Predicting solvatochromic shifts and colours of a solvated organic dye: The example of Nile Red

T. J. Zuehlsdorff,[1,2,*] P. D. Haynes,[3,4] M. C. Payne,[2] and N. D. M. Hine[5]

[1]*School of Natural Sciences, University of California Merced, 5200 N. Lake Road, Merced, CA 95343, USA*

[2]*Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, UK*

[3]*Department of Physics, Imperial College London, Exhibition Road, London SW7 2AZ, UK*

[4]*Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, UK*

[5]*Department of Physics, University of Warwick, Coventry, CV4 7AL, UK*

(Dated: March 9, 2017)

Abstract

The solvatochromic shift, as well as the change in colour of the simple organic dye nile red is studied in two polar and two non-polar solvents in the context of large-scale time-dependent density-functional theory (TDDFT) calculations treating large parts of the solvent environment from first principles. We show that an explicit solvent representation is vital to resolve absorption peak shifts between nile red in *n*-hexane and toluene, as well as acetone and ethanol. The origin of the failure of implicit solvent models for these solvents is identified as being due to the strong solute-solvent interactions in form of π-stacking and hydrogen bonding in the case of toluene and ethanol. We furthermore demonstrate that the failures of the computationally inexpensive PBE functional in describing some features of the excited state potential energy surface of the S_1 state of nile red can be corrected for in a straightforward fashion, relying only on a small number of calculations making use of more sophisticated range-separated hybrid functionals. The resulting solvatochromic shifts and predicted colours are in excellent agreement with experiment, showing the computational approach outlined in this work to yield very robust predictions of optical properties of dyes in solution.

*Electronic address: tzuehlsdorff@ucmerced.edu

I. INTRODUCTION

The reliable prediction of solvatochromic shifts and colours of organic dyes in solution is challenging for a number of different reasons. On the one hand, it is important to accurately compute the excited states of the solute in isolation, on the other hand the influence of the solvent molecules, both through direct interactions with the solute like hydrogen bonding, as well as polarisation effects in the ground and the excited state, need to be correctly accounted for. If one is interested in the shape of the absorption spectrum rather than just the magnitude of the solvatochromic shift, it is necessary to correctly describe the temperature-dependent dynamics of the dye in solution in order to obtain a realistic sample over possible conformations which make up the spectral shape.

A number of these requirements are mutually exclusive. For example, while higher order quantum chemistry methods allow for the precise computation of excited states in small organic dyes, the high computational cost associated with them generally makes it prohibitive to treat any significant part of the solvent at the same level of theory. Thus in practice, a compromise has to be found between the correct treatment of the dye and the appropriate treatment of the environment, in order to maximise the usefulness of the method in making practical predictions.

A highly popular approach for achieving an efficient compromise between computational cost and accuracy for solvated system comes in the form of mixed quantum mechanical/classical methods (QM/MM)[1, 2], where the chromophore and parts of its surroundings are treated at a quantum mechanical level, while the long range electrostatic screening of the environment is introduced via a classical model. QM/MM approaches have been successfully applied to to the modelling of chromophores in complex environments, such as pigments in pigment-protein complexes[3–5]. However, recent work has shown that a range of properties of interest of systems embedded in complex environments, including the absorption properties of small solvated dyes, require the explicit inclusion of large parts of the embedding environment into the QM region in order to obtain consistently converged results[6–11]. In a recent work we demonstrated that to converge the solvatochromic shift of alizarin in water, system sizes of $O(1000)$ atoms have to be treated fully quantum mechanically[11] in order to capture both short range polarisation and long range electrostatic effects on the excitation energies. The necessity of treating large parts of the solvent environment on the same

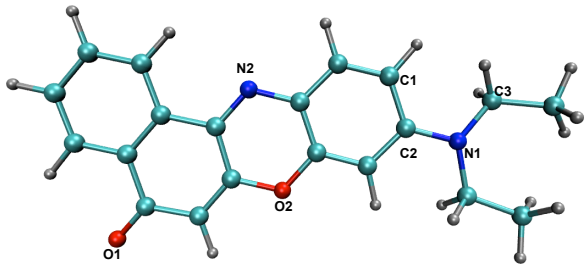


FIG. 1: Ball and stick representation of nile red. The dihedral angle responsible for rotation of the diethylamino group with respect to the aromatic ring system is defined via C1-C2-N1-C3. The figure was produced using VMD[24].

footing with the solute causes the calculations of optical properties of solvated systems to be computationally very challenging.

In recent years a large number of computational studies aiming at predicting solvatochromic shifts and colours of dyes in complex environments[11–15] have relied on time-dependent density-functional theory (TDDFT)[16–18] as the method of choice for calculating optical absorption spectra. While TDDFT has a variety of well known deficiencies in describing certain types of excitations, most notably Rydberg-type states, long-range charge-transfer (CT) states and transitions with double-excitation character[19], it has become the method of choice for a wide variety of applications in fields ranging from materials science to computational biology. The success of TDDFT is largely due to its favourable balance between accuracy and computational efficiency for most common choices of exchange-correlation functionals[19]. Furthermore, recently developed linear-scaling implementations[20–22] allow for the efficient computation of systems sizes reaching thousands of atoms[23], making these approaches an ideal tool for computing optical properties of solvated systems, while treating solvent polarisation effects on a fully quantum mechanical level.

In this work, we consider nile red, a standard fluorescent dye soluble in organic solvents and alcohol, whose absorption and fluorescence properties are highly sensitive to its environment[25–27]. The dye consists of a 2-diethylamino group connected to an aromatic ring system via a single C-N bond (see Fig. 1) and its absorption spectrum in the visible region is dominated by the S_1 state corresponding to a HOMO-LUMO transition with strong oscillator strength. Studies with both TDDFT[28] and higher order quantum chemistry

methods [29] have shown that rotation around the C-N bond is a key feature influencing the dye's absorption and fluorescence spectra. Nile red is very stable in alkaline aqueous environments[25] compared to other dyes like phenol blue[30] and its fluorescent properties have proven to be useful in a number of biological applications, for example as a stain for the detection of intracellular lipid droplets [31]. For this reason, the absorption and emission properties of nile red are well characterised experimentally in a large number of different solvent environments[25, 32].

Nile red has been the focus of a considerable number of theoretical studies[7, 28, 29, 33–36], aimed at explaining its absorption and fluorescence properties in solution. For the purpose of this work, we investigate how well both the solvatochromic shift and the colour of the dye in solution can be predicted using state-of-the-art large-scale TDDFT calculations for the solute and large parts of the solvent environment. The paper is organised as follows: Section II contains a short discussion of the available experimental results for solvatochromic shifts in nile red. We then introduce the relevant theoretical background for this work, including the representation of the solvent environment in different computational approaches, the prediction of colours from TDDFT results and a full outline of the computational method used in this work. Section IV outlines the computational details of the study carried out on nile red, including details on the generation of representative snapshots of the dye in different solvent environments. In Section V the results are presented, demonstrating that an explicit representation of the solvent environment at the TDDFT level is vital in order to obtain correct solvatochromic shifts. We also show that while semi-local exchange correlation functionals fail to predict a number of features of the S_1 excited state of the dye, these shortcomings can be corrected for by applying a *spectral warping* technique[12] that relies on only a small number of calculations using more sophisticated hybrid functionals. The resulting method is demonstrated to yield solvatochromic shifts and colours that are in excellent agreement with experimental results for a range of polar and non-polar solvents studied.

II. EXPERIMENTALLY OBSERVED SOLVATOCHROMIC SHIFTS IN NILE RED

The experimental absorption spectrum of nile red in the visible region is characterised by a single strong feature. The absorption maximum undergoes a spectral shift that is roughly correlated with increasing solvent polarity[25]. Furthermore, in some non-polar solvents like cyclohexane, the single absorption feature is found to exhibit two separate identifiable peaks, while in solvents with stronger polarity only a single peak is found. There are a number of further subtle solvent-dependent changes to the spectra that go beyond a simple shift of the peak position, with nonpolar solvents cyclohexane and 2,2,4-trimethylpentane showing an asymmetric absorption peak profile with a tail in the blue wavelengths, while more polar solvents like acetone produce a more broadened peak profile[25] (see also section 8 in the supplementary material). The significant changes to the absorption spectrum in the visible region cause solutions stained with nile red to differ strongly in colour depending on the solvent used.

The fact that some dyes exhibit solvatochromic shifts whose magnitudes correlate with increasing solvent polarity has lead to a number of theoretical models aimed at describing and predicting the spectral shifts through macroscopic solvent properties alone. Using perturbation theory and making a number of limiting approximations (such as assuming that the solvent forms a spherical cavity around the solute and neglecting any direct solvent-solute interactions through, for example, hydrogen bonding) it is possible to derive a simple expression relating the solvatochromic shift of a dye to only the relative static permittivity ϵ_0 and the relative optical permittivity ϵ_{opt} of the solvent[30, 37]:

$$\Delta E_{\text{solv}} = A\frac{\epsilon_{\text{opt}} - 1}{2\epsilon_{\text{opt}} + 1} + B\left(\frac{\epsilon_0 - 1}{\epsilon_0 + 2} - \frac{\epsilon_{\text{opt}} - 1}{\epsilon_{\text{opt}} + 2}\right) \tag{1}$$

Here, A and B are constants that are dependent on properties of the solute alone, namely the change in dipole moment between the ground state and the excited state of interest and the radius of the solvent cavity.

It was noted by Davis *et al.* that while nile red exhibits a spectral shift whose magnitude correlates with solvent polarity, there are also a number of clear outliers from this general trend[25]. To demonstrate this we take a selection of 20 solvents for which the peak absorption wavelengths of nile red are published[32], ranging from non-polar to strongly polar, and

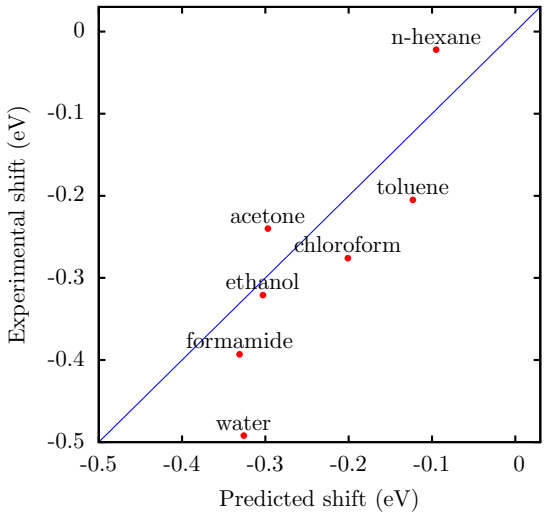


FIG. 2: Solvatochromic shifts of nile red in selected polar and nonpolar dyes as predicted by Eqn. 1 plotted against the actual experimental solvatochromic shifts. The constants in Eqn. 1 were obtained by fitting to a selection of 20 solvents and were found to be A=-0.51 eV and B=-0.32 eV.

fit constants A and B to the data (see Section 1 of the supplementary material for details regarding the fitting). A plot of the solvatochromic shift for selected solvents as predicted by Eqn. 1 after fitting constants A and B to the 20 solvents, in comparison to the actual solvatochromic shifts measured in experiments can be found in Fig. 2.

As can be seen from Fig. 2 the overall trend of increasing solvatochromic shift with increasing solvent polarity is followed, but there are notable outliers. For example, formamide is predicted to produce a slightly larger solvatochromic shift than water, but in experimental results water actually exhibits a stronger solvatochromic redshift. Furthermore, an interesting anomalous behaviour is found for n-hexane and toluene as well as acetone and ethanol, which show a very similar predicted shift following Eqn. 1, but exhibit a strong difference in shifts in the experimental data. More precisely, the simple classical model of the solvent based on perturbation theory predicts a difference in solvatochromic shifts between acetone and ethanol of the order of less than 10 meV, while the experimental data shows that nile red in ethanol produces a peak that is redshifted by more than 80 meV compared to the peak in acetone. In the non-polar solvents n-hexane and toluene, the discrepancy between the model and the experimental results is even more pronounced, with the predicted difference in solvatochromic shifts being of the order of 30 meV compared to the 180 meV found in

experimental results.

The shortcomings of the simple model in Eqn. 1 with regards to reproducing the correct solvatochromic shifts for a number of solvents can most likely be ascribed to effects ignored in the fully macroscopic description of the solvent, primarily any strong solute-solvent interactions. This effect is very pronounced in toluene and n-hexane, as well as ethanol and acetone, where the very similar predicted solvatochomic shifts suggest that the macroscopic dielectric properties of the solvents are very similar, but their interactions with nile red differ profoundly. To be considered reliable, a computational method to predict solvatochromic shifts of organic dyes in solution must therefore be able to correctly resolve these effects and yield a correct description of the splittings observed between the absorption peak positions in varying solvents. For the remainder of this work, we will therefore limit ourselves to study nile red in n-hexane, toluene, acetone, and ethanol, since the solvents show clear anomalous behaviour in the simple semiclassical model and are standard examples of polar and nonpolar solvents.

III. THEORETICAL BACKGROUND

In this section, we will briefly outline the most common computational approaches for dyes in solution in the context of TDDFT as the method of choice for obtaining excitation energies. A special focus will be placed on the prediction of optical absorption spectra and colours, as well as the treatment of solvent polarisation effects. Furthermore, we will outline defining features of the approach used for predicting these properties within the ONETEP code[38, 39], which will be the computational method of choice for the remainder of this work.

A. Treating the environment: Implicit and Explicit Solvation models

An explicit QM treatment of the solute and large parts of the solvent at the TDDFT level has a number of advantages when attempting to compute absorption spectra of solvated systems. The explicit inclusion of solvent atoms allows for the consistent treatment of solute-solvent interactions, for example through hydrogen bonding, as well as polarisation effects on the ground state of the solute. In the excited state, a QM treatment of the

solvent atoms is crucial to account for the self-screening of the TDDFT transition dipole moment due to the delocalisation of fractions of the transition density onto nearest neighbour solvent molecules[11]. The increasing availability of computational resources, as well the development of efficient algorithms[22, 44] in recent years have made an explicit treatment of large parts of the solvent environment at the DFT/TDDFT level feasible[11, 12, 14]. However, QM/MM approaches to computing optical properties of solvated system remain computationally challenging, both due to the need for full TDDFT calculations on large system sizes and the fact that a number of solute-solvent conformations have to be averaged over in order to obtain the average effect of solvent molecules on the absorption properties of the dye. A computationally much cheaper approach to computing optical properties of solvated systems comes in the form of implicit solvent models, where the entire solvent environment is represented through a uniform dielectric medium matching the macroscopic dielectric properties of the solvent.

In implicit solvation models, excitation energies are obtained by assuming a separation of time scales. While the electronic ground state of the molecule in question is screened by a solvent modelled through the relative static permittivity ϵ_0, it is assumed that only the fast electronic degrees of freedom of the solvent can react to the perturbation caused by an excitation. These fast degrees of freedom are modelled through the relative optical permittivity ϵ_{opt}. The contribution to the excitation energy caused through screening by ϵ_{opt} can be accounted for in broadly two different ways, the so-called linear-response approach (LR)[40] and the state-specific approach (SS)[41, 42]. In the LR approach the solvent polarisation as reaction to the excitation I is dependent on its transition density $\rho_I^{\{1\}}$, while in the SS approach it is evaluated using the density difference $\Delta\rho_I$ between the ground state and the excited state. Since the SS approach takes into account relaxation effects of the excited state density, it generally produces more accurate results for excited states with a strong charge transfer character but is not expected to significantly improve the description of localised states of interest here whilst being markedly more computationally intensive. Both the LR and the SS method constitute a computationally highly efficient way to account for solvent effects, but can encounter difficulties for non-polar solvents[43] or in systems with strong solute-solvent interactions.

In this work, we will rely on an explicit treatment of large regions of solvent environment at the TDDFT level in order to fully capture solute-solvent interactions and polarisation

effects in the ground and excited state from first principles. However, a number of calculations will be performed by making use of an LR implicit solvation model for the solvent degrees of freedom in order to compare and contrast the influence of the different treatments of the solvent environment on the calculated absorption properties of nile red in solution.

B. Colour prediction

Predicting the colour of a dye in solution can be seen as a more stringent test of the quality of computational approach chosen in predicting optical properties of dyes in complex environment than just focussing on predicting the solvatochromic shift, as the colour is sensitive to the absorption spectrum in the entire visible region. To predict changes of colours of dyes in different solutions correctly, it is necessary to predict both absolute spectral shifts, as well as solvent-dependent broadenings or other effects that go beyond simple rigid shifts of the entire absorption peak.

In order to obtain a prediction of the colour of a dye in solution from a TDDFT absorption spectrum, we follow the method of tristimulus colorimetry theory, as outlined by Cysewski *et al.*[45] and Baroni *et al.*[14] in the context of TDDFT. We define the intensity $I_{\text{trans}}(\lambda)$ of light transmitted through a sample of the dye in solution via

$$I_{\text{trans}}(\lambda) = I_0(\lambda)e^{-\kappa(\lambda)x} \tag{2}$$

where x is the sample thickness, $\kappa(\lambda)$ is the absorption coefficient of the dye in solution obtained from TDDFT and $I_0(\lambda)$ is the spectrum of a standard illuminant representing the light incident on the sample. In line with previous work on dyes in solution[14], we choose the D65 illuminant representing standard daylight for the purpose of this work. From the standard tristimulus colourmatching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$, the three colour indices X, Y and Z can be computed via[12]

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = N \int I_{\text{trans}}(\lambda) \begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} \mathrm{d}\lambda \tag{3}$$

where N is a normalisation factor. The values X, Y and Z can then be transformed into an RGB colour code[14] giving the colour of the dye in solution under the chosen illuminant.

C. Computing spectra of solvated dyes using the ONETEP code

The computational approach used for predicting solvatochromic shifts and colours of nile red in the four selected solvents in this work utilises the linear-scaling DFT code ONETEP[38, 39]. In ONETEP, the single particle Kohn-Sham density matrix is represented in terms of atom-centered, strictly localised functions $\{\phi_\alpha\}$ (referred to as nonorthogonal generalised Wannier functions or NGWFs) such that[38]

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) P^{\{v\}\alpha\beta} \phi_\beta(\mathbf{r}') \tag{4}$$

where $\mathbf{P}^{\{v\}}$ is the valence density kernel, the ground state density matrix in the representation of the localised orbitals $\{\phi_\alpha\}$. The orbitals themselves are expanded in an underlying basis equivalent to plane waves[46]. This allows for the NGWFs to be optimised *in situ* during a calculation to ideally represent the ground state density matrix of the system, thus only requiring a minimal number of localised functions to yield a representation of the occupied Kohn-Sham space to the same accuracy as plane-wave codes[38, 39]. The ONETEP formalism has been extended to the treatment of unoccupied Kohn-Sham states, where a second set of NGWFs $\{\chi_\beta\}$ is optimised in order to yield an ideal representation of a low energy subset of the conduction space[47]. Together, $\{\chi_\beta\}$ and $\{\phi_\alpha\}$ form a very efficient representation for excited states calculations using TDDFT[22, 23]. The minimal size of the NGWF representation, as well as its strict localisation properties, allow for highly efficient calculations of low lying optical excitations in systems containing thousands of atoms, making the methodology ideal for the treatment of organic dyes in solution[11, 48].

The key feature of the computational approach used in this work is to perform large-scale TDDFT calculations in order to obtain absorption spectra of snapshots of dyes in solution generated with classical molecular dynamics (MD), where a large part of the surrounding solvent environment is treated at the same level of theory as the dye. The full absorption spectrum is then constructed by averaging over the individual MD snapshots. The explicit inclusion of the solvent environment in the TDDFT calculation allows us to fully capture both ground state and excited state polarisation effects from first principles.

While a classical force field is likely not able to reproduce the ground state potential energy surface of the solute to the same accuracy as *ab initio* MD, it allows us to access larger length- and timescales, thus facilitating the generation of fully uncorrelated solvent

snapshots as well as eliminating potential errors in the solvent structure around nile red due to periodic boundary conditions. For the TDDFT calculation, an effective system of nile red together with an extended solvation shell is cut out of the much larger classical MD periodic box, where the extent of the solvation shell is defined through a cutoff radius R_{shell} centered on all heavy atoms in nile red and solvent molecules are retained for the TDDFT system if their centre of mass lies within the volume $V(R_{\text{shell}})$ defined by the combined volume of the spheres on all heavy atoms. To avoid a spurious closure of the band gap[49] and reduce spurious charge transfer states in polar solvents that have been reported for both (semi)-local and hybrid exchange-correlation functionals[9, 11], the entire TDDFT system containing the solute and the explicit solvation shell is placed in an implicit solvent model[50] with ϵ_0 set to the relative static permittivity of the solvent in question. This guarantees that the explicit solvent molecules on the edge of the explicit solvent volume are correctly screened and the long range continuum effects of the solvent are accounted for.

As discussed in previous work[11, 51], the representation of the TDDFT response density matrix in terms of localised atom-centered orbitals $\{\chi_\beta\}$ and $\{\phi_\alpha\}$ allows us remove certain unwanted excitations from the subspace of solutions to the TDDFT eigenvalue problem by placing appropriate constraints on the sparsity pattern of the response density matrix. This technique can be used to eliminate charge-transfer excitations from the solute to the edge of the solvation shell included in the TDDFT calculation for polar solvents[9, 11], which occur due to the underestimation of excitation energies for states with long-range charge transfer character. Here, we make use of the truncation method by defining a second cutoff radius R_{CT}, with $R_{CT} < R_{\text{shell}}$. The truncation of the response density matrix is then used to prevent any charge transfer excitations from the volume $V(R_{CT})$ to the region included in $V(R_{\text{shell}})$ but not $V(R_{CT})$. Previous results on alizarin in water suggest that low energy excited states of organic dyes in solution retain a relatively localised character and that $R_{CT} \approx 7 - 8\,\text{Å}$ yields results that are very well converged compared with equivalent calculations where no truncation of the response density matrix is applied[11].

IV. COMPUTATIONAL SETUP

In order to generate a number of uncorrelated snapshots of nile red in *n*-hexane, toluene, ethanol and acetone, we make use of the classical molecular dynamics package AMBER[52].

As previously mentioned, rotations around the dihedral angle linking the aromatic ring system with the 2-diethylamino group (defined through C1-C2-N1-C3 in Fig. 1) have a strong influence on the absorption properties of the dye. The reason for this can be seen in the excited state potential energy surface of the S_1 state, which shows two distinct minima, one at 0° twist angle corresponding to a locally excited state and one at a twist angle of 90° with considerable charge transfer (CT) character as computed (semi)-local exchange-correlation functionals. For a correct description of the absorption properties it is therefore essential to correctly describe the ground state potential energy surface of nile red with respect to rotations around the dihedral angle, as the structure of the dye evidently has a significant influence on the character of the S_1 excited state of interest. Thus in accordance with previous studies[36] we find it necessary to reparameterise the standard AMBER force field for the dihedral angle in question (see Section 2 of the supplementary material for further details).

After a suitable reparameterised force field for nile red has been found, we use AMBER to create configurations of the dye in its solvated state. As a first step, a large solvent box is created for each of the solvents considered in this work, with nile red placed at the centre, where we aim to have a similar number of atoms in each of the four solvent boxes. The boxes chosen contain 716 solvent molecules for *n*-hexane, 1046 solvent molecules for toluene, 1329 molecules for acetone, and 1855 for ethanol. A 20 ps temperature equilibration in the NVT ensemble is performed on each solvent box, where the temperature is raised gradually from 0 to 300 K, using the Langevin thermostat with a collision frequency of 1 ps^{-1}. This is followed by a 320 ps pressure equilibration in the NPT ensemble, where the pressure is kept constant at 1 atm. Finally, we perform a 200 ps production run in the NVT ensemble on the fully equilibrated system. Snapshots of the production run are taken every 4 ps, generating 50 snapshots in total for each solvent considered. The last 35 of these snapshots, corresponding to a 140 ps trajectory, are taken as an input for the TDDFT calculations that form the centre of this work. All calculations are performed using a time step of 2 fs.

We furthermore perform an AMBER MD calculation of nile red in vacuum in order to obtain benchmark results with respect to which the calculated solvatochromic shift in a given solvent can be measured. For the vacuum calculations, we again carry out a 20 ps temperature equilibration. However, due to the lack of solvent molecules we can skip the pressure equilibration step, so that the 20 ps temperature equilibration is directly followed by

	Average N_{mol}	Average N_{atoms}
n-hexane	55	1142
toluene	67	1047
acetone	98	1022
ethanol	122	1140

TABLE I: Table specifying average number of solvent molecules, as well as the average total number of atoms treated at the TDDFT level for nile red in all four solvent environments.

a 200 ps production run. Again, the last 35 of the 50 snapshots generated in the production run are taken as input for the TDDFT calculations.

We then use ONETEP to perform full TDDFT calculations on all 35 snapshots for each solvent to obtain fully converged averaged excitation spectra. In line with previous investigations carried out with ONETEP[11] and studies on nile red in water[7] we consider a large shell of solvent molecules explicitly in the TDDFT calculation. Here, we choose $R_{\mathrm{shell}} = 11\,\text{Å}$ for all systems (see Section III C), corresponding to roughly 1050 to 1150 atoms per snapshot (see Tab. I). In order to remove any issues regarding spurious charge-transfer states to the edge of the explicit solvent region in the polar solvents acetone and ethanol, we also select $R_{CT} = 9\,\text{Å}$ for these two systems. All ONETEP calculations are performed using the PBE functional[53] and norm-conserving pseudopotentials. A further discussion of the ONETEP calculation parameters can be found in the supplementary material (Section 4). Here we just mention that that the calculation parameters have been shown to yield fully converged TDDFT excitation energies[11, 22, 23] for low energy excitations of small organic dyes that are comparable in quality to those obtained from calculations using large, diffuse Gaussian-type basis sets[54] like aug-cc-pVTZ.

For the TDDFT calculations on the 35 snapshots of nile red in vacuum, we use the same ONETEP calculation parameters as in the solvated system. We furthermore make use of an implicit solvation model, with $\epsilon_0 = 1$. This has the advantage that the system is effectively treated in open boundary conditions[55] and the calculations are fully consistent with the treatment of the long range solvent environment in the solvated systems.

While the PBE semi-local exchange-correlation functional has become a popular choice for the explicit treatment of dyes in solution due to the low computational cost associated

with it, especially in the case where large parts of the solvent environment are considered explicitly[12, 14], it has a number of well-known shortcomings, primarily when considering excitations with charge-transfer character. It has been demonstrated previously, that some of the failures of semi-local functionals in connection to the computation of optical spectra in solution can be corrected for by comparing to a small number of calculations using a more sophisticated functional in a spectral warping procedure[12]. The goal of the spectral warping approach is to generate an approximation to the TDDFT spectrum in solution using a more expensive hybrid functional, by applying a linear transformation to the excitation energies of the solvated system generated with a computationally cheaper semi-local functional. The parameters of the linear transformation are derived by comparing the TDDFT excitation energy of the solute in vacuum as obtained by the semi-local functional of choice to that of a chosen hybrid functional. In this work, we apply a variant of the warping technique and therefore carry out a number of benchmark calculations on the nile red dye using a range-separated hybrid functional in the form of CAM-B3LYP[56]. All calculations using CAM-B3LYP are carried out with the NWChem code[57] and a 6-311++G** Gaussian basis set.

V. RESULTS

The averaged absorption spectra of nile red in vacuum and the four solvents, generated from the full 140 ps classical MD trajectory, and treating the full solvent environment at the TDDFT level, can be found in Fig. 3a. We note that our results predict an increasing redshift of the absorption maximum in going from hexane to ethanol, in accordance with experimental results. The absolute positions of the computed absorption maxima are all strongly redshifted compared to the experimental results (by more than 100 nm in the case of ethanol; see also section 7 and 8 of the supplementary material), which can be attributed to the semi-local PBE functional underestimating the S_1 excitation energy of nile red by a significant amount. However, while the experimental position of the absorption maxima can evidently not be reproduced by the PBE functional, the solvatochromic shift, being a relative measure of the energy difference between the peak absorption energy in vacuum and in solvent, is expected to be in much better agreement with experimental results.

The calculated solvatochromic shift for each solvent, as measured with respect to the vac-

uum absorption maximum, can be found in the third column of Tab. II (labelled "Explicit"). As can be seen from these results, the computed solvatochromic shifts for ethanol and acetone are in relatively good agreement with experimental results, although the absolute shift is underestimated by 20 meV for both solvents. This constitutes a major improvement over the results predicted by the semiclassical model of Eqn. 1 (see the first column of Tab. II). However, for toluene and *n*-hexane, the results are less convincing. The shift of toluene is underestimated by 60 meV, while the shift of hexane is overestimated by about 90 meV, leading to a splitting between the solvatochromic shifts of hexane and toluene that is much too small compared to experiment. Furthermore there is only a small improvement to the absolute solvatochromic shifts for hexane and toluene over those obtained through Eqn. 1.

The apparent failure of the TDDFT procedure outlined in this work to reproduce the correct solvatochromic shift for *n*-hexane can be linked to the fact that experimentally, nile red is observed to have a double peak structure in a number of non-polar solvents including *n*-hexane (see experimental spectra by Davis *et. al*[25], as well as Section 8 of the supplementary material), with the blue peak being higher in oscillator strength (thus constituting the absorption maximum). This double peak structure is not reproduced in the results obtained here, although we note that the solvatochromic shift of our peak is in very close agreement to the second, red-shifted peak predicted experimentally. In fact, it has been demonstrated that the origin of the double peak structure in nile red is due to vibronic excitations[28], meaning that this spectral feature is inherently non-reproducible with the calculation method we rely on in this work, as the Born-Oppenheimer approximation is applied throughout. However, even though the strong overestimation of the solvatochromic shift in *n*-hexane can be explained through the absence of the experimentally observed double peak structure, the splitting between the toluene and hexane solvatochromic shifts is still considerably underestimated, mainly due to the underestimation of the absolute solvatochromic shift for toluene.

Regarding the general shape of the spectra obtained, it can be noted that both *n*-hexane and toluene as solvent environments yield relatively narrow absorption peaks, which broaden considerably for acetone and ethanol, in agreement with experimental data. However, as mentioned earlier, experimentally *n*-hexane and toluene are predicted to produce an asymmetric absorption peak profile with a sharp onset in the red and a long tail in the blue. In our results, the opposite trend can be observed and all spectra tend to have considerable

	Eqn. 1	Implicit	Explicit	Warped	Exp.[26, 32]
ethanol	-0.303	-0.218	-0.303	-0.314	-0.321
acetone	-0.297	-0.178	-0.223	-0.242	-0.240
toluene	-0.123	-0.123	-0.143	-0.177	-0.205
n-hexane	-0.095	-0.111	-0.114	-0.114	-0.022 (-0.122)

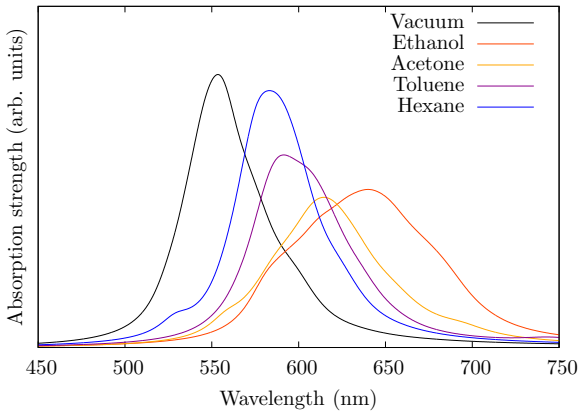
TABLE II: Calculated solvatochromic shift (in eV) of nile red in the four different solvents, as computed with the semiclassical model of Eqn.1 fitted to experiment, the implicit solvent model, the explicit solvent model, as well as the shift obtained from the warped explicit solvent model spectrum, in comparison to experimental shifts[26, 32]. Experimentally, nile red in n-hexane shows a double peak structure, with the position of the second peak relative to the vacuum level indicated in brackets. The absolute peak positions as computed in the implicit, explicit and warped explicit solvent model as compared with experimental values are reported in Section 7 of the supplementary material.

tails in the red wavelengths, which are not found experimentally.

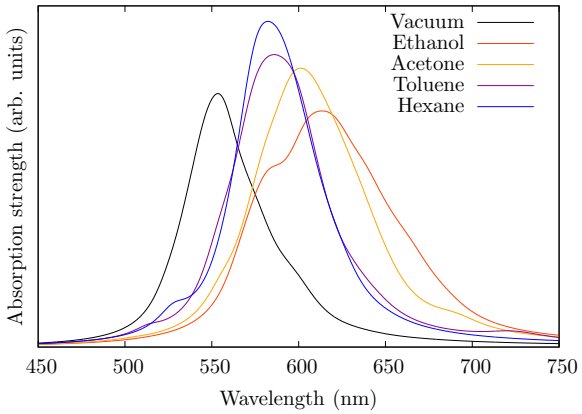
A. The role of the solvent environment

While there are some discrepancies between the obtained results for nile red in the four selected solvents, it is encouraging that the TDDFT method treating a large part of the solvent environment quantum mechanically produces sizeable splittings between ethanol and acetone, where the predicted results from the simple semi-classical model dependent on dielectric properties of the solvents alone (Eqn. 1) show almost no shift. Given that the semiclassical model ignores any explicit solvent-solute interactions it is evident that the predicted splittings in the TDDFT approach chosen here for solvents with very similar dielectric properties must be due to the explicit treatment of the solvent environment. In this section, we study the effect of the solvent environment in some more detail and discuss the origin of the strong difference in solvatochromic shifts.

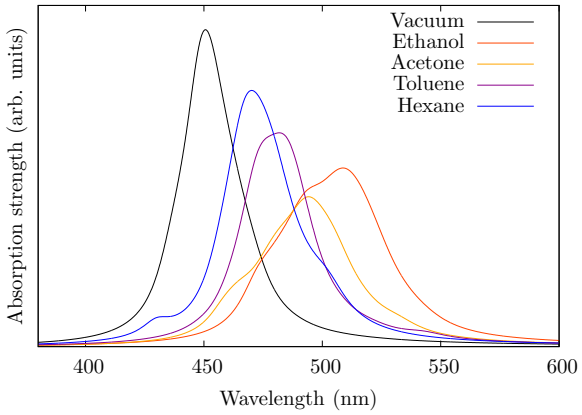
In order to determine the direct effect of the solvent environment on the observed absorption spectra, we take all 35 snapshots for all four solvents, strip away the solvent environment and replace it by an implicit solvent model in the LR formulation. We set ϵ_0 to the



a) Explicit solvent



b) Implicit solvent



c) Explicit warped

FIG. 3: Averaged absorption spectra for nile red in vacuum and the four selected solvents, in explicit solvent representation (Fig. 3a), when the solvent environment is replaced by an implicit solvent model (Fig. 3b), as well as in the explicit solvent representation after applying the spectral warping correction (Eqn. 5) to the TDDFT results (Fig. 3c).

experimental relative static permittivity and for the linear-response TDDFT calculations in implicit solvent we use $\epsilon_{\text{opt}} = n^2$ where n is the experimental refractive index. The averaged absorption spectra can be found in Fig. 3b. As can be seen, the implicit solvent results for toluene and n-hexane produce absorption spectra that are almost on top of each other. There is a small shift in going from toluene to acetone, but the shift between acetone and ethanol is much reduced and the two curves overlap substantially. The predicted solvatochromic shifts for all four solvents as obtained with an implicit solvent treatment can be found in the second column of Tab. II (labelled "Implicit"). In agreement with the spectral shapes, there is almost no splitting between toluene and hexane, the splitting between acetone and ethanol is severely underestimated and the absolute shift of ethanol is underestimated by 0.1 eV. It becomes evident that the strong difference in solvatochromic shifts between different solvent environments with similar macroscopic dielectric properties is indeed due to the presence of the explicit solvent environment directly influencing the excitation energies and not due to conformal changes of the nile red molecule itself in reaction to the surrounding solvent environment.

The most obvious origin of the discrepancy between the results predicted by both the semiclassical model and the implicit solvent model in comparison to those predicted by the fully explicit treatment of the solvent environment is the fact that direct solute-solvent interactions must necessarily be ignored in a fully macroscopic description of the solvent. In the case of toluene and ethanol, the solute-solvent interaction ignored in implicit solvent approaches is due to π-stacking between the toluene and the aromatic rings in nile red and hydrogen bonding between ethanol and the N and O atoms in nile red. Hexane and acetone, on the other hand, do not strongly interact with the nile red, causing the solvatochromic shifts predicted by the implicit solvent model for these two solvents to be somewhat closer to the case where the entire solvent environment is treated explicitly.

We have performed a in-depth analysis of the character of the solvent environment around nile red for ethanol and toluene, and further details can be found in the supplementary material (Section 6). Here we just mention that the first solvation shell in both solvents is indeed highly ordered. In the case of ethanol, 70 % of all snapshots studied show hydrogen bonding between ethanol and nile red, while more than 90 % of all toluene snapshots have at least one solvent molecule in a π-stacking conformation.

To determine the effect of the solute-solvent interaction on the excitation energy, we take

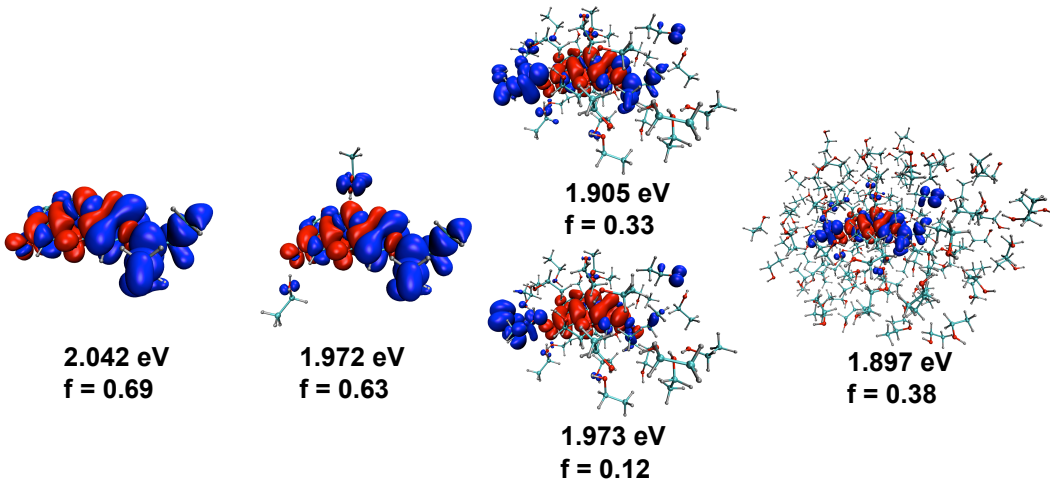


FIG. 4: Figure showing a plot of the electron hole density, as well as energies and oscillator strengths of the S_1 state of nile red in ethanol for a selected snapshot with increasing amount of explicit solvent environment. The first solvent representation is simply given by the implicit solvent, the second contains two ethanol molecules, the third has 26 ethanol molecules (corresponding to $R_{\text{shell}} = 5.5$ Å) and the last one, representing the full system used in previous results, consists of 121 ethanol molecules. In the third representation, the single excited state splits into two states of very similar character.

two specific snapshots, one for ethanol and one for toluene, that show strong evidence of solute-solvent interactions. We gradually increase the size of the explicit solvent environment and compute the excitation energies, while the entire system is surrounded by an implicit solvent model in each calculation. The results can be found in Fig. 4 and Fig. 5. As can be seen, the inclusion of the toluene molecules involved directly in π-stacking and the ethanol molecules involved in hydrogen bonding significantly lowers the excitation energy of the S_1 state compared to the pure implicit solvent model, and there is some delocalisation of fractions of the electron-hole density onto the solvent molecules. These delocalisations are due to polarisation effects, where the excited state dynamically screens its own transition dipole moment, yielding lower excitation energies, consistent with our previous study[11]. The results also show that it is not sufficient to only include the solvent molecules directly interacting with the solute in order to obtain converged excitation energies. Rather, it becomes necessary to fully include the first few solvation shells[7, 11, 33] in order to shift the energy levels associated with the solvent molecule to the right place and thus converging

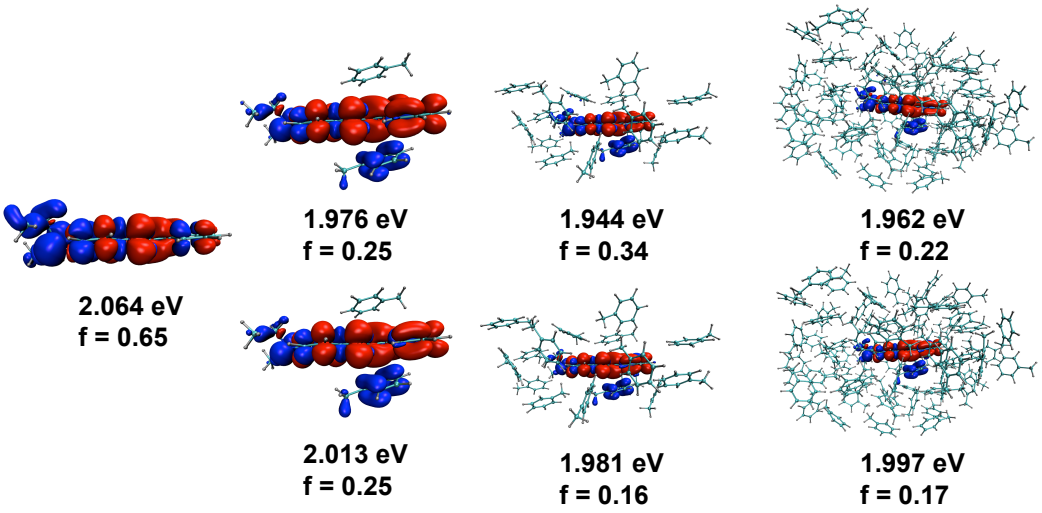


FIG. 5: Figure showing a plot of the electron hole density, as well as energies and oscillator strenghts of the S_1 state of nile red in toluene for a selected snapshot with increasing amount of explicit solvent environment. The first solvent representation is simply given by the implicit solvent, the second contains two toluene molecules, the third has 17 toluene molecules (corresponding to $R_{\text{shell}} = 6$ Å) and the last one, representing the full system used in previous results, consists of 69 toluene molecules. In the second, third and fourth representation, the single excited state splits into two states of very similar character.

the net effect of the dynamic polarisation of the TDDFT response density. For example, in the snapshot in ethanol considered above, including the two ethanol molecules directly involved in the hydrogen bonding in the TDDFT calculation only accounts for about half of the red-shift between the implicit solvent results and the fully converged TDDFT energies for the solvated system. The results for toluene show that the convergence of the excitation energy with respect to the solvent environment does not necessarily follow a simple trend and the inclusion of larger explicit solvent environments can yield a red-shift or a blue-shift of excitation energies. Note however, that the excited state in each case retains a fairly localised character and the delocalisation and polarisation effects are confined to a region close to the nile red molecule, again in agreement with previous results[11].

We point out that an inclusion of explicit solvent environment tends to lower the oscillator strengths of excitations compared with implicit solvent models. The reason for this can be seen in the different way the two approaches account for the dynamic polarisation effects in the solvent environment. In an explicit solvent model, the transition dipole moment

of the excitation self-screens through delocalising small fractions of the transition density to neighbouring solvent atoms. This lowers the excitation energy and dipole moment and thus the oscillator strength of the excitation of interest. In the implicit solvent approach, the dynamic screening of the transition density is modelled through the relative dynamic permittivity ϵ_{opt}, while the density must necessarily stay fully confined on the dye. The effect of the dynamic screening through ϵ_{opt} is to reduce the energy cost associated with the perturbation created by the excitation, thus allowing for a larger dipole moment to be sustainable on the dye while simultaneously lowering the excitation energy.

B. Colour prediction and spectral warping

As mentioned in previous sections, the strong solvatochromism of nile red causes solutions stained with the dye to strongly differ in colour depending on the solvent used. The experimental colours for nile red in ethanol, acetone, benzene and cyclohexane are found in Fig. 6 and range from magenta (ethanol) over red (acetone and benzene) to yellow (cyclohexane). While we do not have access to the experimental spectra for toluene and *n*-hexane, the colours for benzene and cyclohexane can be expected to be highly similar both judging from the solvent structure and the highly similar experimental solvatochromic shifts. Attempting to reproduce these trends forms a good test to determine whether the computational approach outlined in this work yields reasonable spectral shapes. However, before an attempt can be made to predict the colours of nile red in the four different solvents, two major deficiencies of our calculated results, already mentioned in previous sections, have to be addressed.

First of all, the PBE exchange-correlation functional significantly underestimates the S_1 excitation energy of nile red, both in comparison to experimental results and to more sophisticated range-separated hybrid functionals. Therefore, all theoretical spectra obtained in this work are strongly redshifted compared to experimental results and thus cannot yield correct colours. It has been pointed out by Ge and coworkers[12] that the deficiencies of the semi-local PBE functional compared to range-separated hybrids can be corrected by applying a *spectral warping* technique. In this technique, a linear transformation is applied to the peak positions of each snapshot of the fully solvated system computed with the PBE functional, that moves the positions of the peaks to their respective positions that would be

obtained with the range-separated hybrid functional. The linear transformation is defined by comparing the excitation energies of the solute in vacuum as computed with PBE and the range-separated hybrid functional of choice. The advantage of this approach is that the comparatively expensive hybrid functional calculations only have to be performed on the solute, while the computationally cheap PBE functional is used for the solvated system, allowing for a fully first-principles treatment of large solvent environments. The implicit assumption of the approach is that the PBE functional is capable of correctly describing the dynamic polarisation of the excited state in solution, an assumption that has been shown to hold for numerous small dyes studied[11, 12, 14].

The second deficiency of the calculated spectra that needs to be addressed is the long absorption tail in the red produced by most solvents that is not found in the experimental spectra and that is likely to influence the predicted colour. The origin of this spectral feature can be seen in the description of the excited state potential energy surface of the S_1 state with respect to rotations around the dihedral angle C1-C2-N1-C3 (see Fig. 1). Any rotation around the C-N bond that moves the diethylamino groups out of the molecular plane causes the S_1 excitation to have an increasing amount of charge transfer character. In contrast to higher order quantum chemistry methods, semi-local functionals like PBE predict the excitation energy of the S_1 state to decrease for more twisted conformations, while range-separated hybrid functionals like CAM-B3LYP correctly predict an increase[28, 29]. This causes the PBE functional to predict excited state energies that are significantly too low for more twisted conformations, producing a spurious tail in the red wavelengths for the computed spectra in this work. Given that the more polar solvents acetone and ethanol favour a more twisted conformation of nile red (see Section 3 in the supplementary material) the spectral broadening into the red wavelengths becomes stronger in the case of those two solvents. Similarly, while toluene and hexane both strongly favour relatively flat conformations of nile red, toluene has an increased likelihood of strongly twisted conformations as judged from the snapshots studied in this work (Section 3 of the supplementary material), yielding a more evident tail in the red.

However, when comparing a scan of the excited state potential energy surface for rotations around the dihedral angle C1-C2-N1-C3 in nile red as computed with CAM-B3LYP and PBE, it is found that the angle-dependent discrepancy between the two curves follows a simple quadratic form over almost the entire relevant twist angle range occurring in snapshots

studied in this work (see section 5 of the supplementary material). We therefore aim to correct both the spurious underestimation of excitation energies and the wrong description of the S_1 potential energy surface by the PBE functional by applying a transformation to the S_1 peak positions in every snapshot and every solvent. This transformation can be seen as a variant of the spectral warping approach introduced in other work[12, 14] and takes the form of

$$\omega_i^{\text{trans}} = \omega_i + \beta + \alpha\phi_i^2 \tag{5}$$

where ω_i^{trans} is the transformed excitation energy of snapshot i, ω_i is the original excitation energy of snapshot i and β is a constant shift that is independent of the snapshot. The angle ϕ_i is defined as the average of the dihedral angle C1-C2-N1-C3 and its equivalent angle for the other diethylamino group of nile red in a given snapshot i, where we define an angle of 0° as the flat conformation and an angle of 90° as the fully twisted conformation, and α is a snapshot-independent constant.

We stress that the parameters β and α in the above transformation are derived fully from *ab-initio* calculations: we set β equal to the energy difference between the S_1 state of nile red in its ground state structure in vacuum as computed with the PBE functional and with the CAM-B3LYP functional for the same structure, while α is derived from a fit to the angle-dependent difference of the S_1 potential energy surface between PBE and CAM-B3LYP (see Fig. 5 of the supplementary material). Thus only a small number of selected calculations on nile red in vacuum using the CAM-B3LYP functional have to be carried out in order to parameterise Eqn. 5, where we find $\beta = 0.487$ eV and $\alpha = 0.138$ eV$(\text{rad})^{-2}$. Once the two parameters β and α are determined, we use them to transform the spectra for nile red in vacuum and all four solvents and all snapshots as computed with PBE and the full explicit solvent environment. The results obtained from this transformation can then be seen as the best attempt to predict the spectra one would obtain for CAM-B3LYP in the full explicit solvent environment, without having to perform full hybrid functional calculations on prohibitively large system sizes.

A plot of the transformed absorption spectra for nile red in solution can be found in Fig. 3c. We note that while the positioning of the absorption maxima is slightly blue-shifted compared to experimental results (the vacuum absorption peak is predicted to be at 450 nm while experimentally the gas phase absorption peak is expected to occur at approximately 480 nm[32]; see also section 7 in the supplementary material), the agreement

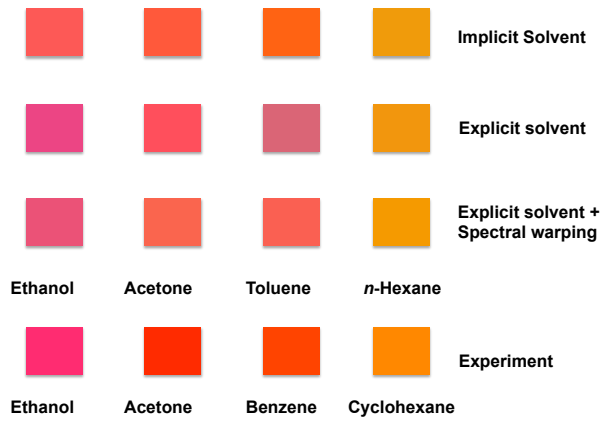


FIG. 6: Predicted colour of nile red in different solvents, for different treatment of the solvent environment, as well as the warped spectrum. The angle-independent constant shift derived in the spectral warping procedure is applied to all spectra, even the ones that were not explicitly warped. This is to ensure that the peaks of the spectra using different treatments of the solvent environment occur at roughly the same wavelength and are thus comparable. The experimental colours are obtained by using the experimental spectra[25] (see also section 8 of the supplementary material) as an input to the tristimulus colorimetry procedure.

is much better than for the unshifted PBE results. The spectral shape of all peaks is also considerably improved due to the dihedral angle-dependent transformation procedure. Nile red in toluene shows a relative sharp peak with a flat top, which compares very well with the experimental spectral shape for the closely related solvent benzene[25] (see also section 8 of the supplementary material). Both acetone and ethanol also produce considerably less broad absorption peaks in line with experimental results.

The general improvement in the absorption peak shape for all solvents also yields an improvement in the predicted solvatochromic shifts, which can be found in the fourth column of Tab. II (labelled "Warped"). While there is almost no change to the predicted solvatochromic shift for n-hexane, the predicted shift for toluene increases, yielding a larger splitting between the two in line with experimental results. Ignoring the absorption peak of n-hexane that is due to vibronic effects that cannot be reproduced with the methods used in this work, the largest discrepancy between predicted and experimental solvatochromic shifts is now less than 30 meV, which can be considered an excellent agreement.

Given the much improved spectra obtained through the transformation procedure fully

defined through *ab-initio* calculations, it is interesting to test whether this approach is capable of producing the correct colours for nile red in all four solvents. Furthermore, it is also worth determining the net effect of the explicit solvent environment and the angle-dependent transformation on the predicted colour. We therefore compute the colour of nile red as obtained using the absorption spectra calculated from the implicit solvent, the explicit solvent and the warped explicit solvent procedure. However, since the implicit and explicit solvent spectra that do not have the transformation of Eqn. 5 applied to them show absorption peaks that are strongly red-shifted due to the underestimation of the S_1 energy by the PBE functional, they produce colours that are not directly comparable to the warped spectrum. In order to isolate the effects of the explicit solvent environment and the angle-dependent transformation in the warping procedure on the predicted colours, we apply the constant shift β to the explicit and implicit solvent spectra as well. The predicted colours for the shifted implicit and explicit solvent spectra, as well as the fully warped and the experimental spectra can be found in Fig. 6.

As can be seen, the implicit solvent model fails to show a strong change in colour in going from toluene to ethanol, while the explicit solvent and the warped explicit solvent model produce a significant change from orange for *n*-hexane to magenta for ethanol. It becomes evident that the explicit inclusion of the solvent environment is the most important factor in yielding a strong difference in colour between each of the solvents. However, the warped spectrum does show some further improvement over the explicit solvent spectrum, as nile red in toluene is incorrectly predicted to have a violet colour in the explicit solvent spectrum and a more red colour in the warped spectrum. In general, the warped spectrum produces colours that are in very good agreement with experiment with the small exception that the colours are not quite as bright as the experimental ones. This failure can possibly be ascribed to the failure in reproducing the long absorption tail in the blue wavelengths in nile red most likely due to vibronic effects, as well as the fact that the CAM-B3LYP results are generally slightly blue-shifted compared to experimental results.

It can be summarised that a full treatment of large parts of the explicit solvent environment is essential in order to predict the solvatochromic shifts and colours of nile red in the four different solvents studied. The large explicit solvent boxes used in this work can be effectively treated using the linear-scaling code ONETEP and the PBE functional. While the PBE functional has a number of clear deficiencies when considering this system, most

notably the systematic underestimation of the S_1 excited state and the wrong behaviour of the excited state potential energy surface with respect to rotations around the C-N bond connecting the diethylamno groups with the aromatic ring system, these deficiencies can be largely corrected for by applying an angle-dependent transformation to the calculated S_1 excitation energies. The excellent agreement of both the solvatochromic shifts and the colours obtained from the explicit solvent spectra transformed according to Eqn. 5 can be seen as clear evidence that the PBE functional correctly accounts for the dynamic polarisation effects of the solvent environment reacting the the perturbation caused by the excitation, in line with previous studies[11, 12, 14].

VI. CONCLUSION

In conclusion, we have successfully applied a computational approach relying on the explicit representation of large parts of the solvent environment to nile red in four common non-polar and polar solvents. We have demonstrated that the explicit representation of the solvent environment is crucial in creating the correct discrepancies in shifts between non-polar solvents *n*-hexane and toluene, as well as polar solvents acetone and ethanol. The origin of the failure of the implicit solvent approach in the case studied here can be seen in the fact that toluene and ethanol show strong solute-solvent interactions through π-stacking and hydrogen bonding, while *n*-hexane and acetone are only interacting weakly with nile red. We furthermore demonstrated that while the PBE semi-local functional used throughout in this work for reasons of computational efficiency systematically underestimates excitation energies and cannot describe the correct excited state energy surface of the S_1 state of interest, these failures can be relatively easily corrected for by comparing to a small number of calculations of nile red in vacuum using the CAM-B3LYP functional. The resulting transformed spectrum yields both solvatochromic shifts and colours in excellent agreement with experimental results.

We furthermore find nile red to be an excellent test system in general for computational approaches aimed at predicting absorption spectra of solvated systems, due to the large number of different effects that need to be correctly accounted for in order to produce correct answers. These effects include: the ground state potential energy surface under rotations around the C-N bond; solvent polarity-dependent changes in the relative likelihood of the

twisted and the flat conformation; the large variety of possible solute-solvent interactions like π-stacking and hydrogen bonding and their influence on both the electronic ground state and the excited states of nile red, as well as considerations regarding the required system size to fully converge polarisation effects; the correct description of the change in character of the excited state S_1 when changing from a flat to a twisted conformation; and finally the influence of vibronic excitations, creating the double peak feature in non-polar solvents as well as contributing to the long absorption tail in the blue wavelengths.

While our results show that a computationally much cheaper implicit solvent treatment is insufficient to resolve the differences between the solvents studied in this work, we note that this failure in the specific case studied here is not surprising. We deliberately selected solvents that show a large difference in solvatochromic shifts experimentally while having very similar macroscopic dielectric properties. Given that the implicit solvent model used throughout this work represents solvents through the macroscopic static dielectric constant and refractive index only, it is clear that strong shifts between, for example, *n*-hexane and toluene cannot be reproduced correctly as their dielectric properties are too similar. For the same reason we would expect solvatochromic shifts that are already well reproduced by the semiclassical model introduced in Section II to be correctly predicted by state-of-the-art implicit solvent approaches.

The explicit solvent treatment that forms the centre of this work is able to correctly predict solvatochromic shifts in the cases where the implicit solvent treatment is shown to be lacking. The large-scale TDDFT code used in this work allows for the highly efficient calculations of large solvent environments, thus yielding fully converged solvent polarisation effects. However, the correct prediction of shifts and colours following the computational approach outlined here relies on a few manual adjustments that are potentially undesirable in high-throughput, routine prediction of optical spectra of dyes in solution. One issue, the need to carefully check and potentially reparameterise the classical force field used for the dye in question against *ab initio* methods, could be alleviated by using a QM/MM approach to generate the solute-solvent configurations, where the solute itself could be treated fully quantum mechanically. The second one, due to the potential shortcomings of the semi-local exchange-correlation functional in describing the excited state potential energy surfaces of the isolated dye, is less straightforwardly addressed. In many cases where the excited state does not change character under typical geometry changes in solution, it is expected that

a simple rigid shift in excitation energy derived from the difference between the TDDFT energy with semi-local and hybrid functionals in vacuum is sufficient to correct for the main deficiencies of PBE and yield robust solvatochromic shifts, peak positions and colours. In the case of nile red the excited state changes character under rotation of the C-N bond but the failures of the semi-local functional in describing this change are easily corrected for. In cases where the excited state potential energy surface undergoes substantial changes under solvent-driven geometry changes that are not easily described by semi-local functionals or corrected for in a simple manner, the approach used in this work might encounter some difficulties. However, for a large class of small dyes, we expect the computational approach discussed in this work to yield very robust predictions of peak positions, solvatochromic shifts and colours.

VII. SUPPLEMENTARY MATERIAL

Details on the reparameterisation of the classical force field parameters for nile red, ONETEP calculation parameters, the analysis of hydrogen bonding and π-stacking interactions between the solvent environment and the nile red, as well as calculations of the S_1 excited state potential energy surface of nile red under rotation around the C-N bond can be found in the supplementary material.

Acknowledgments

TJZ acknowledges the support of EPSRC Grant EP/J017639/1. MCP and PDH acknowledge the support of EPSRC grant EP/J015059/1. The underlying data of this publication can be accessed via the following persistent URI: http://dx.doi.org/10.17863/CAM.5960

[1] A. Warshel, and M. Levitt, J. Mol. Biol. **103**, 227-249 (1976).

[2] H. Lin, and D. G. Truhlar, Theor. Chem. Acc. **117** (2), 185199 (2007).

[3] H. Houjou, Y. Inoue, and M. Sakuray, J. Phys. Chem. B **105** (4), 867-879 (2001).

[4] A. Altun, S. Yokoyama, and K. Morokuma, J. Phys. Chem. B **112** (22), 6814-6827 (2008).

[5] J. Gao, W.-J. Shi, J. Ye, X. Wang, H. Hirao, and Y. Zhao, J. Phys. Chem. B, **117** (13), 3488-3495 (2013).

[6] N. A. Murugan, P. C. Jha, Z. Rinkevicius, K. Ruud, and H. Ågren, J. Chem. Phys. **132**, 234508 (2010).

[7] N. A. Murugan, Z. Rinkevicius, Z. and H. Ågren, Int. J. Quantum Chem. **111**, 1521-1530 (2010).

[8] C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker, and T. J. Martínez, J. Chem. Theory Comput. **8**, 5092-5106 (2012).

[9] C. M. Isborn, B. D. Mar, B. F. E. Curchod, I. Tavernelli, and T. J. Martínez, J. Phys. Chem. B **117**, 12189-12201 (2013).

[10] M. R. Provorse, T. Peev, C. Xiong, and C. M. Isborn, J. Phys. Chem. B **120** (47), 12148-12159 (2016).

[11] T. J. Zuehlsdorff, P. D. Haynes, F. Hanke, M. C. Payne, and N. D. M. Hine, J. Chem. Theory Comput. **12** (4), 1853-1861 (2016).

[12] X. Ge, I. Timrov, S. Binnie, A. Biancardi, A. Calzolari, and S. Baroni, J. Phys. Chem. A **119** (16), 3816-3822 (2015).

[13] T. Wada, H. Nakano, and H. Sato, J. Chem. Theory Comput. **10** (10), 4535-4547 (2014).

[14] O. B. Malcioğlu, A. Calzolari, R. Gebauer, D. Varsano, and S. Baroni, J. Am. Chem. Soc. **133**, 1425-15433 (2011).

[15] D. H. Douma, B. M'Passi-Mabiala, and R. Gebauer, J. Chem. Phys. **137**, 154314 (2012).

[16] E. Runge, and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[17] M. Petersilka, U. J. Grossmann, and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).

[18] M. E. Casida, J. Mol. Struc.-THEOCHEM **914**, 3-18 (2009).

[19] N. T. Maitra, J. Chem. Phys. **144**, 220901 (2016).

[20] C. Yam, S. Yokojima, and G. Chen, Phys. Rev. B, **68**, 153105 (2003).

[21] S. Tretiak, C. M. Isborn, A. M. Niklasson, and M. Challacombe, J. Chem. Phys. **130**, 054111 (2009).

[22] T. J. Zuehlsdorff, N. D. M. Hine, J. S. Spencer, N. M. Harrison, D. J. Riley, and P. D. Haynes, J. Chem. Phys. **139**, 064104 (2013).

[23] T. J. Zuehlsdorff, N. D. M. Hine, M. C. Payne, and P. D. Haynes, J. Chem. Phys. **143**, 204107 (2015).

[24] W. Humphrey, A. Dalke, and D. Schulten, J. Molec. Graphics **14**, 33-38 (1996).

[25] M. M. Davis, and H. B. Helzer, Anal. Chem. **38** (3), 451-461 (1966).

[26] A. K. Dutta, K. Kamada, and K. Ohta, J. Photochem. Photobiol. A: Chem. **93** (1), 57-64 (1996).

[27] A. K. Dutta, K. Kamada, and K. Ohta, Chem. Phys. Lett. **258**, 369-375 (1996).

[28] A. C. Guido, B. Mennucci, D. Jacquemin, and C. Adamo, Phys. Chem. Chem. Phys. **12**, 8016-8023 (2010).

[29] A. Y. Freidzon, A. A. Safonov, A. A. Bagaturyants, and M. V. Alfimov, Int. J. Quantum Chem. **112**, 3059-3067 (2012).

[30] O. Kolling, and J. Goodnight, Anal. Chem. **45** (1), 160-164 (1973).

[31] P. Greenspan, E. P. Mayer, and S. D. Fowler, J. Cell Biol. **100** (3), 965-973 (1985).

[32] J. F. Deye, and T. A. Berger, Anal. Chem. **62**, 615-622 (1990).

[33] L. C. Dias Jr., R. Custodio, and F. B. T. Pessine, Int. J. Quantum Chem. **106**, 2624 (2006).

[34] W.-G. Han, T. Lui, F. Himo, A. Toutchkine, D. Bashford, K. M. Hahn, and L. Noodleman, Chem. Phys. Chem. **4** (10), 1084-1094 (2003).

[35] E. A. Nikitina, A. V. Odinokov, F. V. Grigoriev, M. V. Bsilevsky, A. A. Khlebunov, V. A. Sazhnikov, and M. V. Alfimov, J. Phys. Chem. B, **111** (15), 3953-3959 (2007).

[36] G. Singh, A. C. Chamberlin, H. R. Zhekova, S. Y. Noskov, and D. P. Tieleman, J. Chem. Theory Comput. **12**, 364-371 (2016).

[37] E. G. McRae, J. Phys. Chem. **61** (5), 562-572 (1957).

[38] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, J. Chem. Phys. **122**, 084119 (2005).

[39] P. D. Haynes, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, Psi-k Newsletter **72**, 78-91 (2005).

[40] R. Cammi, and B. Mennucci, J. Chem. Phys. **110**, 9877-9886 (1999).

[41] M. Caricato, B. Mennucci, J. Tomasi, F. Ingrosso, R. Cammi, S. Corni, and G. Scalmani, J. Chem. Phys. **124**, 124520-124513 (2006).

[42] A. V. Marenich, C. J. Cramer, D. G. Truhlar, C. A. Guido, B. Mennucci, G. Scalmani, and M. J. Frisch, Chem. Sci. **2**, 2143-2161 (2011).

[43] L. Cupellini, C. Amovilli, and B. Mennucci, J. Phys. Chem. B **119** (29), 8984-8991 (20015).

[44] D. Rocca, R. Gebauer, Y. Saad, and S. Baroni, J. Chem. Phys. **128**, 154105 (2008).

[45] P. Cysewski, T. Jelińsky, M. Przybyłek, and A. Shyichuk, New J. Chem. **36**, 1836-1843 (2012).

[46] A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, M. C. Payne, Comput. Phys. Commun. **147**, 788-802 (2002).

[47] L. E. Ratcliff, N. D. M. Hine, and P. D. Haynes, Phys. Rev. B. **84**, 165131 (2011).

[48] J.-H. Li, T. J. Zuehlsdorff, M. C. Payne, and N. D. M. Hine, Phys. Chem. Chem. Phys. **17**, 12065-12079 (2015).

[49] G. Lever, D. J. Cole, N. D. M. Hine, P. D. Haynes, and M. C. Payne, J. Phys. Condens. Matter **25**, 152101 (2013).

[50] J. Dziedzic, H. H. Helal, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, Europhys. Lett. **95**, 43001 (2011).

[51] T. J. Zuehlsdorff, *Computing the Optical Properties of Large Systems*; Publisher: Springer International Publishing, Switzerland, 2015; 97-165.

[52] D. A. Case, T. J. Berryman, R. M. Betz, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York, and P. A. Kollman, AMBER 2015, University of California, San Francisco (2015).

[53] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[54] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

[55] N. D. M. Hine, J. Dziedzic, P. D. Haynes, and C.-K. Skylaris, J. Chem. Phys. **135**, 204103 (2011).

[56] T. Yanai, D. P. Tew, and N. C. Handy, Chem. Phys. Lett. **393**, 51-57 (2004).

[57] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, Comput. Phys. Commun. **181**, 1477 (2010).