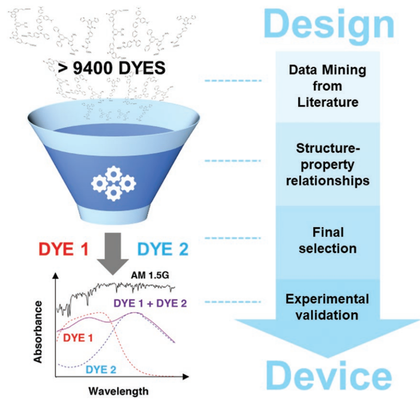
Q1

xxxx



A design-to-device study, based on algorithmic encodings of structure–property relationships, is used to identify new materials with panchromatic optical absorption. 9431 dyes are mined from literature and optimally paired together to afford co-sensitizing dyes with complementary optical absorption properties. Promising combinations are experimentally verified in dye-sensitized solar cells and novel methods for characterizing dye aggregation in co-sensitized devices are presented.

Design-to-Device Approach Affords Panchromatic Co-Sensitized Solar Cells

Q2 Christopher B. Cooper, Edward J. Beard, Álvaro Vazquez-Mayagoitia, Liliana Stan, Gavin B. G. Stenning, Daniel W. Nye, Julian A. Vigil, Tina Tomar, Jingwen Jia, Govardhana Babu Bodedla, Song Chen, Lucía Gallego Aibar, Santiago Franco Ontaneda, Antonio Carella, K. R. Justin Thomas, Song Xue, Xunjin Zhu, and Jacqueline M. Cole*

Data-driven materials discovery has become increasingly important in identifying materials that exhibit specific, desirable properties from a vast chemical search space. Synergic prediction and experimental validation are needed to accelerate scientific advances related to critical societal applications. A design-to-device study that uses high-throughput screens with algorithmic encodings of structure–property relationships is reported to identify new materials with panchromatic optical absorption, whose photovoltaic device applications are then experimentally verified. The data-mining methods source 9431 dye candidates, which are auto-generated from the literature using a custom text-mining tool. These candidates are sifted via a data-mining workflow that is tailored to identify optimal combinations of organic dyes that have complementary optical absorption properties such that they can harvest all available sunlight when acting as co-sensitizers for dye-sensitized solar cells (DSSCs). Six promising dye combinations are shortlisted for device testing, whereupon one dye combination yields co-sensitized DSSCs with power conversion efficiencies comparable to those of the high-performance, organometallic dye, N719. These results demonstrate how data-driven molecular engineering can accelerate materials discovery for panchromatic photovoltaic or other applications.

1. Introduction

Data-driven materials discovery[1,2] allows researchers to mine a vast chemical search space and identify materials that exhibit specific, desirable properties. These high-throughput, automatic approaches have accelerated scientific discovery in important research areas such as photovoltaics, water splitting, and gas capture.[3–9] This paper presents and utilizes a materials discovery approach to predict and then experimentally realize panchromatic solar cells, a factor critical to photovoltaic performance.[10–12]

The approach exploits co-sensitization, in the field of dye-sensitized solar cells (DSSCs), which offers a promising means to achieve the desired panchromatic solar cells for a variety of applications. DSSCs can exhibit efficiencies as high as 28.9% in ambient lighting, outperforming GaAs devices.[13] Q5
Their transparency makes DSSCs

C. B. Cooper, E. J. Beard, Dr. J. M. Cole
Cavendish Laboratory
Department of Physics
University of Cambridge
Q3 J. J. Thomson Avenue, Cambridge CB3 0HE, UK
Q4 E-mail: jmc61@cam.ac.uk

C. B. Cooper, Dr. J. M. Cole
Department of Chemical Engineering and Biotechnology
University of Cambridge
West Cambridge Site
Philippa Fawcett Drive
Cambridge CB3 0AS, UK

C. B. Cooper, Dr. J. M. Cole
Research Complex at Harwell
Rutherford Appleton Laboratory
Harwell Oxford
Didcot, Oxon OX11 0FA, UK

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aenm.201802820.

DOI: 10.1002/aenm.201802820

Dr. Á. Vazquez-Mayagoitia, Dr. L. Stan, Dr. J. M. Cole
Argonne National Laboratory
9700 South Cass Avenue, Argonne, IL 60439, USA

Dr. G. B. G. Stenning, Dr. D. W. Nye, Dr. J. M. Cole
ISIS Neutron and Muon Facility
STFC Rutherford Appleton Laboratory
Harwell Science and Innovation Campus
Didcot OX11 0QX, UK

J. A. Vigil
Department of Chemistry
University of Cambridge
Lensfield Rd, Cambridge CB2 1EW, UK

T. Tomar, Dr. K. R. J. Thomas
Organic Materials Laboratory
Department of Chemistry
Indian Institute of Technology Roorkee
Roorkee 247667, India

J. Jia, Dr. S. Xue
Tianjin Key Laboratory of Organic Solar Cells and Photochemical Conversion
School of Chemistry & Chemical Engineering
Tianjin University of Technology
391 Binshui Xidao, Xiqing District, Tianjin 300384, P. R. China

optimal devices for solar windows,[14] while their ability to be fabricated on flexible substrates or as fibers enables passive energy harvesting in wearable devices and textiles.[15–18] DSSCs can also be manufactured at low cost using scalable techniques such as roll-to-roll processing,[19] inkjet printing,[20] and ultrafast sensitization,[21,22] which are necessary to reach competitive price-to-performance ratios.

Thus far, co-sensitization has helped afford the world-record DSSC efficiency of over 14% under full illumination;[23] however, the lack of a rational, automated method to select combinations of dyes from a large database of light-harvesting chromophores limits further progress. Despite numerous computational studies complementing experimental work on singly sensitized DSSCs, only a few studies have attempted to computationally predict and analyze co-sensitized DSSCs;[24–27] and up until now, no study has offered a full design-to-device materials discovery approach for co-sensitized DSSCs. This paper presents and validates such a method.

A database of dye candidates was compiled via automated text-mining of published journal articles. This custom-made database was then mined using high-throughput screening methods which employed algorithmic encodings of structure–property relationships to identify five promising organic dyes that could act together as co-sensitizers, with six possible co-sensitization pairings. The predicted dyes, which had never been co-sensitized, were then synthesized and characterized experimentally. The dye combination that performed best within a DSSC device exhibited a power conversion efficiency that is comparable to that of the high-performance, organometallic dye, N719. Furthermore, surface characterization via atomic force microscopy (AFM) and X-ray reflectometry (XRR) provided, for the first time, a quantitative analysis of how co-sensitization affected dye aggregation and adsorption onto TiO_2. These results offer a promising example of how a materials discovery approach can accelerate and improve scientific advances related to panchromatic solar cells or other applications.

2. Results and Discussion

2.1. Materials Prediction of Co-Sensitizers

Figure 1 provides a schematic of the computational workflow that predicts optimal dye combinations for co-sensitization. First, we auto-generated a database of 9431 dye candidates (including their chemical structure, maximum absorption wavelengths, and molar extinction coefficients) from academic literature, using the text-mining software ChemDataExtractor.[28] Initial screens then removed small molecules, organometallic dyes, and chemicals not absorbing in the solar spectrum, leaving 3053 organic dyes remaining.

Second, we screened dyes based on two key structure–property relationships: the presence of a carboxylic acid group and a sufficiently large molecular dipole moment. The former ensures that the selected dyes contain a high-performance DSSC anchoring group[29] which enables them to effectively adsorb onto TiO_2 surfaces to create working electrodes. The latter is required for effective intramolecular charge transfer after photoexcitation. After selecting only dyes with carboxylic acid groups via substructure searching and eliminating dyes with a molecular dipole moment less than 5 Debye,[30] 309 dyes remained in the shortlist.

Next, we employed an algorithm to predict dye combinations for co-sensitization based on their optical absorption properties. We provide an overview here with full details given in the Supporting Information. Using the maximum absorption wavelengths and extinction coefficients gathered by ChemDataExtractor,[28] we ranked each potential dye combination using a quality score. Algorithm metrics producing a high quality score comprised a large overlap factor, absorption fraction, and relative change. These ensured that the dye combination i) did not have significant optical absorption overlap between dyes, ii) exhibited panchromatic absorbance, and iii) improved significantly from the addition of each dye. This yielded a shortlist of 33 dyes.

We then checked the highest-occupied molecular orbital–lowest-unoccupied molecular orbital (HOMO–LUMO) energy levels of the 33 shortlisted dyes using Density Functional Theory (6311G** basis set and B3LYP functional) to confirm that the LUMO energy levels were greater than those of the conduction band edge of anatase TiO_2 (−3.74 eV vs vacuum)[31] and that the HOMO energy levels were below the redox potential of I^-/I_3^- (−4.85 eV vs vacuum).[32] These are necessary energetic properties for device integration into a standard DSSC, though these checks could be modified for integration with other semiconductors or redox couples. This screen reduced the shortlist to 29 dyes.

From here, we manually evaluated each dye and considered practical constraints such as ease of synthesis or availability. This afforded a set of five dyes for experimental validation: C1,[33] 8c,[34] XS6,[35] 15,[36] and H3.[37] **Figure 2** provides the 2D and 3D chemical structures of the dyes, with molecular dimensions annotated as a reference for the surface characterization work discussed later. The maximum optical absorption wavelengths and corresponding molar extinction coefficients were 457 nm (1.00×10^5 L mol^{-1} cm^{-1}), 414 nm (3.27×10^4 L mol^{-1} cm^{-1}), 432 nm (1.25×10^5 L mol^{-1} cm^{-1}), 573 nm (3.36×10^4 L mol^{-1} cm^{-1}), and 585 nm (2.87×10^4 L mol^{-1} cm^{-1}), for dyes C1, 8c, XS6, 15, and H3, respectively.[33–37] Co-sensitizing any of the first three dyes (C1, 8c, and XS6) with either of the last two dyes (15 and H3) should create DSSCs with broad optical absorbance.

2.2. Experimental Validation of Predicted Dyes

We experimentally validated and characterized these six potential co-sensitizations using UV–vis absorption spectroscopy and

Dr. G. B. Bodedla, Dr. S. Chen, Dr. X. Zhu
Department of Chemistry and Institute of Molecular Functional Materials
Hong Kong Baptist University
Kowloon Tong, Hong Kong, P. R. China
L. G. Aibar, Dr. S. F. Ontaneda
Departamento de Química Orgánica
ICMA
Universidad de Zaragoza-CSIC
Zaragoza 50009, Spain
Dr. A. Carella
Department of Chemical Sciences
University of Naples Federico II
Via Cintia 21, Naples 80126, Italy

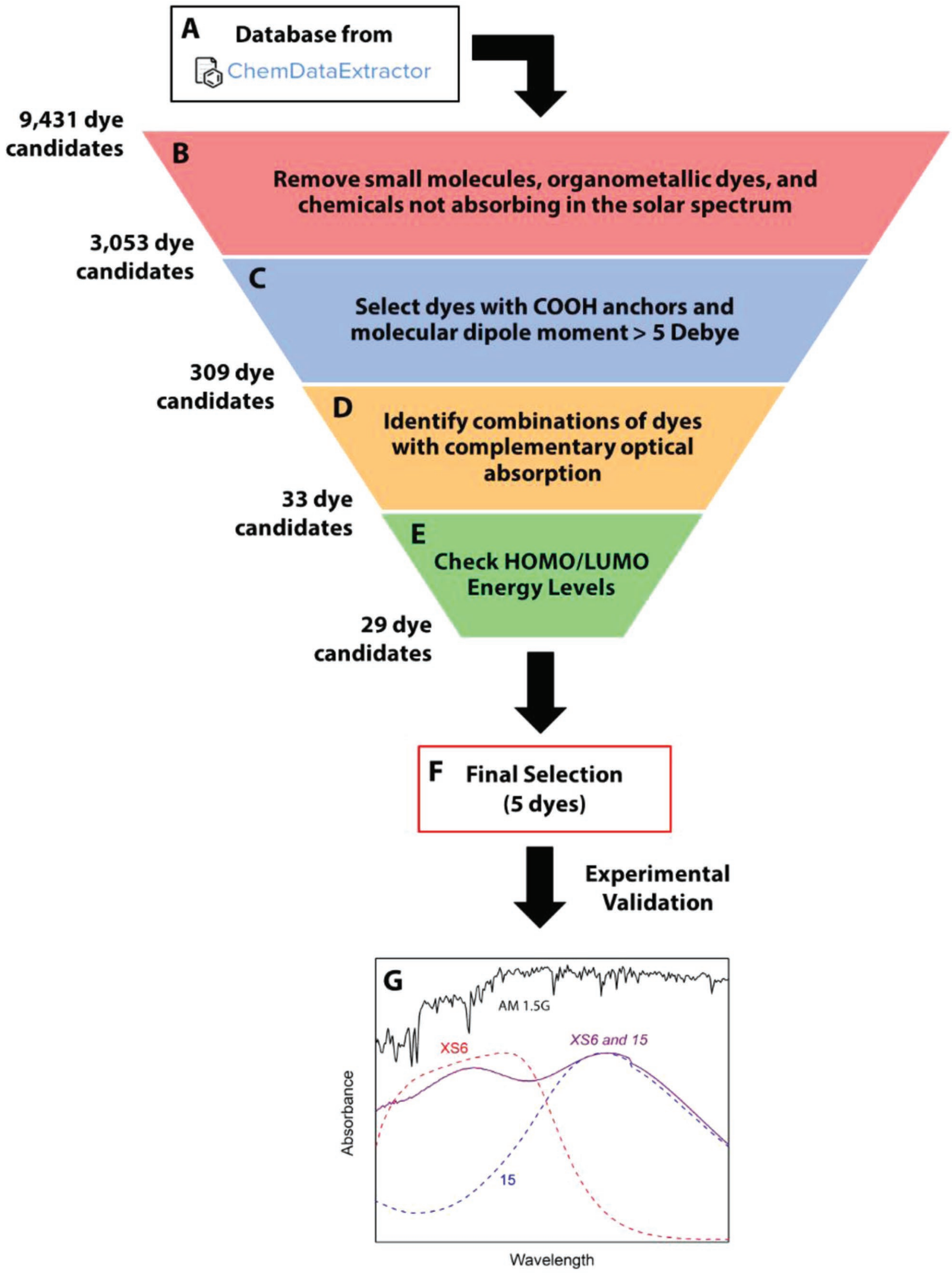


Figure 1. Design-to-device approach to create panchromatic solar cells via co-sensitization. A) A starting database of 9431 dye candidates from the academic literature is auto-generated using the text-mining tool, ChemDataExtractor. B) Initial screens remove small molecules, organometallic dyes, and chemicals not absorbing in the solar spectrum (350–1000 nm) to reduce the number of dyes to 3053. C) Substructure searching and semi-empirical calculations are used to select dyes with a carboxylic acid anchor and a molecular dipole moment > 5 Debye. D) A novel algorithm predicts optimal combinations of dyes with complementary optical absorption spectra and high molar extinction coefficients, narrowing the shortlist to 33 dyes. E) HOMO and LUMO energy levels of each dye are checked using DFT to ensure proper integration into a DSSC. F) A final set of five dyes is selected for experimental verification based on practical constraints such as ease of synthesis and availability. G) Experimental validation illustrates the benefits of co-sensitization and shows how the best performing combination of two dyes with complementary optical absorption spectra, XS6 (red) and 15 (blue), affords a co-sensitized DSSC, *XS6 and 15* (purple), with broad absorbance. The AM 1.5G solar emission spectrum (black) is offset above for reference.

photovoltaic device testing. **Figure 3**A gives the optical absorption spectrum of each individual dye in dichloromethane (DCM). The dyes absorb throughout the visible spectrum, with C1, 8c, and XS6 absorbing primarily in the 300–500 nm range; 15 absorbing primarily in the 500–700 nm range; and H3 exhibiting broad absorbance with a gap between 425 and 525 nm. Figure 3B gives the optical absorption spectra of each dye adsorbed onto TiO_2. Both C1 and 8c exhibit wider optical absorption spectra compared to their absorbance in DCM while 15 and H3 display a 52 and 26 nm blue shift in maximum absorbance, respectively.

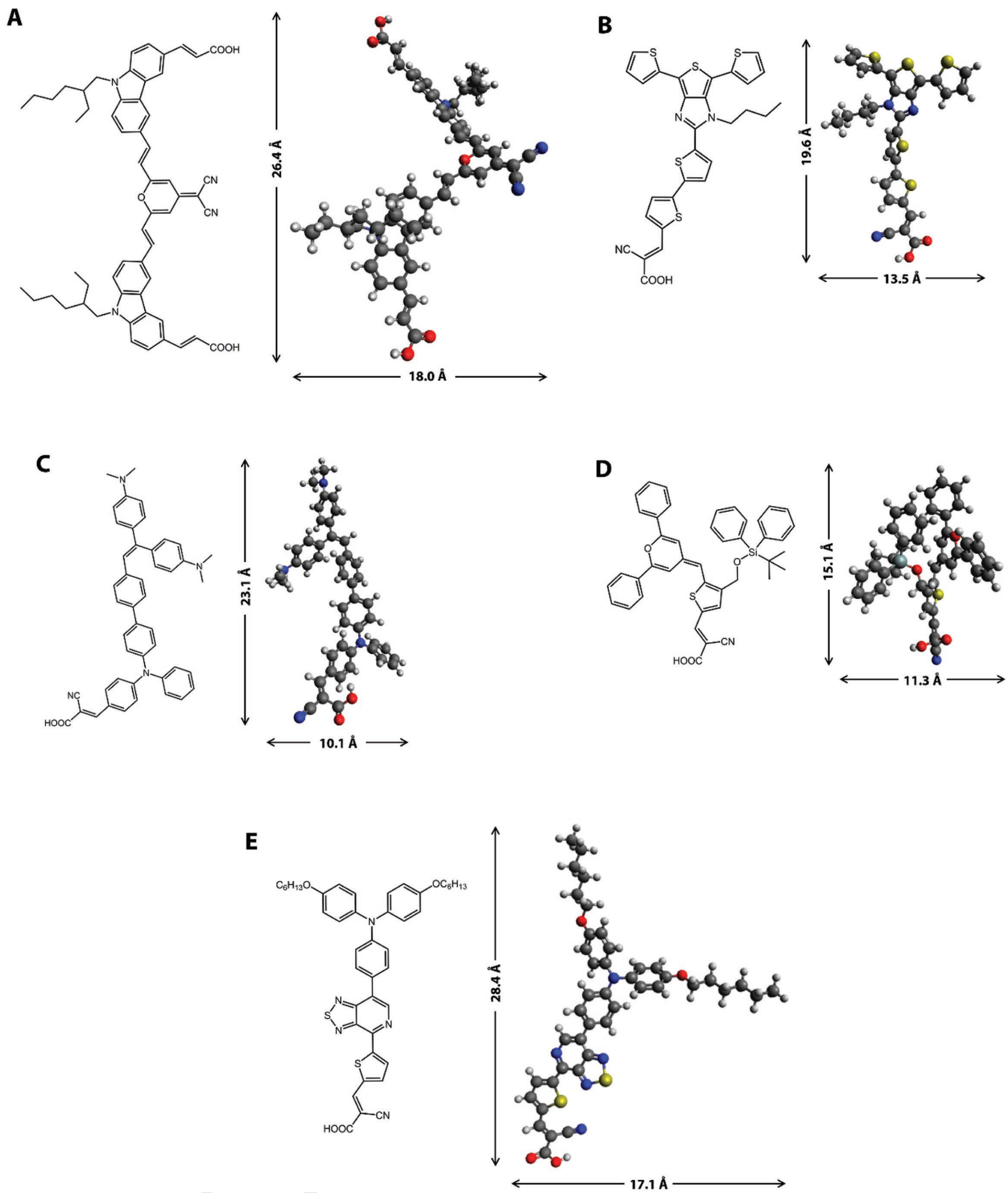


Figure 2. Chemical structures of predicted dyes. The 2D and 3D chemical structures of dyes C1 A), 8c B), XS6 C), 15 D), and H3 E) with annotated molecular length and width marked for each dye. Molecular length is defined as the largest atom-to-atom distance projected from either oxygen in the carboxylate anchor of each dye. The molecular width is defined as the largest atom-to-atom distance perpendicular to the molecular length. 3D structures are optimized with PM7 semi-empirical calculations.

For each dye combination, we identified a sequential and cocktail method that afforded co-sensitized working electrodes (WE) with panchromatic optical absorption (Table S1, Supporting Information). Samples fabricated via the sequential and cocktail method are referred to as "*Dye 1 then Dye 2*" and "*Dye 1 and Dye 2,*" respectively. For simplicity, we use these sample names throughout the letter to refer to WEs sensitized under the specific conditions described in Table S1 in the Supporting Information. Compared to the spectra of the individual dyes on TiO_2, the co-sensitized WEs exhibit broad absorbance throughout the visible region (400–700 nm), indicating that adsorption of both dyes onto TiO_2 has been achieved (Figure S1, Supporting Information). We found that C1 significantly desorbs 15, and thus, we adjusted each sensitization method to achieve adequate adsorption of both dyes (Figure S2, Supporting Information).

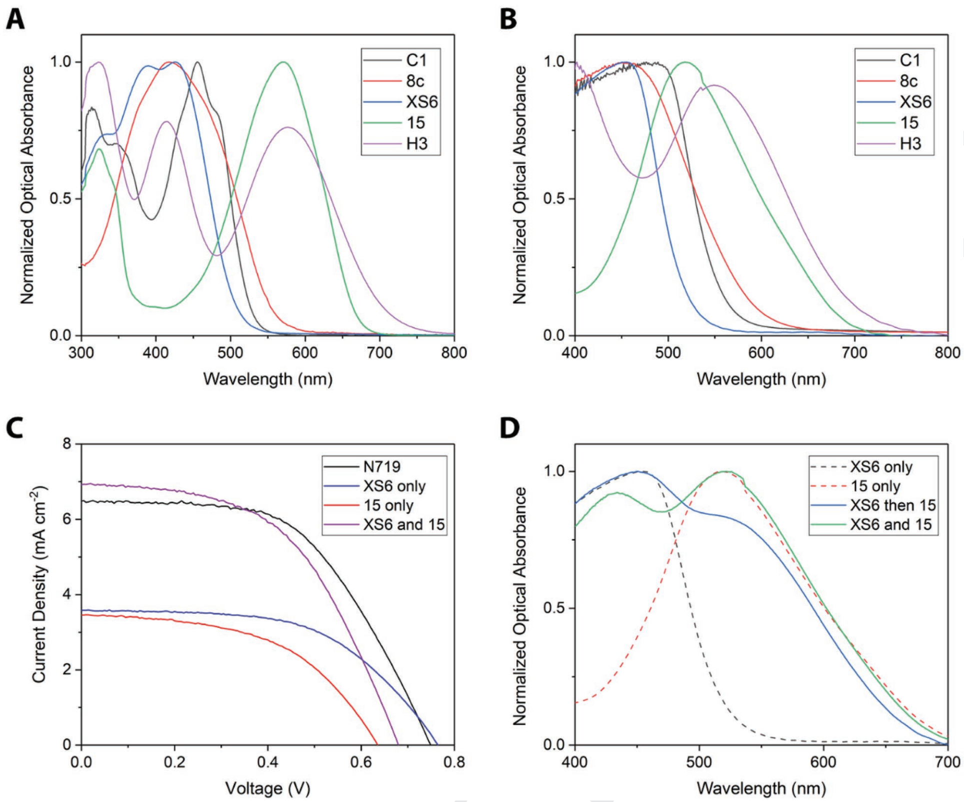


Figure 3. Optical absorption and photovoltaic performance. Optical absorption of the predicted dyes in DCM solution A) and adsorbed onto TiO_2 B). C) *J–V* curves of DSSCs sensitized with XS6 only (blue), 15 only (red), and both XS6 and 15 (purple) compared to an N719 reference (black). The *XS6 and 15* co-sensitized DSSC demonstrates dramatic improvements compared to the singly sensitized DSSC with performance comparable to the N719 reference. D) Optical absorption of co-sensitized WEs *XS6 then 15* and *XS6 and 15* compared to their XS6 and 15 singly sensitized counterparts. Both co-sensitized WEs exhibit broad absorbance compared to the singly sensitized WEs, with that of *XS6 and 15* having a higher concentration ratio of 15:XS6 adsorbed onto the TiO_2 surface as indicated by the shift in maximum absorbance.

We then tested the photovoltaic performance of singly sensitized and co-sensitized DSSCs compared to a reference sample sensitized with the organometallic N719 dye. Reporting with the η_{dye}: η_{N719} ratio method permits effective comparison between power conversion efficiencies, η, published in the literature under a range of experimental conditions. This method has already been adopted in over 250 journal articles.[38] Table S2 in the Supporting Information provides the photovoltaic device performance for each sample, averaged across three different DSSCs. All measured *J–V* curves are given in Figure S3 in the Supporting Information.

Figure 3C presents the *J–V* curve for the best-performing co-sensitization, *XS6 and 15*, which exhibited a 38% increase in η compared to the corresponding singly sensitized DSSCs. Moreover, its η_{dye}: η_{N719} ratio of 0.92 demonstrates performance comparable to that of the high-performance, organometallic N719 dye. Similarly, *XS6 then 15* increased η by 23%, obtaining a promising η_{dye}: η_{N719} ratio of 0.82. Both *XS6 and 15* and *XS6 then 15* exhibit high open circuit voltages (V_{oc}) of 700 and 685 mV, respectively, indicating that electron recombination has been minimized. *XS6 and 15* achieves a higher short-circuit current density (J_{sc}) than *XS6 then 15*, 6.5 mA cm^{-2} compared to 5.5 mA cm^{-2}. Comparing the UV–vis absorption spectra of *XS6 then 15* and *XS6 and 15* (Figure 3D) suggests that the increase in J_{sc} arises from the adsorption of more molecules of 15 onto TiO_2 achieved via the cocktail approach.

C1 then 15 and *C1 and 15* showed a modest, but not statistically significant, gain in η from co-sensitization, 6% and 7%, respectively, with η_{dye}: η_{N719} ratios of 0.54. Both 8c and H3 afforded dramatically lower J_{sc} values and slightly lower V_{oc} values than the other dyes when singly sensitized, leading to deleterious effects whenever they were co-sensitized. Calculated HOMO and LUMO energy levels for these dyes show that 8c and H3 have the lowest predicted LUMO energy levels and highest predicted HOMO energy levels. These smaller bandgaps imply lower driving forces for electron injection and dye regeneration, possibly explaining their poor performance.

2.3. Surface Characterization of Co-Sensitized DSSCs

To better understand the molecular origins of these photovoltaic results, we characterized the surface structure of singly sensitized and co-sensitized WEs, using AFM and XRR. While previous studies have used either AFM or XRR to determine dye aggregation effects, dye coverage, inter-dye spacing, and dye-layer thicknesses in singly sensitized DSSCs,[39–41] here we present the first study of AFM or XRR

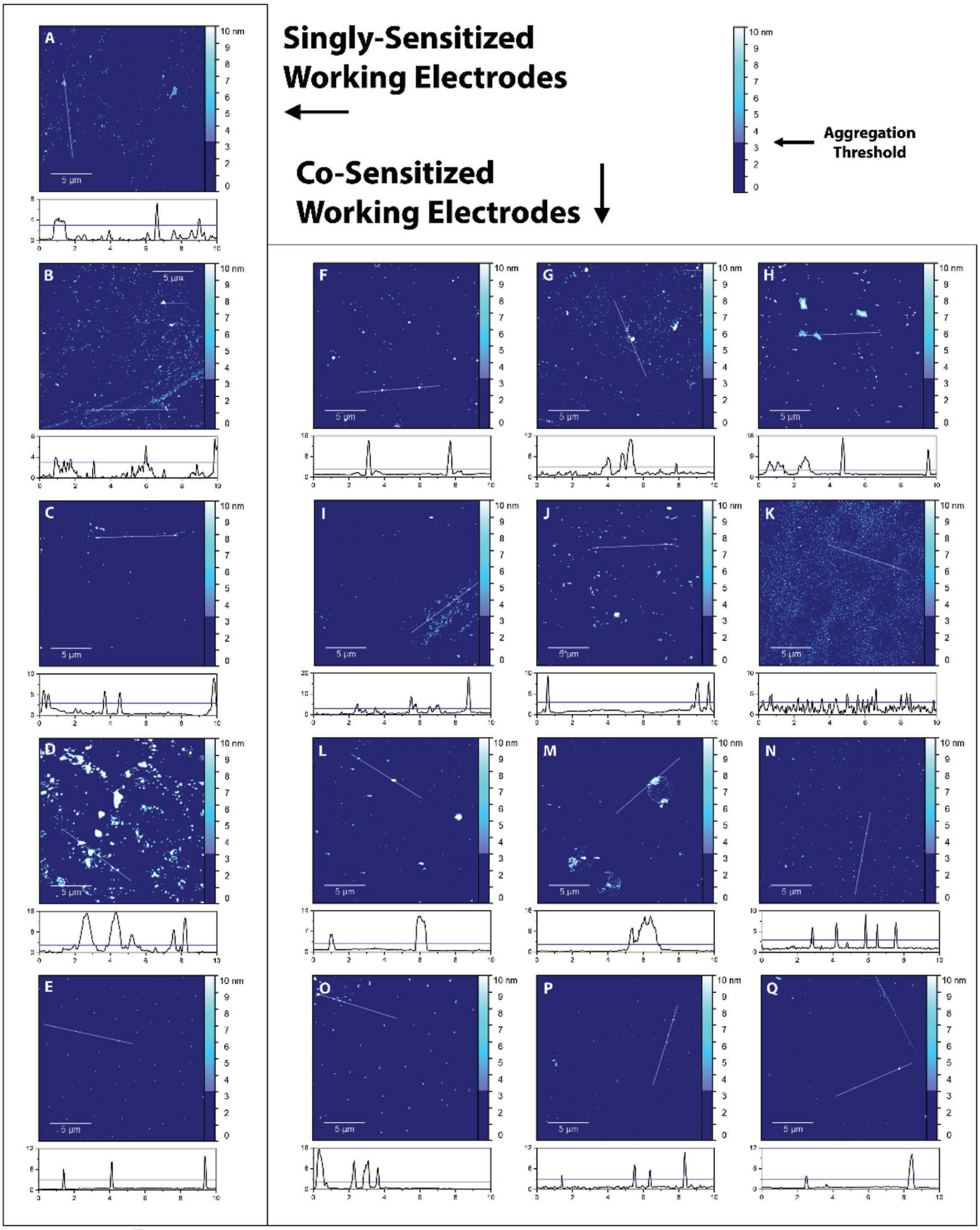


Figure 4. Surface characterization via AFM. Representative 20 μm × 20 μm AFM images for dyes C1 A), 8c B), XS6 C), 15 D), and H3 E). Representative 20 μm × 20 μm AFM images for dye combinations C1 then 15 F), C1 and 15 G), H3 then C1 H), C1 and H3 I), 8c then 15 J), 8c and 15 K), H3 then 8c L), 8c and H3 M), XS6 then 15 N), XS6 and 15 O), XS6 then H3 P), and XS6 and H3 Q). The color bar is solid below the 3 nm aggregation threshold. Below each AFM image is a randomly selected height profile (black) that corresponds to the surface features highlighted by the white trace on the AFM image (from left to right). A blue line showing the aggregation threshold of 3 nm is also included as a visual aid to see features included in the data analysis. Units on the abscissa and ordinate of the height profile are μm and nm, respectively. All AFM images are provided in the Supporting Information.

on co-sensitized WEs and provide a quantitative analysis of how co-sensitization affects dye aggregation and adsorption onto TiO_2.

We selected an AFM base height of 3 nm as the dye aggregation threshold, since this is greater than the molecular length of any dye (Figure 2). Thereby, any continuous areas greater

Table 1. Surface characterization of singly sensitized and co-sensitized working electrodes.

Sample name	AFM parameters				XRR parameters			
	Mean height [nm]	Max height [nm]	Aggregate coverage [%]	Number of aggregates [µm^{-2}]	Dye layer thickness [Å]	SLD_{dye} [× 10^{-6} Å^{-2}]	Surface roughness [Å]	Surface coverage [%]
Singly sensitized working electrodes								
C1 only	5 ± 1	7 ± 2	3 ± 6	2 ± 3	43.5 ± 0.9	6.6 ± 0.5	5.6 ± 0.7	55 ± 4
8c only	5 ± 1	6 ± 2	3 ± 2	3 ± 2	26.6 ± 0.9	5.1 ± 0.9	3.3 ± 0.8	39 ± 7
XS6 only	4.9 ± 0.4	6.0 ± 0.7	1.0 ± 0.1	0.3 ± 0.2	23.6 ± 0.5	8.7 ± 0.4	3.7 ± 0.5	73 ± 3
H3 only	9 ± 1	15 ± 3	0.3 ± 0.1	0.18 ± 0.05	27 ± 1	6.7 ± 0.5	3.7 ± 0.5	55 ± 4
15 only	8 ± 2	15 ± 3	7 ± 2	1.1 ± 0.4	24.3 ± 0.3	7.8 ± 0.4	2.7 ± 0.3	62 ± 3
Co-sensitized working electrodes								
C1 then 15	6 ± 2	10 ± 3	1.3 ± 0.5	0.7 ± 0.2	33.7 ± 0.5	5.9 ± 0.7	3.1 ± 0.6	49 ± 6
C1 and 15	7 ± 2	12 ± 4	2.0 ± 0.5	0.9 ± 0.6	21.5 ± 0.8	6.3 ± 0.9	3.8 ± 0.7	52 ± 7
H3 then C1	8 ± 2	16 ± 4	3 ± 3	0.4 ± 0.2	42 ± 1	6.0 ± 0.6	5.2 ± 0.7	49 ± 5
C1 and H3	5 ± 1	8 ± 3	2 ± 1	2 ± 2	25.4 ± 0.4	8.5 ± 0.4	3.0 ± 0.5	69 ± 3
8c then 15	6 ± 1	9 ± 2	1.1 ± 0.2	0.7 ± 0.6	30.9 ± 0.4	6.9 ± 0.4	3.9 ± 0.6	54 ± 3
8c and 15	4.6 ± 0.3	5.8 ± 0.4	12 ± 9	16 ± 5	31 ± 2	5.7 ± 0.5	7 ± 2	45 ± 4
H3 then 8c	5.5 ± 0.7	8 ± 1	3 ± 2	1 ± 1	37.2 ± 0.2	9.0 ± 0.7	2.9 ± 0.4	70 ± 5
8c and H3	5.2 ± 0.7	7 ± 2	2 ± 2	1 ± 1	27.5 ± 0.4	8.0 ± 0.4	3.3 ± 0.6	63 ± 3
XS6 then 15	6 ± 1	8 ± 2	0.7 ± 0.3	0.8 ± 0.3	18.8 ± 0.3	8.7 ± 0.5	3.6 ± 0.4	72 ± 4
XS6 and 15	7.8 ± 0.7	11 ± 1	0.3 ± 0.1	0.24 ± 0.09	18.6 ± 0.3	8.8 ± 0.5	3.4 ± 0.4	73 ± 4
XS6 then H3	5.5 ± 0.7	7.6 ± 0.8	0.3 ± 0.1	0.25 ± 0.04	21.0 ± 0.3	9.6 ± 0.5	4.1 ± 0.4	79 ± 4
XS6 and H3	5.3 ± 0.8	7 ± 1	0.3 ± 0.1	0.2 ± 0.1	21.6 ± 0.6	8.7 ± 0.6	4.0 ± 0.5	71 ± 5

in height than this threshold were classified as aggregates in the AFM images. For each sample, we obtained five distinct 20 µm × 20 µm AFM images (see representative images in **Figure 4**) and characterized the aggregates based on mean height, max height, coverage, and number of aggregates (**Table 1**).

Co-sensitized WEs, *XS6 then 15, XS6 and 15, XS6 then H3*, and *XS6 and H3* exhibited the lowest amount of dye aggregation with low aggregate coverage (0.3–0.7%) and low total number of aggregates (0.2–0.8 µm^{-1}). *XS6 and 15* specifically exhibited aggregate coverage of 0.3% and 0.2 aggregates per micrometer, both an order of magnitude lower than the aggregation observed in 15 only. The minimal aggregation exhibited by *XS6 then 15* and *XS6 and 15* suggests that a dye monolayer has formed on TiO_2 and partially explains their optimal photovoltaic performance. Similarly, both *C1 then 15* and *C1 and 15* exhibited less aggregation than their singly sensitized counterparts.

Overall, co-sensitized WEs show reduced aggregation compared to their singly sensitized counterparts for seven out of 12 samples. C1, 8c, and 15 singly sensitized WEs all show significant aggregation whose coverages are 3%, 3%, and 7%, respectively. C1 and 8c display many small aggregates manifested by the low mean and max heights (5–7 nm) and high number of aggregates (2–3 µm^{-2}); this indicates that the dyes aggregate longitudinally (i.e., side-by-side). In contrast, 15 exhibits relatively large aggregates with higher mean height (8 nm) and max height (15 nm) but a lower number of aggregates (1.1 µm^{-2}), suggesting a combination of both longitudinal and lateral (i.e., stacked) dye aggregation. Both XS6 and H3 show minimal aggregation, with aggregate coverages of 1% and 0.3%, respectively, and a low total number of aggregates (0.2–0.3 µm^{-2}). For XS6, this minimal aggregation could arise from its twisted π-conjugation, while for H3 it could result from the bulky hexyloxy chains. Both of these properties have reduced the aggregation of other dye molecules.[42]

Next, we employed XRR to obtain structural information about the adsorbed dye layer. Fitting data collected from each WE to a model based on calculated molecular dimensions and scattering length densities (SLDs) (Table S3, Supporting Information) revealed estimates of the dye-layer thickness, SLD_{dye}, surface roughness, and surface coverage (Table 1). Additional fitted parameters, all raw data and fitted models, and all calculations are given in the Supporting Information.

Co-sensitized WEs *XS6 then 15* and *XS6 and 15* exhibit low dye-layer thicknesses (19 Å) together with high surface coverages (above 70%). Consistent with the minimal aggregation seen in the corresponding AFM images, these results strongly suggest the formation of a tightly packed monolayer on the TiO_2 surface, corroborated by the high V_{oc} of both dye combinations (>685 mV). Relatively poor surface coverage (49% and 52%) is observed in *C1 then 15* and *C1 and 15*, which are the other DSSCs to prospect any gain in η from co-sensitization (Table S2, Supporting Information).

In common with the AFM results, XRR models for *XS6 then H3* and *XS6 and H3* display some of the lowest dye-layer thicknesses (21–22 Å) and highest surface coverages (>70%), despite their poor photovoltaic performance. Additionally, singly sensitized XS6 and H3 working electrodes have thicknesses near their molecular lengths, indicating that they have formed monolayers on the TiO_2 surface. We calculated the intermolecular spacing in the XS6 and H3 dye monolayers to be 3.7 and 2.8 Å, respectively, implying each present a tightly packed monolayer that prevents I_3^- (molecular length of ≈5.28 Å) from reaching the TiO_2 surface to cause electron recombination issues (see Supporting Information).[40] The high surface coverage (73%) and V_{oc} (730 mV) observed for XS6, the highest of the singly sensitized dyes in both cases, corroborates the idea of a packed monolayer. The inferior photovoltaic performance of H3, despite its minimal aggregation in AFM images (0.3%), low intermolecular spacing (2.8 Å), and high surface coverage (62%), suggests poor electron injection by H3 into TiO_2.

3. Conclusions

In this work, we have presented and experimentally validated a design-to-device approach that employs structure–property relationships in a computational workflow to achieve panchromatic solar cells. The results, especially for *XS6 then 15* and *XS6 and 15*, offer a promising example of accelerated materials discovery for photovoltaics, given that they yield power conversion efficiencies which are comparable to that of N719, the high-performance organometallic dye that acts as the industry standard for DSSCs. This accomplishment is despite having deliberately restricted our search to organic dyes that historically produce lower DSSC efficiencies but are environmentally superior. This demonstrates the power of our approach.

Our work, thus, offers a rare example of a full cycle of data-driven materials discovery, which is difficult to achieve owing to a dearth in demonstrable methods. Moreover, our methods are distinguished by their success, showing that co-sensitization of DSSCs can be tailored rationally to afford solar-cell devices that perform to world-recognized photovoltaic standards.

4. Experimental Section

Assembly of the Parent Database: The text-mining software tool, ChemDataExtractor,[28] was used to auto-generate a custom database of dye candidates for this project by sourcing matched quantities of chemical, optical absorption properties from the academic literature. Each data field comprised the chemical structure of a molecule in simplified molecular-input line-entry system (SMILES) format[43] along with its optical absorption peak wavelength, λ_{max}, and molar extinction coefficients, ε. SMILES were resolved from their chemical names using OPSIN[44] while OpenBabel[45] was used to read the SMILES structure of each chemical molecule and check for duplicates. After narrowing the shortlist to 309 dyes (i.e., prior to implementing the dye matching algorithm), manual verification of the maximum absorption peak wavelength and molar extinction coefficient for each shortlisted dye was completed and erroneous data were corrected. Data auto-extraction employed the supercomputing resources at the Argonne Leadership Computing Facility, USA.

Initial Screens and Identification of Suitable Anchoring Group: The RDKit Library[46] in Python was utilized for basic filtering of the dye candidates (i.e., removal of small molecules or organometallic dyes). Molecules without a maximum absorption peak between 350 and 1000 nm were removed. RDKit was also used for substructure searching in which only dyes with an identified carboxylic acid group (COOH) in their structure were kept.

Molecular Dipole Moment Calculations: To accurately estimate the molecular dipole moment of each dye candidate in a computationally efficient manner, 3D coordinates were generated for each dye candidate from its corresponding SMILES structure via a weighted rotor search (as defined in OpenBabel) to identify five low-energy conformers. The geometries of the five selected conformers were then further optimized using PM7 semi-empirical geometry optimization executed in MOPAC.[47] PM7 was selected due to its previous use with organic molecules.[48] The molecular dipole moment of each dye was taken from the PM7 results of the lowest energy conformer.

HOMO/LUMO Energy Level Calculations: HOMO and LUMO energy levels were estimated for each remaining dye candidate using a single point calculation with Density Functional Theory (DFT) with the 6311G** basis set and B3LYP functional on the previously PM7-optimized geometry of the lowest energy conformer to reduce computational cost.[49–51] All DFT calculations were completed using NWChem software via the supercomputing resources at the Argonne Leadership Computing Facility, USA.[52]

Dye Synthesis and Characterization: The research groups who originally made each dye synthesized the predicted dyes as a collaboration specifically for this project, according to their previously reported methods.[33–37] Reproducibility was verified for each dye by nuclear magnetic resonance (NMR) spectroscopy. ^1H NMR spectra were recorded on a Bruker 400 MHz DCH cryoprobe spectrometer at room temperature. Chemical shifts for ^1H spectra were referenced to residual signals from the deuterated solvent.

C1. ^1H NMR (DMSO-d_6, 400 MHz): δ/ppm = 12.27 (bs, 1H), 8.72 (s, 2H), 8.52 (s, 2H), 8.03–7.96 (m, 4H), 7.88–7.83 (m, 2H), 7.81 (s, 1H), 7.77 (s, 1H), 7.73–7.68 (m, 2H), 7.66–7.61 (m, 2H), 7.43 (bs, 1H), 7.39 (bs, 1H), 6.84–6.81 (m, 2H), 6.57 (d, J = 15.6 Hz, 2H), 4.37–4.29 (m, 4H), 2.05–1.97 (m, 2H), 1.40–1.12 (m, 16H), 0.87 (t, J = 7.2 Hz, 6H), 0.78 (t, H = 7.2 Hz, 6H).

8C. ^1H NMR (DMSO-d_6, 400 MHz): δ/ppm = 8.43 (s, 1H), 7.95 (d, J = 3.9 Hz, 1H), 7.75–7.70 (m, 2H), 7.66–7.63 (m, 1H), 7.62–7.58 (m, 1H), 7.57–7.53 (m, 1H), 7.52–7.48 (m, 1H), 7.39–7.36 (m, 1H), 7.23–7.20 (m, 1H), 7.17–7.13 (m, 1H), 4.31 (m, 2H), 1.50–1.39 (m, 2H), 1.11–0.99 (m, 2H), 0.66 (t, J = 7.4 Hz, 3H).

XS6. ^1H NMR (DMSO-d_6, 400 MHz): δ/ppm = 8.00 (s, 1H), 7.86 (d, J = 8.8 Hz, 2H), 7.66 (m, 2H), 7.48–7.39 (m, 4H), 7.26–7.08 (m, 9H), 6.99–6.92 (m, 4H), 6.78 (s, 1H), 6.74–6.71 (m, 2H), 6.70–6.66 (m, 2H), 2.93 (s, 6H), 2.91 (s, 6H).

15. ^1H NMR (THF-d_8, 400 MHz): δ/ppm = 8.28 (s, 1H), 8.02–7.97 (m, 2H), 7.88–7.84 (m, 2H), 7.78 (s, 1H), 7.76–7.72 (m, 4H), 7.56–7.37 (m, 13H), 6.75 (bd, J = 2.0 Hz, 1H), 6.12 (s, 1H), 4.87 (s, 2H), 1.09 (s, 9H).

H3. ^1H NMR (DMSO-d_6, 400 MHz): δ/ppm = 8.81 (s, 1H), 8.65 (d, J = 4.1 Hz, 1H), 8.37 (s, 1H), 8.00 (d, J = 4.1 Hz, 1H), 7.98–7.94 (m, 2H), 7.13–7.08 (m, 4H), 6.97–6.92 (m, 4H), 6.90–6.86 (m, 2H), 3.95 (t, J = 6.5 Hz, 4H), 1.75–1.67 (m, 4H), 1.46–1.38 (m, 4H), 1.36–1.26 (m, 8H), 0.88 (t, J = 7.0 Hz, 6H).

UV–vis Absorption Spectroscopy: The optical absorption spectra of the fabricated WEs (see Supporting Information for fabrication details) and of the prepared dye solutions (in DCM, 3×10^{-5} M) were acquired using a Shimadzu UV-1800 Spectrophotometer. All solutions were tested in 10 mm pathlength quartz precision cells (SUPRASIL, Hellma Analytics).

Photovoltaic Performance Testing: The current–voltage characteristics of the singly sensitized and co-sensitized DSSCs (see Supporting Information for fabrication details) were measured with an Ivium CompactStat potentiostat under constant illumination by a Newport Oriel Xenon 150 W solar light simulator (100 mW cm^{-2}, AM1.5G and IR water filters, $\lambda < 400$ nm), calibrated with a Newport Optical power meter (Model 1916-R). Solar cells had an active area of 0.30 cm^2 and were masked with an 8 mm × 8 mm aperture. Linear scanning voltammetry was performed at room temperature in ambient air at 50 mV s^{-1} with a 5 s equilibrium time between forward and backward scans. No pre-conditioning of the devices was completed.

Using the measured J–V curves, the short-circuit current density (J_{sc}), open-circuit voltage (V_{oc}), and fill-factor (FF) were determined for each fabricated cell. The photovoltaic efficiency of the cell was then calculated by Q6

$$\eta = \frac{J_{sc}\, V_{oc}\, FF}{P_{in}} \tag{1}$$

where P_{in} is the power of incident-light radiation. Reported DSSC parameters were found by averaging across three individually tested cells. For co-sensitized DSSCs, the percentage change in efficiency was calculated by comparing each co-sensitized DSSC to the best performing singly sensitized DSSC of the dyes used. The *J–V* curves for all tests are given in Figure S3 in the Supporting Information.

Atomic Force Microscopy: The surfaces of singly sensitized and co-sensitized WEs (see Supporting Information for fabrication details) were imaged using a Bruker Dimension 3100 Atomic Force Microscope with a monolithic silicon AFM probe (Tap300-G, Budget Sensors) with a tip radius less than 10 nm, a resonance frequency of 300 kHz, and a force constant of 40 N m^{-1}. Tapping mode was used to produce five 20 μm × 20 μm images of different areas of each sample to ensure a representative measurement. All AFM images were processed using Gwyddion software.[53] All AFM images are given at the end of the Supporting Information.

X-Ray Reflectometry: A Rigaku SmartLab X-Ray Diffractometer equipped with a 9 kW rotating anode with a Cu X-ray source (λ = 1.54 Å) and Ge (220 × 2) monochromator was utilized to take XRR measurements. Data were collected from 0.1° to 10° at a speed of 0.25° min^{-1} with a 0.02° step size. The GenX reflectivity software package was used to analyze the data and fit the structural parameters.[54] Similar to previous studies,[40,41] a three-layer approach of native silicon oxide, TiO_2, and dye was employed to fit the XRR data. To minimize the number of parameters fit in the model, the thickness and SLD of the native oxide layer were fixed at 5 Å and 18.9×10^{-6} Å^{-2}, respectively. The substrate of the model was Si wafer with a constant SLD of 20.1×10^{-6} Å^{-2}. Errors were calculated based on the change in parameter needed to result in a greater than 5% worsening in the model figure of merit. All collected XRR data and the corresponding model fits are given in Figure S4 in the Supporting Information. See the Supporting Information for all calculations.

Acknowledgements

C.B.C. and J.A.V. gratefully acknowledge funding from the Winston Churchill Foundation of the United States. S.F.O. acknowledges financial support from MINECO (CTQ2014-52331R) and Gobierno de Aragón-FEDER-Fondo Social Europeo 2014-2020 (E14_17R). X.Z. acknowledges support from Hong Kong Research Grants Council (HKBU22304115-ECS and C5015-15GF), Areas of Excellence Scheme ([AoE/P-03/08]), and the Hong Kong Baptist University (RC-ICRS/15-16/02E, RC-ICRS/1617/02C-CHE, and RC-IRMS/16/17/02CHEM). J.M.C. thanks the 1851Royal Commission of the Great Exhibition for the 2014 Fellowship in Design, hosted by Argonne National Laboratory where work was done supported by DOE Office of Science, Office of Basic Energy Sciences, and research resources from the Center of Nanoscale Materials and the Argonne Leadership Computing Facility, which are DOE Office of Science Facilities, all under contract no. DE-AC02-06CH11357. The authors thank Dr. Erwin Reisner, Department of Chemistry, University of Cambridge, for use of laboratory space and equipment. The authors thank the Science and Technology Facilities Council (STFC) for access to facilities at the ISIS Materials Characterisation Laboratory, STFC Rutherford Appleton Laboratory (RAL), and its funding provision for research carried out at the Research Complex at Harwell at RAL. J.M.C. is also indebted to the ISIS Facility at RAL and to Tessella, for financial support (for E.J.B.). C.B.C. and J.M.C. conceived and designed the project. E.J.B. and J.M.C. generated the initial database of dyes. C.B.C., J.M.C., and A.V-M. designed and executed the computational workflow. T.T. and K.R.J.T. synthesized dye 8c. J.J. and S.X. synthesized dye XS6. G.B.B., S.C., and X.Z. synthesized dye H3. L.G.A. and S.F.O. synthesized dye 15. A.C. synthesized dye C1. L.S. and J.M.C. fabricated Si wafers with atomic layer deposited TiO_2. C.B.C. performed experimental validation for all of the synthesized dyes. J.A.V. assisted with the photovoltaic measurements and completed NMR measurements. G.B.G.S. and D.W.N. assisted with the AFM and XRR measurements. C.B.C. and J.M.C. wrote the manuscript. All authors edited the manuscript. Q7

Conflict of Interest

The authors declare no conflict of interest.

Keywords

co-sensitization, data-mining, dye-sensitized solar cell, materials discovery, photovoltaic device

Received: September 10, 2018
Revised: November 13, 2018
Published online:

[1] A. Zunger, *Nat. Rev. Chem.* **2018**, *2*, 0121.
[2] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, O. Levy, *Nat. Mater.* **2013**, *12*, 191.
[3] R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T. O. L. Sunde, D. Chon, K. R. Poeppelmeier, A. Zunger, *Nat. Chem.* **2015**, *7*, 308.
[4] F. Yan, X. Zhang, Y. G. Yu, L. Yu, A. Nagaraja, T. O. Mason, A. Zunger, *Nat. Commun.* **2015**, *6*, 7308.
[5] A. Zakutayev, X. Zhang, A. Nagaraja, L. Yu, S. Lany, T. O. Mason, D. S. Ginley, A. Zunger, *J. Am. Chem. Soc.* **2013**, *135*, 10048.
[6] L. Yu, A. Zunger, *Phys. Rev. Lett.* **2012**, *108*, 068701.
[7] L.-C. Lin, A. H. Berger, R. L. Martin, J. Kim, J. A. Swisher, K. Jariwala, C. H. Rycroft, A. S. Bhown, M. W. Deem, M. Haranczyk, B. Smit, *Nat. Mater.* **2012**, *11*, 633.
[8] I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, K. W. Jacobsen, *Energy Environ. Sci.* **2012**, *5*, 5814.
[9] C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, *Nat. Chem.* **2012**, *4*, 83.
[10] A. Peddapuram, H. Cheema, R. E. Adams, R. H. Schmehl, J. H. Delcamp, *J. Phys. Chem. C* **2017**, *121*, 8770.
[11] P. Brogdon, H. Cheema, J. H. Delcamp, *ChemSusChem* **2018**, *11*, 86.
[12] B. E. Hardin, H. J. Snaith, M. D. McGehee, *Nat. Photonics* **2012**, *6*, 162.
[13] M. Freitag, J. Teuscher, Y. Saygili, X. Zhang, F. Giordano, P. Liska, J. Hua, S. M. Zakeeruddin, J.-E. Moser, M. Grätzel, A. Hagfeldt, *Nat. Photonics* **2017**, *11*, 372.
[14] J. Gong, K. Sumathy, Q. Qiao, Z. Zhou, *Renewable Sustainable Energy Rev.* **2017**, *68*, 234.
[15] K. Yoo, J.-Y. Kim, J. A. Lee, J. S. Kim, D.-K. Lee, K. Kim, J. Y. Kim, B. Kim, H. Kim, W. M. Kim, J. H. Kim, M. J. Ko, *ACS Nano* **2015**, *9*, 3760.
[16] Z. Wen, M.-H. Yeh, H. Guo, J. Wang, Y. Zi, W. Xu, J. Deng, L. Zhu, X. Wang, C. Hu, L. Zhu, X. Sun, Z. L. Wang, *Sci. Adv.* **2016**, *2*, https://doi.org/10.1126/sciadv.1600097. Q8
[17] M. J. Yun, S. I. Cha, S. H. Seo, D. Y. Lee, *Sci. Rep.* **2015**, *4*, 5322.
[18] M. Peng, X. Yu, X. Cai, Q. Yang, H. Hu, K. Yan, H. Wang, B. Dong, F. Zhu, D. Zou, *Nano Energy* **2014**, *10*, 117.
[19] H. C. Weerasinghe, F. Huang, Y.-B. Cheng, *Nano Energy* **2013**, *2*, 174.
[20] S. G. Hashmi, M. Özkan, J. Halme, S. M. Zakeeruddin, J. Paltakari, M. Grätzel, P. D. Lund, *Energy Environ. Sci.* **2016**, *9*, 2453.
[21] P. J. Holliman, M. L. Davies, A. Connell, B. V. Velasco, T. M. Watson, *Chem. Commun.* **2010**, *46*, 7256.

[22] P. J. Holliman, M. Mohsen, A. Connell, M. L. Davies, K. Al-Salihi, M. B. Pitak, G. J. Tizzard, S. J. Coles, R. W. Harrington, W. Clegg, C. Serpa, O. H. Fontes, C. Charbonneau, M. J. Carnie, *J. Mater. Chem.* **2012**, *22*, 13318.
[23] K. Kakiage, Y. Aoyama, T. Yano, K. Oya, J. Fujisawa, M. Hanaya, *Chem. Commun.* **2015**, *51*, 15894.
[24] S. L. Bayliss, J. M. Cole, P. G. Waddell, S. McKechnie, X. Liu, *J. Phys. Chem. C* **2014**, *118*, 14082.
[25] F. A. Y. N. Schröder, J. M. Cole, P. G. Waddell, S. McKechnie, *Adv. Energy Mater.* **2015**, *5*, 1401728.
[26] G. Pepe, J. M. Cole, P. G. Waddell, J. R. D. Griffiths, *M. Syst. Des. Eng.* **2016**, *1*, 402.
[27] G. Pepe, J. M. Cole, P. G. Waddell, S. McKechnie, . *Syst. Des. Eng.* **2016**, *1*, 86.
[28] M. C. Swain, J. M. Cole, *J. Chem. Inf. Model.* **2016**, *56*, 1894.
[29] L. Zhang, J. M. Cole, *ACS Appl. Mater. Interfaces* **2015**, *7*, 3427.
[30] J. M. Cole, K. S. Low, H. Ozoe, P. Stathi, C. Kitamura, H. Kurata, P. Rudolf, T. Kawase, *Phys. Chem. Chem. Phys.* **2014**, *16*, 26684.
[31] F. Labat, C. Adamo, *J. Phys. Chem. C* **2007**, *111*, 15034.
[32] G. Boschloo, A. Hagfeldt, *Acc. Chem. Res.* **2009**, *42*, 1819.
[33] C. Maglione, A. Carella, R. Centore, S. Fusco, A. Velardo, A. Peluso, D. Colonna, A. Di Carlo, *J. Photochem. Photobiol., A* **2016**, *321*, 79.
[34] D. Karthik, V. Kumar, K. R. Justin Thomas, C.-T. Li, K.-C. Ho, *Dyes Pigm.* **2016**, *129*, 60.
[35] L. Zhang, Y. Liu, Z. Wang, M. Liang, Z. Sun, S. Xue, *Tetrahedron* **2010**, *66*, 3318.
[36] R. Pérez-Tejada, N. Martínez de Baroja, S. Franco, L. Pellejà, J. Orduna, R. Andreu, J. Garín, *Dyes Pigm.* **2015**, *123*, 293.
[37] Y. Hua, H. Wang, X. Zhu, A. Islam, L. Han, C. Qin, W.-Y. Wong, W.-K. Wong, *Dyes Pigm.* **2014**, *102*, 196.
[38] V. Venkatraman, R. Raju, S. P. Oikonomopoulos, B. K. Alsberg, *J. Cheminf.* **2018**, *10*, 18.
[39] A. B. Nepomnyashchii, B. A. Parkinson, *Langmuir* **2013**, *29*, 9362.
[40] J. McCree-Grey, J. M. Cole, P. J. Evans, *ACS Appl. Mater. Interfaces* **2015**, *7*, 16404.
[41] J. McCree-Grey, J. M. Cole, S. A. Holt, P. J. Evans, Y. Gong, *Nanoscale* **2017**, *9*, 11793.
[42] L. Zhang, J. M. Cole, *J. Mater. Chem. A* **2017**, *5*, 19541.
[43] D. Weininger, *J. Chem. Inf. Model.* **1988**, *28*, 31.
[44] D. M. Lowe, P. T. Corbett, P. Murray-Rust, R. C. Glen, *J. Chem. Inf. Model.* **2011**, *51*, 739.
[45] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *J.Cheminf.* **2011**, *3*, 33.
[46] *RDKit: Open-Source Cheminformatics*, **2017**.
[47] J. Stewart, *MOPAC2016*, Stewart Computational Chemistry, Colorado Springs, CO, USA **2016**.
[48] R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *Sci. Data* **2014**, *1*, 140022.
[49] S. A. Rice, M. S. Zhao, *Phys. Rev. B* **1998**, *57*, 13501.
[50] T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. V. R. Schleyer, *J. Comput. Chem.* **1983**, *4*.
[51] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[52] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, W. A. de Jong, *Comput. Phys. Commun.* **2010**, *181*, 1477.
[53] D. Nečas, P. Klapetek, *Open Phys.* **2011**, *10*, 181.
[54] M. Björck, G. Andersson, *J. Appl. Crystallogr.* **2007**, *40*, 1174.
[55] F. Schiffmann, J. VandeVondele, J. Hutter, R. Wirz, A. Urakawa, A. Baiker, *J. Phys. Chem. C* **2010**, *114*, 8398.
[56] G. Pepe, J. M. Cole, P. G. Waddell, J. I. Perry, *Mol. Syst. Des. Eng.* **2016**, *1*, 416. Q9