

An Open Access Data Set Highlighting Aggregation of Dyes on Metal Oxides

Vishwesh Venkatraman * and Lethesh Kallidanthiyil Chellappan 

Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway; lethesh.k.chellappan@ntnu.no

* Correspondence: vishwesh.venkatraman@ntnu.no

Received: 2 April 2020; Accepted: 11 May 2020; Published: 13 May 2020



Abstract: The adsorption of a dye to a metal oxide surface such as TiO₂, NiO and ZnO leads to deprotonation and often undesirable aggregation of dye molecules, which in turn impacts the photophysical properties of the dye. While controlled aggregation is useful for some applications, it can result in lower performance for dye-sensitized solar cells. To understand this phenomenon better, we have conducted an extensive search of the literature and identified over 4000 records of absorption spectra in solution and after adsorption onto metal oxide. The total data set comprises over 3500 unique compounds, with observed absorption maxima in solution and after adsorption on the semiconductor electrode. This data may serve to provide further insight into the structure-property relationships governing dye-aggregation behaviour.

Dataset: The data set can be queried at <https://vishwesh.github.io/dyeaggregation>. Search results can be exported as CSV files.

Dataset License: CC BY 4.0.

Keywords: dye sensitized solar cell; aggregation; database; absorption; solvent polarity

1. Background & Summary

Aggregation in small molecules has been shown to alter the shape of the absorption bands [1]. This phenomenon has been used advantageously in organic photoconductive materials for xerography [2], sensors [3], colour filters for fluorescence detection [4], organic photovoltaics [5,6] and optoelectronics [7–9]. For dye sensitized solar cells (DSSCs) in particular, aggregation often reduces device efficiency. This is largely due to dye-dye interactions on the metal oxide (during sensitization) surface that facilitate the formation of aggregates [1,6]. These events in turn lead to shifts in the absorption spectrum compared to the dye in solution. While H-aggregates lead to a hypsochromic/blue shift, J-aggregates result in bathochromic/red shift of the absorption spectra. Such shifts can be either substantial (indicating considerable aggregation) or negligible [10]. While J-aggregation can enhance the device performance (due to broadening of the absorption spectra), H-aggregation invariably lowers the light-harvesting capability.

In identifying new dyes for DSSCs, multiple criteria such as broad absorption spectra, stable metal oxide binding and reduced aggregation need to be considered. Among these, aggregation is commonly occurring phenomenon that is not easily eliminated. In many cases, cholic acid derivatives such as chenodeoxycholic acid (CDCA) are used [11–14] to prevent $\pi - \pi$ stacking, thereby reducing aggregation but has the unwanted side effect of lowering the surface concentration (dye loading) of the sensitizer. Owing to the planar and extended π configurations, organic dyes show an increased tendency to aggregate leading to inefficient electron injection and lower performance. In order to

impede the process, modifications to the dye structure include addition of long-chain alkyl groups or bulky substituents. In a recent study, CDCA was attached covalently to triarylamine sensitizers with encouraging results [15].

We have recently proposed a machine learning route to predict whether dye adsorption on titania is likely to induce a change in its absorption characteristics [10]. In this article, we present a new data set containing experimentally determined UV/vis absorption maxima in solution and on the metal oxide. This work builds on previously published research on dye sensitized solar cells [16] and aims to enable the scientific community in conducting new studies into data-driven materials discovery for photovoltaics and optoelectronic applications.

2. Methods

2.1. Data Acquisition

The data was extracted from over 1500 literature articles. For each dye, the absorption maxima in solution ($\lambda_{\max}^{\text{soln}}$) and after deposition on the metal oxide ($\lambda_{\max}^{\text{MO}}$) were taken from tabulated values, or alternatively from the images of the spectra. Where possible, the chemical name to structure parser (OPSIN [17]) was used, failing which, structures were drawn using MarvinSketch [18]. Experimental data in the presence of additives such as chenodeoxycholic acid were excluded. The structures were subsequently saved in the SMILES format. The overall format of the data records is shown in Table 1.

Table 1. Description of data records. For solvent mixtures, the volume ratio was noted. The type of aggregation is determined based on Equation (2).

Key	Description	Data Type
DOI	Source document Digital Object Identifier	String
STRUCTURE	Dye molecules stored in SMILES format	
$\lambda_{\max}^{\text{soln}}$	Observed maximum wavelength in solvent (in nm)	Float
$\lambda_{\max}^{\text{MO}}$	Observed maximum wavelength in the metal oxide (in nm)	Float
SOLVENT	Solvent used for recording spectra	String
METAL OXIDE	Semiconductor metal oxide	String
AGGREGATION	Hypsochromic/Bathochromic/Unchanged	String

The type of aggregation was determined based on difference between the solution phase and solid-state maxima given by:

$$\Delta\lambda = \lambda_{\max}^{\text{soln}} - \lambda_{\max}^{\text{MO}} \quad (1)$$

In order to provide a reasonable distinction between dyes that show low levels of aggregation from those that show large red/blue shifts, the values of $\Delta\lambda$ were grouped into three categories—hypsochromic, bathochromic and unchanged with respect to the following criteria:

$$\lambda_{\text{Agg}} = \begin{cases} \text{Unchanged} & \text{if } |\Delta\lambda| \leq 10 \text{ nm} \\ \text{Bathochromic} & \text{if } \Delta\lambda < -10 \text{ nm} \\ \text{Hypsochromic} & \text{otherwise} \end{cases} \quad (2)$$

The selected value is an arbitrary choice, but provides a simple scheme that enables easy distinction between H and J-aggregates. The value is also set to a small number to allow for limited experimental variation. Thus, an absorption shift of ± 10 nm implies a low degree of molecular aggregation onto the nanoparticle surface, while larger values on either side suggest moderate to substantial hypsochromic/bathochromic aggregation. This categorization was further used to create machine learning models [10] that were able to predict the nature of the shift with good accuracy.

2.2. Data Analysis

The data set contains a total of 4035 entries, spanning 3685 unique dyes. While studies have largely focused on titania (TiO₂), other metal oxides such as ZnO and NiO have also been used (see Table 2). Spectroscopic properties show a significant solvent dependence. To account for the solvatochromic effects, the data set collates information on the different solvents used for spectroscopic measurements. A total 37 solvents that include 18 pure solvents and 19 solvent mixtures have been studied experimentally. Table 3 gives a solvent-wise distribution of the dyes tested. For reasons of simplicity and readability, we do not make any distinction between mixtures containing the same solvents in different volume ratios. The sensitization solvents are often chosen based on solubility of the dyes. In general, poorly soluble dyes can exhibit aggregation [19]. Solvent polarity can exert considerable on aggregation. For example, high polarity solvents may cause a red shift of the absorption peaks [6,10,20,21]. Figure 1 provides an overview of the impact of polarity on aggregation. In the case of solvents such as N-Methyl-2-Pyrrolidone and acetic acid, only single entries were available. For N-Methyl-2-Pyrrolidone, the spectra was hypsochromically shifted while for acetic acid negligible change with respect to titania was observed.

Table 2. Number of entries with respect to the metal oxide semiconductors.

Metal Oxide	#Entries
TiO ₂	3988
ZnO	28
NiO	16
Al ₂ O ₃	3

Table 3. Distribution of structures according to the solvent used.

Solvent	#Entries	Solvent	#Entries
1,2-dichlorobenzene	7	acetic acid/dichloromethane	22
1,2-dichloroethane	4	acetonitrile/ethanol	4
1,4-dioxane	39	acetonitrile/tert-butanol	42
acetic acid	1	chloroform/ethanol	9
acetone	13	chloroform/methanol	66
chlorobenzene	10	chloroform/methanol/triethylamine	3
chloroform	455	dichloromethane/ethanol	27
dichloromethane	1218	dichloromethane/dimethylformamide	1
dimethyl sulfoxide	59	dichloromethane/methanol	15
dimethylformamide	312	dichloromethane/tetrahydrofuran	21
ethanol	339	dimethyl sulfoxide/acetonitrile	6
ethyl acetate	11	dimethyl sulfoxide/ethanol	8
hexane	7	dimethylformamide/ethanol	5
methanol	92	dimethylformamide/methanol	3
N-Methyl-2-Pyrrolidone	1	dimethylformamide/water	3
tetrahydrofuran	953	ethanol/tetrahydrofuran	2
toluene	96	ethyl acetate/methanol	4
		tetrahydrofuran/toluene	13

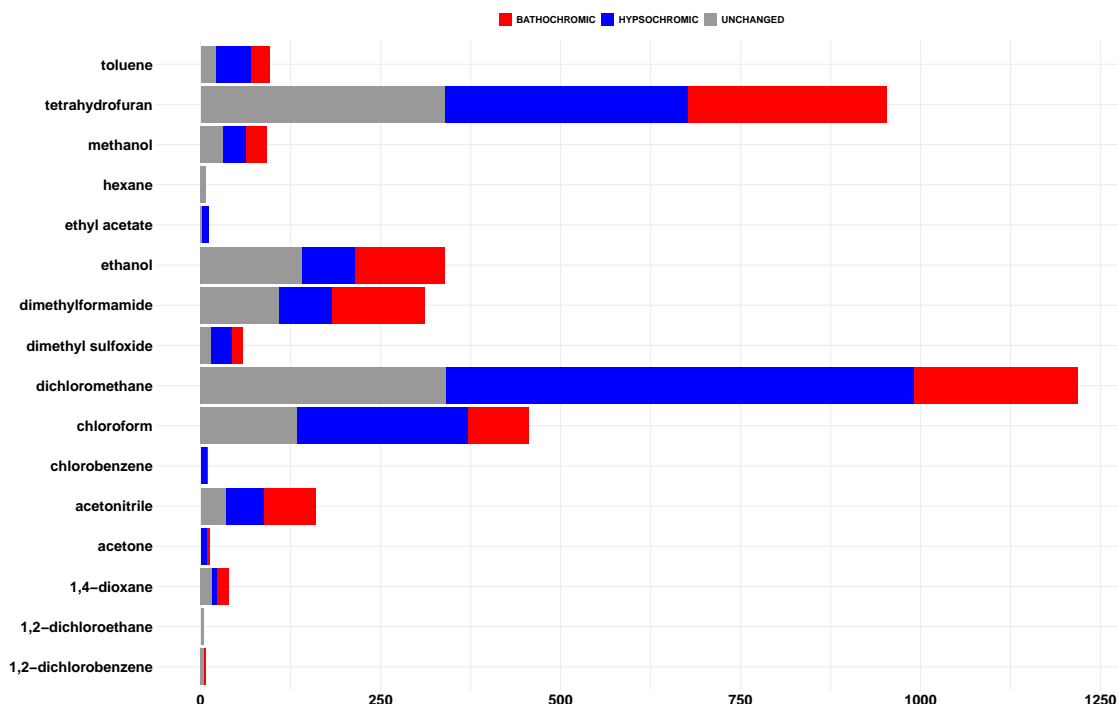


Figure 1. Impact of solvent polarity on aggregation. For each pure solvent, the stacked barplot shows the number of blue/red-shifted or unchanged categories (as defined by Equation (2)). N-Methyl-2-Pyrrolidone and acetic acid are not shown as only single entries were available for these solvents.

In order to uncover frequently occurring moieties, we used the MolBlocks [22] software where the minimum size for a fragment was set to four atoms. A total of 2126 fragments were generated by applying RECAP [23] rules. The fragments include a number of donor scaffolds such as triphenylamine, phenothiazine, phenoxazine, bodipy, ruthenium, carbazole, porphyrin, julolidine, indoline, coumarin and functional groups that include various benzene derivatives (phenol, anisole, aniline, naphthalene, anthracene), anchoring/acceptor moieties (cyanoacrylic acid, benzoic acid, 2-cyanoprop-2-enoic acid, malononitrile), alkyl chains of varying lengths (butyl, hexyl, octyl, dodecyl) and conjugation groups such as pyrene, quinoxaline, thiophene, furan, pyrrole, pyridopyrazine [24]. Selected fragments that are found in many dyes are shown in Figure 2.

Among these fragments, we focused on the aggregation behaviour of the different donor classes and anchoring groups. The pie charts shown in Figure 3A,B summarize the data in terms of the dye class and type of anchoring groups, respectively. While zinc porphyrin and ruthenium dyes show a lower tendency to aggregate (as seen by the relatively small proportion of dyes showing hypsochromic/bathochromic shifts), metal free dyes clearly show a higher tendency to aggregate. Further examination of Figure 3A shows that dyes containing imidazole form J-aggregates while other dye classes tend to exhibit both blue and red shifts. Analysis of the anchoring groups (see Figure 3B) does not suggest a clear explanation of the behaviour, echoing what was seen for the donors. Nonetheless, the dyes containing catechol show a significant propensity for J-aggregation which may be attributed to the fact that these groups bind more strongly to the metal oxide than other anchoring groups [25].

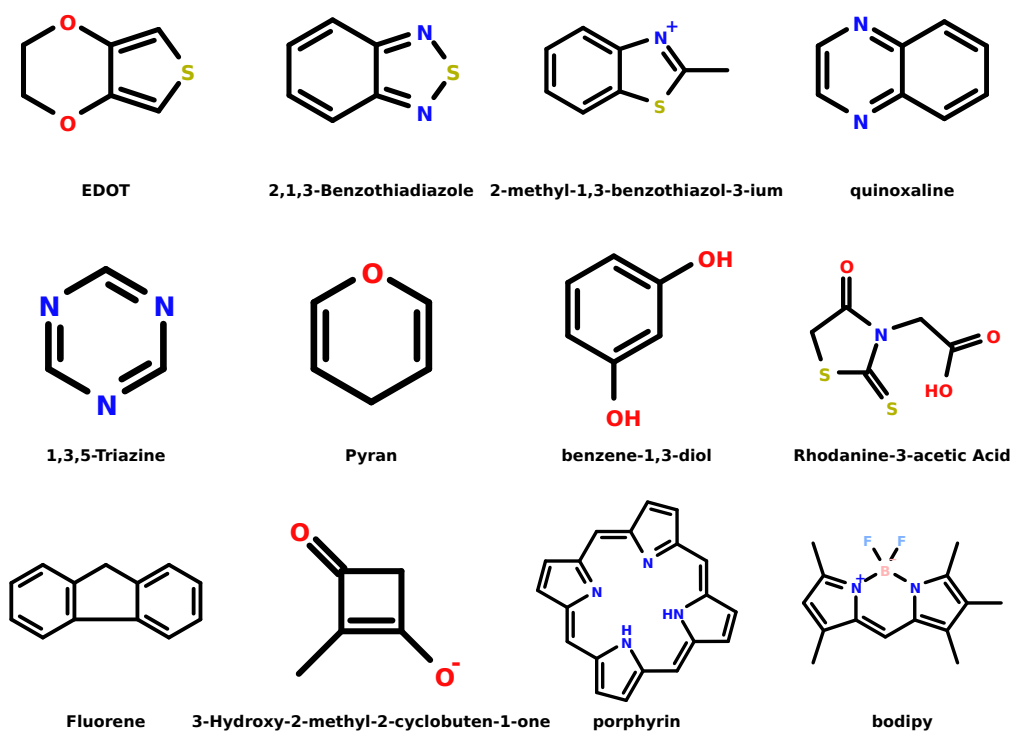


Figure 2. Frequently occurring fragments in the dyes.

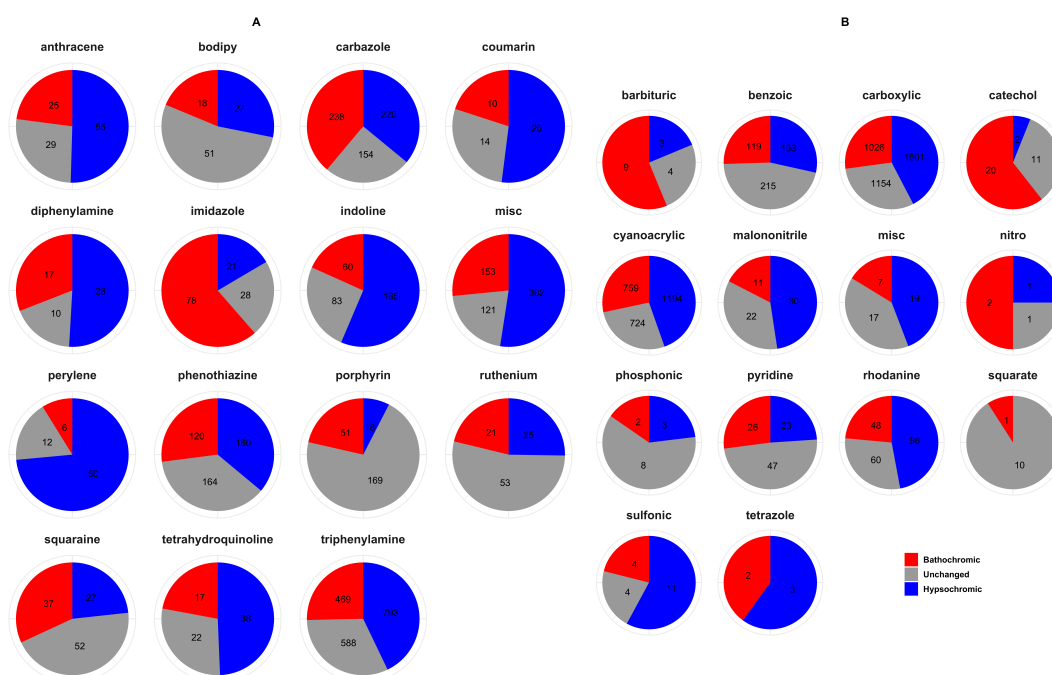


Figure 3. Pie charts showing the distribution of the absorption shifts (irrespective of the solvent) based on the (A) class of the dyes and (B) the anchoring groups used. (A) The “misc” category includes various dyes containing fluorene, phenoxazine, truxene, N,N-dialkylaniline, julolidine and other donor classes. (B) The “misc” category includes dyes anchored to the metal oxide via thiazolidinedione, aldehyde, hydantoin, hydroxybenzonnitrile, alkoxy silane and other groups.

3. Usage Notes

The data set can be queried using an online search utility (see Figure 4) that can be accessed at <https://vvishwesh.github.io/dyeaggregation>. Search queries can be based on solvent, class of the dye, anchoring group, metal oxide type ($\text{TiO}_2/\text{ZnO}/\text{NiO}/\text{Al}_2\text{O}_3$), and absorption maxima range. Alternatively, a structure/substructure-based search can be performed. The comparison and search for molecule structures on web browsers has been enabled using the open source JavaScript library Kekule.js [26]. The JavaScript Molecule Editor [27] (JSME) is used to provide interactive editing of molecules. The results are displayed as a table with molecular structures displayed using the SmilesDrawer [28] JavaScript library and can be saved as a tab-separated file for future use.

The Dye Aggregation Database

Search the database for dyes satisfying criteria

λ_{max}	STRUCTURE	λ_{min}	METAL_OXIDE	SOLVENT	AGGREGATION	DOI
516		512	TiO2	chloroform	UNCHANGED	10.1021/acsomega.7c01387
509		498	TiO2	chloroform	HYPOCHROMIC	10.1021/acsomega.7c01387
568		551	TiO2	chloroform	HYPOCHROMIC	10.1021/acsomega.7c01387

(a) Query form.

(b) Search results.

Figure 4. Browser-based search utility. (a) User interface for searching the aggregation data set. Users can select different options on the form for narrowing the search. (b) Results are displayed as a table. A maximum of 100 records can be displayed. The results can be retrieved as a semicolon delimited CSV file.

Author Contributions: V.V. and L.K.C. performed the data curation. V.V. performed the data analysis models. V.V. wrote the paper with contributions from L.K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported through a grant (Grant No. 262152) from the Research Council of Norway.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DSSC	Dye Sensitized Solar Cells
JSME	JavaScript Molecule Editor
RECAP	Retrosynthetic Combinatorial Analysis Procedure
OPSIN	Open Parser for Systematic IUPAC nomenclature
SMILES	simplified molecular-input line-entry system

References

1. Hestand, N.J.; Spano, F.C. Expanded Theory of H- and J-Molecular Aggregates: The Effects of Vibronic Coupling and Intermolecular Charge Transfer. *Chem. Rev.* **2018**, *118*, 7069–7163. doi:10.1021/acs.chemrev.7b00581. [[CrossRef](#)] [[PubMed](#)]
2. Law, K.Y. Organic photoconductive materials: recent trends and developments. *Chem. Rev.* **1993**, *93*, 449–486. doi:10.1021/cr00017a020. [[CrossRef](#)]
3. Ajayaghosh, A. Chemistry of Squaraine-Derived Materials: Near-IR Dyes, Low Band Gap Systems, and Cation Sensors. *Acc. Chem. Res.* **2005**, *38*, 449–459. doi:10.1021/ar0401000. [[CrossRef](#)] [[PubMed](#)]
4. Yamazaki, M.; Hofmann, O.; Ryu, G.; Xiaoe, L.; Lee, T.K.; deMello, A.J.; deMello, J.C. Non-emissive colour filters for fluorescence detection. *Lab Chip* **2011**, *11*, 1228. doi:10.1039/c0lc00642d. [[CrossRef](#)] [[PubMed](#)]
5. Deing, K.C.; Mayerhöffer, U.; Würthner, F.; Meerholz, K. Aggregation-dependent photovoltaic properties of squaraine/PC61BM bulk heterojunctions. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8328. doi:10.1039/c2cp40789b. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, L.; Cole, J.M. Dye aggregation in dye-sensitized solar cells. *J. Mater. Chem. A* **2017**, *5*, 19541–19559. doi:10.1039/c7ta05632j. [[CrossRef](#)]
7. Gsänger, M.; Kirchner, E.; Stolte, M.; Burschka, C.; Stepanenko, V.; Pflaum, J.; Würthner, F. High-Performance Organic Thin-Film Transistors of J-Stacked Squaraine Dyes. *J. Am. Chem. Soc.* **2014**, *136*, 2351–2362. doi:10.1021/ja409496r. [[CrossRef](#)]
8. Cappello, D.; Therien, D.A.B.; Staroverov, V.N.; Lagugné-Labarthe, F.; Gilroy, J.B. Optoelectronic, Aggregation, and Redox Properties of Double-Rotor Boron Difluoride Hydrazone Dyes. *Chem. Eur. J.* **2019**, *25*, 5994–6006. doi:10.1002/chem.201900383. [[CrossRef](#)]
9. Kim, S.O.; An, T.K.; Chen, J.; Kang, I.; Kang, S.H.; Chung, D.S.; Park, C.E.; Kim, Y.H.; Kwon, S.K. H-Aggregation Strategy in the Design of Molecular Semiconductors for Highly Reliable Organic Thin Film Transistors. *Adv. Funct. Mater.* **2011**, *21*, 1616–1623. doi:10.1002/adfm.201002367. [[CrossRef](#)]
10. Venkatraman, V.; Yemene, A.E.; de Mello, J. Prediction of Absorption Spectrum Shifts in Dyes Adsorbed on Titania. *Sci. Rep.* **2019**, *9*. doi:10.1038/s41598-019-53534-2. [[CrossRef](#)]
11. Choi, I.T.; Ju, M.J.; Kang, S.H.; Kang, M.S.; You, B.S.; Hong, J.Y.; Eom, Y.K.; Song, S.H.; Kim, H.K. Structural effect of carbazole-based coadsorbents on the photovoltaic performance of organic dye-sensitized solar cells. *J. Mater. Chem. A* **2013**, *1*, 9114. doi:10.1039/c3ta11508a. [[CrossRef](#)]
12. Magne, C.; Urien, M.; Ciofini, I.; Tugsuz, T.; Pauporté, T. Amphiphilic acids as co-adsorbents of metal-free organic dyes for the efficient sensitization of nanostructured photoelectrode. *RSC Adv.* **2012**, *2*, 11836. doi:10.1039/c2ra22121g. [[CrossRef](#)]
13. Zhang, F.; Wang, S.R.; Li, X.G.; Xiao, Y.; Guo, J.J. The effect of coadsorbent and solvent on the photovoltaic performance of 2, 9, 16, 23-Tetrakis(7-coumarinoxy-4-methyl)-phthalocyaninatocopper-sensitized solar cells. *J. Mol. Struct.* **2016**, *1107*, 329–336. doi:10.1016/j.molstruc.2015.11.074. [[CrossRef](#)]
14. Neale, N.R.; Kopidakis, N.; van de Lagemaat, J.; Grätzel, M.; Frank, A.J. Effect of a Coadsorbent on the Performance of Dye-Sensitized TiO₂Solar Cells: Shielding versus Band-Edge Movement. *J. Phys. Chem. B* **2005**, *109*, 23183–23189. doi:10.1021/jp0538666. [[CrossRef](#)]
15. Buene, A.F.; Almenningen, D.M.; Hagfeldt, A.; Gautun, O.R.; Hoff, B.H. First Report of Chenodeoxycholic Acid-Substituted Dyes Improving the Dye Monolayer Quality in Dye-Sensitized Solar Cells. *Sol. RRL* **2020**, 1900569. doi:10.1002/solr.201900569. [[CrossRef](#)]
16. Venkatraman, V.; Raju, R.; Oikonomopoulos, S.P.; Alsberg, B.K. The dye-sensitized solar cell database. *J. Cheminform.* **2018**, *10*. doi:10.1186/s13321-018-0272-0. [[CrossRef](#)]

17. Lowe, D.M.; Corbett, P.T.; Murray-Rust, P.; Glen, R.C. Chemical Name to Structure: OPSIN, an Open Source Solution. *J. Chem. Inf. Model.* **2011**, *51*, 739–753. doi:10.1021/ci100384d. [[CrossRef](#)]
18. ChemAxon. *Marvin 5.9.3*; ChemAxon: Budapest, Hungary, 2012.
19. Pescitelli, G.; Bari, L.D.; Berova, N. Application of electronic circular dichroism in the study of supramolecular systems. *Chem. Soc. Rev.* **2014**, *43*, 5211–5233. doi:10.1039/c4cs00104d. [[CrossRef](#)]
20. Calogero, G.; Citro, I.; Crupi, C.; Marco, G.D. Absorption spectra and photovoltaic characterization of chlorophyllins as sensitizers for dye-sensitized solar cells. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2014**, *132*, 477–484. doi:10.1016/j.saa.2014.04.196. [[CrossRef](#)]
21. Fang, H.; Xu, B.; Li, X.; Kuhn, D.L.; Zachary, Z.; Tian, G.; Chen, V.; Chu, R.; DeLacy, B.G.; Rao, Y.; et al. Effects of Molecular Structure and Solvent Polarity on Adsorption of Carboxylic Anchoring Dyes onto TiO₂ Particles in Aprotic Solvents. *Langmuir* **2017**, *33*, 7036–7042. doi:10.1021/acs.langmuir.7b01442. [[CrossRef](#)]
22. Gherzi, D.; Singh, M. molBLOCKS: decomposing small molecule sets and uncovering enriched fragments. *Bioinformatics* **2014**, *30*, 2081–2083. doi:10.1093/bioinformatics/btu173. [[CrossRef](#)] [[PubMed](#)]
23. Lewell, X.Q.; Judd, D.B.; Watson, S.P.; Hann, M.M. RECAP—Retrosynthetic Combinatorial Analysis Procedure: A Powerful New Technique for Identifying Privileged Molecular Fragments with Useful Applications in Combinatorial Chemistry. *J. Chem. Inf. Model.* **1998**, *38*, 511–522. doi:10.1021/ci970429i. [[CrossRef](#)]
24. Hagfeldt, A.; Boschloo, G.; Sun, L.; Kloo, L.; Pettersson, H. Dye-Sensitized Solar Cells. *Chem. Rev.* **2010**, *110*, 6595–6663. doi:10.1021/cr900356p. [[CrossRef](#)] [[PubMed](#)]
25. Sarker, A.K.; Kang, M.G.; Hong, J.D. A near-infrared dye for dye-sensitized solar cell: Catecholate-functionalized zinc phthalocyanine. *Dyes Pigments* **2012**, *92*, 1160–1165. [[CrossRef](#)]
26. Jiang, C.; Jin, X.; Dong, Y.; Chen, M. Kekule.js: An Open Source JavaScript Chemoinformatics Toolkit. *J. Chem. Inf. Model.* **2016**, *56*, 1132–1138. doi:10.1021/acs.jcim.6b00167. [[CrossRef](#)] [[PubMed](#)]
27. Bienfait, B.; Ertl, P. JSME: a free molecule editor in JavaScript. *J. Cheminform.* **2013**, *5*. doi:10.1186/1758-2946-5-24. [[CrossRef](#)] [[PubMed](#)]
28. Probst, D.; Reymond, J.L. SmilesDrawer: Parsing and Drawing SMILES-Encoded Molecular Structures Using Client-Side JavaScript. *J. Chem. Inf. Model.* **2018**, *58*, 1–7. doi:10.1021/acs.jcim.7b00425. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).