Multi-scale simulations of the UV-vis absorption spectra of organic chromophores in condensed phases

Zhou, Xiuwen

UNIVERSITÉ DE GENÈVE
Section de chimie et biochimie
Département de chimie physique

FACULTÉ DES SCIENCES

Professeur Tomasz A. Wesołowski

Multi-scale Simulations of the UV-Vis Absorption Spectra of Organic Chromophores in Condensed Phases

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention chimie

par

Xiuwen ZHOU (周秀文)

de

Chine

Thèse N° 4705

GENÈVE

2014

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention chimie

Thèse de *Madame Xiuwen ZHOU*

intitulée :

"Mutli-scale Simulations of the UV-Vis Absorption Spectra of Organic Chromophores in Condensed Phases"

La Faculté des sciences, sur le préavis de Monsieur T. A. WESOLOWSKI, professeur associé et directeur de thèse (Département de chimie physique), Monsieur E. VAUTHEY, professeur ordinaire (Département de chimie physique) et Monsieur M. E. CASIDA, professeur (Laboratoire de chimie théorique, Département de chimie moléculaire, Institut de chimie moléculaire de Grenoble, Université Joseph Fourier, Grenoble, France), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 21 août 2014

Thèse - 4705 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

To Wen-jun

"The days that make us happy make us wise."
— John Masefield

University of Geneva, Switzerland
Doctoral dissertation in Computational Chemistry

Written in LaTeX
Printed by Ateliers d'impression de l'Université de Genève

Contents

Acknowledgments

What would I have done in Geneva without the people who have helped me? At this moment, too many thanks I would like to express.

I would like to start with expressing my great gratitude to my thesis supervisor, professor Tomasz A. Wesolowski. Thank you for encouraging me to enter into the field of computational chemistry. You have taught me not only the way of doing science but also the right attitude to science. You are a mentor, but also a friend, sometime like a father to me. The "lamb" parties at your home are so unforgettable! They bind together the members in our group and leave us so many delightful memories.

I have to express my special thanks to professor Mark E. Casida, for many reasons. First, thank you for being one of my thesis reviewers and your very useful and detailed comments on my thesis. Thank you also for being a very nice collaborator. You are always initiative and friendly. More importantly, thank you for being a friend, sharing your knowledge and experiences with me.

I would like to thank professor Eric Vauthey. Thanks for being one of my thesis reviewers. Thanks also for your wonderful lectures, inspiring discussions, and your strong support in my post doctoral grant application.

I am also very thankful to all my collaborators. Gion Calzaferri, thank you for initiating the "zeolite" project and being a reference on my CV. Ettore Fois, thanks for your great patience while working with me. Ville R.I. Kaila, thanks for being an excellent example as a young scientist. I appreciate your respects to everyone and your passion in science. Dage Sundholm, it was a great pleasure for me to work with you on the manuscript of our "retinal" project. Mark E. Casida, again, my thanks to you for our "flugi" project. Ivano Tavernelli, thanks a lot for instructing me on CPMD code and conveying me the knowledge on molecular dynamics. You are a patient and wonderful teacher! I also enjoyed working with Gloria Tabacchi and André Devaux.

I would like to thank all members in the group. I really enjoyed the four years spent here. Francesco Aquilante, I am so indebted to you for your kind help on many things. Thank you for proofreading my thesis and my research project, tutoring me on theories in quantum chemistry and on programing (can I call you the wikipea in the group?), and being a good friend. Your profound knowledge and Italian humors are so impressive to me! Jakub Kaminski, kuba, thank you so much for your help at the beginning stage of my Ph.D project. Without your guide, I would need to spend several months more for it. Piotr de Silva, thanks for your inspiring discussions and sharing your knowledge. Giovanni LiManni, thanks for the inspiring discussions, lunch brainstorming, afternoon coffee breaks, Friday dinners, and so on. Special thanks should be given to the French speakers in the group, César Beuchat and Marie Humbert-Droz. Thanks for helping me with letters in French, checking the French document I wrote. Sapana Shedge, thank you for being so kind and helpful to me. Georgios Fradelos, thanks for your criticism, from what I can realize my weak points and improve myself. Nirmal Ram, Lyuben Zhechkov, I enjoyed the short but nice period of working with you.

I wish to thank Max Lawson and Yann Sagon for the help on using the clusters “pulstar”, “artemis”, and “baobab”. I would like also to thank all the colleagues in the Department of Physical Chemistry. You are so friendly and the atmosphere here is so pleasant. I would like to thank Swiss National Science Foundation and the University of Geneva for financial support on my research.

I would like to thank all my friends in Geneva. They are too many to mention one by one. However, I would feel guilty without mentioning Laura, Lucie, Sandra, and the big Chinese community in Geneva. Without your accompany, my friends, I cannot enjoy so much the life in Europe.

Last but not least, I would like to thank my family. Thanks to my parents, Ji-hua and Xiang-Jun, for your forever supports and love devoted to me. Thanks to my dear husband, my darling, Wen-jun, for everything and for being there. You are a gift to my life.

Xiuwen Zhou
Geneva, August, 2014.

Abstract

The effect of a condensed-phase environment on the ultraviolet-visible (UV-vis) absorption spectra of chromophores is of fundamental interest in many phenomena, such as color tuning in photobiology and the solvatochromic shift in chemistry. The experimental measurement of UV-vis absorption spectra can provide the information on the spectral region of the absorption of the sample. While computer simulation of the UV-vis absorption spectra can provide the microscopic interpretation of the absorption bands. Moreover, computer simulations can be done for the cases that are impractical for experimental measurements.

This thesis work aims to develop proper computational strategies to calculate the UV-vis absorption of organic chromophores in a condensed-phase environment, and to interpret the experimental results or to tackle problems that are difficult for experimental approaches. Simulation of the UV-vis absorption spectra typically involves modeling of the structure of the target system and the calculation of its vertical electronic excitation energies and oscillator strengths of the interesting part of the system. The large size of the condensed phase makes both parts of the simulation very challenging. In this thesis work, *frozen-density embedding theory*, one of the appropriate approaches on the market for tackling large systems, combined with *linear-response time-dependent density-functional theory*, is used to calculate the vertical electronic excitation energies and the oscillator strengths of the chromophores in condensed phases.

This thesis work investigates the effect of condensed-phase environments on the UV-vis absorption spectra of organic chromophores for several systems, including a functional host-guest material (fluorenone in zeolite-L channel), a chemical solvation system (coumarin 153 in various solvents), and fundamental biological systems in the eye (retinal in rhodopsin and in three visual cone pigments).

Résumé (Abstract in French)

L'effet de l'environnement en phase condensée sur les spectres d'absorption ultraviolet-visible (UV-vis) de chromophores sont d'un intérêt fondamental dans de nombreux phénomènes, comme le réglage de la couleur dans photobiologie et le changement solvatochrome en chimie. La mesure expérimentale des spectres d'absorption UV-vis peut fournir des informations sur la région spectrale de l'absorption de l'échantillon. La simulation par ordinateur de ces spectres peut fournir une interprétation microscopique des bandes d'absorption. En outre, ce genre de simulations informatiques peut être utilisée lorsque les mesures expérimentales ne sont pas pratiques.

Ce travail de thèse vise à développer des stratégies de calcul propres à calculer l'absorption UV-vis des chromophores organiques en phase condensée, dans le but d'interpréter les résultats expérimentaux ou de s'attaquer aux problèmes qui sont trop difficiles pour une approche expérimentale, La simulation des spectres d'absorption UV-vis implique typiquement deux parties: la modélisation de la structure du système cible, et le calcul des l'énergies et les forces oscillatrices d'excitation électroniques du partie système qui nous intéresse. La grande taille de la phase condensée rend les deux parties de la simulation très difficile. Dans ce travail de thèse, nous combinons la très prometteuse théorie dite "frozen density embedding" avec la réponse linéaire de la théorie de la fonctionnelle de la densité dépendante du temps, afin de calculer les énergies et les forces oscillatrices d'excitation électroniques verticales de chromophores en phase condensée.

Ce travail de thèse étudie l'effet de l'environnement en phase condensée sur les spectres de chromophores organiques pour plusieurs systèmes, tels que des matériaux fonctionnels "hôte-invité" (fluorénone dans un canal zéolite L), des molecules solvatée (coumarine 153 dans divers solvants), et des systèmes biologiques fondamentaux pour la vision (le rétinal dans l'enzyme rhodopsine et dans trois pigments de cônes visueles).

List of publications

- **Xiuwen Zhou**, Dage Sundholm, Tomasz A. Wesołowski and Ville R. I. Kaila, *Spectral Tuning of Rhodopsin and Visual Cone Pigments*, J. Am. Chem. Soc., 2014, 136, 2723–2726.
- **Xiuwen Zhou**, Tomasz A. Wesolowski, Gloria Tabacchi, Ettore Fois, Gion Calzaferri and André Deveaux, *First-principles simulation of absorption bands of fluorenone in zeolite L*, Physical Chemistry Chemical Physics, 2013, 15, 159-167.
- **Xiuwen Zhou**, Jakub Kaminski, Tomasz A. Wesolowski, *Multi-scale modelling of solvatochromic shifts from frozen-density embedding theory with non-uniform continuum model of the solvent: the coumarin 153 case*, Physical Chemistry Chemical Physics, 2011, 13, 10565-10576.
- Marie Humbert-Droz, **Xiuwen Zhou**, Sapana V. Shedge and Tomasz A. Wesolowski, *How to choose the frozen density in Frozen-Density Embedding Theory based numerical simulations?*, Theoretical Chemistry Accounts, 2013, 132, 1405.

List of symbols and abbreviations

Φ_{SD}	single Slater determinant
ϕ	molecular orbital, or one-electron spin orbital
ρ	electron density
$\mathbf{A}$, $\mathbf{a}$	boldface letter for matrix or vector notation
B3LYP	Becke 3-parameter Lee-Yang-Parr functional
CI	configuration interaction
DFT	density-functional theory
FDET	frozen-density embedding theory
FDET/LR-TDDFT	frozen-density embedding theory combined with linear-response time-dependent density-functional theory
GGA	generalized gradient approximation
HEG	homogeneous electron gas
HF	Hartree-Fock
KS	Kohn-Sham
LDA	local-density approximation
LR-TDDFT	linear response approach in time-dependent density-functional theory
MCSCF	multi-configurational self-consistent field
MD	molecular dynamics
MM	molecular mechanics
MO	molecular orbital
NDRE	neglect of dynamic response of the environment
PES	potential energy surface
QM	quantum mechanics
TDDFT	time-dependent density-functional theory
xc	exchange-correlation

Chapter 1
Introduction

"Everywhere is walking distance if you have the time."
— Bertrand Russell

1.1 Chromophores in condensed-phase environment

Chromophores are chemical groups with characteristic optical absorptions and their presence is often responsible for the colors of many substances. The color arises when a molecule absorbs certain wavelengths of visible light and transmits or reflects others. The absorbed light can result in a specific electronic transition in which an electron is excited from one molecular orbital to another. There are three typical types of electronic transitions causing the colors of the molecules including [1]: (1) d–d transitions in a d-metal complex. In a d-metal complex, the immediate environment of the metal atom is no longer spherical, the d orbitals are not all degenerate, and therefore the electrons can absorb energy by making transitions between the d orbitals. (2) charge-transfer transitions in a d-metal complex, where the transfer of an electron from the ligands into the d orbitals of the central atom, or vice versa. (3) $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions in a conjugated system. This thesis work focuses on the third type of electronic transitions.

The ability of capturing or detecting light energy of chromophores makes them play an important role in some processes in nature. One important example is the *photosynthesis* that takes place in the plants. The chromophores (chlorophyll and carotenoids) are responsible for harvesting light energy which is converted into chemical energy stored in carbohydrate molecules that can be released later to support the organisms' activities. Another important example is the *visual phototransduction* process in the eye of humans and animals. The visual phototransduction process starts from the light absorption by the retinal chromophore, then the absorbed light is converted into electrical signals in the rod cells, cone cells, and photosensitive ganglion cells of the retina of the eyes, which are used by the brain to interprets different colors. These natural processes have inspired humans to harvest solar energy and transfer it into useful outputs. Chromophores are commonly used as the light harvester in light-to-electricity *energy transducers*, such as the dye-sensitized solar cells (DSSCs).

In realistic systems, chromophores are usually in condensed phases (i.e., liquids or solids) instead of in gas phase. The condensed-phases environment can affect the equilibrium and kinetics of the molecules as well as distinct electronic states of the molecules. In some cases, the effect of condensed-phase environments on the chromophores are of fundamental interest. One of the important cases is *color tuning* in photobiology, which refers when a photofunctional chromophore shows a wide variation of photoabsorption or emission energies depending on different protein environments. This phenomenon can be observed in rhodopsin, human visual cone pigments, firefly luciferase, and red fluorescent proteins. Retinal is the chromophore in rhodopsin and three human visual cone pigments (red-, green-, and blue-cone pigments). It absorbs ultraviolet light

in aqueous solution, while absorbing visible light in vision proteins, which enables the perception of light and color. Another important case of this effect is *solvatochromic shift*, which refers to a strong dependence of absorption and emission spectra of the solvated molecules on the polarity of the solvent medium. The ability of a molecule to change the color due to a change in solvent polarity is called solvatochromism. Molecules with significant solvatochromic shift property can be in principle used in sensors [2] and for the construction of molecular switches in molecular electronics [3]. In summary, it is very important to understand the interactions between the condensed-phase environment with the target chromophores as well as the environment induced changes in the light absorption of the chromophores.

1.2 UV-vis absorption spectroscopy

The color of the molecules can be studied with UV-vis absorption spectroscopy, which refers to the absorption spectroscopy in the ultraviolet-visible spectral region. UV-vis absorption spectroscopy can be used to study the relationship between the color and the structure of compounds and it is routinely used for the quantitative determination of specific chemical species, such as transition metal ions, highly conjugated organic compounds, and biological macromolecules.

Nowadays, as the development of the computer technology and the increased efficiency of the electronic structure methods, the simulation of the UV-vis absorption spectra becomes feasible even for chemical compounds in condensed phase which usually involves systems of large size. This allows the simulation of the spectra to become a complementary tool to experimental measurements of the spectra when investigating a given phenomena. The measured spectra can serve as references for the simulation of the spectra, while the simulated spectra can provide additional information that is difficult to extract from the measured spectra, such as microscopic interpretation of the absorption bands and the interaction between the chromophore and the environment. Moreover, to analyze a target problem, some computer simulations can be performed for cases where the experimental measurements are impractical. For example, the effect of a complex protein environment on the spectroscopic properties of a chromophore can be decomposed into components from each individual amino acid by computer simulation.

1.2.1 Measurement of the UV-vis absorption spectra

The intensity of absorption at a given wavelength can be measured by using the incident and final intensities of a light beam, which is related to the concentration of the absorbing species by the *Beer-Lambert law* [1]:

$$I = I_0 10^{-\varepsilon[J]L}, \tag{1.1}$$

where I_0 and I are the incident and transmitted intensities, respectively, L and $[J]$ is the length and the concentration of the absorber sample, respectively, and ε is the *molar absorption coefficient* (also referred as extinction coefficient). The dimensionless quantity $\varepsilon[J]L$ is called the absorbance (A) of the sample.

The Beer-Lambert law shows that the intensity of the absorption of radiation at a particular wavelength is proportional to the concentration of the absorbing species in the solution and the path length. Therefore the measured UV-vis absorption spectra can be used to determine the concentration of the absorber in a solution.

The maximum value of the molar absorption coefficient, ε_{max}, is an indicator of the intensity of the transition. However, the absorption band spreads across a range of wavenumbers, so a single wavenumber may not reflect the true intensity of the transition. The integrated absorption coefficient, α, is the sum of the molar absorption coefficients over the entire band, and corresponds to the area under the plot of molar absorption coefficient as a function of wavenumber,

$$\alpha = \int_{band} \varepsilon(\tilde{\nu}) d\tilde{\nu}. \tag{1.2}$$

As a crude, but useful approximation, the area of the average absorption band in solution can be characterized in terms of the half-width of the absorption band, $\tilde{\nu}_{1/2}$. This is the width at which $\varepsilon = 1/2\varepsilon_{max}$. Assuming a band is symmetrical about ε_{max}, we have

$$\int_{band} \varepsilon(\tilde{\nu}) d\tilde{\nu} \approx \varepsilon_{max} \Delta\tilde{\nu}_{1/2}. \tag{1.3}$$

1.2.2 Simulation of the UV-vis absorption spectra

The UV-vis absorption spectrum of molecules is characterized by the vertical excitation energies, according to the classic *Franck-Condon principle* [1], which states that the nuclear configuration does not change significantly during the energy absorption process because the rearrangement of electrons is much faster than the motion of nuclei. The vertical excitation energies corresponds to the energy difference between the excited state potential energy curve and the ground state potential energy curve at the minimum point of ground state geometry.

The oscillator strength expresses the probability of the light absorption in electronic transitions between energy levels of an atom or molecule. It is related to the integrated absorption coefficient α by [4]

$$f = \frac{4m_e c\varepsilon_0 \ln 10}{N_A e^2}\alpha = \frac{4m_e c\varepsilon_0 \ln 10}{N_A e^2}\int_{band} \varepsilon(\tilde{\nu}) d\tilde{\nu}, \tag{1.4}$$

where m_e is the mass of the electron, c is the speed of the light in vacuum, ε_0 is vacuum permittivity, N_A is the Avogadro constant, and e is the electric charge of the electron.

From *Fermi's golden rule*, within the *dipole approximation*, the oscillator strength of a electronic transition is [5]

$$f = \frac{2}{3}\frac{m_e}{\hbar^2}(E_f - E_i)|\boldsymbol{\mu}|^2, \tag{1.5}$$

where $(E_f - E_i)$ is the electronic excitation energy, $\boldsymbol{\mu}$ is the transition dipole moment.

Within the Franck-Condon approximation and Born-Oppenheimer approximation (the separation of the total molecular wave function into an electronic and a nuclear wave function), neglecting the translation part of the nuclear wave function, the total transition dipole moment

of an electronic transition from an initial state to a final state in a molecule can be written as [1]

$$\boldsymbol{\mu} = \langle \Psi_f^e | \hat{\boldsymbol{\mu}}_e | \Psi_i^e \rangle \langle \nu_f | \nu_i \rangle, \tag{1.6}$$

where $\langle \Psi_f^e | \hat{\boldsymbol{\mu}}_e | \Psi_i^e \rangle$ is the electronic transition dipole moment; and ν_i and ν_f are the nuclear vibrational wave functions of the initial state and final state, respectively; $\langle \nu_f | \nu_i \rangle$ is called the Franck-Condon factor or Franck-Condon integral. For the simulation of the UV-vis absorption spectrum of molecules in condensed phases with quantum-mechanical method, the vibrational progression is often not considered (without computing Franck-Condon integral) to reduce the computational difficulty.

The simulation of the UV-vis absorption spectrum usually involves two steps:

- modeling the structure of the system, which is often represented by a finite-temperature structure to include the dynamical effect. The finite-temperature structure is typically obtained by Monte-Carlo sampling or molecular dynamics trajectory simulation. The energy minimum geometry is also often used to represent the structure of the system for the simplification of the calculation.

- calculating the vertical electronic excitation energies and the corresponding oscillator strengths for each configuration of the finite-temperature structure of the system or the energy minimum geometry of the system.

1.3 Aims of the thesis work

This thesis work focuses on the theoretical studies of the effect of condensed-phase environment on the UV-vis absorption spectra of several organic chromophores. The aims of this work are:

- To develop proper computational strategies to calculate the UV-vis absorption of organic chromophores in various condensed-phase environments (solid, solvent, and proteins) using existing computational-chemistry methods and taking the experimental data as references.

- Using the validated computational strategies, to interpret the experimental results and to tackle problems that are impractical for transitional experiments to solve.

1.4 Structure of the thesis

The thesis is organized as follows. Firstly, the theory background of the popular computational-chemistry methods is briefly reviewed in Chapter 2, including the wave function based and the density functional based quantum-mechanical methods, molecular-mechanical methods, and hybrid quantum-mechanics and molecular-mechanics methods. The basic concepts, strengths, and limitations of each methods are briefly described.

In Chapter 3, the frozen-density embedding theory is introduced, which is the basis of the computational methods used in the thesis work for multi-scale simulations for the molecules in condensed phases.

The original research work of this thesis is presented in Chapter 4, Chapter 5, and Chapter 6. Each chapter presents a research project on molecules in different type of environments, for which results have been published. Each chapter is organized by including an overall presentation together with the reprint of the published article.

Finally, the general conclusions of the thesis work and future perspectives are discussed in Chapter 7.

1.5 Units used in the thesis

Hartree atomic units (a.u.) are used throughout this thesis, except Chapter 1 (introduction). This convenient system of units is obtained by assigning a value of 1 to the free-electron mass and charge and to the reduced Planck's constant, i.e., $m_e = e = \hbar = 1$.

Chapter 2
Theory background: Computational chemistry

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that exact application of these laws leads to equations much too complicated to be soluble."
— Paul A. M. Dirac

Computer simulation for realistic system of molecules in condensed phase usually involves various levels of computational-chemistry methods, due the large size of the models needed for describing the condensed phase. Therefore, an overview on the popular methods in computational chemistry related to this thesis work will be given in this chapter. The basic concepts of these computational-chemistry methods are presented here based upon the book by Jensen (*Introduction to Computational Chemistry*) [6], the book by Cramer (*Essentials of Computational Chemistry: Theories and Models*) [7], the book by Ullrich (*Time-Dependent Density Functional Theory: Concepts and Applications*) [8], as well as some other cited literature.

2.1 Describing the chemical system

In order to describe a chemical system, we need firstly to choose the fundamental units ("particles") of the system. The choice of "particles" is made by what we want to describe. For example, if we want to describe the overall structure of a protein but not the details of atomic movements, we may choose amino acids as the building blocks. If we want to describe molecular structures but not the details of the electron distribution, we can choose atoms as the building blocks. If we want to describe atoms and molecules, then we need to choose atomic nuclei and electrons as our building blocks. After choosing the particles and the starting condition, we need to analyze the interactions between the particles and write the dynamical equations, which determine how the system evolves in the phase space, and from which we can get the information about the structure and properties of the system.

2.1.1 Dynamical equations of motion

The dynamical equation of motion can be divided into four regimes depending on the mass and velocity of the particles. Light particles display both wave and particle characteristics (with the borderline being approximately the mass of a proton) are described by quantum mechanics. Heavy particles can be described by classical mechanics. For particles moving at high speed (comparable with the speed of light), relativistic effects should be considered in both quantum mechanics and classical mechanics.

In chemistry, atoms and molecules can be treated with classical mechanics because they behave essentially as classical particles (with a few exceptions such as hydrogen, the lightest nucleus). Thus Newton's equation of motion applies:

$$-\frac{\partial V}{\partial \mathbf{r}} = m\frac{\partial^2 \mathbf{r}}{\partial t^2}, \tag{2.1}$$

where the l.h.s. (left-hand side of the equation) is the force, the r.h.s. (right-hand side of the equation) is the mass times the acceleration, and V denotes the potential energy.

For particles with high speed, Newton's equation is formally unchanged, but the particle mass becomes a function of the velocity.

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}} \tag{2.2}$$

The light particles in chemistry is primarily electrons, which should be treated by quantum mechanics. For light particles in low speed, the system is described by the time-dependent Schrödinger equation:

$$\hat{H}\Psi = i\frac{\partial \Psi}{\partial t}, \tag{2.3}$$

where $\hat{H}$ is the Hamiltonian operator and Ψ is the wave function. The Hamiltonian operator is the sum of kinetic and potential energy operators,

$$\hat{H}_{Schrödinger} = \hat{T} + \hat{V} = -\frac{\mathbf{p}^2}{2m} + \hat{V}, \tag{2.4}$$

where $\mathbf{p}$ is the momentum, $\mathbf{p} = -i\hbar\bigtriangledown$.

For light particles moving with high velocities, the Dirac equation applies to take into account the relativistic effect. Its form is analogous to Schrödinger equation (Eq. 2.3), but the Hamiltonian operator is significantly more complicated. It reads

$$\hat{H}_{Dirac} = (c\boldsymbol{\alpha} \cdot \mathbf{p} + \boldsymbol{\beta} mc^2) + \hat{V}, \tag{2.5}$$

where the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are 4×4 matrices. The relativistic wave function consequently has four components. Traditionally, they are labelled as the large and small components, each having an α and β spin function. The large component describes the electronic part of the wave function, while the small component describes the positronic (electron antiparticle) part of the wave function. The $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ matrices couple these components. In the limit of $c \to \infty$, the Dirac equation reduces to the Schrödinger equation, and the two large components of the wave function reduce to the α and β spin-orbitals in the Schrödinger picture. Computational chemists often attempt to include relativistic effect by adding corrections on non-relativistic Schrödinger equation, instead of solving the fully relativistic Dirac equation. For example, effective core potentials can be used to represent relativistic effects, which are largely confined to the core. Both the scalar (spin-free) relativistic effects and spin-orbit (spin-dependent) relativistic effects may be included in effective potentials.

In computational chemistry, the methods to solve Newton's equation of motion (Eq. 2.1) are called *molecular-mechanics* (MM) methods; and those solving the Schrödinger equation (Eq. 2.3)

are referred to quantum-mechanics (QM) or *quantum-chemistry* methods. The quantum-chemistry methods include wave function based and density functional based methods which will be introduced in Section 2.2 and Section 2.3, respectively. The MM methods will be introduced Section 2.4, and the hybrid methods in Section 2.5.

2.1.2 Foundations for quantum-chemistry methods

The common theory foundations for both wave-function-based and density-functional-based quantum-chemistry methods is introduced in this sub-section.

For the bound system, the Hamiltonian operator is time-independent and the system remains stationary. The energy is a constant, depending on only the space variables. The time-independent Schrödinger equation is written as,

$$\hat{H}\Psi = E\Psi. \tag{2.6}$$

In quantum chemistry, one of the primary goals is to solve the time-independent Schrödinger equation for molecules (including an atom as possibility), and to determine the electronic structures of atoms and molecules. Once the electronic structure is known, a wide range of chemically and physically important properties can be determined. For example, by finding the minimum of the potential energy surface (PES) of a stable molecule, that is, the electronic energy plus the nucleus-nucleus repulsion energy, it is possible to calculate its bond lengths and bond angles in its equilibrium structure. Force constants and vibrational frequencies can also be determined from gradients of the PES.

2.1.2.1 Electronic Hamiltonian

An essential part of the solution of the time-independent Schrödinger equation for the complete molecular (atomic) system is separating the electronic and nuclear coordinates by the *Born-Oppenheimer approximation* (BOA). In the BOA, electrons are treated as moving in a fixed nuclear framework, due to the high ratio between nuclear and electronic masses ($m/M \approx 10^{-4}$), i.e., the coupling between the nuclei and electronic motion is neglected. Therefore the electronic Schrödinger equation can be split from that for the complete system. Firstly, for the convenience, the molecular time-independent Schrödinger equation is rewritten as:

$$\begin{aligned} \hat{H}_{\mathrm{tot}}\Psi_{\mathrm{tot}}(\mathbf{R},\mathbf{r}) &= E_{\mathrm{tot}}\Psi_{\mathrm{tot}}(\mathbf{R},\mathbf{r}) \\ \hat{H}_{\mathrm{tot}} &= \hat{T}_{\mathrm{e}} + \hat{T}_{\mathrm{n}} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{ne}} + \hat{V}_{\mathrm{nn}}, \end{aligned} \tag{2.7}$$

where $\mathbf{R}$ denotes the nuclear coordinates and $\mathbf{r}$ denotes the electron coordinates; $\hat{T}$ is the kinetic energy operator; $\hat{V}_{\mathrm{ee}}$ is the electron-electron interaction operator; $\hat{V}_{\mathrm{ne}}$ is the nucleus-electron interaction operator; and $\hat{V}_{\mathrm{nn}}$ is the nucleus-nucleus interaction operator.

The *electronic Schrödinger equation* separated from Eq. 2.7 by BOA is:

$$\begin{aligned}
\hat{H}_{\mathrm{e}}\Psi_{\mathrm{e}}(\mathbf{R},\mathbf{r}) &= E_{\mathrm{e}}(\mathbf{R})\Psi_{\mathrm{e}}(\mathbf{R},\mathbf{r}) \\
\hat{H}_{\mathrm{e}} &= \hat{T}_{\mathrm{e}} + \hat{V}_{\mathrm{ne}} + \hat{V}_{\mathrm{ee}} \\
\hat{T}_{\mathrm{e}} &= -\sum_{i}^{N_{elec}} \frac{1}{2} \bigtriangledown_i^2 \\
\hat{V}_{\mathrm{ne}} &= -\sum_{a}^{N_{nuclei}} \sum_{i}^{N_{elec}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} \\
\hat{V}_{\mathrm{ee}} &= \sum_{i}^{N_{elec}} \sum_{j>i}^{N_{elec}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}
\end{aligned} \tag{2.8}$$

where Ψ_{e} is the electronic wave function, N_{elec} is the total number of the electrons, and N_{nuclei} is the total number of nuclei. The convention adopted here does not include $\hat{V}_{\mathrm{nn}}$ in the electronic Hamiltonian operator $\hat{H}_{\mathrm{e}}$. The nucleus-nucleus repulsion energy V_{nn} is added as a classical term to the electronic energy $E_{\mathrm{e}}(\mathbf{R})$ at the end of the calculation, to provide the potential energy of the total system.

The BOA is an important tool of quantum chemistry, all computations of molecular wave functions for larger molecules make use of it. It is usually a very good approximation. However, it breaks down near a point where two electronic states acquire very close energy, for example, the system in the conical intersection region during a photoreaction. In this case, the non-adiabatic corrections will be large and the quantum nature of the nuclei may need to be taken into account. Starting from here, the quantum-chemistry methods mentioned in this thesis refer to methods for solving the non-relativistic electronic Schrödinger equation.

2.1.2.2 Many-electron wave function

The many-electron wave function in quantum-chemistry methods is generally constructed by taking the one-electron wave functions as basis. The one-electron wave functions in a molecular system are called molecular orbitals (MOs). The pure electronic energy eigenvalue associated with each MO is the energy of the electron in that orbital. A MO is given as the product of a spatial orbital and a spin function (α or β), also known as *spin-orbitals*. A guess wave function of each MO (ϕ_i) is constructed as a linear combination of a set of finite number of known basis functions (χ_μ), which are conventionally called atomic orbitals (AOs) (though they are generally not the solutions of atomic Schrödinger equation). This is called the linear combination of atomic orbitals (LCAO) basis set approximation approach. A MO (ϕ_i) is expanded with AOs (χ_μ) as

$$\phi_i = \sum_{\mu}^{M_{basis}} C_{\mu i}\chi_\mu, \tag{2.9}$$

where μ (numbered 1 to M_{basis}) represents which atomic orbital is combined in the term, and the coefficient $C_{\mu i}$ is the weight of the contribution of the atomic orbital χ_μ to the molecular orbital. The type of the basis functions can be *exponential*, *Gaussian*, *polynomial*, *cube functions*, *wavelets*, *plane waves*, etc.

Slater type orbitals (STO) and Gaussian type orbitals (GTO) are two types of basis functions commonly used in quantum-chemistry electronic structure calculations. Slater type orbitals are exponential functions and have the functional form,

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\xi r}, \tag{2.10}$$

where N is the normalisation constant and $Y_{l,m}$ are spherical harmonic functions. Gaussian type orbitals have the functional form,

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\xi r^2}. \tag{2.11}$$

There are two guidelines for choosing the type of the basis functions for a given problem. The first one is that the behavior of basis functions should agree with the physics of the problem to facilitate the convergence, because, the better a single basis function is able to reproduce the unknown function, the fewer basis functions are necessary for achieving a given level of accuracy. The second one is that the chosen functions should make it easy to calculate all the required integrals.

Therefore based on these two guidelines, for molecular systems, exponential functions (STOs) located on the nuclei is suggested as basis functions because these functions are known to be exact solutions for the hydrogen atom. However exponential functions are computationally difficult. Gaussian functions are also commonly used as basis functions for molecular systems because they are computationally much easier to handle (less computationally demanding for the integrals required in the calculation of electron-electron interaction), although they are poorer at describing the electronic structure on a one-to-one basis (wrong behavior near the nucleus and far away from the nucleus). For periodic systems, plane waves (the exact solutions for a free electron) is suggested as basis functions due to the infinite nature of the boundary condition to the problem.

2.2 Wave function based methods

2.2.1 Introduction

In this section, the Hartree-Fock method is firstly introduced, which is the branching point for wave-function based quantum-chemistry methods. It is the central starting point of most of the more accurate wave-function based methods, and it can also lead to semi-empirical methods by invoking additional approximations. Some of the post-Hartree-Fock methods are introduced after the Hartree-Fock method, but review of the semi-empirical methods is out of the goal of this thesis. This section is presented based upon the book by Jensen [6].

2.2.2 The Hartree-Fock method

The Hartree-Fock (HF) method is an approximate method for solving the many-electron Schrödinger equation (see Eq. 2.8). In this method, the electronic Hamiltonian is expressed as:

$$\begin{aligned} \hat{H}_{\mathrm{e}} &= \sum_{i}^{N_{elec}} \hat{h}_i + \sum_{j>i}^{N_{elec}} \hat{g}_{ij} \\ \hat{h}_i &= -\frac{1}{2} \bigtriangledown_i^2 - \sum_{a}^{N_{nuclei}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} \\ \hat{g}_{ij} &= \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \tag{2.12}$$

where $\hat{h}_i$ is the one-electron operator describing the motion of electron i in the field of all the nuclei, $\hat{g}_{ij}$ is a two-electron operator giving the electron-electron repulsion.

The total electronic wave function is approximated as a product of one-electron wave functions. Each one-electron wave function is described by a MO. The antisymmetric character of the total wave function (it changes sign upon the interchange of any two sets of electron coordinates) is achieved by a single *Slater determinant*:

$$\Psi = \Phi_{\mathrm{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_{\mathrm{N}}(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_{\mathrm{N}}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_{\mathrm{N}}(N) \end{vmatrix}, \tag{2.13}$$

where ϕ_{i} is the one-electron function that is given as the product of a spatial orbital and a spin orbital, Φ_{SD} denotes the single Slater determinant.

The total electron energy of the system is thus:

$$E = \langle \Phi_{SD} | \hat{H}_{\mathrm{e}} | \Phi_{SD} \rangle \tag{2.14}$$

The *variational principle* of quantum mechanics states that any approximate wave function has an energy above or equal to the exact energy, which is used to optimize the HF wave function (determine the best set of orbitals giving the lowest energy) within the restriction of the wave function being a single Slater determinant. This constrained optimization can be handled by means of Lagrange multipliers with the restriction that that a small change in the orbitals should not change the Lagrange function. The HF equations can be derived to find the optimized wave function:

$$\begin{aligned} \hat{F}_i \phi_i^{'} &= \varepsilon_i \phi_i^{'} \\ \hat{F}_i &= \hat{h}_i + \sum_{j}^{N_{elec}} (\hat{J}_j - \hat{K}_j) \\ \hat{h}_i &= -\frac{1}{2} \bigtriangledown_i^2 - \sum_{a}^{N_{nuclei}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} \\ \hat{J}_j |\phi_i(1)\rangle &= \langle \phi_j(1) | \frac{1}{\mathbf{r}_1 - \mathbf{r}_2} | \phi_j(1) \rangle | \phi_i(2) \rangle \\ \hat{K}_j |\phi_i(1)\rangle &= \langle \phi_j(1) | \frac{1}{\mathbf{r}_1 - \mathbf{r}_2} | \phi_i(1) \rangle | \phi_j(2) \rangle \end{aligned} \tag{2.15}$$

where i and j denote the orbitals; 1 and 2 denote the electrons; $\hat{F}_i$ is the one-electron Fock operator; $\phi_i^{'}$ is a special set of MOs, constructed via unitary transformation from the one-electron wave function; $\hat{h}_i$ is the one-electron part of the Fock operator; $\hat{J}$ (Coulomb operator) and $\hat{K}$ (exchange operator) are the two-electron interaction part of the Fock operator.

The optimized set of MOs can be obtained by the eigensolutions to the Fock operator at the energy minimum, therefore finding the HF one-electron wave functions is equivalent to solving the HF equations. The HF equations are a set of pseudo-eigenvalue equations because the Fock operator depends on their own solutions and must therefore be solved iteratively. Therefore the HF method is also a type of self-consistent field (SCF) method.

In the HF method, the electron-electron repulsion is approximated with a mean field because the many-electron wave function is approximated by a single Slater determinant. Therefore the HF total energy is not exact. The HF wave function can account for around 99% of the total energy; but the remaining ∼1% is very important for describing chemical phenomena. The difference in energy between the HF and the "exact" energy (the lowest possible energy in the given basis set) is called the *electron correlation* energy. The electron correlation can be classified as dynamic and static. The dynamic contribution is associated with the "instant" correlation between electrons, such as between those occupying the same spatial orbital. The static part is associated with electrons avoiding each other on a more "permanent" basis, such as those occupying different spatial orbitals. To achieve the *chemical accuracy* (error below 1 kcal/mol), more accurate methods are developed by including the electron correlation.

2.2.3 Post-Hartree-Fock methods

A set of approaches were developed to improve on HF methods which add electron correlation to include the electron-electron repulsion more accurately than in the HF method. They are collectively called *post-Hartree-Fock methods* (also referred as *electron correlation methods* [6]). Usually, post-Hartree-Fock methods give more accurate results than HF calculations. However the added accuracy comes at the price of dramatically increased computational cost. Currently, their application is still restricted to very small molecular systems (typically less than 100 atoms). Therefore they are usually not used to model condensed phase, but they can produce high-quality results which can be taken as references for benchmarking the lower-level methods.

Møller-Plesset perturbation theory

The Møller-Plesset (MP) perturbation theory takes into account the electron correlation effects by means of Rayleigh-Schrödinger perturbation theory [9]. In MP theory, the correlation potential is treated as a perturbation to the unperturbed Hamiltonian which is taken as the sum of the one particle Fock operators. The zeroth-order (unperturbed) wave function is an exact eigenfunction of the Fock operator. The total Hamiltonian is written as:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}, \tag{2.16}$$

where λ is an arbitrary real parameter that controls the size of the perturbation, the correlation potential operator $\hat{V}$ is the difference of the exact electron Hamiltonian operator and the Fock operator.

The energy can therefore be corrected by n-th order MP theory (MPn). The 1st order MP correction gives the HF energy and the 2nd order perturbations gives the HF energy plus the second-order MP correction $E \approx E_{HF} + E_{MP2}$.

The computational cost scaling of the MPn energy is $O(N^{n+3})$, where N is the number of basis function. MP2 is a relatively cheap form of correlation, but the higher orders become comparatively very expensive. Perturbation theory also relies on the starting wave function being close to the exact wave function. For these reasons it can be expected that MP theory will become less popular.

Configuration Interaction

Configuration interaction (CI) takes account into the electron correlation by using linear combination of Slater determinants as many-electron wave function,

$$\Psi = \sum_{I=0} c_I \Phi_I = c_0 \Phi_0 + \sum_{i}^{occ} \sum_{a}^{vir} c_i^a \Phi_i^a + \sum_{i<j}^{occ} \sum_{a<b}^{vir} c_{ij}^{ab} \Phi_{ij}^{ab} + \cdots, \tag{2.17}$$

where Φ_0 is normally HF determinant, the rest of determinants are excitation determinants. The singly excited wave function Φ_i^a is formed by replacing spin-orbital i with spin orbital a in Φ_0

(one electron is excited), The doubly excited wave function Φ_{ij}^{ab} means two electrons are excited (two spin orbitals differ from those in Φ_0), etc.

The set of N spin orbital occupancies in an N-electron Slater determinant is often referred to as a "configuration", thus "configuration interaction" method is just the matrix mechanics solution of the time-independent non-relativistic electronic Schrödinger equation in its most straightforward implementation. The linear coefficients are determined variationally via diagonalization of the Hamiltonian in the given subspace of determinants.

If the expansion of Ψ includes all possible determinants of the appropriate symmetry, then this is a full CI procedure which exactly solves the electronic Schrödinger equation within the space spanned by the one-particle basis set. Full CI calculations are normally used in benchmarking approximate quantum-chemistry methods.

The simplest standard CI method is a CI that adds all singly and doubly substituted determinants (CISD) to the reference determinant. The CISD wave function has become less popular because the quality of truncated CI wave functions degrades for larger molecules. MP2 is a less expensive alternative which gives results similar to those of CISD for small molecules, but the quality of MP2 does not degrade for larger molecules. Coupled-cluster singles and doubles (CCSD), which will be introduced in the following sub-section, is another size-extensive alternative. It typically provides significantly more accurate results with only slightly more costly computations than CISD.

Since the CI eigenvalues are the energies of the ground and some electronically excited states, it is possible to calculate energy differences (excitation energies) with CI methods. Excitation energies of truncated CI methods are generally too high, because the excited states are not as well correlated as the ground state.

Coupled Cluster

Similar to CI, coupled-cluster (CC) also use more complicated many-electron wave functions than HF to include the effects of electron correlation. However the CC wave function uses an exponential expansion of the reference wave function instead of the linear expansion used in CI. For a N-electron system, CC wave function is expressed by the exponential *ansatz* using an exponential cluster operator,

$$\begin{aligned} \Psi = e^{\hat{T}}\Phi_0 = (1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \cdots)\Phi_0 \\ \hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N \\ \hat{T}_1\Psi = \sum_{i}^{occ}\sum_{a}^{vir} t_i^a \Phi_i^a \\ \hat{T}_2\Psi = \sum_{i<j}^{occ}\sum_{a<b}^{vir} t_{ij}^{ab}\Phi_{ij}^{ab}, \end{aligned} \tag{2.18}$$

where i, j refer the occupied orbitals and a, b refer the virtual orbitals.

The reference wave function Φ_0 is typically a HF Slater determinant. $\hat{T}$ is the exponential cluster operator. Inserting the coupled-cluster wave function into the time-independent electronic Schrödinger equation, the coupled-cluster equations can be obtained, which determine the coefficients (t_i^a,t_{ij}^{ab} ,...) for the solution of Ψ.

For a N-electron system, in the limit of including up to $\hat{T}^N$ in the cluster operator, the CC theory is equal to full CI, thus can also produce the exact energy within the basis set used in the wave function. The approximate CC methods truncates the expansion of the wave function by the definition of $\hat{T}$ (control the highest level of excitations allowed). For examples, coupled-cluster doubles (CCD) has $\hat{T} = \hat{T}_2$; coupled-cluster singles and doubles (CCSD) has $\hat{T} = \hat{T}_1 + \hat{T}_2$. CCSD(T) means the connected triples contribution on CCSD is calculated non-iteratively using many-body perturbation theory.

The choice of the exponential *ansatz* guarantees the size extensivity of the solution, but the complexity of equations and the corresponding computer codes (thus the cost of the computation) increases sharply with the highest level of excitation. CCSD offers a computationally affordable method that performs often better than MP2 and CISD, but it doesn't provide sufficient accuracy except for the smallest systems (approximately 2 to 4 electrons), and often an approximate treatment of triples is needed. The most well-known coupled cluster method that provides an estimate of connected triples is CCSD(T), which is rather expensive for molecules with more than a dozen heavy atoms or so. However, when this method is affordable, it provides very high quality results in most cases.

Multi-configurational self-consistent field

For a system with a significant multi-configurational character (where nearly degenerate determinants build the ground state), such as the dissociation process of hydrogen molecule, the HF determinant is qualitatively wrong and so are the CI wave functions and energies. In this case, it is better to use the multi-configurational self-consistent field (MCSCF) method. MCSCF can be considered as the combination of CI and HF, which uses a linear combination of configuration determinants to approximate the exact electronic wave function, however the coefficients of both the determinants and the MOs that build the determinants are optimized by the variational principle (it is simply HF if the "configuration" is only one).

The major issue of MCSCF methods is selecting which configurations are necessary to include for the property of interest. The particularly important MCSCF method is complete active space SCF (CASSCF), where the linear combination of determinants includes all that arise from a particular number of electrons in a particular number of orbitals. Another important MCSCF method is complete active space 2nd-order perturbation theory (CASPT2), where the perturbation theory is applied to CASSCF wave function (analogous to the MP2 perturbation for the HF wave function).

The computational cost of MCSCF calculations increases with the number of configurations included, therefore the size of MCSCF wave functions that can be treated is smaller than that of CI methods. But if it is affordable, CASPT2 can usually be used to reliably calculate molecular ground- and excited states if all other methods fail.

2.3 Density functional based methods

2.3.1 Introduction

Although post-Hartree-Fock methods are widely used for accurate calculations by quantum chemists today, they are restricted to small molecular systems due to the computational difficulties of performing accurate calculations with large basis sets for large systems. Density functional theory (DFT) is a very popular alternative to post-Hartree-Fock methods. It can be used to do calculations on large molecular systems with acceptable accuracy but in significantly less time than those post-Hartree-Fock methods. Moreover, for systems with d-block metals, DFT often yields results which agree more closely with experiment than post-Hartree-Fock calculations do. [1] The basic concept of DFT is that the properties of an N-electron system can be obtained from the electron density (depending on 3 spatial variables), rather than the N-electron wave function (depending on 3N spatial variables) that is even difficult to store if N is large.

Both DFT that deals with ground-state properties and the time-dependent DFT (TDDFT) that deals with excited-state properties are introduced in this section.

2.3.2 Density-functional theory

The electron density ρ of an N-electron system is related to its wave function as:

$$\rho(\mathbf{r}_1) = \int |\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)|^2 d\mathbf{r}_2 \cdots d\mathbf{r}_N. \tag{2.19}$$

The concept of a density functional for the energy was the basis of some early but useful approximate models such as the Thomas-Fermi model. Then the Hohenberg-Kohn theorems gave a formal proof that the ground-state energy and all other ground-state electronic properties are uniquely determined by the electron density. However it proves only that such a functional exists but does not provide the form of the energy density functional. The Kohn-Sham method makes it practical to obtain the ground-state electron density, which treats a fictitious system with non-interacting electrons having the same exact ground state electron density as that of the real system.

2.3.2.1 The Thomas-Fermi Model

There were already some early attempts to approximate the electron energy as functional of the electron density soon after the discovery of the Schrödinger equation in 1926. The most well known one is probably Thomas-Fermi model developed independently by both Thomas [10] and Fermi [11] in 1927. They used a homogeneous electron gas (HEG) model (electron distributed uniformly in phase space) to approximate the distribution of electrons in the atoms. In this approach, the kinetic energy is approximated as:

$$T = \langle \Psi | \hat{T} | \Psi \rangle \approx T_{TF}[\rho] = C_{TF} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r}, \tag{2.20}$$

where the constant $C_{TF} = \frac{3}{10}(3\pi^2)^{2/3}$. The electron-electron interaction energy is approximated by the expression for a classically repulsive gas (including only the classical Coulomb repulsion energy):

$$V_{ee} = \langle \Psi | \hat{V_{ee}} | \Psi \rangle \approx J[\rho], \tag{2.21}$$

where $J[\rho]$ is the classical Coulomb repulsion energy

$$J[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2. \tag{2.22}$$

In this expression of electron-electron interaction energy, the electron exchange and electron correlation are completely neglected. The total energy of the system is then given as:

$$E = T + V_{ee} + V_{ext} \approx E_{TF} = T_{TF}[\rho] + J[\rho] + \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} \tag{2.23}$$

Later, an exchange energy functional for the electron-electron interaction energy was added by Dirac in 1928 [12]. However, the Thomas-Fermi-Dirac theory was rather inaccurate for most applications, which is due to the crude approximation for the kinetic energy, the errors in the exchange energy, and the complete neglect of electron correlation.

2.3.2.2 The Hohenberg-Kohn Theorems

Although there were already early efforts for finding energy functionals, the basic existence proof of density functional theory was not given until the publication of a landmark paper by Hohenberg and Kohn in 1964 [13], which proved two well known theorems, known as the Hohenberg-Kohn theorems.

Hohenberg-Kohn theorem 1: In a finite, interacting N-elecron system with a given particle-particle interaction there exists a one-to-one correspondence between the external potential ($v(\mathbf{r})$) and the ground-state density ($\rho(\mathbf{r})$), i.e., the external potential is a unique functional of the ground-state density, up to an arbitrary additive constant.

This theorem can be proved easily by two steps. First, it can be proved that the relationship between the potentials and wave functions is unique. Then the one-to-one correspondence between the wave function and the electron density can be proved using the Rayleigh-Ritz variational principle, thus $\Psi = \Psi[\rho]$. Therefore a unique correspondence exists between potentials and ground-state densities, which can be written as $v[\rho](\mathbf{r})$.

From the first Hohenberg Kohn theorem, we know that the Hamiltonian is also a functional of electron density, $\hat{H}[\rho]$ (because the kinetic operator and electron-interaction operator are fixed for a N-electron system). Moreover, we can see that all energy eigenstates of the system become density functionals as well (not only the ground state wave function). This means all ground-

and excited-state properties of a given N-electron system are in principle determined entirely by the ground-state electron density.

Hohenberg-Kohn theorem 2: The total energy functional ($E_{HK}[\rho]$) reaches minimum at the non-degenerate ground state density (ρ_0).

The second Hohenberg-Kohn theorem provides a variational principle for the energy functional using the electron density as the variational quantity instead of the wave function, and in principle this can be used to obtain the ground-state electron density. The Hohenberg-Kohn total energy functional associated with a given external potential $v(\mathbf{r})$ can be written by

$$E_v^{HK}[\rho] = \langle\Psi[\rho]|\hat{T}+\hat{V}_{ee}+\hat{V}_{ext}|\Psi[\rho]\rangle = T[\rho]+V_{ee}[\rho]+\int\rho(\mathbf{r})v(\mathbf{r})d^3\mathbf{r} = F^{HK}[\rho]+\int\rho(\mathbf{r})v(\mathbf{r})d^3\mathbf{r}, \tag{2.24}$$

where $F^{HK}[\rho]$ is a universal functional (the same for an N-electron system with the same electron-electron interaction) and is defined as

$$F^{HK}[\rho] = \langle\Psi[\rho]|\hat{T}+\hat{V}_{ee}|\Psi[\rho]\rangle = T[\rho]+V_{ee}[\rho]. \tag{2.25}$$

According to the second Hohenberg-Kohn theorem, the exact ground-state density $\rho_0(\mathbf{r})$ can be found from the Euler-Lagrange equation

$$\frac{\delta E_v^{HK}[\rho]}{\delta\rho(\mathbf{r})} = \frac{\delta F^{HK}[\rho]}{\delta\rho(\mathbf{r})}+v(\mathbf{r}) = \mu, \tag{2.26}$$

where μ is a Lagrange multiplier which ensures the correct total number of electrons. However the exact form of $F^{HK}[\rho]$ is unknown, and in practice the variational theorem is only applied with approximations to it. Besides, there is another difficult problem that the existence of the energy functional derivative in Eq. 2.26 requires that ρ belongs to some external potential (v-representability) [14].

The v-representability is still an unsolved problem. Fortunately the *constrained search* formalism proposed by Levy [14] and Lieb [14] overcomes most of the difficulties related to the functional $E_v^{HK}[\rho]$. In the constrained search, the variational principle is expressed as a two-step minimization. For a N-electron system, it can be expressed as

$$E_0 = \min_{\Psi}\langle\Psi|\hat{T}+\hat{V}_{ee}+\hat{V}_{ext}|\Psi\rangle = \min_{\rho}\{\min_{\Psi\to\rho}\langle\Psi|\hat{T}+\hat{V}_{ee}+\hat{V}_{ext}|\Psi\rangle\}, \tag{2.27}$$

which means we first search over all antisymmetric N-electron wave functions (Ψ) that yield a given density ρ, then determine the ground-state density ρ_0 as the density that gives the lowest energy. This leads to another way of defining the universal functional,

$$F^{LL}[\rho] = \min_{\Psi\to\rho}\langle\Psi|\hat{T}+\hat{V}_{ee}|\Psi\rangle. \tag{2.28}$$

The advantage of the constrained search formalism is that it gives an operational definition of the universal functional in the form of a constructive procedure. This constrained search procedure

to search the ground-state electron density is not very practical (needs to search over infinite number of wave-functions), but it plays an important formal and conceptual role in DFT.

2.3.2.3 The Kohn-Sham formalism

Kohn-Sham (KS) method, which is developed by Kohn and Sham in 1965 [15], transforms DFT into a practical scheme. It is nowadays the basis of most applications of DFT. The key idea in the Kohn-Sham method is that the ground-state electron density of an interacting N-electron system can be obtained by solving a fictitious non-interacting N-electron problem.

For a non-interacting N-electron system, the Hamiltonian is comprised by only the kinetic-energy operator and the external potential energy operator, and the exact solution to the Schrodinger equation (the many-body ground-state wave function) can be given as a single Slater determinant Φ_{SD} composed of (molecular) orbitals, ϕ_j. According to HK theorems, its total-energy functional can be written as

$$E_s[\rho] = T_s[\rho] + \int \rho(\mathbf{r}) v_s(\mathbf{r}) d^3\mathbf{r}, \tag{2.29}$$

where $T_s[\rho]$ is the kinetic-energy functional of the non-interacting N-electron system, $v_s(\mathbf{r})$ is the external potential of the the non-interacting N-electron system. The ground-state density can be found either by solving the Euler-Lagrange equation

$$\frac{\delta E[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + v_s(\mathbf{r}) = \mu, \tag{2.30}$$

or by solving the single-electron Schrodinger equations

$$\left(\frac{\nabla^2}{2} + v_s(\mathbf{r}) \right) \phi_j(\mathbf{r}) = \varepsilon_j \phi_j(\mathbf{r}), \tag{2.31}$$

$$\rho_s(\mathbf{r}) = \sum_{j=1}^{N} |\phi_j(\mathbf{r})|^2, \tag{2.32}$$

where $\phi_j(\mathbf{r})$ is the single-electron wave function in Ψ_s, and the ground-state density $\rho_s(\mathbf{r})$ is obtained from the N lowest (occupied) single-electron orbitals.

For an interacting N-electron system, the total-energy functional $E[\rho]$ can be written in a clear way (to compare with that for the non-interacting N-electron system):

$$\begin{aligned} E[\rho] &= T[\rho] + V_{ee}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d^3\mathbf{r} \\ &= T_s[\rho] + J[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d^3\mathbf{r} + E_{xc}[\rho] \end{aligned} \tag{2.33}$$

where $J[\rho]$ is the classical Coulomb energy (also called Hartree energy) defined in Eq. 2.22, $v(\mathbf{r})$ is the external potential of the interacting N-electron system, and $E_{xc}[\rho]$ is called the exchange-

correlation energy which is defined as

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]. \tag{2.34}$$

The Euler-Lagrange equation to obtain ground-state electron density of the interacting N-electron system is

$$\frac{\delta E[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} = \mu, \tag{2.35}$$

Eq. 2.35 is formally identical to Eq. 2.30, if we identify an effective single-particle potential as

$$v_s[\rho](\mathbf{r}) = \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + v(\mathbf{r}) + v_{xc}[\rho](\mathbf{r}), \tag{2.36}$$

where $v_{xc}[\rho](\mathbf{r})$ is defined as

$$v_{xc}[\rho](\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \tag{2.37}$$

Then solving the Eq. 2.35 for the interacting system is equivalent to solve the single-electron Schrödinger equation of a fictitious non-interacting system:

$$\left(-\frac{\nabla^2}{2} + v_s[\rho](\mathbf{r}) \right) \phi_j(\mathbf{r}) = \varepsilon_j \phi_j(\mathbf{r}), \tag{2.38}$$

and the ground-state electron density is obtained by summing the squares of the lowest N occupied orbitals:

$$\rho_0(\mathbf{r}) = \sum_{j=1}^{N} |\phi_j(\mathbf{r})|^2. \tag{2.39}$$

Eq. 2.38 and Eq. 2.39 are called the Kohn-Sham equations.

In practice, the Kohn-Sham equations are almost always implemented in their spin-resolved form. For an N-electron system comprised by $N_\uparrow$ spin-up and $N_\downarrow$ spin-down electrons, $N = N_\uparrow + N_\downarrow$. The total ground-state electron density is a sum of spin-up and spin-down densities. The Kohn-Sham equations have the form

$$\rho_0(\mathbf{r}) = \rho_{0\uparrow}(\mathbf{r}) + \rho_{0\downarrow}(\mathbf{r}) = \sum_{\sigma=\uparrow,\downarrow} \sum_{j=1}^{N_\sigma} |\phi_{j\sigma}(\mathbf{r})|^2, \tag{2.40}$$

$$\left(-\frac{\nabla^2}{2} + v_{s\sigma}[\rho_\uparrow, \rho_\downarrow](\mathbf{r}) \right) \phi_{j\sigma}(\mathbf{r}) = \varepsilon_{j\sigma} \phi_{j\sigma}(\mathbf{r}). \tag{2.41}$$

The Kohn-Sham effective potential is written as

$$v_{s\sigma}[\rho_\uparrow, \rho_\downarrow](\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + v_\sigma(\mathbf{r}) + v_{xc\sigma}[\rho_\uparrow, \rho_\downarrow](\mathbf{r}), \tag{2.42}$$

where the xc energy and potential are defined as functionals of the individual spin densities,

$$v_{xc\sigma}[\rho_\uparrow, \rho_\downarrow](\mathbf{r}) = \frac{\delta E_{xc}[\rho_\uparrow, \rho_\downarrow]}{\delta \rho_\sigma(\mathbf{r})}. \tag{2.43}$$

In practice, the functional dependence of E_{xc} on the spin-up and spin-down densities is better suited for the construction of approximations.

Although Kohn-Sham theory is exact in principle, it is approximate in practice because the exchange-correlation functional is unknown. Kohn-Sham theory makes the unknown contribution to the total energy (the exchange-correlation energy) as small as possible. In fact it is the smallest part of the total energy for many systems. However E_{xc} is still extremely important because its contribution can be significant to the binding energy of some systems. Therefore an accurate description of exchange and correlation is crucial for the prediction of binding properties.

2.3.2.4 Approximated exchange-correlation functionals

Since the exact form of exchange-correlation (xc) energy functional E_{xc} is unknown, approximated functionals must be developed for practical calculations. Many approximate functionals exist. A given xc functional may work very well for some properties and some classes of systems, but not in other situations. Therefore it is extremely important to choose a proper functional for the desired properties of a target system. Xc functionals essentially try to model the exchange-correlation hole, which is a quantum-mechanical zone surrounding every electron in an interacting system that reduces the probability of finding other electrons within the immediate vicinity. Here the basic xc functionals and those used in this thesis work are introduced.

The local-density approximation The local-density approximation (LDA) is the oldest approximation of DFT originally proposed by Kohn and Sham in 1965 [15]. In LDA, the true exchange-correlation (xc) energy of a system is approximated by the xc energy associated with a homogeneous electron gas, which is the only system with the form of the xc energy known precisely. The xc energy in LDA depends on only the local density, and has the form:

$$E_{xc}^{LDA} = \int e_{xc}^{HEG}\left(\rho(\mathbf{r})\right) d^3\mathbf{r}, \tag{2.44}$$

where $e_{xc}^{HEG}(\rho)$ is the xc energy density (the xc energy per unit volume) that corresponds to a homogeneous electron gas. The exchange energy in LDA is expressed by the exchange energy of the homogeneous electron gas which is known exactly:

$$E_x^{LDA} = \int \left(-\frac{3}{4}(\frac{3}{\pi})^{1/3}\rho(\mathbf{r})^{4/3}\right) d^3\mathbf{r}. \tag{2.45}$$

The exact expression of the correlation energy of the homogeneous electron gas is unknown but highly accurate numerical results exist. Based on the available numerical data, high-precision analytical parametrizations have been developed. Various correlation energy functionals of LDA were generated with the analytic forms for correlation energy of the homogeneous electron gas, such as the Vosko-Wilk-Nusair (VWN) [16], Perdew-Zunger (PZ81) [17], and Perdew-Wang (PW92) [18] parametrizations.

In practice, the LDA is implemented in its more general spin-polarized form (the LSDA) which has the form,

$$E_{xc}^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \int e_{xc}^{HEG}\left(\rho_\uparrow(\mathbf{r}), \rho_\downarrow(\mathbf{r})\right) d^3\mathbf{r}. \tag{2.46}$$

Despite the simplicity of LDA, it works surprisingly well for many energetic and structural properties in a wide range of materials. It performs almost always better than the HF approximations and predicts the right physical trends. But it has also some failings. One must be aware of the limitations in its application. It has a tendency to favor more homogeneous systems and over-binds molecules and solids. These errors are exaggerated and bond lengths are too short in weakly bonded systems (such as hydrogen bond).

The generalized gradient approximation The generalized gradient approximation (GGA) is a class of xc functionals that depend not only on the local density itself but also its gradients. The GGA functionals have the general form (spin-polarized form):

$$E_{xc}^{GGA}[\rho_\downarrow, \rho_\uparrow] = \int e_{xc}^{GGA}\left(\rho_\downarrow(\mathbf{r}), \rho_\uparrow(\mathbf{r}), \bigtriangledown\rho_\downarrow(\mathbf{r}), \bigtriangledown\rho_\uparrow(\mathbf{r})\right) d^3\mathbf{r}, \tag{2.47}$$

where the arrows indicate the spin directions, and the key idea of GGAs is to construct explicit mathematical expressions for e_{xc}^{GGA}.

Today there exists hundreds of GGA functionals. The most widely used GGA functionals include: exchange functional B88 [19] by Becke; correlation energy functionals P86 [20] by Perdew, and LYP [21] by Lee, Yang and Parr; xc energy functionals BP86 composed of B88 exchange and P86 correlation, PW91 [22] by Perdew and Wang, and PBE [23] by Perdew, Burke and Ernzerhof.

The GGAs improve much over the LDAs in several instances. The most notable outcome was the significant reduction in the LDA overbinding error for solids and molecules. GGAs are currently the most popular exchange-correlation functionals in condensed matter physics.

Although the successes of LDA and GGA functionals in the calculation of electronic structure of materials, they are not sufficiently accurate to meet the chemical accuracy required in solving problems in chemistry, for example, the prediction of chemical reactions. Therefore efforts have been made to develop more accurate xc functionals, including *meta-GGA*, *hybrid functionals*, *double hybrids*, etc.

Hybrid functionals The hybrid approach for constructing density functional approximations was introduced by Becke in 1993 [24]. Hybrid functionals incorporate a portion of exact exchange from Hartree-Fock theory in the xc energy functionals. The exact exchange energy functional is expressed in terms of the Kohn-Sham orbitals. They are constructed by mixing a fraction of the exact exchange energy functional with a standard LDA or GGA and usually have the general expression:

$$E_{xc}^{hybrid} = aE_x^{exact} + (1-a)E_x^{standard} + E_c^{standard}, \tag{2.48}$$

where a is a semi-empirical constant. A typical value of a is 0.25. PBE0 is the functional obtained by substituting the functionals in Eq. 2.48 with PBE functionals.

The most popular hybrid functional today is the Becke 3-parameter Lee-Yang-Parr (B3LYP) [25] which is written as:

$$E_{xc}^{B3LYP} = aE_x^{exact} + (1-a)E_x^{LDA} + bE_x^{B88} + cE_c^{LYP} + (1-c)E_c^{LDA}, \tag{2.49}$$

where $a = 0.20$, $b = 0.72$ and $c = 0.81$. B3LYP outperforms most GGA functionals and LDA for the structural and energetic properties of molecules close to equilibrium. It also performs excellently (the experimental results as reference) for the calculations on the vertical electronic excitation energies of some organic chromophores. It is nowadays the most successful functional for chemical applications and accounts for about 80% of DFT applications. [26] It is of course not perfect. For example, it fails for "free electron like" metallic systems because the LYP correlation functional does not reproduce the correct limit for homogeneous systems; it fails to describe the dissociation of the H_2 molecule correctly due to the lack of static correlation; it usually does not work for electronic transitions with change-transfer character in time-independent DFT calculations.

Range-separated hybrid exchange functionals The range-separated hybrid exchange (RSHX) functionals were developed to fix the problem of the hybrid functionals in describing bond dissociation and change-transfer excitations due to their poorly described long-range electron-electron interaction. Let us consider the exchange energy in terms of the density, $\rho(\mathbf{r})$ and of the exchange hole, $h_x(\mathbf{r}, \mathbf{r}')$ as

$$E_x[\rho] = \frac{1}{2}\int\int \rho(\mathbf{r}) v_{ee}(\mathbf{r}, \mathbf{r}') h_x(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'. \tag{2.50}$$

The Coulomb interaction ($v_{ee}(\mathbf{r}, \mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-1}$) is separated into a short-range and a long-range part,

$$v_{ee}(\mathbf{r}, \mathbf{r}') = v_{ee}^{sr}(\mathbf{r}, \mathbf{r}') + v_{ee}^{lr}(\mathbf{r}, \mathbf{r}') = \frac{f(\mu|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} + \frac{1 - f(\mu|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}, \tag{2.51}$$

where μ is the range separation parameter, and the function f has the properties $f(x \to \infty) = 0$ and $f(x \to 0) = 1$. The above decomposition of the Coulomb operator leads to an exact partition of the exchange energy to short and long range components,

$$E_x[\rho] = E_x^{sr,\mu}[\rho] + E_x^{lr,\mu}[\rho]. \tag{2.52}$$

At short range, the standard LDA or GGA exchange is proved behaving reasonably. For the long range component of the exchange energy, the Hartree exchange is used to ensure its correct asymptotic behavior. The RSHX energy functional is constructed by combining the short and long range exchange components with a total correlation functional,

$$E_{xc}^{RSHX}[\rho] = E_x^{sr,\mu,DF}[\rho] + E_x^{lr,\mu,HF}[\rho] + E_c^{DF}[\rho]. \tag{2.53}$$

A crude estimate of the range separation parameter can be given as the average distance of the electrons, i.e. twice the Seitz radius, around 1-2 a.u. for valence electrons.

The Coulomb-attenuating method corrected B3LYP (CAM-B3LYP) [27] is one of the popular Range-seperated functionals, which is the long range corrected version of B3LYP. It performs well for charge transfer excitations in molecules which B3LYP underestimates enormously.

Range-seperated functionals improves significantly properties such as the polarizabilities of long-chain molecules and bond dissociations. It is particularly important in TDDFT calculations for Rydberg and charge-transfer excitations.

Other functionals The statistical average of (model) orbital potentials (SAOP) [28] is an approach on accurate model exchange-correlation orbital-dependent potentials, which uses a statistical average of different model potentials for the highest occupied Kohn-Sham orbital and the rest of the occupied Kohn-Sham orbitals to get the proper Coulomb asymptotic behavior of the xc potential that is very important for calculating the response properties.

SAOP potential can give accurate one-particle eigenspectra and their derived properties such as electronic excitations and (hyper)polarisabilities, etc. For spectroscopic properties, they usually give results superior to those obtained with LDA or GGA potentials and it produces results comparable with to exact (Hartree-Fock) exchange based methods but with higher computational efficiency for electronic excitations in the condensed phase.

2.3.3 Time-dependent density-functional theory

The properties and dynamics of many-body systems in the presence of time-dependent potentials (such as electric or magnetic fields) are investigated by the excited state methods. The general excited state methods include higher roots in MCSCF and CI calculations, propagator method (also called Green's function approach or equation-of-motion method), and TDDFT. The linear response approach in TDDFT (LR-TDDFT) [29], is a very popular approach for extracting features like excitation energies, frequency-dependent response properties, and photoabsorption spectra. In this section, the general formalism of TDDFT and its practical scheme LR-TDDFT are introduced using the book by Ullrich

The conceptual and computational foundations of TDDFT are analogous to those of ground-state DFT. In TDDFT, it is shown that the time-dependent wave function is equivalent to the time-dependent electronic density, and then the effective potential of a fictitious non-interacting system is derived to produces the same density as any given interacting system. But the construction of such a system is more complex than ground-state DFT. Most notably because the time-dependent effective potential at any given instant depends on value of the density at all previous times. The development of time-dependent approximations for the implementation of TDDFT is still far behind that of DFT. In practical calculations, this memory requirement is usually ignored by invoking the adiabatic approximation (see Eq 2.60).

2.3.3.1 The Runge-Gross existence theorem

The Runge-Gross theorem (1984) [30] is the formal foundation of TDDFT, just like the first Hohenberg-Kohn theorem to ground-state DFT.

The Runge-Gross theorem: Two densities $\rho(\mathbf{r},t)$ and $\rho'(\mathbf{r},t)$, evolving from a common initial many-body state Ψ_0 under the influence of two different potentials $v(\mathbf{r},t)$ and $v'(\mathbf{r},t) \neq v'(\mathbf{r},t)+c(t)$ (both assumed to be Taylor-expandable around t_0), will start to become different infinitesimally later than t_0. Therefore, there is a one-to-one correspondence between densities and potentials, for any fixed initial many-body state.

The Runge-Gross theorem guarantees a rigorous one-to-one correspondence between time-dependent densities and time-dependent external potentials. The one-to-one correspondence holds both for fully interacting systems and for non-interacting particles.

2.3.3.2 Time-dependent variational principles

In ground-state DFT, the second Hohenberg-Kohn theorem plays a prominent role by which the ground state of a system can be obtained by minimizing the expectation value of the Hamiltonian using the Rayleigh-Ritz variational principle. In the time-dependent case, the time evolution of the systems can be derived with other variational principles called *stationary-action principles*, which are analogous to the ground-state minimum principle. There are several different stationary-action principles established in TDDFT. In 1984, Runge and Gross introduced a stationary-action principle which is closely related to Dirac-Frenkel principle in quantum mechanics. The action A is defined as a functional of the many-body wave function $\Psi(t)$ between an initial time t_0 and a final time t_1:

$$A[\Psi] = \int_{t_0}^{t_1} \langle \Psi(t)| i\frac{\partial}{\partial t} - \hat{H}(t)|\Psi(t)\rangle dt \tag{2.54}$$

The true time evolution of the system is that the action is stationary to a small variation of the wave function:

$$\delta A[\Psi] = 0 \tag{2.55}$$

However this principle later turned out to be conflict with causality (a change in the potential at a given time can not affect the density at earlier times). To solve this problem, Van Leeuven [31, 32] later defined an action functional on Keldysh contour. Recently (2008), Vignale [33] proposed a straightforward alternative formulation of TDDFT action principle. The differences of these action principles are discussed in Ref. [8].

2.3.3.3 The Van Leeuwen theorem

The time-depedent Kohn-Sham (TDKS) equation and the time-dependent xc potential can be derived with the Van Leeuwen theorem [31], without recourse to an action-based variational principle.

The van Leeuwen theorem: For a time-dependent density $\rho(\mathbf{r},t)$ associated with a many-body system with a given particle-particle interaction $v(|\mathbf{r}-\mathbf{r}'|)$, external potential $v_{ext}(\mathbf{r},t)$, and initial state Ψ_0, there exists a different many-body system featuring an interaction $v'(|\mathbf{r}-\mathbf{r}'|)$ and a

unique external potential $v'_{ext}(\mathbf{r},t)$ (up to a purely time-dependent $c(t)$) which reproduces the same time-dependent density. The initial state Ψ'_0 in this system must be chosen such that it correctly yields the given density and its time derivative at the initial time.

The Van Leeuwen theorem guarantees that we can find an auxiliary non-interacting system that reproduces the same time-dependent density as a given interacting system. This proof of this theorem is not presented here, for interested readers, it could be found in Ref. [8].

2.3.3.4 The time-dependent Kohn-Sham scheme

According to van Leeuwen theorem, the time-dependent density can be produced by a correspondent non-interacting Kohn-Sham system. The time-dependent density is obtained by summing over occupied Kohn-Sham single-electron states:

$$\rho(\mathbf{r},t) = \sum_{j=1}^{N} |\phi_j(\mathbf{r},t)|^2, \tag{2.56}$$

The time-dependent Kohn-Sham one-electron equation is written as:

$$\left(-\frac{\nabla^2}{2} + v_s[\rho](\mathbf{r},t)\right)\phi_j(\mathbf{r},t) = i\frac{\partial}{\partial t}\phi_j(\mathbf{r},t), \tag{2.57}$$

with the initial condition

$$\phi_j(\mathbf{r},t_0) = \phi_j^0(\mathbf{r}), \tag{2.58}$$

where $\phi_j^0(\mathbf{r})$ is the ground-state Kohn-Sham orbital. The effective potential $v_{eff}(\mathbf{r},t)$ is given by

$$v_s[\rho](\mathbf{r},t) = \int \frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}' + v_{ext}(\mathbf{r},t) + v_{xc}[\rho](\mathbf{r},t), \tag{2.59}$$

where $v_{xc}[\rho](\mathbf{r},t)$ is the time-dependent xc potential. The exact expression of this functional of the density is unknown and the construction of approximations is needed. In contrast to ground-state DFT, where very good xc functionals exist, approximations to $v_{xc}[\rho](\mathbf{r},t)$ are still in their infancy. However, there is a very simple procedure that allows the use of existing xc functionals for ground-state DFT in the time-dependent theory by using the *adiabatic approximation*, in which the time-dependent xc potential is approximated as the ground-state potential at a given ground-state density which equals to the instant time-dependent density:

$$v_{xc}[\rho](\mathbf{r},t) \approx v_{xc}^{AA}(\mathbf{r},t) = v_{xc}^0[\rho_0](\mathbf{r})|_{\rho_0(\mathbf{r})=\rho(\mathbf{r},t)}, \tag{2.60}$$

where $v_{xc}^0[\rho_0](\mathbf{r})$ is the ground-state xc potential functional. The adiabatic approximation work surprisingly well although most situations of practical interest are nonadiabatic at least to some degree. However due to the lack of memory (it is a local approximation in time), it has also failures such as the optical properties of solids and long conjugated molecules, double excitations, and charge-transfer excitations.

2.3.3.5 Linear-response time-dependent density-functional theory

The time-dependent Kohn-Sham equation is general and can be applied to essentially any time-dependent situation. There are two main ways to solve it: real-time propagation and linear-response theory. If the time-dependent potential is strong (e.g.: the applied field is strong laser field), the real-time propagation [34] of the system should be used (a full solution of the Kohn-Sham equations) to describe non-linear phenomena. If the time-dependent external potential is weak, as in normal spectroscopic experiments, perturbation theory applies and it is thus sufficient to use to linear-response theory to study the system, through which the electronic excitations can be extracted (thus the optical absorption spectra). The linear-response approach for the time-dependent Kohn-Sham equation is introduced in this section.

Linear-response of the Kohn-Sham system In LR-TDDFT, the linear response of the density to the applied field is a crucial quantity. The linear response of the density of an interacting (real) system (χ) can be derived from that of its corresponding non-interacting (Kohn-Sham) system (χ_s), which form can be known from the general linear-response theory. Therefore the density response related properties can be extracted. This section will show how χ and χ_s are related, which is summarized based on Ref. [8] and Ref. [35].

If a time-dependent field ($v_{appl}(t)$) is applied to an unperturbed interacting system, this will cause a change in the time-dependent density. The linear changes of the density can be found from the interacting point-wise *density-density response function* (χ, also called the *susceptibility*):

$$\delta\rho(\mathbf{r},t) = \int\int \chi(\mathbf{r},\mathbf{r}^{'},t-t^{'})\delta v_{appl}(\mathbf{r}^{'},t)dt^{'}d^{3}\mathbf{r}^{'}, \tag{2.61}$$

which means if a small change in the external potential is made at point $\mathbf{r}^{'}$ and time $t^{'}$, the density-density response function ($\chi(\mathbf{r},\mathbf{r}^{'},t-t^{'})$) tells how the density will change at point $\mathbf{r}$ and later time t.

When the field $v_{appl}(t)$ is applied to the unperturbed Kohn-Sham system which corresponds to the interacting system, it will cause a variation in the SCF potential (sum of Hartree potential and the xc potential), thus the linearized total variation of the Kohn-Sham effective potential is:

$$\delta v_{eff}(\mathbf{r},t) = \delta v_{appl}(\mathbf{r},t) + \delta v_{H}(\mathbf{r},t) + \delta v_{xc}(\mathbf{r},t) \tag{2.62}$$

where

$$\delta v_{H}(\mathbf{r},t) = \int \frac{\delta\rho(\mathbf{r}^{'},t)}{|\mathbf{r}-\mathbf{r}^{'}|}d^{3}\mathbf{r}^{'}, \tag{2.63}$$

and

$$\delta v_{xc}(\mathbf{r},t) = \int f_{xc}(\mathbf{r},\mathbf{r}^{'},t-t^{'})\delta\rho(\mathbf{r}^{'},t^{'})dt^{'}d^{3}\mathbf{r}^{'} \tag{2.64}$$

The time-dependent xc kernal (f_{xc}) is a key quantity of TDDFT in linear-response regime. It is a functional of the correspoding ground-state density,

$$f_{xc}[\rho](\mathbf{r},\mathbf{r}^{'},t-t^{'}) = \frac{\delta v_{xc}[\rho](\mathbf{r},t)}{\delta\rho(\mathbf{r}^{'},t^{'})}. \tag{2.65}$$

Then the linearized density variation of the Kohn-Sham system to the applied field can be calculated via the non-interacting Kohn-Sham response function (χ_s):

$$\delta\rho(\mathbf{r},t) = \int\int \chi_s(\mathbf{r},\mathbf{r}^{'},t-t^{'})\delta v_{eff}(\mathbf{r},t^{'})dt^{'}d^3\mathbf{r}^{'}. \tag{2.66}$$

According to Kohn-Sham assumption, the density response in Eq. 2.61 and Eq. 2.66 should be equal. This will result in an expression connecting χ and χ_s in real time space. For convenience, the equation is transformed to the frequency representation via Fourier transform and is written as:

$$\chi(\mathbf{r},\mathbf{r}^{'},\omega) = \chi_s(\mathbf{r},\mathbf{r}^{'},\omega) + \int\int \chi_s(\mathbf{r},\mathbf{r}_1,\omega)\left(\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} + f_{xc}(\mathbf{r}_1,\mathbf{r}_2,\omega)\right)\chi(\mathbf{r}_2,\mathbf{r}',\omega)d^3\mathbf{r}_1 d^3\mathbf{r}_2, \tag{2.67}$$

where all terms are functionals of the ground-state density. This equation contains the key to electronic excitations via TDDFT. When the energy frequency of the applied field (ω) matches a true transition frequency of the system, the response function χ has a pole as a function of ω.

The frequency-dependent, noninteracting Kohn-Sham density response function is known as:

$$\chi_s(\mathbf{r},\mathbf{r}^{'},\omega) = 2\lim_{\eta\to 0^+}\sum_q\left(\frac{\xi_q(\mathbf{r})\xi_q^*(\mathbf{r}^{'})}{\omega-\omega_q+i\eta} - \frac{\xi_q^*(\mathbf{r})\xi_q(\mathbf{r}^{'})}{\omega+\omega_q-i\eta}\right), \tag{2.68}$$

where q is a double index (a,i), representing a transition from an occupied Kohn-Sham orbital (ϕ_i) to an unoccupied Kohn-Sham orbital (ϕ_a); ω_q is transition energy of this transition which equals the difference of the Kohn-Sham orbital energy,

$$\omega_q = \varepsilon_a - \varepsilon_i, \tag{2.69}$$

and $\xi_q(\mathbf{r})$ denotes

$$\xi_q(\mathbf{r}) = \phi_i^*(\mathbf{r})\phi_a(\mathbf{r}). \tag{2.70}$$

If the Hartree-exchange-correlation effects is absent (when $v_{eff} = v_s$), $\chi = \chi_s$, and so the allowed transitions of χ_s are exactly those of the interacting system. But in reality, the presence of the Hartree and xc kernel shifts the transitions away from the Kohn-Sham values to the true values.

The Casida equation Casida (1995) [29] first derived the essential equations to extract the exact excitation energies of the interacting system from the density-density response function. He reformulated the calculation of the poles (the transition energies) of χ into a generalized Hermitian eigenvalue problem. This approach is nowadays implemented in many electronic-structure codes, and it is particularly popular for calculating optical properties.

Casida showed that, finding the poles of χ is equivalent to solving the psudo-eigenvalue problem:

$$\sum_{q'}\Omega_{qq'}F_q = \omega_{Iq}^2 F_q, \tag{2.71}$$

where q (a,i) and q' (b,j) are double indexes, q denotes an electronic transition from an occupied Kohn-Sham orbital ϕ_i to an unoccupied Kohn-Sham orbital ϕ_a, and q' denotes an electronic

transition from an occupied Kohn-Sham orbital ϕ_j to an unoccupied Kohn-Sham orbital ϕ_b; and the $\Omega_{qq'}$ matrix is defined as

$$\Omega_{qq'}(\omega) = \omega_q^2 \delta_{q,q'} + 4\sqrt{\omega_q \omega_{q'}} K_{q,q'}(\omega). \tag{2.72}$$

$K_{q,q'}(\omega)$ is the frequency-dependent coupling matrix

$$K_{q,q'}(\omega) = \int\int \xi_q^*(\mathbf{r}) f_{Hxc}(\mathbf{r}, \mathbf{r}', \omega) \xi_q(\mathbf{r}') d^3\mathbf{r} d^3\mathbf{r}', \tag{2.73}$$

where $f_{Hxc}(\mathbf{r}, \mathbf{r}', \omega)$ is the response kernel including the Hatree kernal and xc kernel,

$$f_{Hxc}(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}(\mathbf{r}, \mathbf{r}', \omega), \tag{2.74}$$

and the frequency-dependent xc kernel is given by:

$$f_{xc}(\mathbf{r}, \mathbf{r}', \omega) = \int e^{i\omega(t-t')} f_{xc}(\mathbf{r}, \mathbf{r}^{'}, t - t^{'}) d(t - t'). \tag{2.75}$$

The eigenvalues of Eq. 2.72 provide the excitation energies ω_I; and the eigenvectors can be used to obtain the oscillator strengths (see Eq. 1.5) that are related to the frequency-dependent polarizability $\alpha(\omega)$, and to assign the symmetry of each transition.

2.4 Molecular-mechanics methods

2.4.1 Introduction

The quantum-mechanical methods described in preceding sections are very computationally demanding for large systems such as biomolecular systems, solvation systems and material assemblies with many thousands of atoms or more. Molecular-mechanics methods are more computationally efficient and can be used to study larger system with many thousands to millions of atoms.

The computational efficiency of molecular-mechanics methods comes from their two features. Firstly, the potential energy of the system is treated as a parametric function of the nuclear coordinates, which is called the *force field*. Their parameters are defined by fitting to experimental or higher level computational data. This treatment bypasses the costly computation for solving the electronic Schrödinger equation to get the electronic energy of a given nuclear configuration. Secondly, the quantum feature of the nuclear motion are neglected, which means that the dynamics of the atoms is treated by classical mechanics, i.e., Newton's second law (see Eq. 2.1 in Section 2.1). The discussion on the MM methods in this section is based upon the book by Jensen [6].

2.4.2 Force field

In MM methods, each particle is assigned a radius (typically the van der Waals radius), polarizability, and a constant net charge. Bonded interactions are treated as "springs" with an equilibrium distance equal to the experimental or calculated bond length. The force field is used to calculate the molecular system's potential energy E (denoted as V in Newton's equation of motion) in a given conformation as a sum of individual energy terms. The force field energy has the form typically as:

$$E_{FF} = E_{stretch} + E_{bend} + E_{torsion} + E_{vdw} + E_{elec} + E_{cross}. \tag{2.76}$$

Each term corresponds the energy required for distorting a molecule in a specific way. The first three terms describe the bonded atom–atom interactions: $E_{stretch}$ is the energy function for stretching a bond between two atoms, E_{bend} represents the energy required for bending an angle, and $E_{torsion}$ is the torsional energy for rotation around a bond. While E_{vdw} and E_{elec} describe the non-bonded atom–atom interactions. The final term E_{cross} describes coupling between the first three terms.

The exact functional form of the force field depends on the particular simulation program being used. Generally, the $E_{stretch}$ and E_{bend} terms are modeled as harmonic potentials centered (written as a Taylor expansion) around the equilibrium bond-length (values derived from experiment or theoretical calculations). The simplest form of $E_{stretch}$ and E_{bend} can be written as:

$$E_{stretch} = k_r(\Delta r^{AB})^2, \tag{2.77}$$

and

$$E_{bend} = k_\theta(\Delta\theta^{ABC})^2, \tag{2.78}$$

where k_r is the force constant for bond between atoms A and B, Δr^{AB} is the deviation of the distance between atoms A and B from their equilibrium bond length, k_θ is the bond angle formed by three atoms A, B and C , and $\Delta\theta^{ABC}$ is the bond angle deviation. For accurate reproduction of vibrational spectra, the Morse potential can be used for $E_{stretch}$. The torsional energy $E_{torsion}$ typically have multiple minima and thus cannot be modeled as harmonic oscillators and their specific force field form varies with the implementation.

E_{vdw} is the van der Waals energy describing the repulsion or attraction between atoms that are not directly bonded. It is very repulsive (positive) at small distances, has a minimum that is slightly attractive (negative) at a distance corresponding to the two atoms just "touching" each other, and approaches zero when the distance becomes large. It is typically modeled using a *6–12 Lennard-Jones potential*:

$$E_{LJ}(R) = \varepsilon\left((\frac{R_0}{R})^{12} - 2(\frac{R_0}{R})^6\right), \tag{2.79}$$

where R_0 is the minimum energy distance and ε is the depth of the minimum potential well.

The electrostatic terms E_{elec} is the non-bonded interaction due to internal (re)distribution of the electrons that create positive and negative parts of the molecule. At the lowest approximation, this can be modeled by assigning (partial) charges to each atom, or the bond may be assigned a

bond dipole moment. The interaction between point charges is given by the Coulomb potential. The multipole moments interaction and atomic polarization is usually neglected.

The first five terms in the general energy expression in Eq. 2.76 are common to all force fields. The last term E_{cross} is typically a term to take account into the coupling between bond length and bond angle.

2.4.3 Solving the Newton's equation of motion

Once the force field for potential energy is chosen, the Newton's equation of motion is ready to be solved. The potential energy is a function of the atomic positions of all the atoms in the system. There is no analytical solution to the equations of motion due to the complicated nature of the potential function. There are many numerical algorithms developed for integrating this equation. *Verlet algorithm* is one of the frequently used numerical method. It assumes the position, velocity and acceleration can be approximated by a Taylor series expansion with time as the variable. Consider a set of particles described by a position vector r_i at a given time t_i. A small time step Δt later, the positions can be written by

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_i(\Delta t) + \frac{1}{2}\mathbf{a}_i(\Delta t)^2 + \frac{1}{6}\mathbf{b}_i(\Delta t)^3 + \cdots . \tag{2.80}$$

The positions of a small time step Δt earlier were

$$\mathbf{r}_{i-1} = \mathbf{r}_i - \mathbf{v}_i(\Delta t) + \frac{1}{2}\mathbf{a}_i(\Delta t)^2 - \frac{1}{6}\mathbf{b}_i(\Delta t)^3 + \cdots . \tag{2.81}$$

Addition of these two equations gives a way for predicting the positions a time step Δt later from the current and previous positions, and the current acceleration:

$$\mathbf{r}_{i+1} = (2\mathbf{r}_i - \mathbf{r}_{i-1}) + \mathbf{a}_i(\Delta t)^2 + \cdots . \tag{2.82}$$

All odd terms in the Verlet algorithm disappear, i.e., the algorithm is correct to third order in the time step. The acceleration can be calculated from the force or the potential. The time step Δt is an important control parameter for a simulation. The largest value of Δt is determined by the fastest process occurring in the system, typically being an order of magnitude smaller than the fastest process. For simulating nuclear motions, the fastest process is the motion of hydrogen (proton, the lightest nucleus). Hydrogen vibrations occur with a typical frequency of $\sim$1014 s^{-1} (corresponding to 3000 cm^{-1}), and therefore the time step should be chosen with the order of one femtosecond (10^{-15} s).

2.4.4 Application and limitations

There are numerous types of force fields, which are designed for different purposes. MM2 [36] was developed by Allinger, primarily for conformational analysis of hydrocarbons and other

small organic molecules. Consistent force field (CFF) [37] was developed by Warshel, Lifson and their coworkers as a general method for unifying studies of energies, structures and vibration of general molecules and molecular crystals. Empirical conformational energy program for Peptides (ECEPP) was developed specifically for modeling of peptides and proteins. AMBER [38], CHARMM [39] and GROMOS [40] have been developed primarily for molecular dynamics of macromolecules, although they are also commonly applied for energy minimization. OPLS [41] (optimized potential for liquid simulations) was developed by Jorgensen, which was optimized to fit experimental properties of liquids.

The application of MM can be for time-independent or time-dependent phenomena. For the time-independent case, MM is usually used for energy minimization (geometry optimization), that is finding the lowest energy conformation of a molecule or identifying a set of low-energy conformers that are in equilibrium with each other. Determining the conformation associated with the global minimum is often a challenging task (more difficult than finding the local minima) due to the large number of degrees of freedom in macromolecules. The transition state also can be optimized by energy minimization, which is actually reduced to calculating the energy at a given geometry using the force field parameters. Therefore MM can be used to calculate binding constants, protein folding kinetics, etc. For the time-dependent case, MM can be used to simulate physical movements of atoms and molecules (to get their trajectory) and this computation simulation is called molecular dynamics (MD). The application of MD using MM is in several areas. In biochemistry and biophysics, MD is frequently used to refine three-dimensional structures of proteins and other macromolecules based on experimental constraints from X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. In physics, MD is used to examine the dynamics of atomic-level phenomena that cannot be observed directly, such as thin film growth and ion-subplantation. It is also used to examine the physical properties of nanotechnological devices that have not or cannot yet be created.

Although MM (or classical) methods are fast, they suffer from several limitations. They requires extensive parameterization; energy estimates obtained are not very accurate; they cannot be used to simulate reactions where covalent bonds are broken or formed; and they are limited in their abilities for providing accurate details regarding the chemical environment.

2.5 Hybrid quantum-mechanics/molecular-mechanics methods

2.5.1 Introduction

As discussed in the preceding sections, QM method are accurate but expensive thus limiting application to small systems; while MM methods are fast (cheap) but inappropriate for the description of bond-making and bond-breaking. There are cases where our interest is in chemical processes in solution and in proteins, where the environment has important effects thus requiring an explicit representation. This is the motivation for the development of hybrid quantum-mechanics/molecular-mechanics (QM/MM) methods, which combine the strengths of the QM

(accuracy) and MM (speed) methods. The QM/MM approach was firstly introduced in the paper of Warshel and Levitt in 1976. [42] Later, along with M. Karplus, they won the 2013 Nobel Prize in Chemistry for *"the development of multiscale models for complex chemical systems"*. In this section, the basic concept of QM/MM and popular QM/MM approaches are introduced using the book by Cramer [7] as main reference.

The general idea of QM/MM methods is that large chemical systems may be partitioned into two regions: an electronically important region (QM region) that is treated quantum chemically, and the residual part (MM region) which is described with classical methods and acts on the QM region in a perturbative fashion. The Hamiltonian for the entire system is a hybrid of QM and MM methodologies:

$$\hat{H}_{tot} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM}, \tag{2.83}$$

where $\hat{H}_{QM}$ accounts for the full interaction energy of all quantum mechanical particles with one other, $\hat{H}_{MM}$ accounts for the full interaction energy of all classical particles with one other and it doesn't act on the electronic wave function, and $\hat{H}_{QM/MM}$ accounts for the energy of all interactions between one quantum mechanical particle and one classical particle.

The QM/MM methods can be classified by the boundary separating the QM region from the MM region: it cuts the non-bonded QM/MM regions or bonded QM/MM regions.

2.5.2 Non-bonded QM/MM regions

In the cases that the QM/MM boundary cuts no bonds, the coupling of the QM and MM regions can be represented in a straightforward way.

Unpolarized QM/MM interactions

If the polarization effect of the MM region is not considered (i.e., the Hamiltonian $\hat{H}_{QM/MM}$ doesn't act on the electronic wave function), the QM/MM interaction energy is computed in a way closely resembling the standard approach for MM non-bonded interactions. Then the total energy of the system is computed as:

$$E_{tot} = \langle \Psi | \hat{H}_{QM} | \Psi \rangle + E_{MM} + E_{QM/MM}, \tag{2.84}$$

where Ψ is electronic wave function. An example of this type of method is the AM1/OPLS/CM1 (AOC) method for simulations of molecules in solvent developed by Kaminski and Jorgensen in 1998. Since the polarization effect of the MM region is neglected, this approach it is not very successful at predicting solvation effects on these equilibria in non-polar solvents.

Polarized QM/unpolarized MM

To include the polarization effect of the MM region on the QM region, the environment-induced relaxation of the QM wave function should be taken into account explicitly. This can be included by taking account the explicit electrostatic interaction between the QM electrons and the MM atoms. The coupling Hamiltonian can be divided into two parts:

$$\hat{H}_{QM/MM} = \hat{H}^1_{QM/MM} + \hat{H}^2_{QM/MM}, \tag{2.85}$$

where $\hat{H}^1_{QM/MM}$ accounts for the Coulomb interaction between the QM electrons and the MM atoms and it acts on the electronic wave function, and $\hat{H}^2_{QM/MM}$ accounts for the Coulomb and van der Waals interaction between the QM nuclei and the MM atoms and it doesn't act on the electronic wave function. Thus the total energy is written as:

$$E_{tot} = \langle\Psi|\hat{H}_{QM}|\Psi\rangle + E_{MM} + \langle\Psi|\hat{H}^1_{QM/MM}|\Psi\rangle + E^2_{QM/MM}. \tag{2.86}$$

2.5.3 Bonded QM/MM regions

There are cases where a non-bonded separation for QM/MM regions is not practical, for example, the QM region consists of the substrate for a large enzyme and at least one atom from a side chain residue in the active site. Then the coupling term is more complicated. Three popular approaches are introduced here.

Linear combinations of model compounds

In the linear combination of model compounds scheme, the QM region is capped with hydrogen atoms (or bulky groups). The energy of the whole system is calculated as a linear combination of model compounds of different size and at different levels of theory and has the simplest form:

$$E_{tot} = E^{large}_{MM} + (E^{small}_{QM} - E^{small}_{MM}), \tag{2.87}$$

where *small* denotes the QM region that is computed at both the MM and QM levels, *large* denotes the whole system that is usually a macromolecule, which is described at the MM level. In this strategy, the quantum effects are captured in the small system, and the steric energy[1] associated with MM region is captured as an "embedding" energy (the difference between the MM energy of the small system and the large system).

[1] Here the steric energy is not considered as its fundamentally meaning (electronic exchange-repulsion), but as a distinction from more classical electrostatic interactions.

Link atom method

There are situations that the influence of the MM region on the QM region to which it is bonded cannot be regarded simply as steric. For example, in a large protein, polar and possibly charged residues in an MM region will polarize a QM region in the same protein. The link atoms approach might be used to take account of the strong coupling between the two regions. In this approach, the QM region is truncated and capped with hydrogen atoms at every bond cut by the QM/MM boundary. But the QM region and MM region are coupled explicitly via dummy atoms. The total energy evaluation is more complicated, which can be calculated with Eq. 2.86 plus the bonds-cutting energy and bonding energy by the QM/MM boundary. However, the link atoms cause instability in MD simulations at the QM/MM level because of the stiff force constants maintaining linearity of bonds crossing the boundary and the large electrostatic interactions involving atoms near the boundary.

Frozen orbital method

The frozen orbital method is developed to overcome the instability in the link atom approach. A buffer layer is introduced between the polarizable QM region and the point-charge-represented MM region. This buffer layer is called the auxiliary region, which is represented by a continuous unpolarizable charge density and characterized by nuclei having their normal nuclear charges, and electron density expressed in some set of basis functions. The Hamiltonian of the total system can written as

$$\hat{H}_{tot} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} + \hat{H}_{aux} + \hat{H}_{QM/aux} + \hat{H}_{aux/MM}. \tag{2.88}$$

The $\hat{H}_{aux}$ and $\hat{H}_{aux/MM}$ terms are entirely classical classical terms. The former is s simply the electrostatic interaction within the auxiliary region, while the latter is the electrostatic interaction between the auxiliary region (the frozen density and its nuclei) and the MM point charges and non-bonded Lennard-Jones terms between the two regions. The $\hat{H}_{QM/aux}$ term adds two-electron integrals with the orbitals for one electron being frozen to the Fock operator. Therefore compared with Eq. 2.86, the total energy of the system contains three new terms

$$E_{tot} = \langle\Psi|\hat{H}_{QM}|\Psi\rangle + E_{MM} + \langle\Psi|\hat{H}^1_{QM/MM}|\Psi\rangle + E^2_{QM/MM} + E_{aux} + \langle\Psi|\hat{H}_{QM/aux}|\Psi\rangle + E_{aux/MM}. \tag{2.89}$$

The QM/MM methods are powerful for their computing efficiency and availability of quantum-mechanical study for the QM region. However, there are also some challenging problems include: the parameterization required for the boundary treatment, the choice of the size of the QM region, the polarization effect of the QM region on the MM region, etc.

Chapter 3
Computational methods in this thesis work

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong."
— Richard P. Feynman

3.1 Introduction

Frozen-density embedding theory (FDET) combined with LT-TDDFT (see Section 2.3.3.5) is used in this thesis work to study the vertical transitions of organic chromophores in condensed phase. In this chapter, an introduction to FDET for ground-state electronic structure calculation is given firstly. Then it is followed by the extension of FDET to excited state calculations. Finally, the approximations for FDET in practical calculations are introduced.

3.2 Frozen-density embedding theory: for ground states

In FDET, the effect of the environment on the electronic structure of an embedded system is taken into account by an effective embedding potential, which was introduced by Wesolowski and Warshel [43]. This embedding potential depends explicitly on electron densities corresponding to the embedded subsystem and its environment. FDET allows accurate and efficient descriptions of the coupling between the embedded system and its environment. It is in principle exact and does not rely on empirical parametrization that is usually needed in QM/MM methods. Therefore FDET is a proper approach for tackling studies on large systems especially molecules in condensed-phase environment.

3.2.1 Representation of the electron density of the total system

In FDET, the total system is treated as two subsystems: an embedded system that is the part of interest (e.g., a solvated molecule), and the environment (e.g., solvent). In this thesis, the embedded system is referred as subsystem A and the environment is referred as subsystem B. The electron density of the total system (ρ) is represented as the sum of the electron densities of the two subsystems,

$$\rho(\mathbf{r}) = \rho_A(\mathbf{r}) + \rho_B(\mathbf{r}). \tag{3.1}$$

The electron densities of the two subsystems are allowed to overlap. But they are subject to the condition that both integrate to an integer number of electrons, and the sum of these two integer numbers should be equal to the total number of the electrons in the total system. If a N-electron total system is considered, these conditions can be expressed as:

$$\int \rho_A(\mathbf{r}) d\mathbf{r} = N_A \tag{3.2}$$

$$\int \rho_B(\mathbf{r}) d\mathbf{r} = N_B \tag{3.3}$$

$$N_A + N_B = N. \tag{3.4}$$

The electron density of the embedded system ($\rho_A(\mathbf{r})$) is described at quantum-mechanical level. It is typically represented using one the following auxiliary quantities:

- the occupied orbitals of a non-interacting reference system $\{\phi_i^A(\mathbf{r}), i = 1, \cdots, N_A\}$ ($\phi_i^A(\mathbf{r})$ is the one-electron orbital) , [43]
- the occupied and unoccupied orbitals of a non-interacting reference system, [44]
- the interacting wave function, [45]
- the one-particle density matrix. [46]

In this thesis work, only the first two strategies for $\rho_A(\mathbf{r})$ are used.

The electron density of the environment ($\rho_B(\mathbf{r})$) is called *frozen density*, as it is a given fixed function for a given electronic problem.

3.2.2 Total energy functional

FDET concerns minimizing the total energy for the N-electron system in the external potential $v(\mathbf{r})$ in the presence of the constraint $\rho(\mathbf{r}) \geq \rho_B(\mathbf{r})$:

$$E_{emb}[\rho_B] = \min_{\substack{\rho(\mathbf{r}) \geq \rho_B(\mathbf{r}) \geq 0 \\ \int \rho(\mathbf{r}) d\vec{r} = N}} E_v^{HK}[\rho] \tag{3.5}$$

where $E_v^{HK}[\rho]$ is the Hohenberg–Kohn total energy functional (see Eq. 2.24) and ρ_B is a given function.

The Hohenberg–Kohn total energy functional can be expressed as the Kohn-Sham total energy functional ($E_v^{KS}[\rho]$), which can be written as a bi-functional of the densities of the two subsystems [47]:

$$\begin{aligned} E_v^{HK}[\rho] &= E_v^{KS}[\rho] \\ &= E[\rho_A, \rho_B] \\ &= T_s[\rho_A] + T_s[\rho_B] + T_s^{nad}[\rho_A, \rho_B] + V[\rho_A + \rho_B] + J[\rho_A + \rho_B] + E_{xc}[\rho_A + \rho_B], \end{aligned} \tag{3.6}$$

where $V[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d^3\mathbf{r}$ is the interaction energy with an external potential $v(\mathbf{r})$ (in the case of no applied field, V is the nuclear attraction energy), $J[\rho]$ is the Coulomb electron-electron repulsion energy functional defined in Eq. 2.22, and $T_s[\rho]$ is the kinetic energy functional of the non-interacting electron system and is defined as

$$T_s[\rho] = \min_{\Psi_s \to \rho} \langle \Psi_s | \hat{T} | \Psi_s \rangle, \tag{3.7}$$

where Ψ_s are single-determinantal wave functions constructed using $\{\phi_i\}$ yielding the given electron density ρ. $T_s^{nad}[\rho_A, \rho_B]$ is the non-additive kinetic energy defined as:

$$T_s^{nad}[\rho_A, \rho_B] = T_s[\rho_A + \rho_B] - T_s[\rho_A] - T_s[\rho_B]. \tag{3.8}$$

A non-additive xc energy is also defined in a similar way as the non-additive kinetic energy for later use:

$$E_{xc}^{nad}[\rho_A, \rho_B] = E_{xc}[\rho_A + \rho_B] - E_{xc}[\rho_A] - E_{xc}[\rho_B]. \tag{3.9}$$

3.2.3 Kohn-Sham-like equations

For a given frozen electron density ρ_B, the electron density of subsystem A (ρ_A) can be determined by minimizing the total energy bi-functional in Eq. 3.6 with respect to ρ_A, under the constraint described in Eq. 3.5. This constrained search can be performed by solving the Euler-Lagrange equation:

$$\frac{\delta E[\rho_A, \rho_B]}{\delta \rho_A} = \mu. \tag{3.10}$$

If ρ_A determined from this minimization is v-representable, it can be obtained from a set of Kohn-Sham orbitals

$$\rho_A = \sum_i^{N_A} |\phi_i^A|^2, \tag{3.11}$$

where the orbitals ϕ_i^A are determined by a set of Kohn-Sham equations with constrained electron density (KSCED equations),

$$\left(-\frac{\nabla^2}{2} + v_{eff}^{KS}[\rho_A](\mathbf{r}) + v_{emb}[\rho_A, \rho_B; \mathbf{r}] \right) \phi_i^A(\mathbf{r}) = \varepsilon_i^A \phi_i^A(\mathbf{r}), i = 1, \cdots, N_A, \tag{3.12}$$

where $v_{eff}^{KS}[\rho_A](\mathbf{r})$ is the Kohn-Sham effective potential defined in Eq. 2.36, and $v_{emb}[\rho_A, \rho_B; \mathbf{r}]$ is the effective embedding potential describing the interaction between the subsystem A and the environment (representing the effect of the environment on the subsystem A) and it reads

$$v_{emb}[\rho_A, \rho_B; \mathbf{r}] = v_{ext}^B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta T_s^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})} + \frac{\delta E_{xc}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}, \tag{3.13}$$

where $v_{ext}^{B}(\mathbf{r})$ is the potential generated by the nuclei in the environment in the absence of externally applied field. The FDET embedding potential $v_{emb}[\rho_A, \rho_B](\mathbf{r})$ defined in Eq. 3.13 has the following features:

- it is a local potential,
- it is expressed as a functional of $\rho_B(\mathbf{r})$,
- it depends on $\rho_A(\mathbf{r})$.

The last feature is important because it means that the corresponding energy is not a simple integral of the product of the embedding potential and the density. This distinguishes the FDET embedding potential qualitatively from the electrostatic potential generated by the environment that is frequently used as the embedding potential in QM/MM schemes.

3.3 Frozen-density embedding theory: extension to excited states

FDET is combined with LR-TDDFT to evaluate the environment effect on the electronic excitations of the embedded system. The generalization of FDET for excited states within LR-TDDFT framework has been derived by Casida and Wesolowski [48]. In this approach, the embedded orbitals for subsystem A (ϕ_i^A) obtained from Eq. 3.12 are used to construct the response of the system. The total response kernel includes the terms arising from the embedding potential besides the Hartree and xc kernel (see Eq. 2.74) and it reads (in frequency-independent from)

$$f^{total}(\mathbf{r}, \mathbf{r}') = f^{free}(\mathbf{r}, \mathbf{r}') + f^{emb}(\mathbf{r}, \mathbf{r}'), \tag{3.14}$$

where $f^{free}(\mathbf{r}, \mathbf{r}')$ is the response kernel of isolated subsystem A, and $f^{emb}(\mathbf{r}, \mathbf{r}')$ is the embedding kernel. Within the neglect of dynamic response of the environment (NDRE) approximation [44] and the adiabatic approximation in TDDFT, the embedding kernel has the form

$$f^{emb}(\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E_{xc}^{nad}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')} + \frac{\delta^2 T_s^{nad}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')}. \tag{3.15}$$

FDET with NDRE approximation is referred as uncoupled frozen-density embedding (FDE(u)) in Ref. [49]. This NDRE approximation is valid only for the interfaces where excitations are localized only on the embedded system and in the cases where the two subsystems does not absorb in the same spectral range. If NDRE approximation cannot be applied, as in the case of environment and embedded subsystem absorbing at similar frequency, the method based on generalization of ground-state subsystem DFT [48] is indispensable as shown comprehensively in Ref. [49].

If the same approximations are used for $E_{xc}^{nad}[\rho_A, \rho_B]$ and $T_s^{nad}[\rho_A, \rho_B]$ in expressions for the energy, embedding potential, and response kernel, also the excitation energies are self-consistent with other properties of the embedded system.

3.4 Approximations in practical frozen-density embedding calculations

To solve the Kohn-Sham-like equations defined in Eq. 3.12, $\rho_B(\mathbf{r})$ is given as an *input quantity* generated with an adequately chosen method (introduced in detail in Subsection 3.4.2), and the remaining unknown terms should be approximated:

- the xc functional for $v_{eff}^{KS}[\rho_A](\mathbf{r})$,
- the non-additive xc energy functional $E_{xc}^{nad}[\rho_A, \rho_B]$ for $v_{emb}[\rho_A, \rho_B; \mathbf{r}]$,
- the non-additive kinetic energy functional $T_s^{nad}[\rho_A, \rho_B]$ for $v_{emb}[\rho_A, \rho_B; \mathbf{r}]$.

The first two terms above can be approximated with the existing xc functionals. Moreover, it is possible to use different xc approximations for these two terms in practical calculations. The last term can be approximated with several strategies introduced in the following subsection.

3.4.1 Approximating the non-additive kinetic energy bi-functional

3.4.1.1 Thomas-Fermi approximation

In the original work of Wesolowski and Warshel [43], $T_s^{nad}[\rho_A, \rho_B]$ was constructed with several existing kinetic energy functionals for $T_s[\rho]$. A simple strategy is to use Thomas-Fermi expression for kinetic energy functional,

$$T_s[\rho] \approx T_{TF}[\rho] = C_{TF} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r}, \tag{3.16}$$

where the constant $C_{TF} = \frac{3}{10}(3\pi^2)^{2/3}$. The Thomas-Fermi approximation to $T_s^{nad}[\rho_A, \rho_B]$ is written as

$$\begin{aligned} T_s^{nad}[\rho_A, \rho_B] &\approx T_s^{nad(TF)}[\rho_A, \rho_B] \\ &= C_{TF} \int \left((\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))^{5/3} - \rho_A(\mathbf{r})^{5/3} dr - \rho_B(\mathbf{r})^{5/3} \right) d\mathbf{r}, \end{aligned} \tag{3.17}$$

3.4.1.2 Gradient-dependent approximations

The gradient-dependent approximations to $T_s^{nad}[\rho_A, \rho_B]$ are derived from a group of gradient-dependent approximations to $T_s[\rho]$,

$$T_s^{nad}[\rho_A, \rho_B] \approx T_s^{nad(GGA)}[\rho_A, \rho_B] = T_{GGA}[\rho_A + \rho_B] - T_{GGA}[\rho_A] - T_{GGA}[\rho_B]. \tag{3.18}$$

The gradient-dependent approximations to $T_s[\rho]$ ($T_{GGA}[\rho]$) have the general analytic expression

$$T_{GGA}[\rho] = C_{TF} \int \rho(\mathbf{r})^{5/3} F(s) d\mathbf{r}, \tag{3.19}$$

where $F(s)$ is the enhancement factor that depends on a dimensionless quantity s. Various analytic forms of $F(s)$ are proposed. The enhancement factor proposed by Perdew and Wang for the as the exchange energy functional [22] and reparametrized for the kinetic energy by Lembarki and Chermette [50] leads to GGA97 approximation to $T_s^{nad}[\rho_A, \rho_B]$, which was considered the most accurate approximation to $T_s^{nad}[\rho_A, \rho_B]$ for small overlap ρ_A and ρ_B [51].

3.4.1.3 Non-decomposable-second-derivatives approximation

The non-decomposable second derivatives (NDSD) approximation to $\delta T_s^{nad}[\rho_A, \rho_B](\mathbf{r})$ was developed by Wesolowski and coworkers. [52] This approximation was designed to to improve the local behavior of the effective embedding potential near the nuclei in the environment. In its construction, the exact behavior of $\delta T_s^{nad}[\rho_A, \rho_B](\mathbf{r})/\delta\rho_A(\mathbf{r})$ at $\rho_A \to 0$ and the condition $\int \rho_B(\mathbf{r}) d\mathbf{r} = 2$ given in Eq. A6 of Ref. [52] are taken into account. It provides the analytical form of the corresponding approximation to the functional $T_s^{nad}[\rho_A, \rho_B](\mathbf{r})$, which has the form of $T_s^{nad(TF)}[\rho_A, \rho_B]$ plus the corrections (see Eq. 20 in Ref. [52]). The analysis of the accuracy of the embedding potential based on this approximation shows that NDSD is of the same or superior quality as GGA97. [52] However, it is non-decomposable, i.e., it cannot be used to obtain the analytic expression for the kinetic energy functional $T_s[\rho]$; and the numerical values of only first- and second derivatives of density are needed, similarly to the GGA approximations. Its name (non-decomposable second derivatives) reflects these two features.

Thomas-Fermi, GGA97, and NDSD are all semi-local approximations for $T_s^{nad}[\rho_A, \rho_B]$ (and the corresponding $\delta T_s^{nad}[\rho_A, \rho_B](\mathbf{r})/\delta\rho_A(\mathbf{r})$. The potential $\delta T_s^{nad}[\rho_A, \rho_B](\mathbf{r})/\delta\rho_A(\mathbf{r})$ is determined by the densities ρ_A and ρ_B, and their first- and second derivatives at the same point $\mathbf{r}$. Such potentials were shown to fail if the overlap between ρ_A and ρ_B is large [53–55], which are the cases of an embedded subsystem covalently bound to the environment. For such cases, an alternative strategy to approximate $\delta T_s^{nad}[\rho_A, \rho_B](\mathbf{r})/\delta\rho_A(\mathbf{r})$ is becoming increasingly popular which is based on numerical inversion of the Kohn–Sham potential [55–57].

3.4.2 Approximating the electron density of the environment (ρ_B)

In FDET, any observable evaluated for the embedded species is a functional of the frozen density corresponding to the environment (ρ_B). Approximating ρ_B is a key issue for practical applications of FDET, because the choice of the method to generate ρ_B significantly affects the accuracy of the calculations and the computational efforts required. In this subsection, the popular ways of generating ρ_B as well as their strengths and limitations in practical calculations are introduced based on this thesis work.

3.4.2.1 Kohn-Sham calculations for the isolated environment

When it is computationally affordable, it is preferable to obtain ρ_B via solving Kohn-Sham equations for the entire isolated environment (usually less expensive approximations used than those for subsystem A). If the subsystem A is neutral, the polarization of subsystem B by the field generated by subsystem A can be negligible [58]. However, if the subsystem A is charged, the neglect of the polarization of subsystem B by subsystem A may lead to only qualitatively good results (see Ref. [59]). In this case, using a pre-polarized ρ_B [59] or performing the *freeze-and-thaw* calculations [60] can be used to get more accurate results.

In this strategy, mutual polarization of molecules in the environment is taken into account. Ref. [58] shows that ρ_B generated by inexpensive Kohn-Sham-based methods is usually an adequate approximation for hydrogen-bonded environments. However, this strategy becomes impractical when the size of the environment is very large or large amount of FDET calculations are required. In this thesis work, this strategy is used in highly accurate calculations for the electronic excitation of retinal in protein environment.

3.4.2.2 Superposition of molecular densities

If the environment consists of many weakly interacting (non-bonded) molecules, approximating ρ_B by means of a sum of molecular density (electron density of the molecule) is an appealing strategy to save the computational efforts. But this strategy neglects the mutual polarization among the molecules, therefore the accuracy of the obtained results can be affected. The presence of hydrogen-bonded chains of molecules in the environment is the "worst scenario" as far as the applicability of this strategy. In fact, to avoid the apparent errors, one can approximate ρ_B by the sum of the electron densities of molecular clusters. In this way, the bonded molecules can be considered as a molecular cluster to avoid the neglect of the mutual polarization among them. For the constraints in the usage of this strategy, it is less popular than the first one introduced. However, it can be an efficient and practical strategy for qualitative analysis of the electron density of the protein environment.

3.4.2.3 Superposition of atomic densities

For ionic-solid environment where the ions are weakly interacted, the sum of atomic densities (defined below) can be considered as an efficient and reasonable accurate approximation to ρ_B. This strategy can expressed as

$$\rho_B(\mathbf{r}) \approx \sum_{i=1}^{N_{atomB}} (n_B^i - Z_B^i)\rho_B^i(\mathbf{r}), \tag{3.20}$$

where i indicates the atom in subsystem B, N_{atomB} is the number of atoms in subsystem B, $\rho_B^i(\mathbf{r})$ denotes the spherically symmetric electron density integrating to the total charge equal to the atomic number Z_B^i, and n_B^i denotes the net charge of the atom.

When the effect of the environment on the electronic structure of subsystem A is dominated by confinement effect, this strategy can be also used to approximate ρ_B by taking $n_B^i = 0$.

3.4.2.4 Statistically averaged electron density for solvent environment

For solvation system where the solvent environment has a flexible structure, a statistically averaged electron density (denoted as $\langle\rho_B\rangle$ here) of a solvent can be used in FDET to evaluate the average effect of the solvent on molecular properties of the embedded solute. This can be done by replacing ρ_B and v_{ext}^B in FDET equations with $\langle\rho_B\rangle$ and $\langle v_{ext}^B\rangle$ (the ensemble averaged potential generated by the nuclei in the environment), respectively. A straightforward way to obtain $\langle\rho_B\rangle$ and $\langle v_{ext}^B\rangle$ is averaging ρ_B and v_{ext}^B from explicit atomic level simulations, from a MD trajectory representing the statistical ensemble of the system, for instance. $\langle\rho_B\rangle$ and $\langle v_{ext}^B\rangle$ can be also obtained by a procedure based on a classical statistical-mechanics theory of liquids. Kaminski et al. [61] introduced such a procedure where the solvatchromism was studied with the use of $\langle\rho_B\rangle$ and $\langle v_{ext}^B\rangle$ in FDET. In this procedure, the three-dimensional reference interaction site model (3D-RISM) integral equation theory [62] with Hirata-Kovalenko closure [63] is used to obtain the classical site distributions of a molecular solvent around a solute of arbitrary shape, that is, the probability of finding a particular atom in a given volume element. From the force-field parameters for both the solute and the solvent and other primary conditions such as the number density of the solvent, one can obtain solvent site distributions with this theory. Note that 3D-RISM site distributions do not represent any snapshot in a trajectory and they are related to the nuclei rather than the electrons. Kaminski et al. [61] proposed a way to approximate electron density from 3D-RISM site distributions. The key ideas of 3D-RISM theory and the way of approximating $\langle\rho_B\rangle$ from 3D-RISM site distributions are introduced below.

The solute-solvent 3D-RISM integral equation can be derived from the six-dimensional, molecular Ornstein-Zernike integral equation for the solute–solvent correlation functions by averaging out the orientation degrees of freedom of solvent molecules while keeping the orientation of the solute macromolecule that is described at the three-dimensional level. It is written as,

$$h_\gamma^{\mathrm{uv}}(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}' c_\alpha^{\mathrm{uv}}(\mathbf{r}-\mathbf{r}')\chi_{\alpha\gamma}^{\mathrm{vv}}(r)'. \tag{3.21}$$

In Eq. 3.21, $h_\gamma^{\mathrm{uv}}(\mathbf{r})$ is the 3 dimensional total correlation function of solvent site γ around the solute macromolecule, where the index γ denotes the solvent interaction sites around the solute and the superscripts u and v denote the solute and solvent species, respectively. $h_\gamma^{\mathrm{uv}}(\mathbf{r})$ gives the normalized deviation of the solvent density from its bulk value. It is related to the 3 dimensional solute–solvent site distribution function (g_γ^{uv}) as, $g_\gamma^{uv}(\mathbf{r}) = h_\gamma^{uv}(\mathbf{r}) + 1$. $c_\alpha^{\mathrm{uv}}(\mathbf{r})$ is the 3 dimensional direct correlation function representing "direct" correlations between the solute and solvent site α. It has the long-range asymptotic behavior of the 3 dimensional solute-solvent site interaction potential (u^{uv}), $c_\gamma^{\mathrm{uv}}(\mathbf{r}) \sim -\beta u_\gamma^{\mathrm{uv}}(\mathbf{r})$, where $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature with the Boltzmann constant k_{B}. $\chi_{\alpha\gamma}^{\mathrm{vv}}$ is the site-site susceptibility of pure solvent. It gives the response of site γ to the presence of site α at separation r in terms of the distributions in pure bulk solvent, $\chi_{\alpha\gamma}^{\mathrm{vv}}(r) = \omega_{\alpha\gamma}^{\mathrm{vv}}(r) + \rho_\alpha^{\mathrm{v}} h_{\alpha\gamma}^{\mathrm{vv}}(r)$. $\omega_{\alpha\gamma}^{\mathrm{vv}}(r)$ is the intramolecular matrix that specify the intramolecular correlations of solvent molecules. ρ_α^{v} is the bulk solvent site number density. $h_{\alpha\gamma}^{\mathrm{vv}}(r)$ is the site-

site radial correlation functions of pure bulk solvent and it is obtained in advance from the dielectrically consistent reference interaction site model (DRISM) theory developed by Perkyns and Pettitt [64] that provides a consistent description of the dielectric properties for ions in polar solvent.

To solve the 3D-RISM integral equation denoted in Eq. 3.21, a closure for relating $h_\gamma^{\mathrm{uv}}(\mathbf{r})$ and $c_\gamma^{\mathrm{uv}}(\mathbf{r})$ is needed. In the work of Kaminski et al. [61], the closure developed by Kovalenko and Hirata is used. It is written as

$$\begin{aligned} g_\gamma^{\mathrm{uv}}(\mathbf{r}) &= \begin{cases} \exp\Big(d_\gamma^{\mathrm{uv}}(\mathbf{r})\Big), & \text{for } d_\gamma^{\mathrm{uv}}(\mathbf{r}) \leq 0 \\ 1 + d_\gamma^{\mathrm{uv}}(\mathbf{r}), & \text{for } d_\gamma^{\mathrm{uv}}(\mathbf{r}) > 0 \end{cases}; \\ d_\gamma^{\mathrm{uv}}(\mathbf{r}) &= -\beta u_\gamma^{\mathrm{uv}}(\mathbf{r}) + h_\gamma^{\mathrm{uv}}(\mathbf{r}) - c_\gamma^{\mathrm{uv}}(\mathbf{r}). \end{aligned} \tag{3.22}$$

To obtain the average electron density of the solvent, we need a second step called "dressing up" the site distributions with electrons. $\langle \rho_B(\mathbf{r}) \rangle$ is approximated as a sum of spherically symmetric electron density of each type of atom in the solvent system,

$$\langle \rho_B(\mathbf{r}) \rangle = \sum_\gamma g_\gamma^{uv}(\mathbf{r}) q_\gamma^v \rho_\gamma^v, \tag{3.23}$$

where $g_\gamma^{uv}(\mathbf{r})$ is the 3D solute-solvent site distribution function; q_γ^v denotes the spherically symmetric electron density of an atom in the solvent; and ρ_γ^v is the bulk solvent site number density.

The strategy using $\langle \rho_B \rangle$ in FDET for solvation systems takes account into specific solvent-solute interactions in a statistical manner. It has been proved as a rather accurate and efficient way to study solvatochromism. [61] However, it introduces additional computational difficulties by requiring force-field parameters for both the solvent and solute species in solving the 3D-RISM integral equations.

Chapter 4
Absorption band shapes of fluorenone in zeolite-L channel

"Everybody is ignorant, only on different subjects."
— Will Rogers

4.1 Overall presentation of the article

Zeolite L has arrays of parallel channels with small pore openings and cage diameters imposing severe space restrictions and geometrical constraints to the molecules in the channel. As a result, very high concentrations of well-oriented photoactive molecules can be obtained and the entire system may exhibit peculiar properties (described in the introduction of the original paper). Therefore it is very interesting to study the zeolite-L environment effect on the properties of the guest molecules in its channel.

The aims of this work includes: (1) developing the computational strategy for simulating the UV-vis absorption bands of fluorenone in zeolite-L channel, using the measured absorption bands as references; (2) analyzing the different components of the environment effects on the absorption bands of fluorenone via the analysis for the simulated UV-vis absorption spectra; (3) confirming the orientation of fluorenone in the zelite-L channel via the analysis of the simulated UV-vis absorption spectra.

Through this work, the zeolite-L and water molecules in the channel were found causing a redshift in the lowest absorption band of fluorenone. This effect may originated from the changes in the internal geometry of fluoreone and the orientation of fluoreone with respect to the channel, and the electrostatic interaction between the zeolite-L framework and the fluorenone molecule. Moreover, the fluorenone dye is confirmed to orientate with its long axis parallel to the channel of the zeolite-L channel.

As mentioned in the introduction of the thesis (Chapter 1), the simulation of UV-vis absorption spectrum usually involves two steps: modeling the structure of the target system, and calculating the excitation energy and oscillator strength of the vertical transitions where the measured spectra are often taken as references. This project was completed in a collaborative work. The finite structures (MD trajectories) of the fluorenone-zeolite-L system were calculated by Gloria Tabacchi and Ettore Fois. The referent measured spectra were obtained by André Devaux and Gion Calzaferri. The calculation of the vertical electronic transitions and the simulation of the absorption band shape were performed by Xiuwen Zhou and Tomasz A. Wesolowski.

4.2 Reprint of the article

PCCP

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

First-principles simulation of the absorption bands of fluorenone in zeolite L†

Cite this: *Phys. Chem. Chem. Phys.*, 2013, **15**, 159

Xiuwen Zhou,[a] Tomasz A. Wesolowski,*[a] Gloria Tabacchi,[b] Ettore Fois,*[b] Gion Calzaferri*[c] and André Devaux[d]

The absorption spectrum of fluorenone in zeolite L is calculated from first-principles simulations. The broadening of each band is obtained from the explicit treatment of the interactions between the chromophore and its environment in the statistical ensemble. The comparison between the simulated and measured spectra reveals the main factors affecting the spectrum of the chromophore in hydrated zeolite L. Whereas each distinguishable band is found to originate from a single electronic transition, the bandwidth is determined by the statistical nature of the environment of the fluorenone molecule. The $K^+\cdots O{=}C$ motif is retained in all conformations. Although the interactions between K^+ and the fluorenone carbonyl group result in an average lengthening of the C=O bond and in a redshift of the lowest energy absorption band compared to gas phase or non-polar solvents, the magnitude of this shift is noticeably smaller than the total shift. An important factor affecting the shape of the band is fluorenone's orientation, which is strongly affected by the presence of water. The effect of direct interactions between fluorenone and water is, however, negligible.

Received 6th August 2012,
Accepted 16th October 2012

DOI: 10.1039/c2cp42750h

www.rsc.org/pccp

1 Introduction

The one-dimensional channel framework of zeolite L,[1,2] acting as a host for a large variety of guests, has encouraged syntheses of guest–host composites with remarkable organizational patterns.[3–18] These are currently investigated for applications in different fields, ranging from sensing in analytical chemistry, biology, and diagnostics,[15–18] to applications in optical and electro optical devices[3–5,19–21] and extending to solar energy utilization, *e.g.* in the form of luminescent solar concentrators.[3,4,22–25] Advanced optical microscopy techniques have enabled researchers to obtain detailed information regarding the properties of the guests.[11,26–31]

However, modeling studies needed for interpreting and understanding the experimental data remain challenging, mainly because of the considerable extension of the systems to be handled in the calculations. We have, nevertheless, recently been successful in explaining experimental observations by performing first-principles calculations on systems consisting of hundreds of atoms.[32–34]

Important organic molecules used in guest–zeolite-L composites bear carbonyl groups. Fluorenone was found to be an excellent probe molecule for studying the behavior and interactions of the carbonyl group of such guests inside the channels of zeolite L, both experimentally and theoretically.[33,34] The previously reported computer simulation studies of fluorenone inside dry and hydrated zeolite L channels focused on the structure and interactions of fluorenone with the channels and the charge compensating cations as well as with the water molecules present as co-guests.[32] It was comprehensively shown that the water molecules were incapable of displacing the fluorenone out of the zeolite L channel because of the strong interactions between fluorenone's carbonyl group and the extra-framework potassium cations. The $K^+\cdots O{=}C$ arrangement was found to be the leitmotif determining the behavior of fluorenone in the channel. Without breaking this arrangement, the fluorenone molecule can, however, adopt various orientations owing to the sufficiently large diameter of the channel. The *ab initio* molecular dynamics simulations revealed that the reorientation time of fluorenone in hydrated zeolite L is about twice as long as that in dry zeolite L. Turning to the absorption spectra of fluorenone, the bands in the 200–420 nm range observed in gas phase and in non-polar media are rather well understood. They all originate from four electronic transitions of $\pi \rightarrow \pi^*$ character.[35] This interpretation was confirmed by Fois *et al.*[32] The shape of the electronic

[a] *Département de Chimie Physique, Université de Genève, 30 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland. E-mail: tomasz.wesolowski@unige.ch*
[b] *Department of Science and High Technology, University of Insubria, and INSTM, Via Lucini 3, I-22100 Como, Italy. E-mail: fois@fis.unico.it*
[c] *Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland. E-mail: gion.calzaferri@iac.unibe.ch*
[d] *Institute of Inorganic Chemistry, University of Fribourg, CH-1700 Fribourg, Switzerland*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2cp42750h

absorption spectrum of fluorenone in zeolite L and in non-polar solvent (cyclohexane) is similar (see Fig. 5 in ref. 33 and the data shown in the present work). However, the former is less structured and the maxima of the bands are red-shifted. It is tempting, therefore, to assign the observed bands of fluorenone in zeolite L to the same electronic transitions as the ones in the gas phase and in cyclohexane. The disappearance of the vibrational features can be due to the structural fluctuations of the environment of fluorenone in zeolite L. In contrast to cyclohexane, the zeolite environment includes components strongly interacting with the chromophore: water molecules, extra framework cations, and the zeolite framework. The shifts in the maxima of the absorption bands could be attributed to the average effect of the interactions between fluorenone and charged atoms (potassium cations) or polar molecules (water) as well as to the effect of long-range interactions between the fluorenone and the zeolite framework. The overall effect of the environment on the absorption spectrum can be considered as resulting from two factors: one being *electronic* and the other *structural*. Interactions with the environment affect fluorenone's electronic structure at any instantaneous geometry. On the other hand, these interactions determine both the average internal geometry of the chromophore and its relative orientation in the zeolite channel in a given statistical ensemble. The orientation of fluorenone is especially interesting as it is directly related to the properties of fluorenone as a component of the energy relay system in artificial antennas. Discriminating the direct effects induced by the presence of water molecules in the zeolite channel, which are due to the potential they generated at a given geometry, from the indirect ones such as their impact on the orientation of fluorenone in the channel, is crucial for understanding the experimental data concerning the electronic properties and excited state dynamics of the composite as well as the whole energy relay mechanism in artificial antennae systems based on zeolite L. Experimental estimation of the importance of these two types of effects is rather difficult.[3,4] However, availability of an adequate sample of instantaneous geometries generated in numerical simulations for the corresponding statistical ensembles makes it possible to assess directly the relative importance of these two effects. The electronic effect due to the water molecules can be easily monitored by means of comparing the spectrum evaluated from geometries corresponding to the hydrated zeolite L with the one evaluated from the same set of instantaneous geometries but with the water molecules removed. Their structural effect on the other hand, can be also easily monitored once two samples of instantaneous geometries are available: one for the hydrated and the other for the dry composite. The difference between the spectra evaluated without water but for instantaneous geometries corresponding to these two different statistical ensembles provides a direct measure of the structural effect.

In ref. 32, excitation energies were calculated on molecular models (clusters) extracted from the minimum energy structures of the fluorenone–zeolite system containing fluorenone and the nearest atoms of the zeolite L environment. Due to the statistical nature of the system, we extend these studies of electronic excitations for embedded fluorenone by simulating the theoretical spectra for a model consisting of more than one thousand of atoms in which the structural fluctuations of both the chromophore and the environment are taken into account. The principal objective of the present work is a qualitative and quantitative analysis of the relative significance of the electronic and structural effects of the environment on the absorption bands of fluorenone inside the hydrated zeolite. To this aim, the theoretical absorption spectrum of fluorenone in a hydrated zeolite is obtained as a first step. Subsequently, the relative importance of structural and electronic factors is investigated by selectively switching off and on different effects. The observed broadening of the absorption bands and the ultimate shape of the absorption spectrum result mainly from vibrational effects, *i.e.*, fluctuations of the environment and of the internal geometry of the chromophore. Computer simulation of the band shapes involves, therefore, a large number of evaluations of excitation energies for a relatively large system. In this case, the fluctuation of the absorption energies of the chromophore should be mainly related to its instantaneous environment (zeolite framework, water molecules, cations) which does not absorb light in the investigated spectral range. The embedding strategy is, therefore, an optimal way to evaluate reliably the excitation energies for the corresponding statistical ensemble. In this strategy, the quantum mechanical descriptors are used only for the chromophore, whereas the effect of the environment is taken into account by an additional operator. According to the basic result of Frozen-Density Embedding Theory (FDET),[36–40] such an operator has the form of a local potential, which is uniquely determined by the charge distribution in the environment and the electron density of the chromophore. Moreover, recent benchmarking studies demonstrated that the precision of the environment-induced shift of the electronic excitation energy obtained from FDET was remarkable for environments not interacting covalently with the chromophore. The differences between the environment-induced shifts of the excitation energies and their high-quality reference counter-parts are typically in the 150 cm^{-1} range, corresponding to 0.02 eV.[41,42] Such accuracy of the shift values is expected to be retained for the case studied here as the overlap between the electron densities of the chromophore and its environment is rather small. The secondary objective of the present work is to assess the usefulness of the used computational protocol based on FDET as a first-principles method to simulate the *shape* of the absorption bands. Currently, in order to simulate the absorption spectra of organic chromophores in condensed phases, most of the methods obtain the excitation energies by first-principles based calculations combined with empirical widening of each band.[43] In the present work, the inhomogeneous widening of the bands is also obtained from first principles.

2 Computational details

2.1 First-principles molecular dynamics

Density functional theory (DFT) calculations on the dry fluorenone–zeolite-L composite (simulation cell stoichiometry: $K_{18}[Al_{18}Si_{54}O_{144}]Fl$) and on a hydrated system modeling low water loading conditions (simulation cell stoichiometry: $K_{18}[Al_{18}Si_{54}O_{144}]Fl\cdot(H_2O)_4$) were performed by adopting the PBE approximation[44] and periodic boundary conditions, as described

in detail in ref. 32. The simulation cell was twice the hexagonal experimental unit cell of the zeolite host along c($a = b = 18.466$ Å; $c = 2 \times 7.476$ Å; $\beta = 120^{\circ}$).[45] Wavefunctions were expanded in planewaves up to a 25 Ry cutoff (200 Ry for the density).[46–48] Electron–ion cores interactions were calculated with ultra-soft Vanderbilt pseudopotentials for H, C, O; norm-conserving pseudopotentials for Si, Al, K (semi-core).[49–52] Such a calculation scheme for the electronic structure provided a proper description of other large organic–inorganic systems.[53–60]

First-principles molecular dynamics (FPMD)[61–63] simulations were performed on both hydrated and dry systems. The room temperature FPMD trajectory of the dry composite was described in ref. 32. In the case of the hydrated system, two trajectories were performed. One (12 ps) was at room temperature, and the fluorenone molecular axis was roughly parallel to the channel axis as described in ref. 32. The other (5 ps) was performed at a 40 K temperature starting with fluorenone oriented approximately perpendicular to the channel axis, in order to sample different fluorenone orientations in the zeolite and to investigate their effect on the electronic spectrum. The target temperature of 40 K (NVT ensemble) was chosen because the perpendicular orientation of fluorenone is not stable under room temperature conditions. Actually, in a 300 K simulation starting with perpendicular fluorenone, the molecule was observed to recover in one picosecond to its energetically favored orientation in the zeolite channel, *i.e.* approximately parallel to the zeolite channel axis. As in ref. 32, a time step of 0.121 fs and an inertia parameter of 500 au for the electronic states coefficients were adopted.

2.2 Excitation energies of embedded fluorenone from Frozen-Density Embedding Theory based calculations

One of the clusters used for calculating the excitation energies of fluorenone in the hydrated zeolite taken from the trajectory representing the statistical ensemble is shown in Fig. 1. The same cluster size was used to calculate the excitation energies for fluorenone in the dry zeolite. For each instantaneous geometry of the system, the excitation energy of fluorenone was calculated using the method[64] combining FDET[36–40] with linear-response time-dependent DFT framework.[78] This FDET framework was implemented[38,65] into the ADF code (2009.01 version)[67] allowing for evaluation of various properties of embedded systems and making possible to introduce additional approximations for the frozen density as the ones discussed in ref. 38, 66, 74 for instance. Fluorenone was treated as an embedded system, described by electron density $\rho_A(\vec{r})$ and embedded orbitals, while the rest of the system (zeolite L framework, cations and water molecules) was treated as an embedding system, described by its electron density $\rho_B(\vec{r})$.

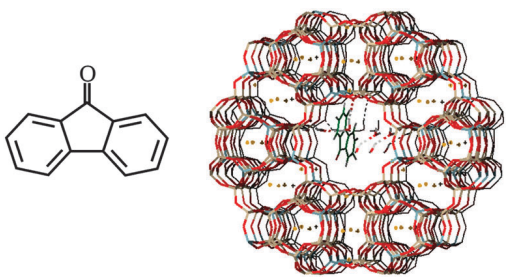


Fig. 1 Left panel: the fluorenone molecule. Right panel: the cluster used for calculating excitation energies of fluorenone in the hydrated zeolite: $K_{71}[Al_{108}Si_{324}O_{819}]Fl\cdot(H_2O)_{12}$ (1380 atoms in total). Color code: red, O; white, H; light blue, Al; grey, Si; yellow, K^+; green, C.

In the excitation energy calculation, the frequency-dependent response was expanded using only embedded orbitals associated with the chromophore (neglect of the dynamic response of the environment, NDRE, called also uncoupled FDE in the literature[68]) as described in ref. 38. The STO-type DZP basis set was applied for all the calculations. The Statistical Averaged Orbital Potential (SAOP)[69–71] model was used to approximate the exchange–correlation potential evaluated for the density corresponding to fluorenone (embedded or isolated). For the orbital-free embedding potential (eqn (3) in ref. 72):

$$v^{\mathrm{eff}}_{\mathrm{emb}}[\rho_A,\rho_B;\vec{r}] = v^{B}_{\mathrm{ext}}(\vec{r}) + \int \frac{\rho_B(\vec{r})}{|\vec{r}-\vec{r}|}d\vec{r} + \frac{\delta T_s^{\mathrm{nad}}[\rho_A,\rho_B]}{\delta\rho_A(\vec{r})} + \frac{\delta E_{xc}^{\mathrm{nad}}[\rho_A,\rho_B]}{\delta\rho_A(\vec{r})} \quad (1)$$

the following approximations were used:

-NDSD approximation[73] for the non-additive kinetic energy dependent component of the embedding potential;

-Dirac- and Vosko–Wilk–Nusair expressions were used for the non-additive exchange energy and correlation energy, respectively.

-The choice of ρ_B used in eqn (1) is a key approximation in any FDET based simulation and lies at the origin of great computational savings compared to the corresponding Kohn–Sham calculations for the whole system. The electronic density of the environment (zeolite L, cations, and water) ρ_B was approximated by the superposition of spherically symmetric electron densities of individual atoms each integrating to the following net charge: silicon +4e, aluminum +3e, oxygen (zeolite) −2e, oxygen (water) −0.8e, potassium +1e, and hydrogen (water) +0.4e. The net charge of each spherically symmetric atomic density corresponds to the formal ionic charges of the zeolitic material. The applied procedure to generate ρ_B was validated by performing the following test. For a smaller cluster comprising only the nearest neighbors of fluorenone ($Si_{50}Al_{16}K_{16}O_{163}H_{62}$, 307 atoms in total), two sets of ρ_B were generated. The first one was obtained using the above procedure (consisting of superposing spherically symmetric atomic densities) and the other one was obtained from conventional Kohn–Sham calculations. In the subsequent step, the two frozen densities were used in FDET based calculations to evaluate the excitation energies for fluorenone embedded in the small cluster. A graphical representation of the small cluster and the calculated values of the excitation energies are reported in the ESI.† The excitation energies obtained with the two sets of frozen densities (ρ_B) agree within about 0.03 eV, validating thus the chosen strategy to generate ρ_B in the larger system simulation.

2.3 Simulation of the absorption spectra

The simulated spectrum was generated through averaging over several hundred configurations representing the corresponding

statistical ensemble. For each configuration, a "stick spectrum" was obtained from LR-TDDFT calculations providing vertical excitation energies and intensities. Each simulated spectrum reported in this work is just the envelope of a large number of superimposed "stick spectra". In this way, the inhomogeneous broadening of each absorption band due to variations of both internal degrees of freedom and intermolecular ones was taken into account. This approach is essentially the same as the one used in ref. 66, which focused on solvatochromism. In the present work, we attempt to simulate not only the solvatochromic shifts of the first excited states but rather the whole shape of the absorption spectrum in the UV/Vis range.

For a given energy point ε, the averaged oscillator strength $f_{\mathrm{av}}(\varepsilon)$ was calculated by summing the oscillator strengths of the excitations for each instantaneous geometry:

$$f_{\mathrm{av}}(\varepsilon) = A \sum_{j}^{N_{\mathrm{config}}} \sum_{i}^{N_{\mathrm{excit}}} f(\varepsilon_{ij})\chi(\varepsilon_{ij}), \qquad (2)$$

where A is the normalization factor; i is the index of the excitations (N_{excit} = 10); j is the index of the configurations; N_{config} is the total number of the instantaneous configurations in the sample, and $\chi(\varepsilon_{ij})$ is the indicator function defined as

$$\chi(\varepsilon_{ij}) = \begin{cases} 1 & \text{if } \varepsilon_{ij} \in \left(\varepsilon - \frac{1}{2}a, \varepsilon + \frac{1}{2}a\right) \\ 0 & \text{if } \varepsilon_{ij} \notin \left(\varepsilon - \frac{1}{2}a, \varepsilon + \frac{1}{2}a\right) \end{cases}$$

The summation range a was set to be 0.08 eV. The simulated spectra considered in the present work were obtained for N_{config} in the range of 600. Numerical experience with the considered system shows that an ensemble of such size obtained from the equilibrated molecular dynamics trajectory assures a stable shape of the simulated spectrum. Spectra obtained using trajectories with different lengths are provided in ESI.† For samples of this size, the chosen value of a = 0.08 eV ensures smoothness of the spectrum without losing finer features of the bands. Varying the value of the parameter a from 0.06 to 0.14 eV marginally changes the overall shape of the spectrum (see ESI†). This procedure corresponds to finding the envelope of a stick spectrum such as the one shown in Fig. 8 in the Results section. We notice that eqn (2) corresponds to essentially the same procedure to simulate spectra as the one applied in ref. 75, in which the parameter a was set to be 0.01 eV for studies of gas and condensed phase spectra of amides. The narrower width of the energy bin used in ref. 75 was needed because gas-phase spectra of amides are much more structured than the spectra considered in the present work. In this context, it is worthwhile pointing out that the used value of the parameter a exceeds the natural width of the absorption band due to the finite lifetime of the excited state. The shape function for each line in the "stick spectrum" is, therefore, not considered explicitly in such averaging procedures.

In our strategy for simulating the shape of the absorption bands, the excitation energy of the embedded species is obtained as the sum of the excitation energy of the isolated chromophore and of the environment induced shift. For fluorenone embedded in the environment at a given instantaneous geometry $\{\bar{R}_i\}$, its excitation energy is obtained as:

$$\varepsilon_{\mathrm{emb}}^{\mathrm{calc}}(\{\bar{R}_i\}) = \varepsilon_{\mathrm{iso}}^{\mathrm{calc}}(\{\bar{R}_i\}) + \Delta\varepsilon^{\mathrm{FDET}}(\{\bar{R}_i\}) \qquad (3)$$

The accuracy of each of these two components is determined by different factors. Concerning the term $\varepsilon_{\mathrm{iso}}^{\mathrm{calc}}(\{\bar{R}_i\})$, LR-TDDFT excitation energies for small organic chromophores are known to lead to errors in the range of 0.1–0.5 eV if semilocal approximations are used for the exchange–correlation potential.[76,77] As far as $\Delta\varepsilon^{\mathrm{FDET}}(\{\bar{R}_i\})$ is concerned, the benchmark calculations show that the errors of the FDET calculated environment induced shifts are about 0.02 eV (the relative errors of the FDET shifts lie around 10%).[41] It can be expected, therefore, that the shape of the absorption bands of fluorenone in a fluctuating environment can be described rather accurately by means of FDET shifts evaluated for a large number of configurations representing the considered statistical ensemble, whereas the position of the center of the bands depends on the method used to evaluate $\varepsilon_{\mathrm{iso}}^{\mathrm{calc}}(\{\bar{R}_i\})$. Using LR-TDDFT/FDET excitation energies to simulate the absorption spectra by means of eqn (3) involves also another approximation, namely the neglect of the difference between the position of the maximum of the absorption band and the corresponding vertical excitation energy. Obtaining the numerical value of such a difference from the theoretical model exceeds the scope of the present work. By taking into account these factors affecting the accuracy of the simulated band shapes, the values of $\varepsilon_{\mathrm{emb}}^{\mathrm{calc}}(\{\bar{R}_i\})$ needed in eqn (2) were evaluated not by using eqn (3) but by adopting its empirically corrected formulation:

$$\varepsilon_{\mathrm{emb}}^{\mathrm{calc}}(\{\bar{R}_i\}) = \varepsilon_{\mathrm{iso}}^{\mathrm{SAOP}}(\{\bar{R}_i\}) + \Delta\varepsilon_{\mathrm{iso}} + \Delta\varepsilon^{\mathrm{FDET}}(\{\bar{R}_i\}) \qquad (4)$$

The constant shift $\Delta\varepsilon_{\mathrm{iso}}$ = 0.8 eV was used for all the simulated spectra presented in this work. This value takes into account both inaccuracies of the LR-TDDFT/SAOP excitation energies of fluorenone in the gas phase and the difference between the true vertical excitation energy and the maximum of the corresponding absorption band. Such a value was chosen to reproduce the experimental position of the maximum of the lowest energy band of fluorenone in hydrated zeolite L.

2.4 Excitation energies of fluorenone in gas phase

The geometry of fluorenone in the gas phase was optimized at the MP2/6-311++g(d,p) level. Excitation energies were calculated using the LR-TDDFT methodology[78] with a STO-type DZP basis set and the SAOP[69–71] exchange–correlation potential. Note that all the parameters in the reported calculations were set in such a way that the gas-phase excitation energies from LR-TDDFT calculations are the *same* as the FDET/LRTDDFT energies obtained when the embedding potential in eqn (1) is set to zero.

3 Results and discussions

The Results section is organized as follows. The characterization of the excited states of isolated fluorenone in its equilibrium geometry is followed by the analysis of the simulated absorption spectrum. Different factors which may affect the shape of the absorption bands of fluorenone in hydrated zeolite L are discussed in the

subsequent sections. In particular, the effect of water molecules on the calculated absorption spectrum of the composite is analysed in detail by decomposing it into a component due to the fluorenone–water direct interactions at each instantaneous geometry and a component due to the effect of water on fluorenone's orientation in the channel. Additionally, the effect of fluctuations in the internal geometry of fluorenone due to interactions with its environment is analyzed in detail.

3.1 Electronic transitions for isolated fluorenone in its equilibrium geometry

The results of LR-TDDFT calculations for isolated fluorenone in its equilibrium geometry and the corresponding results reported in the literature are collected in Table 1. The $n \rightarrow \pi^*$ transitions are symmetry forbidden and the observed bands correspond to $\pi \rightarrow \pi^*$ transitions. The energies of the allowed transitions are systematically lower than their LRTDDFT/B3LYP and PPP counterparts. Such an underestimation is in line with the literature that the SAOP approximation underestimates the energy of the lowest transitions in organic chromophores.[77] For this reason, the LR-TDDFT and semiempirical PPP results are expected to be better estimates of the exact excitation energies of the isolated fluorenone. Compared to PPP excitation energies, which are most likely the best estimate, the LR-TDDFT/SAOP excitation energies are underestimated by 0.3 to 0.58 eV (see Table 1).

3.2 Absorption spectra of fluorenone in hydrated zeolite L

Fig. 2 shows the absorption spectra of fluorenone measured in cyclohexane and in hydrated zeolite L as reported originally in ref. 33. Due to the small values of the molecular extinction coefficient at long wavelengths, the intensities are shown on the logarithmic scale in all the figures subsequently discussed in this work.

The simulated spectrum of fluorenone in the hydrated zeolite together with its experimental counterpart is shown in Fig. 3.

The simulated and measured spectra of fluorenone in hydrated zeolite L feature four intense bands in the low-energy part of the spectrum. The shapes of the two spectra match quite well. The position of the lowest energy band was built into our model to simulate spectra by means of the chosen value of $\Delta\varepsilon_{iso}$ in eqn (4). The correct positions and relative intensities of the three subsequent clearly distinguishable bands indicate the strength of the used strategy to simulate absorption band shapes using the FDET embedding potential given in eqn (1) in LR-TDDFT calculations of excitation energies combined with the non-empirical treatment of line broadening (eqn (2)).

The non-shifted (eqn (3)) excitation energies for embedded fluorenone are collected in Table 2. The analysis of electronic excitations in the instantaneous geometries shows that each electronic transition generates only one peak in the spectrum. The width of each band results from the different excitation energy values due to the fluctuating environment of fluorenone in hydrated zeolite L. Table 2 collects also the range at which each type of excitation occurs in the analyzed sample of conformations.

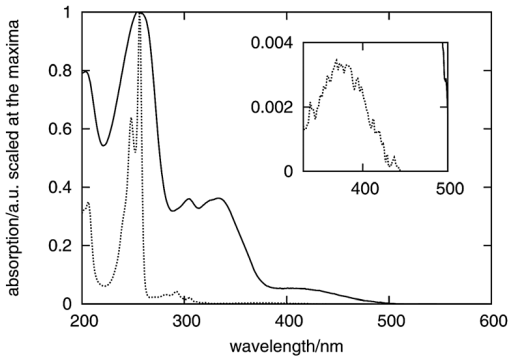


Fig. 2 Absorption spectra of fluorenone measured in cyclohexane (dashed curve) and in hydrated zeolite L (solid curve). The extinction coefficients are scaled to 1 at the absorption maximum. The inset shows an enlargement of the weak band at 375 nm in cyclohexane.

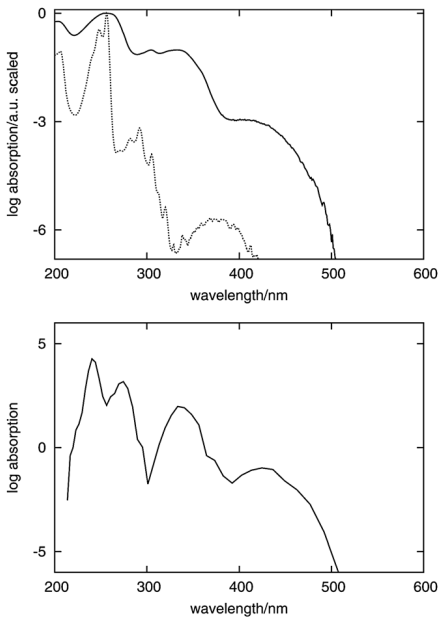


Fig. 3 Top: experimental spectra of fluorenone in cyclohexane solution (dashed curve) and in hydrated zeolite L (solid curve). Bottom: simulated spectrum of fluorenone in hydrated zeolite L. The intensity is given on the logarithmic scale in both measured and simulated spectra to visualize the weak bands.

Table 1 Calculated excitation energies (ε, in [eV]) and oscillator strengths (f) of isolated fluorenone. The excitation energies of the corresponding transitions reported in the literature are also given

Type of excitation	ε^a	f^a	ε^b	ε^c
$\pi \rightarrow \pi^*$	2.75	0.29×10^{-2}	3.09	3.26
$n \rightarrow \pi^*$	3.04	0.35×10^{-9}	3.13	—
$2\pi \rightarrow \pi^*$	3.56	0.31×10^{-1}	4.00	4.00
$3\pi \rightarrow \pi^*$	3.98	0.19×10^{-1}	4.44	4.28
$n \rightarrow 2\pi^*$	4.06	0.74×10^{-10}	4.82	—
$4\pi \rightarrow \pi^*$	4.38	0.23×10^{-1}	5.07	4.96
$\pi \rightarrow 2\pi^*$	4.62	0.69	4.89	—
$3\pi \rightarrow 2\pi^*$	4.78	0.17×10^{-2}	—	—
$2\pi \rightarrow 2\pi^*$	4.90	0.18×10^{-1}	—	—
$n \rightarrow 3\pi^*$	4.95	0.26×10^{-3}	—	—

[a] This work (LR-TDDFT/SAOP). [b] Ref. 32 (LR-TDDFT/B3LYP). [c] Ref. 35 (PPP semi-empirical calculations).

Table 2 Excitation energies (in [eV]) calculated from eqn (3) for the lowest symmetry allowed transitions of fluorenone in gas phase and hydrated zeolite L

State	ε_{iso}	ε_{zeo}^{peak}	$\varepsilon_{zeo}^{min} - \varepsilon_{zeo}^{max}$
1	2.75	2.07	1.61–2.63
2	3.04	2.89	2.37–3.27
3	3.56	3.15	2.67–3.53
4	3.98	3.69	3.25–3.97
5	4.06	3.81	3.41–4.11
6	4.38	3.95	3.61–4.35
7	4.62	4.25	3.71–4.47
8	4.78	4.31	3.91–4.53
9	4.90	4.35	4.17–4.75
10	4.95	4.53	4.27–5.01

The order of the excited states corresponds to the LR-TDDFT result for isolated fluorenone. The correspondence of the transitions obtained for fluorenone in different environments is arbitrary because of the large number of single-electron excitations contributing to a given transition in LR-TDDFT based methods. Such attribution is unambiguous only if there is a clear dominant excitation, as it is the case in the lowest transition. ε_{iso} is the calculated excitation energy of isolated fluorenone. ε_{zeo}^{peak} is the position of the peak of the simulated absorption band for fluorenone in the hydrated zeolite; ε_{zeo}^{min} and ε_{zeo}^{max} are the lowest and highest values of the calculated excitation energy of a given transition in the sample, respectively.

The good agreement between the simulated and observed absorption spectra suggests a more detailed analysis of the factors affecting band shapes. Such analyses are reported in the subsequent sections.

3.3 Absorption spectra of fluorenone in hydrated and dry zeolites

In this section, the spectrum of fluorenone in the dry zeolite is analyzed. Such a spectrum was obtained by using a different sample of configurations corresponding to a different statistical ensemble. In this case, the electronic density of the environment ρ_B used to evaluate the embedding potential from eqn (1) did not include the contributions from water molecules because they were absent in the corresponding FPMD simulation.

Fig. 4 shows the simulated absorption spectra of fluorenone in hydrated and in dry zeolite L. The overall effect (structural and electronic) of the water molecules is significantly smaller than that of the zeolite framework and the counter ions. This is in line with the analysis given in ref. 32 pointing out that the zeolite environment affects fluorenone's excitations primarily by the direct interaction between K^+ cations and fluorenone's carbonyl group. The lowest energy band is broader in the dry zeolite, indicating that the structural fluctuations in the system are reduced upon hydration.

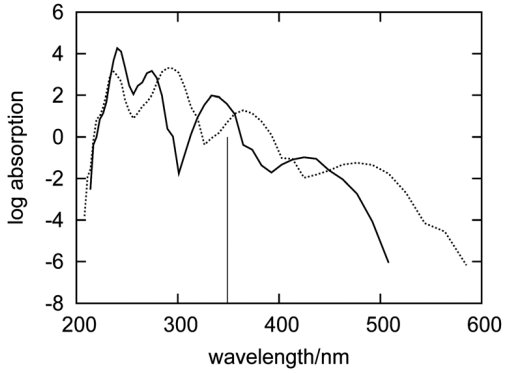


Fig. 4 Simulated absorption spectra of fluorenone in dry (dashed curve) and hydrated (solid curve) zeolite L. The vertical line indicates the calculated wavelength of the lowest excitation for isolated fluorenone in its equilibrium geometry.

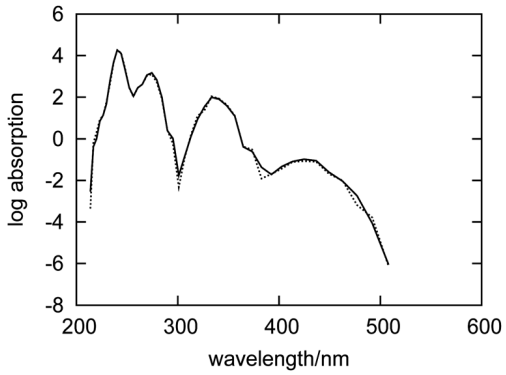


Fig. 5 Simulated absorption spectrum for an artificial system comprising only fluorenone, counterions and the zeolite L framework with geometries taken from the FPMD of fluorenone in hydrated zeolite L (dashed curve). The simulated spectrum for fluorenone in hydrated zeolite L is also shown for comparison (solid curve).

The differences between the absorption spectra of fluorenone in dry and wet zeolites might originate from direct interactions between fluorenone and water, but also from differences in the overall geometry of the dye–zeolite system in the two statistical ensembles. To distinguish between these two effects, the following computational experiment was made. The spectra were evaluated using the trajectory corresponding to the hydrated zeolite, but with the embedding potential (eqn (1)) evaluated without contributions from water molecules. Both spectra are shown in Fig. 5.

The two spectra appear to be almost identical, indicating therefore that the direct interactions between water and fluorenone hardly affect the electronic structure of fluorenone. Actually, during the simulation, water molecules are not directly hydrogen bonded to the fluorenone carbonyl group, but rather preferentially coordinated to other potassium cations in the zeolite channel.[32] The differences between the absorption spectra of fluorenone in hydrated and in dry zeolite L shown in Fig. 4 should, therefore, arise from structural factors, *i.e.*, changes of the fluorenone's preferential orientation and average geometry due to the presence of water molecules.

3.4 The effect of the environment on the internal geometry of fluorenone and its consequences on the absorption bands

The absorption spectrum of fluorenone in zeolite L is affected by both fluctuations in the environment and changes of the internal geometry of fluorenone. As highlighted in Fig. 6, the internal geometry of fluorenone fluctuates following the fluctuating environment during the FPMD trajectory. The fact that

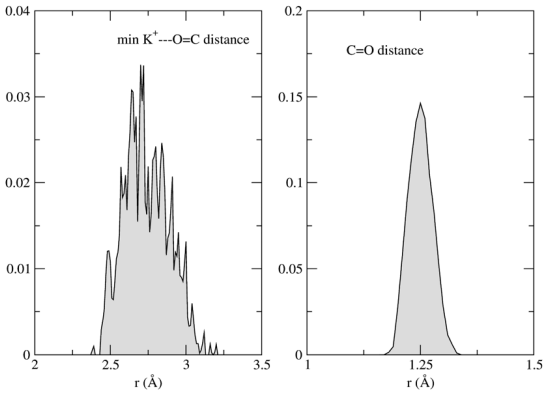


Fig. 9 Left: shortest $K^+\cdots O{=}C$ distance histogram; right: C═O distance histogram.

the instantaneous environment changes in the sampled simulation time is evidenced by the time evolution of the $K^+\cdots O{=}C$ distances, indicating that the carbonyl oxygen of fluorenone switches coordination between two distinct nearest neighbor potassium cations. As an effect, the average C═O bond length, 1.250 Å, is greater than that of isolated fluorenone at the equilibrium geometry (1.220 Å),[32] fluctuates between 1.180 Å and 1.325 Å, and the oscillation amplitude is correlated with the shortest $K^+\cdots O{=}C$ distances. It is instructive, therefore, to estimate the importance of such fluctuations compared to the other factors discussed in the previous sections. To this aim, yet another artificial absorption spectrum of fluorenone was generated following a strategy similar to the one adopted to investigate the significance of the direct electronic effects of water molecules. In this case, the FPMD trajectory corresponding to the hydrated zeolite was used, but the embedding potential was switched off. The obtained spectrum is shown in Fig. 7 together with the positions of the excitation levels in the gas phase. The simulated spectrum of fluorenone in hydrated zeolite L is shown as well for reference purposes. Fig. 8 shows the stick spectrum, *i.e.*, the distribution of the narrow absorption lines corresponding to the lowest excitation arising from

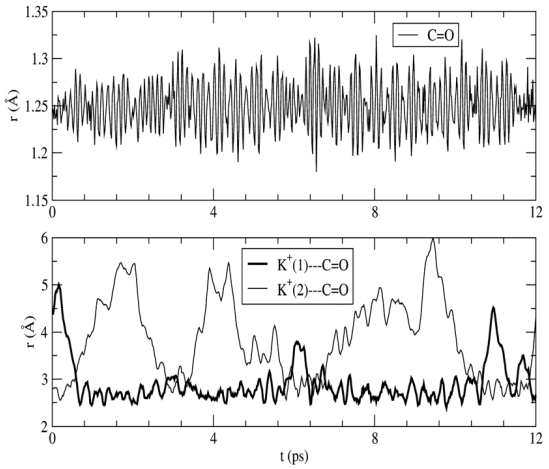


Fig. 6 Top: fluorenone C═O distance *vs.* simulation time. Bottom: $K^+\cdots O{=}C$ distances *vs.* time for the two K^+ closest to the fluorenone carbonyl group.

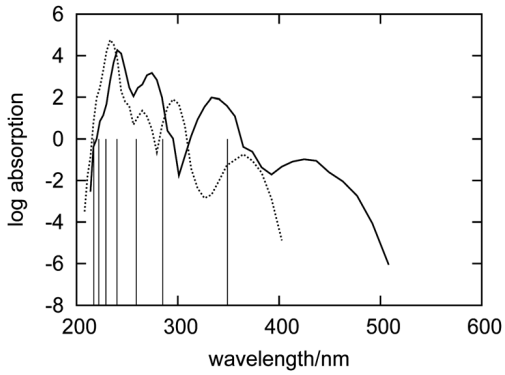


Fig. 7 Simulated absorption spectra of an artificial system comprising only fluctuating fluorenone. The geometries are taken from the FPMD of fluorenone in hydrated zeolite L (dashed curve). The simulated spectrum of fluorenone in hydrated zeolite L is also shown for comparison (solid curve). The vertical lines indicate the calculated wavelengths of the excitations of isolated fluorenone in its equilibrium geometry. The forbidden weak transitions are not shown.

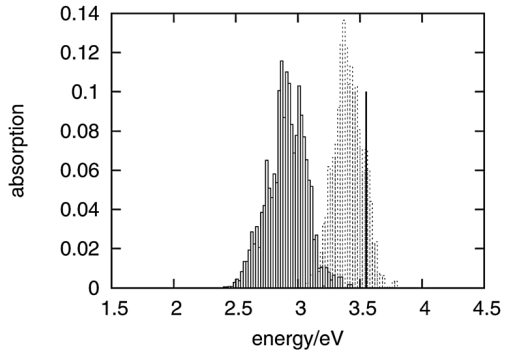


Fig. 8 Distribution of the lowest excitation energies for the artificial system of fluctuating fluorenone (dashed boxes) and for fluorenone in hydrated zeolite L (solid boxes) corresponding to the spectra in Fig. 7. The black vertical line indicates the calculated lowest excitation energy for isolated fluorenone in its equilibrium geometry.

the instantaneous geometries. The fluctuation of the fluorenone's internal geometry results in the broadening of the band and in a shift of the band maximum for the lowest excitation. In Fig. 8, the spectrum changing from a single line (excitation energy and intensity at equilibrium geometry) to a band (distribution of intensities for the isolated chromophore with fluctuating geometry), can be attributed to vibronic broadening, as the distribution of oscillator strengths correlates with the fluctuations of the C═O bond length. However, both broadening and shift effects are much smaller than those induced by the full hydrated zeolite L environment. Interestingly, the histograms in Fig. 8 correlate well with the histograms of the C═O bond length and of the shortest $K^+\cdots O{=}C$ distances sampled along the trajectory, reported in Fig. 9. Indeed, the fluctuating C═O bond length mirrors the band broadening due to the fluctuating internal geometry of fluorenone, while the $K^+\cdots O{=}C$ distance distribution reproduces the multipeak shape of the calculated

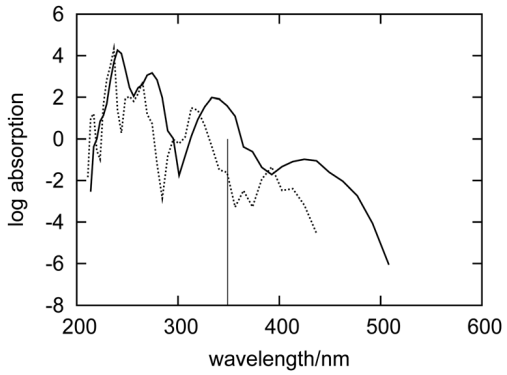


Fig. 10 Simulated absorption spectra of fluorenone in hydrated zeolite L with its long axis perpendicular (dashed curve) and parallel (solid curve) to the zeolite L channel. The vertical line indicates the calculated wavelength of the lowest excitation for isolated fluorenone in its equilibrium geometry.

absorption band for embedded fluorenone. In summary, the changes in fluorenone's internal geometry provide a minor contribution to the whole effect of the hydrated zeolite L environment on fluorenone's absorption spectrum.

3.5 The effect of the environment on the orientation of fluorenone and its consequences on the absorption bands

The relatively large diameter of the channels in zeolite L leaves some degrees of freedom for fluorenone's orientation. In hydrated zeolite L, however, the fluorenone molecule has a strong preference to orient along the channel instead of perpendicular to the channel.[32] In particular, in the case of the hydrated system, a distribution centered at 21° was obtained from the FPMD trajectory for the angle formed by fluorenone's long axis with the zeolite channel axis.[32] In order to investigate the effect of fluorenone's preferential orientation on the absorption spectrum, we also calculated the spectrum of fluorenone with its long axis perpendicular to the channel from a 5 ps trajectory. The FPMD simulation of this trajectory was performed at the temperature of 40 K starting with fluorenone oriented approximately perpendicular to the channel. As mentioned in Section 2.1, such a low temperature is needed because the perpendicular orientation is not stable under room temperature conditions.

Fig. 10 shows the spectra of fluorenone with different orientations in hydrated zeolite L. The spectrum of fluorenone with perpendicular orientation is redshifted by around 0.2 eV with respect to that of fluorenone with parallel orientation, although the band shapes of the two spectra are not much different. This finding is easily rationalized by taking into account that along the trajectory with fluorenone perpendicular to the channel, the carbonyl oxygen of fluorenone is not bound to potassium cations, but directly interacts with a water molecule *via* a hydrogen bond. Therefore both the calculated spectra and FPMD results indicate that fluorenone has an orientation preference in the hydrated zeolite and that the coexistence of parallel/perpendicular arrangements should rather be excluded. Such coexistence would result in the appearance of additional bands not visible in the experimental spectrum.

4 Conclusions

The simulated absorption spectrum reproduces all the main features of the measured spectrum of fluorenone in zeolite L in the region of 250–550 nm. Each distinguishable band in the long wavelength part of the absorption spectrum of fluorenone in zeolite L can be attributed to a single electronic excitation with its width determined by fluctuations of the geometry of the whole system. The adopted computational model provides, therefore, an adequate tool for simulating and interpreting the shape of the absorption bands in such systems. The overall effect of the fluorenone environment on its absorption bands can be decomposed into geometrical factors (orientation of the fluorenone with respect to the zeolite framework, mobility of counter ions and water, and variations of the internal geometry of fluorenone) and the direct electronic effect. The adopted strategy allows one to estimate the relative significance of these factors by selectively switching them off and on. Results obtained using such modified models lead to the following key observations:

(i) The overall redshift of the spectrum observed in hydrated zeolite L compared to cyclohexane originates mainly from the interactions between the counter-ions of the zeolite and the carbonyl group of fluorenone. These interactions lie at the origin of the rather conservative arrangement of the $K^+\cdots O{=}C$ motif in the statistical ensemble.

(ii) The direct effect of water molecules on the excitation energies of fluorenone is negligible at any instantaneous geometry. Water molecules rather affect fluorenone's orientation inside the zeolite channel by stabilizing the parallel arrangement.

(iii) The fluctuations of the internal geometry of fluorenone along FPMD trajectories provide a minor effect on the shape of the absorption bands.

In summary, the simulated absorption bands and the analysis of the factors affecting their position and width support the interpretation that the preferential orientation of fluorenone in hydrated zeolite L is parallel to the zeolite channel. The coexistence of parallel/perpendicular arrangements is rather to be excluded.

Finally, the present study shows that the FDET based simulations of the absorption band shapes can provide an accurate model of the experimental band shapes of chromophores in condensed phases. Such a model is based on first principles except for the offset on the energy scale ($\Delta\varepsilon_{iso}$), which is empirically adjusted due to the intrinsic inaccuracy of the method used to evaluate the vertical excitation energies and the difference between the vertical excitation energy and the absorption energy of the chromophore in the gas phase. The first contribution to $\Delta\varepsilon_{iso}$ is strongly dependent on the computational method and the chromophore in question whereas the first-principles based computation part of the second contribution represents a major computational challenge.

Acknowledgements

This work was supported by the grant 200020-134791 from the Fonds National Suisse de la Recherche Scientifique. CINECA supercomputing center (Bologna, Italy) is acknowledged for computing time (ISCRA project 2011 ''ZeoDye'', HP10B8FNIH).

References

1 International Zeolite Association, http://www.izastructure.org.
2 T. Ohsuna, B. Slater, F. Gao, J. Yu, Y. Sakamoto, G. Zhu, O. Terasaki, D. E. W. Vaughan, S. Qiu and C. R. A. Catlow, *Chem.–Eur. J.*, 2004, **10**, 5031–5040.
3 G. Calzaferri, S. Huber, H. Maas and C. Minkowski, *Angew. Chem., Int. Ed.*, 2003, **42**, 3732–3758.
4 G. Calzaferri, *Langmuir*, 2012, **28**, 6216–6231.
5 A. Devaux, F. Cucinotta, S. Kehr and L. De Cola, *Functional supramolecular architectures for organic electronics and nanotechnology*, Wiley, 2011, vol. 1, pp. 283–242.
6 P. Cao, Y. Wang, H. Li and X. Yu, *J. Mater. Chem.*, 2011, **21**, 2709–2714.
7 P. Cao, H. Li, P. Zhang and G. Calzaferri, *Langmuir*, 2011, **27**, 12614–12620.
8 M. Woerdemann, S. Gläsener, F. Hörner, A. Devaux, L. De Cola and C. Denz, *Adv. Mater.*, 2010, **22**, 4176–4179.
9 F. Cucinotta, Z. Popovic, E. A. Weiss, G. M. Whitesides and L. De Cola, *Adv. Mater.*, 2009, **21**, 1142–1145.
10 A. Guerrero-Martinez, S. Fibikar, I. Pastoriza-Santos, L. Liz-Marzan and L. De Cola, *Angew. Chem., Int. Ed.*, 2009, **48**, 1169.
11 S. Hashimoto, K. Samata, T. Shoji, N. Taira, T. Tomita and S. Matsuo, *Microporous Mesoporous Mater.*, 2009, **117**, 220–227.
12 K. B. Yoon, *Acc. Chem. Res.*, 2007, **40**, 29–40.
13 I. Lopez-Duarte, L.-Q. Dieu, I. Dolamic, M. V. Martinez-Diaz, T. Torres, G. Calzaferri and D. Brühwiler, *Chem.–Eur. J.*, 2011, **17**, 1855–1862.
14 D. Brühwiler, G. Calzaferri, T. Torres, J. H. Ramm, N. Gartmann, L.-Q. Dieu, I. Lopez-Duarte and M. V. Martinez-Diaz, *J. Mater. Chem.*, 2009, **19**, 8040–8067.
15 Z. Popovic, M. Otter, G. Calzaferri and L. De Cola, *Angew. Chem., Int. Ed.*, 2007, **46**, 6188–6191.
16 M. M. Tsotsalas, K. Kopka, G. Luppi, S. Wagner, M. P. Law, M. Schäfers and L. De Cola, *ACS Nano*, 2010, **4**, 342–348.
17 Z. Li, G. Luppi, A. Geiger, H.-P. Josel and L. De Cola, *Small*, 2011, **7**, 3193–3201.
18 J. El-Gindi, K. Benson, L. De Cola, H.-J. Galla and N. Seda Kehr, *Angew. Chem., Int. Ed.*, 2012, **51**, 3716–3720.
19 L. Gartzia-Rivero, J. Banuelos-Prieto, V. Martinez-Martinez and I. Lopez- Arbeloa, *ChemPlusChem*, 2012, **77**, 61–70.
20 V. Vohra, G. Calzaferri, S. Destri, M. Pasini, W. Porzio and C. Botta, *ACS Nano*, 2010, **4**, 1409–1416.
21 V. Vohra, A. Bolognesi, G. Calzaferri and C. Botta, *Langmuir*, 2010, **26**, 1590–1593.
22 G. Calzaferri, R. Méallet-Renault, D. Brühwiler, R. Pansu, I. Dolamic, T. Dienel, P. Adler, H. Li and A. Kunzmann, *ChemPhysChem*, 2011, **12**, 580–594.
23 T. Dienel, C. Bauer, I. Dolamic and D. Brühwiler, *Sol. Energy*, 2010, **84**, 1366–1369.
24 G. Calzaferri, D. Brühwiler and A. Kunzmann, *Patent*, 2008, CH 698333 and WO 2010/009560.
25 M. G. Debije and P. P. C. Verbunt, *Adv. Energy Mater.*, 2012, **2**, 12–35.
26 S. Megelski, A. Lieb, M. Pauchard, A. Drechsler, S. Glaus, C. Debus, A. J. Meixner and G. Calzaferri, *J. Phys. Chem. B*, 2001, **105**, 25–35.
27 M. Busby, C. Blum, M. Tibben, S. Fibikar, G. Calzaferri, V. Subramaniam and L. De Cola, *J. Am. Chem. Soc.*, 2008, **130**, 10970–10976.
28 C. Blum, Y. Cesa, M. Escalante and V. Subramaniam, *J. R. Soc. Interface*, 2009, **6**, S35–S43.
29 A. Gasecka, L.-Q. Dieu, D. Brühwiler and S. Brasselet, *J. Phys. Chem. B*, 2010, **114**, 4192–4198.
30 M. Busby, A. Devaux, C. Blum, V. Subramaniam, G. Calzaferri and L. De Cola, *J. Phys. Chem. C*, 2011, **115**, 5974–5988.
31 K. F. Domke, J. P. R. Day, G. Rago, T. A. Riemer, M. H. F. Kox, B. M. Weckhuysen and M. Bonn, *Angew. Chem., Int. Ed.*, 2012, **51**, 1343–1347.
32 E. Fois, G. Tabacchi and G. Calzaferri, *J. Phys. Chem. C*, 2010, **114**, 10572–10579.
33 A. Devaux, C. Minkowski and G. Calzaferri, *Chem.–Eur. J.*, 2004, **10**, 2391–2408.
34 E. Fois, G. Tabacchi and G. Calzaferri, *J. Phys. Chem. C*, 2012, **116**, 16784–16799.
35 A. Kuboyama, *Chem. Phys. Lett.*, 1976, **41**, 544–546.
36 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050–8053.
37 T. A. Wesolowski, *One-electron equations for embedded electron density: challenge for theory and practical payoffs in multi-level modelling of soft condensed matter*, World Scientific, 2006, vol. X, pp. 1–82.
38 T. A. Wesolowski, *J. Am. Chem. Soc.*, 2004, **126**, 11444–11445.
39 T. A. Wesolowski, *Phys. Rev. A*, 2008, **77**, 012504-1–012504-9.
40 K. Pernal and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2009, **109**, 2520–2525.
41 G. Fradelos, J. J. Lutz, T. A. Wesolowski, P. Piecuch and M. Woch, *J. Chem. Theory Comput.*, 2011, **7**, 1647–1666.
42 G. Fradelos, J. J. Lutz, T. A. Wesolowski, P. Piecuch and M. Wloch, *Advances in the Theory of Quantum Systems in Chemistry and Physics*, Springer Netherlands, 2012, vol. 22, pp. 219–248.
43 R. Improta, V. Barone and F. Santoro, *Angew. Chem.*, 2007, **119**, 409–412.
44 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1997, **78**, 1396.
45 J. M. Newsam, *J. Phys. Chem.*, 1989, **93**, 7689–7694.
46 G. Bandoli, D. Barreca, A. Gasparotto, R. Seraglia, E. Tondello, A. Devi, R. A. Fischer, M. Winter, E. Fois, A. Gamba and G. Tabacchi, *Phys. Chem. Chem. Phys.*, 2009, **11**, 5998–6007.
47 D. Barreca, E. Fois, A. Gasparotto, R. Seraglia, E. Tondello and G. Tabacchi, *Chem.–Eur. J.*, 2011, **17**, 10864–10870.
48 D. Barreca, G. Carraro, A. Devi, E. Fois, A. Gasparotto, R. Seraglia, C. Maccato, C. Sada, G. Tabacchi, E. Tondello, A. Venzo and M. Winter, *Dalton Trans.*, 2012, **41**, 149–155.
49 D. Vanderbilt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **41**, 7892–7895.
50 L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.*, 1982, **48**, 1425–1428.
51 D. R. Hamann, M. Schlüter and C. Chiang, *Phys. Rev. Lett.*, 1979, **43**, 1494–1497.
52 N. Troullier and J. L. Martins, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **43**, 1993–2006.
53 E. Fois, A. Gamba, G. Tabacchi, S. Quartieri and G. Vezzalini, *J. Phys. Chem. B*, 2001, **105**, 3012–3016.
54 E. Fois, A. Gamba, G. Tabacchi, S. Quartieri and G. Vezzalini, *Phys. Chem. Chem. Phys.*, 2001, **3**, 4158–4163.
55 E. Fois, A. Gamba, G. Tabacchi, R. Arletti, S. Quartieri and G. Vezzalini, *Am. Mineral.*, 2005, **90**, 28–35.
56 E. Spano', G. Tabacchi, A. Gamba and E. Fois, *J. Phys. Chem. B*, 2006, **110**, 21651–21661.
57 G. Tabacchi, E. Gianotti, E. Fois, G. Martra, L. Marchese, S. Coluccia and A. Gamba, *J. Phys. Chem. C*, 2007, **111**, 4946–4955.
58 A. Gamba, G. Tabacchi and E. Fois, *J. Phys. Chem. A*, 2009, **113**, 15006–15015.
59 A. Tilocca and E. Fois, *J. Phys. Chem. C*, 2009, **113**, 8683–8687.
60 E. Fois, G. Tabacchi, D. Barreca, A. Gasparotto and E. Tondello, *Angew. Chem., Int. Ed.*, 2010, **49**, 1944–1948.
61 R. Car and M. Parrinello, *Phys. Rev. Lett.*, 1985, **55**, 2471–2474.
62 J. Marx and D. Hutter, *Ab Initio Molecular Dynamics*, Cambridge University Press, 2009.
63 CPMD code, MPI fur Festkorperforschung: Stuttgart, Germany; IBM Zurich Research Laboratory: Zurich, Switzerland, www.cpmd.org, 1990–2012.
64 M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2004, **96**, 577–588.
65 C. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011.
66 J. Neugebauer, C. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805–7814.
67 ADF, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.
68 J. Neugebauer, *J. Chem. Phys.*, 2009, **131**, 084104-1–084104-12.
69 O. Gritsenko, P. Schipper and E. Baerends, *Chem. Phys. Lett.*, 1999, **302**, 199–207.
70 P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344–1352.
71 L. Bernasconi, *J. Chem. Phys.*, 2010, **132**, 184513.
72 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71–76.
73 J. M. G. Lastra, J. W. Kaminski and T. A. Wesolowski, *J. Chem. Phys.*, 2008, **129**, 074107-1–074107-15.
74 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 094115-1–094115-13.
75 N. A. Besley, M. T. Oakley, A. J. Cowan and J. D. Hirst, *J. Am. Chem. Soc.*, 2004, **126**, 13502–13511.
76 L. Goerigk and S. Grimme, *J. Chem. Phys.*, 2010, **132**, 184103-1–184103-9.
77 D. P. Chong and Y. Takahata, *Chem. Phys. Lett.*, 2006, **418**, 286–291.
78 M. E. Casida, in *Time-dependent density-functional response theory for molecules*, ed. D. P. Chong, World Scientific, Singapore, 1995, p. 155.

Supplementary Material (ESI) for PCCP
This journal is © the Owner Societies 2012

First-Principles Simulation of the Absorption Bands of Fluorenone in Zeolite L

Xiuwen Zhou,[a] Tomasz A. Wesolowski,[a] Gloria Tabacchi,[b] Ettore Fois,[b] Gion Calzaferri,[c] André Devaux [d]

[a] Département de Chimie Physique, Université de Genève, 30 quai Ernest-Ansermet, CH-1211 Gen`eve 4, Switzerland.

[b] Department of Science and High Technology, University of Insubria, and INSTM, Via Lucini 3, I-22100 Como, Italy.

[c] Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland.

[d] Institute of Inorganic Chemistry, University of Fribourg, CH-1700 Fribourg, Switzerland.

6th September 2012

Author for correspondence:
e-mail:tomasz.wesolowski@unige.ch

The contents:

Supplementary Material (ESI) for PCCP
This journal is © the Owner Societies 2012

1. Test of the value of "N_{config}" (the total number of instantaneous configurations in the sample) used in Equation 2, N_{config} =200, 300, 400, 500, 600, 650.

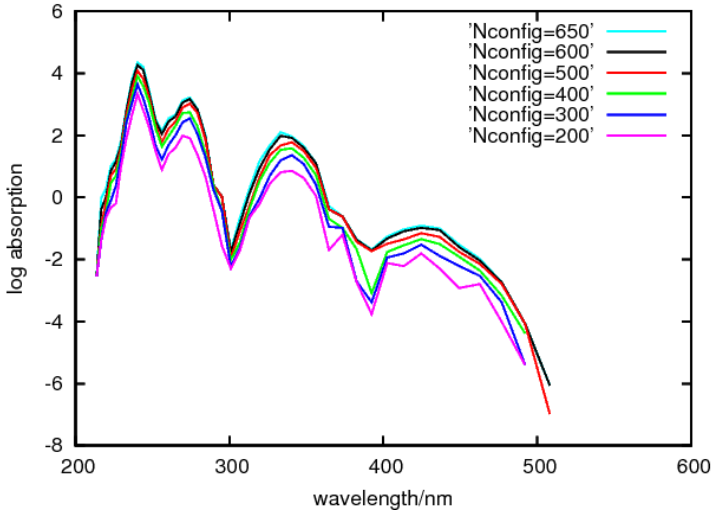


Figure 1. Spectra simulated from different numbers of configurations.

2. Test of the value "*a*" (summation range) used in Equation 2, a = 0.06 eV, 0.08 eV, 0.10 eV, 0.14 eV.

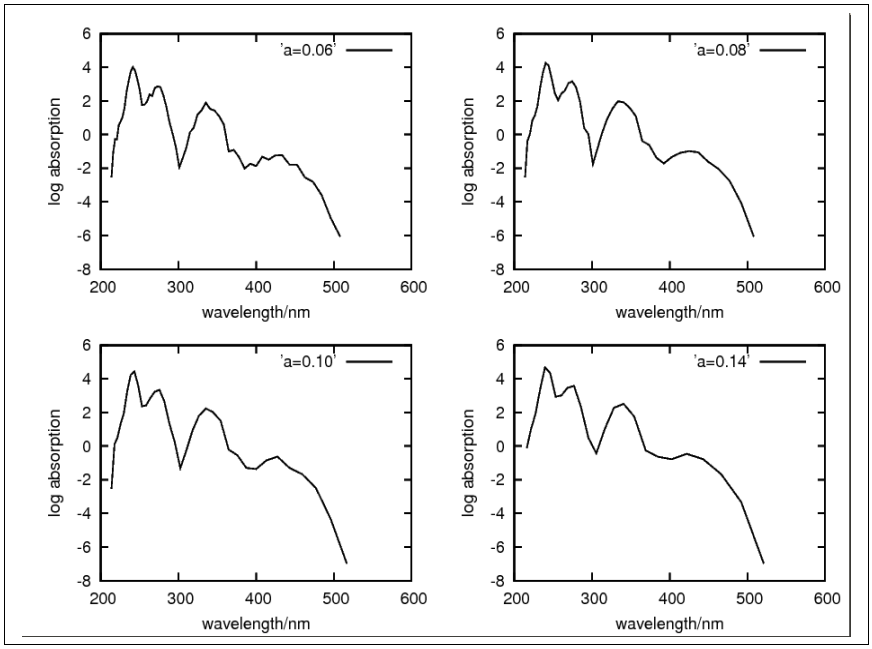


Figure 2. Spectra simulated using different values of the parameter *a* in Equation 2.

Supplementary Material (ESI) for PCCP
This journal is © the Owner Societies 2012

3. The effect of the choice of ρ_B in Equation 1 on calculated excitation energies

Model:

Fluorenone: $C_{13}H_8O$ (22 atoms),

Environment: $Si_{50}Al_{16}K_{16}O_{163}H_{62}$ (307 atoms),

Total net charge=0.

The cut bonds are saturated by hydrogen atoms.

The method to generate electron density of

the $Si_{50}Al_{16}K_{16}O_{163}H_{62}$ cluster.

a) ρ_B from Kohn-Sham calculations

b) ρ_B as superposition of spherically symmetric atomic densities

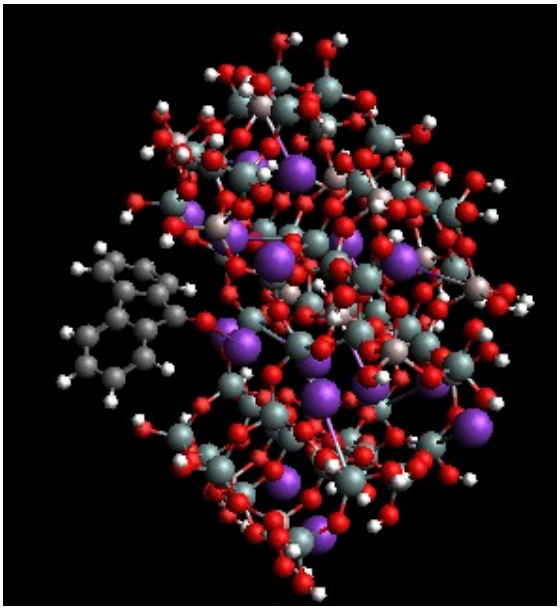


Figure 3. Structure of the cluster $Si_{50}Al_{16}K_{16}O_{163}H_{62}$(Fl)

Supplementary Material (ESI) for PCCP
This journal is © the Owner Societies 2012

Excitation energies of fluorenone embedded in the small cluster:

a) ρ_B from Kohn-Sham calculations for the $Si_{50}Al_{16}K_{16}O_{163}H_{62}$ cluster

For each excitation: excitation energies E in a.u. and eV, oscillator strengths f in a.u., dE wrt previous cycle

no.	E/a.u.	E/eV	f	dE/a.u.
1	0.85125E-01	**2.3164**	0.16012E-02	0.54E-08
2	0.11786	3.2072	0.26010E-02	0.32E-08
3	0.12334	3.3562	0.26969E-01	0.17E-07
4	0.13718	3.7328	0.57932E-02	0.25E-07
5	0.15105	4.1102	0.46147E-01	0.46E-07
6	0.16029	4.3616	0.20362	0.22E-07
7	0.16272	4.4279	0.37489	0.23E-07
8	0.16615	4.5211	0.11445E-01	0.26E-07
9	0.17526	4.7690	0.46239E-02	0.45E-06
10	0.17772	4.8361	0.30761E-01	0.20E-06

b) ρ_B as superposition of spherically symmetric atomic densities in the $Si_{50}Al_{16}K_{16}O_{163}H_{62}$ cluster

For each excitation: Excitation energies E in a.u. and eV, oscillator strengths f in a.u., dE wrt previous cycle

no.	E/a.u.	E/eV	f	dE/a.u.
1	0.86309E-01	**2.3486**	0.21357E-02	0.17E-08
2	0.11683	3.1791	0.67695E-02	0.60E-08
3	0.12151	3.3066	0.30147E-01	0.21E-07
4	0.13774	3.7480	0.64472E-02	0.14E-07
5	0.15148	4.1220	0.77947E-01	0.39E-07
6	0.15773	4.2921	0.71092E-01	0.12E-07
7	0.16021	4.3596	0.21680E-01	0.37E-07
8	0.16306	4.4371	0.42017	0.13E-07
9	0.17292	4.7054	0.56213E-02	0.50E-08
10	0.17882	4.8660	0.47209E-01	0.48E-06

Chapter 5
Solvatochromic shifts in the absorption of coumarin 153

"Experience is what you get when you don't get what you want."
— Dan Stanford

5.1 Overall presentation of the article

The solvatochromic shift refers to a strong dependence of absorption or emission spectra with the solvent polarity. A change in the solvent polarity may lead to a change in the energy gap between these electronic states due to different stabilization of the ground and excited states, since the polarities of the ground and excited state of a chromophore are different.

The solvatochromic shift in absorption can be investigated by simulating the position and intensity of the maximum of the absorption band, via calculating the excitation energy and oscillator strength of the corresponding electronic transition. This type of simulation in general involves modeling the structure of the solvation system and calculate the vertical excitation energy and oscillator strength of the corresponding electronic transitions using a proper electronic structure method that can be QM/MM methods or fully QM methods.

There are typically three strategies to model the solute-solvent system: (1) Implicit model, where the solvent is represented by a structureless polarizable uniform continuous medium, and the solute is placed in a cavity in this medium. This model is referred as *uniform continuum model* in this work. This strategy of treating solute-solvent system is computationally efficient but it is sensitive to the cavity size and shape and does not take account into specific solute-solvent interactions. (2) Explicit model, where the solvent is represented by explicit solvent molecules. The solute-solvent structure is usually represented by Monte-Carlo sampling or molecular dynamics trajectories (generated by solution of Newton's equations of motion). This strategy accounts for explicit solute-solvent interactions but it is computationally demanding. (3) Reference interaction site model (RISM), an intermediate level between explicit and implicit models. In this model, the solvent is represented by an average structure described by pair correlation functions. The solvent is still treated as a continuum but the specific solvent-solvent interactions are treated in a statistical manner. This model is referred as *non-uniform continuum model* in this work. This strategy is appealing because it describes the solvent more accurately than the uniform continuum model while with moderate computational efforts.

RISM strategy is used in this work for obtaining the statistically averaged structure of the solvent around the solute, and FDET/LR-TDDFT is used to investigate the electronic excitations of the solvated solute molecule.

The strategy used in this work is advantageous for studying solvatochromic shift by taking account into specific solute-solvent interactions in a statistical way with moderate computational efforts. However it is limited to the study of the solute molecules with relative rigid structure,

because the structure of the solute is assumed unchanged in the solvent. Besides, the force field parameters for both the solvent and the solute are needed in the RISM calculation.

In this work, the calculations were performed by Xiuwen Zhou with the guide of Jakub W. Kaminski and Tomasz A. Wesolowski.

5.2 Reprint of the article

PCCP

View Online
Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 10565–10576

www.rsc.org/pccp

PAPER

Multi-scale modelling of solvatochromic shifts from frozen-density embedding theory with non-uniform continuum model of the solvent: the coumarin 153 case

Xiuwen Zhou, Jakub W. Kaminski and Tomasz A. Wesolowski*

Received 14th December 2010, Accepted 15th April 2011
DOI: 10.1039/c0cp02874f

For nine solvents of various polarity (from cyclohexane to water), the solvatochromic shifts of the lowest absorption band of coumarin 153 are evaluated using a computational method based on frozen-density embedding theory [Wesolowski and Warshel, *J. Chem Phys.*, 1993, **97**, 9050, and subsequent articles]. In the calculations, the average electron density of the solvent $\langle\rho_B(\vec{r})\rangle$ is used as the frozen density. $\langle\rho_B(\vec{r})\rangle$ is evaluated using the statistical-mechanical approach introduced in Kaminski *et al.*, *J. Phys. Chem. A*, 2010, **114**, 6082. The small deviations between experimental and calculated solvatochromic shifts (the average deviation equals to about 0.02 eV), confirm the adequacy of the key approximations applied: (a) in the evaluation of the average effect of the solvent on the excitation energy, using the average density of the solvent instead of averaging the shifts over statistical ensemble and (b) using the approximant for the bi-functional of the non-electrostatic component of the orbital-free embedding potential, are adequate for chromophores which interact with the environment by non-covalent bonds. The qualitative analyses of the origin of the solvatochromic shifts are made using the graphical representation of the orbital-free embedding potential.

1. Introduction

Multi-scale computational methods, which are often included under the label QM/MM (for representative overviews see ref. 1–4 for instance), are used frequently to target such quantities as: equilibrium structures, transition states, conformational equilibria. The quality of the type of properties obtained in QM/MM simulations depends on the approximations used for the embedding operator (see eqn (2) below) as well as on the various corrections added *a posteriori* to the total energy to account for exchange–repulsion, charge-transfer effects, for instance. Including the non-electrostatic contributions to the total energy as *a posteriori* corrections and not as components of the embedding operator results, however, in the lack of self-consistency between the energy and the embedded wavefunction.† The overall quality of the potential-energy surface obtained in such calculations results from many simplifications and approximations and the lack of self-consistency is just one of them. If, however, the modelling targets the effects of the environment on properties which are directly related to the electronic structure such as UV/Vis-, NMR-, or ESR spectra, the quality of the obtained results is determined only by the approximation for the embedding operator. The most common approach in simulation such properties consists also of adding the non-electrostatic contributions to the energy as *a posteriori* corrections and retaining only the electrostatic component of the embedding operator. Such an approach leads to efficient computational methods especially if small basis sets are used (see ref. 5–9 for instance), but which are prone to numerical instabilities if the basis sets extend to the environment.[10–12] In methods based on the frozen-density embedding theory (FDET),[13–17] however, both the energy and the embedding operator are evaluated self-consistently, which makes them especially suitable for studies of the effects of the environment on the electronic structure of the solute. In FDET, the environment of the embedded species is characterized exclusively by the nuclear and charge densities which determine the *exact embedding operator*.‡ The FDET based methods are, therefore, especially suited for multi-scale modelling. The electron density of the environment (ρ_B) is a well-defined and observable quantity at any length-scale. In the present work, we provide an extensive analysis of the applicability of FDET with a particular method to generate the frozen

Department of Physical Chemistry, University of Geneva, 30 quai Ernest-Ansermet, 1211 GENEVE, Switzerland. E-mail: tomasz.wesolowski@unige.ch

† Note that self-consistency between the embedding operator and the embedded wavefunction is usually assured in QM/MM calculations regardless the form of the embedding operator.

‡ Note that the charge density also suffices to generate the approximated embedding operator obtained by neglecting all terms beyond electrostatics.

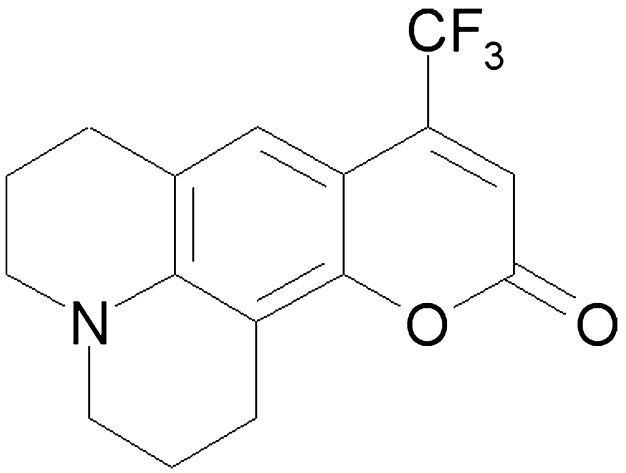


Fig. 1 Coumarin 153.

electron density ($\rho_B(\vec{r})$), on which all FDET results depend parametrically. The frozen density is not obtained from any quantum-mechanical treatment of the solvent but from a model based on classical statistical theory of liquids. Instead of treating the electronic structure of the solvent explicitly, *i.e.*, using a statistical ensemble of instantaneous electron densities of the environment $\rho_{B(i)}$ (where i denotes the configurations of the solvent) in repetitive evaluation of instantaneous excitation energies (conventional strategy applied for instance in our previous works[18,19]), the excitation energy is evaluated only once in calculations performed for $\langle\rho_B\rangle$—average electron density of the solvent surrounding a chromophore. The solute molecule considered in this study is coumarin 153 (see Fig. 1). Coumarins are widely used as laser dyes.[20] Coumarin 153 is considered as a very good probe to study solute–solvent interfaces as it has a rigid structure and is characterized by a single low-lying excited state.[21,22] Nine solvents ranging in polarity from cyclohexane to water are included in the present study.

2. Methods

The computational method applied in this work was introduced in ref. 23. The basic equations of FDET, which depend parametrically on the density ρ_B assumed for the environment, are solved using the ensemble averaged electron density of the solvent (denoted with $\langle\rho_B\rangle$ in the present work). For the sake of completeness, the key elements of the applied theoretical frameworks are given in separate sections below.

2.1 Frozen-density embedding theory

The frozen-density embedding theory[13–17] provides the formal basis for computational methods (see also other representative papers[19,24–27] besides our own work), in which both the energy as well as electronic properties are evaluated in a self-consistent manner as the effective potential and the total energy are related by simple functional derivation. Below, we outline the basic elements of the frozen-density embedding theory:

• **Basic variables:** The total investigated system is characterized by two quantities: the density $\rho_B(\vec{r})$, which for a given electronic problem is a given function, and the density $\rho_A(\vec{r})$, which is represented using auxiliary quantities: occupied orbitals of non-interacting reference system $\{\varphi_A^i(\vec{r})\}$,[13] occupied and unoccupied orbitals of non-interacting reference system,[15] interacting wave-function,[16] or one-particle density matrix.[17]

• **Constrained search:** The optimal electron density $\rho_A(\vec{r})$ is obtained by performing the following search:

$$E_{\mathrm{emb}}[\rho_B] = \min_{\rho_A\geq 0} E^{HK}[\rho_A+\rho_B] \text{ for } \int \rho_A(\vec{r})d\vec{r} = N_A$$
$$= \min_{\rho\geq\rho_B\geq 0} E^{HK}[\rho] \quad (1)$$

where ρ_B is a given electron density such that $\int\rho_B(\vec{r})d\vec{r} = N_B$

• **Performing the constrained search by modifying the external potential:** The search for the optimal ρ_A defined in eqn (1) is conducted in practice by the following equation:

$$(\hat{H}_o + \hat{V}_{\mathrm{emb}})\Psi = E_{\mathrm{emb}}\Psi \quad (2)$$

in which $\hat{H}_o$ is the environment-free Hamiltonian $\hat{H}_o$ and the $\hat{V}_{\mathrm{emb}}(\vec{r})$ has the form of a **local potential** ($v_{\mathrm{emb}}^{\mathrm{eff}}(\vec{r})$, which is determined by the pair of densities $\rho_A(\vec{r})$ and $\rho_B(\vec{r})$ hence orbital-free embedding potential.

• **Orbital-free embedding potential**: The relation between the embedding potential on the densities $\rho_A(\vec{r})$ and $\rho_B(\vec{r})$ depends on the choice of the quantum mechanical descriptor used to generate $\rho_A(\vec{r})$ in the search procedure given in eqn (1).[13,16,17] For the following descriptors, orbitals of non-interacting reference system, a wavefunction of the full Configuration Interaction form, and one-particle density matrix, the orbital-free embedding potential reads:

$$v_{\mathrm{emb}}^{\mathrm{eff}}[\rho_A,\rho_B;\vec{r}] = v_{\mathrm{ext}}^{B}(\vec{r}) + \int\frac{\rho_B(\vec{r}')}{|\vec{r}'-\vec{r}|}d\vec{r}' + \left.\frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\right|_{\rho=\rho_A+\rho_B} - \left.\frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\right|_{\rho=\rho_A} + \left.\frac{\delta T_s[\rho]}{\delta\rho}\right|_{\rho=\rho_A+\rho_B} - \left.\frac{\delta T_s[\rho]}{\delta\rho}\right|_{\rho=\rho_A} \quad (3)$$

The correspondence given in eqn (3) involves density functionals known in the Kohn–Sham formulation[28] of density functional theory:[29] the functional of the exchange–correlation energy ($E_{\mathrm{xc}}[\rho]$) and the functional of the kinetic energy in a non-interacting system ($T_s[\rho]$). The pair of functional derivatives of the functional $T_s[\rho]$ arises from the non-additivity of this functional and represents a potential denoted as $v_t^{\mathrm{nad}}[\rho_A,\rho_B](\vec{r})$ in the present work.

In this context, it is useful to relate the frozen-density embedding theory to the subsystem formulation of density functional theory (SDFT)[30,31] and to the recently developed partition density functional theory (PDFT).[32] Both SDFT and PDFT lead to the exact ground-state electron density and energy of the whole investigated system in an alternative way to the conventional Kohn–Sham framework. In SDFT, the charges of each subsystem are assumed to be integral (similarly as in FDET), whereas fractional charges of subsystems are allowed in PDFT. The FDET targets not the ground-state electron density of the total system but the density minimizing the Hohenberg–Kohn energy functional for the total system

with presence of constraints. FDET, therefore, can lead to the same total ground-state density as SDFT, Kohn–Sham DFT, or PDFT, only for particular constraints[14] (see also below).

In the case of two subsystems, SDFT is based on the following variational principle:

$$E_o = \min_{\rho_A \geq 0, \rho_B \geq 0} E^{HK}[\rho_A + \rho_B] \quad (4)$$

where the search is performed among subsystem densities which are pure-state non-interacting *v*-representable densities ρ_A and ρ_B such that $\int \rho_A(\vec{r})d\vec{r} = N_A$ and $\int \rho_B(\vec{r})d\vec{r} = N_B$.

The sufficient condition for reaching the exact ground-state density in SDFT, it is that it can be decomposed as a sum of two pure-state non-interacting *v*-representable densities comprising integer number of electrons N_A and N_B (see the discussions in ref. 14). FDET does not target the ground-state of the total system but the density, which minimizes the total ground-state energy in presence of the following constraint:

$$\rho \geq \rho_B \quad (5)$$

which is given in advance. The total density obtained in FDET (*i.e.*, in eqn (1)) is, therefore, not equal to the exact ground-state density except for a particular case, *i.e.*, when the difference between $\rho_o^{\mathrm{tot}}(\vec{r})$ and the assumed $\rho_B(\vec{r}')$ is representable using one of the auxiliary descriptors mentioned above: orbitals of the non-interacting reference system,[13] interacting wavefunction,[16] or one particle-density matrix.[17] On the virtue of Hohenberg–Kohn theorems, FDET can lead only to the upper bound of the ground-state energy:

$$E_{\mathrm{emb}}[\rho_B] \geq E_o \quad (6)$$

We underline that the choice for ρ_B is the essential feature of any multi-level type of calculations based on FDET. Compared to empirical QM/MM type of methodologies, which involve many parameters describing the interactions between the quantum mechanical subsystem and its environment, the only subjective choice in constructing the FDET model is the used ρ_B. The procedure to generate ρ_B must reflect, therefore, the physical and chemical properties of the subsystems and might involve various computational costs. Obviously, the number of the electrons in each subsystem must be chosen accordingly to what is known about the systems under investigation. Using the fully variational FDET calculations to attribute the number of electrons to each subsystem is not practical as it would involve costly calculations comparable to DFT treatment of the whole system. Note also that multiple solutions of fully variational and exact FDET calculations are possible.[14] Once the number of electrons is properly assigned to each subsystems, the choice for actual shape of ρ_B must be made. The simplest procedure, which was applied in our first multi-level simulation based on FDET[33] consisted on superposing individual electron densities of isolated solvent molecules. On the other hand, the completely relaxed ρ_B taking into account also the electronic polarization of the subsystem *B* by the embedded species can be obtained in costly *freeze-and-thaw* calculations.[33] This is especially relevant if the embedded species are charged. But even in such a case, the costly *freeze-and-thaw* calculations can be avoided without noticeable deterioration of the accuracy. In the studies on embedded lanthanide cations, we demonstrated that fully variational treatment of both the cation and the environment can be replaced by much simpler procedure consisting of superposing pre-polarized individual atoms in the environment.[34] In the present study, the applied procedure to construct ρ_B does not take into account explicitly the electronic polarization of the environment by the embedded (uncharged) species.

Any numerical implementation of FDET can be easily converted to a method for solving coupled Kohn–Sham-like equations in SDFT. In fact, the first numerical implementation of SDFT for intermolecular complexes used the "freeze-and-thaw" cycle[35] which was applied in a number of subsequent studies (see for instance[36–38]). In the original numerical studies based on SDFT concerning atoms in solids,[30,31] and in the recent numerical implementation of SDFT for molecular liquids,[39] the coupled Kohn–Sham equations are solved simultaneously. We have also shown recently that the "freeze-and-thaw" cycle can be performed simultaneously with displacing nuclear position accelerating the SDFT based geometry optimization.[40]

The "freeze-and-thaw" cycle to solve the coupled Kohn–Sham like equations is used by us in methodological studies on approximants to the bi-functional of the non-additive kinetic potential $v_t^{\mathrm{nad}}[\rho_A,\rho_B]$ (see for instance[41–43]) or in preparation stages for large-scale simulations, in which the search given in eqn (1) is performed for smaller model systems in order to establish the adequacy of the simplified $\rho_B(\vec{r})$ in large-scale simulations. It should be noted that the relaxation of ρ_B is accompanied by the errors introduced by the approximant to the bi-functional of the non-additive kinetic potential $v_t^{\mathrm{nad}}[\rho_A,\rho_B]$ and when the expected polarization effect is small the relaxation should be avoided.

If a non-interacting reference system is used to perform the search given in eqn (1), the corresponding orbitals (ϕ_i^A) are obtained from the following Kohn–Sham-like equations (eqn (20)–(21) in ref. 13):

$$[-\tfrac{1}{2}\nabla^2 + v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_A,\vec{r}] + v_{\mathrm{emb}}^{\mathrm{eff}}[\rho_A,\rho_B;\vec{r}]]\phi_i^A = \varepsilon_i^A \phi_i^A \; i = 1,N^A \quad (7)$$

where $v_{\mathrm{emb}}^{\mathrm{eff}}[\rho_A,\rho_B;\vec{r}]$ is given in eqn (3).

The effectiveness of methods based on eqn (7) for the calculation of the shifts in the electronic structure arising due to the interactions between the embedded system and its environment, was demonstrated for: vertical excitation energies,[15,44] ESR hyperfine coupling constants,[45,46] ligand-field splittings of *f*-levels in lanthanide impurities,[34] NMR shieldings,[47] dipole and quadrupole moments and electronic excitation energies and frequency dependent polarizabilities.[48]

Eqn (7) can be easily used in the general framework of linear-response time-dependent density functional theory (LR-TDDFT)[49] to obtain excitation energies of a system embedded in a frozen electron density ρ_B. The total effective potential in eqn (7) is an explicit functional of ρ_A which makes the evaluation of its functional derivative with respect to ρ_A straightforward leading to efficient computational method introduced in ref. 15. The calculations following such framework, however, represent an additional approximation of neglecting the dynamic, *i.e.*, frequency-dependent response of electron density of the environment (ρ_B) to the external

time-dependent field (Neglect of Dynamic Response of the Environment-NDRE approximation). NDRE is a very adequate approximation in cases of small spectral overlap between the embedded species and the molecules in the environment.[15,18,50,51] On the formal level, using atomic orbitals localized in the environment to construct the occupied and unoccupied embedded orbitals would take into account also the dynamic response of the environment to some extent. But such expansion is hardly attractive from the practical point of view as the computational effort of such calculations would be comparable to full LR-TDDFT treatment of the whole system. The Casida-Wesolowski generalization[53] of Cortona's subsystem formulation of DFT[30] to excited states provides a formal framework going beyond NDRE as it allows for all parts of the total system to respond dynamically. Calculations following the Casida-Wesolowski framework without further approximations such as NDRE are not practical because they would be even more costly than conventional LR-TDDFT calculations for the whole system. Neugebauer introduced a formal framework allowing for a compromise solution lying between NDRE (localized dynamic response) and Casida-Wesolowski cases (no limits on the dynamic response localization) in which selected transitions localized in different subsystems are coupled. Such calculations were shown to efficiently overcome the deficiencies of NDRE approximation in several cases where similar or identical chromophores are localized in different subsystems.[52]

It is worthwhile to underline that in FDET based methods, the effect of the environment on the electronic structure of the embedded species is taken into account by means of modifying the effective Kohn–Sham potential without changing the number of quantum particles treated explicitly (see eqn (7)). As a result, the FDET based approaches are particularly suitable for evaluation of environment-induced shifts on any property depending directly on the embedded orbitals such as the solvatochromic shifts investigated in the present work.

2.2 Classical site distributions from the 3D-RISM-KH theory

The classical site distributions for a molecular solvent around a solute of arbitrary shape are obtained by using the 3D-RISM-KH molecular theory of solvation.[54,55] This method as well as its combination with orbital-free embedding formalism are described in detail in our previous work.[23] Here we will outline the most important points.

The 3D-RISM integral equation can be derived from the six-dimensional, molecular Ornstein-Zernike integral equation[56] for the solute–solvent correlation functions by averaging out the orientation degrees of freedom of solvent molecules while keeping the orientation of the solute macromolecule described at the three-dimensional level.[54,55] It has the form

$$h_\gamma^{\mathrm{uv}}(\vec{r}) = \sum_\alpha \int d\vec{r}\,' c_\alpha^{\mathrm{uv}}(\vec{r}-\vec{r}\,')\chi_{\alpha\gamma}^{\mathrm{vv}}(r'), \qquad (8)$$

where $h_\alpha^{\mathrm{uv}}(\vec{r})$ is the 3D total correlation function of solvent site γ around the solute macromolecule (the superscripts "u" and "v" denoting the solute and solvent, respectively) giving the normalized deviation of the solvent density from its bulk value, which is related to the 3D solute–solvent site distribution function $g_\gamma(\vec{r}) = h_\gamma(\vec{r}) + 1$, and $c_\alpha^{\mathrm{uv}}(\vec{r})$ is the 3D direct correlation function representing "direct" correlations between the solute and solvent site γ, which has the long-range asymptotic behavior of the 3D solute–solvent site interaction potential: $c_\gamma^{\mathrm{uv}}(\vec{r}) \sim \beta u_\gamma^{\mathrm{uv}}(\vec{r})$, where $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature with the Boltzmann constant k_{B}. The site-site susceptibility of pure solvent $\chi_{\alpha\gamma}^{\mathrm{vv}}(r) = \omega_{\alpha\gamma}^{\mathrm{vv}}(r) + \rho_\alpha^{\mathrm{v}} h_{\alpha\gamma}^{\mathrm{vv}}(r)$ gives the response, in terms of the distributions in pure bulk solvent, of site γ to the presence of site α at separation r. It consists of the intramolecular matrix $\omega_{\alpha\gamma}^{\mathrm{vv}}(r) = \delta(r - l_{\alpha\gamma}^{\mathrm{vv}})/(4\pi(l_{\alpha\gamma}^{\mathrm{vv}})^2)$ specifying the intramolecular correlations of solvent molecules with the geometry given by the z-matrix of site separations $l_{\alpha\gamma}^{\mathrm{vv}}$, and the intermolecular part given by the bulk solvent site number density ρ_α^{v} times the site-site radial correlation functions of pure bulk solvent $h_{\alpha\gamma}^{\mathrm{vv}}(r)$. The latter is obtained in advance to the 3D-RISM calculations from the dielectrically consistent RISM theory (DRISM) developed by Perkyns and Pettitt[57] which provides a consistent description of the dielectric properties for ions in polar solvent.

To solve eqn (8), it needs to be complemented with the closure relating functions $h_\gamma^{\mathrm{uv}}(\vec{r})$ and $c_\gamma^{\mathrm{uv}}(\vec{r})$. In literature many different approximate closure relations have been proposed, which application depends on the studied system and/or property. In our work we use closure developed by Kovalenko and Hirata (KH) which is appropriate and successful in the description of the solvation structure and thermodynamics of various inorganic and organic solutes and macromolecules with multiple partial charges in different non-polar and polar liquids, mixtures, and electrolyte solutions, as well as solid–liquid interfaces:[54,55]

$$g_\gamma^{\mathrm{uv}}(\vec{r}) = \begin{cases} \exp(\mathscr{X}_\gamma^{\mathrm{uv}}(\vec{r})) & \text{for } \mathscr{X}_\gamma^{\mathrm{uv}}(\vec{r}) \le 0 \\ 1 + \mathscr{X}_\gamma^{\mathrm{uv}}(\vec{r}) & \text{for } \mathscr{X}_\gamma^{\mathrm{uv}}(\vec{r}) > 0 \end{cases}, \qquad (9)$$

$$\mathscr{X}_\gamma^{\mathrm{uv}}(\vec{r}) = -\beta u_\gamma^{\mathrm{uv}}(\vec{r}) + h_\gamma^{\mathrm{uv}}(\vec{r}) - c_\gamma^{\mathrm{uv}}(\vec{r}),$$

where $g_\gamma^{\mathrm{uv}}(\vec{r}) = 1 + h_\gamma^{\mathrm{uv}}(\vec{r})$ is the radial distribution function.

To account for the solute–solvent coupling in the 3D-RISM-KH calculations aimed at obtaining the site distributions g^{uv}, we adopt the same approach as the one introduced in our previous work on embedding a Kohn–Sham system in the average solvent potential generated by the 3D-RISM-KH method.[58,59] In this approach the electronic structure of the solute is calculated applying KS-DFT formalism, with the presence of the solvent included in the effective Kohn–Sham potential. The whole system is thus described by means of the Helmholtz free energy functional

$$\langle A\rangle[\langle\rho_A\rangle,\{\rho_\gamma^{\mathrm{v}}\}] = \langle E_A\rangle[\langle\rho_A\rangle,\langle\rho_B\rangle] + \Delta\mu_{\mathrm{solv}}^{\mathrm{KH}}[\{\rho_\gamma^{\mathrm{v}}\}] \qquad (10)$$

of the mean electronic density of the embedded molecule $\langle\rho_A(\vec{r})\rangle$ and the set of the classical 3D solvent site density distributions $\rho_\gamma^{\mathrm{v}}(\vec{r}) = \rho_\gamma^{\mathrm{v}} g_\gamma^{\mathrm{uv}}(\vec{r})$ for all solvent sites γ. At this stage, assuming the coupling between the embedded subsystem A and environment B to be weak enough we apply the same above-described FDET functionals to the mean densities $\langle\rho_A(\vec{r})\rangle$ and $\langle\rho_B(\vec{r})\rangle$. For simplicity, we will drop the ensemble averaging brackets in all notations below, keeping in mind that all quantities are mean values averaged over the ensemble of the environment.

2.3 "Dressing up" classical site distributions in electron density

The 3D-RISM-KH method is used to obtain the equilibrium density distributions of their classical interaction sites γ. Such distributions do not suffice to evaluate the orbital-free embedding potential as they are essentially related to nuclei and not electrons. For this reason, the new additional approximation, namely the procedure to "dress up" the site distributions with the electrons is introduced.

First of all, it is assumed that the electron density attributed to each site ($q_\gamma^{\mathrm{v}}(\vec{r})$) is "rigid". This assumption was used already in the first application of the orbital-free embedding potential in simulating solvated system[33] where it was assumed that a "rigid" electron density cloud moves together with the nuclei. In other words, the inhomogeneous average electron density of solvent around the solute molecule is obtained as the convolution of the inhomogeneous probability density of the solvent nuclei and the electron density around each of them. Note that the electron density around each nucleus is assumed to be translationally and rotationally invariant. Therefore, inhomogeneities of the average solvent charge density arise due to inhomogeneities of the nuclear distributions. Such construction of $\langle\rho_B(\vec{r})\rangle$ neglects, therefore, any instantaneous fluctuations of the electron density of each solvent molecule due to fluctuations of its induced dipole moments. Note, however, that the "rigid" density used for each solvent molecule does not correspond to the gas phase but to the liquid. The averaged electron density of the solvent is obtained thus as:

$$\langle\rho_B(\vec{r})\rangle = \sum_\gamma \int d\vec{r}\,' q_\gamma^{\mathrm{v}}(|\vec{r}-\vec{r}\,'|)\rho_\gamma^{\mathrm{v}} g_\gamma^{\mathrm{uv}}(\vec{r}\,'), \tag{11}$$

where the angle brackets $\langle\ldots\rangle$ denote the statistical ensemble average, and $g_\gamma^{\mathrm{uv}}(\vec{r})$ is the 3D solute–solvent site distribution function obtained from the 3D-RISM-KH theory described in the previous section.

Such simplified construction of the electron density of the environment was used in a number of our subsequent applications in liquids.[18,19] In the present work, a further simplification is introduced: the "rigid" electron densities are orientationally averaged, spherically symmetric distributions centered at atoms (O, H in the case of water, for instance) or groups of atoms (O, H, and CH_3, in the case of methanol, for instance). The additional approximation made in the procedure to dress-up the RISM site distributions concerns the evaluation of the convolution (eqn (11)). We note that the functions $q_\gamma^{\mathrm{v}}(\vec{r})$ are short ranged whereas the spacing of the RISM grid, at which the site distributions are evaluated, is large (0.5 Å). Therefore, to a good approximation, the integration can be replaced by summation involving only diagonal elements, with the weights corresponding to the number of electrons associated to each site (8.8 e for oxygen in water and 0.6 e for a hydrogen in water for instance). This assumption means that the weights represent the average charge distribution in the solvent molecule which is not affected by the instantaneous configuration of the solvent. Lifting this simplification by using an orientation-dependent charge density $q_\gamma^{\mathrm{v}}(r,\Omega)$ would involve the orientation-dependent site distributions $g_\gamma^{\mathrm{uv}}(\vec{r},\Omega)$. As a consequence, solving 3D-RISM equations would become significantly more involved. It is important to underline that the average charge distribution differs, however, from its gas-phase counterpart.

The above assumptions lead to the following expression for $\langle\rho_B(\vec{r})\rangle$ at the 3D-RISM grid, which reads:

$$\langle\rho_B(\vec{r})\rangle = \sum_\gamma q_\gamma^{\mathrm{v}}\rho_\gamma^{\mathrm{v}} g_\gamma^{\mathrm{uv}}(\vec{r}). \tag{12}$$

The environment electron density obtained from the 3D-RISM theory in the form of eqn (12) can be used right away to calculate the non-electrostatic components $v_{\mathrm{xc}}^{\mathrm{nad}}[\rho_A,\rho_B](\vec{r})$ and $v_t^{\mathrm{nad}}[\rho_A,\rho_B](\vec{r})$ of the embedding potential $v_{\mathrm{emb}}^{\mathrm{KSCED}}[\rho_A,\rho_B;\vec{r}]$ defined in eqn (3). Furthermore, the electrostatic terms in the embedding potential $v_{\mathrm{emb}}^{\mathrm{KSCED}}[\rho_A,\rho_B;\vec{r}]$ explicitly dependent on the positions of the environment nuclei in the original FDET are replaced in our ensemble approach by the statistical-mechanical average of the electrostatic potential of solvent sites acting on the solute, defined as a variational derivative of the system free energy eqn (10) with respect to the embedded density $\rho_A(\vec{r})$:

$$v_{\mathrm{solv}}^{\mathrm{elec}}(\vec{r}) \equiv \frac{A[\rho_A,\{\rho_\gamma^{\mathrm{v}}\}]}{\delta\rho_A(\vec{r})} = \sum_\gamma \rho_\gamma^{\mathrm{v}} \int d\vec{r}\,' v_\gamma^{\mathrm{el}}(|\vec{r}-\vec{r}\,'|)h_\gamma^{\mathrm{uv}}(\vec{r}\,'), \tag{13}$$

where $v_\gamma^{\mathrm{el}}(\vec{r})$ is the electrostatic potential created by the site electronic charge density $q_\gamma^{\mathrm{v}}(\vec{r})$, and $h_\gamma^{\mathrm{uv}}(\vec{r}) = g_\gamma^{\mathrm{uv}}(\vec{r}) - 1$ is the 3D solute–solvent site total correlation function obtained from the 3D-RISM-KH theory. Note that the expression (13) is obtained by definition with the variational differentiation of the expression for the solvation chemical potential of the embedded molecule with respect to its electronic density $\rho_A(\vec{r})$.[55]

The RISM method has been previously coupled with electronic structure methods to study the solvent effects on the electronic structure of solute. Such combinations include SCF-RISM approach introduced by Kato,[60–62] where RISM is used to embed Hartree–Fock or MC-SCF wavefunction, as well as proposed by Gusarov *et al.*[58,59] self-consistent combination of statistical-theory of liquids with Kohn–Sham formulation of density functional theory. In both of this approaches, the Fock matrix of the isolated solute is augmented by the solvent terms, which constitute only terms arising from electrostatic interactions in the solute–solvent interface. The embedding potential given in eqn (3), used in present work, accounts also for Pauli repulsion terms, which where shown to be of a great importance to correctly describe the environment of the chromophore.[51]

2.4 Evaluation of the solvatochromic shifts at averaged electron density of the environment $\langle\rho_B\rangle$

The basic approximation introduced in ref. 23 and used also in the present work is given in eqn (14):

$$\begin{aligned} I(\omega) &= \left\langle \sum_i f_i[\rho_B(\{R_j^{B(n)}\})]\delta(\omega-\omega_i[\rho_B(\{R_j^{B(n)}\})]) \right\rangle_n \\ &\approx \sum_i f_i[\langle\rho_B\rangle]\delta(\omega-\omega_i[\langle\rho_B\rangle]) \end{aligned} \tag{14}$$

For each instantaneous geometry of the solvent, LR-TDDFT leads to excitation energies and the corresponding intensities which are determined by the electron density of the solvent at this geometry (ρ_B). Averaging the excitation energies and the intensities over the ensemble leads to absorption bands of large width depending on the fluctuation of the solvent. The second line in eqn (14) represent an approximation as the averaging is made for at the level of the density ρ_B. The same type of approximation is known in the QM/MM methods where instead of density and the orbital-free embedding potential, the electrostatic potential generated by the solvent is averaged (see ref. 63 and 64). In either cases (averaging the whole the orbital-free embedding potential or only its electrostatic component) results in neglecting any instantaneous coupling between the structure of the environment and the electronic structure of the embedded species. Such approximation-if applicable-leads to the enormous computational savings as the LR-TDDFT calculations are performed only once. One cannot, however, obtain any realistic band-shapes but only the shifts of the absorption bands as the excitation energies and the intensities are calculated at only *only one* density ρ_B denoted with $\langle\rho_B\rangle$ for obvious reasons. The absorption bands in such a case is represented as a series of spikes at resonance frequencies ω_i. It is worthwhile to notice that $\langle\rho_B\rangle$ is an observable quantity. Although $\langle\rho_B\rangle$ cannot be associated with ground-state density of a system in Born–Oppenheimer approximation it can be used as a parameter in all FDET equations for both ground and excited states.

3. Computational details

The same geometry was used for both isolated and solvated coumarin 153 optimized in Kohn–Sham calculations using the Becke–Perdew (BP)[65,66] exchange–correlation functional.

To obtain the solvent distributions and solute atomic charges, the Kohn–Sham DFT/3D-RISM-KH scheme was used as implemented in ADF program package[58,67,68] using the following main options: (i) the partial charges used for the solute are taken from the multipole derived population analysis (labeled as mdc-q in ADF output)[69] calculated applying STO(TZP) basis set with modified fit functions to give better multipole moments,[69] (ii) the local-density approximation[70–72] (LDA) was used for the exchange–correlation energy, (iii) the integration parameter was set to 6.0, (iv) the size of the 3D-FFT grid to calculate solvent distributions was chosen to $64 \times 64 \times 64$ points in cell size of 32 Å $\times$ 32 Å $\times$ 32 Å, (v) the van der Waals parameters for solvents required to solve 3D-RISM-KH equations were taken from ADF manual and, if not available, the OPLS force field[73,74] was used with charges from mdcm-q population analysis.

The effective embedding potential given in eqn (3) is evaluated for the pair of densities, $\rho_A(\vec{r})$ and $\langle\rho_B(\vec{r})\rangle$, using expressions derived from the local density approximation for each of its non-electrostatic components: the Thomas–Fermi expression for the kinetic energy,[75,76] the Dirac expression for the exchange energy,[70] and the Vosko-Wilk-Nusair expression for the correlation energy.[71,72] The electrostatic component was evaluated using the monomer expansion of the 3D-RISM potential using all the centers of the 3D-RISM grid (see eqn (13)).

The exchange–correlation component of the total effective potential in eqn (7), which is generated by the embedded density ρ_A, *i.e.*, $\frac{\delta E_{xc}[\rho_A]}{\delta\rho_A}$, is approximated using the SAOP potential.[77,78] Such hybrid treatment of the total exchange–correlation potential is motivated by the fact that SAOP potential has correct asymptotic behavior and, as a consequence, is applicable to wider class of excitations in isolated molecules. Note that SAOP cannot be used for approximating the corresponding component of the orbital-free embedding potential because it depends explicitly on orbitals which are not available for $\rho_A + \rho_B$.

4. Results and discussions

4.1 Net-charge densities from dressed RISM probabilities

Fig. 2 and 3 show the net average charge density of the solvent for water and diethyl ether, respectively. The average net charge is the sum of the average nuclear charge obtained as a direct results of 3D-RISM and the average electron $\langle\rho_B\rangle$. It is evaluated as:

$$\langle\rho_B^{\text{net}}(\vec{r})\rangle = \sum_\gamma q_\gamma^{\text{v,net}}\rho_\gamma^{\text{v}}g_\gamma^{\text{uv}}(\vec{r}). \qquad (15)$$

where $q_\gamma^{\text{v,net}}$ is the net charge localized on the given site γ. In the case of water, $q_\gamma^{\text{v,net}}$ of the oxygen is equal to -0.8476e and 0.4238e for hydrogen. For water, the positive charge concentrates in the vicinity of the carbonyl group of coumarin 153 (three blue contours on the density map shown in Fig. 2). These localized positively charged domains correspond to hydrogens (the only atoms in the solvent with a positive net charge). The position of these three areas correspond to three possible geometries of water molecules hydrogen bonded to the carbonyl group: one water molecule forming a linear hydrogen bond, two simultaneously bound water molecules forming bent hydrogen bonds, or an occurrence of three water molecules simultaneously bound to the carbonyl group by three hydrogen bonds. The statistical nature of 3D-RISM theory does not make it possible to distinguish among these three possibilities. The extend of these positive regions near the carbonyl group, although localized, indicates that the hydrogen bonded water molecules possess certain degree of structural flexibility. There are no other regions (even around the CF_3 group) featuring

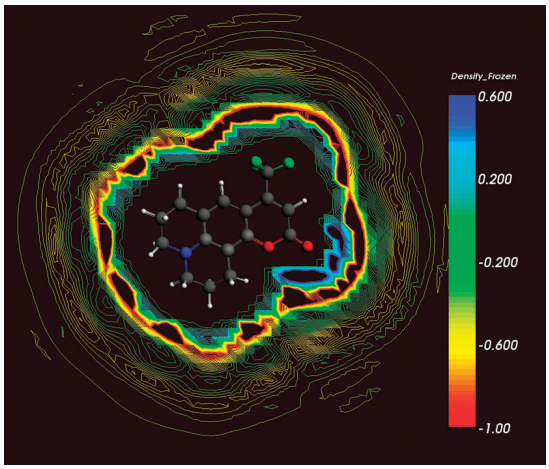


Fig. 2 Net-charge solvent density: coumarin 153 in water.

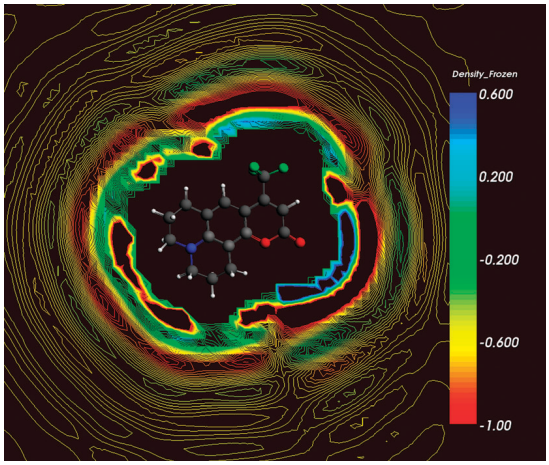


Fig. 3 Net-charge solvent density: coumarin 153 in diethylether.

such localization of the positive charge. The second shell, adjacent to the hydrogen shell, is composed of the oxygen sites as the net charge is negative (shown in red). With the increasing distance from coumarin 153, the positively and negatively charged domains occur alternatively revealing thus the structural order in the solvent.

The net-charge of the diethyl ether solvent is shown in Fig. 3. Due to the fact that two different sites in the diethyl ether molecule are negatively charged, the correspondence between the sign of the net charge and the atoms of the solvent molecule is not direct as the one in the case of water (the net charges $q_\gamma^{\mathrm{v,net}}$ are -0.148961, 0.7206805, -1.143440 for CH_3, CH_2 and O, respectively.) The positively charged CH_2 groups are located in the vicinity of the carbonyl group. The large negative peak is located on the opposite side of the coumarin 153. As it is rather unlikely that the solvent CH_3 groups interact directly with the chromophore, the negatively charged areas represent rather the oxygen atoms. The identification of the groups constituting the second solvation shell is impossible from the net charge maps. The arrangement of the contour lines indicates, nevertheless, that the structure of the solvent is highly ordered. Such interpretation of the positively and negatively areas around the carbonyl group is corroborated by the fact that the positively charged domains are located further from the carbonyl group than those in the case of water. Compared to water (see Fig. 2), the net charge for the diethylether solvent is more structured. This organization reflects the fact that diethyl ether is a larger molecule than water and the structure of the net-charge reflects intramolecular correlations.

4.2 Correspondence of net-charge densities and orbital-free embedding potential

The net-charge density provides information on the structure of the solvent molecules around the solute. A more direct quantity related to the solvatochromism is the orbital-free embedding potential which is determined by the averaged charge of the environment (nuclear and electron) shown in Fig. 4 and 5 for water and diethyl ether, respectively. Before discussing the shape of the orbital-free embedding potential, which determines the solvatochromic shifts, it is useful to indicate the form of the orbitals of coumarin 153 mainly involved in the studied electronic transition: the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). For the isolated coumarin 153, these orbitals are of the π type and are shown in Fig. 8 and 9. HOMO is localized in the carbonyl group of the coumarin 153 but extends also over the whole molecule. LUMO, however, is more localized near the carbonyl group.

We start the analysis with the embedding potential for coumarin 153 in water (Fig. 4 and 5). It can be clearly seen, that the embedding potential is attractive in the vicinity of the carbonyl group. As a consequence, it stabilizes LUMO localized in this attractive region. Fig. 4 and 5 feature also the zero line (shown in orange) which divides the part of the molecule where the orbitals are stabilized from that where the orbitals are destabilized by the solvent. The zero line passes through the center molecule and on the left side of coumarin 153. As a consequence, HOMO can be expected to be destabilized by the solvent whereas LUMO is expected to be even more stabilized. This results in the overall reduction of the HOMO–LUMO gap.

Immediately outside the exclusion volume, the orbital-free embedding potential features strongly attractive areas. According to our previous analyses of the net charge density (Fig. 2), these areas are due to the hydrogen atoms of water in the first solvation shell. The potential $v_{\mathrm{emb}}^{\mathrm{eff}}[\rho_A,\rho_B;\vec{r}]$ is attractive near the CF_3 group (see Fig. 5 where it is shown in the plane comprising the fluorine atoms). The possible stabilizing effect on LUMO is, however, not strong enough to notably enhance the stabilizing effects near the carbonyl group.

In the case of coumarin 153 solvated in diethyl ether, the embedding potential differs significantly from that of water (see Fig. 6 and 7). In the vicinity of the carbonyl group, $v_{\mathrm{emb}}^{\mathrm{eff}}[\rho_A,\rho_B;\vec{r}]$ is weakly attractive and is weakly repulsive at the opposite end of the chromophore. The zero line passes through the center of the chromophore. This behavior results in a smaller effect on the orbital energy for either HOMO or LUMO than the one in the case of water solvent The solvatochromic shift can be, therefore, expected to be smaller.

For other solvents considered in the present work, a similar qualitative analyze of the relation between the shape

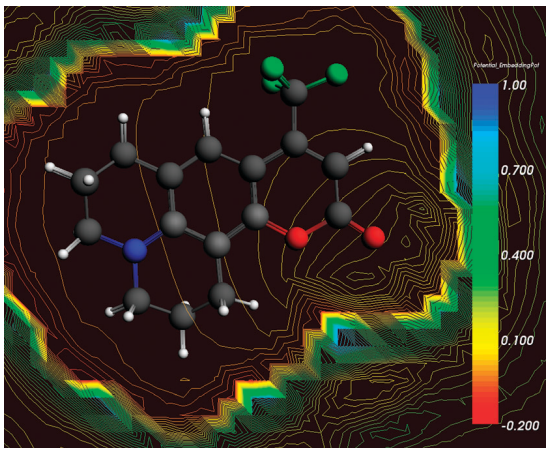


Fig. 4 Orbital-free embedding potential for coumarin 153 in water represented with $\langle\rho_B\rangle$ (projection on the plane of the molecule).

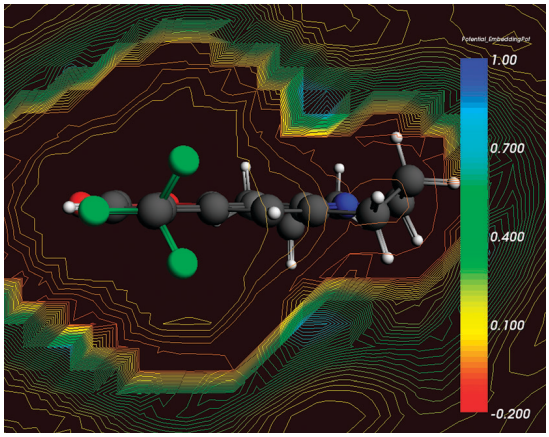


Fig. 5 Orbital-free embedding potential for coumarin 153 in water represented with $\langle \rho_B \rangle$ (projection on the plane comprising the fluorine atoms).

of the orbital-free embedding potential and the orbitals involved in the studied electronic transition reveal that the solvatochromic shifts involve stabilization of LUMO. The quantitative discussion of the calculated solvatochromic shifts is made in the subsequent section.

4.3 Solvatochromic shifts

The experimental reports[22] indicate, that the lowest transition in this chromophore is of $\pi \rightarrow \pi^*$ character and it is red-shifted upon interaction with the polar environment. The qualitative analysis for the two solvents presented in the previous section confirms the experimental picture. The calculated solvatochromic shifts are in excellent agreement with the experimental data (see Table 1). The maximum deviations do not exceed 0.05 eV (calculations for coumarin 153 in 2-propanol using the STO TZ2P basis set). Addition of the diffuse function to the basis set (STO augmented TZP) does not significantly affect the calculated shifts. The magnitude of this effect is about 0.01 eV for most of the solvents except for those water and acetonitrile, for which the effect reaches 0.02 eV (see Table 2).

The small deviations between the calculated and the experimental solvatochromic shifts (relative errors in the 10% range)

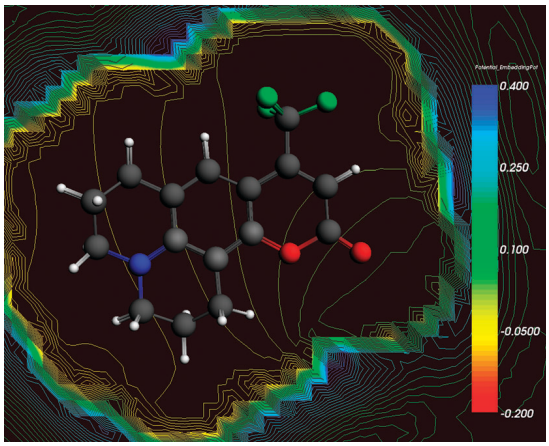


Fig. 6 Orbital-free embedding potential for coumarin 153 in diethyl-ether represented with $\langle \rho_B \rangle$ (projection on the plane of the molecule).

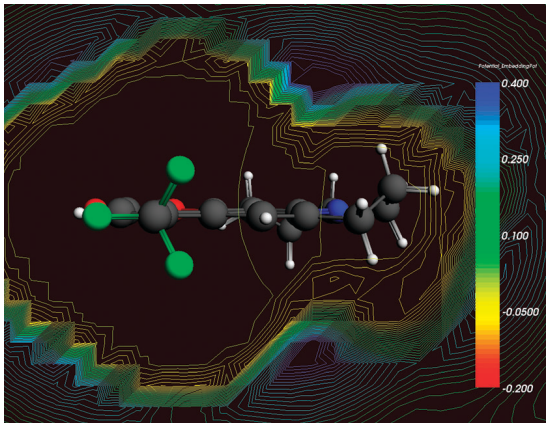


Fig. 7 Orbital-free embedding potential for coumarin 153 in diethyl-ether represented with $\langle \rho_B \rangle$ (projection on the plane comprising the fluorine atoms).

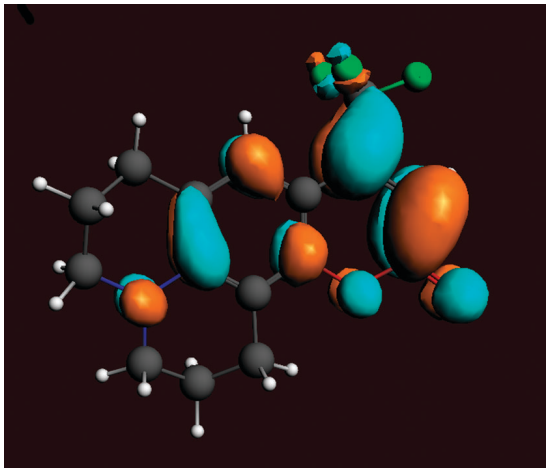


Fig. 8 LUMO of isolated coumarin 153.

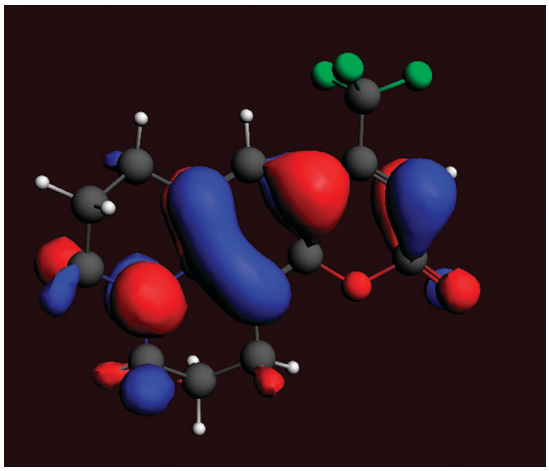


Fig. 9 HOMO of isolated coumarin 153.

call for further analysis of the factors determining the shifts. According to our previous numerical experience[15,51] this indicates that the non-electrostatic component of the orbital-free embedding potential is indispensable. The non-additive exchange–correlation and kinetic components represent the Pauli

Table 1 Solvatochromic shifts[a,b,e] ($\Delta\varepsilon$ in [eV]) of $\pi \rightarrow \pi^*$ absorption band of coumarin 153. The corresponding excitation energy for isolated coumarin 153 is 2.80 eV

Solvent	ε_{diel}	$\Delta\varepsilon_{FDET}$	$\Delta\varepsilon_{Coul}$	$\Delta\varepsilon_{COSMO}$	$\Delta\varepsilon_{exp}$
Water	78	−0.29	−0.30	−0.20	−0.27
Methanol	33	−0.24	−0.24	−0.19	−0.24
Ethanol	25	−0.21	−0.21	−0.19	−0.24
1-Propanol	20	−0.20	−0.20	—	−0.23
2-Propanol	20	−0.18	−0.20	−0.19	−0.23
Acetone	21	−0.21	−0.21	−0.19	−0.19
Acetonitrile	38	−0.17	−0.17	−0.19	−0.20
Diethylether	4	−0.14	−0.14	−0.14	−0.10
Cyclohexane	2	0	0	−0.08	—
AD[c]		0.007	0.001	0.072	
AAD[d]		0.024	0.020	0.043	

[a] Calculated and measured shifts are given with respect to corresponding value for isolated chromophore. [b] All calculations apply the STO TZ2P basis sets. [c] AD is the average deviation between calculated and experimental solvatochromic shifts. [d] AAD is the average absolute deviation between calculated and experimental solvatochromic shifts. [e] Experimental data taken from ref. 22 and 85.

exclusion principle.[79] It assures that the chromophore is not over-polarized by the polar environment. Neglecting the non-electrostatic component, leads to numerical instabilities especially if the basis set is flexible enough to scan the areas near the nuclei in the environment.[42,80,81] In the present work, the environment electric charge (both nuclear and electronic) are not as localized as in the case of the explicit solvent models. $\langle\rho_B\rangle$ is actually smeared over the space around the chromophore. Neglecting the non-electrostatic component (two last terms in eqn (3)) can be expected to be acceptable. Indeed, neglecting this component and retaining only the electrostatic (classical Coulomb) terms (see Table 1 and 2) does not seem to affect noticeably the calculated shifts for the smaller considered basis set (STO TZ2P). For the larger among the considered basis sets (STO augmented TZP), however, neglecting the role of the non-electrostatic term affects the calculated shifts more significantly. Moreover, the electrostatic-only embedding leads to larger deviations from experimental shifts especially for the most polar solvents. These tendencies reflect the variational nature of the FDET based methods.

Table 2 Solvatochromic shifts[a,b,e] ($\Delta\varepsilon$ in [eV]) of $\pi \rightarrow \pi^*$ absorption band of coumarin 153. The corresponding excitation energy for isolated coumarin 153 is 2.86 eV

Solvent	ε_{diel}	$\Delta\varepsilon_{FDET}$	$\Delta\varepsilon_{Coul}$	$\Delta\varepsilon_{COSMO}$	$\Delta\varepsilon_{exp}$
Water	78	−0.31	−0.33	−0.24	−0.27
Methanol	33	−0.25	−0.26	−0.24	−0.24
Ethanol	25	−0.22	−0.23	−0.23	−0.24
1-Propanol	20	−0.21	−0.21	—	−0.23
2-Propanol	20	−0.19	−0.21	−0.23	−0.23
Acetone	21	−0.22	−0.22	−0.23	−0.19
Acetonitrile	38	−0.19	−0.19	−0.24	−0.20
Diethylether	4	−0.15	−0.15	−0.17	−0.10
Cyclohexane	2	0	0	−0.10	—
AD[c]		−0.005	−0.011	−0.012	
AAD[d]		0.022	0.022	0.035	

[a] Calculated and measured shifts are given with respect to the corresponding values for isolated chromophore. [b] All calculations apply the STO augmented TZ2P basis sets. [c] AD is the average deviation between calculated and experimental solvatochromic shifts. [d] AAD is the average absolute deviation between calculated and experimental solvatochromic shifts. [e] Experimental data taken from Ref. 22 and 85.

The specific interactions and the structuring of the solvent are reflected as positive or negative domains in the average net charge plots discussed in the previous sections. In this way, although the solvent molecules are not represented explicitly, the specific interactions are taken into account in the model applied in this work. The global effect of electric polarization, including both specific solvent-chromophore interactions, long-range polarization due to reorientation of the solvent molecules, as well as the electronic polarization, is taken into account by the average net charge distributions as well. It is tempting, therefore, to compare two continuum models: the one applied in this work where the solvent is represented by a non-uniform continuum of electric charge and the uniform dielectric continuum one, commonly used in modelling solvatation effects.[82–84] In dielectic-continuum types of models, the embedding potential includes, however, *only electrostatic* component.§ Since the Coulomb embedding was shown in the previous section leads to reasonable shifts, comparison of shifts obtained with Coulomb embedding to that derived with the uniform continuum dielectric models is expected to reveal the role of specific interactions. To this end, the shifts obtained using commonly used Conductor-like Screening Model (COSMO)[84] are also given in Table 1 and 2. The standard implementation of COSMO into the same code (ADF) as the one used for FDET calculations applying the same exchange–correlation functionals and basis sets was used for this purpose. Comparing the COSMO and FDET results, is made here to identify the importance of replacing the uniform continuum description (COSMO) of the solvent by the non-uniform one (this work) keeping *the same* electrostatic-only embedding potential. The uniform continuum dielectric model leads to the correct trends as far as the dependence of the solvatochromic shift on the solvent polarity is concerned. Taking into account the uniformity of the solvent is, however, crucial for quantitative description of the solvatochromism at two extreme cases. For most polar solvents, for instance, the shifts derived from the uniform dielectric model tend to saturate and do not change much if the dielectric constant increases from 20 to 80.

Turning back to explicit solvent models, reliability of any observable depends critically on two factors: the number of the molecules representing the solvent and the representability of the sample of configuration used in averaging.[19] The relative importance of each of these factors depends on the solvent and the strength of the solute–solvent interactions. In the case of the considered solvents, reducing the solvent model to the nearest neighbors (1–3 molecules) and the sample of configuration to just one the local minimum provides a very poor model of the solvent (see Table 3). The considered conformations include solvent molecules hydrogen bonded to the carbonyl group of the coumarin 153. The solvatochromic shift critically depends on the number of explicitly treated solvent molecules

§ Note that the non-electrostatic components to the energy are obviously inclued in such models by means of *a posteriori* added corrections to the energy (cavitation energy, dispersion energy, *etc.*) but they are not taken into account in the embedding potential.

Table 3 Solvatochromic shifts[a,b] ($\Delta\varepsilon$ in [eV]) of $\pi \rightarrow \pi^*$ absorption band of coumarin 153

Model of solvent	$\Delta\varepsilon$(water)	$\Delta\varepsilon$(methanol)
Explicit:1molecule	−0.03	−0.09
Explicit:2molecules	−0.11	−0.12
Explicit:3molecules	−0.04	−0.06
3D-RISM	−0.31	−0.25
Experiment[c]	−0.27	−0.24

[a] Calculated and measured shifts are given with respect to the corresponding values for isolated chromophore. [b] All calculations apply the STO TZ2P basis sets. [c] Experimental data taken from ref. 22 and 85.

in this case. Such trend is in agreement with earlier studies by Neugebauer *et al.*,[18] where convergence of the solvent shift in acetone as the function of number of water molecules in environment was examined. In that case, the effect of the bulk solvent was reached including at least 50 H_2O molecules.

Finally, it is worthwhile to notice that the FDET shifts are less affected by the change of the basis set than the absolute values of the excitation energy. For the isolated chromophore, addition of diffuse functions affects the calculated vertical excitation energy by 0.06 eV shifting it from 2.80 to 2.86 eV, whereas the effect on the shifts is significantly smaller. It does not exceed 0.02 eV (see Table 1 and 2). A more detailed discussion of the stability of the FDET results with respect to the choice of the basis sets is addressed elsewhere.[51] The calculated LR-TDDFT vertical excitation energy for the isolated chromophore (either 2.80 or 2.86 eV) is, however, lower than the estimates for this quantity reported in the literature. Goerigk and Grimme[8] reported the value of 3.51 eV based on a combined analysis of experimental energy of the 0–0 excitation in solavted aminocoumarine 153 and theoretical estimates for various contributions due to the solvent. The value reported by Improta *et al.*[9] (3.18 eV) is also higher than the vertical excitation energy of the isolated chromophore obtained in the present work. This suggests that the absolute values of all vertical excitation energies for isolated and embedded chromophore are underestimated due to deficiencies of the used approximations in the LR-TDDFT calculations.

5. Conclusions

The underlying approximation used in this work consists of the simplification made in averaging instantaneous excitations over the statistical ensemble introduced in our previous work.[23] Instead of the conventional strategy, in which the instantaneous excitations for each conformation are averaged over the statistical ensemble, the average excitation energy is evaluated only once for the solvent represented by a non-uniform average charge density ($\langle\rho_B(\vec{r})\rangle$). This approximation is given in eqn (14) and has a clear interpretation owing to the fact that all the equations in the frozen-density embedding theory[13–17,53] for both ground and excited states depend parametrically on the electron density of the environment ρ_B. Opposite to most of other applications of frozen-density embedding theory based methods,[19,24–27] the frozen density used in the present work $\langle\rho_B(\vec{r})\rangle$ does not correspond to the ground-state electron density of any molecular system. $\langle\rho_B(\vec{r})\rangle$ is the ensemble (or time) average of the electron density of the solvent surrounding the chromophore—a quantity observable at both microscopic and macroscopic scales.

It is comprehensively shown that the used method leads to accurate solvatochromic shifts. The average absolute deviation between the calculated and measured shifts amounts to about 0.02 eV. The calculated values are numerically stable and are not significantly affected even by inclusion of diffuse functions in the atomic basis sets. The plots of the net charge density of the solvent reveal the ordering effect of the chromophore on the solvent molecules. The topology of the positively and negatively charged regions around the chrompohore derived from "dressed up" 3D-RISM site probabilities reflects the structure of the solvatation shells. In the case of solvents capable of hydrogen-bonding to the chromophore, the solvent is highly structured around the carbonyl group of the coumarin, which shows as strongly localized positively charged regions, which in turn stabilize the LUMO. For all studied solvents, stabilizing LUMO appears to lie at the origin of the redshift in the lowest-energy absorption band. Interestingly, the solvent charge density is less structured near the -CF_3 the group of coumarin 153.

The non-uniformity of the average charge density of the solvent leads to strong variations of the embedding potential (the orbital-free embedding potential), which includes both electrostatic and non-electrostatic component. The plots of this highly non-uniform potential combined with the analysis of the localization of HOMO and LUMO are shown to be very useful in qualitative analyses of the effect of the solvent on the excitation energies.

Calculations using explicit models of the solvent (clusters including 1–3 solvent molecules most tightly bound to the chromophore) made for water and for methanol, lead also to the red-shifted excitation energies but the magnitude of the shifts varies strongly with the number and geometry of explicit solvent molecules.

The results of the present work indicate clearly that the statistical nature of the solvatochromc shifts has to be taken into account in modelling the solvatochromic shifts. The proposed approach to deal with the flexible solvents is closely related to that used by Galvan *et al.*[63,64] Opposite, however, to the approach by Galvan *et al.*, where *only electrostatic component* of the whole embedding operator is averaged, whereas *the whole orbital-free embedding potential* is averaged in this work. This involves more costly calculations as due its additional non-electrostatic component which is a bi-functional of two electron densities: that of the solvent and that of the solute. As a result, it changes if the solute density changes which leads to an additional contribution to the response kernel in LR-TDDFT calculations. This expense, does not affect the numerical results significantly if a small basis sets are used for the solute. It results, however, with increased stability of the calculated shifts with respect to the basis sets reflecting thus the variational origin of the frozen-density embedding theory.

Finally, we notice that the approximants for the non-additive components of the orbital-free embedding potential (exchange–correlation and kinetic) in the present work are derived from local density approximation. If ρ_B is an atomic or molecular electron density, such approximants for the non-additive

kinetic energy (and potential) are known to be less adequate than the ones involving electron density gradients.[42,86] The averaged density $\langle \rho_B \rangle$ used in the present work is much more uniform than any molecular electron density as it is smeared all over around the solute and does not include nuclear cusps causing particular concern in FDET based methods.[42,80,87] If ρ_B corresponds to the ground-state electron density for some polyatomic system in Born–Oppenheimer approximation, this problem can be avoided either by enforcing the total emebedding potential to the desired asymptotic behavior[88] or designing an approximation for the non-additive kinetic potential[42] fulfilling the known exact condition of relevance near nuclear cusps. The advantages of using local density approximation for this purpose is both practical (small computational effort) and formal (the orbital-free embedding potential is obtained from the first principles as a functional of a pair of electron densities).

Acknowledgements

This work was supported by the grant 200020-124817 from the Fonds National Suisse de la Recherche Scientifique.

References

1 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227.
2 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198.
3 J. Sauer, P. Ugliengo, E. Garrone and V. Sounders, *Chem. Rev.*, 1994, **94**, 2095.
4 J. Gao, Methods and applications of combined quantum mechanical and molecular mechanical potentials, in *Reviews in Computational Chemistry*, VCH, New York, 1996, vol. 7, p. 119–185.
5 J. Gao and K. Byun, *Theor. Chem. Acc.*, 1997, **96**, 151.
6 M. E. Martin, A. M. Losa, I. Fdez-Galvan and M. A. Aguilar, *J. Chem. Phys.*, 2004, **121**, 3710.
7 K. Aidas, A. Mogelhoj, E. J. K. Nilsson, M. S. Johnson, K. V. Mikkelsen, O. Christiansen, P. Soderhjelm and J. Kongsted, *J. Chem. Phys.*, 2008, **128**, 194503.
8 L. Goerigk and S. Grimme, *J. Chem. Phys.*, 2010, **132**, 184103.
9 R. Improta, V. Barone and F. Santoro, *Angew. Chem.*, 2007, **119**, 409.
10 A. Laio, J. VandeVondele and U. Rothlisberger, *J. Chem. Phys.*, 2002, **116**, 6941.
11 G. Fradelos and T. A. Wesolowski, *J. Chem. Theory Comput.*, 2011, **7**, 213.
12 G. Fradelos and Wesolowski, to be published.
13 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050.
14 T. A. Wesolowski, *One-Electron Equations for Embedded Electron Density: Challenge for Theory and Practical Payoffs in Multi-level Modelling of Soft Condensed Matter*, World Scientific, 2006, vol. X, pp. 1–82.
15 T. A. Wesolowski, *J. Am. Chem. Soc.*, 2004, **126**, 11444.
16 T. A. Wesolowski, *Phys. Rev. A*, 2008, **77**, 012504.
17 K. Pernal and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2009, **109**, 2520.
18 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 094115.
19 J. Neugebauer, C. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805.
20 R. J. Cave and E. W. C. Jr., *J. Phys. Chem. A*, 2002, **106**, 12117.
21 M. Maroncelli and G. R. Fleming, *J. Chem. Phys.*, 1987, **86**, 6221.
22 M. L. Horng, J. A. Gradecki, A. Papazyan and M. Maroncelli, *J. Phys. Chem.*, 1995, **99**, 17311.
23 J. W. Kaminski, S. Gusarov, T. A. Wesolowski and A. Kovalenko, *J. Phys. Chem. A*, 2010, **114**, 6082.
24 E. V. Stefanovich and T. N. Truong, *J. Chem. Phys.*, 1996, **104**, 2946.
25 N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 1999, **110**, 7677.
26 M. Hodak, W. Lu and J. Bernholc, *J. Chem. Phys.*, 2008, **128**, 014101.
27 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353.
28 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.
29 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.
30 P. Cortona, *Phys. Rev. B*, 1991, **44**, 8454.
31 G. Senatore and K. Subbaswamy, *Phys. Rev. B*, 1986, **34**, 5754.
32 P. Elliott, M. H. Cohen, A. Wasserman and K. Burke, *J. Chem. Theory Comput.*, 2009, **5**, 827.
33 T. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1994, **98**, 5183.
34 M. Zbiri, A. Atanasov, C. Daul, J. M. Garcia-Lastra and T. A. Wesolowski, *Chem. Phys. Lett.*, 2004, **397**, 441.
35 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71.
36 T. A. Wesolowski and F. Tran, *J. Chem. Phys.*, 2003, **118**, 2072.
37 R. Kevorkyants, M. Dulak and T. A. Wesolowski, *J. Chem. Phys.*, 2006, **124**, 024104.
38 M. Dulak, J. W. Kaminski and T. A. Wesolowski, *J. Chem. Theory Comput.*, 2007, **3**, 735.
39 M. Iannuzzi, B. Kirchner and J. Hutter, *Chem. Phys. Lett.*, 2006, **421**, 16.
40 M. Dulak, J. W. Kaminski and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2009, **109**, 1886.
41 T. A. Wesolowski and J. Weber, *Int. J. Quantum Chem.*, 1997, **61**, 303.
42 J. M. G. Lastra, J. W. Kaminski and T. A. Wesolowski, *J. Chem. Phys.*, 2008, **129**, 074107.
43 Y. A. Bernard, M. Dulak, J. W. Kaminski and T. A. Wesolowski, *J. Phys. A*, 2008, **41**, 055302.
44 G. Fradelos, J. W. Kaminski, T. A. Wesolowski and S. Leutwyler, *J. Phys. Chem. A*, 2009, **113**, 9766.
45 T. A. Wesolowski, *Chem. Phys. Lett.*, 1999, **311**, 87.
46 J. Neugebauer, M. J. Louwerse, P. Belanzoni, T. A. Wesolowski and E. J. Baerends, *J. Chem. Phys.*, 2005, **123**, 114101.
47 C. R. Jacob and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 194104.
48 C. J. Jacob, J. Neugebauer, L. Jensen and L. Visscher, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2349.
49 M. E. Casida, Time-dependent density-functional response theory for molecules, in *Recent Advances in Density Functional Methods*, World Scientific, Singapore, 1995, vol. 1, p. 155–192.
50 G. Fradelos, J. W. Kaminski, T. A. Wesolowski and S. Leutwyler, *J. Phys. Chem. A*, 2009, **113**, 9766.
51 G. Fradelos, J. Lutz, M. Wloch, T. A. Wesolowski and P. Piecuch, to be published.
52 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.
53 M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2004, **96**, 577.
54 A. Kovalenko, *Three-dimensional RISM theory for molecular liquids and solid–liquid interfaces in "Understanding Chemical Reactivity"*, Kluwer Acedemic Publishers, Dordrecht, The Netherlands, 2003, vol. 24, pp. 169–275.
55 A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 1999, **110**, 10095.
56 J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic, London, 2nd edn, 1986.
57 J. S. Perkyns and B. M. Pettitt, *Chem. Phys. Lett.*, 1992, **97**, 7656.
58 S. Gusarov, T. Ziegler and A. Kovalenko, *J. Phys. Chem. A*, 2006, **110**, 6083.
59 D. Casanova, S. Gusarov, A. Kovalenko and T. Ziegler, *J. Chem. Theory Comput.*, 2007, **3**, 458.
60 S. Ten-no, F. Hirata and S. Kato, *Chem. Phys. Lett.*, 1993, **214**, 391.
61 S. Ten-no, F. Hirata and S. Kato, *J. Chem. Phys.*, 1994, **100**, 7443.
62 H. Sato, F. Hirata and S. Kato, *J. Chem. Phys.*, 1996, **105**, 1546.
63 I. F. Galvan, M. L. Sanchez, M. E. Martin, F. J. O. del Valle and M. A. Aguilar, *Comput. Phys. Commun.*, 2003, **155**, 244.
64 K. Coutinho, H. C. Georg, T. L. Fonseca, V. Ludwig and S. Canuto, *Chem. Phys. Lett.*, 2007, **437**, 148.
65 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
66 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822.
67 E. J. Baerends, J. Autschbach, A. Bèrces, F. M. Bickelhaupt, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. Fonseca Guerra, S. J. A. van Gisbergen,

A. W. Gotz, J. A. Groeneveld, O. V. Gritsenko, M. Grning, F. E. Harris, P. van den Hoek, C. R. Jacob, H. Jacobsen, L. Jensen, G. van Kessel, F. Kootstra, M. V. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, J. Neugebauer, V. P. Nicu, V. P. Osinga, S. Patchkovskii, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, J. I. Rodriguez, P. Ros, P. R. T. Schipper, G. Schreckenbach, J. G. Snijders, M. Solá, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. L. Yakovlev and T. Ziegler, *ADF2008.01, SCM, Theoretical Chemistry*, Vrije Universiteit, Amsterdam, The Netherlands, 2008, http://www.scm.com.

68 C. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011.
69 M. Swart, P. T. van Duijnen and J. G. Snijders, *J. Comput. Chem.*, 2001, **22**, 79.
70 P. A. M. Dirac, *Math. Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.
71 D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, 1980, **45**, 566.
72 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200.
73 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.
74 W. L. Jorgensen, J. D. Madure and C. J. Swenson, *J. Am. Chem. Soc.*, 1984, **106**, 6638.
75 L. H. Thomas, *Math. Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
76 E. Fermi, *Z. Phys.*, 1928, **48**, 73.
77 P. Schipper, O. Gritsenko, S. van Gisbergen and E. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344.
78 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Chem. Phys. Lett.*, 1999, **302**, 199.
79 A. Savin and T. A. Wesolowski, *Prog. Theor. Chem. Phys.*, 2009, **19**, 327.
80 M. Dulak and T. A. Wesolowski, *J. Chem. Phys.*, 2006, **124**, 164101.
81 C. R. Jacob, S. M. Beyhan and L. Visscher, *J. Chem. Phys.*, 2007, **126**, 234116.
82 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999.
83 J. Tomasi, *Theor. Chem. Acc.*, 2004, **112**, 184.
84 A. Klamt and G. Shuurmann, *J. Chem. Soc. Trans.*, 1993, **2**, 799.
85 P. Hazra, D. Chakrabarty and N. Sarkar, *Chem. Phys. Lett.*, 2003, **371**, 553.
86 T. A. Wesolowski, *J. Chem. Phys.*, 1997, **106**, 8516.
87 M. Dulak, T. A. Wesolowski, 2006, unpublished.
88 C. R. Jacob, S. M. Beyhan and L. Visscher, *J. Chem. Phys.*, 2007, **126**, 234116.

Chapter 6
Spectral tuning of rhodopsin and visual cone pigments

"If I have seen farther, it is by standing on the shoulders of giants."
— Sir Isaac Newton

6.1 Overall presentation of the article

As described in the introduction of this thesis work (Chapter 1), the spectral tuning of the retinal chromophore in different protein environments (rhodopsin pigment, red-, green-, and blue-cone pigments) provides the basis of the visual perception of human. This work aims to understand the origin of the spectral tuning by investigating the $\pi \rightarrow \pi^*$ vertical electronic excitations of retinal for different models of retinal in the protein environments, and to investigate the effect of amino acid mutation in rhodopsin thus to provide a basis for rational photobiological design of proteins with specifically tuned absorption properties.

In this work, global/local energy minimum structures for each models are used for the calculating the vertical electronic excitations. The modeling of the structures were performed by Ville Kaila and Dage Sundholm. The calculation of the vertical electronic excitations were done by Xiuwen Zhou and Tomasz A. Wesolowski.

6.2 Reprint of the article

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Spectral Tuning of Rhodopsin and Visual Cone Pigments

Xiuwen Zhou,[†] Dage Sundholm,[‡] Tomasz A. Wesołowski,[†] and Ville R. I. Kaila*[,§]

[†]Département de Chimie Physique, Université de Genève, 30 quai Ernest-Ansermett, CH-1211 Genève 4, Switzerland

[‡]Department of Chemistry, P.O. Box 55, A. I. Virtanens plats 1, University of Helsinki, FIN-00014 Helsinki, Finland

[§]Department Chemie, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

Supporting Information

ABSTRACT: Retinal is the light-absorbing biochromophore responsible for the activation of vision pigments and light-driven ion pumps. Nature has evolved molecular tuning mechanisms that significantly shift the optical properties of the retinal pigments to enable their absorption of visible light. Using large-scale quantum chemical calculations at the density functional theory level combined with frozen density embedding theory, we show here how the protein environment of vision pigments tunes the absorption of retinal by electrostatically dominated interactions between the chromophore and the surrounding protein residues. The calculations accurately reproduce the experimental absorption maxima of rhodopsin and the red, green, and blue color pigments. We further identify key interactions responsible for the color-shifting effects by mutating the rhodopsin structure *in silico*, and we find that deprotonation of the retinyl is likely to be responsible for the blue-shifted absorption in the blue cone vision pigment.

Retinal is a conjugated polyene that occurs in the light-capturing unit of several photobiological systems. In vision pigments[1] and bacterial light-driven proton pumps,[2] retinal is covalently linked to the protein by a lysine residue, forming a Schiff base (SB). The protein environment shifts the absorption maximum of retinal from 365−430 nm (2.80−3.40 eV) in aqueous solution to 420−560 nm (2.20−2.95 eV) in the vision proteins,[1] enabling their absorption of visible light. The exact molecular mechanism of the spectral shift has remained elusive for more than half a century. It has been suggested that the tuning may arise from an altered conjugation of the polyene,[3] by specific electrostatic interactions between protein residues and the retinal,[4] and by charge transfer and polarization effects.[5] The development of accurate electronic structure theory methods open up new ways of addressing the molecular mechanism of spectral tuning.

In this study, we investigate the protein-induced spectral shifts of retinal in rhodopsin and its homologous color cone pigments using large-scale quantum chemical calculations. Rhodopsin is a protein in the rod cells of the vertebrate eye, responsible for dim vision.[1] Color vision takes place in the cone cells and is catalyzed by three color pigment proteins, responsible for the absorption of red, green, and blue photons, respectively.[1b] Light absorption by these G-protein coupled receptors leads to an 11-*cis* to all-*trans* isomerization of the retinyl side chain, activating a G-protein-mediated signaling cascade that triggers the vision process.[1]

Photobiological systems face unique computational challenges due to their complex chromophore−protein environment, which must be explicitly considered using large computational models.[6] Although *ab initio* methodologies can accurately predict optical transitions in molecules, most such methods are inapplicable to photobiology due to their high computational costs. Recent developments, such as the restricted virtual space approach in combination with low-order correlation methods, increase the possibility of treating large photobiological systems.[4c,7] However, due to the high computational scaling of such methods, extensive studies of the chromophore−protein interactions beyond the immediate chromophore vicinity are demanding.

We use here a frozen-density embedding theory (FDET)[8] based method to compute the vertical excitation energies of retinal embedded in large protein surroundings, within the linear-response time-dependent density functional theory (TDDFT)[9] framework. In these calculations, we treat the chromophore region as an active system that is quantum chemically embedded in a frozen electron density of surrounding protein residues. Due to the large computational savings introduced by treating the surroundings as a frozen electron density, the FDET approach allows the modeling of the chromophore−protein interactions at full quantum mechanical level, using system sizes comprising ∼400 atoms, usually beyond the capabilities of conventional TDDFT methods, especially when a large number of calculations are necessary, as in this work.

Molecular models of rhodopsin and of the red, green, and blue cone pigments were constructed on the basis of coordinates of the crystal structure from *Bos taurus*[11] and the homology models obtained from Brookhaven Protein Data Bank (PDB IDs: 1U19, 1KPX, 1KPN, 1KPW).[12] The models comprised 329−370 atoms, with the retinal surrounded by 25−30 residues nearest to the chromophore binding pocket (see SI Table 1, SI Figures 1 and 2). All amino acid residues were cut at the C_β atoms, which were saturated by hydrogen atoms. The models were structure optimized using the BP86 functional with the RI-MARIJ approximation and def2-SVP basis sets.[13] The retinal side chain and hydrogen atoms in the surrounding residues were allowed to fully relax in the structure optimization. To study the saturation of these models, the CHARMM27 force field[14] was used to embed the quantum chemical models in the point charge surroundings of the protein residues beyond the QM model

Received: November 21, 2013
Published: January 14, 2014

   dx.doi.org/10.1021/ja411864m | *J. Am. Chem. Soc.* 2014, 136, 2723−2726

systems. Based on structural alignment, sequence comparison (SI Figure 1, SI Table 1), and the X-ray structure of rhodopsin,[11] we also constructed blue (Rh:Δblue) and red mutant (Rh:Δred) models to probe the function of key residues responsible for the tuning process.[1b] The Rh:Δblue model comprised the *in silico* mutations W265Y, Y191W, E122L, H211C, G90C, A124T, A292S, A295S, and A299C, whereas the Rh:Δred model comprised E181H, E122L, F208M, H211C, F212C, and optimized similar to the rhodopsin and cone pigment models (Cartesian coordinates are given in the SI). To study alternative protonation states of the retinal, we performed local optimizations of the SB proton with the proton constrained to reside either on Glu-113 or on the retinyl side chain. The optimizations were performed using a hybrid quantum/classical mechanics (QM/MM) approach at the B3LYP/def2-SVP/CHARMM27 level of theory.[14,15] Only the Glu-113/SB retinal was modeled for the QM system, and the remaining system was treated classically. After the QM/MM optimization, the structure of the Glu-113/SB pair was incorporated back into the large full-QM models. For computation of vertical excitation energies, the optimized structures were separated into an embedded active system and an environment region. The embedded active system comprised (i) the retinyl chromophore or (ii) the retinal and Glu-113, which were studied using the B3LYP functional[15] and Slater-type orbitals (STO) of double-zeta quality augmented with polarization functions (DZP).[16] The frozen density of the environment is generated by a Kohn−Sham calculation for the isolated environment at the BP86/DZP level. The electron density of the embedded subsystem, ρ_A, and the charge density of the embedding protein subsystem, ρ_B and ρ_B^{pos}, indicating electrons and nuclei, respectively, uniquely determine the embedding potential (v_{emb}), within the FDET framework:[8]

$$v_{emb}[\rho_A, \rho_B; r] = \int \frac{\rho_B^{pos}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' + \frac{\delta E_{xc}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})} + \frac{\delta T_s^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}$$

The nonadditive exchange-correlation component of the embedding potential was approximated using the local-density approximation for $E_{xc}[\rho]$,[17] whereas the nonadditive kinetic component was approximated using the NDSD bifunctional.[18] This leads to a robust computational protocol that reduces possible errors due to the approximations in the nonadditive kinetic energy.[19] For FDET/TDDFT calculations beyond the Neglect of Dynamic Response of the Environment approximation, the reader should consult the recent comprehensive review by Neugebauer.[20] Since the charge-transfer excitations are more sensitive to the choice of frozen density in FDET calculations than local excitations, additional calculations were performed using the CAM-B3LYP[21] functional as implemented in ADF.[22] We did not observe noticeable TDDFT charge-transfer problems for any of studied systems with retinal comprising the embedded active region (see also SI Table 2). Moreover, test calculations on a *cis*-retinal model at the coupled-cluster approximate singles and doubles (CC2) level[23] (SI Table 3) suggest that the long-range corrected density functional CAM-B3LYP consistently overestimates the excitation energies by ∼0.3 eV in comparison to the CC2 and experimental data (SI Tables 2−4). We thus treated the chromophore at the B3LYP level in all reported calculations. The electronic excitation energies of the embedded retinal subsystem were obtained using

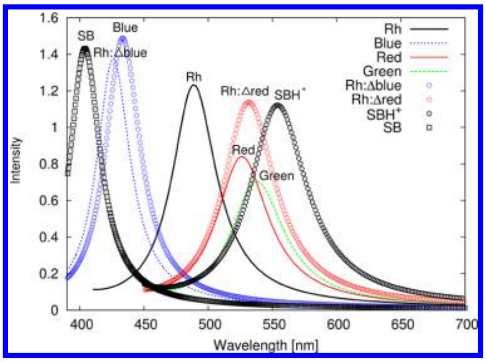


Figure 1. Computed absorption spectra of retinal models in vacuum (SBH^+/SB), and embedded in the protein surroundings of rhodopsin (Rh), the red, green, and blue photopigments (red/green/blue), and *in silico* constructed red and blue mutant pigments of rhodopsin (Rh:Δred/Δblue). The vertical excitation energies and oscillator strengths were obtained using FDET/TDDFT calculation at the B3LYP/DZP level. The intensities are Lorentz broadened with a width that is 0.5% of the frequency range and based on computed oscillator strengths.

the FDET/TDDFT method implemented[24] in ADF[22] versions 2012.01 and 2013.01. The structure optimizations were performed using TURBOMOLE[25] version 6.3 and CHARMM/Q-Chem version 4.0.[26]

The computed absorption spectra for the isolated and protein embedded retinal models are shown in Figure 1. We obtain an excitation energy of 2.54 eV (488 nm) for retinal embedded in rhodopsin, which agrees well with the experimental absorption maximum of 2.49 eV (498 nm).[29] For the red and green pigments, we obtain excitation energies of 2.36 eV (525 nm) and 2.30 eV (540 nm), respectively, whereas the blue pigment absorbs at 2.91 eV (426 nm), which is obtained by deprotonation of the SB, consistent with our previous study.[4c] We also explored different embedding strategies; the results are shown in SI Table 2. We find that the red pigment model has the largest apparent error of ∼0.16 eV in the excitation energy, within the expected error limit of TDDFT/B3LYP calculations.[7c] However, it is likely that uncertainties in the excitation energies may originate from the use of homology models for the cone pigments, for which there are uncertainties in the exact position of residues. Our excitation energies for rhodopsin obtained at the FDET/CAM-B3LYP level of theory are similar to values obtained in a recent detailed QM/MM study (see SI Table 2).[30]

The FDET calculations suggest that the protein surroundings of rhodopsin and the cone pigments absorbing red and green light *electrostatically* blue-shift the retinyl absorption by 0.3−0.45 eV (Table 1) relative to the absorption maximum of retinal in vacuum (2.27 eV/546 nm). The protein induced excitation-energy shift (ΔE_{tot}) is obtained as the difference between the energies calculated for the completely relaxed chromophore in vacuum and in the protein. The electrostatic shift (E_{elec}) is determined by removing the frozen electron density of the surrounding protein residues and keeping the chromophore structure unchanged. This also removes the small Pauli repulsion, which must be included in the FDET embedding potential to obtain meaningful interaction energies.[24] The *steric tuning contribution* (E_{steric}) is obtained by subtracting the electrostatic contribution from the total protein shift. The steric contribution red-shifts the absorption energies by 0.07−0.27 eV in all models, due to a destabilization of the ground state relative to the excited state. This suggest that electrostatic effects

dx.doi.org/10.1021/ja411864m | J. Am. Chem. Soc. 2014, 136, 2723−2726

Table 1. Calculated (E_{FDET}) and Experimental (E_{exp})[3,34] Vertical Excitation Energies (VEE, in eV) and Protein-Induced Shift (ΔE) of Models of Rhodopsin, the Red, Green, and Blue Cone Pigments, and the Blue- (Rh:Δblue) and Red-Shifted (Rh:Δred) *in Silico* Mutant Models of Rhodopsin[a]

system	state	E_{FDET}	E_{isol}	E_{exp}	ΔE_{tot} [E_{elec}/E_{steric}]
Rh	SBH^+/E^-	2.54 (488)	2.09 (593)	2.49 (498)	+0.27 [+0.45/−0.18]
red	SBH^+/E^-	2.36 (525)	2.04 (608)	2.21 (560)	+0.09 [+0.32/−0.23]
green	SBH^+/E^-	2.30 (539)	2.00 (620)	2.32 (534)	+0.03 [+0.30/−0.27]
blue	SB/EH	2.91 (426)	2.90 (428)	2.92 (425)	−0.14 [+0.01/−0.15]
Rh:Δred	SBH^+/E^-	2.33 (532)	2.20 (564)	–	+0.06 [+0.13/−0.07]
Rh:Δblue	SB/EH	2.86 (434)	2.86 (434)	–	−0.19 [+0.00/−0.19]
retinal	SBH^+	–	2.27 (546)	2.03 (610)	–
retinal	SB	–	3.05 (407)	–	–

[a]The calculated values are obtained at the FDET/TDDFT (B3LYP/DZP) level for protein models, with retinal in the embedded active system. E_{isol} is the VEE of the protein-environment-free chromophore, at TDDFT level. SH^+ and S refer to the protonated and deprotonated Schiff base retinal, and E^-/EH to the protonation state of Glu-113.

dominate the spectral tuning of the protonated SB retinal, consistent with earlier results by Coto et al.[31] and Hasegawa et al.[32] However, our calculations indicate that the steric contribution becomes dominating for the blue-cone model, which is electrostatically tuned by only 0.01 eV due to deprotonation of the SB retinal.

The electrostatic tuning mechanism originates from the photophysical properties of the retinal chromophore. The frontier orbitals involved in the excitation, shown in Figure 2, suggest that the retinal excitation has a $\pi\rightarrow\pi^*$ character and results in a redistribution of the charge with the positive charge of the SB transferred toward the β-ionine ring (SI Table 5), a well-known property for retinal.[4,33] This charge transfer makes the retinyl susceptible for electrostatic stabilization by charges and dipoles in the surrounding protein residues. Positive charges or dipoles near the SB, or negative charges near the β-ionine group are found to cause a red shift by stabilizing the excited state with respect to the ground state. In contrast, negative charges and dipoles near the SB, or positive charges near the β-ionine group lead to a blue shift by stabilizing the ground state. Consistent with these findings, we observe that in the *in silico* mutated Rh:Δred model, an increased number of cystein and methionine residues near the β-ionine ring has a red-shifting effect of 0.21 eV, shifting the absorption maximum from 488 to 532 nm. Consistently, we find in the *in silico* mutated Rh:Δblue model that replacing nonpolar residues (G90, A292, A295) near the SB with polar serine and threonine residues has a blue-shifting effect of 0.32 eV, shifting the absorption maximum from 488 to 434 nm.

The blue-shifting effect is strengthened by removing polar residues near the β-ionine and decreasing the aromatic stacking of the retinyl by the W265Y substitution, which may lead to a decreased π-cation interaction in the excited state. However, the interpretation of the tuning effects in Rh:Δblue is more complex as compared to the Rh:Δred model. Similarly to the blue cone pigment, the retinyl SB in the Rh:Δblue becomes deprotonated in the geometry optimization, which most likely has the largest

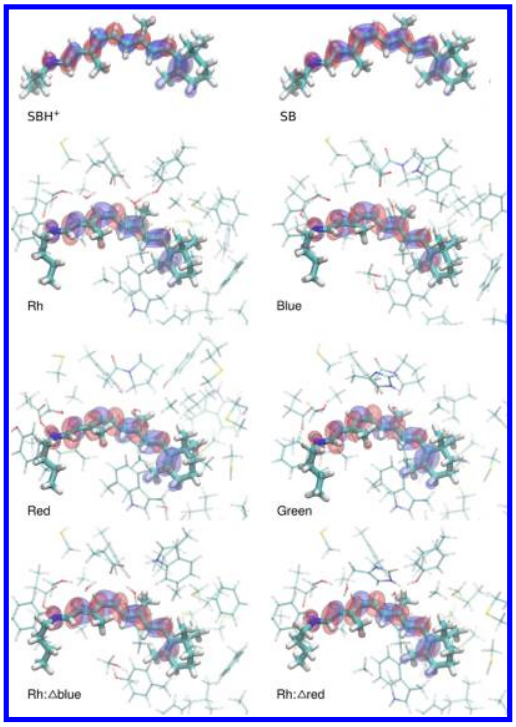


Figure 2. Frontier orbitals involved in the $\pi\rightarrow\pi^*$ photoexcitations of retinal models in vacuum (SBH^+/SB), and embedded in the protein surroundings of rhodopsin (Rh), the red, green, and blue photopigments (red/green/blue), and *in silico* constructed red and blue mutant pigments of rhodopsin (Rh:Δred/Δblue). The figure was prepared using VMD.[27]

blue-shifting effect. The excitation in the deprotonated SB retinal leads to a significantly smaller charge separation than that for the protonated retinal (SI Table 5). Thus, the deprotonated retinal is less sensitive to electrostatic tuning effects, consistently with the larger steric tuning contribution shown in Table 1. The red and blue shifts achieved by substituting the adjacent amino acids are consistent with the spectral shifts observed in previous site-directed mutagenesis experiments.[1b,28] Moreover, the observed spectral shifts and tuning effects of the red-pigment model are also consistent with the retinal−charge/dipole model calculations of the protonated SB retinal shown in SI Figure 2.

To better understand the balance between red- and blue-shifting effects in the tuning process, we performed additional FDET calculations on protein models with Glu-113 removed. In agreement with the pioneering study of rhodopsin tuning by Coto et al.,[31] we find that Glu-113 imposes a blue-shifting effect, while the remaining protein environment causes a red shift (SI Table 6). Our FDET calculations suggest that the blue-shifting effect of Glu-113 varies between 0.2 and 0.4 eV (SI Table 6), thus indicating that the remaining protein environment tunes the blue-shifting effect of this residue. We analyzed the electrostatic potential (ESP) charges of Glu-113 in the different models, and found a variation of up to 0.2*e* (SI Figure 3), suggesting that the remaining protein environment indeed imposes a secondary polarization effect, which may in turn modulate the blue-shifting effect of Glu-113. Different electrostatic polarization of the retinyl side chain was previously described for the pigment models by Hasegawa et al.,[32] who suggested that the polarization of the ESP along the retinyl backbone causes a shift in the LUMO energy, thus changing the excitation energy. They found a somewhat higher blue-shifting effect of 0.7 eV for Glu-113, but

dx.doi.org/10.1021/ja411864m | J. Am. Chem. Soc. 2014, 136, 2723−2726

only a small variation in this effect among the different cone pigments, which may relate to the smaller QM region used in their calculations.

In summary, we discuss here a molecular basis for understanding the mechanism of the spectral tuning in vision pigments. Using large-scale quantum chemical FDET calculations and the TDDFT formalism, our computed vertical excitation energies are in quantitative agreement with experimental absorption maxima. We find that the protein-induced shifts are dominated by electrostatic interactions for the models with a protonated SB retinal, and that for them the observed tuning effects can be explained in terms of an electrostatic interaction model. We find that negative charges and dipoles near the β-ionine ring stabilizes the excited state causing a red-shifting effect, and that negative protein charges and dipoles near the SB stabilize the ground state relatively to the excited state leading to a blue shift on the absorption spectrum. We also find that the strongly blue-shifting effect of Glu-113 is modulated by the remaining protein environment in the different visual pigments. Moreover, the calculations predict that a deprotonated chromophore is responsible for the photon absorption of the blue-cone pigment. The presented large-scale quantum chemical calculations may form a basis for a rational photobiological design of proteins with specifically tuned absorption properties.[35]

■ ASSOCIATED CONTENT

Supporting Information

SI Tables 1−5, SI Figures 1−3, and coordinates of studied systems. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ville.kaila@ch.tum.de

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The Biowulf cluster at NIH and CSC—the Finnish IT Center for Science—are acknowledged for computer time. This research was supported by the Academy of Finland through projects (137460 and 266227) and its Computational Science Research Programme (258258). Grants from Swiss National Science Foundation (200020/134791/1 FNRS), COST Action CM1002, and Magnus Ehrnrooth Foundation are greatly appreciated.

■ REFERENCES

(1) (a) Palczewski, K. *Annu. Rev. Biochem.* **2006**, *75*, 743. (b) Kochendoerfera, G. G.; Lin, S. W.; Sakmar, T. P.; Mathies, R. A. *Trends Biochem. Sci.* **1999**, *24*, 300.

(2) Lanyi, J. K. *Annu. Rev. Physiol.* **2004**, *66*, 665.

(3) Blatz, P. E.; Liebman, P. *Exp. Eye Res.* **1973**, *17*, 573.

(4) (a) Wanko, M.; et al. *J. Phys. Chem. B* **2005**, *109*, 3606. (b) Andruniow, T.; Ferre, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17908. (c) Kaila, V. R. I.; Send, R.; Sundholm, D. *J. Phys. Chem. B* **2012**, *116*, 2249.

(5) Baasov, T.; Friedman, N.; Sheves, M. *Biochemistry* **1987**, *26*, 3210.

(6) Neugebauer, J. *Chem. Phys. Chem.* **2009**, *10*, 3148.

(7) (a) Send, R.; Kaila, V. R. I.; Sundholm, D. *J. Chem. Phys.* **2011**, *134*, 214114. (b) Kaila, V. R. I.; Send, R.; Sundholm, D. *Phys. Chem. Chem. Phys.* **2013**, 4491. (c) Send, R.; Kaila, V. R. I.; Sundholm, D. *J. Chem. Theory Comput.* **2011**, *799*, 2473.

(8) (a) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050. (b) Wesolowski, T. A. *Computational Chemistry: Reviews of Current Trends*, Vol. *X*; World Scientific, 2006. (c) Wesolowski, T. A. *Phys. Rev. A* **2008**, *77*, 012504. (d) Pernal, K.; Wesolowski, T. A. *Int. J. Quantum Chem.* **2009**, *109*, 2520.

(9) (a) Casida, M. E. In *Recent Advances in Density-Functional Methods, Part I: Time-dependent density-functional response theory for molecules*; Chong, D. P., Ed.; World Scientific: Singapore, 1995. (b) Casida, M. E. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 3. (c) Casida, M. E.; Wesolowski, T. A. *Int. J. Quantum Chem.* **2004**, *96*, 577.

(10) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009.

(11) Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V. *J. Mol. Biol.* **2004**, *342*, 571.

(12) Stenkamp, R. E.; Filipek, S.; Driessen, C. A.; Teller, D. C.; Palczewski, K. *Biochim. Biophys. Acta* **2002**, *1565*, 168.

(13) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822. (c) Sierka, M.; Hogekamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136. (d) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

(14) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586.

(15) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(16) van Lenthe, E.; Baerends, E. J. *J. Comput. Chem.* **2003**, *24*, 1142.

(17) (a) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864. (b) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(18) Garcia Lastra, J. M.; Kaminski, J. W.; Wesolowski, T. A. *J. Chem. Phys.* **2008**, *129*, 074107.

(19) Humbert-Droz, M.; Zhou, X.; Shedge, S. V.; Wesolowski, T. A. *Theor. Chem. Acc.* **2013**, *132*, 1405.

(20) Neugebauer, J. *Phys. Rep.* **2010**, *489*, 1.

(21) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.

(22) ADF, SCM: Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.

(23) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.

(24) (a) Wesolowski, T. A. *J. Am. Chem. Soc.* **2004**, *126*, 11444. (b) Dulak, M.; Wesolowski, T. A. *Int. J. Quantum Chem.* **2005**, *101*, 543. (c) Jacob, C. R.; Neugebauer, J.; Visscher, L. *J. Comput. Chem.* **2008**, *29*, 1011.

(25) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. (http://www.turbomole.com)

(26) (a) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545. (b) Shao, Y. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172. (c) Woodcock, H. L., III; et al. *J. Comput. Chem.* **2007**, *28*, 1485.

(27) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(28) Lin, S. W.; Kochendoerfer, G. G.; Carroll, K. S.; Wang, D.; Mathies, R. A.; Sakmar, T. P. *J. Biol. Chem.* **1998**, *273*, 24583.

(29) Wald, G.; Brown, P. K. *J. Gen. Physiol.* **1953**, *37*, 189.

(30) Valsson, O.; Campomanes, P.; Tavernelli, I.; Rothlisberger, U.; Filippi, C. *J. Chem. Theory Comput.* **2013**, *9*, 2441.

(31) Coto, P. B.; Strambi, A.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1714.

(32) Fujimoto, K.; Hasegawa, J.; Nakatsuji, H. *Chem. Phys. Lett.* **2008**, *462*, 318.

(33) (a) Bonacic-Koutecky, V.; Köhler, K.; Michl, J. *Chem. Phys. Lett.* **1984**, *104*, 440. (b) Birge, R. R.; Murray, L. P.; Pierce, M. M.; Akita, H.; Balogh-Nair, V.; Findsen, L. A.; Nakanishi, K. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4117.

(34) (a) Nathans, J. *Sci. Am.* **1989**, *260*, 42. (b) Oprian, D. D.; Asenjo, A. B.; Lee, N.; Pelletier, S. L. *Biochemistry* **1991**, *30*, 11367. (c) Merbs, S. L.; Nathans, J. *Nature* **1992**, *356*, 433. (d) Nielsen, I. B.; Lammich, L.; Andersen, L. H. *Phys. Rev. Lett.* **2006**, *743*, 018304.

(35) Wang, W.; et al. *Science* **2012**, *338*, 1340.

Supporting Information:
Spectral Tuning of Rhodopsin and Visual Cone Pigments

Xiuwen Zhou[1], Dage Sundholm[2], Tomasz A. Wesołowski[1] and Ville R. I. Kaila[3,*]

[1]Département de Chimie Physique, Université de Genève, 30 quai Ernest-Ansermett, CH-1211 Genève 4, Switzerland.
[2]Department of Chemistry, University of Helsinki, P.O. Box 55, A. I. Virtanens plats 1, FIN-00014 Helsinki, Finland.
[3]Department of Chemie, Technische Universität München (TUM), Lichtenbergstrasse 4, D-85747 Garching, Munich, Germany.
[*]Corresponding author: Tel: +49 89 289 13612 Fax +49 89 289 13622 E-mail: ville.kaila@ch.tum.de

SI Table 1. Protein residues included in the computational models of rhodopsin (Rh), and the Red (R), Green (G) and Blue (B) cone pigments.

Rh	G90	E113	G114	A117	T118	G121	E122	L125	C167	Y178
R	A106	E129	G130	V133	S134	G137	I138	L141	W183	Y194
G	A106	E129	G130	V133	S134	G137	I138	L141	Y194	H197
B	S87	E110	G111	G114	T115	G118	L119	G122	V164	F175

Rh	E181	S186	C187	G188	I189	Y191	M207	F208	H211	F212
R	S202	C203	G204	P205	Y219	M220	L223	M224	C227	C228
G	S202	C203	G204	P205	L223	M224	C227	C228	L232	F277
B	E178	S183	C184	P186	W188	L204		F205	C208	F209

Rh	L216	F261	W265	Y268	A269	A272	A292	F293	A295	K296
R	L232	Y277	W281	Y284	T285	A288	A308	Y309	A311	K312
G		W281	Y284	A285	A288	A308	F309	F310	A311	K312
B		F258	Y262	Y265	A266	A269	S289	F290	S292	K293

SI Table 2. Sensitivity of system embedding and saturation of the FDET calculations. The table also shows how the first bright excitation energy is affected by the long-range corrected functional, CAM-B3LYP,[21] and alternative protonation states of the retinyl chain. For systems with both the protonated retinal and Glu in the embedded region, low-lying charge transfer states appear below the first bright excited state. SBH^+ and SB refer to protonated and deprotonated Schiff base retinals, respectively.

System	State	Embedded system	Excitation energy (eV)	N_{atoms}	Exp. excitation energy (eV)	Method
Rhodopsin	SBH^+/Glu^-	SBH^+	2.54	366	2.49	B3LYP/DZP
	SBH^+/Glu^-	SBH^+, point charges	2.54			B3LYP/DZP/MM
	SBH^+/Glu^-	SBH^+	2.73			CAM-B3LYP/DZP
	SBH^+/Glu^-	SBH^+, Glu^-	2.66			B3LYP/DZP
	SB/GluH	SB	2.85			B3LYP/DZP
	SB/GluH	SB, point charges	2.82			B3LYP/DZP/MM
	SB/GluH	SB	3.04			CAM-B3LYP/DZP
Blue cone	SB/GluH	SB	2.91	356	2.92	B3LYP/DZP
	SB/GluH	SB, point charges	2.95			B3LYP/DZP/MM
	SB/GluH	SB	3.17			CAM-B3LYP/DZP
	SB/GluH	SB, GluH	2.83			B3LYP/DZP
	SBH^+/Glu^-	SBH^+	2.52			B3LYP/DZP
	SBH^+/Glu^-	SBH^+, point charges	2.60			B3LYP/DZP/MM
	SBH^+/Glu^-	SBH^+	2.82			CAM-B3LYP/DZP
Red cone	SBH^+/Glu^-	SBH^+	2.36	370	2.10	B3LYP/DZP
	SBH^+/Glu^-	SBH^+, point charges	2.37			B3LYP/DZP/MM
	SBH^+/Glu^-	SBH^+	2.67			CAM-B3LYP/DZP
	SBH^+/Glu^-	SBH^+, Glu^-	2.63			B3LYP/DZP
	SB/GluH	SB	2.82			B3LYP/DZP
	SB/GluH	SB, point charges	2.82			B3LYP/DZP/MM
	SB/GluH	SB	3.11			CAM-B3LYP/DZP
Green cone	SBH^+/Glu^-	SBH^+	2.30	329	2.32	B3LYP/DZP
	SBH^+/Glu^-	SBH^+, point charges	2.30			B3LYP/DZP/MM
	SBH^+/Glu^-	SBH^+	2.66			CAM-B3LYP/DZP
	SBH^+/Glu^-	SBH^+, Glu^-	2.50			B3LYP/DZP
	SB/GluH	SB	2.85			B3LYP/DZP
	SB/GluH	SB, point charges	2.82			B3LYP/DZP/MM
	SB/GluH	SB	3.16			CAM-B3LYP/DZP

SI Table 3. Benchmarking the performance of TDDFT on a truncated 11-*cis* retinal model (PSB11) calculated at the coupled-cluster approximate singles and doubles level (CC2)[23] and using Becke's three parameter functional, B3LYP,[13] and the long-range corrected CAM-B3LYP functional.[21]

System	Basis	CC2[a]	B3LYP	CAM-B3LYP	Experiment[b]
$PSB11^+$	def2-SVP	2.24	2.29	2.42	2.03
	def2-TZVP	2.17	2.27	2.40	
	def2-QZVP	2.16	2.26	2.39	

[a] Ref.[7] (main text)
[b] Ref.[34] (main text)

SI Table 4. Excitation energies of the isolated chromophore in their fixed protein structures used in the embedding calculations at the TDDFT/DZP level of theory using Becke's three parameter functional, B3LYP,[13] and the long-range corrected CAM-B3LYP functional.[21]

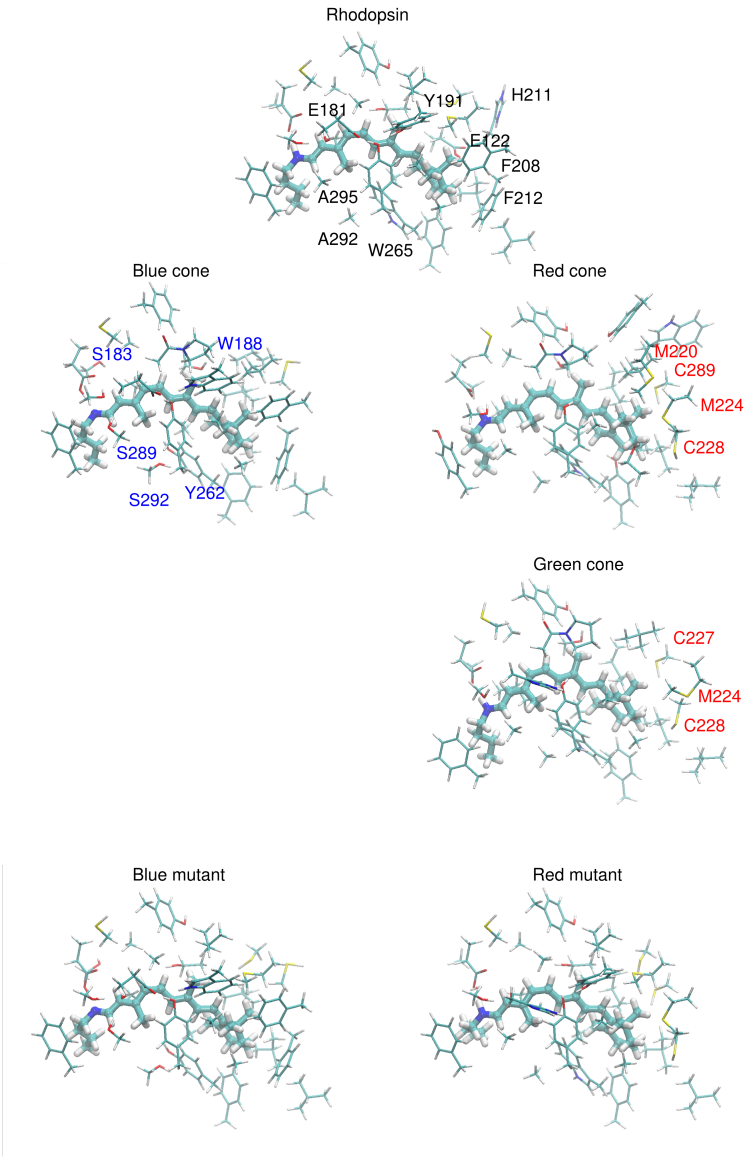
Model	**State**	**B3LYP**	**CAM-B3LYP**
Rh	SBH^+	2.09	2.34
Red	SBH^+	2.04	2.35
Green	SBH^+	2.00	2.34
Blue	SB	2.90	3.16
Rh:Δred	SBH^+	2.20	2.35
Rh:Δblue	SB	2.86	3.09
SBH^+ (BP86 optimization)	SBH^+	2.27	2.36
SB (BP86 optimization)	SB	2.90	3.15
SBH^+ (B3LYP optimization)	SBH^+	2.27	2.40
SB (B3LYP optimization)	SB	3.05	3.33

SI Table 5. Ground (GS) and excited state (ES) charge distribution of the retinal obtained from TDDFT calculations. SBH^+ and SB refer to protonated and deprotonated Schiff base retinals, respectively.

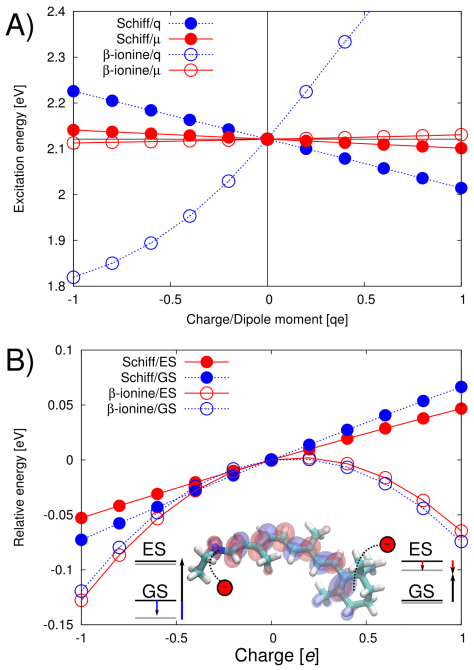
System	**Schiff base**		**Retinyl**		**β-ionine**	
	GS	ES	GS	ES	GS	ES
Retinal (SBH^+)	0.15	0.14	0.78	0.61	0.07	0.25
Retinal (SB)	0.14	0.11	-0.02	0.02	-0.12	-0.13

SI Table 6. Calculated FDET shifts of retinal VEEs (in eV) due to Glu-113.

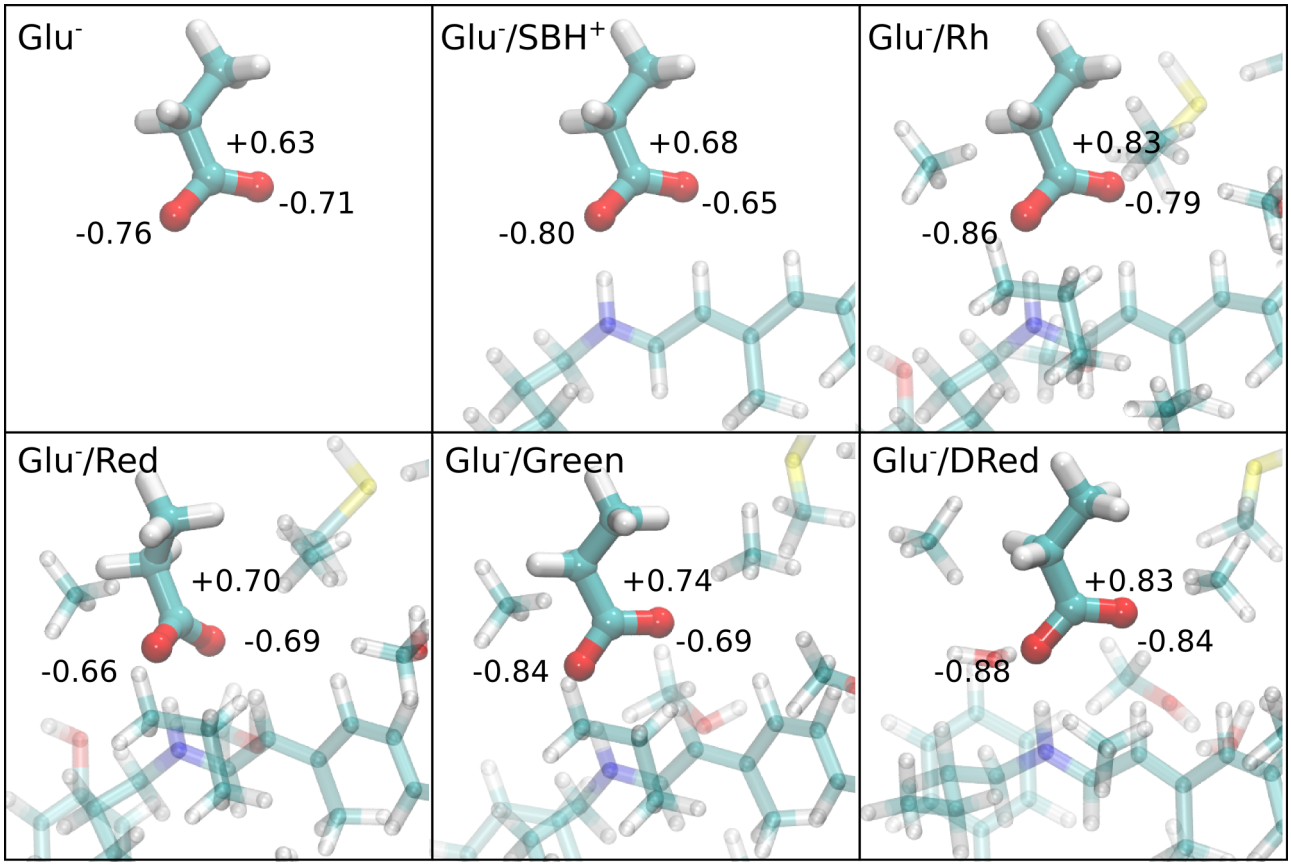
System	*Complete model*	*Glu removed*	ΔE
Rh	2.54	2.26	0.28
Red	2.36	1.98	0.38
Green	2.30	1.93	0.37
Rh:Δred	2.33	2.14	0.19
Retinal	2.45	2.27	0.18



SI Figure 1. Molecular models of Rhodopsin (top), the Blue-, Red, and Green-cone pigments (center), as well as Blue- and Red-mutant models (bottom). Residues in blue and red are expected to have blue- and red-shifting effect on the excitation energy, respectively (see text).



SI Figure 2. An electrostatic tuning model of retinal. A) TDDFT/B3LYP/def2-SVP calculations of the excitation energies of retinal interacting with a scaled point charge (q) or dipole (μ) place 4 Å from Schiff-base or β-ionine group. A negative point-charge or a dipole with its negative end oriented perpendicular to the retinal plane, blueshift the excitation energy, whereas a negative point-charge at the β-ionine end has a red-shifting effect. B) Negative charges near the Schiff base stabilizes the ground state (GS, blue filled circles) relatively to the excited state (ES, red filled circles), leading to a blueshift, whereas negative charges at the β-ionine end stabilize the ES (red open cicles) relatively to the GS (blue open circles), which leads to a redshift of the excitation energy.



SI Figure 3. Polarization of Glu-113 in different visual pigment models. The figure shows the Electrostatic Potential Charges (ESP) of Glu-113 in vacuum, and when bound to retinal, and rhodopsin (top row), and in the Red, Green and ΔRed models (bottom row). The strong polarization of Glu-113 suggests that the remaining protein environment tunes the blueshifting properties Glu-113 as indicated by FDET excitation energies of retinal in Table 2 (main text).

Chapter 7
Conclusions and perspectives

"Unless we change directions, we will end up where we are headed."
— Confucius

Conclusions

This thesis work is about the theoretical studies of the effect of condensed-phases environment on the UV-vis light absorption of organic chromophores. Three distinct types of systems have been investigated, including a functional host-guest material (fluorenone in zeolite-L channel), a prototype of solvated chemical species (coumarin 153 in various solvents), and fundamental biological systems present in the eye (retinal in rhodopsin and in three visual cone pigments).

There are two main challenges in general for this type of research. The first one is modeling the structure of molecules in condensed phases, since the size of the involved system is usually very large that commonly comprises thousands of atoms or more. Therefore simulation methods should be carefully chosen and a lot of computational effort is required. Moreover, modeling different target systems requires related knowledge and experiences. Some of the work on structure simulation reported in this thesis was done by collaborators. The second challenge is calculating the electronic excitation energy and oscillator strength of a target system, which becomes very challenging if the system is large, because quantum-mechanical methods are needed to describe the electronic excitations that are known to be limited to small systems. In this thesis work, a multi-scale method, frozen-density embedding theory, is applied to describe the electronic structure of the molecules in condensed phases. The interesting part of the total system (usually the chromophore molecule, sometimes together with its covalently bonded molecules in the environment) is treated quantum-mechanically, whereas the rest part of the system is generally described with a lower level of theory and it is characterized by its electronic density that can be approximated flexibly with various strategies according to the need of the targeted problem. Frozen-density embedding theory combined with linear-response time-dependent density-functional theory allows an accurate inclusion of the environment effect on the electronic excitations of the chromophore molecule.

For fluorenone zeolite-L system, we found that the zeolite-L frame together with the water molecules in the zeolite-L channel cause a significant red shift on the UV-vis absorption bands (around 0.6 eV for the lowest band). The red shift is found to originate from both the environment induced structural changes in fluorenone (mainly due to the elongation of C=O bond in fluorenone) and electrostatic interactions of the environment with fluorenone. For solvated coumarin 153, solvents with different polarity were shown to lead to different stabilization of the highest occupied molecular orbital and lowest unoccupied molecular orbital of coumarin 153, and thus lead to different shifts in $\pi \rightarrow \pi^*$ transition energies of coumarin 153. For retinal in the four biological pigments, the protein environments enable retinal absorb visible light mainly

via a lysine residue covalently linked to retinal and forming a Schiff base. In different pigments, the protonation state of the Schiff base and the electrostatic interaction between retinal and its neighboring amino acids are found to play important roles in shifting the absorption of retinal to different visible light spectral regions.

This thesis work confirms that frozen-density embedding theory is a successful multi-scale approach for studying molecules embedded in condensed phases. It is advantageous for several factors. First, it can treat the entire system quantum-mechanically and thus the condensed-phase effect on the electronic excitation of the embedded species can be included accurately. Second, much computational effort is saved by treating the environment as a frozen electron density. In addition, when frozen-density embedding theory is combined with LR-TDDFT, it simplifies the analysis and interpretation of the calculated results by limiting the LR-TDDFT calculations on the embedded species, which removes spurious chromophore-environment charge-transfer excitations due the approximations in exchange-correlation potential and the corresponding kernel in LR-TDDFT.

Despite the accuracy and efficiency of the approaches based on frozen-density embedding theory, we must pay careful attention to the approximations used in practical calculations. Firstly, the NDRE approximation, which leads to the reduction in the number of occupied and unoccupied Kohn–Sham orbitals for solving LR-TDDFT equations, is applicable only for systems that have little overlap in the energy levels of the two subsystems. Secondly, semi-local approximations for the non-additive kinetic energy are not applicable for embedded subsystem covalently bound to the environment, that is, the overlap between ρ_A and ρ_B is large. Thirdly, approximating the electron density of the environment should be considered as a key issue in practical frozen-density embedding theory calculations. Several summaries on this subject are made based on this thesis work since it covers several types of environment including ionic solids, solvents, and proteins. The Kohn-Sham electron density is a reliable strategy for representing the solvents and proteins but it may be very expensive for large systems. The superposition of atomic densities can be a practical and reliable strategy for generating the electron density of the ionic solids. The superposition of molecular density can be an efficient and practical strategy for qualitative analysis of the electron density of the protein environment. The statistically averaged electron density for solvents is an efficient but rather accurate way of representing the solvents for studying solvatochromism. The issue on generating ρ_B in frozen-density embedding theory calculations is comprehensively discussed in a paper [58] by the author and coworkers that is included as an appendix of this thesis. Finally, frozen-density embedding theory can account for the environment induced shift in excitation energy, but the absolute value of the excitation energy is dependent on the approximations of the functional used in LR-TDDFT. B3LYP is found to perform well for studying local excitations of the conjugated molecules investigated in this thesis work.

Perspectives

Emission is one of the ways for the excited electrons in atoms and molecules to fall back to the ground state. One of the author's on-going projects is about the solvatochromic shift in

the emission energy and the Stokes shift, which is the difference between positions of the band maxima of the absorption and emission spectra of the same electronic transition.

Modeling the structure of the molecules in condensed-phase environment is very important, because it is the prerequisite for simulating the UV-vis absorption spectra, understanding the equilibrium and energetics of the entire system, and for the analysis of the interaction between the molecules and the environment, etc. Therefore it is one of the areas in which the author would like to pursue researches in the future.

Bibliography

[1] P. W. Atkins and J. de Paula. *Physical Chemistry*. Oxford University Press, 2010.

[2] C. Pietsch, U. S. Schubert, and R. Hoogenboom. Aqueous polymeric sensors based on temperature-induced polymer phase transitions and solvatochromic dyes. *Chem. Commun.*, 47:8750–8765, 2011.

[3] Y. J. Bin, F. Q. Zhao, Z. H. Chen, and F. S. Zhang. Photo-driven molecular switch based on the photochroism and solvatochromism of diarylethene dialdehyde. *Chinese Science Bulletin*, 53(12):1813–1816, 2008.

[4] P. W. Atkins. *Quanta: a handbook of concepts*. Clarendon Press Oxford, 1974.

[5] G. C. Schatz and M. A. Ratner. *Quantum Mechanics in Chemistry*. Dover Books on Chemistry. Dover Publications, 2002.

[6] F. Jensen. *Introduction to Computational Chemistry*. Wiley, 2013.

[7] C. J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. Wiley, 2005.

[8] C. Ullrich. *Time-Dependent Density-Functional Theory: Concepts and Applications*. Oxford Graduate Texts. OUP Oxford, 2012.

[9] Erwin Schrödinger. Quantisierung als eigenwertproblem (quantification of the eigenvalue problem). *Annalen der physik (in German)*, 385(13):437–490, 1926.

[10] L. H. Thomas. The calculation of atomic fields. *Math. Proc. Cambridge Philos. Soc.*, 23:542–548, 1 1927.

[11] E. Fermi. Un metodo statistico per la determinazione di alcune priorieta dell'atome. *Rend. Accad. Naz. Lincei*, 6(602-607):32, 1927.

[12] P. A. M. Dirac. The quantum theory of the electron. *Proc. R. Soc. Lond. A*, pages 610–624, 1928.

[13] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, 1964.

[14] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci.*, 76(12):6062–6065, 1979.

[15] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[16] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin-density calculations - a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980.

[17] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048, 1981.

[18] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45(23):13244, 1992.

[19] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic-behavior. *Phys. Rev. A*, 38(6):3098–3100, SEP 1988.

[20] J. P. Perdew and W. Yue. Accurate and simple density functional for the electronic exchange energy - generalized gradient approximation. *Phys. Rev. B*, 33(12):8800–8802, JUN 1986.

[21] C. T. Lee, W. T. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B*, 37(2):785–789, JAN 1988.

[22] J. P. Perdew, P. Ziesche, and H. Eschrig. *Electronic structure of solids' 91*, volume 11. Akademie Verlag, Berlin, 1991.

[23] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, OCT 1996.

[24] A. D. Becke. A new mixing of hartree–fock and local density-functional theories. *J. Chem. Phys.*, 98(2):1372–1377, 1993.

[25] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98(45):11623–11627, 1994.

[26] S. F. Sousa, P. A. Fernandes, and M. J. Ramos. General performance of density functionals. *J. Phys. Chem. A*, 111(42):10439–10452, 2007.

[27] T. Yanai, D. P. Tew, and N. C. Handy. A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp). *Chem. Phys. L*, 393(1):51–57, 2004.

[28] P.R.T. Schipper O.V. Gritsenko and E.J. Baerends. Approximation of the exchange-correlation kohn–sham potential with a statistical average of different orbital model potentials. *Chem. Phys. L*, 302(3–4):199 – 207, 1999.

[29] M. E. Casida. *Time-dependent density-functional response theory for molecules.* Recent Advances in Density-Functional Methods. World Scientific, Singapore, Ed. D. P. Chong, 1995.

[30] E. Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. L*, 52(12):997, 1984.

[31] R. van Leeuwen. Causality and symmetry in time-dependent density-functional theory. *Phys. Rev. L*, 80(6):1280, 1998.

[32] R. Van Leeuwen. Key concepts in time-dependent density-functional theory. *Int. J. Mod. Phys. B*, 15(14):1969–2023, 2001.

[33] G. Vignale. Real-time resolution of the causality paradox of time-dependent density-functional theory. *Phys. Rev. A*, 77(6):062511, 2008.

[34] A. Castro, M. A. L. Marques, and A. Rubio. Propagators for the time-dependent kohn–sham equations. *J. Am. Chem. Soc.*, 121(8):3425–3433, 2004.

[35] F. Nogueira A. Rubio K. Burke E.K.U. Gross M.A.L. Marques, C.A. Ulrich. *Time-Dependent Density-Functional Theory, Lecture Notes in Physics*, volume 706. Springer, Berlin/Heidelberg, 2006.

[36] N. L. Allinger. Conformational analysis. 130. mm2. a hydrocarbon force field utilizing v1 and v2 torsional terms. *J. Am. Chem. Soc.*, 99(25):8127–8134, 1977.

[37] A. Warshel, M. Levitt, and S. Lifson. Consistent force field for calculation of vibrational spectra and conformations of some amides and lactam rings. *J. Mol. Spectrosc.*, 33(1):84–99, 1970.

[38] C. I. Bayly I. R. Gould K. M. Merz Jr. D. M. Ferguson D. C. Spellmeyer T. Fox J. W. Caldwell W. D. Cornell, P. Cieplak and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.

[39] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, and S. et al. Ha. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.

[40] W. R. P. Scott, P. H. Hünenberger, I. G Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren. The gromos biomolecular simulation program package. *J. Phys. Chem. A*, 103(19):3596–3607, 1999.

[41] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.

[42] A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103(2):227–249, 1976.

[43] T. A. Wesolowski and A. Warshel. Frozen density-functional approach for ab-initio calculations of solvated molecules. *J. Phys. Chem.*, 97(30):8050–8053, JUL 1993.

[44] T. A. Wesolowski. Hydrogen-bonding-induced shifts of the excitation energies in nucleic acid bases: An interplay between electrostatic and electron density overlap effects. *J. Am. Chem. Soc.*, 126(37):11444–11445, SEP 2004.

[45] T. A. Wesolowski. Embedding a multideterminantal wave function in an orbital-free environment. *Phys. Rev. A*, 77:012504, Jan 2008.

[46] K. Pernal and T. A. Wesolowski. Orbital-free effective embedding potential: Density-matrix functional theory case. *Int. J. Quant. Chem.*, 109(11, Sp. Iss. SI):2520–2525, 2009.

[47] P. Cortona. Self-consistently determined properties of solids without band- structure calculations. *Phys. Rev. B*, 44(16):8454–8458, OCT 1991.

[48] M. E. Casida and T. A. Wesolowski. Generalization of the kohn-sham equations with constrained electron density formalism and its time-dependent response theory formulation. *Int. J. Quantum Chem.*, 96(6):577–588, FEB 2004.

[49] J. Neugebauer. Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory. *J. Chem. Phys.*, 126(13):134116, 2007.

[50] A. Lembarki, F. Rogemond, and H. Chermette. Gradient-corrected exchange potential with the correct asymptotic-behavior and the corresponding exchange-energy functional obtained from the virial-theorem. *Phys. Rev. A*, 52(5):3704–3710, NOV 1995.

[51] T. A. Wesolowski. *One-electron equations for embedded electron density: challenge for theory and practical payoffs in multi-level modelling of soft condensed matter*, volume X of *Computational Chemistry: Reviews of Current Trends*. World Scientific, 2006.

[52] J. M. G. Lastra, J. W. Kaminski, and T.A. Wesolowski. Orbital-free effective embedding potential at nuclear cusps. *J. Chem. Phys.*, 129(7), 2008.

[53] T. A. Wesolowski. Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds. *J. Chem. Phys.*, 106(20):8516–8526, MAY 1997.

[54] Y. A. Bernard, M. Dulak, J. W. Kaminski, and T. A. Wesolowski. The energy-differences based exact criterion for testing approximations to the functional for the kinetic energy of non-interacting electrons. *J. Phys. A-Math. Theor.*, 41(5), 2008.

[55] S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher, and M. Reiher. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *J. Chem. Phys.*, 132(16), 2010.

[56] O. Roncero, M. P. de Lara-Castells, P. Villarreal, F. Flores, J. Ortega, M. Paniagua, and A. Aguado. An inversion technique for the calculation of embedding potentials. *J. Chem. Phys.*, 129(18):184104, 2008.

[57] J. D. Goodpaster, N. Ananth, F. R. Manby, and T. F. Miller. Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.*, 133(8):084103, 2010.

[58] M. Humbert-Droz, X. Zhou, S. V. Shedge, and T. A. Wesolowski. How to choose the frozen density in frozen-density embedding theory-based numerical simulations of local excitations? *Theor. Chem. Acc.*, 133(1):1–20, 2014.

[59] M. Zbiri, M. Atanasov, C. Daul, J. M. Garcia-Lastra, and T. A. Wesolowski. Application of the density orbital-free embedding potential functional theory derived to calculate the splitting energies of lanthanide cations in chloroelpasolite crystals. *Chem. Phys. Lett.*, 397(4-6):441–446, OCT 2004.

[60] T. A. Wesolowski and J. Weber. Kohn-sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.*, 248(1):71–76, 1996.

[61] J. W. Kaminski, S. Gusarov, T. A. Wesolowski, and A. Kovalenko. Modeling solvatochromic shifts using the orbital-free embedding potential at statistically mechanically averaged solvent density. *J. Phys. Chem. A*, 114:6082, 2010.

[62] A. Kovalenko. Three-dimensional rism theory for molecular liquids and solid-liquid interfaces. In Fumio Hirata, editor, *Molecular theory of solvation*, volume 24, pages 169–275. Springer Netherlands, 2003.

[63] A. Kovalenko and F. Hirata. Self-consistent description of a metal–water interface by the kohn–sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.*, 110(20):10095–10112, 1999.

[64] J. S. Perkyns and B. Montgomery Pettitt. A dielectrically consistent interaction site theory for solvent—electrolyte mixtures. *Chem. Phys. Lett.*, 190(6):626–630, 1992.

Appendix

"The whole of science is nothing more than a refinement of everyday thinking."
— Albert Einstein

How to choose the frozen density in Frozen-Density-Embedding-Theory-based numerical simulations for local excitations?

In Section 3.4.2, we have discussed a key issue for practical applications of FDET, that is, approximating the electron density of an environment (ρ_B). In the article by Marie Humbert-Droz, Xiuwen Zhou, Sapana V. Shedge, and Tomasz A. Wesolowski, several strategies to generate ρ_B for calculating local excitations were discussed. In this work, the environment effects on the local excitation of various systems were investigated. Xiuwen Zhou studied the solvatochromic shift in the absorption of a realistic solvation system, a 4-hydroxybenzylidene-2,3-dimethylimidazolinone anion in water cluster that comprises 50 water molecules (see Section 3.2.3 in the following reprint of this article).

Reprint of the article

Theor Chem Acc (2014) 133:1405
DOI 10.1007/s00214-013-1405-1

FEATURE ARTICLE

How to choose the frozen density in Frozen-Density Embedding Theory-based numerical simulations of local excitations?

Marie Humbert-Droz · Xiuwen Zhou · Sapana V. Shedge · Tomasz A. Wesolowski

Received: 2 July 2013 / Accepted: 2 October 2013 / Published online: 6 November 2013

Abstract According to Frozen-Density Embedding Theory, any observable evaluated for the embedded species is a functional of the frozen density (ρ_B—the density associated with the environment). The environment-induced shifts in the energies of local excitations in organic chromophores embedded in hydrogen-bonded environments are analyzed. The excitation energies obtained for ρ_B, which is derived from ground-state calculations for the whole environment applying medium quality basis sets (STO–DZP) or larger, vary in a narrow range (about 0.02 eV which is at least one order of magnitude less than the magnitude of the shift). At the same time, the ground-state dipole moment of the environment varies significantly. The lack of correlation between the calculated shift and the dipole moment of the environment reflects the fact that, in Frozen-Density Embedding Theory, the partitioning of the total density is not unique. As a consequence, such concepts as "environment polarization" are not well defined within Frozen-Density Embedding Theory. Other strategies to generate ρ_B (superposition of densities of atoms/molecules in the environment) are shown to be less robust for simulating excitation energy shifts for chromophores in environments comprising hydrogen-bonded molecules.

Keywords Frozen-Density Embedding Theory · Linear-response time-dependent density functional theory · Solvatochromism · Molecular clusters · Multi-level simulations

Electronic supplementary material The online version of this article (doi:10.1007/s00214-013-1405-1) contains supplementary material, which is available to authorized users.

M. Humbert-Droz · X. Zhou · S. V. Shedge · T. A. Wesolowski (✉)
Département de Chimie Physique, Université de Genève, 30, quai Ernest-Ansermet, 1211 Geneva 4, Switzerland
e-mail: Tomasz.Wesolowski@unige.ch

1 Introduction

The use of local potentials to couple a system described at quantum mechanical level with its environment has a long history in quantum chemistry. A large majority of methods use electrostatic potentials to this end (for review see Refs. [1–5]). The energy components not represented in the embedding potential are added as a posteriori corrections known under the name as exchange-repulsion. The use of electrostatic-only embedding potential leads, however, to well known problems such as spurious leak of charge density from the embedded subsystem to environment [6, 7] or erratic behavior of the results if the basis set changes [8, 9]. This problem originates from the fact that limiting the embedding potential to its electrostatic component neglects the Pauli exclusion principle (Pauli repulsion or confinement effects). These formal deficiencies of electrostatic-only embedding potentials, which show up in fully variational calculations (see for instance Refs. [8, 9]), are usually dealt with in a pragmatic manner. Atomic basis sets are centered only on the atoms of the embedded systems not the environment (Refs. [1–5]). As a result, the quality of the embedding potential near the nuclei of the environment plays a secondary role. Moreover, the net atomic charges (or sometimes multipoles) are used to generate the electrostatic potential. The use of net charges instead of exact total electric charge (negative) and the total nuclear charge (singular and positive) results in less attractive electrostatic potential near the nuclei in the environment. As a result, the possible unphysical distribution of the electron density between the embedded system and its

environment is less likely to take place. In the structurally flexible environment, the non-physical solutions due to the neglect of the Pauli repulsion in the embedding potential can be avoided by using the statistically averaged electrostatic potential [10–12]. Averaging of the fluctuating position of positively charged centers in the environment results in the electrostatic potential being smeared out over extended larger space. The averaged electrostatic potential is less attractive and leads to better numerical stability of the results with respect to the basis set used in variational calculations. An alternative solution to the problem of numerical instabilities originating from the neglect of Pauli repulsion consists of attenuating the Coulomb terms near the nuclei of the atoms in the environment [7] or imposing that the electron density near nuclei in the environment disappears [6]. The resulting additional terms in the embedding potential in such methods can be seen as simple local pseudopotentials (see below). Last but not least, most applications of methods applying electrostatic-only embedding potential aim at obtaining the key parameters of the ground-state potential energy surface of relevance to reactivity and/or conformational equilibria. In such a case, the energy of the system is corrected by terms added a posteriori [1–5], which are usually associated with exchange-repulsion, dispersion, and/or charge transfer represented by empirically parametrized energy contributions or terms derived from first principles [13–15]. As a result, the energy of the embedded system and its wavefunction (or a similar quantum mechanical descriptor such as a wavefunction of the reference non-interacting system used in DFT-based methods) are not self-consistent.

At the formal level, there are two ways to go beyond electrostatic-only embedding potentials in order to take into account the fermion statistics of electrons and obtain both the energy and the wavefunction which are self-consistent. Either by the use of projection operators (constructed as pseudopotentials or frozen orbitals) enforcing Pauli exclusion principle for all electrons in the whole system [5, 6, 16–19] or by the use of the non-electrostatic components in the coupling potential terms derived in Frozen-Density Embedding Theory[1] (FDET) [20–23]. The first strategy involves non-local operators and requires that the quantum mechanical descriptors, i.e., pseudopotentials or frozen orbitals, are used also for the environment. The FDET strategy, on the other hand, involves approximating some terms in the embedding potentials by means of bi-functionals depending on the charge densities associated with the embedded subsystem and the environment. Additional key issue in FDET-based simulations, which is dealt with in detail in the present work, is the choice of the frozen electron density associated with the environment (throughout this work, $\rho_B(\mathbf{r})$ denotes this component of the total electron density). Each of these two strategies has its optimal domains of applicability. The pseudopotential strategy is commonly applied to represent an environment which is a solid. For each element and oxidation state, a transferable pseudopotential can be constructed [18, 19]. For environment, which consists of molecules, the pseudopotential or frozen orbital strategies require constructing transferable pseudopotentials for each molecule (or molecular fragment) [5, 16, 17]. Concerning the FDET strategy, however, the domain of possible applications is determined by the availability of acceptable approximation for the bi-functionals by means of which the embedding potential is expressed and by the availability of an adequate inexpensive procedure to generate an adequate electron density of the environment (ρ_B).

In view of the increasing interest in using FDET-based numerical simulation methods [24–29, 31–33], this work focuses on the dependence of FDET results on $\rho_B(\mathbf{r})$ (for a recent comprehensive review see [30]). Some information about the dependency of the properties of embedded system on ρ_B can be found in the literature [34–43]. A comprehensive and systematic investigation of such dependencies was, however, not reported so far. Here we focus on the evaluation of environment-induced shifts in the energies of local excitations evaluated following the method based on generalization of the linear-response time-dependent DFT strategy [44] for responding systems embedded in a frozen density [37]. Within Neglect of Dynamic Response of the Environment (NDRE) approximation, the quality of such shifts is entirely determined by the accuracy of the embedding potential. If NDRE approximation cannot be applied, as in the case of environment and embedded subsystem absorbing at similar frequency, the method based on generalization of ground-state subsystem DFT [45] is indispensable as shown comprehensively in Ref. [46]. The systems considered in the present work concern environments and chromophores, which do not absorb in the same spectral range, i.e., cases where NDRE approximation can be applied. Note that FDET calculations in which NDRE approximation is applied are also referred to as uncoupled FDE or FDE(u) in the literature [46]. The other approximation made in any FDET-based calculations concerns the non-additive kinetic potential (see the next section). The present work does not concern this approximation. We build up upon the gathered numerical experience reported in the literature. The investigated systems are the same or very similar to the ones for which the adequacy of the used approximations was already put to scrutiny.

[1] In the present work any reference to Density Functional Theory (DFT), subsystem DFT, Kohn–Sham DFT, and Frozen-Density Embedding Theory (FDET), concerns the exact formalisms and not approximate methods based on such formalisms.

The present work is organized as follows. We start with a nutshell presentation of Frozen-Density Embedding Theory which is followed by the overview of the literature concerning the dependence of the FDET results on the choice of the frozen density. A dedicated section deals with the issue of electronic polarization of the environment seen from the FDET perspective. The results section deals with the dependence on energies of local excitations on the chosen ρ_B. In the first part, clusters consisting of *cis*-7-hydroxyquinoline (7HQ) and from one-to three hydrogen-bonded molecules are used to test three strategies to generate ρ_B (superposition of atomic densities, superposition of molecular densities, ground-state Kohn–Sham calculations for the whole environment). For these clusters, shifts of the excitation energies obtained in gas-phase experiments, benchmark quality EOM–FDET (for the smallest ones only) and FDET/LR-TDDFT calculations are available (see data collected in Refs. [8, 47]). These three types of shifts do not differ from each other significantly and provide reference data to investigate the effect of the choice of ρ_B in the FDET/LR-TDDFT calculations. The second part of the results section concerns larger systems, for which the third strategy to generate the frozen density is applied extensively and the effect of varying ρ_B on the calculated shifts is investigated in detail.

1.1 Frozen-Density Embedding Theory

Frozen-Density Embedding Theory [20–23] concerns minimizing the total energy for a system comprising N_{AB} electrons in the external potential $v(\mathbf{r})$ in the presence of the constraint $\rho \geq \rho_B$:

$$E_{\mathrm{emb}}[\rho_B] = \min_{\substack{\rho(\mathbf{r}) \geq \rho_B(\mathbf{r}) \geq 0 \\ \int \rho(\mathbf{r})\mathrm{d}\mathbf{r} = N_{AB}}} E_v^{\mathrm{HK}}[\rho] \tag{1}$$

where $E_v^{\mathrm{HK}}[\rho]$ is the Hohenberg–Kohn functional of the total energy [48] and ρ_B is a given function.

There are no other constraints for $\rho_B(\mathbf{r})$ than the ones given in Eq. 1. If the integral $\int \rho_B(\mathbf{r})\mathrm{d}\mathbf{r}$ is an integer (denoted by N_B) then the difference $N_A = N_{AB} - N_B$ is also an integer, and the above definition can be written alternatively:

$$E_{\mathrm{emb}}[\rho_B] = \min_{\substack{\rho_A(\mathbf{r}) \geq 0 \\ \int \rho_A(\mathbf{r})\mathrm{d}\mathbf{r} = N_A}} E_v^{\mathrm{HK}}[\rho_A + \rho_B] \tag{2}$$

We mention here the closely related formal framework of partition DFT [49, 50], in which the integrals $\int \rho_A(\mathbf{r})\mathrm{d}\mathbf{r}$ and $\int \rho_B(\mathbf{r})\mathrm{d}\mathbf{r}$ can be fractional numbers which add up to an integer (N_{AB}). In FDET, the integral $\int \rho_B(\mathbf{r})\mathrm{d}\mathbf{r}$ can be also a fractional number. In such a case, however, the definition of $E_{\mathrm{emb}}[\rho_B]$ given in Eq. 2 is not applicable and $E_{\mathrm{emb}}[\rho_B]$ is only defined in Eq. 1. FDET provides a practical strategy to perform such a search by constructing an appropriate embedding potential ($v_{\mathrm{emb}}(\mathbf{r})$) assuring satisfaction of the constraint $\rho \geq \rho_B$. The embedding potential is determined uniquely by the following quantities: electron density of the environment (denoted as ρ_B throughout this work), electron density of the embedded system ($\rho_A(\mathbf{r}) = \rho(\mathbf{r}) - \rho_B(\mathbf{r}) \geq 0$ by construction), and the density of the positive charge of the environment ($\rho_B^{\mathrm{pos}}(\mathbf{r})$ is usually the sum of nuclear charges) which generates the electrostatics potential $v_B(\mathbf{r}) = \int \frac{\rho_B^{\mathrm{pos}}(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}\mathrm{d}\mathbf{r}'$. The form of this correspondence was derived for embedding various quantum mechanical systems: reference system of non-interacting electrons (the Kohn–Sham system [51]) [20, 21], interacting wavefunction [22], and one-particle reduced density matrix [23] and reads:[2]

$$v_{\mathrm{emb}}[\rho_A, \rho_B, v_B](\mathbf{r}) = v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}\mathrm{d}\mathbf{r}' + \frac{\delta E_{\mathrm{xc}}^{\mathrm{nad}}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})} + \frac{\delta T_s^{\mathrm{nad}}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})} \tag{3}$$

The non-additive bi-functionals occurring in the last two terms in the above equation are defined through the functionals for the exchange-correlation energy $E_{\mathrm{xc}}[\rho]$ and for the kinetic energy in the non-interacting reference system $T_s[\rho]$ known in Kohn–Sham formulation of Density Functional Theory [48, 51]. In particular, the constrained search definition [53, 54] of the bi-functional $T_s^{\mathrm{nad}}[\rho_A, \rho_B]$ reads:

$$T_s^{\mathrm{nad}}[\rho_A, \rho_B] = \min_{\Psi_s \to \rho_A + \rho_B} \langle \Psi_s | \hat{T} | \Psi_s \rangle - \min_{\Psi_s \to \rho_A} \langle \Psi_s | \hat{T} | \Psi_s \rangle - \min_{\Psi_s \to \rho_B} \langle \Psi_s | \hat{T} | \Psi_s \rangle \tag{4}$$

where, Ψ_s denotes a trial function of the form of a single-determinant. The index s used in $T_s[\rho]$ and all subsequently defined quantities is used to indicate that the concerned definitions involve the reference system of non-interacting electrons for which the ground-state wavefunction has this form.

In practice, the quality of calculated environment-induced shifts of a given property of the embedded system obtained from FDET-based multi-level simulations, which is calculated either as the expectation value of the operator associated to this property or from response theory-based calculations using embedded wavefunction (like FDET/

[2] In the case of embedded interacting wavefunction of the truncated Configuration Interaction form, an additional term in the embedding potential is needed [22] but it is a matter of convention whether this term is considered a part of the embedding potential or the potential for subsystem A (see also the relevant discussion in Ref. [52]).

LR-TDDFT calculations reported in the present work) hinges on the following factors: (1) choice of $\rho_B(\mathbf{r})$ (electron density of the environment), (2) approximation for the bifunctional $\frac{\delta E_{xc}^{nad}[\rho_A,\rho_B]}{\delta\rho_A(\mathbf{r})}$ (the exchange-correlation part in the RHS of Eq. 3 denoted by $v_{xc}^{nad}[\rho_A,\rho_B](\mathbf{r})$ in the present work), (3) approximation for the bifunctional $\frac{\delta T_s^{nad}[\rho_A,\rho_B]}{\delta\rho_A(\mathbf{r})}$ denoted by $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$ in the present work. It is worthwhile noticing that the quality of the energy, on the other hand, depends additionally on the error in the functionals used to generate the embedding potential (for the detailed discussion see Ref. [55]).

This work focuses on the first issue. Concerning approximating $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$, a short overview of the situation is given below. In the original and subsequent works by us and others, the bi-functional $T_s^{nad}[\rho_A,\rho_B]$ (and the corresponding functional derivative—$v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$) was approximated by means of one of the known approximations for the functional $T_s[\rho]$.

$$T_s^{nad}[\rho_A,\rho_B] \approx \tilde{T}_s^{nad(decomposable)}[\rho_A,\rho_B] = \tilde{T}_s[\rho_A+\rho_B]-\tilde{T}_s[\rho_A]-\tilde{T}_s[\rho_B] \quad (5)$$

We refer to such approximations as "decomposable" for obvious reasons. In our original work on FDET [20], local-density approximation (Thomas–Fermi functional [56, 57]) and second-order [58] gradient expansion [59] for the kinetic energy was used to construct a decomposable approximation for the bi-functional $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$. One of the surprising results reported in Ref. [20] was that the improvement in approximation to $T_s[\rho]$ by adding to the Thomas–Fermi functional the second-order gradient-expansion correction worsens the FDET interaction energies which are determined rather by the accuracy of the bifunctional $T_s^{nad}[\rho_A,\rho_B]$. This observation was investigated further [60] showing that worsening the interaction energies can be attributed to the errors introduced into the potential $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$. These observations lead us to the GGA97 bi-functional [61, 62], which is currently the most commonly used approximation for $T_s^{nad}[\rho_A,\rho_B]$ in FDET-based numerical simulations. The GGA97 approximation for the bifunctional for $T_s^{nad}[\rho_A,\rho_B]$ (and the corresponding $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$) is obtained directly from the Lembarki–Chermette approximation for $T_s[\rho]$ [63]. In the Lembarki–Chermette approximation for $T_s[\rho]$, the second-order gradient-expansion contribution is attenuated locally depending on the magnitude of the reduced density gradient. This eliminates to some extent the errors in $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$ originating in the second-order term in gradient-expansion approximation for $T_s[\rho]$. NDSD approximation (non-decomposable bi-functional using on up to second derivatives of density [64]) for $T_s^{nad}[\rho_A,\rho_B]$, which does not have the decomposable form given in Eq. 5, is constructed to satisfy the exact asymptotic conditions for the bi-functional $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$ of the greatest relevance near nuclear cusps. It has to be underlined that LDA, GGA97, and NDSD are all semi-local approximations for $T_s^{nad}[\rho_A,\rho_B]$ (and the corresponding $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$). At each point $\mathbf{r}$, the potential $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$ is determined by the densities $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ and their first- and second derivatives at the same point. Such potentials were comprehensively shown to fail if the overlap between ρ_A and ρ_B is large [42, 60, 62, 65–67] as it is the case of an embedded subsystem covalently bound to the environment. For such cases an alternative strategy to approximate $v_t^{nad}[\rho_A,\rho_B](\mathbf{r})$ is becoming increasingly popular which is based on numerical inversion of the Kohn–Sham potential [32, 33, 68, 69]. This work does not deal with such methods and aims at obtaining the exact properties based on semi-local approximation to $T_s^{nad}[\rho_A,\rho_B]$, which are known to be applicable in case where the embedded subsystem is not covalently bound to the environment.

1.2 Interpretation of the electronic polarization of the environment in Frozen-Density Embedding Theory

Let us consider two choices for ρ_B : $\rho_B^{(1)}$ and $\rho_B^{(2)}$. In such a case,

$$E_{emb}[\rho_B^{(1)}] = \min_{\rho_A\geq 0,\ \int\rho_A(\mathbf{r})d\mathbf{r}=N_A} E_v^{HK}[\rho_A+\rho_B^{(1)}] = E_v^{HK}[\rho_{A(opt)}^{(1)}+\rho_B^{(1)}] \quad (6)$$

$$E_{emb}[\rho_B^{(2)}] = \min_{\rho_A\geq 0,\ \int\rho_A(\mathbf{r})d\mathbf{r}=N_A} E_v^{HK}[\rho_A+\rho_B^{(2)}] = E_v^{HK}[\rho_{A(opt)}^{(2)}+\rho_B^{(2)}] \quad (7)$$

where $\rho_{A(opt)}^{(1)}$ and $\rho_{A(opt)}^{(2)}$ are the optimized densities of subsystem A obtained for these two choices for ρ_B. On the virtue of the first Hohenberg–Kohn theorem [48], $E_{emb}[\rho_B^{(1)}] = E_{emb}[\rho_B^{(2)}]$ if $\rho_{A(opt)}^{(1)}+\rho_B^{(1)} = \rho_{A(opt)}^{(2)}+\rho_B^{(2)}$. Such situation is not uncommon. Let us consider $\rho_B^{(1)}$ and $\rho_B^{(2)}$ chosen in such a way that $\rho_B^{(1)} \leq \rho_{tot}$ and $\rho_B^{(2)} \leq \rho_{tot}$, where ρ_{tot} is the ground-state density of the whole system. Either Eq. 6 or 7 lead to the same total density (actually it is a ground-state density because the constraint $\rho \geq \rho_B$ is automatically satisfied for the chosen $\rho_B^{(1)}$ and $\rho_B^{(2)}$). The difference $\rho_B^{(1)} - \rho_B^o$, where ρ_B^o is the density of the isolated subsystem B, represents the electronic polarization of subsystem B due to interactions with subsystem A. But

since $\rho_B^{(1)}$ and $\rho_B^{(2)}$ are different, also $\rho_B^{(2)} - \rho_B^o$ differs from $\rho_B^{(1)} - \rho_B^o$! Despite the fact that they lead to the same total density, the two choices for ρ_B^o lead to different polarization (measured as the difference $\rho_B - \rho_B^o$). As a result, the term "electronic polarization of the environment" is not defined uniquely in FDET.

Even for such $\rho_B^{(1)}$ and $\rho_B^{(2)}$, for which $E_{\mathrm{emb}}[\rho_B^{(1)}]$ and $E_{\mathrm{emb}}[\rho_B^{(2)}]$ are not the same, the difference cannot be attributed only to different representation of electronic polarization of subsystem B. In FDET, it can originate from the fact that imposing the condition $\rho \geq \rho_B^{(1)}$ and $\rho \geq \rho_B^{(2)}$ involves different energy penalties, which is not directly related to the way the polarization is reflected in the choice for ρ_B. The same consideration apply for the interpretation of electronic polarization in subsystem A by switching the roles of A and B in the above considerations. The lack of unique definition of the "electronic polarization of the environment" in FDET makes FDET-based embedding methods different from empirical QM/MM type of approaches, in which the electronic polarization enters as a well defined contribution in the total energy expression which is frequently essential for accuracy (for a recent review see Ref. [5]).

The situation is quite different if approximate density functionals are used. We will discuss the differences in the case of subsystem densities optimized for both subsystems as it is made in Cortona's formulation of DFT [70]:

$$E_o = \min_{\substack{\rho_A \geq 0 \\ \int \rho_A(\mathbf{r})\mathrm{d}\mathbf{r}=N_A}} \min_{\substack{\rho_B \geq 0 \\ \int \rho_B(\mathbf{r})\mathrm{d}\mathbf{r}=N_B}} E_v^{\mathrm{HK}}[\rho_A + \rho_B] \tag{8}$$

The possibility of existence of several pairs yielding the same total energy has been long recognized [21, 71]. The same concerns FDET, where ρ_A is optimized whereas ρ_B is not. The analytical examples of such pairs are provided in Refs. [66, 71–73] whereas numerical examples for molecular systems are provided in Ref. [74]. It is useful to point out here the partition DFT formalism [49, 50], which can be seen as generalization of Cortona's subsystem DFT [70] allowing for a unique partitioning based on chemically motivated additional constraints.

If, however, approximations are used (denoted by tildes in the formulas below), the search given in Eq. 8 is not performed (see also the detailed discussion of this issue given by Gritsenko in Ref. [75]). Instead, another functional is minimized:

$$\tilde{E}_o = \min_{\substack{\rho_A \geq 0 \\ \int \rho_A(\mathbf{r})\mathrm{d}\mathbf{r}=N_A}} \min_{\substack{\rho_B \geq 0 \\ \int \rho_B(\mathbf{r})\mathrm{d}\mathbf{r}=N_B}} \left(E_v^{\mathrm{HK}}[\rho_A + \rho_B] + \Delta\tilde{E}_{\mathrm{xc}}[\rho_A + \rho_B] + \Delta\tilde{T}_s^{\mathrm{nad}}[\rho_A, \rho_B]\right) \tag{9}$$

where Δ denotes the error of a given functional (the difference between the exact functional and its approximated counterpart). Among all possible pairs yielding the same total electron density, the above search picks up that pair $\tilde{\rho}_A^{\mathrm{opt}}$ and $\tilde{\rho}_B^{\mathrm{opt}}$ which minimizes the error $(\Delta\tilde{T}_s^{\mathrm{nad}}[\rho_A, \rho_B])$ in the used approximation to the non-additive kinetic energy. The partitioning is thus unphysical. Note that the use of approximations into the exchange-correlation density functional cannot lead to uniqueness of partitioning because this energy component (and its errors) depends on the total density regardless how it is decomposed. The numerical practice using the *freeze-and-thaw* procedure [65] to perform minimization given in Eq. 9 shows that the use of $\tilde{T}_s^{\mathrm{nad}}[\rho_A, \rho_B]$ instead of $T_s^{\mathrm{nad}}[\rho_A, \rho_B]$ leads to a unique partitioning (for the overview see the articles quoted in Ref. [21]). In contrast with the situation in the formal framework of partition DFT [49, 50], partitioning obtained from Eq. 9 is unphysical.

If an approximation is used for $T_s^{\mathrm{nad}}[\rho_A, \rho_B]$, ρ_B which is optimized in the *freeze-and-thaw* procedure [65], the unique optimized ρ_B might differ from the density of the isolated subsystem B not only due to electronic polarization but also to the error on the non-additive kinetic potential. This error might result in artificial charge-redistribution between subsystems or qualitatively wrong polarization of subsystems [41, 60, 64, 76]. For this reason, we make a distinction between "polarization of subsystem B due to interactions with subsystem A" and "variational relaxation of subsystem B" in subsystem DFT calculations. The latter includes both the physical effect and the effect due to the error in the approximation used for the non-additive kinetic potential. The numerical importance of the two effects cannot be distinguished in practice. Based on our numerical experience, we believe that the error in currently used approximations for the non-additive kinetic potential affects the ρ_B optimized in the *freeze-and-thaw* procedure [65] more than neglecting the electronic polarization in the case of non-charged subsystem A non-covalently linked to subsystem B. The situation is different if subsystem A is charged. In such a case the relaxation can be attributed mainly to the electronic polarization (see also the discussion in the following section).

1.3 Previous studies of the dependence of molecular properties derived from FDET-based calculations on the choice for ρ_B

There are a practically infinite number of strategies to generate ρ_B in any multi-scale simulations based on FDET. They can be made based on "chemical intuition" or on dedicated tests made for model systems. In the previous section, the possible pitfalls associated with the notion of electronic polarization were identified on the formal grounds. The most straightforward strategy to generate ρ_B

is to use ground-state density obtained with some (usually less expensive) quantum mechanical method applied for the environment. Such calculations can be made in the absence of the embedded species or can take into effect the modification of the electron density of the environment by interactions with the embedded species. Numerical examples reported in the literature show frequently that optimizing the environment electron density (ρ_B) by means of the *freeze-and-thaw* cycle does not improve the calculated properties of non-charged embedded species. In the case of chromophores hydrogen bonded to the molecules in the environment, relaxing ρ_B results in the relative change of the excitation energy shift which does not exceed 25 % (the worst detected case) [37, 47, 77]. Moreover, the FDET hydrogen-bonding-induced shifts evaluated with non-relaxed ρ_B agree very well (i.e., within 200 cm^{-1} or better) with high-level wavefunction-based benchmark values [47, 77]. This suggests that the change in ρ_B following the "freeze-and-thaw" energy minimization is rather an effect due to the error in the bi-functional for the non-additive kinetic potential than to the physical effect of the polarization of the environment by the chromophore (see the discussion in the introduction section).

The use of non-relaxed ρ_B is, therefore, expected to be an adequate approximation for this type of weak interactions with the environment. Some properties are insignificantly affected by the relaxation of ρ_B even for charged embedded species. The shifts of the isotropic component of the hyperfine tensor A_{iso} of Mg^+ cation due to embedding in the noble gas matrix, evaluated for the *freeze-and-thaw* optimized and non-optimized ρ_B differ by less than 10 % [34]. This indicates that the effect of the noble gas matrix on the hyperfine tensor of embedded Mg^+ cation originates from the intermolecular Pauli repulsion (confinement effect) rather than from the electric polarization of the matrix by the cation. Concerning the notion of "polarization of the environment", there are cases where it might be helpful in constructing a good approximation for ρ_B such as a highly charged species embedded in highly polarizable environment. The ligand-field splitting of *f*-levels for a trivalent rare-earth cation impurity in a host lattice (chloroelpasolite) might be considered as an extreme example. The nearest ligands of the impurity are Cl^- anions and the use of non-relaxed ρ_B to represent the six ligands seems a very crude approximation in FDET calculations. Indeed, only the qualitative trends for the ligand-field splitting parameters along the lanthanide series are reproduced using non-relaxed ρ_B in KSCED calculations [36]. The FDET splittings calculated at non-relaxed ρ_B are, however, underestimated by a factor of about two. The use of *freeze-and-thaw* optimized ρ_B results in about twofold increase of the splittings, which brings them close to experimental values. Ref. [36] shows also how the rather expensive *freeze-and-thaw* calculations can be avoided even in such extreme case (highly charged subsystem *A* and highly polarisable subsystem *B*) by means of "pre-polarized" ρ_B. Generating the "pre-polarized" ρ_B is quantitatively less expensive than the *freeze-and-thaw* relaxation because it involves just a simple calculation of the ligands in the electric field generated by the embedded cation. The splittings obtained using *freeze-and-thaw* optimized ρ_B and "pre-polarized" ρ_B are almost the same (they agree within 10 cm^{-1}).

The above examples taken from the literature indicate that the choice of ρ_B in FDET-based calculations is the issue calling for a special attention especially since the notion of "polarization" and "relaxation" are not equivalent in FDET as discussed in the previous section. It is a crucial issue in multi-level simulations in which ρ_B represents a system of significantly larger size than subsystem *A* for which quantum mechanical descriptors (orbitals or the non-interacting reference system [20, 21], embedded interacting wavefunction [22], or one-particle reduced density matrix [23]) are constructed. It is, therefore, desirable that generation of an adequate ρ_B involves smaller computational costs than the optimization of ρ_A and the subsequent evaluation of quantum mechanical observables for subsystem *A*. Even in the cases, for which avoiding the optimization of ρ_B is an acceptable approximation, it is highly desirable to use a protocol to generate ρ_B which involves the smallest computational effort. The aim of the present work is a systematic analysis of strengths and weaknesses of possible practical strategies to generate non-relaxed ρ_B for FDET calculations for chromophores in hydrogen-bonded molecular environments.

2 Computational details

2.1 Strategies to generate the frozen electron density—ρ_B

For each considered system, the excitation energies are evaluated using densities of the environment (ρ_B) obtained by means of different computational techniques. The approximations and technical parameters concerning subsystem *A* remain, however, unchanged: (1) approximations for $E_{xc}[\rho_A]$, $T_s^{nad}[\rho_A, \rho_B]$, and $E_{xc}^{nad}[\rho_A, \rho_B]$; (2) atomic basis set used for subsystem *A*; and (3) parameters of technical nature (grids, fitting functions, convergence criteria, etc.) as in the reference calculations. This approach provides direct information concerning the dependence of the FDET calculated excitation energies on ρ_B.

The following simplified methods to generate ρ_B are considered:

(A) *Superposition of atomic densities*

The simplest approximation for the electron density of the environment ρ_B is to use a superposition of atomic densities.

$$\rho_B(\mathbf{r}) = \sum_{i=1}^{N\text{atom}B} (n_B^i - Z_B^i)\rho_B^i(\mathbf{r}) \quad (10)$$

where i indicates the atom in subsystem B, Natom$B$ is the number of atoms in subsystem B, $\rho_B^i(\mathbf{r})$ denotes the spherically symmetric electron density integrating to the total charge equal to the atomic number Z_B^i, and n_B^i denotes the net charge of the atom. The above approximation for ρ_B can be expected to take into account Pauli repulsion between electrons, which are localized in closed shells of atoms in the environment, and the electrons in subsystem A even if $\{n_B^i = 0\}$ (electrically neutral atoms in the environment). In the case when the effect of the environment on the electronic structure is dominated by confinement, choosing $\{n_B^i = 0\}$ can be expected an adequate approximation. For a polar or charged molecule in the environment, taking $\{n_B^i = 0\}$ is most likely a very poor approximation because it would neglect the dominant (i.e. electrostatic) term in the FDET embedding potential at long range. The choice of n_B^i becomes, in such case, the key issue. This approximation for ρ_B was already applied for chromophores in ionic solids [38]. In such a case, a natural choice for n_B^i, is to use the formal ionic charges. Indeed, the numerical tests made for fluorenone embedded in zeolite L reported in Ref. [38] confirm the adequacy of the approximation given in Eq. 10 in the case of the environment being an ionic solid.

(B) *Superposition of molecular densities*

Approximating $\rho_B(\mathbf{r})$ by means of a sum of molecular densities seems to be a particularly appealing strategy if the environment consists of many weakly interacting molecules. In such a case, $\rho_B(\mathbf{r})$ is obtained as:

$$\rho_B(\mathbf{r}) = \sum_{i=1}^{N\text{molec}B} \rho_B^i(\mathbf{r}) \quad (11)$$

where Nmolec$B$ is the number of molecules in subsystem B, $\rho_B^i(\mathbf{r})$ denotes the electron density of the isolated ith molecule.

Approximation given in Eq. 11 was used in the first multi-scale simulations based on FDET concerning ground-state properties of solvated molecules [78, 79]. The adequacy of this approximation for simulating the hydrogen transfer reaction in carbonic anhydrase was investigated in Ref. [35].

Increasing the strength of the interactions between the molecules in subsystem B might invalidate the approximation given in Eq. 11. In the present work, the applicability of Eq. 11 to molecular environment is investigated using complexes of *cis*-7-hydroxyquinoline and several molecules which are hydrogen bonded to the chromophore but also interact via hydrogen bonds among themselves. If the molecules in the environment form a hydrogen-bonded chain, the mutual polarization of the neighboring molecules in the chain, leads to non-additive increase of the total dipole moment. The obtained results can, therefore, be considered as good estimate of maximum error made due to the neglect of such cooperative effects in hydrogen-bonded molecular environments. The presence of hydrogen-bonded chains of molecules in the environment is the "worst scenario" as far as applicability of Eq. 11 is concerned.

(C) *Less expensive Kohn–Sham calculations for the whole environment*

In practice, it is appealing to use less expensive approximations to solve Kohn–Sham equations for the entire environment (subsystem B) than for subsystem A. In contrast with the approximations given in Eqs. 10 and 11, mutual polarization of molecules in the environment is taken into account.

2.2 Miscellaneous computational details

All reported results are obtained using the ADF code [80] into which the ground-state and linear-response time-dependent excited-state FDET calculations were implemented as described in Refs. [37, 81] with subsequent improvements described in detail in Refs. [82]. Slater type orbitals (STO) [83] are used for both subsystems (A and B). Unless specified otherwise, STO–aug-TZP basis set is used for the embedded subsystem (subsystem A). The use of diffuse functions for subsystem A increases the sensibility of the FDET results in change of the embedding potential which might occur as a consequence of changes in ρ_B. For ρ_B or for its components, the basis sets are ranging from STO–SZ to STO–aug-TZ2P and they are specified in each case. The following approximations for density (or orbital) functionals are used: (1) SAOP [84, 85] for $v_{\text{xc}}[\rho_A](\mathbf{r})$, i.e., the exchange-correlation potential in subsystem A, (2) the decomposable approximation using the PW91 [86] exchange-correlation functional for $v_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B](\mathbf{r})$, i.e., the non-additive exchange-correlation bi-functional in Eq. 3, (3) either NDSD functional [64] or the decomposable approximation GGA97 [61] obtained from the Lembarki–Chermette functional for the kinetic energy [63] for $v_t^{\text{nad}}[\rho_A, \rho_B](\mathbf{r})$, the non-additive kinetic potential in Eq. 3. In the case of other approximations, they are specified in the relevant part of the "Results" section. The technical

parameters used in the present study are not optimized for speed of calculations but are set up to eliminate or to keep constant other factors (than the choice for ρ_B), which affect the FDET results. In particular, the exchange-correlation- and the non-additive components of the embedding potential are evaluated at "exact" densities, i.e., electron densities evaluated from orbitals without the use of fitting functions.

3 Results

This section comprises two parts. The first one concerns the same chromophore—*cis*-7-hydroxyquinoline—in environments comprising up to three hydrogen-bonded molecules. The three strategies to generate ρ_B are applied and the corresponding environment-induced shifts of the $\pi \rightarrow \pi^*$ excitation energies are discussed in order to select the most robust strategy to generate ρ_B in large-scale multi-level simulations for similar chromophores in soft condensed matter. In the second part, the most robust strategy (Kohn–Sham calculations for the whole environment) is applied for larger environments of other chromophores in order to verify the applicability of this strategy.

3.1 *cis*-7-Hydroxyquinoline in hydrogen-bonded complexes

These systems are ideal objects for studying the dependence of the calculated complexation induced shifts of the excitation energy on the choice made for ρ_B in FDET/LR-TDDFT calculations. The studied excitations are local. They fall, therefore, in the domain of applicability of FDET. The molecules do not absorb noticeably within the considered spectral range, which makes the NDRE approximation applicable. Moreover, benchmark quality excitation energies obtained from high-level wavefunction methods are available for the smallest of these complexes [47, 77]. It was shown that FDET calculations reproduce very accurately both the reference benchmark shifts of the excitation energy and the experimental shifts, provided ρ_B is obtained from Kohn–Sham calculations for the isolated environment [47, 77]. These FDET excitation energy shifts are used as a reference for the present work in which ρ_B is generated by means of alternative (simpler) methods. The geometries of the considered "microsolvated" *cis*-7-hydroxyquinoline clusters are also taken from Refs. [47, 87].

3.1.1 *Superposition of atomic densities for the environment*

The applicability of approximation given in Eq. 10 is studied for the complexes *cis*-7-hydroxyquinoline-H_2O and *cis*-7-hydroxyquinoline-NH_3. In the considered complexes, the search of the optimal n_B^i is especially straightforward because once the net atomic charge on the non-hydrogen atom is fixed, the charges on hydrogens are uniquely determined by the condition of neutrality of the whole molecule ($n_B^{\mathrm{H}} = -n_B^{\mathrm{O}}/2$ for H_2O and $n_B^{\mathrm{H}} = -n_B^{\mathrm{N}}/3$ for NH_3). The search for the optimal n_B^i can be performed, therefore, in one dimension by varying the net charge on non-hydrogen atom from 0 (neutral atom) to -1 e (complete charge transfer from hydrogens). The principal orbitals contributing to the lowest excitation are shown in Fig. 1.

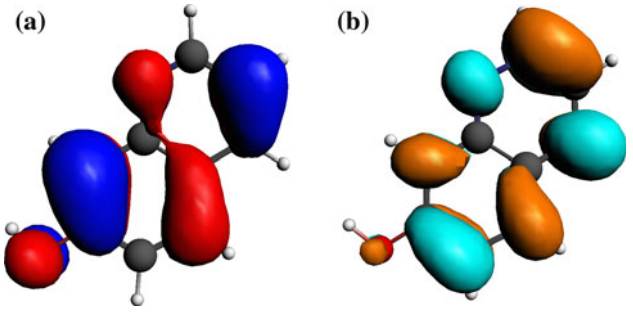


Fig. 1 The pair of molecular orbitals in isolated *cis*-7-hydroxyquinoline, which provides the dominant contribution to the lowest excitation (HOMO on the *left side* and LUMO on the *right side*)

Figures 2 and 3 show the excitation energy shifts calculated for the considered range of atomic net charges. The shifts depend critically on the chosen values of n_B^i in Eq. 10 in both complexes. The increase of the polarity of the molecule in the environment increases also the magnitude of the red shift. The benchmark results for the 7HQ + H_2O complex are reproduced at the net charge at oxygen of about -0.7 e. Interestingly, the oxygen net charge obtained from the Mulliken population analysis lies within this range. The use of any other common methods to decompose the molecular charge into atomic contributions leads to slightly worse shifts but seems also to be a reasonable choice.

Unfortunately, none of the common methods to obtain the net charges in ammonia seems to be useful to generate ρ_B by means of Eq. 10. The best agreement with the reference values for the shift occurs for $n_B^{\mathrm{N}} < -1.0$ e. None of the standard methods to partition the molecular electron density into atomic contributions considered here leads to such a value of the net charge. Similar conclusions can be drawn using smaller basis sets for subsystem *A* such as STO–DZP, STO–TZP (data available from the authors upon request). Most likely, using the spherical atomic charges is too crude as an approximation to represent properly the electric field generated by the lone pair of nitrogen. A universal procedure to generate net atomic charges to be used in Eq. 10 is hardly in view. Obtaining

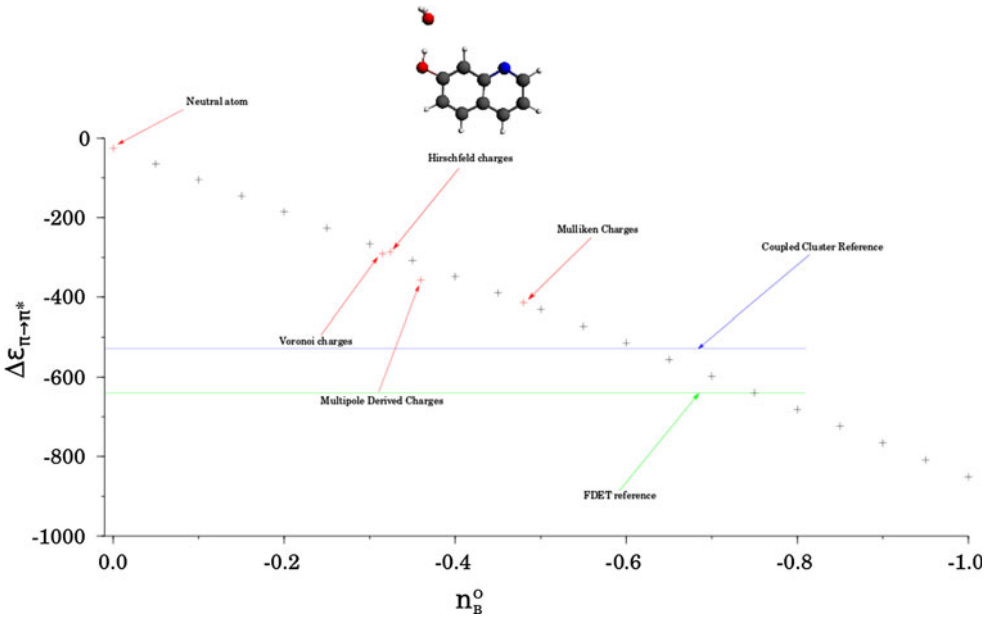


Fig. 2 Dependence of the complexation induced shift (in cm^{-1}) of the lowest $\pi \to \pi^*$ excitation in the complex of *cis*-7-hydroxyquinoline with water on the net atomic charges used in Eq. 10 for generating the frozen density $\rho_B = \rho_{H_2O} \cdot n_B^O$ is the net charge on oxygen and the charge on each hydrogen equals to $-n_B^O/2$. The wavefunction and FDET reference values taken from Ref. [47] are indicated as *horizontal lines*. Results obtained using common methods for generating net atomic charges: Mulliken population analysis [88], Hirschfeld population analysis [89], Voronoi charges [90], multipole derived [91] are also indicated. STO–DZP basis set is used for generation of ρ_B

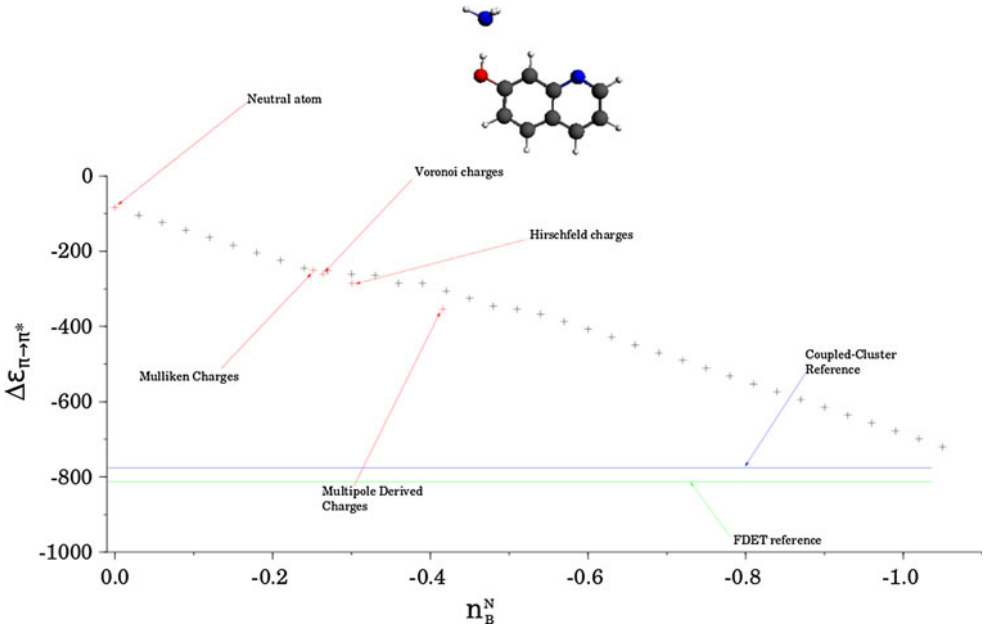


Fig. 3 Dependence of the complexation induced shift (in cm^{-1}) of the lowest $\pi \to \pi^*$ excitation in the complex of *cis*-7-hydroxyquinoline with ammonia on the net atomic charges used in Eq. 10 for generating the frozen density $\rho_B = \rho_{NH_3} \cdot n_B^N$ is the net charge on nitrogen and the charge on each hydrogen equals to $-n_B^N/3$. The wavefunction and FDET reference values taken from Ref. [47] are indicated as *horizontal lines*. Results obtained using common methods for generating net atomic charges: Mulliken population analysis [88], Hirschfeld population analysis [89], Voronoi charges [90], multipole derived [91] are also indicated. STO–DZP basis set is used for generation of ρ_B

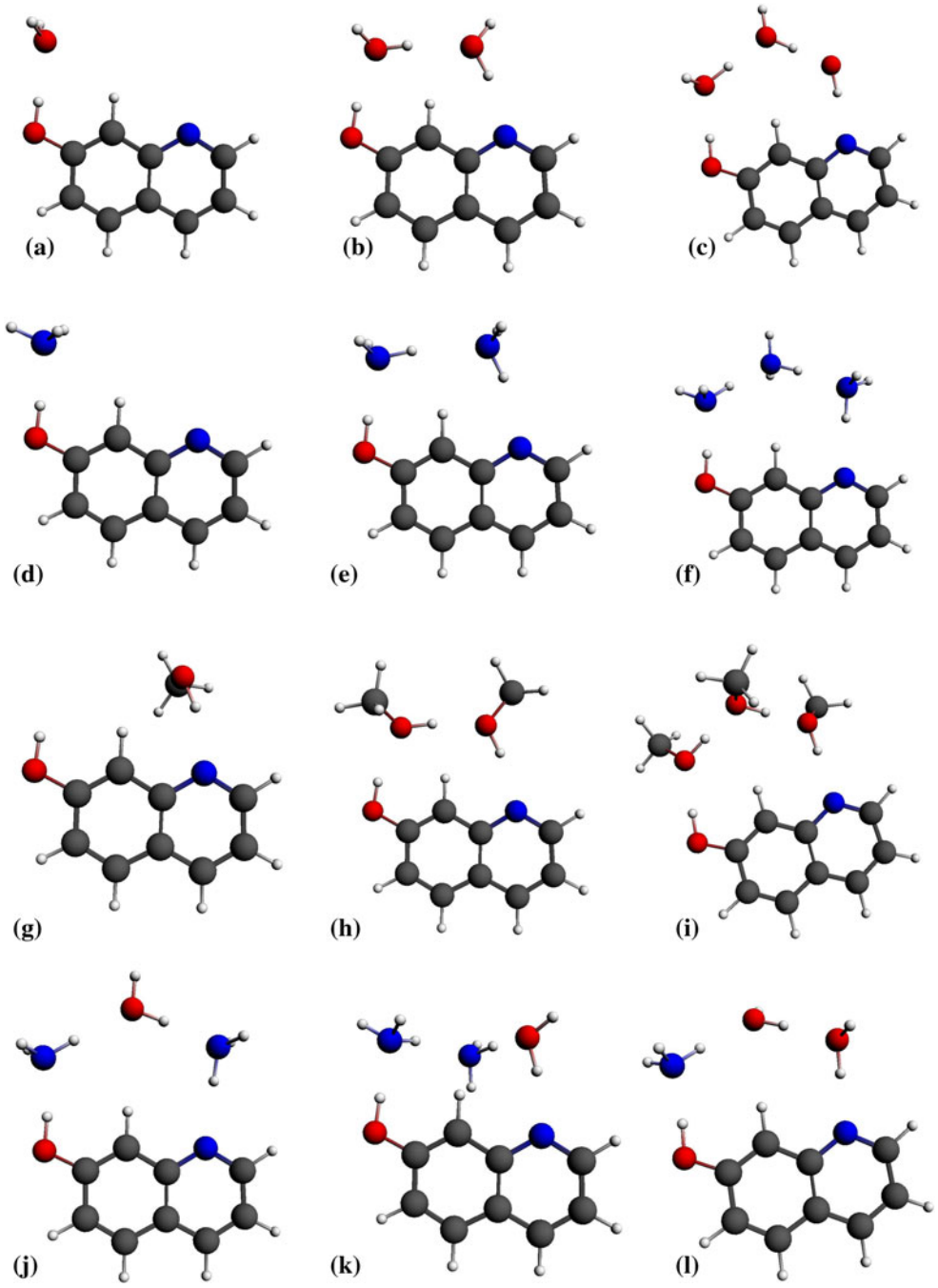


Fig. 4 *cis*-7-Hydroxyquinoline (7HQ) in complexes investigated in this work: **a** 7HQ–H_2O, **b** 7HQ–$2H_2O$, **c** 7HQ–$3H_2O$, **d** 7HQ–NH_3, **e** 7HQ–$2NH_3$, **f** 7HQ–$3NH_3$, **g** 7HQ–CH_3OH, **h** 7HQ–$2CH_3OH$, **i** 7HQ–$3CH_3OH$, **j** 7HQ–NH_3–H_2O–NH_3, **k** 7HQ–NH_3–NH_3–H_2O, **l** 7HQ–NH_3–H_2O–H_2O

ρ_B from Eq. 10 is, therefore, not practical for environments comprising polar molecules despite its usefulness in the case of ionic solids [38].

Turning back to the dependence of the shift on the net charge, we notice that the calculated shifts are very small (a few cm^{-1} for 7HQ + NH_2O and less than 100 cm^{-1} for 7HQ + NH_3) at the neutral atom limit ($n_B^i = 0$) at which the long-range electrostatic component of the embedding potential disappears. This indicates that confinement is a minor factor determining the shifts. The non-electrostatic components of the embedding potentials are present regardless of which value of the net atomic charge is used whereas the electrostatic component is proportional to the magnitude of the net atomic charges (see Eqs. 3, 10). It is also worthwhile noticing the almost linear dependence of the shifts on the net charge supporting the electrostatic interpretation of the origin of the shift.

3.1.2 *Superposition of molecular densities*

The results obtained in the previous section indicate that the simplest strategy (superposition of atomic densities)

Table 1 Environment-induced shifts ($\Delta\varepsilon_{\pi-\pi^*}$) of the lowest $\pi \rightarrow \pi^*$ excitation energy for *cis*-7-hydroxyquinoline in various environments evaluated using either the reference ρ_B (Kohn–Sham calculations for the whole environment) or ρ_B obtained as superposition of molecular densities (Eq. 11). The differences between the excitation energies obtained at these two choices for ρ_B ($\varepsilon[\rho_B^{\text{Eq.11}}] - \varepsilon[\rho_B^{\text{KS}}]$) are given in parentheses. STO–aug-TZ2P basis set is used for ρ_A generation. STO–DZP basis set is used for ρ_B generation

Choice for ρ_B	Environment					
	$\Delta\varepsilon_{\pi-\pi^*}$ (cm^{-1})					
	$2H_2O$	$3H_2O$	$2NH_3$	$3NH_3$	$2CH_3OH$	$3CH_3OH$
$\rho_B = \rho_B^{\text{Eq.11}}$	−1493	−1176	−1158	−1154	−1309	−1275
$\rho_B = \rho_B^{\text{KS}}$	−1614	−1545	−1225	−1392	−1450	−1625
	(121)	(369)	(67)	(238)	(141)	(350)

Table 2 Environment-induced shifts ($\Delta\varepsilon_{\pi-\pi^*}$) of the lowest $\pi \rightarrow \pi^*$ excitation energy for *cis*-7-hydroxyquinoline in various environments evaluated using either the reference ρ_B (Kohn–Sham calculations for the whole environment) or ρ_B obtained as superposition of molecular densities (Eq. 11). The differences between the excitation energies obtained at these two choices for ρ_B ($\varepsilon[\rho_B^{\text{Eq.11}}] - \varepsilon[\rho_B^{\text{KS}}]$) are given in parentheses. STO–aug-TZ2P basis set is used for ρ_A generation. STO–DZP basis set is used for ρ_B generation

Choice for ρ_B	Environment		
	$\Delta\varepsilon_{\pi-\pi^*}$ (cm^{-1})		
	NH_3–H_2O–NH_3	NH_3–NH_3–H_2O	NH_3–H_2O–H_2O
$\rho_B = \rho_B^{\text{Eq.11}}$	−1483	−1716	−1636
$\rho_B = \rho_B^{\text{KS}}$	−1735	−1817	−1856
	(252)	(101)	(220)

Table 3 Environment-induced shifts ($\Delta\varepsilon_{\pi-\pi^*}$) of the lowest $\pi \rightarrow \pi^*$ excitation energy for *cis*-7-hydroxyquinoline environments comprising broken-hydrogen chains evaluated using either the reference ρ_B (Kohn–Sham calculations for the whole environment) or ρ_B obtained as superposition of molecular densities (Eq. 11). The differences between the excitation energies obtained at these two choices for ρ_B ($\varepsilon[\rho_B^{\text{Eq.11}}] - \varepsilon[\rho_B^{\text{KS}}]$) are given in parentheses. STO–aug-TZ2P basis set is used for ρ_A generation. STO–DZP basis set is used for ρ_B generation

Choice for ρ_B	Environment		
	$\Delta\varepsilon_{\pi-\pi^*}$ (cm^{-1})		
	$2H_2O$	$2NH_3$	$2CH_3OH$
$\rho_B = \rho_B^{\text{KS}}$	−899	−900	−1010
$\rho_B = \rho_B^{\text{Eq.11}}$	−881	−888	−993
	(17)	(12)	(18)

might lead to erratic FDET shifts if the environment is a polar molecule. The results are very sensitive to the arbitrary choice of the procedure to generate net atomic charges. As the ammonia example shows, it might even not be possible to generate appropriate net charges.

For environments consisting of several polar molecules the second strategy (Eq. 11) appears as the only option. It is investigated for complexes comprising *cis*-7-hydroxyquinoline embedded in such environments as: $2H_2O$, $3H_2O$, $2NH_3$, $3NH_3$, $2CH_3OH$, and $3CH_3OH$. The complexes are shown in Fig. 4. The geometries of these complexes are taken from Refs. [47, 87], which provide also reference benchmark shifts from EOM-CC and FDET calculations. The chosen systems represent well the worst scenario as far as the approximation introduced in Eq. 11 is concerned. In each case, the molecules in the environment form chain-like structures (see Fig. 4) and the effect of mutual polarization can be expected to be large. Indeed, according to our own studies reported in Ref. [8], this mutual polarization lies at the origin of the positive cooperativity in the solvatochromic shift if the hydrogen-bonded chains consists of more than two hydrogen-bonded molecules.

Using ρ_B as the superposition of electron densities of isolated individual molecules from the environment instead of the Kohn–Sham density obtained for the whole environment affects the shifts noticeably—from 67 cm^{-1} in the case of two ammonia molecules to as much as 369 cm^{-1} in the case of three water molecules forming a chain (see Tables 1, 2). The relative errors in the shifts due to the approximation given in Eq. 11 can reach up to 25 % (the case of three membered chains). Therefore, it can not be considered as a generally adequate approximation.

However, in some cases, generating ρ_B using Eq. 11 is an acceptable approximation. The following computational experiment concerns such a case. The shifts are calculated for the longest chains (three molecules in the environment) but with the central molecule removed to break the chain. The mutual polarization of the molecules in the environment is reduced as the result. As expected, the differences between shifts obtained using superposition of molecular densities and densities derived from Kohn–Sham

Table 4 Environment-induced shifts ($\Delta\varepsilon_{\pi-\pi^*}$) of the lowest $\pi \rightarrow \pi^*$ excitation energy for *cis*-7-hydroxyquinoline environments comprising broken-hydrogen chains evaluated using either the reference ρ_B (Kohn–Sham calculations for the whole environment) or ρ_B obtained as superposition of molecular densities (Eq. 11). The differences between the excitation energies obtained at these two choices for ρ_B ($\varepsilon[\rho_B^{\text{Eq.11}}] - \varepsilon[\rho_B^{\text{KS}}]$) are given in parentheses. STO–aug-TZ2P basis set is used for ρ_A generation. STO–DZP basis set is used for ρ_B generation

Choice for ρ_B	Environment		
	$\Delta\varepsilon_{\pi-\pi^*}$ (cm^{-1})		
	NH_3–NH_3	NH_3–H_2O	NH_3–H_2O
$\rho_B = \rho_B^{\text{KS}}$	−1172	−1545	−1309
$\rho_B = \rho_B^{\text{Eq.11}}$	−1161	−1529	−1297
	(11)	(16)	(12)

calculations for the whole environment are negligible (see Tables 3, 4).

The results indicate clearly that the computationally attractive strategy in which ρ_B is approximated as a sum of electron densities in the isolated molecules is a very good pragmatic solution for hydrogen-bonded clusters in the absence of direct hydrogen bonding between the molecules in the environment. Replacing one Kohn–Sham calculation for the whole environment by several less expensive calculations for each individual molecule in the environment affects the excitation energies by at the most 20 cm^{-1}. In the case of direct hydrogen bonding between the molecules in the environment, the errors due to neglect of mutual polarization between the molecules in the environment are significantly larger (up to 400 cm^{-1}). Such errors might be considered acceptable for some applications (they still do not exceed 25 % of the total shifts). If a better accuracy is needed and the presence of the hydrogen-bonded chains cannot be excluded, the mutual polarization of the molecules in the environment must be taken into account. The subsequent section concerns such a case.

Although the principal interest of the present work lies in the dependency of FDET excitation energies on ρ_B, Tables 10, 11, 12 and 13 in the Supporting Information collect also the excitation energy shifts obtained using other basis sets for subsystem *A*. The approximation given in Eq. 11 leads to the effect on the shifts which is practically independent on the basis set used for subsystem *A*. This justifies attributing the difference between the excitation energy obtained with Eq. 11 and with full Kohn–Sham treatment of subsystem *B* to a physical effect—mutual polarization of the molecules in the environment. Data collected in Tables 10 and 11 of Supporting Information demonstrate remarkable numerical stability of the shifts with respect to the changes of the basis set used for subsystem *A*. Upon changing the basis set for subsystem *A*, the shifts vary within about 20 cm^{-1}. Such small variation is negligible compared to the magnitude of the shifts, which are about one order of magnitude larger. This stability of the shifts reflects the variational origin of the FDET embedding potential and the adequacy of the used approximation for the non-additive kinetic potential.

3.1.3 Kohn–Sham calculations for the whole environment

As shown in the previous section, neglecting the mutual polarization of the molecules in the environment (approximation made in Eq. 11) leads to significant errors in the complexation induced shifts of excitation energy (up to 25 % of the total shift) if the molecules in the environment form longer hydrogen-bonded chains. The errors are not significant in the absence of such chains. This suggests that, in order to generate ρ_B in the case of modeling chromophores in hydrogen-bonded environments such as liquid water, the whole environment should be treated at the molecular orbital level of description. The chosen method should be inexpensive and take into account the mutual polarization of the molecules in the environment properly. Ground-state Kohn–Sham calculations for the whole environment can be, therefore, considered as a practical option. Depending on the choice of the functional and the basis set, such calculations lead to different ρ_B, which in turn leads to different dipole moment of the environment. In the present section, the effect of these choices on the calculated excitation energy shifts is investigated. For each among the five clusters 7HQ − $(H_2O)_n$ (for $n = 1$–3) and 7HQ − $(NH_3)_n$ (for $n = 1$–2) three series of ρ_B were generated by means of Kohn–Sham calculations for the whole environment. In each series, a different approximation for the exchange-correlation potential was used (LDA [92, 93], PBE [94], or SAOP [84, 85]) and the following slater type atomic basis sets: STO–SZ, STO–DZP, STO–TZP, STO–TZ2P, and STO–aug-TZP [83]. In all calculations, the same basis set was used for subsystem *A* (STO–aug-TZ2P) and the same approximation for the exchange-correlation energy of subsystem *A* (SAOP [84, 85]). Figure 5 shows that the complexation induced shifts are almost independent on the dipole moment of subsystem *B*.

The shifts show remarkable stability in each case. In the H_2O case, changing the basis set and the approximation for the exchange-correlation potential results in variation of the magnitude of the dipole moment from 1.77 to 1.94 Debye which is accompanied by a small variation of the excitation energy (from 3.7062 to 3.7123 eV). Since the excitation energy for the isolated chromophore is the same in each case (the same basis set and approximation for the exchange-correlation potential for subsystem *A*) the variation of the shifts is the same (they vary within 0.0061 eV).

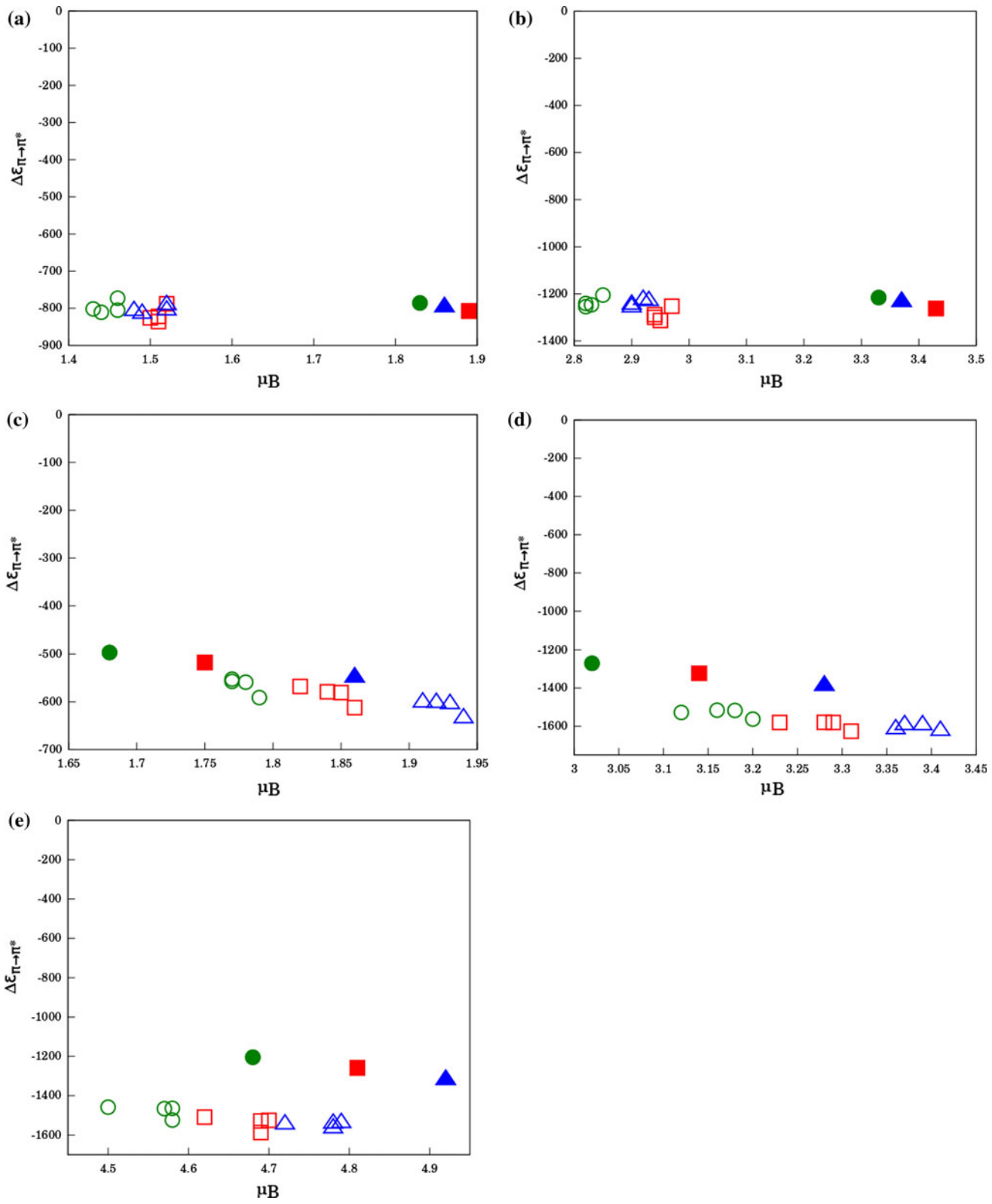


Fig. 5 Complexation induced shifts of the excitation energy ($\Delta\varepsilon_{\pi-\pi^*}$) in *cis*-7-hydroxyquinoline (7HQ) in hydrogen-bonded clusters evaluated at frozen densities (ρ_B) differing in the dipole moment (μ_B). ρ_B generated in Kohn–Sham calculations for isolated environment applying LDA (*squares*), PBE (*circles*), and SAOP (*triangles*) exchange-correlation potentials and basis sets ranging from STO–SZ to STO–aug-TZP): **a** 7HQ–NH_3, **b** 7HQ–$(NH_3)_2$, **c** 7HQ–H_2O, **d** 7HQ–$(H_2O)_2$, and **e** 7HQ–$(H_2O)_3$. The results obtained with minimal basis set (STO–SZ) are indicated with *full symbols*. The zero line, i.e., $\Delta\varepsilon = 0$, corresponds to the isolated chromophore, i.e., $v_{\text{emb}}[\rho_A, \rho_B, v_B](\mathbf{r}) = 0$

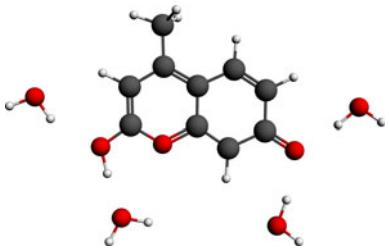


Fig. 6 The cluster of keto-7-hydroxy-4-methylcoumarin and four water molecules

The change of the dipole moment of the environment by 0.17 Debye (10 % relative change) is followed by the change of the excitation energy by 0.0061 eV (which is only 3 % of the shift). Calculations for other environments show a similar trend. Despite the fact that it is electrostatics, which provides the dominant contribution to the shifts, hardly any correlation between the dipole moment of the environment and the calculated shift in the excitation energy can be observed.

The results obtained for embedded *cis*-7-hydroxyquinoline suggest the optimal strategy to generate ρ_B consisting of using a low-end Kohn–Sham calculations for the whole isolated subsystem B which assures taking properly into account the mutual polarization of the molecules in the environment. Among the two factors determining the shifts: neglecting the mutual polarization of the molecules in the environment (the physical approximation) and the use of medium quality basis set (implementation related effect), the first one appears to be clearly more significant. In the following two sections, this recommendation is put to scrutiny using other hydrated chromophores.

3.2 Dependence of the environment-induced shifts of the excitation energies on ρ_B obtained from the Kohn–Sham calculations for the whole environment

In computational studies on solvatochromism, a large number of molecules of the solvent is explicitly included. The issue of inexpensive generation of ρ_B is critical for saving computation time. This section deals with the issue: *How far one can go in simplifying the Kohn–Sham calculations for the whole environment without deteriorating the FDET calculated shifts in the energies of local excitations?*

3.2.1 Microsolvated keto-7-hydroxy-4-methylcoumarin

The excited-state properties of keto-7-hydroxy-4-methylcoumarin (7H4MC) are widely studied experimentally and

Table 5 Excitation energies (ε) and environment-induced spectral shifts ($\Delta\varepsilon$) of the three lowest excitations obtained using different choices for ρ_B (Kohn–Sham calculations with three choices for exchange-correlation potentials and four choices for the basis sets) for the 7-hydroxy-4-methylcoumarin complex with four H_2O molecules. The corresponding excitation energy for isolated chromophore are: $\varepsilon_{\pi\to\pi^*(1)} = 2.9245$ eV, $\varepsilon_{\pi\to\pi^*(2)} = 4.4061$ eV, and $\varepsilon_{n\to\pi^*} = 2.7340$ eV. The corresponding dipole moment (in Debye units) of the environment is also given

Generation of ρ_B		Excitation Energy (eV)			Spectral shift (cm^{-1})			μ_B (Debye)
Xc-potential	Basis set	$\varepsilon_{\pi\to\pi^*(1)}$	$\varepsilon_{\pi\to\pi^*(2)}$	$\varepsilon_{n\to\pi^*}$	$\Delta\varepsilon_{\pi\to\pi^*(1)}$	$\Delta\varepsilon_{\pi\to\pi^*(2)}$	$\Delta\varepsilon_{n\to\pi^*}$	
SAOP	STO–SZ	2.7376	4.4058	3.3832	−1507	−2	5236	5.23
	STO–DZ	2.9824	3.4543	3.7178	466	−462	5810	7.62
	STO–DZP	3.0243	4.5178	3.4349	805	901	5653	5.44
	STO–TZP	3.0155	4.5113	3.4232	734	848	5559	5.39
	STO–TZ2P	3.0233	4.5168	3.4196	797	893	5530	5.42
	STO–aug-TZP	3.0241	4.5118	3.4184	803	852	5520	5.25
PBE	STO–SZ	2.8288	4.4252	3.3404	−772	154	4891	4.68
	STO–DZ	3.0106	3.7320	3.4047	694	−384	5410	7.09
	STO–DZP	3.0206	4.5109	3.4003	775	845	5374	4.94
	STO–TZP	3.0066	4.5004	3.3880	662	761	5275	4.99
	STO–TZ2P	2.9935	4.4993	3.3822	556	751	5228	5.03
	STO–aug-TZP	3.0119	4.4992	3.3712	705	751	5139	5.06
LDA	STO–SZ	2.9231	4.4632	3.3585	−11	460	5037	4.94
	STO–DZ	3.0086	3.4250	3.7256	677	−399	5573	7.50
	STO–DZP	3.0167	4.5142	3.4081	744	872	5437	5.13
	STO–TZP	3.0183	4.5111	3.3986	793	847	5360	5.19
	STO–TZ2P	3.0185	4.5114	3.3935	758	849	5319	5.22
	STO–aug-TZP	3.0152	4.5087	3.3838	732	827	5241	5.25

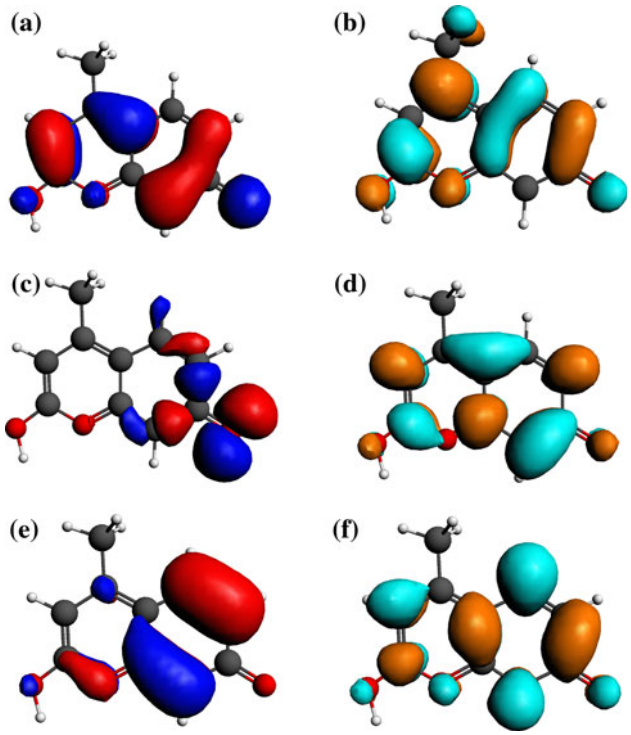


Fig. 7 The molecular orbitals in isolated keto-7-hydroxy-4-methylcoumarin, which provide the dominant contributions to the lowest excitations: **a** HOMO, **b** LUMO, **c** HOMO−1, **d** LUMO+1, **e** HOMO−2, and **f** LUMO+2

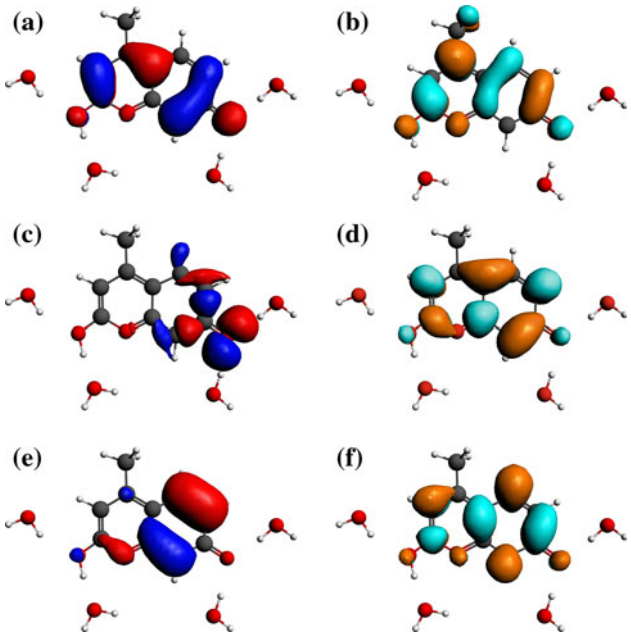


Fig. 8 The embedded orbitals in keto-7-hydroxy-4-methylcoumarin-$(H_2O)_4$ complex, which provide the dominant contributions to the lowest excitations: **a** HOMO, **b** LUMO, **c** HOMO−1, **d** LUMO+1, **e** HOMO−2, and **f** LUMO+2

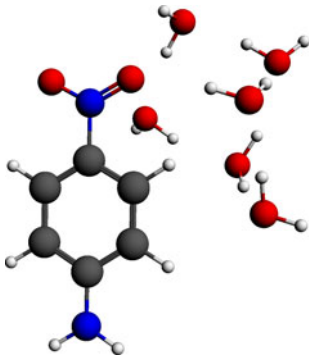


Fig. 9 *P*-Nitro aniline-$(H_2O)_6$ complex

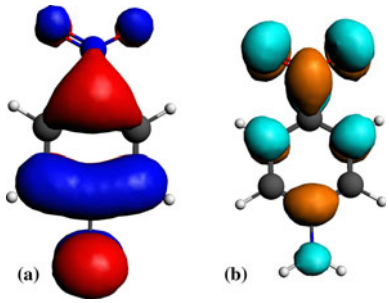


Fig. 10 The pair of molecular orbitals in isolated *P*-nitro aniline, which provides the dominant contribution to the lowest excitation: **a** HOMO, **b** LUMO

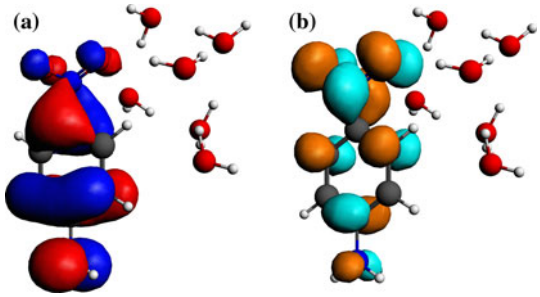


Fig. 11 The pair of molecular orbitals in the *P*-nitro aniline–$6H_2O$ complex, which provides the dominant contribution to the lowest excitation: **a** HOMO, **b** LUMO

theoretically [95–97] due to its remarkable photophysical properties. The theoretical investigations of this molecules have helped to solve the ambiguity of nature of excitation [98]. The keto form of 7H4MC in a cluster comprising four water molecules is shown in Fig. 6. The ground-state geometry of the complex is chosen from Ref. [98] to study the sensitivity of the FDET excitation energy shifts for the choice for ρ_B. Three lowest transitions (two $\pi \rightarrow \pi^*$ and one $n \rightarrow \pi^*$ excitations) are reported in Table 5. For subsystem A, STO–aug-TZP basis set and the SAOP [84, 85] approximation for the exchange-correlation potential is

used for all calculations in this subsection. ρ_B on the other hand is generated using several STO type atomic basis sets and three different approximations for the exchange-correlation potential.

The orbitals for isolated and embedded chromophore are shown in Figs. 7 and 8, respectively. The analyzed excitations are dominated by the following transitions: $\pi \rightarrow \pi^*(1)$ (HOMO to LUMO), $\pi \rightarrow \pi^*(2)$ (HOMO to LUMO−1), and $n \rightarrow \pi^*$ (HOMO−1 to LUMO). In line with the trends observed for embedded *cis*-7-hydroxyquinoline reported in the previous section, the shifts of the excitation energies depend weakly on the dipole moment of the environment provided the STO–DZ or larger basis set is used in generation of ρ_B (see Table 5).

3.2.2 Microsolvated P-nitro aniline

For *P*-nitro aniline (PNA) in six water molecules three local excitations are analyzed: the lowest singlet (S), and triplet (T) $\pi \rightarrow \pi^*$. The geometry of the complex was taken

Table 6 Excitation energies (ε) and environment-induced spectral shifts ($\Delta\varepsilon$) of a lowest singlet and triplet excitations obtained using different choices for ρ_B (Kohn–Sham calculations with three choices for exchange-correlation potentials and four choices for the basis sets) for the PNA with six H_2O molecules. The corresponding excitation energy for isolated chromophore are: $\varepsilon_{\pi\rightarrow\pi^*(S)} = 3.3622$ eV, $\varepsilon_{\pi\rightarrow\pi^*(T)} = 2.5086$ eV, and the corresponding dipole moment (in Debye units) of the environment is also given

Generation of ρ_B		Excitation energy (eV)		Spectral shift (cm^{-1})		μ_B (Debye)
Xc-potential	Basis set	$\varepsilon_{\pi\rightarrow\pi^*(S)}$	$\varepsilon_{\pi\rightarrow\pi^*(T)}$	$\Delta\varepsilon_{\pi\rightarrow\pi^*(S)}$	$\Delta\varepsilon_{\pi\rightarrow\pi^*(T)}$	
SAOP	STO–SZ	3.2549	2.3546	−865	−1242	2.52
	STO–DZ	3.2245	2.3292	−1111	−1447	3.39
	STO–DZP	3.2678	2.3611	−761	−1190	2.54
	STO–TZP	3.2653	2.3624	−782	−1179	2.56
	STO–TZ2P	3.2631	2.3615	−799	−1186	2.57
	STO–aug-TZP	3.2576	2.3603	−844	−1196	2.58
PBE	STO–SZ	3.2639	2.3647	−793	−1161	2.29
	STO–DZ	3.2265	2.3198	−1095	−1523	3.18
	STO–DZP	3.2523	2.3703	−886	−1115	2.35
	STO–TZP	3.2647	2.3700	−786	−1118	2.39
	STO–TZ2P	3.2676	2.3688	−763	−1127	2.40
	STO–aug-TZP	3.2668	2.3727	−769	−1096	2.41
LDA	STO–SZ	3.2523	2.3604	−886	−1195	2.41
	STO–DZ	3.2227	2.3280	−1125	−1457	3.35
	STO–DZP	3.2672	2.3736	−766	−1089	2.44
	STO–TZP	3.2626	2.3731	−803	−1093	2.48
	STO–TZ2P	3.2481	2.3723	−920	−1099	2.49
	STO–aug-TZP	3.2688	2.3760	−753	−1069	2.48

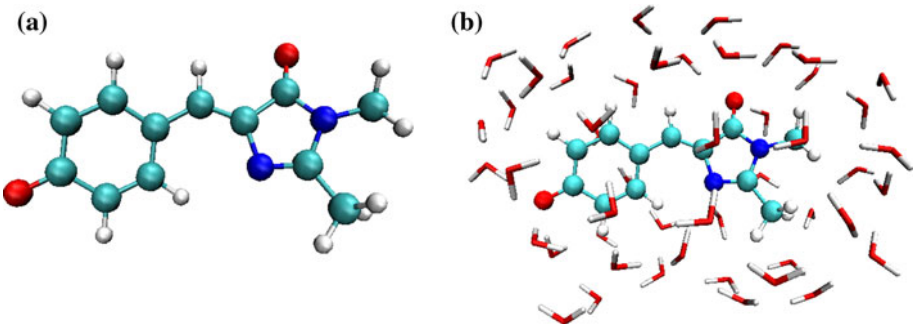


Fig. 12 Isolated (*left*) and hydrated by 49 water molecules (*right*) 4-hydroxybenzylidene-2,3-dimethylimidazolinone anion

Table 7 Dependence of the two lowest electronic excitation energies (ε_1 and ε_2) and the corresponding spectral shifts ($\Delta\varepsilon_1$ and $\Delta\varepsilon_2$) on the basis set choices for generating the Kohn–Sham electronic density of the solvent. The corresponding electronic excitation energies of isolated chromophore are: ε_1 = 2.3827 eV and ε_2 = 2.9291 eV. The corresponding dominant molecular orbital (MO) transitions are HOMO-1 → LUMO for ε_1 and HOMO → LUMO for ε_2. The magnitude of the dipole moment of the environment is also given

Basis set for ρ_B	Excitation energy (eV)		Spectral shift (cm^{-1})		μ_B (Debye)
	ε_1	ε_2	$\Delta\varepsilon_1$	$\Delta\varepsilon_2$	
STO–SZ	2.7907	3.0142	3291	686	11.72
STO–DZ	3.0054	3.0508	5022	982	13.24
STO–DZP	2.9742	3.0340	4771	846	10.66
STO–TZP	2.9605	3.0318	4660	828	10.78
STO–TZ2P	2.9556	3.0306	4621	819	10.81
STO–aug-TZP	2.9415	3.0289	4507	805	10.79

from Ref. [99] and is shown in Fig. 9. The π and π^* orbitals are shown in Figs. 10 and 11 for isolated chromophore and complex respectively. These are HOMO and LUMO for both isolated and embedded PNA. Table 6 collects the excitation energies and solvatochromic shifts for singlet and triplet $\pi \rightarrow \pi^*$ transition. The excitation energies and solvatochromic shifts for both the transitions do not vary significantly with the change of basis set or exchange correlation functional. The FDET shifts evaluated using ρ_B obtained with the STO–DZP or any larger basis set are numerically equivalent.

3.2.3 *Hydrated 4-hydroxybenzylidene-2,3-dimethylimidazolinone anion*

In this subsection, the estimation of the sensitivity of the calculated shifts on ρ_B is made for a realistic model of a chromophore in water solvent, following the same procedure as the one used in the previous subsection. The chromophore investigated is 4-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI) anion, which is the chromophore in green fluorescent protein. The cluster used for the calculations consists of the HBDI anion in the *cis* conformation and 49 water molecules (see Fig. 12), which geometry is taken from Ref. [100]. The water molecules in the cluster represents the complete first solvation shell.

STO–DZP basis sets are used for the HBDI anion. The non-additive kinetic potential in Eq. 3 is approximated using the NDSD bi-functional [64]. ρ_B is obtained from Kohn–Sham calculations for isolated subsystem B applying the LDA exchange-correlation functional and various types of STO orbitals (SZ, DZ, DZP, TZP, TZ2P, and aug-TZP).

Table 7 shows the dependence of the two lowest electronic excitation energies (ε_1 and ε_2) and the corresponding spectral shifts ($\Delta\varepsilon_1$ and $\Delta\varepsilon_2$) on the basis set choices for generating the Kohn–Sham electronic density of the solvent. The corresponding dominant molecular orbital (MO) transitions are HOMO-1 → LUMO for ε_1 and HOMO → LUMO for ε_2. The involved orbitals (both isolated and solvated HBDI anion) are shown in Fig. 13.

The results collected in Table 7 indicate that, starting from STO–DZP, increasing further the basis sets does not affect significantly the calculated spectral shifts.

The solvated 4-hydroxybenzylidene-2,3-dimethylimidazolinone anion is the largest system investigated in the present work. The CPU timings of the FDET calculations for this system are collected in Table 8 to illustrate the main computational advantage of the FDET-based methods. Solving LR-TDDFT equations for getting excitation energies requires the same time regardless which method is used to generate ρ_B (The exceptions for the last two lines in Table 8 originate from different number of iterations in the Davidson procedure to diagonalize the largest matrix in LR-TDDFT calculations. In this particular case, ρ_B is

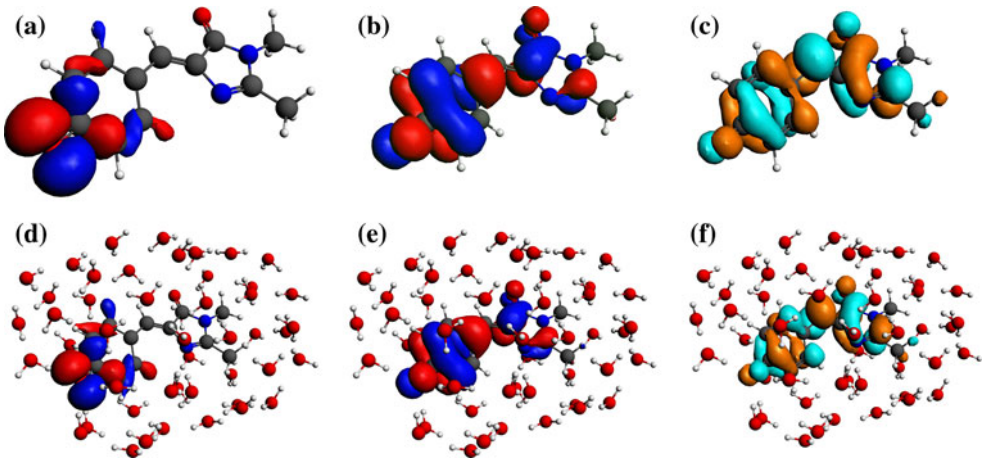


Fig. 13 The molecular orbitals, which provide the dominant contributions to the lowest two excitations. **a** HOMO−1, **b** HOMO, and **c** LUMO for the isolated 4-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI) anion; **d** HOMO−1, **e** HOMO, and **f** LUMO for the embedded HBDI anion

Table 8 The dependence of the CPU time[a] (hour:minute) of FDET calculations on the choice of basis sets used to generate ρ_B for 4-hydroxybenzylidene-2,3-dimethylimidazolinone anion in the environment consisting 49 water molecules

Basis set for ρ_B	Generation of ρ_B	FDET (ground state)	FDET/LR-TDDFT (excited state)	Total
STO–SZ	00:03	00:18	00:17	00:39
STO–DZ	00:05	00:33	00:16	00:55
STO–DZP	00:13	01:22	00:17	01:54
STO–TZP	00:17	01:36	00:16	02:10
STO–TZ2P	00:33	02:06	00:51[b]	03:30
STO–aug-TZP	01:28	03:03	00:25	04:56

[a] Each calculation is run on one node consisting of eight 3.0 GHz cores

[b] The CPU time is proportional to the number of iterations in the Davidson procedure to diagonalize the largest matrix in LR-TDDFT calculations

obtained by using ground-state Kohn–Sham calculations whereas the excited-state calculations involve the same space of occupied and unoccupied orbitals). In contrast with solving Casida's equations for excited-state FDET calculations, the time needed for ground-state FDET calculations increases with the basis set size for ρ_B. This is rather a specific feature of the ADF program concerning the evaluation of the Coulomb terms, the timing of which depends on the site of the orbital- and auxiliary- basis sets used for ρ_B. Concerning the generation of ρ_B, its scaling reflects the conventional implementation of the Kohn–Sham calculations.

4 Conclusions

According to FDET, every observable calculated at a given choice for ρ_B is a functional of ρ_B. The unique correspondence between the observable of interest and ρ_B involves the embedding potential defined in Eq. 3 which changes if ρ_B changes. The present work concerns these correspondences for a particular case of energies of local excitations which are evaluated by means of applying FDET within LR-TDDFT framework. In such calculations, the shifts are determined by the embedding potential (and its functional derivative with respect to ρ_A). The examples provided in the present work form a series constructed to address the issue: *How simple can be the method to generate ρ_B in hydrogen bonded clusters without significant deterioration of the calculated shifts in energies of local excitations?* The following conclusions/recommendations emerge from the reported studies on various organic chromophores, excitations, and hydrogen-bonded environments. Generating ρ_B as a superposition of spherically symmetric atomic charges, which was shown to be adequate in environments being ionic crystals, does not seem to be robust enough for environments consisting of molecules. The FDET excitation energies results depend critically on the choice of net atomic charges. The example of ammonia-7HQ complex shows that none of the commonly used techniques to assign atomic net charges in a molecule is applicable. This originates probably from the fact that such simple approximation is incapable of reproducing the directional character of atoms with lone pairs. Generating ρ_B as a superposition of molecular densities is a more reliable strategy but still might lead to relative errors in the shifts exceeding 25 %. Such errors occur if the molecules within the environment interact strongly with each other as in the shown examples of environments consisting of hydrogen-bonded chains of molecules. The most robust strategy to generate ρ_B emerging from the present study consists of using less expensive Kohn–Sham calculations for the whole environment (or at least for all molecules near the chromophore). STO–DZP basis set or larger and any among the three investigated exchange-correlation potentials (LDA, PBE, or SAOP) lead to similar results. The relative shift vary usually less than 0.02 eV (or about 150 cm^{-1}) whereas the relative excitation energy shifts change by no more than 10 % regardless which approximation for the exchange-correlation potential is used or which basis set (STO–DZP or larger) is used. These results can be considered as estimations of errors made due to arbitrary choices of ρ_B made in multi-scale computer simulations based on FDET.

The embedding potential given in Eq. 3 applies not only for embedding a non-interacting reference system, which is considered in the present work, but also for methods using other quantum mechanical descriptors for ρ_A as shown in Ref. [22, 23]. Such FDET-based simulations are gaining increasing popularity [24–29, 31–33]. The effect of varying ρ_B on the embedding potential and subsequently for shifts in excitation energies (and other observables directly related to the electronic structure of subsystem A) can be expected to be similar for such methods.

Concerning the relative stability of FDET calculated shifts in energies of local excitations with respect to the choices for ρ_B, we underline that in the commonly used methods of the QM/MM type, the embedding potential comprises ONLY electrostatic component. Therefore, the embedded wavefunction or other quantum mechanical descriptor obtained in such calculations depends critically on the distribution of charges and higher electric moments in the environment. Including or not the electrostatic field generated by the induced dipole moments on atoms in the environment might affect the results significantly. The present work shows that, in the FDET calculations which

are based on variational principle, the results depend less on the choice for ρ_B. Several cases, where two choices for the density of the environment differing in the dipole moment by more than 10 % lead to almost identical shifts in the excitation energy, are reported in the present work. We interpret this robustness of the FDET calculated excitation energy shifts with respect to the choice of the density of the environment (ρ_B) as originating in the variational character of FDET and the existence of multiple pairs of ρ_A and ρ_B adding up to the same total density in the case of exact theory. This numerical result shows that the use of the notion of "polarization of the environment by the embedded subsystem" might be misleading in the FDET calculations at least for such weak interactions as the hydrogen bond between the chromophore and the molecule in the environment.

Finally, it is worthwhile noticing that performing FDET-based simulations for excited states of embedded chromophores as an alternative to treat the whole system at the same quantum mechanical level leads to significant computational savings which result from two approximations: (a) neglect of dynamic response of the environment and (b) the use of inexpensive methods to generate the ρ_B—a ground-state Kohn–Sham electron density for instance. None of them is general. The first approximation leads to the reduction of the number of Kohn–Sham orbitals (occupied and unoccupied) in solving LR-TDDFT equations and is applicable in the absence of overlap in the energy levels in the two subsystems. In such cases going beyond NDRE is indispensable [46]. The second approximation is also not general. The present work, indicates that using inexpensive Kohn–Sham-based methods to generate ρ_B such as using local-density approximation for exchange-correlation potential and STO–DZP basis set, seems a universally adequate approximation for hydrogen-bonded environments. This general recommendation complements our conclusions emerging from studies concerning other environments such as ionic solids [38] or cases where confinement dominate the environment-induced shifts of properties of embedded molecules [34], for which it was found that approximating the density of the solid by a sum of ionic densities is an adequate approximation.

5 Supporting information

Results obtained with other basis sets, numerical values of excitation energies.

Acknowledgments Grants from Swiss National Science Foundation (200020/134791/1 FNRS) and COST (CODECS) are greatly appreciated.

References

1. Åquist J, Warshel A (1993) Chem Rev 93:2523
2. Bakowies D, Thiel W (1996) J Phys Chem 100(25): 10580–10594
3. Sauer J, Ugliengo P, Garrone E, Sounders VR (1994) Chem Rev 94:2095
4. Cramer CJ, Truhlar DG (1999) Chem Rev 99(8):2161–2200
5. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV (2012) Chem Rev 112(1):632–672
6. Zheng H (1993) Phys Rev B 48:14868–14883
7. Laio A, VandeVondele J, Rothlisberger U (2002) J Chem Phys 116(16):6941–6947
8. Fradelos G, Wesolowski TA (2011) J Phys Chem A 115(35):10018–10026
9. Fradelos G, Wesolowski TA (2011) J Chem Theory Comput 7(1):213–222
10. Sanchez ML, Aguilar MA, Olivares del Valle FJ (1997) J Comput Chem 18(3):313–322
11. Coutinho K, Georg HC, Fonseca TL, Ludwig V, Canuto S (2007) Chem Phys Lett 437(1–3):148–152
12. Casanova D, Gusarov S, Kovalenko A, Ziegler T (2007) J Chem Theory Comput 3(2):458–476
13. Engkvist O, strand P-O, Karlstrm G (2000) Chem Rev 100(11):4087–4108
14. Nanda KD, Beran GJO (2012) J Chem Phys 137(17)
15. Gresh N, Cisneros GA, Darden TA, Piquemal J-P (2007) J Chem Theory Comput 3(6):1960–1986
16. Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Krauss M, Garmer D, Basch H, Cohen D (1996) J Chem Phys 105(5): 1968–1986
17. Assfeld X, Rivail JL (1996) Chem Phys Lett 263(1–2):100–106
18. Winter NW, Pitzer RM, Temple DK (1987) J Chem Phys 86(6):3549–3556
19. Moriarty JA, Phillips R (1991) Phys Rev Lett 66(23):3036–3039
20. Wesolowski TA, Warshel A (1993) J Phys Chem 97(30): 8050–8053
21. Wesolowski TA (2006) One-electron equations for embedded electron density: challenge for theory and practical payoffs in multi-level modelling of soft condensed matter. In: Leszczynski J (ed) Computational chemistry: reviews of current trends, vol X. World Scientific, Singapore, pp 1–82
22. Wesolowski TA (2008) Phys Rev A 77(1):012504(1–8)
23. Pernal K, Wesolowski TA (2009) Int J Quant Chem 109(11, Sp. Iss. SI):2520–2525
24. Stefanovich EV, Truong TN (1996) J Chem Phys 104(8): 2946–2955
25. Govind N, Wang YA, da Silva AJR, Carter EA (1998) Chem Phys Lett 295(12):129–134
26. Trail JR, Bird DM (2000) Phys Rev B 62(24):16402–16411
27. Hodak M, Lu W, Bernholc J (2008) J Chem Phys 128:014101
28. Neugebauer J, Baerends EJ (2006) J Phys Chem A 8786–8796
29. Lahav D, Kluner T (2007) J Phys Cond Matt 19:226001
30. Neugebauer J (2010) Phys Rep 489:1-87
31. Gomes ASP, Jacob CR, Visscher L (2008) Phys Chem Chem Phys 10:5353
32. Roncero O, de Lara-Castells MP, Villarreal P, Flores F, Ortega J, Paniagua M, Aguado A (2008) J Chem Phys 129(18):184104
33. Goodpaster Jason D, Ananth Nandini, Manby Frederick R, Miller Thomas F III (2010) J Chem Phys 133(8):084103
34. Wesolowski TA (1999) Chem Phys Lett 311(1–2):87–92
35. Hong GY, Strajbl M, Wesolowski TA, Warshel A (2000) J Comput Chem 21(16):1554–1561
36. Zbiri M, Atanasov M, Daul C, Garcia-Lastra JM, Wesolowski TA (2004) Chem Phys Lett 397(4–6):441–446

37. Wesolowski TA (2004) J Am Chem Soc 126(37):11444–11445
38. Zhou X, Wesolowski Tomasz A, Tabacchi G, Fois E, Calzaferri G, Devaux A (2013) Phys Chem Chem Phys 15:159–167
39. Neugebauer J, Louwerse MJ, Baerends EJ, Wesolowski TA (2005) J Chem Phys 122(9):094115
40. Jacob CR, Visscher L (2006) J Chem Phys 125:194104
41. Jacob CR, Beyhan SM, Visscher L (2007) J Chem Phys 126:234116
42. Gotz AW, Beyhan SM, Visscher L (2009) J Chem Theory Comput 5(12):3161–3174
43. Kiewisch K, Eickerling G, Reiher M, Neugebauer J (2008) J Chem Phys 128(4):044114(1–15)
44. Casida ME (1995) Time-dependent density-functional response theory for molecules. In: Chong DP (ed) Recent advances in density-functional methods. World Scientific, Singapore
45. Casida ME, Wesolowski TA (2004) Int J Quantum Chem 96(6):577–588
46. Neugebauer J (2007) J Chem Phys 97:134116
47. Fradelos G, Lutz JJ, Wesolowski TA, Piecuch P, Woch M (2011) J Chem Theory Comput 7(6):1647–1666
48. Hohenberg P, Kohn W (1964) Phys Rev 136(3B):B864–B871
49. Elliott P, Cohen MH, Wasserman A, Burke K (2009) J Chem Theory Comput 5(4):827–833
50. Elliott P, Burke K, Cohen MH, Wasserman A (2010) Phys Rev A 82:024501
51. Kohn W, Sham LJ (1965) Phys Rev 140(4A):A1133–A1138
52. Severo Pereira Gomes A, Jacob CR (2012) Annu Rep Prog Chem Sect C Phys Chem 108:222–277
53. Levy M (1979) Proc Natl Acad Sci U S A 76(12):6062–6065
54. Levy M (1982) Phys Rev A 26(3):1200–1208
55. Tran F, Wesolowski TA (2003) J Chem Phys 118:2072–2080
56. Thomas LH (1927) Proc Camb Philos Soc 23:542
57. Fermi E (1928) Z Phys 48:73
58. von Weizsäcker CF (1935) Z Phys 96:431
59. Kirzhnits DA (1957) Sov Phys JETP 5:64
60. Wesolowski TA, Weber J (1997) Int J Quantum Chem 61(2):303–311
61. Wesolowski TA, Chermette H, Weber J (1996) J Chem Phys 105(20):9182–9190
62. Wesolowski TA (1997) J Chem Phys 106(20):8516–8526
63. Lembarki A, Rogemond F, Chermette H (1995) Phys Rev A 52(5):3704–3710
64. Garcia Lastra JM, Kaminski JW, Wesolowski TA (2008) J Chem Phys 129(7):074107(1–15)
65. Wesolowski TA, Weber J (1996) Chem Phys Lett 248(1–2): 71–76
66. Bernard YA, Dulak M, Kaminski JW, Wesolowski TA (2008) J Phys A Math Theor 41(5):055302(1–19)
67. Fux S, Kiewisch K, Jacob CR, Neugebauer J, Reiher M (2008) . Chem Phys Lett 461(46):353–359
68. Fux S, Jacob CR, Neugebauer J, Visscher L, Reiher M (2010) J Chem Phys 132(16):164101(1–18)
69. Huang C, Pavone M, Carter EA (2011) J Chem Phys 134(15):154110(1–11)
70. Cortona P (1991) Phys Rev B 44(16):8454–8458
71. Wesolowski TA (2005) Mol Phys 103(6-8):1165–1167
72. Savin A, Wesolowski TA (2009) Prog Theor Chem Phys 19:327–339
73. Savin A, Wesolowski TA (2013) Non-additive kinetic energy and potential in analytically solvable systems and their approximate counterparts. In: Wesolowski TA, Wang YA (eds) Recent progress in orbital-free density functional theory. Recent Progress in Computational Chemistry, vol 6. World Scientific, Singapore, pp 275–295
74. de Silva P, Wesolowski TA (2012) J Chem Phys 137(9):094110
75. Gritsenko OV (2013) On the principal difference between the exact and approximate frozen-density embedding theory. In: Wesolowski TA, Wang YA (eds) Recent progress in computational chemistry, vol 6. World Scientific, Singapore, pp 355–365
76. Dulak M, Wesolowski TA (2006) J Chem Phys 124:164101
77. Fradelos G, Lutz JJ, Wesolowski TA, Piecuch P, Wloch M (2012) Shifts in excitation energies induced by hydrogen bonding: A comparison of the embedding and supermolecular time-dependent density functional theory calculations with the equation-of-motion coupled-cluster results. In: Hoggan PEE, Brandas EJJ, Maruani J, Piecuch P, Delgado-Barrio G (eds) Advances in the theory of quantum systems in chemistry and physics. Progress in theoretical chemistry and physics, vol 22. Springer, Netherlands, pp 219–248
78. Wesolowski T, Warshel A (1994) J Phys Chem 98(20): 5183–5187
79. Wesolowski T, Muller RP, Warshel A (1996) J Phys Chem 100(38):15444–15449
80. Adf, scm, theoretical chemistry. Vrije Universiteit, Amsterdam, the Netherlands. http://www.scm.com
81. Dulak M, Wesolowski TA (2005) Int J Quantum Chem 101(5):543–549
82. Jacob CR, Neugebauer J, Visscher L (2008) J Comput Chem 29:1011
83. van Lenthe E, Baerends EJ (2003) J Comput Chem 24(9): 1142–1156
84. Gritsenko OV, Schipper PRT, Baerends EJ (1999) Chem Phys Lett 302(34):199–207
85. Schipper PRT, Gritsenko OV, van Gisbergen SJA, Baerends EJ (2000) J Chem Phys 112(3):1344–1352
86. Perdew JP (1991) Electronic structure of solids'91, vol X. Academie Verlag, Berlin
87. Fradelos G, Kaminski JW, Wesolowski TA, Leutwyler S (2009) J Phys Chem A 113:9766
88. Mulliken RS (1955) J Chem Phys 23:1833–1840
89. Hirschfeld FL (1977) Theor Chim Acta 44:129–138
90. Baerends EJ, Bickelhaupt FM, Guerra CF, Handgraaf JW (2003) J Comp Chem 25:189–210
91. Snijders JG, Swart M, Van Duijnen PTh (2001) J Comp Chem 22:79–88
92. Dirac PAM (1930) Proc Camb Philos Soc 26:376–385
93. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58(8): 1200–1211
94. Perdew JP, Burke K, Ernzerhof M (1997) Phys Rev Lett 78(7):1396–1396
95. Seixas de Melo J, Fernandes PF (2001) J Mol Struct 565–566: 69–78
96. Novak I, Kovač B (2001) J Phoochem Photobiol 113:913
97. Kovač B, Novak I (2002) Spectrochim Acta A 58:14831488
98. Georgieva I, Trendafilova N, Aquino A, Lischka H (2005) J Phys Chem A 109:11860–11869
99. Slipchenko LV (2010) J Phys Chem A 114:8824–8830
100. Polyakov I, Epifanovsky E, Grigorenko B, Krylov AI, Nemukhin A (2009) J Chem Theory Comput 5(7):1907–1914