Supplementary Information (SI). In silico strategy to design an efficient organic photoswitch based on Excited-State Cation Transfer

Laure de Thieulloy[a], Cédric Mongin[b], Isabelle Leray[b], Clément Guerrin[c], Stéphane Aloïse[c], Guy Buntinx[c], Aurélie Perrier[a,d,∗]

[a] *Chimie ParisTech, PSL Research University, CNRS, Institute of Chemistry for Life and Health Sciences (i-CLeHS), UMR 8060, F-75005 Paris, France.*
[b] *Université Paris-Saclay, ENS Paris-Saclay, CNRS, PPSM, 91190, Gif-sur-Yvette, France.*
[c] *Univ. Lille, CNRS, UMR 8516 - LASIRE - LAboratoire de Spectroscopie pour les Interactions, la Réactivité et l'Environnement F-59000 Lille, France.*
[d] *Université de Paris, F-75206 Paris, France.*

S1. Benchmark study: determination of the computational scheme to study the structural and optical properties of PyB derivatives

S1.1. PyB molecule: structural properties

At that stage, we are first interested in the **PyB** molecule depicted in Fig. S1 along with the PyB-Aza molecule. The atom numbering and the definition of key bond distances (d_i with i=1...11), angles (θ_j with j=1...4) and dihedral angles (ϕ_k with k=1...3) are also given.

We have first optimized the ground state (GS) using the $6-31+$G(d) basis set within an IEF-PCM solvation model in acetonitrile (ACN). We have tested seven different exchange-correlation (XC) functionals: a hybrid XC functional, PBE0 [1, 2], the same functional with the Grimme's dispersion [3], PBE0-GD3BJ, the B3LYP functional with GD3BJ dispersion, B3LYP-GD3BJ [4], two meta$-$GGA, M06 [5] and M06-2X [6], a range-separated hybrid, ωB97X-D [7] and a long-range separated hybrid functional, CAM-B3LYP [8].

Table S1 summarizes the key structural parameters obtained after the GS optimization and compares them with the data obtained from X-ray diffraction analysis [9].

First, we can observe that the structures obtained with PBE0 and PBE0-GD3BJ are quite similar, and we can conclude that the introduction of dispersion corrections does not modify the

---

∗Corresponding author
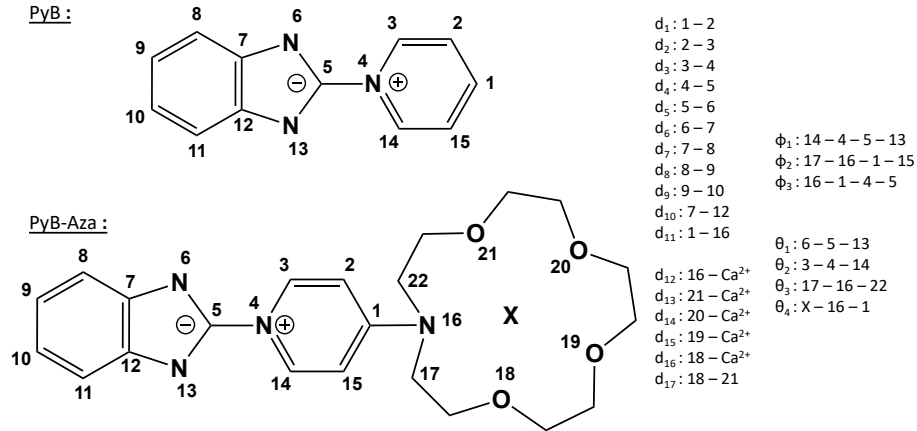*Email address:* `aurelie.perrier-pineau@u-paris.fr` (Aurélie Perrier)

Figure S1: **PyB** and **PyB-Aza** molecules: definition of the atom numbering and of the key structural parameters. **X** stands for the center of mass of the azacrown.

.

Table S1: **PyB** molecule: Selected structural parameters obtained with the $6-31+$G(d) basis set in ACN (IEF-PCM) with different exchange-correlation functionals. The distances, d, are in Å and the dihedral angles, ϕ, are in degrees (°). See Figure S1 for the definition of these parameters. The Absolute Percent Error (APE, in %) between the calculated and experimental values are given in brackets. The Mean APE (MAPE) is also given for the bond distances.

	Exp. [9]	PBE0	PBE0-GD3BJ	B3LYP-GD3BJ	M06	M062X	ωB97XD	CAM-B3LYP
d_1	1.366	1.394 *(2.0)*	1.393 *(2.0)*	1.398 *(2.4)*	1.393 *(1.9)*	1.394 *(2.0)*	1.393 *(2.0)*	1.392 *(1.9)*
d_2	1.373	1.380 *(0.5)*	1.379 *(0.5)*	1.382 *(0.7)*	1.378 *(0.4)*	1.381 *(0.6)*	1.379 *(0.5)*	1.378 *(0.4)*
d_3	1.333	1.353 *(1.5)*	1.353 *(1.5)*	1.363 *(2.2)*	1.354 *(1.6)*	1.352 *(1.4)*	1.352 *(1.4)*	1.351 *(1.4)*
d_4	1.450	1.433 *(1.2)*	1.431 *(1.3)*	1.423 *(1.9)*	1.440 *(0.7)*	1.445 *(0.4)*	1.444 *(0.4)*	1.444 *(0.5)*
d_5	1.355	1.332 *(1.7)*	1.332 *(1.7)*	1.340 *(1.1)*	1.331 *(1.8)*	1.329 *(1.9)*	1.329 *(1.9)*	1.328 *(2.0)*
d_6	1.356	1.372 *(1.2)*	1.372 *(1.2)*	1.372 *(1.2)*	1.374 *(1.3)*	1.375 *(1.4)*	1.377 *(1.6)*	1.376 *(1.5)*
d_7	1.417	1.404 *(0.9)*	1.403 *(0.9)*	1.408 *(0.6)*	1.402 *(1.0)*	1.406 *(0.8)*	1.404 *(0.9)*	1.403 *(1.0)*
d_8	1.359	1.388 *(2.1)*	1.388 *(2.1)*	1.387 *(2.0)*	1.387 *(2.0)*	1.386 *(2.0)*	1.387 *(2.1)*	1.385 *(1.9)*
d_9	1.391	1.414 *(1.7)*	1.414 *(1.7)*	1.422 *(2.3)*	1.413 *(1.6)*	1.417 *(1.9)*	1.415 *(1.7)*	1.414 *(1.7)*
d_{10}	1.423	1.429 *(0.5)*	1.430 *(0.5)*	1.442 *(1.3)*	1.426 *(0.2)*	1.426 *(0.2)*	1.423 *(0.0)*	1.422 *(0.0)*
MAPE (%)		*(1.3)*	*(1.3)*	*(1.6)*	*(1.3)*	*(1.3)*	*(1.2)*	*(1.2)*
ϕ_1	0.0	0.0	-0.0	0.0	0.0	0.0	-1.1	0.0

structural parameters under investigation. In addition, the structures obtained with M06 and M06-2X, on the one hand, and ωB97X-D and CAM-B3LYP, on the other hand, are similar. The Mean Absolute Percent Error (MAPE) values calculated for the selected bond distances show that the XC functionals provide satisfactory values ranging between 1.2% and 1.3%, except for B3LYP-GD3BJ that provides a larger MAPE (1.6 %). If we focus our attention on the Donor-Acceptor bond, d_4, we note that all the XC functionals slightly underestimate this parameter. The MAPE between calculated and experimental values are negligible with M06-2X (0.4%), ωB97X-D (0.4%), CAM-B3LYP (0.5%) and M06 (0.7%).

We expect the molecule to be planar as shown by the experimental value of the ϕ_1 torsion angle between the benzo-imidazole ring and the pyridinium moiety, which is the case for all the calculated structures.

Therefore, except for B3LYP-GD3BJ, all tested XC functionals are suitable to describe the GS geometry of PyB.

S1.2. PyB molecule: computed absorption and emission properties

We have considered the same set of XC functionals and, for each calculation scheme, we have carried out both the geometry optimization and the TD-DFT calculation with the same functional. Thus, 7 different TD calculation schemes have been set up (for instance: TD-PBE0/6-31+G(d)//PBE0/6-31+G(d)). First, we will discuss the results obtained with the Linear-Response (LR) - PCM model in ACN that is to say that we have calculated the vertical absorption energies, $\Delta Er^{vert-a}(LR, neq)$ in its non-equilibrium limit relyiong on the followin definition:

$$\Delta E^{vert-a}(LR, neq) = E^{ES}(R^{GS}, LR, neq) - E^{GS}(R^{GS}, eq) \tag{S1}$$

with $E^{ES}(R^{GS}, LR, neq)$ the energy of the excited state (ES) calculated at the ground state (GS) optimized structure R^{GS} and $E^{GS}(R^{GS}, eq)$ the energy of the GS corresponding to the same structure

From $\Delta E^{vert-a}(LR, neq)$, one determines the computed vertical absorption wavelength, λ_{abs}. This value, provided in Table S2, is defined as the excitation wavelength with the highest f value over the energy range considered. λ_{abs} is compared to the experimental maximum absorption λ_{max}. One should note that these values do not necessarily coincide, and a more refined computational

model should take into account the vibronic structure of the electronic absorption spectra. The deviation between the calculated and experimental excitation energy, ΔE, is given in eV. If ΔE is positive (respectively negative), the calculation scheme underestimates (resp. overestimates) the λ_{abs} value.

Table S2: **PyB** molecule: computed vertical absorption wavelength λ_{abs} (in nm) calculated with 7 different XC functionals within the LR-PCM acetonitrile solvation model. The energy difference between the calculated and the experimental excitation energies ΔE is given in eV. The orbitals involved in the electronic transition and the corresponding electronic states are also given.

	λ_{abs} (nm)	Electronic excitation	State	ΔE (eV)
PBE0	404	HOMO-1$\rightarrow$LUMO	S_2	-0.08
PBE0-GD3BJ	404	HOMO-1$\rightarrow$LUMO	S_2	-0.08
B3LYP-GD3BJ	427	HOMO-1$\rightarrow$LUMO	S_2	-0.25
M06	413	HOMO-1$\rightarrow$LUMO	S_2	-0.16
M06-2X	342	HOMO-1$\rightarrow$LUMO	S_1	0.47
ωB97X-D	330	HOMO-1$\rightarrow$LUMO	S_1	0.61
CAM-B3LYP	341	HOMO-1$\rightarrow$LUMO	S_1	0.49
Exp [10]	393			

Table S2 first shows that large deviations are obtained with M06-2X (ΔE=0.47 eV), CAM-B3LYP (0.49 eV) and ωB97X-D (0.61 eV). Indeed RSH functionals such as CAM-B3LYP and ωB97X-D are known to provide excitation energies and absorption maxima that are always higher in energy than the results obtained with global hybrids [11, 12]. These 3 XC functionals are thus not able to reproduce the absorption properties of **PyB** , and we have not considered them in the following.

B3LYP-GD3BJ provides a larger deviation value (ΔE =-0.25 eV) than the other XC functionals, respectively -0.08 eV for PBE0 and PBE0-GD3BJ, and -0.16 eV for M06. Since the performance of this functional to describe the structural properties of **PyB** was also less satisfying than the others, B3LYP-GD3BJ has not been taken into account in the following. The three remaining functionals, PBE0, PBE0-GD3BJ and M06, provide a LR-PCM λ_{abs} with an energy deviation within the acceptable TD-DFT error margin of 0.20 eV from the experiment and they are thus suitable candidates.

Table S3: **PyB** molecule: calculation of the maximum absorption wavelength λ_{abs} and vertical theoretical fluorescence wavelength λ_{em}: comparison between the LR-PCM and the cLR-PCM solvation models, and comparison with experiments. The 6-31+G(d) basis set has been used in the course of the geometry optimizations and TD-DFT calculations.

	Absorption				Emission			
	LR-PCM		cLR-PCM		LR-PCM		cLR-PCM	
	λ_{abs} (nm)	ΔE (eV)	λ_{abs} (nm)	ΔE (eV)	λ_{em} (nm)	ΔE (eV)	λ_{em} (nm)	ΔE (eV)
PBE0	404	-0.08	420	-0.20	530	0.53	732	-0.11
PBE0-GD3BJ	404	-0.08	420	-0.20	530	0.53	681	0.02
M06	413	-0.16	434	-0.30	540	0.49	673	0.04
Exp (nm) [10]	393				687			

For these 3 XC functionals, we have then compared the maximum absorption wavelength calculated within the LR and the cLR formalisms. The vertical absorption energies used to calculate λ_{abs} respectively correspond to $\Delta E^{vert-a}(LR, neq)$ (eq. S1) and $\Delta E^{vert-a}(cLR, neq)$. The results are gathered in Table S3. We can note that, for the 3 XC functionals, the cLR formalism leads to λ_{abs} values with a larger energy deviation compared to the experimental value.

In the same vein, the vertical theoretical fluorescence wavelength λ_{em} has been obtained by optimizing the first excited-state using the LR-PCM model in its equilibrium limit and then calculating the excitation energies, $\Delta E^{vert-f}(LR, neq)$ and $\Delta E^{vert-f}(cLR, neq)$. Either the LR or the corrected Linear-response (cLR)[13] solvation model is used to calculate the energy of the ES and GS at the optimized ES geometry R^{ES}:

$$\Delta E^{vert-f}(LR, neq) = E^{ES}(R^{ES}, LR, neq) - E^{GS}(R^{ES}, eq) \tag{S2}$$

$$\Delta E^{vert-f}(cLR, neq) = E^{ES}(R^{ES}, cLR, eq) - E^{GS}(R^{ES}, neq) \tag{S3}$$

Table S3 shows that the cLR model greatly improves the description of the emission wavelength. Indeed, from a deviation ΔE of 0.53 eV obtained with PBE0 and PBE0-GDB3J within the LR formalism, the cLR method gives a deviation of, respectively, -0.11 eV and 0.02 eV. A similar trend is observed with M06.

Therefore, in the following, the maximum absorption wavelength λ_{abs} will be calculated using the LR-PCM solvation model and vertical theoretical fluorescence wavelength λ_{em} will be obtained within the cLR-PCM formalism.

Then, the influence of the solvent polarity on the calculated maximum absorption wavelength has been studied and compared to the solvatochromic effect which was experimentally observed [10]. Table S4 reports the variation of the excitation energy corresponding to the maximum absorption wavelength, from ACN to toluene (TLN) (for instance -0.37 eV with PBE0), and from ACN to tetrahydrofuran (THF) (-0.22 eV with PBE0). We can note that the 3 XC functionals lead to the same trend: when the solvent polarity decreases, the maximum absorption energy also decreases, that is to say that the maximum absorption wavelength λ_{abs} is red-shifted. Furthermore, the 3 XC functionals reproduce the variation of the excitation energy from ACN→TLN with a deviation from the experimental results ranging between 0.00 eV and 0.02 eV. The experimental energy variation observed from ACN→THF (-0.21 eV) is also reproduced for the three functionals, PBE0 providing the best agreement with experiments. As a conclusion, the 3 XC functionals under consideration are able to describe the solvatochromic effect observed experimentally.

Table S4: PyB molecule: variation of the excitation energy (in eV) corresponding to the λ_{abs} wavelength, from ACN to TLN and from ACN to THF. The optical ΔE^{0-0} transition energy in ACN is also given. $\Delta\Delta E^{0-0}$ is the deviation between the calculated and experimental ΔE^{0-0} energies (in eV).

	Solvent effect (eV)		Optical properties (eV)	
	ACN→TLN	ACN→THF	ΔE^{0-0}	$\Delta\Delta E^{0-0}$
PBE0	-0.37	-0.22	2.12	-0.36
PBE0-GD3BJ	-0.36	-0.12	2.10	-0.38
M06	-0.38	-0.13	2.26	-0.22
Exp (eV) [10]	-0.36	-0.21	2.48	

The optical 0-0 transitions energies in ACN [14] have also been computed with PBE0, PBE0-GD3BJ and M06. The LR-PCM method has been considered for the absorption properties and the cLR formalism for the emission, namely, $\Delta E^{vert-a}(LR, neq)$ (Eq. S1) and $\Delta E^{vert-f}(cLR, neq)$ (Eq. S3) for the description of the ΔE^{0-0} energy.

From the optical properties, the deviation between the calculated and experimental ΔE^{0-0}

energy ranges between -0.38 eV for PBE0-GD3BJ and -0.22 eV for M06.

In conclusion, at that stage, among PBE0, PBE0-GD3BJ and M06, there is no XC functional outperforming the others and we cannot exclude one of these 3 functionals.

S1.3. Influence of the basis set.

To determine the influence of the basis set on the structural and optical properties, we have first optimized the structure of **PyB** using the 6-311+G(d,p) basis set with PBE0, PBE0-GD3BJ and M06, and compared the results with the ones obtained with 6-31+G(d).

If we compare the distances obtained with $6-31+G(d)$ (see Table S1) with the ones obtained with the larger basis set (Table S5), we can note that the optimization with the larger basis set results in a smaller MAPE and a slightly better description of the bond distance values. For instance, with PBE0, the MAPE is 1.3% with the 6-31+G(d) basis set and 0.4% with 6-311+G(d,p).

Table S5: **PyB** molecule: Selected structural parameters obtained with the 6-311+G(d,p) basis set basis set in ACN (IEF-PCM) with different exchange-correlation functionals. The Absolute Percent Error (APE, in %) between the calculated and experimental values are given in brackets. The Mean APE (MAPE) is also given. See Table S1 for more details.

	Experiment [9]	PBE0	PBE0-GD3BJ	M06
d_1	1.366	1.390 *(-1.8)*	1.390 *(-1.8)*	1.393 *(-1.9)*
d_2	1.373	1.376 *(-0.2)*	1.376 *(-0.2)*	1.378 *(-0.4)*
d_3	1.333	1.351 *(-1.4)*	1.351 *(-1.3)*	1.354 *(-1.6)*
d_4	1.450	1.433 *(1.2)*	1.431 *(1.4)*	1.440 *(0.7)*
d_5	1.355	1.329 *(1.9)*	1.329 *(1.9)*	1.331 *(1.8)*
d_6	1.356	1.370 *(-1.1)*	1.371 *(-1.1)*	1.374 *(-1.3)*
d_7	1.417	1.401 *(1.1)*	1.401 *(1.1)*	1.402 *(1.0)*
d_8	1.359	1.385 *(-1.9)*	1.385 *(-1.9)*	1.387 *(-2.0)*
d_9	1.391	1.411 *(-1.5)*	1.411 *(-1.4)*	1.413 *(-1.6)*
d_{10}	1.423	1.427 *(-0.3)*	1.427 *(-0.3)*	1.426 *(-0.2)*
MAPE		*(-0.4)*	*(-0.4)*	*(-0.6)*

For each structure, the maximum absorption wavelength has then been computed with the

same computational scheme. For instance, at the TD-PBE0/6-311+g(d,p)//PBE0/6-311+g(d,p) level (with a LR-PCM solvation scheme considered for both geometry optimization and TD-DFT calculations with ACN), the computed λ_{abs} value is 405 nm (see Table S6). The energy difference between the calculated and experimental λ_{abs} is thus -0.09 eV for PBE0 and PBE0-GD3BJ, and -0.16 eV for M06, and these results are similar to those obtained with the smaller basis set, $6-31+G(d)$. Therefore, using a larger basis set does not bring additional precision when describing the structural and absorption properties of the **PyB** system and we will continue the benchmark and the study using the $6-31+G(d)$.

Table S6: **PyB** molecule: computed vertical absorption wavelength λ_{abs} (in nm) computed with PBE0, PBE0-GD3BJ and M06 in combination with the 6-311+g(d,p) basis set (LR-PCM solvation model). ΔE (eV) is the energy difference between the calculated and the experimental excitation energies.

		PBE0		PBE0-GD3BJ		M06	
	Exp (nm) [10]	λ_{abs} (nm)	ΔE (eV)	λ_{abs} (nm)	ΔE (eV)	λ_{abs} (nm)	ΔE (eV)
ACN	393	405	-0.09	405	-0.09	414	-0.16
TLN	446	460	-0.08	460	-0.08	475	-0.17

S1.4. PyB-Aza molecule: structural and optical properties

To decide which functional will be used, extra tests have been carried out on the **PyB-Aza** molecule, a derivative of **PyB** with an azacrown connected to the pyridinium in the para position. The bond numbering is given in Figure S1.

In Table S7, the analysis of the structural parameters shows that the value of the d_4 bond distance, between the benzo-imidazole and the pyridinium moieties, decreases by 0.01 Å when the azacrown is added. Furthermore, this bond is slightly shorter after optimization with PBE0 and PBE0-GD3BJ than for the structure obtained with M06. This is in agreement with the value of the torsion angle ϕ_1: the structure of the **PyB** moiety is more planar and thus more conjugated with PBE0 and PBE0-GD3BJ compared to M06.

The position and the orientation of the azacrown with respect to the pyridinium group are described by the bond distance d_{11}, the angle θ_4 and the dihedral ϕ_3. Table S7 shows that the
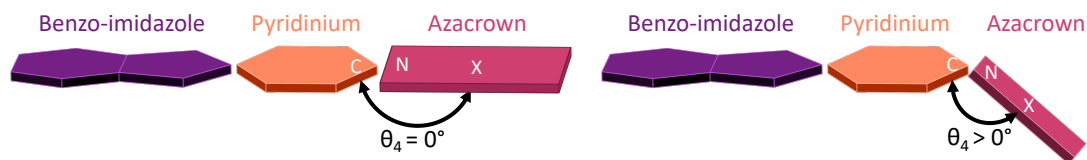
Figure S2: **PyB-Aza** molecule: Schematic representation of the θ_4 angle

.

choice of the XC functional used in the course of the geometry optimization has a small impact on the value of d_{11} while, from the value of the angle ϕ_3, one can note that M06 leads to a more planar organization. Finally, the angle θ_4 schematized in Figure S2 depicts the pendulum movement between the pyridinium moiety and the azacrown. Table S7 indicates that for the structure obtained with M06, θ_4 is slightly more planar than for the structures obtained with the XC functionals PBE0 and PBE0-GD3BJ ($\theta_4 = 114.5°$ and $113.7°$ respectively).

The impact of the XC functional on the absorption and emission properties has then been investigated. As shown by Table S7, PBE0, PBE0-GD3BJ and M06 lead to calculated λ_{abs} and λ_{em} wavelengths, as well as ΔE^{0-0} energies in good agreement with experimental features.

Table S7: **PyB-Aza** molecule: main structural parameters obtained in ACN (IEF-PCM), see Figure S1 for the definition of the parameters. The maximum absorption wavelength λ_{abs} (in nm, LR-PCM formalism), the fluorescence wavelength λ_{em} (in nm, cLR solvation scheme), the ΔE^{0-0} energy (in eV) and the deviations between the calculated and experimental values ΔE (in eV) are also given.

	Structural parameters						Optical properties (eV)					
	d_4 (Å)	d_{11} (Å)	ϕ_1 (°)	ϕ_2 (°)	ϕ_3 (°)	θ_4 (°)	λ_{abs} (nm)	ΔE (eV)	λ_{em} (nm)	ΔE (eV)	ΔE^{0-0}	$\Delta\Delta E^{0-0}$ (eV)
PBE0	1.429	1.349	0.65	6.95	-6.58	114.5	371	-0.11	479	0.26	2.85	0.05
PBE0-GD3BJ	1.427	1.348	0.23	6.65	-3.29	113.7	359	0.00	494	0.18	2.81	0.09
M06	1.435	1.351	1.01	5.28	0.65	116.2	367	-0.07	494	0.18	2.78	0.12
Exp [10]	–	–	–	–	–	–	359		533		2.90	

Finally, the three functionals nicely reproduce the experimental shift observed for the maximum absorption wavelength λ_{abs} when changing the solvent polarity (Table S8).

Table S8: **PyB-Aza** molecule: variation of the absorption energy (in eV) corresponding to the λ_{abs} wavelength, from ACN to TLN and from ACN to THF (LR-PCM solvation model).

	solvent effect (eV)	
	ACN→TLN	ACN→THF
PBE0	0.25	0.07
PBE0-GD3BJ	0.36	0.10
M06	0.36	0.10
Exp (eV)	0.21	0.15

S1.5. Conclusion: definition of the computational protocol to determine the structural and optical properties of PyB derivatives.

To conclude, we have performed a first benchmark using 7 different XC functionals, B3LYP-GD3BJ, PBE0, PBE0-GD3BJ, ωB97X-D, M06 and M06-2X, in combination with the $6-31+$G(d) basis set.

For the **PyB** and **PyB-Aza** molecules, the analysis of the bond distances shows that all XC functionals, except B3LYP-GD3BJ, provide satisfactory structural properties. From the calculation of the maximum absorption wavelength λ_{abs}, we have shown that three functionals provide absorption properties in better agreement with experiments, namely PBE0, PBE0-GD3BJ and M06. The 3 latter XC functionals also provide λ_{em} wavelengths (calculated within the corrected-Linear-Response formalism) in good agreement with experiments and are able to reproduce the solvatochromic effects experimentally observed.

Finally, another benchmark has been performed on the basis set. The structural and optical properties obtained with the 6-311+G(d,p) larger basis are similar with the one obtained with $6-31+$G(d).

Therefore, in the following, we have decided to carry out the geometry optimizations and the calculation of the absorption properties at the LR-PCM-(TD-)M06/$6-31+$G(d) level (energy $\Delta E^{vert-a}(LR, neq)$). To obtain the vertical fluorescence wavelengths λ_{em}, we will rely on the cLR-PCM-TD-M06/$6-31+$G(d) computational scheme (energy $\Delta E^{vert-f}(cLR, neq)$).

S2. Charge Transfer parameters

To analyse the CT character of the different excited states, we have used three different descriptors.

First, the variation of electronic density $\Delta\rho(\vec{r})$ is calculated on a grid of points using the cubegen utility provided by the Gaussian package:

$$\Delta\rho(\vec{r}) = \rho^{ES}(\vec{r}) - \rho^{GS}(\vec{r}) \tag{S4}$$

with $\rho^{GS}(\vec{r})$ and $\rho^{ES}(\vec{r})$ the electron density of the fundamental state and the considered excited-state respectively.

Secondly, the density derived indexes namely the D_{CT} index and the norm of the dipole moment corresponding to the CT observed upon excitation $\|\vec{\mu}_{CT}\|$ are computed using the DctViaCube [15, 16] program. In particular, the gain, $\rho^+(\vec{r})$, and the loss, $\rho^-(\vec{r})$, of electron density are described as :

$$\rho^+(\vec{r}) = \begin{pmatrix} \Delta\rho(\vec{r}) & if \ \Delta\rho > 0 \\ 0 & if \ \Delta\rho < 0 \end{pmatrix}; \ \rho^-(\vec{r}) = \begin{pmatrix} \Delta\rho(\vec{r}) & if \ \Delta\rho < 0 \\ 0 & if \ \Delta\rho > 0 \end{pmatrix} \tag{S5}$$

These densities are associated with barycenters and are defined as follows:

$$R_+ = \frac{\int r\rho^+(\vec{r})d\vec{r}}{\int \rho^+(\vec{r})d\vec{r}} = (x_+, y_+, z_+); R_- = \frac{\int r\rho^-(\vec{r})d\vec{r}}{\int \rho^-(\vec{r})d\vec{r}} = (x_-, y_-, z_-) \tag{S6}$$

The distance between these barycenters is known as the D_{CT} index:

$$D_{CT} = |R_+ - R_-| \tag{S7}$$

The norm of variation of dipole moment between the ground and excited-state is:

$$\mu_{CT} = \|\vec{\mu}_{CT}\| = D_{CT} \int \rho^+(\vec{r})d\vec{r} = -D_{CT} \int \rho^-(\vec{r})d\vec{r} = D_{CT}.q_{CT} \tag{S8}$$

with q_{CT} the amount of charge transferred.

Finally, the two centroids of charges associated to the positive and negative density regions such as:

$$C_+(\vec{r}) = A_+ e\left(-\frac{(x-x_+)^2}{2\sigma_{+x}^2} - \frac{(y-y_+)^2}{2\sigma_{+y}^2} - \frac{(z-z_+)^2}{2\sigma_{+z}^2}\right) \tag{S9}$$

$$C_-(\vec{r}) = A_+ e \left(-\frac{(x - x_-)^2}{2\sigma_{-x}^2} - \frac{(y - y_-)^2}{2\sigma_{-y}^2} - \frac{(z - z_-)^2}{2\sigma_{-z}^2} \right) \tag{S10}$$

where A_+ and A_- are the normalization factors, defined as:

$$A_+ = \frac{\int \rho_+(\vec{r}) d\vec{r}}{\int e \left(-\frac{(x-x_+)^2}{2\sigma_{+x}^2} - \frac{(y-y_+)^2}{2\sigma_{+y}^2} - \frac{(z-z_+)^2}{2\sigma_{+z}^2} \right) d\vec{r}} \tag{S11}$$

$$A_- = \frac{\int \rho_+(\vec{r}) d\vec{r}}{\int e \left(-\frac{(x-x_-)^2}{2\sigma_{-x}^2} - \frac{(y-y_-)^2}{2\sigma_{-y}^2} - \frac{(z-z_-)^2}{2\sigma_{-z}^2} \right) d\vec{r}} \tag{S12}$$

and σ is the root-mean square deviation along the three axis computed as :

$$\sigma_{a,j} = \sqrt{\frac{\Sigma_i \rho_a(r_i)(j_i - j_a)^2}{\Sigma_i \rho_a(r_i)}} \qquad a = +, - \text{ and } j = x, y, z \tag{S13}$$

S3. Determination of the ground-state complexation constant K.

We are first interested in the determination of the complexation constant K. Namely, the formation of the 1:1 complex corresponds to the equilibrium:

$$\text{Mol} + \text{M}^{\text{n}+} \underset{\longleftarrow}{\overset{\Delta G^*_{sol}}{\longrightarrow}} \text{Mol-M}^{\text{n}+} \tag{S14}$$
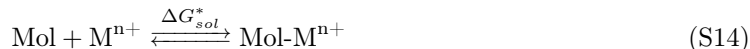
where Mol is a **PyB-Aza** derivative, M^{n+} a metal cation (for instance Ca^{2+}) and Mol-M^{n+} the 1:1 complex. At that stage, Mol, M^{n+} and Mol.M^{n+} are in their electronic ground state. The complexation constant K can be expressed as:

$$K = \frac{[\text{Mol-M}^{\text{n}+}]}{[\text{Mol}].[\text{M}^{\text{n}+}]} \tag{S15}$$

One generally computes the complexation constant K through:

$$K = -exp\left(\frac{\Delta G^*_{sol}}{RT}\right) \tag{S16}$$

This complexation constant is usually expressed in the form of a negative logarithm that constitutes the definition of pK:

$$pK = -logK = \frac{\Delta G^*_{sol}}{RTln(10)} \tag{S17}$$

To determine the complexation constant K and pK, several methods are available and we have tested three of them that are described below: (1) a Born-Haber (BH) thermodynamic cycle (TC), method TC-1; (2) the same BH TC with gas-phase energies computed with a more accurate level of theory, method TC-2[17]; (3) a direct model, method Direct[18]. Moreover, for each method, we consider two different continuum solvation models, IEFPCM [19, 20] and SMD [21].

These different approaches are used to determine the complexation constant of **PyB** and **PyB-Aza** molecules with Ca^{2+}. For **PyB-Aza**, there are two different complexation sites, near the imidazole bridge and in the azacrown. The complexation constant will thus be determined for these two sites.

S3.1. Method TC-1: Born-Haber Thermodynamic Cycle

The thermodynamic cycle (TC) commonly used for calculating complexation constants is shown in Figure S3.
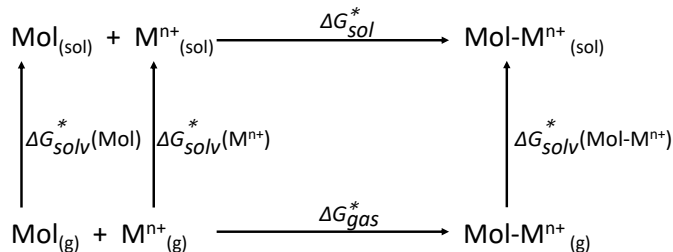
Figure S3: Born-Haber thermodynamic cycle used for computing the complexation constant pK values.

The complexation free energy (or Gibbs energy) $\Delta G^*_{sol}(TC-1)$ associated with this cycle is:

$$\Delta G^*_{sol}(TC-1) = \Delta G^*_{gas} + \Delta G^*_{solv}(\text{Mol-M}^{n+}) - \Delta G^*_{solv}(\text{Mol}) - \Delta G^*_{solv}(\text{M}^{n+}) \tag{S18}$$

where ΔG^*_{gas} is the gas-phase free energy of complexation:

$$\Delta G^*_{gas} = G^\circ_{gas}(\text{Mol-M}^{n+}) - G^\circ_{gas}(\text{Mol}) - G^\circ_{gas}(\text{M}^{n+}) + \Delta G^{0\to*} \tag{S19}$$

In equation S18, $\Delta G^*_{solv}(\text{Mol-M}^{n+})$, $\Delta G^*_{solv}(\text{Mol})$ and $\Delta G^*_{solv}(\text{M}^{n+})$ are the solvation free energies. $\Delta G^{0\to*}$ stands for the Gibbs free energy change for converting between standard states: the superscript "$\circ$" denotes that the thermochemical properties are computed with a standard state of 1 mol.L^{-1} and "$*$" corresponds to a standard state of 1 atm. Hence:

$$\Delta G^{0\to*} = RTln\left(\frac{RT}{P}\right) = 1.89 \text{ kcal.mol}^{-1} \tag{S20}$$

The solution phase and gas phase components of the solvation free energies are evaluated on the geometry optimized in their respective phases. The effect of the geometrical relaxation is thus included.

In equations S18 and S19, the gas phase energies, thermal corrections and solvation free energies are calculated at the same level of theory, namely with the M06/6-31+G(d) computational scheme.

S3.2. Method TC-2: TC with with gas-phase energies computed with a more accurate level of theory

Usually, continuum solvation models are parametrized to provide accurate free energy of solvation. However, the low levels of theory at which they are typically implemented may not be accurate enough to reproduce the solution phase free energy [18]. Therefore, another methodology based on the TC was proposed by Ho and coworkers[17, 18]. In this approach, one considers high-level

calculations in the gas phase to improve the accuracy of the resulting reaction energies. More in details, the geometries are still optimized in the gas phase at a lower level of theory (L), typically with a double zeta basis set, and high-level (H) single point calculations are then carried out.

This is resumed in the following equations:

$$\Delta G_{sol}^*(TC-2) = \Delta E_{sol}^L + \Delta G_{corr,sol}^{\circ L} + \Delta E_{gas}^H - \Delta E_{gas}^L + \Delta G^{\circ \to *} \tag{S21}$$

with

$$\Delta E_{sol}^L = E_{sol}^L(\text{Mol-M}^{n+}) - E_{sol}^L(\text{Mol}) - E_{sol}^L(\text{M}^{n+}) \tag{S22}$$

$$\Delta G_{corr,sol}^{\circ L} = G_{corr,sol}^{\circ}(\text{Mol-M}^{n+}) - G_{corr,sol}^{\circ}(\text{Mol}) - G_{corr,sol}^{\circ}(\text{M}^{n+}) \tag{S23}$$

$$\Delta E_{gas}^H = E_{gas}^H(\text{Mol-M}^{n+}) - E_{gas}^H(\text{Mol}) - E_{gas}^H(\text{M}^{n+}) \tag{S24}$$

and

$$\Delta E_{gas}^L = E_{gas}^L(\text{Mol-M}^{n+}) - E_{gas}^L(\text{Mol}) - E_{gas}^L(\text{M}^{n+}) \tag{S25}$$

$G_{corr,sol}^{\circ L}$ are the thermal corrections to reaction Gibbs free energy that are computed in the solution phase.

For PyB derivatives, the lower level of theory L relies on the M06 XC functional in combination with the 6-31+G(d) double zeta basis set, while high level single point calculations are performed at the M06/Def2TZVP level.

S3.3. Method "Direct"

In the direct approach, the solution phase reaction energies are directly computed within a dielectric continuum model as shown on Figure S4.

Moreover, high-level calculations are carried out on the solution phase optimized geometry in the presence of the solvent reaction to improve the accuracy of solution phase energies:

$$\Delta G_{sol}^*(Direct) = \Delta E_{sol}^H + \Delta G_{corr,sol}^{\circ L} + \Delta G^{\circ \to *} \tag{S26}$$

with

$$\Delta E_{sol}^H = E_{sol}^H(\text{Mol-M}^{n+}) - E_{sol}^H(\text{Mol}) - E_{sol}^H(\text{M}^{n+}) \tag{S27}$$

and $\Delta G_{corr,sol}^{\circ L}$ defined in Equation S23.

$$Mol_{(sol)} + M^{n+}_{(sol)} \xrightarrow{\Delta G^*_{sol}(Direct)} Mol\text{-}M^{n+}_{(sol)}$$

$\Delta G^*_{solv}Mol \quad \Delta G^*_{solv}M^{n+} \qquad \qquad \Delta G^*_{solv}Mol\text{-}M^{n+}$

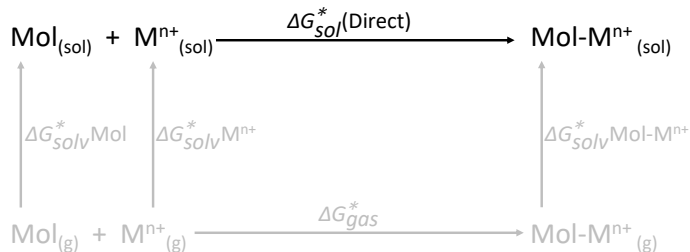$$Mol_{(g)} + M^{n+}_{(g)} \xrightarrow{\Delta G^*_{gas}} Mol\text{-}M^{n+}_{(g)}$$

Figure S4: Illustration of the "Direct" method.

S3.4. Results

Table S9 gathers the log K calculated using the methods TC-1, TC-2 and Direct. We have investigated the complexation constant of $M^{n+}= Ca^{2+}$ with the molecules **PyB** and **PyB-Aza**. For **PyB** , there is only one complexation site, the imidazole bridge, and the Mol-M^{n+} is denoted **PyB** -$Ca^{2+}@N_{13}$. For **PyB-Aza**, there are two possible complexation sites, on the imidazole bridge (**PyB-Aza**-$Ca^{2+}@N_{13}$), and in the azacrown (**PyB-Aza**-$Ca^{2+}@Aza$). Experimentally, the binding constants obtained from spectral titrations are only available for the imidazole moiety, namely for **PyB** -$Ca^{2+}@N_{13}$ and **PyB-Aza**-$Ca^{2+}@N_{13}$ [10]. In presence of an increasing concentration of $Ca(ClO_4)_2$ in solution, a complexation between the imidazole bridge of **PyB** and Ca^{2+} was shown to take place with a complexation constant of 3.4 log units and 6.8 log units for the system **PyB-Aza**-$Ca^{2+}@N_{13}$. Hence, we will first focus our attention on these two systems and discuss later the results obtained for the complexation in the azacrown. For the latter site, the complexation constant of Ca^{2+} in the azacrown usually ranges between 2 to 5 log units [22, 23, 24, 25].

A. Binding to the imidazole moiety.

Focusing on the method TC-1, one can see for the two models of solvation, PCM and SMD, that a negative complexation constant is obtained for **PyB** -$Ca^{2+}@N_{13}$ and **PyB-Aza**-$Ca^{2+}@N_{13}$. This indicates that the complexation is not and these results are thus inconsistent with the experimental observations.

The method TC-2 provides a positive value of log K. With the PCM model, the calculated values are overestimated when compared to experiments. For instance, the log K value computed for **PyB** -$Ca^{2+}@N_{13}$ is 5.3, that is to say that we obtain an overestimation of $+1.9$ log units compared to the

Table S9: Determination of the log K value for the **PyB** -Ca^{2+}@N_{13}, **PyB-Aza**-Ca^{2+}@N_{13}, **PyB-Aza**-Ca^{2+}@Aza systems. The methods TC-1, TC-2 and Direct are compared

| | **PyB** -Ca^{2+}@N_{13} | | **PyB-Aza**-Ca^{2+}@N_{13} | | **PyB-Aza**-Ca^{2+}@Aza | |
	PCM	SMD	PCM	SMD	PCM	SMD
Method TC-1	-2.8	-3.3	-1.0	-4.0	0.1	-7.2
Method TC-2	5.3	4.2	8.3	5.0	2.3	-8.2
Method Direct	9.8	7.1	11.4	8.4	19.0	6.0
Exp [10]	3.4 ± 0.1		6.8 ± 0.1		2 to 5	

experimental value. Nevertheless, the variation of the complexation constant Δ(log K) observed when going from **PyB** -Ca^{2+}@N_{13} to **PyB-Aza**-Ca^{2+}@N_{13} (+ 3.0 log units) is in agreement with the pK variation experimentally observed (+3.4 log units) [10].

With the SMD model, the calculated values of log K are in better agreement with experiments, there is an overestimation of +0.8 log units for **PyB** -Ca^{2+}@N_{13} and an underestimation of -1.4 log units for **PyB-Aza**-Ca^{2+}@N_{13}. However, from **PyB** -Ca^{2+}@N_{13} to **PyB-Aza**-Ca^{2+}@N_{13}, the variation Δ(log K) is +0.8 log units which is far from the experimental trend (+ 3.0 log units).

The use of the method "Direct" in combination with the PCM solvation model leads to a significant overestimation of the complexation constant of 6.4 log units for **PyB** -Ca^{2+}@N_{13} and 4.6 log units for **PyB-Aza**-Ca^{2+}@N_{13}. Altogether, the computed variation Δ(log K) is +4.7 log units, far from the experimental trend (+ 3.0 log units). Δ(log K) decreases strongly with the SMD model (+1.3 log units) but is still far from experiments. Within this approach, the calculated values of log K are overestimated compared to experimental ones, respectively +3.7 and +1.6 log units for **PyB** -Ca^{2+}@N_{13} and **PyB-Aza**-Ca^{2+}@N_{13}.

To conclude, no method can provide at the same time a quantitative prediction of (i) the log K values for PyB derivatives and (ii) the log K variation observed upon structural modification of the PyB skeleton. In the following, we will be interested in the design of PyB derivatives presenting an enhancement of the log K value compared to our reference molecule **PyB-Aza**. Therefore, we have decided to use the methodology that is able to provide the most accurate prediction of the log K variation, Δ(log K), namely the method TC-2 in combination with the PCM solvation model. Hereafter, this methodology will be referred to as TC-2/PCM. The error is in the range of +2 pK units. In other words, we do not intend to provide quantitative results with the TC-2/PCM method.

B. Complexation in the azacrown cavity.

Table S9 shows that the log K value computed for **PyB-Aza**-Ca^{2+}@Aza with the TC-2/PCM methodology is 2.3 log units that is to say in the range of the log K values that are usually observed for complexation in aza-15-crown-5 moieties [23, 24, 25, 22]. We remind here that this value was not determined experimentally for **PyB-Aza**. Moreover, with the TC-2/PCM method, there is a decrease of the computed log K value from **PyB-Aza**-Ca^{2+}@N_{13} to **PyB-Aza**-Ca^{2+}@Aza (-6.0 log units) and this is in agreement with the experimental observation: there is no the possibility for the cation to be trapped inside the azacrown cavity [10].

In Table S9, a unique **PyB-Aza**-Ca^{2+}@Aza structure is considered. A Ca^{2+} cation is added in the azacrown cavity of the **PyB-Aza** optimized structure and the protocols described above to calculate log K are considered. However, the azacrown cavity is well known to present a flexible architecture. Using the AMBER18 package [26, 27], we have decided to carry on a molecular dynamics simulation (MD) to take into account all the conformations and complexation possibilities of Ca^{2+} inside this cavity.

The simulation box corresponds to one **PyB-Aza**-Ca^{2+}@Aza molecule surrounded by 7436 solvent molecules of chloroform[1] within a 56 Å cubic box, two chloride ions are added for the neutrality of the system. The General Amber Force Field (GAFF) [27] is used to describe **PyB-Aza**-Ca^{2+}@Aza and the chloroform molecules are described with the default parameters implemented in AMBER. For **PyB-Aza**-Ca^{2+}@Aza, the atomic charges are obtained from the parametrization procedure applied in GAFF, i.e using the HF/6-31G(d) RESP charges. The particle mesh Ewald (PME) method [28, 29] is used to model the electrostatic interactions and a cut-off of 8 Å is considered. Three steps are then used successively. First (i) the minimization and heating process: using the steepest descent algorithm, the energy of the system is minimized then the system is then heated from 0 to 300 K during 100 ps in the NVT ensemble. Then, (ii) the system is equilibrated in the NPT ensemble for 2 ns with a 1 fs time step. The reference pressure (1 bar) and the temperature (300K) are incorporated with the Berendsen method [30]. Finally, (iii) the molecular dynamics simulation is performed for 10 ns in the NVT ensemble. To conclude, the simulation is

---

[1]Along the MD simulation carried out in acetonitrile with Ca^{2+} in the azacrown, the cation did not remain in the cavity but is captures either by the bulk or by the imidazole site. This is why we used the chloroform solvent.

analyzed with the CPPTRAJ software [31] included in the AMBER18 package, and 50 structures are randomly extracted along the trajectory.
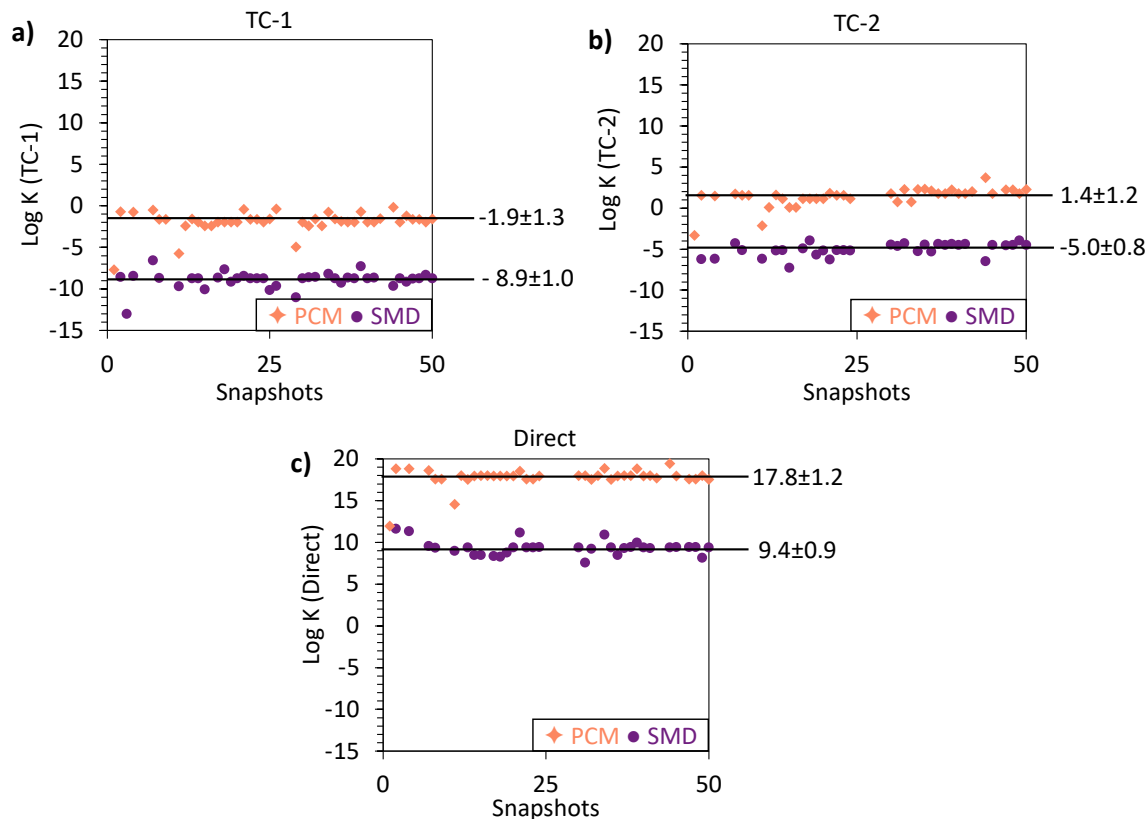

Figure S5: **PyB-Aza**-Ca^{2+}@Aza system: Determination of the log K value for **PyB-Aza**-Ca^{2}+@Aza for each structure extracted along the molecular dynamics simulation. The methods **a)** TC-1, **b)** TC-2 and **c)** Direct are compared either with the PCM (purple) or the SMD (orange) solvation models.

For each **PyB-Aza**-Ca^{2}+@Aza extracted structure, the following steps are applied: (1) Optimization and frequency calculation in ACN solvated phase with M06/6-31+G(d) method, followed by a single point calculation on the optimized structure using M06/def2TZVP method in ACN. (2) Starting from the newly optimized structure: optimization and frequency calculation in gas phase with M06/6-31+G(d) followed by a single point frequency calculation with the high-level method, M06/def2TZVP. (3) Then Ca^{2+} is removed from each structure and optimization and frequency

calculations of **PyB-Aza** are run in vacuum or ACN (PCM) using M06/6-31+G(d. Once the structures are optimized, a single-point calculation is performed on each structure in their respective phase with the M06/def2TVP method.

Then, the complexation constant is computed according to the three methods, TC-1, TC-2 and Direct previously described. Like for the site 1, the solvent is treated with both the PCM and SMD solvation models. The results obtained are gathered in the three graphics depicted in Figure S5.

Starting with the TC-1/PCM method, we remind here that the log K computed for a single structure (Table S9) is 0.1 units. This value thus underestimates the expected value of the complexation constant. From the MD simulation, we obtain an average value of the complexation constant equal to (-1.9 ± 1.3) units (See Figure S5a).

The TC-1/SMD method leads to a more negative value than the PCM solvation model with log K = -7.2 units for a single structure (Table S9) and (-8.9 ± 1.0) units for the snapshots extracted from the MD trajectory. These values indicate that Ca^{2+} is not able to complex in the azacrown cavity, which is not expected. Nevertheless, the results obtained from the MD simulations and the unique optimized structure are in good agreement with each other.

Next, the TC-2/PCM method provides a complexation constant of 2.3 units for the unique structure. As stated above, this value is in good agreement with the experiment. The complexation constant computed along the MD simulation is of (1.4±1.2) units therefore, the reference value of 2.3 units is within the confidence interval.

Concerning the TC-2/SMD method, the complexation constant in Table S9 is set to -8.2 units for a unique structure while the one resulting from the MD trajectory is equal to (-5.0±0.8) units. Although these results are close and thus consistent with each other, they do not reproduce the experiment (the log K value is negative) and therefore the TC-2/SMD method cannot be used to determine the complexation constant.

Finally the Direct/PCM method overestimates significantly the complexation constant with a computed result of 19.0 units (see Table S9).

In the same vein, the complexation constant determined from the MD structures greatly overestimates the expected experimental value, with a value of (17.8±1.2) units the value in Table S9 is

within the confident interval and both results are not able to reproduce the experimental trend. Focusing on the Direct/SMD method, the unique structure overestimates slightly the experimental value with a complexation constant of 6.0 units (see Table S9). The results obtained from the MD simulation is log K = (9.4±0.9) units, which overestimates the reference by more than 3 units.

To conclude, the results obtained with MD simulation provide the same trend as the one obtained when a single structure is optimized (see Table S9). The TC-2/PCM method correctly reproduces the experimental trend while the TC-1 method considerably underestimates the experimental value of the log K and, on the contrary, the direct method leads to an overestimation of several units.

Finally, one should note that the solvent used during the MD simulations is the chloroform and not the acetonitrile. This is due to the fact that during the MD simulations carried out in acetonitrile with Ca^{2+} in the azacrown, the cation did not remain in the cavity but was captured either by the bulk or by the imidazole site.
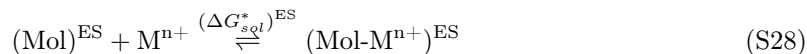
C. Conclusion: pros and cons of the TC-2 PCM method.

As a conclusion, the TC-2/PCM method is able to reproduce the variation of the log K value observed upon (1) the structural modification of the PyB moiety (**PyB** *versus* **PyB-Aza**) and (2) the modification of the binding site for **PyB-Aza** (imidazole bridge *versus* azacrown).

In this method, one needs to optimize the structures of the **PyB-Aza** and **PyB-Aza** -Ca^{2+} derivatives in the gas and solvent phases. These two optimizations steps thus increase the burden of the computational protocol and the structure optimization of **PyB-Aza** -Ca^{2+} in the gas phase can turn out to be a complicated calculation step. In that case, we will rely on the Direct/SMD method which only requires the geometry optimization to be carried out in the solution phase. Indeed, as shown in Table S9, for **PyB-Aza**, the variation of the log K value observed upon modification of the binding site is reproduced with the Direct/SMD method: the log K value decreases of -2.4 log units from **PyB-Aza**-Ca^{2+}@N_{13} to **PyB-Aza**-Ca^{2+}@Aza which is in qualitative agreement with experimental trends.

S4. Determination of the excited-state complexation constant K*

If we consider the excited state "ES", we have:

$$(\text{Mol})^{\text{ES}} + \text{M}^{\text{n}+} \overset{(\Delta G^*_{sol})^{\text{ES}}}{\rightleftharpoons} (\text{Mol-M}^{\text{n}+})^{\text{ES}} \tag{S28}$$

where (Mol) and (Mol-M$^{\text{n}+}$) are in their electronically excited state (ES). The complexation constant at the excited state, K* is:

$$K^* = \frac{[(\text{Mol-M}^{\text{n}+})^{\text{ES}}]}{[(\text{Mol})^{\text{ES}}].[\text{M}^{\text{n}+}]} \tag{S29}$$

and

$$K^* = -exp\left(\frac{(\Delta G^*_{sol})^{\text{ES}}}{RT}\right) \tag{S30}$$

$$pK^* = -logK^* = \frac{(\Delta G^*_{sol})^{\text{ES}}}{RTln(10)} \tag{S31}$$

The experimental approach for the determination of the excited state acid dissociation constants, requires a complex instrumental installation to take into account ultrafast kinetics process. To overcome this problem, Weller and Förster [32, 33], developed in the 60s a new mathematical approach for the determination of the photoacidity pK$_a^*$ for the photoacids. This approach is based on the absorption and emission wavelength which are easier to obtain, this model is known as the indirect Förster cycle.

We have transposed here the theory developed through the Förster cycle to estimate the photoacidity of photoacids to the determination of the K^* value. The adapted Förster cycle is depicted on Figure S6 [34, 35, 36].

For the absorption energies, from Figure S6, we have:

$$\Delta E^{vert-a}(\text{Mol}) + \Delta H = \Delta E^{vert-a}(\text{Mol-M}^{\text{n}+}) + \Delta H^* \tag{S32}$$

The enthalpy change is thus defined as:

$$\Delta H^* - \Delta H = \Delta E^{vert-a}_{\text{Mol}} - \Delta E^{vert-a}_{\text{Mol-M}^{\text{n}+}} \tag{S33}$$
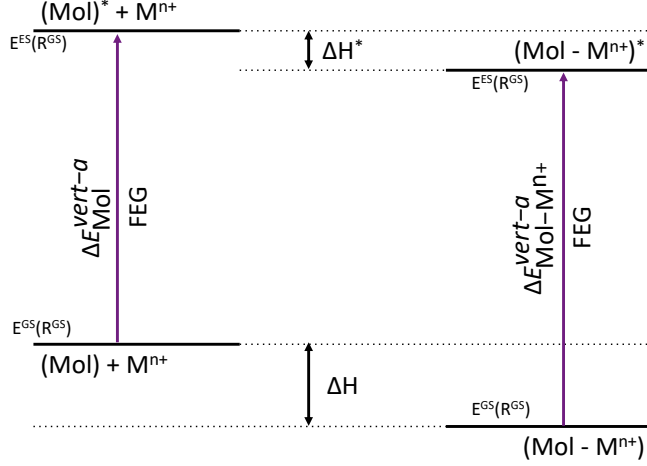
Figure S6: Förster's relationship of enthalpy changes to electronic transitions (Ref [34]). FEG is the Förster energy gap.

If one considers the solution is dilute enough to approximate ΔH to the standard value, ΔH°, and if the entropy does not change appreciably upon excitation, one can suppose that:

$$\Delta H^{\circ} \approx \Delta G^{\circ} = ln10.RT \times pK \tag{S34}$$

Therefore, using the Förster indirect cycle based on the absorption energies, Eq. S33 becomes:

$$\Delta pK^{*} = pK^{*} - pK = N_{\mathrm{A}} \times \frac{\Delta E_{\mathrm{Mol}}^{vert-a} - \Delta E_{\mathrm{Mol\text{-}M^{n+}}}^{vert-a}}{ln10.RT} = \frac{\Delta\Delta E^{vert-a}}{ln10.RT} \tag{S35}$$

With ΔpK^{*} the shift between pK^{*} and pK, $\Delta E_{\mathrm{Mol}}^{vert-a}$ and $\Delta E_{\mathrm{Mol\text{-}M^{n+}}}^{vert-a}$ are the vertical absorption energies, respectively of the molecule, **PyB-Aza** (Mol), and its complexed form, **PyB-Aza**-$Ca^{2+}@N_{13}$ (Mol-M^{n+}).

The ΔE^{vert-a} energies are computed within the cLR^2 solvation model[37, 38] :

$$\Delta E^{vert-a}(cLR^2, neq) = E^{ES}(R^{GS}, cLR, neq) + E^{ES}(R^{GS}, LR, neq) - E^{\omega_0} - E^{GS}(R^{GS}, eq) \tag{S36}$$

with R^{GS} the geometry of the Ground state and E^{ω_0} the transition energy computed with solvation orbitals but not accounting for the coupling between the medium and the transition [38].

S5. Spin-orbit coupling (SOC)

The Breit-Pauli operator is widely used for the expression of SOC operator such as:

$$\hat{H}_{SO}^{BP} = \frac{1}{2m_e^2 c^2} \sum_\alpha \sum_i \frac{Z_\alpha}{r_{i\alpha}^3} (\hat{r}_{i\alpha} \times \hat{p}_i)\hat{s}_i - \frac{1}{2m_e^2 c^2} \sum_i \sum_{j \neq i} \frac{1}{r_{ij}^3} (\hat{r}_{i\alpha} \times \hat{p}_i)(\hat{r}_{ij} \times \hat{p}_i).(\hat{s}_i + 2\hat{s}_j) \quad \text{(S37)}$$

$\hat{r}_{i\alpha} \times \hat{p}_i$ is the orbital angular momentum operator of electron i with respect to nucleus α while $\hat{r}_{ij} \times \hat{p}$ is the angular momentum of electron i with respect to electron j, and $\hat{s}_i$ is the spin-operator of electron i.

The first term of eq.S37, is the one-electron term that involves the spin-same-orbit coupling only while the second is the two-electron term.

The BP spin-orbit matrix element were calculated using PySOC [39], a program interface based on MolSOC [40] and implemented within MOMAP package [41, 42, 43]. PySOC uses the first term of eq.S37, that is to say the SOC element is implemented as a single-electron effective operator with an effective charge approximation such as:

$$\hat{H}_{SO}^{eff} = \frac{e^2}{2m_e^2 c^2} \sum_i^{N_{el}} \sum_\alpha^{N_A} Z_\alpha^{eff} \left(\frac{\hat{l}_{i\alpha}}{r_{i\alpha^3}} \right).\hat{s}_i \quad \text{(S38)}$$

where e is the electronic charge, m_e is the mass of electron, c is the speed of light, Z_α^{eff} is the effective nuclear change, $\hat{l} = \vec{r}_{ij} * \vec{p}$

Finally, SOC term between a particular singlet state S_j and a triplet manifold T_k was evaluated by summing the three matrix elements such as:

$$V_{S_j, T_k}^{SO} = \sqrt{|\langle S_j|\hat{H}_{SO}^{eff}|T_k^1 \rangle|^2 + |\langle S_j|\hat{H}_{SO}^{eff}|T_k^0 \rangle|^2 + |\langle S_j|\hat{H}_{SO}^{eff}|T_k^{-1} \rangle|^2} \quad \text{(S39)}$$

S6. Molecules 1-14: detailed analysis of the absorption properties

Group I: absorption properties

The excitation energies associated with their oscillator strengths and CT parameters ($\|\vec{\mu}_{CT}\|$ and Δq_I) are gathered in Table S10.

Without Ca^{2+}, the transition characterized by the largest oscillator strength is $S_2 \leftarrow S_0$ for all the systems. This transition is characterized by a HOMO-1→LUMO electronic transition. As seen in Figure S7 and for all the systems, the LUMO is strongly delocalised on the pyridinium moiety. The HOMO-1 is strongly delocalised on the benzo-imidazole and slightly delocalized on the pyridine, and the substituent is not involved in the electronic delocalisation for **1** and **4**. The substituents are involved in the electronic delocalisation for **3** and **6**. This orbital (HOMO-1) is similar to the HOMO of **PyB-Aza**.

Considering the energy of these molecular orbitals, one notes that the energy of the LUMO does not vary compared to the energy of the LUMO of **PyB-Aza**. If one considers molecule **1**, the addition of the OCH_3 substituent leads to a stabilization of the occupied molecular orbitals involved in the considered transition. Consequently, the energy difference between the **PyB-Aza** frontier orbitals increases by 0.07 eV with respect to **PyB-Aza**, which is consistent with the absorption wavelength blue-shift calculated for **1** ($\lambda_{abs} = 360$ nm). The same trend is obtained for molecule **3** for which $\lambda_{abs} = 361$ nm. Looking at molecule **4**, the HOMO-1 molecular orbital is close to the HOMO of **PyB-Aza**. Consequently, the absorption wavelength is close to the **PyB-Aza** one ($\lambda_{abs} = 367$ nm). Due to the additional electronic delocalization on the substituents of molecule **3**, the energy of the HOMO-1 increases; so, the absorption wavelength increases with respect to **PyB-Aza** ($\lambda_{abs} = 383$ nm).

When Ca^{2+} is added near the imidazole bridge (position N_{13}), one notes that for **1**-Ca^{2+}@N_{13}, the dihedral angle ϕ_1 decreases by 3.6° with respect to **PyB-Aza**-Ca^{2+}@N_{13}. The absorption wavelength is characterized by the $S_2 \leftarrow S_0$ transition, which corresponds to a HOMO-1→LUMO electronic excitation. The occupied molecular orbital is localized on the benzo-imidazole-pyridinium only. Thus, this orbital is stabilized compared to the one of **PyB-Aza**. Therefore, the gap between the PyB-Aza frontier orbitals increases, and the absorption wavelength is blue-shifted down to $\lambda_{abs} = 334$ nm. In the same vein, the main structural modification between **PyB-Aza**-Ca^{2+}@N_{13}
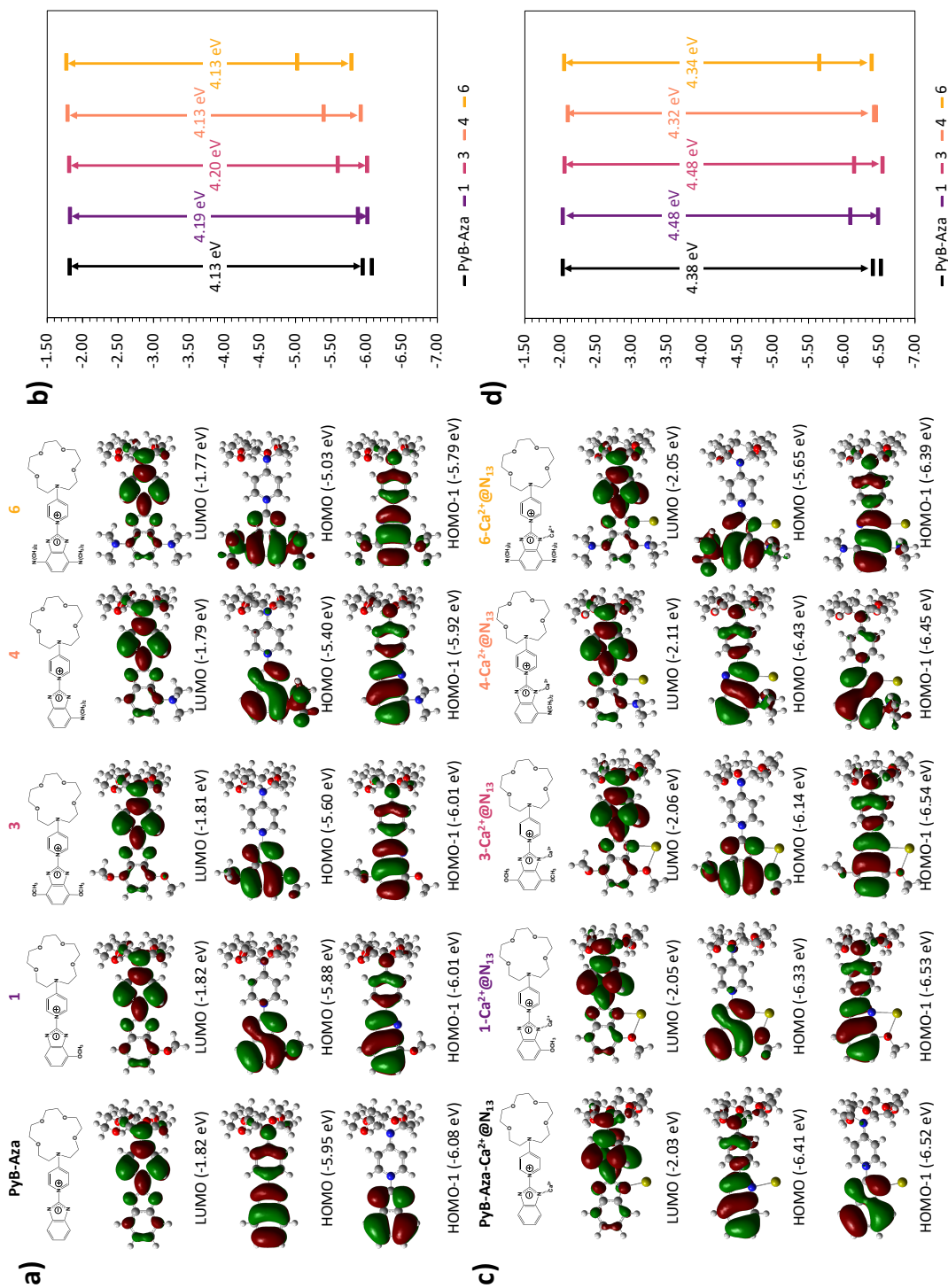
Figure S7: (a) and (c) Frontier orbitals of **PyB-Aza**, **1**, **3**, **4** and **6** without and in presence of Ca^{2+}@N_{13}. (b) and (d) Energy of the frontier orbitals.

S-27

Table S10: Absorption and photoejection properties: absorption wavelengths (λ, nm) and its associated oscillator strength (f) for the transitions $S_1 \leftarrow S_0$ and $S_2 \leftarrow S_0$ without and with $Ca^{2+}@N_{13}$. Photoejection descriptors, such as CT dipole moment ($\|\vec{\mu}_{CT}\|$, Debye), and MK charge variation from the benzo-imidazole group to the pyridinium moiety (Δq_I, e) are given. Complexation constant ($\log K_1$) associated to site 1 ($@N_{13}$) and the pK^* variation are also provided.

			PyB-Aza	Group I						Group II						Group III	
				1	2	3	4	5	6	7	8	9	10	11	12	13	14
No Ca^{2+}	$S_1 \leftarrow S_0$	λ	367	373	401	401	424	428	476	355	366	356	363	357	362	360	380
		(f)	0.661	0.154	0.525	*0.052*	0.116	0.450	*0.018*	0.517	0.618	0.606	0.668	0.62	0.695	0.700	0.154
	$S_2 \leftarrow S_0$	λ	354	360	351	361	367	357	383	349	342	347	340	347	340	349	365
		(f)	0.012	0.508	0.119	0.596	0.507	0.162	0.509	0.184	0.073	0.132	0.066	0.129	0.063	0.048	0.558
$Ca^{2+}@N_{13}$	$S_1 \leftarrow S_0$	λ	344	349	380	365	349	415	422	329	358	325	342	329	341	340	342
		(f)	0.556	0.164	0.384	*0.015*	0.628	0.323	*0.084*	0.660	0.644	0.614	0.533	0.671	0.548	0.741	0.697
		$\vec{\mu}_{CT}$	15.8	19.1	20.1	23.5	13.9	24.1	23.4	13.6	16.2	13.9	17.4	13.3	17.7	14.5	15.6
		Δq_I	-0.530	-0.606	-0.550	-0.707	-0.439	-0.649	-0.305	-0.450	-0.526	-0.456	-0.533	-0.444	-0.526	-0.411	-0.456
	$S_2 \leftarrow S_0$	λ	332	334	343	335	343	376	(S_3) 348	325	325	322	317	322	316	320	325
		(f)	0.029	0.495	0.035	*0.656*	0.074	0.035	(S_3) *0.540*	0.013	0.097	0.019	0.095	0.022	0.093	0.057	0.036
		$\vec{\mu}_{CT}$	21.3	17.9	18.6	13.7	21.3	20.1	(S_3) 15.26	21.6	20.2	21.5	20.5	21.2	20.7	20.4	21.7
		Δq_I	-0.758	-0.582	-0.411	-0.450	-0.691	-0.744	(S_3) -0.433	-0.725	-0.706	-0.714	-0.714	-0.708	-0.695	-0.642	0.671
	$\log K_1$		8.3	8.9	7.6	10.3	9.2	8.6	11.4	3.8	7.2	6.7	6.9	5.8	8.4	8.8	8.9
	ΔpK^*		-6.3	-5.0	-4.5	-6.3	-7.9	-3.2	-6.3	-5.3	-5	-6.6	-5	-5.7	-5.2	-5.5	-5.3

and 4-$Ca^{2+}@N_{13}$ is the decrease by 14.3° of ϕ_1. Concerning the electronic properties, the absorption wavelength is characterized by the $S_1 \leftarrow S_0$ transition, arising from a HOMO→LUMO electronic excitation. Being delocalised on the benzo-imidazole-pyridinium and the substituent, the HOMO has a larger electronic delocalisation than **PyB-Aza** which results in the destabilization of this orbital compared to the HOMO of **PyB-Aza**. In addition, the energy of the first vacant orbital slightly decreases in energy compared to **PyB-Aza**. Therefore, the gap between the **PyB-Aza** frontier orbitals decreases from 4.38 eV to 4.32 eV **PyB-Aza**-$Ca^2+@N_{13}$ and 4-$Ca^{2+}@N_{13}$, respectively. As a consequence, the absorption wavelength is red-shifted up to 349 nm for 4-Ca^{2+}.

Group II: Halogen

The excitation energies along with their oscillator strengths and charge transfer parameters ($\|\vec{\mu}_{CT}\|$ and Δq_I) ofin acetonitrile (IEF-PCM) solvent are gathered in Table S10.

At first glance, one can note that the introduction of a halogen atom does not lead to strong variations of the absorption properties: For instance, without Ca^{2+}, the maximum absorption wavelength decreased by less than 10 nm compared to **PyB-Aza**.

Considering the molecules **6**, **10**, and **12**, the variation of λ_{abs} is even smaller than 5 nm compared to **PyB-Aza**. The same trend is obtained when Ca^{2+} is added. Consequently, since the

evolution of the absorption properties is the same with the F, Cl, and Br substituents and due to the expected and well-known heavy-atom effect of the Bromine, the results related to the molecules **11** and **12** respectively will be exclusively discussed hereafter.

Considering the system without Ca^{2+}, one notes that the substitution in R_1 leads to a small blue shift of the maximum absorption wavelength, associated with a slight decrease of the oscillator strength. For instance, we obtain $\lambda_{abs} = 367$ nm ($f = 0.661$) and $\lambda_{abs} = 357$ nm ($f = 0.620$) for **PyB-Aza** and **11**, respectively. This variation corresponds to a blue shift of $\Delta E = 0.09$ eV only. Regarding the molecule **12**, this variation is even less significant with a blue shift of only $\Delta E = 0.05$ eV, with an absorption wavelength of $\lambda_1 = 362$ nm.

When Ca^{2+} is complexed to the imidazole bridge ($@N_{13}$), the absorption wavelength is blue-shifted with respect to the isolated molecule. This corresponds to a variation of 0.23 eV for **PyB-Aza**. As can be seen in Table S10, with respect to the isolated molecule, this behaviour is consistent with the modification of the structural parameters and the increase of the dihedral angle ϕ_1 by more than $35°$ (from 1 to $-37°$). The same trend is observed for the molecule **11** for which the absorption wavelength is also blue-shifted when Ca^{2+} is complexed to the imidazole bridge with a λ_{abs} computed value equal to 329 nm. Upon complexation with Ca^{2+}, for the molecule **11** the hypsochromic shift ($\Delta E = 0.30$ eV) is more significant than the one obtained for **PyB-Aza** ($\Delta E = 0.23$ eV).

For the molecule **12**, the addition of Ca^{2+} induces a smaller blue-shift than molecule **11** and similar to PyB-Aza with a deviation equal to 0.21 eV (Figure S8a and b).

Group III

Without Ca^{2+}, the absorption wavelength of the molecule **13** is similar to the one obtained for **1**, *i.e.*, $\lambda_{abs} = 360$ nm. It corresponds to the $S_1 \leftarrow S_0$ transition (HOMO-1→LUMO electronic excitation). The absorption properties (Table S10 as well as the molecular orbitals depicted in Figure S9 are similar and again the introduction of Br does not modify the absorption properties.

The same conclusion can be drawn when Ca^{2+} is added. For instance, **1**-$Ca^{2+}@N_{13}$ and **13**-$Ca^{2+}@N_{13}$ present the same electronic and absorption properties.

If we are now interested in the optical properties of **14**, one notes that they are like the one
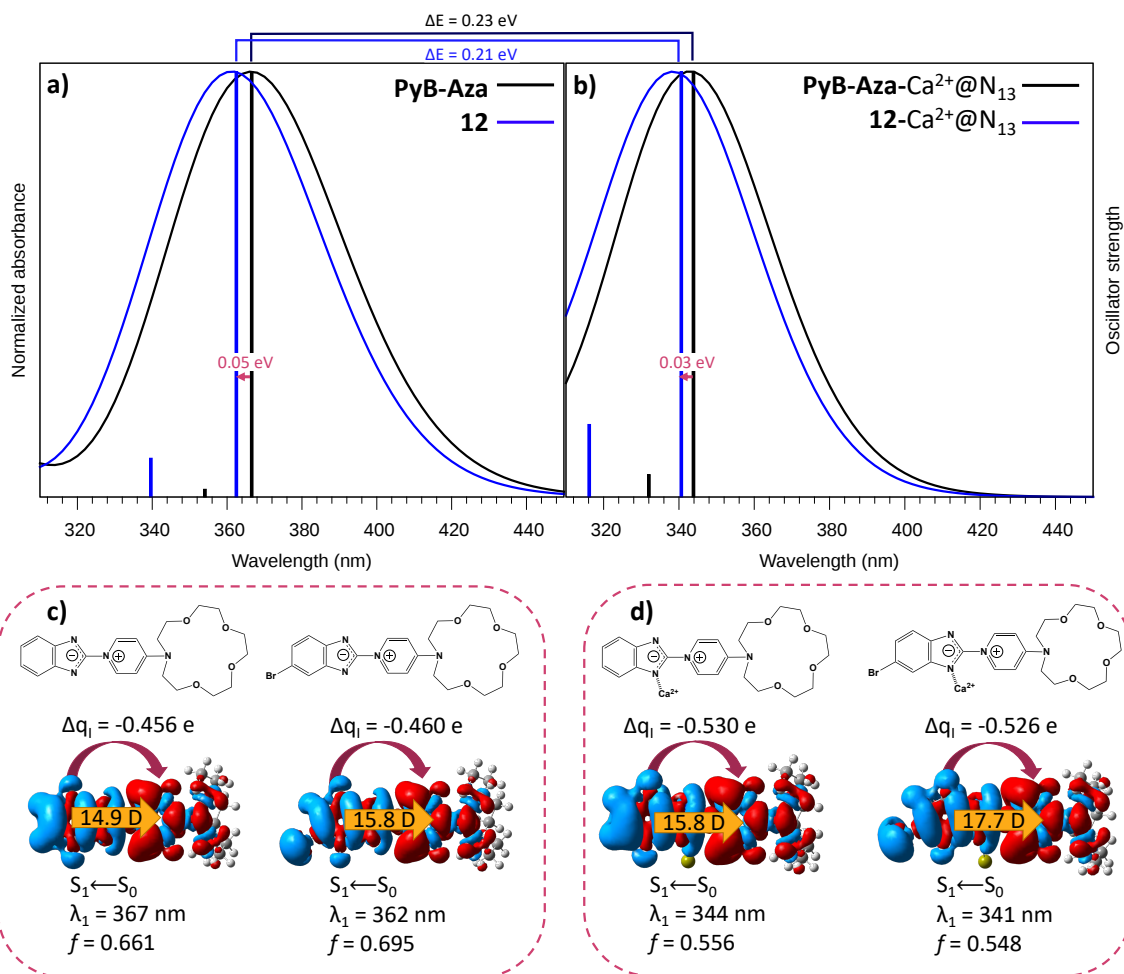
Figure S8: Comparison of the absorption spectra of **PyB-Aza** and **12** (a) without and (b) in the presence of Ca^{2+}. (c) and (d) CT parameters.
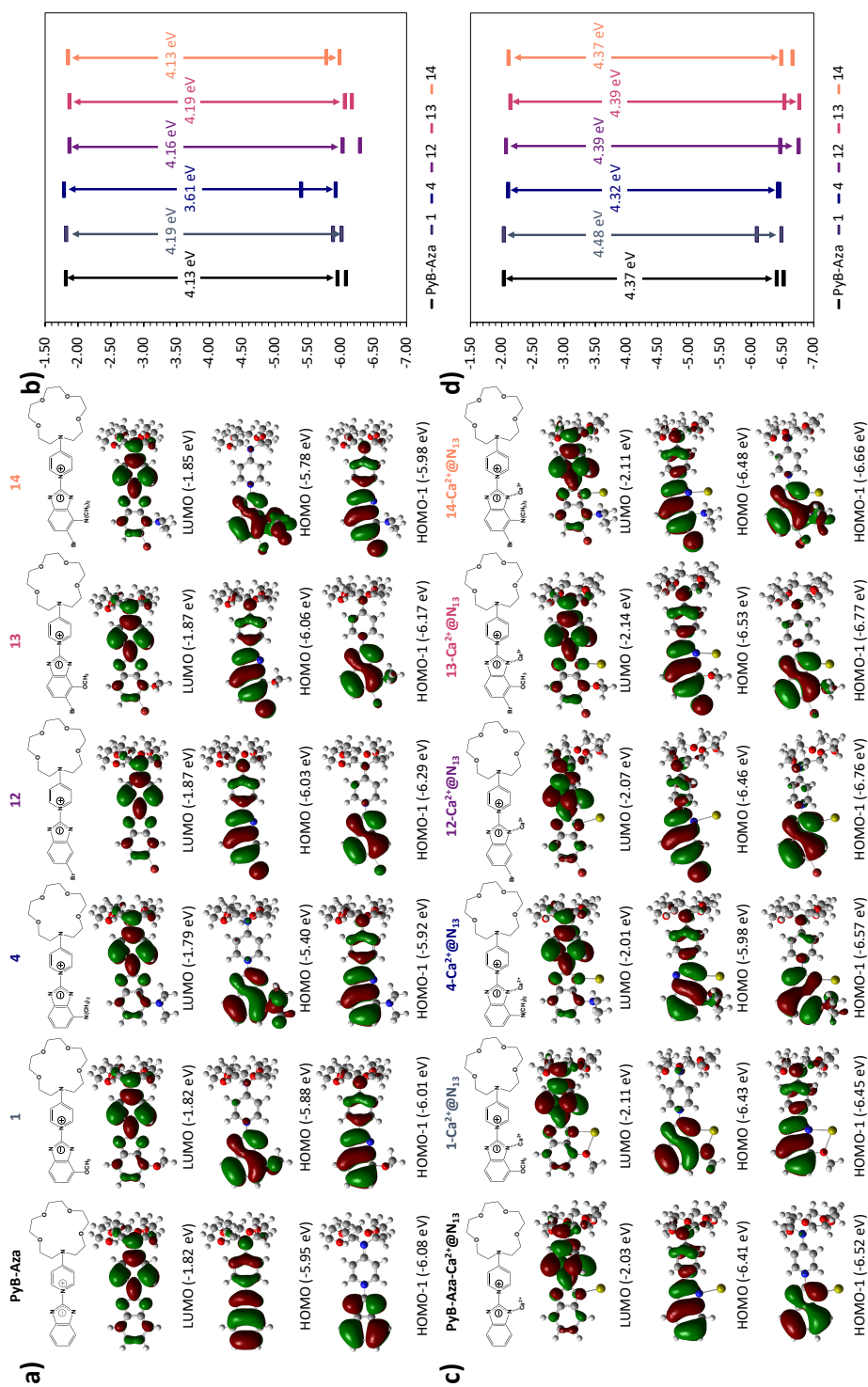
Figure S9: (a) and (c) Frontier orbitals of **PyB-Aza**, **1**, **4** , **12**, **13** and **14** without and in presence of $Ca^{2+}@N_{13}$. (b) and (d) Energy of the frontier orbitals.

obtained for **4**. Again, the Br atom has no influence on the absorption properties.

Concerning **14**-$Ca^{2+}@N_{13}$, the transition with the largest oscillator strength is $S_1 \leftarrow S_0$ (HOMO$\rightarrow$LUMO electronic excitation). The HOMO is similar to that of the **12**-$Ca^{2+}@N_{13}$ system, *i.e.* delocalised on the whole benzo-imidazole-pyridinium but not on the $N(CH_3)_2$ substituent, in contrast to the HOMO of **4**-$Ca^{2+}@N_{13}$ for which the substituent was involved in the delocalisation.

S7. Molecule 12: Frontier orbitals.

Concerning **PyB-Aza**-Ca^{2+}@N_{13}, the frontier orbitals are alike independently of the geometry. The HOMO is mainly delocalised on the benzo-imidazole, while the LUMO and SOMO are delocalised on the pyridinium.

Focusing on **12**, for the triplet state, we observe the SOMO-1 is localized on both the Br atom and the benzo-imidazole moiety. The SOMO is similar to the LUMO obtained for **PyB-Aza**-Ca^{2+}@N_{13}.
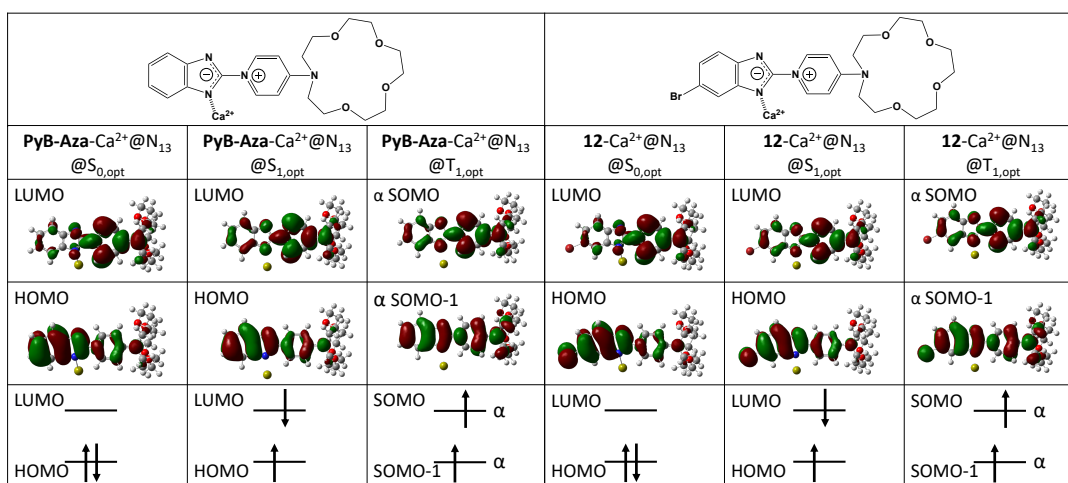
PyB-Aza-Ca^{2+}@N_{13} @$S_{0,opt}$	PyB-Aza-Ca^{2+}@N_{13} @$S_{1,opt}$	PyB-Aza-Ca^{2+}@N_{13} @$T_{1,opt}$	12-Ca^{2+}@N_{13} @$S_{0,opt}$	12-Ca^{2+}@N_{13} @$S_{1,opt}$	12-Ca^{2+}@N_{13} @$T_{1,opt}$
LUMO	LUMO	α SOMO	LUMO	LUMO	α SOMO
HOMO	HOMO	α SOMO-1	HOMO	HOMO	α SOMO-1

Figure S10: **PyB-Aza**-Ca^{2+}@N_{13} and **12**-Ca^{2+}@N_{13} molecules: Frontier molecular orbitals obtained from the equilibrium geometries at $S_{0,opt}$ and $S_{1,opt}$ optimized geometries. The α-SOMO and α-SOMO-1 are represented for the $T_{1,opt}$ optimized geometries. (isodensity values = 0.02 a.u).
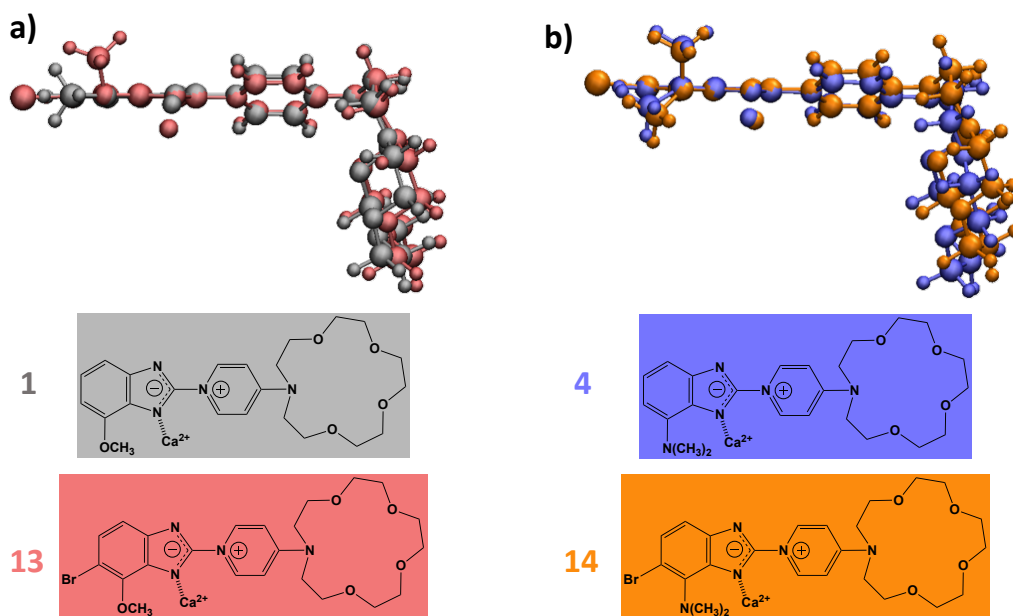
S8. Group III : extra information.

Figure S11: Group III: Superposition of the molecules **1** and **13** (left) and **4** and **14** (right).

References

[1] J. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys*, 1996, **105**, 9982–9985.

[2] C. Adamo and V. Barone, *J. Chem. Phys*, 1999, **110**, 6158–6170.

[3] S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

[4] C. Lee, W. Yang and R. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[5] Y. Zhao and D. G. Truhlar, *Theor Chem Account*, 2008, **120**, 215–241.

[6] Y. Zhao and D. Truhlar, *J. Chem. Phys*, 2006, **125**, 194101.

[7] J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

[8] T. Yanai, D. Tew and N. Handy, *Chem. Phys. Lett*, 393, **2004**, 51–57.

[9] E. Alcalde, I. Dinares, J. Elguero, J. P. Fayet, M. C. Vertut, C. Miravitlles and E. Molins, *J. Org. Chem.*, 1987, **52**, 5009–5015.

[10] S. Aloïse, Y. Ruan, I. Hamdi, A. K. Tiwari, G. Buntinx, C. Azarias, A. Perrier and I. Leray, *Phys. Chem. Chem. Phys.*, 2016, **18**, 15384–15393.

[11] D. Jacquemin, V. Wathelet, E. A. Perpete and C. Adamo, *J. Chem. Theory Comput.*, 2009, **5**, 2420–2435.

[12] T. Le Bahers, E. Bremond, I. Ciofini and C. Adamo, *Phys. Chem. Chem. Phys.*, 2014, **16**, 14435–14444.

[13] M. Caricato, B. Mennucci, J. Tomasi, F. Ingrosso, R. Cammi, S. Corni and G. Scalmani, *J. Chem. Phys.*, 2006, **124520**,.

[14] D. Jacquemin, A. Planchat, C. Adamo and B. Mennucci, *J. Chem. Theory Comput.*, 2012, **8**, 2359–2372.

[15] D. Jacquemin, T. Le Bahers, C. Adamo and I. Ciofini, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5383–5388.

[16] T. Le Bahers, C. Adamo and I. Ciofini, *J. Chem. Theory Comput.*, 2011, **7**, 2498–2506.

[17] J. Ho and M. Z. Ertem, *J. Phys. Chem. B*, 2016, **120**, 1319–1329.

[18] J. Ho, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2859–2868.

[19] J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

[20] C. A. Guido, G. Scalmani, B. Mennucci and D. Jacquemin, *J. Chem. Phys.*, 2017, **146**, 204106.

[21] A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.

[22] M. V. Rusalov, B. M. Uzhinov, M. V. Alfimov and S. P. Gromov, *Russ. Chem. Rev.*, 2011, **79**, 1099–1121.

[23] Y. Ruan, *Theses*, École normale supérieure de Cachan - ENS Cachan, 2012.

[24] E. N. Ushakov, V. B. Nazarov, O. A. Fedorova, S. P. Gromov, A. V. Chebun'kova, M. V. Alfimov and F. Barigelletti, *J. Phys. Org. Chem.*, 2003, **16**, 306–309.

[25] J.-F. Letard, R. Lapouyade and W. Rettig, *Pure Appl. Chem.*, 1993, **65**, 1705–1712.

[26] D. Case, D. Cerutti, T. I. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, D. York and P. Kollman, *University of California : San Francisco CA*, 2016.

[27] J. Wang, R. Wolf, J. Caldwell, P. Kollman and D. Case, *J.Comput. Chem.*, 2004, **25**, 1157–1174.

[28] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

[29] T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

[30] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

[31] D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

[32] A. Weller, *Prog. React. Kinet.*, 1961, **1**, 187–214.

[33] T. Förster, *Zeitschrift für Elektrochemie und angewandte physikalische Chemie*, 1950, **54**, 531–535.

[34] J. Ireland and P. Wyatt, *Advances in Physical Organic Chemistry*, Elsevier, 1976, vol. 12, pp. 131–221.

[35] C. Scharnagl and R. A. Raupp-Kossmann, *J. Phys. Chem. B*, 2004, **108**, 477–489.

[36] Y. Houari, D. Jacquemin and A. D. Laurent, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11875.

[37] P. M. Vérité, C. A. Guido and D. Jacquemin, *Phys. Chem. Chem. Phys.*, 2019, **21**, 2307–2317.

[38] C. A. Guido, A. Chrayteh, G. Scalmani, B. Mennucci and D. Jacquemin, *J. Chem. Theory Comput.*, 2021, **17**, 5155–5164.

[39] X. Gao, S. Bai, D. Fazzi, T. Niehaus, M. Barbatti and W. Thiel, *J. Chem. Theory Comput.*, 2017, **13**, 515–524.

[40] S. G. Chiodo and M. Leopoldini, *Computer Physics Communications*, 2014, **185**, 676–683.

[41] Z. Shuai, *Chin. J. Chem.*, 2020, **38**, 1223–1232.

[42] Z. Shuai and Q. Peng, *Phys. Rep.*, 2014, **537**, 123–156.

[43] Z. Shuai and Q. Peng, *National Science Review*, 2017, **4**, 224–239.