Quantitative Structure-Property Relationship Modelling for the Prediction of Singlet Oxygen Generation by Heavy-Atom-Free BODIPY Photosensitizers**

Andrey A. Buglak,[a] Asterios Charisiadis,[b] Aimee Sheehan,[c] Christopher J. Kingsbury,[b] Mathias O. Senge,*[d] and Mikhail A. Filatov*[c]

Abstract: Heavy-atom-free sensitizers forming long-living triplet excited states *via* the spin-orbit charge transfer intersystem crossing (SOCT-ISC) process have recently attracted attention due to their potential to replace costly transition metal complexes in photonic applications. The efficiency of SOCT-ISC in BODIPY donor-acceptor dyads, so far the most thoroughly investigated class of such sensitizers, can be finely tuned by structural modification. However, predicting the triplet state yields and reactive oxygen species (ROS) generation quantum yields for such compounds in a particular solvent is still very challenging due to a lack of established quantitative structure-property relationship (QSPR) models. In this work, the available data on singlet oxygen generation quantum yields (Φ_Δ) for a dataset containing >70 heavy-atom-free BODIPY in three different solvents (toluene, acetonitrile, and tetrahydrofuran) were analyzed. In order to build reliable QSPR model, a series of new BODIPYs were synthesized that bear different electron donating aryl groups in the *meso* position, their optical and structural properties were studied along with the solvent dependence of singlet oxygen generation, which confirmed the formation of triplet states via the SOCT-ISC mechanism. For the combined dataset of BODIPY structures, a total of more than 5000 quantum-chemical descriptors was calculated including quantum-chemical descriptors using density functional theory (DFT), namely M06-2X functional. QSPR models predicting $\Phi\Delta$ values were developed using multiple linear regression (MLR), which perform significantly better than other machine learning methods and show sufficient statistical parameters ($R = 0.88–0.91$ and $q^2 = 0.62–0.69$) for all three solvents. A small root mean squared error of 8.2% was obtained for Φ_Δ values predicted using MLR model in toluene. As a result, we proved that QSPR and machine learning techniques can be useful for predicting $\Phi\Delta$ values in different media and virtual screening of new heavy-atom-free BODIPYs with improved photosensitizing ability.

Introduction

Photosensitizers (PSs) which efficiently form long-lived triplet excited states are crucially important in such fields as photoredox catalysis,[1] photodynamic therapy (PDT)[2] and triplet-triplet annihilation upconversion.[3] Common approach for enhancing the triplet state yield (Φ_T) in organic chromophores relies on the introduction of heavy atoms, such as halogens (Br or I) or transition metals (e.g., Ru, Pd or Pt) in the structure (Figure 1A), which promote the intersystem crossing process via spin-orbit coupling interactions.[4] However, introduction of heavy atoms often requires tedious synthesis and thus leads to high cost of such photosensitizers.[5] Moreover, the presence of heavy atoms can also result in shortening the triplet state lifetimes.[6] The replacement of costly transition metal complexes in industrial-scale photocatalytic processes with organic PSs has drawn considerable attention,[7] and has stimulated a search for alternative methods to promote ISC not relying on

[a] *Dr. A. A. Buglak*
Faculty of Physics
Saint-Petersburg State University
Universiteteskaya Emb. 7–9, 199034 St. Petersburg (Russia)

[b] *Dr. A. Charisiadis, Dr. C. J. Kingsbury*
Chair of Organic Chemistry
School of Chemistry
Trinity Biomedical Sciences Institute
Trinity College Dublin
The University of Dublin
152-160 Pearse Street, Dublin 2 (Ireland)

[c] *A. Sheehan, Dr. M. A. Filatov*
School of Chemical and Pharmaceutical Sciences
Technological University Dublin
City Campus, Kevin Street, Dublin 8 (Ireland)
E-mail: mikhail.filatov@tudublin.ie

[d] *Prof. Dr. M. O. Senge*
Institute for Advanced Study (TUM-IAS)
Technical University of Munich
Lichtenberg-Str. 2a, 85748 Garching (Germany)
E-mail: sengem@tcd.ie

[**] *A previous version of this manuscript has been deposited on a preprint server (https://doi.org/10.26434/chemrxiv.14199335.v1).*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202100922

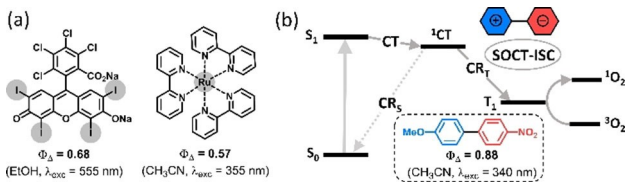

Figure 1. (a) Examples of heavy-atom-containing dyes. Φ_Δ-singlet oxygen quantum yield (solvent and excitation wavelength are given). (b) Jablonski diagram illustrating the formation of triplet excited state *via* the SOCT-ISC mechanism.

the heavy atom effect, for example, using a spin converter,[8] exciton coupling,[9] doubly-substituted excited states,[10] twisting of the aromatic systems[11] and radical-induced ISC have been very actively studied in recent years.[12]

The formation of triplet excited states in electron donor-acceptor dyads *via* the process of spin-orbit charge transfer intersystem crossing (SOCT-ISC) does not require introduction of transition metals or other heavy atoms into the molecule. So far, it has been observed in donor-acceptor dyads based on BODIPYs[13] and other difluoroboron complexes,[14] metal dipyrrins,[15] phenoxazines,[16] biphenyls,[17] naphthalene and perylene imides.[18] In these molecules, photoinduced electron transfer between the donor (D) and acceptor (A) subunits leads to formation of a charge-transfer state (CT), which further undergoes a non-radiative charge recombination into the ground state (CR_s) or into the lowest triplet excited state (CR_T) *via* SOCT-ISC (Figure 1B).[19] The latter process is commonly observed for dyads with orthogonal mutual arrangement of donor and acceptor subunits, which induces a large variation of the orbital magnetic momentum during the recombination process and thus compensates the change of spin magnetic momentum.[20] High yields of triplet state formation are observed for molecules in which the CT state lies close in energy to the lowest singlet excited state (S_1). In this case, the CR_s process falls within the Marcus inverted region[21] and is relatively slow due to a large negative value of the free energy change (ΔG_{CRS}). Under such circumstances, the CR_T process is considerably faster due to a smaller energy gap between the CT state and the lowest triplet excited state (T_1).[22]

Many reported molecular systems undergoing SOCT-ISC exhibit triplet state yields and singlet oxygen generation quantum yields (Φ_Δ) values that are comparable or even higher than those of transition metal complexes and halogenated dyes. Such dyes have other important advantages, for example, synthetic accessibility, high phototoxicity in cells (nM–μM range) with negligible dark toxicity,[23] long triplet excited state lifetimes (hundreds of μs),[24] and intense absorption in the 400–600 nm region (extinction coefficients up to 10^5).[25]

BODIPY dyes have been actively employed in the design of SOCT-ISC photosensitizers due to their excellent photophysical properties, ease of derivatization and predictable structural parameters (1464 crystal structures reported in the CCDC CSD in 2020).[26] In particular, a series of BODIPY dyads containing different electron donors, such as anthracene,[27] pyrene,[28] perylene,[29] phenothiazine,[30] carbazole,[31] and phenoxazine[32] has been systematically investigated. Recently, we reported biocompatible derivatives of such dyads, bearing polar solubilizing groups for enhancing cell penetration and studied their toxicity as well as fluorogenic response to singlet oxygen in cancer cells.[33] A library of BODIPYs having high Φ_Δ values (up to 70%) has been screened by us as candidates for PDT with potential clinical relevance.[34] Concurrently, we explored application of these sensitizers in the process of triplet-triplet annihilation upconversion (TTA-UC) and showed their unprecedented dual performance, namely the ability to play a role of either a sensitizer or an emitter component, depending on the media polarity.[35] We also found another unique feature of BODIPY dyads -a relatively strong fluorescence from the CT state and, using this property, developed a method for precise determination of TTA-UC quantum efficiency.[36]

As has been shown in previous works, the key factors which affect the efficiency of SOCT-ISC in BODIPY dyads are: 1) mutual orientation of the donor and acceptor subunits; 2) values of the driving force for the charge separation and recombination processes (ΔG_{CT} and ΔG_{CR}, respectively); 3) the ratio between the rates of charge recombination into the ground state and into the lowest triplet excited state (k_{CRS} and k_{CRT}, respectively). Orthogonal geometry between the electron donor and the acceptor, which is beneficial for efficient SOCT-ISC, can be secured by introduction of substituents in positions 1 and 7 of the BODIPY core. On the other hand, the rates of charge transfer and recombination steps can be modulated by varying the number of alkyl substituents, which affect the reduction potentials of the BODIPY core.[37] Since the energy of the CT state, and consequently ΔG_{CT} and ΔG_{CR} values, is strongly affected by polarity of the media, triplet state formation and generation of singlet oxygen by SOCT-ISC sensitizers is solvent-dependent. Generally, most of the BODIPY dyads studied so far showed higher Φ_T and Φ_Δ values in polar solvents, where the CT state is stabilized and the charge transfer process is energetically favorable. However, strong stabilization of the CT state in polar media leads to a reduced CT-S_0 energy gap that causes enhancement of the ground state recombination rate. For this reason, many dyads show reduced Φ_T and Φ_Δ values in highly polar solvents, such as acetonitrile or water.[38] In particular, for dyads bearing electron accepting groups in the BODIPY core, which stabilize the CT states, photosensitization process is more efficient in non-polar solvents (hexane, toluene).[39] BODIPY dimers showed the highest Φ_Δ values in solvents of intermediate polarity (chloroform, THF).[40] Overall, predicting Φ_T and Φ_Δ in a specific solvent is difficult and relies on test-and-trial approach. This largely limits the potential of SOCT-ISC sensitizers, for instance in advanced PDT utilizing controlled generation of cytotoxic singlet oxygen in target cells.[41]

Computational methods for predicting the SOCT-ISC efficiency in media of a given polarity could be valuable for pre-synthetic screening of potential sensitizer structures. However, accurate computations of charge transfer excited states and triplets formation by charge recombination using first-principle techniques are rather challenging and time-demanding.[42] In recent years, quantitative structure-activity relationship (QSAR)

and quantitative structure-property relationship (QSPR) modelling have emerged as a useful tool for the design of fluorophores and photoactive materials.[43] QSAR/QSPR analysis has been applied in the studies of photophysics[44] and photodynamic activity of BODIPYs,[45] proving that "big data" approach can provide a powerful platform for the design of new photosensitizers. Yet, QSAR modelling for predicting $^1O_2$ generation quantum yields is rare.[46] However, earlier we showed that QSPR may be used for the analysis of $^1O_2$ production by pterins and flavins,[47] psoralens and angelicins,[48] porphyrins and metalloporphyrins.[49] In this regard, applying QSPR for the prediction of Φ_Δ values for heavy-atom-free BODIPYs in solvents of different polarity offers a means to estimate SOCT-ISC efficiency in these systems. Hence, we started to search BODIPY structures which undergo SOCT-ISC to correlate molecular features to singlet oxygen quantum yields with the aid of machine learning methods. To develop a reliable model we analyzed a dataset which includes several classes of compounds (Figure 2): 1) parent BODIPY compounds with different substitution patterns; 2) BODIPY dyads with different nature, position, or number of donor/acceptor subunits; 3) symmetrical and asymmetrical BODIPY dimers; 4) a validation set comprised of a newly synthesized library of BODIPYs with various *meso*-methoxyphenyl groups as electron donors. Previously reported and experimentally obtained data on Φ_Δ for these molecules in solvents of different polarity – toluene (non-polar), tetrahydrofuran (moderately polar) and acetonitrile (strongly polar) were used for analysis. We performed QSPR analysis and optimized models for each solvent to ensure sufficient statistical parameters for photosensitizing ability prediction. To our knowledge, this is the first report of a QSPR model for predicting the activity of polarity-sensitive photosensitizers. The simplicity and versatility of this approach allows for virtual screening of structures with singlet oxygen quantum yields optimized for a desired range of polarities. SOCT-ISC photosensitizers can be activated by recognition of the appropriate environment, for example, cell membranes[50] and certain proteins.[51] We believe that the results obtained in this work will unlock a more target-oriented exploration of this class of photosensitizers in biomedical applications relying on a controlled ROS generation.

Results and Discussion

Compound dataset

The general dataset used here to build QSPR models includes compounds reported in experimental studies on triplets' formation in heavy-atom-free BODIPYs via SOCT-ISC. We examined all related works published before September 2020 and combined experimental values of Φ_Δ measured in various solvents using chemical trapping method and phosphorescence of singlet oxygen (Table S1).[52] Several reference compounds (structures **1**–**3**, Figure 2) were included into the dataset to guarantee the reliability of models in cases when Φ_Δ values are very low. Other compounds in the dataset (**4**–**64**) include donor-acceptor dyads and dimers, with various substitution patterns of the BODIPY core and nature of electron donating or electron accepting subunits. Values of Φ_Δ measured in toluene ($\varepsilon_r = 2.4$), tetrahydrofuran ($\varepsilon_r = 7.6$), acetonitrile ($\varepsilon_r = 37.5$) were used for analysis since these solvents have been employed to study charge transfer and $^1O_2$ generation for the highest number of compounds in the dataset.

To ensure the reliability of the QSPR models developed in this work for the development of practical sensitizers, several new BODIPYs derivatives were also synthesized and investigated by us. BODIPY dyads bearing electron-donating aryl groups such as aminophenyl[53] or methoxyphenyl[54] groups were previously shown to undergo charge transfer and generate $^1O_2$ both in polar and non-polar solvents. Such dyads are potentially interesting for application in PDT as alternatives

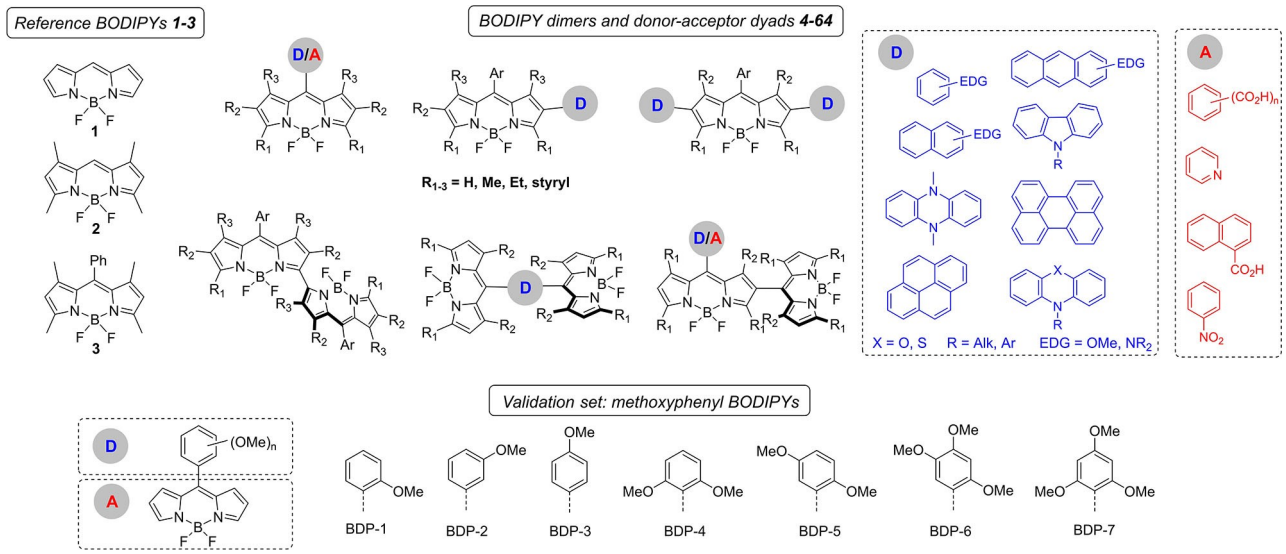

Figure 2. General structures of BODIPY dimers, donor-acceptor (D–A) dyads and reference compounds investigated in this work.

to cyclic tetrapyrroles (porphyrins, chlorins, and bacteriochlorins), which often require tedious synthesis. Moreover, these common PDT agents are rather large molecules which must be injected and as a result retain in the bloodstream for a long time, leading to photodermatosis among other side-effects.[55] Thus there is a growing interest in non-porphyrin sensitizers, based on small photoactive molecules with higher absorption and excretion rates. With this in mind, we prepared compounds **BDP 1–7** (Figure 2) bearing different methoxyphenyl groups in the *meso*-position of the BODIPY core.

These seven BODIPY compounds (**BDP 1–7**) were synthesized in the stepwise manner as shown in Scheme S1. In the first step, a standard dipyrromethane synthesis was performed of each aldehyde and pyrrole in the presence of trifluoroacetic acid.[56] Dipyrromethanes (**DPM 1–7**) were prepared and purified by following earlier reports.[57] The second step involved the well-known one-pot two-step oxidation-deprotonation-complexation reaction for the synthesis of the desired BODIPY compounds.[58] More specifically, **DPM 1–7** derivatives were oxidized to dipyrromethenes using DDQ, which were subsequently treated with triethylamine (TEA) and $BF_3 \cdot OEt_2$. All BODIPY products were isolated by silica gel column chromatography and final purification was performed through recrystallization in a MeOH/H_2O mixture in good to high yields. Monomethoxy substituted BODIPY derivatives (**BDP-1**, **BDP-2** and **BDP-3**) were prepared as reference compounds following previously published procedures.[59]

All BODIPY compounds (**BDP 1–7**) were fully characterized through ^{1}H, ^{13}C, ^{19}F, ^{11}B NMR spectroscopy and HRMS (see Supporting Information). The spectroscopic data of the three reference derivatives (**BDP 1–3**) were identical with those reported previously.[60] The successful formation of the final BODIPY products was confirmed by the appearance of the corresponding molecular ion peaks in their HRMS spectra. All ^{1}H and ^{13}C NMR spectroscopic data are in agreement with the proposed structures of the four newly reported compounds (**BDP 4–7**). Regarding the ^{19}F NMR spectra of these four final products, the typical doublet of doublets signal[61] was observed for both **BDP-4** and **BDP-7** at approximately −145 ppm (Figures S3 and S18). On the other hand, the remaining two BODIPY compounds (**BDP-5** and **BDP-6**) presented two multiplets at ~−144 and −146 ppm (Figures S8 and S13), which can be attributed to the single *ortho*-methoxy substituent forming an inequivalent environment for the F-atoms.[58,62] Finally, all compounds showed a typical triplet peak (~0.3 ppm) in their ^{11}B NMR spectra, as expected.[60]

Furthermore, the structures of all four newly synthesized donor-acceptor dyads (**BDP 4–7**) were analyzed through single crystal X-ray crystallography. Compounds **BDP-4**, **BDP-5** and **BDP-7** were crystallized from dichloromethane at 4 °C, while **BDP-6** was crystallized from acetonitrile by slow evaporation; details of crystallographic refinement for these representative species are provided in the Supporting Information (Table S2). The crystal structures for **BDP 1–3** have been previously reported.[58]

Two crystallographically inequivalent molecules of **BDP-4** (Figure 3a), co-crystallize with a dichloromethane solvate; nearly

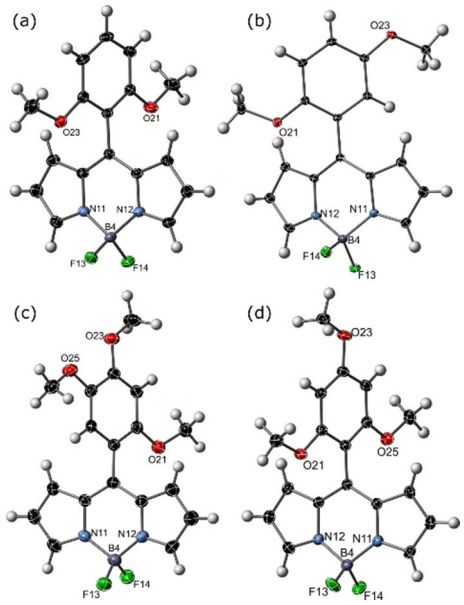

Figure 3. The individual molecular units of compound (a) **BDP-4**·$^{1}/_{2}$**DCM** (solvate omitted), (b) **BDP-5**, (c) **BDP-6** and (d) **BDP-7**; thermal ellipsoids are shown at the 50% probability level, H-atoms are represented as spheres of fixed radius.

perpendicular aryl components (81.54(5)° and 82.129(5)° for the two inequivalent molecules) are as expected for the steric bulk of the bis(*o*-methoxy) units. This dihedral angle has been shown to be critically important for efficient SOCT-ISC process.[28] In **BDP-5** (Figure 3b), the aryl ring is inclined at 50.163(13)° to the plane of the BODIPY core. Molecules were found to adopt a dimeric arrangement in the solid state, mediated by weak aryl C–H···F of 3.292 Å (C20-F13) shown in Figure 4; the formation of dimers can modulate fluorescence behavior.[38] The related compound **BDP-6** (Figure 3c) shows an aryl-BODIPY angle of 66.14(6)°; although crystallizing in the chiral space-group $P2_12_12_1$, this compound exhibits no permanent chirality due to equivalency of the two rotamers. Close contact C–H···O interactions (3.28 Å C···O) form one-dimensional chains, shown in Figure S56. The aryl group in compound **BDP-7** (Figure 3d) is

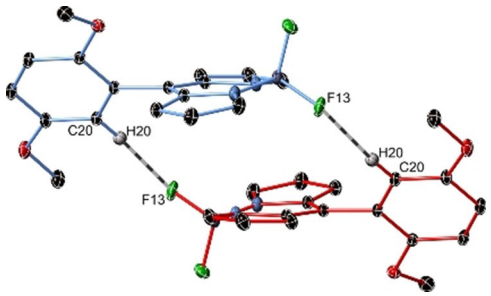

Figure 4. The dimeric C–H···F of 3.292 Å (C20-F13) interaction observed within the crystal structure of **BDP-5**; the symmetry-equivalent red and blue molecules are related by an inversion center. Thermal ellipsoids are shown at the 50% probability level, H-atoms not involved in this interaction have been omitted.

rotated 72.04(3)° to the BODIPY component. Molecules of **BDP-7** exhibit close π-π stacks of pyrrole moieties, with a plane separation of 3.23 Å; directional C–H···F interactions are observed, at 3.316 Å C3···F14. Unit cell packing diagrams are shown in Figures S52-S55.

Increasing the number of methoxy substituents leads to reduced oxidation potentials[63] of the aryl subunit in these dyads and, as a consequence, promotes intramolecular charge transfer. The absorption and fluorescence emission parameters of the prepared dyads in different solvents are given in Table 1. Changing the solvent polarity significantly impacts the emission properties. For instance, dyad **BDP-7** in toluene exhibits fluorescence emission with quantum yield of 0.661, associated with $S_0 \rightarrow S_1$ transition. On the other hand, in polar acetonitrile, the fluorescence is strongly quenched leading to quantum yield of < 0.001. This behavior is observed for other dyads and is the signature of intramolecular charge transfer.

Table 1. Spectroscopic data and singlet oxygen quantum yield for compounds **BDP 1–7**.

Compound	Solvent	λ_{abs} [nm]	λ_{em} [nm][a]	Φ_{em}[b]	Φ_{Δ}[c]
BDP-1	toluene	506	523	0.381	0.009
	THF	503	519	0.238	0.011
	ACN	500	516	0.104	0.041
BDP-2	toluene	504	522	0.057	0.005
	THF	501	518	0.029	0.009
	ACN	497	515	0.014	0.004
BDP-3	toluene	501	518	0.093	0.005
	THF	498	513	0.044	0.021
	ACN	494	512	0.017	0.007
BDP-4	toluene	507	524	0.988	0.022
	THF	504	521	0.904	0.036
	ACN	500	516	0.062	0.059
BDP-5	toluene	506	525	0.028	0.251
	THF	502	513	0.001	0.061
	ACN	499	505	0.001	0.003
BDP-6	toluene	507	521	0.002	0.158
	THF	503	517	0.0006	0.011
	ACN	499	505	0.0013	0.004
BDP-7	toluene	507	522	0.661	0.110
	THF	504	519	0.316	0.363
	ACN	500	518	0.0003	0.193

[a] The fluorescence was excited at the vibrational shoulder of the BODIPY absorption. Excitation wavelengths: 470–490 nm for **BDP 1–7**. [b] Fluorescence quantum yields were measured using Rhodamine 6G ($\Phi_{em} = 0.95$ in ethanol). [c] Quantum yields were measured using 1,9-dimethylanthracene as a $^1O_2$ trap and 2,6-diiodo-8-phenylBODIPY as a reference photosensitizer ($\Phi_{\Delta} = 0.85$ in toluene).

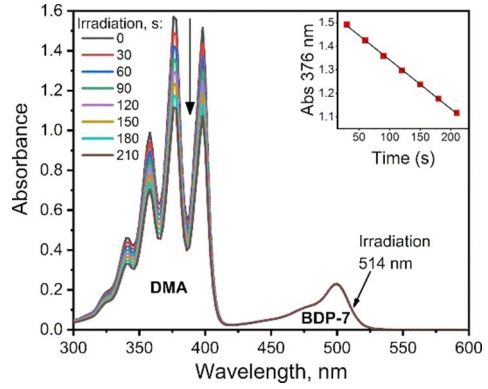

Figure 5. Photosensitized oxidation of 1,9-dimethylanthracene in the presence of **BDP-7** in air saturated acetonitrile solution irradiated with 514 nm laser (12 mW cm^{-2}). Inset: change of absorbance at 376 nm with time.

Singlet oxygen sensitization by **BDP 1–7** was evaluated using chemical trapping method with 1,9-dimethylanthracene (DMA) in toluene, THF and acetonitrile to ensure the consistency of the resulting dataset. Upon irradiation of air-saturated solutions containing each BODIPY compound at 514 nm, DMA selectively reacts with singlet oxygen forming corresponding endoperoxide. The BODIPY absorption shows no change during irradiation, while the DMA absorption decreases. The change of DMA absorbance with time is linear (Figure 5, inset), allowing to obtain Φ_{Δ} value from comparison with reference sensitizers (2,6-diiodo-8-phenylBODIPY). As shown in Table 1, the resulting Φ_{Δ} values vary depending on the solvent and the structure of electron-donating aryl group. The most efficient $^1O_2$ sensitization was observed for **BDP-7** in THF ($\Phi_{\Delta} = 0.363$) that correlates with efficient charge transfer process in this solvent. This correlates with efficient charge transfer process in this solvent. Dyads **BDP-5** and **BDP-6** showed Φ_{Δ} values of 0.25 and 0.16, respectively, in toluene and low $^1O_2$ sensitization ability in the remaining two solvents. Dyads **BDP 1–4** showed very modest Φ_{Δ} values in all solvents.

The resulting general dataset combining previously studied compounds and dyads **BDP 1–7** was divided into the training set (80% of the compounds) and the test set (20%) using random number generation. The activity of compounds, expressed as Log Φ_{Δ}, was used as a dependent variable during QSPR. Four newly synthesized compounds were put into the so-called external set. The toluene dataset was divided into the training set (35 compounds), the test set (9 compounds), and the external set (4 compounds). The acetonitrile dataset consisted of 33 compounds in the training set, 8 compounds in the test set, and 4 compounds in the external set. The tetrahydrofuran dataset was divided into the training set (30 compounds), the test set (7 compounds), and the external set (4 compounds).

Molecular descriptors and model search

The QSPR approach employed here is based on the assumption that the efficiency of triplet state formation and singlet oxygen generation by the photosensitizer molecule depends on its structure and attempts to formulate mathematical relationship between calculated features of the structure (known as molecular descriptors) and its singlet oxygen quantum yield value.

The contribution of each descriptor to the model was estimated using Equation (1):

$$\alpha(x_1) = \frac{R(x_1)}{R(x_1) + \ldots + R(x_n)} \times 100 \quad (1)$$

where $\alpha(x_1)$ is the relative contribution of the descriptor x_1 to the model with several descriptors, $R(x_n)$ is the correlation coefficient of the nth descriptor towards $\log\Phi_\Delta$.

Three different machine learning methods were used for models search: support vector regression (SVR), multiple linear regression (MLR), and random forest regression (RFR). QSPR models were selected on the basis of statistical parameters, such as R^2 (determination coefficient of the training set) and R^2_{test} (predictive R^2 for the test set of compounds). Predicting ability of the obtained models was evaluated for a test set of compounds, R^2_{test} parameter was used for model validation and comparison. A QSPR model is considered to be predictive if the following conditions are met: $R^2 > 0.6$, $q^2 > 0.6$, and $R^2_{test} > 0.5$.[64] Among the different machine learning methods applied, only MLR models showed satisfactory statistical parameters and thus was used in further analysis. Statistical parameters for toluene, acetonitrile, and tetrahydrofuran (THF) models are summarized in Table 2 and a detailed description of the statistical parameter calculation is given in the Experimental Section.

Comparison of the root-mean-square (RMS) error for Models 1–3 and those previously reported for other types of phosensitizers demonstrates good predictive ability of models developed in this work. For instance, models reported for pteridines, furicoumarins and porphyrins showed RMS Φ_Δ errors of 5.7–24.8%.[47–49] It should be noted that in these studies the values of singlet oxygen quantum yields were measured in identical conditions for all compounds in the dataset. In the current study we used larger datasets (>40 BODIPYs for each model) with Φ_Δ values measured using different $^1O_2$ traps and reference photosensitizers. As one can see, the resulting RMS error shows comparable values: 8.2–18.3%. Therefore, we conclude that possible inconsistencies in Φ_Δ values which might be due to the difference in experimental conditions did not significantly influence predictive ability of the developed models.

The resulting MLR equations have the following form [Eq. (2)]:

$$y = c + a_1{}^*x_1 + \ \ldots\ + a_n{}^*x_n \tag{2}$$

where y is the dependent variable $\log\Phi_\Delta$, c is a regression constant, a_1 and a_n are regression coefficients, x_1 and x_n are independent variables. The equation obtained for toluene model includes seven descriptors, whereas in the case of acetonitrile and THF the MLR models include six descriptors.

Table 2. Statistical parameters for multiple linear regression (MLR) models predicting the quantum yield of singlet oxygen generation by BODIPYs in toluene, acetonitrile, and tetrahydrofuran (THF).

Parameter	Toluene	Acetonitrile	THF
R	0.882	0.890	0.906
R^2	0.778	0.792	0.820
$R^2_{adjusted}$	0.739	0.744	0.773
SEE	0.282	0.319	0.325
RMSE	0.240	0.283	0.285
q^2	0.686	0.693	0.620
SDEP	0.306	0.344	0.414
R^2_{test}	0.800	0.823	0.879
$R^2_{external}$	0.635	0.722	0.584
RMS Φ_Δ error (Φ_Δ in %)	8.2	18.3	12.0

[a] R^2, $R^2_{adjusted}$, SEE, and RMSE relate to the training set. SDEP and q^2 relate to leave-one-out cross validation of the training set. R^2_{test} relates to the test set whereas $R^2_{external}$ describes the external set of compounds. Root-mean-square (RMS) Φ_Δ error relates to the whole dataset, which combines the training set, the test set and the external set.

Model 1 (toluene)

The model for predicting activity of photosensitizers in toluene was developed using data for the highest number of compounds - 48 BODIPYs from the general dataset. The MLR equation for the model includes seven descriptors [Eq. (3), Model 1]:

$$\begin{aligned}\mathrm{Log}\,\Phi_\Delta &= 1.5424(+/-0.2559) - 1.4403\\ &(+/-0.3471)\,\mathrm{DLS\_04} - 0.4718\\ &(+/-0.1024)\,\mathrm{Mor24m} + 0.536(+/-0.0995)\\ &\mathrm{ATSC6e} - 0.1479(+/-0.0529)\,\mathrm{F09[N-N]} - 4.1433\\ &(+/-0.9008)\,\mathrm{R6s} + 0.8055(+/-0.1948)\,\mathrm{S1\_fosc} + 0.012\\ &(+/-0.0031)\,\mathrm{TPSA(Tot)}\end{aligned} \tag{3}$$

Among the seven descriptors involved in Model 1, the two most influential are *R6s+* (relative contribution α is equal to −29.9% and *ATSC6e* ($\alpha = 25.8\%$). *R6s+* is a GETAWAY (GEometry, Topology, and Atom-Weights AssemblY) descriptor, an R maximum autocorrelation of lag 6 (a path length, or topological distance between atoms) weighted by intrinsic state (I-state).[65] The I-state of the *i*-th atom is calculated by Equation (4):

$$I_i = \frac{(2/L_i)^2\delta_i^v + 1}{\delta_i} \tag{4}$$

where L_i is the principal quantum number (2 for C, N, O, F, 3 for S), δ_i^v is the number of valence electrons, and δ_i is the number of sigma electrons of the *i*-th atom.

The I-state of an atom can be considered as the ratio of n and lone pair electrons to the count of the σ bonds. Thus, the I-state evaluates the possible partitioning of non-sigma electrons influence along the paths starting from the regarded atom; with less partitioning of the electron influence, the more available are the valence electrons for intermolecular interactions.[66] I-state is higher for electron withdrawing groups: for example, for $-CH_2-$, –NH–, and –O– groups I-state is equal to 1.5 and 2.5, and 3.5, respectively. Hence, the descriptor captures the through-bond effects of the rest of the molecule on the electron density of atoms that can be involved in the charge transfer process.

ATSC6e descriptor is defined as the centered Broto-Moreau autocorrelation-lag 6/weighted by the Sanderson electronegativities.[67] In general, the presence of electron withdrawing groups, such as F, Cl, and –COOH decrease *ATSC6e*

Table 3. Values of the descriptors used in Model 1 (toluene).

Compound	DLS_04	Mor24m	ATSC6e	F09[N–N]	R6s+	S1_fosc	TPSA(Tot)
2	0.8	0.362	0.509	0	0.254	0.5982	8.81
3	0.8	0.466	0.489	0	0.187	0.5741	8.81
21	0.8	0.429	0.493	0	0.131	0.3868	8.81
25	0.4	0.424	0.48	0	0.184	0.0147	46.1
26 [test a]	0.6	0.507	0.788	0	0.142	0.1667	46.1
27	0.6	0.678	0.486	0	0.0646	0.3985	8.81
28	0.4	0.737	0.505	0	0.057	0.3966	8.81
29 [test a]	0.4	0.568	0.552	0	0.0996	0.8794	8.81
30	0.6	−0.00895	0.909	0	0.0869	0.0265	6.48
31	0.5	−0.12	0.505	0	0.101	0.3872	8.81
32	0.4	0.171	0.112	0	0.0967	0.3404	8.81
33	0.5	0.158	0.516	0	0.0729	0.7044	8.81
34	0.6	−0.761	0.827	0	0.065	0.1024	35
35	0.6	0.0461	0.566	0	0.0532	0.0313	42
36 [test a]	0.8	−0.581	0.661	0	0.0635	0.2843	19
37	0.5	0.531	0.297	1	0.186	0.117	26.9
38	0.7	−0.6	0.995	0	0.0625	0.4448	19
39	0.4	0.633	0.828	0	0.0702	0.0787	26.9
40	0.5	0.665	0.807	2	0.0695	0.0095	74.5
41	0.5	0.482	0.56	0	0.103	0.4147	26.9
42	0.7	0.473	0.709	1	0.132	0.6649	26.9
43 [test a]	0.6	0.707	0.251	0	0.109	0.0835	18.7
44	0.8	0.546	0.597	0	0.0815	0.0011	18.7
46	0.7	1.06	1.15	0	0.0756	0.0843	27.5
47	0.6	0.89	0.118	0	0.0881	0.2915	13.7
48	0.8	0.878	0.518	0	0.0653	0.4534	13.7
49	0.4	0.868	0.597	1	0.112	0.701	13.7
50	0.4	0.734	0.608	1	0.122	0.8677	13.7
51	0.4	0.784	0.752	2	0.0695	0.2483	18.7
52 [test a]	0.4	0.35	0.782	2	0.0641	0.2885	15.3
53	0.4	0.267	0.752	2	0.0845	0.0589	75.2
55	0.6	1.56	0.644	2	0.0836	0.9164	18.7
56	0.6	1.69	0.847	2	0.0342	0.8045	36.7
57	0.7	0.739	2.15	0	0.0826	0.3811	17.6
58 [test a]	0.6	0.728	2.14	0	0.0603	0.3454	17.6
59	0.6	0.768	2.14	0	0.0609	0.4058	17.6
60 [test a]	0.8	0.91	2.03	0	0.0644	0.4318	17.6
61	0.7	0.976	1.47	0	0.0796	0.698	17.6
62 [test a]	0.7	0.972	1.64	0	0.09	0.1135	63.4
63	0.7	0.89	1.27	0	0.0911	0.0819	43.6
64 [test a]	0.7	0.709	1.08	4	0.0699	0.0011	17.6
BDP-1 [ext b]	0.6	0.346	0.356	0	0.199	0.3557	18
BDP-2 [ext b]	0.6	0.576	0.308	0	0.244	0.276	18
BDP-3	0.6	0.55	0.253	0	0.245	0.2689	18
BDP-4	0.6	0.609	0.666	0	0.093	0.3084	27.3
BDP-5 [ext b]	0.6	0.362	0.51	0	0.113	0.3352	27.3
BDP-6	0.6	0.483	0.727	0	0.0884	0.3005	36.5
BDP-7 [ext b]	0.6	0.665	0.987	0	0.093	0.0017	36.5

[a] [test] designates compounds belonging to the test set. [b] [ext] relates to the molecules of the external set

values, whereas electron-donating alkyl and alkoxy groups increase the value of *ATSC6e*. In the obtained model, the descriptor is inversely correlated with Log Φ_Δ. Symmetrical BODIPY dimer **57** has the highest value of *ATSC6e* equal to 2.15, whereas BODIPY-perylene dyad **32** possesses the lowest value equal to 0.112.

TPSA(Tot) ($\alpha = 13.3\,\%$) is a total topological polar surface area, corresponding to the polar surface area derived from polar contributions of N, O and S atoms and is related to the hydrogen bonding ability of the compound.[68] *TPSA(Tot)* is directly proportional to Log Φ_Δ of the studied BODIPYs. Polar substituents increase the value of this descriptor whereas non-polar substituents decrease the value of *TPSA(Tot)*. Compound **53**, containing two phenothiazine groups, has the highest value of the descriptor equal to 75.4 whereas for BODIPY-anthracene dyad **30** *TPSA(Tot)* is equal to 6.48.

DLS_04 ($\alpha = -12.2\,\%$) descriptor is related to drug-like score indices, similar to drug-like filters (7 rules) described by Chen et al.[69] This descriptor considers several characteristics: partition coefficient, the amount of hydrogen bond donors, number of hydrogen bond acceptors, molecular weight, ratio of the number of C(sp^3) atoms over the total number of non-halogen heavy atoms, the ratio of H atoms to non-halogen heavy atoms, and the ratio of the molecular unsaturation over the total number of non-halogen heavy atoms. In the studied dataset the values of *DLS 04* fall within the range of 0.4-0.8. Lower values of the *DLS 04* index indicates that the compound is not good for drug-like purposes and is in accordance with the aim

of our study, since efficient photosensitization of $^1O_2$ leads to high cytotoxicity.

Mor24 m (α=-10.3%) is a 3D-MoRSE-signal 24/weighted by atomic masses,[70] which was found to be useful in predicting the toxicity of drugs.[71] 3D-MoRSE descriptors (Molecular Representation of Structures based on Electronic diffraction) are often regarded as "black box", although the mathematical formula for the MoRSE descriptors is rather simple [Eq. (5)]:

$$Mor = \sum_{i=2}^{N}\sum_{j=1}^{i-1} A_i A_j \frac{sinsr_{ij}}{sr_{ij}} \quad (5)$$

where *s* is the scattering parameter, $r_{ij}$is the euclidean distance between *i*-th and *j*-th atoms, *N* is the total number of atoms, A_i and A_j are different atomic properties used as weights,[70] in the case of *Mor03s* and *Mor18s*-I-state.

The presence of *Mor24m* in the Model 1 reveals a significant dependence of Φ_Δ on the size of a molecule. The descriptor is inversely correlated with Log Φ_Δ. Compound **56** has the highest molecular weight among the studied compounds and its *Mor24m* value is the highest in the dataset. BODIPY-phenothiazine **34**, on the contrary, has the lowest value of *Mor24m* descriptor.

Two descriptors have minor contributions to Model 1 ($\alpha <$ 10%). *F09[N-N]* ($\alpha = -5.3$%) is a frequency of topological distance N–N equal to 9. The descriptor is inversely correlated with Log Φ_Δ. Phenylene-separated BODIPY dimer **64** possesses the highest value of *F09[N-N]* equal to 4, while for most compounds in the analyzed dataset it is equal to 0.

S1_fosc (α=3.3%) is an oscillator strength of the $S_0 \rightarrow S_1$ transition. This parameter correlates with quantum yield of the fluorescence from the lowest singlet excited state (Φ_{fl}). As was shown in previous works, fluorescence from the S1 state competes with the charge transfer process in BODIPY dyads and molecules with high Φ_{fl} values show inefficient SOCT-ISC.[38]

Model 2 (acetonitrile)

A total of 45 compounds from the general dataset are involved in this model. The MLR equation includes six descriptors [Eq. (6), Model 2]:

$$\begin{aligned}&\text{Log}\,\Phi_\Delta = 3.2341(+/-0.4672) + 0.0872(+/-0.0456)\\ &\text{Mor18 s} + 0.0342(+/-0.0079)\,\text{Mor03s} + 0.1974\\ &(+/-0.0648)\,\text{F06[C–B]} - 0.0304\\ &(+/-0.0138)\,\text{RDF065m} - 5.6375\\ &(+/-0.9105)\,\text{MATS6s} - 76.4904\\ &(+/-12.1236)\,\text{R8u+}\end{aligned} \quad (6)$$

Table 2 demonstrates that Model 2 performed as the most internally stable one (q^2=0.693) and the best in predicting the properties for the external set of compounds ($R^2_{external}$=0.722). The predicted versus experimental values of Log Φ_Δ are presented in Figure 6b and in Table S4 (Supporting Information).

MATS6s (α=-28.9) is a 2D Moran autocorrelation of lag 6 / weighted by the I-state as described by Todeschini and Consonni.[72] The correlation coefficient between *MATS6s* and Log Φ_Δ is high and equal to 0.569. The negative regression coefficient before *MATS6s* in Model 2 equation indicates that increased autocorrelation of six-membered structural graphs is unfavorable for high quantum yield Φ_Δ values. The importance of topological distance 6 is apparently due to the presence of aryl fragments (electron donors/acceptors) in most of the BODIPYs structures included in the dataset. The Moran coefficients usually fall within the [−1,+1] interval, whereas in the analyzed dataset most compounds possess *MATS6 s* in the [−0.2,−0.1] interval. BODIPY **16** bearing dimethylaminophenyl group in the *meso* position has the highest value of *MATS6s* equal to 0.0676 whereas 1,3,5,7-tetramethylBODIPY **2** has the

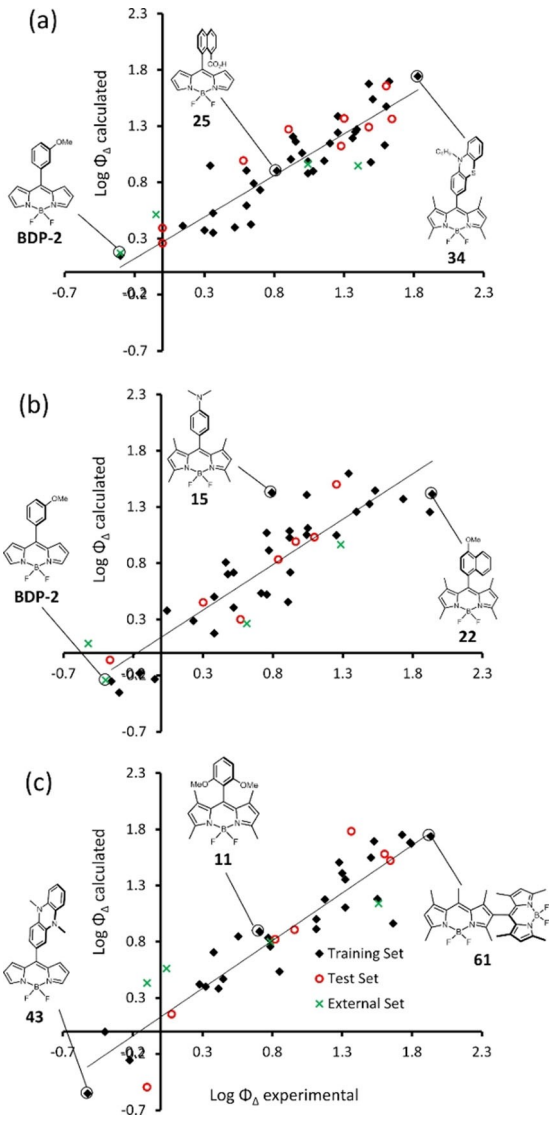

Figure 6. Experimental vs. predicted Log Φ_Δ values for BODIPYs in toluene (a), acetonitrile (b), and THF (c). Three sample molecules from the general dataset are highlighted for each model.

minimal *MATS6s* value in the acetonitrile dataset (−0.214) (Table 4). The reason for such diverse values of the descriptor for these two compounds is the I-state values of certain atom types: for example, **16** contains tertiary amine N atom having rather low I-state value of 1.0882, whereas **2** is a relatively small molecule and contains four methyl groups: I-state equals to 2.0 which is high compared to other types of compounds.[73]

R8u+ (α = −22.2) is an R maximal autocorrelation of lag 8 / unweighted. The descriptor is inversely correlated with Log Φ_Δ. *meso*-Pyridyl-substituted BODIPY **8** has the highest value of *R8u+* equal to 0.0392, whereas dihydrophenazine-separated BODIPY dimer **45** has the lowest value of the descriptor (0.013). Here, the importance of GETAWAY descriptors is demonstrated once again: Model 1 included *R6s+* descriptor with high relative contribution. Apparently, as in the case of *R6s+*, the maximum leverage influences the activity and *R8u+* value decreases with increasing the number of atoms.

Model 2 involves two 3D-MoRSE descriptors weighted by I-state: *Mor18s* (α = 20.3) and *Mor03s* (α = 10.3). Both descriptors are directly proportional to Log Φ_Δ. Dyad **54**, containing N-phenylcarbazol as an electron donating subunit, has the lowest value of *Mor18s* (−8.79) whereas **BPD-1** possesses the highest value (−1.7) in the analyzed dataset. All compounds in the validation set (**BDP 1**–**7**) possess high values of *Mor18s*, which is apparently caused by the presence of methoxy groups in their structures (–O– atom has a high I-state equal to 3.5). In comparison, dyad **54** is characterized by the presence of multiple aromatic rings with carbon atoms having I-state values of 1.5–2.0.

Regarding *Mor03s*, dyad **16** showed the highest value of −15.8, whereas the lowest value (−63) was obtained for BODIPY-anthracene dyad **30**, which is a relatively small molecule and contains four six-membered rings with carbon atoms having low I-state values. Although the structure of **16** is

Table 4. Values of the descriptors used in Model 2 (acetonitrile).

Compound	R8u+	MATS6s	F06[C–B]	Mor03s	RDF065m	Mor18s
2[test [a]]	0.0349	−0.214	0	−19.5	1.16	−2.72
3	0.0351	−0.142	2	−22.7	7.86	−5.08
4[test [a]]	0.0328	−0.124	2	−22.6	10.2	−3.28
5	0.0331	−0.121	2	−35.9	11.8	−5.11
6[test [a]]	0.0311	−0.0955	2	−28.4	13.3	−5.45
7[test [a]]	0.0304	0.0526	2	−29.9	15.7	−3.34
8	0.0392	−0.126	2	−20.9	6.67	−2.85
9[test [a]]	0.0362	−0.142	2	−22	7.88	−4.24
10	0.0243	−0.127	2	−21.7	10.5	−4.32
11[test [a]]	0.017	−0.147	2	−21.9	14.5	−5.36
12[test [a]]	0.0211	−0.109	2	−24.1	11.9	−4.74
13	0.0236	−0.13	2	−29.7	14.1	−4.55
14	0.0215	−0.197	2	−24.8	16.2	−4.91
15	0.019	−0.117	2	−18.3	11.3	−5.03
16	0.0185	0.0676	2	−15.8	10.1	−5.2
17	0.0179	−0.109	2	−29.1	9.85	−5.67
18	0.0161	0.0515	2	−25.2	9.11	−5.14
19	0.0182	−0.111	2	−30.8	9.35	−5.72
20	0.0163	0.0555	2	−27.7	8.6	−5.23
21	0.0239	−0.134	3	−23.9	11.6	−5.89
22	0.0177	−0.126	3	−25.3	14.5	−5.3
23	0.0283	−0.141	2	−26.5	12.8	−5.84
24	0.0195	0.0262	2	−23.2	11.9	−5.25
25	0.0235	−0.051	3	−28.1	9.81	−3.89
26[test [a]]	0.0249	−0.143	3	−26.2	10.8	−5.87
27	0.0158	−0.133	4	−27.1	11.1	−8.06
28	0.0146	−0.126	4	−27.8	22.3	−8.42
29	0.0166	−0.115	6	−32.1	21.8	−7.21
30	0.0163	−0.0437	4	−63	23.5	−5.81
31	0.0183	−0.131	3	−24.5	18.8	−3.55
33	0.0131	−0.124	3	−27.2	27.6	−8.05
45	0.013	0.0582	4	−52.7	25.9	−3.94
47	0.0159	0.0351	2	−33.3	9.88	−3.16
48	0.0131	−0.108	2	−28.2	13.5	−5.63
54	0.0145	−0.121	2	−28.5	21.5	−8.79
61	0.0173	−0.159	5	−51.1	13.3	−4.43
63	0.0157	−0.127	7	−49.4	22.8	−7.31
64	0.0156	−0.128	4	−43.9	29.4	−5.22
BDP-1[ext [b]]	0.0252	0.0179	2	−26.5	9.32	−1.7
BDP-2[ext [b]]	0.0292	0.0549	2	−27.1	7.78	−1.88
BDP-3	0.0294	0.0562	2	−25.2	7.5	−1.76
BDP-4	0.019	−0.0496	2	−28.3	12.1	−2.36
BDP-5[ext [b]]	0.0237	0.0429	2	−28.4	10.5	−2.29
BDP-6	0.0258	2.30E−04	2	−32.2	12.9	−2.04
BDP-7[ext [b]]	0.0236	−0.132	2	−29.8	13.5	−1.99

[a] [test] designates compounds belonging to the test set. [b] [ext] relates to the molecules of the external set.

rather compact, it contains nitrogen atom and four methyl groups, each having high I-state of 2.

F06[C-B] ($\alpha = 12.2\%$) is a frequency of topological distance C–B equal to 6. The descriptor is directly correlated with Log Φ_Δ. BODIPY dimer **63** possesses the highest value (7) of *F06[C-B]*, whereas compound **2** shows the lowest value equal to 0. It can be concluded that the descriptor allows to distinguish some cases when a more complicated molecular topology is favorable for efficient ISC.

RDF065m is a descriptor with minor relative contribution to Model 2 ($\alpha = -6.3\%$). It is a radial distribution function at 6.5 Å/ weighted by atomic masses. In simple terms, *RDF065m* can be regarded as a contribution of the atomic masses within the 6.5 Å radius of the molecule center. The lowest value is equal to 1.16 for compound **2**, having the lowest molecular weight in the dataset, whereas **64** is one of the largest molecules in the analyzed dataset and possesses the highest value of *RDF065m* equal to 29.4.

Model 3 (THF)

The total number of compounds involved in this model is 41. The MLR equation includes six descriptors [Eq. (7), Model 3]:

This model possessed the best statistical parameters in terms of correlation coefficient of the training set ($R = 0.906$) and predicting ability towards the test set ($R^2_{test} = 0.879$). The predicted versus experimental values of Log Φ_Δ are presented on Figure 6c and in Table S5 (Supporting Information).

N-071 ($\alpha = 29.3$) has a major contribution to the model. The descriptor counts the number of tertiary nitrogen atoms attached to aromatic carbons. Only nine such compounds (**15**–**20** and **44**–**46**) are present in the dataset. The lone pair of nitrogen contributes to the electron π-system and causes charge transfer process in these dyads, ultimately leading to SOCT-ISC.[53]

CATS3D_07_PL ($\alpha = 24.0$) is a Chemically Advanced Template Search (CATS) 3D descriptor, namely the frequency of polar and non-polar groups separated by 7–8 Å. The descriptor is directly proportional to Log Φ_Δ. For most of the compounds in the dataset *CATS3D_07_PL* is equal to 0. Dimer **64** has the highest value of the descriptor equal to 3 (Table 5). Seemingly, in the case of **64** the descriptor evaluates the frequency of pairs between nitrogen atoms and methyl groups.

F03[C-N] ($\alpha = -20.2$) is a frequency of C and N atoms at a topological distance of 3. The higher the value of the descriptor, the lower the quantum yield of $^1O_2$ generation. Parent BODIPY **1** has the lowest value of *F03[C-N]* equal to 4, whereas dimer **46** possesses the highest value of 24.

S2_nm ($\alpha = -19.7$) is the energy of the second lowest singlet excited state in nanometers. The descriptor is inversely correlated with the quantum yield logarithm. Therefore, high S_2 state energy (in eV) is favorable for high quantum yield. Compound **61** has the highest value of the descriptor equal to 370 nm, whereas **2** possesses the lowest value equal to 237 nm. BODIPYs are known to emit from the S_2 state;[74] apparently this process can compete with intersystem crossing and triplet state formation which leads to lower Φ_Δ values.

Two descriptors have a minor contribution to Model 3: *LLS_01* ($\alpha = -4.3$) and *ATSC7i* ($\alpha = 2.1$). *LLS_01* is a lead-like score derived from the rules proposed by Congreve et al.[75] It takes into account the number of H-bond donors, number of H-bond acceptors, molecular weight, lipophilicity (log P), rotatable bond number, and polar surface area. Compounds with high *LLS_01* (**BDP-1**, **BDP-2**, and **BDP-3**) show low photosensitization efficiency, in contrast to compounds with the low descriptor value like dyad **12**, which have Φ_Δ value of 0.46 in this solvent. *ATSC7i* is a centered Broto-Moreau autocorrelation of lag 7 / weighted by ionization potential. Dimer **46** has the highest *ATSC7i* equal to 4.65 whereas the *ATSC7i* of **1** is equal to 0.0498. The descriptor is directly proportional to Log Φ_Δ. Apparently, structures with high numbers of nitrogen, oxygen, and halogen atoms have high *ATSC7i* values due to higher ionization potential of such atoms.

Conclusion

Herein we present for the first time a QSPR approach for predicting the efficiency of singlet oxygen generation by heavy-atom-free BODIPYs in solvents of different polarity. Models developed using multiple linear regression (MLR) are quantitatively accurate and show good statistical parameters ($R = 0.88$–0.91 and $q^2 = 0.62$–0.69) for Φ_Δ prediction in all solvents, outperforming other machine learning methods, such as support vector regression (SVR) and random forest regression (RFR). The models were built using a combination of alvaDesc and quantum-chemical descriptors which can be obtained via simple calculations in a highly efficient manner.

For acetonitrile and toluene models topological descriptors, such as GETAWAY which tries to match 3D-molecular geometry provided by the molecular influence matrix and atom relatedness by molecular topology, Broto-Moreau, and Moran autocorrelations were the most influential (see Results and Discussion section for details). Model 3 (THF) is dominated by N-071 (constitutional descriptor) and CATS3D_07_PL (frequency of polar and non-polar groups separated by 7–8 Å). Therefore, Models 1 and 2 are slightly different from Model 3. In terms of atomic properties, intrinsic state of the atoms played a significant role for the used descriptors and strongly correlates with the photosensitizing ability of BODIPYs. Interestingly, quantum-chemical descriptors such as frontier molecular orbital energies, HOMO-LUMO gap, energies of the lowest singlet and triplet excited states, as well as the singlet-triplet energy gap (ΔE_{ST}) were found to have a low relative contribution to the models. BODIPY structures involved in our study possess rather large substituents, which strongly influence mutual orientation between the donor and acceptor subunits (one of the key factors determining SOCT-ISC efficiency) and probably due to this reason topological and 3D descriptors play more significant role than QC descriptors.

However, when smaller subsets of BODIPY structures are considered, strong correlations between quantum-chemical

Table 5. Values of the descriptors used in Model 3 (THF).

Compound	ATSC7i	LLS_01	N-071	F03[C–N]	CATS3D_07_PL	S2_nm
1	0.0498	1	0	4	0	277.52
2[test [a]]	0.443	1	0	6	0	237.18
3	1.7	0.667	0	8	0	293.08
4	2	0.5	0	11	0	340.14
5	1.86	0.5	0	10	0	336.33
6	1.76	0.5	0	10	0	334.9
7	2.05	0.333	0	13	0	369.32
8[test [a]]	1.55	0.833	0	11	0	321.06
9	1.66	0.833	0	9	0	299.85
10	1.99	0.667	0	8	0	335.9
11	2.35	0.667	0	8	0	345
12	2.16	0.667	0	8	0	325.8
13	2.32	0.5	0	8	0	320.85
14	2.53	0.5	0	8	0	368.96
15	1.84	0.667	1	10	0	294.79
16	1.42	0.833	1	8	0	301.55
17	2.05	0.667	1	11	0	293.3
18[test [a]]	1.63	0.667	1	9	0	304.64
19	1.97	0.667	1	10	0	295.06
20	1.55	0.667	1	8	0	301.8
21	2.24	0.667	0	8	0	335.93
22[test [a]]	2.4	0.667	0	8	2	335.77
23[test [a]]	2.01	0.667	0	8	2	337.59
24	1.57	0.667	0	6	2	339.04
25	1.63	0.667	0	6	1	337.27
26[test [a]]	2.46	0.5	0	8	0	365.96
27	2.89	0.667	0	8	0	357.51
31	2.48	0.667	0	8	1	339.7
33	2.79	0.667	0	8	1	364.79
43	1.72	0.667	0	14	0	368.72
44	2.51	0.5	0	16	0	368.04
45[test [a]]	3.08	0.667	0	20	0	366.31
46	4.65	0.5	0	24	0	367.39
61	4.1	0.5	0	17	3	369.51
BDP-1[ext [b]]	1.21	1	0	6	0	320.26
BDP-2[ext [b]]	1.07	1	0	6	0	324.36
BDP-3	0.973	1	0	6	0	321.65
BDP-4	1.55	0.667	0	6	0	326.55
BDP-5[ext [b]]	1.37	0.667	0	6	0	319.95
BDP-6	1.53	0.5	0	6	0	324.17
BDP-7[ext [b]]	1.72	0.5	0	6	0	314.59

[a] [test] designates compounds belonging to the test set. [b] [ext] relates to the molecules of the external set.

descriptors and Log Φ_Δ can be observed. For compounds **10**–**14** (BODIPYs containing methoxy-substituted phenyl group in the *meso* position) from the THF dataset, there is a strong correlation ($R^2 = 0.694$) between the HOMO energy and Log Φ_Δ. For BODIPYs **15**–**20** (containing amino-substituted phenyl group in the *meso* position), we found a determination coefficient between HOMO-LUMO gap and Log Φ_Δ equal to 0.683. However, when the whole THF dataset is regarded determination coefficients are very low: 0.214 and 0.094 for HOMO energy and HOMO-LUMO gap, respectively. A possible explanation is the structural diversity of molecules involved in the dataset. Previously, for a dataset containing 26 furocoumarins we have shown that Log Φ_Δ strongly correlates with the triplet state energy.[48] For pteridines, Log Φ_Δ correlated with ionization potential ($R^2 = 0.806$) and electronegativity ($R^2 = 0.840$).[47] However, in both cases datasets were smaller and more homogenous (having less structural diversity).

These results show that challenging first-principle computations of triplets formation by charge recombination for predicting SOCT-ISC efficiency can to a certain extent be replaced by QSPR involving AlvaDesc molecular descriptors which are computationally much faster to calculate and can be determined "on click" even for rather large molecules. The generated predictive models can serve as a simple and effective tool for guiding the design of SOCT-ISC photosensitizers with singlet oxygen quantum yield values optimized for desired range of polarity. Instead of randomly synthesizing donor-acceptor structures, QSPR models can be employed for in silico screening of large virtual libraries. Based on these predictions, a focused series of photosensitizers with tailored properties may be specifically selected and synthesized. As a proof of concept, in this work we succeeded to accurately predict $^1O_2$ yields for several newly synthesized BODIPYs. Such approach is well-established in medicinal chemistry, where QSPR allow to achieve much higher success rate and speed up the search of lead compounds. However, in the field of photochemistry and photosensitizers design, QSPR methods are still largely unexplored. The results reported here, in conjunction with our

previous studies, demonstrate that QSPR prediction of $^1O_2$ generation is applicable to various types of organic dyes and that this methodology requires minimal resources and time for accurate prediction of the photosensitizers' activity.

Experimental Section

Synthesis of BODIPY

Compounds **BDP 1–3** were prepared according to published procedures.[57,59] The four newly reported BODIPY compounds (**BDP 4–7**) were prepared according to a previously described general procedure.[58] All synthetic details and analytical data for new compounds are presented in the Supporting Information.

Computations

Conformational analysis of BODIPY molecules was performed using Spartan v. 16 modeling software from Wavefunction, Inc. (www.wavefun.com). Generation of low-energy conformers was performed using the MMFF force field.[76] Geometry optimization was done using Density Functional Theory. M06-2X functional[77] and 6-31G (d,p) basis set in Spartan v. 16 were used.

A total of 5305 molecular descriptors were calculated using alvaDesc v.1.0.22 program and online chemical modeling environment OCHEM.[78] The number of alvaDesc descriptors was reduced to 87 in the case of toluene, to 101 for acetonitrile, and to 107 for THF using Generic Algorithm v. 4.1 developed by the DTC lab and Kunal Roy.[79] Twenty three quantum-chemical descriptors, such as dipole moment, frontier molecular orbital energies, HOMO-LUMO gap, electronegativity, polarizability, and partial atomic charges were obtained at M06-2X/6-31G(d,p) level of theory for the gas phase. Energies of the two lowest singlet excited states (S_1 and S_2), energies of the two lowest triplet states (T_1 and T_2), as well as the singlet-triplet energy gap (ΔE_{ST}) and oscillator strengths were calculated using Time-Dependent Density Functional Theory (TD-DFT).

Machine learning (ML) was performed using scikit-learn library of Python programming language. SVR and RFR models search was performed in scikit-learn using grid-search method and 5-fold cross validation. During SVR model search three parameters were varied: C, epsilon, and kernel (linear, polynominal, sigmoid or radial basis function). RFR search was performed by changing the values of two parameters: number of estimators (trees) and maximal depth. Other parameters were used by default. In the case of MLR, Genetic Algorithm v. 4.1[79] was used for the model search. Data pretreatment included variance cut-off equal to 0.001 and intercorrelation cut-off equal to 0.9. Default values of all the parameters in Genetic Algorithm were used, except the equation length: the number of iterations/generations was equal to 100, the mutation probability was equal to 0.3, and the number of equations selected in each generation was equal to 30.

Statistical parameters: $R^2_{adjusted}$ is an adjusted coefficient of determination [Eq. (8)]:

$$R^2_{adjusted} = 1 - \frac{(1-R^2)(n-1)}{n-p-1} \tag{8}$$

where n is the total number of compounds in the training set; p is the number of predictors used by the model. Standard error of estimate (SEE) was calculated using Equation 9:

$$SEE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-p-1}} \tag{9}$$

where y_i and $\hat{y}_i$ are the actual and predicted Log Φ_Δ value of the i-th molecule in the training set. To calculate the root mean square error (RMSE), a similar Equation 10 was used:

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}} \tag{10}$$

To calculate q^2 (leave-one-out (LOO) cross-validation parameter) each molecule in the training set was excluded once and Log Φ_Δ of the excluded molecule was predicted by using the model developed for the remaining compounds. q^2 describes the internal stability of a model and was calculated using Equation (11):

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_{i,-i})^2}{\sum (y_i - y_{mean})^2} \tag{11}$$

where y_i and $\hat{y}_{i,-i}$ are the actual and predicted Log Φ_Δ values of the ith molecule in the training set, respectively; y_{mean} is the average Log Φ_Δ of all compounds in the training set. To calculate standard deviation error of prediction (SDEP) during LOO cross-validation Equation (12) was used:

$$SDEP = \sqrt{\frac{\sum (y_i - \hat{y}_{i,-i})^2}{n}} \tag{12}$$

where y_i and $\hat{y}_{i,-i}$ are the actual and predicted Log Φ_Δ values of the i-th molecule, respectively.

The activity of each compound in the test set was predicted using the model developed with the training set. R^2_{test} reflects the predictive ability of the model and was calculated using Equation (13):

$$R^2_{test} = 1 - \frac{\sum (y_{act} - y_{pred})^2}{\sum (y_{act} - y_{mean})^2} \tag{13}$$

where y_{act} and y_{pred} are the actual and predicted activity of the i-th compound in the test set, respectively; y_{mean} is the average value of Log Φ_Δ in the training set. Both summations are over all compounds in the test set. $R^2_{external}$ was calculated in a similar way as R^2_{test}, but for compounds included in the external set.

Singlet oxygen quantum yield determination

The singlet oxygen quantum yield measurements were performed according to the literature.[80] Solutions of the $^1O_2$ trap, 1,9-dimethylanthracene (DMA), with an optical density of around 1.4 in air-saturated solvent (acetonitrile, toluene, and tetrahydrofuran respectively) were employed. Corresponding BODIPY was added to the cuvette, and its absorbance was adjusted to around 0.29 at the wavelength of irradiation. The solutions in the cuvette were irradiated with 514 nm laser light at a constant power density of 12 mW cm^{-2}. The absorption spectra of the solutions were measured every 30–90 s. The slope of plots of absorbance of DMA at 376 nm vs. irradiation time for each photosensitizer was calculated.

Singlet oxygen quantum yields were calculated based on Equation (14):

$$\Phi_{\Delta} = \Phi_{\Delta}^{ref} \frac{k}{k_{ref}} \frac{I_{abs}^{ref}}{I_{abs}} \quad (14)$$

where Φ_{Δ} is the singlet oxygen quantum yield; the superscript ref stands for 2,6-diiodoBODIPY (0.85 in toluene);[81] k is the slope of the curves of DMA absorption (376 nm) change vs. irradiation time; I_{abs} represents the absorption correction factor which is given by $I = 1–10^{-OD}$ (OD is the optical density at 514 nm).

X-ray crystallography

Diffraction patterns were collected using CuK_{α} and MoK_{α} radiation (Bruker Duo, Bruker AXS package);[82] solved with direct methods (ShelXT)[83] and refined with ShelxI[84] in the shelxle GUI.[85] Non-H atoms were refined with anisotropic thermal parameters; H-atoms were placed at geometrically ideal positions with riding thermal parameters. Full details are given in the Supporting Information.

Deposition numbers 2069560, 2069563, 2069562, and 2069561 contain the supplementary crystallographic data for this paper. These data are provided free of charge by the joint Cambridge Crystallographic Data Centre and Fachinformationszentrum Karlsruhe Access Structures service.

Acknowledgements

This work was prepared with the support of funding from the European Union's Horizon 2020 research and innovation programme under the FET-OPEN grant agreement No.828779 and the Technical University of Munich – Institute for Advanced Study through a Hans Fischer Senior Fellowship. M.A.F. and A.S. acknowledge the TU Dublin Research Scholarship programme for support of this work. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: BODIPY · machine learning · photosensitization · structure–property relationship · singlet oxygen

[1] a) N. A. Romero, D. A. Nicewicz, *Chem. Rev.* **2016**, *116*, 10075–10166; b) M. H. Shaw, J. Twilton, D. W. C. MacMillan, *J. Org. Chem.* **2016**, *81*, 6898–6926; c) F. Glaser, O. S. Wenger, *Coord. Chem. Rev.* **2020**, *405*, 213129.
[2] a) S. Kwiatkowski, B. Knap, D. Przystupski, J. Saczko, E. Kedzierska, K. Knap-Czop, J. Kotlinska, O. Michel, K. Kotowski, J. Kulbacka, *Biomed. Pharmacother.* **2018**, *106*, 1098–1107; b) S. Callaghan, M. O. Senge, *Photochem. Photobiol. Sci.* **2018**, *17*, 1490–1519.
[3] a) F. N. Castellano, C. E. McCusker, *Dalton Trans.* **2015**, *44*, 17906–17910; b) J. Zhou, Q. Liu, W. Feng, Y. Sun, F. Li, *Chem. Rev.* **2015**, *115*, 395–465; c) P. Bharmoria, H. Bildirir, K. Moth-Poulsen, *Chem. Soc. Rev.* **2020**, *49*, 6529–6554.
[4] J. Zhao, K. Xu, W. Yang, Z. Wang, F. Zhong, *Chem. Soc. Rev.* **2015**, *44*, 8904–8939.
[5] L. Huang, X. Cui, B. Therrien, J. Zhao, *Chem. Eur. J.* **2013**, *19*, 17472–17482.
[6] a) J. Zhao, W. Wu, J. Sun, S. Guo, *Chem. Soc. Rev.* **2013**, *42*, 5323–5351; b) J. Zhao, K. Xu, W. Yang, Z. Wang, F. Zhong, *Chem. Soc. Rev.* **2015**, *44*, 8904–8939.
[7] a) S. P. Pitre, C. D. McTiernan, H. Ismaili, J. C. Scaiano, *J. Am. Chem. Soc.* **2013**, *135*, 13286–13289; b) E. Speckmeier, T. G. Fischer, K. Zeitler, *J. Am. Chem. Soc.* **2018**, *140*, 15353–15365; c) N. J. Treat, H. Sprafke, J. W. Kramer, P. G. Clark, B. E. Barton, J. Read de Alaniz, B. P. Fors, C. J. Hawker, *J. Am. Chem. Soc.* **2014**, *136*, 16096–16101; d) S. M. Sartor, B. G. McCarthy, R. M. Pearson, G. M. Miyake, N. H. Damrauer, *J. Am. Chem. Soc.* **2018**, *140*, 4778–4781; e) D. A. Nicewicz, T. M. Nguyen, *ACS Catal.* **2014**, *4*, 355–360; f) C. Li, Y. Xu, W. Tu, G. Chen, R. Xu, *Green Chem.* **2017**, *19*, 882–899; g) J. Zhao, K. Xu, W. Yang, Z. Wang, F. Zhong, *Chem. Soc. Rev.* **2015**, *44*, 8904–8939.
[8] J. Zhao, W. Wu, J. Sun, S. Guo, *Chem. Soc. Rev.* **2013**, *42*, 5323–5351.
[9] M. Bröring, R. Krüger, S. Link, C. Kleeberg, S. Köhler, X. Xie, B. Ventura, L. Flamigni, *Chem. Eur. J.* **2008**, *14*, 2976–2983.
[10] Y. Cakmak, S. Kolemen, S. Duman, Y. Dede, Y. Dolen, B. Kilic, Z. Kostereli, L. T. Yildirim, A. L. Dogan, D. Guc, E. U. Akkaya, *Angew. Chem. Int. Ed.* **2011**, *50*, 11937–11941; *Angew. Chem.* **2011**, *123*, 12143–12147.
[11] a) K. Nagarajan, A. R. Mallia, K. Muraleedharan, M. Hariharan, *Chem. Sci.* **2017**, *8*, 1776–1782; b) Z. Wang, L. Huang, Y. Yan, A. M. El-Zohry, A. Toffoletti, J. Zhao, A. Barbon, B. Dick, O. F. Mohammed, G. Han, *Angew. Chem. Int. Ed.* **2020**, *59*, 16114–16121; *Angew. Chem. Int. Ed.* **2020**, *132*, 16248–16255.
[12] D. Sasikumar, A. T. John, J. Sunny, M. Hariharan, *Chem. Soc. Rev.* **2020**, *49*, 6122–6140.
[13] a) Z. Wang, J. Zhao, *Org. Lett.* **2017**, *19*, 4492–4495; b) S. Kim, Y. Zhou, N. Tohnai, H. Nakatsuji, M. Matsusaki, M. Fujitsuka, M. Miyata, T. Majima, *Chem. Eur. J.* **2018**, *24*, 636–645.
[14] a) J. H. Golden, L. Estergreen, T. Porter, A. C. Tadle, D. M. R. Sylvinson, J. W. Facendola, C. P. Kubiak, S. E. Bradforth, M. E. Thompson, *ACS Appl. Mater. Interfaces* **2018**, *1*, 1083–1095; b) W. Huang, X. Zhang, B. Chen, H. Miao, C. O. Trindle, Y. Wang, Y. Luo, G. Zhang, *Chem. Commun.* **2019**, *55*, 67–70; c) J.-X. Wang, L.-Y. Niu, P.-Z. Chen, Y.-Z. Chen, Q.-Z. Yang, R. Boulatov, *Chem. Commun.* **2019**, *55*, 7017–7020.
[15] a) C. Trinh, K. O. Kirlikovali, S. Das, M. E. Ener, H. B. Gray, P. I. Djurovich, S. Bradforth, M. E. Thompson, *J. Phys. Chem. C* **2014**, *118*, 21834–21845; b) S. Das, W. G. Thornbury, A. N. Bartynski, M. E. Thompson, S. E. Bradforth, *J. Phys. Chem. Lett.* **2018**, *9*, 3264–3270; c) J. Karges, U. Basu, O. Blacque, H. Chao, G. Gasser, *Angew. Chem. Int. Ed.* **2019**, *58*, 14334–14340; *Angew. Chem.* **2019**, *131*, 14472–14478.
[16] S. M. Sartor, B. G. McCarthy, R. M. Pearson, G. M. Miyake, N. H. Damrauer, *J. Am. Chem. Soc.* **2018**, *140*, 4778–4781.
[17] Y. Tsuga, M. Katou, S. Kuwabara, T. Kanamori, S. Ogura, S. Okazaki, H. Ohtani, H. Yuasa, *Chem. Asian J.* **2019**, *14*, 2067–2071.
[18] a) Y. Zhao, X. Li, Z. Wang, W. Yang, K. Chen, J. Zhao, G. G. Gurzadyan, *J. Phys. Chem. C* **2018**, *122*, 3756–3772; b) Y. Zhao, R. Duan, J. Zhao, C. Li, *Chem. Commun.* **2018**, *54*, 12329–12332.
[19] J. W. Verhoeven, *J. Photochem. Photobiol. C* **2006**, *7*, 40–60.
[20] a) Z. E. X. Dance, S. M. Mickley, T. M. Wilson, A. B. Ricks, A. M. Scott, M. A. Ratner, M. R. Wasielewski, *J. Phys. Chem. A* **2008**, *112*, 4194–4201; b) Z. E. X. Dance, Q. Mi, D. W. McCamant, M. J. Ahrens, M. A. Ratner, M. R. Wasielewski, *J. Phys. Chem. B* **2006**, *110*, 25163–25173; c) D. J. Gibbons, A. Farawar, P. Mazzella, S. Leroy-Lhez, R. M. Williams, *Photochem. Photobiol. Sci.* 136–158.
[21] G. Grampp, *Angew. Chem. Int. Ed.* **1993**, *32*, 691–693; *Angew. Chem.* **1993**, *105*, 724–726.
[22] J. T. Buck, A. M. Boudreau, A. DeCarmine, R. W. Wilson, J. Hampsey, T. Mani, *Chem* **2019**, *5*, 138–155.
[23] V.-N. Nguyen, Y. Yim, S. Kim, B. Ryu, K. M. K. Swamy, G. Kim, N. Kwon, C.-Y. Kim, S. Park, J. Yoon, *Angew. Chem. Int. Ed.* **2020**, *59*, 8957–8962; *Angew. Chem.* **2020**, *132*, 9042–9047.
[24] Z. Mahmood, N. Rehmat, S. Ji, J. Zhao, S. Sun, M. Di Donato, M. Li, M. Teddei, Y. Huo, *Chem. Eur. J.* **2020**, *26*, 14912–14918.
[25] a) J. Deckers, T. Cardeynaels, H. Penxten, A. Ethirajan, M. Ameloot, M. Kruk, B. Champagne, W. Maes, *Chem. Eur. J.* **2020**, *26*, 15212–15225; b) Y. Dong, A. Elmali, J. Zhao, B. Dick, A. Karatay, *ChemPhysChem* **2020**, *21*, 1388–1401.
[26] I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, R. Taylor, *Acta Crystallogr.* **2002**, B*58*, 389–397.
[27] a) Z. Wang, J. Zhao, *Org. Lett.* **2017**, *19*, 4492–4495; b) M. A. Filatov, S. Karuthedath, P. M. Polestshuk, S. Callaghan, K. Flanagan, M. Telitchko, T. Wiesner, F. Laquai, M. O. Senge, *Phys. Chem. Chem. Phys.* **2018**, *20*, 8016–8031; c) Y. Hu, Y. Hou, Z. Wang, Y. Li, J. Zhao, *J. Chem. Phys.* **2020**, *153*, 224304.

[28] M. A. Filatov, S. Karuthedath, P. M. Polestshuk, S. Callaghan, K. J. Flanagan, T. Wiesner, F. Laquai, M. O. Senge, *ChemPhotoChem* **2018**, *2*, 606–615.
[29] a) Z. Wang, J. Zhao, M. Di Donato, G. Mazzone, *Chem. Commun.* **2019**, *55*, 1510–1513; b) Z. Wang, M. Ivanov, Y. Gao, L. Bussotti, P. Foggi, H. Zhang, N. Russo, B. Dick, J. Zhao, M. Di Donato, G. Mazzone, L. Luo, M. Fedin, *Chem. Eur. J.* **2020**, *26*, 1091–1102.
[30] K. Chen, W. Yang, Z. Wang, A. Iagatti, L. Bussotti, P. Foggi, W. Ji, J. Zhao, M. Di Donato, *J. Phys. Chem. A* **2017**, *121*, 7550–7564.
[31] Y. Hou, I. Kurganskii, A. Elmali, H. Zhang, Y. Gao, L. Lv, J. Zhao, A. Karatay, L. Luo, M. Fedin, *J. Chem. Phys.* **2020**, *152*, 114701.
[32] Y. Dong, A. A. Sukhanov, J. Zhao, A. Elmali, X. Li, B. Dick, A. Karatay, V. K. Voronkova, *J. Phys. Chem. C* **2019**, *123*, 22793–22811.
[33] M. A. Filatov, S. Karuthedath, P. M. Polestshuk, H. Savoie, K. J. Flanagan, C. Sy, E. Sitte, M. Telitchko, F. Laquai, R. W. Boyle, M. O. Senge, *J. Am. Chem. Soc.* **2017**, *139*, 6282–6285.
[34] a) S. Callaghan, M. A. Filatov, H. Savoie, R. W. Boyle, M. O. Senge, *Photochem. Photobiol. Sci.* **2019**, *18*, 495–504; b) S. Callaghan, B. E. Vindstad, K. J. Flanagan, T. B. Melø, M. Lindgren, K. Grenstad, O. A. Gederaas, M. O. Senge, *ChemPhotoChem* **2021**, *5*, 131–141.
[35] N. Kiseleva, M. A. Filatov, M. Oldenburg, D. Busko, M. Jakoby, I. A. Howard, B. S. Richards, M. O. Senge, S. M. Borisov, A. Turshatov, *Chem. Commun.* **2018**, *54*, 1607–1610.
[36] N. Kiseleva, D. Busko, B. S. Richards, M. A. Filatov, A. Turshatov, *J. Phys. Chem. Lett.* **2020**, *11*, 6560–6566.
[37] A. B. Nepomnyashchii, A. J. Bards, *Acc. Chem. Res.* **2012**, *45*, 1844–1853.
[38] M. A. Filatov, *Org. Biomol. Chem.* **2020**, *18*, 10–27.
[39] X.-F. Zhang, X. Yang, *J. Phys. Chem. B* **2013**, *117*, 9050–9055.
[40] N. Epelde-Elezcano, E. Palao, H. Manzano, A. Prieto-CastaCeda, A. R. Agarrabeitia, A. Tabero, A. Villanueva, S. de la Moya, C. Ljpez-Arbeloa, V. Martinez-Martinez, M. J. Ortiz, *Chem. Eur. J.* **2017**, *23*, 4837–4848.
[41] J. F. Lovell, T. W. B. Liu, J. Chen, G. Zheng, *Chem. Rev.* **2010**, *110*, 2839–2857.
[42] a) C. M. Isborn, B. D. Mar, B. F. E. Curchod, I. Tavernelli, T. J. Martínez, *J. Phys. Chem. B* **2013**, *117*, 12189–12201; b) S. Bhandari, B. D. Dunietz, *J. Chem. Theory Comput.* **2019**, *15*, 4305–4311.
[43] J. Shi, F. Luan, H. Zhang, M. Liu, Q. Guo, Z. Hu, B. Fan, *QSAR Comb. Sci.* **2006**, *25*, 147–155.
[44] A. Schüller, G. B. Goh, H. Kim, J.-S. Lee, Y.-T. Chang, *Mol. Inf.* **2010**, *29*, 717–729.
[45] E. Caruso, M. Gariboldi, A. Sangion, P. Gramatica, S. Banfi, *J. Photochem. Photobiol. B* **2017**, *167*, 269–281.
[46] J. Zhao, Y. Zhou, C. Li, Q. Xie, J. Chen, G. Chen, W. J. G. M. Peijnenburg, Y. N. Zhang, J. Qu, *Sci. Total Environ.* **2020**, *712*, 136450
[47] A. A. Buglak, T. A. Telegina, M. S. Kritsky, *Photochem. Photobiol. Sci.* **2016**, *15*, 801–811.
[48] A. A. Buglak, A. I. Kononov, *New J. Chem.* **2018** *42*, 14424–14432.
[49] A. A. Buglak, M. A. Filatov, M. A. Hussain, M. Sugimoto, *J. Photochem. Photobiol. A* **2020**, *403*, 112833.
[50] I. S. Turan, G. Gunaydin, S. Ayan, E. U. Akkaya, *Nat. Commun.* **2018**, *9*, 805.
[51] T. Yogo, Y. Urano, A. Mizushima, H. Sunahara, T. Inoue, K. Hirose, M. Iino, K. Kikuchi, T. Nagano, *Proc. Nat. Acad. Sci. U. S. A.* **2008**, *105*, 28–32.
[52] C. Schweitzer, R. Schmidt, *Chem. Rev.* **2003**, *103*, 1685–1758.
[53] a) H. Liang, S. Sun, M. Zafar, Z. Yuan, Y. Dong, S. Ji, Y. Huo, M.-D. Li, J. Zhao, *Dyes Pigm.* **2020**, *173*, 108003; b) Y. Hou, J. Liu, N. Zhang, J. Zhao, *J. Phys. Chem. A* **2020**, *124*, 9360–9374.
[54] W. Hu, Y. Lin, X.-F. Zhang, M. Feng, S. Zhao, J. Zhang, *Dyes Pigm.* **2019**, *164*, 139–147.
[55] a) A. E. O'Connor, W. M. Gallagher, A. T. Byrne, *Photochem. Photobiol.* **2009**, *85*, 1053–1074; b) T. M. Baran, *Lasers Surg. Med.* **2018**, *50*, 476–482.
[56] a) W.-S. Cho, H.-J. Kim, B. J. Littler, M. A. Miller, C.-H. Lee, J. S. Lindsey, *J. Org. Chem.* **1999**, *64*, 7890–7901; b) C.-H. Lee, J. S. Lindsey, *Tetrahedron* **1994**, *50*, 11427–11440.
[57] a) M. Elisa Milanesio, F. S. Morán, E. Ines Yslas, M. Gabriela Alvarez, V. Rivarola, E. N. Durantini, *Bioorg. Med. Chem.* **2001**, *9*, 1943–1949; b) R. R. Kavali, B. Chul Lee, B. Seok Moon, S. Dae Yang, K. Soo Chun, C. Woon Choi, C.-H. Lee, D. Yoon Chi, *J. Labelled Compd. Radiopharm.* **2005**, *48*, 749–758; c) B. J. Littler, M. A. Miller, C.-H. Hung, R. W. Wagner, D. F. O'Shea, P. D. Boyle, J. S. Lindsey, *J. Org. Chem.* **1999**, *64*, 1391–1396; d) A. Ogawa, K. Oohora, T. Hayashi, *Inorg. Chem.* **2018**, *57*, 14644–14652; e) A. J. Pistner, D. A. Lutterman, M. J. Ghidiu, Y.-Z. Ma, J. Rosenthal, *J. Am. Chem. Soc.* **2013**, *135*, 6601–6607; f) W. Senapak, R. Saeeng, J. Jaratjaroonphong, T. Kasemsuk, U. Sirion, *Org. Biomol. Chem.* **2016**, *14*, 1302–1310.
[58] a) G. Ulrich, R. Ziessel, A. Harriman, *Angew. Chem. Int. Ed.* **2008**, *47*, 1184–1201; *Angew. Chem.* **2008**, *120*, 1202–1219; b) R. W. Wagner, J. S. Lindsey, *Pure Appl. Chem.* **1996**, *68*, 1373–1380.
[59] a) F. Frank, L. M. Alice, P. Mauker, A. A. Alsimaree, P. G. Waddell, M. R. Probert, T. J. Penfold, J. G. Knight, M. J. Hall, *Tetrahedron* **2020**, *76*, 131113; b) T.-I. Kim, S. B. Maity, J. Bouffard, Y. Kim, *Anal. Chem.* **2016**, *88*, 9259–9263; c) P. E. Kesavan, R. N. Behera, S. Mori, I. Gupta, *J. Fluoresc.* **2017**, *27*, 2131–2144.
[60] R. I. Roacho, A. J. Metta-Magaña, E. Peña-Cabrera, K. H. Pannell, *J. Phys. Org. Chem.* **2013**, *26*, 345–351.
[61] T. V. Goud, A. Tutar, J.-F. Biellmann, *Tetrahedron* **2006**, *62*, 5084–5091.
[62] M. R. Rao, R. Bolligarla, R. J. Butcher, M. Ravikanth, *Inorg. Chem.* **2010**, *49*, 10606–10616.
[63] P. Luo, E. C. Feinberg, G. Guirado, S. Farid, J. P. Dinnocenzo, *J. Org. Chem.* **2014**, *79*, 9297–9304.
[64] A. Golbraikh, A. Tropsha, *J. Mol. Graphics Modell.* **2002**, *20*, 269–276.
[65] L. H. Hall, B. Mohney, L. B. Kier, *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 76–81.
[66] V. Consonni, R. Todeschini, M. Pavan, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 682–692.
[67] L. B. Kier, L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*. Academic Press, New York, **1976**.
[68] P. Ertl, B. Rohde, P. Selzer, *J. Med. Chem.* **2000**, *43*, 3714–3717.
[69] G. Chen, S. Zheng, X. Luo, J. Shen, W. Zhu, H. Liu, C. Gui, J. Zhang, M. Zheng, C. Puah, K. Chen, H. Jiang, *J. Comb. Chem.* **2005**, *7*, 398–406.
[70] O. Devinyak, D. Havrylyuk, R. Lesyk, *J. Mol. Graphics Modell.* **2014**, *54*,194-203.
[71] M. Doga Ertürk, M. Türker Sacan, M. Novic, N. Minovski, *J. Mol. Graphics Modell.* **2012**, *38*, 90–100.
[72] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH: Weinheim, Germany, **2000**
[73] L. T. Qin, S. S. Liu, H. L. Liu, *Mol. Diversity* **2010**, *14*, 67–80.
[74] D. W. Cho, M. Fujitsuka, J. H. Ryu, M. H. Lee, H. K. Kim, T. Majima, C. Im, *Chem. Commun.* **2012**, *48*, 3424–3426.
[75] M. Congreve, R. Carr, C. Murray, H. Jhoti, *Drug Discovery Today* **2003**, *8*, 876–877.
[76] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.
[77] a) Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241; b) Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.* **2008**, *41*, 157–167.
[78] I. Sushko, S. Novotarskyi, R. Körner, A. K. Pandey, M. Rupp, W. Teetz, S. Brandmaier, A. Abdelaziz, V. V. Prokopenko, V. Y. Tanchuk, R. Todeschini, A. Varnek, G. Marcou, P. Ertl, V. Potemkin, M. Grishina, J. Gasteiger, C. Schwab, I. I. Baskin, V. A. Palyulin, E. V. Radchenko, W. J. Welsh, V. Kholodovych, D. Chekmarev, A. Cherkasov, J. Aires-de-Sousa, Q.-Y. Zhang, A. Bender, F. Nigsch, L. Patiny, A. Williams, V. Tkachenko, I. V. Tetko, *J. Comput.-Aided Mol. Des.* **2011**, *25*, 533–554.
[79] K. Roy, *J. Indian Chem. Soc.* **2018**, *95*, 1497–1502.
[80] L. Gou, C. N. Coretsopoulos, A. B. Scranton, *J. Polym. Sci. Part A* **2004**, *42*, 1285–1292.
[81] Y. Zhao, R. Duan, J. Zhao, C. Li, *Chem. Commun.* **2018**, *54*, 12329–12332.
[82] APEX 3 and SAINT. Bruker AXS Inc., Madison, Wisconsin, USA, **2015**.
[83] G. M. Sheldrick, *Acta Crystallogr.* **2008**, A*64*, 112–122.
[84] G. M. Sheldrick, *Acta Crystallogr.* **2015**, C*71*, 3–8.
[85] C. B. Hübschle, G. M. Sheldrick, B. Dittrich, *J. Appl. Crystallogr.* **2011**, *44*, 1281–1284.

Manuscript received: March 13, 2021
Accepted manuscript online: April 20, 2021
Version of record online: May 26, 2021

