



Cite this: *J. Mater. Chem. C*,  
2024, 12, 18886

## Calibration of several first excited state properties for organic molecules through systematic comparison of TDDFT with experimental spectra†

Xia Wu, \*<sup>a</sup> Xiaoyu Xie \*<sup>b</sup> and Alessandro Troisi \*<sup>a</sup>

Time-dependent density functional theory (TDDFT) is a powerful computational tool for investigating excitation properties in organic electronics, and it holds significant potential for high-throughput virtual screening (HTVS) in this field. While most benchmarks focus on excitation energies, less attention has been paid to evaluating the accuracy of computed oscillator strengths and exciton reorganization energies against experimental data. In this work, we provide a systematic approach to evaluate in parallel the accuracy of these three quantities on the basis of a suitable fitting of the experimental absorption spectra of 71 molecules in solution. After considering 18 computational methodologies, the results from the M06-2X/def2-TZVP/PCM method demonstrate the strongest correlation with experimental data across the desired properties. For HTVS, the M06-2X/6-31G(d)/PCM method appears to be a particularly convenient choice among all methodologies due to its balance of computational efficiency and accuracy. Our results provide an additional benchmark needed before employing TDDFT methods for the discovery and design of organic electronic molecules.

Received 16th August 2024,  
Accepted 11th October 2024

DOI: 10.1039/d4tc03511a

[rsc.li/materials-c](http://rsc.li/materials-c)

## Introduction

Time-dependent density functional theory (TDDFT) has gained significant traction in high-throughput virtual screening, enabling researchers to evaluate and predict the properties of molecules efficiently.<sup>1–7</sup> Numerous studies have delved into the performance of TDDFT in calculating excitation energies, which is a crucial aspect of understanding the optical and electronic characteristics of materials.<sup>8–19</sup> Naturally, the accuracy of these calculations needs to be thoroughly evaluated by comparing the computed excitation energies against experimental data.<sup>20–23</sup> For example, Jacquemin and coworkers investigated the performance of various density functional theory methods in calculating the excitation energies of over 100 organic dyes from major chromophore classes.<sup>24</sup> With an integrated design process involving theoretical calculations, machine learning, and experimental validation, highly efficient organic light-emitting diode molecules can be discovered from 1.6 million molecules with external quantum efficiencies up to 22%.<sup>2</sup> In another recent example, a dataset of 48 182 organic

semiconductors containing benchmarks of organic molecules was presented, providing relevant electronic properties and demonstrating its potential for repurposing known molecules.<sup>25</sup> Thus, TDDFT provides a computationally efficient approach to modelling excited-state energies, facilitating the exploration of potential candidates for applications such as organic light-emitting diodes,<sup>26,27</sup> photovoltaics,<sup>28,29</sup> and photocatalysis.<sup>30,31</sup>

Nevertheless, for organic electronics, achieving optimal performance goes beyond merely aligning energy levels. In the absence of competing photophysical processes, the oscillator strength ( $f$ ) determines the rate of light absorption and is strongly correlated with the efficiency of light emission, both of which are critical aspects in the performance of optoelectronic devices.<sup>32</sup> Additionally, the exciton reorganization energy ( $\lambda$ ), a measure of the structural changes accompanying the formation or dissociation of excitons, significantly influences phenomena such as exciton transport, emission efficiency and narrowness of absorption/emission peaks.<sup>33–35</sup> Tailoring these properties through molecular design or external manipulation becomes imperative for applications such as organic light-emitting diodes, organic photovoltaics, and organic field-effect transistors, where efficient light-matter interactions and charge carrier mobilities are paramount.

Over the decade, numerous studies have benchmarked oscillator strengths using various functionals and basis sets.<sup>36–39</sup> Recently, a few studies appeared that explore the accuracy of the computed oscillator strength of organic

<sup>a</sup> Department of Chemistry, University of Liverpool, Liverpool, L69 3BX, UK.  
E-mail: [xiawu@liverpool.ac.uk](mailto:xiawu@liverpool.ac.uk), [A.Troisi@liverpool.ac.uk](mailto:A.Troisi@liverpool.ac.uk)

<sup>b</sup> Qingdao Institute for Theoretical and Computational Sciences, School of Chemistry and Chemical Engineering, Shandong University, Qingdao, Shandong 266237, China. E-mail: [xiaoyuxie@sdu.edu.cn](mailto:xiaoyuxie@sdu.edu.cn)

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4tc03511a>



molecules in comparison to experimental data. In Cole's work, the oscillator strength is discussed as a computational property analogous to the experimental extinction coefficient, and they also analyzed the agreement between sTDA/TDDFT calculated  $f$  values and experimental intensity values.<sup>40</sup> Gozem's team fitted and integrated 100 experimental UV-visible spectra of solvated organic molecules to obtain experimental oscillator strengths ( $f_{\text{exp.}}$ ), compared them to computed oscillator strengths ( $f_{\text{comp.}}$ ) from TDDFT, finding that while  $f_{\text{comp.}}$  overestimates  $f_{\text{exp.}}$ , there is good overall correlation between the experimental and computed values.<sup>41</sup> Alipour and coworkers explored the use of the optimally tuned range-separated hybrid density functionals combined with a polarizable continuum model (OT-RSHs-PCM) and their screened versions (OT-SRSHs-PCM) to reliably predict oscillator strengths for organic compounds.<sup>42</sup> To the best of our knowledge, no prior studies have investigated the impact of  $\lambda$  with experimental data in the context of high-throughput virtual screening (HTVS), despite its crucial importance.

It is important to note that highly accurate methodologies have been developed over the years to reproduce the finest details of the vibronically-resolved spectra, including the effect of solvent,<sup>43–45</sup> Duschinsky effect,<sup>46,47</sup> and some of these results in publicly available software like FCClasses and ezFCF.<sup>48,49</sup> Santoro and coworkers presented a mixed quantum-classical approach for computing vibronic absorption spectra of molecular aggregates, reproducing changes in absorption spectrum upon aggregation.<sup>50</sup> Thiel's team analysed vertical and adiabatic Franck–Condon approaches using time-independent or time-dependent methods to simulate absorption spectra of three flavin compounds, finding that VFC with IMDHO-FA provides superior accuracy in spectral predictions.<sup>51</sup> These studies, while suitable for a few molecules at a time, often require selecting the most appropriate quantum chemistry methodology. Consequently, they are not particularly suitable for comprehensive comparisons across a large set of molecules. Also, it is common for the work focusing on vibronic coupling to ignore energy level accuracy and oscillator strength.<sup>52–54</sup> (there are also a large number of works comparing a range of quantum chemical methods among themselves and with experimental vertical excitation energy).<sup>20–25</sup> These can be considered complementary to the work presented here, where we instead focus on a restricted number of methods popular in HTVS.

For TDDFT to be an effective screening protocol, it is crucial to evaluate the accuracy of all three components (excitation energy,  $f$ , and  $\lambda$ ) *simultaneously*. This comprehensive assessment becomes particularly important when dealing with a large number of compounds, as one typically aims to employ a single efficient method for the screening process. HTVS involves in principle millions of molecules, making it impractical to select the optimal method for each dye individually. Moreover, when the exciton–phonon coupling is very strong (large  $\lambda$ ), it becomes challenging to accurately extract the oscillator strength  $f$  without fitting the vibronic side peaks, which are affected by  $\lambda$ . In other words, it is not only useful but also more accurate to

assess  $f$  and  $\lambda$  at the same time. In this work, we analysed 71 experimental absorption spectra and developed a fitting procedure to extract the excitation energy of the lowest excited state  $S_1$  ( $E_{S_1}$ ), the corresponding oscillator strength ( $f_{S_1}$ ) and  $\lambda$ . Subsequently, we compared the parameters derived experimentally and those computed with 18 different calculation levels to determine the most convenient approach and its expected accuracy.

## Methods

Usually, there are vibronic peaks for the  $S_1$  absorption spectrum. The intensity of these peaks can be evaluated as

$$I_{0-m} = I \prod_i |\langle S_0; 0_i | S_1; m_i \rangle|^2, \quad (1)$$

here,  $m = (m_1, m_2, \dots, m_N)$  contains occupation numbers of  $N$  modes,  $|S_k; m_i\rangle$  is a vibronic state for mode  $i$  at  $S_k$  potential energy surface state with  $m_i$  being the occupation number and  $I$  is the total intensity of  $S_1$  states, which is determined by its oscillator strength.

Under the harmonic approximation and Condon approximation, overlap in eqn (1) can be expressed as:

$$|\langle S_0; 0_i | S_1; m_i \rangle|^2 = e^{-s_i} \frac{s_i^{m_i}}{m_i!}, \quad (2)$$

here,  $s_i$  is the Huang–Rhys (HR) factor of normal mode  $i$ , measuring the displacement of the excited state along that particular mode.<sup>55</sup> Moreover, the product in eqn (1) can be further approximated as one overlap using one effective mode with HR factor  $s$  and vibrational frequency  $\omega$ .

Then, the total spectrum can be decomposed as a set of normal distributions with centres  $e_{0-m}$  being,

$$e_{0-m} = e_0 + m\omega, \quad (3)$$

and intensities  $I_{0-m}$  being,

$$I_{0-m} = I \times e^{-s} \frac{s^m}{m!}, \quad (4)$$

where  $e_0$  is the 0–0 transition energy (or adiabatic excitation energy) between  $S_1$  and  $S_0$ , and  $m$  is the vibrational quantum number of the effective mode. For the broadening  $\sigma_{0-m}$  of normal distribution, two additional parameters are included considering the broadening of the pure electronic part  $\sigma_0$  and vibrational part  $\Delta\sigma$ ,

$$\sigma_{0-m} = \sigma_0 + m\Delta\sigma, \quad (5)$$

Therefore, there are six parameters for the spectrum fitting: adiabatic excitation energy ( $e_0$ ), frequency of the effective mode ( $\omega$ ), HR factor of the effective mode ( $s$ ), the total intensity of  $S_1$  absorption ( $I$ ) and two broadening parameters  $\sigma_0$  and  $\Delta\sigma$



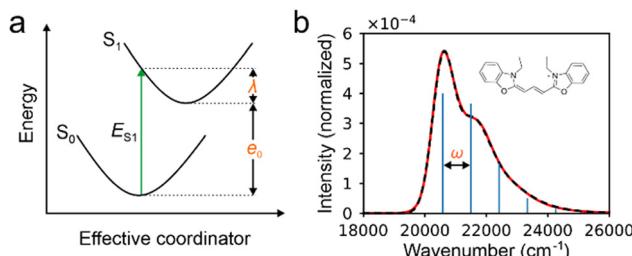


Fig. 1 (a) Illustration of the relationship among the energy of the first excited state ( $E_{S_1}$ ), the reorganization energy ( $\lambda$ ) and adiabatic excitation energy ( $e_0$ ) in the one mode model used to fit the experimental spectra of absorption only. (b) The experiment (the red line) and fitted absorption spectra (the dashed black line) for the molecule in the inset as example, resulting  $e_0 = 20573.63\text{ cm}^{-1}$ ,  $\omega = 920.38$ ,  $s = 0.91$ ,  $\sigma_0 = 345.574\text{ cm}^{-1}$ ,  $\Delta\sigma = 207.24\text{ cm}^{-1}$ . The blue bars represent positions and intensities of vibrational peaks before broadening.

(see Fig. 1 for the illustration of the fitting), and the total fitting absorption spectrum is given by,

$$\begin{aligned} \varepsilon_{\text{fit}}(e) &= \sum_m I_{0-m} \cdot \frac{1}{\sqrt{2\pi\sigma_{0-m}^2}} e^{-\frac{(e-e_{0-m})^2}{2\sigma_{0-m}^2}} \\ &= I \sum_m e^{-\frac{S^m}{m!}} \cdot \frac{1}{\sqrt{2\pi(\sigma_0 + m\Delta\sigma)^2}} e^{-\frac{(e-e_0-m\omega)^2}{2(\sigma_0 + m\Delta\sigma)^2}} \end{aligned} \quad (6)$$

In practice, the sum over occupation number is truncated to 4 ( $m < 5$ ), considering that HR factors of organic molecules are usually small.

After the fitting, oscillator strength can be calculated using the total intensity  $I$  obtained by fitting,<sup>32</sup>

$$f_0 = 10^3 \ln 10 \frac{m_e c^2}{N_A \pi \epsilon_0^2} \int \varepsilon_{\text{fit}}(e) de = I \cdot 4.32 \times 10^{-9} \text{ cm}^{-1}, \quad (7)$$

Here,  $m_e$  and  $e$  is the mass and charge of electron, respectively;  $c$  is speed of light in vacuum and  $N_A$  is the Avogadro's number.

However, since the experiments were conducted in different solvents, we incorporated the refractive index of the solvent ( $n$ ) when fitting  $f$  by employing an alternative equation for more accurate calibration results, as suggested by the, ref. 41,42,56 and 57

$$f_{S_1} = nf_0. \quad (8)$$

In general, the adiabatic energy differs from 0–0 energy depending on whether the zero-point energies (ZPE) in the S<sub>0</sub> and S<sub>1</sub> states are considered. However, our model (displaced harmonic oscillator) assumes the same curvature of the potential energy surfaces in the ground and excited states, resulting in identical ZPE. Consequently, within our model, the adiabatic energy and 0–0 energy are the same. The actual difference between ZPE was verified for few selected molecules reported in Table S4 of the ESI.<sup>58</sup>

In this work, the reorganization energy is defined as the difference between the energy of the vertical excitation from the equilibrium energy of S<sub>0</sub> state and the lowest energy of the

potential energy of S<sub>1</sub> state. Thus, the exciton reorganization energy can be calculated by

$$\lambda = s\omega. \quad (9)$$

If we could fit a more complex vibronic model from the experiment, we could extract a more accurate experimental reorganization energy. An example of benchmark work considering of this type (not focused on organic electronics) was given in ref. 51. However, the low-resolution spectra we have prevent us from deriving such a model.

Then, the vertical excitation energy of S<sub>1</sub> state can be evaluated as

$$E_{S_1} = e_0 + \lambda. \quad (10)$$

The error of the fitting was computed as:

$$\text{error} = \sqrt{\int \Delta\varepsilon(e)^2 de}. \quad (11)$$

The  $\Delta\varepsilon(e)$  represents the difference between experimental and fitting absorption spectrum (normalized).

The code for the fitting processes is now shared in GitHub by [https://github.com/XiaoyuUoL/spectrum\\_fitting](https://github.com/XiaoyuUoL/spectrum_fitting).

The UV/Vis<sup>+</sup> photochemistry database of dyes is used in this work to investigate the performance of our fitting.<sup>59</sup> All experimental data in the database are sourced from a single laboratory, with detailed experimental conditions to ensure consistency and avoid any potential instrumental errors. 103 dyes from the database with isolated S<sub>1</sub> absorption peak or mirrored absorption/emission signal are selected for simultaneously evaluating of  $e_0$ ,  $f_{S_1}$  and  $\lambda$ . Besides, the spectra of the database were recorded in different solvents including ethylene glycol ( $n = 1.43$ ), ethanol ( $n = 1.36$ ), cyclohexane ( $n = 1.43$ ), methanol ( $n = 1.33$ ) and dimethyl sulfoxide ( $n = 1.48$ ). The corresponding  $n$  was used for the fitting of  $f_{S_1}$ , and the specific solvent was employed in the calculations including solvent effects. Since the original dataset focuses on dyes with relatively high oscillator strength, to improve its representativeness, several organic systems (named 061, 068, 077, 085, and 099) with low UV/Vis intensity are also collected from ref. 41. We use the names as given in the original database with full chemical detail provided in the ESI.<sup>†</sup>

To achieve more accurate fitting results, we select 71 molecules for the final dataset, each with an error below 0.03 cm<sup>0.5</sup>. This selection includes 66 molecules from the original 103 dyes and 5 additional molecules with low UV/Vis intensity (the molecular structures are shown in Fig. S1, ESI<sup>†</sup>). All results of fitting are presented in Table S1 (ESI<sup>†</sup>), Fig. 1b and Fig. S13–S82 (ESI<sup>†</sup>). The molecules with poor fitted results are almost invariably affected by the presence of a nearby excited state. Therefore, we believe there is no significant bias in excluding these molecules. Moreover, we verified that the S<sub>2</sub> state in 71 molecules is mostly “dark” with much lower  $f$  values than the S<sub>1</sub> state, and the energy gap between the S<sub>1</sub> state and the next “bright” state exceeds 1 eV for most dyes, resulting no overlap between states for the dyes in our study.



All DFT and TDDFT calculations were carried out using the Gaussian 16A suite of programs.<sup>60</sup> Geometry optimises 10 conformers using PM7 semi-empirical method. For the lowest energy PM7 conformer, geometry optimizations were carried out at the BLYP35/3-21G(d) level.<sup>4</sup> Corresponding TDDFT calculations were utilizing the B3LYP, M06-2X, and  $\omega$ B97XD functionals in conjunction with the 6-31G(d), def2-SVP, and def2-TZVP basis sets, total 18 methods. Solvation effects were all taken into account based on the polarizable continuum model (PCM), using the solvent as the same as that employed in corresponding experiments (see Table S1, ESI<sup>†</sup>).<sup>61</sup> Gauge-invariance corrections are not included in our calculations. Due to stability issues, we didn't consider diffuse functions which are not suitable for HTVS, where the goal is to achieve a high yield of converged calculations without human intervention. The molecular excitation properties were also investigated by the hole-electron analysis using Multiwfns 3.6.<sup>62</sup> For the calculation of  $\lambda$ , we obtained results for 70 dyes, excluding the dye named 'bfl' (details in Fig. S1, ESI<sup>†</sup>) due to unsuccessful convergence in the  $S_1$  excited state.

## Results and discussion

To begin, we computed  $E_{S_1}$  of 71 dyes in the dataset along with  $f_{S_1}$  using a total of 18 calculation levels (Table S2 and Fig. S2–S7, ESI<sup>†</sup>). The best results are obtained with the functional M06-2X in conjunction with the larger basis set (def2-TZVP), accounting for solvent effects. The calculated values for both  $E_{S_1}$  and  $f_{S_1}$  demonstrated a remarkable consistency with the experimental fitted data, yielding the coefficient of determination ( $R^2$ ) values of 0.948 and 0.844, respectively (Fig. 2a and b). Moreover, their slope of the linear fitted by M06-2X/def2-TZVP/PCM was both close to 1 (0.967 and 1.048, separately), indicating this method exhibits excellent accuracy in predicting  $E_{S_1}$  and  $f_{S_1}$ . Computed  $f_{S_1}$  also exhibits a good correlation with experimental fitted  $f_0$  ( $R^2 = 0.86$ ), but the slope of the correlation is approximately 1.5 (Fig. S8, ESI<sup>†</sup>). This discrepancy prompted us to use  $n f_{\text{exp.}}$  instead of  $f_{\text{exp.}}$  to mitigate the inaccuracy as proposed in ref. 33,34 and 36 where  $n$  is the refractive index of the solution.

To illustrate the relative significance of functionals, basis sets, and solvent effects, we compared the results of 18 methods in Fig. 3. The choice of functional is crucial, as M06-2X and  $\omega$ B97XD functionals yielded similar and superior performance compared to B3LYP, with the  $R^2$  exceeding 0.9 for both functionals. Basis set selection influences the results moderately, but its impact is less important. A more cost-effective calculation with the 6-31G(d) basis set can be recommended. Notably, it is noteworthy that the solvent effects are essential for accurately predicting both the  $E_{S_1}$  and the corresponding  $f_{S_1}$ . The calculated values obtained using the PCM model exhibit noticeably better agreement with experimental data compared to gas-phase calculations employing the same functional. However, the impact of solvent effects is more pronounced on the calculations of  $E_{S_1}$  compared to  $f_{S_1}$ .

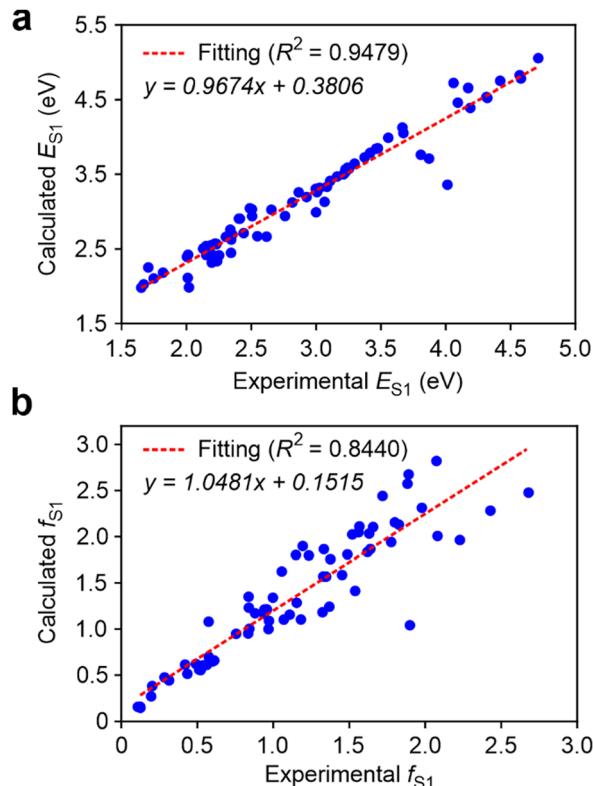


Fig. 2 The correlation between the experimental fitted and calculated (a) the first vertical excited state energy  $E_{S_1}$  and (b) the corresponding oscillator strength  $f_{S_1}$  under M06-2X/def2-TZVP/PCM.

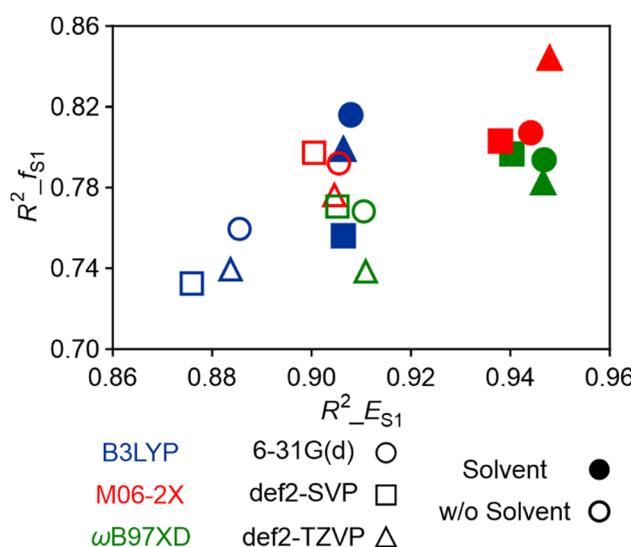


Fig. 3  $R^2$  values of  $E_{S_1}$  and  $f_{S_1}$  between the experimental fitted and calculated data under 18 computational methods.

Regarding  $\lambda$ , which involves the expensive geometry optimization of the  $S_1$  state, our initial approach utilized the M06-2X/6-31G(d) method to expedite the calculations, as its correlation for  $E_{S_1}$  and  $f_{S_1}$  was deemed acceptable. However, the obtained



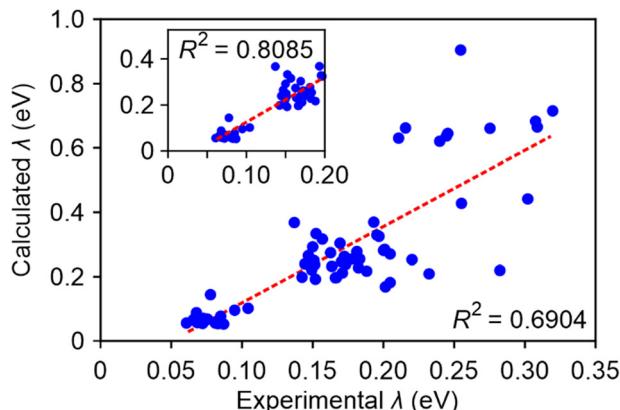


Fig. 4 The correlation between the experimental fitted and calculated  $\lambda$  using M06-2X/def2-TZVP/PCM (equation:  $y = 2.3690x - 0.1181$ ). The inset shows the results for small experimental fitted  $\lambda$  at 0–0.2 eV region (equation:  $y = 1.9650x - 0.0712$ ).

results for  $\lambda$  proved unsatisfactory, with an  $R^2$  of only 0.344 (Fig. S9a, ESI<sup>†</sup>). This poor performance can be primarily attributed to the inaccurate predictions for charged molecules, where the computed values deviated significantly from the experimental fitted data. Furthermore, even for few neutral molecules, the results exhibited large errors, further decreasing the  $R^2$  value for  $\lambda$ .

Consequently, we employed our best-performing method, M06-2X/def2-TZVP/PCM, and observed a favourable correlation between the calculated and the experimental fitted  $\lambda$ , with an  $R^2$  value of approximately 0.7 (Fig. 4). For charged molecules, the computed values also aligned well with the fitted line. Within the experimental fitted  $\lambda$  range of 0 to 0.2 eV, the correlation strength is further improved, yielding an  $R^2$  value of 0.809. This underscores the substantial impact of incorporating solvent effects on the accuracy of reorganization energy calculations. Additionally, it is worth noting that the theoretical values tended to overestimate the experimental results by a factor of approximately 2.

The unsatisfactory outcomes associated with the M06-2X/6-31G(d) method prompted us to investigate the solvent effect on our systems more closely. Firstly, to avoid possible state crossing during the process of  $S_1$  state optimization, we checked all the molecules and ensure their  $S_1$  state has the same character between the Franck–Condon state and adiabatic excited state with and without PCM. Fig. S9a (ESI<sup>†</sup>) highlights an intriguing observation: dyes such as **dasbt**, **pyr1**, **pyr2**, and the **sty** series (molecular details provided in the ESI<sup>†</sup>) exhibit notably low calculated  $\lambda$  compared to experimental fitted data without solvent effects. Interestingly, these molecules share a common characteristic – they are all charged species. However, when solvent effects are introduced into the calculations (Fig. S9b, ESI<sup>†</sup>), the results align more closely with the experimental data, suggesting a significant influence of charges on the system behaviour, especially for  $\lambda$ . This is illustrated more in detail in Table 1, showing that both **pyr1** and **sty6** exhibit similar trends under solvated and gas-phase conditions, with relative errors

Table 1 The experimental fitted and calculated  $\lambda$  using different methods and corresponding relative CPU time during the computation

Dyes	Experimental $\lambda$ (eV)	Methods	Calculated $\lambda$ (eV)	Relative CPU time
<b>pyr1</b>	0.195	M06-2X/6-31G(d)	0.046	1.000
		M06-2X/def2-TZVP	0.044	13.894
		M06-2X/6-31G(d)/PCM	0.318	1.048
		M06-2X/def2-TZVP/PCM	0.329	13.095
<b>sty6</b>	0.205	M06-2X/6-31G(d)	0.050	1.503
		M06-2X/def2-TZVP	0.053	11.873
		M06-2X/6-31G(d)/PCM	0.161	1.926
		M06-2X/def2-TZVP/PCM	0.182	17.894

becoming much smaller when the effect of the solvent is included.

Nevertheless, certain charged molecules remain unaffected by solvent effects, as evidenced by their consistent placement near the fitting curve, irrespective of the presence or absence of solvent environments. Examples include **rh6g** and **hidei**, where the calculated  $\lambda$  exhibit good correlations with the experimental fitted  $\lambda$  regardless of the inclusion of solvent effects (Table S2, ESI<sup>†</sup>). Therefore, it becomes evident that the presence of charges alone does not directly influence the calculated  $\lambda$  results. To gain deeper insights, we investigated the distance between the centroid of the electron and hole distributions, as illustrated in Fig. S11 and Table S2 (ESI<sup>†</sup>). Molecules like **pyr1** and **sty6** exhibit a larger separation between the electron and hole centroids (larger than 2 Å) compared to **rh6g** and **hidei** (less than 2 Å). This characteristic of spatially separated charge distributions makes them more susceptible to the stabilizing effects of the solvent environment.

Table 1 also reports the relative computational costs by using four cores for the different methods employed, which are representative for molecules of these sizes. These sizes, in turn, are typical for molecules considered in virtual screening protocols. Clearly, the significantly improved accuracy of calculations for  $\lambda$  including the implicit solvent effect carries a very small additional computational cost. However, the inclusion of a large basis set does not offer a major advantage in accuracy, despite being substantially more expensive computationally. When considered alongside the results for the excitation energy and oscillator strength it seems that M06-2X/6-31G(d)/PCM offers among the best compromises between accuracy and speed across different quantities to evaluate.

## Conclusions

In summary, we proposed a procedure to fit automatically experimental absorption spectra of medium-sized organic molecules in solution and obtain parameters that can be accessed experimentally: the excitation energy of the first excited state, corresponding oscillator strength, and exciton reorganization energy. By comparing computed and experimental parameters for a dataset of 71 molecules we can determine the accuracy expected from a range of TDDFT methods routinely used in the digital design of molecules for



organic electronics. Our study demonstrates a robust correlation between calculated and experimental  $S_1$  energy, with most  $R^2$  values exceeding 0.9, particularly achieving  $R^2 = 0.948$  with the M06-2X/def2-TZVP/PCM method. The same method also produces very good quality predictions of oscillator strength ( $R^2 = 0.844$ ) and reasonable predictions ( $R^2 = 0.690$ ) of the excitation reorganization energy (the quality of the latter prediction improves for technologically relevant dyes with reorganization energy  $< 0.2$  eV). This suggests that the M06-2X/def2-TZVP/PCM method exhibits the best overall performance among the 18 computational methods tested. Comparison of several methodologies highlights that the inclusion of implicit solvent effects is essential to achieve good predictivity while extended basis sets are not as critical: the results obtained with the smaller 6-31G(d) basis set seems to be sufficiently accurate for high-throughput screening applications. These benchmarks can be used to determine the expected accuracy of computational predictions and to introduce suitable corrections to the systematic errors inherent to the computational methods. Our procedure will yield increasingly more accurate and informative results as larger and homogenous experimental data sets become available in the future.

## Data availability

The code for the fitting processes is now shared in GitHub by [https://github.com/XiaoyuUoL/spectrum\\_fitting](https://github.com/XiaoyuUoL/spectrum_fitting).

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We are grateful for the financial support from the European Union (European Innovation Council, project no. 101057564). A. T. also thanks the support from the European Research Council (grant no. 101020369). We thank Dr Ömer Omar for useful discussions.

## Notes and references

- 1 G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller and O. A. Von Lilienfeld, *New J. Phys.*, 2013, **15**, 095003.
- 2 R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams and A. Aspuru-Guzik, *Nat. Mater.*, 2016, **15**, 1120–1127.
- 3 K. Zhao, Ö. H. Omar, T. Nematiaran, D. Padula and A. Troisi, *J. Mater. Chem. C*, 2021, **9**, 3324–3333.
- 4 Ö. H. Omar, X. Xie, A. Troisi and D. Padula, *J. Am. Chem. Soc.*, 2023, **145**, 19790–19799.
- 5 Y. Wu, S. Lu, Q. Zhou, M.-G. Ju, X. C. Zeng and J. Wang, *Adv. Funct. Mater.*, 2022, **32**, 2204579.
- 6 V. Bhat, C. P. Callaway and C. Risko, *Chem. Rev.*, 2023, **123**, 7498–7547.
- 7 J. T. Blaskovits, R. Laplaza, S. Vela and C. Corminboeuf, *Adv. Mater.*, 2024, **36**, 2305602.
- 8 C. Jamorski Jödicke and H. P. Lüthi, *J. Am. Chem. Soc.*, 2003, **125**, 252–264.
- 9 H. Qian, M. E. Cousins, E. H. Horak, A. Wakefield, M. D. Liptak and I. Aprahamian, *Nat. Chem.*, 2017, **9**, 83–87.
- 10 M. Albota, D. Beljonne, J.-L. Brédas, J. E. Ehrlich, J.-Y. Fu, A. A. Heikal, S. E. Hess, T. Kogej, M. D. Levin, S. R. Marder, D. McCord-Maughon, J. W. Perry, H. Röckel, M. Rumi, G. Subramaniam, W. W. Webb, X.-L. Wu and C. Xu, *Science*, 1998, **281**, 1653–1656.
- 11 P. K. Samanta, D. Kim, V. Coropceanu and J.-L. Brédas, *J. Am. Chem. Soc.*, 2017, **139**, 4042–4051.
- 12 N. Aizawa, Y.-J. Pu, Y. Harabuchi, A. Nihonyanagi, R. Ibuka, H. Inuzuka, B. Dhara, Y. Koyama, K.-I. Nakayama, S. Maeda, F. Araoka and D. Miyajima, *Nature*, 2022, **609**, 502–506.
- 13 Z. An, C. Zheng, Y. Tao, R. Chen, H. Shi, T. Chen, Z. Wang, H. Li, R. Deng, X. Liu and W. Huang, *Nat. Mater.*, 2015, **14**, 685–690.
- 14 J. Liang, X. Feng, D. Hait and M. Head-Gordon, *J. Chem. Theory Comput.*, 2022, **18**, 3460–3473.
- 15 M. Isegawa, R. Peverati and D. G. Truhlar, *J. Chem. Phys.*, 2012, **137**, 244104.
- 16 A. DeFusco, N. Minezawa, L. V. Slipchenko, F. Zahariev and M. S. Gordon, *J. Phys. Chem. Lett.*, 2011, **2**, 2184–2192.
- 17 S. Grimme and F. Neese, *J. Chem. Phys.*, 2007, **127**, 154116.
- 18 M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **129**, 104103.
- 19 D. Mester and M. Källay, *J. Chem. Theory Comput.*, 2019, **15**, 4440–4453.
- 20 N. M. O'Boyle, C. M. Campbell and G. R. Hutchison, *J. Phys. Chem. C*, 2011, **115**, 16200–16210.
- 21 D. Padula, Ö. H. Omar, T. Nematiaran and A. Troisi, *Energy Environ. Sci.*, 2019, **12**, 2412–2416.
- 22 R. Grotjahn, T. M. Maier, J. Michl and M. Kaupp, *J. Chem. Theory Comput.*, 2017, **13**, 4984–4996.
- 23 P.-F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama and D. Jacquemin, *J. Chem. Theory Comput.*, 2020, **16**, 1711–1741.
- 24 D. Jacquemin, V. Wathelet, E. A. Perpète and C. Adamo, *J. Chem. Theory Comput.*, 2009, **5**, 2420–2435.
- 25 Ö. H. Omar, T. Nematiaran, A. Troisi and D. Padula, *Sci. Data*, 2022, **9**, 54.
- 26 S.-J. Zou, Y. Shen, F.-M. Xie, J.-D. Chen, Y.-Q. Li and J.-X. Tang, *Mater. Chem. Front.*, 2020, **4**, 788–820.
- 27 Y. Zhu, S. Vela, H. Meng, C. Corminboeuf and M. Fumanal, *Adv. Opt. Mater.*, 2022, **10**, 2200509.
- 28 M. Riede, D. Spoltore and K. Leo, *Adv. Energy Mater.*, 2021, **11**, 2002653.
- 29 R. S. Sprick, J.-X. Jiang, B. Bonillo, S. Ren, T. Ratvijitvech, P. Guiglion, M. A. Zwijnenburg, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2015, **137**, 3265–3270.



30 B. Yang, Y. Lu, D. Jiang, Z. Li, Y. Zeng, S. Zhang, Y. Ye, Z. Liu, Q. Ou, Y. Wang, S. Dai, Y. Yi and J. Huang, *Adv. Mater.*, 2020, **32**, 2001227.

31 V. Bhat, P. Sornberger, B. S. S. Pokuri, R. Duke, B. Ganapathysubramanian and C. Risko, *Chem. Sci.*, 2023, **14**, 203–213.

32 N. J. Turro, *Modern molecular photochemistry*, University science books, 1991.

33 X. Xie and A. Troisi, *J. Phys. Chem. Lett.*, 2023, **14**, 4119–4126.

34 J. Aragó and A. Troisi, *Adv. Funct. Mater.*, 2016, **26**, 2316–2325.

35 J. R. Reimers, *J. Chem. Phys.*, 2001, **115**, 9103–9109.

36 R. Sarkar, M. Boggio-Pasqua, P.-F. Loos and D. Jacquemin, *J. Chem. Theory Comput.*, 2021, **17**, 1117–1132.

37 M. Caricato, G. W. Trucks, M. J. Frisch and K. B. Wiberg, *J. Chem. Theory Comput.*, 2011, **7**, 456–466.

38 A. Chrayteh, A. Blondel, P.-F. Loos and D. Jacquemin, *J. Chem. Theory Comput.*, 2021, **17**, 416–438.

39 D. Jacquemin, I. Duchemin, A. Blondel and X. Blase, *J. Chem. Theory Comput.*, 2016, **12**, 3969–3981.

40 E. J. Beard, G. Sivaraman, Á. Vázquez-Mayagoitia, V. Vishwanath and J. M. Cole, *Sci. Data*, 2019, **6**, 307.

41 A. S. Tarleton, J. C. Garcia-Alvarez, A. Wynn, C. M. Awbrey, T. P. Roberts and S. Gozem, *J. Phys. Chem. A*, 2022, **126**, 435–443.

42 M. Soltani Nejad and M. Alipour, *Phys. Chem. Chem. Phys.*, 2024, **26**, 879–894.

43 J. Cerezo, F. J. Avila Ferrer, G. Prampolini and F. Santoro, *J. Chem. Theory Comput.*, 2015, **11**, 5810–5825.

44 E. G. McRae, *J. Phys. Chem.*, 1957, **61**, 562–572.

45 A. Halpin, P. J. M. Johnson, R. Tempelaar, R. S. Murphy, J. Knoester, T. L. C. Jansen and R. J. D. Miller, *Nat. Chem.*, 2014, **6**, 196–201.

46 A. Baiardi, J. Bloino and V. Barone, *J. Chem. Theory Comput.*, 2013, **9**, 4097–4115.

47 F. Santoro, A. Lami, R. Improta, J. Bloino and V. Barone, *J. Chem. Phys.*, 2008, **128**, 224311.

48 J. Cerezo and F. Santoro, *J. Comput. Chem.*, 2023, **44**, 626–643.

49 S. Gozem and A. I. Krylov, *WIREs Comput. Mol. Sci.*, 2022, **12**, e1546.

50 A. Segalina, D. Aranda, J. A. Green, V. Cristino, S. Caramori, G. Prampolini, M. Pastore and F. Santoro, *J. Chem. Theory Comput.*, 2022, **18**, 3718–3736.

51 B. Karasulu, J. P. Götze and W. Thiel, *J. Chem. Theory Comput.*, 2014, **10**, 5549–5566.

52 C. Fang, B. Oruganti and B. Durbeej, *J. Phys. Chem. A*, 2014, **118**, 4157–4171.

53 H. Haghshenas and F. C. Spano, *J. Chem. Theory Comput.*, 2024, **20**, 4790–4803.

54 J. D. Schultz, J. Y. Shin, M. Chen, J. P. O'Connor, R. M. Young, M. A. Ratner and M. R. Wasielewski, *J. Am. Chem. Soc.*, 2021, **143**, 2049–2058.

55 K. Huang and A. Rhys, *Proc. R. Soc. London, Ser. A*, 1950, **204**, 406–423.

56 P. M. Hare, E. A. Price, C. M. Stanisky, I. Janik and D. M. Bartels, *J. Phys. Chem. A*, 2010, **114**, 1766–1775.

57 J. C. Garcia-Alvarez and S. Gozem, *J. Chem. Theory Comput.*, 2024, **20**, 7227–7243.

58 L. Goerigk and S. Grimme, *J. Chem. Phys.*, 2010, **132**, 184103.

59 A. Noelle, A. C. Vandaele, J. Martin-Torres, C. Yuan, B. N. Rajasekhar, A. Fahr, G. K. Hartmann, D. Lary, Y.-P. Lee, P. Limão-Vieira, R. Locht, K. McNeill, J. J. Orlando, F. Salama and R. P. Wayne, *J. Quant. Spectrosc. Radiat. Transfer*, 2020, **253**, 107056.

60 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Rev. C.01*, Gaussian Inc., Wallingford, CT, 2016.

61 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

62 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.

