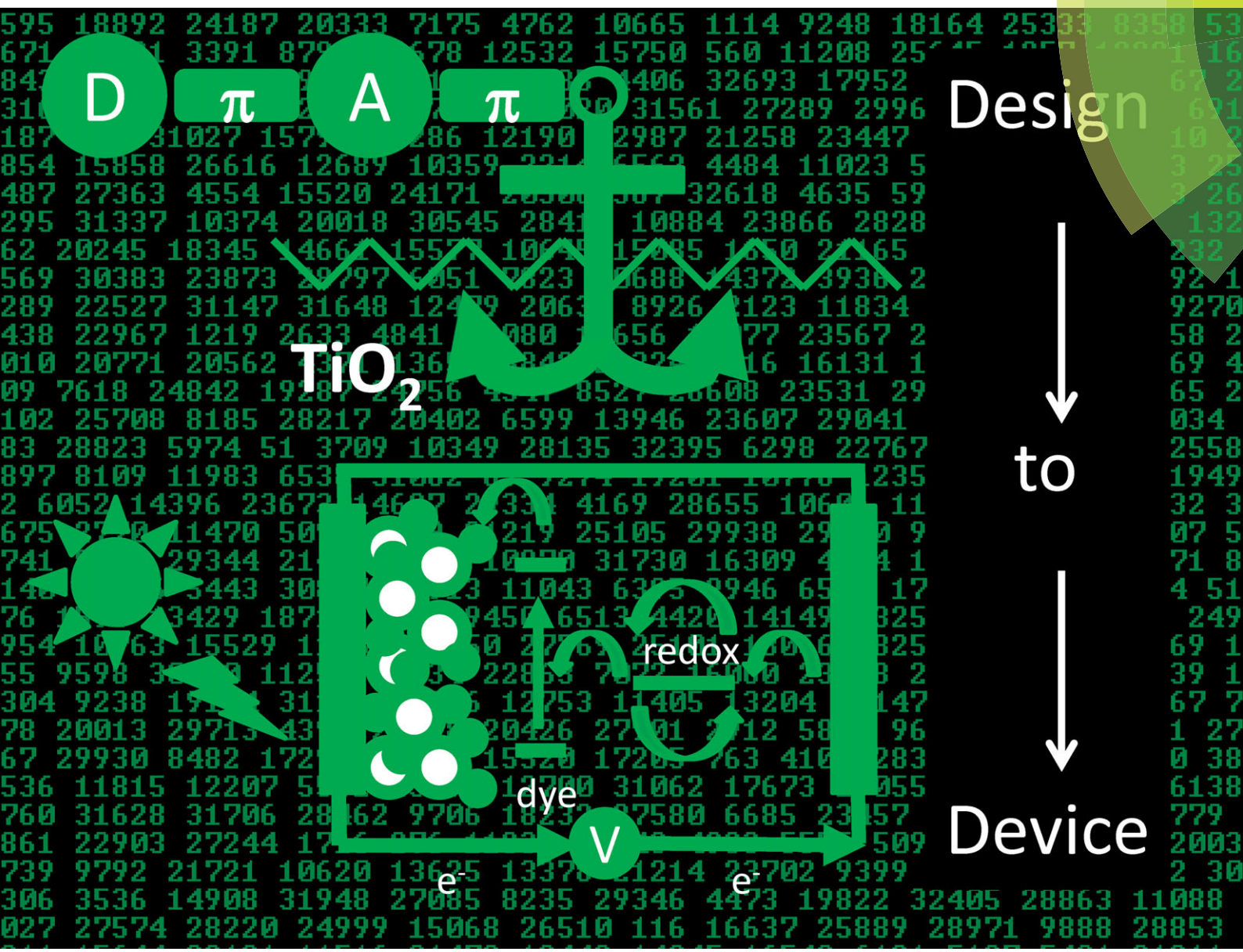
Volume 16 | Number 48 | 28 December 2014 | Pages 26393–27136

PCCP

Physical Chemistry Chemical Physics
www.rsc.org/pccp



ISSN 1463-9076

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION
Jacqueline M. Cole *et al.*
Data mining with molecular design rules identifies new class of dyes for dye-sensitised solar cells

PCCP

COMMUNICATION

View Article Online
View Journal | View Issue

Data mining with molecular design rules identifies new class of dyes for dye-sensitised solar cells†

Cite this: *Phys. Chem. Chem. Phys.*, 2014, **16**, 26684

Received 16th June 2014,
Accepted 24th June 2014

DOI: 10.1039/c4cp02645d

www.rsc.org/pccp

Jacqueline M. Cole,*[abc] Kian Sing Low,[a] Hiroaki Ozoe,[d] Panagiota Stathi,[e] Chitoshi Kitamura,[d] Hiroyuki Kurata,[f] Petra Rudolf[e] and Takeshi Kawase[d]

A major deficit in suitable dyes is stifling progress in the dye-sensitised solar cell (DSC) industry. Materials discovery strategies have afforded numerous new dyes; yet, corresponding solution-based DSC device performance has little improved upon 11% efficiency, achieved using the N719 dye over two decades ago. Research on these dyes has nevertheless revealed relationships between the molecular structure of dyes and their associated DSC efficiency. Here, such structure–property relationships have been codified in the form of molecular dye design rules, which have been judiciously sequenced in an algorithm to enable large-scale data mining of dye structures with optimal DSC performance. This affords, for the first time, a DSC-specific dye-discovery strategy that predicts new classes of dyes from surveying a representative set of chemical space. A lead material from these predictions is experimentally validated, showing DSC efficiency that is comparable to many well-known organic dyes. This demonstrates the power of this approach.

1. Introduction

There is currently a world-wide deficit in suitably efficient materials for photovoltaic applications, despite their crucial role in realising new generation power sources. This is particularly relevant to the field of dye-sensitised solar cells (DSCs), which is one of the strongest contenders for next generation photovoltaic technology. Although less efficient than silicon-based solar cells, DSCs are far more cost-effective to the extent that their price-to-performance ratio achieves 'grid-parity' status, *i.e.* they are competitive with fossil-fuel energy production.

DSCs arose from the pioneering work of O'Regan and Grätzel[1] and are based upon a chemical photosynthetic redox process. The molecular dye is a particularly critical component of a DSC since it is responsible for both the light-harvesting of energy from the sun, and electron injection that initiates the chemical redox reaction of the solar cell. Consequently, there have been extensive efforts to discover new materials that outperform the ruthenium-based dye, N719 (10–10.4% solar-cell efficiency[2]), which remained the world's most efficient dye for DSCs for over 20 years. A major breakthrough finally came in 2011 with the report[3] of a zinc porphyrin-based dye which broke this world record, affording 12.3% under 1 sun illumination, when coupled with a cobalt-based electrolyte in a Grätzel cell. Last year, a solid-state DSC employing a lead-iodide based pigment surpassed solution-based DSC records with a 15% efficiency.[4]

This said, metal-based dyes are extremely expensive relative to organic dyes, and are less environmentally sound. Given that price-to-performance ratio governs 'grid parity', economic incentives are heavily motivating the development of efficient organic dye alternatives. In this regard, carbazole-based organic dyes with ferrocene-based electrolytes have received particular acclaim, demonstrating 7.5% solar-cell efficiency;[5] 10.3% DSC efficiency has even been reported using a thiophene-based dye, albeit in a highly volatile electrolyte.[6] But clearly there is still a long way to go in this area.

Dye materials discovery is still based on serendipity or iterative chemical substitution methodologies that are based on an *a priori* known DSC active chemical exemplar. Such methods are very valuable within a specific area of synthetic chemistry; see for example, the work by Lee *et al.*[7] on designing a porphyrin class of dyes by systematic chemical substitution. However, this approach can only ever afford improvements within a currently known class of DSC dyes. Computational efforts have similarly restricted themselves to predicting or

[a] *Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge, CB3 0HE, UK. E-mail: jmc61@cam.ac.uk*
[b] *Argonne National Laboratory, 9700 S Cass Avenue, Argonne, IL 60439, USA*
[c] *International Institute for Complex Adaptive Matter, University of California, Davis, CA 95616, USA*
[d] *Graduate School of Engineering, University of Hyogo, 2167 Shosha, Himeji, Hyogo 671-2280, Japan*
[e] *Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands*
[f] *Department of Chemistry, Graduate School of Science, Osaka University, Machikaneyama-Cho 1-1, Toyonaka, Osaka, 560-0043, Japan*

† Electronic supplementary information (ESI) available: Preparation and chemical characterisation of **1** and **2**; Density Functional Theory and Time-Dependent Density Functional Theory of **1** and **2**; X-ray Photoelectron Spectroscopy of **1** and **2**. See DOI: 10.1039/c4cp02645d

rationalising new dyes within the framework of an existing class of chemical dyes. Such methods therefore have no capacity to reveal entirely new classes of suitable DSC dyes. If one is going to make a step-change in dye performance in DSCs, a distinct shift in the approach to dye design is therefore urgently needed that specifically targets the materials discovery of new classes of dyes.

To this end, we present a systematic large-scale data-mining procedure that successfully identifies new classes of dyes, whose intrinsic molecular characteristics predict good DSC functionality. Its underlying search methodology relies on a judicious encoding of molecular design rules for dyes that have recently begun to develop in the DSC field.

In that regard, molecular design rules are developing *via* the application of chemical intuition to empirical surveys of DSC dye discovery in the literature. Certain chemical groups and molecular architectures that commonly feature in a well-performing DSC dye have been spotted and translated into generic design principles.[8,9] For example, the importance of having an electron donor (D) and acceptor (A), separated *via* a π-conjugating framework, in an organic DSC dye was quickly realised. This D–π–A molecular architecture enables a charge-separated resonance structure, from which an electron–hole pair can be generated. Within this molecular design structure, A then has the additional role of injecting the electronic charge into the conduction band of the TiO_2 semiconductor electrode, to which the dye is adsorbed *via* A. Thus, the electrical circuit is initiated.

Recent work has shown that A need not also be the substituent that adsorbs to TiO_2, with the advantage that a natural adsorbing substituent (Ads), most commonly a carboxyl group, is not the best acceptor.[10,11] To this end, a D–π–A–π_{phenyl}–Ads chemical motif has demonstrated in excess of 6 × DSC efficiency improvements over its analogous D–π–A molecular architecture.[12] The judicious choice of a phenyl ring to separate A and Ads provides sufficient aromatic ring stabilisation to inhibit the back-transfer of electronic charge from TiO_2 to A, thereby circumventing undesirable electron recombination.[9]

This decoupling of A and Ads has exciting implications for systematic molecular design. For the first time, we have the opportunity to discover new classes of dyes by searching chemical space for molecules that contain strong D–π–A moieties, devoid of the constraints that Marcus theory[10] imposes on Ads.

2. Algorithmic dye prediction strategy

This work realises this opportunity by automatically sampling a diverse range of 118 465 organic molecules with an algorithm (see Fig. 1), designed to first identify molecules containing D–π–A moieties, and then rank them according to D–π–A strength using bond-length alternation[13] and molecular dipole moment as criteria.

To this end, the Cambridge Structural Database[14] provided a naturally diverse set of chemical space (118 465 organic molecules) for this study. Data quality control was pre-assured using the crystallographic $R_1 < 0.07$ statistic as a compound inclusion threshold for data mining, while all disordered materials,

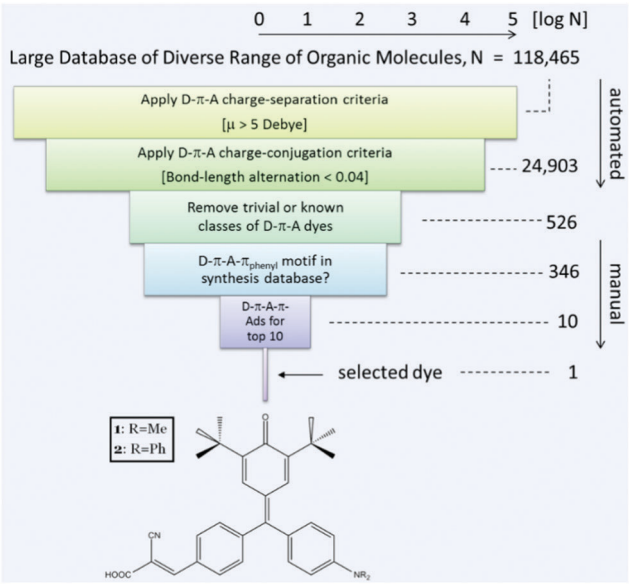


Fig. 1 Step-wise filtering algorithm to identify trial DSC dye candidates from large sets of chemical data. A series of DSC molecular design criteria were sequentially applied to an initial set of 118 465 molecules: (a) compounds containing a D–π–A structural motif were automatically sifted from the initial chemical set *via* charge (i) separation (threshold $\mu > 5$ Debye: 24 903 compounds) and (ii) delocalisation (bond-length alternation) considerations, yielding an initial short-list of 526 molecules. (b) Manual intervention then removed (i) trivial or known D–π–A classes of molecules and (ii) those where there are no examples of D–π–A–π_{para} in a synthesis database; (c) the 10 most novel chemical motifs in the resulting final short-list underwent 'molecular mutation' to create the target D–π–A–π–Ads (Ads = cyanoacrylate) compounds; DFT was employed to test their energetic feasibility for application as DSC dyes. The lead material, **1**, was identified; it was then synthesised, so that its DSC dye prospects could be experimentally validated.

polymers and ionic salts were excluded. Semi-empirical (AM1) calculations on all 118 465 molecules provided their molecular dipole moments, μ, automatically by coupling the search engine of the database to MOPAC.[15] To ensure that prospective dyes have sufficient charge-separation, a $\mu > 5$ Debye data filter was imposed. This minimum threshold was selected for two reasons. Firstly, μ ~ 6–7 Debye is commonly observed in mid-sized DSC dyes such as coumarins,[16] and μ generally increases when creating larger dye motifs such as D149.[17] Secondly, this minimum corresponds closely to the charge separation of an electron and proton separated by 1 Å (1 eÅ = 4.8 Debye), suggesting that a molecule with $\mu > 5$ Debye can sustain an electron–hole pair separated by at least one organic chemical bond length (typically ~1–1.5 Å). The resulting 24 903 molecules were then passed through a second data filter that exclusively selects dyes with good D–π–A intramolecular charge transfer (ICT) characteristics. To this end, Hammett constants[18] were employed to identify molecules containing an electron donor, D, and the bond-length alternation (BLA) parameter[13] then determined the level of π-delocalisation between D and A; a BLA < 0.04 Å was used as the threshold for good ICT. This series of calculations employed an algorithm similar to that of Cole & Weng,[15] but with one important exception: the electron acceptor, A, was not defined by Hammett constants;

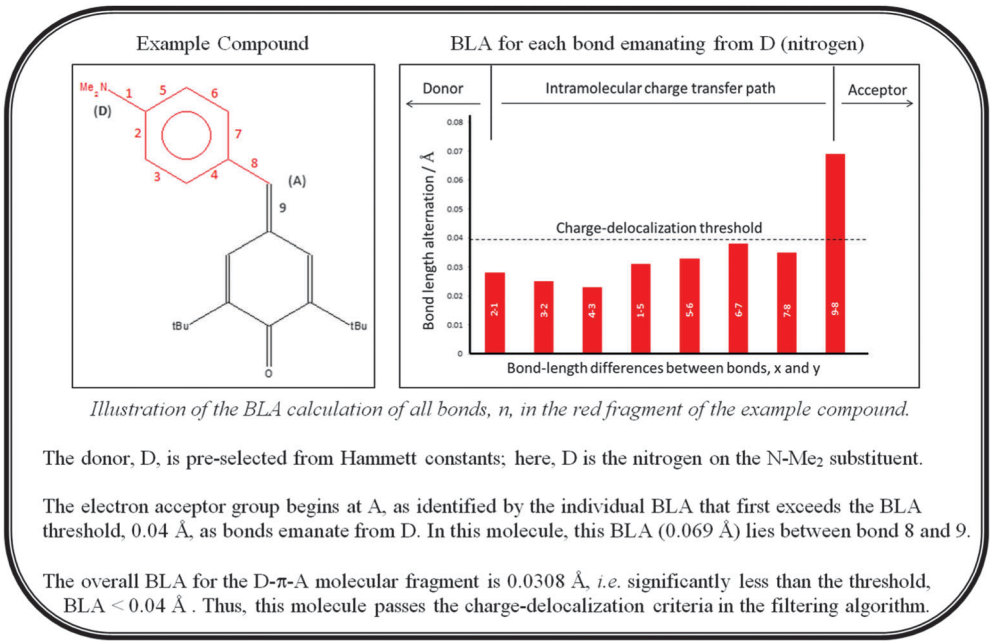


Fig. 2 Example of BLA calculation encoded to automatically evaluate D–π–A intramolecular charge transfer. The compound chosen as the exemplar is the 'molecular mutation' precursor to the lead material, **1**.

rather, BLA was calculated sequentially along all bond paths emanating from D until a BLA $>$ 0.04 Å was exceeded; the atom at which this inequality falters is deemed to be an acceptor, A. If BLA $<$ 0.04 Å was not realised after the second BLA calculation along the bond path from D, it is discarded as a donor in that molecule. This ICT calculation is illustrated in Fig. 2 using, as an exemplar, the molecular dye precursor that ultimately identified the target material.

Trivial molecules and known classes of D–π–A dye motifs were then manually removed from the resulting 526 initial short-list of molecules. This left 346 candidate dye precursors for which SciFinder® acted as a further filter to target molecules where a known synthetic route for realising a target with a *para*-phenyl group substituted to A was available; this provided a suitable electronic pathway to generate a D–π–A–π–X motif (X = any chemical substituent). 10 dye motifs comprised the final short-list, where this synthetic feasibility was assured and which, by inspection, appear most distinct chemically from known DSC dyes. These 'top 10' molecules then underwent simulated 'molecular mutation' where X was replaced by a cyanoacrylate (Ads) group using DFT calculations. This group was selected as the Ads moiety given its particularly good adsorption characteristics.[19] DFT then assessed the energetic suitability of the resulting D–π–A–π–Ads target molecules to DSC function. All DFT calculations used a B3LYP functional and a 6-31++G(d,p) basis set. This final step identified **1** (Fig. 1 bottom) as the lead material.

3. Dye candidate selection

Compound **1** stood out from the final short-list of molecules *via* manual inspection of its respective HOMO–LUMO charge-transfer distributions and associated band gaps, generated *via* Density Functional Theory (DFT) calculations (see ESI†). A particular attraction of compound **1** was its near-complete shift of charge-density distribution from the D–π–A moiety in the HOMO to the π–Ads moiety in the LUMO (Fig. 3). On the one hand, this indicates that the charge-separated state is stable so that **1** can avoid undesirable electron recombination from TiO_2-to-dye electron back transfer. On the other hand, this reveals that Ads in **1** is electron-rich for efficient electron injection into the TiO_2 electrode. The associated HOMO and LUMO energy levels of **1** are also favourable as illustrated in Fig. 2; with the LUMO residing $\sim$0.8 eV *above* the TiO_2 conduction band edge, and the HOMO lying $\sim$0.8 eV *below* that of the driving voltage of the iodide/triiodide redox couple, the most commonly used DSC electrolyte. A minimum driving potential difference of $\Delta V \sim 0.2$ eV is required in each case, to ensure good electron injection[20] and dye regeneration,[21] respectively. Premature dye regeneration has also been associated with structural aspects of dye molecules; in particular, conjugation length and the type of side chain on the dye have demonstrated relationships.[9] These factors were not considered by the algorithm presented herein. However, it has been shown that such structural influences have only minor effects if $\Delta V_{\text{electrolyte–HOMO}} > 0.2$–$0.25$ eV[22] as in **1**.

Compound **1** was therefore synthesised (see Methods and ESI†) in order that its dye performance in a DSC could be experimentally validated. A related compound, **2**, was also synthesised by substituting the π-NMe_2 unit of **1** for a triarylamine group. This enabled a check for possible dye aggregation issues in **1** *via* a DSC performance comparison of **1** and **2**: triarylamines are proving very popular as bulky donor groups in dyes to offset intrinsic DSC photocurrent losses owing to dye aggregation.[25]

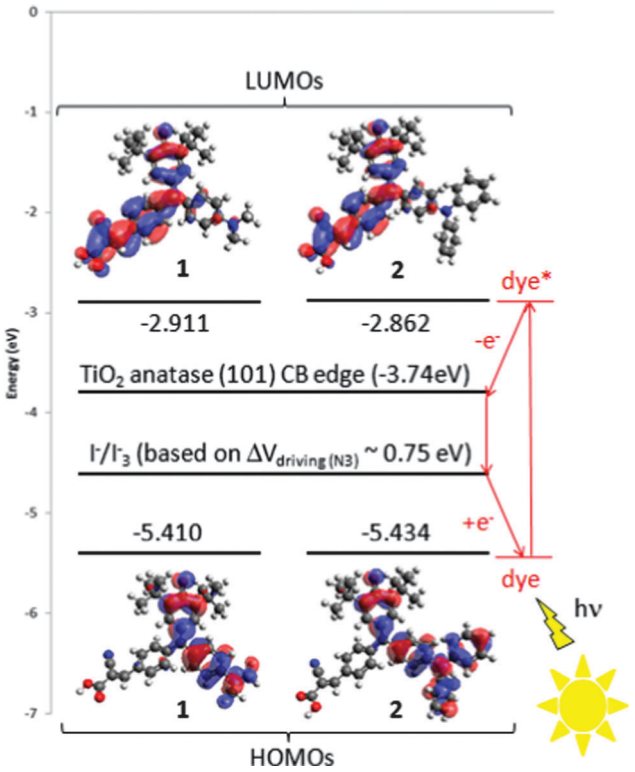


Fig. 3 The energy relationship between frontier molecular orbitals of **1** and **2** and key quantum levels in other DSC components (left), which drives device operation (right, in red). DFT calculations afforded highest-occupied and lowest-unoccupied molecular orbitals (HOMOs and LUMOs) of **1** and **2**. These computational results are backed up by cyclic voltammetry experiments (see ESI†). Results show that a near-complete shift of orbital population from the D–π–A group to the π_{phenyl}–Ads moiety occurs in the HOMO-to-LUMO transition for both dyes. This prospects good electron injection efficiency and minimal dye-to-TiO_2 electron back transfer. The 0.8 eV energy gap between the LUMO of the dye and TiO_2 conduction-band edge, assuming preferential dye adsorption to the (101) face of TiO_2 anatase,[23] is also favorable for electron injection. A similar driving voltage between the HOMO of each dye and the redox level of the I^-/I^-_3 electrolyte is also indicative of good dye regeneration; the redox level is calculated with reference to the maximum 0.75 eV TiO_2-to-I^-/I^-_3 driving voltage, *cf.* for the cationic N719, *i.e.* N3.[24] All of these factors suggest that **1** and **2** would be attractive DSC dyes.

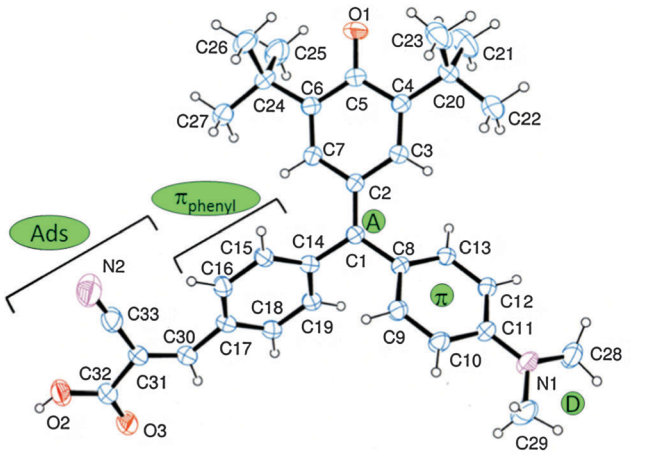


Fig. 4 Molecular structure of **1**. Bond lengths evidence (i) high levels of D–π–A intramolecular charge transfer (ICT), as predicted, which enable stable charge-separation; (ii) low levels of π-conjugation in the quinodimethene group which shows that this moiety is largely isolated from ICT, *i.e.* it does not play a significant electronic role in DSC function; (iii) π-conjugation is somewhat hindered between the π_{phenyl}–Ads and D–π–A moieties; indeed, the respective phenyl rings lie out-of-plane with respect to each other, by 72.1(1)°. As such, the likelihood of undesirable TiO_2-to-dye electron back-transfer is diminished.

4. Experimental validation of dye candidate

4.1 Materials characterisation of 1 and 2

The molecular structure of **1** (Fig. 4) was confirmed by single-crystal X-ray diffraction (see ESI† for full details). This data also evidenced experimentally the high D–π–A intramolecular charge-transfer (ICT), predicted by DFT. To this end, the NMe_2 group of **1** (D) carries a definite δ+ charge in its ground-state given that all three N–C bonds display partial double-bonded character, with that of N1–C11 being particularly marked. π-density from the three phenyl rings of **1** conjugates with the intersecting carbon, C1 (A), resulting in its heavily δ− character. In fact, the C1–C2 bond length (1.381 Å) is particularly short since it forms part of a classically bonded quinodimethene unit, whose high bond-length alternation[13] somewhat isolates it from the primary ICT pathway; this takes place in the delocalised regions of a molecule, *i.e. via* the other two (more delocalised) phenyl rings as indicated *via* HOMO-to-LUMO charge redistributions (Fig. 3). In the ground-state, the charge is essentially held within the D–π–A moiety between the highly δ+ and δ− atom centres; π-conjugation from the adjoining π_{phenyl}–Ads moiety is also hindered, being it twisted out of the plane. Fig. 3 shows that photo-absorption enables the ICT of **1** to channel almost entirely through to the π_{phenyl}–Ads moiety, thereby rendering a stable charge-separated excited state. This suggests good resistance of **1** to premature dye regeneration.

The crystal structure of **2** was not available. However, one assumes that the substitution of NMe_2 for NPh_2 would place the nitrogen in a less δ+ environment owing to the lower electron donating ability of Ph relative to Me; as such, **2** is expected to possess slightly less pronounced charge-separation. Indeed, this effect is manifest in the DFT charge-distributions of Fig. 3.

Prior to incorporating **1** or **2** into a full DSC device, their optical absorption characteristics, when adsorbed onto an open TiO_2 electrode, were evaluated. UV/vis absorption spectroscopy revealed a good panchromatic nature of both **1** and **2**, in methanol solution and when dye-sensitised onto TiO_2 nanoparticles (Fig. 5a). The minor red/blue shifting observed between the solution and dye-sensitised TiO_2 spectra indicate successful dye adsorption to TiO_2. This was further evidenced by X-ray Photoelectron Spectroscopy (XPS) of dye-sensitised TiO_2 as compared with bare TiO_2; the 2p Ti XPS peaks displayed a −0.2 to 0.3 eV shift upon dye-sensitisation for both **1** and **2**, as is characteristic of adsorption of organic chromophores to TiO_2.[26,27] Corresponding shifts in 1s O XPS spectra, reminiscent of surface Ti···HO–OC interactions in **1** and **2**, were also identified (see ESI† for further XPS details).

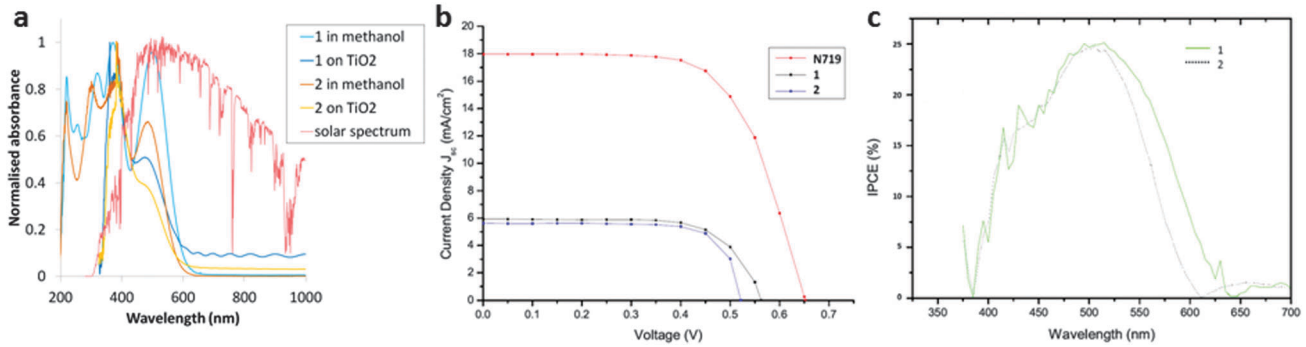


Fig. 5 (a) UV/vis absorption spectra of **1** and **2** in solution and adsorbed onto TiO_2; these show that **1** and **2** capture well the solar spectrum (overlaid) up to ~600 nm, as is common in DSC dyes. (b) *J–V* curves from the DSC photovoltaic tests. (c) Corresponding IPCE plots of **1** and **2** in the functioning DSC.

Time-dependent (TD)-DFT calculations permitted the assignment of optical transitions to the three main optical absorption peaks, λ_{max}, of **1** and **2**, respectively: HOMO to LUMO charge-transfer excitation (499, 481 nm); a HOMO − 1 to LUMO π–π* transition (372, 371 nm); a HOMO to LUMO + 1 π–π* transition (321, 300 nm). The greater amount of charge shifting from the HOMO in **2**, relative to **1**, results in its slightly lower λ_{max} values for the first and third peaks.

4.2 DSC device fabrication and testing of 1 and 2

Full DSC device fabrication and testing was then undertaken, employing the doctor-blade method to lay the nanoporous TiO_2 and Pt electrodes onto transparent conducting glass substrates, sensitising the TiO_2 by its overnight soaking in a methanol solution of the dye, and heat-sealing the two electrodes together using a thermoplastic resin (Surlyn), while placing a I^-/I_3^- redox couple betwixt to act as an electrolyte. A full description of these methods is given in the ESI.†

The photocurrent–voltage (*J–V*) characteristics under standard solar simulation reporting conditions (AM 1.5G illumination at 1000 W m^{-2} at 298 K) were determined for each DSC, yielding its power conversion efficiency (PCE, η) according to:

$$\eta = P_{out(max)}/P_{in} = J_{sc} \times V_{oc} \times FF/P_{in}$$

where J_{sc} is the short-circuit current density, V_{oc} is the open-circuit voltage, FF is the fill factor (maximum power observed in the *J–V* curve (Fig. 5b) divided by J_{sc}, V_{oc}), and P_{in} is the intensity of simulated sunlight impinging on the active area of the solar cell.

Results in Table 1 show that DSC devices employing **1** and **2** deliver ~30% of the solar-cell performance efficiency realised using N719. The results of **1** and **2** were normalised to in-house results on DSCs containing N719, since interpretations based on this $\eta_{Dye}:\eta_{N719}$ ratio, rather than η, obviate the common variability in DSC efficiency reporting. Such variability is caused by the renowned acute sensitivity of J_{sc} to the fabrication conditions;[28] while $J_{sc(dye)}:J_{sc(N719)}$ is generally consistent between duplicate runs, as in this case. Absolute V_{oc} and FF values are consistent and typical for organic DSCs bearing an I^-/I^-_3 electrolyte;[25] thus good electron injection and DSC device quality are implied, respectively.

Table 1 Optoelectronic and photovoltaic characteristics of **1** and **2**. The DSC performance efficiency for devices containing **1** and **2** can be ascribed to 30% of that of N719. Interpretations are based on this $\eta_{Dye}:\eta_{N719}$ ratio rather than η in order to avoid the common variability in absolute DSC efficiency reporting.[21] The DSC device fabrication and testing for **1**, **2** and N719 were all obtained from the same in-house fabrication and testing methods; conducted in triplicate to ensure good consistency. V_{oc} and FF values evidence efficient electron injection and DSC device quality, respectively

Dye	V_{oc} (V)	J_{sc} (mA cm^{-2})	FF	η (%)	$\eta_{dye}:\eta_{N719}$ (%)
1	0.56	5.92	0.69	2.32	30.8
2	0.52	5.62	0.75	2.20	29.2
N719	0.65	17.96	0.64	7.54	100

Corresponding Incident-Photon-to-Current Efficiency (IPCE) measurements were undertaken from 375 to 700 nm in 5 nm increments. Results (Fig. 5c) on each dye show a red shift relative to the dye in solution or as adsorbed on TiO_2 (open electrode), as is commonly observed.[29] **1** retains a greater panchromatic absorption over **2** in the DSC by an extent of ~45 nm at long-wavelengths, judging from the IPCE spectra.

The same $\eta_{dye}:\eta_{N719}$ ratio observed for both **1** and **2** suggests that **1** exhibits no significant dye aggregation problems; recall that the presence of the bulky donor group, triarylamine, in **2** was used to test for this issue. The molecular structure of **1** holds a possible explanation for its lack of dye aggregation, despite hosting a non-bulky donor group: its three rings are somewhat twisted relative to each other (see ESI†), affording a distinctly 3-D molecule. π···π intermolecular interactions that typically cause dye aggregation[9] in planar dye molecules therefore cannot easily form. The two *tert*-butyl groups in **1** will further hinder such intermolecular interactions with neighbouring molecules.

There are two common structural causes of DSC performance losses that were not considered by the molecular design approach presented herein. One of these is electron recombination owing to undesirable TiO_2-to-electrolyte electron back transfer. To some extent, this has been implicitly tested *via* experimental validation of **2** since the triarylamine offers a hydrophobic substitute for the more commonly used long alkyl chains that physically block the approach of the electrolyte to

the TiO_2 surface. Finally, long-term chemical and photo-chemical stability is naturally a key asset of a dye for DSCs. While it is too soon to assess the long-term photo-chemical stability of **1** or **2**, they present unusually high melting points (182 and 208 °C) for a small organic material which suggests good chemical stability.

5. Conclusions

In summary, we have discovered a new class of dyes for DSC applications, which were predicted from large-scale data-mining that encodes molecular design rules. Two dye specifications, **1** and **2**, exhibit solution-based DSC efficiencies of ~1/3rd that of those containing world's best performing dyes. This is very encouraging, given that these findings emanate from the first experimental validation phase of this algorithmic approach to DSC dye materials-by-design. Moreover, **1** does not appear to suffer from significant dye aggregation which is a common drawback of DSC dyes; **1** also shows good chemical stability. Therefore, this class of dyes now has the potential to be fully developed by the manual, but chemically intuitive substitution strategies, as has transpired for many classes of *a priori* known organic dyes; such developments would lead to a brand new range of dyes for DSCs, with the prospect that some new specifications are likely to be more superior than the D–π–A–π–Ads building blocks, **1** or **2**. Indeed, one would expect better performing dyes to unfold from these ideal 'ICT molecular cores'; in the same way that organic DSC dyes, such as MK2,[30] have been generated by extension of their core molecular building blocks (*e.g.* MK44 for MK2). The likes of MK44 had, of course, been identified by means of chemical intuition. This new algorithm-based molecular design rule strategy has the innate strength to successfully predict DSC dye molecular building blocks before even turning to the chemical bench; and its systematic nature already offers many short-cuts for the synthetic chemist; with further refinement, the predictive power and reach of this computational approach stands to deliver even greater dye design foresight and versatility. Indeed, the large-scale data-mining nature of these DSC dye predictions somewhat parallels computational efforts in related fields of energy research; for example, in organic photovoltaics[31,32] and in inorganic energy materials.[33]

New structure–property relationships are continuously evolving from DSC-related chemical research, which can also be fed into the next steps in the algorithmic prediction for DSC applications. To this end, next steps in algorithmic development will not only embed further dye design rules based on structure, but will combine both structural and energetic considerations in a fully automated and concerted manner. This will naturally extend into the prediction of not only the dye, but all DSC material components simultaneously, since their roles are so entangled; this has the tantalising scope of realising new electrolytes, electrodes, or substrates as well as the dyes themselves. Ultimately, we seek an algorithmic abstraction of the material influences on DSC performance efficiency, whose verisimilitude to its device functionality is sufficient to yield a superior DSC technology.

6. Experimental methods

Synthesis

1 and **2** were synthesised by reacting 4-bromo-2,6-di(*t*-butyl)-1-(trimethylsiloxy)benzene with 4-*N*,*N*-dialkylamino-4′-formylbenzophenone ethylene acetal (where alkyl = methyl[34] or phenyl[35]) in the presence of *n*-butyl lithium in tetrahydrofuran at −78 °C, followed by a Knoevenagel condensation reaction with cyanoacetic acid in acetonitrile.[14] Full synthetic details are given in the ESI.†

DSC fabrication and testing

DSCs were fabricated using Dyesol TiO_2 DSL 18NR-T paste. The TiO_2 was deposited onto cleaned fluorine-doped tin oxide (FTO) conductive glass (Solaronix S. A.) *via* the doctor-blade method. The composite was then successively sintered for 10 minutes at 100 °C, 10 minutes at 150 °C, 30 minutes at 325 °C, 5 minutes at 400 °C, and 30 minutes at 300 °C, before allowing to cool to 70 °C where it was dipped into a 0.5 mM dye solution using methanol for **1** and **2**, and a 1 : 1 ratio of acetonitrile : *tert*-butanol solvent for N719. The platinum counter-electrode was prepared by spin coating a H_2PtCl_6 solution (52 mM in isopropanol) onto the FTO coated glass and then annealing at 450 °C. The electrolyte comprised 50 mM iodide/triiodide in acetonitrile (Iodolyte AN-50 from Solaronix S. A.). The photo-electrode and counter-electrode were sealed together using 25 mm thick Surlyn. *J–V* characteristics of the resulting DSCs were determined under the equivalent of 100 mW cm^{-2} AM 1.5G illumination with a calibrated ABET Sun 2000 solar simulator, corrected for spectral mismatch.

Authors contribution

J.M.C. conceived the project, designed and carried out the data-mining work, interpreted the primary results according to DSC molecular design criteria, and drafted the paper. K.S.L. conducted the UV/vis absorption spectroscopy, DSC fabrication and testing, and detailed DFT and TD-DFT calculations on **1** and **2**, under the supervision of J.M.C. P.S. undertook the XPS characterisation and associated data interpretation of **1** and **2**, under the direction of P.R. H.O. synthesised **1** and **2** and did the cyclic voltammetry; C.K. determined the crystal structure of **1** from X-ray diffraction; H.K. performed the mass spectrometry for **1** and **2**. H.O., C.K., and H.K. were supervised by T.K.

Acknowledgements

J.M.C. is indebted to the ICAM Branches Cost Sharing Fund, the Fulbright Commission for a UK-US Fulbright Scholar Award, and to Argonne National Laboratory where work done was supported by DOE Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. J.M.C. would also like to thank Peter Littlewood at Argonne for helpful discussions.

K.S.L. acknowledges the EPSRC for a DTA PhD studentship (EP/P504120/1). Partial financial support also came from the ''Top Research School'' program of the Zernike Institute for Advanced Materials under the Bonus Incentive Scheme (BIS) of the Netherlands' Ministry of Education, Science, and Culture and from the 'Stichting voor Fundamenteel Onderzoek der Materie (FOM)', which is financially supported by the 'Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)'. This study was also supported by the Hyogo prefecture and the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1 B. O'Regan and M. Grätzel, *Nature*, 1991, **353**, 737.
2 M. K. Nazeeruddin, A. Kay, I. Rodicio, R. Humphry-Baker, E. Mueller, P. Liska, N. Vlachopoulos and M. Grätzel, *J. Am. Chem. Soc.*, 1993, **115**, 6382.
3 A. Yella, H.-W. Lee, H. N. Tsao, C. Yi, A. K. Chandiran, Md. K. Nazeeruddin, E. W.-G. Diau, C.-Y. Yeh, S. M. Zakeeruddin and M. Grätzel, *Science*, 2011, **334**, 629.
4 J. Burschka, N. Pellet, S. J. Moon, R. Humphry-Baker, P. Gao, M. K. Nazeeruddin and M. Grätzel, *Nature*, 2013, **499**, 316.
5 T. Daeneke, T. H. Kwon, A. B. Holmes, N. W. Duffy, U. Bach and L. Spiccia, *Nat. Chem.*, 2011, **3**, 211.
6 W. Zeng, Y. Cao, Y. Bai, Y. Wang, Y. Shi, M. Zhang, F. Wang, Y. Pan and P. Wang, *Chem. Mater.*, 2010, **22**, 1915.
7 C.-W. Lee, H.-P. Lu, C.-M. Lan, Y.-L. Huang, Y.-R. Liang, W.-N. Yen, Y.-C. Liu, Y.-S. Lin, E. W.-G. Diau and C.-Y. Yeh, *Chem. – Eur. J.*, 2009, **15**, 1403.
8 A. Mishra, M. K. R. Fischer and P. Bäuerle, *Angew. Chem., Int. Ed.*, 2009, **48**, 2474.
9 B.-G. Kim, K. Chung and J. Kim, *Chem. – Eur. J.*, 2013, **19**, 5220.
10 R. A. Marcus, *J. Chem. Phys.*, 1956, **24**, 979.
11 H. Tian, X. Yang, R. Chen, Y. Pan, L. Li, A. Hagfeldt and L. Sun, *Chem. Commun.*, 2007, 3741.
12 S. Haid, M. Marszalek, A. Mishra, M. Wielopolski, J. Teuscher, J. E. Moser, R. Humphry-Baker, S. M. Zakeeruddin, M. Grätzel and P. Bäuerle, *Adv. Funct. Mater.*, 2012, **22**, 1291.
13 C. B. Gorman and S. R. Marder, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 11297.
14 F. H. Allen, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58**, 380.
15 J. M. Cole and Z. F. Weng, *Adv. Mater. Res.*, 2010, **123–125**, 959.
16 M. S. Zakerhamidi, A. Ghanadzadeh and M. Moghadam, *Spectrochim. Acta, Part A*, 2011, **78**, 961.
17 A. El-Zohry, A. Orthaber and B. Zietz, *J. Phys. Chem. C*, 2012, **116**, 26144.
18 L. P. Hammett, *J. Am. Chem. Soc.*, 1937, **59**, 96.
19 D. P. Hagberg, T. Marinado, K. M. Karlson, K. Nonomura, P. Qin, G. Boschloo, T. Brinck, A. Hagfeldt and L. Sun, *J. Org. Chem.*, 2007, **72**, 9550.
20 K. Hara, T. Sato, R. Katoh, A. Furube, Y. Ohga, A. Shinpo, S. Suga, K. Sayama, H. Sugihara and H. Arakawa, *J. Phys. Chem. B*, 2003, **107**, 597.
21 S. Wenger, P. A. Bouit, Q. Chen, J. Teuscher, D. Di Censo, R. Humphry-Baker, J. E. Moser, J. L. Delgado, N. Martin, S. M. Zakeeruddin and M. Grätzel, *J. Am. Chem. Soc.*, 2010, **132**, 5164.
22 T. Daeneke, A. J. Mozer, Y. Uemura, S. Makuta, M. Fekete, Y. Tachibana, N. Koumura, U. Bach and L. Spiccia, *J. Am. Chem. Soc.*, 2012, **134**, 16925.
23 F. Labat and C. Adamo, *J. Phys. Chem. C*, 2007, **111**, 15034.
24 G. Boschloo and A. Hagfeldt, *Acc. Chem. Res.*, 2009, **42**, 1819.
25 M. Liang and J. Chen, *Chem. Soc. Rev.*, 2013, **42**, 3453.
26 E. M. J. Johansson, M. Hedlund, H. Siegbahn and H. Rensmo, *J. Phys. Chem. B*, 2005, **109**, 22256.
27 E. M. J. Johansson, M. Hedlund, M. Odelius, H. Siegbahn and H. Rensmo, *J. Chem. Phys.*, 2007, **126**, 244303.
28 C. Y. Jen, L. V. Munukutla, S. Radhakrishnan, A. M. Kannan and A. Htun, *J. Nanosci. Nanotechnol.*, 2012, **12**, 1829.
29 M. Grätzel, *Inorg. Chem.*, 2005, **44**, 6841.
30 N. Koumura, Z.-S. Wang, S. Mori, M. Miyashita, E. Suzuki and K. Hara, *J. Am. Chem. Soc.*, 2006, **128**, 14256.
31 I. Y. Kanal, S. G. Owens, J. S. Bechtel and G. R. Hutchison, *J. Phys. Chem. Lett.*, 2013, **4**, 1613.
32 J. Hachmann, R. Olivares-Amaya, A. Jinich, A. L. Appleton, M. A. Blood-Forsythe, L. R. Seress, C. Roman-Salgado, K. Trepte, S. Atahan-Evrenk, S. Er, S. Shrestha, R. Mondal, A. Sokolov, Z. Bao and A. Aspuru-Guzik, *Energy Environ. Sci.*, 2014, **7**, 698.
33 S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito and O. Levy, *Nat. Mater.*, 2013, **12**, 191.
34 T. Kumagai, T. Anki, T. Ebi, A. Konishi, K. Matsumoto, H. Kurata, T. Kubo, K. Katsumoto, C. Kitamura and T. Kawase, *Tetrahedron*, 2010, **66**, 8968.
35 G. P. Schiemenz and G. Stein, *Tetrahedron*, 1970, **26**, 2007.

