Supplementary Information

Shotgun sequencing of a 512-mer copolyester allows random access to stored information

Heejeong Jang[1], Hyunseon Chu[1,†], Hyojoo Noh[1,†], and Kyoung Taek Kim[1,*]

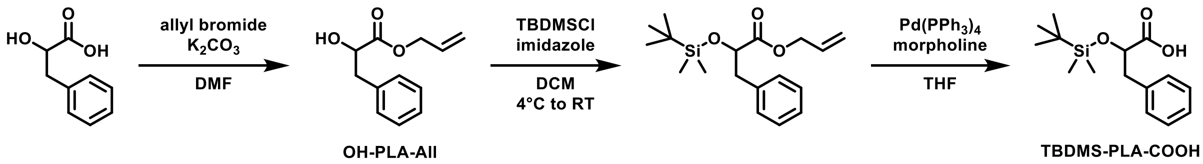
[1] Department of Chemistry, Seoul National University, Seoul 08826, Korea

[†] These authors contributed equally

*Correspondence to: ktkim72@snu.ac.kr

Table of Contents

1. Preparation of dimers precursors



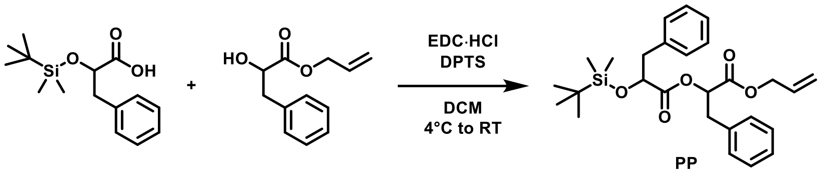
Supplementary Scheme 1. Synthesis of OH-PLA-All and TBDMS-PLA-COOH

allyl 2-hydroxy-3-phenylpropanoate (**HO-PLA-All**)

A round bottom flask was charged with *rac*-phenyllactic acid (300 mmol), allyl bromide (600 mmol, 2 eq.), and potassium carbonate (600 mmol, 2 eq.), and DMF (300 mL). The mixture was stirred at room temperature for 12 h. After DMF was removed under reduced pressure the residue was taken up in DCM and washed with H_2O and brine. The combined organic layer was dried over $MgSO_4$, and the solvent was removed. The crude mixture was purified by column chromatography. The product was obtained as colorless oil.

2-((*tert*-butyldimethylsilyl)oxy)-3-phenylpropanoic acid (**TBDMS-PLA-COOH**)

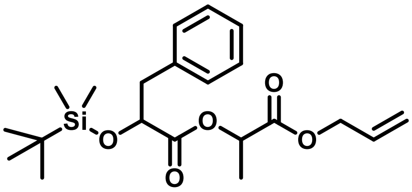
HO-PLA-All (140 mmol) was dissolved in DCM (200 mL), and *tert*-butyldimethylsilyl chloride (154 mmol, 1.1 eq.) and imidazole (168 mmol, 1.2 eq.) were added at 0℃. The reaction was stirred at room temperature for 12 h. The mixture was washed with water, 1M HCl aqueous solution, and brine. The combined organic layer was dried over $MgSO_4$, and the solvent was removed under pressure. Subsequently, the resulting organic compound and 3 mol % tetrakis(triphenylphosphine)palladium(0) were dissolved in THF. Morpholine (147 mmol, 1.05 eq.) was slowly added and the mixture was stirred at room temperature for 2 h. Upon completion of the reaction, the organic solvent was changed into DCM followed by washed with water, 1 M HCl aqueous solution, and brine. The solution was dried over $MgSO_4$, and the product with small amount of triphenylphosphine was obtained as yellow liquid.



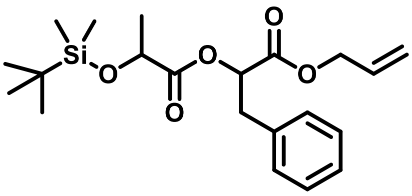
Supplementary Scheme 2. Synthesis of PP

PP

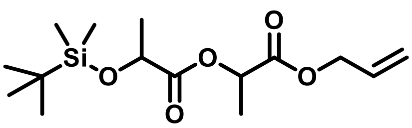
The equimolar mixture of **TBDMS-PLA-COOH** (120 mmol) and **OH-PLA-All** (120 mmol) were dissolved in DCM, and the solution was cooled to at 0°C. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (144 mmol, 1.2 eq.) and 4-(dimethylamino)pyridinium 4-toluenesulfonate (12 mmol, 0.1 eq.) were added and the reaction mixture was stirred at room temperature for 12 h. Upon completion of the reaction, the solution was washed with water and brine. The combined organic layer was dried over $MgSO_4$ followed the solvent was removed under reduced pressure. The crude mixture was purified with column chromatography and the dimer was obtained as colorless oil. 46 g, 80%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33 – 7.04 (m, 10H), 5.96 – 5.69 (m, 1H), 5.43 – 5.10 (m, 3H), 4.62 (t, J = 6.6 Hz, 2H), 4.33 (ddd, J = 34.5, 9.2, 3.3 Hz, 1H), 3.34 – 2.96 (m, 2H), 2.91 – 2.59 (m, 2H), 0.74 (d, J = 11.3 Hz, 9H), -0.08 – -0.51 (m, 6H).



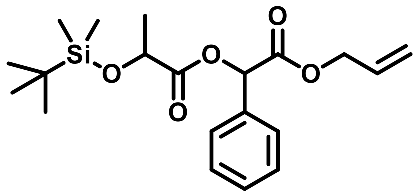
PL 101 g, 84%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 7.10 (m, 5H), 5.90 (ddt, J = 16.3, 10.8, 5.7 Hz, 1H), 5.38 – 5.22 (m, 2H), 5.15 (p, J = 7.0 Hz, 1H), 4.64 (dq, J = 5.8, 1.4 Hz, 2H), 4.44 – 4.35 (m, 1H), 3.30 – 2.76 (m, 2H), 1.57 – 1.46 (m, 3H), 0.82 – 0.75 (m, 9H), 0.05 – -0.46 (m, 6H).



LP 59 g, 81%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.50 – 6.97 (m, 5H), 5.84 (dddd, J = 20.5, 17.4, 10.1, 5.5 Hz, 1H), 5.34 – 5.19 (m, 3H), 4.61 (ddt, J = 12.9, 5.0, 1.3 Hz, 2H), 4.34 (dq, J = 18.3, 6.8 Hz, 1H), 3.40 – 2.96 (m, 2H), 1.33 (dd, J = 62.2, 6.8 Hz, 3H), 0.87 (d, J = 5.6 Hz, 9H), 0.15 – -0.13 (m, 6H).

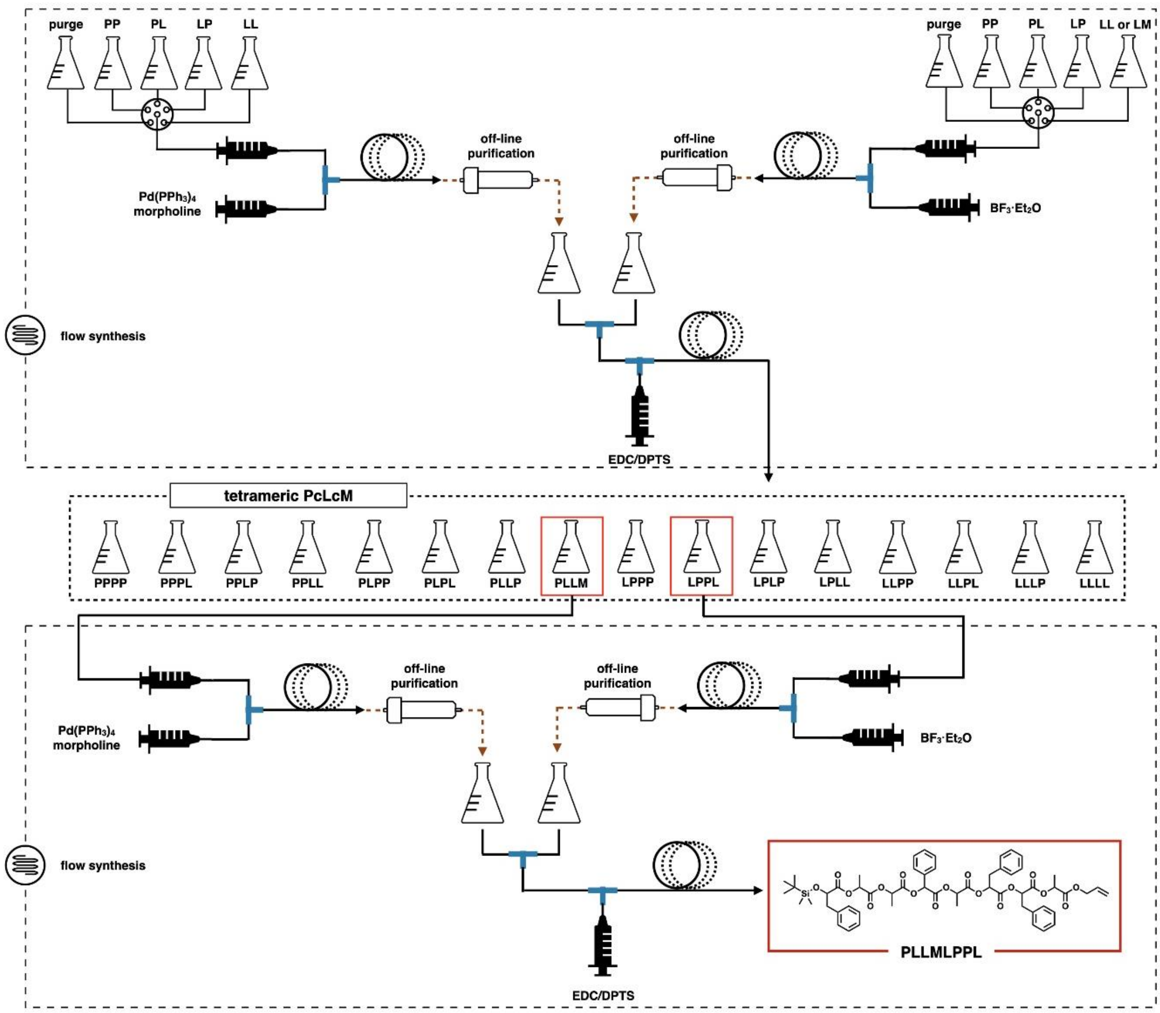


LL 33 g, 75%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.89 (ddd, J = 16.6, 8.0, 3.5 Hz, 1H), 5.40 – 5.21 (m, 2H), 5.14 (q, J = 7.2 Hz, 1H), 4.71 – 4.56 (m, 2H), 4.40 (q, J = 7.0 Hz, 1H), 1.55 – 1.42 (m, 6H), 0.90 (q, J = 1.6 Hz, 9H), 0.10 (dt, J = 9.6, 1.6 Hz, 6H).



LM 17 g, 80%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.65 – 7.29 (m, 5H), 6.04 – 5.92 (m, 1H), 5.95 – 5.65 (m, 1H), 5.44 – 4.95 (m, 3H), 4.76 – 4.41 (m, 2H), 1.66 – 1.36 (m, 3H), 1.09 – 0.84 (m, 9H), 0.30 – 0.01 (m, 6H).

2. Semi-automated flow synthesis for sequence-defined oligoesters



Supplementary Figure 1. Schematic illustrations of a semi-automated flow process for synthesis of sequence-defined tetramers and 8-mers.

Synthesis of tetramers via a semi-automated flow system

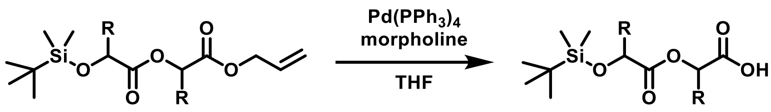
The synthesis method is the same as our previous study. The detailed procedure is as follows.

The all permutations of tetramers were generated by a semi-automated flow system, which is operated via 6-way syringe pumps connected with programmable controller. 2 mmol of a dimer (1 M in DCM) and $BF_3 \cdot Et_2O$ (7 M in DCM) were injected at 0.1 mL/min and mixed through T-mixer. The mixture was allowed to react in the reaction loop (volume of 2 mL). Simultaneously, 2 mmol of a dyad (1 M in THF) and $Pd(PPh_3)_4$ and morpholine mixture (0.03 M and 1.05 M in THF) were injected at 0.1 mL/min and mixed through T-mixer. The mixture was allowed to react in the reaction loop (volume of 2 mL). Deprotected dimers, hydroxyl and carboxylic acid, were purified by automated flash column chromatography using HEX/EA and ether/MeOH eluents, respectively. Subsequently, the purified dimers (1 M in DCM, 0.1 mL/min) were reinjected and mixed with EDC·HCl and DPTS (0.7 M and 0.07 M in DCM, 0.2 mL/min). The mixture was allowed to react in reaction loop (volume of 6 mL). After synthetic step, DCM solvent (6 mL) was purged to flow reactor for cleaning reaction loop. Synthetic and cleaning cycle were repeated 16 times for generating every permutation of tetramers. Collected tetramers were purified with automated flash column chromatography (86–90% yield)

Synthesis of 8, 16, 32-mers via a semi-automated flow system

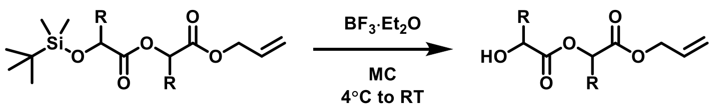
The 8, 16, 32-mers of PcLs were obtained by flow synthesis following the same conditions described for the synthesis of tetramers. The building block (0.25–0.5 M in DCM) and BF_3·Et_2O (7 M in DCM) were injected at 0.1 mL/min and mixed through T-mixer. The mixture was allowed to react in the reaction loop (volume of 2 mL). Simultaneously, the building block (0.25–0.5 M in THF) and $Pd(PPh_3)_4$ and morpholine mixture (0.03 M and 1.05 M in THF) were injected at 0.1 mL/min and mixed through T-mixer. The mixture was allowed to react in the reaction loop (volume of 2 mL). Deprotected building blocks were purified by automated flash column chromatography. The purified building blocks (0.25–0.5 M in DCM, 0.1 mL/min) were reinjected and mixed with EDC·HCl and DPTS (0.7 M and 0.07 M in DCM, 0.3 mL/min). The mixture was allowed to react in reaction loop (volume of 9 mL). Collected product was purified with automated flash column chromatography (86–91% yield).

3. General procedure for sequence-defined polyesters



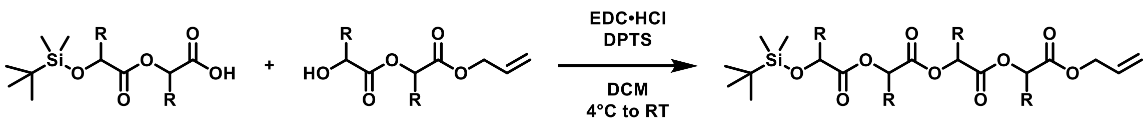
Supplementary Scheme 3. General procedure for deprotection of All group

A building block protected with TBDMS and All group was dissolved in THF (0.25 M), and $Pd(PPh_3)_4$ (3 mol%) and morpholine (1.05 eq) were added. The reaction was stirred at room temperature for 2 h. After THF was removed under reduced pressure, the residue was taken up in DCM and washed with water, 1M HCl aqueous solution, and brine. The combined organic layer was dried over $MgSO_4$, and the solvent was removed. The crude mixture was purified by column chromatography.



Supplementary Scheme 4. General procedure for deprotection of TBDMS group

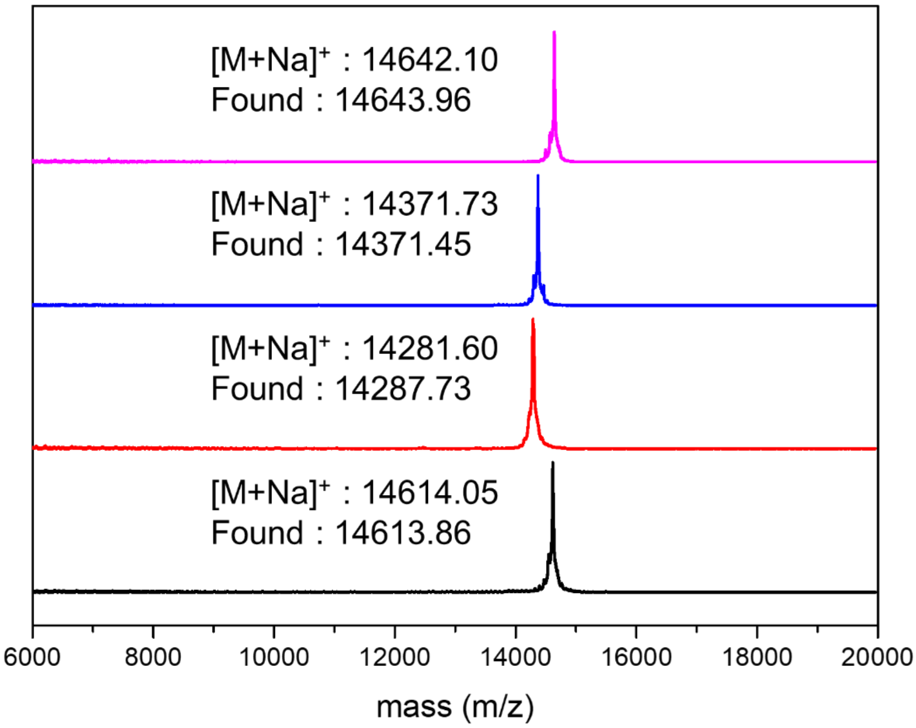
A building block protected with TBDMS and All group was dissolved in DCM (0.2 M). The solution was cooled to 0°C on an ice bath, and boron trifluoride diethyl etherate (BF3·Et2O, 4 eq) was added dropwise. The reaction mixture was stirred at room temperature for 4 h. Upon completion of the reaction, the reaction was quenched with saturated $NaHCO_3$ followed by dilution with water. The organic layer was separated and washed with brine. The combined organic layer was dried over $MgSO_4$, and the solvent was removed under reduced pressure. The crude product was purified by automated column chromatography.



Supplementary Scheme 5. General procedure for esterification reaction

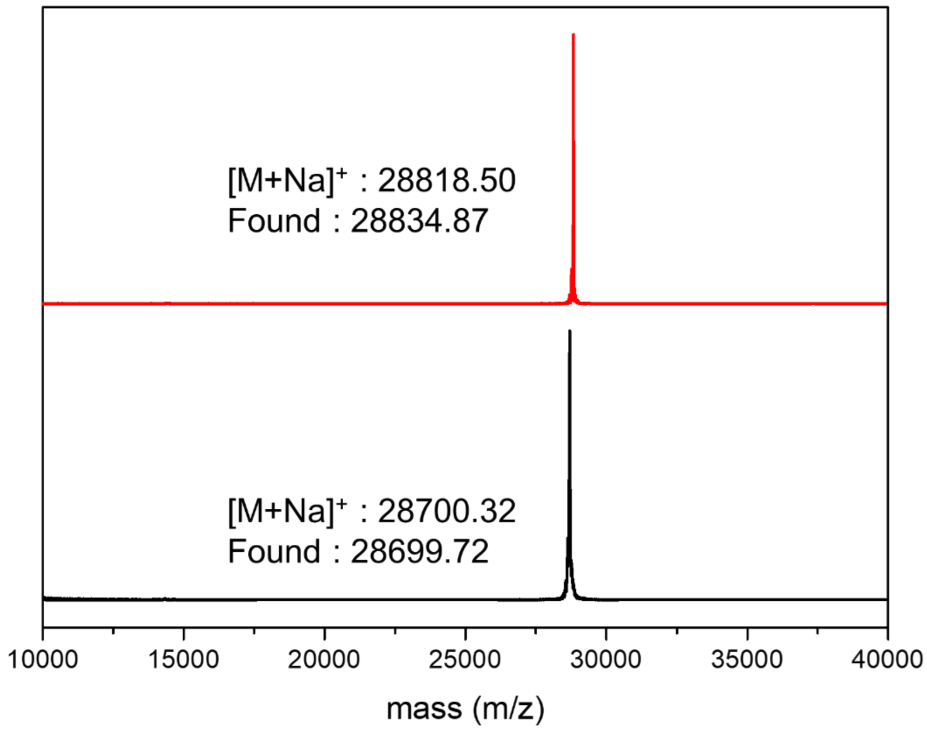
Alcohol and carboxylic acid were dissolved in DCM (0.5 M), and the mixture was cooled to 0°C on an ice bath. To the mixture, 4-(dimethylamino)pyridine (DMAP, 0.15 eq.) and 1-(3-dimethylaminopropyl)- 3-ethylcarbodiimide hydrochloride (EDC·HCl, 1.2 eq.) were added. The reaction mixture was stirred overnight at room temperature. Upon completion of the reaction, the reaction mixture was washed with water and brine. The combined organic layer was dried over $MgSO_4$, and the solvent was removed under reduced pressure. Products with 64 repeating units or more were purified by preparative-size exclusion chromatography (prep-SEC) with a series of columns using chloroform as an eluent.

4. MALDI-TOF spectra of 128, 256, and 512-mer



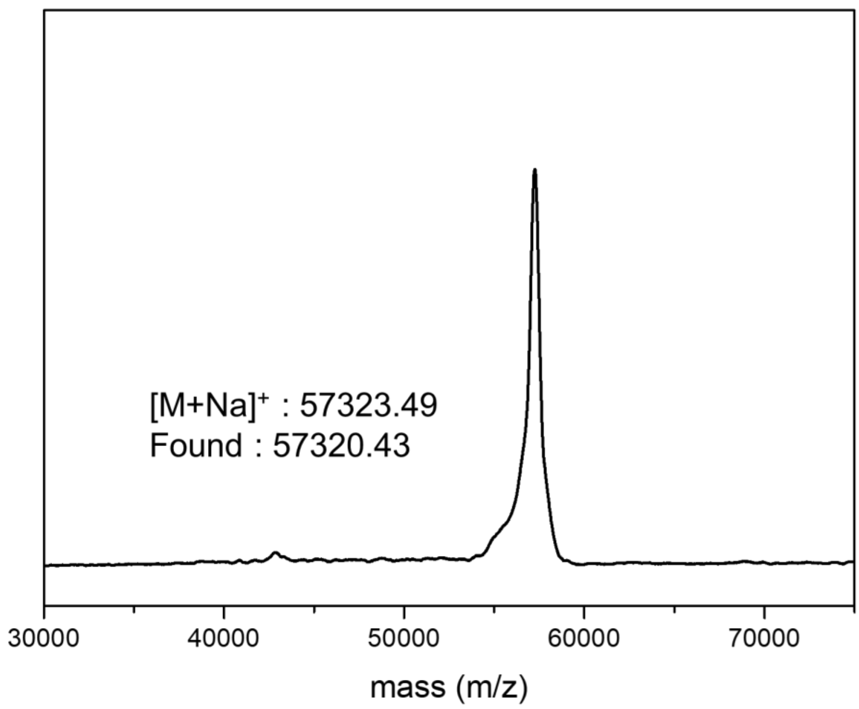
Supplementary Figure 2. MALDI-TOF spectra of 128-mer

128-mer_repeat unit 1-128 (black, calculated mass: 14614.05 Da; found mass: 14613.86 Da), 128-mer_repeat unit 129-256 (red, calculated mass: 14281.60 Da; found mass: 14287.73 Da), 128-mer_repeat unit 257-384 (blue, calculated mass: 14371.73 Da; found mass: 14371.45 Da), and 128-mer_repeat unit 385-512 (pink, calculated mass: 14642.10 Da; found mass: 14643.96 Da).



Supplementary Figure 3. MALDI-TOF spectra of 256-mer

256-mer_repeat unit 1-256 (black, calculated mass: 28700.32 Da; found mass: 28699.72 Da), and 256-mer_repeat unit 257-512 (red, calculated mass: 28818.50 Da; found mass: 28834.87 Da).



Supplementary Figure 4. MALDI-TOF spectrum of 512-mer (calculated mass: 57323.49 Da; found mass: 57320.43 Da).

5. Computer code for the identification, decoding, and reconstruction of mass data (CIDER)

Step 1. identification of fragments by counting the number of repeat units and end groups from MALDI-TOF spectra

After hydrogenolysis of PcL, fragments have one of three types of end groups: (i) fragments having a benzyl group at the α-terminus and a carboxylic acid at the ω-terminus, (ii) fragments having TBDMS group at the α-terminus and a carboxylic acid at the ω-terminus, (iii) fragments having a benzyl group at the α-terminus and propyl ester at the ω-terminus. The number of repeat units and end groups were analyzed by using the following equation:

$$M_{[M+Na]^+} = aM_{TBDMS\text{-}O} + bM_{Benzyl} + cM_P + dM_L + eM_M + fM_H + gM_{propyl} + M_{Na} \qquad \text{(Eq. 1)}$$

where M_x is mass of x, $a + b = 1$, $f + g = 1$, and a, b, c, d, e, f, and g are integer numbers equal or greater than 0. The CIDER just needs $M_{[M+Na]^+}$ input values for identification of fragments and outputs c, d, e values and end group type. c, d, and e represent the number of phenyllactic acid, lactic acid, and mandelic acid of fragment, respectively. If $a = 1$, and $f = 1$, the fragment has TBDMS group at the α-terminus and a carboxylic acid at the ω-terminus. On the other hand, $b = 1$, and $g = 1$ indicate that the fragment has a benzyl group at the α-terminus and propyl ester at the ω-terminus. Except for these two cases, fragments have a benzyl group at the α-terminus and a carboxylic acid at the ω-terminus, which is $b = 1$, and $f = 1$.

```
def CIDER1(x):
  # x = M[M +Na]+
  closest = float('inf')
  endgroup = ''
  values = (0, 0, 0)

  for c in range(513):
    for d in range(513):
      for e in range(18):
        #value1: end group (i), value2: end group (ii), value3: end group (iii)
        value1 = 135.0446 + c * 148.0524 + d * 72.0211 + e * 134.0367 + 1.0078 + 22.9897
        value2 = 131.0891 + c * 148.0524 + d * 72.0211 + e * 134.0367 + 1.0078 + 22.9897
        value3 = 135.0446 + c * 148.0524 + d * 72.0211 + e * 134.0367 + 43.0547 + 22.9897

        diff1 = abs(value1 - x)
        diff2 = abs(value2 - x)
        diff3 = abs(value3 - x)
        # PLLM: fragmentation code
        if d >= 2 * e and c >= e and (diff1 < closest or diff2 < closest or diff3 < closest):
          closest = min(diff1, diff2, diff3)
          if diff1 == closest:
            endgroup = 'benzyl+COOH'
          elif diff2 == closest:
            endgroup = 'Si+COOH'
          else:
            endgroup = 'benzyl+propyl'
          values = (c, d, e)

  return endgroup, values
```

Supplementary Figure 5. CIDER for identification of fragments

Step 2. sequencing of individual fragments by reading the mass difference in MS^2 spectra

The fragment sequence was decoded by MALDI-TOF/TOF MS^2 which breaks C(a)-O bond of PcL backbone. Fragmentation produces two types fragments: (i) ai fragments that have the original α-terminus and a new carboxylic acid terminus, (ii) yi fragments that contain the original ω-terminus and a new alkene terminus. The mass difference between adjacent peaks was converted to the sequence, P (148 Da) or L (72 Da).

After measuring MS^2 of the fragment, we input $M_{[M+Na]^+}$ and a text table listing all mass and intensity values of MS^2 spectrum into CIDER. To distinguish between signal and noise, CIDER selects only values of a text table above a threshold intensity. Sequencing proceeds in two directions: α-terminus to ω-terminus and ω-terminus to α-terminus. When sequencing in the α to ω direction, the α-terminus end group is subtracted from the input mass value, and in the opposite direction, the ω-terminus end group is subtracted. After that, if a mass value-148 exists, it is a P sequence, and if a mass value-72 exists, it is an L sequence. Since both the mass value-148 and the mass value-72 can exist, a function that subtracts 72 first must also be performed. Repeat this to finally output the sequence when the mass value becomes 0 or less. By overlapping the sequences obtained in two directions, the fragment sequence was decoded.

a

```
# 1: α-terminus to ω-terminus direction sequencing
def CIDER2_1(x, data):
    # x = M[M+Na]+, select rows where the intensity is greater than 0.1
    data[data.iloc[:, 1] > 0.1]
    sequence = []
    current_mass = x - 135.04

    while current_mass > 0:
        diff_minus_72 = abs(current_mass - data.iloc[:, 0] - 72.02)
        diff_minus_148 = abs(current_mass - data.iloc[:, 0] - 148.05)

        if diff_minus_72.min() < diff_minus_148.min():
            sequence.append('L')
            current_mass -= 72.02
        else:
            sequence.append('P')
            current_mass -= 148.05

    return sequence
```

b

```
# 2: ω-terminus to α-terminus direction sequencing
def CIDER2_2(x, data):
    # x = M[M+Na]+, select rows where the intensity is greater than 0.1
    data[data.iloc[:, 1] > 0.1]
    sequence = []
    current_mass = x - 1.01

    while current_mass > 0:
        diff_minus_72 = abs(current_mass - data.iloc[:, 0] - 72.02)
        diff_minus_148 = abs(current_mass - data.iloc[:, 0] - 148.05)

        if diff_minus_72.min() < diff_minus_148.min():
            sequence.append('L')
            current_mass -= 72.02
        else:
            sequence.append('P')
            current_mass -= 148.05

    return sequence
```

Supplementary Figure 6. CIDER for sequencing fragments having a benzyl group at the α-terminus and a carboxylic acid at the ω-terminus. a α-terminus to ω-terminus direction sequencing **b** ω-terminus to α-terminus direction sequencing

Step 3. reconstruction of a full-chain sequence by assembling fragment sequences using CRC

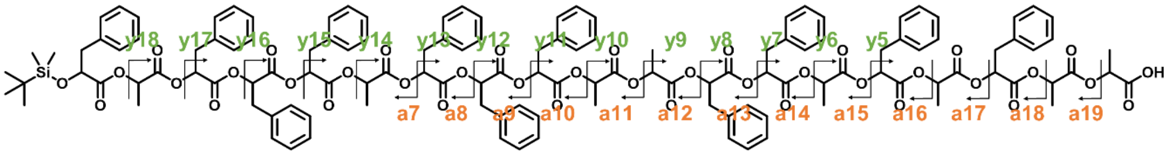
The number of P, L, and M of 512-mer PcL could be known by using the following equation:

$$M_{[M+Na]^+} = M_{TBDMS} + aM_P + bM_L + cM_M + M_{All} + M_{Na} \qquad \text{(Eq. 2)}$$

where M_x is mass of x, a, b, and c are integer numbers equal or greater than 0. If we enter $M_{[M+Na]^+}$, CIDER outputs the values of a, b, and c. a, b, and c represent the number of phenyllactic acid, lactic acid, and mandelic acid of 512-mer PcL, respectively. The number of mandelic acid (c) indicates that c+1 fragments would appear on the activation of the fragmentation code embedded in the 512-mer PcL. Find duplicate fragments so that the number of fragments known until **Step 2** matches c+1.

The CRC-16 (Modbus) is located in the last 16 bits of the last fragment which has propyl ester at the ω-terminus. The reconstruction of the fragment sequence is repeated until the calculated CRC matches the original CRC. The CRC-16 (Modbus) calculator uses already known code.

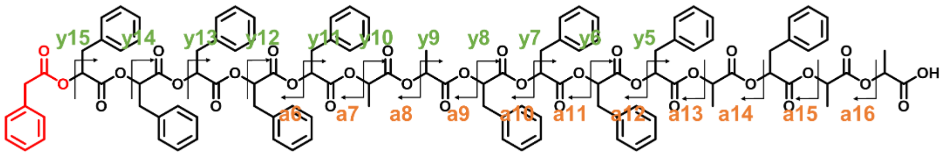
6. Sequencing of oPcLs by MALDI-TOF/TOF MS^2



Supplementary Figure 7. Fragment pattern of F1 of 512-mer

COOH to Si	Calc. m/z	Found m/z	Diff.	sequence	Si to COOH	Calc. m/z	found m/z	Diff.	sequence
[M+Na]+	2359.83	2361.259			[M+Na]+	2359.83	2361.259		
y18	2289.239	2289.155	72.104	L	a19	2081.109	2087.48	273.779	P-Si
y17	2217.219	2217.172	71.983	L	a18	2009.089	2015.639	71.841	L
y16	2069.169	2068.11	149.062	P	a17	1861.039	1863.378	152.261	P
y15	1997.149	1996.185	71.925	L	a16	1712.989	1711.183	152.195	P
y14	1849.099	1847.63	148.555	P	a15	1564.939	1563.433	147.75	P
y13	1777.079	1775.676	71.954	L	a14	1492.919	1491.365	72.068	L
y12	1629.029	1627.693	147.983	P	a13	1344.869	1344.348	147.017	P
y11	1480.979	1479.816	147.877	P	a12	1196.819	1199.261	145.087	P
y10	1408.959	1407.66	72.156	L	a11	1048.769	1051.321	147.94	P
y9	1336.939	1335.741	71.919	L	a10	976.749	975.159	76.162	L
y8	1188.889	1187.449	148.292	P	a9	904.729	903.396	71.763	L
y7	1040.839	1039.489	147.96	P	a8	756.679	755.306	148.09	P
y6	892.789	891.526	147.963	P	a7	608.629	611.481	143.825	P
y5	820.769	819.347	72.179	L	a6	536.609	535.17	76.311	L

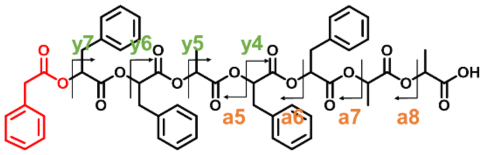
Supplementary Table 1. Decoding table of F1 of 512-mer



Supplementary Figure 8. Fragment pattern of F2 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	1999.67	1999.338			$[M+Na]^+$	1999.67	1999.338		
y15	1864.298	1864.797	134.541	M	a16	1926.308	1928.007	71.331	L
y14	1716.248	1717.812	146.985	P	a15	1854.288	1855.837	72.17	L
y13	1568.198	1569.48	148.332	P	a14	1706.238	1707.693	148.144	P
y12	1420.148	1420.574	148.906	P	a13	1634.218	1635.954	71.739	L
y11	1272.098	1271.44	149.134	P	a12	1486.168	1487.636	148.318	P
y10	1124.048	1123.346	148.094	P	a11	1338.118	1339.737	147.899	P
y9	1052.028	1051.287	72.059	L	a10	1190.068	1191.545	148.192	P
y8	980.008	979.491	71.796	L	a9	1042.018	1043.481	148.064	P
y7	831.958	831.387	148.104	P	a8	969.998	971.513	71.968	L
y6	683.908	683.392	147.995	P	a7	897.978	899.457	72.056	L
y5	535.858	535.309	148.083	P	a6	749.928	751.312	148.145	P

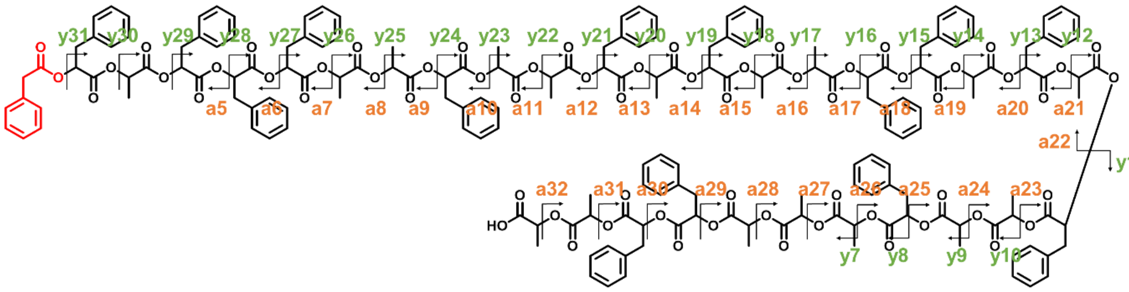
Supplementary Table 2. Decoding table of F2 of 512-mer



Supplementary Figure 9. Fragment pattern of F3, F8 and F13 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	967.32	967.35			$[M+Na]^+$	967.32	967.35		
y7	831.3	830.098	137.252	M	a8	895.33	895.601	71.749	L
y6	683.25	683.214	146.884	P	a7	823.31	823.34	72.261	L
y5	535.2	535.225	147.989	P	a6	675.26	675.38	147.96	P
y4	463.18	463.197	72.028	L	a5	527.21	527.264	148.116	P

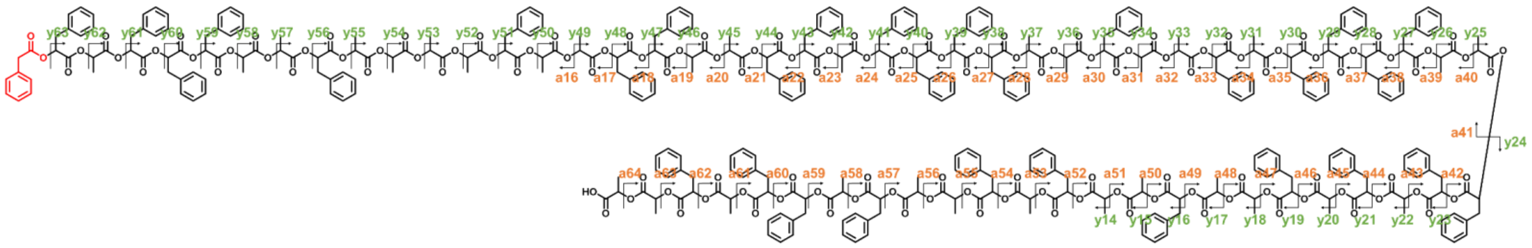
Supplementary Table 3. Decoding table of F3, F8, and F13 of 512-mer



Supplementary Figure 10. Fragment pattern of F4 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	3456.14	3457.794			$[M+Na]^+$	3456.14	3457.794		
y31	3321.744	3322.714	135.08	M	a32	3385.774	3386.187	71.607	L
y30	3173.694	3173.531	149.183	P	a31	3313.754	3315.681	71.506	L
y29	3101.674	3102.194	71.337	L	a30	3165.704	3166.657	149.024	P
y28	2953.624	2953.102	149.092	P	a29	3017.654	3018.794	147.863	P
y27	2805.574	2804.267	148.835	P	a28	2945.634	2946.277	72.517	L
y26	2657.524	2655.884	148.383	P	a27	2873.614	2873.783	72.494	L
y25	2585.504	2586.176	69.708	L	a26	2801.594	2800.627	73.156	L
y24	2513.484	2514.656	71.52	L	a25	2653.544	2653.592	147.035	P
y23	2365.434	2367.586	147.07	P	a24	2581.524	2582.386	71.206	L
y22	2293.414	2294.757	72.829	L	a23	2509.504	2509.329	73.057	L
y21	2221.394	2223.405	71.352	L	a22	2361.454	2362.293	147.036	P
y20	2073.344	2075.718	147.687	P	a21	2289.434	2289.557	72.736	L
y19	2001.324	2005.569	70.149	L	a20	2141.384	2142.176	147.381	P
y18	1853.274	1856.441	149.128	P	a19	2069.364	2069.944	72.232	L
y17	1781.254	1785.173	71.268	L	a18	1921.314	1922.197	147.747	P
y16	1709.234	1712.334	72.839	L	a17	1773.264	1774.186	148.011	P
y15	1561.184	1564.489	147.845	P	a16	1701.244	1700.853	73.333	L
y14	1413.134	1416.166	148.323	P	a15	1629.224	1630.143	70.71	L
y13	1341.114	1345.218	70.948	L	a14	1481.174	1481.73	148.413	P
y12	1193.064	1196.993	148.225	P	a13	1409.154	1409.691	72.039	L
y11	1121.044	1125.089	71.904	L	a12	1261.104	1261.602	148.089	P
y10	972.994	976.857	148.232	P	a11	1189.084	1188.689	72.913	L
y9	900.974	904.798	72.059	L	a10	1117.064	1117.452	71.237	L
y8	828.954	832.913	71.885	L	a9	969.014	968.576	148.876	P
y7	680.904	684.813	148.1	P	a8	896.994	896.659	71.917	L

Supplementary Table 4. Decoding table of F4 of 512-mer

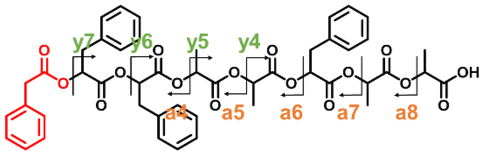


Supplementary Figure 11. Fragment pattern of F5 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	7053.34	7057.36			$[M+Na]^+$	7053.34	7059.36		
y63	6917.29	6921.722	135.638	M	a64	6981.32	6985.405	73.955	L
y62	6769.24	6774.118	147.604	P	a63	6909.3	6913.836	71.569	L
y61	6697.22	6700.263	73.855	L	a62	6761.25	6764.429	149.407	P
y60	6549.17	6553.148	147.115	P	a61	6689.23	6693.271	71.158	L
y59	6401.12	6406.002	147.146	P	a60	6541.18	6544.529	148.742	P
y58	6329.1	6335.002	71	L	a59	6393.13	6397.442	147.087	P
y57	6257.08	6261.156	73.846	L	a58	6321.11	6325.979	71.463	L
y56	6109.03	6113.164	147.992	P	a57	6173.06	6178.588	147.391	P
y55	6037.01	6041.774	71.39	L	a56	6101.04	6106.085	72.503	L
y54	5964.99	5970.526	71.248	L	a55	6029.02	6033.251	72.834	L
y53	5892.97	5898.402	72.124	L	a54	5880.97	5885.832	147.419	P
y52	5820.95	5824.952	73.45	L	a53	5808.95	5815.324	70.508	L
y51	5672.9	5675.342	149.61	P	a52	5660.9	5668.034	147.29	P
y50	5600.88	5604.727	70.615	L	a51	5588.88	5595.735	72.299	L
y49	5528.86	5532.771	71.956	L	a50	5516.86	5522.498	73.237	L
y48	5380.81	5384.134	148.637	P	a49	5368.81	5374.534	147.964	P
y47	5232.76	5236.301	147.833	P	a48	5296.79	5302.387	72.147	L
y46	5160.74	5164.193	72.108	L	a47	5224.77	5228.867	73.52	L
y45	5088.72	5092.819	71.374	L	a46	5076.72	5081.515	147.352	P
y44	4940.67	4944.706	148.113	P	a45	5004.7	5011.029	70.486	L
y43	4792.62	4795.014	149.692	P	a44	4856.65	4862.888	148.141	P
y42	4720.6	4723.402	71.612	L	a43	4784.63	4790.99	71.898	L
y41	4572.55	4575.251	148.151	P	a42	4636.58	4643.598	147.392	P
y40	4424.5	4427.179	148.072	P	a41	4488.53	4496.166	147.432	P
y39	4276.45	4279.707	147.472	P	a40	4416.51	4423.081	73.085	L
y38	4128.4	4131.013	148.694	P	a39	4344.49	4352.045	71.036	L
y37	4056.38	4059.121	71.892	L	a38	4196.44	4203.603	148.442	P

y36	3984.36	3987.998	71.123	L	a37	4048.39	4055.127	148.476	P
y35	3836.31	3839.759	148.239	P	a36	3900.34	3907.127	148	P
y34	3764.29	3767.747	72.012	L	a35	3752.29	3759.078	148.049	P
y33	3692.27	3695.251	72.496	L	a34	3680.27	3686.686	72.392	L
y32	3544.22	3547.766	147.485	P	a33	3532.22	3539.342	147.344	P
y31	3472.2	3475.92	71.846	L	a32	3460.2	3465.999	73.343	L
y30	3324.15	3327.906	148.014	P	a31	3388.18	3393.63	72.369	L
y29	3176.1	3179.239	148.667	P	a30	3240.13	3247.963	145.667	P
y28	3028.05	3031.259	147.98	P	a29	3168.11	3174.741	73.222	L
y27	2880	2883.209	148.05	P	a28	3096.09	3103.303	71.438	L
y26	2807.98	2810.466	72.743	L	a27	2948.04	2955.521	147.782	P
y25	2735.96	2738.805	71.661	L	a26	2799.99	2807.673	147.848	P
y24	2587.91	2590.326	148.479	P	a25	2651.94	2659.584	148.089	P
y23	2439.86	2442.437	147.889	P	a24	2503.89	2511.077	148.507	P
y22	2367.84	2370.727	71.71	L	a23	2431.87	2439.112	71.965	L
y21	2219.79	2222.916	147.811	P	a22	2283.82	2290.745	148.367	P
y20	2147.77	2150.356	72.56	L	a21	2135.77	2142.763	147.982	P
y19	1999.72	2002.98	147.376	P	a20	2063.75	2070.31	72.453	L
y18	1927.7	1930.928	72.052	L	a19	1991.73	1998.051	72.259	L
y17	1855.68	1858.691	72.237	L	a18	1843.68	1850.385	147.666	P
y16	1707.63	1710.415	148.276	P	a17	1695.63	1701.575	148.81	P
y15	1635.61	1638.399	72.016	L	a16	1623.61	1629.741	71.834	L
y14	1563.59	1566.653	71.746	L	a15	1551.59	1558.243	71.498	L
y13	1415.54	1418.838	147.815	P	a14	1403.54	1409.616	148.627	P

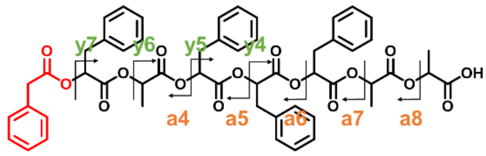
Supplementary Table 5. Decoding table of F5 of 512-mer



Supplementary Figure 12. Fragment pattern of F6 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	891.28	891.454			$[M+Na]^+$	891.28	891.454		
y7	756.414	754.263	137.191	M	a8	819.434	819.528	71.926	L
y6	608.364	607.181	147.082	P	a7	747.414	747.375	72.153	L
y5	460.314	459.233	147.948	P	a6	599.364	599.316	148.059	P
y4	388.294	387.163	72.07	L	a5	527.344	527.327	71.989	L
y3	316.274	315.184	71.979	L	a4	455.324	455.327	72	L

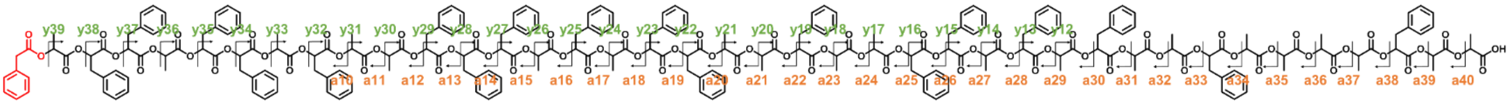
Supplementary Table 6. Decoding table of F6 of 512-mer



Supplementary Figure 13. Fragment pattern of F7 and F15 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	967.32	967.48			$[M+Na]^+$	967.32	967.48		
y7	831.436	830.217	137.263	M	a8	895.466	895.437	72.043	L
y6	683.386	683.285	146.932	P	a7	823.446	823.614	71.823	L
y5	611.366	611.283	72.002	L	a6	675.396	675.345	148.269	P
y4	463.316	463.226	148.057	P	a5	527.346	527.264	148.081	P
y3	315.266	315.15	148.076	P	a4	379.296	379.222	148.042	P

Supplementary Table 7. Decoding table of F7 and F15 of 512-mer

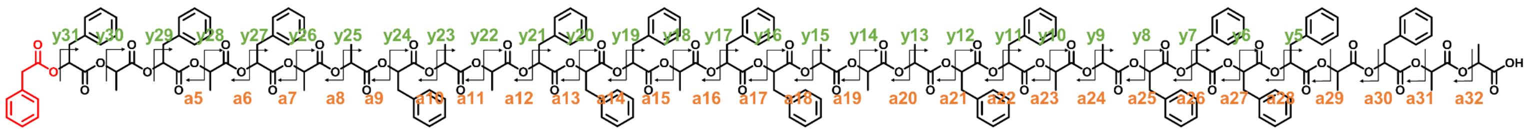


Supplementary Figure 14. Fragment pattern of F9 and F16 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
[M+Na]+	4336.43	4339.978			[M+Na]+	4336.43	4339.978		
y39	4203.928	4201.052	138.926	M	a40	4267.958	4267.969	72.009	L
y38	4131.908	4130.331	70.721	L	a39	4195.938	4196.185	71.784	L
y37	3983.858	3982.245	148.086	P	a38	4047.888	4047.352	148.833	P
y36	3835.808	3833.36	148.885	P	a37	3975.868	3976.384	70.968	L
y35	3763.788	3761.405	71.955	L	a36	3903.848	3904.675	71.709	L
y34	3615.738	3615.226	146.179	P	a35	3831.828	3833.36	71.315	L
y33	3467.688	3467.462	147.764	P	a34	3759.808	3759.405	73.955	L
y32	3395.668	3395.612	71.85	L	a33	3611.758	3611.696	147.709	P
y31	3247.618	3247.813	147.799	P	a32	3539.738	3539.987	71.709	L
y30	3175.598	3175.703	72.11	L	a31	3467.718	3467.462	72.525	L
y29	3103.578	3103.837	71.866	L	a30	3319.668	3319.515	147.947	P
y28	2955.528	2955.498	148.339	P	a29	3247.648	3248.046	71.469	L
y27	2807.478	2807.344	148.154	P	a28	3099.598	3099.744	148.302	P
y26	2659.428	2659.868	147.476	P	a27	3027.578	3026.998	72.746	L
y25	2587.408	2586.761	73.107	L	a26	2879.528	2879.428	147.57	P
y24	2439.358	2438.655	148.106	P	a25	2731.478	2731.366	148.062	P
y23	2367.338	2367.098	71.557	L	a24	2659.458	2659.309	72.057	L
y22	2219.288	2220.493	146.605	P	a23	2587.438	2586.967	72.342	L
y21	2071.238	2071.707	148.786	P	a22	2439.388	2439.188	147.779	P
y20	1999.218	1998.639	73.068	L	a21	2367.368	2366.901	72.287	L
y19	1927.198	1926.494	72.145	L	a20	2295.348	2294.996	71.905	L
y18	1779.148	1778.749	147.745	P	a19	2147.298	2146.676	148.32	P
y17	1707.128	1706.396	72.353	L	a18	1999.248	1998.699	147.977	P
y16	1635.108	1634.451	71.945	L	a17	1927.228	1926.788	71.911	L
y15	1487.058	1485.629	148.822	P	a16	1779.178	1778.637	148.151	P
y14	1339.008	1337.277	148.352	P	a15	1707.158	1706.507	72.13	L
y13	1266.988	1264.175	73.102	L	a14	1559.108	1558.414	148.093	P
y12	1118.938	1117.306	146.869	P	a13	1411.058	1410.345	148.069	P
y11	1046.918	1045.339	71.967	L	a12	1263.008	1262.363	147.982	P

y10	898.868	897.052	148.287	P	a11	1190.988	1190.275	72.088	L
y9	826.848	825.401	71.651	L	a10	1118.968	1118.262	72.013	L

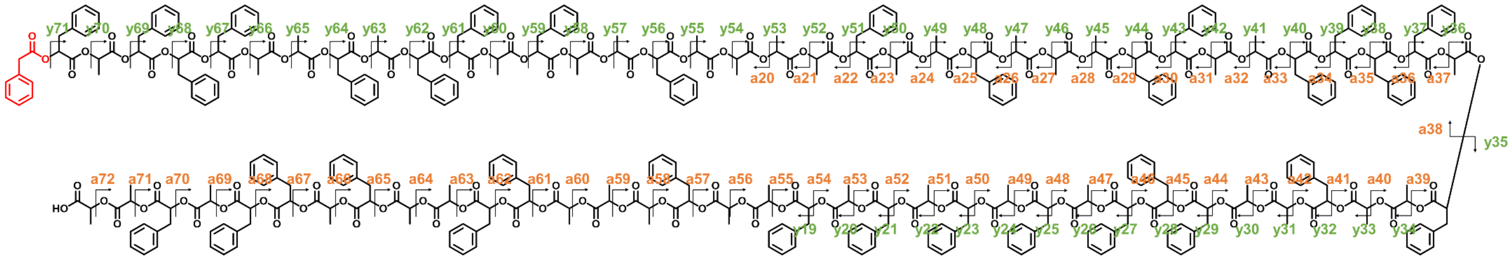
Supplementary Table 8. Decoding table of F9 and F16 of 512-mer



Supplementary Figure 15. Fragment pattern of F10 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	3608.2	3611.634			$[M+Na]^+$	3608.2	3611.634		
y31	3475.584	3478.072	133.562	M	a32	3539.614	3539.163	72.471	L
y30	3327.534	3330.838	147.234	P	a31	3467.594	3467.842	71.321	L
y29	3255.514	3258.443	72.395	L	a30	3319.544	3318.784	149.058	P
y28	3107.464	3110.953	147.49	P	a29	3247.524	3246.445	72.339	L
y27	3035.444	3037.481	73.472	L	a28	3099.474	3099.534	146.911	P
y26	2887.394	2889.455	148.026	P	a27	2951.424	2950.97	148.564	P
y25	2815.374	2816.799	72.656	L	a26	2803.374	2802.99	147.98	P
y24	2743.354	2745.679	71.12	L	a25	2655.324	2654.2	148.79	P
y23	2595.304	2596.795	148.884	P	a24	2583.304	2582.093	72.107	L
y22	2523.284	2524.47	72.325	L	a23	2511.284	2509.941	72.152	L
y21	2451.264	2452.64	71.83	L	a22	2363.234	2362.667	147.274	P
y20	2303.214	2302.89	149.75	P	a21	2215.184	2213.983	148.684	P
y19	2155.164	2156.415	146.475	P	a20	2143.164	2142.254	71.729	L
y18	2007.114	2009.154	147.261	P	a19	2071.144	2070.181	72.073	L
y17	1935.094	1938.749	70.405	L	a18	1999.124	1997.488	72.693	L
y16	1787.044	1789.399	149.35	P	a17	1851.074	1849.868	147.62	P
y15	1638.994	1641.405	147.994	P	a16	1703.024	1701.891	147.977	P
y14	1566.974	1567.808	73.597	L	a15	1631.004	1629.09	72.801	L
y13	1494.954	1494.514	73.294	L	a14	1482.954	1481.649	147.441	P
y12	1422.934	1420.963	73.551	L	a13	1334.904	1332.799	148.85	P
y11	1274.884	1273.563	147.4	P	a12	1186.854	1184.868	147.931	P
y10	1126.834	1125.558	148.005	P	a11	1114.834	1113.712	71.156	L
y9	1054.814	1052.6	72.958	L	a10	1042.814	1040.571	73.141	L
y8	982.794	980.693	71.907	L	a9	894.764	892.555	148.016	P
y7	834.744	832.542	148.151	P	a8	822.744	820.487	72.068	L
y6	686.694	684.638	147.904	P	a7	750.724	748.774	71.713	L
y5	538.644	536.469	148.169	P	a6	602.674	600.606	148.168	P

Supplementary Table 9. Decoding table of F10 of 512-mer

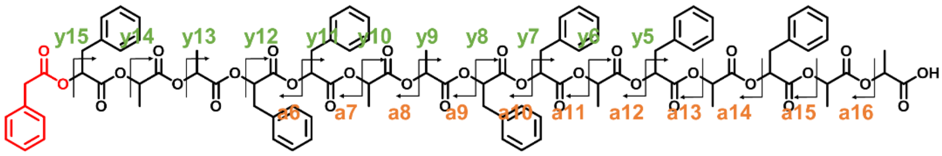


Supplementary Figure 16. Fragment pattern of F11 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
[M+Na]+	7781.58	7782.936			[M+Na]+	7781.58	7782.936		
y71	7646.886	7647.715	135.221	M	a72	7710.916	7710.844	72.092	L
y70	7498.836	7499.954	147.761	P	a71	7638.896	7638.715	72.129	L
y69	7426.816	7427.351	72.603	L	a70	7490.846	7492.309	146.406	P
y68	7278.766	7279.075	148.276	P	a69	7418.826	7418.762	73.547	L
y67	7130.716	7131.237	147.838	P	a68	7270.776	7271.229	147.533	P
y66	6982.666	6983.913	147.324	P	a67	7122.726	7123.237	147.992	P
y65	6910.646	6910.082	73.831	L	a66	7050.706	7052.598	70.639	L
y64	6838.626	6839.177	70.905	L	a65	6902.656	6904.972	147.626	P
y63	6690.576	6691.539	147.638	P	a64	6830.636	6833.177	71.795	L
y62	6618.556	6619.297	72.242	L	a63	6758.616	6759.595	73.582	L
y61	6470.506	6471.791	147.506	P	a62	6610.566	6613.409	146.186	P
y60	6322.456	6323.279	148.512	P	a61	6462.516	6464.386	149.023	P
y59	6250.436	6251.849	71.43	L	a60	6390.496	6392.634	71.752	L
y58	6102.386	6103.027	148.822	P	a59	6318.476	6320.279	72.355	L
y57	6030.366	6030.525	72.502	L	a58	6246.456	6247.849	72.43	L
y56	5958.346	5960.337	70.188	L	a57	6098.406	6099.921	147.928	P
y55	5810.296	5812.782	147.555	P	a56	6026.386	6026.525	73.396	L
y54	5738.276	5739.783	72.999	L	a55	5954.366	5954.443	72.082	L
y53	5666.256	5668.169	71.614	L	a54	5806.316	5807.886	146.557	P
y52	5594.236	5594.658	73.511	L	a53	5734.296	5735.887	71.999	L
y51	5522.216	5522.145	72.513	L	a52	5586.246	5590.658	145.229	P
y50	5374.166	5375.194	146.951	P	a51	5514.226	5519.145	71.513	L
y49	5302.146	5302.796	72.398	L	a50	5366.176	5371.094	148.051	P
y48	5230.126	5230.8	71.996	L	a49	5294.156	5299.994	71.1	L
y47	5082.076	5083.633	147.167	P	a48	5146.106	5150.131	149.863	P

y46	5010.056	5012.356	71.277	L	a47	5074.086	5078.181	71.95	L
y45	4938.036	4940.462	71.894	L	a46	4926.036	4931.129	147.052	P
y44	4866.016	4869.903	70.559	L	a45	4777.986	4782.902	148.227	P
y43	4717.966	4722.757	147.146	P	a44	4629.936	4634.405	148.497	P
y42	4569.916	4574.553	148.204	P	a43	4557.916	4562.462	71.943	L
y41	4497.896	4502.389	72.164	L	a42	4485.896	4491.298	71.164	L
y40	4425.876	4430.441	71.948	L	a41	4337.846	4345.235	146.063	P
y39	4277.826	4282.438	148.003	P	a40	4265.826	4270.255	74.98	L
y38	4129.776	4133.623	148.815	P	a39	4193.806	4198.505	71.75	L
y37	3981.726	3984.232	149.391	P	a38	4045.756	4051.97	146.535	P
y36	3833.676	3836.779	147.453	P	a37	3973.736	3979.147	72.823	L
y35	3761.656	3763.131	73.648	L	a36	3825.686	3830.696	148.451	P
y34	3613.606	3615.005	148.126	P	a35	3677.636	3683.209	147.487	P
y33	3541.586	3542.248	72.757	L	a34	3529.586	3536.248	146.961	P
y32	3469.566	3470.334	71.914	L	a33	3381.536	3387.398	148.85	P
y31	3321.516	3323.791	146.543	P	a32	3309.516	3315.443	71.955	L
y30	3249.496	3251.913	71.878	L	a31	3237.496	3242.177	73.266	L
y29	3177.476	3178.968	72.945	L	a30	3089.446	3095.279	146.898	P
y28	3029.426	3031.664	147.304	P	a29	2941.396	2947.794	147.485	P
y27	2881.376	2883.293	148.371	P	a28	2869.376	2875.27	72.524	L
y26	2733.326	2734.564	148.729	P	a27	2797.356	2802.012	73.258	L
y25	2661.306	2663.453	71.111	L	a26	2725.336	2731.147	70.865	L
y24	2513.256	2514.265	149.188	P	a25	2577.286	2583.584	147.563	P
y23	2441.236	2442.178	72.087	L	a24	2505.266	2511.265	72.319	L
y22	2293.186	2294.733	147.445	P	a23	2433.246	2438.656	72.609	L
y21	2221.166	2222.105	72.628	L	a22	2285.196	2290.797	147.859	P
y20	2073.116	2074.882	147.223	P	a21	2213.176	2219.168	71.629	L
y19	2001.096	2003.394	71.488	L	a20	2141.156	2146.292	72.876	L

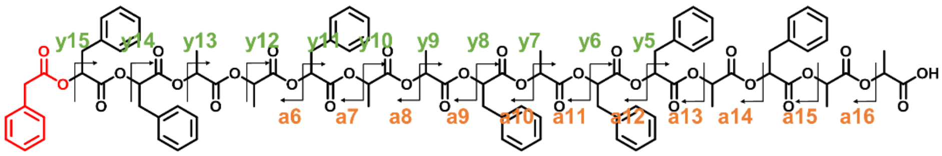
Supplementary Table 10. Decoding table of F11 of 512-mer



Supplementary Figure 17. Fragment pattern of F12 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	1771.58	1771.73			$[M+Na]^+$	1771.58	1771.73		
y15	1635.68	1634.802	136.928	M	a16	1699.71	1699.91	71.82	L
y14	1487.63	1490.691	144.111	P	a15	1627.69	1627.897	72.013	L
y13	1415.61	1414.834	75.857	L	a14	1479.64	1479.79	148.107	P
y12	1343.59	1343.787	71.047	L	a13	1407.62	1407.863	71.927	L
y11	1195.54	1195.623	148.164	P	a12	1259.57	1259.639	148.224	P
y10	1047.49	1048.51	147.113	P	a11	1187.55	1187.605	72.034	L
y9	975.47	975.397	73.113	L	a10	1039.5	1039.55	148.055	P
y8	903.45	903.593	71.804	L	a9	891.45	891.495	148.055	P
y7	755.4	755.487	148.106	P	a8	819.43	819.456	72.039	L
y6	607.35	607.38	148.107	P	a7	747.41	747.452	72.004	L
y5	535.33	535.346	72.034	L	a6	599.36	599.38	148.072	P

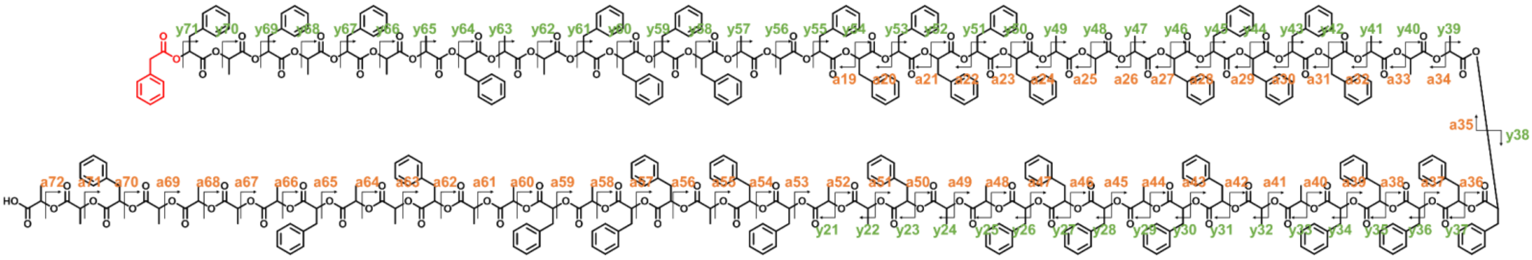
Supplementary Table 11. Decoding table of F12 of 512-mer



Supplementary Figure 18. Fragment pattern of F14 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
$[M+Na]^+$	1771.58	1771.73			$[M+Na]^+$	1771.58	1771.73		
y15	1635.68	1633.802	137.928	M	a16	1699.71	1699.91	71.82	L
y14	1487.63	1490.691	143.111	P	a15	1627.69	1627.897	72.013	L
y13	1339.58	1343.537	147.154	P	a14	1479.64	1479.79	148.107	P
y12	1267.56	1267.523	76.014	L	a13	1407.62	1407.863	71.927	L
y11	1195.54	1195.623	71.9	L	a12	1259.57	1259.639	148.224	P
y10	1047.49	1048.51	147.113	P	a11	1111.52	1111.523	148.116	P
y9	975.47	975.397	73.113	L	a10	1039.5	1039.55	71.973	L
y8	903.45	903.593	71.804	L	a9	891.45	891.495	148.055	P
y7	755.4	755.487	148.106	P	a8	819.43	819.456	72.039	L
y6	683.38	683.451	72.036	L	a7	747.41	747.452	72.004	L
y5	535.33	535.346	148.105	P	a6	599.36	599.38	148.072	P

Supplementary Table 12. Decoding table of F14 of 512-mer

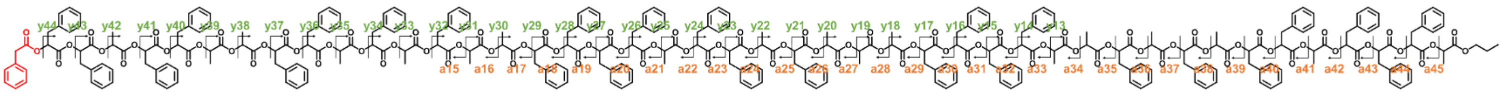


Supplementary Figure 19. Fragment pattern of F17 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
[M+Na]+	8161.73	8165.068			[M+Na]+	8161.73	8165.068		
y71	8029.018	8034.953	130.115	M	a72	8093.048	8093.066	72.002	L
y70	7880.968	7885.324	149.629	P	a71	8021.028	8020.054	73.012	L
y69	7808.948	7813.746	71.578	L	a70	7872.978	7871.987	148.067	P
y68	7660.898	7665.387	148.359	P	a69	7800.958	7801.867	70.12	L
y67	7588.878	7593.546	71.841	L	a68	7728.938	7729.022	72.845	L
y66	7440.828	7445.514	148.032	P	a67	7656.918	7657.952	71.07	L
y65	7368.808	7373.729	71.785	L	a66	7584.898	7585.546	72.406	L
y64	7296.788	7301.358	72.371	L	a65	7436.848	7437.514	148.032	P
y63	7148.738	7154.453	146.905	P	a64	7364.828	7364.044	73.47	L
y62	7076.718	7081.281	73.172	L	a63	7292.808	7293.474	70.57	L
y61	7004.698	7010.32	70.961	L	a62	7144.758	7146.338	147.136	P
y60	6856.648	6861.635	148.685	P	a61	7072.738	7074.281	72.057	L
y59	6708.598	6713.099	148.536	P	a60	7000.718	7003.549	70.732	L
y58	6560.548	6565.998	147.101	P	a59	6852.668	6854.748	148.801	P
y57	6412.498	6417.785	148.213	P	a58	6780.648	6784.603	70.145	L
y56	6340.478	6345.254	72.531	L	a57	6632.598	6637.64	146.963	P
y55	6268.458	6272.369	72.885	L	a56	6484.548	6489.085	148.555	P
y54	6120.408	6123.639	148.73	P	a55	6412.528	6418.675	70.41	L
y53	5972.358	5977.468	146.171	P	a54	6264.478	6269.477	149.198	P
y52	5824.308	5830.385	147.083	P	a53	6116.428	6122.639	146.838	P
y51	5676.258	5683.899	146.486	P	a52	6044.408	6049.07	73.569	L
y50	5528.208	5536.852	147.047	P	a51	5972.388	5977.362	71.708	L
y49	5380.158	5388.653	148.199	P	a50	5824.338	5830.385	146.977	P
y48	5308.138	5316.045	72.608	L	a49	5752.318	5757.051	73.334	L
y47	5236.118	5241.406	74.639	L	a48	5680.298	5684.796	72.255	L

y46	5164.098	5168.788	72.618	L	a47	5532.248	5536.751	148.045	P
y45	5016.048	5020.75	148.038	P	a46	5384.198	5388.753	147.998	P
y44	4867.998	4872.54	148.21	P	a45	5236.148	5241.7	147.053	P
y43	4719.948	4727.993	144.547	P	a44	5164.128	5168.691	73.009	L
y42	4571.898	4579.907	148.086	P	a43	5016.078	5020.942	147.749	P
y41	4423.848	4432.263	147.644	P	a42	4868.028	4872.54	148.402	P
y40	4351.828	4360.41	71.853	L	a41	4796.008	4800.539	72.001	L
y39	4279.808	4285.781	74.629	L	a40	4723.988	4727.993	72.546	L
y38	4207.788	4214.266	71.515	L	a39	4575.938	4579.907	148.086	P
y37	4059.738	4066.377	147.889	P	a38	4427.888	4433.611	146.296	P
y36	3911.688	3918.016	148.361	P	a37	4279.838	4285.781	147.83	P
y35	3763.638	3774.073	143.943	P	a36	4131.788	4136.895	148.886	P
y34	3615.588	3625.806	148.267	P	a35	3983.738	3990.215	146.68	P
y33	3467.538	3476.75	149.056	P	a34	3835.688	3844.936	145.279	P
y32	3319.488	3329.165	147.585	P	a33	3763.668	3772.156	72.78	L
y31	3247.468	3257.449	71.716	L	a32	3691.648	3698.545	73.611	L
y30	3175.448	3182.268	75.181	L	a31	3619.628	3625.806	72.739	L
y29	3027.398	3033.231	149.037	P	a30	3471.578	3477.223	148.583	P
y28	2879.348	2889.771	143.46	P	a29	3323.528	3329.474	147.749	P
y27	2807.328	2817.06	72.711	L	a28	3175.478	3180.536	148.938	P
y26	2659.278	2668.059	149.001	P	a27	3027.428	3033.818	146.718	P
y25	2511.228	2521.279	146.78	P	a26	2879.378	2886.914	146.904	P
y24	2363.178	2372.826	148.453	P	a25	2807.358	2815.507	71.407	L
y23	2291.158	2300.323	72.503	L	a24	2735.338	2742.249	73.258	L
y22	2219.138	2226.208	74.115	L	a23	2663.318	2669.059	73.19	L
y21	2071.088	2076.395	149.813	P	a22	2515.268	2521.412	147.647	P
y20	1999.068	2009.043	67.352	L	a21	2367.218	2372.053	149.359	P
y19	1927.048	1928.332	80.711	L	a20	2219.168	2224.333	147.72	P
y18	1778.998	1784.205	144.127	P	a19	2071.118	2076.455	147.878	P

Supplementary Table 13. Decoding table of F17 of 512-mer



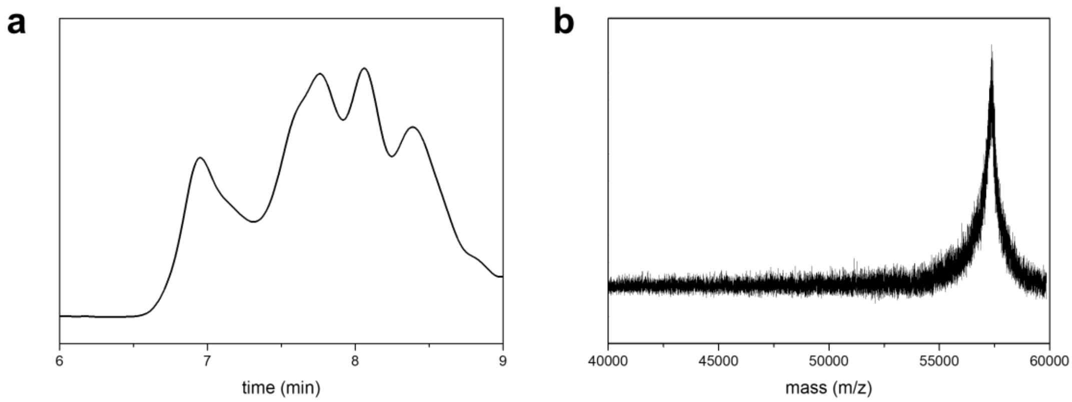
Supplementary Figure 20. Fragment pattern of F18 of 512-mer

M to COOH	Calc. m/z	Found m/z	Diff.	sequence	COOH to M	Calc. m/z	Found m/z	Diff.	sequence
[M+Na]+	5346.83	5348.768			[M+Na]+	5346.83	5348.768		
y44	5212.718	5211.855	136.913	M	a45	5235.708	5233.568	115.2	L-All
y43	5064.668	5063.026	148.829	P	a44	5087.658	5086.007	147.561	P
y42	4916.618	4916.245	146.781	P	a43	4939.608	4938.017	147.99	P
y41	4844.598	4843.889	72.356	L	a42	4791.558	4791.819	146.198	P
y40	4696.548	4697.076	146.813	P	a41	4719.538	4720.204	71.615	L
y39	4548.498	4549.818	147.258	P	a40	4571.488	4572.142	148.062	P
y38	4476.478	4475.959	73.859	L	a39	4423.438	4424.581	147.561	P
y37	4404.458	4404.623	71.336	L	a38	4351.418	4352.929	71.652	L
y36	4256.408	4256.51	148.113	P	a37	4203.368	4205.127	147.802	P
y35	4108.358	4109.976	146.534	P	a36	4131.348	4133.721	71.406	L
y34	4036.338	4036.804	73.172	L	a35	3983.298	3984.677	149.044	P
y33	3888.288	3889.042	147.762	P	a34	3911.278	3910.813	73.864	L
y32	3816.268	3816.372	72.67	L	a33	3839.258	3838.435	72.378	L
y31	3668.218	3669.329	147.043	P	a32	3691.208	3690.939	147.496	P
y30	3596.198	3596.829	72.5	L	a31	3543.158	3543.583	147.356	P
y29	3524.178	3524.692	72.137	L	a30	3395.108	3396.006	147.577	P
y28	3376.128	3377.143	147.549	P	a29	3247.058	3247.512	148.494	P
y27	3228.078	3229.246	147.897	P	a28	3175.038	3175.007	72.505	L
y26	3080.028	3081.489	147.757	P	a27	3103.018	3103.147	71.86	L
y25	2931.978	2932.438	149.051	P	a26	3030.998	3032.236	70.911	L
y24	2859.958	2861.046	71.392	L	a25	2882.948	2883.821	148.415	P
y23	2711.908	2713.139	147.907	P	a24	2810.928	2811.494	72.327	L
y22	2563.858	2565.322	147.817	P	a23	2662.878	2663.793	147.701	P
y21	2491.838	2494.037	71.285	L	a22	2514.828	2516.22	147.573	P
y20	2343.788	2345.056	148.981	P	a21	2442.808	2444.144	72.076	L
y19	2271.768	2272.139	72.917	L	a20	2294.758	2295.881	148.263	P
y18	2199.748	2200.384	71.755	L	a19	2146.708	2147.153	148.728	P
y17	2127.728	2128.31	72.074	L	a18	1998.658	1999.055	148.098	P
y16	1979.678	1980.144	148.166	P	a17	1850.608	1851.066	147.989	P

y15	1831.628	1832.23	147.914	P	a16	1778.588	1778.968	72.098	L
y14	1683.578	1684.071	148.159	P	a15	1706.568	1707.4	71.568	L
y13	1535.528	1535.247	148.824	P	a14	1558.518	1559.241	148.159	P

Supplementary Table 14. Decoding table of F18 of 512-mer

7. Random access of PcL



Supplementary Figure 21. Partially fragmentation of 512-mer. a GPC spectrum after hydrogenolysis of 512-mer for 30 min **b** MALDI-TOF spectrum after hydrogenolysis of 512-mer for 30 min

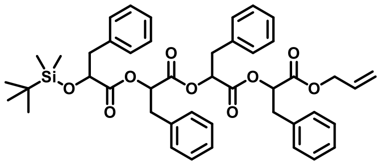
High molecular weight fragments of 512-mer generated after partially hydrogenolysis could not be measured due to the resolution and sensitivity limit of the MALDI-TOF mass spectrometer used in this study.

Found m/z	Calc. m/z	fragment	Found m/z	Calc. m/z	fragment
28707.54	28702.34	starting	4337.993	4336.43	F9
26362.11	26364.79	F(2~11)	3609.771	3608.2	F10
24389.54	24388.73	F(3~11)	3457.724	3456.14	F4
23447.79	23445.77	F(4~11)	2942.624	2942.02	F(2,3)
20014.98	20012.31	F(5~11)	2776.575	2775.79	F(6,7,8)
18497.56	18492.07	F(3~9)	2360.575	2359.83	F1
17558.41	17556.75	F(4~9)	2326.489	2325.79	F11
16613.67	16613.03	F(1~6)	2000.435	1999.67	F2
15748.01	15746.16	F(1~5)	1910.411	1908.92	F(7,8)
12980.46	12979.17	F(6~11)	1833.371	1832.83	F(6,7)
12114.61	12112.3	F(7~11)	967.222	967.32	F3, F7, F8
11170.64	11169.33	F(8~11)	891.187	891.28	F6
10230.46	10226.36	F(9~11)			
9809.249	9808.93	F(7~10)			
8869.763	8865.97	F(8~10)			
8715.088	8713.02	F(1~4)			
7923.286	7923.00	F(9,10)			
7059.057	7058.15	F5			
6221.617	6223.28	F(7,8,9)			
5911.003	5912.01	F(10,11)			
5279.954	5279.56	F(1~3)			

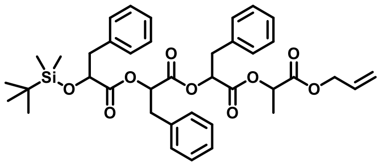
Supplementary Table 15. MALDI-TOF mass list and fragments of 256-mer_repeat unit 1-256.

8. Characterization of PcLs

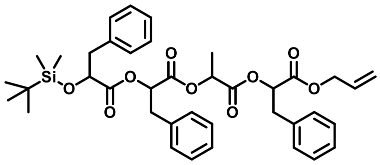
Characterization of tetramers



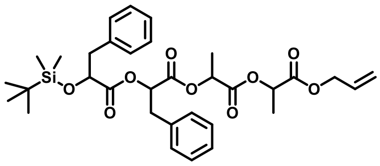
PPPP 4.6 g, 88%; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.36 – 6.90 (m, 22H), 5.83 (ddt, J = 16.7, 11.0, 5.7 Hz, 1H), 5.46 – 5.19 (m, 5H), 4.60 (d, J = 6.0 Hz, 2H), 4.35 – 4.19 (m, 1H), 3.33 – 2.56 (m, 9H), 0.76 – 0.66 (m, 9H), -0.18 – -0.44 (m, 6H).



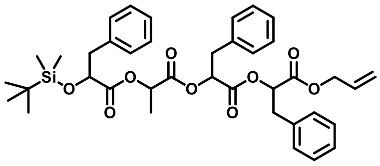
PPPL 10.9 g, 90%; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.37 – 6.95 (m, 16H), 6.01 – 5.81 (m, 1H), 5.47 – 5.07 (m, 5H), 4.70 – 4.55 (m, 2H), 4.37 – 4.16 (m, 1H), 3.43 – 2.56 (m, 6H), 1.58 – 1.44 (m, 4H), 0.78 – 0.67 (m, 9H), -0.16 – -0.43 (m, 6H).



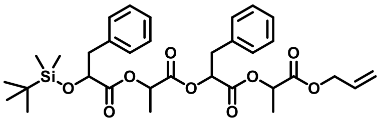
PPLP 12.8 g, 89%; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.34 – 7.00 (m, 17H), 5.91 – 5.75 (m, 1H), 5.41 – 5.13 (m, 5H), 4.65 – 4.53 (m, 2H), 4.39 – 4.23 (m, 1H), 3.38 – 2.58 (m, 6H), 1.60 – 1.28 (m, 4H), 0.71 (dd, J = 11.4, 2.6 Hz, 9H), -0.13 – -0.42 (m, 6H).



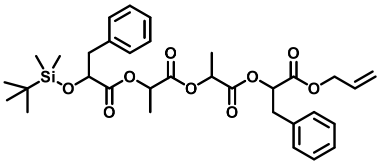
PPLL 12.3 g, 88%; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.33 – 7.01 (m, 12H), 5.88 (tdd, J = 13.1, 6.5, 4.3 Hz, 1H), 5.41 – 5.12 (m, 5H), 4.63 (ddp, J = 8.5, 4.0, 1.4 Hz, 2H), 4.38 – 4.23 (m, 1H), 3.46 – 2.59 (m, 4H), 1.64 – 1.50 (m, 7H), 0.72 (dt, J = 11.3, 1.7 Hz, 9H), -0.15 – -0.40 (m, 6H).



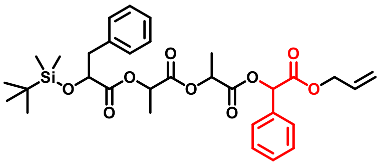
PLPP 17 g, 87%; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.37 – 6.93 (m, 16H), 5.89 – 5.77 (m, 1H), 5.43 – 5.09 (m, 5H), 4.59 (d, J = 6.3 Hz, 2H), 4.40 – 4.27 (m, 1H), 3.34 – 2.72 (m, 6H), 1.50 – 1.24 (m, 4H), 0.76 (dd, J = 6.7, 3.3 Hz, 9H), -0.08 – -0.35 (m, 6H).



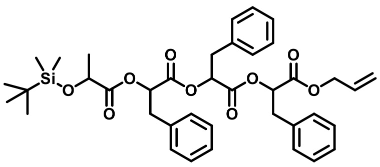
PLPL 25.4 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 – 7.17 (m, 11H), 5.97 – 5.82 (m, 1H), 5.44 – 5.07 (m, 5H), 4.69 – 4.60 (m, 2H), 4.42 – 4.31 (m, 1H), 3.42 – 2.75 (m, 5H), 1.56 – 1.28 (m, 7H), 0.76 (dq, J = 6.6, 1.9 Hz, 9H), -0.09 – -0.35 (m, 6H).



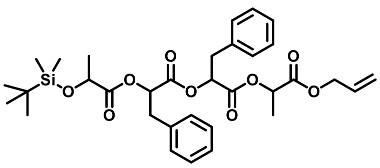
PLLP 53.2 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.25 (td, J = 11.5, 6.1 Hz, 14H), 5.81 (ddt, J = 16.7, 11.0, 5.7 Hz, 1H), 5.36 – 5.07 (m, 5H), 4.65 – 4.53 (m, 2H), 4.38 (td, J = 8.5, 3.3 Hz, 1H), 3.30 – 2.80 (m, 4H), 1.62 – 1.34 (m, 9H), 0.84 – 0.68 (m, 9H), -0.06 – -0.32 (m, 6H).



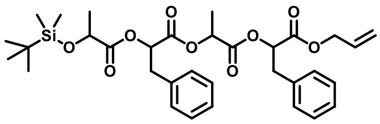
PLLM 16.8 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 – 7.19 (m, 11H), 6.02 – 5.94 (m, 1H), 5.81 (ddt, J = 16.6, 10.9, 5.4 Hz, 1H), 5.37 – 5.08 (m, 4H), 4.62 (t, J = 4.8 Hz, 2H), 4.38 (td, J = 9.3, 4.9 Hz, 1H), 3.25 – 2.83 (m, 2H), 1.74 – 1.36 (m, 7H), 0.84 – 0.71 (m, 9H), -0.08 – -0.32 (m, 6H).



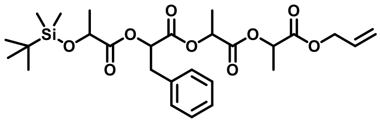
LPPP 5.4 g, 86%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.91 (m, 17H), 5.83 (ddt, J = 16.5, 10.2, 5.8 Hz, 1H), 5.45 – 5.20 (m, 5H), 4.64 – 4.58 (m, 2H), 4.29 (dq, J = 12.4, 6.7 Hz, 1H), 3.32 – 2.80 (m, 7H), 1.38 – 1.18 (m, 3H), 0.86 (dd, J = 4.9, 2.0 Hz, 10H), -0.05 (s, 6H).



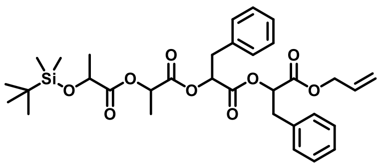
LPPL 10.3 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.95 (m, 11H), 5.90 (ddq, J = 21.0, 10.7, 5.4 Hz, 1H), 5.44 – 5.07 (m, 5H), 4.68 – 4.59 (m, 2H), 4.38 – 4.18 (m, 1H), 3.43 – 2.82 (m, 4H), 1.63 – 1.19 (m, 7H), 0.86 (dd, J = 4.4, 1.7 Hz, 9H), -0.05 (s, 6H).



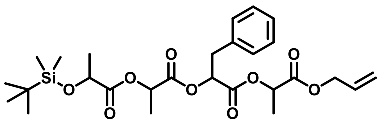
LPLP 1.4 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 – 7.12 (m, 11H), 5.82 (ddt, J = 16.6, 11.1, 5.8 Hz, 1H), 5.37 – 5.13 (m, 5H), 4.65 – 4.55 (m, 2H), 4.33 (qd, J = 8.9, 6.0 Hz, 1H), 3.38 – 3.02 (m, 4H), 1.63 – 1.16 (m, 6H), 0.86 (q, J = 2.4 Hz, 9H), 0.02 (dt, J = 8.5, 1.9 Hz, 6H).



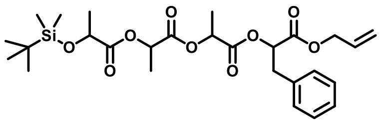
LPLL 1.2 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32 – 7.20 (m, 7H), 5.90 (ddt, J = 16.5, 10.9, 5.8 Hz, 1H), 5.39 – 5.11 (m, 5H), 4.69 – 4.58 (m, 2H), 4.39 – 4.25 (m, 1H), 3.43 – 3.06 (m, 2H), 1.64 – 1.20 (m, 11H), 0.86 (dd, J = 5.9, 2.7 Hz, 9H), 0.11 – -0.14 (m, 6H).



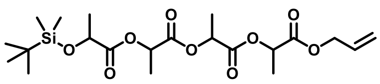
LLPP 8.2 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 – 6.92 (m, 11H), 5.91 – 5.76 (m, 1H), 5.40 – 5.07 (m, 5H), 4.64 – 4.55 (m, 2H), 4.33 (dq, J = 29.9, 6.6 Hz, 1H), 3.31 – 2.90 (m, 4H), 1.49 – 1.24 (m, 7H), 0.89 (q, J = 1.7 Hz, 9H), 0.14 – 0.03 (m, 6H).



LLPL 2.9 g, 87%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33 – 7.20 (m, 6H), 5.99 – 5.82 (m, 1H), 5.39 – 5.22 (m, 3H), 5.15 (ddtd, J = 11.8, 9.9, 6.5, 3.2 Hz, 2H), 4.71 – 4.59 (m, 2H), 4.34 (dtd, J = 18.6, 7.4, 5.9 Hz, 1H), 3.37 – 3.11 (m, 2H), 1.55 – 1.33 (m, 10H), 0.93 – 0.86 (m, 9H), 0.08 (dt, J = 8.7, 2.3 Hz, 6H).

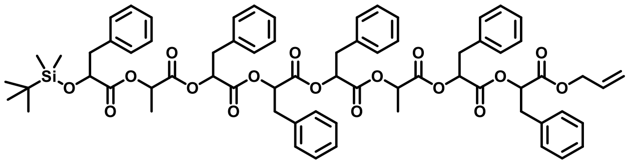


LLLP 7.7 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34 – 7.18 (m, 5H), 5.91 – 5.77 (m, 1H), 5.37 – 5.07 (m, 5H), 4.61 (ddt, J = 19.7, 5.8, 1.3 Hz, 2H), 4.40 (q, J = 6.6 Hz, 1H), 3.29 – 3.09 (m, 2H), 1.58 – 1.35 (m, 9H), 0.91 (s, 9H), 0.15 – 0.06 (m, 6H).

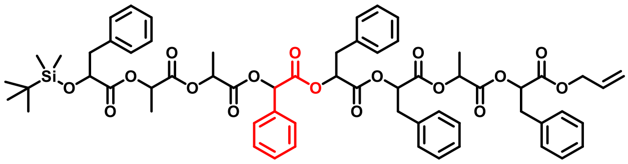


LLLL 6 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.89 (ddt, J = 16.5, 10.9, 5.8 Hz, 1H), 5.37 – 5.23 (m, 2H), 5.23 – 5.09 (m, 3H), 4.63 (ddt, J = 6.0, 3.3, 1.5 Hz, 2H), 4.44 – 4.35 (m, 1H), 1.62 – 1.42 (m, 12H), 0.96 – 0.86 (m, 9H), 0.10 (d, J = 10.1 Hz, 6H).

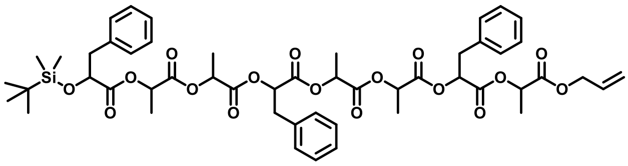
Characterization of 8-mers



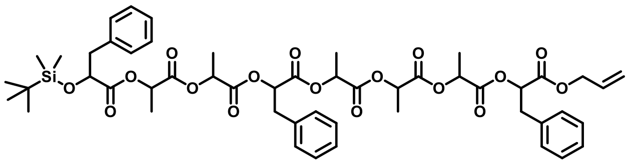
PLPPPLPP 1.2 g, 87%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.74 (m, 33H), 5.81 (td, *J* = 10.5, 5.0 Hz, 1H), 5.42 – 5.04 (m, 9H), 4.58 (q, *J* = 6.4 Hz, 2H), 4.35 (dt, *J* = 11.8, 6.0 Hz, 1H), 3.39 – 2.73 (m, 12H), 1.51 – 1.18 (m, 6H), 0.75 (dd, *J* = 5.9, 3.4 Hz, 9H), -0.07 – -0.36 (m, 6H).



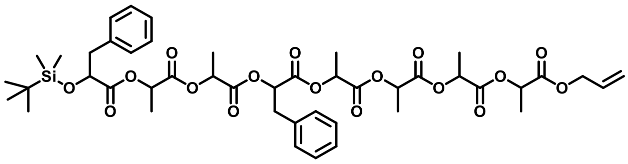
PLLMPPLP 4.1 g, 91%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.41 – 6.97 (m, 26H), 6.08 – 5.97 (m, 1H), 5.81 (ddt, *J* = 16.1, 10.5, 5.5 Hz, 1H), 5.38 – 5.03 (m, 8H), 4.59 (dt, *J* = 11.7, 5.8 Hz, 2H), 4.38 (d, *J* = 9.6 Hz, 1H), 3.33 – 2.75 (m, 8H), 1.55 – 1.25 (m, 9H), 0.83 – 0.70 (m, 9H), -0.06 – -0.34 (m, 6H).



PLLPLLPL 2.5 g, 89%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.35 – 7.18 (m, 17H), 5.90 (ddt, *J* = 16.6, 10.2, 6.1 Hz, 1H), 5.40 – 5.04 (m, 9H), 4.64 (d, *J* = 5.6 Hz, 2H), 4.43 – 4.32 (m, 1H), 3.43 – 2.78 (m, 6H), 1.56 – 1.27 (m, 15H), 0.77 (dt, *J* = 7.2, 1.6 Hz, 9H), -0.05 – -0.32 (m, 6H).

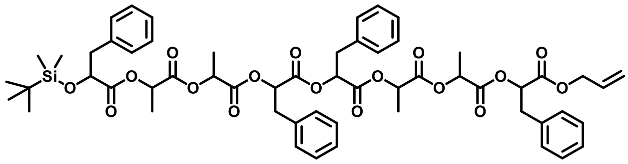


PLLPLLLP 3.2 g, 89%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.37 – 7.15 (m, 16H), 5.82 (ddq, *J* = 16.6, 11.1, 5.5 Hz, 1H), 5.40 – 5.07 (m, 9H), 4.60 (dd, *J* = 18.6, 5.3 Hz, 2H), 4.38 (td, *J* = 8.6, 3.1 Hz, 1H), 3.44 – 2.77 (m, 6H), 1.67 – 1.26 (m, 15H), 0.85 – 0.65 (m, 9H), -0.05 – -0.32 (m, 6H).

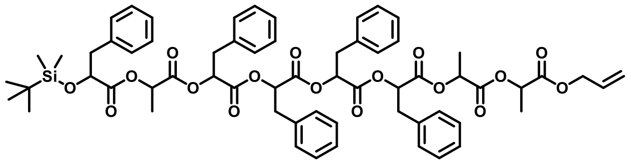


PLLPLLLL 7 g, 90%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.34 – 7.18 (m, 11H), 5.89 (ddt, *J* =

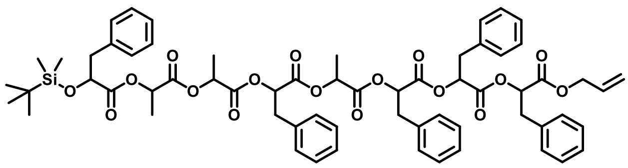
16.6, 11.0, 5.8 Hz, 1H), 5.40 – 5.07 (m, 9H), 4.62 (d, *J* = 5.6 Hz, 2H), 4.38 (td, *J* = 8.8, 4.7 Hz, 1H), 3.43 – 2.78 (m, 4H), 1.66 – 1.28 (m, 18H), 0.77 (dd, *J* = 7.1, 1.7 Hz, 9H), -0.08 – -0.32 (m, 6H).



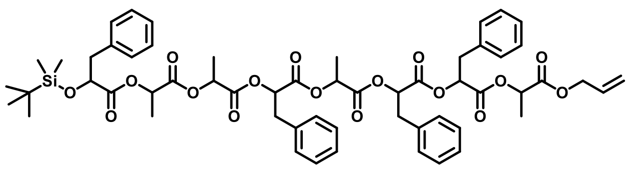
PLLPPLLP 1.1 g, 91%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.24 (dt, *J* = 28.3, 8.3 Hz, 20H), 5.89 – 5.74 (m, 1H), 5.39 – 5.05 (m, 9H), 4.66 – 4.53 (m, 2H), 4.39 (dd, *J* = 9.6, 3.6 Hz, 1H), 3.37 – 2.79 (m, 8H), 1.57 – 1.23 (m, 12H), 0.77 (d, *J* = 5.5 Hz, 9H), -0.07 – -0.33 (m, 6H).



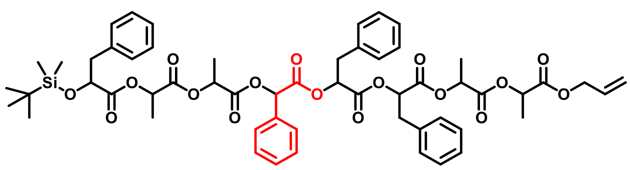
PLPPPPLL 1.2 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 6.75 (m, 28H), 5.96 – 5.82 (m, 1H), 5.42 – 5.04 (m, 9H), 4.61 (d, *J* = 7.8 Hz, 2H), 4.40 – 4.25 (m, 1H), 3.52 – 2.67 (m, 10H), 1.64 – 1.21 (m, 9H), 0.75 (dd, *J* = 5.8, 3.1 Hz, 9H), -0.11 – -0.38 (m, 6H).



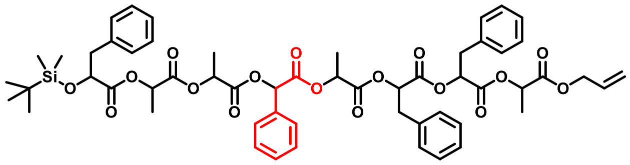
PLLPLPPP 1.5 g, 91%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 6.86 (m, 27H), 5.82 (ddt, *J* = 16.9, 11.4, 5.9 Hz, 1H), 5.42 – 5.01 (m, 9H), 4.59 (d, *J* = 6.3 Hz, 2H), 4.38 (d, *J* = 9.5 Hz, 1H), 3.35 – 2.78 (m, 10H), 1.57 – 1.18 (m, 9H), 0.84 – 0.70 (m, 9H), -0.09 – -0.33 (m, 6H).



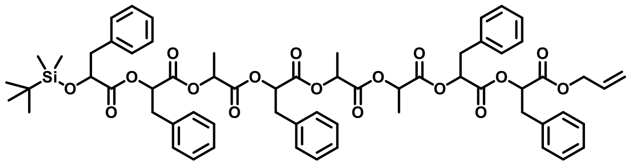
PLLPLPPL 4.3 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.41 – 6.82 (m, 22H), 5.90 (tt, *J* = 11.0, 5.5 Hz, 1H), 5.44 – 5.04 (m, 9H), 4.64 (t, *J* = 6.5 Hz, 2H), 4.38 (d, *J* = 8.7 Hz, 1H), 3.48 – 2.79 (m, 8H), 1.57 – 1.19 (m, 12H), 0.77 (d, *J* = 6.2 Hz, 9H), -0.08 – -0.34 (m, 6H).



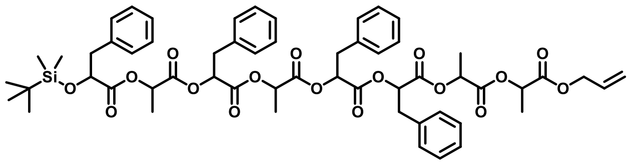
PLLMPPLL 2.8 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.43 – 6.58 (m, 22H), 6.10 – 5.82 (m, 2H), 5.40 – 5.04 (m, 8H), 4.63 (t, *J* = 6.2 Hz, 2H), 4.36 (t, *J* = 9.5 Hz, 1H), 3.41 – 2.74 (m, 6H), 1.65 – 1.30 (m, 12H), 0.84 – 0.67 (m, 9H), -0.06 – -0.34 (m, 6H).



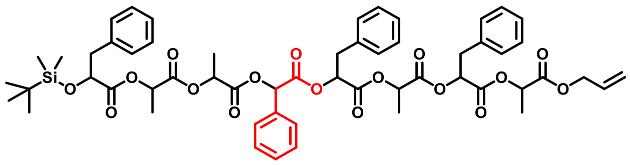
PLLMLPPL 2.2 g, 87%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 – 6.78 (m, 28H), 6.00 (d, *J* = 8.4 Hz, 1H), 5.89 (td, *J* = 11.0, 5.4 Hz, 1H), 5.39 – 5.05 (m, 8H), 4.64 (s, 2H), 4.38 (s, 1H), 3.41 – 2.63 (m, 6H), 1.68 – 1.16 (m, 12H), 0.77 (t, *J* = 3.4 Hz, 9H), -0.06 – -0.32 (m, 6H).



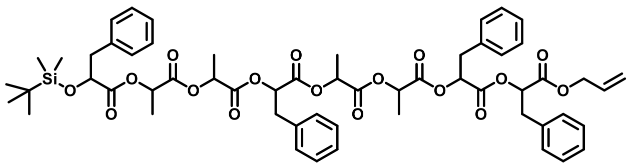
PPLPLLPP 3 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.90 (m, 28H), 5.82 (tt, *J* = 10.7, 5.5 Hz, 1H), 5.44 – 5.05 (m, 9H), 4.60 (t, *J* = 8.2 Hz, 2H), 4.38 – 4.21 (m, 1H), 3.40 – 2.59 (m, 10H), 1.56 – 1.19 (m, 9H), 0.80 – 0.62 (m, 9H), -0.12 – -0.42 (m, 6H).



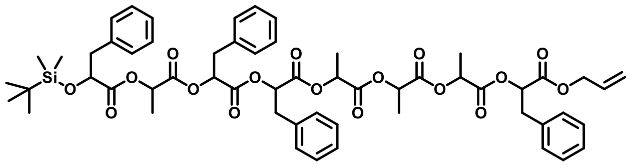
PLPLPPLL 2.4 g, 86%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 7.12 (m, 26H), 5.87 (dt, *J* = 10.8, 5.7 Hz, 1H), 5.41 – 5.07 (m, 9H), 4.63 (s, 2H), 4.35 (s, 1H), 3.42 – 2.71 (m, 8H), 1.67 – 1.20 (m, 12H), 0.82 – 0.68 (m, 9H), -0.07 – -0.38 (m, 6H).



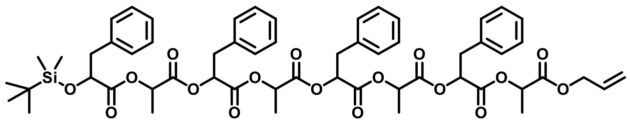
PLLMPLPL 2.1 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.42 – 6.87 (m, 22H), 6.11 – 5.84 (m, 2H), 5.43 – 4.99 (m, 8H), 4.70 – 4.60 (m, 2H), 4.42 – 4.34 (m, 1H), 3.42 – 2.79 (m, 6H), 1.58 – 1.08 (m, 12H), 0.78 (ddd, *J* = 7.2, 3.5, 2.1 Hz, 9H), -0.08 – -0.32 (m, 6H).



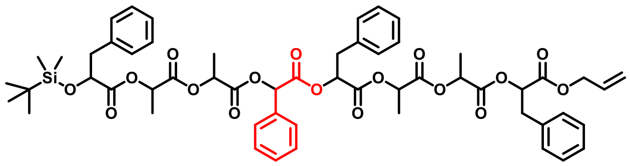
PLLPLLPP 4 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36 – 6.95 (m, 21H), 5.83 (ddq, *J* = 16.8, 11.3, 5.0 Hz, 1H), 5.43 – 5.05 (m, 9H), 4.61 (dd, *J* = 14.0, 4.9 Hz, 2H), 4.43 – 4.36 (m, 1H), 3.38 – 2.81 (m, 8H), 1.58 – 1.24 (m, 12H), 0.83 – 0.75 (m, 9H), -0.07 – -0.30 (m, 6H).



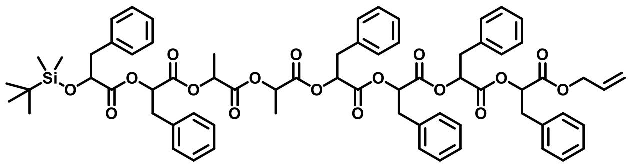
PLPPLLLP 1.3 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 6.83 (m, 21H), 5.80 (ddq, J = 16.5, 11.0, 5.5 Hz, 1H), 5.43 – 5.04 (m, 9H), 4.65 – 4.52 (m, 2H), 4.34 (td, J = 8.9, 3.6 Hz, 1H), 3.43 – 2.73 (m, 8H), 1.63 – 1.24 (m, 12H), 0.85 – 0.69 (m, 9H), -0.06 – -0.37 (m, 6H).



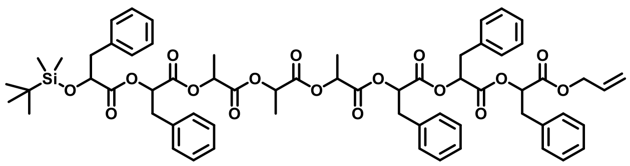
PLPLPLPL 1.8 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36 – 7.16 (m, 22H), 5.91 (ddd, J = 16.6, 10.7, 5.5 Hz, 1H), 5.42 – 5.09 (m, 9H), 4.65 (d, J = 6.3 Hz, 2H), 4.37 (dt, J = 10.4, 5.1 Hz, 1H), 3.45 – 2.73 (m, 8H), 1.57 – 1.24 (m, 12H), 0.83 – 0.71 (m, 9H), -0.07 – -0.34 (m, 6H).



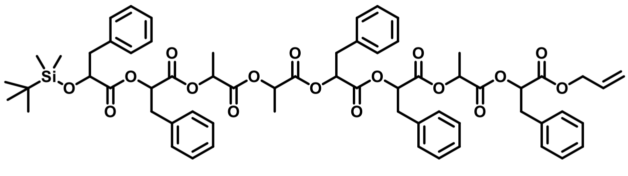
PLLMPLLP 1.5 g, 87%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 – 6.84 (m, 25H), 6.12 – 5.93 (m, 1H), 5.79 (s, 1H), 5.34 – 4.95 (m, 8H), 4.56 (d, J = 6.1 Hz, 2H), 4.36 (s, 1H), 3.31 – 2.75 (m, 6H), 1.67 – 1.17 (m, 12H), 0.76 (dd, J = 4.7, 2.7 Hz, 9H), -0.10 – -0.36 (m, 6H).



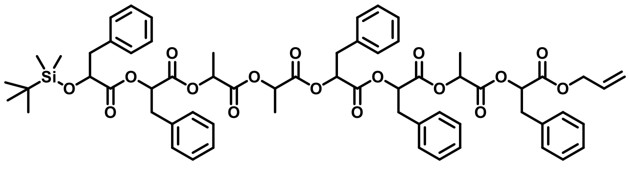
PPLLPPPP 1.9 g, 89%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 – 6.75 (m, 31H), 5.90 – 5.74 (m, 1H), 5.43 – 5.02 (m, 9H), 4.60 (d, J = 12.5 Hz, 2H), 4.38 – 4.19 (m, 1H), 3.45 – 2.51 (m, 12H), 1.53 – 1.20 (m, 6H), 0.77 – 0.62 (m, 9H), -0.13 – -0.45 (m, 6H).



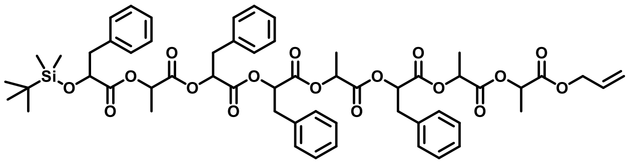
PPLLLPPP 2.7 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.82 (m, 26H), 5.82 (ddd, J = 16.9, 11.2, 5.7 Hz, 1H), 5.39 – 5.03 (m, 9H), 4.58 (d, J = 6.4 Hz, 2H), 4.38 – 4.21 (m, 1H), 3.45 – 2.55 (m, 10H), 1.55 – 1.21 (m, 9H), 0.80 – 0.65 (m, 9H), -0.13 – -0.44 (m, 6H).



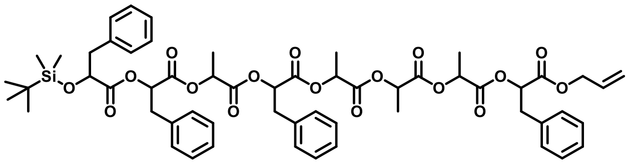
PPLLPPLP 1.3 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 – 6.82 (m, 26H), 5.82 (ddt, *J* = 16.1, 11.1, 5.5 Hz, 1H), 5.40 – 5.05 (m, 9H), 4.60 (dd, *J* = 12.0, 5.7 Hz, 2H), 4.40 – 4.21 (m, 1H), 3.44 – 2.59 (m, 10H), 1.57 – 1.25 (m, 9H), 0.72 (dd, *J* = 10.8, 2.3 Hz, 9H), -0.12 – -0.43 (m, 6H).



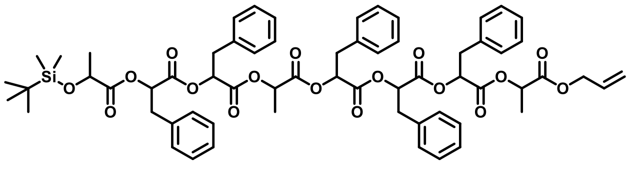
PPLLPPLP 1.1 g, 86%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 6.95 (m, 24H), 5.81 (dd, *J* = 11.3, 5.8 Hz, 1H), 5.42 – 5.02 (m, 9H), 4.64 – 4.52 (m, 2H), 4.37 – 4.20 (m, 1H), 3.47 – 2.56 (m, 10H), 1.62 – 1.29 (m, 9H), 0.70 (dd, *J* = 10.8, 2.1 Hz, 9H), -0.10 – -0.46 (m, 6H).



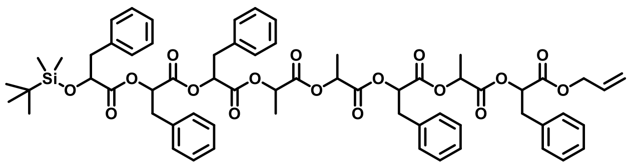
PLPPLPLL 1.5 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 – 6.87 (m, 21H), 5.97 – 5.83 (m, 1H), 5.41 – 5.06 (m, 9H), 4.68 – 4.60 (m, 2H), 4.43 – 4.28 (m, 1H), 3.47 – 2.73 (m, 8H), 1.66 – 1.24 (m, 12H), 0.78 (dd, *J* = 8.0, 3.7 Hz, 9H), -0.08 – -0.35 (m, 6H).



PPLPLLLP 1.2 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 – 6.96 (m, 21H), 5.80 (ddd, *J* = 16.7, 10.9, 5.6 Hz, 1H), 5.42 – 5.08 (m, 9H), 4.67 – 4.51 (m, 2H), 4.38 – 4.17 (m, 1H), 3.45 – 2.54 (m, 8H), 1.68 – 1.22 (m, 12H), 0.70 (dt, *J* = 10.6, 1.7 Hz, 9H), -0.12 – -0.46 (m, 6H).

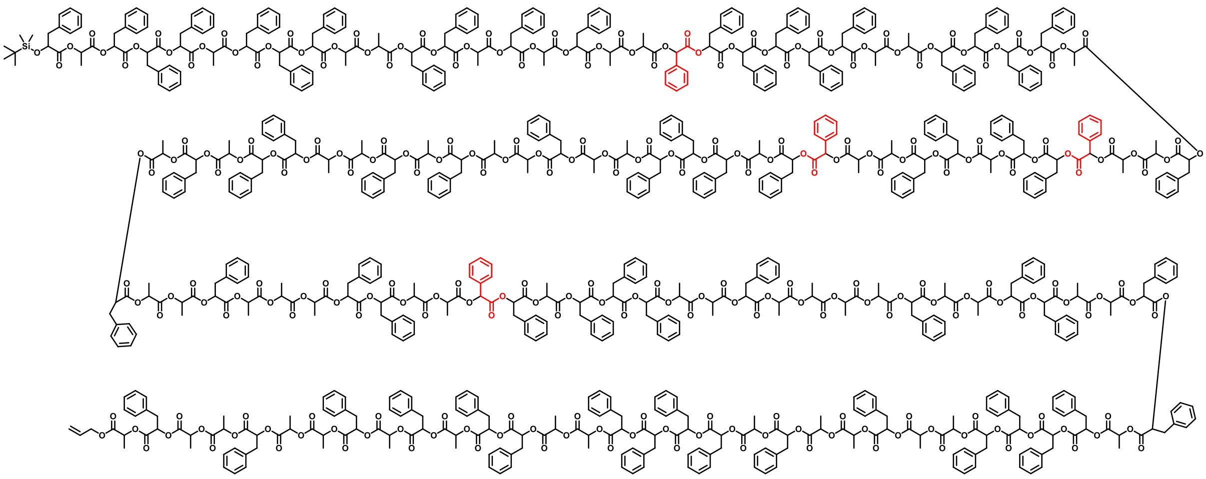


LPPLPPPL 1.3 g, 88%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 – 6.75 (m, 28H), 5.89 (ddq, *J* = 16.5, 11.2, 5.7 Hz, 1H), 5.46 – 4.99 (m, 9H), 4.63 (t, *J* = 5.4 Hz, 2H), 4.27 (dq, *J* = 19.9, 9.8 Hz, 1H), 3.46 – 2.77 (m, 10H), 1.55 – 1.16 (m, 9H), 0.86 (d, *J* = 3.2 Hz, 9H), 0.01 (d, *J* = 51.2 Hz, 6H).

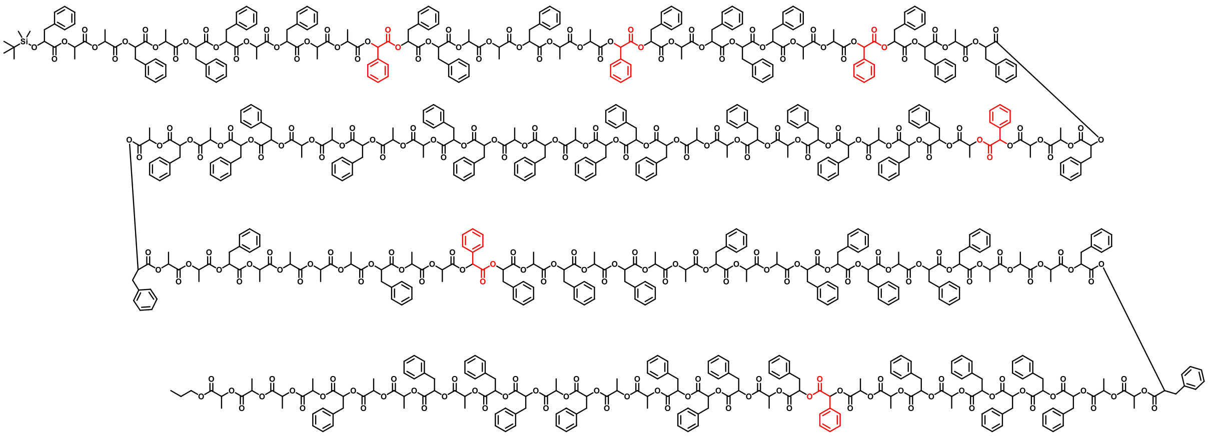


PPPLLPLP 1.5 g, 90%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.42 – 6.93 (m, 26H), 5.95 – 5.68 (m, 1H), 5.46 – 5.00 (m, 9H), 4.60 (dd, *J* = 14.6, 5.7 Hz, 2H), 4.39 – 4.16 (m, 1H), 3.46 – 2.56 (m, 10H), 1.56 – 1.22 (m, 9H), 0.71 (dd, *J* = 7.4, 5.3 Hz, 9H), -0.13 – -0.44 (m, 6H).

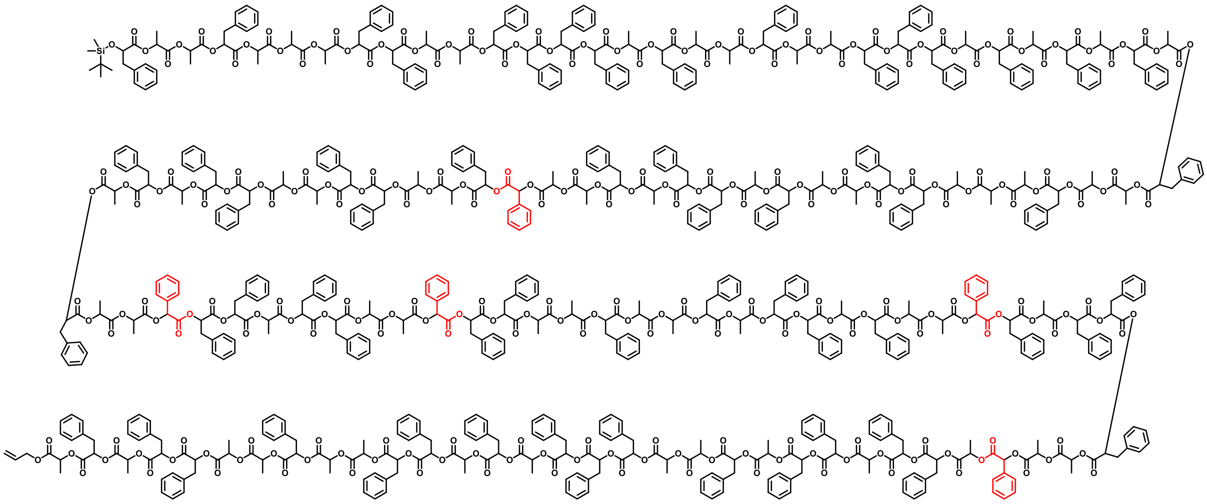
Characterization of 128-mers, 256-mers and 512-mer



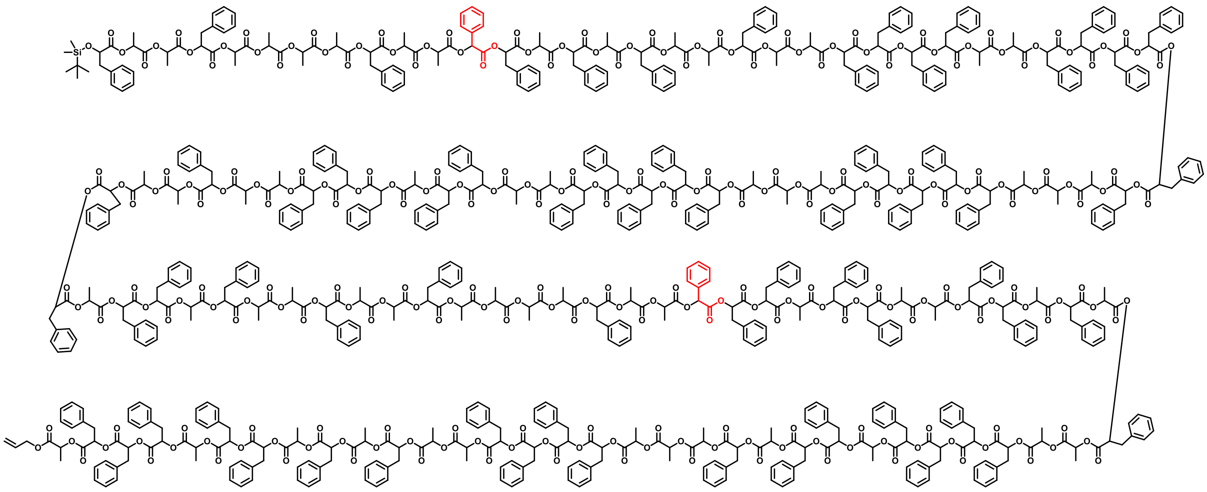
128-mer_repeat unit 1-128 1.2 g, 81%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 – 6.59 (m, 352H), 6.04 (d, J = 5.6 Hz, 5H), 5.46 – 5.03 (m, 119H), 4.71 – 4.60 (m, 2H), 4.37 (ddd, J = 12.7, 9.3, 3.4 Hz, 1H), 3.48 – 2.75 (m, 127H), 1.57 – 1.10 (m, 166H), 0.83 – 0.71 (m, 9H), -0.10 – -0.32 (m, 6H).



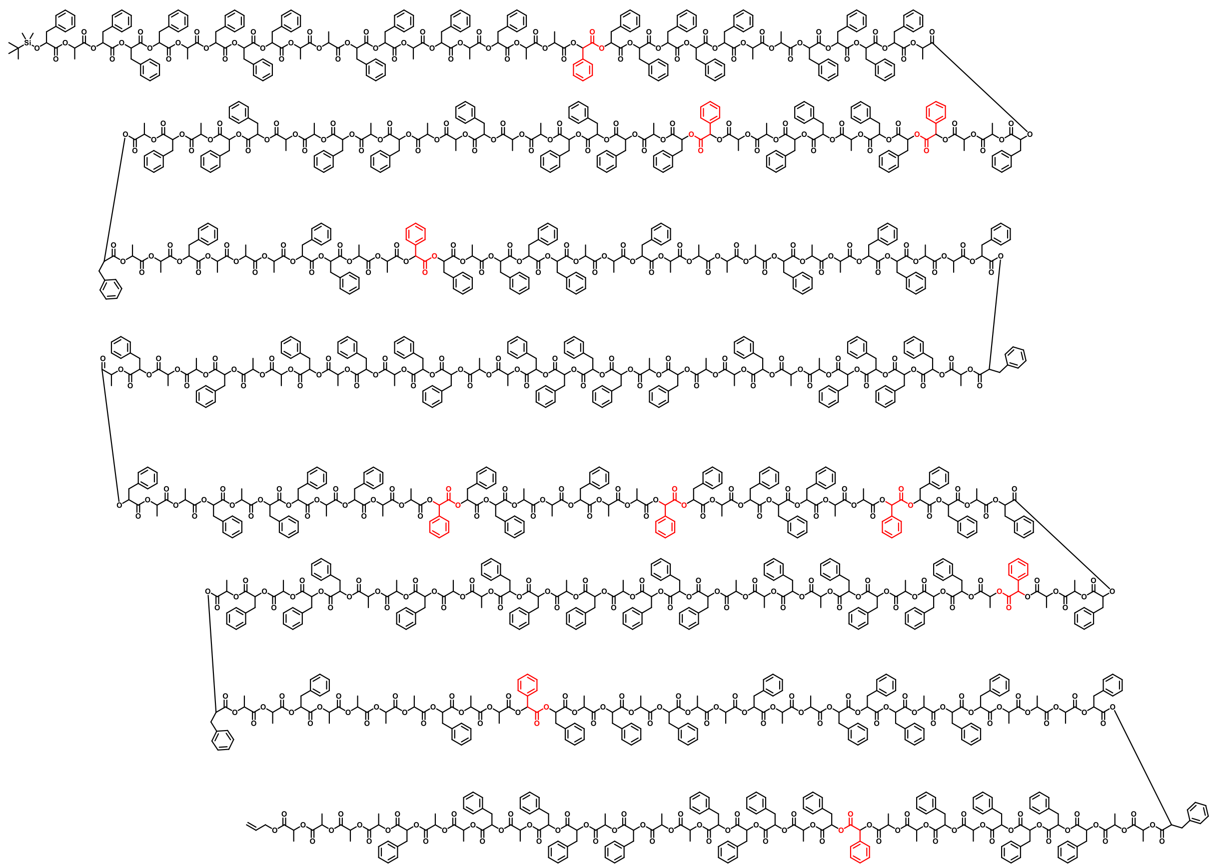
128-mer_repeat unit 129-256 1.8 g, 83%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.54 – 6.66 (m, 345H), 6.03 (d, J = 32.0 Hz, 6H), 5.45 – 5.06 (m, 108H), 4.64 (d, J = 5.8 Hz, 3H), 4.45 – 4.36 (m, 2H), 3.18 (d, J = 10.9 Hz, 108H), 1.46 (s, 208H), 0.84 – 0.74 (m, 9H), -0.06 – -0.29 (m, 6H).



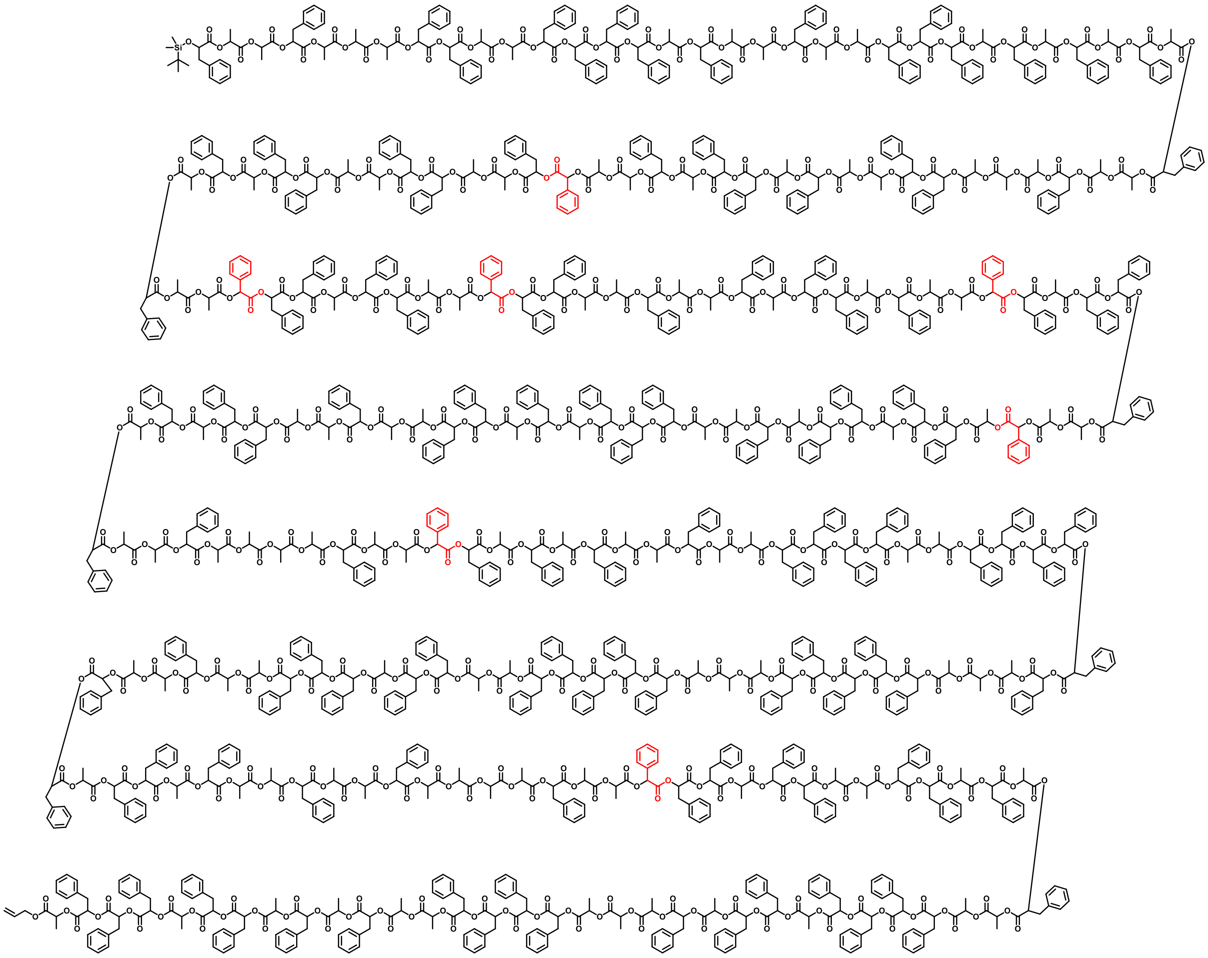
128-mer_repeat unit 257-384 1.3 g, 79%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.59 – 6.54 (m, 348H), 6.09 – 5.79 (m, 6H), 5.55 – 4.88 (m, 125H), 4.63 (s, 2H), 4.39 (d, J = 7.4 Hz, 1H), 3.46 – 2.76 (m, 126H), 1.57 – 1.14 (m, 191H), 0.77 (d, J = 6.2 Hz, 9H), -0.06 – -0.32 (m, 6H).



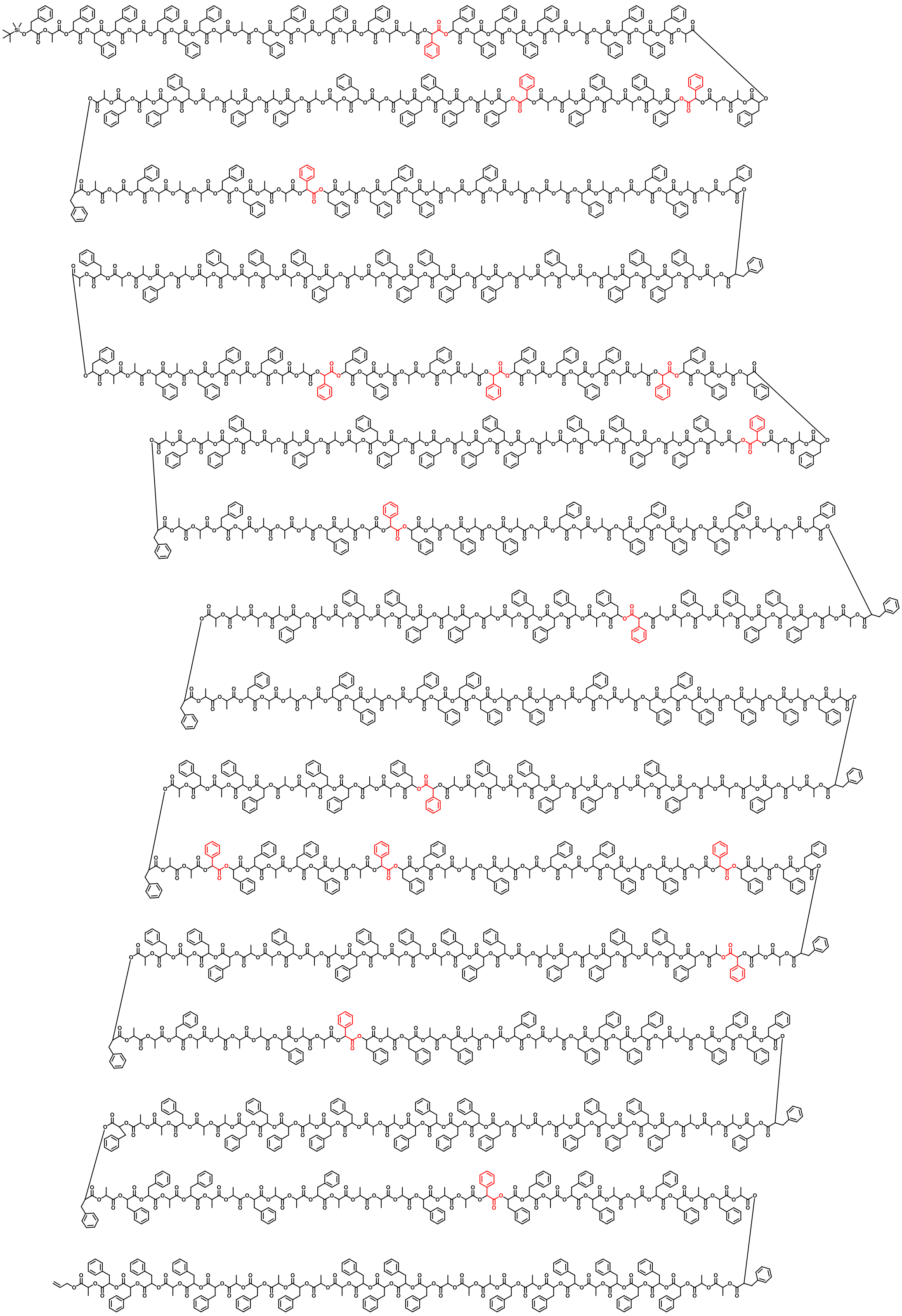
128-mer_repeat unit 385-512 1.5 g, 80%; [1]H NMR (500 MHz, $CDCl_3$) δ 7.50 – 6.52 (m, 410H), 6.01 (d, J = 21.6 Hz, 2H), 5.20 (d, J = 74.0 Hz, 136H), 4.63 (s, 3H), 4.38 (s, 2H), 3.48 – 2.70 (m, 147H), 1.55 – 1.11 (m, 190H), 0.77 (d, J = 6.1 Hz, 10H), 0.00 – -0.36 (m, 6H).



256-mer_repeat unit 1-256 1.5 g, 75%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.56 – 6.57 (m, 647H), 6.13 – 5.84 (m, 11H), 5.55 – 4.89 (m, 228H), 4.70 – 4.59 (m, 2H), 4.43 – 4.27 (m, 1H), 3.50 – 2.71 (m, 231H), 1.79 – 1.02 (m, 380H), 0.78 (dd, J = 7.3, 4.2 Hz, 9H), -0.05 – -0.37 (m, 6H).

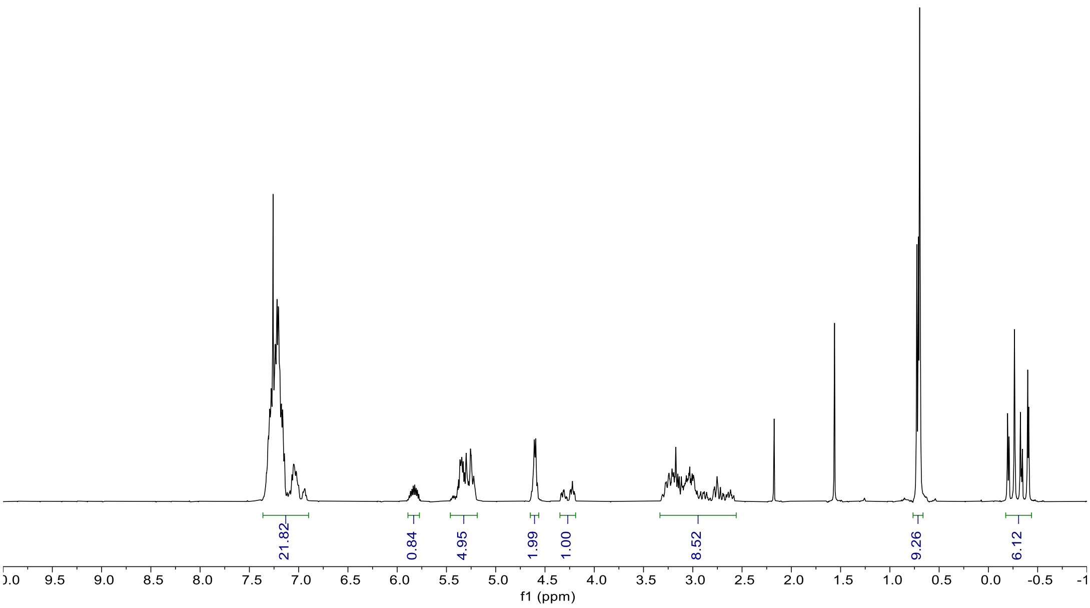


256-mer_repeat unit 257-512 1.3 g, 75%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.60 – 6.57 (m, 799H), 5.98 (s, 8H), 5.53 – 4.85 (m, 277H), 4.63 (s, 3H), 4.38 (s, 1H), 3.10 (d, J = 60.2 Hz, 291H), 1.74 – 1.11 (m, 525H), 0.77 (d, J = 6.1 Hz, 11H), -0.06 – -0.34 (m, 7H).

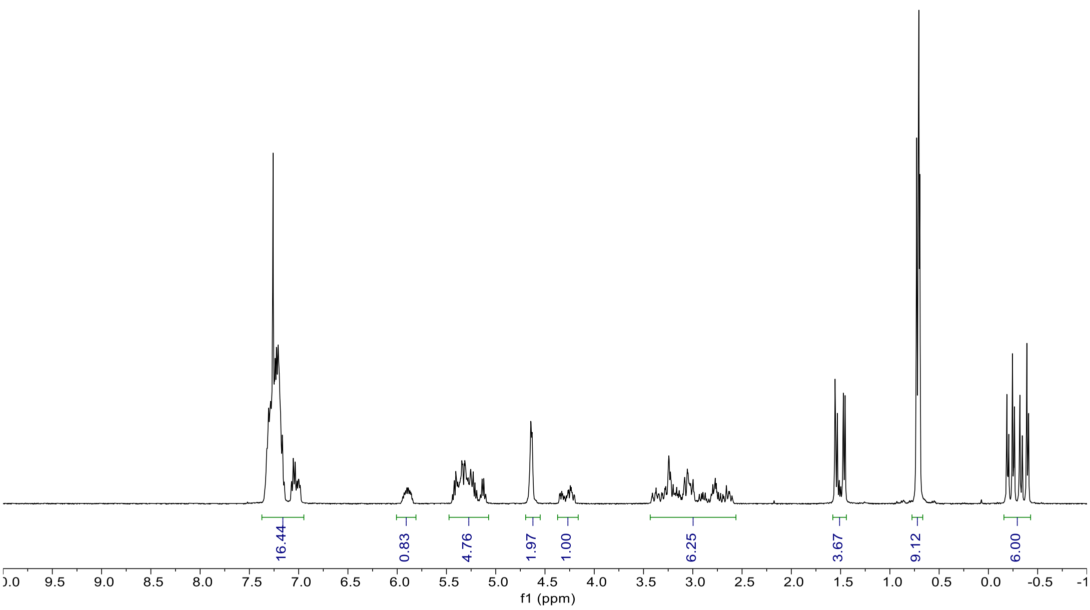


512-mer 0.2 g, 15%; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.22 (dd, J = 27.8, 14.1 Hz, 1933H), 6.01 (d, J = 22.0 Hz, 23H), 5.47 – 4.88 (m, 673H), 4.63 (s, 2H), 4.34 (s, 1H), 3.78 – 2.70 (m, 695H), 1.78 – 1.07 (m, 1180H), 0.76 (dd, J = 5.6, 3.3 Hz, 10H), -0.06 – -0.39 (m, 6H).

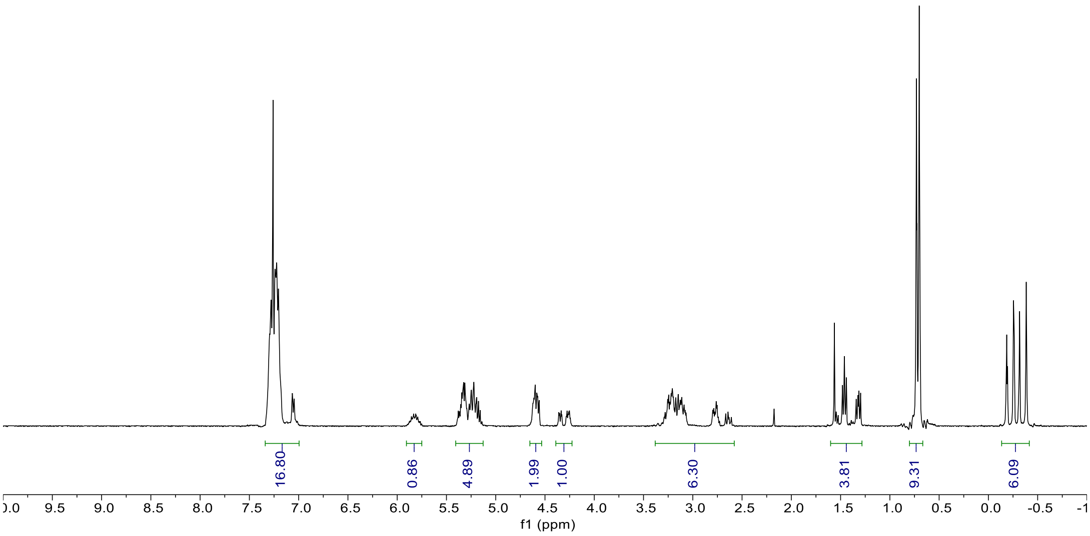
^{1}H NMR spectra of tetramers



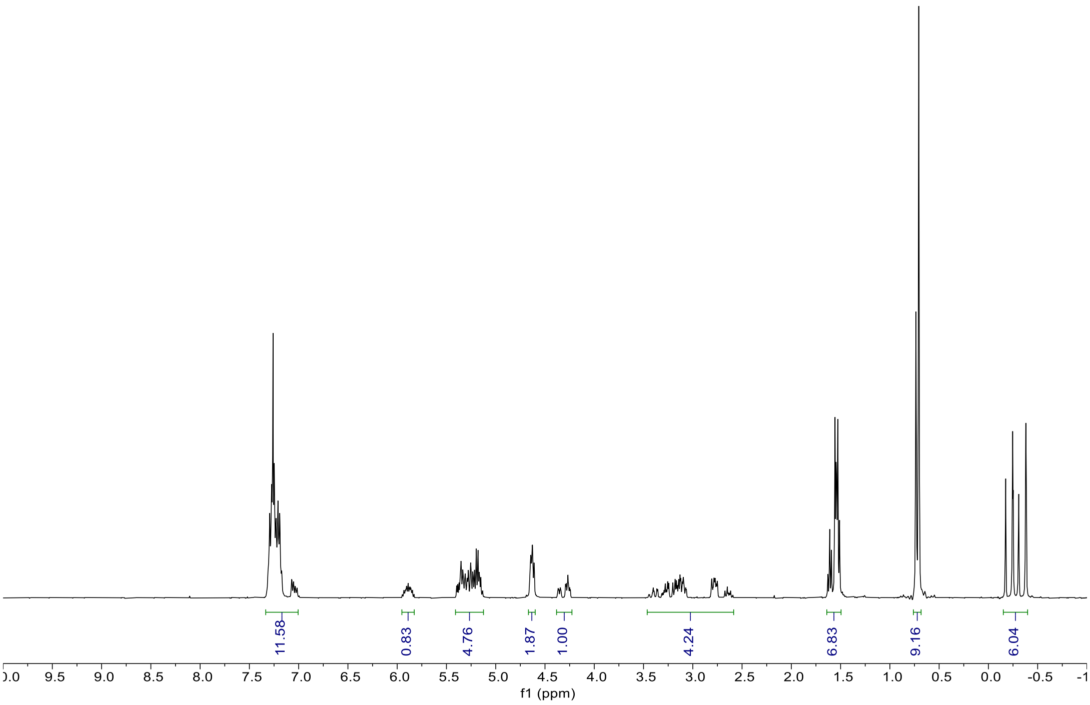
Supplementary Figure 22. ^{1}H NMR spectrum of PPPP.



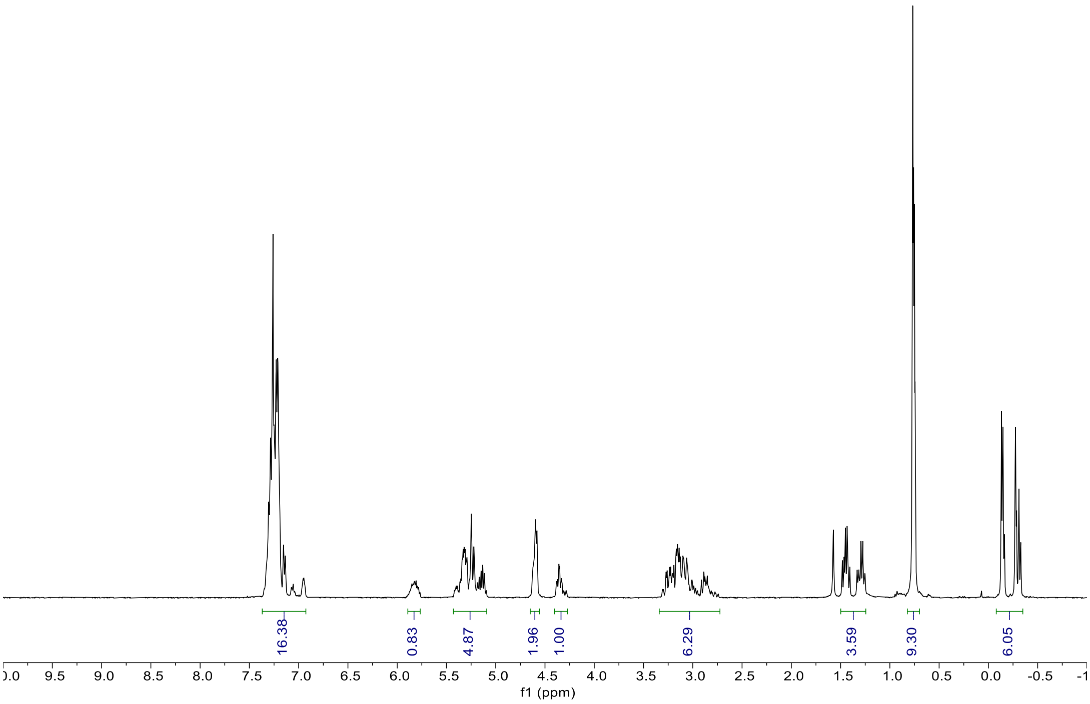
Supplementary Figure 23. ^{1}H NMR spectrum of PPPL.



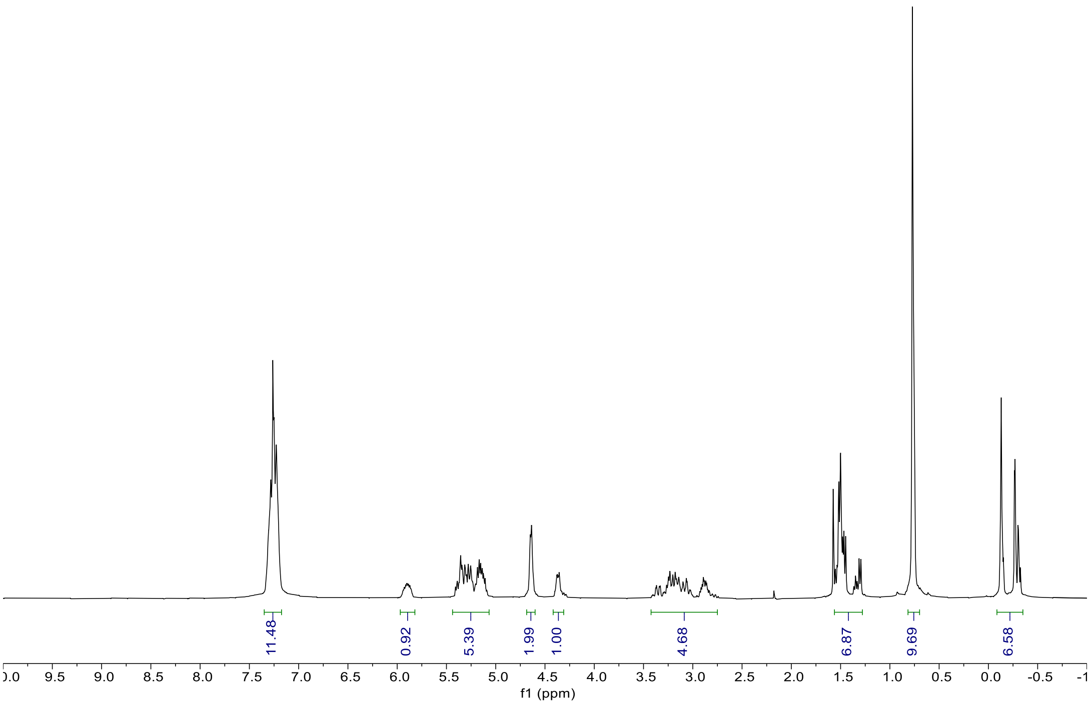
Supplementary Figure 24. ^{1}H NMR spectrum of PPLP.



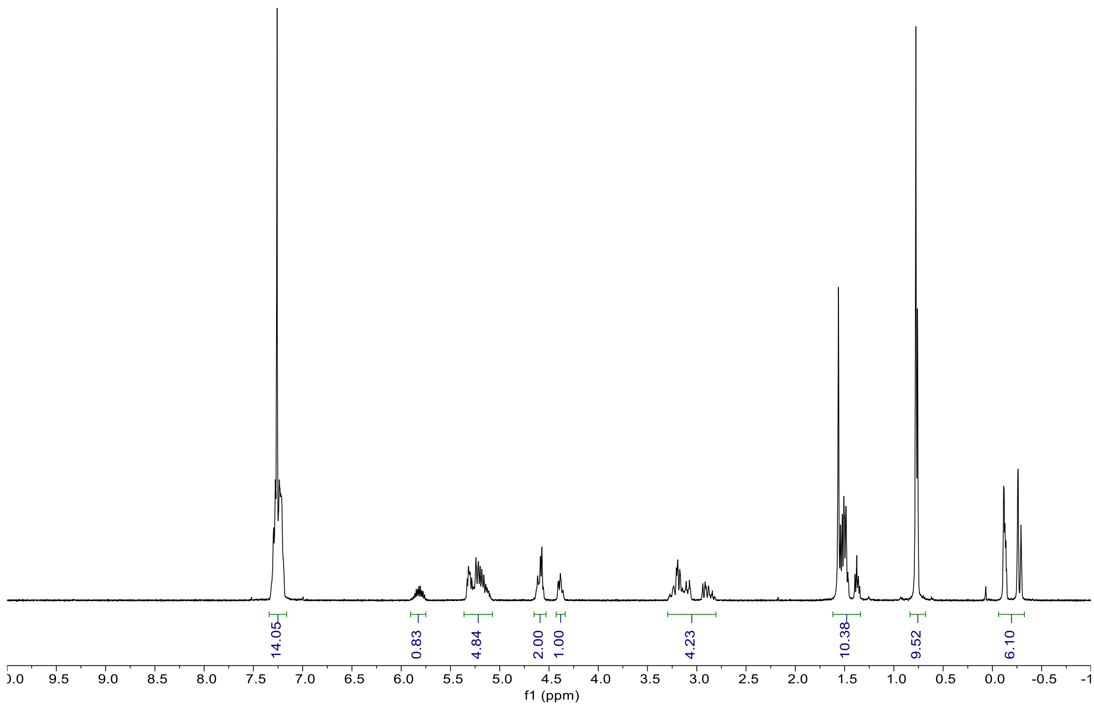
Supplementary Figure 25. ^{1}H NMR spectrum of PPLL.



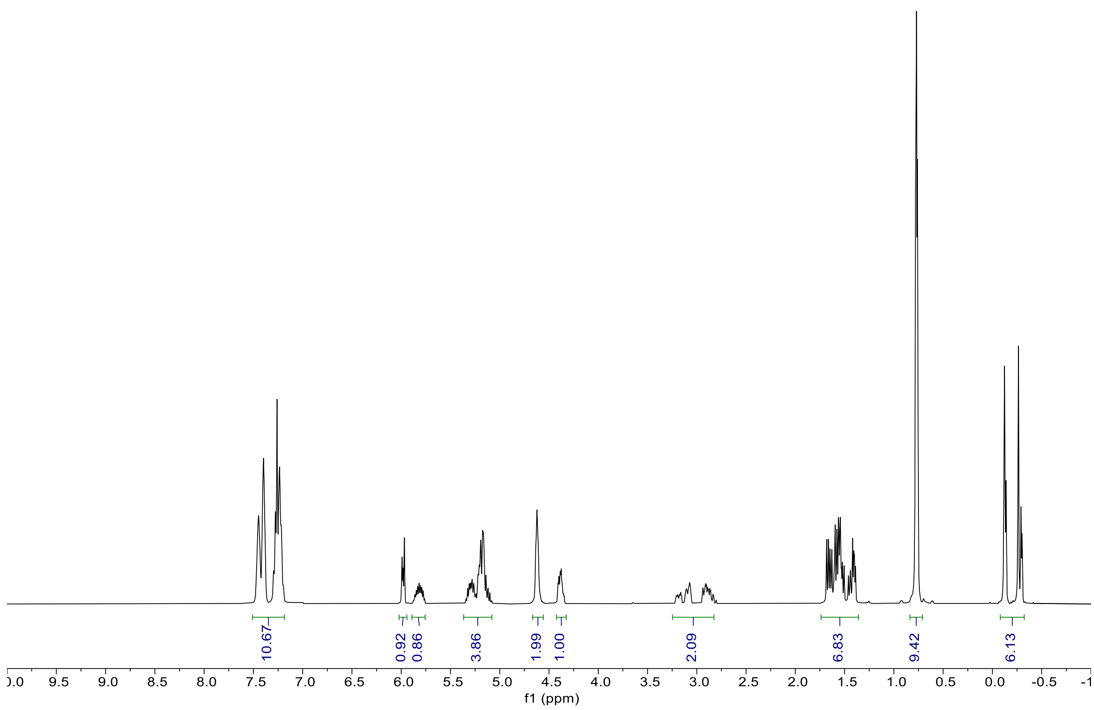
Supplementary Figure 26. ^{1}H NMR spectrum of PLPP.



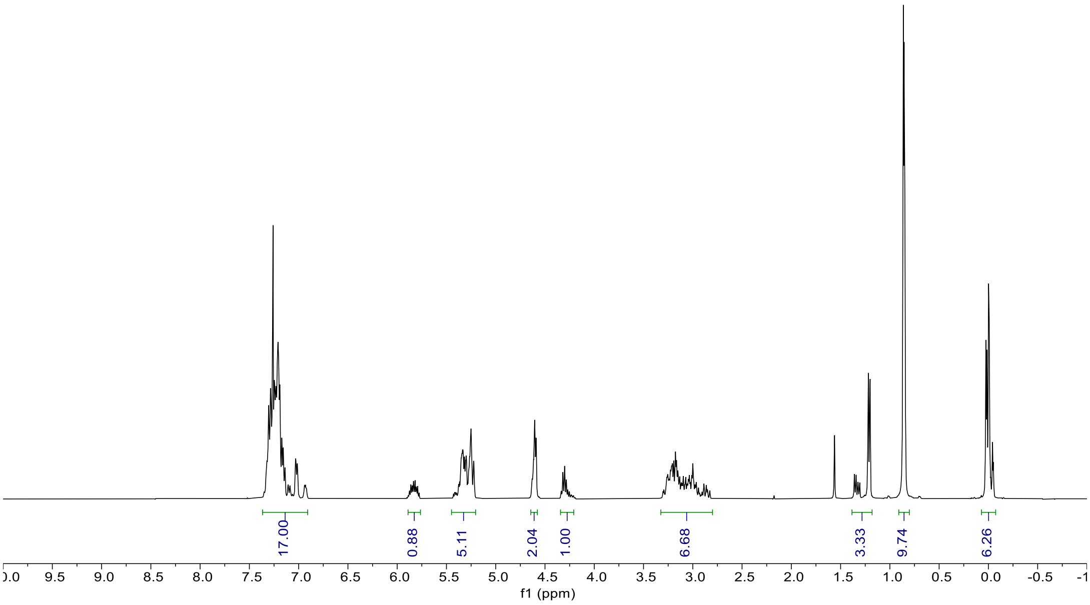
Supplementary Figure 27. ^{1}H NMR spectrum of PLPL.



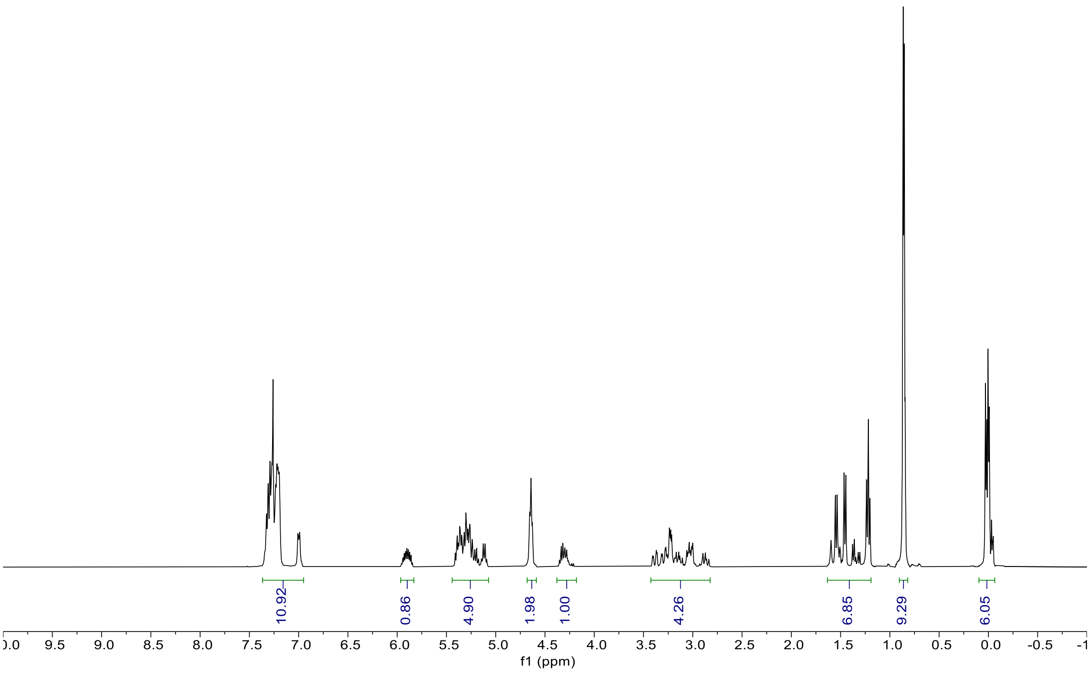
Supplementary Figure 28. ^{1}H NMR spectrum of PLLP.



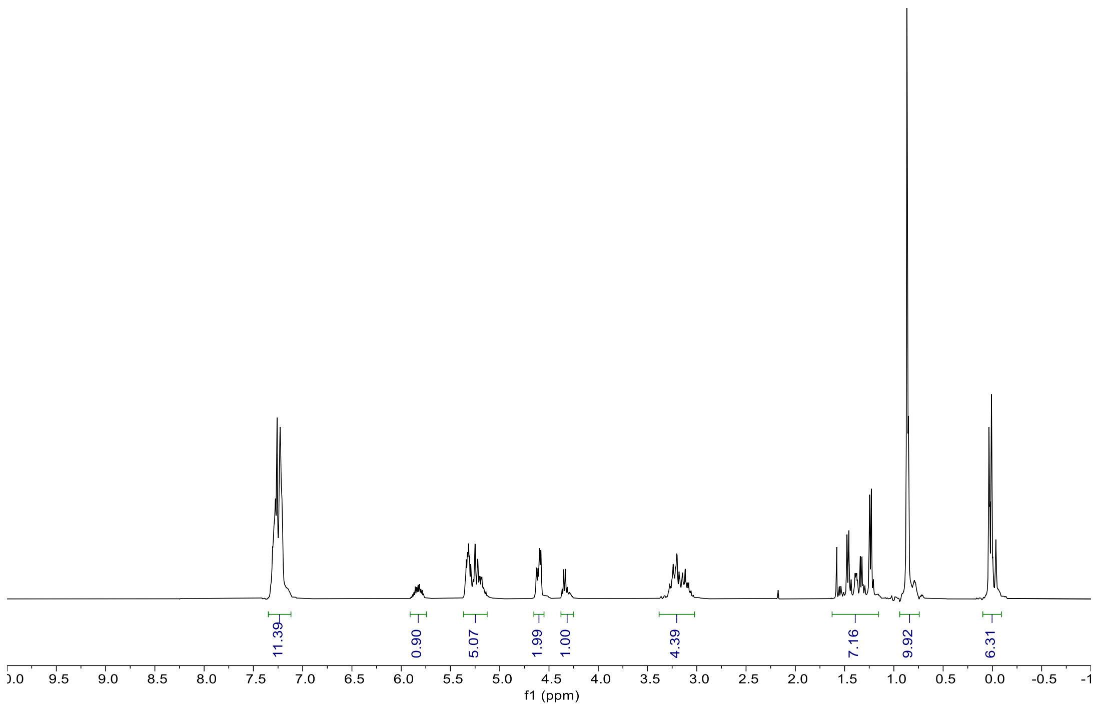
Supplementary Figure 29. ^{1}H NMR spectrum of PLLM.



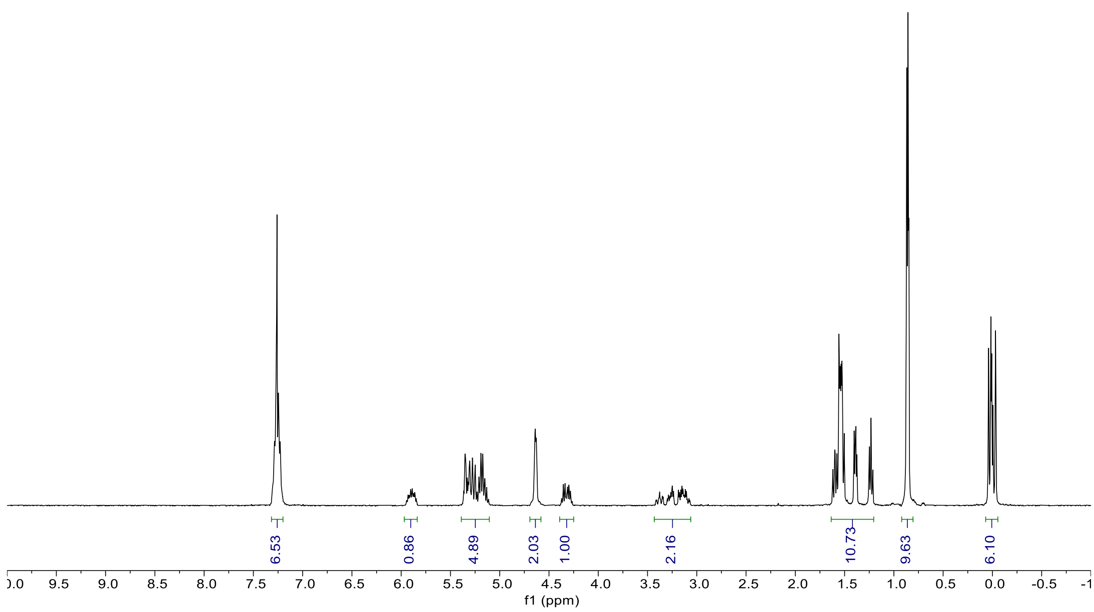
Supplementary Figure 30. ^{1}H NMR spectrum of LPPP.



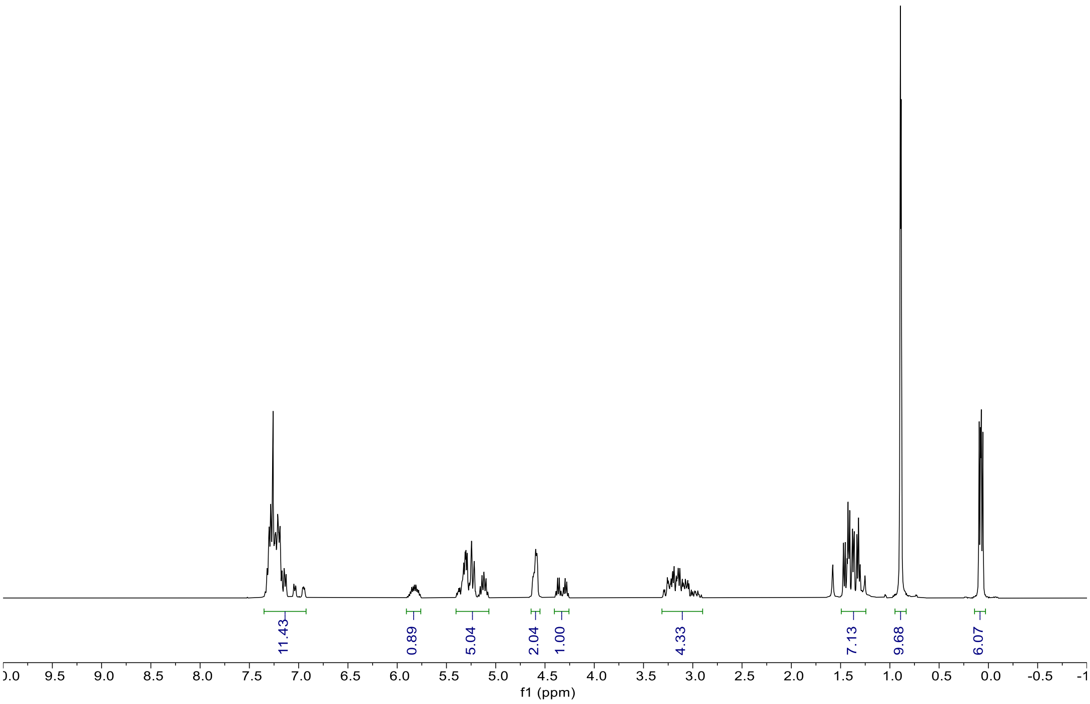
Supplementary Figure 31. ^{1}H NMR spectrum of LPPL.



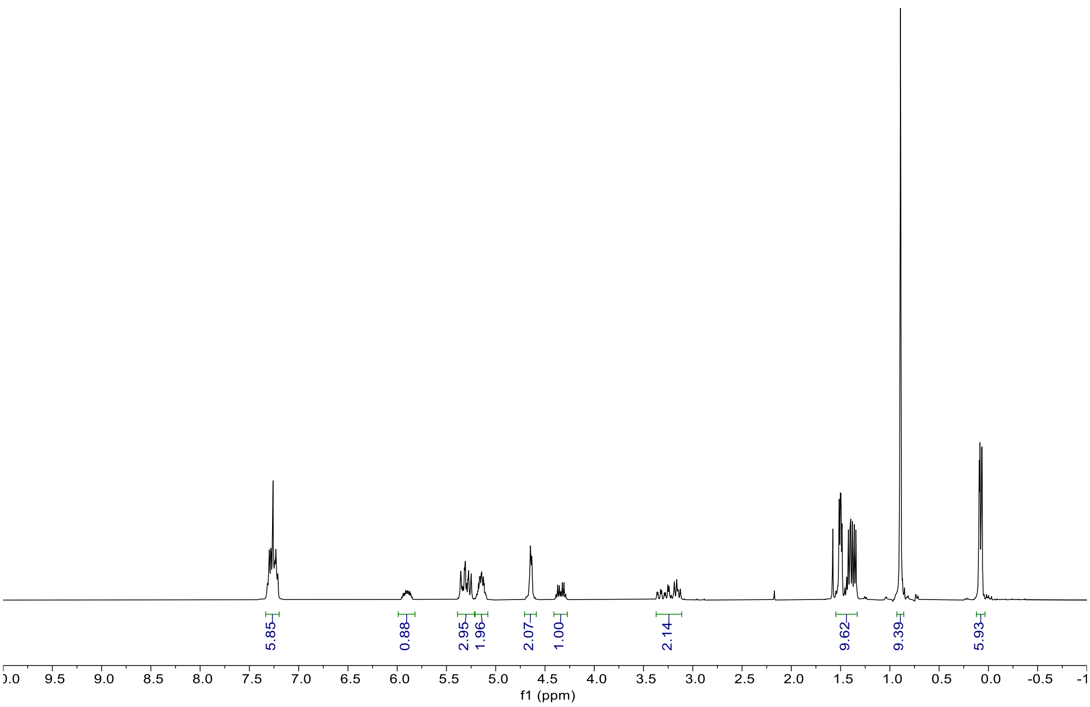
Supplementary Figure 32. ^{1}H NMR spectrum of LPLP.



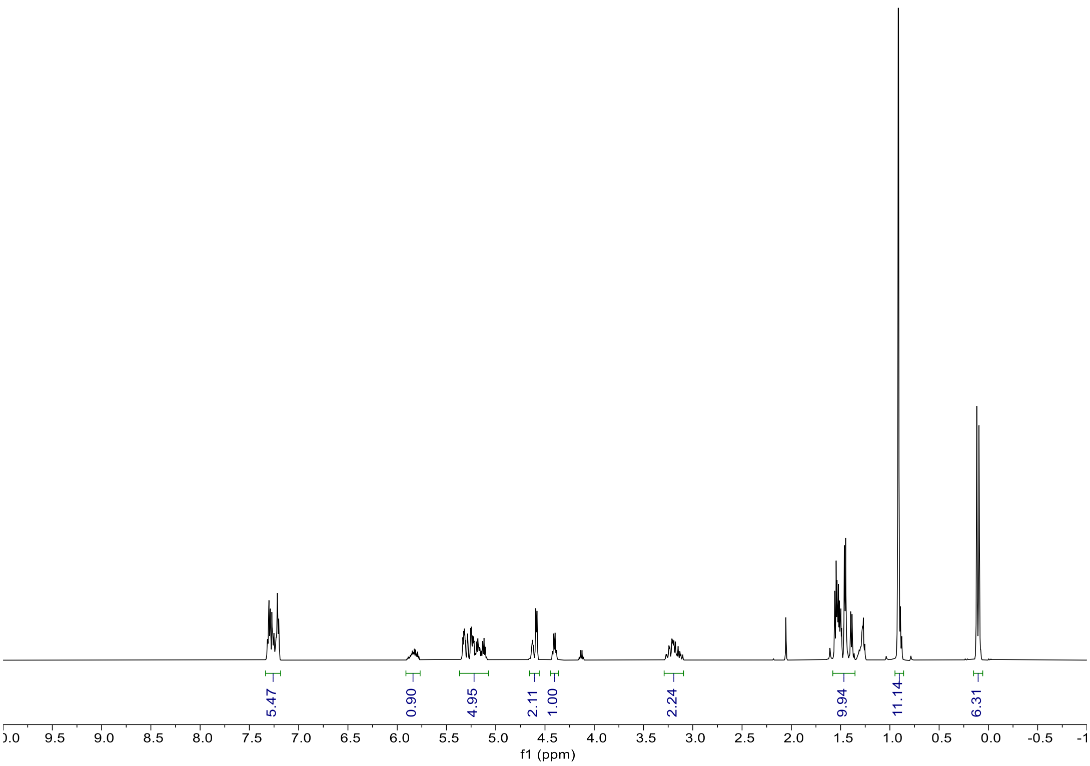
Supplementary Figure 33. ^{1}H NMR spectrum of LPLL.



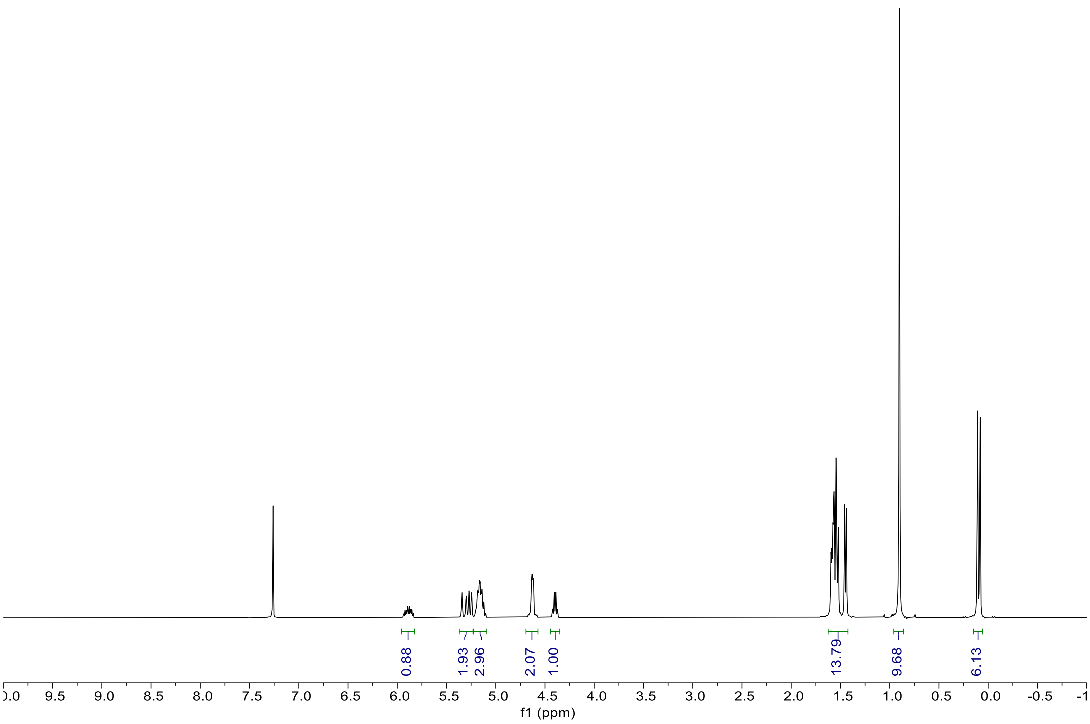
Supplementary Figure 34. ^{1}H NMR spectrum of LLPP.



Supplementary Figure 35. ^{1}H NMR spectrum of LLPL.

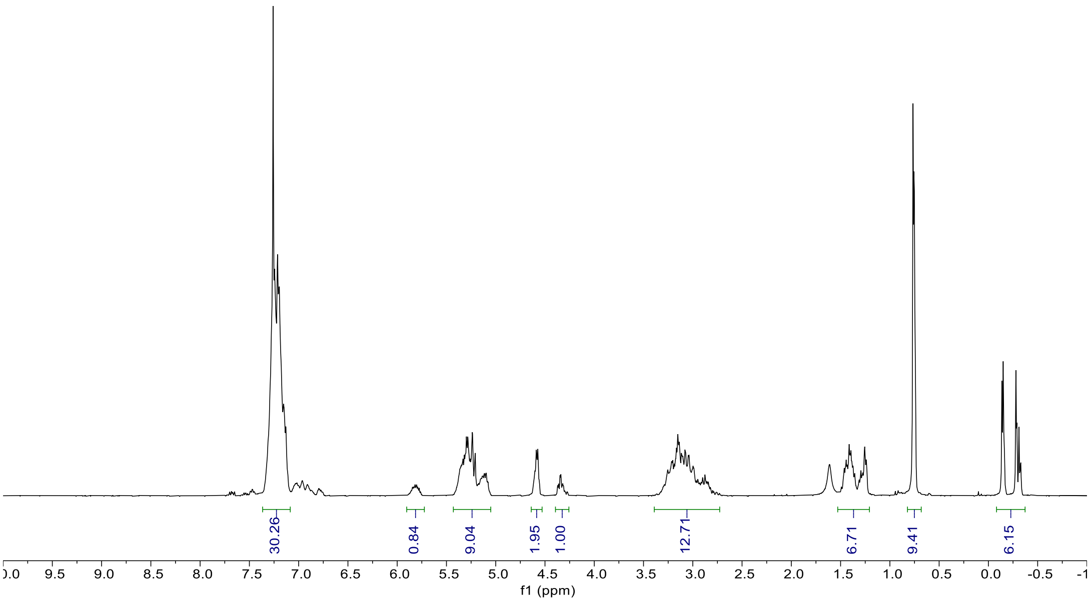


Supplementary Figure 36. ^{1}H NMR spectrum of LLLP.

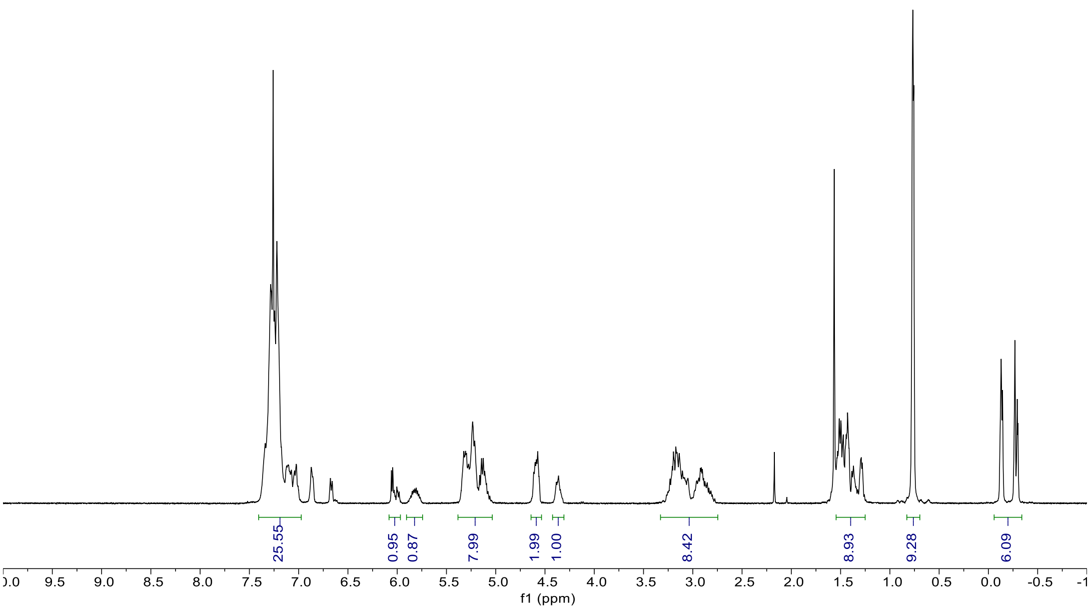


Supplementary Figure 37. ^{1}H NMR spectrum of LLLL.

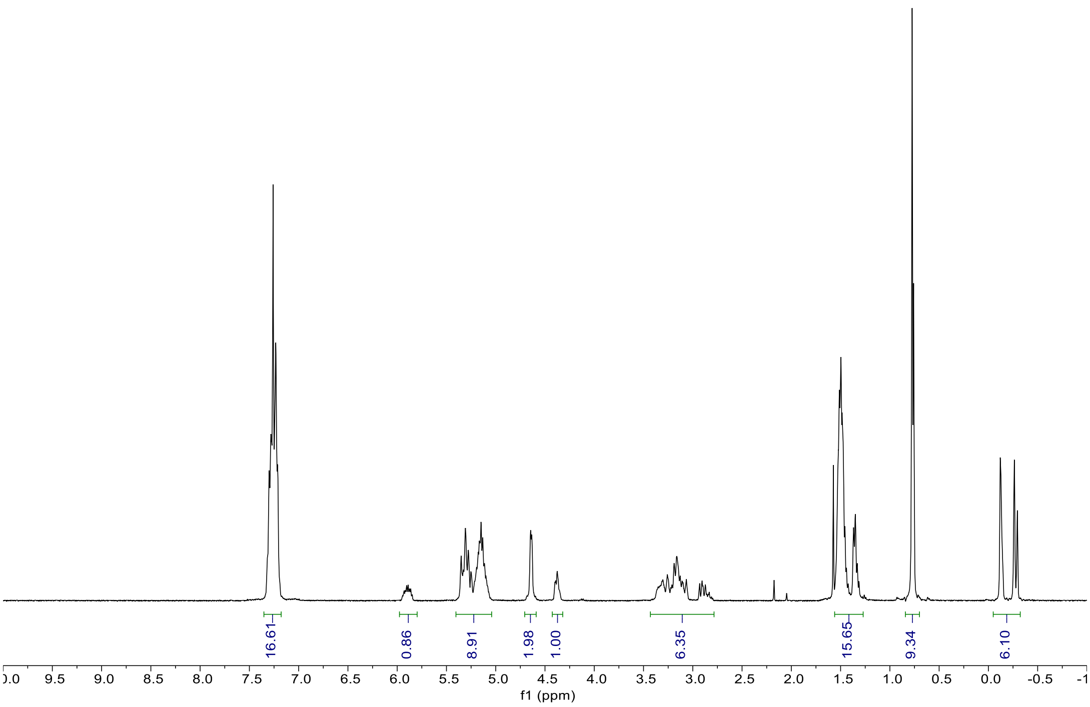
[1]H NMR spectra of octamers



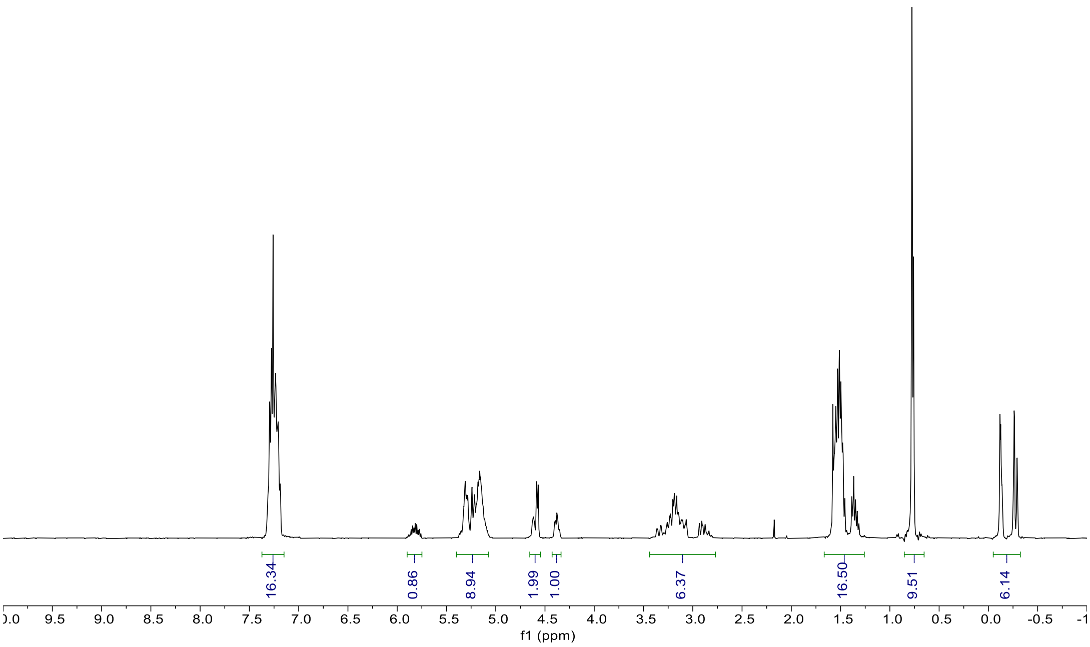
Supplementary Figure 38. ^{1}H NMR spectrum of PLPPPLPP.



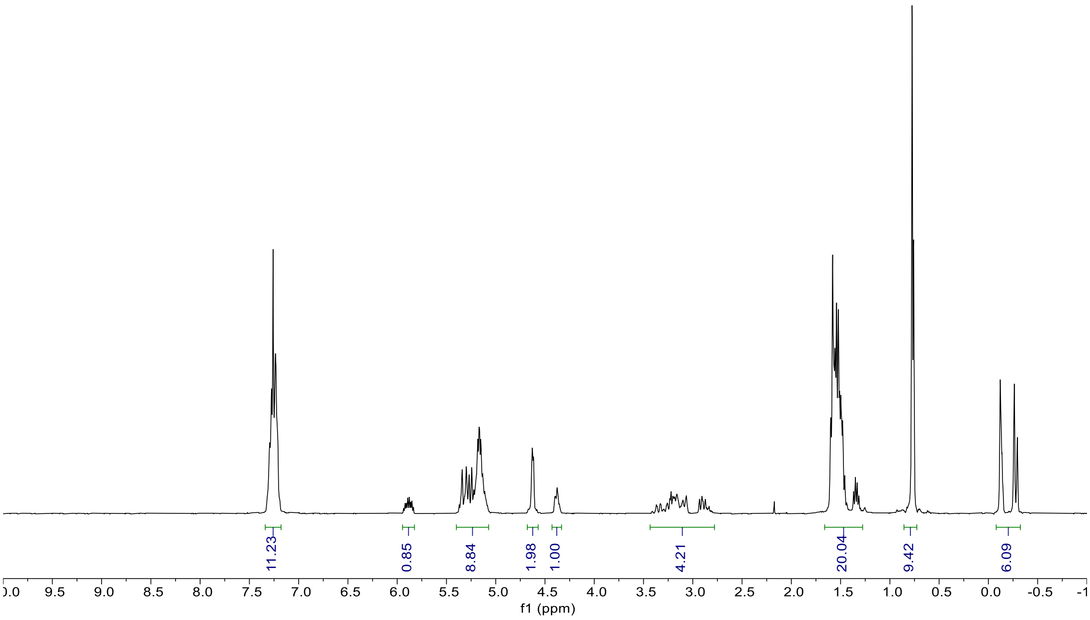
Supplementary Figure 39. ^{1}H NMR spectrum of PLLMPPLP.



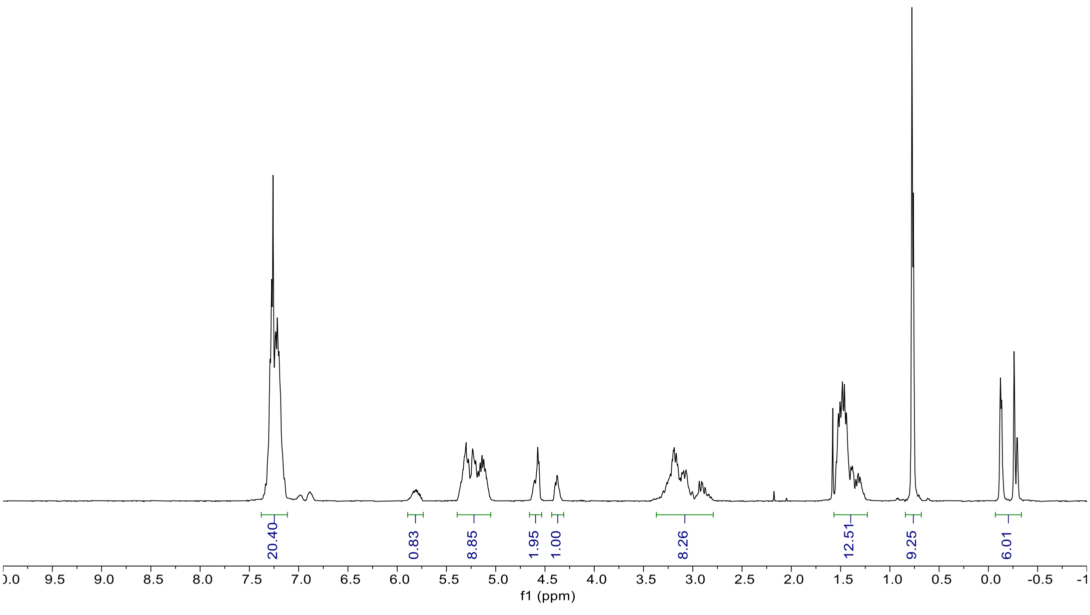
Supplementary Figure 40. ^{1}H NMR spectrum of PLLPLLPL.



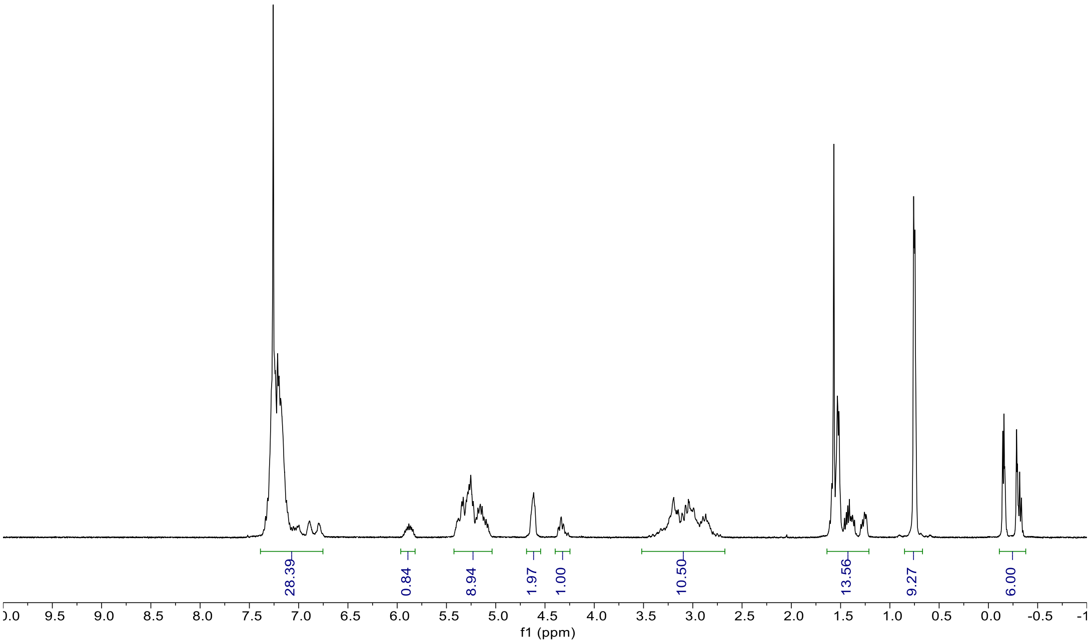
Supplementary Figure 41. ^{1}H NMR spectrum of PLLPLLLP.



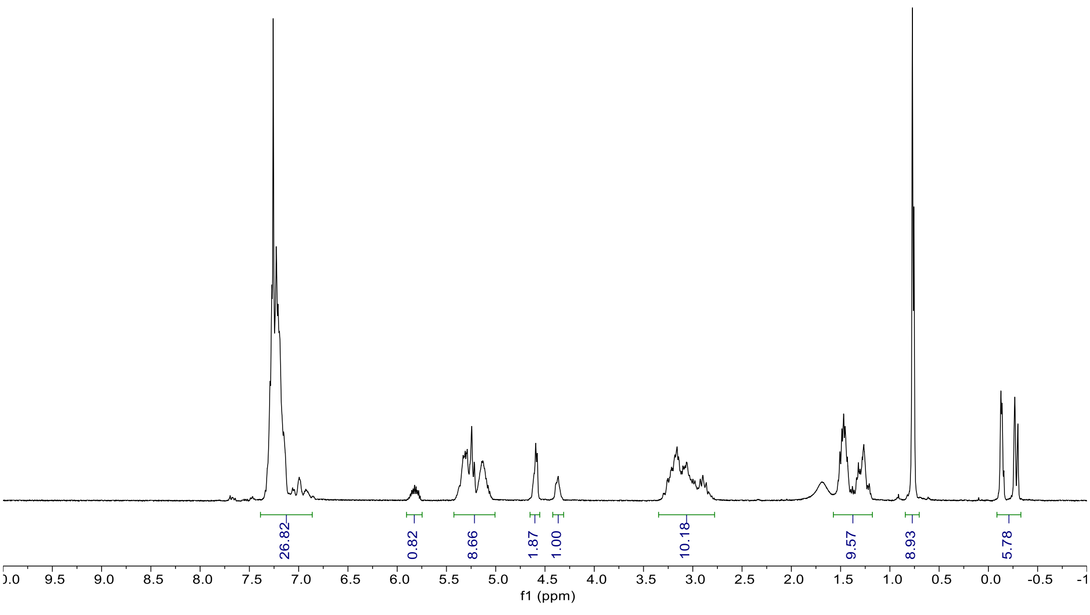
Supplementary Figure 42. ^{1}H NMR spectrum of PLLPLLLL.



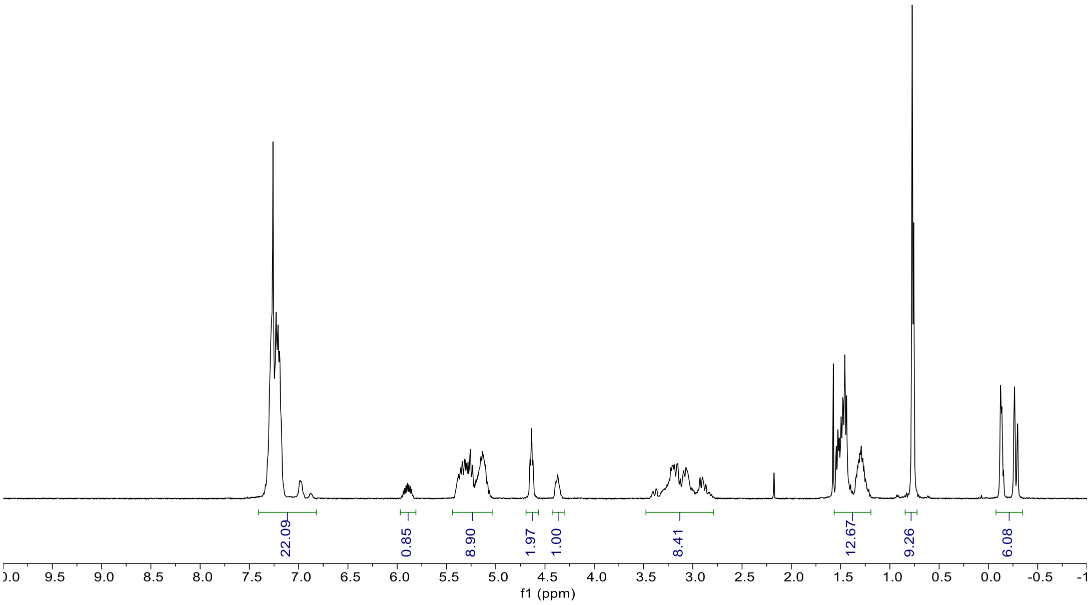
Supplementary Figure 43. ^{1}H NMR spectrum of PLLPPLLP.



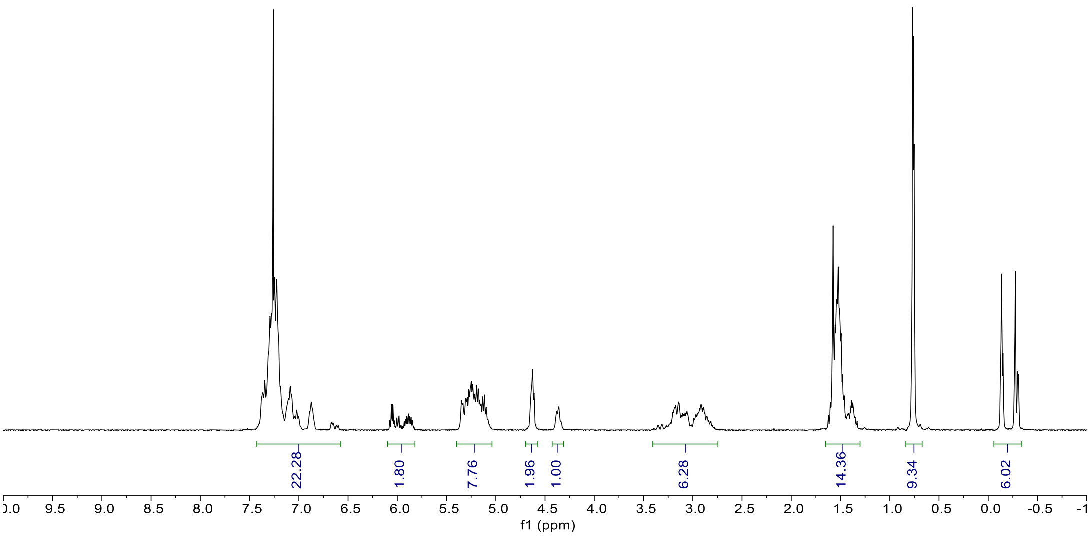
Supplementary Figure 44. ^{1}H NMR spectrum of PLPPPPLL.



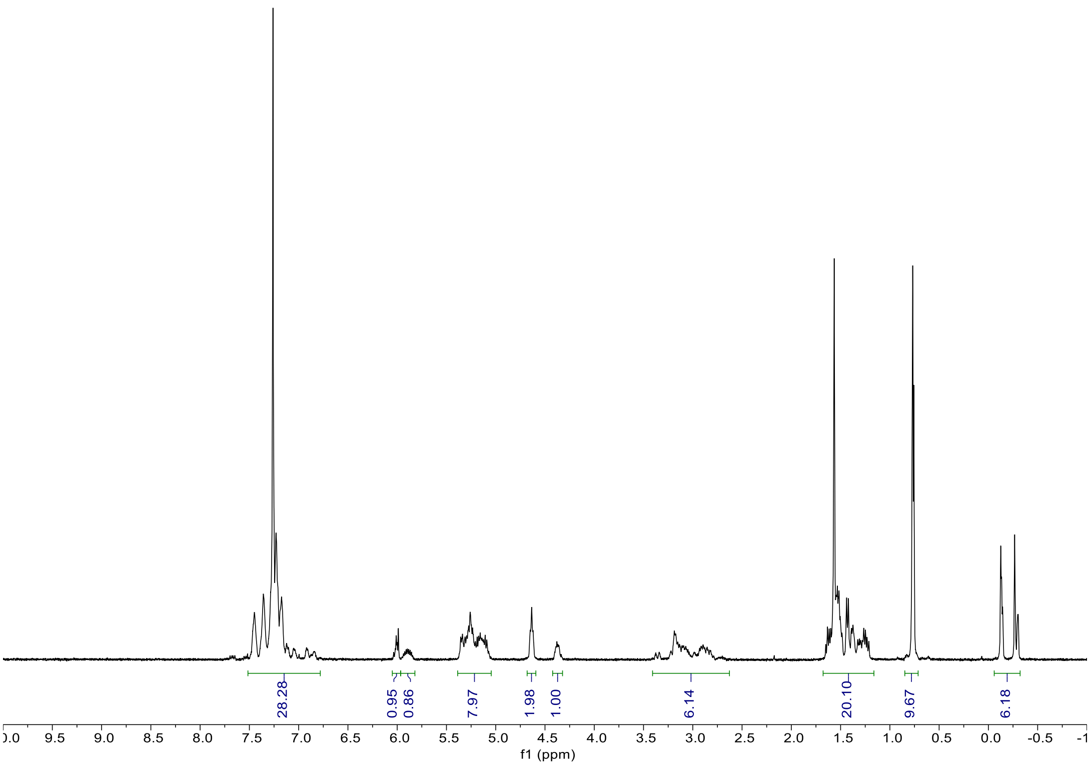
Supplementary Figure 45. ^{1}H NMR spectrum of PLLPLPPP.



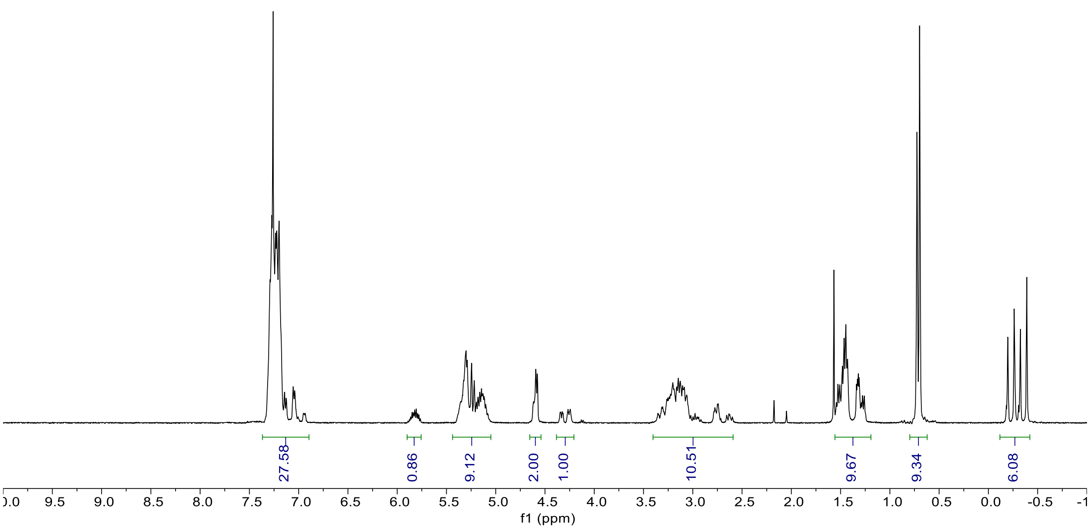
Supplementary Figure 46. ^{1}H NMR spectrum of PLLPLPPL.



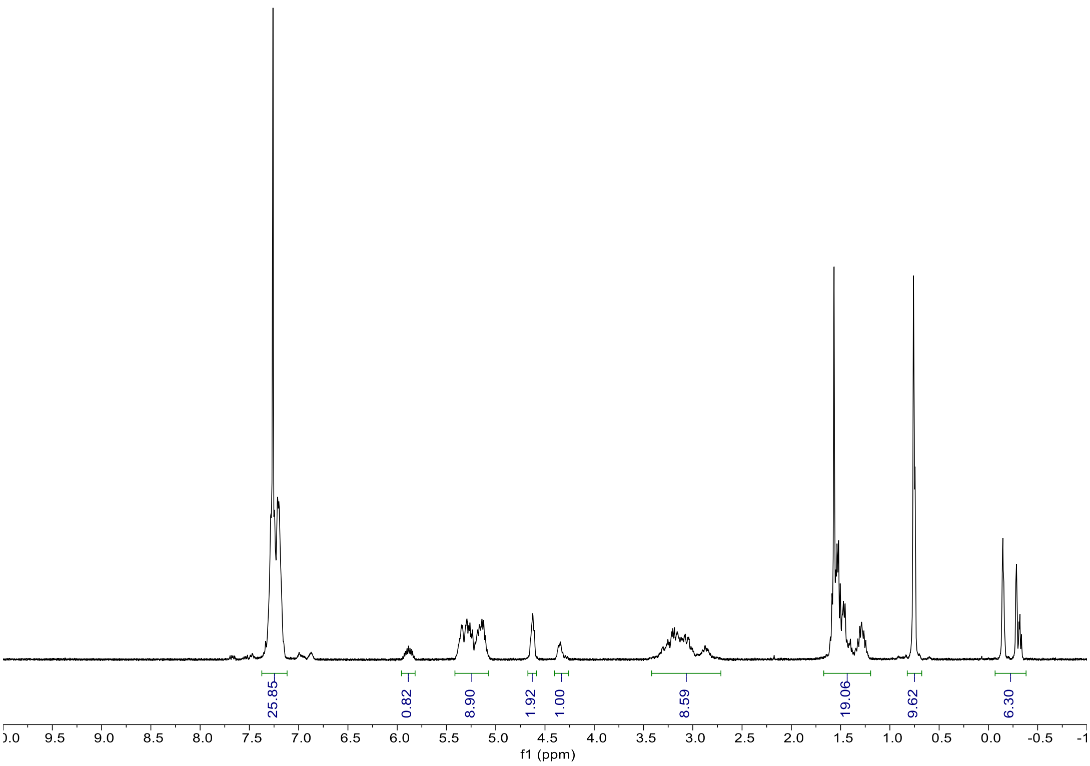
Supplementary Figure 47. ^{1}H NMR spectrum of PLLMPPLL.



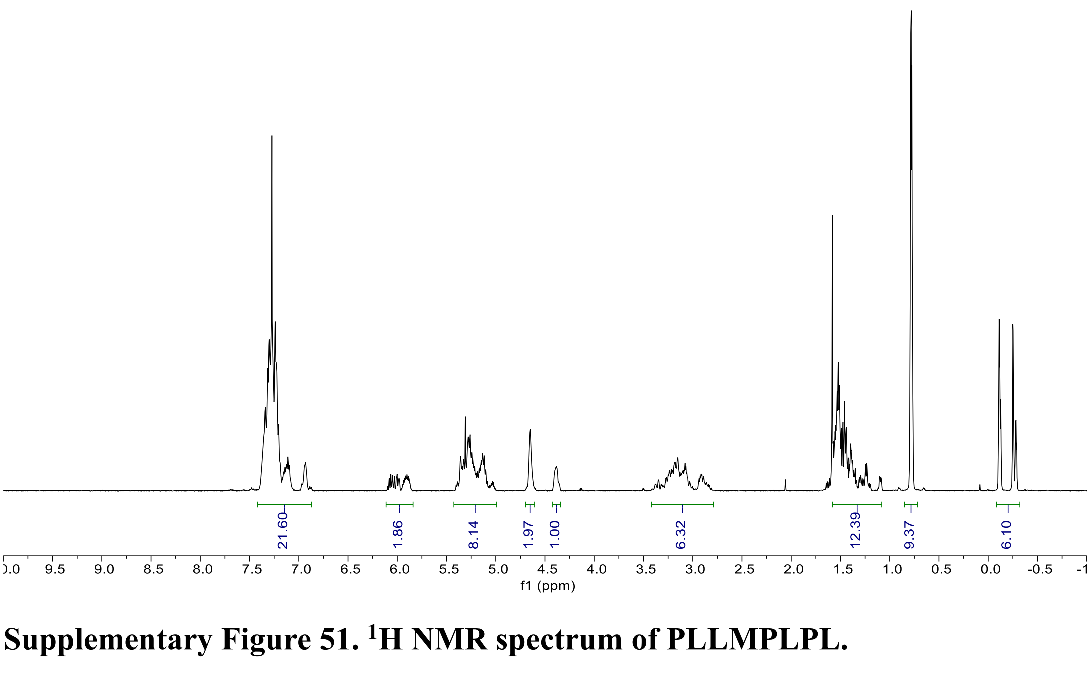
Supplementary Figure 48. ^{1}H NMR spectrum of PLLMLPPL.



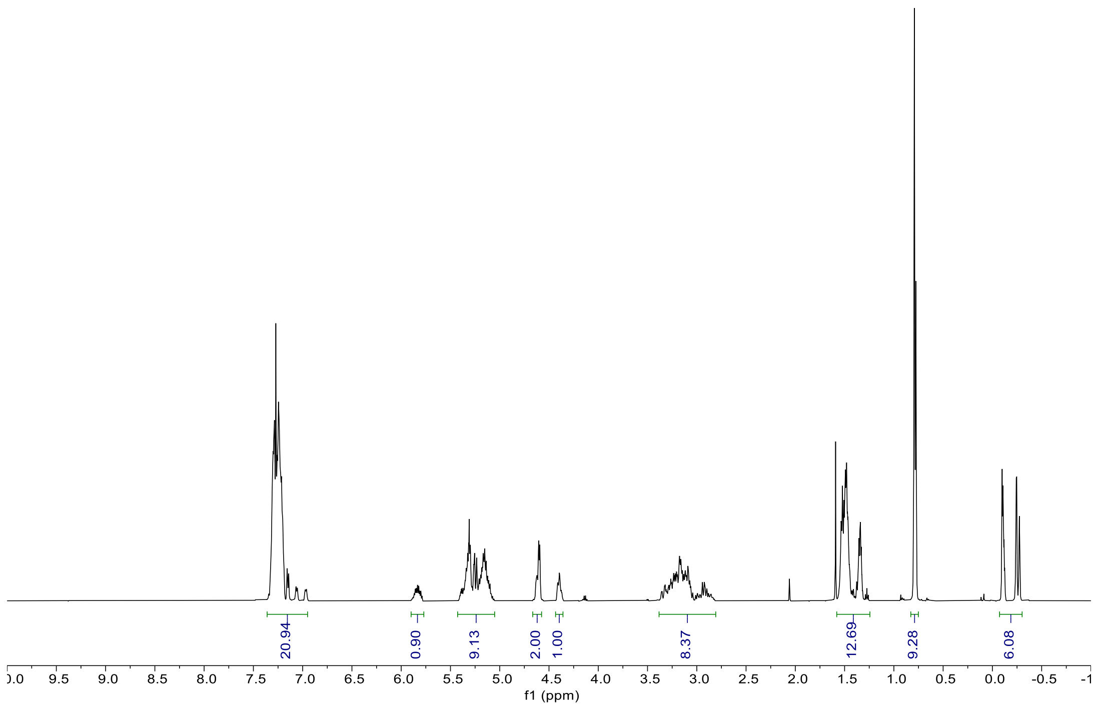
Supplementary Figure 49. ^{1}H NMR spectrum of PPLPLLPP.



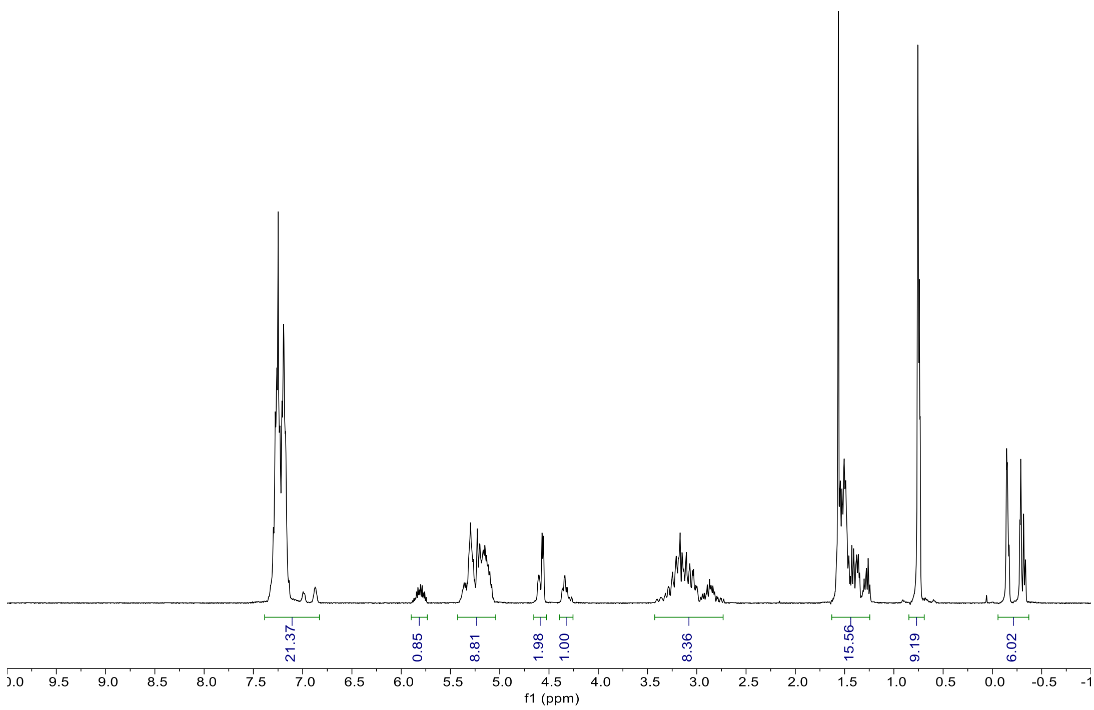
Supplementary Figure 50. ^{1}H NMR spectrum of PLPLPPLL.



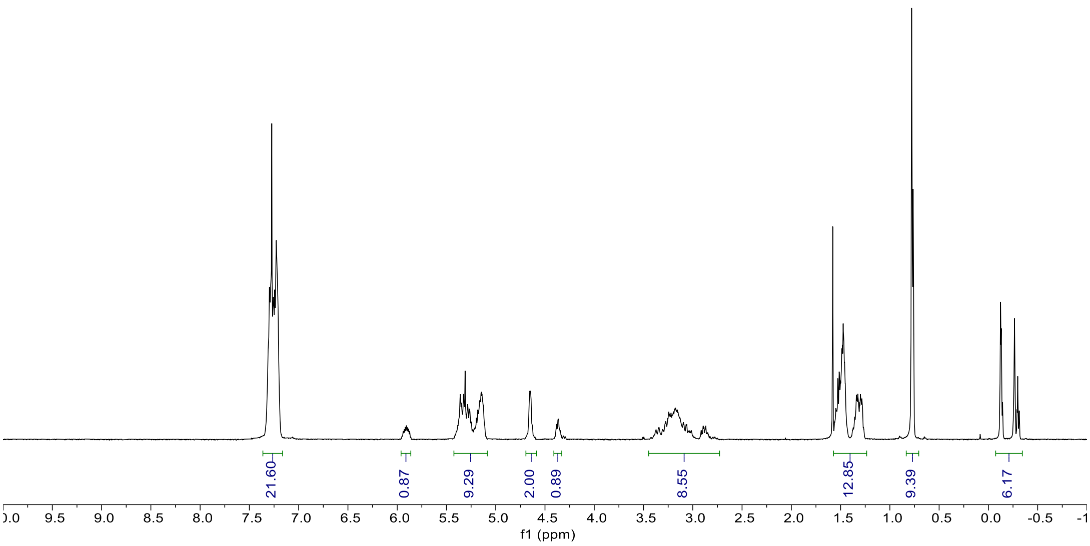
Supplementary Figure 51. ^{1}H NMR spectrum of PLLMPLPL.



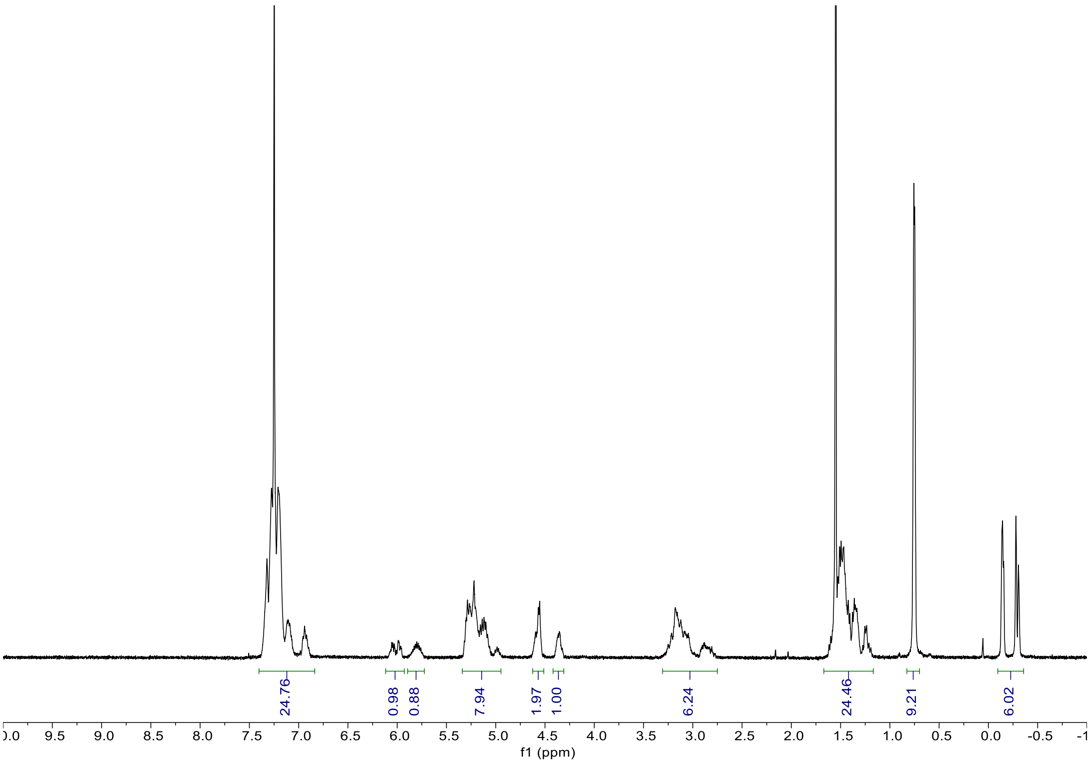
Supplementary Figure 52. ^{1}H NMR spectrum of PLLPLLPP.



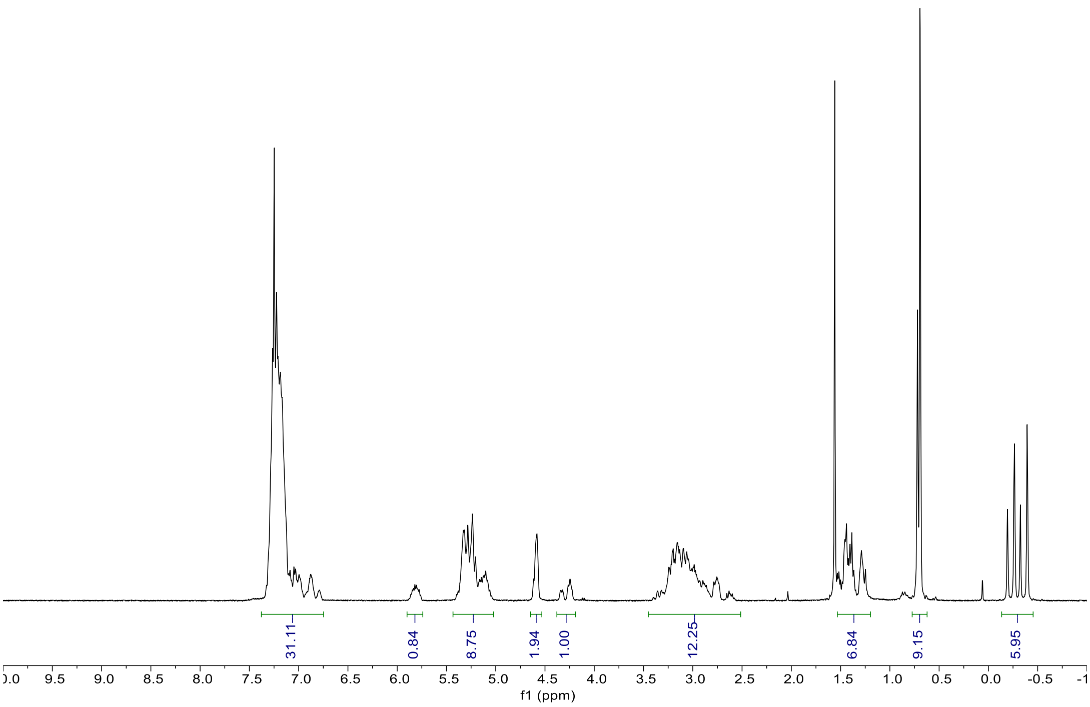
Supplementary Figure 53. ^{1}H NMR spectrum of PLPPLLLP.



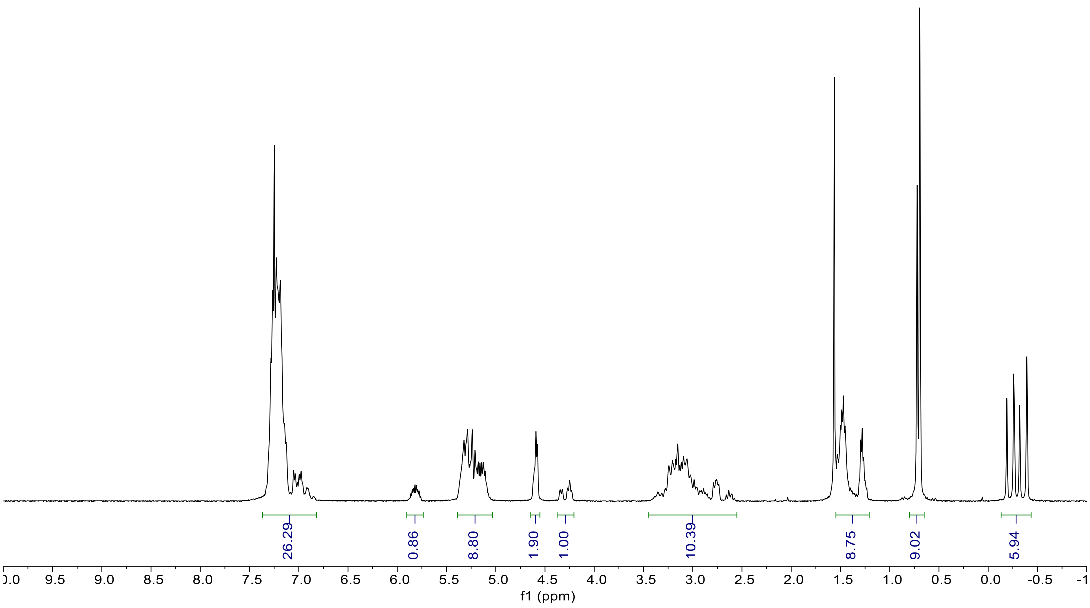
Supplementary Figure 54. ^{1}H NMR spectrum of PLPLPLPL.



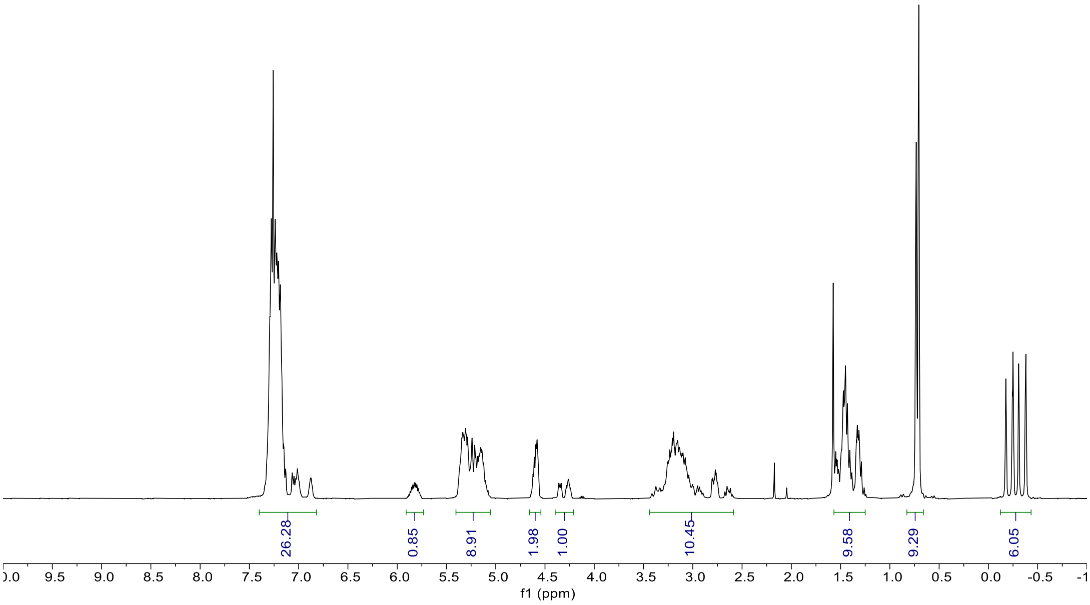
Supplementary Figure 55. ^{1}H NMR spectrum of PLLMPLLP.



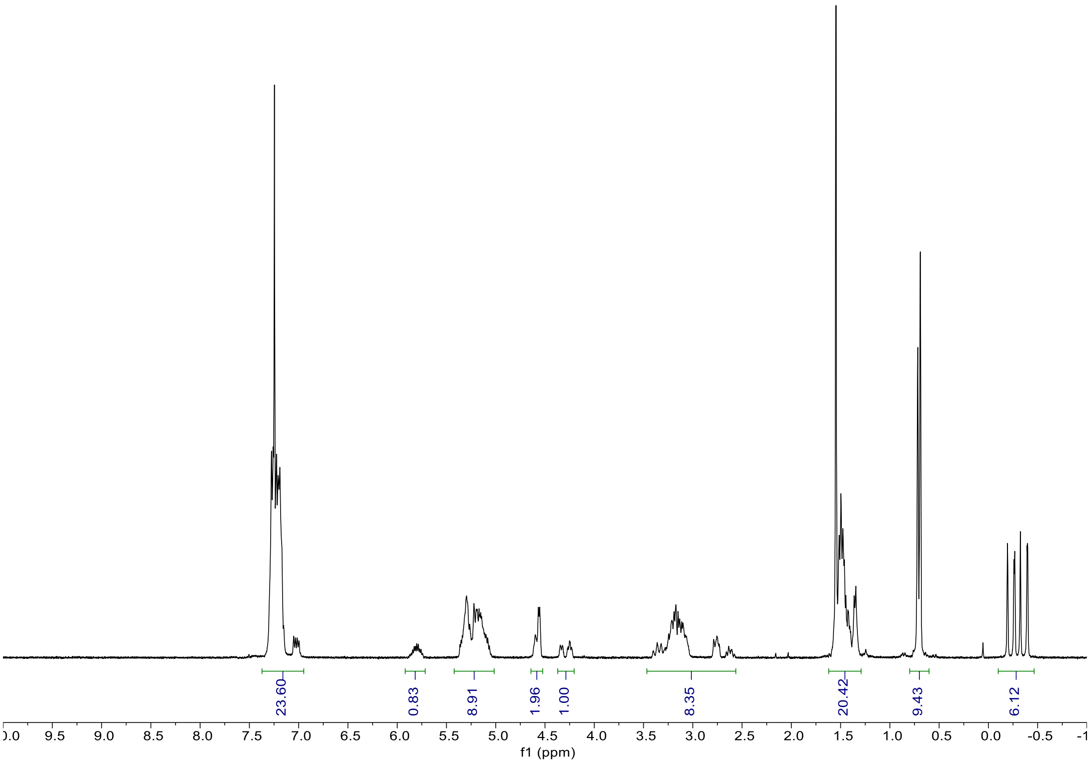
Supplementary Figure 56. ^{1}H NMR spectrum of PPLLPPPP.



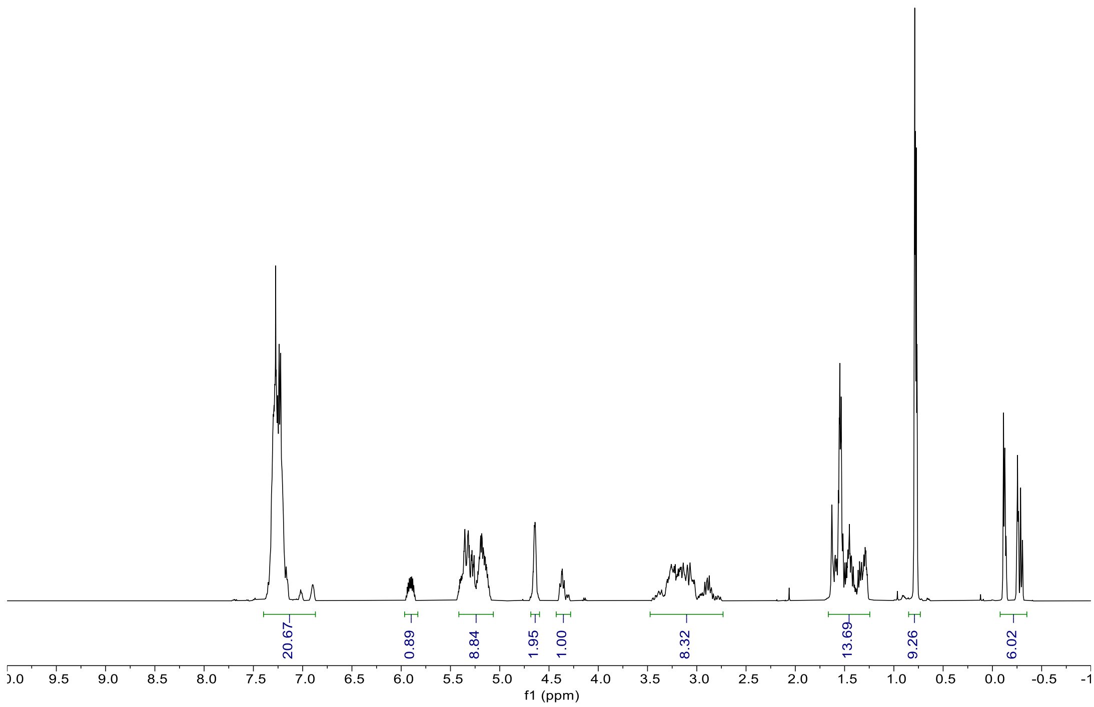
Supplementary Figure 57. ^{1}H NMR spectrum of PPLLLPPP.



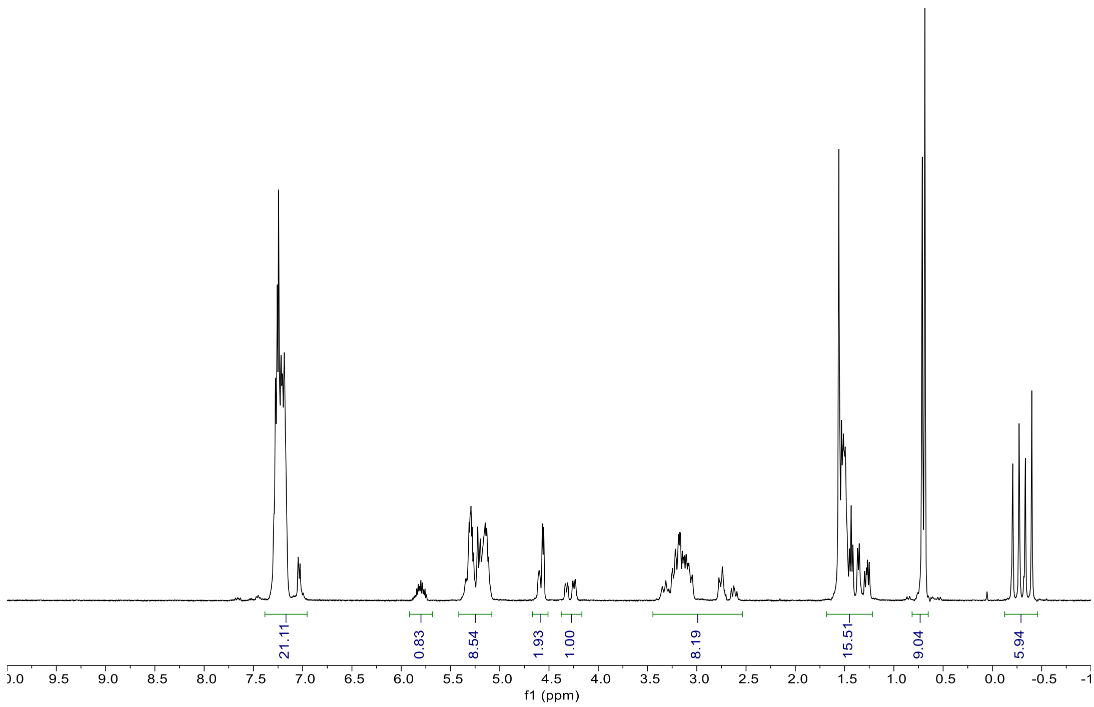
Supplementary Figure 58. ¹H NMR spectrum of PPLLPPLP.



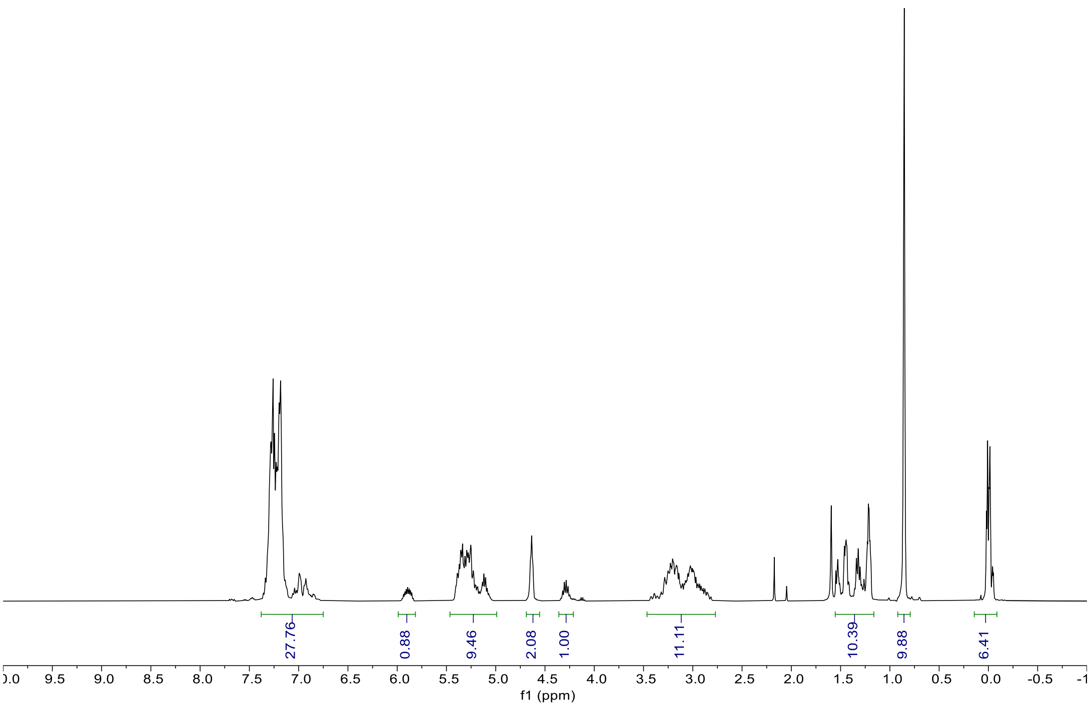
Supplementary Figure 59. ¹H NMR spectrum of PPLLPLLP.



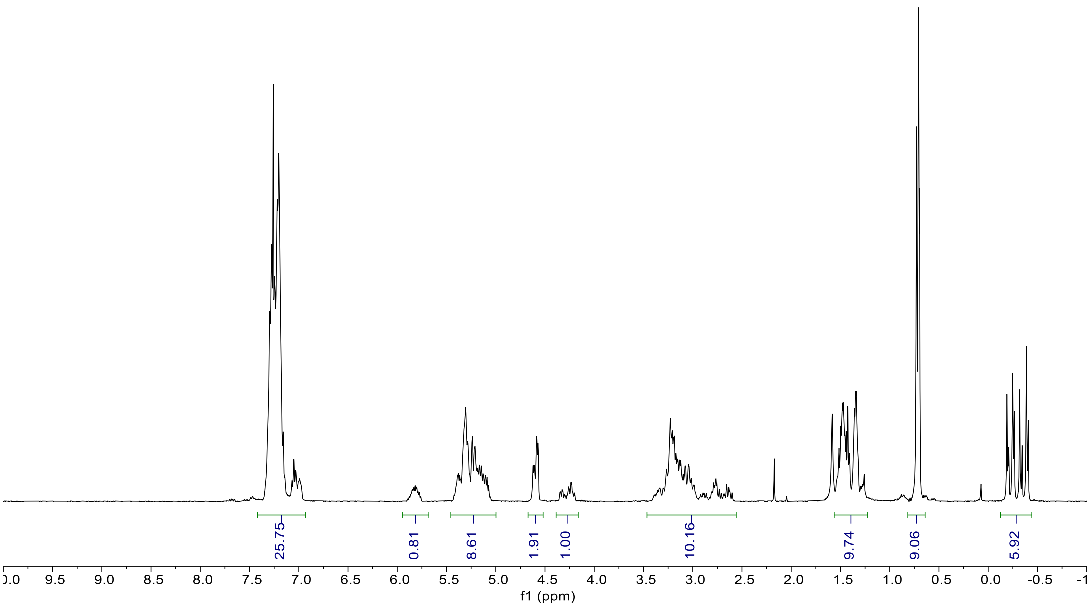
Supplementary Figure 60. ^{1}H NMR spectrum of PLPPLPLL.



Supplementary Figure 61. ^{1}H NMR spectrum of PPLPLLLP.

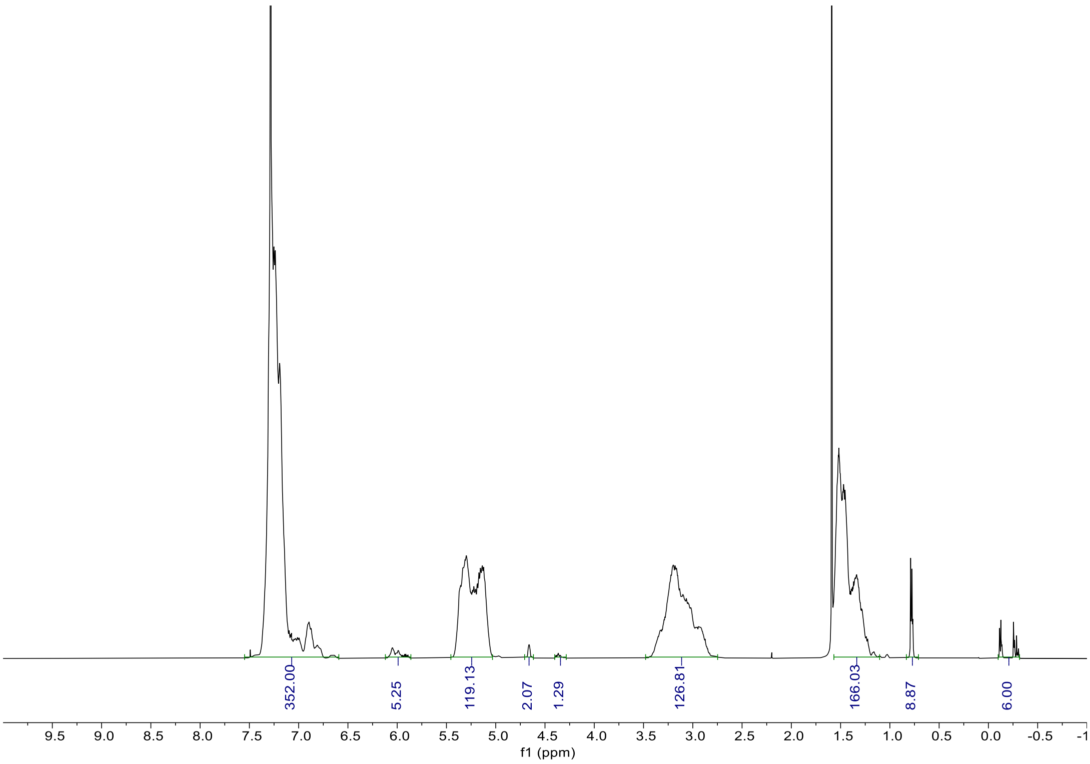


Supplementary Figure 62. ^{1}H NMR spectrum of LPPLPPPL.

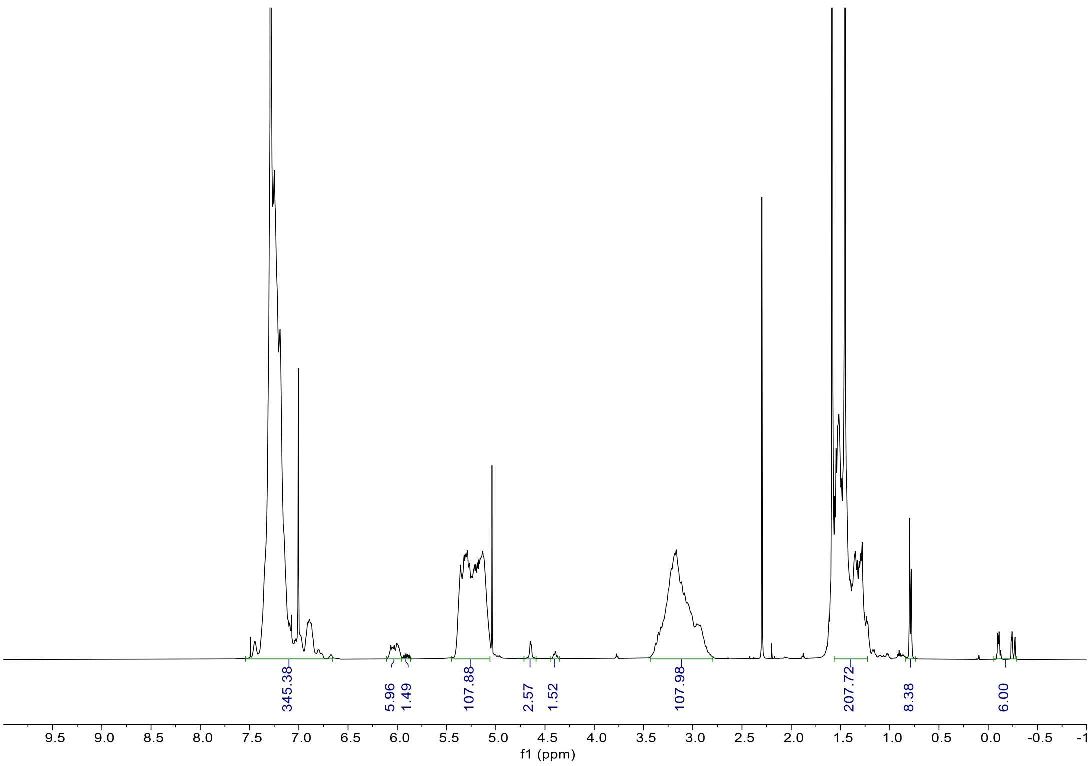


Supplementary Figure 63. ^{1}H NMR spectrum of PPPLLPLP.

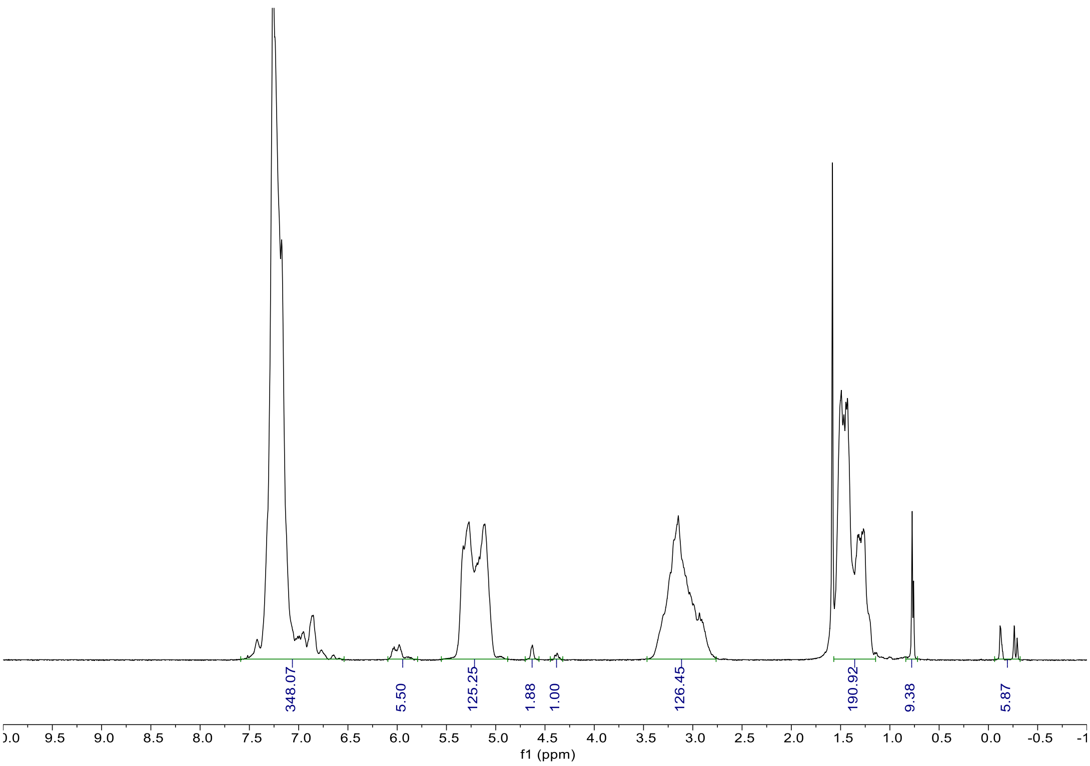
^{1}H NMR spectra of 128mers



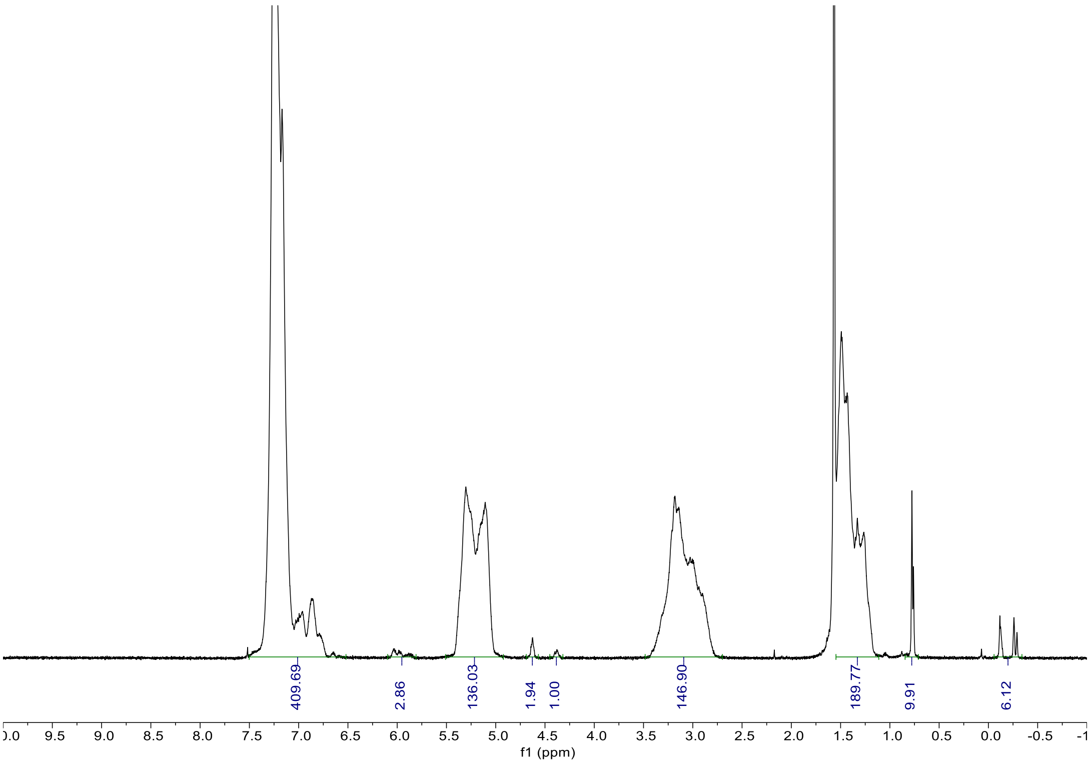
Supplementary Figure 64. ^{1}H NMR spectrum of 128-mer 1st.



Supplementary Figure 65. ^{1}H NMR spectrum of 128-mer 2nd.

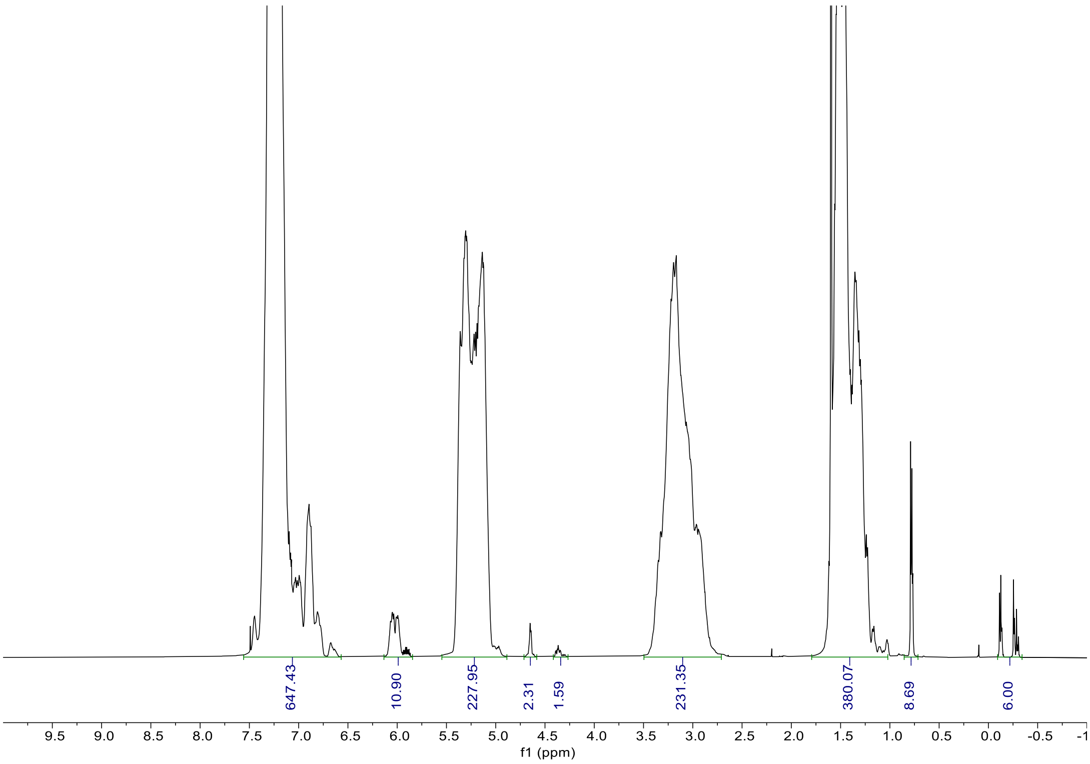


Supplementary Figure 66. ^{1}H NMR spectrum of 128-mer 3rd.

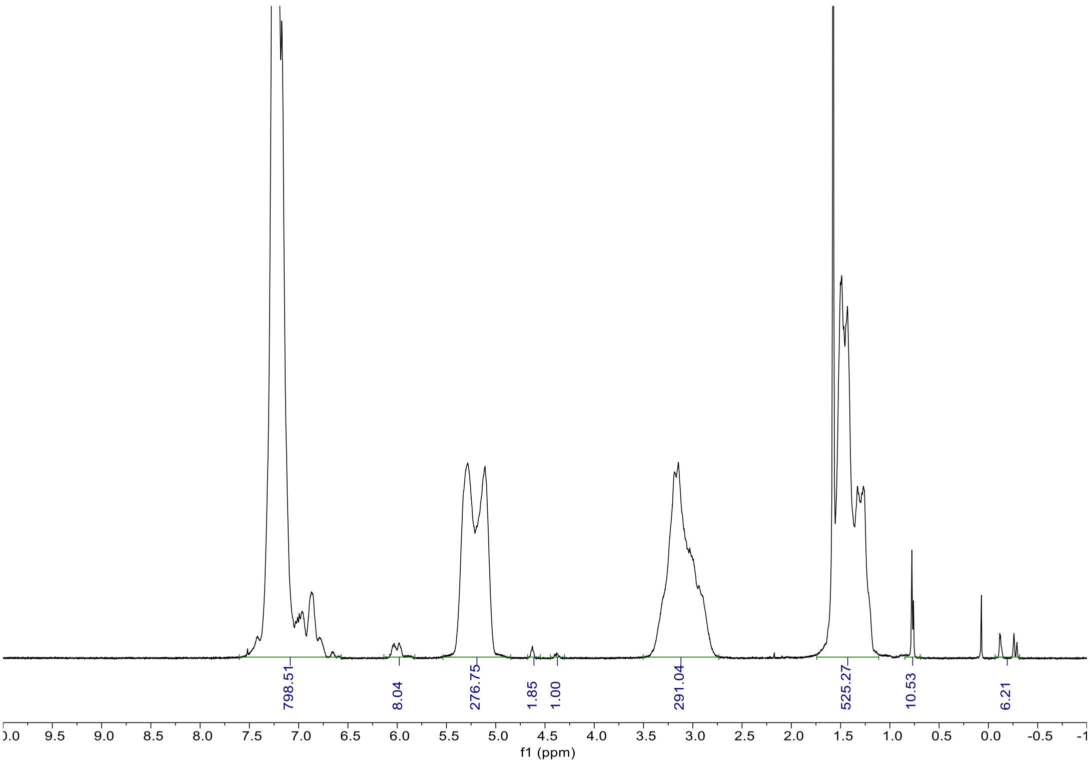


Supplementary Figure 67. ^{1}H NMR spectrum of 128-mer 4th.

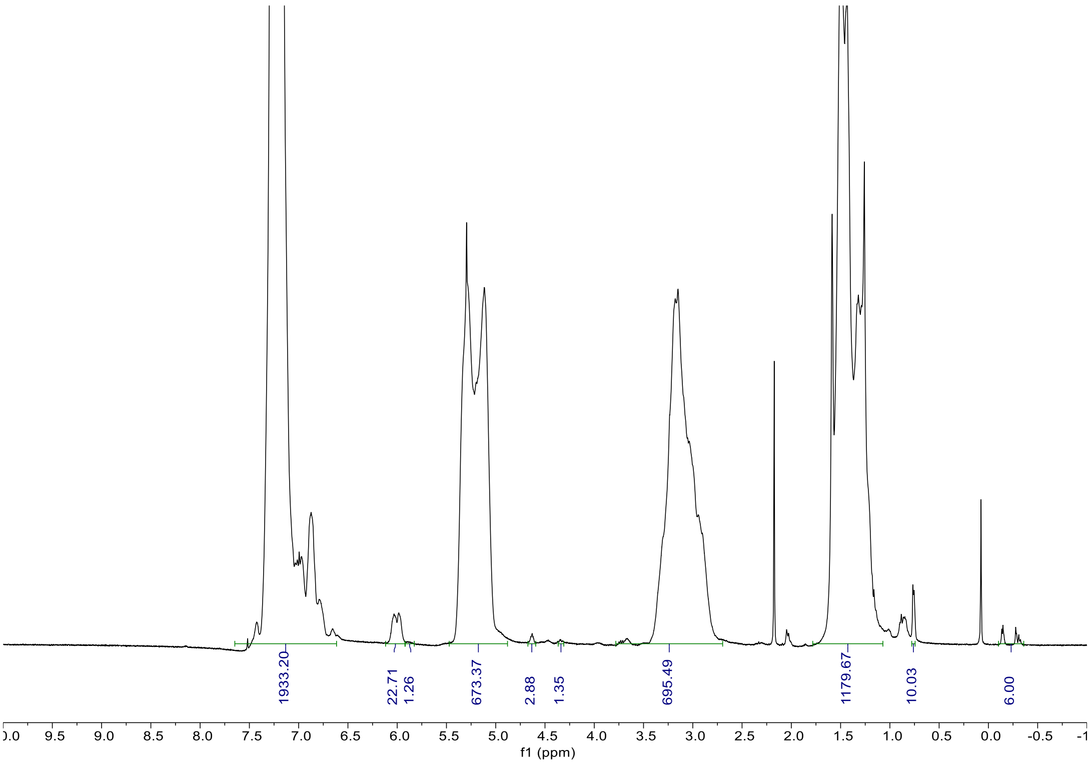
^{1}H NMR spectra of 256mers and 512mer



Supplementary Figure 68. ^{1}H NMR spectrum of 256-mer 1st.

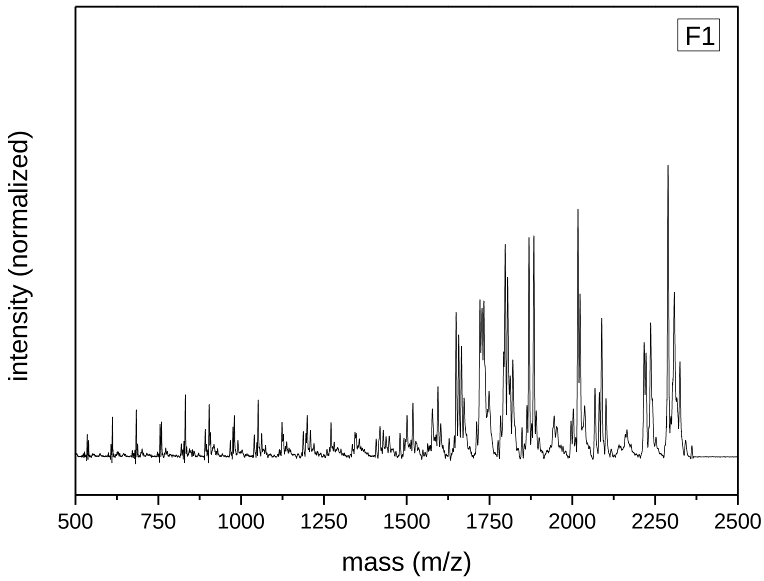


Supplementary Figure 69. ^{1}H NMR spectrum of 256-mer 2nd.

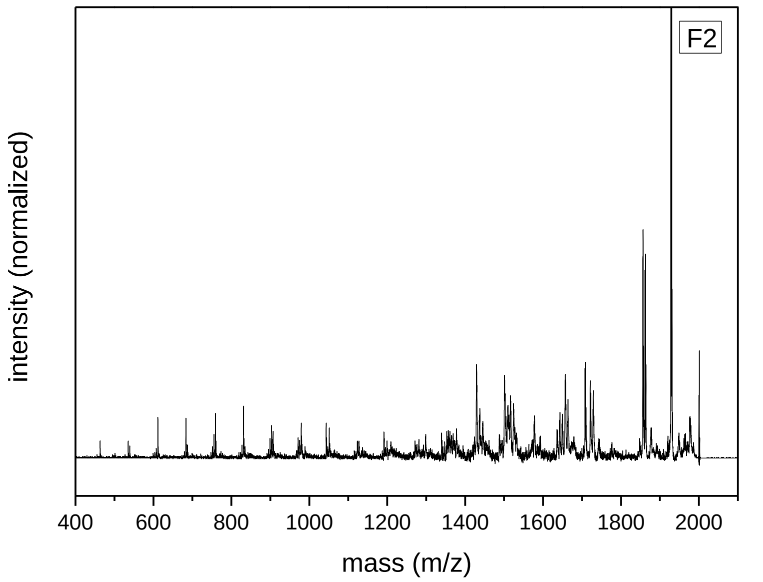


Supplementary Figure 70. **^{1}H NMR spectrum of 512-mer**

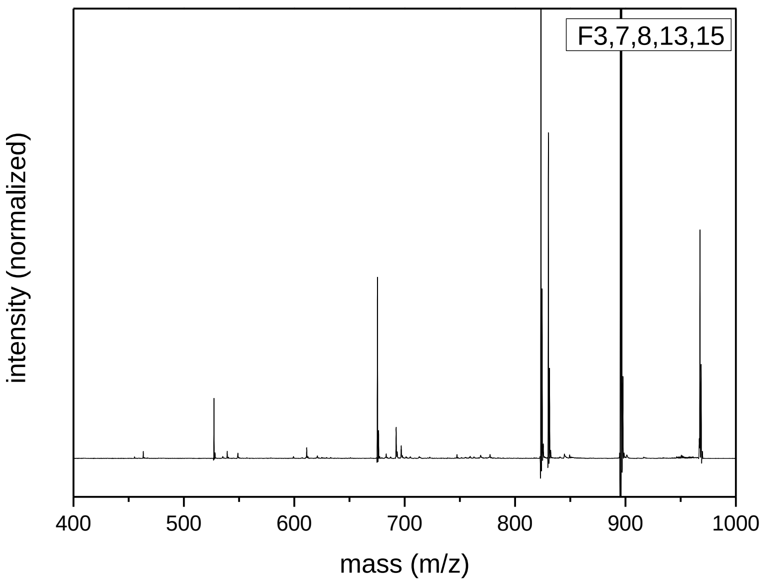
9. MALDI-TOF/TOF MS^2 spectra of oPcLs



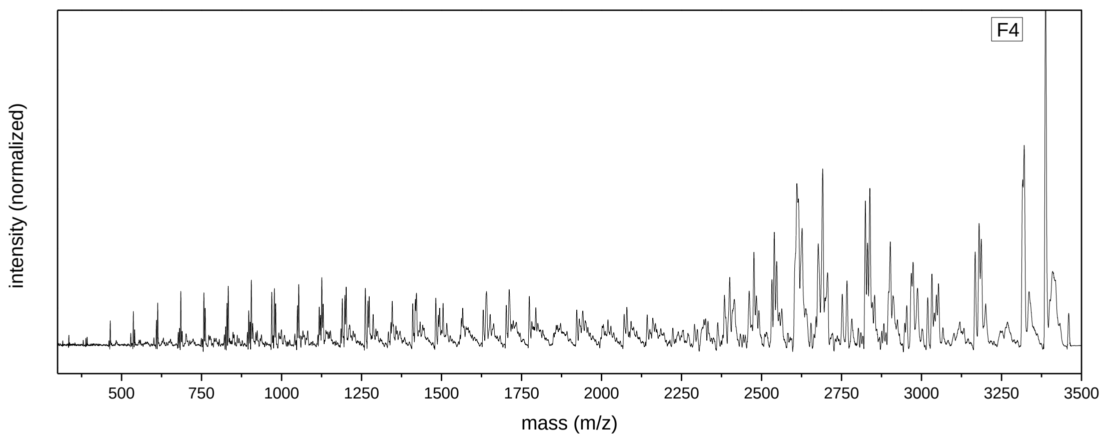
Supplementary Figure 71. MALDI-TOF/TOF MS^2 spectrum of F1 of 512-mer



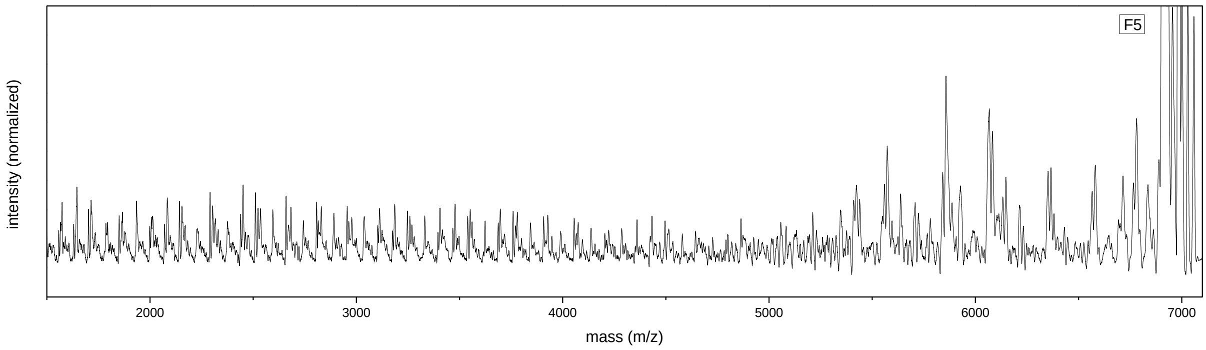
Supplementary Figure 72. MALDI-TOF/TOF MS^2 spectrum of F2 of 512-mer



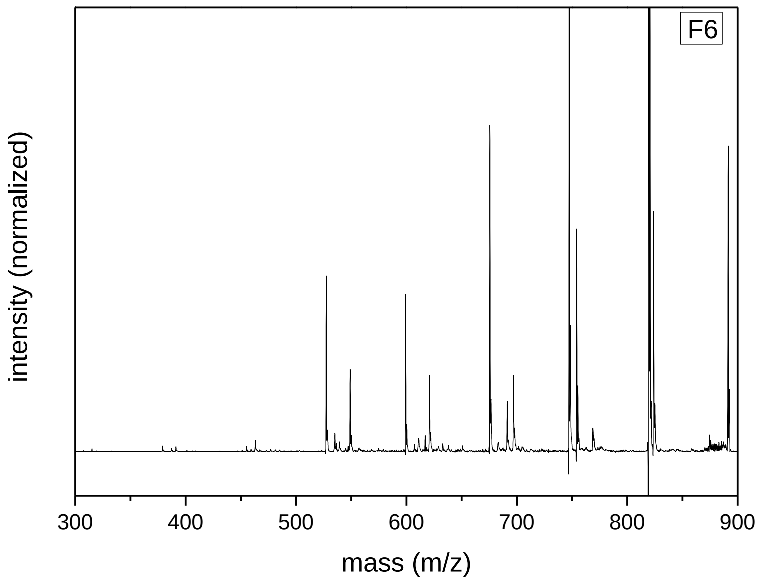
Supplementary Figure 73. MALDI-TOF/TOF MS^2 spectrum of F3, F7, F8, F13, and F15 of 512-mer



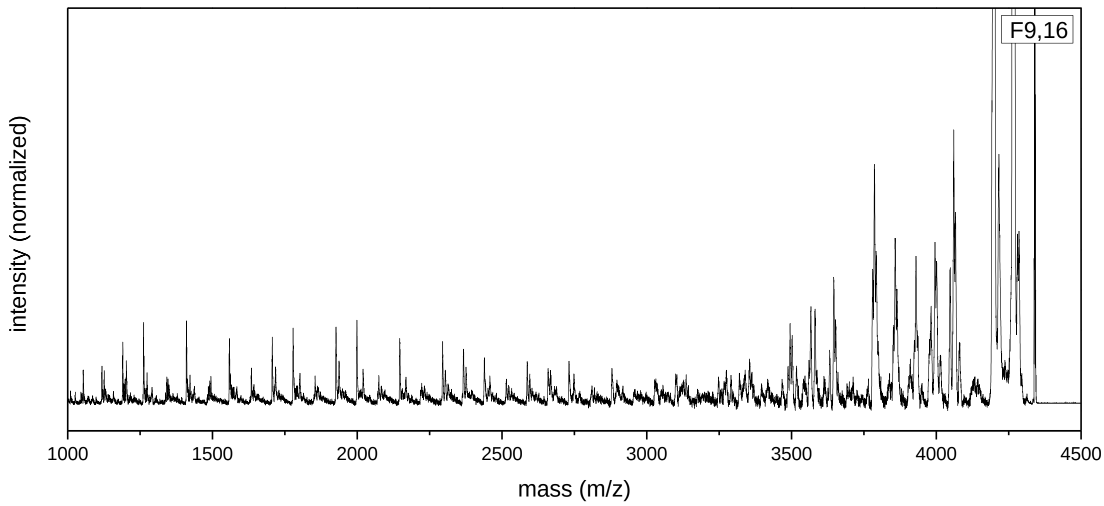
Supplementary Figure 74. MALDI-TOF/TOF MS^2 spectrum of F4 of 512-mer



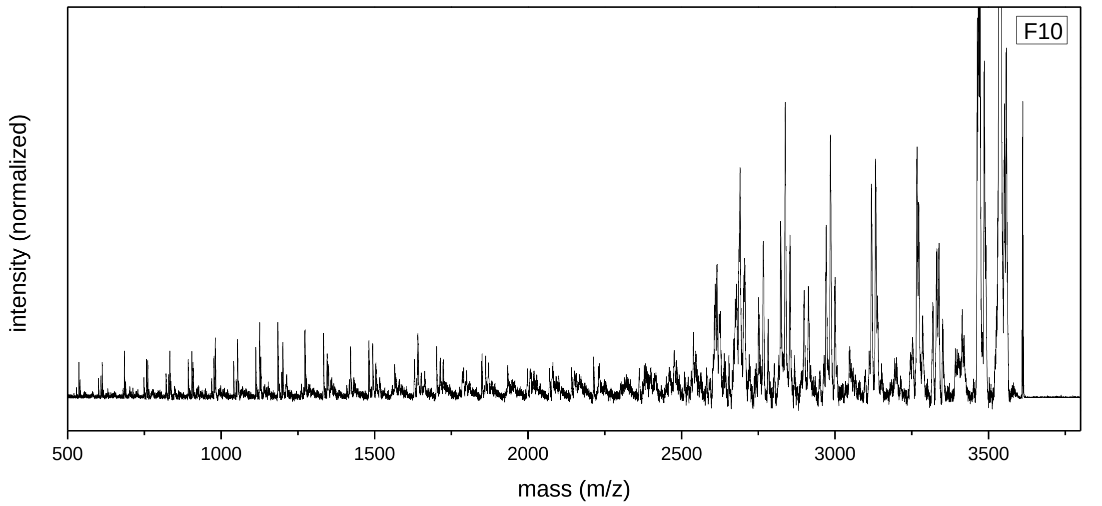
Supplementary Figure 75. MALDI-TOF/TOF MS^2 spectrum of F5 of 512-mer



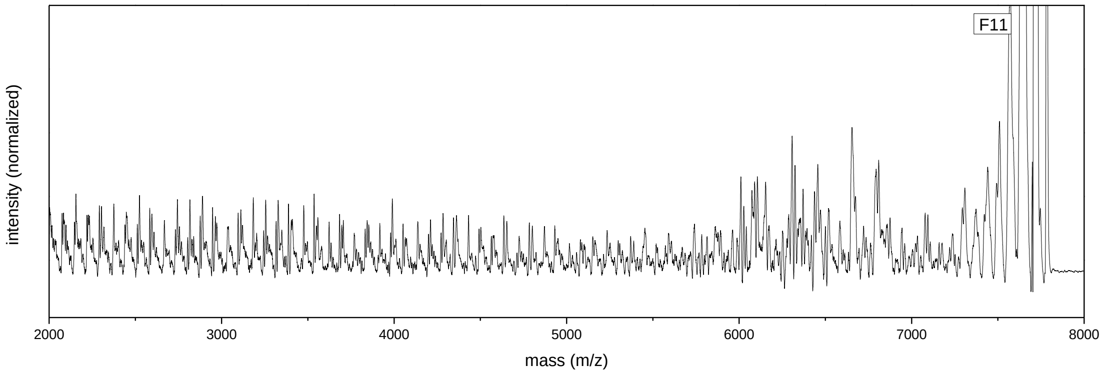
Supplementary Figure 76. MALDI-TOF/TOF MS^2 spectrum of F6 of 512-mer



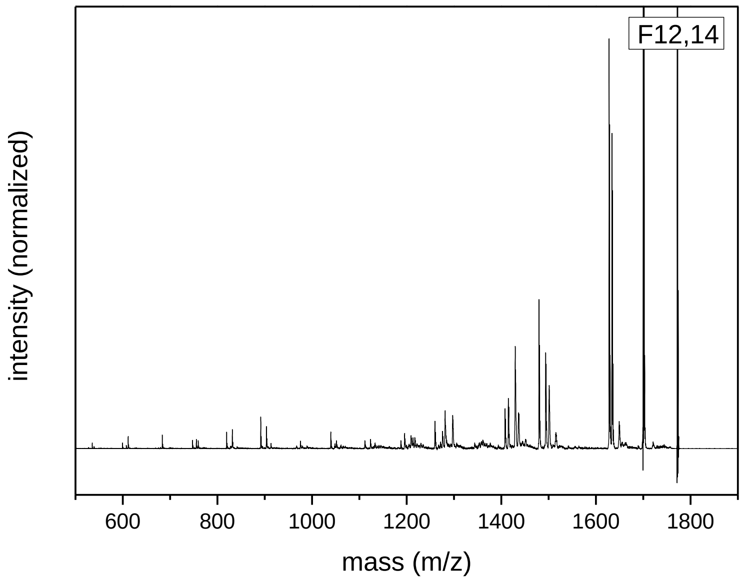
Supplementary Figure 77. MALDI-TOF/TOF MS^2 spectrum of F9 and F16 of 512-mer



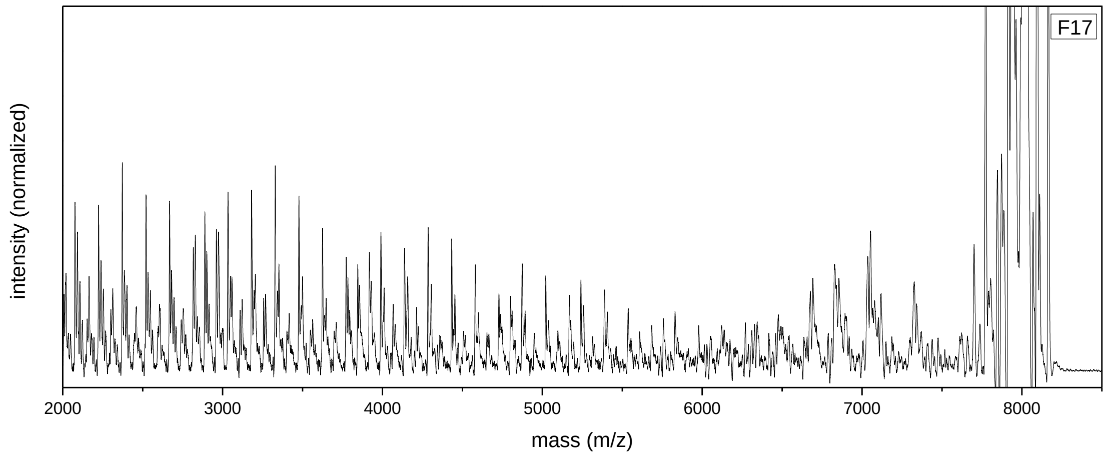
Supplementary Figure 78. MALDI-TOF/TOF MS^2 spectrum of F10 of 512-mer



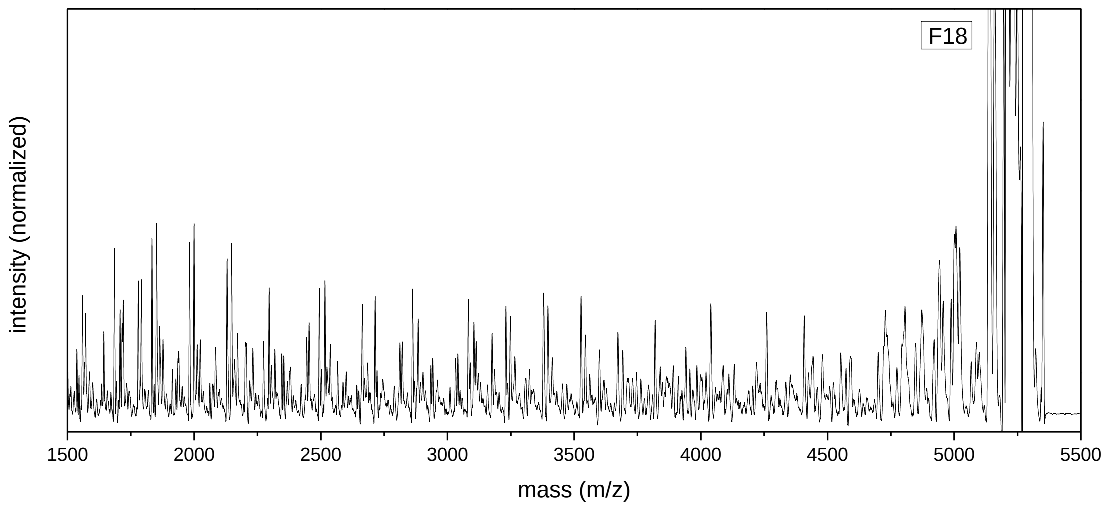
Supplementary Figure 79. MALDI-TOF/TOF MS^2 spectrum of F11 of 512-mer



Supplementary Figure 80. MALDI-TOF/TOF MS^2 spectrum of F12 and F14 of 512-mer



Supplementary Figure 81. MALDI-TOF/TOF MS^2 spectrum of F17 of 512-mer



Supplementary Figure 82. MALDI-TOF/TOF MS^2 spectrum of F18 of 512-mer