

Embedding-theory-based simulations using experimental electron densities for the environment

Niccolò Ricardi,^a Michelle Ernst,^b Piero Macchi^c and Tomasz Adam Wesolowski^{a*}

Received 5 May 2020

Accepted 16 June 2020

Edited by U. Grimm, The Open University, UK

Keywords: quantum crystallography; density embedding; multi-scale simulations; electronic structure; chromophores.

^aDepartment of Physical Chemistry, University of Geneva, 30, Quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland,

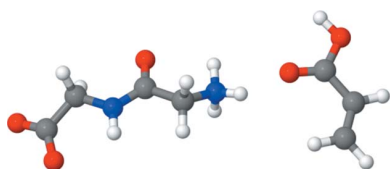
^bUniversity of Bern, Freiestraße 3, 3012 Bern, Switzerland, and ^cDepartment of Chemistry, Materials and Chemical Engineering, Polytechnic of Milan, via Mancinelli 7, Milano 20131, Italy. *Correspondence e-mail:

tomasz.wesolowski@unige.ch

The basic idea of frozen-density embedding theory (FDET) is the constrained minimization of the Hohenberg–Kohn density functional $E^{\text{HK}}[\rho]$ performed using the auxiliary functional $E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B]$, where Ψ_A is the embedded N_A -electron wavefunction and $\rho_B(\mathbf{r})$ is a non-negative function in real space integrating to a given number of electrons N_B . This choice of independent variables in the total energy functional $E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B]$ makes it possible to treat the corresponding two components of the total density using different methods in multi-level simulations. The application of FDET using $\rho_B(\mathbf{r})$ reconstructed from X-ray diffraction data for a molecular crystal is demonstrated for the first time. For eight hydrogen-bonded clusters involving a chromophore (represented as Ψ_A) and the glycylglycine molecule [represented as $\rho_B(\mathbf{r})$], FDET is used to derive excitation energies. It is shown that experimental densities are suitable for use as $\rho_B(\mathbf{r})$ in FDET-based simulations.

1. Introduction

Frozen-density embedding theory (FDET) is the Hohenberg–Kohn theorems-based formal framework for multi-level simulations (Wesolowski, 2004). The total electron density is built up from two components, $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$, of which only the first is constructed from quantum-mechanical descriptors. FDET was originally formulated for variational methods used to obtain such descriptors of the embedded species as: (i) a non-interacting reference system described with a Kohn–Sham determinant (Wesolowski & Warshel, 1993), (ii) an interacting system described with a multi-determinant wavefunction (Wesolowski, 2008), and (iii) a one-particle density matrix (Pernal & Wesolowski, 2009). An extension of FDET for non-variational methods has recently been formulated (Zech *et al.*, 2019). Extensions of FDET for excited states can be made based on the response theory for either non-interacting (Wesolowski, 2004) or interacting (Höfener *et al.*, 2012) systems. Another possibility for describing excited states relies on the Perdew–Levy theorem on extrema of the ground-state energy functional (Perdew & Levy, 1985). It makes it possible to interpret other-than-the-lowest-energy stationary embedded wavefunctions obtained in FDET as excited states, as pointed out by Khait & Hoffmann (2010). In any of these variants of FDET, the embedded wavefunction depends on the chosen $\rho_B(\mathbf{r})$. Several computational methods sharing some elements with FDET but differing in some key aspects, such as the choice of independent variables, self-consistency between the embedding potential and the



OPEN ACCESS

embedded wavefunction, locality of the embedding potential *etc.*, have been developed by various groups. We address the reader to reviews concerning – besides the methods based on FDET – related computational approaches (Wang & Carter, 2000; Wesolowski, 2006; Jacob & Neugebauer, 2014; Wesolowski *et al.*, 2015; Krishtal *et al.*, 2015).

At the present state of development of approximations for the FDET embedding functional [see equation (8) below], applications of FDET are limited to such systems where $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ do not overlap significantly (Wesolowski *et al.*, 1996; Bernard *et al.*, 2008). As a rule of thumb, FDET-based methods are only applicable to such cases where the environment is not covalently bound to the embedded species (Götz *et al.*, 2009; Goodpaster *et al.*, 2010; Fux *et al.*, 2010). In such cases, the overlap between $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ is small and simple local and semi-local approximations are sufficiently accurate. FDET-based simulations can be seen as a variant of quantum mechanics/molecular mechanics simulations, in which the modeller decides on the procedure to generate $\rho_B(\mathbf{r})$ instead of parametrizing the force-field parameters describing the energy contributions due to the interactions between the quantum system and its environment. Various system- and property-specific protocols for generating $\rho_B(\mathbf{r})$ for FDET-based simulations are possible. Some examples of different treatments of the environment density are given below. If the environment comprises several weakly bound molecules, the corresponding $\rho_B(\mathbf{r})$ can be obtained either from quantum-mechanical calculations for the whole cluster comprising all molecules in the environment or, in a simplified manner, as a superposition of molecular densities derived from some quantum-mechanical method (Wesolowski & Warshel, 1994; Humbert-Droz *et al.*, 2014). If $\rho_B(\mathbf{r})$ is localized in a pre-defined part of space, the effect of electronic polarization of the environment by the embedded species can be taken into account by optimizing $\rho_B(\mathbf{r})$ also (Wesolowski & Weber, 1996) or by ‘pre-polarizing’ it using simpler techniques (Zbiri *et al.*, 2004; Ricardi *et al.*, 2018). FDET can also be used to set up a multi-physics simulation in which $\rho_B(\mathbf{r})$ represents a statistical ensemble-averaged electron density [$\langle\rho_B(\mathbf{r})\rangle$] represented as a continuum derived using classical statistical-mechanics-based approaches (Kaminski *et al.*, 2010; Laktionov *et al.*, 2016). Such methods are especially useful for studying the electronic structure of solvated molecules (Shedge *et al.*, 2014).

The above examples show clearly that the choice of the procedure to generate $\rho_B(\mathbf{r})$ is the key element of any FDET-based simulation. This can be made in an ‘automatic’ way by making some system-independent procedures, choices or approximations, or made in a system-dependent manner involving user-provided information about $\rho_B(\mathbf{r})$ such as: (i) using as $\rho_B(\mathbf{r})$ the ground-state density of some system obtained without putting in any information about the embedded species, (ii) localizing $\rho_B(\mathbf{r})$ in a pre-defined region of space by choosing a limited set of atom-centred basis functions, (iii) allowing it to spread over the whole system, (iv) optimizing $\rho_B(\mathbf{r})$ by means of the ‘freeze-and-thaw’ minimization of the total energy (Wesolowski & Weber, 1996), or (v) any combination of the above. In principle, the density

$\rho_B(\mathbf{r})$ obtained from the unique partitioning of the total density using the approach developed by Carter and collaborators (Huang *et al.*, 2011; Huang & Carter, 2011) could be used as a possible ‘automatic’ procedure to generate $\rho_B(\mathbf{r})$ in FDET.

The strategy by which $\rho_B(\mathbf{r})$ is obtained from quantum-mechanical calculations for the environment only, *i.e.* in the absence of the embedded species, is particularly attractive. Both our experience and work by other researchers show that obtaining $\rho_B(\mathbf{r})$ from an isolated calculation yields the dominant contribution to the complexation-induced shifts of the excitation energies, especially for excitations with shifts of large magnitude [see Fradelos *et al.* (2011), Zech *et al.* (2018) and Ricardi *et al.* (2018)]. Daday and co-workers showed that the effects of the optimization of $\rho_B(\mathbf{r})$, either for the ground state alone or for both ground and excited state, are secondary, albeit numerically non-negligible: the excitation energy for methylenecyclopropene solvated by 17 water molecules (for which the reference shift obtained from the reference calculations for the whole cluster equals 0.86 eV) is fairly well reproduced (0.82 eV) with such a choice for $\rho_B(\mathbf{r})$ (Daday *et al.*, 2014). In the case of $n-\pi^*$ excitations for acrolein in water, the corresponding shifts are 1.42 and 1.10 eV, respectively. Although the difference between the FDET with such a choice of $\rho_B(\mathbf{r})$ and the reference shifts cannot be attributed to the ‘neglect of the electronic polarization of the environment’ within the formal framework of FDET, the optimization of $\rho_B(\mathbf{r})$ (Daday *et al.*, 2014) or pre-polarization (Ricardi *et al.*, 2018) usually reduces this difference. This secondary importance of the explicit treatment of the polarization of the environment is due to the variational character of FDET and the fact that the partitioning of the total density of the complex into $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ is not unique in exact FDET, resulting in a better capacity to approach the exact total density [see the discussions by Wesolowski *et al.* (2015) and Humbert-Droz *et al.* (2014)].

The present work concerns yet another possibility for generating $\rho_B(\mathbf{r})$ for FDET simulations of embedded species in a given environment consisting of non-covalently bound molecules, in which $\rho_B(\mathbf{r})$ is obtained from experimental data concerning a different system: a molecular crystal of the environment molecule. Recent years have brought a number of reports showing that both electron densities (Hansen & Coppens, 1978) and wavefunctions (Jayatilaka, 2012) can be reconstructed from X-ray diffraction data. It is tempting, therefore, to explore these new possibilities to generate $\rho_B(\mathbf{r})$ for use in FDET-based simulations. We have to emphasize that several approximations may undermine the use of X-ray-based densities for FDET. The most important issues can be listed as follows: (i) the link to any experimental quantity is of course affected by experimental errors, which are unavoidable and may affect both precision and accuracy; (ii) the electron density and wavefunction that are extracted from experiment are static, whereas atoms are not steady in the crystal; (iii) the sampling of the diffraction in reciprocal space is necessarily incomplete; (iv) only the intensity of the diffracted ray is measured, and not the phase; and (v) the crystal sample is

imperfect. For these reasons, the possibility of using experimental densities as $\rho_B(\mathbf{r})$ in FDET hinges critically on robust and numerically stable protocols to generate such densities, and it is thus important to investigate the dependence of FDET results on such procedures. The current state of the art in wavefunction and electron-density reconstruction from X-ray diffraction data encourages this attempt. In particular, the above-mentioned pitfalls may be tackled as follows: (i) modern instrumentation enables measurement of the diffraction intensities with high precision; (ii) the deconvolution of thermal motion is reliable if the measurements are carried out at a sufficiently low temperature and if the resolution of the diffraction is sufficiently large; (iii) complementary input from theory can compensate for the missing information; (iv) appropriate modelling enables the phasing of the diffracted rays; and (v) data correction from the ideal kinematic theory of diffraction allows for sufficiently accurate data. The present work reports an exploratory study of the use of densities from X-ray restricted wavefunctions in FDET.

Concerning a particular variant of FDET and system to be investigated, we have chosen to evaluate the excitation energies obtained from LinearizedFDET (Wesolowski, 2014; Zech *et al.*, 2015) for several organic chromophores, each hydrogen-bonded to its environment. Our extensive benchmarking of the performance of FDET for such cases indicates that the errors in FDET excitation energies due to the approximations used for the explicit density functional for non-electrostatic components of the FDET embedding potential (see the next section) are small. In a benchmark set of embedded organic chromophores, the average deviation from the reference amounts to about 0.04 eV (Zech *et al.*, 2018). This magnitude of the deviation defines the threshold for complexation-induced shifts in the excitation energy above which analysis of the dependence of the shift on $\rho_B(\mathbf{r})$ is meaningful. In the embedded chromophores chosen for the present study, these shifts vary between 0.15 and 0.6 eV.

2. Embedded chromophores

Concerning the molecules for which $\rho_B(\mathbf{r})$ is generated, we have chosen glycylglycine (GlyGly). For this exploratory study, it is crucial that the molecule(s) corresponding to $\rho_B(\mathbf{r})$ are capable of forming hydrogen bonds with the chromophore. GlyGly satisfies this condition. Moreover, the molecular density of GlyGly reconstructed from X-ray diffraction data reflects features arising from intermolecular hydrogen bonds present in the crystal structure (Genoni *et al.*, 2018). Fig. 1 shows the GlyGly molecule, together with its nearest neighbours in the crystal.

The densities reconstructed from experimental data on glycylglycine are used in the present work as $\rho_B(\mathbf{r})$ in FDET calculations of excitation energies for eight different hydrogen-bonded complexes formed by one organic chromophore (acrolein, acrylic acid or acetone) and glycylglycine. Fig. 2 shows the considered clusters.

The hydrogen-bonding networks shown in Figs. 1 and 2 are not the same. In the crystal, all donors and acceptors are

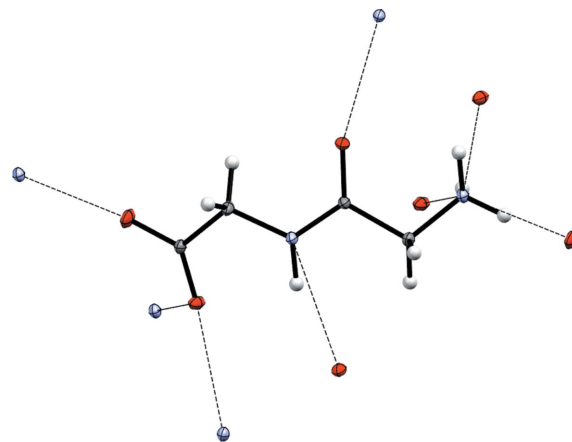


Figure 1
The hydrogen-bonding pattern (dashed lines) for the glycylglycine molecule in the crystal, taken from Dos Santos *et al.* (2014). Only nearest atoms involved in hydrogen bonding are shown: oxygen (red) and nitrogen (blue).

involved in hydrogen bonding, which is not the case in the investigated clusters. Nevertheless, each individual hydrogen bond in the clusters has its corresponding partner in the crystal. It can be expected, therefore, that the effect of the hydrogen bonding on $\rho_B(\mathbf{r})$ in the cluster is also reflected in the density obtained from the crystal.

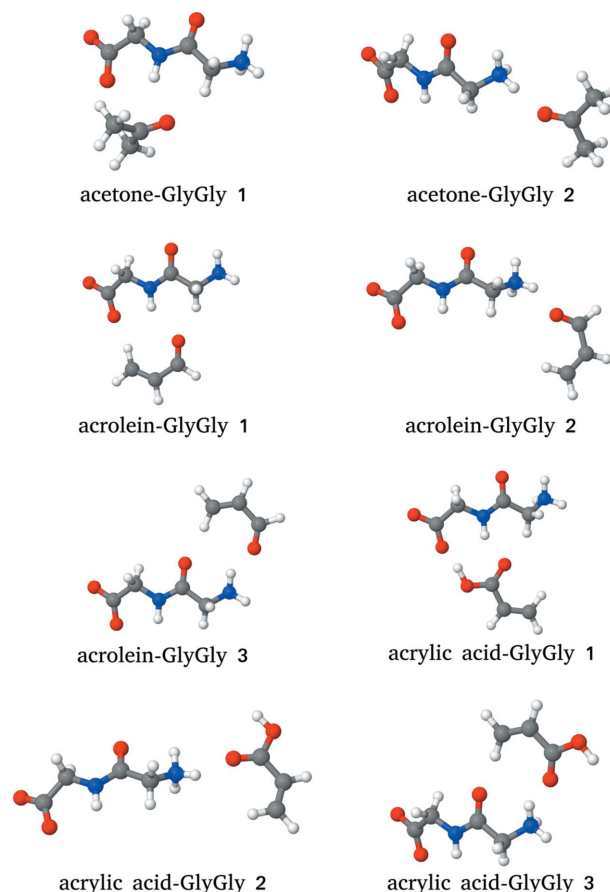


Figure 2
The complexes of three different chromophores with glycylglycine.

3. FDET approach to multi-level simulations

For a system comprising N_{AB} electrons in an external potential $v_{AB}(\mathbf{r})$, the functional $E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B]$ is defined to satisfy by construction the following relation:

$$\min_{\Psi_A} E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B] = E_{v_{AB}}^{\text{FDET}}[\Psi_A^0, \rho_B] = E_{v_{AB}}^{\text{HK}}[\rho_A^0 + \rho_B], \quad (1)$$

where $E_{v_{AB}}^{\text{HK}}[\rho]$ is the Hohenberg–Kohn ground-state energy functional (Hohenberg & Kohn, 1964) and

$$\rho_A^0(\mathbf{r}) = \left\langle \Psi_A^0 \left| \sum_i^{N_A} \delta(\mathbf{r} - \mathbf{r}_i) \right| \Psi_A^0 \right\rangle.$$

By virtue of the second Hohenberg–Kohn theorem, equation (1) leads to

$$E_{v_{AB}}^{\text{FDET}}[\Psi_A^0, \rho_B] \geq E_0, \quad (2)$$

where $E_0 = E_{v_{AB}}^{\text{HK}}[\rho_0]$ and $\rho_0(\mathbf{r})$ are the ground-state energy and density of the total system, respectively. Equality is reached for a large class of densities $\rho_B(\mathbf{r})$,

$$E_{v_{AB}}^{\text{FDET}}[\Psi_A^0, \rho_B] = E_0 \quad \text{if} \quad \forall \mathbf{r} \quad [\rho_0(\mathbf{r}) > \rho_B(\mathbf{r})]. \quad (3)$$

Using conventional density functionals representing components of the total energy, and arbitrary partitioning of the external potential $v_{AB}(\mathbf{r}) = v_A(\mathbf{r}) + v_B(\mathbf{r})$ leads to a form of $E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B]$ more suitable for further discussions,

$$\begin{aligned} E_{v_{AB}}^{\text{FDET}}[\Psi_A, \rho_B] = & \left\langle \Psi_A \left| \hat{H}_A \right| \Psi_A \right\rangle + V_B[\rho_A] + J_{AB}[\rho_A, \rho_B] \\ & + E_{xcT}^{\text{nad}}[\rho_A, \rho_B] + \Delta F[\rho_A] \\ & + E_{v_B}^{\text{HK}}[\rho_B] + V_A[\rho_B] + V_{N_A N_B}, \end{aligned} \quad (4)$$

where

$$\begin{aligned} V_A[\rho_B] &= \int v_A(\mathbf{r}) \rho_B(\mathbf{r}) \, d\mathbf{r}, \\ V_B[\rho_A] &= \int v_B(\mathbf{r}) \rho_A(\mathbf{r}) \, d\mathbf{r} \\ J_{AB}[\rho_A, \rho_B] &= \iint \frac{\rho_A(\mathbf{r}) \rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}', \end{aligned}$$

and $V_{N_A N_B}$ is the interaction energy between the nuclei defining $v_A(\mathbf{r})$ and $v_B(\mathbf{r})$. The non-additive bi-functional $E_{xcT}^{\text{nad}}[\rho_A, \rho_B]$ is related to the functionals $E_{xc}[\rho]$ and $T_s[\rho]$ defined in the constrained-search formulation of the Kohn–Sham formalism (Levy, 1979). It is defined as

$$\begin{aligned} E_{xcT}^{\text{nad}}[\rho_A, \rho_B] = & E_{xc}[\rho_A + \rho_B] - E_{xc}[\rho_A] - E_{xc}[\rho_B] \\ & + T_s[\rho_A + \rho_B] - T_s[\rho_A] - T_s[\rho_B]. \end{aligned} \quad (5)$$

The functional $\Delta F[\rho]$, on the other hand, depends on the form of the wavefunction Ψ used in equation (1) and is also defined via the constrained search (Wesołowski, 2008). For instance, if Ψ_A is a single determinant (Φ), it reads

$$\begin{aligned} \Delta F^{\text{SD}}[\rho] = & \min_{\Phi \rightarrow \rho} \left\langle \Phi \left| \hat{T}_{N_A} + \hat{V}_{N_A}^{ee} \right| \Phi \right\rangle - T[\rho] - V_{ee}[\rho] \\ = & \left\langle \Phi^0[\rho] \left| \hat{T}_{N_A} + \hat{V}_{N_A}^{ee} \right| \Phi^0[\rho] \right\rangle - T[\rho] - V_{ee}[\rho], \end{aligned} \quad (6)$$

where $\hat{V}_{N_A}^{ee}$ is the electron–electron repulsion operator, $V_{ee}[\rho]$ is the density functional of the electron–electron repulsion energy, and $\Delta F^{\text{SD}}[\rho]$ is just the correlation functional ($\Delta F^{[\rho]} \equiv E_c[\rho]$) defined in the constrained-search formulation of density functional theory (Levy, 1979; Baroni & Tuncel, 1983). For Ψ of the full CI form, $\Delta F^{\text{FCI}}[\rho] = 0$ by definition.

Euler–Lagrange optimization of Ψ_A leads to the Schrödinger-like equation

$$(\hat{H}_A + \hat{v}_{\text{emb}}) \Psi_A = \lambda \Psi_A, \quad (7)$$

where

$$\begin{aligned} v_{\text{emb}}[\rho_A, \rho_B, v_B](\mathbf{r}) = & v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}' \\ & + v_{xcT}^{\text{nad}}[\rho_A, \rho_B](\mathbf{r}) + v_F[\rho_A](\mathbf{r}), \end{aligned} \quad (8)$$

with $v_{xcT}^{\text{nad}}[\rho_A, \rho_B](\mathbf{r})$ and $v_F[\rho_A](\mathbf{r})$ being the first functional derivatives of $E_{xcT}^{\text{nad}}[\rho, \rho_B]$ and $\Delta F[\rho]$, respectively.

The lowest-energy solution of equation (7) will be denoted as Ψ_A^{EL} . Note that the energy is given not by the Lagrange multiplier λ but by equation (4). For exact density functionals, any variational method can be used to obtain Ψ_A^{EL} and the corresponding density $\rho_A^{\text{EL}}(\mathbf{r})$, which satisfy by construction the basic FDET equality given in equation (1).

3.1. Reconstruction of $\rho_B(\mathbf{r})$ from X-ray diffraction data

X-ray restrained wavefunctions (XRW), in the literature commonly (but incorrectly) termed X-ray constrained wavefunctions, were initially developed by Jayatilaka and co-workers (Jayatilaka, 2012; Jayatilaka & Grimwood, 2001; Grimwood & Jayatilaka, 2001). Within the XRW paradigm, instead of applying the variational principle, like in conventional SCF, a special functional L is defined, based on a classical Hamiltonian and a function of the square difference between calculated and experimentally measured structure factors, which is ideally distributed with χ^2 statistics. Here,

$$\chi^2 = \frac{1}{N_r - N_p} \sum_k^{N_r} \frac{(F_k - F_k^{\text{exp}})^2}{\sigma_k^2}, \quad (9)$$

with N_r and N_p being, respectively, the number of experimental data and the number of parameters in the model. $(F_k - F_k^{\text{exp}})$ is the difference between the structure factors from the trial wavefunction and the experimental ones, and σ_k is the experimental standard deviation. Thus, the minimization of L implies finding the minimal energy AND the best agreement with experiment. Of course, these cannot be achieved simultaneously and a parameter λ_j must be defined in order to weight the two parts of the functional. Therefore, the functional takes the form

$$L = E + \lambda_j \chi^2. \quad (10)$$

This procedure allows the construction of molecular wavefunctions from experimental observations in crystals. By increasing λ_j , both long- and short-range interactions in the crystal are progressively taken into account. In this work, we used structure factors measured for GlyGly to calculate X-ray restrained wavefunctions with λ_j values from 0.0 to 1.0, since

for higher values the SCF procedure does not converge. We stress that such a value of $\lambda_j = 1.0$ has no specific meaning because the electronic energy of the Hamiltonian and the electron-density difference in the χ^2 function have two different units; thus λ_j is not dimensionless but depends on the number of electrons, the molecular volume and the diffraction resolution. Moreover, the structure factors in the χ^2 function are weighted by the variance of their measurement statistics. The aforementioned wavefunctions were then used to calculate $\rho_B(\mathbf{r})$.

3.2. Computational details

The following approximations were used in the reported FDET calculations: (i) ADC(2) treatment (Schirmer, 1982) of correlation for embedded N_A electrons as implemented by Prager *et al.* (2016), (ii) decomposable approximations for $v_I^{\text{nad}}[\rho_A, \rho_B](\mathbf{r})$ (LDA) and $v_{xc}^{\text{nad}}[\rho_A, \rho_B](\mathbf{r})$ (note that in the LinearizedFDET used here, approximations for the energy components $E_{xcT}^{\text{nad}}[\rho_A, \rho_B]$ and $\Delta F[\rho_A]$ are not used at all), (iii) neglect of the $v_F[\rho_A]$ contribution to the embedding potential, (iv) monomer expansion of $\rho_A(\mathbf{r})$ (only atomic basis sets centred on the chromophore), (v) monomer expansion of $\rho_B(\mathbf{r})$ (only atomic basis sets centred on GlyGly), and (vi) chromophore-independent generation of $\rho_B(\mathbf{r})$ using one of the following methods for the isolated GlyGly: Hartree–Fock, first-order Møller–Plesset (MP) perturbation theory, Kohn–Sham (KS) with PBE (Perdew *et al.*, 1996) approximation for $E_{xc}[\rho]$, coupled clusters singles and doubles (CCSD) or the reconstruction from experimental structure factors (see the next section).

The FDET results for each cluster are compared with the reference obtained from ADC(2) calculations. The reported reference shifts in the excitation energy are evaluated as $\Delta\epsilon_{\text{ref}} = \epsilon_{AB}^{\text{ADC}(2)} - \epsilon_{A(B)}^{\text{ADC}(2)}$, where AB denotes the complex and $A(B)$ denotes the chromophore alone but with the basis set expanded by the functions localized on GlyGly [similar to how it is done in the counterpoise technique of Boys & Bernardi (1970) for intermolecular interaction energy]. In all calculations, including also the reconstruction of the electron density of GlyGly from the X-ray structure factors, the 6-311G basis set was used.

At $\lambda_j = 0$, the wavefunction obtained from the X-ray diffraction data is just the Hartree–Fock molecular wavefunction. The numerical results should be identical, regardless of what software is used to generate $\rho_B(\mathbf{r})$. We used this fact to check the numerical soundness of the procedures to export–import densities $\rho_B(\mathbf{r})$ obtained with different software. *Tonto* (Grimwood *et al.*, 2003) was used for X-ray restrained wavefunction calculations, *Psi-4* (Parrish *et al.*, 2017) to generate the CCSD GlyGly density, and *Q-Chem* (Shao *et al.*, 2015), with its *ADCMAN* (Wormit *et al.*, 2014) and *FDEMAN* (Prager *et al.*, 2016) modules, for all other calculations, including FDET/ADC(2) ones.

Throughout this article, $\epsilon_{\text{emb}}[\rho_B^{\text{method}}]$ (and $\Delta\epsilon_{\text{emb}}[\rho_B^{\text{method}}]$) denote the FDET-derived excitation energy (and environ-

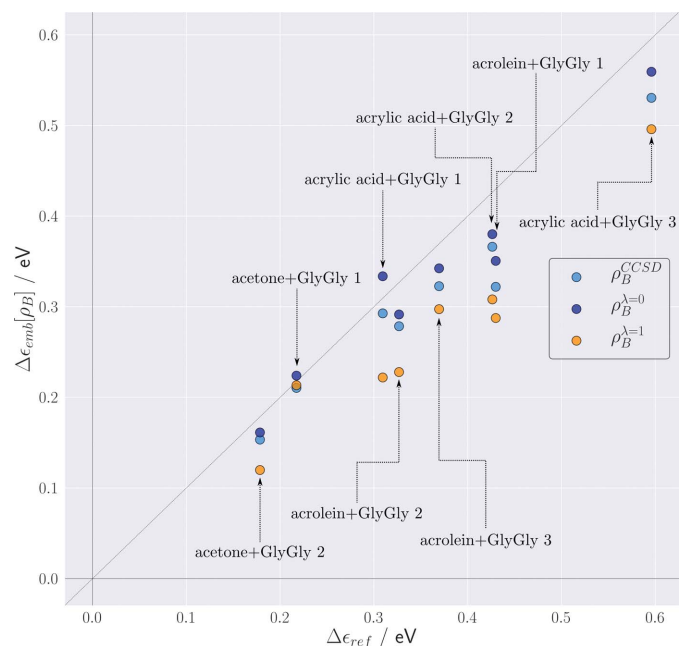


Figure 3

Complexation-induced shifts of the excitation energy ($\Delta\epsilon_{\text{emb}}[\rho_B]$) for eight chromophores hydrogen-bonded to GlyGly. For each complex, FDET calculations are made [embedded ADC(2)] using as $\rho_B(\mathbf{r})$ the electron density of GlyGly obtained from three different methods: Hartree–Fock ($\lambda_j = 0$) or CCSD for the isolated GlyGly, or density reconstructed from X-ray structure factors for the GlyGly molecular crystal at $\lambda_j = 1$. Reference values ($\Delta\epsilon_{\text{ref}}$) are obtained from ADC(2) calculations for the whole complex.

ment-induced shift), where the superscript in $\rho_B^{\text{method}}(\mathbf{r})$ specifies the method used to generate $\rho_B(\mathbf{r})$.

4. Results

For the eight considered clusters, the lowest excitation energies obtained from FDET/ADC(2) calculations ($\epsilon_{\text{emb}}[\rho_B]$) using several choices for $\rho_B(\mathbf{r})$ are shown in Fig. 3, together with the corresponding reference supermolecular ADC(2) results. These excitations have $n-\pi^*$ character and are blue-shifted due to the interactions with the environment. The magnitude of the reference shift falls in the 0.15–0.6 eV range, which makes the shift in these complexes a suitable observable for discussing the effect of the ρ_B -dependency of the FDET results. For this type of excitation, the combined effect of the approximation used for the FDET embedding potential and the use of the isolated environment density as $\rho_B(\mathbf{r})$ results in an average error in the excitation energy of magnitude 0.04 eV (Zech *et al.*, 2018; Ricardi *et al.*, 2018).

We start with an analysis of the results obtained without taking any experimental information from the molecular crystal, *i.e.* $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_j=0}]$. $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_j=0}]$ corresponds to a ‘standard’ FDET protocol in which the Hartree–Fock density of the isolated environment is used as $\rho_B(\mathbf{r})$. The deviations from the reference are small and their magnitude is consistent with the benchmark results published elsewhere (Zech *et al.*, 2018). The effect of correlation on $\rho_B(\mathbf{r})$ [see the shifts obtained with $\rho_B^{\text{CCSD}}(\mathbf{r})$] results in a slight reduction of the shifts in all cases.

At $\lambda_J > 0$, both the correlation and the crystal-field effects are taken into account in $\rho_B(\mathbf{r})$, leading to a further reduction in the shifts. The deviations of FDET shifts from the reference increase (see the values of $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J=1}]$ in Fig. 3).

As previously mentioned, we could not extend the restraint to values of λ_J larger than 1, because the procedure became numerically unstable (Genoni *et al.*, 2018). Unfortunately, although the effects of correlation and polarization by the crystal field are reflected in $\rho_B^{\lambda_J}(\mathbf{r})$, they cannot be separated. Moreover, the environment of GlyGly in the molecular crystal and in the clusters analysed in the present work are different. As a result, even if the reconstruction of the density of GlyGly from X-ray structure factors were exact, this would not guarantee that such density would yield the best FDET results for the clusters under investigation. The values $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J}]$ at $\lambda_J = 0$ and $\lambda_J = 1$ represent, therefore, a good estimate of the maximum scatter (minimal and maximal bounds) of the FDET results due to the ρ_B -dependency of the FDET embedding potential. Within these bounds, the deviations from the reference do not exceed 0.1 eV (or 30% in terms of the relative error). This also points out the need for a thorough analysis of the disentangled effects of correlation and polarization in $\Delta\epsilon[\rho_B^{\lambda_J}]$.

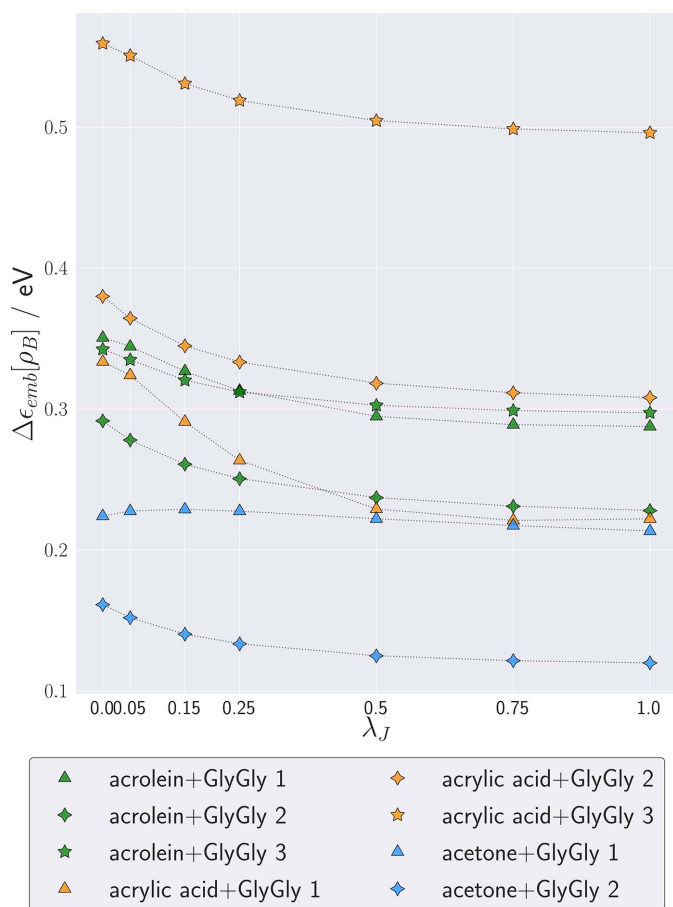


Figure 4 Complexation-induced shifts of the excitation energy ($\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J}]$) at various values of λ_J for eight clusters.

The subsequent part concerns the numerical stability of the FDET-derived complexation-induced shifts of the lowest excitation energy with respect to variations of $\rho_B(\mathbf{r})$ corresponding to the change in the parameter λ_J from 0 to 1.

Fig. 4 shows the dependence of the calculated shifts $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J}]$ on the parameter λ_J for each complex. The dependence of $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J}]$ on λ_J is smooth and monotonic. Above $\lambda_J = 0.5$ up to the maximum value used in this study, $\lambda_J = 1$, $\Delta\epsilon_{\text{emb}}[\rho_B^{\lambda_J}]$ remains almost constant (it changes by as little as about 0.01 eV). The magnitude of the solvatochromic shift decreases for all but one system (acetone + GlyGly 1) when λ_J increases. This can be ascribed to a decrease in dipole-moment magnitude for the XRW densities when λ_J increases. A similar trend appears for correlated methods, which, as is known, tend to yield lower dipole moments.

Turning back to practical applications, we notice that large-scale simulations usually apply the monomer expansion for both $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$, and using the isolated environment density as $\rho_B(\mathbf{r})$ already assures good accuracy of the FDET-derived environment-induced shifts. In such simulations, the modeller has a wide range of available methods to generate $\rho_B(\mathbf{r})$ (see the *Introduction*). The data collected in Fig. 5 show how the FDET results depend on the method used to generate $\rho_B(\mathbf{r})$, including Hartree–Fock, MP1, CCSD and KS-DFT(PBE). For reference purposes, the values of $\Delta\epsilon_{\text{emb}}$ obtained from X-ray diffraction data at $\lambda_J = 0.25$ are also given.

The results collected in Fig. 5 indicate clearly that X-ray-derived molecular densities are suitable for generating $\rho_B(\mathbf{r})$ for FDET calculations following the conventional protocol [LinearizedFDET, monomer expansion of $\rho_A(\mathbf{r})$, monomer

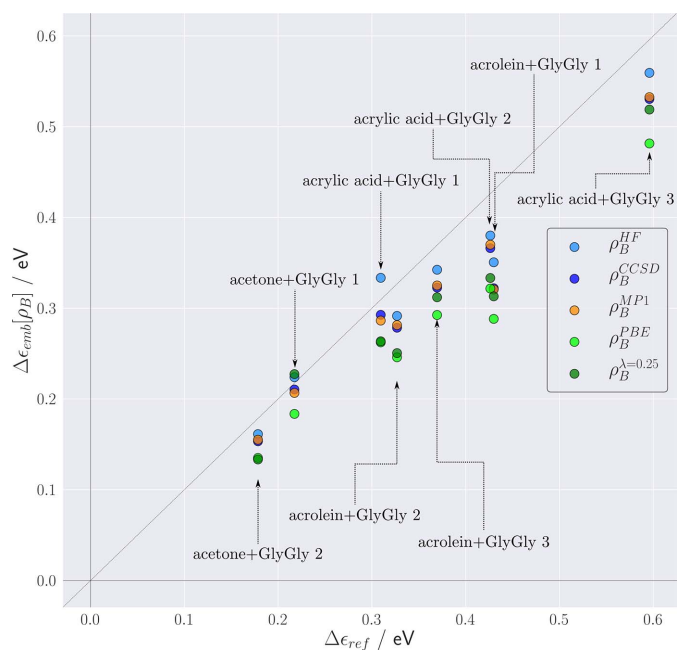


Figure 5 FDET [embedded ADC(2)] derived complexation-induced shifts of the excitation energy ($\Delta\epsilon_{\text{emb}}$) obtained for eight intermolecular complexes with different choices for $\rho_B(\mathbf{r})$. Reference values ($\Delta\epsilon^{\text{ref}}$) are obtained from ADC(2) calculations for the whole complex.

expansion for $\rho_B(\mathbf{r})$, lack of explicit treatment of $\rho_B(\mathbf{r})$ polarization by the chromophore]. The deviations from the reference are, however, larger than if Hartree–Fock or correlated isolated GlyGly densities are used for this purpose. This does not bear direct relevance to the quality of these densities. The overall error of the FDET-derived excitation energy results from the balance of the errors in two FDET embedding potentials evaluated at two different pairs of densities, $v_{\text{emb}}[\rho_A^{\text{ES}}, \rho_B, v_B](\mathbf{r})$ and $v_{\text{emb}}[\rho_A^{\text{GS}}, \rho_B, v_B](\mathbf{r})$, where ρ_A^{GS} and ρ_A^{ES} denote the ground and excited state, respectively, in which the non-electrostatic contributions are approximate.

It is worth noting that the use of X-ray-derived densities as $\rho_B(\mathbf{r})$ leads to smaller errors than if the Kohn–Sham PBE calculations are used for this purpose.

5. Conclusions

Recent developments in techniques to reconstruct the electron density from X-ray diffraction data have made it possible not only to determine the maxima of the electron density (coinciding with the positions of nuclei) but also to reveal its more detailed features (Genoni *et al.*, 2018). For a molecular crystal, the reconstruction yields a localized density of a single molecule but taking into account its chemical environment. In FDET-based simulations, the molecular densities are used as an input quantity providing the complete quantum-mechanical descriptor of the environment of the embedded species. In the present work, we explored the possibility of using the molecular density of glycylglycine derived from X-ray diffraction data collected for the molecular crystal in FDET calculations of the complexation-induced shift of the excitation energy in eight intermolecular complexes, each consisting of an organic chromophore hydrogen-bonded to one glycylglycine molecule.

The usability of such densities for this purpose was not evident before the present study was made. Several factors could, in principle, invalidate such practical applications of X-ray reconstructed densities. First of all, glycylglycine in the crystal and in the complexes analysed in the present work has different environments. This might result in different polarization of such localized molecular densities, and as a consequence, contribute to errors in the FDET results. Other factors relate rather to the reconstruction procedure. It cannot be made perfect due to (i) errors in the experimental measurements, (ii) the very basic assumption according to which the average of a dynamic quantity (electron density) is represented using an intermediate object, namely a static single-determinant wavefunction, (iii) incompleteness of the experimental data, (iv) errors in the phasing procedures, and (v) crystal defects. It might be expected, therefore, that the reconstructed densities would indeed not deviate significantly from the constraint-free one and, in turn, yield similar excitation energies if used as $\rho_B(\mathbf{r})$. The primary objective of this work was the verification of this expectation. The obtained results demonstrate, indeed, that X-ray reconstructed densities are suitable to be used as $\rho_B(\mathbf{r})$ in FDET on a par with possible alternative techniques.

Despite the fact that the X-ray restrained wavefunction procedure does not yield a unique solution, but rather a range of densities parametrized by λ_J , the scatter of the excitation energies obtained using the whole range of this parameter is rather narrow. For all but one complex (acetone + GlyGly 1, characterized by a remarkably short hydrogen bond of only 1.3 Å) the excitation energies vary within approximately 0.05 eV, depending on the details of the reconstruction procedure. This scatter in calculated shifts is small compared with the range of variation in solvatochromic shifts (Reichardt, 1994; Improta *et al.*, 2016), making FDET simulations using X-ray-derived molecular densities an attractive tool for making quantitative predictions and interpreting experimental results. Further reduction of this scatter is probably possible through disentangling the effect of crystal-field polarization and the correlation effect on the electron density of a molecule in a molecular crystal. We intend to deal with this issue in our subsequent work. Also here the experimentally derived $\rho_B(\mathbf{r})$ might prove more useful than alternative techniques. This is the case when the molecules associated with $\rho_B(\mathbf{r})$ have similar neighbours in the cluster to be investigated and in the molecular crystal used to generate $\rho_B(\mathbf{r})$.

For the first time, we adopted an experimental density for the environment and tested this on spectral shifts for valence excitations. At present, a density calculated via an X-ray restrained procedure is the only possibility for this approach. The numerical examples in this work in which we applied the proposed procedure concern microsolvated clusters. For finite systems, many alternatives to generate $\rho_B(\mathbf{r})$ involving similar or even lower computational cost are possible (Hartree–Fock or Kohn–Sham densities of isolated environments including, or not, their optimization or pre-polarization). This numerical validation is the first stage in our long-standing interests and plans aimed at modelling the electronic structure of species in the condensed phase such as neat or doped molecular crystals (Meirzadeh *et al.*, 2018). We plan to apply the same strategy to generate the FDET embedding potential using experimentally derived $\rho_B(\mathbf{r})$ for modelling other spectroscopic properties (core excitations, NMR shifts, two-photon absorption, hyperpolarizabilities *etc.*) that are evaluated from embedded wavefunctions. In infinite systems, the generation of $\rho_B(\mathbf{r})$ from first principles might face serious difficulties if done in a similar way as for the studies of clusters, for the following reasons. Firstly, a density obtained using a straightforward application of the simplest protocol [generation of $\rho_B(\mathbf{r})$ in an artificial system with a void in place of the part described by means of Ψ_A] might be unphysical or even impossible to obtain due to convergence problems or the need for a much larger supercell. Secondly, taking into account the effect of electronic correlation on the electron density in periodic systems is generally limited to Kohn–Sham types of methods. A final point is worth discussing. By using a wavefunction restrained to fit the electron density of a molecule in a crystal, the environment density used in the approach we propose here implicitly includes not only the effects of intermolecular interactions but also the long-range electrostatic effects of crystalline matter.

Funding information

Funding for this research was provided by: Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (grant No. 200020-172532).

References

- Baroni, S. & Tuncel, E. (1983). *J. Chem. Phys.* **79**, 6140–6144.
- Bernard, Y. A., Dulak, M., Kamiński, J. W. & Wesolowski, T. A. (2008). *J. Phys. A Math. Theor.* **41**, 055302.
- Boys, S. F. & Bernardi, F. (1970). *Mol. Phys.* **19**, 553–566.
- Daday, C., König, C., Neugebauer, J. & Filippi, C. (2014). *ChemPhysChem*, **15**, 3205–3217.
- Dos Santos, L. H. R., Genoni, A. & Macchi, P. (2014). *Acta Cryst.* **A70**, 532–551.
- Fradelos, G., Lutz, J. J., Wesolowski, T. A., Piecuch, P. & Włoch, M. (2011). *J. Chem. Theory Comput.* **7**, 1647–1666.
- Fux, S., Jacob, C. R., Neugebauer, J., Visscher, L. & Reiher, M. (2010). *J. Chem. Phys.* **132**, 164101.
- Genoni, A., Bučinský, L., Claiser, N., Contreras-García, J., Dittrich, B., Dominiak, P. M., Espinosa, E., Gatti, C., Giannozzi, P., Gillet, J. M., Jayatilaka, D., Macchi, P., Madsen, A., Massa, L., Matta, C. F., Merz, K. M., Nakashima, P. N., Ott, H., Ryde, U., Schwarz, K., Sierka, M. & Grabowsky, S. (2018). *Chem. Eur. J.* **24**, 10881–10905.
- Goodpaster, J. D., Ananth, N., Manby, F. R. & Miller, T. F. (2010). *J. Chem. Phys.* **133**, 084103.
- Götz, A. W., Beyhan, S. M. & Visscher, L. (2009). *J. Chem. Theory Comput.* **5**, 3161–3174.
- Grimwood, D. J., Bytheway, I. & Jayatilaka, D. (2003). *J. Comput. Chem.* **24**, 470–483.
- Grimwood, D. J. & Jayatilaka, D. (2001). *Acta Cryst.* **A57**, 87–100.
- Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* **A34**, 909–921.
- Höfener, S., Severo Pereira Gomes, A. & Visscher, L. (2012). *J. Chem. Phys.* **136**, 044104.
- Hohenberg, P. & Kohn, W. (1964). *Phys. Rev.* **136**, B864–B871.
- Huang, C. & Carter, E. A. (2011). *J. Chem. Phys.* **135**, 194104.
- Huang, C., Pavone, M. & Carter, E. A. (2011). *J. Chem. Phys.* **134**, 154110.
- Humbert-Droz, M., Zhou, X., Shedge, S. V. & Wesolowski, T. A. (2014). *Theor. Chem. Acc.* **133**, 1405.
- Improta, R., Santoro, F. & Blancafort, L. (2016). *Chem. Rev.* **116**, 3540–3593.
- Jacob, C. R. & Neugebauer, J. (2014). *WIREs Comput. Mol. Sci.* **4**, 325–362.
- Jayatilaka, D. (2012). *Modern Charge-Density Analysis*, edited by C. Gatti & P. Macchi, ch. 6, pp. 213–257. Dordrecht: Kluwer.
- Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* **A57**, 76–86.
- Kaminski, J. W., Gusarov, S., Wesolowski, T. A. & Kovalenko, A. (2010). *J. Phys. Chem. A*, **114**, 6082–6096.
- Khait, Y. G. & Hoffmann, M. R. (2010). *J. Chem. Phys.* **133**, 044107.
- Krishtal, A., Sinha, D., Genova, A. & Pavanello, M. (2015). *J. Phys. Condens. Matter*, **27**, 183202.
- Laktionov, A., Chemineau-Chalaye, E. & Wesolowski, T. A. (2016). *Phys. Chem. Chem. Phys.* **18**, 21069–21078.
- Levy, M. E. L. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 6062–6065.
- Meirzadeh, E., Weissbuch, I., Ehre, D., Lahav, M. & Lubomirsky, I. (2018). *Acc. Chem. Res.* **51**, 1238–1248.
- Parrish, R. M., Burns, L. A., Smith, D. G., Simmonett, A. C., DePrince, A. E., Hohenstein, E. G., Bozkaya, U., Sokolov, A. Y., Di Remigio, R., Richard, R. M., Gonthier, J. F., James, A. M., McAlexander, H. R., Kumar, A., Saitow, M., Wang, X., Pritchard, B. P., Verma, P., Schaefer, H. F., Patkowski, K., King, R. A., Valeev, E. F., Evangelista, F. A., Turney, J. M., Crawford, T. D. & Sherrill, C. D. (2017). *J. Chem. Theory Comput.* **13**, 3185–3197.
- Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865–3868.
- Perdew, J. P. & Levy, M. (1985). *Phys. Rev. B*, **31**, 6264–6272.
- Pernal, K. & Wesolowski, T. A. (2009). *Int. J. Quantum Chem.* **109**, 2520–2525.
- Prager, S., Zech, A., Aquilante, F., Dreuw, A. & Wesolowski, T. A. (2016). *J. Chem. Phys.* **144**, 204103.
- Reichardt, C. (1994). *Chem. Rev.* **94**, 2319–2358.
- Ricardi, N., Zech, A., Gimbal-Zofka, Y. & Wesolowski, T. A. (2018). *Phys. Chem. Chem. Phys.* **20**, 26053–26062.
- Schirmer, J. (1982). *Phys. Rev. A*, **26**, 2395–2416.
- Shao, Y., Gan, Z., Epifanovsky, E., Gilbert, A. T. B., Wormit, M., Kussmann, J., Lange, A. W., Behn, A., Deng, J., Feng, X., Ghosh, D., Goldey, M., Horn, P. R., Jacobson, L. D., Kaliman, I., Khaliullin, R. Z., Kuš, T., Landau, A., Liu, J., Proynov, E. I., Rhee, Y. M., Richard, R. M., Rohrdanz, M. A., Steele, R. P., Sundstrom, E. J., Woodcock, H. L. III, Zimmerman, P. M., Zuev, D., Albrecht, B., Alguire, E., Austin, B., Beran, G. J. O., Bernard, Y. A., Berquist, E., Brandhorst, K., Bravaya, K. B., Brown, S. T., Casanova, D., Chang, C., Chen, Y., Chien, S. H., Closser, K. D., Crittenden, D. L., Diedenhofen, M., DiStasio, R. A. Jr, Do, H., Dutoi, A. D., Edgar, R. G., Fatehi, S., Fusti-Molnar, L., Ghysels, A., Golubeva-Zadorozhnaya, A., Gomes, J., Hanson-Heine, M. W. D., Harbach, P. H. P., Hauser, A. W., Hohenstein, E. G., Holden, Z. C., Jagau, T., Ji, H., Kaduk, B., Khistyayev, K., Kim, J., Kim, J., King, R. A., Klunzinger, P., Kosenkov, D., Kowalczyk, T., Krauter, C. M., Lao, K. U., Laurent, A. D., Lawler, K. V., Levchenko, S. V., Lin, C. Y., Liu, F., Livshits, E., Lochan, R. C., Luenser, A., Manohar, P., Manzer, S. F., Mao, S., Mardirossian, N., Marenich, A. V., Maurer, S. A., Mayhall, N. J., Neuscammann, E., Oana, C. M., Olivares-Amaya, R., O'Neill, D. P., Parkhill, J. A., Perrine, T. M., Peverati, R., Prociuk, A., Rehn, D. R., Rosta, E., Russ, N. J., Sharada, S. M., Sharma, S., Small, D. W., Sodt, A., Stein, T., Stück, D., Su, Y., Thom, A. J. W., Tsuchimochi, T., Vanovschi, V., Vogt, L., Vydrov, O., Wang, T., Watson, M. A., Wenzel, J., White, A., Williams, C. F., Yang, J., Yeganeh, S., Yost, S. R., You, Z., Zhang, I. Y., Zhang, X., Zhao, Y., Brooks, B. R., Chan, G. K. L., Chipman, D. M., Cramer, C. J., Goddard, W. A. III, Gordon, M. S., Hefre, W. J., Klamt, A., Schaefer, H. F. III, Schmidt, M. W., Sherrill, C. D., Truhlar, D. G., Warshel, A., Xu, X., Aspuru-Guzik, A., Baer, R., Bell, A. T., Besley, N. A., Chai, J., Dreuw, A., Dunietz, B. D., Furlani, T. R., Gwaltney, S. R., Hsu, C., Jung, Y., Kong, J., Lambrecht, D. S., Liang, W., Ochsenfeld, C., Rassolov, V. A., Slipchenko, L. V., Subotnik, J. E., Van Voorhis, T., Herbert, J. M., Krylov, A. I., Gill, P. M. W. & Head-Gordon, M. (2015). *Mol. Phys.* **113**, 184–215.
- Shedge, S. V., Zhou, X. & Wesolowski, T. A. (2014). *CHIMIA Int. J. Chem.* **68**, 609–614.
- Wang, Y. A. & Carter, E. A. (2000). *Theoretical Methods in Condensed Phase Chemistry*, pp. 117–184. Dordrecht: Kluwer Academic Publishers.
- Wesolowski, T. A. (2004). *J. Am. Chem. Soc.* **126**, 11444–11445.
- Wesolowski, T. A. (2006). *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczynski, Vol. 10, pp. 1–82. Singapore: World Scientific.
- Wesolowski, T. A. (2008). *Phys. Rev. A*, **77**, 012504.
- Wesolowski, T. A. (2014). *J. Chem. Phys.* **140**, 18A530.
- Wesolowski, T. A., Chermette, H. & Weber, J. (1996). *J. Chem. Phys.* **105**, 9182–9190.
- Wesolowski, T. A., Shedge, S. & Zhou, X. (2015). *Chem. Rev.* **115**, 5891–5928.
- Wesolowski, T. A. & Warshel, A. (1993). *J. Phys. Chem.* **97**, 8050–8053.
- Wesolowski, T. A. & Warshel, A. (1994). *J. Phys. Chem.* **98**, 5183–5187.
- Wesolowski, T. A. & Weber, J. (1996). *Chem. Phys. Lett.* **248**, 71–76.
- Wormit, M., Rehn, D. R., Harbach, P. H., Wenzel, J., Krauter, C. M., Epifanovsky, E. & Dreuw, A. (2014). *Mol. Phys.* **112**, 774–784.

- Zbiri, M., Atanasov, M., Daul, C., Garcia-Lastra, J. M. & Wesolowski, T. A. (2004). *Chem. Phys. Lett.* **397**, 441–446.
- Zech, A., Aquilante, F. & Wesolowski, T. A. (2015). *J. Chem. Phys.* **143**, 164106.
- Zech, A., Dreuw, A. & Wesolowski, T. A. (2019). *J. Chem. Phys.* **150**, 121101.
- Zech, A., Ricardi, N., Prager, S., Dreuw, A. & Wesolowski, T. A. (2018). *J. Chem. Theory Comput.* **14**, 4028–4040.