SUPPORTING INFORMATION FOR

Cyclodextrin - peptide conjugates for sequence specific DNA binding

Yara Ruiz García, Jan Zelenka, Y. Vladimir Pabon, Abhishek Iyer, Miloš Buděšínský, Tomas Kraus, C. I. Edvard Smith and Annemieke Madder*

TABLE OF CONTENTS

1. Materials and Methods

1.1 Peptide Synthesis

Automated peptide syntheses were performed on a fully-automated SYRO Multiple Peptide Synthesizer robot, equipped with a vortex unit for the 24-reactor block (MultiSynTech GmbH). Reactions were open to the atmosphere and executed at ambient temperature.

1.2 Analysis & Equipment

Reversed-Phase HPLC analysis and purification was performed on an Agilent 1100 Series instrument with diode array detector (at 214, 254, 280, 310, 360 nm), equipped with a Phenomenex Luna C18(2) 100 Å column (250 x 4.6 mm, 5 µm, at 35 °C) or a Phenomenex Jupiter C4 300 Å column (250 x 4.6 mm, 5 µm, at 35 °C). Linear gradient elutions were performed from 0 to 100% buffer B in 15 minutes by using a binary solvent system composed of buffer A: 0.1% TFA in H_2O and B: MeCN) with a flow of 1.0 mL/min at 35°C. ESI-MS spectra were recorded on a quadrupole ion trap LC mass spectrometer (Thermo Finnigan MAT LCQ), equipped with electrospray ionization. MeOH/H_2O (4/1 ± 0.1 % formic acid) was used as carrier solution. All reported data were collected in the positive mode, at 250 °C. MALDI-TOF-MS data were acquired on an Applied Biosystems Voyager-DE STR Biospectrometry Workstation, equipped with a high performance nitrogen laser (337 nm). All spectra were recorded in the positive and reflector mode, with delayed extraction. LC-TIC-MS data (reversed phase) were recorded on an Agilent 1100 Series instrument with diode array detector (at 214, 254, 280, 310, 360 nm), equipped with Phenomenex Kinetex C18 100 Å (150 x 4.6 mm, 5 µm, at 35 °C), hyphenated to an Agilent ESI-single quadrupole MS detector type VL. Mass detection operated in the positive mode. Linear gradient elutions were performed by using a binary solvent system composed of buffer A: 0.1% formic acid in H_2O and B: MeCN) with a 1.0 mL/min flow. Pre- and post-flushing is included in the LC-MS and RPHPLC results. A solution of 4-5 mg α-cyano-4-hydroxycinnamic acid in 500 µL MeCN, 490 µL mQ, 10 µL 1M ammoniumcitrate, 1 µL TFA was used as matrix for MALDI-TOF-MS. Peptide concentrations were determined by ^{1}H-NMR ERETIC (Electronic REference To access In vivo Concentrations)[1] with a high-resolution 500 MHz NMR-spectrometer. DNA concentrations were determined with the Trinean dropsense multichannel spectrophotometer.

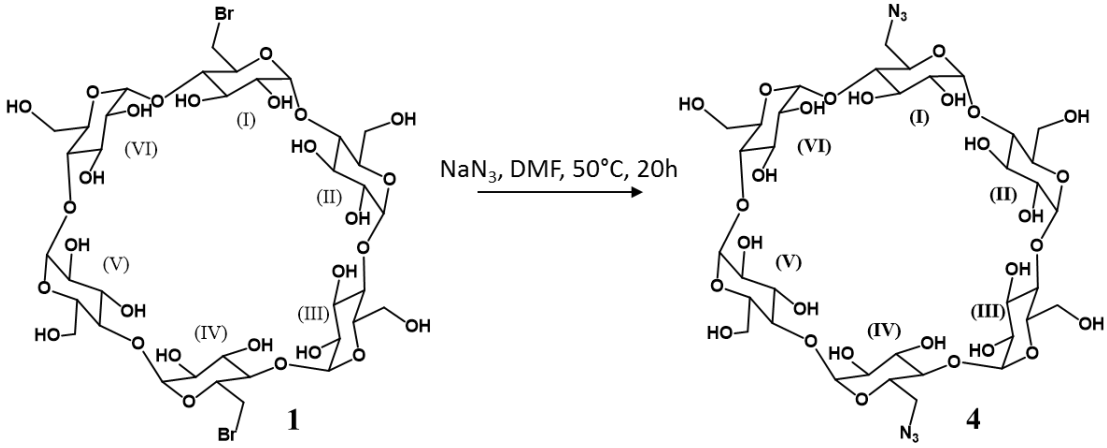
2. Synthesis of scaffolds

2.1 Cyclodextrin template molecules:

General experimental procedures: Preparative reversed–phase (RP) chromatography was carried out using medium pressure columns containing C-18 modified silica (Phenomenex Luna, 15 µm). Thin-layer (TLC) and reversed-phase thin-layer chromatography (RPTLC) were performed with precoated Silica Gel 60F and RP-18 F plates (E. Merck) respectively, which were visualized by spraying with an aqueous solution of phosphomolybdic acid containing 5% of H_2SO_4 and heating. All chemicals used were commercially available. Compounds **1, 2** and **3** were prepared according to the known procedure[2–4]. Satisfactory elemental analysis could not be obtained for hydrophilic compounds **4, 5** and **6** unless variable numbers of water molecules

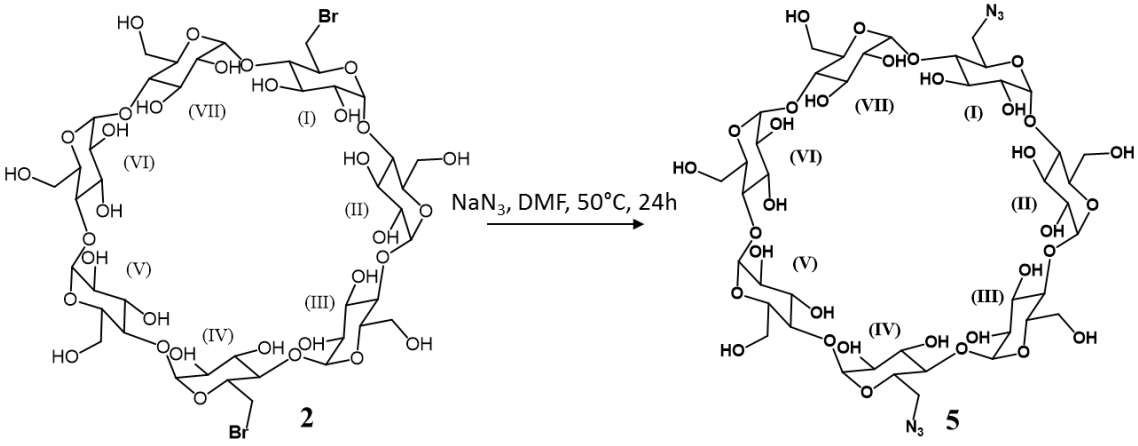
were taken into account. Thus, calculations based on weights of these compounds (molarity, yield) are related to the hydrated molecules.

6^{I},6^{IV}-Dideoxy-6^{I},6^{IV}-diazido-α-cyclodextrin **(4)**.



To a solution of 6^{I},6^{IV}-dideoxy-6^{I},6^{IV}-dibromo-α-cyclodextrin trihydrate **(1)** (0.367 g, 0.318 mmol) in DMF (5 ml) sodium azide (0.130 g, 2 mmol) was added. The mixture was stirred at 50 °C for 20 h under argon atmosphere and then the solvent was evaporated under reduced pressure. The solid material was dissolved in a warm mixture of methanol-water (8:2, 5 ml) and charged onto RP column. Gradient elution with a mixture of methanol-water from 1:9 to 4:6 followed with evaporation of selected fractions and drying of the isolated material over phosphorus pentoxide *in vacuo* gave compound **4** as white amorphous material (0.305 g, 90%, calcd for dihydrate). R_f = 0.3 (RP-TLC, methanol-water 1:1). Elemental analysis: Calcd for $C_{36}H_{58}N_6O_{28}$. $2H_2O$: C, 40.83, H 5.90; N 7.94; Found: C, 40.70, H, 5.77, N 7.63. ESI-MS calcd for $[M+Na]^+$: 1045.3; found: 1045.3. ^{1}H NMR (see Table S1); ^{13}C NMR (see Table **S2**).

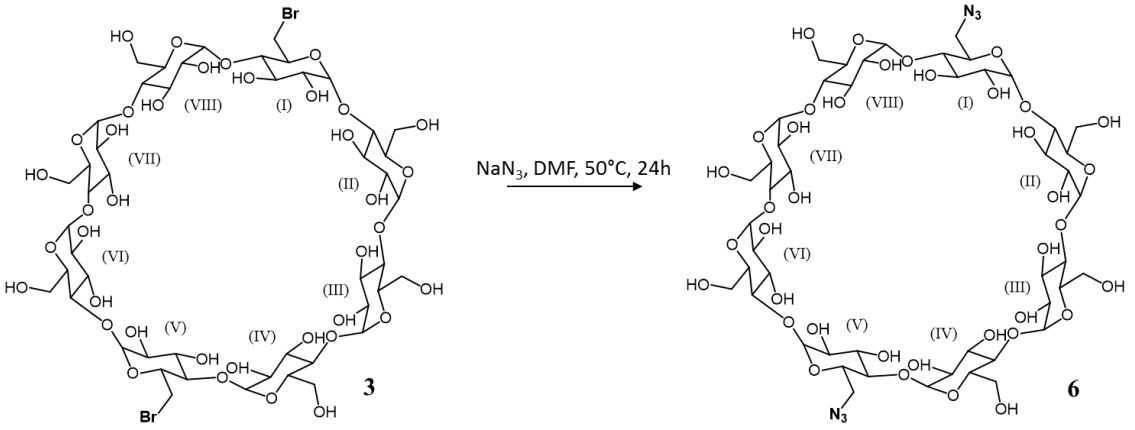
6^{I},6^{IV}-Dideoxy-6^{I},6^{IV}-diazido-β-cyclodextrin **(5)**.



To a solution of 6^{I},6^{IV}-dideoxy-6^{I},6^{IV}-dibromo-β-cyclodextrin tetrahydrate **(2)** (0.533 g, 0.399 mmol) in DMF (15 mL) sodium azide (0.156 g, 2.40 mmol) was added and the mixture was stirred at 50 °C for 24 h under argon atmosphere. Then the solvent was evaporated under reduced pressure to about a half of the volume and acetone (90 mL) was added under stirring. The precipitate, which formed upon addition of acetone, was collected on a glass frit, washed

with acetone and dried *in vacuo*. Subsequently, it was dissolved in a warm mixture of methanol-water (9:1, 9 ml) and charged onto RP column. Gradient elution with a mixture of methanol-water from 1:9 to 4:6 followed with evaporation of selected fractions gave compound **5** (0.444 g, 87 %, calcd as pentahydrate) as white amorphous material after drying the isolated material over phosphorus pentoxide *in vacuo*. R_f = 0.3 (RP-TLC, methanol-water 1:1). Elemental analysis: Calcd for $C_{42}H_{68}N_6O_{33}.5H_2O$: C, 39.56; H, 6.17; N, 6.59; Found: C, 39.83, H, 6.15, N 6.14. ESI-MS calcd for $[M+Na]^+$: 1207.4; found 1207.4. ^{1}H NMR (see Table S1); ^{13}C NMR (see Table S2).

$6^I,6^V$-Dideoxy-$6^I,6^V$-diazido-γ-cyclodextrin **(S6)**.

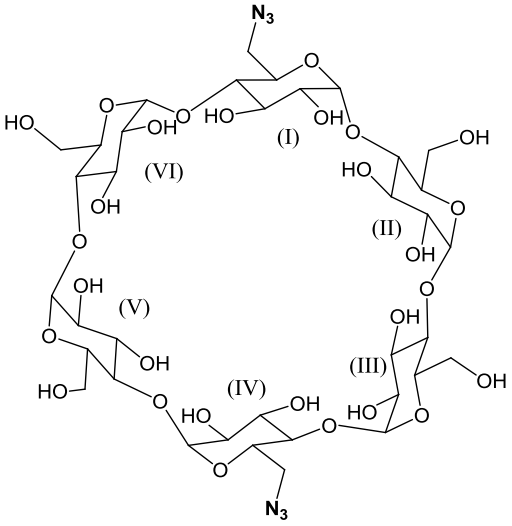


To a solution of $6^I,6^V$-dideoxy-$6^I,6^V$-dibromo-γ-cyclodextrin pentahydrate **(3)** (43 mg, 28.4 µmol) in DMF (1 ml) sodium azide (11 mg, 0.171 mmol) was added. The mixture was stirred at 50 °C for 24 h under argon atmosphere and then solvents were evaporated under reduced pressure. The residue was dissolved in a mixture of methanol-water (8:2, 3 ml) and charged onto RP column (10 g, Phenomenex Luna 15 µm). Gradient elution with a mixture of methanol-water from 1:9 to 4:6 followed with evaporation of selected fractions and drying of the isolated material over phosphorus pentoxide *in vacuo* gave compound **6** as white amorphous material (32 mg, 77%, calcd for hexahydrate). R_f = 0.4 (RP-TLC, methanol-water 4:6). Elemental analysis: Calcd for $C_{48}H_{78}N_6O_{38}.6H_2O$: C, 39.62, H 6.23, N 5.78; Found: C, 39.73, H, 5.91, N 5.39. HR-MALDI-MS calcd for $[M+Na]^+$: 1369.4248; found 1369.4266. ^{1}H NMR (see Table S1); ^{13}C NMR (see Table S2).

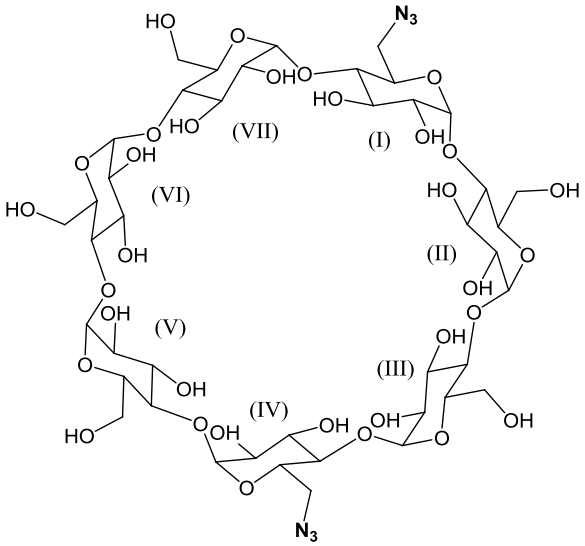
2.2 NMR spectra of cyclodextrin template molecules

The NMR spectra of compounds **4, 5** and **6** were measured on Bruker AVANCE-600 instrument (^{1}H at 600.13 MHz and ^{13}C at 150.9 MHz) with a cryoprobe in CD_3OD at 25 °C. Spectra were referenced to solvent peak (using δ_H (CHD_2OD) = 3.31 ppm; δ_C(CD_3OD) = 49.0 ppm). Structural assignment of proton and carbon signals was achieved combining 1D-^{1}H and ^{13}C-spectra with homonuclear 2D-H,H-COSY, 2D-H,H-TOCSY, 2D-H,H-ROESY and heteronuclear 2D-H,C-HSQC and 2D-H,C-HMBC spectra. The ^{1}H and ^{13}C NMR data are summarized in Tables S1 and S2. Complete sequential structural assignment of proton and carbon signals has been achieved only for "symmetrical" α-cyclodextrin derivative **4**. For β- and γ-cyclodextrin derivatives **5** and **6**, respectively, only partial structural assignment was possible due to the poor separation of signals. The sequential positions of glucose units in Table S1 and S2 are labeled with numbers I

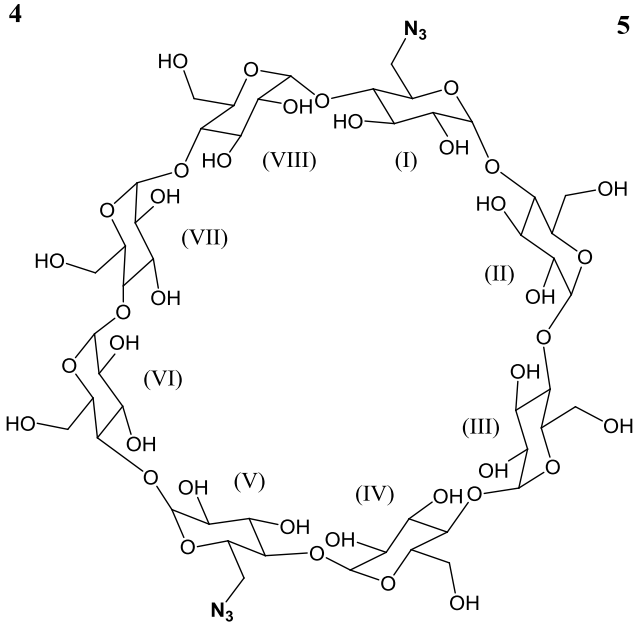
to VIII as it is shown in Figure S1. All observed coupling constants *J*(H,H) of ring protons of glucose units (not given in Table 1) showed the values characteristic for $^4C_1$ conformation (*J*(1,2) = 3.5 – 4.0 Hz, *J*(2,3) ~ *J*(3,4) ~ *J*(4,5) = 8.5 – 10.0 Hz).



4



5



6

Figure S1 Structures and glucose residue numbering of cyclodextrin derivatives

Table 1. Proton chemical shifts of compounds **4, 5** and **6** in CD_3OD at 600.13 MHz

Comp.	Glucose residue	H-1 (d)	H-2 (dd)	H-3 (dd)	H-4 (dd)	H-5 (ddd)	H-6a (dd)	H-6b (dd)
4	I, IV	4.971	3.509	3.916	3.447	3.961	3.728	3.612
	II, V	4.940	3.494	3.935	3.517	3.821	~3.882 (2H)	
	III, VI	4.907	3.494	3.911	3.469	3.793	~3.833 (2H)	

5[a]	I	4.994	3.53-3.50 (7H)	3.88-3.83 (7H)	3.456	3.891 (2H)	3.713	3.590
	IV	4.994			3.450		3.713	3.580
	II	4.975			3.428	3.76-3.71 (5H)	3.88 – 3.83 (10H)	
	III	4.968			3.424			
	V	4.962			3.51 (3H)			
	VI	4.943						
	VII	4.937						
6[a]	I, V	5.032	3.528	3.826	3.424	3.878	3.695	3.568
	II, VI	5.013	3.516	3.839	3.510	3.742	3.80-3.89 (12H)	
	III, VII	5.004	3.516	3.833	3.493	3.738		
	IV, VIII	4.978	3.522	3.817	3.462	3.705		

[a] Chemical shift values in columns may be interchanged (except of H-6a,H-6b).

Table 2. Carbon-13 chemical shifts of compounds **4-6** in CD_3OD at 150.9 MHz

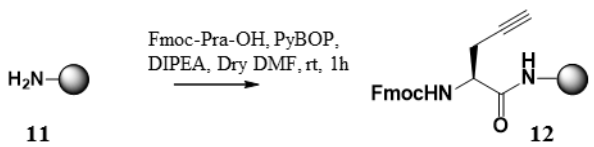
Comp.	Glucose residue	C-1	C-2	C-3	C-4	C-5	C-6
4	I, IV	103.43	73.84	74.92	84.36	72.50	52.82
	II, V	103.67	73.82	75.15	83.38	73.75	61.93
	III, VI	103.71	73.76	75.23	83.40	73.89	61.99
5[a]	I	103.99	74.21	74.82 (2C)	84.26	73.85	52.76
	IV	103.97	74.17 (3C)		84.21	73.81	52.73
	II	103.92		74.77	83.23	73.70	61.95 (2C)
	III	103.91		74.68	83.20	73.67	
	V	103.84	74.14	74.65	83.04 (2C)	73.63	61.87 (2C)
	VI	103.53 (2C)	74.09	74.55		72.42 (2C)	
	VII		74.07	74.52	83.02		61.84
6[a]	I, V	103.44	74.42 (2C)	74.50	83.95	72.49	52.71
	II, VI	103.84		74.56	82.64	73.79	61.99
	III, VII	103.90	74.47	74.62	82.82	73.74	61.92 (2C)
	IV, VIII	104.02	74.50	74.65	82.87	73.95	

[a] Chemical shift values in columns may be interchanged (except of C-6).

3. Peptide synthesis

3.1 Manual coupling protocols

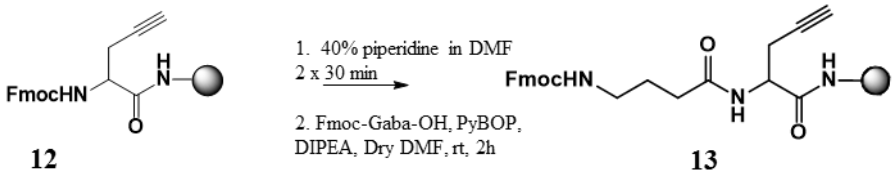
Immobilization of Fmoc-Pra-OH on Rink amide ChemMatrix resin



11 12

To a suspension of resin **11** (100 mg, 0.54 mmol/g) in DMF (dry, 10 mL/g resin), were added Fmoc-Pra-OH (54.5 mg, 0.162 mmol), PyBOP (84.3 mg, 0.162 mmol) and DIPEA (0.056 mL, 0.324 mmol). The mixture was shaken at room temperature for 1 h. After reaction, the resin was washed with DMF, ACN and MeOH. The resin was then capped with acetic anhydride (0.031 mL, 0.324 mmol) and DIPEA (0.056 mL, 0.324 mmol) in dry DMF (2 mL) 2 times for 30 min.

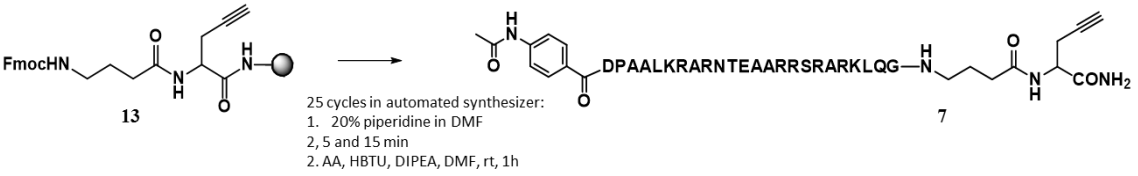
Fmoc deprotection and coupling of Fmoc-GABA-OH

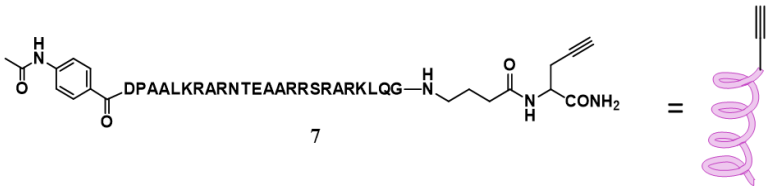


*NHFmoc deprotection of **12***. After an initial DMF washing step, resin **12** (0.054 mmol) was successively treated twice for 30 min with a piperidine solution in DMF (40 % v/v, 2 mL) at ambient temperature, applying intermediate filtration under reduced pressure and washing with DMF, while the final resin is additionally washed with ACN and DCM.

Coupling of Fmoc-GABA-OH. After Fmoc deprotection, Fmoc-GABA-OH (70.2 mg, 0.216 mmol, 0.5 M), PyBOP (112.4 mg, 0.216 mmol, 0.5 M) and DIPEA (0.038 mL, 0.216 mmol, 2 M) were added to a suspension of resin in dry DMF (3 mL). The mixture was shaken at room temperature for 2 h. After the reaction, the resin was washed again with DMF/MeOH/DCM/Et_2O/DMF.

3.2 Automated synthesis of linear peptide Aba-DPAALKRARNTEAARRDRARKLQG-Gaba-Pra





Automated solid phase peptide synthesis was carried out on a Syro synthesizer from Biotage using standard Fmoc/tBu chemistry with HBTU as coupling reagent and 20% piperidine in NMP as deprotection reagent.

Fmoc protected resin **13** (0.054 mmol) was subjected to automated synthesis where solutions of Fmoc-Nα-protected amino acids (0.5 M in NMP) were prepared. Each coupling reaction lasted 1h (amino acids 0.250 mmol; HBTU 0.250 mmol in DMF, 0.5 M; DIPEA 0.250 mmol 2 M) and was followed by Fmoc deprotection with 40% piperidine in NMP. After each reaction the

resin was washed with NMP (9 x). The peptide was then cleaved for the resin and deprotected with a cocktail of TFA/TIS/water = 95: 2.5: 2.5 for 4h. After precipitation in cold ether, the peptide was analyzed by RP-HPLC and MALDI.

3.3 Analysis of linear peptide 7

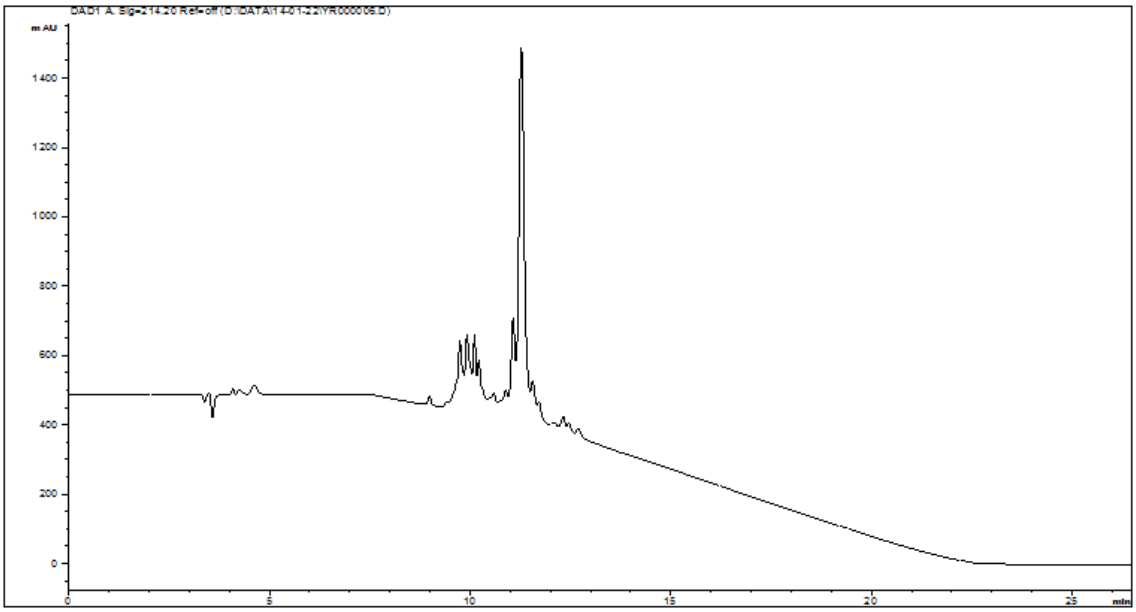


Figure S2: RP-HPLC Chromatogram of crude compound **7** (Jupiter C4, 300Å column using a gradient from 0 to 100 % CH_3CN in 15 minutes)

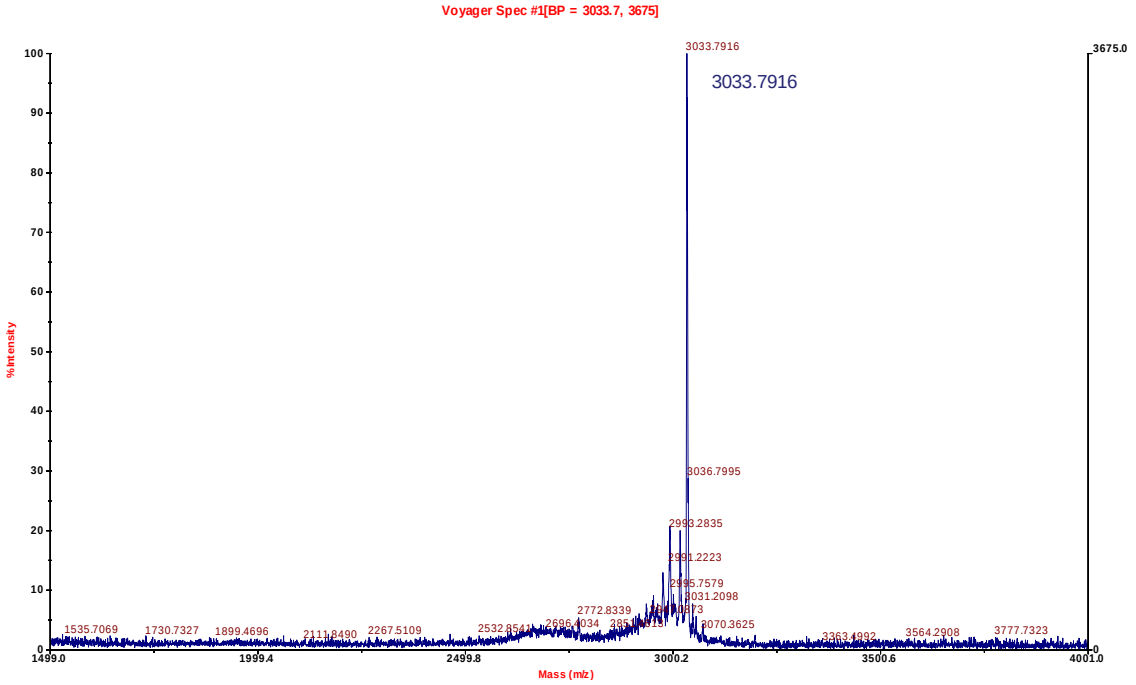
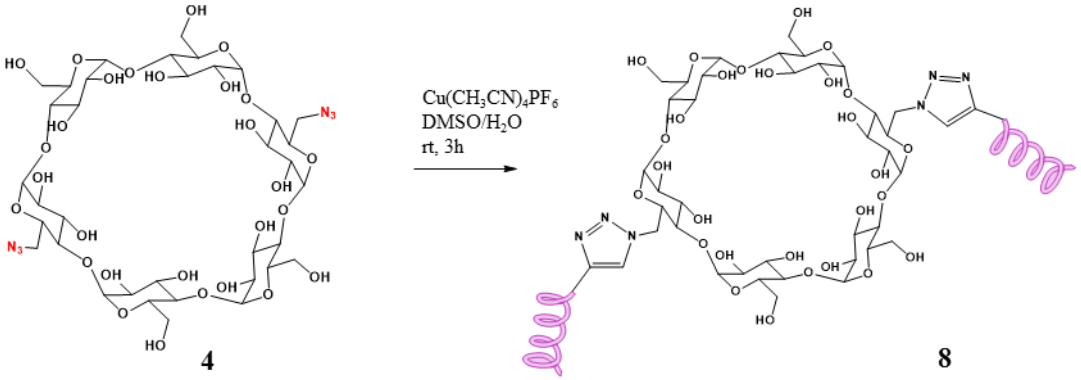


Figure S3: MALDI-TOF Spectrum of crude compound **7**. Calculated m/z (100%) = 3032.68; m/z (91.8%) = 3033.68. Found: 3033.7916 = M + H^+.

4. CuAAC Conjugation

4.1 Peptide-Cyclodextrin conjugation protocol via CuAAC

Using alfa cyclodextrin



Chemical Formula: $C_{292}H_{494}N_{106}O_{100}$
Exact Mass: 7085,68
Molecular Weight: 7089,70

$6^{I},6^{IV}$-Dideoxy-$6^{I},6^{IV}$-diazido-α-cyclodextrin **4** (6.75 mg, 6.6 µmol) was dissolved in 0.2 mL dry DMSO in a 5 mL round bottom flask. Peptide **7** (5 mg, 1.6 µmol) was dissolved in 0.1 mL milliQ water and added it to the reactor. $Cu(CH_3CN)_4PF_6$ (6.15 mg, 16 µmol) was dissolved in 75 µL dry DMSO and added to the reaction mixture. The reaction was stirred for 3h at room temperature under argon. The reaction was monitored by RP-HPLC and purified by fraction collection in RP-HPLC to obtain compound **8**. Fractions were lyophilized and analyzed by RP-HPLC and MALDI-TOF.

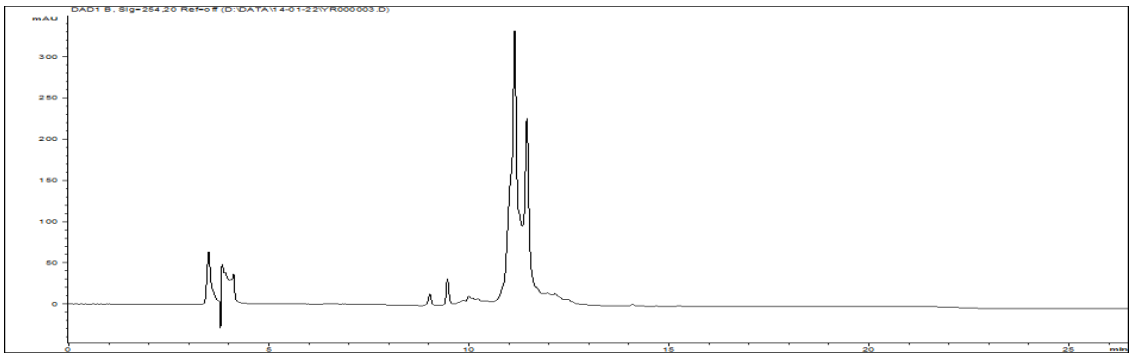
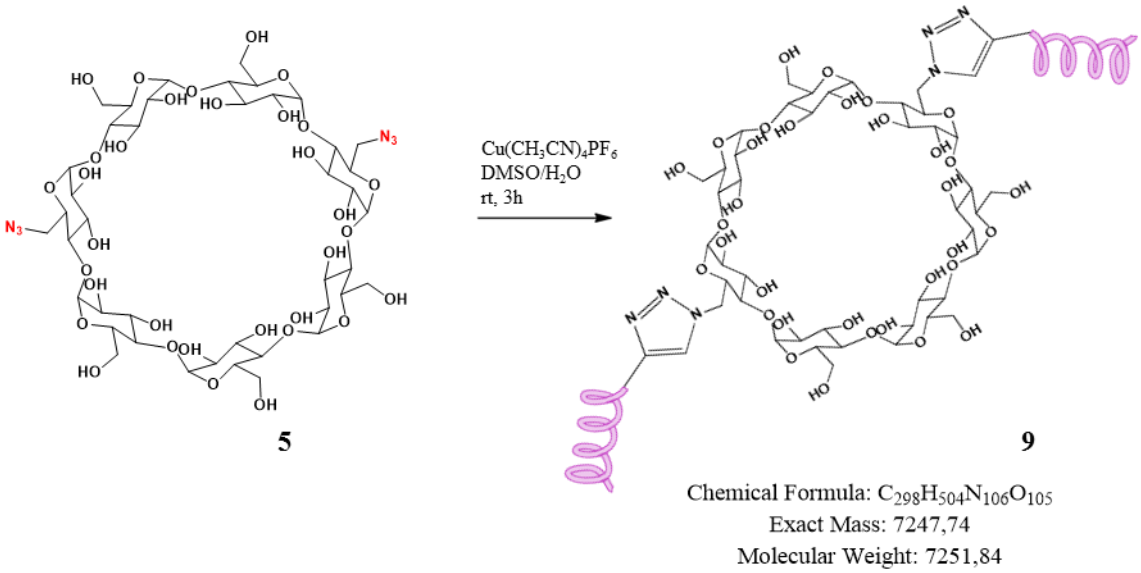


Figure S4: RP-HPLC (0-100% in 15 min on Jupiter C4, 300 Å) of reaction mixture after 3 h.

The chromatogram shows two main peaks corresponding to the unreacted peptide (11.1 min) and desired conjugate (11.5 min). The sharpness of the peak of the conjugate allowed for easy purification by fraction collection.

Using beta cyclodextrin



6^{I},6^{IV}-Dideoxy-6^{I},6^{IV}-diazido-β-cyclodextrin **5** (6.25 mg, 5 µmmol) was dissolved in 0.2 mL dry DMSO in a 5 mL round bottom flask. Peptide **7** (4 mg, 1.3 µmol) was dissolved in 0.1 mL milliQ water and added it to the reactor. $Cu(CH_3CN)_4PF_6$ (4.91 mg, 13 µmol) was dissolved in 75 µL dry DMSO and added to the reaction mixture. The reaction was stirred for 3h at room temperature under argon. The reaction was monitored by RP-HPLC and purified by fraction collection in RP-HPLC to obtain compound **9**. Fractions were lyophilized and analyzed by RP-HPLC and MALDI-TOF.

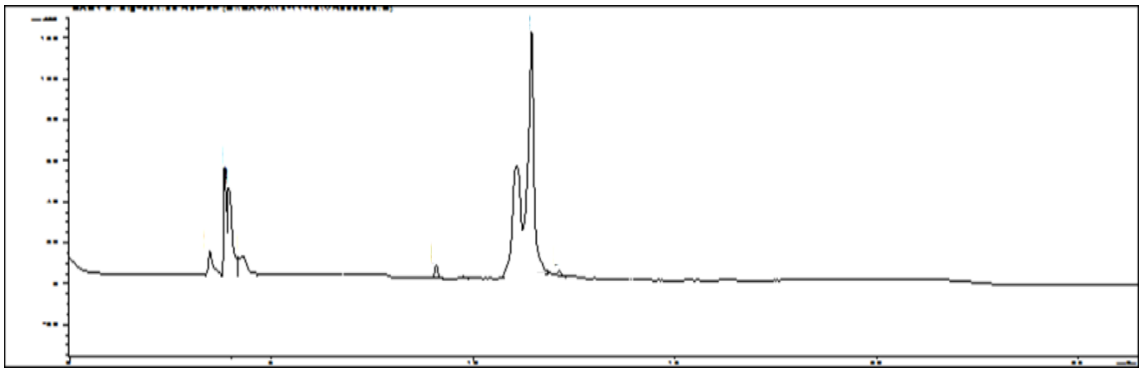
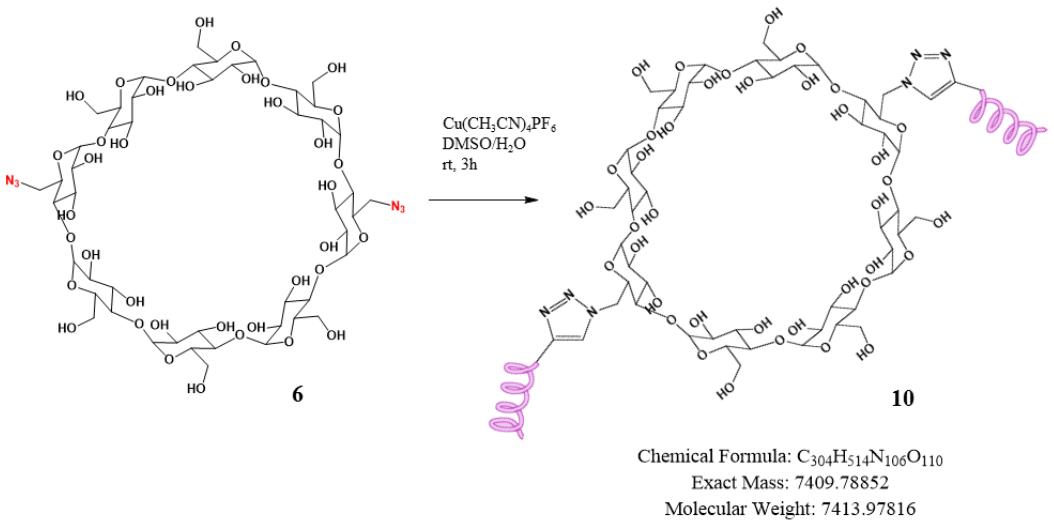


Figure S5: RP-HPLC (0-100% in 15 min on Jupiter C4, 300 Å) of reaction mixture after 3 h

The chromatogram shows two main peaks corresponding to the unreacted peptide (11.1 min) and desired conjugate (11.4 min). The conjugation of the peptide to the β-cyclodextrin derivative proceeded better than to α and γ ones, as the major peak is the desired compound. The sharpness of the peak of the conjugate allowed for easy purification by fraction collection.

Using gamma cyclodextrin



6^{I},6^{V}-Dideoxy-6^{I},6^{V}-diazido-γ-cyclodextrin **6** (8.9 mg, 6.6 µmol) was dissolved in 0.2 mL dry DMSO in a 5 mL round bottom flask. Peptide **7** (5 mg, 1.6 µmol) was dissolved in 0.1 mL milliQ water and added it to the reactor. $Cu(CH_3CN)_4PF_6$ (6.15 mg, 16 µmol) was dissolved in 75 µL dry DMSO and added to the reaction mixture. The reaction was stirred for 3h at room temperature under argon. The reaction was monitored by RP-HPLC and purified by fraction collection in RP-HPLC to obtain compound **10**. Fractions were lyophilized and analyzed by RP-HPLC and MALDI-TOF.

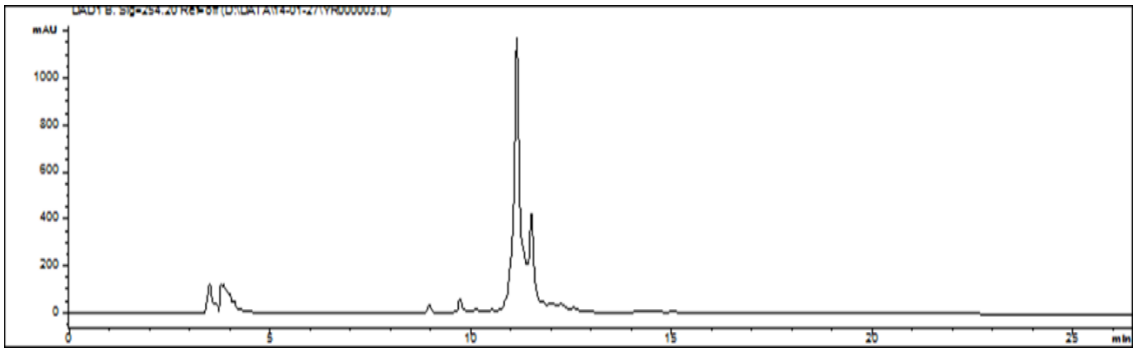


Figure S6: RP-HPLC (0-100% in 15 min on Jupiter C4, 300 Å) of reaction mixture after 3 h

The chromatogram shows two main peaks corresponding to the unreacted peptide (11.1 min) and desired conjugate (11.5 min). The sharpness of the peak of the conjugate allowed for easy purification by fraction collection.

4.2 Analytical data of purified peptide-cyclodextrin conjugates

Compound **8** (Alfa conjugate)

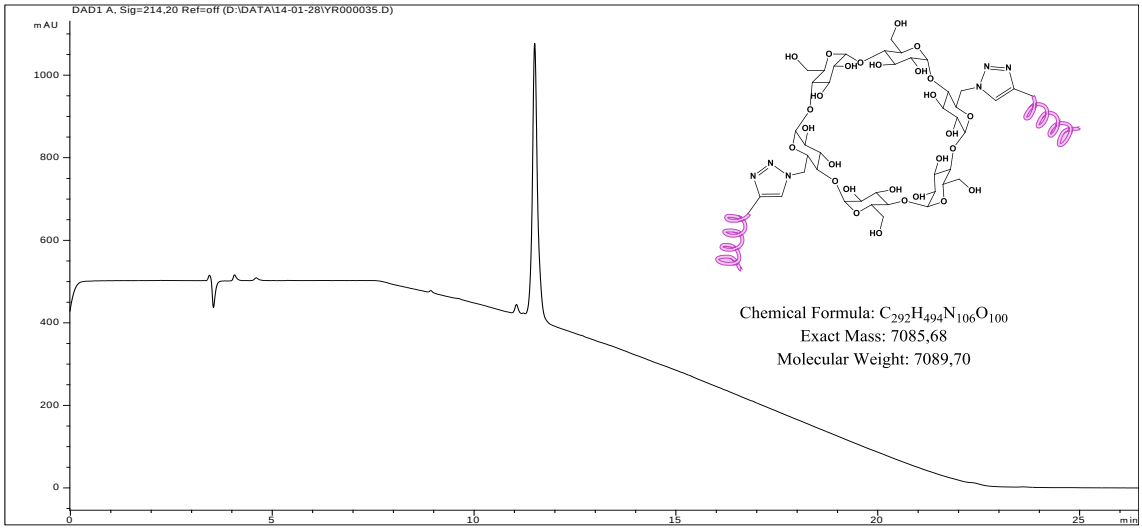


Figure S7: RP-HPLC Chromatogram of pure compound **8** (C4, 300Å column using a gradient from 0 to 100 % CH_3CN in 15 minutes)

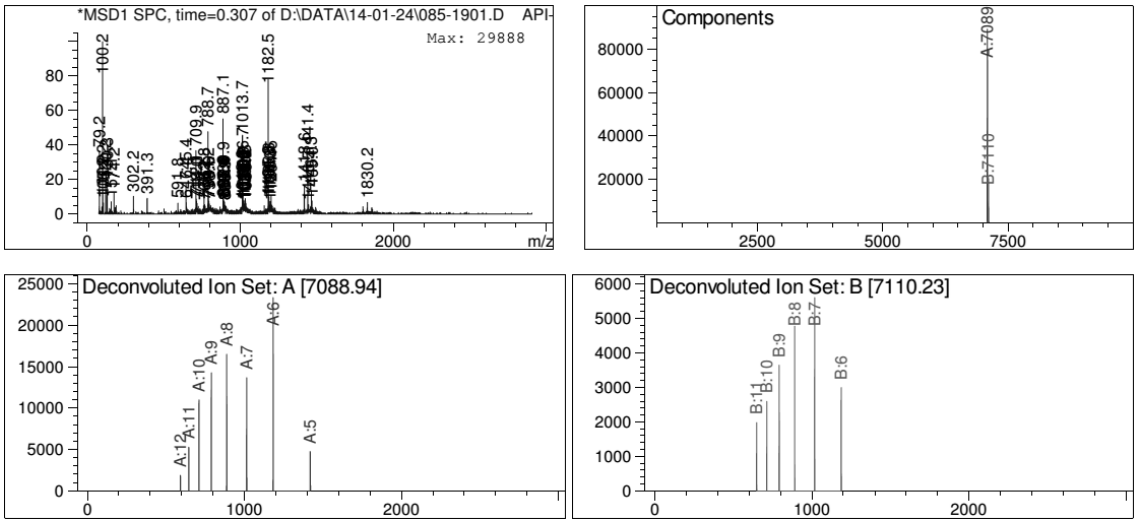


Figure S8: ESI-MS of compound **8**. Calcd. E.M. = 7085.68, M + Na^+ = 7108.68, Obsd. M = 7088.94, M + Na^+ = 7110.23

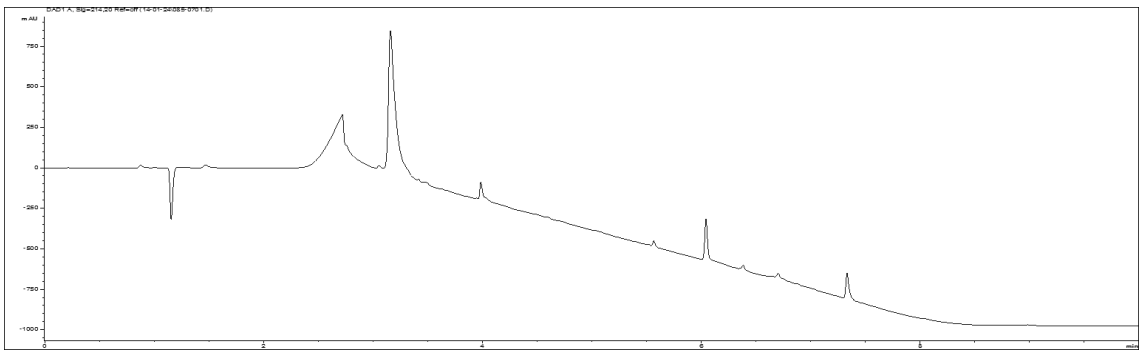


Figure S9: HPLC trace of compound **8** from LC-MS (0-100% ACN in 6 min on Kinetex C18 100 Å, 150 x 2.1 mm, 2.6 µm, at 35 °C)

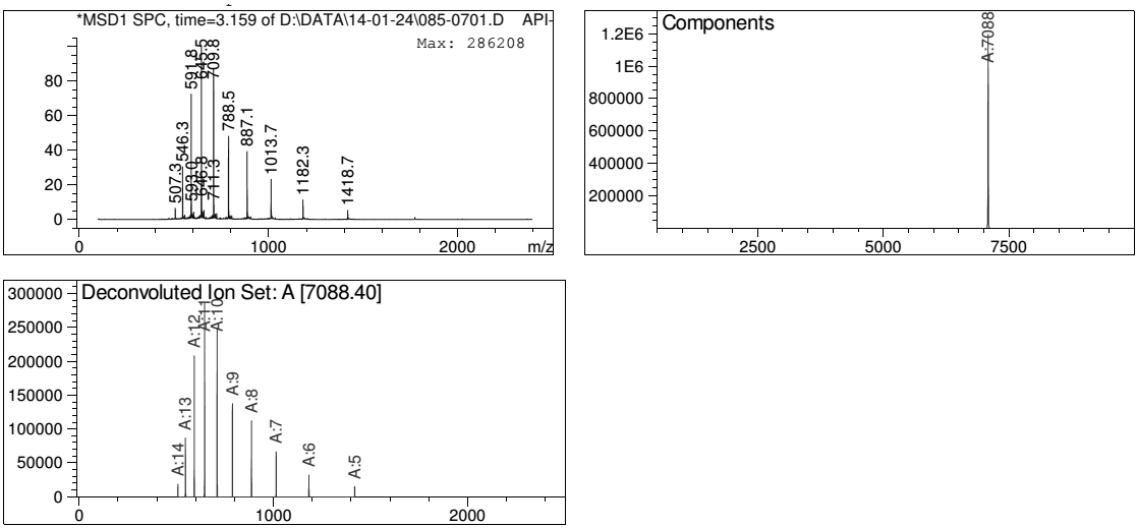


Figure S10: ESI-MS from LC-MS at r.t. = 3.159 min for purified compound **8**. E.M = 7085.68 and deconvoluted mass found 7088.40.

Compound **9** (Beta conjugate)

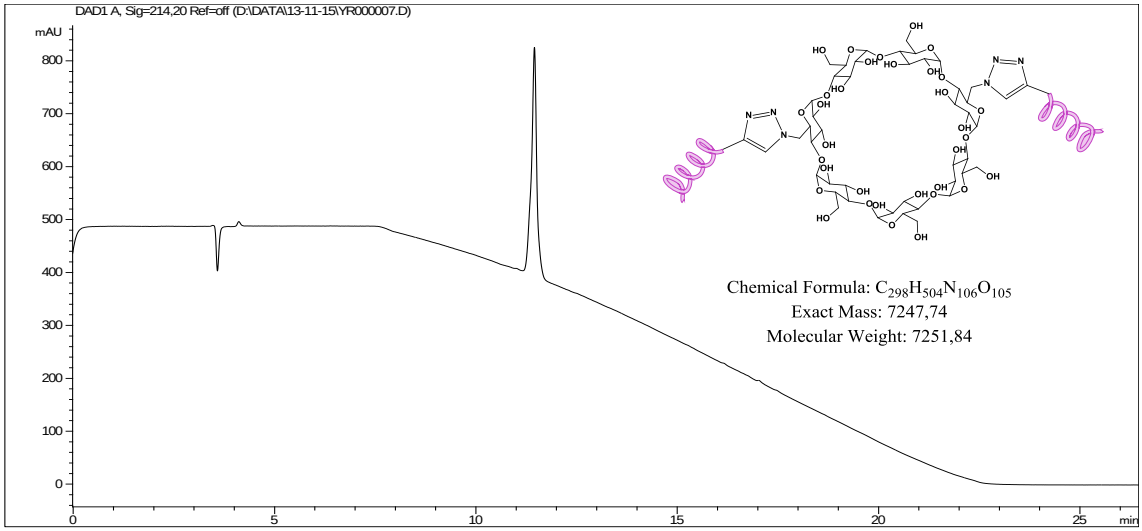


Figure S11: RP-HPLC Chromatogram of pure compound **9** (Jupiter C4, 300Å column using a gradient from 0 to 100 % CH_3CN in 15 minutes)

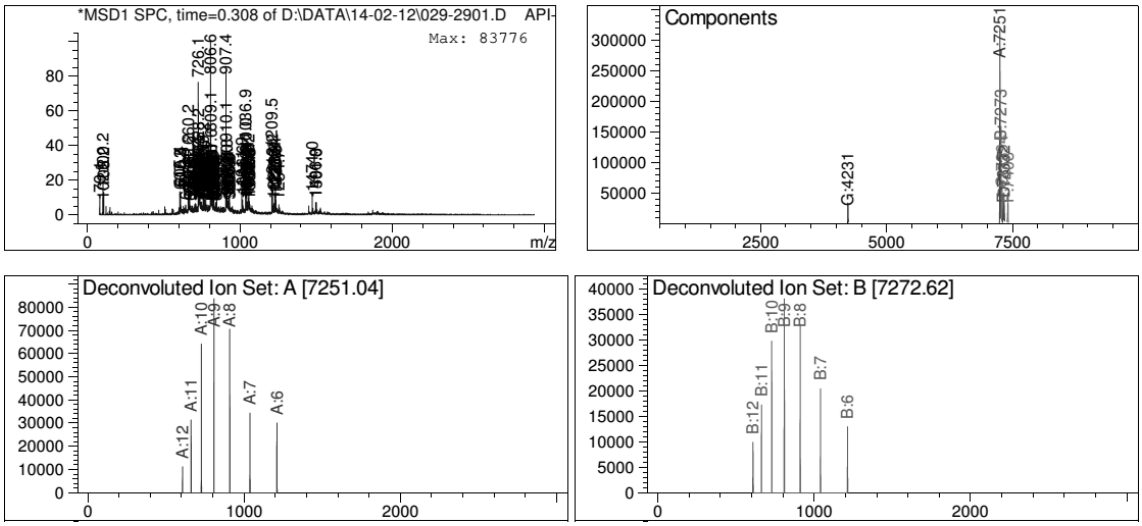


Figure S12: ESI-MS of compound **9**. Calcd. E.M. = 7247.74, M + Na^+ = 7270.74 and deconvoluted masses found 7251.04 and 7272.62 + other minor impurities.

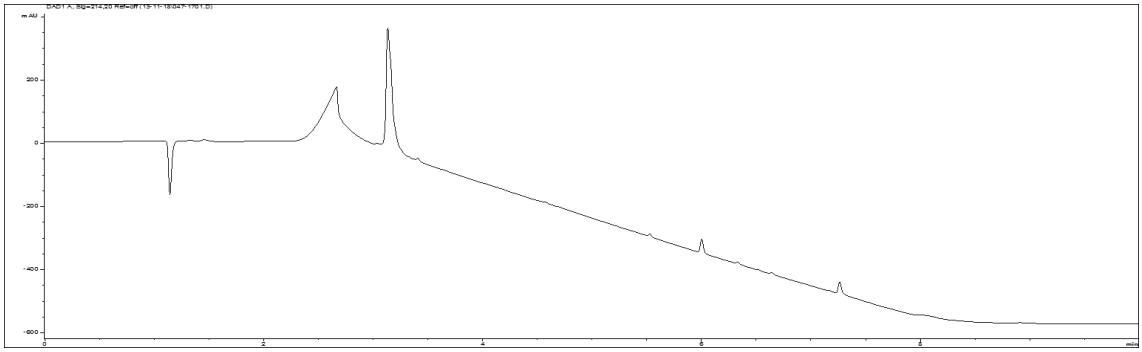


Figure S13: HPLC trace of compound **9** from LC-MS (0-100% ACN in 6 min on Kinetex C18 100 Å, 150 x 2.1 mm, 2.6 µm, at 35 °C)

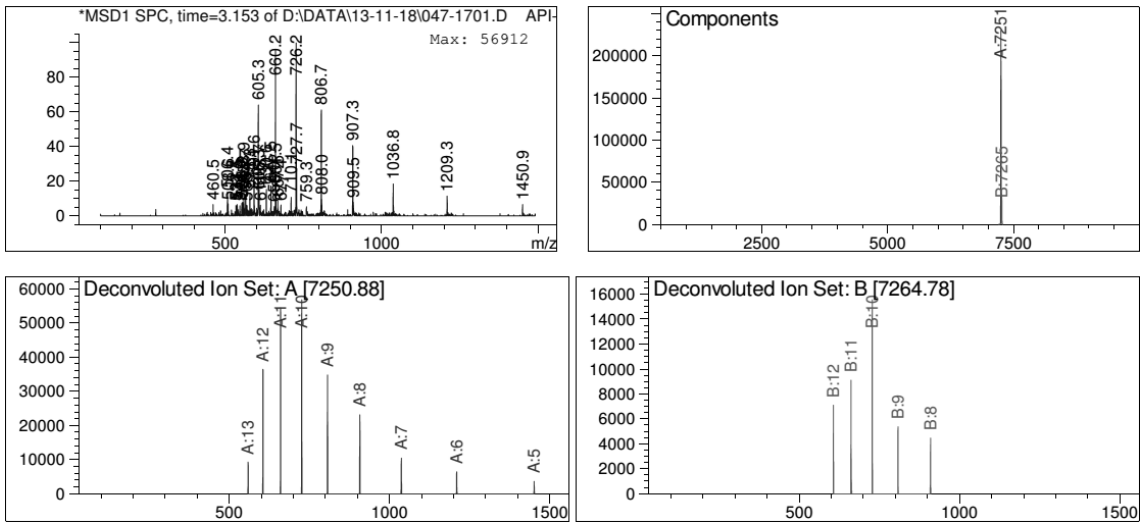


Figure S14: ESI-MS from LC-MS at r.t. = 3.153 min for purified compound **9**. E.M. = 7247.74, M+NH_4^+ = 7265.74 and deconvoluted mass found 7250.88 and 7264.78.

Compound S13 (Gamma conjugate)

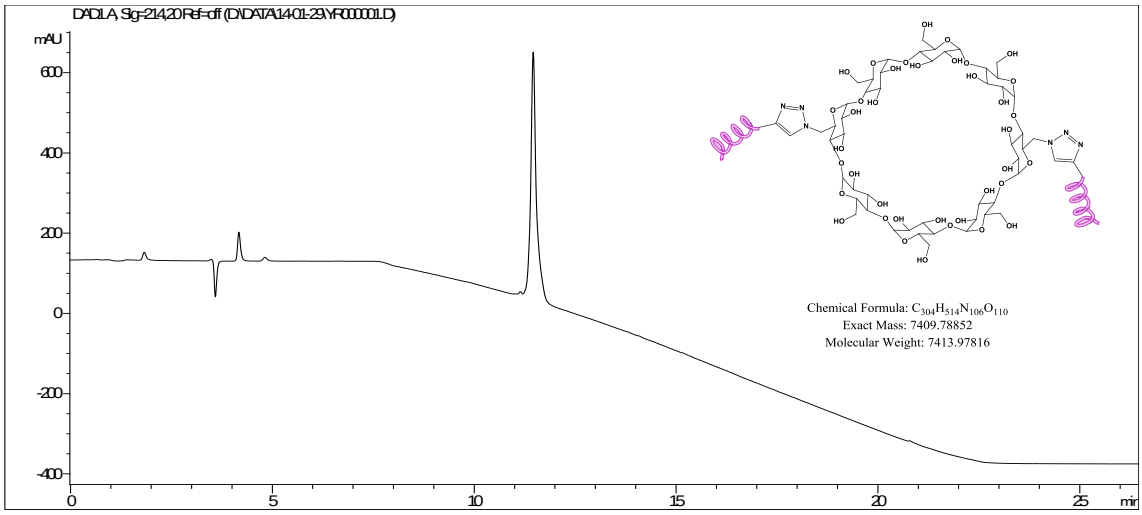


Figure S15: RP-HPLC Chromatogram of pure compound **10** (Jupiter C4, 300Å column using a gradient from 0 to 100 % CH_3CN in 15 minutes)

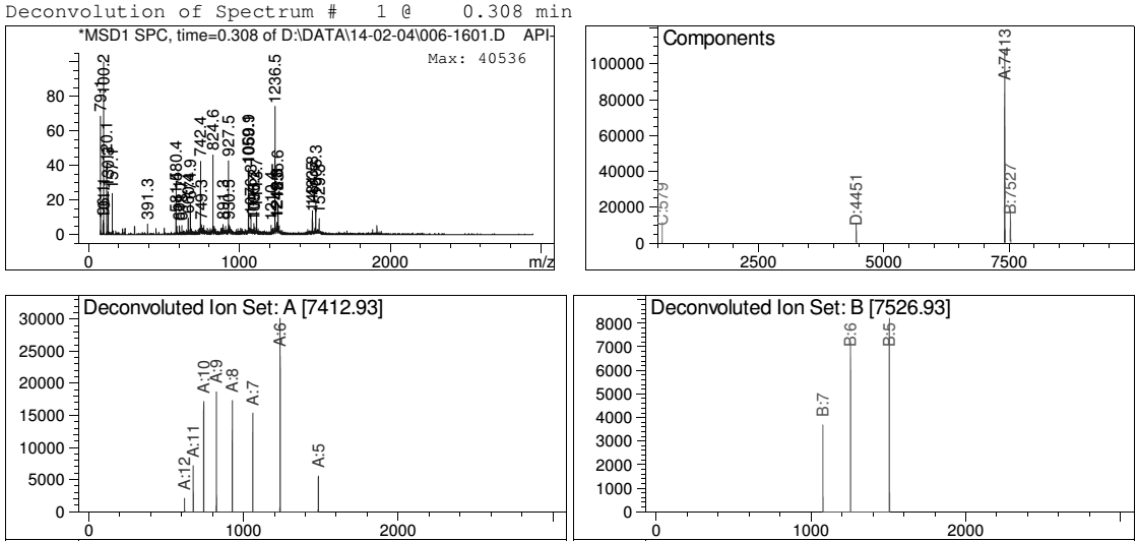


Figure S16: ESI-MS of compound **10**. Calcd. E.M. = 7409.79, M + 114.02 (TFA salt) = 7523.81 and deconvoluted masses found 7412.93 and 7526.93 + other minor impurities.

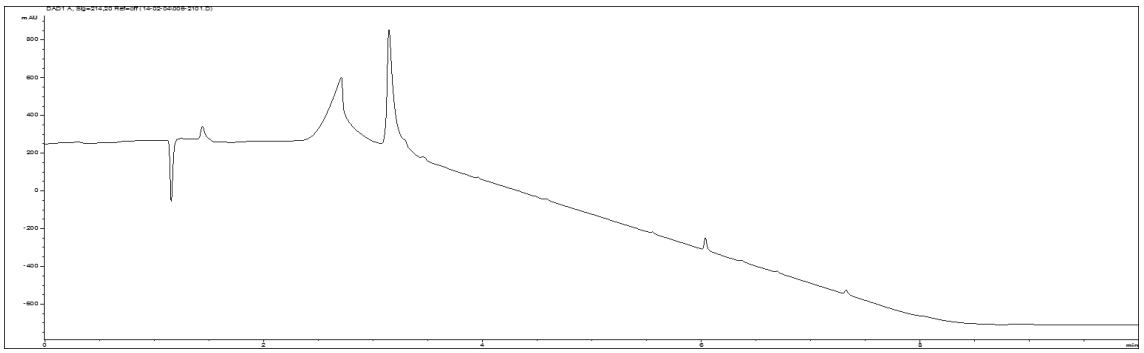


Figure S17: HPLC trace of compound **10** from LC-MS (0-100% ACN in 6 min on Kinetex C18 100 Å, 150 x 2.1 mm, 2.6 µm, at 35 °C)

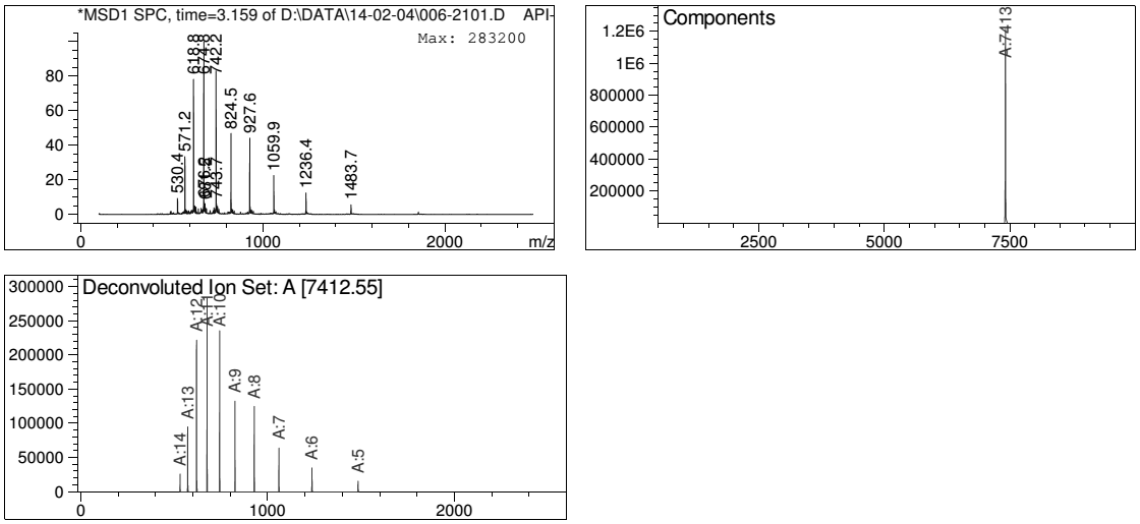


Figure S18: ESI-MS from LC-MS at r.t. = 3.159 min. Cacld. E.M. = 7409.79 and deconvoluted mass found 7412.55

5. Electrophoretic Mobility Assay

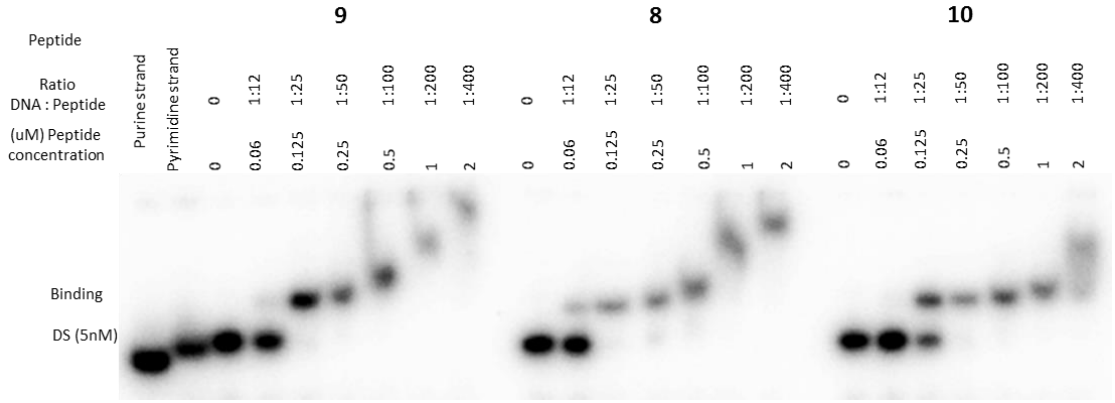


Figure S21: EMSA titration of the dipodal peptidocyclodextrin conjugates to the 5'-labeled ^{32}P-CRE sequence: Concentrations mentioned above. First and second lane: purine and pyrimidine strand respectively. Gels from left to right: **9, 8** and **10**. Bands at the bottom correspond to the excess of isotope and P^{32}-radiolabelled ssDNA.

As it is observed in figure S21, all 3 compounds binds to the DNA. In addition, compound **9** presents the best binding capability, as at a concentration of 0.125 µM all dsDNA are bound by the peptide. On the other hand, compound **10** has the lowest affinity for the target sequence compared to **8** and **9**. At a concentration of 0.125 µM only half of the free dsDNA is consumed.

Moreover, in order to describe quantitatively the binding capacity of the 3 derivatives, the determination of the dissociation constant is needed. However, a optimal binding pattern was obtained only for compound **10**, while aggregation took place at high concentrations for compounds **8** and **9**. The determination of the dissociation constant required a gradual binding pattern, in which it could be observed how the band corresponding to the dsDNA disappeared gradually while the band of the peptide-DNA complex started to be visible until its intensity becomes predominant in the gel at higher concentrations of peptide. Therefore, a new set of

experiments was needed of these compounds. Optimization of the concentration range was performed and a 10x range concentrations of 0, 0.5, 0.625, 0.75, 0.875, 1, 1.125 and 1.25 µM for **8** and **9** was considered for Kd calculations.

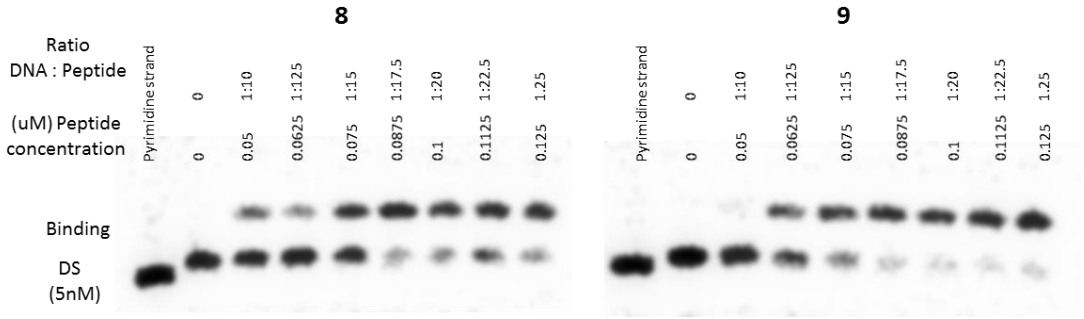


Figure S22: EMSA titration of the dipodal peptidocyclodextrin conjugates to the 5'-labeled ^{32}P-CRE sequence: Concentrations mentioned above. First lane in all the gels: pyrimidine strand. Gels from left to right: **8** and **9**.

5.2 Calculation of the dissociation constant

5.2.1 Kd calculations

The data were fitted with the following equation: KD=[Peptide][D]/[SB] where [D] and [SB] are the concentrations of duplex and shifted band, respectively, and by assuming a 1:1 stoichiometry of Peptide and duplex and an unknown KD. Because the Peptide is present in large excess over the target duplex, we assume the concentration of Peptide as constant and equal to its initial concentration C0. The equation used for the curve fit is %D=100/(1+C0/KD).

Using these calculations, we obtained the Kd values for **8**, **9** and **10**; 50 ± 20, 30 ± 20 and 100 ± 60 nM respectively.

5.2.2 Previous models

For better understanding where the constructs synthesized in the OBCR group stand in terms of relative binding affinity to the CRE we compared it to previously published models by various groups[5]. Due to several variations in the length of the peptide chain and conditions used for gel electrophoresis including different DNA concentrations, buffers, acrylamide percentages in gels, etc. a direct comparison is almost impossible. To obtain a realistic comparison, we have only chosen those models which have the same peptide length as the one synthesized by us.

The construct synthesized by Morii[6] which is based on the non-covalent interaction between cyclodextrin and adamantine, has a Kd < 100 nm. Although a Kd has not been stated in the article, this is quite evident from the EMSA data published. This data is comparable to our models, although based on our calculations the beta cyclodextrin is marginally better than the construct by Morii and is also a simpler design.

The mimic from the lab of Mascareñas[7] based on the photo switchable diazobenzene linker in the cis form has a Kd < 5 nm as stated in the article. The dissociation constant of this construct

is lower than our best construct (Kd = 30 nm). However, whether the synthetic complexity of the linker is worth the improvement in binding affinity if this construct is to be used for future applications is something to be considered.

Yet another TF design from the lab of Mascareñas[8] with the ability to bind selectively either CRE or its inverse sequence by changing the reaction conditions, has a Kd of 299 ± 26 nM for CRE. This ability of modulating the DNS recognition makes the design unique. Given the nature of the construct, the affinity values are quite good, but of a higher order of magnitude than what we have synthesized by using derivatized CD as scaffolds.

The fourth and final construct we have used for comparison is actually a shortened version of the dimer from Kim[9] which has been successfully modified and published by the labs of Mascareñas and Vázquez. The Kd, although not reported is <150 nm based on the EMSA data from the articles, once again making our models slightly better in terms of binding affinity to the CRE.

5.3 Electrophoresis mobility shift assay with random DNA

Sample preparation:

The following stock solutions were prepared (fresh each time, except for DNA and peptide):

DNA: 1.67 µM prepared from Random sequence (5' – CGG ATG ACG TCA TTT TTT TTC – 3') & Random complement (5' – GAA AAA AAA TGA CGT CAT CCG – 3') DNA solutions (obtained from Eurogentec purified over gold cartridge) by diluting with 20 µL 0.5 M Tris, pH = 8, 40 µL 2.5 M NaCl, 40 µL 0.025 M EDTA and then adding milliQ water such that the total volume is 1 mL. DNA was annealed by heating from room temperature to 95°C for 24 min in a Thermomixer. The machine was then turned off and the sample was allowed to cool down slowly.

Loading buffer: 20 µL Tris 1 M, pH = 7.6, 20 µL KCl 0.2 M, 20 µL $MgCl_2$ 0.1 M, 40 µL EDTA 0.025 M.

Sucrose: 30% sucrose in milliQ (300 mg/mL)

Peptides: 10 µL stock solutions (10x) were prepared in MilliQ water. Concentrations in lanes 1-10 in µM: 0, 0.167, 0.501, 0.668, 0.751, 0.835, 1.002, 1.169, 1.336, 1.67.

Loading mixture: The loading mixture comprised of: 10 µL milliQ, 4 µL sucrose, 2 µL loading buffer, 2 µL DNA, 2 µL peptide. The loading mixture was prepared only 1-2 hr prior to running of gels and kept on ice as soon as ready.

Preparation of Gels (for 2 Gels):

In a clean falcon tube the following solutions were added (in given order): 15.595 mL milliQ, 0.4 mL TBE buffer, 4.005 mL of 40% acrylamide solution, 200 µL APS(10% w/w in milliQ). The solution was mixed by sonication to remove any air bubbles and cooled to 0°C (1 h under ice). 20 µL of TEMED was then added to the mixture and was again mixed properly before pouring it gently along parallel glass plates. The glass plates were tapped gently to ensure removal of

all air bubbles and the markers were squeezed between the plates to ensure uniform width of each well. Sufficient time was given for polymerization (~1 h).

Gel Electrophoresis:

A pre-run of the gels was performed prior to loading them. Care was taken to see that the gels were properly immersed in 0.2x TBE buffer (non-denaturing gel, without urea) and the loading wells were free from any air bubbles. Instrument settings: 150 V, 100 mA, 19 W for 30 mins at 4°C. The wells were washed after the pre-run. 5 µL of the loading mixture was then loaded onto the wells. Instrument settings: 150 V, 100 mA, 19 W for 45 mins at 4°C.

Staining of gels:

After the run, the gels were removed from the glass and were stained using 100 mL of 0.2x TBE buffer + 10 µL SybrGold® (Life Technologies™) stock solution 10,000X in DMSO. The gels were then washed twice with milliQ and gently placed under a UV lamp (dark room) to observe the gel pattern.

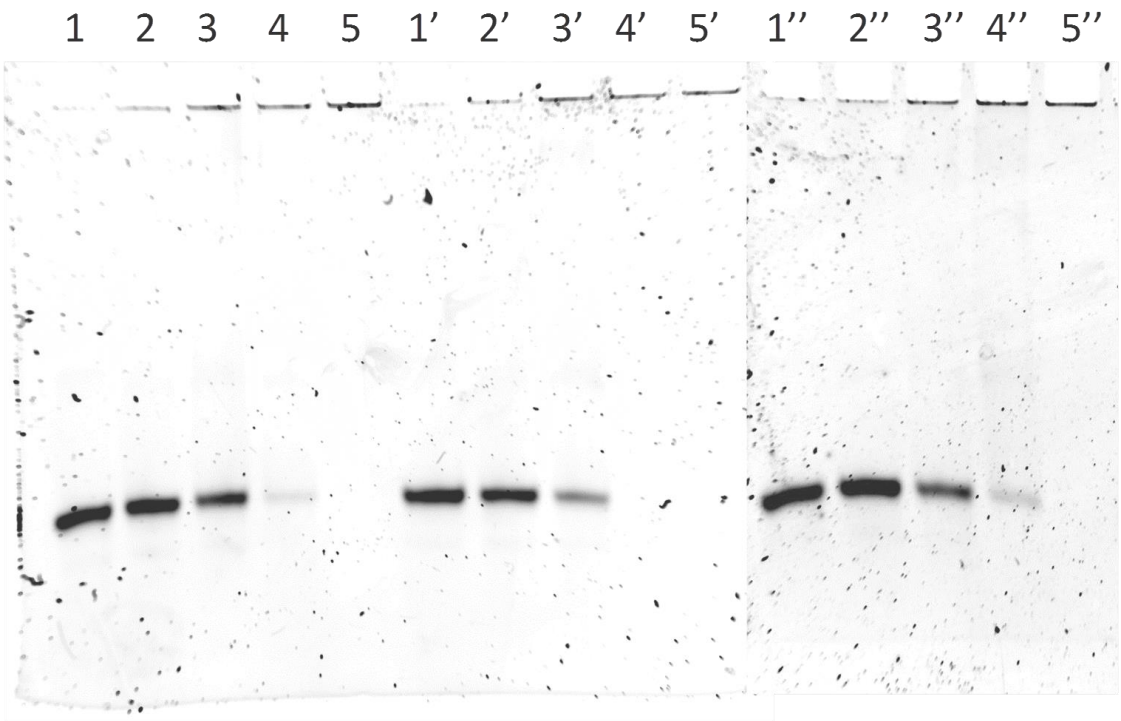


Figure S23: EMSA titration of the dipodal peptidocyclodextrin conjugates to the random sequence: Concentrations: 0.167 µM dsDNA; Lanes 1-5: 0. 0.167, 0.501, 0.668, 1.67 µM. Lanes 1-5 for compound **8** (alfa conjugate). Lanes 1’-5’ compound **9** (beta conjugate). Lanes 1’’-5’’ Compound **10** (gamma conjugate).

It can be concluded from figure S23 that compounds **8**, **9** and **10** are sequence-selective towards the target sequence CRE.

REFERENCES:

(1) Akoka, S.; Barantin, L.; Trierweiler, M. *Anal. Chem.* **1999**, *71*, 2554.

(2) Volkov, S.; Kumprecht, L.; Buděšínský, M.; Dušek, M.; Kraus, T. *Org. Biomol. Chem. Accept. Publ.* **2014**.

(3) Kumprecht, L.; Budesínský, M.; Vondrásek, J.; Vymetal, J.; Cerný, J.; Císarová, I.; Brynda, J.; Herzig, V.; Koutník, P.; Závada, J.; Kraus, T. *J. Org. Chem.* **2009**, *74*, 1082.

(4) Grishina, A.; Stanchev, S.; Kumprecht, L.; Buděšínský, M.; Pojarová, M.; Dušek, M.; Rumlová, M.; Křížová, I.; Rulíšek, L.; Kraus, T. *Chem. - A Eur. J.* **2012**, *18*, 12292.

(5) Pazos, E.; Mosquera, J.; Vázquez, M. E.; Mascareñas, J. L. *Chembiochem* **2011**, *12*, 1958.

(6) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575.

(7) Caamano, A. M.; Vazquez, M. E.; Martinez-Costas, J.; Castedo, L.; Mascarenas, J. L. *Angew. Chem., Int. Ed.* **2000**, *39*, 3104.

(8) Mosquera, J.; Jimanez-Balsa, A.; Dodero, V. I.; Vazquez, M. E.; Mascarenas, J. L. *Nat. Commun.* **2013**, *4*, ncomms2825.

(9) Jiménez-Balsa, A.; Pazos, E.; Martínez-Albardonedo, B.; Mascareñas, J. L.; Vázquez, M. E. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 8825.