

Cite this: *Digital Discovery*, 2025, 4, 2156

# Pursuit of singlet fission fulvene candidates using inverse design†

Irene Casademont-Reig, \*ab Roger Monreal-Corona, ac Eline Desmedt, a Freija De Vleeschouwer \*a and Mercedes Alonso \*a

The pursuit of higher solar cell efficiency has sparked considerable interest in singlet fission materials, which offer the possibility to surpass the Shockley–Queisser limit. Singlet fission (SF) is a multiexciton process in which a singlet exciton is converted into two triplet excitons, doubling the number of low energy excitons per absorbed photon. A central bottleneck in the development of SF-enhanced solar cells is the limited number of chromophores fulfilling the strict energy-matching conditions required for SF. Tetracene and pentacene remain the most investigated SF materials despite their poor photostability. To address this challenge, we propose an inverse molecular design protocol to accelerate the discovery of suitable SF candidates based on fulvene derivatives. The excited-state energies of fulvenes can be modulated by the electronic nature of substituents and they behave as aromatic molecular chameleons with the ability to arrange their  $\pi$ -electrons to exhibit aromaticity in both the ground and the lowest triplet state. Using a best-first search algorithm, we explored the combinatorial chemical compound space to identify optimal substitution patterns satisfying the primary thermodynamic SF condition (*i.e.*,  $E(S_1) \approx 2E(T_1)$ ). This approach represents a significant improvement over traditional design methods. Database analysis revealed 17 fulvene candidates fulfilling all the conditions for SF-enhanced silicon solar cells. Next to identifying potential SF chromophores, we derived general design rules highlighting the critical role of strategically positioning electron-donating and electron-withdrawing groups across the six substitution sites. The proposed protocol represents a significant step forward in the quest for efficient SF materials, offering a systematic approach to navigating chemical compound spaces and optimizing molecular structures with desired properties.

Received 11th December 2024  
Accepted 1st July 2025

DOI: 10.1039/d4dd00389f

rsc.li/digitaldiscovery

## Introduction

Solar cells are a very important cornerstone in the transition of the world's energy production from a fossil-based source to a CO<sub>2</sub> neutral future.<sup>1,2</sup> To push the performance of silicon solar cells beyond their thermodynamic efficiency limit, the sensitization of photovoltaic cells using singlet fission materials has attracted tremendous interest.<sup>3–5</sup> The singlet fission (SF) process offers the possibility to surpass the Shockley–Queisser limit<sup>6</sup> and achieve power conversion efficiencies of approximately 45% due to the ability of doubling the number of low energy excitons per absorbed photon.<sup>4,7</sup> In the SF process, a singlet exciton ( $S_1$ ) is converted into two separated triplets ( $T_1$ ), involving a correlated triplet pair state  $^1(T_1T_1)$  with singlet

multiplicity.<sup>8,9</sup> As a spin-allowed process, SF can proceed in the ultrafast scale, outcompeting most other loss mechanisms and reaching quantum yields approaching 200%.<sup>10</sup> Notably, triplet excitons exhibit longer lifetimes compared to singlet ones, thereby facilitating their migration towards the donor/acceptor interface.<sup>11</sup> Nonetheless, the dearth of SF chromophores combining an efficient SF rate with suitable chemical stability hampers the realization of commercial SF-enhanced solar cells.<sup>12</sup>

The SF process requires chromophores that satisfy strict energy-matching conditions regarding the arrangement of the lowest-energy excited states. First, the energy of  $S_1$  should be at least equal to twice the energy of  $T_1$  to subsequently generate two triplet states:<sup>13,14</sup>

$$E(S_1) \geq 2E(T_1) \quad (1)$$

Ideally,  $E(S_1) - 2E(T_1)$  must be close to zero to minimize the energy loss during the SF process.<sup>8</sup> To hamper  $S_1$ -to- $T_2$  intersystem crossing,<sup>15</sup> it is desirable that a second thermodynamic condition is also fulfilled:

$$E(S_1) \leq E(T_2) \quad (2)$$

\*Research group of General Chemistry (ALGC), Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussels, Belgium. E-mail: irene.casademont.reig@vub.be; freija.de.vleeschouwer@vub.be; mercedes.alonso.giner@vub.be

<sup>b</sup>Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain

<sup>c</sup>Institut de Química Computacional i Catalisi, Departament de Química, Universitat de Girona, C/Maria Aurèlia Capmany 69, 17003 Girona, Catalonia, Spain

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4dd00389f>



From these two energy criteria it can be deduced that  $E(T_2) > 2E(T_1)$ . If the latter is satisfied, triplet–triplet recombination into the second triplet excited state will be prevented.<sup>16</sup> All these conditions combined imply that  $2E(T_1) \leq E(S_1) \leq E(T_2)$  for efficient SF chromophores. Finally, sufficiently high triplet energies are also required to facilitate triplet exciton transfer to the semiconductor. For silicon,<sup>8</sup>

$$E(T_1) \geq 1.11 \text{ eV} \quad (3)$$

Besides the strict energy conditions, strong visible absorption, long triplet lifetimes, photo- and thermal stability, and cost-effectiveness constitute important factors in device engineering.<sup>12,17</sup> Stability remains a critical challenge for organic solar cells compared to inorganic counterparts,<sup>18</sup> with issues such as thermal, photochemical, and oxidative degradation affecting their long-term performance.<sup>8</sup> Recent work on SF-sensitized silicon solar cell architectures has demonstrated their potential to extend device lifetimes by lowering the operating temperature compared to conventional silicon solar cells.<sup>19</sup> This study estimated that incorporating tetracene as a sensitizer could significantly reduce thermal degradation, extending the silicon cell lifetime by approximately 3.7 years. In addition, SF chromophores like tetracene become transparent upon degradation, allowing the underlying silicon cell to continue operating with reasonable efficiency.

Unfortunately, most organic molecules fail to meet the energy-matching conditions primarily due to their relatively high  $T_1$  energies,<sup>20</sup> which significantly limits the pool of SF chromophores.<sup>12</sup> Tetracene and pentacene remain the most intensively investigated SF materials,<sup>5,21,22</sup> despite their poor photostability and sensitivity to oxygen.<sup>8,23</sup> To address these limitations, various chemical strategies, such as steric protection, the introduction of electron-withdrawing substituents, and the captodative effect, have been employed to enhance the intrinsic stability of SF dyes.<sup>24–26</sup> As a result, the search for new SF chromophores is a very active field together with the development of general design guidelines<sup>7,27</sup> and innovative strategies<sup>28</sup> to accelerate their discovery driven by the urgency to switch to renewable energy sources.<sup>2</sup>

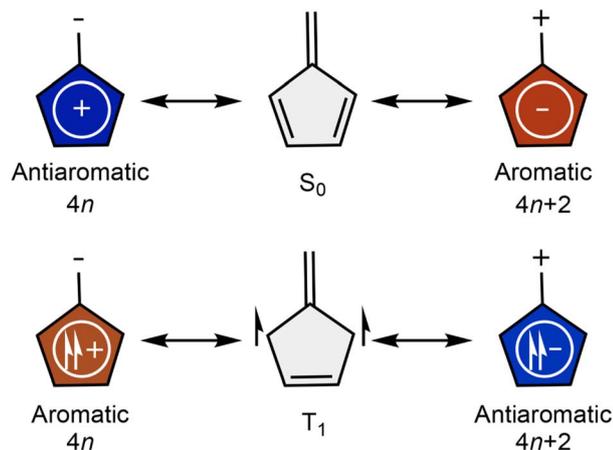
The search for SF materials is largely driven by computational approaches by evaluating the energy-matching conditions outlined above.<sup>29–31</sup> From these studies, the ground-state diradical character has emerged as the underlying concept for rendering efficient SF materials.<sup>32</sup> Overall, molecules with the proper amount of open-shell singlet diradical character ( $y_0 > 0.1$ ) benefit from lower triplet energies, fulfilling the SF thermodynamic requirements.<sup>33</sup> In addition, the tetradical character ( $y_1$ ) needs to be sufficiently small (*i.e.*,  $y_0/y_1 > 5$ ) for a molecule to possess the best electronic arrangement for SF. Both indices together compose the multiradical character index ( $y_0, y_1$ ) which is a useful metric for identifying novel SF chromophores, as verified for a pool of 241 potential SF candidates compiled from the Cambridge Structural Database (CSD).<sup>34</sup> However, while open-shell diradical character can enhance SF efficiency, it might also reduce photochemical stability, limiting the practical use of certain SF chromophores in device

fabrication.<sup>24</sup> A recent study reveals that a balance between the diradical character and stability is possible for a series of *o*- and *p*-quinomethides and the relationship can be understood through Clar's sextet theory.<sup>35</sup> In addition to diradical character, intrinsic stability is influenced by several factors, including molecular structure, functionalization, and steric protection, all of which play crucial roles in determining the reactivity and long-term stability of organic compounds. To enhance the chemical stability of SF chromophores, the concept of excited-state aromaticity has emerged as an alternative guiding principle to engineer the excited states of organic materials without compromising their stability.<sup>36–39</sup> Yet, recent analyses underscore the importance of solid computational assessments to render the Baird-aromaticity concept as a useful tool for tuning the energies of molecular excited states.<sup>40–42</sup>

Notwithstanding these discoveries, it is important to highlight how difficult it is to find organic chromophores that satisfy the energy-matching conditions. Out of  $\sim 40$  k molecules compiled from the CSD database, only 241 structures were identified as potential SF chromophores representing a success rate of less than 1%.<sup>43</sup> Accordingly, it is important to introduce alternative design approaches to explore more efficiently the vast and complex chemical compound space (CCS).<sup>44</sup> Machine learning models have been recently extended to the field of SF to assess the thermodynamic driving force (eqn (1)) of larger databases of polycyclic aromatic hydrocarbons and cibalackrot derivatives.<sup>31,45,46</sup> Recent computational screening studies have investigated extensive molecular datasets, typically comprising a broad chemical space, with only a few focusing on a common core undergoing systematic substitution. Despite these efforts, success rates have remained low due to the strict energetic requirements of SF and the dependence on the choice of the functional.<sup>31,47–49</sup> In this work, we present an alternative inverse molecular design (ID) framework to accelerate the discovery of SF materials. Inverse design approaches reverse the conventional design paradigm by starting with the desired properties and producing a molecule that fulfills them as an output.<sup>50,51</sup> Hence, inverse design involves the optimization of the target property as a function of the chemical structure.<sup>52,53</sup> Remarkably, only a small fraction of a predefined CCS needs to be explored using ID to identify new and largely improved structures, as we recently proved for the discovery of nonlinear optical (NLO) switches.<sup>54–56</sup>

To test the potential of our ID strategy to identify SF chromophores fulfilling the energy-matching conditions, we focus on fulvenes with different substitution patterns. The properties and aromaticity of fulvenes are strongly influenced by the electronic nature of the substituents at the exocyclic position.<sup>57</sup> Strikingly, they behave as aromatic molecular chameleons since they can rearrange the  $\pi$ -electrons to exhibit aromatic characteristics in both the ground state ( $S_0$ ) and the lowest triplet state ( $T_1$ ) according to the Hückel and Baird rules, respectively (Scheme 1).<sup>58,59</sup> The adaptive nature of fulvenes can be exploited to modulate the energies of the relevant states for SF as well as to assess the role of ground- and excited-state aromaticity on the properties of excited states.<sup>37</sup> Fulvenes are typically highly reactive,<sup>60</sup> but several studies have shown that appropriate





**Scheme 1** Resonance structures that influence the  $S_0$  and  $T_1$  states of fulvenes, illustrating its behaviour as dipolar aromatic chameleons. Adapted from ref. 66.

substitution patterns can markedly enhance their stability by promoting a delocalized  $\pi$ -electronic structure.<sup>37,61,62</sup> Persistent substituted fulvenes have been synthesized and fully characterized,<sup>57,63,64</sup> with 6,6-dicyanofulvenes identified as suitable  $n$ -type additives in bulk heterojunction solar cells.<sup>65</sup>

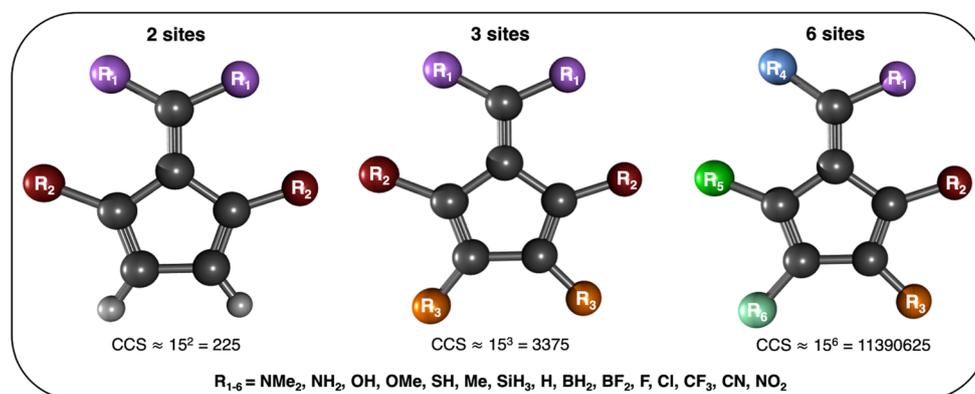
Our main goal is to establish an inverse design protocol to accelerate the discovery of substituted fulvenes for singlet fission. As a test bed, we consider symmetric fulvenes with two and three modifiable sites as well as asymmetric fulvenes containing up to six different substituents (Fig. 1). The substituent library contains 15 different functional groups with varied electronic nature, including electron-donating (EDG), neutral, and electron-withdrawing (EWG) groups. Ottosson *et al.* proved that these substituents induce a large variation on the  $S_1$  and  $T_1$  energies relative to the parent fulvene.<sup>37,59</sup> First, a proof-of-concept study is performed to unveil the impact of the substitution at each position independently ( $R_1$ ,  $R_2$ , and  $R_3$ ) on the lowest-lying excited-state energies as well as on the aromaticity of the  $S_0$  and  $T_1$  states. This subset is employed to assess the accuracy of our computational protocol and the performance of electronic, structural, and magnetic criteria to quantify ground-

and excited-state aromaticity. Second, we search for the optimal fulvene functionalizations for distinct substitution patterns (*cf.* Fig. 1) using the best-first search (BFS) algorithm.<sup>51,56,67,68</sup> Finally, we analyse the database generated from the inverse design runs to understand the key functionalization patterns required to fulfill the SF conditions in fulvene derivatives.

## Inverse molecular design algorithm

Since our goal is the combinatorial exploration of the fulvene chemical space towards suitable SF candidates, we employed the BFS algorithm to find optimal functionalizations for differently substituted fulvenes (Fig. 1).<sup>51</sup> BFS involves a site-by-site optimization during which chemical functionalizations from a predefined library are introduced into the selected molecular framework. Chemical modifications to the molecular structure are accepted if they lead to an improvement of the target property.<sup>54</sup> The BFS algorithm is a discrete optimization method implying that only structures corresponding to ‘real’ chemical species are considered for the property optimization, *i.e.*, functionalized fulvene derivatives.<sup>68</sup> Furthermore, the BFS method is quite efficient for both single-property and multiple-property optimizations,<sup>51</sup> and delivers crucial insights on the influence of the type and location of each functionalization on the target property since for all the sites the full substituent library is systematically screened.<sup>55</sup>

Fig. 2 illustrates schematically how the BFS algorithm works applied on a symmetric fulvene with three modifiable sites ( $R_1$ ,  $R_2$ , and  $R_3$ ) and a functionalization library of 5 substituents depicted as blue-, cyan-, green-, yellow-, and orange-colored circles. The size of the combinatorial CCS is determined by the number of functionalizations per modifiable site (for this example,  $5^3$ ) since all the sites share the same substituent library. The algorithm starts generating a random starting structure. Then, site 1 ( $R_1$ ) is optimized by evaluating the 5 possible structures according to the functionalization library. Here, it is important to note that full geometry optimizations and vibrational frequency calculations at the selected level of theory are performed for all generated fulvenes, followed by the excited-state calculations (Fig. S14<sup>†</sup>). For each structure, the



**Fig. 1** Fulvene test bed with different substitution patterns investigated in this work and the selected substituent library.



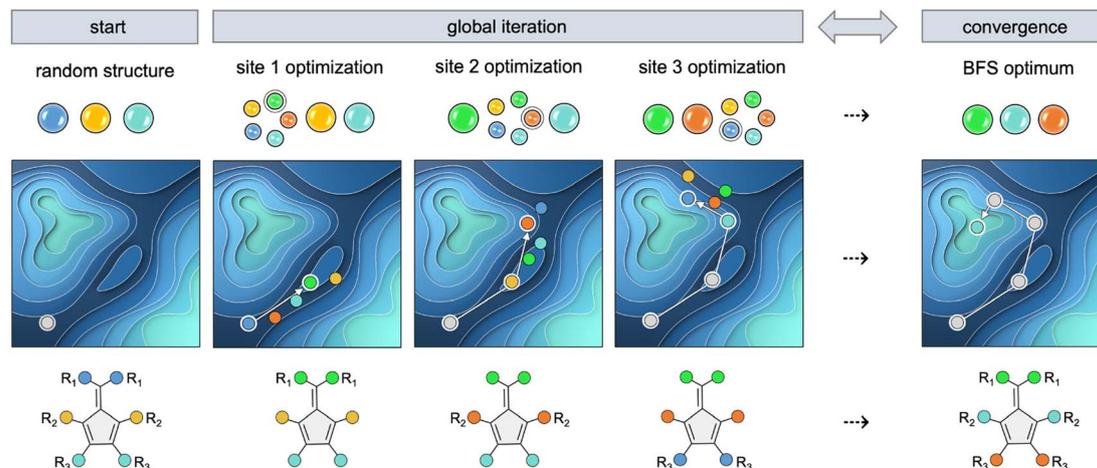


Fig. 2 Schematic workflow of the BFS optimization on a symmetric fulvene with three optimization sites ( $R_1$ ,  $R_2$ , and  $R_3$ ) and a functionalization library of 5 substituents represented as colored circles. Light-colored regions in the property map are more optimal. Adapted from ref. 56.

target property is computed, and its values are used to determine the largest gradient, as a finite difference approximation, among the 5 available substituents. BFS presumes that the largest gradient corresponds to the most promising optimization route and accordingly, the substituent leading to the largest property improvement is implemented. This modified structure is subsequently used for the optimization of site 2 ( $R_2$ ), while keeping fixed the other two positions. This procedure is repeated until all sites have been optimized. Then, the algorithm checks if the input (random starting structure unless otherwise stated) and output structure are identical. If yes, convergence is reached and the BFS optimization is complete. If the input and output structures are different, the BFS reprises with the output structure as the new input. Generally, the BFS algorithm takes several global iterations until convergence has been reached. This iterative process circumvents the often incorrect assumption that the sites can be optimized independently and ensures that the inter-site influences are partially considered.

Optima are found quite efficiently with this greedy algorithm, as we have previously shown in the optimization of the HOMO–LUMO gap in diamondoids<sup>52,69</sup> and the NLO contrasts of hexaphyrin-based switches.<sup>54–56</sup> Nevertheless, when the target function has multiple local optima, BFS usually converges to the nearest (local) optimum since it follows the largest gradient. Therefore, in this work, multiple BFS runs are conducted varying the starting structure and the site sequence, to explore different regions of CCS. As we aim to derive structure–property relationships for SF, BFS offers the significant advantage of searching the CCS locally at each iteration step by only changing one site at a time while evaluating the full substituent library.

As a target property for the SF chromophore inverse design, we rely on the first energy-matching condition (eqn (1)), which is the most fundamental thermodynamic condition to ensure an efficient SF process.<sup>8,12,15</sup> This figure-of-merit is commonly used in the screening of chromophores for SF applications.<sup>43,45,46</sup> Initially, we aim to maximize the SF function (eqn

(4)) as this metric helps to achieve the greatest disparity between the lowest-lying excited energy states. The advantage of this approach is that it tests the boundary of the function, while potentially generating a sufficient number of structures with  $f_{\text{SF}} = 0$ .

$$f_{\text{SF}} = \Delta E_{\text{SF}} = E(S_1) - 2E(T_1) \quad (4)$$

Subsequently, to further refine our target function, we target a value for  $E(S_1) - 2E(T_1)$  around 0.1–0.2 eV which is optimal for reducing energy losses during the SF process.<sup>8,37</sup> Accordingly, for the latest BFS runs on six modifiable sites, we minimize the absolute value of the SF function. The additional SF conditions outlined in the introduction (eqn (2) and (3)) will be imposed through filter functions on the generated fulvene database.

## Computational details

The fulvene geometries have been fully optimized and characterized using harmonic vibrational frequency calculations with the M06-2X<sup>70</sup> functional in combination with the 6-311+G(d,p) basis set<sup>71,72</sup> in gas phase. The geometry optimizations and frequencies, calculated within the DFT framework, were performed with the Gaussian 16 software package.<sup>73</sup> The M06-2X hybrid functional has been chosen, as it has been demonstrated to provide relative energies close to chemical accuracy for various  $\pi$ -conjugated systems,<sup>74</sup> when compared to the gold standard canonical CCSD(T) at the extrapolated basis set limit.<sup>75–77</sup> This choice also helps to mitigate the impact of delocalization error in the description of the electronic structure of  $\pi$ -conjugated molecules.<sup>78,79</sup> Based on the optimized structures, linear-response TD-DFT calculations were performed at the same level of theory to obtain the vertical and adiabatic excited-state energies based on previous benchmarks devoted to the screening of SF materials.<sup>37,80</sup> Notably, a TD-DFT protocol utilizing M06-2X combined with a triple- $\zeta$  basis set demonstrated good accuracy in predicting the  $S_1$  and  $T_1$  excitation energies of a diverse set of 100 molecules as compared to



experimental data (Fig. S5†).<sup>43</sup> Nonetheless, we caution that the heavily-parameterized M06-2X *meta*-GGA hybrid may not be optimal for all the excitation energies, but a comprehensive benchmarking study is beyond the scope of this work.

The accuracy of our excitation energies was further evaluated through comparison with multiconfigurational methods. Single-point multireference calculations were conducted using the ORCA software package,<sup>81</sup> employing N-electron valence state perturbation theory (NEVPT2)<sup>82</sup> applied to complete active space self-consistent field CASSCF(12,12) wavefunctions.<sup>83,84</sup> These computations were performed with the def2-TZVP basis set,<sup>85,86</sup> along with the corresponding auxiliary basis sets, on the previously optimized DFT geometries.

The diradical character was calculated using the Yamaguchi method (eqn (5)) to determine the  $y_0$  parameter.<sup>87,88</sup> Single-point calculations at the (U)HF/6-31G(d) level of theory were employed, as this method effectively reproduces the trends in diradical character derived from CASSCF(6,6) calculations for a wide variety of predicted SF materials.<sup>34</sup>

$$y_i = 1 - \frac{2T_i}{1 + T_i^2} \text{ and } T_i = \frac{n_{\text{HONO}-i} - n_{\text{LUNO}+i}}{2} \quad (5)$$

where  $n_{\text{HONO}}$  and  $n_{\text{LUNO}}$  correspond to the occupation number of the highest occupied natural orbital and lowest unoccupied natural orbital, respectively. The diradical character ( $y_0$  parameter) is obtained by eqn (5) with  $i = 0$  and lies between 0 and 1, representing closed- and open-shell nature, respectively. Higher  $i$  values allow for the estimation of tetra-, hexa-, and higher-order radical character.

The calculation of the electronic aromaticity indices (FLU, BOA, and MCI) relies on QTAIM atomic partitioning performed by the AIMAll software.<sup>89</sup> The atomic overlap matrices (AOM) resulting from this partition and the molecular geometries serve as input for the ESI-3D program,<sup>90–92</sup> which provides the aromatic fluctuation index (FLU),<sup>93,94</sup> the bond-order alternation (BOA),<sup>74</sup> the harmonic oscillator model of aromaticity (HOMA),<sup>95</sup> the bond-length alternation (BLA),<sup>74,96,97</sup> the multicenter based index ( $I_{\text{ring}}$ ),<sup>98,99</sup> and the multicenter index (MCI).<sup>100</sup> The magnetic current density and the current strengths were obtained using the gauge including atomic orbital (GIAO) method<sup>101</sup> with the GIMIC program.<sup>102,103</sup> The density current vector plots were generated at 0.5 Å above the molecular plane. The Paraview 5.10.0 program was used for the visualization of the current densities.<sup>104</sup> For more details, see the ESI.†

Inverse design procedures were performed with the best-first search algorithm<sup>105–108</sup> implemented in the in-house CINDES program.<sup>109,110</sup> Based on a previous work,<sup>37</sup> we selected a fragment library of 15 different substituents going from strongly electron-donating groups (EDG) to strongly electron-withdrawing groups (EWG): NMe<sub>2</sub>, NH<sub>2</sub>, OH, OMe, Me, SiH<sub>3</sub>, H, F, BH<sub>2</sub>, BF<sub>2</sub>, SH, Cl, CF<sub>3</sub>, CN, and NO<sub>2</sub>. The structures generated by CINDES are automatically optimized and characterized through harmonic vibrational analysis, using the same level of theory as described above. After ensuring no imaginary frequencies are encountered for the optimized geometries, the excited-state calculations were performed within the framework of linear response TD-DFT, at the same level of theory. In the

inverse design procedure, vertical energies were used to scrutinize different candidates with respect to the first energy-matching condition (eqn (1) and (4)). For a detailed explanation about the accuracy of our computational approach including the use of vertical and adiabatic energies, we refer to the ESI.†

## Results and discussion

### Proof-of-concept: the role of aromaticity on the S<sub>1</sub> and T<sub>1</sub> energies of substituted fulvenes

As previously noted by Zeng *et al.*<sup>111</sup> and Ottosson *et al.*,<sup>37</sup> the parent fulvene is unsuitable for singlet-fission applications due to its low-lying T<sub>2</sub> state. Additionally, it undergoes efficient S<sub>1</sub>/S<sub>0</sub> radiationless decay through two possible conical intersections.<sup>112,113</sup> Using our TD-DFT protocol, the SF function (eqn (4)) for the parent fulvene is estimated at −1.37 eV, which is significantly far away than the optimal range of 0.1–0.2 eV needed for effective SF. Starting with the parent fulvene, in theory, three strategies can be considered to reduce the SF function value: (i) significantly increase  $E(S_1)$  relative to  $E(T_1)$ , (ii) lowering  $E(T_1)$  while minimizing the impact on  $E(S_1)$ , or (iii) reducing both  $E(S_1)$  and  $E(T_1)$  until  $E(S_1)$  is nearly twice as large as  $E(T_1)$ .

The first step in designing our protocol for the efficient search of fulvenes as promising singlet-fission materials was the rational fine-tuning of  $E(T_1)$  and  $E(S_1)$  using excited-state Baird and ground-state Hückel aromaticity concepts. For that, we used our library of 15 different substituents. These substituents were placed on three different positions on the fulvene scaffold (R<sub>1</sub>, R<sub>2</sub>, and R<sub>3</sub>) yielding 45 symmetric fulvenes in our archetype set (Fig. 3). Accordingly, we assessed at each position independently the substituent effect on the excited-state energies and aromaticity. As reported by Ottosson and coworkers,<sup>37</sup> the exocyclic R<sub>1</sub> position (R<sub>1</sub>-fulvenes subset) has the greatest impact on the ground-state aromaticity of fulvenes, with NICS(±1)<sub>zz,S<sub>0</sub></sub> values spanning a wide range from −19.6 to 9.2 ppm.

In Table 1, we present the ranges of the vertical excitation energies, NICS(±1)<sub>zz,S<sub>0</sub></sub>, NICS(±1)<sub>zz,T<sub>1</sub></sub>, MCI for the ground and the lowest triplet excited state, the SF function, and the Yamaguchi index across the different subsets. The exocyclic position (R<sub>1</sub>) exerts the largest influence on both the excitation energies and aromaticity descriptors, with the impact diminishing across the other substituted positions (R<sub>1</sub> > R<sub>2</sub> > R<sub>3</sub>). Although the R<sub>3</sub> position has the least impact on most of the parameters, it significantly affects the value of the SF function (eqn (4)), similar to the effect of the R<sub>2</sub> position. This trend is also observed for the aromatic character of both the ground state and the lowest triplet excited state, with variations decreasing from R<sub>1</sub> to R<sub>3</sub>.

The R<sub>1</sub>-fulvene subset exhibits the widest range of energies for the lowest-lying excited states (S<sub>1</sub> and T<sub>1</sub>), resulting in significant variations in the SF function. This group includes the fulvenes that present the highest S<sub>1</sub> and T<sub>1</sub> energy values among all the fulvenes in our archetype set. Additionally, the same subset of symmetric fulvenes shows the greatest variation



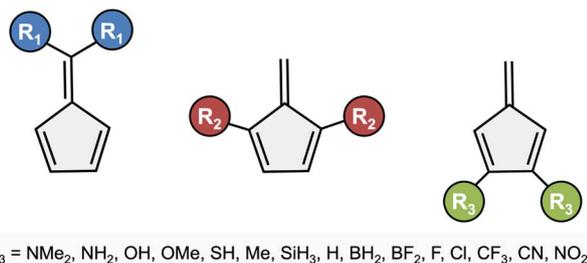


Fig. 3 Structures of archetype fulvenes with substitution positions labeled  $R_1$ ,  $R_2$ , and  $R_3$ . The selected substituent library is specified at the bottom of the figure.

in ground- and excited-state aromaticity, with Baird-aromatic fulvenes being the most optimal for the SF process (see Tables S4–S12<sup>†</sup> for specific energy and aromaticity index values). Our results confirm that  $\pi$ -electron withdrawing groups at the exocyclic position lead to low-lying  $T_1$  and  $S_1$  states, as they enhance the Baird-aromatic character of these states (Fig. 4).

In Fig. 4, the energies of the lowest-lying excited states of the  $R_1$ -fulvene subset are plotted according to their aromatic character, ranging from highly Hückel aromatic to highly Hückel antiaromatic.  $\text{NICS}(\pm 1)_{zz,S_0}$  values were used for the color coding as it is the aromatic index with the most diverse values, facilitating the distinction between aromatic and antiaromatic fulvenes. Similar plots for  $R_2$  and  $R_3$  positions are provided in the ESI (see Fig. S6 and S7).<sup>†</sup> For the ground state, all the structural and electronic indices correlate quite well with the magnetic descriptor except for the bond-order alternation index (Table 2). However, for the first triplet state, the correlations between the different descriptors drop down significantly reflecting the fact that these indices were originally designed to quantify ground-state aromaticity (see Table S13<sup>†</sup>).

Compared to the parent fulvene, the fulvenes in the  $R_1$  subset with the strongest EDGs, such as  $\text{NMe}_2$ ,  $\text{NH}_2$ ,  $\text{OH}$ , and  $\text{OMe}$ , exhibit higher  $S_1$  and  $T_1$  energies. In contrast, those with strong EWGs like  $\text{NO}_2$ ,  $\text{CN}$ ,  $\text{CF}_3$ , and  $\text{Cl}$  display the lowest  $S_1$  and  $T_1$  energies. This variation is closely related to their aromatic character: the strongest EWGs increase ground-state antiaromaticity, while the strongest EDGs enhance ground-state aromaticity. In other words, exocyclic EDGs enhance the contribution of the zwitterionic resonance structure in  $S_0$  with

Hückel aromatic  $6\pi$ -electron five-membered ring, while EWGs favour the contribution of the zwitterionic resonance structure with  $4\pi$ -electron five-membered ring. Interestingly, this difference in aromaticity of the ground state will have an opposite effect on the  $T_1$  state's aromaticity, in line with the Baird aromaticity rule. These results demonstrate that the  $S_1$  and  $T_1$  energies can be precisely tuned by the nature of substituents. Assuming that the  $S_1$  and  $T_1$  states are described by the same electron configuration, except for the multiplicity difference, a particular EWG at the exocyclic position of a fulvene will have the same stabilizing effect on  $S_1$  as on  $T_1$  when compared to the parent fulvene as a reference.<sup>37</sup> Under this situation,  $\Delta E(S_1-T_1)$  will remain constant within the series as observed in Fig. 4.

In the  $R_2$  subset, fulvenes with strong EDGs still tend to lower the energies of the lowest-lying excited states compared to the parent fulvene, although there is no clear trend for the fulvenes with strong EWGs. Additionally, in this subset, the relationship between aromaticity and the nature of the substituents is reversed compared to the  $R_1$  subset (see Fig. S6<sup>†</sup>). Here, fulvenes with strong EDGs are more ground-state antiaromatic, while those with EWGs show enhanced aromaticity in the  $S_0$  state. In the final subgroup ( $R_3$ ), there is no significant correlation between substituent type and the energy values compared to the unsubstituted fulvene (see Fig. S7<sup>†</sup>). The relationship between aromaticity and the nature of the substituents is similar to that in the  $R_2$  subgroup, but with less variation in the  $\text{NICS}(\pm 1)_{zz,S_0}$  values, indicating smaller differences in aromaticity induced by the nature of the substituents in these positions.

To further validate the findings derived from the NICS and MCI descriptors, we conducted calculations using the gauge-including magnetically induced currents (GIMIC) method, which reinforced the initial observations. As shown in Table 1, a notable reduction in the range of  $\text{NICS}(\pm 1)_{zz}$  values between the most aromatic and the most antiaromatic fulvenes was observed for the ground state when progressing from the  $R_1$  to  $R_2$ , and  $R_3$  modifications (*i.e.*, ranges of the  $\text{NICS}(\pm 1)_{zz}$  amount to 28.8, 22.3, and 5.3 ppm for  $R_1$  to  $R_2$ , and  $R_3$ ). A similar trend was found for the MCI ranges, which decrease from 0.012 to 0.009 to 0.006 for  $R_1$  to  $R_2$ , and  $R_3$ , respectively. With the trend aligned according to the  $\text{NICS}(\pm 1)_{zz}$  values in  $S_0$ , GIMIC currents were calculated for the most aromatic and antiaromatic fulvenes within each subset, revealing differences in

Table 1 Range of the computed properties for fulvenes with differently substituted positions ( $R_1$ ,  $R_2$ , and  $R_3$ ). The table displays vertical excitation electronic energies ( $E_v(S_1)$  and  $E_v(T_1)$ ), the SF function value ( $f_{\text{SF}}$ ),  $\text{NICS}(\pm 1)_{zz}$ , and MCI values for the ground and triplet states, and the diradical character index ( $\gamma_0$ )

Parameter	$R_1$	$ \Delta R_1 $	$R_2$	$ \Delta R_2 $	$R_3$	$ \Delta R_3 $
$E_v(S_1)$ (eV)	[2.64, 4.15]	1.51	[2.57, 3.89]	1.32	[3.34, 3.89]	0.55
$E_v(T_1)$ (eV)	[1.86, 3.00]	1.14	[1.71, 2.64]	0.93	[2.20, 2.80]	0.60
$f_{\text{SF}}$ (eV)	[-1.86, -1.02]	0.84	[-1.39, -0.64]	0.75	[-1.78, -1.06]	0.72
$\text{NICS}(\pm 1)_{zz,S_0}$ (ppm)	[-19.6, 9.2]	28.8	[-12.7, 9.6]	22.3	[-7.7, -2.4]	5.3
$\text{NICS}(\pm 1)_{zz,T_1}$ (ppm)	[-8.4, 17.0]	25.4	[-11.2, 14.7]	25.9	[3.7, 16.3]	12.6
MCI ( $S_0$ )	[0.003, 0.034]	0.031	[0.006, 0.015]	0.009	[0.007, 0.013]	0.006
MCI ( $T_1$ )	[0.016, 0.038]	0.022	[0.009, 0.026]	0.017	[0.018, 0.023]	0.005
$\gamma_0$	[0.00, 0.06]	0.06	[0.01, 0.06]	0.05	[0.00, 0.04]	0.04



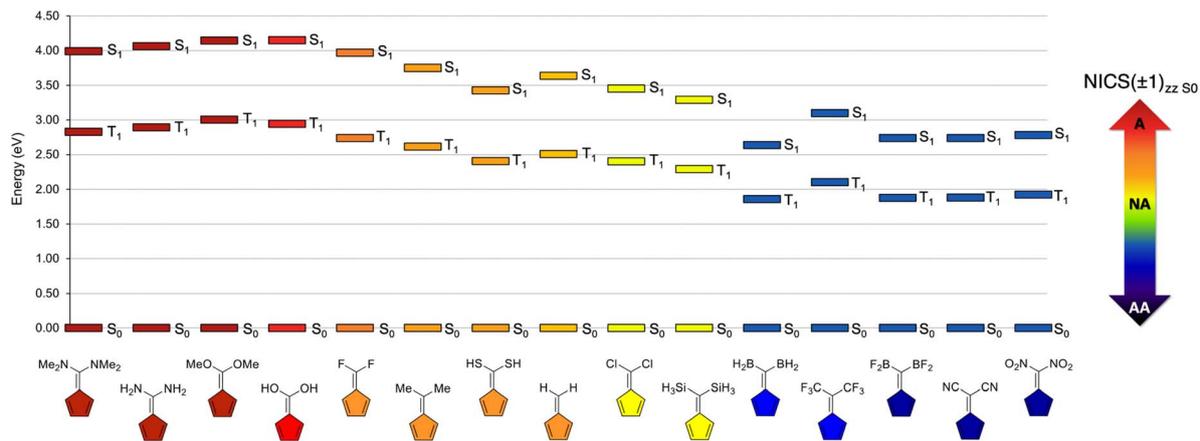


Fig. 4 Electronic energy of the different states of each  $R_1$ -substituted fulvene ordered according to  $\text{NICS}(\pm 1)_{zz,S_0}$  values.

Table 2 Correlation ( $R^2$  values) between the different aromaticity indices for the archetype fulvene set in the ground state ( $S_0$ )

	FLU	BOA	HOMA	BLA	$I_{\text{ring}}$	MCI	$\text{NICS}(\pm 1)_{zz}$
FLU	1						
BOA	0.866	1					
HOMA	0.928	0.787	1				
BLA	0.931	0.817	0.921	1			
$I_{\text{ring}}$	0.818	0.552	0.845	0.833	1		
MCI	0.859	0.681	0.881	0.839	0.938	1	
$\text{NICS}(\pm 1)_{zz}$	0.781	0.554	0.790	0.737	0.723	0.723	1

the net ring current strength ( $\text{RCS}_{\text{net}}$ ) between the most aromatic and most antiaromatic fulvene of 10.4, 8.3, and 1.4 nA  $\text{T}^{-1}$  for  $R_1$ ,  $R_2$ , and  $R_3$ , respectively (Fig. 5). We refer to the ESI (Fig. S11–S13)<sup>†</sup> for the GIMIC plots corresponding to the  $T_1$  state and the RCS profiles, respectively.

Overall, for the complete archetype fulvene set of 45 compounds, we observe a strong correlation between the low-lying excited-state energies and the degree of (anti)aromaticity in the ground state (Fig. 6), in line with previous findings for a different subset of di- and tetrasubstituted fulvene derivatives.<sup>37</sup> Fig. 6 shows that the  $T_1$  becomes stabilized for Hückel antiaromatic fulvenes. Given the good correlation between the  $\text{NICS}(\pm 1)_{zz}$  index computed for the  $S_0$  and  $T_1$  states (Fig. S10<sup>†</sup>), this implies that the  $T_1$  state becomes aromatic for Hückel antiaromatic fulvenes in the ground state. Accordingly, the presence of excited-state aromaticity appears to be a guiding concept to explain the stabilization of the  $T_1$  and  $S_1$  states induced by certain substituents on the fulvene moiety. It is important to note that similar trends in the variation of excitation energies as a function of the substituent are observed when using both vertical and adiabatic  $T_1$  excitation energies (Fig. S4<sup>†</sup>). This is important for the inverse design protocol, as bypassing excited-state geometry optimizations significantly reduces the computational cost to evaluate the target SF function (eqn (4)).

To further evaluate the accuracy of our TD-DFT protocol in determining the excitation energies of fulvene derivatives, we

conducted comparisons with both experimental data and multireference NEVPT2 calculations (see ESI<sup>†</sup>). The TD-DFT calculations at the M06-2X/6-311+G(d,p) level of theory show good agreement with the available experimental measurements with deviations ranging from 0.01 to 0.31 eV for both  $S_1$  and  $T_1$  energies (Table S1<sup>†</sup>). Similarly, our computed excitation energies align well with the NEVPT2  $S_1$  and  $T_1$  energies with deviations smaller than 0.3 eV for most of the investigated fulvenes (Table S2<sup>†</sup>). Although the overall trends are preserved, the deviations between TD-DFT and NEVPT2 excitation energies will inevitably impact the computed values of  $f_{\text{SF}}$ . Nevertheless, given the high computational cost of the inverse design protocol, TD-DFT remains a practical and effective choice for exploring a broader chemical compound space.

A symmetric fulvene has three distinct positions that can be functionalized, while an asymmetric fulvene has six functionalization sites. By exploring substitutions at a single position, we found that none of the 45 archetypal fulvenes meet the first energy-matching condition for SF. Thus, the computed values of  $f_{\text{SF}}$  are far from zero. Consequently, we propose that an effective SF fulvene candidate can be realized by leveraging the additive or even synergistic effects between the various functionalizable positions. Considering the six potential substitution sites in an asymmetric fulvene, a chemical compound space of approximately 11.4 million ( $15^6$ ) derivatives is estimated for our library containing 15 substituents. Examining such a vast number of fulvenes using traditional methods is exceedingly time-consuming. Therefore, we propose to use inverse molecular design to streamline the design process and save computational time. Even by exploring only a small part of the CCS, this protocol aids to identify optimal functionalizations to render fulvenes as promising chromophores for SF.

### Inverse design

Using traditional direct molecular design approaches, only small parts of the enormous chemical compound space are explored. In contrast, ID techniques investigate boarder regions, strategically leading to better structures compared to the starting one. Although inverse design may not find the



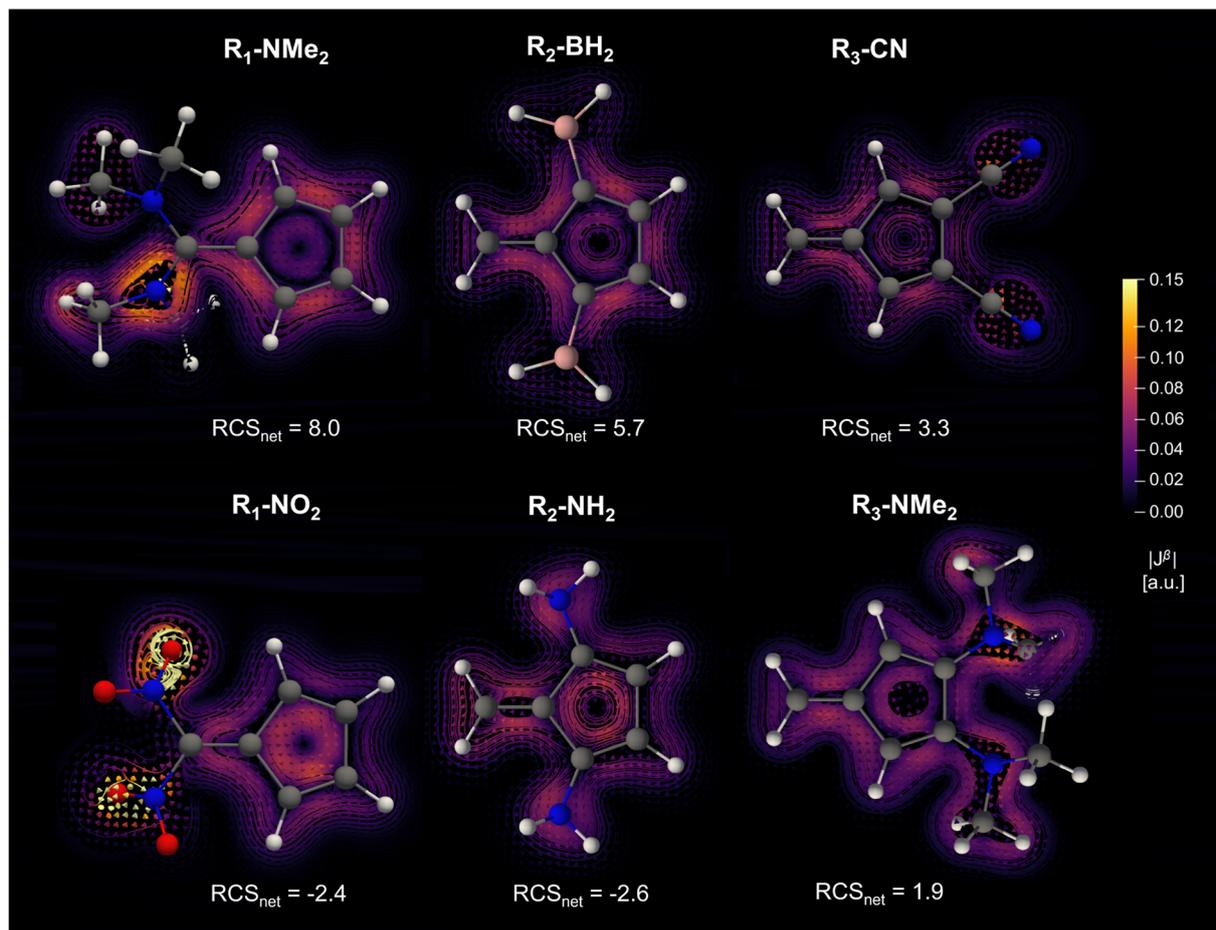


Fig. 5 Current density analysis of the most aromatic (first row) and most antiaromatic (second row) archetypal fulvenes in the ground state. Plot of vectors at 0.5 Å above the molecular plane with the integrated ring current strength through the integration plane, which bisects the carbon-carbon bond opposite the exocyclic position.  $RCS_{net}$  values are given in  $nA T^{-1}$ .

ultimate performer, its end structure still improves the target property.

In this work, we studied three different cases using ID techniques. Considering symmetric fulvenes, we explore fulvenes with 2 or 3 substituted positions. As a last case, asymmetric fulvenes are also considered, *i.e.*, having 6 modifiable sites (see Fig. 1). For the sake of simplicity, we will deeply analyse the results of the first BFS procedure with 2 substituted positions and discuss the remaining two cases more generally. All generated structures, their vertical energies of the lowest-lying excited states ( $S_1$  and  $T_1$ ), and the function value can be found in Table S15.† In the final part of this section, we will estimate the importance and significance of each functionalization towards the fulvene's performance by using a steepest descent approach.

**Case I: ID procedure with two substituted positions.** A random site order was implemented in the ID protocol. In the initial global iteration, the algorithm starts by modifying the  $R_2$  position while keeping the F substituent in the first position (Fig. 7). The  $BH_2$  substituent results in the most significant improvement, achieving a function value reduced to half the function of the parent fulvene ( $f_{SF} = -0.7$  eV). During the

subsequent step,  $BH_2$  is fixed in the  $R_2$  position, and variations are made to the  $R_1$  position, where the  $NO_2$  substituent performs most effectively ( $f_{SF} = -0.04$  eV). Moreover, positioning  $NO_2$  at the  $R_1$  induces a more substantial change in the function value than placing  $BH_2$  at  $R_2$ . The second global iteration begins by keeping the  $NO_2$  in  $R_1$  position and testing all the different substituents from the library in the  $R_2$  position, reaffirming  $BH_2$  as the optimal choice. The final step, optimizing the  $R_1$  position, is redundant as convergence has been reached.

As a mono-substituent in the fulvene structure at the  $R_1$  position,  $NO_2$  reduces the  $S_1$  energy more than the  $T_1$  energy compared to the parent fulvene (*cf.* green bars in Fig. 8). Conversely,  $BH_2$  as mono-substituent at  $R_2$  does not affect  $S_1$  but lowers  $T_1$  (*cf.* red bars in Fig. 8). Compared to  $H_{BH_2}$ , introducing  $NO_2$  at  $R_1$  lowers  $S_1$  substantially more than  $T_1$ , which agrees with the  $R_1 = NO_2$  mono-substitution result. Interestingly, the combination of  $NO_2$  at  $R_1$  and  $BH_2$  at  $R_2$  results in a more significant decrease in  $S_1$  compared to only having  $NO_2$  at  $R_1$ , to a greater extent than any other substituent (Fig. 7, BFS step 1). This reduction is even more pronounced for the  $T_1$  energy. This aligns with the third strategy to lower the SF



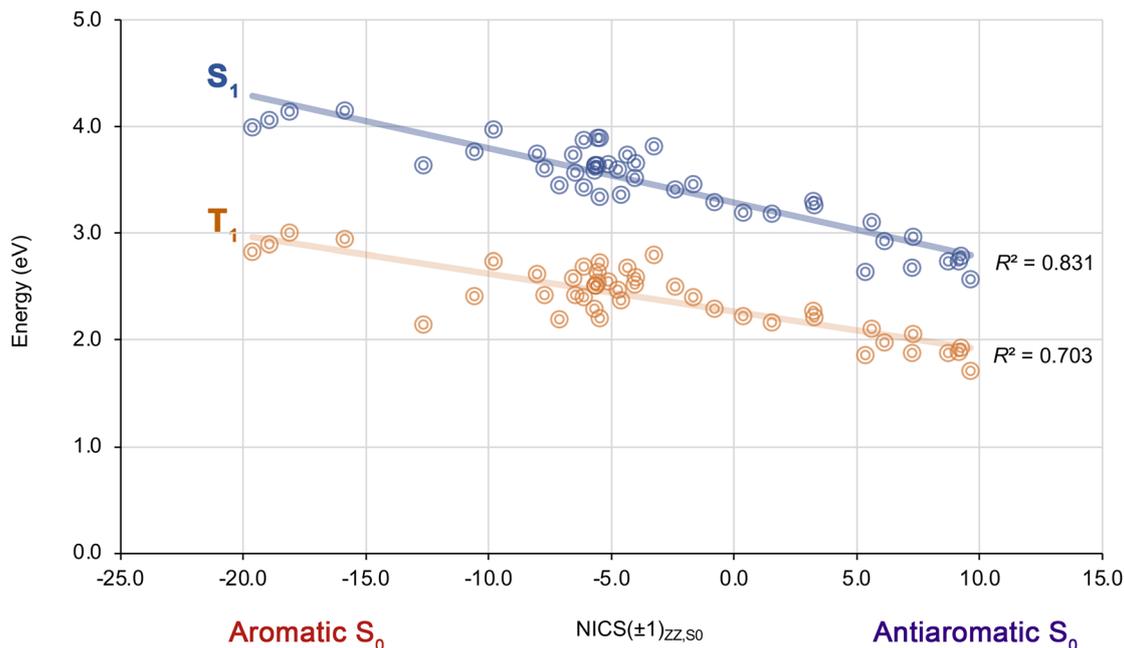


Fig. 6 Plots of excited-state energies of  $S_1$  and  $T_1$  against degree of (anti)aromaticity in  $S_0$  for the complete archetype fulvene set.

function as stated earlier, *i.e.*, low excitation energy values for both  $S_1$  and  $T_1$ .

This seeming synergy between both substituents becomes evident when comparing the  $\text{NO}_2\text{-BH}_2$  result with its additive version (equal to  $f_{\text{SF}}(\text{parent}) - \Delta f_{\text{SF}}(\text{parent}, \text{H-BH}_2) - \Delta f_{\text{SF}}(\text{parent}, \text{NO}_2\text{-H})$ ), indicated as purple bars for the former and yellow bars for the latter in Fig. 8. This synergistic effect is larger for  $S_1$  than for  $T_1$  and amounts to a significant 1.55 eV and 0.92 eV, respectively.

It turns out that this effect of the  $(\text{NO}_2, \text{BH}_2)$  couple is quite unique. Based on the mono-substitution values ( $R_1$ -fulvene archetype set), alternatives to  $\text{NO}_2$  for the  $R_1$  position include CN and  $\text{BF}_2$ . However, they do not show the same behavior. This can be nicely witnessed from Fig. 9, in which the  $S_1$  and  $T_1$  excitation energies of the last global iteration with either  $\text{NO}_2$  on

$R_1$  ( $\text{NO}_2\text{-R}_2$  series) or  $\text{BH}_2$  on  $R_2$  ( $\text{R}_1\text{-BH}_2$  series) are plotted against the energies of the corresponding archetype analogues ( $\text{H-R}_2$  and  $\text{R}_1\text{-H}$  series, respectively). Except for the  $S_1$  of  $\text{R}_1\text{-H}$  vs.  $\text{R}_1\text{-BH}_2$ , we can appreciate good linear correlations in the energy of the lowest excited states. The couple ( $\text{H-BH}_2, \text{NO}_2\text{-BH}_2$ ) is the only outlier. The synergistic effect is demonstrated by the large horizontal shift to lower values of  $S_1$  and  $T_1$ , indicated by red arrows in Fig. 9.

We observed that the algorithm converged very quickly due to few synergetic functional groups. We also affirmed that the  $R_1$  position is more critical than  $R_2$  for identifying a possible candidate for the SF process.

To assess whether we could further improve the function beyond the optimum of the first run, we conducted a second

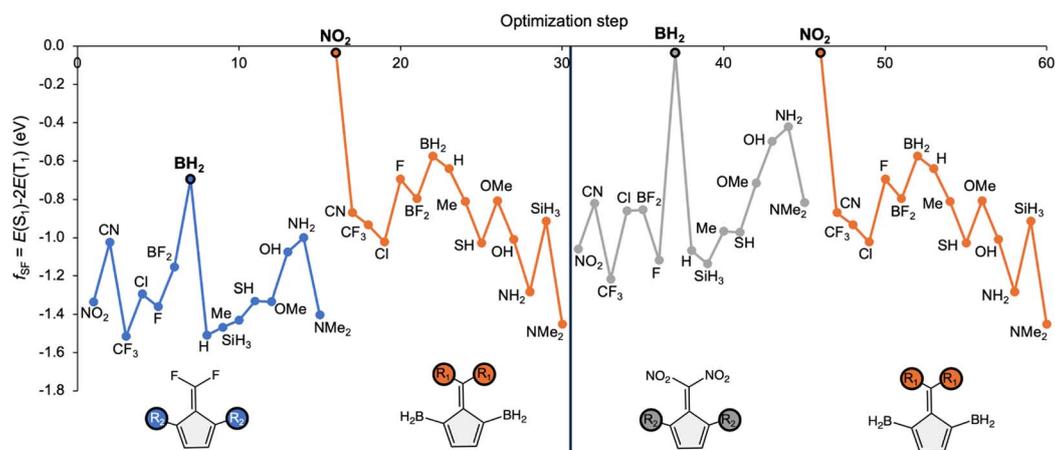


Fig. 7 Chronological site order of the BFS for maximizing the function  $f_{\text{SF}} = E(S_1) - 2E(T_1)$ , starting from a random symmetric structure with two substituted sites. The black line separates the global iterations. The site order varies in every global iteration. The current best substituent of each site optimization, selected for the next iteration, is circled in black.



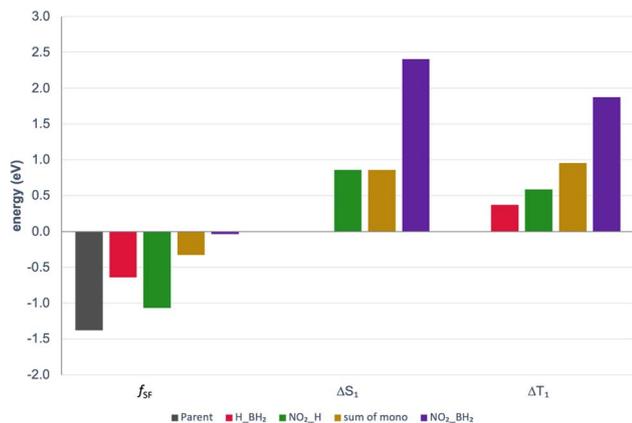


Fig. 8 Energy differences (eV) in the function value ( $f_{SF}$ ),  $S_1$  and  $T_1$  states for the parent,  $H\_BH_2$ ,  $NO_2\_H$ , addition of mono-substitutions, and  $NO_2\_BH_2$  fulvenes.

BFS run with 2 substituted positions starting with the  $BH_2\_NO_2$  symmetric fulvene (see Fig. S15<sup>†</sup>). We purposely selected  $NO_2$  for the  $R_2$  positions due to its dominant role on  $R_1$  as observed in the first BFS run. In this second BFS run, we can see that both global iterations are the same, meaning that the optimum was already identified in the first iteration step corresponding to  $BF_2\_NO_2$ . However, it is evident that the associated SF function value ( $-0.69$  eV) remains far from the ideal. The  $NO_2$  substituent at the  $R_2$  positions performs best but not substantially (*e.g.*,  $-0.80$  eV for  $BF_2\_BH_2$ ) and a similar conclusion applies to  $BF_2$  at the  $R_1$  positions (*e.g.*,  $-0.85$  eV for  $SH\_NO_2$ ).

As a mono-substituent at  $R_1$ ,  $BF_2$  lowers  $S_1$  more than  $T_1$ , comparable to  $NO_2$  and  $CN$ . However, due to the presence of  $NO_2$  on  $R_2$ , this behavior reverses with now a larger decrease in the  $T_1$  energy compared to  $H\_NO_2$  and the parent fulvene. By examining the correlation between the  $R_1\_H$  and  $R_1\_NO_2$  fulvene series in Fig. S16,<sup>†</sup> we identify  $BF_2\_NO_2$  as an outlier for the  $S_1$  energies with a smaller  $S_1$  decrease than expected (*cf.* shift

to the right). The other outlier, for both  $S_1$  and  $T_1$ , is  $R_1 = BH_2$ . Although some other functional groups (*e.g.*,  $CN$ ,  $NO_2$ , and  $SH$ ) can further decrease  $T_1$ , this is counterbalanced by an excessive reduction in  $S_1$ . Thus, achieving a balance between both excitation energy reductions is crucial for fulfilling the first energy-matching condition.

Conversely,  $NO_2$  as a mono-substituent at  $R_2$  has an almost negligible effect on  $S_1$  and moderately lowers the  $T_1$  energy compared to the parent fulvene. When combining the best performing substituents,  $S_1$  increases significantly with the inclusion of  $NO_2$  at  $R_2$  (compared to  $BF_2\_H$ ), with only  $NMe_2$  and  $CF_3$  causing a greater increase. A more moderate increase is observed for the  $T_1$  energy, which makes that the function value of  $BF_2\_NO_2$  is closer to zero than that of  $BF_2\_H$ . Note that there is still a reduction in the  $S_1$  energy and a notable decrease in the  $T_1$  energy compared to the unsubstituted fulvene. While larger decreases in the  $T_1$  energy can be obtained with other functional groups, they are again offset by an even larger decrease in the  $S_1$  energy. Comparing  $H\_R_2$  and  $BF_2\_R_2$  fulvene series, we observe that there is no correlation between the two series for either the  $S_1$  or  $T_1$  energies (see Fig. S16<sup>†</sup>). In conclusion, there is no evidence from this BFS run of a synergistic effect. The optimum merely found the best balance between  $S_1$  and  $T_1$  reduction.

**Case II: ID procedure with three substituted positions.** For this second case, similar to the first, we conducted two BFS runs starting from a random structure (see Fig. S17 and S18<sup>†</sup>). The first one commenced with the  $NO_2\_R_2\_CF_3$  structure, while the second began with  $R_1\_BH_2\_CF_3$ . In both runs, we observe rapid convergence to the same optimum,  $NO_2\_BH_2\_SH$ , which closely resembles the best performing fulvene from Case I. In addition, Case II suggests that sites  $R_3$  are the least important, as previously noted by Ottosson *et al.*,<sup>37</sup> particularly when  $NO_2$  and  $BH_2$  are present on  $R_1$  and  $R_2$ , respectively.

The first global iteration of the first BFS run (see Fig. S17<sup>†</sup>) starts by modifying the  $R_2$  positions, again identifying  $BH_2$  at  $R_2$  as optimal, consistent with the best-performing structure from

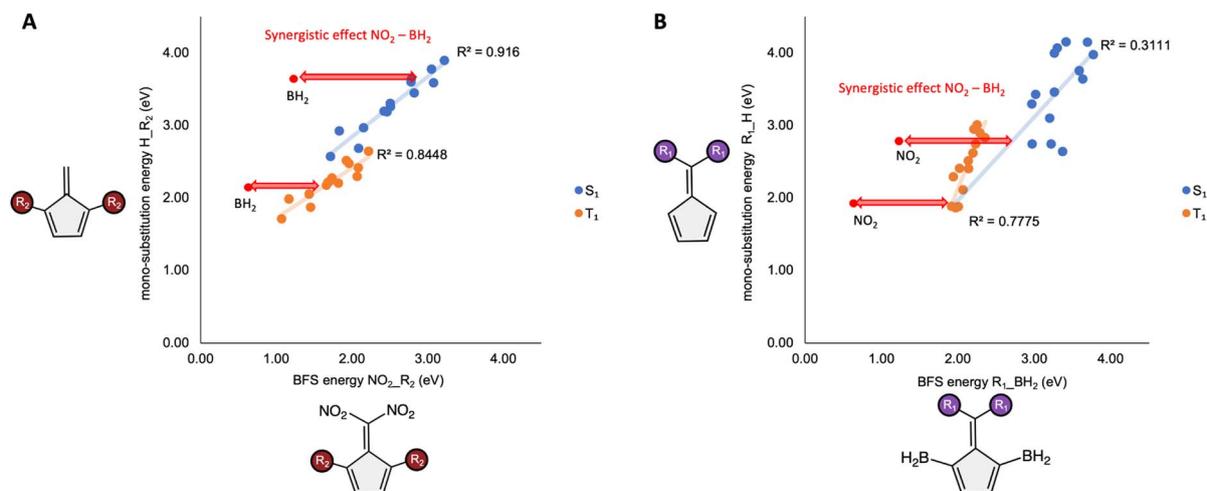


Fig. 9 Correlation between (A) mono-substituted fulvene series  $H\_R_2$  and  $NO_2\_R_2$  and (B) mono-substituted fulvene series  $R_1\_H$  and  $R_1\_BH_2$ . The outlier, indicated in red, is identified as the BFS optimum  $NO_2\_BH_2$ .  $S_1$  excitation energies are highlighted in blue, while the  $T_1$  excitation energies are marked in orange.



the two-site optimization. In the second step, the optimum has already been found. With NO<sub>2</sub> at R<sub>1</sub> and BH<sub>2</sub> at R<sub>2</sub>, all S<sub>1</sub> energies fall below 2 eV (except for NO<sub>2</sub>\_BH<sub>2</sub>\_NO<sub>2</sub>). Removing NO<sub>2</sub> from R<sub>1</sub> increases the S<sub>1</sub> energy by 1.3 eV, on average, while removing BH<sub>2</sub> from R<sub>2</sub> results in an average increase of 0.6 eV. Additionally, having NO<sub>2</sub> at R<sub>1</sub> and BH<sub>2</sub> at R<sub>2</sub> results in the T<sub>1</sub> energy being at or below 1 eV (again except for NO<sub>2</sub>\_BH<sub>2</sub>\_NO<sub>2</sub>). A similar effect is observed as with S<sub>1</sub> when either NO<sub>2</sub> or BH<sub>2</sub> is absent. This is a remarkable result, as none of the Case I fulvenes have these characteristics, with the exception of NO<sub>2</sub>\_BH<sub>2</sub>\_H. The substantial reduction of both S<sub>1</sub> and T<sub>1</sub> energies seen for the NO<sub>2</sub>\_BH<sub>2</sub>\_R<sub>3</sub> series has as a consequence that 11 out of 15 R<sub>3</sub>-substituents (or 73%) have a function value larger than -0.25 eV. Furthermore, functionalizing R<sub>3</sub> has a minor impact on the  $f_{\text{SF}}$  value, with a minimum–maximum range of only 0.6 eV (except for the outlier NO<sub>2</sub>\_BH<sub>2</sub>\_NO<sub>2</sub>). Substituents such as SH, OMe, BH<sub>2</sub>, OH, and CN yield a SF function value below that of the unsubstituted R<sub>3</sub>, with the first three having very low T<sub>1</sub> energies.

Although our focus was on maximizing the function, to see where its boundary lies, our maximum remains close to the ideal function value of 0 eV. During the two optimization runs, we identified 13 unique compounds within a 0.5 eV range around zero (*i.e.*, -0.25 eV to 0.25 eV). Nonetheless, it is evident that for structures with NO<sub>2</sub> at R<sub>1</sub> and BH<sub>2</sub> at R<sub>2</sub>, substitution at R<sub>3</sub> is unnecessary to fulfill the first strategy. For the three-site substitution, no clear correlation with the archetype set trends is observed. Notably, NO<sub>2</sub>\_BH<sub>2</sub>\_SH stands out as an outlier in terms of S<sub>1</sub> and T<sub>1</sub> energies compared to all other R<sub>1</sub>\_BH<sub>2</sub>\_SH structures.

**Case III: ID procedure with six substituted positions (asymmetric structures).** Following the procedure used in the two previous cases, we conducted two complete BFS runs, this time considering the fulvene structure without symmetry, which allows for six distinct positions to be substituted (see Fig. S19 and S20†). The first one was initiated using a random structure, while the second began from a structure containing one of the least visited substituents from the first BFS run for each position: NMe<sub>2</sub>\_SiH<sub>3</sub>\_Me\_OH\_NMe<sub>2</sub>\_OMe. Note that in the second BFS run we minimize the SF function in absolute value (Fig. 10). As with the previous cases, optima are reached quickly. The optima from both BFS runs (BF<sub>2</sub>\_NH<sub>2</sub>\_NH<sub>2</sub>\_SiH<sub>3</sub>\_NH<sub>2</sub>\_CN and NO<sub>2</sub>\_OH\_OH\_CN\_H\_BH<sub>2</sub>, with SF function values -0.20 eV and 0.00 eV, respectively) include substituents we have discussed previously (see optimal fulvenes structures in Fig. 11). An asymmetric fulvene structure increases the potential for discovering viable SF candidates, even taking into account that the property landscape contains more local optima.

From the three cases we investigated, it is important to note that maximizing the SF function enables the identification of fulvenes suitable for the SF process. For example, in Case II we encountered several fulvenes, *e.g.*, NO<sub>2</sub>\_BH<sub>2</sub>\_OH with  $f_{\text{SF}} = 0.00$  eV and NO<sub>2</sub>\_BF<sub>2</sub>\_SH with  $f_{\text{SF}} = 0.04$  eV, exhibiting function values closer to zero than our BFS optimum. Therefore, to prevent large overshooting of the 0 eV objective when there is more functionalization flexibility as in Case III, we decided to improve our BFS protocol by changing the optimization

objective to the minimization of the absolute value of the SF function (eqn (4), in absolute value) in our final run. Moreover, by minimizing  $|f_{\text{SF}}|$  we also reduce the possibility of having triplet instability problems. Finally, also local screening of other functionalized fulvenes was carried out to enlarge our fulvene database and include larger portions of CCS.

### Steepest descent

To better understand to what extent the substituents of asymmetric fulvene structures contribute to the SF function, we employed a steepest descent approach. This method involves following the path of the largest gradient, in a discrete sense, beginning with the parent fulvene and progressing to the optimal fulvene by introducing the most impactful functionalization of the structure first. The process requires six steps to reach the minimum, corresponding to the six substitution sites involved in the optimization.

Table 3 provides detailed information, including S<sub>1</sub> and T<sub>1</sub> energies used to decompose the SF function minimum based on the position of functionalization of one of our best performing fulvenes: NO<sub>2</sub>\_Me\_NMe<sub>2</sub>\_CF<sub>3</sub>\_NH<sub>2</sub>\_NO<sub>2</sub> (Fig. 12). Two measures have been included. The first measure, % Imp(p → b), tells us for each structure to which degree the 100% improvement (going from parent to best performing structure) has been reached. In order to quantify the contribution of a particular site/functionalization combination, the difference in % Imp(p → b) between the current combination and the previous best can be taken. The second measure, % Best it., shows how well a certain site/functionalization combination performs in that iteration step, compared to the iteration's best.

Compared to the parent fulvene, all substitutions result in a decrease in the SF function. Following the steepest descent path, the first three substitutions yield a similar improvement to the parent structure, approximately 23–27% (Fig. 12). Note that these are also the three best performing in iteration step 1. Initially, there is a preference for an EDG, followed by an EWG, and then another EDG. The need for EDG/EWG combination can also be witnessed from iteration step 2, where these types of associations are preferred over EDG/EDG couples. The fourth functionalization, an EWG, brings the SF function value to -0.16 eV. It is important to note that there is no clear preference for a single site/functionalization for most steps, having always one or more alternatives present at minimally 80% of the iteration's best. Moreover, we can see that there is a slight preference for functionalizing one of each of the three different substituted positions first (as if we had symmetry in the structure), and not all of them being at the same side. Despite the exocyclic positions being expected to have the highest impact on the SF function value, they are not the preferred positions for the first substitution.

When we break down the function contributions to the changes in the S<sub>1</sub> and T<sub>1</sub> energies relative to the current best fulvene at each step (see Fig. S21†), both S<sub>1</sub> and T<sub>1</sub> energies decrease substantially for the first two functionalizations, with the former being more pronounced. In the next two functionalizations, S<sub>1</sub> energies increase while T<sub>1</sub> energies continue to



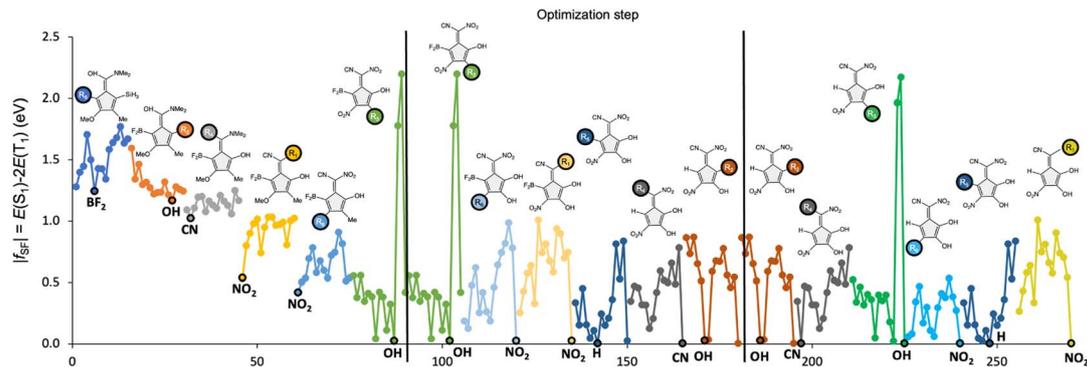


Fig. 10 Chronological site order of the BFS for minimization of the function,  $|f_{\text{SF}}| = E(S_1) - 2E(T_1)$ , starting from  $\text{NMe}_2\text{-SiH}_3\text{-Me-OH-NMe}_2\text{-OMe}$  fulvene with six substituted positions, without taking into account the symmetry of the structure. The black line separates the global iterations. The site order varies in every global iteration. The optimum of each site optimization, selected for the next iteration, is circled in black. There are four fulvenes ( $\text{NO}_2\text{-OH-NMe}_2\text{-CN-BF}_2\text{-NO}_2$ ,  $\text{NO}_2\text{-OH-NH}_2\text{-CN-BF}_2\text{-NO}_2$ ,  $\text{NO}_2\text{-OH-NMe}_2\text{-CN-H-NO}_2$ , and  $\text{NO}_2\text{-OH-NH}_2\text{-CN-H-NO}_2$ ) that present triplet instabilities and should not be considered as potential SF chromophores.

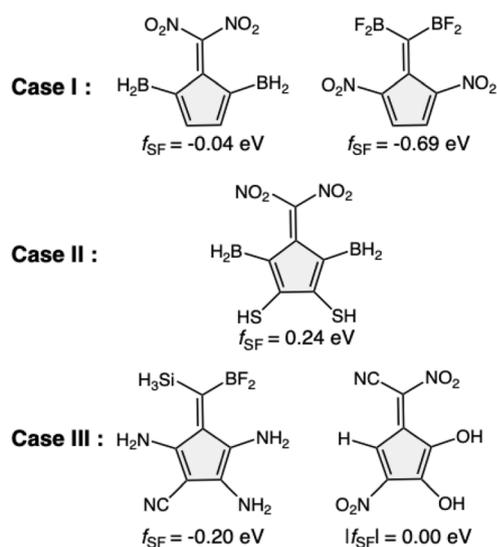


Fig. 11 Optimal fulvenes structures for all the three cases of BFS procedures. The optimum of Case II was the same in the two different runs.

decrease. Despite the very different changes in  $S_1$  and  $T_1$ , the third functionalization has a similar effect on the SF function as the first two. Think back on the three strategies that can be considered to bring the SF function value close to 0 eV: (i) significantly increasing  $E(S_1)$  relative to  $E(T_1)$ , (ii) decreasing  $E(T_1)$  while minimizing the impact on  $E(S_1)$ , or (iii) decreasing both  $E(S_1)$  and  $E(T_1)$ , with a sufficiently large reduction in  $E(T_1)$ . According to these options, the different functionalizations, following the order in the graph, adopt the subsequent strategy to lower the SF function to zero: third, third, second, first, third, and third, respectively.

In step 1, a large reduction of  $S_1$  energy is favored, but it should be accompanied by a sufficiently large decrease in  $T_1$  energy. This holds true for the second step as well. In the third step, all potential functionalizations increase the  $S_1$  energy, which is favorable for a reduction in the function value, further

enhanced when there is a reduction in the  $T_1$  energy. In the fourth step, a significant increase in  $S_1$  energy combined with a small decrease in  $T_1$  is preferred over a large reduction in  $T_1$ , which seems insufficient to counterbalance the decrease in  $S_1$  (cf.  $\text{NO}_2$  on  $R_6$  vs. Me on  $R_2$  and  $\text{CF}_3$  on  $R_4$ ).

We can also observe that the inclusion of functionalizations in each step can modify how a new substituent will affect the  $S_1$  and  $T_1$  energies. One example is  $\text{NMe}_2$ , which in step 1 had a significantly reducing effect on  $S_1$  and  $T_1$ , in step 2 had a negligible effect, and finally, in the third step increased  $S_1$  while decreasing  $T_1$ . This demonstrates three different impacts due to the inclusion of other functional groups on the fulvene scaffold. Another example is Me, which exhibits opposite behavior in the third step compared to all other steps. Another steepest descent example is described in the ESI.†

### Database analysis

Following the completion of all BFS runs, and incorporating the structures generated from local CCS screenings, we constructed a fulvene database containing 1658 systems. All energetic data regarding the database can be found in the ESI (see Table S15).† As mentioned before, a promising SF chromophore is characterized by having  $|f_{\text{SF}}|$  values close to 0 eV. Considering the plausible errors of the computational methods on the evaluation of the  $S_1$  and  $T_1$  excitation energies, we narrowed our focus to fulvene derivatives with an absolute value of the SF function ( $|f_{\text{SF}}|$ ) between 0 and 0.5 eV, resulting in 495 systems within this range (30% of our database). From this subset, we selected the top 100 candidates based on their proximity to  $|f_{\text{SF}}| = 0.00 \text{ eV}$  (see Fig. 13), corresponding to 6% of the full dataset.

To identify the most promising SF candidates, we applied a rigorous set of selection criteria. First, fulvenes with a triplet ground state were excluded, narrowing the pool to 98 systems. Next, we applied the second thermodynamic condition for SF to avoid intersystem crossing (*i.e.*,  $E(T_2) > E(S_1)$ ), further refining the selection to 65 fulvene derivatives. Lastly, we filtered the derivatives based on the triplet energies to ensure efficient injection of the triplet exciton into the silicon semiconductor



**Table 3** Steepest descent of NO<sub>2</sub>-Me-NMe<sub>2</sub>-CF<sub>3</sub>-NH<sub>2</sub>-NO<sub>2</sub> fulvene. % Imp(p → b) is the percentage-wise improvement of each steepest descent step taking the parent as the reference structure. % Best it. stands for the percentage-wise improvement of each steepest descent step taking the current best in each step as the reference

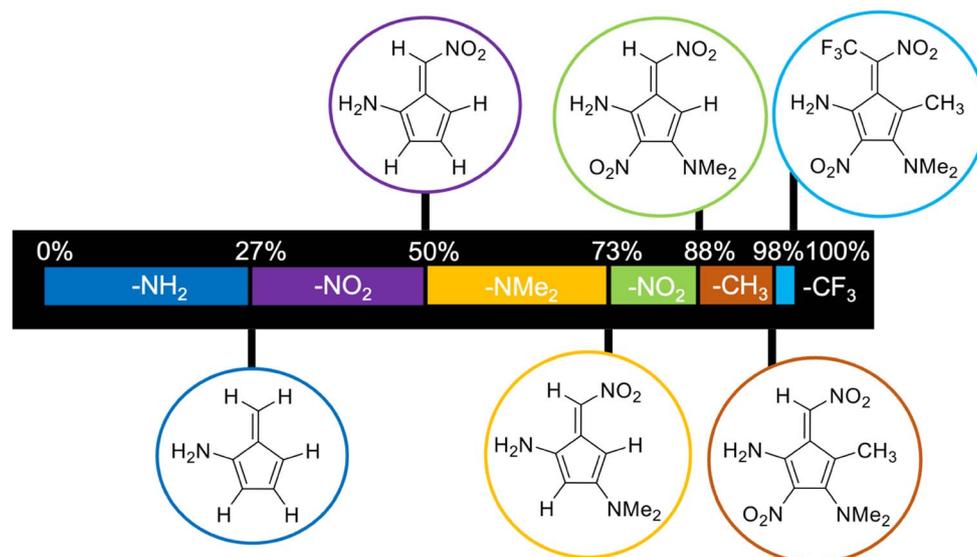
Step	Fulvene	S <sub>1</sub> (eV)	T <sub>1</sub> (eV)	f <sub>SF</sub> (eV)	% Imp(p → b)	% Best it.
Parent	H_H_H_H_H_H	3.64	2.51	-1.37		
1	NO <sub>2</sub> _H_H_H_H_H	2.92	2.03	-1.14	17	62
	H_Me_H_H_H_H	3.46	2.38	-1.30	5	19
	H_H_NMe <sub>2</sub> _H_H_H	3.36	2.22	-1.08	21	80
	H_H_H_CF <sub>3</sub> _H_H	3.40	2.33	-1.25	8	32
	<b>H_H_H_H_NH<sub>2</sub>_H</b>	<b>2.92</b>	<b>1.96</b>	<b>-1.00</b>	<b>27</b>	<b>100</b>
	H_H_H_H_H_NO <sub>2</sub>	3.59	2.45	-1.30	5	18
2	<b>NO<sub>2</sub>_H_H_H_NH<sub>2</sub>_H</b>	<b>2.20</b>	<b>1.44</b>	<b>-0.69</b>	<b>50</b>	<b>100</b>
	H_Me_H_H_NH <sub>2</sub> _H	2.82	1.90	-0.98	28	56
	H_H_NMe <sub>2</sub> _H_NH <sub>2</sub> _H	2.93	1.97	-1.01	26	53
	H_H_H_CF <sub>3</sub> _NH <sub>2</sub> _H	2.50	1.62	-0.73	47	93
	H_H_H_H_NH <sub>2</sub> _NO <sub>2</sub>	2.93	1.89	-0.85	38	76
	NO <sub>2</sub> _Me_H_H_NH <sub>2</sub> _H	2.32	1.57	-0.82	40	55
3	<b>NO<sub>2</sub>_H_NMe<sub>2</sub>_H_NH<sub>2</sub>_H</b>	<b>2.27</b>	<b>1.32</b>	<b>-0.37</b>	<b>73</b>	<b>100</b>
	NO <sub>2</sub> _H_H_CF <sub>3</sub> _NH <sub>2</sub> _H	2.26	1.44	-0.63	54	74
	NO <sub>2</sub> _H_H_H_NH <sub>2</sub> _NO <sub>2</sub>	2.24	1.39	-0.54	60	83
	NO <sub>2</sub> _Me_NMe <sub>2</sub> _H_NH <sub>2</sub> _H	2.02	1.11	-0.21	85	96
	NO <sub>2</sub> _H_NMe <sub>2</sub> _CF <sub>3</sub> _NH <sub>2</sub> _H	2.10	1.20	-0.30	78	88
	<b>NO<sub>2</sub>_H_NMe<sub>2</sub>_H_NH<sub>2</sub>_NO<sub>2</sub></b>	<b>2.40</b>	<b>1.28</b>	<b>-0.16</b>	<b>88</b>	<b>100</b>
5	<b>NO<sub>2</sub>_Me_NMe<sub>2</sub>_H_NH<sub>2</sub>_NO<sub>2</sub></b>	<b>2.17</b>	<b>1.10</b>	<b>-0.03</b>	<b>98</b>	<b>100</b>
	NO <sub>2</sub> _H_NMe <sub>2</sub> _CF <sub>3</sub> _NH <sub>2</sub> _NO <sub>2</sub>	2.36	1.28	-0.20	86	87
6	<b>NO<sub>2</sub>_Me_NMe<sub>2</sub>_CF<sub>3</sub>_NH<sub>2</sub>_NO<sub>2</sub></b>	<b>2.08</b>	<b>1.04</b>	<b>0.00</b>	<b>100</b>	<b>100</b>

( $E(T_1) \geq 1.11$  eV), yielding a final set of 17 fulvenes fulfilling the strict energy conditions for SF-enhanced silicon solar cells (Fig. 14), representing approximately 1% of the original database.

Our inverse design strategy to find new SF chromophores offers a satisfactory success rate compared to previously reported high-throughput approaches.<sup>31,43,45,47-49</sup> As excitation energies and success rates are inherently dependent on the choice of functional and considering the stringent energy requirements that must be fulfilled, the resulting success rate

remains significant and highlights the effectiveness of our inverse design strategy within these realistic limitations.

Several computational screening studies have been reported in the singlet fission field, although most are based on large and structurally diverse molecular databases. Only a few focus on a common core with systematic substitution, as done in the present work. For example, Padula *et al.*<sup>43</sup> investigated 40 K structurally diverse compounds using traditional screening techniques and reported a 3% success rate, including candidates that narrowly missed the first SF criterion by up to 0.4 eV.



**Fig. 12** Scheme of the steepest descent of one of our best performing structures ( $f_{SF} = 0.00$  eV): NO<sub>2</sub>-Me-NMe<sub>2</sub>-CF<sub>3</sub>-NH<sub>2</sub>-NO<sub>2</sub>.



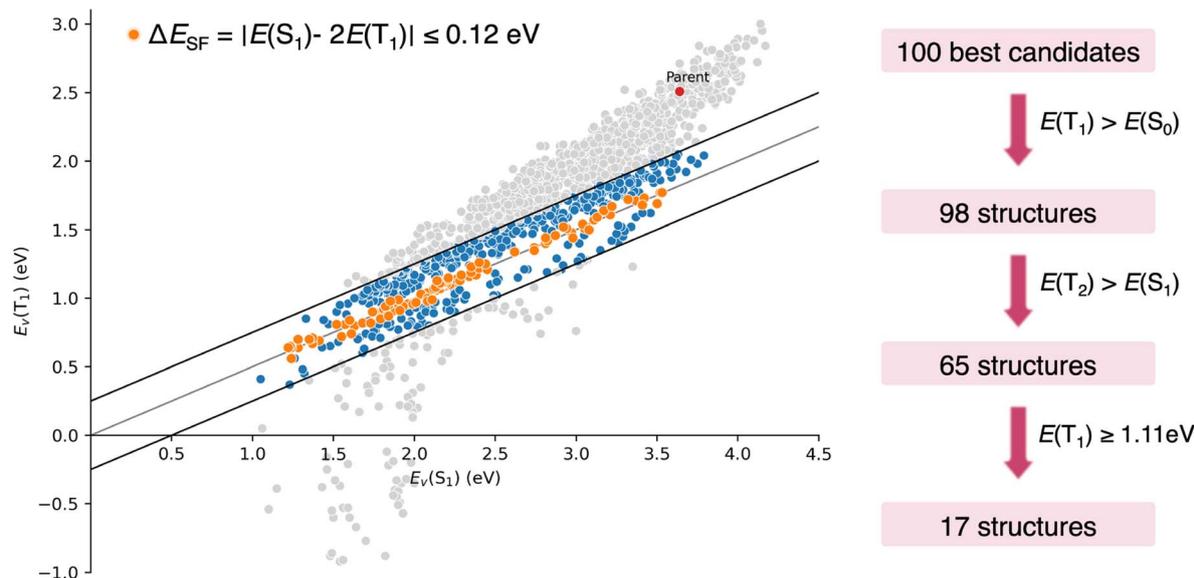


Fig. 13 Vertical energy of  $T_1$  state plotted against vertical energy of  $S_1$  for the complete fulvenes database. Black lines delimitate the area where  $|f_{SF}| \leq 0.5$  eV. Orange dots represent the 100 best candidates ( $|f_{SF}| \leq 0.12$  eV), further filtered according to the scheme on the right. The parent fulvene has been highlighted in red.

In another study, López-Carballeira and Polcar<sup>47</sup> screened nearly 30 K species, identifying 254 molecules (0.9%) that met the first two energetic criteria for singlet fission ( $E(S_1) \geq 2E(T_1)$  and  $E(S_1) \leq E(T_2)$ ). After applying additional practical

constraints such as synthetic accessibility and stability, the final set was reduced to 24 compounds (0.1%), with amino-anthraquinone derivatives emerging as particularly promising. Recently, Schaufelberger *et al.* presented an uncertainty-

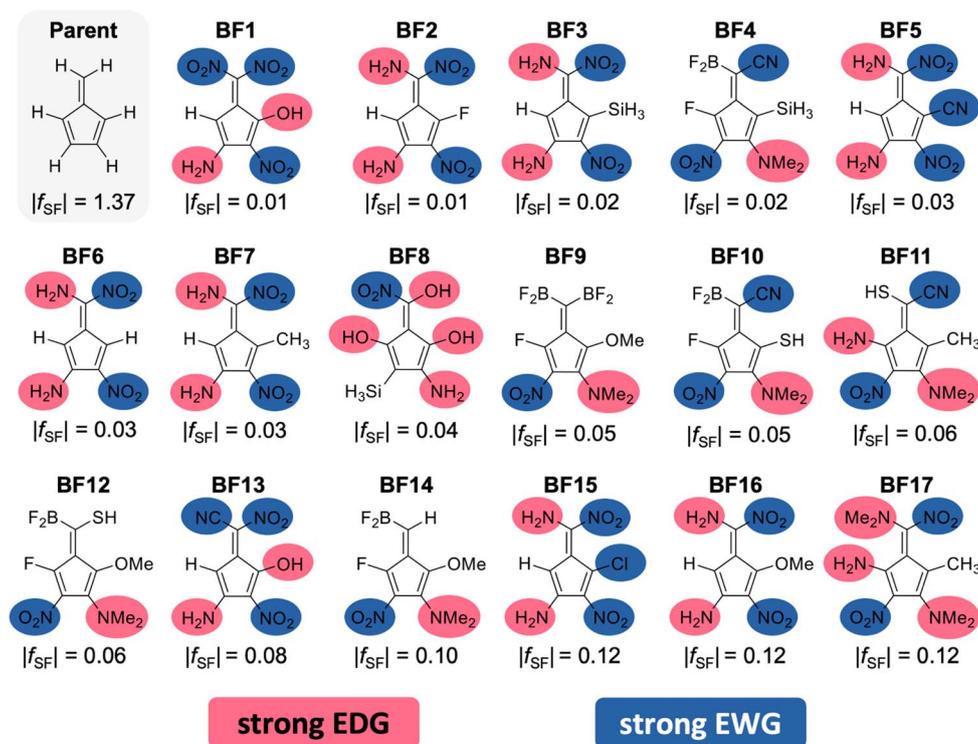


Fig. 14 Structures of the top fulvene candidates identified in this study are depicted. The parent fulvene is included for reference. Each structure is labeled with the corresponding absolute value of SF function (eV). Additionally, strong EDGs and EWGs are highlighted in pink and blue, respectively.



Table 4 Aromaticity indices for the 17 best candidates in the singlet ground state ( $S_0$ ) at M06-2X/6-311+G(d,p) level of theory

Label	System	FLU	BOA	HOMA	BLA	$I_{ring}$	MCI	NICS(0) <sub>zz</sub>	NICS( $\pm 1$ ) <sub>zz</sub>
BF1	NO <sub>2</sub> _OH_NO <sub>2</sub> _NO <sub>2</sub> _H_NH <sub>2</sub>	0.049	0.283	-0.431	0.089	0.011	0.005	33.5	10.1
BF2	NO <sub>2</sub> _F_NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.037	0.316	0.142	0.079	0.016	0.013	18.4	-2.0
BF3	NO <sub>2</sub> _SiH <sub>3</sub> _NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.034	0.333	0.085	0.083	0.018	0.014	20.4	-4.5
BF4	CN_SiH <sub>3</sub> _NMe <sub>2</sub> _BF <sub>2</sub> _F_NO <sub>2</sub>	0.047	0.372	-0.319	0.097	0.013	0.007	28.5	7.4
BF5	NO <sub>2</sub> _CN_NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.029	0.279	0.311	0.070	0.019	0.017	15.5	-7.3
BF6	NO <sub>2</sub> _H_NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.034	0.338	0.114	0.083	0.018	0.014	21.5	-3.2
BF7	NO <sub>2</sub> _CH <sub>3</sub> _NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.035	0.336	0.004	0.087	0.016	0.013	22.4	-1.4
BF8	OH_OH_NH <sub>2</sub> _NO <sub>2</sub> _OH_SiH <sub>3</sub>	0.066	0.444	-0.795	0.117	0.009	0.002	32.3	10.4
BF9	BF <sub>2</sub> _OMe_NMe <sub>2</sub> _BF <sub>2</sub> _F_NO <sub>2</sub>	0.053	0.381	-0.381	0.108	0.011	0.007	27.1	8.6
BF10	CN_SH_NMe <sub>2</sub> _BF <sub>2</sub> _F_NO <sub>2</sub>	0.048	0.354	-0.284	0.095	0.012	0.008	28.6	8.2
BF11	CN_CH <sub>3</sub> _NMe <sub>2</sub> _SH_NH <sub>2</sub> _NO <sub>2</sub>	0.044	0.254	-0.212	0.081	0.011	0.007	32.2	8.4
BF12	SH_Ome_NMe <sub>2</sub> _BF <sub>2</sub> _F_NO <sub>2</sub>	0.045	0.328	-0.064	0.089	0.014	0.011	19.4	2.6
BF13	NO <sub>2</sub> _OH_NO <sub>2</sub> _CN_H_NH <sub>2</sub>	0.048	0.285	-0.395	0.089	0.011	0.005	34.3	10.8
BF14	H_Ome_NMe <sub>2</sub> _BF <sub>2</sub> _F_NO <sub>2</sub>	0.053	0.378	-0.376	0.106	0.011	0.008	25.8	7.5
BF15	NO <sub>2</sub> _Cl_NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.033	0.316	0.140	0.079	0.017	0.014	19.7	-3.1
BF16	NO <sub>2</sub> _OMe_NO <sub>2</sub> _NH <sub>2</sub> _H_NH <sub>2</sub>	0.039	0.302	0.130	0.077	0.015	0.012	18.7	-2.2
BF17	NO <sub>2</sub> _CH <sub>3</sub> _NMe <sub>2</sub> _NMe <sub>2</sub> _NH <sub>2</sub> _NO <sub>2</sub>	0.045	0.256	-0.196	0.081	0.010	0.007	32.8	9.1

controlled genetic algorithm powered by machine learning to efficiently explore a chemical space of over  $10^{12}$  possible molecules for the discovery of novel singlet-fission materials, a task impossible with high-throughput screening.<sup>49</sup> Despite focusing on a narrower chemical space, the proportion of promising candidates identified in our study underscores the effectiveness of the inverse design strategy implemented here.

Having identified the most promising SF fulvene candidates through the combinatorial CCS, we further characterize the underlying electronic structure of this subset by computing the vertical energies of the lowest-lying excited states ( $E_v(S_1)$ ,  $E_v(T_1)$ , and  $E_v(T_2)$ ), analyzing the diradical character (Table S16<sup>†</sup>), and assessing their aromatic character in both the ground state and lowest triplet state (see Tables 4 and S17<sup>†</sup>). Within this subset, the  $T_1$  vertical excitation energies show minimal variation, ranging from 1.11 to 1.26 eV, while the  $S_1$  vertical energies display slightly more variability within a range from 2.14 to 2.45 eV, resulting in a  $\Delta(S_1 - T_1)$  range of 1.01 to 1.24 eV. Fig. 15 clearly illustrates that the  $T_1$  state is positioned midway between the  $S_1$  and  $S_0$  states, in agreement with our applied selection rule.

To further validate the accuracy of the TD-DFT predictions, we computed NEVPT2 excitation energies for the top 17 candidates (Table S3<sup>†</sup>). The deviation between TD-DFT and NEVPT2 is generally larger for the triplet states (0.35–0.68 eV) than for the singlets (0.06–0.39 eV), which results in a  $\Delta(S_1 - T_1)$  range of 0.25 to 0.90 eV. These discrepancies highlight the known limitations of TD-DFT in capturing multiconfigurational effects, which are better described within the multireference NEVPT2 framework.<sup>114</sup> Nevertheless, TD-DFT reliably captures the relative trends in  $S_1$  and  $T_1$  excitation energies across the series with  $R^2$  values of 0.902 and 0.964, respectively (Fig. S3<sup>†</sup>), making it suitable for the efficient screening of large chemical spaces in the context of inverse design.

Using the Yamaguchi method,<sup>87,88</sup> we evaluated the diradical character index  $y_0$  which show values between 0.03 and 0.10, indicating minimal diradical character with zero spin contamination for all candidates (Table S16<sup>†</sup>). While open-shell diradical character might reduce photochemical stability, other factors such as functionalization and steric protection also play important roles in determining the stability and reactivity of the

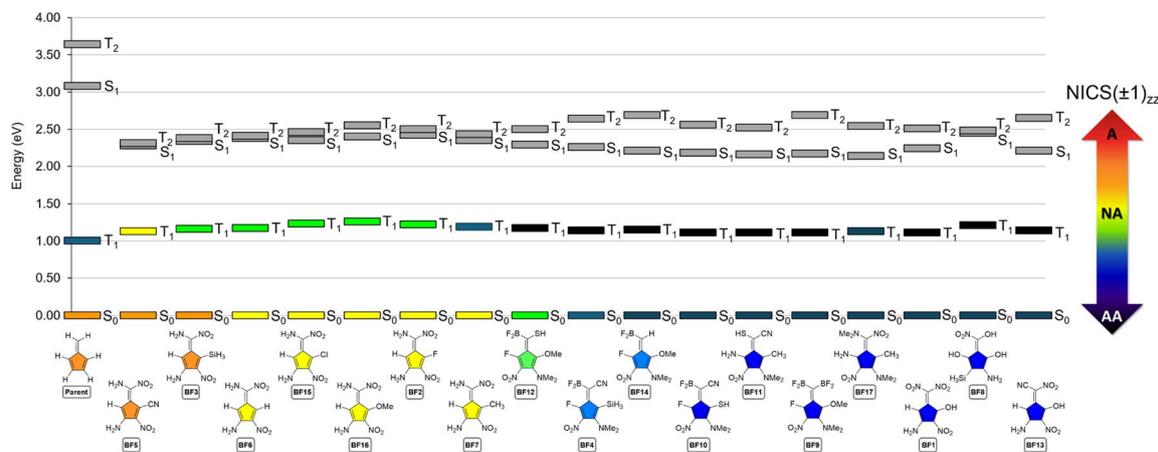


Fig. 15 Electronic energy of the different states of the best 17 SF candidates ordered according to NICS( $\pm 1$ )<sub>zz</sub>.  $S_0$ ,  $S_0$  and  $T_1$  states are colored according their respective aromatic character.



proposed fulvene derivatives. For example, the combination of exocyclic ( $R_1$ ) and endocyclic ( $R_6$ ) groups with opposite electronic properties can markedly increase the aromaticity and stability of fulvenes.<sup>61</sup>

Further analysis of the electronic and geometric aromaticity indices revealed that the ground state of the best fulvene derivatives predominantly exhibits non-aromatic character or weak antiaromaticity (Table 4). According to GIMIC calculations, we observe similar patterns in the paramagnetic and diamagnetic currents of the five-membered ring in these derivatives (Fig. S26–S29†). Upon integration of the induced ring currents, the net ring current strength values range from  $-2.5$  to  $3.7$  nA T<sup>-1</sup> in  $S_0$  and from  $-1.1$  to  $3.8$  nA T<sup>-1</sup> in the  $T_1$  state (see Fig. S30–S33†). These results prove the non-aromatic nature of the examined fulvenes. Nevertheless, the magnetic NICS( $\pm 1$ )<sub>zz</sub> index show a large variation in aromaticity among the candidates, as shown in Fig. 15. Moving to the aromaticity of the lowest triplet state, all the aromaticity indices, except NICS( $\pm 1$ )<sub>zz</sub>, indicate that  $T_1$  is more aromatic than the ground state (see Fig. 15, S24, S25 and Table S17†).

Concerning the nature of the substituents of the most promising fulvenes (Fig. 14), we establish a general condition for the fulvene substitution pattern required for SF. The presence of a strong EDG paired with a potent EWG in the  $R_3$  position is desired, if not indispensable, to satisfy the energy-matching conditions for SF. The exocyclic  $R_1$  position has a strong preference for at least one strong EWG, in many cases coupled with a strong EDG at  $R_1$ . The less influential positions are  $R_2$ , for which even a hydrogen atom can sufficiently fulfill this role.

By means of our ID computational study the best fulvene candidates for single fission have been identified, but the preparation of functionalized fulvenes poses several challenges. These compounds are often highly reactive,<sup>8,23</sup> and their synthesis requires precise control over reaction conditions to achieve the desired substitution patterns without unwanted side reactions. The methodologies to prepare functionalized fulvenes can be categorized into two main strategies:<sup>57</sup> (a) classical fulvene synthesis, consisting of the reaction of cyclopentadiene with various aldehydes or ketones in the presence of a base. This method allows for the introduction of various functional groups depending on the aldehyde or ketone used; (b) transition metal-catalyzed reactions using precious metals like palladium or rhodium to couple cyclopentadiene derivatives with other organic halides or pseudohalides, or to promote  $[2 + 2]$  or  $[4 + 2]$  cycloaddition reactions to form the fulvene core.

Controlling selectivity using the described methodologies is challenging, making the synthesis of some of the most promising fulvene candidates difficult and, in some cases, unachievable. This difficulty underscores the need for advancements in synthetic techniques. Recognizing this issue, it is anticipated that new experimental methodologies will emerge in the near future. These innovative approaches may include improved catalyst design, more efficient protecting group strategies, and enhanced radical and cycloaddition reactions. Such advancements will aim to provide greater

control over reaction pathways, thereby enabling the selective synthesis of these optimal fulvene candidates.

## Conclusions

In summary, our study has clearly demonstrated the efficiency of inverse design approaches in uncovering promising fulvene derivatives for singlet fission process. By applying a stepwise selection strategy based on increasingly strict energetic criteria, we narrowed our initial set from 1658 to 17 compounds that fully satisfy the SF energy requirements, representing a final success rate of approximately 1%. While 30% of the database meets the first and most important condition ( $E(S_1) \geq 2E(T_1)$ ) at the M06-2X level of theory, further filters reduce this number significantly, highlighting the complexity of identifying viable SF chromophores. Nonetheless, this outcome compares favorably with previously reported screening strategies, which often involve chemical diversity and lower final success rates.<sup>43,47,49</sup> Our results reaffirm and underscore the immense potential of computational methods in accelerating the discovery of materials tailored for SF photovoltaic cells.

This study has established a robust protocol for the application of inverse design in the search for novel SF candidates. From the database generated, we have derived optimal substitution patterns to render fulvene as a promising SF chromophore fulfilling the strict energy-matching conditions. As a rule, the  $R_3$  position must be occupied by at least one strong EDG and one strong EWG, while at least one strong EWG is required in the exocyclic  $R_1$  position. The position  $R_2$ , while influential, suggests that even a hydrogen atom can effectively serve this role.

These insights not only offer practical guidelines for future material design but also deepen our understanding of the factors governing the excited state energies relevant for singlet fission, including the diradical character and the ground- and excited-state aromaticity. Through the integrated use of computational methodologies with foundational chemical insights, we are set to drive the advancement of efficient and sustainable SF-based photovoltaic technologies. We plan to apply this protocol to the study of larger and more complex systems, further expanding the frontier of SF materials research.

## Data availability

All data supporting the findings of this study are available in the published article and/or the ESI of this article.† Additionally, a public Git repository is available at <https://gitlab.com/fdeveles/cindes> and <https://github.com/fdeveles/CINDES>, DOI: <https://doi.org/10.5281/zenodo.15753864>, which includes the CINDES program and a complete example illustrating the full inverse design process used in this study (example: singlet fission). The version of the code employed for this study is version 3.0. The repository also contains the Jupyter notebook Fulvene\_analysis.ipynb, used for analyzing the full dataset, along with a README file providing instructions on the usage of the program and the specific workflow applied in this



work. All software and code used in this study are appropriately cited within the manuscript and/or the ESI.†

## Author contributions

I. C.-R. and M. A. conceptualized the project. I. C.-R., R. M.-C., and E. D. performed all quantum chemical calculations. I. C.-R. and M. A. supervised the project. The first draft of the manuscript was written by I. C.-R., M. A., and F. D. V. All authors were involved in future editing and reviewing process. All authors have read and agreed to the published version of the manuscript. F. D. V. possesses the authorship of the CINDES code.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

M. A., F. D. V., and I. C. R. wish to acknowledge the VUB for a Strategic Research Program awarded to ALGC. The resources and services used in this work were provided by the Flemish Supercomputer Center (VSC), funded by the Research Foundation – Flanders (FWO), and the Flemish Government. I. C. R. acknowledges co-funding from the European Union's Horizon 2020 research and innovation Marie Skłodowska-Curie Actions, under grant agreement number 945380. R. M. C. thanks the Government of Spain for the predoctoral fellowship FPU20/00707 and the mobility grant EST23/00507. E. D. thanks the Fund for Scientific Research-Flanders (FWO-11E0321N) for financial support.

## References

- D. Bogdanov, J. Farfan, K. Sadovskaia, A. Aghahosseini, M. Child, A. Gulagi, A. S. Oyewo, L. de Souza Noel Simas Barbosa and C. Breyer, *Nat. Commun.*, 2019, **10**, 1–16.
- M. Bregnhøj, and P. R. Ogilby, Molecular oxygen in photoresponsive organic materials, in *Emerging Strategies to Reduce Transmission and Thermalization Losses in Solar Cells: Redefining the Limits of Solar Power Conversion Efficiency*, Springer, 2022, pp. 121–148.
- M. Einzinger, T. Wu, J. F. Kompalla, H. L. Smith, C. F. Perkinson, L. Nienhaus, S. Wieghold, D. N. Congreve, A. Kahn, M. G. Bawendi, *et al.*, *Nature*, 2019, **571**, 90–94.
- B. Daiber, K. van den Hoven, M. H. Futscher and B. Ehrler, *ACS Energy Lett.*, 2021, **6**, 2800–2808.
- T. Wang, H. Liu, X. Wang, L. Tang, J. Zhou, X. Song, L. Lv, W. Chen, Y. Chen and X. Li, *J. Mater. Chem. A*, 2023, **11**, 8515–8539.
- A. Rao and R. H. Friend, *Nat. Rev. Mater.*, 2017, **2**, 1–12.
- T. Hasobe, S. Nakamura, N. V. Tkachenko and Y. Kobori, *ACS Energy Lett.*, 2021, **7**, 390–400.
- A. J. Baldacchino, M. I. Collins, M. P. Nielsen, T. W. Schmidt, D. R. McCamey and M. J. Tayebjee, *Chem. Phys. Rev.*, 2022, **3**, 021304.
- K. Miyata, F. S. Conrad-Burton, F. L. Geyer and X.-Y. Zhu, *Chem. Rev.*, 2019, **119**, 4261–4292.
- T. Sharma, P. Mahajan, M. Adil Afroz, A. Singh, Yukta, N. Kumar Tailor, S. Purohit, S. Verma, B. Padha, V. Gupta, *et al.*, *ChemSusChem*, 2022, **15**, e202101067.
- C. Gao, Z. Miao, W. W. Wong, T. A. Smith, S.-C. Lo, W. Hu, E. B. Namdas and H. Dong, *Fundam. Res.*, 2024, DOI: [10.1016/j.fmre.2024.05.009](https://doi.org/10.1016/j.fmre.2024.05.009).
- T. Ullrich, D. Munz and D. M. Guldi, *Chem. Soc. Rev.*, 2021, **50**, 3485–3518.
- M. B. Smith and J. Michl, *Chem. Rev.*, 2010, **110**, 6891–6936.
- M. B. Smith and J. Michl, *Annu. Rev. Phys. Chem.*, 2013, **64**, 361–386.
- A. Japahuge and T. Zeng, *ChemPlusChem*, 2018, **83**, 146–182.
- R. J. Hudson, A. N. Stuart, D. M. Huang and T. W. Kee, *J. Phys. Chem. C*, 2022, **126**, 5369–5377.
- T. Wang, B.-Y. Zhang and H.-L. Zhang, *Macromol. Rapid Commun.*, 2022, **43**, 2200326.
- C. M. Nkinyam, C. O. Ujah, K. C. Nnakwo and D. V. Kallon, *Unconv. Resour.*, 2024, **5**, 100121.
- Y. Jiang, M. P. Nielsen, A. J. Baldacchino, M. A. Green, D. R. McCamey, M. J. Tayebjee, T. W. Schmidt and N. J. Ekins-Daukes, *Prog. Photovolt. Res. Appl.*, 2021, **29**, 899–906.
- D. Sasikumar, A. T. John, J. Sunny and M. Hariharan, *Chem. Soc. Rev.*, 2020, **49**, 6122–6140.
- C. Hetzer, D. M. Guldi and R. R. Tykwinski, *Chem.–Eur. J.*, 2018, **24**, 8245–8257.
- J. Li, H. Cao, Z. Zhang, S. Liu and Y. Xia, *Photonics*, 2022, **9**, 689.
- J. E. Anthony, *Angew. Chem., Int. Ed.*, 2008, **47**, 452–483.
- K. J. Thorley, H. Le, Y. Song and J. E. Anthony, *J. Mater. Chem. C*, 2022, **10**, 15861–15871.
- J. M. Wasikiewicz, L. Abu-Sen, A. B. Horn, J. M. Koelewijn, A. V. Parry, J. J. Morrison and S. G. Yeates, *J. Mater. Chem. C*, 2016, **4**, 7309–7315.
- J. Wen, Z. Havlas and J. Michl, *J. Am. Chem. Soc.*, 2015, **137**, 165–172.
- S. Ito, T. Nagami and M. Nakano, *J. Photochem. Photobiol., C*, 2018, **34**, 85–120.
- J. D. Green, E. G. Fuemmeler and T. J. Hele, *J. Chem. Phys.*, 2022, **156**, 180901.
- D. Casanova, *Chem. Rev.*, 2018, **118**, 7164–7207.
- F. J. Hernández and R. Crespo-Otero, *Annu. Rev. Phys. Chem.*, 2023, **74**, 547–571.
- X. Wang, S. Gao, Y. Luo, X. Liu, R. Tom, K. Zhao, V. Chang and N. Marom, *J. Phys. Chem. C*, 2024, **128**, 7841–7864.
- M. Nakano, *Chem. Rec.*, 2017, **17**, 27–62.
- T. Minami, S. Ito and M. Nakano, *J. Phys. Chem. Lett.*, 2013, **4**, 2133–2137.
- Ö. H. Omar, D. Padula and A. Troisi, *ChemPhotoChem*, 2020, **4**, 5223–5229.
- V. Petakova, M. Nedyalkova, J. Stoycheva, A. Tadjer and J. Romanova, *Symmetry*, 2021, **13**, 1448.



- 36 K. J. Fallon, P. Budden, E. Salvadori, A. M. Ganose, C. N. Savory, L. Eyre, S. Dowland, Q. Ai, S. Goodlett, C. Risko, *et al.*, *J. Am. Chem. Soc.*, 2019, **141**, 13867–13876.
- 37 O. El Bakouri, J. R. Smith and H. Ottosson, *J. Am. Chem. Soc.*, 2020, **142**, 5602–5617.
- 38 W. Zeng, O. El Bakouri, D. W. Szczepanik, H. Bronstein and H. Ottosson, *Chem. Sci.*, 2021, **12**, 6159–6171.
- 39 A. V. Girija, W. Zeng, W. K. Myers, R. C. Kilbride, D. T. Toolan, C. Zhong, F. Plasser, A. Rao and H. Bronstein, *J. Am. Chem. Soc.*, 2024, **146**, 18253–18261.
- 40 W. Zeng, D. W. Szczepanik and H. Bronstein, *J. Phys. Org. Chem.*, 2023, **36**, e4441.
- 41 J. Yan, T. Slanina, J. Bergman and H. Ottosson, *Chem.–Eur. J.*, 2023, **29**, e202203748.
- 42 S. Escayola, C. Tonnelé, E. Matito, A. Poater, H. Ottosson, M. Solà and D. Casanova, *Angew. Chem., Int. Ed.*, 2021, **60**, 10255–10265.
- 43 D. Padula, Ö. H. Omar, T. Nemataram and A. Troisi, *Energy Environ. Sci.*, 2019, **12**, 2412–2416.
- 44 O. A. von Lilienfeld, K.-R. Müller and A. Tkatchenko, *Nat. Rev. Chem.*, 2020, **4**, 347–358.
- 45 X. Liu, X. Wang, S. Gao, V. Chang, R. Tom, M. Yu, L. M. Ghiringhelli and N. Marom, *npj Comput. Mater.*, 2022, **8**, 70.
- 46 F. Weber and H. Mori, *npj Comput. Mater.*, 2022, **8**, 176.
- 47 D. López-Carballeira and T. Polcar, *J. Comput. Chem.*, 2021, **42**, 2241–2249.
- 48 D. López-Carballeira and T. Polcar, *Comput. Theor. Chem.*, 2023, **1229**, 114343.
- 49 L. Schaufelberger, J. T. Blaskovits, R. Laplaza, K. Jorner and C. Corminboeuf, *Angew. Chem., Int. Ed.*, 2025, **137**, e202415056.
- 50 B. Sanchez-Lengeling and A. Aspuru-Guzik, *Science*, 2018, **361**, 360–365.
- 51 F. De Vleeschouwer, P. Geerlings and F. De Proft, *ChemPhysChem*, 2016, **17**, 1414–1424.
- 52 J. L. Teunissen, F. De Proft and F. De Vleeschouwer, *J. Chem. Theory Comput.*, 2017, **13**, 1351–1365.
- 53 S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković and A. W. Rodriguez, *Nat. Photonics*, 2018, **12**, 659–670.
- 54 E. Desmedt, T. Woller, J. L. Teunissen, F. De Vleeschouwer and M. Alonso, *Front. Chem.*, 2021, **9**, 786036.
- 55 E. Desmedt, D. Smets, T. Woller, M. Alonso and F. De Vleeschouwer, *Phys. Chem. Chem. Phys.*, 2023, **25**, 17128–17142.
- 56 E. Desmedt, L. Serrano Gimenez, F. De Vleeschouwer and M. Alonso, *Molecules*, 2023, **28**, 7371.
- 57 P. Preethalayam, K. S. Krishnan, S. Thulasi, S. S. Chand, J. Joseph, V. Nair, F. Jaroschik and K. Radhakrishnan, *Chem. Rev.*, 2017, **117**, 3930–3989.
- 58 H. Möllerstedt, M. C. Piqueras, R. Crespo and H. Ottosson, *J. Am. Chem. Soc.*, 2004, **126**, 13938–13939.
- 59 S. Yadav, O. El Bakouri, K. Jorner, H. Tong, C. Dahlstrand, M. Solà and H. Ottosson, *Chem.–Asian J.*, 2019, **14**, 1870–1878.
- 60 E. Swan, K. Platts and A. Blencowe, *Beilstein J. Org. Chem.*, 2019, **15**, 2113–2132.
- 61 P. A. Wieczorkiewicz, K. K. Zborowski, T. M. Krygowski and H. Szatyłowicz, *J. Org. Chem.*, 2023, **88**, 14775–14780.
- 62 M. Neuenschwander, *Helv. Chim. Acta*, 2015, **98**, 763–784.
- 63 T. L. Andrew, J. R. Cox and T. M. Swager, *Org. Lett.*, 2010, **12**, 5302–5305.
- 64 A. D. Finke, B. O. Jahn, A. Saithalavi, C. Dahlstrand, D. Nauroozi, S. Haberland, J.-P. Gisselbrecht, C. Boudon, E. Mijangos, W. B. Schweizer, *et al.*, *Chem.–Eur. J.*, 2015, **21**, 8168–8176.
- 65 T. L. Andrew and V. Bulovic, *ACS Nano*, 2012, **6**, 4671–4677.
- 66 M. Rosenberg, H. Ottosson and K. Kilså, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12912–12919.
- 67 M. Wang, X. Hu, D. N. Beratan and W. Yang, *J. Am. Chem. Soc.*, 2006, **128**, 3228–3232.
- 68 D. Balamurugan, W. Yang and D. N. Beratan, *J. Chem. Phys.*, 2008, **129**, 174105.
- 69 J. L. Teunissen, F. De Proft and F. De Vleeschouwer, *J. Chem. Inf. Model.*, 2019, **59**, 2587–2599.
- 70 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
- 71 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.
- 72 W. J. Hehre, L. Radom, P. von R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
- 73 M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. Petersson, H. Nakatsuji, *et al.*, *Gaussian 16*, Revision A.01, 2016.
- 74 I. Casademont-Reig, E. Ramos-Cordoba, M. Torrent-Sucarrat and E. Matito, *Molecules*, 2020, **25**, 711.
- 75 T. Woller, A. Banerjee, N. Sylvetsky, G. Santra, X. Deraet, F. De Proft, J. M. Martin and M. Alonso, *J. Phys. Chem. A*, 2020, **124**, 2380–2397.
- 76 N. Sylvetsky, A. Banerjee, M. Alonso and J. M. Martin, *J. Chem. Theory Comput.*, 2020, **16**, 3641–3653.
- 77 M. Torrent-Sucarrat, S. Navarro, F. P. Cossio, J. M. Anglada and J. M. Luis, *J. Comput. Chem.*, 2017, **38**, 2819–2828.
- 78 I. Casademont-Reig, T. Woller, J. Contreras-García, M. Alonso, M. Torrent-Sucarrat and E. Matito, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2787–2796.
- 79 D. W. Szczepanik, M. Solà, M. Andrzejak, B. Pawełek, J. Dominikowska, M. Kukułka, K. Dydach, T. M. Krygowski and H. Szatyłowicz, *J. Comput. Chem.*, 2017, **38**, 1640–1654.
- 80 R. Grotjahn, T. M. Maier, J. Michl and M. Kaupp, *J. Chem. Theory Comput.*, 2017, **13**, 4984–4996.
- 81 F. Neese, *Wiley Interdiscip. Rev.:Comput. Mol. Sci.*, 2018, **8**, e1327.
- 82 C. Angeli, R. Cimiraglia and S. Evangelisti, *J. Phys. Chem.*, 2001, **114**, 10252.
- 83 D. Hegarty and M. A. Robb, *Mol. Phys.*, 1979, **38**, 1795–1812.
- 84 R. H. Eade and M. A. Robb, *Chem. Phys. Lett.*, 1981, **83**, 362–368.
- 85 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 86 F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057–1065.
- 87 M. Nakano, *Top. Curr. Chem.*, 2017, **375**, 1–67.
- 88 K. Yamaguchi, *Chem. Phys. Lett.*, 1975, **33**, 330–335.



- 89 T. A. Keith, AIMALL, (Version 14.11.23), TK Gristmill Software, Overland Park KS, USA, 2014.
- 90 E. Matito, M. Solà, P. Salvador and M. Duran, *Faraday Discuss.*, 2007, **135**, 325–345.
- 91 E. Matito, M. Solà, M. Duran and P. Salvador, *J. Phys. Chem. A*, 2006, **110**, 5108–5113.
- 92 E. Matito, ESI-3D: Electron Sharing Indices Program for 3D Molecular Space Partitioning, *Institute of Computational Chemistry and Catalysis*, University of Girona, Catalonia, Spain, 2015, (available upon request: [ematito@gmail.com](mailto:ematito@gmail.com)).
- 93 E. Matito, M. Duran and M. Solà, *J. Chem. Phys.*, 2005, **122**, 014109.
- 94 E. Matito, M. Duran and M. Solà, *J. Chem. Phys.*, 2006, **125**, 059901.
- 95 J. Kruszewski and T. M. Krygowski, *Tetrahedron Lett.*, 1972, **13**, 3839–3842.
- 96 J.-L. Bredas, *J. Chem. Phys.*, 1985, **82**, 3808–3811.
- 97 C. B. Gorman and S. R. Marder, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 11297–11301.
- 98 M. Giambiagi, M. S. de Giambiagi, C. D. dos Santos Silva and A. P. de Figueiredo, *Phys. Chem. Chem. Phys.*, 2000, **2**, 3381–3392.
- 99 C. G. Bollini, M. Giambiagi, M. S. d. Giambiagi and A. P. de Figueiredo, *Struct. Chem.*, 2001, **12**, 113.
- 100 P. Bultinck, R. Ponc and S. Van Damme, *J. Phys. Org. Chem.*, 2005, **18**, 706–718.
- 101 K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.*, 1990, **112**, 8251–8260.
- 102 H. Fliegl, S. Taubert, O. Lehtonen and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20500–20518.
- 103 D. Sundholm, H. Fliegl and R. Berger, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2016, **6**, 10–1002.
- 104 U. Ayachit, *The paraview guide: a parallel visualization application*, Kitware, Inc., 2015.
- 105 J. Pearl, *Heuristics: intelligent search strategies for computer problem solving*, Addison-Wesley Longman Publishing Co., Inc., 1984.
- 106 J. Pearl and R. E. Korf, *Annu. Rev. Comput. Sci.*, 1987, **2**, 451–467.
- 107 D. B. Gordon and S. L. Mayo, *Structure*, 1999, **7**, 1089–1098.
- 108 R. Dechter and J. Pearl, *J. Assoc. Comput. Mach.*, 1985, **32**, 505–536.
- 109 J. L. Teunissen, PhD thesis, Vrije Universteit Brussel, 2019.
- 110 J. L. Teunissen, CINDES, <https://gitlab.com/jlteunissen/CINDES>, Accessed on June 30, 2022, 2016.
- 111 T. Zeng, N. Ananth and R. Hoffmann, *J. Am. Chem. Soc.*, 2014, **136**, 12638–12647.
- 112 M. J. Bearpark, F. Bernardi, M. Olivucci, M. A. Robb and B. R. Smith, *J. Am. Chem. Soc.*, 1996, **118**, 5254–5260.
- 113 D. Mendive-Tapia, B. Lasorne, G. A. Worth, M. J. Bearpark and M. A. Robb, *Phys. Chem. Chem. Phys.*, 2010, **12**, 15725–15733.
- 114 I. Schapiro, K. Sivalingam and F. Neese, *J. Chem. Theory Comput.*, 2013, **9**, 3567–3580.

