Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Exploring the chemical space in the search for improved Azoheteroarene-based photo-switches

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

Table of Contents

S1. Computational Benchmark

The first ten singlet excited states of Azobenzene (*case 1*) and three azoheteroarenes (*cases 2-4*) have been computed using an array of TDDFT functionals and basis sets. We have explored the hybrid functionals ωB97-XD, CAM-B3LYP, B3LYP, M06, M06-2X and PBE0. Notice that the former two functionals include long-range corrections. All along this benchmark, we have taken as a reference the state energies obtained with CC2 and the TZVP basis set. We have done so for two reasons. First, because the vertical energies obtained with TDDFT do not correspond to the state stabilities accessible experimentally due to the absence of zero-point corrections.[1] Second, because a direct comparison with experiment would imply the adoption of an implicit-solvent method, which would imply another layer of computational benchmark, with no clear insight on the properties studied in this work. Moreover, solvent corrections are not yet standard in the ADC2, CC2 and CIS(D) wavefunction-based calculations employed herein (some progress on this topic[2]). In all those calculations, the initial geometry was computed at the PBE0/def2-TZVP level.

Case 1: Azobenzene

Excitation of *E*-Azobenzene from S_0 to S_1 leads to a weakly-allowed nπ* 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.[3] In turn, excitation from S_0 to S_2 leads to an intense symmetry-allowed ππ* transition in the UV range ($\lambda \approx 320$ nm). The error in excitation wavelength of S_1 and S_2 obtained with the six functionals and seven basis sets is shown in Figure S1, with respect to the CC2 values. For S_1, CAM-B3LYP and ωB97-XD give outstanding results, with errors limited to few nm, whereas the error in the remaining functionals raises up to 50-70 nm. In turn, for S_2 we observe that M06-2X joins CAM-B3LYP and ωB97-XD in the group of functionals that perform well, whereas PBE0, B3LYP and M06 perform poorly, with errors of *ca.* 40 nm.

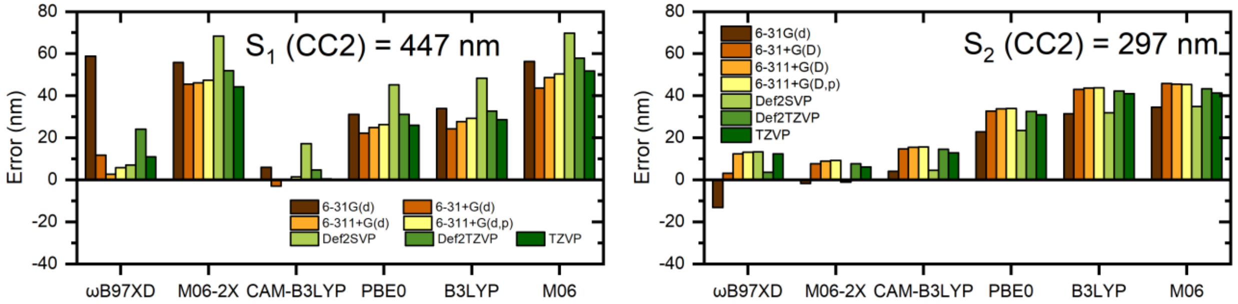


Figure S1. Error in the wavelength associated with the computed S_0-S_1 and S_0-S_2 transitions in E-azobenzene using different DFT functionals and basis sets. The reference value obtained with CC2 is highlighted on top.

All functionals describe an S_0-S_1 transition of nπ* character, centered in the Azo group (see Figure S2) and a S_0-S_2 transition of ππ* character with both energy and charge transfer components.

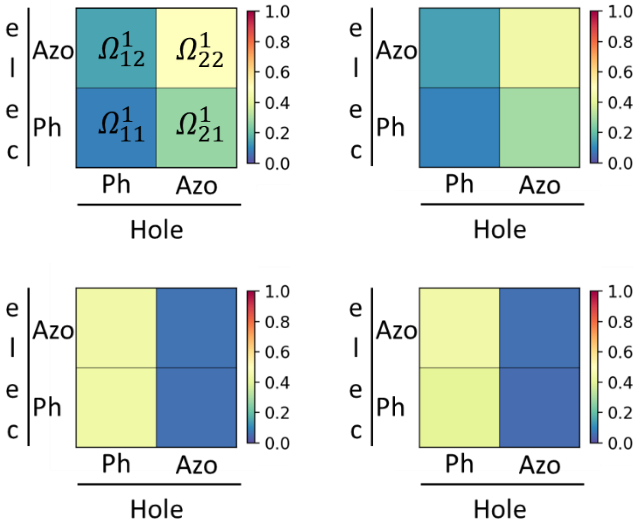


Figure S2. Excitation fingerprints associated with the S_1 (top) and S_2 (below) states of E-Azobenzene obtained with ωB97-XD (left) and CC2 (right).

The nπ* transition of Z-Azobenzene appears in the visible region ($\lambda \approx 450$ nm), similar to the E-isomer. In turn, the ππ* transition appears blue-shifted with respect to the E-isomer ($\lambda \leq 300$ nm). Both trends are captured by both the CC2 method and by the array of TDDFT calculations (see Figure S3).

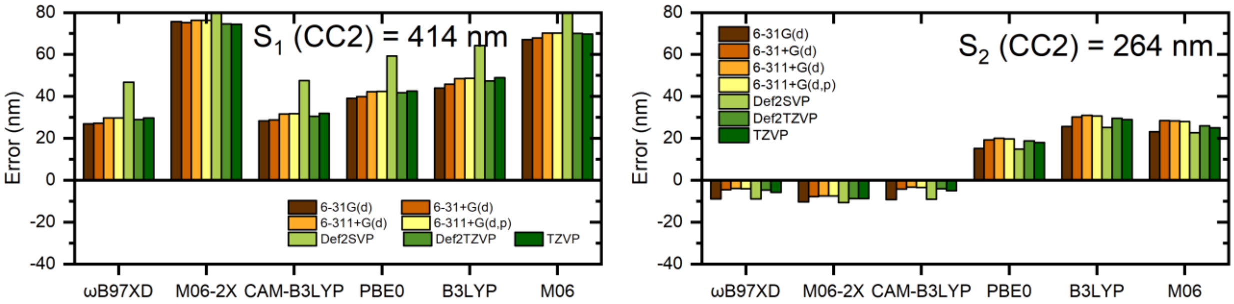


Figure S3. Error in the wavelength associated with the computed S_0-S_1 and S_0-S_2 transitions in Z-azobenzene using different DFT functionals and basis sets. The reference value obtained with CC2 is highlighted on top.

In this case, however, we observe differences in the EF obtained from TDDFT and CC2 calculations. The nature of the excitations is correctly described, and the usual features of both the nπ* (*i.e. strong* Ω^1_{22}) and ππ* (*i.e.* strong Ω^2_{11} and Ω^2_{12}) excitations can be easily identified (see Figure S4). The main difference lies in the amount of energy transfer within the Ph group in S_0-S_2 (Ω^2_{11}), which seems surprisingly large in CC2. On the contrary, ωB97-XD (and other functionals) lead to a ππ* excitation more spread into the Ph and Azo group, in a way that the excitation fingerprint is comparable to that of the E-isomer.

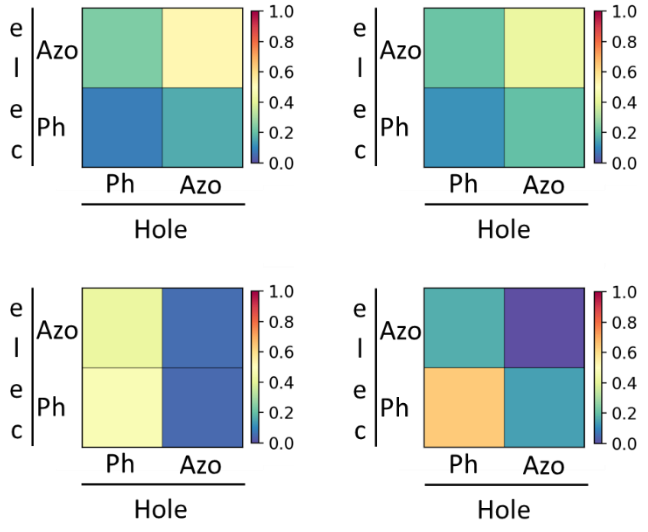


Figure S4. Excitation fingerprints associated with the S_1 (top) and S_2 (below) states of Z-Azobenzene obtained with ωB97-XD (left) and CC2 (right).

Case 2: AAP-12. Solvent vs. Gas Phase

The second compound is one of the arylazopyrazoles (AAPs) recently synthesized by Ravoo and coworkers.[4] We selected AAP-12 because it is one of the compounds highlighted in the paper (together with AAP-2, -5 and -13). Therein, the nπ* and ππ* bands in the E-isomer appear at 400 and 347 nm, respectively, in water solvent. CC2 in gas-phase finds these transitions at 389 and 305 nm. Similar to the case of Azobenzene, CAM-B3LYP and ωB97-XD provide the best energy for S_1, and these two functionals, together with M06-2X, are able to describe S_2 correctly (see Figure S5). We note, however, that even in the best functionals incur in a significant red-shift in the energy of the nπ* transition.

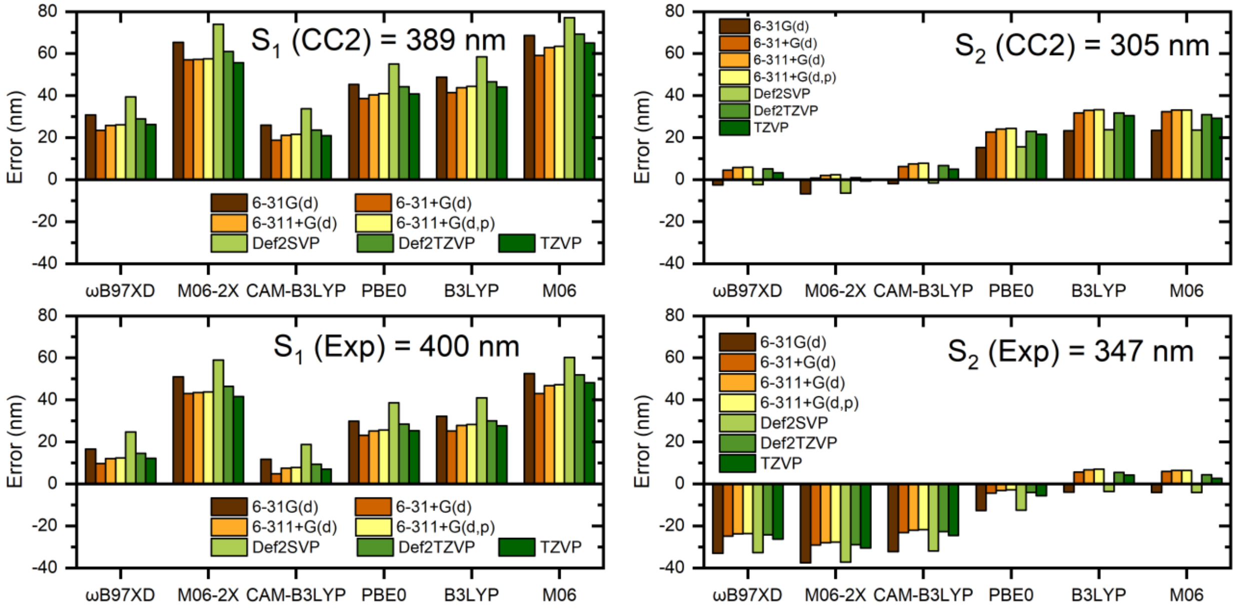


Figure S5. Error in the wavelength associated with the computed S_0-S_1 and S_0-S_2 transitions in E-**AAP-12** using different DFT functionals and basis sets. Gas-phase values (top) are compared with the reference CC2 values, whereas the solvent-phase (below) are compared with experiment. Reference values are highlighted on top.

All DFT functionals and CC2 show the expected EF of S_1 and S_2, with a strong Azo-centered nπ* excitation (*i.e.* Ω^1_{22}) and an Azo-to-Ph/Pz ππ* transition (*i.e.* Ω^2_{23} and Ω^2_{43}) (see Figure S6).

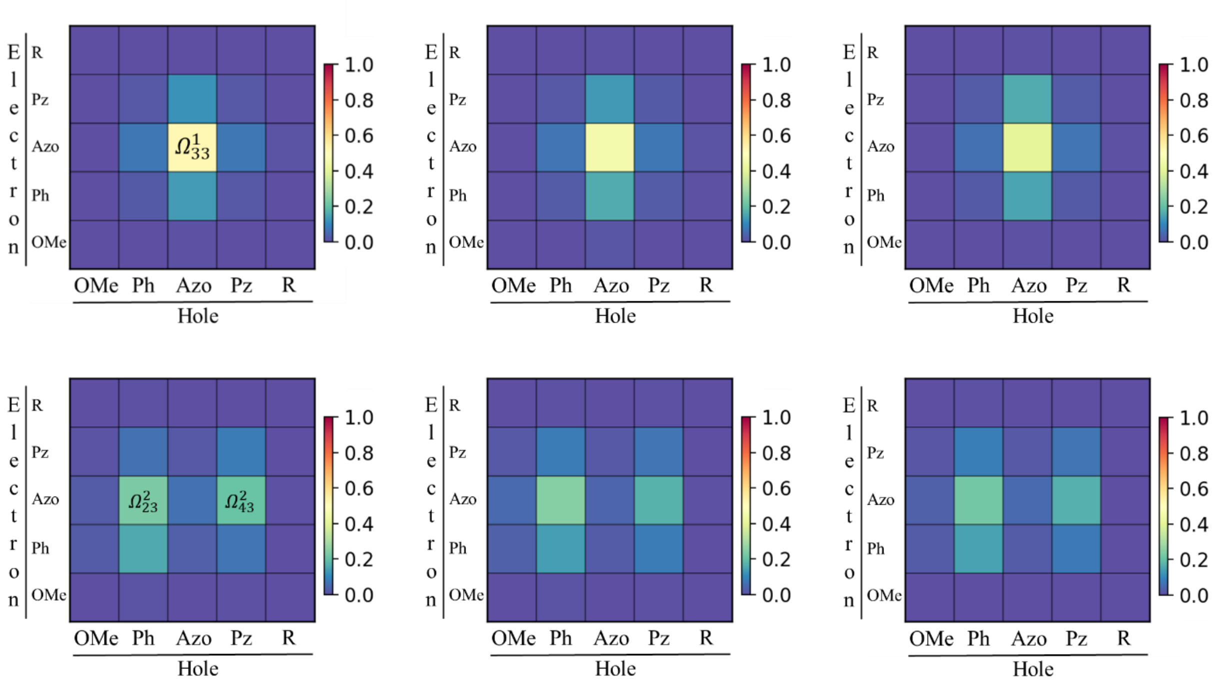


Figure S6. Excitation fingerprints associated with the S_1 (top) and S_2 (below) states of **AAP-12** obtained with ωB97-XD (left), PBE0 (middle) and CC2 (right), and the TZVP basis set.

However, the actual values for some of the larger Ω^l_{AB} differ significantly between functionals (see Figure S7). For instance, Ω^1_{33} is 0.52, 0.51, 0.45 and 0.40 with ωB97-XD, CAM-B3LYP, PBE0 and CC2 respectively (TZVP basis set). It is interesting that the functionals providing most similar energetics to CC2 (ωB97-XD and CAM-B3LYP) are also the ones with a larger difference in Ω^1_{33}. The same behavior is observed for Ω^2_{23} and Ω^2_{43}.

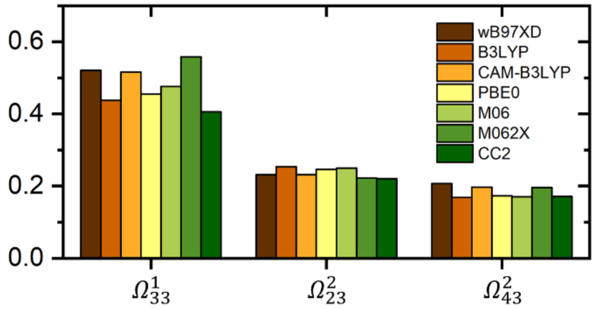


Figure S7. Charge-transfer numbers highlighted in Figure S6 computed with different methods and the TZVP basis set.

As we previously mentioned, the experimental absorption spectrum of AAP-12 was collected in water. Being a polar solvent, we decided to evaluate the how the gas-phase properties are modified in the presence of solvent. To do so, the same analysis above has been carried out using the CPCM implicit solvent model as implemented in Gaussian 09. The energy of the bands only slightly shifted. For the E-isomer, the nπ* band is shifted from 415 nm in gas-phase to 412 in water, whereas the ππ* band is shifted from 308 in gas-phase to 320 nm in water. For the Z-isomer, the shifts associated with the solvent are −5 (nπ*) and +10 nm (ππ*). These results have been obtained with ωB97-XD/TZVP, but other functionals and basis sets provide very similar deviations. Finally, the EF obtained under the influence of water solvent are not modified with respect to the gas-phase ones. Therefore, it can be concluded that the excited state properties of AAP-12 are not significantly affected by the presence of solvent.

Case 3: 3pzH

The third compound studied in this benchmark is the 3-pyrazole azoheteroarene (3pzH) recently synthesized by Fuchter and coworkers.[5] Similar to the case of AAP-12, compound 3pzH was highlighted in the original paper due to its extraordinary properties as a photoswitch. The nπ* and ππ* bands in its E-isomer appear at 425 and 320 nm, respectively, in acetonitrile. CC2 in gas-phase finds these transitions at 421 and 269 nm. Similar to the cases of Azobenzene and AAP-12 described above, CAM-B3LYP and ωB97-XD provide the best energy for S_1, and these two functionals, together with M06-2X, are able to describe S_2 with the closest agreement with CC2 (see Figure S8), yet it must be mentioned that this reference value seems to overestimate the energy of S_2 with respect to the experiment (269 *vs.* 320 nm).

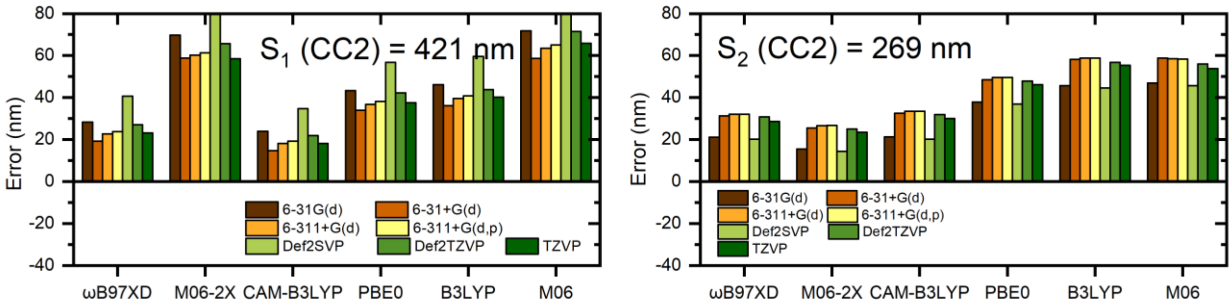


Figure S8. Error in the wavelength associated with the computed S_0-S_1 and S_0-S_2 transitions in E-**3PzH** using different DFT functionals and basis sets. The reference value obtained with CC2 is highlighted on top.

All functionals describe the nπ* and ππ* transitions in the same way than CC2 (see Figure S9). The former is characterized by a large energy transfer within the azo, and by a small charge transfer from the azo to the Pz and Ph rings. The latter is, in turn, characterized by energy transfer within the Ph and Pz groups, and by a small charge transfer from the Ph to the Pz.

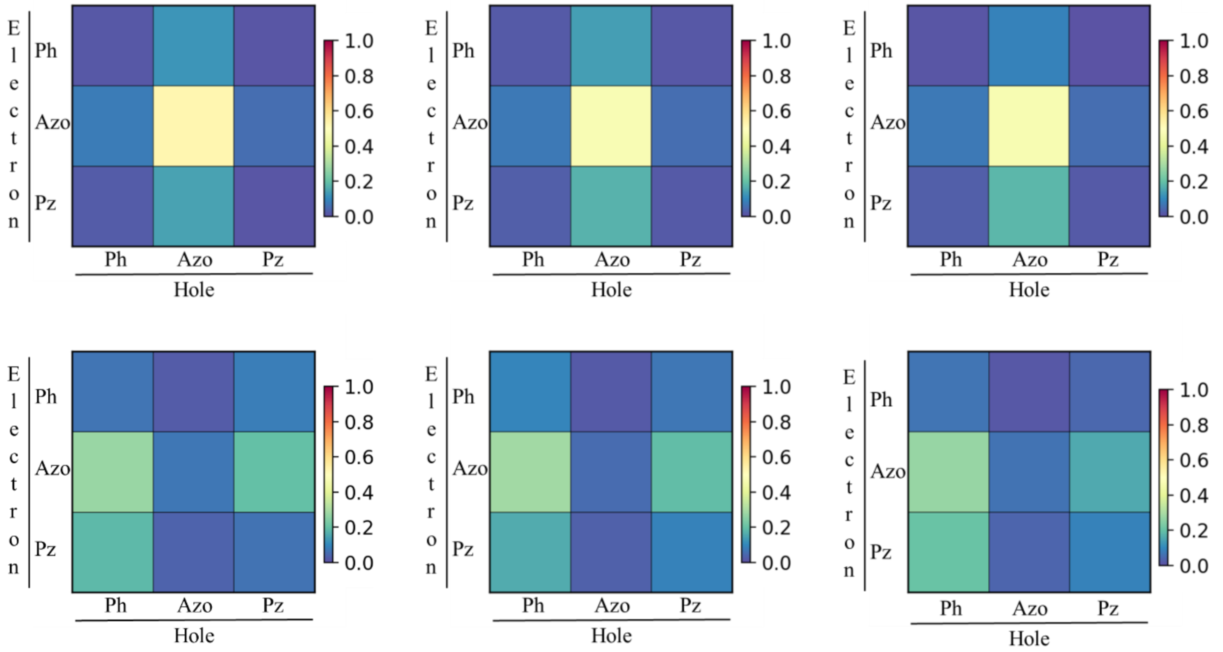


Figure S9. Excitation fingerprints associated with the S_1 (top) and S_2 (below) states of **3pzH** obtained with ωB97-XD (left), PBE0 (middle) and CC2 (right), and the TZVP basis set.

*Case 4: C**534***

The fourth compound studied in this benchmark is compound C**534**, first reported in this manuscript. Compound C**534** combines 2-py in R_1 (R_1=**5**), Thiophene in R_2 (R_2=**3**) and Imidazole in R_3 (R_3=**4**, see Table 1). It is therefore an "ortho" substituted azoheteroarene (*o*-R_1). Once again, CAM-B3LYP and ωB97-XD provide the "best" (*i.e.* closer to CC2) energy for S_1, and these two functionals, together with M06-2X, give the energy in closest agreement with CC2 (see Figure S10).

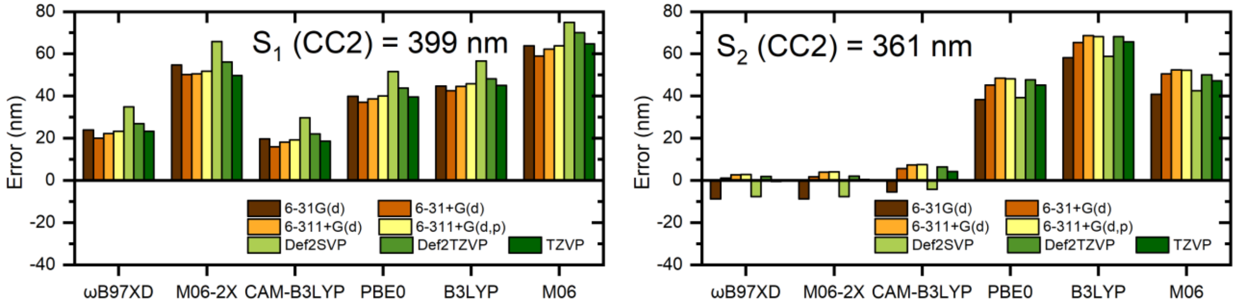


Figure S10. Error in the wavelength associated with the computed S_0-S_1 and S_0-S_2 transitions in C**322** using different DFT functionals and basis sets. The reference value obtained with CC2 is highlighted on top.

Proceeding to analyze the EF of S_1 and S_2, we observe once again that the S_0-S_1 transition is correctly captured by both PBE0 and ωB97-XD (see Figure S11). In turn, the S_0-S_2 transition is correctly described by ωB97-XD but not by PBE0, which switches the order of the S_2 and S_3 states. We have confirmed that CAM-B3LYP gives excitation fingerprints that are equivalent to those obtained with ωB97-XD. At this stage we need to mention that we will continue referring to the $\pi\pi^*$ transition as the S_2 state, even if it actually is the third singlet state (S_3) when computed with PBE0. The three lowest singlet state energies are, respectively: 2.94, 3.44 and 3.84 eV for ωB97-XD, 3.10, 3.43 and 3.72 eV for CC2 and 2.83, 3.05 and 3.20 eV for PBE0. Therefore, we can conclude that PBE0 provides a moderately good energy for S_2 (3.20 *vs.* 3.43 eV with CC2) but indeed underestimates the energy of S_3 (3.05 *vs.* 3.72 eV with CC2). By looking at the excitation fingerprints of S_3, we observed that it is an electronic transition characterized by a dominant electron transfer from R_1 and R_3 to the azo group.

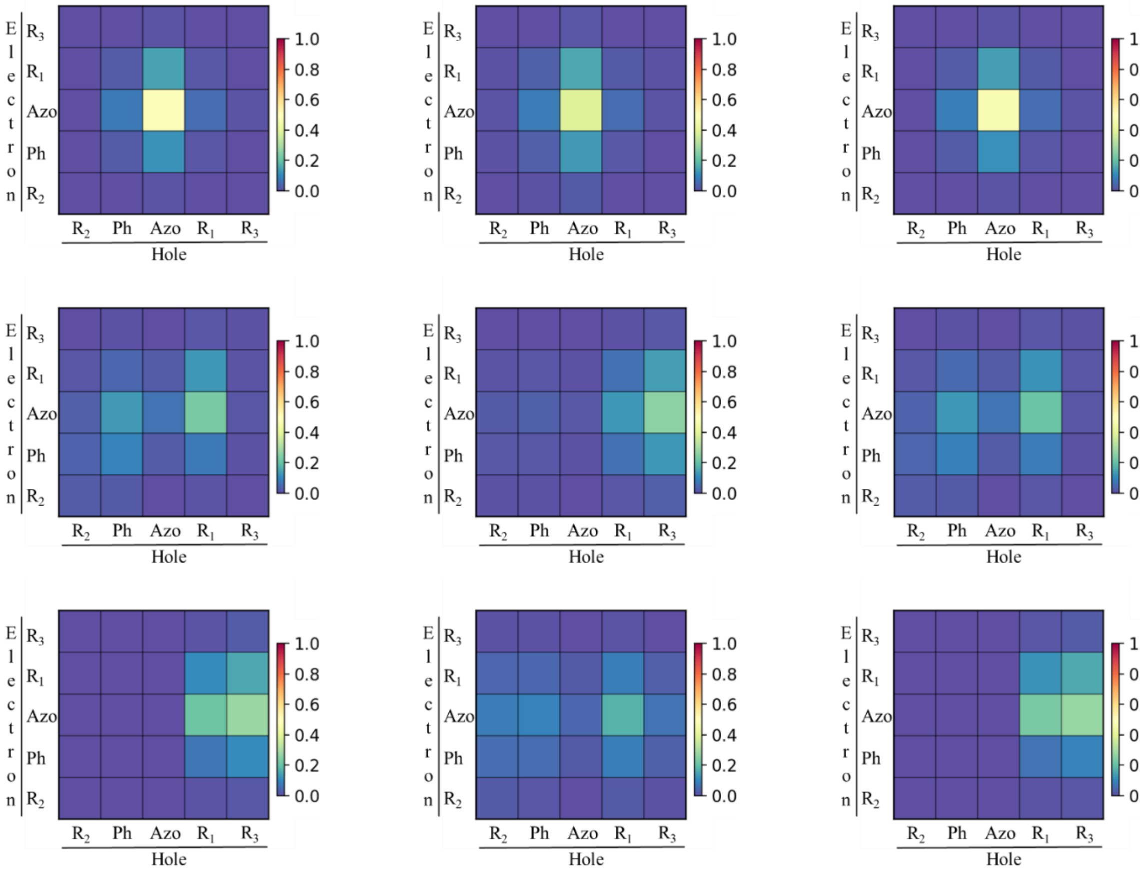


Figure S11. Excitation fingerprints associated with the S_1 (top) and S_2 (middle) and S_3 (below) states of C**322** obtained with ωB97-XD (left), PBE0 (middle) and CC2 (right), and the TZVP basis set.

Conclusion

In this supporting section we have benchmarked the energies and nature (*i.e.* fragment-based excitation fingerprints) of the relevant excited states (S_1 and S_2) of azobenzene, AAP-12, 3pzH and C**534**. From the results described above, we conclude that the range-separated functionals CAM-B3LYP and ωB97-XD are the best functionals to describe the relevant excited states (S_1 and S_2). Indeed, the agreement of S_2 energies with CC2 is particularly good. We therefore coincide with the conclusion extracted from previous works.[6, 7] We can also conclude that most of the basis sets studied are suitable (except def2-SVP), being TZVP the best one. In any case, the basis set seems to be a secondary parameter, much less important than the DFT functional. For these reasons, the TDDFT calculations described in this manuscript have been performed at the ωB97-XD/TZVP level.

S2. Structure of E- and Z-isomers

A. Conformers. H-bond formation

A striking observation in some figures of the main text is the different behaviour of compounds with R_1=**5**,**7**,**8** *vs.* those with R_1=**6** when 4-im is employed in R_3 (R_3=**4**). The reason is that each set of compounds displays a different conformer of the θ angle. The former group has the N-H group of 4-imi pointing towards the azo group, while the other display the 180º-rotated one, with the C-H group pointing towards the azo (see Figure S12). This behaviour is captured by an energy scan along θ, which demonstrates that the conformers obtained in each case are the minimum-energy ones. The different behavior of R_1=**6** stems from the fact that R_1 has a N-atom next to the R_3-substitution, while R_1=**5**,**7**,**8** have a CH group in this position, which generates a repulsion with the NH group of the 4-imidazole (see left-part of blue curve in Figure S12).

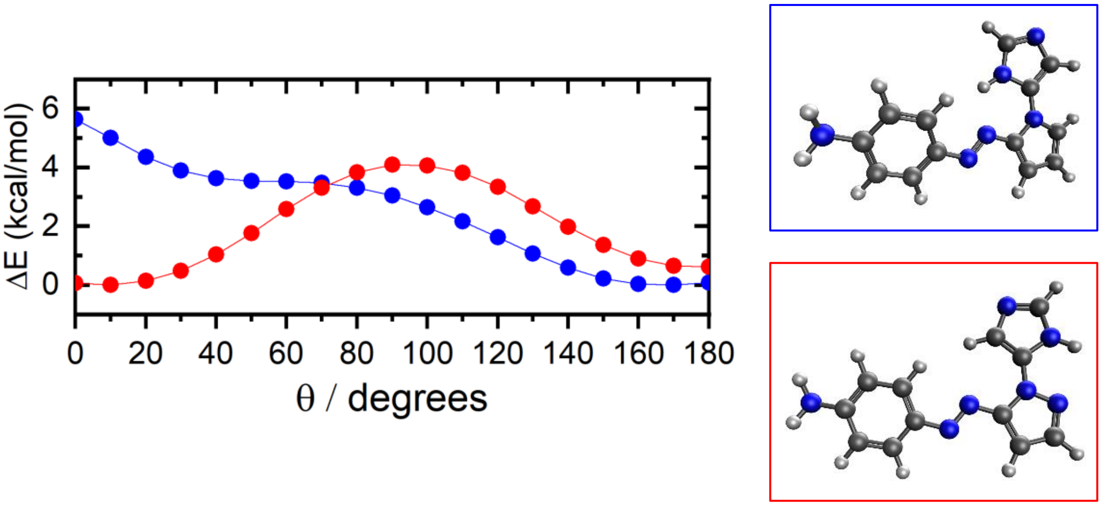


Figure S12. Energy profile along a rotation about θ in compounds C**514** (blue curve) and C**614** (red curve), whose minimum energy structures are shown at the right. In the plot, 0º refers to the conformation with C-atom pointing towards the azo, and 180º is the one with the N-atom.

Only the most stable conformer of θ and φ has been considered for all compounds. An alternative would have been to Boltzmann-weight the properties of each conformer.[5] However, this would have increased the computational cost for little gain. First, because the potential impact of different R_2 (φ) and R_3 (θ) conformers in the studied properties is, in general, very small. For instance, the R_2 substitution in *para* position of the benzene limits significantly the impact of a rotation about φ. The same applies to the θ rotation in compounds with *m*-R_1. In principle, a significant effect could be expected in *o*-R_1, given the proximity of the R_3 fragment to the azo group. However, in practice one conformer is typically much more stable than the other. Indeed, as we have just seen, some conformers are not even a minimum energy structure and hence are not attainable. In the case discussed in this section, the normalized Boltzmann weight for the minor conformers of C**514** and C**614** at 300K is 0.1% and 25%, respectively. It thus follows that, even in this case, our description using a single conformer is a good approximation.

B. Geometry of E- and Z-isomers

To facilitate the comparison and the statistical treatment of the dihedral angles, we have numerically treated them as deviations from planarity, irrespectively of this deviation being clockwise (positive values) or anticlockwise (negative values). For instance, the value of the β angle that we give for the E-isomers in Table S1 is of around 0º, indicating the deviation from planarity, instead of the expected 180º.

E-Isomers:

Table S1. Partial averages of the five torsion angles defining the structure of the E-isomers (see Scheme 2). For instance, the entry R_1=**1** indicates the average value for all compounds with R_1=**1**, irrespectively of R_2 and R_3.

R_1	α (º)	β (º)	γ (º)	φ (º)	θ (º)	R_2	α (º)	β (º)	γ (º)	φ (º)	θ (º)	R_3	α (º)	β (º)	γ (º)	φ (º)	θ (º)
1	0.4	0.1	0.3	8.3	5.4	**0**	2.6	0.5	2.5		21.4	**0**	0.3	0.1	0.1	8.2	
2	0.6	0.1	0.3	9.5	24.3	**1**	1.7	0.6	2.8		20.2	**1**	0.5	0.1	0.3	10.1	
3	0.6	0.1	0.4	9.2	6.4	**2**	1.9	0.6	2.7		20.7	**2**	2.0	0.7	2.9	8.4	27.8
4	0.6	0.1	0.4	9.4	26.9	**3**	2.2	0.6	2.4	21.5	23.0	**3**	1.4	0.6	3.0	9.0	27.6
5	3.7	0.9	3.8	9.4	31.7	**4**	2.3	0.6	2.6	0.6	21.6	**4**	6.5	1.2	2.2	9.2	18.0
6	3.0	1.1	5.3	8.8	20.3	**5**	2.0	0.6	2.3	19.9	22.8	**5**	1.9	0.7	3.9	8.1	32.0
7	4.3	1.1	4.4	9.0	27.9	**6**	2.2	0.6	2.5	13.9	22.9	**6**	2.1	0.7	3.9	8.6	31.4
8	3.6	1.1	5.3	9.0	32.1	**7**	1.7	0.6	2.5	5.8	22.4	**7**	2.0	0.7	4.0	11.0	31.1

Notice that all regioisomers of pyrazole (R_1=**1**, **3**, **5**) lead to smaller values of θ due to the presence of an N atom (instead of a CH) adjacent to the position where R_3 is attached. Furane (R_2=**4**) and Thiazine (R_2=**7**) groups favour planarity in the R_2-Ph torsion, while 4-imidazole (R_3=**4**) favours planarity of the R_1-R_3 torsion (θ ~ 18º).

Z-Isomers:

Table S2. Partial averages of the five torsion angles defining the structure of the Z-isomers (see Scheme 2). For instance, the entry R_1=**1** indicates the average value for all systems with R_1=**1**, irrespectively of R_2 and R_3.

R_1	α (º)	β (º)	γ (º)	φ (º)	θ (º)	R_2	α (º)	β (º)	γ (º)	φ (º)	θ (º)	R_3	α (º)	β (º)	γ (º)	φ (º)	θ (º)
1	51.6	9.0	29.0	19.5	21.7	**0**	69.2	5.5	14.5		24.4	**0**	68.2	5.3	10.9	19.3	
2	71.4	4.8	10.5	20.3	28.1	**1**	48.7	9.8	24.2		23.1	**1**	64.9	6.4	14.0	17.6	
3	79.2	2.6	6.1	20.3	10.3	**2**	57.4	8.2	20.8		25.8	**2**	58.5	8.1	20.1	17.1	
4	45.1	11.5	24.2	15.9	35.1	**3**	63.9	7.1	17.3	24.7	28.2	**3**	58.8	7.9	20.1	17.2	12.5
5	66.5	6.3	15.1	16.4	30.0	**4**	61.7	7.5	17.9	1.5	28.6	**4**	60.0	7.6	16.4	16.7	18.7
6	52.4	9.8	28.1	13.0	25.3	**5**	62.1	7.4	17.8	24.6	28.5	**5**	55.3	8.9	23.8	16.1	30.9
7	70.1	5.0	12.9	17.8	29.8	**6**	62.6	7.2	18.0	16.6	28.6	**6**	55.9	8.7	23.0	16.0	31.0
8	39.0	13.1	27.7	12.3	29.1	**7**	49.8	9.6	23.2	6.9	22.1	**7**	53.7	9.2	25.3	15.4	30.7

We notice that the "*twisted*" structure has not been obtained in significant amounts. This structure is characteristic of bis-ortho-substituted azoheteroarenes because, in these compounds, both conformers of θ lead to steric hindrance with the Ph ring and the heteroarene has no other option than to rotate.[5, 8] However, being our compounds only mono-substituted, they can adopt the conformer in which R_3 is pointing outwards (*i.e.* far from the Ph group, see Figures 1 and S13b)

C. Thermal Stability

A computational estimation of the half-life times ($t_{1/2}$) would require obtaining the relative energy difference between the minima (*i.e.* ΔH_{EZ}), and also the transition state structure, giving access to the thermodynamic and kinetic profiles. While computational works are increasingly providing both profiles to estimate $t_{1/2}$,[5, 6, 9] in this work we could not automatize such process for the 512 compounds. An alternative to this approach is to quantify the strength of the azo N=N bond in the Z-isomers through the distance and/or the bond order. The reason is that the thermal isomerization through the rotation mechanism is favored when this bond acquires some single-bond character (through either *push-pull resonance* or *tautomerism*[10, 11]). As demonstrated by Füchter and coworkers, the quantification of the bond order (*e.g.* through the Wiberg indices, WI) provides a qualitative trend in the expected $t_{1/2}$ at almost no cost. In their analysis, the authors find that the compounds in which the correlation is worse are the ortho-substituted ones in which the Z-isomer is either (i) displaying the *T-shape* structure (larger $t_{1/2}$ than predicted), or (ii) too sterically-destabilized (shorter $t_{1/2}$ than predicted). The first type is not common in our dataset (see paragraph above), and virtually none of our systems fits the second type. Therefore, we decided to use this metric in conjunction with ΔH_{EZ} to analyze the expected half-live of our compounds. For testing purposes, ΔG_{EZ} was also evaluated, yielding similar values to ΔH_{EZ} in average (see Table S3), but more spread, which made much more difficult the identification of patterns.

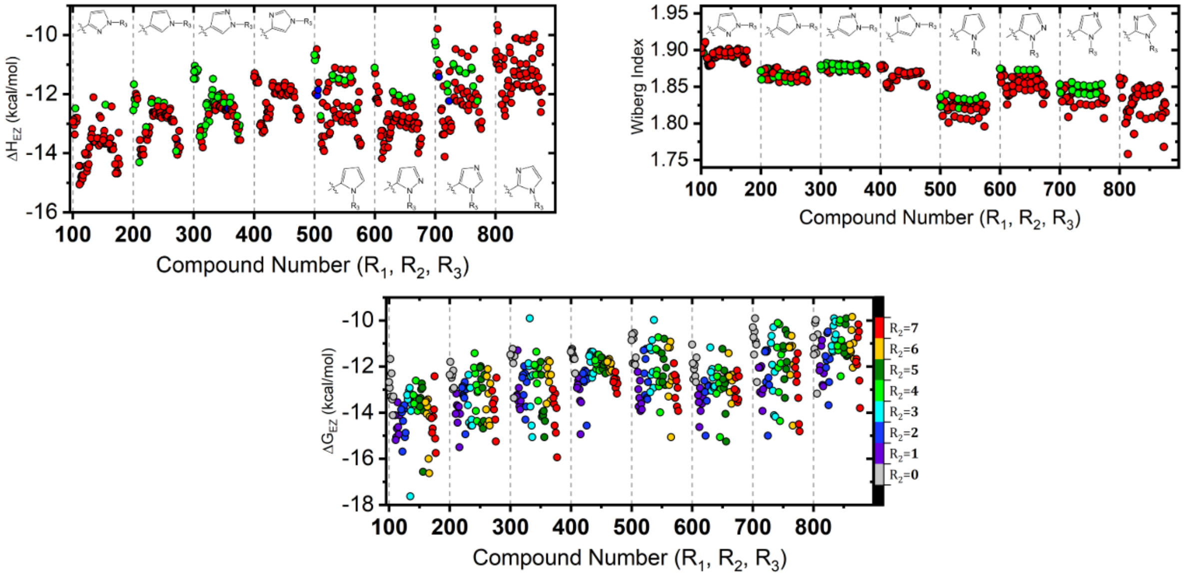

(a) (b)

Figure S13. (a) Relative thermal stability, and (b) WI *vs.* structure type (red=*folded*, green=*T-shape*, blue=*twisted*). (c) Thermal stability evaluated using the free energy difference ΔG_{EZ}.

Finally, our results indicate that, in general, Z-isomers adopting the *T-shape* structure are more stable (*i.e.* less negative ΔH_{EZ}) than those adopting the *folded* structure (see Figure S13a).

Table S3. Partial averages of ΔH_{EZ} depending on chemical composition. For comparison, the Free-Energy differences (ΔG_{EZ}) are also provided.

R_1	ΔH_{EZ}	ΔG_{EZ}	R_2	ΔH_{EZ}	ΔG_{EZ}	R_3	ΔH_{EZ}	ΔG_{EZ}
1	−13.7	−14.0	**0**	−11.5	−11.8	**0**	−12.3	−12.1
2	−13.0	−13.3	**1**	−13.4	−13.3	**1**	−12.2	−12.3
3	−12.5	−12.9	**2**	−12.8	−12.9	**2**	−12.7	−12.7
4	−12.2	−12.2	**3**	−12.1	−12.4	**3**	−12.3	−12.3
5	−12.3	−12.3	**4**	−12.2	−12.3	**4**	−12.1	−12.0
6	−12.8	−12.8	**5**	−12.3	−12.5	**5**	−12.8	−13.2
7	−11.9	−12.0	**6**	−12.3	−12.5	**6**	−12.8	−13.4
8	−11.2	−11.3	**7**	−13.1	−13.1	**7**	−12.6	−12.8

D. Wiberg Indices

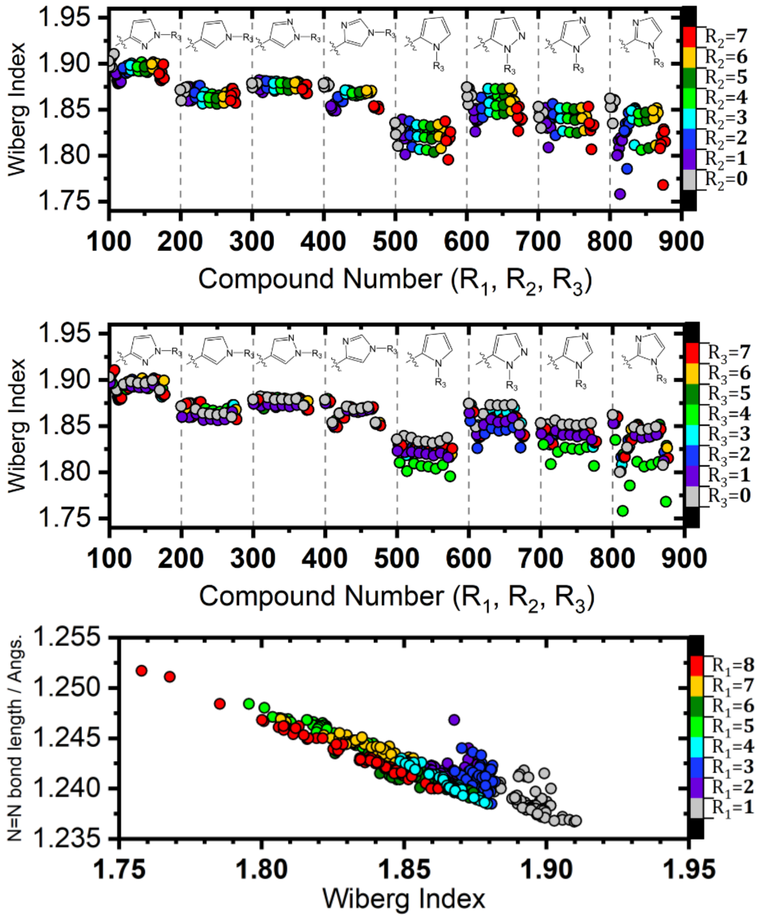


Figure S14. Wiberg Index associated with the computed Z-isomers. The color code indicates the (top) R_2 or (middle) R_3 fragment. Compounds are ordered using its three-digit identifier. (below) Correlation between the N=N bond length and the Wiberg index. In this case, the color code indicates R_1.

Table S4. Wiberg indices (WI) averaged by chemical composition.

R_1	WI	R_2	WI	R_3	WI
1	1.89	**0**	1.87	**0**	1.86
2	1.86	**1**	1.85	**1**	1.85
3	1.88	**2**	1.86	**2**	1.86
4	1.86	**3**	1.86	**3**	1.86
5	1.82	**4**	1.86	**4**	1.85
6	1.85	**5**	1.86	**5**	1.86
7	1.84	**6**	1.86	**6**	1.86
8	1.83	**7**	1.85	**7**	1.86

S3. Photochemistry of the E-isomers

A. Energy of the nπ and ππ* transitions*

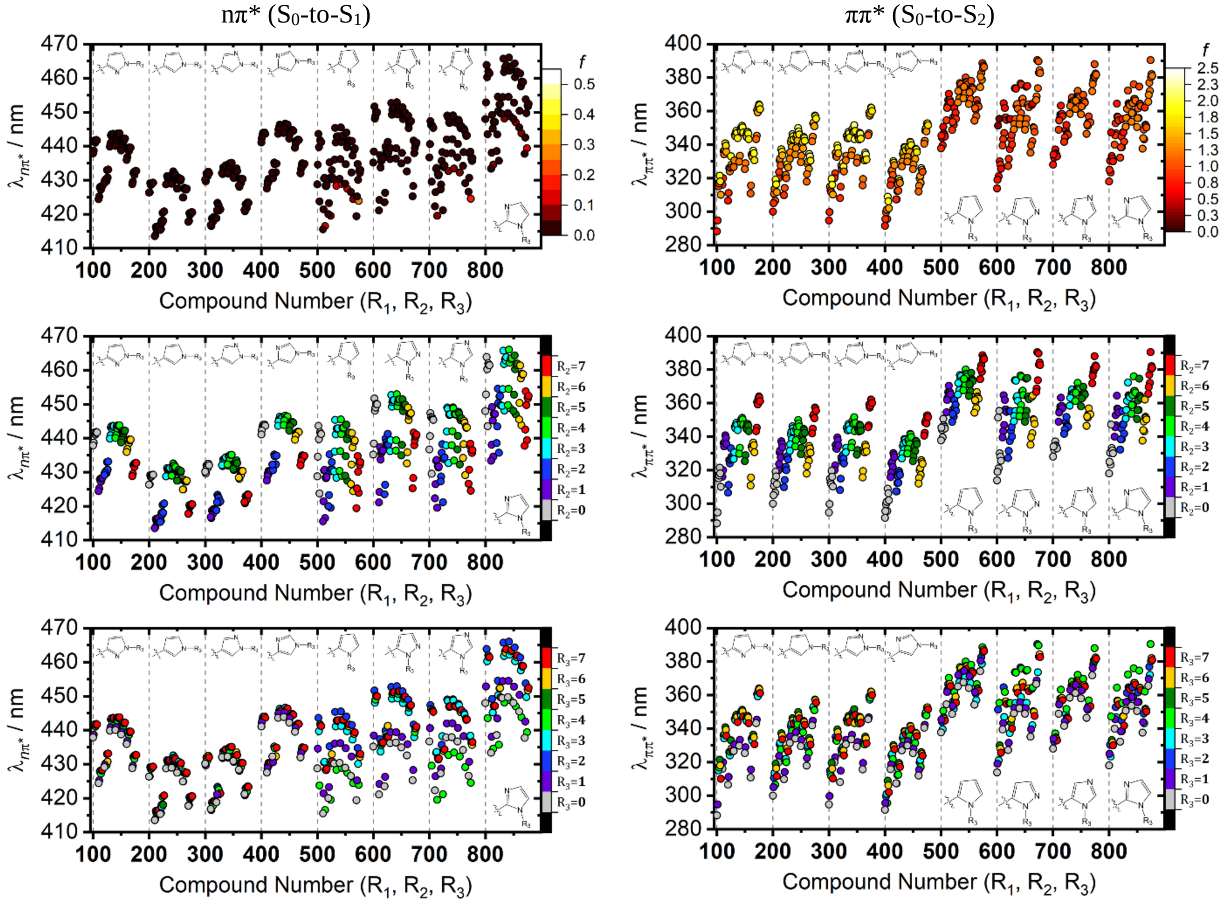
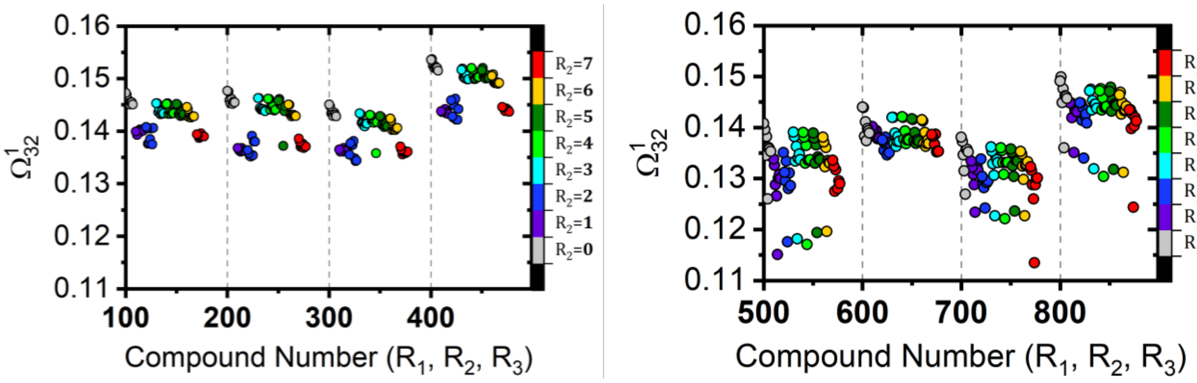


Figure S15. Excitation wavelength associated with the (left) nπ* and (right) ππ* transitions in the E-isomer of the studied compounds, ordered using its three-digit identifier. Color code indicates (top) oscillator strength, (middle) R_2 index and (bottom) R_3 index.

B. Relationship between relevant Ω_{AB}^{l} values the energy and intensity of the nπ transition (S_0-S_1).*



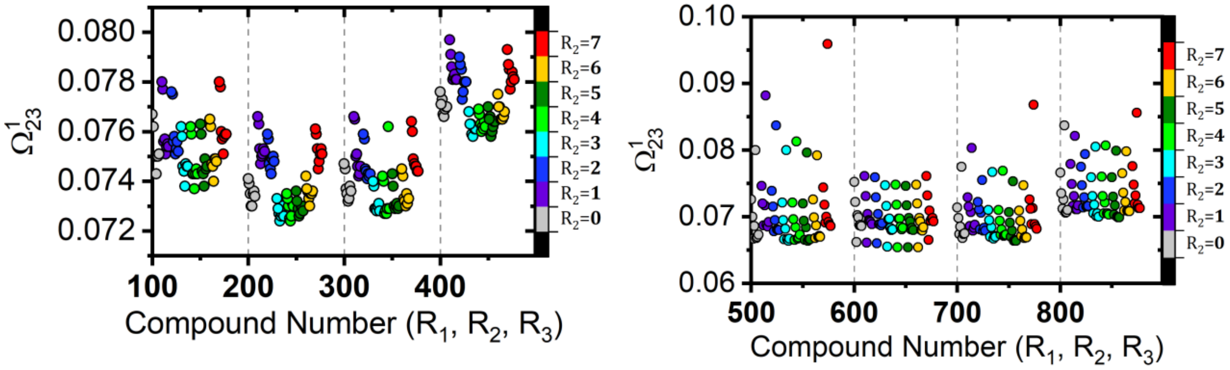


Figure S16. (top) Azo-to-Ph (Ω^1_{32}) (bottom) and Ph-to-azo (Ω^1_{23}) charge-transfer components associated with the nπ* transition of the E-isomers. There is a clear relationship between those values and the excitation energies depicted in Figure 4, especially for compounds with meta-substituted R_1 (left). Notice the larger point dispersion within same-R_1 values in ortho-substituted R_1 (right) showing the larger effect of R_3.

S4. Photochemistry of the Z-isomers

A. Energy of the nπ and ππ* transitions*

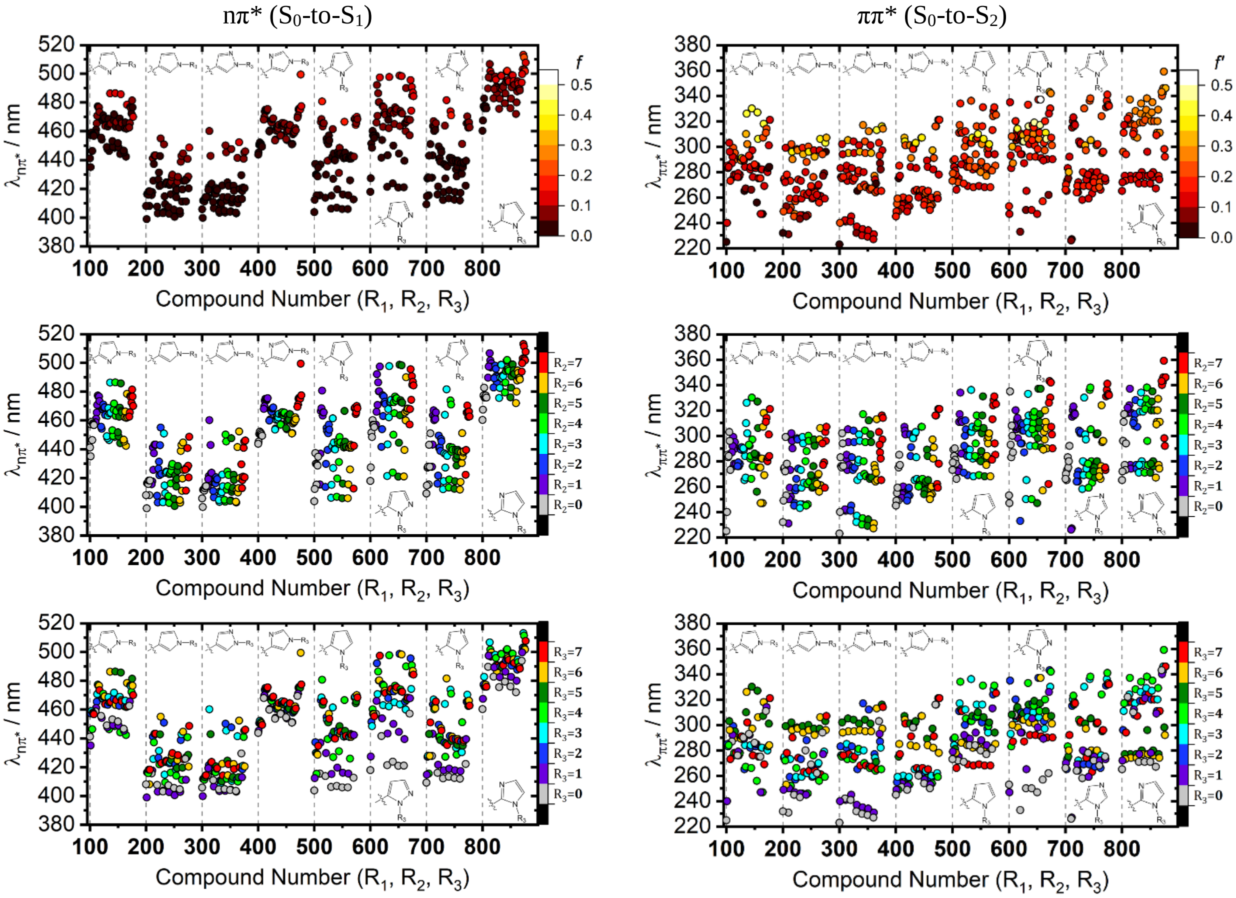


Figure S17. Excitation wavelength associated with the (left) nπ* and (right) ππ* transitions in the Z-isomer of the studied compounds, ordered using its three-digit identifier. Color code indicates (top) oscillator strength (*f* or *f'*, see discussion in Section S4.C), (middle) R_2 index and (bottom) R_3 index.

B. Relationship between structure and excitation wavelength of the S_0-S_1 (nπ) transition*

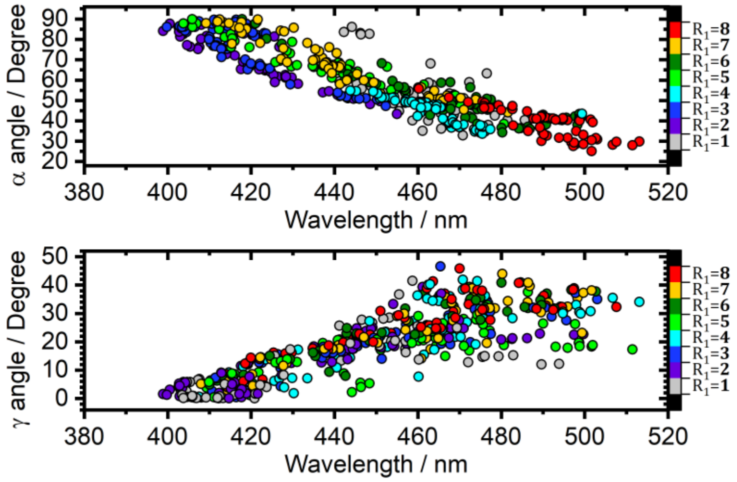


Figure S18. Relationship between the excitation wavelength of the nπ* transition in Z-isomers, and the torsion angles (top) α and (below) γ.

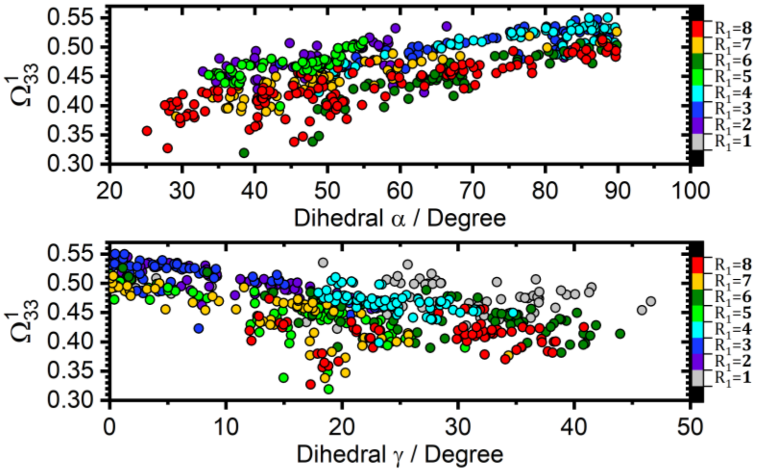


Figure S19. Relationship between the Ω^1_{33} component of the nπ* transition in Z-isomers, and the torsion angles (top) α and (below) γ.

C. Identification of the ππ transition.*

As we mention in the main text, the identification of the *productive ππ** transition in the Z-isomer is not straightforward. In principle, the *ππ** transition refers to a bright transition of *ππ** character involving the S_2 state that should lead to isomerization of the azoheteroarene (it is *productive*). In practice, in our set of compounds the transition to S_2 is often dark, with the brightest one involving either S_2, S_3 or even S_4 and above. Also, some of the low-energy *ππ** transitions that we found are, indeed, completely localized in the outer molecular fragments (energy transfer within R_3), so they are not likely to lead to isomerization. Notice that we faced a similar problem when discussing the E-isomer, but it was restricted to few compounds, so we could handpick the transition. Given that in this case it is a general feature, we had to define a criterion of wide applicability based on the available data (*i.e.* intensity, energy and character), to select which of the low-energy transitions displayed by our set of compounds is comparable to the "*expected*" – or "*productive*"- ππ* transition. Finally, we identified this transition as the one that is (i) low in energy. (ii) has significant intensity, and (iii) the electron is brought to the azo group. Mathematically, this is done using Ω_{AB}^{l}-weighted intensities (referred to as f'), in which the regular intensity *f* is multiplied by the total amount of electron density that is transferred to the azo (from any fragment, *i.e.* Ω_{X3}^{l}). The lowest energy transition for which *f'* is larger than a threshold of 0.1 is selected. Of course, the *nπ** state is excluded from this analysis. We are aware that a relationship between Ω_{X3}^{l} and the isomerization quantum yield has not been proven. However, in doing so we obtain a unified mathematical criterion to treat multiple types of *ππ** excitations, and excludes transitions that are likely to have small quantum yields.

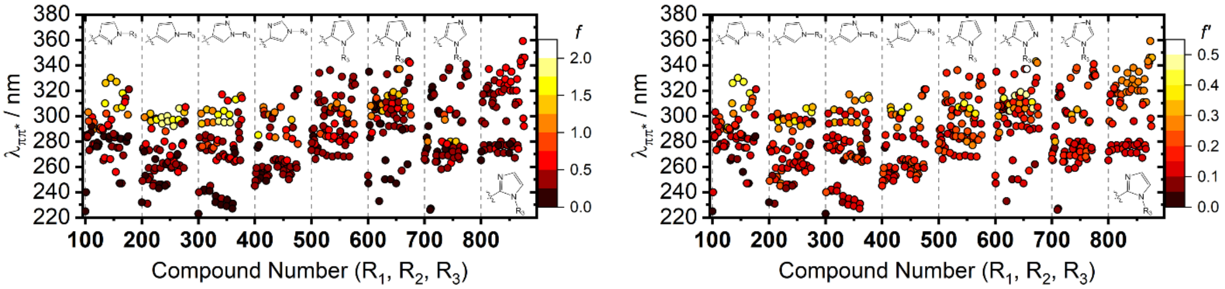


Figure S20. Comparison between (left) regular intensities (*f*) and (right) Ω_{AB}^{l}-weighted intensities associated with the *ππ** transition of the Z-isomer.

Under this definition, our results indicate that 38% of the Z-isomers display the lowest-energy *productive* ππ* transition in S_2, 29% in S_3, 24% in S_4 and 10% in S_5. Concerning its character, three main types could be identified and are discussed in the main text. The intensity associated with each type of *ππ** transition reveals that the third type has a much larger *f* (~1.0) than the first (*f* = 0.4) and second (*f* = 0.3) types. As a result, the range of regular intensities (*f*) is rather wide (see Figure S20-left), which might give the erroneous impression that some compounds present a much larger potential towards larger isomerization quantum yields. Once the intensity is weighted by Ω_{X3}^{2}, the resulting Ω_{AB}^{l}-weighted intensities are very similar (f'=0.2), and the range of values is much more homogeneous (see Figure S20-right).

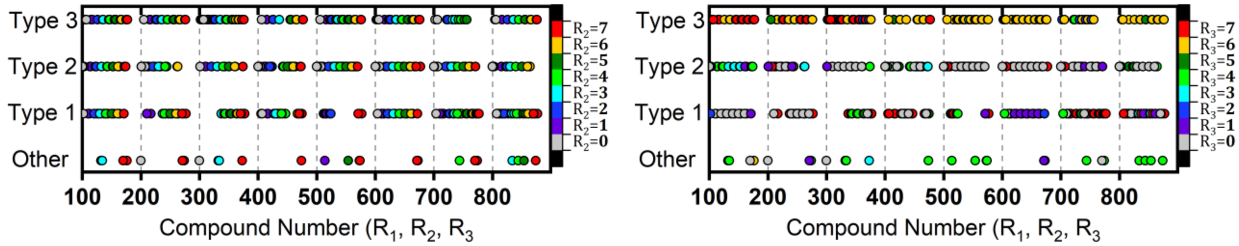


Figure S21. Representation of which transition type is observed for each compound. The color code indicates the (left) R_2 index or (right) R_3 index.

We observe (i) that the first type of transition dominates in compounds with R_1=**6-8**, and appears with similar probability for all R_2 and R_3, (ii) that the second type appears mostly in compounds with R_1=**4-5** and with R_2=R_3=H, and (iii) that the third is almost exclusive of compounds with electron-acceptors in R_3 (R_3=**5-7**) (see Figure S21). Finally, we have confirmed that the three transitions types are not an artefact of our computational method. For three representative compounds (C**112**, C**103** and C**115**), we confirmed that we obtain the same type of transition when using CC2 or CAM-B3LYP as an alternative to ωB97-XD (see Figures 7 and S21).

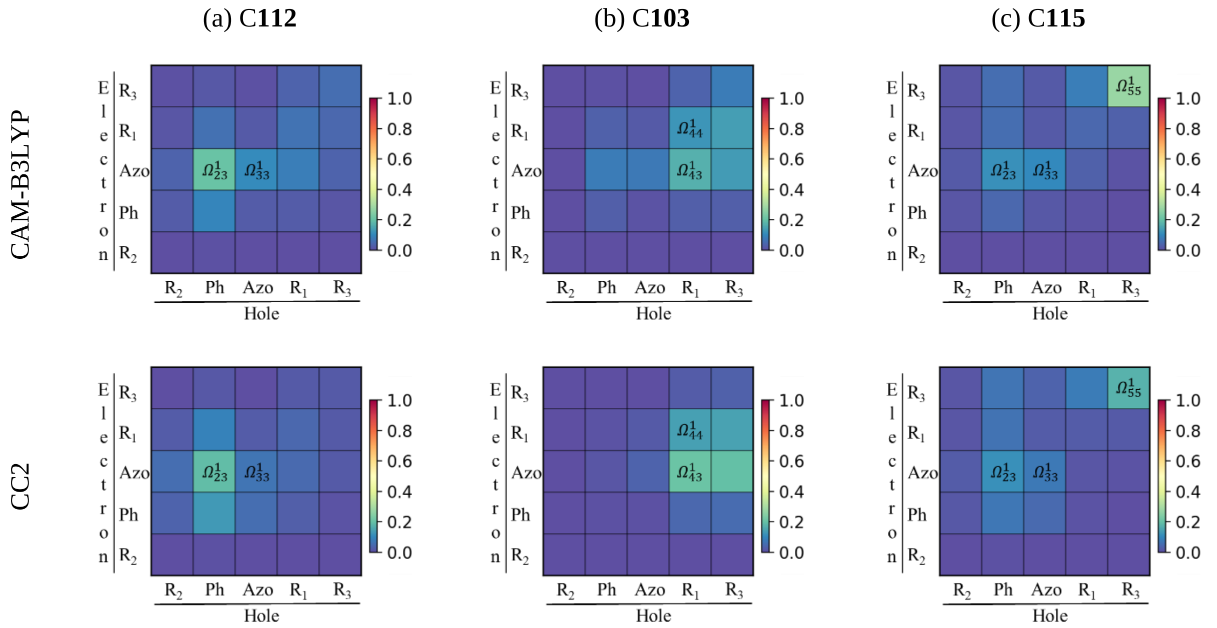
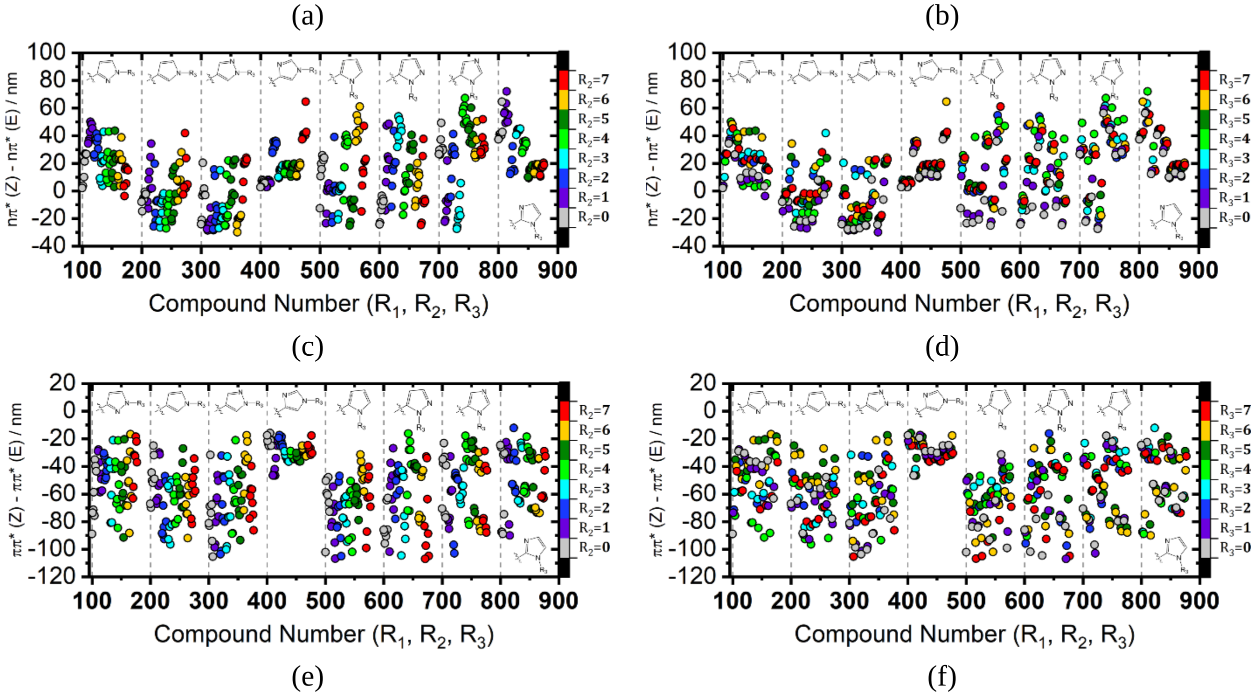


Figure S22. The three different patterns associated with the ππ* transition in the Z-isomer of the studied azoheteroarenes computed with CAM-B3LYP (top) and CC2 (below).

S5. Band Separation



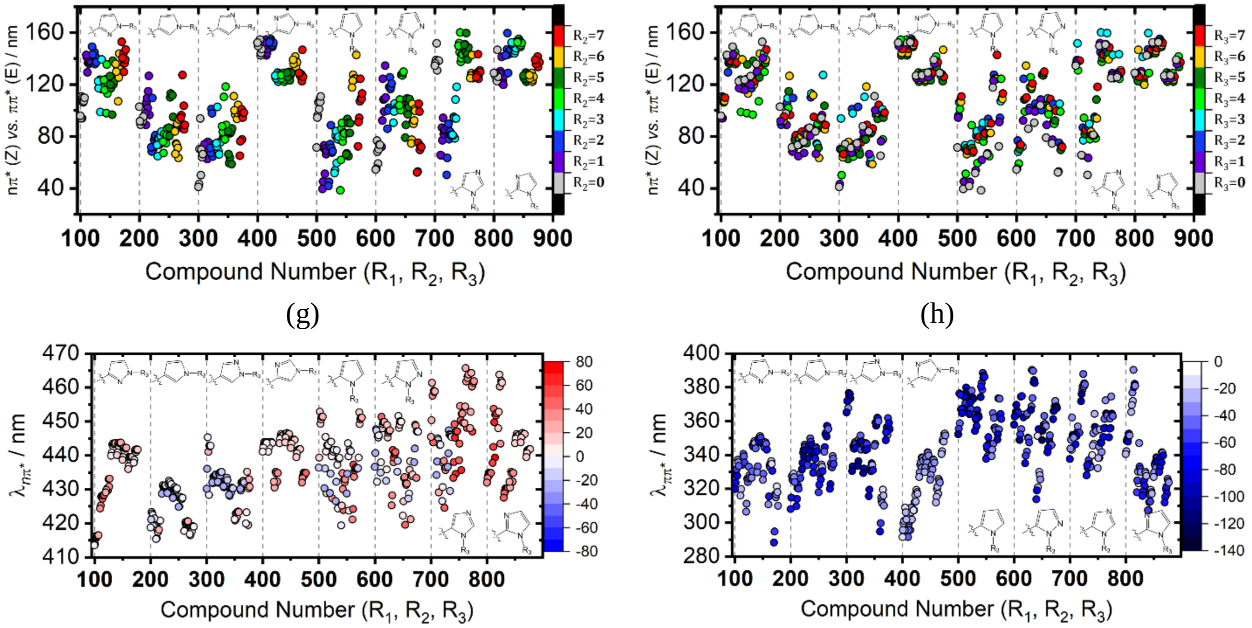


Figure S23. Separation between the (a-b) nπ* and (c-d) ππ* transitions in the E- and Z- isomers. (e-f) Separation between the nπ* band of the Z-isomer and the ππ* band of the E-isomer. The color code indicates the (left) R_2 index or the (right) R_3 index. Compounds are ordered using its three-digit identifier.

Table S5. Partial average of the absolute band separations depending on the fragments in $R_{1\text{-}3}$.

R_1	**nπ***	**ππ***	**nπ*(Z)−ππ*(E)**	R_2	**nπ***	**ππ***	**nπ*-ππ***	R_3	**nπ***	**ππ***	**nπ*-ππ***
1	25	50	127	**0**	22	48	117	**0**	26	59	102
2	14	60	89	**1**	29	53	118	**1**	23	53	111
3	17	60	85	**2**	21	55	110	**2**	22	49	106
4	23	51	136	**3**	19	65	96	**3**	23	56	109
5	16	71	75	**4**	18	67	95	**4**	21	58	109
6	26	55	110	**5**	18	64	100	**5**	21	64	99
7	19	68	84	**6**	24	60	103	**6**	20	68	98
8	37	50	137	**7**	24	53	103	**7**	21	57	108

S6. Comparison with previous works

As we mention in section 2.1 in the main text, few compounds that originate from the combination of the R_{1-3} in our fragment library have already been studied in previous works. The purpose of this section is to offer a clear comparison between the results obtained as a result from our screening, and those available in the literature. This should be also understood as an opportunity to compare and validate our computational approach.

A. Compound list

Table S6. List of compounds reported in the existing literature

Compound Number	corresponds to…	in reference:	Compound Number	corresponds to…	in reference:
C**100**	*3pzH*	5	C**501**	**2**	8
C**101**	*3pzMe**	5	C**501**	*2pyMe**	5
C**200**	*3pyH*	5	C**501**	**2**	6
C**201**	*3pyMe**	5	C**600**	*5pzH*	5
C**300**	*4pzH*	5	C**601**	*5pzMe**	5
C**301**	**8**	8	C**701**	**2**	[12]
C**301**	*4pzMe**	5	C**701**	**7**	[13]
C**301**	**8**	6	C**721**	**2b**	[12]
C**321**	*pzAzo ether 1*	[14]	C**800**	*Pai−H*	8
C**401**	**5****	2	C**800**	*Pai−H*	[15]
C**500**	*2pyH*	5	C**801**	*Pai−Me*	8
			C**801**	*Pai−Me*	[15]

*Two *vs.* one methyl substituent in the original *vs.* this paper. **C**401** is a conformer of the representation of **5** shown in Scheme 1 of ref. 2.

B. Wavelength of $n\pi^*$ and $\pi\pi^*$ transitions

Table S7. Comparison of the band position for compounds studied in this work that have been previously reported in the literature. Experimental values inside brackets.

Compound	$n\pi^*$ (E)	$\pi\pi^*$ (E)	$n\pi^*$ (Z)	$\pi\pi^*$ (Z)	Reference
C**100**	438 (425)[b]	288 (320)[b]	441 (422)[b]	259 (272)[b]	5
C**101**	439 (428)[b]	295 (325)[b]	435 (438)[b]	258 (286)[b]	5
C**200**	426 (~410)[b]	300 (363)[b]	408 (405)[b]	261 (311)[b]	5
C**201**	426 (~410)[b]	305 (345)[b]	399 (462)[b]	249 (300)[b]	5
C**300**	430 (417)[b]	295 (328)[b]	406 (403)[b]	243 (275)[b]	5
C**301**	430 (425)[b]	300 (335)[b]	400 (441)[b]	240 (296)[b]	5
C**301**	430 (426)	300 (327)	400 (397)	240 (262)	6
C**301**	430 (417)[b]	300 (328)[b]	400 (403)[b]	240 (275)[b]	8
C**321**	421 (-)	315 (342)[b]	406 (415)[b]	245 (276)[b]	[14]
C**401**	442	296 (336)[a]	445	274 (~310)[a]	[13]
C**500**	429 (~413)[b]	338 (385)[b]	404 (423)[b]	273 (333)[b]	5
C**501**	435 (422)	345 (356)	412 (408)	279 (300)	6
C**501**	435 (~430)[b]	345 (394)[b]	412 (479)[b]	279 (346)[b]	5
C**501**	435 (~413)[b]	345 (385)[b]	412 (423)[b]	279 (333)[b]	8
C**600**	435 (425)[b]	314 (341)[b]	418 (430)[b]	253 (289)[b]	5
C**601**	438 (435)[b]	319 (340)[b]	427 (451)[b]	269 (296)[b]	5
C**701**	441	334 (362)[a]	415	270 (312)[a]	[13]
C**800**	448 (435)[a]	318 (362)[a]	460	294	[15]
C**800**	448	318 (362)[a]	460 (445)[a]	294 (306)[a]	[16]
C**801**	453 (435)[a]	325 (363)[a]	467	293	[15]
C**801**	453	325 (363)[a]	467 (454)[a]	293 (329)[a]	[16]

[a]In toluene, [b]In acetonitrile

C. Band Separation

The ππ* band separation values reported for 4-pz (R_1=**3**) in this work lie mostly within −40 and −60 nm, slightly larger than those reported experimentally by Ravoo and coworkers[4] for a different set of compounds of the same family (*ca.* −30 nm). In turn, the nπ*(Z)−ππ*(E) band separation for the 4-pz (R_1=**3**) compounds included in this work lies mostly within 60 and 100 nm, slightly smaller than those reported by Ravoo and coworkers (*ca.* 100 nm).[4]

Table S8. Comparison between the band separation computed here (this work) and the computational and experimental estimation provided in reference 5.

Compound		nπ* (Z) – nπ* (E)			ππ* (Z) – ππ* (E)			nπ* (Z) – ππ* (E)		
Name	Index	**This Work**	**Ref. 5**		**This Work**	**Ref. 5**		**This Work**	**Ref. 5**	
			Comp.	**Exp.**		**Comp.**	**Exp.**		**Comp.**	**Exp.**
3pzH	C**100**	3	−13	−3	−29	−19	−48	153	143	102
*3pzMe**	C**101**	−4	−2	10	−37	−6	−39	140	117	113
3pyH	C**200**	−18	−9	−5	−39	−32	−52	108	90	42
*3pyMe**	C**201**	−27	22	51	−56	−9	−46	94	148	115
4pzH	C**300**	−24	−38	−14	−52	−36	−53	111	94	75
*4pzMe**	C**301**	−30	15	16	−60	−30	−39	100	141	106
2pyH	C**500**	−25	50	49	−65	−29	−48	66	135	85
*2pyMe**	C**501**	−22	16	10	−66	−38	−52	67	113	38
5pzH	C**600**	−17	6	5	−61	−23	−52	104	135	89
*5pzMe**	C**601**	−11	28	16	−50	−12	−44	108	150	111

*Two *vs.* one methyl substituent in the original *vs.* this paper.

Concerning the previous work of Fuchter and coworkers, in Table S7 we compare the experimental and computational results offered in ref. 5 with our results. We notice that the error in the nπ* separation is particularly large for C**201** and C**500**. In our work, we found these transitions at 399 and 404 nm, respectively, very similar to most compounds in Table S8. Experimentally, these are assigned at 462 and 479 nm, respectively, which implies that this transition is considerably red-shifted with respect to the other compounds studied therein, and in general, to other azoheteroarenes. For the remaining compounds, our results have a similar error than the computations included in ref. 5.

S7. Summary

A. Summarized Raw data

Table S9. Maximum, minimum and average values within each R_1 subset for the adiabatic energy difference (ΔH_{EZ}). The range is calculated as the difference between the maximum and minimum values. The “bare” entry refers to compounds with R_1=R_3=**0**

R_1	Max	Min	Range	Average	Bare
1	−15.1	−12.1	−2.9	−13.7	−12.9
2	−14.3	−11.4	−2.9	−12.8	−12.6
3	−13.6	−11.0	−2.6	−12.4	−11.5
4	−13.3	−11.3	−2.0	−12.2	−10.7
5	−13.9	−10.5	−3.5	−12.3	−11.1
6	−14.2	−11.1	−3.1	−12.8	−10.2
7	−14.1	−9.8	−4.3	−11.9	−10.8
8	−12.7	−9.7	−3.1	−11.2	−11.4

Table S10. Maximum, minimum and average values within each R_1 subset for the nπ* and $\pi\pi$* transitions of the **E-isomers**. The range is calculated as the difference between the maximum and minimum values. The “bare” entry refers to compounds with R_1=R_3=**0**

	nπ*					$\pi\pi$*				
R_1	λ_{MAX}	λ_{MIN}	Range	λ_{AVG}	λ_{BARE}	λ_{MAX}	λ_{MIN}	Range	λ_{AVG}	λ_{BARE}
1	444	424	19	437	438	364	288	76	335	288
2	446	414	32	426	426	359	300	59	334	300
3	440	416	23	428	430	362	295	67	336	295
4	447	428	19	440	429	353	292	61	326	292
5	445	416	30	434	435	387	338	49	366	338
6	453	421	32	442	437	390	314	76	355	314
7	449	420	29	438	448	387	328	59	358	328
8	466	433	33	453	441	388	318	70	355	318

Table S11. Maximum, minimum and average values within each R_1 subset for the nπ* and $\pi\pi$* transitions of the **Z-isomers**. The range is calculated as the difference between the maximum and minimum values. The “bare” entry refers to compounds with R_1=R_3=**0**

	nπ*					$\pi\pi$*				
R_1	λ_{MAX}	λ_{MIN}	Range	λ_{AVG}	λ_{BARE}	λ_{MAX}	λ_{MIN}	Range	λ_{AVG}	λ_{BARE}
1	486	435	51	463	441	332	258	74	300	259
2	455	399	56	422	408	314	247	67	285	261
3	460	400	60	419	406	318	240	78	285	243
4	499	444	56	463	444	335	274	61	299	274
5	481	404	77	440	404	353	273	80	305	273
6	499	418	81	466	418	372	253	119	315	253
7	482	409	72	440	409	341	264	77	303	266
8	513	460	53	491	460	359	293	66	321	294

Table S12. Maximum, minimum and average values within each R_1 subset for the three band separations explored. The range is calculated as the difference between the maximum and minimum values. The “bare” entry refers to compounds with R_1=R_3=**0**

	nπ* (Z) – nπ* (E)				ππ* (Z) – ππ* (E)				nπ* (Z) – ππ* (E)			
R_1	Max	Min	Average	Bare	Max	Min	Average	Bare	Max	Min	Average	Bare
1	−4	50	26	3	−64	−17	−35	−29	97	153	128	153
2	−27	34	−4	−18	−74	−14	−49	−39	62	127	88	108
3	−30	42	−10	−24	−78	−15	−51	−52	59	127	83	111
4	3	65	24	3	−37	−16	−28	−18	122	155	137	152
5	−25	64	7	−25	−93	−24	−60	−65	38	111	75	66
6	−17	61	24	−17	−84	3	−40	−61	72	143	111	104
7	−27	40	2	−27	−86	−31	−55	−62	50	118	82	81
8	12	70	38	12	−45	−12	−34	−24	120	160	136	142

B. Criteria used in compound ranking.

In Figure 10 of the main text, all compounds are ranked according to the properties of their nπ* and $\pi\pi$* transitions. For the nπ* transition, we highlight all compounds whose $\lambda^{n\pi*}$ lies in the top 25% (*i.e.* more red-shifted) for both the E- and Z-isomers, the associated intensity (f) is larger than 0.0, and the band separation ($\Delta\lambda^{n\pi*}$) is larger than 30 nm. For the $\pi\pi$* transition, we highlight the compounds whose $\lambda^{\pi\pi*}$ lies in the top 25% for both the E- and Z-isomers, has strong intensities (f and f' are top 25% for E- and Z-, respectively), and whose $\Delta\lambda^{\pi\pi*}$ is larger than 30 nm. These criteria are adopted with the aim at identifying 10-20 compounds for each of the two relevant transitions.

References

1. P.-F. Loos, N. Galland and D. Jacquemin, *J. Phys. Chem. Lett.*, 2018, **9**, 4646-4651.

2. J. Heuser and S. Höfener, *J. Chem. Theory Comput.*, 2018, DOI: 10.1021/acs.jctc.8b00369.

3. T. Cusati, G. Granucci, M. Persico and G. Spighi, *J. Chem. Phys.*, 2008, **128**, 194312.

4. L. Stricker, M. Böckmann, T. M. Kirse, N. L. Doltsinis and B. J. Ravoo, *Chem. Eur, J.*, 2018, **0**.

5. 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.

6. T.-T. Yin, Z.-X. Zhao and H.-X. Zhang, *New J. Chem.*, 2017, **41**, 1659-1669.

7. D. Jacquemin, E. A. Perpete, G. E. Scuseria, I. Ciofini and C. Adamo, *J. Chem. Theory Comput.*, 2007, **4**, 123-135.

8. 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.

9. S. Devi, M. Saraswat, S. Grewal and S. Venkataramani, *J. Org. Chem.*, 2018, **83**, 4307-4322.

10. J. Garcia-Amorós, S. Nonell and D. Velasco, *Chem. Commun. (Cambridge, U. K.)*, 2011, **47**, 4022-4024.

11. N. A. Simeth, S. Crespi, M. Fagnoni and B. König, *J. Am. Chem. Soc.*, 2018, **140**, 2940-2946.

12. C. Schütt, G. Heitmann, T. Wendler, B. Krahwinkel and R. Herges, *J. Org. Chem.*, 2016, **81**, 1206-1215.

13. T. Wendler, C. Schütt, C. Näther and R. Herges, *J. Org. Chem.*, 2012, **77**, 3284-3287.

14. Z.-y. Zhang, Y. He, Y. zhou, C. Yu, L. Han and T. Li, *Chem. Eur, J.*, 2019, **0**.

15. J. Otsuki, K. Suwa, K. Narutaki, C. Sinha, I. Yoshikawa and K. Araki, *J. Phys. Chem. A*, 2005, **109**, 8064-8069.

16. J. Otsuki, K. Suwa, K. K. Sarker and C. Sinha, *J. Phys. Chem. A*, 2007, **111**, 1403-1409.