

# Exploring the chemical space in search for improved Azoheteroarene-based photoswitches

*Sergi Vela, Constantin Krüger and Clémence Corminboeuf\**

*Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL),  
Laboratory for Computational Molecular Design, CH-1015 Lausanne, Switzerland*

## Abstract

In the quest for improved photo switches, azoheteroarenes have emerged as a potential alternative to azobenzene. However, to date the number and types of these species that have subjected to study is insufficient to provide an in-depth understanding of the photochemical effects brought about by different substituents. Here, we computationally screen the optical properties and thermal stabilities of 512 azoheteroarenes that consist of eight different N-containing heteroarenes combined with 64 substitution patterns. The most promising compounds are identified and their properties rationalized based on the nature of the azoheteroarene core and the location and type of substitution patterns.

## 1. Introduction

The design of molecular switches able to reversibly modify their physicochemical properties is a relevant topic in physical organic chemistry. Those that are triggered by light -so called photoswitches- are much appealing in that regard due to the significant advantages of light over other stimuli. Photoswitches are being exploited with great success in a number of applications encompassing molecular motors and actuators,<sup>1-3</sup> sensors,<sup>4</sup> and memory devices.<sup>5, 6</sup> While the ideal characteristics depend on its field of application, four general criteria have been proposed<sup>7</sup> to evaluate a photoswitch: (1) the wavelength of maximum light absorption and the associated quantum yield, (2) the relative thermal stability of the switch states, (3) the absence of degradation mechanisms (*i.e.*, fatigue) and (4) the dominant presence of either state in the photostationary distribution (PSD). Dyes based on the azo group are among of the mostly-studied family of photoswitches (together with diarylethenes and spiropyranes), not only because they fulfill the four requirements mentioned above, but also due to the ease of their synthesis,<sup>8</sup> the large number of functionalization sites, and the significant change in molecular shape upon photoswitch, which is highly appreciated in the context of molecular actuators. The most prototypical molecule of this family is Azobenzene, which can undergo a reversible photo-isomerization between its E- and Z-isomers.<sup>9-11</sup> Not surprisingly, a large variety of photoswitches has been designed and synthesized based on the azobenzene core.<sup>12, 13</sup> The quest for better azobenzene photoswitches prompted and inspired the investigation of azoheteroarenes. These are compounds in which one or both benzene rings of azobenzene is substituted by a heterocycle, which leads to changes in the structure and electronic properties, leading to interesting physicochemical properties.<sup>14-17</sup>

Despite the interest in this “new” class of compounds, the exploration of the chemical space has been rather limited, especially in comparison with azobenzene. Among five-membered azoheteroarenes, few examples of azopyrazoles,<sup>18, 19</sup> azoimidazoles,<sup>20-25</sup> and azopyrroles<sup>26</sup> have been reported. In these studies, it is quite common the comparison of unsubstituted azoheteroarenes and their methyl-substituted counterparts.<sup>18, 21, 25,</sup>

<sup>26</sup> This is because the bulkier methyl group modifies the conformation adopted when in the Z-isomer, leading

to reduced half-life times and a better PSD. Beyond methyl groups, the use of bulkier substituents has been explored by Herges and coworkers, both in isolated azoheteroarenes, and in the context of the LD-CISSS (*i.e.*, light-driven coordination-induced spin state switch) strategy.<sup>22</sup> In this very same work, the authors explored the use of substituents in the “para” position (with respect to the azo moiety) of the benzene group, to which they attached simple electron-donating and -withdrawing groups such as F, OMe and NMe<sub>2</sub>. The benzene substitution has been further explored by Ravoo and coworkers in azopyrazoles, in which 4-pyrazole had always two methyl groups in the two ortho positions.<sup>19</sup> Despite the range of known azoheteroarenes is continuously increasing, it is still unclear how -and how much- their photoswitching properties can be improved. Rather than blindly searching for promising azoheteroarenes, computational tools can quickly and reliably explore the chemical space.<sup>27-30</sup> Having this goal in mind, in this work we present a systematic computational screening of azoheteroarene photoswitches. Our library covers a total of 512 compounds, representing 64 different substitution patterns for eight N-containing heteroarenes. We discuss how the chemical functionalization modifies the thermal stability and the properties (*i.e.* energy, intensity, character) of the  $n\pi^*$  and  $\pi\pi^*$  bands of the E- and Z- isomers. Based on these properties, and on their connection with the criteria described in the first paragraph, we identify the most promising heteroarene families and substitution patterns.

## 2. Methodology and Computational Details

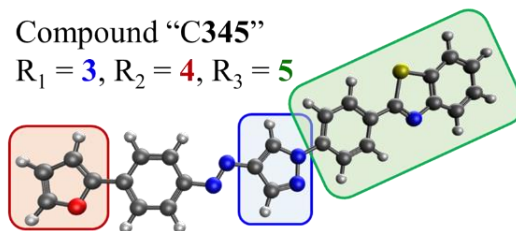
### 2.1. Fragment library

The azoheteroarenes database is built from the combination of three different fragments, labelled R<sub>1-3</sub> (see Scheme 1). The first is the heteroaryl ring, R<sub>1</sub>. Following the existing literature, the library of R<sub>1</sub> includes all regioisomers of pyrazole (3-pz, 4-pz and 5-pz), pyrrole (2-py and 3-py) and imidazole (2-imid and 4-imid and 5-imid). The second fragment (R<sub>2</sub>) is the benzene substituent. We have only considered substitution in position “para”, that

is, opposite to the Azo group, to maximize resonance effects while avoiding potential steric conflicts associated with the “ortho” substitution. The library of R<sub>2</sub> includes (i) the amine and methoxy groups, which are considered good donors *via* resonance, (ii) five-membered conjugated rings such as thiophene, furane, pyrazole and imidazole, and (iii) the six-membered ring thiazine, a good donor molecule that, in contrast to amine and methoxy, has a pi-conjugated system. Finally, the third fragment (R<sub>3</sub>) is the functionalization of the heteroaryl (*i.e.* R<sub>1</sub>). In four R<sub>1</sub> families, R<sub>3</sub> lies adjacent to the azo group (*i.e.* “ortho”, *o*-R<sub>1</sub>), and in four more it is in “meta” position (*m*-R<sub>1</sub>). The library of R<sub>3</sub> includes (i) a methyl group, in line with the array of azoheteroarenes recently reported,<sup>19, 26</sup> (ii) the five-membered rings thiophene, furane and imidazole, and (iii) the pi-conjugated electron acceptors Tz (2-phenyl-benzothiazole), BP (diphenyl-methanone) and DPO (2,5-diphenyl-1,3,4-oxadiazole). For completeness, H atoms in R<sub>2</sub> and R<sub>3</sub> are also considered, which corresponds to having a non-substituted benzene (case of R<sub>2</sub>=H) or a non-substituted heteroarene (case of R<sub>3</sub>=H). Compounds in which both R<sub>2</sub> and R<sub>3</sub> are H atoms (R<sub>2</sub>=R<sub>3</sub>=0) are called “bare” azoheteroarenes along the text. The design choice of having donors in R<sub>2</sub> and acceptors in R<sub>3</sub> stems for the potential to combine them generating *push-pull* azoheteroarenes. Being natural electron-acceptors themselves, the heteroarenes are enhanced when further substituted with an acceptor. Concerning the five-membered rings that are not strong donors/acceptors, they have been included in our library for their ubiquity in chemistry

Compound “C345”

R<sub>1</sub> = 3, R<sub>2</sub> = 4, R<sub>3</sub> = 5



Scheme 1.

and in biological systems, the possibility to further substitute these groups thus extending the  $\pi$ -conjugation, their reduced steric effects with respect to six-membered rings, and for their potential use as coordination points to a metal (cases of pyrazole and imidazole). All 512 possible  $R_{1-3}$  fragment combinations are built (see Table 1). They are labelled using the string “C +  $R_1$  +  $R_2$  +  $R_3$ ” in a way that, for instance, **C345** refers to the combination of  $R_1=3$ ,  $R_2=4$  and  $R_3=5$  (see Scheme 1). Some of the 512 compounds have already been reported in the literature. A comparison of previous experimental or computational results with the ones offered in this manuscript is in section S6.

**Table 1.** List and code of the fragments  $R_1$  (left),  $R_2$  (middle) and  $R_3$  (right) included in our library. Those fragments are combined following Scheme 1 to obtain the 512 compounds studied in this work.

$R_1$	$R_2$	$R_3$
<b>1 = 3-pz</b> 	<b>0 = H</b>	<b>0 = H</b>
<b>2 = 3-py</b> 	<b>1 = Amine</b> 	<b>1 = Me</b> 
<b>3 = 4-pz</b> 	<b>2 = Methoxy</b> 	<b>2 = Thiophene</b> 
<b>4 = 4-im</b> 	<b>3 = Thiophene</b> 	<b>3 = Furane</b> 
<b>5 = 2-py</b> 	<b>4 = Furane</b> 	<b>4 = 4-im</b> 
<b>6 = 5-pz</b> 	<b>5 = 4-im</b> 	<b>5 = Tz<sup>a</sup></b> 
<b>7 = 5-im</b> 	<b>6 = 5-pz</b> 	<b>6 = BP<sup>b</sup></b> 
<b>8 = 2-im</b> 	<b>7 = Thiazine</b> 	<b>7 = DPO<sup>c</sup></b> 

<sup>a</sup>Tz=2-phenyl-benzothiazole, <sup>b</sup>BP=diphenyl-methanone, <sup>c</sup>DPO=2,5-diphenyl-1,3,4-Oxadiazole

## 2.2. Computational Details

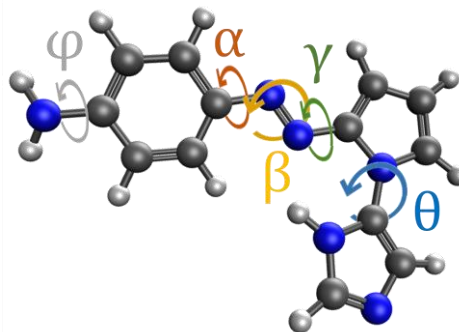
The  $R_{1-3}$  fragments defined above are initially built and handled in the SMILES (Simplified Molecular-Input Line-Entry System) notation, and put together using local scripts. The conversion from SMILES to the 3D coordinates is then performed with OpenBabel version 2.4.1<sup>31</sup> using the MMFF94 force field.<sup>32</sup> Further geometry optimizations were performed at the PBE0-D3BJ/def2-TZVP level using Gaussian 09 (G09)<sup>33</sup> until convergence, followed by a frequency calculation to ensure that all structures belong to a

minimum energy structure. Linear-Response TDDFT computations<sup>34, 35</sup> were then performed at these structures at the  $\omega$ B97-XD/TZVP level, for the reasons discussed in the computational benchmark (see Section S1). Ten singlet states were systematically computed. A comparison between experimental measurements and our computations for a limited set of compounds suggests a deviation of about 20 nm (*i.e.* red-shift) in the estimation of the  $n\pi^*$  band, and of  $-30$  nm for the  $\pi\pi^*$  band (*i.e.* blue-shift) of the E-isomer with respect to experiment (see Figure S5 and Table S8).

The identification and analysis of the electronic transitions was carried out using the charge-transfer matrices (Excitation Fingerprints (EF)). The EF represent the localization of the electron/hole of a given electronic transition ( $l$ ), using the charge transfer numbers between two molecular fragments (A and B) ( $\Omega_{AB}^l$ ). In the present study, the molecules were split into five components following Scheme 1 from left to right: R<sub>2</sub>, Ph, Azo, R<sub>1</sub>, R<sub>3</sub>. As a result, the EF have the form of 5x5 matrices of  $\Omega_{AB}^l$  values, numbered from bottom-left to top-right. The diagonal terms correspond to the energy-transfer components (*i.e.* hole and electron in the same fragment), while the off-diagonal terms are the charge-transfer components of the electronic excitation. To identify the relevant  $\pi\pi^*$  transition, we have used the total electron transfer towards the azo fragment (from any fragment,  $\Omega_{X3}^l$ ) (see Section 4.C). The  $\Omega_{AB}^l$  values were obtained from the G09 outputs using cclib<sup>36</sup> (for parsing) and the analysis of the transition density matrix<sup>37</sup> provided in TheoDORE version 1.7.1,<sup>38</sup> and subsequently plotted using a modified version of the python scripts included therein. See ref. 39 for further information on  $\Omega_{AB}^l$ . The Wiberg indices have been computed on the optimized structures using the Natural Bond Order analysis (version 3.1) implemented in G09.

### 2.3. Geometry and Conformers

In the upcoming section, the minima are often characterized using the torsion angles as defined in Scheme 2. Beyond the Z/E isomerization, characterized by  $\beta$ , rotation about the angles leads to several conformers, whose optical properties might differ. Generating and analyzing all possible conformers is beyond the scope of this work. Instead, the same conformers of  $\alpha$  and  $\gamma$  were systematically generated and optimized without constrain on  $\theta$  and  $\varphi$ . *o*-R<sub>1</sub> (R<sub>1</sub>=**4-8**) with 4-imidazole as R<sub>3</sub> (R<sub>3</sub>=**4**) were the only systems for which rotation around one of the angle ( $\theta$ ) had an effect on the properties (*vide infra*, Section S2.A).



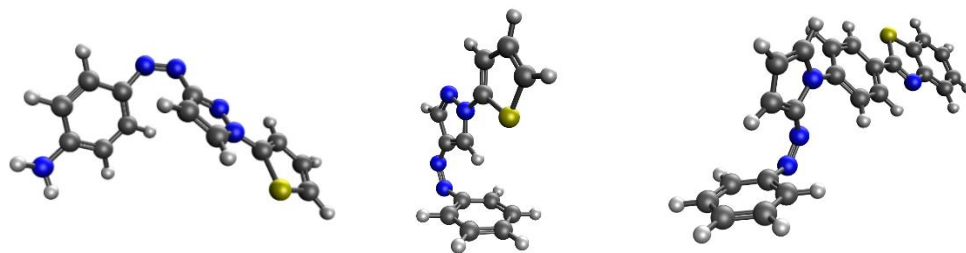
Scheme 2.

## 3. Results and discussion

### 3.1. Structure and Thermal stability

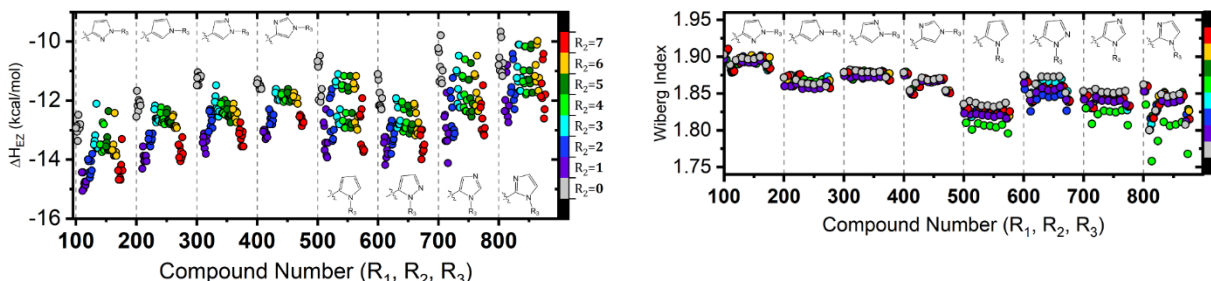
**Structure.** Prior to analyzing the thermal stability and optical properties of the compounds, the key structural features need to be mentioned. A general observation is that for the E-configuration, the *m*-R<sub>1</sub> derivatives behave slightly differently from the *o*-R<sub>1</sub>. The former compounds are mostly planar with no major steric constraints. Their optical properties are essentially modulated by the electronic structure of the R<sub>1</sub>-R<sub>3</sub> substituents. In contrast, larger geometrical variations are observed for the *o*-R<sub>1</sub> compounds, which are more sensitive to the combination of R<sub>1-3</sub> fragments (see Table S1). The Z-isomers exhibit significantly-different torsion angles depending on the compound (see Table S2). Overall, all geometries can be grouped into three

categories: most compounds fall into the so-called “*folded*” conformation (see Figures 1 and S13), a minority of compounds (some with  $R_1=2$  (3-py), and most  $R_1=3$  (4-pz) derivatives) displays the “*T-shape*” conformation. The “*twisted*” form is only favored occasionally.



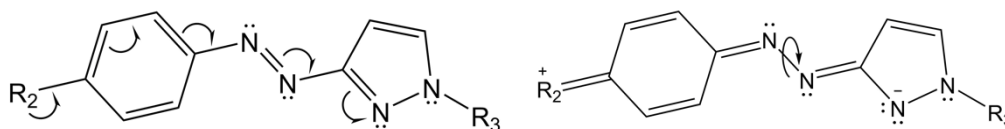
**Figure 1.** Z-isomer of compounds (left) C112, (middle) C302 and (right) C505, representative of “*folded*”, *T-shape*” and “*twisted*” structures, respectively.

**Thermal Stability.** In azobenzene derivatives, the Z-isomer is generally less stable than the E-form. Depending on the energy difference, and on the switching mechanism (inversion *vs.* rotation), the Z-isomer might undergo thermal- (instead of photo-) isomerization, hence reducing its half-life time ( $t_{1/2}$ ). A good estimation of  $t_{1/2}$  can be extracted from (i) the relative energy difference between the E- and Z-isomers minima ( $\Delta H_{EZ}$ ), and (ii) the strength of the azo N=N bond in the Z-isomers, quantified using the Wiberg indices (WI) (see Figure 2 and Section S2.C).<sup>26</sup> 3-pz derivatives ( $R_1=1$ ) show the most-stable E-isomers, while 2-im ( $R_1=8$ ) leads to the least-unstable Z-isomers. “*Bare*” heteroarenes decrease the relative stability of the E-isomer (compare grey *vs.* colored circles), while good donors ( $R_2=1, 2$  and  $7$ ) lead to the opposite. These substituents favor *push-pull resonance* (see Figure 3), which stabilizes the planar E-configurations. The  $R_3$  substitution has a smaller impact on  $\Delta H_{EZ}$ , which is mostly due to minor geometrical changes.



**Figure 2.** (left) Thermal stability between the E- and Z- isomers. Negative  $\Delta H_{EZ}$  values indicate that the E- isomer is more stable. (right) Wiberg indices of the azo N=N bond of the computed Z-isomers. The color code indicates  $R_2$ .

Alternatively, larger WI values are obtained for the *m*- than *o*- $R_1$  derivatives (see Figure 2), in agreement with the trend in  $t_{1/2}$  observed in ref. 26. The reason is the partial *vs.* complete  $\pi$ -conjugation in the *m*- $R_1$  *vs.* *o*- $R_1$  compounds (see Figure 3). The influence of  $R_2$  and  $R_3$  on the WI is larger for *o*- $R_1$  (see also Figure S14 and Table S4) and systematically leads to a smaller WI with respect to the *bare* azoheteroarenes, the most striking case being the *o*- $R_1$  compounds (except  $R_1=6$ , see Section S2.A) with 4-im in  $R_3$  ( $R_3=4$ ) (green circles in Figure 2). These compounds generate a short contact between the lone-pair of an azo N atom (contributing to the  $n$  orbital) and the N-H group of  $R_3$ . This H-bond promotes a tautomeric form in which the azo-group is protonated and the N atoms are connected through a single bond. Finally, Z-isomers with a *T-shape* conformation possess larger WI (see Figure S13), in line with their larger experimental half-lives.<sup>26</sup>

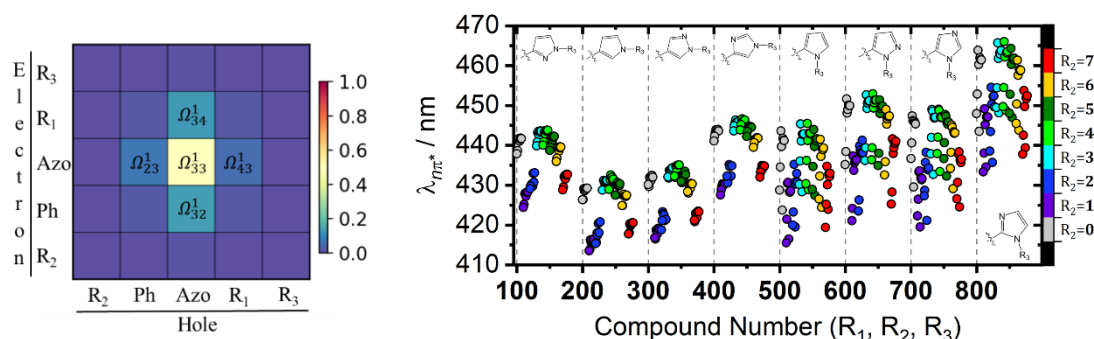


**Figure 3.** Push-pull resonance in azoheteroarenes, which allows free rotation in the azo bond. The case of  $R_1=3\text{-py}$  is taken as an example. The negative charge in  $R_1$  is partially delocalized in  $m\text{-}R_1$  and completely delocalized in  $o\text{-}R_1$ .

### 3.2. Photochemistry of the *E*- and *Z*-isomers

Excitation of *E*-Azobenzene from  $S_0$  to  $S_1$  leads to a weakly-allowed  $n\pi^*$  transition in the visible region ( $\lambda \approx 450$  nm), whose bright component stems from vibrationally-accessible non-planar structures that break the center of inversion.<sup>40</sup> Excitation from  $S_0$  to  $S_2$  leads to an intense symmetry-allowed  $\pi\pi^*$  transition in the UV range ( $\lambda \approx 320$  nm). Once excited, Azobenzene undergoes photo-isomerization of the  $N=N$  bond from either  $S_1$  or  $S_2$  through either inversion-assisted rotation about the azo group,<sup>41, 42</sup> or by in-plane inversion, when rotation is sterically hindered.<sup>43</sup> The isomerization mechanism of azoheteroarenes is expected to be very similar, including the competition between the rotation and inversion mechanism.<sup>44, 45</sup> Therefore, our discussion focuses on the  $S_0$ -to- $S_1$  ( $n\pi^*$ ) and  $S_0$ -to- $S_2$  ( $\pi\pi^*$ ) electronic transitions.

***E*-isomer. First Excited State ( $n\pi^*$ ).** The  $n\pi^*$  transition is a combination of an energy transfer within the azo group ( $\Omega_{33}^1$ ), and charge-transfer components from the azo to either the Ph ring ( $\Omega_{32}^1$ ) or  $R_1$  ( $\Omega_{34}^1$ ) (see EF in Figure 4). Complementary representations of the excitation energies are provided in Figures 4 and S15. The most striking observation in Figure 4 is that compounds with 2-imidazole ( $R_1=8$ ) display, by far, the most red-shifted  $n\pi^*$  transition and this for two reasons. These derivatives, as well as all ortho-substituted  $R_1$  ( $o\text{-}R_1$ ) compounds, are fully conjugated and planar, in contrast to the partially conjugated meta-substituted species ( $m\text{-}R_1$ ) (see Figure 3).<sup>26</sup> Conjugation stabilizes the electron density that is transferred to  $R_1$ , thus red-shifting the transition. The second reason is related to the large number of N atoms in the heteroarene ring and their location. When comparing  $R_1$  derivatives with the same number of N atoms, those in which the N atoms are closer to the azo group are more red-shifted. With two N atoms next to the azo group, compounds with  $R_1=8$  display the most red-shifted transition.



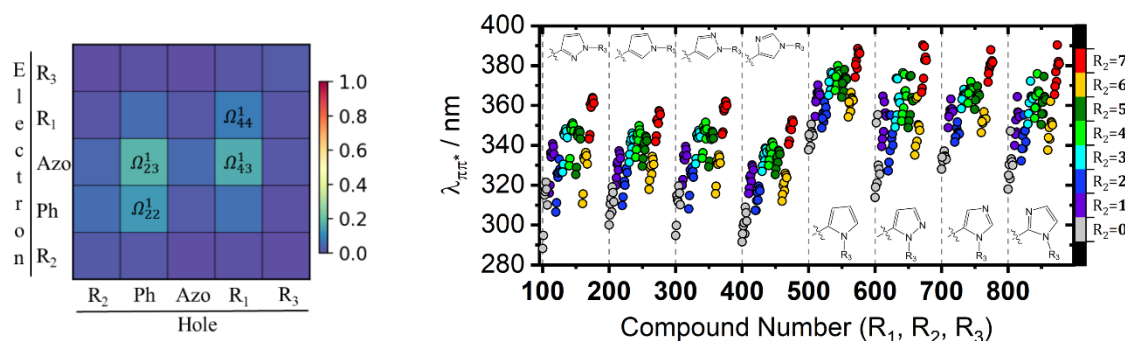
**Figure 4.**  $n\pi^*$  transition in the *E*-isomers. (left) Average Excitation Fingerprint (EF) representing the character of the transition. (right) Excitation wavelength for each compound, ordered using its three-digit identifier. To identify the patterns discussed in the text, vertical dashed lines separate groups with the same  $R_1$ , and the color code indicates the  $R_2$  index. Transition probabilities (*i.e.* osc. strength,  $f$ ) are shown in Figure S15.

A second pattern arises from the effect of the chemical substitution of the Ph and  $R_1$  groups (*i.e.*,  $R_2$  and  $R_3$ , respectively). These fragments modulate the nature of the transition, either by increasing the amount of



charge transfer to the Ph ring ( $\Omega_{32}^1$ ) or to R<sub>1</sub> ( $\Omega_{34}^1$ ), or by reducing the back-transfer to the azo group ( $\Omega_{23}^1$  and  $\Omega_{43}^1$ ). In general, the recipe for a red-shifted transition is clear: the portion of electron density brought away from the azo group must increase. This is mostly achieved by avoiding electron-donors as R<sub>2</sub> (R<sub>2</sub>=**1**, **2** and **7**), since they decrease  $\Omega_{32}^1$  and increase  $\Omega_{23}^1$ , thus blue-shifting the transition (see Figure S16). The influence of R<sub>3</sub> acceptors is much smaller. In other words, the *push* effect is more important than the *pull*. This is especially evident in *m*-R<sub>1</sub> compounds where electronic-structure effects dominate. Finally, the set of compounds that have an H-bond between R<sub>3</sub> and the azo (see Sections 3.1 and S2.A), present  $n\pi^*$  transitions that are systematically higher in energy, and slightly bright (osc. strength,  $f \approx 0.1$ , see Figure S15).

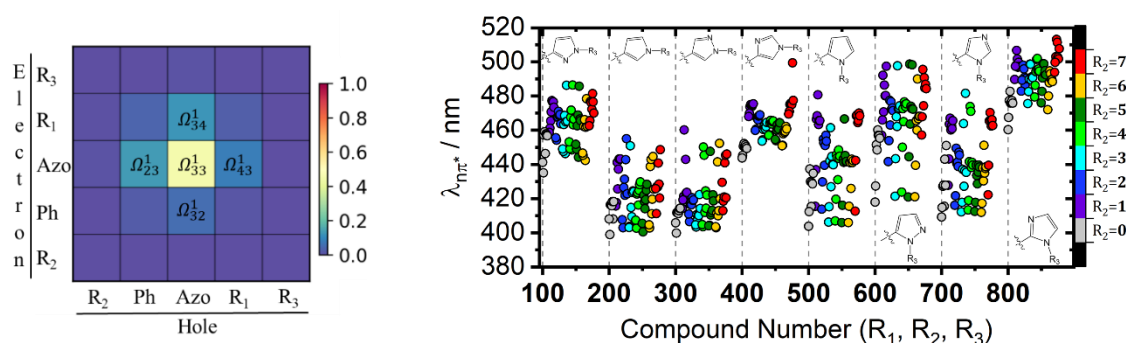
**E-isomer. Second Excited State ( $\pi\pi^*$ ).** The  $\pi\pi^*$  transition is, in most cases, the S<sub>0</sub>-S<sub>2</sub> transition. The EF shows dominant charge-transfer components from the Ph and R<sub>1</sub> groups to the azo ( $\Omega_{23}^2$  and  $\Omega_{43}^2$ ), together with some energy transfer within the Ph and R<sub>1</sub> groups ( $\Omega_{22}^2$  and  $\Omega_{44}^2$ ) (see Figure 5). Compounds **C107**, **C207**, **C307**, **C407**, **C427** and **C467** are an exception. In these six compounds, S<sub>2</sub> corresponds to a  $\pi\pi^*$  excitation centered in R<sub>3</sub>=**7** (*i.e.* DPO). Given that this transition is likely to be *non-productive* (*i.e.* not leading to isomerization), the relevant  $\pi\pi^*$  transition (in this case S<sub>3</sub>) is discussed instead (*i.e.* the one displaying the EF in Figure 3). It remains to be determined whether such localized transition in R<sub>3</sub> could quench the photo-isomerization. As discussed below, a similar scenario -albeit more complex- is found for the Z-isomer.



**Figure 5.**  $\pi\pi^*$  transition in the E-isomers. (left) Average Excitation Fingerprint (EF) representing the character of the transition. (right) Excitation wavelength for each compound, ordered using its three-digit identifier. To identify the patterns discussed in the text, vertical dashed lines separate groups with the same R<sub>1</sub>, and the color code indicates the R<sub>2</sub> index. Transition probabilities (*i.e.* osc. strength,  $f$ ) are shown in Figure S15.

Two clear patterns are identified (see Figures 5 and S15). The first is identical to the previous  $n\pi^*$  transition namely that the  $\pi\pi^*$  transition in the *o*-R<sub>1</sub> derivatives are systemically red-shifted (and less bright) than the *m*-R<sub>1</sub>. The second is associated with the effect of R<sub>2</sub> and is apparent in each R<sub>1</sub>=**1-8** subset. The  $\pi\pi^*$  excitation energy depends on the amount of electron density that is transferred into the azo group, and where it comes from. The Ph substituent (*i.e.* R<sub>2</sub>) modulates the Ph-to-azo component ( $\Omega_{23}^2$ ) in a similar way as in the  $n\pi^*$  transition. For the electron-donors R<sub>2</sub>=**1** (NH<sub>2</sub>) and R<sub>2</sub>=**2** (OMe),  $\Omega_{23}^2$  increases, and the transition is red-shifted with respect to the bare Ph (*i.e.* with R<sub>2</sub>=**0**). For  $\pi$ -conjugated substituents (R<sub>2</sub>=**3-7**), the  $\pi$ -orbitals involved in the transition is delocalized across the entire  $\pi$ -system and, as a result, a portion of the electron density that is transferred to the azo group comes, from R<sub>2</sub> ( $\Omega_{13}^2$ ) and not from the Ph ring ( $\Omega_{23}^2$ ). Given that more delocalized transition are associated with lower energies,  $\pi$ -conjugated R<sub>2</sub> red-shift the transitions with thiazine, being also a good donor, being the best among them.

**Z-isomer. First Excited State ( $n\pi^*$ ).** Similar to the E-configurations, the  $S_0$ - $S_1$  transition of  $n\pi^*$  character is mainly localized in the azo group (see Figures 6 and S17). In general, the key aspects discussed above (for the E-isomer) remain valid for the Z-isomer. For instance, the same  $R_1$ -derivatives ( $R_1=8$ ) display the most red-shifted transition. There is however a difference within each  $R_1$  family. In the E-state, the extent of density leaving the azo group was modulated by the electron donating/withdrawing ability of  $R_2$  and  $R_3$ . In the Z-isomer, such electronic effect is not observed, the key factor being the structure of the azoheteroarene core (Ph-azo- $R_1$ ). There exists a clear correlation between the value of  $\alpha$  and  $\gamma$ , and the excitation energy: red-shifted transitions coincide with these angles lying at around  $45^\circ$  (see Figure S18). This angle favors the mixing of  $\sigma$ -molecular orbitals in the azo group (including  $n$ ), with those of  $\pi$ -character in Ph and  $R_1$ , which facilitates the delocalization of the hole density over these neighboring rings, diminishing  $\Omega_{33}^1$  (see Figure S19). This increased  $\pi$ -character is also associated to a larger transition intensity. It is therefore not by chance that the most red-shifted  $n\pi^*$  transitions are the brightest ones (see Figure S17). This analysis is in agreement with the relationship between intensity and  $\alpha$  angle found by Fuchter and coworkers (see Figure 10 of ref. 26).



**Figure 6.**  $n\pi^*$  transition in the Z-isomers. (left) Average Excitation Fingerprint (EF) representing the character of the transition. (right) Excitation wavelength for each compound, ordered using its three-digit identifier. To identify the patterns discussed in the text, vertical dashed lines separate groups with the same  $R_1$ , and the color code indicates the  $R_2$  index. Transition probabilities (*i.e.* osc. strength,  $f$ ) are shown in Figure S17.

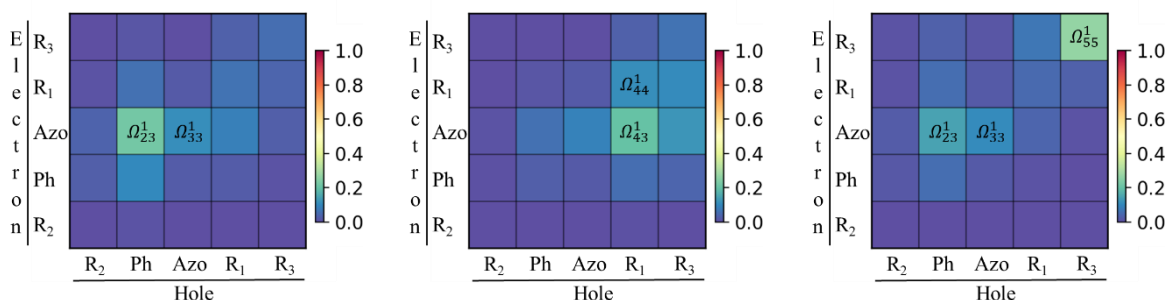
**Z-isomer. Second Excited State ( $\pi\pi^*$ ).** Formally, this transition refers to the lowest-lying bright  $\pi\pi^*$  transition ( $S_2$ - $S_5$  for this series of compounds) that might lead to isomerization. The *productive* transition was identified as the one with larger  $\Omega_{AB}^l$ -weighted intensity (referred to as  $f'$ ) (see Section S4.C). Its nature differs significantly depending on the compound, with up to three main types of EFs identified (see Figure 7). The most common one (35% of the cases) is a  $\pi\pi^*$  transition dominated by an electron transfer from the Ph group to the azo, and some energy transfer in the azo group (see Figure 7a). Notice that this type of EF resembles the left-side of the EF associated with the  $\pi\pi^*$  transition in the E-isomer (see Figure 3). The right-side counterpart is, indeed, the second type (33%) for which the electron transfer mainly from  $R_1$  to the azo with some intra-fragment contributions in  $R_3$  and the azo itself (see Figure 7b). Finally, the third type of transition (23% of cases) shows a very-large intra-fragment component within  $R_3$  and a smaller contribution of the azo (see Figure 7c). The appearance of each type of transition depending on  $R_{1-3}$  is analyzed in Section S4.B, where it is also demonstrated that the three types of transition are not an computational artefact (see Figure S22).

(a) C112

(b) C103

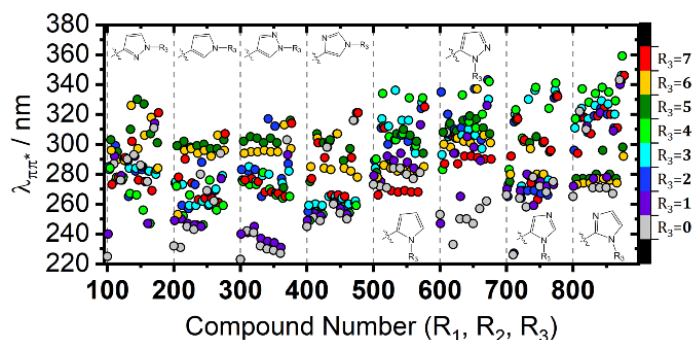
(c) C115





**Figure 7.** The three different patterns associated with the  $\pi\pi^*$  transition in the Z-isomers, exemplified by compounds **C112**, **C103** and **C115**.

The excitation energies are very similar for all heteroarene families (see Figures 8 and S17). The *electronic* effects induced by  $R_2$  and  $R_3$ , which we discussed for the E-isomer, re-appear clearly in compounds with *m*- $R_1$ . Within this group, red-shifted transitions are obtained with  $\pi$ -conjugated acceptors Tz ( $R_3=5$ ) and BP ( $R_3=6$ ), while blue-shifted transitions are obtained with the non- $\pi$ -conjugated and electronically-innocent fragments  $R_3=0$  and **1** (see Figure S17). These trends are masked -or even cancelled- in *o*- $R_1$  given the larger structural differences within this group. In some cases, structure seems to be the key, since the most red-shifted transitions are generally obtained for  $R_3=3$  and **4**, which are the fragments that lead to a more planar  $R_1$ - $R_3$  torsion angle ( $\theta$ ). In other cases, the donor/acceptor character of the substituents also appears as relevant, but not systematically for all compounds. Unfortunately, the fact that we found different EF for this transition makes much more complicate the analysis of the  $\Omega_{AB}^l$  values, and hence the identification of patterns.

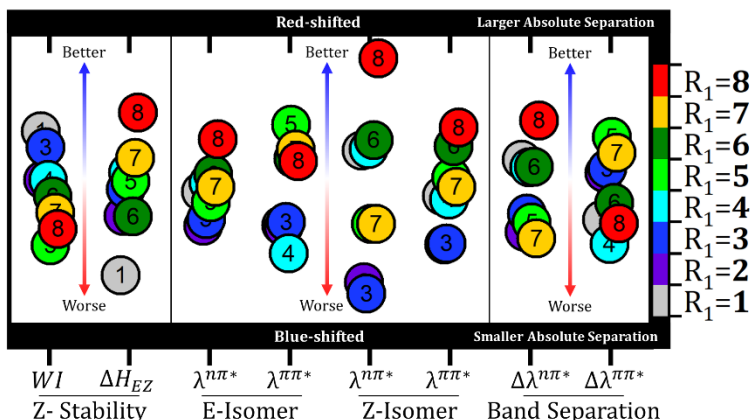


**Figure 8.** Excitation wavelength associated with the  $\pi\pi^*$  transition in the Z-isomer of the studied compounds, which are ordered using its three-digit identifier.

## 4. Summary

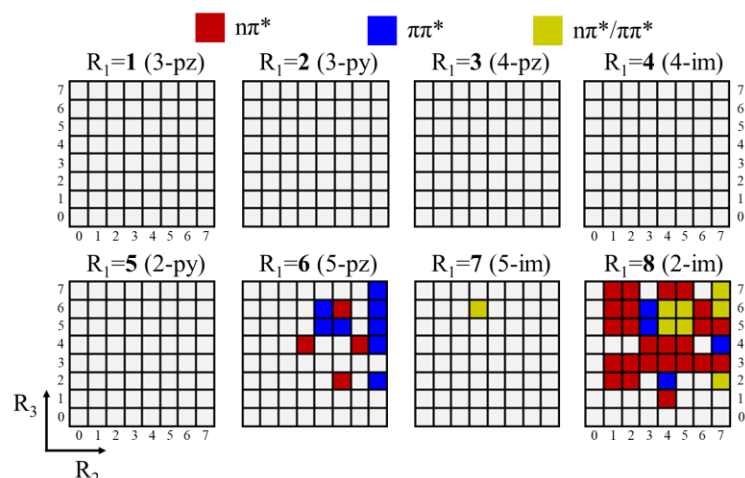
The introduction mentioned the four criteria that define a good photoswitch. We discussed how the first criterion can be evaluated using  $\Delta H_{EZ}$  and the WI, and the second with the  $n\pi^*$  and  $\pi\pi^*$  excitation wavelengths ( $\lambda^{n\pi^*}$  and  $\lambda^{\pi\pi^*}$ ). An assessment of the third criterion requires the description of competing non-adiabatic processes through molecular dynamics, which falls beyond the scope of the current work. The fourth criterion refers to the photostationary distribution (PSD). The PSD depends on the ratio between the (i) E-to-Z and Z-to-E isomerization quantum yields at a given irradiation wavelength, and (ii) the overlap of the bands that trigger such processes. When the quantum yields are not available, the PSD can be assessed through the band overlap. When it is small, the PSD has a larger probability of being close to 0 or 1,

indicating a complete isomerization (either forwards or backwards) and the possibility to address the photoswitch in a selective manner. Herein, the overlap is evaluated through the separation of the  $\pi\pi^*$  and  $\pi\pi^*$  bands in each isomer, together with the separation between the  $n\pi^*$  of the Z-isomer and the  $\pi\pi^*$  of the E-isomer (see Section S5).



**Figure 9.** Performance comparison for each azoheteroarene family according to  $R_1$ , which is indicated using both the color code, and the  $R_1$  number inside circles. For clarity, all values are given as weighted deviations from the average. See raw data in Tables S9-12.

Once the thermal stabilities, excitation wavelengths and band separations are known, it is possible to *rank* the (i) azoheteroarene families (*i.e.*  $R_1$ ), and (ii) the individual compounds according to these descriptors (average values of the relevant descriptors are compared in Figure 9). 2-im ( $R_1=8$ ) presents the most red-shifted  $n\pi^*$  transition (*i.e.* largest average  $\lambda^{n\pi^*}$ ) for both the E- and Z- isomers, and also the largest average  $\pi\pi^*$  band separation (38 nm). Thus, compounds based on 2-im ( $R_1=8$ ) are a promising family of azoheteroarenes when it comes to their  $\pi\pi^*$  transitions, although they might display shorter  $t_{1/2}$  according to the computed WI. The largest average  $\lambda^{\pi\pi^*}$  corresponds to 2-im ( $R_1=8$ ) for the Z-isomer, and to 2-py ( $R_1=5$ ) for the E-isomer. The latter family is also the one with largest  $\pi\pi^*$  band separation, but its average  $\lambda^{\pi\pi^*}$  is only moderately good. A second ranking representation highlighting the individual compounds for which  $n\pi^*$  (red),  $\pi\pi^*$  (blue), or both (gold) transitions are particularly well-suited can be found in Figure 10 (see Section S7.B). Once again, the superiority of 2-im ( $R_1=8$ ) over all other heteroarene families is clear, as it gathers most of the  $n\pi^*$  and  $\pi\pi^*$  top-performers.  $R_1=6$  is a relevant alternative, with a dozen candidates. With respect to the  $R_2$  and  $R_3$  substitution, bare azoheteroarenes ( $R_2=R_3=0$ ) underperform in comparison to the substituted ones (not a single compound with  $R_2$  or  $R_3=0$  is highlighted).  $R_2$  has little influence on the  $n\pi^*$  band, but the  $\pi$ -conjugation and electron-donor ability of  $R_2=3-7$  improves the  $\pi\pi^*$  transition. Finally, having electron-acceptors groups in  $R_3$  improves the  $\pi\pi^*$  transition (*blue* and *gold* top-performers are clustered in the top).



**Figure 10.** Performance comparison of compounds based on the criterion described in the main text. Top  $n\pi^*$  ( $\pi\pi^*$ ) performers are highlighted in red (blue). Gold is used when a compound excels in both transitions.

## 5. Conclusions

In search for improved azoheteroarene photoswitches, this computational work analyzed the photochemical properties of 512 substituted ( $R_2$  and  $R_3$ ) azoheteroarenes ( $R_1$ ) based on pyrazole (3-pz, 4-pz and 5-pz), pyrrole (2-py and 3-py) and imidazole (3-im, 4-im and 5-im). Emphasis was placed on identifying derivatives with lower  $n\pi^*$  and  $\pi\pi^*$  energy transitions, larger band separations and extended half-life times. A related objective was to provide key guidelines determining of how much these properties can be tuned through chemical substitution.

While azoheteroarenes based on 2-im ( $R_1=8$ ) are found to be promising photoswitches working at both low- and high- excitation wavelengths through their  $n\pi^*$  and  $\pi\pi^*$  bands, they exhibit one of the shortest half-life times among all the heteroarene derivatives explored. Alternatives based on 5-pz ( $R_1=6$ ) are also appealing. Overall, the use of donor and acceptor substituents in  $R_2$  and  $R_3$  was especially relevant to modulate the  $\pi\pi^*$  transition. Some compounds feature a low-energy  $\pi\pi^*$  transition localized on the ( $R_3$ ) substituent whose effect on the isomerization is not clear. Further work is needed to clarify this point.

## Acknowledgements

The authors are grateful to the EPFL for the allocation of computer time. S.V. acknowledges funding from the *European Union's H2020 research and innovation programme* (grant agreement No 794519) and assistance from Dr. B. Meyer and R. Fabregat in setting the analysis tools.

## Dataset

The dataset will be available upon publication at DOI:10.5281/zenodo.3271200

## Funding sources

MSCA-IF-2017 (#794519)

## Conflicts of interest

There are no conflicts to declare

## References

1. Y. Yu, M. Nakano and T. Ikeda, *Nature*, 2003, **425**, 145-145.
2. T. Muraoka, K. Kinbara and T. Aida, *Nature*, 2006, **440**, 512-515.
3. M. Yamada, M. Kondo, J.-i. Mamiya, Y. Yu, M. Kinoshita, C. J. Barrett and T. Ikeda, *Angew. Chem. Int. Ed.*, 2008, **47**, 4986-4988.
4. M. Natali and S. Giordani, *Chem. Soc. Rev.*, 2012, **41**, 4010-4029.
5. Z. F. Liu, K. Hashimoto and A. Fujishima, *Nature*, 1990, **347**, 658-660.
6. S. Hvilsted, C. Sánchez and R. Alcalá, *J. Mater. Chem.*, 2009, **19**, 6641-6648.
7. S. Crespi, N. A. Simeth and B. König, *Nature Reviews Chemistry*, 2019, **3**, 133-146.
8. E. Merino, *Chem. Soc. Rev.*, 2011, **40**, 3835-3853.
9. H. M. D. Bandara and S. C. Burdette, *Chem. Soc. Rev.*, 2012, **41**, 1809-1825.
10. A. A. Beharry, O. Sadovski and G. A. Woolley, *J. Am. Chem. Soc.*, 2011, **133**, 19684-19687.
11. D. Bléger and S. Hecht, *Angew. Chem. Int. Ed.*, 2015, **54**, 11338-11349.
12. D. Bléger, J. Schwarz, A. M. Brouwer and S. Hecht, *J. Am. Chem. Soc.*, 2012, **134**, 20597-20600.
13. A. A. Beharry and G. A. Woolley, *Chem. Soc. Rev.*, 2011, **40**, 4422-4437.
14. S. Crespi, N. A. Simeth and B. König, *Nature Reviews Chemistry*, 2019, DOI: 10.1038/s41570-019-0074-6.
15. E. V. Brown and G. R. Granneman, *J. Am. Chem. Soc.*, 1975, **97**, 621-627.
16. N. A. Simeth, S. Crespi, M. Fagnoni and B. König, *J. Am. Chem. Soc.*, 2018, **140**, 2940-2946.
17. R. Travieso-Puente, S. Budzak, J. Chen, P. Stacko, J. T. B. H. Jastrzebski, D. Jacquemin and E. Otten, *J. Am. Chem. Soc.*, 2017, **139**, 3328-3331.
18. C. E. Weston, R. D. Richardson, P. R. Haycock, A. J. P. White and M. J. Fuchter, *J. Am. Chem. Soc.*, 2014, **136**, 11878-11881.
19. L. Stricker, M. Böckmann, T. M. Kirse, N. L. Doltsinis and B. J. Ravoo, *Chem. Eur. J.*, 2018, **0**.
20. M. Endo, K. Nakayama, Y. Kaida and T. Majima, *Tetrahedron Lett.*, 2003, **44**, 6903-6906.
21. J. Otsuki, K. Suwa, K. Narutaki, C. Sinha, I. Yoshikawa and K. Araki, *J. Phys. Chem. A*, 2005, **109**, 8064-8069.
22. C. Schütt, G. Heitmann, T. Wendler, B. Krahwinkel and R. Herges, *J. Org. Chem.*, 2016, **81**, 1206-1215.
23. T. Wendler, C. Schütt, C. Näther and R. Herges, *J. Org. Chem.*, 2012, **77**, 3284-3287.
24. N. Fukuda, J. Y. Kim, T. Fukuda, H. Ushijima and K. Tamada, *Japanese Journal of Applied Physics*, 2006, **45**, 460-464.
25. J. Otsuki, K. Suwa, K. K. Sarker and C. Sinha, *J. Phys. Chem. A*, 2007, **111**, 1403-1409.
26. J. Calbo, C. E. Weston, A. J. P. White, H. S. Rzepa, J. Contreras-García and M. J. Fuchter, *J. Am. Chem. Soc.*, 2017, **139**, 1261-1274.
27. K. D. Collins, T. Gensch and F. Glorius, *Nature Chem.*, 2014, **6**, 859.
28. B. Meyer, B. Sawatlon, S. Heinen, O. A. von Lilienfeld and C. Corminboeuf, *Chemical Science*, 2018, **9**, 7069-7077.
29. K.-H. Lin, A. Prlj and C. Corminboeuf, *J. Phys. Chem. C*, 2017, **121**, 21729-21739.
30. Y. Bai, L. Wilbraham, B. J. Slater, M. A. Zwijnenburg, R. S. Sprick and A. I. Cooper, *J. Am. Chem. Soc.*, 2019, DOI: 10.1021/jacs.9b03591.
31. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *Journal of Cheminformatics*, 2011, **3**, 33.
32. T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 490-519.
33. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, 2009.
34. E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, 1984, **52**, 997-1000.
35. M. E. Casida, in *Recent Advances in Density Functional Methods*, DOI: 10.1142/9789812830586\_0005, pp. 155-192.

- 36.N. M. O'boyle, A. L. Tenderholt and K. M. Langner, *J. Comput. Chem.*, 2008, **29**, 839-845.
- 37.F. Plasser and H. Lischka, *J. Chem. Theory Comput.*, 2012, **8**, 2777-2789.
- 38.F. Plasser, *Journal*, 2017.
- 39.S. Mai, F. Plasser, J. Dorn, M. Fumanal, C. Daniel and L. González, *Coord. Chem. Rev.*, 2018, **361**, 74-97.
- 40.T. Cusati, G. Granucci, M. Persico and G. Spighi, *J. Chem. Phys.*, 2008, **128**, 194312.
- 41.E. M. M. Tan, S. Amirjalayer, S. Smolarek, A. Vdovin, F. Zerbetto and W. J. Buma, *Nature Communications*, 2015, **6**, 5860.
- 42.P. Tavadze, G. Avendaño Franco, P. Ren, X. Wen, Y. Li and J. P. Lewis, *J. Am. Chem. Soc.*, 2018, **140**, 285-290.
- 43.H. Rau and E. Lueddecke, *J. Am. Chem. Soc.*, 1982, **104**, 1616-1620.
- 44.L. Zhao, J. Liu and P. Zhou, *J. Phys. Chem. A*, 2017, **121**, 141-150.
- 45.T.-T. Yin, Z.-X. Zhao and H.-X. Zhang, *New J. Chem.*, 2017, **41**, 1659-1669.