Electronic Supplementary Material (ESI) for RSC Advances.
This journal is © The Royal Society of Chemistry 2017

pH-Controlled recognition of amino acids by urea derivatives of β-cyclodextrin

Pawel Stepniak,[a] Bruno Lainer,[ab] Kazimierz Chmurski,[b] Janusz Jurczak[a*]

[a] Institute of Organic Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224 Warsaw, Poland

[b] Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

jurczak_group@icho.edu.pl

Supporting Information

<u>Table of Contents</u>

General information…………………..S3

Synthetic procedures………………….S4 – S8

Isothermal titration calorimetry………S9 – S11

DLS measurements…………………...S12

NMR spectra………………………….S13 – S15

General Information:

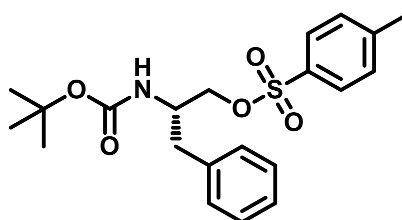
All solvents were used as received, unless stated otherwise. Purification of products was performed using chromatography on silica gel (Merck Kieselgel 60, 230-400 mesh) with mixtures of hexane/ethyl acetate, methanol/dichloromethane, acetonitrile/water/ammonia or using gel filtration on styrene resin (Diaion® HP-20) with water/methanol gradient. Thin-layer chromatography (TLC) was performed on silica gel plates (Merck Kieselgel 60 F254). Visualization of the developed chromatogram was accomplished using UV light or ninhydrin and cerium molybdate stains. Reported NMR spectra were recorded in $CDCl_3$ or $(CD_3)_2SO$ using a Varian Unity Plus 200 MHz and Agilent 300 MHz and 500 MHz spectrometers. Chemical shifts of ^1H NMR and ^{13}C NMR are reported as δ values relative to TMS (δ=0.00) and $CDCl_3$ (δ=77.0) or $(CD_3)_2SO$ (δ=39.5), respectively. The following abbreviations are used to indicate the multiplicity: s - singlet; d - doublet; t - triplet; q - quartet; m – multiplet; dm - doublet of multiplets. Mass spectra were measured on a Shimadzu LCMS-IT-TOF using ESI technique.

Distilled water (>18 MΩ/cm grade) was supplied by Mili-Q water system. Phosphate buffers were prepared by mixing specified amounts of $NaH_2PO_4 \cdot 2H_2O$ and $Na_2HPO_4 \cdot H_2O$ and subsequent titration with NaOH, to desired pH, using Elmetron CP-401 pH-meter.

Dialysis was performed using Spectrum Labs Biotech Grade Cellulose Ester Dialysis Membrane, MWCO: 100-500 Da, preserved with 0.05% sodium azide.

Synthetic Procedures:

(S)-O-Tosyl-2-(Boc-amino)-3-phenylpropanol



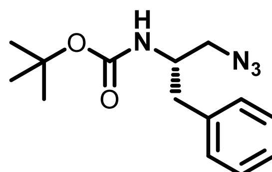
N-Boc-L-Phenylalaninol (10.8 g, 43.0 mmol) was dissolved in dichloromethane (100 ml) and cooled to 0°C. Triethylamine (1.1 equiv., 47.3 mmol, 6.59 mL) and p-tosyl chloride (1.05 equiv., 45.1 mmol, 7.77 g) were added and the resulting mixture was stirred overnight (allowing the mixture to warm up to rt). It was next washed with saturated sodium bicarbonate (100 ml) and brine (20 ml). Organic layer was then dried over $MgSO_4$ and evaporated under reduced pressure. Procuct was purified on silica gel using 10:90 ethyl acetate:hexane resulting with 10.1 g of **2** (24.9 mmol, 58%).

Careful purification of the aminoalcohol **2** before the tosylation process is crucial, as mixed anhydride (formed prior to the reduction step) acts as an inhibitor.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.78 (d, *J* = 8.3 Hz, 2H), 7.50 – 6.88 (m, 7H), 4.71 (s, 1H), 4.20 – 3.71 (m, 2H), 2.95 – 2.56 (m, 2H), 2.46 (s, 3H), 1.40 (d, *J* = 7.2 Hz, 9H).

^{13}C NMR (75 MHz, $CDCl_3$) δ 145.07, 136.74, 132.52, 129.98, 129.24, 128.61, 128.02, 126.74, 79.83, 70.01, 37.18, 28.28, 21.67.

(S)-2-(Boc-amino)-1-azido-3-phenylpropane (3)



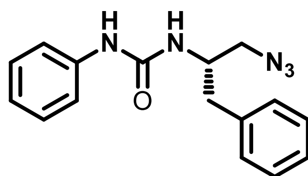
2 (6.6 g, 16.0 mmol) was dissolved in DMF (20 ml) and NaN_3 (3 equiv., 48.1 mmol, 3.1 g) was added. Addition of more than 3 equivs of NaN_3 caused the elimination of the tosylated aminoalcohol due to the basicity of azide ions in DMF. Resulting mixture was stirred at 60°C for 24 h. It was then poured into water (300 ml) and extracted (3 x 100 ml) with diethyl ether. Combined organic layers were then evaporated under reduced pressure. Oil residue was dried under vacuum until it solidified resulting with 3.67g (13.3 mmol, 83%) of **3**.

^{1}H NMR (200 MHz, $CDCl_3$) δ 8.01 (s, 1H), 7.46 – 6.99 (m, 5H), 3.95 (s, 1H), 3.36 (qd, *J* = 12.4, 4.6 Hz, 2H), 2.91 – 2.67 (m, 2H), 1.42 (s, 9H).
^{13}C NMR (50 MHz, $CDCl_3$) δ 155.18, 137.23, 136.05, 129.35, 129.07, 128.72, 127.31, 126.80, 69.67, 53.87, 53.23, 51.43, 41.53, 38.24, 36.57, 28.39.

Synthesis of urea 4

(S)-2-(Boc-amino)-1-azido-3-phenylpropane (600 mg, 2.2 mmol) was dissolved in 1:1 mixture of TFA and dichloromethane (5 ml) and the resulting mixture was stirred for 1h, after which all volatiles were evaporated. In order to neutralize residual TFA and deprotonate the amine, the solid residue wastreated with triethylamine (1 mL), which excess was subsequentlyevaporated. Dichloromethane (20 ml) was added and the solution was treated with respective isocyanate (1.1 equiv, 2.4 mmol). After 24h of stirring overnight the mixture was washed with saturated sodium bicarbonate (50 ml) and brine (50 ml). Organic layer was dried with $MgSO_4$ and the product was purified on silica gel using 3:97 mixture of methanol:dichloromethane, yielding, after evaporation and vacuum drying, urea **4**.

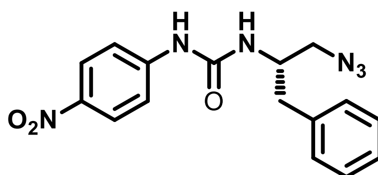
(S)-1-(1-azido-3-phenylpropan-2-yl)-3-phenylurea (**4a**)



Yield: 507 mg (1.72 mmol, 78%)

^{1}H NMR (200 MHz, $CDCl_3$) 7.64 (s, 1H), , 7.40 – 7.14 (m, 10H), 5.55 (d, *J* = 8.2 Hz, 1H), 4.32 – 4.13 (m, 1H), 3.58 – 3.29 (m, 2H), 3.01 – 2.75 (m, 2H).

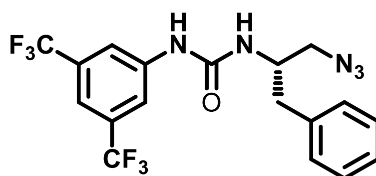
(S)-1-(1-azido-4-methylpentan-2-yl)-3-(4-nitrophenyl)urea (**4b)**



Yield: 523 mg (1.5 mmol, 70%)

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.18 – 8.08 (m, 2H), 7.64 (s, 1H), 7.53 – 7.44 (m, 2H), 7.40 – 7.14 (m, 5H), 5.55 (d, *J* = 8.2 Hz, 1H), 4.32 – 4.13 (m, 1H), 3.58 – 3.29 (m, 2H), 3.01 – 2.75 (m, 2H).

(S)-1-(1-azido-4-methylpentan-2-yl)-3-(3,5-di(trifluoromethyl)-phenyl)urea (**4c**)



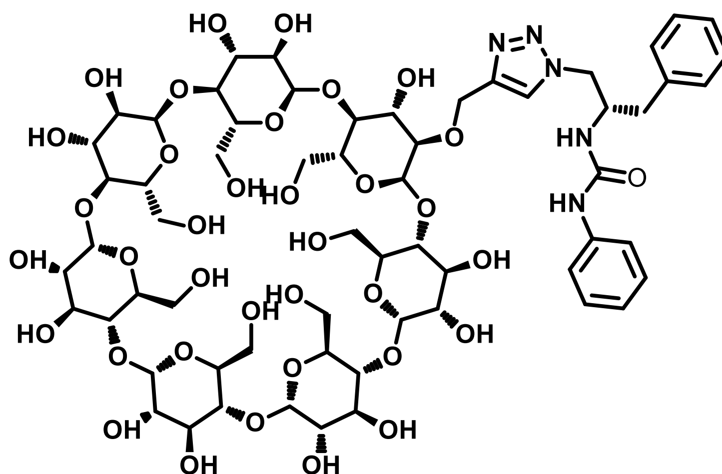
Yield: 776 mg (1.8 mmol, 80%)

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.28 (s, 1H), 7.55 (s, 2H), 7.37 (s, 1H), 7.30 – 7.09 (m, 5H), 6.15 (d, J = 8.5 Hz, 1H), 4.36 – 4.04 (m, 1H), 3.63 – 3.19 (m, 2H), 2.99 – 2.69 (m, 2H).

Synthesis of receptors 1d-f

Mono-2-propargyl-β-CD **1** (1 g, 0.85 mmol) and corresponding azide (1.1 equiv., 0.94 mmol) were dissolved in 4:1 mixture of DMSO and water (30 ml). Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, ca. 20 mg). Mixture was deaerated by passing argon through solution. CuSO4 (0.1 equiv, 0.085 mmol, aqueous solution) and sodium ascorbate (0.2 equiv, 0.17 mmol, aqueous solution) were added dropwise into stirred reaction mixture, which was further stirred for 24 h at 50°C. After the removal of the catalyst (via the addition of ammonia to the reaction mixture and its passage through a thin pad of silica gel) Solvents were evaporated under reduced pressure and crude product was dissolved in water (5 ml) and precipitated with acetone (500 ml). Precipitate was filtered and air dried, yielding pure compound **1**.

(S)-2-*O*-((1-(2-(3-phenylureido)-3-phenylpropyl)-1*H*-1,2,3-triazol-4-yl)methyl)-β-CD (**1d**)



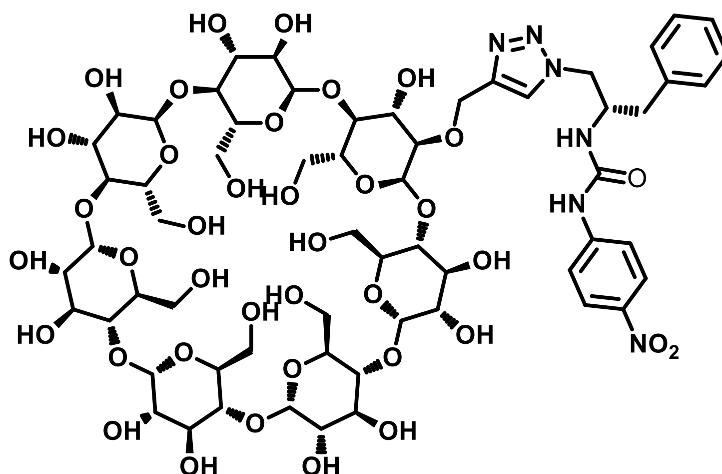
Yield: 1160 mg (0.78 mmol, 92%)

^{1}H NMR (500 MHz, DMSO) δ 8.40 (s, 1H), 8.17 – 7.97 (m, 1H), 7.49 – 7.05 (m, 8H), 6.88 (t, *J* = 7.3 Hz, 2H), 6.15 (s, 1H), 6.03 – 5.53 (m, 11H), 5.03 – 4.64 (m, 7H), 4.48 (tdd, *J* = 20.4, 18.2, 9.1 Hz, 7H), 3.80 (dd, *J* = 28.9, 19.3 Hz, 1H), 3.75 – 3.49 (m, 23H), 3.46 – 3.06 (m, 31H).

^{13}C NMR (126 MHz, DMSO) δ 129.19, 128.65, 128.38, 126.37, 124.99, 121.24, 117.77, 101.96, 101.59, 100.25, 82.07, 81.56, 81.24, 79.52, 73.28, 73.07, 72.98, 72.72, 72.40, 72.19, 72.05, 71.79, 71.71, 64.47, 59.94, 52.60, 50.89, 37.57, 29.56.

HRMS (ESI TOF) calcd for $C_{61}H_{89}N_5O_{36}Na$ $[M+Na]^+$: 1490.5185 found: 1490.5161.

(S)-2-*O*-((1-(2-(3-(4-nitrophenyl)ureido)-3-phenylopropyl)-1*H*-1,2,3-triazol-4-yl)methyl)-β-CD (**1e**)

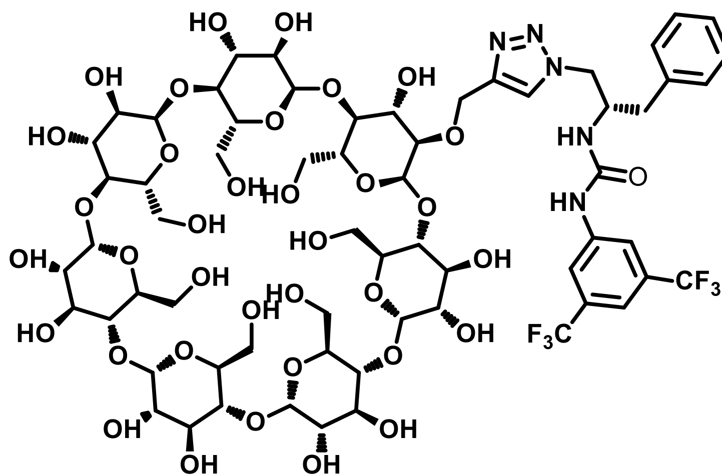


Yield: 1170 mg (0.76 mmol, 90%)

^{1}H NMR (500 MHz, DMSO) δ 9.21 (d, *J* = 22.1 Hz, 1H), 8.12 (dd, *J* = 12.2, 9.8 Hz, 3H), 7.56 (dd, *J* = 21.5, 9.3 Hz, 2H), 7.25 (qd, *J* = 14.6, 7.4 Hz, 5H), 6.53 (d, *J* = 40.4 Hz, 1H), 6.03 – 5.52 (m, 13H), 4.95 – 4.70 (m, 7H), 4.65 – 4.26 (m, 7H), 3.84 (dd, *J* = 28.2, 19.0 Hz, 1H), 3.77 – 3.49 (m, 23H), 3.49 – 3.14 (m, 20H).

^{13}C NMR (126 MHz, DMSO) δ 153.84, 146.79, 140.54, 137.76, 129.14, 128.40, 126.42, 125.10, 116.99, 101.95, 81.57, 73.06, 72.39, 72.02, 59.93.
HRMS (ESI TOF) calcd for $C_{61}H_{88}N_6O_{38}Na$ $[M+Na]^+$: 1535.5036 found: 1535.5029.

(S)-2-*O*-((1-(2-(3-(3,5-bis(trifluoromethyl)phenyl)ureido)-3-phenylpropyl)-1*H*-1,2,3-triazol-4-yl)methyl)-β-CD (**1f**)



Yield: 1310 mg (0.80 mmol, 95%)

^{1}H NMR (500 MHz, DMSO) δ 9.19 (s, 2H), 8.14 (d, *J* = 10.9 Hz, 1H), 7.98 (s, 2H), 7.54 (s, 1H), 7.25 (qd, *J* = 14.5, 7.4 Hz, 5H), 6.57 (s, 1H), 6.03 – 5.56 (m, 12H), 4.95 – 4.69 (m, 7H), 4.63 – 4.38 (m, 7H), 3.93 – 3.78 (m, 1H), 3.78 – 3.48 (m, 23H), 3.48 – 3.15 (m, 20H).

^{13}C NMR (126 MHz, DMSO) δ 129.16, 128.35, 126.35, 125.07, 123.42, 117.40, 113.69, 101.94, 101.56, 100.21, 82.05, 81.57, 81.49, 81.43, 81.20, 79.43, 73.23, 73.04, 72.93, 72.68, 72.49, 72.38, 72.18, 72.02, 71.76, 71.68, 64.38, 59.92, 52.39, 51.26, 37.39.

HRMS (ESI TOF) calcd for $C_{63}H_{87}N_5O_{36}NaF_6$ $[M+Na]^+$: 1626.4933 found: 1626.4904.

Isothermal titration calorimetry:

Prior to binding measurements, compounds 5 were purified via dialysis. 500 mg of compound **5** was dissolved in 10 ml of water and placed in a dialysis tube. Sealed tube was placed in 3L of water, which was stirred during the purification process. Water was replaced with a fresh one after 2, 12 and 24 hours. Finally the content of the bag was lyophilized.

Isothermal titration calorimetry (ITC) was performed using TA Instruments NanoITC calorimeter. Solution of all hosts and guests were prepared in 50mM phosphate buffer (pH=8). In a typical experiment calorimeter cell (0.949 ml) was filled with 2 mM solution of host **5**. Cell solution was stirred at 300 rpm. 10-16 mM amino acid solution (250 µl) was titrated into the calorimeter cell in 25 injections, with 300 s intervals between injections, at 25°C. Measured raw heats were corrected with the heat of dilution of amino acid titrated into phosphate buffer. Heats were fitted to n-independent-sites binding model ($H + nG = HG_n$), which allowed to determine K_a (affinity constant), ΔH (effective molar enthalpy of the reaction) and n (stoichiometry of guest:host interaction). Mean parameters and their uncertainties were calculated of at least 3 independent experiments.

Model:

ITC measurements were conducted by the titration of amino acids into solutions of receptors. The concentrations were restricted by the heat response upon titration and solubility of the receptors. We have applied a model where one of the interacting species has n identical and independent binding sites (n-independent-sites binding model), which allowed to compare the strength of the interaction between the receptor and particular guests (amino acids).

Interaction parameter $\bar{\nu}$ can be defined as number of moles of guest (G) bound to one mole of host (H), or in terms of concentration:

$$\bar{\nu} = \frac{[G]_b}{[H]_T} = \frac{\sum_{i=1}^{n} i \cdot \beta_i \cdot [G]^i}{\sum_{i=0}^{n} \beta_i \cdot [G]^i}$$

Transformation of this relation allows to link $\bar{\nu}$ with number of active sites of the host (n), microscopic affinity constant K (defined for elementary reaction G+H=GH) and unbounded guest concentration:

$$\bar{\nu} = \frac{[G]_B}{[H]_T} = \frac{nK[G]}{1 + K[G]}$$

Heat q_i, exchanged during single injection can be defined as:

$$q_i = \Delta H_{app} \cdot \Delta n_G$$

where Δn_G is number of moles of guest titrated into the cell during single injection and ΔH_{app} is apparent enthalpy of interactions in the system. In terms of guest concentration this equation can be written as:

$$q_i = \Delta H_{app} \cdot V_C \cdot \left([G]_{b,i} - [G]_{b,i-1}\right) = \Delta H_{app} \cdot V_C \cdot \left(\bar{\nu}_i \cdot [H]_{T,i} - \bar{\nu}_{i-1} \cdot [H]_{T,i-1}\right)$$

where V_C is cel volume, $[H]_T$ is total host concentration in the cell. In consequence after N injextions exchanged heat Q can be defined as:

$$Q = \sum_{i=1}^{n} q_i = \Delta H_{app} \cdot V_C \cdot [H]_T \cdot \bar{\nu} = \Delta H_{app} \cdot V_C \cdot [H]_T \cdot \frac{nK[G]}{1 + K[G]}$$

Therefore, after substitution of $[G]$ with known or measurable parameters:

$$[G] = [G]_T - [G]_b = [G]_T - \frac{Q}{V_C - \Delta H_{app}}$$

We have quadratic equation which solutions gives expresion for total exchanged heat:

$$Q = \frac{V_C \Delta H_{app}}{2K}\left[1 + K[G]_T + nK[H]_T - \sqrt{(1 + K[G]_T + nK[H]_T)^2 - 4nK^2[G]_T[H]_T}\right]$$

Single injection heat exchange can be infered by estimating derivative:

$$\frac{1}{V_C} \cdot \frac{dQ}{d[G]_T} \approx \frac{1}{V_C} \cdot \frac{\Delta Q}{\Delta [G]_T} = \frac{\Delta H_{app}}{2}\left[1 - \frac{1 + [G]_T - nK[H]_T}{\sqrt{(1 + K[G]_T + nK[H]_T)^2 - 4nK^2[G]_T[H]_T}}\right]$$

Values of K, ΔH_{app} and n were calculated using nonlinear regresion program NanoAnalyzer supplied by TA Instruments. Errors in estimation were calculated using Monte Carlo analysis and included in measurement uncertainties analysis.

	K_a (M^{-1})					
	Ala		**Phe**		**Trp**	
Receptor	Value	Error	Value	Error	Value	Error
1a	2300	400	7000	1600	9800	5600
1b	2100	200	20200	1900	34400	2000
1c	5040	910	22400	1400	52400	3000
1d	4600	550	16500	2600	19100	2000
1e	2100	470	5300	1500	23400	6500
1f	5100	1160	23000	10000	54800	16000

Receptor	ΔH (kJ/mol)					
	Ala		**Phe**		**Trp**	
	Value	Error	Value	Error	Value	Error
1a	2.99	0.32	4.18	0.79	4.02	2.51
1b	4.89	0.27	4.90	0.09	5.42	0.07
1c	3.72	0.28	5.51	0.07	4.46	0.05
1d	2.92	0.09	4.10	0.13	5.19	0.17
1e	4.23	0.54	3.10	0.40	4.18	0.36
1f	3.08	0.27	1.69	0.31	1,55	0.34

Receptor	*n*					
	Ala		**Phe**		**Trp**	
	Value	Error	Value	Error	Value	Error
1a	0.509	0.039	0.380	0.050	0.349	0.092
1b	0.773	0.030	0.481	0.006	0.393	0.003
1c	0.922	0.026	0.520	0.002	0.457	0.002
1d	0.932	0.024	0.628	0.016	0.416	0.008
1e	0.358	0.036	0.108	0.025	0.129	0.009
1f	0.459	0.022	0.271	0.019	0.243	0.013

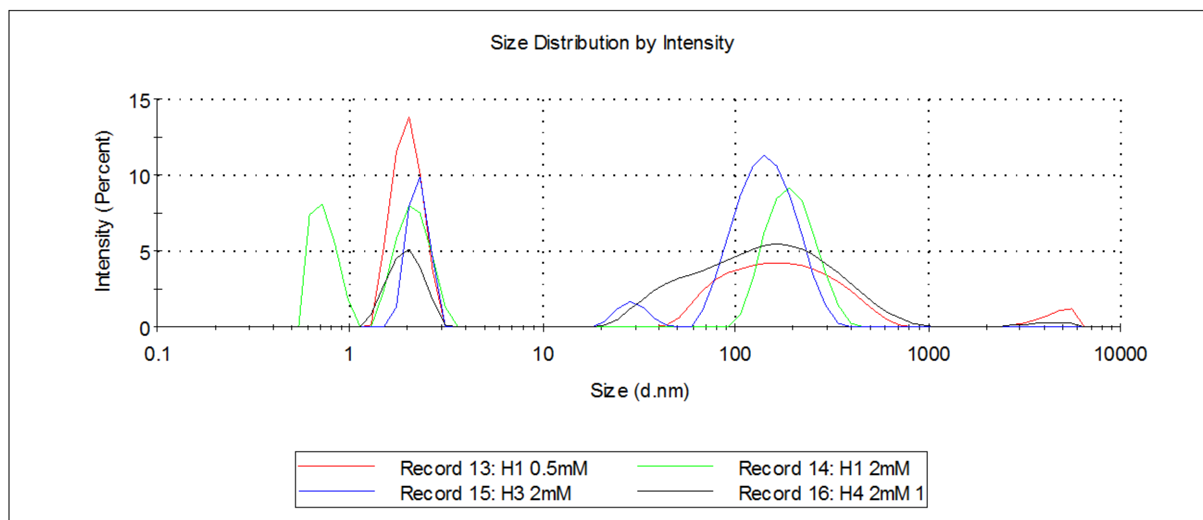
Parameter	Guest vs. receptor **1c**										
	Phe		**Tyr**		**PhGly**		**Ser**		**α-Ph-But**		**Man**
	Value	Error	Value	Error	Value	Error	Value	Error	Value	Error	Value
K_a (M^{-1})	22400	1400	13330	5645	4550	1017	19650	2673	443	nd	<100
ΔH (kJ/mol)	5.51	0.07	3.86	1.55	3.83	0.18	3.43	0.10	-0.79	nd	nd
n	0.52	0.002	0.44	0.042	0.77	0.023	0.53	0.008	1	nd	nd

Receptor	(S)-Ibuprofen vs. receptors					
	K_a (M^{-1})		ΔH (kJ/mol)		*n*	
	Value	Error	Value	Error	Value	Error
β-CD	8900	420	-14.0	0.12	0.75	0.005
1a	2100	120	-13.3	0.26	0.70	0.046
1c	1900	60	-17.7	0.29	0.70	0.009
1d	500	130	-11.4	4.23	0.54	0.136

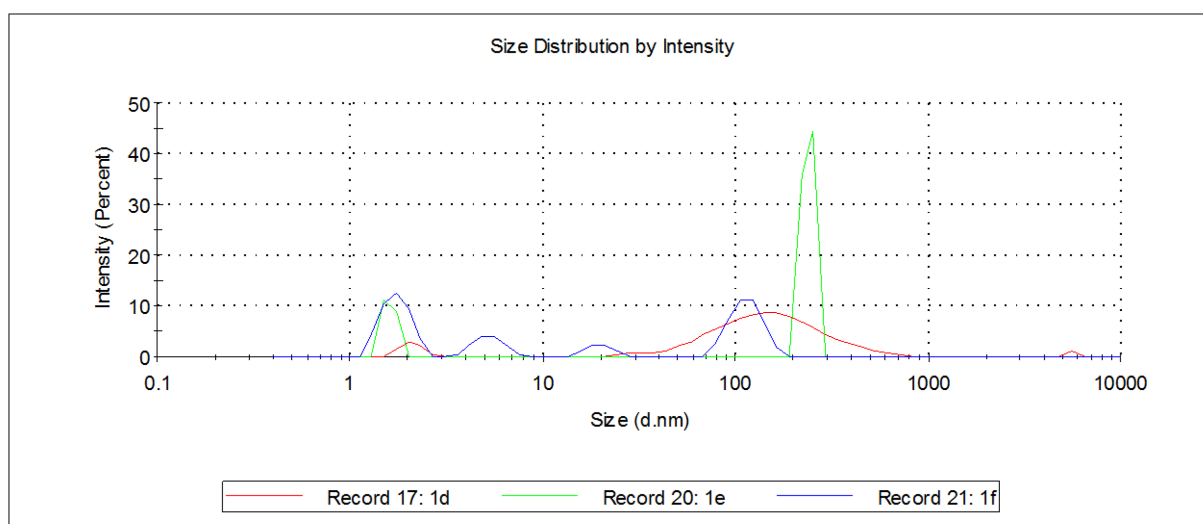
Aa	Parameter	pH (measured for receptor **1c**)							
		6		7		7.5		8	
		Value	Error	Value	Error	Value	Error	Value	Error
Phe	K_a (M^{-1})	585	80	1181	170	8044	1030	22400	1400
	n	0.413	0.079	0.687	0.046	0.662	0.009	0.577	0.002
	ΔH (kJ/mol)	4.92	nd	4.896	0.52	4.873	0.17	5.51	0.07
Trp	K_a (M^{-1})	381	40	6020	650	12559	2640	52400	3000
	n	0.5	nd	0.627	0.009	0.545	0.019	0.419	0.002
	ΔH (kJ/mol)	4.5	nd	3.767	0.13	5.001	0.37	4.46	0.05

<u>DLS measurements:</u>

Dynamic light scattering (DLS) measurements were conducted with a Zetasizer nanoseries Nano-ZS-ZEN3600 instrument equipped with a laser of 633 nm using a ZEN-112 cuvette at 25 °C. For each measurement the number of scans was 10, the run duration was 35 s, the quilibration time was 140 s, and the time delay was 4 s. Before measurements, the samples were incubated for 2 h, then filtered through a 0.45 μm PTFE filter, sonicated for 1 min, and finally filtered again.



Receptors: **5a** (0.5mM – red line, 2mM – green line), **5b** (blue line), **5c** (black line).



Spectra:

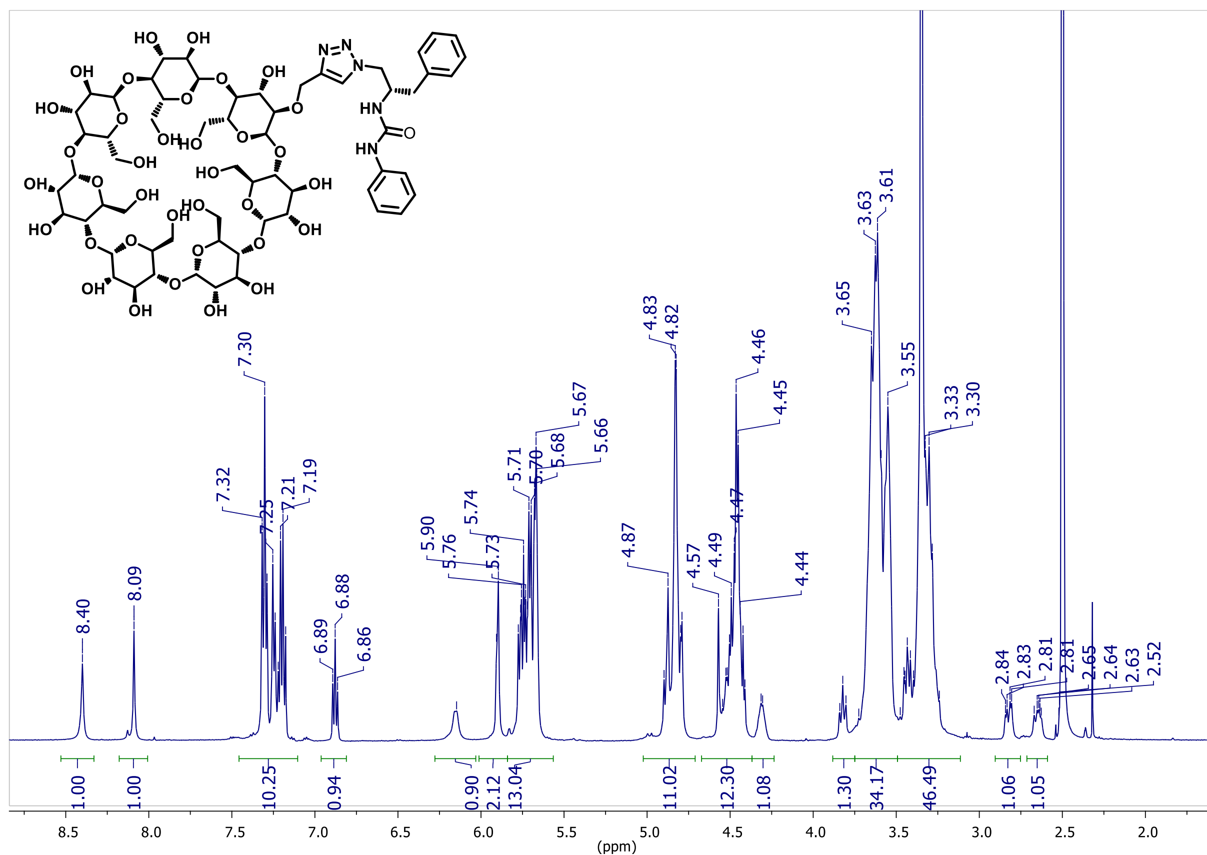


Fig. S1 ^{1}H NMR spectrum of compound **1d**.

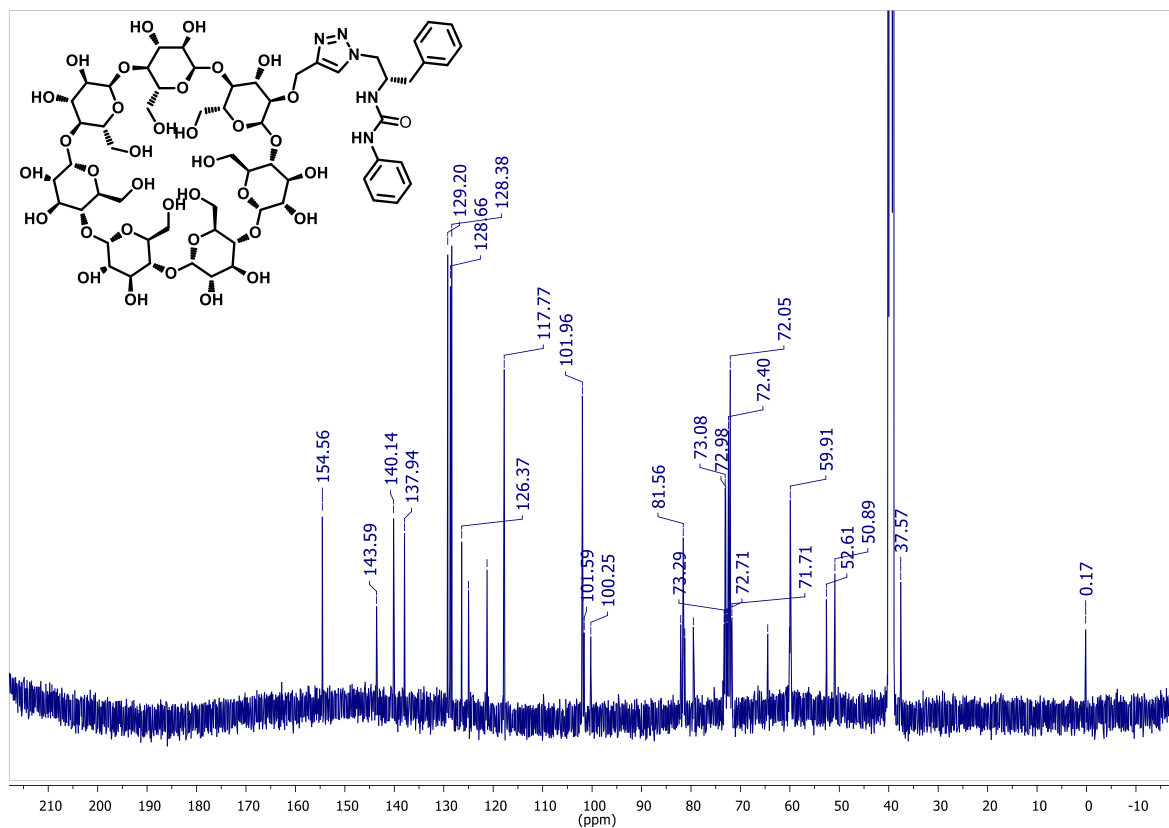


Fig. S2 ^{13}C NMR spectrum of compound **1d**.

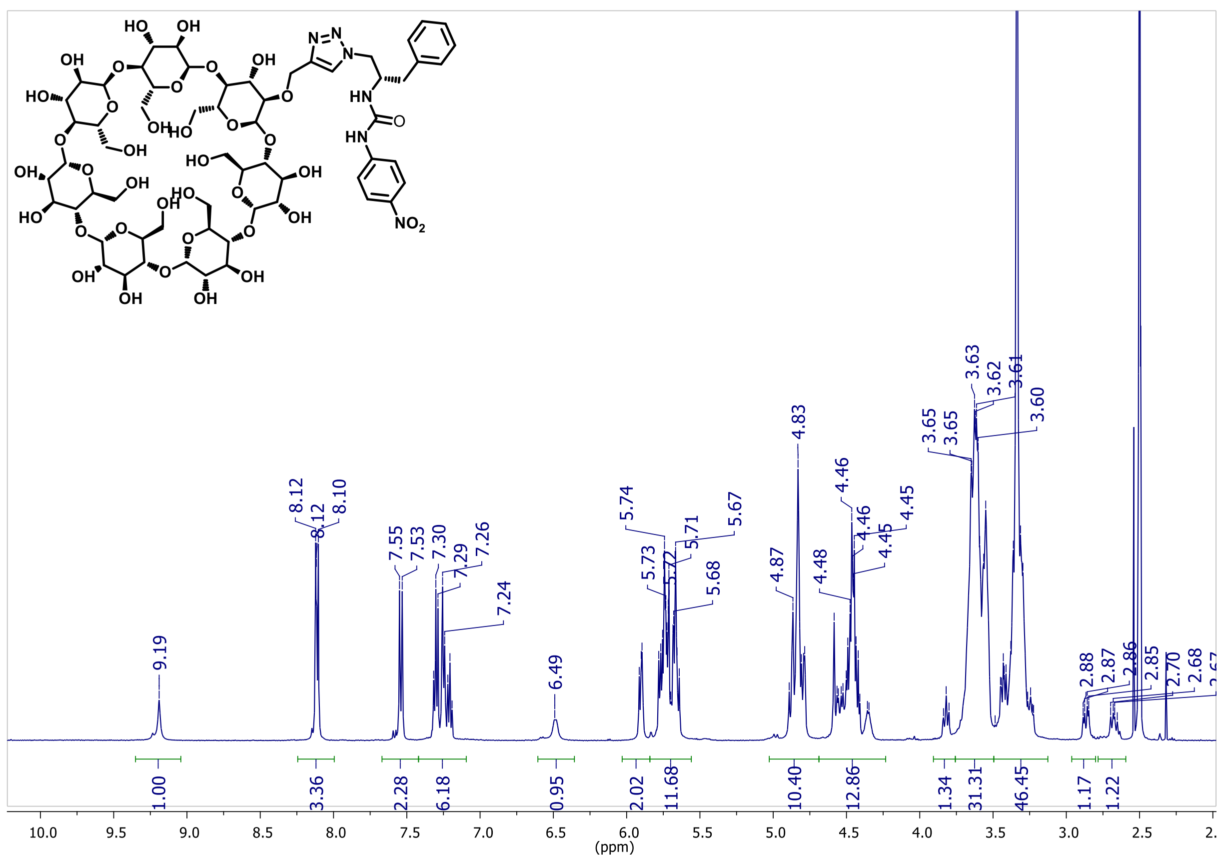


Fig. S3 ^{1}H NMR spectrum of compound **1e**.

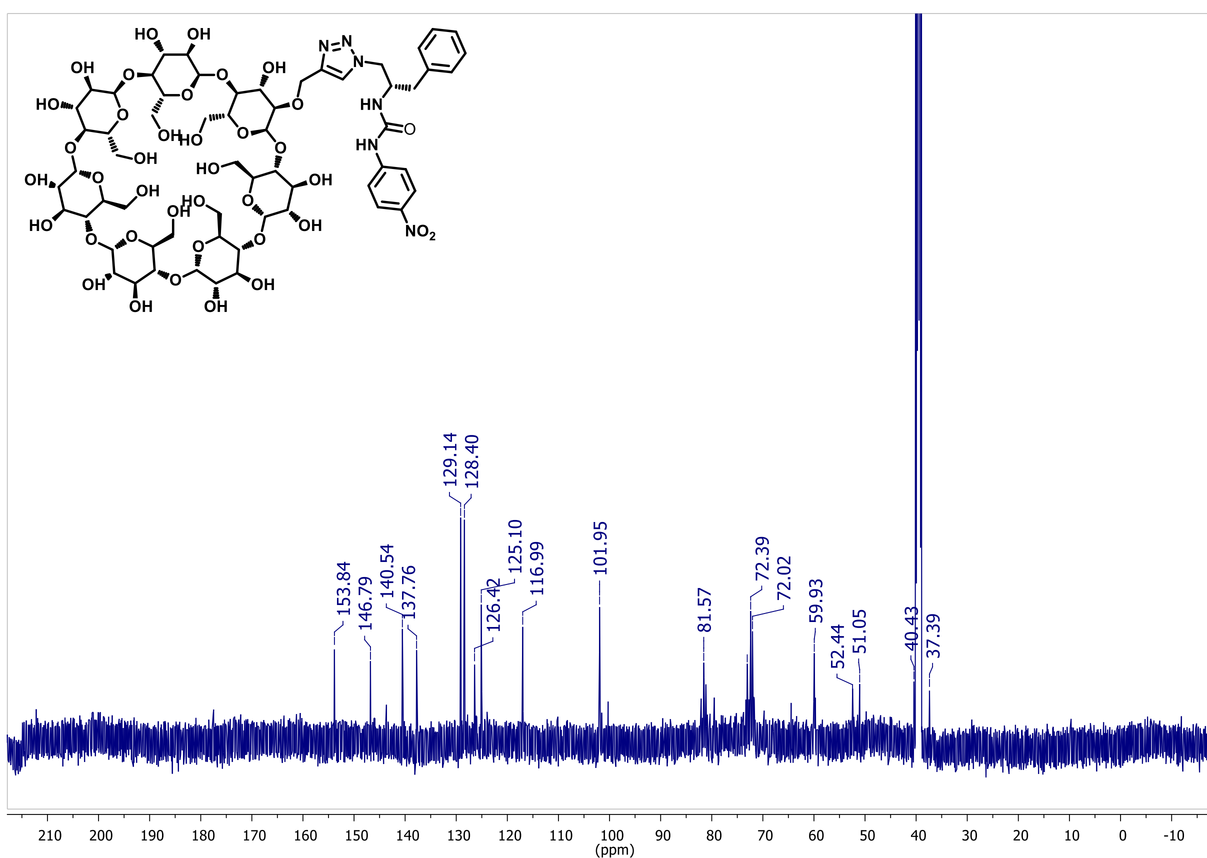


Fig. S4 ^{13}C NMR spectrum of compound **1e**.

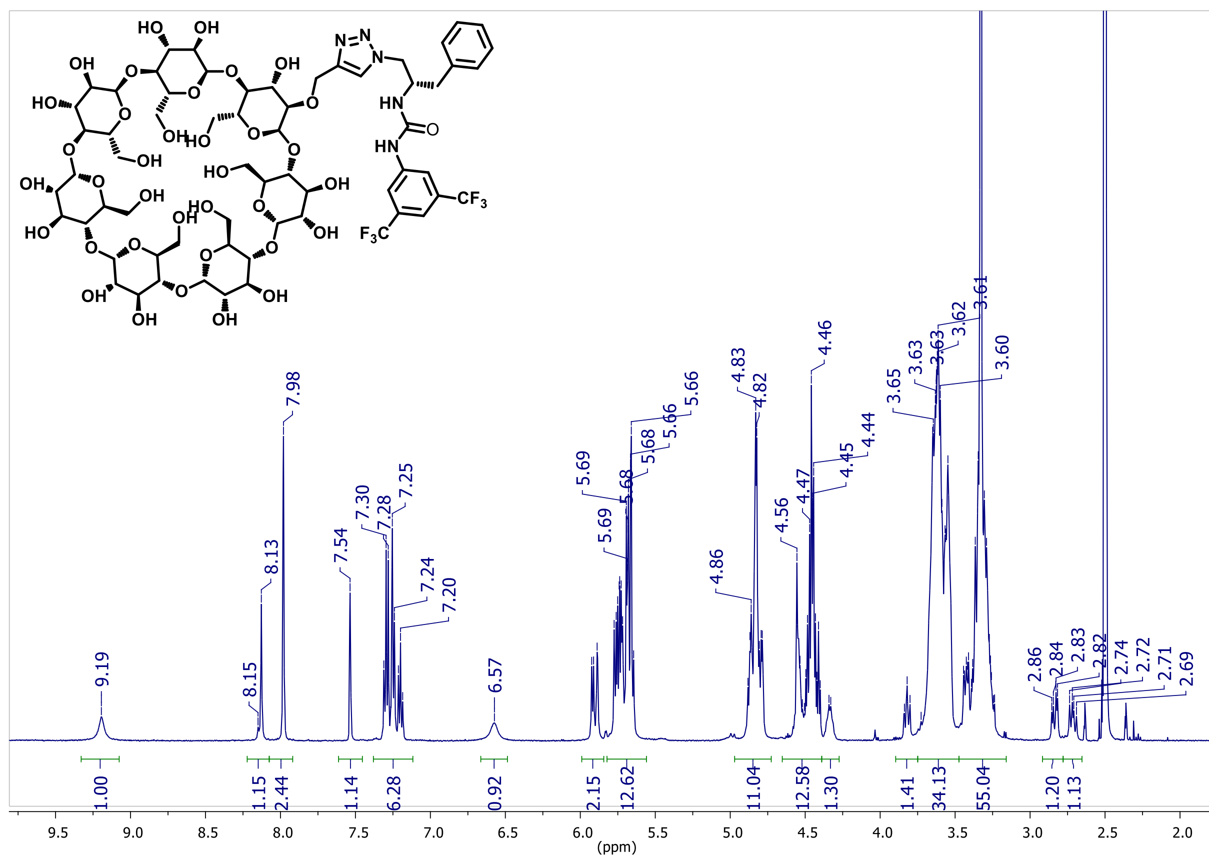


Fig. S5 ^{1}H NMR spectrum of compound **1f**.

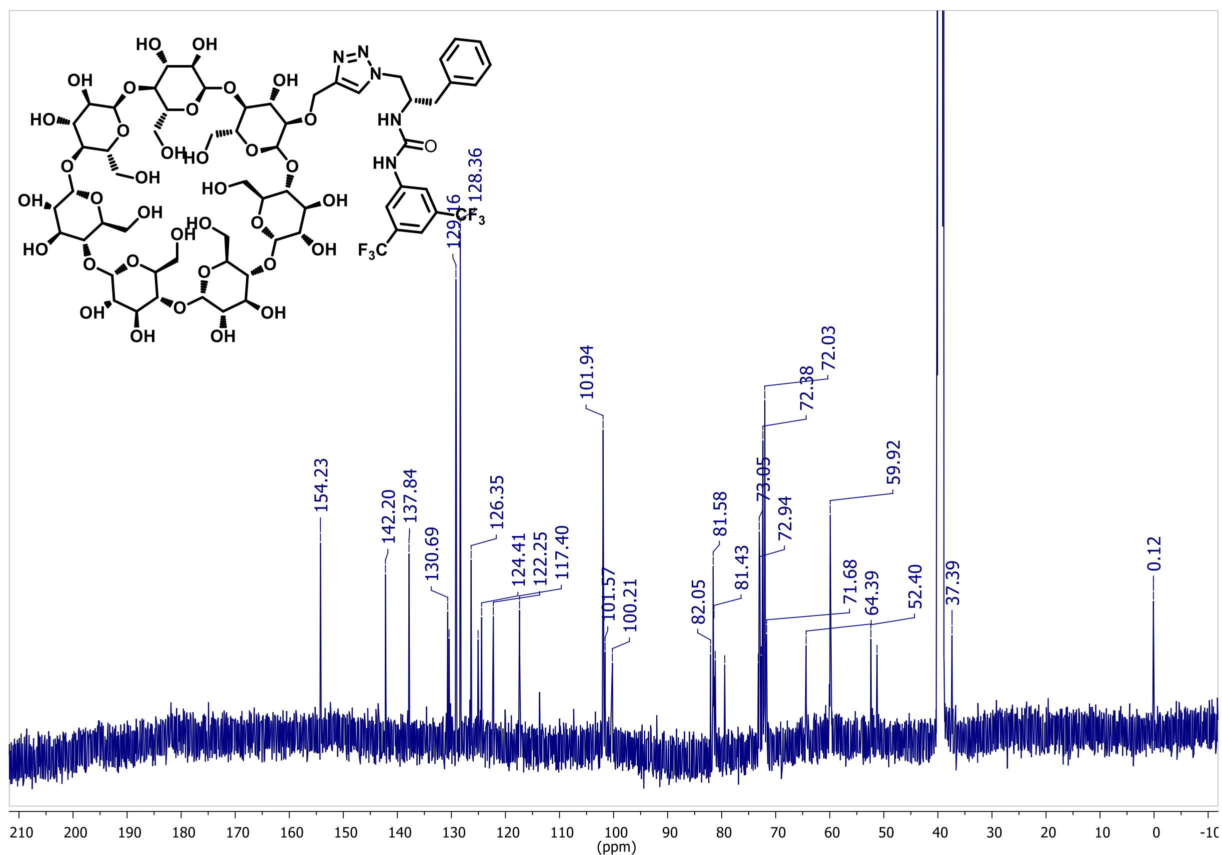


Fig. S6 ^{13}C NMR spectrum of compound **1f**.