Supporting Information

Access to a stabilized *i*-motif DNA structure through four successive ligation reactions on a cyclopeptide scaffold

Alexandre Devaux, Laureen Bonnat, Thomas Lavergne and Eric Defrancq

Table of contents

1. Abbreviations

Alloc: Allyloxycarbonyl; **Boc**: tert-butoxycarbonyl; **CD**: Circular Dichroism; **CPG**: Controlled Pore Glass; **CuAAC**: Copper Catalyzed Alkyne-Azide Cycloaddition; **DCM**: dichloromethane; **Dde**: bis-N-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl]; **DIEA**: diisopropylamine; **DMF**: dimethylformamide; **DTT**: dithiothreitol; **EDTA**: Ethylenediaminetetraacetic acid; **Eei**: *Nε*-ethoxyethylidene; **ESI-MS**: Electrospray Ionization Mass Spectrometry; **ISAHC**: imidazole-1-sulfonylazide hydrochloride **MALDI-Tof-MS**: Matrix Assisted Laser Desorption Ionization - Time of Flight Mass Spectrometry; **OL**: oxime ligation; **PBS**: phosphate buffer saline; **RP-HPLC**: Reverse-Phase High Performance Liquid Chromatography; **Rt**: Retention time; **SEC**: Size Exclusion Chromatography; **SPOS**: solid-phase oligonucleotide syntheses; **SPPS**: solid-phase peptide synthesis; **TC**: thiol-iodoacetamide SN2 reaction; **TCEP:** tris(2-carboxyethyl)phosphine, **TFA**: trifluoroacetic acid; **THPTA**: Tris-(hydroxypropyltriazolylmethyl)amine; **TIS**: triisopropylsilane; **Tm**: melting temperature; **TRIS**: 2-amino-2-hydroxymethyl-propane-1,3-diol; **UPLC**: Ultra Performance Liquid Chromatography; **UV**: Ultraviolet.

2. General details

ESI mass spectra were performed on an Esquire 3000 spectrometer from Bruker or on an Acquity UPLC/MS system from Waters equipped with a SQ Detector 2 within a column heater set at 60°C. MALDI-Tof mass spectra were performed on an Autoflex Bruker using hydropiccolinic acid (HPA, 45 mg; ammonium citrate 4 mg in 500 µL H_2O/CH_3CN) as matrix. Peptides were analyzed in positive mode and oligonucleotides and conjugates in negative mode. The mass spectra display either the relative abundance of ion signals (total ion counts) against the *m/z* ratios or the total ion counts against the *m/z* ratios. All solvents and reagents used were of highest purity commercially available.

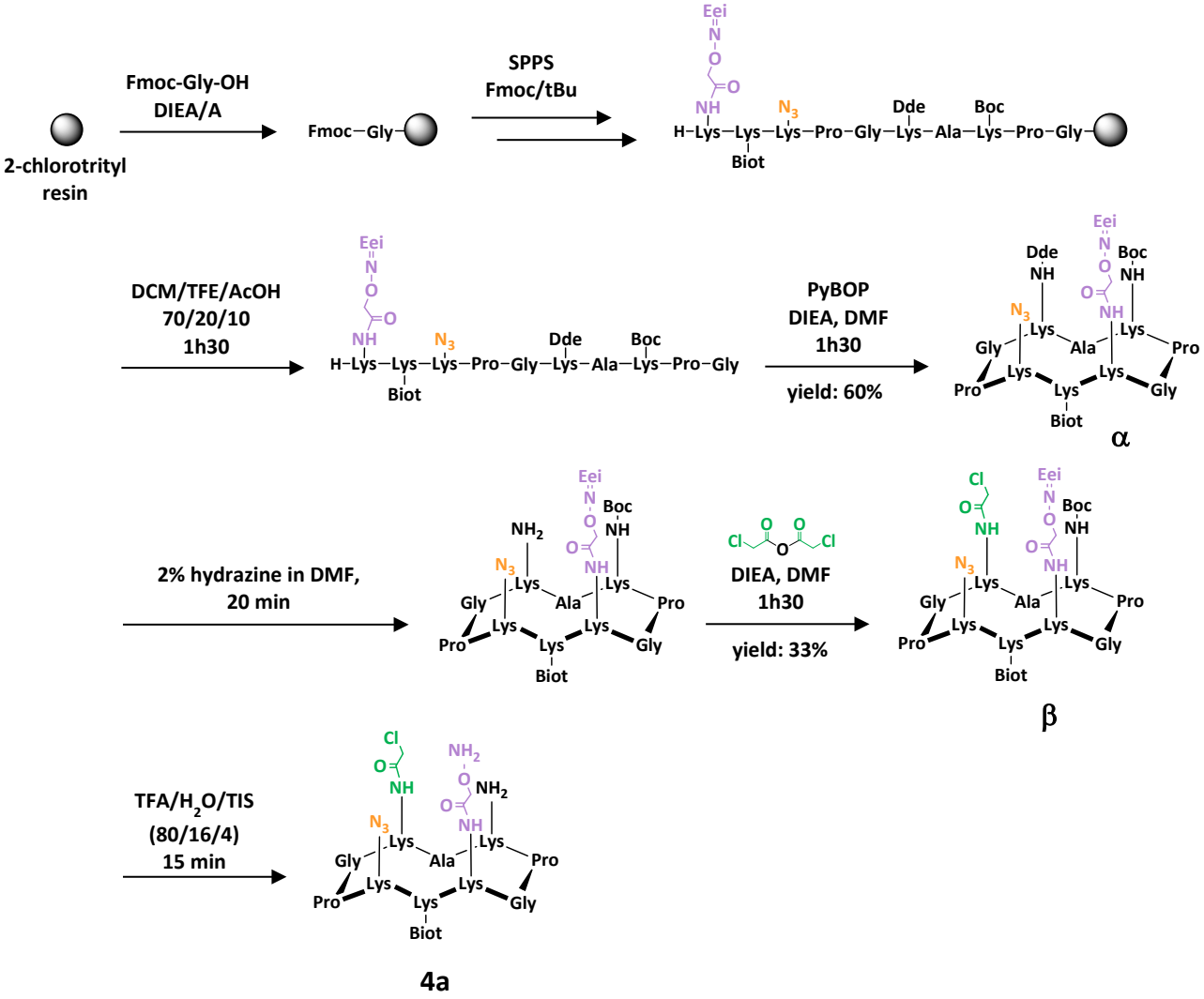
3. Peptide synthesis

General details

The synthesis was performed on a Syro II synthesizer using Fmoc/tBu strategy on a 2-Chlorotrityl resin.

The courses of the reactions were monitored by using UPLC system Waters, it includes reverse phase chromatography using Nucleosil C18 column (130 Å, 2.1 x 50 mm, 1.7 µm) and detection by UV at 214 nm and 250 nm. A 0.6 mL/min flow linear gradient from 95% solvent A (0.1% formic acid in water) and 5% solvent B (0.1% formic acid in acetonitrile) to 100% B for 3 minutes was applied. The purified products were analyzed by the same UPLC system and the chromatograms display the UV absorbance at 214 nm against time.

RP-HPLC purifications were performed on a Gilson system with Nucleosil C18 column (100 Å, 250 x 21 mm, 7 µm) with UV monitoring at 214 nm and 250 nm. A 22 mL/min flow linear gradient was applied from 95% solvent A (0.1% formic acid in water) and 5% solvent B (0.1% formic acid in acetonitrile) to 100% B for 15 minutes.



Scheme S1. Synthesis of peptide **4a**.

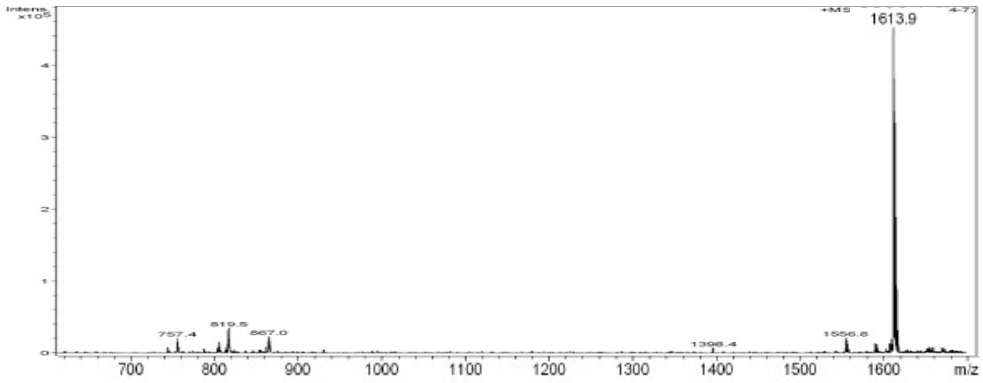


Figure S1. ESI mass spectrum of intermediate peptide β: ESI-MS (+): *m/z* calcd for $C_{70}H_{116}N_{20}O_{18}SCl$: 1591.8, *m/z* found: 1613.9 $[M+Na]^+$.

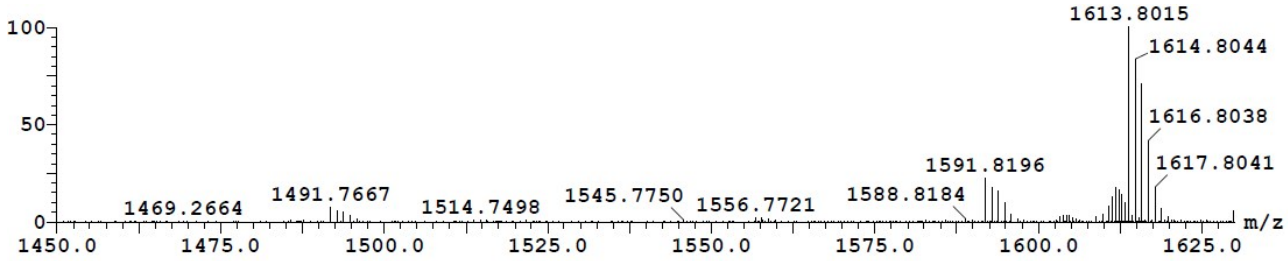


Figure S2. HR-ESI mass spectrum of intermediate peptide β: ESI-MS (+): *m/z* calcd for $C_{70}H_{116}N_{20}O_{18}SCl$: 1591.8186, *m/z* found: 1591.8196.

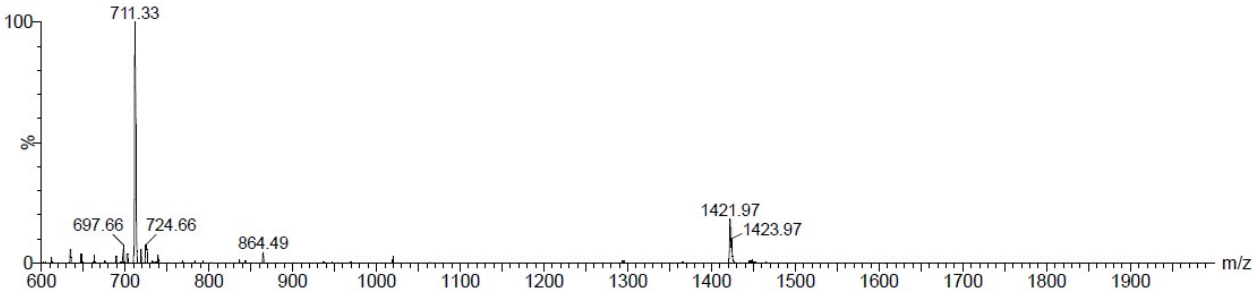
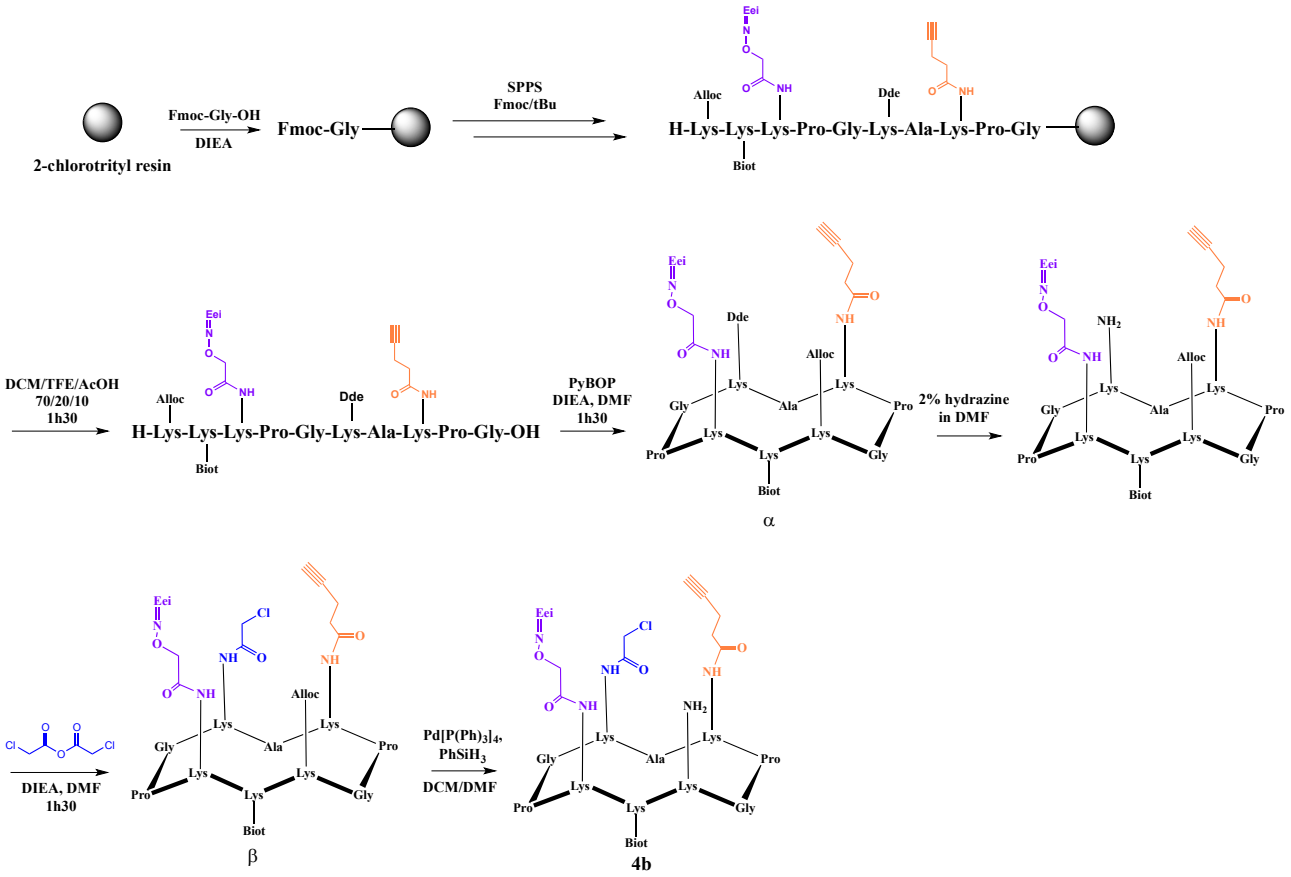


Figure S3. ESI mass spectrum of peptide **4a** ESI-MS (+): *m/z* calcd for $C_{61}H_{102}N_{20}O_{15}SCl$: 1421.7, *m/z* found: 1421.9 $[M+H]^+$.



Scheme S2. Synthesis of peptide **4b**.

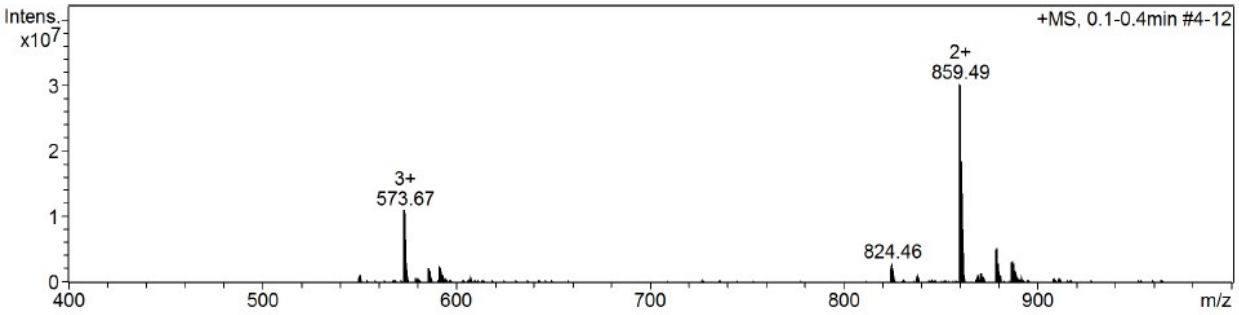


Figure S4. ESI mass spectrum of intermediate **α :** ESI-MS (+): *m/z* calcd for $C_{82}H_{128}N_{18}O_{20}S$: 1716.92, *m/z* found: 1716.97 $[M+H]^+$.

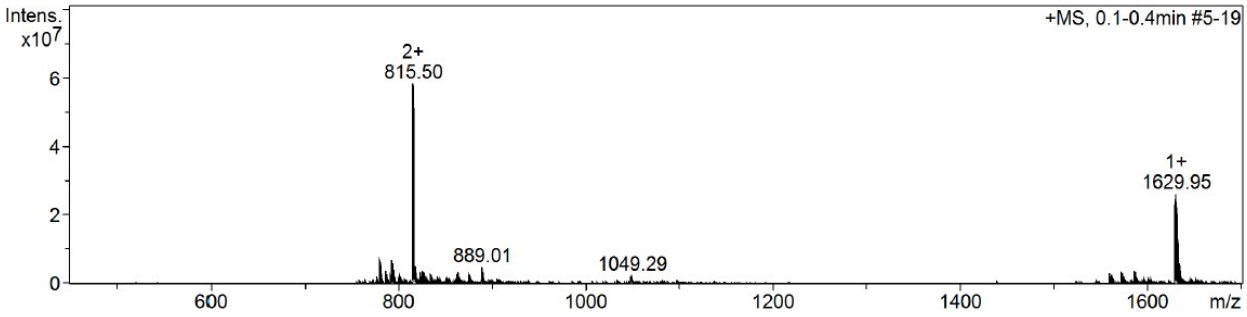


Figure S5. ESI mass spectrum of intermediate **β:** ESI-MS (+): *m/z* calcd for $C_{74}H_{118}N_{18}O_{19}SCl$: 1629.34, *m/z* found: 1629.95 $[M+H]^+$.

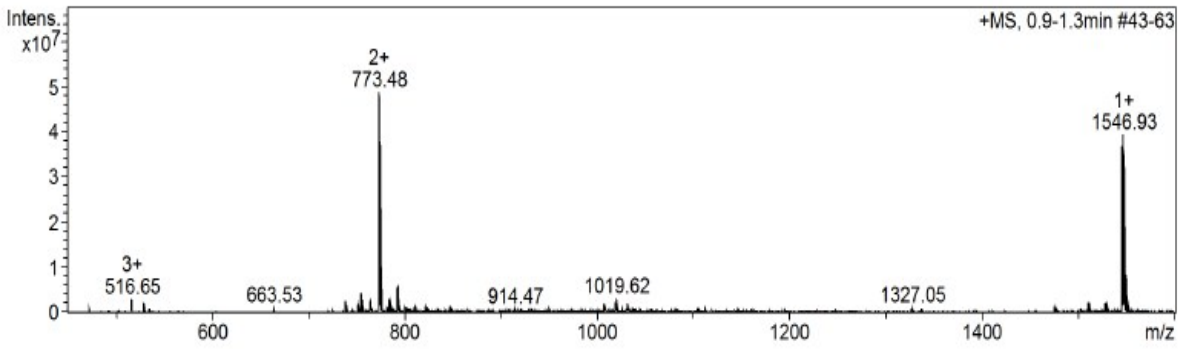


Figure S6. ESI mass spectrum of peptide **4b:** ESI-MS (+): *m/z* calcd for $C_{70}H_{115}N_{18}O_{17}SCl$: 1546.27, *m/z* found: 1546.93 $[M+H]^+$.

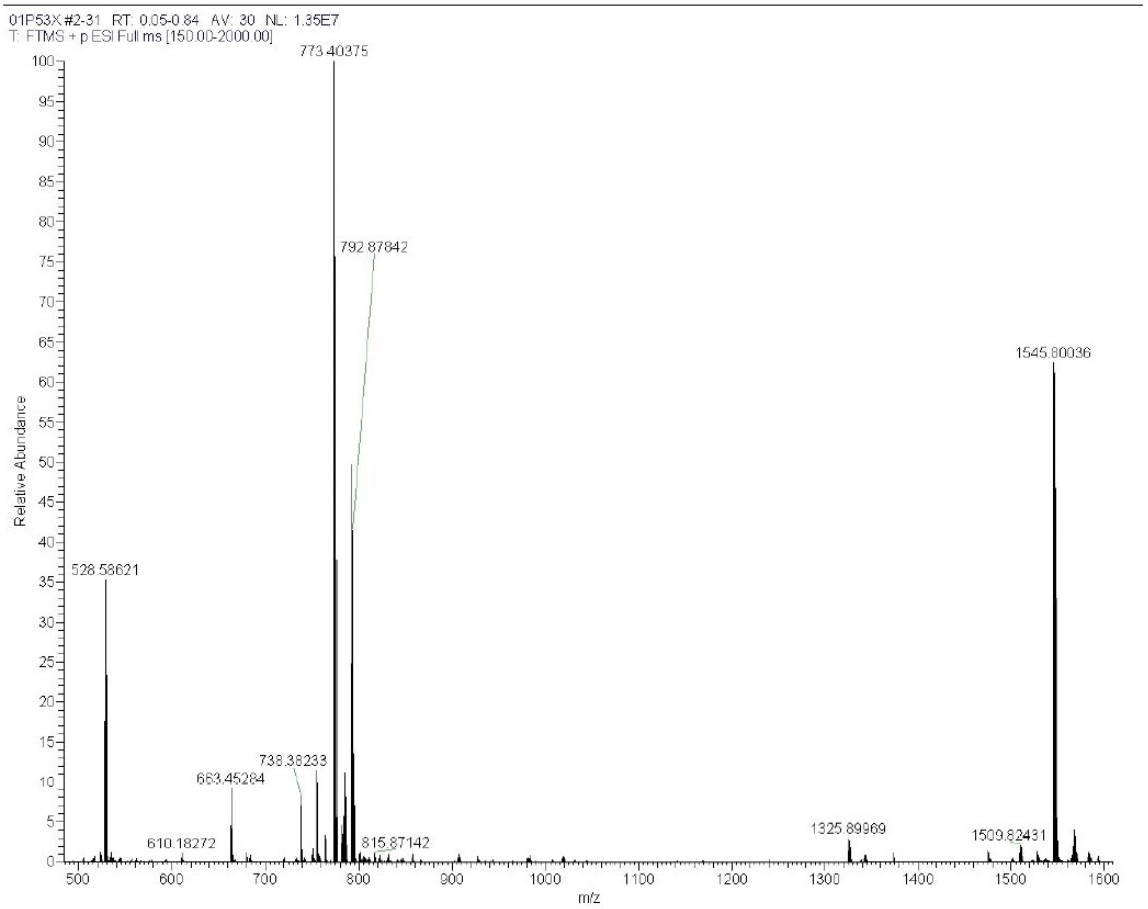


Figure S7. ESI mass spectrum of peptide **4b:** ESI-HRMS (+): *m/z* calcd for $C_{70}H_{114}N_{18}O_{17}SCl$: 1545.8018, *m/z* found: 1545.8003 $[M]^+$.

4. Oligonucleotide synthesis

General details

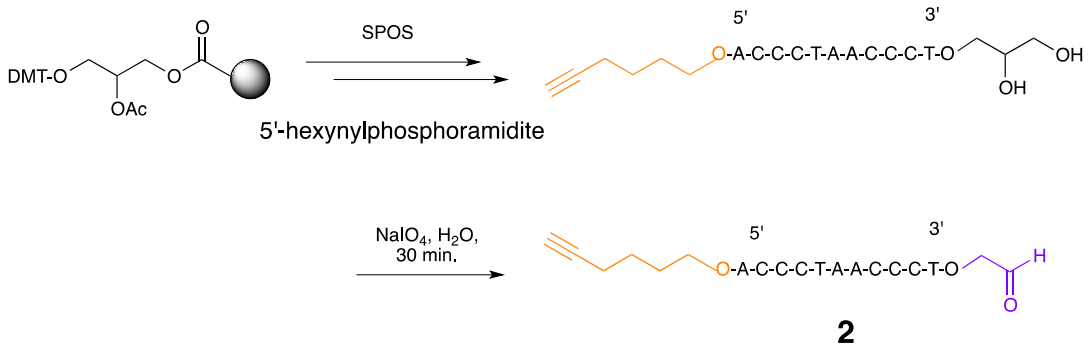
Oligonucleotides were prepared using β-cyanoethylphosphoramidite chemistry on a 3400 DNA synthesizer at 1 µmol scale.

RP-HPLC analyses were performed on a Waters HPLC system using C18 Nucleosil column (Macherey-Nagel, 100 Å, 250 x 4.6 mm, 5 µm) with UV-monitoring at 260 nm and 280 nm. A 1 mL/min flow linear gradient was applied. Solvent A (50 mM triethylammonium acetate buffer with 5% acetonitrile) and solvent B (acetonitrile with 5% water) were used. A stepwise gradient of 0-30% B in 20 min then from 30 to 100% B in 10 min was applied for the gradient. The RP-HPLC purifications of oligonucleotides were performed on a Gilson system with Nucleosil C-18 colunm (Macherey-Nagel, 100 Å, 250 x 10 mm, 7 µm) with UV-monitoring at 260 nm and 280 nm using 4 mL/min flow linear gradient. A stepwise gradient of 0-30% B in 20 min then from 30 to 100% B in 10 min was applied.

Desalting of oligonucleotides was performed by SEC on NAP 25 cartridge using manufacturer's protocol.

Quantification of oligonucleotides was performed at 260 nm using Nanodrop apparatus (molar extinction ε_{260nm} was estimated according to the nearest neighbor model).

a) Synthesis of oligonucleotide 2



Scheme S3. Synthesis of oligonucleotide **2**.

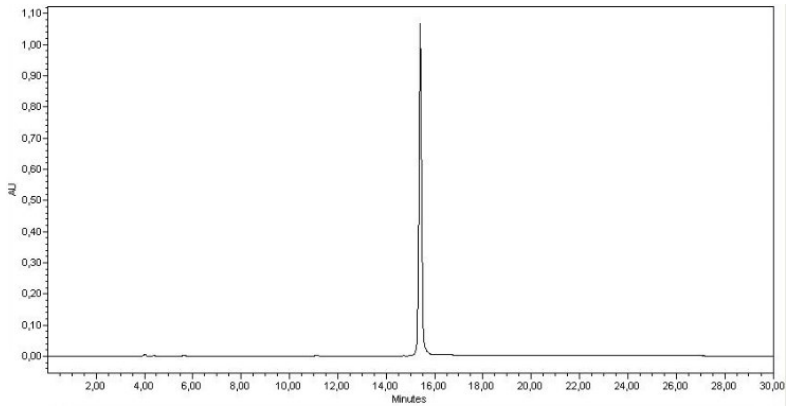


Figure S8. RP-HPLC chromatogram of pure oligonucleotide **2**.

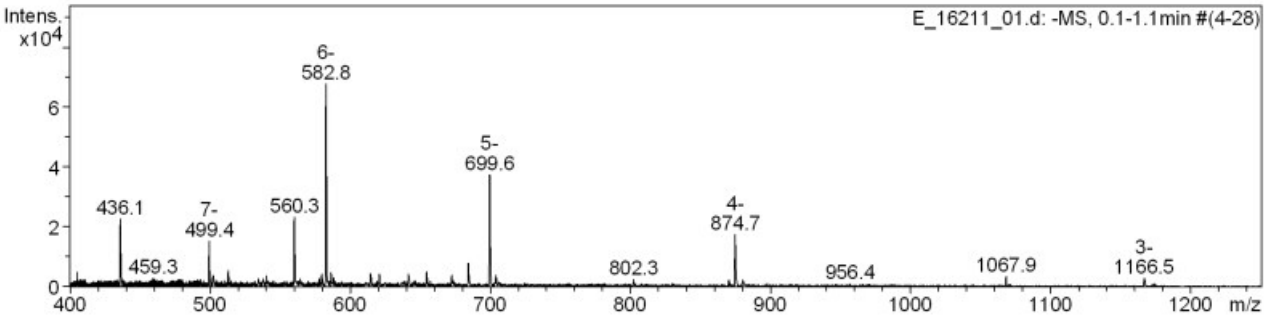
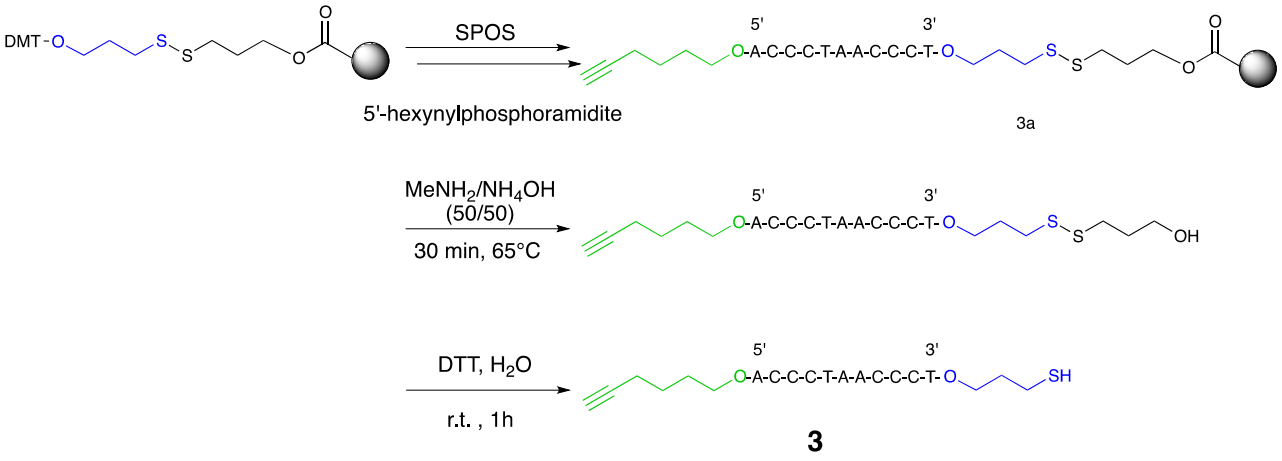


Figure S9. ESI mass spectrum of oligonucleotide **2.** ESI-MS (-): *m/z* calcd for $C_{112}H_{148}N_{37}O_{70}P_{12}$: 3502.60, *m/z* found: 3502.68 $[M-H]^-$.

b) Synthesis of oligonucleotide 3



Scheme S4. Synthesis of oligonucleotide **3**.

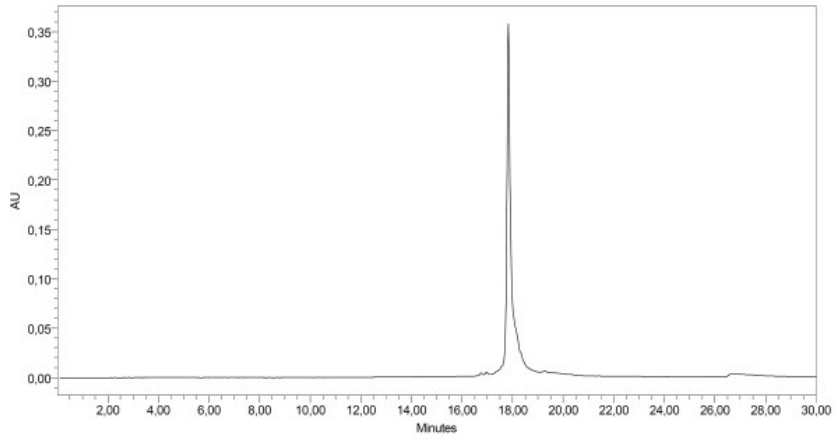


Figure S10. RP-HPLC chromatogram of pure oligonucleotide **3**.

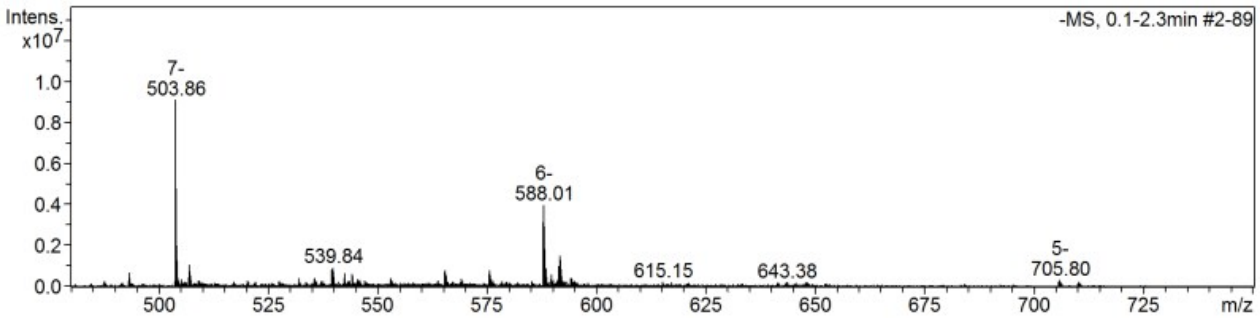
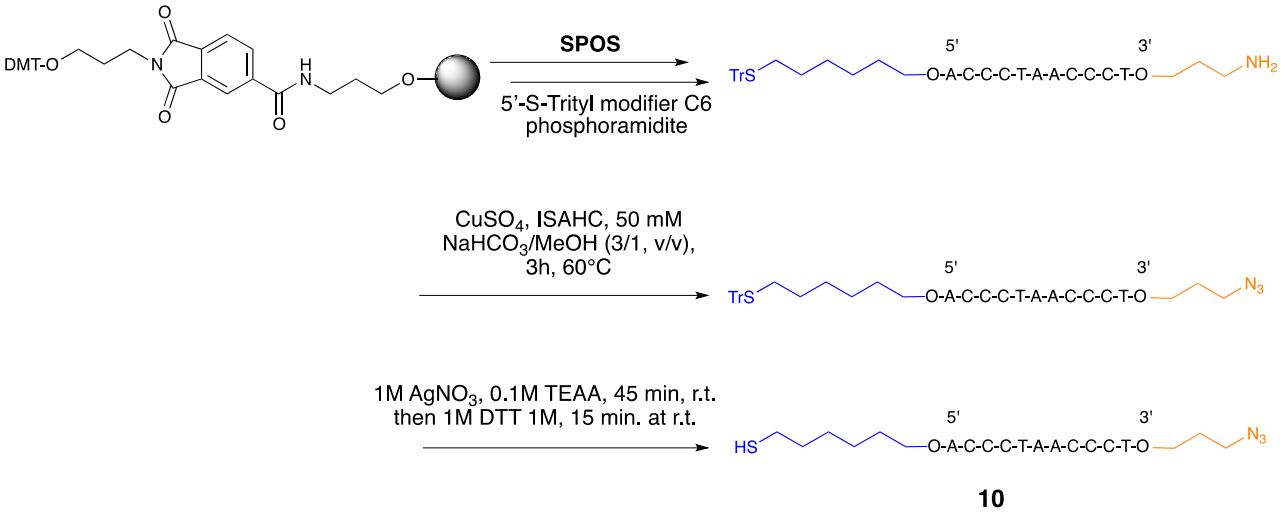


Figure S11. ESI mass spectrum of oligonucleotide **3.** ESI-MS (-): *m/z* calcd for $C_{113}H_{151}N_{37}O_{69}P_{12}S$: 3534.4, *m/z* found: 3534.9 $[M-H]^-$.

c) Synthesis of oligonucleotide 10



Scheme S5. Synthesis of oligonucleotide **10**.

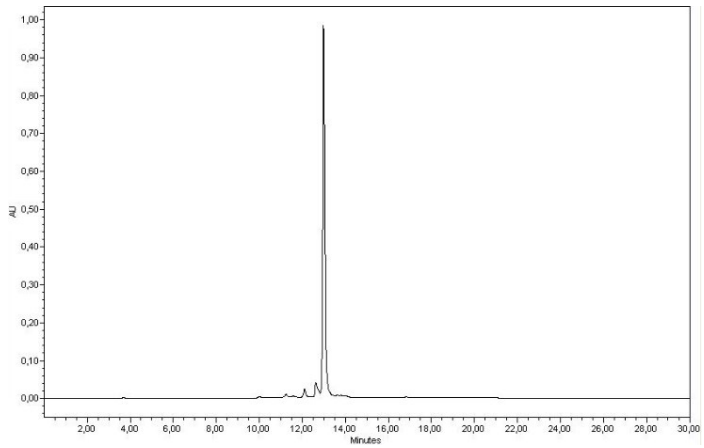


Figure S12. RP-HPLC chromatogram of oligonucleotide **10**.

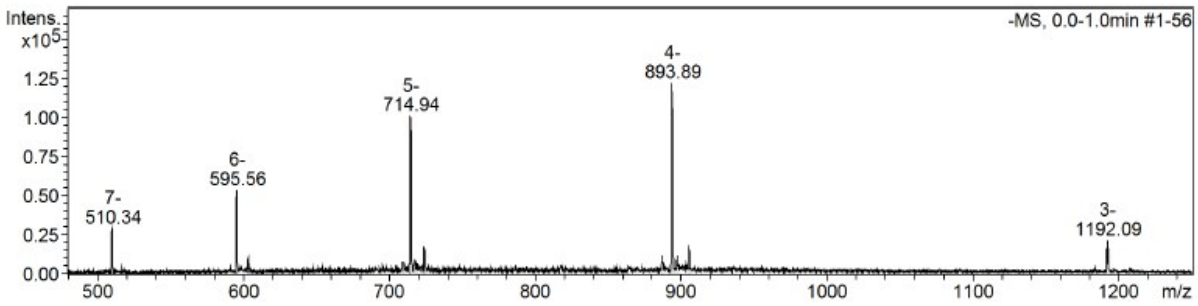


Figure S13. ESI mass spectrum of oligonucleotide **10.** ESI-MS (-): *m/z* calcd for $C_{113}H_{155}N_{40}O_{69}P_{12}S$: 3579.63, *m/z* found: 3579.58 $[M-H]^-$.

5. Conjugates synthesis

General details

RP-HPLC analyses were performed on a Waters HPLC system using C18 Nucleosil column (Macherey-Nagel, 100 Å, 250 x 4.6 mm, 5 µm) with UV-monitoring at 260 nm and 280 nm. A 1 mL/min flow linear gradient was applied. Solvent A (50 mM triethylammonium acetate buffer with 5% acetonitrile) and solvent B (acetonitrile with 5% water) were used. Two gradients have been used: gradient I (0-30% B in 30 minutes) and gradient II (0-50% B in 30 minutes).

The course of reactions were monitored by using UPLC-MS system Waters, it includes reverse phase chromatography using Nucleosil C18 column (130 Å, 2.1 x 50 mm, 1.7 µm) equipped with a column heater set at 60°C and detection by UV at 260 nm and 280 nm and by electron spray ionization mass spectrometry. A 0.3 mL/min flow linear gradient with solvent C (triethylamine (15mM) and hexafluoro-2-propanol (50mM) in water) and solvent D (triethylamine (15mM) and hexafluoro-2-propanol (50mM) in methanol) was used.

The RP-HPLC purifications of conjugates were performed on a Gilson system with Nucleosil C-18 column (Macherey-Nagel, 100 Å, 250 x 10 mm, 7 µm) with UV-monitoring at 260 nm and 280 nm using 4 mL/min flow linear gradient. Solvent A (50 mM triethylammonium acetate buffer with 5% acetonitrile) and solvent B (acetonitrile with 5% water) were used.

a) Synthesis of conjugate 5

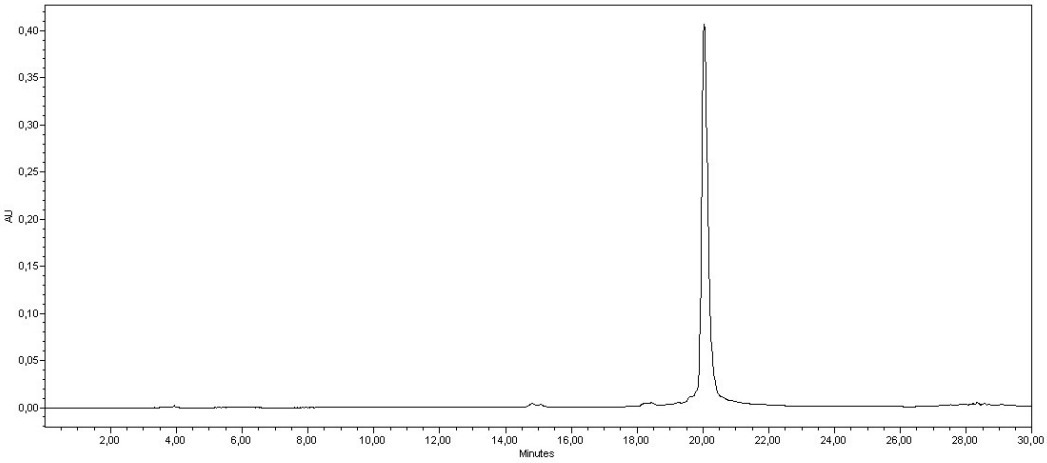


Figure S14. RP-HPLC chromatogram of conjugate **5** (*oxime ligation*).

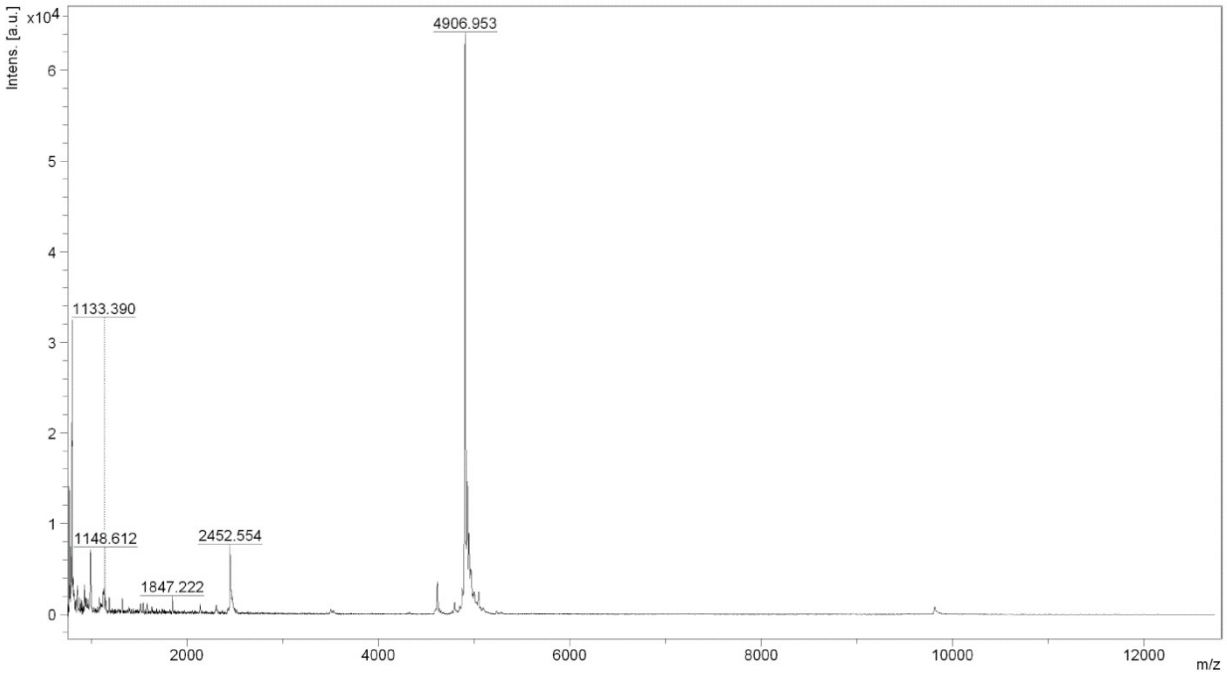


Figure S15. MALDI-Tof-MS mass spectrum of oligonucleotide **5.** MALDI-MS (-): *m/z* calcd for $C_{173}H_{245}N_{57}O_{84}P_{12}SCl$: 4906.3, *m/z* found: 4906.9 $[M-H]^-$.

b) Synthesis of conjugate 6

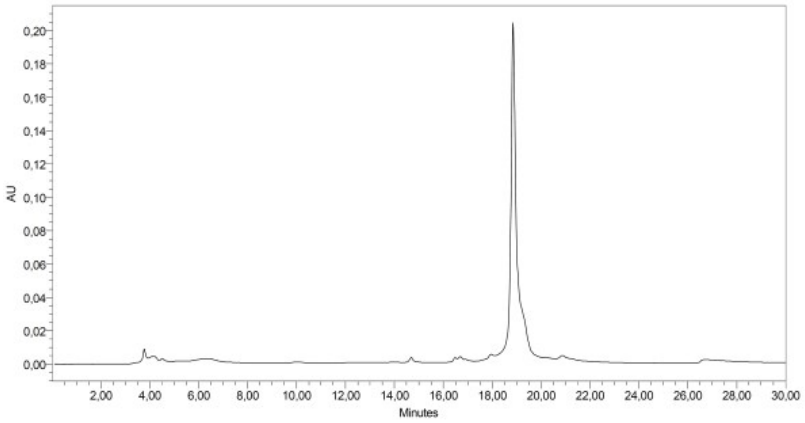


Figure S16. RP-HPLC chromatogram of conjugate **6** (*CuAAC reaction*).

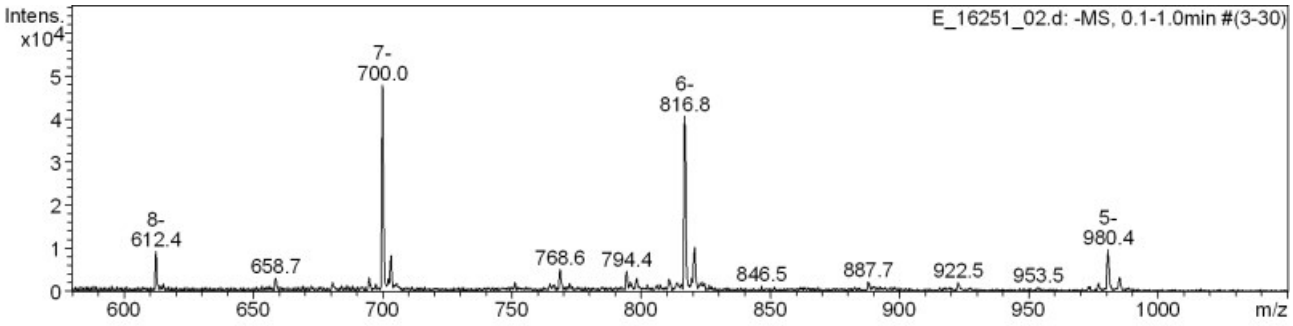


Figure S17. ESI mass spectrum of oligonucleotide **6.** ESI-MS (-): *m/z* calcd for $C_{173}H_{245}N_{57}O_{84}P_{12}SCl$: 4906.3, *m/z* found: 4907.2 $[M-H]^-$.

c) Synthesis of conjugate 7

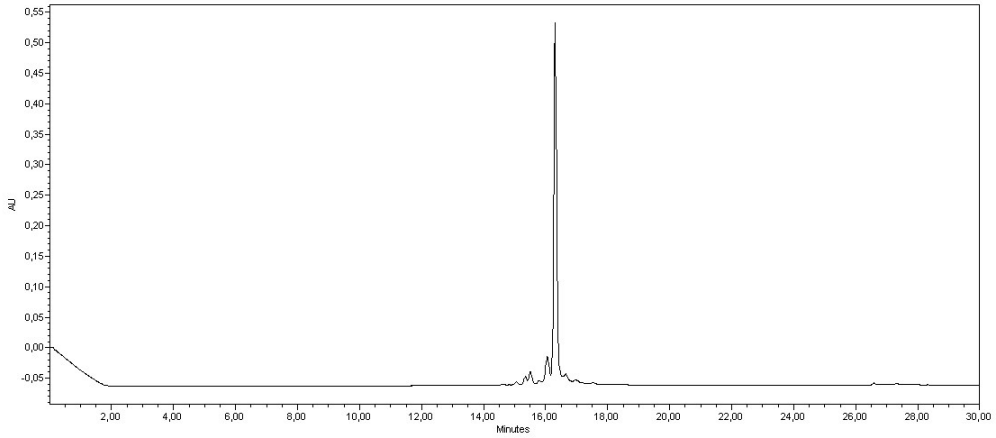


Figure S18. RP-HPLC chromatogram of conjugate **7** (*TC reaction*).

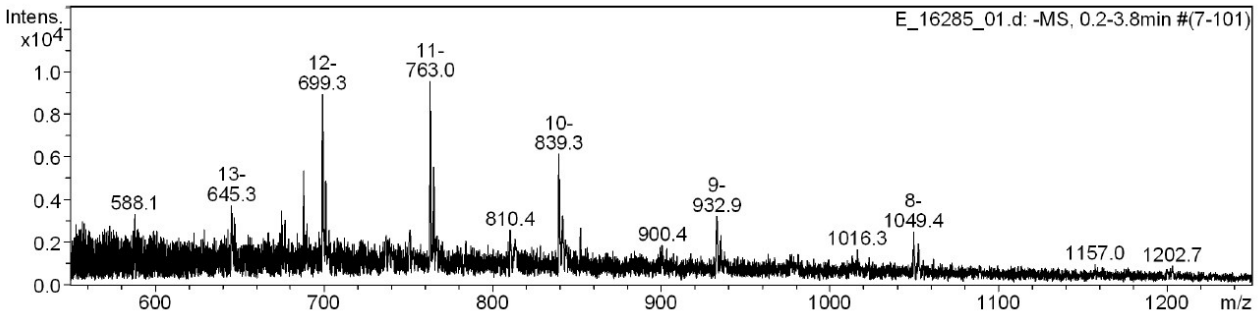


Figure S19. ESI mass spectrum of oligonucleotide **7.** ESI-MS (-): *m/z* calcd for $C_{286}H_{395}N_{94}O_{153}P_{24}S_2$: 8405.2, *m/z* found: 8403.5 $[M-H]^-$.

d) Synthesis of conjugate 8

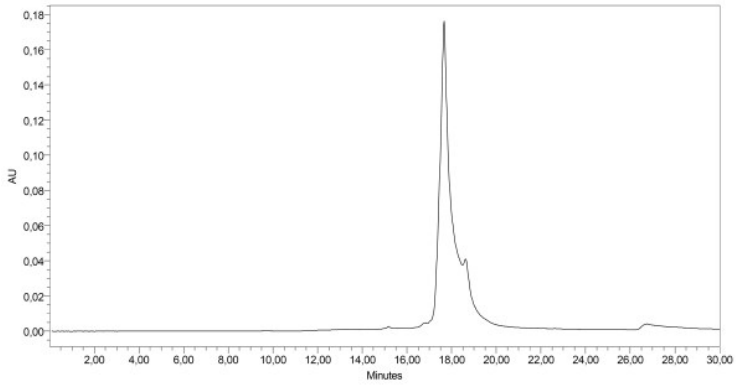


Figure S20. RP-HPLC chromatogram of conjugate **8**.

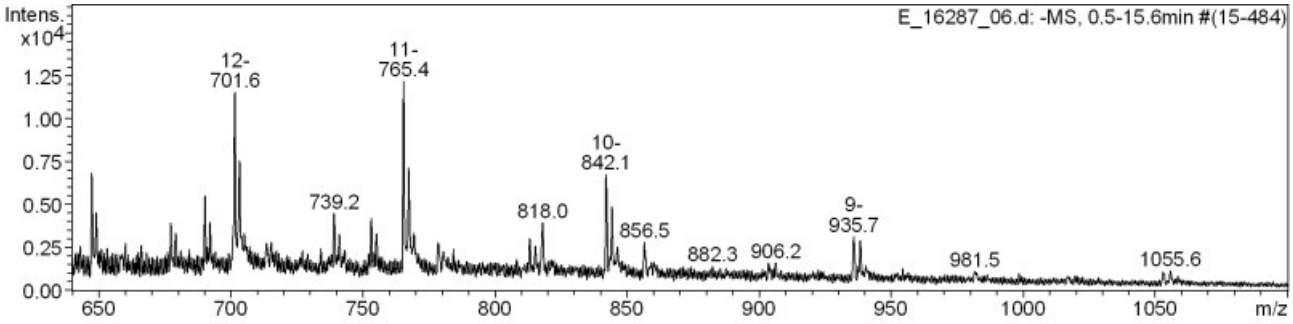


Figure S21. ESI mass spectrum of oligonucleotide **8.** ESI-MS (-): *m/z* calcd for $C_{286}H_{393}N_{96}O_{153}P_{24}S_2$: 8430.8, *m/z* found: 8433.8 $[M-H]^-$.

e) Synthesis of conjugate 9

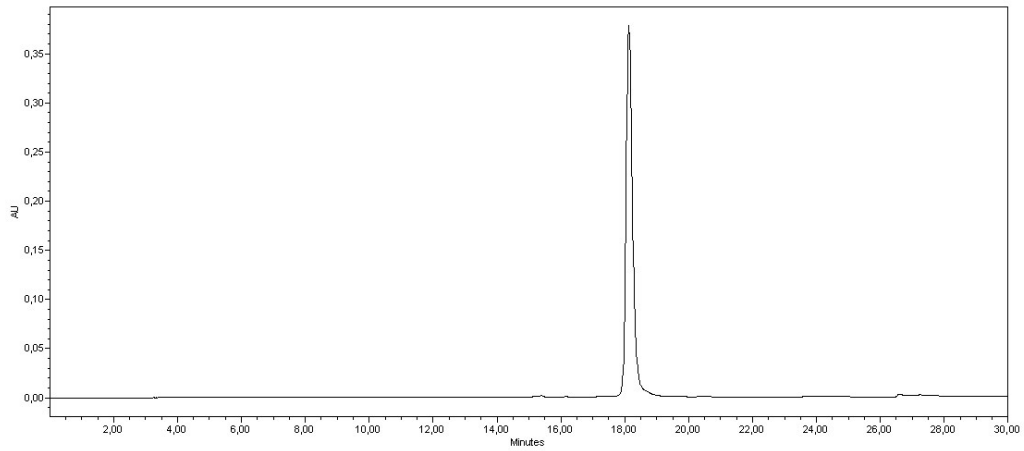


Figure S22. RP-HPLC chromatogram of conjugate **9** (*diazo transfer reaction*).

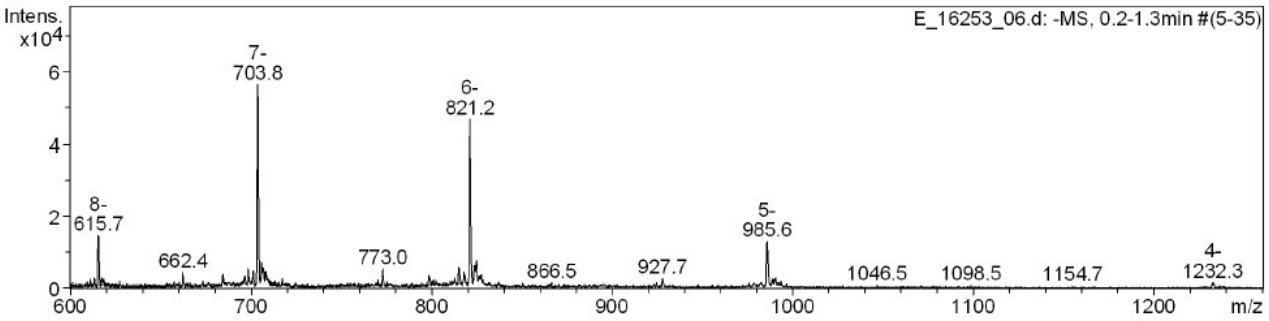


Figure S23. ESI mass spectrum of oligonucleotide **9.** ESI-MS (-): *m/z* calcd for $C_{173}H_{243}N_{59}O_{84}P_{12}SCl$: 4932.3, *m/z* found: 4934.4 $[M-H]^-$.

f) Synthesis of conjugate 11

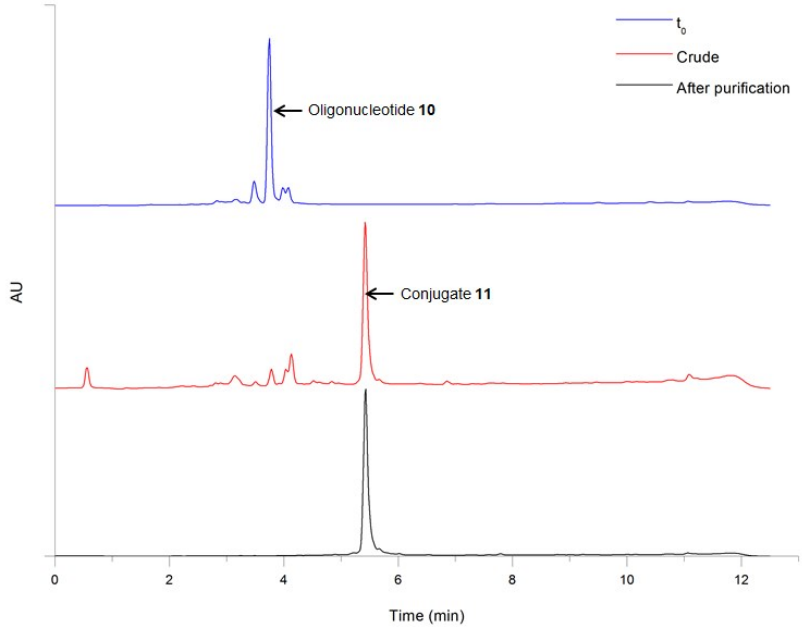


Figure S24. Reaction monitoring (UPLC-MS, negative mode) of the formation of conjugate **11** from oligonucleotide **10** (*TC reaction*).

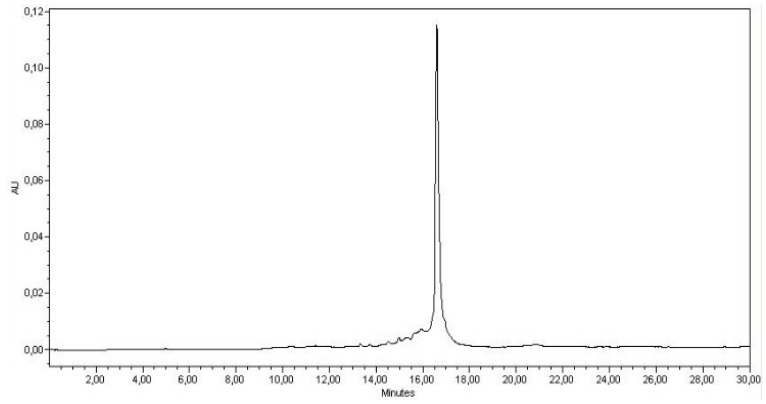


Figure S25. RP-HPLC chromatogram of pure conjugate **11**.

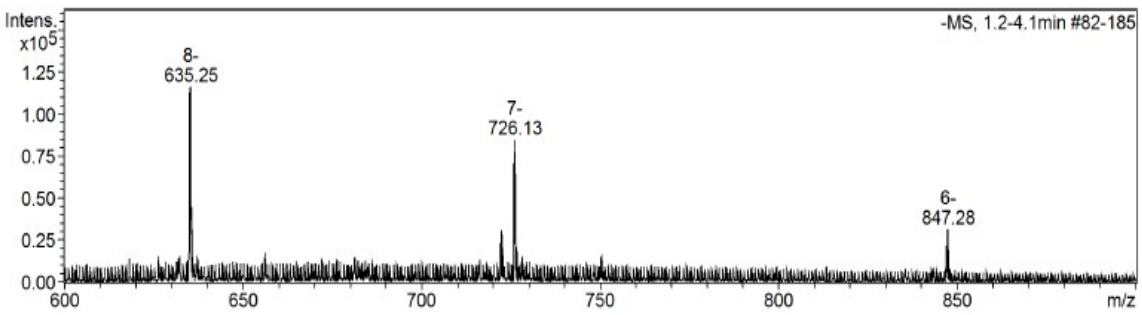


Figure S26. ESI mass spectrum of conjugate **11:** ESI-MS (-): *m/z* calcd $C_{183}H_{268}N_{58}O_{86}P_{12}S_2$: 5089.46, *m/z* found: 5089.68 $[M-H]^-$.

g) Synthesis of conjugate 12

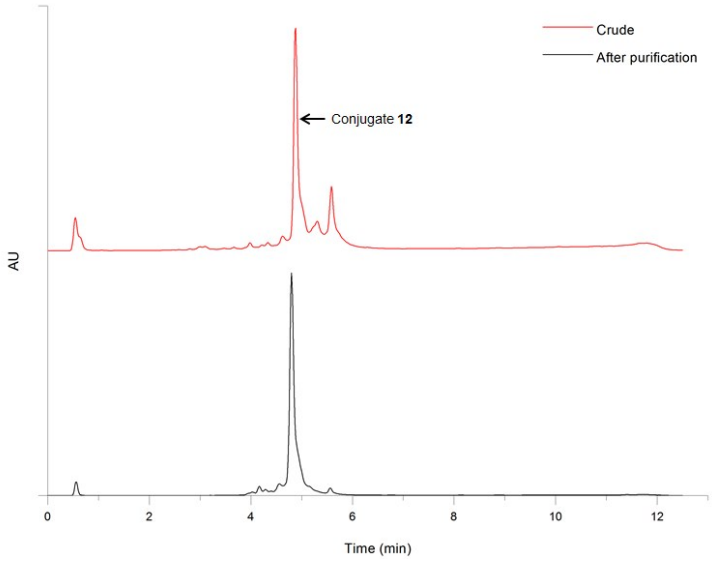


Figure S27. Reaction monitoring (UPLC-MS, negative mode) of the formation of conjugate **12** from **11** (*CuAAC reaction*).

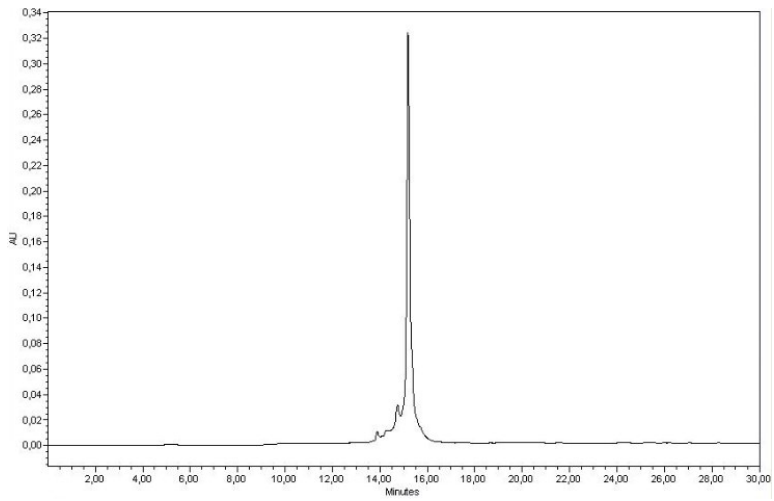


Figure S28. RP-HPLC chromatogram of pure conjugate **12**.

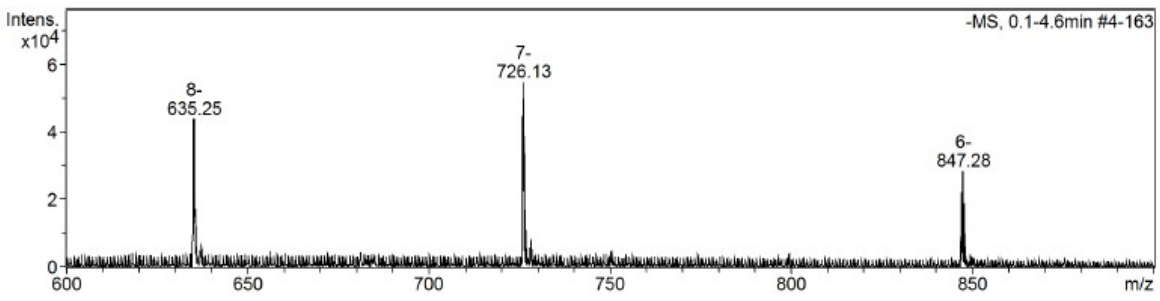


Figure S29. ESI mass spectrum of conjugate **12:** ESI-MS (-): *m/z* calcd $C_{183}H_{268}N_{58}O_{86}P_{12}S_2$: 5089.46, *m/z* found: 5089.62 $[M\text{-}H]^-$.

h) Synthesis of conjugate 13

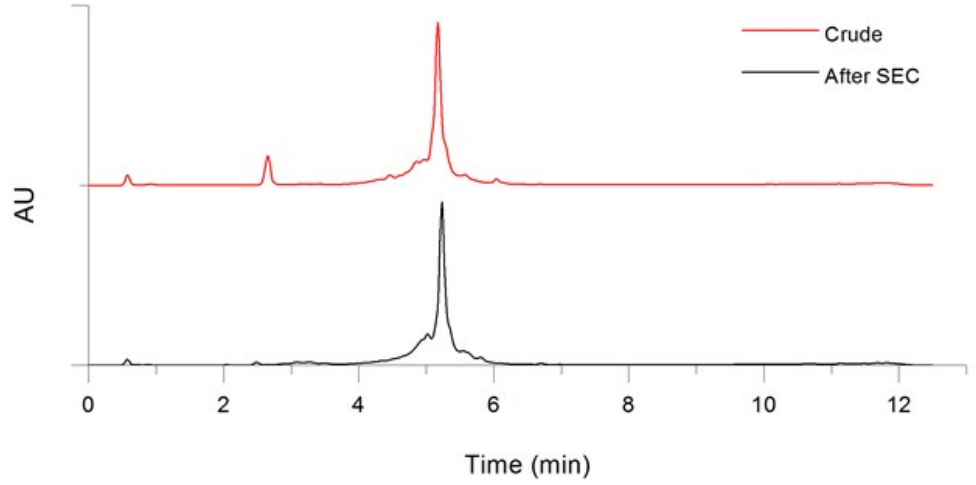


Figure S30. Reaction monitoring (UPLC-MS, negative mode) of the formation of conjugate **13** from **12** (*diazo transfer reaction*).

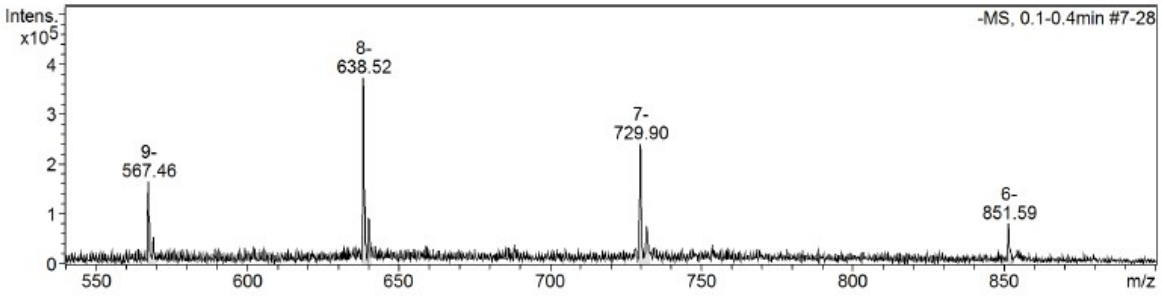


Figure S31. ESI mass spectrum of conjugate **13:** ESI-MS (-): *m/z* calcd $C_{183}H_{267}N_{60}O_{86}P_{12}S_2$: 5116.45, *m/z* found: 5116.21 $[M-H]^-$.

i) Synthesis of conjugate 14

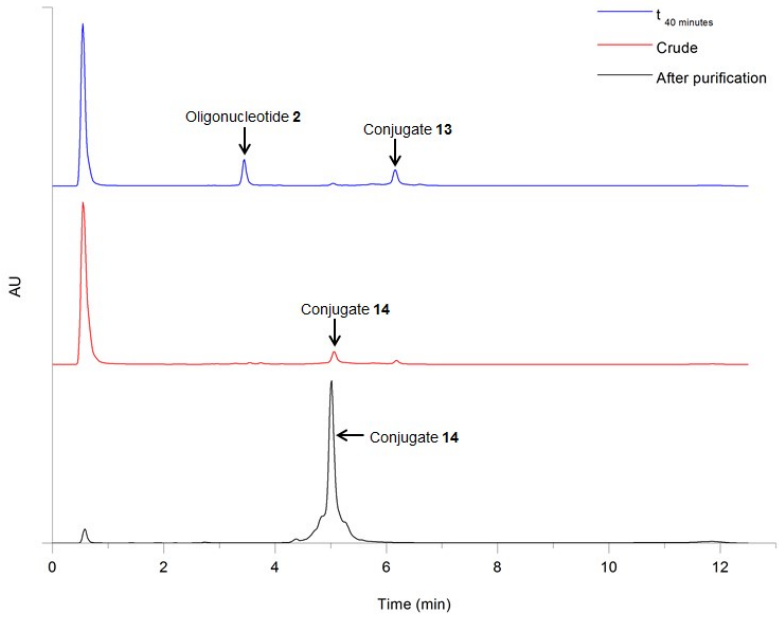


Figure S32. Reaction monitoring (UPLC-MS, negative mode) of the formation of conjugate **14** from **13** (*CuAAC reaction*).

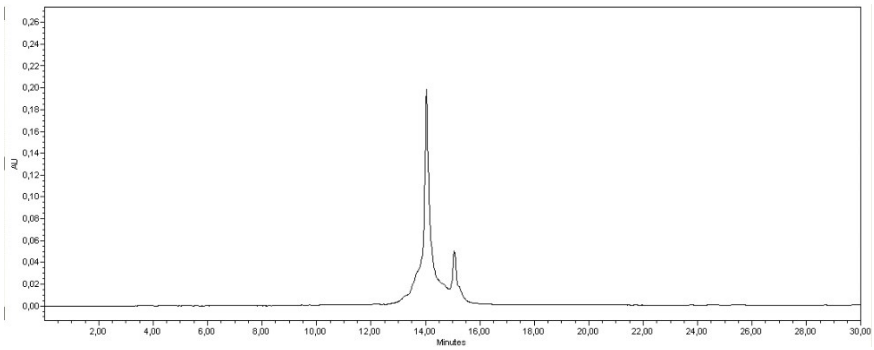


Figure S33. RP-HPLC chromatogram of pure conjugate **14**.

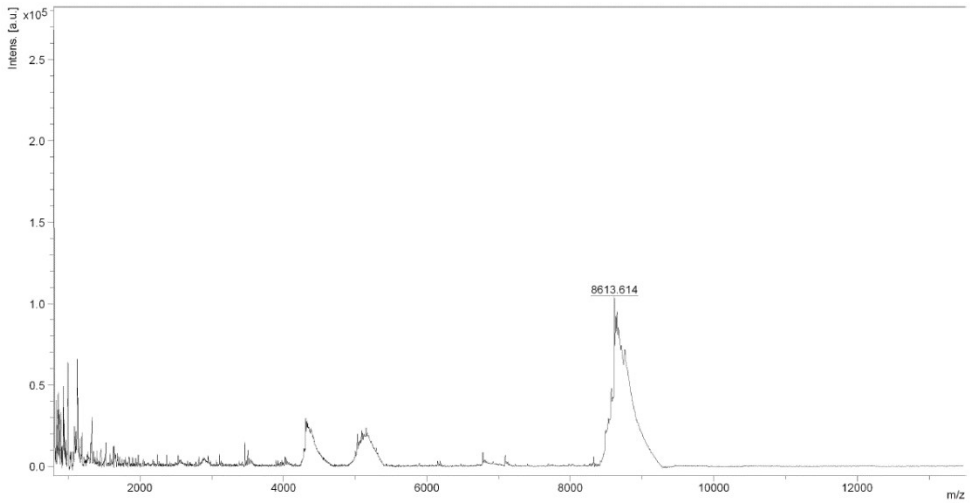


Figure S34. MALDI-Tof mass spectrum of conjugate **14.** MALDI-Tof-MS (-): *m/z* calcd $C_{295}H_{413}N_{97}O_{156}P_{24}S_2$: 8617.05, *m/z* found: 8613.61 $[M-H]^-$.

j) Synthesis of conjugate 1b

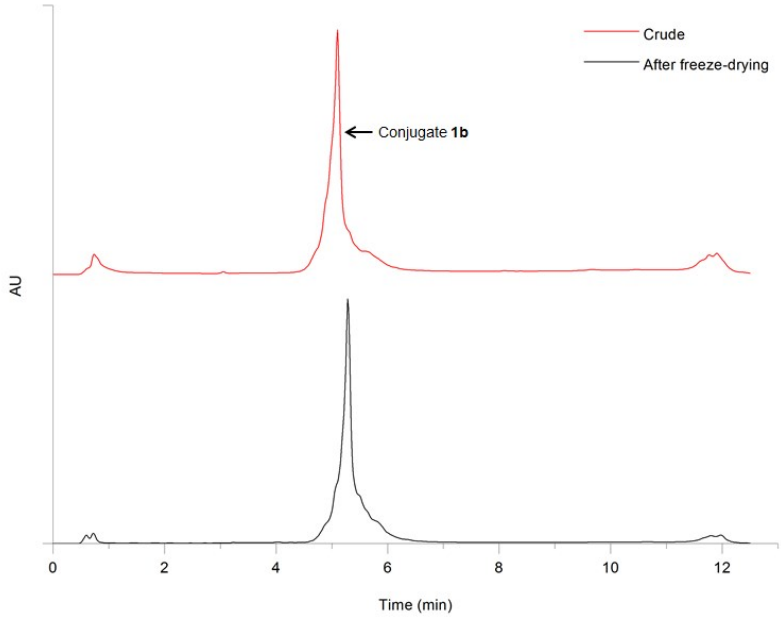


Figure S35. Reaction monitoring (UPLC-MS, negative mode, 60°C) of the formation of conjugate **1b** from **14** (*oxime ligation*). The top chromatogram is the crude reaction. The bottom chromatogram is the pure conjugate **1b** after Size Exclusion filtration (NAP 25 cartridge) and freeze-drying to remove volatile TFA reagent and the water byproduct and solvent. The UPLC was performed at 60°C in conditions where **1b** do not form the i-motif topology.

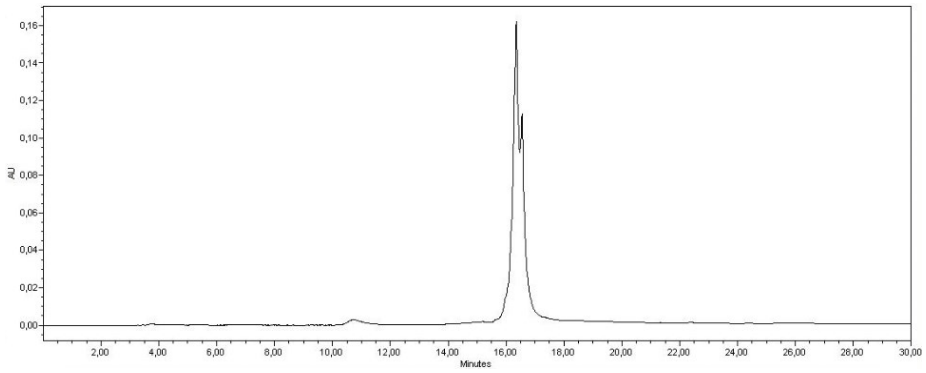


Figure S36. RP-HPLC chromatogram of conjugate **1b** at room temperature. This chromatogram shows two peaks which corresponds to a mixture of conformers, one peak corresponding to the folded i-motif structure and the other to one or more unfolded forms of the conjugate **1b**.

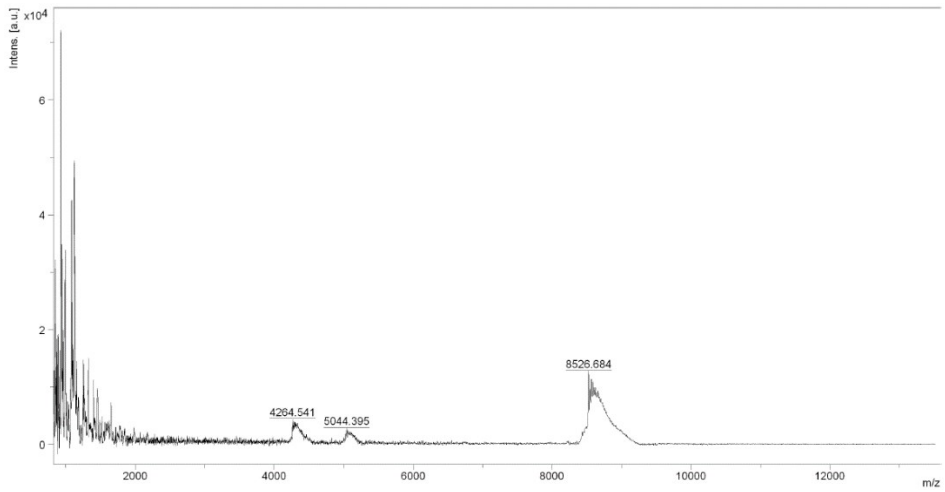


Figure S37. MALDI-Tof mass spectrum of conjugate **1b:** MALDI-Tof-MS (-): *m/z* calcd $C_{291}H_{405}N_{97}O_{154}P_{24}S_2$: 8528.99, *m/z* found: 8526.68 $[M-H]^-$.

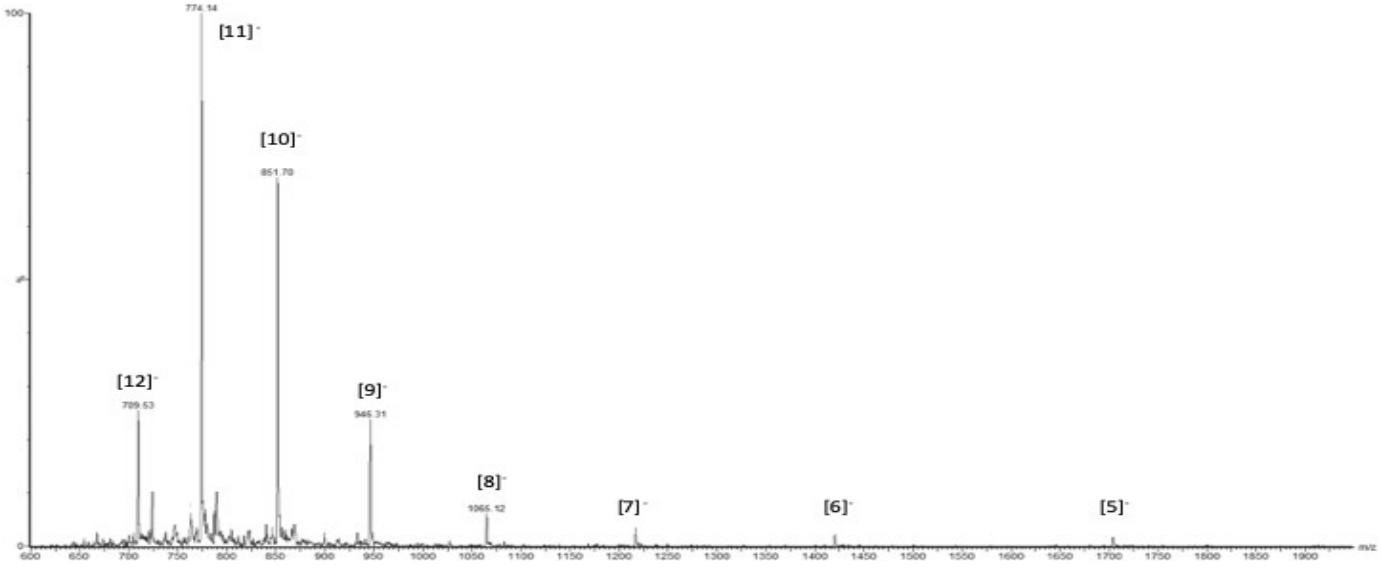


Figure S38. mass spectrum of conjugate **1b from** LC-MS analysis**:** *m/z* calcd $C_{291}H_{405}N_{97}O_{154}P_{24}S_2$: 8528.99, *m/z* found: 8528.00 $[M-H]^-$.

6. Circular Dichroism studies

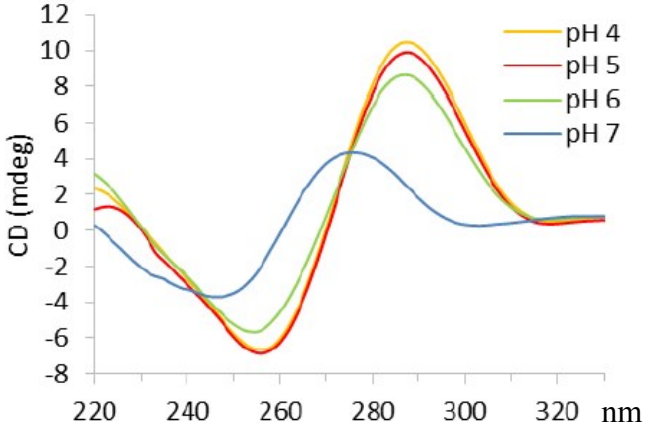


Figure S39. CD spectra of conjugate **8** at different pH (PBS buffer) at room temperature.

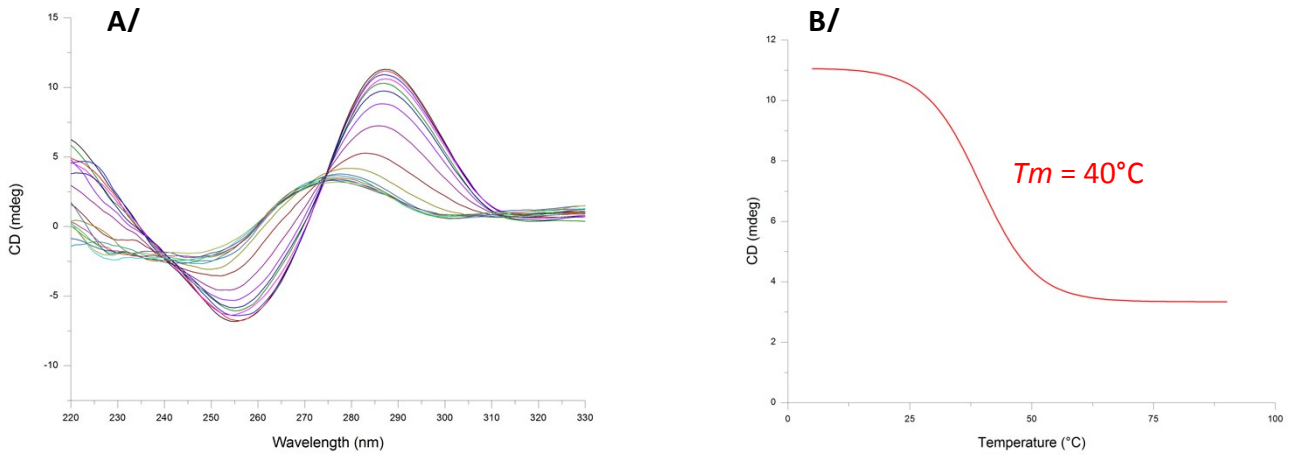


Figure S40. A/ Raw CD traces for conjugate **1b** (2 µM) in 10 mM PBS buffer at pH 6, Temperature varied from 5°C to 90°C, **B/** CD signal intensities at λ_{max} *versus* temperature.

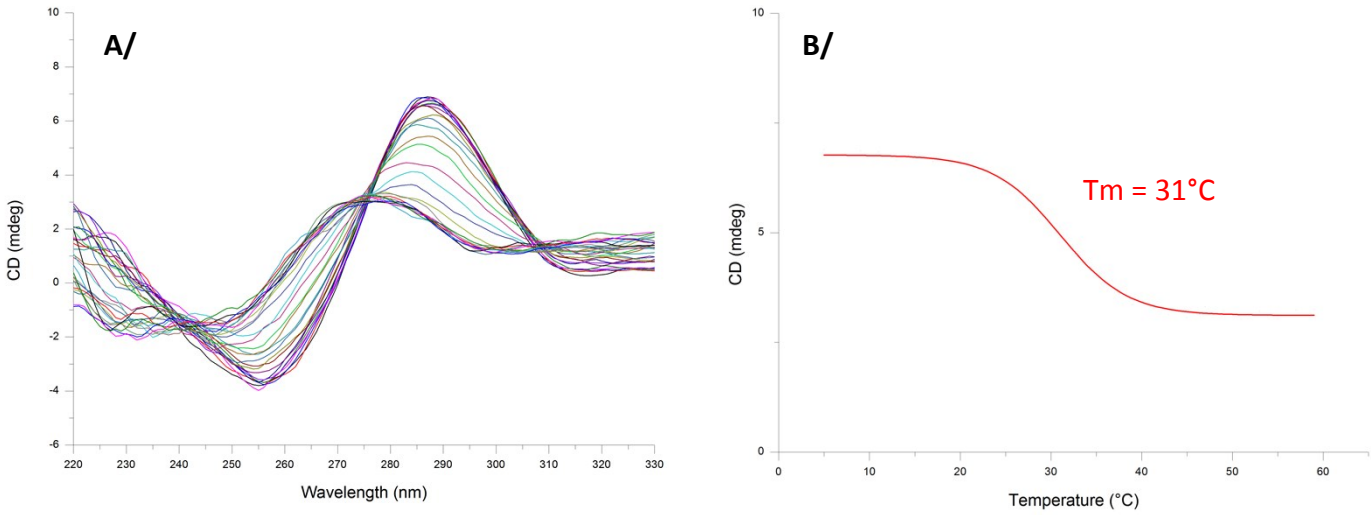


Figure S41. A/ Conjugate **1b** (2 µM) in 10 mM PBS buffer at pH 7, Temperature varied from 5°C to 60°C. **B/** CD signal intensities at λ_{max} *versus* temperature.