Electronic Supplementary Material (ESI) for Polymer Chemistry.

Supporting information

"DNA-Teflon" sequence-controlled polymers

Donatien de Rochambeau, Maciej Barłóg, Thomas GW Edwardson, Johans J Fakhoury, Robin S Stein, Hassan S Bazzi, Hanadi F Sleiman *

Contents

SI-I General

All starting materials were obtained from commercial suppliers and used without further purification unless otherwise noted. Magnesium chloride, triethylamine, tris(hydroxymethyl)-aminomethane (Tris), urea, EDTA, glycerol, formamide, diethylpyrocarbonate (DEPC), triethylamine tetrahydrofluoride and solvents were used as purchased from Sigma-Aldrich. Acetic acid and boric acid were purchased from Fisher Scientific and used without further purification. GelRed™ nucleic acid stain was purchased from Biotium Inc. Concentrated ammonium hydroxide, acrylamide/Bis-acrylamide (40% 19:1 solution) and TEMED were obtained from Bioshop Canada Inc. and used as supplied. 1 µmol Universal 1000Å LCAA-CPG supports and standard reagents used for automated DNA and RNA synthesis were purchased through Bioautomation. DMT-hexaethyloxy glycol (cat.# CLP-9765) phosphoramidites were purchased from Chemgenes. Sephadex G-25 (super fine, DNA grade) and sulfurizing reagent II were purchased from Glen Research. AFM cantilevers (model SCANASYST-AIR) were purchased from Bruker and RubyRed mica from Electron Microscopy Sciences. TAMg buffer is composed of 40 mM Tris and 7.6 mM $MgCl_2 \cdot 6H_2O$ with pH adjusted to 8.0 using glacial acetic acid. TBE buffer is 90mM Tris, 90mM boric acid and 1.1mM EDTA with a pH of 8.0. HepG2 cells (Human hepatocellular carcinoma) were purchased from ATCC.

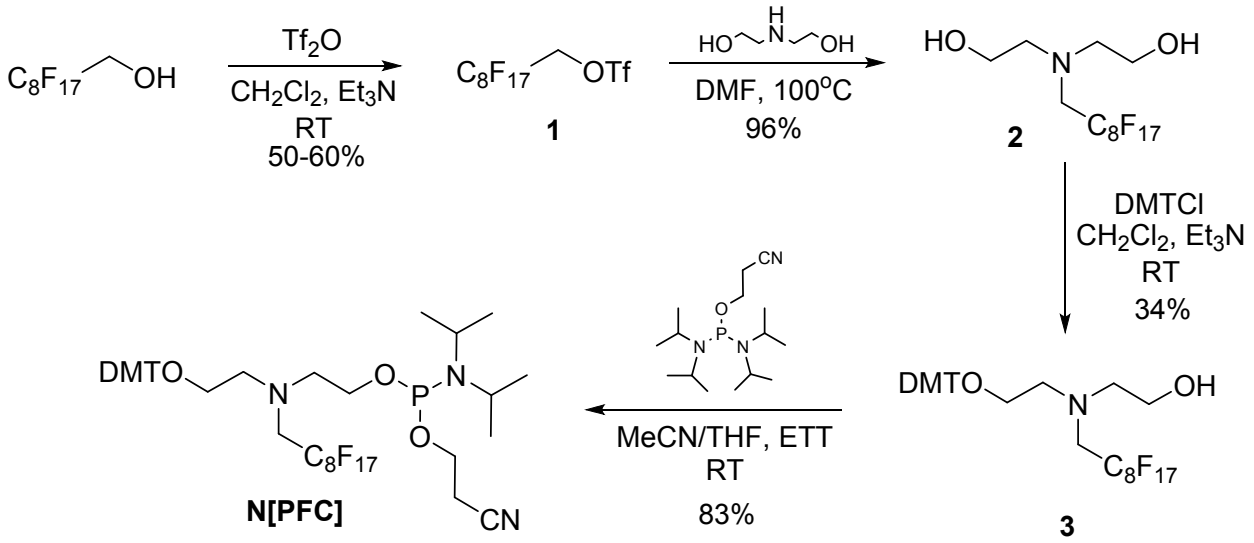
SI-II Instrumentation

Standard automated solid-phase synthesis was performed on a Mermade MM6 synthesizer from Bioautomation. HPLC purification was carried out on an Agilent Infinity 1260. DNA quantification measurements were performed by UV absorbance with a NanoDrop Lite spectrophotometer from Thermo Scientific. A Varian Cary 300 Bio spectrophotometer was used for melting temperature studies. Polyacrylamide Gel Electrophoresis (PAGE) experiments were carried out on a 20 X 20 cm vertical Hoefer 600 electrophoresis unit while agarose gel electrophoresis (AGE) were performed with an Owl Mini gel electrophoresis unit. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Mass determination of the phosphoramidite was carried out using Electron-Spray Ionization – Ion Trap - Mass Spectrometry (MS) on a Finnigan LCQ Duo device. Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC-ESI-MS) was carried out using Dionex Ultimate 3000 coupled to a Bruker MaXis Impact™ QTOF. Dynamic Light Scattering (DLS) experiments were carried out using a DynaPro™ Instrument from Wyatt Technology. AFM was performed with a MultiModeTM MM8 SPM connected to a NanoscopeTM controller, from the Digital Instruments Veeco Metrology Group. Oxygen and Moisture sensitive experiments were carried out in a Vacuum Atmospheres Co. glove

box. The NMR spectra were recorded at 400 MHz for ^{1}H and ^{13}C at 100.6 MHz, with chloroform-d_1 (δ 7.26, ^{1}H; δ 77.0, ^{13}C), acetone–d_6 d_1 (δ 2.04, ^{1}H; δ 29.8, ^{13}C) as internal lock solvent and chemical shift standard unless otherwise indicated. ^{19}F spectra were acquired using a Bruker AVIIIHD spectrometer equipped with a BBFO+ Smartprobe operating at 470.7 MHz. Chemical shift referencing used the lock solvent.

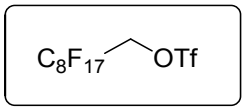
SI-III Synthesis, Purification and Characterization of the different oligomers.

IIIa. N[PFC] synthesis



Supporting Scheme SS1: Synthetic pathway for the preparation of N[PFC] phosphoramidite.

1. 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-Heptadecafluoro-1-nonanol trifluoromethanesulfonate



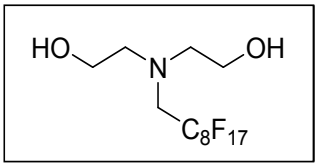
The fluorinated alcohol (3g, 6.67 mmol) was dissolved in dry dichloromethane (25 mL) in a round bottom flask (RBF), followed by dry triethylamine (3 mL) and cooled down to 0°C. Trifluoromethanesulphonate anhydride (2.82g, 10 mmol) was added dropwise over 15 minutes turning the reaction mixture dark. The

reaction mixture was stirred at 0°C for 1 hour, than allowed to warm up to RT and stirred for 1 more hour. The reaction was quenched with sat. $NaHCO_3$ (50 mL) and product extracted with dichloromethane (3 x 50 mL). Organic fractions were combined, dried with $MgSO_4$ and solvent removed under vacuum to produce dense black oil. Crude material was purified by flash chromatography with 0-5% EtOAc/hexanes mixture to produce dense colourless/yellowish oil solidifying on standing. Typical yield 50-60%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.82 (t, *J* = 12.3 Hz, 2H).

Note: Synthesis of the mesyl and tosyl equivalents of molecule **1** led to lower yields.

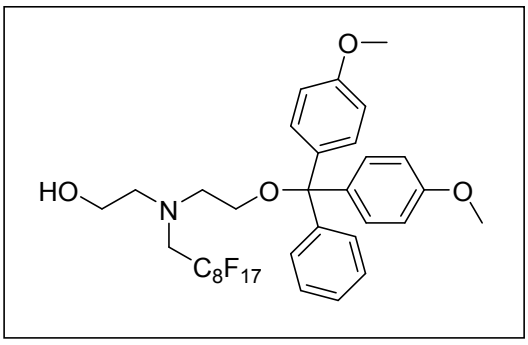
2. 2,2'-(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-Heptadecafluoro-1-nonylazanediyl)diethanol.



1 (2.7 g, 4.64 mmol) was added to a solution of diethanolamine (1.53 g, 9.28 mmol) in dry DMF (6 mL). The reaction mixture was placed in 100°C oil bath and stirred for 2 hours. It was cooled down, dissolved in water (50 ml) and the product extracted with dichloromethane (3 x 50 mL). The organic fractions were combined, dried with $MgSO_4$ and the solvent was removed under vacuum to produce a dense yellowish oil solidifying on standing (2.42 g, 96%). The crude product was analysed with ^{1}H NMR, determined as analytically pure and used in the next step without any further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 2.68 (br. s, 2H) 2.92 (t, *J* = 5.1 Hz, 4H), 3.34 (t, *J* = 16.9 Hz, 2H), 3.65 (t, *J* = 5.1 Hz, 4H);

3. 2-((2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononyl)amino)ethan-1-ol

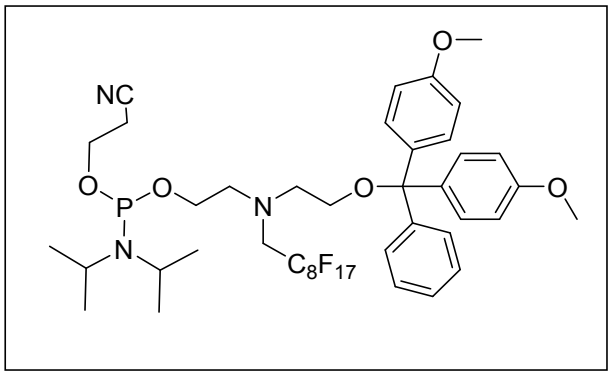


2 was dissolved in dry dichloromethane (15 mL) and triethylamine (2 mL). DMTCl (1.51g, 4.45 mmol) was added portionwise and the reaction mixture was stirred at RT for 2 hours. The solvents were evaporated under vacuum resulting a yellow oil purified by column chromatography on triethylamine pre-treated silica with slow gradient of EtOAc/hexanes (0-15%) mixture to produce a dense yellow oil 1.31g (34%). Unreacted **2** can be recovered from the column by washing with pure EtOAc.

^{1}H NMR (400 MHz, acetone-d_6) δ 2.87 (t, J = 5.9 Hz, 2H), 3.02 (t, J = 5.7 Hz, 2H), 3.26 (t, J = 5.7 Hz, 2H), 3.46 (br. s, 1H), 3.57 (t, J = 17.1 Hz, 2H), 3.62 (t, J = 6.0 Hz, 2H), 3.77 (s, 6H), 6.88 (d, J = 8.8 Hz, 4H), 7.22 (d, J = 7.4 Hz, 1H), 7.30 (t, J = 7.6 Hz, 2H), 7.36 (d, J = 8.8 Hz, 4H), 7.49 (d, J = 7.7 Hz, 2H).

^{13}C NMR (100 MHz, acetone-d_6) δ 48.1 (CH_2), 56.4 (2 x CH_3), 57.2 (CH_2), 59.4 (CH_2) 61.9 (CH_2), 63.9 (CH_2), 88.2 (C), 114.8 (4 x CH), 128.5 (CH), 129.5 (2 x CH), 130.0 (2 x CH), 131.9 (4 x CH), 138.2 (2 x C), 147.4 (C), 160.6 (2 x C).

4 (N[PFC] phosphoramidite). 2-((2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononyl)amino)ethyl (2-cyanoethyl) diisopropylphosphoramidite



An oven-dried round bottom flask was charged with **3** (860.10 mg, 1.02 mmol, 1eq.) dissolved in dry THF (3.3 mL). 5-(ethylthiotetrazole) (ETT) (6.15 mL, 0.25M in acetonitrile, 1.54 mmol, 1.5 eq.) was added under argon and stirred vigorously followed by addition of 3-((bis(diisopropylamino)phosphanyl)oxy)propanenitrile. The reaction mixture was left stirring for 6 hours at room temperature under argon. The solvent was evaporated under reduced pressure. The crude product was purified by column chromatography under a positive pressure of argon with mobile phase: degassed

Hexanes/ethyl acetate/triethylamine TEA (90:10:2). **4** was isolated as a colourless oil: 880 mg. Yield: 83%. Purity (calculated through ^{31}P NMR and ^{1}H NMR): 99%.

MS: Calc.exact mass: 1039.30 g/mol. Measured (positive mode): 1062.18 (M+23), 1078.16 (M+39) g/mol.

TLC (Hexanes/ethyl acetate/TEA: 90:10:2): Rf=0.16.

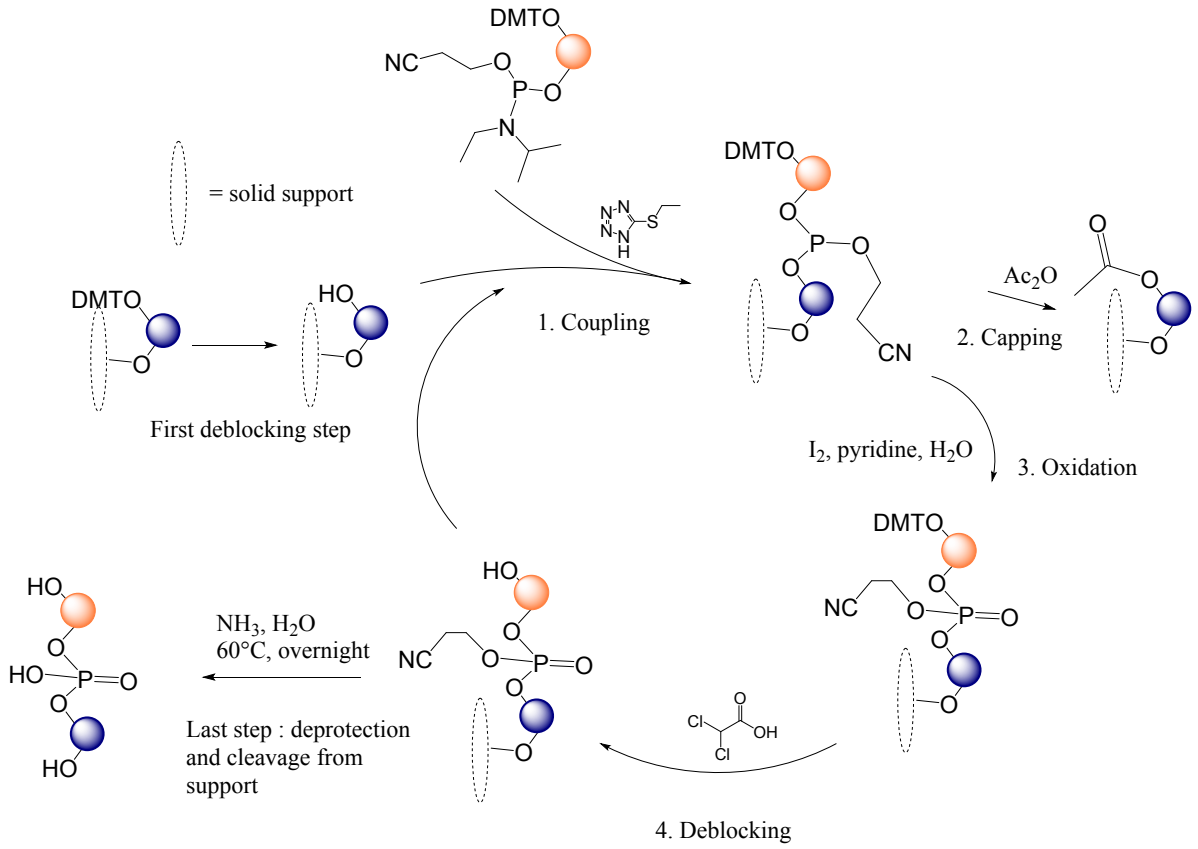
^{31}P NMR (400 MHz, acetone-d_6) δ (ppm) = 147.4 (s).

^{1}H NMR (400 MHz, acetone-d_6): δ (ppm) = 7.46 (d, *J*=8Hz, 2H), 7.29-7.35 (m, 6H), 7.23-7.25 (m, 1H), 6.87 (d, *J*=12Hz, 4H), 3.78 (s, 6H), 3.57-3.79 (m, 6H), 3.48 (t, *J*=16Hz, 2H), 3.17 (t, *J*=8Hz, 2H), 2.95 (m , 4H), 2.61 (t, *J*=4Hz, 2H)m, 1.16 (dd, *J*=16, 4Hz, 12H).

^{13}C NMR (125 MHz, acetone-d_6) δ 19.8(CH_2), 24.0(4 x CH_3), 42.8 (2 x CH), 54.6 (2 x CH_3), 55.2 (CH_2), 55.9 (CH_2), 58.5 (CH_2), 61.9 (CH), 62.0 (CH), 62.2 (CH_2), 86.2 (C), 112.9 (4 x CH), 118.2 (C), 126. (CH), 127.7 (2 x CH), 128.0 (2 x CH), 130.0 (4 x CH), 136.3 (2 x C), 145.5 (C), 158.7 (2 x C).

Note: automated column chromatography and solid loading for phosphoramidite purification led to lower yields.

III.b. Solid phase synthesis



Supporting Figure SF1: Synthetic cycle performed by the automated synthesizer. Orange and blue spheres represent the monomers used.

DNA synthesis was performed on a 1 µmole scale, starting from a universal 1000 Å LCAA-CPG solid support. Coupling efficiency was monitored after removal of the dimethoxytrityl (DMT) 5'-OH protecting groups. DMT-hexaethyloxy glycol amidite and N[PFC] were respectively dissolved in acetonitrile and acetonitrile/THF (90:10) under a nitrogen atmosphere in a glove box (<0.04 ppm oxygen and <0.5 ppm trace moisture). For DMT-hexaethyloxy glycol (0.1M) and N[PFC] (0.08M) amidites, extended coupling times of 10 minutes were used. For the addition of each RNA nucleoside phosphoramidite extended coupling time of 6 minutes was used. For 2' OMe modified RNA phosphoramidites, under a nitrogen atmosphere, coupling was done using the 'syringe' technique: the amidite solution (200 µl, 0.1 M) is mixed with the usual activator solution (200µl, 0.25 M) in presence of the CPG using syringes. After twenty minutes, the solution was removed from the columns and the strands underwent capping, oxidation and deblocking steps in the synthesizer. Removal of the DMT

protecting group was carried out using 3% dichloroacetic acid in dichloromethane on the DNA synthesizer. Sulfurizing reagent was used according to standard procedures.

Deprotection procedure (except RNA): Completed syntheses were cleaved from the solid support and deprotected in 28% aqueous ammonium hydroxide solution for 16-18 hours at 60°C. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60°C. This crude solid was re-suspended in 1mL Millipore water. Filtration with 0.22µm centrifugal filter was then performed prior to HPLC purification. The resulting solution was quantified by absorbance at 260nm.

RNA deprotection procedure: Completed 1 µmol syntheses were deprotected in 1 mL of a 1:1 v/v mixture of 40% aqueous Methylamine and 28% aqueous ammonium hydroxide solution for 30 minutes at room temperature, followed by 3 hours at 65° C. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60°C. This crude solid was re-suspended in 150 µL of a desilylation solution containing triethylamine, N-methylpyrrolidone, and triethylaminetrihydrofloride (3:2:1.5) and heated to 65˚C for 2 hours, to remove the 2'-OH *tert*-butyldimethylsilyl protecting groups. This desilylation step was then quenched by the addition of 100 µL of 3 M sodium acetate (pH 5.5) and vortexed. RNA precipitation was induced by addition of 1 mL of cold butanol and left for 30 minutes at -20˚C. A pellet appeared through centrifugation (20 minutes, 12000 x g, 4˚C) and the supernatant was removed. The pellet was washed a second time with 500 µL butanol, and dried under reduced pressure at 60 °C. This crude product was resuspended in DEPC-treated sterile water and quantified by absorbance at 260nm.

HPLC purification: Solvents (0.22µm filtered): 50mM triethylammonium acetate (TEAA) buffer (pH 8.0) and HPLC grade acetonitrile. Elution gradient: 3-95% acetonitrile over 40 minutes at 60°C. Column: Hamilton PRP-C18 5µm 100 Å 2.1 x 150mm. For each analytical separation approximately 0.5 OD_{260} of crude DNA was injected as a 20-50µL solution in Millipore water. Detection was carried out using a diode-array detector, monitoring absorbance at 260nm.

Supporting Table ST1. Yields and ESI-MS characterization of different amphiphilic polymers $\mathbf{A_{n,m}}$.

Number of HEG (n) and N[PFC] units (m)	HPLC yields[a] [%]	Calculated exact mass [(M-2)/2] [g/mol]	Found exact mass [g/mol]
n=6, m=0	81	2305.83[b]	2305.82[b]
n=6, m=2	78	1751.43	1751.44
n=6, m=3	75	2050.94	2050.95
n=8, m=0	83	1496.54	1496.55
n=8, m=2	73	2095.56	2095.57
n=8, m=4	68	2694.57	2694.91

[a] Calculated through the integration of the peak associated to the expected product (260 nm detection). [b] (M-1) is reported

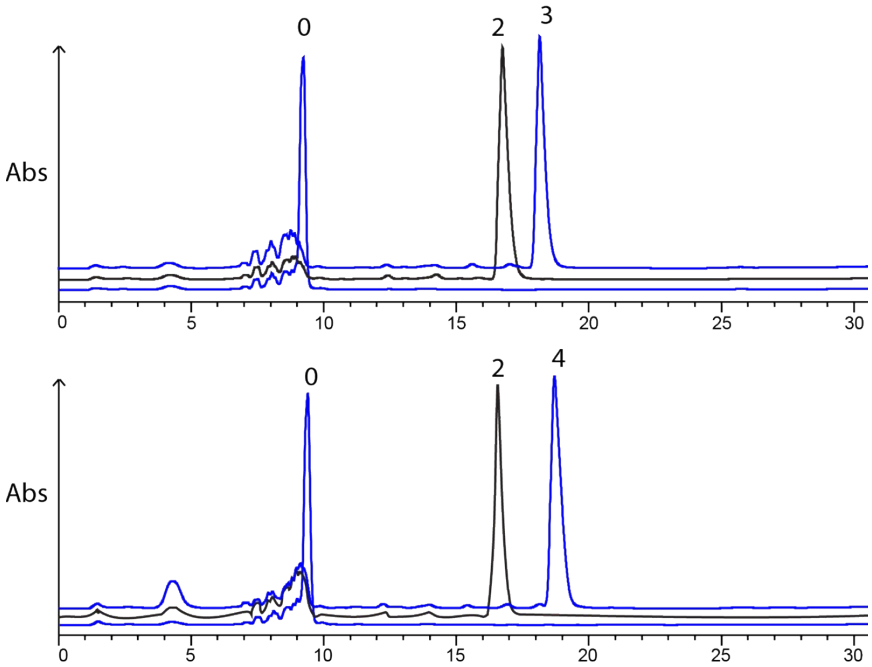


Figure SF2. Reverse-phase HPLC traces (UV detection, 260 nm) from crude mixtures of $\mathbf{A_{n,m}}$ polymers. Top: n =6, bottom: n=8. Numbers on the peaks are the number (m) of N[PFC]. Byproducts are in low quantities. They are almost exclusively related to HEG coupling as the byproducts peaks are visible before the polymer retention time for m=0.

Gel electrophoresis purification: In the case of DNA/RNA internal or 3' end modification, purification was carried out through gel electrophoresis instead of RP-HPLC. In that case, crude products were

purified on 19% polyacrylamide gels, supplemented with 8M urea (loading up to 20 OD260 of crude DNA per gel, 500 V field applied). Electrophoresis was run at lower voltage for the first 30 minutes. Following electrophoresis, the gel was wrapped in plastic and visualized by UV shadowing over a fluorescent TLC plate. The full-length product was quickly excised, then crushed and incubated in ~10 mL of autoclaved water (treated with DEPC in the case of RNA purification) at 55°C overnight. The supernatant was then concentrated to 1.0 mL, and desalted using size exclusion chromatography (Sephadex G-25). Sephadex was treated with DEPC in the case of RNA purification. Strands were then quantified (OD260) and converted to micromolar concentrations using the extinction coefficients obtained on IDT technology website (http://www.idtdna.com/calc/analyzer).

Table ST2: DNA, RNA and conjugates. Lower case letters indicate RNA, uppercase letters indicate DNA. CpAT is complementary to the AT sequence. * indicates a phosphorothioate linkage. Letters in bold indicate 2' OMe modifications

Molecule	**Sequence (5'-xx-3')**
AT	TTTTTCAGTTGACCATATA
cpAT	TATATGGTCAACTGAAAAA
AT-(N[PFC])$_n$ or B$_n$	(N[PFC])$_n$TTTTTCAGTTGACCATATA
(N[PFC]$_n$)-cpAT	TATATGGTCAACTGAAAAA(N[PFC])$_n$
AT-(N[PFC]$_n$)p10	TTTTTCAGT(N[PFC])$_n$TGACCATATA
cpAT-(N[PFC]$_n$)p11	TATATGGTCA(N[PFC])$_n$ACTGAAAAA
AT-N[PFC]p7p14	TTTTTC(N[PFC])AGTTGAC(N[PFC])CATATA
cpAT-N[PFC]p7p14	TATATG(N[PFC])GTCAACT(N[PFC])GAAAAA
ApoB-anti	auugguauucagugugauga**c*a***c
ApoB-sense	gucaucacacugaauaccaa*u
ApoB-sense-N[PFC]	gucaucacacugaauaccaa*u*N[PFC]

Internal N[PFC] modification has also been carried out successfully on these sequences (data not shown):

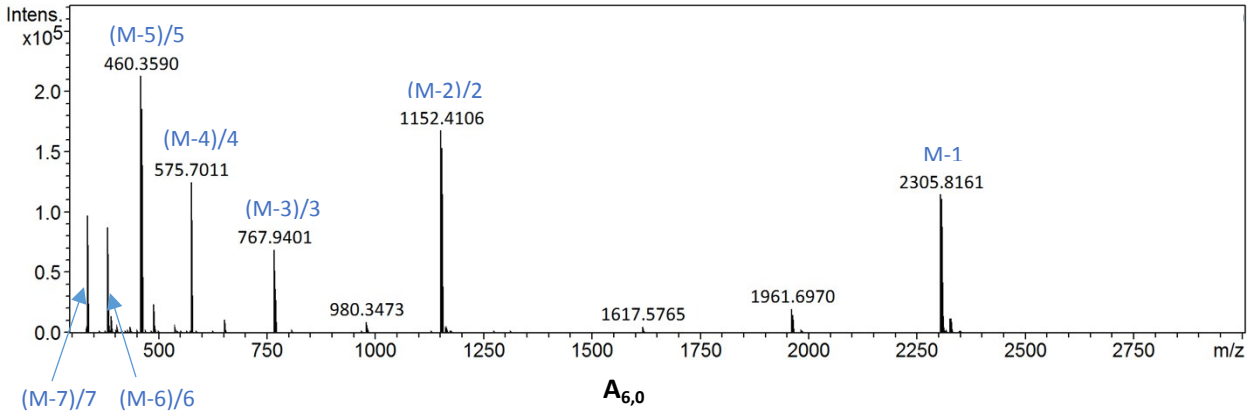
5' – CGTGTGCCTCACCGACCAATGC - 3'

5' – GCCAGCGTAGTGGATGTCTGCG – 3'

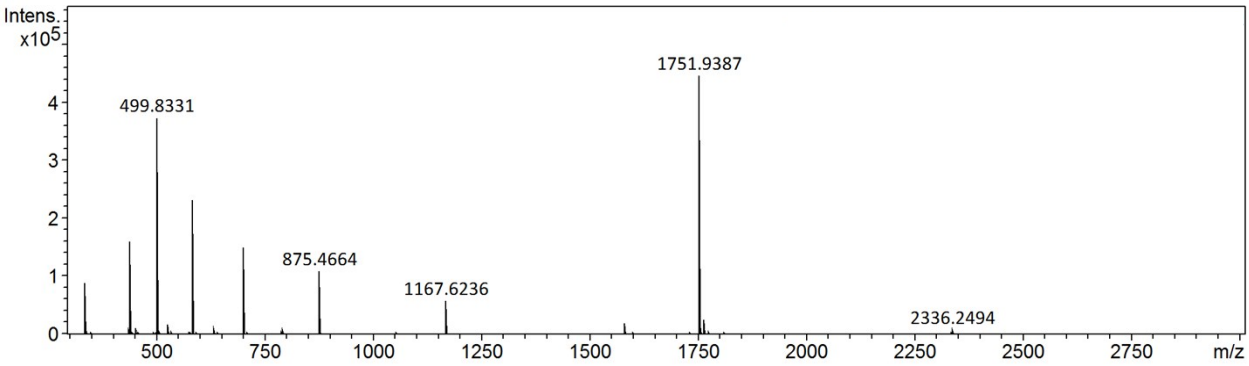
5' – TCAGTGGCTACGGCACCTGT- 3'.

IIIc. LC-ESI-MS characterization

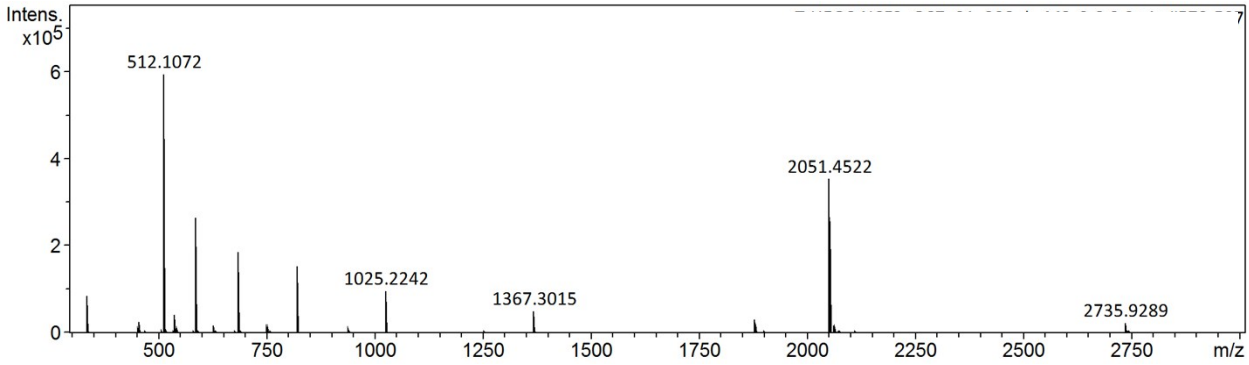
The oligomers were analyzed by LC-ESI-MS in negative ESI mode. Samples were run through an Acclaim RSLC 120 C18 column (2.2µm, 120Å 2.1 x 50mm) using a gradient of mobile phase A (100mM 1,1,1,3,3,3-hexafluoro-2-propanol and 5mM triethylamine in water) and mobile phase B (Methanol) in 8 minutes (2% to 100% B). Liquid chromatography was performed as a control for strand purity. Traces are not reported here since, as expected, they are very clean after RP-HPLC purification.



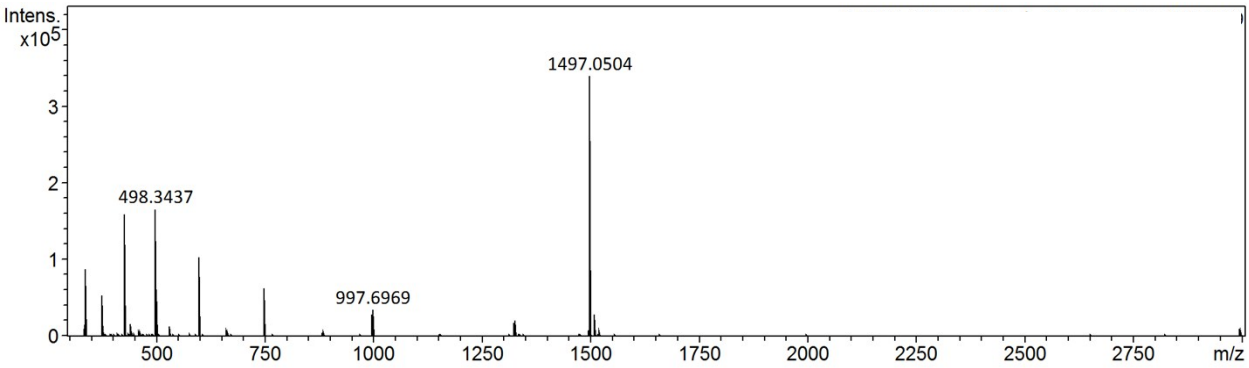
$A_{6,0}$



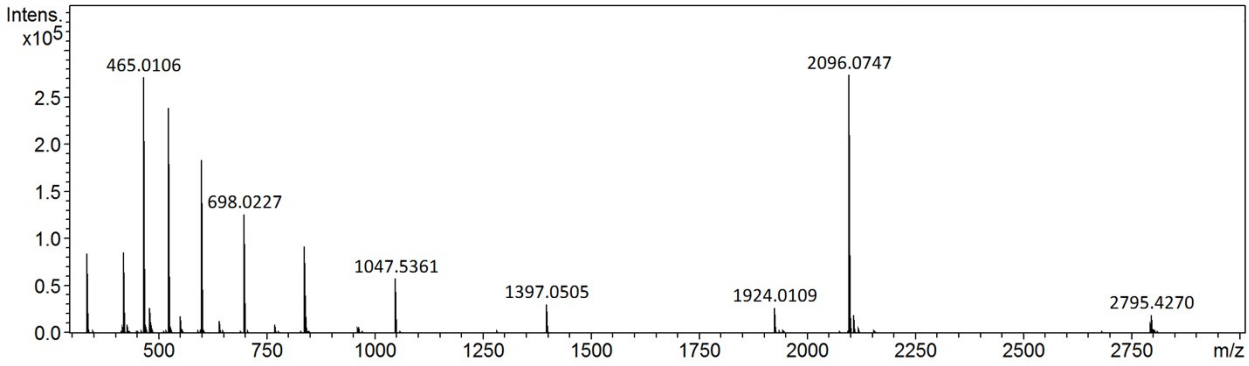
$A_{6,2}$


Intens.
x10^5
512.1072
2051.4522
1025.2242
1367.3015
2735.9289
m/z

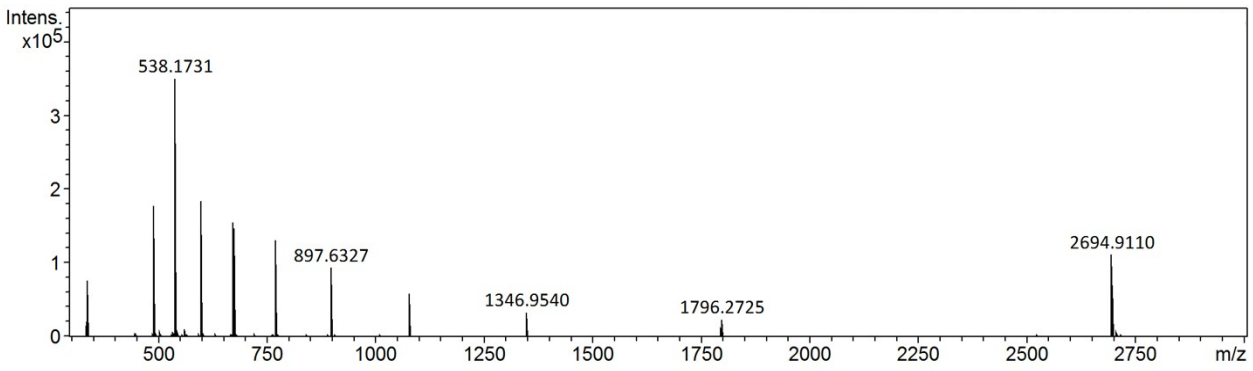
$A_{6,4}$


Intens.
x10^5
1497.0504
498.3437
997.6969
m/z

$A_{8,0}$

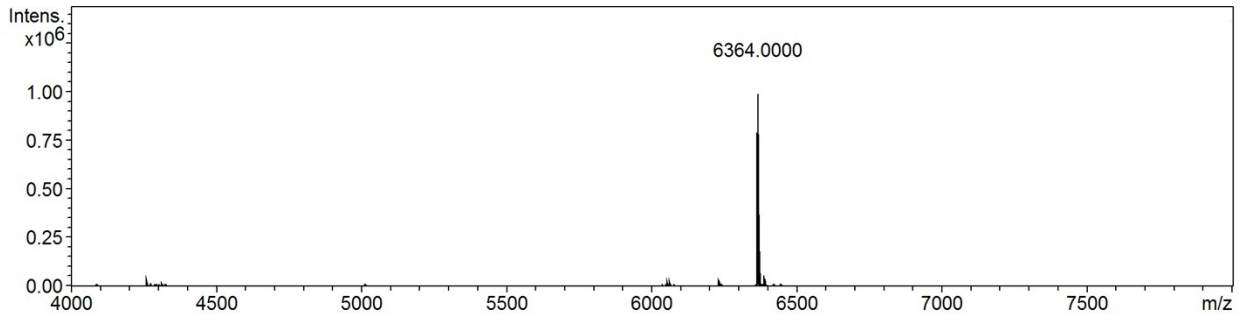

Intens.
x10^5
465.0106
2096.0747
698.0227
1047.5361
1397.0505
1924.0109
2795.4270
m/z

$A_{8,2}$

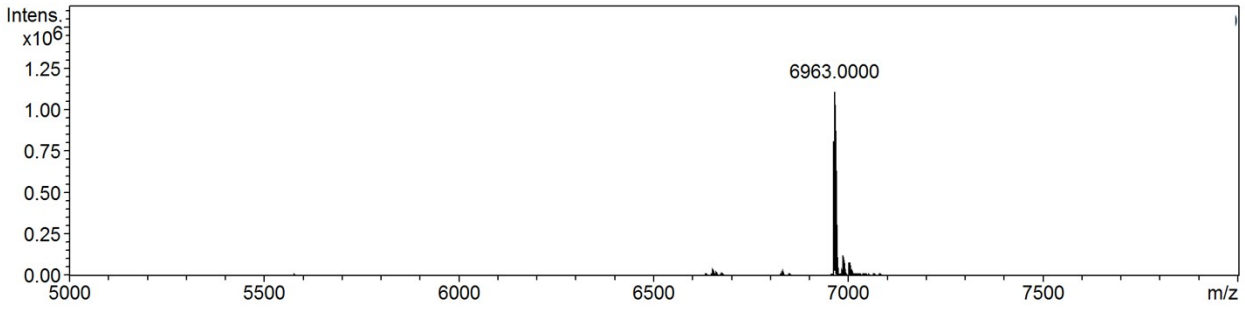


$A_{8,4}$

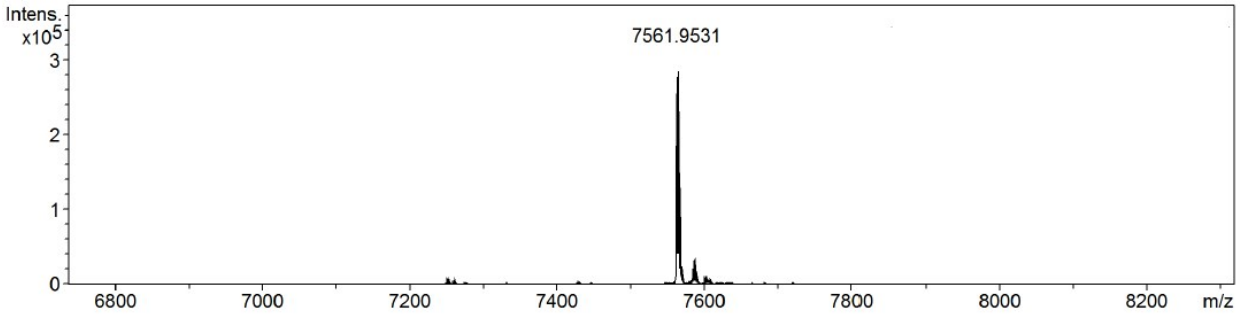
Figure SF3: MS data for $\mathbf{A_{n,m}}$ oligomers. Almost all peaks can be associated with a (M-x)/x anion, as illustrated on the first spectrum. $\mathbf{A_{n,m}}$ stands for T-HEG_n-N$[PFC]_m$.



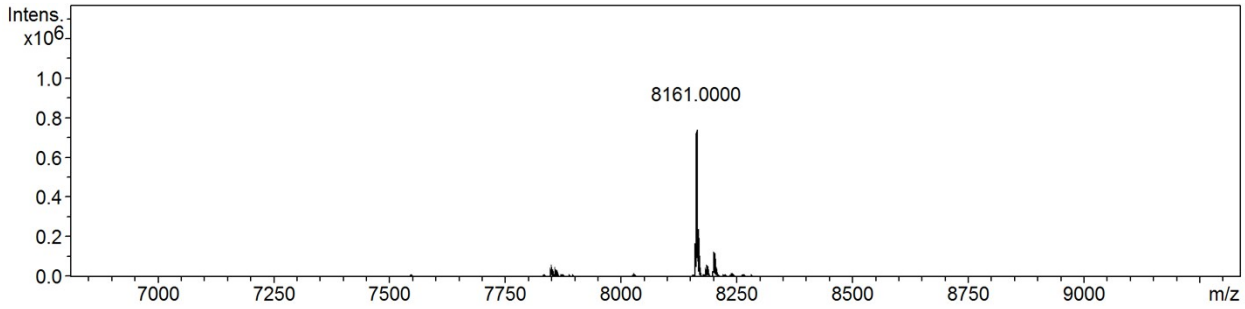
B_1



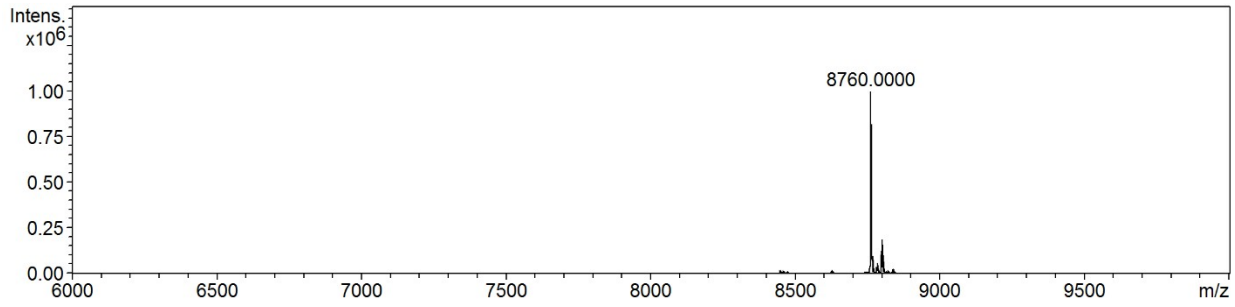
B_2


Intens.
x10^5
3
2
1
0
7561.9531
6800
7000
7200
7400
7600
7800
8000
8200
m/z

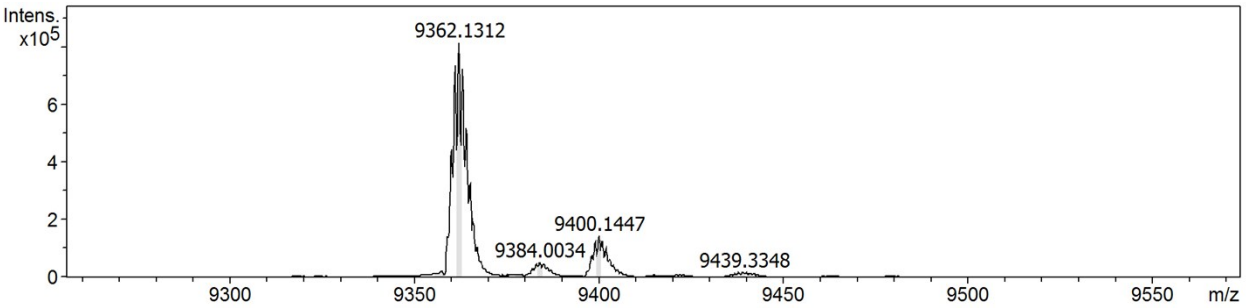
B_3


Intens.
x10^6
1.0
0.8
0.6
0.4
0.2
0.0
8161.0000
7000
7250
7500
7750
8000
8250
8500
8750
9000
m/z

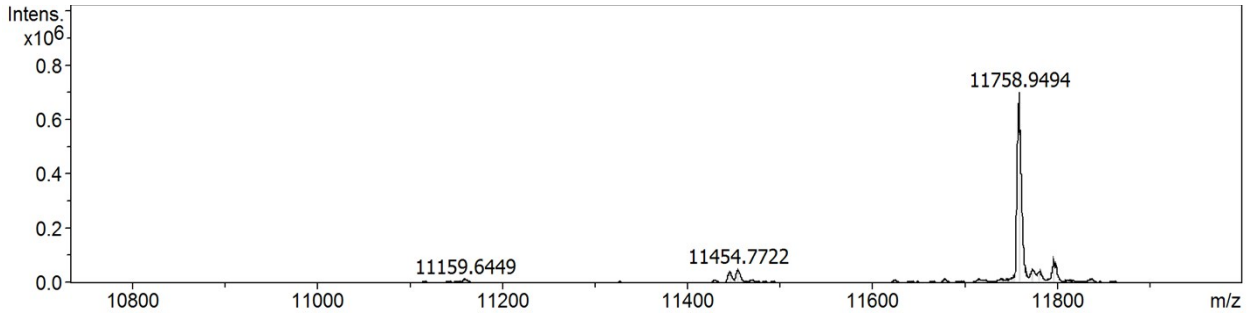
B_4


Intens.
x10^6
1.00
0.75
0.50
0.25
0.00
8760.0000
6000
6500
7000
7500
8000
8500
9000
9500
m/z

B_5


Intens.
x10^5
6
4
2
0
9362.1312
9400.1447
9384.0034
9439.3348
9300
9350
9400
9450
9500
9550
m/z

B_6

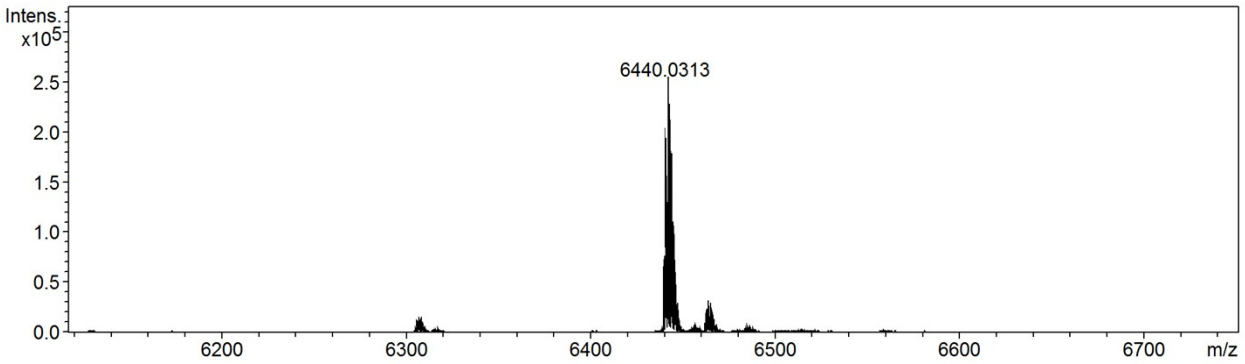


B_{10}

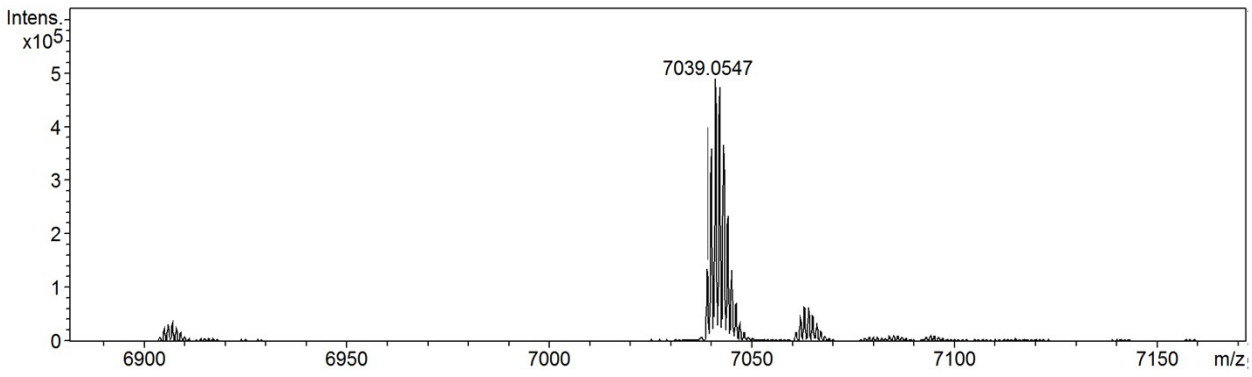
Figure SF4: LC/MS data for DNA (sequence is called AT)-"Teflon" polymers (**B_n**). The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Masses reported are exact masses except for big conjugates (>10kDa).

Table ST3: LC/MS results for DNA with internal modifications or 3'end modification.

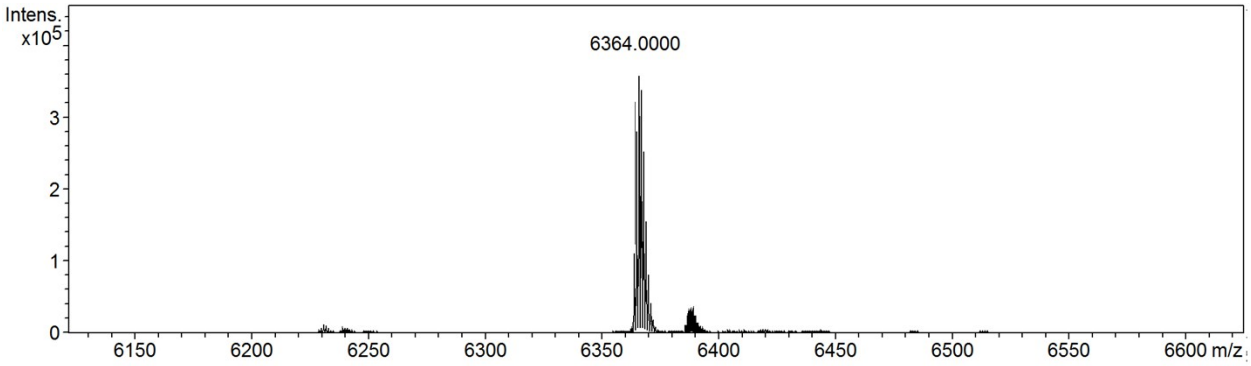
Strand name	Calculated exact mass (g/mol)	Measured mass (g/mol)
(N[PFC])-cpAT	6440.05	6440.03
(N[$PFC]_2$)-cpAT	7039.06	7039.05
AT-(N[PFC])p10	6364.01	6364.00
AT-(N[$PFC]_2$)p10	6963.02	6962.97
AT-(N[$PFC]_3$)p10	7562.03	7561.95
cpAT-(N[PFC])p11	6440.05	6440.03
cpAT-(N[$PFC]_2$)p11	7039.06	7039.09
cpAT-(N[$PFC]_3$)p11	7638.07	7638.01
AT-N[PFC]p7p14	6963.02	6962.97
cpAT-N[PFC]p7p14	7039.06	7039.09



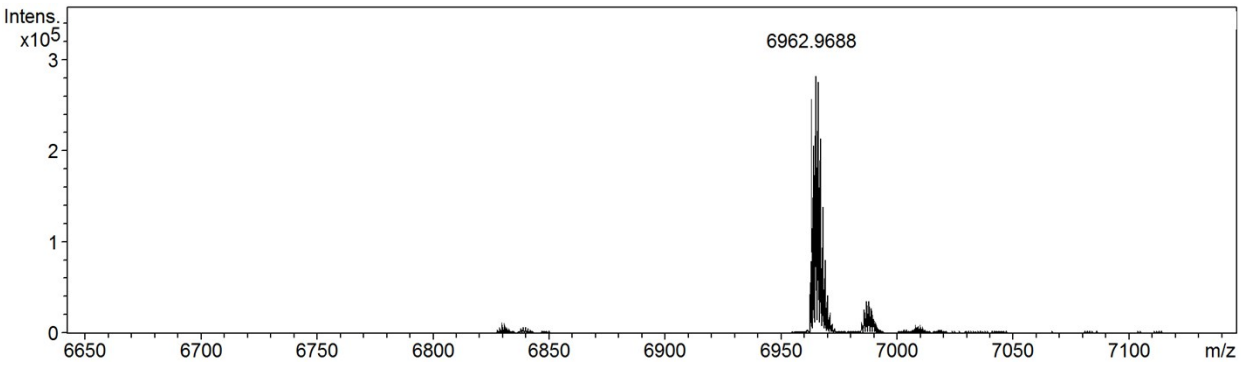
(N[PFC])-cpAT



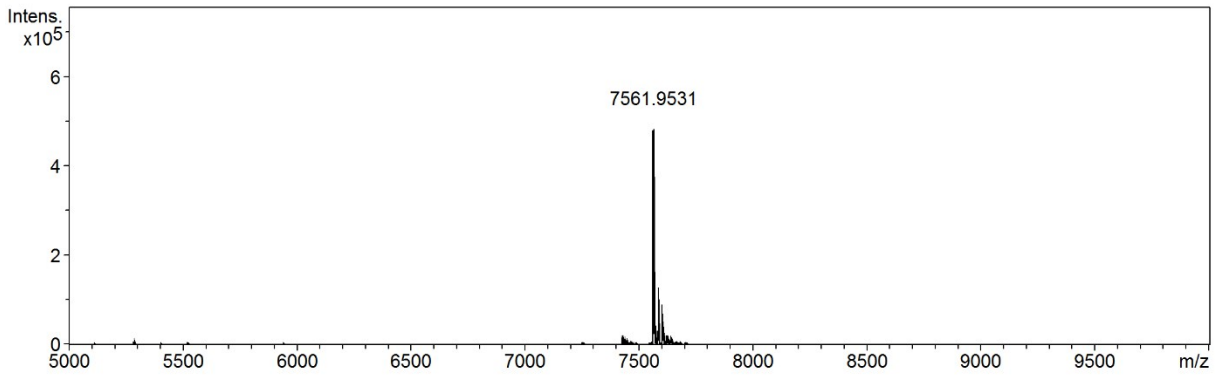
(N[PFC])$_2$-cpAT



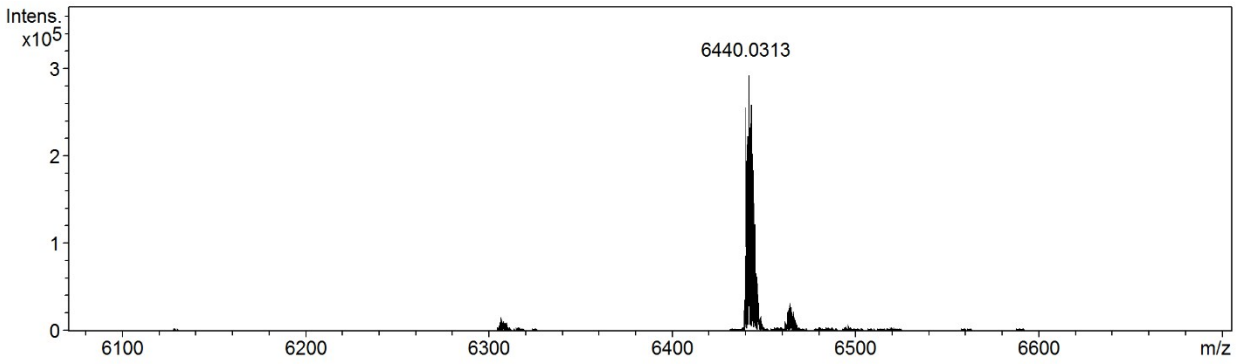
AT-(N[PFC])p10



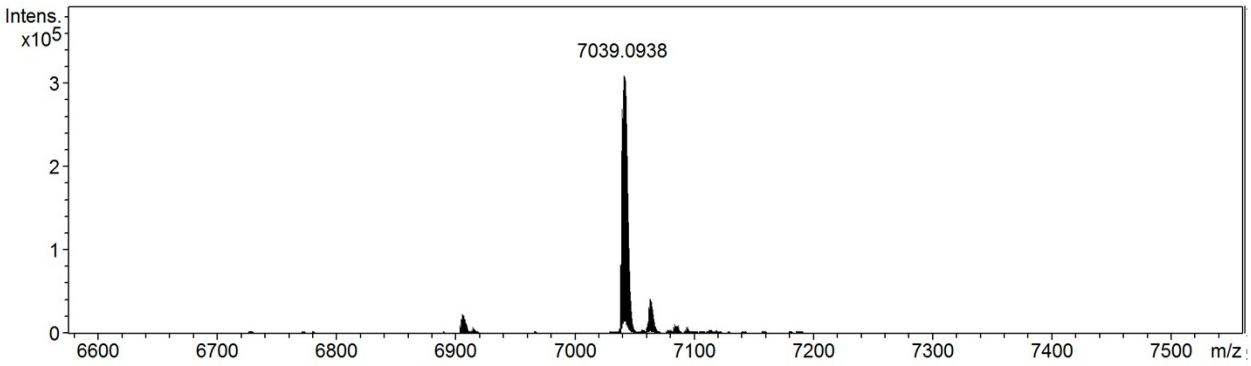
AT-(N[PFC])$_2$p10



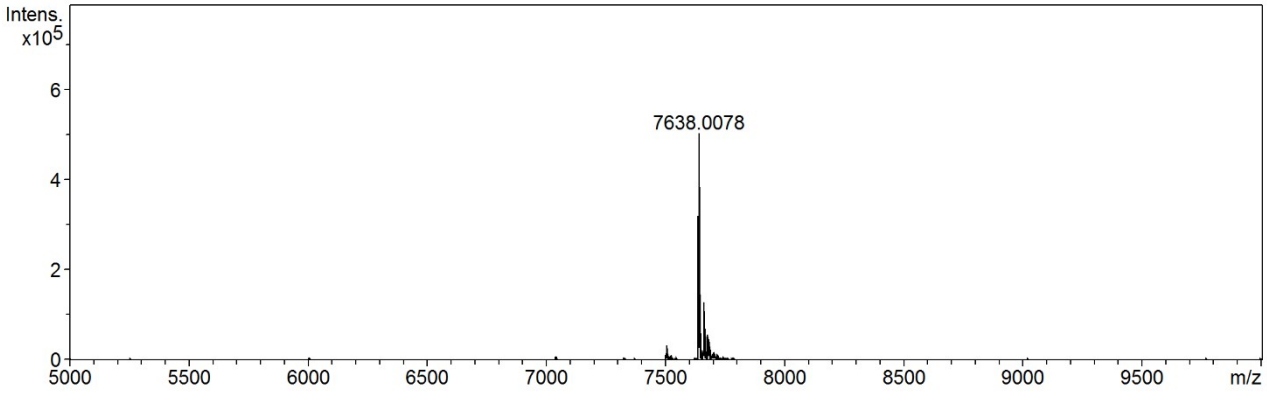
AT-(N[PFC])$_3$p10



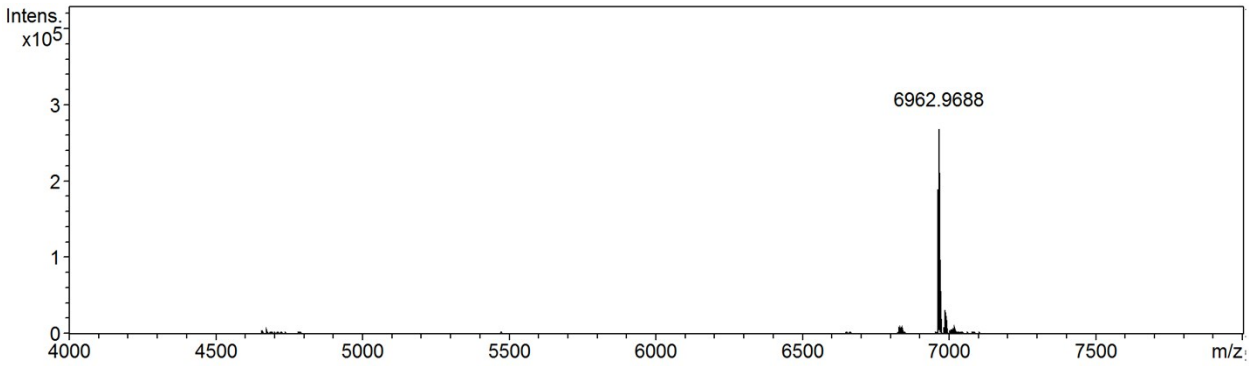
cpAT-(N[PFC])p11



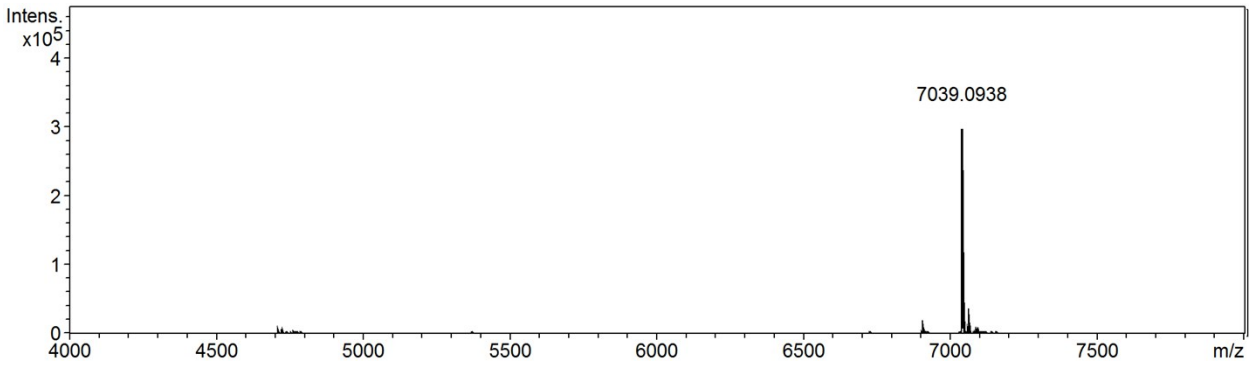
cpAT-(N[PFC])$_2$p11



cpAT-(N[PFC])$_3$p11



AT-N[PFC]p7p14



cpAT-N[PFC]p7p14

Figure SF5: LC/MS data for DNA with internal modifications or 3'end modification. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1.

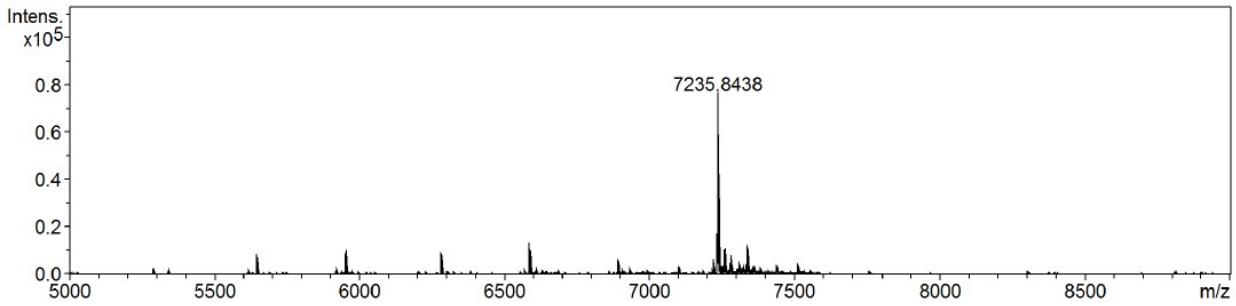


Figure SF6: LC/MS data for RNA with 3'end N[PFC] modification. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Calculated exact mass: 7235.94, Obtained: 7235.84.

SI-IV Self-assembly investigation

a. Electrophoretic mobility assays

15% denaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out at room temperature for 30 minutes at 250V followed by 1 hour at 500V. TBE buffer (1X) was used and the concentration of urea in the gel was 7M. For each lane 5 µL of sample (2µM) in water was added to 5uL of 8M urea. The DNA bands for all gels were visualized by incubation with GelRed™.

1: B_0

2: B_1

3: B_2

4: B_3

5: B_4

6: B_5

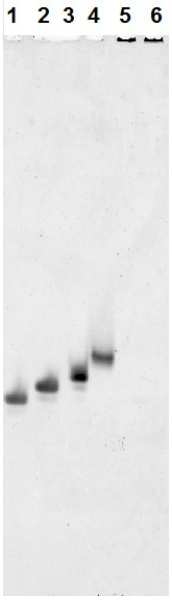


Figure SF7: 15% denaturing PAGE analysis of DNA-Teflon polymers. The DNA sequence is the AT one and the length of the PFC tail vary from 0 to 5 PFC units ($\mathbf{B_n}$ polymers). From n=4, self-assembly occurs and the material does not seem able to penetrate in the gel (non-penetrating band).

2.5% agarose gel electrophoresis (AGE) was carried out in two different buffers. The first one, TAE, does contain EDTA whereas the second one, TAMg contains Mg^{2+} cations. They were carried out at 4°C for respectively 2h and 2h15 at 80V. Gel was cast in the appropriate buffer and the samples were at a concentration of 1 µM in the appropriate buffer. 2µl of glycerol were added to the samples before loading. The DNA bands for all gels were visualized by incubation with GelRed™.

1: Ladder
2: B_0
3: B_1
4: B_2
5: B_3
6: B_4
7: B_5
8: B_6
9: B_{10}

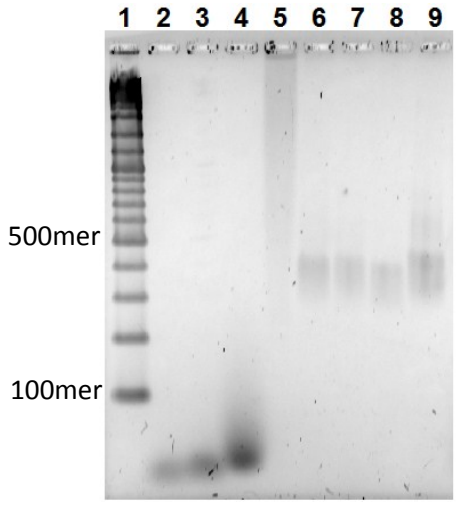


Figure SF8: 2.5% AGE analysis of the different AT-(N[PFC])$_n$ strands in TAE. Aggregation seems to occur from n=3 but cleaner assembled products appear from n=4. According to the ladder used, their hydrodynamic volume is greater than that of a 400mer. Mobility difference between each band of the ladder corresponds to a 100nucleotide difference.

1: Ladder
2: B_0
3: B_1
4: B_2
5: B_3
6: B_4
7: B_5
8: B_6
9: B_{10}

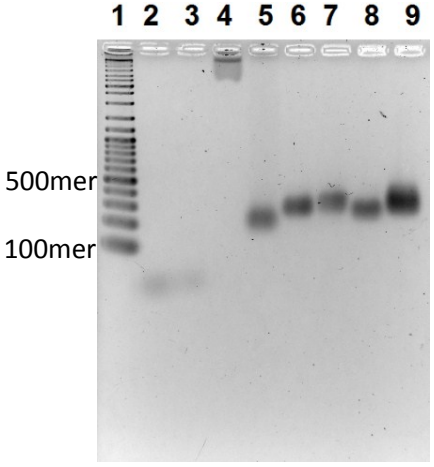
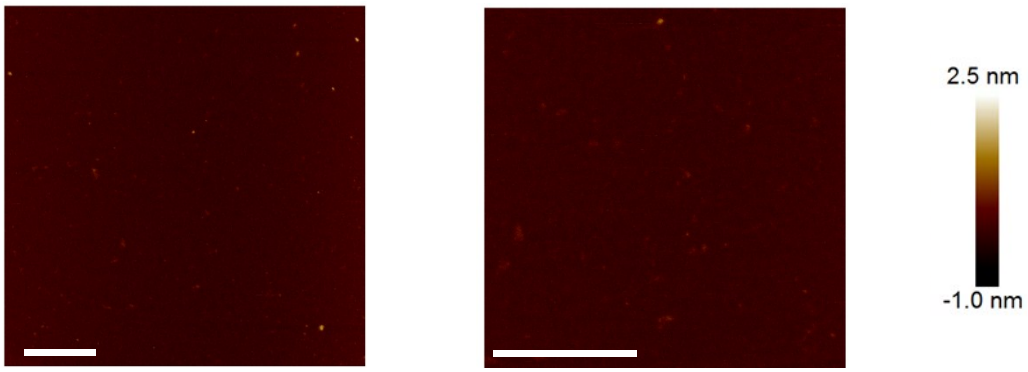


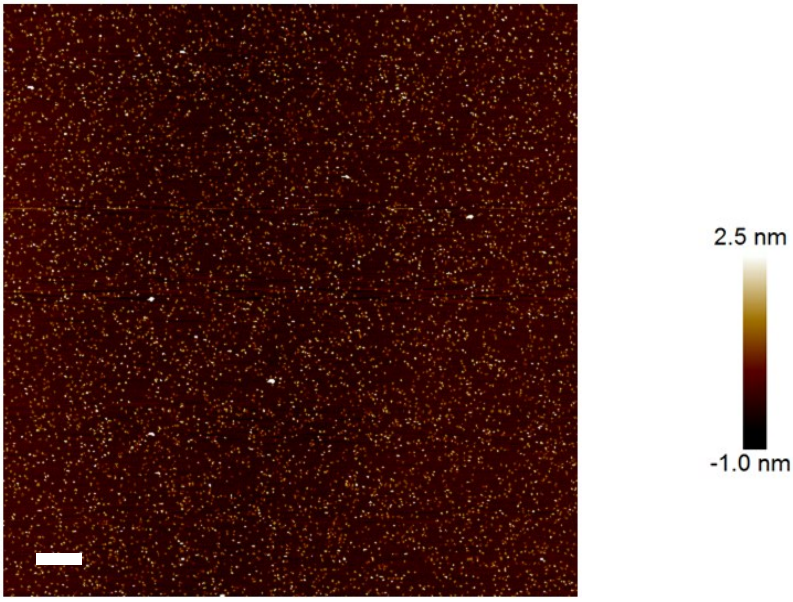
Figure SF9: 2.5% AGE analysis of the different AT-(N[PFC])$_n$ strands in TAMg. Aggregation seems to occur from n=2 but clean assembled products appear from n=3. According to the ladder used, their hydrodynamic volume is greater than that of a 400mer. Mobility difference between each band of the ladder corresponds to a 100nucleotide difference.

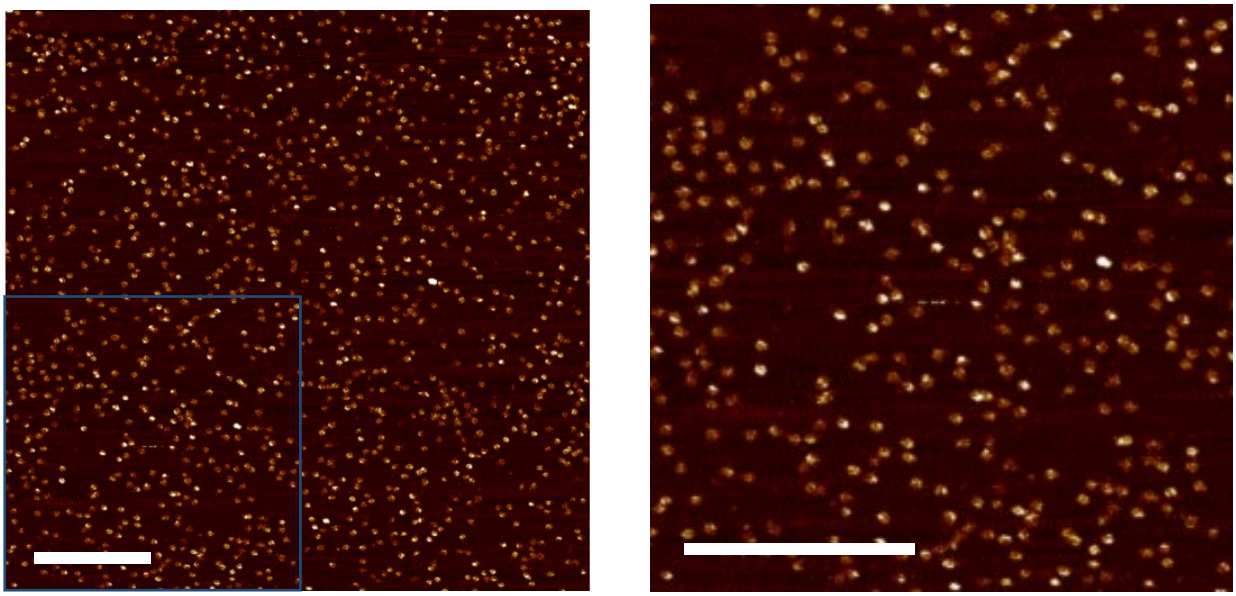
b. Atomic force microscopy

Samples were diluted to 1μM in TAMg buffer and 4μL of this solution was deposited on a freshly cleaved mica surface (ca. 7 x 7mm) and allowed to adsorb for 1-2 seconds. Then 50μL of 0.22μm filtered Millipore water was dropped on the surface and instantly removed with filter paper. The surface was then washed four times with 80μL of water and the excess removed with a strong flow of nitrogen. Samples were dried under vacuum for 15-30 minutes prior to imaging.

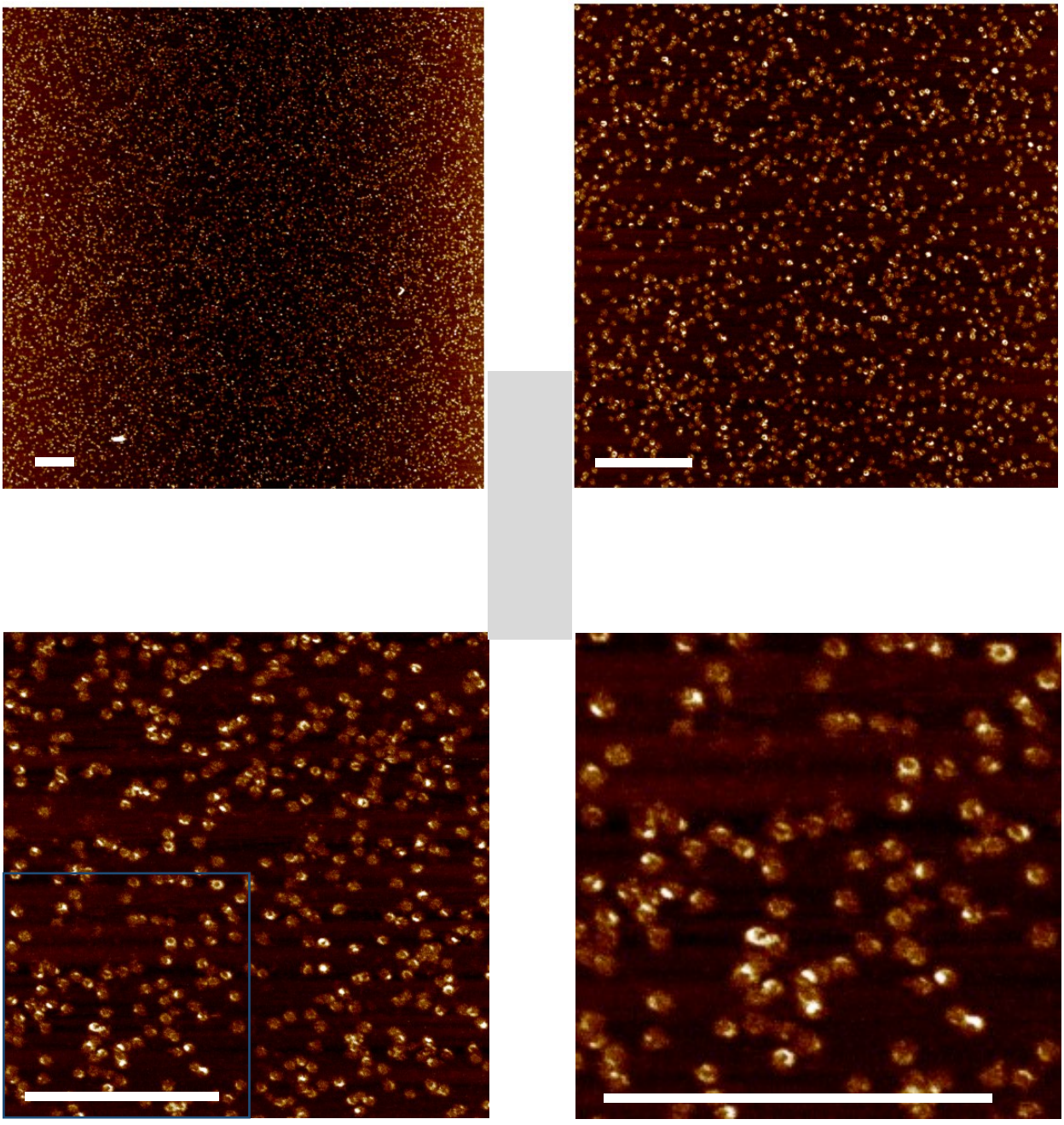


AT strand as a control





AT-(N[PFC])$_5$ micelles. The image in the bottom-right corner is a zoom-in of the bottom- left image (blue square).



AT-(N[PFC])$_{10}$ micelles. The image in the bottom-right corner is a zoom-in of the bottom- left image (blue square).

Figure SF10: Typical height AFM images of AT sequence as a control, **B$_5$** and **B$_{10}$** micelles. White bars represent 400 nm. Statistics were realized by hand using the software Image J. We found an average diameter of 19.4 ± 3.1 nm for AT-(N[PFC])$_5$ micelles and 20.9 ± 3.3 nm for AT-(N[PFC])$_{10}$ micelles. We suppose the structures we see are flat and wide due to the drying effect while depositing the micelles on mica.

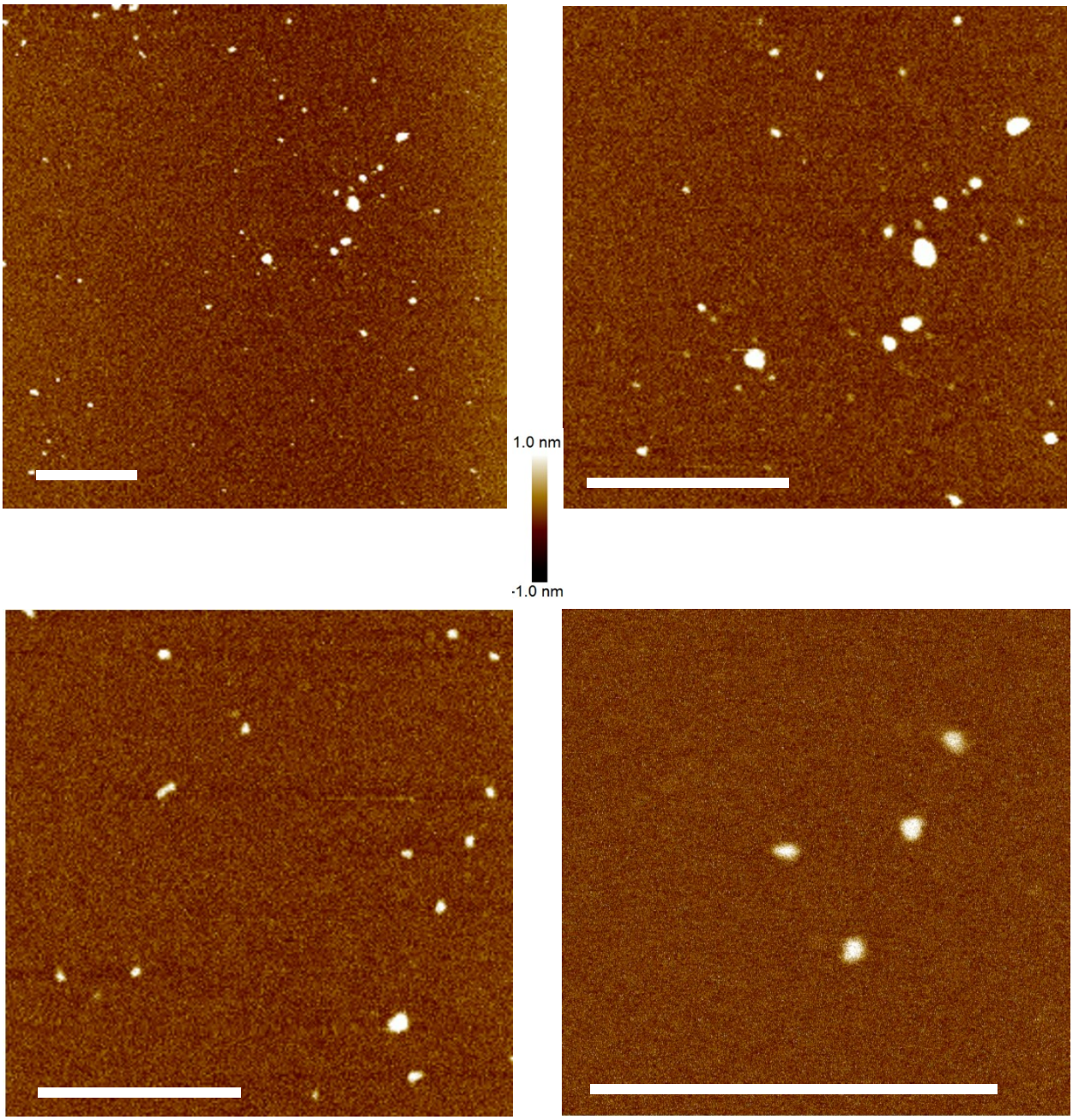


Figure SF11: Typical height AFM images of $\mathbf{B_2}$. White bars represent 400 nm. Again, we suppose the structures we see are flat and wide due to the drying effect while depositing the micelles on mica. A clean background was difficult to obtain as the height of the micelles is lower than with more N[PFC] units. Aggregation was clear on each acquired image but we could not obtain monodisperse micelle population (DLS shows narrow polydispersity). It is probably due to the smaller PFC core.

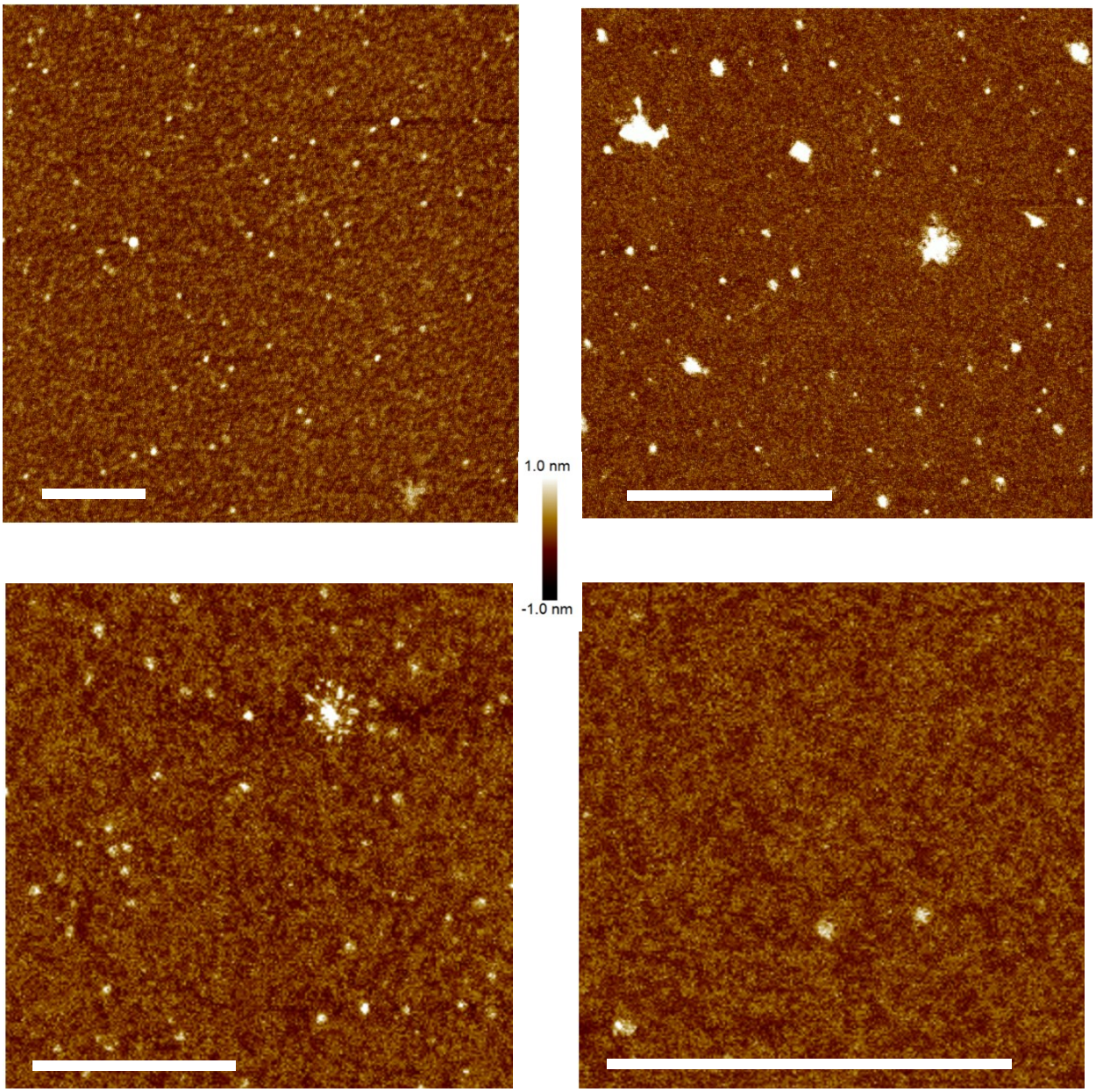


Figure SF12: Typical height AFM images of $\mathbf{C_6}$ (AT-(N[PFC])$_2$/(N[PFC])$_2$-cpAT duplex) aggregation. White bars represent 400 nm. Aggregation was clear on each acquired image but we could not obtain monodisperse micelle populations (DLS shows narrow polydispersity). It is probably due to the smaller PFC core.

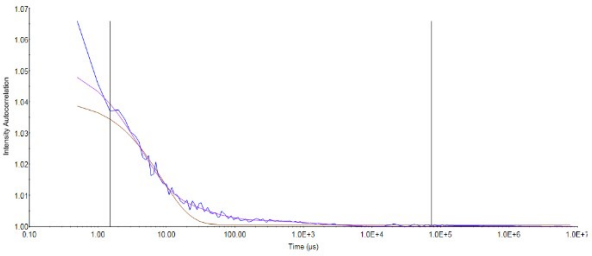
c. Dynamic light scattering

A cumulants fit model was used to confirm the presence and determine the size of a monomodal population of micellar aggregates. Sterile water and TAMg were filtered using a 0.2 µm nylon syringe filter before use for DLS sample preparation. All measurements were carried out at 20°C. Concentration

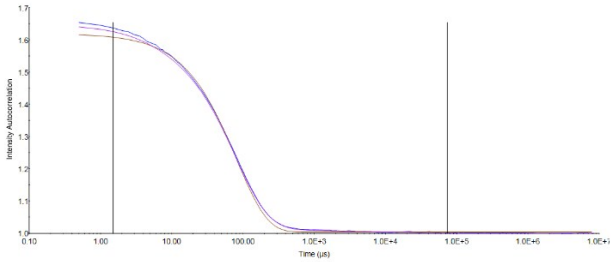
of the sample is 10µM if not specified. All the measurements were at least triplicated. For experiments with Mg^{2+}, samples were directly diluted in TAMg or Mg^{2+} ions were added after solubilizing our constructs in sterile water : both methods led to similar results.

Table ST4. Measured micelles radii as a function of number of PFC units attached to DNA and Mg^{2+} concentration at 20°C. Standard deviation are reported. At 37°C, the dynamic radius of AT-(N[PFC]4) micelles is 7.5 ± 0.4 nm in H_2O.

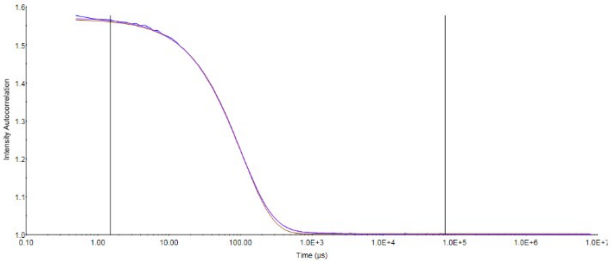
Number of N[PFC] units (n)	Presence of Mg^{2+} (7.6 mM)	Measured dynamic radius (nm)
0-1	+/-	NO
2	-	NO
	+	6.2 ± 0.1
4	-	8.1 ± 0.4
	+	6.3 ± 0.3
5	-	8.6 ± 0.3
	+	6.5 ± 0.4
10	-	8.5 ± 0.3
	+	8.5 ± 0.1



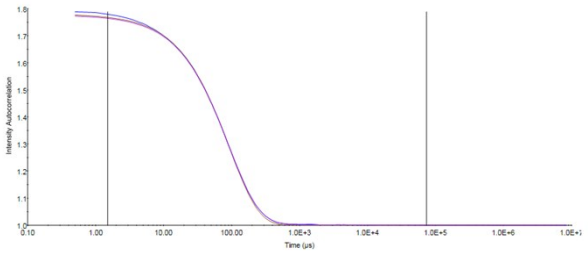
AT (Mg^{2+})



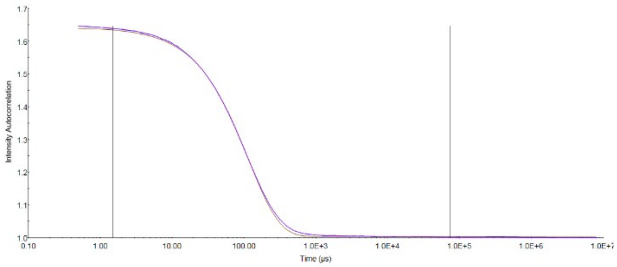
AT-$(N[PFC])_2$ (Mg^{2+})



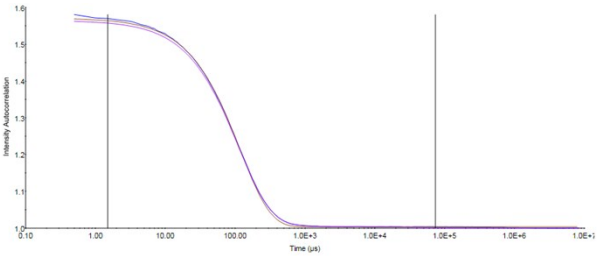
AT-$(N[PFC])_4$



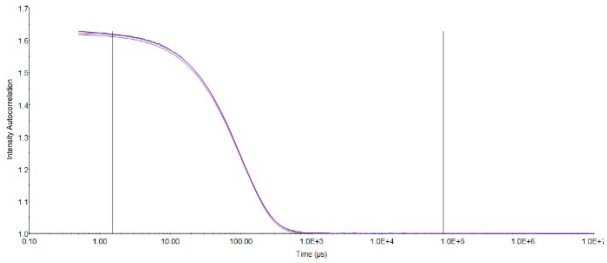
AT-$(N[PFC])_4$ (Mg^{2+})



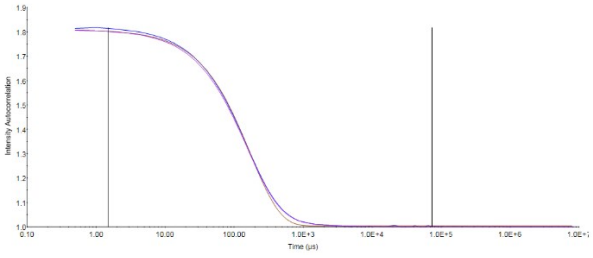
AT-$(N[PFC])_5$



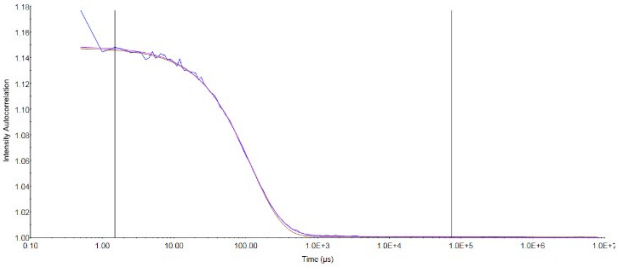
AT-$(N[PFC])_5$ (Mg^{2+})



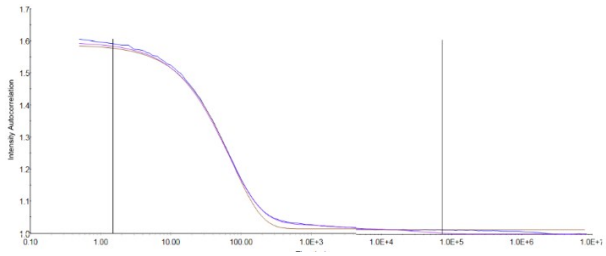
AT-$(N[PFC])_{10}$



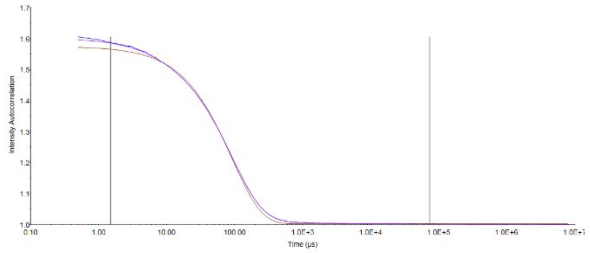
AT-$(N[PFC])_{10}$ (Mg^{2+})



T-HEG$_8$-(N[PFC])$_4$



AT-(N[PFC])$_4$ 37°C



AT-(N[PFC])$_2$ double stranded

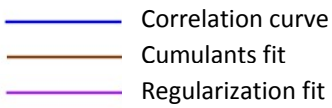


Figure SF13: Dynamic light scattering on sequence controlled polymers. Representative DLS intensity correlation functions for 10µM solutions. The low scattering intensity and poor correlation functions measured for AT is characteristic of individual molecules in solution. In contrast the data for self-assembling oligomers reveals excellent correlation except for T-HEG$_8$-(N[PFC])$_4$ for which data is less convincing but still satisfying. Polydispersity in the case of self-assembled material was **never exceeding 16%** showing the narrow polydispersity of the observed structures. (Mg^{2+}) means Mg^{2+} has been added to reach a concentration of 7.6 mM.

d. NMR study

^{19}F NMR spectra of the following molecules were recorded for quantitation using the PULCON method: molecule **2** in DMSO-d$_6$ (external reference), and **B$_1$**, **B$_4$**, and **B$_{10}$** in D_2O. The concentrations of **B$_1$**, **B$_4$**, and **B$_{10}$** were also determined by UV absorption at 260nm. The NMR analysis focused on the CF_3 signal of molecule **2** related to the CF_3 end group of the C_8F_{17} chain as the CF_3 signals were the most intense. T_1 and T_2 values were measured using an inversion recovery sequence and a CPMG sequence, respectively, using between 5 and 8 appropriate tau values. Fitting was done using the Bruker Dynamics Center.

Table ST5. ^{19}F relaxation times in D_2O.

Molecule	T_1 (s)	T_2 (s)
2 (in DMSO)	1.28 ± 0.09	
B_1	0.7 ± 0.1	0.185 ± 0.02
B_4	0.161 ± 0.05	0.007 ± 0.002
B_{10}	0.148 ± 0.02	0.0020 ± 0.0004

All quantitative ^{19}F spectra were acquired using the same experimental conditions, with the exception of the transmitter offset, which was always placed on the CF_3 peak (-80.1 ppm for molecule **2**, -80.8 ppm for **B_1** and -83.3 ppm for **B_4**, and **B_{10}**. The DEPTH method[1] along with a pre-acquisition delay of 35 µs and the Bruker baseopt method for baseline correction at acquisition time were used to suppress the large ^{19}F background arising from probe components. A slight baseline hump was still present in the samples with the lowest S/N ratios (B_4 and B_{10}) and the Bruker routine cryoproc1d was used to replace the first points of those FIDs with backwards linearly predicted points. Automatic tuning and matching were performed before each experiment. Because signal-to-noise was poor, the ^{19}F 90° pulse was not measured before each experiment, but the ^{1}H 90°pulse was measured for all samples and found not to vary, so it was assumed that the ^{19}F was also consistent for all samples. In the quantitative NMR experiments, a recycle delay of 6 s was used for best reproducibility (the T_1 of molecule **2** was found to be 1.3 s). The temperature was controlled at 300K during the experiments. The receiver gain was kept constant at 203 for all experiments.

Concentration calculations were done within the ERETIC 2 implementation in Bruker TopSpin 3.5 pl 2 software, based on the PULCON (pulse length based concentration determination) method[2]. The external reference, molecule **2**, was prepared using careful weighing and dilution. The results of the measurement as applied to the CF_3 peaks of each sample are shown in Table ST6.

Table ST6. Quantitative NMR data.

Molecule	$[B_n]_{UV}$[a] (μM)	$[B_n]_{NMR}$[c] (μM)	Ratio $[B_n]_{NMR}/[B_n]_{UV}$
2	800[b]		
$\mathbf{B_1}$	71	81	1.1
$\mathbf{B_4}$	39	35	0.9
$\mathbf{B_{10}}$	46	40	0.9

[a] Concentration determined by UV absorption at 260nm except as noted.[b] Determined by weighing and precisely diluting the compound. [c] Concentration determined by NMR. No more significant figures are given, especially as the UV measurements could be influenced by self-assembly.

Sensitivity measurements: We measured the signal-to-noise ratios (S/N) of $\mathbf{B_4}$ and $\mathbf{B_{10}}$ in samples of the same concentration (46 μM). Conditions were the same as explained above except that the acquisition time was reduced to 28 ms, the recycle delay was reduced to 0.46 s, and the number of scans was set to 1800, with 8 dummy scans, for a total experimental time of 15 min 33s. To further increase the S/N of the 46 μM $\mathbf{B_{10}}$ sample, the acquisition time was reduced to 14 ms, the recycle delay was reduced to 0.43 s, and the number of scans was set to 1900 with 8 dummy scans, for a total experiment time of 15 min 10 s.

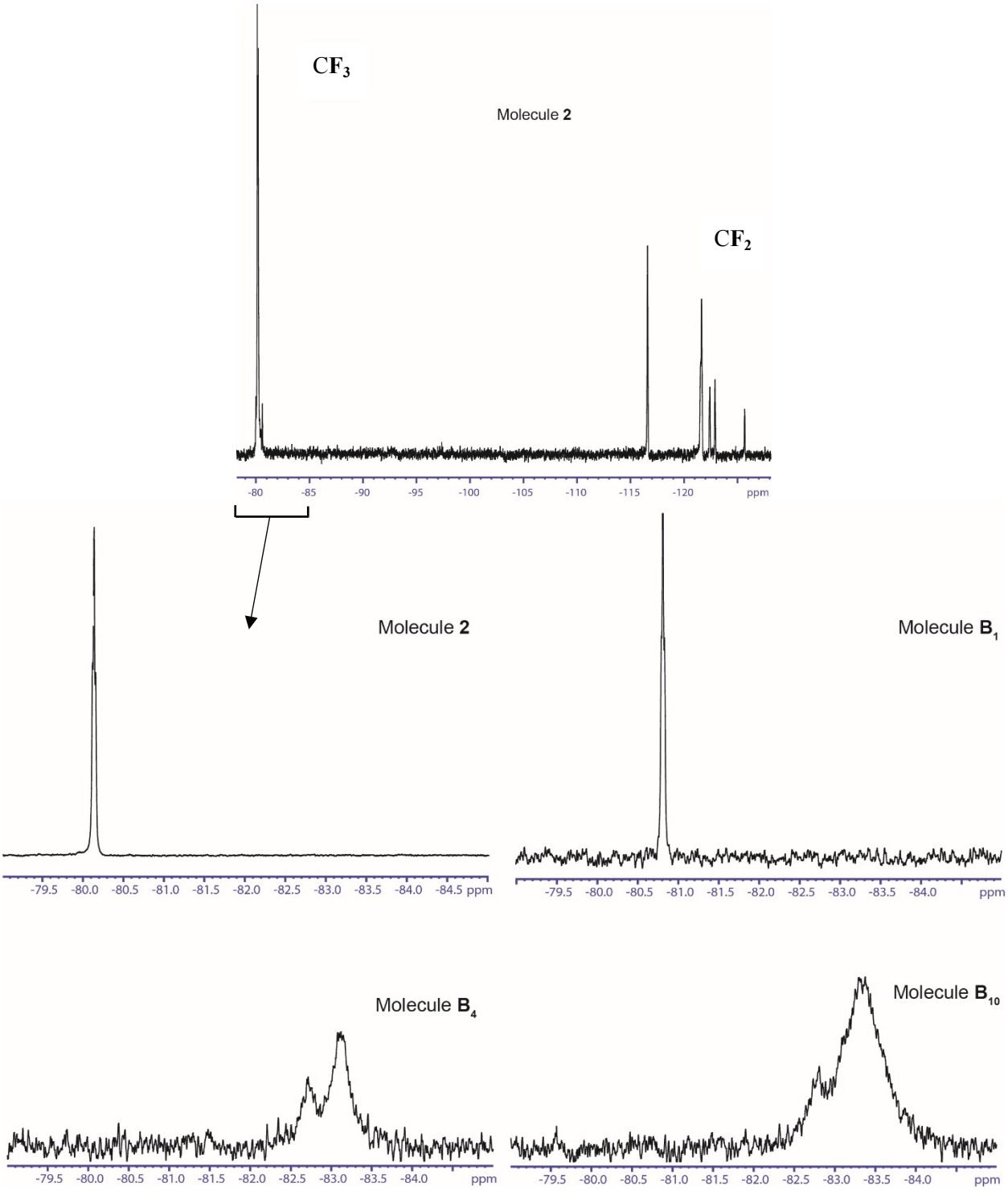


Figure SF14. Top: ^{19}F NMR spectrum of molecule **2** (80 μM in DMSO-*d6*). Bottom: ^{19}F NMR spectra of molecules **2** (800 μM in DMSO-*d6*), $\mathbf{B_1}$ (71 μM in D_2O), $\mathbf{B_4}$ (39 μM in D_2O), $\mathbf{B_{10}}$ (46 μM in D_2O). Only the strong signal from CF_3 is visible at low concentrations.

SI-V Duplex hybridization investigation and properties

a. Electrophoretic mobility assays

8% Native PAGE was carried out at room temperature for 2.5 hours at a constant voltage of 250V in 1X TAMg buffer. Sample loading was 0.01 nmol ssDNA or 0.002 nmol dsDNA per lane (12 µl samples in 1 X TAMg, including 2µl of glycerol).

1' – C_1: AT/cpAT

2' – C_5: AT-N[PFC]/N[PFC]-cpAT

3' – C_2: AT-N[PFC]p10/cpAT-N[PFC]p11

4' – C_3: AT-N[PFC]$_2$p10/cpAT-N[PFC] 211

5' – C_6: AT-N[PFC]$_2$/ N[PFC]$_2$-cpAT

6' – AT-(N[PFC])$_3$p10/ cpAT-(N[PFC])$_3$p11

7' – C_4: AT-N[PFC]p7p14/cpAT-N[PFC]p7p14

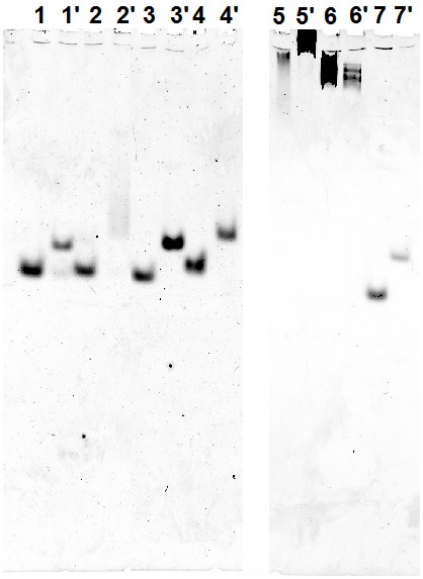


Figure SF15: Electrophoretic mobility of DNA duplexes (19mer) modified with PFCs. Numbers without ' represent the corresponding single strand with the AT sequence. Lanes 5 to 7' belong to a different gel. The ssDNA and duplexes behave well except in lanes 2', 5, 5', 6 and 6' for which self-assembly occurs. For duplexes 2 and 5, we believe aggregation occurs but DNA hybridization remains (cf. melting curves).

b. Melting temperature determination

Experiments were carried out in quartz cuvettes (rectangular, 10mm, 80µL) in triplicate. Absorbance was measured at 260nm and detected in increments of 1°C from 20°C to 50°C and from 75°C to 90°C of 0.4°C from 50°C to 75°C. Concentration of DNA duplexes was 4 µM in 1XTAMg buffer.

T_M are calculated by taking the temperatures corresponding to the derivative maxima of the curves obtained. Representative curves, melting temperature averages with standard deviations are shown on figure 6.

Table ST7. Melting temperatures of the duplexes analyzed

Duplex name	Strands in the duplex	Melting temperature (°C)
C_1	AT/cpAT	62.6 ± 0.5.
C_2	AT-(N[PFC])p10/cpAT-(N[PFC])p11	63.1 ± 0.4
C_3	AT-$(N[PFC])_2$p10/cpAT-$(N[PFC])_2$p11	70.6 ± 0.5
C_4	AT-(N[PFC])p7/cpAT-(N[PFC])p14	63.3 ± 0.5
C_5	AT-(N[PFC])/N[PFC])-cpAT	65.3 ± 0.7
C_6	AT-$(N[PFC])_2$/$N[PFC])_2$-cpAT	82.6 ± 0.8

c. Nuclease assays

For degradation studies, DNA duplexes were concentrated to a stock solution of 40 μM in 1XTAMg buffer. As an example, duplex AT/cpAT (40 μM, 2 μL) was first diluted with DMEM media (88 μL). To this mixture was added a fresh sample of undiluted FBS (10 μL) with slight mixing to make the overall % of FBS 10% (v/v). An aliquot was immediately taken out (10 μL), formamide (10 μL) added and then stored at -20 °C as the t =0 h time point. The remaining sample was then incubated at 37 °C and similar aliquots were removed and treated as described above at time points of 35 min, 1, 2, 4, 6, 8.5 and 24 hr. Digested products were analyzed by denaturing PAGE (20%, 15 mA, 250V during 30 min followed by 500V 2 hr).

Higher band intensity was extracted and divided by the intensity at t=0. These experiments were run in triplicate. Data was analyzed using GraphPad Prism with a one phase decay exponential fit. The general equation is : `Y=(Y0-Plateau)*exp(-K*X)+Plateau.` Here, Y is the relative band intensity, X is time, the plateau is supposed to be 0 while Y0 is equal to one. Half-life is defined as ln(2)/K.

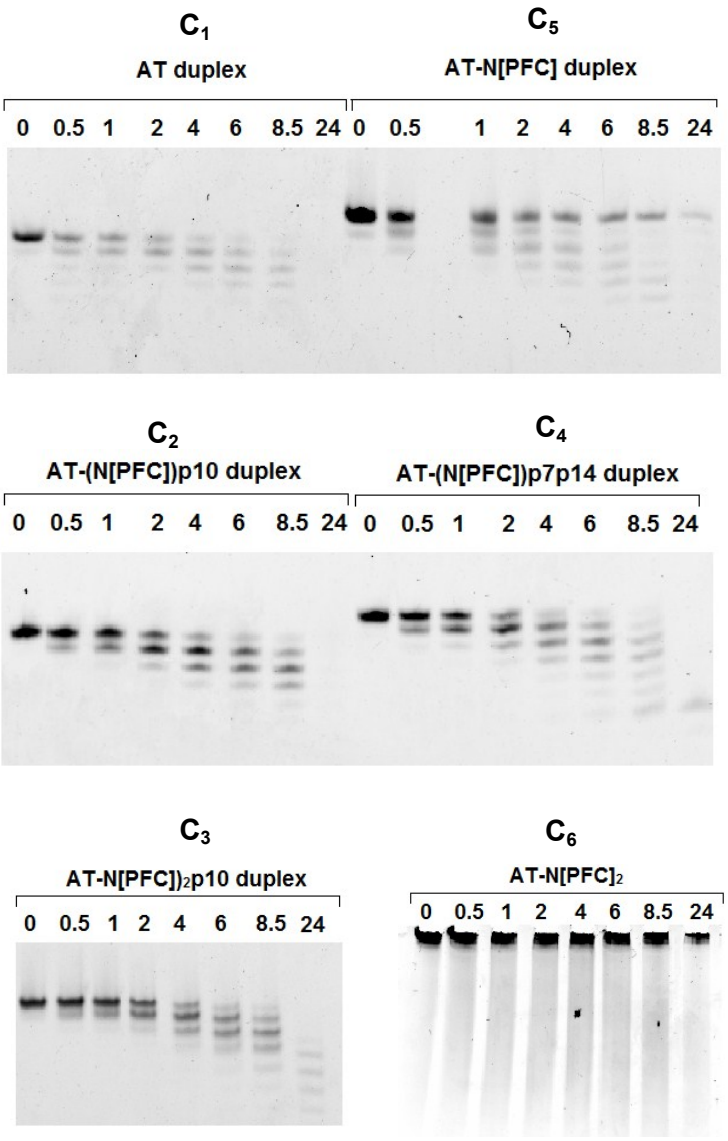


Figure SF16A. Representative gels after FBS degradation assays. The name of the analyzed strand is on top of each gel and numbers correspond to the time the aliquot were taken (h). One would notice that C_6 leads to non-penetrating bands only. This is probably due to self-assembly in DMEM+10%FBS (presence of mono/divalent cations).

Table ST8. Summary of exponential decay analysis.

Duplex name	R squared	Decay constant K (h^{-1})	Half-life (h)
C_1	0.99	0.75	0.9
C_2	0.98	0.34	2.0
C_3	0.90	0.26	2.7
C_4	0.97	0.45	2.2
C_5	0.96	0.57	1.2
C_6	NO	NO	NO

The one-phase decay model may not be fully adapted to a few strands with poor R square but this fit was used in all cases allowing comparisons. Everytime, Y0 and plateau are respectively very close to 1 and 0 except for strand C_5 for which plateau is equal to 0.12. This result highlights that the model chosen is not adapted to this strand.

Only non-penetrating material was observed in the case of $\mathbf{C_6}$ analysis. We hypothesized that the absorbance in the well is proportional to the total amount of oligonucleotides. This quantity was compared to the total intensity in each line during $\mathbf{C_1}$ analysis. It appeared that almost 50% of the intensity remains after 24h while 3% was detected after 24h in the case of $\mathbf{C_1}$ and 5% in the case of $\mathbf{C_6}$. After 8.5h, more than 75% of the intensity remains for $\mathbf{C_6}$ compared to 20% in the case of $\mathbf{C_1}$.

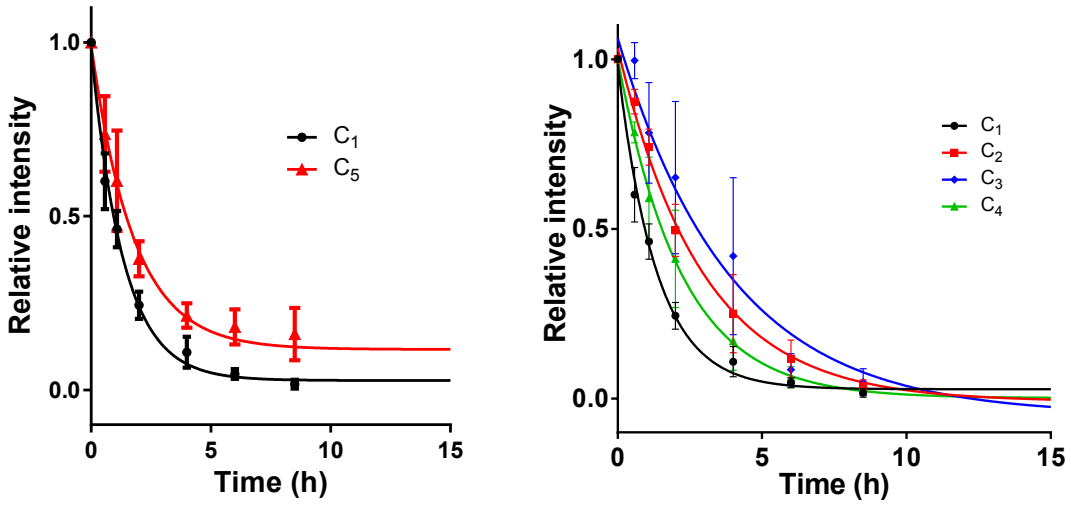


Figure SF16B. Average of degradation curves with error bars. One can see that the full length product is still visible after 24h in the case of $\mathbf{C_5}$ which makes the one phase decay model very inaccurate.

SI-VI RNA synthesis and gene silencing assay

a. Transfection of ApoB-siRNA and N[PFC]-ApoB-siRNA

HepG2 cells (Human hepatocellular carcinoma) were seeded in a 24-well plate at a density of $5x10^4$ cells. Cells were allowed to adhere overnight. Transfection was achieved by diluting 2 µM stocks of pre-annealed siRNA with OptiMEM (Life) and Oligofectamine (Life) to attain a final concentration of 20 nM of nucleic acid therapeutics. After an incubation period of 24 hours, we proceeded to isolate the RNA.

b. RNA Isolation

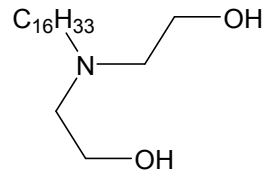
Total RNA was isolated from the 24-well plate by using an RNeasy mini kit (QIAGEN, Hilden, Germany) as described by the manufacturer. Genomic DNA was eliminated by RNase-free DNase I treatment during the isolation procedure. Reverse transcription was performed using the iScript Advanced cDNA Synthesis Kit (Bio-Rad) according the manufacturer's protocol. In a typical reaction 500 ng of RNA was mixed with 4 µl of reaction buffer and 1 µl of reverse transcriptase and the volume made up to 20 µl with nuclease-free water. Reverse transcription was performed at 42°C for 30 min and inactivated at 85°C for 5 minutes.

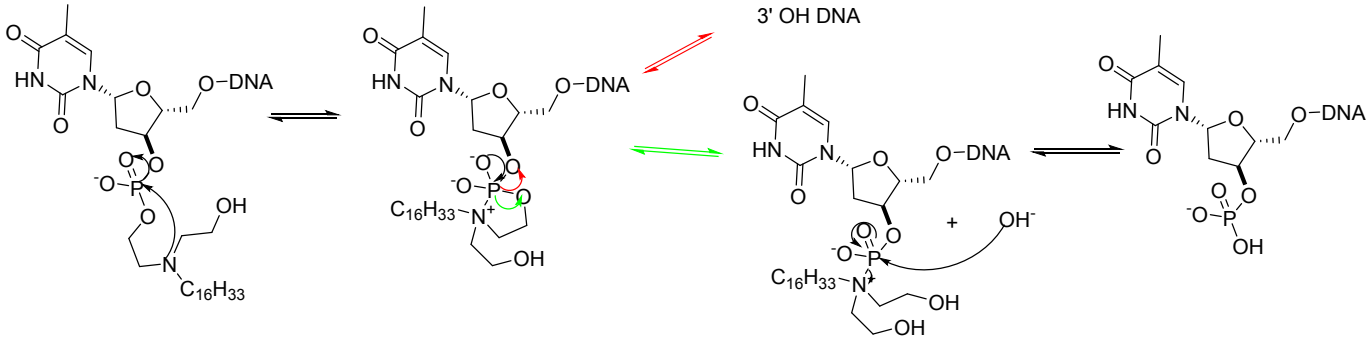
c. Quantitative real-time PCR.

Gene quantification was performed with a Step-One Plus (Life). Primers for ApoB used: Forward - 5'-TTTGCCCTCAACCTACCAAC-3' and Reverse - 5'-TGCGATCTTGTTGGCTACTG-3'. GAPDH was used as an endogenous control with the following primers: Forward -5'-GGAGCGAGATCCCTCCAAAAT-3' and Reverse - 5'-GGCTGTTGTCATACTTCTCATGG-3'. Each PCR was performed in a 20-µl reaction mixture containing 10 µl of SsoAdvanced universal SYBR Green supermix (Bio-Rad) and 250 nM of each primer. The thermal cycling conditions were as follows: 30 sec at 95°C, followed by 40 cycles of 15 s at 95°C, 30 s at 60°C. Data collection was performed during each extension phase. A negative control (distilled water), and RT-negative controls (total RNA sample) were included in each run. For each of the RNA extractions, measurements of gene expression were obtained in triplicate, and the mean of these values was used for further analysis using ΔΔCt method for relative quantification.

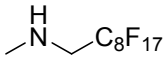
Note: pKa of N[PFC] molecule

Figure SF17: Potential degradation mechanism of the so called N[C16] modification.

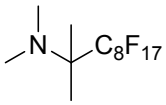




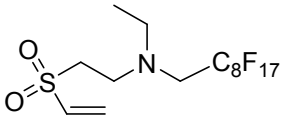
pKa of molecule **2** should be under 6.5 after comparison with SciFinder® theoretical values for very similar compounds:



pKa=6.44



pKa=5.35



pKa=3.79

[1] D. G. Cory, W. M. Ritchey, *Journal of Magnetic Resonance (1969)* **1988**, *80*, 128-132.

[2] G. Wider, L. Dreier, *J. Am. Chem. Soc.* **2006**, *128*, 2571-2576.