Electronic Supplementary Material (ESI) for RSC Advances.
This journal is © The Royal Society of Chemistry 2024

Supporting Information

Novel Diamine-Scaffold based *N*-Acetylgalactosamine (GalNAc)-siRNA Conjugate: Synthesis and *in Vivo* Activities

Qiang Li[a,b,*], Mingxin Dong[a,*], Pu Chen[b,c,*]

[a]Department of Medicinal Chemistry, School of Pharmacy, Qingdao University, Qingdao 266021, China

[b]Research and Development Department, NanoPeptide (Qingdao) Biotechnology Ltd., Qingdao, China

[c]Department of Chemical Engineering and Waterloo Institute for Nanotechnology, University of Waterloo, Waterloo, ON, Canada

*Corresponding authors. Qiang Li, sdhyliqiang@163.com; Mingxin Dong, Mxdong64@qdu.edu.cn; Pu Chen, p4chen@uwaterloo.ca.

Materials and methods

General. All chemicals were purchased from commercial supplier and used with no further purification. Organic solvents for reactions were dried overnight over an appropriate drying agent. Analytical thin-layer chromatography (TLC) was carried out on YINLONG silica gel HSGF254 plates. Column chromatography was carried out on silica gel (200-300 mesh). ^{1}H and ^{31}P NMR spectra of small molecules were obtained from a Bruker ASCEND 500 MHz. Mass spectras (MS) were performed on Thermo Scientific™ ISQ™ EC Single Quadrupole Mass Spectrometer.

Notice: All animal experiments were performed in compliance with the relevant guidelines.

Synthesis of compound 3.

To a solution of **2** (23.5 g, 98.5 mmol, 1.0 eq.) in 1 L acetonitrile (ACN) was added K_2CO_3 (81.8 g, 592 mmol, 6.0 eq.), NaI (44.4 g, 296 mmol, 3.0 eq.) and **1** (73 g, 306 mmol, 3.1 eq.) at room temperature. After stirring 16 h at 50 ℃, the reaction mixture was filtered through a Celite pad. The resulting filtrate was concentrated and redissolved with 700 mL ethyl acetate, washed with water (300 mL × 2) and brine (300 mL). The organic layer was dried over sodium sulfate, filtered and concentrated on vacuum. The residue was purified on a silica gel column (2%-10% MeOH in DCM). The obtained crude was dissolved with 50 mL ethyl acetate. To above solution, HCl in ethyl acetate (2 M, 205 mL) was added in 4 portions. After stirring 3 h at 25 ℃, the reaction mixture was filtered and the filter cake was washed with 500 mL ethyl acetate and dispersed in 50 mL ethyl acetate. The mixture was concentrated and dried on vacuum to give light yellow solid **3**(27 g, 65%, with 3HCl). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.42 – 7.28 (m, 5H), 4.50 (s, 2H), 4.20 – 4.01 (m, 2H), 3.83 (t, *J* = 4.9 Hz, 2H), 3.58 (hept, *J* = 2.7 Hz, 4H), 3.52 – 3.30 (m, 14H), 3.01 (d, *J* = 5.8 Hz, 2H), 2.90 (q, *J* = 5.9 Hz, 6H). LCMS (ESI): $C_{25}H_{47}N_8O_5$, $[M + H]^+$ calc. 839.52, found, 839.54.

Synthesis of compound 6.

A mixture of **4** (50 g, 213.4 mmol, 1.0 eq.) and Galactosamine pentaacetate (80.5 g, 219.8 mmol, 1.03 eq.) was dissolved in anhydrous DCM (750 mL). TMSOTf (14.25 g, 64 mmol, 0.3 eq.) in 250 mL anhydrous DCM was added dropwise. The mixture was stirred at 25 ℃ for 72 hours to afford compound **5**, when TLC showed the absence of starting compound. TFA (150 mL) was then added. After stirring 12 hours at room temperature, the reaction mixture was concentrated and redissolved in 750 mL water. After adjusting the pH to 8-9 by slowly adding solid $NaHCO_3$, the aqueous solution was extracted with DCM (300 mL × 2). The pH of aqueous phase was readjusted to 3-4 by slowly adding HCl (3 Mol/L). The above acidic aqueous solution was reextracted with DCM (300 m L × 4). The combined organic phase was dried with anhydrous Na_2SO_4 and evaporated. The obtained reside was purified by column chromatography (silica gel, DCM/MeOH = 50:1~20:1) to give compound **6** (72 g, 141 mmol, 66%) as a white solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.45 (d, *J* = 9.4 Hz, 1H), 5.28 (d, *J* = 2.9 Hz, 1H), 5.05 (dd, *J* = 11.1, 3.4 Hz, 1H), 4.71 (d, *J* = 8.6 Hz, 1H), 4.21 (dt, *J* = 11.0, 9.0 Hz, 1H), 4.11 (dd, *J* = 8.6, 6.7 Hz, 2H), 3.89 (t, *J* = 6.9 Hz, 1H), 3.86 –

3.81 (m, 2H), 3.80 – 3.72 (m, 2H), 3.62 – 3.54 (m, 5H), 2.58 (td, *J* = 5.7, 2.4 Hz, 2H), 2.12 (s, 3H), 2.01 (s, 3H), 1.97 (s, 3H), 1.95 (s, 3H). LCMS (ESI): $C_{21}H_{32}NO_{13}$, $[M - H]^-$ calc. 506.2, found, 506.2.

Synthesis of compound **7**.

A mixture of compound **6** (36.5 g, 71.9 mmol, 1 eq.), NHS (16.6 g, 143.8 mmol, 2 eq.) and EDCI (27.6 g, 143.8 mmol, 2.0 eq.) was dissolved in anhydrous DCM (750 mL). Diisopropylethylamine (DIPEA, 27.9 g, 215.8 mmol. 3.0 eq.) was added at 0 ℃. After stirring 16 hours at room temperature, the reaction mixture was quenched and washed by dilute HCl (0.5 M, 250 mL). The resulting aqueous phase was extracted with DCM (250 mL). The combined organic phase was washed with saturated $NaHCO_3$ (250 mL) and brine (250 mL), dried and concentrated. The obtained residue was purified by flash column chromatography (silica gel, DCM/MeOH = 50:1) to give NHS activated GalNAc intermediate **I** (32 g, 53 mmol). The intermediate **I** obtained from above step was redissolved in anhydrous DCM (100 mL) and added dropwise to a solution of **3** (11.1 g, 17.1 mmol) and DIPEA (22 g, 171 mmol) in 500 mL DCM. After stirring 16 hours at room temperature, LCMS showed the starting compound **3** was consumed completely and the mixture was washed with saturated $NaHCO_3$ (250 mL) and brine (250 mL). The organic layer was dried with anhydrous Na_2SO_4 and evaporated. The obtained residue was purified by column chromatography (silica gel, DCM/MeOH = 50:1~7:1) to give intermediate **II** (24.0 g, 12 mmol) as a white solid. The intermediate **II** obtained from above step was redissolved in AR grade MeOH (120 mL). Pd/C (purity 10%, 2.8 g, 20% wt) and formic acid (14 g) was added. After stirring 16 hours at 25 ℃, LCMS showed starting intermediate **II** was consumed completely. The reaction mixture was filtered over Celite and the filtrate was evaporated. The obtained residue was redissolved in DCM (150 mL) and added dropwise to Methyl tert-butyl ether (MTBE, 250 mL). After stirring for 30 min, the mixture was filtered and a white solid was obtained. The solid was purified by RP-HPLC (C18 column, phase A: 0.05% TFA, phase B: ACN, 0~5 min, 5%~20% B, 5~60 min, 20%~40% B.). The combined solution was freeze-dried to give compound **7** (12.6 g, 6.6 mmol, 45% from compound **3**) as a white solid. LCMS (ESI): $C_{40.5}H_{67.5}N_{5.5}O_{20.5}$, $[0.5M + H]^+$ calc. 958.9, found, 959.1.

Synthesis of compound **8**.

Compound **7** (10 g, 5.2 mmol, 1.0 eq.) was co-evaporated with anhydrous MeCN three times, and then dissolved in anhydrous DCM (50 mL). 5-Ethylmercapto-1*H*-tetrazole (677 mg, 5.2 mmol, 1.0 eq.) was added, followed by 2-Cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropyl-phosphorodiamidite (2.48 ml, 7.8 mmol) at room temperature. The reaction mixture was stirred under N_2 atmosphere at room temperature for 4 hours, when TLC indicated the absence of starting material. The reaction mixture was diluted with 50 mL DCM, washed with saturated $NaHCO_3$ (50 × 2 mL) and brine (50 mL), dried over sodium sulfate, followed by solvent removal. The crude product was purified by silica gel column chromatography using a 15:1 mixture of dichloromethane : methanol and 1% triethylamine. The obtained

product was co-evaporated with anhydrous MeCN (30 mL) twice to give compound **8** (6.9 g, 3.2 mmol, 63%) as a white solid. LCMS (ESI): $C_{45}H_{76}N_{6.5}O_{21}P_{0.5}$, $[0.5M + H]^+$ calc. 1058.9, found, 1059.0.

Synthesis of compound **10**.
Compound **7** (5.0 g, 2.6 mmol, 1.0 eq.) was dissolved in anhydrous DMF (35 mL). The reaction flask was flushed with nitrogen three times, and 1,1'-carbonyldiimidazole (634 mg, 3.9 mmol, 1.3 eq.) was added. The mixture was stirred for 16 hours under nitrogen atmosphere. 1-amino-3-O-(4,4'-dimethoxytrityl)-2,3-propanediol (compound **9**) (2.26 g, 5.74 mmol, 2.2 eq.) was added at room temperature. After stirring overnight, the mixture was purified by flash column chromatography (silica gel, DCM/MeOH = 8:1), followed RP-HPLC (C18 column, phase A: 0.1% TEA, phase B: ACN, 0~50 min, 10%~70% B) to afford compound **10** (4.0 g, 1.7 mmol, 65%) as a white solid. LCMS (ESI): $C_{53}H_{79}N_6O_2Na$, $[0.5M + Na]^+$ calc. 1090.5, found, 1190.6.

Synthesis of compound **11**.
Compound **10** (872 mg, 0.37 mmol, 1.0 eq.) was co-evaporated with anhydrous MeCN three times, and then dissolved in dry DCM (10 mL). Triethylamine (0.16 mL, 1.12 mmol, 3.0 eq.), 4-dimethylaminopyridine (10 mg, 0.076 mmol, 0.2 eq.) was added, followed by succinic anhydride (75 mg, 0.74 mmol, 2.0 eq.) at room temperature. After stirring overnight, the reaction mixture was diluted with DCM (40 mL), extracted with saturated $NaHCO_3$ (10 mL) and brine (20 mL), dried and evaporated. Compound **11** (820 mg, 0.33 mmol, 90%) was obtained as a white solid, which was used without any further purification.

Synthesis of CPG **12**.
A mixture of compound **11** (500 mg, 0.2 mmol, 1.0 eq.), *O*-(7-azabenzotriazollyl)-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluophosphate (HBTU, 152 mg, 0.4 mmol, 2.0 eq.) and DIPEA (103 mg, 0.8 mmol, 4.0 eq.) in anhydrous MeCN (5 mL) was stirred for 10 min at room temperature. Aminoalkyl CPG (360 mg, 0.068 mmol, 500 Å, purchased from He Bei DNA chem, LOT: 3006-500) was added. After stirring overnight, the mixture was filtered and the solid was washed with DCM:MeOH (V/V=10:1) (15 mL × 2) and MeCN (15 mL × 2). The obtained residue was dried in a drying oven at 60 °C for 2 hours and dispersed in an acetic anhydride/pyridine mixed solution (V/V=1:3, 4 mL). After stirring overnight, the mixture was refiltered, rewashed with DCM:MeOH (V/V=10:1) (15 mL × 2) and MeCN (15 mL × 2) and redried in a drying oven at 60 °C for 2 hours to give CPG **12** (300 mg) as a light yellow solid. The loading amount of CPG **12** was determined to be 0.053 mmol/g using HPLC by detecting the released DMTr group.

The synthesis of siRNA.
The RNA oligomers **13-15** and **17-22** were synthesized via standard solid phase RNA synthesis using a Biolytic Lab Performance Dr.Oligo 48 synthesizer in 4 μmol scale. The 2'-OMe and 2'-F modified phosphoramidates which purchased from commercial supplier, were 2 eq. and performed common coupling time (10 min), while others were 4 eq. with extended coupling time (30~60 min). After synthesis, solid supports were treated with 4

mL mixyure of 28% NH_3`H_2O and 40% CH_3NH_2 (V/V, 1:1) at 37 °C for 12 hours. The supernatant was then separated from the solid supports, diluted with water, purified by anion-exchange chromatography column (PS-15Q, phase A, 10 mM PBS, phase B, 10 mM PBS with 2 M NaCl). The desired eluent (the UV purities of oligomers more than 85%) were collected and desalted by ultrafiltration using an ultrafiltration tube.
As for RNA strand **16**, C7-NH2 CPG was used. After the regular synthesis, deprotection and purification, The NH2 group at 3'-terminal was conjugated with commercial Alexa Fluor™ 647 NHS Ester (Succinimidyl Ester) under the catalysis of $NaHCO_3$ in mixed H_2O/ACN. The pure 16 was purified by reverse phase C18 chromatography (phase A, 10 mM TEAA, phase B, ACN)
The integrity of these oligomers was characterized by ESI mass spectroscopy.

Annealing of siRNA

The obtained sense and antisense strand after ultrafiltration was diluted to 3 mg/mL with RNase-free water. The sense strand combined with 1 equipment of antisense strand was heated for 5 min at 90 °C in a water bath. The mixture was naturally cooled down to room temperature. The integrity of siRNA was detected by MS. All annealed siRNAs were lyophilized. Once annealed and lyophilized, double stranded siRNA can be safely stored frozen at -20 °C in a freezer.

***In vivo* biodistribution study (fluorescence experiment)**.

All AF647 labeled siRNA was dissolved with sterile PBS buffer to 0.5 mg/mL. Experiments were performed on 3 BALB/c nude mice for negative control and siRNA **24-25**, while 2 BALB/c nude mice for siRNA **26-27**. All mice were randomly divided into groups, fed normally, and administered subcutaneously with siRNA (2 mg/kg, each~90 µL). Plannar scintigraphic images were acquired at 0.25, 0.5, 1, 2, 6, 24 hours post-injection separately using a small *in vivo* imaging system (IVIS lumina XRMS Series III). Before collection, mice were anesthetized using a small animal inhalation anesthesia machine (Matrx UIP3000 2.5 by MIDMARK 0.5). At the middle (9 hours) of image acquisition, one mouse of each group was euthanized, and samples from heart, brain, lung, uterus, spleen, kidney and liver were collected and subjected to fluorescence imaging for the assessment of biodistribution profiles. At the end of image acquisition, the same procedure above was performed for the remaining mice in all groups.
The fluorescence imaging data of average radiant efficiency were analyzed by Live image 4.5.2 system.

Cell viabilities study

Human liver cancer Hep3B cells obtained from the Shanghai Cell Bank, Chinese Academy of Sciences, were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, US) supplemented with 10% fetal bovine serum (FBS) (Gibco, US) at 37°C in a 5% CO2 atmosphere (Thermo Fisher Scientific). Following a 24-hour incubation period, transfection was carried out. The cells underwent digestion using 0.25% Trypsin (Gibco, US), were quantified, and then placed in a 96-well plate at a concentration of 5000 cells per well with a volume of 100 µL per well. Utilize RNAiMAX (Thermo Fisher) as the transfection reagent

and follow the standard protocol for transfection, using 0.25 μL of RNAiMAX per well. The concentrations of siRNA in the final solution are 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, and 1 nM, 0 nM. Following a 72-hour transfection period, the supernatant was aspirated, and CCK8 (Solebao) was introduced. Subsequently, the mixture was allowed to incubate for 1 to 2 hours, and the absorbance was quantified at 450 nm utilizing a full-wavelength microplate reader (SYNERGY H1, BioTek). The CCK8 value of the culture wells without siRNA added in the transfection system was normalized to 100%, and the cell survival rate at each concentration was calculated as the average of two independent repeated experimental measurements. SiRNA targeting Polo-like kinase 1 (PLK1) was used as the positive control (SS: 5’-CCUUGAUGAAGAAGAUCACdTdT-3’; AS: 5’-AAGUGAUCUUUCUUCAUCAAGG-3’)

***In vivo* gene silence study**.

The experimental wild-type C57BL/6J mice were randomly divided into groups (n = 7 for each indicated time) and administered subcutaneously with 1 mg/kg test siRNA, prepared in an injection of 10 μL/g body weight in PBS. At the indicated time (Day 0, 14, 39), the mice were killed and liver tissues was obtained. The whole RNA was extracted from the liver tissue using trizol or magnetic beads. 1 μg RNA was chosen and reverse transcription was performed by Reverse transcription kit (HiScriptRII 1st Strand cDNA Synthesis Kit R211-01/02) to give cDNA. The cDNA samples were diluted and quantified by method of SYBR green QPCR. Mouse Ttr forward: 5’-CTG CTG TAG ACG TGG CTG TAA-3’; reverse: 5’-CTT CCA GTA CGA TTT GGT GTC C-3’. Mouse Gapdh forward: 5’-TGT GTC CGT CGT GGA TCT GA-3’，reverse：5’-TTG CTG TTG AAG TCG CAG GAG-3’. The results of samples were analyzed by method of delta delta CT.

Supplementary Figures

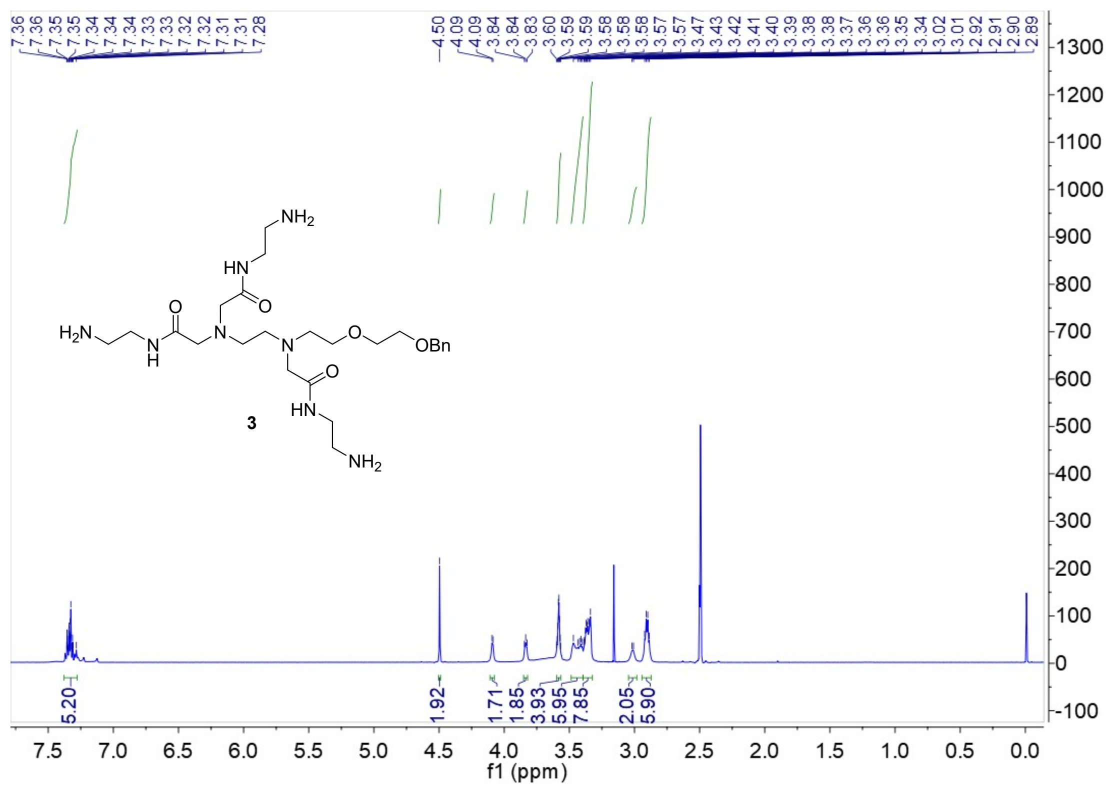


Figure S1. ^{1}H NMR spectra of compound **3**.

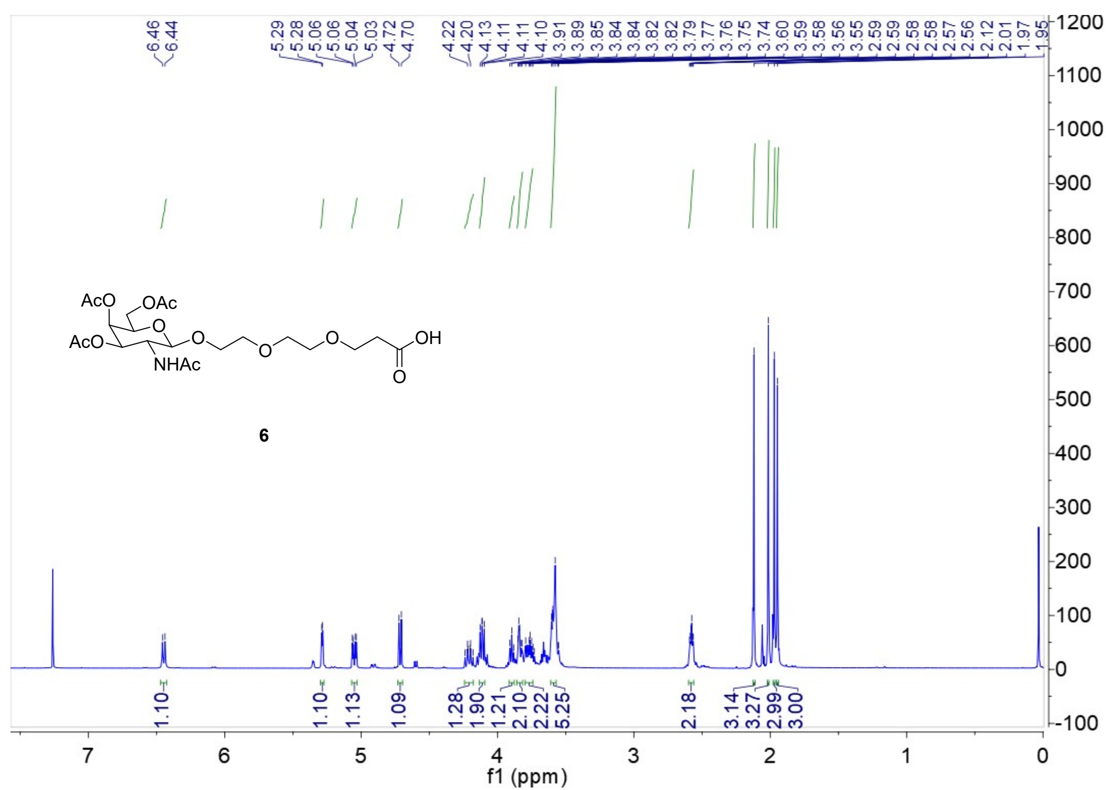


Figure S2. ^{1}H NMR spectra of compound **6**.

Oligomer	Sequence (5’-3’)	Cal. MS	Act. MS
13	gcucaacaUAUuugaucagu*a	6916.5	6915.7
14	**TG1***gcucaacaUAUuugaucagu*a	8532.4	8532.4
15	**NAG37***invAb*gcucaacaUAUuugaucagua*invAb	8645.5	8645.5
16	g*cucaacaUAUuugaucagua***TG2**	8549.0	8548.7
17	u*A*c*UgAuCaAaUaUgUuGaGc*C7-NH2-AF647	7970.7	7969.1
18	c*a*guguUCUugcucuauaa**L96**	8011.2	8011.1
19	**TG1***caguguUCUugcucuau*a*a	7839.0	7839.0
20	c*a*guguUCUugcucuauaa**TG2**	7940.0	7939.9
21	c*a*guguUCUugcucuauaa***TG2**	7856.0	7856.0
22	c*a*guguUCUugcucuaua*a***TG2**	7872.0	7872.1
23	u*U*auaGagcaagaAcAcug*u*u	7015.8	7014.9

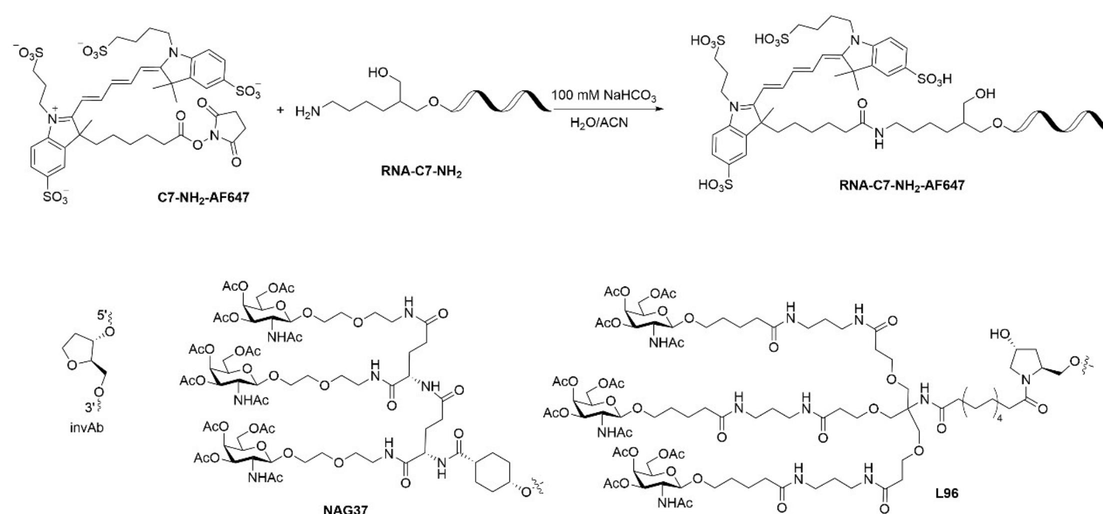


Figure S3. Up: Calculated and observed masses of sense and antisense strands used in this study. Middle: Synthesis of C7-AF647 modified oligomer **17**. Bottom: structure details of conjugated moietis (NAG37, invAb and L96) in oligomer **15** and **18**.

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
13	gcucaacaUAUuugaucagu*a	6916.5	6915.7

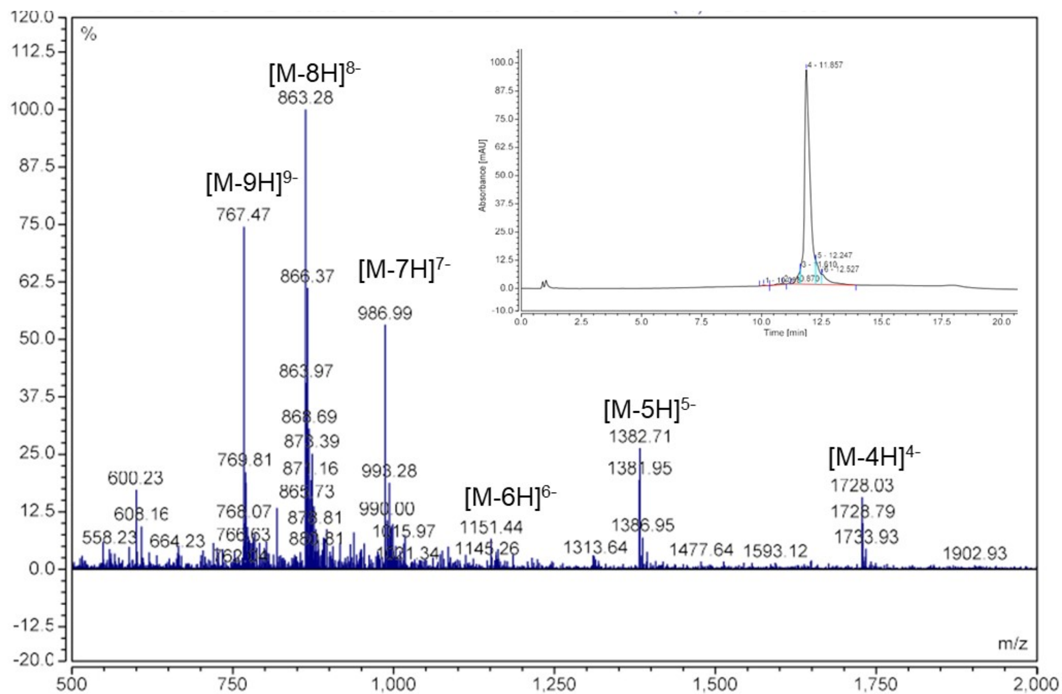


Figure S4. Mass spectrum of oligomer **13** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
14	**TG1***gcucaacaUAUuugaucagu*a	8532.4	8532.4

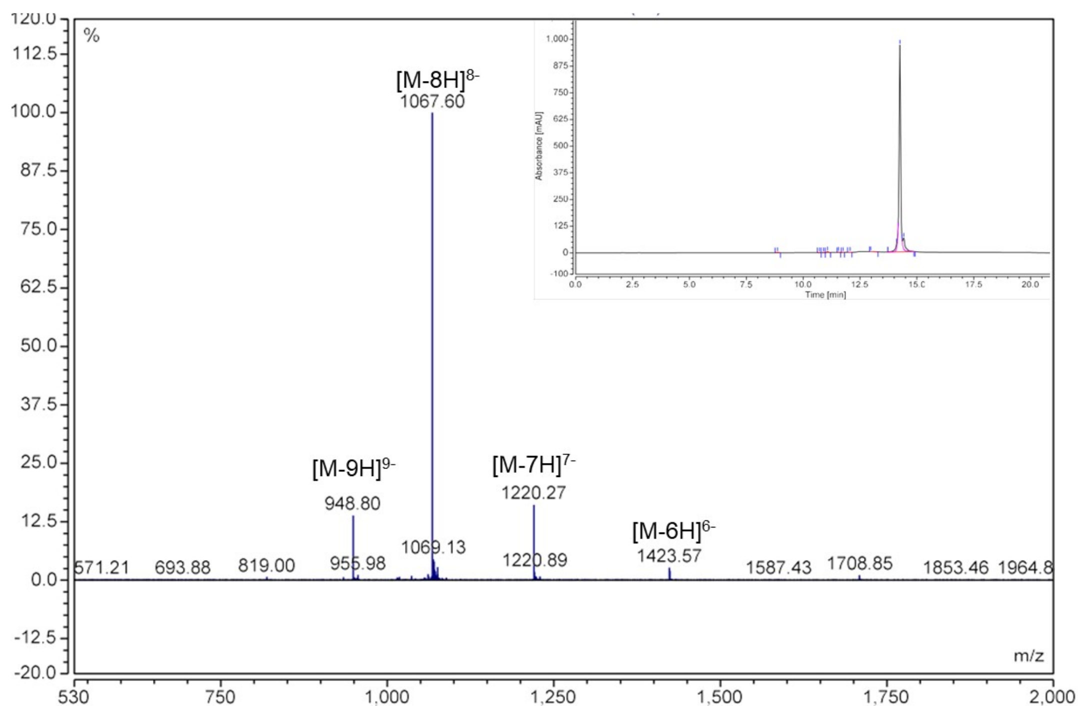


Figure S5. Mass spectrum of oligomer **14** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
15	**NAG37***invAb*gcucaacaUAUuugaucagua*invAb	8645.5	8645.5

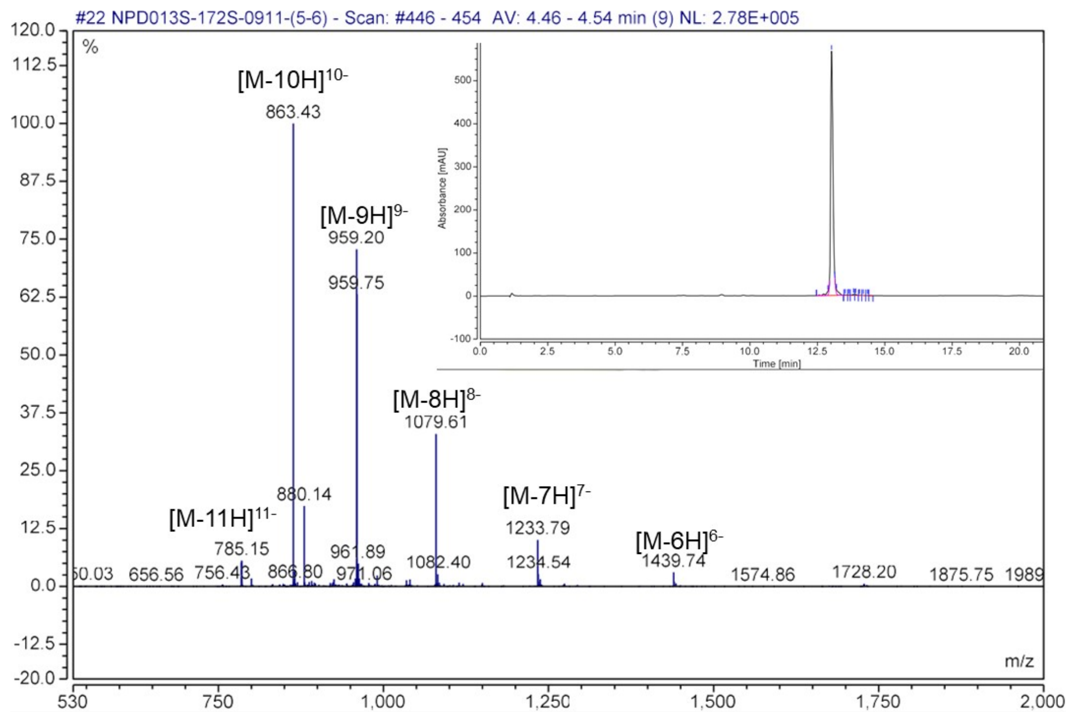


Figure S6. Mass spectrum of oligomer **15** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
16	g*cucaacaUAUuugaucagua***TG2**	8549.0	8548.7

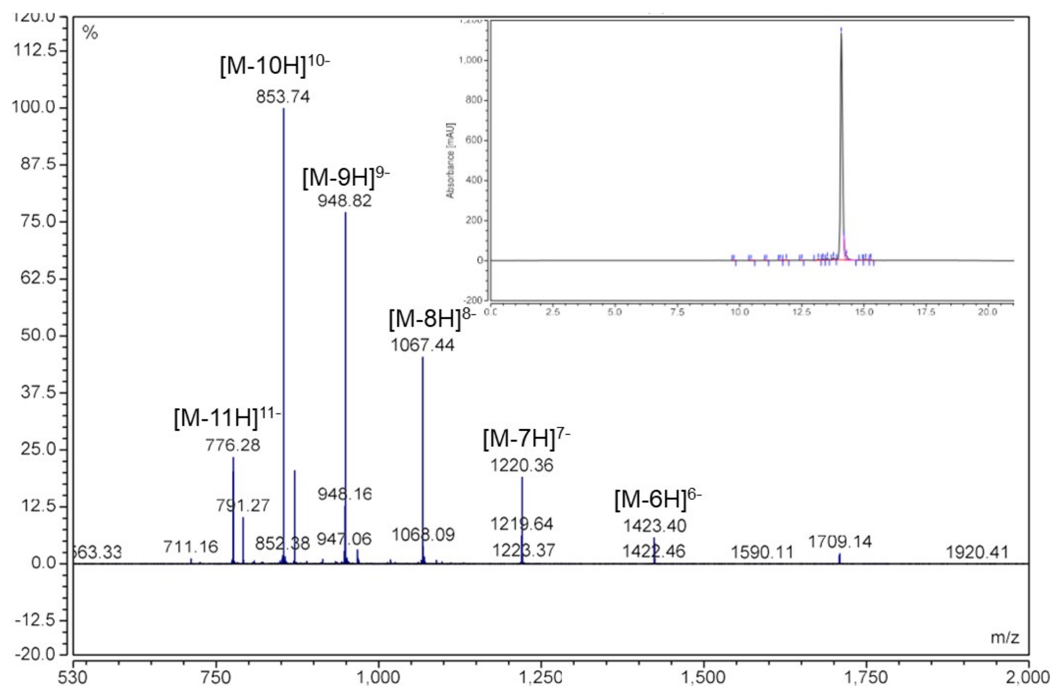


Figure S7. Mass spectrum of oligomer **16** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
17	u*A*c*UgAuCaAaUaUgUuGaGc*C7-NH2-AF647	7970.7	7969.1

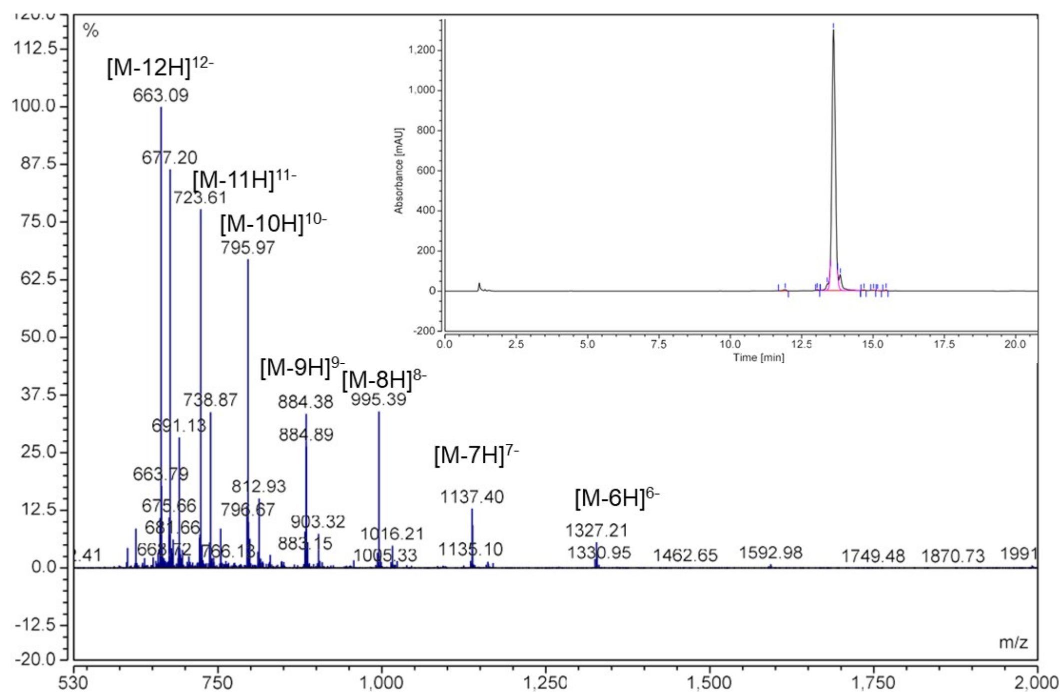


Figure S8. Mass spectrum of oligomer **17** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
18	c*a*guguUCUugcucuauaa**L96**	8011.2	8011.1

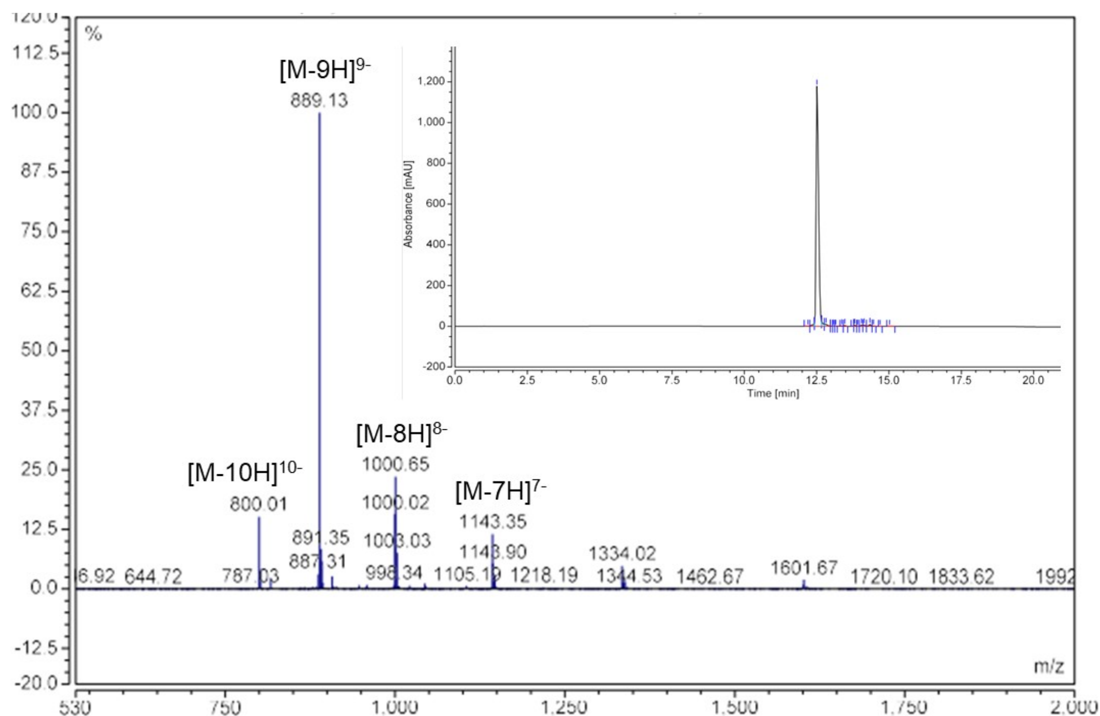


Figure S9. Mass spectrum of oligomer **18** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
19	**TG1***caguguUCUugcucuau*a*a	7839.0	7839.0

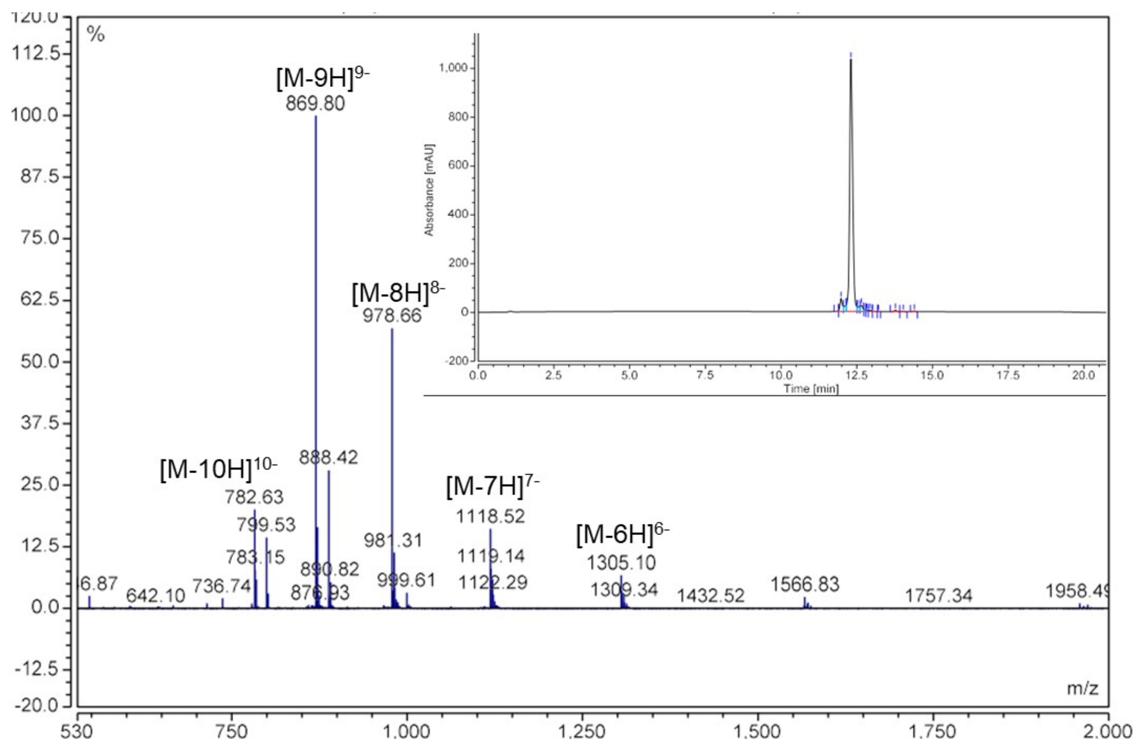


Figure S10. Mass spectrum of oligomer **19** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
20	c*a*guguUCUugcucuauaa**TG2**	7940.0	7939.9

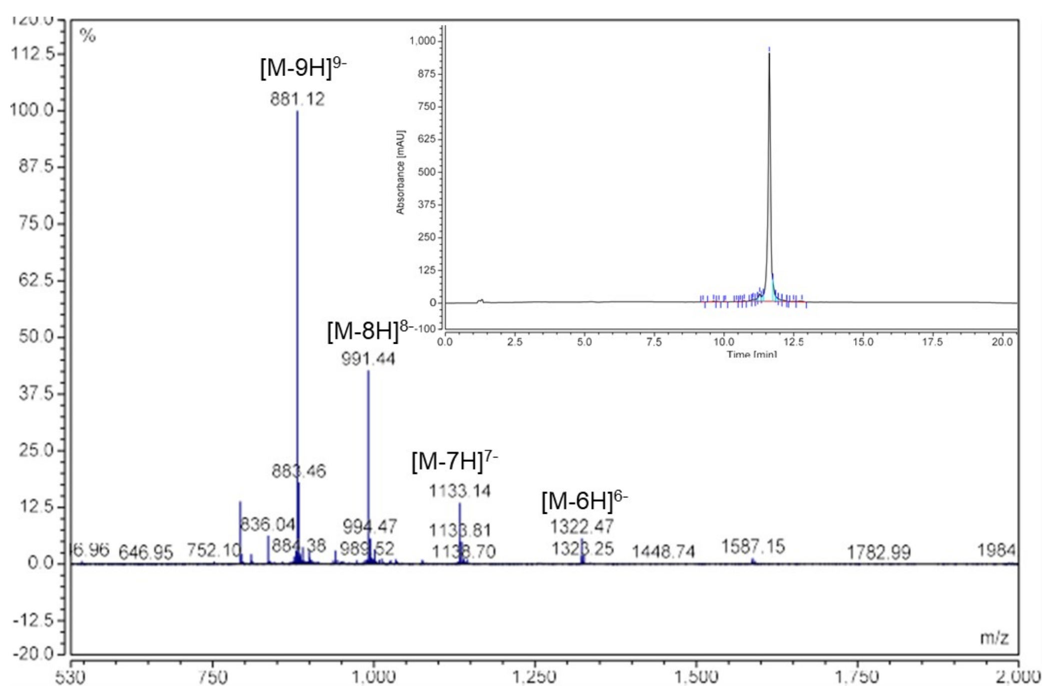


Figure S11. Mass spectrum of oligomer **20** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
21	c*a*guguUCUugcucuauaa***TG2**	7856.0	7856.0

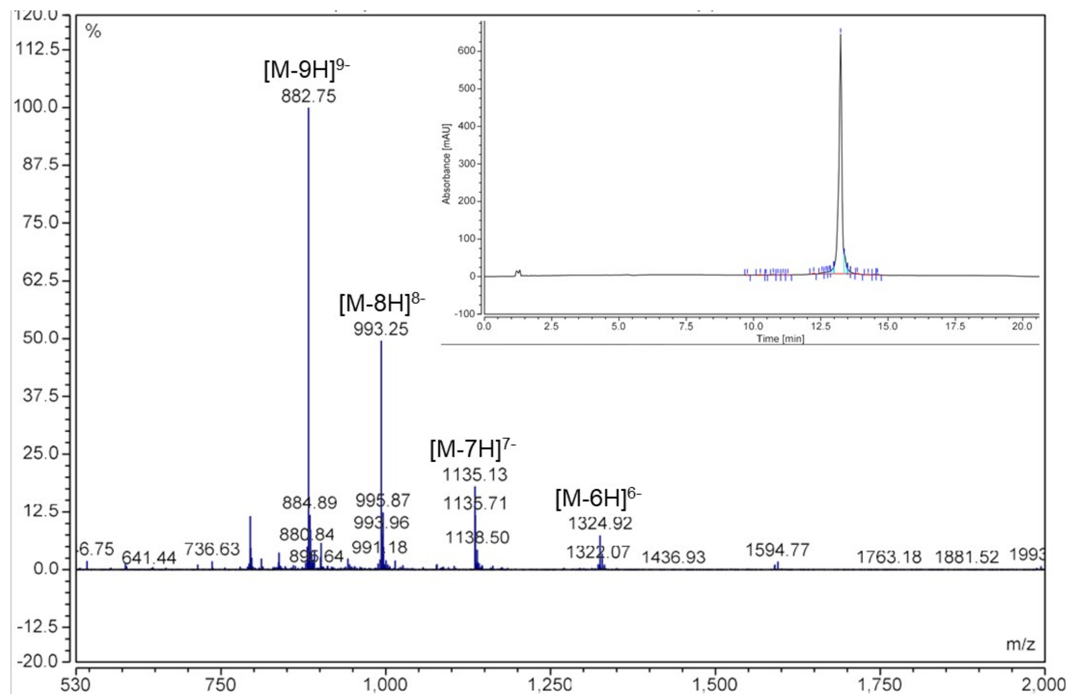


Figure S12. Mass spectrum of oligomer **21** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
22	c*a*guguUCUugcucuaua*a***TG2**	7872.0	7872.1

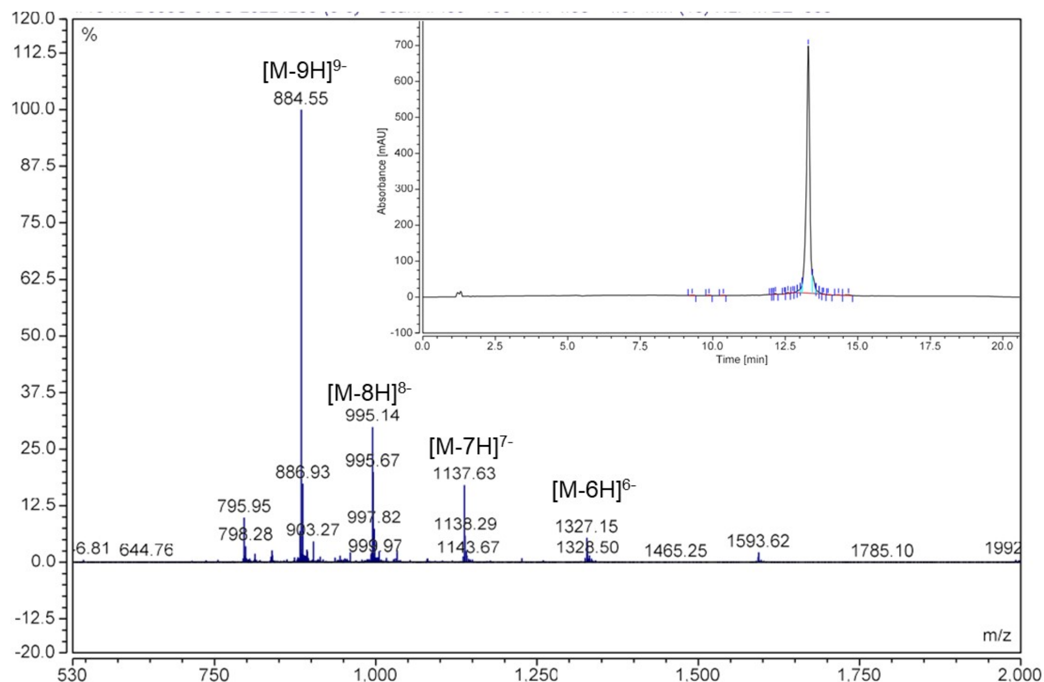


Figure S13. Mass spectrum of oligomer **22** (with HPLC spectrum inside).

Oligomer	Sequence (5'-3')	Cal. MS	Act. MS
23	u*U*auaGagcaagaAcAcug*u*u	7015.8	7014.9

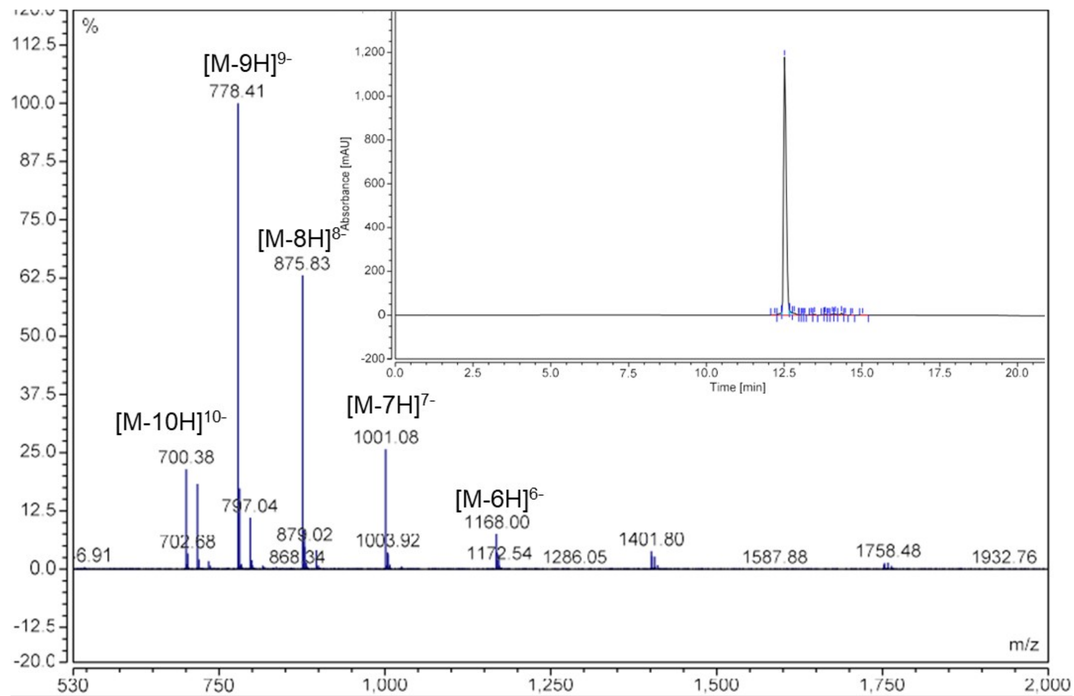


Figure S14. Mass spectrum of oligomer **23** (with HPLC spectrum inside).

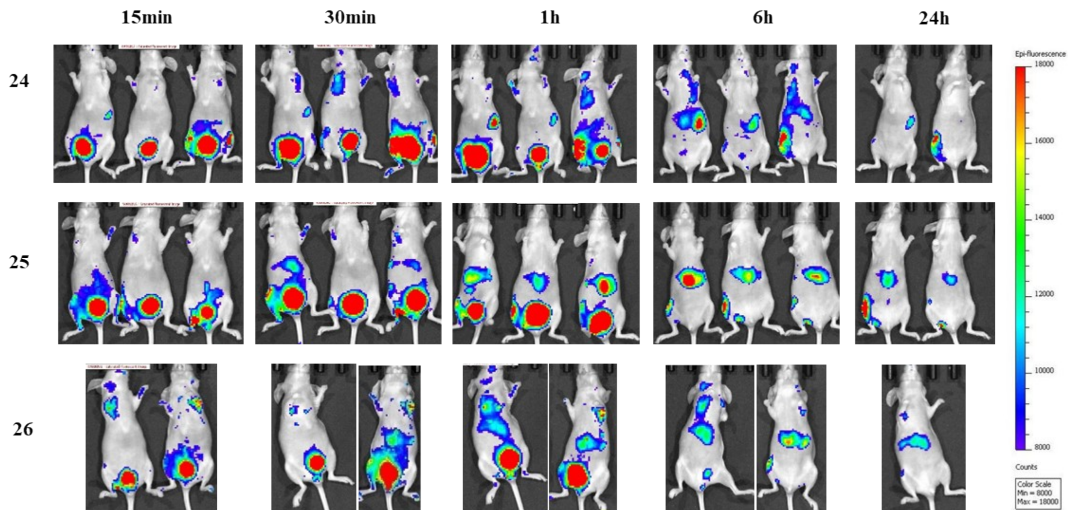


Figure S15. Imaging results of *in vivo* uptake of siRNA **24-26** at various times post-injection.

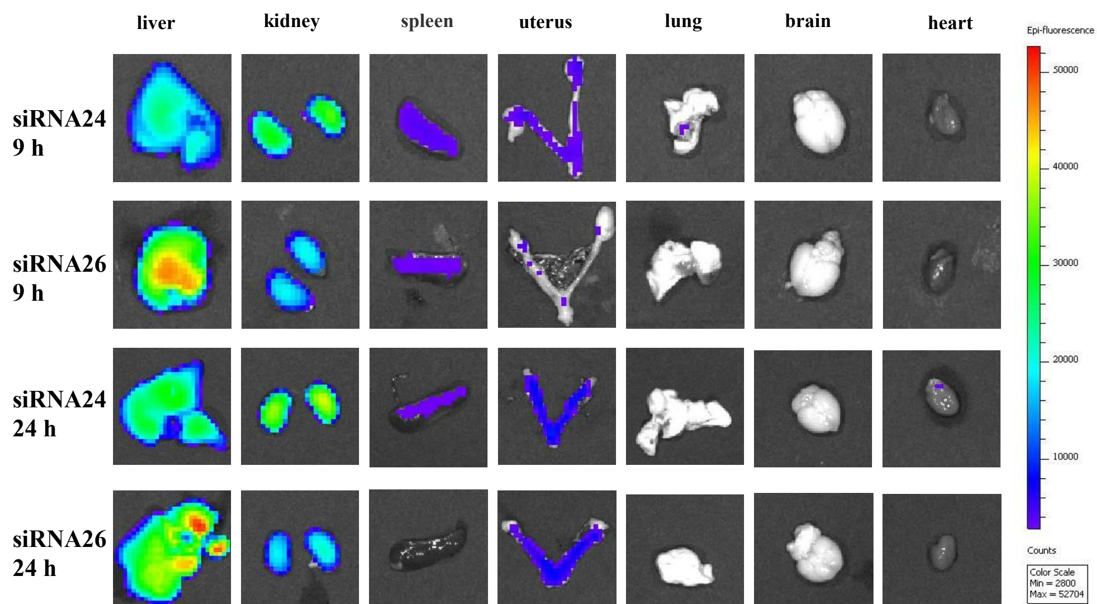


Figure S16. Representative organ images showing target-delivery capabilities of TG1 conjugated siRNA **24** and NAG37 conjugated siRNA **26** (9 h or 24 h post-injection into mice). Organ type and fluorescence signal intensity are located above and below the graph, respectively

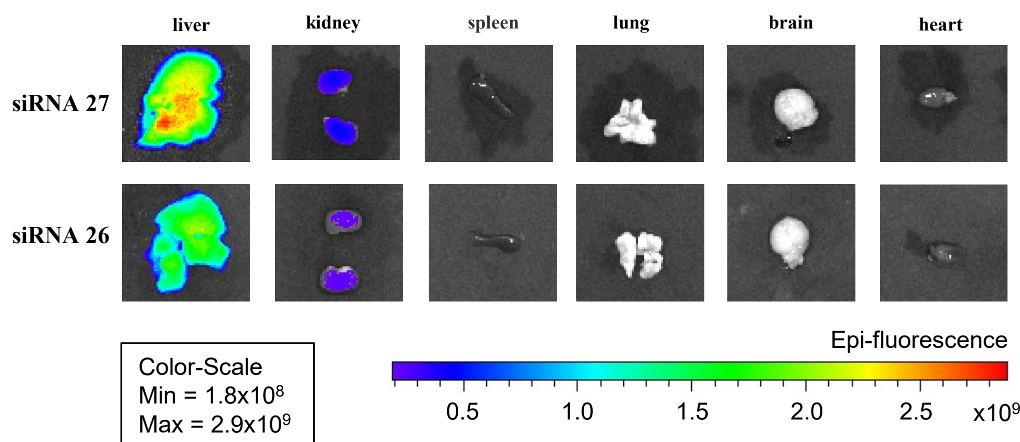


Figure S17. Representative organ images showing target-delivery capabilities of TG2 conjugated siRNA **27** and NAG37 conjugated siRNA **26** (24 h post-injection into mice). Organ type and fluorescence signal intensity are located above and below the graph, respectively.

siRNA	S/AS[a]	Sequence (5'-3')[b]	Target [c]
no GalNAc	**13**	gcucaacaUAUuugaucagu*a	ANGPTL3
	17b	u*A*c*ugAuCaAaUaUgUuGaG*c	
5'-TG1	**14**	**TG1***gcucaacaUAUuugaucagu*a	ANGPTL3
	17b	u*A*c*ugAuCaAaUaUgUuGaG*c	
3'-TG2	**16**	g*cucaacaUAUuugaucagua***TG2**	ANGPT3
	17b	u*A*c*ugAuCaAaUaUgUuGaG*c	
PLK1	**SS**	*CCUUGAUGAAGAAGAUCAC*dTdT	PLK1
	AS	*AAGUGAUCUUUCUUCAUCAAGG*	

[a]S indicates sense strands and AS indicates antisense strands. [b]Italics, upper-case, and lower-case letters represent unmodified, 2'-deoxy-2'-fluoro (2'-F), and 2'-*O*-methyl (2'-OMe) sugar-modified uridine (U), adenosine (A), guanosine (G), and cytidine (C), respectively. * Indicates a phosphorothioate (PS) linkage. **TG1** and **TG2** represent different GalNAc ligands. [c]ANGPTL3, angiopoietin-like protein 3, PLK, Polo-like kinase 1.

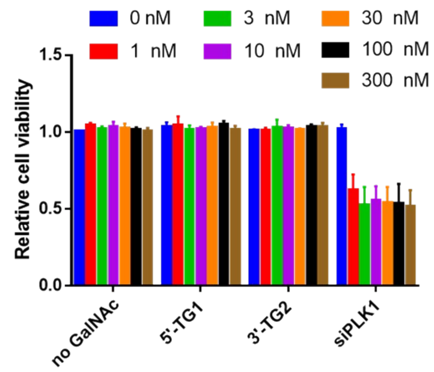


Figure S18. Cell viability of 5'-TG1- and 3'-TG2 conjugated siRNA with the sequence depicted on the left.

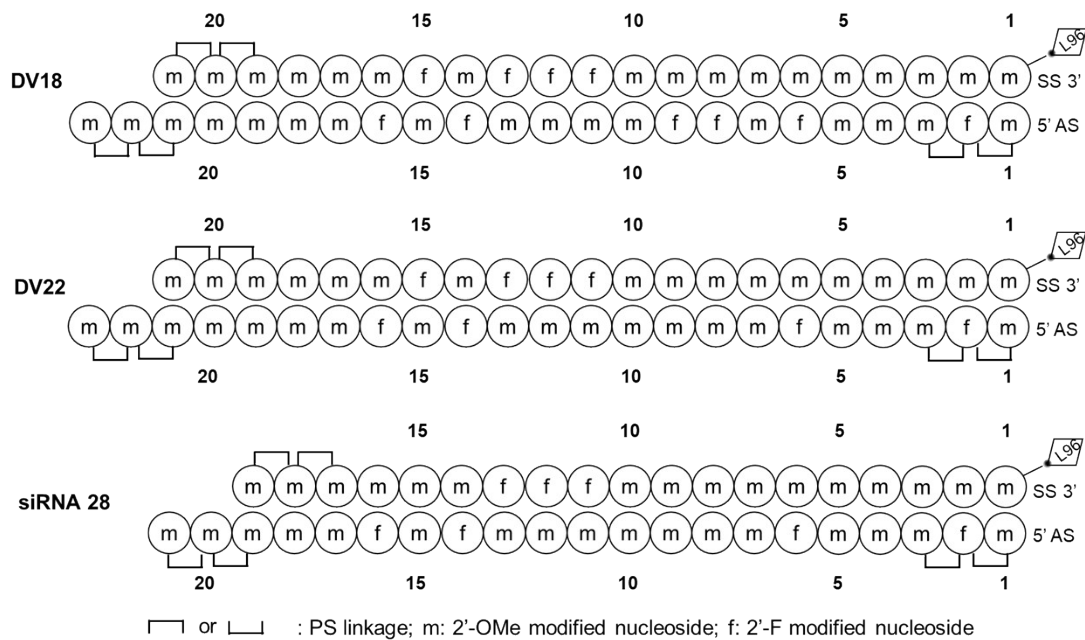


Figure S19. SiRNA design details of DV18, DV22 and siRNA **28**.

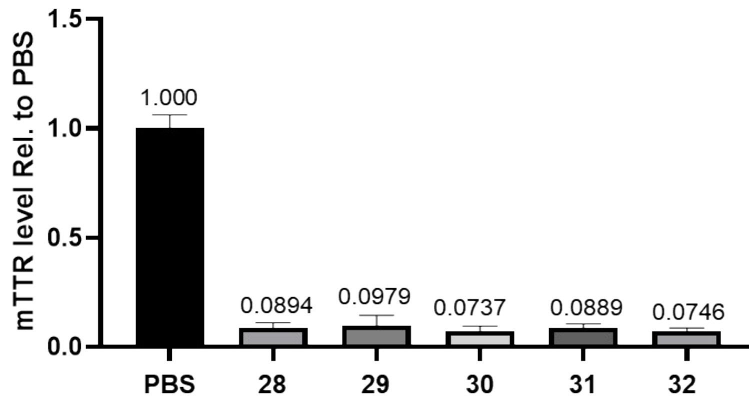
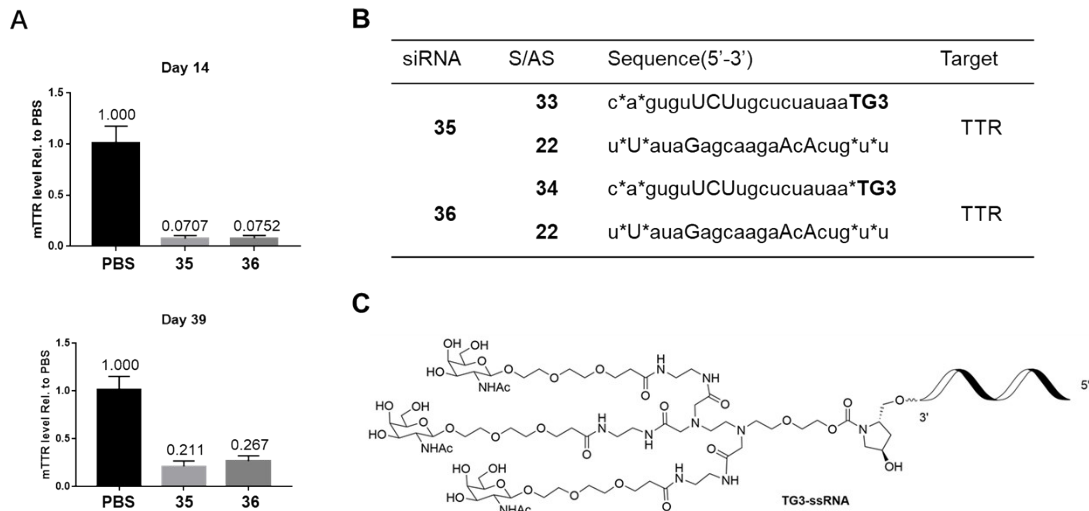


Figure S20. TTR gene silencing results of 14 days after single dose SC at 1 mg/kg with siRNA **28-32** in wild-type C57BL/6 mice (n = 7). Error bars are standard errors for TTR mRNA (mTTR) measurements.



siRNA	S/AS	Sequence(5'-3')	Target
35	**33**	c*a*guguUCUugcucuauaa**TG3**	TTR
	22	u*U*auaGagcaagaAcAcug*u*u	
36	**34**	c*a*guguUCUugcucuauaa***TG3**	TTR
	22	u*U*auaGagcaagaAcAcug*u*u	

Figure S21. (A)TTR gene silencing results of 14 (up) and 39 (down) days after single dose SC at 1 mg/kg with siRNA **35, 36** in wild-type C57BL/6 mice (n = 7). Error bars are standard errors for TTR mRNA (mTTR) measurements. (B) Sequence information for siRNA **35** and **36**. (C) Structural details for TG3.