Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers.

Supporting Information

Dibenzopyridoimidazocinnolinium cations: a new family of light-up fluorescent DNA probes

Pedro Bosch,[1] David Sucunza,[1,*] Francisco Mendicuti,[2] Alberto Domingo,[3] Juan J. Vaquero[1,*]

[1]*Departamento de Química Inorgánica y Química Orgánica,* [2]*Departamento de Química Analítica, Química Física e Ingeniería Química,* [3]*Departamento de Biología de Sistemas, Universidad de Alcalá, 28805- Alcalá de Henares, Madrid, Spain.*

david.sucunza@uah.es, juanjose.vaquero@uah.es

Table of Content

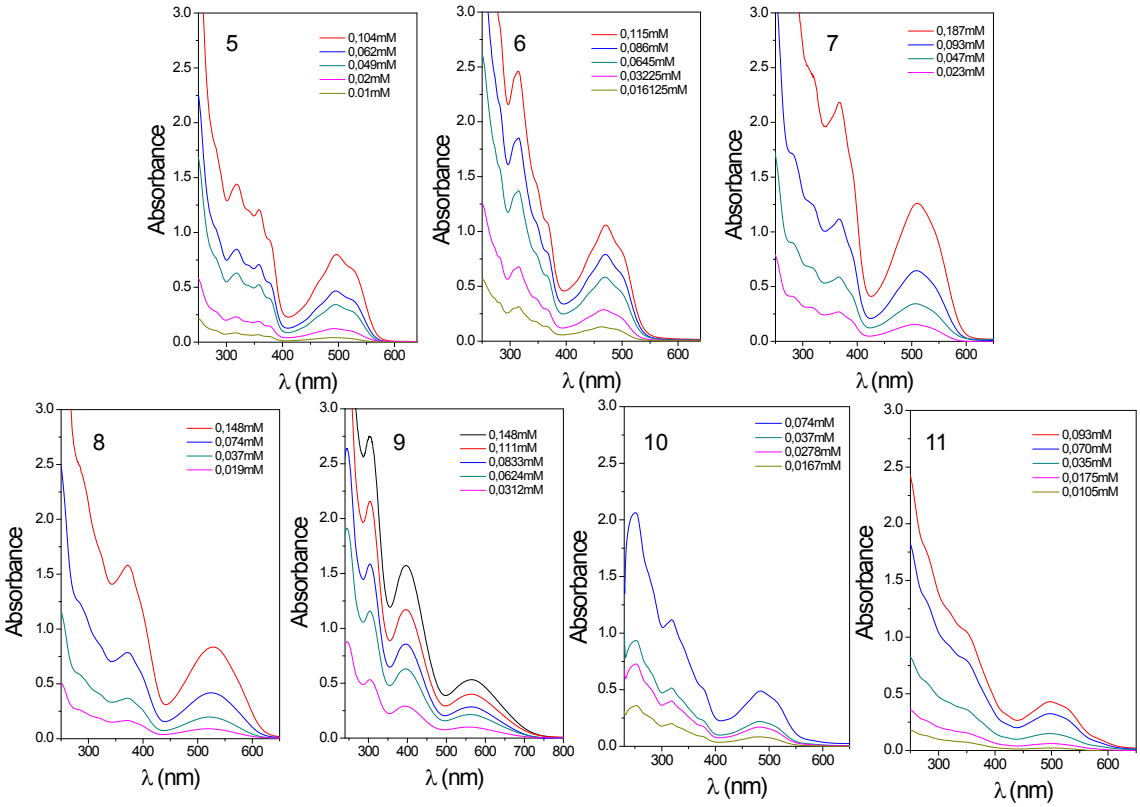


Figure 1S. Absorption spectra for (upper panel) **5**, **6** and **7** (bottom panel) **8**, **9**, **10** and **11** in 5%DMSO/water at 25ºC and several concentrations.

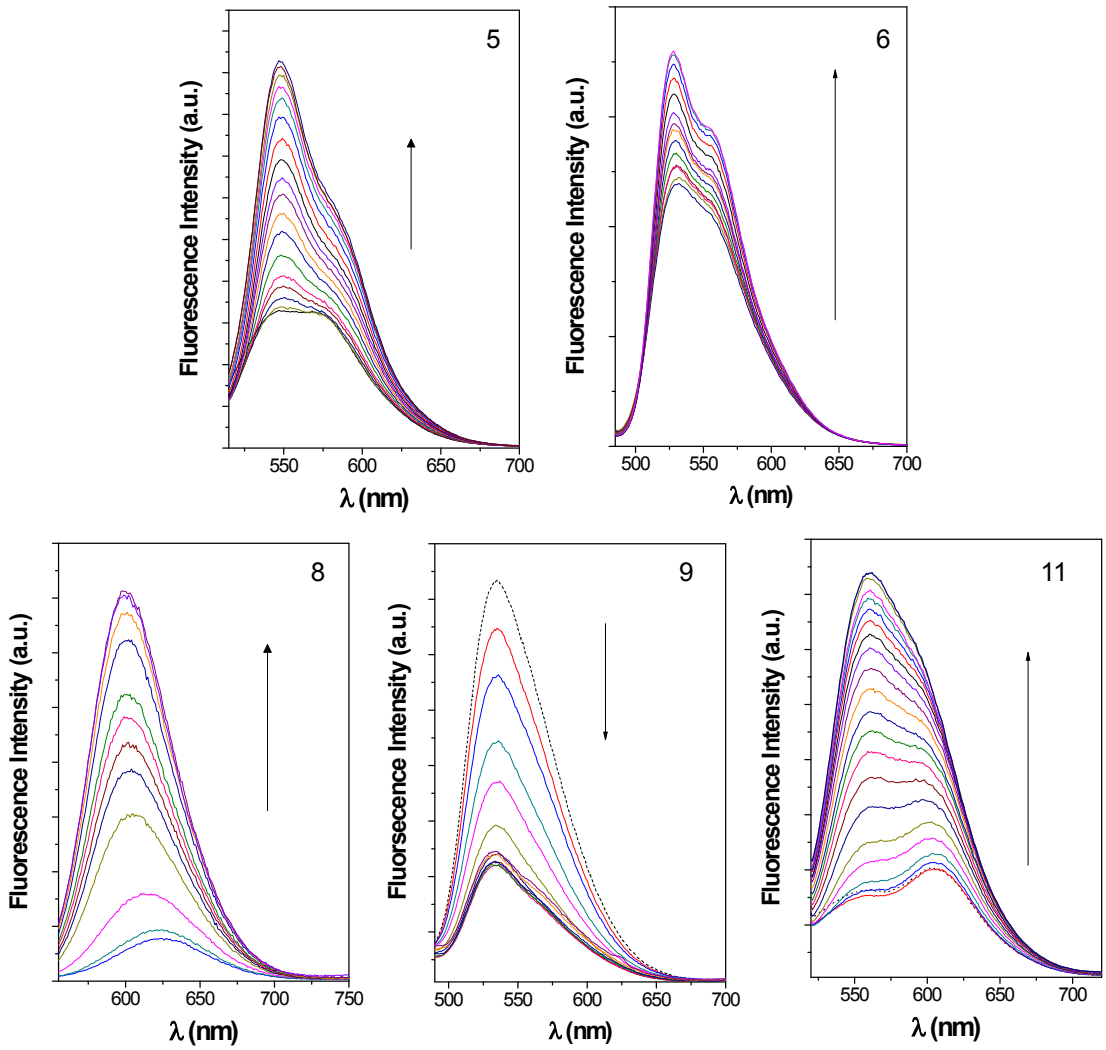


Figure 2S. Emission spectra for (upper panel) **5** and **6** (bottom panel) **8**, **9** and **11** dilute solutions in 5%DMSO/water at 25ºC and upon DNA addition.

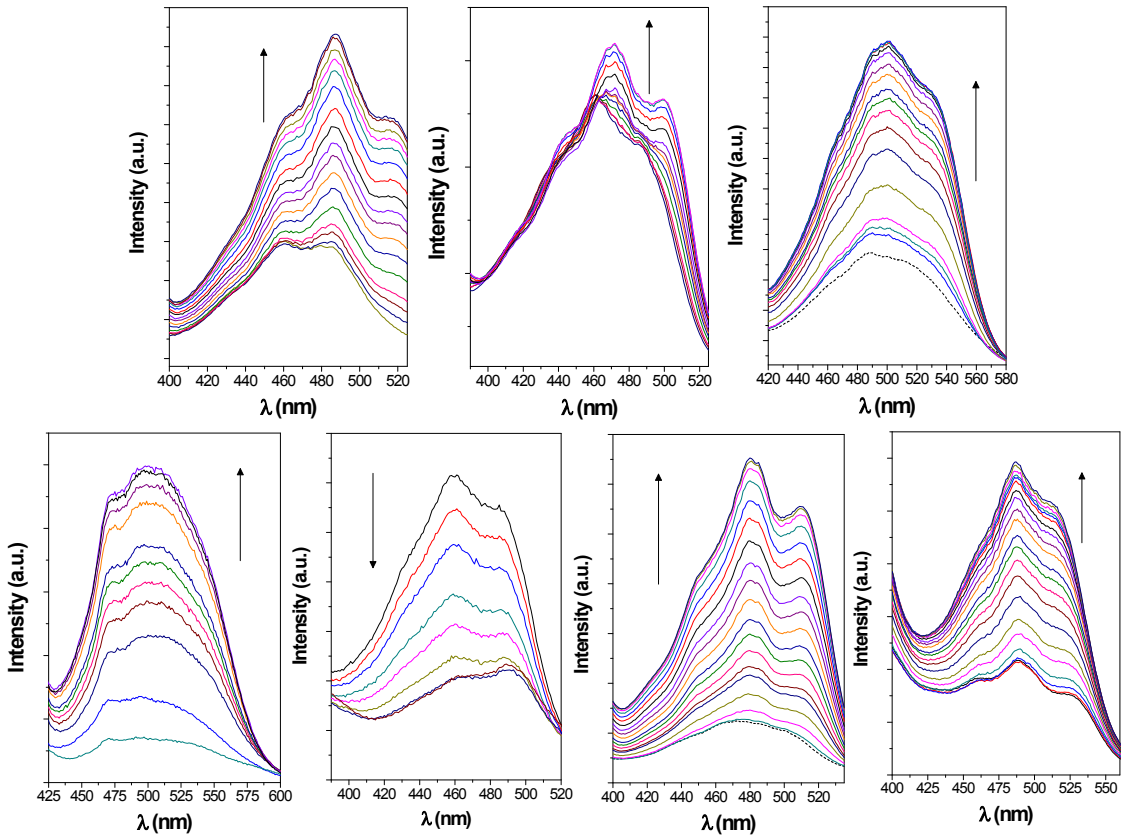


Figure 3S. Excitation spectra for (upper panel, left to right) **5**, **6**, **7** and (bottom panel, left to right) **8**, **9**, **10** and **11** dilute solutions in 5% DMSO/water at 25ºC upon the adding of DNA aliquots (λ_{em} were fixed for at the maximum of emission intensities for each ligand).

Thermodynamics of the ligand-macromolecule binding

For the 1:1 binding of a ligand molecule (L) to a single binding site of a large molecule (M), whose association is described by the equilibrium,

$$L + M \underset{}{\overset{K}{\rightleftharpoons}} LM \qquad (1S)$$

the association constant, *K* is given by,

$$K = \frac{[LM]}{[L][M]} \qquad (2S)$$

where [L], [M], and [LM] symbolize the concentration of each species at the equilibrium. For higher order 1:n (L:M) stoichiometries, *i.e.,* the ligand is bound to n effective binding sites of M. Assuming that these n binding sites of M per ligand (L) are independent and equivalent (named multiple independent binding sites) the total M and L concentration in solution can be obtained by the mass balances:

$$[L]_0 = L + [LM] \qquad (3S)$$

$$[M]_0 = [M] + n[LM] \qquad (4S)$$

From 2 and the mass balances 3 and 4, the following equation is derived:

$$nK[L]_0\left(\frac{[M]}{[L]_0}\right)^2 - \left(1 + nK[L]_0 + K[M]_0\right)\frac{[M]}{[L]_0} + K[M]_0 = 0 \qquad (5S)$$

and the fraction of the ligand bound to the macromolecule $[LM]/[L]_0$ is calculated as:

$$\frac{[LM]}{[L]_0} = \frac{\left(1 + n\,K\,[L]_0 + K[M]_0\right) \pm \sqrt{\left(1 + nK[L]_0 + K\,[M]_0\right)^2 - 4\,nK^2\,[L]_0[M]_0}}{2K\,n\,[L]_0} \qquad (6S)$$

By assuming two fluorescent species at the equilibrium, the L and the LM, the fluorescence intensity (measured as the area under the emission spectra) can be related to the initial $[M]_0$ and $[L]_0$ concentrations and *K* by the following equation:

$$I = I_0 + (I_\infty - I_0) \times \frac{(1 + n\,K\,[L]_0 + K[M]_0) - \sqrt{(1 + nK[L]_0 + K\,[M]_0)^2 - 4\,nK^2\,[L]_0[M]_0}}{2K\,n\,[L]_0} \qquad (7S)$$

and $\Delta I/I_0$, the normalized difference between the intensity of fluorescence (I) for L in the presence of M and in its absence, I_0, can by written as:

$$\frac{\Delta I}{I_0} = \left(\frac{I_\infty - I_0}{I_0}\right) \times \frac{(1 + n\,K\,[L]_0 + K[M]_0) - \sqrt{(1 + nK[L]_0 + K\,[M]_0)^2 - 4\,nK^2\,[L]_0[M]_0}}{2K\,n\,[L]_0} \qquad (8S)$$

where I_∞ represents the fluorescence intensity for the totally complexed chromophore guest. Equation 8 can be modified including the $[M]_0/[L]_0$ molar ratio *R* as:

$$\frac{\Delta I}{I_0} = \left(\frac{I_\infty - I_0}{I_0}\right) \times \frac{(1/[L]_0 + n\,K + K\,R) - \sqrt{(1/[L]_0 + nK + K.R)^2 - 4\,nK^2 R}}{2K\,n} \qquad (9S)$$

In our case L is the ligand and M is the DNA. Thus $[L]_0$ is the initial concentration of the ligand, $[M]_0$ is the DNA concentration per mol of base pairs and n is the number of base pairs per bound ligand molecule.

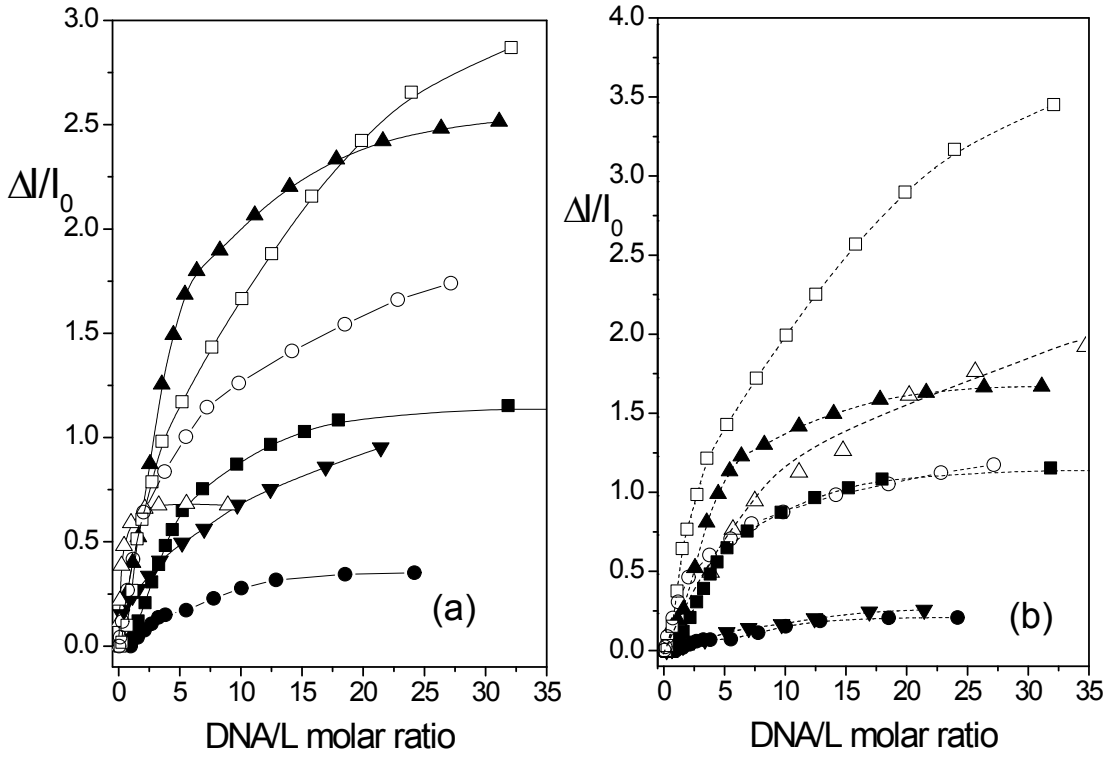


Figure 4S. Normalized changes of the fluorescence intensities for ligands 5%DMSO/water solutions 25ºC obtained from (a) emission spectra (b) excitation spectra versus DNA/ligand molar ratios during titrations. Symbols are 5(○), 6 (ᔓ), 7(□), 8(❑), 9(❒), 10 (●) and 11(○).

Fluorescence anisotropy, *r*

The fluorescence anisotropy *r* was obtained from the fluorescence polarization measurements by using the *L-format* method.[1] The anisotropy *r* was defined as:

$$r = (I_{VV} - GI_{VH})/(I_{VV} + 2GI_{VH}) \qquad (10S)$$

where I_{xy} is the intensity of the emission that is measured when the excitation polarizer is in position x (V for vertical, H for horizontal), the emission polarizer is in position y, and the G factor (= I_{HV}/I_{HH}) corrects for any depolarization produced by the optical system.

Weighted average lifetimes $<\tau>$

The weighted average lifetime of a multiple-exponential decay function was then defined as

$$\langle\tau\rangle = \frac{\sum_{i=1}^{n} A_i \tau_i^2}{\sum_{i=1}^{n} A_i \tau_i} \qquad (11S)$$

where A_i is the pre-exponential factor of the component with a lifetime τ_i of the multi-exponential decay function, $I(t) = \sum_{i=1}^{n} A_i \tau_i$.[2,3]

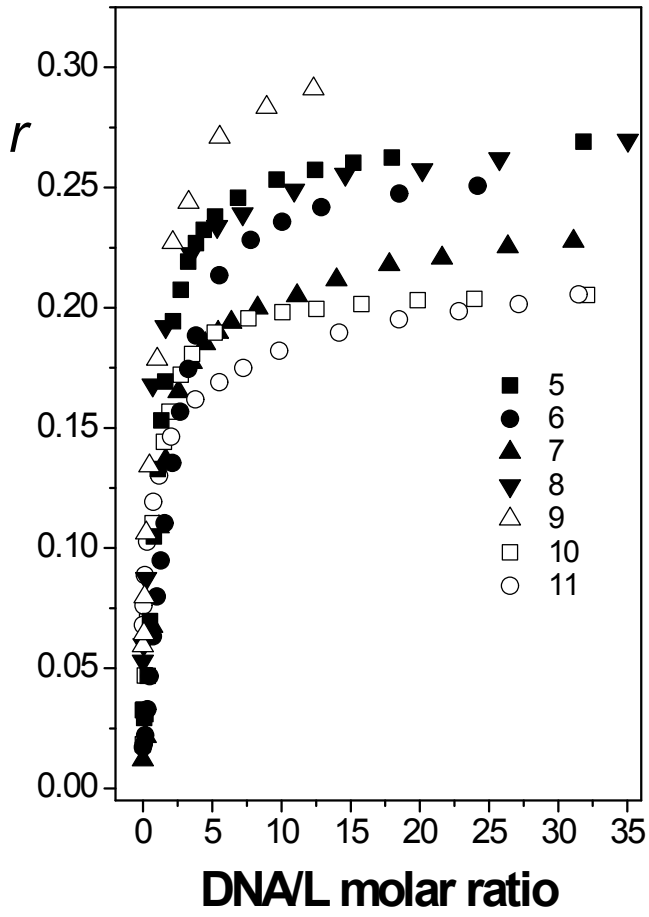


Figure 5S. Fluorescence anisotropies (*r*) changes for ligands in 5%DMSO/water solutions 25ºC during titrations versus DNA molar concentration. Measurements were performed in 1cm path cuvettes.

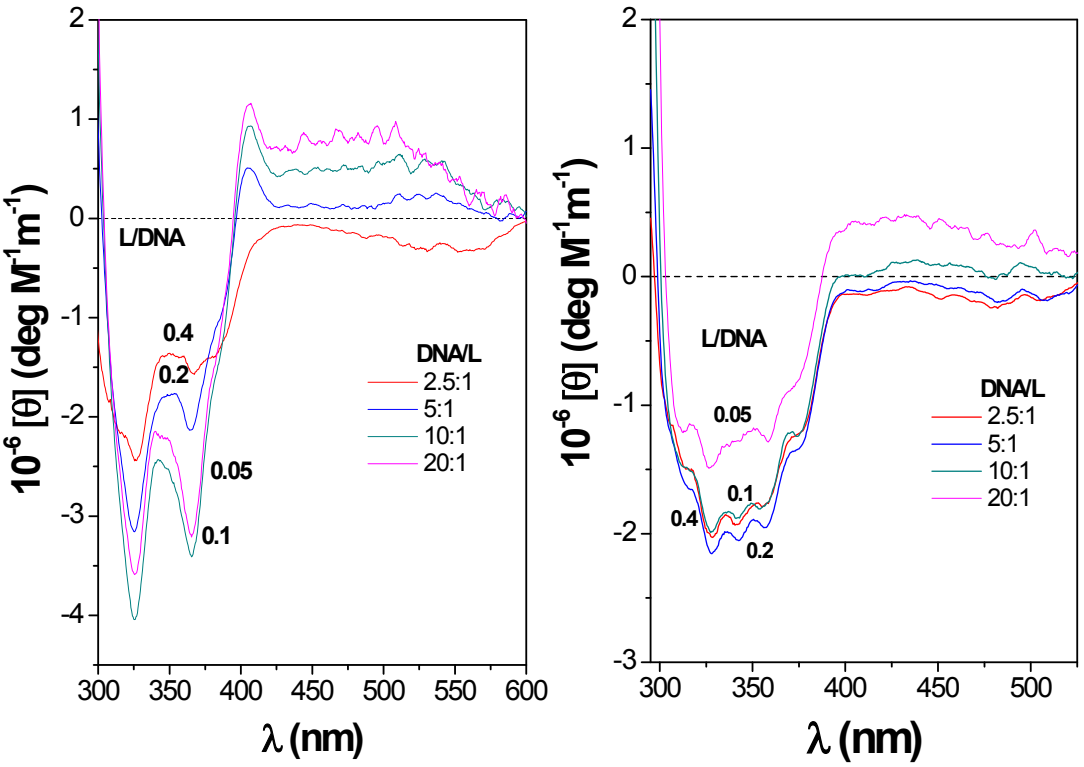


Figure 6S. Molar ellipticities ([θ]) measured at 25ºC for cationic ligands **7** (left) and **10** (right) in 5% DMSO/water solutions in the presence of DNA ([DNA]=1.95×10^{-4} M and 3.0×10^{-4} M respectively) at several ligand/DNA molar ratios. Measurements were performed in 1cm path cuvettes.

Theoretical protocols on the binding of a Ligand to DNA

Molecular Mechanics (MM) and Molecular Dynamics (MD) were performed to study the forces responsible for the binding of ligands to DNA and the structure and stability of the system. Calculations were performed with Sybyl X-2.0[4] and the Tripos Force Field.[5] B-DNA helical fragments, whose charges were derived by using the Gasteiger and Marsili method, contained twelve nucleotides with a CGCGAATTCGCG sequence (% of the calf-thymus DNA bases). Charges for the ligands were obtained by using the Gaussian suite of quantum chemical programs at the HF/6-31G(d) level of theory.[6] A relative permittivity ε=1 was used for electrostatic contributions in the presence of explicit water (ε=3.5 for some preliminary calculations in the vacuum). Water solvation was performed by using the Molecular Silverware (MS) algorithm and periodic boundary conditions (PBC).[7] Non-bonded cut-off distances for MM, as well as for MD were set at 12 Å. Optimizations were carried out by the simplex algorithm, and the conjugate gradient was used as a termination method with gradients of 3.0 Kcal/molÅ for the calculations carried out in water).[8]

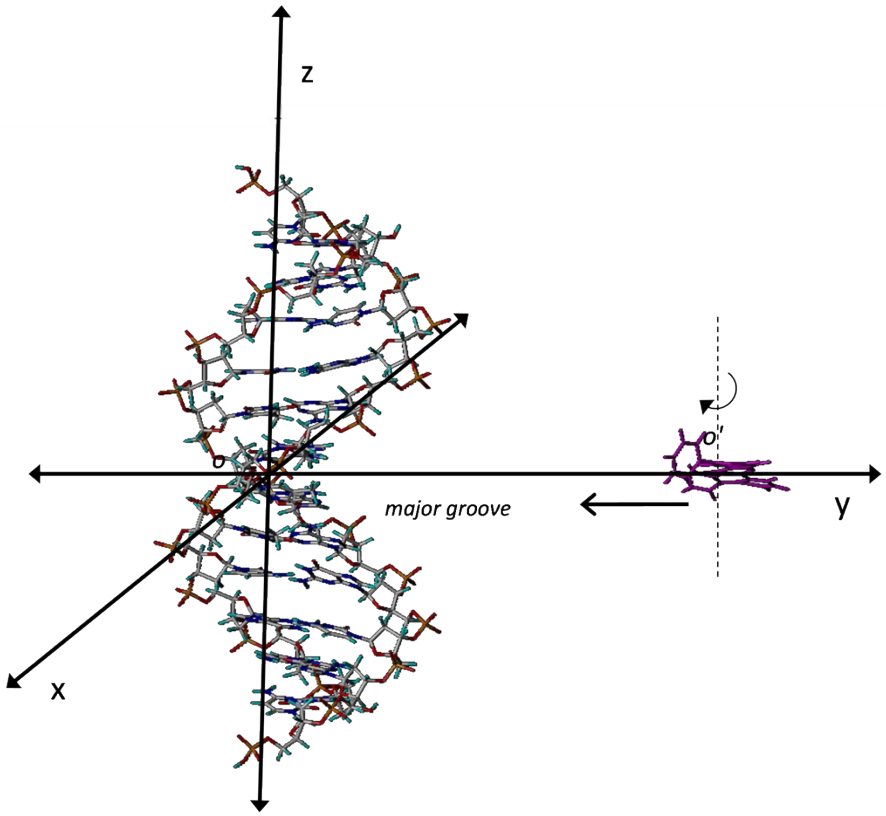


Figure 7S. Cartesian system used to define the L-to-DNA approaching. The pair of central AT and TA bases were almost parallel to the xy ligand plane. The most favorable L-to-DNA relative orientation for approaching is depicted.

The DNA was placed with its center of mass at the origin of a coordinate system oriented as depicted in Figure 1S. In addition, to maintain the regular helical structure of the pair of DNA chains and avoid the unwinding of the end portions of this short DNA helix during MD simulations, the N…HN hydrogen bond distances between each pair of nucleobases were kept constant. For this purpose a harmonic penalty energy function was added to the force field equations for those atoms which were involved in the constraint. This function can be written as $E=k\ (r-r_i)^2$, where k=200 kcal/molÅ^5, r_i and r are the initial distance (equilibrium) and at each time of the MD trajectory respectively.

To investigate the structure of the L-DNA system and the type of interactions responsible for its stabilization, L was approached to the DNA in 0.5 Å steps along the *y* coordinate from *y* = +30 to -30 (Å) respectively. This means that L located in the xy plane coming from the major groove side intercalated in the DNA (between both central TA and AT base pairs), passed through it and left through the minor groove. Because of the experimentally obtained ICD sign, the possibility of binding to grooves was not considered.

In a first step, the most favorable orientations for L approaching were achieved from the L-DNA binding energy calculations in the vacuum for all structures generated by rotating the ligand around the axis, which passing through (*o*') was perpendicular to the ring plane, and changing the *o-o'* distance (measured along the y coordinate) in the -180º-180º and -30Å to +30 Å ranges in small steps of 30º and 2Å respectively. Fixed a L relative to DNA favorable orientation, L was approached to the DNA helix in the same range of distances (measured along the y coordinate) in small 0.5 Å steps as stated in the previous paragraph. Each structure generated in the latter protocol was solvated (MS and PBC) and optimized (gradient 3 kcal/molÅ). The minima binding energy (MBE) structures for the L/DNA complexes were optimized once again (gradient 0.5 kcal/molÅ) and used as starting conformations for 1.0 ns MD simulations following the same protocols described earlier for the complexation of other systems.[9]

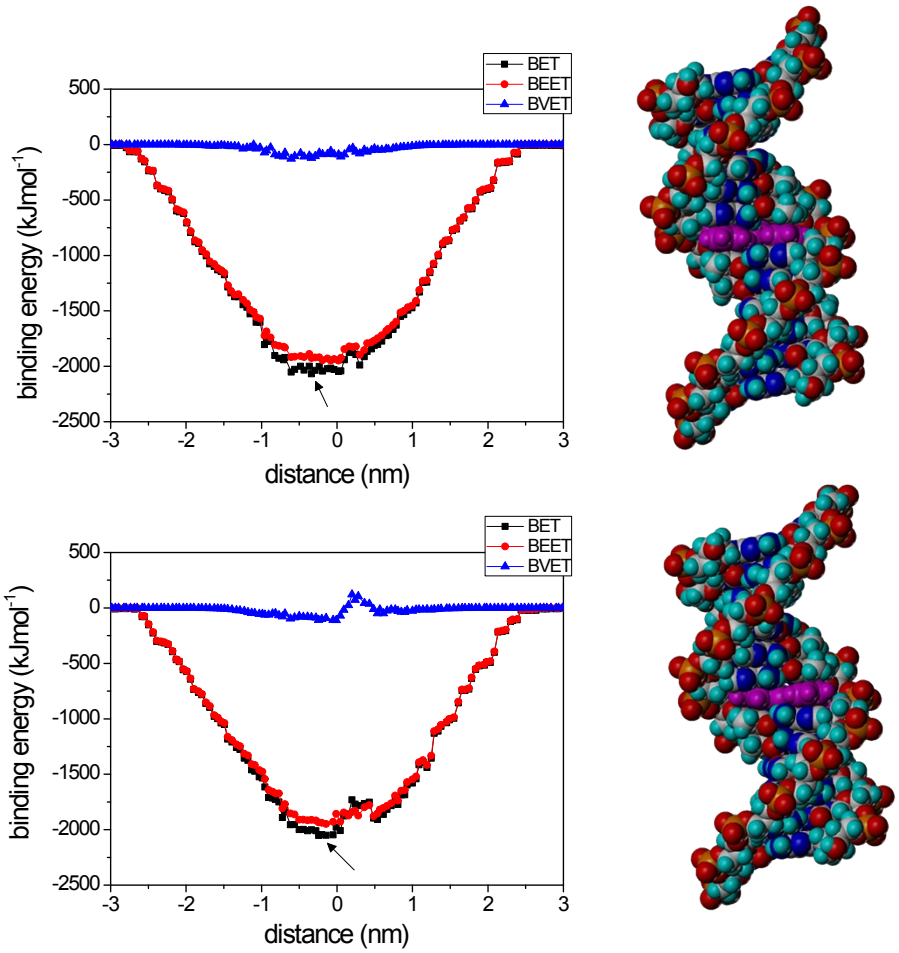


Figure 8S. Total binding energies (O), and van der Waals (□) and electrostatics (ᔓ) contributions as a function of the *oo'* distance measured along the *y* coordinate for (upper) ligands **7** and (bottom) **10** approaching the DNA as intercalators with their long axis parallel to AT, TA base pairs pocket. Represented are the MBE structures (pointed out buy the arrows) for both ligands, once minimized again, were used as starting conformations for the MD simulations.

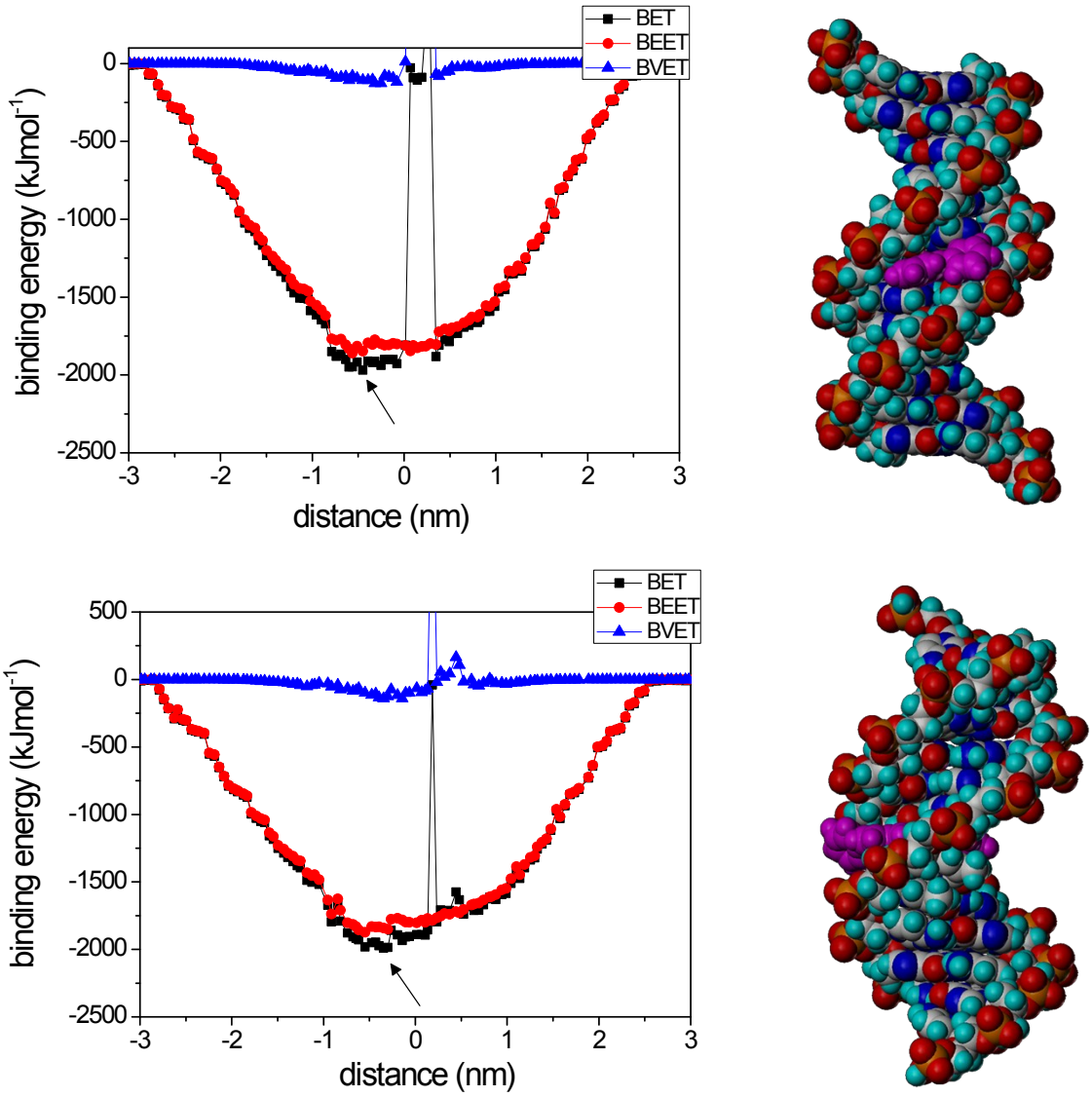


Figure 9S. Total binding energies (O), and van der Waals (□) and electrostatics (ꝏ) contributions as a function of the *oo'* distance measured along the *y* coordinate for (upper) ligands **7** and (bottom) **10** approaching the DNA as intercalators with their long axis perpendicular to AT, TA base pairs pocket . Represented are the MBE structures (pointed out buy the arrows) for both ligands, once minimized again, were used as starting conformations for the MD simulations.

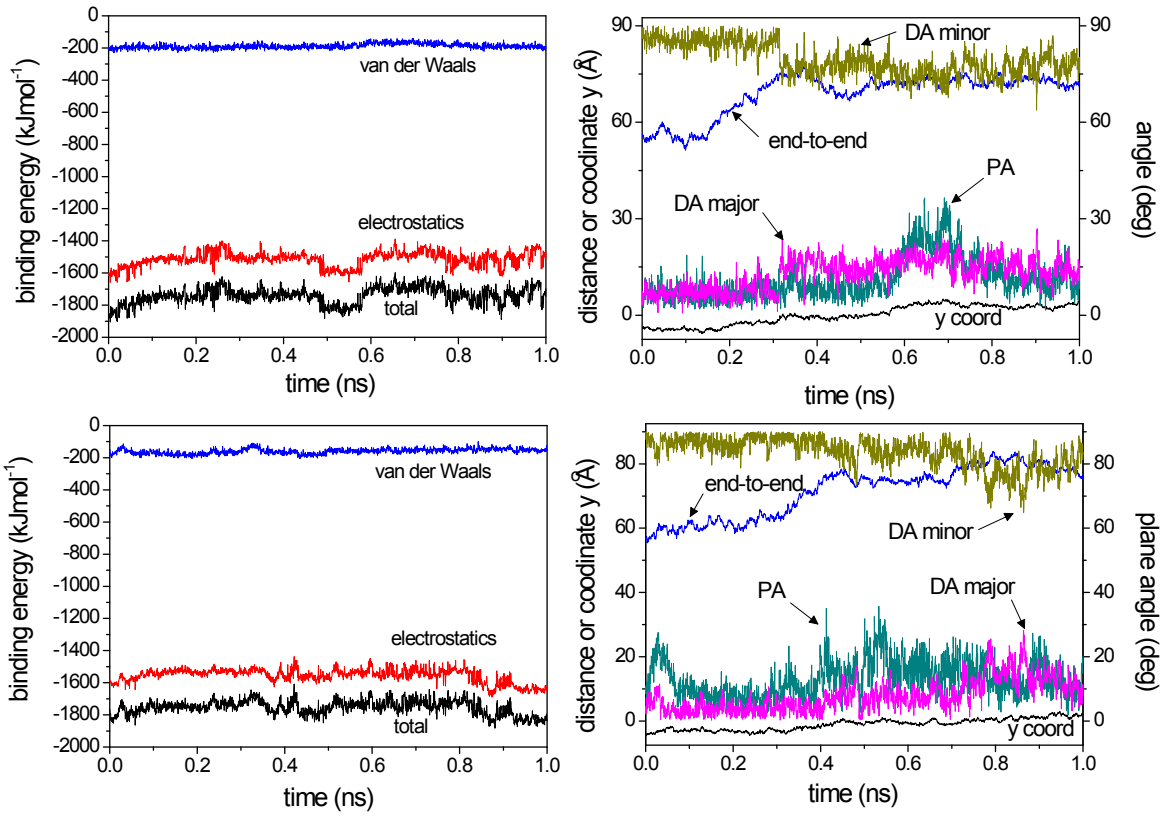


Figure 10S. (Left panels) Histories for binding energies and electrostatics and van der Waals contributions; (right panels) histories for the *y* coordinate (y coord) of the centre of ligands, for the end-to-end twelve nucleotide DNA fragment distance (end-to-end), for the averaged angles between the plane of both AT and TA pairs of bases and the plane of ligands rings (PA) and the angles between the pocket bases axis and the long (DA major) or the short ligand axis (DA minor) obtained from the analysis of the 1 ns MD trajectory at 300K starting from the minimized MBE structure of the **7**/DNA complex (upper panels) and **10**/DNA (bottom panels) by approaching ligand with its long axis parallel to the base pairs pocket .

Table 1S. Average of some energetic and geometrical parameters obtained from the analysis of 1ns MD trajectories for **7**/DNA and **10**/DNA complexes by approaching ligand with its long or main axis parallel (||) or perpendicular (⊥) to the base pairs pocket axis.

Parameter	**7** (||)	**7** (⊥)	**10** (||)	**10** (⊥)
$Etotal_{bin}$ ($kJmol^{-1}$)	-1741.3 ± 52.2	-1633.4 ± 79.5	-1754.8 ± 44.5	-1654.7 ± 38.5
EE_{bin}	-1514.6 ± 46.4	-1439.7 ± 77.8	-1555.7 ± 42.3	-1477.1 ± 33.9
EvW_{bin}	-188.6 ± 15.4	-155.5 ± 13.9	-160.4 ± 16.7	-139.1.± 13.0
y (Å)	+0.1 ± 3.0	-2.5 ± 2.5	-1.1 ± 1.8	-4.2 ± 1.1
distance (Å)	3.8 ± 1.0	4.2 ± 0.9	2.1 ± 1.1	4.7 ± 1.2
end-to-end distance (Å)	68.9 ± 6.6	65.2 ± 9.2	71.6 ± 7.9	63.9 ± 8.1
plane angle (deg)	11.2 ± 6.5	13.2 ± 5.9	12.9 ± 7.9	11.9 ± 4.3
main axis angle (deg)	12.9 ± 4.9	82.2 ± 5.8	7.8 ± 5.1	72.9 ± 5.9
short axis angle (deg)	79.5 ± 5.0	12.6 ± 7.4	83.6 ± 4.8	20.4 ± 5.8

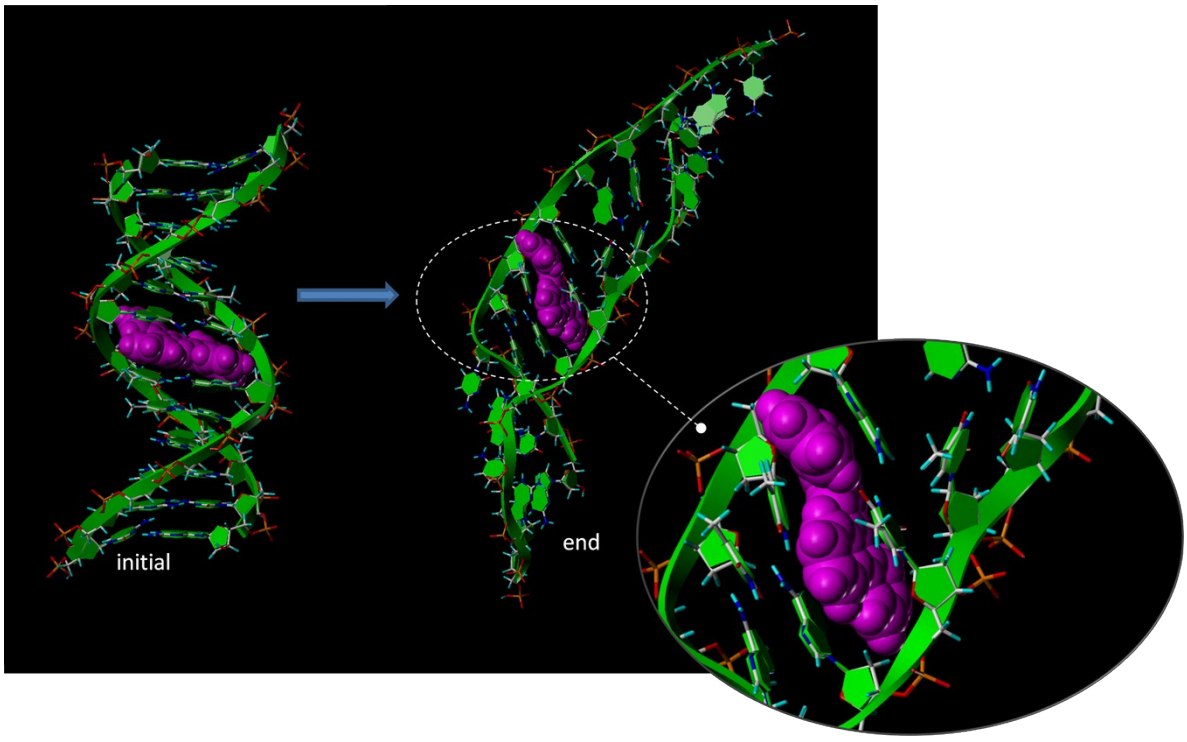


Figure 11S. Snapshots of the 1 ns MD trajectory at 300K showing the location of ligand PBM460 at time = 0 ns *i.e.,* starting minimized MBE structure upon major axis parallel to the pocked bases axis approaching, and at 1 ns from the beginning of the trajectory (right).

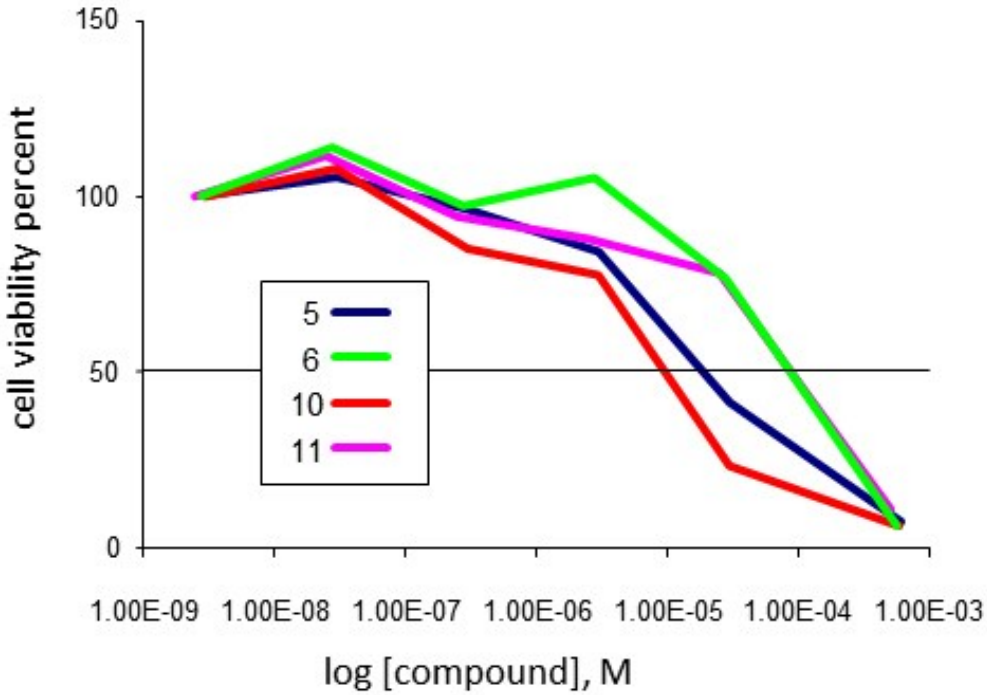


Figure 12S. Line graph of cytotoxicity profile (plot of cell viability percent vs molar concentration of compound in log scale) determined by the standard MTT assay using HeLa cells in culture treated with compounds **5**, **6**, **10** and **11** at different concentrations, after 24 hours of incubation. Each data point represents the mean of three experiments. The estimated apparent lethal dose 50% (LD50) is 2.2×10^{-4} M for compounds **6** and **11**, 1.5×10^{-5} M for **5** and 2.3×10^{-5} M for **10**.

References

[1] Lakowicz JR. Principles of Fluorescence Spectroscopy. 3rd ed. New York: Springer; 2006, Ch 10, p. 361.

[2] Lakowicz, JR. In *Principles of Fluorescence Spectroscopy;* 3rd ed. New York: Springer; 2006; Ch 4, p. 97.

[3] O'Connor DV, Ware WR, Andre JC. J Phys Chem 1979; 83:1333-43.

[4] Sybyl-X 2.0, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.

[5] Clark M, Cramer RD III, Van ON. J. Comput. Chem. 1989; 10, 982-1012.

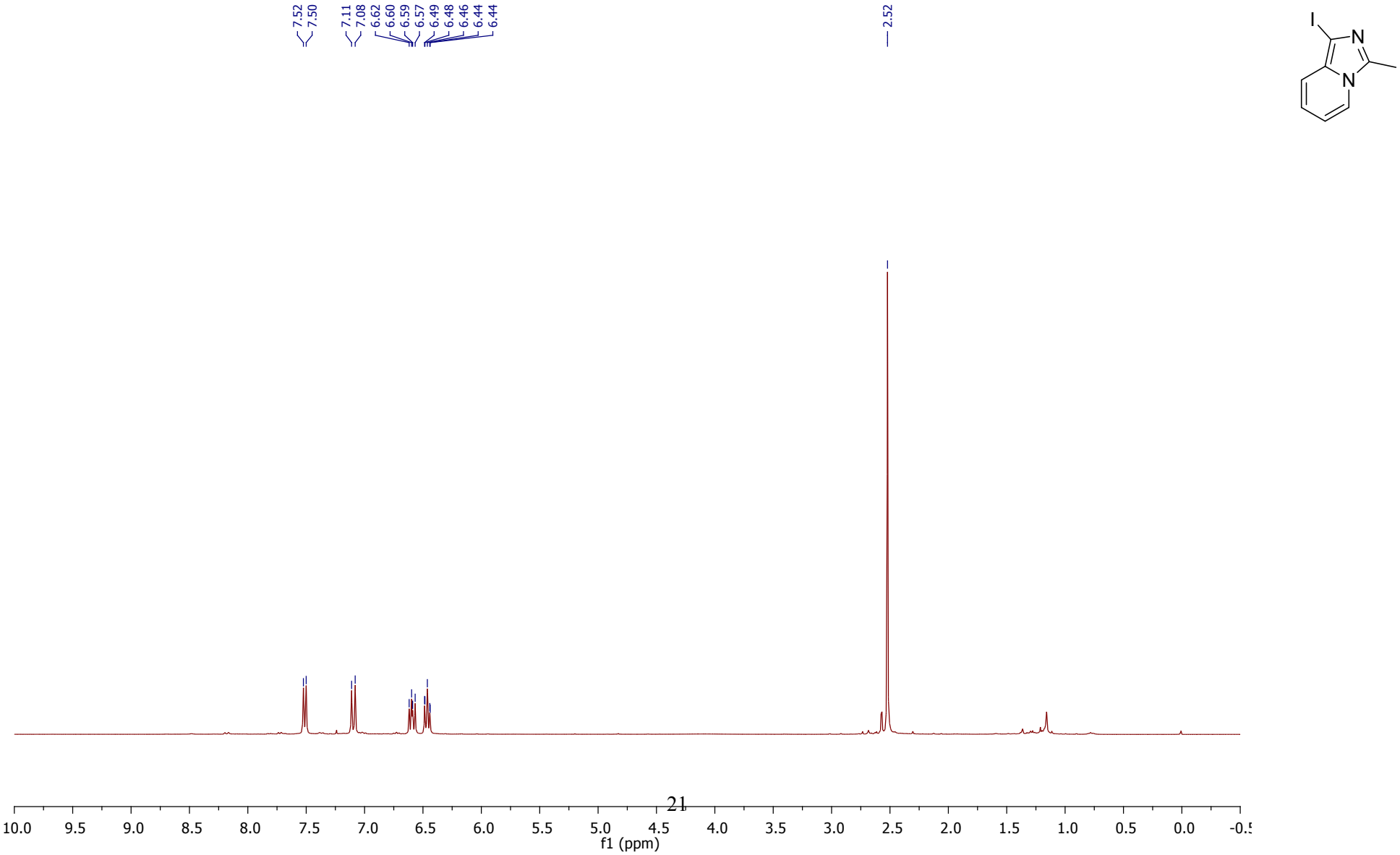
[6] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta J E, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand, J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ. Gaussian 09, Revision A.1. ed. Wallingford, CT, USA: Gaussian, Inc., 2009.

[7] Blanco M, J Comput Chem 1991; 12, 237-47

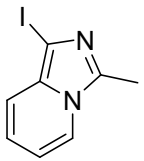
[8] Brunel Y, Faucher H, Gagnaire D, Rassat A, Tetrahedron, 1975; 31, 1075-91.

[9] (a) Bosch P, García V, Bilen BS, Sucunza D, Domingo A, Mendicuti F, Vaquero JJ. Dye & Pigments, 2017, 138, 135-146. (b) Suárez RM, Bosch P, Sucunza D, Cuadro AM, Domingo A, Mendicuti F, Vaquero JJ. Org Biomol Chem 2015; 13:527–38. (c) Abarca B, Custodio R, Cuadro AM, Sucunza D, Domingo A, Mendicuti F, J. Alvarez-Builla J, Vaquero JJ. Org Lett 2014, 16 (13): 3464–3467. (d) Carmona T, Cañeque T, Custodio R, Cuadro AM, Vaquero JJ, Mendicuti F. Dyes Pigments 2014; 103: 106-117.

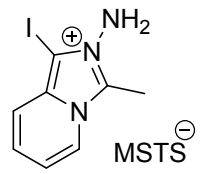
1-iodo-3-ethylimidazo[1,5-a]pyridine ^{1}H-RMN (1) (300 MHz, $CDCl_3$)



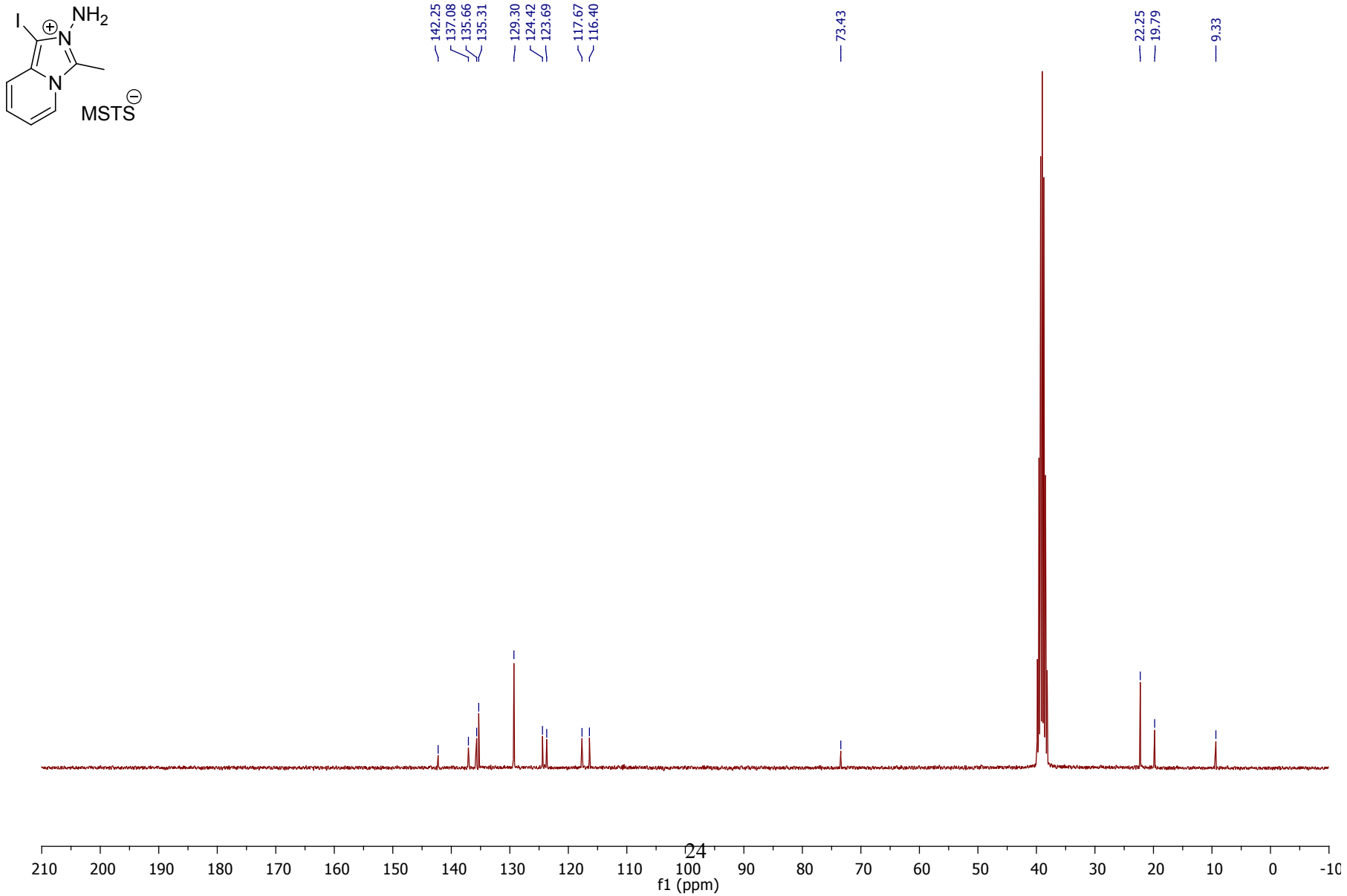
1-iodo-3-ethylimidazo[1,5-a]pyridine ^{13}C-RMN (1) (75 MHz, $CDCl_3$)



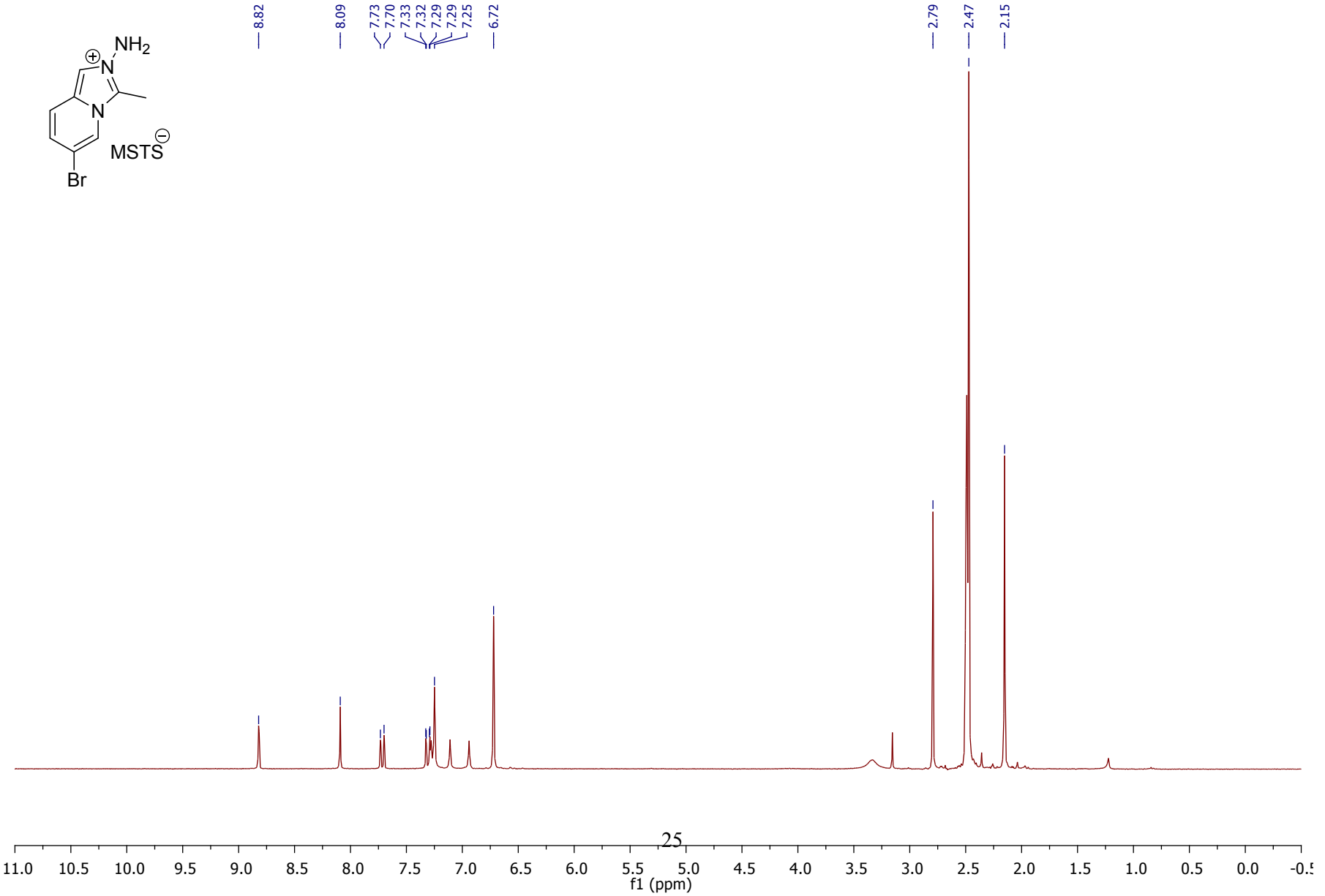
2-amino-1-iodo-3-methylimidazo[1,5-a]pyridin-2-ium 2,4,6-trimethylbenzenesulfonate ^{1}H-RMN (3) (300 MHz, DMSO d_6)



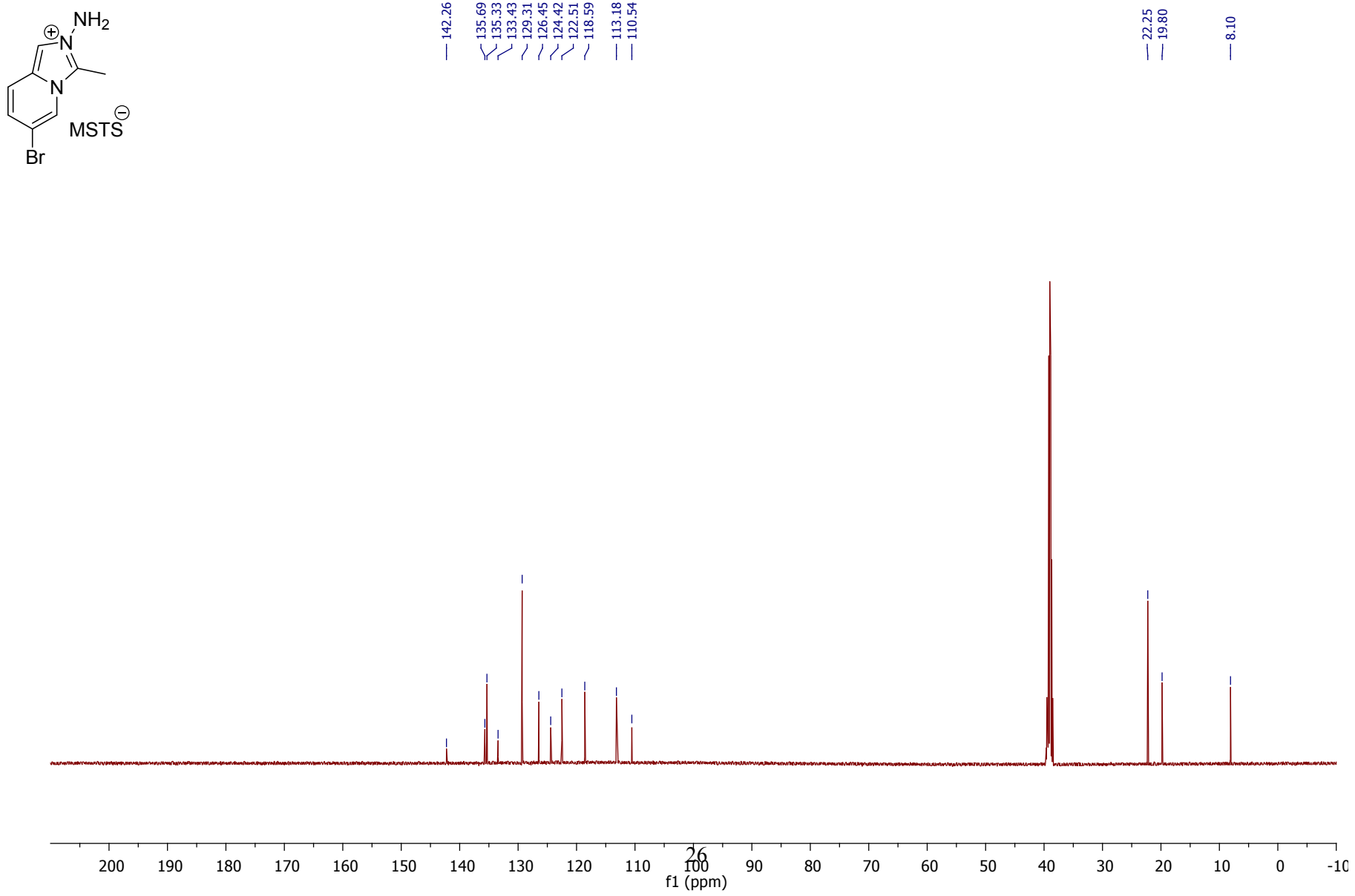
2-amino-1-iodo-3-methylimidazo[1,5-a]pyridin-2-ium 2,4,6-trimethylbenzenesulfonate ^{13}C-RMN (3) (75 MHz, DMSO d_6)



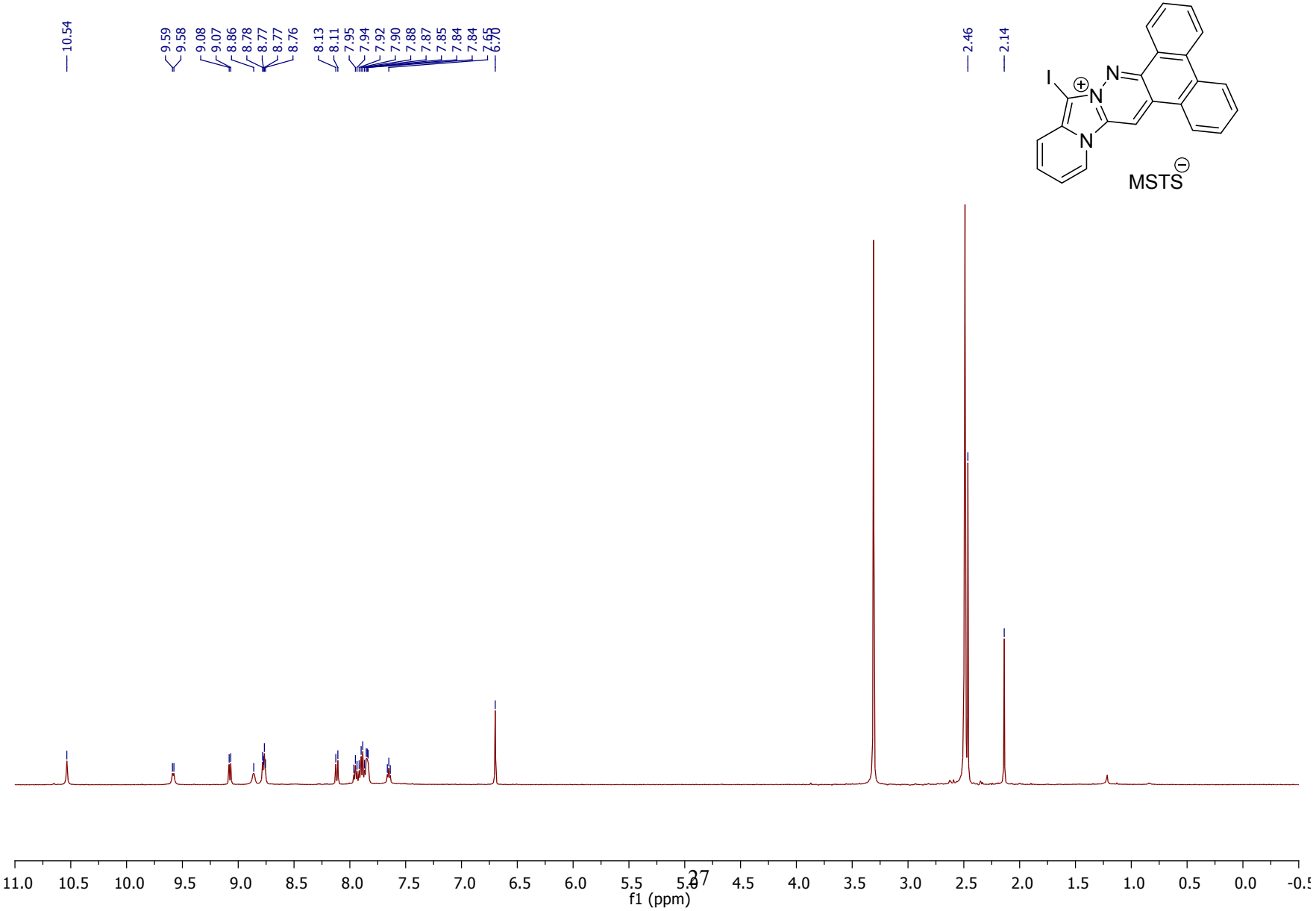
2-amino-6-bromo-3-methylimidazo[1,5-a]pyridin-2-ium 2,4,6-trimethylbenzenesulfonate ^{1}H-RMN (4) (300 MHz, DMSO d_6)



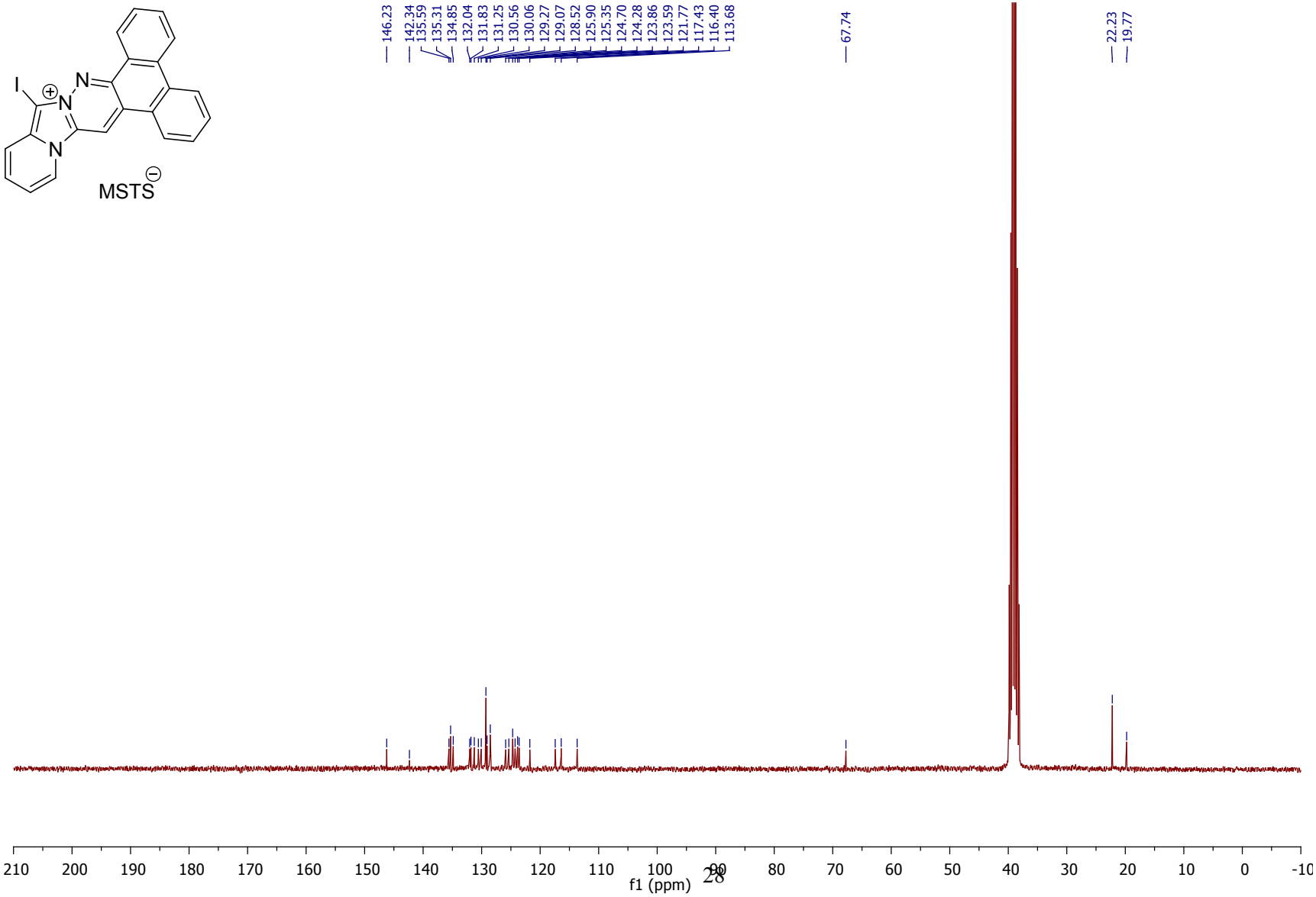
2-amino-6-bromo-3-methylimidazo[1,5-a]pyridin-2-ium 2,4,6-trimethylbenzenesulfonate ^{13}C-RMN (4) (126 MHz, DMSO d_6)



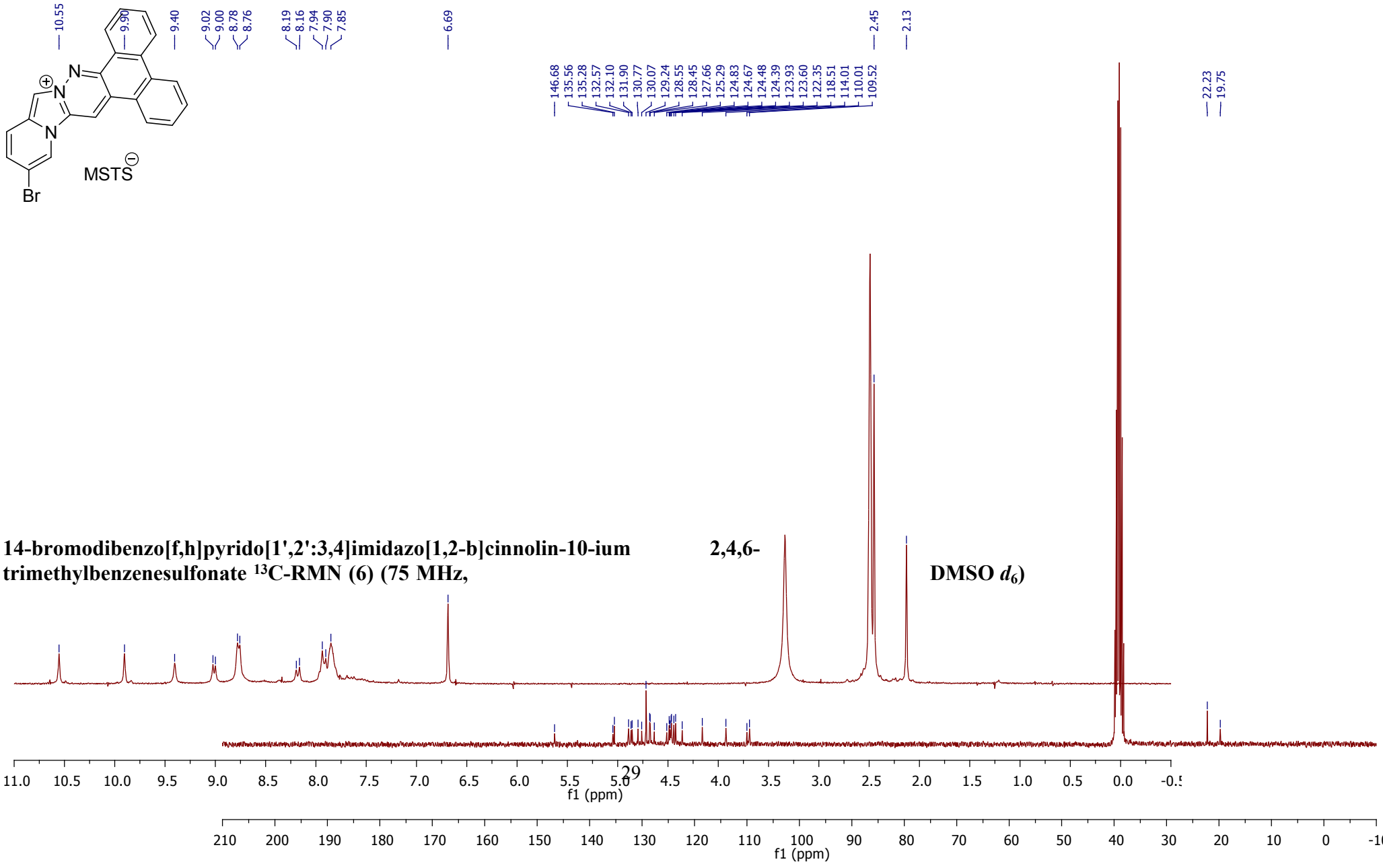
11-iododibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium 2,4,6-trimethylbenzenesulfonate ^{1}H-RMN (5) (500 MHz, DMSO d_6)



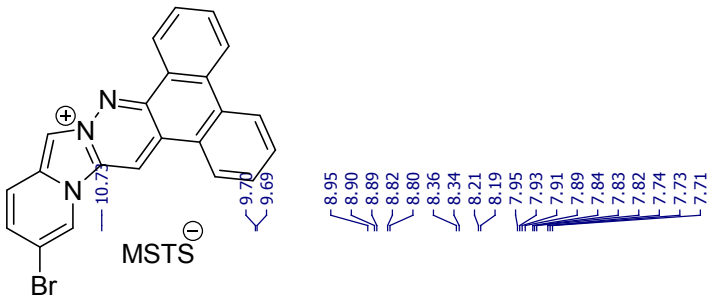
11-iododibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium 2,4,6-trimethylbenzenesulfonate ^{13}C-RMN (5) (75 MHz, DMSO d_6)



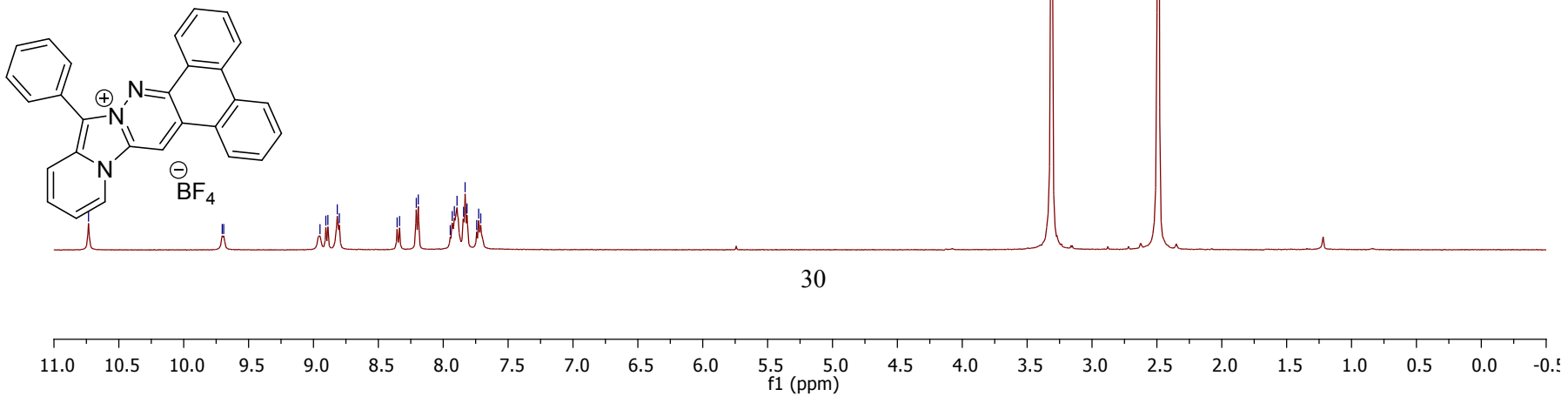
14-bromodibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium 2,4,6-trimethylbenzenesulfonate ^{1}H-RMN (6) (300 MHz, DMSO d_6)



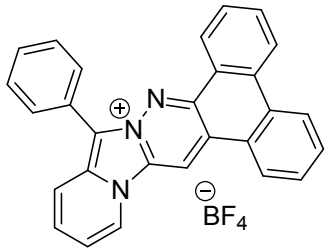
14-bromodibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium 2,4,6-trimethylbenzenesulfonate ^{13}C-RMN (6) (75 MHz, DMSO d_6)



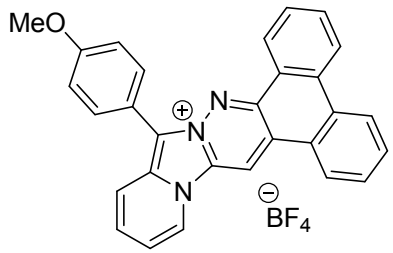
11-phenyldibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{1}H-RMN (7) (500 MHz, DMSO d_6)



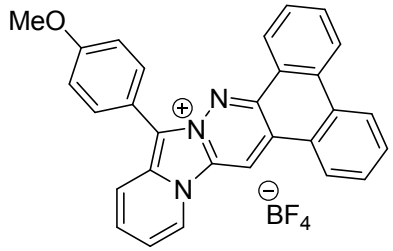
11-phenyldibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{13}C-RMN (7) (75 MHz, DMSO d_6)



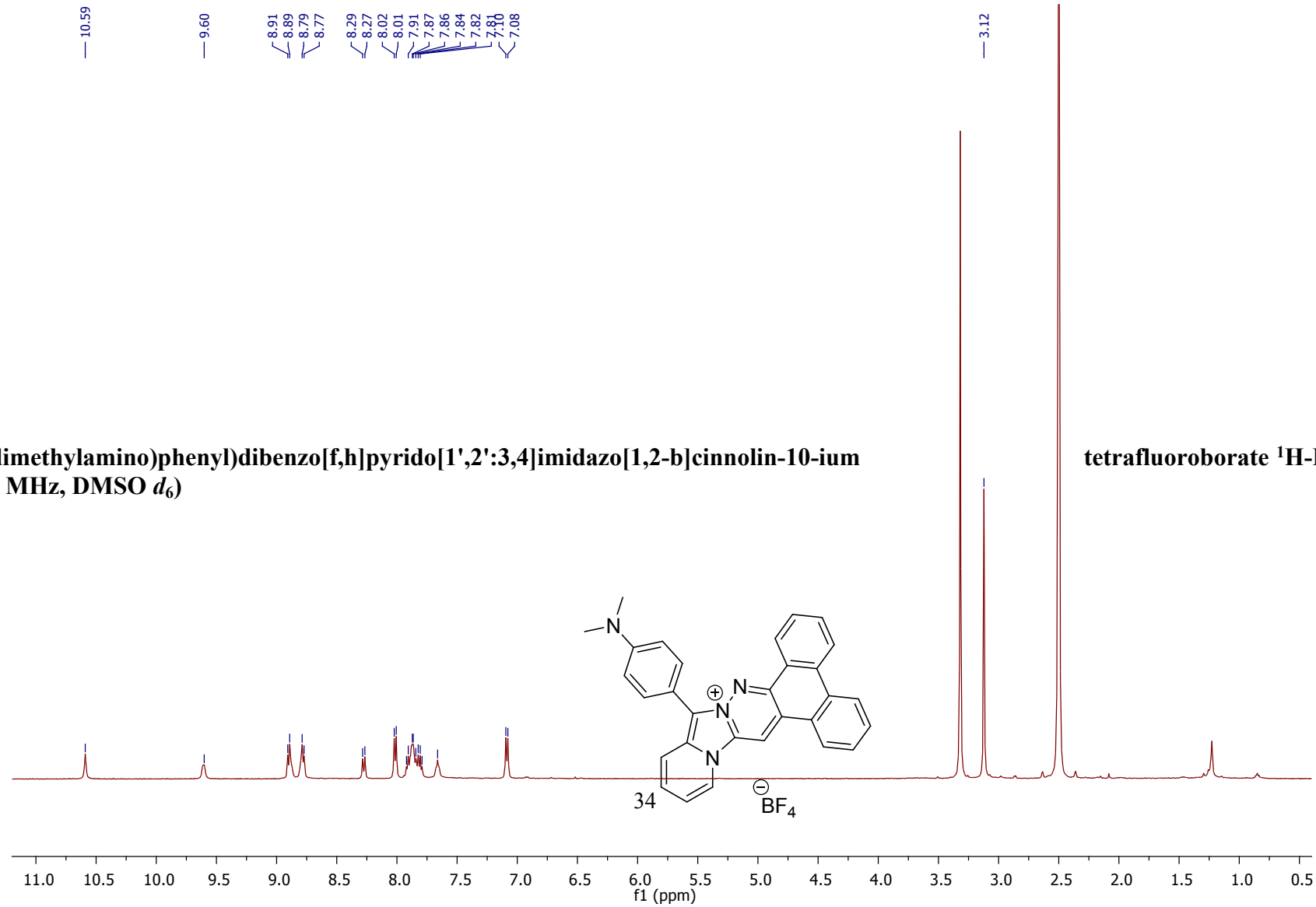
11-(4-methoxyphenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{1}H-RMN (8) (500 MHz, DMSO d_6)



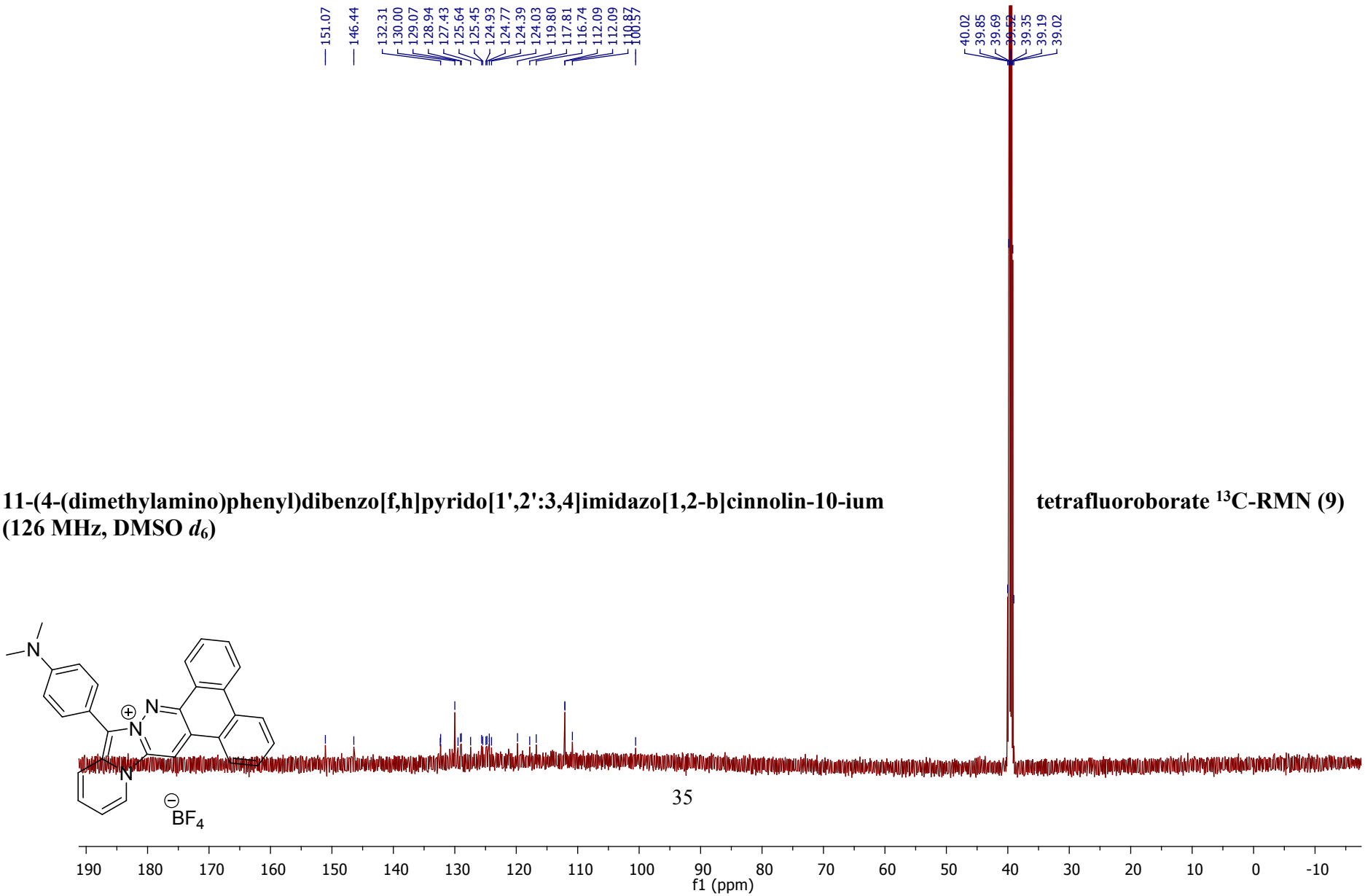
11-(4-methoxyphenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{13}C-RMN (8) (126 MHz, DMSO d_6)



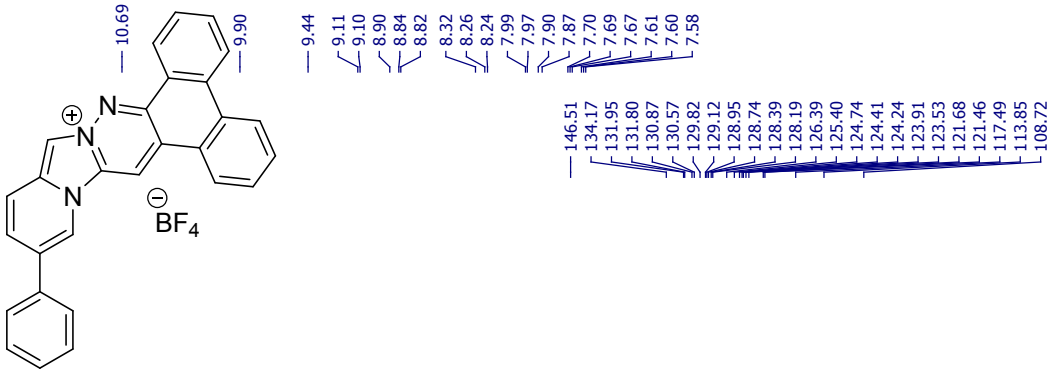
11-(4-(dimethylamino)phenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{1}H-RMN (9) (500 MHz, DMSO d_6)



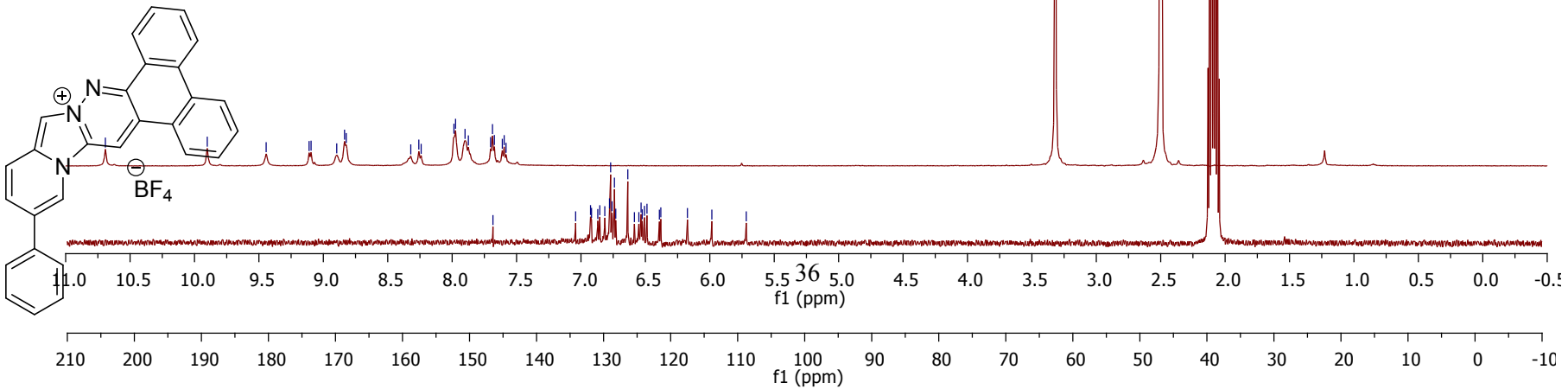
11-(4-(dimethylamino)phenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{13}C-RMN (9)
(126 MHz, DMSO d_6)



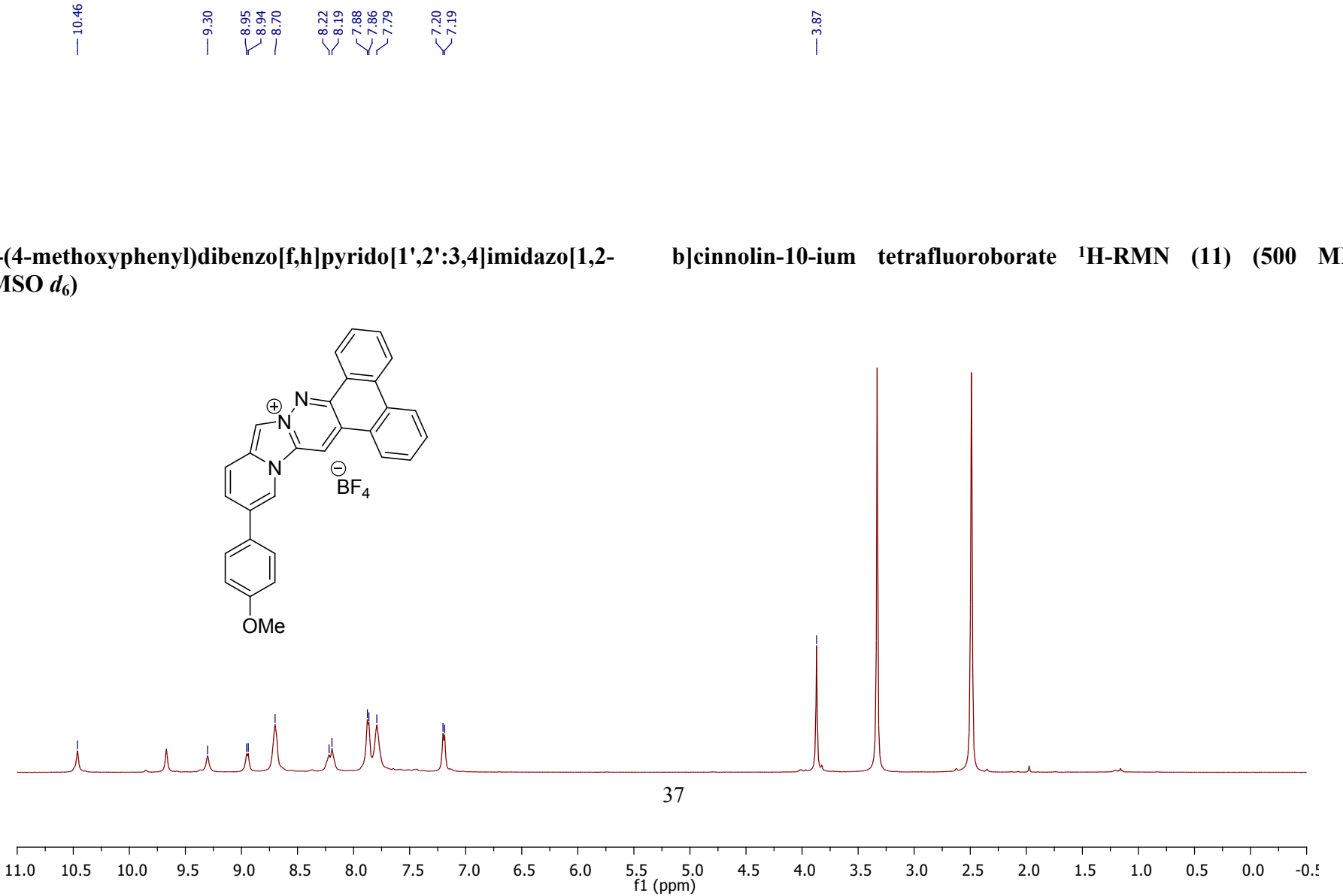
14-phenyldibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{1}H-RMN (10) (500 MHz, DMSO d_6)



14-phenyldibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{13}C-RMN (10) (75 MHz, DMSO d_6)



14-(4-methoxyphenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{1}H-RMN (11) (500 MHz, DMSO d_6)



14-(4-methoxyphenyl)dibenzo[f,h]pyrido[1',2':3,4]imidazo[1,2-b]cinnolin-10-ium tetrafluoroborate ^{13}C-RMN (11) (75 MHz, DMSO d_6)