MJFAS
Malaysian Journal of Fundamental and Applied Sciences

RESEARCH ARTICLE

Computational Insights into Nonlinear Optical Properties of Azobenzene Derivatives: A DFT-QSPR Approach

Mamoona Jillani[a], Suhaila Sapari[a], Syahrul Imran Abu Bakar[b], Fazira Ilyana Abdul Razak[a]*

[a]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia; [b]Atta-ur-Rahman Institute for Natural Products Discovery (AuRIns), Universiti Teknologi MARA, Puncak Alam Campus, 42300 Bandar Puncak Alam, Selangor Darul Ehsan, Malaysia

Abstract Theoretical analysis and predictions have become indispensable in guiding experimental advancements in nonlinear optics (NLO). This study investigates the NLO properties of azobenzene systems using computational strategies, specifically density functional theory (DFT) and quantitative structure-property relationship (QSPR) methods. While azobenzene's quantitative properties are well-documented, its NLO potential remains relatively unexplored, motivating this research. Azobenzene systems are promising for NLO applications due to their photoisomerization properties, enabling tuneable optical responses. In this study, DFT calculations, employing the B3LYP functional with the 6-31G(d,p) basis set, were used to determine the first hyperpolarizability (β) of various azobenzene derivatives. QSPR models were developed using descriptors derived from molecular geometry and electronic properties. Our DFT calculations revealed that azobenzene derivatives with electron-donating groups exhibited significantly higher hyperpolarizability, with β values reaching 12414 x 10^{-30} esu. The QSPR models demonstrated strong correlations between molecular descriptors and β values, providing a rapid screening tool for predicting NLO properties. This study provides a comparative analysis of DFT and QSPR methodologies in predicting NLO properties of azobenzene derivatives, offering insights for material optimization and highlighting the potential of azobenzene-based materials for NLO applications.

Keywords: Azobenzene derivatives, nonlinear optical properties, density functional theory, quantitative structure property relationship.

Introduction

Nonlinear optical (NLO) materials are essential for advancing technologies in telecommunications, information processing, and optical imaging. The ability to predict and design compounds with desired NLO properties is crucial for these fields. Computational chemistry, by analyzing molecular structures and properties, offers a robust method to understand the complex relationship between structure and NLO behavior. Combining computational chemistry with quantitative structure-property relationship (QSPR) modeling promises to enhance material design [1,2,3]. NLO materials exhibit optical phenomena, such as frequency doubling, harmonic generation, and electro-optic modulation, that are not linearly proportional to the applied light intensity. These phenomena are influenced by molecular structure and electronic characteristics. However, predicting and understanding the intricate relationships between molecular structure and NLO properties, which involve complex electronic and vibrational interactions, remain challenging [4,5].

Computational chemistry, utilizing quantum mechanical methods and simulations, provides powerful insights for exploring NLO properties. It allows for precise calculations of molecular electronic structures, response properties, and energy surfaces. Energy surfaces refer to the potential energy of a molecule as a function of its atomic coordinates, providing insights into molecular stability and reactivity. By

***For correspondence:**
fazirailyana@utm.my

Received: 23 Sept. 2024
Accepted: 04 March 2025

applying these techniques, researchers can derive molecular descriptors, which are numerical representations of molecular structure and properties, such as dipole moments, polarizabilities, hyperpolarizabilities, and excited-state characteristics [6,7].

Combining computational chemistry with QSPR modeling offers valuable opportunities for predicting NLO properties. QSPR models establish quantitative relationships between computed molecular descriptors and NLO properties obtained experimentally or theoretically. These models enable the estimation of NLO properties for new compounds or those difficult to study experimentally. Additionally, QSPR models reveal structural motifs and characteristics that significantly influence NLO behavior, aiding in material design and optimization [8,9].

However, while Density Functional Theory (DFT) provides high quantum mechanical accuracy in calculating NLO properties, it is computationally intensive, limiting its application to large datasets. QSPR, on the other hand, enables rapid property prediction through empirical correlations but may lack the accuracy of DFT for complex systems. Therefore, a synergistic approach combining DFT and QSPR is necessary to achieve both accuracy and efficiency in predicting NLO properties. Limited studies have explored this combined approach to enhance predictive accuracy, particularly for azobenzene systems. Previous studies, such as Dudek *et al*. [56], have explored the linear and nonlinear optical properties of azobenzene derivatives using DFT, showing the influence of specific substituents. Wu *et al*. [10] have reviewed the evolution of QSPR in drug discovery and materials science, highlighting the importance of descriptor selection. However, a comprehensive study that bridges the accuracy of DFT with the efficiency of QSPR for predicting NLO properties in azobenzene derivatives is lacking. This research gap motivates our study, which aims to develop and validate a combined DFT-QSPR approach to accurately and efficiently predict NLO properties in azobenzene systems.

Creating accurate and reliable QSPR models for NLO properties requires careful selection of suitable molecular descriptors and application of advanced statistical and machine learning methods. Techniques such as regression analysis, decision trees, support vector machines, and neural networks can be employed to develop robust and predictive models [10].

In this study, we aim to investigate the connection between computational chemistry and QSPR in NLO properties. By analyzing molecular descriptors and their associations with NLO behavior, we seek to identify critical structural and electronic factors influencing nonlinear optical responses. Our goal is to provide valuable insights and predictive tools for designing and discovering new NLO materials with enhanced performance.

Azobenzene–Based Compounds

The azobenzene moiety is one of the most extensively studied classes of NLO chromophores due to its variable polarizability within the π-system. Azobenzene (see Figure 1) organic materials can react to electromagnetic fields with frequencies up to 10^{14} Hz, making them suitable for applications such as optical switching, where the speed of information processing relies on this rapid optical response. Azobenzene exists in two isomeric forms, E (trans-) and Z (cis-), which are crucial for its photo-switching properties and contribute to its high optical nonlinearity. [11,12,13].

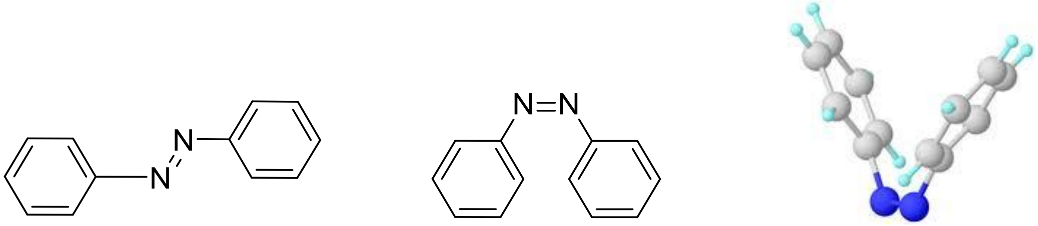


Figure 1. General structure of Azobenzene

Due to the unique cis-trans isomerization property of azobenzene chromophore, it has shown wide varieties of applications in different fields like photorefractive polymers, optical data storage on photochromism, liquid crystals, reversible optical waveguide and in medicine. In recent developed strategies, azobenzene photo switches show application in biological functions with long wavelength light operation to get the control of cell receptors and channels by penetrating deep into the tissues with lower photodamage using cis-state thermal stability [14,15].

Density Functional Theory (DFT) Parameters

In today's world of technology, theoretical chemistry is considered as a main tool to predict the behavior of broad range of chemical, physical, and biological phenomena of different material systems and has been applied to study nonlinear optical materials, molecular activation, organometallic reactivity, bioinorganic chemistry, and modeling of spectroscopic properties. Computational study offers help in understanding the molecular electronic structure, specifically in designing NLO materials. For quantitative calculation of transition metals, quantum chemical method is considered as a standard tool. Data obtained from computer simulations, such as geometry optimization, frequencies, bond dissociation energy, and electronic correlations, have proven to be reliable and comparable to experimental results. Quantum mechanical calculation like density functional theory (DFT) and semi-empirical methods simplify the experimental data in terms of electronic wave function and density [16,17,18].

DFT Approach in Designing NLO Materials

The first investigation of frequency-dependent polarizability, first and second order hyperpolarizability of organic compounds by using ab-initio method was done by Karna (1991). It showed good equivalence with the experimental values but the values of hyperpolarizability were lower in comparison with experimental values. So, it was recommended to use basis sets with diffused functions and electron correlations to get more accurate results [19,20,21].

To discover new NLO materials and identify effective NLO chromophores with strong hyperpolarizability values, DFT methods appear to be a promising approach for studying the NLO properties of complexes [22,23]. Materials having push-pull systems consisting of polarizable organic framework bonded to metal core that helps in increasing the conjugation within the structural framework considered as good candidates for NLO application. High NLO responses are shown by organic compounds having donor or acceptor groups that can increase the nonlinearity based on conjugation effects due to delocalized π-electrons within the framework.

Alyar [24] explored the NLO properties associated with the conjugation effects of a naphthalene derivative using various DFT methods and basis sets. The study, conducted with the B3LYP/6-31G basis set, identified the compound as a promising NLO candidate, with a total hyperpolarizability (βtot) of 1856.8 a.u. and a dipole moment (μ) of 2.78 D (Figure 2).

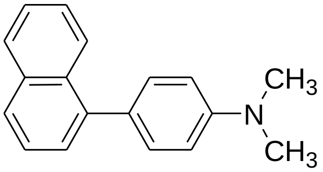


Figure 2. Structure of 1-(4- dimethylaminophenyl) naphthalene

Lakshmi *et al*. [25] examined the first hyperpolarizability (β_{tot}) of (2*E*)-3-[4-(dimethylamino) phenyl]-1-(naphthalen-2-yl) prop-2-en-1-one using the B3LYP/6-31G (d, p) basis set, which yielded a value of 63.02 x 10^{-30} esu significantly higher than the theoretical value of 0.13 x 10^{-30} esu, attributed to the increased conjugation within the molecule. The study also reported a high dipole moment (5.84 D), due to the strong donor-acceptor configuration, contributing to the compound's enhanced NLO activity (Figure 3).

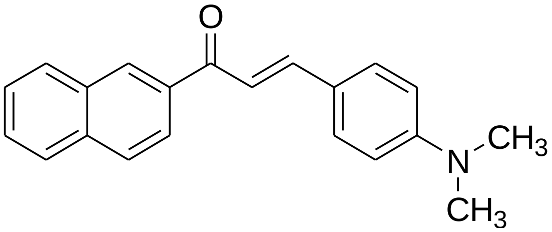


Figure 3. Structure of (2*E*)-3-[4-(dimethylamino) phenyl]-1- (naphthalen-2-yl) prop-2-en-1-one

Li *et al*. [26] investigated the structure, electronic absorption spectra, vibrational absorption spectra, and hyperpolarizability values of cis- and trans-azo dye 4-hydroxyazobenzene using DFT methods. The study was performed in both the gas phase and acetonitrile solvent, employing the 6-31G(d,p) basis set to calculate static electronic β and γ values. The results indicated that the trans-isomer exhibits a higher degree of third-order NLO response. The research also examined the functional dependence of hyperpolarizability values by comparing the *cis*- and *trans*-isomers. It was observed that increasing the Hartree-Fock (HF) exchange in conventional hybrid functionals such as O3LYP, B3LYP, B3PW91, and PBE0 leads to a decrease in both β and γ values.

In the study by Babić and co-workers [27], the geometries of dicyclopalladated complexes were optimized using the time-dependent DFT (TD-DFT) method to calculate adiabatic excited state energies and oscillatory strengths, utilizing the B3LYP functional and the 6-311+G(d, p) basis set, except for the palladium metal (Figure 4). The study concluded that solvation effects in DMSO shift the π-π* transition to the lowest energy state, resulting in the n-π* transition occurring in the second excited state. The absorption bands in monocyclopalladated azobenzenes exhibited a red shift that observed in the free ligands, indicating an enhancement in π-π* character as revealed by TD-DFT calculations. Additionally, in the presence of DMSO complexes, fluorescence activity was detected with emissions at 670 nm and 720 nm, following an excitation at 640 nm.

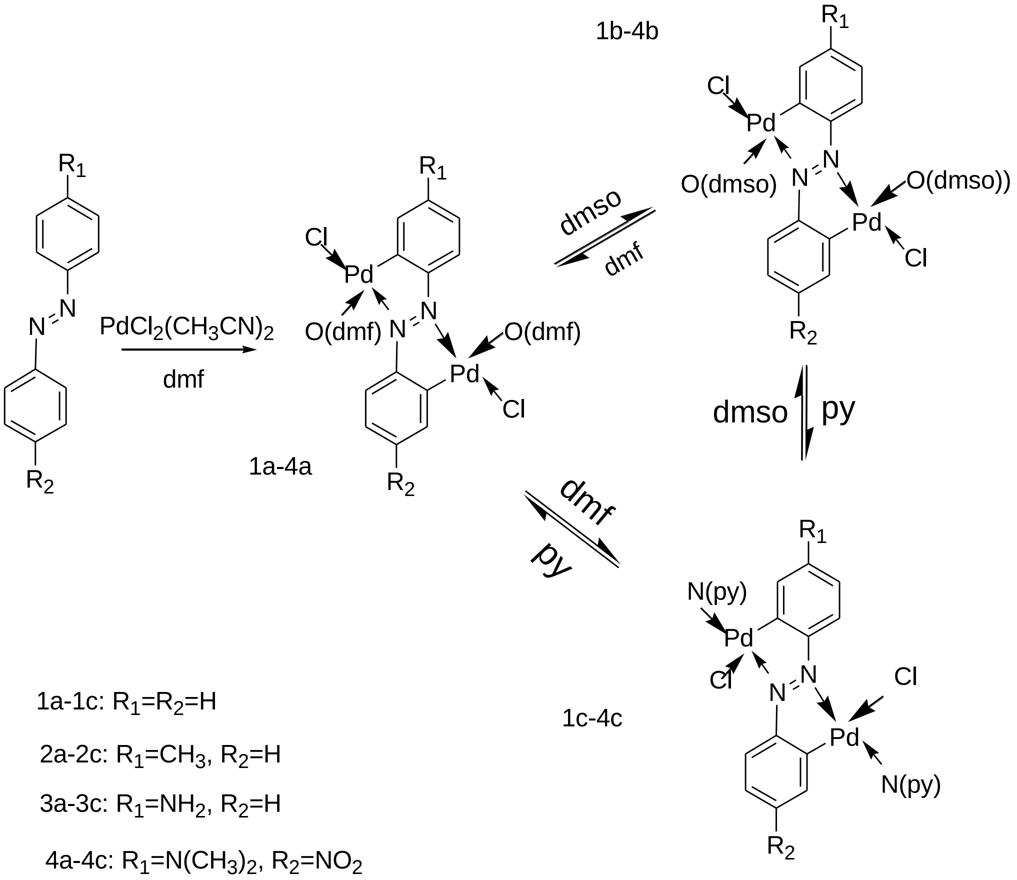
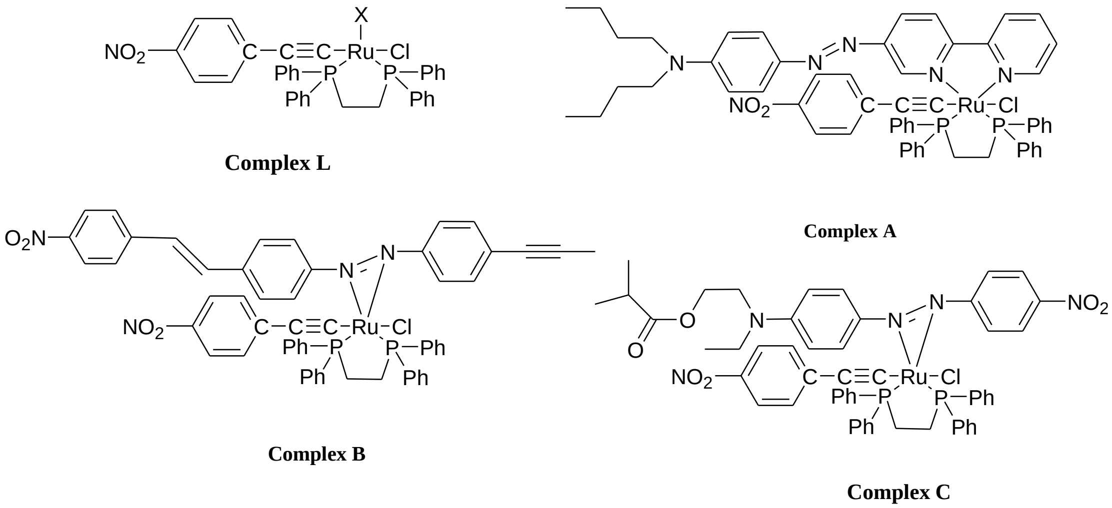


Figure 4. Reaction scheme synthesis for dicyclopalladated complexes of azobenzene derivatives

Woon *et al*. [28] studied the NLO of various azobenzene systems containing metals like ruthenium (Figure 5). The study involved comparing the Hartree-Fock (HF) method at the 3-21G level with density functional theory (DFT) using the LANL2DZ/6-31G basis set to examine parameters such as geometry optimization, bond lengths, and bond angles. The DFT method was employed to explore the NLO properties, specifically focusing on the total frequency-dependent first hyperpolarizability, βtot, at a wavelength of 1064 nm, which is influenced by electron delocalization. Among the three complexes studied, complex C exhibited the highest NLO property, with a βtot value of 12414.87 × 10^{-30} esu.



Complex L

Complex A

Complex B

Complex C

Figure 5. Studied complexes of azobenzene derivatives

Jaunet-lahary *et al*. [29] studied the NLO switching potentials of different azobenzene dimers and monomers using strong donating groups like NH_2, NMe_2, OMe, NPh_2, $N(Ph\text{-}OMe)_2$ and $N(Ph\text{-}NMe)_2$ (Figure 6). The static polarizability (α) and first hyperpolarizability (β) calculations were performed using DFT *w*B97X-D^{58}/6-311++G(d,p). In solvent effects, β values increase depending on the solvents used. β_{ratio} values for asymmetric azobenzene derivatives having different functional groups were found to be dimethylformamide, DMF: NO_2-CF_3 (2.12) < NMe_2-H (2.20) < H-CF_3 (2.37) < H-NO_2 (2.50) < NMe_2-CF_3 (2.54) < NMe_2-NO_2 (3.47). Hyperpolarizability value in DMF showed good percentage of nonlinearity β^{DMF} = +42 % as compared to other solvents. Polar solvents could stabilize the dipolar compounds which helps in increasing nonlinear behavior. Hyperpolarizability values (7.54×10^{-30} esu) < (16.43×10^{-30} esu) < (20.53×10^{-30} esu) of azobenzene with NPh_2, $N(Ph\text{-}OMe)_2$ and $N(Ph\text{-}NMe)_2$ in symmetry showed that the electron donor groups increase the nonlinear response in threefold.

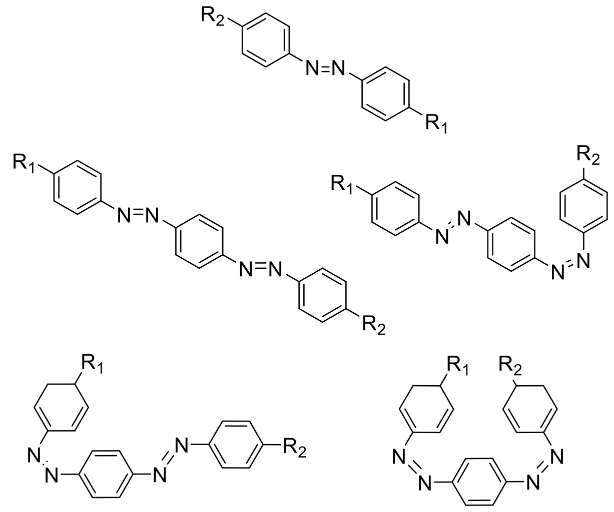


Figure 6. Azobenzene systems with electron donating groups (azobenzene derivatives having substituents NH_2, NMe_2, OMe, NPh_2, $N(Ph\text{-}OMe)_2$, $N(Ph\text{-}NMe_2)_2$, benzothiadiazole, azobenzenethiadiazole- relabelling and rearrangement

Liu *et al*. [30] studied the optical properties of azobenzene derivatives within D-π-A systems. Using DFT methods with the B3LYP/6-31G+(d) basis set, they modified the structures in the gas phase by extending conjugation lengths and increasing dipole moments, and further examined these modifications in tetrahydrofuran (THF) and methanol (MeOH) solvents. The results indicated that these structural changes could enhance both linear and second-order hyperpolarizability (Table 1).

Table 1. Second-order polarizabilities (β_0, in 10^{-30} esu) calculated at B3LYP/6–31+G(d) basis set, R_1=CF_3, R_2=H

Compound	Gas	THF	MeOH
1a	30	72 (52)	103
2a-d	44	130 (117)	198
3a-d	22	115	214
4a-d	135	314	425
5a-d	214	753	930

Chromophores with strong electron-donating groups (such as amino) and strong electron-accepting groups (like acetyl) exhibit higher second-order polarizability values compared to those with halogen-substituted fragments (R =–CF_3) > (R =–F, Cl, Br). Additionally, increasing the solvent polarity can further enhance the NLO properties (Figure 7). A computational study on bond lengths and bond angles was conducted for the non-centrosymmetric chalcone derivative, (2*E*)-3-(4-fluorophenyl)-1-(4-{[(1*E*)-(4-fluorophenyl)methylene]amino}phenyl)prop-2-en-1-one (FAMFC). The molecule was optimized using the B3LYP/6-31G basis set, yielding a carbon-oxygen bond length of 1.261 Å and a carbon-nitrogen bond length of 1.411 Å, both of which showed good agreement with experimental data [31].

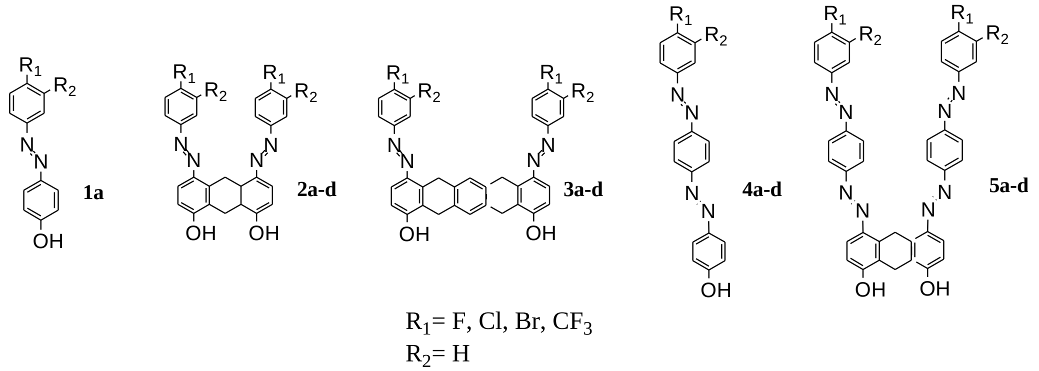


Figure 7. D-π-A systems of azobenzene

Quantitative Structure Property Relationship Parameters

Another computational approach that can be used to explore the relationship between experimental and theoretical aspects of materials is the quantitative structure-property relationship (QSPR). This method has gained importance with the rapid advancement of theoretical quantum chemical studies. The goal of QSPR is to derive mathematical equations using descriptors based on molecular properties. While there are thousands of potential descriptors, only a select few can accurately correspond to the desired physical, chemical, and biological properties (Table 2). If a strong correlation is identified, it allows for the prediction of properties for compounds that have not yet been synthesized or discovered [32].

Table 2. Summary of various descriptors explored in QSPR studies

Chemical properties
Boiling Point
Dielectric constant
Diffusion coefficient
Dissociation constant
Melting point
Reactivity
Solubility
Stability
Thermodynamic properties
Viscosity

QSPR Model

QSPR method is applied as mathematical models to anticipate the properties of physical characteristics and chemical structures of different compounds (Figure 8). Though few literatures are found for nonlinear behavior of azobenzene fragments, some of the work has been done on azobenzene as dyes and nonlinear behavior of other compounds [33,34,35].

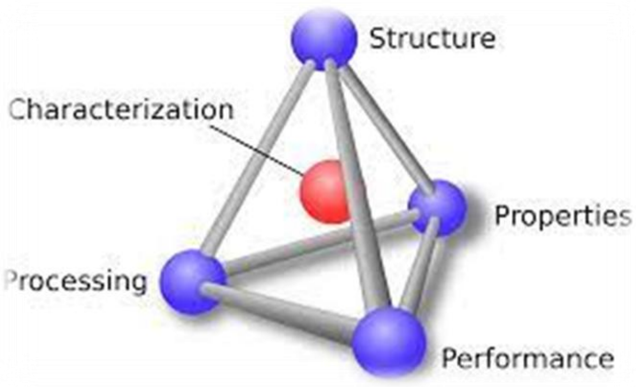


Figure 8. QSPR study flow diagram

QSPR Insights on Azobenzene Dyes

Azobenzenes were used as dyes and pigments after their first preparation in the mid-19th century. It usually absorbs light in the blue region and therefore a vast majority of azobenzene compounds show yellow, orange, or red color. A wide variety of substituents can be introduced to modify their color or increase their adsorption on different materials. As a result, azobenzenes probably become the most significant class of dye stuff which are still contributing half of the market for organic dyes and pigments in present days [36,37,38]

Schuurmann and Timofei [39] used MLR model to observe the affinity of azobenzene dyes for fibers. Statistical prediction showed that HOMO energy is comprehended in both linear and nonlinear form which is a major part of molecular electronegativity.

Fayet *et al*. [40] presented a model to approximate the absorption maxima of 22 azobenzene followed by DFT framework using MLR based on PCM-PBE0/6-31G calculations to quantify descriptors at different levels of theory. This study allows rapid and efficient dye designing.

Table 3 summarizes the parameters, software, and chemical properties of both DFT and QSPR models for understanding purpose.

Table 3. Summary of DFT and QSPR methodologies

DFT	QSPR
Gaussian software	Spartan, NWChem, MOLCAS, Q-Chem, MOPAC, Dragon software
Basis-sets	Descriptors
Optimization	Model development
Characterizations	Multiple linear regression
Dipole moment, HOMO-LUMO Band gap, Metal effects, Molecular mass	Dielectric constant, Reactivity, Solubility, Stability, Thermodynamic properties

Limitations of DFT and QSPR Methods

Both Density Functional Theory (DFT) and Quantitative Structure-Property Relationship (QSPR) modeling, despite their utility, possess inherent limitations. DFT's computational burden restricts its application to large systems, its accuracy depends heavily on functional and basis set selection, and it struggles with Van der Waals interactions and complex excited-state properties. Conversely, QSPR's accuracy relies on high-quality, diverse training data, faces challenges in descriptor selection and risks overfitting, lacks mechanistic insight, has a limited applicability domain, and is vulnerable to inaccuracies in the training dataset, necessitating careful consideration and validation.

Benefits and Applicability

Despite these limitations, DFT and QSPR remain indispensable tools in computational chemistry and materials science due to their unique strengths. DFT provides a fundamental, quantum mechanical understanding of electronic structure and properties, enabling accurate predictions of crucial properties like electronic transitions and hyperpolarizabilities, and its applicability expands with advancements in algorithms and hardware. QSPR excels in rapid screening and trend identification, guiding optimization processes through quantitative structure-property relationships. Furthermore, the synergistic use of DFT and QSPR mitigates individual limitations, with DFT providing accurate training data for robust QSPR models and QSPR enabling efficient exploration of chemical space, rendering them invaluable for scientific research and technological development.

QSPR Model Approach in Designing NLO Materials

QSPR applications are used to model the properties of the compounds with structure. QSPR relates to finding the relationship between absorption, hyperpolarizability or thermal stability with the structure of the compounds. Here we explain some of the QSPR models used for different activities.

QSPR applications are also reported to analyze the relationship between λ_{max} and structure of chromophores. Buttingsrud *et al*. [41] observed the relation between electronic structure and absorption maxima for azobenzene dyes using MLR method and nonlinear radial basis function neural network (RBFNN) method to conclude that bond length is important in regulating the colour of dyes.

Few attempts were made to design a relationship between physiochemical properties and organic compounds framework for second - order NLO studies. Oberg *et al*. [42] developed a QSPR model to calculate the second and third - order (β and γ) NLO quantities employing ab initio method to generate variables (5 chemical descriptors) and used partial least square analysis to build a model to predict the nonlinearities of 22 chromophores. Zeng *et al*. [43] predicted the nonlinear relationship of thirty-two para-disubstituted benzenes using a 3-parameter correlation. Li *et al*. [44] used a PM3 computation method to observe the data of 31 azodyes by studying the relationship between experimental λ_{max} and structural parameters. Xu *et al*. [45] established a linear QSPR model of 72 chromophores to study the second order NLO maximum absorption wavelength. They also investigated a set of 90 second order NLO fragments by studying the relationship between the descriptors of molecular structures and thermal decomposition temperatures (T_d).

Some other attempts are also made to study the nonlinear optical properties of different compounds. Alyar [24] has calculated nonlinear optical effects of naphthalene using extended basis sets followed by calculating molecular volume and miller-savchik polarizability using AM1 semi empirical method with HyperChem v7, resulting in a correlation coefficient value of 0.95 with DFT. Labidi [46] calculated optical properties that are empirical polarizabilities and hyperpolarizability values of polyacetylene chains by using AM1 followed by MOPAC 2000 in Hyperchem v7 showing regression coefficient of 0.91 and correlation coefficient of 0.93 between molecular calculations and Ab initio polarizability. Luan *et al.* [47] calculated the thermal stabilities of different second order NLO fragments using descriptors which are easy to read. To build a model, molecules were optimized in MOPAC6.0 software using semi-empirical PM3 method. Thirty-five constitutional descriptors were calculated using CODESSA software. Models built using multiple linear regression MLR method showed the nonlinear model with the correlation coefficient $R_2 = 0.494$ for training set and $R_2 = 0.867$ for test set.

Significance of the Previous Studies

The significance of the various studies mentioned lies in their contributions to the field of nonlinear optics and the understanding of the relationships between chemical structure and nonlinear optical (NLO) behavior. Each study addresses different aspects and compounds, adding valuable knowledge to this field. Oberg's [42] QSPR model for chromophores (β and γ) used ab initio methods and partial least squares analysis offers a systematic and theoretically grounded approach to understanding and predicting NLO effects in these compounds. Zeng's [43] correlation study on para-disubstituted benzenes established a 3-parameter correlation simplifies the prediction process for a specific class of compounds and helps identify key structural features affecting NLO properties. Li's [44] study on azodyes sheds light on the relationship between structural parameters and experimental λmax for azodyes. Understanding this connection is vital for optimizing the design of azodyes for specific NLO applications, such as optoelectronic devices. Xu's [45] linear QSPR model for chromophores expands the scope to 72 chromophores and focuses on the second order NLO maximum absorption wavelength. This broader dataset can contribute to a more comprehensive understanding of NLO properties in various chromophores. Xu's study on second order NLO fragments and thermal decomposition temperatures (Td) investigated the relationship between molecular descriptors and thermal decomposition temperatures for second order NLO fragments is valuable for materials science and safety considerations when designing NLO materials. Alyar's [24] study on naphthalene combines extensive basis sets, molecular volume calculations, and empirical polarizabilities to predict NLO effects in naphthalene. Achieving a high correlation coefficient with DFT signifies the reliability of the methodology and highlights its potential application in larger, more complex systems. Labidi's [46] study on polyacetylene chains investigation of empirical polarizabilities and hyperpolarizability values in polyacetylene chains provides insights into the NLO properties of extended conjugated systems. This knowledge can be valuable in the development of materials for NLO applications.

These studies collectively contribute to the understanding of NLO properties in various compounds. They use different approaches, such as QSPR models, empirical correlations, and computational methods, to explore these properties in different classes of molecules. Their significance lies in advancing the knowledge base and providing tools for predicting and optimizing nonlinear optical behavior in materials and compounds, which has practical applications in fields like photonics and optoelectronics.

Limitations in the Previous Work

The research conducted by Oberg, Jie Xu, Zheng Li, and Alyar also has source of errors in model testing, and correlations between different factors.

Variation in Experimental Conditions (Oberg's Study):

Oberg's study mentions that the responses were measured under different experimental conditions, which can introduce potential errors in their QSPR model. They believe that these errors are of minor importance because their model captures the responses well. This suggests that their model may be robust and capable of accommodating variations in conditions.

Testing the Model (Jie Xu's Study):

Jie Xu tested the model size and algorithm used by Oberg with randomized dependent variables. The model performed poorly (SEE of 54.6) with random data, indicating that it doesn't generate accurate predictions when the data is not structured, further validating the effectiveness of the real model for standard experiments.

Neglect of Solvent Effects (Zheng Li's Study):

Zheng Li's study points out that a source of error in predicting the maximum absorption wavelength (λmax) by their method is the neglect of solvent effects. Solvent effects can affect energy levels and the shape of absorption spectra, potentially shifting the λmax. This source of error emphasizes the need to account for solvation effects in certain cases.

Correlations by Alyar:

Alyar performed least-squares correlations between various sets of results for polarizability predictions. The correlations showed that these procedures are reliable predictors for smaller molecules but less reliable for larger molecules. The acceptable R-values for correlations of β and γ suggest that Jie Xu's model accounts for experimental errors well.

Dipole Moment and Polarizability Correlations with Calculation Methods by Labidi:

There are good least squares correlations between Ab initio (highly accurate) results and cheaper computational procedures, such as calculated molecular volume and Miller empirical polarizability models. However, semi-empirical AM1 models tend to underestimate the normal component of the polarizability tensor, suggesting limitations in this method.

Outlook

The researchers in their studies also acknowledged the various sources of error in their models, such as variations in experimental conditions, neglect of solvent effects, and model robustness. They also highlight the importance of accurate modeling and validation when dealing with complex experimental conditions and properties. The overall significance lies in improving the reliability and accuracy of predictive models for nonlinear optical properties.

QSPR Model Approach in Predicting NLO Response

The studies illustrate the diverse methodologies employed in the field of nonlinear optics, ranging from quantum mechanical calculations to empirical correlations and computational modeling. Each approach has its strengths and limitations. Their work enhances our understanding of the relationships between molecular structure and various properties, opening doors to potential applications in materials science and optics. However, the studies themselves provide a comprehensive QSPR model that can be directly applied to a wide range of compounds, also they contribute to the knowledge and methodology used in building such models.

Here's what these studies offer in terms of QSPR model development.

Software

Molecular descriptors are the chemical information of the molecules that signifies its physicochemical properties like chemical, constitutional, electronic, geometrical, solubility and quantum chemical descriptors which are encoded within the molecular structure for transforming the sets of algorithms. Such descriptors can be calculated using different software such as Gaussian, Spartan, NWChem, MOLCAS, Q-Chem, MOPAC or specialized software DRAGON, CODESSA, ADRIANA.Code and RECON [33,34]

Mathematical Modeling

Main concept of deriving a QSPR model through mathematical basis of chemical properties is correlated to the structure. Here, the activity describes the chemical properties "*x*": $y = f(x)$ by the function "*f*". To find "*x*", a few chemical moieties of the studied activity *(y)* with known values is required. A sequence of parameters called chemical descriptors is calculated for each chemical moiety. After that, an algorithm that gives an accurate value related to experimental result is calculated to check the validity of the model by observing if the algorithm is competent to calculate the properties of other chemical compounds. The validation phase is the key in predicting the correct statistical properties of the model [48,49].

Dataset and Molecular Descriptor Generation

Linear and nonlinear methods were investigated in QSPR for predicting the property of the newly designed proposed compounds. Descriptive and statistical studies (prediction of models) and evaluation of the suggested model of the novel compounds were conducted. Chemical descriptors can be of many types, but a few chemical descriptors will be discussed here that could help in studying NLO properties. In this study, two broad categories of molecular descriptors are involved: electronic descriptors (e.g., LUMO, HOMO, dipole moment, hyperpolarizabilities) and physicochemical or constitutional descriptors (e.g., molar refractivity, molecular weights) that can be calculated from various software. Descriptors can be directional and non-directional in terms of size, shape, symmetry, and distribution of atoms. A training set and test set are formed from the dataset. To obtain a nonlinear model, descriptors are submitted to a multiple nonlinear regression (MNLR) method, and for the linear model, the descriptors are submitted to a multiple linear regression (MLR) method. The models are then validated using internal (correlation coefficients r or r^2) and external (using the test set) validation techniques. Finally, evaluating and studying the proposed models with the desired properties is conducted. Crucially, the selection of molecular descriptors for NLO applications in QSPR modeling is not arbitrary but is guided by the underlying physics and chemistry of nonlinear optical phenomena. Specifically, electronic descriptors are vital, including the energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), as the HOMO-LUMO gap is directly related to electronic transition energies and influences the NLO response, with a smaller gap generally indicating higher polarizability and hyperpolarizability. The dipole moment, reflecting the charge distribution, is essential for understanding the molecule's response to an electric field, where higher dipole moments often correlate with stronger NLO responses. Hyperpolarizabilities (β, γ), direct measures of the nonlinear optical response calculated via quantum chemical methods, are used as both dependent and independent variables. Descriptors quantifying charge distribution and atomic charges provide insights into molecular polarization, and electronic polarizability reflects how easily the electron cloud distorts under an electric field [50,51,52].

Descriptive and Statistical Interpretation

Statistical methods include MLR which is the most simple and popular method of QSPR used to interpret the features of the descriptors by the help of the algorithm that is available and makes the predictions transparent. The second method is Partial Least Square (PLS) which is the abstraction of MLR that can interpret the data with strong correlation. Third method is MNLR used to build a model that is nonlinear by using the descriptors of MLR related to dataset (training set) [53,54]

Model Development and Validation

Model strength can be done by Leave-one-out cross validation technique in which dependent variables are randomly mixed up and used in the experiment. Model strength can be evaluated using the Leave-one-out cross-validation technique, where dependent variables are randomly mixed and used in the experiment. This process builds models and determines the cross-validation coefficient, $R^2_{training}$. For a model with good predictive ability, $R^2_{training}$ should be greater than 0.5. However, while $R^2_{training}$ provides insight into the model's internal consistency, external validation offers a more accurate representation of its true predictive power. External validation assesses the model's ability to predict the properties of not only the training set molecules but also, crucially, the test set, which can then be compared to experimental values. To achieve a reliable QSPR model, the external cross-validation coefficient, R^2_{test}, should also be greater than 0.5.

To further address the model validation process, it's essential to specify the specific steps taken to ensure the predictive accuracy of the QSPR models. For instance, in addition to R^2 values, other statistical metrics such as the root mean squared error (RMSE), mean absolute error (MAE), and the concordance correlation coefficient (CCC) were used to evaluate the model's performance. The RMSE and MAE provide information about the average prediction error, while the CCC assesses the agreement between predicted and observed values [55,47].

Conclusion

Overall, DFT and QSPR calculations are effective methods for studying molecular orbital energies and electronic spectra of organometallic NLO complexes. Computational simulations generally exhibit strong agreement with experimental results, underscoring the importance of selecting reliable methods with appropriate basis sets and levels of theory for accurately understanding and predicting chemical reactivity, catalytic activity, bioactivity, and linear and nonlinear optical behavior. Specifically, this study demonstrated that DFT calculations accurately predict molecular hyperpolarizability, with the B3LYP/6-31G(d) method showing deviations within 5-10% of experimental data. Furthermore, QSPR models successfully identified key molecular descriptors influencing NLO properties.

However, it is challenging to draw definitive conclusions due to the limited research conducted on azobenzene-based complexes for NLO applications. Nonetheless, the potential of these materials for NLO applications is significant, and computational approaches are crucial for predicting their performance. The practical implications of these findings for designing new NLO materials are substantial, as they provide a pathway to rapidly screen and optimize azobenzene derivatives with enhanced NLO properties before experimental synthesis.

To further advance this field, future research directions include integrating machine learning with QSPR models to enhance prediction accuracy and computational efficiency. Additionally, experimental validation of high-performing azobenzene derivatives is strongly recommended to confirm theoretical predictions. Continued research in this area is highly encouraged, as azobenzene-based materials possess promising characteristics for NLO applications, and computational prediction of their performance is both crucial and beneficial.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgement

This study was supported by the Ministry of Higher Education, Malaysia, under grant number FRGS/1/2021/STG04/UTM/02/7, and by Universiti Teknologi Malaysia with grant number UTM CG, PY/2024/02436/Q.J130000.3054.05M08, also Computing Information Centre Technology (CICT) for providing High Performance Computer in Gaussian 16 simulation.

References

[1] Zuo, J., & You, X. (2001). Chemical studies on the nonlinear optics of coordination compounds. *Chinese Science Bulletin, 46*(3), 178–184.

[2] Pate, B. D., Thorne, J. R., Click, D. R., Chisholm, M. H., & Denning, R. G. (2002). Third-order nonlinear optical properties of complexes with MM triple and quadruple bonds (M = Mo, W) at 1064 nm by degenerate four-wave mixing. *Inorganic Chemistry, 41*(7), 1975–1978.

[3] Rajasekhar, B., Bodavarapu, N., Sridevi, M., Thamizhselvi, G., RizhaNazar, K., Padmanaban, R., & Swu, T. (2018). Nonlinear optical and G-Quadruplex DNA stabilization properties of novel mixed ligand copper (II) complexes and coordination polymers: Synthesis, structural characterization, and computational studies. *Journal of Molecular Structure, 1156*, 690–699.

[4] Dalton, L. R., Günter, P., Jazbinsek, M., Sullivan, P. A., & Kwon, O. P. (2015). *Organic electro-optics and photonics: Molecules, polymers and crystals*. Cambridge University Press.

[5] Leuthold, J., Koos, C., Freude, W., Alloatti, L., Palmer, R., Korn, D., & Wehrli, S. (2013). Silicon-organic hybrid electro-optical devices. *IEEE Journal of Selected Topics in Quantum Electronics, 19*(6), 114–126.

[6] Luan, F., Liu, H., Gao, Y., & Zhang, X. (2009). QSPR model to predict the thermal stabilities of second-order nonlinear optical (NLO) chromophore molecules. *Molecular Simulation, 35*(3), 248–257.

[7] Berhanu, W. M., Pillai, G. G., Oliferenko, A. A., & Katritzky, A. R. (2012). Quantitative structure–activity/property relationships: The ubiquitous links between cause and effect. *ChemPlusChem, 77*(7), 507–517.

[8] McCoy, E. F., & Sykes, M. J. (2003). Quantum-mechanical QSAR/QSPR descriptors from momentum-space wave functions. *Journal of Chemical Information and Computer Sciences, 43*(2), 545–553.

[9] Santos, C. B., Lobato, C. C., Braga, F. S., Morais, S. S., Santos, C. F., Fernandes, C. P., & Carvalho, J. C. (2014). Application of Hartree-Fock method for modeling of bioactive molecules using SAR and QSPR. *Computational Molecular Bioscience, 4*(1), 1–9.

[10] Wu, K., Natarajan, B., Morkowchuk, L., Krein, M., & Breneman, C. M. (2013). From drug discovery QSAR to predictive materials QSPR: The evolution of descriptors, methods, and models. In *Materials science and engineering* (Chapter 16). Elsevier Inc.

[11] Chen, Z., Dong, S., Zhong, C., Zhang, Z., Niu, L., Li, Z., & Zhang, F. (2009). Photoswitching of the third-order nonlinear optical properties of azobenzene-containing phthalocyanines based on reversible host–guest interactions. *Journal of Photochemistry and Photobiology A: Chemistry, 206*(2–3), 213–219.

[12] Krawczyk, P., Kaczmarek, A., Zaleśny, R., Matczyszyn, K., Bartkowiak, W., Ziółkowski, M., & Cysewski, P. (2009). Linear and nonlinear optical properties of azobenzene derivatives. *Journal of Molecular Modeling, 15*(6), 581–590.

[13] Liaros, N., Couris, S., Maggini, L., De Leo, F., Cattaruzza, F., Aurisicchio, C., & Bonifazi, D. (2013). NLO response of photoswitchable azobenzene-based materials. *ChemPhysChem, 14*(13), 2961–2972.

[14] Kondo, T., Kanai, T., & Uosaki, K. (2001). Control of the charge-transfer rate at a gold electrode modified with a self-assembled monolayer containing ferrocene and azobenzene by electro- and photochemical structural conversion of *cis* and *trans* forms of the azobenzene moiety. *Langmuir, 17*(20), 6317–6324.

[15] Feng, W., Luo, W., & Feng, Y. (2012). Photo-responsive carbon nanomaterials functionalized by azobenzene moieties: Structures, properties, and application. *Nanoscale, 4*(20), 6118–6134.

[16] Siegbahn, P. E., & Blomberg, M. R. (2000). Transition-metal systems in biochemistry studied by high-accuracy quantum chemical methods. *Chemical Reviews, 100*(2), 421–438.

[17] Zhao, Y., & Truhlar, D. G. (2006). A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Physics, 125*(19), 194101.

[18] Zhao, Y., & Truhlar, D. G. (2008). 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. *Theoretical Chemistry Accounts, 120*(1), 215–241.

[19] Karna, S. P., Prasad, P. N., & Dupuis, M. (1991). Nonlinear optical properties of *p*-nitroaniline: Ab

[20] Sałek, P., Vahtras, O., Helgaker, T., & Ågren, H. (2002). Density-functional theory of linear and nonlinear time-dependent molecular properties. *Journal of Chemical Physics, 117*(21), 9630–9645.

[21] Scalmani, G., Frisch, M. J., Mennucci, B., Tomasi, J., Cammi, R., & Barone, V. (2006). Geometries and properties of excited states in the gas phase and in solution: Theory and application of a time-dependent

density functional theory polarizable continuum model. *Journal of Chemical Physics, 124*(9), 094107.

[22] Kodikara, M. S., Stranger, R., & Humphrey, M. G. (2018). Computational studies of the nonlinear optical properties of organometallic complexes. *Coordination Chemistry Review, 375*, 389–409.

[23] Águila, J. C., & Trejo-Durán, M. (2018). Theoretical study of the second-order nonlinear optical properties of ionic liquids. *Journal of Molecular Liquids, 269*, 833–838.

[24] Alyar, H. (2013). A review on nonlinear optical properties of donor-acceptor derivatives of naphthalene and azanaphthalene. *Review of Advanced Material Science, 34*, 79–87.

[25] Lakshmi, C. S. N., Balachandran, S., Arul, D. D., Ronaldo, A. A., & Hubert, J. I. (2019). DFT analysis on spectral and NLO properties of (2E)-3-[4-(dimethylamino)phenyl]-1-(naphthalen-2-yl)prop-2-en-1-one; a d-π-A chalcone derivative and its docking studies as a potent hepatoprotective agent. *Chemical Data Collection, 20*, 100205.

[26] Li, H. P., Bi, Z. T., Fu, W. Y., Xu, R. F., Zhang, Y., Shen, X. P., & Han, K. (2017). Theoretical study of the spectroscopic and nonlinear optical properties of trans- and cis-4-hydroxyazobenzene. *Journal of Molecular Modeling, 23*(3), 79.

[27] Babic, D., Curic, M., Molčanov, K. I., Ilc, G., & Plavec, J. (2008). Synthesis and characterization of dicyclopalladated complexes of azobenzene derivatives by experimental and computational methods. *Inorganic Chemistry, 47*(22), 10446–10454.

[28] Azobenzena, M. (2020). Computational studies on nonlinear optical properties of metal complexes containing azobenzene. *Malaysian Journal of Analytical Science, 24*(5), 719–726.

[29] Jaunet-Lahary, T., Chantzis, A., Chen, K. J., Laurent, A. D., & Jacquemin, D. (2014). Designing efficient azobenzene and azothiophene nonlinear optical photochromes. *Journal of Physical Chemistry C, 118*(49), 28831–28841.

[30] Liu, C. G., Guan, X. H., & Su, Z. M. (2011). Computational study on redox-switchable 2D second-order nonlinear optical properties of push−pull mono-tetrathiafulvalene-bis(salicylaldiminato) Zn(II) Schiff base complexes. *Journal of Physical Chemistry C, 115*(13), 6024–6032.

[31] Maidur, S. R., Patil, P. S., Ekbote, A., Chia, T. S., & Quah, C. K. (2017). Molecular structure, second- and third-order nonlinear optical properties and DFT studies of a novel non-centrosymmetric chalcone derivative: (2E)-3-(4-fluorophenyl)-1-(4-{[(1E)-(4-fluorophenyl)methylene]amino}phenyl)prop-2-en-1-one. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 184*, 342–354.

[32] Karelson, M., Lobanov, V. S., & Katritzky, A. R. (1996). Quantum-chemical descriptors in QSAR/QSPR studies. *Chemical Reviews, 96*(3), 1027–1044.

[33] Nantasenamat, C., Isarankura-Na-Ayudhya, C., Naenna, T., & Prachayasittikul, V. (2009). A practical overview of quantitative structure-activity relationship. *EXCLI Journal, 8*, 74–88.

[34] Katritzky, A. R., Kuanar, M., Slavov, S., Hall, C. D., Karelson, M., Kahn, I., & Dobchev, D. A. (2010). Quantitative correlation of physical and chemical properties with chemical structure: Utility for prediction. *Chemical Reviews, 110*(10), 5714–5789.

[35] Ojha, P. K., Mitra, I., Das, R. N., & Roy, K. (2011). Further exploring rm² metrics for validation of QSPR models. *Chemometrics and Intelligent Laboratory Systems, 107*(1), 194–205.

[36] Reuter, R. (2011). Synthesis, properties, and applications of macrocyclic azobenzenes. *University of Basel*.

[37] Hiroto, S. (2019). Synthesis of π-functional molecules through oxidation of aromatic amines. *Chemistry – An Asian Journal, 14*(15), 2514–2523.

[38] Reuter, R., Hostettler, N., Neuburger, M., & Wegner, H. A. (2010). Oligoazobenzenes—Switching in a new dimension. *Chimia, 64*(3), 180–183.

[39] Schüürmann, G., & Funar-Timofei, S. (2003). Multilinear regression and comparative molecular field analysis (CoMFA) of azo dye−fiber affinities. 2. Inclusion of solution-phase molecular orbital descriptors. *Journal of Chemical Information and Computer Sciences, 43*(5), 1502–1512

[40] Reuter, R., Hostettler, N., Neuburger, M., & Wegner, H. A. (2010). Oligoazobenzenes-switching in a new dimension. *Chimia*, 64(3), 180–183.

[41] Buttingsrud, B., Alsberg, B. K., & Åstrand, P. O. (2007). Quantitative prediction of the absorption maxima of azobenzene dyes from bond lengths and critical points in the electron density. *Physical Chemistry Chemical Physics, 9*(18), 2226–2233.

[42] Öberg, K., Berglund, A. U., & Eliasson, B. (2001). Prediction of nonlinear optical responses of organic compounds. *Journal of Chemical Information and Computer Sciences, 41*(3), 811–814.

[43] Xu, J., Zheng, Z., Chen, B., & Zhang, Q. (2006). A linear QSPR model for prediction of maximum absorption wavelength of second-order NLO chromophores. *QSAR & Combinatorial Sciences, 25*(4), 372–379.

[44] Li, G. Z., Yang, J., Song, H. F., Yang, S. S., Lu, W. C., & Chen, N. Y. (2004). Semiempirical quantum chemical method and artificial neural networks applied for λmax computation of some azo dyes. *Journal of Chemical Information and Computer Sciences, 44*(6), 2047–2050.

[45] Xu, J., Chen, B., Xu, W., Zhao, S., Yi, C., & Cui, W. (2007). 3D-QSPR modeling and prediction of nonlinear optical responses of organic chromophores. *Chemometrics and Intelligent Laboratory Systems, 87*(2), 275–280.

[46] Labidi, N. S. (2016). Semi-empirical and ab initio methods for calculation of polarizability (α) and the hyperpolarizability (β) of substituted polyacetylene chain. *Arabian Journal of Chemistry, 9*, S1252–S1259.

[47] Liu, L., & Özsu, M. T. (Eds.). (2009). *Encyclopedia of database systems* (Vol. 6). Springer.

[48] Roy, K., Kar, S., & Das, R. N. (2015). *A primer on QSAR/QSPR modeling: Fundamental concepts*. Springer.

[49] Dearden, J. C. (2017). The history and development of quantitative structure-activity relationships (QSARs). In *Oncology: Breakthroughs in research and practice* (pp. 67–117). IGI Global.

[50] Chtita, S., Larif, M., Ghamali, M., Adad, A., Rachid, H., Bouachrine, M., & Lakhlifi, T. (2013). Studies of two different cancer cell lines activities (MDAMB-231 and SK-N-SH) of imidazo[1,2-a]pyrazine derivatives by combining DFT and QSAR results. *Studies, 2*(11).

[51] Chtita, S., Bouachrine, M., & Lakhlifi, T. (2016). Basic approaches and applications of QSAR/QSPR methods.

Revue Interdisciplinaire, 1(1).

[52] Mammino, L. (2020). Computational chemistry for green design in chemistry and pharmacy: Building awareness in the classroom. *Sustainable Chemistry and Pharmacy, 18*, 100283.

[53] Cronin, M. T. (2010). Quantitative structure–activity relationships (QSARs)–applications and methodology. In *Recent advances in QSAR studies* (pp. 3–11). Springer.

[54] Polishchuk, P. (2017). Interpretation of quantitative structure–activity relationship models: Past, present, and future. *Journal of Chemical Information and Modeling, 57*(11), 2618–2639.

[55] Ekins, S., Bravi, G., Binkley, S., Gillespie, J. S., Ring, B. J., Wikel, J. H., & Wrighton, S. A. (2000). Three-and four-dimensional-quantitative structure activity relationship (3D/4D-QSAR) analyses of CYP2C9 inhibitors. *Drug Metabolism and Disposition, 28*(8), 994–1002.

[56] Dudek, M., Kaczmarek-Kędziera, A., Deska, R., Trojnar, J., Jasik, P., Młynarz, P., Samoc, M., & Matczyszyn, K. (2022). Linear and nonlinear optical properties of azobenzene derivatives modified with an (amino)naphthalene moiety. *The Journal of Physical Chemistry B, 126*(32), 6063–6073.