Supplementary Information

Carboxylate-functionalized foldamer inhibitors of HIV-1 integrase and Toposiomerase 1: artificial analogues of DNA mimic proteins

Valentina Corvaglia[1,2,†], Daniel Carbajo[2,‡,†], Panchami Prabhakaran[2], Krzysztof Ziach[2], Pradeep Kumar Mandal[1,2], Victor Dos Santos[3], Carole Legeay[3], Rachel Vogel[3], Vincent Parissi[4], Philippe Pourquier[5] and Ivan Huc[1,2,*]

[1] Department of Pharmacy and Center for Integrated Protein Science, Ludwig-Maximilians-Universität, München 81377, Germany

[2] Université de Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN (UMR 5248), Institut Européen de Chimie et Biologie, Pessac 33600, France

[3] Sanofi recherche & développement, Montpellier 34184, France

[4] Université de Bordeaux, CNRS, Laboratoire de Microbiologie Fondamentale et Pathogénicité (UMR 5234), Bordeaux 33146, France

[5] INSERM U1194, Institut de Recherche en Cancérologie de Montpellier & Université de Montpellier, Montpellier 34298, France

[‡] Current address: Department of Biological Chemistry Institute of Advanced Chemistry of Catalonia, IQAC-CSIC, Barcelona 08034, Spain

This PDF file includes:

1. Synthetic Schemes S2

2. Methods for chemical synthesis and characterizations S3

2.1 Monomer synthesis S3

2.2 Oligomer synthesis S6

3. Solid state X-ray Crystallography S7

4. Biological assays S13

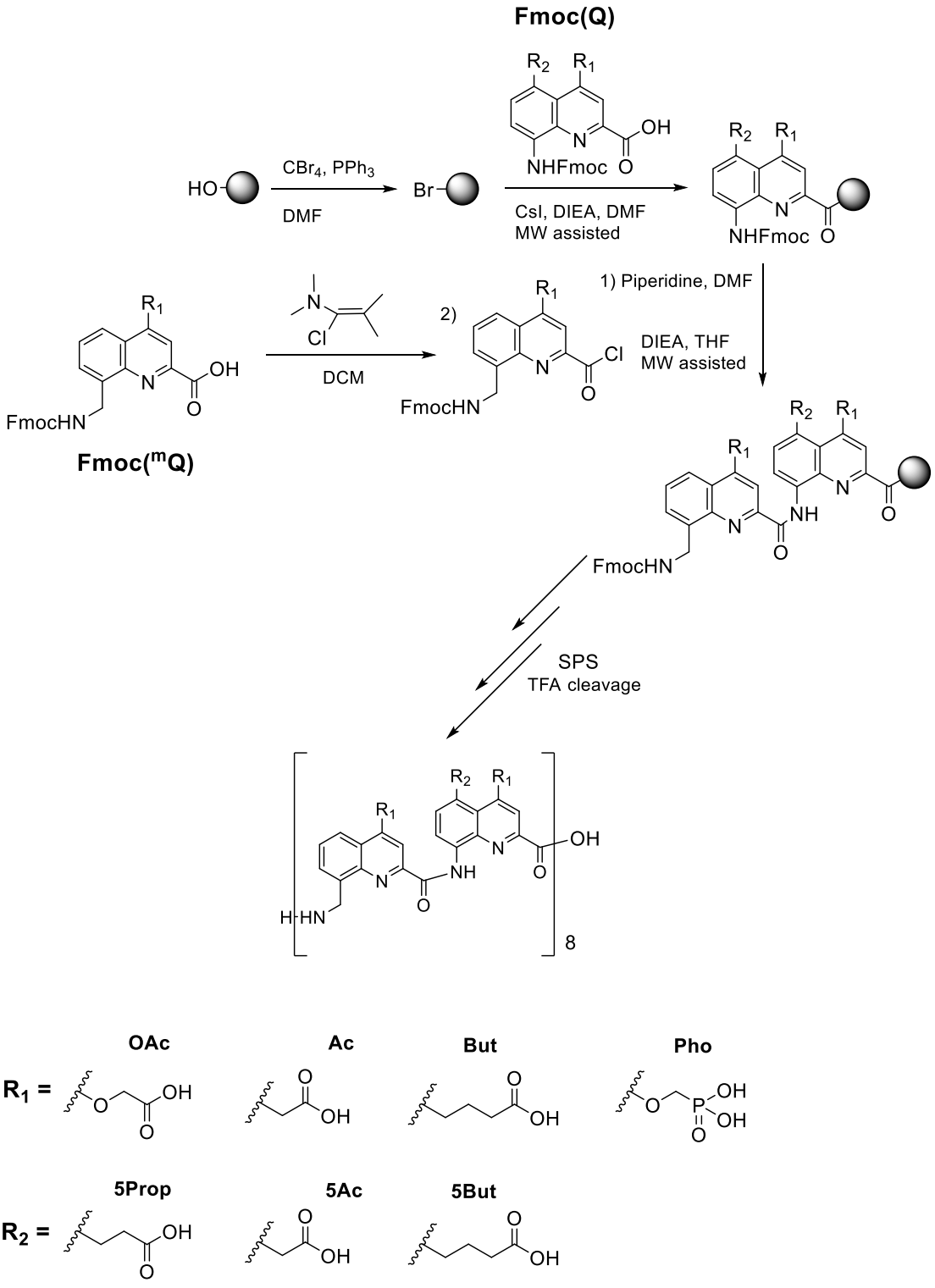
5. ^{1}H NMR, ^{13}C NMR and ^{31}P spectra of purified new synthetic compounds S14

6. HPLC traces of purified new synthetic compounds S31

7. Mass spectrometry analysis of new synthetic compounds S36

8. References S39

1. Synthetic schemes



Supplementary Scheme S1. Solid Phase Synthesis of foldamers **1-6**.

2. Methods for chemical synthesis and characterizations

2.1 Monomer synthesis

Fmoc($\mathbf{Q^{OAc}}$) and **Fmoc**($\mathbf{Q^{5Ac}}$) were synthesized as previously described (1,2). **Fmoc**($\mathbf{^{m}Q^{OAc}}$), **Fmoc**($\mathbf{^{m}Q^{Ac}}$), **Fmoc**($\mathbf{^{m}Q^{But}}$), **Fmoc**($\mathbf{Q^{5Prop}}$), and **Fmoc**($\mathbf{Q^{5But}}$) were synthesized by Sanofi Recherche and Development and their synthetic protocols will be described elsewhere. The synthesis of **Fmoc**($\mathbf{Q^{Pho}}$) and **Fmoc**($\mathbf{^{m}Q^{Pho}}$) is described in the following protocols. Compounds **7** and **10** were synthesized based on the reported procedures (3).

Di-*tert*-butyl hydroxymethylphosphonate was prepared according the reported procedures (4). mp: 100-102 °C; ^{1}H NMR ($CDCl_3$, 300 MHz, δ in ppm) δ: 3.76-3.71 (t, *J* = 6.0 Hz, 2H), 1.51 (s, 18H); ^{13}C NMR ($CDCl_3$, 75 MHz, δ in ppm) δ: 82.6, 82.5, 60.7, 58.5, 30.42, 30.37; ^{31}P NMR ($CDCl_3$, 121 MHz, δ in ppm) δ: 17 ppm.

Benzyl 4-((di-tert-butoxyphosphoryl)methoxy)-8-nitroquinoline-2-carboxylate, 8. To a stirred solution of compound **7** (6.1 g, 18.8 mmol, 1 equiv.) in anhydrous THF under nitrogen, was added di-*tert*-butyl hydroxymethylphosphonate (4.64, 20.7 mmol, 1.1 equiv.), and triphenyl phosphine (5.91 g, 22.5 mmol, 1.2 equiv.) and the reaction was cooled to 0°C using an ice bath. Di-isopropyl azodicarboxylate (4.44 mL, 22.5 mmol, 1.2 equiv.) was added dropwise to the reaction mixture which was stirred for 30-45 min at 0°C, 1 hour at r.t and then 16 hours at 50°C. Solvent was removed under reduced pressure and the crude product was purified by silica gel chromatography (eluent: EtOAc/cyclohexane 7:3, R_f = 0.4) affording a crystalline yellow solid. Yield: 75%; ^{1}H NMR ($CDCl_3$, 300 MHz, δ in ppm) δ: 8.51-8.48 (dd, *J* = 8.49 Hz, 1.14 Hz, 1H), 8.13-8.10 (dd, *J* = 7.50 Hz, 1.38 Hz, 1H), 7.72 (s, 1H), 7.70-7.65 (m, 1H), 7.54-7.51 (m, 2H), 7.43-7.32 (m, 3H), 5.49 (s, 2H), 4.46-4.43 (d, *J* = 9.79 Hz, 2H), 1.55 (s, 18H); ^{13}C NMR ($CDCl_3$, 75 MHz, δ in ppm) δ: 164.7, 162.6, 162.4, 151.3, 148.5, 140.1, 135.4, 128.6, 128.3, 128.2, 126.3, 126.1, 125.1, 123.0, 102.5, 84.2, 84.1, 67.8, 66.5, 64.2, 30.5, 30.4; ^{31}P NMR ($CDCl_3$, 121 MHz, δ in ppm) δ: 6.42 Hz (referenced to PPh_3O: 27.0 ppm); ESI MS for calculated for $C_{26}H_{32}N_2O_8P$: 531.1896; Found: 531.1886 $(M+H)^+$; Calculated for $C_{26}H_{31}N_2NaO_8P$: 553.1716; Found: 553.1703 $(M+Na)^+$. Melting point: 155-156 °C.

8-amino-4-((di-tert-butoxyphosphoryl)methoxy)quinoline-2-carboxylic acid, 9. Compound **8** (0.78 g, 1.4 mmol) was dissolved in ethyl acetate (10 mL) in a 100 mL two neck round bottom flask which was previously evacuated and backfilled with nitrogen (twice). To this, Pd-C (10%, 78 mg) was added as a solid and the reaction mixture was allowed to stir under hydrogen (balloon pressure) at room temperature. The reaction was monitored by TLC and after complete consumption of starting material (approx. 4 hours), the reaction mixture was filtered over a celite pad, washed with ethyl acetate several times, then with THF until no more product was detected in the filtrate (TLC). The greenish solid obtained after removal of solvent under reduced pressure was used for the next step without further purification. Yield: quantitative. ^{1}H NMR ($CDCl_3$, 300 MHz, δ in ppm) δ: 7.64-7.60 (m, 2H), 7.46-7.41 (t, *J* = 7.62 Hz, 1H), 7.03-6.99 (dd, *J* = 7.58 Hz, 1.19 Hz, 1H), 4.44-4.40 (d, *J* = 9.97 Hz, 2H), 1.57 (s, 18

H); ^{13}C NMR ($CDCl_3$, 75 MHz, δ in ppm) δ: 164.7, 163.5, 163.3, 144.3, 143.9, 136.6, 129.3, 123.1, 112.0, 110.2, 99.1, 84.2, 84.1, 66.2, 63.9, 30.6, 30.5; ESI MS calculated for $C_{19}H_{26}N_2O_6P$: 409.1528; Found: 409.1534 $(M-H)^-$. Melting Point: 121-122 °C.

Fmoc(Q^{Pho}). To a stirred solution of compound **9** (4 g, 9.8 mmol, 1 equiv.) in 120 mL dioxane in a 500 mL round bottom flask, was added 10% wt/vol $NaHCO_3$ solution (17.2 g, 21 equiv.) and the resulting mixture was cooled to 0 °C using an ice bath. To this, was added a solution of Fmoc-Cl (3.02 g, 11.7 mmol, 1.2 equiv.) in dioxane (70 mL) drop wise over 1 hour maintaining the temperature of the reaction mixture to 0°C. It was further stirred for 1 hour at 0°C and then overnight at rt. Water was then added and the mixture acidified with saturated $KHSO_4$ solution to pH 3-4 and the product was extracted with DCM. The aqueous layer was extracted twice with DCM and the combined organic layers were washed twice with water, then with brine and finally dried over an. Na_2SO_4. The crude product obtained after removal of solvent under reduced pressure was purified by silica gel chromatography (eluent: 5% MeOH in DCM with the addition of 0.1% AcOH, R_f = 0.3) yielded a light yellow solid. After column chromatography, AcOH was removed by extraction with water and lyophilization. Yield: 50%; ^{1}H NMR (DMSO-d_6, 300 MHz, δ in ppm) δ: 13.52 (s, 1H), 10.44 (s, 1H), 8.34(bs, 1H), 7.94-7.92 (m, 2H), 7.83-7.77 (m, 4H), 7.68-7.62 (t, *J* = 6.24 Hz, 1H), 7.46-7.41 (m, 2H), 7.39-7.34 (td, *J* = 7.39 Hz, 1.25Hz, 2H), 4.67-4.60 (m, 4H), 4.47-4.43 (t, *J* = 6.72 Hz, 1H), 1.47(s, 18 H); ^{13}C NMR (DMSO-d_6, 75 MHz, δ in ppm) δ: 165.2, 162.4, 162.2, 153.4, 146.5, 143.6, 140.7, 137.4, 135.6, 128.5, 127.6, 125.0, 121,7, 120.1, 116.4, 114.1, 101.4, 82.7, 82.6, 66.3, 66.1, 63.8, 46.5, 29.9, 29.86, 26.2; ESI MS calculated for $C_{34}H_{36}N_2O_8P$: 631.2209; Found: 631.2217 $(M-H)^-$; calculated for $C_{34}H_{37}N_2O_8P$: 632.2288; Found: 632.2249 $(M)^+$. HPLC purity: 99.6% at 300 nm. Melting point: 73-75 ^{o}C.

Benzyl 8-cyano-4-((di-tert-butoxyphosphoryl)methoxy)quinoline-2-carboxylate, 11. A solution containing compound **10** (4.9 g 16.1 mmol, 1 equiv.), triphenyl phosphine, (5.48 g, 20.9 mmol, 1.3 equiv.) and di-*tert*-butyl hydroxymethylphosphonate (3.97 g, 17.7 mmol, 1.1 equiv.) in dry THF (45 mL) was stirred at 0°C under nitrogen. To this, di-isopropyl azodicarboxylate (4.12 mL, 20.9 mmol, 1.3 equiv.) was added drop wise and the resulting orange colored solution was stirred for 30-45 min at 0°C, 1 hour at r.t and then 16 hours at 50°C. Solvent was removed under reduced pressure and the crude product was purified by silica gel chromatography (eluent: ethylacetate, R_f = 0.5). Yield: 75 %; ^{1}H NMR ($CDCl_3$, 300 MHz, δ in ppm) δ: 8.52-8.49 (dd, *J* = 8.45 Hz, 1.41 Hz, 1H), 8.19-8.16 (dd, *J* = 7.20 Hz, 1.41 Hz, 1H), 7.71 (s, 1H), 7.68-7.64 (m, 1H), 7.58-7.56 (m, 2H), 7.44-7.31 (m, 3H), 5.52 (s, 2H), 4.45-4.42 (d, *J* = 9.78 Hz, 2H), 1.54 (s, 18H); ^{13}C NMR ($CDCl_3$, 75 MHz, δ in ppm)δ: 164.9, 162.9, 162.7, 151.2, 147.8, 136.7, 135.4, 128.6, 128.3, 128.2, 126.9, 126.8,122.3, 116.6, 114.0, 102.6, 84.2, 84.0, 67.9, 66.5, 64.1, 30.5, 30.4; ^{31}P NMR ($CDCl_3$, 121 MHz, δ in ppm) δ: 6.08 ppm (Ref: PPh_3O at 27 ppm); ESI MS calculated for $C_{27}H_{32}N_2O_6P$: 511.1998; Found: 511. 1995 $(M+H)^+$. Melting point: 89-90°C.

8-(aminomethyl)-4-((di-tert-butoxyphosphoryl)methoxy)quinoline-2-carboxylic acid, 12. Compound **11** (1 g, 1.96 mmol) was dissolved in THF (10 mL). The flask was evacuated and backfilled with nitrogen. To this solution, Pd-C (10%, 100 mg) was added as a solid and the reaction mixture was

allowed to stir under hydrogen balloon pressure at rt. After 5 hours, hydrogenolysis was complete according to TLC monitoring. The reaction mixture was evacuated and backfilled with nitrogen and raney nickel (approx. 1g) was added into the reaction mixture and again the resulting mixture was vigorously stirred at rt. under hydrogen balloon pressure for 16 hours. It was then filtered over a celite pad and washed with THF several times. The crude compound isolated after the removal of solvent under reduced pressure was used for the next step without further purification. ESI MS calculated for $C_{20}H_{30}N_2O_6P$: 425.1836; Found: 425.1837 $(M+H)^+$.

Fmoc($^{m}Q^{Pho}$). To a stirred solution of compound **12** (4 g, 9.8 mmol, 1 equiv.) in 120 mL dioxane in a 500 mL round bottom flask, was added 10% wt/vol $NaHCO_3$ solution (17.2 g, 21 equiv.) and the resulting mixture was cooled to 0°C using an ice bath. To this, was added a solution of FmocCl (3.02 g, 11.7 mmol, 1.2 equiv.) in dioxane (70 mL) drop wise over 1h maintaining the temperature of the reaction mixture to 0°C. It was further stirred for 1 hour at 0°C and then overnight at r.t. Water was then added and the mixture acidified with saturated $KHSO_4$ solution to pH 3-4 and the product was extracted with DCM. The aqueous layer was extracted twice with DCM and the combined organic layers were washed twice with water, then with brine and finally dried over an. Na_2SO_4. The crude product obtained after removal of solvent under reduced pressure was purified by silica gel chromatography (eluent: 5% MeOH in DCM with the addition of 0.1% AcOH) yielded a white solid. After column chromatography, AcOH was removed by extraction with water and lyophilisation. Yield: 50%; ^{1}H NMR (DMSO-d_6, 300 MHz, δ in ppm) δ: 8.07-8.05 (m, 1H), 7.90-7.86 (m, 2H), 7.72-7.69 (m, 3H), 7.62-7.53 (m, 2H), 7.44-7.39 (t, *J* = 7.33 Hz, 2H), 7.34-7.29 (t, *J* = 7.2 Hz, 2 H), 4.87-4.85 (d, *J* = 5.76 Hz, 2H), 4.54-4.51 (d, *J* = 9.03 Hz, 2H), 4.39-4.37 (d, *J* = 6.6 Hz, 2H), 4.27-4.22 (t, *J* = 6.6 Hz, 1H), 1.47 (s, 18H); ^{13}C NMR (DMSO-d_6, 75 MHz, δ in ppm) δ: 165.9, 162.1, 162.0, 156.4, 148.5, 145.2, 143.7, 140.7, 137.8, 128.2, 127.5, 126.9, 124.98, 121.2, 119.98, 101.3, 82.7, 82.6, 65.9, 65.2, 63.6, 46.7, 29.9, 29.87, 26.2; ESI MS for calculated for $C_{35}H_{39}N_2O_8P$: 646.2444; Found: 646.2446 $(M)^+$; calculated for $C_{35}H_{38}N_2O_8P$: 645.2360; Found: 645.2377 $(M-H)^-$. HPLC purity: 99.0% at 300 nm. Melting point: 93-94°C.

Fmoc(Q^{OAc}): ^{1}H NMR (300MHz, DMSO-d_6): δ 10.46 (1H, s), 8.38 (1H, brot), 7.93 (2H, d, J=7.3Hz), 7.83 (3H, m), 7.65 (1H, tr, J=8.1Hz), 7.62 (1H, s), 7.45 (2H, tr, 7.3Hz), 7.37 (2H, tr, J=7.3Hz), 5.15 (2H, s), 4.63 (2H, d, J=6.7Hz), 4.47 (1H, tr, 6.7Hz), 1.46 (9H, s). ^{13}C NMR (300MHz, DMSO): 167.3, 165.8, 162.3, 154.0, 147.0, 144.2, 141.3, 138.1, 136.2, 129.2, 128.3, 127.7, 125.6, 122.2, 120.7, 117.1, 115.1, 101.5, 82.6, 66.9, 66.3, 47.1, 28.1. HRMS (ESI) calculated for $C_{31}H_{29}N_2O_7$ $(M+H)^+$: 541.1969. Found: 541.1963. HPLC purity: 99% at 254 nm. Melting point: 126-128°C.

Fmoc(Q^{5Prop}): ^{1}H NMR (300MHz, DMSO-d_6): δ 10.4 (1H, s), 8.73 (1H, d, J=8.8Hz), 8.24 (2H, d, J=8.8Hz), 7.93 (2H, d, J=7.4Hz), 7.79 (2H, d, J=7.4Hz), 7.52 (1H, d, J=8Hz), 7.46 (2H, tr, J=7.4Hz), 7.37 (2H, tr, 7.4Hz), 4.62 (2H, d, 6.7Hz), 4.46 (1H, tr, J=6.7Hz), 3.28 (2H, tr, J=7.3Hz), 2.61 (2H, tr, J=7.3Hz), 1.36 (9H, s). ^{13}C NMR (300MHz, DMSO): 171.8, 165.8, 154.0, 145.5, 144.2, 141.3, 137.5, 135.5, 134.8, 130.8, 129.4, 128.3, 128.14, 127.7, 125.6, 121.0, 120.7, 116.4, 80.3, 66.8, 47.1, 36.1, 28.2, 26.8. HRMS (ESI) calculated for $C_{32}H_{31}N_2O_6$ $(M+H)^+$: 539.2177. Found: 539.2180. HPLC purity:

98.7% at 254 nm. Melting point: 192-195°C.

Fmoc($^{m}Q^{OAc}$): ^{1}H NMR (300MHz, DMSO-d_6): δ 8.17 (1H, d, 7.9Hz), 7.9 (3H, d, 7.5Hz), 7.65 (4H, m), 7.52 (1H, s), 7.43 (2H, tr, J=7,3Hz), 7.32 (2H, tr, J=7.3Hz), 5,14 (2H, s), 4.86 (2H, d, 6Hz), 4.42 (2H, d, 6.7Hz), 4.26 (1H, tr, 6.5Hz), 1,46 (9H, s). ^{13}C NMR (300MHz, DMSO): 167.4, 166.4, 162.1, 157.0, 148.7, 145.9, 144.4, 141.3, 138.4, 128.9, 128.1, 127.5, 125.6, 121.8, 120.9, 120.6, 101.4, 82.6, 66.2, 65.8, 47.3, 28.2. HRMS (ESI) calculated for $C_{32}H_{31}N_2O_7$ $(M+H)^+$: 555.2126 Found: 555.2128. HPLC purity: 98.7% at 254 nm. Melting point: 182-186°C.

Fmoc(Q^{5But}): ^{1}H NMR (300MHz, DMSO-d_6): δ 10.4 (1H, s), 8.74 (1H, d, J=8.8Hz), 8.23 (2H, d, 8.8Hz), 7.93 (2H, d, J=7.3Hz), 7.78 (2H, d, J=7.3Hz), 7.4 (5H, m), 4.62 (2H, d, J=6.6Hz), 4.45 (1H, tr, J=6.6Hz), 3.02 (2H, tr, J=7.5Hz), 2.30 (2H, tr, J=7.5Hz), 1.85 (2H, m), 1.4 (9H, s). ^{13}C NMR (300MHz, DMSO): 172.5, 165.9, 154.0, 145.6, 144.2, 141.3, 137.6, 135.5, 134.6, 131.9, 129.6, 128.3, 128.2, 125.6, 120.9, 120.7, 116.5, 80.1, 66.8, 47.1, 34.7, 30.8, 28.3, 26.5. HRMS (ESI) calculated for $C_{33}H_{33}N_2O_6$ $(M+H)^+$: 553.2333. Found: 553.2355. Purity by HPLC: 99.7% at 254nm. Melting point: 111-112°C.

Fmoc($^{m}Q^{But}$): ^{1}H NMR (300MHz, DMSO-d_6): δ 8.18 (1H, d, J=8.2Hz), 8.03 (1H, s), 7.90 (2H, d, J=7.4Hz), 7.70 (4H, m), 7.43 (2H, tr, J=7.4Hz), 7.32 (2H, tr, J=7Hz), 4.9, (2H, d, J=5.4Hz), 4.42 (2H, 6.7Hz), 4.26 (1H, tr, J=6.7Hz), 3.19 (2H, 7.9Hz), 2.34 (2H, tr, J=7.1Hz), 1.93 (2H, m), 1.4 (9H, s) ^{13}C NMR (300MHz, DMSO): 172.4, 166.6, 157.0, 150.5, 147.2, 145.1, 144.4, 141.3, 139.2, 128.7, 128.4, 128.1, 127.5, 125.6, 123.3, 120.7, 120.6, 80.2, 65.8, 47.3, 34.6, 31.4, 28.2, 25.9. HRMS (ESI) calculated for $C_{34}H_{35}N_2O_6$ $(M+H)^+$: 567.2490 Found: 567.2491. HPLC purity: 98.9% at 254 nm. Melting point: 100-102°C.

Fmoc(Q^{5Ac}): ^{1}H NMR (300MHz, DMSO-d_6): δ 13.63 (1H, s, COOH), 10.47 (1H, s, NH), 8.62 (1H, d J=8.5Hz), 8.27 (2H, d, J=8.8Hz), 7.94 (2H, d, J=7.4Hz), 7.80 (2H, d, J=7.4Hz), 7.58 (1H, d, J=8.1Hz), 7.44 (2H, tr, J=7.4Hz), 7.39 (2H, tr, J=7.4Hz), 4.64 (2H, d, 6.7Hz), 4.47(1H, tr, 6.7Hz), 4.06 (2H, s), 1.38 (9H, s). ^{13}C NMR (300MHz, DMSO): 170.7, 165.8, 154.0, 145.6, 144.2, 141.3, 137.5, 136.1, 135.5, 131.3, 128.8, 128.3, 127.7, 125.8, 125.6, 121.0, 120.7, 116.3, 81.1, 66.9, 47.1, 38.7, 28.1. HRMS (ESI) calculated for $C_{31}H_{29}N_2O_6$ $(M+H)^+$: 525.2020. Found: 525.2037. HPLC purity: 99.7% at 254 nm. Melting point: 108-109°C.

Fmoc($^{m}Q^{Ac}$): ^{1}H NMR (300MHz, DMSO-d_6): δ 13.18 (1H, s, COOH), 8.12 (1H, s), 8.02 (1H, d, J=8.2Hz), 7.92 (2H, d, J=7.2Hz), 7.7 (4H, m), 7.45 (2H, tr, J=7.2Hz), 7.32 (2H, tr, J=7.2Hz), 4.91 (2H, d, J=7Hz), 4.42 (2H, d, J=6Hz), 4.27 (3H, m), 1.40 (9H, s). ^{13}C NMR (300MHz, DMSO): 169.8, 166.5, 157.1, 145.0, 144.4, 144.0, 141.3, 139.2, 128.9, 128.8, 128.1, 127.5, 125.6, 123.5, 122.3, 120.6, 81.5, 65.8, 47.4, 41.0, 28.1. HRMS (ESI) calculated for $C_{32}H_{31}N_2O_6$ $(M+H)^+$: 539.2177. Found: 539.2180. Purity by HPLC: 98.3% at 254nm. Melting point: 143-145ºC.

2.2 Oligomer synthesis

Solid phase synthesis of foldamers **1-6** (Supplementary Scheme S1) was performed by using low loading Wang resin (0.38 mmol/g) following previously described procedures (1,2,5). It was converted

to a bromomethyl Wang resin. Each first monomer was anchored to the bromomethyl Wang resin in the presence of cesium iodide. The introduction of the subsequent monomers was done *via* acid chloride activation. Coupling reactions were performed at 50°C, 50 W, for 5-15min. Fmoc deprotection was performed at room temperature with a 20% vol/vol solution of piperidine in DMF. Resin cleavage was performed with a standard TFA cocktail (TFA/triisopropylsilane/H_2O 95:2.5:2.5) freshly prepared. The detailed characterization is described in the supplementary information.

Oligomer 2 ($^{m}Q^{OAc}Q^{OAc}$)$_8$**.** Crude compound was obtained in 72% purity. The compound was purified by semi-prep HPLC; Gradient: 5-16% of solvent B in A in 45 minutes. Final oligomer was 96% pure by HPLC and MS confirmed its composition: ESI MS calculated for $C_{200}H_{145}N_{32}O_{65}$ $[M-H]^{-}$: 4033.9103; Found: 1344.5832 $[M-3H]^{3-}$, 1008.1801 $[M-4H]^{4-}$.

Oligomer 3 ($^{m}Q^{OAc}Q^{5Prop}$)$_8$**:** Crude compound was obtained in 83% purity. The compound was purified by semi-prep HPLC; Gradient: 5-16% of solvent B in A in 45 minutes. Final oligomer was 97% pure at 254 nm and MS confirmed its composition: ESI MS calculated for $C_{208}H_{161}N_{32}O_{57}$ $[M-H]^{-}$: 4018.0762; Found: 2009.4777 $[M-2H]^{2-}$, 1339.3089 $[M-3H]^{3-}$.

Oligomer 4 ($^{m}Q^{Ac}Q^{5Ac}$)$_8$: Crude compound was obtained in 73% purity. The compound was purified by semi-prep HPLC; gradient 7-14% of solvent B in A in 32 minutes. Final oligomer was 97.3% pure by HPLC and MS confirmed its composition: ESI MS calculated for $C_{200}H_{145}N_{32}O_{49}$ $[M-H]^{-}$: 3777.9916; Found: 1259.2795 $[M-3H]^{3-}$, 944.2014 $[M-4H]^{4-}$.

Oligomer 5 ($^{m}Q^{But}Q^{5But}$)$_8$: Crude compound was obtained in 80% purity. The compound was purified by semi-prep HPLC; gradient 7-18% of solvent B in A in 32 minutes. Final oligomer was 98% pure by HPLC and MS confirmed its composition: ESI MS calculated for $C_{232}H_{209}N_{32}O_{49}$ $[M-H]^{-}$: 4226.4924; Found: 2113.6838 $[M-2H]^{2-}$, 1408.7785 $[M-3H]^{3-}$, 1056.3253 $[M-4H]^{4-}$.

Oligomer 6 ($^{m}Q^{OAc}Q^{Pho}$)$_8$: Crude compound was obtained in 78% purity. The compound was purified by semi-prep HPLC; gradient 7-12% of solvent B in A in 32 minutes. Final oligomer was 99.1% pure by HPLC and MS confirmed its composition: ESI MS calculated for $C_{192}H_{153}N_{32}O_{73}P_8$ $[M-H]^{-}$: 4321.7223; Found: 1440.5176 $[M-3H]^{3-}$, 1080.1314 $[M-4H]^{4-}$.

3. Solid state X-ray Crystallography

Crystals of **4** was observed to have large volume fractions of disordered solvent molecules and side chains, weak diffraction intensity, incompleteness of data and moderate resolution. A number of A- and B - level alerts were detected and are explicitly listed below and have been divided into two groups. They are inherent to the data and refinement procedure. They illustrate the limited practicality of the checkcif tool for medium sized molecule crystallography. The first group illustrates the weak quality of the data and refinement statistics when compared to that expected for small molecule structures from highly diffracting crystals. The second group is connected with decisions made during refinement and explained below.

Group 1 alerts

THETM01_ALERT_3_A: The value of sine(theta_max)/wavelength is less than 0.550

Calculated sin(theta_max)/wavelength = 0.5263

PLAT029_ALERT_3_B _diffrn_measured_fraction_theta_full value Low 0.949 Note

PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) ..... 0.39 Report

High R1 Value .....0.13 Report

Group 2 alerts

PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors of O1C Check

This alert concerns a side chain atom; however, this does not indicate an incorrect atom-type assignment.

PLAT242_ALERT_2_B Low 'MainMol' Ueq as Compared to Neighbors of Ca5D Check

This alert concerns a Ca^{2+} ion, coordinated to a side chain and solvent molecules.

PLAT430_ALERT_2_B Short Inter D...A Contact O15A .. O16D .. 2.69 Ang.

PLAT430_ALERT_2_B Short Inter D...A Contact O15D …O21C … 2.76 Ang.

PLAT430_ALERT_2_B Short Inter D...A Contact O16A …O17D … 2.58 Ang.

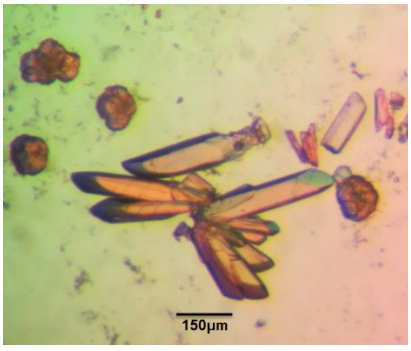
PLAT430_ALERT_2_B Short Inter D...A Contact O16D …O26C … 2.71 Ang.

PLAT430_ALERT_2_B Short Inter D...A Contact O17A … O17D … 2.73 Ang.

PLAT430_ALERT_2_B Short Inter D...A Contact O17A … O18D … 2.81 Ang.

These alerts concern solvent molecules coordinated to the Ca^{2+} ions and **4**.

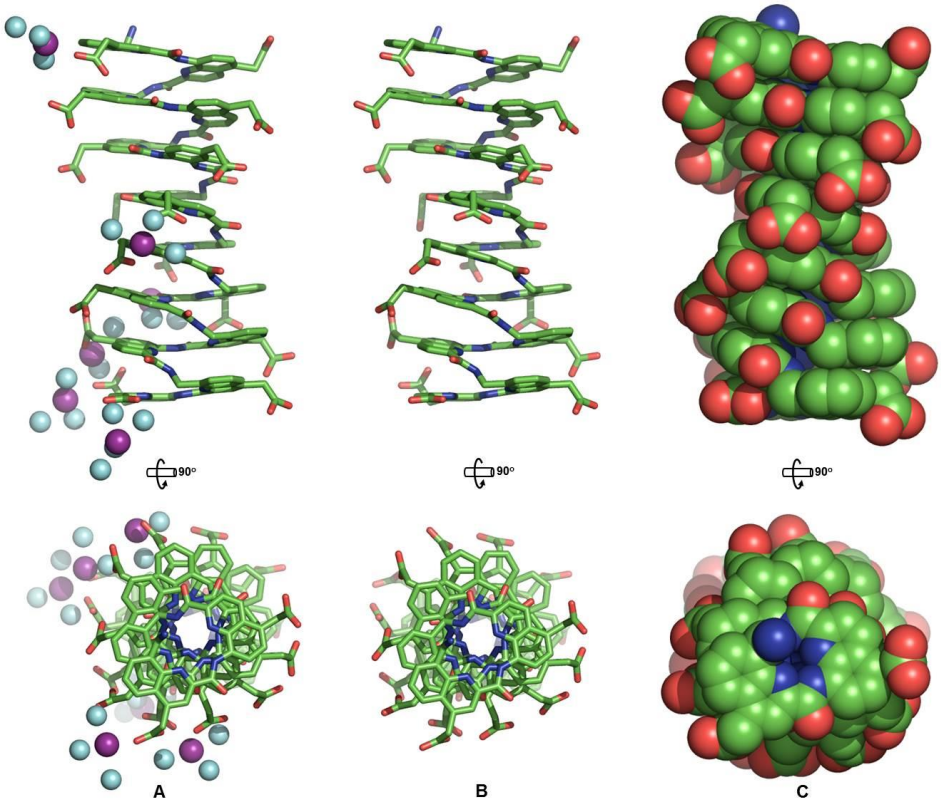
Atomic coordinates and structure factors for the crystal structure has been deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession code 1527431. The data is available free of charge upon request (www.ccdc.cam.ac.uk/). All figures of the model were prepared in *PyMOL*.(6)



Supplementary Figure S1. Aqueous drop containing crystals of the DNA mimic **4**. The crystals appeared after 5 days and were left to grow for 10 days before data collection.

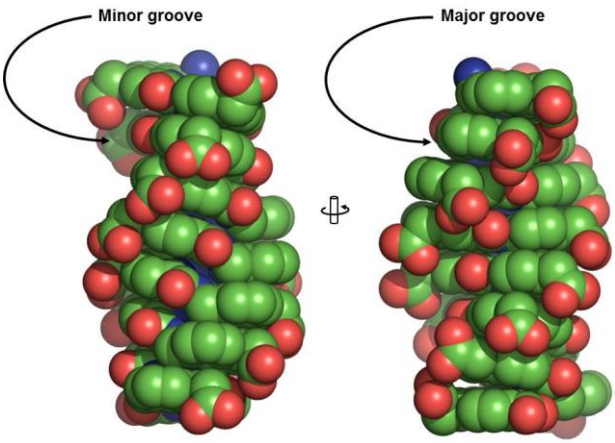
Crystal structure of the DNA mimic 4 from aqueous solution

The crystallographic asymmetric unit consisted of a single helix of **4** along with six Ca^{2+} ions (used in crystallization drop as additive to counter the poly-anionic species, Supplementary Figure S2). Previous reports of aromatic oligoamides derived from 8-amino-2-quinoline carboxylic acid revealed single helical conformations in crystal structure(7) and solution.(8) The single helical folding is driven by electrostatic repulsions, hydrogen bonds (between amide and endocyclic nitrogen atoms) and stacking between aromatic units.

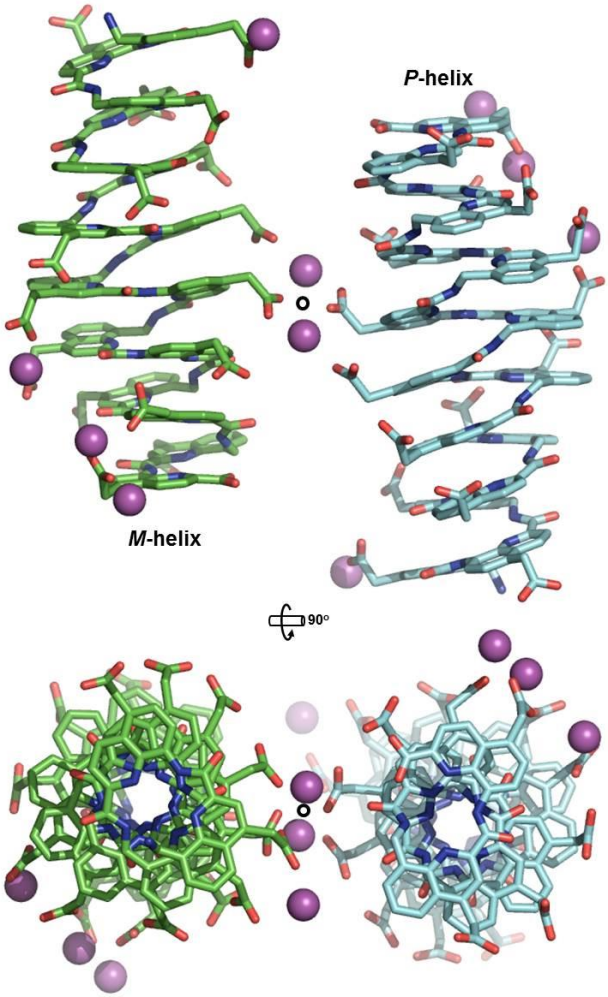


Supplementary Figure S2. Crystal structure of the water soluble DNA mimic **4**: top and bottom view are across and down the helical axis, respectively. (A) The asymmetric unit consists of a single stranded helix (stick representation), six Ca^{2+} ions (magenta spheres) and ordered water molecules (cyan spheres); (B, C) stick and sphere representations respectively (ions and water omitted for clarity).

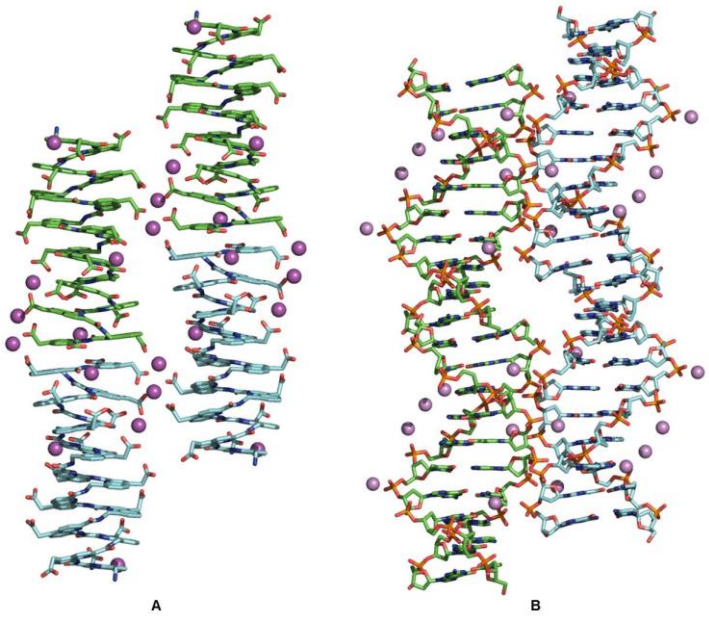
The design of DNA mimic **4** ensures major and minor grooves (Figures S3-S4). The presence of Ca^{2+} ions and associated water molecules in the minor groove causes narrowing of the groove towards the C-terminus Ca^{2+} ions and associated water molecules were observed to have direct and water-mediated interactions to the side-chains, backbone and C-terminus of the mimic. The presence of Ca^{2+} is often observed in the structures of nucleic acids(9-11) and other poly-anionic species.(12) In this structure, the Ca^{2+} ions were identified in the minor groove and to the side chains that mimic the phosphate backbone in B-DNA. The Ca^{2+} ions were also observed in the inter-helical bridging positions (Supplementary Figure S4).



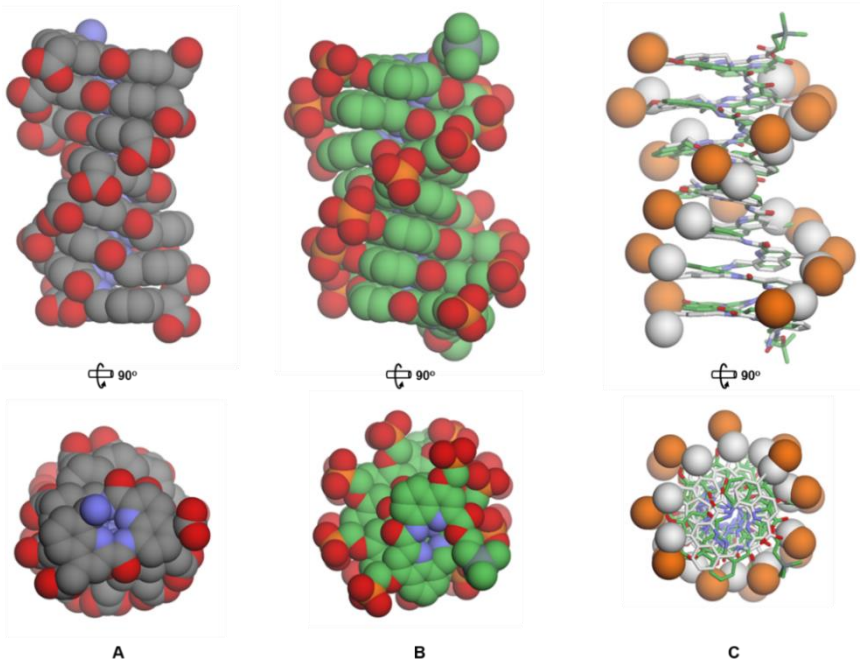
Supplementary Figure S3. Comparison of minor groove and major groove in the single helical conformation of **4**.



Supplementary Figure S4. Racemic crystal structure of **4** illustrating the relation between the left-handed (*M*) and right- handed (*P*) helices (coloured green and cyan, respectively). O represents the centre of inversion, Ca^{2+} ions are shown as magenta spheres and water molecules were omitted for clarity.



Supplementary Figure S5. Comparison between the racemic crystal packing of (A) single stranded DNA mimic **4** and (B) double stranded 10bp B-DNA duplex d(CCGGTACCGG)$_2$ (PDB ID 4R49)(13) illustrating a pseudo-continuous packing arrangement of left-handed helices (coloured green) and right-handed helices (coloured cyan). In both the racemic structures, Ca^{2+} ions (shown as magenta spheres) were identified. Pseudo-continuity of the left- and right-handed helices are alternating in **4**, whereas in the B-DNA helical columns were enantiopure.



Supplementary Figure S6. Comparison between the crystal structures of (A) DNA mimics **4** and (B) (**^{m}Q^{Pho}Q^{Pho}**)$_8$ with phosphonate side chains protected as TMSE esters in position 4. (C) Overlay of the two crystal structures in stick representation except that the phosphorus atoms (orange) and the carbon atoms (white) of the side chains are shown as spheres.

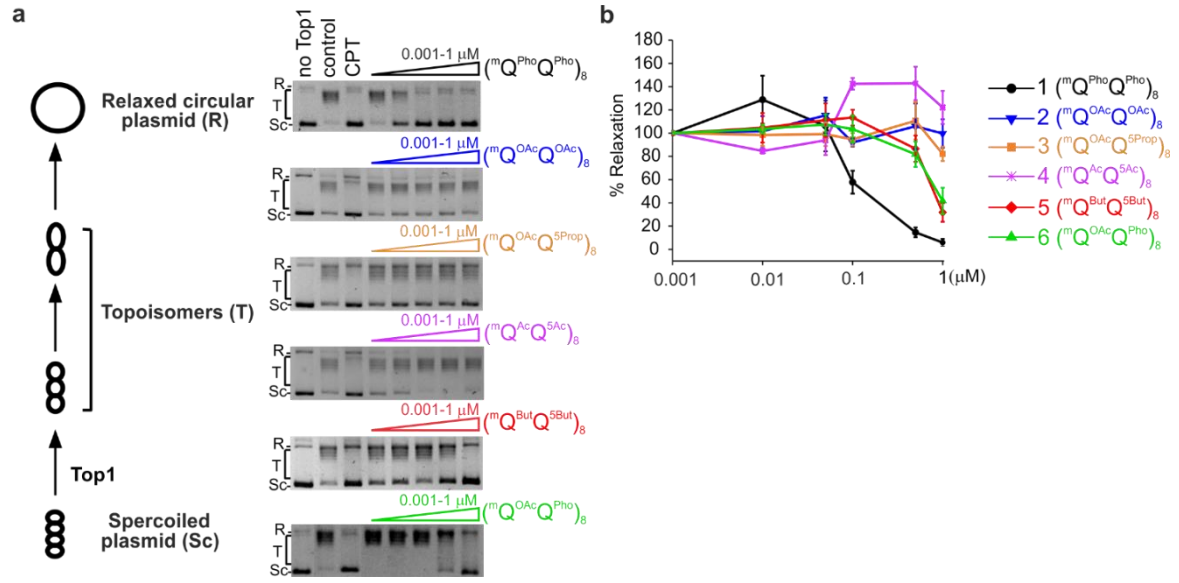
Supplementary Table S1: Crystallographic data of **4**.

Solvent	H_2O	T/K	100
Formula	$C_{200}H_{129}Ca_{4.5}N_{32}O_{66}$	ρ/g cm^{-3}	1.049
M	4216.70	Shape and colour	Rectangular rods, Yellow
Crystal system	Triclinic	size (mm)	0.3 x 0.05 x 0.05
Z	2	λ/Å	1.54178
Space group	*P-1*	μ/mm^{-1}	1.416
a/Å	20.852(4)	Absorption correction	Multi-scan
b/Å	21.727(4)	Collected reflections	30907
c/Å	31.330(6)	unique data [*Fo*>2□□*Fo*)]	15607
α/°	83.61 (3)	R_{int} %	0.0662
β/°	75.07(3)	parameters/restraints	2353/240
γ/°	77.15(3)	R_1, wR_2 ($I > 2\sigma(I)$)	0.1328, 0.3866
U/Å^3	13349(5)	goodness of fit	1.236
CCDC Number	1527431		

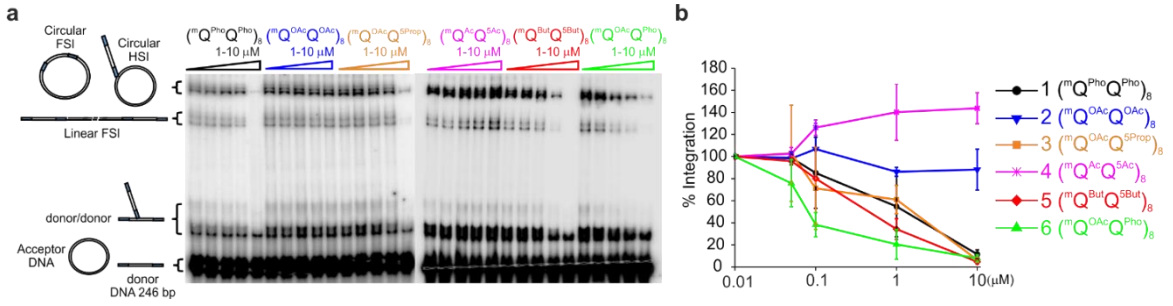
Supplementary Table S2: Crystallographic data of **Fmoc**(^{m}**Q**^{Pho}).

Solvent	AcOEt / *n*-hexane	T/K	143
Formula	$C_{35}H_{39}N_2O_8P$	ρ/g cm^{-3}	1.294
M	646.65	Shape and colour	block, colorless
Crystal system	Triclinic	size (mm)	0.12×0.10×0.08
Z	2	λ/Å	1.54187
Space group	*P-1*	μ/mm^{-1}	1.185
a/Å	10.2817(7)	Absorption correction	Multi-scan
b/Å	12.1899(8)	Collected reflections	4845
c/Å	13.6767(9)	unique data [*Fo*>2□□*Fo*)]	5881
α/°	79.719	R_{int} %	0.0221
β/°	79.875	parameters/restraints	474/24
γ/°	89.901	R_1, wR_2 ($I > 2\sigma(I)$)	0.0430, 0.1257
U/Å^3	1659.57(19)	goodness of fit	1.019
CCDC Number	1474392		

4. Biological Assays

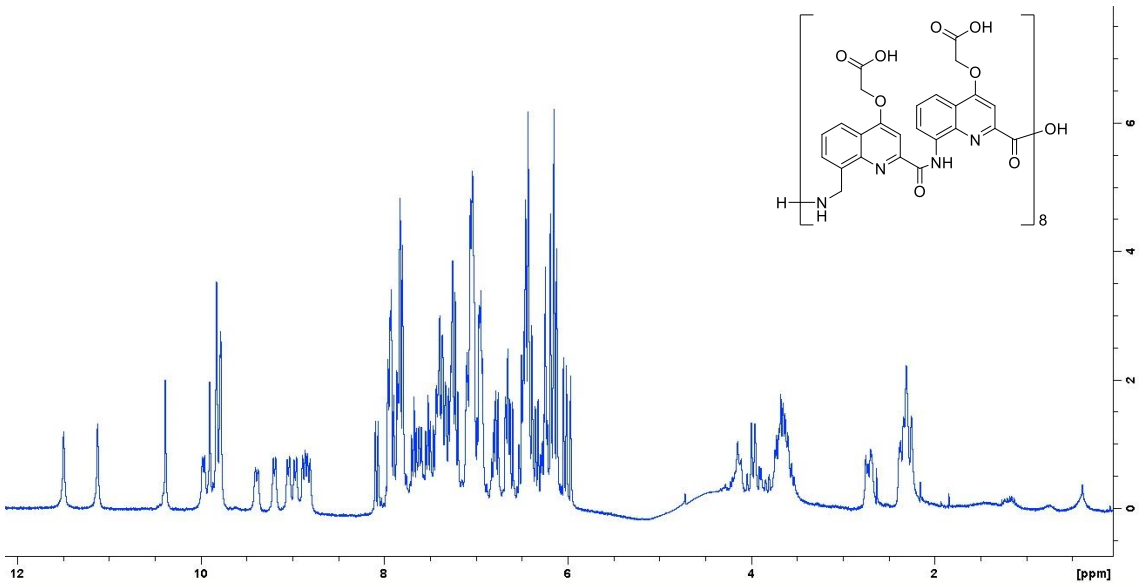


Supplementary Figure S7. **Effect of $(^{m}Q^{Pho}Q^{Pho})_8$ and its derivatives on Top1-mediated relaxation of supercoiled DNA *in vitro***. (a) Representative electrophoreses of reaction products obtained with increasing concentrations of each DNA mimic (10, 50, 100, 500 and 1000 nM, lanes 4 to 8 respectively). Lanes 1: DNA alone, lanes 2: DNA+Top1, Lanes 3: same as lanes 2 in the presence of 50 µM camptothecin (CPT). (b) Quantitation of the results shown in panel (a). Results are the mean ± standard deviation of three independent sets of experiments.

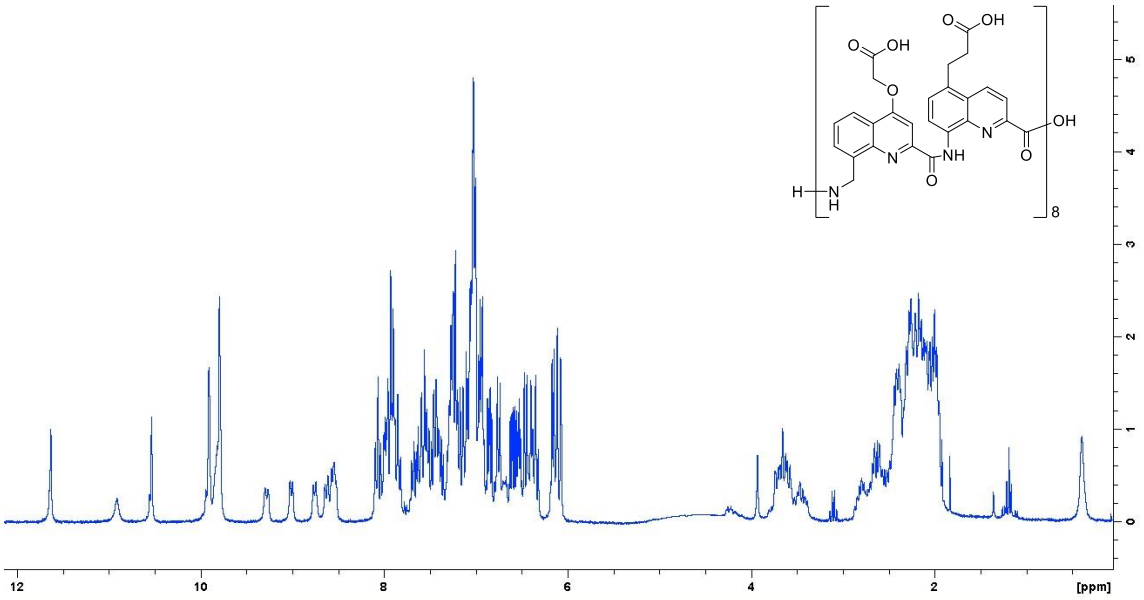


Supplementary Figure S8. **Effect of $(^{m}Q^{Pho}Q^{Pho})_8$ and its derivatives on *in vitro* HIV-IN concerted integration activity.** (a) Increasing concentrations of compounds were added to a typical concerted integration assay, respectively 0, 0.05, 0.1, 1 and 10 µM. The reaction products were loaded onto 1.4 % agarose gel and a representative set of experiments are shown in the figure (a). The position and structures of the donor substrate and different products obtained after half-site, full-site (FSI) and donor/donor integration (d/d) are displayed. (b) Quantitation of the results shown in panel (a). The circular FSI + HSI and linear FSI products were quantified on gel using the ImageJ software and the heterointegration (circular and linear HIS + FSI) was reported as the percentage of heterointegration quantified in the lack of drug. All the values are shown as the mean ± standard deviation of three independent sets of experiments.

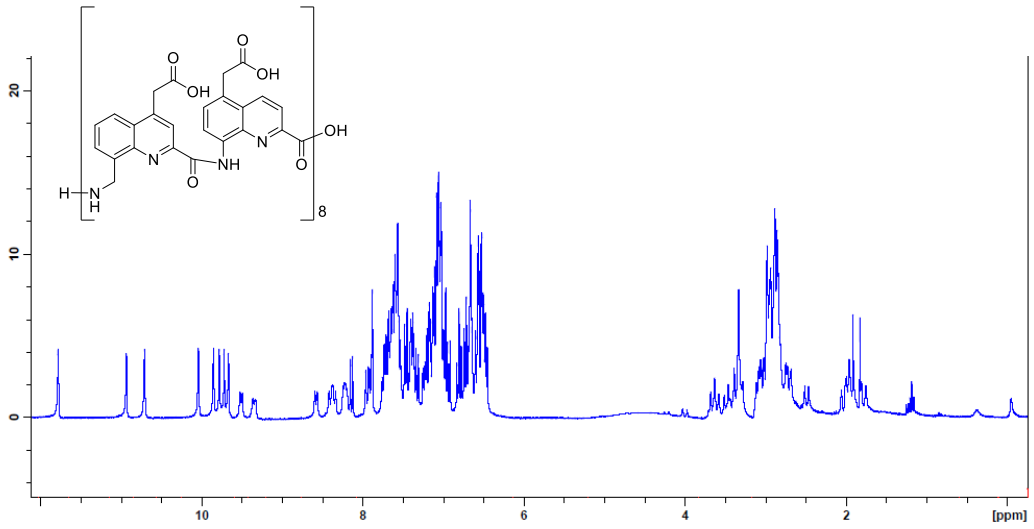
5. ^{1}H NMR,^{13}C NMR and ^{31}P spectra of purified new synthetic compounds



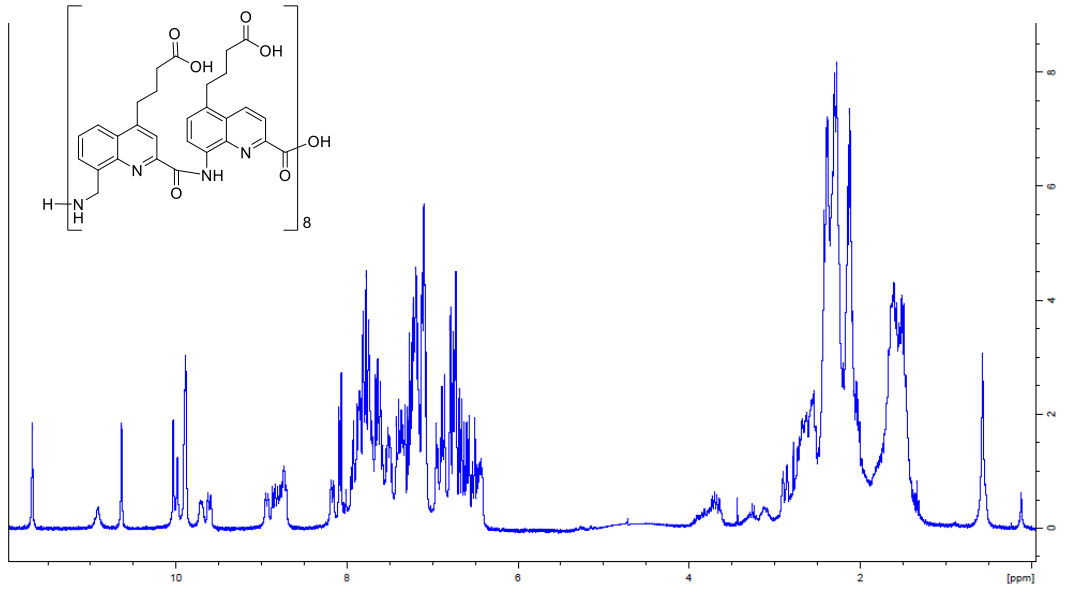
Supplementary Figure S9. ^{1}H NMR of **oligomer 2** (300 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3, 'watergate' water suppression applied at 4.71 ppm which may cause errors in the observed peak intensities in the vicinity of the suppressed peak).



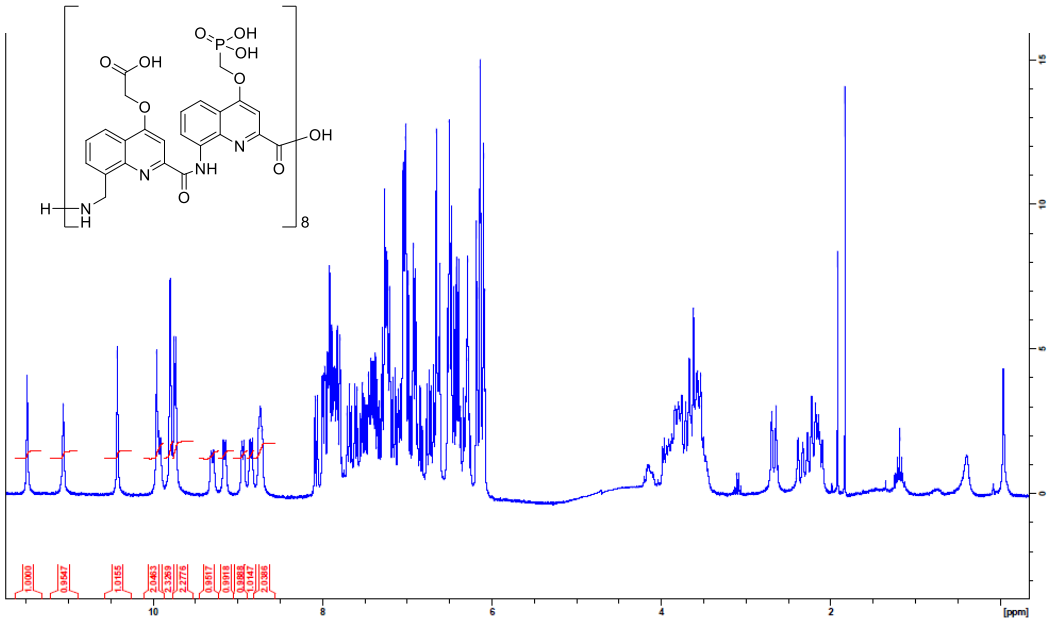
Supplementary Figure S10. ^{1}H NMR of **oligomer 3** (300 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3, 'watergate' water suppression applied at 4.71 ppm which may cause errors in the observed peak intensities in the vicinity of the suppressed peak).



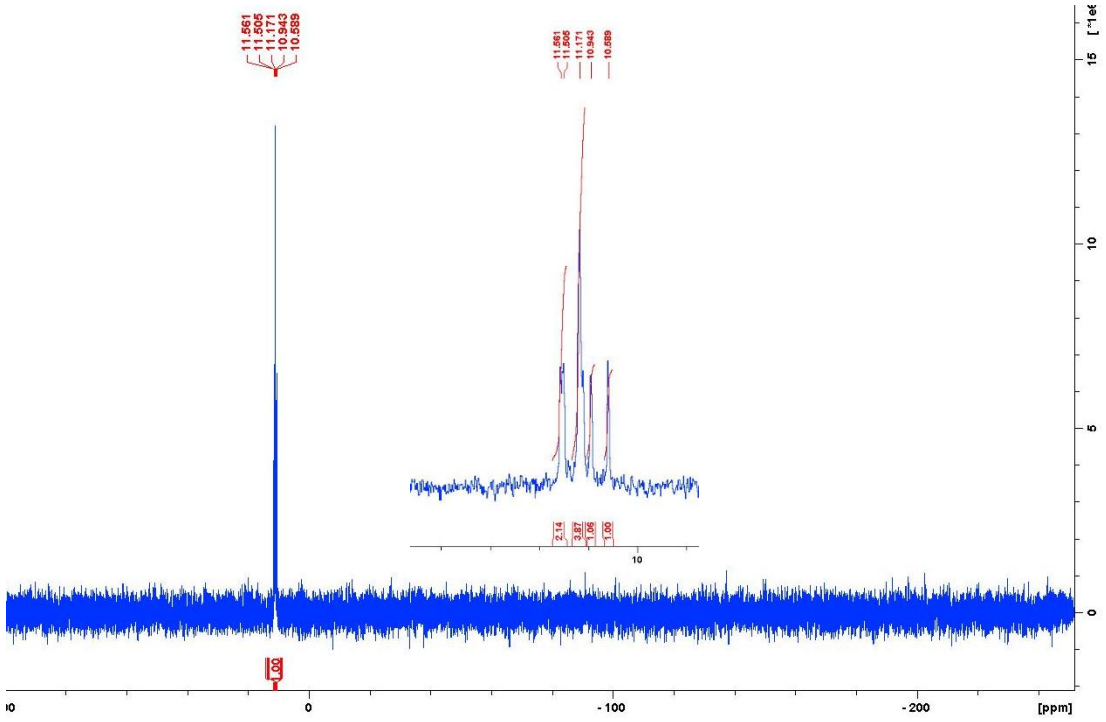
Supplementary Figure S11. ^{1}H NMR of **oligomer 4** (300 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3, 'watergate' water suppression applied at 4.71 ppm which may cause errors in the observed peak intensities in the vicinity of the suppressed peak).



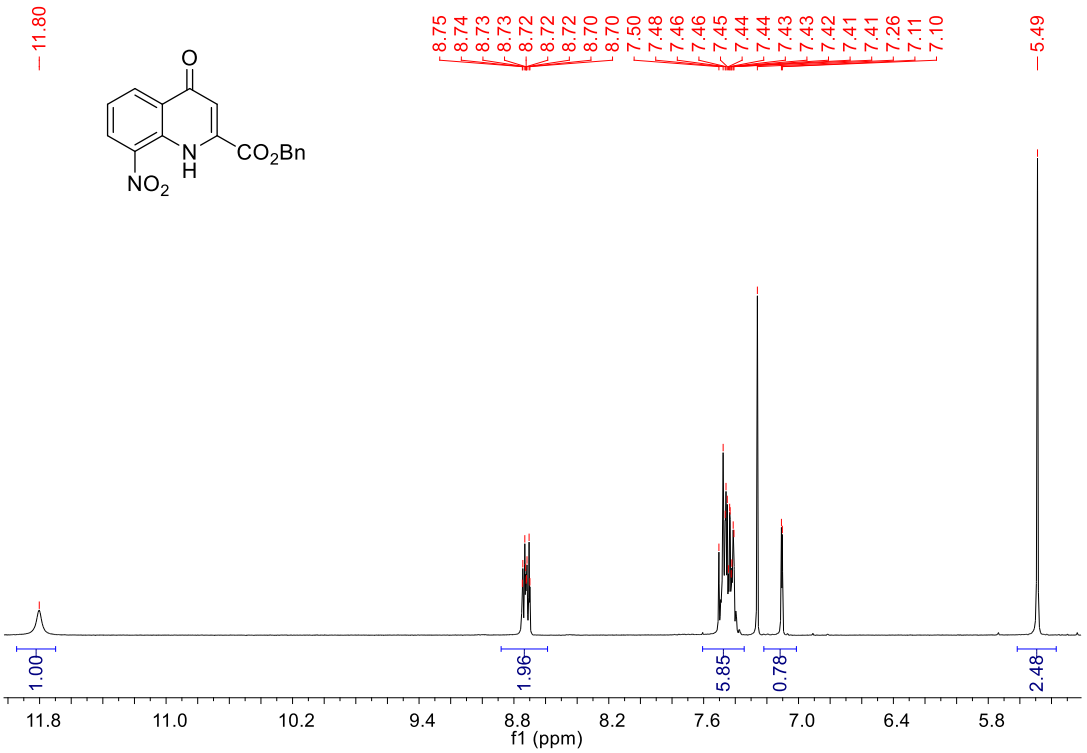
Supplementary Figure S12. ^{1}H NMR of **oligomer 5** (300 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3, 'watergate' water suppression applied at 4.71 ppm which may cause errors in the observed peak intensities in the vicinity of the suppressed peak).



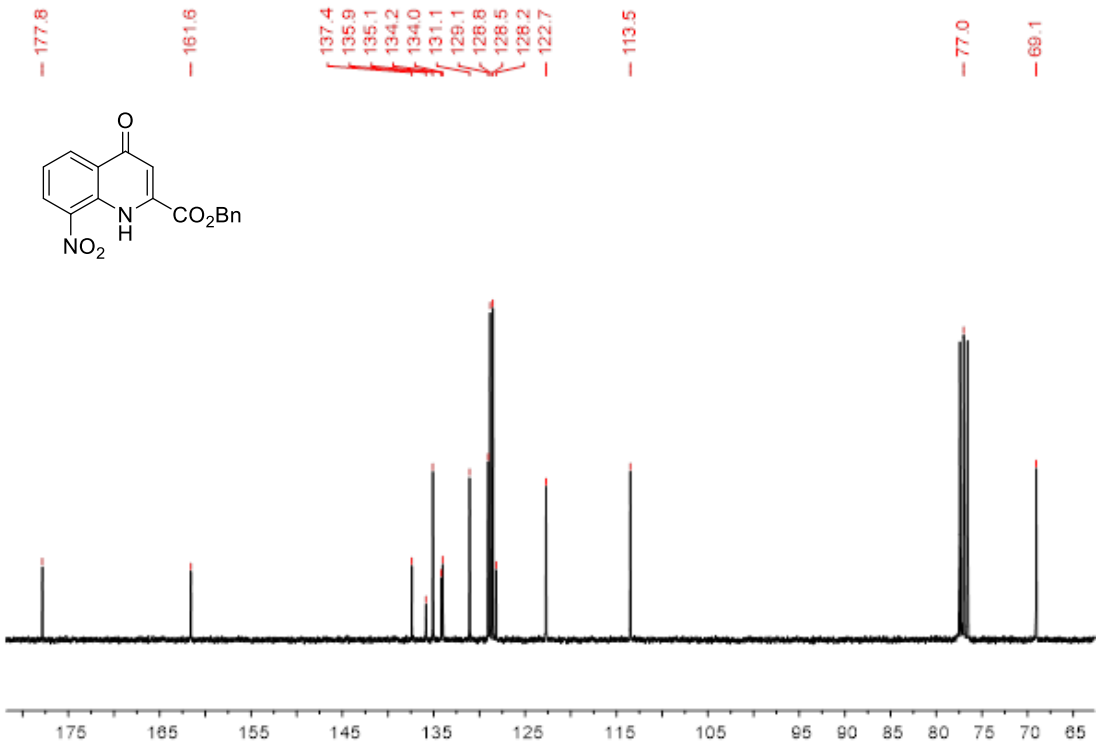
Supplementary Figure S13. ^{1}H NMR of **oligomer 6** (300 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3, 'watergate' water suppression applied at 4.71 ppm which may cause errors in the observed peak intensities in the vicinity of the suppressed peak).



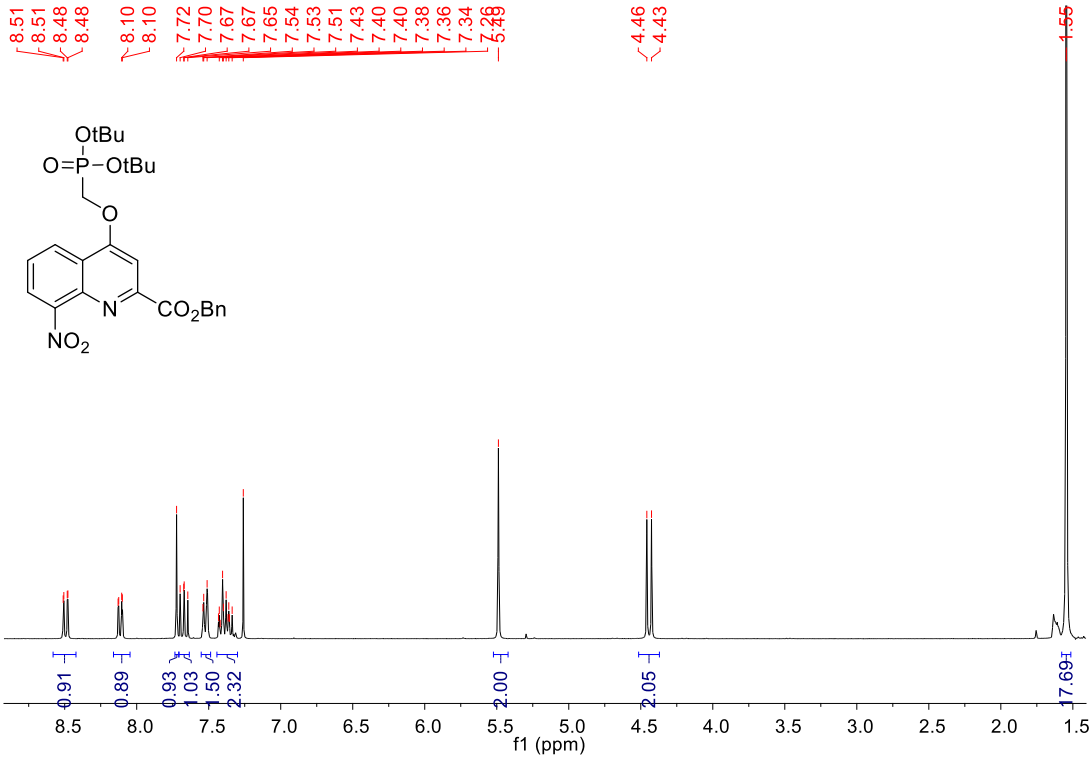
Supplementary Figure S14. ^{31}P NMR of **oligomer 6** (122 MHz, H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3).



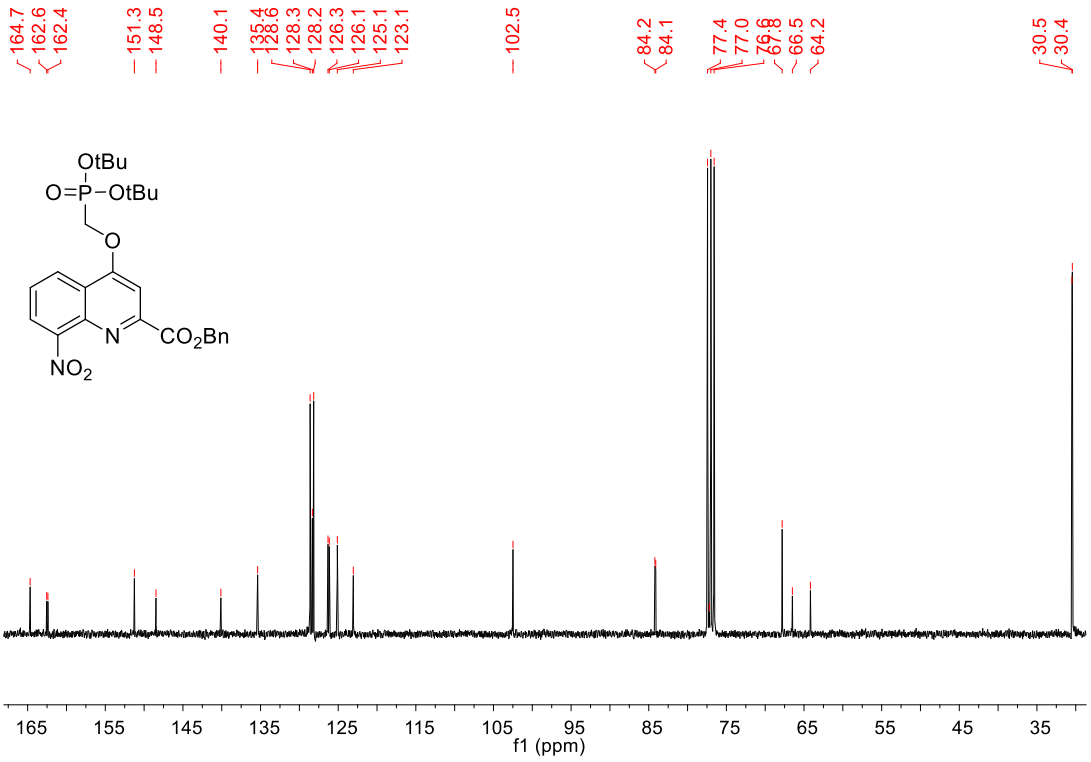
Supplementary Figure S15. ^{1}H NMR of **7** (300 MHz, $CDCl_3$).



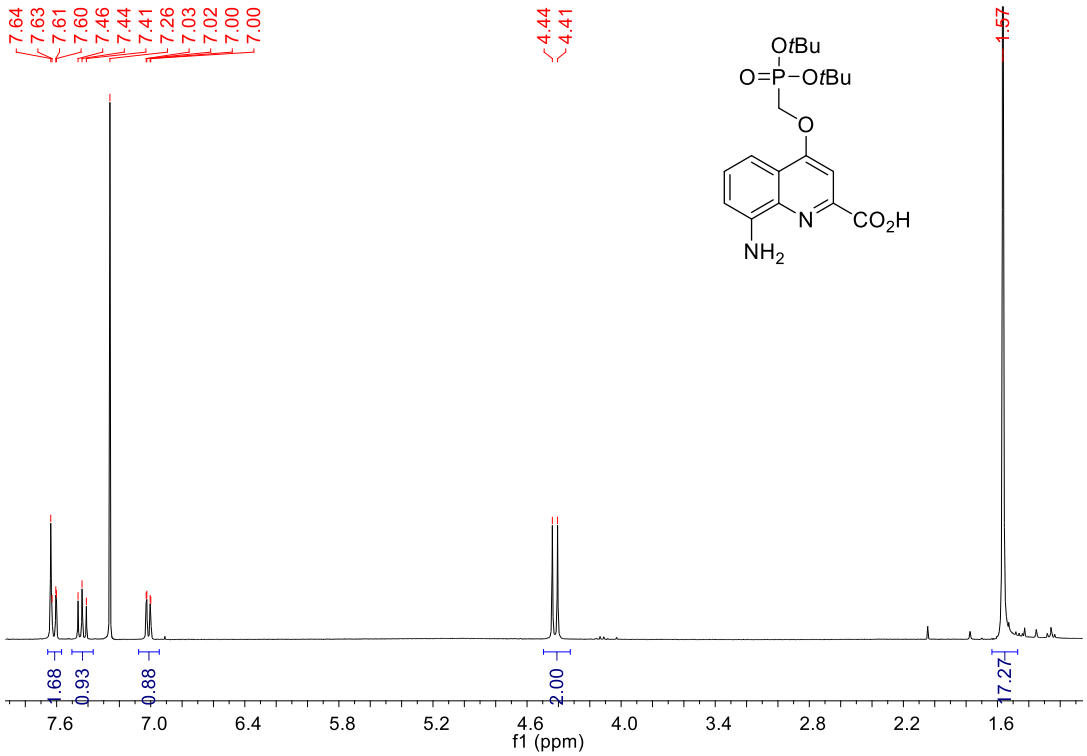
Supplementary Figure S16. ^{13}C NMR of **7** (75 MHz, $CDCl_3$).



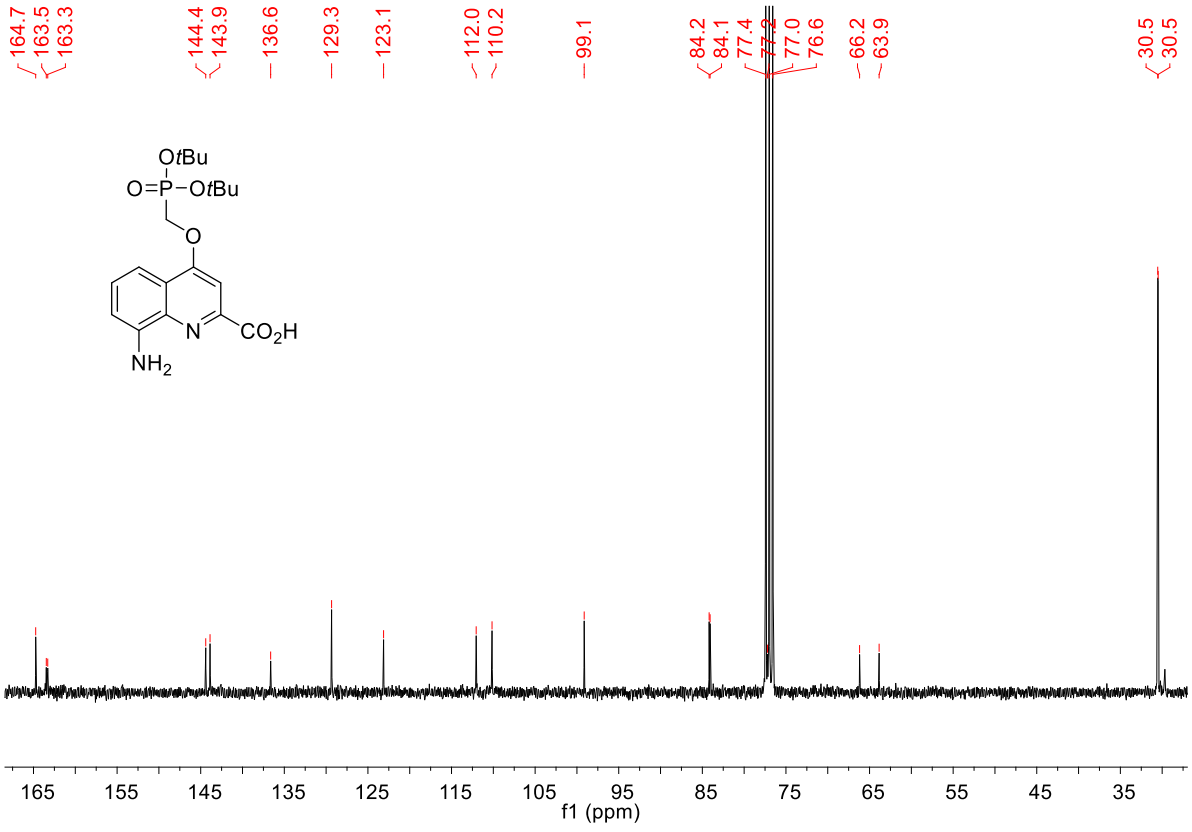
Supplementary Figure S17. ^{1}H NMR of **8** (300 MHz, $CDCl_3$).



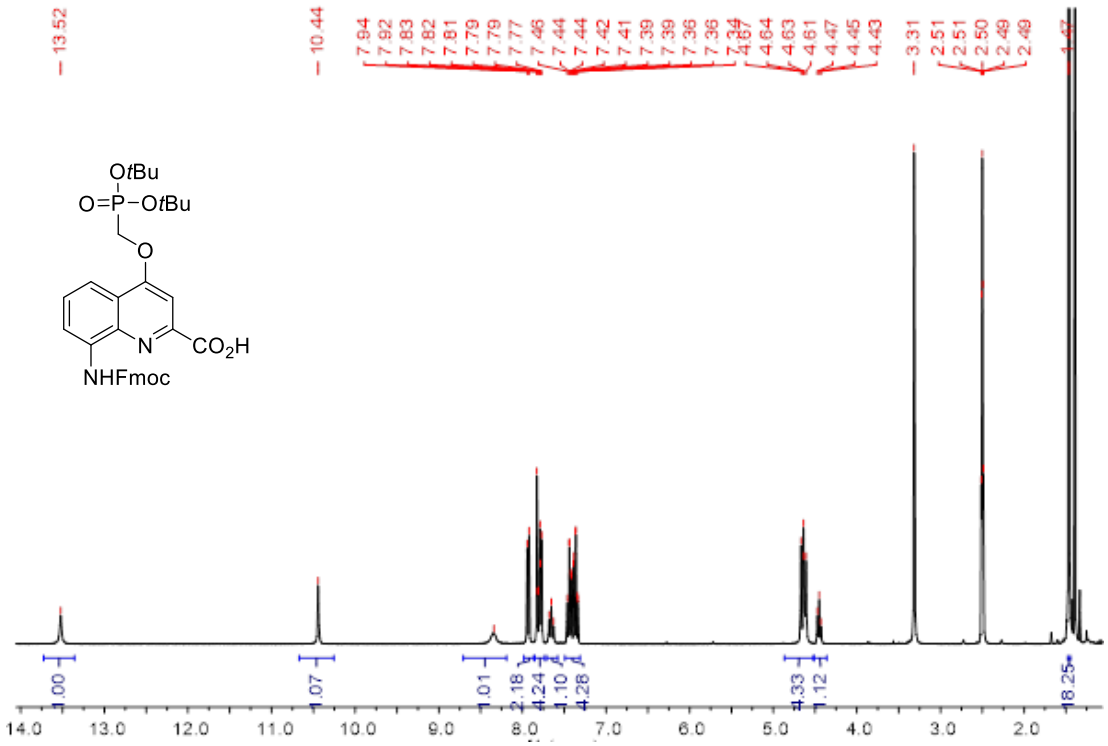
Supplementary Figure S18. ^{13}C NMR of **8** (75 MHz, $CDCl_3$).



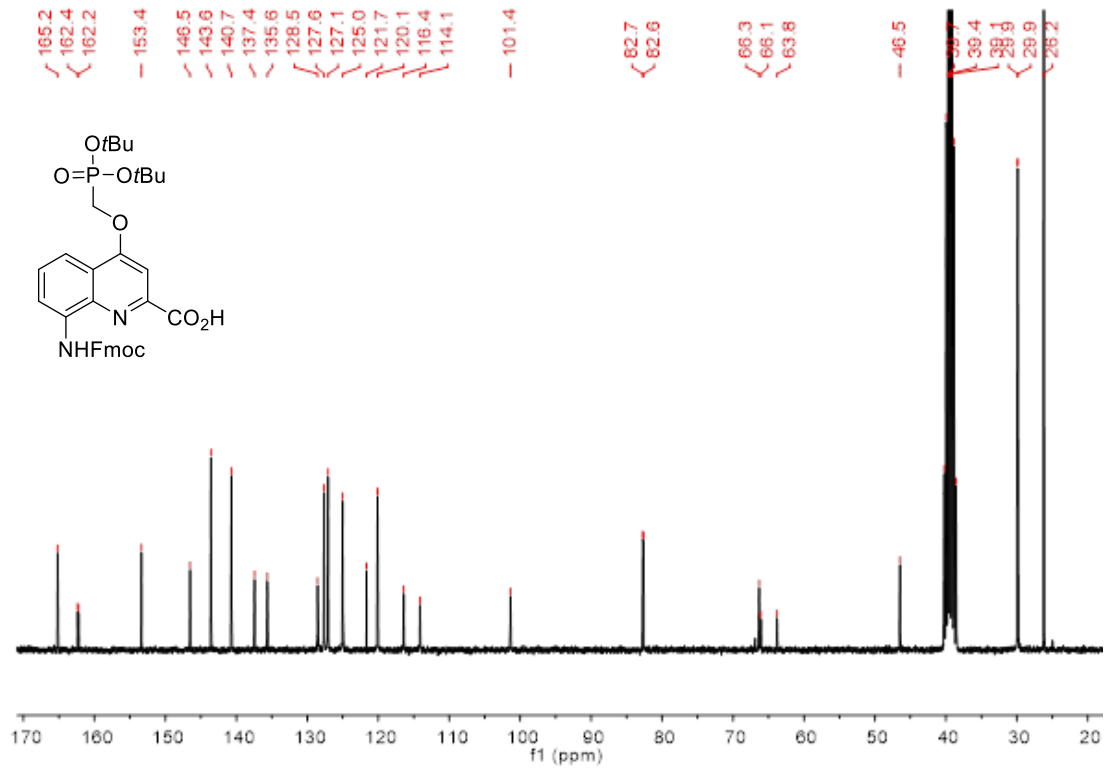
Supplementary Figure S19. ^{1}H NMR of **9** (300 MHz, $CDCl_3$).



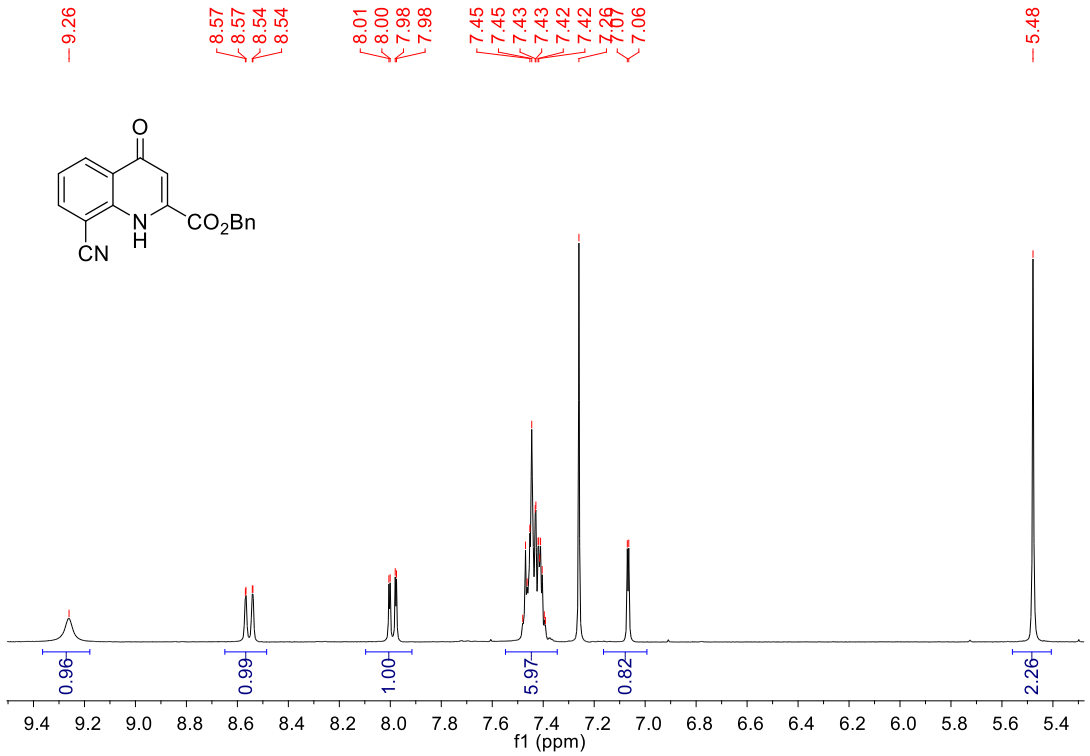
Supplementary Figure S20. ^{13}C NMR of **9** (75 MHz, $CDCl_3$).



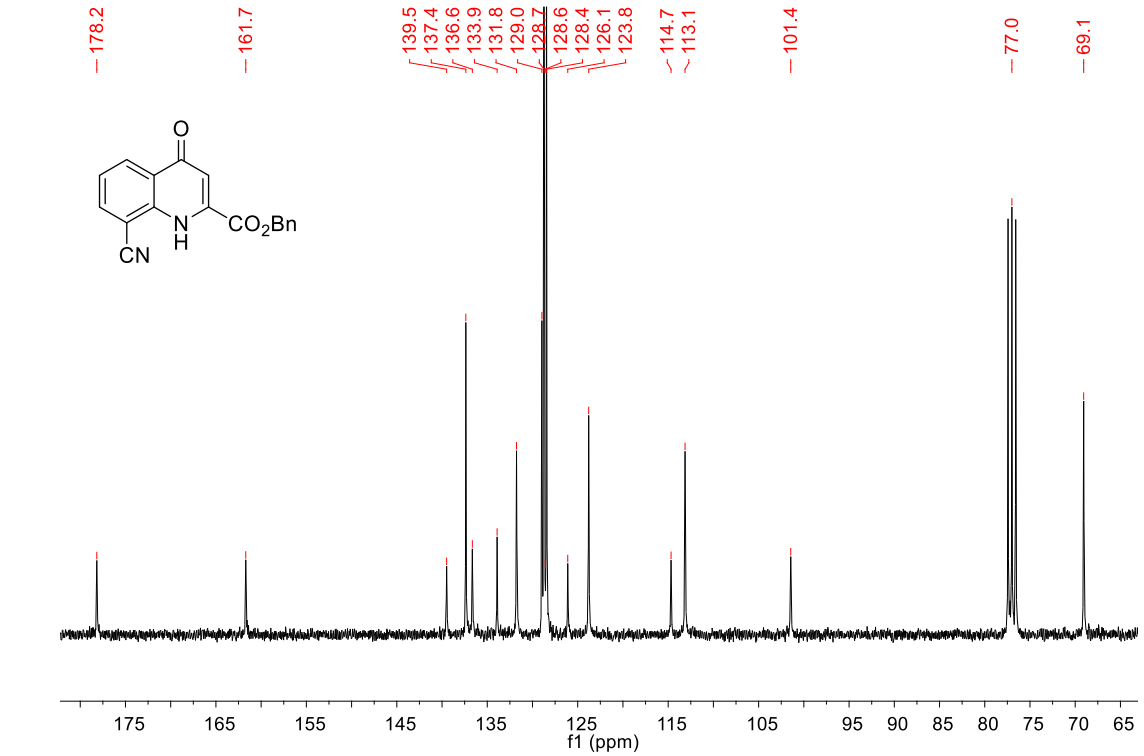
Supplementary Figure S21. ^{1}H NMR of **Fmoc(Q^{Pho})** (300 MHz, DMSO-d$_6$).



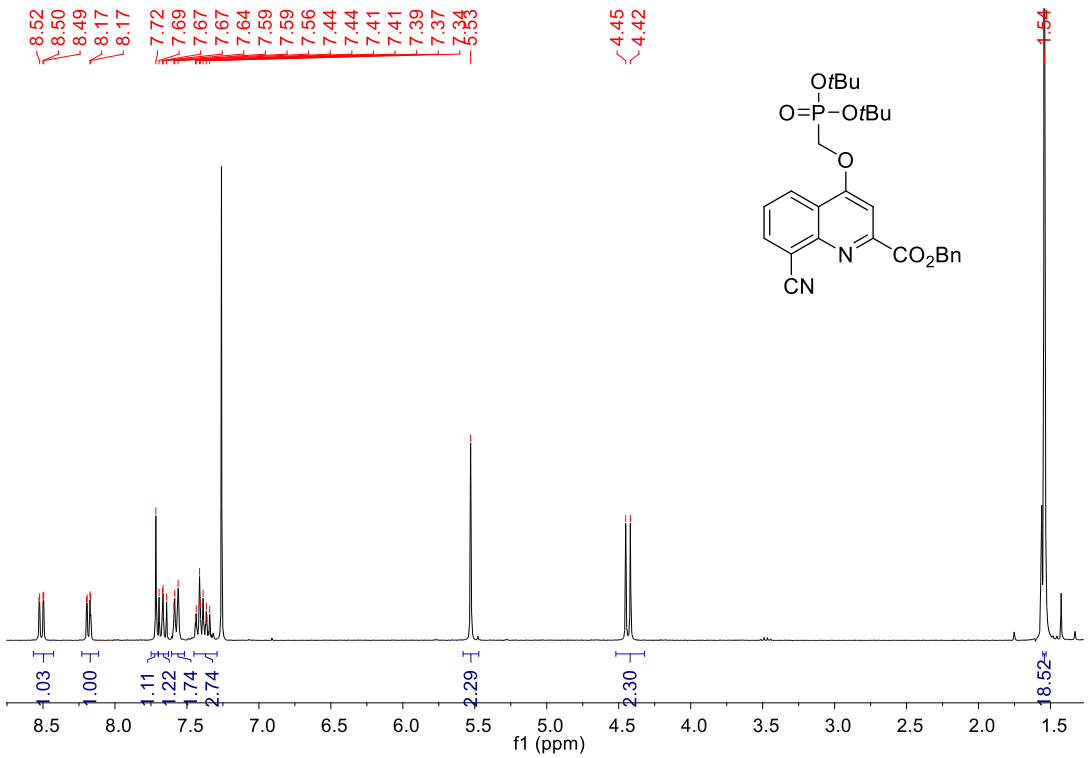
Supplementary Figure S22. ^{13}C NMR of **Fmoc(Q^{Pho})** (75 MHz, DMSO-d$_6$).



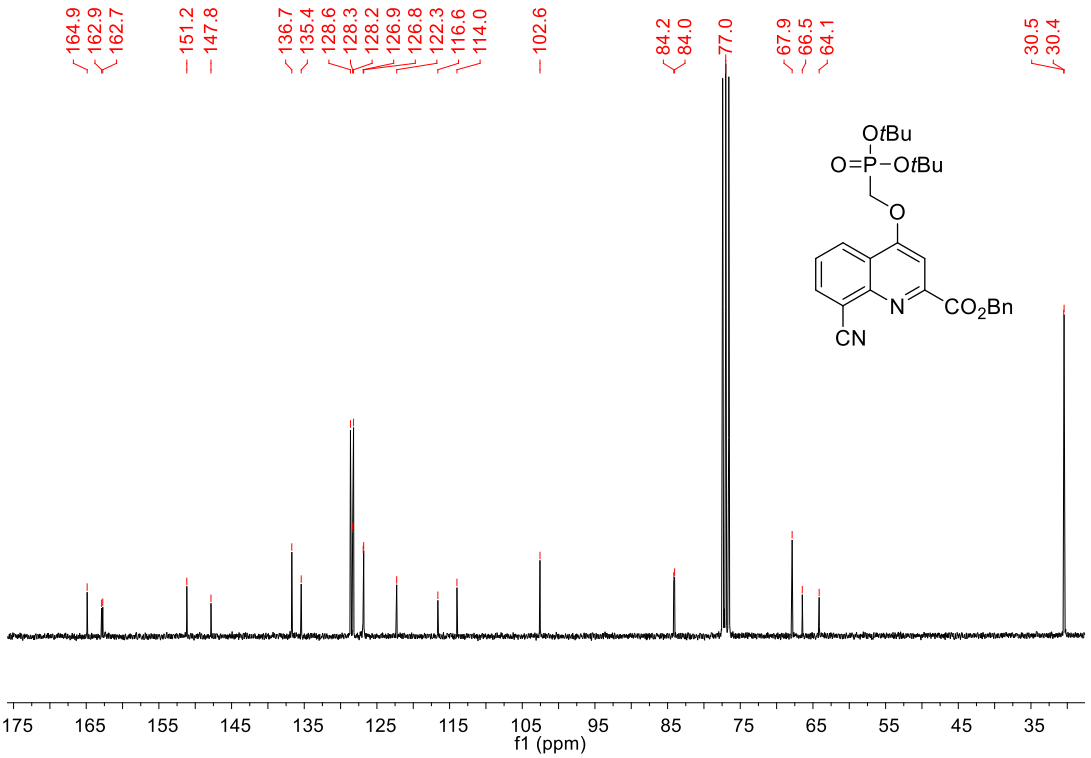
Supplementary Figure S23. ^{1}H NMR of **10** (300 MHz, $CDCl_3$).



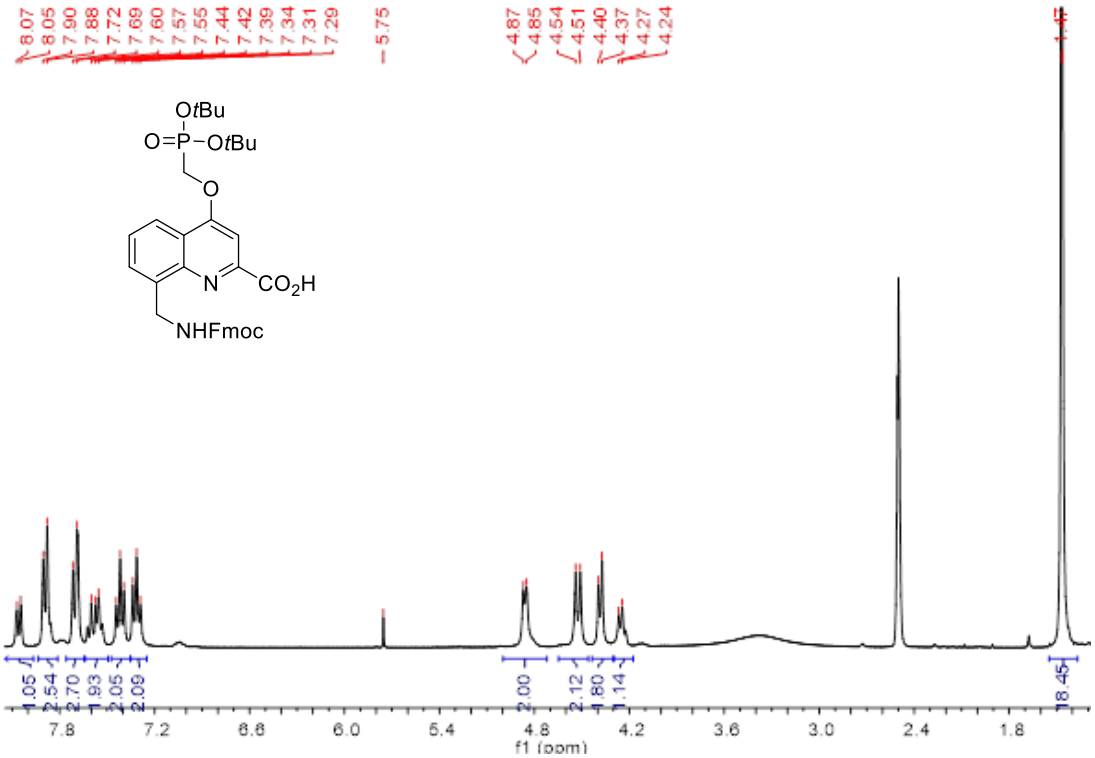
Supplementary Figure S24. ^{13}C NMR of **10** (75 MHz, $CDCl_3$).



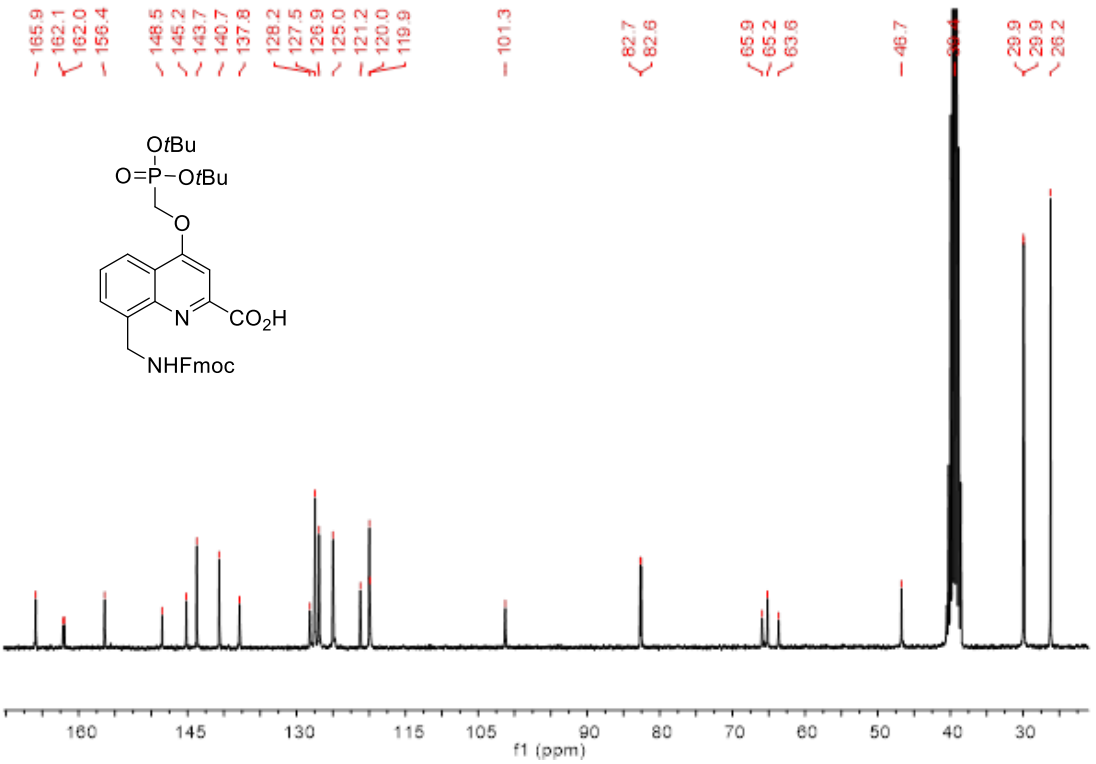
Supplementary Figure S25. ^{1}H NMR of **11** (300 MHz, $CDCl_3$).



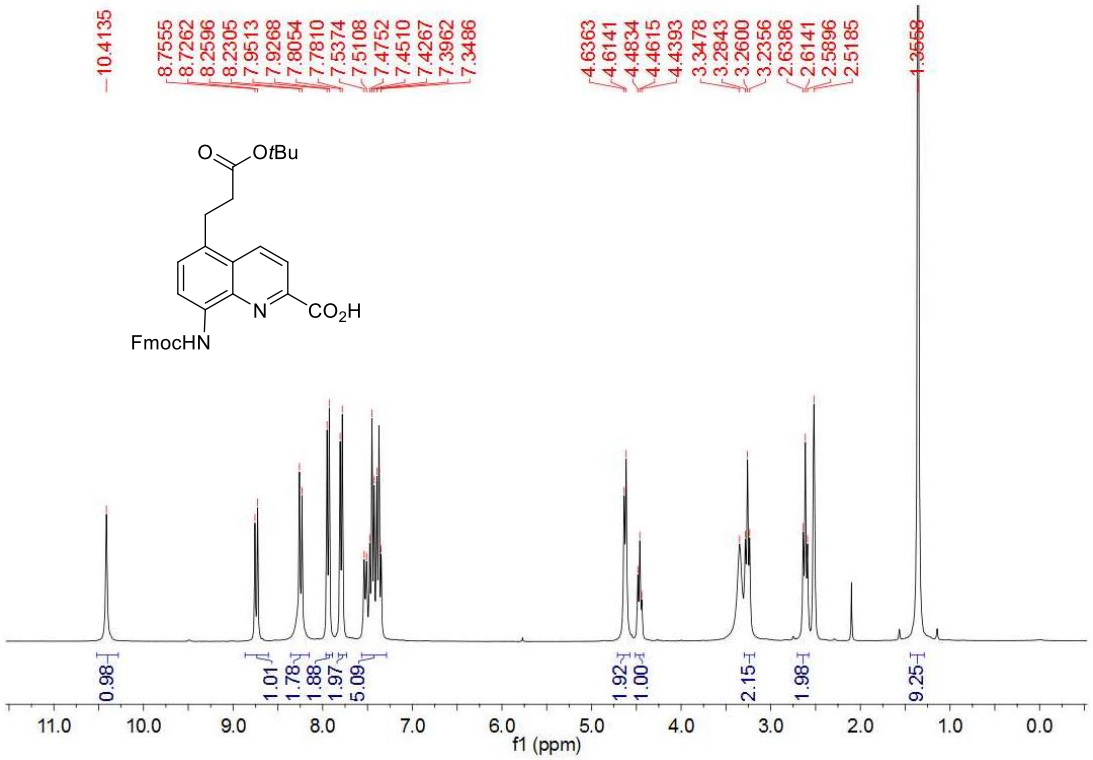
Supplementary Figure S26. ^{13}C NMR of **11** (75 MHz, $CDCl_3$).



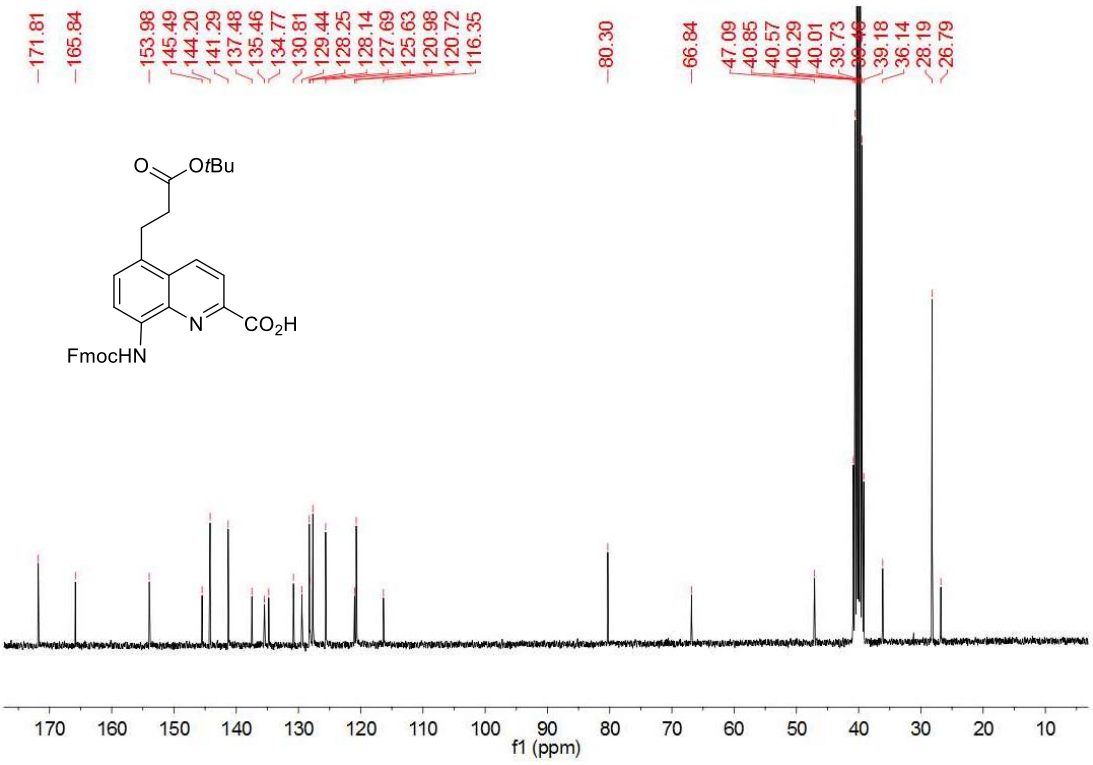
Supplementary Figure S27. ^{1}H NMR of **Fmoc(**^{m}**Q**^{Pho}**)** (300 MHz, DMSO-d_6).



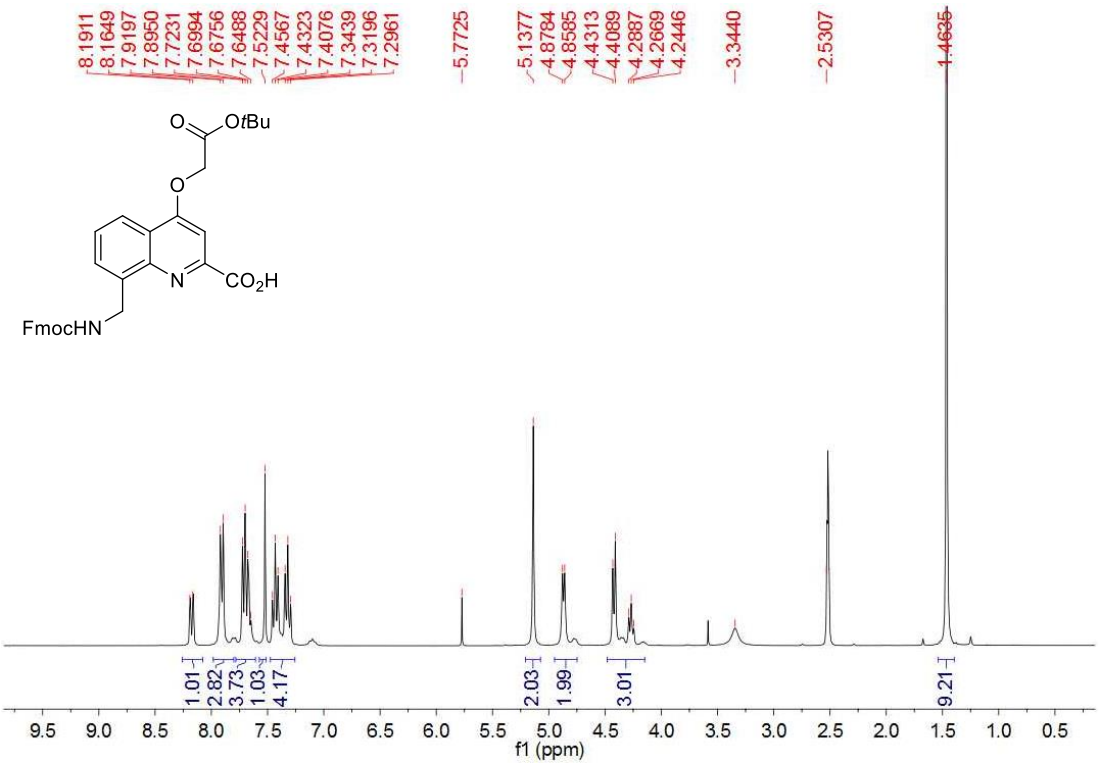
Supplementary Figure S28. ^{13}C NMR of **Fmoc(**^{m}**Q**^{Pho}**)** (75 MHz, DMSO-d_6).



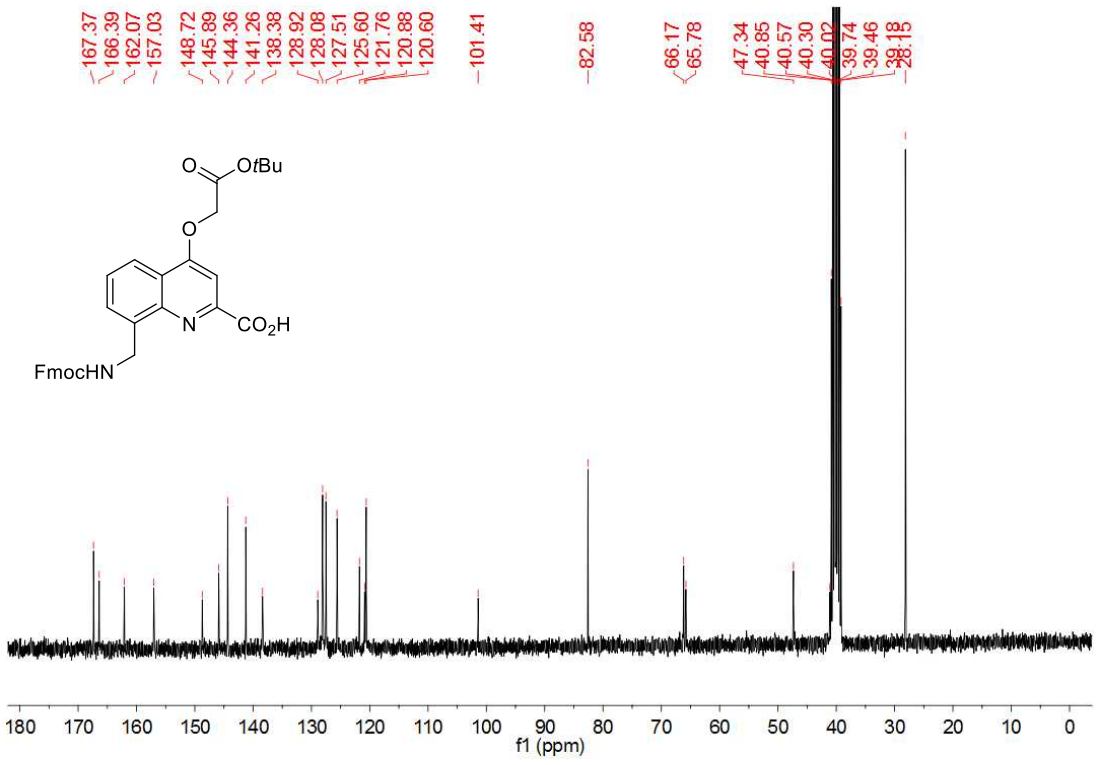
Supplementary Figure S29. ^{1}H NMR of **Fmoc**(**Q^{5Prop}**) (300 MHz, DMSO-d$_6$).



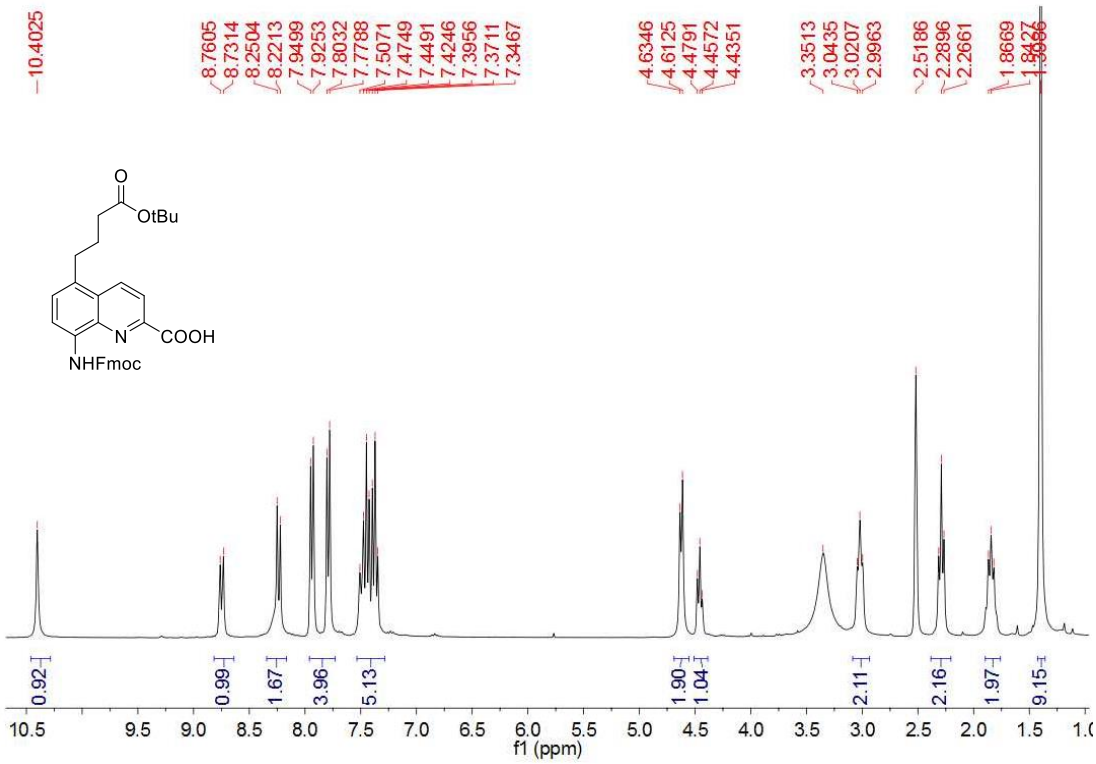
Supplementary Figure S30. ^{13}C NMR of **Fmoc**(**Q^{5Prop}**) (75 MHz, DMSO-d$_6$).



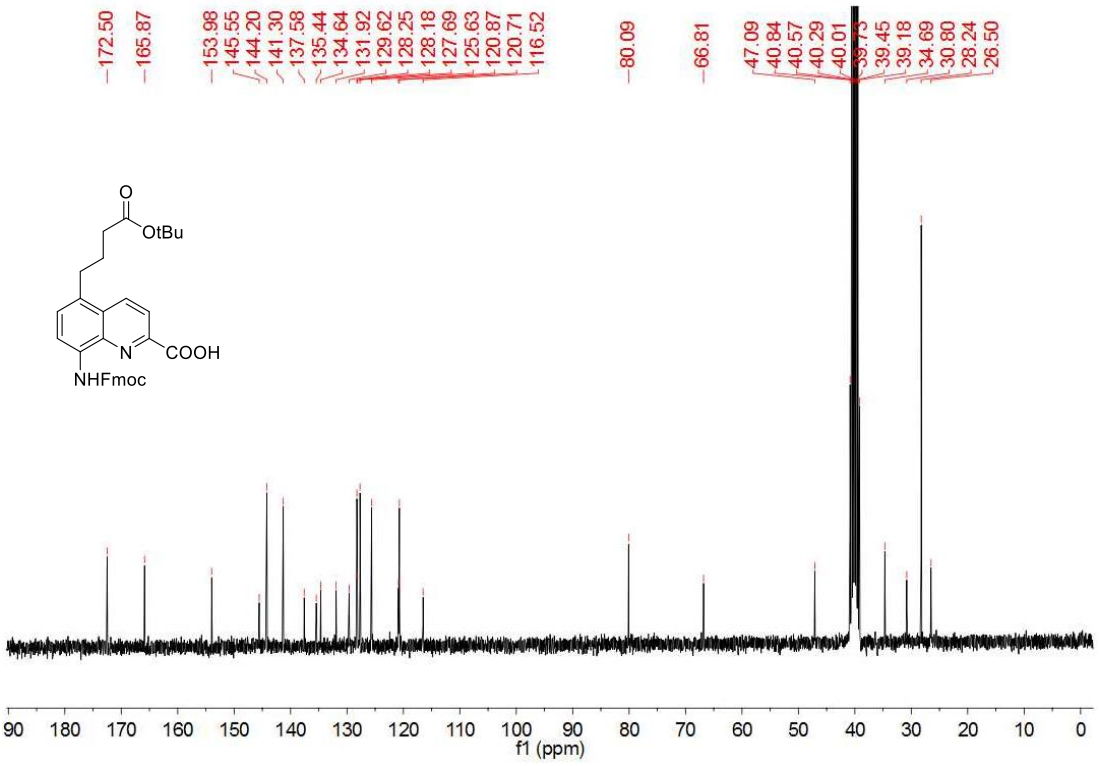
Supplementary Figure S31. ^{1}H NMR of **Fmoc**(**^{m}Q^{OAc}**) (300 MHz, DMSO-d_6).



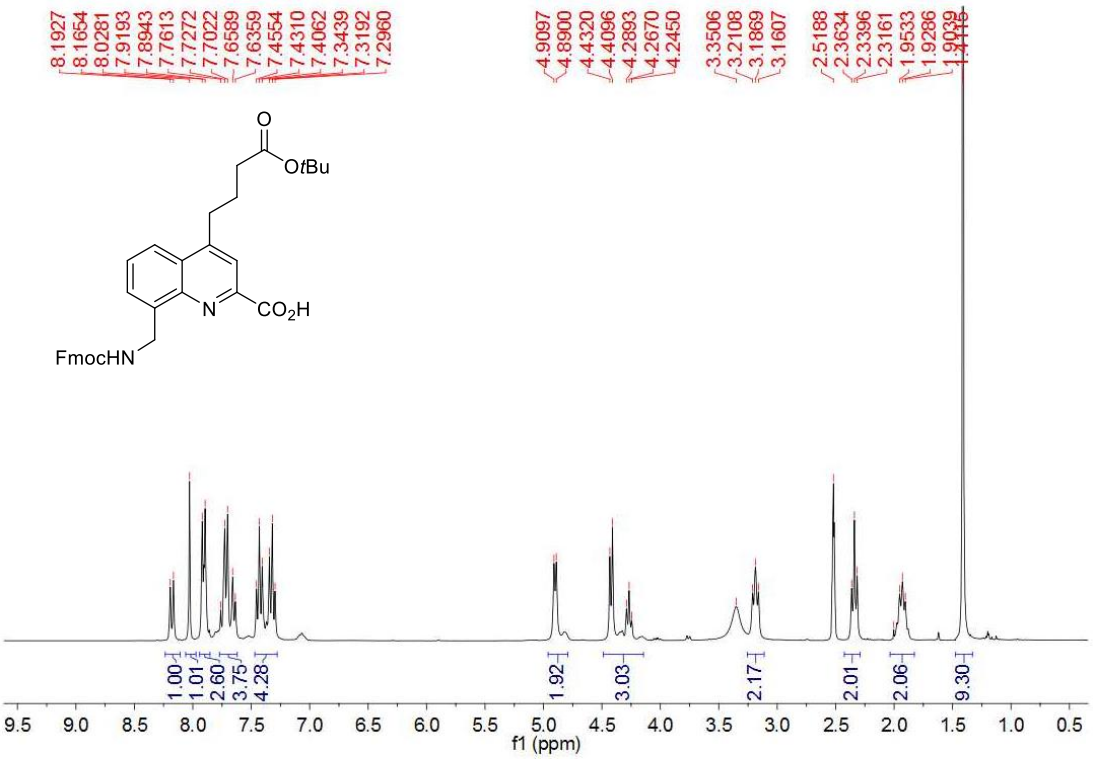
Supplementary Figure S32. ^{13}C NMR of **Fmoc**(**^{m}Q^{OAc}**) (75 MHz, DMSO-d_6).



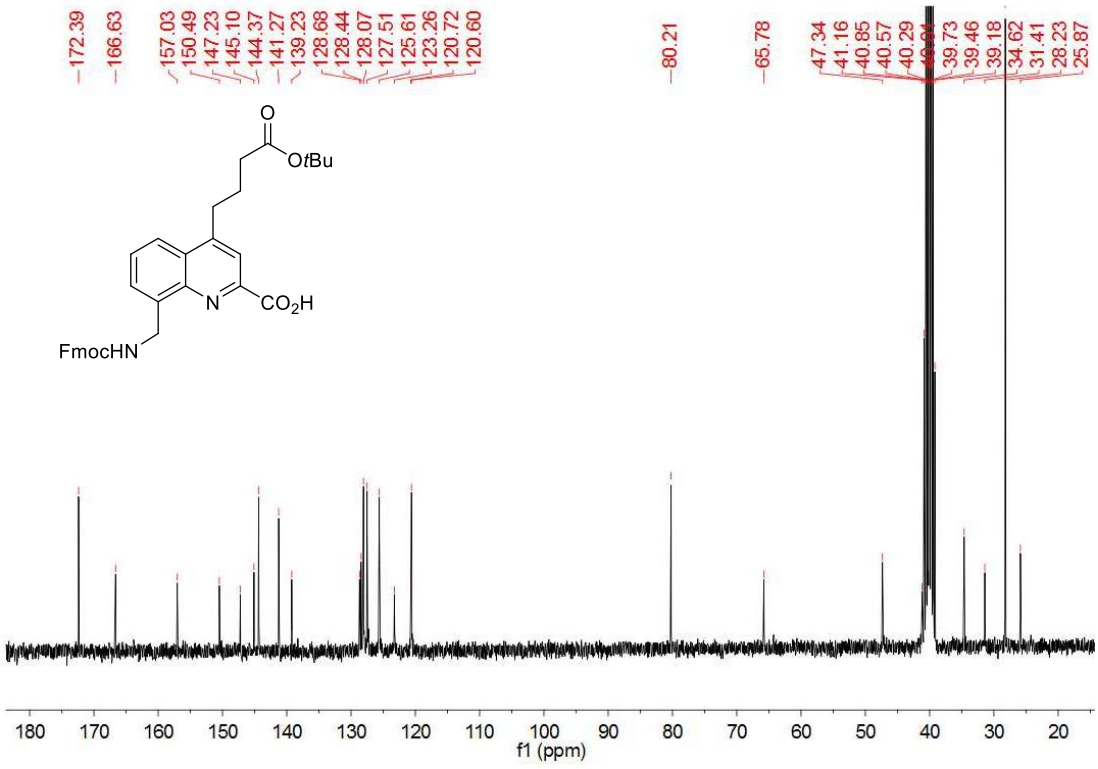
Supplementary Figure S33. ^{1}H NMR of **Fmoc(Q^{5But})** (300 MHz, DMSO-d$_6$).



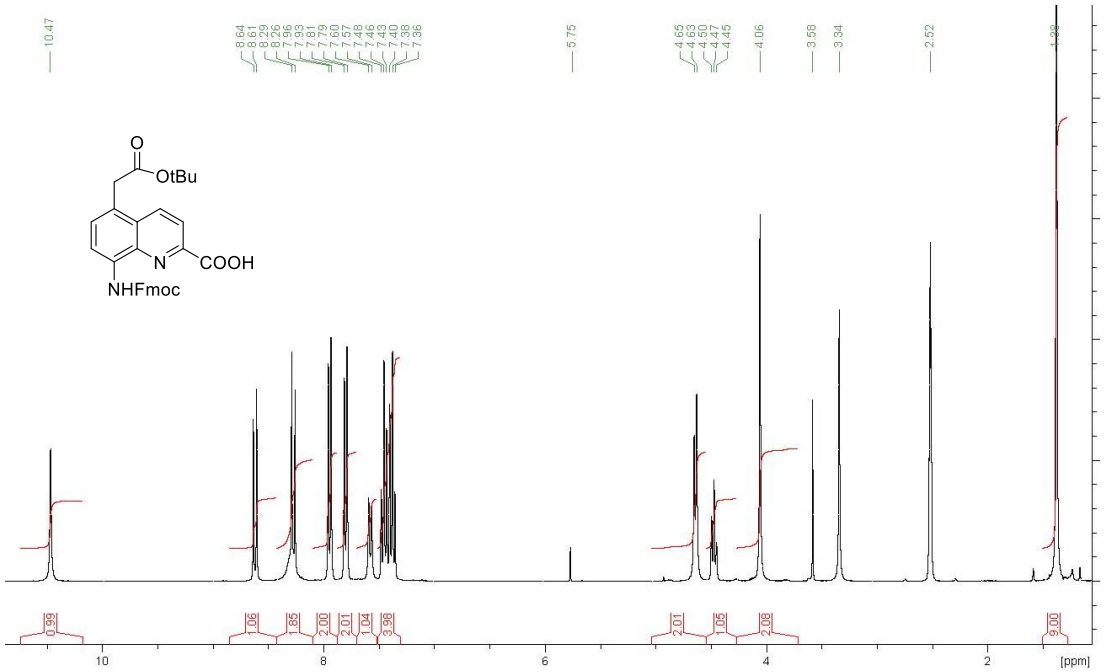
Supplementary Figure S34. ^{13}C NMR of **Fmoc(Q^{5But})** (75 MHz, DMSO-d$_6$).



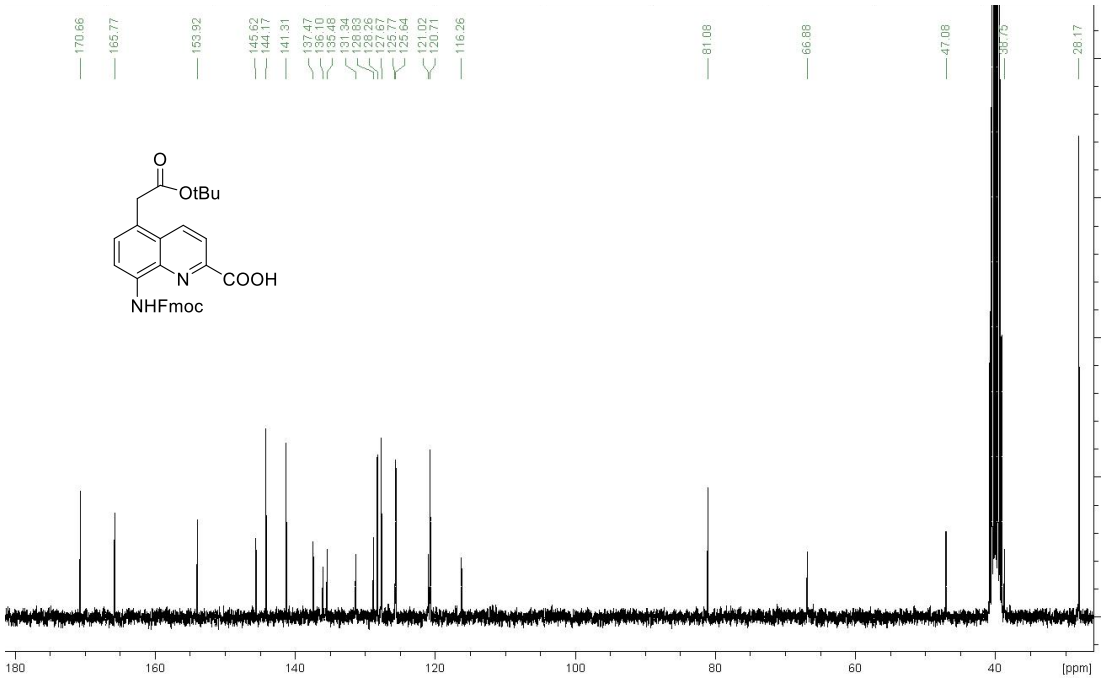
Supplementary Figure S35. ^{1}H NMR of **Fmoc(**^{m}**Q**^{But}**)** (300 MHz, DMSO-d_6).



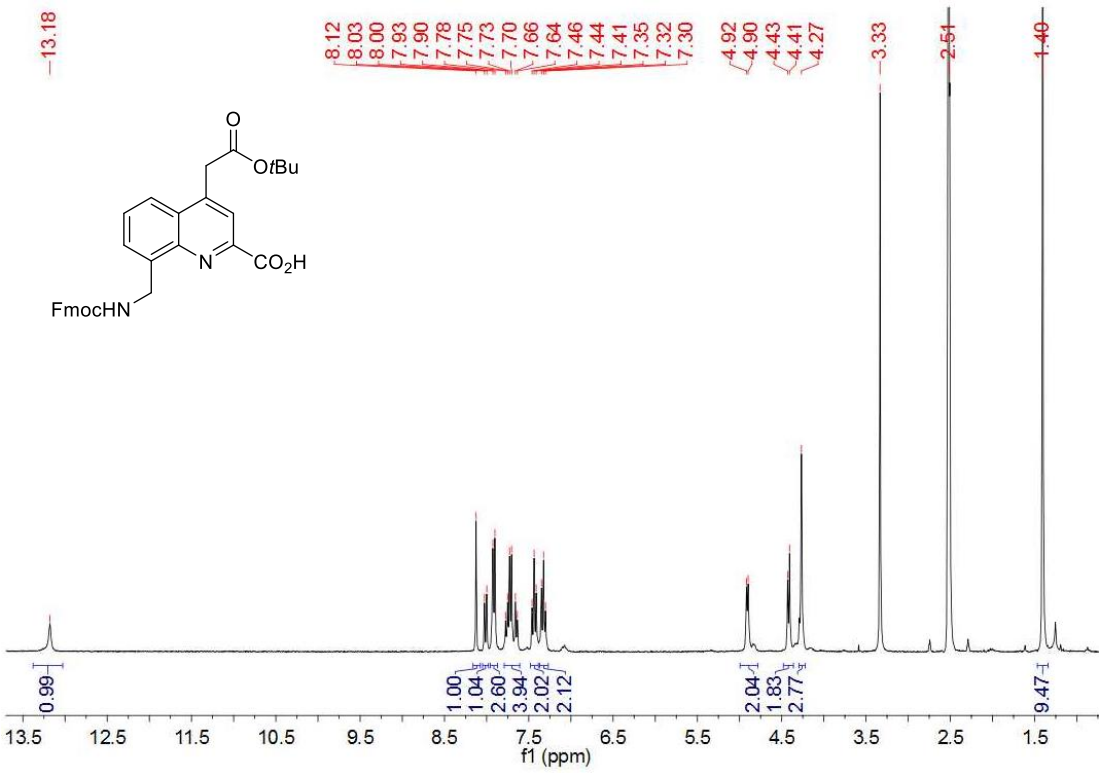
Supplementary Figure S36. ^{13}C NMR of **Fmoc(**^{m}**Q**^{But}**)** (75 MHz, DMSO-d_6).



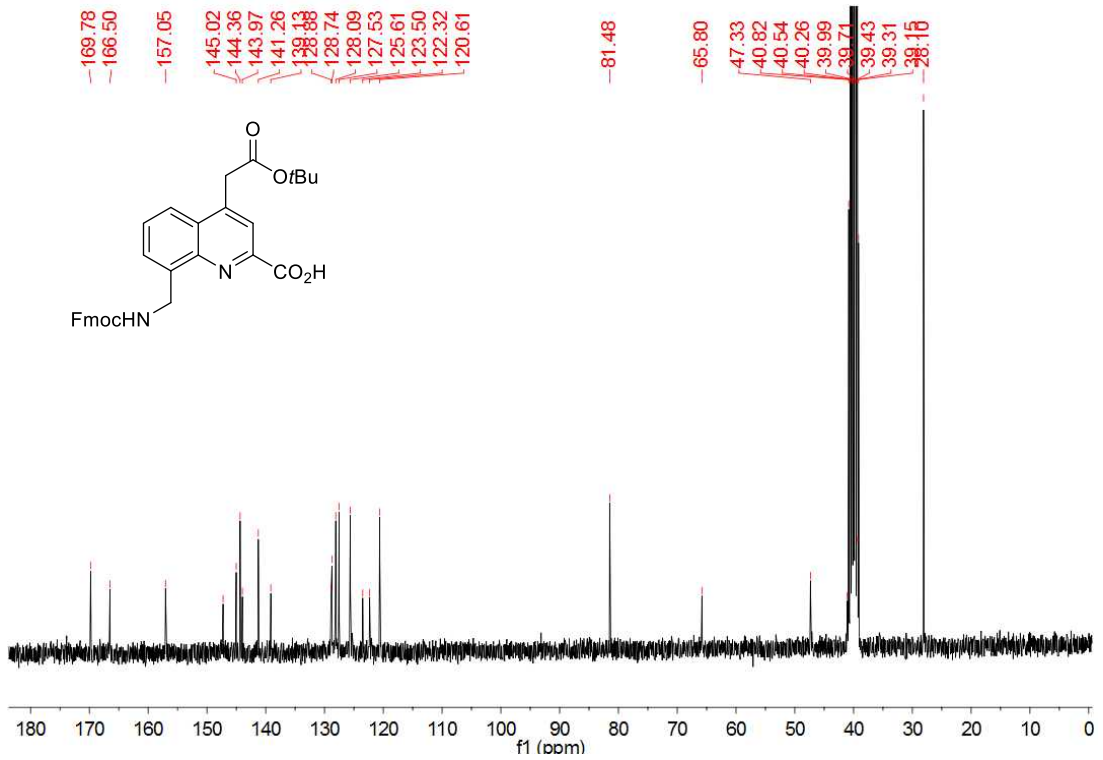
Supplementary Figure S37. ^{1}H NMR of **Fmoc**(**Q^{5Ac}**) (300 MHz, DMSO-d$_6$).



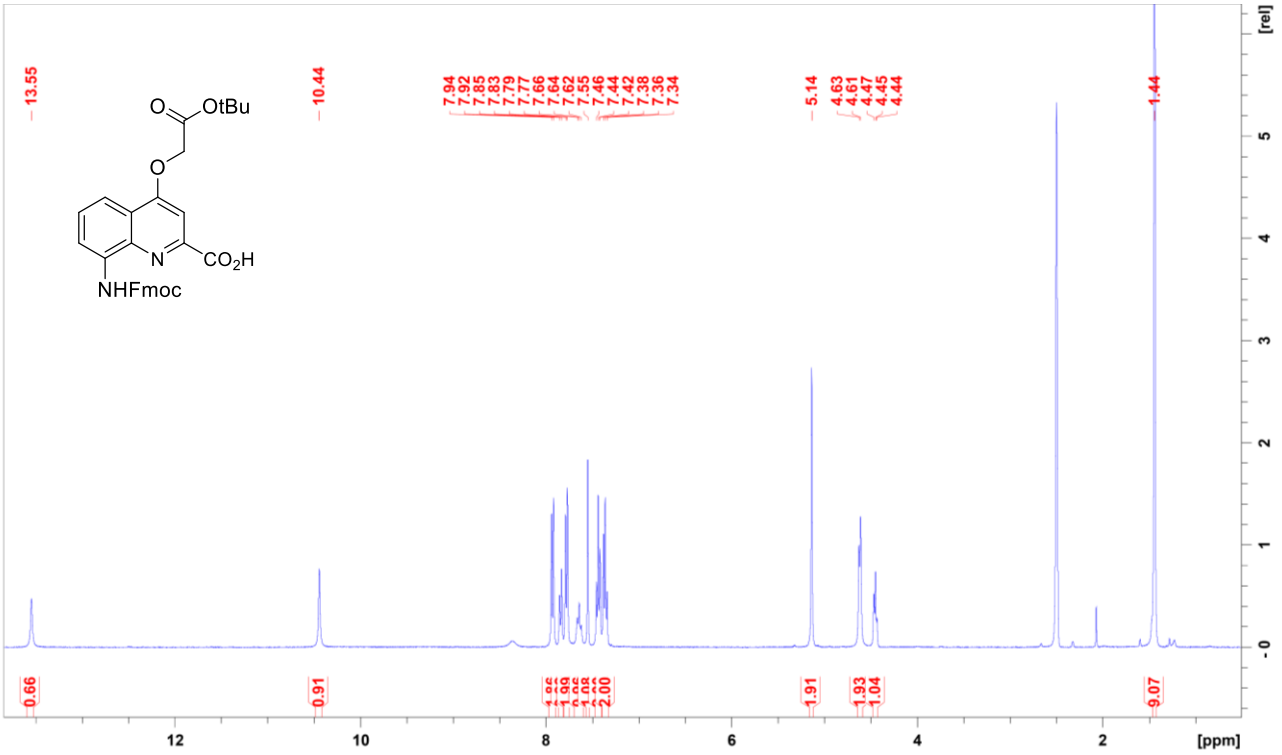
Supplementary Figure S38. ^{13}C NMR of **Fmoc**(**Q^{5Ac}**) (75 MHz, DMSO-d$_6$).



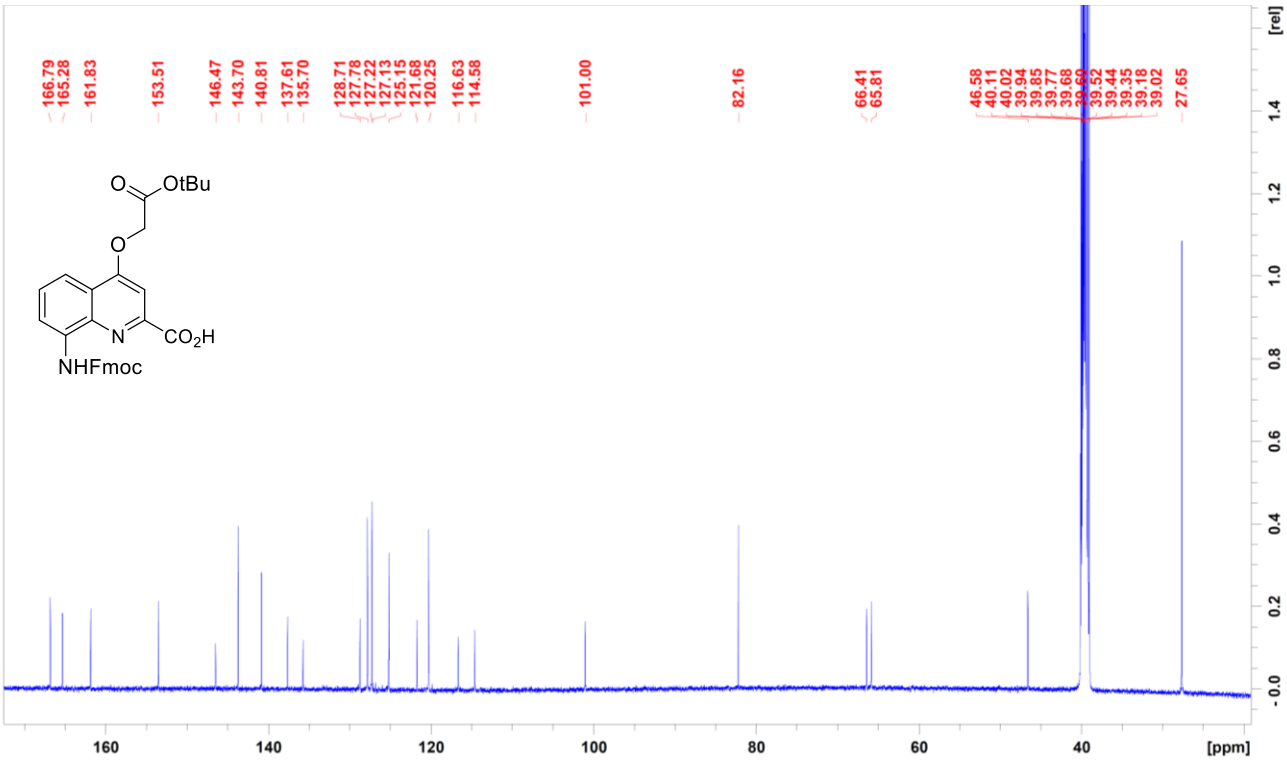
Supplementary Figure S39. ^{1}H NMR of **Fmoc(**^{m}**Q**^{Ac}**)** (300 MHz, DMSO-d_6).



Supplementary Figure S40. ^{13}C NMR of **Fmoc(**^{m}**Q**^{Ac}**)** (75 MHz, DMSO-d_6).

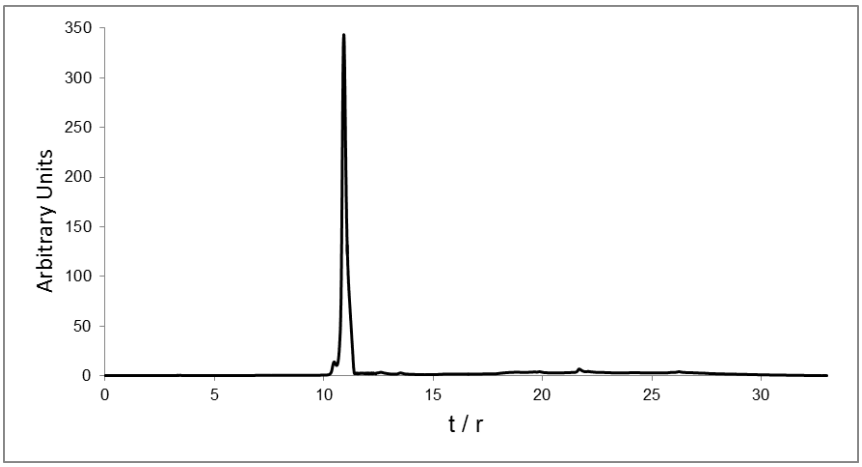


Supplementary Figure S41. ^{1}H NMR of **Fmoc(Q^{OAc})** (400 MHz, DMSO-d$_6$).

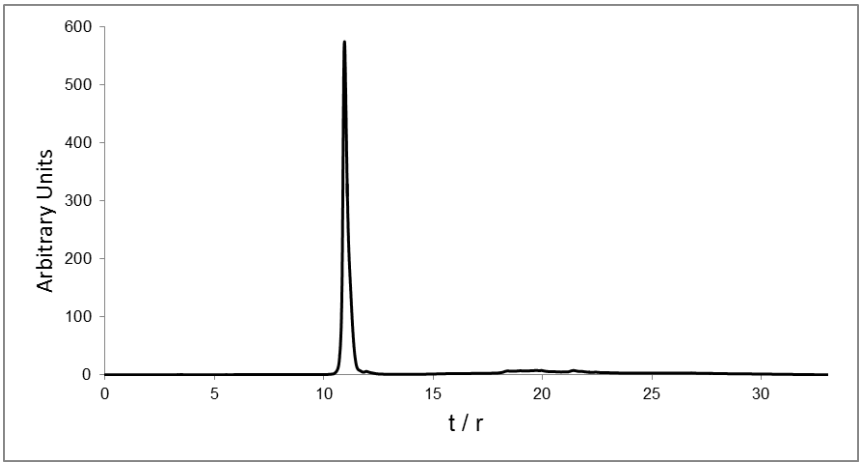


Supplementary Figure S42. ^{13}C NMR of **Fmoc(Q^{OAc})** (125 MHz, DMSO-d$_6$).

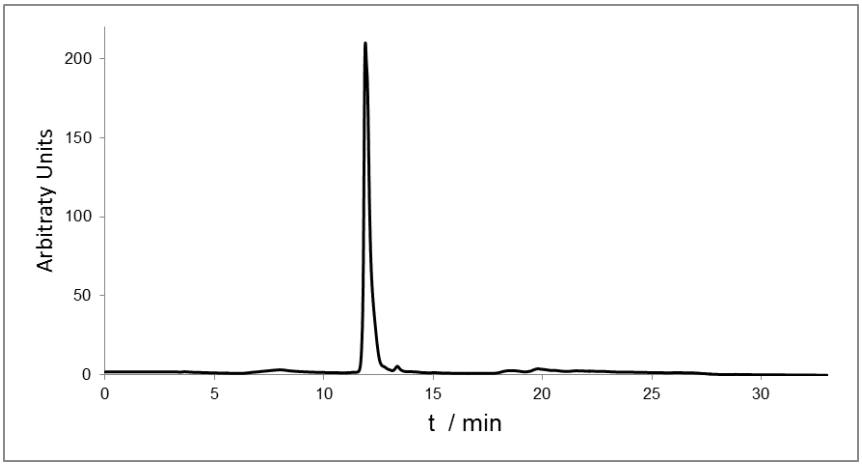
6. HPLC traces of purified new synthetic compounds



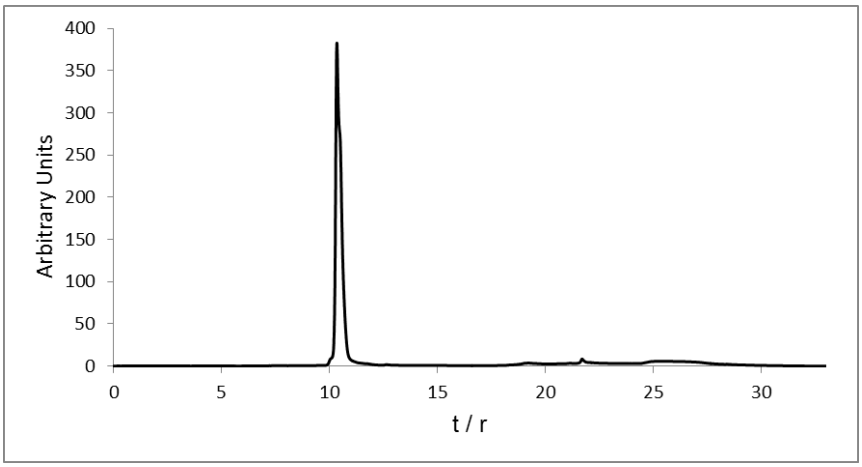
Supplementary Figure S43. Analytical RP-HPLC profile of **oligomer 2**. Elution conditions: Varian Pursuit C18 column (250 x 4.6 mm, 5 µm); solvents A and B were prepared as follows: a stock 50 mM aqueous triethyl-ammonium acetate buffer solution at pH 8.7 was prepare by dissolving 3 mL of glacial acetic acid in 950 mL water and, while mixing, adding freshly distilled triethylamine (6.5 mL). After adjusting the pH to 8.7 with trimethylamine, the volume was finally adjusted to 1 L with water. Solvent A was prepared by dilution of the stock buffer solution with water 1:3 (vol/vol) to a final concentration of 12.5 mM triethyl-ammonium acetate, pH 8.7, in water. Solvent B was prepared by dilution of the stock buffer solution with acetonitrile 1:3 (vol/vol) to a final composition of 12.5 mM triethyl-ammonium acetate, pH 8.7, in water:acetonitrile 3:1 vol/vol.**;** gradient program: from 0% to 20% of solvent B in 20 min; UV detection at λ = 300 nm.



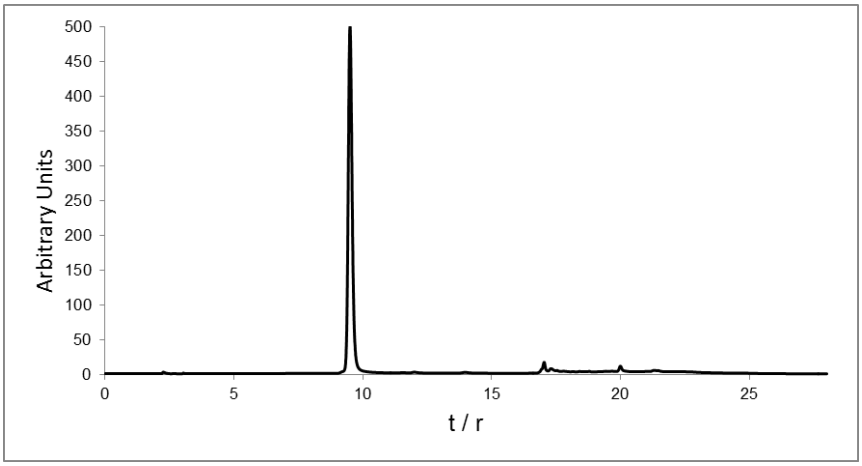
Supplementary Figure S44. Analytical RP-HPLC profile of **oligomer 3**. Elution conditions are the same as for oligomer **2**.



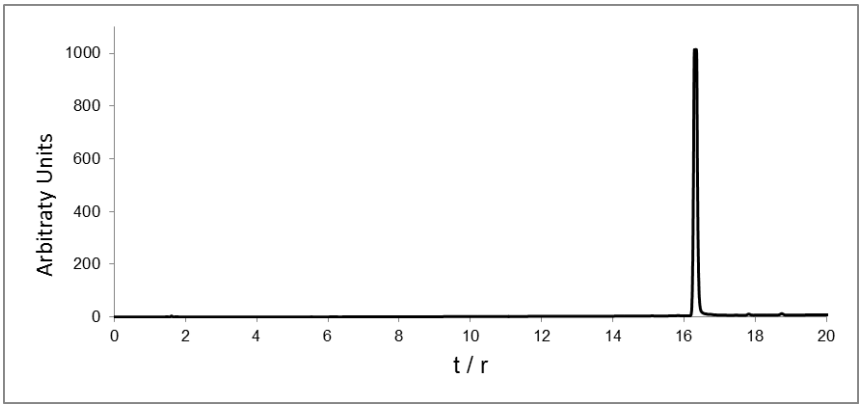
Supplementary Figure S45. Analytical RP-HPLC profile of **oligomer 4**. Elution conditions are the same as for oligomer **2**.



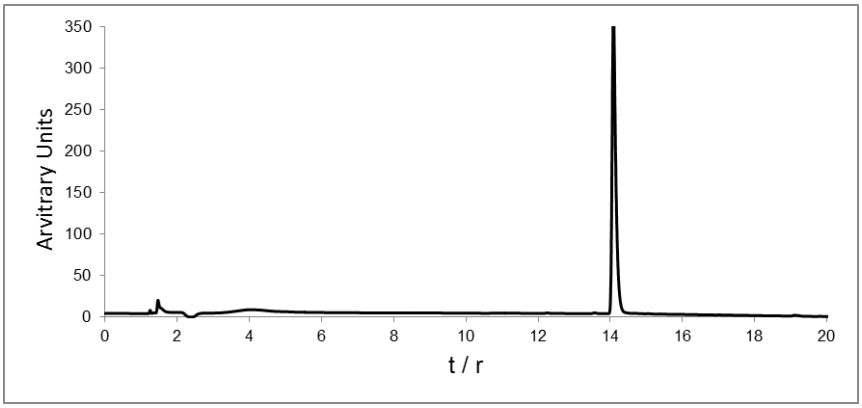
Supplementary Figure S46. Analytical RP-HPLC profile of **oligomer 5**. Elution conditions are the same as for oligomer **2**.



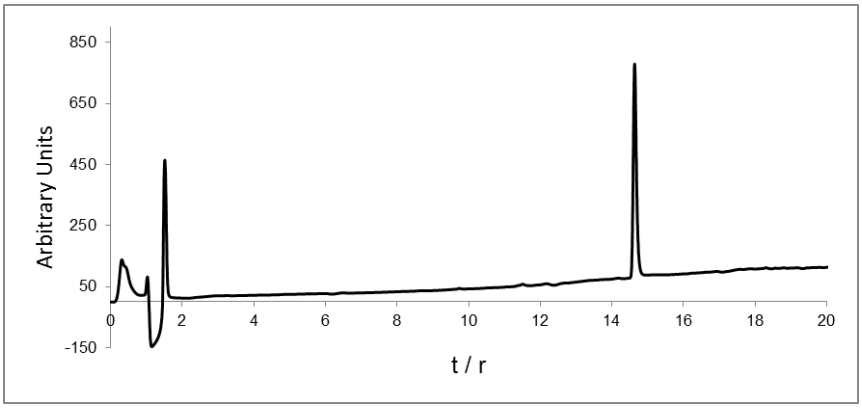
Supplementary Figure S47. Analytical RP-HPLC profile of **oligomer 6**. Elution conditions are the same as for oligomer **2**.



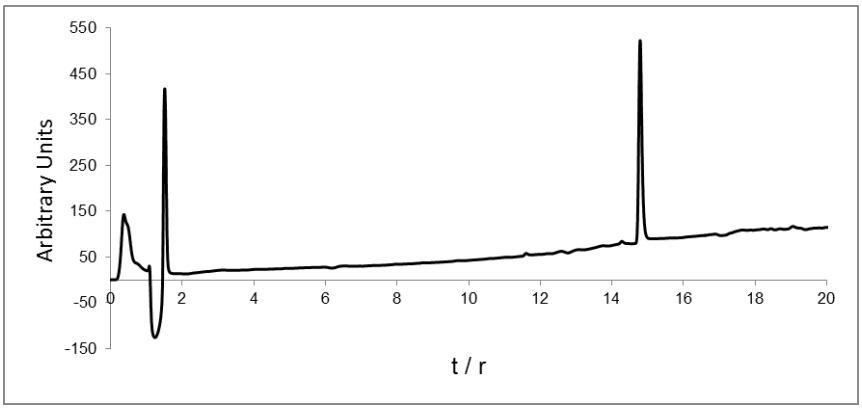
Supplementary Figure S48. Analytical RP-HPLC profile of **Fmoc**(**Q^{Pho}**). Elution conditions: Nucleodur C8 column (120 x 4.6 mm, 5 µm); eluents C and D, where solvent C was H_2O + 0.1% TFA and solvent D was ACN + 0.1% TFA; gradient program: from 30% to 100% of solvent B in 20 min; UV detection at λ = 300 nm.



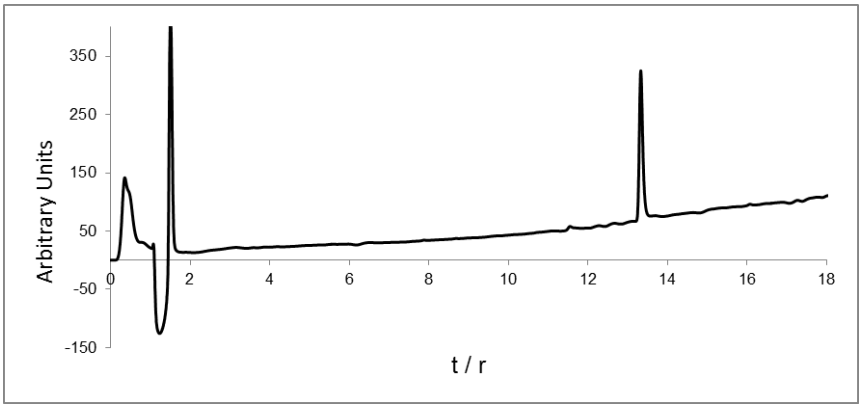
Supplementary Figure S49. Analytical RP-HPLC profile of **Fmoc**(**^{m}Q^{Pho}**). Elution conditions: Nucleodur C8 column (120 x 4.6 mm, 5 µm); eluents C and D, where solvent C was H_2O + 0.1% TFA and solvent D was ACN + 0.1% TFA; gradient program: from 5% to 100% of solvent B in 20 min; UV detection at λ = 300 nm.



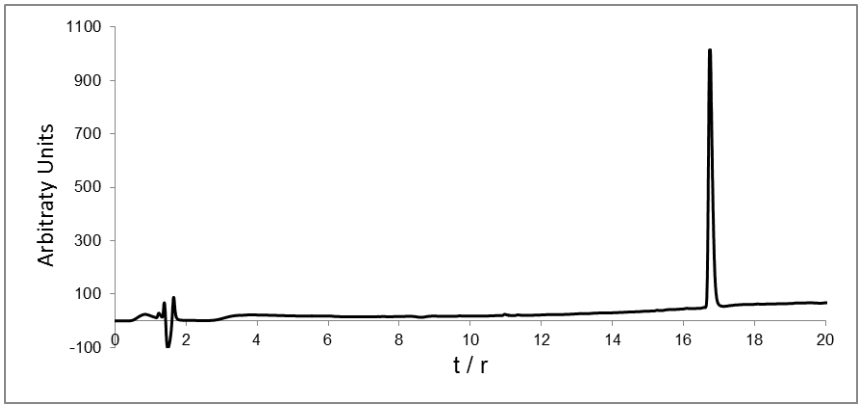
Supplementary Figure S50. Analytical RP-HPLC profile of **Fmoc**(**Q^{OAc}**). Elution conditions are the same as for **Fmoc**(**Q^{Pho}**).



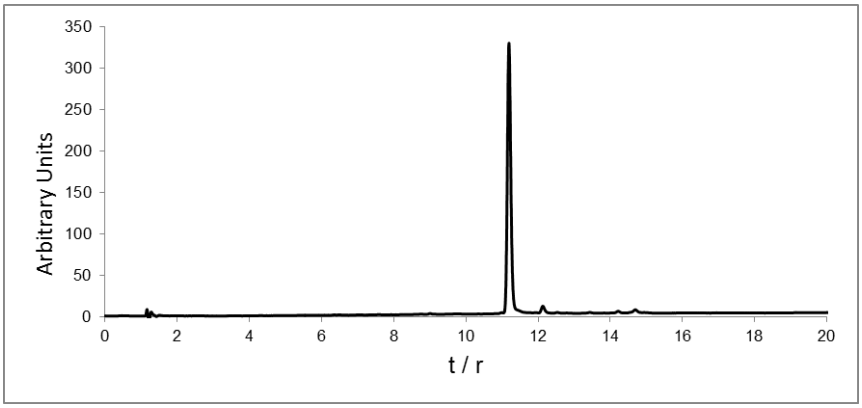
Supplementary Figure S51. Analytical RP-HPLC profile of **Fmoc**(**Q^{5Prop}**). Elution conditions are the same as for **Fmoc**(**Q^{Pho}**).



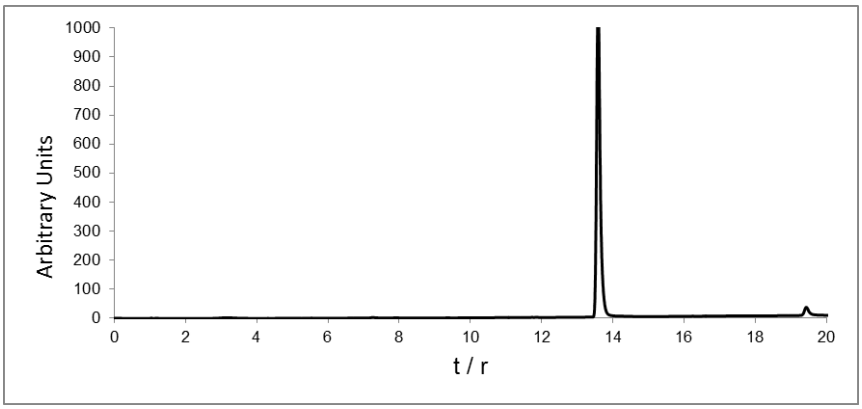
Supplementary Figure S52. Analytical RP-HPLC profile of **Fmoc**(**^{m}Q^{OAc}**). Elution conditions are the same as for **Fmoc**(**Q^{Pho}**).



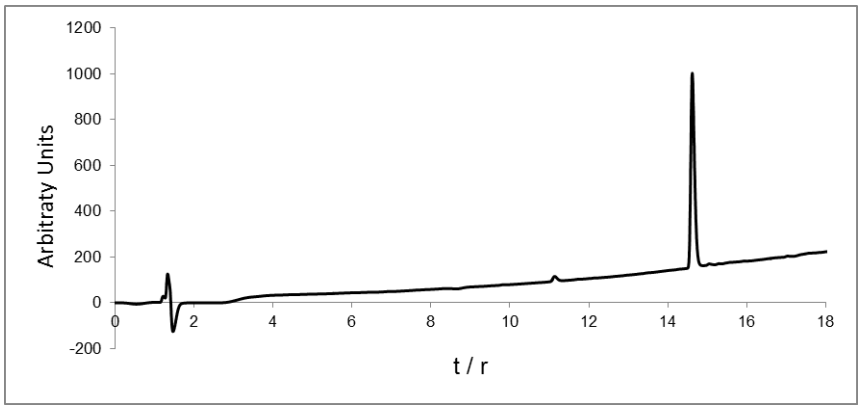
Supplementary Figure S53. Analytical RP-HPLC profile of **Fmoc**(**Q^{5But}**). Elution conditions are the same as for **Fmoc**(**Q^{Pho}**).



Supplementary Figure S54. Analytical RP-HPLC profile of **Fmoc**(**$^{m}Q^{But}$**). Elution conditions are the same as for **Fmoc**(**Q^{Pho}**).

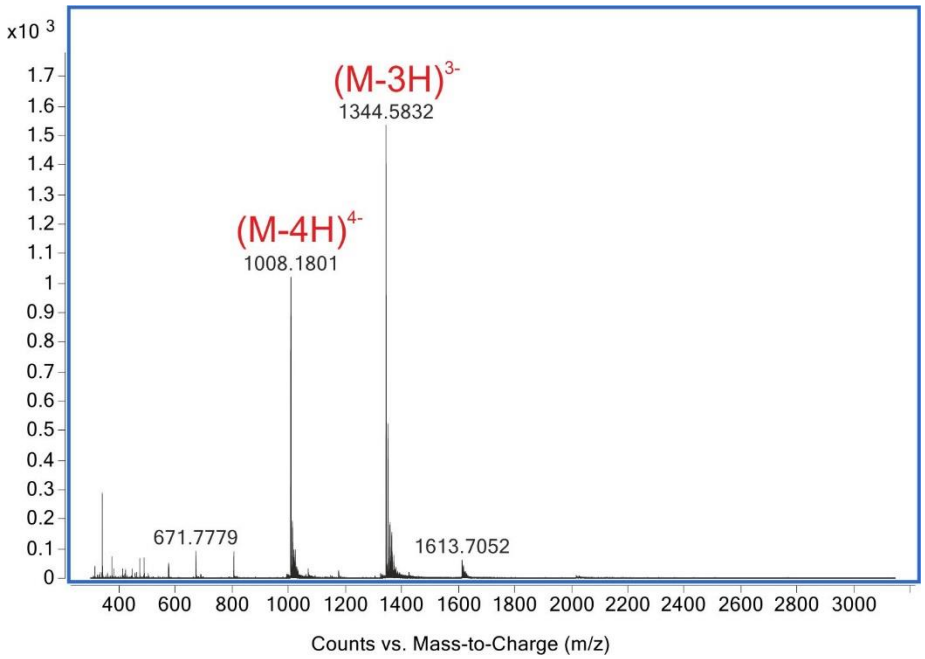


Supplementary Figure S55. Analytical RP-HPLC profile of **Fmoc**(**Q^{5Ac}**). Elution conditions: Nucleodur C8 column (120 x 4.6 mm, 5 µm); eluents C and D, where solvent C was H2O + 0.1% TFA and solvent D was ACN + 0.1% TFA; gradient program: from 30% to 55% of solvent B in 20 min; UV detection at λ = 300 nm.

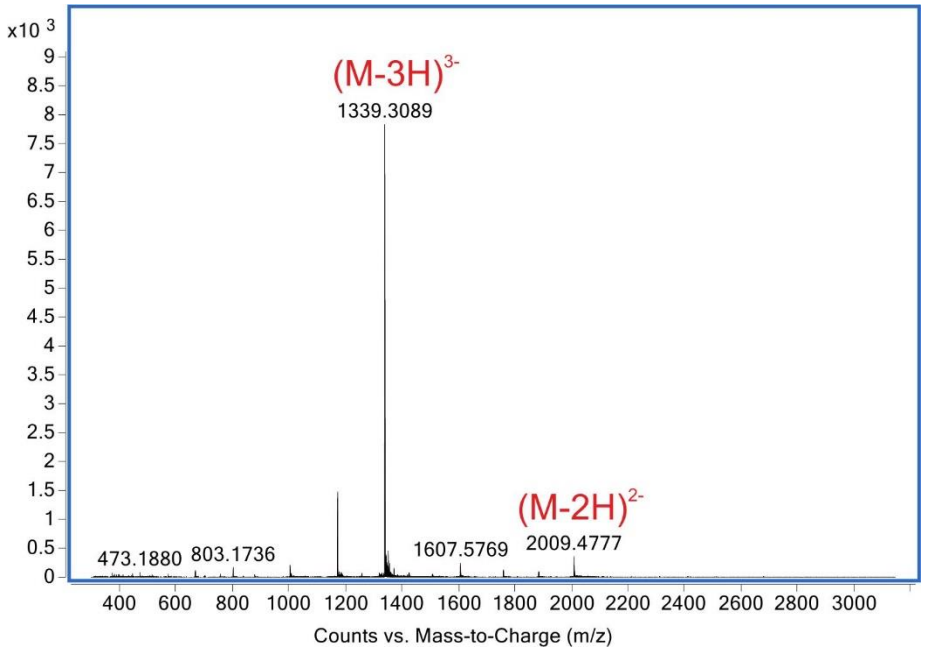


Supplementary Figure S56. Analytical RP-HPLC profile of **Fmoc**(**$^{m}Q^{Ac}$**). Elution conditions are the same as for **Fmoc**(**Q^{5Ac}**).

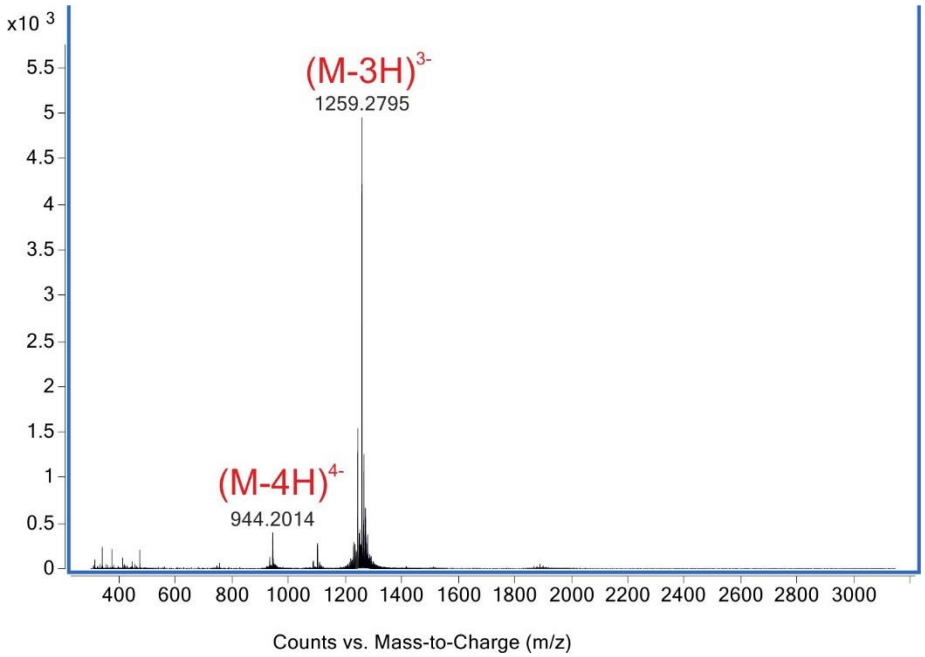
7. Mass spectrometry analysis of new synthetic compounds



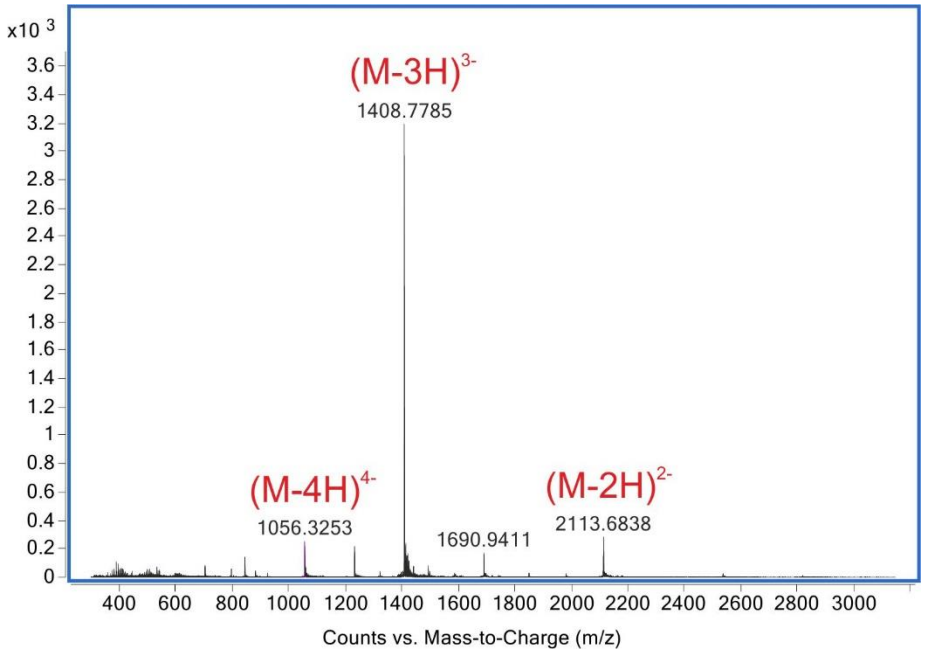
Supplementary Figure S57. Multicharged species observed by HRMS ESI MS (anionic mode) of **oligomer 2**.



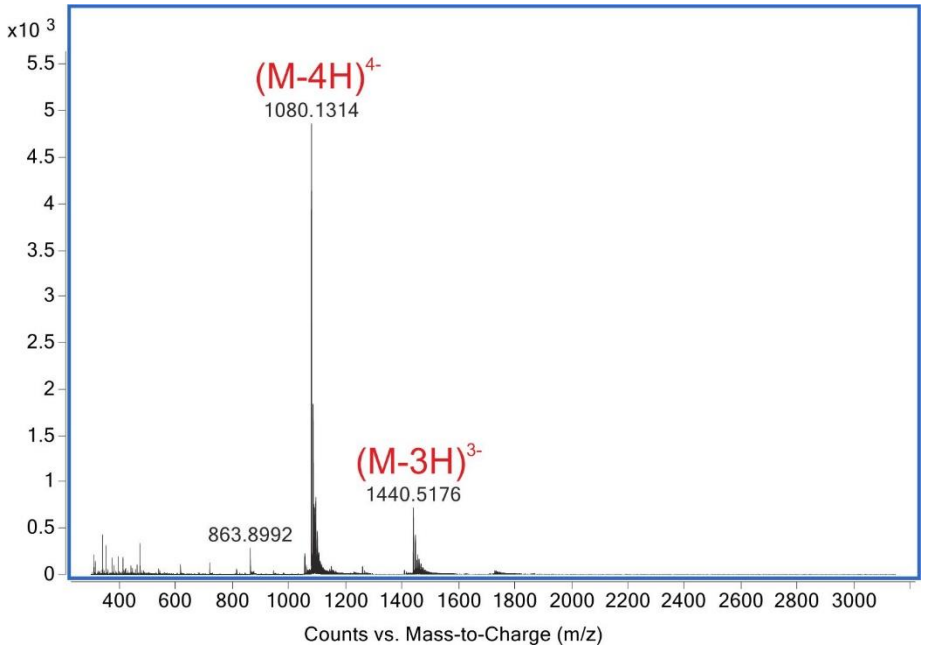
Supplementary Figure S58. Multicharged species observed by HRMS ESI MS (anionic mode) of **oligomer 3**.



Supplementary Figure S59. Multicharged species observed by HRMS ESI MS (anionic mode) of **oligomer 4**.



Supplementary Figure S60. Multicharged species observed by HRMS ESI MS (anionic mode) of **oligomer 5**.



Supplementary Figure S61. Multicharged species observed by HRMS ESI MS (anionic mode) of **oligomer 6**.

8. References

1. Baptiste, B., Douat-Casassus, C., Laxmi-Reddy, K., Godde, F. and Huc, I. (2010) Solid Phase Synthesis of Aromatic Oligoamides: Application to Helical Water-Soluble Foldamers. *J. Org. Chem.*, **75**, 7175-7185.
2. Hu, X., Dawson, S.J., Mandal, P.K., de Hatten, X., Baptiste, B. and Huc, I. (2017) Optimizing side chains for crystal growth from water: a case study of aromatic amide foldamers. *Chem. Sci.*, **8**, 3741-3749.
3. Qi, T., Deschrijver, T. and Huc, I. (2013) Large-scale and chromatography-free synthesis of an octameric quinoline-based aromatic amide helical foldamer. *Nat. Protoc.*, **8**, 693-708.
4. Phillion, D.P. (1988) Patent US4740608 A.
5. Dawson, S.J., Hu, X., Claerhout, S. and Huc, I. (2016) Chapter Thirteen - Solid Phase Synthesis of Helically Folded Aromatic Oligoamides. *Methods Enzymol.*, **580**, 279-301.
6. DeLano, W.L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, USA.
7. Jiang, H., Léger, J.-M. and Huc, I. (2003) Aromatic δ-Peptides. *J. Am. Chem. Soc.*, **125**, 3448-3449.
8. Dolain, C., Grélard, A., Laguerre, M., Jiang, H., Maurizot, V. and Huc, I. (2005) Solution Structure of Quinoline- and Pyridine-Derived Oligoamide Foldamers. *Chem. - Eur. J.*, **11**, 6135-6144.
9. Hays, F.A., Teegarden, A., Jones, Z.J.R., Harms, M., Raup, D., Watson, J., Cavaliere, E. and Ho, P.S. (2005) How sequence defines structure: A crystallographic map of DNA structure and conformation. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 7157-7162.
10. Narayana, N. and Weiss, M.A. (2009) Crystallographic Analysis of a Sex-Specific Enhancer Element: Sequence-Dependent DNA Structure, Hydration, and Dynamics. *J. Mol. Biol.*, **385**, 469-490.
11. Mandal, P.K., Venkadesh, S. and Gautham, N. (2011) Structure of d(CGGGTACCCG)4 as a

four-way Holliday junction. *Acta Cryst. F*, **67**, 1506-1510.

12. Cnudde, S.E., Prorok, M., Dai, Q., Castellino, F.J. and Geiger, J.H. (2007) The Crystal Structures of the Calcium-Bound con-G and con-T[K7γ] Dimeric Peptides Demonstrate a Metal-Dependent Helix-Forming Motif. *J. Am. Chem. Soc.*, **129**, 1586-1593.

13. Mandal, P.K., Collie, G.W., Kauffmann, B. and Huc, I. (2014) Racemic DNA Crystallography. *Angew. Chem. Int. Ed.* , **53**, 14424-14427.