Politechnika Wrocławska

Universitat de Girona

Field of science: Natural sciences
Discipline of science: Chemical sciences

Doctoral Dissertation

Computer-aided design of two-photon-absorbing organic chromophores

Lizaveta Petrusevich

Supervisors:
Dr. Robert Zaleśny
Dr. Josep Maria Luis

Keywords: computational chemistry, computational spectroscopy, two-photon absorption, organoboron complexes, fluorescent probes

Wrocław 2023

I'm deeply grateful to my supervisors Dr. Robert Zaleśny and Dr. Josep Maria Luis, who suggested the topic of my thesis and gave me an opportunity to participate in multiple fruitful collaborations. Their support and professional guidance on multi-photon processes and vibrational structure simulations were essential for my PhD studies.

I want to express my gratitude to Prof. Borys Ośmiałowski, who synthesized the majority of compounds investigated in my thesis, and to Prof. Marcin Nyk, who performed Z-scan two-photon absorption measurements which are discussed in my work. I'm thankful to Prof. Denis Jacquemin whose expertise on various aspects of computational chemistry gave me multiple hints for research performed in this thesis.

I gratefully acknowledge the support from the National Science Centre, Poland (Grant No. 2018/30/E/ST4/00457). All discussed below calculations were performed using computational resources generously provided by the Wroclaw Center for Networking and Supercomputing.

Contents

List of abbreviations

1PA one-photon absorption

2PA two-photon absorption

2PACS two-photon action cross section

2SM two-state model

3SM three-state model

ADC algebraic-diagrammatic construction

AH adiabatic Hessian

AHAS adiabatic Hessian after step

AS adiabatic shift

CPU central processing unit

CT charge transfer

dAH displaced adiabatic Hessian

DFA density functional approximation

DFT density functional theory

EE electrostatic embedding

EFP effective fragment potential

EOM-EE-CCSD equation-of-motion for excitation energies coupled-cluster with single and double substitutions

FC Franck–Condon

FCF Franck–Condon factor

FE full embedding

FQY fluorescence quantum yield

FWHM full width at half maximum

GFSM generalized few-state model

HT Herzberg–Teller

HWHM half width at half maximum

IMDHO independent mode displaced harmonic oscillator

LCM linear coupling model

MAE mean absolute error

MAPE mean absolute percentage error

MD molecular dynamics

ME mechanistic embedding

ML machine learning

MM molecular mechanics

PAH polyaromatic hydrocarbon

PCM polarizable continuum model

PE polarizable embedding

PES potential energy surface

PT1 first-order perturbation theory

PT2 second-order perturbation theory

QM quantum mechanics

QM/MM hybrid quantum mechanics/molecular mechanics

RAM random-access memory

RI-CC2 approximate coupled cluster singles and doubles model using the resolution-of-identity approximation

TD time-dependent

TD-DFT time-dependent density functional theory

TI time-independent

TPE tetraphenylethylene

VG vertical gradient

VH vertical Hessian

XMCQDPT2 extended multiconfiguration quasidegenerate perturbation theory

1 Introduction

Nowadays two-photon absorption (2PA) phenomenon has a significant influence on the development of molecular biology, medicine, optics, and information technology. Indispensable fundamental stages in the formation of the multiphoton absorption field were theoretical prediction by Maria Göppert-Mayer during her Ph.D. studies [1], and experimental demonstration after thirty years — in 1961 by Wolfgang Kaiser and Charles Geoffrey Blythe Garrett [2], preceded by the invention of the laser in 1960 [3]. The discovery of 2PA induced numerous momentous applications, namely Doppler-free 2PA spectroscopy [4–6], three-dimensional data storage [7, 8], optical power limiting [9, 10], microfabrication [11, 12], two-photon photodynamic therapy [13, 14], and two-photon-excited fluorescence microscopy [15]. Two-photon absorbers for some of these applications can be designed using computational chemistry toolbox and, in fact, one has been witnessing growing number of *in silico* studies in this field. The present dissertation also contributes to these efforts with emphasis on probes for two-photon-excited fluorescence microscopy. In the remainder of this section, a review of the latter application will be presented, followed by a presentation of computational methods used in modeling of two-photon absorption phenomenon.

Invented in 1990, two-photon excitation microscopy has wide practical usage in health-related studies today: cancer imaging [16–18], neurobiology and brain studies [19–21], visualisations of the eye [22–24], studies of skin [25, 26], imaging of kidney [27], imaging of dynamic processes in cells and tissues [28, 29], studies of the drug delivery [30], to name a few. The main mechanism underlying the application of 2PA is the fluorescence induced by the two-photon excitation. The successful usage of the two-photon excitation in bioimaging is caused by the transparency of biological materials in the range 650–1100 nm (i.e. the first biological window), and that the fitting of excitation wavelengths into the desirable range is attainable by minor modifications of 2PA probes. Due to the transparency of the biological environment, two-photon excitation has several advantages over one-photon excitation in fluorescence microscopy, particularly, larger penetration depth (up to 1000 μm) of excitation light into the sample [31–33], higher spatial resolution [32, 34], and the decrease of photobleaching and photodamaging [35, 36]. Moreover, currently, novel machine learning (ML) based approaches utilizing deep learning are being developed for denoising of 2PA images [37].

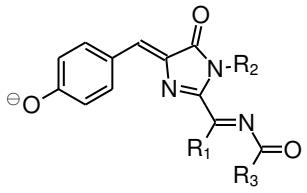
Numerous studies of 2PA microscopy techniques allowed the formulation of basic requirements for 2PA probes. Firstly, absorption of 2PA probes should lay in the first biological window to achieve transparency of the biological environment. Secondly, dyes should have a high 2PA cross section and high fluorescence quantum yield (FQY) to obtain bright images. The two latter characteristics may be evaluated together through the two-photon action cross section (2PACS) parameter — the product of 2PA cross section and FQY, which should be higher than 50 GM [38]. However, these requirements can be addressed not to the probe itself, but to its complex formed in the biological environment. For example, the contrast between the weak fluorescence of the pure probe in solution and the high fluorescence of its metal complex can be applied for the detection of metal ions [39, 40]. Since this aspect cannot be easily addressed using computational tools, in the remainder of the thesis the focus will be put on 2PA efficiency and excitation wavelengths. Dyes for 2PA bioimaging and their bioconjugated complexes should be also photostable. Besides the photophysical requirements listed above, 2PA probes should demonstrate high specificity towards the target, good solubility in water, and low cytotoxicity. There is a big variety of applicable 2PA materials, e.g., biological molecules [41, 42], inorganic materials [43–45], metalorganic complexes [46, 47] and small organic molecules [48,

49] — all these systems find usage in actual bioimaging techniques.

A family of interesting 2PA-active dyes is the inorganic materials. Perovskite materials are showing particularly high 2PA cross sections — up to 10^6 GM per nanoparticle [50]. However, due to the low stability of these crystals in water and their high toxicity [51], the potential of perovskites as 2PA probes for bioimaging has not been fully investigated yet. The silica (SiO_2) coating of perovskites increases the resistance of probes in aqueous environment and holds toxic ions inside nanoparticles [52, 53]. This strategy allowed to use $CsPbBr_3$ perovskites for real 2PA bioimaging techniques [43, 45]. The encapsulation of perovskites into phospholipids is another approach to prepare aqua-stable non-toxic probes for bioimaging [54]. Thus, nanocrystals are promising probes for 2PA bioimaging, but this field requires additional studies, and currently, it is rapidly developing.

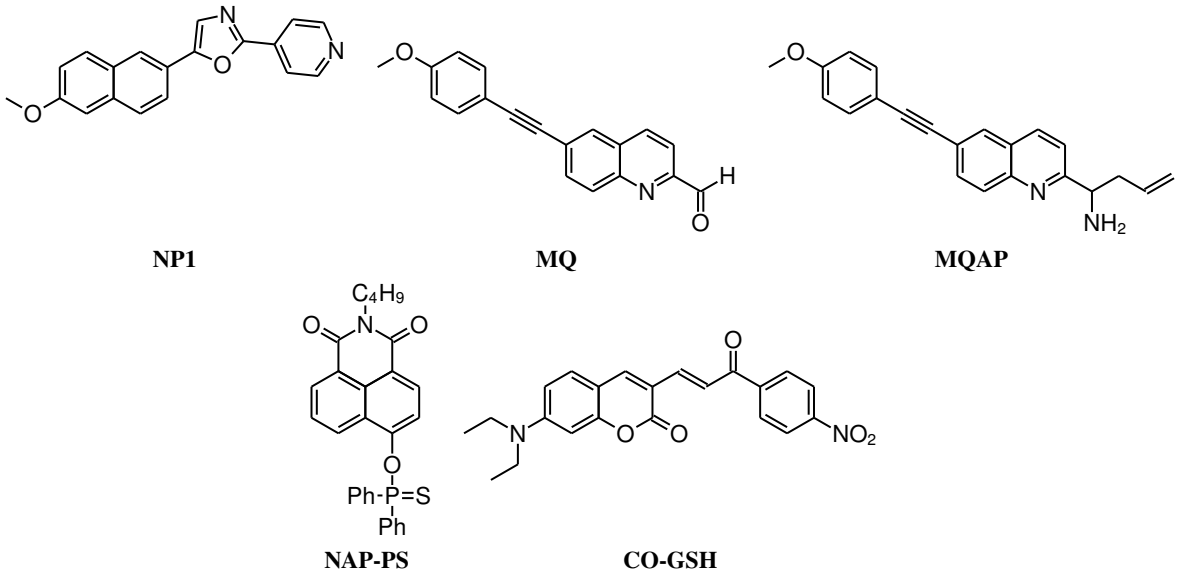
Fluorescent proteins contain a photoactive chromophore which is covalently bonded to the main sequence of amino acids. A good example is the tdKatushka2 protein, whose chromophore is shown in Scheme 1. This protein has high 2PA cross section in the first biological window (645 GM per protein molecule at 710 nm) and good FQY (equal to 0.37) [42, 55]; such excellent optical properties allowed to use this protein in multiple 2PA microscopy studies [56, 57]. The diversity of fluorescent proteins permits to use of a wide range of excitation energies, and it is possible to use a combination of green and red fluorescent proteins for multicolour imaging. For instance, EGFP (2PA cross section equal to 54 GM at 911 nm, and FQY equal to 0.76 at 510 nm [58]) and tdTomato (2PA cross section equal to 316 GM at 684 nm, and FQY equal to 0.65 at 581 nm [42, 59]) proteins are used for multichromatic microscopy [60]. A significant advantage of fluorescent proteins is their relatively low toxicity [61, 62]. Moreover, fluorescent proteins are soluble and stable in biological environments. However, the optical properties of fluorescent proteins are challenging to accurately predict due to the complex interactions between the chromophore and its surroundings. As far as the geometry of the chromophore is stabilized inside the protein by multiple hydrogen bonds, it is important to take into account the environment of such chromophore. The strong dependency of photophysical properties on the chromophore environment was observed for the first time together with the first reported crystal structures of green fluorescent proteins, where a single mutation of the protein sequence could increase excitation wavelength by 20 nm [63, 64]. Proteins mKate and mKate2 have the same chromophore, and only 5 mutations in the amino acids sequence result in significant differences between the optical properties of these two proteins, namely, mKate2 has 2PA cross section 200% higher and FQY 30% higher than mKate protein [55]; and eventually only mKate2 is suitable for bioimaging. The key reason is stabilization of either *cis*- or *trans*-isomer of the chromophore by the protein environment, i.e., the dark *cis*-form is more stable in mKate, while due to mutations the photoactive *trans*-form is dominant in mKate2. Besides, depending on the surroundings, the chromophore may switch between neutral, zwitterionic, cationic or anionic form through *cis-trans* isomerisation. Furthermore, the anionic chromophore may be stabilized by the environment as benzenoid or quinoid form. All described modifications lead to changes in the photophysical properties [65, 66]. Therefore, the design of novel 2PA probes with target optical properties based on fluorescent proteins is somewhat hampered by the required thoughtful consideration of the chromophore environment.

Small organic molecules also have several advantages as two-photon-excited fluorescent dyes for bioimaging. The compactness of these molecules allows them to penetrate through biological membranes. Synthesis of such dyes is simple, and their photophysical properties, solubility, and conjugation to specific molecules in live tissues and cells can be easily adjusted to the desirable values by adding the correct functional groups. Additionally, such organic dyes can



Scheme 1: Structure of the chromophore in tdKatushka2

be integrated into various nanoparticles to increase the solubility of dyes in the biological environment [67]. Several examples of fluorescent dyes that are used as probes in 2PA bioimaging are given in Figure 2. Dye **NP1** is used for measurements of pH in human tissues by 2PA microscopy since the protonation of this compound leads to a change of fluorescence colour [68]. Compound **MQ** can be applied for the detection of cysteine and homocysteine in living cells based on a cyclization reaction, which leads to switching off intramolecular charge transfer and, consequently, to the shift of emission wavelength [69]. For 2PA-bioimaging-based detection of formaldehyde one can use **MQAP** probe, which undergoes Aza-Cope reaction in the presence of target formaldehyde, which causes the formation of structure **MQ** and consequent changes of the emission spectrum [70]. Compound **NAP-PS** significantly changes both absorption and emission spectra in the presence of Hg^{2+}, therefore, this probe is suitable for colourimetric and ratiometric monitoring of mercury in living cells and tissues [71]. Finally, dye **CO-GSH** is applied as 2PA fluorescent probe for the detection of glutathione in living cells [72].



Scheme 2: Structures of typical 2PA probes for bioimaging

Multiple studies of 2PA materials allowed to pinpoint basic strategies for design of organic dyes with high 2PA cross sections [73–75]. Two-photon absorption cross section is proportional to the imaginary part of cubic hyperpolarizability. For this reason, basic design strategies include a combination of electron-accepting and electron-donating groups to form dipolar, quadrupolar and octupolar structures. Possible molecular arrangements are schematically demonstrated in Figure 1. In these structures, electron-donating and electron-accepting moieties are connected by π-

conjugated linkers that ease charge transfer between active groups upon excitation. All structural features described for the mentioned 2PA probes are shown in Scheme 2. The chromophore presented in 2PA-active fluorescent proteins (Scheme 1) also follows the same rules, namely, the chromophore is an extended π-conjugated system containing both electron-accepting and electron-donating groups, and the 2PA strength is controlled by the environment.

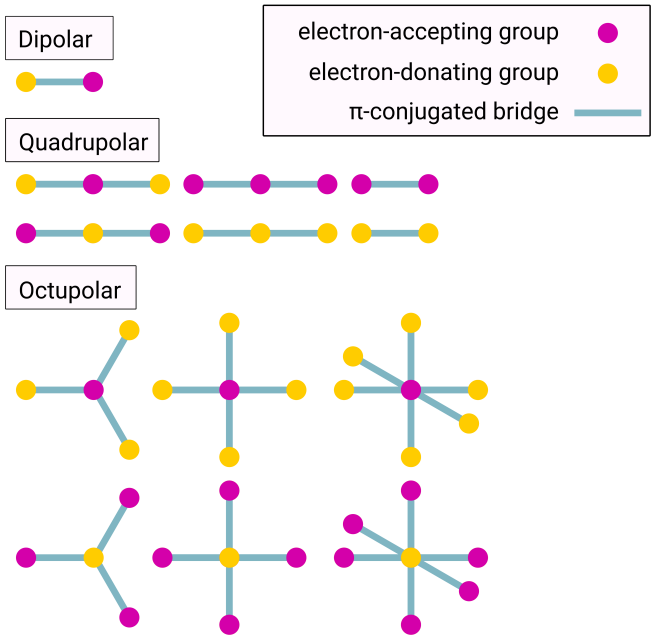
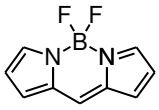


Figure 1: Schematic illustration of dipolar, quadrupolar, and octupolar structures

There are several standard ways to improve 2PA properties of small organic dyes. Firstly, 2PA cross section may increase on passing from dipolar to quadrupolar and octupolar dyes. In Ref. [76] two-photon absorption properties were studied for promising 2PA bioimaging dyes composed of electron-accepting 1,3,5-triazine center mono-, bi-, and tri-substituted with electron-donating groups. Dipolar mono-substituted derivative showed 2PA cross section equal to 34 GM, quadrupolar — 143 GM, and octupolar — 186 GM, confirming such strategy. In the paper [77] another strategy was adopted, namely, the central group was an electron-donating triphenylamine surrounded by electron-accepting triarylborane groups; likewise, dipolar structure demonstrated 2PA cross section equal to 91 GM, and octupolar structure — 335 GM. The second approach of 2PA strength improvement consists of using stronger electron-donating and electron-accepting groups. Following this route it is possible to increase 2PA cross section significantly [78–81]. This tactic has potential disadvantages, e.g., in aqueous environment strong electron-donating groups may lead to formation of hydrogen bonds and aggregation with further decrease of 2PA cross section; this phenomenon was described in several works both theoretically and experimentally [82–84]. However, some studies present contradictory conclusions, although all of them were observed in non-aqueous solvents. For example, the hydroxypyrimidine derivative studied in [85] remains in its monomeric form in tetrahydrofuran, but tends to form hydrogen bonds and aggregate in chloroform. Then, in chloroform, 2PA cross section of aggregated dye is equal to 1531 GM, while the 2PA cross section of the monomeric form in tetrahydrofuran is about half — 647 GM. Thus, strong electron-donating substituents should be applied carefully — taking into account possible undesirable effects caused by the formation of intermolecular hydrogen bonds. And finally, extending π-path between active electron-donating and electron-accepting groups leads to increase

of (hyper)polarizabilities and, consequently, to improvement of 2PA cross sections of dyes [86–89]. As an illustration, the results presented in [90] show how the introduction of two additional double bonds in between electron-donating and electron-accepting groups in dipolar structure leads to an increase of 2PA cross section by a factor of 4. Multiple studies demonstrated that dyes with π-bridge containing triple bond exhibit 2PA cross section lower than their analogues with double bond [91, 92]. However, a double bond may cause photoisomerisation of dye, which would corrupt the emission process, thus, leading to the decrease of imaging brightness.

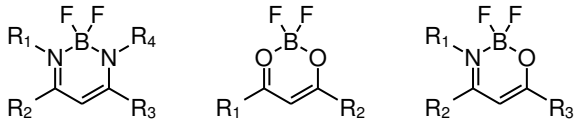
Organic dyes containing the BF_2 group (organoboron complexes, hereafter also referred to as difluoroborates/difluoroboranes) deserve more elaborate discussion. In these 2PA probes boron-containing moiety plays the role of the electron-accepting group, and the introduction of additional electron-donating groups increases 2PA cross section of these dyes, making them suitable for bioimaging [93, 94]. Below is given a detailed description of some famous organoboron families. Nowadays BODIPY-based family of dyes (Scheme 3) is highly popular amongst known probes for 2PA bioimaging. BODIPY derivatives were reported for the first time in 1968 by A. Treibs and F.-H. Kreuzer [95], and these compounds quickly became interesting due to their excellent photophysical properties. For example, BODIPY substituted with one tetraphenylethylene (TPE) group encapsulated into nanomicelles presented a 2PA cross section of 2.9×10^6 GM at 1040 nm, showed good biocompatibility and was successfully applied for 2PA bioimaging of blood vessels in the mouse [96]. In the paper [97], the authors compared BODIPY substituted with one, two and three TPE fragments, with dipolar, quadrupolar and octupolar structures, respectively; and, as expected, octupolar structure presented the highest 2PA cross section (256 GM per molecule) and was applied for further bioimaging of breast cancer cells. Structures containing two BODIPY fragments connected by truxene moiety presented 2PA cross sections over 5000 GM per molecule at 800-850 nm [98]. Furthermore, photophysical properties of BODIPY dyes can be adjusted by direct modifications of the core moiety; for example, the introduction of additional nitrogen atom to the BODIPY-core increased 2PA cross sections up to 2000 GM per molecule [99], and replacement of substituents at boron atom allowed to increase solubility of BODIPY dyes in water [100].



Scheme 3: BODIPY structure

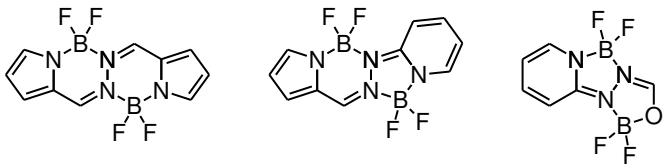
Following the outstanding strategy of the design of BODIPY derivatives with excellent photophysical properties, based on the skeleton with tetra-coordinated boron atom, dyes constructed with difluoroborane β-diiminate, difluoroborane β-diketonate, and difluoroborane β-ketoiminate groups (illustrated in Scheme 4) are also good 2PA biomarkers. Curcumine derivatives containing difluoroborane β-diketonate fragment presented 2PA cross sections exceeding 5000 GM in the first biological window [101]. The work [102] shows how conformational changes of difluoroborane β-diketonate dyes might be used for detection of anions, namely, in anion-free environment the dye has a straight conformation with the high 2PA cross section equal to 1500 GM at 800 nm, while the presence of anions leads to the formation of the complex with a bent conformation of the dye which has five times smaller 2PA cross section but preserves FQY values. In general, various modifications of difluoroborane β-diketonate are widely used in bioimaging and show a potential therapeutic effect on Alzheimer's Disease progression [103, 104]. Difluoborane β-ketoiminate

dyes also show properties suitable for bioimaging [105–107], for example, quadrupolar dye containing two N–BF_2–O groups has 2PACS higher than 800 GM in the first biological window [108]. In the paper [109], a comparison amongst several difluoroborane β-ketoiminates with dimethylaminophenyl electron-donating group shows that the introduction of a single double bond between electron-accepting and electron-donating group increases 2PA cross section by a factor of two. Thus, electron-accepting groups with slight modifications of the environment coordinating boron atom preserve good characteristics of 2PA properties.



Scheme 4: Basic structures of difluoroborane-based dyes

Recently, inspired by the excellent optical properties of BODIPY derivatives, two new highly fluorescent chromophores were proposed: BOPHY in 2014 [110] and BOPPY in 2018 [111] (Scheme 5). The literature lacks experimental data corresponding to 2PA properties of BOPHY dyes. However, BOPPY derivatives presented 2PA cross section up to 1000 GM at 770 nm; in addition, BOPPY showed low cytotoxicity, and 2PA bioimaging in living cells was performed using these dyes, confirming their applicability in imaging. Reported in 2021, the dye BOAPH (Scheme 5) does not show a high 2PA cross section (about 26 GM at 830 nm), but it was suitable for 2PA bioimaging of cells [112]. This dye was the starting point of the development of multiple novel core structures with two BF_2 groups, as, for instance, recently proposed dyes BOAPY [113], BOPYPY [114], BOPAHY [115], to name a few. These dyes are easy to synthesize and show excellent FQY values, but 2PA properties of these structures have not been reported yet. To sum up, listed above novel families of difluoroborates are promising as 2PA probes, but require further investigation.



Scheme 5: New core structures containing two BF_2 fragments

To conclude this discussion about organoboron complexes, the optical characteristics of such dyes can be controlled by including different electron-donating groups, by substitution at boron atom, by using a dipolar, quadrupolar or octupolar as the skeleton of the dyes, by the introduction of several BF_2 groups to the same molecule, by the extension of π-conjugation in structure, etc. Moreover, due to their potential for bioimaging, good optical properties and biocompatibility, boron difluorates deserve further investigation. In fact, recent study by Hajda et al. demonstrates that BF_2-functionalized benzothiazoles can be used as amyloid markers [116].

As discussed above there are numerous modifications and a plethora of possible 2PA probes. At the same time, wide usage of two-photon fluorescence microscopy requires a rapid design of new fluorescent dyes. Solid knowledge

about dependencies of 2PA properties on molecular architecture is helpful for faster construction of novel promising dyes, and currently "structure–property" databases are being developed by multiple experimental groups. Moreover, these databases can be implemented for further ML-based design of new dyes (for example, currently deep learning approaches can predict optical peaks of dyes with mean absolute error (MAE) of less than 7 nm [117]). Currently, there is a great need for a cost-effective procedure for the preselection of 2PA dyes. Furthermore, the wide usage of experimental techniques for these purposes is problematic, because syntheses of potential dyes are time-consuming, and Z-scan measurements and two-photon-induced fluorescence microscopy are still difficult to approach and cannot be routinely applied due to the expensive and complicated laser equipment required. Furthermore, experimentally obtained 2PA cross sections for the same compound are different from one experimental setup to another (even up to a factor of 11 [118]); and due to significant step in experimental measurements (up to 25 nm), resolution of recorded spectra is low. It is even possible that measured frequencies do not overlap with physical 2PA cross section maxima, and then the experimental data underestimate the magnitude of the 2PA process, causing inconsistencies in the "structure–property" database.

Computational chemistry can be an invaluable tool for the prediction of photophysical properties at a relatively small cost. With the help of computational chemistry, it is possible to predict vertical excitation energies (absorption maxima) and two-photon transition strengths (proportional to 2PA cross sections) and to determine vibrationally-resolved spectra. For example, in Ref. [119] authors presented density functional theory (DFT) calculations of 2PA properties for 280 large molecules with two N–BF_2–O groups in the gas phase, pinpointing the promising main structural modifications for enhancement of 2PA properties of dyes. This study revealed the following patterns: all dyes have excitation energy laying in the first biological window, vinylene linkers lead to higher 2PA cross sections, while phenylene moiety leads to the improvement of 2PA cross sections only for two out of three families of dyes. Such knowledge obtained thanks to the computational chemistry may significantly speed up the design of new 2PA probes. Despite some studies being performed at the gas phase [119], it is possible to perform calculations taking into account the influence of solvent on the biological environment. Nevertheless, it is still difficult to use theoretical chemistry for simulations of 2PA spectra, because trustworthy quantum theories require high computational costs and efficient approaches do not provide reliable results. For example, the relatively fast level of theory applied in computational study [119] shows generally good performance in the prediction of experimental trends, but recently it was demonstrated that such computational method provides doubtful results for difluoroboranes with certain electron-donating groups [120]. Evaluation of theoretical approaches through comparison with experimental 2PA spectra requires the consideration of the vibronic structure of multiple absorption bands, which is also a non-trivial task today. Finally, precise prediction of 2PA spectra is still challenging, since it cannot be limited by a narrow area of calculations, and requires the development of new simulation protocols which take into consideration multiple aspects.

Nowadays, electronic-structure calculations of two-photon transition strength (proportional to 2PA cross section) are implemented at various levels of theory, the most popular are time-dependent density functional theory (TD-DFT) [121, 122], coupled-cluster theories [123, 124], in particular, approximate coupled cluster singles and doubles model using the resolution-of-identity approximation (RI-CC2), equation-of-motion for excitation energies coupled-cluster with single and double substitutions (EOM-EE-CCSD) [125], algebraic-diagrammatic construction (ADC) [126], extended multiconfiguration quasidegenerate perturbation theory (XMCQDPT2) [127]. Any of these theories can be used

with generalized few-state model (generalized few-state model (GFSM)) [128, 129], which allows for the linking of 2PA strength with electronic-structure parameters. The comparison of previous approaches requires considering their computational costs and accuracy in the theoretical prediction of 2PA activity. The dyes that require the use of more than 1500 contracted Gaussian basis functions can be treated using TD-DFT without demanding extraordinary computational resources, while the above number is the limit for RI-CC2. On the other hand, EOM-EE-CCSD can only be used for calculations up to several hundreds of contracted Gaussian basis functions. XMCQDPT2 provides highly accurate 2PA cross sections and vertical excitation energies. Work [130] demonstrated that XMCQDPT2 in combination with effective fragment potential (EFP) lead to reliable 2PA predictions for the same anionic chromophore in different protein environments. The main disadvantage limiting the wide application of XMCQDPT2 is the huge random-access memory (RAM) allocation required. Because 2PA probes showing potential for bioimaging are usually represented by elongated π-conjugated structures containing at least several dozens of atoms (which translates into roughly 1000 contracted Gaussian basis functions when using augmented double-ζ atomic basis sets), the practical calculations of 2PA properties are limited to RI-CC2 and TD-DFT levels of theory.

For the prescreening of photophysical properties, the selected level of theory should provide prompt and reliable results. Desired rapidity can be achieved by TD-DFT calculations, and it is often used for scanning 2PA dyes. The key question is which density functional approximation (DFA) amongst the known ones should be used for estimation of 2PA strength. A study by Badaeva et. al. [131] used TD-DFT (B3LYP [132, 133]) to study quadrupolar structures that contain cyclohexanone (or piperidone) as the central moiety connected to different donor substituents through π-paths of different length. They found out that for all studied compounds calculated 2PA cross sections overestimated experimental values by 1.5-3 times. Nevertheless, the trends of calculated 2PA cross sections qualitatively reproduce experimental data, in particular, dyes with weaker –OCH_3 electron-donating group have smaller 2PA cross sections than dyes with stronger –$N(C_2H_5)_2$ electron-donating group, and the dyes with one double bond in π-bridge have smaller 2PA cross section than dyes with two double bonds. However, the experimental 2PA cross section of the dye with one bond is 1.64 times lower than of the analogous dye but with two bonds, while in the corresponding DFT calculations, this factor is equal to 3.81. Thus, the selected DFA is insufficient for the reproduction of quantitative patterns. Nevertheless, comparison between B3LYP 2PA cross sections and experimental values for a wide palette of 2PA-active dyes demonstrated a very good performance of B3LYP in simulations of excitation energies, and, in general, acceptable values of calculated 2PA cross sections [134]. SAOP functional [135] was used in other two works [136, 137] for simulations of 2PA spectra of quadrupolar and octupolar organic dyes. In both studies, the comparison between simulations and experiment shows that simulated excitation energies are unreliable, and simulated 2PA cross section values are overestimated by a factor of 10. Albeit, experimental patterns of 2PA changes are frequently correctly reproduced. The CAM-B3LYP [138] functional handles well charge-transfer excitations [139] and this explains its widespread use in the calculations of 2PA properties. For instance, it was used in a study of hexahelicene derivatives [140], and calculated 2PA cross sections significantly overestimated experimental values. In calculations of 2PA properties of BODIPY-like structures containing N–BF_2–O group [141] CAMB3LYP provides 2PA cross sections of the same magnitude as experimental spectra, but theoretical results exhibit different trends with respect to the experimental cross sections.

As explained above, although TD-DFT provides doubtful results in simulations of 2PA cross sections, it still may

be helpful for fast prescreening of changes of 2PA properties in big sets of dyes. Nevertheless, to avoid possible errors caused by the incorrect selection of a reliable DFA, it is needed to evaluate the performance of different functionals. Discrepancies between simulated and experimental magnitudes of 2PA can be explained by the incorrect selection of the broadening function and the broadening parameters, by the overlap of several absorption bands, and due to the neglect of the vibronic structure. Therefore, more solid benchmarking of DFAs should be performed using reference values calculated by accurate theoretical approximations rather than experimental cross sections. Multiple works performed comparison amongst known DFAs using as reference RI-CC2. It is important to mention that RI-CC2 provides a solid and trustworthy prediction of 2PA strengths. RI-CC2 is less accurate than CCSD, CC3, and EOM-EE-CCSD approaches, but it still can be used as a reliable reference in the validation of DFAs for simulations of organic molecules. For a set of small molecules (formaldehyde, diacetylene, and water) 2PA strengths computed by RI-CC2 showed a strong dependence on the atomic basis set, and RI-CC2 method generally gave values that are overestimated with respect to CC3 and CCSD reference by a factor of two or three [142]. Comparison between RI-CC2 and EOM-EE-CCSD based on the set of four neutral chromophores demonstrated that RI-CC2 slightly overestimates EOM-EE-CCSD (up to a factor 1.4). However, RI-CC2 may be considered as a good alternative to EOM-EE-CCSD, since RI-CC2 has satisfactory accuracy and smaller computational costs [143]. Importantly, the same work demonstrates that CAM-B3LYP functional significantly underestimated EOM-EE-CCSD up to a factor of 3. Moreover, CC2, CCSD, and CC3 and several DFA approaches were compared in calculations of 2PA strengths for retinal and its smaller prototypes [144], and coupled cluster models predicted similar results (the difference between CC2, CCSD, and CC3 do not exceed 10%), while, unfortunately, all studied DFAs underestimated 2PA cross sections about by a factor of 10. Therefore, though RI-CC2 is not the most accurate amongst known theories for 2PA calculations, it still can be a reliable reference method for the selection of DFAs.

There are several publications comparing the performance of B3LYP, CAM-B3LYP, and RI-CC2. A validation of CAM-B3LYP, B3LYP, and BP86 DFAs in comparison with RI-CC2 was performed based on 6 compounds containing $–NH_2$ and $–NO_2$ groups separated by different π-bridges in work [124]. Despite the generally poor performance of TD-DFT observed in this study, it is worth highlighting that amongst the investigated set of functionals, CAM-B3LYP is the only one showing acceptable results for both charge transfer (CT) and non-CT states. Another work compared CAM-B3LYP, B3LYP, and RI-CC2 in calculations of 2PA strengths using combined quantum mechanics (QM) and molecular mechanics (MM) approach for the description of three fluorophores in their biological surrounding [145]. In this study, all three methods demonstrated roughly the same average accuracy. However, B3LYP produced non-systematic errors in calculations of excitation energies and 2PA cross sections. Multiple works [120, 129, 146, 147] compared two-photon transition strengths computed by a wide palette of DFAs (CAM-B3LYP, B3LYP, BLYP [148, 149], LC-BLYP [150], BH&HLYP [151], PBE [152], PBE0 [153]) and by RI-CC2, and all of them arrived to the same conclusion: the most appropriate DFA for simulations of 2PA properties is CAM-B3LYP, which shows a pleasingly high Pearson's correlation coefficient (as large as 0.99). Note that there are currently no implementations allowing for computing 2PA strengths using meta-GGA functionals. Aiming at establishing "structure–property" relationships, 138 difluoroborates — 5 BF_2-containing electron-accepting cores with various electron-donating substituents — were theoretically investigated in Ref. [154]. In particular, this work included the comparison between CAM-B3LYP/aug-cc-pVDZ and RI-CC2/aug-cc-pVDZ levels of theory. It showed that CAM-B3LYP always underestimates 2PA cross

sections by a factor of 2-3, compared to RI-CC2, although CAM-B3LYP reliably predicts relative changes of 2PA cross sections upon structural modifications of dyes, even varying core structures. However, a benchmark performed in the recent study [120] showed that CAM-B3LYP also gives some erroneous values of 2PA cross sections. For instance, 2PA cross sections calculated by CAM-B3LYP for molecules containing cyano- group, and pentafluorophenyl group are significantly underestimated and do not follow trends predicted by RI-CC2. Nevertheless, other functionals investigated in the mentioned benchmarking showed even worse results for compounds with these groups. To sum up, currently amongst known DFAs there is a lack of functional providing as accurate results as RI-CC2 in calculations of 2PA transition strength. However, CAM-B3LYP is suitable for the prompt scanning of 2PA properties for big sets of large dyes, not relying on exact values, but considering only relative changes upon structural modifications. However, CAM-B3LYP should be applied carefully since it has been shown that it is not a suitable approach for dyes with certain substituent groups.

Interestingly, there are some attempts to develop simplified TD-DFT approaches for the estimation of two-photon absorption spectra [155]. Verifying obtained results with EOM-EE-CCSD and RI-CC2, new methods are showing acceptable performance taking into account their significantly reduced computational costs (20000 times faster than original TD-DFT). These methods allow the calculation 2PA properties of systems up to 2000 atoms using purely quantum theory. Nevertheless, such methodologies require the parametrization of multiple preset semiempirical integrals. Furthermore, the simplified TD-DFT without tuned parameters to match reference EOM-EE-CCSD or RI-CC2 values do not reproduce the general trends in the changes of experimental 2PA cross sections.

Usually 2PA probes are applied in solution or in biological environment. For this reason, precise calculations of 2PA cross sections should be performed taking into account surroundings, while initial efficient prescreening of 2PA cross sections can be performed in the gas phase. Indeed, 2PA cross sections of organic dyes are very sensitive to the solvent. For example, 2PA cross section of Reichardt's dye is equal to 38 GM in chloroform, 26 GM in dimethyl formamide, and 25 GM in dimethyl sulfoxide [156]. Currently, there are multiple theoretical solvation models available. The simplest and fastest model is polarizable continuum model (PCM), which was proposed 40 years ago [157]. PCM does not take into account the real atomistic structure of the environment, but reproduces the effects of the environment through distribution dielectric functions. The first studies that applied PCM for solvent description (water and cyclohexane) in simulations of 2PA properties demonstrated the strong dependency of cross sections and excitation energies on solvent [158]. In general, PCM is widely used to simulate the 2PA [159–161]. However, the PCM method might neglect some important interactions affecting the electronic structure of the dye, like for example the formation of hydrogen bonds or π-π stacking. Besides, this method can be applied only for simple solutions, but not for more complex environments, for example, proteins, where the 2PA properties crucially depend on the chromophore surrounding.

It is possible to overcome these challenges by combining QM and MM approaches, where 2PA-active dye is characterized by a precise quantum theory, while the solvent or biological microenvironment is described by simple molecular mechanics principles and the interactions between MM part and QM part are implemented into the QM effective Hamiltonian. As it has been explained above, there are many possible ways for the theoretical description of QM part (TD-DFT, RI-CC2, EOM-EE-CCSD, etc.). And the influence of the MM part on the dye can be introduced by various approximations:

- mechanistic embedding (ME) — interactions between QM and MM parts are described by molecular mechanics,

this approach is mostly used for studies of geometrical configurations;

- electrostatic embedding (EE) — the MM system polarize the QM system [162];
- polarizable embedding (PE) — the MM and QM systems polarize each other self-consistently [163];
- full embedding (FE) — Van der Waals interactions are also described by quantum mechanics.

A refined method for the description of MM part is the EFP approach [164, 165], where the whole MM part is divided into small fragments (with frozen geometries), and force field parameters of these fragments are obtained from *ab initio* calculations. All listed QM/MM methods are time-demanding in comparison with PCM, and they require extensive preoptimization of the microenvironment by molecular dynamics (MD) simulations, which frequently needs non-standard force fields. Statistically converged spectroscopic properties require a sequential QM/MM approach, in which the property of interest is calculated for multiple atomistic representations obtained from MD calculations [166, 167]. Taking into account the significant computational resources required, it is important to evaluate the improvement of QM/MM 2PA cross sections in comparison with PCM ones. For 4-dimethylamino-4'-nitrostilbene 2PA properties computed by PCM and QM/MM follow the similar patterns [168]. However, QM/MM better reproduced the shifts in experimentally observed vertical excitation energies. Values of 2PA cross section produced with PCM are up to 4 times higher than the ones obtained by EE approximation, and 2 times higher than the ones obtained by PE approach, but the ratio of 2PA cross sections in different solvents are roughly conserved by each solvation model.

Investigation of the neutral and anionic forms of the chromophore in green fluorescent protein was performed earlier at CAM-B3LYP/PE level of theory [169]. This study showed the problems of the PE model. The authors compared two approaches: *i)* water molecules (as found in the crystal structure) included to the MM part (simulated by PE model), and *ii)* water molecules included directly to the QM part (simulated together with chromophore molecule using DFT). For the neutral form of the chromophore, the study showed that 2PA cross section calculated using the approach *ii)* is higher by a factor of 3 than 2PA cross section calculated using the approach *i)*. This difference is caused by water molecules hydrogen bonded to the imidazole ring, since this interaction is not characterized correctly by PE model. However, in the case of the anionic chromophore, both QM and PE descriptions of water showed differences in the 2PA cross section smaller than 5%. Computed by the same level of theory one-photon absorption (1PA) spectrum nicely reproduced the experimental spectrum. Nevertheless, the comparison of computed and experimental 2PA spectra demonstrated that the ratio of the anionic form 2PA cross section to the neutral form 2PA cross section is significantly overestimated, what might be caused by imperfect performance of CAM-B3LYP approximation in simulations of 2PA properties.

In the study [170] the authors used CAM-B3LYP level of theory for the description of the QM part, and they compared gas phase, PCM, EE and PE models in the simulations of one- and two-photon electronic structure of rhodamine 6G. Again, the excitation energies computed by QM/MM approaches demonstrated better agreement with the experiment. All four studied models correctly predicted the brightest 2PA transition, but there were some discrepancies with the experiment, which according to the authors are due to the the lack of vibronic structure in the simulations. Moreover, even though the CAM-B3LYP predictive power is the best amongst known DFAs, CAM-B3LYP is not a fully reliable approach for comparison amongst precise solvent models due to the underestimated excited-state dipole moments.

In contrast, in the study [171] the authors used the highly accurate EOM-EE-CCSD approach to perform simulations of 2PA properties for microsolvated (in water) *para*-nitroaniline, thymine, and anionic chromophore of a yellow fluorescent protein. Solvent effects were included in simulations in several ways: *i)* all water molecules included in QM part (these 2PA cross sections were used as reference), *ii)* solvent is described by EFP, *iii)* solvent is described by EFP but including only electrostatic interactions, and *iv)* solvent is described by electrostatic embedding with point charges of atoms from a classic force field. For *para*-nitroaniline and thymine all three latter approaches give nearly the same 2PA cross sections, and comparing with the fully quantum simulations *i)* the errors did not exceed 20%. Thus, consideration of solvent polarization did not crucially affect results. Importantly, the inclusion of one or two explicit water molecules (forming hydrogen bonds with studied dyes) to the QM part allowed to decrease errors down to 3%. However, in the case of the anionic chromophore, the contribution of polarization effects was more significant, it allowed to decrease the errors in calculated 2PA cross sections by 5%, because this approach more accurately reproduces dipole-moment differences.

Thus, the microenvironment of a dye has a significant influence on its 2PA cross section, and therefore the selected solvent model can drastically affect simulations of 2PA properties. Amongst known models, PCM is the fastest approach, which shows correct trends in simple cases, but it is not applicable for accurate calculations and for complex (heterogeneous) environments. PCM problems can be overcome by refined computationally expensive QM/MM models which provide a more reliable prediction of 2PA properties.

All described above calculations were performed accounting only the electronic counterpart of response to electromagnetic perturbation, i.e. the effect of molecular vibrations was either neglected or treated classically by sampling over various conformations. However, as was already mentioned, the vibronic structure may significantly affect the shape of the spectrum. For example, if the intensity of the vibronic 0-0 transition is not dominant, the absorption maximum of the band will be shifted to higher frequencies. A historically important example is the two-photon transition in benzene which is electronically forbidden, but through the Herzberg–Teller (HT) effect in the vibronic structure, this transition is observed in spectrum [172]. To estimate the vibrationally-resolved spectrum it is needed to account for vibronic transitions and broadening mechanisms (inhomogeneous broadening, finite lifetime of excited states, etc.). One possible protocol to include these effects is schematically illustrated in Figure 2. The simulations of the vibronic structure require the estimation of equilibrium ground and excited state geometries and the corresponding Hessians (in some vibrational-structure theories). Therefore, the QM approach for estimation of vibronic properties has a significant influence on the final vibrationally-resolved spectra. In the case of fluorescent 2PA probes, the usage of accurate approaches (for example, RI-CC2) is restricted due to the high computational resources needed for simulations of excited-state Hessians. First, one has to choose a reliable TD-DFT approximation for this task. There are multiple publications benchmarking the performance of TD-DFT approximation in the evaluation of the band shape [173–176], which are mostly underlying a good performance of range-separated CAM-B3LYP, ωB97X-D [177], and LC-BLYP functionals. Recently, a deeper theoretical analysis of TD-DFT approaches demonstrated that some functionals (for instance, ωB97X [178]) present spurious oscillations in the profiles of energy derivatives with respect to the nuclear coordinates displacements [179]. On the contrary, LC-BLYP and CAM-B3LYP show high insensitivity to these numerical instabilities.

Nowadays, there are numerous available approaches for simulating the vibronic transitions, which can be classified

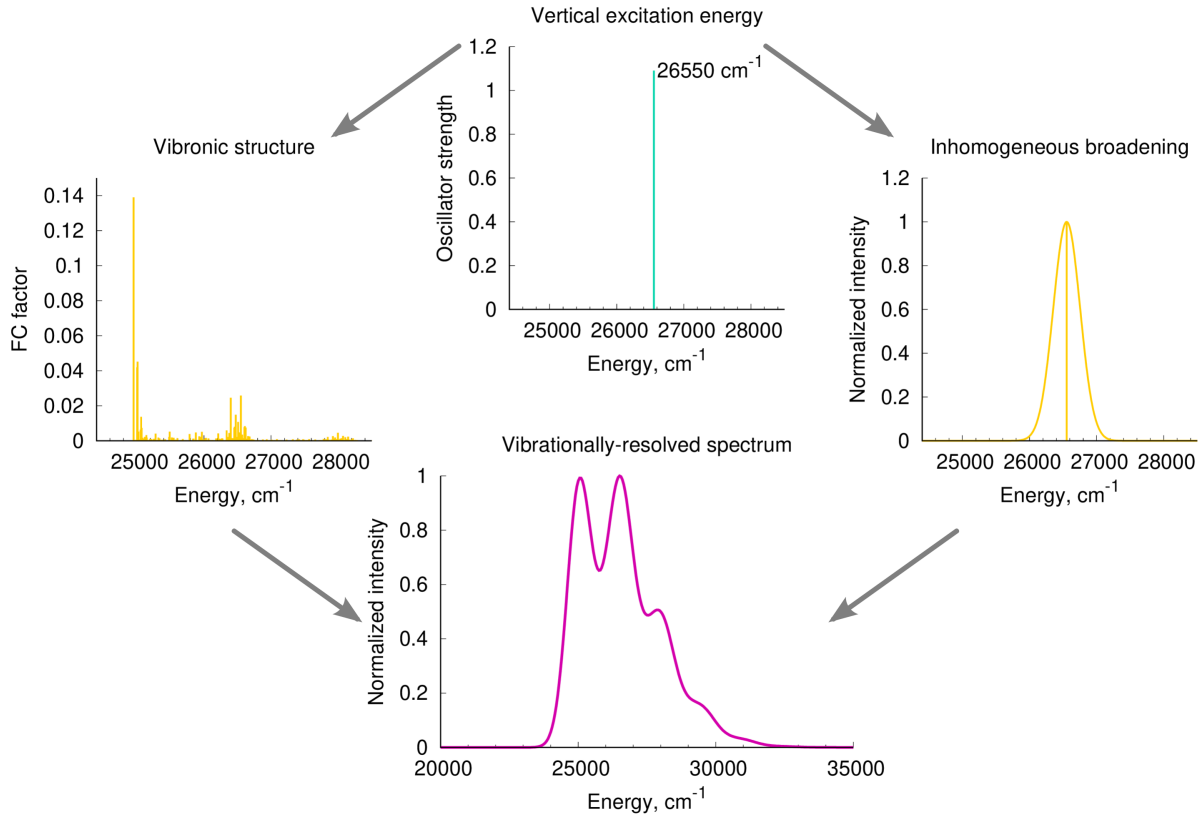


Figure 2: Schematic illustration of the computational route for estimation of vibrationally-resolved spectra

in several ways:

- according to their description of the vibrational space: adiabatic models (where the used geometry for the final state corresponds to the minima of the excited state potential energy surface (PES)) and vertical models (where the used geometry for the final state is estimated by a Newton-Raphson optimization step from the ground state geometry using the excited state gradient) [180];
- according to their description of electronic transition dipole μ: pure Franck–Condon (FC) theory (for bright transitions with high electronic transition dipole μ_{g} at the ground state geometry) or inclusion of HT corrections (for weak transitions with low μ_{g}) [181–183];
- according to their description of the shapes of the PES of the ground and excited states: harmonic or anharmonic [184].

First, it is worth mentioning that the assessment of the first and third classifications can be performed based on 1PA spectra because these approaches differ only in the mechanical treatment of vibrations, and within FC approximation these models would provide the same vibronic lineshape in 1PA and 2PA spectra. The comparison between simulated and experimental vibrationally-resolved spectra based on the 1PA process significantly simplifies the required calculations. Following this approach is not necessary to use as a reference the experimental 2PA spectra whose low resolution

is usually not sufficient for observation of vibronic structure. However, the importance of HT effects might be different for 1PA and 2PA spectra [182].

The literature contains extensive comparisons amongst vibrationally-resolved spectra simulated by the mentioned above models for small compounds. However, due to their big computational costs, there are not so many studies describing the vibronic spectrum of fluorescent dyes based on organoboron complexes. The latter analysis — referring to the anharmonic effects — has not been explored for fluorescent dyes yet. Several works showed a good agreement of vertical gradient (VG) approach with experiment for BF_2 dyes [185–187]. However, there is a lack of studies comparing vertical and adiabatic models for this family of dyes. Thus, it is hard to predict whether adiabatic approaches would outperform vertical models for dyes with tetra-coordinated boron atoms. Nonetheless, for fluorescent dyes with non-flexible structures adiabatic Hessian (AH) approach demonstrates a good performance [174].

In Ref. [188] the authors explored the vibrationally-resolved spectrum of BODIPY derivatives using AH approach and including HT contributions to the electronic structure. In that study the AH approach is in good agreement with experimental data, but in many cases full width at half maximum (FWHM) of the simulated spectra is bigger than the experimental one. In that work, it was demonstrated that for bright 1PA transitions in BODIPY dyes, HT contributions do not significantly improve upon FC approximation. Taking into account the high computational costs of HT approach (it requires numerical calculations of electronic transition dipole derivatives), the neglect of HT corrections is acceptable trade-off in the simulation of band shapes corresponding to bright one-photon transitions. In the Ref. [189] the authors assessed HT corrections in simulations of vibronic structure in one- and two-photon absorption bands for centrosymmetric dye with two BF_2 groups. This study confirmed the general conclusion, i.e., HT effects have a significant influence on weak 2PA transitions, and do not improve 2PA spectra of bright 2PA transitions. In conclusion, it should be highlighted that number of studies assessing the importance of HT contributions for the shapes of bands in 2PA is scarce and this hampers arriving at general conclusions regarding this important aspect.

To obtain the full vibrationally-resolved spectrum, the stick vibronic structure may be convoluted with a corresponding broadening function. For dyes in solution, the broadening is represented mainly by the inhomogeneous distribution of solvent molecules around the solute, and other broadening factors can be neglected. Solvent fluctuations within the time of spectrum measurements induce perturbations of the dye's electronic structure, which causes inhomogeneous broadening of spectrum [190, 191]. Inhomogeneous broadening has a wide range of possible values depending on solvent-solute interactions [192, 193]. There are two main approaches for estimation of inhomogeneous broadening, e.g., through the simulations of multiple discrete solute-solvent configurations [194, 195] or using the Marcus model [196–200].

An important technical issue is the lineshape function for the description of the inhomogeneous broadening. The two simplest models are Lorentzian and Gaussian distributions. Lorentzian lineshape was applied originally for the description of spectral lines in the gas phase and aggregates [201]. However, when there are no strong interactions between solute and solvent, Gaussian distribution gives the most suitable lineshape for the description of dye spectrum in a solvent environment because solvent undergoes slow fluctuations [202–204].

Marcus model is computationally attractive since it does not require simulations of large sets of solute-solvent configurations. However, it predicts solvent influence on solute's absorption broadening based on interactions between the solvent and the dye using classical statistics of the solvent fluctuations. Marcus formula demonstrated a good

performance in multiple works [205, 206]. For example, vibrationally-resolved absorption spectra simulated using a combination of AH approach (including HT effect) with Marcus formula demonstrated an excellent agreement with experiment for coumarin C153 and 4-aminophthal-imide dyes [207]. However, the use of this model is limited only to simple solvent microenvironments, and it can not be applied in heterogeneous environments, biological systems, or solid states. The mentioned restrictions can be bypassed through classic approaches based on the sampling of MD trajectory and hybrid quantum mechanics/molecular mechanics (QM/MM) simulations of solute in the solvent using a discrete representation of solvent molecules. In Ref. [195], the authors simulated inhomogeneous broadenings for 500 solute-solvent configurations at RI-CC2/PE level of theory. For 2PA spectra, vibronic structure simulated by the vertical model taking into account HT corrections and convoluted with obtained inhomogeneous broadening presented a very good agreement with the experimental data. Moreover, in several works, this approach was successfully used for simulations of fluorescent dyes with tetra-coordinated boron atoms. For instance, in the Ref. [208] the authors used the same model for the estimation of inhomogeneous broadenings for dyes with N–BF_2–O group in chloroform solvent. Even without accounting vibronic structure, the experimental large broadenings are well-reproduced by their calculations. In the Ref. [186] the VG vibronic model was convoluted with broadening obtained from 1000 solute-solvent snapshots calculated at CAM-B3LYP/PE level of theory, obtaining results in good agreement with experiment.

A new promising approach for simulating vibrationally-resolved spectra of flexible molecules in condensed phase is adiabatic-molecular dynamics combined with the generalized vertical Hessian (VH) approach, which was published in 2020 [209, 210]. This approach is based on the separation of "soft" and "stiff" vibrational modes, where "soft" modes are treated with MM approximations, while "stiff" modes are considered as much faster and described at QM level. The authors obtained geometrical configurations around "soft" vibrational modes from MD simulation, calculated the Hessians for these snapshots, excluded "soft" vibrational modes from internal coordinates, and for each geometrical configuration simulated VH vibronic structure based on only "stiff" modes. The final absorption spectrum is the result of averaging the vibronic spectra obtained for each mode. As far as simulations are performed assuming the solvent environment, inhomogeneous broadening is already included in the final spectra. This model requires high computational resources, however, it demonstrated a good agreement with the experimental data.

To sum up, the current state of computational chemistry allows the reproduction and rationalization of the experimental studies of 2PA fluorescent probes. However, some limitations (high computational costs, gaps in literature, lack of simulation protocols) remain and constrain the broad applicability of computational chemistry in the design of 2PA fluorescent dyes and the prescreening of photophysical properties.

2 Objectives

The described interest in new two-photon-excited fluorescent dyes for imaging applications and the room for improvement of existing simulation protocols motivated the research presented in this dissertation. Two main research objectives are:

- the computer-aided design of fluorescent dyes with refined one- and two-photon absorption properties that can be suitable for 2PA-based bioimaging applications;
- the development of simulation protocols for the estimation of vibrationally-resolved spectra with improved efficiency and accuracy.

The description of the research performed during my PhD studies is split into chapters according to these two main objectives. Boron-containing dyes were selected for presented studies due to their high potential in 2PA-based bioimaging, as it was formerly discussed.

The first objective encompasses the investigation of four sets of organoboron complexes presenting a large variety of structural motifs: various electron-donating groups, additional heteroatom in the electron-accepting BF_2-containing moiety, other substituents at boron atom, presence of a center of inversion or benzannulation. These studies aim to contribute to the current knowledge of "structure–property" relationships, which can hint at novel design strategies of fluorescent probes for 2PA-based bioimaging.

The second objective is devoted to the development of a fully non-empirical and efficient computational protocol for simulations of vibrationally-resolved absorption spectra. To this end, several aspects will be considered: *i)* the assessment of vertical and adiabatic vibrational spaces in the harmonic approximation; *ii)* the evaluation of mechanical anharmonicity and Herzberg-Teller effects; *iii)* the comparison between EE and PE models in simulations of inhomogeneous broadenings, and *vi)* the development of a novel machine-learning–based approach for efficient estimations of inhomogeneous broadening. The ultimate goal behind these efforts is to combine the outcomes of these studies to design a reliable computational route for the prescreening of photophysical properties of organic dyes.

3 Theoretical background

3.1 Electronic-structure theories

Two-photon absorption cross section

Two-photon absorption cross section is one of the key values for 2PA probes. It can be calculated from the two-photon transition strength in the following way:

$$\sigma^{\mathrm{2PA}}(\omega) = \frac{4\pi^3 \alpha a_0^5 \omega^2}{c} g(2\omega) \delta^{\mathrm{2PA}} \tag{1}$$

where ω is the two-photon excitation energy, α is the fine structure constant, a_0 is the Bohr radius, $g(2\omega)$ is the lineshape function, and δ^{2PA} is the two-photon transition strength. In the first part of the dissertation, the non-Hermitian coupled-cluster theory framework was used to determine the rotationally averaged two-photon transition strength (between electronic states 0 and J) assuming one source of linearly polarized monochromatic light beam:

$$\delta_{0J}^{\mathrm{2PA}} = \frac{1}{15} \sum_{\mu} \sum_{\nu} \left[M_{J\leftarrow 0}^{\mu\mu} M_{0\leftarrow J}^{\nu\nu} + M_{J\leftarrow 0}^{\mu\nu} M_{0\leftarrow J}^{\mu\nu} + M_{J\leftarrow 0}^{\mu\nu} M_{0\leftarrow J}^{\nu\mu} \right] \quad \mu, \nu \in x, y, z \tag{2}$$

where $M_{J\leftarrow 0}^{\mu\mu}$ and $M_{0\leftarrow J}^{\nu\nu}$ stand for the left and right second-order transition moments, respectively. For one source of photons, *i.e.*, $\omega = \frac{1}{2}\omega_J$, they are given by:

$$M_{0\leftarrow J}^{XY} = \sum_{K} \left(\frac{\langle 0|X|K\rangle\langle K|Y|J\rangle}{\frac{1}{2}\omega_J - \omega_K} + \frac{\langle 0|Y|K\rangle\langle K|X|J\rangle}{\frac{1}{2}\omega_J - \omega_K} \right) \tag{3}$$

$$M_{J\leftarrow 0}^{XY} = \sum_{K} \left(\frac{\langle J|X|K\rangle\langle K|Y|0\rangle}{\frac{1}{2}\omega_J - \omega_K} + \frac{\langle J|Y|K\rangle\langle K|X|0\rangle}{\frac{1}{2}\omega_J - \omega_K} \right) \tag{4}$$

where ω_K (ω_J) is the excitation energy between the electronic states 0 and K (J).

In simplest cases, the lineshape function $g(\omega)$ (in Equation 1) can be represented by mathematical models, for example, Lorentzian distribution $L(2\omega)$ or Gaussian function $G(2\omega)$:

$$L(2\omega) = \frac{1}{\pi} \frac{\Gamma}{(2\omega - \omega_0)^2 + (\Gamma)^2} \tag{5}$$

$$G(2\omega) = \frac{\sqrt{\ln 2}}{\Gamma\sqrt{\pi}} \exp\left[-\ln 2 \left(\frac{2\omega - \omega_0}{\Gamma} \right)^2 \right] \tag{6}$$

In these equations Γ is the broadening factor (half width at half maximum (HWHM)), ω_0 is the excitation energy, ω is the photon energy. Two-photon absorption cross section is proportional to the two-photon transition strength, and assuming that standard deviation does not undergo significant changes from one compound to another inside studied sets, the two-photon absorption cross section would follow the same trends as the two-photon transition strength.

Generalized few-state model

The previously mentioned GFSM [211] allows not only to estimate 2PA transition strengths but also to pinpoint the key states in studied transitions and to understand deeper the nature of the nonlinearity. Inserting Equations 3 and 4

into Equation 2, one can derive the expression for a generalized few-state model for non-Hermitian theories, where the left and right transition moments are different. The corresponding expression for the two-photon absorption strength is given by:

$$\begin{aligned}
\delta_{0J}^{\mathrm{GFSM}} &= \sum_K \sum_L \delta_{0JKL}, \\
\delta_{0JKL} &= \frac{2}{15\Delta E_K \Delta E_L}(\alpha+\beta), \\
\alpha &= |\mu^{JK}||\mu^{K0}||\mu^{0L}||\mu^{LJ}| \left(\cos\theta_{JK}^{K0}\cos\theta_{0L}^{LJ} + \cos\theta_{JK}^{0L}\cos\theta_{K0}^{LJ} + \cos\theta_{JK}^{LJ}\cos\theta_{K0}^{0L}\right) \\
\beta &= |\mu^{JL}||\mu^{L0}||\mu^{0K}||\mu^{KJ}| \left(\cos\theta_{JL}^{L0}\cos\theta_{0K}^{KJ} + \cos\theta_{JL}^{0K}\cos\theta_{L0}^{KJ} + \cos\theta_{JL}^{KJ}\cos\theta_{L0}^{0K}\right)
\end{aligned} \tag{7}$$

In the above expression, the superscripts distinguish between left ($L0$) and right moments ($0L$) and $\Delta E_K = \frac{1}{2}\omega_J - \omega_K$. The term $\theta_{\mathrm{PQ}}^{\mathrm{RS}}$ in Equation 7 represents the angle between the transition dipole moment vectors μ^{PQ} and μ^{RS}. In the case of theories with Hermitian structure, *i.e.* where the left and right moments are equal, the above expression reduces to the one for Hermitian theories:

$$\begin{aligned}
\delta_{0J}^{\mathrm{GFSM}} &= \sum_K \sum_L \frac{4}{15\Delta E_K \Delta E_L} \times \\
&\quad |\mu^{0K}||\mu^{KJ}||\mu^{0L}||\mu^{LJ}| \left(\cos\theta_{JK}^{0K}\cos\theta_{0L}^{LJ} + \cos\theta_{JK}^{0L}\cos\theta_{0K}^{LJ} + \cos\theta_{JK}^{LJ}\cos\theta_{0K}^{0L}\right)
\end{aligned} \tag{8}$$

Any number of intermediate states K and L can be chosen in the generalized few-state model expressions in Equations 7 and 8. In this thesis, we will use two-state model (2SM) and three-state model (3SM). In 2SM, K and L can be either the ground state 0 or the final excited state J, whereas in 3SM an additional state is also included for both K and L. For instance, four terms contribute to the 2SM expression for $\delta(\mathrm{2SM})$ are δ_{0J00}, δ_{0J0J}, δ_{0JJ0} and δ_{0JJJ} Similarly, in 3SM with intermediate state I, nine terms contribute to $\delta(\mathrm{3SM})$ are δ_{0J00}, δ_{0J0I}, δ_{0J0J}, δ_{0JI0}, δ_{0JII}, δ_{0JIJ}, δ_{0JJ0}, δ_{0JJI}, δ_{0JJJ}.

3.2 Vibrational-structure theories

Harmonic approximation

To facilitate the discussion of results obtained within the thesis, the theoretic basics of applied approaches are given below [180]. Firstly, one needs to build the vibrational space. The harmonic PES of the initial state V_{g} can be written in the matrix form as:

$$V_{\mathrm{g}}(\boldsymbol{Q}_{\mathrm{g}}) = \frac{1}{2}\boldsymbol{Q}_{\mathrm{g}}^{\mathrm{T}}\boldsymbol{\Omega}_{\mathrm{g}}^{2}\boldsymbol{Q}_{\mathrm{g}} \tag{9}$$

where $\boldsymbol{\Omega}_{\mathrm{g}}$ is the diagonal matrix of the normal frequencies of vibration, and $\boldsymbol{Q}_{\mathrm{g}}$ is the vector of N associated normal coordinates:

$$\boldsymbol{Q}_{\mathrm{g}} = \boldsymbol{L}_{\mathrm{g}}^{\mathrm{T}}\boldsymbol{M}^{1/2}(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{g}}) \tag{10}$$

where $\boldsymbol{x}$ is the $3N_a$ column vector of Cartesian coordinates (N_a is the number of atoms in the studied system), $\boldsymbol{x}_\mathrm{g}$ corresponds to the equilibrium geometry of the initial state, $\boldsymbol{M}$ is the diagonal matrix of atomic masses, and $\boldsymbol{L}_\mathrm{g}$ is the unitary matrix that transforms the mass-weighted Cartesian displacement coordinates to normal coordinates.

In AH model the equivalent equations are used to describe the PES of the final state V_f:

$$V_\mathrm{f}(\boldsymbol{Q}_\mathrm{f}) = E_\mathrm{ad} + \frac{1}{2}\boldsymbol{Q}_\mathrm{f}^\mathrm{T}\boldsymbol{\Omega}_\mathrm{f}^2\boldsymbol{Q}_\mathrm{f} \tag{11}$$

$$\boldsymbol{Q}_\mathrm{f} = \boldsymbol{L}_\mathrm{f}^\mathrm{T}\boldsymbol{M}^{1/2}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{f}) \tag{12}$$

where E_ad represents the minimum energy of the final state PES with respect to the minimum energy of the initial state PES, and $\boldsymbol{x}_\mathrm{f}$ corresponds to the equilibrium geometry of the final state. The variable $\boldsymbol{x}$, which is common for the initial and final states, allows to express $\boldsymbol{Q}_g$ as a function of $\boldsymbol{Q}_f$:

$$\boldsymbol{Q}_\mathrm{g} = \boldsymbol{J}\boldsymbol{Q}_\mathrm{f} + \boldsymbol{K} \tag{13}$$

$$\boldsymbol{J} = \boldsymbol{L}_\mathrm{g}^{-1}\boldsymbol{L}_\mathrm{f} \tag{14}$$

$$\boldsymbol{K} = \boldsymbol{L}_\mathrm{g}^{-1}\boldsymbol{M}^{1/2}(\boldsymbol{x}_\mathrm{f} - \boldsymbol{x}_\mathrm{g}) \tag{15}$$

$\boldsymbol{J}$ is Duschinsky rotation matrix [212], which rotates the normal modes of the final state, and $\boldsymbol{K}$ is a displacement vector. An important step here is the elimination of translational and rotational normal modes coordinates by excluding their mixing with vibrational coordinates. In the end, the final state PES can be written as a function of normal coordinates of the initial state $\boldsymbol{Q}_\mathrm{g}$:

$$V_f(\boldsymbol{Q}_g) = E_\mathrm{ad} + \frac{1}{2}\boldsymbol{K}^\mathrm{T}\boldsymbol{\Omega}_\mathrm{f}\boldsymbol{J}^\mathrm{T}\boldsymbol{K} - \boldsymbol{K}^\mathrm{T}\boldsymbol{\Omega}_\mathrm{f}\boldsymbol{J}^\mathrm{T}\boldsymbol{Q}_g + \frac{1}{2}\boldsymbol{Q}_g^\mathrm{T}\boldsymbol{\Omega}_\mathrm{f}\boldsymbol{J}^\mathrm{T}\boldsymbol{Q}_g \tag{16}$$

In VG approach [213] (also referred to as linear coupling model (LCM) [214] and independent mode displaced harmonic oscillator (IMDHO) [215]) initial and final states have the same shape of PES. Both states share the same normal modes and frequencies, but they have different equilibrium geometries. The geometry of the final state is predicted based on a Newton-Raphson step from the ground state geometry using the excited state gradient. Therefore, in VG approach, the vectors of normal modes of the final state are the same as for the ground state, $\boldsymbol{Q}_\mathrm{VG} = \boldsymbol{Q}_\mathrm{g}$; the diagonal matrix of normal frequencies of the final state is the same as the one for the ground state, $\boldsymbol{\Omega}_\mathrm{VG} = \boldsymbol{\Omega}_\mathrm{g}$; the Duschinsky matrix is equal to the identity matrix, $\boldsymbol{J}_\mathrm{VG} = 1$; and the displacement vector is given by: $\boldsymbol{K}_\mathrm{VG} = -(\boldsymbol{\Omega}_\mathrm{g})^{-2}\boldsymbol{L}_g^{-1}\boldsymbol{M}^{1/2}\boldsymbol{g}_x$ (where $\boldsymbol{g}_x$ is the final state energy gradient in Cartesian coordinates).

The internal coordinates [216] describe vibrations in terms of bonds, angles and dihedrals. Compared to the Cartesian coordinates, the internal ones provide an easier exact separation between translational, rotational, and vibrational normal modes. The internal coordinates $\boldsymbol{s}$ can be expressed through Cartesian coordinates $\boldsymbol{x}$ around equilibrium geometry $\boldsymbol{x}_0$ using Taylor expansion:

$$\boldsymbol{s}(\boldsymbol{x}) = \boldsymbol{s}_0 + \boldsymbol{B}(\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^T\boldsymbol{\beta}(\boldsymbol{x} - \boldsymbol{x}_0) + ... \tag{17}$$

where $\boldsymbol{s}_0 = \boldsymbol{s}(\boldsymbol{x}_0)$, $\boldsymbol{B} = \boldsymbol{s}'(\boldsymbol{x}_0)$, $\boldsymbol{\beta} = \boldsymbol{s}''(\boldsymbol{x}_0)$, and primes denote derivatives of the internal coordinates with respect to the Cartesian coordinates.

Now we will introduce the theory required to perform the simulations of spectral lineshape. In time-independent (TI) formalism, the stick vibronic spectrum can be described using sum-over-states expressions to describe the transitions from $|e_\mathrm{g}\rangle|\boldsymbol{w}_\mathrm{g}\rangle$ to $|e_\mathrm{f}\rangle|\boldsymbol{w}_\mathrm{f}\rangle$. $|\boldsymbol{w}_\mathrm{g}\rangle$ and $|\boldsymbol{w}_\mathrm{f}\rangle$ are vibrational states that correspond to electronic states $|e_\mathrm{g}\rangle$ and $|e_\mathrm{f}\rangle$, respectively. In this case, the spectrum intensity (depending on the frequency ω) is:

$$S(\omega) = C\omega \sum_{\boldsymbol{w}_\mathrm{g},\boldsymbol{w}_\mathrm{f}} \rho_{\boldsymbol{w}_\mathrm{g}}(T) P(\boldsymbol{w}_\mathrm{g},\boldsymbol{w}_\mathrm{f}) \delta(E_{e_\mathrm{g}\boldsymbol{w}_\mathrm{g}} - E_{e_\mathrm{f}\boldsymbol{w}_\mathrm{f}} + \hbar\omega) \tag{18}$$

In this equation C is a multiplicative coefficient [217], $\rho_{\boldsymbol{w}_\mathrm{g}}(T)$ is the Boltzmann population of state $\boldsymbol{w}_\mathrm{g}$ at temperature T, $E_{e_\mathrm{g}\boldsymbol{w}_\mathrm{g}}$ and $E_{e_\mathrm{f}\boldsymbol{w}_\mathrm{f}}$ are energies of corresponding states, δ is the Dirac delta function (which is replaced by broadening function), and $P(\boldsymbol{w}_\mathrm{g},\boldsymbol{w}_\mathrm{f})$ is a quantity which rules the intensity of of the stick band. In 1PA spectrum:

$$P(\boldsymbol{w}_\mathrm{g},\boldsymbol{w}_\mathrm{f}) = |\langle \boldsymbol{w}_\mathrm{g}|\boldsymbol{\mu}_{e_\mathrm{g},e_\mathrm{f}}|\boldsymbol{w}_\mathrm{f}\rangle|^2 \tag{19}$$

where $\boldsymbol{\mu}_{e_\mathrm{g},e_\mathrm{f}} = \langle e_\mathrm{g}|\boldsymbol{\mu}|e_\mathrm{f}\rangle$ is the electronic transition dipole. Expanding it into a Taylor series:

$$\boldsymbol{\mu}_{e_\mathrm{g},e_\mathrm{f}} = \boldsymbol{\mu}_\mathrm{g}^0 + \dot{\boldsymbol{\mu}}_\mathrm{g}^T \boldsymbol{Q}_\mathrm{g} + ... \tag{20}$$

where $\boldsymbol{\mu}_\mathrm{g}^0 = \boldsymbol{\mu}_{e_\mathrm{g},e_\mathrm{f}}$ at $\boldsymbol{Q}_\mathrm{g} = 0$ (at the initial state equilibrium geometry $\boldsymbol{x}_\mathrm{g}$), and $\dot{\boldsymbol{\mu}}_\mathrm{g}^j = (\partial\mu_\mathrm{g}\partial Q_\mathrm{g}^j)_{(\boldsymbol{Q}_\mathrm{g}=0)}$. If the electronic transition dipole at the equilibrium geometry of the initial state ($\boldsymbol{\mu}_\mathrm{g}^0$) is large, then higher-order terms of Eq. 20 can be neglected. This is known as FC approximation. However, when the intensity of the transition is weak, $\boldsymbol{\mu}_\mathrm{g}^0$ is small, and the influence of higher-order terms should be taken into account. The first correction to the FC approximation is HT [218, 219], which includes the linear terms of Eq. 20. It can be obtained for 2PA spectrum through the sum-over-state model and two-photon transition matrix elements [183].

As an alternative to the TI formalism (Equation 18), time-dependent (TD) approach can be used for the description of the spectrum [220]. In this case, considering the zero-temperature spectrum, where only the ground vibronic state is populated, the expression of the spectrum intensity is given by the following Fourier transform:

$$S(\omega) = C\omega \frac{1}{2\pi} \int dt \exp\left[i(E_g + \hbar\omega)t\right] \chi(t,T) g'(t) \tag{21}$$

where $g'(t)$ is a damping function that leads to the broadening function, and $\chi(t,T)$ is the transition dipole correlation function. Thus, TD approach lacks information about individual FC factors.

Anharmonic corrections

Now we will briefly summarize the equations used for performing the simulations of FC factors taking into account vibrational anharmonicity. These simulations were performed based on FC Code [221]. Referring to the ground state as "g" and the final state as "e", the Schrödinger equations for nuclear motion are given by:

$$\widehat{H}^g|\psi_{\nu_g}^g\rangle = E_{\nu_g}^g|\psi_{\nu_g}^g\rangle \tag{22}$$

$$\widehat{H}^e|\psi_{\nu_e}^e\rangle = E_{\nu_e}^e|\psi_{\nu_e}^e\rangle \tag{23}$$

In equations above $\widehat{H}^g$, $|\psi^g_{\nu_g}\rangle$ and $E^g_{\nu_g}$ are ground state vibrational Hamiltonian, wave functions, and energies; and $\widehat{H}^e$, $|\psi^e_{\nu_e}\rangle$ and $E^e_{\nu_e}$ are the electronic excited state vibrational Hamiltonian, wave functions, and energies. Multiplying Eq. 22 by $\langle\psi^e_{\nu_e}|$ and 23 by $\langle\psi^g_{\nu_g}|$ leads to:

$$\langle\psi^e_{\nu_e}|\widehat{H}^g|\psi^g_{\nu_g}\rangle = E^g_{\nu_g} S_{\nu_e\nu_g} \tag{24}$$

$$\langle\psi^g_{\nu_g}|\widehat{H}^e|\psi^e_{\nu_e}\rangle = E^e_{\nu_e} S_{\nu_g\nu_e} \tag{25}$$

Here $S_{\nu_e\nu_g} = S_{\nu_g\nu_e} = \langle\psi^g_{\nu_g}|\psi^e_{\nu_e}\rangle = \langle\psi^e_{\nu_e}|\psi^g_{\nu_g}\rangle$ are FC overlap integrals. Combining equations 24 and 25:

$$\langle\psi^g_{\nu_g}|\widehat{H}^g - \widehat{H}^e|\psi^e_{\nu_e}\rangle = (E^g_{\nu_g} - E^e_{\nu_e}) S_{\nu_g\nu_e} \tag{26}$$

Considering that the vibrational eigenfunctions for the excited electronic state form a complete set, we can write:

$$\langle\psi^g_{\nu_g}|\widehat{H}^g - \widehat{H}^e|\psi^e_{\nu_e}\rangle = \sum_{\mu_e}\langle\psi^g_{\nu_g}|\psi^e_{\mu_e}\rangle\langle\psi^e_{\mu_e}|\widehat{H}^g - \widehat{H}^e|\psi^e_{\nu_e}\rangle = \sum_{\mu_e} S_{\nu_g\mu_e}\langle\psi^e_{\mu_e}|\widehat{H}^g - \widehat{H}^e|\psi^e_{\nu_e}\rangle \tag{27}$$

Therefore, from two latter equations:

$$\sum_{\mu_e} S_{\nu_g\mu_e}(\langle\psi^e_{\mu_e}|\widehat{H}^g - \widehat{H}^e|\psi^e_{\nu_e}\rangle + \delta_{\mu_e\nu_e}(E^g_{\nu_g} - E^e_{\nu_e})) = 0 \tag{28}$$

The total nuclear kinetic energy operator $\widehat{T}$ is the same in both Hamiltonians: $\widehat{H}^g = \widehat{T} + \widehat{V}^g$ and $\widehat{H}^e = \widehat{T} + \widehat{V}^e$, that is why:

$$\sum_{\mu_e} S_{\nu_g\mu_e}(\langle\psi^e_{\mu_e}|\widehat{V}^g - \widehat{V}^e|\psi^e_{\nu_e}\rangle + \delta_{\mu_e\nu_e}(E^g_{\nu_g} - E^e_{\nu_e})) = 0 \tag{29}$$

For a given ν_g, the latter equation forms an infinite set of linear equations. At this step one needs an algorithm to truncate to a finite set of equations including only the FC factors larger than a given threshold (see [221]). Then, the surviving M terms given by Eq. 29 are divided by $S_{\nu_g\lambda_e}$ (λ_e is arbitrary as long as $S_{\nu_g\lambda_e} \neq 0$):

$$\sum_{\mu_e}^{M} r^{\mu_e}_{\lambda_e}(\langle\psi^e_{\mu_e}|\widehat{V}^g - \widehat{V}^e|\psi^e_{\nu_e}\rangle + \delta_{\mu_e\nu_e}(E^g_{\nu_g} - E^e_{\nu_e})) = 0, \quad \{\nu_e\} = \{\mu_e\}, \forall\nu_e \tag{30}$$

$$r^{\mu_e}_{\lambda_e} = S_{\nu_g\mu_e}/S_{\nu_g\lambda_e} \tag{31}$$

There are M linear equations, but only $M-1$ unknown ratios, and then one equation is redundant. Any equation can be removed, and the authors of this method chose the remove the equation corresponding to ν_e=λ_e. Solving the previous system of equations, one will determine the ratios $r^{\mu_e}_{\lambda_g}$. To obtain FC integrals $S_{\nu_g\mu_e}$ one can use the normalization condition:

$$\sum_{\mu_e}^{M} S^2_{\nu_g\mu_e} = \sum_{\mu_e}^{M}\langle\psi^g_{\nu_g}|\psi^e_{\mu_e}\rangle\langle\psi^e_{\mu_e}|\psi^g_{\nu_g}\rangle = 1 \tag{32}$$

Dividing the previous equation by $S^2_{\nu_g\lambda_e}$ leads to the following equation for $S_{\nu_g\lambda_e}$ integral:

$$S_{\nu_g\lambda_e} = \frac{1}{\sqrt{\sum_{\mu_e}^{M}(r^{\mu_e}_{\lambda_e})^2}} \tag{33}$$

The remaining FC integrals can be obtained as:

$$S_{\nu_g\mu_e} = r^{\mu_e}_{\lambda_e} S_{\nu_g\lambda_e} \tag{34}$$

To introduce mechanical anharmonicity, one can use the perturbation theory using the harmonic oscillator Hamiltonian as zero-order Hamiltonian 29 [222]. The first-order perturbation theory (PT1) is given by:

$$\sum_{\mu_e} S^{(1)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(0)}+\sum_{\mu_e} S^{(0)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(1)}=0 \tag{35}$$

It is again represented by an infinite number of equations, which can truncated to the same finite set of excitations M as used in harmonic approximation. Here cubic terms in V^g and V^e are taken to be first-order. The first-order corrections to the integrals with potential operator difference are given by:

$$\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle^{(1)}=\langle\psi^{e(1)}_{\mu_e}|(\widehat{V}^g-\widehat{V}^e)^{(0)}|\psi^{e(0)}_{\nu_e}\rangle+\langle\psi^{e(0)}_{\mu_e}|(\widehat{V}^g-\widehat{V}^e)^{(1)}|\psi^{e(0)}_{\nu_e}\rangle+\langle\psi^{e(0)}_{\mu_e}|(\widehat{V}^g-\widehat{V}^e)^{(0)}|\psi^{e(1)}_{\nu_e}\rangle \tag{36}$$

where:

$$|\psi^{e(1)}_{\nu_e}\rangle=-\sum_{\mu_e\neq\nu_e}^{M}\frac{\langle\psi^{e(0)}_{\nu_e}|\widehat{V}^{e(1)}|\psi^{e(0)}_{\mu_e}\rangle|\psi^{e(0)}_{\mu_e}\rangle}{E^{e(0)}_{\mu_e}-E^{e(0)}_{\nu_e}} \tag{37}$$

and, including Equation 13 for description of vibronic space:

$$\begin{aligned}(\widehat{V}^g-\widehat{V}^e)^{(1)}=\frac{1}{6}\sum_{i,j,k=1}^{3N-6}\left(\frac{\partial^3V^g}{\partial Q^g_i\partial Q^g_j\partial Q^g_k}\right)_{Q^g=0}\times\\ \times\left[K_iK_jK_k+3K_iK_j\sum_{l=1}^{3N-6}J_{kl}Q^e_l+3K_i\sum_{l,m=1}^{3N-6}J_{jl}J_{km}Q^e_lQ^e_m+\sum_{l,m,n=1}^{3N-6}J_{il}J_{jm}J_{kn}Q^e_lQ^e_mQ^e_n\right]-\\ -\frac{1}{6}\sum_{i,j,k=1}^{3N-6}\left(\frac{\partial^3V^e}{\partial Q^e_i\partial Q^e_j\partial Q^e_k}\right)_{Q^e=0}Q^e_iQ^e_jQ^e_k\end{aligned} \tag{38}$$

The column vector $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ with components $S^{(1)}_{\nu_g\mu_e}$ can be expressed as:

$$\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}=\boldsymbol{S}'^{(1)}_{\boldsymbol{\nu}^g}+\boldsymbol{S}''^{(1)}_{\boldsymbol{\nu}^g}=\boldsymbol{S}'^{(1)}_{\boldsymbol{\nu}^g}+f\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g} \tag{39}$$

where $\boldsymbol{S}'^{(1)}_{\boldsymbol{\nu}^g}$ is the component of $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ orthogonal to $\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g}$, and $\boldsymbol{S}''^{(1)}_{\boldsymbol{\nu}^g}$ is the component of $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ parallel to $\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g}$. The first term on the left side of Equation 35 vanishes if we substitute $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ for $\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g}$. Therefore, $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ is a solution of Equation 35 for any f. Let us choose $f=0$ so that $\boldsymbol{S}^{(1)}_{\boldsymbol{\nu}^g}$ is orthogonal to $\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g}$, and satisfies the first-order normalization conditions:

$$2\sum_{\mu_e}S^{(1)}_{\nu_g\mu_e}S^{(0)}_{\nu_g\mu_e}=2\boldsymbol{S}^{(1)^\dagger}_{\boldsymbol{\nu}^g}\boldsymbol{S}^{(0)}_{\boldsymbol{\nu}^g}=0 \tag{40}$$

Now, solving the set of equations 35, the first-order corrections to the Franck–Condon factor (FCF)s are:

$$F^{(1)}_{\nu_g\mu_e} = 2S^{(0)}_{\nu_g\mu_e}S^{(1)}_{\nu_g\mu_e} \tag{41}$$

The same approach can be performed for second-order perturbation theory (PT2) corrections, solving the following equation:

$$\begin{aligned}\sum_{\mu_e}^{M} S'^{(2)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(0)}+\\ +\sum_{\mu_e}^{M} S''^{(2)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(0)}+\\ +\sum_{\mu_e}^{M} S^{(1)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(1)}+\\ +\sum_{\mu_e}^{M} S^{(0)}_{\nu_g\mu_e}\left[\langle\psi^e_{\mu_e}|\widehat{V}^g-\widehat{V}^e|\psi^e_{\nu_e}\rangle+\delta_{\mu_e\nu_e}(E^g_{\nu_g}-E^e_{\nu_e})\right]^{(2)}=0\end{aligned} \tag{42}$$

In the case of PT2 approximation, $V^{e(2)}$ and $V^{g(2)}$ contain the quartic terms. Repeating an equivalent procedure to PT1 approximation, one will obtain following PT2 corrections to FCFs:

$$F^{(2)}_{\nu_g\mu_e} = 2S^{(0)}_{\nu_g\mu_e}S^{(2)}_{\nu_g\mu_e} + S^{(1)}_{\nu_g\mu_e}S^{(1)}_{\nu_g\mu_e} \tag{43}$$

4 Analysis of the electronic structure of fluorescent dyes

In this section, four series of boron-containing dyes are studied theoretically aiming at the establishment of "structure–property" relations. These investigations were performed to expand the current knowledge about these families of dyes and facilitate further design of new probes for 2PA bioimaging. Analysis was performed based on the combination of the accurate RI-CC2 level of theory with GFSM approach. Thus, performed simulations provide reliable results which are helpful for a a rich understanding of 2PA nature in studied compounds. Studies were performed in collaboration with two experimental groups. The research group of prof. Borys Ośmiałowski (Nicolaus Copernicus University, Toruń) performed the synthesis of studied dyes, and the research group of prof. Marcin Nyk (Wrocław University of Science and Technology, Wrocław) performed the Z-scan measurements of 2PA properties.

The protocol applied for the simulations of 2PA properties of fluorescent dyes is schematically illustrated in Figure 3. Since the experimental data was available in a chloroform solution, the computer simulations also were performed taking into account the chloroform environment. The employed computational protocol benefits from the combination of the accurate RI-CC2 level of theory for the description of the dye (QM part) combined with the EE approach for the discrete description of the solvent environment (MM part). Moreover, the electronic structure calculations performed for multiple snapshots along MD simulations allow us to pinpoint the optimal "medium" atomistic composition of the solvent molecules. A detailed description of this methodology is given below.

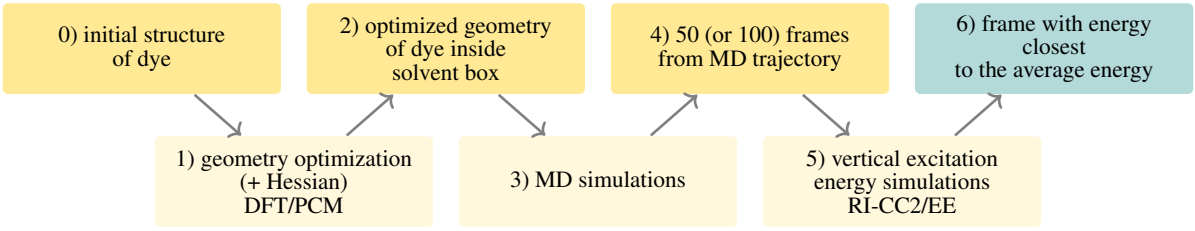


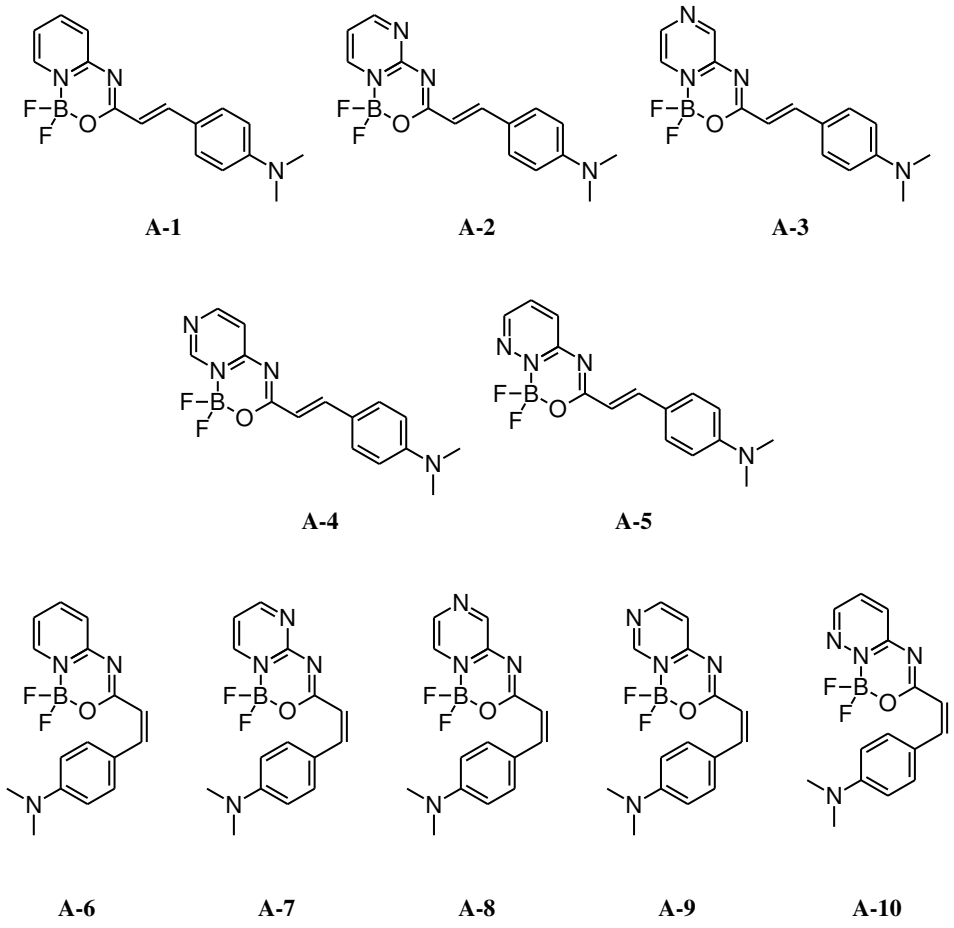
Figure 3: Protocol of simulations with discrete representation of solvent environment

The first stage of the applied protocol consisted of the ground state geometry optimization of the dye with further evaluation of Hessian to confirm that the obtained geometry corresponds to a minimum on the PES. The geometry optimization was performed using DFT for the description of the dye, and the solvent impact was introduced by PCM approach using the integral equation formalism [223]. Thus, the optimization of the ground state geometry did not take into account individual geometries of solvent molecules but modeled the environment as a continuum dielectric medium. Simulations were performed using the Gaussian 16 program [224]. Provided that a reliable ground state geometry of the dye had been obtained, an atomistic model with the discrete representation of solvent molecules was prepared for MD simulations. In this model, the dye was placed in the center of the solvent box, and the number of solvent molecules was selected in a way to preserve density of the solvent at the room temperature. VMD [225] and Packmol [226] programs were used at this step. The third stage of our computational protocol involved MD simulations of the molecules contained in the prepared box. The required force field parameters were prepared using CHARMM force field [227–229]. However, partial charges of atoms were calculated using CHELPG routine [230] implemented in the Gaussian 16 program [224]. The MD simulations started with a minimization followed by an NVT dynamics at 300K using a Langevin thermostat. Periodic boundary conditions were applied at this step. To avoid possible flaws

in the classical force field, the simulations were performed using a frozen geometry for the solute. MD simulations were performed using NAMD program [231]. The main outcome of the third step was the trajectory which contained geometrical configurations of solvent molecules along performed MD simulations. In the fourth and fifth steps, 50 sequential snapshots were taken from the obtained trajectory to perform electronic structure simulations. RI-CC2 [124, 232] was used for the description of the QM part, and the impact of the solvent environment was included by EE approach using the discrete representations of solvent molecules obtained in the MD simulations. Limited by the high computational costs of RI-CC2 approach, the atomic basis set was selected depending on the size of the studied dyes. The augmented basis set aug-cc-pVDZ was applied for small dyes [233], and cc-pVDZ was used for medium-size compounds [234]. However, a significant deterioration of the results obtained using cc-pVDZ is not expected as the previous study [129] demonstrated that the difference in 2PA transition strength calculated by RI-CC2/cc-pVDZ and RI-CC2/aug-cc-pVDZ does not exceed 5%. The described RI-CC2/EE electronic structure calculations were performed using TURBOMOLE v 7.3 [235]. For each dye, the average excitation energy was calculated from computed vertical excitation energies for the set of snapshots. The snapshot corresponding to the energy closest to the average energy was selected as the optimal atomistic model representing the studied system. Based on this snapshot 2PA cross section was calculated assuming Lorentzian broadening equal to 0.2 eV. Even though this value does not accurately reproduce all experimental bandwidths, it was used for the sake of comparison of 2PA efficiency among all studied sets. Moreover, in further steps, the "average" snapshot was used for calculations of dipole and transition moments (using RI-CC2) in order to perform further analysis by the GFSM model (using 2SM and 3SM variants). To sum up, based on the discussed protocol it is possible to estimate excitation energies and 2PA strengths of fluorescent probes in solution and to pinpoint electronic-structure parameters responsible for changes of 2PA activities.

We will now discuss the strategies which were applied in this thesis for the design of new dyes. The effective starting point in the design of difluoroborane dyes for 2PA bioimaging is the selection of the electron-accepting BF_2-containing core structure, which will be "dressed" with electron-donating substituents in further steps. In particular, it is important to analyze how small modifications of the electron-accepting difluroborane group would influence 2PA activity. A recent study demonstrated the enhancement of fluorescent properties in two diazine-based BF_2 compounds compared to the corresponding pyridine-based BF_2 compound [236]. However, the mentioned work described only two possible positions of the nitrogen atom (amongst 4 available), and the two-photon properties of these dyes were not investigated. The observed improvement of fluorescent properties inspired our analysis of how additional nitrogen atom affects 2PA properties of structures with the N–BF_2–O group, studying all 4 possible isomers of benzannulated difluoroborane β-ketoiminates. The first studied set of compounds is presented in Scheme 6. The investigated dyes have a dipolar skeleton, with the N–BF_2–O fragment as the electron-accepting group, and the 4-(*N,N*-dimethylamino)phenyl group as the electron-donating group. These two essential groups are joined by the vinylene linker. The compound **A-1** (which does not possess an additional nitrogen atom) was studied previously, and it demonstrated properties promising for 2PA bioimaging, e.g., in the first biological window 2PA cross section of this structure is equal to 109 GM [109]. Therefore, this structure has the potential for further improvements. Investigated in the same work, the analogue of the structure **A-1**, which lacks the double bond between electron-donating and electron-accepting groups, has more than two times lower 2PA cross section (47 GM). Thus, this double bond is indeed substantial for these dyes. However, the significant disadvantage of the double bond is a possible *cis-trans* photoisomerization. Photoisomerization is an interesting feature

for photochromic applications, but for 2PA bioimaging, this is an undesirable phenomenon as it causes a decrease in the brightness of imaging. For this reason, the set of studied dyes included both *cis*- and *trans*-isomers.



Scheme 6: Studied compounds of series **A**

The theoretical part of this study aimed to support experimentally obtained results with more data which is not feasible to access experimentally. To this end, the excitation energies and 2PA transition strengths were calculated, and state-of-the-art theory was applied to have a deeper understanding of the nature of these dyes. The earlier described computational protocol (Figure 3) was used for this analysis. The optimization of ground state geometries was performed using M06-2X functional [237] combined with 6-31G(d) basis set [238–240] in chloroform solution included through the PCM level of theory. The M06-2X functional shows a good agreement with the experiment for the description of fluorescent dyes containing a tetracoordinated boron atom. In particular, for studied compounds, M06-2X is one of the best functionals for geometry optimizations due to its good description of medium-range electron correlations [241–245]. For the series **A** MD simulations were performed in the solvent box with a side of 40 Å and containinig roughly 470 chloroform molecules. Considered dyes contain up to 40 atoms, therefore, the more accurate aug-cc-pVDZ atomic basis set [233] is computationally accessible, and further electronic structure simulations were performed using RI-CC2/aug-cc-pVDZ/EE level of theory.

Before the discussion of 2PA properties, it is important to focus on the *cis-trans* isomerization. The synthetic

path applied by the experimental group assumed the presence of the *trans*-isomers. Figure 4 illustrates experimentally obtained normalized 1PA spectra of *trans*-isomers **A-1** – **A-5**. Both initial excitation spectra and spectra after pre-irradiation were studied here. The spectrum of the compound **A-1** undergoes significant changes, which can be explained by photoisomerization or decomposition. Note that spectra of the structures **A-2** – **A-5** (compounds with an additional nitrogen atom) remain unchanged. To rationalize the lack of change in the spectra there are two possible justifications: either *i)* spectra of *cis*- and *trans*-isomers are very similar, or *ii)* the compounds **A-2** – **A-5** are photostable and do not undergo structural transformations under irradiation. The computational chemistry helped to shed some light on this phenomenon. Table 1 includes information about experimental and theoretical excitation wavelengths for the $S_0 \rightarrow S_1$ transition. In all pairs of complementary *cis-trans* isomers, the difference in simulated absorption wavelengths is roughly 30 nm, which would be visible in experimental spectra. For compounds **A-1** – **A-5**, the theoretical absorption wavelengths underestimate experimental values by 50 nm, but this is a typical behavior of RI-CC2 [246]. The reason is the approximate treatment of electron correlation effects of the RI-CC2 approach. At the same time, simulated absorption wavelengths precisely follow experimental trends. Therefore, the performed electronic structure calculations are reliable, and the difference between simulated vertical excitation energies for *cis*- and *trans*-form is trustworthy. The second confirmation of this notion is based on the comparison between absorption wavelengths of the structures **A-1** and **A-6**. According to the simulations vertical excitation energy of the *cis*-isomer is red-shifted in contrast with the *trans*-form, and indeed in experimental spectra of **A-1** after irradiation there is a not very bright red-shifted local maximum. Thus, the synergy between theory and experiment allows us to conclude that the compounds **A-2** – **A-5** do not undergo photoconversion. Therefore, the introduction of the additional nitrogen atom increases the photostability of the studied dyes, which is one of the desired properties for 2PA bioimaging.

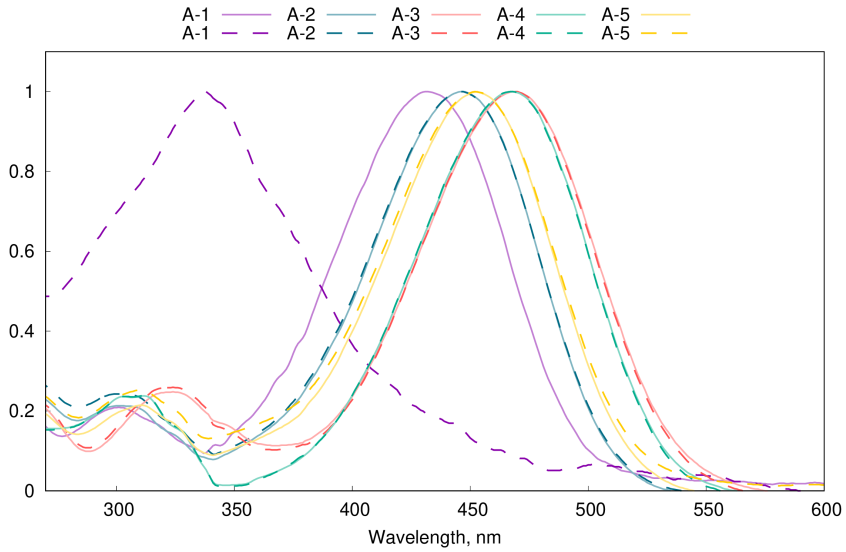


Figure 4: Series **A**, experimental normalized 1PA spectra. The solid line corresponds to the initial spectra, and the dashed line to the spectra after irradiation

Knowing that the *trans*-isomers are photostable, the *cis*-isomers are excluded from further discussion of 2PA properties. For all investigated dyes the absorption wavelength lies in the first biological window, and the introduction of the additional nitrogen atom is responsible for a red-shift of the absorption by 30-50 nm. Thus, one more essential

	experiment						theory					
	ΔE_{abs}, eV	λ_{abs}, nm	λ_{flu}, nm	FQY	σ^{2PA}, GM	2PACS, GM	ΔE_{abs}, eV	λ_{abs}, nm	δ^{2PA}, a.u.	σ^{2PA}, GM	δ_{2SM}, a.u.	δ_{3SM}, a.u.
A-1	2.81	441	527	0.186	109	20	3.15	393	108×10^3	393	139×10^3	139×10^3
A-2	2.65	468	546	0.201	404	81	3.08	403	129×10^3	447	156×10^3	156×10^3
A-3	2.55	487	612	0.023	380	8	2.96	419	138×10^3	443	174×10^3	174×10^3
A-4	2.57	482	564	0.462	367	169	2.97	417	142×10^3	460	208×10^3	210×10^3
A-5	2.63	472	575	0.020	280	6	3.04	407	119×10^3	404	148×10^3	147×10^3
A-6							2.97	417				
A-7							2.83	438				
A-8							2.76	449				
A-9							2.79	444				
A-10							2.83	437				

Table 1: Series **A**, experimental and calculated photophysical properties. One-photon excitation energy (ΔE_{abs}), absorption wavelength (λ_{abs}), emission wavelength (λ_{flu}), fluorescence quantum yield (FQY), 2PA cross section (σ^{2PA}), 2PACS, and response theory two-photon transition strength (δ^{2PA}). Shown are also δ^{2PA} values obtained using two- (δ_{2SM}) and three-state (δ_{3SM}) models. Experimental values for compound **A-1** are from Ref. [109]

requirement of 2PA probes is fulfilled by studied compounds. The spectroscopic 2PA properties of the **A-1** compound were measured in the previous studies. As can be seen in Table 1, the experimental 2PA cross section of the **A-1** structure is 3-4 times lower than that of the compounds **A-2** – **A-5**. The first possible explanation is that indeed the introduction of the additional nitrogen atom causes a great enhancement of the 2PA cross section. However, as it was noted in the Introduction, experimental measurements of 2PA cross sections pose some challenges, and results for the same compound can significantly vary in different experimental measurements. Indeed, the performed simulations demonstrated that the structure **A-1** has the lowest computed 2PA cross section (which is given in Table 1), but the difference between compounds with and without an additional nitrogen atom is less crucial. In an earlier study [109] the authors theoretically investigated 2PA properties of the compound **A-1** at CAM-B3LYP/6-311++G(d,p) level of theory in the gas phase. It turned out that 2PA strength computed by CAM-B3LYP is almost 5 times lower than the value obtained in this work at RI-CC2 level. This result is in line with previous studies which demonstrate that range-separated functionals strongly underestimate the magnitude of the 2PA transition strengths, which emphasizes the importance of using selected RI-CC2 level of theory. Amongst the structures **A-2** – **A-5** theoretically obtained 2PA cross sections have the same magnitude and do not differ more than by a factor of 1.2. The same is observed in the experiment for the compounds **A-2** – **A-5**, where 2PA cross sections vary from 280 GM to 404 GM.

As far as the quality of 2PA probes is described by 2PACS, which is the product of 2PA cross section and FQY, the fluorescent properties of dyes also should be considered to determine the potential of 2PA probes. Interestingly, the FQY of the studied compounds (given in Table 1) vary from 0.020 for the structure **A-5** up to 0.462 for **A-4**. Thus, the nitrogen position is crucial for the 2PACS. While all compounds exhibit large enough and roughly similar 2PA cross sections, small FQY makes structures **A-3** and **A-5** inapplicable for 2PA bioimaging. On the contrary, compounds **A-2** and **A-4** have high enough 2PACS, up to 81 GM and 169 GM respectively. Remarkably, the results reported in the study [236] about the effect of the positions of nitrogen atoms on the FQY do not agree with results obtained for the set **A**. However, the mentioned study described dyes that have one more extra nitrogen atom in the ring with the BF_2 group. Additionally, to understand the nature of 2PA processes in studied dyes, GFSM was applied as 2SM and 3SM approximations. The values of two-photon transition strengths determined based on 2SM and 3SM are given in Table 1. They are in good agreement with the values obtained by the response theory, which confirms the reliability of both investigated models for these dyes. Figure 5 shows the decomposition of 2PA strengths by the 2SM simulations, and it illustrates that decomposed terms undergo significant changes. The term δ_{0111} has the largest absolute value for all compounds, and this term has the greatest variance amongst the studied dyes, namely, from 2.27×10^5 a.u. (**A-1** compound) to 4.35×10^5 a.u. (**A-4** compound). But δ_{0111} is quenched by the negative δ_{0101} term, and therefore the final 2PA strength remains almost unchanged for all five compounds. Analysing the equations delivering the final GFSM transition strengths, namely, to Equation 7 (page 23), the physical explanation of the large δ_{0111} term can be found. This term encompasses the product of the square of $S_0 \rightarrow S_1$ transition moment $|\mu^{01}|^2$ and the square of the dipole moment in the S_1 electronic state $|\mu^{11}|^2$. The corresponding data demonstrated that the variations of the δ_{0111} term are ruled by the values of the S_1 dipole moment $|\mu^{11}|$, and, importantly, all five compounds have large transition dipole moments $|\mu^{01}|$.

To conclude, based on the series **A** it was demonstrated that the introduction of the additional nitrogen atom to the heterocyclic ring fused with the difluoroborane β-ketoiminate group has a high potential in the design of 2PA probes

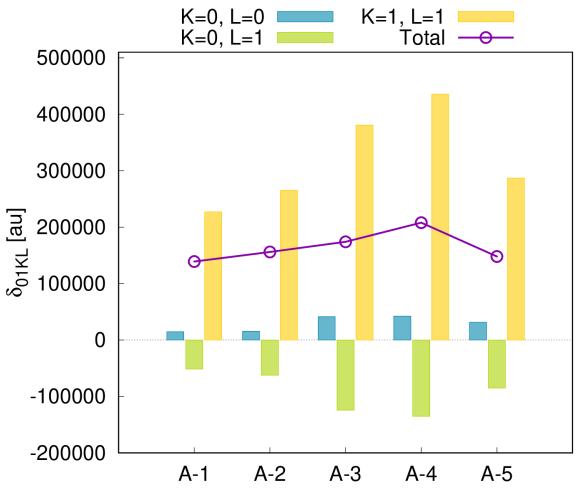
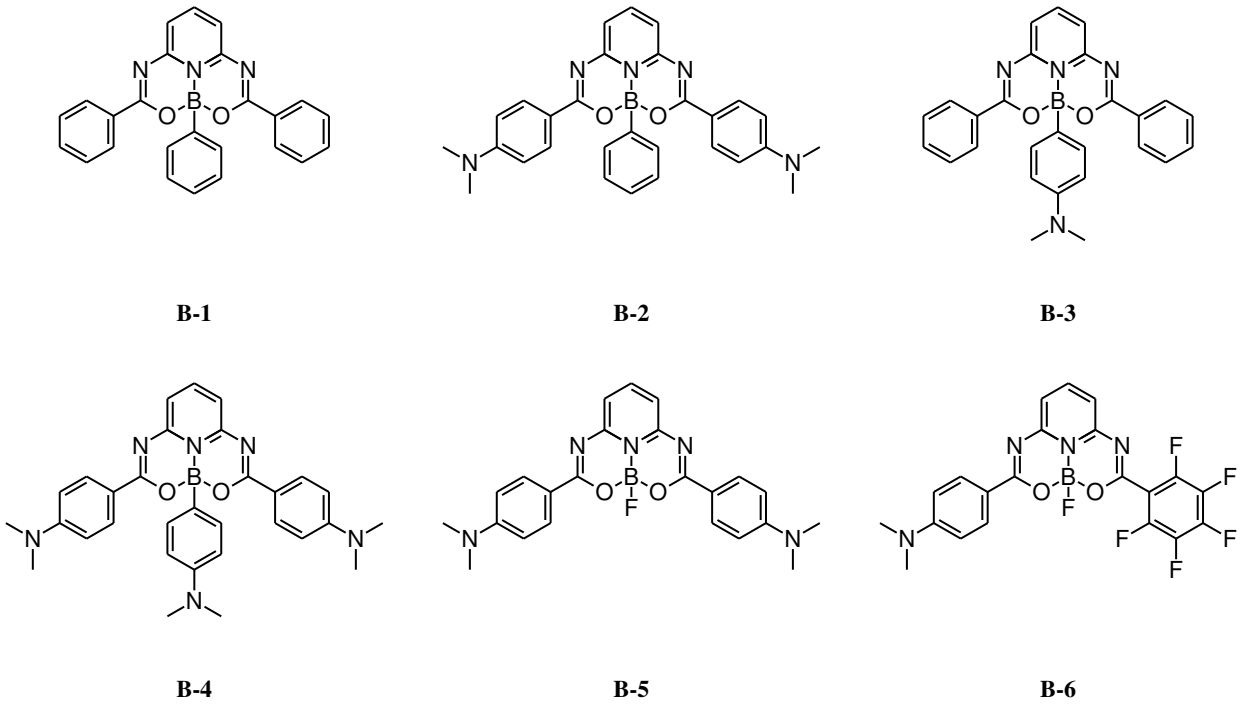


Figure 5: Series **A**, summary of 2SM calculations corresponding to the two-photon $S_0 \rightarrow S_1$ transition

for bioimaging. Firstly, all compounds with the complementary nitrogen atom are photostable in contrast with the molecule **A-1** which lacks the additional nitrogen atom. Secondly, 2PA of all investigated compounds lies in the first biological window, and the introduction of the additional nitrogen atom helps to red-shift excitation wavelength. The third advantage of the additional nitrogen atom is the enhancement of 2PA magnitude. Finally, it was demonstrated that the FQY can be controlled by the placement of the nitrogen atom. It is worth mentioning that two compounds (**A-2** and **A-5**) have high enough 2PACS for further application in 2PA bioimaging.

In what follows we will analyze the influence of the substituent at the boron atom on 2PA cross sections. The studied series **B** is illustrated in Scheme 7. The crystal structures of the compounds **B-1**, **B-3** and **B-6** were reported recently [247]. In these structures, the acceptor groups are almost planar symmetrical π-extended groups, where the boron atom is coordinated by two oxygen and one nitrogen atom. One variable substituent, which is perpendicular to the plane of the central group, is attached to the boron atom. This strategy was used before in several studies [248–251]. Due to the perpendicularity of this substituent, it does not participate in the π-conjugation of the whole molecule. Varying this substituent the Lewis-acidity of the boron atom moiety can be tuned. Additionally, for these dyes, the effect of two peripheral substituents is studied. Thus, compounds containing one (**B-6**), two (**B-5**) and three (**B-1** — **B-4**) electron-donating groups of different strength are investigated.

Let us now discuss the observed photophysical benefits from modifications of the substituents at the boron atom, and, in parallel, to examine the effect of several peripheral substituents. The basic initial prediction proposed is that the implementation of strong electron-donating groups (as $–C_6H_4–NMe_2$ in the compounds **B-2**, **B-4**, **B-5** and **B-6**) would lead to enhancing the 2PA magnitude. Since for this set of dyes experimental data was available in chloroform solvent, the simulations were performed also assuming chloroform environment using the already described protocol presented in Figure 3. Considering that the studied compounds are extended π-conjugated molecules and range-separated functionals well describe ground-state geometries of such compounds (in particular bond length alternation) [252], the CAM-B3LYP functional with cc-pVDZ basis set [234] were chosen for ground-state geometry optimization and evaluation of Hessian [252]. PCM [223] was applied to include the influence of solvent. It should be mentioned that geometries obtained using CAM-B3LYP match crystallographic data very well. The studied molecules **B-1** – **B-6** are



Scheme 7: Studied compounds of series **B**

larger than compounds from the set **A**, and therefore the solvent box prepared for MD simulations was also bigger, with the edge equal to 50 Å. After the MD simulations, one- and two-photon properties were computed for 50 snapshots/dye from the trajectory using RI-CC2/EE approach. Due to the significant size of dyes, in this study, the smaller cc-pVDZ basis set was employed. As was highlighted before, the reduction of the basis set size is not expected to noticeably affect the final results.

Average results obtained by simulations for the $S_0 \rightarrow S_1$, $S_0 \rightarrow S_2$, and $S_0 \rightarrow S_3$ transitions are given in Table 2. From these results it is possible to divide the studied dyes into three subsets according to their photophysical properties:

- the compounds **B-1** and **B-3**, which have the lowest 2PA strengths and blue-shifted absorption wavelengths;
- the compounds **B-2**, **B-4**, and **B-5**, which present the largest 2PA strengths and red-shifted absorption wavelengths;
- the compound **B-6**, which has a reasonable 2PA strength and a medium absorption wavelength.

For all dyes, the calculated excitation wavelengths corresponding to the $S_0 \rightarrow S_1$ transition lay in the first biological window and, therefore, show potential for the 2PA-based imaging. The largest 2PA cross sections are observed for the $S_0 \rightarrow S_2$ and $S_0 \rightarrow S_3$ transitions, and their computed excitation wavelengths are below the transparency range. However, assuming that RI-CC2 tends to underestimate 2PA wavelengths by roughly 50 nm, which was confirmed for the series **A**, it can be concluded that for all states of our interest excitation wavelength fits the desirable biological window. The highest absorption wavelength is observed for the compounds **B-2**, **B-4**, and **B-5**, where peripheral substituents are represented by the strong electron-donating group.

	ΔE, eV	λ, nm	f	$\delta^{2\mathrm{PA}}$, a.u.	$\sigma^{2\mathrm{PA}}$ GM
	$S_0 \rightarrow S_1$				
B-1	3.47	357	0.36	0.0	0
B-2	3.17	391	0.96	7.7×10^3	42
B-3	3.32	373	0.11	0.4×10^3	2
B-4	3.19	389	0.89	6.6×10^3	36
B-5	3.22	385	1.08	8.7×10^3	49
B-6	3.30	376	0.88	26.2×10^3	153
	$S_0 \rightarrow S_2$				
B-1	4.48	277	0.48	0.2×10^3	1
B-2	3.76	330	0.67	84.7×10^3	438
B-3	3.60	344	0.27	0.2×10^3	1
B-4	3.61	343	0.19	13.6×10^3	65
B-5	3.79	327	0.74	84.9×10^3	447
B-6	3.94	315	0.41	25.8×10^3	146
	$S_0 \rightarrow S_3$				
B-1	4.58	271	0.01	0.1×10^3	0
B-2	4.05	306	0.47	24.9×10^3	150
B-3	3.96	313	0.09	3.9×10^3	22
B-4	3.79	327	0.59	67.1×10^3	353
B-5	4.11	302	0.39	23.6×10^3	146
B-6	4.31	288	0.07	30.4×10^3	206

Table 2: Series **B**, summary of electronic-structure calculations. One-photon excitation energies (ΔE) and wavelengths (λ), oscillator strengths (f), and two-photon transition strengths ($\delta^{2\mathrm{PA}}$)

To facilitate the analysis of the 2PA properties, Figure 6 illustrates the calculated two-photon transition strengths for all six compounds. The compounds **B-1** and **B-3** have two phenyl groups (weak, almost neutral, electron-donating groups) as peripheral substituents, and it results in the lowest 2PA strengths for all transitions. The introduction of two strong peripheral electron-donating groups (–C_6H_4–NMe_2) increased 2PA strength of the compounds **B-2**, **B-4** and **B-5** more than by a factor of 10. The compounds **B-2** (weak, almost neutral substituent at boron atom) and **B-5** (electron-accepting substituent at boron atom) have almost the same 2PA strengths for all three considered transitions, with the largest 2PA strength for the $S_0 \rightarrow S_2$ transition. But the compound **B-4**, which has strong electron-donating groups in all three vacant positions, presented the highest 2PA strength for the $S_0 \rightarrow S_3$ transition. The same is observed for the compound **B-3**, which also has –C_6H_4–NMe_2 group at the boron atom. The compound **B-6** is the only dipolar structure presented in the set **B**, and it has approximately the same 2PA strengths for all three considered transitions. Its 2PA strength is lower than the one for the compounds **B-2**, **B-4**, and **B-5** by a factor of 4, but still suitable for 2PA-based

bioimaging.

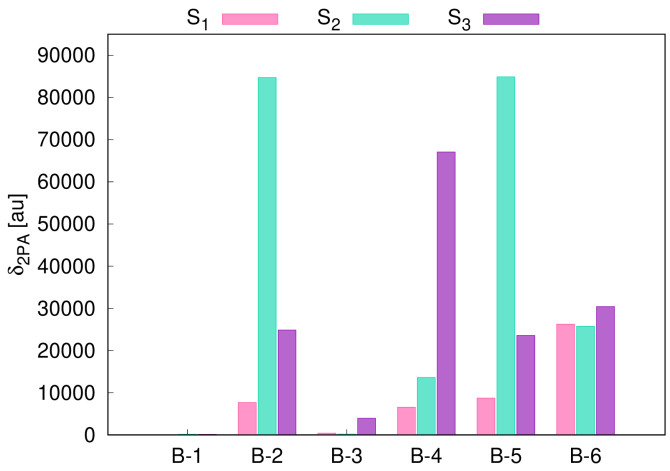


Figure 6: Series **B**, two-photon transition strengths calculated using the RI-CC2 method

Once the computer simulations showed that the proposed set has indeed interesting and promising 2PA properties, experimental studies were performed. Figure 7 illustrates 1PA spectra obtained experimentally. One finds a good agreement between theoretical results and experimental data, and it is possible to highlight the same three subsets of compounds. The compounds **B-1** and **B-3** have a similar shape of 1PA spectra, which is blue-shifted, and the $S_0 \rightarrow S_2$ transition is more intense than the $S_0 \rightarrow S_1$ transition. The same was predicted theoretically as can be seen in the oscillator strength values presented in Table 2). The compounds **B-2**, **B-4**, and **B-5** have almost identical 1PA spectra, where the absorption maxima are red-shifted compared to the rest of the set. The compound **B-6** has an absorption wavelength lying between the first and the second subsets. Thus, the computational methods allowed us to predict the experimental 1PA spectra with a good accuracy. All computated patterns were confirmed by the experimental spectra. At this stage, we can make the following important "structure–property" conclusions: *i)* 1PA spectra do not depend on the orthogonal substituent at the boron atom, and *ii)* strong electron-donating substituents in classic π-conjugated peripheral positions lead to a red-shift of the absorption maxima.

Table 3 provides the summary of experimentally obtained photophysical properties for the series **B**. It confirms again that the absorption wavelength obtained by RI-CC2 underestimates the experimental wavelengths by 40-50 nm. The agreement between theory and experiment for 1PA spectra proved the reliability of performed electronic structure calculations and allowed the subsequent estimation of 2PA spectra for the dyes **B-2**, **B-4**, **B-5**, and **B-6**. The two-photon absorption spectra of these compounds were recorded experimentally using the Z-scan technique. The data presented in Table 3 confirms our initial prediction. Amongst the 4 promising compounds, the dye **B-6** has the lowest 2PA cross section as predicted by calculations. The dyes **B-2** and **B-5** have approximately the same 2PA cross sections, as was also shown by the theoretical studies. On the other hand, the compound **B-4**, which presents a strong electron-donating substituent at the boron atom, has 2PA cross section twice higher than the compounds with electron-accepting or neutral substituents at the boron atom (dyes **B-2** and **B-5**). One more important feature of the 2PA probes is the FQY characteristics, which are also given in Table 3. The compounds **B-2**, **B-4** and **B-5** have high enough FQY, while the corresponding value for the dipolar compound **B-6** is very low. The FQY for the compound **B-4** (with strong

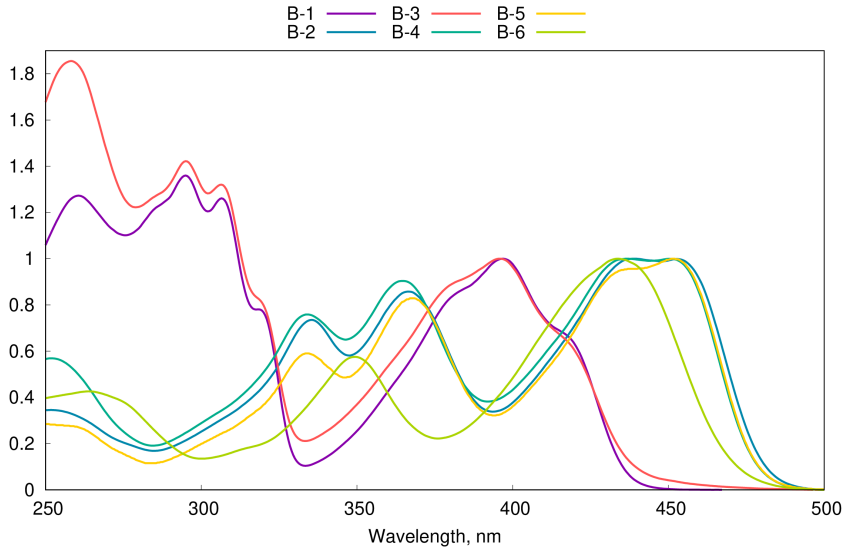


Figure 7: Series **B**, experimental normalized 1PA spectra

electron-donating group at boron atom) is lower than the one for the compounds **B-2** and **B-5** by a factor of 4. Thus, unfortunately, the beneficial value of 2PA cross section for this structure is not paralleled by similarly satisfactory FQY value. The 2PACS (the final value of our interest for bioimaging) of the dye **B-4** is twice smaller than the one for **B-2** and **B-5**.

	1PA				2PA	
	ΔE_{abs}, eV	λ_{abs}, nm	λ_{flu}, nm	FQY	$\lambda_{\text{abs}}^{\text{2PA}}$, nm	σ^{2PA}, GM
B-1	3.12	397	466	0.186		
B-2	2.82	439	485	0.446	725	589
B-3	3.14	395	467	0.015		
B-4	2.84	437	489	0.106	725	1120
B-5	2.74	452	484	0.385	725	651
B-6	2.86	434	525	0.041	800	137

Table 3: Series **B**, experimentally obtained photophysical properties. One-photon excitation energy (ΔE_{abs}), absorption wavelength (λ_{abs}), emission wavelength (λ_{flu}), fluorescence quantum yield (FQY), 2PA cross section (σ^{2PA})

Figure 8 represents normalized 1PA and 2PA spectra. In these plots, the six points denote the results of calculations for 1PA and 2PA spectra and correspond to the vertical $S_0 \rightarrow S_1$ (two right points), $S_0 \rightarrow S_2$ (two middle points), and $S_0 \rightarrow S_3$ (two left points) excitations. Even though the applied computational protocol did not account for the vibronic structure of dyes, we can see that the relative intensities of all peaks were calculated with excellent accuracy. Both theory and experimental results show that for the compounds **B-2** and **B-5** the $S_0 \rightarrow S_2$ transition is the most 2PA-intense, while for the compound **B-4** the $S_0 \rightarrow S_3$ transition is the most bright in the 2PA spectrum, and the structure

B-6 has precisely the same 2PA cross section for all the three investigated transitions.

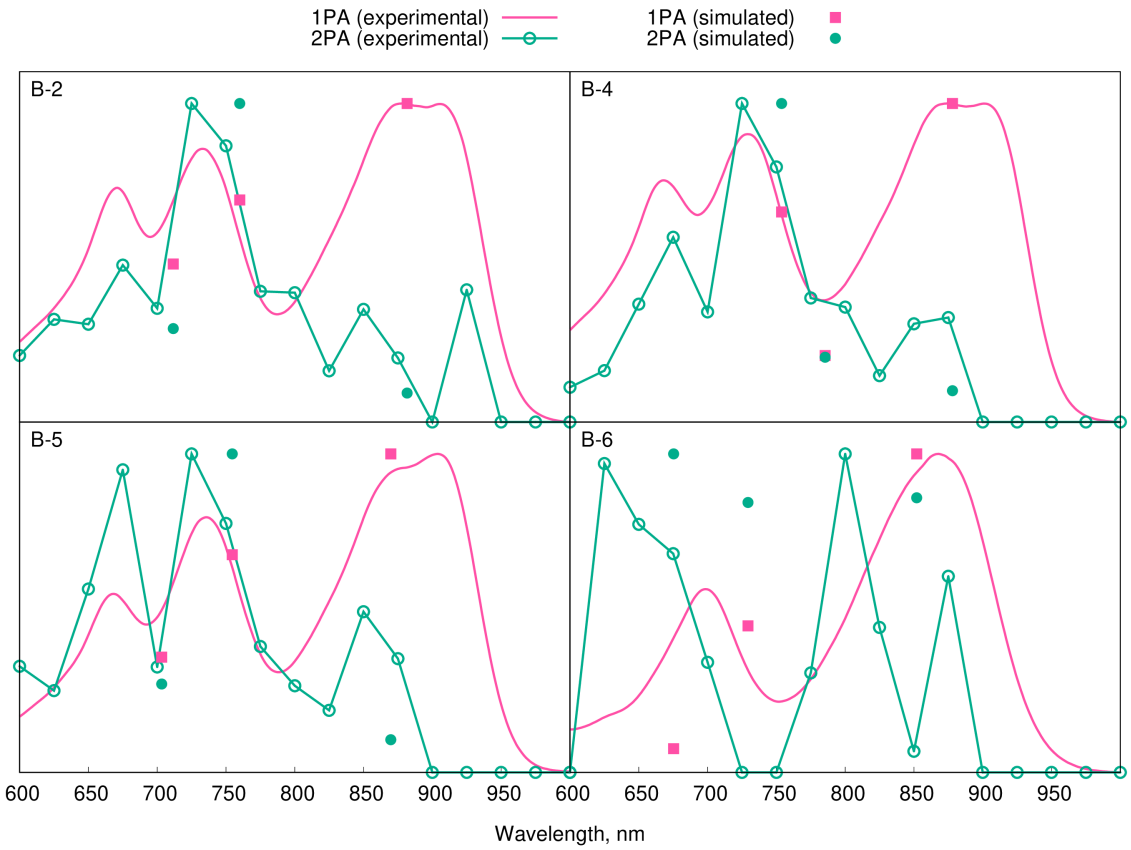


Figure 8: Series **B**, comparison between experimental and theoretical normalized 1PA and 2PA spectra in chloroform. One-photon intensity is plotted against the wavelength multiplied by a factor of two. The theoretical wavelengths were red-shifted by 50 nm (in one-photon spectra values) to match the experimental features. A green line connecting the experimental points (cubic spline) is used to guide the eyes

Table 4 contains a summary of the obtained results, which shows that strong electron-donating group impact in 2PA bioimaging properties is essential. The strong electron-donating groups in both studied placements (at boron atom and in classical peripheral position) are beneficial for the 2PA magnitude, as demonstrated by the quadrupolar structures **B-2**, **B-4** and **B-5**, which have a large 2PA cross sections in 600–1100 GM range. Nevertheless, the introduction of this group at the boron atom causes the decrease of FQY by a factor of 4, and therefore it should be avoided if high brightness is required. Importantly, by changing the character of the substituent at the boron atom it is possible to tune 2PA properties keeping the 1PA spectra intact. The dipolar structure **B-6** is the only compound where the high 2PA cross section is similar in magnitude for all three considered transitions. However, it also exhibits a weak emission, making it inapplicable for bioimaging. Furthermore, it should be highlighted that theoretically performed analyses satisfactorily match the experimental results, confirming that the combination RI-CC2/EE provides reliable 1PA and 2PA properties. Three of the compounds presented in this work (**B-2**, **B-4**, and **B-5**) have the potential for further usage in 2PA imaging techniques.

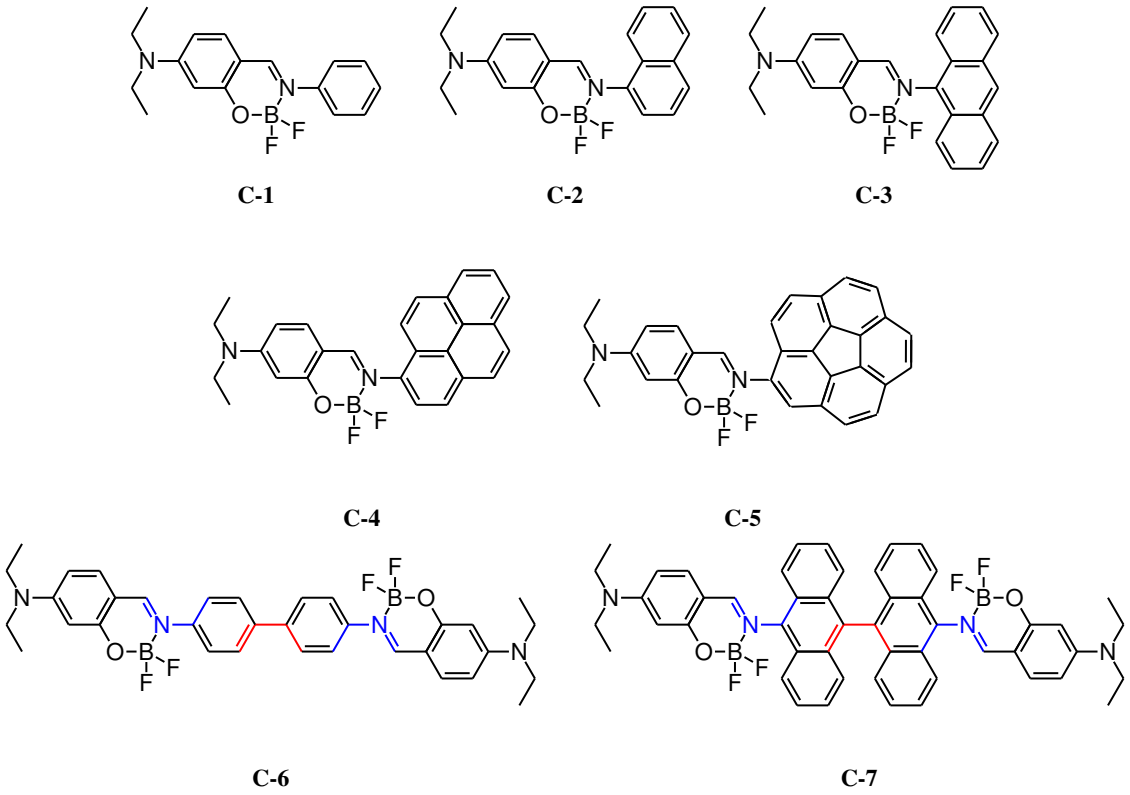
The discussion of the series **A** and **B** was focused on the influence of electron-accepting group modifications on 2PA

property	–C_6H_4–NMe_2 at boron	peripheral –C_6H_4–NMe_2
excitation energy	remains unchanged	red-shifted
1PA	remains unchanged	higher intensity of the first absorption band
2PA	higher 2PA cross section	higher 2PA cross section
emission	lower FQY	higher FQY

Table 4: Impact of the –C_6H_4–NMe_2 substituent in different positions on optical properties of series **B**

properties. Studies of the compounds **B-2**, **B-4**, and **B-5** demonstrated that, unfortunately, in some cases, compounds with large 2PA cross sections have small FQYs. One of the known ways to tune FQY is to attach polyaromatic hydrocarbon (PAH), as in the set **C**, which is illustrated in Scheme 8. In the study [253] we investigated fluorescence properties of the compounds **C-1** – **C-5**. The study in question demonstrated that an increase of PAH size leads to higher FQY, because the smaller flexibility of the rotation around the C—N bond improves the efficiency of the emission. The study [254] demonstrated that asymmetric quadrupolar dye with the same core as **C-1** – **C-5** structures, but with a triphenylamine group instead of the PAH substituent at the nitrogen atom, reaches 2PA cross section equal to 198 GM. The FQY of that dye is roughly 0.1, while the FQY of the compound **C-5** is equal to 0.39 in chloroform. For this reason, it is interesting to assess how the PAH substituents affect 2PA properties of dyes, which will be the first aspect of the discussion. The compounds **C-1** – **C-5** are dipolar dyes with the electron-accepting N–BF_2–O group and the strong electron-donating diethylamino group (the same principles were used in the design of the set **A**). Moreover, the symmetric quadrupolar structures **C-6** and **C-7** were included in this discussion, because as it was mentioned in the Introduction and demonstrated for the set **B**, quadrupolar dyes have 2PACS values larger than dipolar structures.

Calculations in chloroform solution were performed using the protocol described earlier (illustrated in Figure 3). The ground state geometries of the studied compounds were optimized using the B3LYP functional and the 6-31G(d,p) basis set [238–240]. The significant size of the studied compounds (up to 104 atoms) was considered in the choice of the functional. The impact of the solvent was included with the PCM [223] solvation approach. At this step the geometry optimization of the compounds **C-6** and **C-7** was performed within the C_2 point group of symmetry. For each dye, the optimized geometry was placed in the center of atomistic model of chloroform box with the size of $50 \times 50 \times 50$ Å. Subsequently, rigid-body MD simulations were performed, and 100 snapshots were taken from the resulting MD trajectory for further RI-CC2 electronic-structure calculations with EE solvation model for the description of dye-solvent interactions. For the compounds **C-1** – **C-5** the electronic structure calculations were performed using the cc-pVDZ basis set [234], while for the larger compounds **C-6** and **C-7** basis set was limited to the def-SV(P) [255]. To confirm the reliability of the selected basis set, a subset of snapshots for the compounds **C-1** and **C-3** was computed using both def-SV(P) and cc-pVDZ basis sets. The comparison of the computed vertical excitation energies showed that def-SV(P) results overestimate cc-pVDZ values, but the error does not exceed 1%. Additionally, for each investigated structure GFSM analysis was performed for the dye-solvent snapshot with the energy closest to the arithmetic mean value of energy. However, to increase accuracy RI-CC2/cc-pVDZ level of theory was used for all compounds at this



Scheme 8: Studied compounds of series **C**

stage.

The results of electronic-structure calculations are presented in Table 5 where, additionally, experimental absorption maxima are given in brackets. By and large, the calculated vertical excitation energies are in good agreement with the experiment except for the **C-4** structure. As it was shown for the series **A** and **B**, the applied computational protocol leads to underestimated absorption wavelengths by roughly 40-50 nm. It is worth mentioning that for all studied dyes the excitation energy lies in the first biological window. This property is very important to use these compounds as 2PA probes for bioimaging, especially taking into account the high fluorescence of these dyes.

The values of 2PA cross sections computed based on the response theory are also given in Table 5. Firstly, the quadrupolar compounds **C-6** and **C-7** present characteristic symmetry-related features in electronic structure, namely, the $S_0 \rightarrow S_1$ transition has a low 2PA transition strength because it is symmetry-forbidden, while the $S_0 \rightarrow S_2$ transition presents large 2PA transition strength. The same was observed for the quadrupolar-like (Λ-shaped) dyes **B-2** and **B-5**. On the contrary, unexpectedly, the dipolar dyes **C-1** – **C-5** have low 2PA strengths in both considered transitions. Similarly to the compounds of the series **A**, the molecules **C-1** – **C-5** are dipolar structures including an electron-accepting N–BF_2–O group and a strong electron-donating alkyl-substituted amino group. But 2PA strengths of the dyes from the series **A** are higher than those of the dipolar molecules from the series **C** by a factor of 10. To understand the inner nature of observed behavior, 2SM and 3SM analyses of 2PA magnitudes were performed. The results of 2SM calculations for $S_0 \rightarrow S_1$ are presented in Table 5. The dyes **C-6** and **C-7** are excluded from the discussion of this transition due to their low 2PA cross section values. Total 2SM 2PA strengths are in good agreement with the values

	$S_0 \rightarrow S_1$					$S_0 \rightarrow S_2$				
	ΔE, eV	λ_{abs}, nm	f	δ^{2PA}, a.u.	δ_{2SM}, a.u.	ΔE, eV	λ_{abs}, nm	f	δ^{2PA}, a.u.	δ_{3SM}, a.u.
C-1	3.45	359(398)	1.19	11.2×10^3	9.0×10^3	3.91	317	0.02	5.3×10^3	5.5×10^3
C-2	3.54	350(390)	1.15	10.8×10^3	8.5×10^3	3.86	321	0.02	4.2×10^3	4.7×10^3
C-3	3.51	353(396)	1.18	9.8×10^3	6.9×10^3	3.74	332	0.03	3.5×10^3	4.2×10^3
C-4	3.39	366(395)	1.39	4.3×10^3	0.5×10^3	3.71	334	0.01	4.9×10^3	7.3×10^3
C-5	3.30	376(406)	1.44	10.6×10^3	8.1×10^3	3.57	347	0.00	0.4×10^3	0.1×10^3
C-6	3.25	381(424)	3.16	1.0×10^3		3.45	359	0.03	43.6×10^3	52.1×10^3
C-7	3.38	367(409)	2.50	0.2×10^3		3.52	352	0.01	25.1×10^3	18.4×10^3

Table 5: Series **C**, summary of electronic-structure calculations. One-photon excitation energy (ΔE), absorption wavelength (λ_{abs}), oscillator strength (f), and average response theory two-photon transition strength (δ^{2PA}). δ values obtained using two- (δ_{2SM}) and three-state (δ_{3SM}) models are also shown. Values given in brackets correspond to experimental values

obtained by the response theory (except for compound **C-4**, for which an inconsistency with experimental absorption maxima was observed). Figure 9 illustrates the 2SM decomposition of $S_0 \rightarrow S_1$ transition for the set **C**. To better rationalize the results, we can compare the 2SM decomposition performed for the series **C** with the results for the series **A** (Figure 5). The negative term δ_{0101} in the series **A** and **C** has approximately the same order of magnitude, higher than -100×10^3 a.u. and lower than -20×10^{-3} a.u. While in the series **A** the largest term δ_{0111} is in the range 200×10^3 – 400×10^3 a.u., in the series **C** the largest term δ_{0111} is in the range 20×10^3 – 80×10^3, about 5 times weaker than for the set **A**. The destructive interference of the largest positive term δ_{0111} and the negative term δ_{0101} leads to poor total 2PA strengths for dipolar compounds from the series **C**. According to Equation 7 (page 23), the term δ_{0111} (causing the difference between the **A** and **C** sets) incorporates the product of the square of $S_0 \rightarrow S_1$ transition moment and the square of the dipole moment in the S_1 electronic state. The lower value of this product for the series **C** may be due to a shorter π-path between electron-donating and electron-accepting groups, what results into weaker 2PA strengths.

The 2PA strengths of the $S_0 \rightarrow S_2$ transition were analyzed using the 3SM approach. Amongst several tested intermediate states, for all seven compounds total 3SM values with the intermediate state S_1 are in the best agreement with 2PA strengths calculated by the response theory (see Table 5). Figure 10 illustrates that the total 3SM strength is almost equal to the largest term δ_{0211}. The magnitude of this term is proportional to the square of the transition dipole moment between S_0 and S_1 states and the square of the transition dipole moment between S_1 and S_2 states, which can be linked with the high 2PA strengths for quadrupolar dyes **C-6** and **C-7**, and to the low 2PA magnitude for dipolar molecules **C-1** – **C-5**.

The quadrupolar structures **C-6** and **C-7** deserved further investigation, and therefore, our collaborators performed Z-scan measurements of their 2PA cross sections. The **C-6** compound is not enough soluble, and this explains why experimental data is only available for the dye **C-7**, which has a significant experimental 2PA cross section in chloro-

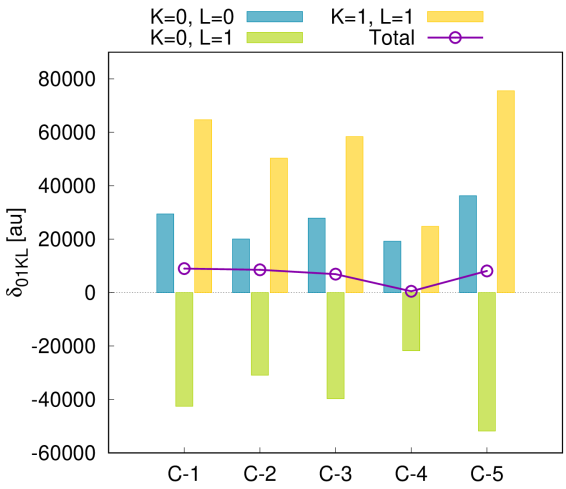


Figure 9: Series **C**, summary of 2SM calculations corresponding to the two-photon $S_0 \rightarrow S_1$ transitions

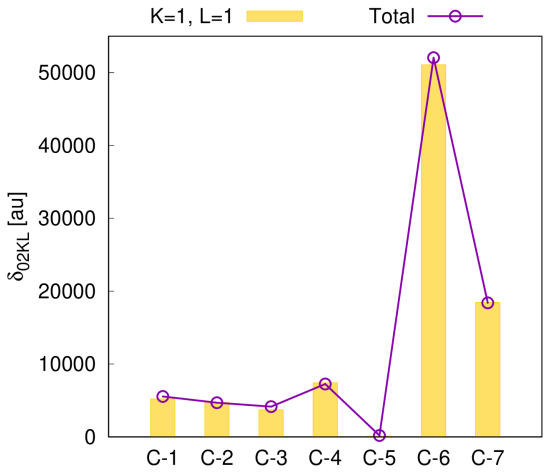
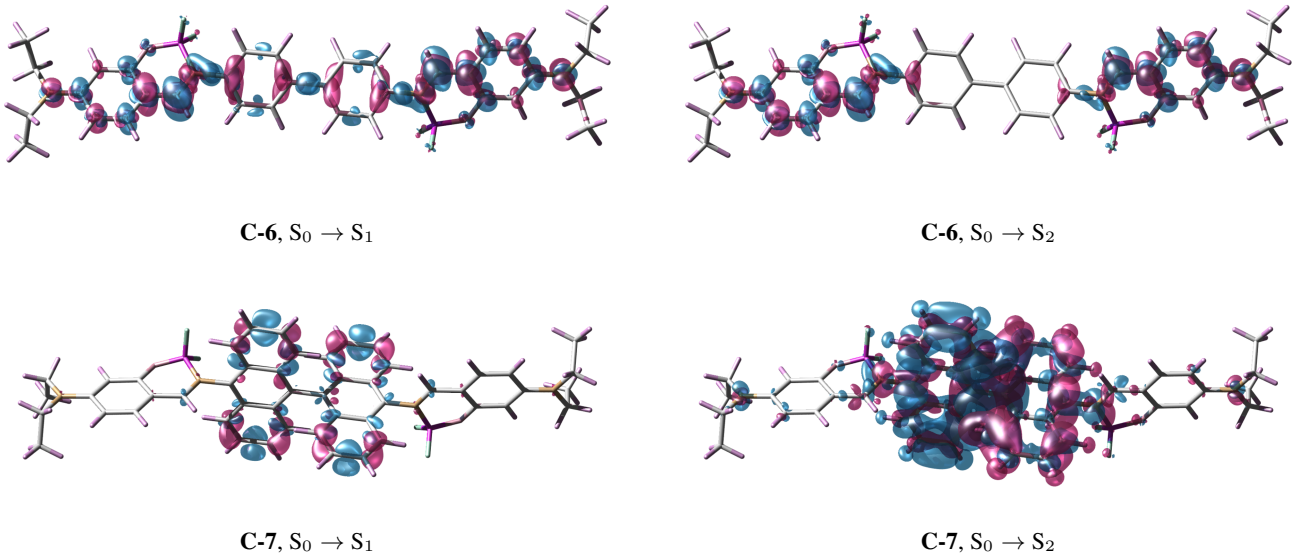


Figure 10: Series **C**, summary of 3SM calculations corresponding to the two-photon $S_0 \rightarrow S_2$ transition

form, 98 GM. The FQY of the dyes **C-6** and **C-7** are 0.58 and 0.21 respectively. Taking into account that simulated 2PA strength of the dye **C-6** is two times larger than that of the dye **C-7**, the compound **C-6** has sufficient 2PACS for 2PA bioimaging. The dyes **C-6** and **C-7** are interesting for exploration of twisting-induced changes of 2PA properties, and this aspect will be analyzed below for the set **C**. The dihedral angle θ, which is interesting for the present analysis, is highlighted with the red color in Scheme 8. This angle is equal to 35.2 deg. and 90.7 deg. in optimized geometries of the compounds **C-6** and **C-7**, respectively, in chloroform solution. To analyze the differences in the electronic structure in similar dyes differing in central arrangement, electronic density difference plots (i.e. $(\Delta\rho(\boldsymbol{r}) = \rho^{S_n}(\boldsymbol{r}) - \rho^{S_0}(\boldsymbol{r}))$) were calculated using "average" snapshots at MN15/cc-pVDZ/EE [234, 256] level of theory (using Gaussian program [224]). The applied level of theory is justified by the recent study that compared the performance of 18 DFAs in simulations of excited state properties of fluorescent dyes (including a big subset of difluoroboranes). In this benchmark study, MN15 outperforms other DFAs [257]. Figure 11 presents resulting plots of $\Delta\rho(\boldsymbol{r})$ for the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions. In the case of the structure **C-7** only the bianthracene moiety is involved in the two considered transitions,

while in the case of the compound **C-6** electron-accepting N–BF_2–O and the electron-donating diethylamino groups also participate in those transitions. These results confirm that by twisting the angle of the central moiety (and C—N bond) the nature of transitions might be significantly tuned.



C-6, $S_0 \rightarrow S_1$ **C-6**, $S_0 \rightarrow S_2$

C-7, $S_0 \rightarrow S_1$ **C-7**, $S_0 \rightarrow S_2$

Figure 11: Compounds **C-6** and **C-7**, electronic density difference plots corresponding to the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions. Presented plots were prepared with the density contour value 0.001

For these two dyes, the twisting around the central bond may lead to changes in the symmetry of the chemical systems. The planar biphenyl (or bianthracene) moiety would lead to C_i symmetry. However, in the optimized geometries, this planarity is not observed due to steric effects caused by hydrogen atoms in bi-arenes. In the case of the compound **C-6**, the planar arrangement of the central biphenyl group is possible, but the distance between hydrogen atoms is small in the planar configuration. On the other hand, for the compound **C-7** planar central moiety is impossible because in this structure hydrogen atoms would simply overlap. Thus, the tracking of the symmetry-relaxation effects on the two-photon transition strengths was accomplished only for the dye **C-6**. It was carried out using the same level of theory, but in the gas phase, which excludes the complicated calculations for including the solvent effects. The optimization of the ground state geometry of the compound **C-6** was performed for the following structural configurations: *i)* assuming C_i symmetry point group, and *ii)* for interplanar angles θ (marked in red in Scheme 8) fixed in the range 0 – 90 deg. with a step of 15 deg. (constrained optimization). These constrained geometries assumed C_1 point group symmetry. Thus, 8 constrained optimized ground state geometries were obtained, for which the 2PA properties were calculated using the response theory at RI-CC2/cc-pVDZ level of theory. Subsequently, the 3SM analysis was performed. Considering that the $S_0 \rightarrow S_1$ transition is not 2PA-bright for the compound **C-6** (as it was confirmed theoretically for all investigated geometries), the further discussion on the dye **C-6** is focused only on the $S_0 \rightarrow S_2$ transition.

The simulated structural parameters are described in Table 6. Firstly, the geometries of the dye **C-6** with C_i point group symmetry and the one with a planar biphenyl group and C_1 point group symmetry have almost the same dihedral angles, which are responsible for the planarity of the dye. The 2PA cross sections for these two geometries calculated by the response theory are also similar, 29.7×10^3 and 30.1×10^3 a.u., respectively. The solvent effects have a

significant role. The 2PA strength calculated taking into account the chloroform environment by EE model is more than 40% higher than the gas-phase value.

	C-6			C-7		
Symmetry	θ	ϕ_1	ϕ_2	θ	ϕ_1	ϕ_2
C_i (gas)	0.0	40.0	40.0	-	-	-
C_1 planar (gas) [constrained opt.]	0.0	39.4	39.4	-	-	-
C_2 (chloroform)	35.2	40.9	40.9	90.7	87.8	87.8
crystal	0.2	43.8	43.8	81.9/80.3	85.7	87.2
				94.8/95.7	74.8	79.9

Table 6: Compounds **C-6** and **C-7**, structural parameters. Dihedral angles θ and dihedral angles ϕ_1 and ϕ_2 (marked with blue in Scheme 8) in the gas phase (simulations), chloroform solution (simulations), and crystal (experiment)

Upon the relaxation of the symmetry, the electronic structure parameters would also change. Figure 12 illustrates the dependence of the 2PA strength and absorption wavelength on the dihedral angle θ. Upon rotation of the dihedral angle θ the excitation wavelength monotonically decreases, but the changes do not exceed the 10 nm. However, the behavior of the 2PA strength calculated by response theory is less expected, since it decreases upon symmetry relaxation for $\theta > 15$ deg. Nevertheless, amongst the studied geometries, the changes in the 2PA strength do not exceed 20%. Thus, rotation around the central bond in biphenyl does not lead to critical changes in the 2PA strength of the $S_0 \rightarrow S_2$ transition. The $S_0 \rightarrow S_1$ transition becomes only weakly 2PA-allowed upon rotation around the central bond for geometry with $\theta = 90$ deg., with an 2PA strength of only 1.3×10^2 a.u.

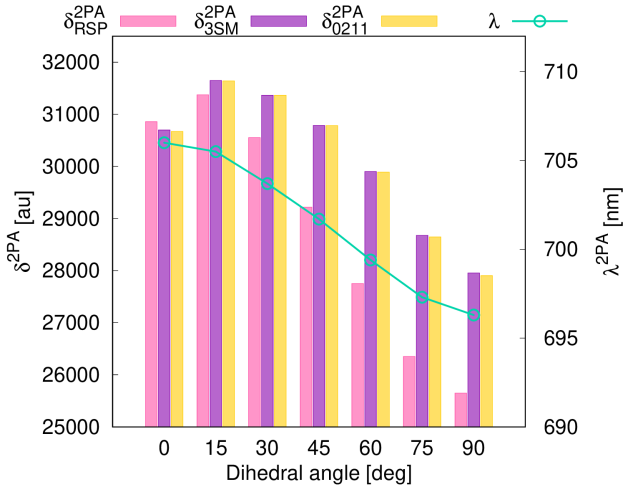


Figure 12: Compound **C-6**, the dependence of two-photon $S_0 \rightarrow S_2$ transition strengths (δ^{2PA}) and two-photon excitation wavelength (λ) on the dihedral angle θ

The two-photon $S_0 \rightarrow S_2$ transition strengths obtained by the response theory and the total 3SM values (with S_1 as

an intermediate state) are in good agreement. Both approaches have local maxima of the 2PA strengths corresponding to the $S_0 \rightarrow S_2$ transition at 15 deg. Therefore, the GFSM analysis performed is reliable. For all geometries, the dominant term is δ_{0211}, which closely follows total 3SM values. According to Equation 7 (page 23) this term is proportional to the following product of transition moments ($|\mu|^{01}|\mu|^{10}|\mu|^{12}|\mu|^{21}$). The product of left and right transition moments, referred to as transition strength (related to oscillator strength), is shown in Figure 13. This figure shows that upon symmetry relaxation the $S_0 \rightarrow S_1$ transition dipole moment $|\mu|^{01}|$ decreases, while the reverse happens to the strength of the $S_0 \rightarrow S_2$ transition. The interplay of S_0/S_1 and S_1/S_2 transition dipole couplings is the reason for observed 2PA strength maxima at 15 deg. Thus, the transition moment dependence on the torsional parameter is a key reason for the 2PA changes for the studied quadrupolar dyes **C-6** and **C-7**.

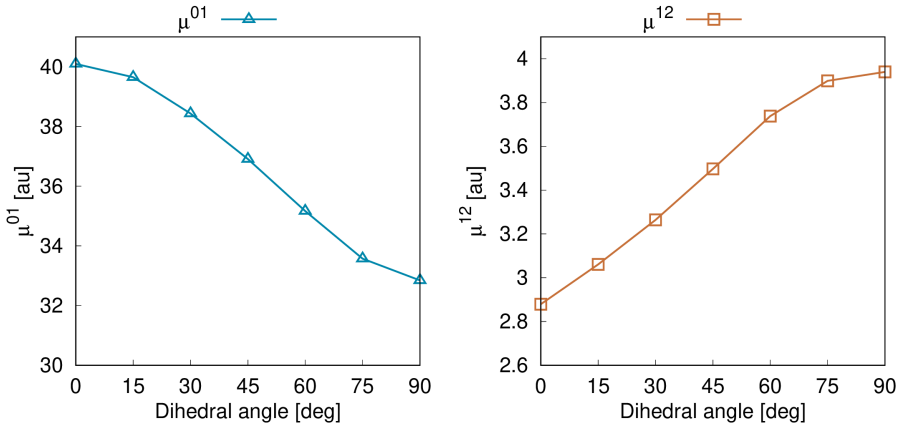
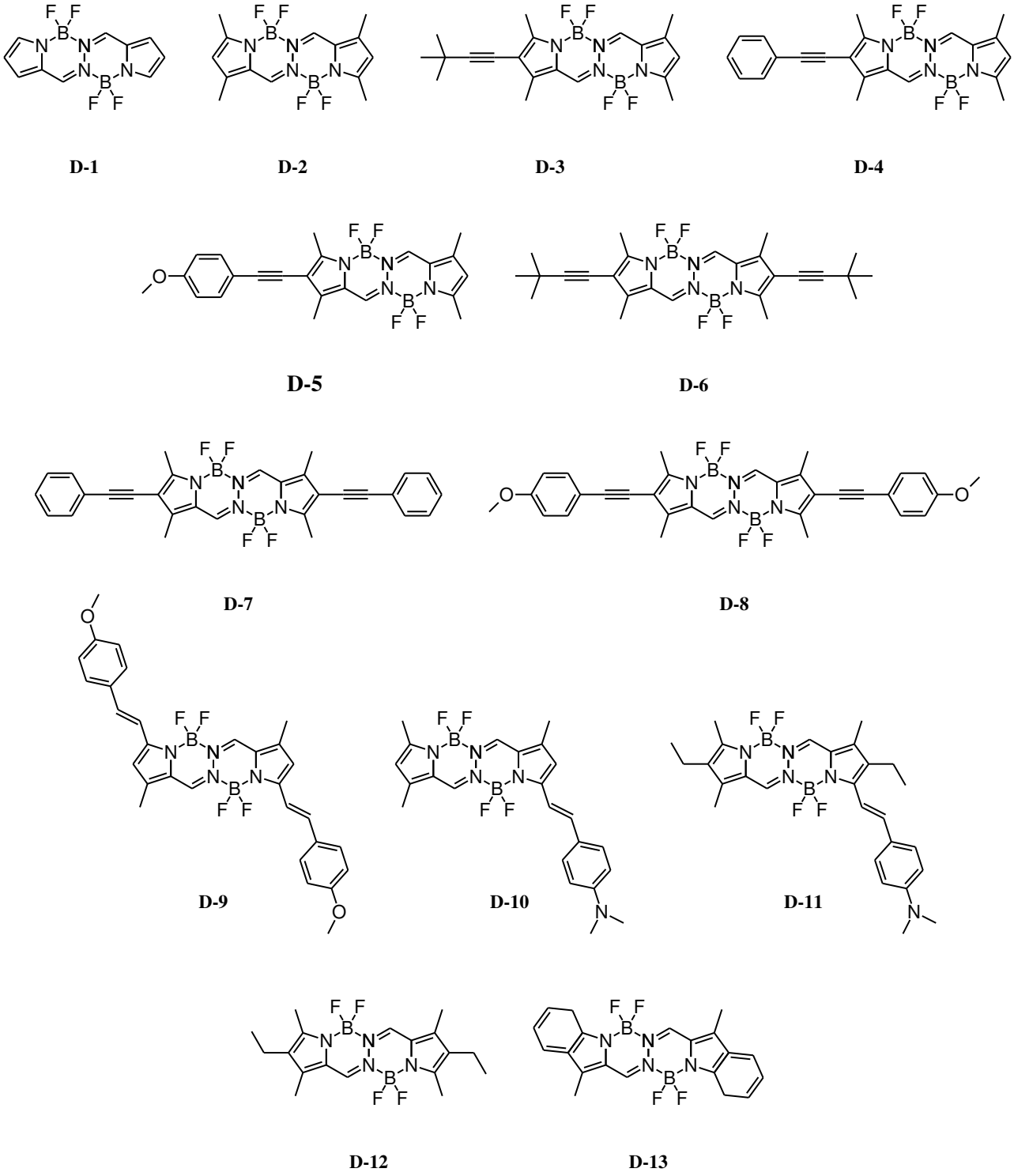


Figure 13: Compound **C-6**, the dependence of one-photon transition strengths (μ) corresponding to $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_2$ electronic transitions on the dihedral angle θ

In conclusion, these analyses demonstrated an effective strategy for the design of highly fluorescent dyes with good 2PACS. While dipolar dyes have weak 2PA magnitudes, their quadrupolar analogues show both sufficient 2PA cross sections and bright fluorescence in chloroform solution. The electronic density difference plots demonstrated that the twisting around the central bond in the studied quadrupolar structures leads to changes in their electronic structure nature. The dependence of 2PA strength on the rotation around the dihedral angle shows that by symmetry relaxation the reduction of 2PA efficiency might be even as large as 20%. The GFSM analysis demonstrated that the observed behavior of 2PA strength upon symmetry relaxation is caused by transition moment dependence on the torsional parameter.

The presented above studies of various electron-donating substituents, dipolar and quadrupolar structures, and planarity relaxation were extended by an investigation of BOPHY derivatives. The design of BOPHY dye was inspired by prosperous BODIPY derivatives which are widely used in multiple 2PA applications. The easy two-step synthesis of BOPHY dyes makes them especially interesting for potential applications. However, unfortunately, even though BOPHY dyes were first published in 2014, the 2PA properties of these dyes had not been reported yet. The aim of these efforts is to investigate theoretically the 2PA properties of BOPHY derivatives containing both dipolar and quadrupolar structures with electron-donating substituents of different strengths. The selected for this work series **D** is given in Scheme 9.



Scheme 9: Studied compounds of series **D**

In experimental studies selected dyes demonstrated high FQY (Table 7), presenting values up to 1.00 in dichloromethane solution [110, 258–261], which is an essential property for potential 2PA applications. Furthermore, for these dyes one-photon absorption wavelength lies in the range of 400-600 nm. Thus, their two-photon excitation energies fit into the

first biological window. Previously the 1PA properties of this set were investigated theoretically by M. Ponce-Vargas et al. [262]. This study included the calculations of 0-0 energies, simulations of 1PA vibronic spectra, and the analysis of charge transfer upon excitation. The current study would extend Ponce-Vargas et al. work by performing simulations of 2PA strengths and rationalizations of obtained "structure–property" relationships aided by GFSM approach combined with the CC2 model.

	λ_{abs}, nm	λ_{flu}, nm	FQY		λ_{abs}, nm	λ_{flu}, nm	FQY
D-1	424	465	0.95	**D-8**	468	536	0.46
D-2	444	485	0.92	**D-9**	579	611	0.47
D-3	453	499	0.83	**D-10**	557	667	0.20
D-4	456	509	0.70	**D-11**	557	670	0.21
D-5	479	523	0.35	**D-12**	478	497	1.00
D-6	464	512	0.73	**D-13**	498	580	0.06
D-7	489	524	0.72				

Table 7: Series **D**, experimental 1PA properties in dichloromethane. Absorption maxima (λ_{abs}), emission maxima (λ_{em}), fluorescence quantum yield (FQY) [110, 258–260]

The selected set is useful for the current discussion of the 2PA properties of BOPHY dyes because it includes various structural modifications essential for the 2PA activity. The compounds **D-1** (BOPHY) and **D-2** (1,3,6,8-tetramethyl BOPHY) are used in this study as the starting point for further extensions. In the subsets **D-3** – **D-5** and **D-6** – **D-8** the strength of electron-donating groups increase. Furthermore, using these subsets one can compare the dipolar dyes **D-3** – **D-5** with their quadrupolar analogues **D-6** – **D-8**. The impact of π-paths represented by one triple bond and one double bond can be evaluated based on the compounds **D-8** and **D-9**, respectively. The simulated 2PA strength of the compounds **D-9** and **D-10** would help in assessing which design strategy is more successful: mediocre electron donating group (OMe, Hammett's substituent constant $\sigma_p = -0.27$ [263]) in quadrupolar structure or strong electron-donating group (NMe_2, $\sigma_p = -0.83$ [263]) in dipolar structure. The pairs **D-10** and **D-11**, **D-2** and **D-12** permit to evaluate the impact of steric effects on 2PA properties since the ethyl substituents could cause distortion of symmetric BOPHY-core. This is interesting because 2PA probes require additional bulky substituents for conjugation with target biomolecules or/and increasing solubility. Finally, by investigation of the structure **D-13** it is possible to understand how benzannulation affects the photophysical properties of BOPHY dye.

Considering that this work includes only a theoretical prescreening of the 2PA properties, the calculations were performed solely in the gas phase, thus, significantly saving the time required for the simulations. The study started with the ground state geometry optimization at the B3LYP/6-31G(d) [238–240] level of theory as implemented at the Gaussian program [224]. After minimum at the PES was confirmed by evaluation of the Hessian, electronic structure calculations were conducted using response theory combined with RI-CC2 model and cc-pVDZ basis set. Turbomole v 7.3 program was used at this step. Additionally, the 2PA strengths were decomposed using GFSM model in its 2SM and 3SM variants. The photophysical properties obtained for the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions are given in Table

8.

molecule	$S_0 \rightarrow S_1$					$S_0 \rightarrow S_2$				
	ΔE, eV	λ_{abs}, nm	f	δ^{2PA}, a.u.	δ_{2SM}, a.u.	ΔE, eV	λ_{abs}, nm	f	δ^{2PA}, a.u.	δ_{3SM}, a.u.
D-1	3.28	378.0	0.99	0.0		4.16	298.0	0.00	1.4×10^4	
D-2	3.09	401.2	1.07	0.0		4.13	300.2	0.00	1.7×10^4	2.3×10^4
D-3	3.00	413.3	1.24	0.1×10^4		3.74	331.5	0.05	2.8×10^4	3.7×10^4
D-4	2.96	418.9	1.38	0.4×10^4		3.64	340.6	0.07	4.7×10^4	6.6×10^4
D-5	2.92	424.6	1.37	1.0×10^4		3.54	350.2	0.12	6.5×10^4	8.6×10^4
D-6	2.92	424.6	1.43	0.0		3.57	347.3	0.00	5.8×10^4	9.4×10^4
D-7	2.84	436.6	1.74	0.0		3.40	364.7	0.00	11.9×10^4	21.3×10^4
D-8	2.79	444.4	1.79	0.0		3.28	378.0	0.00	18.4×10^4	33.0×10^4
D-9	2.62	473.2	1.97	0.0		3.20	387.5	0.00	40.9×10^4	43.9×10^4
D-10	2.70	459.2	1.29	9.4×10^4	12.9×10^4	3.27	379.2	0.47	12.2×10^4	13.9×10^4
D-11	2.72	455.8	1.24	6.7×10^4	9.3×10^4	3.18	389.9	0.51	9.2×10^4	11.0×10^4
D-12	3.04	407.8	1.20	0.0		3.88	319.5	0.00	2.0×10^4	
D-13	2.78	446.0	1.19	0.0		2.94	421.7	0.00	0.9×10^4	

Table 8: Series **D**, calculated photophysical properties. One-photon excitation energy (ΔE), absorption wavelength (λ_{abs}) oscillator strength (f), and average two-photon transition strength (δ^{2PA}). δ values obtained using two- (δ_{2SM}) and three-state (δ_{3SM}) model are also shown

The dependencies of the calculated vertical excitation energies on structural modifications in BOPHY dyes follow well-known design principles. Compared to the initial **D-1** and **D-2** molecules, the introduction of electron-donating groups in the compounds **D-3** – **D-12** causes a red-shift of the excitation wavelengths. In the arrays **D-3** – **D-5** and **D-6** – **D-8** the increase of electron-donating capacity of the substituent results in red-shifted absorption wavelengths. Comparing the dipolar compounds **D-3** – **D-5** with their quadrupolar analogs **D-6** – **D-8**, a red-shift of absorption wavelength is observed. The compound **D-8** with triple bonds in π-path has a lower absorption wavelength than the compound **D-9** with double bonds in π-path. The quadrupolar dye **D-9** with weaker electron-donating substituent has a higher absorption wavelength than the dipolar dyes **D-10** and **D-11** with stronger electron-donating groups. Based on the results for the compounds **D-10** and **D-11**, and **D-2** and **D-12**, we can conclude that possible steric effects do not significantly impact excitation energies. Last but not least, the benzannulation (as in the compound **D-13**) led to a red-shift of energy, in agreement with what was observed for compounds with N–BF_2–O fragment in the work [109]. The calculated $S_0 \rightarrow S_1$ excitation wavelengths parallel experimental findings, even though calculations were performed in the gas phase. However, as was observed for the series **A**, **B** and **C**, the simulated RI-CC2 absorption wavelengths underestimate the experimental data. A significant discrepancy is detected for the quadrupolar compound **D-8**. In the experiment the absorption maxima of **D-8** is blue-shifted compared with its dipolar analog **D-5** and to the compound **D-7** with weaker electron-donating substituent, while in performed RI-CC2 simulations (and in mentioned theoretical

work [262] using TD-DFT/PCM) the vertical excitation energy of the compound **D-8** is red-shifted. Nevertheless, the experimental absorption spectra of BOPHY derivatives have an explicit vibronic structure with several peaks, and an incorrect interpretation of experimental data might be the reason for theory–experiment inconsistency. Assuming that RI-CC2 underestimates absorption wavelength, we can see that for both considered transitions and for all compounds the excitation energies lay in the first biological window, thus, an important bioimaging requirement is fulfilled.

The transition $S_0 \rightarrow S_1$ has a weak 2PA strength for the majority of the studied here dyes. The compounds **D-1**, **D-2**, **D-5** – **D-9**, **D-12**, and **D-13** have a quadrupolar structure and belong to C_i symmetry point group, and for this reason, their $S_0 \rightarrow S_1$ transition is symmetry forbidden in 2PA spectra. Consequently, their 2PA strengths for this transition are insignificant. Amongst the studied dipolar compounds, the dyes **D-3** – **D-5** also have low 2PA strengths. Note that the molecules **D-10** and **D-11** have stronger electron-donating groups than the compounds **D-3** – **D-5**, and this leads to improvements in their 2PA strengths. Due to steric effects causing a distortion of the π-conjugation, the compound **D-11** has slightly lower 2PA strength than the compound **D-10**. The decomposition of 2PA strengths corresponding to the $S_0 \rightarrow S_1$ transition for the molecules **D-10** and **D-11** using 2SM formalism is presented in Figure 14. The total 2SM transition strengths are in good agreement with the values computed by the response theory. The term δ_{0111} is dominant for both dipolar compounds, but the resulting 2PA strength is almost twice smaller due to the negative effect of the term δ_{0101}. The same was observed for the series **A**. According to Equation 7 (page 23) the term δ_{0111} implies the product of the square of the $S_0 \rightarrow S_1$ transition dipole moment and the square of the dipole moment in the S_1 state. The smaller magnitude of the term δ_{0111} of the compound **D-11** is caused by its smaller dipole moment in the S_1 state, which is the result of its π-system deformation.

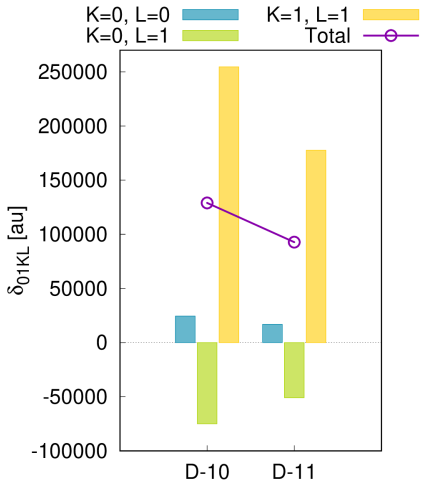


Figure 14: Series **D**, summary of 2SM calculations corresponding to the two-photon $S_0 \rightarrow S_1$ transition

The transition $S_0 \rightarrow S_2$ is 2PA-bright for all the compounds in the series **D**. As expected, the 2PA strength is growing upon the increase of electron-donating properties of the substituent. This is clearly observed for the subsets **D-3** – **D-5** and **D-6** – **D-8**. Additionally, the comparison amongst the dipolar dyes **D-3** – **D-5** and the quadrupolar dyes **D-6** – **D-8** show the increment of 2PA strength in quadrupolar systems with respect to dipolar ones. The compound **D-9** with double bonds in the π-bridge has two times larger 2PA strength than the structure **D-8**, which has similar electron-donating groups but with triple bonds in the π-path. Unfortunately, a double bond might lead to undesirable

photoisomerization (as it was discussed for the series **A**), and this aspect should be explored separately. Due to steric effects, the dye **D-11** has a smaller 2PA strength than **D-10**. However, the 2PA strength of the dye **D-12** is higher than that of the dye **D-2** since the additional ethyl groups do not cause undesirable steric effects, and play the role of complementary electron-donating substituents. Interestingly, the dipolar compound **D-10** with strong electron-donating group has a significantly smaller 2PA strength than the quadrupolar structures **D-7**, **D-8** and **D-9**, which have weaker electron-donating groups, and roughly the same 2PA strength as *t*-butyl-ethynyl-substituted quadrupolar structure **D-6**. Thus, quadrupolar structures with weaker electron-donating groups show higher 2PA magnitudes than dipolar structures with strong electron-donating groups. Moreover, the inspection of Table 7 allows us to conclude that according to experimental data the FQY of the dipolar dye **D-10** is more than twice smaller than of the dye **D-9**. Thus, the compound **D-9** would have 2PACS higher than the compound **D-10** roughly by a factor of 8. The benzannulated dye has the smallest 2PA strength in the set **D**. Thus, if required, it should be applied in combination with strong electron-donating groups.

To shed some light on the inner characteristics responsible for substituent-induced changes of the 2PA properties for the studied dyes, a 3SM analysis was performed. To that end, the first five intermediate states were considered, and finally, the intermediate state of interest was selected as the one that yields the total 3SM transition strength value closest to the response theory value. For all studied compounds the S_1 electronic excited state was chosen as the intermediate state. Figure 15 shows the values of the dominant term δ_{0211} and total 2PA strength delivered by 3SM approach. According to Equation 7 (page 23), the major term δ_{0211} encompasses the product of the square of the transition dipole moment between S_0 and S_1 states and the square of the transition dipole moment between S_1 and S_2 states. In the dipolar structures, this term is slightly higher than the corresponding total 3SM strength. And in the quadrupolar dyes the term δ_{0211} is the only non-zero term due to the symmetry, and, consequently, the total 3SM 2PA strength is equal to the value of this term.

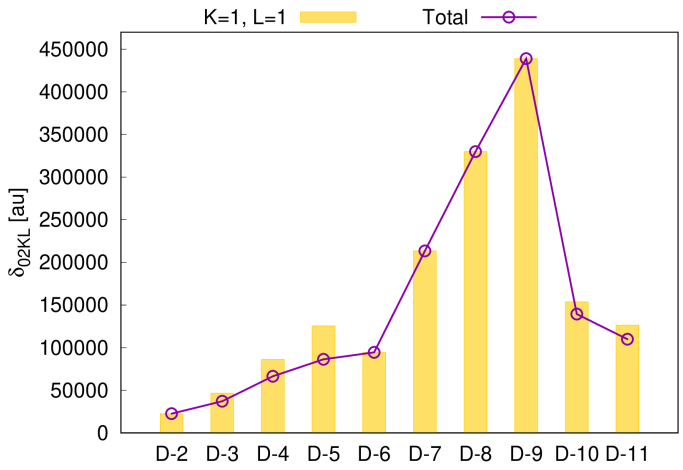


Figure 15: Series **D**, summary of 3SM calculations corresponding to the two-photon $S_0 \rightarrow S_2$ transition

In conclusion, for the first time, the 2PA properties of the highly fluorescent BOPHY dyes were studied. Unsurprisingly, the simulations confirmed that modifications of BOPHY derivatives cause changes of 2PA strengths according to the following classic principles: *i)* the increase of electron-donating power of substituents causes the increase of 2PA

strength; *ii)* quadrupolar structures have higher 2PA strengths than their dipolar analogues; and *iii)* double bonds in π-bridges are more favorable for the 2PA magnitude than triple bonds. The performed studies demonstrated that the quadrupolar BOPHY derivatives with weaker electron-donating groups are more beneficial than the dipolar structures with stronger electron-donating groups. Importantly, based on the set **D**, a route for magnifying 2PA strength by a factor of 30 (compound **D-9**) was established. The discussed strategies are promising for further development of 2PA probes for imaging.

Figure 16 summarizes results obtained for the 4 sets discussed above. It is focused on the calculated 2PA cross section and absorption wavelength, which are two key photophysical properties required for 2PA bioimaging. The excitation energies of the majority of the studied compounds lay in the first biological window. Moreover, structural modifications leading to a red-shift of excitation energy also result in an increase of the 2PA cross section. Importantly, the maximum values of the 2PA cross sections of the investigated dyes reach 1500 GM, and amongst the designed probes (at least one compound from each series) more than 15 compounds have 2PA cross sections higher than 200 GM. The highest 2PA cross sections are observed for the $S_0 \rightarrow S_2$ transition in the series **D**, the BOPHY derivatives, which deserve further experimental investigation. Thus, the design strategies investigated in this chapter have high potential for further implementation into the development of commercial probes for 2PA bioimaging.

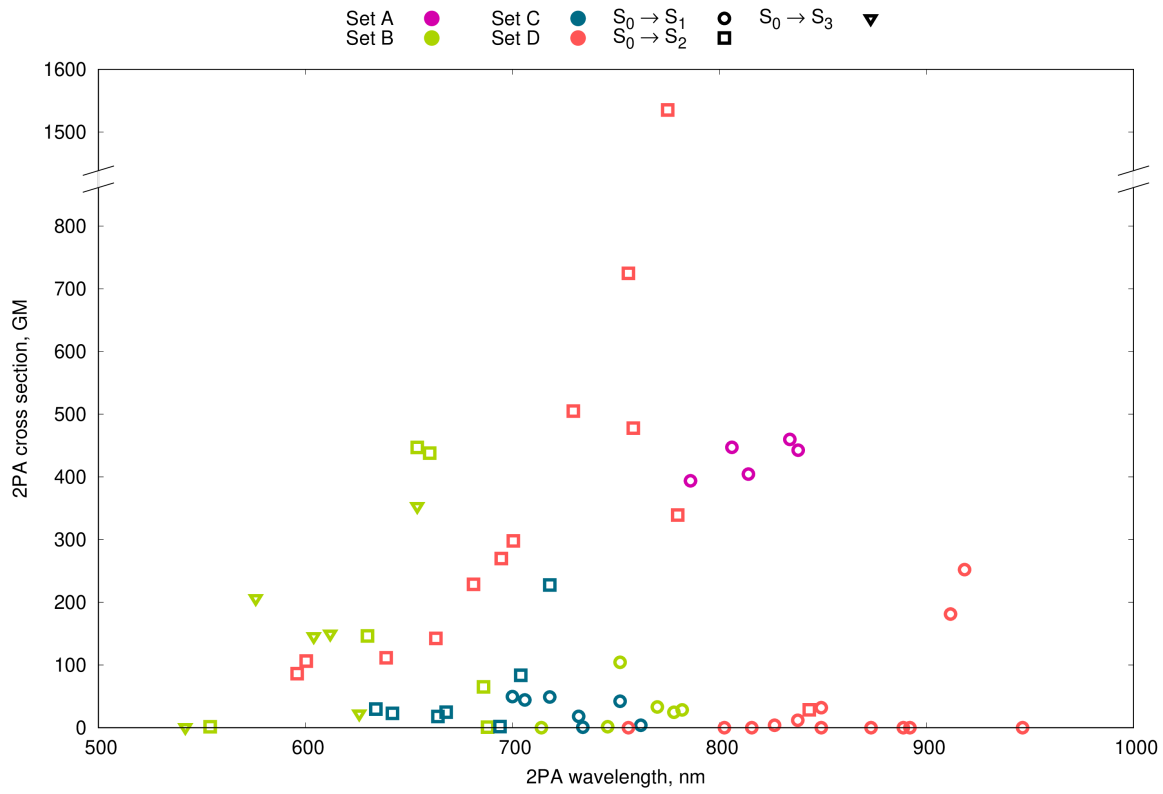


Figure 16: Series **A**, **B**, **C**, and **D**, summary of computed 2PA wavelengths (*X*-axis) and 2PA cross sections (*Y*-axis). 2PA cross sections were calculated assuming Lorentzian broadening equal to 0.2 eV. Dot color defines the series of studied compounds, and dot type defines the transition considered. For series **A**, **B**, and **C** data was computed including the effect of chloroform as solvent, and for series **D** calculations were done in gas phase

5 Simulations of vibrationally-resolved one- and two-photon absorption spectra

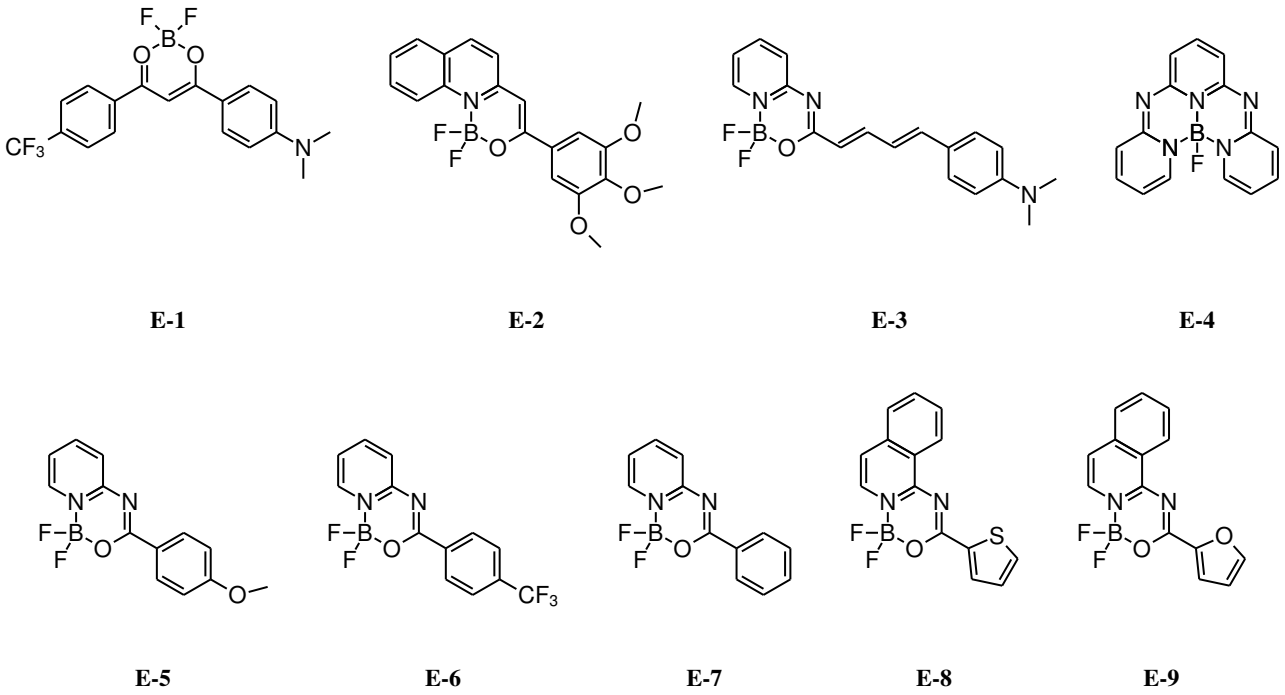
In the previous chapter, the analysis of 1PA and 2PA properties was performed neglecting the effect of molecular vibrations. Note, for example, that the 1PA and 2PA spectra of the series **B** have a complex vibronic structure (Figure 8), with several local absorption maxima observed for the same electronic absorption band. Thus, the calculations of vertical excitation energies might be insufficient for the interpretation of complex band shapes in experimental spectra.

As mentioned in the Introduction, there are multiple works comparing different quantum chemistry theories for simulations of ground and excited state geometries and Hessians and evaluating different approaches for simulations of vibronic couplings as well as some techniques for simulations of inhomogeneous broadening. However, reported works either treat computational aspects separately or do the comparison for a small set of dyes. Unfortunately, selecting an approach for simulations of vibronic structure even for 1PA spectra is by no means an easy task. The literature reports controversial results that are insufficient for an unambiguous choice of vibronic approach. The study [180] demonstrates that VH and AH produce similar spectra and they are superior to VG and adiabatic shift (AS) models for semi-rigid molecules. Ref. [174] states that VG is inapplicable for simulations of the vibronic structure of aminocoumarines, while work [264] demonstrates that AH fails in simulations of BODIPY dyes which have significant differences between ground and excited state geometries. The study [265] reports a good agreement between VG approach and experiment for 19 dyes with N–BF_2–O group, but it does not include a comparison with AH model. Based on available literature it is difficult to predict whether AH approach would outperform VG model for studying the vibronic structure of bands in the spectra of organoboron complexes, or whether it is needed to include the mechanical anharmonicity to accurately reproduce the experimental spectra. Thus, there is still a lack of a robust computational protocol for the theoretical prediction of vibrationally-resolved spectra of fluorescent dyes. Motivated by these gaps, this chapter is focused on the development of a fully nonempirical computational protocol for the simulations of vibrationally-resolved spectra of dyes containing the BF_2 group. Namely, we performed a comparison amongst various vibrational spaces, evaluated the importance of electronic HT and technical anharmonic corrections, and additionally developed a cost-effective approach for the robust estimation of inhomogeneous broadenings. However, the assessment of available theories for simulations of 0-0 transition energies will not be presented, as there are multiple extensive works devoted to this subject [266–270], which came to similar conclusions. They conclude that the coupled clusters approximation provides accurate results, and, efficient TD-DFT combined with B3LYP, CAM-B3LYP, or M06-2X functionals was proved to be reliable. For this reason, we will skip computational estimation of exact 0-0 transition energies, and the spectra presented in this chapter were simply shifted to overlap experimental 0-0 transition (the absorption cross section is proportional to the transition energy and this was taken into account).

5.1 Simulations of vibrationally-resolved spectra

We will start with the analysis of adiabatic and vertical harmonic approximations in simulations of 1PA spectra. In what follows we will study only the bright $S_0 \rightarrow S_1$ transition. The compounds studied in this investigation were selected from a representative set of organoboron complexes, which includes: *i)* the series **E** (Figure 10); *ii)* the compounds **A-1** – **A-6** (Figure 6), which are the molecules of the series **A** for which it was possible to obtain experimental spectra; and *iii)* all dyes from the set **B** (Figure 7), except the **B-3** structure because it presents a very low intensity for the $S_0 \rightarrow$

S_1 transition. The whole set of the studied compounds contains a diversity of structural modifications that might affect the vibronic structure, including the rigid molecule (**E-4**), molecules with unsubstituted phenylene moieties having significant rotational freedom (**B-1**, **B-2**, **E-7**), compounds with a donor-acceptor substituted phenylene moiety (**A-1**, **B-2**, etc.), derivatives with a vinylene linker that are prone to photoisomerization (**A** series), and a molecule with two vinylene linkers (**E-3**). Most of the compounds studied here have superb charge transfer properties and present structural variability that leads to different experimental absorption band shapes and widths. Thus, the proposed set is suitable for the verification of approaches for simulations of the vibronic structure of the bands in the spectra.



E-1 **E-2** **E-3** **E-4**

E-5 **E-6** **E-7** **E-8** **E-9**

Scheme 10: Studied compounds of series **E**

Vibronic structure simulations require the estimation of ground and excited state geometries and Hessians based on electronic structure simulations. For realistic fluorescent probes containing more than 50 atoms the optimization of excited state geometries and calculation of excited state Hessians using highly accurate theories are unfeasible due to significant computational costs. One of the potential solutions is the application of TD-DFT for these calculations, although it is very important to choose a reliable DFA. In the current study the selection of DFAs for refined calculations was performed using the vibrational reorganization energy as a metric. This choice was justified by the close relation of this value to the experimentally observed width of absorption bands and calculated frequencies and displacements of normal modes. The reliability of this approach was confirmed earlier [265]. This vibrational contribution to the reorganization energy is:

$$\lambda^k_{\text{vib}} = \sum_{j=1}^{3N-6} \omega_j s_j^k = \frac{1}{2} \sum_{j=1}^{3N-6} \omega_j (\Delta_j^k)^2, \tag{44}$$

where k labels the excited states, ω_j is the vibrational frequency and s_j^k is the Huang-Rhys factor related to j^{th} normal mode dimensionless displacement Δ_j^k between the electronic state k and the ground state minimum.

Vibrational reorganization energies calculated with the RI-CC2 method were used as reference values for further examination of the DFAs. Together with some medium-size basis set, the RI-CC2 approximation is a solid wavefunction method. RI-CC2 was shown to be trustworthy approach for further vibronic structure simulations [195, 265]. The benchmark performed in the current work included an extensive set of DFAs: semilocal (BLYP), global hybrids (B3LYP, PBE0, MN15, BH&HLYP, M06-2X), and range-separated hybrids (CAM-B3LYP, ωB97X-D, ωB97X, LC-BLYP, and optimally tuned LC-BLYP-OT(J) and LC-BLYP-OT(α) [271]). The performance of the majority DFAs listed here was already tested in vibronic structure simulations, except MN15 and two later optimally-tuned DFAs. Previously, CAM-B3LYP, M06-2X, ωB97X-D, ωB97X, and BH&HLYP functionals demonstrated a good agreement of vibrational reorganization energies with RI-CC2 values for difluoroborates [265]. However, that study included only one family containing N–BF_2–O moiety. The ground state geometry optimization, calculations of the ground state Hessian, and calculations of the excited state gradient at the ground state geometry were performed for each of the listed DFAs. The cc-pVDZ basis set was used for these calculations. TD-DFT simulations were completed using *ultrafine* pruned integration grid as implemented in Gaussian 16 program [224]. RI-CC2 calculations were performed using TURBOMOLE v. 7.3 program [235]. These electronic structure calculations were performed in the gas phase. All harmonic vibronic structure simulations were performed at 300K using FCclasses 3 program [272].

The results of the analysis described above are illustrated in Figure 17, where is shown the map of calculated absolute vibrational reorganization energies and the map of corresponding unsigned relative percentage errors computed for every DFA/compound in comparison with the reference RI-CC2 values. Three functionals are showing average errors lower than 20%: ωB97X (ave. 9.09%, max. 25.91%), LC-BLYP (ave. 10.37%, max. 28.36%), and LC-BLYP-OT(α) (ave. 17.08%, max. 34.85%). MN15 functional is potentially interesting for its excellent description of excited state properties (very accurate vertical excitation energies and excited-state dipole moments [257]), but unfortunately, it has significant errors (ave. 39%) in simulations of vibrational reorganization energies. Figure 18 demonstrates how significant might be the difference in vibronic spectra between "the best" and "the worst" DFAs and confirms that the vibrational reorganization energy and the vibrational fine structure of the absorption band are deeply related to each other, emphasizing the importance of benchmarking performed.

Two functionals were selected for further assessment of the vibronic structure models: ωB97X (which has the best agreement with RI-CC2), and novel LC-BLYP-OT(α) (which has a good agreement with RI-CC2 and is promising for further 2PA studies due to it accurate simulations of second hyperpolarizabilities and reliable chemical rankings for two-photon absorption strengths [120, 271]). Moreover, as demonstrated in the recent studies by Sitkiewicz et al. the LC-BLYP functional does not have the spurious oscillations problem, and thus it is suitable for computing vibronic spectra with mechanical anharmonicity included [179, 273]. In order to reproduce the experimental conditions, all further electronic structure calculations were performed accounting for the chloroform environment. Ground and excited state geometries were reoptimized and Hessians and energy gradients were recomputed using the pruned integration grid *SuperFine*, the aug-cc-pVDZ atomic basis set, and the PCM approach accounting for the influence of the chloroform solvent. Moreover, electronic density difference plots ($\Delta\rho(\boldsymbol{r}) = \rho^{S_1}(\boldsymbol{r}) - \rho^{S_0}(\boldsymbol{r})$)) were calculated for $S_0 \rightarrow S_1$ transition using MN15/aug-cc-pVDZ level of theory with the Gaussian program [224]. The resulting plots for representative molecules are illustrated in Figure 19. They confirm that studied compounds have the same $\pi \rightarrow \pi^*$ character of $S_0 \rightarrow S_1$ transition.

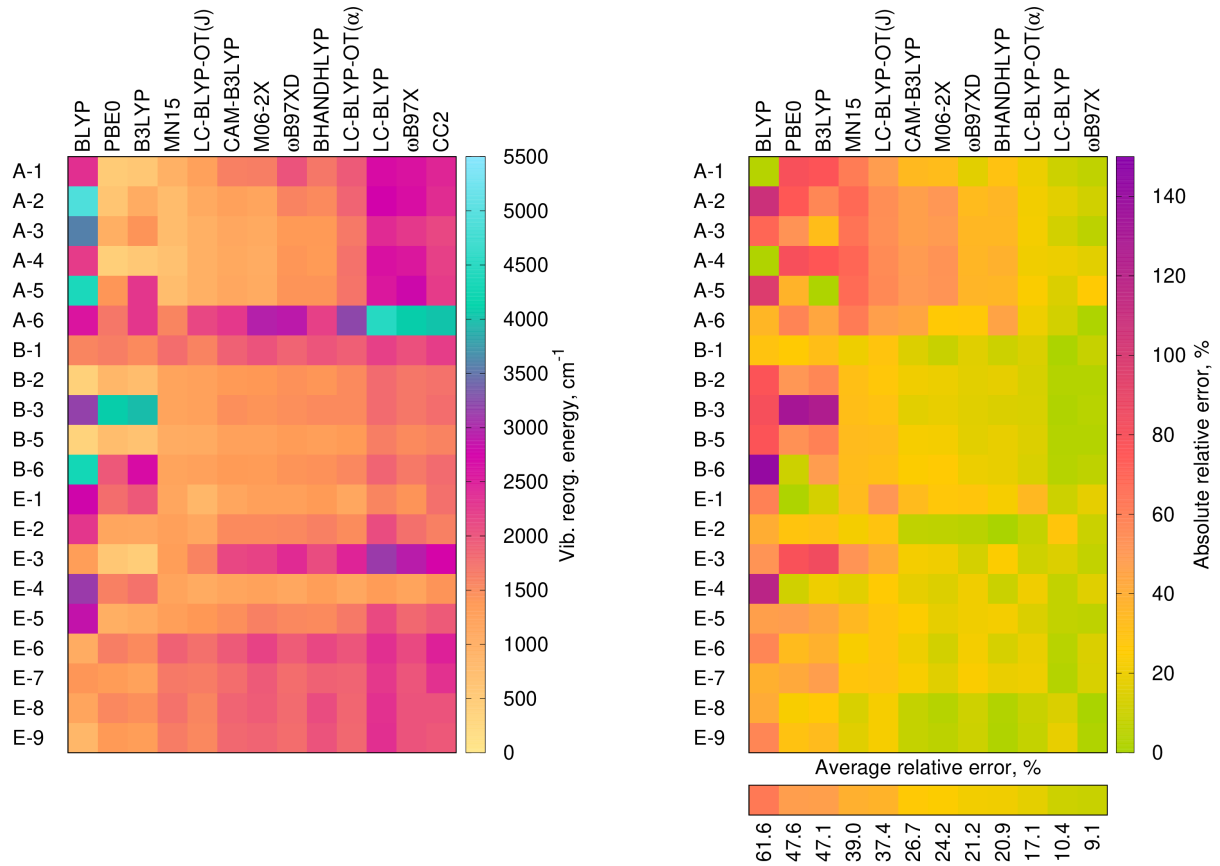


Figure 17: Vibrational reorganization energies for various DFAs and corresponding unsigned relative errors (using RI-CC2 values as reference)

The first vibronic structure calculations were performed aiming at the comparison of the two selected functionals using the VG approach, and the obtained results are shown in Figure 20. The functionals ωB97X and LC-BLYP-OT(α) are in excellent agreement for all structures, which supports the reliability of the functional selection procedure. The further discussion of different approaches for vibronic structure simulations in this section is based on LC-BLYP-OT(α) electronic structure calculations. However, all presented simulations were reproduced using ωB97X, and the spectra obtained by both functionals are similar and lead to the same final conclusions.

For the vibronic structure simulations, geometries and Hessians of the ground and excited states are required to build the vibrational space. A scheme illustrating the difference between methods used in this work is shown in Figure 21. In all approaches the ground state descriptors are the same, the optimized S_0 geometry and the S_0 Hessian calculated at the optimized geometry. The applied methods differ only in the excited state descriptors. The most "candid" method is the AH approach, where the S_1 geometry and the S_1 Hessian at the S_1 geometry have to be computed using quantum chemistry methods. The most computationally cheap method is the VG approach, where excited-state geometry is the first step of the Newton-Raphson optimization from the ground state geometry using the excited state gradient. Note that the excited state Hessian used in VG is equal to the ground state Hessian (S_0 Hessian calculated at S_0 geometry). The VG approach significantly reduces the CPU time compared to the AH model, but previous studies show that the accuracy might become a victim of this simplification [180]. Other approaches used in this work (VH and displaced

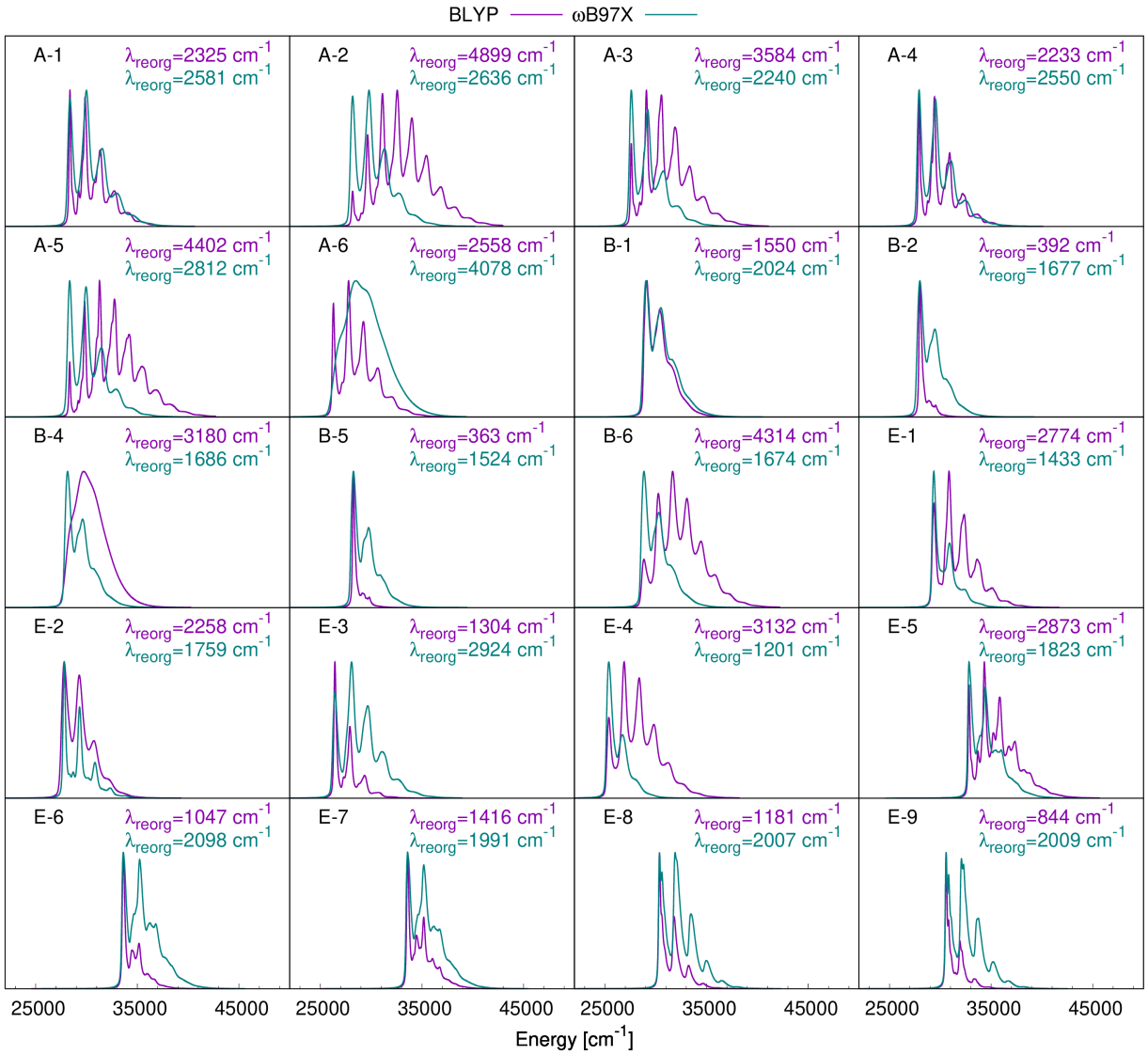


Figure 18: Comparison between ωB97X and BLYP DFAs (gas phase) in simulations of vibrationally-resolved absorption spectra (using VG model). ωB97X and BLYP are "the best" and "the worst" DFAs according to the evaluation of the methods based on the vibrational reorganization energy

adiabatic Hessian (dAH)) are transitional between AH and VG, aiming at the understanding of the difference in vibronic structures obtained by AH and VG. The VH approach uses VG geometry and S_1 Hessian computed at the optimized S_0 geometry, and the dAH approach uses VG geometry and S_1 Hessian computed at VG geometry. Evidently, excited state Hessians calculated for two latter approaches might lead to imaginary frequencies. In VH approach FCclasses program turns all imaginary frequencies into real numbers by taking the module, and the same routine was reproduced for the dAH method. But this trick makes the two latter approaches unreliable, which is why in these studies they are used only to shed some light on the difference between the AH and VG results. A summary of the theoretical basics of harmonic approximations was given above (pages 23-25).

Vibronic structure simulations might be performed in the TI formalism [274–277] or in the TD formalism [278, 279]. The TI method allows the calculation of the FC factors individually for each transition, while the TD approach

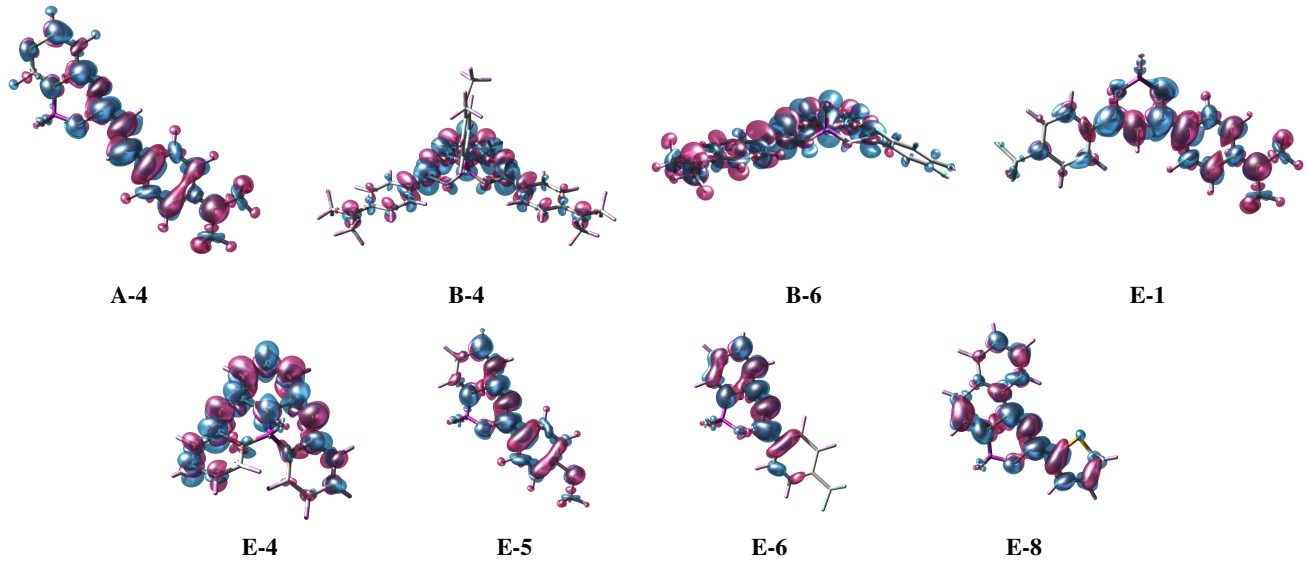


Figure 19: Series **A**, **B** and **E**, electronic density difference plots corresponding to the $S_0 \rightarrow S_1$ transition for representative molecules. The presented plots were prepared with the same density contour value (0.001)

performs Fourier-transformation of the corresponding correlation function. Thus, TD approach saves a lot of time by skipping the screening of the FCFs for billions of possible transitions, but from TD calculations one can not identify FCFs of individual vibronic transitions. For most of the compounds studied here, TI calculations showed a poor convergence since the recovery of FCF spectra (the sum of FCFs which should be ideally equal to 1.00) does not exceed 0.5 even in calculations with 10^{12} integrals considered. Furthermore, TI calculations require significant central processing unit (CPU) time. The information about assignments of FCFs to individual transitions is less important for this study. Therefore, to obtain fully converged vibronic spectra, the TD formalism was employed for the confrontation of adiabatic and vertical models.

Figures 22 – 26 contain the simulated vibronic spectra essential for this discussion. All the spectra presented here were plotted assuming Gaussian broadening with HWHM equal to 100 cm^{-1}, except in Figure 26 where the values of Gaussian HWHM were selected for each structure to fit experimental FWHM. As it was noted before, the spectra shown here were obtained from LC-BLYP-OT(α) electronic structure calculations. There is one more aspect of the simulations that should be considered during vibronic structure simulations. It is possible to perform calculations in internal or Cartesian coordinates, and it turned out to be a crucial parameter for studied dyes. Figure 22 shows the difference between spectra simulated in internal and in Cartesian coordinates using the AH approach (in the VG approach both coordinates give the same result). For many structures simulations in Cartesian coordinates showed an inferior performance, since the obtained bands are unphysically wide. For instance, for the compounds **E-1**, **E-6**, and **E-7** the results obtained with the Cartesian coordinates are clearly erroneous. A significant improvement of vibrationally resolved spectra is achieved using internal coordinates. The inspection of geometrical parameters for molecules, that have the most different spectra when different coordinates are used, shows large changes in the excited state geometry. The main explanation for the poor performance of Cartesian coordinates is the erroneous interpretation of translational and rotational coordinates character, which results in their insufficient elimination from vibrational space. It should

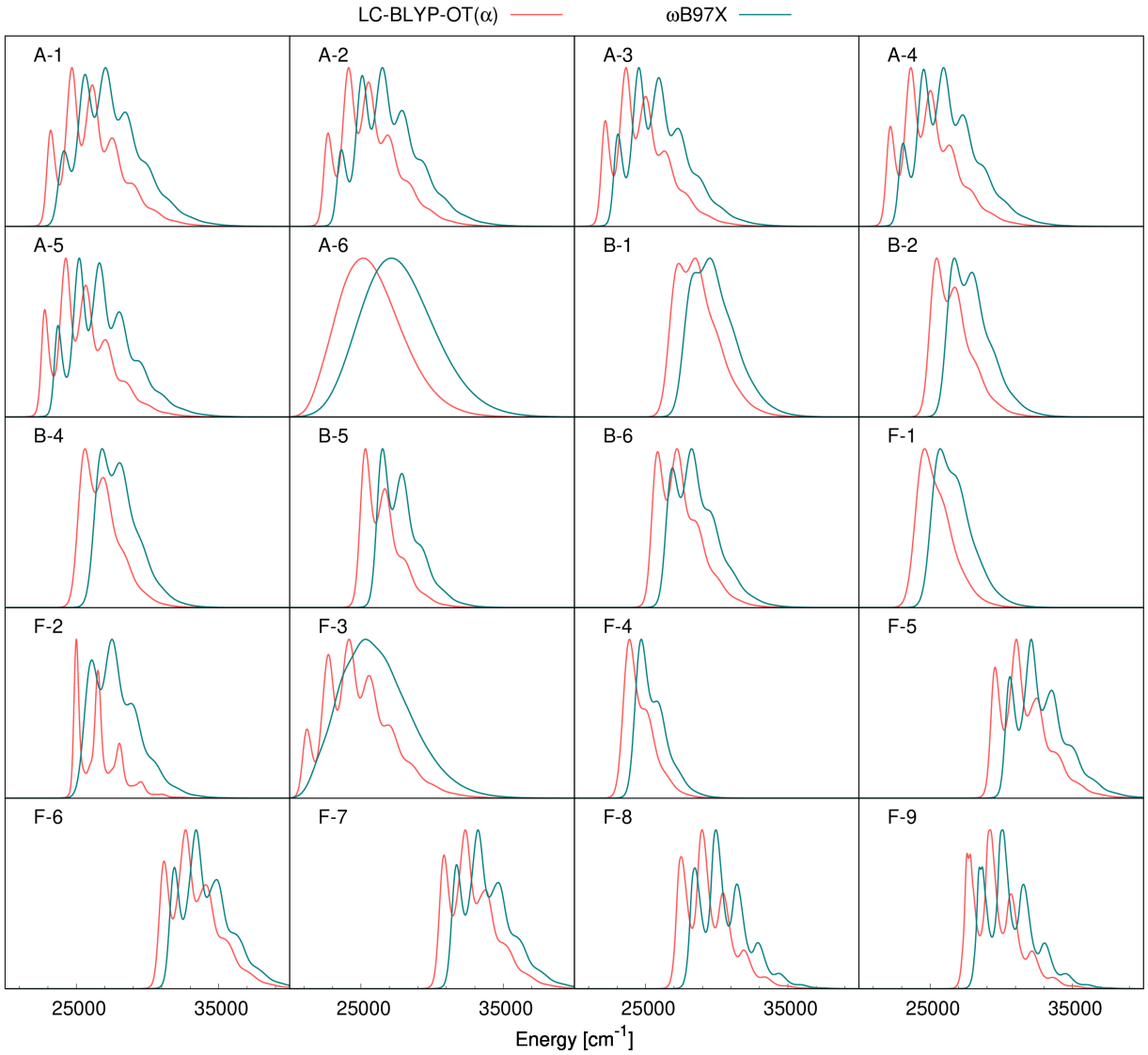


Figure 20: Comparison between ωB97X and LC-BLYP-OT(α) (in chloroform solution) in simulations of vibrationally-resolved absorption spectra using VG model

be highlighted that simulations in internal coordinates do not suffer from this problem since the vibrational space is defined directly through inner bonds, angles, and dihedrals. The literature points out that in vertical models, the internal coordinates are almost systematically superior to Cartesian coordinates [216].

The comparison between VG and AH models (Figure 23) shows that spectra computed by AH approach remains unphysically wide even in internal coordinates. It was found that the excited state geometry is significantly different from the ground state geometry for the compounds with problematic AH spectra. The same was observed for the set of BODIPY dyes before [264]). For example, in the case of the compound **E-6**, upon excitation the trifluoromethylphenyl substituent leaves the plane of BF_2-core. On the contrary, the rigid compound **E-4** does not have visible differences in the ground and excited state geometries, and its VG and AH spectra successfully coincide. Thus, significant differences between geometries of ground and excited state transitions cause long vibrational progressions, producing in this way extremely wide spectra. This also explains why Cartesian coordinates show poor results in VG simulations.

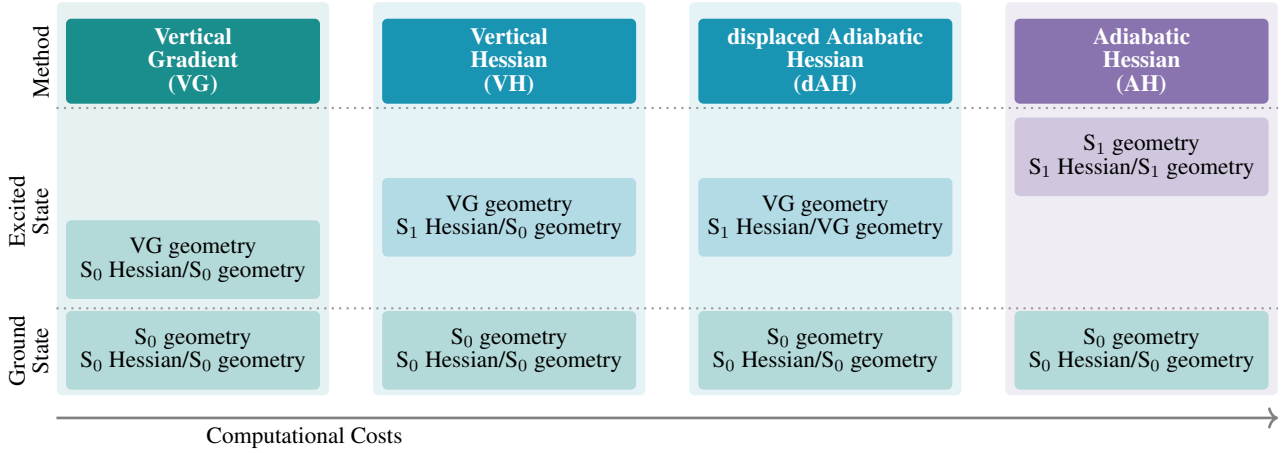


Figure 21: Schematic comparison amongst vibrational descriptors used in computational models applied for the simulations of vibronic spectra (VG, VH, dAH, and AH)

To have a deeper understanding of the unexpectedly bad performance of AH simulations, the results of the VH and dAH approaches were examined. It turned out that for the studied set of compounds the VH approach (Figure 24) reproduces spectra simulated by VG. Comparing technical aspects of these simulations, in this study VG was superior to VH because *i)* VH is more computationally expensive and requires calculations of the excited state Hessian; and *ii)* for all 20 compounds studied here VH excited state Hessians (computed at ground state geometries) exhibit imaginary frequencies, and setting imaginary frequencies into real numbers might lead to biased result. Therefore, as far as VG approach avoids this assumption, it seems to be more robust than VH.

The VG and VH approaches use VG geometry to describe the equilibrium geometry of the excited state that is different from the geometry at which the excited state Hessian was computed. To be consistent, the dAH approach also was verified. It is closer to the AH method since the Hessian used in the excited state corresponds to the geometry used in the excited state. The difference between these two approaches is that AH requires S_1 geometry, while dAH uses VG geometry. This method is also known as the adiabatic Hessian after step (AHAS) [280], but in this thesis, a different label is used to highlight that an in-house code was applied to perform such calculations. The spectra simulated by dAH are shown in Figure 25, and it demonstrates that dAH method reduces the unphysically wide absorption bands as compared with the AH method. This result confirms the statement which relates the poor performance of the AH method to significant differences between the ground and excited state geometries.

The previous spectra were simulated with the same Gaussian broadening (HWHM=100 cm^{-1}) aiming at the comparison of vibronic structures computed at different levels of theory. In the next step, the comparison of simulated spectra with experimental data was performed. For this task, the HWHM was selected individually for each compound to make the final FWHM of the simulated absorption band match the FWHM of this absorption band in experimental spectra. The compound **A-6** was excluded from these studies because its experimental spectrum has a mixture of *cis*- and *trans*-isomers. The compound **E-4** was also removed from the further discussion since it has a bright excitation to the S_2 state, and this band overlaps with the S_1 absorption band. Therefore, their separation would require deconvolutions that rely on certain assumptions. The comparison between experimental spectra and VG results is shown in Figure 26. The experimental spectra of the compounds from the set **A** mostly do not show vibrationally-resolved absorption bands. Therefore, the quality of the selected theoretical approach might be evaluated better by looking at the set **B** and

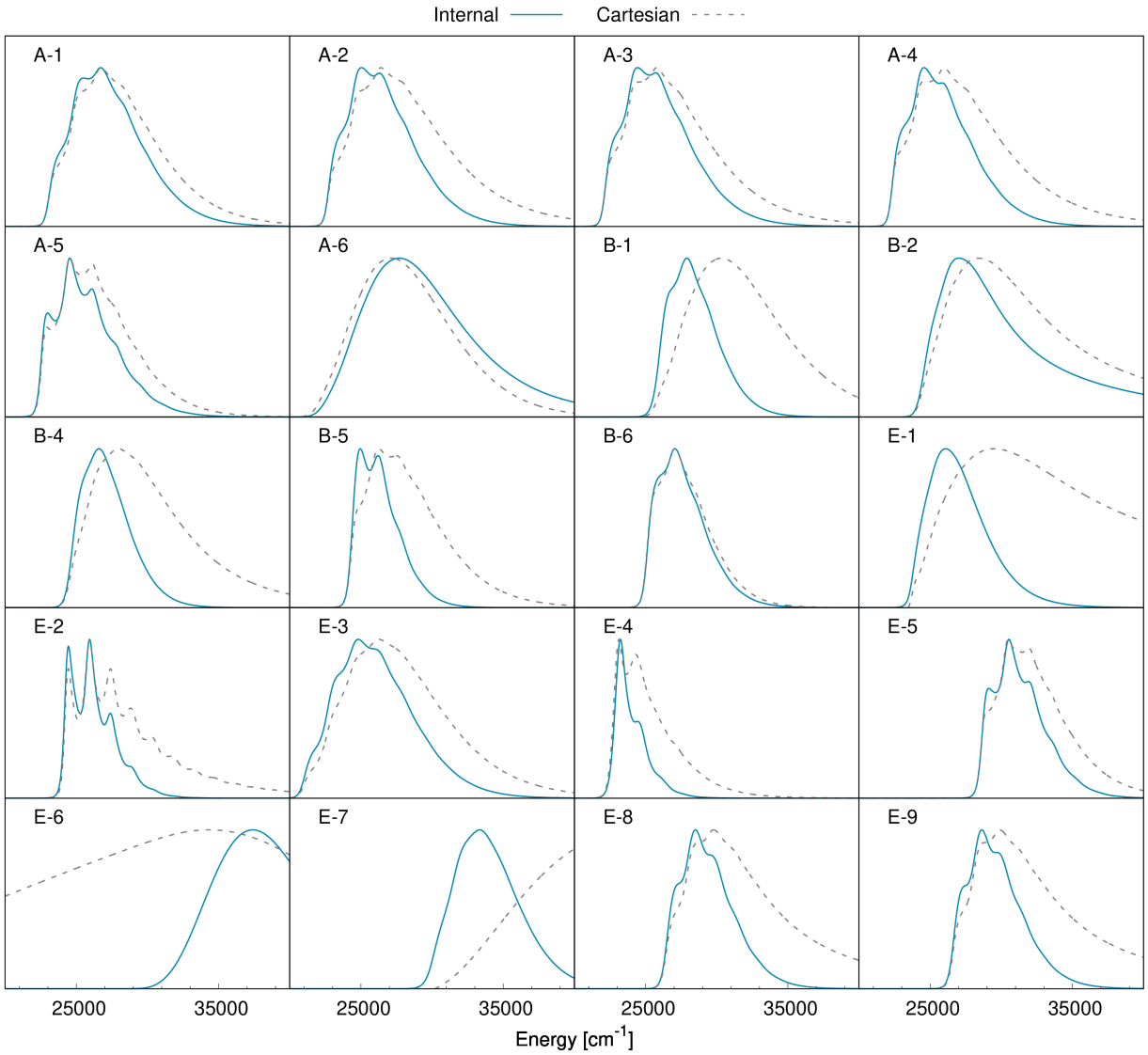


Figure 22: Comparison between internal and Cartesian coordinates in simulations of vibrationally-resolved absorption spectra using AH model based on LC-BLYP-OT(α) electronic structure calculations in chloroform solution

the set **E**. For these two sets the theory has a good consistency with the experiment. The intensity of the 0-0 transition is slightly overestimated (due to the description of the excited state with the VG geometry, which is rather similar to the S_0 geometry than to the real S_1 geometry). The comparison between the AH spectra and the experiment is problematic because even with unphysically small broadenings (HWHM$<$50 cm^{-1}) simulated absorption bands remain too wide.

It is worth mentioning that for the compound **E-7**, which is the smallest compound studied in this work, the electronic structure calculations needed for the VG model (ground geometry optimization, calculations of Hessian and calculations of excited state gradient on the ground state geometry) took 11 times less CPU time than the electronic structure calculation required for the AH model (ground and excited state geometry optimizations and evaluation of the ground and excited state Hessian). Thus, as it was shown in this section for the studied fluorescent dyes, the VG approach demonstrates a good agreement with the experiment, while the AH approach requires 10 times more computational resources and leads to unreliable results.

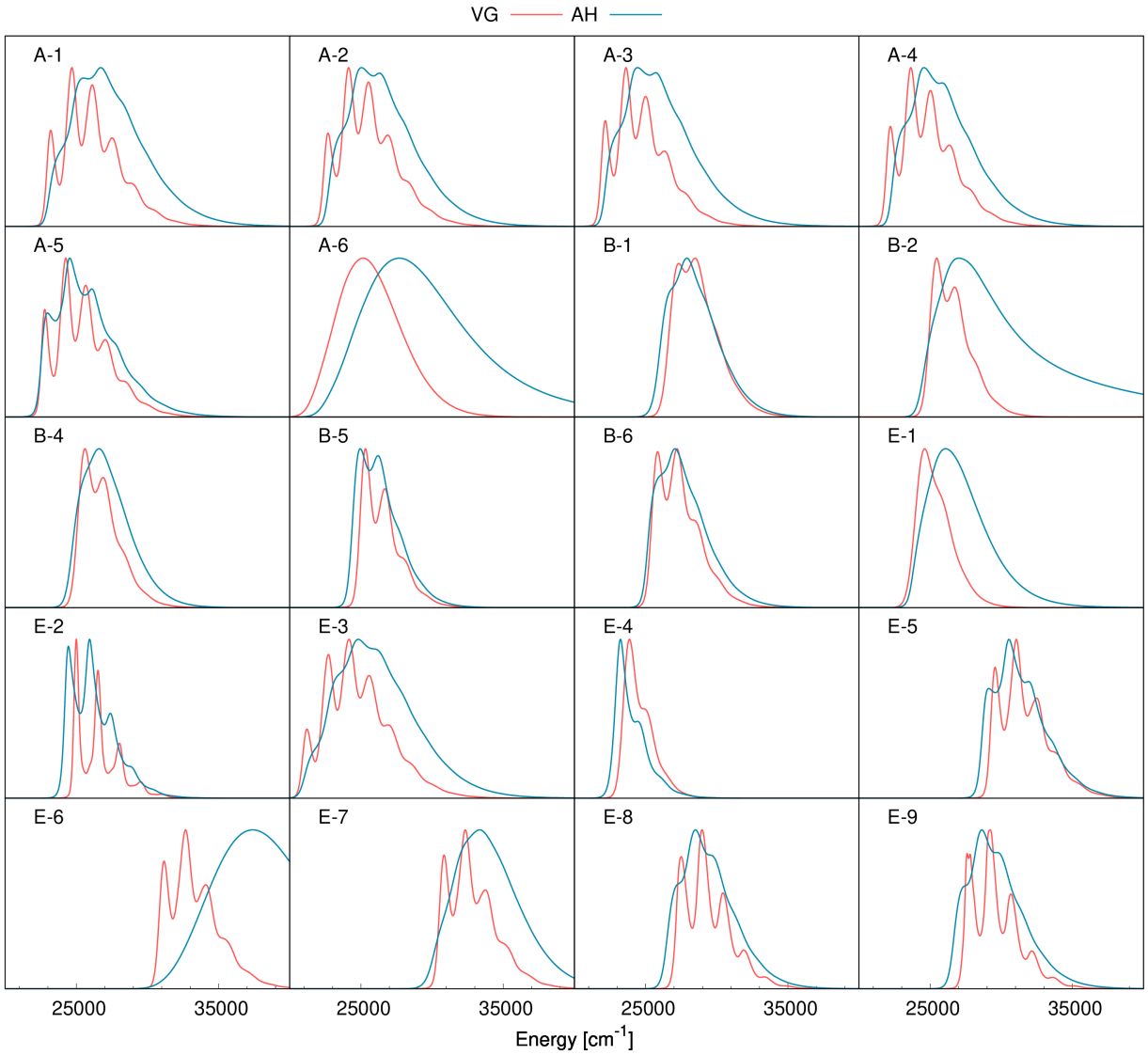


Figure 23: Comparison between AH (internal coordinates) and VG models in simulations of vibrationally-resolved absorption spectra based on LC-BLYP-OT(α) electronic structure calculations in chloroform solution

To summarize, the aim of the above analyses was the assessment of computationally accessible methods, namely, the combination of DFAs for electronic structure calculations and harmonic approaches for simulations of the vibronic structure. To this end, efficiencies of multiple DFAs were compared using the vibrational reorganization energy as a handy metric and taking gas-phase RI-CC2 values as references. As a result, ωB97X, LC-BLYP, and LC-BLYP-OT(α) showed the smallest errors. The comparison between adiabatic and vertical models results in the conclusion that the computationally cheap VG approach is more effective than the AH model, which fails due to the formation of unphysically long vibrational progressions of FC factors due to significant differences between ground and excited state geometries. Using the recently developed LC-BLYP-OT(α) functional and the VG approach one can achieve quite accurate vibronic spectra using relatively small computational resources, which, taking into account the excellent performance of this functional in the calculations of highly accurate second hyperpolarizabilities and two-photon absorption strengths, is important for the further development of the robust protocol for preselection of probes for 2PA

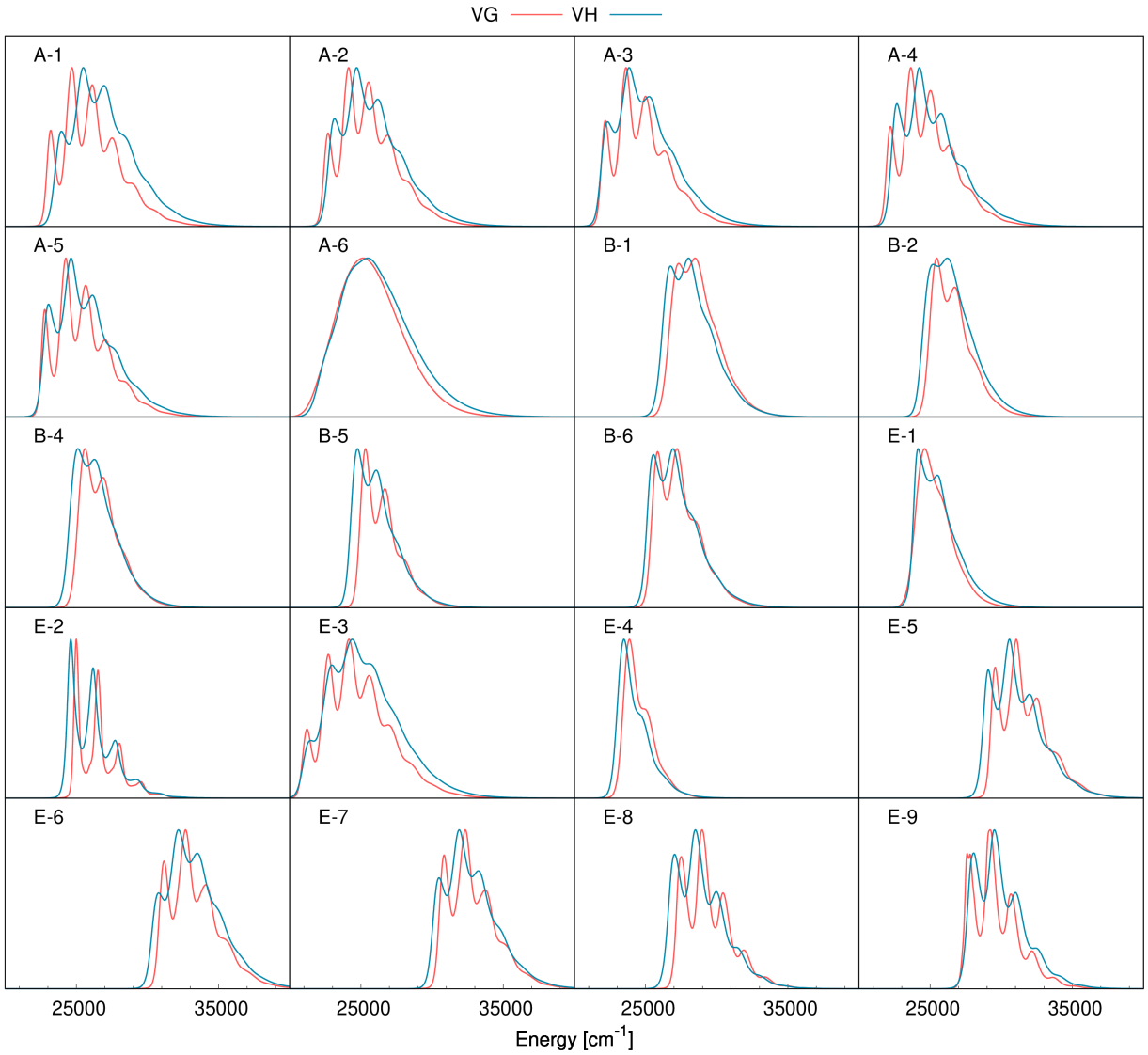


Figure 24: Comparison between VH (internal coordinates) and VG models in simulations of vibrationally-resolved absorption spectra based on LC-BLYP-OT(α) electronic structure calculations in chloroform solution

bioimaging.

VG provides trustworthy results for most of the studied dyes. However, it might be a result of multiple error cancellations. Thus, the next step in our study towards the improvement of simulation protocol was to go beyond harmonic FC approximations and analyze the weight of the anharmonic corrections and HT effects. In these studies the AH model was used for the description of the vibronic structure of absorption bands. For small molecules (up to 10 atoms) calculations taking into account anharmonic couplings are shown to be very accurate in the theoretical prediction of absorption spectra [281–283]. Thus, it can be a useful tool for the prescreening spectral properties with high reliability. However, the inclusion of the anharmonic contributions requires a high computational cost. In what follows we will discuss the results of calculations of anharmonic vibronic spectra for smaller prototypes than the 2PA probes studied above. The set selected for these studies contains 4 molecules and it is shown in Scheme 11. The compound **F-1** is a simplistic difluoroborate, and the structures **F-2**, **F-3** and **F-4** were formed from **F-1** by

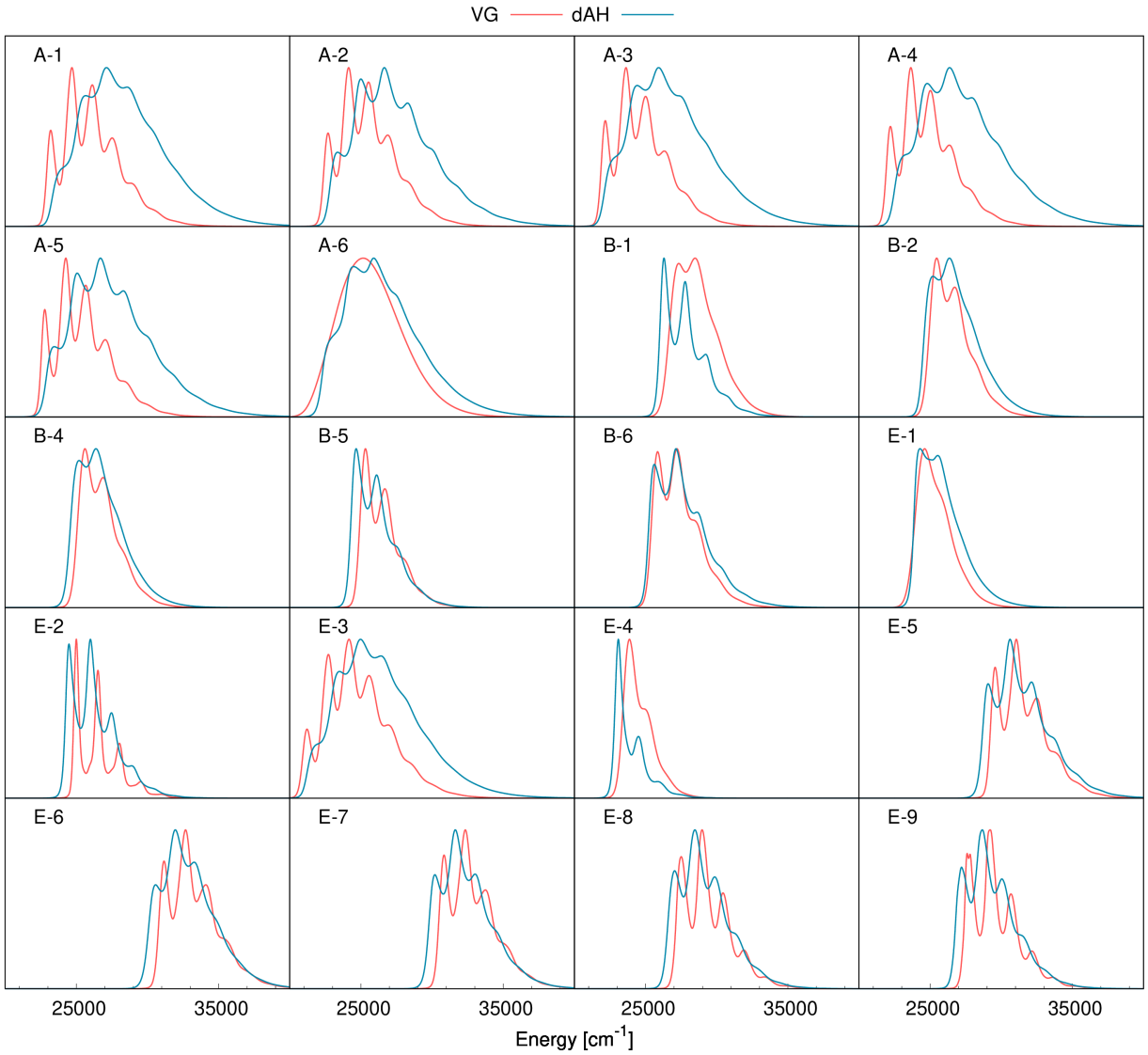


Figure 25: Comparison between dAH (internal coordinates) and VG models in simulations of vibrationally-resolved absorption spectra based on LC-BLYP-OT(α) electronic structure calculations in chloroform solution

benzannulation in all three possible positions. The chosen structures are small compared to the "realistic" bioimaging probes containing functional substituents and long π-bridges as the ones described in the series **A** – **E**. The bright $S_0 \rightarrow S_1$ transition of these four molecules will be discussed below.

Since the hardware requirements for full-dimension anharmonic calculations are difficult to meet for these 4 compounds, PT2 approximation implemented to the FC Code program [221] was used in this study. This approach allows the simulations of anharmonic contributions for a subset of selected modes amongst all vibrational modes of the compound. A summary of the theoretical derivation of this approach was given above (pages 25-28). For example, in the studies of the furan photoelectron spectrum (9 atoms, 21 vibrational modes), only 4 modes are required to accurately reproduce the spectrum [284]. For this reason, considering that we are dealing with two to three times bigger molecules (26 atoms, 72 vibrational modes), we will focus on calculations of anharmonic contributions for some selected modes only. Later we will apply these contributions to the full harmonic FCF spectrum. However, even for selected relatively

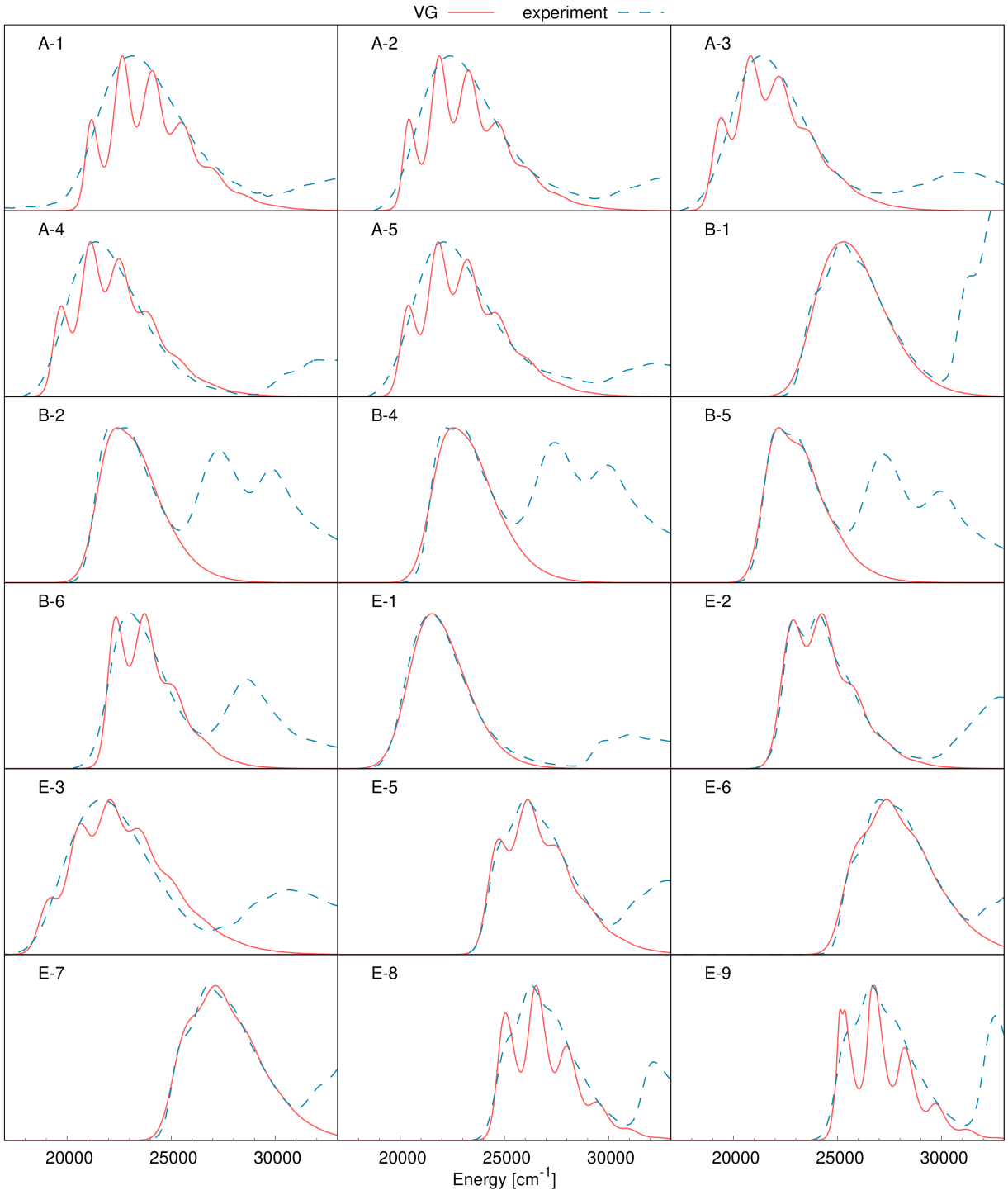
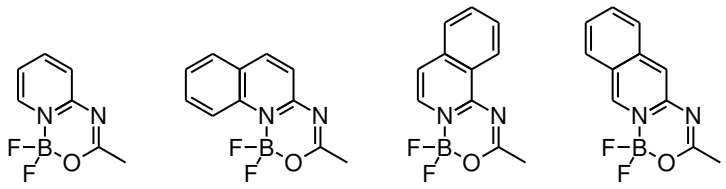


Figure 26: Comparison between experimental spectra (measured in chloroform) and vibrationally-resolved absorption spectra computed using VG model and LC-BLYP-OT(α) functional in chloroform solution

small fluorescent dyes, simulations required significant readjustment of the FC Code algorithm and code. In this section, we will explain step by step the computational protocol developed for the simulations of anharmonic corrections computed in this thesis. This protocol is shown in Figure 27.



F-1 **F-2** **F-3** **F-4**

Scheme 11: Studied compounds of series **F**

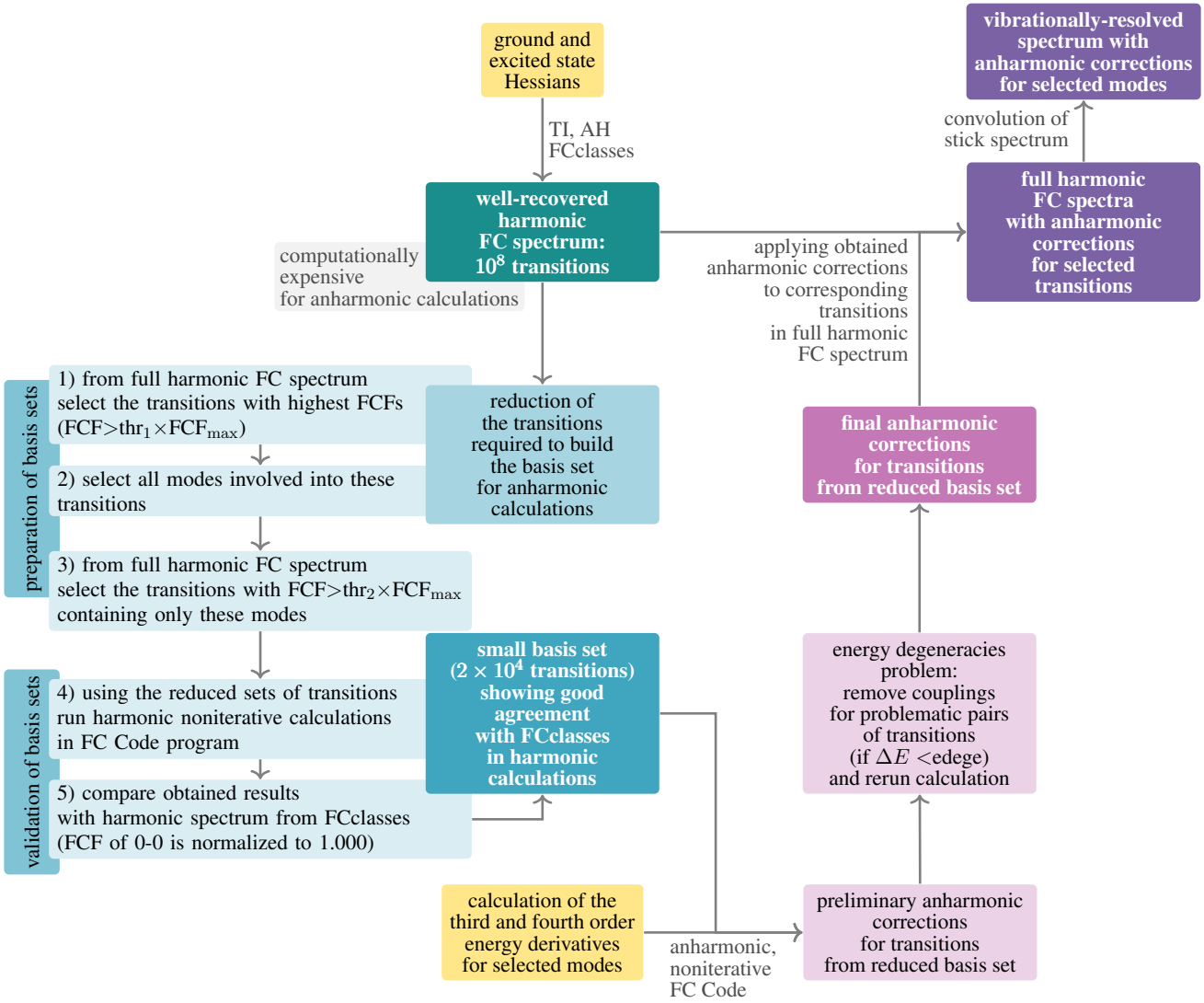
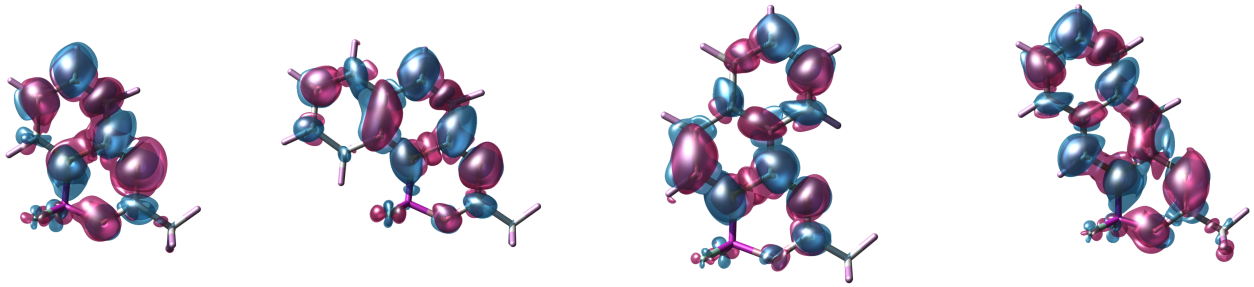


Figure 27: The protocol developed for simulations of vibronic spectra with anharmonic contributions. FCF$_{max}$ denotes the largest FCF presented in the harmonic FC spectrum, thr$_1$ denotes the threshold for selection of important modes, thr$_2$ denotes the threshold for selection of the transition. The reduction of the transitions is performed by selection of optimal thr$_1$ and thr$_2$ parameters

The first task (marked with yellow color) was the calculation of the ground and excited state Hessians, and cubic and quartic force constants. Gaussian 16 program was used for this step [224]. These calculations were performed at the LC-BLYP/aug-cc-pVDZ level of theory. For the current studies, the LC-BLYP functional was selected for its good performance in the calculations of cubic and quartic force constants [273], and because it does not suffer from spurious oscillations [179]. Moreover, as it was discussed earlier (Fig. 17), at the harmonic level the LC-BLYP approach provides reliable vibrational reorganization energies for a wide palette of BF_2-containing molecules. The ground and the first electronic state geometries were optimized with further evaluation of corresponding Hessians. The simulations were performed in the gas phase to avoid additional numerical errors generated by solvent models (e.g., cavity changes upon nuclear displacements might lead to significant errors in numerical normal mode derivatives). The Hessians were calculated at a set of displaced geometries (optimized geometries were displaced along each normal mode with a 0.02 a.u. step) to determine the higher-order derivatives. Furthermore, the electronic density difference plots for representing the $S_0 \rightarrow S_1$ transition difference of densities ($\Delta\rho(\boldsymbol{r}) = \rho^{S_1}(\boldsymbol{r}) - \rho^{S_0}(\boldsymbol{r})$)) were computed using MN15/aug-cc-pVDZ level of theory with Gaussian program [224]. The obtained plots are given in Figure 28. They confirm that the studied compounds have the same character of $S_0 \rightarrow S_1$ transition, i.e. $\pi \rightarrow \pi^*$.



F-1 **F-2** **F-3** **F-4**

Figure 28: Series **F**, electronic density difference plots corresponding to the $S_0 \rightarrow S_1$ transition. All the plots were prepared with the same density contour value (0.001)

In the previous studies at the harmonic level the TD approach was the best election. However, for simulations of the anharmonic corrections, we need to know the assignments of FCFs to individual transitions. For this reason, harmonic spectra were computed using AH approximation in TI formalism at 0 K temperature. The harmonic calculations of the FCF set that give trustworthy absorption spectra are challenging because the total number of possible transitions is infinite. Therefore, a protocol for simulations of vibronic spectra requires a reliable algorithm for selecting which transitions should be considered. For this reason, the FCclasses 3 program was used for initial calculations of full harmonic spectra. FCclasses does not support anharmonic calculations, but the program permits the calculation of harmonic FC spectra in a cost-effective manner [274]. The resulting harmonic FCF spectra are shown in Figure 29. The structure **F-2** has the simplest spectrum amongst other ones studied here, since it includes only 20 significant FC factors, and the shape of the final absorption spectra (obtained after convolution with broadening function) is formed mainly based on these transitions. On the contrary, for compounds **F-1**, **F-3** and **F-4** FCF spectra are very complex. Looking at their obtained FCFs one can not predict the appearance of the final absorption spectra since their shape will not be determined based on a few dominant vibronic transitions, but based on the overall distribution of

FC factors. Moreover, to obtain the converged FC spectrum (with the sum of FCFs > 0.99) for compound **F-2** it is enough to compute 10^8 FCFs, while for other compounds from the set one needs more than 10^{12} FCFs. Therefore, the assessment of anharmonic contributions will be limited to the studies of **F-2** molecule, since it is the one with the least complicated distribution of FCFs amongst investigated difluoroboranes. However, even this molecule has a very complex vibrational absorption spectrum if it is compared with previously studied water and furan photoelectron spectra.

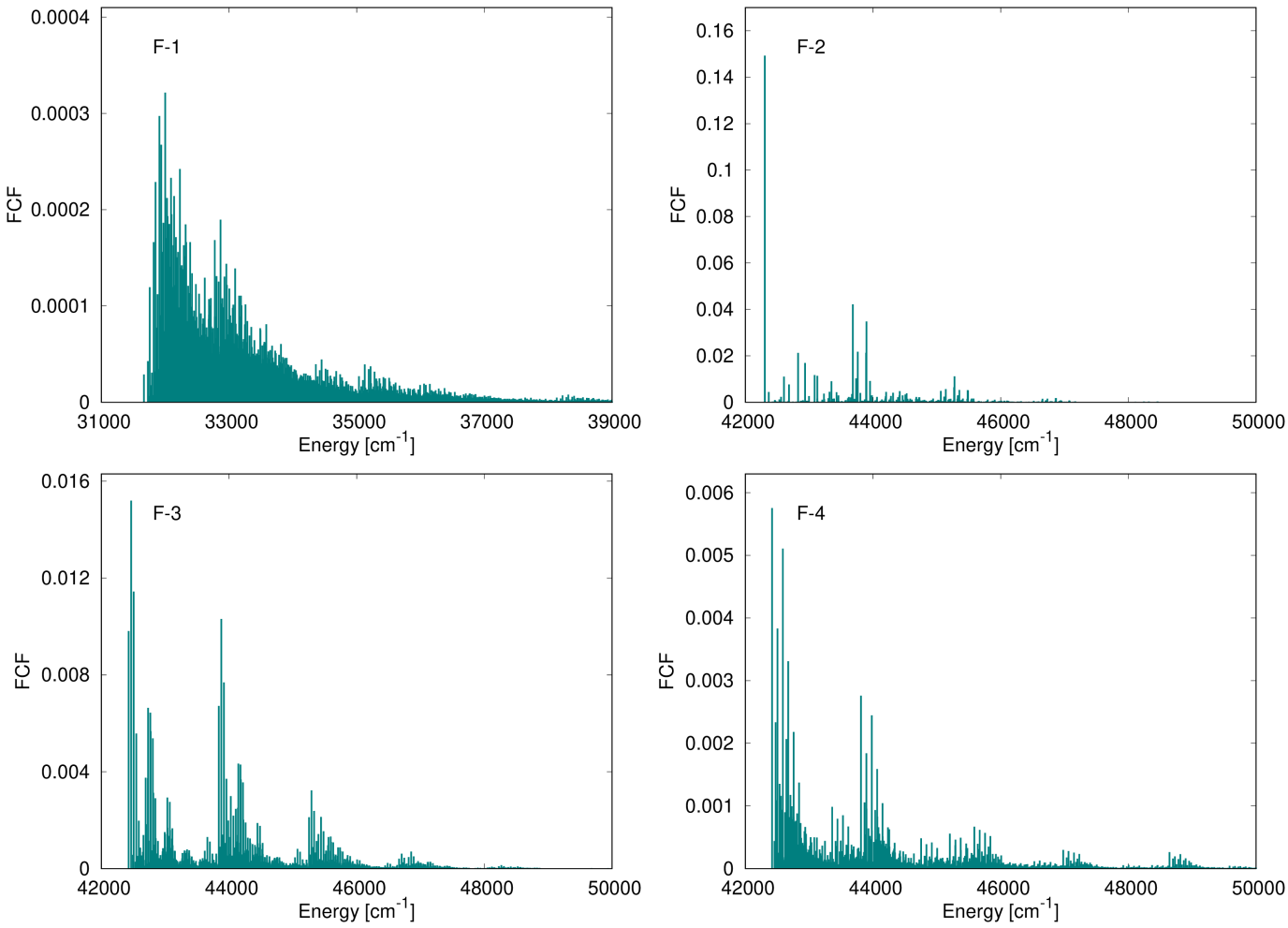


Figure 29: Series **F**, FC spectra calculated in the gas-phase using AH harmonic approximation

For calculations of anharmonic contributions, it was planned to use the FC Code program, which supports calculations only in Cartesian coordinates. It is worth mentioning that, as demonstrated above for harmonic simulations, for the medium-size molecules, the AH approach works better with internal coordinates than with Cartesian ones. For this reason, an additional comparison of spectra obtained by VG and AH (in both Cartesian and internal coordinates) was performed. From Figure 30 it is clear that in the case of the **F-2** compound, internal coordinates and Cartesian coordinates harmonic calculations are in good agreement, and therefore further calculations of anharmonic contributions in Cartesian coordinates are reliable.

Previous studies for smaller molecules showed that to obtain the correct vibronic spectrum one does not need to consider all modes. It is enough to calculate the FC spectrum based only on the few selected modes that have the

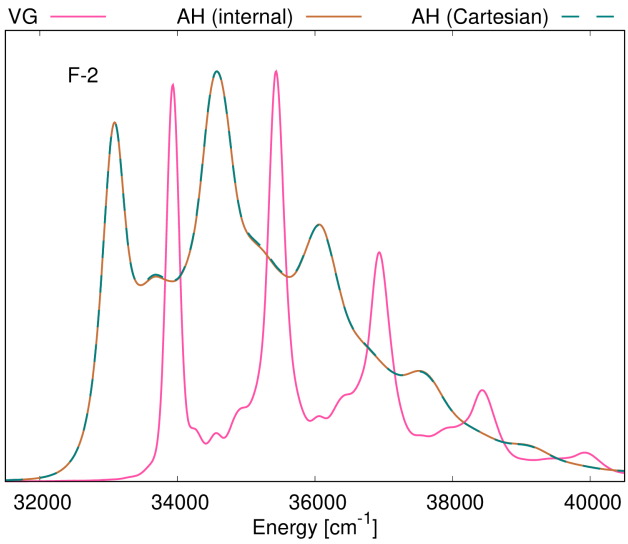


Figure 30: Compound **F-2**, comparison between VH and AH models (using TD approach) in simulations of vibrationally-resolved absorption spectra based on LC-BLYP electronic structure calculations in the gas phase

biggest influence on the final spectrum. From the full set of harmonic FCFs, the most important modes with the highest FCF values were selected. To simplify the FC Code algorithm for the selection of excitations, which is quite time-consuming even in vibrational space limited to 19 modes, a new non-iterative procedure was developed. This new procedure relies on a subset of transitions that contain only the excitations of some selected modes from the full set of harmonic FCFs. The anharmonic calculations were performed using only this selected basis set of transitions. Considering that the FC Code uses a truncated basis set for the estimation of anharmonic corrections, we just changed the code to use our selected basis set, which is read as an input file, instead of using the truncated basis set selected by the FC Code algorithm. The technical details concerning the optimization of the basis set of transitions are given in Figure 27.

For the compound **F-2**, the selected basis set of transitions contains 19 modes and 23905 transitions (amongst the 10^8 transitions required for the converged FCF spectrum). For this basis set the anharmonic contributions were calculated using PT1 and PT2 methods. The calculations required 77 Gb of RAM and took 78 hours (Intel Xeon Platinum 8268 (2.9 GHz)). The high usage of RAM is the most significant limitation of the current approach, confirming that for more complicated cases simulations of anharmonic contributions will require the improvement of the current algorithm.

Resulting PT1 and PT2 FCF spectra including anharmonic contributions are given in Figures 31 and 32. The two upper plots of Figure 31, show that each negative PT1 FCF value has a positive twin with approximately the same energy (for example, see the highlighted peaks **u** and **t**, **w** and **v**, and **x** and **y**). The calculations of perturbation theory corrections may show spurious huge corrections due to very small energy denominators corresponding to the difference between two quasi-degenerate transitions (see $E^{e(0)}_{\mu_e} - E^{e(0)}_{\nu_e}$ denominator in Equation 37 (page 27)). Thus, increasing the number of vibrational normal modes in a studied molecule would cause magnification of the degeneracy problem due to random degeneracies due to accidentally similar energies of excitations. As a preliminary approach, in the current studies such degeneracies were removed. If the difference in energies of two transitions is lower than a threshold named $edege$, then couplings between these two transitions are removed. For PT1 calculations, it was

enough to remove couplings between transitions with energy differences lower than 0.22 cm^{-1} (Figure 31, the plot at the bottom). PT2 calculations are more sensitive, and the $edege$ threshold should be increased to 33 cm^{-1} (see Figure 32). For the remaining three compounds **F-1**, **F-3** and **F-4** the perturbation theory can not be applied, since the energy differences between large FCFs are very small, and then removing couplings between all transitions would simply lead us back to a result very close to the harmonic approach.

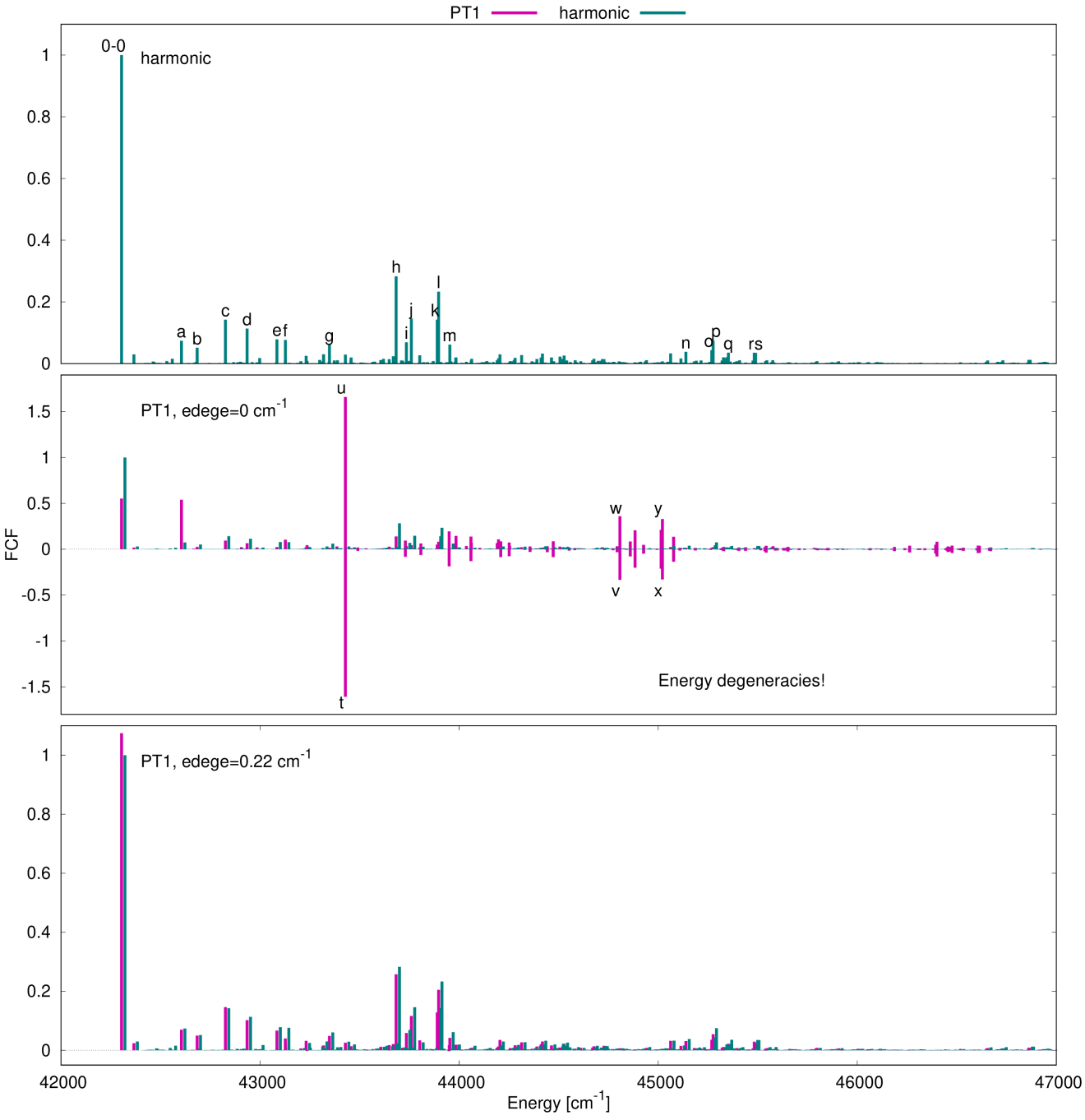


Figure 31: Compound **F-2**, harmonic and first-order anharmonic (PT1) FC spectra

The towering result of both PT1 and PT2 corrections is the increase of the FCF corresponding to the 0-0 transition,

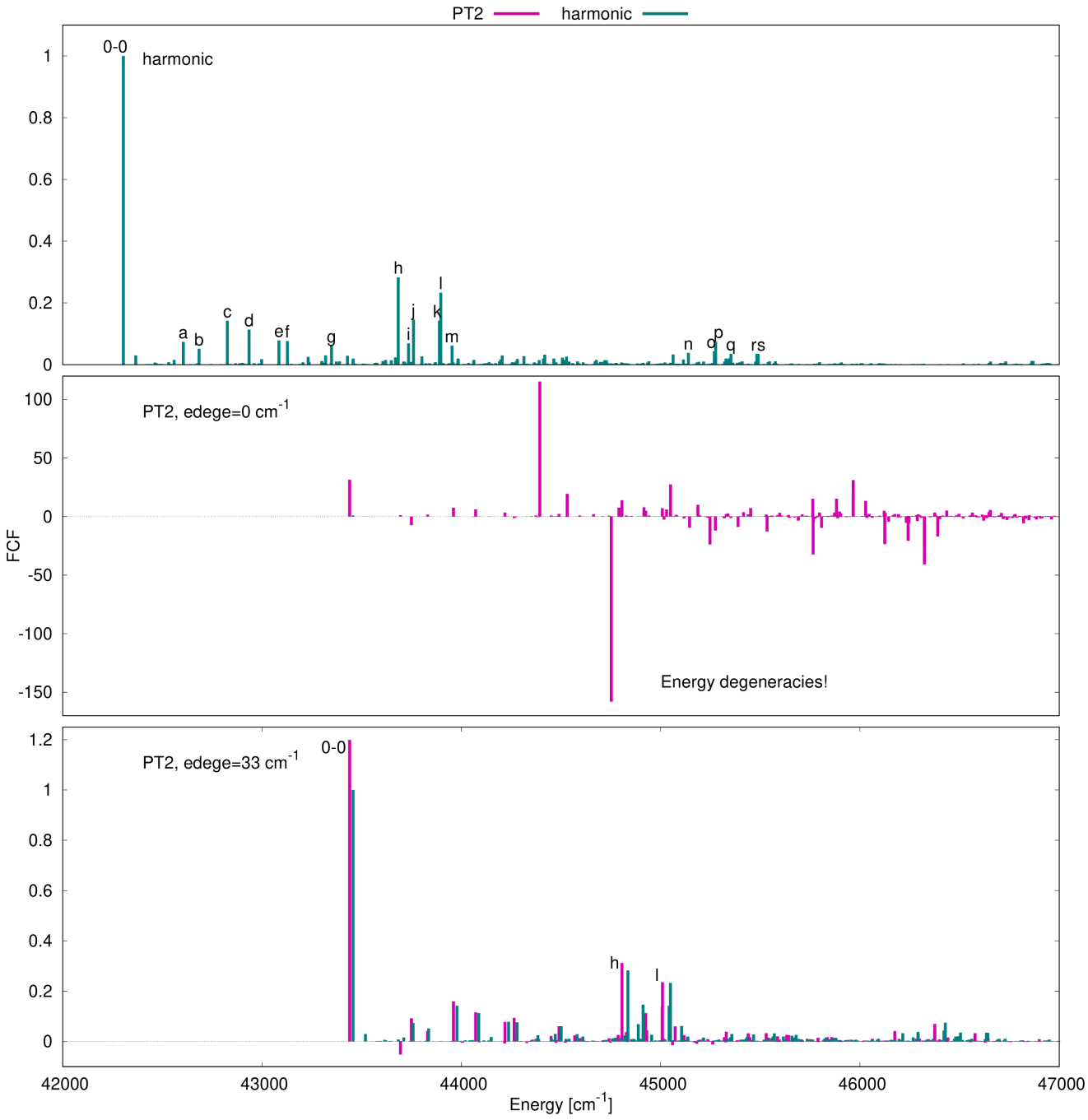


Figure 32: Compound **F-2**, harmonic and second-order anharmonic (PT2) FC spectra

while other transitions remain almost unchanged. The calculated anharmonic corrections were applied to the corresponding transitions of the full harmonic spectrum (10^8 transitions), and the obtained FCFs were convoluted with the arbitrary value of broadening allowing to match FWHM of the experimental spectrum. Applied at this step transition energies correspond to the harmonic model. The result of this approach is presented in Figure 33. Indeed, the proposed protocol to include the most important effects of the anharmonic corrections leads to improvements in the comparison with the experimental spectrum with respect to the harmonic approach. The relative intensities of the two dominant

absorption maxima calculated taking into account the PT2 corrections perfectly reproduce the experimental spectrum, while the harmonic approach underestimates the intensity of the 0-0 transition.

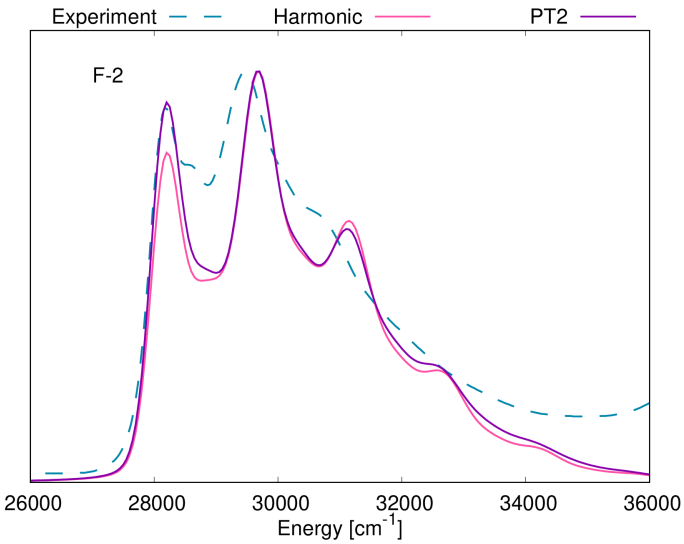


Figure 33: Compound **F-2**, comparison amongst vibrationally-resolved absorption band simulated using the developed second-order anharmonic protocol, vibrationally-resolved absorption band simulated by original harmonic approximation, and normalized experimental absorption spectrum

In summary, we have shown that for medium-size molecules the practical application of perturbation theory is challenging due to energy degeneracies caused by the large amount of overtones with very close wavenumbers. For bigger and more complex structures this problem will be amplified, and therefore the current approach should be improved. From the set of four small boron difluorates, only one compound (**F-2**) had enough simple harmonic FC spectra to compute the anharmonic corrections with the current approach, and even for this molecule it was required the removal of couplings between all pairs of transitions with energy differences up to 33 cm^{-1}. The removal of couplings between problematic transitions is leading to a reduction of the negative and unphysically large FC factors, but unfortunately, there is a chance that this procedure is also removing some important anharmonic corrections. Nevertheless, the anharmonic corrections computed just for 25% of vibronic modes notably improved the accuracy of the simulated spectrum. The next possible extension of these studies is applying the variational approach in calculations of FC spectrum with anharmonic couplings.

Now we will assess the importance of HT contributions on vibronic 1PA and 2PA spectra. Short theoretical explanations of differences between the pure FC approximation and the FC approximation extended with the HT correction were explained above (Equation 20, page 25). The estimation of the HT corrections requires the numerical computation of derivatives of the electric transition dipoles for 1PA spectra and derivatives of the second-order transition moments with respect to normal modes for 2PA spectra. For these reasons, the assessment of HT effects will be performed based on the smaller compounds from the set **F** (Figure 11). The analyses will be performed for the $S_0 \rightarrow S_1$ transition, which is bright in 1PA spectra and dark in 2PA spectra. The LC-BLYP/aug-cc-pVDZ level of theory was used, because, as it was shown previously, even at the harmonic level AH calculations based on Hessians estimated by LC-BLYP provide spectra with satisfactory accuracy for these compounds. The numerical derivatives of the transition dipole moment

(HT corrections to 1PA spectra) and the second-order transition moment (HT corrections to 2PA spectra) with respect to normal modes were computed fully numerically in Cartesian coordinates and were transformed to the normal mode basis. In doing so, calculations for 10 displaced geometries were performed for each Cartesian direction for all atoms, thus, safeguarding the high numerical stability of computed derivatives. The stability was further controlled using the Rutishauser-Romberg algorithm [285]. The total number of electronic-structure calculations for each molecule (for given HT correction type - 1PA or 2PA) was: $10\times3\times N$+1, where N is the number of atoms. The resulting absorption spectra were simulated using AH approximation in internal coordinates. The comparison between the pure FC theory and the HT model is shown in Figure 34.

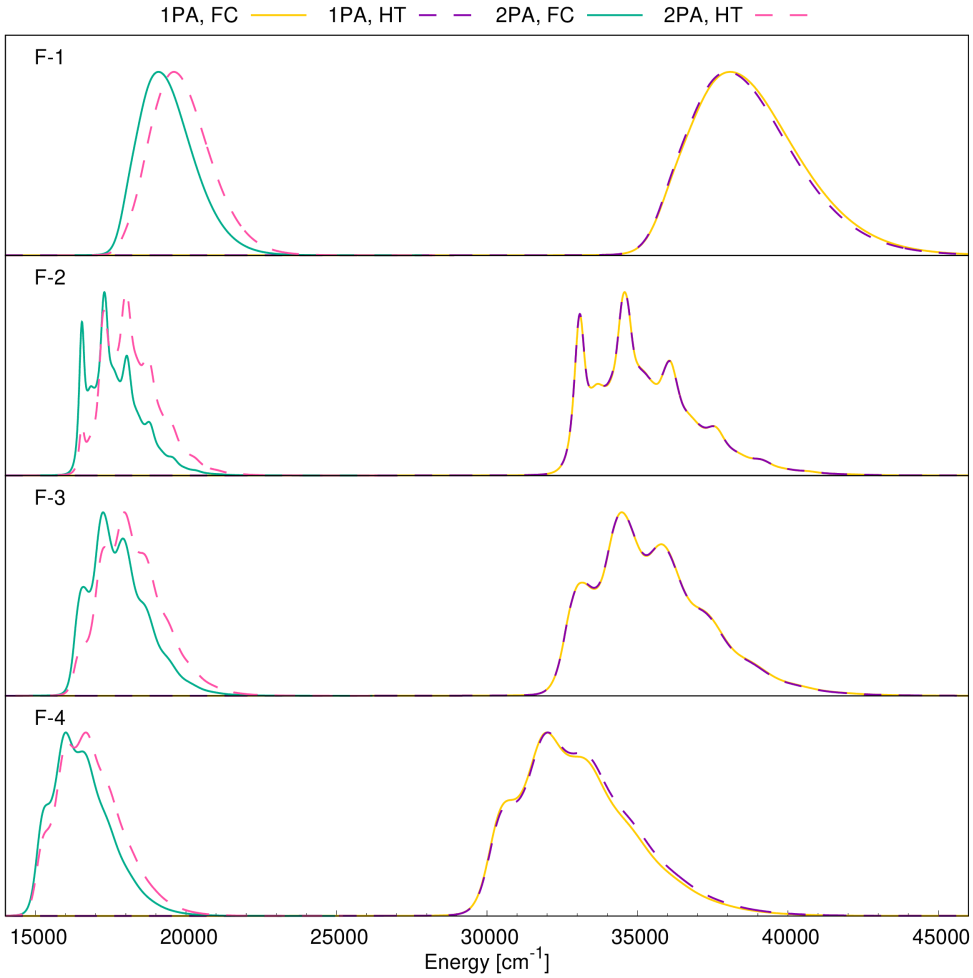


Figure 34: Series **F**, comparison between FC and HT models in simulations of vibrationally-resolved 1PA and 2PA spectra ($S_0 \rightarrow S_1$ transition) based on LC-BLYP electronic structure simulations in the gas phase and AH model (in internal coordinates) with HWHM=100 cm^{-1}

The HT effect does not provide any visible corrections to the vibronic structure of the bright 1PA spectra corresponding to the $S_0 \rightarrow S_1$ transition for all 4 dyes studied here. Numerically it can be explained by large electric transition dipoles, which do not undergo significant changes upon excitation. Then the derivatives of the electric transition dipoles along normal modes are insignificant compared to the transition dipoles themselves. The combination of

electron-donating and electron-accepting groups, which is a basic strategy in the design of 2PA dyes, would cause high electric transition dipoles for bright excitations. That is why the HT corrections will not influence their vibrationally-resolved absorption spectra. On the contrary, the computed 2PA spectra corresponding to the same $S_0 \rightarrow S_1$ transition undergo significant changes by the introduction of the HT effect, which causes a significant decrease in the relative intensity of the 0-0 transition. Despite here is illustrated the normalized spectra, it is worth mentioning that the absolute maximum intensity of the 2PA band becomes up to 4 times higher due to the HT. Thus, electronically forbidden transitions that are invisible in the FC approximation become stronger due to the HT effects. This is because values in transition moment tensor are small, and their derivatives have comparable magnitudes. Experimentally it was shown that all four dyes studied here do not exhibit intense 2PA, which is in agreement with low values in the transition moment tensor. To sum up, HT contributions do not affect the simulated vibronic structure of 1PA bands corresponding to bright transitions for studied compounds, which are of main interest for the computational tool to pre-screen the photophysical properties of dyes. However, they are important for weak 1PA transitions and they are also crucial for simulations of 2PA spectra.

5.2 Estimations of inhomogeneous broadening

In the results presented above, the broadening was simply selected to match the experimental FWHM of the studied absorption band. In the next step of this project, we tried to design a cost-effective way to simulate the HWHM. Here we examine two sets of compounds: set **B** and set **F**. Note that we exclude compound **B-3** because it has a highly solvent-dependent low-intensity $S_0 \rightarrow S_1$ transition.

In previous sections, we learned that the VG method shows the highest accuracy in simulations of the vibronic structure of studied compounds and requires the smallest computational resources. Thus, it was tempting to develop a cost-effective method for simulations of the inhomogeneous broadening. Then, by the convolution of the VG-based vibronic spectrum with the determined broadening function, it is possible to obtain a reliable vibrationally-resolved spectrum at a low computational cost. Actually, for the prediction of the inhomogeneous broadening the palette of machine learning-based methods might be reliable as one needs to correlate excitation energies with the distribution of solvent molecules around the solute. Our proposed approach is based on the ML prediction of vertical excitation energies of dyes using a discrete representation of solvent molecules used as variables. Therefore, firstly it was needed to prepare a data set based on an accurate methodology with a reliable electronic structure theory. In what follows, the accurate solvation PE model and the less computationally expensive EE model were compared.

Let us start with the discussion of the calculations performed for the **B**. Two solvents with largely different dielectric constants were selected for this investigation: less polar chloroform (ϵ=4.7) and more polar acetonitrile (ϵ=35.7). The optimal way to obtain the distribution of solvent molecules around the dye is to perform rigid-body MD simulations. Modeled systems were represented by the dye in the middle of the solvent box. Subsequently, vertical excitation energies were computed at CAM-B3LYP/cc-pVDZ level of theory for 10 000 snapshots from a MD trajectory for each dye/solvent system (in total 100 000 electronic-structure calculations were performed for set **B**). The influence of solvent molecules was considered in these simulations using the EE approach. At this stage, a large amount of calculated snapshots is required for the estimation of the reliable broadening values and the preparation of a sufficiently large data set for further development of the ML approach. Standard deviations and HWHM were computed based on these vertical excitation energies.

Figure 35 shows the dependency of the calculated standard deviation versus the number of snapshots considered. The convergence is reached at about 2000 snapshots. These results confirm the profit of further ML implementation that would significantly decrease the required computational resources. The computed vibrationally-resolved 1PA spectra were obtained by the convolution of the VG spectra obtained using LC-BLYP-OT(α) electronic-structure vibrational parameters with the computed HWHM values. The comparison between the resulting simulated absorption spectra and the experimental spectra is given in Figure 36. Since the theoretically predicted broadening functions correctly reproduce the experimental spectra, we will now explore the possibility of reducing the computational costs of this protocol by applying a machine-learning approach.

The proposed protocol requires a selection of a fingerprint representing the studied systems and a statistical model relying on these fingerprints. Nowadays there is a wide palette of various fingerprints available. From the physical point of view, considering that the EE model was used for the description of the microenvironment, the applied fingerprint should represent the geometrical configuration of solvent molecules around the dye and information about the point charges of the atoms. One of the most potentially interesting fingerprints is the Coulomb matrix $\boldsymbol{M}$ proposed in 2012,

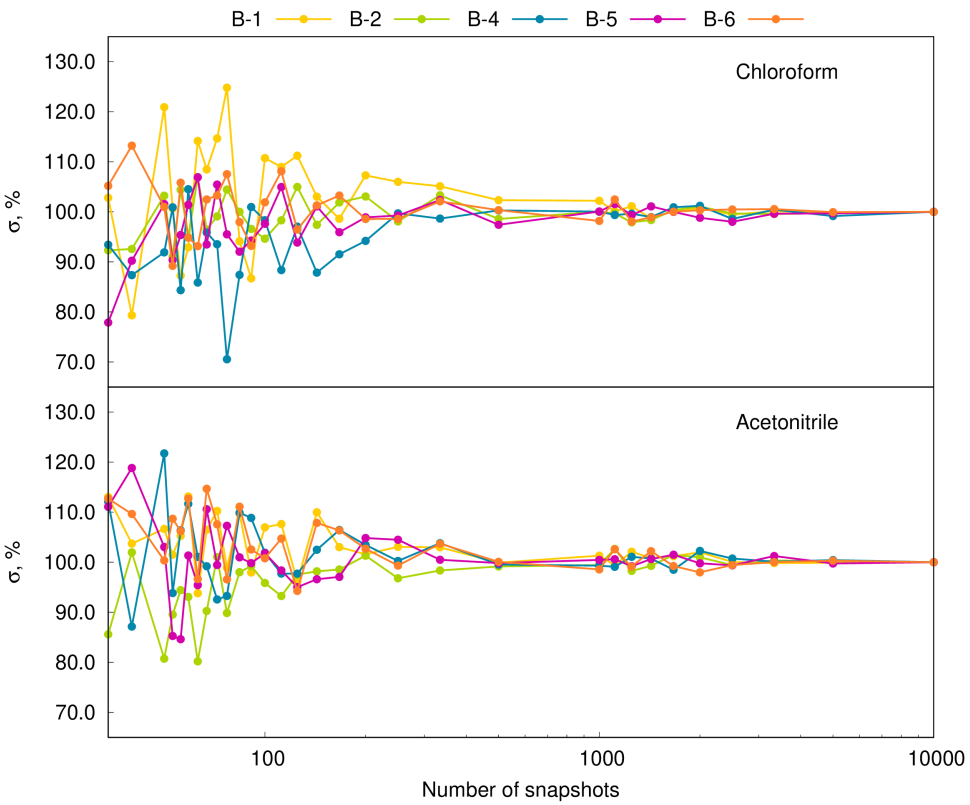


Figure 35: Series **B**, convergence of the standard deviation corresponding to the distribution of vertical excitation energy with respect to the considered number of snapshots in two solvents

which projects the three-dimensional structure of the compound. The elements of this matrix encompass the atomic number, Z_i, of the atom i, and the distance, R_{ij}, between atoms i and j [286]:

$$M_{ij} = \begin{cases} 0.5 Z_i^{2.4} & \text{for } i = j \\ Z_i Z_j / R_{ij} & \text{for } i \neq j \end{cases} \tag{45}$$

The off-diagonal elements correspond to the Coulombic interactions between all pairs of atoms, while diagonal elements encode a polynomial fit of atomic energies to nuclear charge. This fingerprint and its extensions are applied in combination with a kernel ridge regression basically as a measure of the chemical compound space similarity in the prediction of atomization energies [286, 287], the prediction of formation energies [288], the identification of molecules from their rotational spectroscopy [289], and the description of nanoparticles and their properties [290]. However, usage of the original Coulomb matrix and its published derivatives is problematic for systems of our interest. Considering that there are thousands of atoms in the solvent box, for the studied snapshots the Coulomb matrix would include billions of elements, and the simulations of so large fingerprints would require huge RAM allocations (hundreds of GB). For this reason, a new one-dimensional derivative of the Coulomb matrix was proposed by the author of this thesis which includes as many elements M_i as the number of atoms in the dye structure (about 50-70 elements in case of **B** series). One element of such fingerprint is given by:

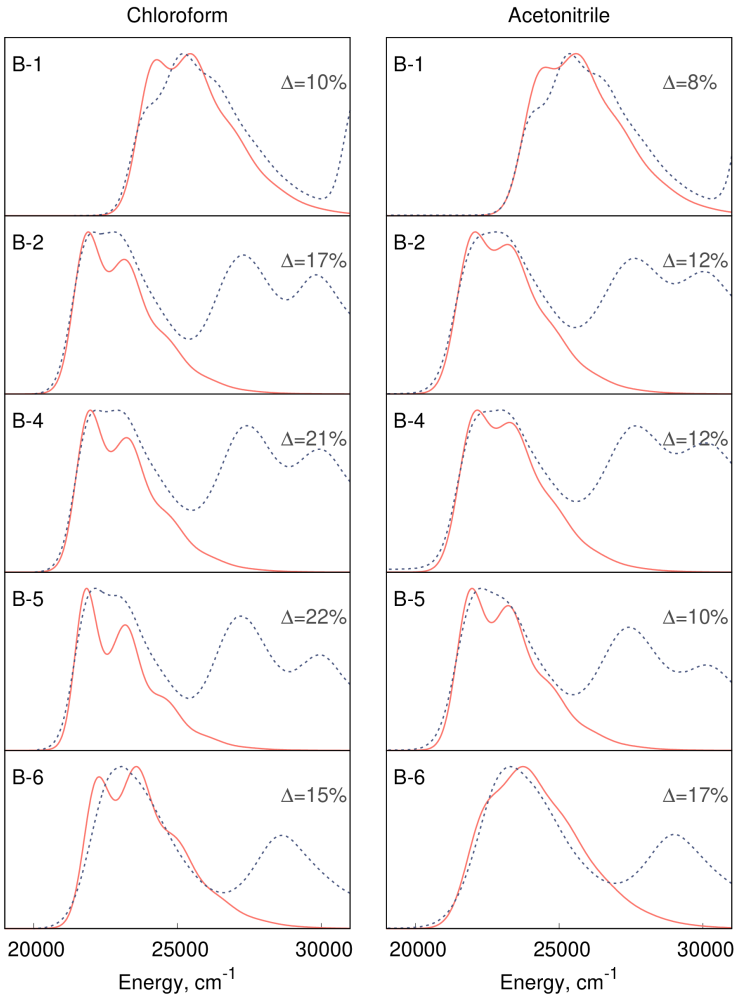


Figure 36: Series **B**, assessment of the computed inhomogeneous broadenings. Comparison between experimental 1PA spectra and computed vibrationally-resolved 1PA spectra (VG spectra convoluted with computed HWHM values) in two solvents

$$M_i = \sum_j \frac{Q_i Q_j}{R_{ij}}, \tag{46}$$

where M_i represents the interaction between the dye's atom i and all solvent atoms j; Q_i and Q_j are, respectively, the net point charges of the dye atom i and the solvent atom j; and R_{ij} is the distance between the solvent atom j and the dye atom i. Compared to the initial Coulomb matrix, the new one-dimensional fingerprint neglects intra-dye, intra-solvent as well as inter-solvent atom-atom interactions. In doing so, we only probe the fluctuations of the local microenvironment of the atomic sites of the dye. Furthermore, partial charges, as in MD simulations and further CAM-B3LYP/EE calculations, were used instead of the nuclear charges of atoms. The size of this novel fingerprint significantly decreases the RAM required for simulations and makes such ML prediction possible to run on the ordinary

laptop.

Gaussian kernel ridge regression [291] implemented into the QML package [292] was used as the statistical model to predict vertical excitation energies for the snapshots represented by proposed one-dimensional fingerprints. The applied regression model for the property y of the system $\widetilde{\boldsymbol{X}}$ can be expressed through the equation:

$$y(\widetilde{\boldsymbol{X}}) = \sum_{i=1}^{n} \alpha_i k(\widetilde{\boldsymbol{X}}, \boldsymbol{X}_i) \tag{47}$$

where α_i stays for regression coefficients, $\boldsymbol{X}_i$ stays for the fingerprint of each element of the training set, n is the number of training samples, and $k(\widetilde{\boldsymbol{X}}, \boldsymbol{X}_i)$ corresponds to the Gaussian kernel function:

$$k(\widetilde{\boldsymbol{X}}, \boldsymbol{X}_i) = \exp\left(-\frac{\|\widetilde{\boldsymbol{X}} - \boldsymbol{X}_i\|_2^2}{2\sigma^2}\right) \tag{48}$$

where σ determines the width of the Gaussian kernel, which can be obtained empirically by additional tests. Based on the training set one can estimate the Gaussian kernel matrix $\boldsymbol{K}$ for further calculations of regression coefficients through the kernel matrix inversion and multiplication with reference y value:

$$\boldsymbol{\alpha} = (\boldsymbol{K} - \lambda \boldsymbol{I})^{-1} \boldsymbol{y} \tag{49}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix, and $\boldsymbol{K}$ is the kernel matrix with elements $K_{ij} = k(\boldsymbol{X}_i, \boldsymbol{X}_j)$. Once the training of the ML approach is done, obtained regression coefficients can be applied for the prediction of properties y based on Equation 47. In our studies y corresponds to the vertical excitation energy of the snapshot, and $\widetilde{\boldsymbol{X}}$ is represented by the fingerprint (Equation 46) of the considered snapshot.

A combination of the developed fingerprint with the kernel ridge regression was used for the prediction of vertical excitation energies. A set of randomly selected 200 snapshots was used to train the proposed ML model for the further prediction of the vertical excitation energies of 9800 snapshots. Figure 37 is a scatter plot comparing energies computed by DFA and predicted by ML. Based on this Figure one can conclude that the proposed method is highly accurate since the energies predicted by ML approach reproduce vertical excitation energies from electronic structure calculations with mean absolute percentage error (MAPE) smaller than 0.2%. Table 9 includes the information about MAE, MAPE, and calculated broadenings. The results show that the proposed approach can predict the HWHM in a wide range of values: starting from 70 cm^{-1} up to 600 cm^{-1}, with the absolute error never exceeding 8 cm^{-1}. Then, the spectra convoluted with broadenings predicted by ML are quasi-identical to the simulated spectra from Figure 36.

To sum up, for the proposed ML approach it is needed to run MD simulations, and then extract more than 2000 snapshots from this trajectory and divide them into two subsets:

- the training set formed by about 200 snapshots, for which vertical excitation energies will be calculated using quantum chemistry methods;
- the prediction set formed by the rest of the snapshots, for which only the geometrical parameters are required.

The vertical excitation energies of the prediction set will be simulated by ML approach. Note that in the classic non-ML approach one has to calculate vertical excitation energies using QM/MM methods for the whole set of snapshots. The

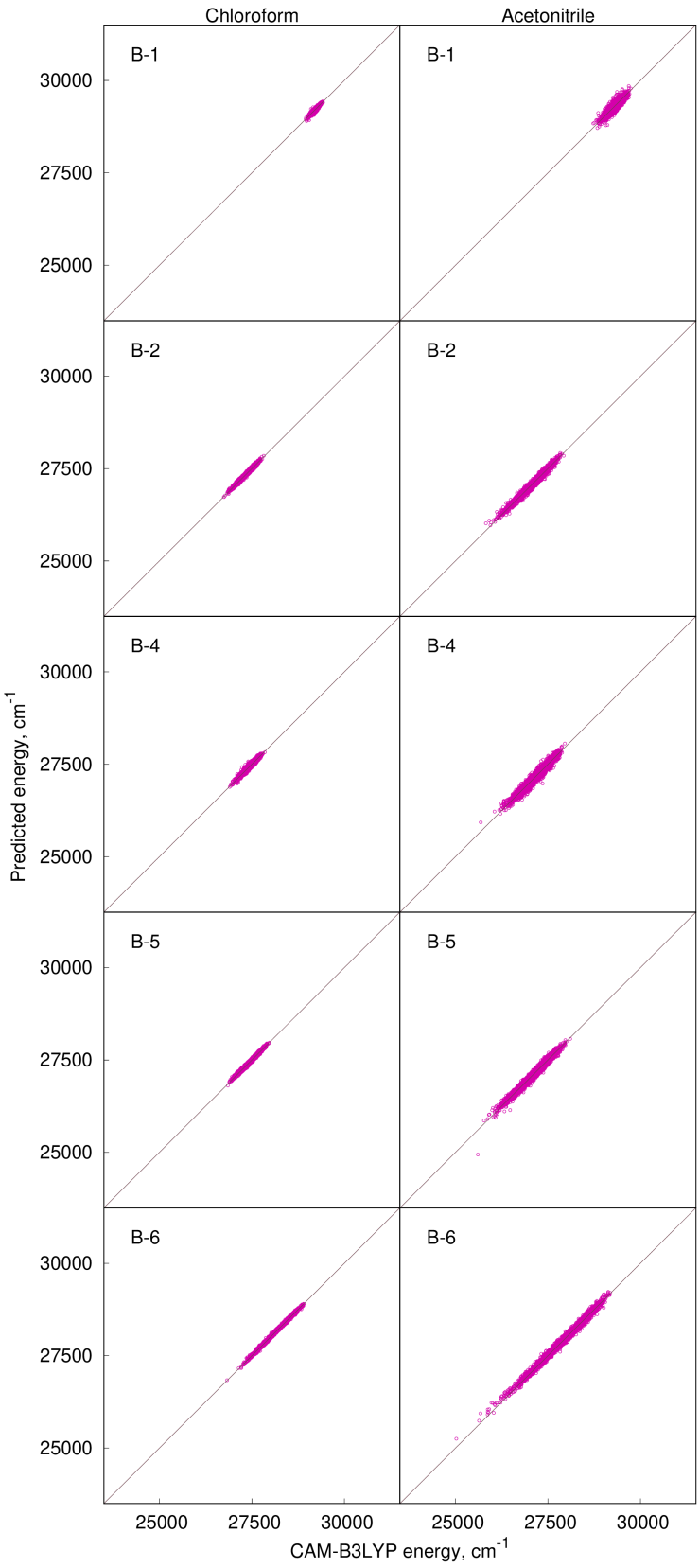


Figure 37: Series **B**, assessment of the developed ML approach. Scatter plot of the calculated by CAM-B3LYP/EE (*X*-axis) and predicted based on the developed ML model (*Y*-axis) values of vertical excitation energy in two solvents

Structure	MAE, cm^{-1}	MAPE, %	TD-DFT/EE		ML	
			σ^{inh}, cm^{-1}	HWHM, cm^{-1}	σ^{inh}, cm^{-1}	HWHM, cm^{-1}
Chloroform						
B-1	13	0.05	56	74	58	76
B-2	15	0.06	137	180	138	182
B-4	21	0.08	122	161	121	159
B-5	14	0.05	162	213	163	215
B-6	13	0.05	247	325	247	325
Acetonitrile						
B-1	41	0.14	123	162	129	170
B-2	32	0.12	268	353	270	356
B-4	43	0.16	246	324	252	332
B-5	35	0.13	303	399	304	400
B-6	35	0.12	463	610	460	606

Table 9: Series **B**, summary of ML-based prediction of the inhomogeneous broadenings. Standard deviation (σ^{inh}) and calculated HWHM correspond to the distribution of the vertical excitation energies

most time-consuming step in both classic and ML approaches is the calculations of excitation energies, while ML step itself takes just a few seconds. Therefore, using ML approach as described above we could decrease the CPU time, which is needed for estimation of vertical excitation energies for 10000 snapshots, by a factor of 50.

I will now focus on the assessment of inhomogeneous broadenings obtained by two solvent models, namely PE and EE. The results presented below correspond to the chloroform solution. In what follows, the quantum part will be described at the RI-CC2 level of theory to retain the high accuracy of the PE solvation model. The set **F** was selected for this study because it contains both polar and non-polar structures. We need a set formed with small molecules because RI-CC2/EE and RI-CC2/PE calculations require significantly higher timings than CAM-B3LYP simulations described above. It was assumed that for non-polar structures the inhomogeneous broadening would be improved using the more accurate PE model, while for polar molecules the inhomogeneous broadening obtained using both models would remain approximately constant. Additionally, the novel ML approach would be also tested for these compounds.

The estimation of the inhomogeneous broadenings was performed based on the same steps as for the set **B**: *i)* rigid-body MD simulations of the dye inside the solvent box; *ii)* calculation of vertical excitation energies for 5000 snapshots/dye with solvent involved by the PE or EE model, and *iii)* analysis of obtained results. As it was shown in the previous studies (Figure 35), to obtain the converged inhomogeneous broadening it is enough to use 2000 snapshots. That is why for the current study calculations of inhomogeneous broadenings were performed based on 5000 snapshots/structure. Figure 38 confirms that, as for the set **B**, to obtain the reliable broadening it is required to know vertical excitation energies of at least 2000 snapshots/structure. This conclusion is valid for both the EE and PE solvation models. Based on vertical excitation energies computed by RI-CC2/EE, ML prediction of inhomogeneous broadenings was performed. Figure 39 includes scatter plots demonstrating that the proposed ML approach (the new

one-dimentional fingerprint and kernel ridge regression) works well also for this set of dyes.

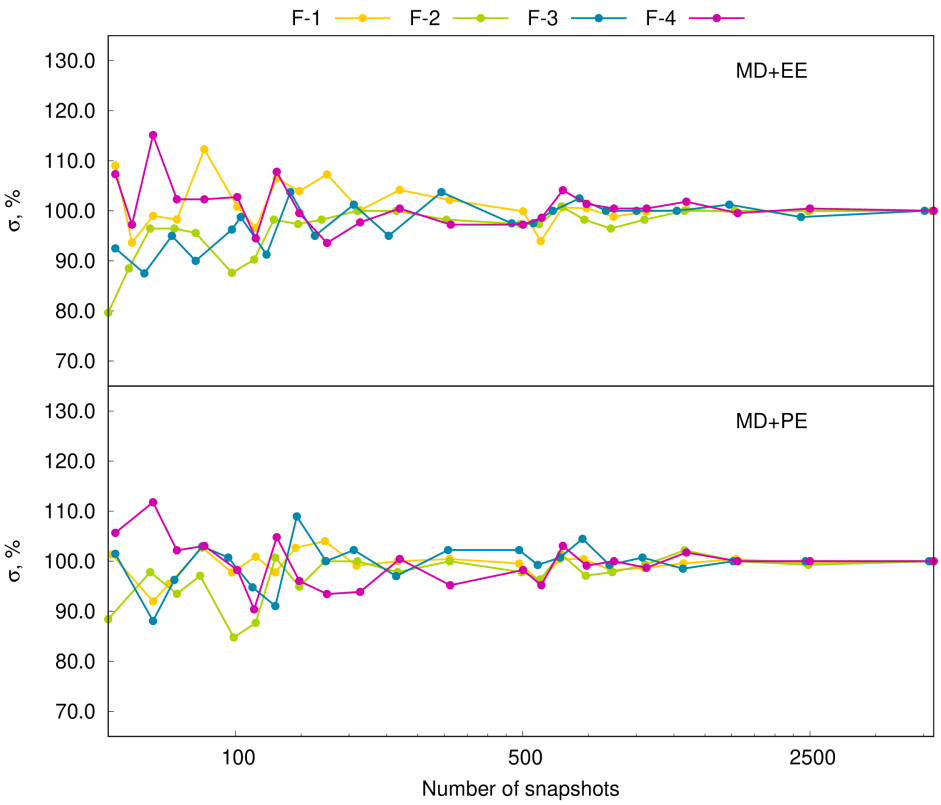


Figure 38: Series **F**, convergence of the standard deviation corresponding to the distribution of vertical excitation energy (computed in chloroform using EE and PE solvation models) with respect to the considered number of snapshots

Table 10 summarizes the results of the inhomogeneous broadenings simulations by the MD/PE, MD/EE approaches, and the ones obtained using the proposed ML based on EE excitation energies. First of all, it is needed to highlight that the ML approach remains stable and accurate for this set since the errors in the prediction of the final HWHM value do not exceed 7 cm^{-1}. For all compounds studied here, PE gives higher values of inhomogeneous broadening than EE. The observed increase in inhomogeneous broadening is in the range of 9 cm^{-1} – 52 cm^{-1}. As expected, the highest difference between EE and PE simulations is observed for less-polar structure **F-3**.

To simulate the vibrational spectra, it is needed to convolute the vibronic structure with function representing inhomogeneous broadening. Thus, the simulation of vibronic structure is an important step in this procedure. Based on the good performance of the vertical approach shown previously, the VG approach was used for the current analyses. The selection of functionals for vibronic structure simulations was performed using the vibrational reorganization energy as metrics (as in the previous section), but in this case, using the more accurate CCSD/aug-cc-pVDZ level of theory as reference. Based on the obtained results, two different functionals were chosen:

- CAM-B3LYP (unsigned relative errors: **F-1** – 17%, **F-2** – 15%, **F-3** – 6%, **F-4** – 39%, average error: 19%);
- LC-BLYP (unsigned relative errors: **F-1** – 1%, **F-2** – 18%, **F-3** – 34%, **F-4** – 31%, average error: 21%).

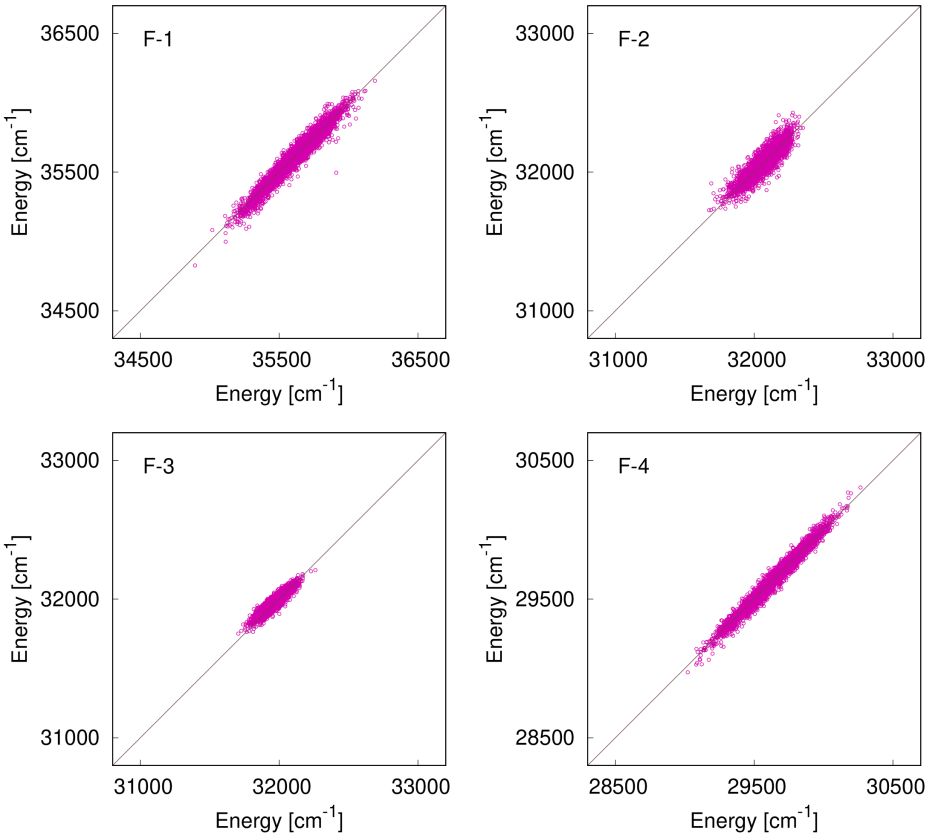


Figure 39: Series **F**, assessment of the developed ML approach. Scatter plot of the calculated by RI-CC2/EE (*X*-axis) and predicted based on the developed ML model (*Y*-axis) values of vertical excitation energy

structure	MD/EE		MD/EE+ML		MD/PE	
	σ^{inh}, cm^{-1}	HWHM, cm^{-1}	σ^{inh}, cm^{-1}	HWHM, cm^{-1}	σ^{inh}, cm^{-1}	HWHM, cm^{-1}
F-1	159	187	164	193	181	213
F-2	91	107	94	110	111	130
F-3	64	75	60	70	108	127
F-4	176	207	181	213	184	216

Table 10: Series **F**, inhomogeneous broadenings (standard deviation σ^{inh} and HWHM) calculated using EE and PE solvation models and predicted based on developed ML approach

LC-BLYP has a slightly higher average error than CAM-B3LYP, but CAM-B3LYP produces the greatest maximum error. The final spectra simulated by the convolution of "stick" vibronic spectra with obtained inhomogeneous broadenings are shown in Figures 40 and 41. In general, both functionals show acceptable agreement with the experiment. CAM-B3LYP is better in the prediction of excitation energies, while LC-BLYP provides a wider vibronic structure of the absorption band, which is in agreement with the conclusion made in previous works [265]. Nonetheless, LC-BLYP looks better in the prediction of relative intensities of vibronic transitions, while CAM-B3LYP tends to overestimate

the intensity of 0-0 transition, which contradicts the previous analysis of the DFAs performance for cognate molecules [265].

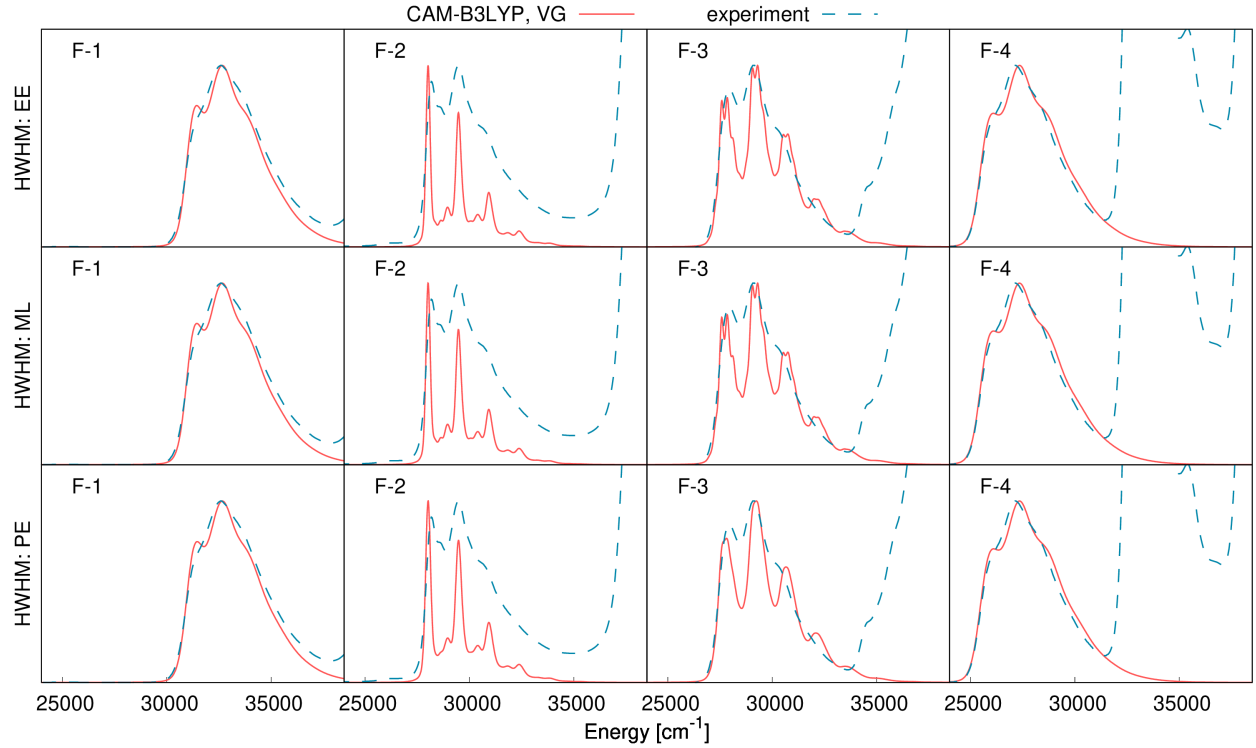


Figure 40: Series **F**, comparison amongst inhomogeneous broadenings computed by EE (first row), ML (second row), and PE (third row) approaches in simulations of vibrationally resolved absorption spectra using VG model based on CAM-B3LYP electronic structure simulations in chloroform solvent

The two upper rows of Figures 40 and 41 correspond to pure EE and combined EE/ML broadenings. There is no difference between these two sets of plots, which confirms the reliability and high accuracy of the developed ML approach. Comparing EE and PE results shows that the more expensive and accurate PE solvation model only appreciably improves the spectrum for the non-polar structure **F-3**, while for the rest of the structures simulated absorption spectra remain visually unchanged. As was shown in the previous section, good 2PA probes for bioimaging combine strong electron-accepting and strong electron-donating groups (resulting in polar molecules), and for the majority of structures that might be potentially designed, it would be enough to use the EE solvation model (or the ML prediction based on the EE solvation model) for the estimation of the inhomogeneous broadening.

The results of this study demonstrate how to perform efficient calculations of the inhomogeneous broadening caused by the solvent microenvironment of absorption bands for fluorescent dyes. For this purpose, the vertical excitation energies of studied dyes were calculated along the MD trajectory taking into account the discrete representation of solvent by using the EE and PE approaches. After the convolution of vibronic stick spectra with obtained inhomogeneous broadening function a good match with the experiment is achieved (i. e. satisfactory visual agreement of band shapes and average FWHM errors not exceeding 20%). The PE model can enhance the EE inhomogeneous broadening up to 70% for non-polar compounds and up to 15% for polar dyes. For polar fluorescent dyes, the spectra simulated with the EE and PE broadenings look similar, and taking into account the significantly higher computational costs of PE, the EE

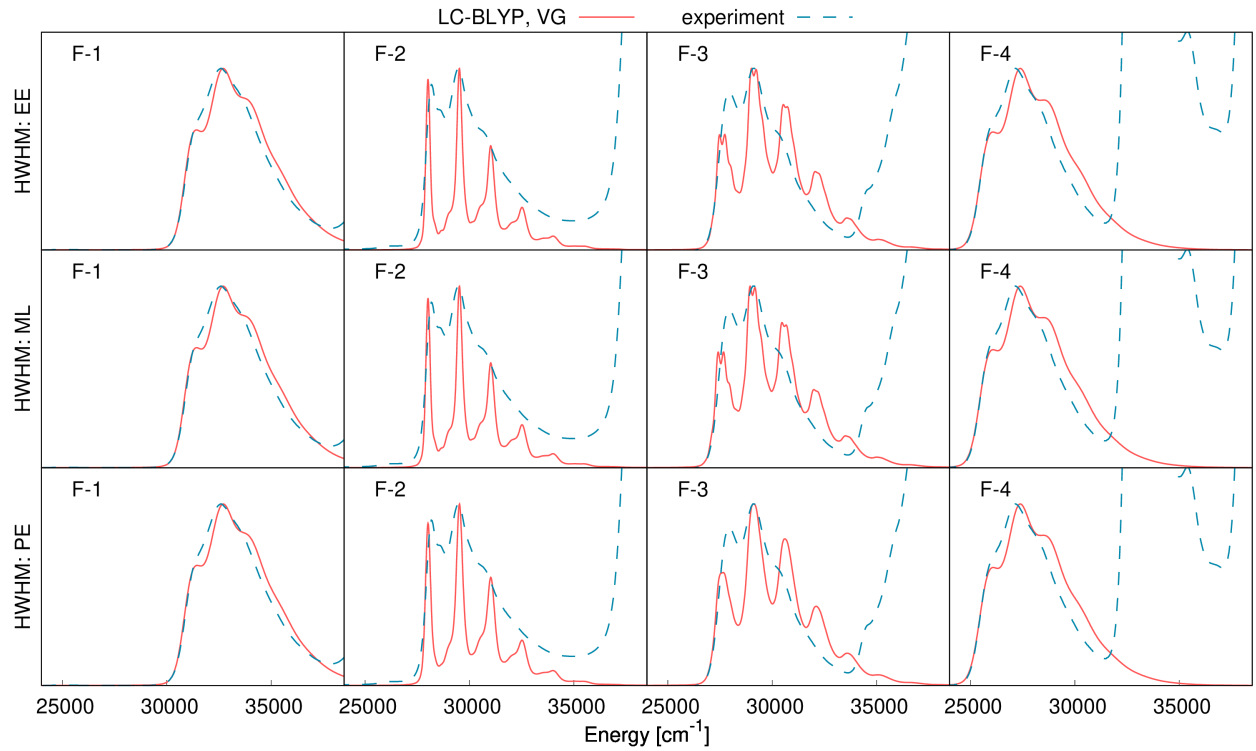


Figure 41: Series **F**, comparison amongst inhomogeneous broadenings computed by EE (first row), ML (second row), and PE (third row) approaches in simulations of vibrationally resolved absorption spectra using VG model based on LC-BLYP electronic structure calculations in chloroform solvent

inhomogeneous broadening is sufficient for a reliable description of polar structures. The new fingerprint developed in this work is 10^5 times smaller than the original Coulomb matrix for such systems. The combination of kernel ridge regression with the proposed one-dimensional fingerprint derived from the Coulomb matrix has an excellent performance, with an error in the prediction of inhomogeneous broadenings smaller than 5%. Using the suggested robust ML approach the CPU time needed for simulations of inhomogeneous broadening can be reduced by 98%. The further possible extension of this ML-based approach could be proposition of efficient fingerprint for PE simulations or the validation of ML/EE approach for chromophores in a more complex environment, for example, in a protein.

6 Conclusions

The research outlined in this dissertation was undertaken to facilitate the efficient and reliable design of fluorescent probes for bioimaging applications with the use of computational tools. In more detail, the emphasis was put on the following two objectives: *a)* the establishment of relations between structural motives and two-photon absorption activity for organoboron complexes; and *b)* the development of an efficient and reliable protocol for simulations of vibrationally-resolved absorption spectra of two-photon-excited probes.

The first goal was addressed by theoretical studies of novel boron-containing probes and thorough analyses of the nature of their two-photon absorption properties. Based on the studies of electronic structure for four investigated series of 2PA-promising compounds, important "structure–property" patterns were established. In more detail, it was shown that the introduction of an additional nitrogen atom to the boron-containing electron-accepting group leads to an increase of two-photon transition strength, red-shift of excitation energies, and improved photostability of studied compounds, which are desired characteristics for bioimaging applications. All derivatives containing two nitrogen atoms in a heterocyclic ring exhibit similar two-photon transition strengths, but the FQY of these dyes significantly depends on the heteroatom placement. Yet another strategy relied on changing the substituent at the boron atom and as we proved the electron-donating p-(dimethylamino)phenyl group at the boron atom can enhance the 2PA cross section by a factor of 2. However, this substitution, unfortunately, reduces the FQY of studied dyes. The research presented in the thesis shows that the design of quadrupolar structures from highly fluorescent dipolar dyes with low 2PA cross section may lead to enhanced 2PA cross sections and obtain FQYs suitable for two-photon microscopy. It was proved that symmetry may be used to control 2PA activity in these dyes through couplings of transition dipole moments. To further contribute to "structure–property" relationships, the two-photon absorption properties of BOPHY derivatives were studied for the first time. The important finding of these investigations is that weaker electron-donating substituents in quadrupolar structures are showing higher two-photon transition strength than stronger electron-donating groups in dipolar structures. This strategy can help to prevent the aggregation of 2PA probes in water solutions. Among the studied four series, at least eleven compounds exhibit potential as probes for two-photon fluorescent microscopy, although not all requirements can be reliably assessed using purely theoretical tools.

As highlighted above, the second goal concerned the development of a robust and cost-effective method to simulate vibrationally-resolved absorption spectra for a wide range of boron-containing fluorescent dyes. Two important aspects related to the modeling of absorption spectra were extensively studied here: vibronic structure and inhomogeneous broadening. It was shown that ωB97X, LC-BLYP, and LC-BLYP-OT(α) functionals have the best performance (amongst twelve examined DFAs) in the electronic structure simulations accounting for the effect of molecular vibrations. The satisfactory performance of the LC-BLYP-OT(α) functional is particularly promising for further investigation due to its excellent accuracy in the calculations of second-order hyperpolarizabilities and two-photon absorption strengths. Several harmonic approaches were compared in simulations of the vibronic structure of organoboron complexes concluding that vertical models show better agreement with the experiment than adiabatic approaches. The latter ones predict unphysically long vibrational progressions. The primary notion put forward in the thesis is that simple Franck-Condon harmonic approximation can not account for significant differences between ground and excited state geometries of big compounds in adiabatic approximations, which was confirmed using VH and dAH approaches. The compensation of errors makes vertical gradient approximation the best candidate for efficient simulations of the

vibronic structure and the results of VG simulations nicely reproduce experimental spectra. A very important benefit of the vertical gradient approach is its low computational cost, a characteristic important for further improvements of fast computational protocols for pre-screening of fluorescent dye properties.

Aiming at improving the reliability of simulations, i.e. to achieve better agreement between theory and experiment, the effect of mechanical anharmonic corrections for selected modes was analyzed. For simulations of anharmonic corrections of compounds with more than 20 atoms, a new noniterative protocol using the FC code based on a preselected basis set of transitions was proposed. It was shown that the perturbation-theory-based approach applied to a subset of vibrational modes can deliver a more accurate vibrationally-resolved spectrum in comparison to harmonic approximation. Nevertheless, this work showed that for more realistic and bigger fluorescent dyes, the current PT2 implementation has two limitations: *i)* large molecules imply using a very large basis set of vibrational transitions, thus leading to high computational demands, and *ii)* problem of energy degeneracies may become too significant. Taken together, these results hint towards the development of a variational approach that could be efficiently used for molecules composed of more than several dozen vibrational modes. Besides studies of mechanical anharmonicity, the analysis of Herzberg-Teller contribution was performed (linear term in the expansion of electric dipole moment wrt. normal modes was included). It was demonstrated that accounting for the HT effect around S_0 geometry does not affect the vibronic structure obtained by the FC approach in the case of bright 1PA spectra corresponding to the $S_0 \rightarrow S_1$ transition in studied organoboron complexes. On the contrary, the HT effect significantly changed the band shape in 2PA spectra corresponding to the very same transition. Summing up this aspect, three main conclusions can be drawn: *i)* the vertical gradient approach is the most suitable for boron-containing fluorescent dyes amongst applied methods for simulations of vibronic structure, *ii)* mechanical anharmonicity might improve theory-experiment agreement in the case of electronic one-photon absorption spectra, *iii)* linear HT term does not change overall band shape corresponding to bright $\pi \rightarrow \pi^{\star}$ transition in one-photon absorption spectra, in contrast to band shape in two-photon absorption spectra.

Another aspect considered in the thesis was devoted to simulations of inhomogeneous broadening of electronic transitions for molecules in solution. Two approaches were employed for the discrete representation of solvent: efficient electronic embedding and, more accurate, polarizable embedding. For the studied compounds, the EE approach gives satisfactory results, i.e. upon its combination with the vertical gradient method, the obtained error in prediction of FWHM of absorption band does not exceed 22% for non-polar chloroform and 17% for acetonitrile. On the other hand, the PE approach provides significant improvements only for non-polar molecules, for which the EE approach underestimates FWHM values. An important contribution to achieving the efficient simulations of band broadenings was the development of a ML approach based on a new proposed 1D fingerprint. This new approach allows to decrease the CPU time needed for simulations by a factor of 50 and yet it predicts inhomogeneous broadenings with error not exceeding 5 %.

Taken together, the results presented in the doctoral dissertation are a significant step towards the development of a reliable and efficient computational protocol for the prescreening of photophysical properties of organic dyes.

References

1. Göppert-Mayer, M. Über elementarakte mit zwei quantensprüngen. *Ann. Phys. (Berl.)* **401,** 273–294 (1931).
2. Kaiser, W. & Garrett, C. G. B. Two-photon excitation in CaF_2: Eu^{2+}. *Phys. Rev. Lett.* **7,** 229–231 (1961).
3. Mairnan, T. Stimulated optical radiation in ruby. *Nature* **187,** 493–494 (1960).
4. Vasilenko, L., Chebotaev, V. & Shishaev, A. Line shape of two-photon absorption in a standing-wave field in a gas. *Sov. J. Exp. Theor. Phys. Lett.* **12,** 113 (1970).
5. Cagnac, B., Grynberg, G. & Biraben, F. Spectroscopie d'absorption multiphotonique sans effet Doppler. *J. Phys. France* **34,** 845–858 (1973).
6. Hänsch, T., Harvey, K., Meisel, G. & Schawlow, A. Two-photon spectroscopy of Na 3s-4d without Doppler broadening using a cw dye laser. *Opt. Commun.* **11,** 50–53 (1974).
7. Parthenopoulos, D. A. & Rentzepis, P. M. Three-dimensional optical storage memory. *Science* **245,** 843–845 (1989).
8. Strickler, J. H. & Webb, W. W. Three-dimensional optical data storage in refractive media by two-photon point excitation. *Opt. Lett.* **16,** 1780–1782 (1991).
9. Ehrlich, J. *et al.* Two-photon absorption and broadband optical limiting with bis-donor stilbenes. *Opt. Lett.* **22,** 1843–1845 (1997).
10. Said, A. *et al.* Third-and fifth-order optical nonlinearities in organic materials. *Chem. Phys. Lett.* **228,** 646–650 (1994).
11. Wu, E.-S., Strickler, J. H., Harrell, W. R. & Webb, W. W. *Two-photon lithography for microelectronic application* in *Optical/Laser Microlithography V* **1674** (1992), 776–782.
12. Maruo, S., Nakamura, O. & Kawata, S. Three-dimensional microfabrication with two-photon-absorbed photopolymerization. *Opt. Lett.* **22,** 132–134 (1997).
13. Lenz, F. Nonlinear optical effects in PDT. *J. Phys. IV France* **4,** 237–240 (1994).
14. Ogawa, K. & Kobuke, Y. Recent advances in two-photon photodynamic therapy. *Anticancer Agents Med. Chem.* **8,** 269–279 (2008).
15. Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248,** 73–76 (1990).
16. Bonapace, L. *et al.* If you don't look, you won't see: Intravital multiphoton imaging of primary and metastatic breast cancer. *J. Mammary Gland Biol. Neoplasia.* **17,** 125–129 (2012).
17. Wang, C.-C. *et al.* Early development of cutaneous cancer revealed by intravital nonlinear optical microscopy. *Appl. Phys. Lett.* **97,** 113702 (2010).
18. Wu, Y.-X. *et al.* Multicolor two-photon nanosystem for multiplexed intracellular imaging and targeted cancer therapy. *Angew. Chem. Int. Ed.* **60,** 12569–12576 (2021).

19. Ricard, C. *et al.* Two-photon probes for in vivo multicolor microscopy of the structure and signals of brain cells. *Brain Struct. Funct.* **223,** 3011–3043 (2018).

20. Bethge, P., Chéreau, R., Avignone, E., Marsicano, G. & Nägerl, U. V. Two-photon excitation STED microscopy in two colors in acute brain slices. *Biophys. J.* **104,** 778–785 (2013).

21. Meng, G. *et al.* High-throughput synapse-resolving two-photon fluorescence microendoscopy for deep-brain volumetric imaging in vivo. *eLife* **8** (eds Marder, E. & Kleinfeld, D.) e40805 (2019).

22. Wang, B.-G., König, K. & Halbhuber, K.-J. Two-photon microscopy of deep intravital tissues and its merits in clinical research. *J. Microsc.* **238,** 1–20 (2010).

23. Schiazza, O. L. & Bille, J. F. High-speed two-photon excited autofluorescence imaging of ex vivo human retinal pigment epithelial cells toward age-related macular degeneration diagnostic. *J. Biomed. Opt.* **13,** 064008 (2008).

24. Qin, Z. *et al.* Adaptive optics two-photon microscopy enables near-diffraction-limited and functional retinal imaging in vivo. *Light Sci. Appl.* **9,** 79 (2020).

25. König, K. *et al.* Translation of two-photon microscopy to the clinic: Multimodal multiphoton CARS tomography of in vivo human skin. *J. Biomed. Opt.* **25,** 014515 (2020).

26. Kröger, M. *et al.* In vivo non-invasive staining-free visualization of dermal mast cells in healthy, allergy and mastocytosis humans using two-photon fluorescence lifetime imaging. *Sci. Rep.* **10,** 14930 (2020).

27. Zhang, K. *et al.* In vivo two-photon microscopy reveals the contribution of Sox9+ cell to kidney regeneration in a mouse model with extracellular vesicle treatment. *J. Biol. Chem.* **295,** 12203–12213 (2020).

28. Bower, A. J. *et al.* Tracking metabolic dynamics of apoptosis with high-speed two-photon fluorescence lifetime imaging microscopy. *Biomed. Opt. Express* **10,** 6408–6421 (2019).

29. Wu, X. *et al.* From static to dynamic: Live observation of the support system after ischemic stroke by two photon-excited fluorescence laser-scanning microscopy. *Neural Regen. Res.* **18,** 2093–2107 (2023).

30. Jeong, S. *et al.* Time-resolved fluorescence microscopy with phasor analysis for visualizing multicomponent topical drug distribution within human skin. *Sci. Rep.* **10,** 5360 (2020).

31. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2,** 932–940 (2005).

32. Ishii, H., Otomo, K., Takahashi, T., Yamaguchi, K. & Nemoto, T. Focusing new light on brain functions: Multi-photon microscopy for deep and super-resolution imaging. *Neurosci. Res.* **179,** 24–30 (2022).

33. Liu, J. *et al.* Ruthenium(II) polypyridyl complexes as mitochondria-targeted two-photon photodynamic anti-cancer agents. *Biomaterials* **56,** 140–153 (2015).

34. Sakadžić, S. *et al.* Two-photon high-resolution measurement of partial pressure of oxygen in cerebral vasculature and tissue. *Nat. Methods* **7,** 755–759 (2010).

35. Platisa, J. *et al.* High-speed low-light in vivo two-photon voltage imaging of large neuronal populations. *Nat. Methods* **20,** 1095–1103 (2023).

36. Zipfel, W. R., Williams, R. M. & Webb, W. W. Nonlinear magic: Multiphoton microscopy in the biosciences. *Nat. Biotechnol.* **21,** 1369–1377 (2003).

37. Vora, N. *et al. Denoising of depth-resolved, label-free, two-photon images using deep-learning-based algorithms* in *Label-free Biomedical Imaging and Sensing (LBIS) 2023* (eds Shaked, N. T. & Hayden, O.) **12391** (SPIE, 2023), 1239107.

38. Kim, H. M. & Cho, B. R. Small-Molecule Two-Photon Probes for Bioimaging Applications. *Chem. Rev.* **115,** 5014–5055 (2015).

39. Hu, B., Lu, P. & Wang, Y. Fluorescent chemosensors based on 9-cycloheptatrienylidene fluorenes (9-CHFs). *New J. Chem.* **37,** 1645–1653 (2013).

40. Poronik, Y. M., Clermont, G., Blanchard-Desce, M. & Gryko, D. T. Nonlinear optical chemosensor for sodium ion based on rhodol chromophore. *J. Org. Chem.* **78,** 11721–11732 (2013).

41. Miller, D. R., Jarrett, J. W., Hassan, A. M. & Dunn, A. K. Deep tissue imaging with multiphoton fluorescence microscopy. *Curr. Opin. Biomed. Eng.* **4,** 32–39 (2017).

42. Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Methods* **8,** 393–399 (2011).

43. Zhong, C.-Y. *et al.* Two-photon photoluminescence and bio-imaging application of monodispersed perovskite-in-silica nanocrystals with high biocompatibility. *Chem. Eng. J.* **431,** 134110 (2022).

44. Pramanik, A., Patibandla, S., Gao, Y., Gates, K. & Ray, P. C. Water triggered synthesis of highly stable and biocompatible 1D nanowire, 2D nanoplatelet, and 3D nanocube CsPbBr3 perovskites for multicolor two-photon cell imaging. *JACS Au* **1,** 53–65 (2020).

45. Chan, K. K., Giovanni, D., He, H., Sum, T. C. & Yong, K.-T. Water-stable all-inorganic perovskite nanocrystals with nonlinear optical properties for targeted multiphoton bioimaging. *ACS Appl. Nano Mater.* **4,** 9022–9033 (2021).

46. Yang, C. *et al.* Nanoscale metal–organic framework based two-photon sensing platform for bioimaging in live tissue. *Anal. Chem.* **91,** 2727–2733 (2019).

47. Chen, Y. *et al.* Two-photon luminescent metal complexes for bioimaging and cancer phototherapy. *Coord. Chem. Rev.* **310,** 16–40 (2016).

48. Juvekar, V., Lee, H. W., Lee, D. J. & Kim, H. M. Two-photon fluorescent probes for quantitative bio-imaging analysis in live tissues. *TrAC, Trends Anal. Chem.* **157,** 116787 (2022).

49. Zhang, Y. *et al.* Activatable organic upconversion nanoprobe for bioimaging of monoamine oxidase B in Parkinson's disease. *Sens. Actuators B Chem.* **389,** 133880 (2023).

50. Xu, Y. *et al.* Two-photon-pumped perovskite semiconductor nanocrystal lasers. *J. Am. Chem. Soc.* **138,** 3761–3768 (2016).

51. Hardman, R. A toxicologic review of quantum dots: Toxicity depends on physicochemical and environmental factors. *Environ. Health Perspect.* **114,** 165–172 (2006).

52. Li, S. *et al.* Water-resistant perovskite nanodots enable robust two-photon lasing in aqueous environment. *Nat. Commun.* **11,** 1192 (2020).

53. Kumar, P. *et al.* Highly luminescent biocompatible $CsPbBr_3$@ SiO_2 core–shell nanoprobes for bioimaging and drug delivery. *J. Mater. Chem. B* **8,** 10337–10345 (2020).

54. Yang, Z. *et al.* Lead halide perovskite nanocrystals–phospholipid micelles and their biological applications: Multiplex cellular imaging and in vitro tumor targeting. *ACS Appl. Mater. Interfaces* **11,** 47671–47679 (2019).

55. Shcherbo, D. *et al.* Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* **418,** 567–574 (2009).

56. Lin, B., Luo, J. & Lehmann, R. Collectively stabilizing and orienting posterior migratory forces disperses cell clusters in vivo. *Nat. Commun.* **11,** 4477 (2020).

57. Sidorenko, P. *et al.* Evaluation of a gain-managed nonlinear fiber amplifier for multiphoton microscopy. *Biomed. Opt. Express* **14,** 2324–2332 (2023).

58. Molina, R. S. *et al.* Blue-shifted green fluorescent protein homologues are brighter than enhanced green fluorescent protein under two-photon excitation. *J. Phys. Chem. Lett.* **8,** 2548–2554 (2017).

59. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **22,** 1567–1572 (2004).

60. Ozbay, B. N. *et al.* Three dimensional two-photon brain imaging in freely moving mice using a miniature fiber coupled microscope with active axial-scanning. *Sci. Rep.* **8,** 8108 (2018).

61. Shcherbo, D. *et al.* Near-infrared fluorescent proteins. *Nat. Methods* **7,** 827–829 (2010).

62. Shcherbakova, D. M., Subach, O. M. & Verkhusha, V. V. Red fluorescent proteins: Advanced imaging applications and future design. *Angew. Chem. Int. Ed.* **51,** 10724–10738 (2012).

63. Ormö, M. *et al.* Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273,** 1392–1395 (1996).

64. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 11984–11989 (2000).

65. Polyakov, I. V., Grigorenko, B. L., Epifanovsky, E. M., Krylov, A. I. & Nemukhin, A. V. Potential energy landscape of the electronic states of the GFP chromophore in different protonation forms: Electronic transition energies and conical intersections. *J. Chem. Theory Comput.* **6,** 2377–2387 (2010).

66. Grabarek, D. & Andruniów, T. The role of hydrogen bonds and electrostatic interactions in enhancing two-photon absorption in green and yellow fluorescent proteins. *ChemPhysChem* **23,** e202200003 (2022).

67. Andrade, C. D. *et al.* Two-photon fluorescence lysosomal bioimaging with a micelle-encapsulated fluorescent probe. *J. Fluoresc.* **21,** 1223–1230 (2011).

68. Park, H. J. *et al.* Measurement of pH values in human tissues by two-photon microscopy. *Angew. Chem. Int. Ed.* **51,** 2673–2676 (2012).

69. Meng, X. *et al.* A ratiometric two-photon fluorescent probe for cysteine and homocysteine in living cells. *Sens. Actuators B Chem.* **201,** 520–525 (2014).

70. Yang, H. *et al.* A ratiometric two-photon fluorescent probe for selective detection of endogenous formaldehyde via Aza-Cope reaction. *Sens. Actuators B Chem.* **270,** 318–326 (2018).

71. Chen, L., Park, S. J., Wu, D., Kim, H. M. & Yoon, J. A two-photon fluorescent probe for colorimetric and ratiometric monitoring of mercury in live cells and tissues. *Chem. Commun.* **55,** 1766–1769 (2019).

72. Xu, G., Tang, Y. & Lin, W. Development of a two-photon ratiometric fluorescent probe for Glutathione and its applications in living cells. *Chem. Res. Chin. Univ.* **34,** 523–527 (2018).

73. Albota, M. *et al.* Design of organic molecules with large two-photon absorption cross sections. *Science* **281,** 1653–1656 (1998).

74. Pawlicki, M., Collins, H. A., Denning, R. G. & Anderson, H. L. Two-photon absorption and the design of two-photon dyes. *Angew. Chem. Int. Ed.* **48,** 3244–3266 (2009).

75. Marder, S. R. Organic nonlinear optical materials: Where we have been and where we are going. *Chem. Commun.,* 131–134 (2006).

76. Zhou, H. *et al.* 1, 3, 5-Triazine-cored derivatives dyes containing triphenylamine based two-photon absorption: Synthesis, optical characterization and bioimaging. *Dyes Pigm.* **94,** 570–582 (2012).

77. Griesbeck, S. *et al.* The effect of branching on the one- and two-photon absorption, cell viability, and localization of cationic triarylborane chromophores with dipolar versus octupolar charge distributions for cellular imaging. *Chem. Eur. J.* **25,** 13164–13175 (2019).

78. Lavanya Devi, C. *et al.* Fluorenylethynylpyrene derivatives with strong two-photon absorption: Influence of substituents on optical properties. *J. Mater. Chem. C* **3,** 3730–3744 (2015).

79. Hrobárik, P. *et al.* Benzothiazoles with tunable electron-withdrawing strength and reverse polarity: A route to triphenylamine-based chromophores with enhanced two-photon absorption. *J. Org. Chem.* **76,** 8726–8736 (2011).

80. Rodrigues, C. A., Mariz, I. F., Maçôas, E. M., Afonso, C. A. & Martinho, J. M. Two-photon absorption properties of push–pull oxazolones derivatives. *Dyes Pigm.* **95,** 713–722 (2012).

81. Canola, S. *et al.* One-and two-photon absorption properties of quadrupolar thiophene-based dyes with acceptors of varying strengths. *Photochem. Photobiol. Sci.* **18,** 2180–2190 (2019).

82. Woo, H. Y. *et al.* Solvent effects on the two-photon absorption of distyrylbenzene chromophores. *J. Am. Chem. Soc.* **127,** 14721–14729 (2005).

83. Zhou, Y., Wang, X.-f., Tan, C.-h. & Wang, C.-k. Effect of a hydrogen bond on molecular probing properties in the solvent. *J. Phys. Chem. A* **124,** 520–528 (2020).

84. Woo, H. Y. *et al.* Water-soluble [2.2]paracyclophane chromophores with large two-photon action cross sections. *J. Am. Chem. Soc.* **127,** 820–821 (2005).

85. Liu, Z. *et al.* Two-photon absorption enhancement induced by aggregation due to intermolecular hydrogen bonding in V-shaped 2-hydroxypyrimidine derivatives. *Chem. Commun.,* 2260–2262 (2008).

86. Michail, E., Schreck, M. H., Holzapfel, M. & Lambert, C. Exciton coupling effects on the two-photon absorption of squaraine homodimers with varying bridge units. *Phys. Chem. Chem. Phys.* **22,** 18340–18350 (2020).

87. Griesbeck, S. *et al.* Tuning the π-bridge of quadrupolar triarylborane chromophores for one- and two-photon excited fluorescence imaging of lysosomes in live cells. *Chem. Sci.* **10,** 5405–5422 (2019).

88. Šimon, P. *et al.* Centripetal triazine chromophores: Towards efficient two-photon absorbers and highly emissive polyimide films. *J. Mater. Chem. C* **11,** 7252–7261 (2023).

89. Coe, B. J. *et al.* Quadratic and cubic nonlinear optical properties of salts of diquat-based chromophores with diphenylamino substituents. *J. Phys. Chem. A* **114,** 12028–12041 (2010).

90. Strehmel, B., Sarker, A. M. & Detert, H. The influence of σ and π acceptors on two-photon absorption and solvatochromism of dipolar and quadrupolar unsaturated organic compounds. *ChemPhysChem* **4,** 249–259 (2003).

91. Kato, S.-i. *et al.* Strongly red-fluorescent novel donor–π-bridge–acceptor–π-bridge–donor (D–π–A–π–D) type 2,1,3-benzothiadiazoles with enhanced two-photon absorption cross-sections. *Chem. Commun.,* 2342–2343 (2004).

92. Kamada, K., Iwase, Y., Sakai, K., Kondo, K. & Ohta, K. Cationic two-photon absorption chromophores with double- and triple-bond cores in symmetric/asymmetric arrangements. *J. Phys. Chem. C* **113,** 11469–11474 (2009).

93. Obregón-Mendoza, M. A. *et al.* Non-cytotoxic dibenzyl and difluoroborate curcuminoid fluorophores allow visualization of nucleus or cytoplasm in bioimaging. *Molecules* **25** (2020).

94. DeRosa, C. A. & Fraser, C. L. in *Main Group Strategies towards Functional Hybrid Materials* 111–140 (John Wiley & Sons, Ltd, 2017). ISBN: 9781119235941.

95. Treibs, A. & Kreuzer, F.-H. Difluorboryl-komplexe von di- und tripyrrylmethenen. *Justus Liebigs Ann. Chem.* **718,** 208–223 (1968).

96. Wang, Y. *et al.* Red emissive AIE nanodots with high two-photon absorption efficiency at 1040 nm for deep-tissue in vivo imaging. *Biomed. Opt. Express* **6,** 3783–3794 (2015).

97. Zhao, Z. *et al.* Red emissive biocompatible nanoparticles from tetraphenylethene-decorated BODIPY luminogens for two-photon excited fluorescence cellular imaging and mouse brain blood vascular visualization. *Part. Part. Syst. Charact.* **31,** 481–491 (2014).

98. Yang, J. *et al.* Synthesis, spectroscopic characterization, one and two-photon absorption properties, and electrochemistry of truxene π-expanded BODIPYs dyes. *Dyes Pigm.* **176,** 108183 (2020).

99. Chang, H.-J. *et al.* Femtosecond spectroscopy and nonlinear optical properties of aza-BODIPY derivatives in solution. *Chem. Eur. J.* **28,** e202104072 (2022).

100. Fan, G., Yang, L. & Chen, Z. Water-soluble BODIPY and aza-BODIPY dyes: Synthetic progress and applications. *Front. Chem. Sci. Eng.* **8,** 405–417 (2014).

101. Kamada, K. *et al.* Boron difluoride curcuminoid fluorophores with enhanced two-photon excited fluorescence emission and versatile living-cell imaging properties. *Chem. Eur. J.* **22,** 5219–5232 (2016).

102. Kita, H. *et al.* Switching of two-photon optical properties by anion binding of pyrrole-based boron diketonates through conformation change. *Chem. Eur. J.* **26,** 3404–3410 (2020).

103. Zhang, X. *et al.* A bifunctional curcumin analogue for two-photon imaging and inhibiting crosslinking of amyloid beta in Alzheimer's disease. *Chem. Commun.* **50,** 11550–11553 (2014).

104. Zhang, X. *et al.* Design and synthesis of curcumin analogues for in vivo fluorescence imaging and inhibiting copper-induced cross-linking of amyloid beta species in alzheimer's disease. *J. Am. Chem. Soc.* **135,** 16397–16409 (2013).

105. Frath, D., Didier, P., Mély, Y., Massue, J. & Ulrich, G. Vectorization and intracellular distribution of a two-photon-absorbing, near-infrared-emitting π-extended boranil dye. *ChemPhotoChem* **1,** 109–112 (2017).

106. Zhao, N. *et al.* Facile construction of boranil complexes with aggregation-induced emission characteristics and their specific lipid droplet imaging applications. *Chem. Commun.* **55,** 8494–8497 (2019).

107. Shao, T. *et al.* Identification of fatty liver disease at diverse stages using two-photon absorption of triphenylamine-based BODIPY analogues. *J. Mater. Chem. B* **7,** 3704–3709 (2019).

108. Zhang, M. *et al.* Ultra-bright intercellular lipids pseudo di-BODIPY probe with low molecular weight, high quantum yield and large two-photon action cross-sections. *Sens. Actuators B Chem.* **261,** 161–168 (2018).

109. Bednarska, J. *et al.* Two-photon absorption of BF2-carrying compounds: Insights from theory and experiment. *Phys. Chem. Chem. Phys.* **19,** 5705–5708 (2017).

110. Tamgho, I.-S., Hasheminasab, A., Engle, J. T., Nemykin, V. N. & Ziegler, C. J. A new highly fluorescent and symmetric pyrrole–BF2 chromophore: BOPHY. *J. Am. Chem. Soc.* **136,** 5623–5626 (2014).

111. Yu, C. *et al.* A family of highly fluorescent and unsymmetric bis(BF2) chromophore containing both pyrrole and N-heteroarene derivatives: BOPPY. *Org. Lett.* **20,** 4462–4466 (2018).

112. Yu, C. *et al.* A family of highly fluorescent and membrane-permeable bis(BF2) acyl-pyridinylhydrazine dyes with strong solid-state emission and large stokes shifts: The BOAPH fluorophores. *J. Org. Chem.* **86,** 11492–11501 (2021).

113. Yu, C. *et al.* A family of BODIPY-like highly fluorescent and unsymmetrical bis(BF2) pyrrolyl–acylhydrazone chromophores: BOAPY. *Org. Lett.* **22,** 4588–4592 (2020).

114. Jiang, X. *et al.* Synthesis, properties and application of novel 5,6,5,6-tetracyclic pyrazine/pyrrole-fused unsymmetric bis(BF2) fluorescent dyes: BOPYPYs. *Chin. Chem. Lett.* **30,** 2271–2273 (2019).

115. Pookkandam Parambil, S. *et al.* BOPAHY: A doubly chelated highly fluorescent pyrrole–acyl hydrazone –BF2 chromophore. *Chem. Commun.* **56,** 5791–5794 (2020).

116. Hajda, A. *et al.* BF2-functionalized benzothiazole amyloid markers: Effect of donor substituents on one- and two-photon properties. *ACS Appl. Bio Mater.* published online 7 December 2023, DOI: 10.1021/acsabm.3c00815, 1–9 (2023).

117. Greenman, K. P., Green, W. H. & Gómez-Bombarelli, R. Multi-fidelity prediction of molecular optical peaks with deep learning. *Chem. Sci.* **13,** 1152–1162 (2022).

118. Oulianov, D., Tomov, I., Dvornikov, A. & Rentzepis, P. Observations on the measurement of two-photon absorption cross-section. *Opt. Commun.* **191,** 235–243 (2001).

119. Knysh, I., Jassar, M. B., Ośmiałowski, B., Zaleśny, R. & Jacquemin, D. In silico screening of two-photon absorption properties of a large set of bis-difluoroborate dyes. *ChemPhotoChem* **6,** e202200137 (2022).

120. Chołuj, M. *et al.* Choosing bad versus worse: Predictions of two-photon-absorption strengths based on popular density functional approximations. *J. Chem. Theory Comput.* **18,** 1046–1060 (2022).

121. Salem, M. A. & Brown, A. Two-photon absorption in fluorescent protein chromophores: TDDFT and CC2 results. *J. Chem. Theory Comput.* **10,** 3260–3269 (2014).

122. Sałek, P. *et al.* Calculations of two-photon absorption cross sections by means of density-functional theory. *Chem. Phys. Lett.* **374,** 446–452 (2003).

123. Hättig, C., Christiansen, O. & Jo/rgensen, P. Coupled cluster response calculations of two-photon transition probability rate constants for helium, neon and argon. *J. Chem. Phys.* **108,** 8355–8359 (1998).

124. Friese, D. H., Hättig, C. & Ruud, K. Calculation of two-photon absorption strengths with the approximate coupled cluster singles and doubles model CC2 using the resolution-of-identity approximation. *Phys. Chem. Chem. Phys.* **14,** 1175–1184 (2012).

125. Nanda, K. D. & Krylov, A. I. Two-photon absorption cross sections within equation-of-motion coupled-cluster formalism using resolution-of-the-identity and Cholesky decomposition representations: Theory, implementation, and benchmarks. *J. Chem. Phys.* **142,** 064118 (2015).

126. Knippenberg, S. *et al.* Calculations of nonlinear response properties using the intermediate state representation and the algebraic-diagrammatic construction polarization propagator approach: Two-photon absorption spectra. *J. Chem. Phys.* **136,** 064107 (2012).

127. Gholami, S. *et al.* Multistate multiconfiguration quantum chemical computation of the two-photon absorption spectra of bovine rhodopsin. *J. Phys. Chem. Lett.* **10,** 6293–6300 (2019).

128. Alam, M. M., Chattopadhyaya, M. & Chakrabarti, S. Solvent induced channel interference in the two-photon absorption process - a theoretical study with a generalized few-state-model in three dimensions. *Phys. Chem. Chem. Phys.* **14,** 1156–1165 (2012).

129. Beerepoot, M. T. P. *et al.* Benchmarking the performance of exchange-correlation functionals for predicting two-photon absorption strengths. *J. Chem. Theory Comput.* **14,** 3677–3685 (2018).

130. Aslopovsky, V. R., Scherbinin, A. V., Kleshchina, N. N. & Bochenkova, A. V. Impact of the protein environment on two-photon absorption cross-sections of the GFP chromophore anion resolved at the XMCQDPT2 level of theory. *Int. J. Mol. Sci.* **24** (2023).

131. Badaeva, E. A., Timofeeva, T. V., Masunov, A. & Tretiak, S. Role of donor-acceptor strengths and separation on the two-photon absorption response of cytotoxic dyes: A TD-DFT study. *J. Phys. Chem. A* **109,** 7276–7284 (2005).

132. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98,** 5648–5652 (1993).

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

134. Masunov, A. & Tretiak, S. Prediction of two-photon absorption properties for organic chromophores using time-dependent density-functional theory. *J. Phys. Chem. B* **108,** 899–907 (2004).

135. Schipper, P. R. T., Gritsenko, O. V., van Gisbergen, S. J. A. & Baerends, E. J. Molecular calculations of excitation energies and (hyper)polarizabilities with a statistical average of orbital model exchange-correlation potentials. *J. Chem. Phys.* **112,** 1344–1352 (2000).

136. Amar, A., Boucekkine, A., Paul, F. & Mongin, O. DFT study of two-photon absorption of octupolar molecules. *Theor. Chem. Acc.* **138,** 1–7 (2019).

137. Richy, N. *et al.* Linear and nonlinear optical properties of quadrupolar bithiophenes and cyclopentadithiophenes as fluorescent oxygen photosensitizers. *Photochem* **3,** 127–154 (2023).

138. Yanai, T., Tew, D. P. & Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393,** 51–57 (2004).

139. Peach, M. J. G., Benfield, P., Helgaker, T. & Tozer, D. J. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.* **128,** 044118 (2008).

140. Díaz, C. *et al.* Two-photon absorption and two-photon circular dichroism of hexahelicene derivatives: A study of the effect of the nature of intramolecular charge transfer. *RSC Adv.* **5,** 17429–17437 (2015).

141. Dipold, J. *et al.* Two-photon absorption properties of BODIPY-like compounds based on BF2–naphthyridine complexes. *Phys. Chem. Chem. Phys.* **21,** 6662–6671 (2019).

142. Paterson, M. J. *et al.* Benchmarking two-photon absorption with CC3 quadratic response theory, and comparison with density-functional response theory. *J. Chem. Phys.* **124,** 054322 (2006).

143. Beerepoot, M. T. P., Friese, D. H., List, N. H., Kongsted, J. & Ruud, K. Benchmarking two-photon absorption cross sections: Performance of CC2 and CAM-B3LYP. *Phys. Chem. Chem. Phys.* **17,** 19306–19314 (2015).

144. Sirimatayanant, S. & Andruniów, T. Benchmarking two-photon absorption strengths of rhodopsin chromophore models with CC3 and CCSD methodologies: An assessment of popular density functional approximations. *J. Chem. Phys.* **158,** 094106 (2023).

145. Murugan, N. A. & Zaleśny, R. Multiscale modeling of two-photon probes for parkinson's diagnostics based on monoamine oxidase B biomarker. *J. Chem. Inf. Model.* **60,** 3854–3863 (2020).

146. Rudberg, E., Sałek, P., Helgaker, T. & Ågren, H. Calculations of two-photon charge-transfer excitations using Coulomb-attenuated density-functional theory. *J. Chem. Phys.* **123,** 184108 (2005).

147. Grabarek, D. & Andruniów, T. Assessment of functionals for TDDFT calculations of one- and two-photon absorption properties of neutral and anionic fluorescent proteins chromophores. *J. Chem. Theory Comput.* **15,** 490–508 (2019).

148. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38,** 3098–3100 (1988).

149. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37,** 785–789 (1988).

150. Iikura, H., Tsuneda, T., Yanai, T. & Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **115,** 3540–3544 (2001).

151. Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **98,** 1372–1377 (1993).

152. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **78,** 1396–1396 (1997).

153. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110,** 6158–6170 (1999).

154. Zaleśny, R., Szczotka, N., Grabarz, A., Ośmiałowski, B. & Jacquemin, D. Design of two-photon-excited fluorescent dyes containing fluoroborylene groups. *ChemPhotoChem* **3,** 719–726 (2019).

155. De Wergifosse, M., Beaujean, P. & Grimme, S. Ultrafast evaluation of two-photon absorption with simplified time-dependent density functional theory. *J. Phys. Chem. A* **126,** 7534–7547 (2022).

156. Wielgus, M. *et al.* Two-photon solvatochromism II: Experimental and theoretical study of solvent effects on the two-photon absorption spectrum of Reichardt's dye. *ChemPhysChem* **14,** 3731–3739 (2013).

157. Miertuš, S., Scrocco, E. & Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **55,** 117–129 (1981).

158. Frediani, L., Rinkevicius, Z. & Ågren, H. Two-photon absorption in solution by means of time-dependent density-functional theory and the polarizable continuum model. *J. Chem. Phys.* **122,** 244104 (2005).

159. Vivas, M. G. *et al.* Molecular structure–optical property relationships for a series of non-centrosymmetric two-photon absorbing push-pull triarylamine molecules. *Sci. Rep.* **4,** 4447 (2014).

160. Uudsemaa, M., Trummal, A., de Reguardati, S., Callis, P. R. & Rebane, A. TD-DFT calculations of one- and two-photon absorption in Coumarin C153 and Prodan: Attuning theory to experiment. *Phys. Chem. Chem. Phys.* **19,** 28824–28833 (2017).

161. Wei, X. *et al.* Theoretical investigation of Ru(II) complexes with long lifetime and a large two-photon absorption cross-section in photodynamic therapy. *J. Med. Chem.* **66,** 4167–4178 (2023).

162. Fox, S. J. *et al.* Electrostatic embedding in large-scale first principles quantum mechanical calculations on biomolecules. *J. Chem. Phys.* **135,** 224107 (2011).

163. Hršak, D., Marefat Khah, A., Christiansen, O. & Hättig, C. Polarizable embedded RI-CC2 method for two-photon absorption calculations. *J. Chem. Theory Comput.* **11,** 3669–3678 (2015).

164. Day, P. N. *et al.* An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.* **105,** 1968–1986 (1996).

165. Gordon, M. S. *et al.* The effective fragment potential method: A QM-based MM approach to modeling environmental effects in chemistry. *J. Phys. Chem. A* **105,** 293–307 (2001).

166. Silva, D., Barreto, R., Lacerda, E., Coutinho, K. & Canuto, S. One- and two-photon absorption of fluorescein dianion in water: A study using S-QM/MM methodology and ZINDO method. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **119,** 63–75 (2014).

167. Coutinho, K. & Canuto, S. Solvent effects in emission spectroscopy: A Monte Carlo quantum mechanics study of the $\pi \leftarrow \pi^*$ shift of formaldehyde in water. *J. Chem. Phys.* **113,** 9132–9139 (2000).

168. Arul Murugan, N. *et al.* Hybrid density functional theory/molecular mechanics calculations of two-photon absorption of dimethylamino nitro stilbene in solution. *Phys. Chem. Chem. Phys.* **13,** 12506–12516 (2011).

169. Steindal, A. H., Olsen, J. M. H., Ruud, K., Frediani, L. & Kongsted, J. A combined quantum mechanics/molecular mechanics study of the one- and two-photon absorption in the green fluorescent protein. *Phys. Chem. Chem. Phys.* **14,** 5440–5451 (2012).

170. Di Remigio, R., Giovannini, T., Ambrosetti, M., Cappelli, C. & Frediani, L. Fully polarizable QM/fluctuating charge approach to two-photon absorption of aqueous solutions. *J. Chem. Theory Comput.* **15,** 4056–4068 (2019).

171. Nanda, K. D. & Krylov, A. I. The effect of polarizable environment on two-photon absorption cross sections characterized by the equation-of-motion coupled-cluster singles and doubles method combined with the effective fragment potential approach. *J. Chem. Phys.* **149,** 164109 (2018).

172. Wunsch, L., Neusser, H. & Schlag, E. Two photon excitation spectrum of benzene and benzene-d6 in the gas phase: Assignment of inducing modes by hot band analysis. *Chem. Phys. Lett.* **31,** 433–440 (1975).

173. Jacquemin, D., Brémond, E., Planchat, A., Ciofini, I. & Adamo, C. TD-DFT vibronic couplings in anthraquinones: From basis set and functional benchmarks to applications for industrial dyes. *J. Chem. Theory Comput.* **7,** 1882–1892 (2011).

174. Muniz-Miranda, F. *et al.* Benchmarking TD-DFT against vibrationally resolved absorption spectra at room temperature: 7-aminocoumarins as test cases. *J. Chem. Theory Comput.* **11,** 5371–5384 (2015).

175. Jacquemin, D., Brémond, E., Ciofini, I. & Adamo, C. Impact of vibronic couplings on perceived colors: Two anthraquinones as a working example. *J. Phys. Chem. Lett.* **3,** 468–471 (2012).

176. Dierksen, M. & Grimme, S. The Vibronic structure of electronic absorption spectra of large molecules: A time-dependent density functional study on the influence of "exact" Hartree-Fock exchange. *J. Phys. Chem. A* **108,** 10225–10237 (2004).

177. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10,** 6615–6620 (2008).

178. Chai, J.-D. & Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **128,** 084106 (2008).

179. Sitkiewicz, S., Matito, E., Luis, J. M. & Zalesny, R. Pitfall in simulations of vibronic TD-DFT spectra: Diagnosis and assessment. *Phys. Chem. Chem. Phys.* **25,** 30193–30197 (2023).

180. Avila Ferrer, F. J. & Santoro, F. Comparison of vertical and adiabatic harmonic approaches for the calculation of the vibrational structure of electronic spectra. *Phys. Chem. Chem. Phys.* **14,** 13549–13563 (2012).

181. Cerezo, J., Mazzeo, G., Longhi, G., Abbate, S. & Santoro, F. Quantum-classical calculation of vibronic spectra along a reaction path: The case of the ECD of easily interconvertible conformers with ppposite chiral responses. *J. Phys. Chem. Lett.* **7,** 4891–4897 (2016).

182. Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane, A. Describing two-photon absorptivity of fluorescent proteins with a new vibronic coupling mechanism. *J. Phys. Chem. B* **116,** 1736–1744 (2012).

183. Horz, M. *et al.* Vibrationally resolved two-photon electronic spectra including vibrational pre-excitation: Theory and application to VIPER spectroscopy with two-photon excitation. *J. Chem. Phys.* **158,** 064201 (2023).

184. Rodriguez-Garcia, V., Yagi, K., Hirao, K., Iwata, S. & Hirata, S. Franck-Condon factors based on anharmonic vibrational wave functions of polyatomic molecules. *J. Chem. Phys.* **125,** 014109 (2006).

185. Zakrzewska, A. *et al.* Substituent effects on the photophysical properties of fluorescent 2-benzoylmethylene-quinoline difluoroboranes: A combined experimental and quantum chemical study. *Dyes Pigm.* **99,** 957–965 (2013).

186. Zaleśny, R. *et al.* Toward fully nonempirical simulations of optical band shapes of molecules in solution: A case study of heterocyclic ketoimine difluoroborates. *J. Phys. Chem. A* **119,** 5145–5152 (2015).

187. Grabarz, A. M. *et al.* Photophysical properties of phenacylphenantridine difluoroboranyls: effect of substituent and double benzannulation. *J. Org. Chem.* **82,** 1529–1537 (2017).

188. Fortino, M. *et al.* On the simulation of vibrationally resolved electronic spectra of medium-size molecules: the case of styryl substituted BODIPYs. *Phys. Chem. Chem. Phys.* **21,** 3512–3526 (2019).

189. Zaleśny, R., Murugan, N. A., Tian, G., Medved', M. & Ågren, H. First-principles simulations of one- and two-photon absorption band shapes of the bis(BF2) core complex. *J. Phys. Chem. B* **120,** 2323–2332 (2016).

190. Nanosecond timescale optical inhomogeneous broadening of dye molecules in liquids at and near room temperature. *Chem. Phys. Lett.* **176,** 159–166 (1991).

191. Berg, M. & Vanden Bout, D. A. Ultrafast Raman echo measurements of vibrational dephasing and the nature of solvent-solute interactions. *Acc. Chem. Res.* **30,** 65–71 (1997).

192. Nemkovich, N. A., Rubinov, A. N. & Tomin, V. I. in *Topics in Fluorescence Spectroscopy: Principles* (ed Lakowicz, J. R.) 367–428 (Springer US, Boston, MA, 2002).

193. Renge, I. Mechanisms of solvent shifts, pressure shifts, and inhomogeneous broadening of the optical spectra of dyes in liquids and low-temperature glasses. *J. Phys. Chem. A* **104,** 7452–7463 (2000).

194. Zuehlsdorff, T. J. & Isborn, C. M. Modeling absorption spectra of molecules in solution. *Int. J. Quantum Chem.* **119,** e25719 (2019).

195. Bednarska, J. *et al.* Nonempirical simulations of inhomogeneous broadening of electronic transitions in solution: Predicting band shapes in one- and two-photon absorption spectra of chalcones. *Molecules* **22** (2017).

196. Marcus, R. A. Interactions in polar media. I. Interparticle interaction energy. *J. Chem. Phys.* **38,** 1335–1340 (1963).

197. Marcus, R. A. Interactions in polar media. II. Continua. *J. Chem. Phys.* **39,** 460–469 (1963).

198. Marcus, R. A. Free energy of nonequilibrium polarization systems. III. Statistical mechanics of homogeneous and electrode systems. *J. Chem. Phys.* **39,** 1734–1740 (1963).

199. Marcus, R. A. On the theory of shifts and broadening of electronic spectra of polar solutes in polar media. *J. Chem. Phys.* **43,** 1261–1274 (1965).

200. Marcus, R. A. Relation between charge transfer absorption and fluorescence spectra and the inverted region. *J. Phys. Chem.* **93,** 3078–3086 (1989).

201. Londergan, C. H., Wang, J., Axelsen, P. H. & Hochstrasser, R. M. Two-dimensional infrared spectroscopy displays signatures of structural ordering in peptide aggregates. *Biophys. J.* **90,** 4672–4685 (2006).

202. Myers, A. B. Molecular electronic spectral broadening in liquids and glasses. *Annu. Rev. Phys. Chem.* **49,** 267–295 (1998).

203. Loring, R. F. Statistical mechanical calculation of inhomogeneously broadened absorption line shapes in solution. *J. Phys. Chem.* **94,** 513–515 (1990).

204. Langhals, H. A re-examination of the line shape of the electronic spectra of complex molecules in solution: log-normal function versus gaussian. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **56,** 2207–2210. ISSN: 1386-1425 (2000).

205. Benassi, E., Cappelli, C., Carlotti, B. & Barone, V. An integrated computational tool to model the broadening of the absorption bands of flexible dyes in solution: cationic chromophores as test cases. *Phys. Chem. Chem. Phys.* **16,** 26963–26973 (2014).

206. Cerezo, J., Avila Ferrer, F. J. & Santoro, F. Disentangling vibronic and solvent broadening effects in the absorption spectra of coumarin derivatives for dye sensitized solar cells. *Phys. Chem. Chem. Phys.* **17,** 11401–11411 (2015).

207. Ferrer, F. J. A., Improta, R., Santoro, F. & Barone, V. Computing the inhomogeneous broadening of electronic transitions in solution: a first-principle quantum mechanical approach. *Phys. Chem. Chem. Phys.* **13,** 17007–17012 (2011).

208. Grabarz, A. M. *et al.* The impact of the heteroatom in a five-membered ring on the photophysical properties of difluoroborates. *Dyes Pigm.* **170,** 107481 (2019).

209. Cerezo, J., Aranda, D., Avila Ferrer, F. J., Prampolini, G. & Santoro, F. Adiabatic-molecular dynamics generalized vertical hessian approach: A mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase. *J. Chem. Theory Comput.* **16,** 1215–1231 (2020).

210. Cerezo, J., García-Iriepa, C., Santoro, F., Navizet, I. & Prampolini, G. Unraveling the contributions to the spectral shape of flexible dyes in solution: insights on the absorption spectrum of an oxyluciferin analogue. *Phys. Chem. Chem. Phys.* **25,** 5007–5020 (2023).

211. Alam, M. M., Beerepoot, M. T. P. & Ruud, K. A generalized few-state model for the first hyperpolarizability. *J. Chem. Phys.* **152,** 244106 (2020).

212. Duschinsky, F. The importance of the electron spectrum in multi atomic molecules. Concerning the Franck-Condon principle. *Acta Physicochim. URSS* **7,** 551–566 (1937).

213. Bloino, J., Biczysko, M., Santoro, F. & Barone, V. General approach to compute vibrationally resolved one-photon electronic spectra. *J. Chem. Theory Comput.* **6,** 1256–1274 (2010).

214. Macak, P., Luo, Y. & Ågren, H. Simulations of vibronic profiles in two-photon absorption. *Chem. Phys. Lett.* **330,** 447–456 (2000).

215. Petrenko, T. & Neese, F. Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model. *J. Chem. Phys.* **137,** 234107 (2012).

216. Cerezo, J. & Santoro, F. Revisiting vertical models to simulate the line shape of electronic spectra adopting cartesian and internal coordinates. *J. Chem. Theory Comput.* **12,** 4970–4985 (2016).

217. Lin, N. *et al.* Theory for vibrationally resolved two-photon circular dichroism spectra. Application to (R)-(+)-3-methylcyclopentanone. *J. Phys. Chem. A* **113,** 4198–4207 (2009).

218. Herzberg, G. & Teller, E. Schwingungsstruktur der elektronenübergänge bei mehratomigen molekülen. *Z. Phys. Chem. Abt. B* **21B,** 410–446 (1933).

219. Small, G. J. Herzberg–Teller vibronic coupling and the Duschinsky effect. *The Journal of Chemical Physics* **54,** 3300–3306 (1971).

220. Lami, A. & Santoro, F. in *Computational Strategies for Spectroscopy* 475–516 (John Wiley & Sons, Ltd, 2011).

221. Luis, J. M., Bishop, D. M. & Kirtman, B. A different approach for calculating Franck–Condon factors including anharmonicity. *J. Chem. Phys.* **120,** 813–822 (2004).

222. Hirschfelder, J. O., Brown, W. B. & Epstein, S. T. Recent developments in perturbation theory. *Adv. Quantum Chem.* **1,** 255–374 (1964).

223. Ab initio study of solvated molecules: A new implementation of the polarizable continuum model. *Chem. Phys. Lett.* **255,** 327–335 (1996).

224. Frisch, M. J. *et al. Gaussian 16 Revision C.01* Gaussian Inc. Wallingford CT. 2016.

225. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38 (1996).

226. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30,** 2157–2164 (2009).

227. MacKerell Jr, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 (1998).

228. Dietz, W. & Heinzinger, K. A molecular dynamics study of liquid chloroform. *Ber. Bunsenges. Phys. Chem.* **89,** 968–977 (1985).

229. Yu, W., He, X., Vanommeslaeghe, K. & MacKerell Jr, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **33,** 2451–2468 (2012).

230. Breneman, C. M. & Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11,** 361–373 (1990).

231. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).

232. Hättig, C. & Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **113,** 5154–5161 (2000).

233. Weigend, F., Köhn, A. & Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **116,** 3175–3183 (2002).

234. Dunning Thom H., J. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).

235. *TURBOMOLE V7.3 2018, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from* `http://www.turbomole.com`.

236. Yamaji, M. *et al.* Blue fluorescence from BF2 complexes of N,O-Benzamide ligands: Synthesis, structure, and photophysical properties. *Inorg. Chem.* **56,** 12514–12519 (2017).

237. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120,** 215–241 (2008).

238. Ditchfield, R., Hehre, W. J. & Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **54,** 724–728 (1971).

239. Hehre, W. J., Ditchfield, R. & Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **56,** 2257–2261 (1972).

240. Hariharan, P. C. & Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **28,** 213–222 (1973).

241. Huh, D. S. & Choe, S. J. Comparative DFT study for molecular geometries and spectra of methyl pheophorbides-a: Test of M06-2X and two other functionals. *J. Porphyr. Phthalocyanines* **14,** 592–604 (2010).

242. Chibani, S., Le Guennic, B., Charaf-Eddin, A., Laurent, A. D. & Jacquemin, D. Revisiting the optical signatures of BODIPY with ab initio tools. *Chem. Sci.* **4,** 1950–1963 (2013).

243. Chibani, S., Charaf-Eddin, A., Le Guennic, B. & Jacquemin, D. Boranil and related NBO dyes: insights from theory. *J. Chem. Theory Comput.* **9,** 3127–3135 (2013).

244. Charaf-Eddin, A., Le Guennic, B. & Jacquemin, D. Excited-states of BODIPY–cyanines: ultimate TD-DFT challenges? *RSC Adv.* **4,** 49449–49456 (2014).

245. Storm, F. E., Hansen, T. & Mikkelsen, K. V. Density functional theory investigation on boron-subphthalocyanine. *J. Nanosci. Res. Rep.* **152,** 2–10 (2023).

246. Jacquemin, D., Wathelet, V., Perpète, E. A. & Adamo, C. Extensive TD-DFT benchmark: Singlet-excited states of organic molecules. *J. Chem. Theory Comput.* **5,** 2420–2435 (2009).

247. Dziuk, B., Ośmiałowski, B., Zarychta, B., Ejsmont, K. & Chęcińska, L. Symmetric fluoroborate and its boron modification: Crystal and electronic structures. *Crystals* **9** (2019).

248. Zhang, B., Wang, S., Tan, J. & Zhang, X. Unique fluorescence of boronic acid derived salicylidenehydrazone complexes with two perpendicular ICT: Solvent effect on PET process. *Dyes Pigm.* **155,** 186–193 (2018).

249. Li, B. *et al.* Novel axially substituted lanthanum phthalocyanines: Synthesis, photophysical and nonlinear optical properties. *Dyes Pigm.* **179,** 108407 (2020).

250. Dowds, M. & Nielsen, M. B. Controlling the optical properties of boron subphthalocyanines and their analogues. *Mol. Syst. Des. Eng.* **6,** 6–24 (2021).

251. Alam, M. M., Chattopadhyaya, M., Chakrabarti, S. & Ruud, K. Chemical control of channel interference in two-photon absorption processes. *Acc. Chem. Res.* **47,** 1604–1612 (2014).

252. Jacquemin, D. *et al.* Assessment of the efficiency of long-range corrected functionals for some properties of large compounds. *J. Chem. Phys.* **126,** 144105 (2007).

253. Głodek, M., Petrusevich, E. F., Plażuk, D., Jacquemin, D. & Ośmiałowski, B. Polyaromatic hydrocarbon antennas as tools for tuning properties of push-pull difluoroborates. *Dyes Pigm.* **212,** 111112 (2023).

254. Shao, T. *et al.* Identification of fatty liver disease at diverse stages using two-photon absorption of triphenylamine-based BODIPY analogues. *J. Mater. Chem. B* **7,** 3704–3709 (2019).

255. Schäfer, A., Horn, H. & Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **97,** 2571–2577 (1992).

256. Yu, H. S., He, X., Li, S. L. & Truhlar, D. G. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.* **7,** 5032–5051 (2016).

257. Grabarz, A. M. & Ośmiałowski, B. Benchmarking density functional approximations for excited-state properties of fluorescent dyes. *Molecules* **26** (2021).

258. Yu, C. *et al.* Highly fluorescent BF2 complexes of hydrazine–Schiff base linked bispyrrole. *Org. Lett.* **16,** 3048–3051 (2014).

259. Huaulmé, Q., Mirloup, A., Retailleau, P. & Ziessel, R. Synthesis of highly functionalized BOPHY chromophores displaying large stokes shifts. *Org. Lett.* **17,** 2246–2249 (2015).

260. Li, X., Ji, G. & Son, Y.-A. Tunable emission of hydrazine-containing bipyrrole fluorine–boron complexes by linear extension. *Dyes Pigm.* **124,** 232–240 (2016).

261. Li, Y. *et al.* A BOPHY probe for the fluorescence turn-on detection of Cu^{2+}. *Sens. Actuators B Chem.* **235,** 33–38 (2016).

262. Ponce-Vargas, M., Azarias, C., Jacquemin, D. & Le Guennic, B. Combined TD-DFT-SOS-CIS(D) study of BOPHY derivatives with potential application in biosensing. *J. Phys. Chem. B* **121,** 10850–10858 (2017).

263. Hansch, C., Leo, A. & Taft, R. W. A survey of hammett substituent constants and resonance and field parameters. *Chem. Rev.* **91,** 165–195 (1991).

264. Shagurin, A. Y., Pogonin, A. E., Ivanova, M. A., Postnikova, D. A. & Marfin, Y. S. Molecular and electronic structure of substituted BODIPY dyes: Quantum chemical study. *Comput. Theor. Chem.* **1212,** 113719 (2022).

265. Bednarska, J. *et al.* Quantifying the performances of DFT for predicting vibrationally resolved optical spectra: Asymmetric fluoroborate dyes as working examples. *J. Chem. Theory Comput.* **13,** 4347–4356 (2017).

266. Jacquemin, D., Planchat, A., Adamo, C. & Mennucci, B. TD-DFT assessment of functionals for optical 0–0 transitions in solvated dyes. *J. Chem. Theory Comput.* **8,** 2359–2372 (2012).

267. Jacquemin, D., Duchemin, I. & Blase, X. 0–0 energies using hybrid schemes: Benchmarks of TD-DFT, CIS(D), ADC(2), CC2, and BSE/GW formalisms for 80 real-life compounds. *J. Chem. Theory Comput.* **11,** 5340–5359 (2015).

268. Loos, P.-F., Galland, N. & Jacquemin, D. Theoretical 0–0 energies with chemical accuracy. *J. Phys. Chem. Lett.* **9,** 4646–4651 (2018).

269. Jacquemin, D., Moore, B. I., Planchat, A., Adamo, C. & Autschbach, J. Performance of an optimally tuned range-separated hybrid functional for 0–0 electronic excitation energies. *J. Chem. Theory Comput.* **10,** 1677–1685 (2014).

270. Loos, P.-F. & Jacquemin, D. Chemically accurate 0–0 energies with not-so-accurate excited state geometries. *J. Chem. Theory Comput.* **15,** 2481–2491 (2019).

271. Besalú-Sala, P., Sitkiewicz, S. P., Salvador, P., Matito, E. & Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **22,** 11871–11880 (2020).

272. Cerezo, J. & Santoro, F. FCclasses3: Vibrationally-resolved spectra simulated at the edge of the harmonic approximation. *J. Comput. Chem.* **44,** 626–643 (2023).

273. Sitkiewicz, S. P., Zaleśny, R., Ramos-Cordoba, E., Luis, J. M. & Matito, E. How reliable are modern density functional approximations to simulate vibrational spectroscopies? *J. Phys. Chem. Lett.* **13,** 5963–5968 (2022).

274. Santoro, F., Improta, R., Lami, A., Bloino, J. & Barone, V. Effective method to compute Franck-Condon integrals for optical spectra of large molecules in solution. *J. Chem. Phys.* **126.** 084509 (2007).

275. Santoro, F., Lami, A., Improta, R. & Barone, V. Effective method to compute vibrationally resolved optical spectra of large molecules at finite temperature in the gas phase and in solution. *J. Chem. Phys.* **126,** 184102 (2007).

276. Santoro, F., Lami, A., Improta, R., Bloino, J. & Barone, V. Effective method for the computation of optical spectra of large molecules at finite temperature including the Duschinsky and Herzberg–Teller effect: The Qx band of porphyrin as a case study. *J. Chem. Phys.* **128,** 224311 (2008).

277. Santoro, F. & Barone, V. Computational approach to the study of the lineshape of absorption and electronic circular dichroism spectra. *Int. J. Quantum Chem.* **110,** 476–486 (2010).

278. Lami, A. & Santoro, F. in *Computational Strategies for Spectroscopy* 475–516 (John Wiley & Sons, Ltd, 2011). ISBN: 9781118008720.

279. Avila Ferrer, F. J., Cerezo, J., Soto, J., Improta, R. & Santoro, F. First-principle computation of absorption and fluorescence spectra in solution accounting for vibronic structure, temperature effects and solvent inhomogenous broadening. *Comput. Theor. Chem.* **1040-1041,** 328–337 (2014).

280. De Souza, B., Neese, F. & Izsák, R. On the theoretical prediction of fluorescence rates from first principles using the path integral approach. *J. Chem. Phys.* **148,** 034104 (2018).

281. Torrent-Sucarrat, M., Anglada, J. M. & Luis, J. M. Role of vibrational anharmonicity in atmospheric radical hydrogen-bonded complexes. *Phys. Chem. Chem. Phys.* **11,** 6377–6388 (2009).

282. Luis, J. M., Torrent-Sucarrat, M., Solà, M., Bishop, D. M. & Kirtman, B. Calculation of Franck–Condon factors including anharmonicity: Simulation of the $C_2H_4^+ \tilde{X}^2B_{3u} \leftarrow C_2H_4\tilde{X}^1A_g$ band in the photoelectron spectrum of ethylene. *J. Chem. Phys.* **122,** 184104 (2005).

283. Bishop, D. M., Luis, J. M. & Kirtman, B. Vibration and two-photon absorption. *J. Chem. Phys.* **116,** 9729–9739 (2002).

284. Bonness, S., Kirtman, B., Huix, M., Sanchez, A. J. & Luis, J. M. Simulation of photoelectron spectra with anharmonicity fully included: Application to the $\tilde{X}^2A_2 \leftarrow \tilde{X}^1A_1$ band of furan. *J. Chem. Phys.* **125,** 014311 (2006).

285. Medved, M., Stachová, M., Jacquemin, D., André, J.-M. & Perpète, E. A. A generalized Romberg differentiation procedure for calcultion of hyperpolarizabilities. *J. Mol. Struct. (THEOCHEM)* **847,** 39–46 (2007).

286. Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **108,** 058301 (2012).

287. Collins, C. R., Gordon, G. J., von Lilienfeld, O. A. & Yaron, D. J. Constant size descriptors for accurate machine learning models of molecular properties. *J. Chem. Phys.* **148,** 241718 (2018).

288. Himanen, L. *et al.* DScribe: Library of descriptors for machine learning in materials science. *Comput. Phys. Commun.* **247,** 106949 (2020).

289. McCarthy, M. & Lee, K. L. K. Molecule identification with rotational spectroscopy and probabilistic deep learning. *J. Phys. Chem. A* **124,** 3002–3017 (2020).

290. Różański, P. T. & Zieliński, M. Exploiting underlying crystal lattice for efficient computation of Coulomb matrix elements in multi-million atoms nanostructures. *Comput. Phys. Commun.* **287,** 108693 (2023).

291. Vovk, V. in *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik* (eds Schölkopf, B., Luo, Z. & Vovk, V.) 105–116 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).

292. Christensen, A. S. *et al.* QML: A Python toolkit for quantum machine learning. *URL https://github.com/qmlcode/qml* (2017).

List of Schemes

List of Figures

List of Tables

List of publications done during PhD studies

i) Ośmiałowski, B., **Petrusevich, L.**, Antoniak, M., Grela, I., Bin Jassar, M. A., Nyk, M. W., Luis. J. M., Jędrzejewska, B., Zaleśny, R. & Jacquemin, D. Controlling two-photon action cross section by changing a single heteroatom position in fluorescent dyes. *J. Phys. Chem. Lett.* **11** (15) 5920-5925 (2020). IF_{2020}=6.475.

ii) Ośmiałowski, B., **Petrusevich, L.**, Nawrot, K., Paszkiewicz, B., Nyk, M. W., Zielak, J., Jędrzejewska, B., Luis, J. M., Jacquemin, D. & Zaleśny, R. Tailoring the nonlinear absorption of fluorescent dyes by substitution at a boron center. *J. Mater. Chem. C* **9** (19) 6225-6233 (2021). IF_{2021}=8.067.

iii) **Petrusevich, L.**, Ośmiałowski, B., Zaleśny, R. & Alam, M. M. Two-photon absorption activity of BOPHY derivatives: insights from theory. *J. Phys. Chem. A* **125** (12) 2581-2587 (2021). IF_{2021}=2.994.

iv) **Petrusevich, L.**, Głodek, M., Antoniak, M., Muzioł, T. M., Plażuk, D., Siomra, A., Nyk, M. W., Ośmiałowski, B. & Zaleśny, R. Difluoroborate-based bichromophores: symmetry relaxation and two-photon absorption. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **295**, 122600 (2023). IF_{2022}=4.400.

v) **Petrusevich, L.,**, Bousquet, M. H. E., Ośmiałowski, B., Jacquemin, D., Luis, J. M. & Zaleśny, R. Cost-effective simulations of vibrationally-resolved absorption spectra of fluorophores with machine-learning-based inhomogeneous broadening. *J. Chem. Theory Comput.* **19** (8), 2304-2315 (2023). IF_{2022}=5.500.

vi) Kosevich, A., **Petrusevich, L.**, Maklakov, S., Naboko, A., Kolesnikov, E., Petrov, D., Zezyulina, P., Pokholok, K., Filimonov, D. & Han, M. Low weight hollow microspheres of iron with thin dielectric coating: synthesis and microwave permeability. *Coatings* **10** (10), 995 (2020). IF_{2020}=2.881.

vii) Gorbachev, D. A., **Petrusevich, L.**, Kabylda, A. M., Maksimov, E. G., Lukyanov, K. A., Bogdanov, A. M., Baranov, M. S., Bochenkova, A. V. & Mishin, A. S. A general mechanism of green-to-red photoconversions of GFP. *Frontiers in Molecular Biosciences* **7**, 176 (2020). IF_{2020}=5.246.

viii) Moshkina, T. N., Nosova, E. V., Lipunova, G. N., Valova, M. S., **Petrusevich, L.**, Zaleśny, R., Ośmiałowski, B. & Charushin, V.N. Substituted 2-(2-hydroxyphenyl)-3H-quinazolin-4-ones and their difluoroboron complexes: synthesis and photophysical properties. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **252**, 119497 (2021). IF_{2021}=4.831.

ix) Dubinina, T. V., Maklakov, S., **Petrusevich, L.**, Borisova, N. E., Trashin, S. A., de. Wael, K. & Tomilova, L. G. Photoactive layers for photovoltaics based on near-infrared absorbing aryl-substituted naphthalocyanine complexes: preparation and investigation of properties. *New J. Chem.* **45** (32) 14815-14821 (2021). IF_{2021}=3.925.

x) Chołuj, M., Behera, R., **Petrusevich, L.**, Bartkowiak, W., Alam M. M. & Zaleśny, R. Much of a muchness: on the origins of two- and three-photon absorption activity of dipolar Y-shaped chromophores. *J. Phys. Chem. A* **126** (5) 752-759 (2022). IF_{2022}=2.900.

xi) Gordel-Wójcik, M., Nyk, M. W., Śmiałek, P., **Petrusevich, L.**, Bartkowiak, W., Samoć, M. & Jędrzejewska, B. Linear and nonlinear absorption properties of hemicyanine dyes containing 5-(4H)-oxazolone as heterocyclic ring. *Dyes Pigm.* **208**, 110772 (2022). IF_{2022}=4.500.

xii) Głodek, M., **Petrusevich, L.**, Plażuk, D., Jacquemin, D., Ośmiałowski, B. Polyaromatic hydrocarbon antennas as tools for tuning properties of push-pull difluoroborates. *Dyes Pigm.* **212**, 111112 (2023). IF_{2022}=4.500.

xiii) Chorobiński, M., Śmiałek, P., Gordel-Wójcik, M., **Petrusevich, L.**, Zych, E., Plażuk, D., Skowronski, Ł. & Jędrzejewska, B. Photophysical and optical properties of 4-(4'-(diphenylamino)-benzylidene)-2-phenyloxazol-5(4H)-ones induced by the peripheral substituent. *J. Lumin.* **255**, 119584 (2023). IF_{2022}=3.600.