Supplementary Material (ESI) for Chemical Communications


# dsDNA-triggered energy transfer and lanthanide sensitization processes. Luminescent probing of specific A/T sequences

*Olalla Vázquez, Mateo I. Sánchez, José L. Mascareñas*, and M. Eugenio Vázquez**

## Supporting Information

*Dept. de Química Orgánica, Facultade de Química*
*Universidade de Santiago de Compostela*
*15782 Santiago de Compostela (Spain)*

*e-mail: eugenio.vazquez@usc.es, joseluis.mascarenas@usc.es*

# Abbreviations

| Abbreviation | Meaning |
|---|---|
| δ | Chemical Shift |
| $^{t}$Bu | Tert-butyl substituent |
| DIEA | *N, N*-diisopropylethylamine |
| DMF | Dimethylformamide |
| DMSO-$d_6$ | Deuterated dimethyl sulfoxide |
| eq. | Equivalents |
| ESI | Electrospray ionization |
| HPLC | High-Performance Liquid Chromatography |
| LC/MSD | Liquid Chromatograph Mass Spectrometer Detector |
| MWD | Multiple Wavelength Detector |
| PyOAP | 7-azabenzotriazol-1-yloxy-tris-(pyrrolidino)phosphonium hexafluorophosphate |
| RP | Reverse Phase |
| r.t. | Room temperature |
| TIS | Triisopropylsilane |

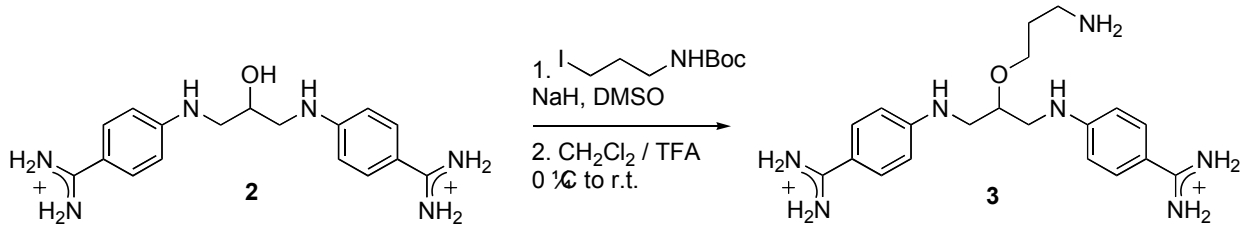
# Synthesis

All reagents were acquired from commercial sources: DMF and TFA were purchased from *Scharlau*, $CH_2Cl_2$ from *Panreac*, $CH_3CN$ from *Merck*, 1,4,7,10-tetraazacyclododecane-1,4,7-tris-*tert*-butyl acetate-10-acetic acid from *Macrocyclics* (Dallas, TX). The rest of reagents were acquired from *Sigma-Aldrich*.

Some reactions were followed by analytical RP-HPLC with an *Agilent 1100* series LC/MS using an *Eclipse XDB-$C_{18}$* (4.6 × 150 mm, 5 μm) analytical column. Standard conditions for analytical RP-HPLC consisted on an isocratic regime during the first 5 min, followed by a linear gradient from 25% to 95% of solvent B for 30 min at a flow rate of 1 mL/min (A: water with 0.1% TFA, B: methanol with 0.1% TFA). Compounds were detected by UV absorption at 220, 270, 304 and 330 nm. Final products were purified on a Büchi *Sepacore* preparative system consisting on a pump manager C-615 with two pump modules C-605 for binary solvent gradients, a fraction collector C-660, and UV Photometer C-635. Purification was made using reverse phase linear gradients of MeOH / $H_2O$ 0.1% TFA in 30 min with a flow rate of 30 mL/min, using a pre-packed preparative cartridge (150 × 40 mm) with reverse phase RP18 silica gel (Büchi order #54863). The fractions containing the products were freeze-dried, and their identity was confirmed by $ESI^+$-MS and NMR. Compounds were isolated as TFA salts.

NMR spectra show multiple signals corresponding to equilibria between protonation states in the imines. Extra signals corresponding to the minor component in the equilibria are not integrated and listed separately.

## *4,4'-(2-(3-aminopropoxy)propane-1,3-diyl)bis(azanediyl)dibenzimidamide (3)*



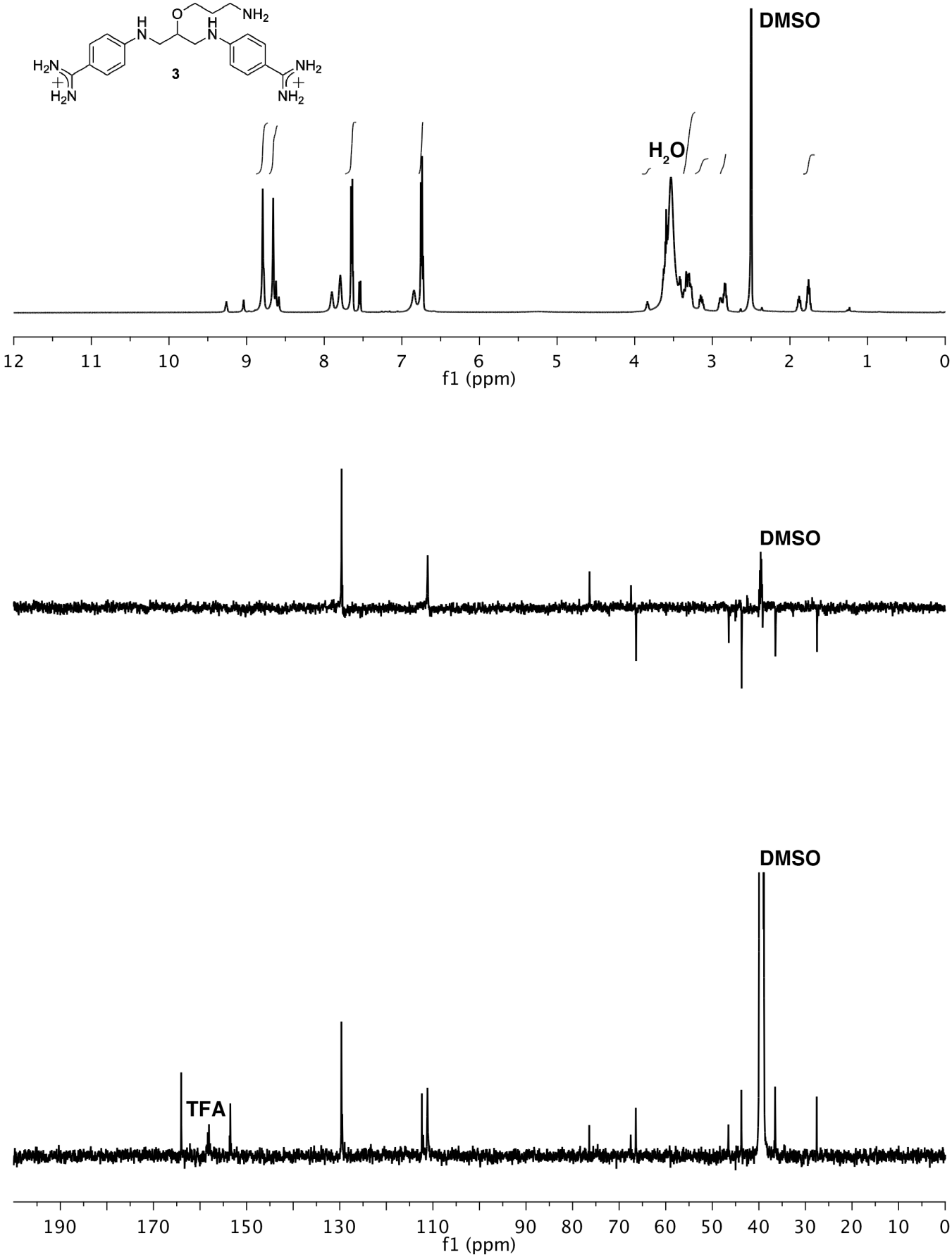

A 60% dispersion of sodium hydride (68 mg, 476 µmol, 4 equiv) was added to a solution of the bis-amino benzamidine **2** (66 mg, 119 µmol, 1 equiv) in dry DMSO (1.5 mL). After 30 min, 1-iodo-*tert*-butyl-2-aminopropylcarbamate (19 mg, 239 µmol, 2 equiv) was added in portions. The reaction mixture was stirred under Ar at rt for 6 h. The cude reaction was directly purified by preparative reverse-phase chromatography (*Büchi Sepacore*) (gradient: 15% B, 5 min; 15% → 95 % B, 30 min.). The combined fractions were concentrated and freeze-dried.

The isolated Boc-protected compound was dissolved in $CH_2Cl_2$ (1 mL) and cooled to 0 ºC. TFA (1 mL) was added dropwise and the resulting solution was stirred at 0 ºC for 1h and at room temperature for other 2 h. The solvent was removed under reduced pressure, and the residual TFA was removed by co-distillation with $CH_2Cl_2$. The residue was purified by preparative reverse-phase chromatography (Büchi Sepacore) (gradient: 0% B, 5 min; 0% → 50 % B, 30 min.). The freeze-dried solid was identified as the desired product **3** (13 mg, 39% overall yield for the 2-step process).

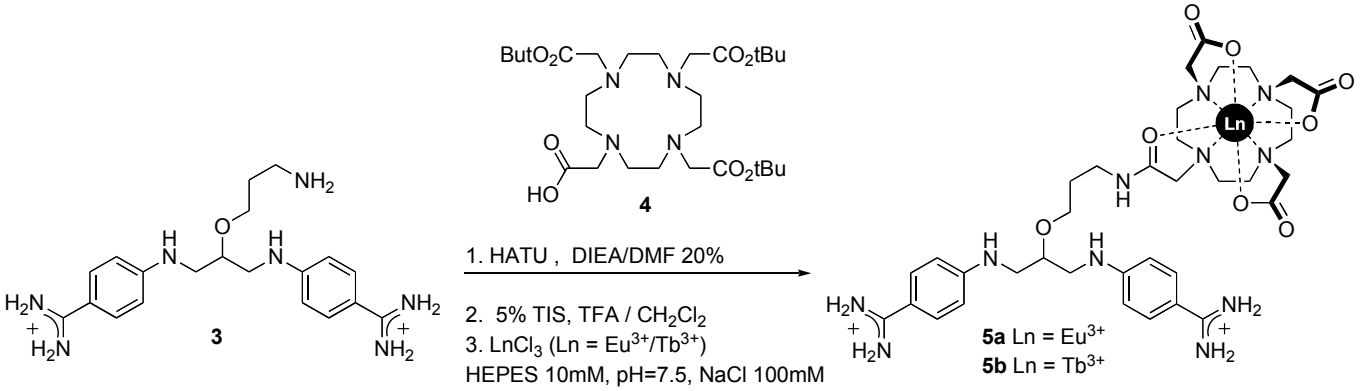
**$^1$H-NMR δ (DMSO-$d_6$):** 1.72-1.82 (m, 2H), 2.81-2.91 (m, 2H), 3.11-3.16 (m, 2H), 3.27-3.40 (m, 4H), 3.78-3.90 (m, 1H), 6.75 (d, *J* = 8.8 Hz, 4H), 7.65 (d, *J* = 8.8 Hz, 4H), 8.66 (s, 4H), 8.80 (s, 4H).

**$^{13}$C-NMR δ (DMSO-$d_6$):** 27.6 ($CH_2$), 36.5 ($CH_2$), 43.8 ($CH_2$), 66.4 ($CH_2$), 76.4 (CH), 111.2 (CH), 112.4 (C), 129.6 (CH), 153.5 (C), 158.4 (C, TFA), 164.1 (C). **ESI$^+$-MS:** $[M+H]^+$calcd. for $C_{20}H_{30}N_7O$ = 384,2506 found 384.2506.

Extra signals from minor tautomer and amine protons: 1.88, 6.83, 7.54, 7.79, 7.90, 8.58, 8.62, 9.03, 9.26


3
DMSO
H2O
DMSO
DMSO
TFA
f1 (ppm)
f1 (ppm)

## *Complexes 5a and 5b*




A solution of 1,4,7,10-Tetraazacyclododecane-1,4,7-tris-*tert*-butyl acetate-10-acetic acid (**4**, 52 mg, 92 µmol, 4 equiv) in DMF/DIEA 0.2M (460 µL, 92 µmol 4 equiv) was activated with HATU (35 mg, 92 µmol, 4 equiv) for 5 min, added over **3** (9 mg, 23 µmol, 1 equiv) in DMF/DIEA 0.2M (230 µL, 46 µmol 2 equiv) and stirred for 2 h. The consumption of the starting material and the appearance of the desired product was confirmed by analytical RP-HPLC. The solvent was removed and the residue was used in the next step without further purification. The protected *tert*-butyl compound was dissolved in $CH_2Cl_2$ (500 µL) and the mixture cooled down to 0 ºC. TFA (500 µL) was slowly added under stirring and the mixture was allowed to reach room temperature (for 3h). The mixture was concentrated, and 5% TIS in TFA (1mL) was added and the solution stirred for 48 h.[1] The residue was purify by preparative reverse-phase chromatography (*Büchi Sepacore*) (gradient: 8% B, 5 min; 8% → 40 % B, 30 min.). The appropriated fractions were collected, concentrated and freeze-dried to obtain the desired conjugate (19 mg, 85 % overall yield for the 2-step process).

**$^1$H-NMR δ (DMSO-$d_6$/$CD_2Cl_2$):** 1.72-1.82 (m, 2H), 2.68 (s, 3H), 2.72 (s, 3H), 2.84 (dd, *J* = 13.1, *J* = 6.61 Hz, 2H), 2.88 (s, 4H), 3.05 (s, 2H), 3.11-3.16 (m, 4H), 3.27- 3.36 (m, 4H), 3.39 (s, 2H), 3.58-3.64 (m, 8H), 3.80-4.50 (b, 1H overlays by water signal), 6.75 (d, *J* = 8.9, 4H) 7.65 (d, *J* = 8.9, 4H), 8.73 (s, 4H), 8.80 (s, 4H).

**$^{13}$C-NMR δ (DMSO- $d_6$/$CD_2Cl_2$):** 27.6 ($CH_2$), 28.6 ($CH_2$), 28.8 ($CH_2$), 30.7 ($CH_2$), 36.6 ($CH_2$), 41.7 ($CH_2$), 43.8 ($CH_2$), 44.9 ($CH_2$), 45.0 ($CH_2$), 47.5 ($CH_2$), 66.5 ($CH_2$), 76.5 (CH), 111.1 (CH), 112.4 (C), 129.7 (CH), 153.5 (C), 158.0 (C, TFA), 164.0 (C), 169.5 (C). **ESI$^+$-MS:** $[M+H+Na]^{2+}$calcd. for $C_{36}H_{56}N_{11}NaO_8$ = 396,7100 found 396,7112.

Extra signals from minor tautomer and amine protons: 6.86, 7.83, 7.95, 8.51. Extra C signals from minor tautomer: 50.0 ($CH_2$), 115.6 (C), 117.8 (C), 151.0 (C), 162.0 (CH)

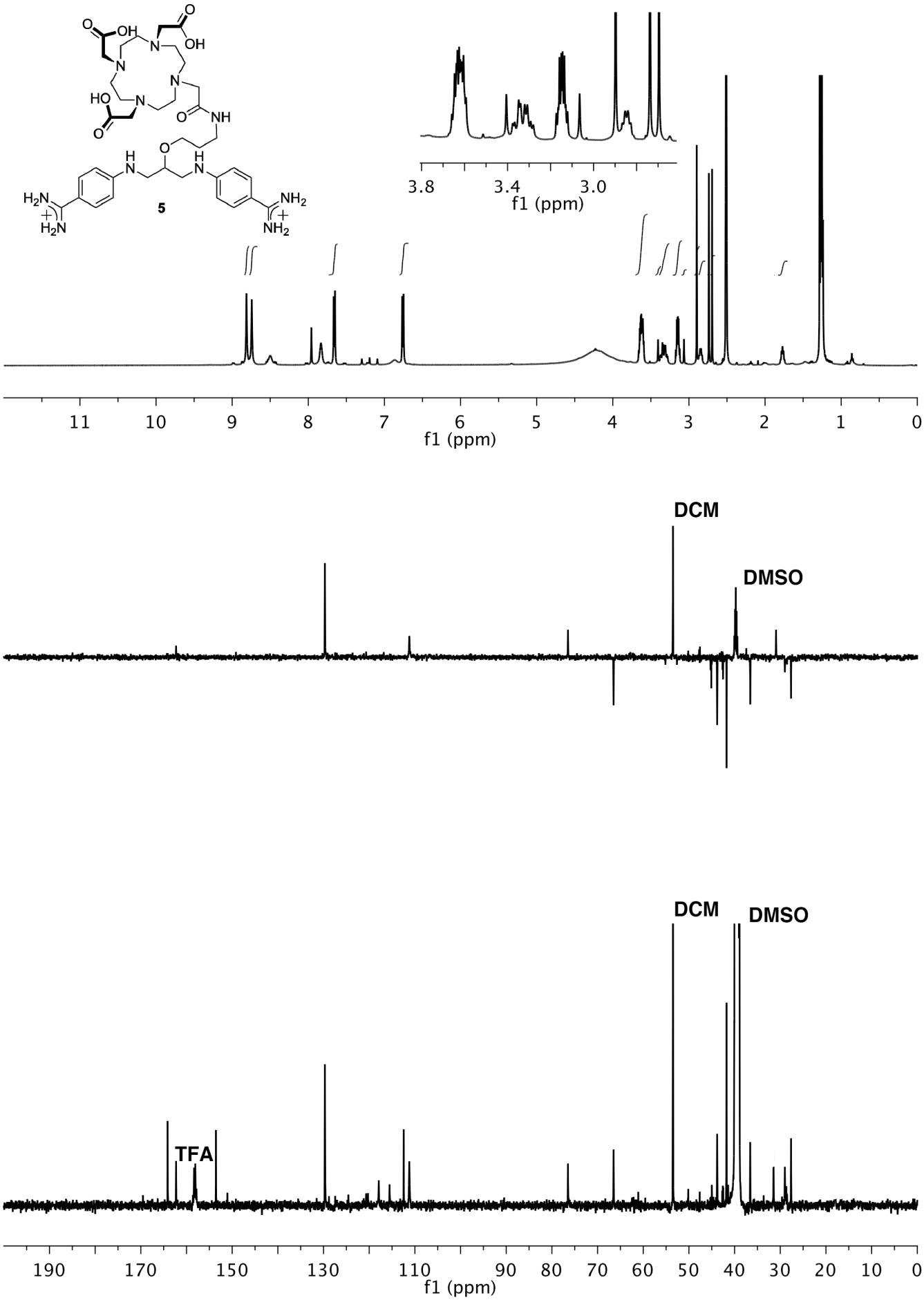
The lanthanide chelation was performed in the same way for both metals: Previous conjugate was dissolved in 600 µL HEPES buffer 10 mM, pH 7.5, NaCl 100mM, and 400 µL of the corresponding lanthanide trichloride (50 mM solution in HCl 1mM) was added. The mixture was shaken for 1h and then purified by analytical RP-HPLC (gradient: 95/5→25/75 over 30 min, water/acetonitrile, 0.1% TFA). The resulting complexes were identified by

MALDI:

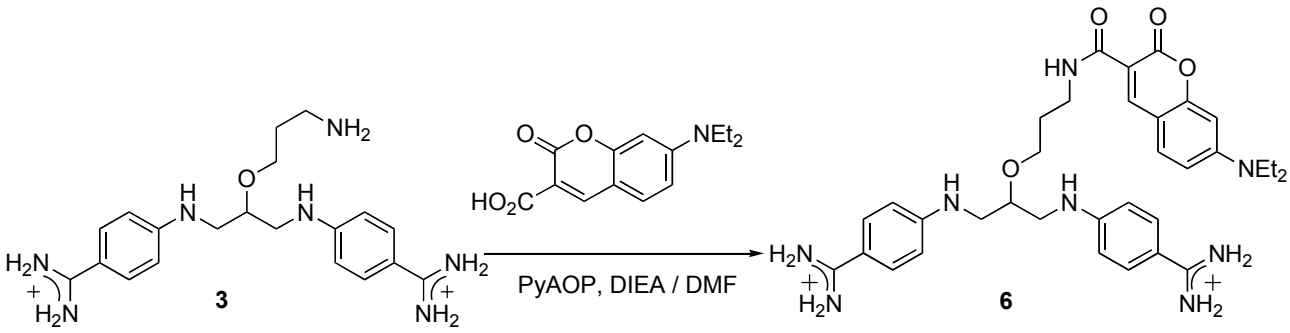
**5a: MALDI/TOF-MS:** $[M+H]^+$calcd. $C_{36}H_{56}Eu\ N_{11}O_8$ for = 920.3 found 920.1

**5b: MALDI/TOF-MS:** $[M+H]^+$calcd. for $C_{36}H_{56}N_{11}O_8Tb$ = 926.3 found 926.3

[1] A. Barge, L. Tei, D. Upadhyaya, F. Fedeli, L. Beltrami, R. Stefanìa, S. Aime, G. Cravotto. *Org. Biomol. Chem*. **2008**, *6*, 1176-1184.


5
3.8
3.4
3.0
f1 (ppm)
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)
DCM
DMSO
DCM
DMSO
TFA
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

## *N-(3-{2-({4-[amino(imino)methyl]phenyl}amino)-1-[({4-[amino(imino)methyl]phenyl}amino) methyl]ethoxy}propyl)-7-(diethylamino)-2-oxo-2H-chromene-3-carboxamide ammoniate (6)*

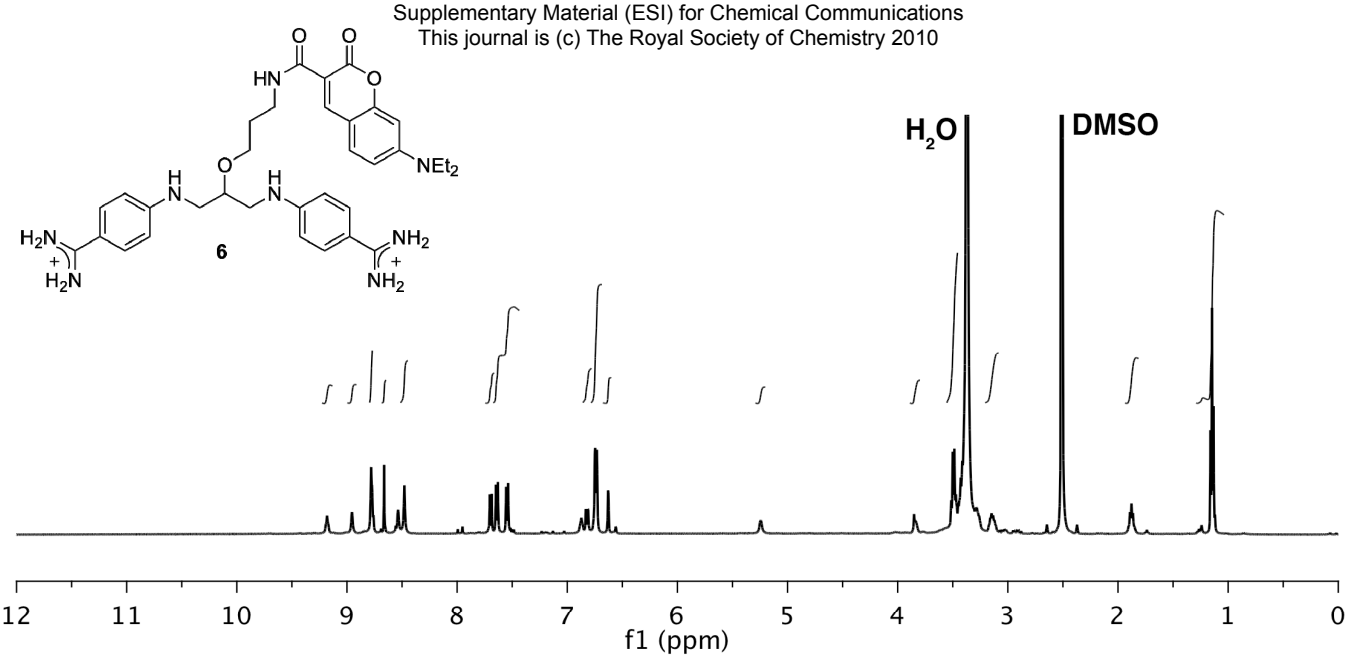



A solution of 7-diethylaminocoumarin-3-carboxylic acid (24 mg, 92 µmol, 4 equiv) in DMF/DIEA 0.2 M (460 µL, 92 µmol 4 equiv) was activated with PyAOP (48 mg, 92 µmol, 4 equiv) and added over **3** (9 mg, 23 µmol, 1 equiv) in DMF/DIEA 0.2 M (230 µL, 46 µmol 2 equiv) and stirred for 2 h. The consumption of the starting material and the appearance of the desired product was confirmed by analytical RP-HPLC. The residue was purified by preparative reverse-phase chromatography (Büchi Sepacore) (gradient: 15% B, 5 min; 15%→95 % B, 30 min.). The appropriated fractions were collected, concentrated and freeze-dried to obtain the desired conjugate (12 mg, 62 %).

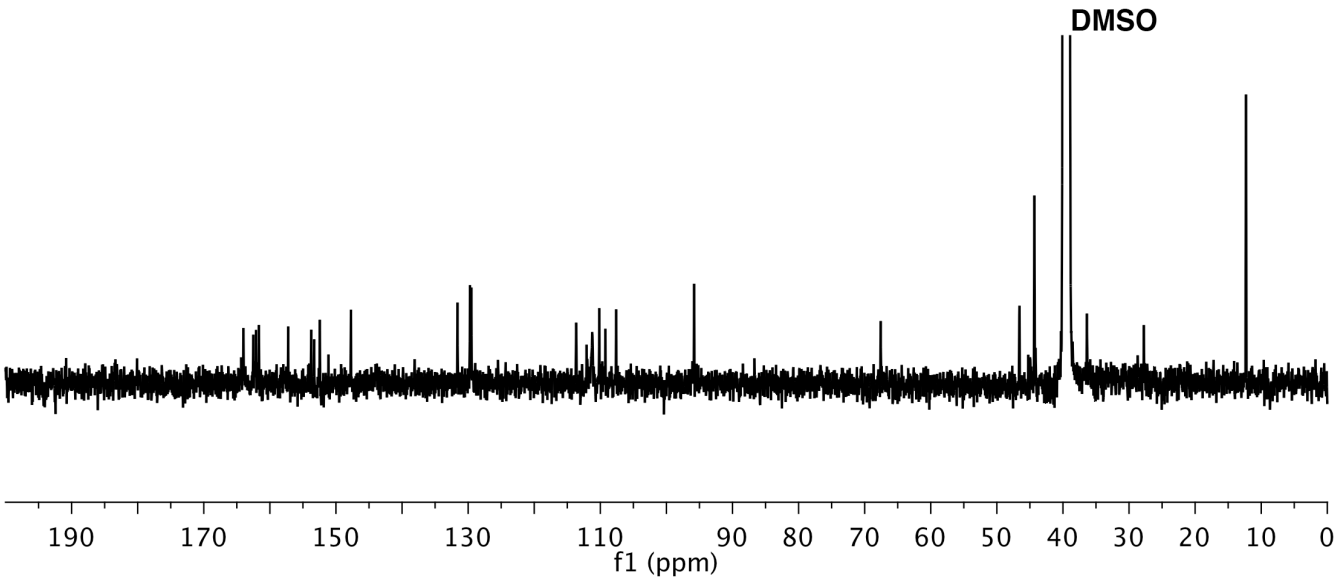
**$^{1}$H-NMR δ (DMSO- $d_6$):** 1.14 (t, $J$ = 6.9 Hz, 6H), 1.87 (q, $J$ = 6.5 Hz, 2H), 3.10-3.14 (m, 2H), 3.22-3.32 (m, 2H), 3.44-3.52 (m, 4H), 3.82-3.86 (b, 1H), 5.23 (s, 1H) 6.62 (s, 1H), 6.73 (d, $J$ = 7.4 Hz, 4H), 6.81 (d, $J$ = 9.0 Hz, 1H), 7.59 (dd, $J_2$ = 46.4, $J_1$ = 8.8 Hz, 4H), 7.68 (d, $J$ = 9.0 Hz, 1H), 8.47 (s, 1H), 8.65 (s, 1H), 8.78 (s, 4H), 8.95 (s, 1H), 9.18 (s, 1H).

**$^{13}$C-NMR δ (DMSO-$d_6$):** 12.3 ($CH_3$), 27.7 ($CH_2$), 36.4 ($CH_2$), 44.1 ($CH_2$), 44.3 ($CH_2$), 46.6 ($CH_2$), 67.6 (CH), 95.5 (CH), 107.6 (C), 109.3 (CH), 110.1 (CH), 111.9 (C), 113.5 (C), 129.7 (CH), 131.6 (CH), 147.7 (CH), 152.4 (C), 153.7 (C), 157.2 (C), 158.2 (C, TFA), 162.1 (C), 162.5 (C), 164 (C). **ESI$^+$-MS:** $[M+H]^+$calcd. for $C_{34}H_{43}N_8O_4$ = 627,3402 found 627,3385.

Extra signals from minor tautomer and amine protons: 6.87 (m), 8.53. Extra C signals from minor tautomer: 111.1 (CH), 129.6 (CH), 151.3 (C), 153.2 (C), 162.3 (C).


H2O
DMSO
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)


DMSO


DMSO
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

# Fluorescence Spectroscopy

Measurements were made with a *Jobin-Yvon Fluoromax-3*, (*DataMax 2.20*), and coupled to a *Wavelength Electronics LFI−3751* temperature controller.

### *Fluorescence experiments of lanthanide chelates (5a and 5b) with DNA*

The fluorometer settings for lanthanide experiments were the following:

**for $Eu^{3+}$:** increment: 1.0 nm; integration time: 0.2 s; excitation slit width: 6.0 nm; emission slit width: 6.0 nm; excitation wavelength 329 nm. The emission spectra were recorded at 20 ºC from 550 to 750 nm with a 500 nm long-pass filter to avoid interference from harmonic doubling.

**for $Tb^{3+}$:** increment: 1.0 nm; integration time: 0.2 s; excitation slit width: 3.0 nm; emission slit width: 6.0 nm; excitation wavelength 329 nm. The emission spectra were recorded from 345 to 600 nm at 20 ºC;

Hairpin oligonucleotides were supplied by *Thermo Fischer Scientific GmbH* and their sequences were:

| DNA | Full sequence (5' to 3') |
|---|---|
| **AATTT** | GGCG *AATTT* CGC TTTTT GCG AAATT CGCC |
| **AATTC** | GGCG *AATTC* AGC TTTTT GCT GAATT CGCC |
| **AATGC** | GGCG *AATGC* AGC TTTTT GCT GCATT CGCC |
| **GGCCC** | GGCA *GGCCC* AGC TTTTT GCT GGGCC TGCC |

In both cases, the fluorescence emission spectrum of a 0.5 µM solution of lanthanide chelate (**5a** or **5b**) in Tris-HCl buffer 20 mM pH 7.5, 100 mM NaCl, were compared with mixtures with 9 equivalents of either **GGCC, AATGC** DNAs or the target **AATTC** and **AATTT** DNAs. The following figure shows the fluorescence emission spectra for the experiments of europium and terbium chelates (**5a/b**). Spectra with **AATGC** are very similar to those with **GGCCC** as three A/T base-pairs is too short to accommodate the propamidine unit.

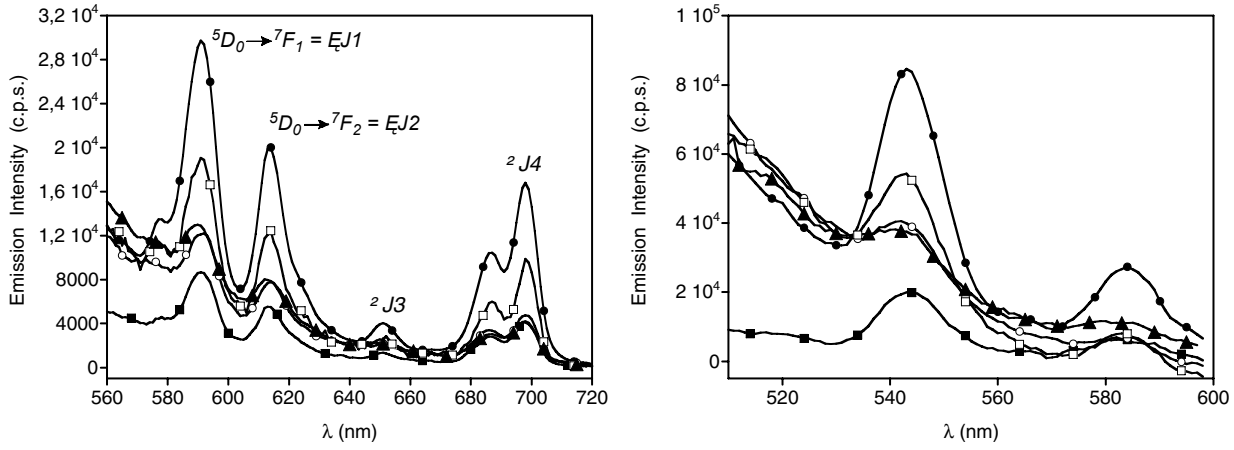



**Figure S1: Left:** Emission spectra of DOTA[$Eu^{3+}$] **5a**; **Right:** DOTA[$Tb^{3+}$] **5b** in Tris-HCl buffer 20 mM pH 7.5, 100 mM NaCl. Spectrum in absence of DNA (■), with 9 equiv of non-specific dsDNA (**GGCCC**, ○) or non-target dsDNA (**AATGC**, ▲), with 9 equiv of sub-optimal dsDNA (**AATTC**, □), and with 9 equiv. of target dsDNA (**AATTT**, ●).

## *Design and fluorescence of coumarin conjugate 6*

A key parameter for efficient energy transfer between two fluorophores is the combination of species with suitable energy levels. This energetic requirement can be easily visualized by the overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor fluorophore. For our particular case, this is reflected in the overlap between the emission of the aminobenzamidine fluorophore (Figure S2, ○), and the absorption spectrum of (Figure S2, dashed line, band b).

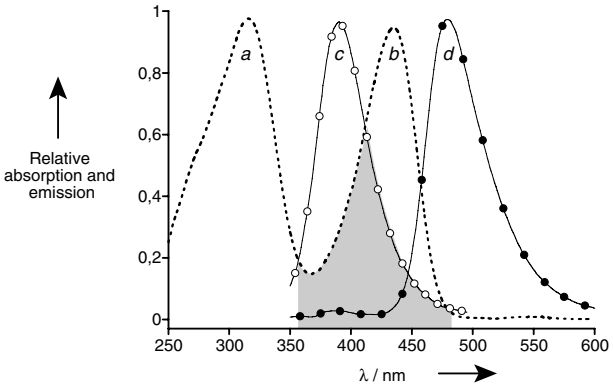



**Figure S2:** Dashed line: UV absorption of **6** composed by two bands from each chromophore in the conjugate: a) corresponds to the absorption of the *aza*-benzamidinium unit, b) results from coumarin absorption. Solid lines: fluorescence emission of **1** (○) and fluorescence emission of **6** (●). Spectra were normalized for comparison.

The excitation spectra of **6** were recorded in buffer, and with 9 equivalents of **AATTT** DNA, monitoring the emission at 472 nm (coumarin emission). The excitation spectrum of conjugate **6** in buffer is very weak, and its maximum excitation corresponds to the expected coumarin excitation band at 432 nm. Upon addition of **AATTT**, a large increase in the excitation spectrum intensity is observed, and the excitation band of the bis(benzamidine) at 329 nm becomes more intense, demonstrating the energy transfer between both fluorophores (exciting at 329 the bis(benzamidine) fluorophore results in increased coumarin emission at 472 nm, Figure S3). The excitation spectra were recorded from 220 to 470 nm at 20 ºC with the following settings: increment: 1.0 nm; integration time: 0.2 s; excitation slit width: 3.0 nm; emission slit width: 4.0 nm; emission at 474. The emission measurements were made with the following settings: increment: 1.0 nm; integration time: 0.2 s; excitation slit width: 3.0 nm; emission slit width: 4.0 nm; excitation wavelengths 329 or 434 nm spectra recorded from 345 to 640 nm.

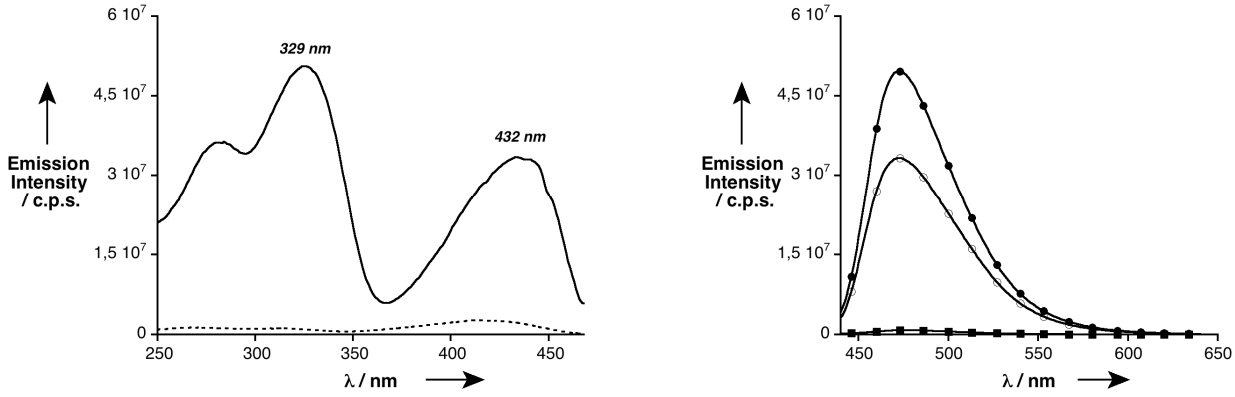



**Figure S3: Left:** Excitation spectrum at 472 nm of 0.5 μM solution of **6** in Tris-HCl buffer 20 mM pH 7.5, 100 mM NaCl (dashed line), and with 9 eq. of **AATTT** DNA (solid line). **Right:** Fluorescence emission of analogue sample **6** in absence of DNA (■); emission spectrum of analogue sample with 9 eq. of **AATTT** DNA and excitation wavelength 432 nm (○); emission spectrum with 9 eq. of **AATTT** DNA; excitation wavelength 329 nm (●).

## *Fluorescence titrations of conjugate 6 with different DNAs*

Measurements were made with a Jobin−Yvon Fluoromax−3, (DataMax 2.20), and coupled to a Wavelength Electronics LFI−3751 temperature controller. Except for the case of the lanthanide experiments, the measurements were made with the following settings: increment: 1.0 nm; integration time: 0.2 s; excitation slit width: 3.0 nm; emission slit width: 6.0 nm; excitation wavelength 329 nm. The emission spectra were recorded from 345 to 640 nm at 20 ºC.

All titrations were made following the same procedure: to 1 mL of a 0.5 μM solution of coumarin conjugate (**6**) in Tris-HCl buffer 20 mM pH 7.5, 100 mM NaCl, aliquots of a ≅ 400 μM stock solution of respective annealed DNA were successively added and the fluorescence spectra were recorded after each addition. The maximum emission wavelength was then used for the determination of the binding constant as detailed below.

The values of $K_D$ found for coumarin conjugate and DNA **AATTT**, **AATGC**, were respectively: (814 ± 90) nM, and (2.38 ±0.08) μM.

The following figures show the fluorescence emission spectra for the different titrations of coumarin conjugate (6) with selected hairpin oligonucleotides reported in the article **AATTT**, **AATGC**.

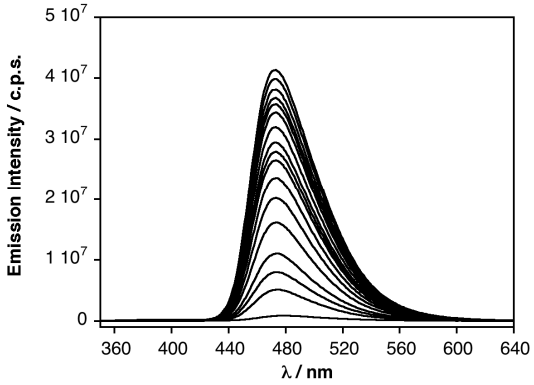



**Figure S4:** Fluorescence spectra of the titration of coumarin conjugate (**6**) with dsDNA **AATTT**

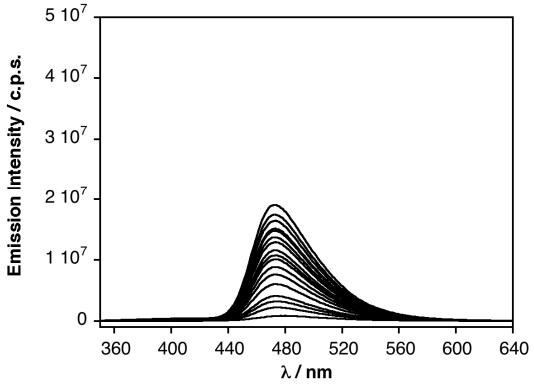



**Figure S5:** Fluorescence spectra of the titration of coumarin conjugate (**6**) with dsDNA **AATGC**.

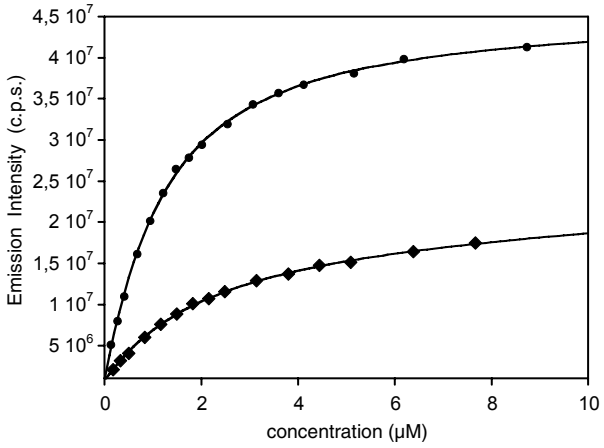



**Figure S6:** Titration profiles of coumarin hybrid **6** in Tris-HCl buffer 20 mM pH 7.5, 100 mM NaCl with the truncated-site dsDNA (**AATGC**, ♦), and with the target dsDNA (**AATTT**, ●). Lines represent the best fit to a 1:1 binding model (see below). The apparent $K_D$ are: (814 ± 90) nM for **AATTT**; (2.38 ± 0.24) and μM for **AATCG**.

## *DNA binding models: Accounting for the residual emission fluorescence of DNA*

The typical equation for a 1:1 binding in which an unlabeled ligand (dsDNA) is added over a fluorescent receptor (coumarin conjugate, **6**) is described by the following equations when non-specific binding is ignored:

$$K_D = \frac{R \cdot L}{C} \qquad (1)$$

$$R_T = R + C \qquad (2)$$

$$L_T = L + C \qquad (3)$$

$$F_T = F_0 + F_C \times C \qquad (4)$$

Where $R$ is the concentration of the free receptor in the equilibrium; $R_T$, total receptor concentration (**6**, kept approximately constant throughout the titration); $L$, concentration of the free ligand in the equilibrium; $L_T$, total concentration of added ligand (DNA); $K_D$, dissociation constant of the interaction between the receptor and the ligand; $C$ concentration of labeled ligand-receptor complex; $F_T$, total observed fluorescence; $F_0$, adjustable parameter accounting for the background fluorescence; $F_C$ adjustable parameter for the labeled ligand-receptor complex molar fluorescence.

$$F_T = \frac{1}{2}[2F_0 + F_C\sqrt{K_D^2 + (L_T - R_T)^2 + 2K_D(L_T + R_T))}] \qquad (5)$$

Solving the system for $F_T$ and eliminating $R$, $L$, and $R_L$, we obtain the well-known equation 5. This equation can be used when the contribution to the total fluorescence of the DNA is small, which only occurs when **6** is efficiently bound to its target DNA sequences and there is a large increase in its fluorescence emission signal.