UNIVERSITY OF WARSAW
Faculty of Chemistry

WCh Faculty of Chemistry University of Warsaw

March 30th, 2024, Warsaw

Dr hab. Tatiana Korona, Univ. Prof.
Quantum Chemistry Laboratory
Pasteura 1, Faculty of Chemistry
University of Warsaw
Warsaw 02-093, Poland
t.korona@uw.edu.pl

Review of doctoral dissertation of Ms. Lizaveta Petrusevich entitled "Computer-aided design of two-photon-absorbing organic chromophores".

The doctoral dissertation written by Ms. Lizaveta Petrusevich and submitted to Wrocław University of Science and Technology, addresses an important topic of in silico design of two-photon absorbing molecules. A computer-aided design of new compounds of desired properties experiences a rapid development nowadays, but there are still many challenges hampering routine applications of theoretical quantum methods to real-life cases, which are caused by large size of potentially interesting molecules, the necessity of modelling them in solution, and to many other factors, which are difficult to include in a balanced way. Ms. Lizaveta Petrusevich deals in her thesis with a particularly complex phenomenon of two-photon absorption processes, for which a general protocol does not exist so far. Therefore the selection of such topic of the thesis is a guaranty of interesting results, which are expected to be very useful for the scientific community.

The dissertation, written in English, is composed of six chapters, preceded with the list of abbreviations. The work contains 41 figures and 10 tables, cites 292 references, and has a separate list of 13 publications of the present author. The consecutive chapters include an introduction, objectives of the dissertation, theoretical backgrounds, results and discussion, and a summary. I found the list of abbreviations at the beginning of the work particularly useful, since it allows to quickly recall the less-known abbreviation meanings without the need of browsing through the manuscript for the explanation. It should be emphasized that the author has an impressive list of papers related to the Ph.D. theme, published in very good journals. All these publications have several authors, therefore the decision of Ms. Lizaveta Petrusevich to write a standalone thesis (instead of a popular method of using publications and writing a short introduction), should be appreciated.

The first chapter describes recent developments in the field of 2PA dyes, lists relevant theoretical methods and presents arguments for the selection of a particular class of dyes, which will be studied in detail in the thesis. At the beginning of the chapter the author proves the importance of the study of the 2PA phenomenon, pointing out to its application for fluorescence induced by the two-photon

excitation, which is better suitable for bioimaging than fluorescence induced by the one-photon excitation because of the transparency of biological materials in the so-called first biological window. The author also lists requirements laid upon such materials, like the above-mentioned absorption in the respective range, high 2PA cross sections, high fluorescence quantum yield, plus the photostability. These considerations are followed by a discussion of known types of 2PA dyes, from which the author selects small organic molecules as candidates with the most potential. Afterwards the author discusses common structures of such molecules, which include one or more electron-accepting and electron-donating groups connected with the π-conjugated bridges, and introduces the main actors of her work, which are organic dyes containing the BF2 group. The final part of this chapter is devoted to theoretical methods with a discussion of their applications found in the literature, what gives the author an opportunity to justify her selections of the main workhorses of her thesis, which for the electronic part are: RI-CC2 for benchmarks and TD-DFT for the remaining calculations. Summarizing, this chapter is very well written - it provides a clear introduction to the topics which will be then investigated by the author.

The unsolved problems and questions raised in the introduction are neatly summarized in a short chapter, containing two main objectives of the dissertation, which are: the computer-aided design of fluorescent dyes with refined 1PA and 2PA properties suitable for the 2PA-based bioimaging applications and the development of simulation protocols for the estimation of vibrationally-resolved spectra with improved efficiency and accuracy.

Chapter 3 introduces a theoretical formalism for describing the 2PA phenomenon and provides formulas for all the quantities presented in the following chapters, including electronic-structure and vibronic-structure theories, as well as anharmonic corrections for the FC factors. Very detailed derivations and exhaustive explanations of the resulting formulas provide a solid ground for the following computational and discussion parts of the thesis.

In Chapter 4 Ms. Lizaveta Petrusevich presents a general protocol of simulations of the electronic spectra, also in the chloroform solution, and describes technical details of each stage of these calculations. This part is followed by a description of the strategy of designing new dyes, starting from the BF2 core moiety, which has been adopted by her in the thesis. The result of this strategy are various candidate molecules, which have been grouped into series from A to E, each containing about 10 molecules. The analysis of the electronic spectra of each series ends with the analysis of advantages and disadvantages of molecules from these series.

Chapter 5 provides a detailed presentation of the inclusion of vibronic effects computed with methods varying in the applied approximation levels, what finally allows for a selection of computationally affordable methods for tested set of molecules. One of the conclusions of this chapter is that the ωB97X and LC-BLYP-OT(α) functionals and the VG approach provide a better accuracy than a more involved AH method, which is important from the point of view of practical

application of the computational protocol for selection of new dyes. After making this conclusion Ms. Lizaveta Petrusevich emphasizes however that such an agreement might be the result of error cancellation and extends the study to include anharmonic corrections and HT effects. The final part of Chapter 5 introduces a novel method of estimations of inhomogeneous broadening of the spectra. The author compares two solvation models, PE and a simpler EE, performs MD simulations, and utilizes the CAM-B3LYP energies calculated for snapshots from the MD trajectory to feed the ML model. The key element in the application of the ML model is to find a suitable fingerprint representing the placement of solvent molecules. The author first presents the so-called Coulomb matrix available in the literature, which however has a huge drawback, as it grows quickly with the size of studied systems. In order to tackle this problem, Ms. Lizaveta Petrusevich proposes a much simpler one-dimensional formula and shows that indeed this formula is a suitable fingerprint within the ML model of spectra broadening due to the solvent effect. The final protocol based on the new fingerprint and reasonably small randomly selected snapshots, on which the ML is trained in order to predict the vertical spectra for the remaining snapshots, is a huge advantage in comparison to the classical way of calculations, which requires taking into account all the snapshots. The tests show that the application of the ML allows to significantly reduce the computational costs with no influence on accuracy. I find this protocol particularly useful for the scientific community.

The final Chapter 6 recapitulates the achievements of the dissertation.

The dissertation is well composed and it reads well in spite of presenting a lot of data, which is meticulously examined in Chapters 4 and 5. My following comments are minor and come from the reviewer's responsibility to point out debatable issues and perceived inaccuracies in the dissertation.

First, I have several questions to the author regarding the data interpretation in the dissertation:

1) From the data presented in Table 6 and the discussion of the planar or nonplanar arrangements of the C-6 molecule it seems that geometry parameters corresponding to the optimization with the C_i group symmetry and the constrained optimization in the C_1 symmetry are very close to each other. Is it possible that they are actually the same minimum and that tighter geometry optimization thresholds would make the optimization procedure to converge to exactly the same geometry in both cases?

2) The author attributes the discrepancy between experimental and theoretical spectra for the molecule D-8 to the incorrect interpretation of experimental data. Why does she consider the experiment less reliable than theory in this case?

3) When discussing Figure 20, where VG results are presented, the author writes that the used functionals are in excellent agreement for all the structures. However, for F-2 and F-3 the agreement is not so perfect in my opinion - the LC-BLYP-OT(α) spectra show a much more

resolved structure with separated peaks than for the ωB97X functional. Therefore, my question is: what makes the author to describe this agreement as "excellent"?

4) It is not clear from the text who wrote the "in-house code" for the dAH approach.

5) It is somewhat disappointing that the most advanced theoretical model (AH) produces worse results than a more approximate dAH. Is this a general rule for AH as compared to dAH?

6) Figure 26 shows a remarkable agreement between the experimental and VG absorption spectra, obtained with the LC-BLYP-OT(α) functional. Have these theoretically simulated spectra been shifted to match the experiment, like it has been suggested earlier for the RI-CC2 method?

7) In the PT methods a problem of random degeneracy was encountered and solved by removing the coupling between two transitions of too similar energies. Wouldn't it be a better idea to apply quasidegenerate PT for such cases?

Minor issues:

-Slash missing in the line following eq. 20 (in the inline formula for the derivative of the transition dipole moment).

-References for B3LYP are missing (p. 41).

-In Scheme 8 the meaning of colored red and blue bonds and atoms should be given in the caption, too.

Misprints etc.:

- double "a" in p. 29

- containinig -> containing p. 31

- computate -> compute p.38

- in -20×10^{-3} a.u. -> should be 10^{3} instead (since it is written that the term δ_{0101} has approximately the same magnitude), p. 43

"Despite here is illustrated the normalized spectra" p. 75 - it seems that this sentence is corrupted.

- a general notice for references in German: nouns in German are written with a capital letter

- "ppposite" -> "opposite" ref. 181

- missing pages in ref. 195

- ref. 254 is a repeated ref. 107

- "calcultion" -> "calculation" ref. 285

The minor flaws listed above do not diminish the value of the present Ph.D. dissertation, which constitutes a well-composed and self-contained entity with all elements smoothly fitting into their places. Also my questions to the author are mostly the expression of the reviewer's curiosity.

Summarizing, I can write with all confidence that the dissertation fulfills all the requirements imposed on doctoral dissertations, according to Article 187, paragraphs 1-2 of the Act of 20 July 2018, "Law on Higher Education and Science".

In addition, in my opinion the dissertation of Ms. Lizaveta Petrusevich exceeds a usual level expected for a Ph.D. dissertation and deserves a distinction (which implies graduation with honors). A particularly outstanding achievement, presented in the thesis, is the development of the machine learning approach to simulate inhomogeneous broadening. For the new approach to be applicable for large molecules, the author proposed a new geometrical fingerprint, describing the interaction between atoms of a dye molecule and all solvent molecules, which is computationally affordable for many-atom cases, and developed a new protocol, which includes the ML step in the process of spectra simulations. The author showed that the new approach could save over 90% of the computational time for such simulations without compromising the accuracy. Summarizing, the new protocol with the ML component makes such calculations accessible for much larger systems than before and represents an important contribution in the field of theoretical simulations of molecular spectra.

Tatiana Kolos

ul. Ludwika Pasteura 1, 02-093 Warsaw, Poland
phone.: 22 55 26211 (Dean's Office), 22 55 26230 (Administration)
e-mail: dziekan@chem.uw.edu.pl, chemia@chem.uw.edu.pl, www.chem.uw.edu.pl