Electronic Supplementary Material (ESI) for ChemComm.

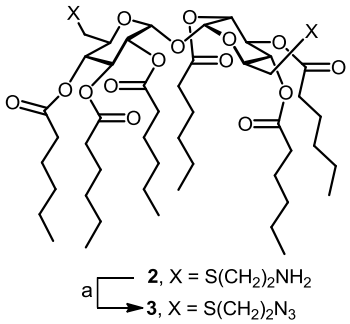
Trehalose-based Siamese twin amphiphiles with tunable self-assembling, DNA nanocomplexing and gene delivery properties

Ana I. Carbajo-Gordillo,[†a] Julio Rodríguez-Lavado, [†b] José Luis Jiménez Blanco, [b] Juan M. Benito, [a] Christophe Di Giorgio, [c] Itziar Vélaz, [d] Concepción Tros de Ilarduya, *[e] Carmen Ortiz Mellet*[b] and José M. García Fernández*[a]

List of contents

General methods. 6,6'-Diazido-6,6'-dideoxy-2,2',3,3',4,4'-hexa-*O*-hexanoyl-α,α'-trehalose (**1**),[1] 2,3,4,2',3',4'-hexa-*O*-hexanoyl-6,6'-bis(2-aminoethylthio)-α,α'-trehalose dihydrochloride (**2**),[1] *N*,*N*-bis(*tert*-butoxycarbonylaminoethyl)propargylamine (**4**),[2] 2-(*N*-*tert*-butoxycarbonylamino)ethyl isothiocyanate (**5**),[3] and 6,6'[4-(2,2-diaminoethylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-6,6'-dideoxy-2,3,4,2',3',4'-hexa-*O*-hexanoyl]-α,α'-trehalose tetrahydrochloride (trehalose gemini facial amphiphile **6**)[1] were obtained according to literature procedures. Optical rotations were measured at 20 ± 2 °C in 1-dm tubes on a Jasco P-2000 polarimeter. Ultraviolet-visible (UV) spectra were recorded in 1-cm tubes on a Jasco V-630 spectrophotometer. Infrared (IR) spectra were recorded on a Jasco ATR MIRacle™ spectrophotometer. ^{1}H (and ^{13}C NMR) spectra were recorded at 300 (75.5), 500 (125.7) MHz with Bruker 300 ADVANCE and 500 DRX. 1D TOCSY, 2D COSY, HMQC and HSQC experiments were used to assist on NMR assignments. Thin-layer chromatography (TLC) was carried out on aluminium sheets coated with Sílica Gel 60 F_{254} Merck with visualization by UV light and by charring with 10% H_2SO_4. With preparative purposes, column chromatography was carried out on Silica Gel 60 F_{254} Merck. Electrospray mass spectra were obtained for samples dissolved in MeOH or H_2O-MeOH mixtures at low μm concentrations. Elemental analyses were carried out at the Instituto de Investigaciones Químicas (Sevilla, Spain) using an elemental analyser Leco CHNS-932 or Leco TruSpec CHN.

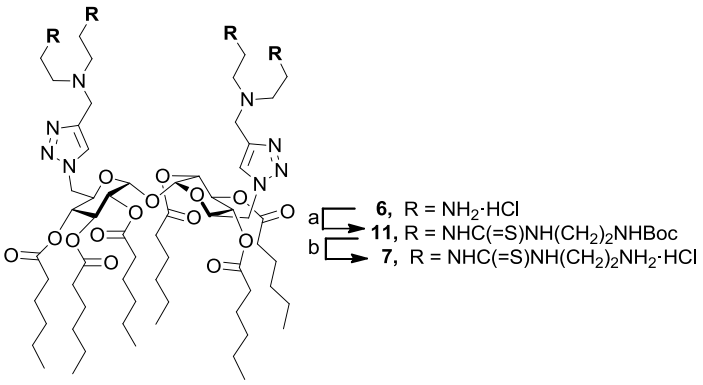
Synthesis of new compounds.



Scheme 1S. Synthesis of the 6,6'-di-(2-azidoethylthio)-α,α'-trehalose derivative **3**. Reagents and conditions: (a) TfN_3, $CuSO_4 \cdot 5\ H_2O$, MeOH, rt, 24 h, 59%.

6,6'-Di-(2-azidoethylthio)-2,3,4,2',3',4'-hexa-*O*-hexanoyl)-α,α'-trehalose (3). To a solution of **2**[1] (956 mg, 0.85 mmol), $CuSO_4 \cdot 5\ H_2O$ (8.5 mg, 0.034 mmol) and $NaHCO_3$ (286 mg, 3.41 mmol) in H_2O (6.25 mL), TfN_3 (10.6 mL) and MeOH (42 mL) were added. The blue mixture was stirred vigorously at rt for 24 h, and concentrated in a rotatory evaporator at rt. The residue was purified by column chromatography (1:8 → 1:4 EtOAc-cyclohexane). Yield: 550 mg (59%). R_f = 0.51 (1:2 EtOAc-cyclohexane); $[\alpha]_D$ = +109.9 (*c* 1.0, DCM). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 5.49 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.6 Hz, H-3), 5.31 (d, 2 H, $J_{1,2}$ = 3.7 Hz, H-1), 5.08 (dd, 2 H, H-2), 5.02 (t, 2 H, $J_{4,5}$ = 9.6 Hz H-4), 3.96 (m, 2 H, H-5), 3.42 (t, 4 H, $^{3}J_{H,H}$ = 7.0 Hz, CH_2N), 2.74 (m, 4 H, CH_2S), 2.63 (m, 4 H, H-6a, H-6b), 2.42-2.10 (m, 12 H, H-2_{Hex}), 1.58 (m, 12 H, H-3_{Hex}), 1.29 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.89 (t, 18 H, $^{3}J_{H,H}$

= 6.0 Hz, H-6_{Hex}). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ = 172.8, 171.7 (CO ester), 91.6 (C-1), 71.3 (C-5), 71.1 (C-2), 69.6 (C-4, C-3), 51.1 (C-6), 33.2 (CH_2N), 32.6 (CH_2S), 34.1 (C-2_{Hex}), 31.2 (C-4_{Hex}), 24.5 (C-3_{Hex}), 22.2 (C-5_{Hex}), 13.9 (C-6_{Hex}). ESIMS: *m/z* = 1123.7 [M + Na]$^+$. Anal. Calcd for $C_{52}H_{88}N_8O_{15}S_2$: C, 56.71; H, 8.05; N, 7.63; S, 5.82. Found: C, 56.79; H, 7.96; N, 7.54; S, 5.68.

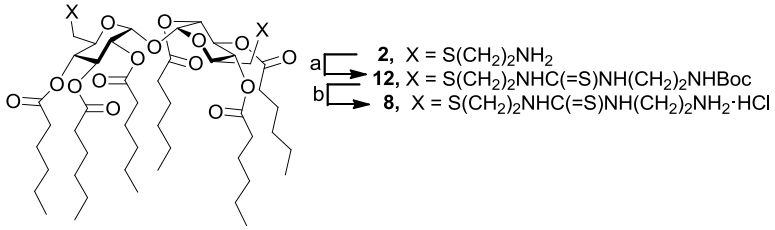


Scheme 2S. Synthesis of trehalose cationic glycolipid derivative **7**. Reagents and conditions: (a) **5**, Et_3N, DCM, rt, 16 h, 69%; (b) TFA, DCM, rt, 94%.

Boc-protected trehalose derivative 11. To a solution of **6**[1] (230 mg, 0.163 mmol) and Et_3N (99 μL, 0.98 mmol) in DCM (17 mL), 2-(*N*-tert-butoxycarbonylamino)ethyl isothiocyanate (**5**;[3] 198 mg, 0.98 mmol) was added and the mixture was stirred at rt overnight. The solvent was removed and the residue was purified by column chromatography (3:1 EtOAc-cyclohexane → 9:1 DCM-MeOH). Yield: 232 mg (69%). R_f = 0.45 (9:1 DCM-MeOH); $[\alpha]_D$ = +30.2 (*c* 1.0, DCM); UV (DCM): λ_{max} = 247 nm (ε_{mM} 32.3); IR: ν_{max} = 3307, 2959, 1752, 1674, 733 cm^{-1}.^{1}H NMR (300 MHz, $CDCl_3$, 303 K): δ = 7.77 (s, 2 H, =CH), 7.25, 7.10 (bs, 4 H, NHCS), 5.46 (bs, 4 H, NHBoc), 5.44 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.7 Hz, H-3), 4.91 (t, 2 H, $J_{4,5}$ = 9.7 Hz, H-4), 4.87 (bs, 2 H, H-1), 4.85 (dd, 2 H, H-2), 4.65 (bs, 4 H, CH_2-triazole), 4.54 (d, 2 H, $J_{6a,6b}$ = 14.0 Hz, H-6a), 4.34 (dd, 2 H, $J_{5,6b}$ = 9.0 Hz, H-6b), 4.07 (bs, 8 H, NCH_2CH_2NHCS), 3.97 (bt, 2 H, H-5), 3.60 (bs, 8 H, NCH_2CH_2NHCS, $NHBocCH_2CH_2NHCS$), 3.30 (m, 8 H, CH_2NHBoc), 2.29 (m, 12 H, H-2_{Hex}), 1.55 (m, 12 H, H-3_{Hex}), 1.41 (s, 36 H, CMe_3), 1.29 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.90 (t, 18 H, $J_{H,H}$ = 6.5 Hz, H-6_{Hex}). ^{13}C NMR (75.5 MHz, $CDCl_3$, 303 K): δ = 183.1 (CS), 173.0-172.1 (CO ester), 156.7 (CO carbamate), 144.2 (C-4 triazole), 128.3 (C-5 triazole), 90.8 (C-1), 79.6 (CMe_3), 69.4 (C-2), 69.3 (C-4), 69.1(C-3), 68.8 (C-5), 53.5 (C-6), 52.3 (NCH_2CH_2NHCS), 50.4 (C-6), 48.2 (CH_2-triazole), 45.9 ($NHBocCH_2CH_2NHCS$), 40.2 (CH_2NHBoc), 33.9 (C-2_{Hex}), 31.2 (C-4_{Hex}), 28.3 (CMe_3), 24.3 (C-3_{Hex}), 22.18 (C-5_{Hex}), 13.9 (C-6_{Hex}). ESIMS: *m/z* = 2135.5 [M + Cu]$^+$. Anal. Calcd for $C_{94}H_{166}N_{20}O_{23}S_4$: C, 54.47; H, 8.07; N, 13.52; S, 6.19. Found: C, 55.54; H, 8.01; N, 13.6; S, 5.97.

Trehalose gemini facial amphiphile 7. Treatment of **11** (160 mg, 0.077 mmol) with 1:1 TFA-DCM (3 mL) and freeze-drying from 10:1 H_2O/0.1 N HCl solution yielded **7**. Yield: 122 mg (94%). $[\alpha]_D$ = +23.9 (*c* 0.89, MeOH); UV (MeOH): λ_{max} = 243 nm (ε_{mM} 34.3); IR: ν_{max} = 3230, 2954, 1752, 1676, 721 cm^{-1}. ^{1}H NMR (300 MHz, CD_3OD): δ = 8.43 (s, 2 H, =CH), 5.49 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.7 Hz, H-3), 5.07 (dd, 2 H, $J_{1,2}$ = 3.9 Hz, H-2), 5.03 (t, 2 H, $J_{4,5}$ = 9.7

Hz, H-4), 4.91 (d, 2 H, H-1), 4.77 (bs, 4 H, CH_2-triazole), 4.67 (dd, 2 H, $J_{5,6a}$ = 2.9 Hz, $J_{6a,6b}$ = 14.7 Hz, H-6a), 4.59 (dd, 2 H, $J_{5,6b}$ = 8.1 Hz, H-6b), 4.15 (ddd, 2 H, H-5), 4.11 (m, 8 H, H-5, C*H*$_2$CH$_2$NH$_2$), 3.89 (t, 8 H, $J_{H,H}$ 6.0 Hz, NCH$_2$C*H*$_2$NHCS), 3.56 (bs, 8 H, C*H*$_2$NH$_2$), 3.23 (t, 8 H, NC*H*$_2$CH$_2$NHCS), 2.50-2.19 (m, 12 H, H-2$_{Hex}$), 1.60 (m, 12 H, H-3$_{Hex}$), 1.33 (m, 24 H, H-4$_{Hex}$, H-5$_{Hex}$), 0.93 (m, 18 H, H-6$_{Hex}$). ^{13}C NMR (75.5 MHz, CD_3OD): δ = 184.5 (CS), 174.2-173.2 (CO), 136.2 (C-4 triazole), 128.1 (C-5 triazole), 90.8 (C-1), 71.1 (C-2), 70.7 (C-3), 70.6 (C-4), 70.4 (C-5), 54.5 (CH_2NH_2), 51.7 (C-6), 48.0 (CH_2-triazole), 42.6 (NH$_2$CH$_2$*C*H$_2$NHCS), 40.7 (N*C*H$_2$CH$_2$NHCS), 40.4 (CH_2NH_2), 34.9 (C-2$_{Hex}$), 32.5-32.4 (C-4$_{Hex}$), 25.6-25.5 (C-3$_{Hex}$), 23.4 (C-5$_{Hex}$), 14.3 (C-6$_{Hex}$). ESIMS: *m/z* = 1733.3 [M + Cu]$^+$, 867.1 [M + Cu]$^{2+}$. Anal. Calcd for $C_{74}H_{134}N_{20}O_{15}S_4$·4 HCl: C, 48.89; H, 7.65; N, 15.41; S, 7.05. Found: C, 48.81; H, 7.35; N, 15.10; 7.17.

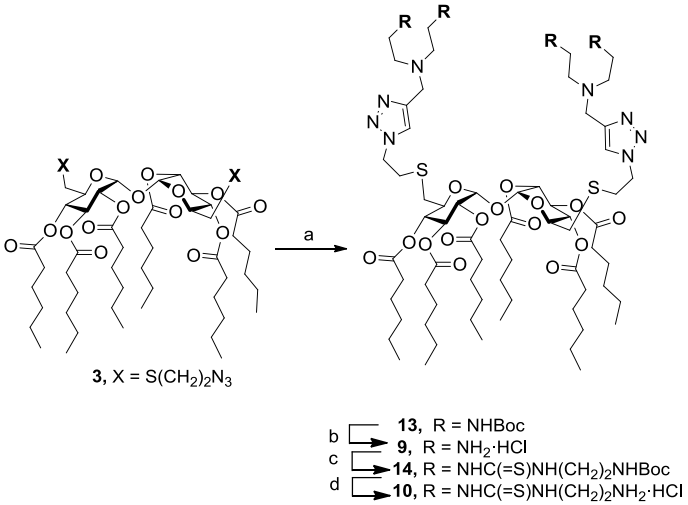


Scheme 3S. Synthesis of trehalose cationic glycolipid derivative **8**. Reagents and conditions: (a) **5**, Et_3N, DCM, rt, 16 h, 73%; (b) TFA, DCM, rt, quant.

Bis[6-(2-(*N´*-(2-(*N*-*tert*-Butoxycarbonylamino)ethyl)thioureido)ethylthio)-2,3-di-*O*-hexanoyl]-α,α'-trehalose (12). To a solution of **2**[1] (140 mg, 0.125 mmol) and Et_3N (51.8 μL, 0.374 mmol) in DCM (12 mL), 2-(*N*-tert-butoxycarbonylamino)ethyl isothiocyanate (**5**;[3] 76 mg, 0.374 mmol) was added and the reaction mixture was stirred at rt for 36 h. The reaction mixture was washed with aqueous diluted HCl (2 x 20 mL), dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography (1:1 EtOAc-cyclohexane). Yield: 185 mg (73%). R_f = 0.63 (2:1 EtOAc-cyclohexane); $[\alpha]_D$ = +62.0 (*c* 1, DCM); UV (DCM): λ_{max} = 248 nm (ε_{mM} 73.1); IR: ν_{max} = 2957, 1749, 1685 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.01, 6.91 (bs, 4 H, NHCS), 5.47 (t, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, 2 H, H-3), 5.4 (bd, 2 H, H-1), 5.16 (bs, 2 H, NHBoc), 5.05 (dd, 2 H, $J_{1,2}$ = 3.9 Hz, H-2), 4.95 (t, 2 H, H-4), 3.94 (m, 2 H, H-5), 3.62 (m, 4 H, SCH$_2$C*H*$_2$NHCS), 3.56 (bs, 4 H, C*H*$_2$NHCS), 3.29 (m, 4 H, C*H*$_2$NHBoc), 2.84 (t, 4 H, $J_{H,H}$ = 7.0 Hz, SC*H*$_2$CH$_2$NHCS), 2.58 (m, 4 H, H-6a, H-6b), 2.34-2.31 (m, 12 H, H-2a$_{Hex}$, H-2b$_{Hex}$), 1.55 (m, 12 H, H-3$_{Hex}$), 1.42 (s, 18 H, C*Me*$_3$), 1.28 (m, 24 H, H-4$_{Hex}$, H-5$_{Hex}$), 0.88 (t, 9 H, $J_{H,H}$ = 6.5 Hz, H-6$_{Hex}$). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ = 173-172.3 (CO ester), 157.8 (CO carbamate), 91.3 (C-1), 80.1 (*C*Me$_3$),72.1 (C-5), 71.0 (C-4), 69.8 (C-2), 69.6 (C-3), 45.5 (*C*H$_2$NHCS), 43.7 (SCH$_2$*C*H$_2$NHCS), 39.9 (*C*H$_2$NHBoc), 32.7 (*C*H$_2$S, C-6), 34.2-34.0 (C-2$_{Hex}$), 31.1 (C-4 $_{Hex}$), 28.5 (C*Me*$_3$) 24.3 (C-3$_{Hex}$), 22.1 (C-5$_{Hex}$), 13.8 (C-6$_{Hex}$). ESIMS: *m/z* = 1487.5 [M + Cl]$^-$. Anal. Calcd for $C_{68}H_{120}N_6O_{19}S_4$: C, 56.17; H, 8.32; N, 5.78; S, 8.82. Found: C, 56.24; H, 8.26; N, 5.59; S, 8.51.

Trehalose gemini facial amphiphile 8. Treatment of **12** (119 mg, 0.082 mmol) with 1:1 TFA-DCM (1.6 mL) and freeze-drying from 10:1 H_2O/0.1 N HCl solution afforded **8**. Yield: 119 mg (100%). $[\alpha]_D$ = +62.1 (*c* 0.8,

MeOH); UV (MeOH): λ_{max} = 245 nm (ε_{mM} 24.1); IR: ν_{max} = 2955, 1750 cm^{-1}. ^{1}H NMR (500 MHz, CD_3OD): δ 5.51 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.5 Hz, H-3), 5.41 (d, 2 H, H-1), 5.12 (dd, 2 H, $J_{1,2}$ = 4.2 Hz, H-2), 5.07 (t, 2 H, $J_{4,5}$ = 9.5 Hz, H-4), 4.00 (m, 2 H, H-5), 3.83 (m, 4 H, NC$H_2$$CH_2$NHCS), 3.66 (bs, 4 H, C$H_2$$CH_2$S), 3.16 (m, 4 H, NCH_2CH_2NHCS), 2.77 (m, 4 H, CH_2CH_2S), 2.75 (m, 4 H, H-6a, H-6b), 2.42-2.26 (m, 12 H, H-2_{Hex}), 1.65-1.54 (m, 12 H, H-3_{Hex}), 1.32 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.92 (m, 18 H, H-6_{Hex}). ^{13}C NMR (125 MHz, CD_3OD): δ = 173.2-172.4 (CO), 90.8 (C-1), 71.1 (C-5), 70.0 (C-4), 69.8 (C-3), 69.6 (C-2), 44.3 (*C*$H_2$$CH_2$S), 40.8 (*C*$H_2$$CH_2$NHCS), 39.4 ($CH_2$*C*$H_2$NHCS), 32.7 ($CH_2$S), 31.8 (C-6), 33.6, 33.8, 31.2 (C-2_{Hex}), 31.1 (C-4 $_{Hex}$), 24.2 (C-3_{Hex}), 22.1 (C-5_{Hex}), 13.1 (C-6_{Hex}). ESIMS: *m/z* = 1253.7 $[M]^+$. Anal. Calcd for $C_{58}H_{104}N_6O_{15}S_4$·2 HCl: C, 52.51; H, 8.05; N, 6.33; S, 9.67. Found: C, 52.34; H, 7.89; N, 6.14; S, 9.33.



Scheme 4S. Synthesis of trehalose cationic glycolipid derivatives **9** and **10**. Reagents and conditions: (a) **4**, CuI·$P(OEt)_3$, DIPEA, 9:1 ^{t}BuOH-acetone, 85 ºC, 16 h, 63%; (b) TFA, DCM, rt, quant, (c) **5**, Et_3N, DCM, rt, 16 h, 80%; (d) TFA, DCM, rt, 96%.

Boc-protected trehalose derivative 13. To a solution of diazide **3** (281 mg, 0.25 mmol) and *N*,*N*-bis(*tert*-butoxycarbonylaminoethyl)propargylamine **4**[2] (192 mg, 0.56 mmol) in a 9:1 ^{t}BuOH-acetone mixture (15 mL), CuI·$P(OEt)_3$ (17.8 mg, 0.05 mmol) and DIPEA (102 μL, 0.6 mol) were added and the reaction mixture was refluxed for 16 h at 85 °C. The catalyst was filtered off, the solvent was removed and the residue was purified by column chromatography (1:1 EtOAc-cyclohexane → 15:1 DCM-MeOH). Yield: 281 mg (63%). R_f = 0.67 (9:1 DCM-MeOH); $[\alpha]_D$ = +55.1 (*c* 1.0, DCM). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.57 (s, 2 H, =CH), 5.47 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.5 Hz, H-3), 5.28 (d, 2 H, $J_{1,2}$ = 4.0 Hz, H-1), 5.17 (bs, 2 H, NHBoc), 5.02 (dd, 2 H, H-2), 4.98 (t, 2 H, $J_{4,5}$ = 9.8 Hz, H-4), 4.48 (t, 4 H, $^3J_{H,H}$ = 6.9 Hz, CH_2N), 3.91 (m, 2 H, H-5), 3.80 (s, 4 H, CH_2-triazole), 3.21 (bd, 8 H, CH_2NHBoc), 3.06 (m, 4 H, CH_2S), 2.55 (t, 8 H, $^3J_{H,H}$ = 5.6 Hz, C$H_2$$CH_2$NHBoc), 2.49 (m, 4 H, H-6a, H-6b), 2.38-2.13 (m, 12 H, H-2_{Hex}), 1.55 (m, 12 H, H-3_{Hex}), 1.43 (s, 36 H, CMe_3), 1.27 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.87 (m, 18 H, H-6_{Hex}). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ = 172.5 (CO ester), 156.2 (CO carbamate), 144.4 (C-4 triazole), 123.4 (C-5 triazole), 88.2 (C-1), 79.1 (*C*Me_3), 71.3 (C-4), 70.9 (C-2), 69.6 (C-5), 69.4 (C-3), 53.1 (CH_2NHBoc), 50.0 (CH_2N

triazole), 38.2 (CH_2-triazole), 33.6 (CH_2S), 34.1 (C-2_{Hex}), 33.1 (C-6), 31.3 (C-4_{Hex}), 28.6 (CMe_3), 24.3 (C-3_{Hex}), 22.3 (C-5_{Hex}), 13.8 (C-6_{Hex}). ESIMS: *m/z* = 1806.8 $[M + Na]^+$. Anal. Calcd for $C_{58}H_{98}N_8O_{15}S_2$: C, 57.85; H, 8.14; N, 7.94; S, 4.54. Found: C, 57.76; H, 8.19; N, 7.76; S, 4.32.

Trehalose gemini facial amphiphile 9. Treatment of **13** (412 mg, 0.23 mmol) with 1:1 TFA-DCM (6 mL) at rt for 30 min followed by evaporation of the solvents and freeze-drying from a diluted HCl solution afforded **9**. Yield: 359 mg (100%). $[\alpha]_D$ = +45.3 (*c* 1.0, MeOH). ^{1}H NMR (300 MHz, CD_3OD): δ = 8.09 (s, 2 H, =CH), 5.49 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.7 Hz, H-3), 5.36 (d, 2 H, $J_{1,2}$ = 4.0 Hz, H-1), 5.09 (dd, 2 H, H-2), 5.04 (t, 2 H, $J_{4,5}$ = 9.7 Hz, H-4), 4.61 (t, 4 H, $^3J_{H,H}$ = 6.5 Hz, CH_2N), 4.00 (m, 2 H, H-5), 3.90 (s, 4 H, CH_2-triazole), 3.15 (m, 12 H, CH_2S, $CH_2CH_2NH_2$), 2.81 (t, 8 H, $^3J_{H,H}$ = 6.0 Hz, C$H_2CH_2NH_2$), 2.63 (m, 4 H, H-6a, H-6b), 2.42-2.21 (m, 12 H, H-2_{Hex}), 1.58 (m, 12 H, H-3_{Hex}), 1.31 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.91 (m, 18 H, H-6_{Hex}). ^{13}C NMR (75.5 MHz, CD_3OD): δ = 174.8, 173.8 (CO), 143.4 (C-4 triazole), 126.1 (C-5 triazole), 92.3 (C-1), 72.6 (C-4), 72.3 (C-2), 71.3 (C-5), 71.1 (C-3), 52.0 (CH_2NH_2), 51.2 (CH_2N-triazole), 47.2 (CH_2-triazole), 38.2 (C$H_2CH_2NH_2$), 35.1 (C-2_{Hex}), 34.9 (CH_2S), 34.0 (C-6), 32.5 (C-4_{Hex}), 25.6 (C-3_{Hex}), 23.1 (C-5_{Hex}), 14.0 (C-6_{Hex}). ESIMS: *m/z* = 1384.0 $[M]^+$, 692.5 $[M + H]^{2+}$. Anal. Calcd for $C_{66}H_{118}N_{12}O_{15}S_2$·4 HCl: C, 51.82; H, 8.04; N, 10.99; S, 4.19. Found: C, 51.59; H, 8.12; N, 10.68; S, 3.82.

Boc-protected trehalose derivative 14. To a solution of **9** (182 mg, 0.12 mmol) and Et_3N (47 μL, 0.71 mmol) in DCM (10 mL), 2-(*N-tert*-butoxycarbonylamino)ethyl isothiocyanate (**5**;[3] 144 mg, 0.71 mmol) was added, and the mixture was stirred at rt for 48 h. The mixture was washed with 0.1 N HCl (2 x 10 mL) and the organic phase was dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography (4:1 EtOAc-cyclohexane → 9:1 DCM-MeOH). Yield: 184 mg (80%). R_f = 0.51 (9:1 DCM-MeOH); $[\alpha]_D$ = +26.6 (*c* 0.7, DMF); UV (DMF-H_2O 0.4% v/v): λ_{max} = 247 nm (ε_{mM} 21.9). ^{1}H NMR (300 MHz, DMSO-d_6): δ = 8.05 (s, 2 H, =CH), 7.59, 7.36 (bs, 4 H, NHCS), 6.80 (s, 2 H, NHBoc), 5.35 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.5 Hz, H-3), 5.25 (d, 2 H, $J_{1,2}$ = 3.7 Hz, H-1), 5.07 (dd, 2 H, H-2), 5.02 (t, 2 H, $J_{4,5}$ = 9.5 Hz, H-4), 4.49 (bt, 4 H, $^3J_{H,H}$ = 6.4 Hz, CH_2N), 3.88 (m, 2 H, H-5), 3.78 (s, 2 H, CH_2-triazole), 3.60-3.50 (m, 16 H, CH_2CH_2NHCS, CH_2CH_2NHBoc), 3.04 (m, 12 H, CH_2S, CH_2NHBoc), 2.73-2.50 (m, 12 H, NCH_2CH_2NHCS, H-6a, H-6b), 2.42-2.10 (m, 12 H, H-2_{Hex}), 1.58 (m, 12 H, H-3_{Hex}), 1.39 (s, 36 H, CMe_3), 1.25 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.85 (m, 18 H, H-6_{Hex}). ^{13}C NMR (75.5 MHz, DMSO-d_6): δ = 172.5, 172.0 (CO ester), 155.6 (CO carbamate), 141.2 (C-4 triazole), 124.5 (C-5 triazole), 90.5 (C-1), 78.7 (CMe$_3$), 70.7 (C-4), 70.5 (C-2), 69.3 (C-5), 69.0 (C-3), 52.0 (NCH$_2$CH$_2$NHCS), 49.0 (CH_2N-triazole), 47.5 (CH_2-triazole), 43.3 (NCH$_2C$H$_2$NHCS), 42.1 (CH_2CH_2NHBoc), 33.5 (C-2_{Hex}), 32.5 (CH_2S), 32.0 (C-6), 31.9 (C-4_{Hex}), 28.5 (CMe_3), 23.5 (C-3_{Hex}), 21.5 (C-5_{Hex}), 13.5 (C-6_{Hex}). ESIMS: *m/z* = 2255.9 $[M + Cu]^+$. Anal. Calcd for $C_{98}H_{174}N_{20}O_{23}S_6$: C, 53.67; H, 8.00; N, 12.77; S, 8.77. Found: C, 53.49; H, 7.88; N, 12.54; S, 8.39.

Trehalose gemini facial amphiphile 10. Treatment of **14** (193 mg, 0.23 mmol) with 1:1 TFA-DCM (4 mL) at rt for 30 min followed by evaporation of the solvents and freeze-drying from a diluted HCl solution afforded **10**. Yield: 167 mg (96%). $[\alpha]_D$ = +28.6 (*c* 1.0, MeOH); UV (MeOH): λ_{max} = 244 nm (ε_{mM} 47.3). ^{1}H NMR (300 MHz, CD_3OD): δ = 8.43 (s, 2 H, =CH), 5.50 (t, 2 H, $J_{2,3} = J_{3,4}$ = 9.5 Hz, H-3), 5.35 (d, 2 H, $J_{1,2}$ = 3.7 Hz, H-1), 5.12 (dd, 2 H,

H-2), 5.07 (t, 2 H, $J_{4,5}$ = 9.5 Hz, H-4), 4.49 (bt, 4 H, $^{3}J_{H,H}$ = 6.2 Hz, CH_2N), 4.07 (bs, 8 H, NCH_2CH_2NHCS), 3.98 (m, 2 H, H-5), 3.87 (bt, 8 H, CH_2NH_2), 3.22 (bt, 8 H, NCH_2CH_2NHCS), 2.67 (m, 4 H, H-6a, H-6b), 2.32 (m, 12 H, H-2_{Hex}), 1.59 (m, 12 H, H-3_{Hex}), 1.32 (m, 24 H, H-4_{Hex}, H-5_{Hex}), 0.91 (m, 18 H, H-6_{Hex}). ^{13}C NMR (75.5 MHz, CD_3OD): δ = 174.8, 173.8 (CO), 143.4 (C-4 triazole), 126.1 (C-5 triazole), 92.3 (C-1), 72.6 (C-4), 72.3 (C-2), 71.3 (C-5), 71.1 (C-3), 52.0 (CH_2NH_2), 51.2 (CH_2N-triazole), 47.2 (CH_2-triazole), 38.2 ($CH_2CH_2NH_2$), 35.1 (C-2Hex), 34.9 (CH_2S), 34.0 (C-6), 32.5 (C-4_{Hex}), 25.6 (C-3_{Hex}), 23.1 (C-5Hex), 14.0 (C-6_{Hex}). ESIMS: *m/z* = 1793 $[M + H]^+$, 896.9 $[M + 2H]^{2+}$. Anal. Calcd for $C_{78}H_{142}N_{20}O_{15}S_6 \cdot 4$ HCl: C, 48.33; H, 7.59; N, 14.45; S, 9.93. Found: C, 47.97; H, 7.22; N, 14.11; S, 9.55.

Formulation of pDNA nanocomplexes. The quantity of trehalose gemini facial amphiphile **6**-**10** used in each formulation was calculated according to the desired DNA concentration, the N/P ratio, the molecular weight and the number of protonable nitrogen atoms in the corresponding cationic derivative. Typically, pDNA was diluted in HEPES (10 mM, pH 7.4) to the desired final concentration as specified hereinafter, and then the desired amount of CD derivative was dispersed in this solution from a stock solution in DMSO (typically 1-10 mM). The resulting mixture (with a final DMSO content below 1% in all cases) was instantly vortexed thoroughly and the complexes were incubated for 1 h prior to subjecting them to characterization or transfection experiments.

Electrophoresis mobility shift assay (EMSA). Each gemini facial amphiphile:pDNA nanocomplex (20 µL, 0.4 µg of pDNA) was submitted to electrophoresis for about 30 min under 150 V through a 0.8% agarose gel in TAE 1X (Tris-acetate-EDTA) buffer and stained by spreading GelRed® Nucleic Acid Stain (Biotium). The DNA was then visualized after photographing on an Alphaimager Mini UV transilluminator. The plasmid integrity in each sample was confirmed by electrophoresis after decomplexation with sodium dodecyl sulfate (SDS, 8%).

pDNA condensation/protection assays. Nanocomplex samples (50 µL of 50 µg/mL preparations) were electrophoresed for 30 min under 150 mV in 0.8% agarose gel. For protection assays, DNAase I (1U/µg pDNA) was added to each sample and stirred for 30 min at 37 ºC. 20 µL of 0.25 M EDTA was added to inactivate DNAase and the sample was vortexed and incubated for 5 min. 20 µL of 25% SDS was then added and further incubated for 5 min. Samples were electrophoresed as described above. Plasmid integrity was compared with free pDNA both treated and untreated.

Particle size and ζ-potential measurements. The size of the CDplexes was measured by dynamic light scattering (DLS), and the overall charge by "Mixed Mode Measurement" phase analysis light scattering (M3-PALS) measurements using a Zetasizer Nano ZS (Malvern Instruments). All measurements were performed in HEPES 10 mM, pH 7.4, in triplicate. Size results are given as volume distribution of the major population by the mean diameter with its standard deviation (Tables S1-S3). Additional comparative DLS measumets were conducted at pH 7.4 and 5.0 with the nanocomplexes formulated with compounds **6** and **10** at N/P 20, as

representative examples. The results confirm that shifting from neutral to acidic pH significantly affect the stability of the nanocomplexes, as inferred from the observed increases in hydrodynamic dimeters and the higher polydispersity (Fig. S13).

Transmission Electron Microscopy (TEM). Formvar-carbon coated grids previously made hydrophilic by glow discharge were placed on top of small drops of the nanocomplexes (DNA 303 μM phosphate) prepared as describe above using N/P 20 ratios. After 1-3 min, grids were negatively stained with a few drops of 1% aqueous solution of uranyl acetate. The grides were then dried and observed with a Philips CM12 electron microscope working under standard conditions. All these experimentes were reproduces twice on each formulation.

pH Buffering capacity. pH Buffering capacities were potentiometrically determined using an adapted acid-base titration procedure over a pH range between ca. 2 and 9. In a typical experiment, the appropriate amount of the hydrochloride salt of each compound (**6**, **7**, **9** or **10**) was accurately weighted to prepare a 50 mM solution (based on the protonable amino groups) in 150 mM NaCl aqueous solution. 1-mL of the solution was transferred to the titration cell and the pH was adjusted (if necessary) below 2 by adding the aliquots of a 0.10 M HCl solution. Then, 10-μL aliquots aq. 0.10 M NaOH (standardized using 2% w/v aq. potassium hydrogen phthalate) were successively added and the pH value was after each addition with a conventional pH-meter. The process was repeated until the pH shifted over 9. A solution of aq. NaCl (150 mM) was used as blank under the same experimental conditions. The buffering capacity is graphically determined as the NaOH amount (meq) consumed to switch pH in the biologically relevant window (from 5 to 7; Fig. S14). The titration curve for compound 8 could not be determined due to its limited solubility under the experimental conditions of the assay.

***In vitro* transfection activity.** The procedure for *in vitro* transfection assays was the same for COS-7, HepG2, BNL-CL2 and RAW 264.7 cell lines. Cells were seeded in medium in 48-well plates (Iwaki Microplate, Japan), and incubated for 24 h at 37 °C in 5% CO_2. The medium was removed and 0.3 mL of complete medium (activated FBS) and 0.2 mL of complexes (containing 1 μg of pDNA) were added to each well. After 4 h incubation the medium was replaced for complete medium and the cells were further incubated for 48 h. Cells were washed with phosphate-buffered saline (PBS) and lysed with 100 μL of Reporter Lysis Buffer (Promega, Madison, WI, USA) at room temperature for 10 min, followed by a freeze-thaw cycle. 20 μL of the supernatant was assayed for total luciferase activity using the luciferase assay reagent (Promega), according to the manufacturer's protocol. A luminometer (Sirius-2, Berthold Detection Systems, Innogenetics, Diagnóstica y Terapéutica, Barcelona, Spain) was used to measure luciferase activity. The protein content of the lysates was measured by de DC protein Assay Reagent (Bio-Rad, Hercules, CA, USA) using bovine serum albumin as the standard. The data were expressed as nanograms of luciferase (based on a standard curve for luciferase activity) per milligram of protein. Citokine levels were obtained using the kit BD OptEIA ELISA Set (Pharmingen,

San Diego, CA, USA) for IL-12 p40 following the manufacturer's instructions. Values were calculated based on a standard curve. Samples were analyzed in a plate spectrophotometer Power Wave XS and a data processor KC junior, BioTek®.

Cell viability, Alamar blue® assay. Cell viability was quantified by a modified Alamar blue® assay (Invitrogen). Briefly, 1 mL of 10% (v/v) Alamar blue dye in complete medium was added to each well 48 h post-transfection. After 2.5 h of incubation at 37 ºC, 200 μL of the supernatant was assayed by measuring the absorbance at 570 and 600 nm. Cell viability (as percentage of control cells; Figure 4S) was calculated according to the formula (A_{570} - A_{600}) of treated cells × 100/(A_{570} - A_{600}) of control cells. Data for compounds **6**-**10** and controls are presented in Fig. S16A.

Cell viability, MTT assay. For MTT assay $3 \cdot 10^3$ cells (COS-7, HepG2, BNL-CL2 or RAW 264.7) were seeded in 96 well plates and incubated at 37 °C, in a humidified atmosphere containing 5% CO_2 till confluence. 50 μL per well of the selected dilution of the vector (6-10 or the controls PEI, Lipofectamine 3000® or ADM70) formulations were then added. The plates were incubated for 24 hours at 37 °C, 5% CO_2. 10 μL per well of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) were then added at the final concentration of 5 mg/mL (MTT can be diluted in sterile PBS or in medium) and after an incubation of 6 h 100 μL per well of a lysant (10% SDS and 0.01 M HCl in H_2O) were added. Plates were sealed and incubated overnight at 37 °C, 5% CO_2. The absorbance was read at 620 or 540 nm after shaking of the plates. The rate of toxicity/survival was calculated respect to the corresponding cell line incubated in absence of the compound. Data for compounds **6**-**10** and controls are presented in Fig. S16B.

***In vivo* transfection activity.** Female Balb-c mice (6-8 weeks of age, 20-25 grams weigh) were purchased from Harlan Ibérica Laboratories. All animals were studied in accordance with guidelines established by Directive 86/609/EEC and with the approval of the Committee on Animal Research at the University of Navarra. Individual mice in groups of eight were injected via the tail vein with 200 μL of CDplexes containing 60 μg of pCMV-Luc VR1216 plasmid DNA at N/P 5 and 10. Naked DNA was injected as control. Twenty four hours after injection the mice were sacrificed. The liver, heart, lungs and spleen were collected and washed with cold PBS. The organs were homogenized with 1 mL lysis buffer using a homogenizer at 5000 rpm (Mini-Beadbeater; BioSpec Products, Inc., Bartlesville, OK, USA) and centrifuged at 10000 rpm for 3 min. 20 μL of the supernatant were analysed for luciferase activity following the same procedure as for *in vitro* assays.

Statistical Analysis. Statistical analyses were performed using SPSS software from SPSS Inc. (Chicago, IL, USA). The analysis of the transfection efficiency of CDplexes was performed with a two-tailed unpaired Student´s t-test. $P < 0.05$ was considered statistically significant.

NMR Spectra.

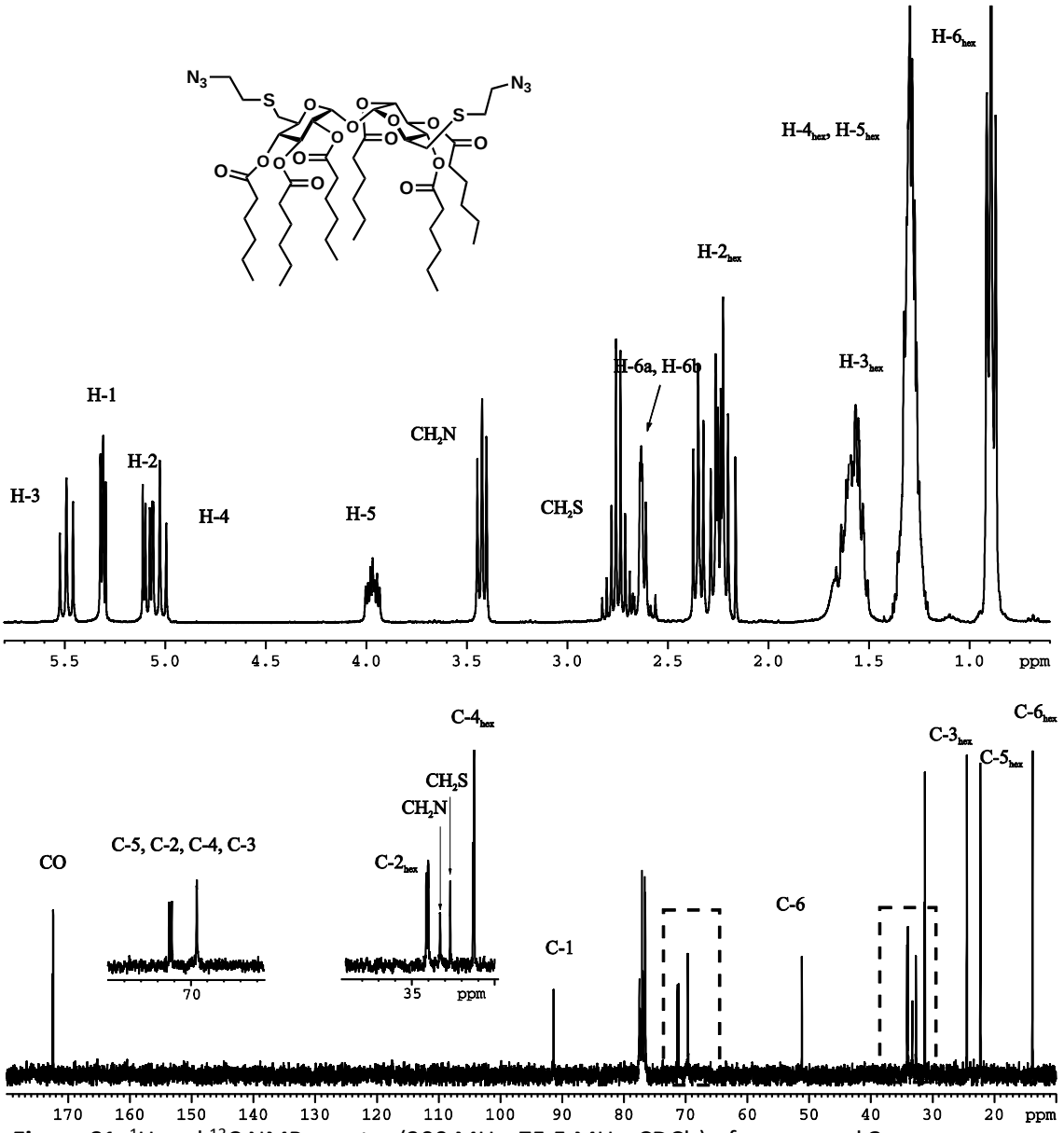


Figure S1. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, $CDCl_3$) of compound **3**.

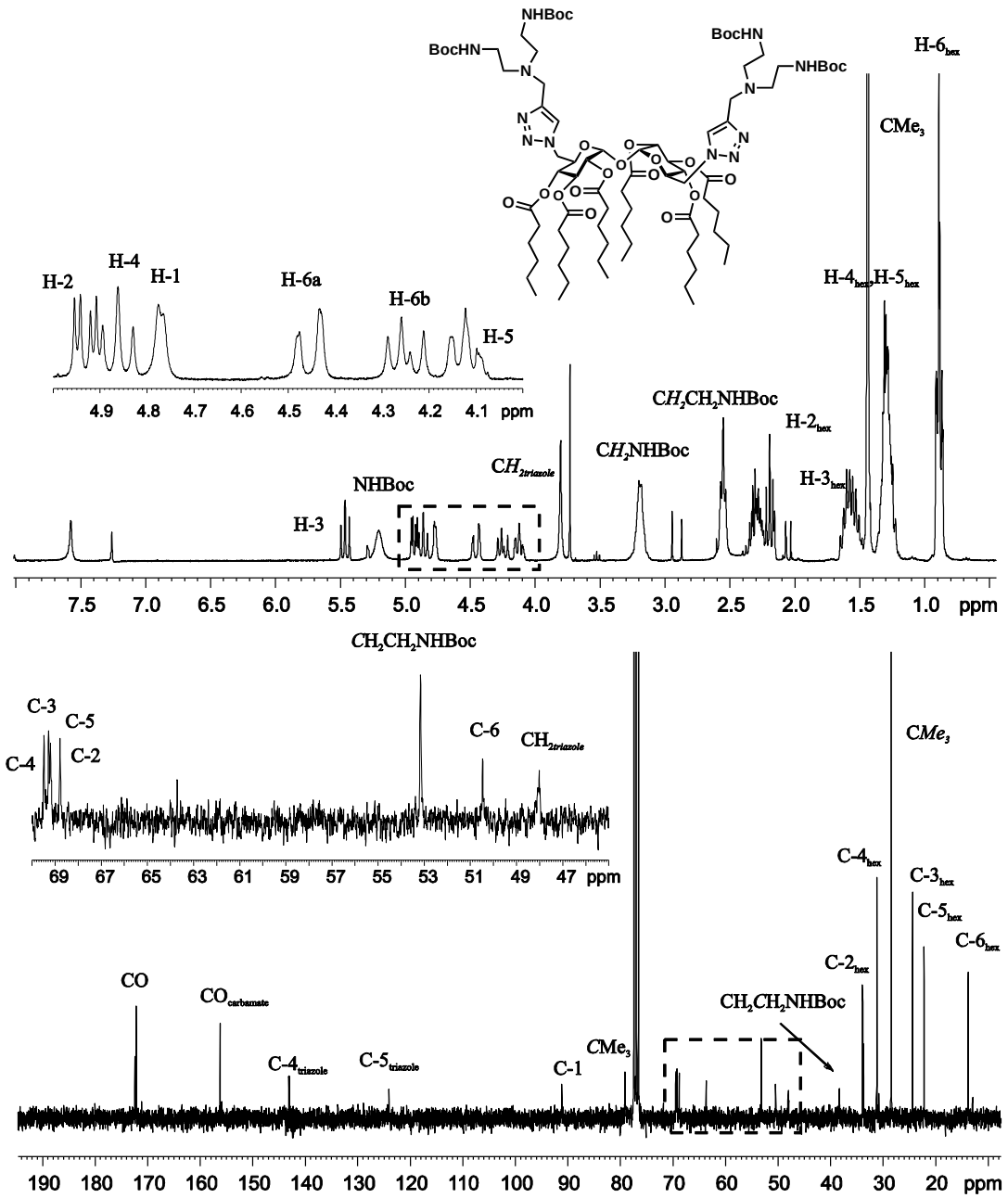


Figure S2. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, $CDCl_3$) of compound **11**.

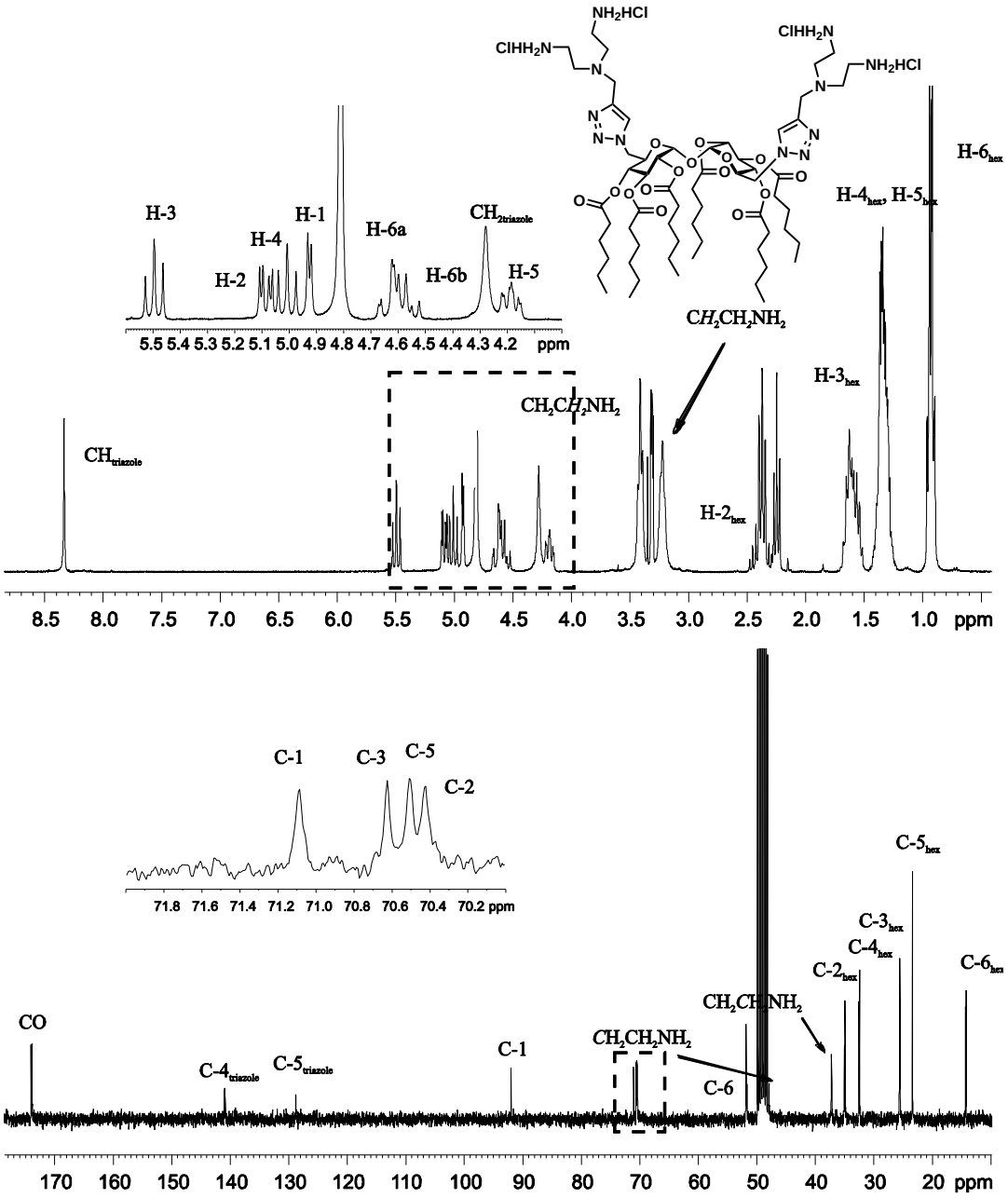


Figure S3. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, CD_3OD) of compound **6**.

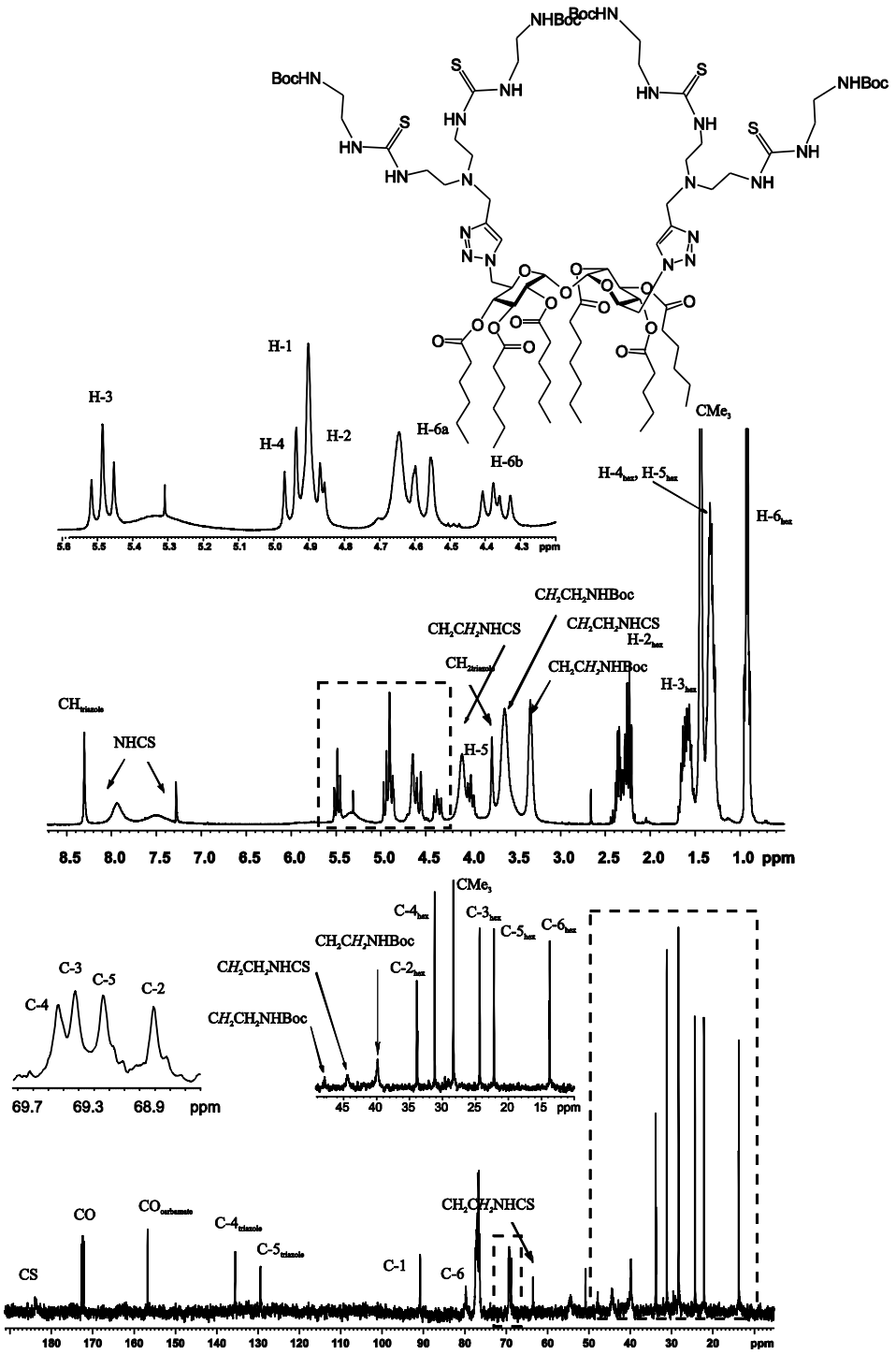


Figure S4. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, $CDCl_3$) of compound **11**

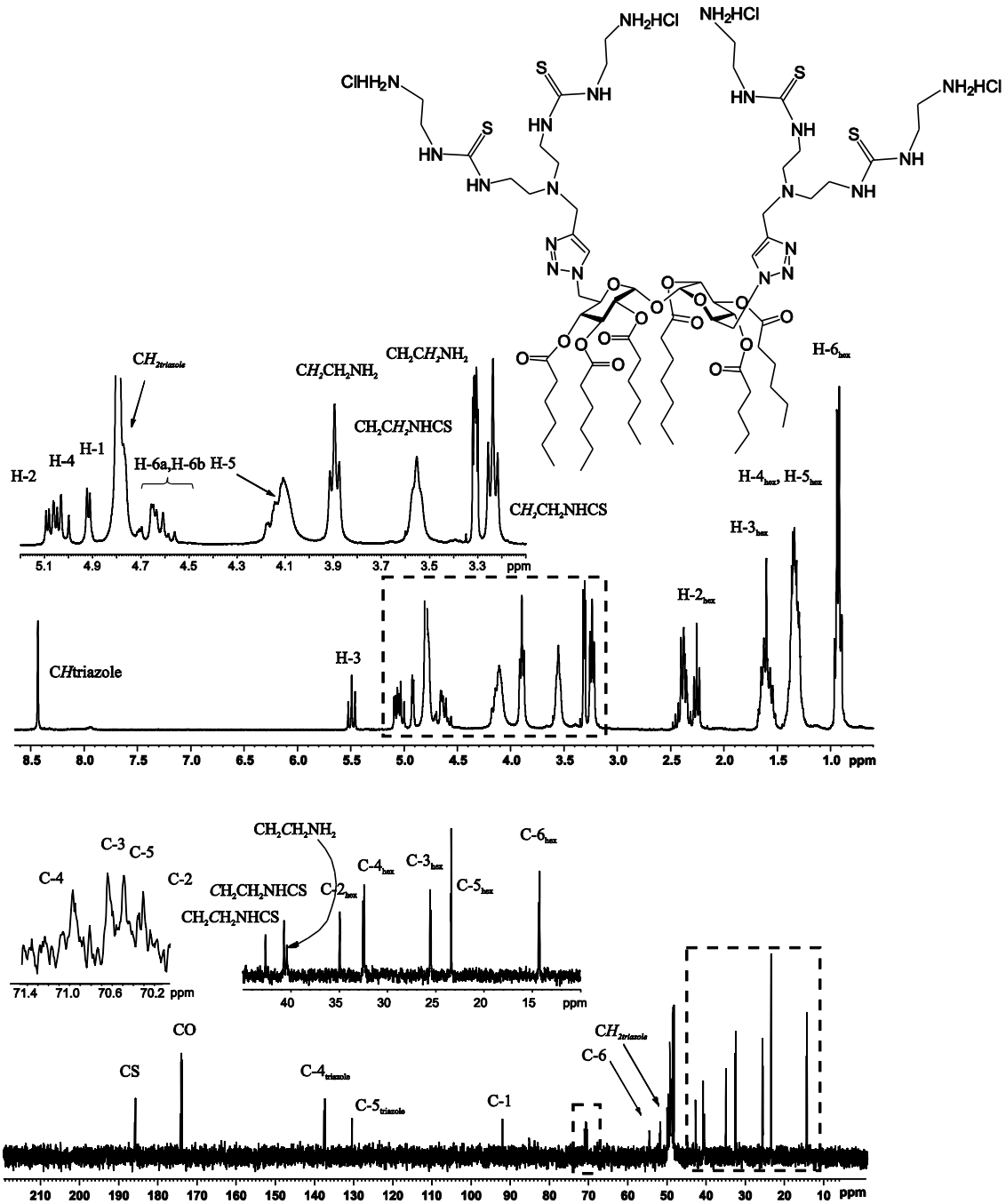


Figure S5. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, CD_3OD) of compound **7**.

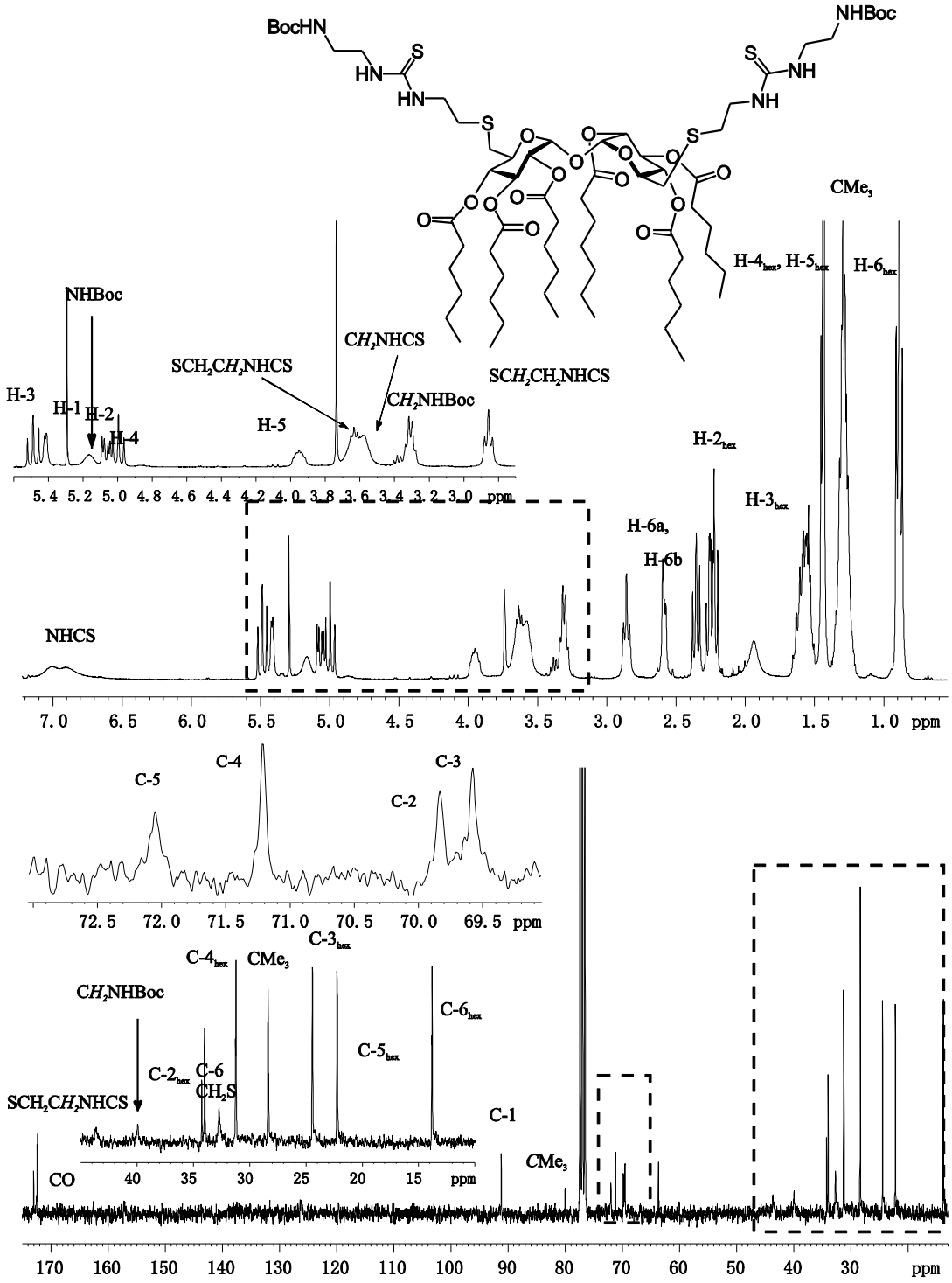


Figure S6. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, $CDCl_3$) of compound **12**

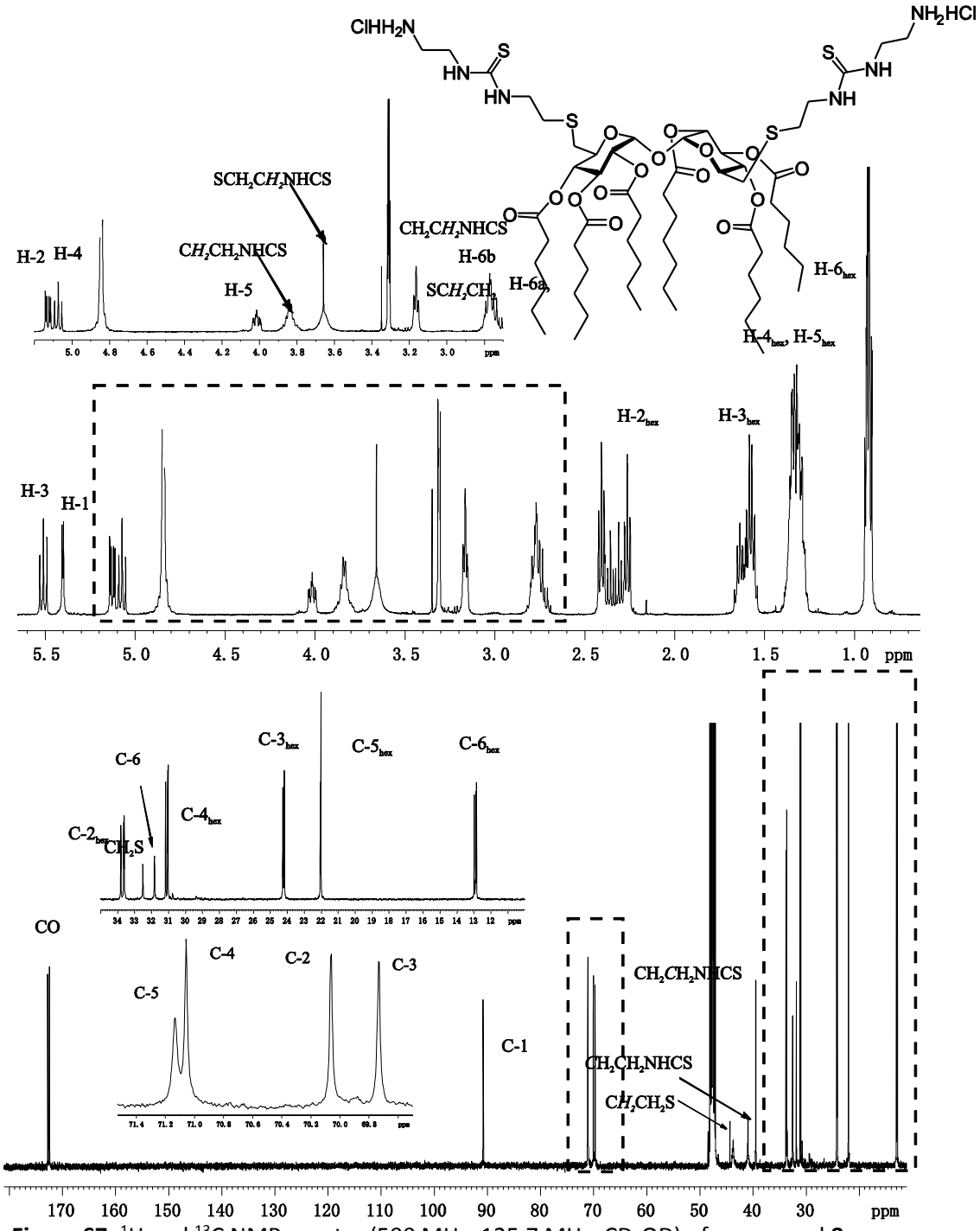


Figure S7. ^{1}H and ^{13}C NMR spectra (500 MHz, 125.7 MHz, CD_3OD) of compound **8**

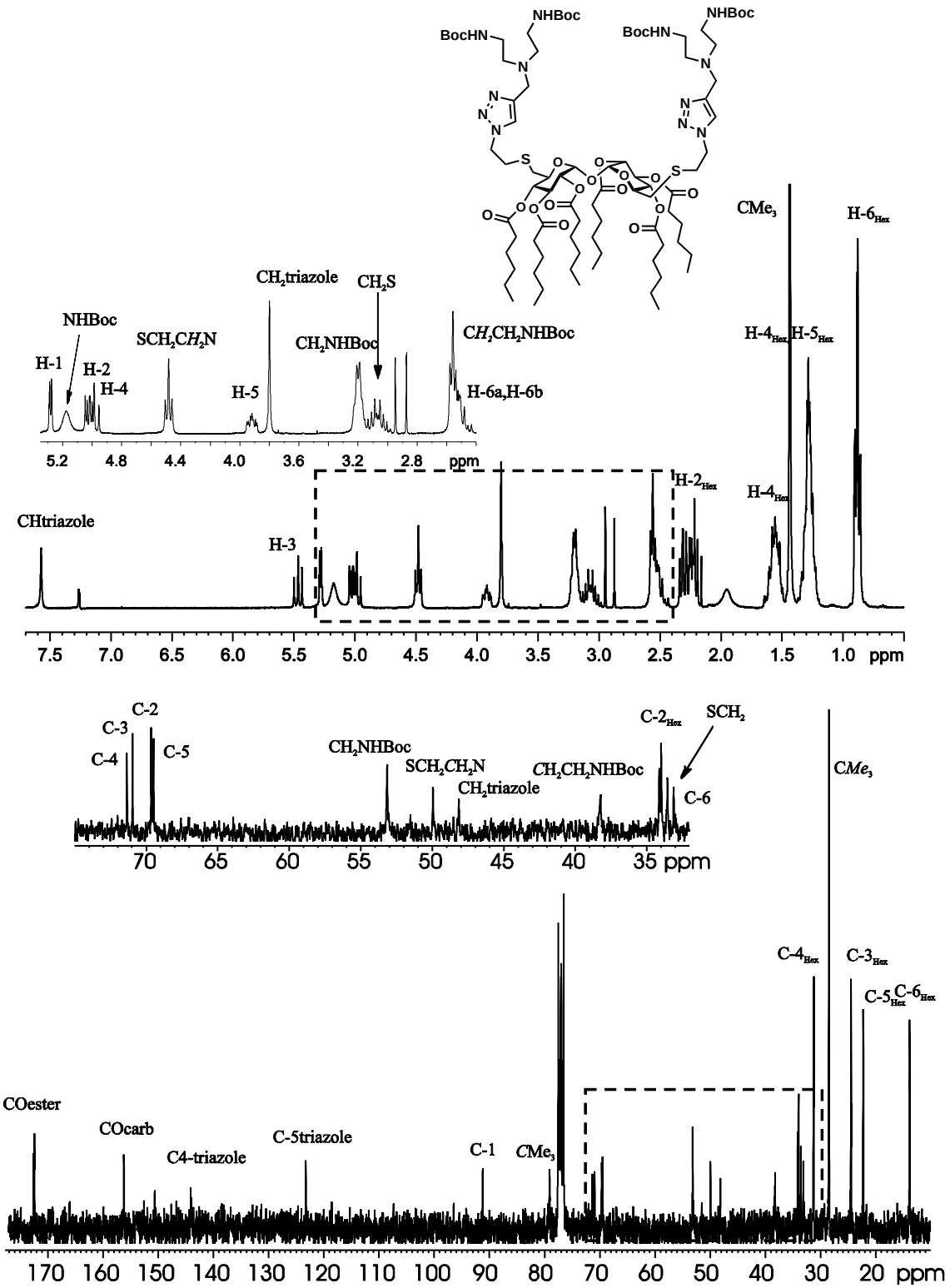


Figure S8. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, CD_3OD) of compound **13**.

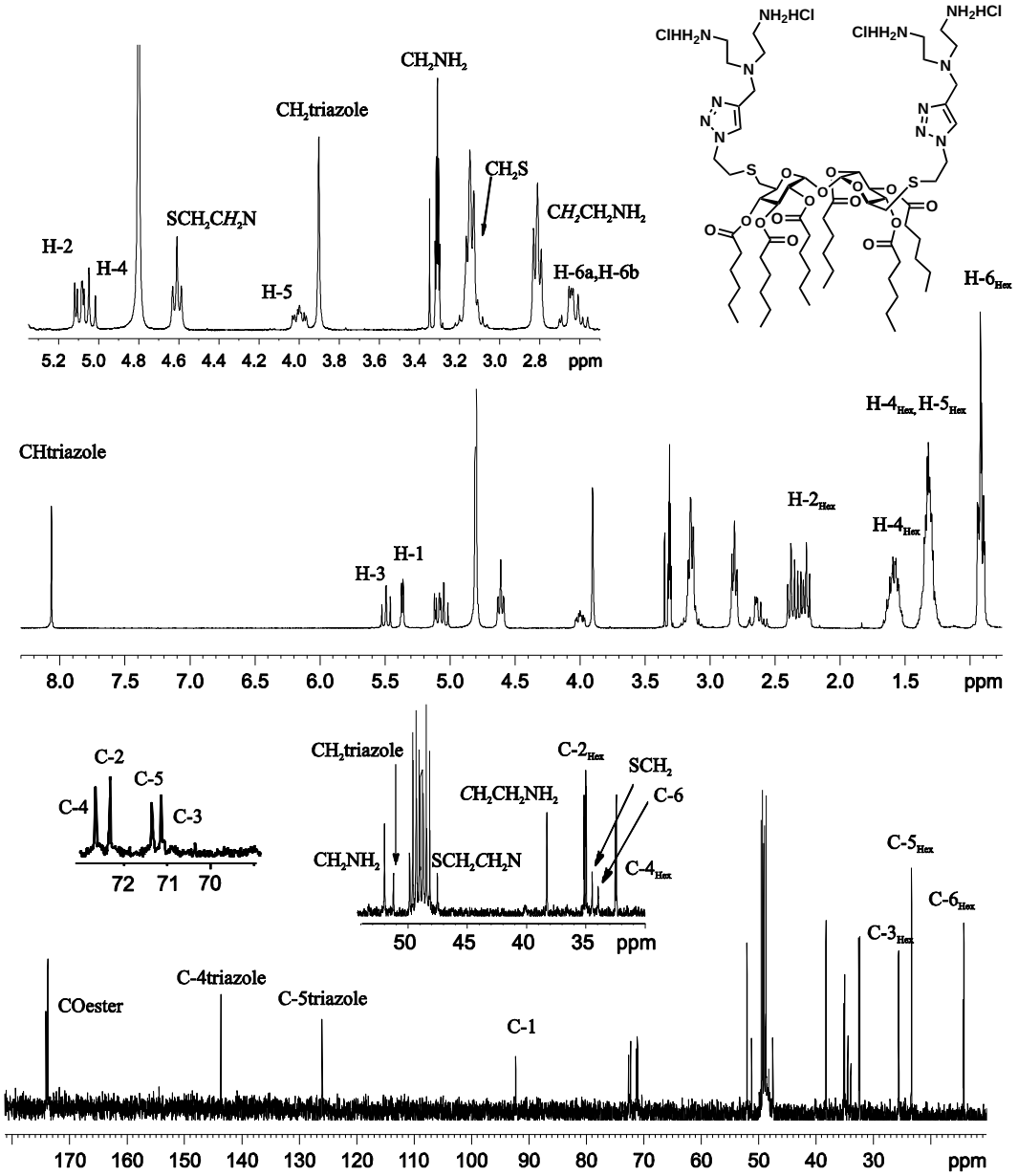


Figure S9. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, $CDCl_3$) of compound **9**.

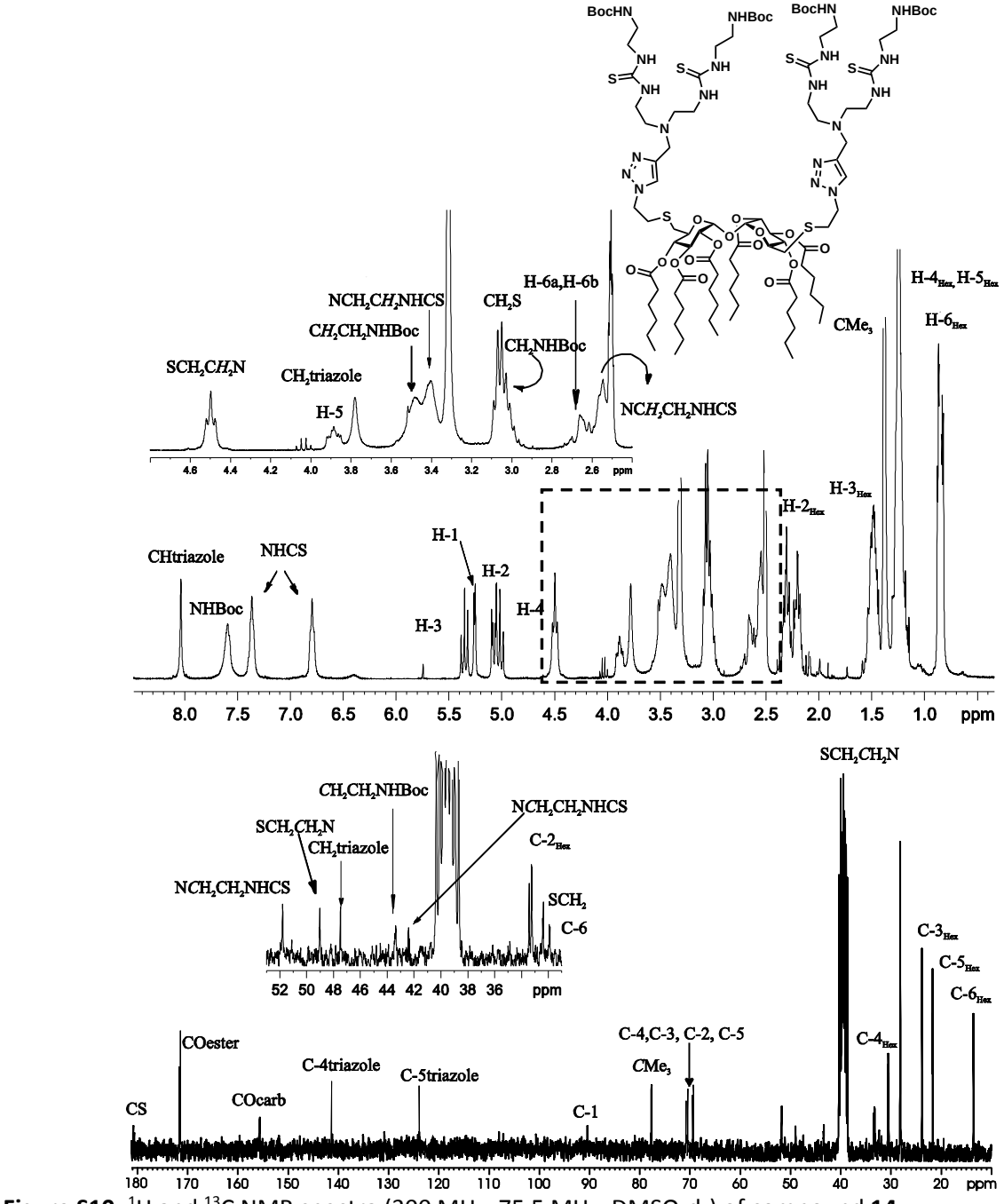


Figure S10. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, DMSO-d_6) of compound **14**.

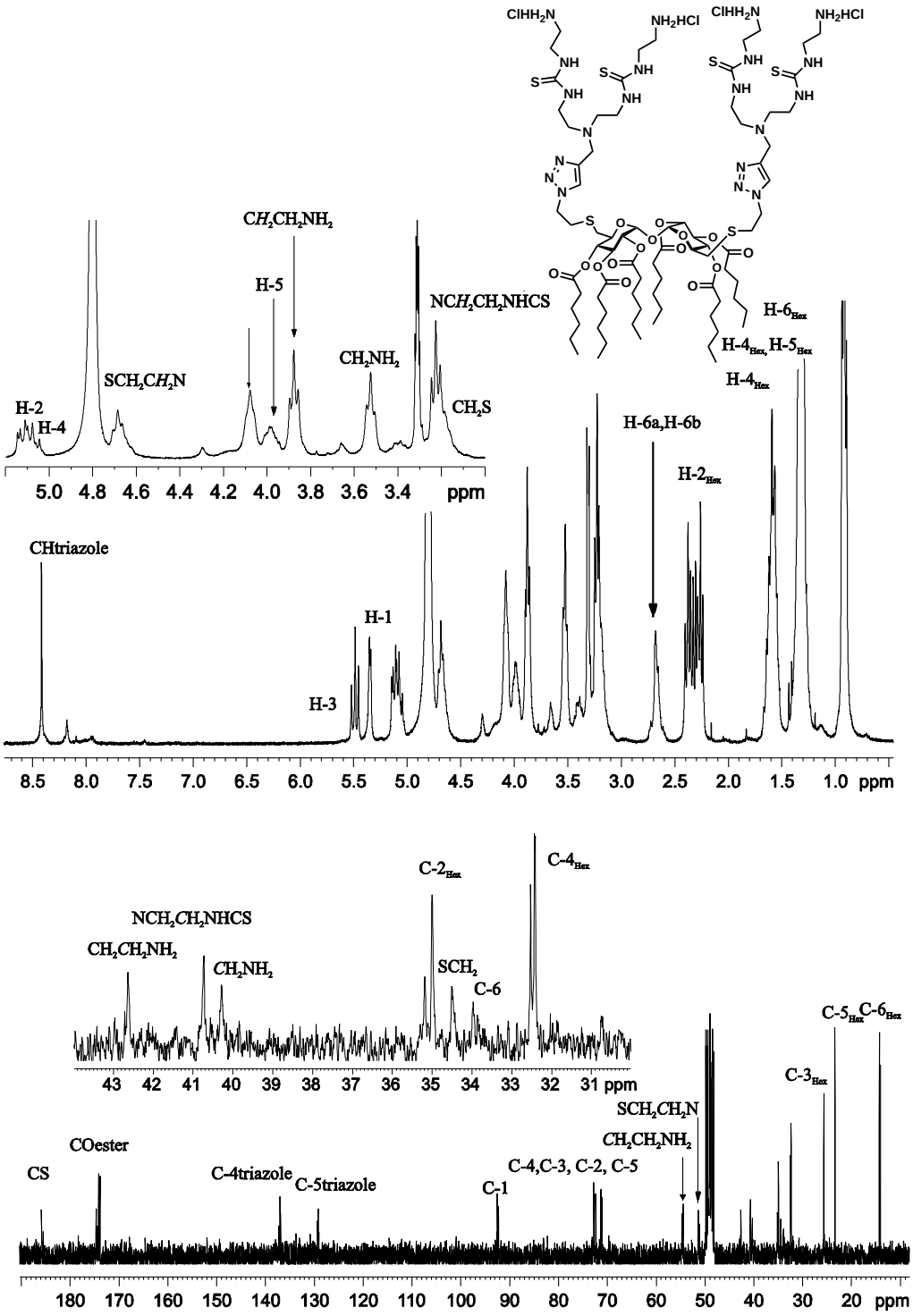


Figure S11. ^{1}H and ^{13}C NMR spectra (300 MHz, 75.5 MHz, CD_3OD) of compound **10**.

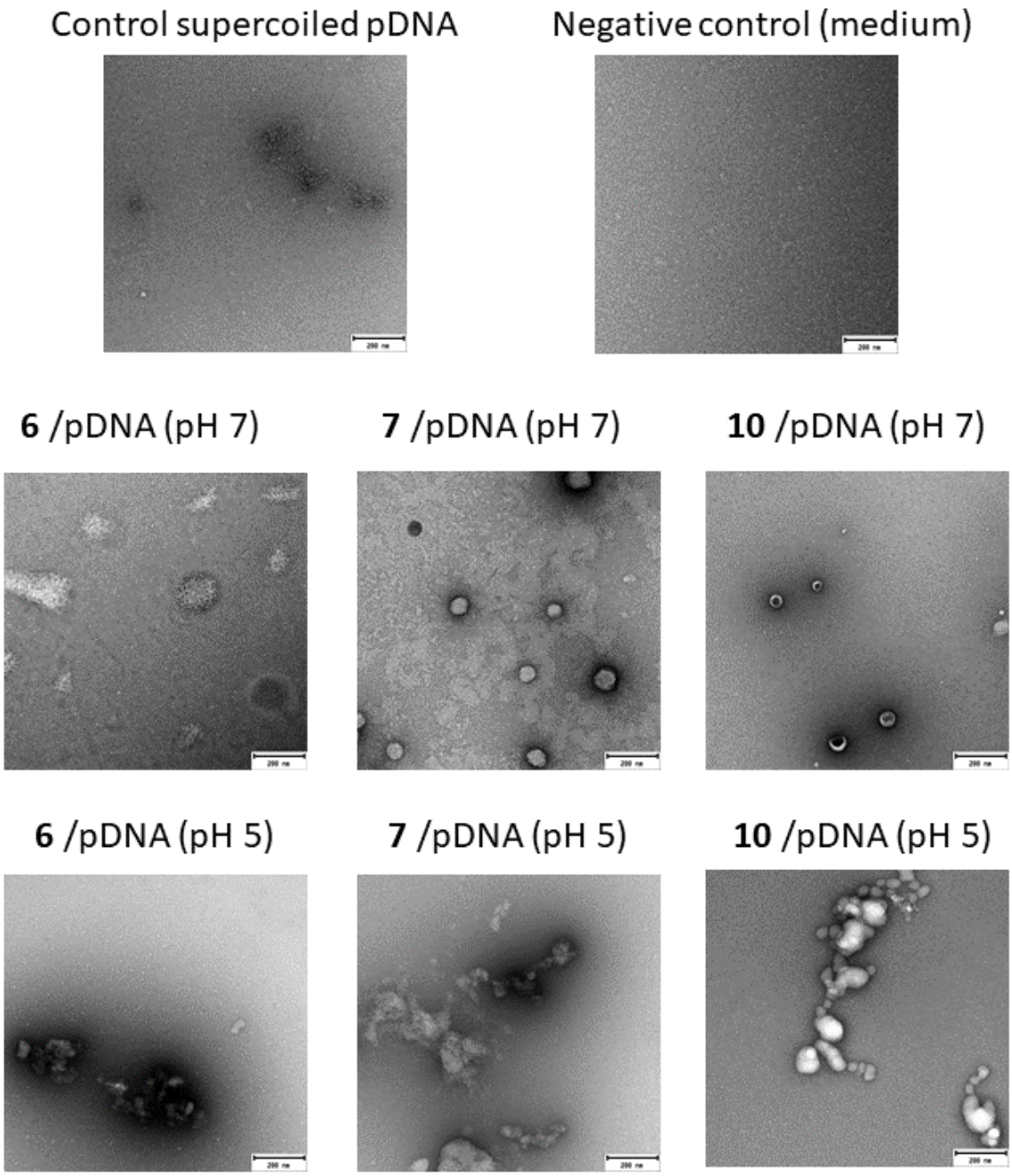


Figure S12. Control TEM experiments (upper panel) and representative TEM micrographs of trehalose vector/pDNA co-assembled nanocomplexes at pH 7 (middle panel) and pH 5 (lower panel).

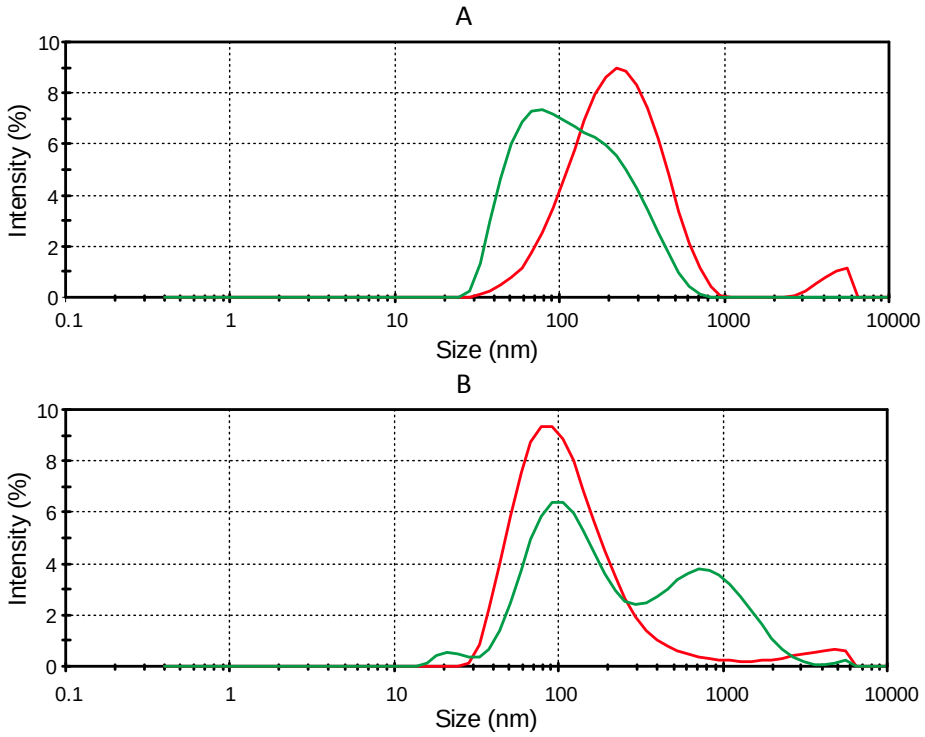


Figure S13. Variations in the hydrodynamic diameters, as determined by DLS, of the nanocomplexes formulated with the Siamese twin vectors **6** (A) and **10** (B) N/P 20 on going from pH 7.4 (red line) to pH 5. The increase in the size of the aggregates and in the degree of polydispersity is consistent with the destabilization of the nanocomplexes upon acidification of the environment.

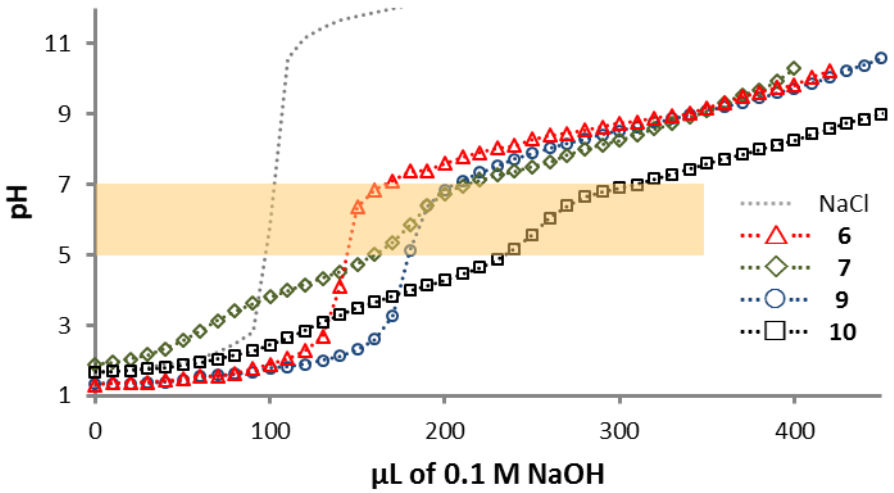


Figure S14. pH Titration curves of compounds **6**-**7** and **9**-**10** (1 mL, 50 mM in protonable amino groups) in the presence of 0.15 M NaCl. Titration plot for NaCl alone (grey dotted line) is included as reference. The titration curve for compound **8** could not be determined due to the limited solubility of the title compound in the experimental conditions. The shadowed region represents the biologically relevant buffering window (from pH 5 to pH 7).

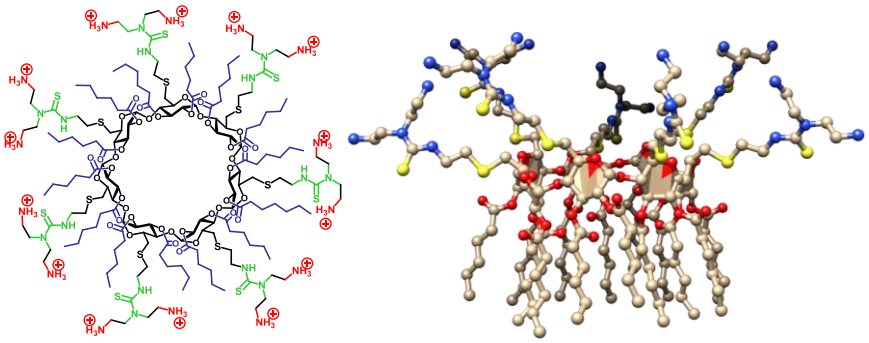


Figure S15. Chemical structure and 3D molecular model of the polycationic amphiphilic cyclodextrin ADM70 used as multivalent molecular vector control in transfection experiments.

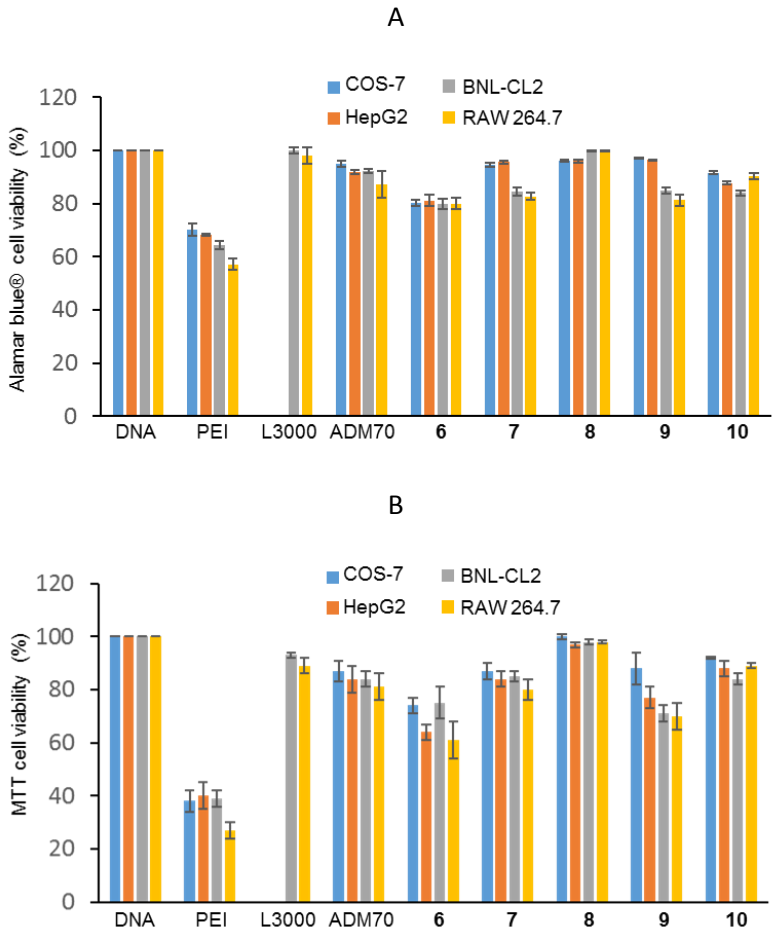


Figure S16. Viability of COS-7, HepG2, BNL-CL2 and RAW 264.7 cells in the presence of nanocomplexes formulated with the Siamese twin vectors **6**-**10** and the luciferase-encoding plasmid pCMV-Luc VR1216 at N/P 20, determined by the Alamar blue® (A) and the MTT protocol (B). Data for polyplexes formulated with polyethyleneimine (PEI), CDplexes formulated with ADM70 and lipoplexes formulated with Lipofectamine 3000® (L3000) are also shown. The data represent the mean ± SD of three wells and are representative of three independent determinations.

Table S1. Hydrodynamic diameter and ζ-potential of the nanoplexes formulated with the trehalose-based gemini facial amphiphiles **6**-**10** and pDNA (pCMV-Luc VR1216) at N/P 5.

Compound	Particle size (nm)	ζ-Potential (mV)	PI
6	Low quality data	--	--
7	126.2 ± 2.8	+6.7 ± 1.5	0.26
8	151.6 ± 22.2	+8.8 ± 1.6	0.29
9	Low quality data	--	--
10	134.0 ± 24.3	+22.4 ± 2.7	0.21

Table S2. Hydrodynamic diameter and ζ-potential of the nanoplexes formulated with the trehalose-based gemini facial amphiphiles **6**-**10** and pDNA (pCMV-Luc VR1216) at N/P 10.

Compound	Particle size (nm)	ζ-Potential (mV)	PI
6	Low quality data	--	--
7	116.3 ± 12.8	+19.3 ± 8.0	0.27
8	127.1 ± 1.3	+27.4 ± 0.8	0.23
9	106.2 ± 5.6	+22.4 ± 1.1	0.24
10	99.5 ± 19.0	+22.7 ± 2.9	0.25

Table S3. Hydrodynamic diameter and ζ-potential of the nanoplexes formulated with the trehalose-based gemini facial amphiphiles **6**-**10** and pDNA (pCMV-Luc VR1216) at N/P 20.

Compound	Particle size (nm)	ζ-Potential (mV)	PI
6	108.9 ± 28.4	+21.1 ± 2.6	0.30
7	109.8 ± 12.5	+32.7 ± 5.5	0.25
8	105.3 ± 14.4	+35.2 ± 2.6	0.22
9	96.0 ± 10.5	+23.7 ± 1.5	0.24
10	102.7 ± 11.9	+37.0 ± 0.9	0.19

[1] J. Rodriguez Lavado, S. E. Sestito, R. Cighetti, E. M. Aguilar Moncayo, A. Oblak, D. Lainscek, J. L. Jimenez Blanco, J. M. García Fernández, C. Ortiz Mellet, R. Jerala, V. Calabrese and F. Peri, *J. Med. Chem.,* 2014, **57**, 9105.

[2] A. Méndez-Ardoy, M. Gómez-García, C. Ortiz Mellet, N. Sevillano, M. D. Girón, R. Salto, F. Santoyo-González and J. M. García Fernández, *Org. Biomol. Chem.*, 2009, **7**, 2681.

[3] D. M. Kneeland, K. Ariga, V. M. Lynch, C. Y. Huang and E. V. Anslyn, *J. Am. Chem. Soc.*, 1993, **115**, 10042.