the wave function of the chromophore is solely described with the basis functions of the QM subsystem. The basis set requirements of QM/MM calculations are thus similar to vacuum calculations where diffuse basis functions are important even for low-lying valence excitations if the accuracy of correlated wave functions methods (<0.3 eV) should be exploited. Thiel and co-workers[55] showed that in the gas phase, CC2 excitation energies that are calculated with the triple-ζ basis TZVP, deviate already for valence states by about 0.2 eV from more accurate results in the aug-cc-pVTZ basis. Therefore, it is desirable to be able to use these basis sets also in QM/MM calculations.

As we show below, ESO is not limited to Rydberg or high-lying excitations but also occurs for local valence excitations ($\pi\pi^*$, $n\pi^*$, etc.) and in the lowest electronically excited states, vide infra. The severeness of ESO depends on, apart from the diffuseness of the basis set, the polarity of the QM system and the MM environment. An example of ESO for local excitations is the lowest $\pi\pi^*$ electronic excitation of the green fluorescent protein (GFP), see the SI.

For Rydberg and mixed-valence Rydberg states, diffuse functions are already unavoidable for a qualitatively correct description. Even though the Rydberg character is partially quenched in solution, the corresponding states remain diffuse in character and require diffuse functions for a physically correct description. In ref 56, some of us have shown that the standard PE scheme without accounting for Pauli repulsion is not sufficient for the electronically excited states of cholesterol in cyclohexane solution. The second-lowest excited state, S_2, highlights the problem. Table 1 summarizes some results from full QM and QM/MM calculations at the ADC(2) level for the second-lowest excited state, S_2, of cholesterol in a cyclohexane cluster. This example highlights the basis set requirements in describing electron excitations and the importance of avoiding ESO to MM environments.

In the full-QM PNO-ADC(2) calculation with the aug-cc-pVTZ (aTZ) basis, the vertical excitation energy of cholesterol within the cyclohexane cluster is 6.40 eV after subtracting the dispersion contribution δE_{Disp} which in this case would cause a red-shift of 0.138 eV. A similar PNO-ADC(2) calculation in the cc-pVTZ (TZ) basis, i.e., leaving out the diffuse functions, gives as S_2 a different excited state with a much higher excitation energy of 7.54 eV that deviates from the aTZ result by 1.1 eV. The PE-ADC(2) QM/MM calculation in the aTZ basis set is, however, strongly affected by ESO which leads to the localization of the virtual natural transition orbital (NTO) on a subunit of the MM environment and an unphysically small excitation energy of 4.5 eV within the post-SCF linear response (LR) reaction field scheme.[25] In contrast to the standard PE technique, PE(ECP)-ADC(2) with ECPs on all MM sites can reproduce in the aTZ basis the NoD-PNO-ADC(2) result very accurately with a remaining small deviation of only 0.01 eV.

To evaluate the transferability and accuracy of the constructed all-electron ECPs for the embedding, we chose for (PE-)ADC(2) a benchmark set that comprises 9 chromophores and in total 15 electronic excitations, 10 local ($\pi\pi^*$, $n\pi^*$), and 5 nonlocal transitions with charge-transfer (CT) or Rydberg character (see the NTOs given in the SI Tables S5−S13). For (PE-)TDHF we included in addition 11 further excitation energies (cmp. Table S15−S16) so that, in total, 26 states are considered at this level. The chromophores are embedded in chemically diverse environments (see Figure 1), including water, organic solvents (cyclohexane and

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

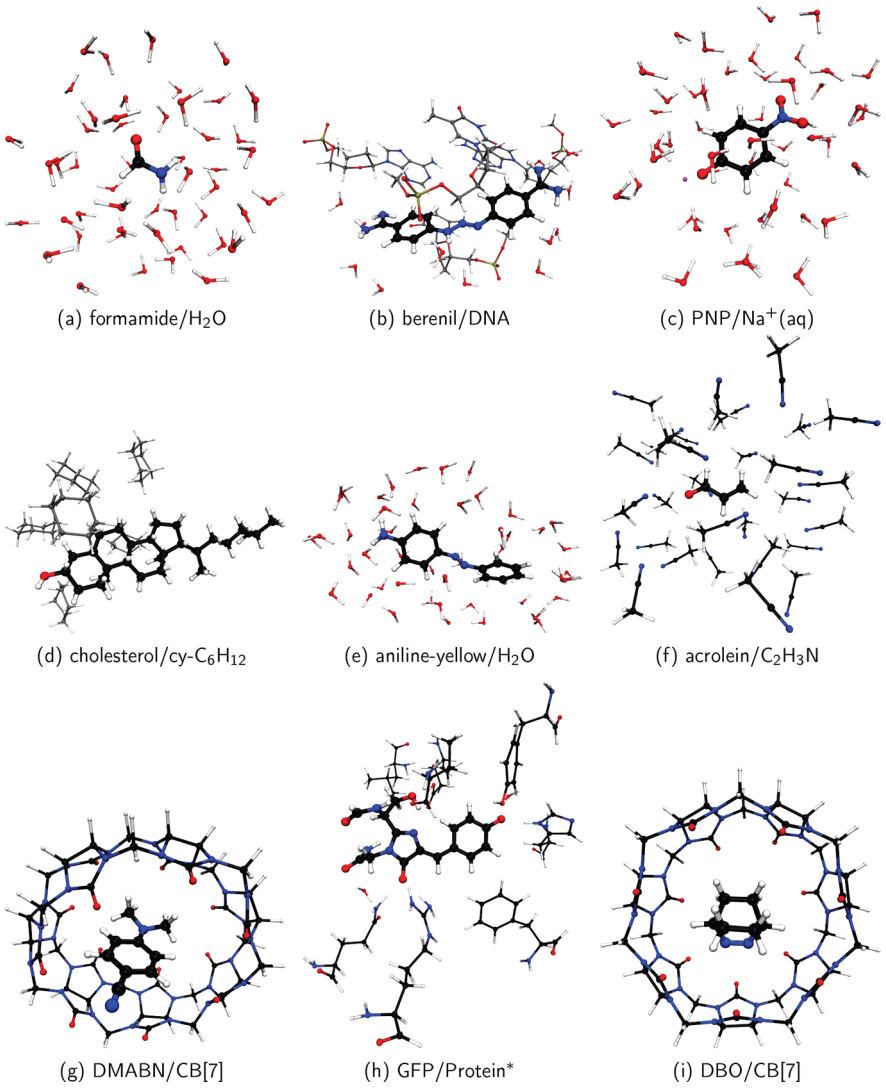


Figure 1. Systems that are considered in this study. The QM and MM subsystems for the QM/MM calculations are distinguished by a "/": (a,f) taken from a classical MD, see the SI, (b) taken from a classical MD, see the SI, (c) taken from ref 54, (d) taken from ref 56, (e) taken from ref 57, (g,i) taken from ref 26, and (h) taken from ref 7 (*with modification).

acetonitrile), biomolecular (protein and DNA) matrices, and a large molecular container (cucurbituril-[7], CB[7]). These examples have been chosen because of their sensitivity to ESO reported in previous excited-state QM/MM studies.[26,56,57] In all QM/MM calculations, we placed only the chromophore in the QM region, and all other molecules are treated at the MM level. In the PE(ECP) calculations, ECPs are added to all atoms of the MM environment. The full-QM TDHF calculations are done using either the cc-pVDZ (DZ) or the aug-cc-pVDZ (aDZ) basis set for all atoms. For the PE-TDHF excitation energies, the embedding parameters, i.e, the atom-centered multipoles and polarizabilities, were determined with the LoProp method[58] using an ANO-recontracted version[59] of the corresponding basis that is used in the full-QM TDHF calculation. The potential calculation was automated using the PyFraME Python package.[60] These calculations are done with the Dalton program.[61]

The full-QM PNO-ADC(2) calculations, PE-ADC(2), and PE(ECP)-ADC(2) calculations are done with the `pnoccsd` and `ricc2` modules of the Turbomole program package.[62] In the full-QM PNO-ADC(2) calculations, mixed basis sets are used with a larger (TZ or aTZ) basis for the chromophore, and a smaller DZ basis set is used for the remaining molecules. Consistently, in the PE-ADC(2) calculations, we used the TZ and aTZ basis for the QM subsystem. For the embedding, we used atomic multipoles obtained from an intrinsic atomic orbital analysis[63] at the B3LYP/DZ level and isotropic polarizabilities from the parameter set for the D3 dispersion correction.[64,65] Furthermore, in the PE- and PE(ECP)-ADC(2) calculations, the LR coupling between the correlated wave function and the polarizable environment is included at the post-SCF LR level.[25,26] In the following, the comparison between excitation energies is based on the NTO analysis in order to ensure that results are compared for states with the same physical character.

Looking first at the statistical measures for the TDHF benchmark set (see Table 2), we see that for a compact basis set, the overall performance between the PE and PE(ECP) models is comparable, with similar mean signed deviation (MSD) (0.05 and 0.04 eV, respectively) and likewise for the mean absolute deviation (MAD). If diffuse basis functions are included, the two embedding models perform very differently.

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 2. Statistical Measures for the Deviations from the Full QM Results of Excitation Energies (in eV) for the QM/MM PE and PE(ECP) Models Combined with TDHF and ADC(2) in the Test Set[a]

	$\Delta E_{\text{full-QM}} - \Delta E_{\text{QM/MM}}$			$\Delta E_{\text{full-QM}} - \delta E_{\text{disp}} - \Delta E_{\text{QM/MM}}$		
method/basis set	MSD	MAD	MAX	MSD	MAD	MAX
PE-TDHF/DZ	0.05	0.06	0.21			
PE(ECP)-TDHF/DZ	0.04	0.05	0.22			
PE-TDHF/aDZ	−0.68	0.69	−3.83			
PE(ECP)-TDHF/aDZ	0.05	0.05	0.21			
PE-ADC(2)/TZ	0.13	0.15	0.48	0.02	0.11	0.34
PE(ECP)-ADC(2)/TZ	0.13	0.14	0.54	0.02	0.10	0.39
PE-ADC(2)/aTZ	−0.95	1.02	−2.91	−1.06	1.07	−3.07
PE(ECP)-ADC(2)/aTZ	0.13	0.14	0.54	0.02	0.07	0.37

[a]The MSD, MAD, and MAX are the mean signed, mean absolute, and maximum deviations, respectively. The TDHF results are reported w.r.t. full-QM TDHF. PE- and PE(ECP)-ADC(2) deviations are compared with the reference PNO-ADC(2) (columns 2−4) and dispersion-corrected NoD-PNO-ADC(2) (columns 5−7) calculations.

Whereas the MSD and mean absolute deviation MAD remain similar to the DZ case for the PE(ECP), the conventional PE approach has much larger errors, with the MAD increased from 0.06 to 0.69 eV. The errors are systematically dominated by highly red-shifted states, as indicated by the large, negative MSD. The MAX of −3.8 eV occurs with the third excited state in the DBO/CB[7] system. Similar large errors are present in the DMABN and cholesterol systems.

The full set of excitation energies is depicted in Figure 2. It clearly demonstrates that the inclusion of ECPs significantly improves the quality of the embedding description when using diffuse functions. Some of the excitation energies that should be above 5 eV collapse with PE to unphysically low values if diffuse basis functions allow the electrons to spill out into the environment. This problem disappears if the ECPs are added. Some states are still accurately described by the original PE scheme even if diffuse functions are included, as evidenced by the clustering of green points near the $y = x$ line. In such cases, including the ECPs has no negative impact on the quality of the description.

Table 2 also shows the statistical measures for the deviations of PE- and PE(ECP)-ADC(2) excitation energies from the reference full-QM PNO-ADC(2) calculations in the test set. For the TZ basis set, the deviations between the full-QM reference and both the PE and the PE(ECP) results are reasonably small with MAD of 0.15 and 0.14 eV, respectively. After excluding the dispersion contribution (NoD-PNO-ADC(2) results), the MADs reduce to 0.11 and 0.10 eV. However, with the aTZ basis, the PE-ADC(2) results have an order of magnitude larger MAD of ≈1.0 eV. The large negative MSD has the same magnitude, indicating again that large red-shifts are responsible for most of the errors in the excitation energies. This dramatic failure is again due to ESO, which we observe not only for nonlocal excitations (see Table 1) but also for local $\pi\pi^*$ and $n\pi^*$ excitations (NTOs are available in the SI).

In contrast, PE(ECP)-ADC(2) performs equally well for basis sets with and without diffuse functions. In combination with the aTZ basis set, the MSD and MAD of PE(ECP)-ADC(2) calculations are 0.13 and 0.14 eV, respectively. Not a single example shows the huge red-shifts from ESO. On the contrary, the MSD between the PE(ECP) and the uncorrected PNO-ADC(2) results is positive, indicating that a significant part of the remaining error is probably due to the missing London dispersion contribution which in most cases lowers the excitation energy. This conclusion is also corroborated by the comparison between the PE(ECP)-ADC(2)/aTZ and the corrected NoD-PNO-ADC(2)/aTZ results for which the MSD is only 0.02 eV, i.e., there is no systematic red- or blue-shift. The MAD and MAX of the PE(ECP)-ADC(2)/aTZ from the NoD-PNO-ADC(2) results are, respectively, 0.07 and 0.37 eV. The MAX is found for the lowest excitation in DMABN/CB[7]. A test calculation revealed that if the PE(ECP)-ADC(2) calculation is done with B3LYP/DZ LoProp-based multipole moments and polarizabilities, the error decreases for this excitation to 0.04 eV. A systematic study of the performance of different parametrizations of the electrostatic embedding will be the subject of forthcoming work.

Figure 3 visualizes the distribution of the results from the QM/MM calculations versus the full QM reference values for the standard TZ and the augmented aTZ basis sets. Both with

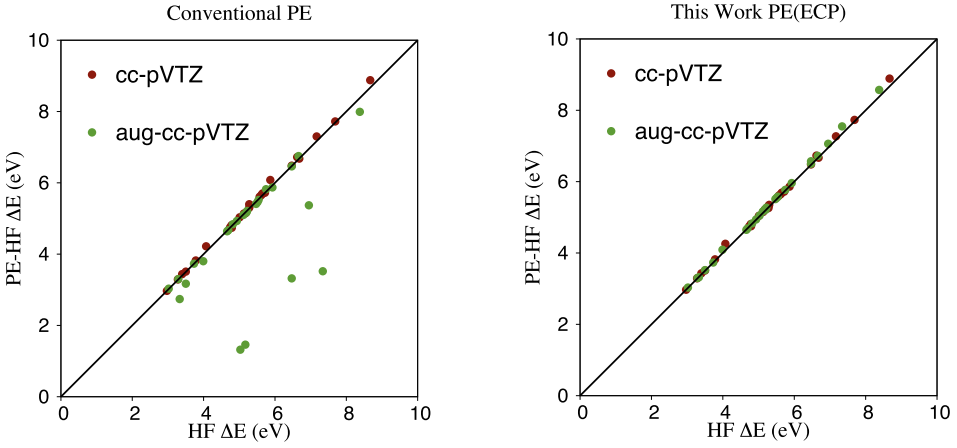


Figure 2. Supermolecular TDHF excitation energies are compared to excitation energies computed with a conventional PE-TDHF (left) or a PE(ECP)-TDHF (right) calculation.

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

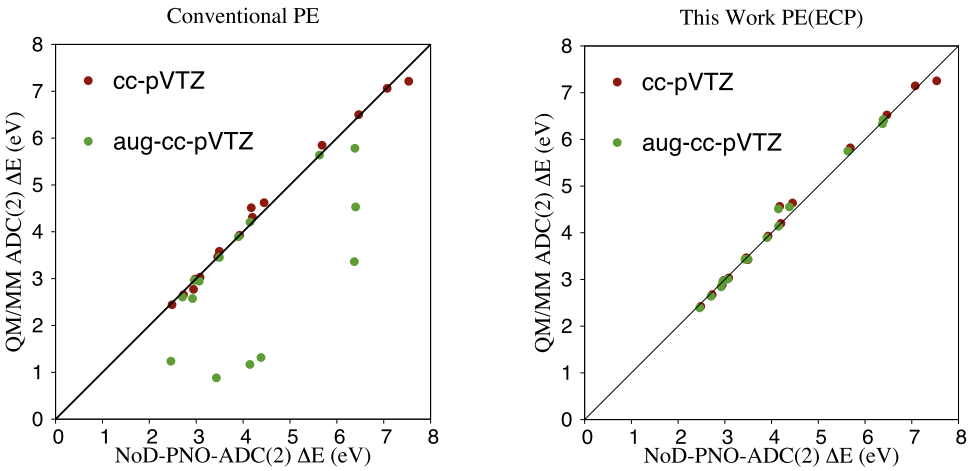


Figure 3. Comparison between (left) PE- and (right) PE(ECP)-ADC(2) QM/MM excitation energies with TZ and aTZ basis sets. The central diagonal line indicates the perfect match between the QM/MM method and the reference NoD-PNO-ADC(2) calculation with the same basis set.

and without including ECPs at the MM sites, the correlation factors, R^2, between the full-QM and QM/MM results at the ADC(2) levels are with the TZ basis 0.99. When including diffuse basis functions, only the PE(ECP) scheme provides physically meaningful results with a correlation coefficient of $R^2 = 0.99$. Due to the high number of cases with severe ESO, the results for the PE scheme without ECPs show essentially no correlation ($R^2 = -0.17$) with the full-QM reference values. In the PE calculations without ECPs, the lack of Pauli repulsion causes ESO at least for half of the cases that were considered in this study.

A comparison between aTZ and TZ results in Figure 3 shows that the full QM PNO-ADC(2)/TZ excitation energies are blue-shifted (on average 0.15 eV) relative to those for the aTZ basis. Typically, the higher the excitation energies are, the more important are the diffuse functions to obtain accurate results and the correct physical character of the excited state. Therefore, the possibility of using extended basis sets in QM/MM calculations is of utmost importance, which requires methods that are unaffected by ESO.

In conclusion, our benchmark study on the excitation energies of several chromophores in different realistic molecular environments reveals that ESO is an important and common pitfall in QM/MM calculations on electronic excitations with basis sets that contain diffuse functions even if the QM subsystem has no bonded interactions into the MM region. The PE(ECP) model with the above-described simple ECPs provides a solution to this problem that is transferable and easy to implement and can be combined with most MM force fields without the need to adapt the parametrization of the electrostatic potential. It is also compatible with the ECPs of the pseudobond approach[47,66−68] and the multicentered valence-electron ECPs that have been introduced by Slavíček and Martinez[48] to treat covalent bonds at the QM/MM boundary. The PE(ECP) model gives a consistently robust performance, even when using diffuse basis functions. At the TDHF level, it provides results with a MAD of less than 0.1 eV from full QM calculations. After excluding the dispersion contribution from the full QM PNO-ADC(2) reference values, also the MAD of the PE(ECP)-ADC(2)/aTZ calculations is for the current test set as small as 0.07 eV, which shows that PE(ECP) parametrization is well-suited for this kind of QM/MM calculations. At this point, the most important remaining error in the current PE(ECP) scheme is, for vertical excitation energies, the missing effect of the dispersion interaction with the environment on the electronic wave function of the QM subsystem. Further tuning of the ECPs, in particular, to fit them for specific MM subsystems seems at this point, at least for vertical excitation energies, not to be necessary. This makes polarizable QM/MM calculations with ECPs a readily accessible, reliable, and cost-effective solution to the electron spill-out problem in excited-state QM/MM calculations. The PE(ECP) approach can easily be integrated into existing codes also for other electronic structures methods like DFT and TDDFT and extended for the calculation of ground- and excited-state molecular gradients[26,69] with no or minor changes to the programs when the required integrals are already available in many program packages.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.9b01162.

ECP parameters, detailed information about computational costs of PE and PE(ECP) models, details of molecular dynamics simulations, and calculated excitation energies and NTO analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Alireza Marefat Khah − *Quantum Chemistry Group, Ruhr University of Bochum, D-44801 Bochum, Germany;* orcid.org/0000-0002-9593-0344; Email: alireza.marefatkhah@rub.de

Peter Reinholdt − *Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense, Denmark;* orcid.org/0000-0003-2406-700X; Email: reinholdt@sdu.dk

Authors

Jógvan Magnus Haugaard Olsen − *Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway;* orcid.org/0000-0001-7487-944X

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Jacob Kongsted − *Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense, Denmark;* orcid.org/0000-0002-7725-2164

Christof Hättig − *Quantum Chemistry Group, Ruhr University of Bochum, D-44801 Bochum, Germany;* orcid.org/0000-0002-5752-2710

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.9b01162

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy − EXC 2033 − 390677874 - RESOLV. C.H. and A.M.K. acknowledge support by the Deutsche Forschungsgemeinschaft within the Priority Program SPP 1807 (HA 2588/10-1). Computations/simulations for the work described herein were supported by the DeIC National HPC Centre, SDU. We acknowledge the Danish Council for Independent Research for financial support (Grant ID: DFF-7014-00050B) and the H2020-MSCA-ITN-2017 COSINE Training network for COmputational Spectroscopy In Natural sciences and Engineering (Project ID: 765739) for financial support. J.M.H.O. acknowledges financial support from the Research Council of Norway through its Centres of Excellence scheme (Project ID: 262695).

■ REFERENCES

(1) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(2) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(3) Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chem. Rev.* **2015**, *115*, 6217−6263.

(4) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited States in Large Molecular Systems Through Polarizable Embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20234−20250.

(5) Morzan, U. N.; de Armiño, D. J. A.; Foglia, N. O.; Ramírez, F.; Lebrero, M. C. G.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in Complex Environments from QM−MM Simulations. *Chem. Rev.* **2018**, *118*, 4071−4113.

(6) Olsen, J. M. H.; List, N. H.; Kristensen, K.; Kongsted, J. Accuracy of Protein Embedding Potentials: An Analysis in Terms of Electrostatic Potentials. *J. Chem. Theory Comput.* **2015**, *11*, 1832−1842.

(7) Nåbo, L. J.; Olsen, J. M. H.; Martínez, T. J.; Kongsted, J. The Quality of the Embedding Potential Is Decisive for Minimal Quantum Region Size in Embedding Calculations: The Case of the Green Fluorescent Protein. *J. Chem. Theory Comput.* **2017**, *13*, 6230−6236.

(8) Thompson, M. A. QM/MMpol: a consistent model for solute/solvent polarization. application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone. *J. Phys. Chem.* **1996**, *100*, 14492−14507.

(9) Gao, J. Energy components of aqueous solution: Insight from hybrid QM/MM simulations using a polarizable solvent model. *J. Comput. Chem.* **1997**, *18*, 1061−1071.

(10) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Mol. Phys.* **2002**, *100*, 1813−1828.

(11) Zhang, Y.; Lin, H.; Truhlar, D. G. Self-Consistent Polarization of the Boundary in the Redistributed Charge and Dipole Scheme for Combined Quantum-Mechanical and Molecular-Mechanical Calculations. *J. Chem. Theory Comput.* **2007**, *3*, 1378−1398.

(12) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.

(13) Olsen, J. M. H.; Kongsted, J. *Advances in Quantum Chemistry*; Elsevier, 2011; Vol. *61*, pp 107−143, DOI: 10.1016/B978-0-12-386013-2.00003-6.

(14) Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Geometry optimization in polarizable QM/MM models: the induced dipole formulation. *J. Chem. Theory Comput.* **2014**, *10*, 1588−1598.

(15) Menger, M. F. S. J.; Caprasecca, S.; Mennucci, B. Excited-state gradients in polarizable QM/MM models: an induced dipole formulation. *J. Chem. Theory Comput.* **2017**, *13*, 3778−3786.

(16) Mortier, W. J.; Genechten, K. V.; Gasteiger, J. Electronegativity Equalization: Application and Parametrization. *J. Am. Chem. Soc.* **1985**, *107*, 829−835.

(17) Chelli, R.; Procacci, P. A Transferable Polarizable Electrostatic Force Field for Molecular Mechanics Based on the Chemical Potential Equalization Principle. *J. Chem. Phys.* **2002**, *117*, 9175−9189.

(18) Wanko, M.; Hoffmann, M.; Frähmcke, J.; Frauenheim, T.; Elstner, M. Effect of Polarization on the Opsin Shift in Rhodopsins. 2. Empirical Polarization Models for Proteins. *J. Phys. Chem. B* **2008**, *112*, 11468−11478.

(19) Lipparini, F.; Barone, V. Polarizable Force Fields and Polarizable Continuum Model: A Fluctuating Charges/PCM Approach. 1. Theory and Implementation. *J. Chem. Theory Comput.* **2011**, *7*, 3711−3724.

(20) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153−168.

(21) Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430−5449.

(22) Boulanger, E.; Thiel, W. Toward QM/MM Simulation of Enzymatic Reactions with the Drude Oscillator Polarizable Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 1795−1809.

(23) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551.

(24) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The Polarizable Embedding Coupled Cluster Method. *J. Chem. Phys.* **2011**, *134*, 104108.

(25) Schwabe, T.; Sneskov, K.; Olsen, J. M. H.; Kongsted, J.; Christiansen, O.; Hättig, C. PERI-CC2: A Polarizable Embeddded RI-CC2Method. *J. Chem. Theory Comput.* **2012**, *8*, 3274−3283.

(26) Marefat Khah, A.; Khani, S. K.; Hättig, C. Excited State Gradients for the QM/MM Polarizable Embedded Second-Order Algebraic Diagrammatic Construction for the Polarization Propagator PE-ADC(2). *J. Chem. Theory Comput.* **2018**, *14*, 4640−4650.

(27) Scheurer, M.; Herbst, M. F.; Reinholdt, P.; Olsen, J. M. H.; Dreuw, A.; Kongsted, J. QM/MM Methods for Biomolecular Systems. *J. Chem. Theory Comput.* **2018**, *14*, 4870−4883.

(28) Hedegård, E. D.; List, N. H.; Jensen, H. J. A.; Kongsted, J. The Multi-Configuration Self-Consistent Field Method within A Polarizable Embedded Framework. *J. Chem. Phys.* **2013**, *139*, No. 044101.

(29) Li, Q.; Mennucci, B.; Robb, M. A.; Blancafort, L.; Curutchet, C. Polarizable QM/MM Multiconfiguration Self-Consistent Field Approach with State-Specific Corrections: Environment Effects on Cytosine Absorption Spectrum. *J. Chem. Theory Comput.* **2015**, *11*, 1674−1682.

(30) Eriksen, J. J.; Sauer, S. P. A.; Mikkelsen, K. V.; Jensen, H. J. A.; Kongsted, J. On the Importance of Excited State Dynamic Response Electron Correlation in Polarizable Embedding Methods. *J. Comput. Chem.* **2012**, *33*, 2012−2022.
(31) Hedegård, E. D.; Reiher, M. Polarizable Embedding Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2016**, *12*, 4242−4253.
(32) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car-Parrinello Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116*, 6941−6947.
(33) Laino, T.; Mohammed, F.; Laio, A.; Parrinello, M. An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.* **2005**, *1*, 1176−1184.
(34) DeFusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling Solvent Effects on Electronic Excited States. *J. Phys. Chem. Lett.* **2011**, *2*, 2184−2192.
(35) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(36) Gökcan, H.; Kratz, E.; Darden, T. A.; Piquemal, J.-P.; Cisnero, G. A. QM/MM Simulations with the Gaussian Electrostatic Model: A Density-based Polarizable Potential. *J. Phys. Chem. Lett.* **2018**, *9*, 3062−3067.
(37) Viquez Rojas, C. I.; Fine, J.; Slipchenko, L. V. Exchange-repulsion energy in QM/EFP. *J. Chem. Phys.* **2018**, *149*, No. 094103.
(38) Minezawa, N.; Zahariev, F.; De Silva, N. Implementation of the analytic energy gradient for the combined time-dependent density functional theory/effective fragment potential method: Application to excited-state molecular dynamics simulations. *J. Chem. Phys.* **2011**, *134*, No. 054111.
(39) 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. *J. Chem. Phys.* **1996**, *105*, 1968−1986.
(40) Krauss, M.; Stevens, W. J. Effective Potentials in Molecular Quantum Chemistry. *Annu. Rev. Phys. Chem.* **1984**, *35*, 357−385.
(41) Hellmann, H. Combined perturbation process in the multiple-electron problem. *C. R. (Dokl.) Acad. Sci. URSS N.S.* **1934**, *4*, 444−446 (in German).
(42) Hellmann, H. A New Approximation Method in the Problem of Many Electrons. *J. Chem. Phys.* **1935**, *3*, 61.
(43) Gombás, P. About the metallic bonding (in German). *Eur. Phys. J. A* **1935**, *94*, 473−488.
(44) Dolg, M.; Cao, X. Relativistic Pseudopotentials: Their Development and Scope of Applications. *Chem. Rev.* **2012**, *112*, 403−480.
(45) Luaña, V.; Pueyo, L. Simulation of ionic crystals: The ab initio perturbed-ion method and application to alkali hydrides and halides. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1990**, *41*, 3800−3814.
(46) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. Variational Optimization of Effective Atom Centered Potentials for Molecular Properties. *J. Chem. Phys.* **2005**, *122*, No. 014113.
(47) Zhang, Y.; Lee, T.-S.; Yang, W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.* **1999**, *110*, 46−54.
(48) Slavíček, P.; Martínez, T. J. Multicentered Valence Electron Effective Potentials: A Solution to The Link Atom Problem for Ground And Excited Electronic States. *J. Chem. Phys.* **2006**, *124*, No. 084107.
(49) DiLabio, G. A.; Hurely, M. M.; Christiansen, P. A. Simple One-Electron Quantum Capping Potentials for Use in Hybrid QM/MM Studies of Biological Molecules. *J. Chem. Phys.* **2002**, *116*, 9578.
(50) Igel-Mann, G.; Stoll, H.; Preuss, H. Pseudopotentials for main group elements (IIIa through VIIa). *Mol. Phys.* **1988**, *65*, 1321−1328.
(51) Helmich, B.; Hättig, C. A Pair Natural Orbital Implementation of The Coupled Cluster Model CC2 for Excitation Energies. *J. Chem. Phys.* **2013**, *139*, No. 084114.
(52) Helmich, B.; Hättig, C. A Pair Natural Orbital Based Implementation of ADC(2)-x: Perspectives and Challenges for Response Methods for Singly And Doubly Excited States in Large Molecules. *Comput. Theor. Chem.* **2014**, *1040−1041*, 35−44.
(53) Development version of TURBOMOLE, 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007. Available from http://www.turbomole.com (accessed Feb 20, 2020).
(54) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modelling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949−5958.
(55) Silva-Junior, M. R.; Sauer, S. P.; Schreiber, M.; Thiel, W. Basis set effects on coupled cluster benchmarks of electronically excited states: CC3, CCSDR(3) and CC2. *Mol. Phys.* **2010**, *108*, 453−465.
(56) Nåbo, L. J.; Olsen, J. M. H.; List, N. H.; Solanko, L. M.; Wüstner, D.; Kongsted, J. Embedding beyond electrostatics−The role of wave function confinement. *J. Chem. Phys.* **2016**, *145*, 104102.
(57) Reinholdt, P.; Nørby, M. S.; Kongsted, J. Modeling of Magnetic Circular Dichroism and UV/Vis Absorption Spectra Using Fluctuating Charges or Polarizable Embedding within a Resonant-Convergent Response Theory Formalism. *J. Chem. Theory Comput.* **2018**, *14*, 6391−6404.
(58) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: The LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494−4500.
(59) Söderhjelm, P.; Krogh, J. W.; Karlström, G.; Ryde, U.; Lindh, R. Accuracy of distributed multipoles and polarizabilities: Comparison between the LoProp and MpProp models. *J. Comput. Chem.* **2007**, *28*, 1083−1090.
(60) Olsen, J. M. H. PyFraME: Python framework for Fragment-based Multiscale Embedding (version 0.2.0); 2018. DOI: 10.5281/zenodo.1443314. Available at https://gitlab.com/FraME-projects/PyFraME.
(61) 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.
(62) TURBOMOLE V7.2.1 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007. Available from http://www.turbomole.com (accessed Feb 10, 2020).
(63) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834−4843.
(64) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(65) Schröder, H.; Schwabe, T. Efficient Determination of Accurate Atomic Polarizabilities for Polarizable Embedding Calculations. *J. Comput. Chem.* **2016**, *37*, 2052−2059.
(66) Zhang, Y. Improved pseudobonds for combined ab initio quantum mechanical/molecular mechanical methods. *J. Chem. Phys.* **2005**, *122*, No. 024114.
(67) Parks, J. M.; Hu, H.; Cohen, A. J.; Yang, W. A pseudobond parametrization for improved electrostatics in quantum mechanical/

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

molecular mechanical simulations of enzymes. *J. Chem. Phys.* **2008**, *129*, 154106.
(68) Chaudret, R.; Parks, J. M.; Yang, W. Pseudobond parameters for QM/MM studies involving nucleosides, nucleotides, and their analogs. *J. Chem. Phys.* **2013**, *138*, No. 045102.
(69) List, N. H.; Beerepooti, M. T. P.; Olsen, J. M. H.; Gao, B.; Ruud, K.; Jensen, H. J. A.; Kongsted, J. Molecular quantum mechanical gradients within the polarizable embedding approach—Application to the internal vibrational Stark shift of acetophenone. *J. Chem. Phys.* **2015**, *142*, No. 034119.

https://dx.doi.org/10.1021/acs.jctc.9b01162
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX