Supporting Information

Design and characterization of reversible head-to-tail boronate-linked macrocyclic nucleotides

Mégane Debiais,[a] Alejandro Gimenez Molina,[a] Jean-Jacques Vasseur,[a] Sabine Müller,[b] Ivan Barvik,[c] Carine Baraguey [*,a] and Michael Smietana [*,a]

[a] *Institut des Biomolecules Max Mousseron, IBMM UMR 5247 CNRS, Université de Montpellier, ENSCM, 1919 route de Mende, 34095 Montpellier, (France).*

[b] *Institut für Biochemie, Ernst-Moritz-Arndt-Universität Greifswald, Felix-Hausdorff-Str. 4, D-17489 Greifswald, Germany*

[c] *Institute of Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 5, Prague 2, 121 16, Czech Republic.*

E-mail address:

Carine.baraguey@umontpellier.fr

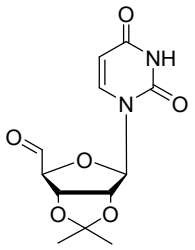
michael.smietana@umontpellier.fr

TABLE OF CONTENTS

1. Experimental part

Except THF that was distilled on calcium with benzophenone, all the anhydrous solvents and reagents were commercial and used without further purification. Thin-layer chromatographies (TLC) were made on silica plate 60 F_{254} Merck and revealed at 254 nm. Purifications on silica gel column chromatography were made with silica 40-63 µm Merck-Millipore. HRMS analysis were realised on Micromass Q-TOF spectrometer with an electrospray ionisation (ESI). NMR analysis were acquired at 25 °C on a Bruker Avance spectrometer operating at 400 MHz as well as on 600 MHz and 500 MHz Bruker Avance III spectrometers equipped with TCI and BBO cryo-probeheads respectively. Chemical shifts in organic solvents were reported in parts per million (ppm) referenced to the solvent residual peaks ($CDCl_3$ δ ^{1}H = 7.26 and δ ^{13}C = 77.16, CD_3OD δ ^{1}H = 3.31 and δ ^{13}C = 49.00, DMF-d_7 δ ^{1}H = 2.75 and δ ^{13}C = 29.76 for the high field methyl signal). Experiments in deuterated water were calibrated using DSS as reference for ^{1}H NMR. Indirect referencing was used for ^{13}C NMR using the ^{13}C/^{1}H ratio = 0.251449530.[1] The pD of the solutions was measured with a Mettler-Toledo micro electrode and was adjusted using solutions of NaOD in D_2O and taking into account that pD = $pH_{reading}$ + 0.4.[2] ^{1}H NMR data were reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, dt = doublet of triplets, m = multiplet, bs = broad singlet), coupling constant (*J*) in Hertz, integration and attribution. ^{11}B NMR data were acquired on the 400 MHz Bruker Avance spectrometer equipped with BBO probe head and the spectra were calibrated from BF_3.EtOEt as external reference. Diffusion Ordered Spectroscopy experiments (DOSY) were run with a sequence including longitudinal Eddy current delay with bipolar gradients and a double stimulated echo for convection compensation (sequence "dtebpgp3s" from the Bruker library). The magnetic field gradients were calibrated so that measuring the diffusion coefficient of water in D_2O at 25 °C yields a value of 1.902 x 10^{-9} m^2s^{-1}.[3] The T_1 relaxation times were measured for the sample in DMF before to run the DOSY to properly set the relaxation delay (3 s). The DOSY spectra were acquired with optimized gradient lengths of 2.6 and 2.8 ms (δ), a diffusion delay of 50 ms (Δ) and with gradient values linearly ranging from 2 to 95 % of the maximal gradient (69.2 G/cm) in 32 steps. Translational diffusion coefficients (D) were extracted by non-linear curve fitting of the intensity decay with increasing gradient strength using the T_1/T_2 processing tool of the Topspin software. The analysis of the intensity attenuation curves was repeated fort at least 3 resonances for each compound in order to get reliable average D values.

5'-aldehyde-2',3'-isopropylideneuridine (5)

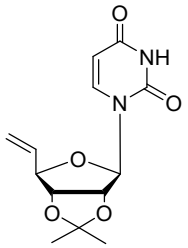


Chemical Formula: $C_{12}H_{14}N_2O_6$
Exact Mass: 282,09

2',3'-*O*-isopropylideneuridine (2.3 g, 8.1 mmol) was dissolved in 100 mL of acetonitrile and IBX (3.4 g, 12.0 mmol) was added (insoluble). The mixture was stirred at 80°C for 5 hours. The progress of the reaction was monitored by TLC. After return to room temperature, the solution was filtered on celite and evaporated under reduced pressure to lead to a white solid corresponding to the compound **5** (2.0 g, 88 %) which did not required purification. **R_f :** 0.55 (DCM/MeOH : 95/5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm) : 1.35 (s, 3H, $CH_{3\ IP}$), 1.52 (s, 3H, $CH_{3\ IP}$), 4.55 (bs, 1H, $H_{4'}$), 5.10 (d, *J* = 6.4 Hz, 1H, $H_{2'}$), 5.20 (d, *J* = 6.0 Hz, 1H, $H_{3'}$), 5.51 (s, 1H, $H_{1'}$), 5.77 (d, *J* = 8.0 Hz, 1H, H_5), 7.29 (d, *J* = 8.0 Hz, 1H, H_6), 9.42 (s, 1H, CHO),

9.74 (s, 1H, H_N). **^{13}C NMR** (125 MHz, $CDCl_3$) δ (ppm) : 24.9 ($CH_{3\ IP}$), 26.5 ($CH_{3\ IP}$), 83.8 ($C_{2'}$), 84.9 ($C_{3'}$), 94.2 ($C_{4'}$), 100.2 (C_5), 102.9 ($C_{1'}$), 113.7 ($C_{IV\ IP}$), 144.5 (C_6), 150.7 (C_2), 163.8 (C_4), 199.6 (CHO). **HRMS (ESI$^+$) :** *m/*z calculated for $C_{12}H_{14}N_2O_6$ $[M+H]^+$: 283.0930, found : 283.0931.

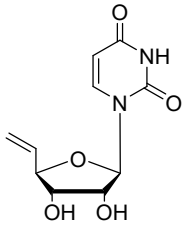
5'-deoxy-5'-vinyl-2',3'-isopropylideneuridine (6)



Chemical Formula: $C_{13}H_{16}N_2O_5$
Exact Mass: 280,11

The Nysted reagent (cyclo-dibromodi-μ-methylene[μ-(tetrahydrofuran)]trizinc, 16.2 mL, 7.08 mmol,) was dissolved in 20 mL of anhydrous THF and BF_3OEt_2 (1.75 mL, 14.0 mmol) was added dropwise at 0°C. The mixture was stirred at 0°C for 15 minutes and compound **5** (2.0 g, 7.08 mmol) dissolved in 5 mL of anhydrous THF was added dropwise. The mixture was stirred at room temperature for 4 hours. The progress of the reaction was followed by TLC. Sodium bicarbonate was added dropwise. 1-Butanol was then added and the organic phase was washed with sodium bicarbonate, dried on sodium sulphate, filtered and evaporated under reduced pressure. The crude material was purified by silica gel column chromatography (DCM/MeOH : 95/5) to lead to a white solid corresponding to the compound **6** (1.0 g, 50 %). **R_f :** 0.60 (EA/Cyclohexane 90/10). **^{1}H NMR** (400 MHz, CD_3OD) δ (ppm) : 1.35 (s, 3H, $CH_{3\ IP}$), 1.55 (s, 3H, $CH_{3\ IP}$), 4.46 (m, 1H, $H_{4'}$), 4.74 (dd, *J* = 6.4 Hz, 4.4 Hz, 1H, $H_{3'}$), 5.04 (dd, *J* = 6.4 Hz, 2.1 Hz, 1H, $H_{2'}$), 5.22 (dt, *J* = 10.4 Hz, 1.4 Hz, 1H, $H_{6'cis}$), 5.33 (dt, *J* = 17.2 Hz, 1.4 Hz, 1H, $H_{6'trans}$), 5.68 (d, *J* = 8.0 Hz, 1H, H_5), 5.78 (d, *J* = 2.1 Hz, 1H, $H_{1'}$), 6.00 (ddd, *J* = 17.2 Hz, 10.4 Hz, 7.2 Hz, 1H, $H_{5'}$), 7.59 (d, *J* = 8.0 Hz, H_6). **^{13}C NMR** (125 MHz, CD_3OD) δ (ppm) : 25.6 ($CH_{3\ IP}$), 27.5 ($CH_{3\ IP}$), 85.5 ($C_{3'}$), 85.8 ($C_{2'}$), 89.7 ($C_{1'}$), 94.5 ($C_{4'}$), 102.8 (C_5), 115.5 ($C_{IV\ IP}$), 118.4 ($C_{6'}$), 136.9 ($C_{5'}$), 144.6 (C_6), 151.9 (C_2), 166.2 (C_4). **HRMS (ESI$^+$) :** *m/*z calculated for $C_{13}H_{16}N_2O_5$ $[M+H]^+$: 281.1137, found : 281.1142.

5'-deoxy-5'-vinyl uridine (7)

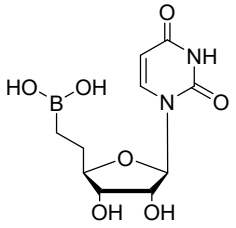


Chemical Formula: $C_{10}H_{12}N_2O_5$
Exact Mass: 240,07

Compound **6** (1.5 g, 5.35 mmol) was dissolved in 60 mL of milli-Q water and TCA (8.74 g, 54.0 mmol) was added. The mixture was stirred at room temperature for 4 hours. The progress of the reaction was monitored by TLC. Water was evaporated under reduced pressure and the crude material was purified by silica gel column chromatography (DCM/MeOH : 90/10) to lead to a white solid corresponding to the compound **7** (1.2 g, 93 %). **R_f :** 0.30 (DCM/MeOH : 90/10). **^{1}H NMR** (400 MHz, CD_3OD) δ (ppm) : 3.96 (t, *J* = 5.8 Hz, 1H, $H_{3'}$), 4.20 (dd, *J* = 5.4 Hz, 3.9 Hz, 1H, $H_{2'}$), 4.34 (tt, *J* = 6.4 Hz, 1.4 Hz, 1H, $H_{4'}$), 5.28 (dt, *J* = 10.5 Hz, 1.3 Hz, 1H, $H_{6'cis}$), 5.40 (dt, *J* = 17.2 Hz, 1.4 Hz, 1H, $H_{6'trans}$), 5.74 (d, *J* = 8.4 Hz, 1H, H_5), 5.83 (d, *J* = 3.8 Hz, 1H, $H_{1'}$), 6.03 (ddd, *J* = 17.2 Hz, 10.5 Hz, 6.6 Hz, 1H, $H_{5'}$), 7.59 (d, *J* = 8.4 Hz, 1H, H_6).

^{13}C NMR (125 MHz, CD_3OD) δ (ppm) : 75.0 ($C_{3'}$ + $C_{2'}$), 85.6 ($C_{4'}$), 91.9 ($C_{1'}$), 102.9 (C_5), 118.4 ($C_{6'}$), 136.7 ($C_{5'}$), 142.4 (C_6), 152.2 (C_2), 166.1 (C_4). **HRMS (ESI$^-$) :** *m/z* calculated for $C_{10}H_{12}N_2O_5$ [M-H]$^-$: 239.0668 found : 239.0669.

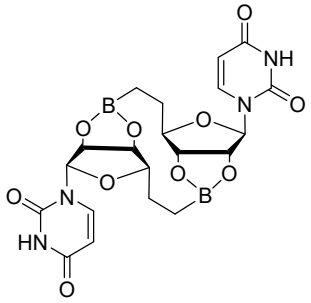
1-(5'-deoxyuridin-5'-yl)methylboronic acid (3)



Chemical Formula: $C_{10}H_{15}BN_2O_7$
Exact Mass: 286,10

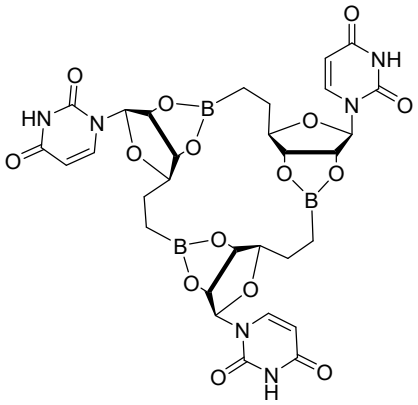
To 5 mL of anhydrous diethylic ether were added dropwise borane dimethysulfide (3.0 mL, 15.0 mmol) and α-pinene (4.8 mL, 30.0 mmol). The mixture was stirred at 35°C for 4 hours. Compound **7** (450.0 mg, 1.9 mmol) dissolved in anhydrous THF was then added dropwise and the resulting mixture was stirred at room temperature for 15 hours. The progress of the reaction was monitored by TLC. The reaction was then quenched by addition of 1 mL of hydrochloric acid 0.1 M and diluted with ethyl acetate. The organic layer was washed with sodium chloride, dried on sodium sulphate, filtered and evaporated under reduced pressure. The crude material was purified by silica gel column chromatography (DCM/MeOH : 85/15) to lead to a white solid corresponding to the compound **3** (170.0 mg, 32 %). **R_f :** 0.25 (DCM/MeOH : 85/15). **^{1}H NMR** (37 mM, 298 K, D_2O, pD 6.5, 400 MHz) δ (ppm) : 7.67 (d, *J* = 8.1 Hz, 1H, H_6), 5.89 (d, *J* = 8.1 Hz, 1H, H_5), 5.85 (d, *J* = 4.6 Hz, 1H, $H_{1'}$), 4.33 (t, *J* = 5.0 Hz, 1H, $H_{2'}$), 4.03 (t, *J* = 5.4 Hz, 1H, $H_{3'}$), 3.97 (dt, *J* = 7.8 Hz, 5.4 Hz, 1H, $H_{4'}$), 1.87 (m, 1H, $H_{5'or\ 5''}$), 1.77 (m, 1H, $H_{5'or\ 5''}$), 0.90 (t, *J* = 8.2 Hz, 2H, $H_{6'/6''}$). **^{13}C NMR** (37 mM, 298 K, D_2O, 125 MHz) δ (ppm) : 166.1 (Cq, C_4), 151.6 (Cq, C_2), 141.9 (CH, C_6), 102.4 (CH, C_5), 89.5 (CH, $C_{1'}$), 85.5 (CH, $C_{4'}$), 73.5 (CH, $C_{2'}$), 72.6 (CH, $C_{3'}$), 27.3 (CH_2, $C_{5'}$), 10.1 (CH_2, $C_{6'}$). **^{11}B NMR** (37 mM, 298 K, D_2O, 128 MHz) δ (ppm) : 32.5 ppm. **HRMS (ESI$^+$) :** *m/z* calculated for $C_{10}H_{16s}N_2O_7$ [M+H]$^+$: 287.10469, found : 287.10509.

Head-to-tail homodimer of 3 (htt_2)



^{1}H NMR (37 mM, 298 K, DMF-d_7, 600 MHz) δ (ppm) : 11.40 (bs, 1H, H_{N3}), 7.85 (d, *J* = 8.0 Hz, 1H, H_6), 5.71 (d, *J* = 8.0 Hz, 1H, H_5), 5.88 (d, *J* = 3.2 Hz, 1H, $H_{1'}$), 5.19 (m, 1H, $H_{2'}$), 4.81 (m, 1H, $H_{3'}$), 4.05 (q, *J* = 6.2 Hz, 1H, $H_{4'}$), 1.93 (m, 2H, $H_{5'/5''}$), 1.10-0.95 (m, 2H, $H_{6'/6''}$). **^{13}C NMR** (37 mM, 298 K, DMF-d_7, 125 MHz) δ (ppm) : 163.6 (Cq, C_4), 151.0 (Cq, C_2), 142.9 (CH, C_6), 102.6 (CH, C_5), 92.4 (CH, $C_{1'}$), 87.4 (CH, $C_{4'}$), 84.9 (CH, $C_{2'}$), 84.6 (CH, $C_{3'}$), 27.7 (CH_2, $C_{5'}$), 6.3 (CH_2, $C_{6'}$). **^{11}B NMR** (37 mM, 298 K, DMF, 128 MHz) : very large peak around 34 ppm.

Head-to-tail homotrimer of 3 (htt_3)



^{1}H NMR (37 mM, 298 K, DMF-d_7, 600 MHz) δ (ppm) : 7.81 (d, J = 8.0 Hz, 1H, H_6), 5.72 (d, J = 8.0 Hz, 1H, H_5), 5.86 (d, J = 2.5 Hz, 1H, $H_{1'}$), 5.18 (m, 1H, $H_{2'}$), 4.80 (m, 1H, $H_{3'}$), 3.90 (q, J = 6.4 Hz, 1H, $H_{4'}$), 1.91 (m, 2H, $H_{5'/5''}$), 1.10-0.95 (m, 2H, $H_{6'/6''}$). **^{13}C NMR** (37 mM, 298 K, DMF-d_7, 125 MHz) δ (ppm) : 163.6 (Cq, C_4), 150.9 (Cq, C_2), 143.2 (CH, C_6), 102.6 (CH, C_5), 92.9 (CH, $C_{1'}$), 88.1 (CH, $C_{4'}$), 85.3 (CH, $C_{2'}$), 83.3 (CH, $C_{3'}$), 27.3 (CH_2, $C_{5'}$), 6.2 (CH_2, $C_{6'}$). **^{11}B NMR** (37 mM, 298 K, DMF-d_7, 128 MHz) : very large peak around 34 ppm.

2. Oligonucleotides synthesis

All the natural oligonucleotides used were synthesized with an automatic DNA synthesizer Applied Biosystems 394 by classical phosphoramidite chemistry with commercial phosphoramidites and controlled-pore glass (lcaa-CPG) linked to 5'-*O*-DMTr-U or 5'-*O*-DMTr-dT through a 3'-*O*-succinyl linker. Analytical and semi-preparative high performance liquid chromatographies (HPLC) in reverse phase were made on a HPLC Dionex 600 system with UV detection at 260 nm and EC 75/4.6 Nucleodur 100-3 C18 (analytical) or VP 250/10 Nucleodur C18 HTec 5 µm (semi-preparative) columns. Buffers used for the analytical HPLC were: 100 % TEAAc 0.05 M (buffer A) and 100 % ACN (buffer B) with a flow of 1 mL/min. Buffers used for the semi-preparative HPLC were : 1 % ACN in a solution of TEAAc 0.05 M (buffer A) and 80 % ACN in a solution of TEAAc 0.05 M (buffer B) with a flow of 4 mL/min. MALDI-TOF spectra were acquired on an Axima assurance spectrometer (Shimadzu Biotech) with a N_2 laser (337 nm). The matrix used was THAP or ATT in a solution of ACN/Ammonium citrate 0.1 M (1/1). The sample was mixed with the matrix with a ratio of 1/5. Oligonucleotides dosages were made on a Varian Cary 300 Bio UV spectrometer with a UV detection at 260 nm.

General procedure for linear oligonucleotides synthesis[4,5]

RNA oligonucleotides **U_2** and **U_3** were prepared by automated synthesis on a 8 and 12 µmol scale respectively using commercial phosphoramidites with a 2'-*O*-PivOM protecting group and 3'-phosphate CPG solid supports under the conditions shown in Table S1. A trichloroacetic acid solution in DCM was used as detritylation reagent, 5-benzylmercapotetrazole (BMT) as activator, a mixture of phenoxyacetic anhydride in THF/pyridine and *N*-methylimidazole in THF as capping solution and an iodine solution as oxidizing agent.

Table S1 : Standard conditions for oligonucleotides synthesis

Step	Reaction	Reagent	Time (s)
1	Detritylation	3 % TCA in DCM	35
2	Coupling	0.1 M amidite in ACN + 0.3 M BMT in ACN	20
3	Capping	Ac_2O/THF/Pyridine + 10 % NMI in THF	8
4	Oxidation	0.1 M I_2/THF/H_2O/Pyridine	15

General procedure for cyclic oligonucleotides synthesis

c(T_2) and **c(T_3)** were synthesized according to the classical phosphoramidite chemistry with commercial phosphoramidites and 3'-phosphate CPG solid supports. The first cycle required a supplementary step of sulfurization with Beaucage reagent (3-phenyl-1,2,4-dithiazoline-5-one) after the first phosphoramidite incorporation (Table S2). Next cycles were executed according to the standard conditions. The column was then percolated with 1 mL of a solution of $(PhO)_3PCH_3I$ (0.5 M in anhydrous DMF) at room temperature for 15 minutes, successively washed with 15 mL of DCM, 10 mL of ACN, 10 mL of DCM and dried with argon. The deprotection and cyclization were both achieved at the same time.

Table S2 : Conditions for insertion of a phosphorothioate in the sequence.

Step	Reaction	Reagent	Time (s)
1	Detritylation	3 % TCA in DCM	35
2	Coupling	0.1 M amidite in ACN + 0.3 M BMT in ACN	20
3	Capping	Ac_2O/THF/Pyridine + 10 % NMI in THF	8
4	Sulfurisation	0.05 M Beaucage reagent in ACN	60
5	Oxidation	0.1 M I_2/THF/H_2O/Pyridine	15

General procedure for deprotection, analysis and purification of oligonucleotides

DNA oligonucleotides were deprotected in a solution of ammonia 28 % at 55°C for 5 hours. RNA oligonucleotides were percolated for 3 minutes in a solution of DBU 1 M in ACN, dried with argon and then deprotected in a solution of ammonia 28 % at 37°C for 3 hours. Ammonia was then evaporated in SpeedVac and oligonucleotides were analysed by analytical HPLC and MALDI-TOF with a ATT matrix and ammonium citrate as co-matrix in negative mode. Oligonucleotides were then purified by semi-preparative HPLC (gradient 0-20-20 for linear **U_2** and **U_3**, 0-25-15 for cyclic **c(T_2)** and **c(T_3)).**

Table S3: *m/z* obtained with MALDI-TOF analysis for pure ONs

Name	Yield (µmol)	*m/z* calcd.[a]	*m/z* found[a]
U_2	2.79 (35 %)[b]	549.36	549.24
U_3	2.46 (12 %)[c]	855.53	854.89
c(T_2)	2.33 (29 %)[b]	623.45	623.45
c(T_3)	2.76 (35 %)[b]	927.64	927.57

[a] *m/z* calculated for $[M-H]^-$. [b] Isolated yields based on 8 µmol scale. [c] Isolated yields based on 12 µmol scale.

3. Estimation of the macrocycles size from diffusion coefficients

The translation diffusion coefficients can be connected to the molecular size through the Stoke-Einstein equation:

$$D = \frac{k_B T}{6\pi\eta f r_h} \quad (1)$$

where T is the sample temperature, k_B is the Boltzman constant, η is the solvent viscosity, r_H the hydrodynamic radius.

For small to medium-size molecules moreover, the equation denominator is overestimated and have to be correct using the *f* factor proposed by Gierer and Wirtz :[6]

$$f_{GW} = \left(\frac{3r_s}{2r} + \frac{r}{r + r_s}\right)^{-1} \quad (2)$$

where r and r_s and the solute and solvent hydrodynamic radii respectively.

It is then possible to relate the molecular weight of the solute *i* (MW_i) to its hydrodynamic radius according to the following expression:

$$r_i = \sqrt[3]{\frac{3 * MW_i}{4\,\pi\, N_A \rho}} \quad (3)$$

where N_A is Avogadro's number and ρ is the solute density.[7]

Considering the polymeric nature of the different assemblies, one can estimate the molecular weight of each oligomer from the ratio of the diffusion coefficients. Assuming that both the monomer and the cyclic species do have spherical shape and relatively close densities and f factors (calculation of f according to equation (2) for the monomer, dimer and trimer respectively gave the following values for f: 0.63, 0.68 and 0.71), the dependence of D on the molecular weight for two molecules can be written :[7]

$$\frac{D_1}{D_2} = \sqrt[3]{\frac{MW_2}{MW_1}} \quad (4)$$

Table S4: ^{1}H-^{1}H Vicinal coupling constants (in Hz).

	Uridine		**3**		
	D_2O [a]	D_2O	Wet DMF-d_7	Anhydrous DMF-d_7 Dimer [b]	Trimer [c]
$^3J_{1'-2'}$	4.4	4.7	5.3	3.2	2.5
$^3J_{2'-3'}$	5.3	5.0	5.3	7.8	nd
$^3J_{3'-4'}$	5.5	5.4	nd	5.1	5.6

[a] from Blackburn B. J. et al.[8] [b] measured on the diluted sample where the trimer **htt$_3$** was almost undetectable [c] measured on the concentrated sample (37 mM).

Table S5: Exocyclic angles $\phi_{i\text{-}j}$ (H_i-C_i-C_j-$H_{j'}$) in degrees calculated for the different models.

	NMR (single model)[a]	QM[b]
$\phi_{1'2'}$	123.0	117.9
$\phi_{2'3'}$	6.2	21.5
$\phi_{3'4'}$	-134.4	-161.2

[a] for P = 52 and ϕ_m = 9°; [b] MP2/6-31G(d)

4. NMR spectra

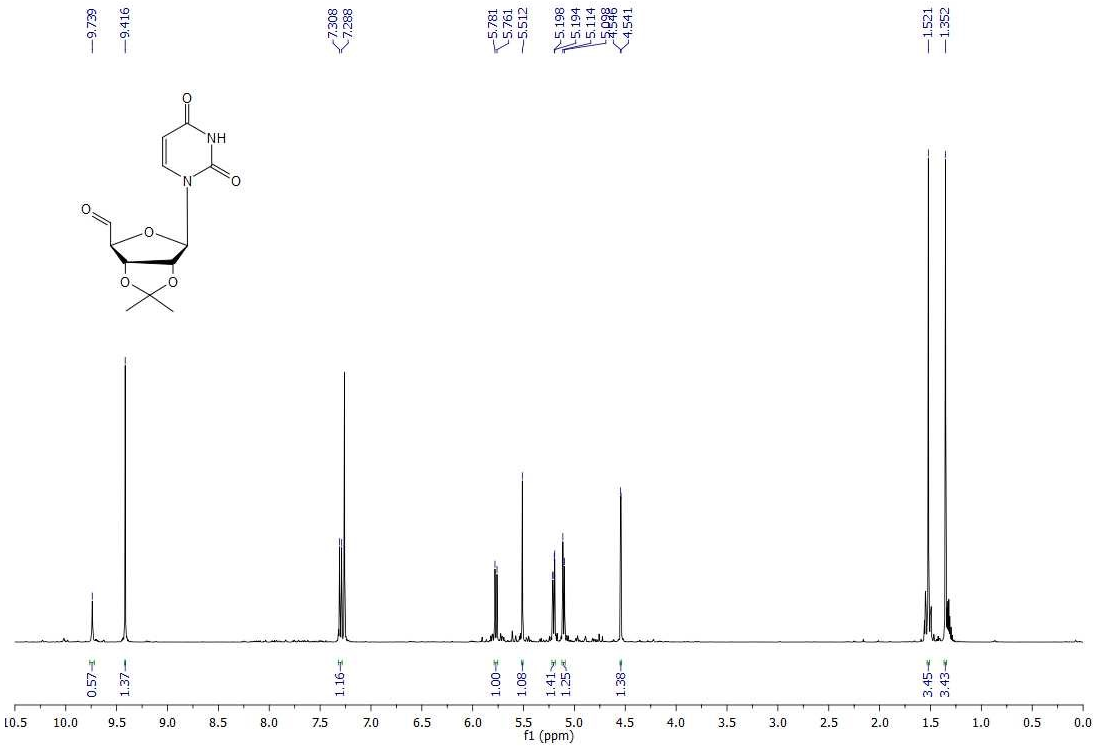


Figure S1 : ^{1}H NMR spectrum of **5** in $CDCl_3$ (400 MHz)

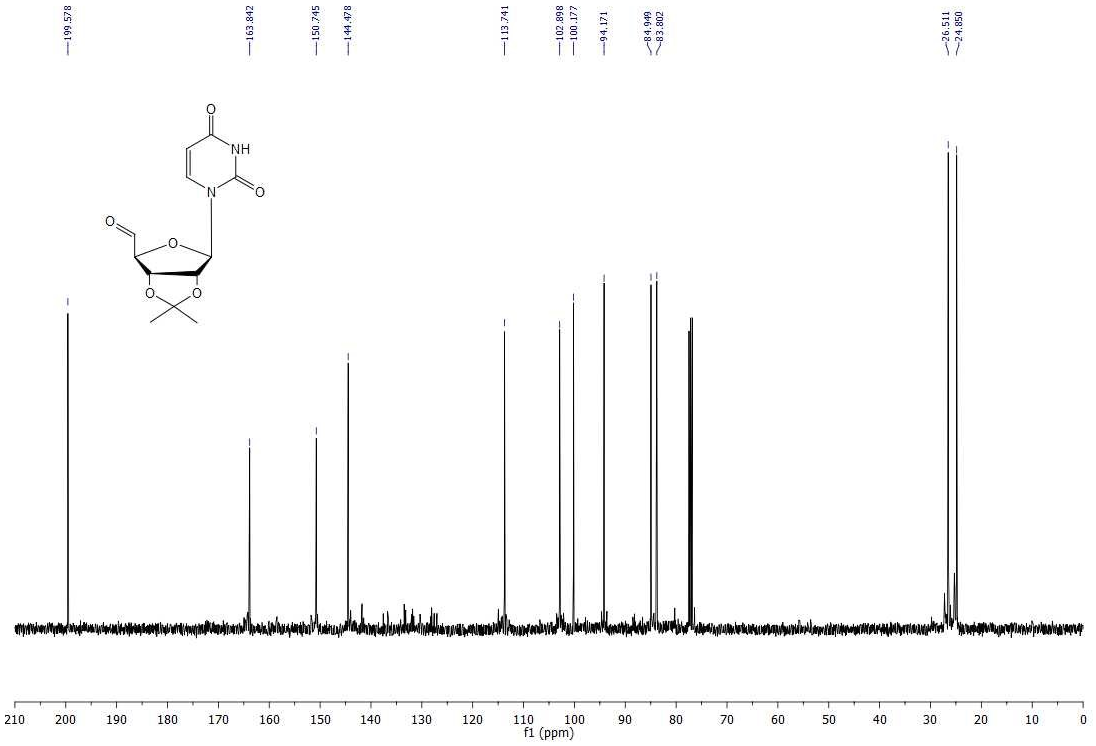


Figure S2 : ^{13}C NMR spectrum of **5** in $CDCl_3$ (125 MHz)

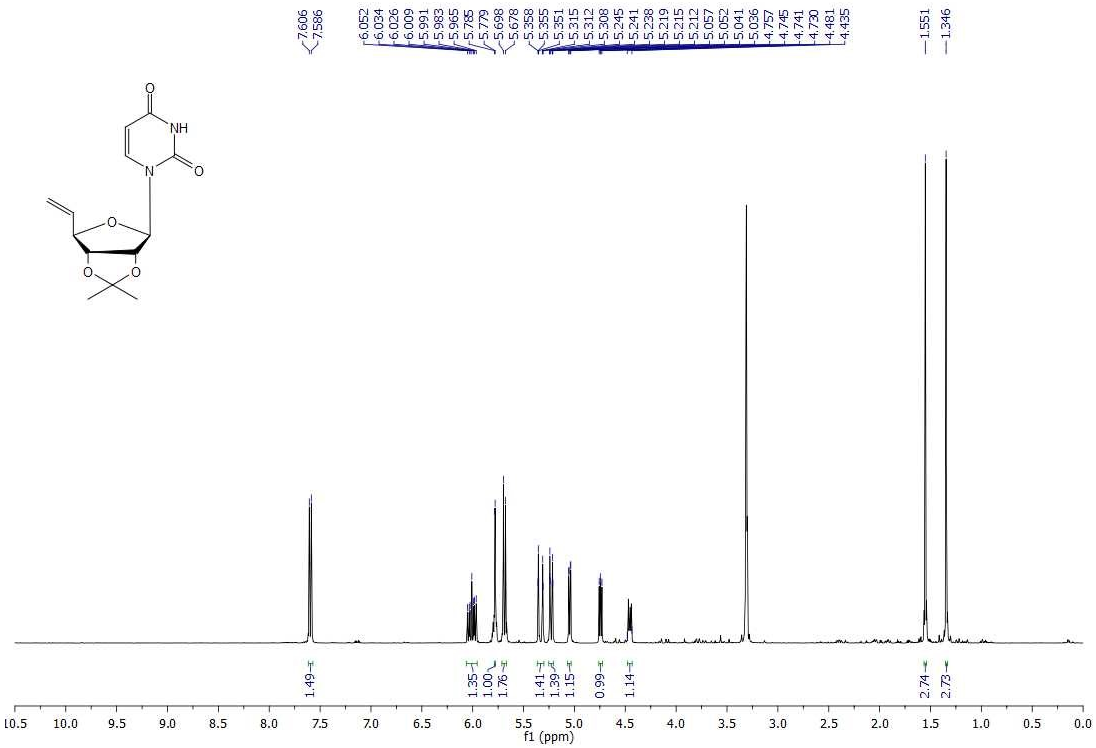


Figure S3 : ^{1}H NMR spectrum of **6** in CD_3OD (400 MHz)

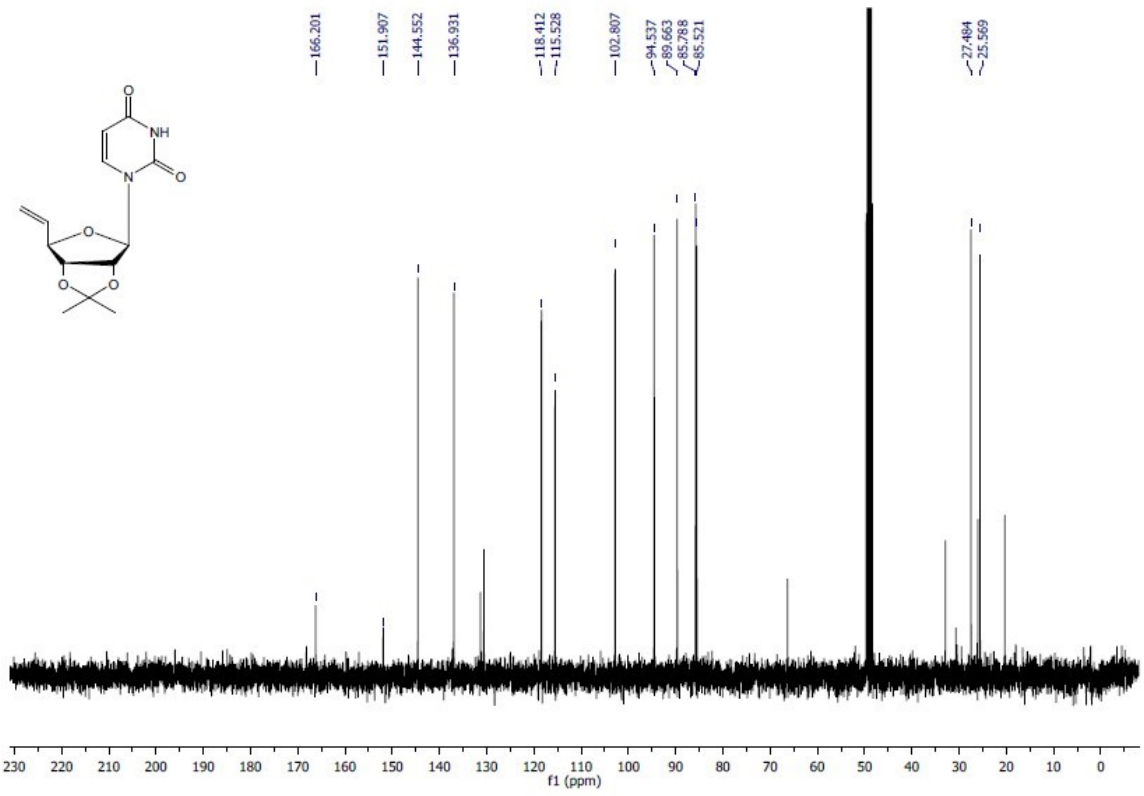


Figure S4 : ^{13}C NMR spectrum of **6** in CD_3OD (125 MHz)

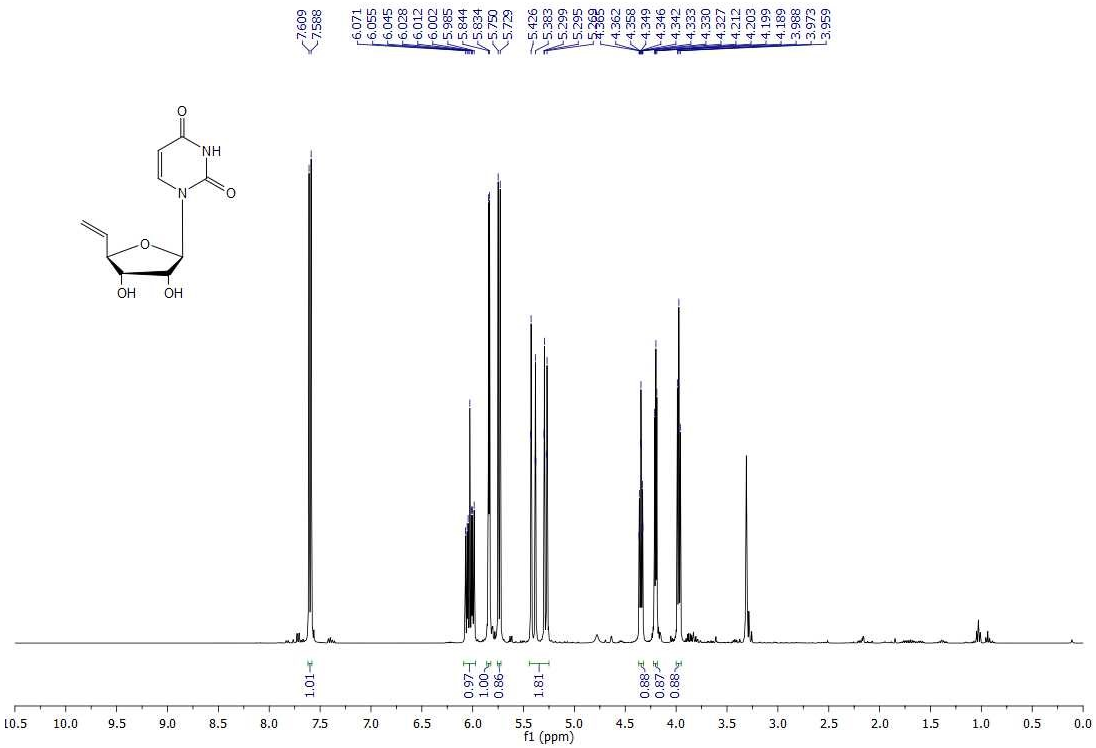


Figure S5 : ^{1}H NMR spectrum of **7** in CD_3OD (400 MHz)

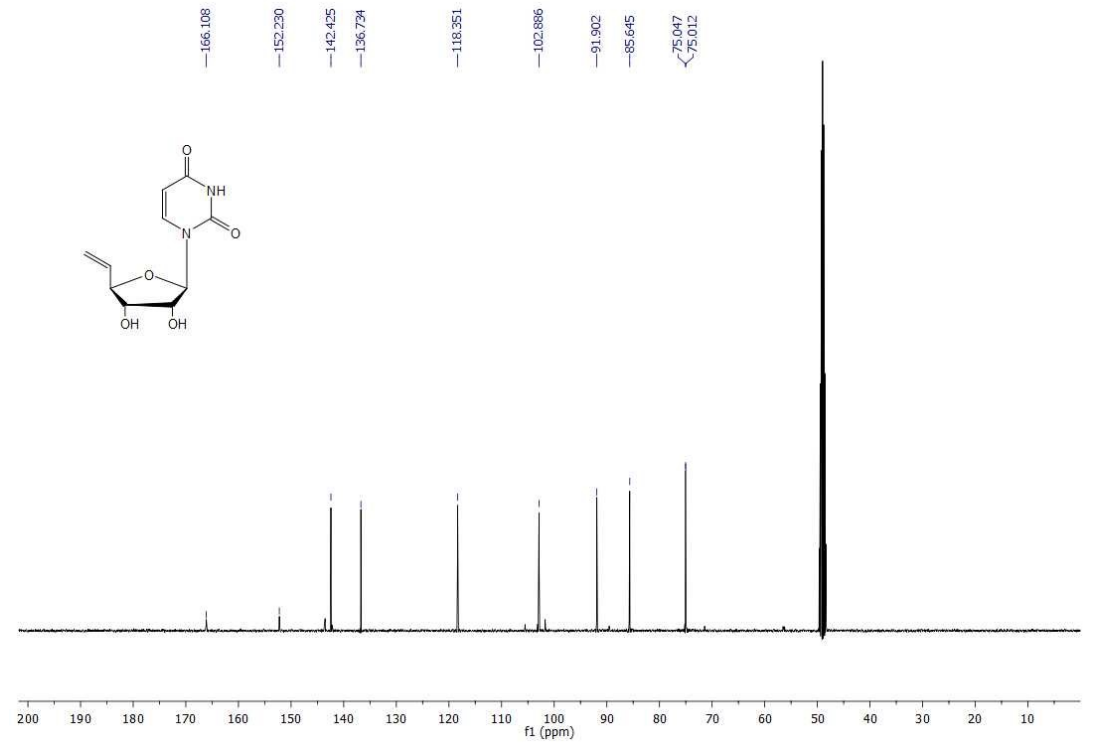


Figure S6 : ^{13}C NMR spectrum of **7** in CD_3OD (125 MHz)

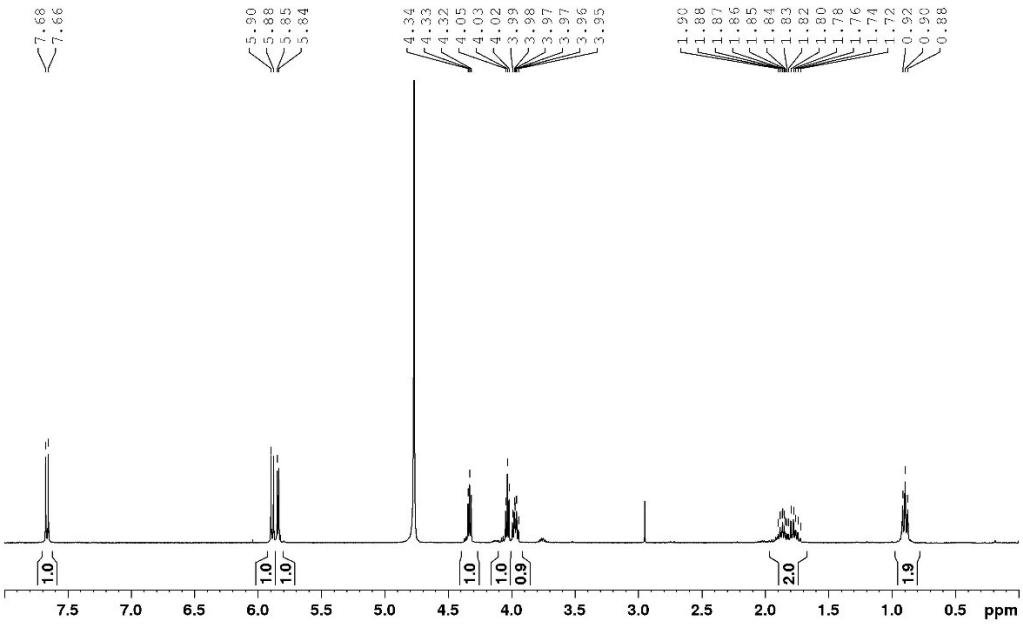


Figure S7 : ^{1}H NMR spectrum of **3** in D_2O (pD 6.5, 400 MHz, 37 mM)

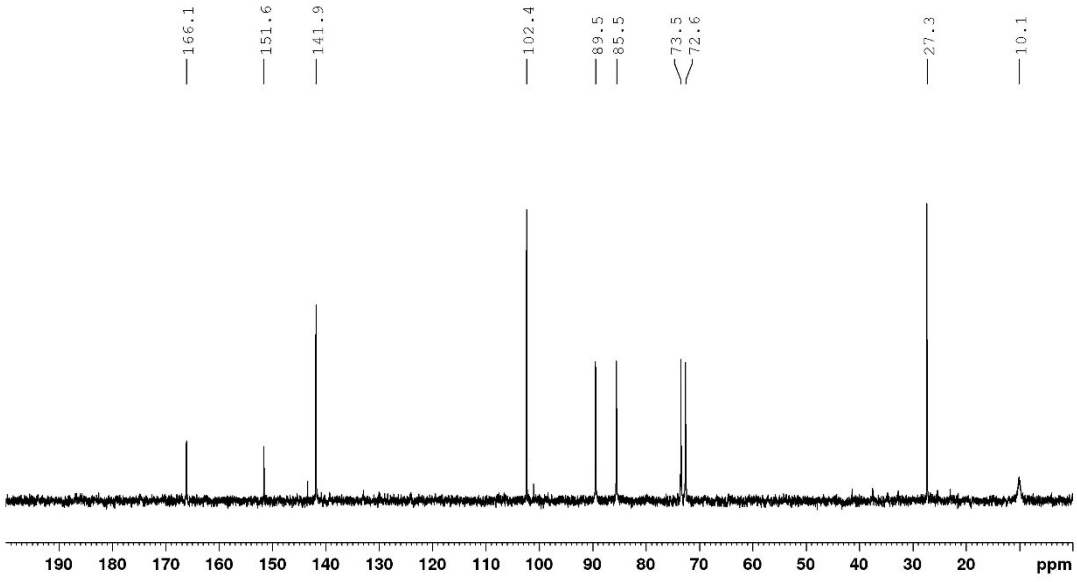


Figure S8 : ^{13}C NMR spectrum of **3** in D_2O (pD 6.5, 125 MHz, 37 mM)

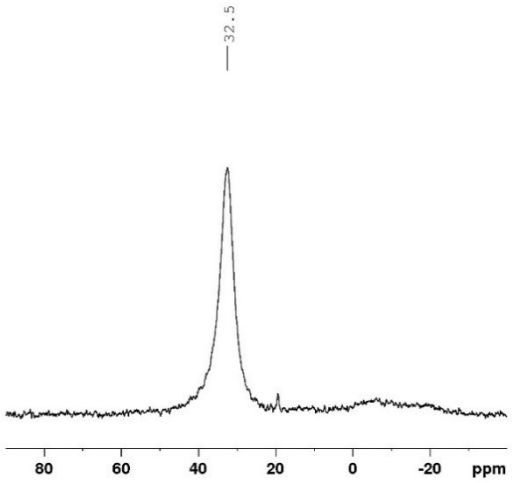


Figure S9 : ^{11}B NMR spectrum of **3** in D_2O (pD 6.5, 128 MHz, 37 mM)

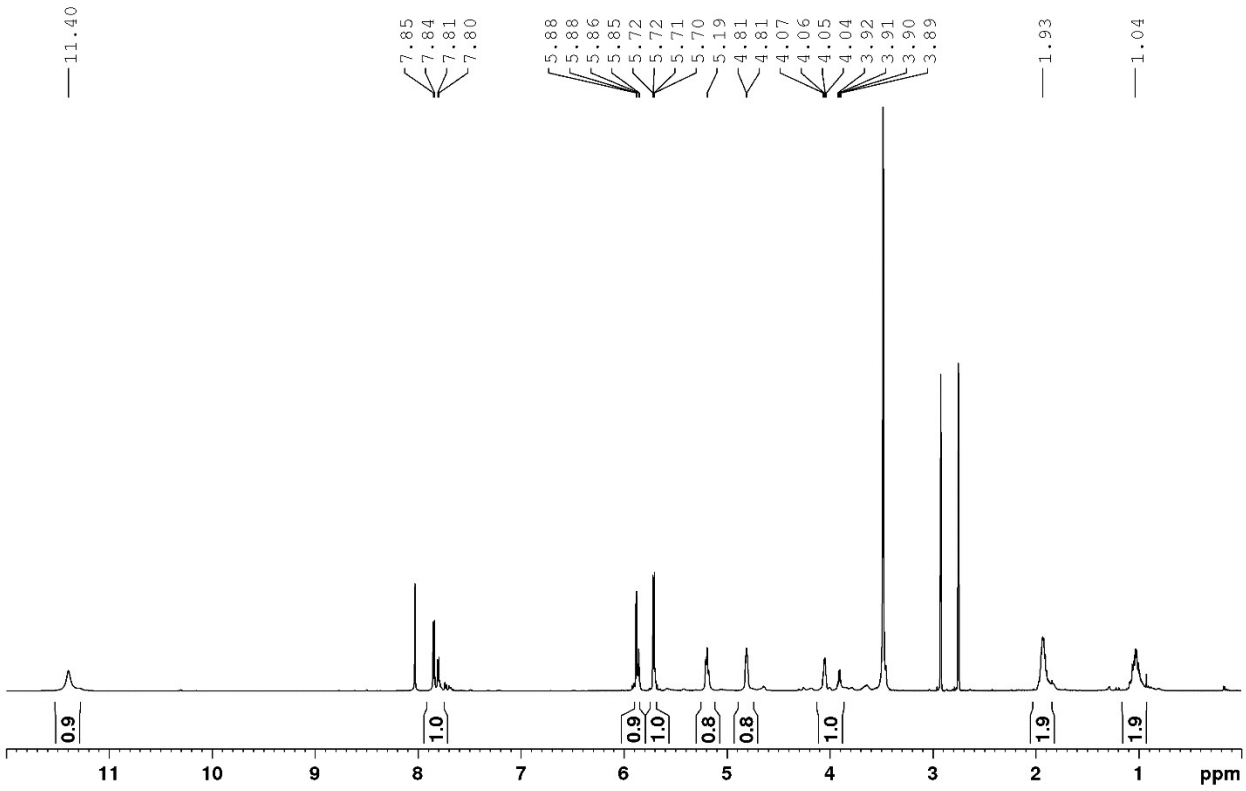


Figure S10 : ^{1}H NMR spectrum of the cyclic homopolymers of **3** (**htt_2** and **htt_3**) in anhydrous DMF-d_7 (600 MHz, 37 mM)

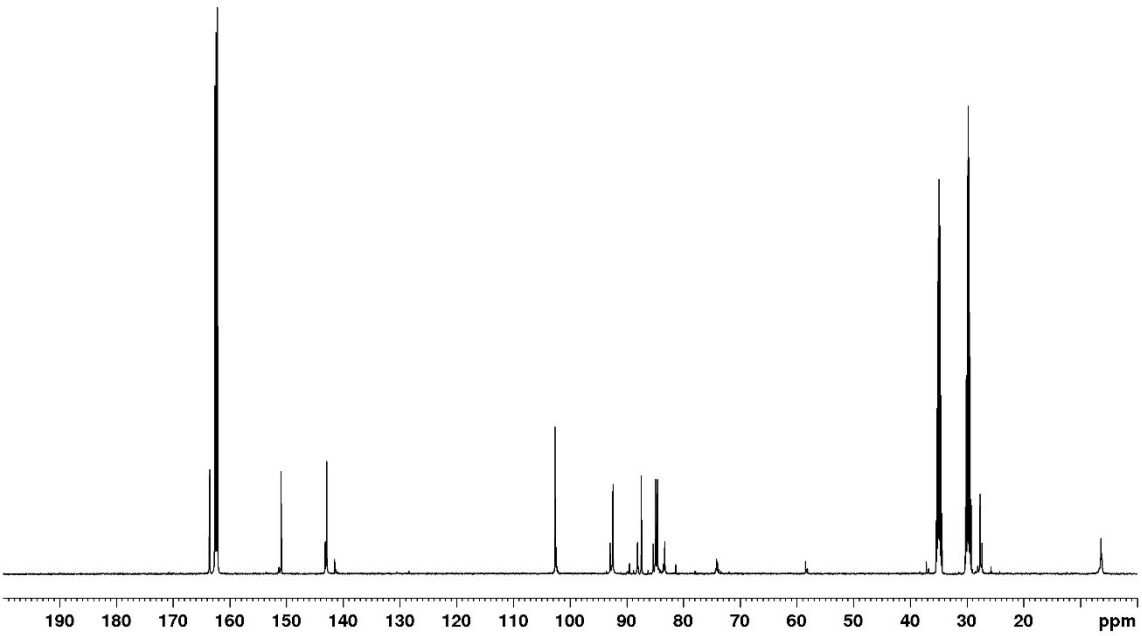


Figure S11 : ^{13}C NMR spectrum of the cyclic homopolymers of **3** (**htt_2** and **htt_3**) in anhydrous DMF-d_7 (125 MHz, 37 mM)

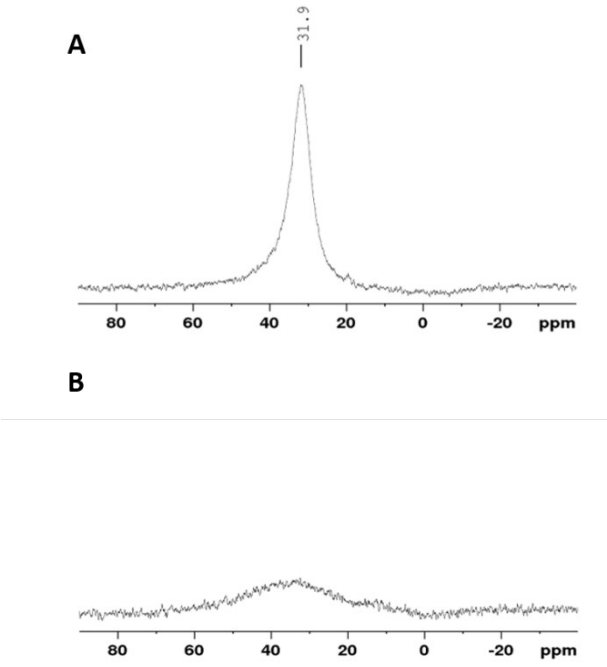


Figure S12 : ^{11}B NMR spectra of : (**A**) compound **3** in wet DMF-d_7 (4% H_2O, 128 MHz, 37 mM) and (**B**) the cyclic homopolymers of **3** (**htt_2** and **htt_3**) in anhydrous DMF-d_7 (128 MHz, 37 mM). Data collected on the same machine and with the same acquisition parameters.

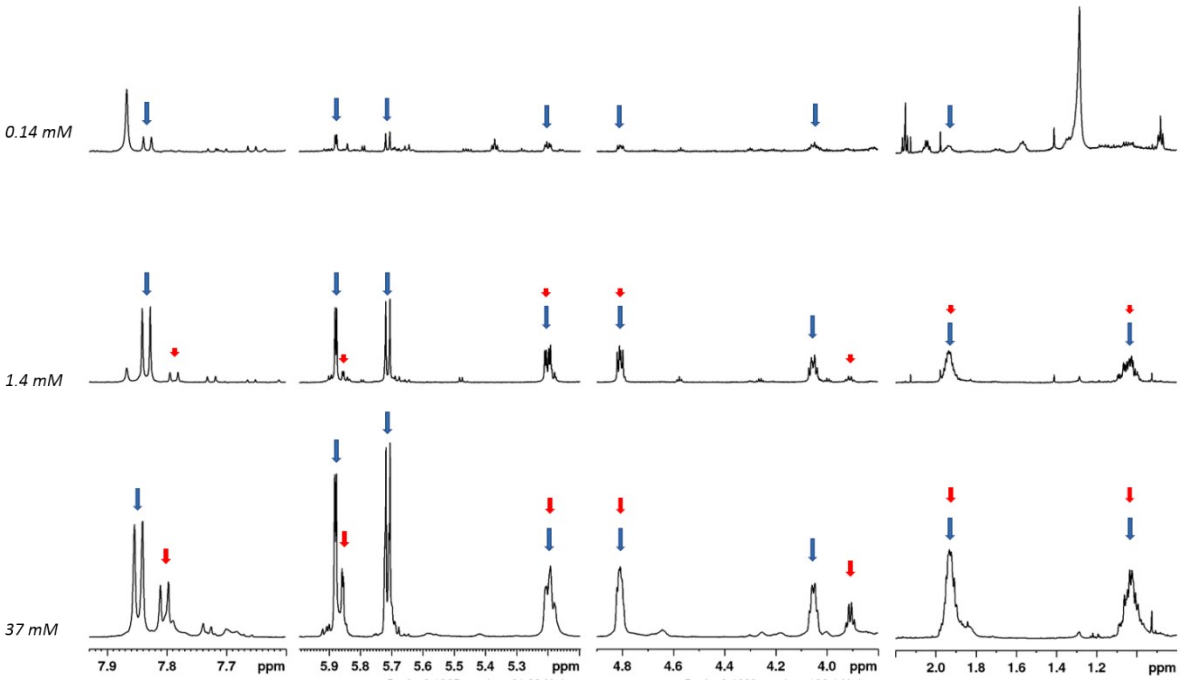


Figure S13 : ^{1}H NMR spectra of the cyclic homopolymers of **3** in anhydrous DMF-d_7 at different concentrations: 37 mM (bottom), 1.4 mM (middle) and 0.14 mM (top). Blue and red arrows identify the signals of the major **htt$_2$** dimer and minor **htt$_3$** trimer respectively.

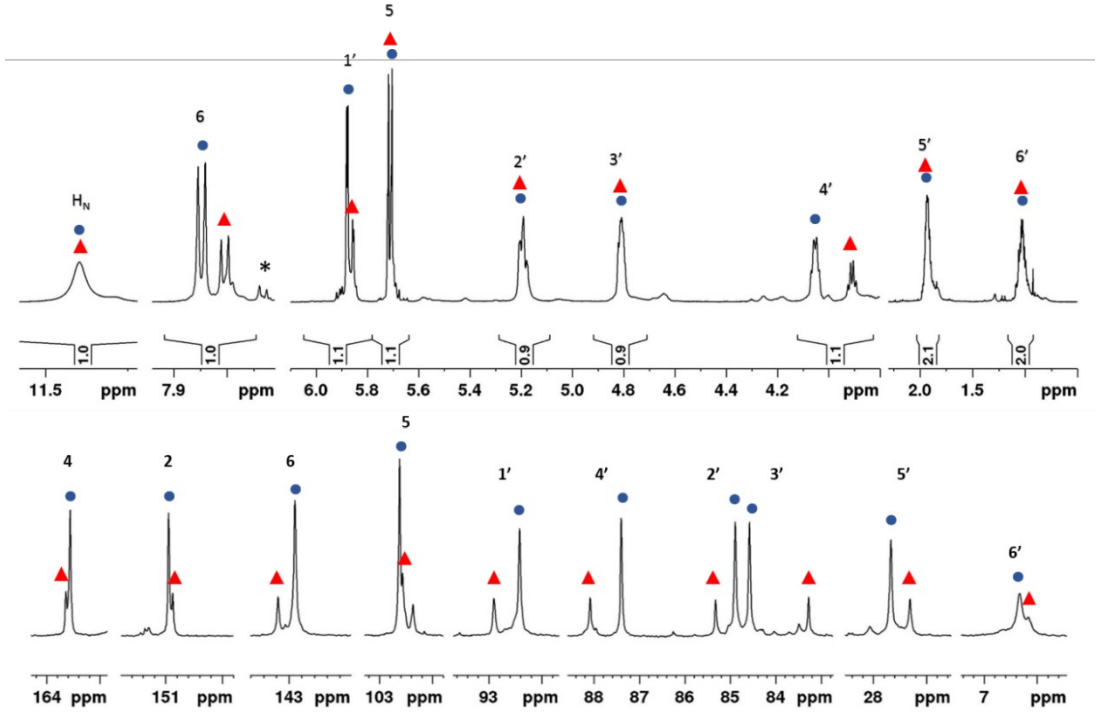


Figure S14 : Assignment of ^{1}H and ^{13}C resonances for the two cyclic homopolymers of **3** in anhydrous DMF-d_7 on the ^{1}H (upper part) and ^{13}C (lower part) NMR spectra . For clarity purpose, parts of the spectra free of signal were removed. Major **htt$_2$** (•) and minor **htt$_3$** (▲) compounds are labelled. * This doublet can reasonably be assigned to the H_6 proton either from the free form or from a small linear assembly.

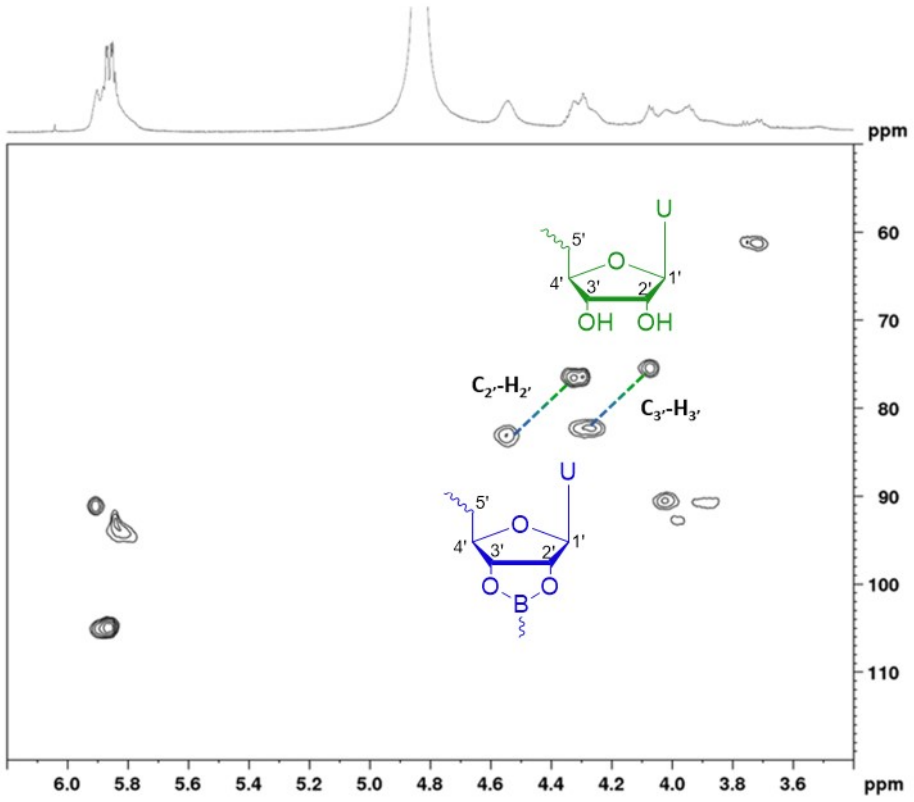


Figure S15 : Heteronuclear ^{1}H-^{13}C NMR spectrum of **3** and cyclic homopolymers of **3** in D_2O (37 mM, pD 9.3). Peaks corresponding to bound methynes at position 2' and 3' are labelled.

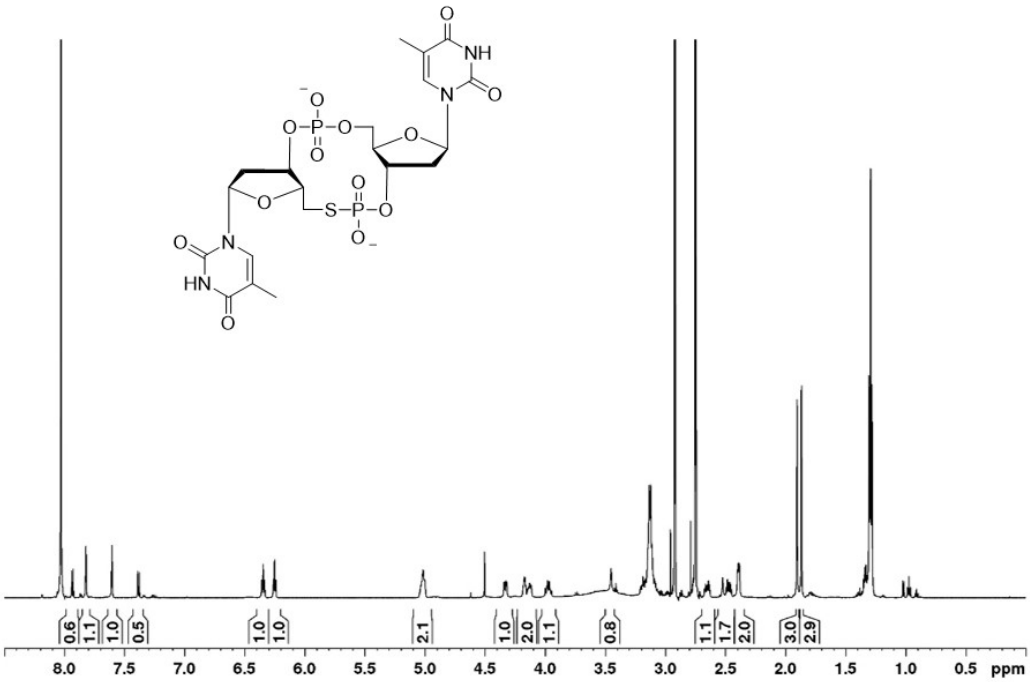


Figure S16 : ^{1}H NMR spectrum of **cT_2** in anhydrous DMF-d_7.

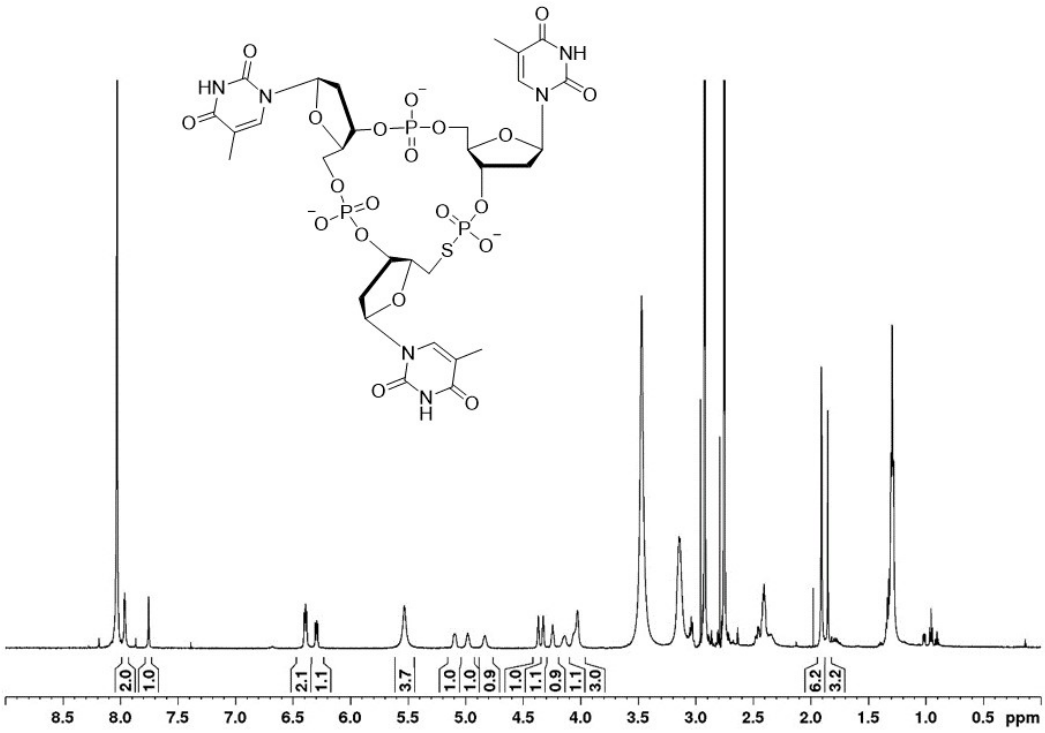


Figure S17 : ^{1}H NMR spectrum of **cT_3** in anhydrous DMF-d_7.

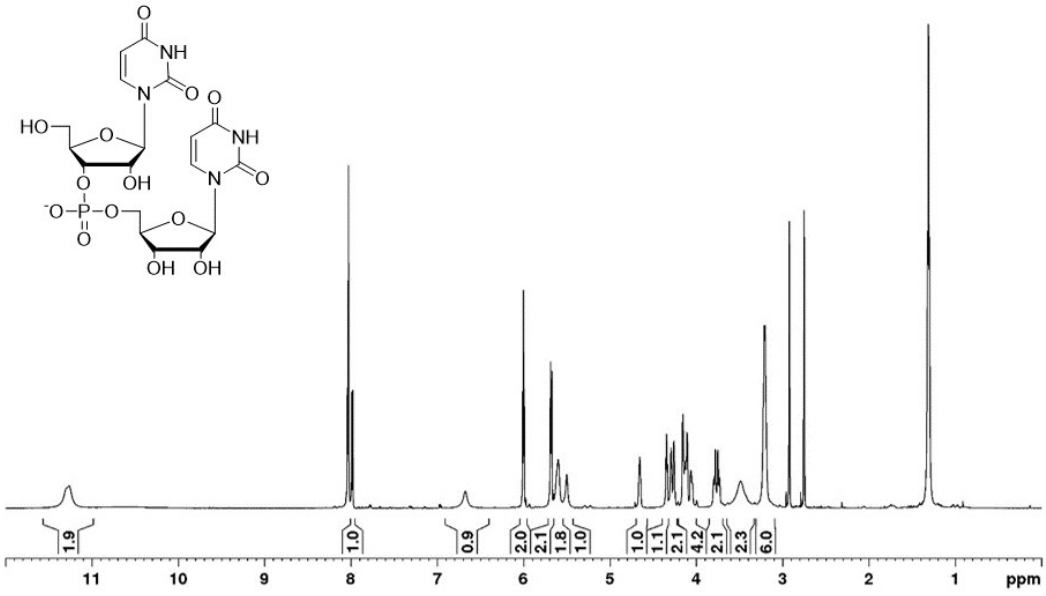


Figure S18 : ^{1}H NMR spectrum of **U_2** in anhydrous DMF-d_7.

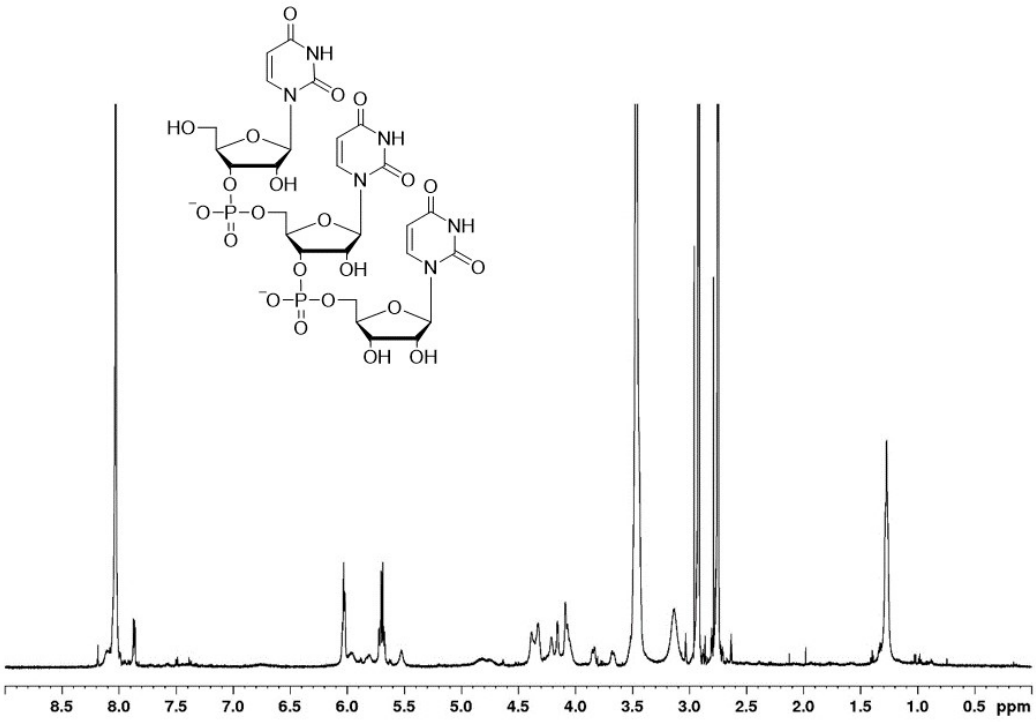


Figure S19 : ^{1}H NMR spectrum of **U_3** in anhydrous DMF-d_7.

5. Gaussian fits of the ^{1}H peaks intensity curve decay

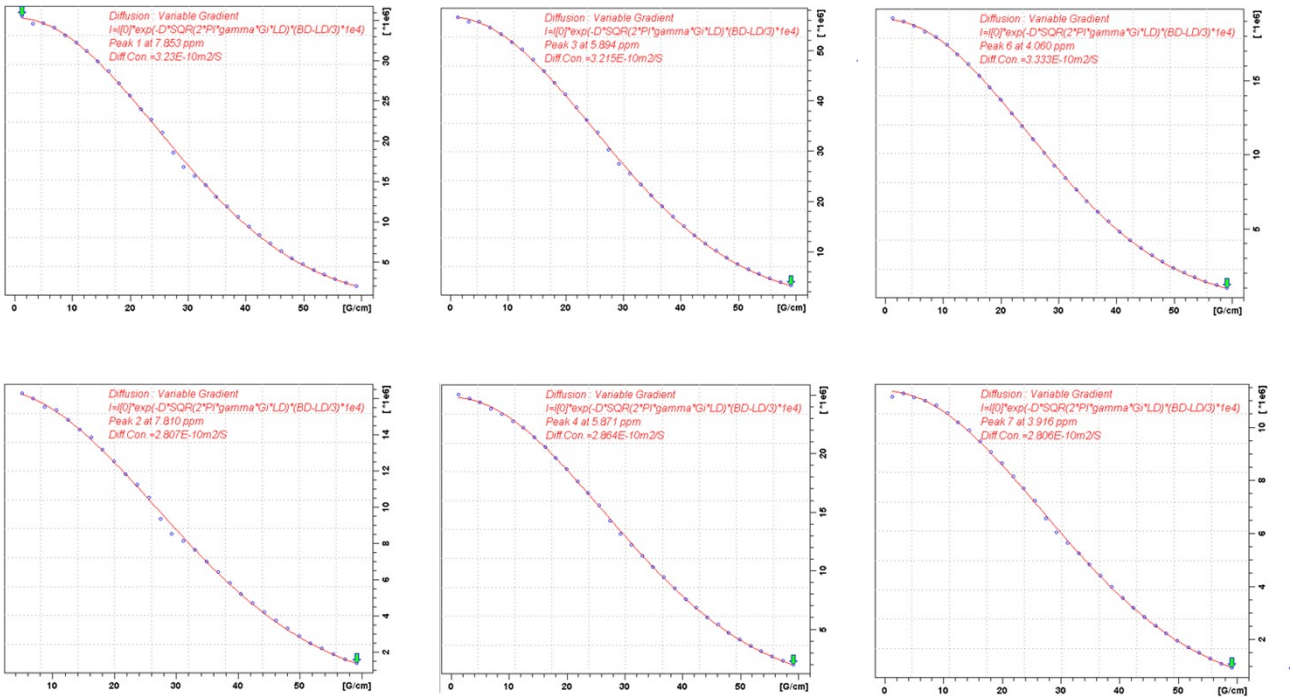


Figure S20 : Gaussian fits of protons H_6, $H_{1'}$ and $H_{4'}$ used to calculate the diffusion coefficients of **htt$_2$** (upper part) and **htt$_3$** (lower part) in anhydrous DMF-d_7 at 37 mM.

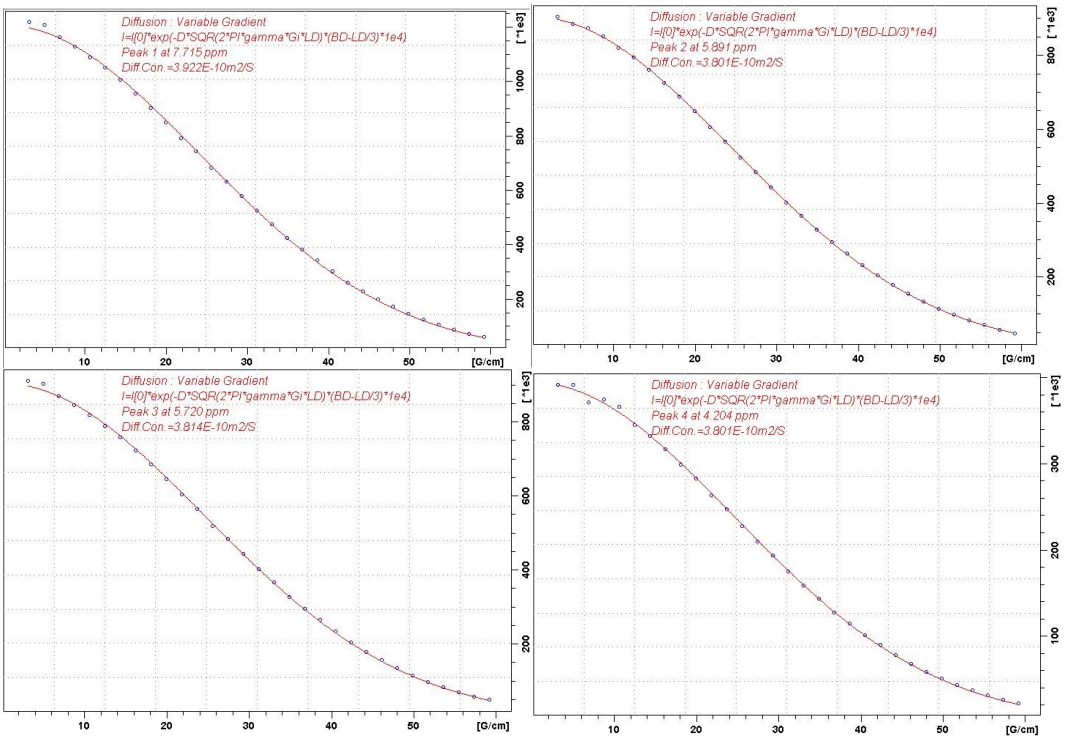


Figure S21 : Gaussian fits of protons H_6, $H_{1'}$, H_5 and $H_{2'}$ used to calculate the diffusion coefficients of **3** in DMF-d_7 with 4 % of water (%v/v) at 34.3 mM.

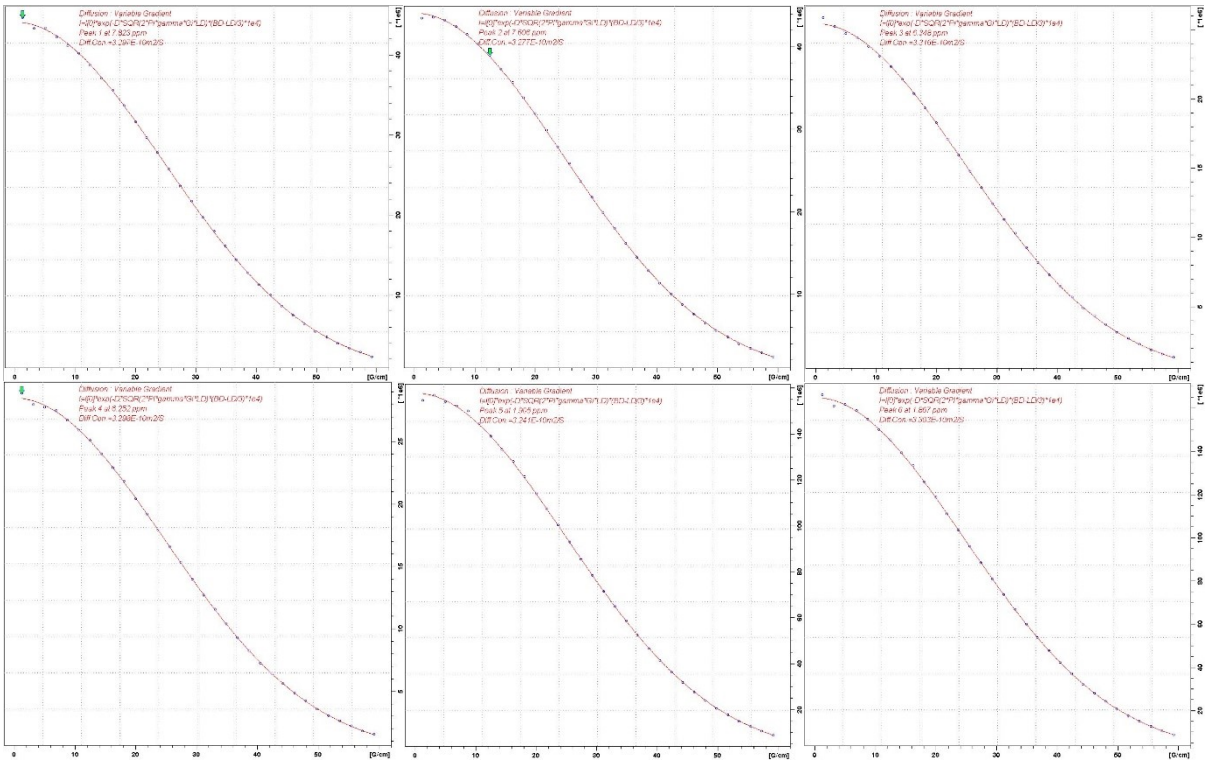


Figure S22 : Gaussian fits to calculate the diffusion coefficients of **c(T_2)** in anhydrous DMF-d_7.

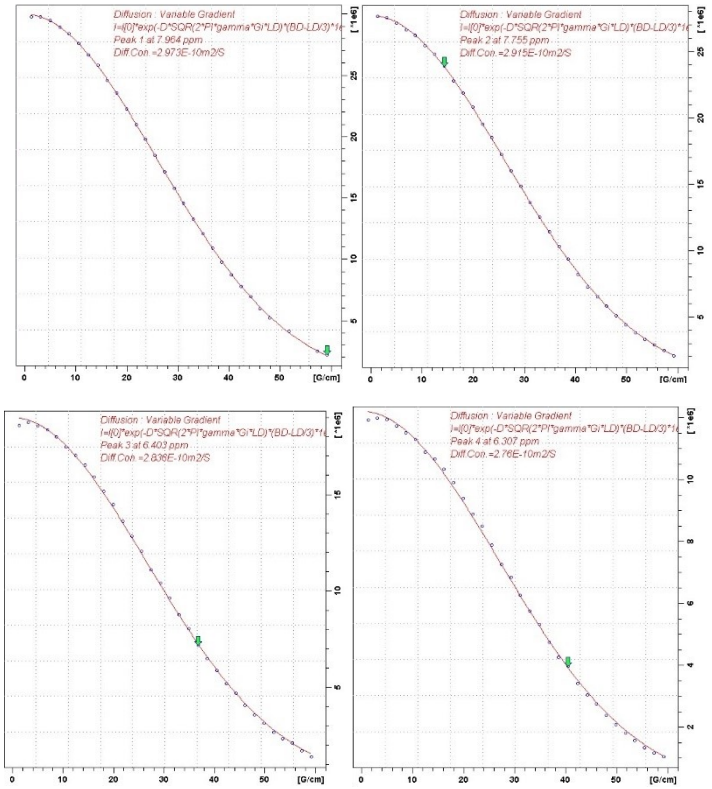


Figure S23 : Gaussian fits to calculate the diffusion coefficients of **c(T_3)** in anhydrous DMF-d_7.

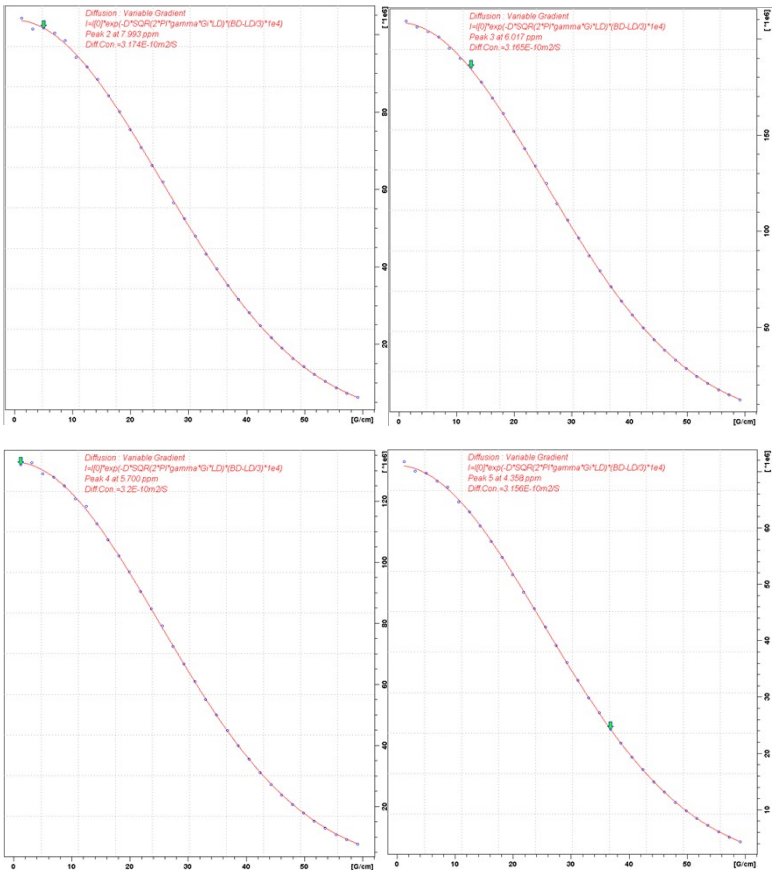


Figure S24 : Gaussian fits to calculate the diffusion coefficients of $\mathbf{U_2}$ in anhydrous DMF-d_7.

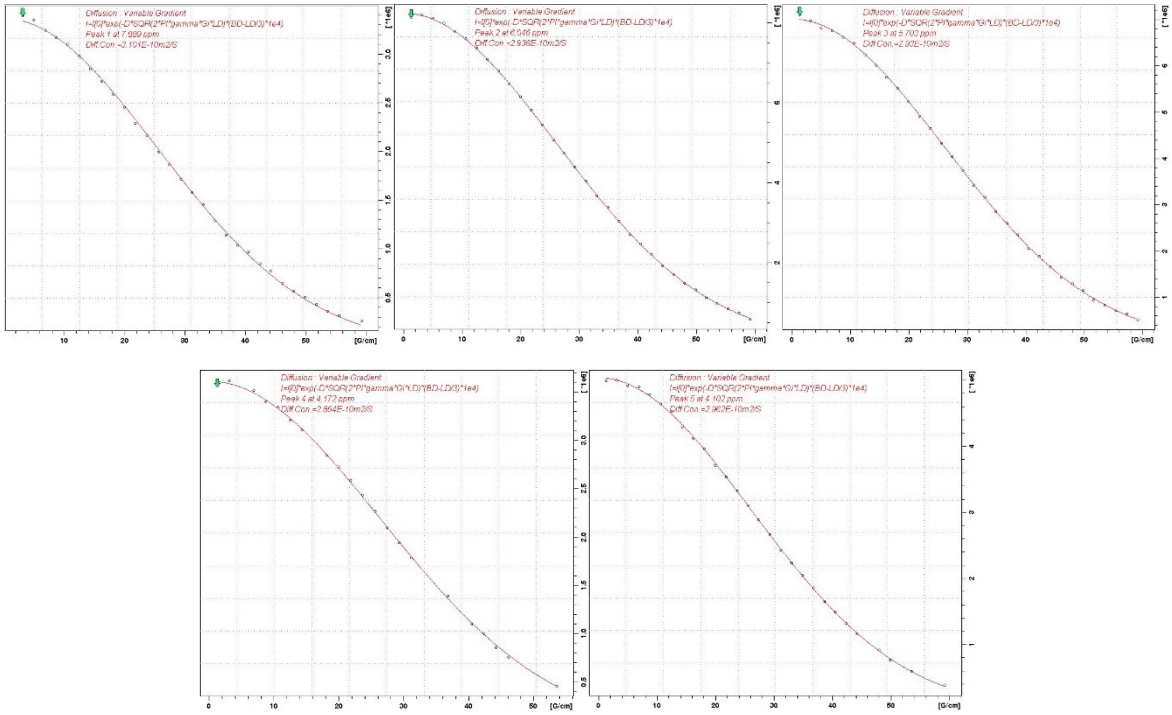


Figure S25 : Gaussian fits to calculate the diffusion coefficients of $\mathbf{U_3}$ in anhydrous DMF-d_7.

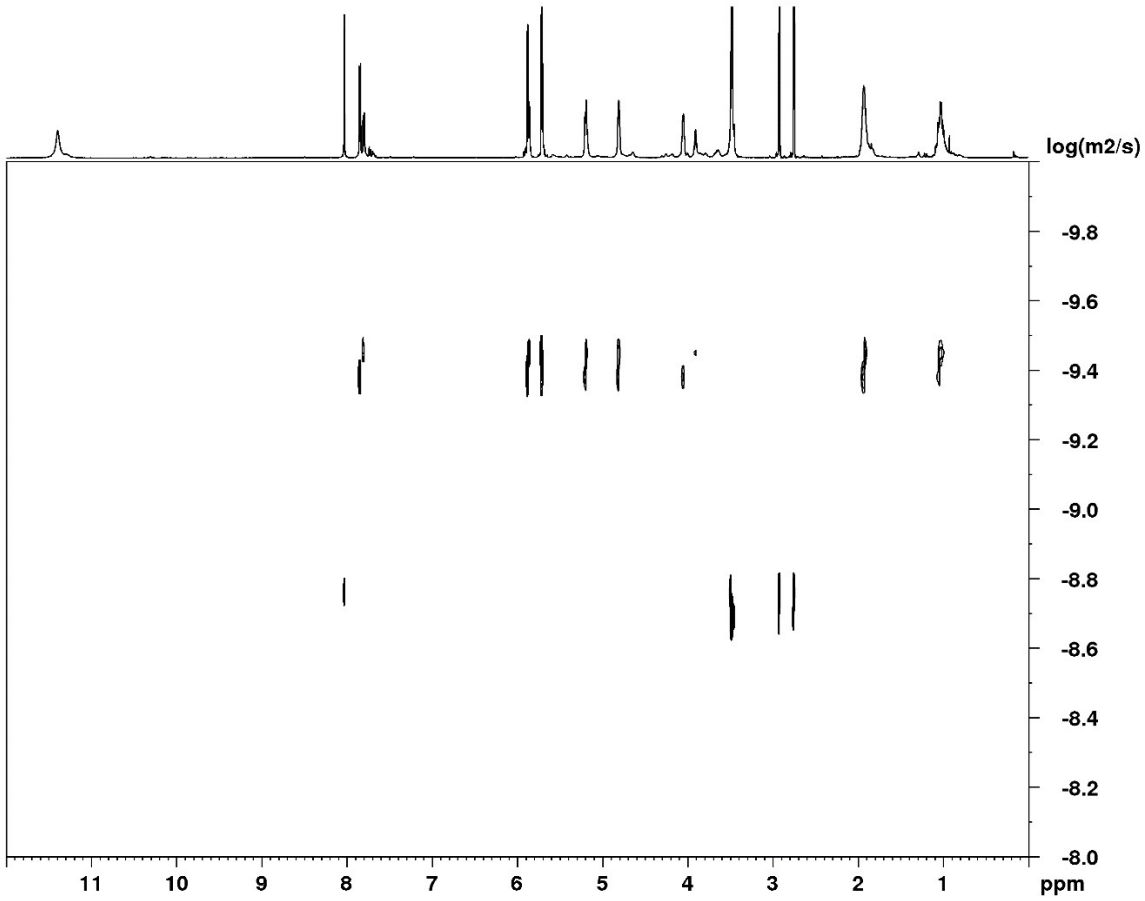


Figure S26 : DOSY NMR spectrum of **3** in anhydrous DMF-d_7.

6. Conformational analysis of the sugar ring from the coupling constants

Coupling constants were calculated from the generalized Karplus equation (1):[9]

$$^3J_{HH} = P_1\cos^2\phi_{HH} + P_2\cos\phi_{HH} + P_3 + \Sigma\Delta\chi_i\{P_4 + P_5\cos^2(\zeta_i\,\phi_{HH} + P_6|\Delta\chi_i|)\} \quad (1)$$

using the empirical parameters $P_1 = 13.24$; $P_2 = -0.91$; $P_3 = 0$; $P_4 = 0.53$; $P_5 = -2.41$; $P_6 = 15.5$; $P_7 = 0.19$ and taking into account the electronegativity of the substituents according to the following equation:

$$\Delta\chi_i = \Delta\chi_i^{\alpha\text{-susbt}} - P_7\,\Sigma\Delta\chi_i^{\beta\text{-susbt}} \quad (2)$$

Following expressions were used to correlate between the parameters governing the conformation of the ribose:

$$\phi_{1'2'} = 123.3° + 1.102\phi_m\cos(P - 144°) \quad (3)$$

$$\phi_{2'3'} = 0.2° + 1.090\phi_m\cos P \quad (4)$$

$$\phi_{3'4'} = -124.9° + 1.095\phi_m\cos(P + 144°) \quad (5)$$

7. Computational Method

As a starting point, structures of related c-di-GMP representing its "closed" and "open" conformers were extracted from the crystal structures 3I5A[10] and 3HV8.[11] The two guanines are on opposite faces/the same face in the open/closed conformer.[12] Here, the boronate linkages bound to the C2' and C3' sugar carbons allowed to construct just a single conformer of the **htt_2** dimer using the Molefacture module of the VMD software package.[13] Finally, the structure of **htt_2** was geometry optimized by means of the Gaussian 03 software package[14] at the MP2/6-31G(d) level of theory.[15]

8. Crystallization tests

Several methods were attempted to crystallize **htt_2** and **htt_3** :

- Evaporation : 10 mg of monomer **3** were dissolved in 1 mL of solvent (methanol, ethanol or water) in an opened test tube. The opened tube was let at room temperature or at 0°C for evaporation and crystallization.
- Diffusion : 10 mg of the monomer **3** were dissolved in 0.5 mL of methanol or DMF and 2 mL of another solvent (acetonitrile, diethyl ether, dichloromethane, dichloroethane, cyclohexane, pentane, toluene, dioxane or tetrahydrofuran) were slowly added. Test tubes were the placed at room temperature, at 0°C or at -20°C for several days.
- Vapour diffusion : 10 mg of the monomer **3** were dissolved in 0.5 mL of methanol or DMF. The test tube containing the mixture was placed into a closed falcon containing another solvent (acetonitrile, diethyl ether, dichloromethane, dichloroethane, cyclohexane or pentane). The falcon was placed at room temperature or 0°C for several days.
- Dean Stark : A suspension of the monomer **3** in a mixture of benzene/DMSO (9/1) or toluene/methanol (8/2) was heated under reflux with the use of a Dean Stark trap. The mixture was then slowly cooled to room temperature or 0°C. This method was also tested with 1 equivalent of triethylamine, 1,4-dipyridyl or 1,2-di(4-pirydil)ethylene.

Unfortunately, none of these methods allowed us to obtain any crystals.

9. Bibliography

[1] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, B. D. Sykes, *Journal of Biomolecular NMR* **1995**, *6*, 135–140.

[2] P. K. Glasoe, F. A. Long, *J. Phys. Chem.* **1960**, *64*, 188–190.

[3] M. Holz, H. Weingartner, *Journal of Magnetic Resonance (1969)* **1991**, *92*, 115–125.

[4] M. D. Matteucci, M. H. Caruthers, *Journal of the American Chemical Society* **1981**, *103*, 3185–3191.

[5] T. Lavergne, J.-R. Bertrand, J.-J. Vasseur, F. Debart, *Chemistry – A European Journal* **2008**, *14*, 9135–9138.

[6] V. A. Gierer, K. Wirtz, *Naturforschung A* **1953**, *8*, 532–538.

[7] A. R. Waldeck, P. W. Kuchel, A. J. Lennon, B. E. Chapman, *Progress in Nuclear Magnetic Resonance Spectroscopy* **1997**, *30*, 39–68.

[8] B. J. Blackburn, A. A. Grey, I. C. P. Smith, F. E. Hruska, *Canadian Journal of Chemistry* **1970**, *48*, 2866–2870.

[9] C. A. G. Haasnoot, F. A. A. M. de Leeuw, H. P. M. de Leeuw, C. Altona, *Organic Magnetic Resonance* **1981**, *15*, 43–52.

[10] N. De, M. V. A. S. Navarro, R. V. Raghavan, H. Sondermann, *Journal of Molecular Biology* **2009**, *393*, 619–633.

[11] M. V. A. S. Navarro, N. De, N. Bae, Q. Wang, H. Sondermann, *Structure* **2009**, *17*, 1104–1116.

[12] J. Wang, J. Zhou, G. P. Donaldson, S. Nakayama, L. Yan, Y. Lam, V. T. Lee, H. O. Sintim, *J. Am. Chem. Soc.* **2011**, *133*, 9320–9330.

[13] W. Humphrey, A. Dalke, K. Schulten, *Journal of Molecular Graphics* **1996**, *14*, 33–38.
[14] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Menucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratman, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Danenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford CT, **2004**.
[15] D. Cremer, *WIREs Comput Mol Sci* **2011**, *1*, 509–530.