Supporting Information

Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition

Khirud Gogoi* and Vaijayanti A Kumar*

Division of Organic Chemistry, National Chemical Laboratory, Pune-411008, India
E-mail: va.kumar@ncl.res.in

S2-S3: Experimental procedures

S4-S6: ^{1}H NMR spectra of **7** and **8**, ^{13}C and DEPT NMR spectra of **8** and Mass spectra of compound **8**

S6-S7: Oligomer synthesis, Solid phase synthesis scheme and structure of ONs synthesized.

S8-S12: Table of ONs synthesized and their RP-HPLC and MALDI-TOF mass spectral data.

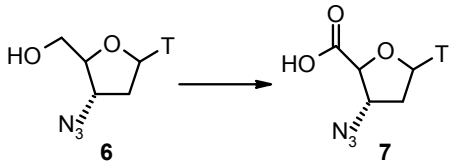
S13-S16: UV-melting experiments and UV-Job's plots.

S16-S17: CD graphs.

Experimental Procedure:

Melting points of samples were determined in open capillary tubes using Buchi Melting point B-540 apparatus and are uncorrected. IR spectra were recorded on an infrared Fourier Transform spectrometer .Column chromatographic separations were performed using silica gel 60-120 and 230-400 mesh, Ethyl acetate, Petroleum ether, Dichloromethane and Methanol as the solvent system. ^{1}H and ^{13}C spectra were obtained using Bruker AC 200 (200MHz) NMR spectrometers.

Synthesis of 3'-azidothymidine-5'-carboxylic acid **7**[1]:

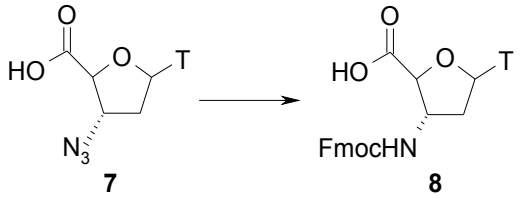


BAIB (3.9g, 12.4 mmol) and TEMPO (180mg, 1.12mmol) and 3'-azido-thymidine **6** (1.5g, 5.6 mmol) were combined in a reaction vessel and to the mixture was added 12 ml of 1:1 acetonitrile-water solution. The reaction mixture was stirred for 3h. Solvents were removed in vacuo and the resulting residue was triturated sequentially with diethyl ether and acetone, filtered and dried in vacuo. Yield: 1.22g, 78%.

M.P. 174-177 ºC FT-IR (Nujol) 3440, 2110, 1748, 1721, 1694, 1610 cm^{-1}

^{1}H NMR (DMSO-d_6) δ 1.78(s, 3H), 2.2(m, 2H), 4.35(d, 1H), 4.64(m, 1H), 6.2(t, 1H), 8.36(s, 1H).

Synthesis of 3'-*N*-Fmoc-3'-aminothymidine-5'-carboxylic acid **8**:



[1] J. B. Epp and T. S. Widlanski, *J. Org. Chem.* 1999, **64**, 293-295

3'-azidothymidine-5'-carboxylic acid **7 (**1.2 g, 4.26 mmol) was dissolved in 10 ml methanol and to it was added 120 mg (10%) Pd-C catalyst. The mixture was subjected to hydrogenation at 40 Psi of hydrogen pressure for 3.5 h. The catalyst was filtered over celite and concentration of the filtrate in vacuo gave 1.1g, 97% of 3'-deoxy-3'-amino thymidine-5'-carboxylic acid.

The amino acid 3'-deoxy-3'-aminothymidine-5'-carboxylic acid (1.1g, 4.3 mmol) was treated with Fmoc-succinimide (1.74g, 5.17 mmol) and $NaHCO_3$ (1.81g, 21.5 mmol) in 10 ml 1:1 acetone-water. After stirring for five hours, acetone was evaporated and the reaction mixture was diluted with 5 ml water. The water layer was extracted with ethyl acetate to remove all the nonpolar unwanted impurities. The water layer was neutralized by adding dilute HCl to bring the pH=7.0. The water layer was then extracted with ethyl acetate. The organic layer was washed with water (1x10ml) and saturated NaCl solution (1x10ml), dried over Na_2SO_4 and evaporated to dryness to afford 1.25g of **8** (61%) as pure white amorphous solid. M.p.147-150°C.

^{1}H NMR ($CDCl_3$ + 1 drop DMSO-d_6, 200 MHz) 1.87(s, 3H), 2.0-2.2(m, 1H), 2.3-2.4(m, 1H), 4.1-4.5(m, 4H), 6.46(t, 3H), 7.2-7.4(m, 5H), 7.6-7.73(m, 5H), 8.1(s, 1H), 10.4(s, 1H).

^{13}C NMR (DMSO-d_6, 50 MHz) 12.4, 36.2, 46.6, 54.5, 65.6, 81.5, 85.1, 109.3, 120.1, 125.1, 127.1, 127.6, 136.4, 140.7, 143.8, 150.5, 155.5, 163.7, 172.2.

Mass calc. 477.42 obs. 499.62(+Na^+)

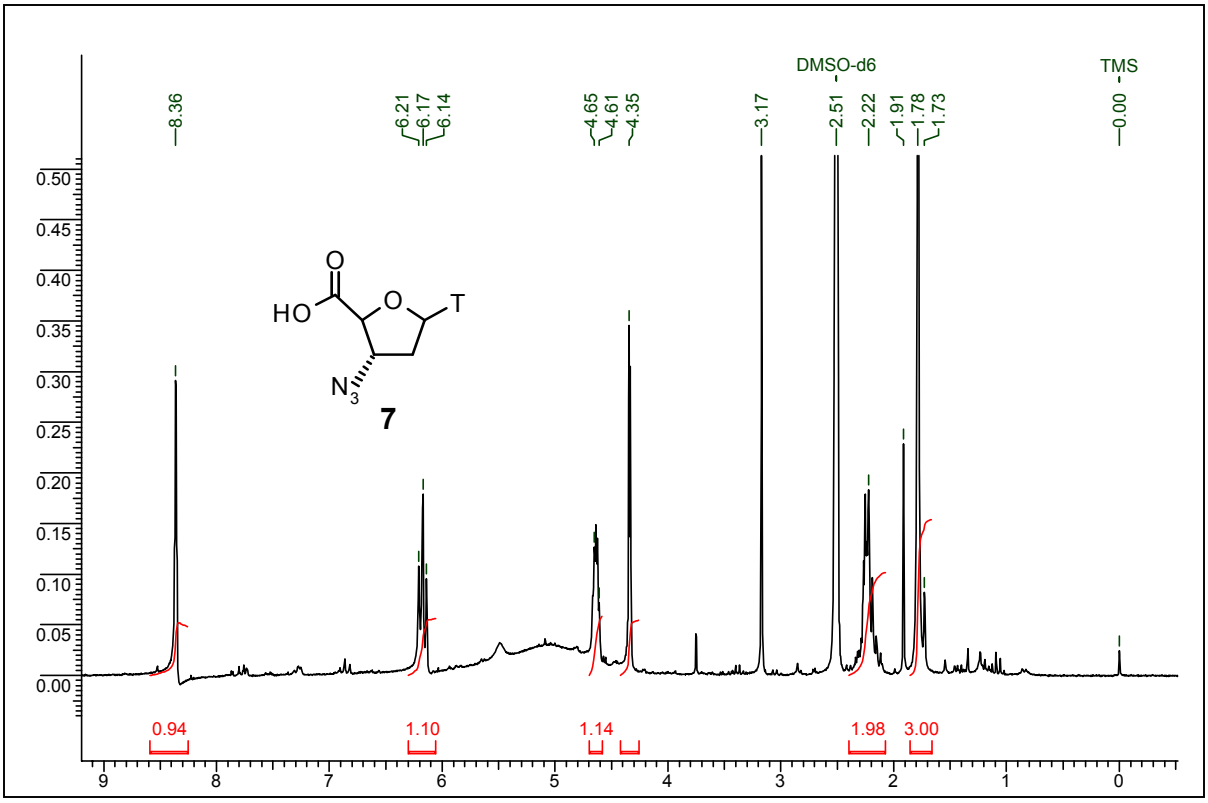


Figure **S1**: 1H NMR of **7**

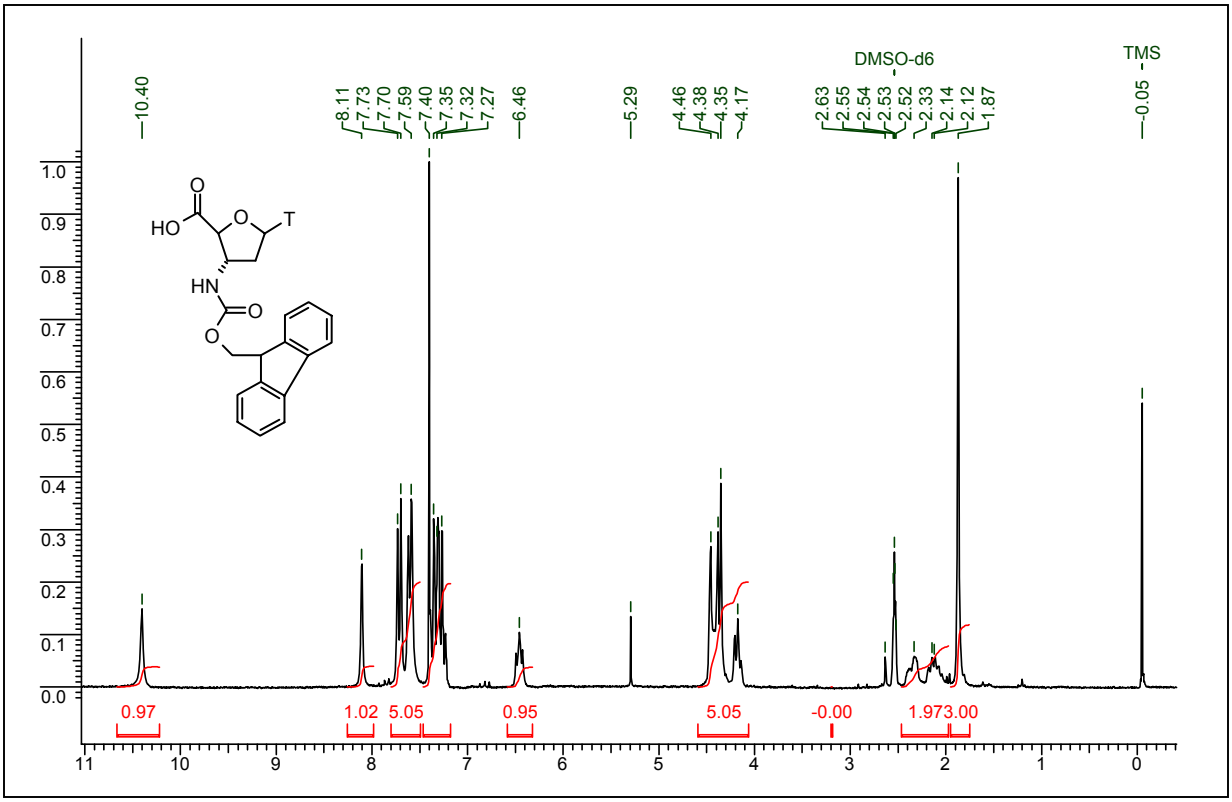


Figure **S2**: ^{1}H NMR of **8**

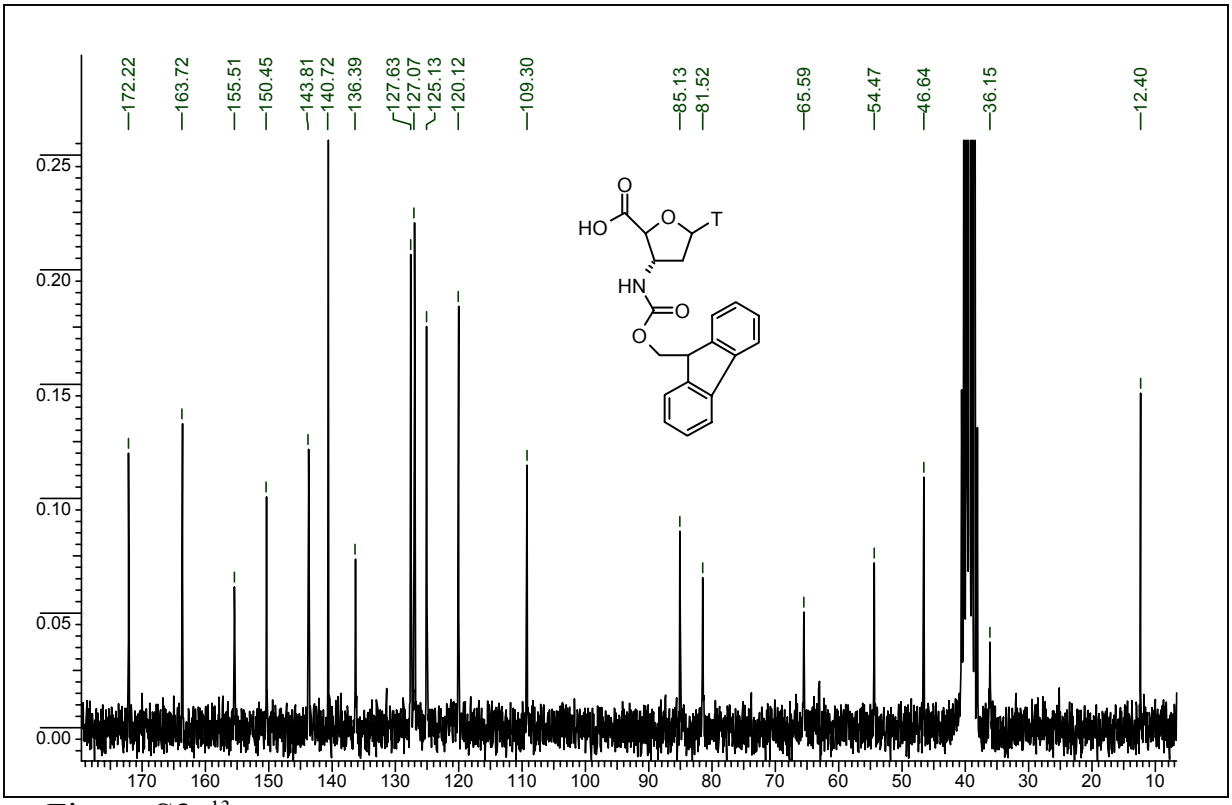


Figure **S3**: ^{13}C NMR of **8**

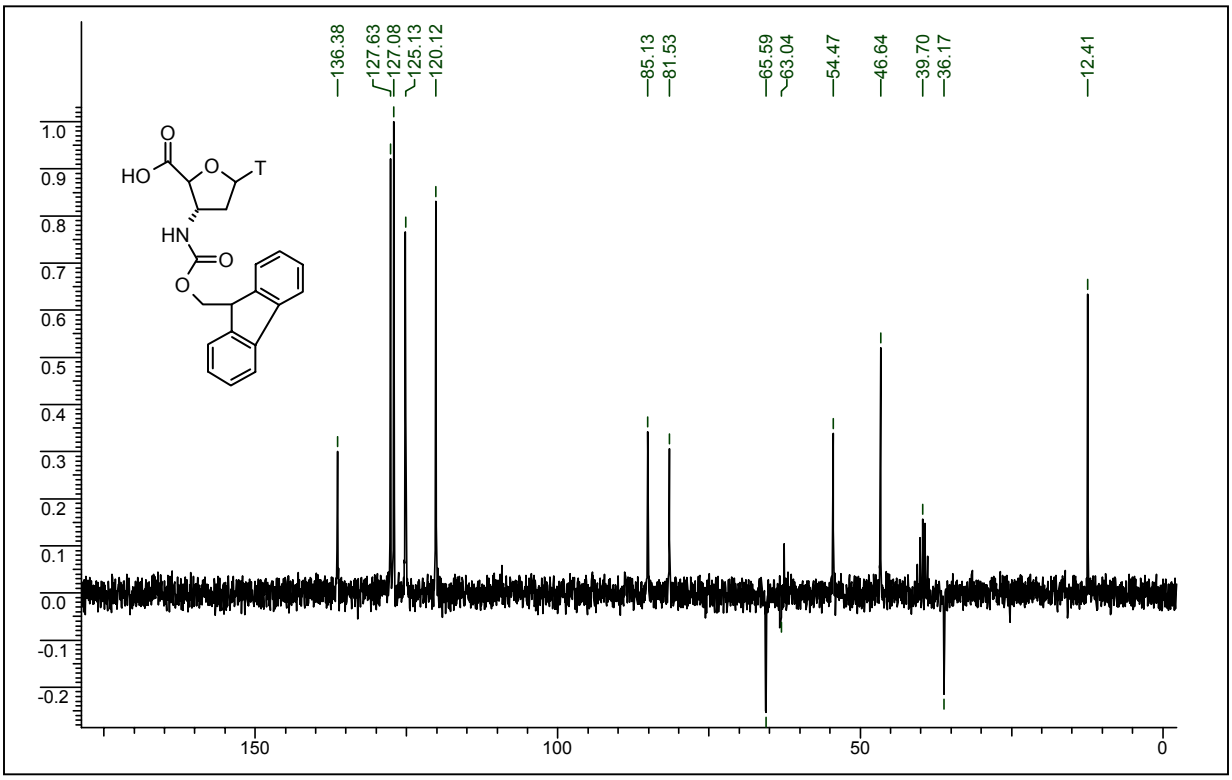


Figure **S4**: DEPT NMR of **8**

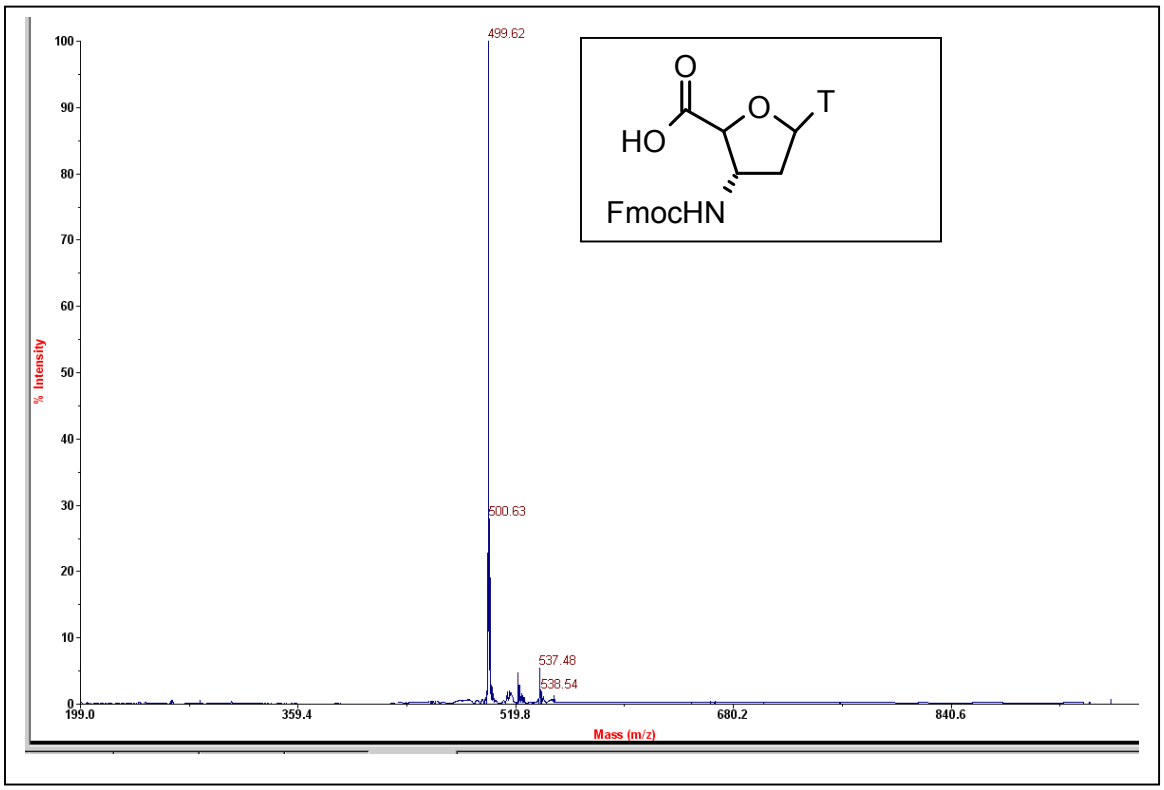


Figure **S5**: Mass spectrum of **8**

Oligomer Synthesis:

The amino acids used for solid-phase oligonucleotides(ON) synthesis were Fmoc -Lys(Boc)-OH, Fmoc-Pro-OH, Fmoc-Met-OH and Fmoc-Sar-OH. All the N^{α}-Fmoc L-amino acids (except Fmoc-Sar-OH) and resins were obtained from Novabiochem (Fmoc= 9-fluorenyl methoxy-carbonyl). Sarcosine (*N*-methylglycine) was obtained from Lancaster chemicals and Fmoc protection was done by standard methods. The amounts of isolated ONs were determined spectrophotometrically at 260 nm wavelength. The molar absorptivities were assumed to be identical to the Deoxyribonucleosides.

Figure **S6**: Solid-Phase synthesis of amino acid backbone modified ONs and structure of ONs synthesized.

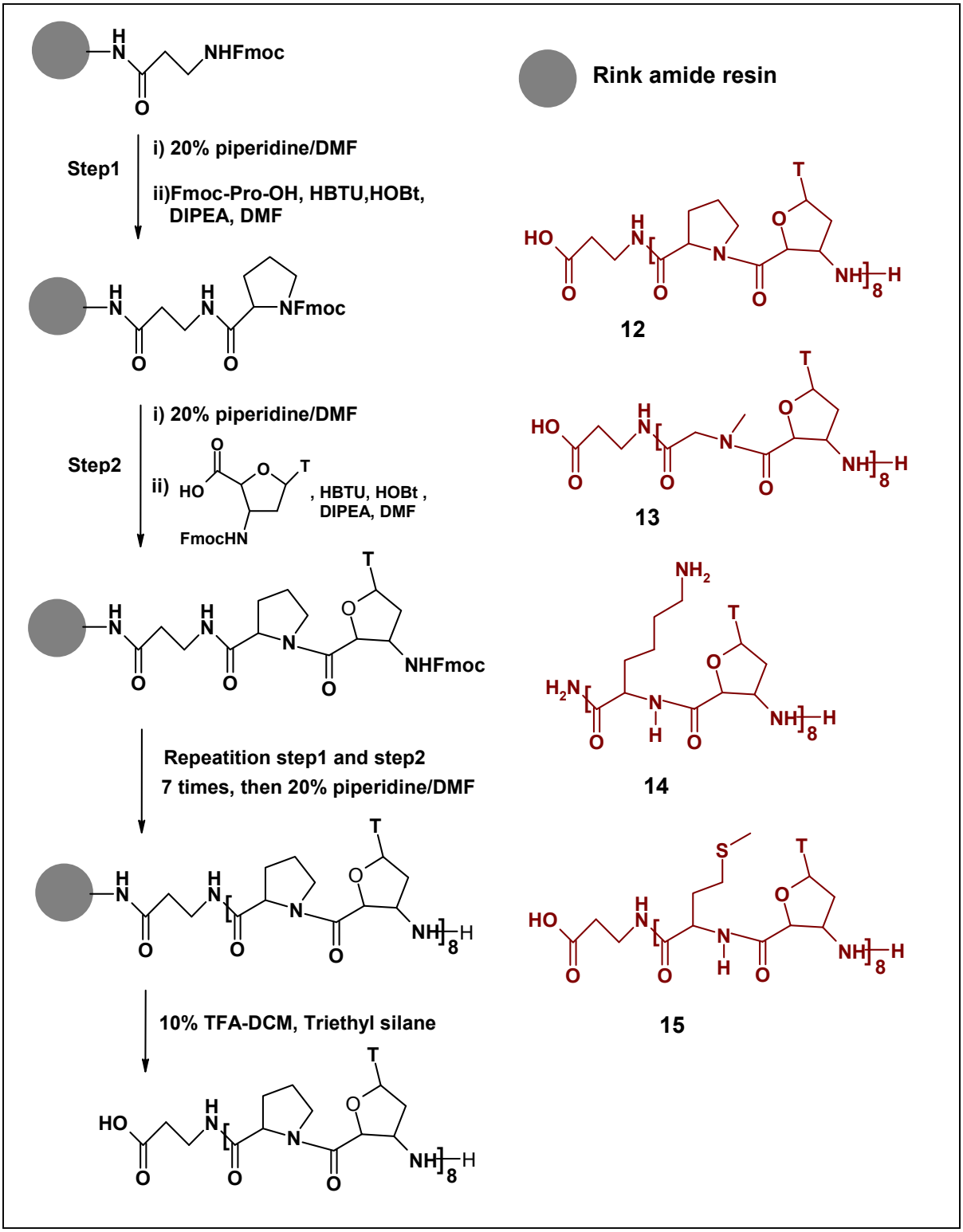


Table **S1**: Amino-acid backbone modified ONs and their HPLC and MALDI-TOF mass characterization.

Entry	Sequence	Mol Formula	HPLC t_R (min)	Mass calculated /observed
1	β-ala-(Pro-**t**)$_8$-H **12**	$C_{123}H_{145}N_{33}O_{42}$	12.7	2763.7/ 2785.6(+Na^+) 2801.6(+K^+)
2	β-ala-(Sar-**t**)$_8$-H **13**	$C_{107}H_{134}N_{33}O_{42}$	12.3	2555.4/ 2578.8(+Na^+) 2594.9(+K^+)
3	(Lys-**t**)$_8$-H **14**	$C_{128}H_{186}N_{40}O_{41}$	11.7	2941.1/ 2941.7
4	β-ala-(Met-**t**)$_8$-H **15**	$C_{123}H_{147}N_{33}O_{42}S_8$	12.6	3036.3/ 3037.0

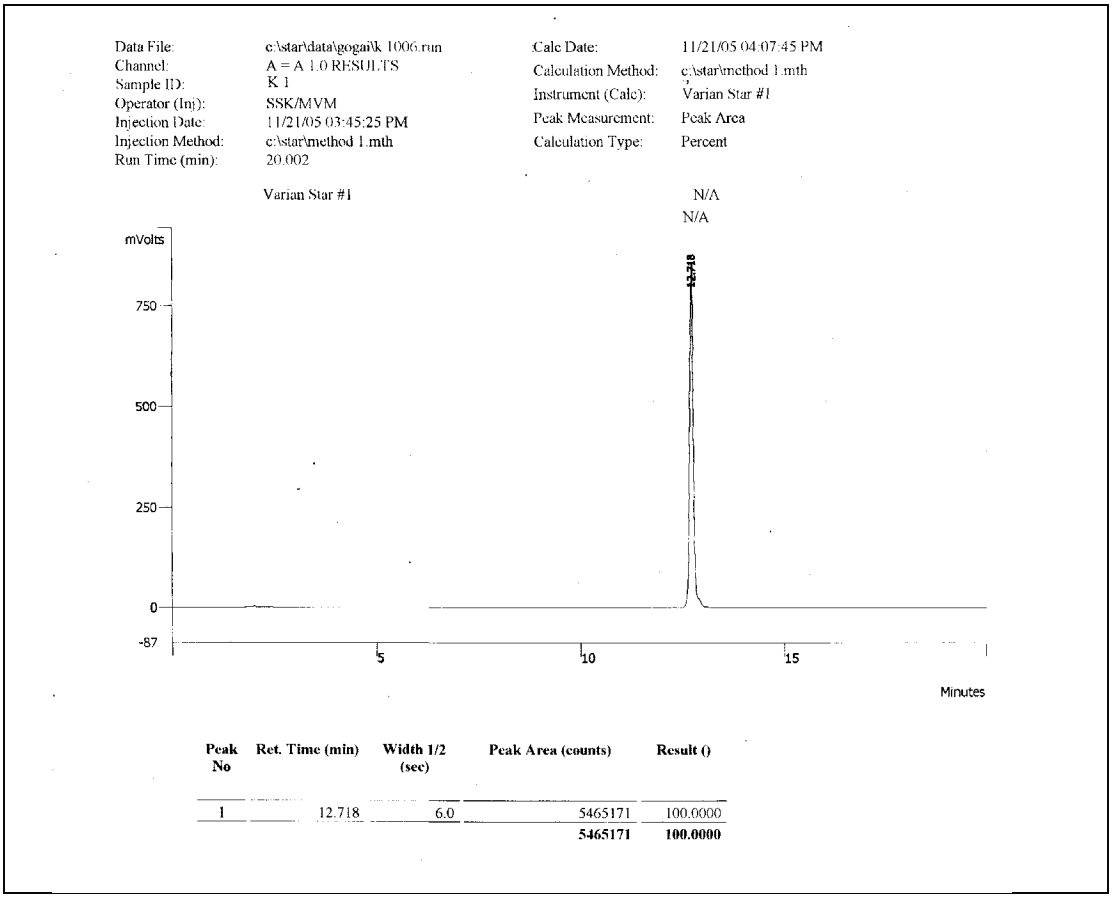
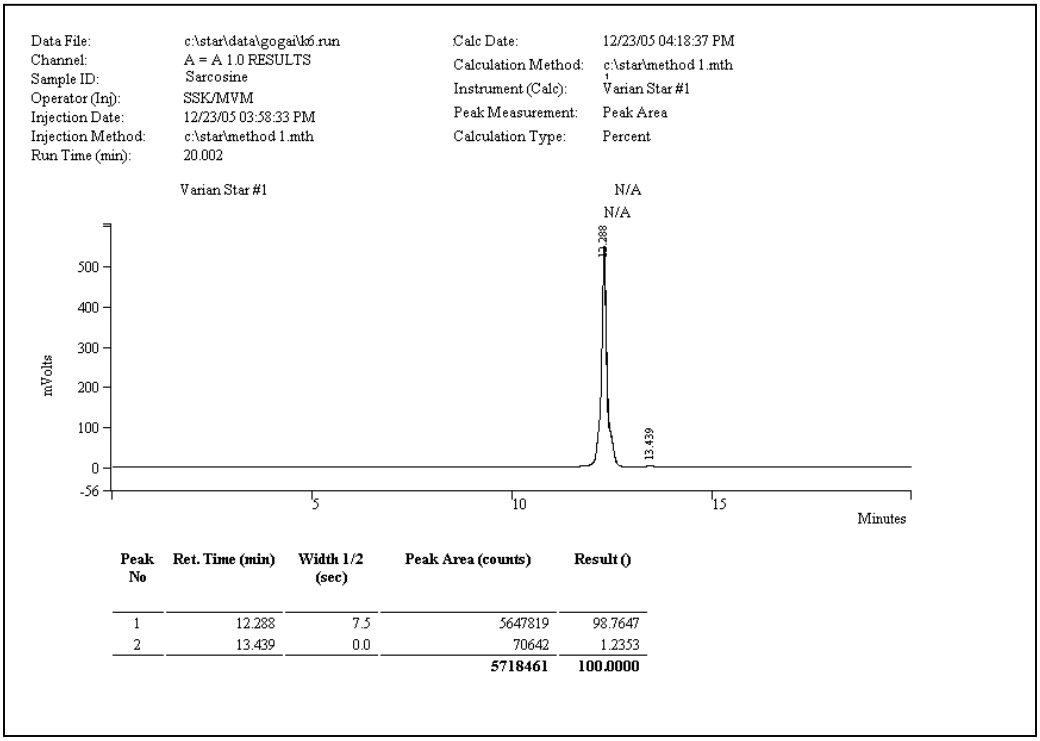
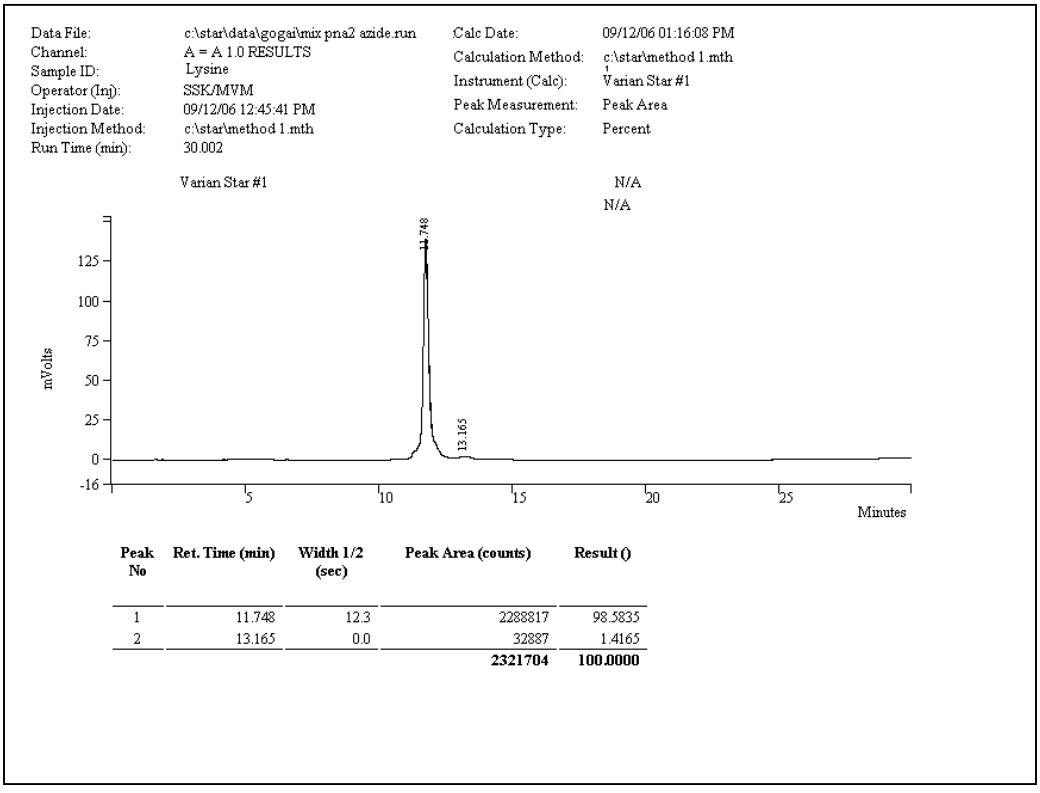


Figure **S7**: RP-HPLC of **12**



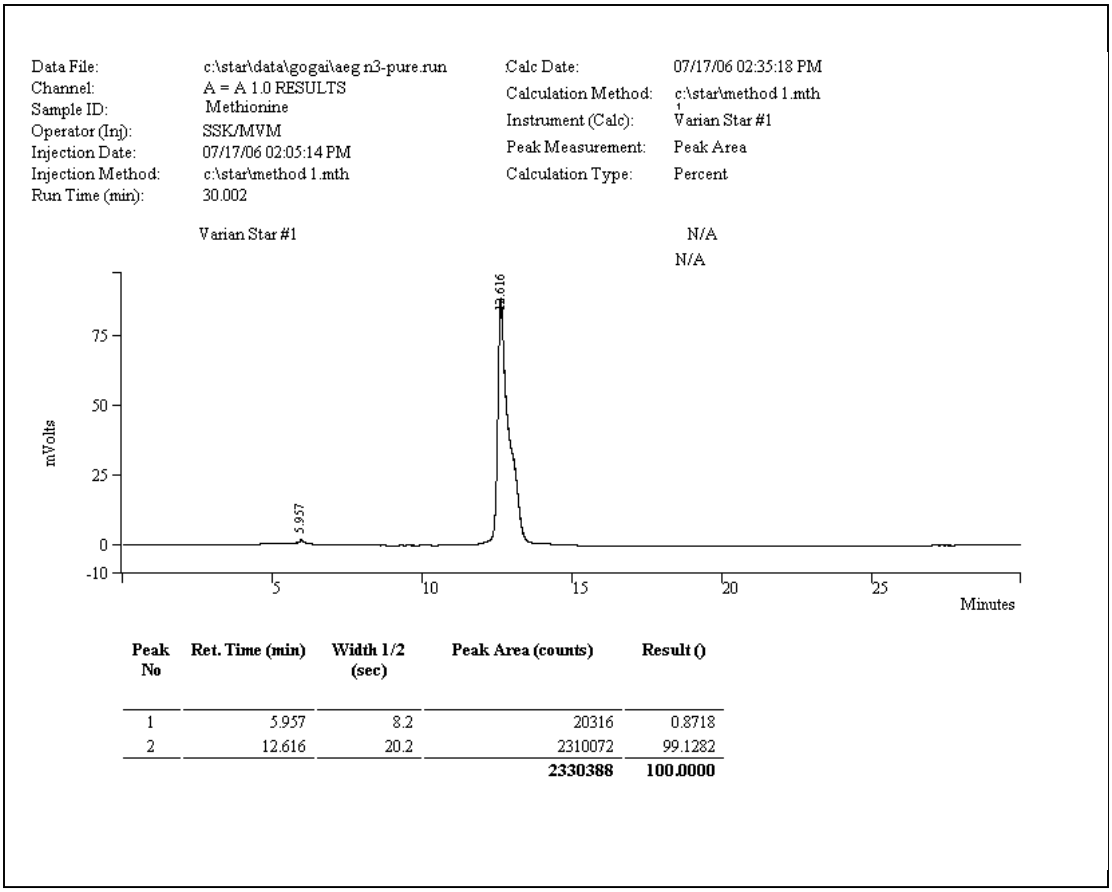
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	12.288	7.5	5647819	98.7647
2	13.439	0.0	70642	1.2353
			5718461	**100.0000**

Figure **S8**: RP-HPLC of **13**



Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	11.748	12.3	2288817	98.5835
2	13.165	0.0	32887	1.4165
			2321704	**100.0000**

Figure **S9**: RP-HPLC of **14**



Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	5.957	8.2	20316	0.8718
2	12.616	20.2	2310072	99.1282
			2330388	**100.0000**

Figure **S10**: RP-HPLC of **15**

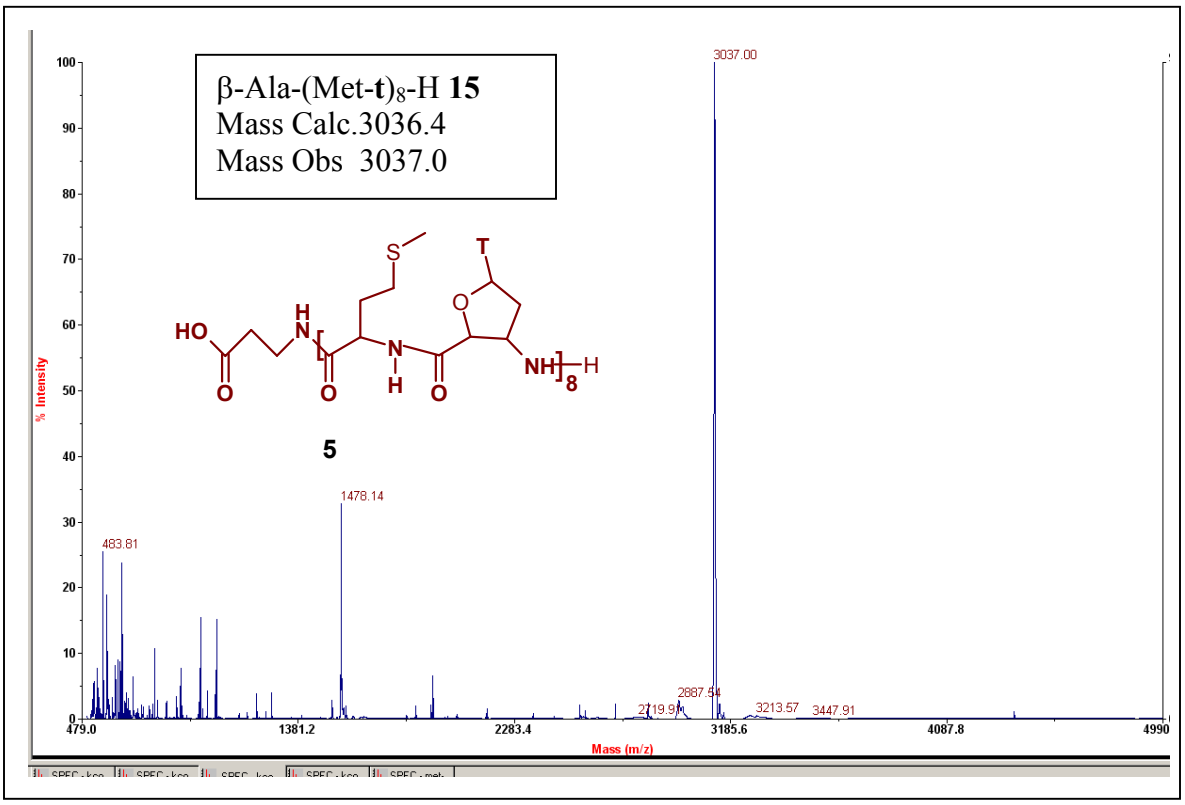


Figure **S11**: MALDI-TOF mass of **15**

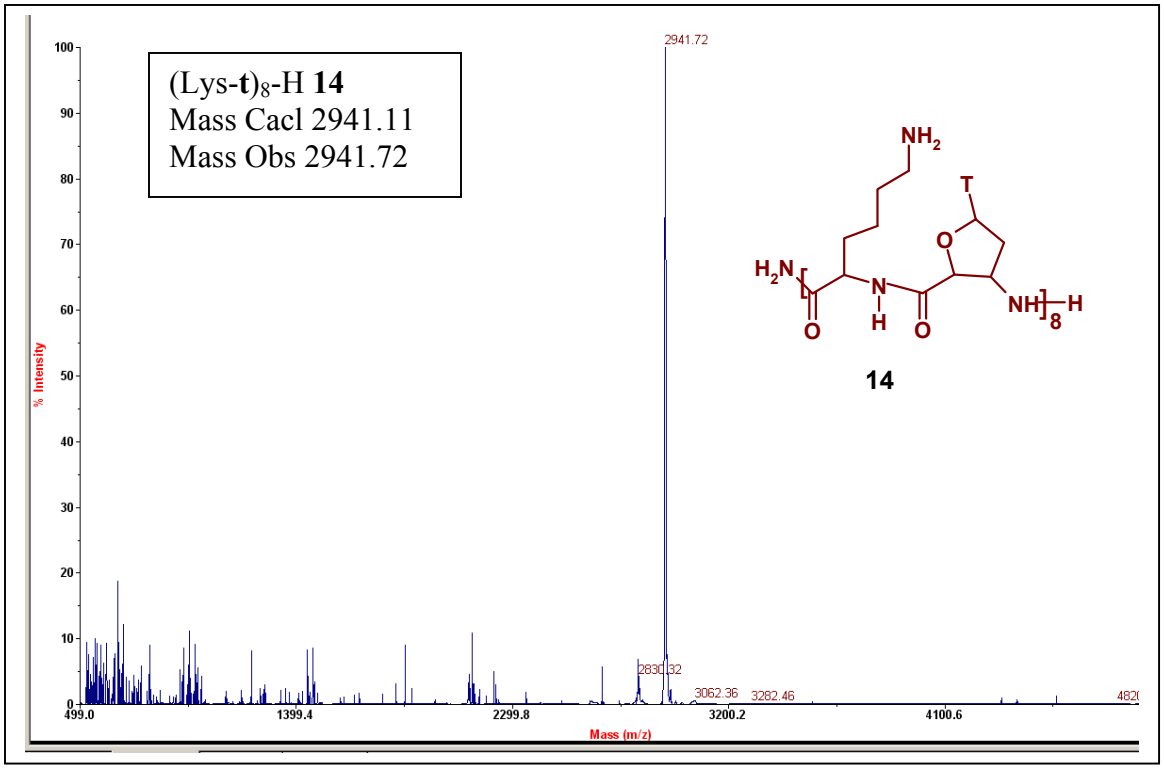


Figure **S12**: MALDI-TOF mass of **14**

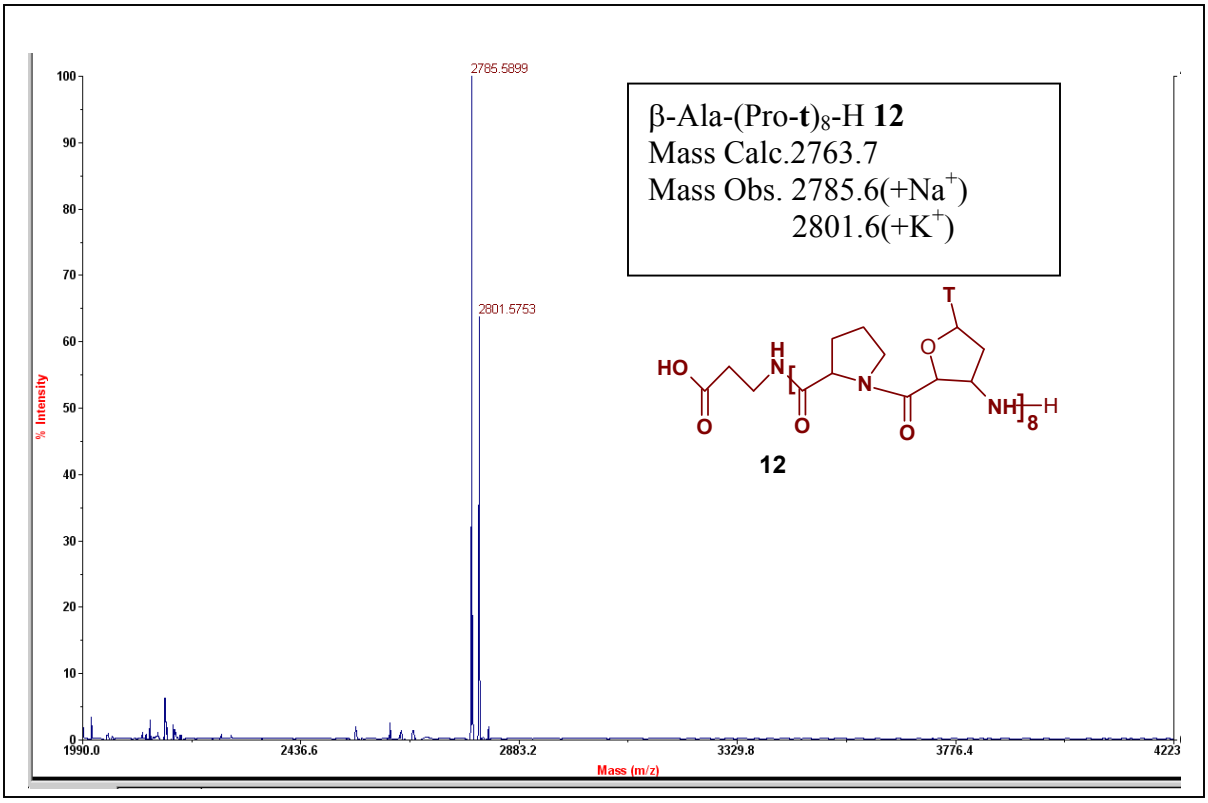


Figure **S13**: MALDI-TOF mass of **12**

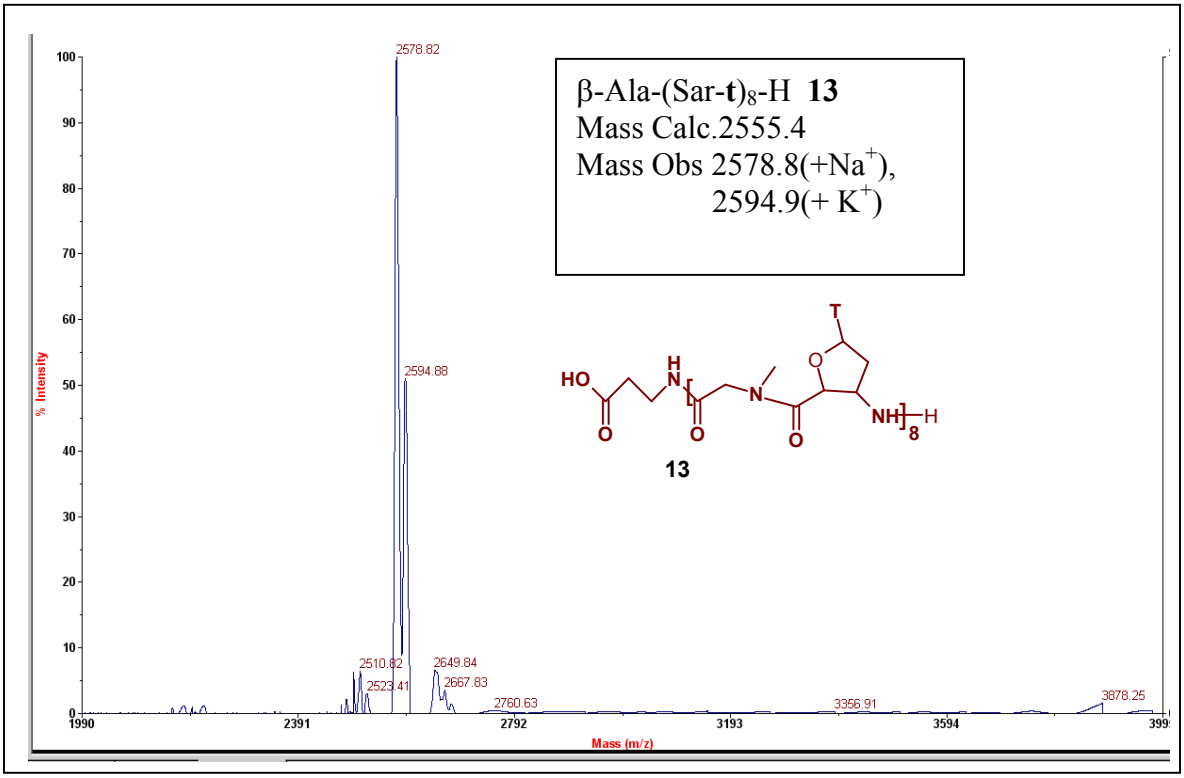


Figure **S14**: MALDI-TOF mass of **13**

UV-*Tm* studies: The complementary DNA and RNA oligomers were synthesized on an Applied Biosystems 3900 DNA Synthesizer. The concentration was calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases. The experiments were performed in 1-2 μM concentrations. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature. Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Perkin-Elmer Lambda 35 spectrophotometer scanning from 5 to 85°C at a ramp rate of 0.5°C per minute. The data were processed using Microcal Origin 6.1 and T_m values derived from the derivative curves. All values are an average of at least 3 experiments and accurate to within ± 0.5ºC.

Table **S2:** *T*m(ºC) values of ONs: DNA / Mismatch DNA

Entry	ON Sequence	DNA **9**	DNA **16**	Δ*T*m (mismatch)
1	5' **TTTTTTTT** 3' **11**	17.8		
2	β-Ala-(Pro-**t**)$_8$-H **12**	49.0	37.5	11.5
3	β-Ala-(Sar-**t**)$_8$-H **13**	49.1	35.3	13.8
4	(Lys-**t**)$_8$-H **14**	56.3	38.5	17.8
5	β-Ala-(Met-**t**)$_8$-H **15**	47.3	32.9	14.4

DNA **9** 5' GCAAAAAAAACG 3' DNA **16** 5' GCAAAATAAACG 3'(Single base mismatch sequence)

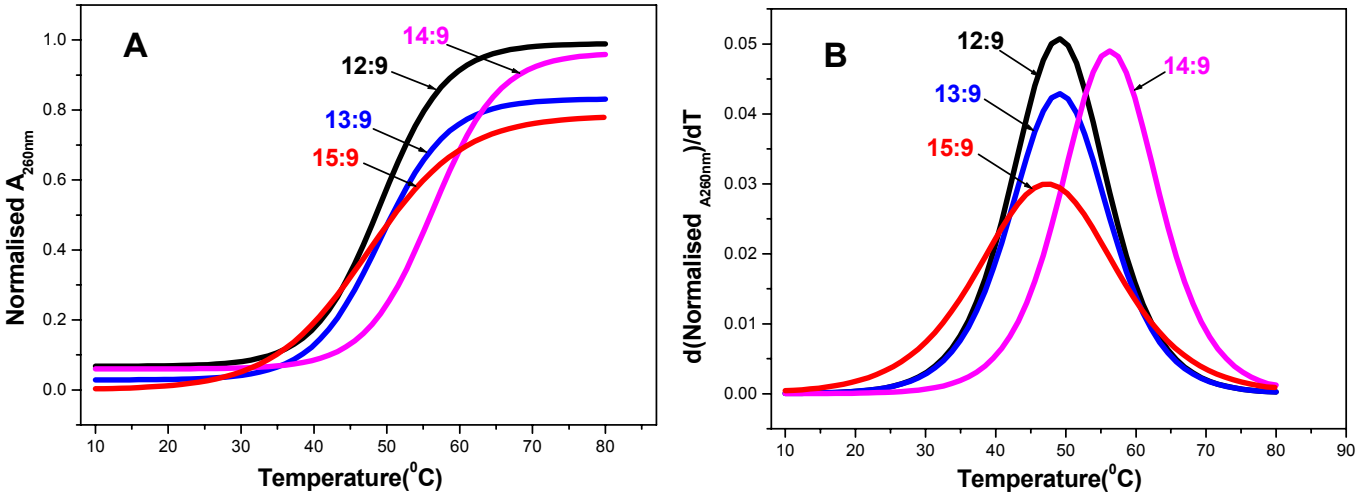


Figure **S15:** **A.**UV-melting Curves of **12**, 1**3, 14** and **15** with complementary DNA **9** and **B.** Corresponding first derivative Curves.

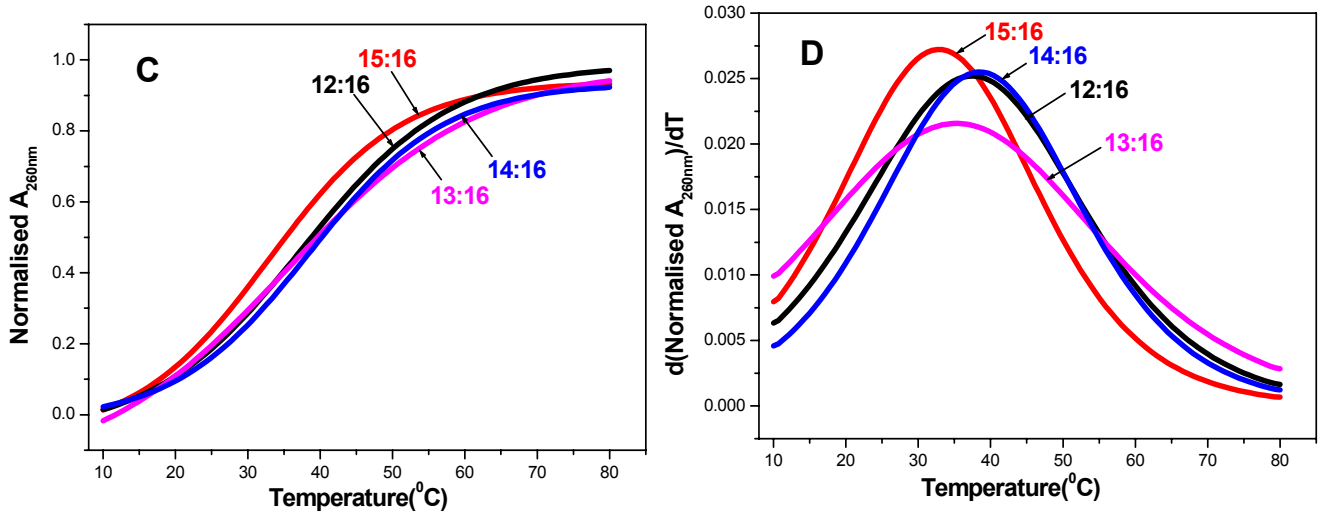


Figure **S16: C.** UV-melting Curves of **12**, **13, 14** and **15** with mismatch DNA **16** and **D**. Corresponding first derivative Curves.

Table **S3:** *T*m(ºC) values of ONs: RNA / Mismatch RNA

Entry	ON Sequence	RNA**10**	RNA **17**	Δ*T*m (mismatch)
1	5' **TTTTTTTT** 3' **11**	15.6		
2	β-Ala-(Pro-**t**)$_8$-H **12**	60.8	44.6	16.2
3	β-Ala-(Sar-**t**)$_8$-H **13**	61.6	45.6	16.0
4	(Lys-**t**)$_8$-H **14**	69.0	55.5	13.5
5	β-Ala-(Met-**t**)$_8$-H **15**	57.3	41.2	16.1

RNA **10** **r**(5' GCAAAAAAAACG 3') RNA **17** **r**(5' GCAAAAUAAACG 3')(single base mismatch sequence).

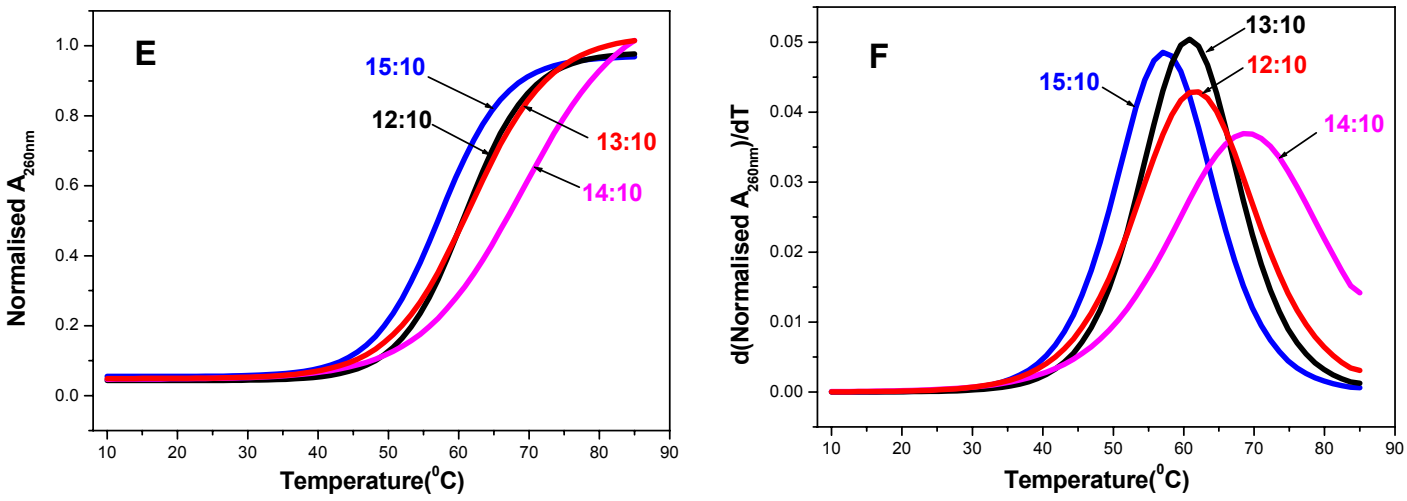


Figure **S17**: **E**.UV-melting Curves of **12**, **13, 14** and **15** with complementary RNA **10** and **F.** Corresponding first derivative Curves.

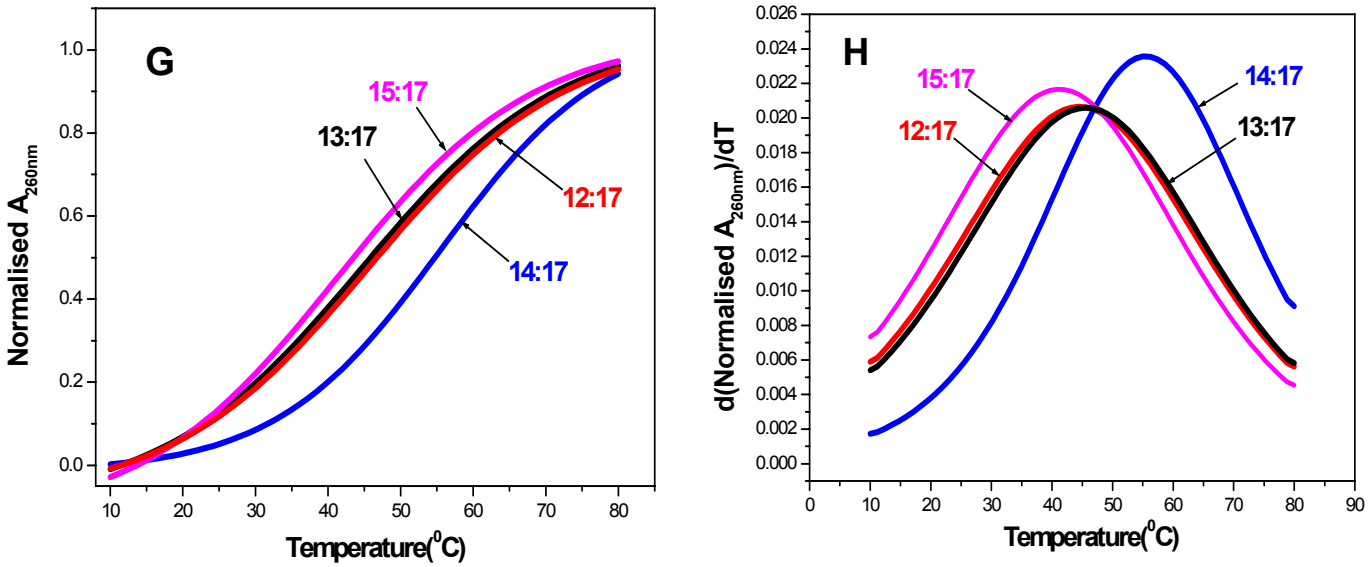


Figure **S18**: **G**.UV-melting Curves of **12**, **13, 14** and **15** with mismatch RNA **17** and **H.** Corresponding first derivative Curves.

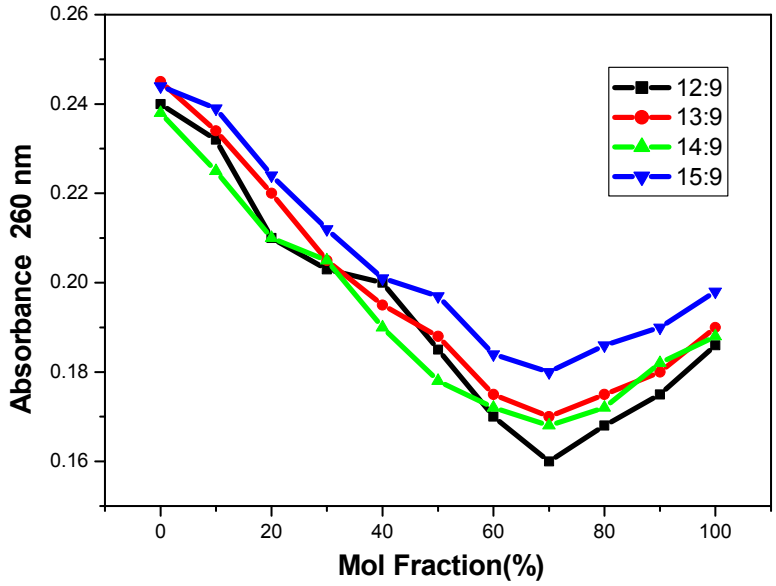


Figure **S19**: Jobs plot of modified ONs **12**, **13**, **14** and **15** with complementary DNA **9** at 25ºC in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA.

Reference: P. Job, *Ann. Chim.*,1928, **9**, 113-203.

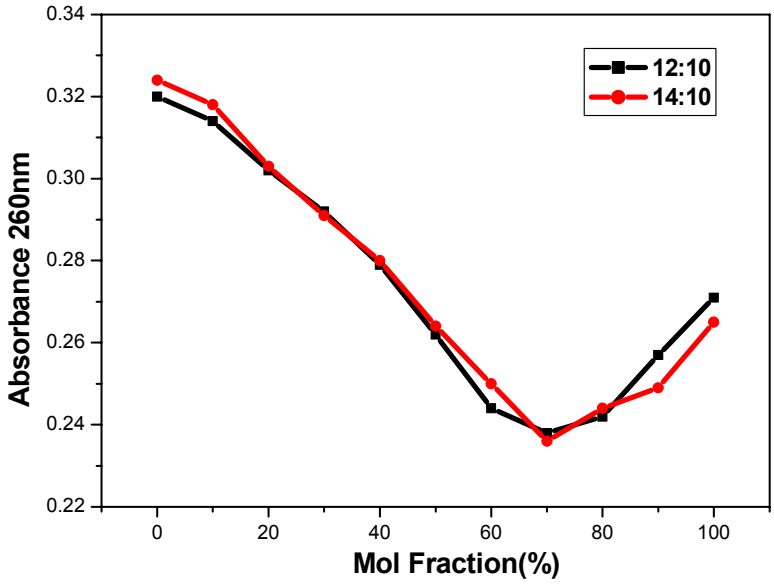


Figure **S20**: Jobs plot of modified ONs **12** and **14** with complementary RNA **10** at 25ºC in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA.

Reference: P. Job, *Ann. Chim.*, 1928, **9**, 113-203.

Circular dichorism studies:

The circular dichorism spectra were obtained in a JASCO J-715 spectrometer. The experiment were performed in 2µM concentration and The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature. The experiments were performed at 25ºC.The data were processed using Microcal Origin 6.1.

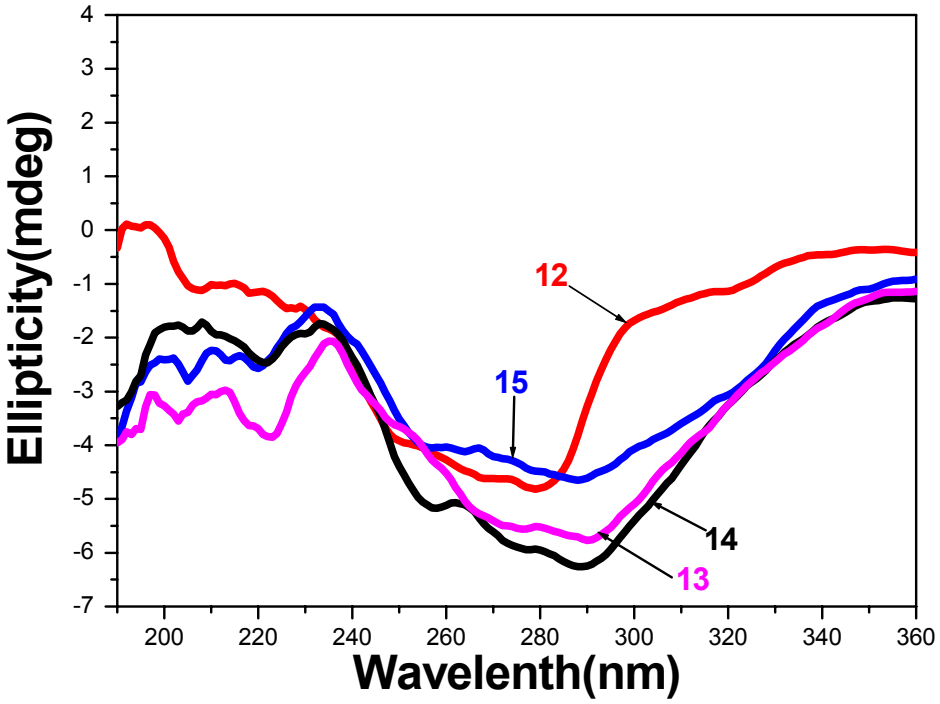


Figure **S21**: CD Curves of ONs **12**, **13**, **14** and **15** in water

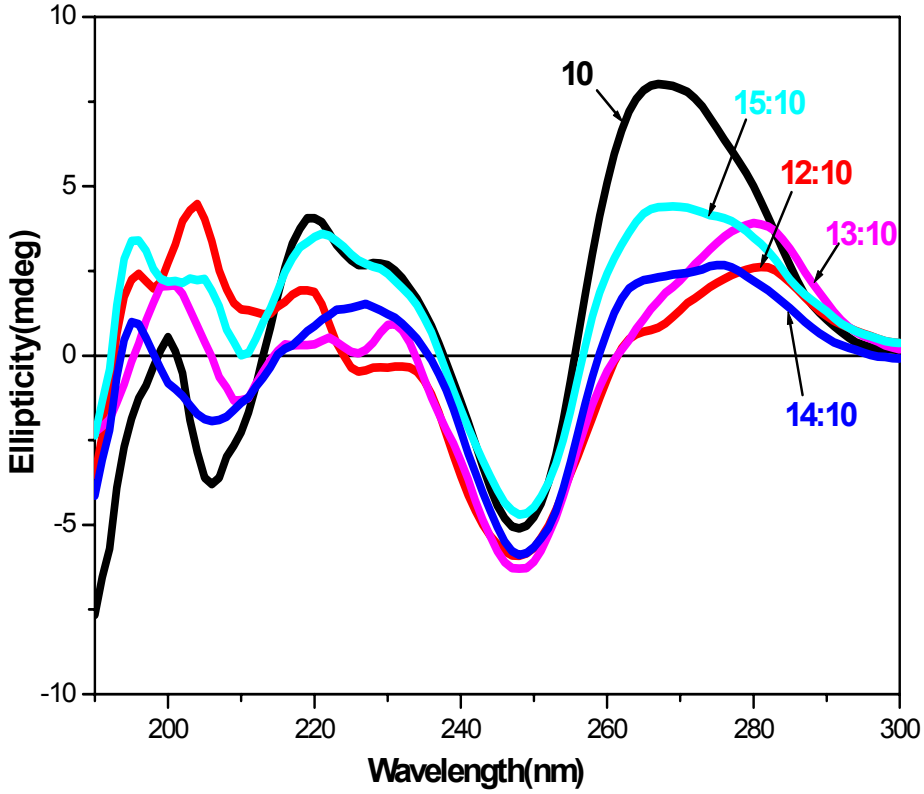


Figure **S23:** CD curves of **12**, **13**, **14** and **15** with complementary RNA **10**