

## Article

# Application of Inverse Design Approaches to the Discovery of Nonlinear Optical Switches

Eline Desmedt , Léa Serrano Gimenez , Freija De Vleeschouwer \*  and Mercedes Alonso \* 

Eenheid Algemene Chemie (ALGC), Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussels, Belgium; eline.louise.desmedt@vub.be (E.D.); lea.serrano.gimenez@vub.be (L.S.G.)

\* Correspondence: freija.de.vleeschouwer@vub.be (F.D.V.); mercedes.alonso.giner@vub.be (M.A.)

**Abstract:** Molecular switches, in which a stimulus induces a large and reversible change in molecular properties, are of significant interest in the domain of photonics. Due to their commutable redox states with distinct nonlinear optical (NLO) properties, hexaphyrins have emerged as a novel platform for multistate switches in nanoelectronics. In this study, we employ an inverse design algorithm to find functionalized **26R** → **28R** redox switches with maximal  $\beta_{HRS}$  contrast. We focus on the role of core modifications, since a synergistic effect with *meso*-substitutions was recently found for the **30R**-based switch. In contrast to these findings, the inverse design optima and subsequent database analysis of **26R**-based switches confirm that core modifications are generally not favored when high NLO contrasts are targeted. Moreover, while push–pull combinations enhance the NLO contrast for both redox switches, they prefer a different arrangement in terms of electron-donating and electron-withdrawing functional groups. Finally, we aim at designing a three-state **26R** → **28R** → **30R** switch with a similar NLO response for both ON states. Even though our best-performing three-state switch follows the design rules of the **30R**-based component, our chemical compound space plots show that well-performing three-state switches can be found in regions shared by high-responsive **26R** and **30R** structures.

**Keywords:** chemical compound space; inverse design; multistate switches; nonlinear optical properties; structure–property relationships



**Citation:** Desmedt, E.; Serrano Gimenez, L.; De Vleeschouwer, F.; Alonso, M. Application of Inverse Design Approaches to the Discovery of Nonlinear Optical Switches. *Molecules* **2023**, *28*, 7371. <https://doi.org/10.3390/molecules28217371>

Academic Editors: Heribert Reis and Josep Maria Luis

Received: 3 October 2023

Revised: 24 October 2023

Accepted: 26 October 2023

Published: 31 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Molecular switches are essential components of artificial molecular machines and responsive materials [1–3]. A molecular switch consists of a single molecule that can shift controllably between two or more stable states triggered by an external stimulus [4]. These smart switching molecules help to engineer new functionalities unattainable with silicon-based technology, leading to a broad range of applications in biomedicine [5], catalysis [6], sensors [7], and single-molecule electronics [8]. Within this context, molecules displaying distinct responses towards multiple external stimuli as well as multistate molecular switches are highly desirable as different switchable functions can be integrated into a single molecule [9–11].

In the field of nonlinear optics, molecules with switchable nonlinear optical (NLO) properties are key building blocks for optoelectronic and photonic devices, like high-density optical memories or logic gates [12–15]. In these dynamic systems, different NLO properties can reversibly be turned “ON” and “OFF”, such as second-harmonic generation (SHG), two-photon absorption (TPA), and third-harmonic generation (THG). At the molecular level, the NLO quantities of interest are the first ( $\beta$ ) and second hyperpolarizability ( $\gamma$ ), while the key factor governing the efficiency of NLO switches is the contrast between the NLO properties of the different states [12]. The amplitude of the NLO contrast can be fine-tuned by chemical design, i.e., varying the donor/acceptor substituents, the length and planarity of the  $\pi$ -conjugated spacer and the diradical character [16].

The quest for novel NLO materials usually relies on complementary experimental and theoretical approaches [12]. Quantum chemistry tools are very helpful to understand the structural and electronic factors governing the magnitude of the molecular hyperpolarizabilities, reducing the trial-and-error approach in the design of efficient NLO switches [17]. Most theoretical studies involve direct design approaches, in which ad hoc variation of a chemical structure is followed by evaluation of its NLO properties [18,19]. Although direct design has led to the identification of NLO switches with high first hyperpolarizability contrasts, alternative design approaches are needed for a larger-scale exploration of chemical compound space (CCS) given its size and complexity [20]. Inverse design is an innovative approach for exploring the CCS, involving the optimization of the property of interest as a function of the chemical structure. Hence, inverse design inverts the design protocol by starting with the desired properties and producing a molecule that satisfies them as an output (Figure 1A) [21]. The main advantage of all inverse design strategies is the fact that only a small fraction of a predefined CCS needs to be explored to identify new and largely improved compounds [22]. Recent advances in computational inverse design approaches have started to reshape the landscape of structures with enhanced optical functionalities available to nanophotonics [23,24].

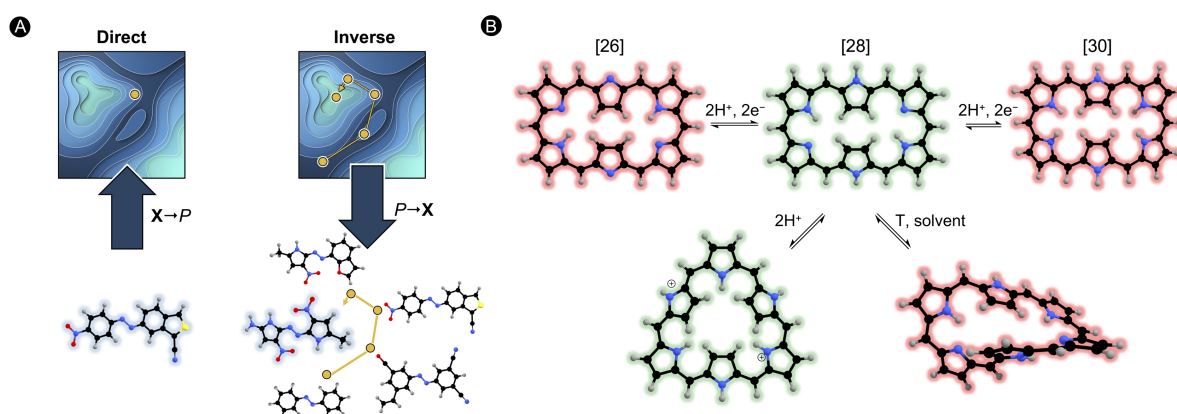


**Figure 1.** (A) Schematic representation of direct and inverse design approaches. Inverse design follows the reverse design protocol starting with the property of interest and ending in the optimal structure. (B) Different structures in hexaphyrin macrocycles interconvertible upon redox and protonation reactions as well as changing solvation and temperature.

Motivated by these findings and building on our expertise in developing and applying property optimization algorithms [22,25,26], we recently extended inverse design to the prominent field of molecular switches for nonlinear optical applications. In our long-standing interest in the design of porphyrinoid-based molecular switches [27–29], we selected hexaphyrin macrocycles as a test bed to assess the potential of inverse design approaches to identify hexaphyrin-based switches with enhanced NLO contrasts. Previous experimental and theoretical work demonstrated that molecular switches based on expanded porphyrins have a multiaddressable and multistate character, leading to exciting nanoelectronic applications such as conductance and optical switching [28,30,31]. On the one hand, hexaphyrins are flexible enough to switch between different  $\pi$ -conjugation topologies, each with distinct electronic properties and aromaticity [32]. On the other hand, hexaphyrins exhibit reversible redox chemistry providing congeneric macrocycles with  $[4n]$  and  $[4n + 2]$   $\pi$ -electrons. Such redox and topology switches can be triggered by different external stimuli (Figure 1B) [33–35]. Furthermore, the NLO properties are easily tuneable in these macrocycles due to their structural versatility with respect to chemical functionalizations such as peripheral substitution or core modifications [28,31].

In previous studies, we applied inverse design to rationally modify redox ( $26R \rightarrow 28R$ ,  $30R \rightarrow 28R$ ) and topological hexaphyrin-based switches ( $28M \rightarrow 28R$ ) with

the help of the best-first search (BFS) algorithm (vide infra) in search for well-performing NLO switches [36,37]. We succeeded in enhancing the NLO efficiency by at least an order of magnitude and identified some key players influencing the NLO contrast, i.e., the electronic character of *meso*-substituents and the centrosymmetry of the OFF state. For the **30R**-based switch, core modifications of the macrocycle were also considered, and a synergistic interplay between *meso*-substituents and core modifications was unveiled. Even though the inclusion of both types of functionalizations unlocks design potential, it also makes the prediction of the NLO contrast of hexaphyrin-based switches highly complicated and reinforces the need for innovative approaches beyond direct design.

Given the importance of core modifications for **30R**-based NLO switches [37], in this study, we revisit the **26R** → **28R** redox switch with the aim to explore the effect of core modifications on its NLO efficiency [36]. To this end, BFS optimizations of the NLO contrast are set up for the substitution pattern termed NH\_Y\_R<sub>1,4</sub>\_R<sub>2,5</sub>\_R<sub>3,6</sub> (Figure 2), with three sets of *meso*-substitution sites and one set of core modification sites combined with a diverse functionalization library. Our objective is to establish general structure–property relationships for the **26R**-based switch and compare them to the design rules devised for the **30R** → **28R** molecular switch [37]. Due to the increasing importance of multifunctional NLO materials, for the first time, we search for the optimal functionalization of a three-state molecular switch (**26R** → **28R** → **30R**) using a newly introduced figure of merit. Finally, CCS plots of the three-level switch database help us to further understand the role that each individual switch plays in the multistate switch.

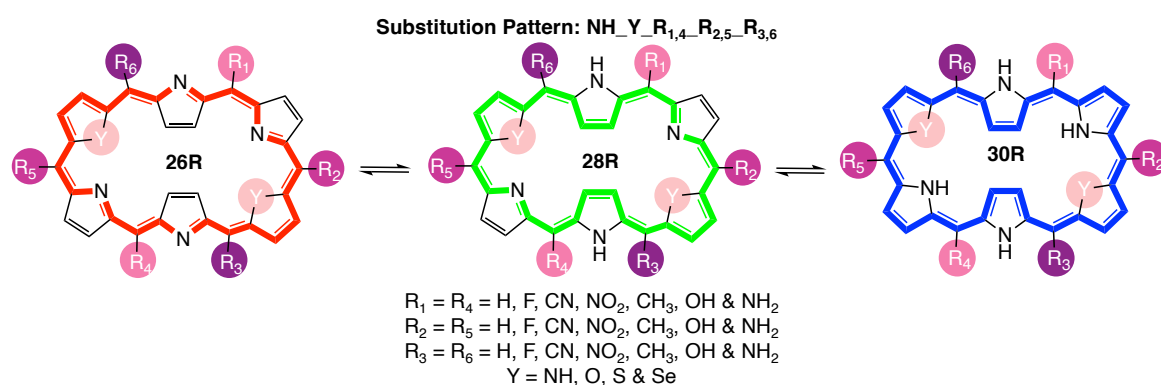


**Figure 2.** Graphical representation of the three-state hexaphyrin-based redox switch **26R** → **28R** → **30R**, indicating the three pairs of *meso*-substitution sites {R<sub>1,4</sub>; R<sub>2,5</sub>; R<sub>3,6</sub>}, the one set of core modification sites {Y} as well as the functionalization library.

## 2. Methodology

### 2.1. Inverse Design Algorithm-Best-First Search

When exploring the combinatorial chemical compound space (CCS) aiming at finding the optimal functionalization of a known molecular framework, several optimization algorithms can be employed. One of them is the greedy best-first search (BFS) approach [38–40], a discrete heuristic search method that optimizes the target property on a site-by-site basis assuming that the sites can be independently optimized. The discrete aspect of BFS translates to the calculation of chemically realistic molecules, i.e., functionalized derivatives of the scaffold of interest, when following the steepest gradient towards the optimum. Taking the azoheteroarene scaffold as an example of a photoswitch with 4 (sets of) sites, R<sub>1</sub>–R<sub>4</sub>, available for modification and 5 possible substituents, depicted as colored circles (Figure 3), it leads to a chemical space region of 625 potential compounds. After generating a random start structure, the first site (R<sub>1</sub>) is visited and 5 possible structures are constructed in accordance with the fragment library size. For each structure, the target property is evaluated and its value serves as input for the determination of the largest gradient among the 5 available substituents. As BFS assumes that the largest gradient points to the most

promising optimization pathway, the corresponding functionalization is implemented, and the next site can be optimized. When all sites are iterated, it is checked whether the input and output structures are identical. If yes, the BFS procedure terminates as convergence is reached; if not, the procedure reprises with the output structure as new input. This reiteration circumvents the initial presumption of the independent-site approximation and partially ensures that the effect of neighbouring sites is accounted for.

As this optimization method is quite greedy, optima are found highly efficiently. However, depending on the roughness of the property landscape, these optima are often local, whereas one generally targets the global optimum. More effective optimization algorithms have been designed for that purpose, such as the genetic algorithm [41,42], particle swarm optimization [43], and Bayesian optimization [44,45]. Nevertheless, also for BFS, the issue of efficacy can be (partially) remedied. In this work, multiple BFS runs are conducted varying the starting structure and the site sequence. By selecting starting structures from different regions of CCS, a large enough portion of CCS is being explored. Despite the shortcoming of becoming trapped in local optima, BFS has the great advantage to search CCS locally at each iteration step by only changing one site at a time (Figure 3). The wealth of consistent structure–property information undeniably aids in deriving general design rules or guiding principles, although multiple BFS runs are advisable to gather more data.

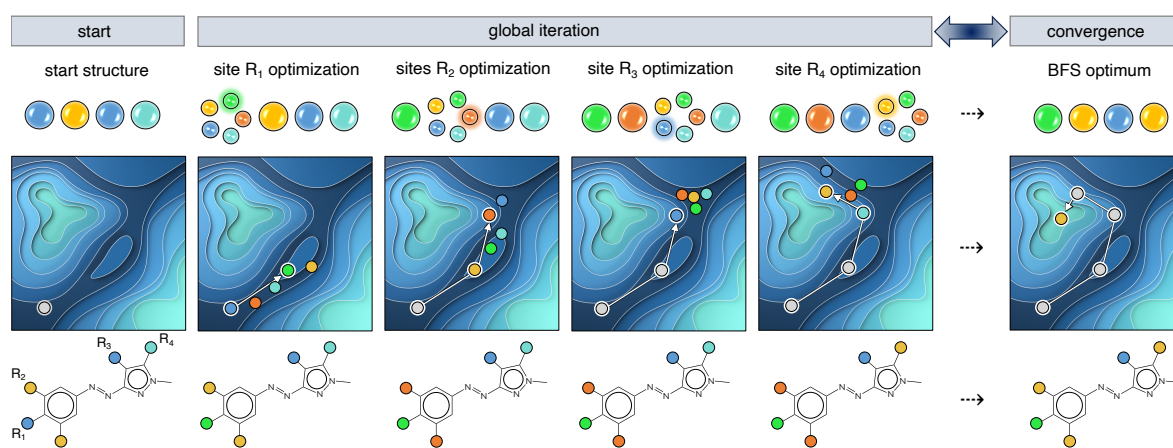


**Figure 3.** Schematic representation of the best-first search algorithm applied on an azoheteroarene photoswitch with 4 modifiable (sets of) sites  $R_1$ – $R_4$  and a fragment library of 5 substituents depicted as blue-, cyan-, green-, yellow- and orange-colored circles.

## 2.2. Figure of Merit for the Design of NLO Switches

The first hyperpolarizability related to the Hyper-Rayleigh Scattering ( $\beta_{HRS}$ ) is an experimentally measurable property to probe the second-order NLO responses of the different states of the switches [12,46–48]. This property is related to the intensity of the incoherent scattered light at twice the frequency of the incident light ( $2\omega$ ). Under the condition that the laser's propagation is perpendicular to the incoherent scattered light, the HRS equations based on the full tensor components are written as Equation (1). The full tensor components can be found in Supporting Information.

$$\beta_{HRS}(-2\omega; \omega, \omega) = \sqrt{\langle \beta_{ZZZ}^2 \rangle + \langle \beta_{ZXX}^2 \rangle}. \quad (1)$$

To estimate the performance of a molecular NLO switch, a figure of merit is required. Most publications rely on the ratio between the  $\beta_{HRS}$  responses of the ON and OFF states of the molecular switch to estimate the switching performance (Equation (2)) [12,28].

$$\text{contrast} = \frac{\beta_{HRS}(\text{ON})}{\beta_{HRS}(\text{OFF})}. \quad (2)$$

Nonetheless, this ratio-based contrast definition is unsuited when the OFF state possesses  $C_i$  symmetry, as happens in certain [28]-hexaphyrins. In this case, the denominator becomes zero and the contrast value approaches infinity. In our previous study, we circumvented this problem within the BFS procedure by assigning an arbitrary value of 0.001 a.u. to the  $\beta_{HRS}(\text{OFF})$  response when the  $\beta_{HRS}(\text{OFF})$  value is below 10 a.u. [36]. This arbitrary OFF-state value in case of (nearly) centrosymmetric structures has a significant impact on the NLO contrast of such hexaphyrin-based switches, as illustrated in Figure 4A. As desired, NLO switches with a centrosymmetric OFF state ( $\beta_{HRS}(\text{OFF}) \triangleq 0.001$  a.u. for most blue-colored marks), like Switch 3, are preferred over those with a high  $\beta_{HRS}(\text{OFF})$  response, such as Switch 2 indicated with an orange mark, although their variance is overestimated. However, switches with a high ON-state response and a relatively low OFF-state response, though larger than 10 a.u., are undervalued by the ratio-based definition. For example, Switch 1 has a  $\beta_{HRS}(\text{ON})$  value of  $1.72 \times 10^4$  a.u., similar to that of System 3 ( $2.07 \times 10^4$  a.u.). Nonetheless, even with a  $\beta_{HRS}(\text{OFF})$  response below 500 a.u. (blue mark), Switch 1 will never be considered using the ratio-based contrast definition despite its potential, as it yields a contrast value as low as 87, in stark contrast to the ratio of  $2.07 \times 10^7$  evaluated for 3.

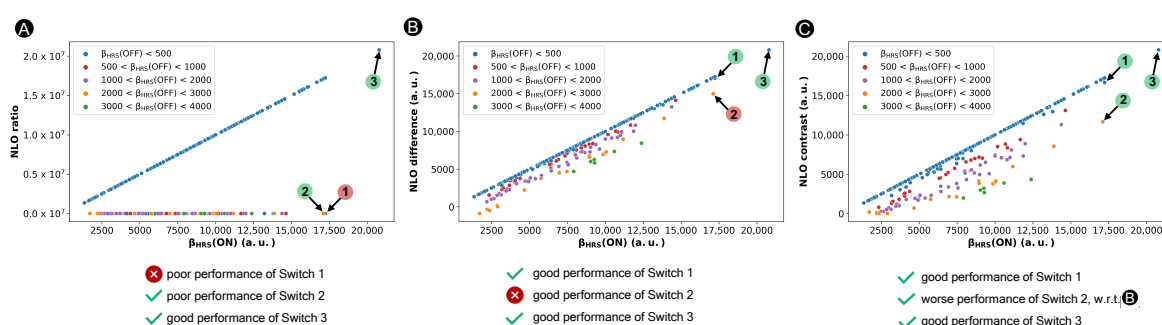


**Figure 4.** Scatter plots of three contrast definitions based on the ratio (A), difference (B) and a revised ratio (C) versus the  $\beta_{HRS}$  of the 30R structures. Three representative switches, Switches 1–3, are highlighted. If the system is categorized correctly by the NLO contrast definition, it is labelled with a green circle, otherwise with a red circle.

A second NLO contrast definition introduced in the literature is based on the difference between the  $\beta_{HRS}$  responses of the ON and OFF states, as written in Equation (3) [31].

$$\text{contrast} = \beta_{HRS}(\text{ON}) - \beta_{HRS}(\text{OFF}). \quad (3)$$

Similar to the ratio-based contrast, a high difference value corresponds to a switch with a high  $\beta_{HRS}(\text{ON})$  and a low  $\beta_{HRS}(\text{OFF})$  response. As illustrated in Figure 4B, 1 is now considered as one of the best performing switches. Thus, the difference-based definition overcomes the issue of the ratio-based contrast overpromoting centrosymmetric OFF states. However, structures with a significant  $\beta_{HRS}(\text{OFF})$  response may also be favored, such as Switch 2 with a  $\beta_{HRS}(\text{OFF})$  value of  $2.13 \times 10^3$  a.u., which conflicts with the more commonly used ratio-based metric in Figure 4A.

To overcome the downsides of the previous ratio- and difference-based contrast definitions, we introduced a new metric to assess the performance of the NLO switches in our most recent work [37]. This definition (Equation (4)) steers our search of well-performing

NLO switches towards a combination of high  $\beta_{HRS}(\text{ON})$  states and low  $\beta_{HRS}(\text{OFF})$  states and simultaneously disfavors switches with excessively high  $\beta_{HRS}(\text{OFF})$  responses.

$$\text{contrast} = \frac{(\beta_{HRS}(\text{ON}) - \beta_{HRS}(\text{OFF}))^2}{\beta_{HRS}(\text{ON}) + \beta_{HRS}(\text{OFF})}. \quad (4)$$

The performance of our new figure of merit for the design of NLO switches is verified in Figure 4C. Systems 1 and 3 still remain one of the overall best-performing structures, while the NLO contrast of 2 is decreased compared to the difference-based metric. Nevertheless, it is important to note that molecular switches characterized by an enhanced ON state and a centrosymmetric OFF state like System 3 will always be dominant, irrespective of the NLO contrast metric used.

In the Results and Discussion section, the revised contrast ratio definition (Equation (3)) is applied throughout.

### 3. Results and Discussion

#### 3.1. Functionalized [26]- and [30]-Hexaphyrin-Based Switches: What Sets Them Apart?

##### 3.1.1. Applying the BFS algorithm on the 26R → 28R and 30R → 28R

In this section and for the first time, the role of core modifications, in combination with *meso*-substitution, is assessed towards the improvement of 26R → 28R switches based on  $\beta_{HRS}$  response. An important measure for the BFS efficacy is the optimum's improvement over the parent, namely the unsubstituted switch with a  $\beta_{HRS}$  contrast of  $2.09 \times 10^3$  a.u. for the 26R-based switches and  $2.18 \times 10^3$  a.u. for the 30R-based switches. Figure 5 summarizes two BFS inverse design procedures, yielding the best 26R → 28R (A) and 30R → 28R (B) NLO switches discovered over all generated BFS procedures. We focus on the patterns shown in Figure 5. In the following sections, the 26R → 28R functionalization patterns are denoted as NH\_Y\_R<sub>1,4</sub>\_R<sub>2,5</sub>\_R<sub>3,6</sub>, and the 30R → 28R patterns as X\_Y\_R<sub>1,4</sub>\_R<sub>2,5</sub>\_R<sub>3,6</sub>, in accordance with 28R functionalized structures. We note that an additional set of modifiable core modification sites (X) is included for the 30R → 28R switch with respect to 26R → 28R, since O/S/Se core modifications on X would lead to charged 26R structures. Both procedures follow the same site iteration sequence and start from similar starting configurations: NH\_O\_NH<sub>2</sub>\_F\_OH and S\_O\_NH<sub>2</sub>\_F\_OH for 26R → 28R and 30R → 28R switches, respectively. Each site optimization step represents the introduction of a functionalization on or in the macrocycle. Strong EWGs (CN and NO<sub>2</sub>) are marked in red, strong EDGs (OH and NH<sub>2</sub>) in green, and unmodified core sites (NH) in blue; all other site functionalizations are colored in black. Individual  $\beta_{HRS}$  responses and the associated contrasts of the generated structures during both BFS procedures can be found in Tables S1 and S2 in Supporting Information.

An elaborate analysis of the BFS optimization path is included for the 26R → 28R switch. For a detailed description of the 30R → 28R's BFS run (Figure 5B), we refer to our recent publication [37]. The algorithm starts off with substituting the R<sub>3</sub>/R<sub>6</sub> *meso*-positions of 26R → 28R (Figure 5A). All OFF states of the switches generated within this site iteration are centrosymmetric. Interestingly, all *meso*-substitutions enhance the NLO response of the ON state compared to the parent 26R → 28R switch. Nonetheless, there is a clear preference for strong EDGs, yielding an NLO contrast almost three-fold that of the EWG-substituted switches and even 10 times higher than the parent switch. The next sites up for modification are R<sub>1</sub>/R<sub>4</sub>. These *meso*-substitution positions behave similarly to the previous site iteration, where EDGs outperform EWGs. On both site combinations, R<sub>1</sub>/R<sub>4</sub> and R<sub>3</sub>/R<sub>6</sub>, NH<sub>2</sub> is favored the most.

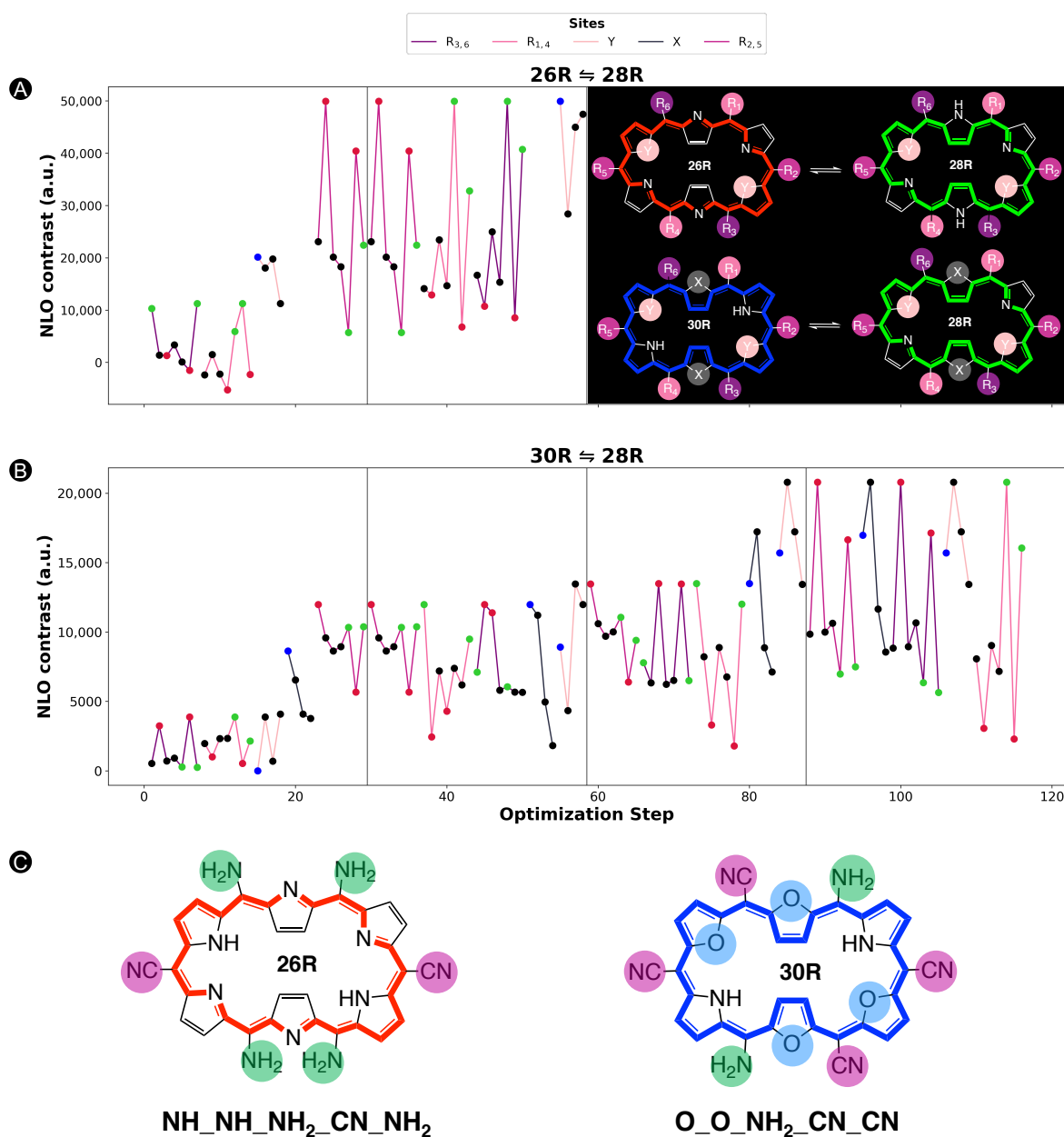


**Figure 5.** Chronological site order of the BFS procedure maximizing the NLO contrast for (A)  $26R \rightarrow 28R$  and (B)  $30R \rightarrow 28R$  switches starting from  $NH\_O\_NH\_2\_F\_OH$  and  $S\_O\_NH\_2\_F\_OH$ , respectively. Black lines separate the global iterations. The data points highlighted in blue, red and green are the sites with NH groups, strong EWGs and strong EDGs, respectively. The gaps in (A) correspond to the sites X that were not considered for the  $26R \rightarrow 28R$  switch. The ON states of the optimal substituted molecular switches for both BFS runs are presented in (C) with indication of the substitution pattern.

The effect of core modifications on the NLO contrast is evaluated during the third site optimization. As all OFF states are, again, centrosymmetric, the NLO contrast mainly depends on the  $\beta_{HRS}(ON)$  response. Thiophene and selenophene rings behave similarly to pyrrole rings regarding the NLO contrast, with pyrrole rings ranked first. Only furan rings underperform to the other possible modifications, and thus O becomes replaced by NH on core modification sites Y. The change to NH is a crucial factor in enhancing the NLO contrast as it increases by almost  $10 \times 10^4$  a.u.

The final site iteration of the first global iteration consists of *meso*-substitution of sites  $R_2/R_5$ , where only the  $NH_2$  substituent results in a noncentrosymmetric OFF state (cf. green mark with lowest NLO contrast value). Remarkably, on  $R_2/R_5$  positions, strong EWGs are essential for further maximizing the performance of the switch, in contrast to the other *meso*-substitution sites. The best-performing molecular switch has the  $NH\_NH\_NH_2\_CN\_NH_2$  pattern, which is a push–pull configuration containing two sets of strong EDGs and one set of strong EWGs, and without core modifications present. The NLO contrast of  $5.99 \times 10^4$  a.u. is 30 times larger than that of the fully unsubstituted switch. As observed from the wide ranges in NLO contrast values during the entire second global iteration, exchanging the optimal *meso*-substitution for other substituents immensely impacts the contrast, but for a handful of functionalizations. The resulting optimal switch is in line with our earlier study on the **26R**  $\rightarrow$  **28R** switch design [36].

When comparing the BFS optimization paths for the **26R**  $\rightarrow$  **28R** and **30R**  $\rightarrow$  **28R** switches (Figure 5A,B), we observe a different preference for certain modifications despite the structural similarity of both switches. Whereas strong EDGs were preferred on the  $R_3/R_6$  *meso*-positions throughout the entire **26R**  $\rightarrow$  **28R** optimization, EWGs are promoted in these positions for **30R**  $\rightarrow$  **28R**. Moreover, the core modification sites Y favour any type of core modification over none. Besides their differences, both redox switches also show similarities.  $R_1/R_4$  and  $R_2/R_5$  share the same preference for the type of *meso*-substituents as the **26R**  $\rightarrow$  **28R** switch, with few exceptions that can be attributed to the presence of a noncentrosymmetric OFF states or an unexpectedly low NLO response of the ON state. Finally, the type of core modification on sites X, only considered for **30R**  $\rightarrow$  **28R**, inclines more towards the same preference as the sites Y of the **26R**  $\rightarrow$  **28R**. This trend, however, collapses during the last two global iterations as inserting the same core modifications on sites X and Y, more specifically two sets of oxygen atoms, further boosts the NLO contrast to a value of  $2.08 \times 10^4$  a.u.

In conclusion, the best-performing patterns resulting from the BFS procedures are  $NH\_NH\_NH_2\_CN\_NH_2$  for the **26R**-based NLO switch and  $O\_O\_NH_2\_CN\_CN$  for the **30R**-based switch. We note that three additional BFS runs using start structures for which a variation in type of core modifiers was imposed resulted in our aforementioned optimal structure. This makes it very plausible that we identified one of the most NLO-efficient **26R**-based switches. The **26R** and **30R** optima confirm our hypothesis that the preferred number of EWGs and EDGs on the macrocycle depends on the oxidation state of the redox switch. Furthermore, the **30R**-based switches favour the incorporation of core modifications, contrary to the **26R**-based switch. Nevertheless, both optima are characterized by an enhanced ON state and a centrosymmetric OFF state, in line with our previous studies [36,37]. Through functionalization, the total NLO contrast can be enhanced by a factor ranging from 10 for **30R**-based switches to even 30 for **26R**-based switches. To assess the role of each functionalization towards the total NLO response, a steepest ascent approach is applied on the two optimal switches.

### 3.1.2. Steepest Ascent of the BFS Optima

The steepest ascent methods start from the unsubstituted molecular switch and gradually builds up towards the optimum, following the largest property gradient. The number of steps is equal to the number of functionalized sites. For the **26R**- and **30R**-based BFS optima, this number is three and five, respectively. Table 1 and Figure 6 summarize our analysis of the steepest ascent results, in which the NLO response of the optima is decomposed with respect to the functionalization's position and type. Our main figure of interest is the %(B-P), which quantifies how much each modification contributes to the overall improvement of the best-performing switch over the parent switch.

In the first step, all possible functionalizations enhance the NLO contrast with respect to the parent structure. The EDGs on  $R_{1,4}$  and  $R_{3,6}$  each contribute around 19% and 15%, respectively, to the best NLO contrast, while the EWG on  $R_{2,5}$  only yields a 6% improvement. Although similar contributions are observed for the two sets of EDGs,  $NH_2$  substituents

on R<sub>1,4</sub> perform slightly better. In Step 2, we observe that adding a second set of NH<sub>2</sub> substituents on R<sub>1,4</sub> leads to the largest enhancement. The combination of strong EDGs on R<sub>1,4</sub> and R<sub>3,6</sub> is therefore preferred over a push–pull pattern, in which NH<sub>2</sub> on R<sub>3,6</sub> is combined with CN on R<sub>2,5</sub>, although the differences are small (5%). The strong EDGs on R<sub>3,6</sub> provide a 26% improvement, i.e., 11% larger than their contribution in Step 1. In other words, the *meso*-substitutions combined act synergistically, with the most dramatic effect for the push–pull-type switch (+15%).

**Table 1.** Steepest ascent of our best-performing 26R-based switch: NH\_NH\_NH<sub>2</sub>\_CN\_NH<sub>2</sub>. %B represents the percentage in the NLO contrast of all switches in a steepest ascent step with respect to the NLO contrast maximum. %(B-P) is similar to %B but takes the parent as a reference. %(CB) stands for the percentage in the NLO contrast of all switches in a steepest ascent step with respect to the current best.

| Step   | X  | Y  | R <sub>1,4</sub> | R <sub>2,5</sub> | R <sub>3,6</sub> | Contrast (a.u.)    | %B  | %(B-P) | %(CB) |
|--------|----|----|------------------|------------------|------------------|--------------------|-----|--------|-------|
| Parent | NH | NH | H                | H                | H                | $2.09 \times 10^3$ | 3   | 0      | -     |
| 1      | NH | NH | NH <sub>2</sub>  | H                | H                | $1.32 \times 10^4$ | 22  | 19     | 100   |
|        | NH | NH | H                | CN               | H                | $5.73 \times 10^3$ | 10  | 6      | 33    |
|        | NH | NH | H                | H                | NH <sub>2</sub>  | $1.09 \times 10^4$ | 18  | 15     | 80    |
| 2      | NH | NH | NH <sub>2</sub>  | CN               | H                | $2.53 \times 10^4$ | 42  | 40     | 89    |
|        | NH | NH | NH <sub>2</sub>  | H                | NH <sub>2</sub>  | $2.83 \times 10^4$ | 47  | 45     | 100   |
| 3      | NH | NH | NH <sub>2</sub>  | CN               | NH <sub>2</sub>  | $5.99 \times 10^4$ | 100 | 100    | 100   |

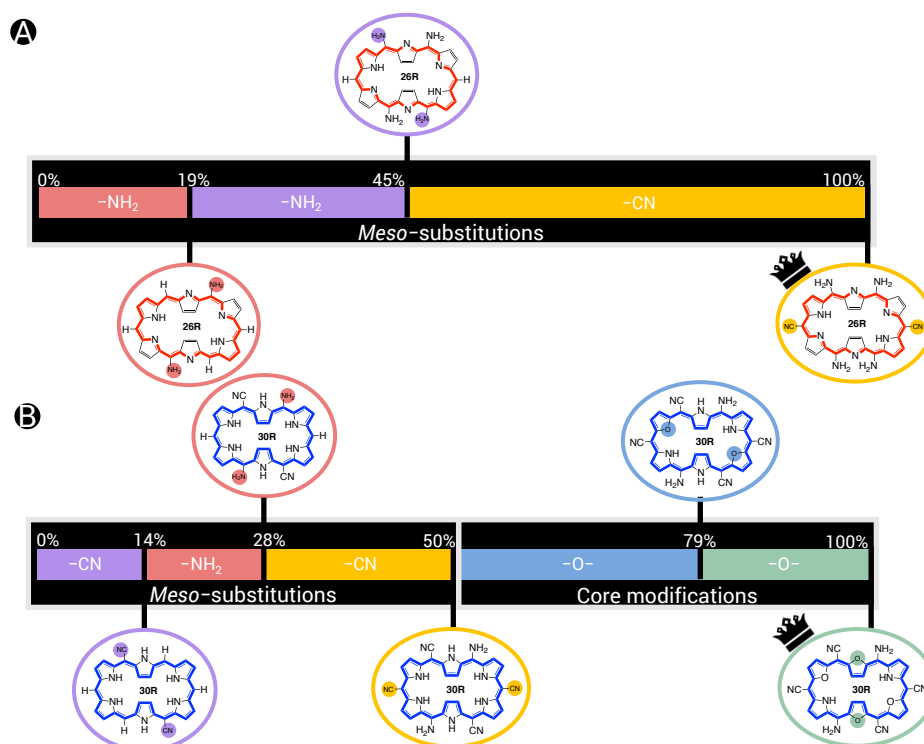


**Figure 6.** Steepest ascent of our best-performing NLO switches (A) NH\_NH\_NH<sub>2</sub>\_CN\_NH<sub>2</sub> for 26R → 28R and (B) O\_O\_NH<sub>2</sub>\_CN\_CN 30R → 28R. The percentage represents the contribution of each step to the maximal NLO response relative to the unsubstituted structure.

Remarkably, the two EDGs placed on their respective positions do not even account for half of the total NLO contrast. Consequently, the EWGs on R<sub>2,5</sub>, finally introducing a push–pull substitution pattern, are responsible for a 55% enhancement to the maximal

NLO contrast. We again take note that these CN groups only become so decisive for the NLO contrast due to the presence of the other  $\text{NH}_2$  substituents. Indeed, only including CN groups on the  $\text{R}_{2,5}$  hardly enhances the NLO contrast compared to the unsubstituted structure, but in combination with one or particularly two sets of  $\text{NH}_2$ , its full potential is shown via a synergistic effect.

For the **30R**-based molecular switch, the main findings for the steepest ascent to the optimal pattern, **O\_O\_NH<sub>2</sub>\_CN\_CN**, are summarized in Figure 6B. Here, all *meso*-substituents only account for 50% of the total NLO contrast. First, inclusion of the strongly electron-withdrawing cyano groups on  $\text{R}_{3,6}$  is favored with an increase of 14% with respect to the parent. The  $\text{NH}_2$  groups on  $\text{R}_{1,4}$  are selected as the next modification, establishing a push–pull effect, but they similarly only account for an additional 14% of the maximal NLO contrast improvement compared to the parent switch. The third set of *meso*-substituents (CN on  $\text{R}_{2,5}$ ) is responsible for a 23% further enhancement of the NLO contrast. The last functionalizations correspond to the core modifications with oxygen on sites Y and X, contributing 29% and 21%, respectively. Thus, core modifications entail a 50% increase in the NLO contrast. Due to the combination of core modifications and *meso*-substitutions, the NLO contrast is synergistically improved as shown by the steepest ascent analysis.

In summary, from our inverse design procedures, two possible functionalization combinations emerge that can synergistically enhance the NLO contrast. The first combination, the push–pull effect, is beneficial for both types of redox switches and consists of the integration of strong EDGs coupled to strong EWGs. The second functionalization combination encompasses the interplay between core modifications and *meso*-substitutions, which can synergistically amplify the NLO contrast in case of **30R** → **28R**.

### 3.1.3. Dataset Comparison: **26R** → **28R** versus **30R** → **28R**

In previous sections, we established that the two redox switches have their own ideal recipe to maximize the NLO contrast. In an attempt to generalize our findings, an analysis of the generated **26R** → **28R** database is carried out to derive guiding principles for the design of high-potential **26R**-based NLO switches, and these findings are directly compared to our previously established design rules for **30R** → **28R** [37]. For the details, we refer to Supporting Information. From our previous studies [36,37] and the BFS runs performed in this work, we collected 242 and 320 different functionalization patterns for the **26R** → **28R** and **30R** → **28R** switches, respectively. Figure 7 summarizes our main findings.

It is clear that much higher NLO contrasts can be achieved for the **26R**-based switch than for the **30R** switch, with the majority of contrast values below  $1.50 \times 10^4$  a.u. for the former and  $8.00 \times 10^3$  a.u. for the latter. However, for both the **26R** and **30R**-based switches, only select sets of 8% and 4% have contrasts larger than  $3.00 \times 10^4$  a.u. and  $1.60 \times 10^4$  a.u., respectively. Of course, the two BFS optima retrieved in Section 3.1.1 fall within these high-NLO contrast groups.

Since our interest lies in obtaining design rules for the best-performing hexaphyrin-based switches, we redistributed our two datasets into four groups based on their respective quartiles. Our target group is primarily Q3–Q4, or the best 25%, as this group contains the patterns with the highest NLO contrasts. We examine, for both hexaphyrin redox switches, which core modifications and *meso*-substitutions are the most populated in the different groups.

First, we remark again that no structures with core modifications on sites X are generated for the **26R** → **28R**, because this site combination would lead to charged ON states. In general, we observe that in our databases, structures with core modifications exceed the number of all-aza structures for both switches (Figure 7B). Nevertheless, for the **26R** → **28R** switches, most patterns without core modifications are positioned in the 50% best-performing NLO switches (Q2–Q3 and Q3–Q4). By contrast, only 6.25% of the best-performing **30R** → **28R** structures in Q2–Q3 and Q3–Q4 are not core modified (Figure 7B), and the number of all-aza structures increases, moving to lower NLO contrasts. As witnessed for the 50% best-performing **30R**-based switches in Figure 7B, there is a

preference to modify only sites Y, with the presence of core modifications on sites X being less favoured.

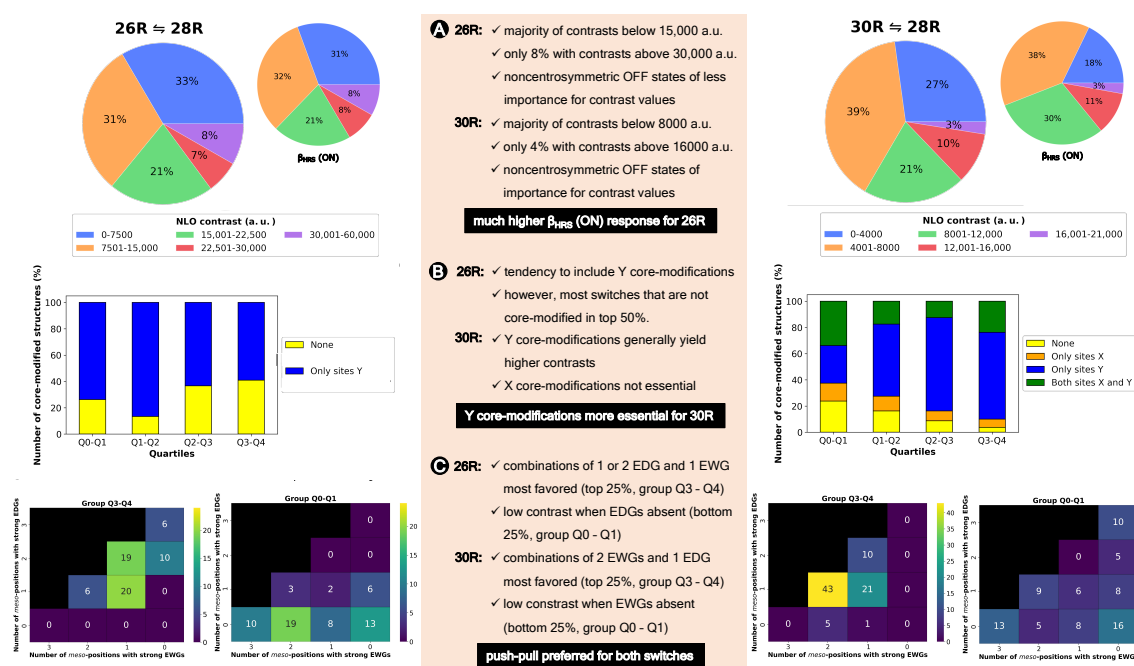


**Figure 7.** (A) Pie diagrams of NLO contrast and  $\beta_{HRS}(ON)$  of the 26R  $\rightarrow$  28R and the 30R  $\rightarrow$  28R divided into groups ranging  $7.50 \times 10^3$  a.u. and  $4.00 \times 10^3$  a.u., respectively. (B) Barplots of the number of core-modified structures in percentage versus quartiles for both switches. (C) Heatmap of the number of functionalized switches against the number of strongly electron-donating groups versus strongly electron-withdrawing groups on the *meso*-positions for the Q3–Q4 and Q0–Q1 quartiles.

Next, we focus on the *meso*-substitutions on R<sub>1,4</sub>, R<sub>2,5</sub>, and R<sub>3,6</sub>. Figure 7C displays the number of structures containing a combination of zero, one, two or three pairs of EDGs and/or EWGs in the top and bottom quartile for both types of switches. In the heatmaps, only strong EDGs and strong EWGs are counted in, so F and CH<sub>3</sub> are left out. For the 26R-based switch, combinations of one or two pairs of EDGs together with one set of EWGs are mostly present in the top 25%. The preference for more EDGs than EWGs is also witnessed from the last column, containing 16 structures, or around 25% of the group size, with a minimum of two pairs of EDGs but no EWGs. On the contrary, the majority of the worst-performing switches in group Q0–Q1 have a *meso*-substitution pattern with only EWGs and no EDGs. Hence, we can conclude that the presence of at least one set of EDGs is crucial in enhancing the NLO contrasts of 26R-based switches. For the 30R-based switches, the Q3–Q4 group consists of combinations of EWGs and EDGs, but here the largest percentage of structures contains two sets of EWGs and one set of EDGs, while such combinations are less populated within the best-performing 26R  $\rightarrow$  28R switches. Note that each switch in the top 25% contains at least one set of EWGs. Three patterns dominate in the low-contrast Q0–Q1 group: patterns with only EDGs, only EWGs and no EWGs/EDGs. Therefore, both hexaphyrin-based NLO switches benefit from a combination of EDGs and EWGs (push–pull effect) to increase the NLO contrast. Nevertheless, the NLO contrast of the [26]-hexaphyrins is also enhanced with only EDGs, while this substitution pattern significantly decreases the NLO contrast of [30]-hexaphyrins.

In summary, two different design rules emerged from the statistical analysis of the databases generated for each redox switch. First, the 26R  $\rightarrow$  28R molecular switch prefers more EDGs than EWGs on their macrocycle, although also fully EDG patterns can enhance the NLO contrast. Second, the 30R  $\rightarrow$  28R molecular switch prefers the opposite peripheral substitution pattern with more EWGs than EDGs.

### 3.2. Towards Optimal Hexaphyrin Based Three-State Molecular Switches

Due to the ability of the redox states to interconvert between each other through two-electron redox reactions [49], a multistate switch can be constructed based on two ON states (**26R** and **30R**) and a single OFF state (**28R**). The following question arises: Which of the established design rules devised for the individual switches prevails during the optimization of the three-state molecular switch? In this section, we perform two BFS optimizations from different starting structures to improve the performance of the three-state NLO switch (**26R** → **28R** → **30R**), with the performance defined as the comparable enhancement of the NLO contrast of both ends of the three-state switch, i.e., the **26R** → **28R** and **30R** → **28R** switches.

#### 3.2.1. Finding the Best Three-State Hexaphyrin Switch with the BFS Algorithm

To evaluate the efficiency of the three-state switch, a new target function must be introduced to be optimized using inverse design. Our goal is to keep the  $\beta_{HRS}(\text{OFF})$  response (**28R**) as low as possible, while maximizing the NLO response of the two ON states (**26R** and **30R**). Furthermore, we aim for a figure of merit that increases the NLO contrasts of both individual switches to a similar value. Hence, we propose Equation (5) as our new metric, with an explicit dependence on the contrast values of the **26R** → **28R** and **30R** → **28R** switches.

$$function = \begin{cases} \frac{contrast(\mathbf{26R} \rightarrow \mathbf{28R}) + contrast(\mathbf{30R} \rightarrow \mathbf{28R})}{2 \times \frac{\beta_{HRS}(\mathbf{26R})}{\beta_{HRS}(\mathbf{30R})}}, & \text{if } \beta_{HRS}(\mathbf{26R}) > \beta_{HRS}(\mathbf{30R}) \\ \frac{contrast(\mathbf{26R} \rightarrow \mathbf{28R}) + contrast(\mathbf{30R} \rightarrow \mathbf{28R})}{2 \times \frac{\beta_{HRS}(\mathbf{30R})}{\beta_{HRS}(\mathbf{26R})}}, & \text{if } \beta_{HRS}(\mathbf{26R}) < \beta_{HRS}(\mathbf{30R}) \\ \frac{contrast(\mathbf{26R} \rightarrow \mathbf{28R}) + contrast(\mathbf{30R} \rightarrow \mathbf{28R})}{2}, & \text{if } \beta_{HRS}(\mathbf{26R}) = \beta_{HRS}(\mathbf{30R}) \end{cases} \quad (5)$$

This metric penalises the average of the NLO contrast of both switches according to the dissimilarity between the  $\beta_{HRS}(\text{ON})$  responses of **26R** and **30R** states defined as the ratio between their  $\beta_{HRS}$  values. In other words, if both ON-state responses are more or less equivalent, the function is approximately equal to the average of the switch contrasts. However, if the ON-state values differ from each other by a significant factor, the contrast average is reduced by the same factor (>1).

Figure 8 summarizes the BFS runs by displaying the optimization steps with respect to the NLO function (A), the **26R** → **28R** NLO contrast (B) and the **30R** → **28R** NLO contrast (C). We select the unsubstituted **26R** → **28R** → **30R** as starting structure of this BFS run, for which four sites (Y, R<sub>1,4</sub>, R<sub>2,5</sub> and R<sub>3,6</sub>) are available for functionalization. Tables with the detailed BFS results for the different global iterations can be found in Supporting Information.

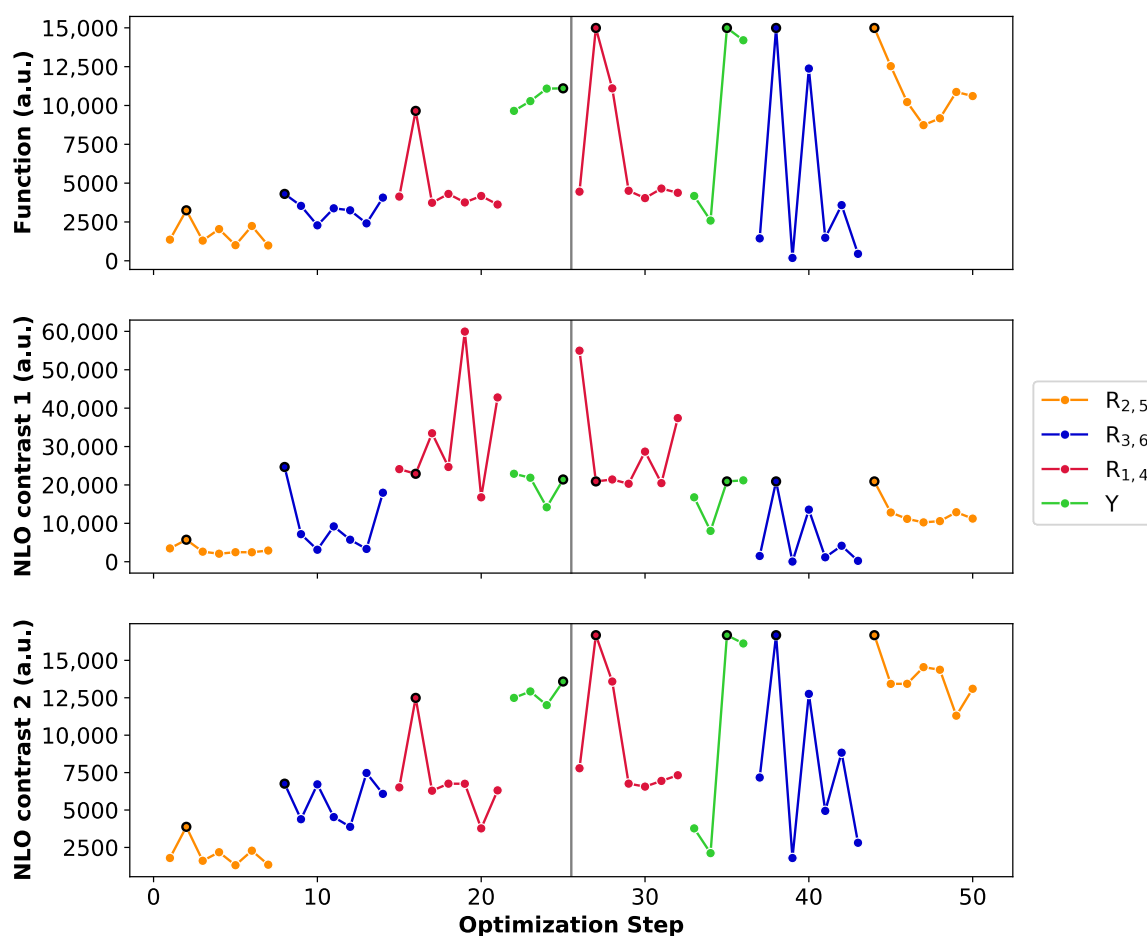


**Figure 8.** Chronological site order of the BFS procedure maximizing the function value of the **26R** → **28R** → **30R** switch starting from the unsubstituted structure. Black lines separate the global iterations. NLO contrast 1 and NLO contrast 2 stand for the individual **26R** → **28R** and **30R** → **28R** switches, respectively. We remark the different y-axis range for the NLO contrast 1.

The first sites to be optimized are the *meso*-positions  $R_2/R_5$ , where the maximum function value ( $3.25 \times 10^3$  a.u.) is obtained for the CN group. The **28R** OFF state is centrosymmetric and both the **26R** and **30R** states exhibit the highest  $\beta_{HRS}(\text{ON})$  values with this strong EWG, ensuring a maximal NLO contrast value for both redox states. Next, the  $\text{NH}_2$  group is incorporated on the  $R_3/R_6$  *meso*-positions creating a push–pull pattern with a function value of  $4.31 \times 10^3$  a.u. The OFF state remains centrosymmetric and the **26R**-based NLO contrast is the highest among all possible substitutions, while the **30R**-based contrast is the second best, after the CN-substituted one. However, as CN performs the worst for **26R** → **28R**, the resulting function value is half as low. This example nicely highlights the clear difference in preferred functionalization between the redox states. Substituting *meso*-positions  $R_1/R_4$  results in all OFF states being centrosymmetric, except for  $\text{NO}_2$ . The maximal contrast value for the **26R** → **28R** switch reaches  $5.99 \times 10^4$  a.u. for  $R_1/R_4 = \text{NH}_2$ , which corresponds to the BFS optimum for the individual **26R** → **28R** switch (Section 3.1.1, Figure 5C). For the **30R**-based switch, the maximal NLO contrast is at  $1.25 \times 10^4$  a.u. containing CN instead of  $\text{NH}_2$  at  $R_1/R_4$ . Since the average NLO contrast for the **30R** → **28R** is much lower than for the **26R** → **28R**, our proposed metric depends more strongly on the contrast of the **30R** → **28R** switch. Indeed, the highest function value is attained for the  $\text{NH\_NH\_CN\_CN\_NH}_2$  *meso*-substituted pattern with a value of  $9.65 \times 10^3$  a.u. We note that the function value indeed becomes closer to the ON state values of both individual switches. Finally, the different core modifications are tested for this *meso*-substitution optimum. All three core modifications (O, S or Se)

provide higher function values than the three-level switch without core modifications, with small differences between the various heteroatoms. The best-performing three-state switch retrieved during this first global iteration has the **NH\_S\_CN\_CN\_NH<sub>2</sub>** pattern with a function value of  $1.12 \times 10^4$  a.u. and has the highest NLO contrast for **30R**. The **26R**-based switch still prefers no core modifications, making the **30R**-based switch the decisive factor for selecting the optimal pattern of the three-state switch.

The second global iteration immediately starts with finding a new optimum with a function value of  $1.50 \times 10^4$  a.u. by replacing the CN group on the *meso*-positions  $R_1/R_4$  by NO<sub>2</sub>. Even though there exist several substituents yielding a higher **26R** → **28R** contrast, the corresponding NLO contrast of **30R** → **28R** becomes significantly lower for these patterns, making them unsuitable for our metric. The remaining three site optimizations (Y,  $R_3/R_6$  and  $R_2/R_5$ ) do not improve the performance of the three-state switch. Remarkably, for these site optimizations, the highest function values align with the highest or near-highest individual **26R**- and **30R**-based NLO contrasts. Also, we observe that introducing NH or O on the sites Y results in noncentrosymmetric OFF states, significantly decreasing the individual contrasts. Additionally, we again find that EDGs on the  $R_3/R_6$  positions are essential for both switches to maintain the push–pull substitution pattern. The BFS optimum for the three-state switch is thus the **NH\_S\_NO<sub>2</sub>\_CN\_NH<sub>2</sub>** pattern with a function value of  $1.50 \times 10^4$  a.u., which is only 1.4 times lower than the optimal **30R**-based switch but 4 times lower than the best-performing **26R**-based switch.

In our previous study on **30R** → **28R**, this pattern, **NH\_S\_NO<sub>2</sub>\_CN\_NH<sub>2</sub>**, was ranked as the sixth best performing pattern generated by the BFS procedures [37]. This observation is in parallel with our proposed function value for the performance of multistate switches being more prone towards better performing **30R** → **28R** than **26R** → **28R** due to the lower contrast values found for the former. To confirm whether an alternative pattern could be found closer to the best-performing patterns for **26R** → **28R**, we carried out an additional BFS optimization starting from the most optimal structure found in Section 3.1.1, namely **NH\_NH\_NH<sub>2</sub>\_CN\_NH<sub>2</sub>**. Convincingly, after four global iterations, the algorithm results in the same optimum as our initial BFS procedure. However, the question still remains whether a synergistic effect exists between the different types of functionalizations in the optimal switch.

### 3.2.2. Steepest Ascent of the Best **26R** → **28R** → **30R** Molecular Switch

To answer this question, we apply the steepest ascent method (Table 2 and Figure 9) following the largest gradient in the function value to build up our optimal three-state switch.

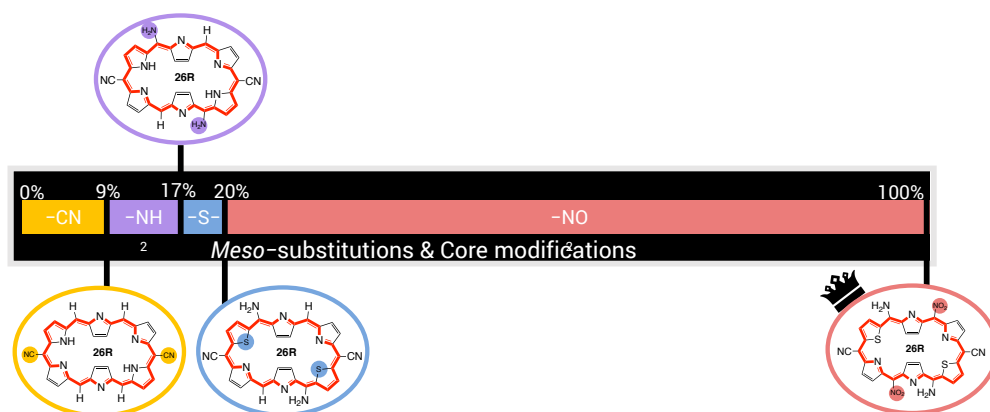


**Figure 9.** Steepest ascent of our best-performing three-state switch for NLO applications with pattern **NH\_S\_NO<sub>2</sub>\_CN\_NH<sub>2</sub>**. The percentage quantifies the contribution of each steepest ascent step to the maximal NLO response as defined by our function value relative to the unsubstituted structure.

**Table 2.** Steepest ascent of our best-performing multistate switch with pattern **NH\_S\_NO<sub>2</sub>\_CN\_NH<sub>2</sub>**. %B represents the percentage in NLO function of all switches in a steepest ascent step with respect to the NLO contrast maximum. %(B-P) is similar to %B but takes the parent as a reference. %(CB) stands for the percentage in NLO function of all switches in a steepest ascent step with respect to the current best.

| Step   | X  | Y  | R <sub>1,4</sub> | R <sub>2,5</sub> | R <sub>3,6</sub> | contrast <sub>26R→28R</sub> (a.u.)   | Contrast <sub>30R→28R</sub> (a.u.)   | Function (a.u.)                      | %B         | %(B-P)     | %(CB)      |
|--------|----|----|------------------|------------------|------------------|--------------------------------------|--------------------------------------|--------------------------------------|------------|------------|------------|
| Parent | NH | NH | H                | H                | H                | $2.09 \times 10^3$                   | $2.18 \times 10^3$                   | $2.04 \times 10^3$                   | 14         | 0          | -          |
| 1      | NH | S  | H                | H                | H                | $1.36 \times 10^3$                   | $2.01 \times 10^3$                   | $1.15 \times 10^3$                   | 8          | -7         | -74        |
|        | NH | NH | NO <sub>2</sub>  | H                | H                | $9.52 \times 10^0$                   | $3.47 \times 10^3$                   | $3.00 \times 10^2$                   | 2          | -13        | -144       |
|        | NH | NH | H                | CN               | H                | <b><math>5.73 \times 10^3</math></b> | <b><math>3.88 \times 10^3</math></b> | <b><math>3.25 \times 10^3</math></b> | <b>22</b>  | <b>9</b>   | <b>100</b> |
|        | NH | NH | H                | H                | NH <sub>2</sub>  | $1.94 \times 10^4$                   | $3.49 \times 10^3$                   | $2.30 \times 10^3$                   | 15         | 2          | 22         |
| 2      | NH | S  | H                | CN               | H                | $3.87 \times 10^3$                   | $1.35 \times 10^4$                   | $3.75 \times 10^3$                   | 25         | 13         | 75         |
|        | NH | NH | NO <sub>2</sub>  | CN               | H                | $1.35 \times 10^3$                   | $4.31 \times 10^3$                   | $1.57 \times 10^3$                   | 10         | -4         | -21        |
|        | NH | NH | H                | CN               | NH <sub>2</sub>  | <b><math>2.47 \times 10^4</math></b> | <b><math>6.76 \times 10^3</math></b> | <b><math>4.31 \times 10^3</math></b> | <b>29</b>  | <b>17</b>  | <b>100</b> |
| 3      | NH | S  | H                | CN               | NH <sub>2</sub>  | <b><math>2.04 \times 10^4</math></b> | <b><math>6.94 \times 10^3</math></b> | <b><math>4.65 \times 10^3</math></b> | <b>31</b>  | <b>20</b>  | <b>100</b> |
|        | NH | NH | NO <sub>2</sub>  | CN               | NH <sub>2</sub>  | $1.68 \times 10^4$                   | $3.77 \times 10^3$                   | $4.17 \times 10^3$                   | 28         | 16         | 82         |
| 4      | NH | S  | NO <sub>2</sub>  | CN               | NH <sub>2</sub>  | $2.09 \times 10^4$                   | $1.67 \times 10^4$                   | <b><math>1.50 \times 10^4</math></b> | <b>100</b> | <b>100</b> | <b>100</b> |

The first build-up step shows that only CN at R<sub>2</sub>/R<sub>5</sub> and NH<sub>2</sub> at R<sub>3</sub>/R<sub>6</sub> yield improved function values compared to the parent switch, with the former leading to the largest function value ( $3.25 \times 10^3$  a.u.), a 9% increase over the parent. Whereas the NH<sub>2</sub> group would be the favored choice for **26R**, CN at R<sub>2</sub>/R<sub>5</sub> is preferred for **30R**. This seems to contradict the results of the steepest ascent on the **30R** → **28R** optimum. However, our previous research revealed that the exact positioning of EDGs and EWGs on the macrocycles (R<sub>1</sub>/R<sub>4</sub> vs. R<sub>3</sub>/R<sub>6</sub>) can also substantially impact the efficiency of the switch [37]. Next, NH<sub>2</sub> is positioned on R<sub>3</sub>/R<sub>6</sub> and is responsible for an additional increase of 8% in the NLO contrast metric. We note that immediately introducing a second strong EWG on the macrocycle is disadvantageous for the performance of the three-state switch. Curiously though, this step follows the **26R** preference and not that of **30R**. In the third step, the core modification is included, but it only contributes 3%. Although the contribution of S is small, its impact mainly relies on the conservation of the centrosymmetric OFF state. Indeed, the  $\beta_{HRS}$  value of **28R(NH\_S\_H\_CN\_NH<sub>2</sub>)** is 0 a.u., while **28R(NH\_NH\_NO<sub>2</sub>\_CN\_NH<sub>2</sub>)** has a  $\beta_{HRS}$  value of  $2.92 \times 10^3$  a.u. Unexpectedly, the three functionalizations so far only contribute 20% to the best-performing three-state switch. That means that 80% of the multistate optimum's enhancement is ascribed to the insertion of NO<sub>2</sub> on R<sub>1</sub>/R<sub>4</sub>. Due to the presence of the S core modifier, introducing NO<sub>2</sub> finally yielded a centrosymmetric OFF state, unlike the NO<sub>2</sub> patterns in the previous steps. The synergistic effect of the Y core modifications manifests itself in the response of the OFF state, which is in stark contrast to our recent findings on the **30R** → **28R** switch where the core modifications were responsible for the enhancement of the  $\beta_{HRS}(\text{ON})$  response [37].

### 3.2.3. Chemical Space Visualization

Altogether, 146 patterns were retrieved for the three-state switch. As mentioned in Section 3.1.1, the NLO contrast of hexaphyrin-based switches is mainly dependent on the  $\beta_{HRS}(\text{ON})$  response. Taking a closer look at the  $\beta_{HRS}$  values over all ON states generated during the BFS procedures on the multistate switch (Figure 10), we remark that the largest values correspond to **26R** ON states. Only 8% of all generated switches have ON responses above  $2.25 \times 10^4$  a.u., and all correspond to **26R** states. The majority of **30R** structures have  $\beta_{HRS}(\text{ON})$  responses in the range of 0–15,000 a.u. Hence, **26R** → **28R** is more likely to outperform the **30R** → **28R**-based switches in terms of individual NLO contrast.

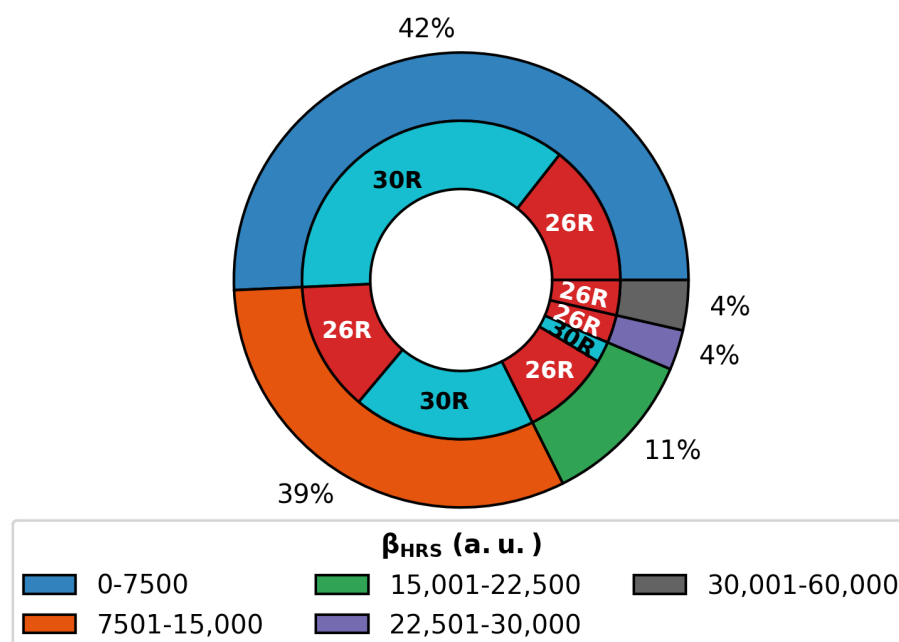


**Figure 10.** Pie diagram of  $\beta_{HRS}$  of the **26R** and **30R** structures divided into groups with ranges of 7500 a.u. For each group, the number of **26R** and **30R** structures is colored in red and blue, respectively.

Because the BFS procedure sequentially visits all modifiable sites, a large set of structures in the vicinity of first the initial structure and later each intermediate optimum is generated, with each optimization step covering larger portions of the CCS. To investigate which regions of the space yield the most efficient three-state switches, we visualized the CCS through t-SNE plots by confining the space through the extended connectivity fingerprints (ECFP) (Figure 11). Different measures were used to color the data points of the t-SNE plots. Regions of interest were aggregated by a colored shape. We note that the t-SNE plots contain all functionalization patterns gathered in this and previous studies from our group and are therefore not restricted to the three-state switches; a complete list is included in Supporting Information.

We started by investigating the regions that contain a certain number of core modifications (Figure 11A). Almost all structures with either none or two sets of core modifications were grouped together in a red oval and blue box, respectively. If we compare the CCS of the **26R**  $\rightarrow$  **28R** (Figure 11B) and **30R**  $\rightarrow$  **28R** switches (Figure 11C), we observe that distinct regions of the CCS are favoured depending on the type of redox switch. A large part of the structures that are not core modified are performing significantly well for **26R**-based NLO contrasts, as represented by the overlap between the red and yellow ovals. We note that there are no **26R**-based switches located in the blue rectangle, as two sets of core modification sites were not examined for this type of switch. On the other hand, for the **30R**-based NLO contrast, we found more switches with high NLO contrasts in this blue region, as indicated by a green oval within the blue rectangle in Figure 11C. Furthermore, the region with no core modifications (red oval) does not contain many structures with high NLO contrasts. Other smaller regions marked by a yellow or green oval were found outside of these blue and red regions. These structures only had one out of two core modification sites modified. Finally, Figure 11D highlights the data points according to their function value as metric for the NLO contrast of the three-state switch. Most of the top-performing three-state switches are positioned in the regions with a high **30R**  $\rightarrow$  **28R** NLO contrast, and some of them reside in the intersection of the yellow and green ovals, regions delimited by a black line in Figure 11D. This demonstrates that, although maximizing the **30R**  $\rightarrow$  **28R** NLO contrast is key for elevating the performance of the three-state switch, it does not imply that these switches have patterns very different from the ideal **26R**  $\rightarrow$  **28R**.

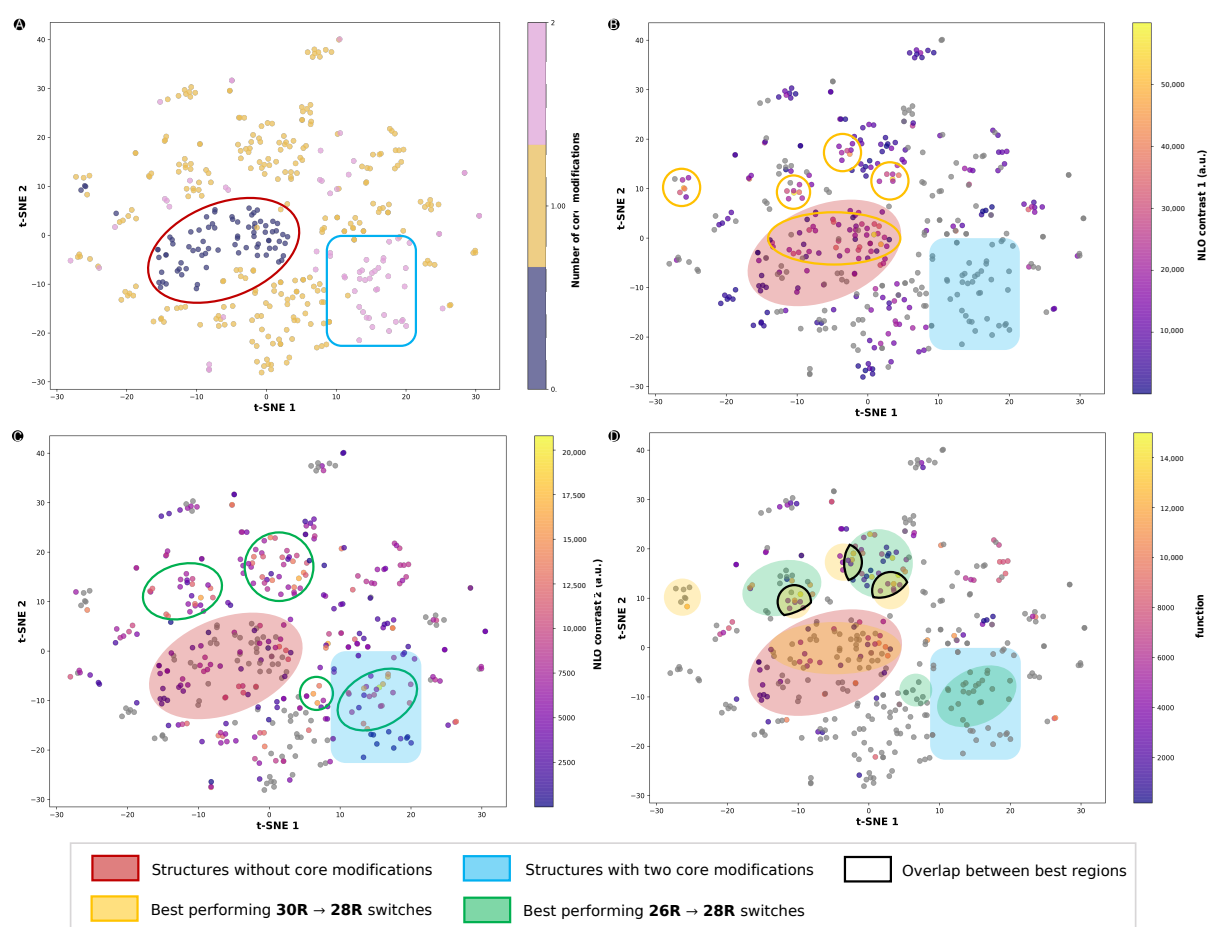


**Figure 11.** t-SNE plots of the chemical compound space described by the ECFP fingerprints with different measures: (A) number of core-modified sites (S, O or Se), (B) the  $26R \rightarrow 28R$  NLO contrast, (C) the  $30R \rightarrow 28R$  NLO contrast and (D) the function value of the three-state switches. NLO contrasts 1 and 2 refer to  $26R$ - and  $30R$ -based switches, respectively. Regions of the CCS highlighted by a red oval are structures without core modifications, while the blue rectangle represents structures with two core-modified sites, X and Y. Green and yellow ovals highlight regions with high NLO contrasts for  $30R \rightarrow 28R$  and  $26R \rightarrow 28R$ , respectively.

#### 4. Conclusions

Building further on our previous work on the design of functionalized hexaphyrin-based molecular switches with enhanced NLO contrast, we investigate the effect of core modifications in combination with *meso*-substitutions for the  $26R \rightarrow 28R$  redox switch with the help of the BFS algorithm. The optimal  $26R$ -based switch combines two sets of EDGs and one set of EWGs, but does not include any core modifiers. In several facets, this optimal functionalization differs from the optimal pattern found for the  $30R \rightarrow 28R$  molecular switch [37], despite the structural similarity of both redox switches. Not only do core modifications synergistically help to enhance the contrasts of  $30R$ -based switches, but these switches are also characterized by a distinct push–pull character, with two pairs of EWGs combined with one pair of EDGs. A thorough analysis of our large  $26R \rightarrow 28R$  and  $30R \rightarrow 28R$  databases confirms that the design principles devised for the BFS optima can be generalized for the overall best-performing hexaphyrin-based switches. Through a steepest ascent analysis of the optima, we also conclude that there is a synergistic effect in amplifying the ON state's NLO response between either the *meso*-substituents of the push–pull patterns or between the core modifications and the push–pull substitution pattern.

Because the three redox states are interconvertible between each other, we additionally optimize the performance of the three-state switch  $26R \rightarrow 28R \rightarrow 30R$  by maximizing

the response for both ON states (**26R** and **30R**) and minimizing the response of the OFF state (**28R**). Our aim is to discover which of the design rules would prevail during the BFS optimization for the multistate switch. Our best-performing three-state switch has a pattern similar to the **30R** → **28R** due to the lower NLO response of the **30R** state compared to that of **26R**. Again, a synergistic effect is responsible for the enhancement of the performance of the three-state switch, although in this case, the synergistic effect manifests itself by keeping the OFF state centrosymmetric instead of enhancing the ON state. Finally, we visualized the chemical space and identified the regions that represent enhanced **26R** → **28R** and **30R** → **28R** switches. Even though high-potential three-state switches are generally found in regions of increased **30R**-based systems, some overlap with the **26R** → **28R** regions is seen, demonstrating the compatibility of the two types of switches despite their dissimilarities.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules28217371/s1>, Table S1: Global iteration structures of the BFS procedure on the maximization of the NLO contrast of the **28R** ⇌ **26R** switch with the X<sub>Y</sub>R<sub>1,4</sub>R<sub>2,5</sub>R<sub>3,6</sub> pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [28]hexaphyrins and [26]hexaphyrins are given in a.u.; Table S2: Global iteration structures of the BFS procedure on the maximization of the NLO contrast of the **28R** ⇌ **30R** switch with the X<sub>Y</sub>R<sub>1,4</sub>R<sub>2,5</sub>R<sub>3,6</sub> pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [28]hexaphyrins and [30]hexaphyrins are given in a.u.; Table S3: Global iteration structures of the BFS procedure on the maximization of the NLO contrast of the **26R** → **28R** → **30R** switch with the X<sub>Y</sub>R<sub>1,4</sub>R<sub>2,5</sub>R<sub>3,6</sub> pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [26]hexaphyrins, [28]hexaphyrins and [30]hexaphyrins are given in a.u.; Table S4: Global iteration structures of the BFS procedure on the maximization of the NLO contrast of the **26R** → **28R** → **30R** switch with the X<sub>Y</sub>R<sub>1,4</sub>R<sub>2,5</sub>R<sub>3,6</sub> pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [26]hexaphyrins, [28]hexaphyrins and [30]hexaphyrins are given in a.u.; Table S5: Quartiles and interquartile range (IQR) of NLO contrasts (in a.u.) of the **26R** and **30R**-based molecular switches.; Table S6: Collection of the dataset generated during the BFS and prestudy containing the individual first hyperpolarizabilities of the OFF- and ON-state, NLO contrast of the **26R** ⇌ **28R** switch pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [26]hexaphyrins and [28]hexaphyrins are given in a.u. Table S7: Collection of the dataset generated during the BFS and prestudy containing the individual first hyperpolarizabilities of the OFF- and ON-state, NLO contrast of the **30R** ⇌ **28R** switch pattern. The static hyper-Rayleigh scattering first hyperpolarizability values of the [30]hexaphyrins and [28]hexaphyrins are given in a.u. Figure S1: (A) Pie diagram of NLO contrast of the **26R** → **28R** divided into groups ranging 7500 a.u., (B) Pie diagram of NLO contrast of the **30R** → **28R** divided into groups ranging 4000 a.u. (C) Violin plot describing the distribution of **26R** → **28R** NLO contrasts. (D) Violin plot describing the distribution of **30R** → **28R** NLO contrasts. (E) Pie diagram of  $\beta_{HRS}(\text{ON})$  of the **26R** → **28R** divided into groups ranging 7500 a.u., (F) Pie diagram of  $\beta_{HRS}(\text{ON})$  of the **30R** → **28R** divided into groups ranging 4000 a.u. Figure S2: Barplots of the number of core-modified structures in percentage *versus* quartiles for (A) **26R** → **28R** and (B) **30R** → **28R**, Figure S3: Barplots of number of structures with core-modifications on sites Y colored per type and distributed within the quartiles for (A) **26R** → **28R** and (B) **30R** → **28R**. Figure S4: Heatmap of the number of functionalized structures against the number of strongly electron-donating groups versus strongly electron-withdrawing groups on the *meso*-positions, grouped per quartile for **26R** → **28R** (A) and **30R** → **28R** (B), respectively. References [50–67] are cited in the Supplementary Materials.

**Author Contributions:** M.A. and F.D.V. conceptualized the project. E.D. and L.S.G. performed all quantum chemical calculations. M.A. and F.D.V. supervised the project. E.D. analyzed the data. The first draft of the manuscript was written by E.D., F.D.V. and M.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Research Foundation Flanders (FWO-11E0321N).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data will be made available upon request.

**Acknowledgments:** F.D.V. and M.A. wish to thank the VUB for the Strategic Research Program awarded to the ALGC research group. E.D. thanks the Fund for Scientific Research-Flanders (FWO-11E0321N) for financial support. In addition, E.D. thanks Michiel Jacobs for brainstorming about the visualization of the t-SNE plots of the chemical space. The resources and services used in this work were provided by the VSC (Flemish Super-computer Center), funded by the Research Foundation—Flanders (FWO) and the Flemish Government.

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

## References

1. Feringa, B.L. The art of building small: From molecular switches to motors (Nobel Lecture). *Angew. Chem. Int. Ed.* **2017**, *56*, 11060–11078. [[CrossRef](#)] [[PubMed](#)]
2. Steen, J.D.; Duijnste, D.R.; Browne, W.R. Molecular switching on surfaces. *Surf. Sci. Rep.* **2023**, *78*, 100596. [[CrossRef](#)]
3. Mondal, A.; Toyoda, R.; Costil, R.; Feringa, B.L. Chemically driven rotatory molecular machines. *Angew. Chem. Int. Ed.* **2022**, *61*, e202206631. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, J.L.; Zhong, J.Q.; Lin, J.D.; Hu, W.P.; Wu, K.; Xu, G.Q.; Wee, A.T.S.; Chen, W. Towards single molecule switches. *Chem. Soc. Rev.* **2015**, *44*, 2998–3022. [[CrossRef](#)] [[PubMed](#)]
5. Lerch, M.M.; Hansen, M.J.; van Dam, G.M.; Szymanski, W.; Feringa, B.L. Emerging targets in photopharmacology. *Angew. Chem. Int. Ed.* **2016**, *55*, 10978–10999. [[CrossRef](#)]
6. Ghorbani-Choghamarani, A.; Taherinia, Z. Recent advances utilized in artificial switchable catalysis. *RSC Adv.* **2022**, *12*, 23595–23617. [[CrossRef](#)]
7. Natali, M.; Giordani, S. Molecular switches as photocontrollable “smart” receptors. *Chem. Soc. Rev.* **2012**, *41*, 4010–4029. [[CrossRef](#)]
8. Sun, L.; Diaz-Fernandez, Y.A.; Gschneidner, T.A.; Westerlund, F.; Lara-Avila, S.; Moth-Poulsen, K. Single-molecule electronics: From chemical design to functional devices. *Chem. Soc. Rev.* **2014**, *43*, 7378–7411. [[CrossRef](#)]
9. Rajeshirke, M.; Sekar, N. Multi-stimuli responsive emissive NLOphoric colorants—A recent trend in research. *Dyes Pigms.* **2019**, *163*, 675–683. [[CrossRef](#)]
10. Chandler, H.J.; Stefanou, M.; Campbell, E.E.B.; Schaub, R. Li@C60 as a multi-state molecular switch. *Nat. Commun.* **2019**, *10*, 2283. [[CrossRef](#)]
11. Beaujean, P.; Sanguinet, L.; Rodriguez, V.; Castet, F.; Champagne, B. Multi-State second-order nonlinear optical switches incorporating one to three benzazolo-oxazolidine units: A quantum chemistry investigation. *Molecules* **2022**, *27*, 2770. [[CrossRef](#)]
12. Castet, F.; Rodriguez, V.; Pozzo, J.L.; Ducasse, L.; Plaquet, A.; Champagne, B. Design and characterization of molecular nonlinear optical switches. *Acc. Chem. Res.* **2013**, *46*, 2656–2665. [[CrossRef](#)] [[PubMed](#)]
13. Coe, B.J. Molecular materials possessing switchable quadratic nonlinear optical properties. *Chem. Eur. J.* **1999**, *5*, 2464–2471. [[CrossRef](#)]
14. Delaire, J.A.; Nakatani, K. Linear and nonlinear optical properties of photochromic molecules and materials. *Chem. Rev.* **2000**, *100*, 1817–1846. [[CrossRef](#)] [[PubMed](#)]
15. Gao, S.; Yi, X.; Shang, J.; Liu, G.; Li, R.W. Organic and hybrid resistive switching materials and devices. *Chem. Soc. Rev.* **2019**, *48*, 1531–1565. [[CrossRef](#)]
16. Nakano, M.; Champagne, B. Nonlinear optical properties in open-shell molecular systems. *WIREs Comput. Mol. Sci.* **2016**, *6*, 198–210. [[CrossRef](#)]
17. Castet, F.; Tonnelé, C.; Muccioli, L.; Champagne, B. Predicting the second-order nonlinear optical responses of organic materials: The role of dynamics. *Acc. Chem. Res.* **2022**, *55*, 3716–3726. [[CrossRef](#)]
18. Hassan, A.U.; Mohyuddin, A.; Güler, C.; Nadeem, S.; Nkungli, N.K.; Hassan, S.U.; Javed, M. Novel pull-push organic switches with D- $\pi$ -A structural designs: computational design of star shape organic materials. *Struct. Chem.* **2023**, *34*, 399–412. [[CrossRef](#)]
19. Avramopoulos, A.; Zalesny, R.; Reis, H.; Papadopoulos, M.G. A computational strategy for the design of photochromic derivatives based on diarylethene and Nickel dithiolene with large contrast in nonlinear optical properties. *J. Phys. Chem. C* **2020**, *124*, 4221–4241. [[CrossRef](#)]
20. Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **2018**, *361*, 360–365. [[CrossRef](#)]
21. Green, J.D.; Fuemmeler, E.G.; Hele, T.J.H. Inverse molecular design from first principles: Tailoring organic chromophore spectra for optoelectronic applications. *J. Chem. Phys.* **2022**, *156*, 180901. [[CrossRef](#)] [[PubMed](#)]
22. Teunissen, J.L.; De Proft, F.; De Vleeschouwer, F. Tuning the HOMO–LUMO energy gap of small diamondoids using inverse molecular design. *J. Chem. Theory Comput.* **2017**, *13*, 1351–1365. [[CrossRef](#)] [[PubMed](#)]
23. Molesky, S.; Lin, Z.; Piggott, A.Y.; Jin, W.; Vucković, J.; Rodriguez, A.W. Inverse design in nanophotonics. *Nat. Photonics* **2018**, *12*, 659–670. [[CrossRef](#)]

24. Pollice, R.; dos Passos Gomes, G.; Aldeghi, M.; Hickman, R.J.; Krenn, M.; Lavigne, C.; Lindner-D'Addario, M.; Nigam, A.; Ser, C.T.; Yao, Z.; et al. Data-driven strategies for accelerated materials design. *Acc. Chem. Res.* **2021**, *54*, 849–860. [\[CrossRef\]](#) [\[PubMed\]](#)
25. De Vleeschouwer, F.; Geerlings, P.; De Proft, F. Molecular property optimizations with boundary conditions through the best first search scheme. *ChemPhysChem* **2016**, *17*, 1414–1424. [\[CrossRef\]](#)
26. Teunissen, J.L.; De Proft, F.; De Vleeschouwer, F. Acceleration of inverse molecular design by using predictive techniques. *J. Chem. Inf. Model.* **2019**, *59*, 2587–2599. [\[CrossRef\]](#)
27. Alonso, M.; Geerlings, P.; De Proft, F. Viability of Möbius topologies in [26]- and [28]hexaphyrins. *Chem. Eur. J.* **2012**, *18*, 10916–10928. [\[CrossRef\]](#)
28. Woller, T.; Geerlings, P.; De Proft, F.; Champagne, B.; Alonso, M. Fingerprint of aromaticity and molecular topology on the photophysical properties of octaphyrins. *J. Phys. Chem. C* **2019**, *123*, 7318–7335. [\[CrossRef\]](#)
29. Yu, D.; Rong, C.; Lu, T.; Geerlings, P.; De Proft, F.; Alonso, M.; Liu, S. Switching between Hückel and Möbius aromaticity: A density functional theory and information-theoretic approach study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4715–4730. [\[CrossRef\]](#)
30. Stuyver, T.; Perrin, M.; Geerlings, P.; De Proft, F.; Alonso, M. Conductance switching in expanded porphyrins through aromaticity and topology changes. *J. Am. Chem. Soc.* **2018**, *140*, 1313–1326. [\[CrossRef\]](#)
31. Torrent-Sucarrat, M.; Navarro, S.; Marcos, E.; Anglada, J.M.; Luis, J.M. Design of Hückel–Möbius topological switches with high nonlinear optical properties. *J. Phys. Chem. C* **2017**, *121*, 19348–19357. [\[CrossRef\]](#)
32. Tanaka, T.; Osuka, A. Chemistry of Meso-Aryl Expand. Porphyrins: Aromat. Mol. Twist. *Chem. Rev.* **2016**, *117*, 2584–2640. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Sung, Y.M.; Oh, J.; Cha, W.Y.; Kim, W.; Lim, J.M.; Yoon, M.C.; Kim, D. Control and switching of aromaticity in various all-aza-expanded porphyrins: Spectroscopic and theoretical analyses. *Chem. Rev.* **2017**, *117*, 2257–2312. [\[CrossRef\]](#)
34. Casademont-Reig, I.; Woller, T.; García, V.; Contreras-García, J.; Tiznado, W.; Torrent-Sucarrat, M.; Matito, E.; Alonso, M. Quest for the most aromatic pathway in charged expanded porphyrins. *Chem. Eur. J.* **2023**, *29*, e202202264. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Bettens, T.; Hoffmann, M.; Alonso, M.; Geerlings, P.; Dreuw, A.; De Proft, F. Mechanochemically triggered topology changes in expanded porphyrins. *Chem. Eur. J.* **2021**, *27*, 3397–3406. [\[CrossRef\]](#)
36. Desmedt, E.; Woller, T.; Teunissen, J.L.; De Vleeschouwer, F.; Alonso, M. Fine-tuning of nonlinear optical contrasts of hexaphyrin-based molecular switches using inverse design. *Front. Chem.* **2021**, *9*, 786036. [\[CrossRef\]](#)
37. Desmedt, E.; Smets, D.; Woller, T.; Alonso, M.; De Vleeschouwer, F. Designing hexaphyrins for high-potential NLO switches: The synergy of core-modifications and Meso-Substitutions. *Phys. Chem. Chem. Phys.* **2023**, *25*, 17128–17142. [\[CrossRef\]](#)
38. Pearl, J.; Korf, R.E. Search techniques. *Ann. Rev. Comp. Sci.* **1987**, *2*, 451–467. [\[CrossRef\]](#)
39. Gordon, D.B.; Mayo, S.L. Branch-and-terminate: A combinatorial optimization algorithm for protein design. *Structure* **1999**, *7*, 1089–1098. [\[CrossRef\]](#)
40. Dechter, R.; Pearl, J. Generalized best-first search strategies and the optimality of A\*. *J. Assoc. Comput. Mach.* **1985**, *32*, 505–536. [\[CrossRef\]](#)
41. Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: New York, NY, USA, 1998.
42. David E. Goldberg, Kalyanmoy Deb, J.H.C. Genetic algorithms, noise, and the sizing of populations. *Complex Syst.* **1992**, *6*, 333–362.
43. Strasser, S.; Goodman, R.; Sheppard, J.; Butcher, S. A new discrete particle swarm optimization algorithm. In Proceedings of the Genetic and Evolutionary Computation Conference 2016, GECCO '16, Denver, CO, USA, 20–24 July 2016.
44. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; The MIT Press: New York, NY, USA, 2005.
45. Häse, F.; Roch, L.M.; Kreisbeck, S.; Aspuru-Guzik, A. Phoenix: A Bayesian optimizer for chemistry. *ACS Cent. Sci.* **2018**, *4*, 1134–1145. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Verbiest, T.; Clays, K.; Rodriguez, V. *Second-Order Nonlinear Optical Characterization Techniques: An Introduction*; CRC Press: Boca Raton, FL, USA, 2009.
47. Clays, K.; Persoons, A. Hyper-rayleigh scattering in solution. *Phys. Rev. Lett.* **1991**, *66*, 2980–2983. [\[CrossRef\]](#)
48. Hendrickx, E.; Clays, K.; Persoons, A. Hyper-Rayleigh scattering in isotropic solution. *Acc. Chem. Res.* **1998**, *31*, 675–683. [\[CrossRef\]](#)
49. Ambhore, M.D.; Basavarajappa, A.; Anand, V.G. A wide-range of redox states of core-modified expanded porphyrinoids. *Chem. Commun.* **2019**, *55*, 6763–6766. [\[CrossRef\]](#)
50. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Petersson, G.; Nakatsuji, H.; et al. *Gaussian 16 Revision A.01*; Gaussian Inc.: Wallingford, CT, USA, 2016.
51. Yanai, T.; Tew, D.; Handy, N. A new hybrid exchange–correlation functional using the coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57. [\[CrossRef\]](#)
52. Hehre, W.J.; Radom, L.; Schleyer, P.v.R.; Pople, J.A. *Ab Initio Molecular Orbital Theory*; Wiley: Hoboken, NJ, USA, 1986.
53. Torrent-Sucarrat, M.; Navarro, S.; Cossío, F.; Anglada, J.; Luis, J. Relevance of the DFT method to study expanded porphyrins with different topologies. *J. Comput. Chem.* **2017**, *38*, 2819–2828. [\[CrossRef\]](#)
54. Woller, T.; Banerjee, A.; Sylvetsky, N.; Santra, G.; Deraet, X.; De Proft, F.; Martin, J.; Alonso, M. Performance of electronic structure methods for the description of Hückel–Möbius interconversions in extended  $\pi$ -systems. *J. Phys. Chem. A* **2020**, *124*, 2380–2397. [\[CrossRef\]](#)

55. Sylvestry, N.; Banerjee, A.; Alonso, M.; Martin, J. Performance of localized coupled cluster methods in a moderately strong correlation regime: Hückel–Möbius interconversions in expanded porphyrins. *J. Chem. Theory Comput.* **2020**, *16*, 3641–3653. [CrossRef]
56. Torrent-Sucarrat, M.; Anglada, J.; Luis, J. Evaluation of the nonlinear optical properties for an expanded porphyrin Hückel–Möbius aromaticity switch. *J. Chem. Phys.* **2012**, *137*, 184306. [CrossRef]
57. Torrent-Sucarrat, M.; Anglada, J.; Luis, J. Evaluation of the nonlinear optical properties for annulenes with Hückel and Möbius topologies. *J. Chem. Theory Comput.* **2011**, *7*, 3935–3943. [CrossRef] [PubMed]
58. Plaquet, A.; Guillaume, M.; Champagne, B.; Castet, F.; Ducasse, L.; Pozzo, J.; Rodriguez, V. In silico optimization of merocyanine-spiropyran compounds as second-order nonlinear optical molecular switches. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6223–6232. [CrossRef] [PubMed]
59. Wergifosse, M.; Champagne, B. Electron correlation effects on the first hyperpolarizability of push–pull  $\pi$ -conjugated systems. *J. Chem. Phys.* **2011**, *134*, 074113. [CrossRef] [PubMed]
60. Lescos, L.; Sitkiewicz, S.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B.; Matito, E.; Castet, F. Performance of DFT functionals for calculating the second-order nonlinear optical properties of dipolar merocyanines. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16579–16594. [CrossRef] [PubMed]
61. Teunissen, J. *Inverse Molecular Design: Optimization and Application of Combinatorial and Stochastic Approaches*; Vrije Universiteit Brussel: Ixelles, Belgium, 2019.
62. Teunissen, J. CINDES. GitHub Repository. Available online: <https://gitlab.com/jlteunissen/CINDES> (accessed on 1 October 2023).
63. RDKit: Open-Source Cheminformatics. Software. Available online: <http://www.rdkit.org> (accessed on 7 September 2023).
64. Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754. [CrossRef]
65. Cihan Sorkun, M.; Mullaj, D.; Koelman, J.; Er, S. ChemPlot, a python library for chemical space visualization. *Chem. Methods* **2022**, *2*, e202200005. [CrossRef]
66. Van der Maaten, L.; Hinton, G. Visualizing high-dimensional data Using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
67. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.