Chem, Volume *11*

Supplemental information

Departure from randomness:

Evolution of self-replicators that can

self-sort through steric zipper formation

Marcel J. Eleveld, Juntian Wu, Kai Liu, Jim Ottelé, Omer Markovitch, Armin Kiani, Lukas C. Herold, Alessia Lasorsa, Patrick C.A. van der Wel, and Sijbren Otto

Supplemental methods

Table of contents

S1. Structures of building blocks

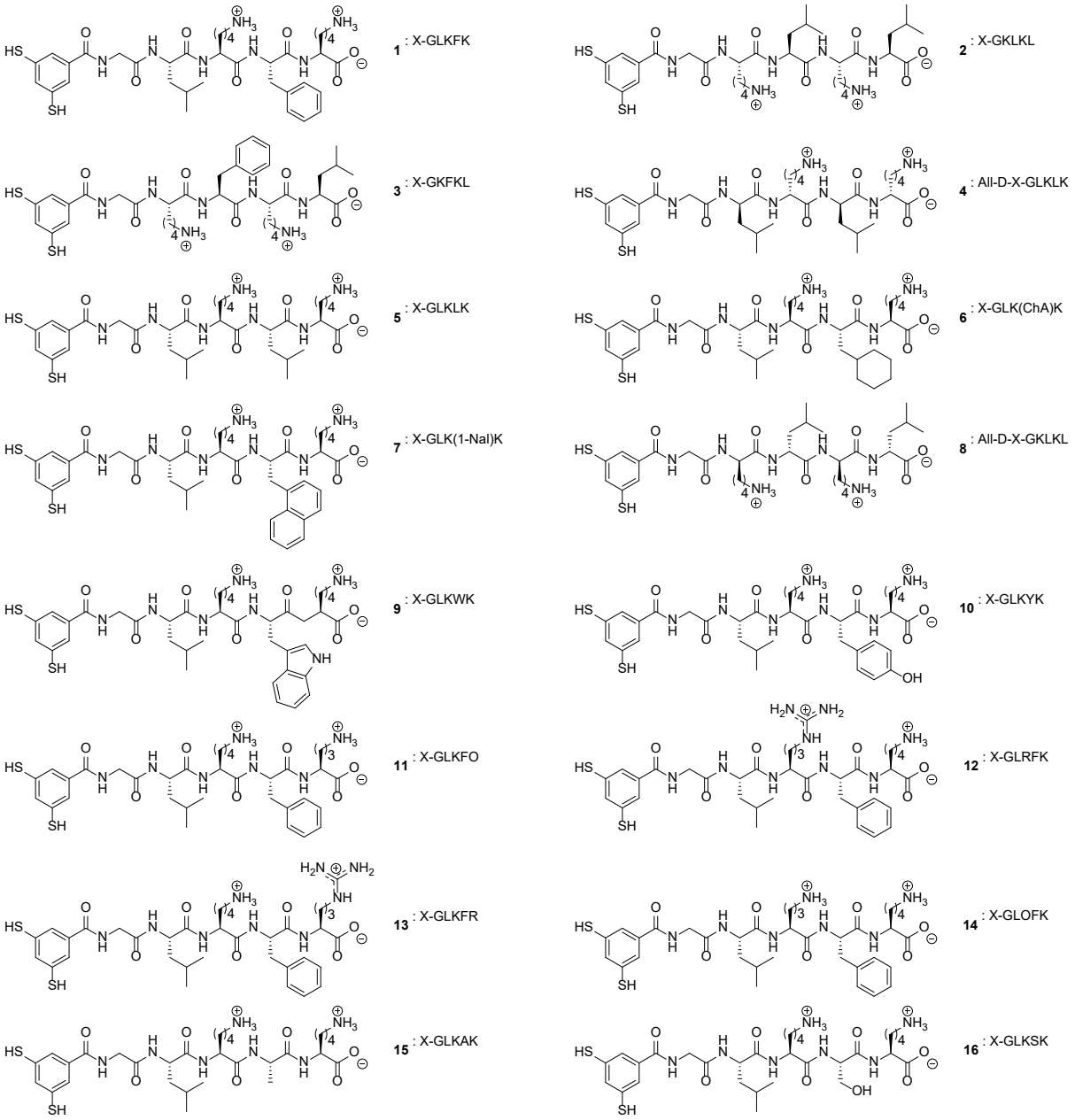


Figure S1. Molecular structures of building blocks and their amino-acid sequences.

S2. Mass spectrometry electron microscopy and spectroscopy methods

MALDI-TOF analysis

Matrix solution was prepared by mixing equal amounts of water and acetonitrile (ULC-MS grade) and adding 0.10 vol% of TFA (trifluoroacetic acid) as a modifier. Small aliquots of DCLs were diluted to 100 µM (in building block concentration) using the matrix solution. A solution was prepared of α-cyano-4-hydroxycinnamic acid (CHCA) by dissolving 10 mg of CHCA in 1.0 mL of matrix solution. The diluted aliquots of the DCL were mixed with the CHCA solution in ratios of 1:5, 1:10 and 1:20. Each sample:CHCA ratio was spotted three times on a MALDI plate (384 Opti-TOF 123 x 81 mm; 1.0 µL per spot). The spots were left to dry for one hour before analysis. Spectra were recorded using an Applied Biosystems 4800 MALDI TOF/TOF spectrometer operated in linear positive mass mode. Various mass ranges were analyzed with 20 subspectra (containing 50 shots) each. The laser was moved to a different position of the spotted sample before every sub-spectrum. Obtained mass spectra were analyzed using mMass (open source mass spectrometry tool) software.

Solid-state NMR analysis

Solid-state NMR experiments were performed on a Bruker AVANCE NEO NMR spectrometer operating at a ^{1}H Larmor frequency of 600 MHz (14.1 Tesla) equipped with a 3.2 mm EFree HCN MAS Probe. The sample (~2.5 mg of total fibrils), in the form of suspension (~16 mg of wet sample), was transferred into a 3.2-mm regular-wall MAS rotor using a home-built packing tool[2] via an ultracentrifugation step at 17,100 g, for 1 h at 4 °C (Optima LE-80K ultracentrifuge with a SW-32 Ti rotor, Beckman Coulter). Additional inserts were used to achieve center-packing of the sample material within the MAS rotor. Temperature was set to 278 K and spinning frequency to 12 kHz.

1D ^{13}C cross polarization (CP) and direct excitation (DE) experiments were measured using 50 kHz ^{13}C nutation frequency, 3 s repetition delay, 2048 scans and 0.02 s acquisition time. During acquisition, 83 kHz two-pulse phase modulation (TPPM) ^{1}H decoupling was applied.[3] For 1D ^{13}C CP, a 70%-100% ramped ^{1}H–^{13}C CP step was used with contact time set to 1000 µs.

1D ^{15}N CP was measured using 50 kHz ^{15}N nutation frequency, 3.5 s repetition delay, 6144 scans and 0.015 s acquisition time. During acquisition, 83 kHz TPPM ^{1}H decoupling was applied. Here, a 70%-100% ramped ^{1}H–^{15}N CP step was used with contact time set to 1000 µs.

The 2D ^{13}C–^{13}C CP-based dipolar-assisted rotational resonance (DARR) spectrum[4] was recorded using the same CP conditions, with DARR ^{13}C-^{13}C mixing time of 25 ms. Here, a 5 µs 90° ^{13}C pulse and 3 µs 90° ^{1}H pulse were used. TPPM ^{1}H decoupling during acquisition was around 83 kHz, repetition delay 3 s and number of scans 32 per datapoint.

A double quantum (DQ) – single quantum (SQ) ^{13}C-^{13}C shift correlation spectrum with SPC-5 DQ excitation/reconversion[5] was recorded using an equal number (i.e. 10) of composite pulses for excitation and reconversion, a 4.8 µs 90° ^{13}C pulse and 3 µs 90° ^{1}H pulse. TPPM ^{1}H decoupling during acquisition was around 100 kHz, recycle delay 3.8 s and number of scans 32 per datapoint. Here spinning frequency was set to 10kHz.

A CP-based ^{13}C detected water-edited spectrum was obtained by setting an initial ^{1}H T_2 (transverse relaxation time) filter, containing two symmetrical delays (τ), to 0.66 ms, and subsequently adding a longitudinal ^{1}H-^{1}H mixing time of 7 ms.[6] Here the CP contact time was set to 200 µs and number of scans was 1024.

Spectra were acquired with Bruker Topspin, processed with NMRPipe,[7] and analyzed with the CCPNmr Analysis Version-2.5.[8] The chemical shifts of ^{1}H and ^{13}C were indirectly referenced to aqueous DSS based on external measurements of the ^{13}C signals of adamantane,[9,10] setting the high-frequency adamantane peak to 40.49 ppm.

Thioflavin-T fluorescence assay

A ThT stock solution (2.2 mM) was prepared in 10 mL borate buffer (100 mM B_2O_3, pH 8.2) and filtered through a 0.2 µm syringe filter. On the day of analysis, 50 µL of the stock solution was diluted into 5 mL phosphate buffer (50 mM phosphate, pH 8.2) to generate the working solution of 22 µM. The fluorescence intensity of 450 µL ThT solution was measured by excitation at 440 nm (slit width 5 nm) and emission between 480-700 nm (slit width 5 nm), averaging 3 accumulations. An aliquot of 80 µL of a DCL (diluted to 80 µM in building block concentration) was added to the HELMA 10*2 mm quartz cuvette, incubated for 2 min, and the intensity was measured over 3 accumulations. All fluorescence measurements were performed on a JASCO FP6200 fluorimeter equipped with a 480 nm high pass cut-off filter on the emission channel to avoid high order diffractions coming from the excitation.

Fourier-transform Infrared Spectroscopy (FTIR)

FTIR spectroscopy of the freeze dried samples was performed using a PerkinElmer FT-IR Spectrum Two spectrometer. The samples were background corrected in air. Measurements were carried out in the spectral range 4000 -450 cm^{-1} with a resolution of 2 cm^{-1}. To increase the signal to noise ratio, each spectrum was an average of four scans.

Transmission Electron Microscopy (TEM)

The specimens were prepared by placing a small aliquot (5 µL) of sample onto a freshly glow-discharged copper grid (400 mesh) covered with a thin carbon film (Van Loenen Instruments). After waiting 30 seconds at room temperature, the droplet was gently blotted on filter paper. The sample was then negatively stained two times with 5 µL solution of 2% uranyl acetate deposited on the grid and blotted on the filter paper after 30 seconds each time. The grids were observed in a Philips CM120 cryo-electron microscope operating at 120 keV. Images were recorded on a slow scan CCD camera.

Circular Dichroism Analysis (CD)

All CD spectra were recorded using a JASCO J715 spectrophotometer and HELMA quartz cuvettes with a path length of 1 mm. Spectra were recorded at room temperature from 190 nm to 400 nm, with a 1 nm step interval and averaged over 3 scans using a scanning speed of 200 nm/min. Solvent spectra were subtracted from all reported spectra. Samples were diluted using borate buffer (100 mM B_2O_3, pH 8.2) to a concentration of 0.15 mM. The concentration is expressed in monomer units.

S3. UPLC analysis

UPLC analyses were performed on Waters Acquity UPLC H-class and H+class systems equipped with a PDA detector. All analyses were performed using a reversed-phase UPLC column (Aeris Peptide 1.7 µm XB-C18 150 × 2.10 mm, purchased from Phenomenex). UV absorbance was monitored at 254 nm. Column temperature was kept at 35 °C. UPLC-MS was performed using a Waters Acquity UPLC H-class system coupled to a Waters Xevo-G2 TOF. The mass spectrometer was operated in the positive electrospray ionization mode. Capillary, sampling cone, and extraction cone voltages were kept at 2.5 kV, 30 V, and 4 V, respectively. Source and desolvation temperatures were set at 150°C.

Solutions containing two different peptides and their oxidation products were prepared by diluting a small aliquot of a DCL (5 µL) in water (45 µL). These diluted samples were analyzed using the methods shown below (all gradients are linear). All DCLs were analyzed using **Method 1**, unless otherwise indicated.

Table S1. UPLC Method 1
Solvent A: ULC-MS grade water (0.1 V/V% trifluoroactic acid), Solvent B: ULC-MS grade acetonitrile (0.1% V/V% trifluoroactic acid)

time (min)	flowrate (mL/min)	% A	% B
0.0	0.30	90.0	10.0
1.0	0.30	90.0	10.0
1.3	0.30	75.0	25.0
3.0	0.30	72.0	28.0
11.0	0.30	60.0	40.0
11.5	0.30	5.0	95.0
12.0	0.30	5.0	95.0
12.5	0.30	90.0	10.0
17.0	0.30	90.0	10.0

Table S2. UPLC Method 2
Solvent A: ULC-MS grade water (0.1 V/V% trifluoroactic acid), Solvent B: ULC-MS grade acetonitrile (0.1% V/V% trifluoroactic acid)

time (min)	flowrate (mL/min)	% A	% B
0.0	0.30	90.0	10.0
1.0	0.30	90.0	10.0
1.3	0.30	75.0	25.0
3.0	0.30	72.0	28.0
7.2	0.30	65.7	34.3
13.1	0.30	64.3	35.7
16	0.30	60.0	40.0
16.5	0.30	5.0	95.0
18.0	0.30	5.0	95.0
18.5	0.30	90.0	10.0
20.0	0.30	90.0	10.0

UPLC chromatograms for single building block DCLs

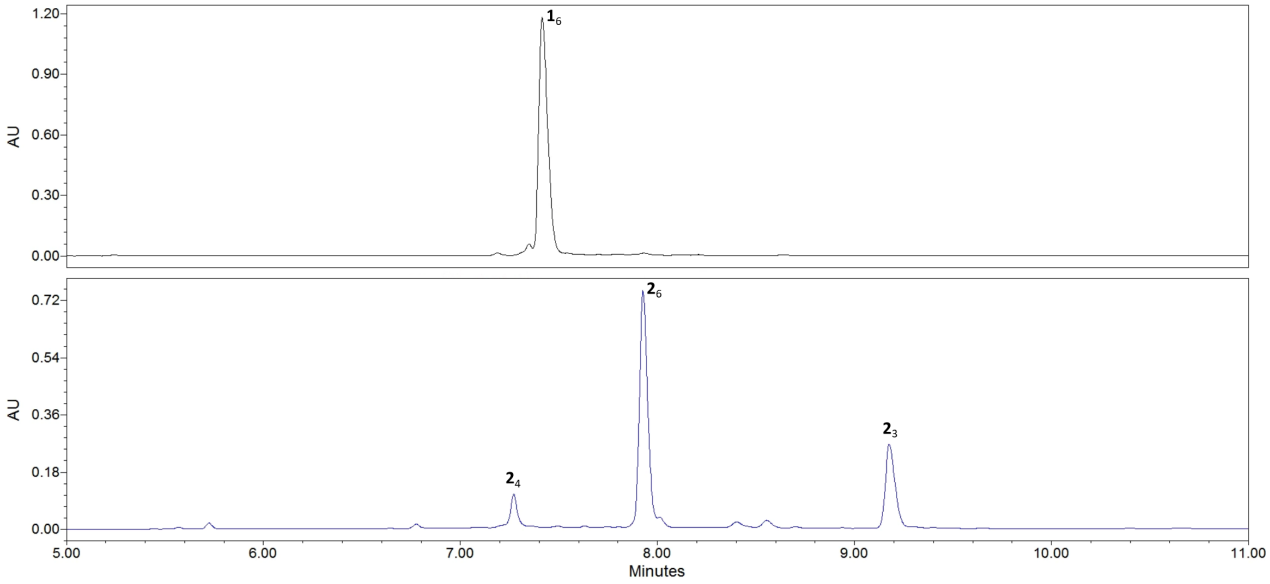


Figure S2. RP-UPLC chromatograms (monitored at 254 nm) of the product mixture obtained from a DCL made from **1** (top) or **2** (bottom) at 45 °C after 7 days.

Total peak area analysis of library containing 1 and 2

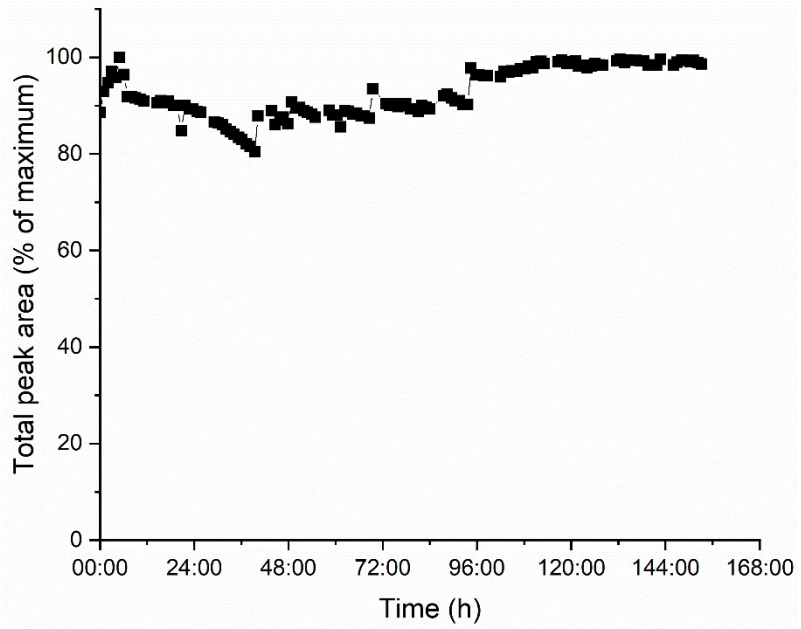


Figure S3. Total peak area obtained from RP-UPLC measurements (monitored at 254 nm) of the DCL formed upon mixing building blocks **1** and **2** (corresponding to Figures 3c+d in the main text). High-throughput measurements were automated using a 3D printed device.[1] The total peak area is normalized to the highest obtained value, which was set to 100%. The total peak area fluctuates between 100% and 80% over the course of the experiment, where the DLC composition changed substantially but is constant enough to justify the quantification of the formed species by comparing their relative peak areas. The relatively constant total peak area implies that the building blocks have the similar molar absorptivity, irrespective of the macrocycle in which they are part of.

S4. Structural study of zipper formation

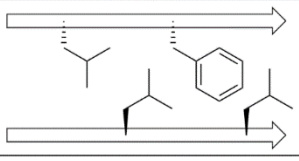
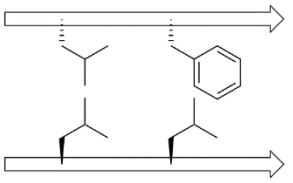
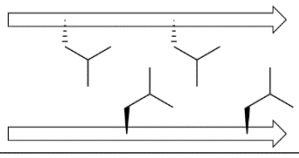
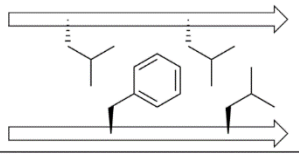
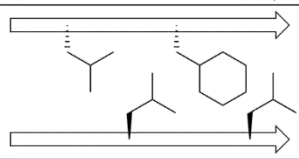
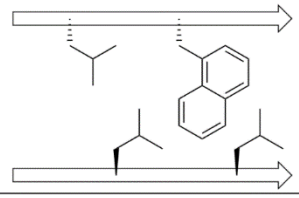
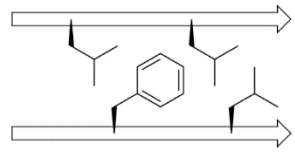
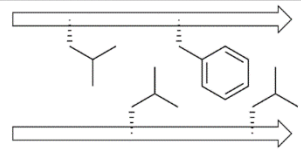
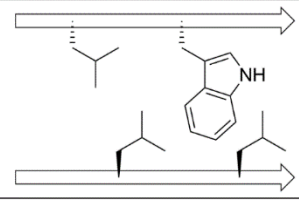
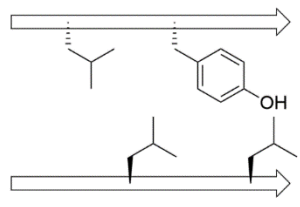
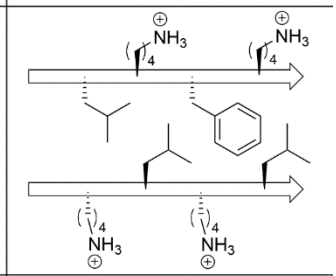
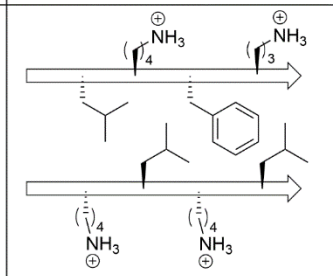
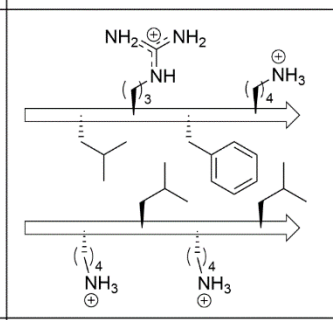
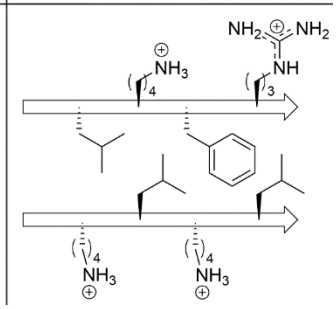
Building blocks	AF	Structure
1 + 2	2.6	
1 + 5	1.0	
5 + 2	1.4	
5 + 3	2.1	
6 + 2	2.2	
7 + 2	1.2	
4 + 3	1.1	
1 + 8	0.7	
9 + 2	1.0	
10 + 2	1.5	

Table S3. Overview of the amplification factors (AF) of $\mathbf{A_3B_3}$ obtained in the final hexamer distributions in DCLs made from equimolar amounts of two different building blocks (**A** and **B**). For each combination of building blocks a schematic drawing of the amino-acid side chains projecting from two adjacent β-sheets is shown. In these drawings the peptide backbone is depicted as an arrow and only the hydrophobic residues (and their stereochemistry) are shown. In the right part of the table also the charged residues on the hydrophobic faces of the β-sheets are shown.

Building blocks	AF	Structure
1 + 2	2.6	
11 + 2	2.4	
12 + 2	1.8	
13 + 2	1.9	

To investigate if there is indeed steric zipper formation at play here, we mixed **1** with **5**, a building block that contains the same amino acid residues as **2**, but with the residues in-register compared to **1**, where they are out-of-register when **1** is facing **2**. Because the hydrophobic residues are now on the 2nd and 4th position (after the dithiol core) in both the peptide chains, it should be more difficult to interdigitate them to form a zipper structure. In addition to this, two adjacent beta-sheets need to be formed in opposite directions (the carbonyl group closest to the dithiol core pointing in opposite directions) in order to have the hydrophobic faces facing each other and the aromatic core needs to be tilted to bring them in close proximity. We prepared a DCL with equimolar amounts of **1** and **5** using the same conditions as in previous experiments and analyzed the library composition when all thiols were oxidized to disulfides. As expected, no enrichment in $\mathbf{1}_3\mathbf{5}_3$ was observed, but a statistical distribution of all possible hexamers was obtained (Figure S31). This indicates that there is no self-sorting occurring in this system and that it is essential to have an offset between the hydrophobic residues in the adjacent beta-sheets in order to allow for the formation of a steric zipper structure.

Next we probed whether if the specific structure of the hydrophobic residues also plays a role. Therefore a DCL containing equimolar amounts of **5** and **2** was prepared. Building block **5** is a variant of **1** where the Phe residue is changed into a Leu residue. Both residues are hydrophobic, but their size and shape is different and Phe has the potential to participate in π-stacking interactions. The final distribution of the hexamers that are formed is enriched in $\mathbf{5}_3\mathbf{2}_3$ with an AF of 1.4 (Figure S32). This is significantly lower than the AF of 2.6 that was found for $\mathbf{1}_3\mathbf{2}_3$. We speculate that the smaller size of the Leu-residue makes interdigitation between the Leu-residues of the opposing peptide chain less efficient.

Knowing the importance of the Phe-residue for the formation of a steric zipper structure between **1** and **2**, we wanted to investigate if the interdigitation of a Phe-residue between two Leu-residues is a more universal motif. We therefore prepared a DCL with equimolar amounts of **5** and **3**. Building block **3** is a variant of **2** where the first Leu-residue is changed to a Phe-residue. Also here it should be possible to interdigitate the Phe-residue in between the Leu-residues of the opposing peptide chain. The final hexamer distribution indeed showed that also here there is a strong enrichment in the self-sorted $\mathbf{5}_3\mathbf{3}_3$ macrocycle, with an AF of 2.1 (Figure S33). The $\mathbf{5}_3\mathbf{3}_3$ macrocycle was isolated by collecting a fraction from the UPLC and it was analyzed by MALDI-TOF mass spectrometry, showing that the isomer where **5** and **3** are alternating is dominant (see Section S8). These results suggest that the interdigitation of a Phe-residue between two Leu-residues could be a general motif for steric zipper formation. From bioinformatics it is known that hydrophobic residues are prominent in tertiary packing of β-sheets, but to the best of our knowledge the interdigitation of a Phe-residue between two Leu-residues in a (parallel) β-sheet is not previously reported as a favorable interaction motif.

In order to rule out that π-interactions are the driving force behind the steric zipper formation a DCL containing equimolar amounts of **6** and **2** was prepared. Building block **6** is a variant of **1** where the Phe-residue is changed into a cyclohexylalanine-residue, which has a similar size compared to Phe, but lacks the possibility to participate in π-interactions. The distribution of the hexamers that are formed when all thiols have been converted to disulfides shows a strong enrichment in the $\mathbf{6}_3\mathbf{2}_3$ macrocycle, with an AF of 2.2 (Figure S34). Because a strong social self-sorting behavior is observed without the possibility of π-interactions in the peptides, this result confirms that the driving force behind steric zipper formation is most likely the interdigitation of a large hydrophobic residue (Phe or Cha) between two Leu-residues of the opposing peptide chain.

Next we investigated if interdigitation also works with larger hydrophobic residues. Therefore DCLs were prepared using equimolar amounts of **7** + **2**, **9** + **2** and **10** + **2**. Building blocks **7**, **9** and **10** are variants of **1**, where the Phe is changed into 1-Nal (1-naphthylalanine), Trp or Tyr, respectively. These amino acids also bear flat and hydrophobic residues, with 1-Nal and Trp being significantly larger than Phe (or Cha) and Tyr with a size in between. The obtained distribution of hexamers for **7** + **2** shows a slight enrichment in the $\mathbf{7}_3\mathbf{2}_3$ macrocycle with an AF of 1.2 (Figure S35) and for **9** + **2** an AF of 1.0 for the $\mathbf{9}_3\mathbf{2}_3$ is obtained (Figure S36). The hexamer distribution obtained for **10** + **2** showed an AF of 1.5 (Figure S37) for the $\mathbf{10}_3\mathbf{2}_3$ hexamer, which indicates that there might be some steric zipper formation between **10** and **2**, but to a far lesser extent than between **1** and **2**. The results suggest that there is an optimal size of the interdigitating hydrophobic residue in steric zipper formation. When the residue is too large (1-Nal, Trp or Tyr) or too small (Leu) it appears to be harder to achieve an optimal packing of hydrophobic residues in the zipper structure.

We also wanted to investigate if the relative orientation of the residues in adjacent peptide chains plays a role in steric zipper formation. We therefore prepared a DCL of **4** and **3** as a comparison to the DCL of **5** and **3**. Building block **4** is a variant of **5**, where the stereochemistry of every amino acid is inverted. The final hexamer distribution in this DCL is very close to a statistical distribution and the $\mathbf{4}_3\mathbf{3}_3$ macrocycle is formed with only a small AF of 1.1 (Figure S38). We also prepared a DCL using equimolar amounts of **1** and **8**, where **8** is the enantiomer of **2**. The final hexamer distribution of this DCL also shows no social self-sorting behavior. The $\mathbf{1}_3\mathbf{8}_3$ macrocycle is only produced with an AF of 0.7 (Figure S38). This is mainly caused by the incomplete incorporation of **8** into the hexamer assembly. Both results show that the relative orientation of the peptide chains and/or the orientation of their hydrophobic residues play a crucial role in steric zipper formation. When in two systems that show a strong tendency to socially self-sort the orientation of the residues in one of the building blocks is changed by using its enantiomer, no social self-sorting was observed. This could potentially be a result of the fact that the adjacent peptides have a different orientation, where one of them has to point up from the aromatic core, while the other has to point down from the core in order for their hydrophobic faces to face each other.

To conclude this structural analysis of steric zipper formation between two different building blocks, we wanted to confirm that indeed the hydrophobic residues were mainly responsible for the observed self-sorting effect. We therefore prepared DCLs of **11** + **2**, **12** + **2** and **13** + **2**. Building blocks **11** and **13** are variants of **1** where the second Lys was changed into an Orn (Ornithine, **11**) or an Arg (**13**). Building block **12** is also a variant of **1** where the first Lys was changed into an Arg. Changing the second Lys into Orn has a minimal effect on the AF of the $\mathbf{A}_3\mathbf{B}_3$ hexamers (AF = 2.6 → AF = 2.4; Figure S39). Changing the first or second Lys into Arg had a stronger effect on the AF, which became 1.8 and 1.9, respectively (Figures S40 and S41, respectively). These effects are, however, still much smaller than the observed drop in AF when changing the size or the position of the hydrophobic residues in the peptide chain or when the stereochemistry of one building block was changed.

S5. Circular Dichroism data

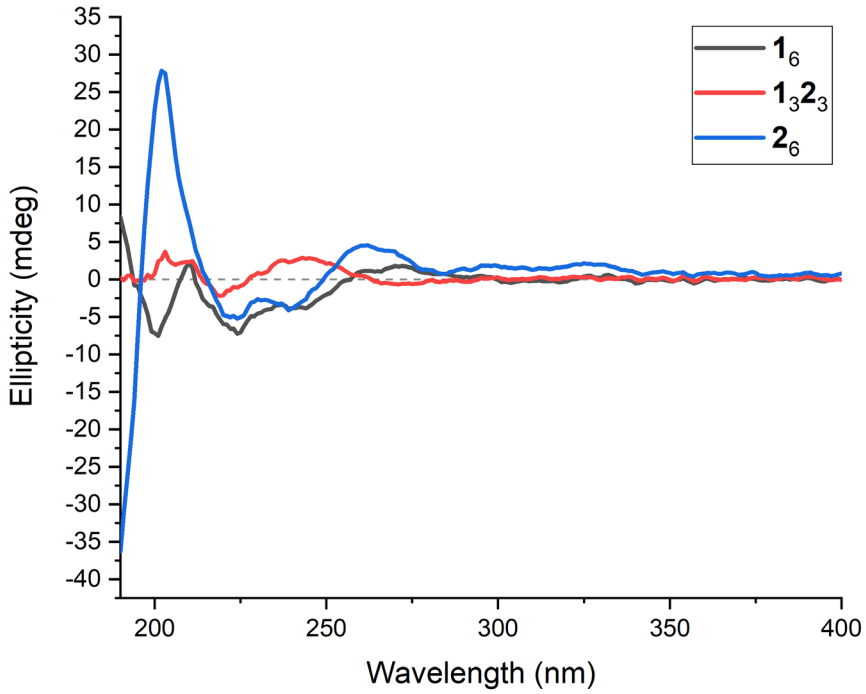


Figure S4. Circular dichroism spectra obtained from DCLs made from only **1** (which formed predominantly $\mathbf{1}_6$), only **2** (which formed predominantly $\mathbf{2}_6$) or equimolar amounts of **1** and **2** (which formed predominantly $\mathbf{1}_3\mathbf{2}_3$) after stirring at 45 °C for 6 days.

S6. Thioflavin T (ThT) Fluorescence Data

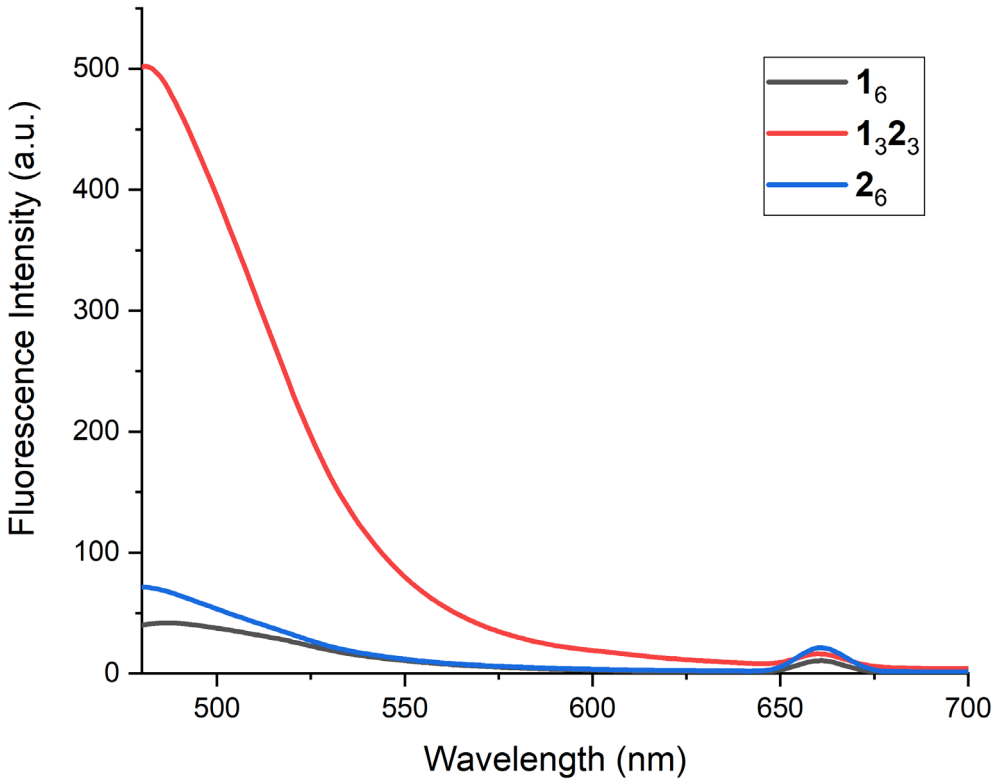


Figure S5. Thioflavin T (ThT) emission spectra obtained from DCLs made from only **1** (which formed predominantly $\mathbf{1}_6$), only **2** (which formed predominantly $\mathbf{2}_6$) or equimolar amounts of **1** and **2** (which formed predominantly $\mathbf{1}_3\mathbf{2}_3$) after stirring at 45 °C for 6 days. The background ThT fluorescence was subtracted from all spectra.

S7. Fourier-transform Infrared Spectroscopy Data

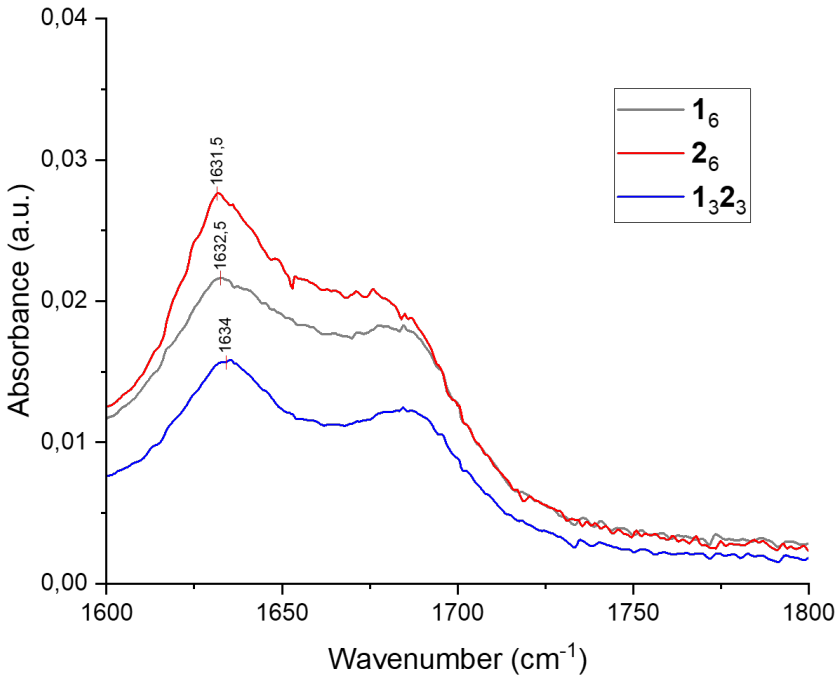


Figure S6. FTIR spectra obtained from freeze dried DCLs made from only **1** (which formed predominantly $\mathbf{1}_6$), only **2** (which formed predominantly $\mathbf{2}_6$) or equimolar amounts of **1** and **2** (which formed predominantly $\mathbf{1}_3\mathbf{2}_3$) after stirring at 45 °C for multiple days. Marked peaks are amide I absorption bands, which are characteristic for β-sheet structures.

S8. Transmission Electron Microscopy (TEM) Data

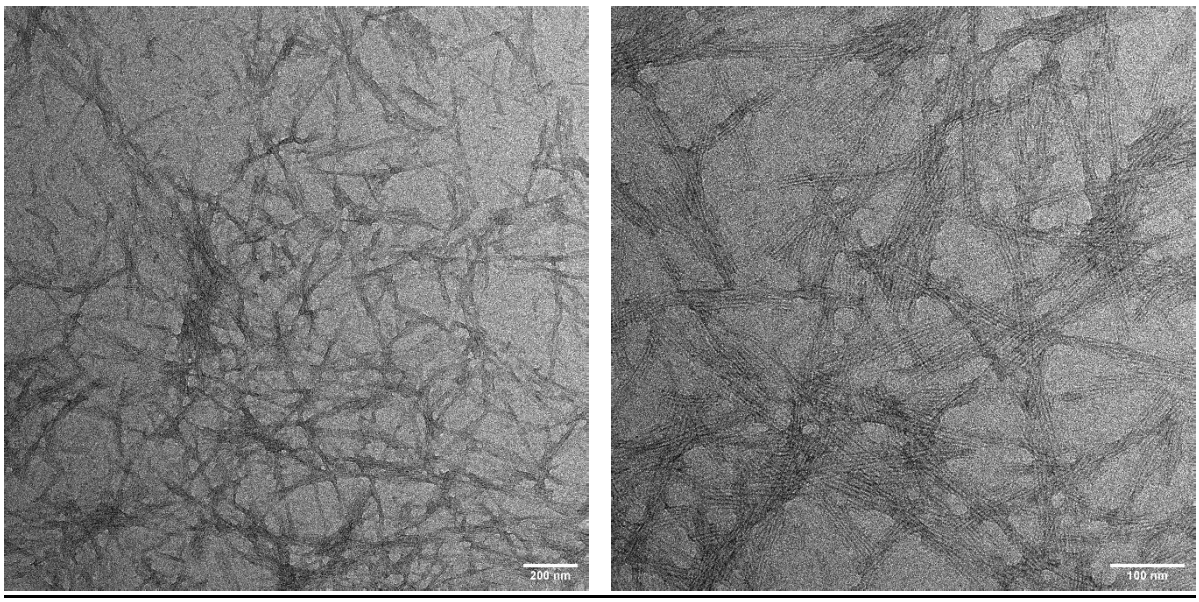


Figure S7. Transmission electron micrographs of a DCL made from **1**, which formed predominantly the $\mathbf{1}_6$ replicator, indicating that $\mathbf{1}_6$ is assembled into fibres.

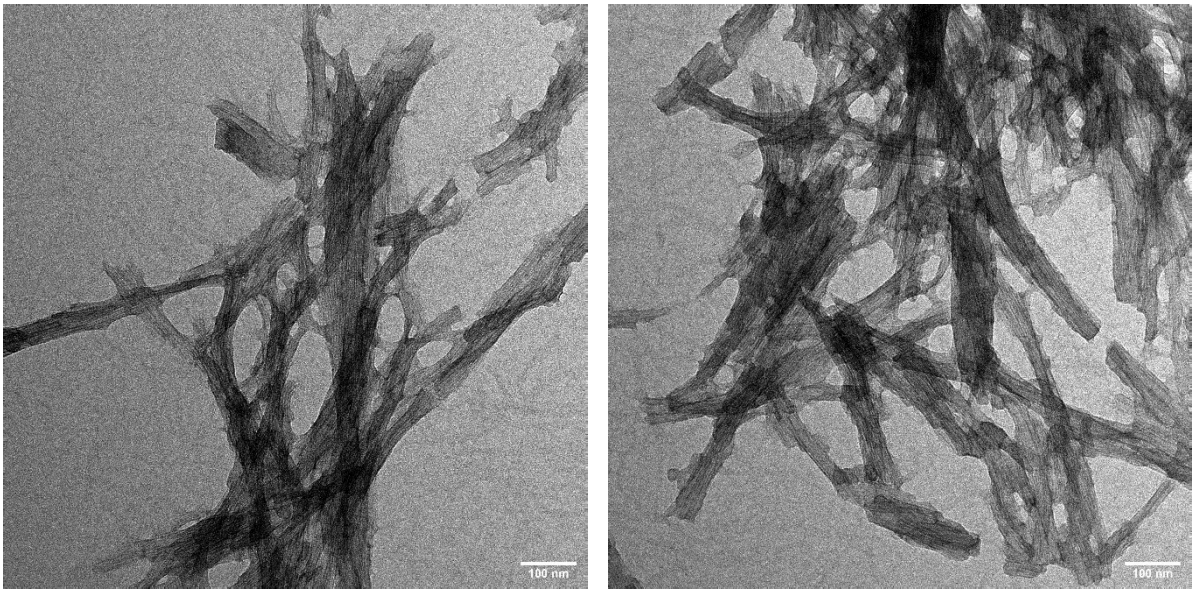


Figure S8. Transmission electron micrographs of a DCL made from **2**, which formed predominantly the $\mathbf{2}_6$ replicator, indicating that $\mathbf{2}_6$ is assembled into fibres.

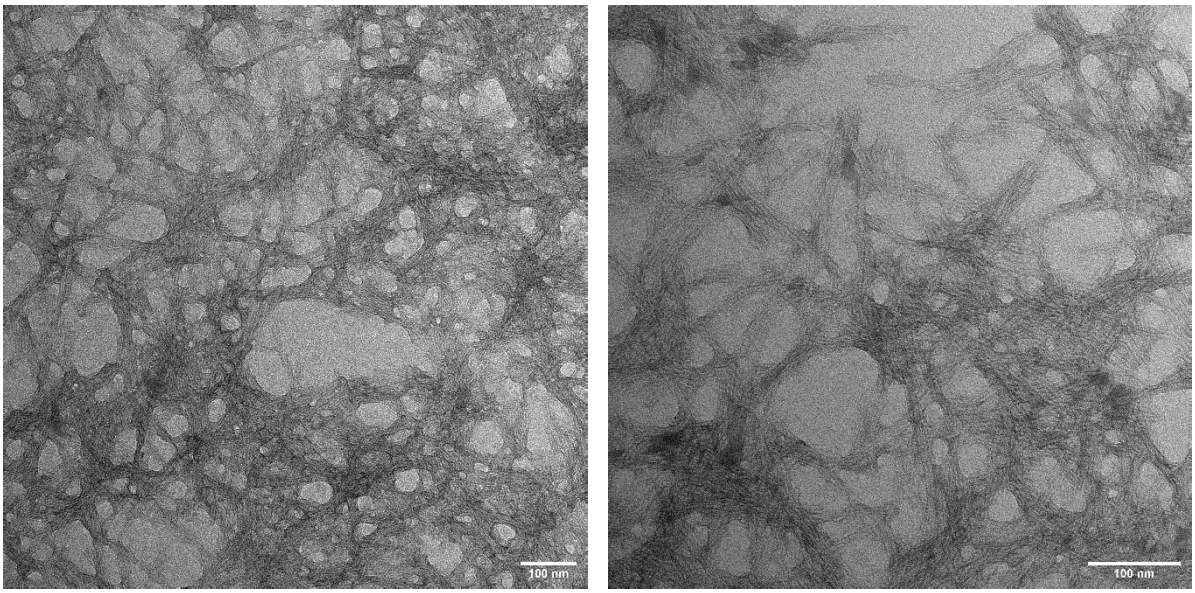


Figure S9. Transmission electron micrographs of a DCL made from equimolar amounts of **1** and **2**, which formed predominantly the $\mathbf{1}_3\mathbf{2}_3$ replicator, indicating that $\mathbf{1}_3\mathbf{2}_3$ is assembled into fibres.

S9. MALDI-TOF fragmentation data

To confirm that the fragmentation of the disulfide bonds in MALDI-TOF is not selective for the composition of the macrocycle a sample containing a statistical distribution of hexamers that was formed in a DCL containing **3** and **4** was analyzed (Table S4). The initial distribution closely matches the expected distribution that follows from the ratios in Pascal's triangle (1:6:15:20:15:6:1 for A6, A5B1… A1B5, B6 respectively). When the chances of fragmentation are the same for all (A-A, A-B and B-B) disulfide bonds, one would expect the ratio of the fragments of a specific size to also follow a statistical distribution. For example for trimers the expected ratio is 1:3:3:1. The observed ratios are quite close to the expected ratios, which confirms the aselectivity of the disulfide fragmentation.

Table S4. Observed areas for all possible species from MALDI-TOF analysis of a DCL consisting of **3** and **4**, which produces a statistical distribution of hexamer replicators. The observed areas are also shown as normalized values for easier comparison with the expected ratios.

Possible species	Observed area	Normalized ratio	Expected ratio
$\mathbf{3}_6$	4043	0,7	1
$\mathbf{3}_5\mathbf{4}_1$	25142	4,3	6
$\mathbf{3}_4\mathbf{4}_2$	75552	13,0	15
$\mathbf{3}_3\mathbf{4}_3$	116147	20,0	20
$\mathbf{3}_2\mathbf{4}_4$	86236	14,8	15
$\mathbf{3}_1\mathbf{4}_5$	25584	4,4	6
$\mathbf{4}_6$	2727	0,5	1
$\mathbf{3}_5$	314	0,1	1
$\mathbf{3}_4\mathbf{4}_1$	5452	2,5	5
$\mathbf{3}_3\mathbf{4}_2$	21561	10,0	10
$\mathbf{3}_2\mathbf{4}_3$	24835	11,5	10
$\mathbf{3}_1\mathbf{4}_4$	8058	3,7	5
$\mathbf{4}_5$	55	0,0	1
$\mathbf{3}_4$	6746	0,6	1
$\mathbf{3}_3\mathbf{4}_1$	35219	3,4	4
$\mathbf{3}_2\mathbf{4}_2$	62292	6,0	6
$\mathbf{3}_1\mathbf{4}_3$	41513	4,0	4
$\mathbf{4}_4$	3961	0,4	1
$\mathbf{3}_3$	74706	1,0	1
$\mathbf{3}_2\mathbf{4}_1$	228064	3,0	3
$\mathbf{3}_1\mathbf{4}_2$	259517	3,4	3
$\mathbf{4}_3$	76162	1,0	1
$\mathbf{3}_2$	56888	0,9	1
$\mathbf{3}_1\mathbf{4}_1$	129168	2,0	2
$\mathbf{4}_2$	58304	0,9	1

It is crucial to rule out that the fragments that are created in the ionization step can exchange building blocks in the gas phase before they reach the detector in the mass spectrometer. We therefore prepared a sample where we mixed preformed $\mathbf{3}_6$ and $\mathbf{4}_6$ together and analyzed the mixture using MALDI-TOF.

In the partial mass spectra the observed fragments for this mixture ($\mathbf{3}_6$ and $\mathbf{4}_6$) is shown in black and the observed fragments from the statistical distribution of **3** and **4** is shown in red (Figure S10 and Figure S11). The absence of the mixed fragments in the black curve indicate that there is no scrambling occurring in the gas phase.

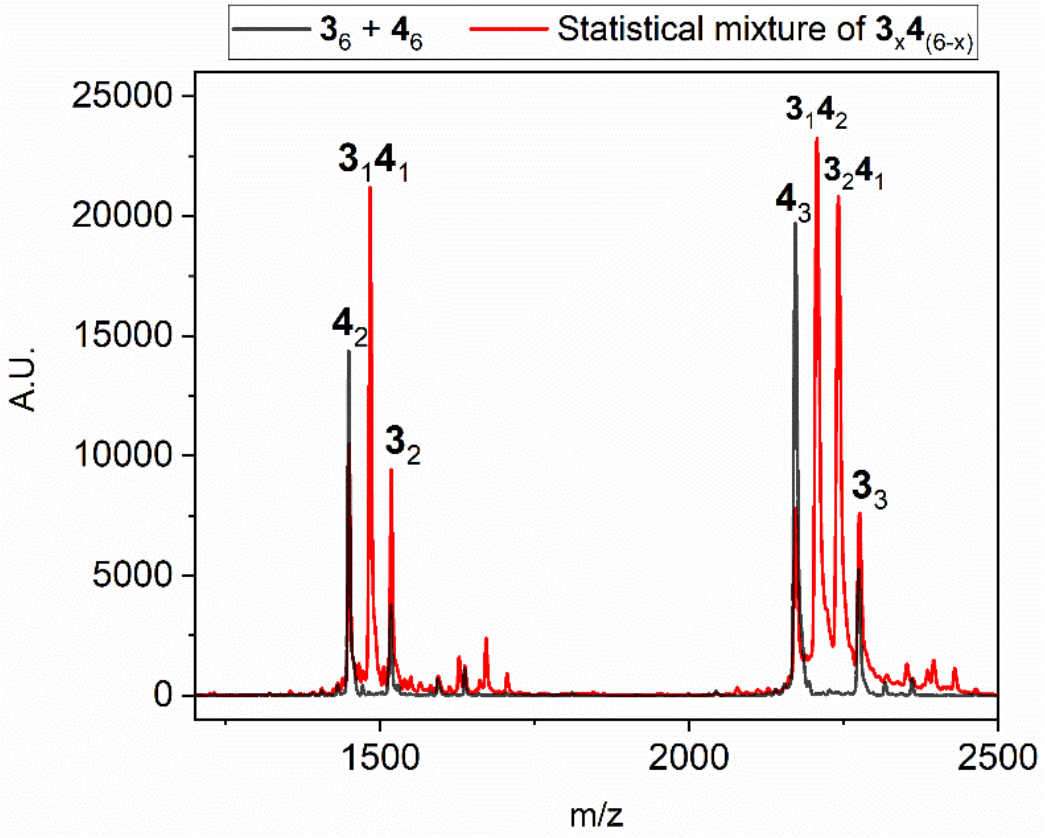


Figure S10. Partial MALDI-TOF mass spectra (showing dimer and trimer fragments) of $\mathbf{3}_6$ and $\mathbf{4}_6$ spotted together (black) and a statistical mixture of hexamers prepared by mixing building blocks **3** and **4** (red). The absence of fragments containing both building blocks in the sample where $\mathbf{3}_6$ was spotted together with $\mathbf{4}_6$ shows that there is no scrambling of fragments.

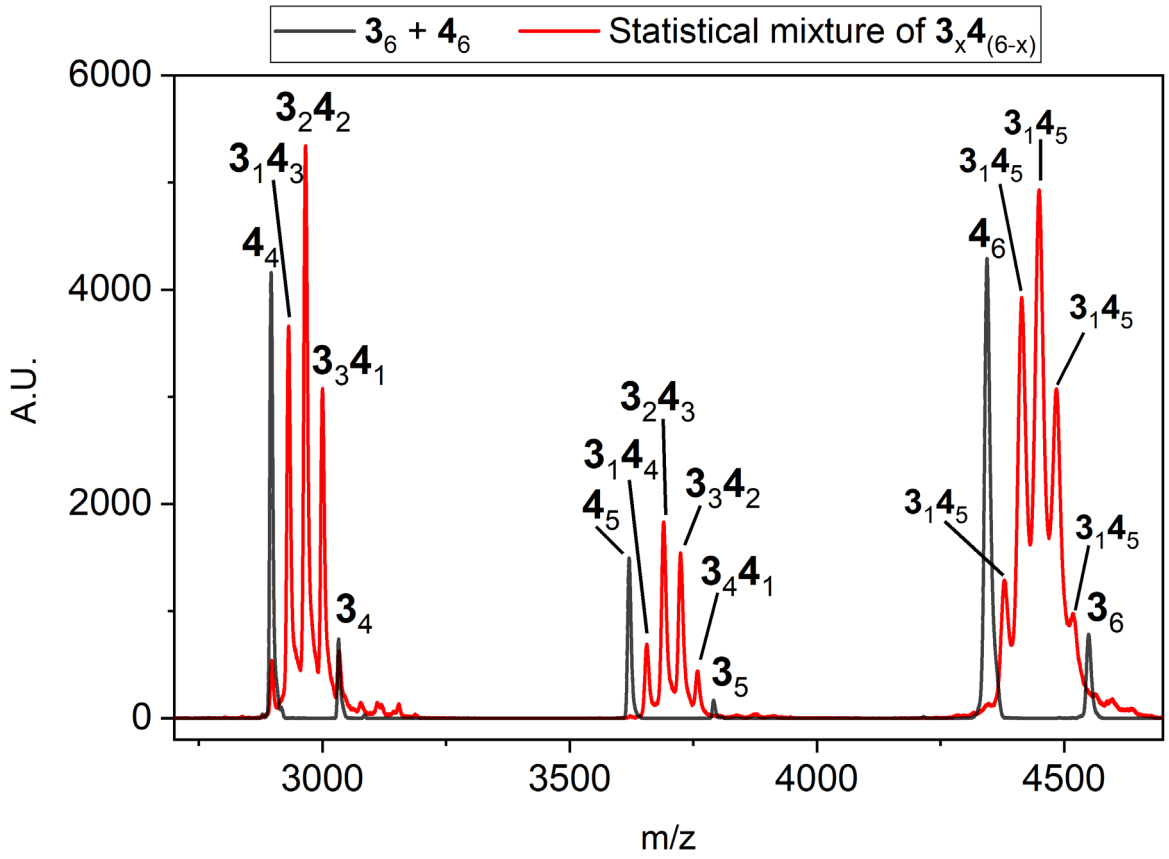


Figure S11. Partial MALDI-TOF mass spectra (showing tetramer and pentamer fragments and hexamers) of $\mathbf{3}_6$ and $\mathbf{4}_6$ spotted together (black) and a statistical mixture of hexamers obtained by mixing building blocks **3** and **4** (red). The absence of fragments containing both building blocks in the sample where $\mathbf{3}_6$ and $\mathbf{4}_6$ we spotted together shows that there is no scrambling of fragments.

In order to probe whether $\mathbf{1}_3\mathbf{2}_3$ hexamer had indeed the **1-2-1-2-1-2** sequence, we isolated $\mathbf{1}_3\mathbf{2}_3$ from the DCL by collecting fractions from the UPLC, followed by fragmentation of the disulfide bonds during MALDI-TOF analysis (efforts to fragment the disulfides directly by UPLC-MS were not successful). If we would have exclusively a $\mathbf{1}_3\mathbf{2}_3$ hexamer where the building blocks are strictly alternating, we would only expect to observe a subset of fragments. For the tetramer fragments, for example, we would expect to only observe $\mathbf{1}_2\mathbf{2}_2$. Similarly we would only expect $\mathbf{1}_1\mathbf{2}_2$ and $\mathbf{1}_2\mathbf{2}_1$ for the trimer fragments and only $\mathbf{1}_1\mathbf{2}_1$ for the dimers. MALDI-TOF analysis confirmed that $\mathbf{1}_3\mathbf{2}_3$ was isolated successfully because no other hexamer masses were observed other than the expected one (Figure S12a). In addition to the hexamer mass also masses were found that correspond to tetramer $\mathbf{1}_2\mathbf{2}_2$, trimers $\mathbf{1}_1\mathbf{2}_2$ and $\mathbf{1}_2\mathbf{2}_1$ and all three dimers: $\mathbf{1}_2$, $\mathbf{1}_1\mathbf{2}_1$ and $\mathbf{2}_3$ (Figure S12b-d). The fact that almost no fragments were observed other than the fragments expected for the alternating structure indicates that there is indeed alternation of **1** and **2** within the $\mathbf{1}_3\mathbf{2}_3$ replicator. For the dimer fragments it was still possible to observe some homodimers ($\mathbf{1}_2$ and $\mathbf{2}_3$) besides the expected heterodimer ($\mathbf{1}_1\mathbf{2}_1$), indicating that the self-sorting is not absolute. This conclusion is in line with the UPLC data which indicates that the $\mathbf{1}_3\mathbf{2}_3$ hexamer was not formed exclusively. As a comparison, also a DCL was prepared using equimolar amounts of **3** and **4**. Building block **3** is a variant of **1** that has opposite stereochemistry to **1** and contains a Leu instead of Phe. Building block **4** is a variant of **2** where the first Leu is changed into a Phe**.** These building blocks (**3** and **4**) have identical masses as **2** and **1**, respectively, but form a statistical distribution of all possible hexamers, containing the statistically expected amount of 31% $\mathbf{3}_3\mathbf{4}_3$ hexamer. The $\mathbf{3}_3\mathbf{4}_3$ hexamer fraction was isolated by UPLC and also analyzed using MALDI-TOF. For $\mathbf{3}_3\mathbf{4}_3$ much more of the

homodimer was observed than for the $\mathbf{1}_3\mathbf{2}_3$. Also fragments could be observed for tetramers $\mathbf{3}_1\mathbf{4}_3$ and $\mathbf{3}_3\mathbf{4}_1$ and trimers $\mathbf{3}_3$ and $\mathbf{4}_3$. The corresponding fragments were not found for the self-sorted $\mathbf{1}_3\mathbf{2}_3$. This indicates that for $\mathbf{3}_3\mathbf{4}_3$ the non-alternating isomers are present in a significantly larger amount than in $\mathbf{1}_3\mathbf{2}_3$ (Figure S12 and Table S5).

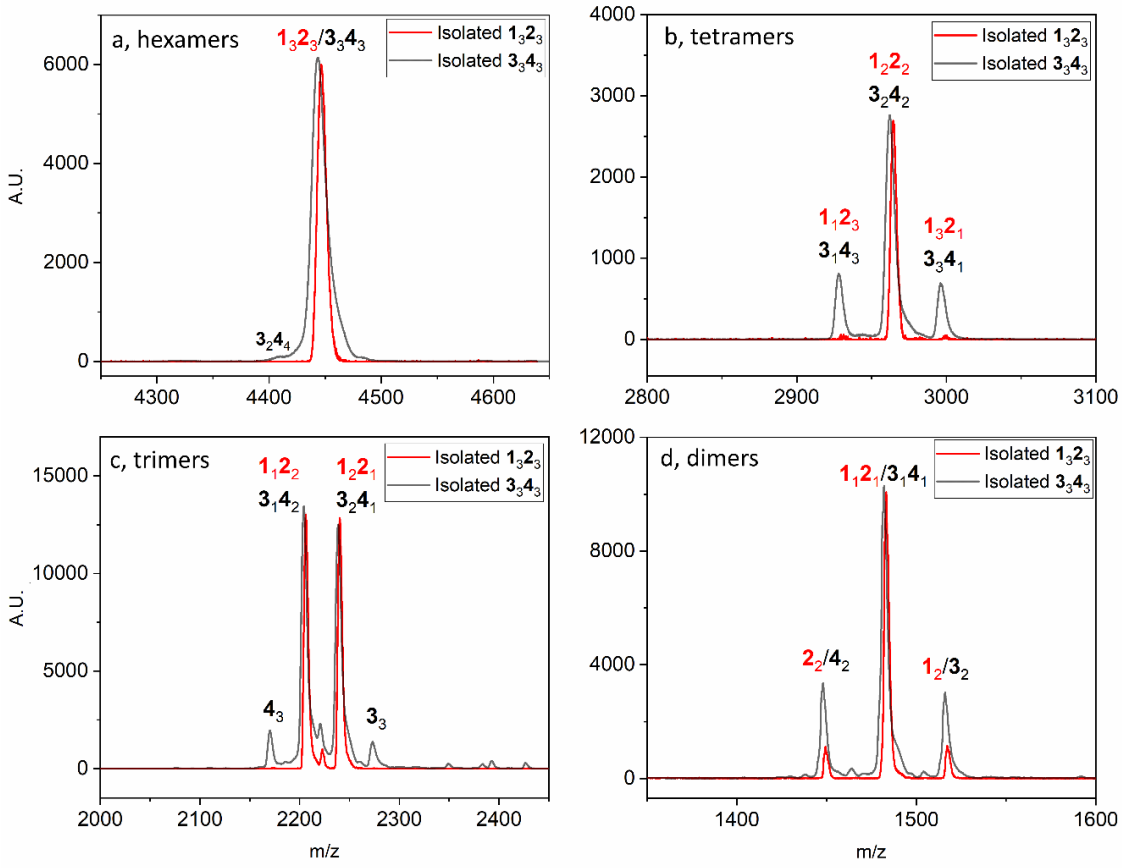


Figure S12. MALDI-TOF mass spectra of isolated $\mathbf{1}_3\mathbf{2}_3$ hexamers from a self-sorted DCL made from equimolar amounts of **1** and **2** (red) and isolated $\mathbf{3}_3\mathbf{4}_3$ hexamers from a DCL made from equimolar amounts of **3** and **4** (black) that makes a statistical distribution of all possible hexamers. The sub-spectra are normalized have a similar intensity for the middle peak(s) for each macrocycle size for easier comparison. Sub-spectrum for hexamers (a) indicates successful isolation of $\mathbf{1}_3\mathbf{2}_3$ and $\mathbf{3}_3\mathbf{4}_3$. The absence of the more homomeric fragments for tetramers (b) and trimers (c) for $\mathbf{1}_3\mathbf{2}_3$ that are present for $\mathbf{3}_3\mathbf{4}_3$ show that $\mathbf{1}_3\mathbf{2}_3$ is highly enriched in the isomer in which **1** and **2** are alternating. The presence of small signals for $\mathbf{1}_2$ and $\mathbf{2}_2$ in the dimer fragments (d) show that the self-sorting is not absolute.

Table S5. Observed areas resulting from integration of MALDI-TOF spectra of isolated $\mathbf{3}_3\mathbf{4}_3$, $\mathbf{1}_3\mathbf{2}_3$ and $\mathbf{5}_3\mathbf{3}_3$. The areas are also shown as normalized values. In the possible species column A corresponds to **1** or **3** and B corresponds to **2**, **4** or **5**.

	Isolated $\mathbf{3}_3\mathbf{4}_3$ from a statistical distribution		Isolated self-sorted $\mathbf{1}_3\mathbf{2}_3$		Isolated self-sorted $\mathbf{5}_3\mathbf{3}_3$		
Possible species	**Observed area**	**Normalized**	**Observed area**	**Normalized**	**Observed area**	**Normalized**	**Expected ratio for alternating A3B3**
A6							
A5B1							
A4B2	1376	0,01					
A3B3	110264	1,00	3286	1,00	8120	1,00	1
A2B4	1812	0,02					
A1B5							
B6							
A5							
A4B1							
A3B2	5937	1,00	7	0,88	33	0,94	1
A2B3	6456	1,09	8	1,00	35	1,00	1
A1B4							
B5							
A4							
A3B1	4912	0,22	6	0,00	28	0,02	
A2B2	22248	1,00	612	1,00	1192	1,00	1
A1B3	5289	0,24	7	0,00	19	0,02	
B4							
A3	10329	0,10	12	0,00	214	0,02	
A2B1	104673	1,00	8215	1,00	14218	1,00	1
A1B2	109404	1,05	8216	1,00	13946	0,98	1
B3	11894	0,11	20	0,00	25	0,00	
A2	14153	0,27	454	0,10	911	0,14	
A1B1	52113	1,00	4627	1,00	6568	1,00	1
B2	13839	0,27	408	0,09	697	0,11	

S10. Estimating the amplification factor (AF) of 1-2-1-2-1-2

The amplification factor for the alternating isomer was determined as follows:

In a statistical mixture the ABABAB (alternating) isomer accounts for 10% of the expected A3B3, where the AABBAB accounts for 60% and AAABBB for 30% (Figure S19). A3B3 accounts for 31.25% of all hexamers formed in a statistical distribution, which means that 3.125% of all hexamers in a statistical distribution are the ABABAB isomer.

For the analysis of the MALDI MS fragmentation data we assumed that all disulfide bonds (A-A, A-B and B-B) have the same chance of being fragmented. The expected ratios between the fragments of each fragment size was calculated for each isomer of A3B3 (Table S6).

Table S6. Expected ratios between the fragments per fragment size for each A3B3 isomer.

Expected fragments	AAABBB	AABBAB	ABABAB
A5	0	0	0
A4B1	0	0	0
A3B2	1	1	1
A2B3	1	1	1
A1B4	0	0	0
B5	0	0	0
A4	0	0	0
A3B1	1	1	0
A2B2	1	4	1
A1B3	1	1	0
B4	0	0	0
A3	1	0	0
A2B1	2	1	1
A1B2	2	1	1
B3	1	0	0
A2	1	1	0
A1B1	1	4	1
B2	1	1	0

For $\mathbf{1}_3\mathbf{2}_3$ the experimentally determined ratio of the fragments matches the expected ratio for ABABAB very closely, except for the dimer fragments. The observed normalized ratio is 0,10 : 1,00 : 0,09 (A2 : A1B1 : B2) where the expected ratio is 0 : 1 : 0.

We assume that the AAABBB isomer is not formed in significant amounts in this system, which can be justified because ABABAB is formed preferentially because of an extra stability for A-B (compared to A-A and B-B) and the AAABBB isomer has the smallest amount of A-B.

Based on the expected ratios for the dimer fragments for AABBAB and ABABAB a linear interpolation between them was made. When 50% of all A3B3 would be AABBAB isomer and the other 50% would be ABABAB isomer, the expected fragmentation pattern would be 0,10 : 1,00 : 0,10. This most closely matches the observed ratio of 0,10 : 1,00 : 0,09.

As A3B3 accounts for 70% of all the 6mers that are formed and the ABABAB isomer accounts for 50% of the A3B3, this means that 35% of all hexamers that are formed are the ABABAB isomer, which results in an AF of 11.2 compared to the statistical distribution.

S11. Varying Stochiometry of 1 and 2

To experimentally confirm if there is indeed an additional secondary interaction between **1** and **2** within a macrocycle, we prepared a series of DCLs with the same total amount of building blocks, but with an increasing molar ratio of **1**, ranging from 0 to 100 mol%. In the absence of any specific secondary interactions within the macrocycle we would expect to find a distribution of different hexamer compositions for each building block ratio that is governed by statistics (Figure S13a). Instead we observed the preferential formation of $\mathbf{1}_3\mathbf{2}_3$ when the building block fraction is around 50%. Other noticeable deviations from statistically expected amounts are observed for $\mathbf{1}_5\mathbf{2}_1$ at high fractions of **1** and $\mathbf{1}_2\mathbf{2}_4$ at low fractions of **1**. These results are consistent with the formation of a secondary interaction between **1** and **2** within the macrocycle. The asymmetry of this figure could be explained by incomplete self-sorting at high fractions of **2**. In those DCLs there is still material in solution (small macrocycles) that is not incorporated into the assembly, which contains mostly **2**. Additionally, it is also beneficial for a replicator to incorporate more **1**, because **1** has a higher tendency to self-assemble than **2** (see also Section S16). This could explain why $\mathbf{1}_5\mathbf{2}_1$ is preferentially formed at high fractions of **1**, where $\mathbf{1}_2\mathbf{2}_4$ is the preferred macrocycle composition at high fractions of **2**.

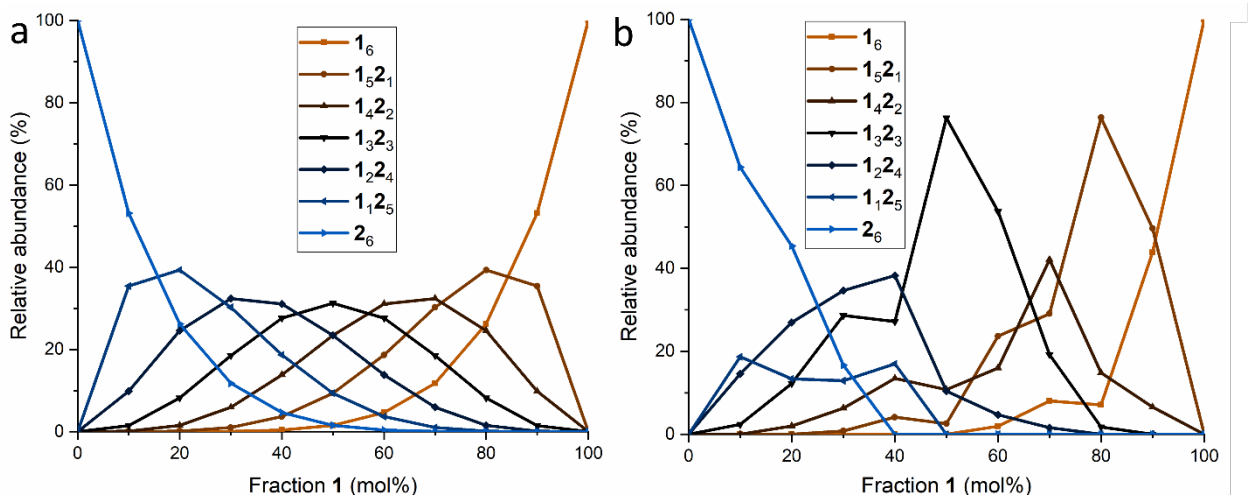


Figure S13. (a) Expected distribution of hexamer compositions for different fractions of **1** and **2** when the incorporation of the building blocks would only be governed by statistics. (b) Observed distribution of hexamer compositions for different fractions of **1** and **2** in a DCL in a closed vial. The observed distribution is different from the expected distribution, suggesting that there are specific interactions between the building blocks.

S12. Solid-state NMR analysis of self-replicators

To perform the solid-state NMR analysis, building block **1** was prepared with the incorporation of a doubly labeled ($^{13}C_6$,^{15}N) L-leucine. A DCL dominated by the $\mathbf{1}_3\mathbf{2}_3$ hexamer was prepared and the resulting material was analyzed by magic-angle-spinning (MAS) solid-state NMR. A single Cα peak was observed in the ^{13}C cross-polarization (CP) 1D spectrum (Figure S14a), indicative of a highly ordered and highly symmetric assembly. A broader peak, revealing three main components, was identified in the Cβ region of the same spectrum, highlighting the existence of three different chemical environments for the leucine residue of **1**.

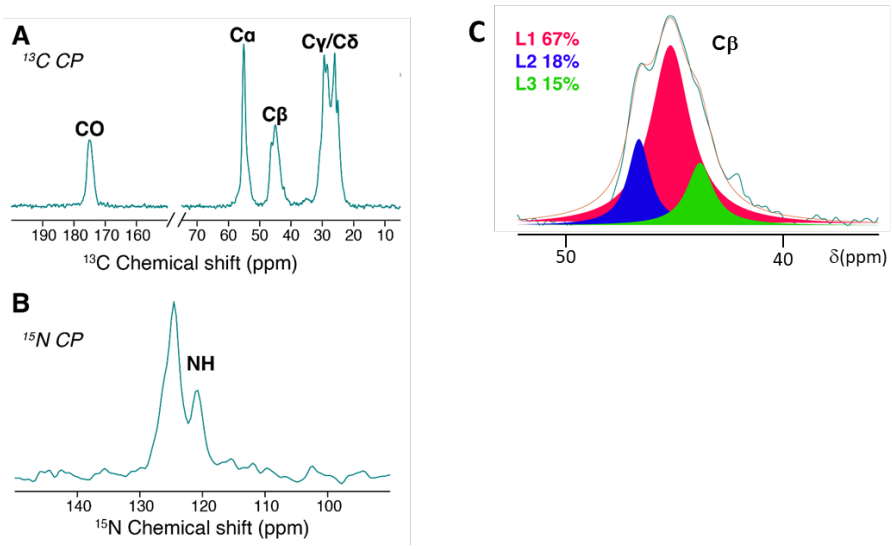


Figure S14. (a) ^{13}C and (b) ^{15}N 1D CP/MAS NMR spectra of the self-sorted $\mathbf{1}_3\mathbf{2}_3$-replicator, with building block **1** containing doubly labelled L-leucine ($^{13}C_6$,^{15}N). (c) Deconvolution of the three spin systems L1-L3 in the ^{13}C signal of the leucine β carbon, reflecting the co-existence of three distinct conformational states of the labeled leucine residue.

The assignment of the resonances of the three overlapping spin systems, indicated as L1, L2 and L3, was performed with the aid of a 2D ^{13}C-^{13}C dipolar assisted rotational resonance (DARR) spectrum (Figure S15a). At 25 ms ^{13}C-^{13}C mixing, predominantly one-bond correlations were observed, along with some two-bond correlations. We focused our attention mainly on one side of the diagonal of the highly asymmetric spectrum. Here, Cα-Cβ and Cβ-Cγ correlations could be easily identified for all three leucine conformers, together with a two-bond correlation Cβ-Cδ1 attributed to L3. When examining the other side-chain shifts, only two cross-peaks could be clearly distinguished, i.e. L1 Cγ-Cδ1 and L2 Cγ-Cδ1. Other cross-peaks fall close to the diagonal, so their assignment is not unambiguous. For this reason, we recorded a 2D double-quantum single-quantum (DQ-SQ) ^{13}C-^{13}C spectrum, which is better suited for resolving correlations between carbons with similar chemical shifts (Figure S15b). With this experiment we could assign L1(2) Cδ2 and L3 Cγ and Cδ2.

The values of the chemical shifts of the Cα, Cβ and CO carbons of all three spin systems are indicative of the formation of a β-sheet assembly, as evident from negative chemical shift differences Cα - Cβ, with respect to their corresponding random coil values.[11,12] The values of these chemical shift differences are -3.1 ppm, -4.1 ppm and -2.3 ppm for forms L1, L2 and L3, respectively. The maximum error on the chemical shift determination is +/- 0.2 ppm.

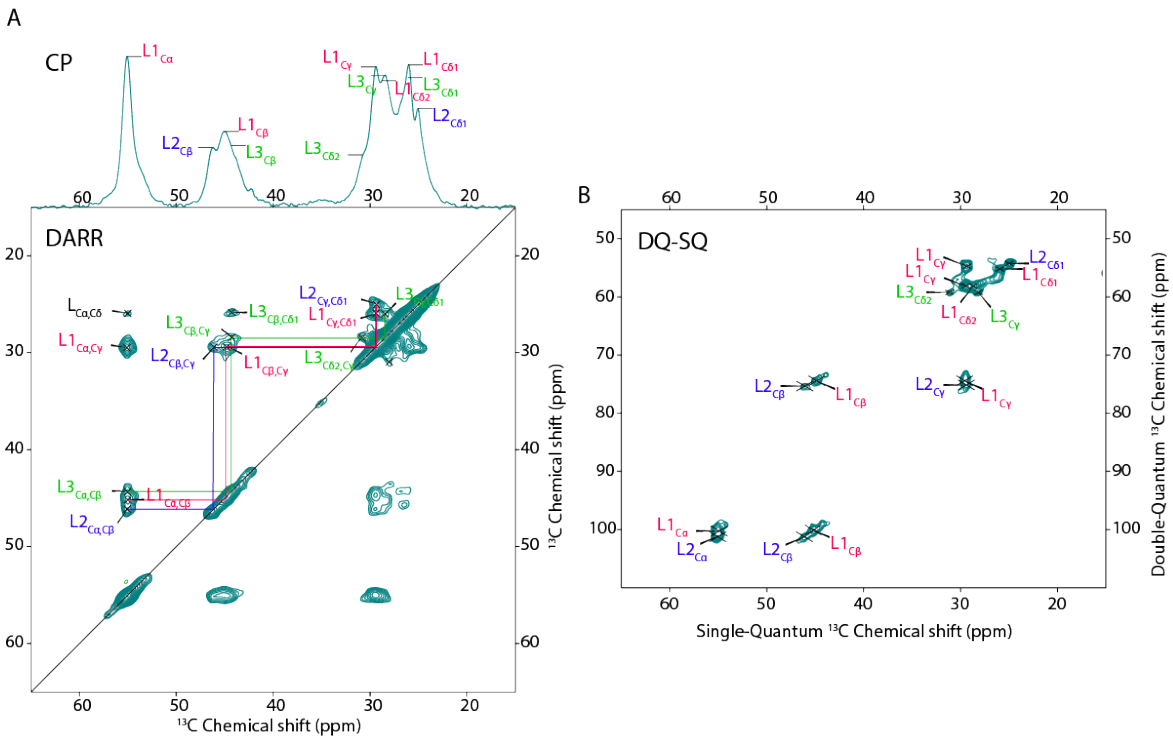


Figure S15. Chemical shift assignment and secondary structure propensity of the self-sorted $\mathbf{1}_3\mathbf{2}_3$-replicator. (a) ^{13}C 1D CP/MAS NMR spectrum and 2D ^{13}C-^{13}C DARR spectrum with DARR ^{13}C-^{13}C mixing time of 25 ms, (b) 2D ^{13}C-^{13}C DQ-SQ spectrum under 10 kHz MAS with 0.8 ms total recoupling time (0.4 ms each for excitation and reconversion). Peaks attributed to the three leucine conformers are marked with L1, L2 and L3 (see also Figure S14).

Figure S16a presents a comparison of the ^{13}C 1D CP/MAS spectrum with a ^{13}C 1D direct excitation (DE)/MAS spectrum. Only minor differences were observable, suggesting that the three leucine conformers have similar dynamics, since CP and DE experiments differ in their sensitivity to dynamics.[13] The 1D ^{13}C INEPT (insensitive-nuclei enhanced-polarization transfer) experiment, which reveals flexible segments, gave no signal after comparable amount of time .[13] Thus, these MAS NMR data indicate the (leucine within) the fibrils to be ordered and relatively rigid.

To probe for differences in water accessibility, we conducted water-edited solid-state NMR experiments (Figure S16b).[6] In this technique, water ^{1}H polarization is transferred to the protons in an (immobilized) molecule of interest and detected via monitoring the ^{13}C signals of the latter. Here, we implemented the 1D version of this experiment, where a ^{1}H T_2 (transverse relaxation time) filter is first used to create a ^{1}H polarization gradient. We retain the polarization from dynamic water molecules (with longer T_2 values), while the polarization of other molecules with shorter T_2 is destroyed. A subsequent ^{1}H-^{1}H mixing time allows the water polarization to be transferred to the other molecules, before a CP-based detection occurs. Faster water polarization transfer was observed to α and β carbons compared to γ and δ carbons indicating that the latter are shielded from contact with water, presumably through being engaged in hydrophobic interactions.

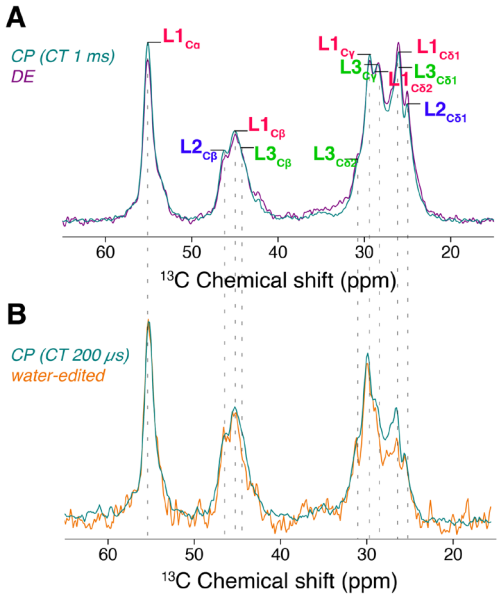


Figure S16. Dynamic comparison of the three leucine conformers and hydration obtained from water-edited MAS NMR. (a) ^{13}C 1D CP/MAS and DE/MAS NMR spectra of self-sorted $\mathbf{1}_3\mathbf{2}_3$-replicator, with building block **1** containing doubly labelled L-leucine ($^{13}C_6$,^{15}N). The lack of differences in peak intensities suggest a similar dynamics for each of the three conformers. (b) ^{13}C 1D CP/MAS, and ^{13}C 1D water-edited CP/MAS experiment performed with an initial ^{1}H T_2 (transverse relaxation time) filter of 0.66 ms and subsequent longitudinal ^{1}H-^{1}H mixing time of 7 ms. The water-edited spectrum is normalised to the $L1_{C\alpha}$ peak of the CP spectrum to facilitate comparison of the relative peak intensities. The relative intensity differences among the different carbon peaks indicate the burial of the methyl groups.

S13. Simulating expected hexamer distributions

Obtaining the distribution of all hexamer mutants in the absence of self-assembly

The basal distribution of all the different structural isomers of all the different sizes of macrocycles was obtained by first enumerating all the different possible oxidation and disulfide exchange reactions that can occur with two different types of monomeric building blocks (labelled as A and B) and then those reactions were integrated into a mass-action kinetics model and the system was simulated, as explained below.

Starting from two types of monomers (A and B) all the different oxidation reactions were enumerated up to a linear species of length 3 (for example, A+B→AB, or AB+A→ABA). Second, all linear species of length 3 (trimers) were assumed to instantaneously close to form cyclic species as an intermolecular oxidation reaction. Third, disulfide exchange between cyclic and linear species was treated as an S_N2 reaction[14] that can also lead to cyclic products of different size than the reactant, by accounting for all the different possible pathways resulting from linear species attacking the cyclic species (Figure S17).

A mass-action kinetics model was constructed by collecting all those reactions into ordinary differential equations representing the rates at which the concentrations of the different isomers change and including $NaBO_3$ as an oxidizing agent.

This kinetics model was simulated deterministically using MATLAB's ode15s numerical integrator, until the system reached equilibrium (Figure S18). The rate-constant for each individual attack by a linear species on a disulfide within a cyclic species is assumed to be identical for all reactions and species. The latter means that, because larger cyclic species have more dithiol cores, larger cyclic species will be exchanged faster.

Simulating the model gave rise to different proportions of isomers of the same A_xB_{6-x} composition (Figure S19 left). The outcome was validated by when grouping together hexamers of same composition which yielded a distribution that matches Pascal's Triangle, as expected for statistical incorporation of building blocks (Figure S19 right).

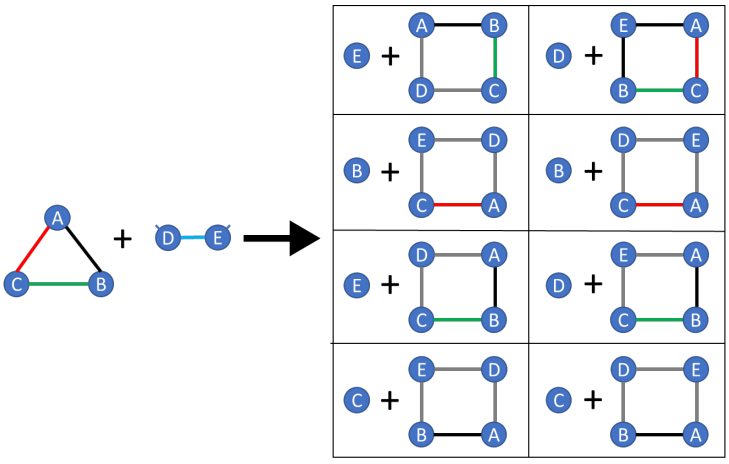


Figure S17. A heuristic example how the exchange pathways were modeled (A-E represent different building blocks). This figure depicts the possible outcomes of a single attack by the linear dimer DE on the A part of the cyclic ABC leading to a cyclic tetramer. This calculation assumes the intermediacy of a 5-long linear species. Disulfide bonds are coloured for brevity.

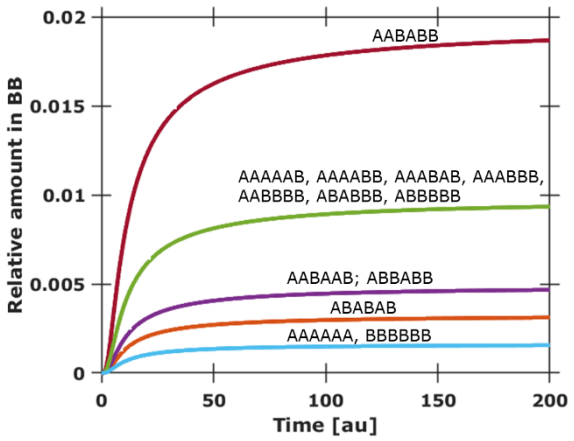


Figure S18. Hexamer isomers emergence kinetics using the mass-action model. Initial concentrations were 50 [arbitrary units] for A and B monomers, 100 [arbitrary units] for $NaBO_3$, and 0 for the other species. The following rate constants we used for the simulations: k_{ox}=1e-3 1/(C*time) and k_{ex}=1e1 1/(C*time).

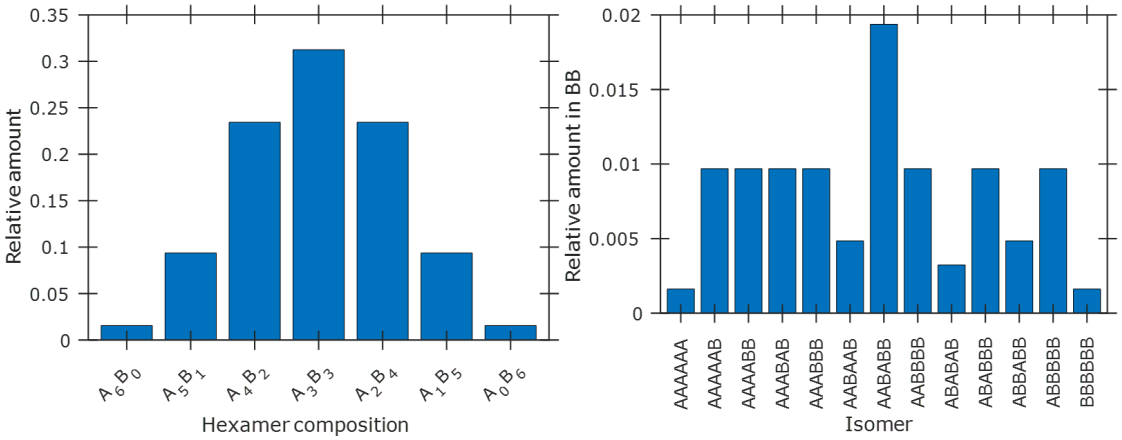


Figure S19. Hexamer and isomers distributions taken at the end of the simulation (Figure S18). (left) Grouped per isomer composition. (right) Individual isomers.

<u>Modeling the hexamer distribution in the presence steric zipper interactions</u>

To evaluate how the mutant distribution of hexamer isomers changes as a result of steric zipper interactions, the mass-action kinetics model described above was expanded. When two neighbouring building blocks within a hexamer sterically complement each other (Figure 1c in the main text), the relevant disulfide bond (A-B type of bond, as oppose to A-A or B-B) was assumed to be more shielded which should diminish the rate of the relevant disulfide bond exchange reaction. To represent this steric hindrance in the kinetics model, the rate constant of all exchange pathways that involves the breakage of an A-B disulfide bond within a hexamer were reduced.

When simulating this expanded model, a narrower mutant distribution was obtained (Figure S20 left), in line with the experimental results (Figure 3f the main text). Closer inspection reveals that not all structural isomers are affected in the same way (Figure S20 right). The isomer ABABAB, which for the composition A_3B_3 is the least abundant one in the basal case (Figure S19 right), became significantly more abundant when the rate constant for A-B cleavage was reduced, until the other two isomers with the same A_3B_3 composition (AAABBB and AABABB) virtually disappeared. This shift towards ABABAB is because in this isomer all the six disulfide bonds are of the A-B type and therefore this macrocycle is least reactive toward disulfide exchange.

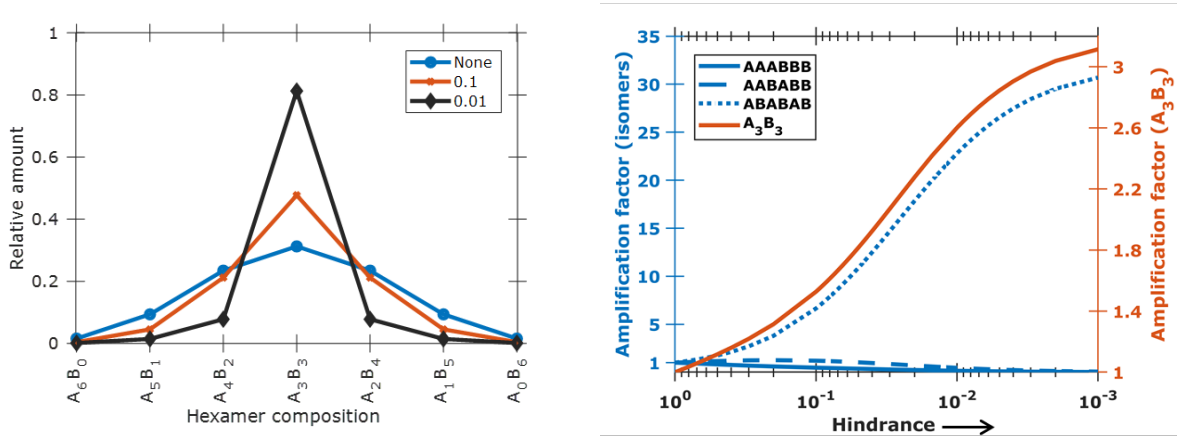


Figure S20. Hexamer distributions under increasing steric zipper effect, represented by a relative decrease of disulfide exchange rate constant for disulfide bonds connecting building block A to B. (left) Examples of the hexamer distribution without a zipper effect (blue; data identical to that in Figure S19) and with a zipper effect, where the rate constant for A-B cleavage was reduced by a factor 10 (red) or 100 (black). (right) Calculated amplification factor of the three isomers of A_3B_3 as a function of increasing efficiency of steric zipper formation.

S14. UPLC data for different stoichiometries of 1 and 2

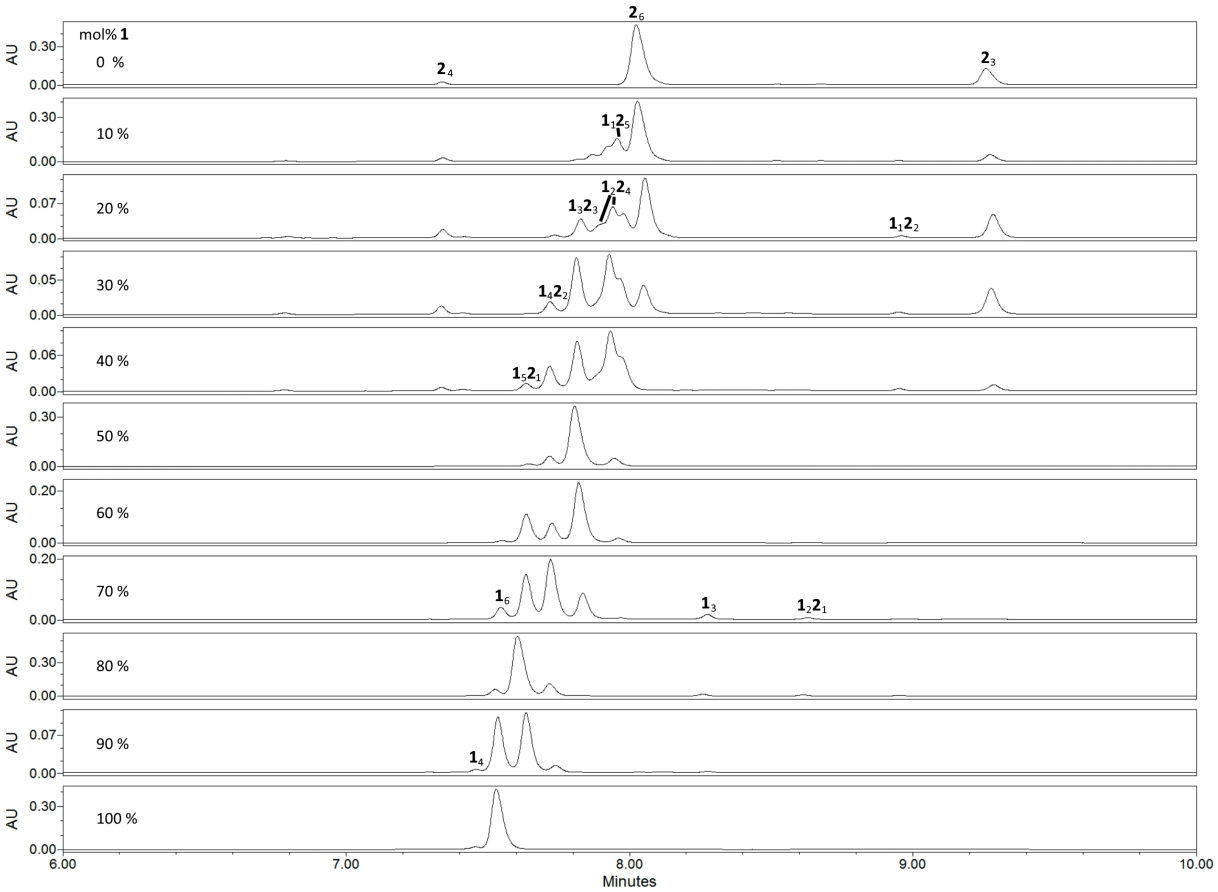


Figure S21. RP-UPLC traces (monitored at 254 nm) of the product mixtures obtained from DCLs made from various molar ratios of **1** and **2**, ranging from 0% to 100% **1**, at 45 °C after 7 days.

S15. Repeats of single building block (cross-)seeding experiments

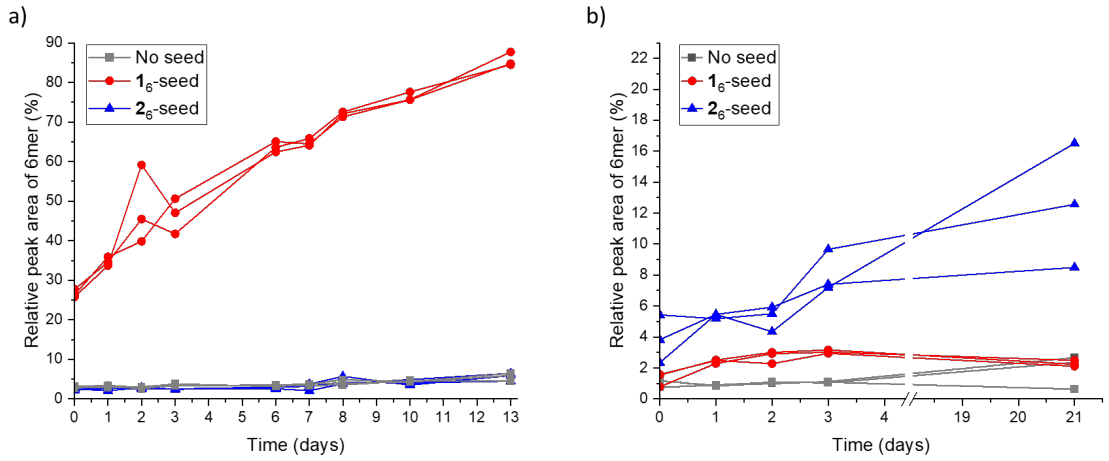


Figure S22. Kinetics of hexamer formation based on UPLC-MS analysis on (a) DCLs made from **1** (2.0 mM in 100 mM B_2O_3 buffer) seeded with 10% preformed $\mathbf{1}_6$, $\mathbf{2}_6$ or no seed at room temperature without stirring and (b) DCLs made from **2** (2.0 mM in 100 mM B_2O_3 buffer) seeded with 10% preformed $\mathbf{1}_6$, $\mathbf{2}_6$ or no seed at 45 °C without stirring.

S16. Seeding experiments on the mixed replicator made from 1 and 2

We prepared two DCLs containing equimolar amounts of **1** and **2** (50 µM each in 25 mM B_2O_3 buffer, pH 8.50). These DCLs were stirred at 40 °C with a low stirring speed (400 rpm)[1] to minimize the fragmentation of the fibers, but ensuring that the DCL is homogeneous to allow analysis by RP-UPLC(-MS). The DCLs were oxidized to 75-80% disulfides using a sodium perborate solution (40 mM). Just before the first measurement 5.0 mol% of preformed hexameric replicators were added: self-sorted $\mathbf{1}_3\mathbf{2}_3$ was added to one and $\mathbf{1}_6$ to the other DCL. After 12 hours a total growth of 32% (in total peak area) was observed for the DCL that was seeded with $\mathbf{1}_3\mathbf{2}_3$. The $\mathbf{1}_5\mathbf{2}_1$, $\mathbf{1}_4\mathbf{2}_2$ and $\mathbf{1}_3\mathbf{2}_3$ were found to be the compositions that grew the most (Figure S23a). The DCL that was seeded with $\mathbf{1}_6$ showed a growth of only 12% after 12h. Also here the $\mathbf{1}_5\mathbf{2}_1$, $\mathbf{1}_4\mathbf{2}_2$ and $\mathbf{1}_3\mathbf{2}_3$ were the compositions that showed the most growth, but the difference with $\mathbf{1}_2\mathbf{2}_4$ was significantly smaller (Figure S23b). These experiments show a similar trend as observed in the previous seeding experiment (Figure 3d), where the **1**-rich hexamers grow most readily at the start of the experiment. The large difference in the total growth upon addition of different seeds does, however, show that the composition of the seed has a substantial influence on the replication process. Replication occurs with a relatively high mutation rate, although some bias due to the nature of seed persists. Upon seeding with $\mathbf{1}_3\mathbf{2}_3$ this macrocycle accounted for 7,4% of the newly formed hexamers, compared to 2,4% when seeded with $\mathbf{1}_6$. The fact that especially $\mathbf{1}_5\mathbf{2}_1$, $\mathbf{1}_4\mathbf{2}_2$ and $\mathbf{1}_3\mathbf{2}_3$ showed an increased growth when seeded with $\mathbf{1}_3\mathbf{2}_3$ rather than $\mathbf{1}_6$, suggest that the macrocycle composition is also (partially) replicated.

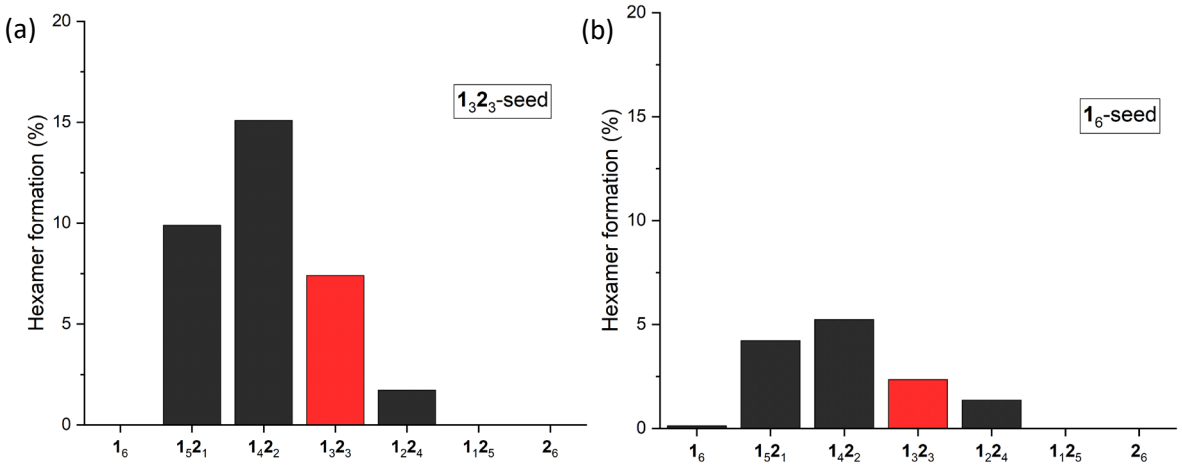


Figure S23. Formation of hexameric replicators (subtraction of peak area of initial composition from the composition found after 12h) with minimal stirring (400 rpm) at 40 °C with the addition of 5.0 mol% preformed seed at the start of the experiment. (a) $\mathbf{1}_3\mathbf{2}_3$ and (b) $\mathbf{1}_6$ were used as seeds. A total growth of 32% and 12% was found for (a) and (b), respectively.

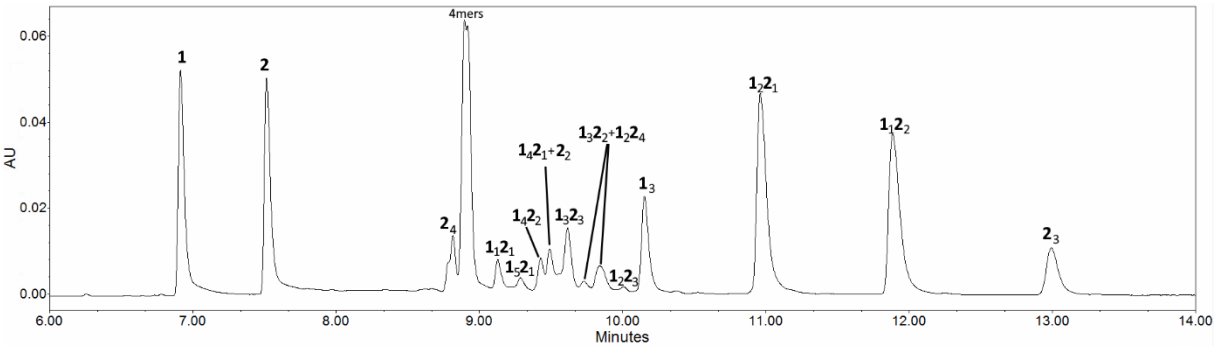


Figure S24. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from a DCL of **1** and **2** after sodium perborate oxidation and addition of the $\mathbf{1}_3\mathbf{2}_3$-seed. This chromatogram corresponds to t = 0 h in the seeding experiment using the $\mathbf{1}_3\mathbf{2}_3$-seed.

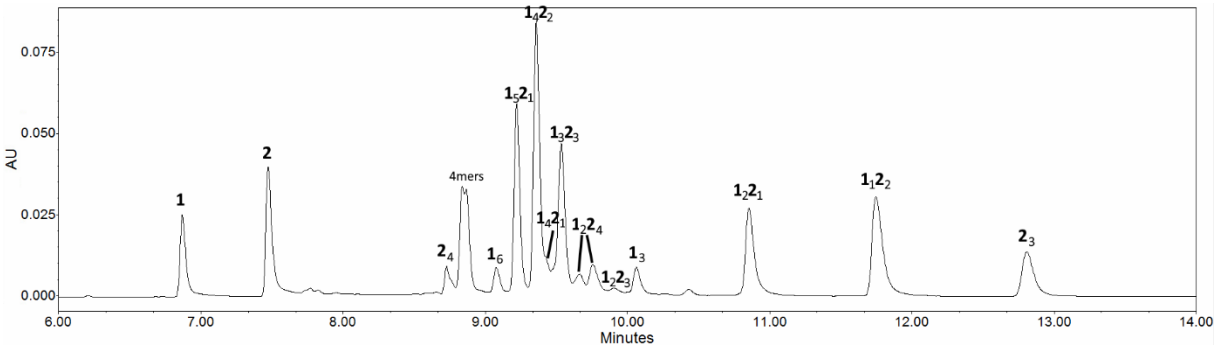


Figure S25. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from a DCL of **1** and **2** with $\mathbf{1}_3\mathbf{2}_3$-seed after 12h at 40 °C with 400 rpm stirring. This chromatogram corresponds to t = 12h in the seeding experiment using the $\mathbf{1}_3\mathbf{2}_3$-seed.

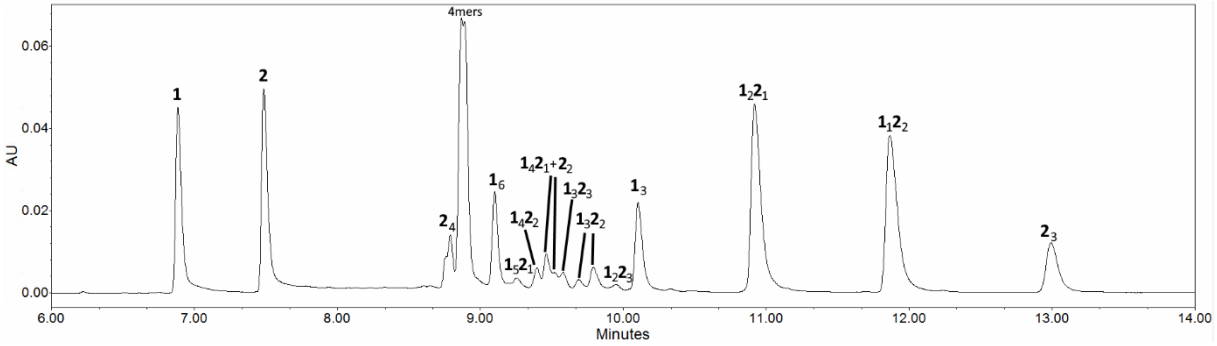


Figure S26. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from a DCL of **1** and **2** after sodium perborate oxidation and addition of the $\mathbf{1}_6$-seed. This chromatogram corresponds to t = 0 h in the seeding experiment using the $\mathbf{1}_6$-seed.

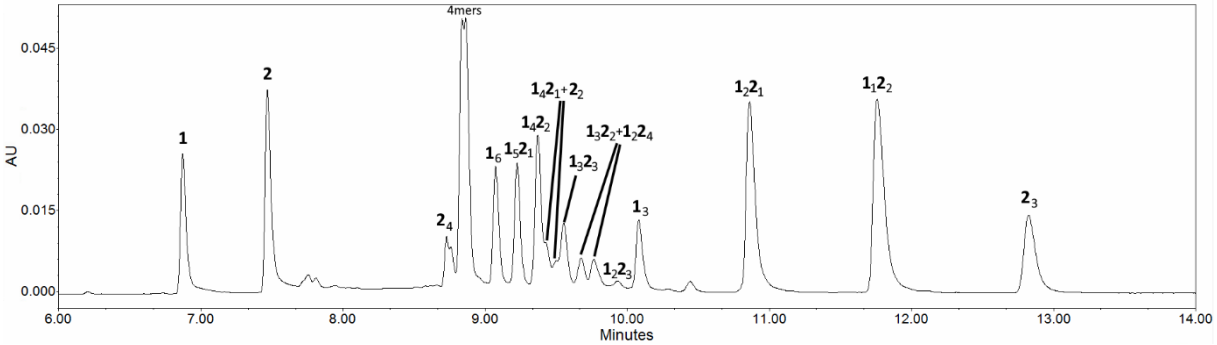


Figure S27. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from a DCL of **1** and **2** with $\mathbf{1}_6$-seed after 12h at 40 °C with 400 rpm stirring. This chromatogram corresponds to t = 12 h in the seeding experiment using the $\mathbf{1}_6$-seed.

Table S7. Composition of the DCLs during a seeding experiment when 5.0 mol% $\mathbf{1}_3\mathbf{2}_3$ is used as a seed. All compositions are listed as their percentage of the total peak area. All peak area that is not listed in this table can be assigned to tetramers, pentamers and dimers.

Time (h)	Total peak area (AU)	**1**	**2**	$\mathbf{1}_6$	$\mathbf{1}_5\mathbf{2}_1$	$\mathbf{1}_4\mathbf{2}_2$	$\mathbf{1}_3\mathbf{2}_3$	$\mathbf{1}_2\mathbf{2}_4$	$\mathbf{1}_3$	$\mathbf{1}_2\mathbf{2}_1$	$\mathbf{1}_1\mathbf{2}_2$	$\mathbf{2}_3$
0	1504303	11,83	11,84	0,00	1,03	1,55	3,79	2,97	5,71	15,66	14,33	4,35
1	1541662	9,38	11,43	0,00	1,62	2,63	4,29	3,20	5,07	15,16	15,18	4,96
2	1568753	8,29	10,87	0,00	3,29	4,93	5,25	3,43	4,52	14,04	14,75	5,10
3	1589064	7,62	9,92	0,00	4,80	7,04	6,06	3,70	4,04	12,79	14,03	5,09
4	1586622	7,27	9,83	0,00	6,14	8,98	7,54	3,87	3,66	11,90	13,21	5,15
5	1589706	6,87	9,61	0,00	7,12	10,32	8,04	4,00	3,30	11,08	12,90	5,08
6	1579725	6,54	9,54	0,00	8,00	11,64	8,65	4,11	3,09	10,33	12,22	5,11
7	1587189	6,26	9,37	0,00	8,74	12,79	9,17	4,18	2,86	9,71	11,84	5,12
8	1624884	5,87	9,04	0,00	9,24	13,81	9,45	4,23	2,68	9,19	12,06	5,07
9	1679237	5,41	8,65	0,00	9,56	13,93	9,72	4,45	2,67	8,52	10,72	4,88
10	1591830	5,43	8,74	0,00	10,17	15,15	10,31	4,44	2,44	8,55	11,15	5,12
11	1597109	5,16	8,49	0,00	10,61	15,97	10,78	4,57	2,33	8,10	10,93	5,13
12	1588463	4,98	8,44	0,00	10,91	16,64	11,19	4,69	2,18	7,81	10,46	5,17

Table S8. Composition of the DCLs during a seeding experiment when 5.0 mol% $\mathbf{1}_6$ is used as a seed. All compositions are listed as their percentage of the total peak area. All peak area that is not listed in this table can be assigned to tetramers, pentamers and dimers.

Time (h)	Total peak area (AU)	**1**	**2**	$\mathbf{1}_6$	$\mathbf{1}_5\mathbf{2}_1$	$\mathbf{1}_4\mathbf{2}_2$	$\mathbf{1}_3\mathbf{2}_3$	$\mathbf{1}_2\mathbf{2}_4$	$\mathbf{1}_3$	$\mathbf{1}_2\mathbf{2}_1$	$\mathbf{1}_1\mathbf{2}_2$	$\mathbf{2}_3$
0	1482083	10,22	11,87	4,95	0,99	1,12	0,99	2,26	5,55	15,81	14,88	4,83
1	1497236	8,99	11,03	4,84	1,36	1,50	1,11	2,42	5,24	15,59	15,77	5,25
2	1506325	8,69	10,41	4,92	1,76	2,01	1,33	2,64	4,96	15,13	15,75	5,41
3	1463142	8,08	10,26	4,93	2,24	2,46	1,55	2,77	4,81	14,83	15,34	5,52
4	1474573	7,73	10,09	4,97	2,59	2,98	1,78	2,89	4,58	14,36	15,17	5,58
5	1461559	7,43	9,95	5,00	2,94	3,34	1,96	2,98	4,41	13,98	15,43	5,63
6	1438663	7,16	9,79	5,17	3,33	3,91	2,20	3,16	4,27	13,69	15,35	5,70
7	1427302	6,90	9,68	5,05	3,69	4,35	2,40	3,23	4,13	13,36	14,81	5,77
8	1422320	6,72	9,59	4,98	3,93	4,78	2,53	3,21	3,93	13,17	15,29	5,90
9	1511903	6,02	8,93	5,15	4,37	5,11	2,80	3,44	3,80	12,01	14,10	5,53
10	1398906	6,27	9,25	4,98	4,59	5,62	2,90	3,39	3,71	12,60	15,03	6,00
11	1378968	6,11	9,09	5,04	4,91	6,19	3,09	3,44	3,62	12,47	14,66	6,11
12	1478620	5,50	8,60	5,07	5,21	6,36	3,34	3,62	3,51	11,20	13,88	5,75

S17. Repeats of the experiments shown in Figure 3

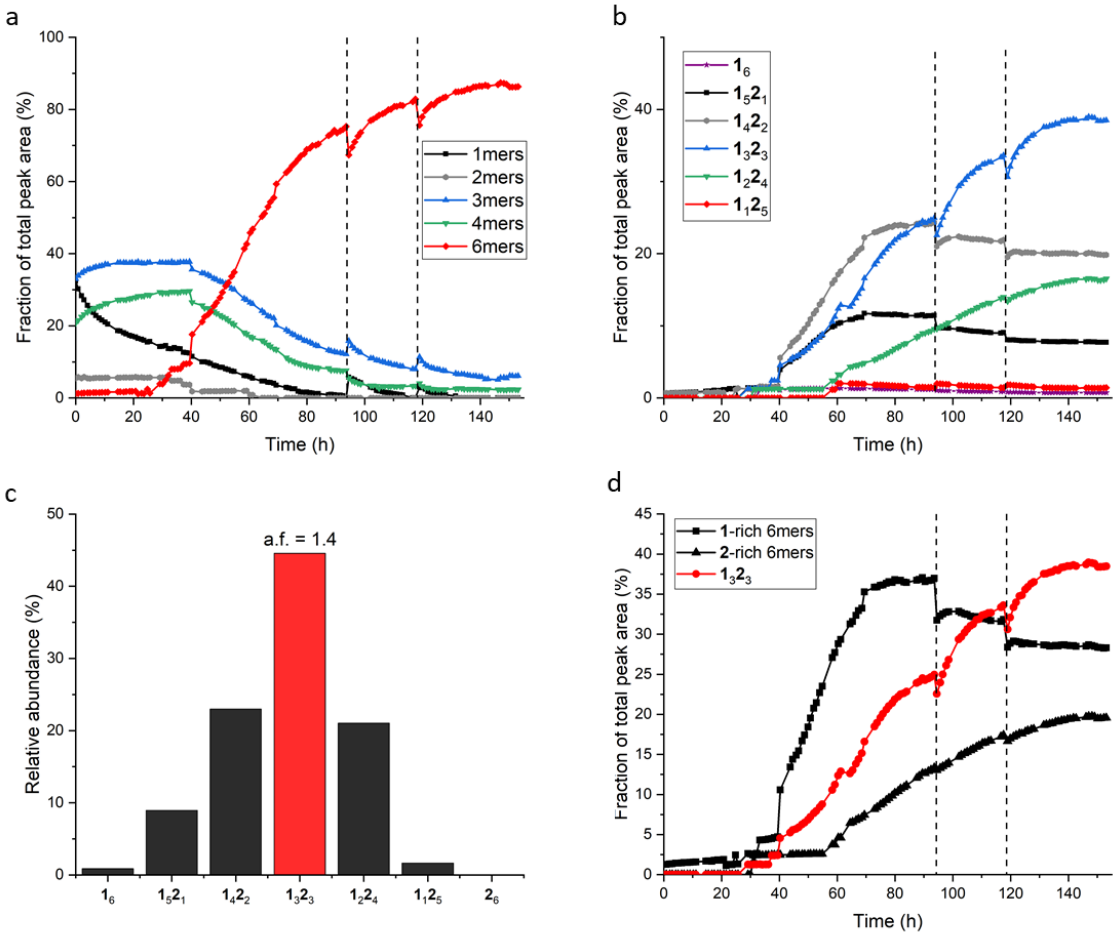


Figure S28. Repeat of the experiments shown in Figure 3. (a) Changes in the distribution of ring sizes in a DCL (agitated by stirring at 1200 rpm at 40 °C) made from equimolar amounts of **1** (1.0 mM) and **2** (1.0 mM). (b) Changes in the amounts of the different hexamer rings in the DCL in panel a. The dotted vertical lines indicate the addition of a mixture of 4.2% **1** and 4.2% **2** to ensure continued disulfide exchange. (c) Relative abundance of hexameric replicators at the end of the experiment in panels a and b. The distribution is enriched in $\mathbf{1}_3\mathbf{2}_3$ with an AF of 1.4; (d) Changes in the distributions of **1**-rich hexamers, **2**-rich hexamers and $\mathbf{1}_3\mathbf{2}_3$-hexamer showing the emergence of the self-sorted $\mathbf{1}_3\mathbf{2}_3$ mixed replicator at the expense of the **1**-rich hexamers.

S18. Repeat of mass-flow experiment shown in Figure 4

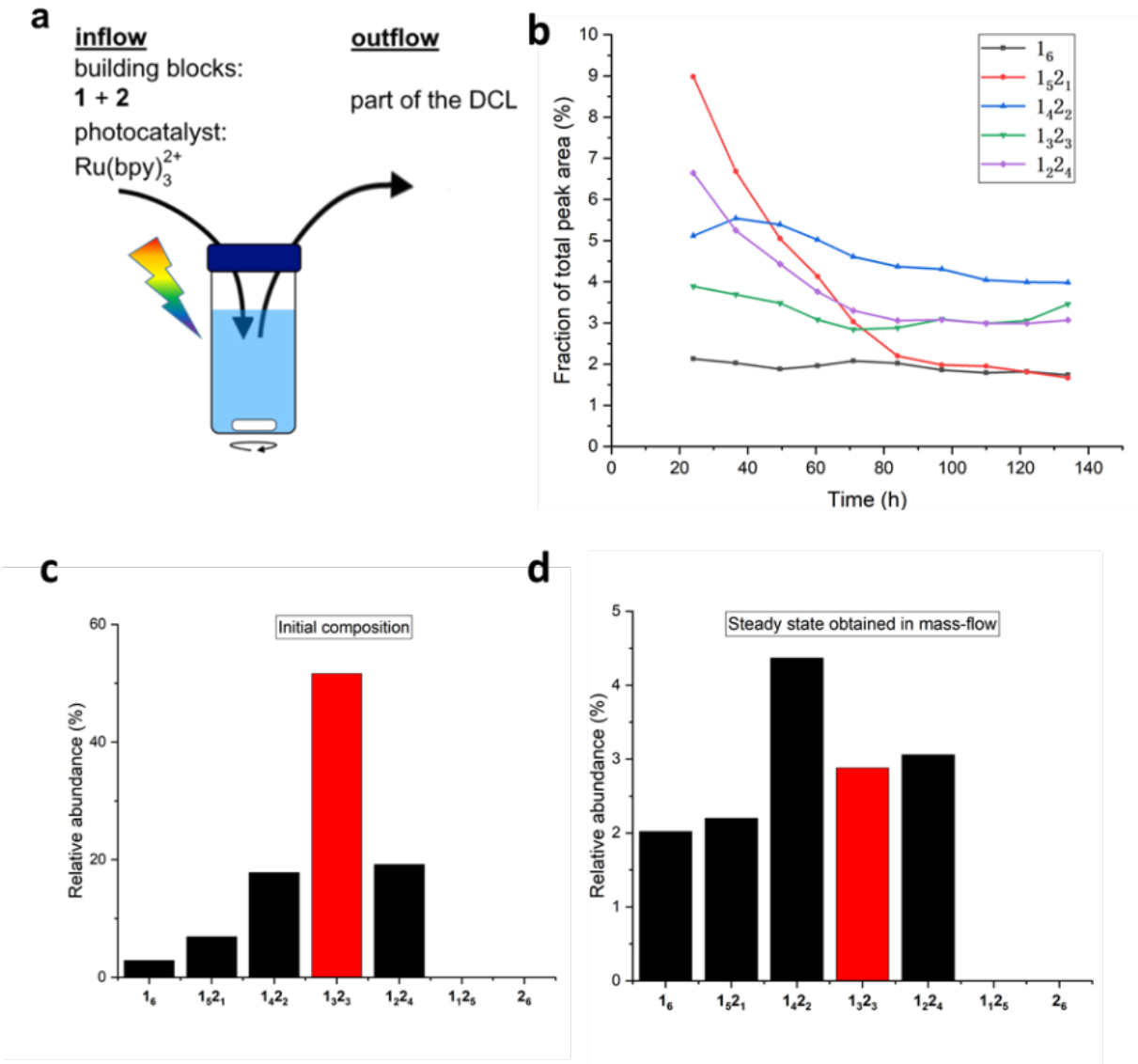


Figure S29. Repeat of the experiment shown in Figure 4. Replicator selection in an out-of-equilibrium regime implemented by mass-flow. (a) Schematic depiction of the flow set-up. (b) Change in relative abundances of hexameric self-replicators with different building block compositions over the course of the flow-experiment. The proportion of self-replicators (especially $\mathbf{1}_3\mathbf{2}_3$) decreased from >90% at the start of the experiment to ~20% at the steady state. (c) Distribution of hexameric self-replicators at the start of the experiment and (d) the corresponding distribution at the steady state.

S19. Selectivity of TCEP reduction of $1_n2_{(6-n)}$ replicators

To investigate the selectivity of TCEP reduction a DCL made from equimolar amounts of **1** and **2** (1.0 mM each; in 100 mM B_2O_3-buffer, pH 8.20), that already produced a self-sorted distribution of hexameric replicators dominated by $\mathbf{1}_3\mathbf{2}_3$, was partially reduced by addition of various amounts of a concentrated TCEP solution (20 mM in 100 mM B_2O_3-buffer, pH 8.20). The difference between the relative abundances of hexamers with different building block composition before (Figure S30a) and after TCEP reduction is shown in Table S9. The relative change in abundance is calculated as follows: (change in abundance)/(relative abundance initial distribution)*100%. The data suggests that the resistance to TCEP reduction increases with the proportion of building block **2** present in the hexamer replicators (Figure S30b-d). The apparent increase of the abundance of $\mathbf{1}_1\mathbf{2}_5$ in the experiments using 10 mol% and 20 mol% TCEP may be due to an overlap of the $\mathbf{1}_1\mathbf{2}_5$ and $\mathbf{2}_2$ peaks that we were unable to separate in the UPLC analysis.

Table S9. Change and relative change in abundance of hexameric replicators with different building block compositions upon TCEP reduction.

	Relative abundance (%)	Change in abundance (%)			Relative change in abundance (%)		
	Initial distribution	5 mol% TCEP	10 mol% TCEP	20 mol% TCEP	5 mol% TCEP	10 mol% TCEP	20 mol% TCEP
$\mathbf{1}_6$	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\mathbf{1}_5\mathbf{2}_1$	3,65	-0,70	-1,16	-1,63	-19,18	-31,78	-44,66
$\mathbf{1}_4\mathbf{2}_2$	13,86	-2,45	-3,77	-5,71	-17,68	-27,20	-41,20
$\mathbf{1}_3\mathbf{2}_3$	65,55	-8,36	-13,25	-21,08	-12,75	-20,21	-32,16
$\mathbf{1}_2\mathbf{2}_4$	8,31	-0,58	-0,94	-1,50	-6,98	-11,31	-18,05
$\mathbf{1}_1\mathbf{2}_5$	2,99	-0,25	0,02	0,02	-8,36	0,67	0,67
$\mathbf{2}_6$	0,00	0,00	0,00	0,00	0,00	0,00	0,00

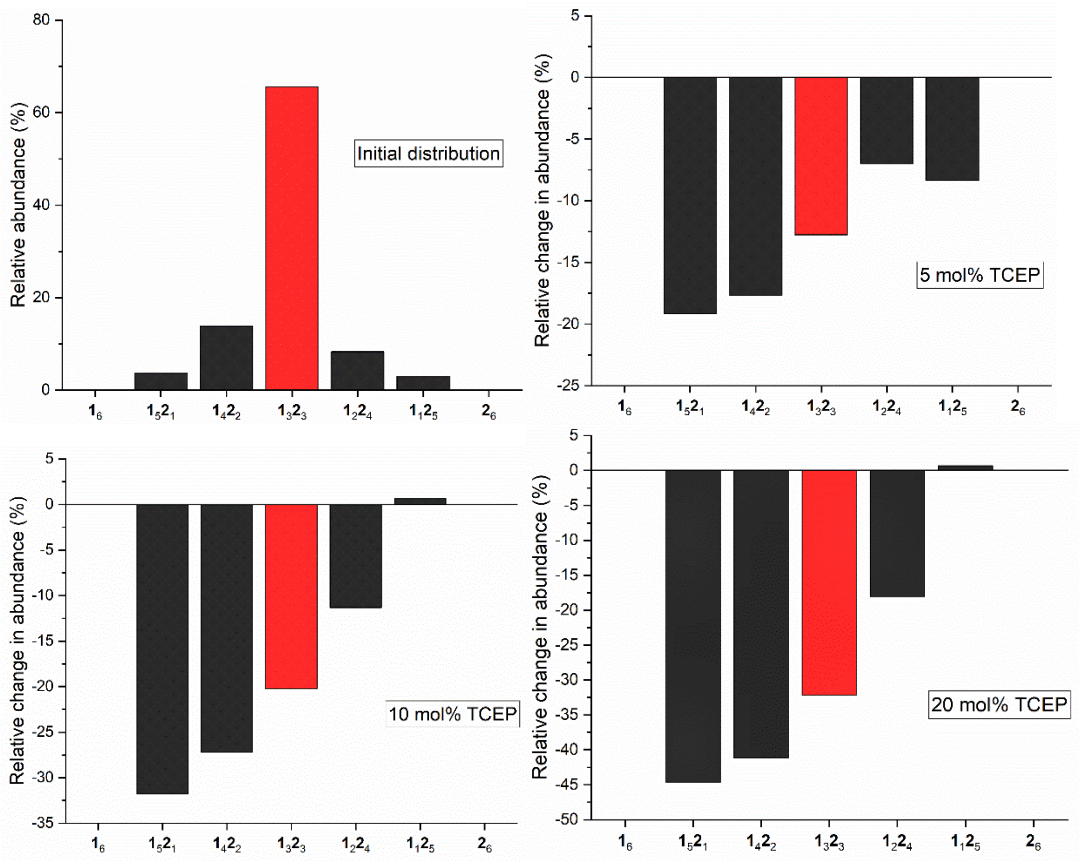


Figure S30. (a) Relative abundance of hexamers in a self-sorted distribution of hexameric replicators in a DCL consisting of equimolar amounts of **1** and **2**. (b-d) Relative change in abundance of the hexamers upon partial reduction by addition of (b) 5 mol%, (c) 10 mol % or (d) 20 mol% TCEP.

S20. Hexamer distributions of the systems shown in Table S3

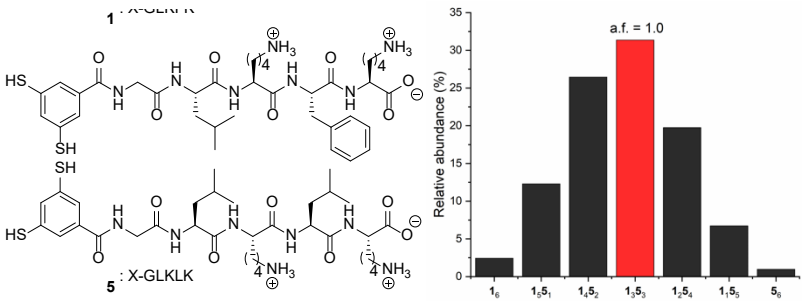


Figure S31. (left) Structures of building blocks **1** and **5**, that both have a peptide chain of alternating hydrophobic and hydrophilic residues, starting with a hydrophobic one. (right) Relative abundance of hexameric replicators in a DCL at 45 °C. No enrichment of $\mathbf{1}_3\mathbf{5}_3$ was observed, indicating that there is no social self-sorting in this system.

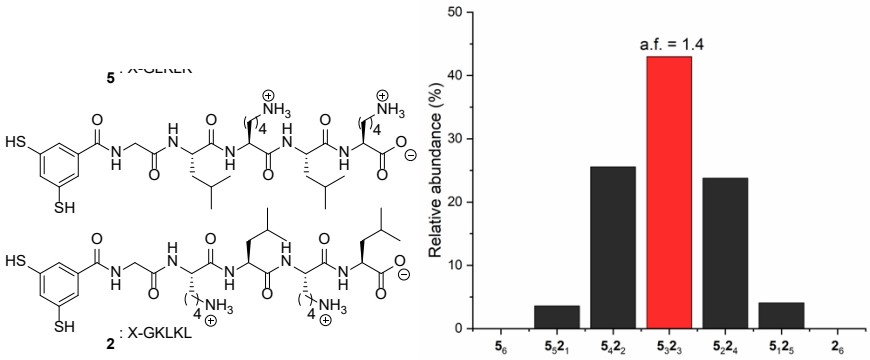


Figure S32. (left) Structures of building blocks **5** and **2**. Building block **5** is a variant of **1** where the Phe residue is changed to a Leu residue (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{5}_3\mathbf{2}_3$ with an AF of 1.4. Note that the assignment of the different hexamers is based on their elution pattern, which is very similar to the elution patterns of similar building blocks. The different compositions could not be confirmed with LC-MS because the masses of **5** and **2** are identical which causes all hexamers to have an identical mass.

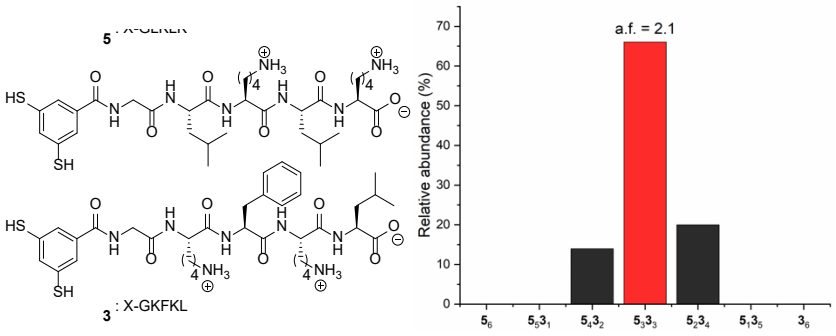


Figure S33. (left) Structures of building blocks **5** and **3**. Building block **5** is a variant of **1** where the Phe residue is changed to a Leu residue and **3** is a variant of **2** where the first Leu residue is changed to Phe residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{5}_3\mathbf{3}_3$ with an AF of 2.1.

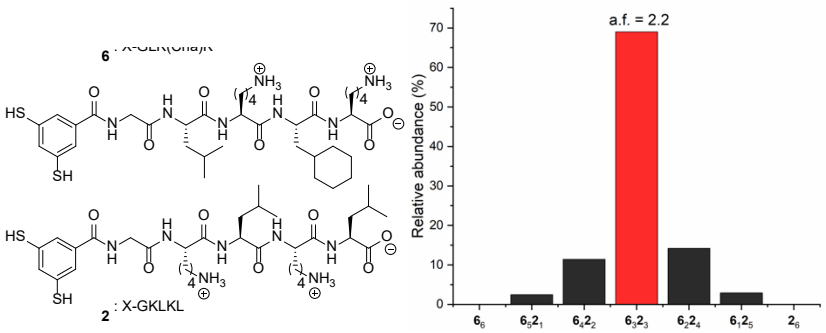


Figure S34. (left) Structures of building blocks **6** and **2**. Building block **6** is a variant of **1** where the Phe residue is changed to a Cha (cyclohexyl alanine) residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{6}_3\mathbf{2}_3$ with an AF of 2.2.

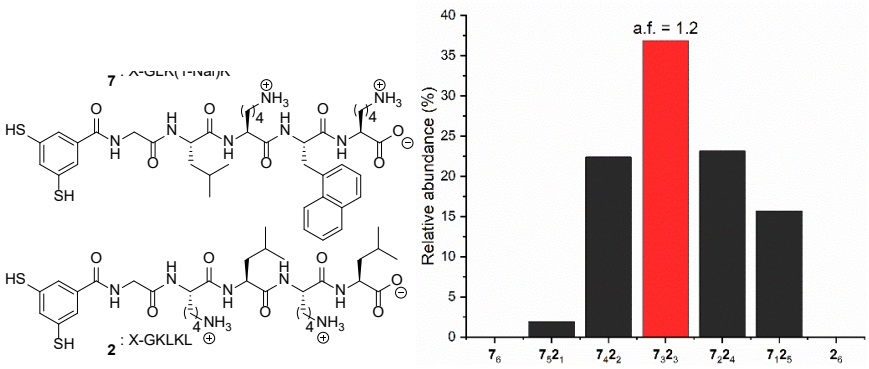


Figure S35. (left) Structures of building blocks **7** and **2**. Building block **7** is a variant of **1** where the Phe-residue is changed to a 1-Nal-residue (1-naphthyl alanine); (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is slightly enriched in $\mathbf{7}_3\mathbf{2}_3$ with an AF of 1.2. The asymmetry of the distribution is a result of an incomplete incorporation of **7** in the assembly of hexamers.

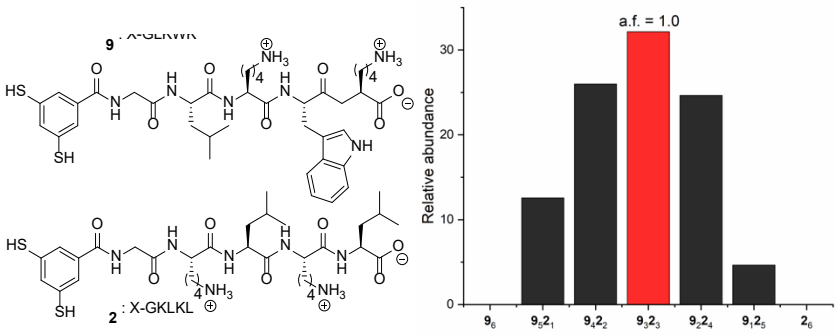


Figure S36. (left) Structures of building blocks **9** and **2**. Building block **9** is a variant of **1** where the Phe-residue is changed to a Trp-residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is not enriched in $\mathbf{9}_3\mathbf{2}_3$ and an AF of 1.0 was obtained.

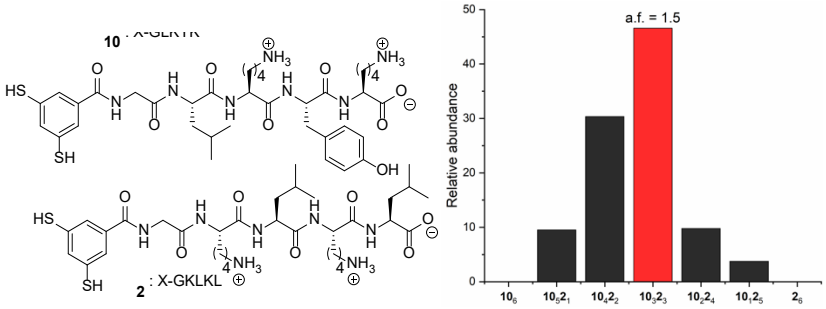


Figure S37. (left) Structures of building blocks **10** and **2**. Building block **10** is a variant of **1** where the Phe-residue is changed to a Tyr-residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{10}_3\mathbf{2}_3$ with an AF of 1.5. The asymmetry of the distribution is a result of an incomplete incorporation of 10 in the assembly of hexamers.

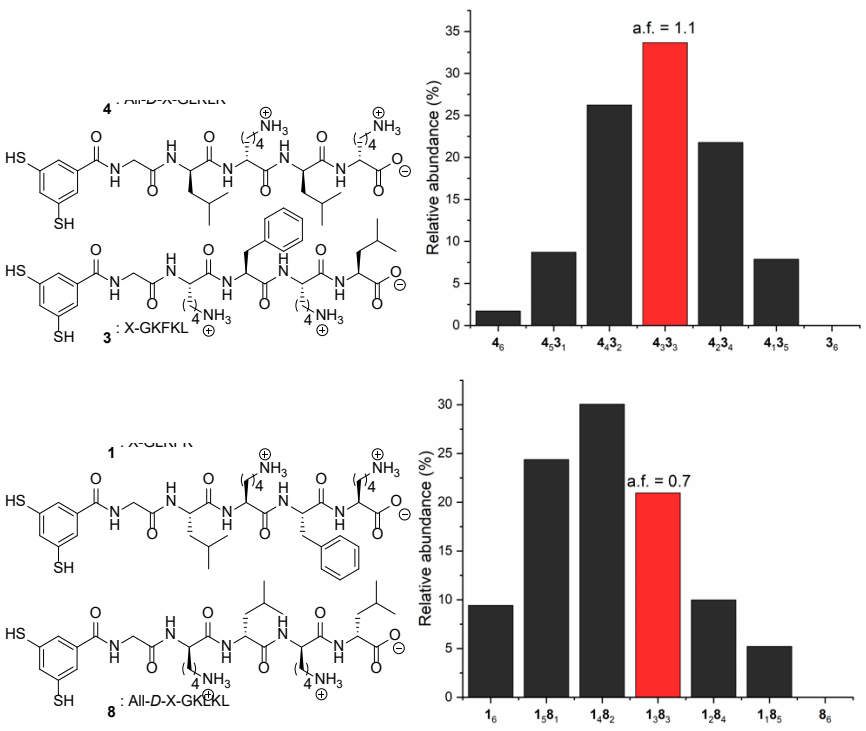


Figure S38. (top left) Structures of building blocks **4** and **3**. Building block **4** is a variant of **5** where the chirality of every amino acid is inverted; (top right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is only slightly enriched in $\mathbf{4}_3\mathbf{3}_3$ with an AF of 1.1. (bottom left) Structures of building blocks **1** and **8**. Building block **8** is a variant of **2** where the chirality of every amino acid is inverted; (bottom right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is skewed towards **1**. This asymmetry is a result of an incomplete incorporation of **8** in the assembly of hexamers. No amplification of $\mathbf{1}_3\mathbf{8}_3$ was observed as it was formed with an AF of only 0.7.

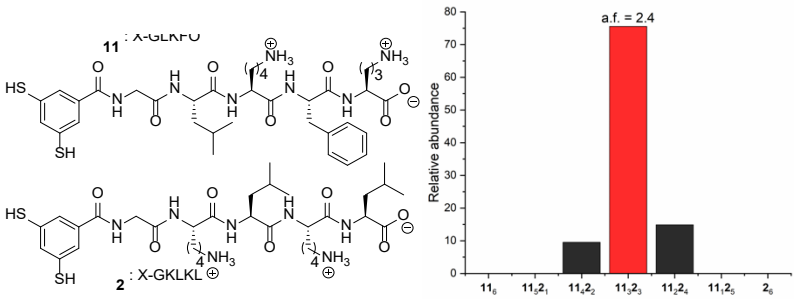


Figure S39. (left) Structures of building blocks **11** and **2**. Building block **11** is a variant of **1** where the second Lys residue is changed to an Orn-residue (Ornithine); (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{11}_3\mathbf{2}_3$ with an AF of 2.4.

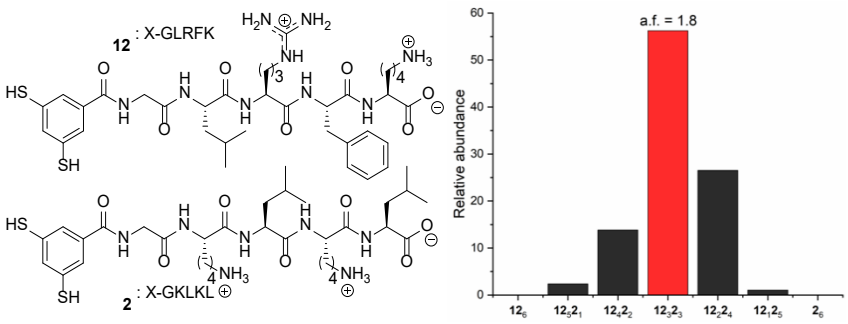


Figure S40. (left) Structures of building blocks **12** and **2**. Building block **12** is a variant of **1** where the first Lys residue is changed to an Arg-residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{12}_3\mathbf{2}_3$ with an AF of 1.8.

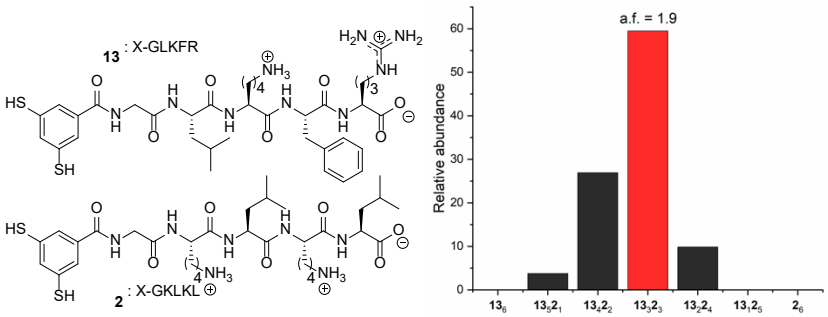


Figure S41. (left) Structures of building blocks **13** and **2**. Building block **13** is a variant of **1** where the second Lys residue is changed to an Arg-residue; (right) Relative abundance of hexameric replicators in a DCL at 45 °C. The distribution is enriched in $\mathbf{13}_3\mathbf{2}_3$ with an AF of 1.9.

S21. Self-assembly is essential for self-sorting

When 3mers and 4mers in a DCL made from building blocks **1** and **2** are molecularly dissolved (not assembled) they do not show self-sorting. A DCL rich in these macrocycles was obtained by not agitating the sample, which will prevent the emergence of self-replicating species (Figure S42). The mixed trimers and tetramers were formed with a distribution of different compositions that is very close to the statistical distribution (Figure S43).

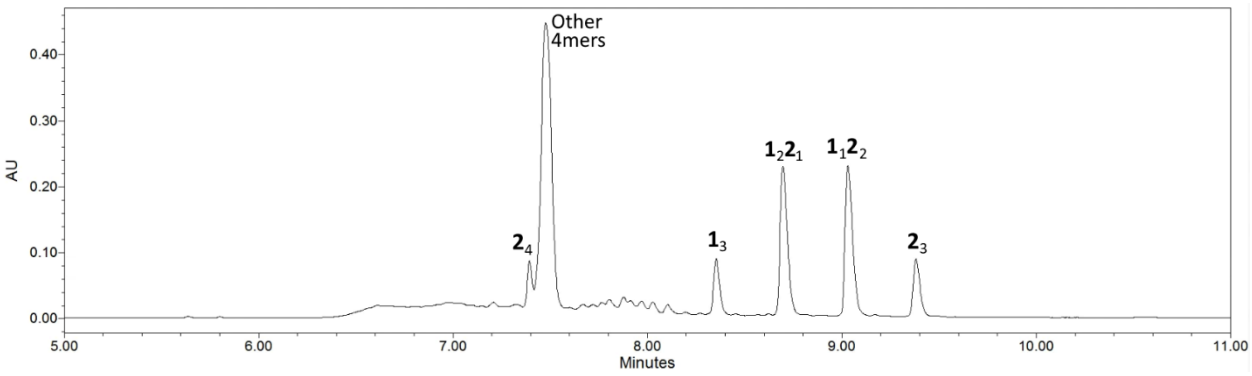


Figure S42. RP-UPLC analysis (monitored at 254 nm) of the product mixture obtained from a DCL of **1** and **2** (that was pre-oxidized to 50% disulfides with sodium perborate) after 5 days at 45 °C without stirring.

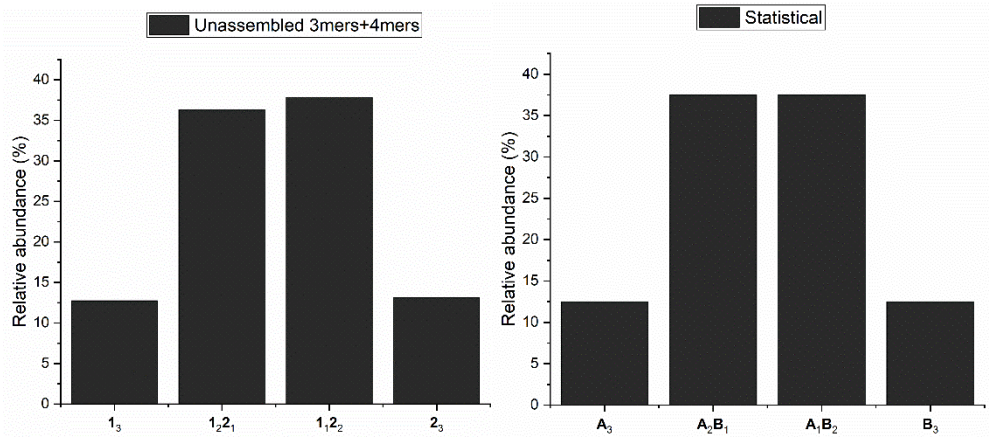


Figure S43. (left) Relative abundance of 3mers with different macrocycle compositions as observed in a non-agitated DCL (Figure S42). (right) Statistical distribution of 3mers with different macrocycle compositions in an equimolar mixture of **A** and **B**, that is expected to form in the absence of specific interactions within the macrocycle.

When a solution of building block **1** is rapidly oxidized by adding >1 eq of perborate, a 3mer self-replicator $\mathbf{1}_3$ is obtained. When the same procedure is performed using **2** no self-assembly (and therefore no self-replication) is observed. Instead the sample produces 3mers and 4mers. When, however, samples are prepared where **1** and **2** are mixed, a self-sorted 3mer replicator ($\mathbf{1}_2\mathbf{2}_1$) is formed with an AF of 2.1 or 1.9 when building block ratios of 1:1 or 2:1, respectively, are used (Figure S44).

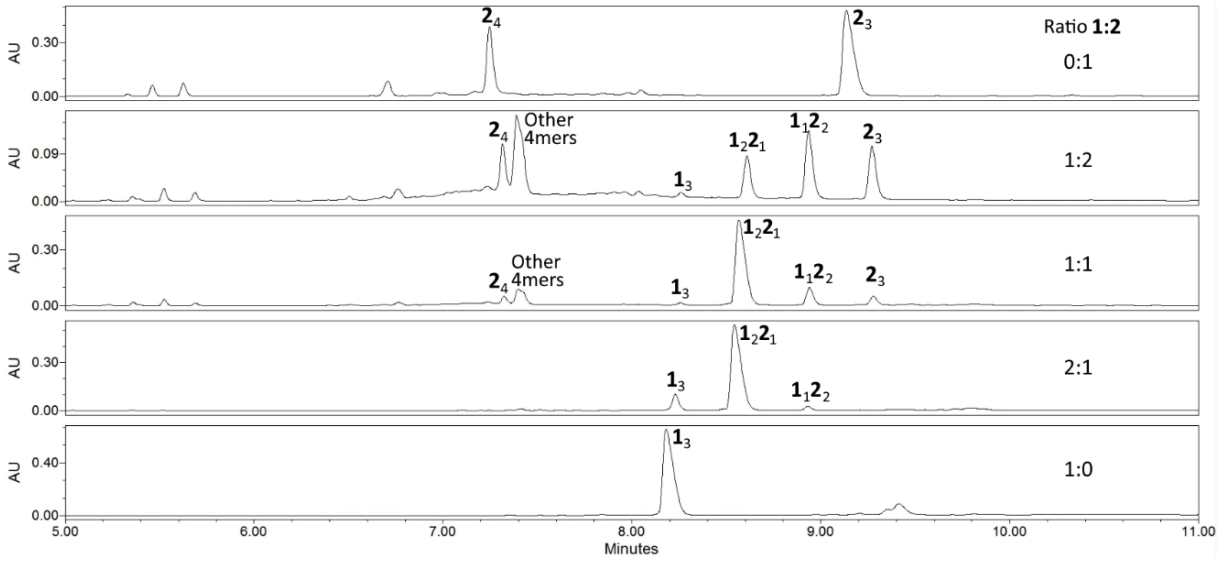


Figure S44. RP-UPLC analyses (monitored at 254 nm) of the product mixture obtained from DCLs made using various ratios of **1** and **2** after 3 days. The DCLs were rapidly oxidized at the start of the experiment with 1.5 eq of sodium perborate and stirred (1200 rpm) at 45 °C for 3 days.

Integration of the chromatograms in Figure S44 gave the relative abundances shown in Figure S45.

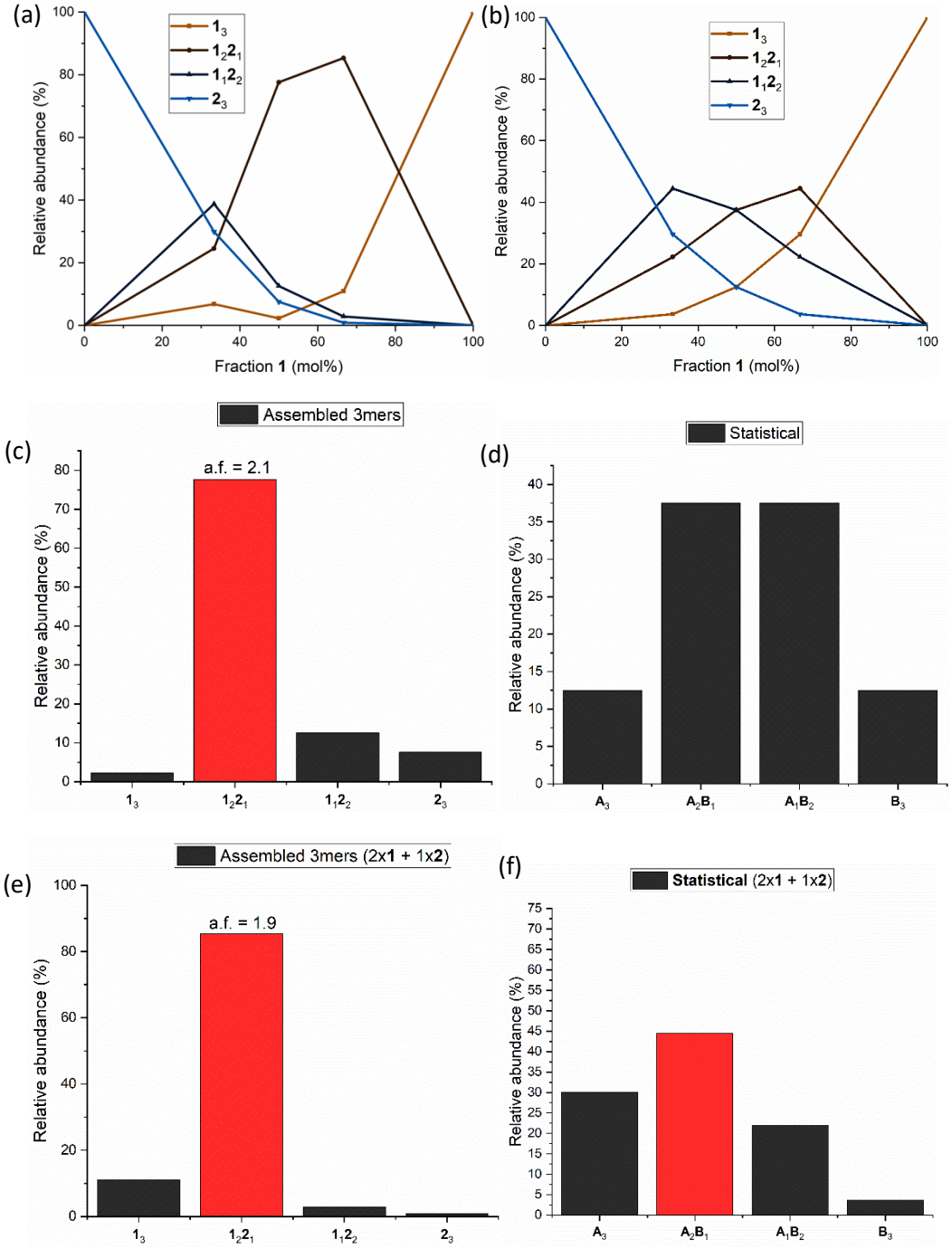


Figure S45. (a) Observed and (b) statistically expected distributions of 3mers with different building block compositions for various ratios of **1** and **2**. (c) Observed and (d) statistically expected distributions of 3mers in a DCL of **1** and **2** with a 1:1 ratio. The $\mathbf{1}_2\mathbf{2}_1$ trimer is formed with an AF of 2.1. (e) Observed and (f) statistically expected distributions of 3mers in a DCL of **1** and **2** with a 2:1 ratio. The $\mathbf{1}_2\mathbf{2}_1$ trimer is formed with an AF of 1.9.

Surprisingly $\mathbf{1}_1\mathbf{2}_2$ is not amplified and seems to be not able to self-assemble. We speculate that this could be a caused by an insufficient driving force of **2** to self-assemble. This is strengthened by the fact that no self-assembly was observed for samples containing only **2** or when **1** and **2** were used with a 1:2 ratio. It appears that at least two peptide chains of **1** are required in order to make 3mers self-assemble.

S22. RP-UPLC and LC-MS analyses

XGLKFK (1) + XGKLKL (2)

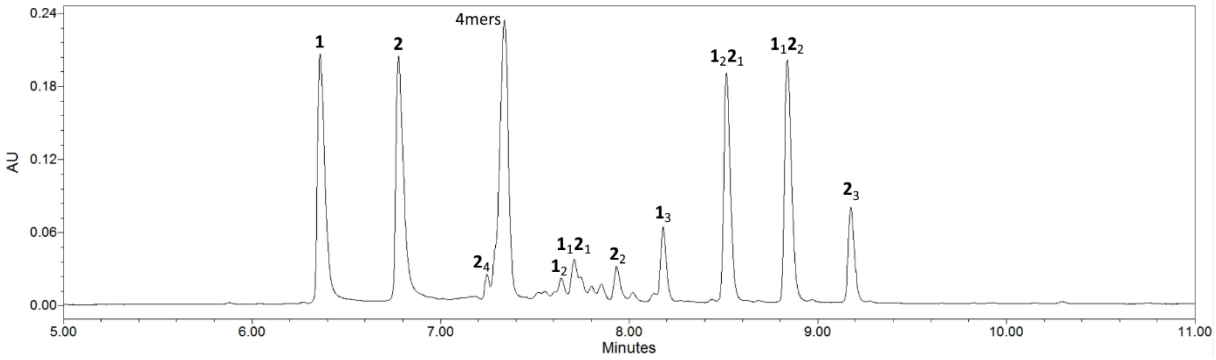


Figure S46. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **2** at 40 °C at the start of the experiment (corresponding to t = 0 h in Figures 3a,b).

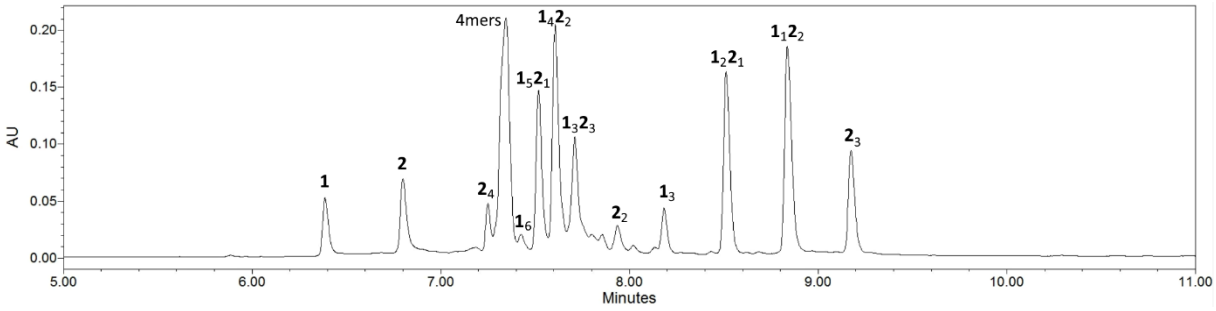


Figure S47. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **2** at 40 °C after 2 days (corresponding to t = 48 h in Figures 3a,b).

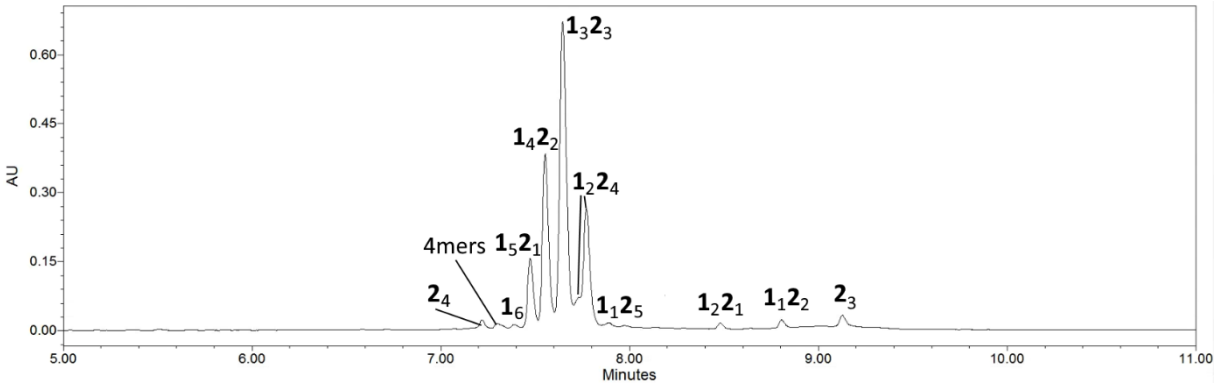


Figure S48. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **2** at 40 °C after 6 days (corresponding to Figure 3c and t = 153 h in Figures 3a,b).

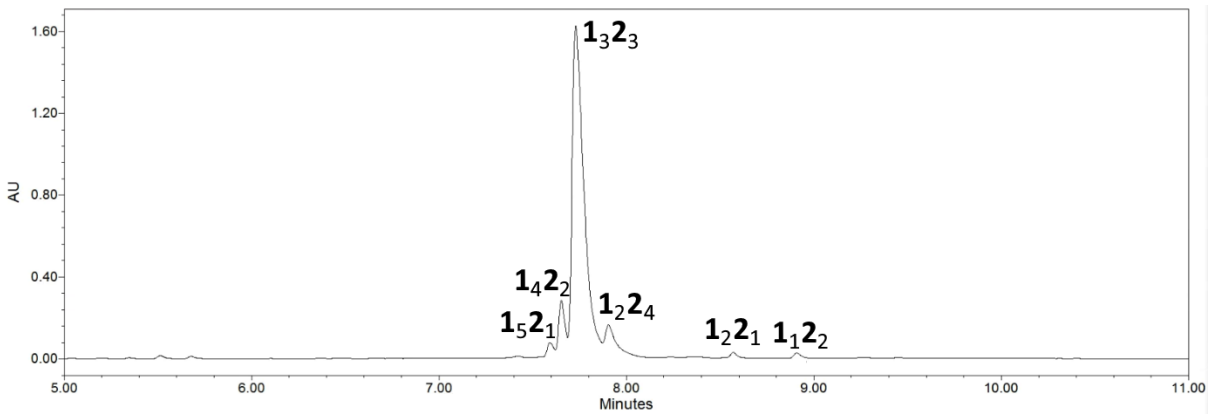


Figure S49. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **2** at 45 °C after 5 days (corresponding to Figure 3f).

1mers

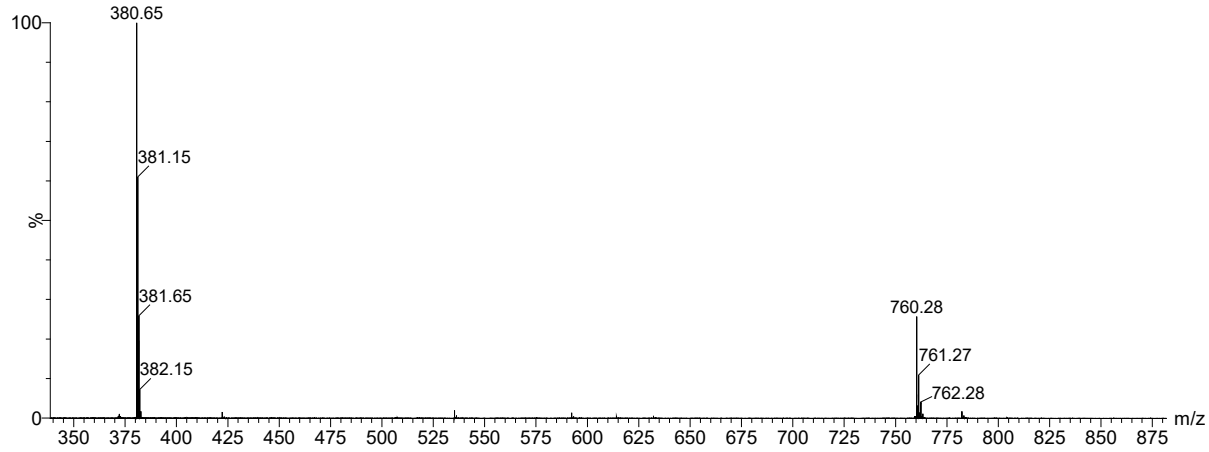


Figure S50. Mass spectrum of monomer **1** from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 380.67 $[M+2H]^{2+}$, 760.34 $[M+1H]^{1+}$; observed m/z: 380.65 $[M+2H]^{2+}$, 760.28 $[M+1H]^{1+}$

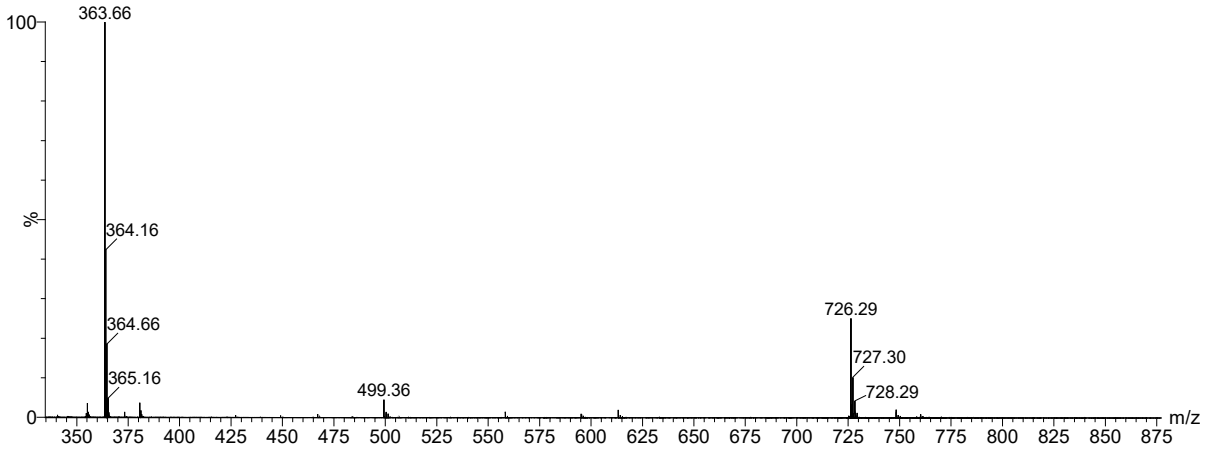


Figure S51. Mass spectrum of monomer **2** from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 363.68 $[M+2H]^{2+}$, 726.36 $[M+1H]^{1+}$; observed m/z: 363.66 $[M+2H]^{2+}$, 726.29 $[M+1H]^{1+}$

2mers

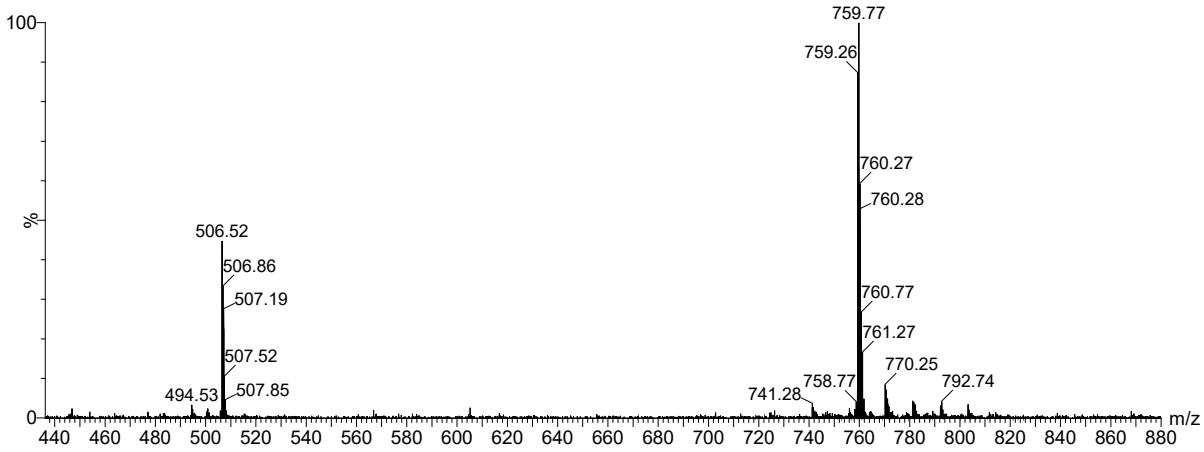


Figure S52. Mass spectrum of dimer $\mathbf{1}_2$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 506.56 $[M+3H]^{3+}$, 759.34 $[M+2H]^{2+}$; observed m/z: 506.52 $[M+3H]^{3+}$, 759.26 $[M+2H]^{2+}$

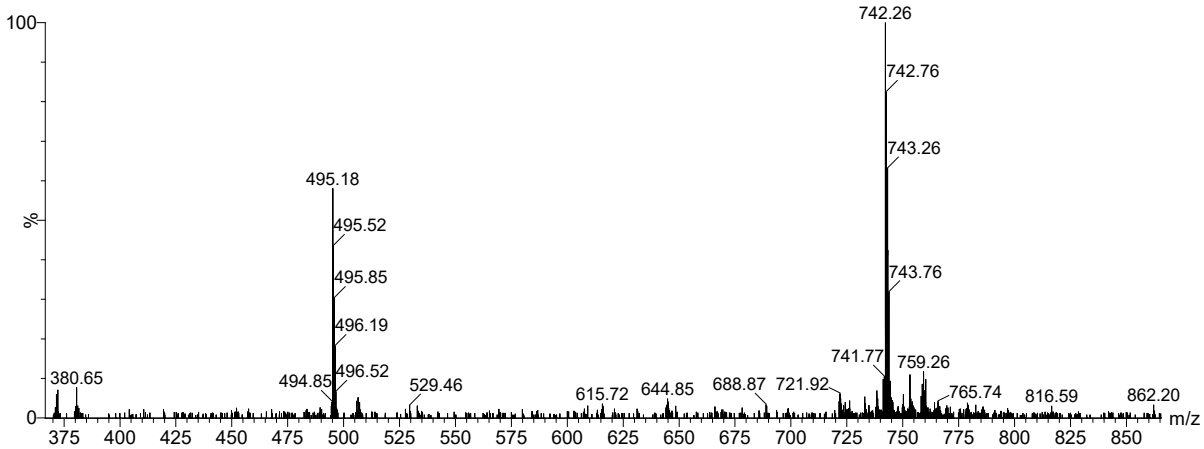


Figure S53. Mass spectrum of dimer $\mathbf{1}_1\mathbf{2}_1$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 495.23 $[M+3H]^{3+}$, 742.35 $[M+2H]^{2+}$; observed m/z: 495.18 $[M+3H]^{3+}$, 742.26 $[M+2H]^{2+}$

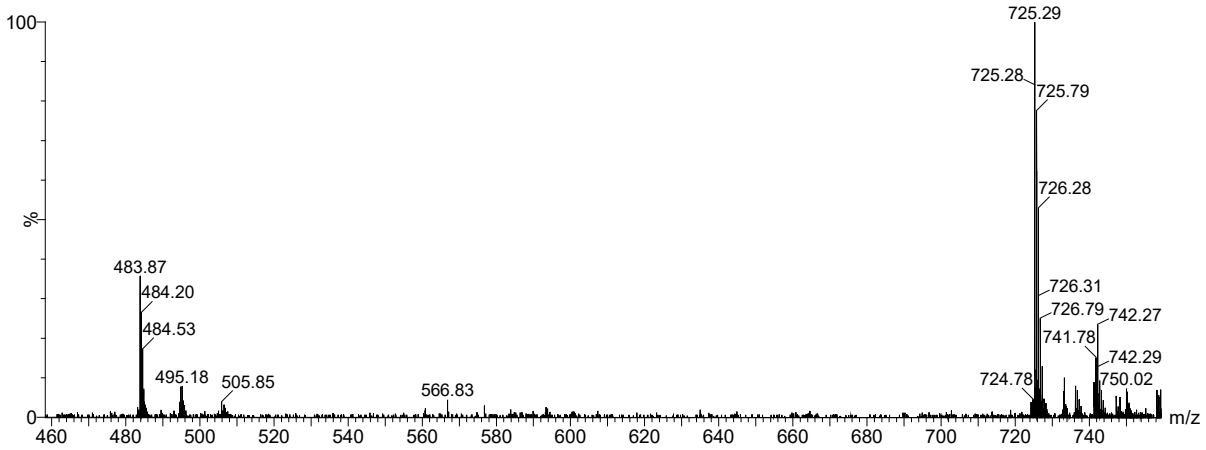


Figure S54. Mass spectrum of dimer $\mathbf{2}_2$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 483.91 $[M+3H]^{3+}$, 725.36 $[M+2H]^{2+}$; observed m/z: 483.87 $[M+3H]^{3+}$, 725.28 $[M+2H]^{2+}$

3mers

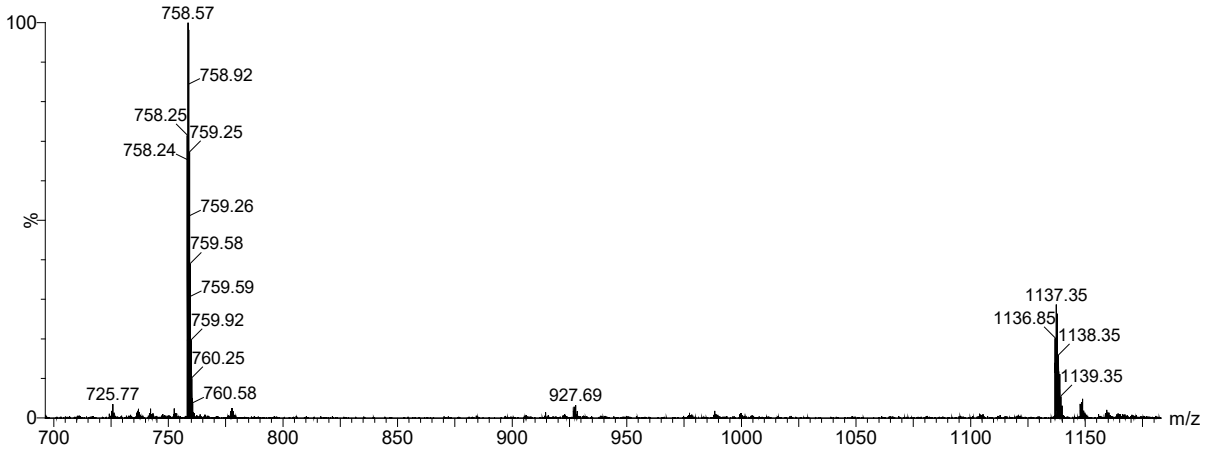


Figure S55. Mass spectrum of trimer $\mathbf{1}_3$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 758.34 $[M+3H]^{3+}$, 1137.01 $[M+2H]^{2+}$; observed m/z: 758.24 $[M+3H]^{3+}$, 1136.85 $[M+2H]^{2+}$

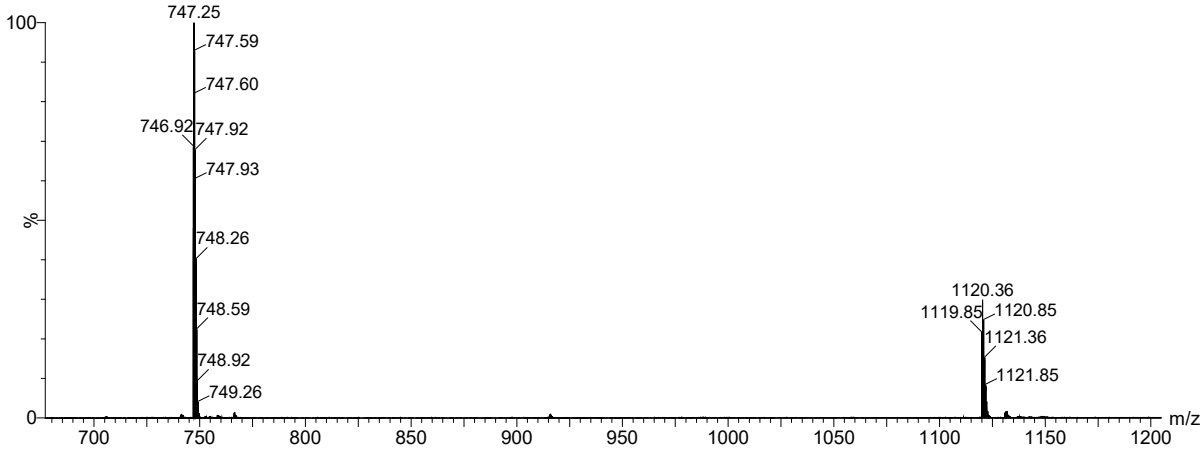


Figure S56. Mass spectrum of trimer $\mathbf{1}_2\mathbf{2}_1$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 747.01 $[M+3H]^{3+}$, 1120.02 $[M+2H]^{2+}$; observed m/z: 746.92 $[M+3H]^{3+}$, 1119.85 $[M+2H]^{2+}$

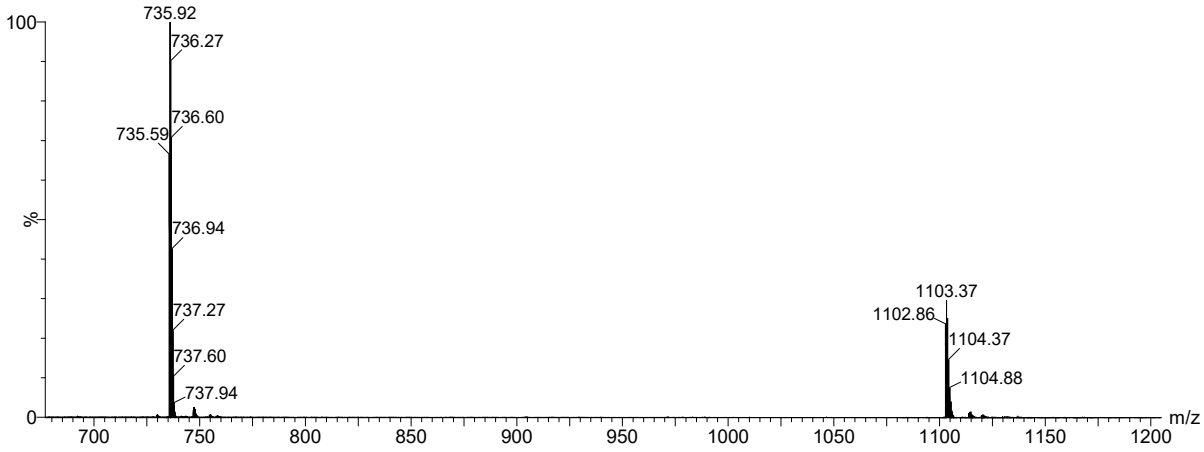


Figure S57. Mass spectrum of trimer $\mathbf{1}_1\mathbf{2}_2$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 735.69 $[M+3H]^{3+}$, 1103.03 $[M+2H]^{2+}$; observed m/z: 735.59 $[M+3H]^{3+}$, 1102.86 $[M+2H]^{2+}$

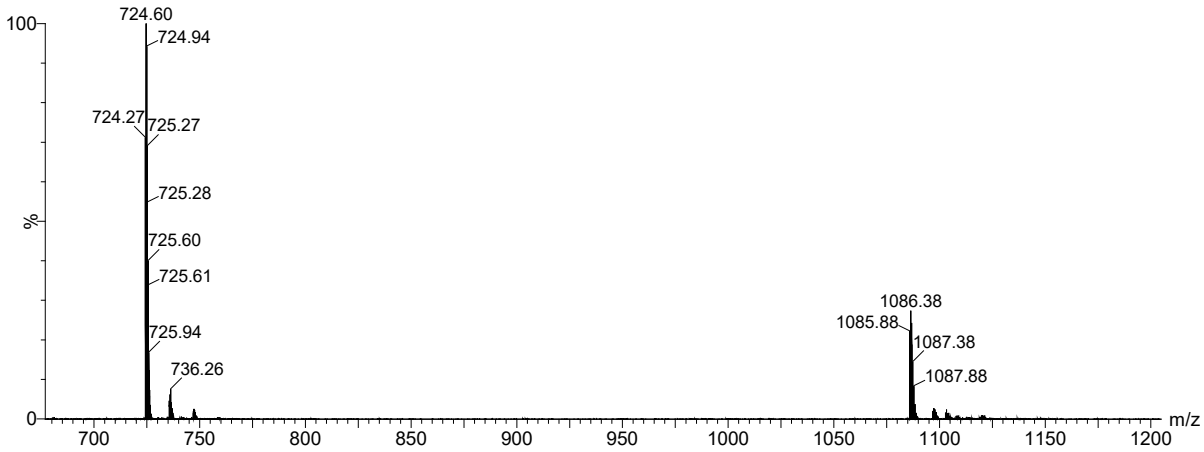


Figure S58. Mass spectrum of trimer $\mathbf{2}_3$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.27 $[M+3H]^{3+}$, 1085.88 $[M+2H]^{2+}$

4mers

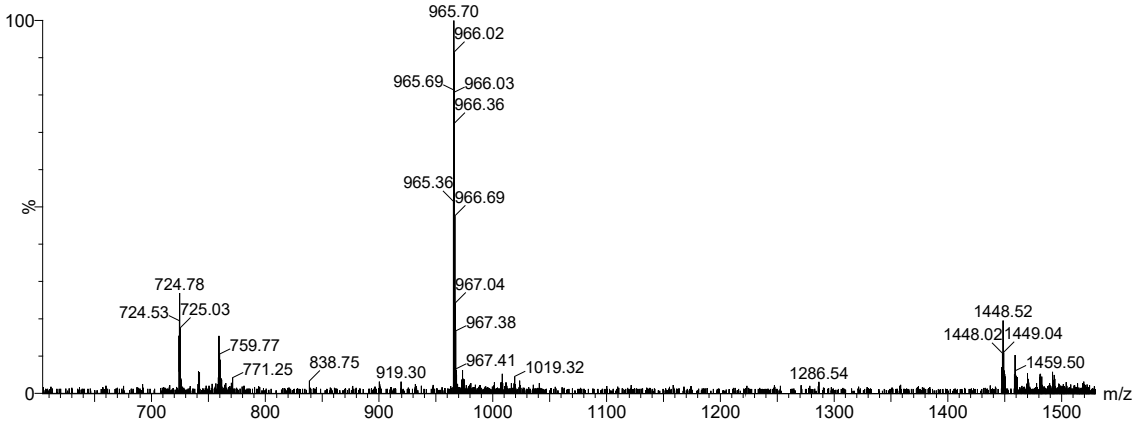


Figure S59. Mass spectrum of tetramer $\mathbf{2}_4$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). Calculated m/z: 724.36 $[M+4H]^{4+}$, 965.48 $[M+3H]^{3+}$, 1447.72 $[M+2H]^{2+}$; observed m/z: 724.53 $[M+4H]^{4+}$, 965.36 $[M+3H]^{3+}$, 1448.02 $[M+2H]^{2+}$

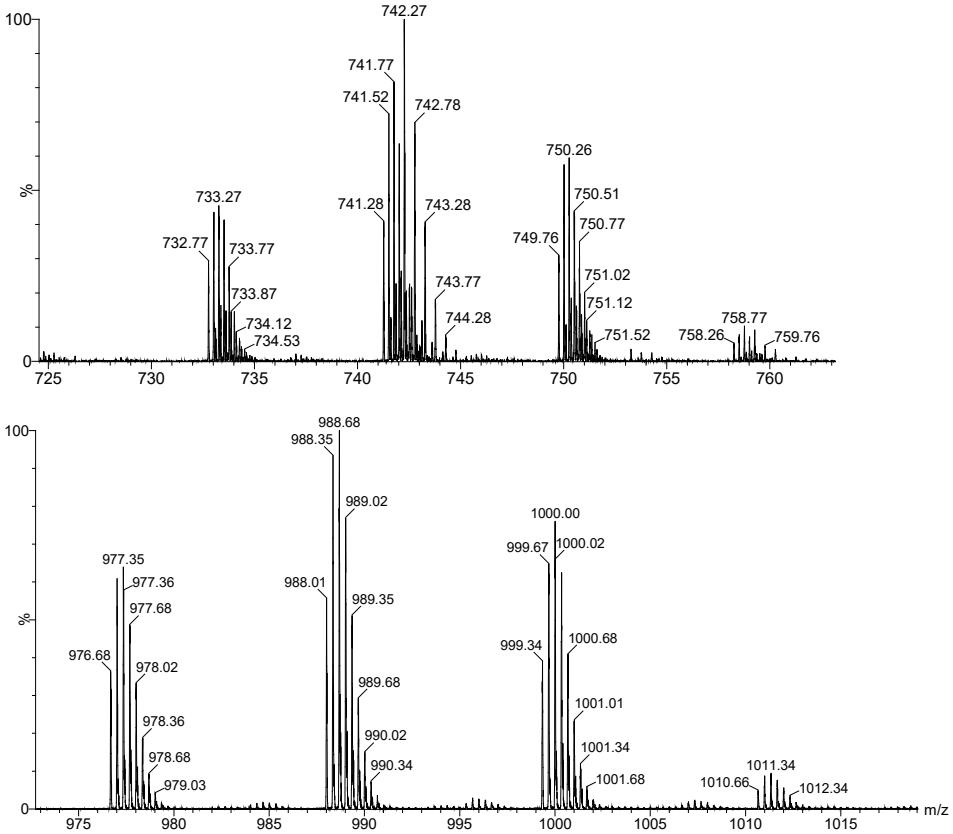


Figure S60. Two partial mass spectra of tetramers $\mathbf{1}_4$, $\mathbf{1}_3\mathbf{2}_1$, $\mathbf{1}_2\mathbf{2}_2$, $\mathbf{1}_1\mathbf{2}_3$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S46). These tetramers elute together in a single peak in the LC-MS analysis. $\mathbf{1}_4$: Calculated m/z: 758.34 $[M+4H]^{4+}$, 1010.79 $[M+3H]^{3+}$; observed m/z: 758.26 $[M+4H]^{4+}$, 1010.66 $[M+3H]^{3+}$; $\mathbf{1}_3\mathbf{2}_1$: Calculated m/z: 749.85 $[M+4H]^{4+}$, 999.46 $[M+3H]^{3+}$; observed m/z: 749.76 $[M+4H]^{4+}$, 999.34 $[M+3H]^{3+}$; $\mathbf{1}_2\mathbf{2}_2$: Calculated m/z: 741.35 $[M+4H]^{4+}$, 988.13 $[M+3H]^{3+}$; observed m/z: 741.28 $[M+4H]^{4+}$, 988.01 $[M+3H]^{3+}$; $\mathbf{1}_1\mathbf{2}_3$: Calculated m/z: 732.86 $[M+4H]^{4+}$, 976.81 $[M+3H]^{3+}$; observed m/z: 732.77 $[M+4H]^{4+}$, 976.68 $[M+3H]^{3+}$

5mers

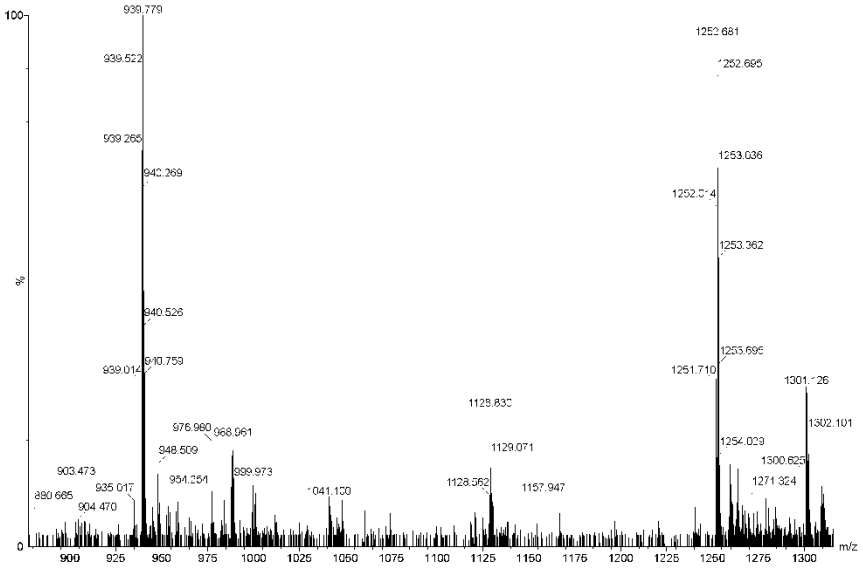


Figure S61. Mass spectrum of pentamer $\mathbf{1}_4\mathbf{2}_1$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S24). Calculated m/z: 939.18 $[M+4H]^{4+}$, 1251.91 $[M+3H]^{3+}$; observed m/z: 939.01 $[M+4H]^{4+}$, 1251.71 $[M+3H]^{3+}$

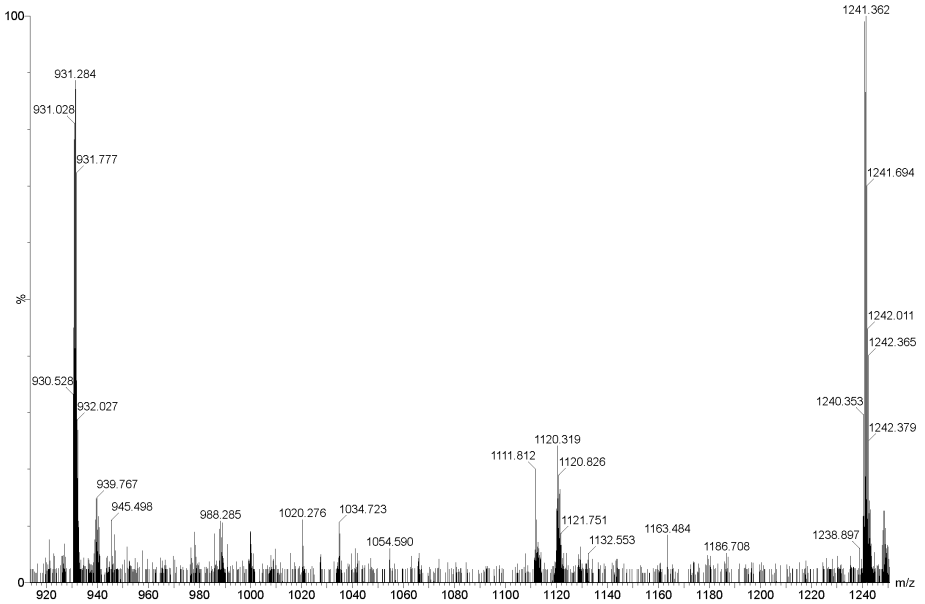


Figure S62. Mass spectrum of pentamer $\mathbf{1}_3\mathbf{2}_2$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S24). Calculated m/z: 930.69 $[M+4H]^{4+}$, 1240.58 $[M+3H]^{3+}$; observed m/z: 930.53 $[M+4H]^{4+}$, 1240.35 $[M+3H]^{3+}$

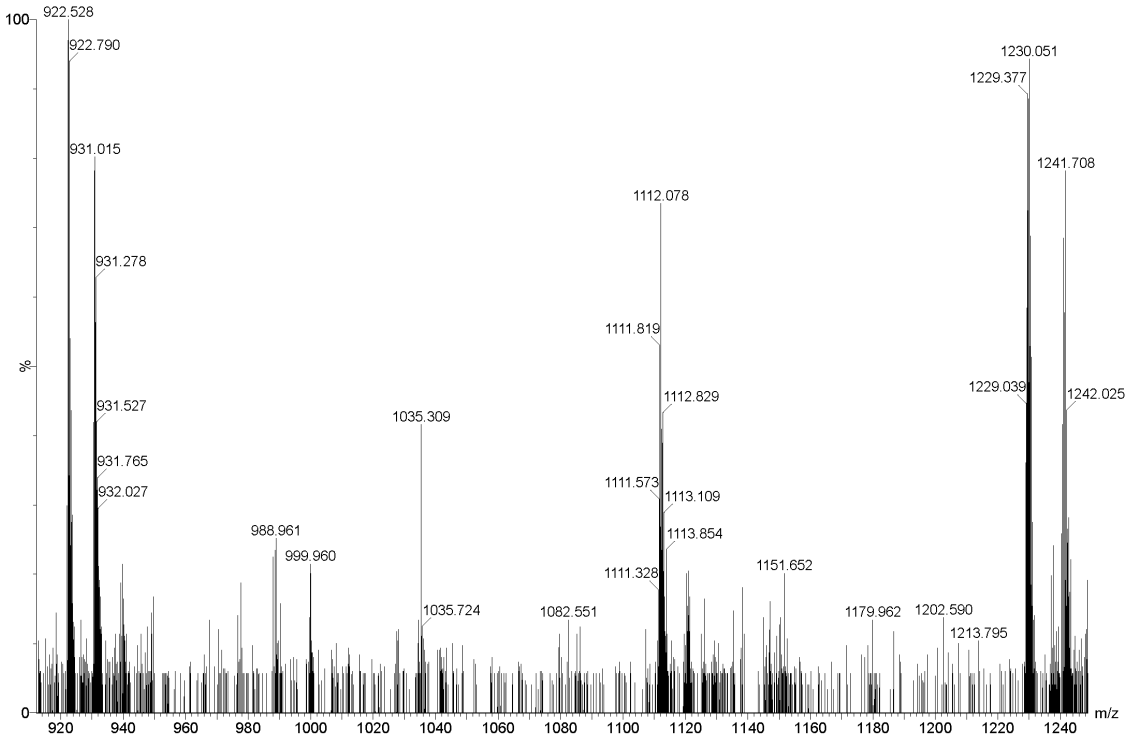


Figure S63. Mass spectrum of pentamer $\mathbf{1}_2\mathbf{2}_3$ from the LC-MS analysis of a library made from peptides **1** and **2** after the initial perborate oxidation (corresponding to Figure S24). Calculated m/z: 922.19 $[M+4H]^{4+}$, 1229.25 $[M+3H]^{3+}$; observed m/z: 922.53 $[M+4H]^{4+}$, 1229.04 $[M+3H]^{3+}$. Additional peaks are detected that correspond to pentamer $\mathbf{1}_3\mathbf{2}_2$ and hexamer $\mathbf{1}_3\mathbf{2}_3$ that elute close to pentamer $\mathbf{1}_2\mathbf{2}_3$.

6mers

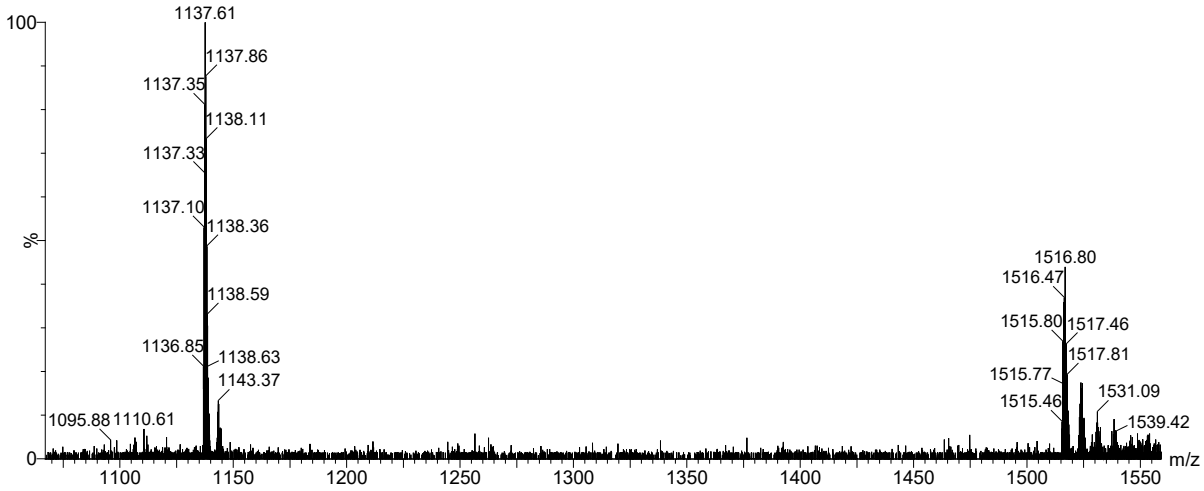


Figure S64. Mass spectrum of hexamer $\mathbf{1}_6$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S47). Calculated m/z: 1137.01 $[M+4H]^{4+}$, 1515.68 $[M+3H]^{3+}$; observed m/z: 1136.85 $[M+4H]^{4+}$, 1515.46 $[M+3H]^{3+}$

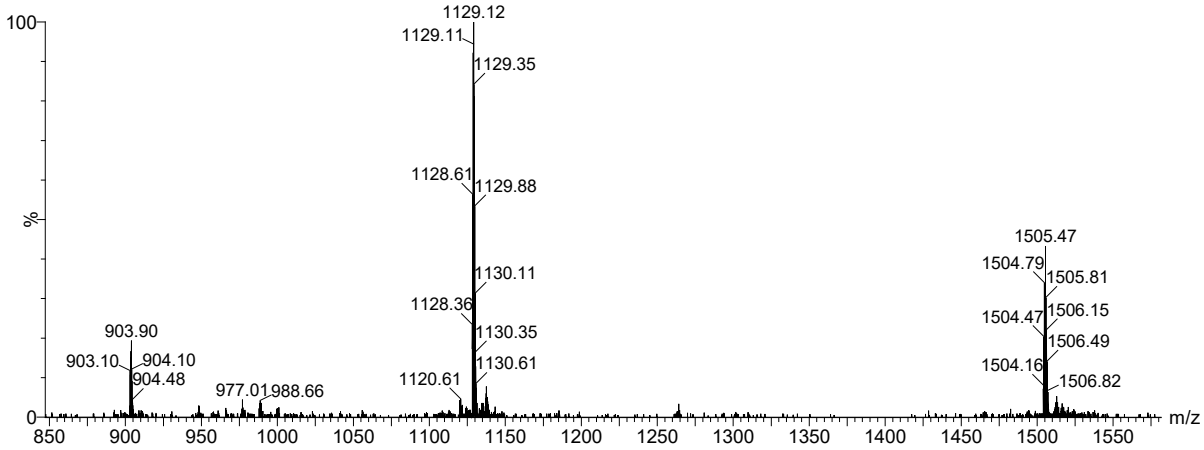


Figure S65. Mass spectrum of hexamer $\mathbf{1}_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S47). Calculated m/z: 903.01 $[M+5H]^{5+}$, 1128.52 $[M+4H]^{4+}$, 1504.35 $[M+3H]^{3+}$; observed m/z: 903.10 $[M+5H]^{5+}$, 1128.36 $[M+4H]^{4+}$, 1504.16 $[M+3H]^{3+}$

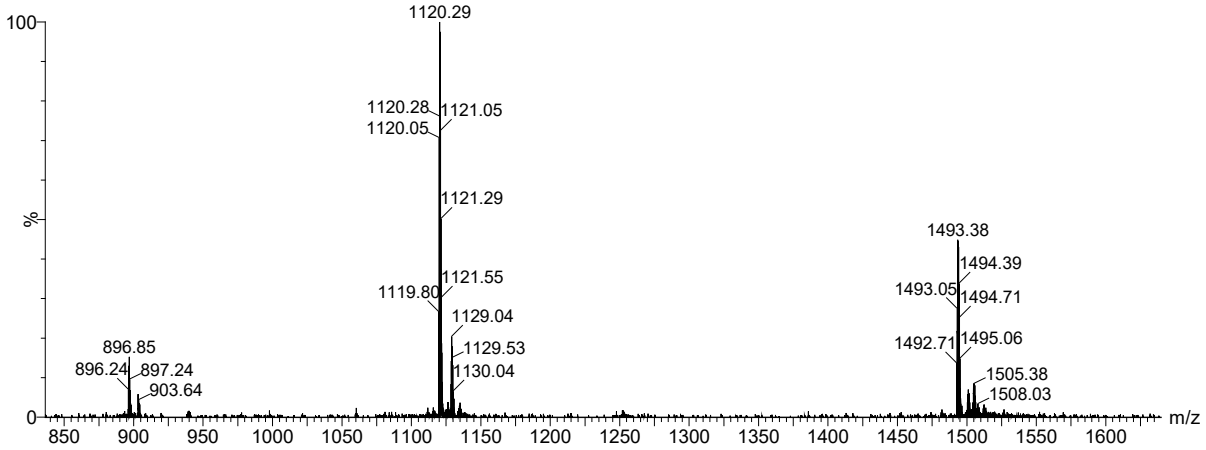


Figure S66. Mass spectrum of hexamer $\mathbf{1}_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S47). Calculated m/z: 896.22 $[M+5H]^{5+}$, 1120.02 $[M+4H]^{4+}$, 1493.03 $[M+3H]^{3+}$; observed m/z: 896.24 $[M+5H]^{5+}$, 1119.80 $[M+4H]^{4+}$, 1492.71 $[M+3H]^{3+}$

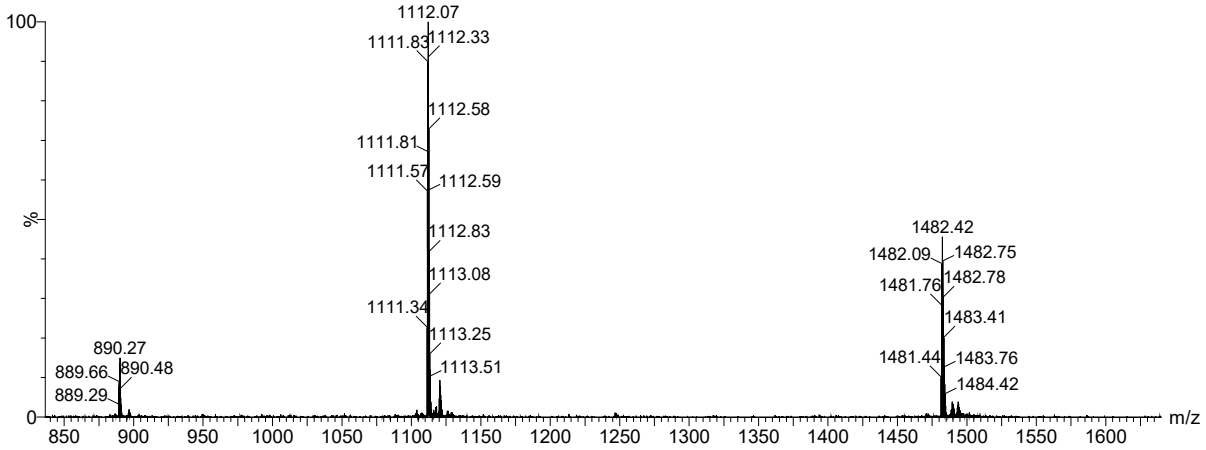


Figure S67. Mass spectrum of hexamer $\mathbf{1}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S48). Calculated m/z: 889.42 $[M+5H]^{5+}$, 1111.53 $[M+4H]^{4+}$, 1481.70 $[M+3H]^{3+}$; observed m/z: 889.29 $[M+5H]^{5+}$, 1111.34 $[M+4H]^{4+}$, 1481.44 $[M+3H]^{3+}$

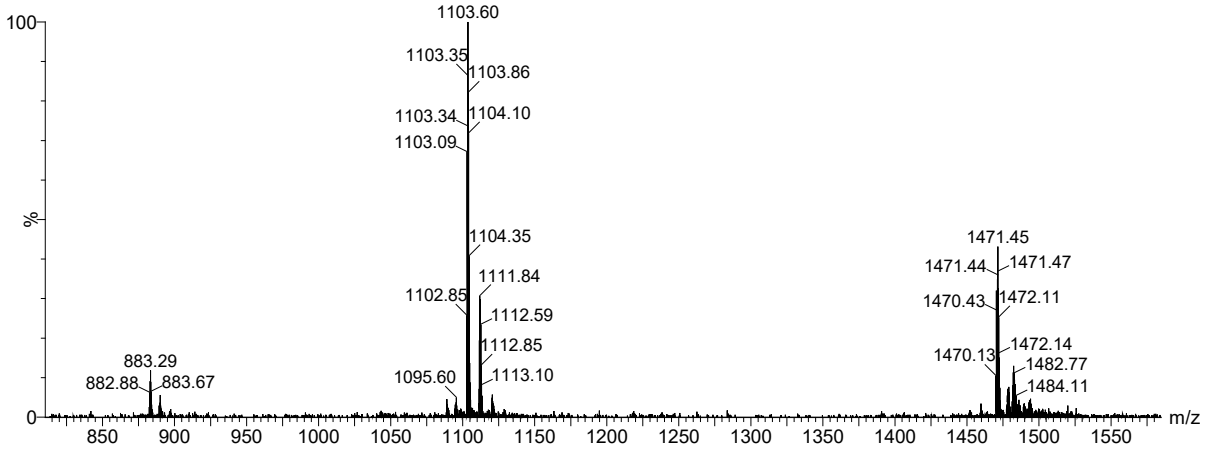


Figure S68. Mass spectrum of hexamer $\mathbf{1}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S48). Calculated m/z: 882.62 $[M+5H]^{5+}$, 1103.03 $[M+4H]^{4+}$, 1470.37 $[M+3H]^{3+}$; observed m/z: 882.88 $[M+5H]^{5+}$, 1102.85 $[M+4H]^{4+}$, 1470.13 $[M+3H]^{3+}$

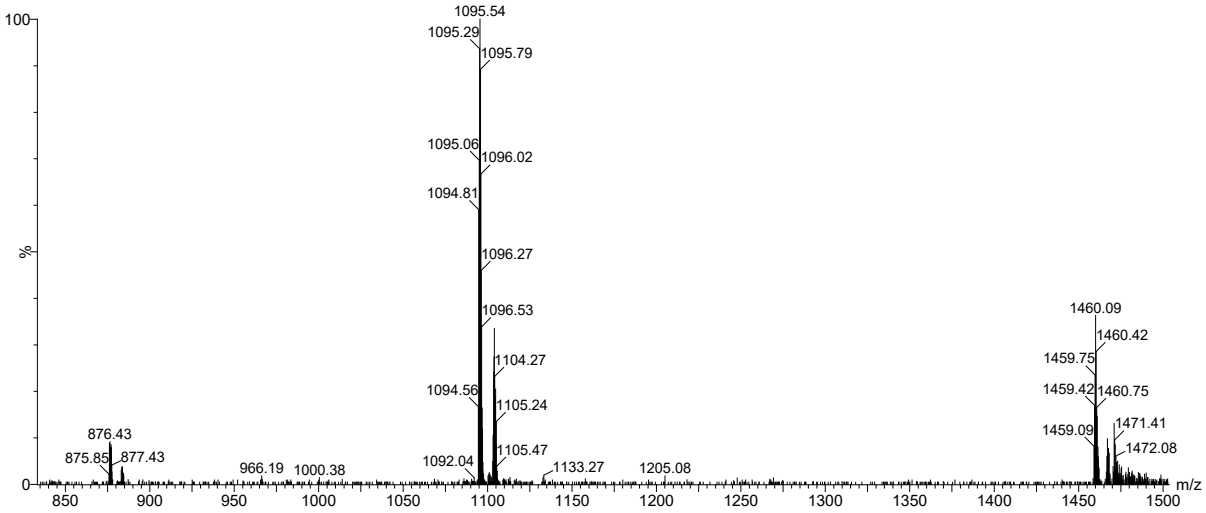


Figure S69. Mass spectrum of hexamer $\mathbf{1}_1\mathbf{2}_5$ from the LC-MS analysis of a stirred library made from peptides **1** and **2** (corresponding to Figure S48). Calculated m/z: 875.83 $[M+5H]^{5+}$, 1094.54 $[M+4H]^{4+}$, 1459.05 $[M+3H]^{3+}$; observed m/z: 875.85 $[M+5H]^{5+}$, 1094.56 $[M+4H]^{4+}$, 1459.09 $[M+3H]^{3+}$

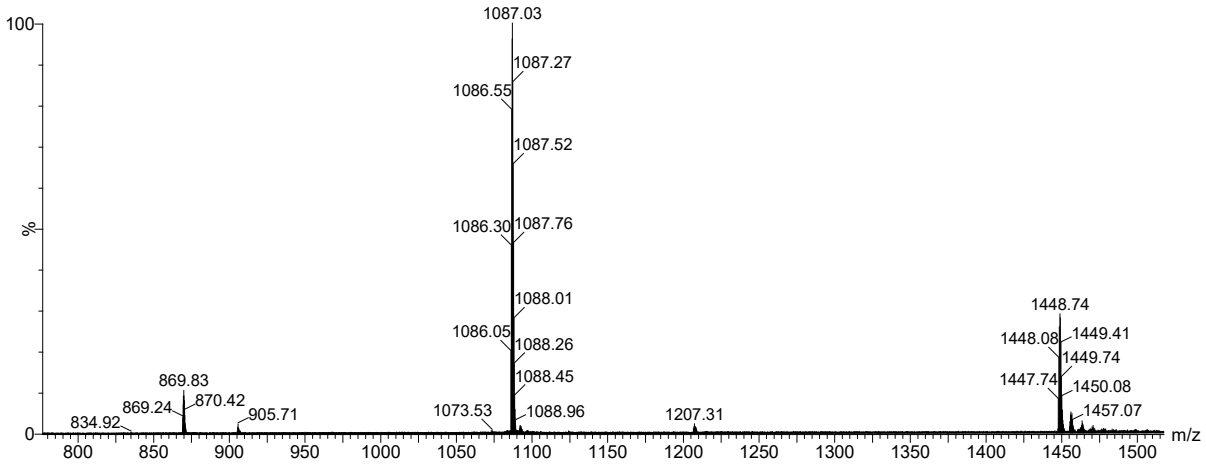


Figure S70. Mass spectrum of hexamer $\mathbf{2}_6$ from the LC-MS analysis of a stirred library made from peptide **2** (corresponding to Figure S2). Calculated m/z: 869.03 $[M+5H]^{5+}$, 1086.04 $[M+4H]^{4+}$, 1447.72 $[M+3H]^{3+}$; observed m/z: 869.24 $[M+5H]^{5+}$, 1086.05 $[M+4H]^{4+}$, 1447.74 $[M+3H]^{3+}$

XGKFKL (3) + XGLKLK (5)

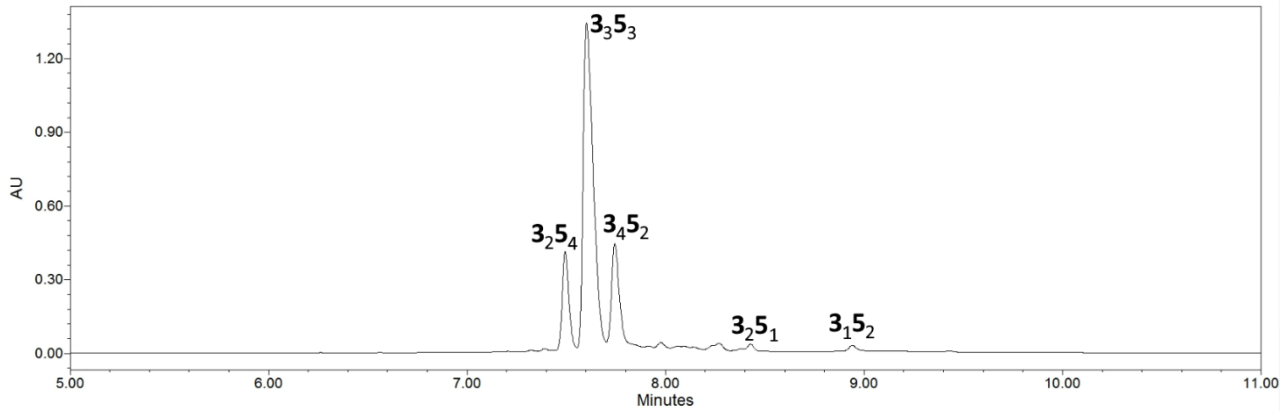


Figure S71. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **3** and **5** at 45 °C after 5 days (corresponding to Figure S33).

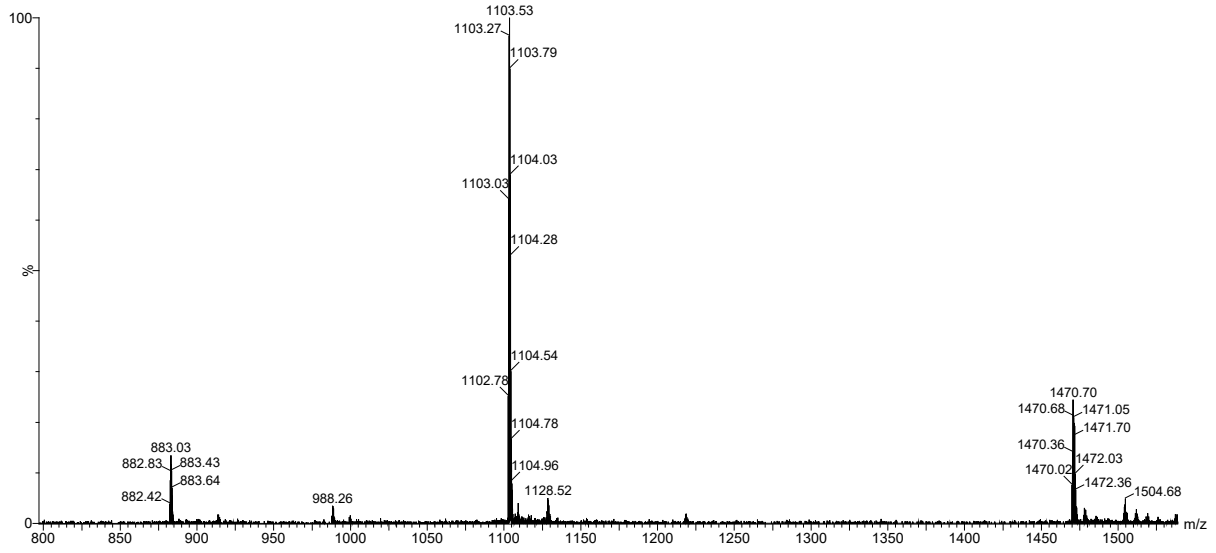


Figure S72. Mass spectrum of hexamer $\mathbf{3}_2\mathbf{5}_4$ from the LC-MS analysis of a stirred library made from peptides **3** and **5** (corresponding to **Figure S71**). Calculated m/z: 882.62 $[M+5H]^{5+}$, 1103.03 $[M+4H]^{4+}$, 1470.37 $[M+3H]^{3+}$; observed m/z: 882.42 $[M+5H]^{5+}$, 1102.78 $[M+4H]^{4+}$, 1470.02 $[M+3H]^{3+}$

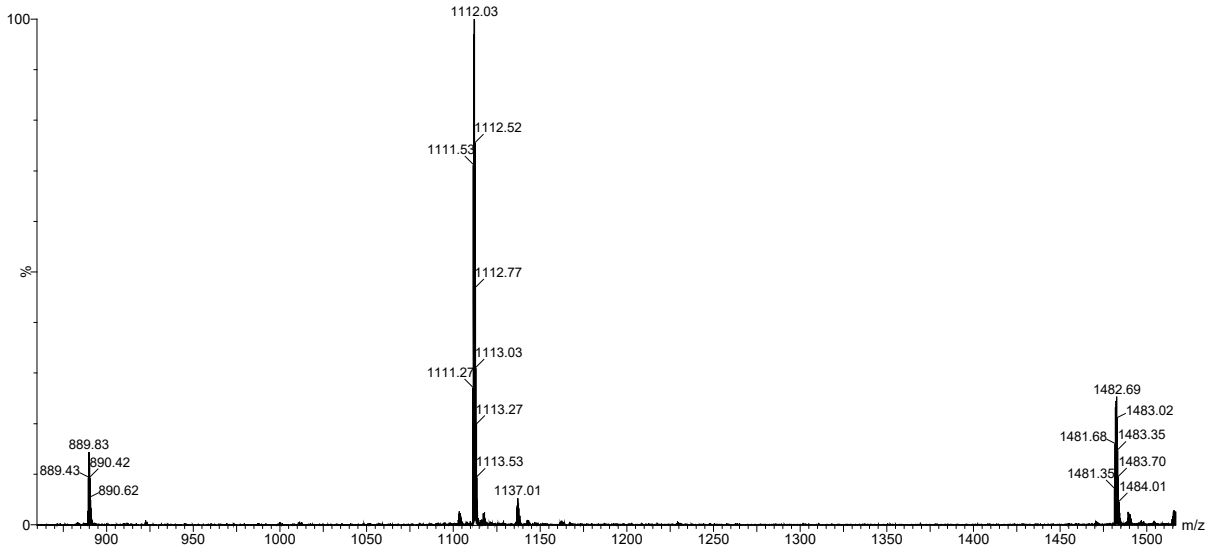


Figure S73. Mass spectrum of hexamer $\mathbf{3}_3\mathbf{5}_3$ from the LC-MS analysis of a stirred library made from peptides **3** and **5** (corresponding to **Figure S71**). Calculated m/z: 889.42 $[M+5H]^{5+}$, 1111.53 $[M+4H]^{4+}$, 1481.70 $[M+3H]^{3+}$; observed m/z: 889.43 $[M+5H]^{5+}$, 1111.27 $[M+4H]^{4+}$, 1481.35 $[M+3H]^{3+}$

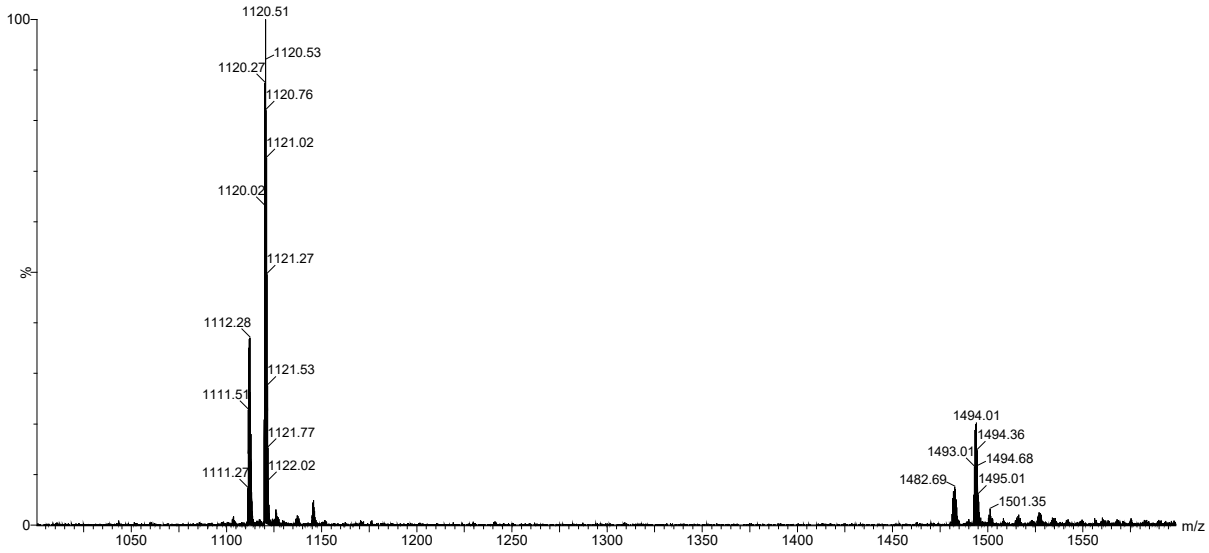


Figure S74. Mass spectrum of hexamer $\mathbf{3}_4\mathbf{5}_2$ from the LC-MS analysis of a stirred library made from peptides **3** and **5** (corresponding to **Figure S71**). This hexamer partially co-elutes with $\mathbf{3}_3\mathbf{5}_3$. Calculated m/z: 1120.02 $[M+4H]^{4+}$, 1493.03 $[M+3H]^{3+}$; observed m/z: 1120.02 $[M+4H]^{4+}$, 1493.01 $[M+3H]^{3+}$

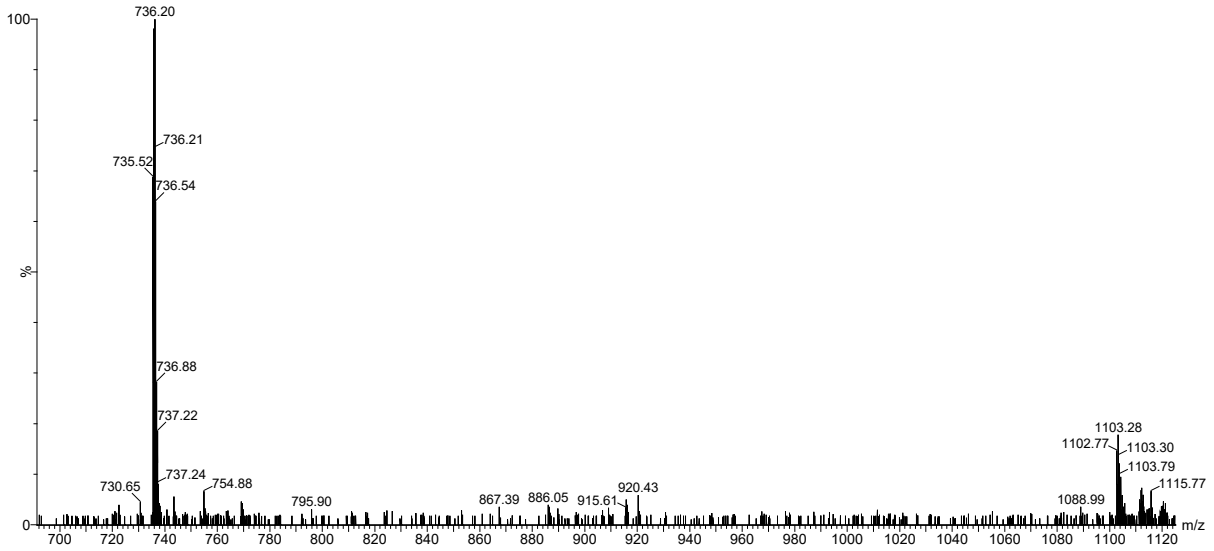


Figure S75. Mass spectrum of trimer $\mathbf{3}_1\mathbf{5}_2$ from the LC-MS analysis of a library made from peptides **3** and **5** after the initial perborate oxidation (corresponding to **Figure S71**). Calculated m/z: 735.69 $[M+3H]^{3+}$, 1103.03 $[M+2H]^{2+}$; observed m/z: 735.52 $[M+3H]^{3+}$, 1102.77 $[M+2H]^{2+}$

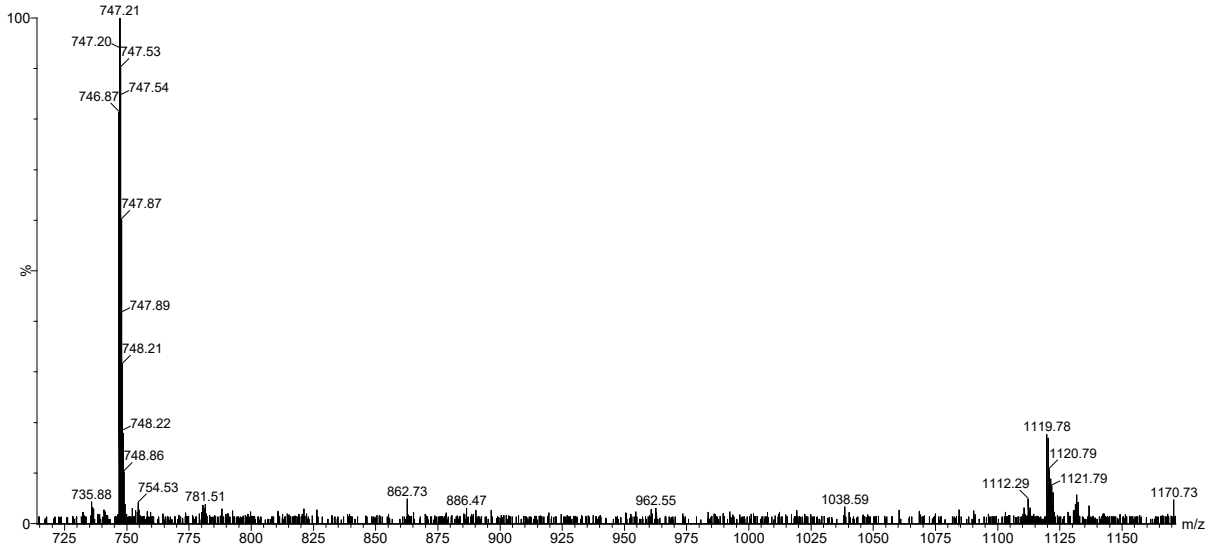


Figure S76. Mass spectrum of trimer $\mathbf{3}_2\mathbf{5}_1$ from the LC-MS analysis of a library made from peptides **3** and **5** after the initial perborate oxidation (corresponding to **Figure S71**). Calculated m/z: 747.01 $[M+3H]^{3+}$, 1120.02 $[M+2H]^{2+}$; observed m/z: 746.87 $[M+3H]^{3+}$, 1119.78 $[M+2H]^{2+}$

XGKFKL (3) + All-D-XGLKLK (4)

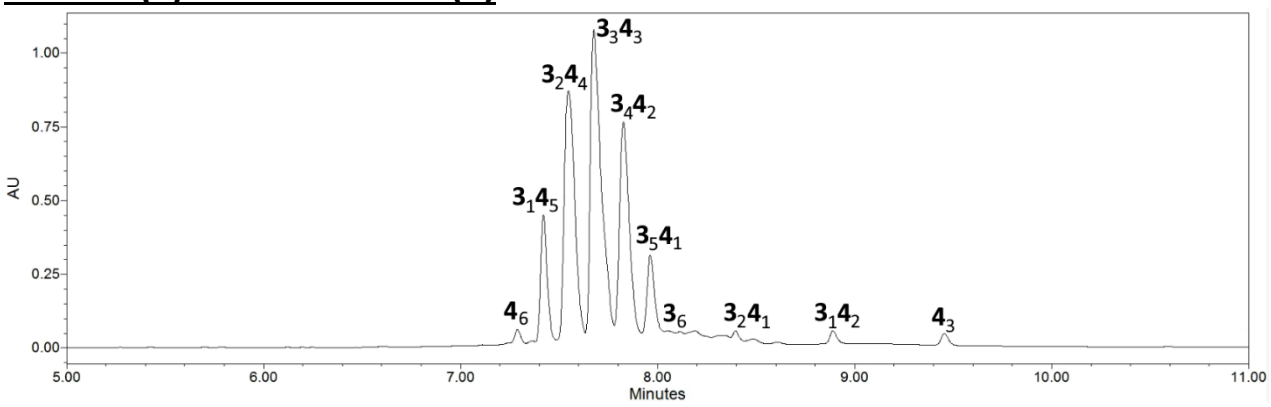


Figure S77. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **3** and **4** at 45 °C after 5 days (corresponding to Figure S38).

3mers

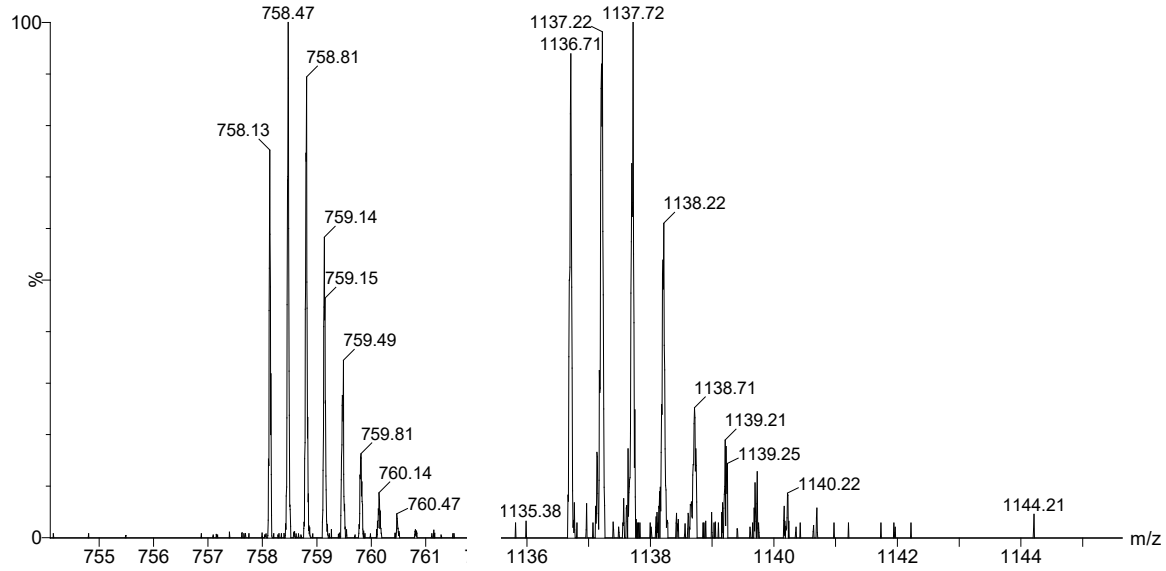


Figure S78. Partial mass spectra of trimer $\mathbf{3}_3$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 758.34 $[M+3H]^{3+}$, 1137.01 $[M+2H]^{2+}$; observed m/z: 758.13 $[M+3H]^{3+}$, 1136.71 $[M+2H]^{2+}$

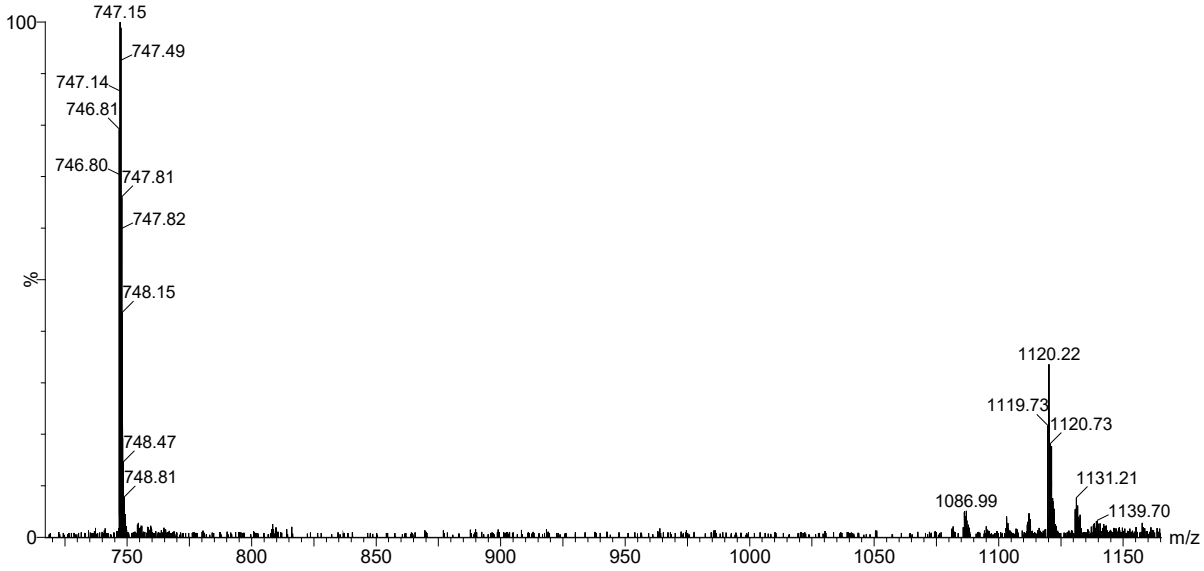


Figure S79. Mass spectrum of trimer $\mathbf{3_2 4_1}$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77)**.** Calculated m/z: 747.01 $[M+3H]^{3+}$, 1120.02 $[M+2H]^{2+}$; observed m/z: 746.80 $[M+3H]^{3+}$, 1119.73 $[M+2H]^{2+}$

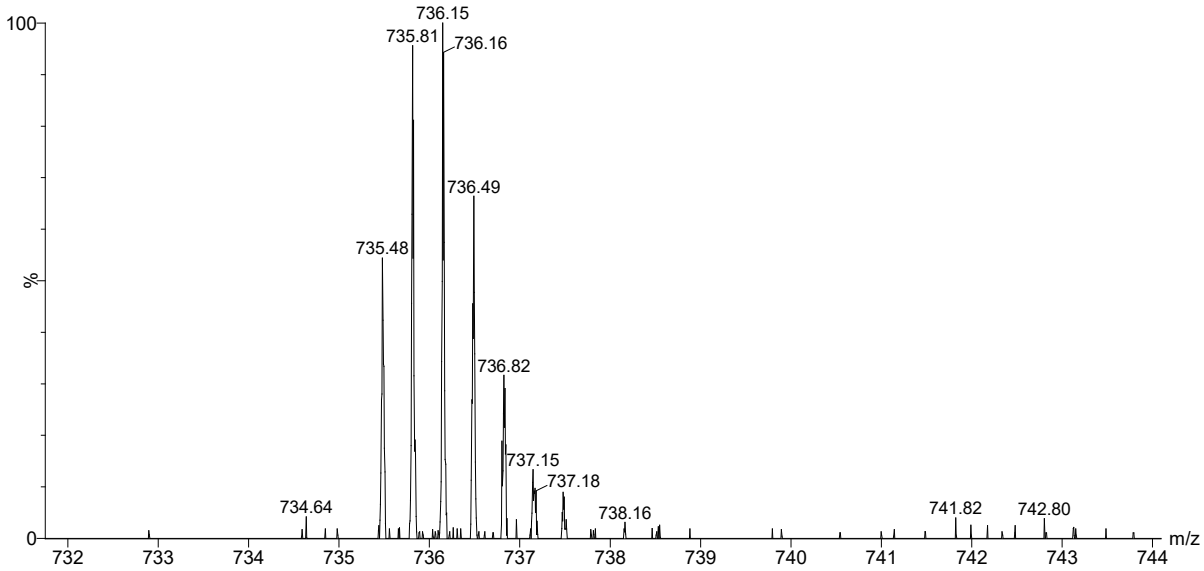


Figure S80. Mass spectrum of trimer $\mathbf{3_1 4_2}$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 735.69 $[M+3H]^{3+}$; observed m/z: 735.48 $[M+3H]^{3+}$

6mers

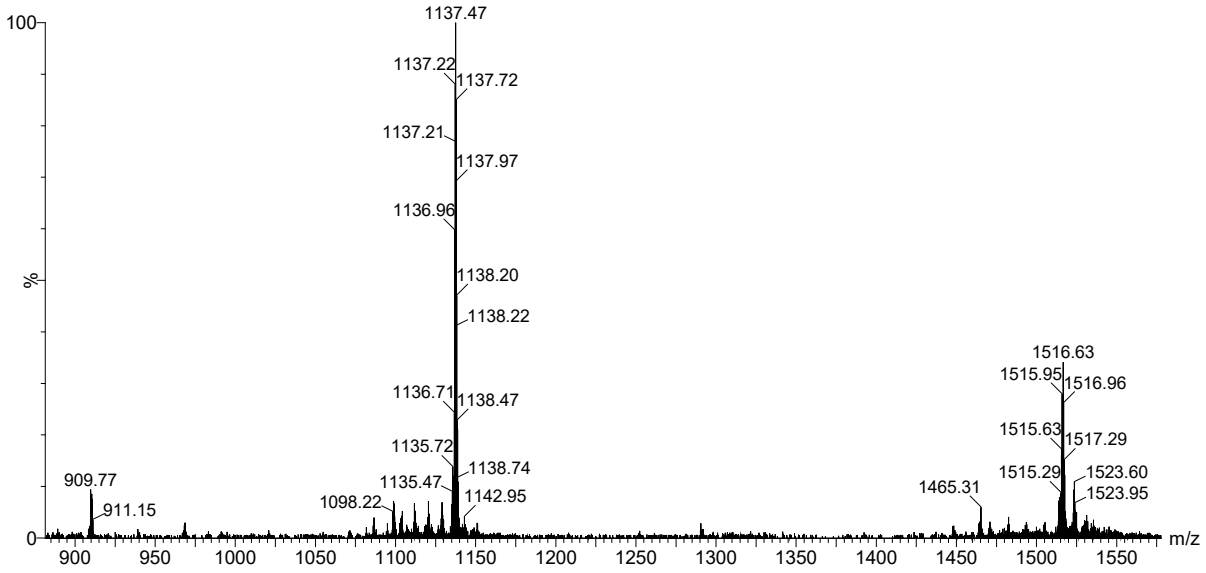


Figure S81. Mass spectrum of hexamer $\mathbf{3}_6$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 909.81 $[M+5H]^{5+}$, 1137.01 $[M+4H]^{4+}$, 1515.68 $[M+3H]^{3+}$; observed m/z: 909.77 $[M+5H]^{5+}$, 1136.71 $[M+4H]^{4+}$, 1515.29 $[M+3H]^{3+}$

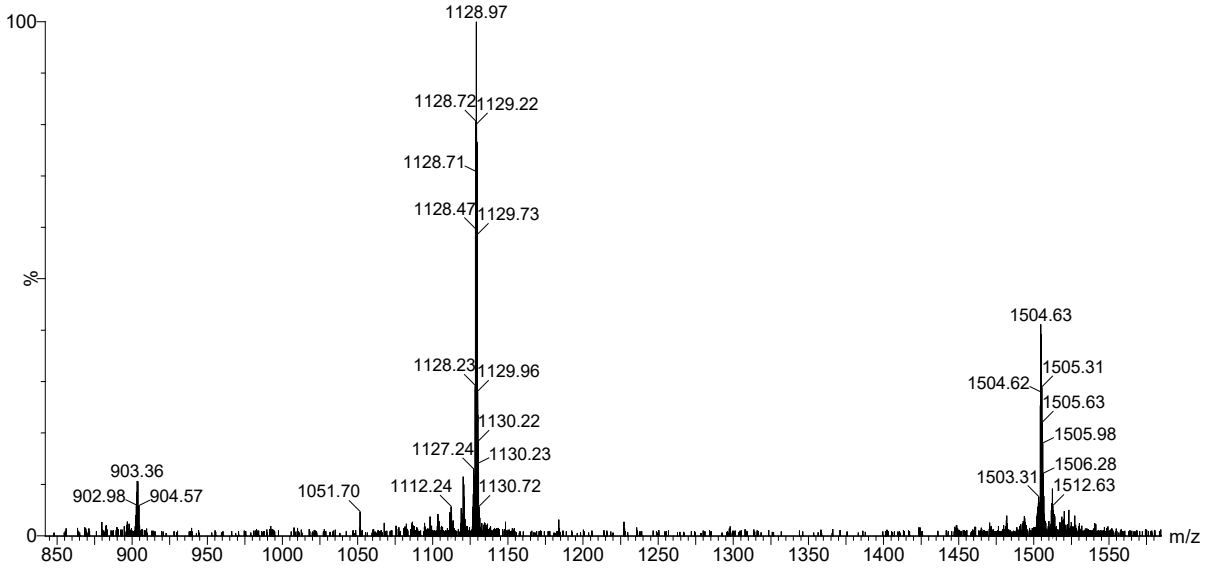


Figure S82. Mass spectrum of hexamer $\mathbf{3}_5\mathbf{4}_1$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 903.01 $[M+5H]^{5+}$, 1128.52 $[M+4H]^{4+}$, 1504.35 $[M+3H]^{3+}$; observed m/z: 902.98 $[M+5H]^{5+}$, 1128.23 $[M+4H]^{4+}$, 1504.62 $[M+3H]^{3+}$

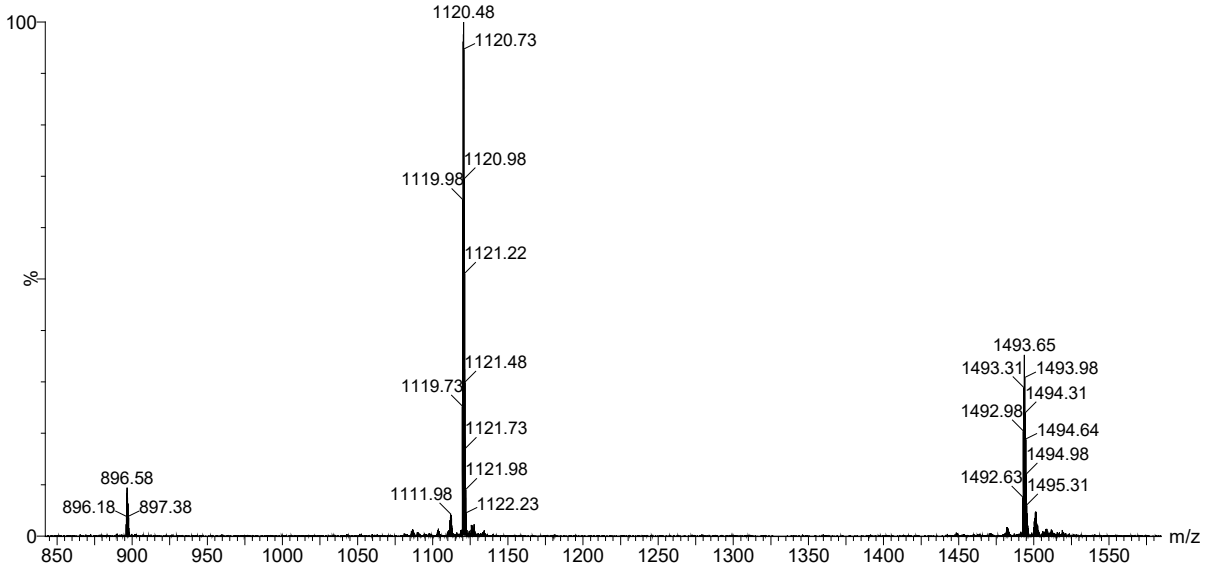


Figure S83. Mass spectrum of hexamer $\mathbf{3}_4\mathbf{4}_2$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 896.22 $[M+5H]^{5+}$, 1120.02 $[M+4H]^{4+}$, 1493.03 $[M+3H]^{3+}$; observed m/z: 896.18 $[M+5H]^{5+}$, 1119.73 $[M+4H]^{4+}$, 1492.63 $[M+3H]^{3+}$

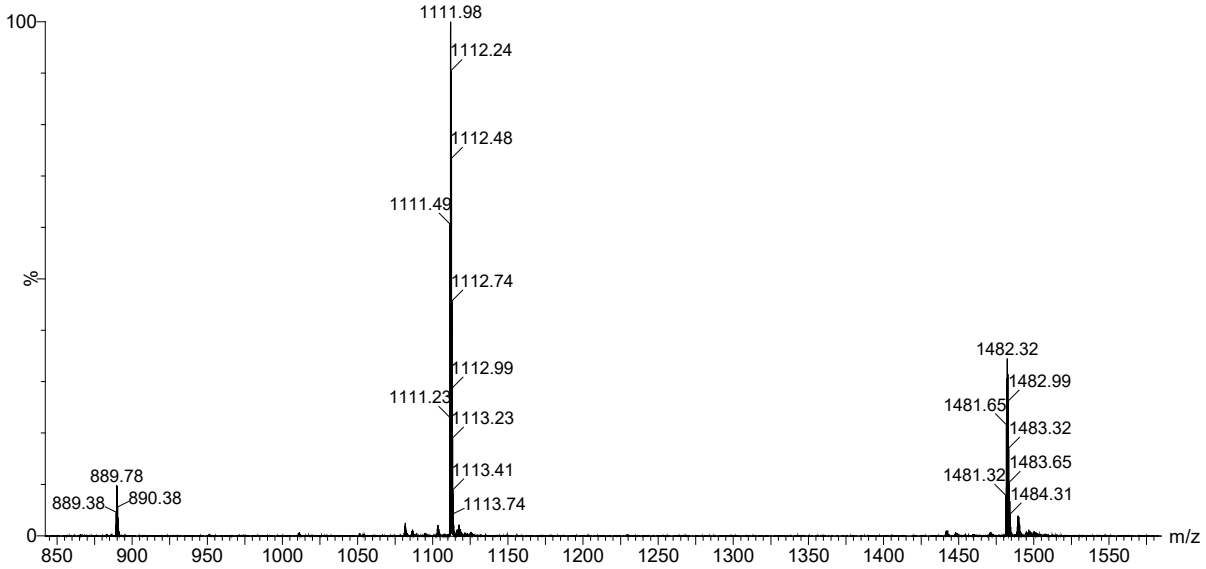


Figure S84. Mass spectrum of hexamer $\mathbf{3}_3\mathbf{4}_3$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 889.42 $[M+5H]^{5+}$, 1111.53 $[M+4H]^{4+}$, 1481.70 $[M+3H]^{3+}$; observed m/z: 889.38 $[M+5H]^{5+}$, 1111.23 $[M+4H]^{4+}$, 1481.32 $[M+3H]^{3+}$

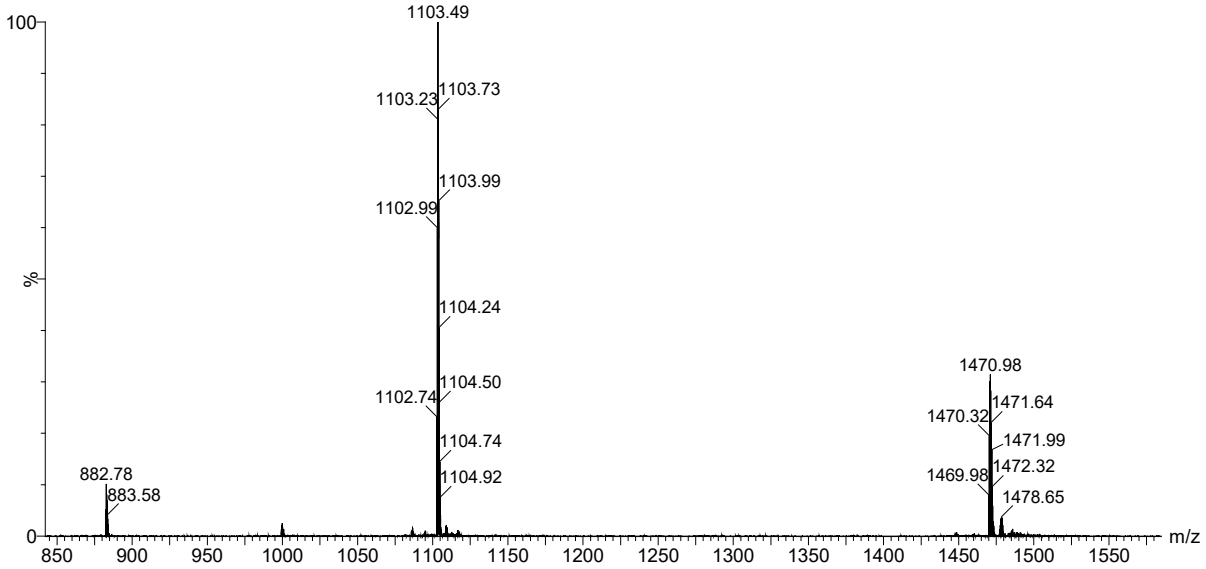


Figure S85. Mass spectrum of hexamer $\mathbf{3}_2\mathbf{4}_4$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 882.62 $[M+5H]^{5+}$, 1103.03 $[M+4H]^{4+}$, 1470.37 $[M+3H]^{3+}$; observed m/z: 882.78 $[M+5H]^{5+}$, 1102.74 $[M+4H]^{4+}$, 1469.98 $[M+3H]^{3+}$

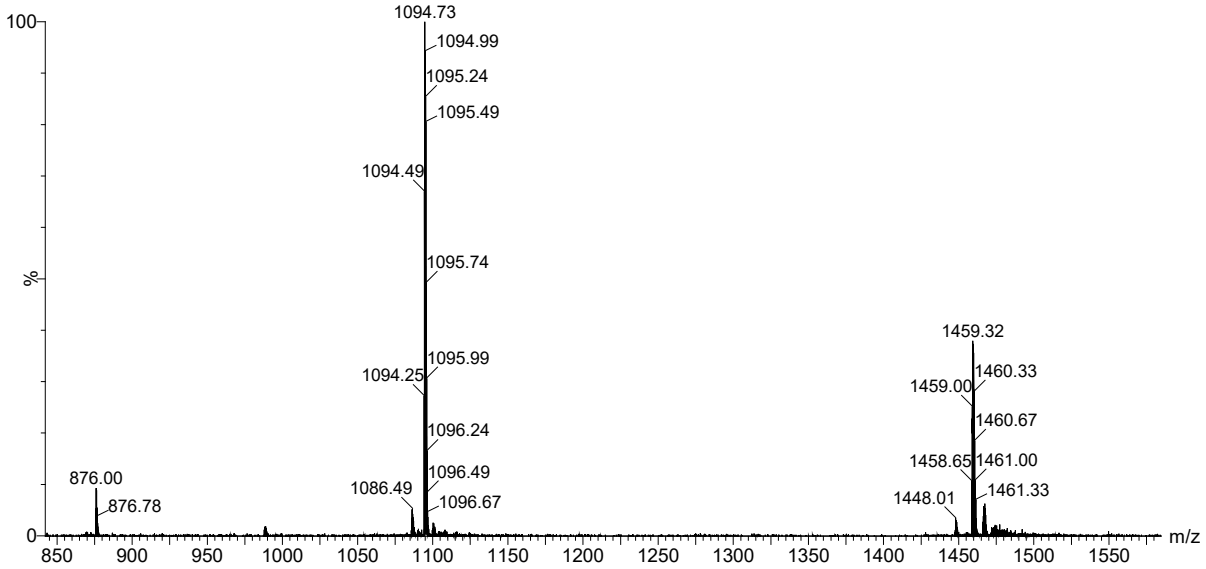


Figure S86. Mass spectrum of hexamer $\mathbf{3}_1\mathbf{4}_5$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 875.83 $[M+5H]^{5+}$, 1094.54 $[M+4H]^{4+}$, 1459.05 $[M+3H]^{3+}$; observed m/z: 876.00 $[M+5H]^{5+}$, 1094.25 $[M+4H]^{4+}$, 1458.65 $[M+3H]^{3+}$

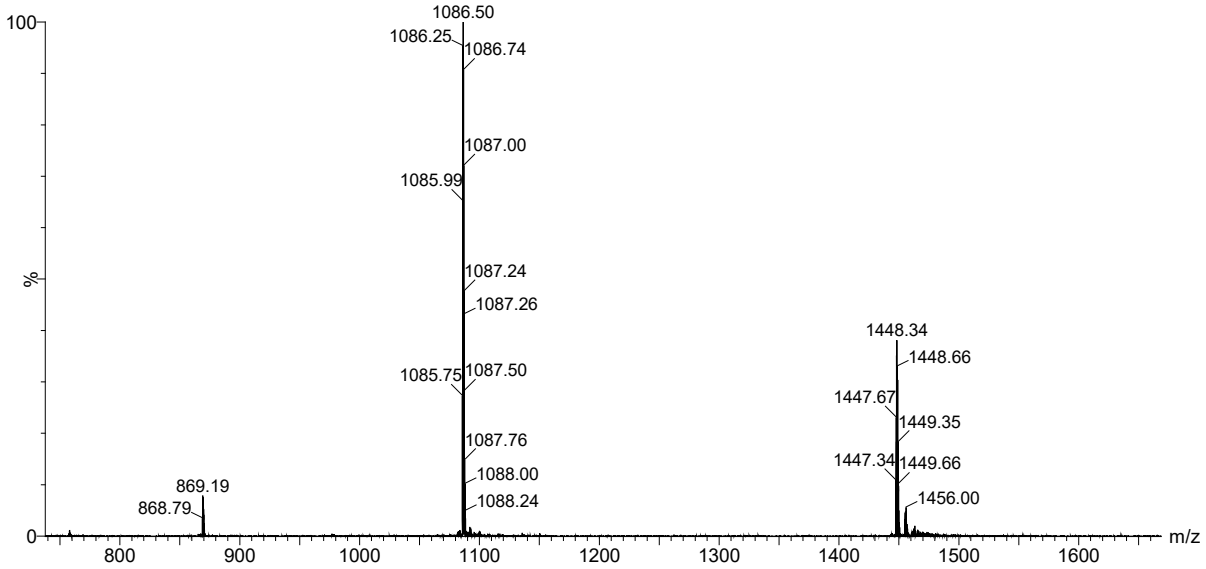


Figure S87. Mass spectrum of hexamer $\mathbf{4}_6$ from the LC-MS analysis of a stirred library made from peptides **3** and **4** (corresponding to Figure S77). Calculated m/z: 869.03 $[M+5H]^{5+}$, 1086.04 $[M+4H]^{4+}$, 1447.72 $[M+3H]^{3+}$; observed m/z: 868.79 $[M+5H]^{5+}$, 1085.75 $[M+4H]^{4+}$, 1447.34 $[M+3H]^{3+}$

XGLKFK (1) + All-D-XGKLKL (8)

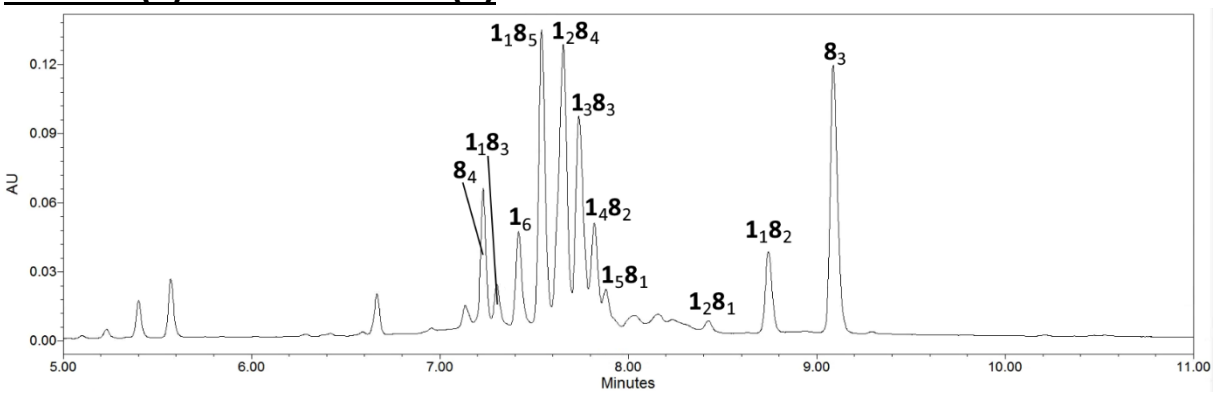


Figure S88. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **8** at 45 °C after 5 days (corresponding to Figure S38).

3mers

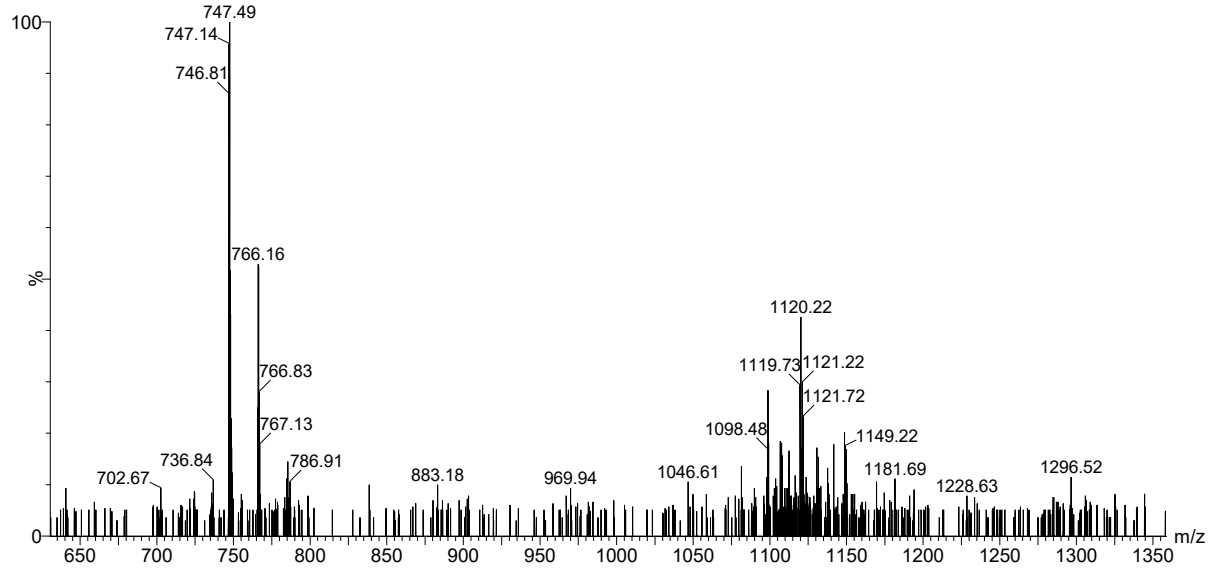


Figure S89. Mass spectrum of trimer $\mathbf{1_2 8_1}$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 747.01 $[M+3H]^{3+}$, 1120.02 $[M+2H]^{2+}$; observed m/z: 746.81 $[M+3H]^{3+}$, 1119.73 $[M+2H]^{2+}$

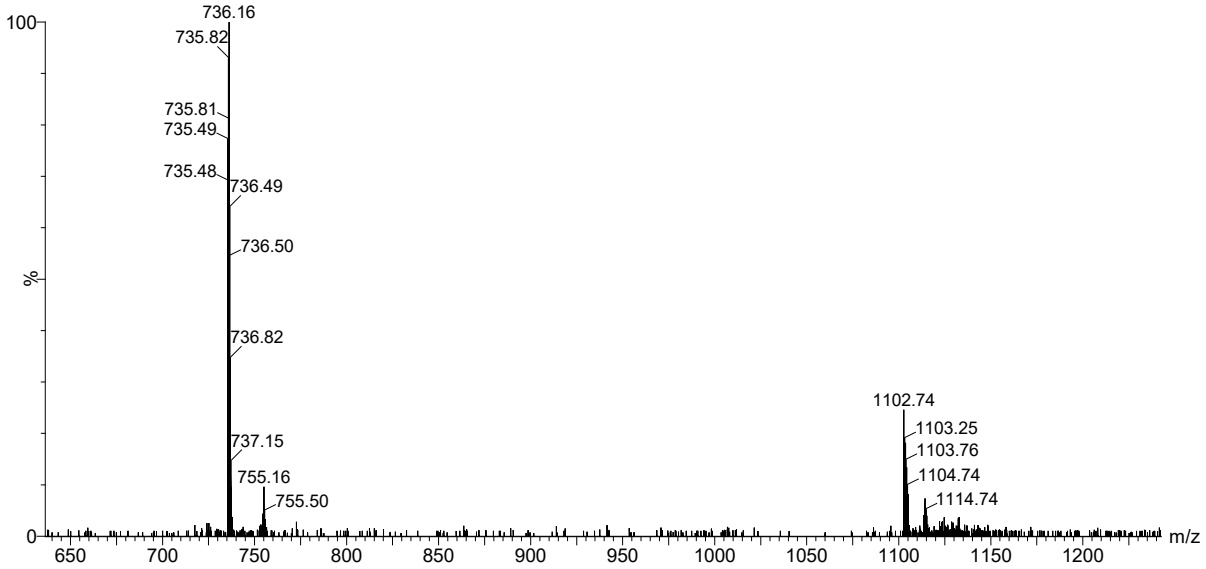


Figure S90. Mass spectrum of trimer $\mathbf{1}_1\mathbf{8}_2$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 735.69 $[M+3H]^{3+}$, 1103.03 $[M+2H]^{2+}$; observed m/z: 735.48 $[M+3H]^{3+}$, 1102.74 $[M+2H]^{2+}$

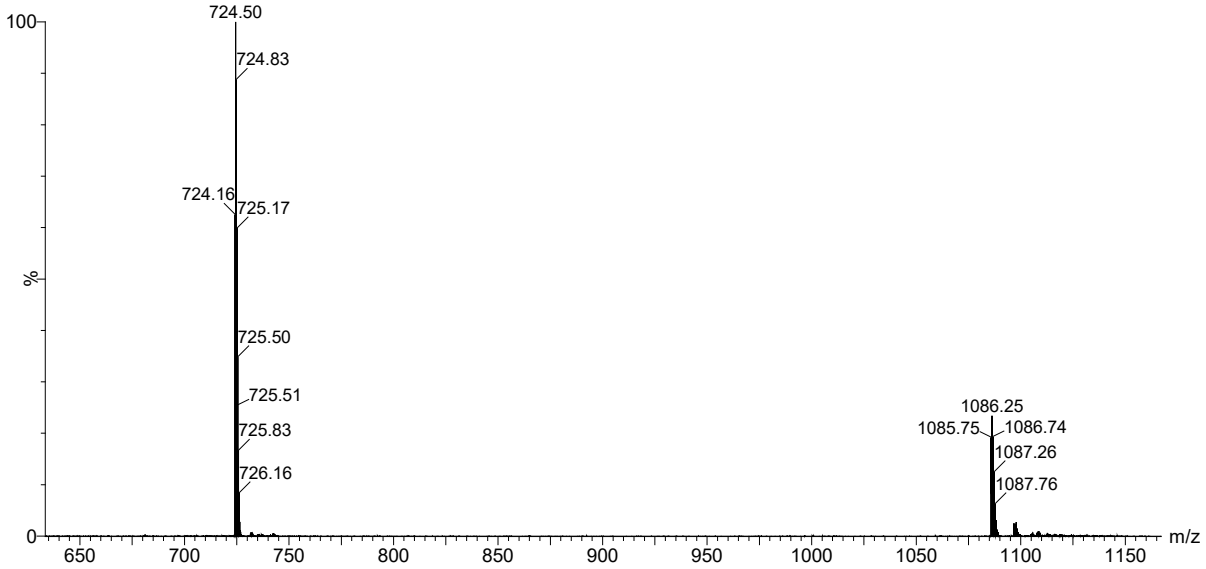


Figure S91. Mass spectrum of trimer $\mathbf{8}_3$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.16 $[M+3H]^{3+}$, 1085.75 $[M+2H]^{2+}$

4mers

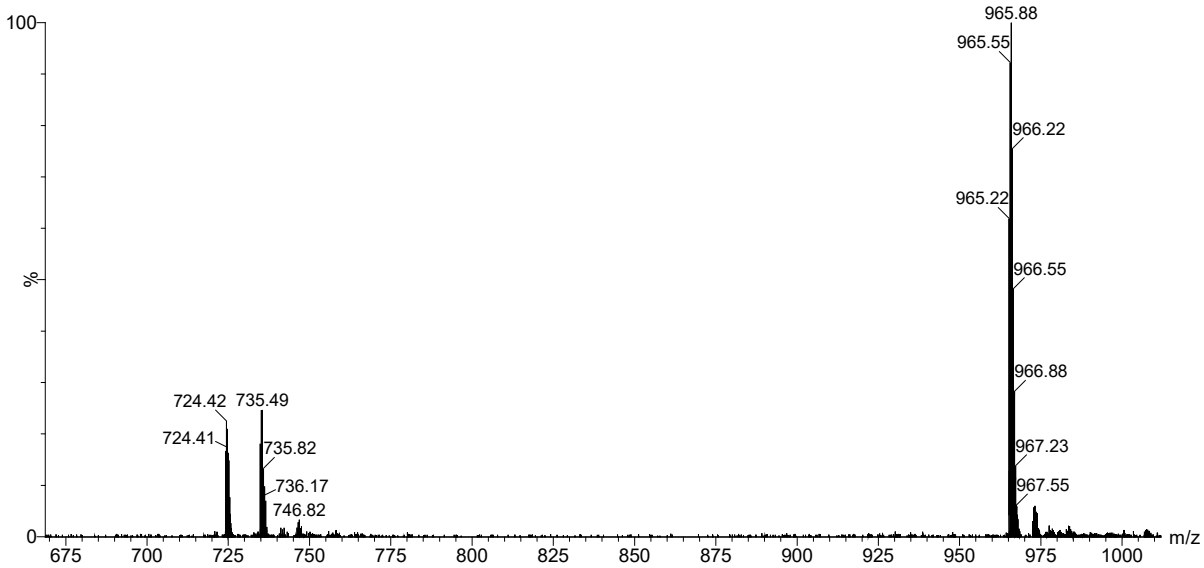


Figure S92. Mass spectrum of tetramer $\mathbf{8}_4$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88). Calculated m/z: 724.36 $[M+4H]^{4+}$, 965.48 $[M+3H]^{3+}$; observed m/z: 724.41 $[M+4H]^{4+}$, 965.22 $[M+3H]^{3+}$

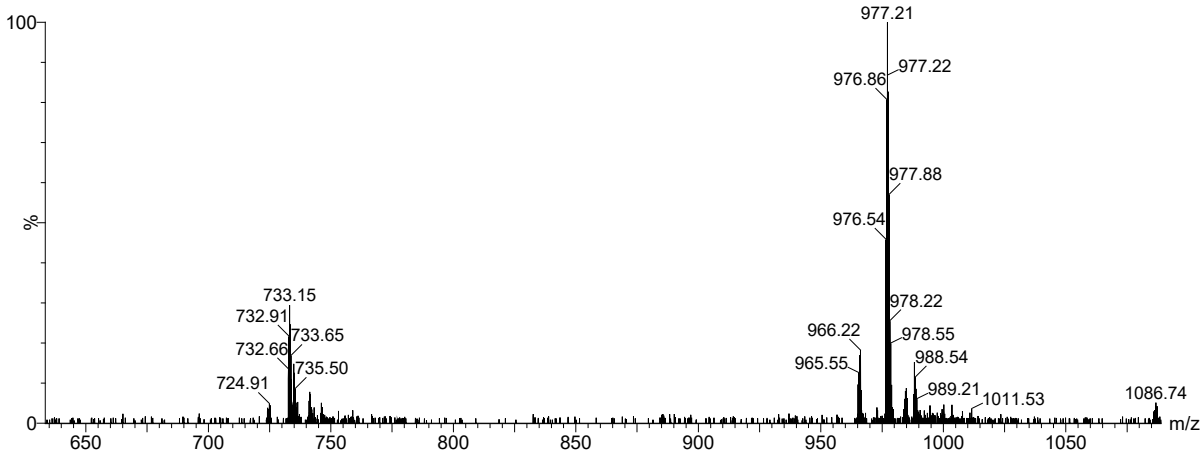


Figure S93. Mass spectrum of tetramer $\mathbf{1}_1\mathbf{8}_3$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 732.86 $[M+4H]^{4+}$, 976.81 $[M+3H]^{3+}$; observed m/z: 732.66 $[M+4H]^{4+}$, 976.54 $[M+3H]^{3+}$

6mers

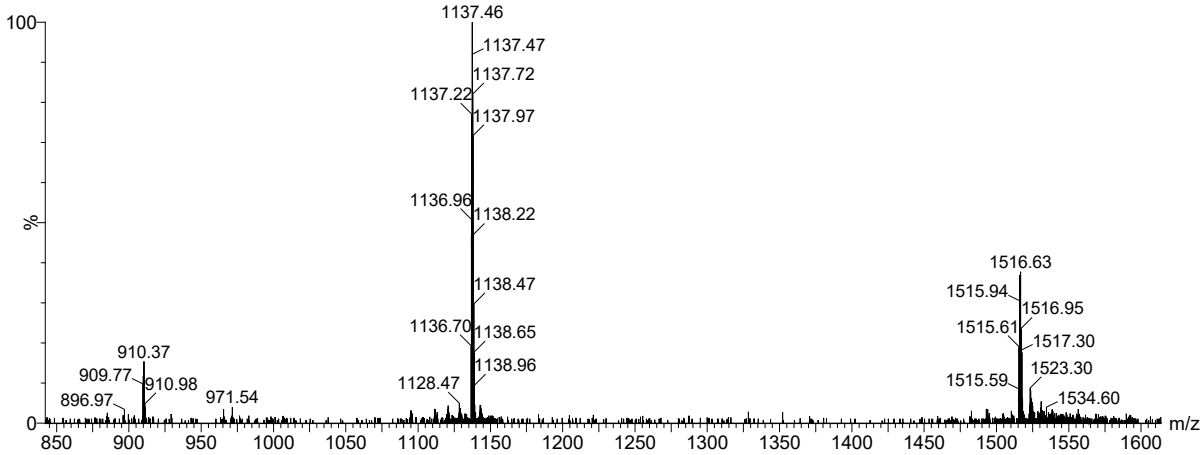


Figure S94. Mass spectrum of hexamer $\mathbf{1}_6$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 909.81 $[M+5H]^{5+}$, 1137.01 $[M+4H]^{4+}$, 1515.68 $[M+3H]^{3+}$; observed m/z: 909.77 $[M+5H]^{5+}$, 1136.70 $[M+4H]^{4+}$, 1515.59 $[M+3H]^{3+}$

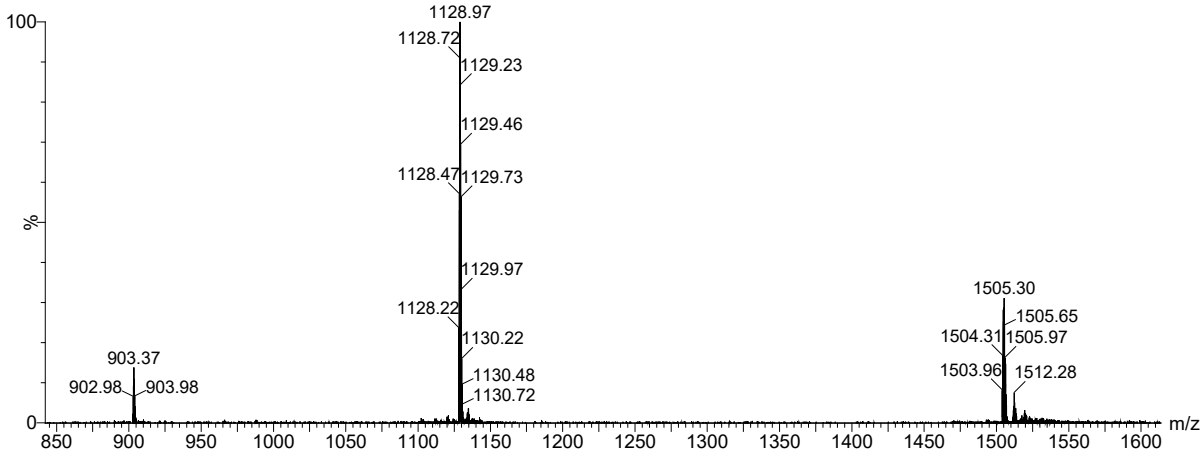


Figure S95. Mass spectrum of hexamer $\mathbf{1}_5\mathbf{8}_1$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 903.01 $[M+5H]^{5+}$, 1128.52 $[M+4H]^{4+}$, 1504.35 $[M+3H]^{3+}$; observed m/z: 902.98 $[M+5H]^{5+}$, 1128.22 $[M+4H]^{4+}$, 1503.96 $[M+3H]^{3+}$

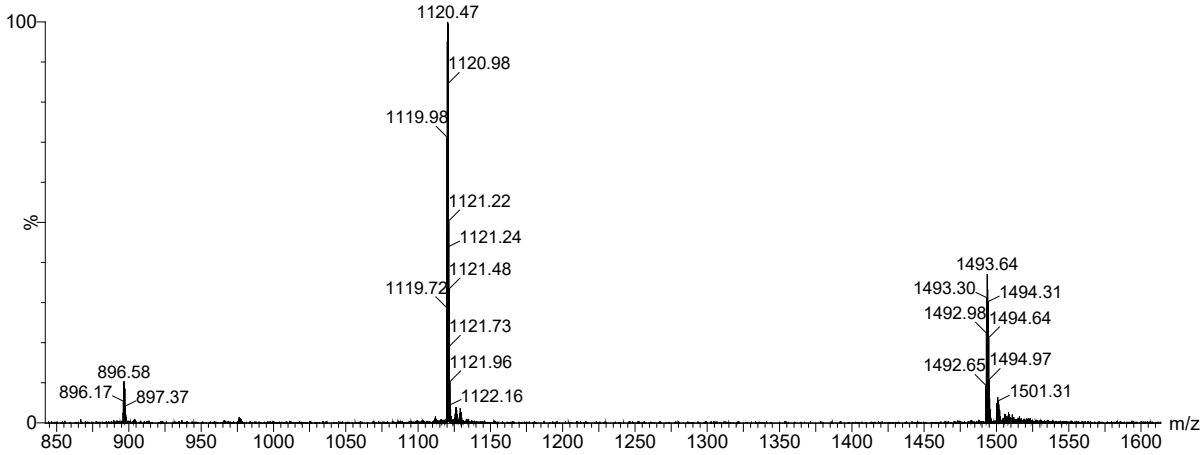


Figure S96. Mass spectrum of hexamer $\mathbf{1}_4\mathbf{8}_2$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 896.22 $[M+5H]^{5+}$, 1120.02 $[M+4H]^{4+}$, 1493.03 $[M+3H]^{3+}$; observed m/z: 896.17 $[M+5H]^{5+}$, 1119.72 $[M+4H]^{4+}$, 1492.65 $[M+3H]^{3+}$

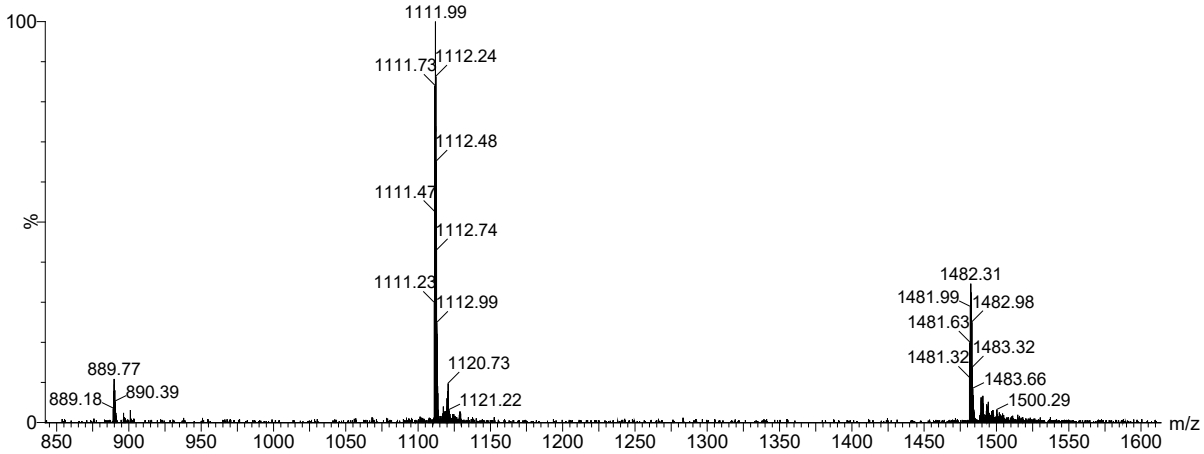


Figure S97. Mass spectrum of hexamer $\mathbf{1}_3\mathbf{8}_3$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 889.42 $[M+5H]^{5+}$, 1111.53 $[M+4H]^{4+}$, 1481.70 $[M+3H]^{3+}$; observed m/z: 889.18 $[M+5H]^{5+}$, 1111.23 $[M+4H]^{4+}$, 1481.32 $[M+3H]^{3+}$

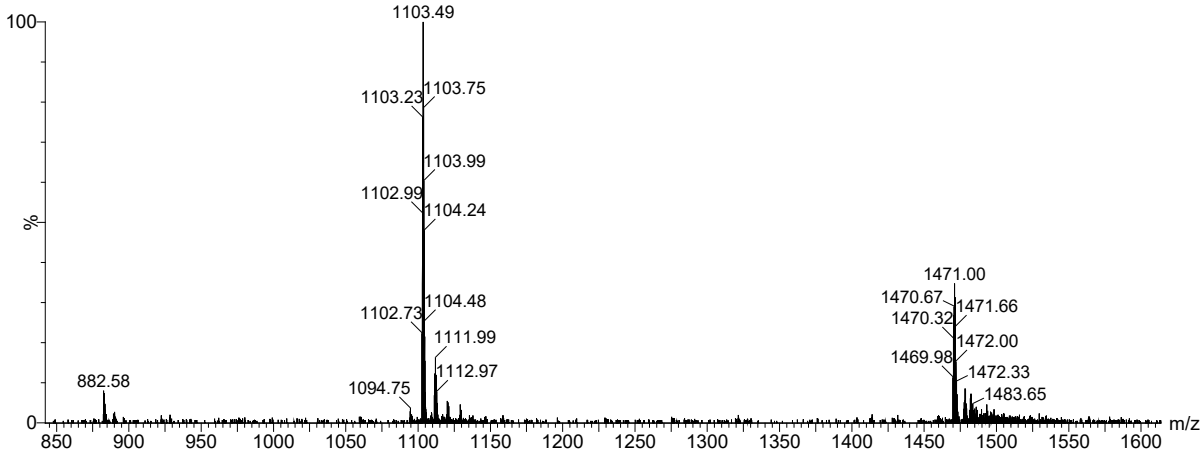


Figure S98. Mass spectrum of hexamer $\mathbf{1}_2\mathbf{8}_4$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 882.62 $[M+5H]^{5+}$, 1103.03 $[M+4H]^{4+}$, 1470.37 $[M+3H]^{3+}$; observed m/z: 882.58 $[M+5H]^{5+}$, 1102.73 $[M+4H]^{4+}$, 1469.98 $[M+3H]^{3+}$

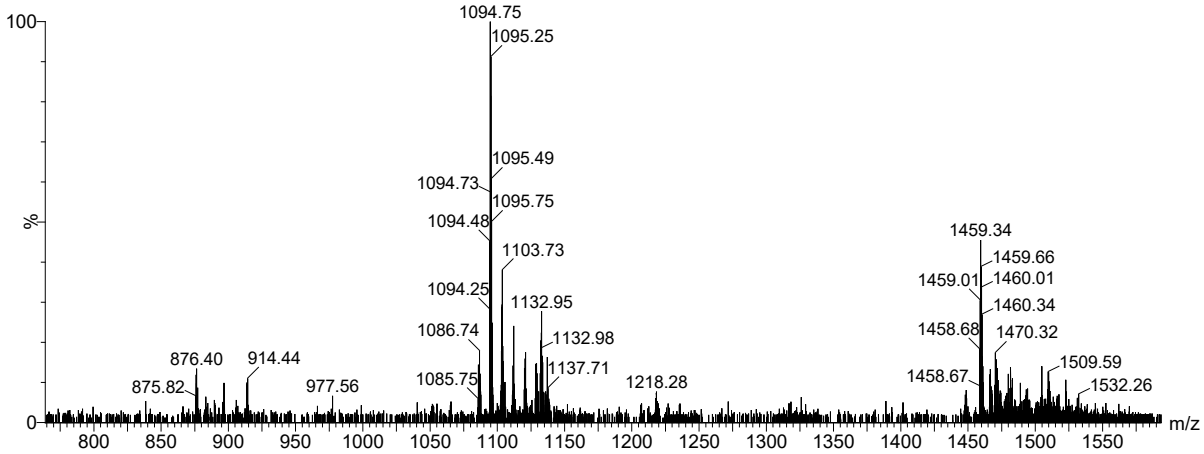


Figure S99. Mass spectrum of hexamer $\mathbf{1}_1\mathbf{8}_5$ from the LC-MS analysis of a stirred library made from peptides **1** and **8** (corresponding to Figure S88)**.** Calculated m/z: 875.83 $[M+5H]^{5+}$, 1094.54 $[M+4H]^{4+}$, 1459.05 $[M+3H]^{3+}$; observed m/z: 875.82 $[M+5H]^{5+}$, 1094.25 $[M+4H]^{4+}$, 1458.68 $[M+3H]^{3+}$

X-GLKLK (5) + X-GKLKL (2)

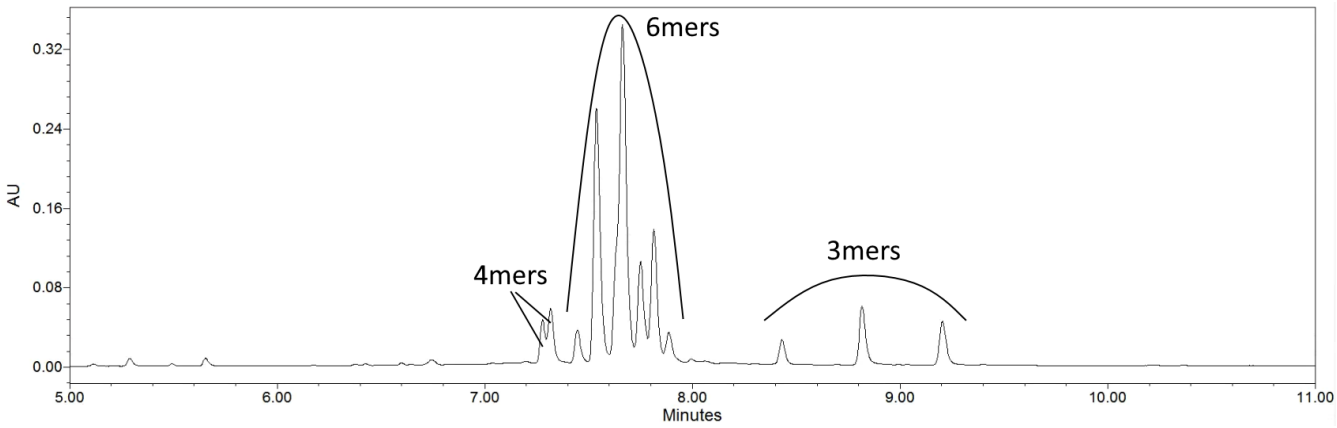


Figure S100. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **2** and **5** at 45 °C after 5 days (corresponding to Figure S32).

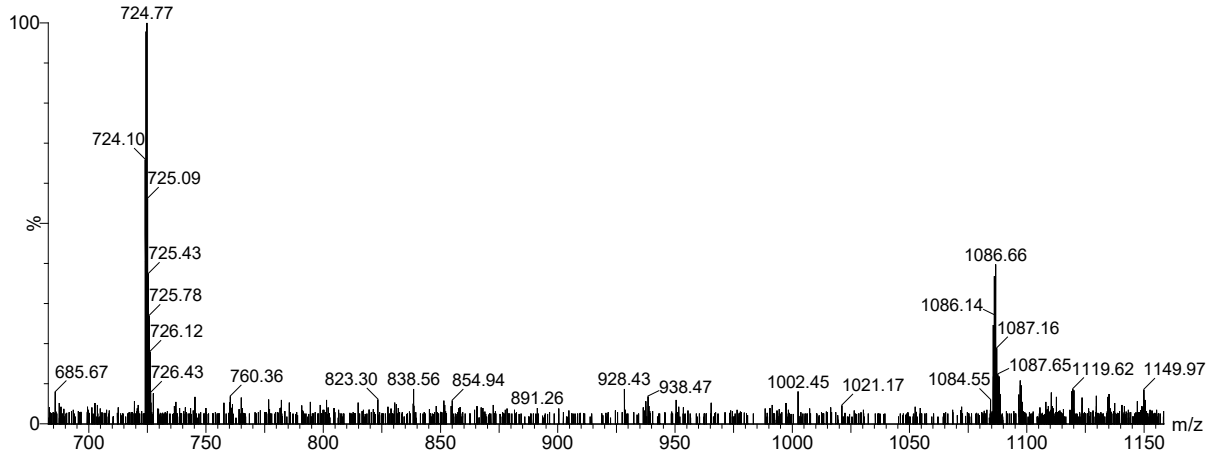


Figure S101. Combined mass spectra of trimer peaks from the LC-MS analysis of a library made from peptides **2** and **5** after 4 days (corresponding to Figure S100). Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.10 $[M+3H]^{3+}$, 1086.14 $[M+2H]^{2+}$

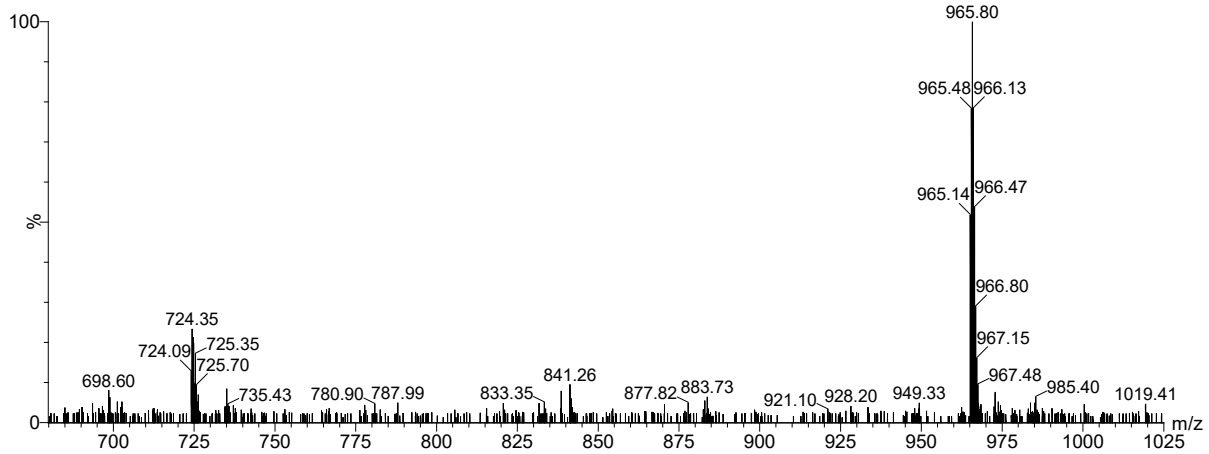


Figure S102. Combined mass spectra of tetramer peaks from the LC-MS analysis of a library made from peptides **2** and **5** after 4 days (corresponding to Figure S100). Calculated m/z: 724.36 $[M+4H]^{4+}$, 965.48 $[M+3H]^{3+}$; observed m/z: 724.09 $[M+4H]^{4+}$, 965.14 $[M+3H]^{3+}$

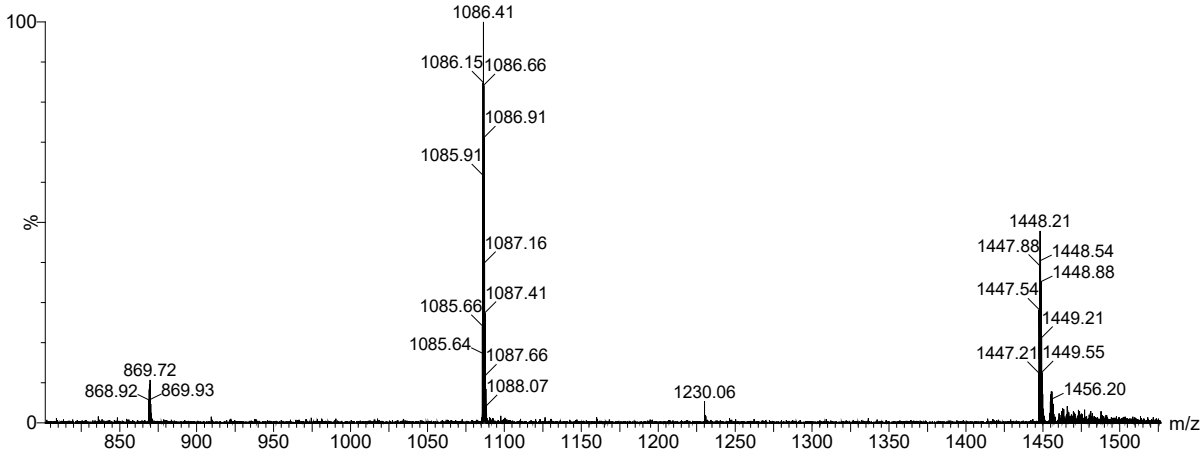


Figure S103. Combined mass spectra of hexamer peaks from the LC-MS analysis of a library made from peptides **2** and **5** after 4 days (corresponding to Figure S100). Calculated m/z: 869.03 $[M+5H]^{5+}$, 1086.04 $[M+4H]^{4+}$, 1447.72 $[M+3H]^{3+}$; observed m/z: 868.92 $[M+5H]^{5+}$, 1085.64 $[M+4H]^{4+}$, 1447.21 $[M+3H]^{3+}$

X-GLK[14]K (6) + XGKLKL (2)

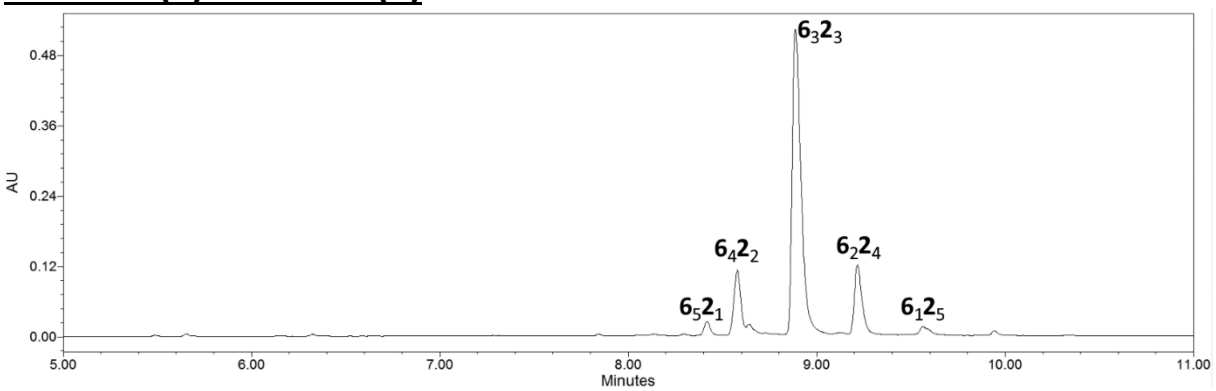


Figure S104. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **2** and **6** at 45 °C after 5 days (corresponding to Figure S34).

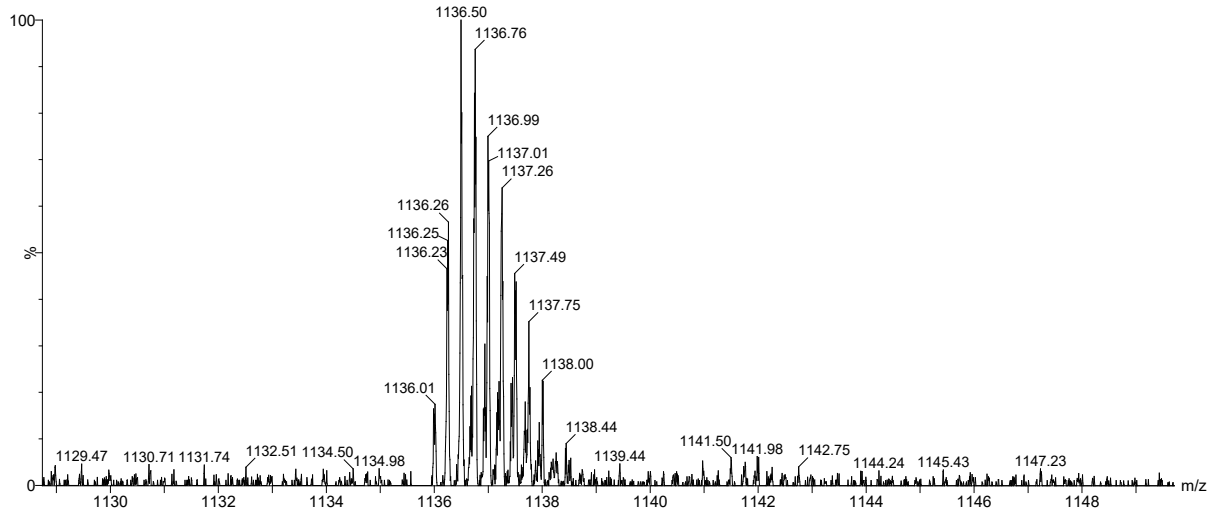


Figure S105. Mass spectrum of hexamer $6_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **2** and **6** (corresponding to Figure S104). Calculated m/z: 1136.08 $[M+4H]^{4+}$; observed m/z:, 1136.01 $[M+4H]^{4+}$

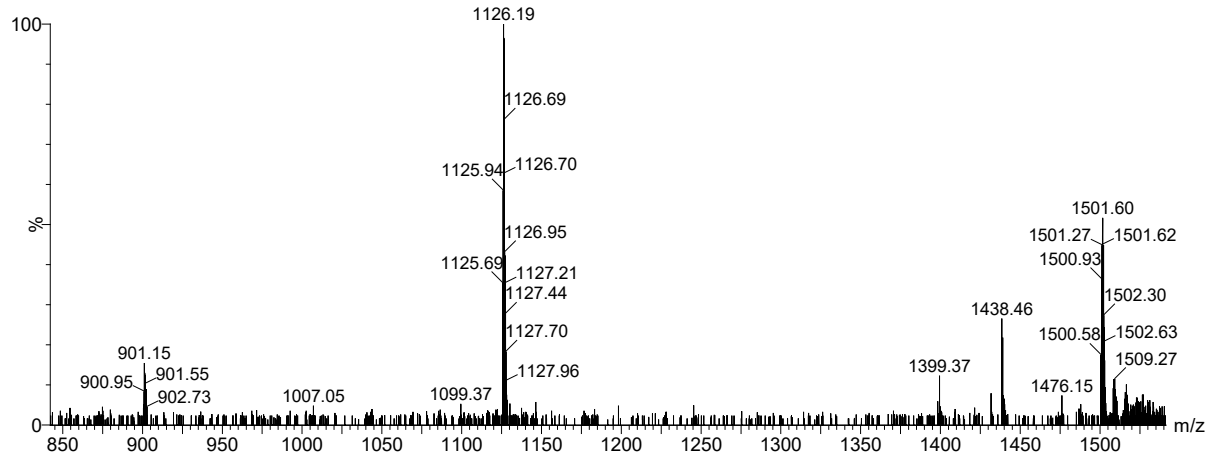


Figure S106. Mass spectrum of hexamer $6_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **2** and **6** (corresponding to Figure S104). Calculated m/z: 901.06 $[M+5H]^{5+}$, 1126.07 $[M+4H]^{4+}$, 1501.09 $[M+3H]^{3+}$; observed m/z: 900.95 $[M+5H]^{5+}$, 1125.69 $[M+4H]^{4+}$, 1500.93 $[M+3H]^{3+}$

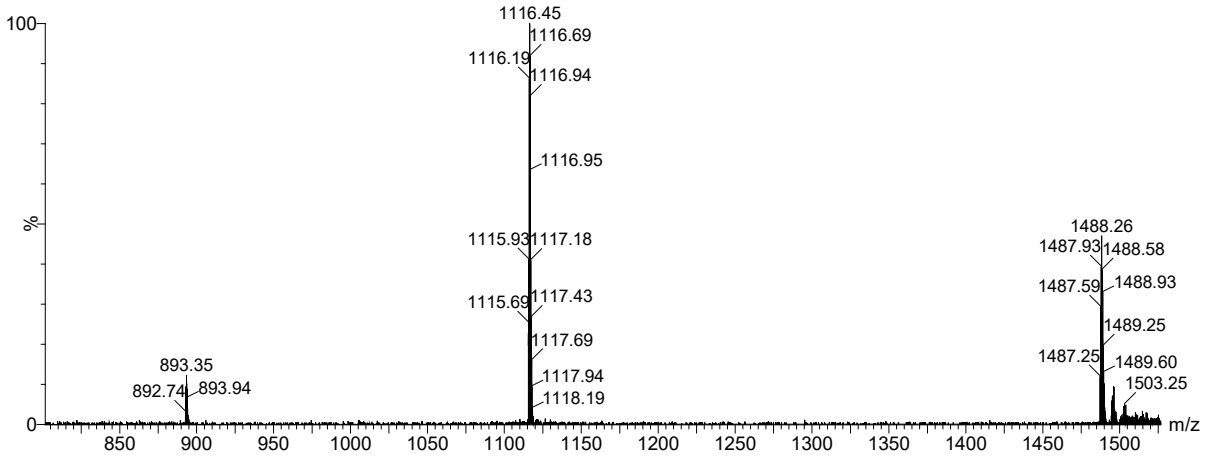


Figure S107. Mass spectrum of hexamer $\mathbf{6}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **2** and **6** (corresponding to Figure S104). Calculated m/z: 893.05 $[M+5H]^{5+}$, 1116.06 $[M+4H]^{4+}$, 1487.75 $[M+3H]^{3+}$; observed m/z: 892.74 $[M+5H]^{5+}$, 1115.69 $[M+4H]^{4+}$, 1487.25 $[M+3H]^{3+}$

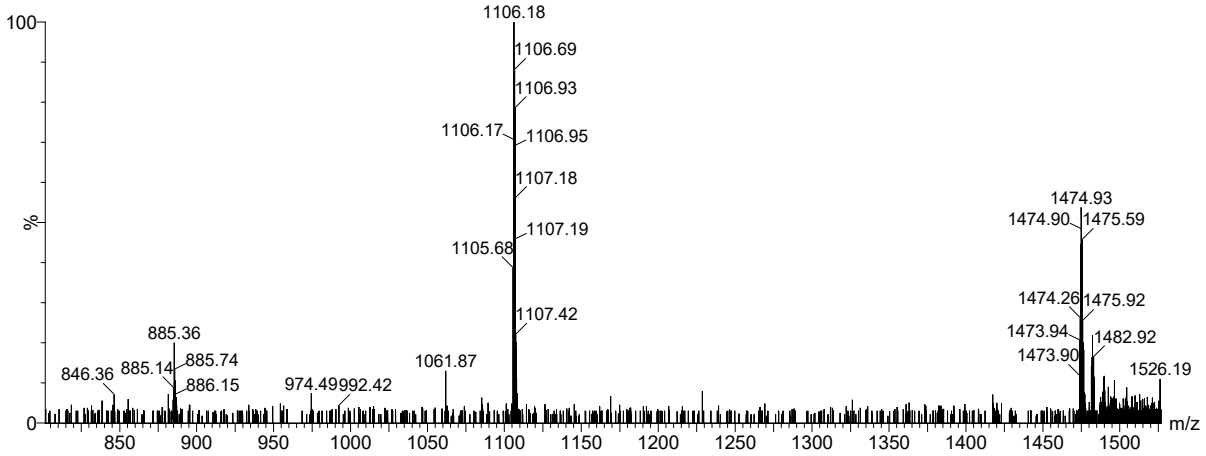


Figure S108. Mass spectrum of hexamer $\mathbf{6}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **2** and **6** (corresponding to Figure S104). Calculated m/z: 885.04 $[M+5H]^{5+}$, 1106.06 $[M+4H]^{4+}$, 1474.41 $[M+3H]^{3+}$; observed m/z: 885.14 $[M+5H]^{5+}$, 1105.68 $[M+4H]^{4+}$, 1473.90 $[M+3H]^{3+}$

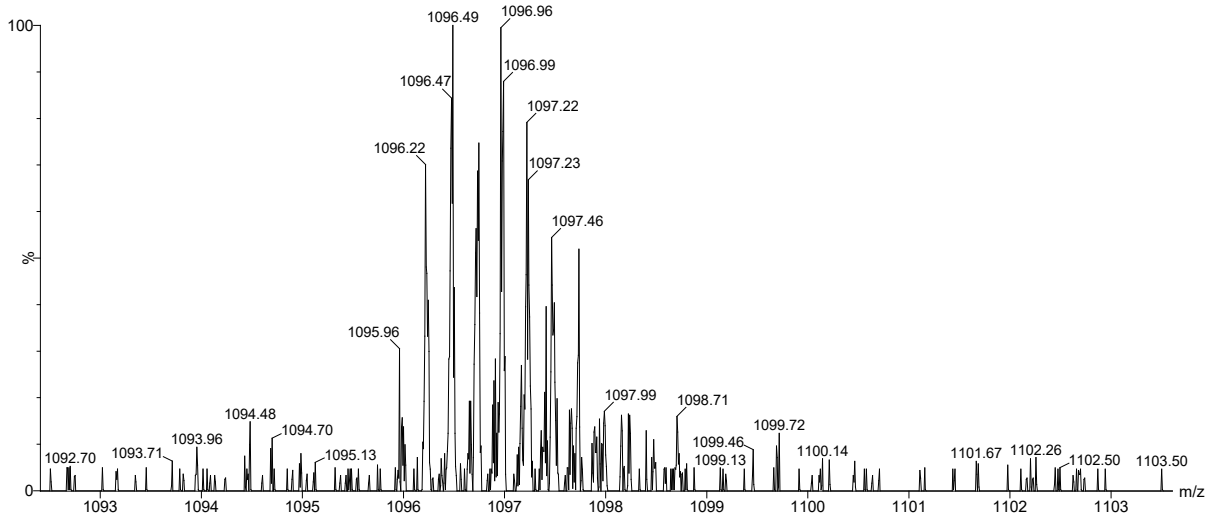


Figure S109. Mass spectrum of hexamer $\mathbf{6}_1\mathbf{2}_5$ from the LC-MS analysis of a stirred library made from peptides **2** and **6** (corresponding to Figure S104). Calculated m/z: 1096.05 $[M+4H]^{4+}$; observed m/z:, 1095.96 $[M+4H]^{4+}$

X-GLK{1-Nal}K (7) + X-GKLKL (2)

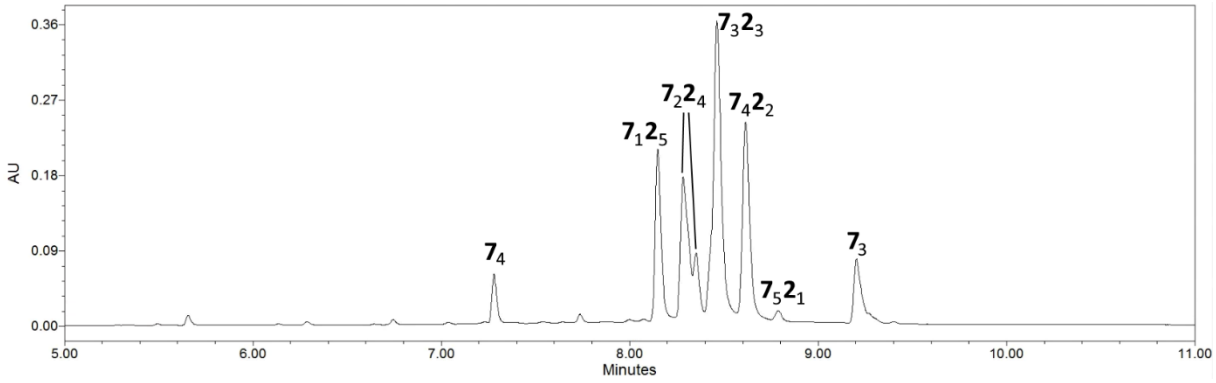


Figure S110. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **2** and **7** at 45 °C after 4 days (corresponding to Figure S35).

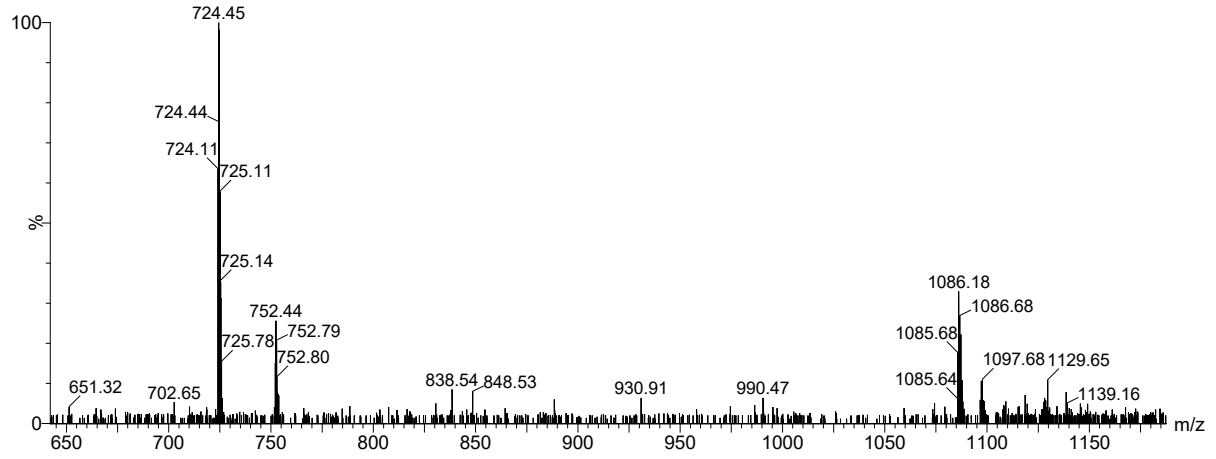


Figure S111. Mass spectrum of trimer $\mathbf{7}_3$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.11 $[M+3H]^{3+}$, 1085.64 $[M+2H]^{2+}$

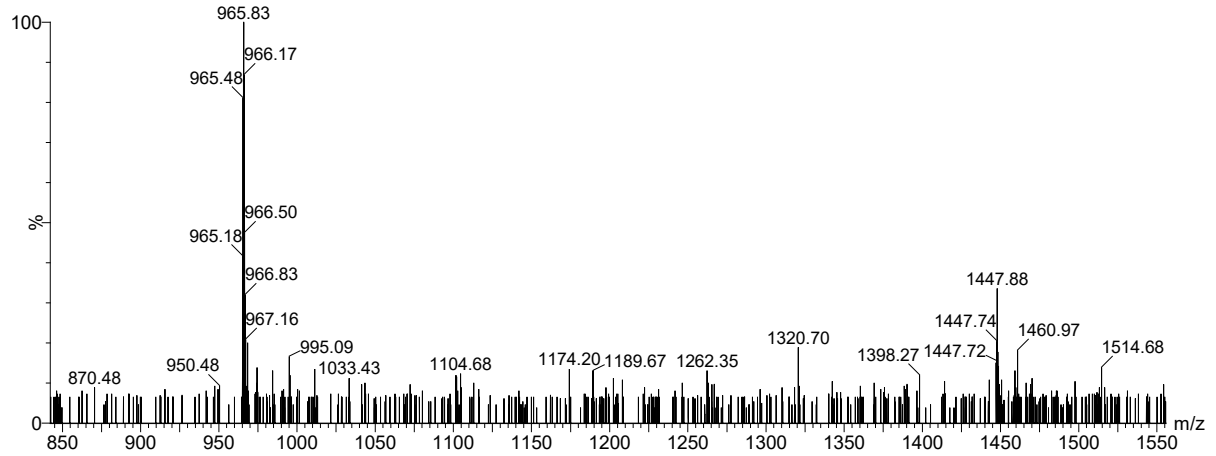


Figure S112. Mass spectrum of tetramer $\mathbf{7}_4$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 965.48 $[M+3H]^{3+}$, 1447.72 $[M+2H]^{2+}$; observed m/z: 965.18 $[M+3H]^{3+}$, 1447.72 $[M+2H]^{2+}$

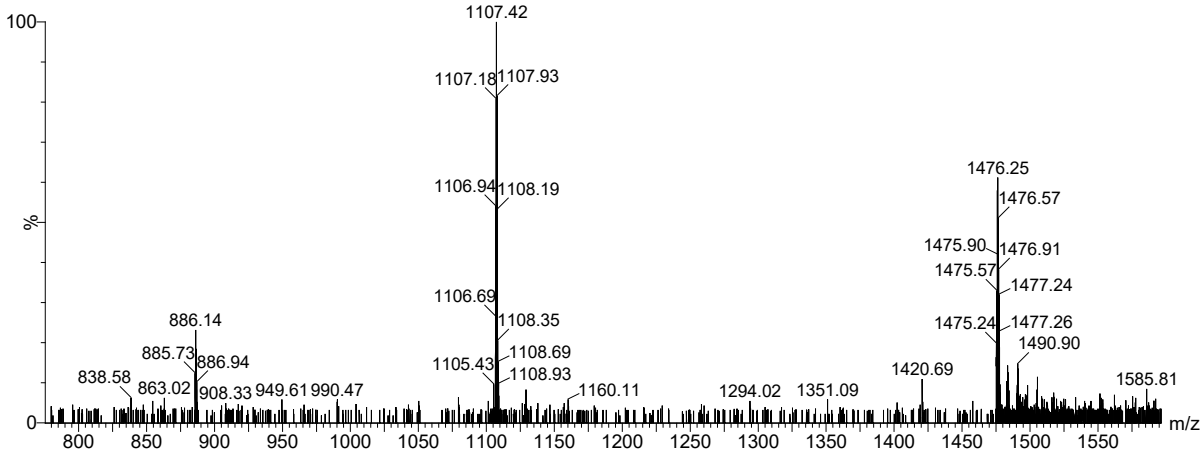


Figure S113. Mass spectrum of hexamer $\mathbf{7}_1\mathbf{2}_5$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 885.83 $[M+5H]^{5+}$, 1107.04 $[M+4H]^{4+}$, 1475.72 $[M+3H]^{3+}$; observed m/z: 885.73 $[M+5H]^{5+}$, 1106.69 $[M+4H]^{4+}$, 1475.24 $[M+3H]^{3+}$

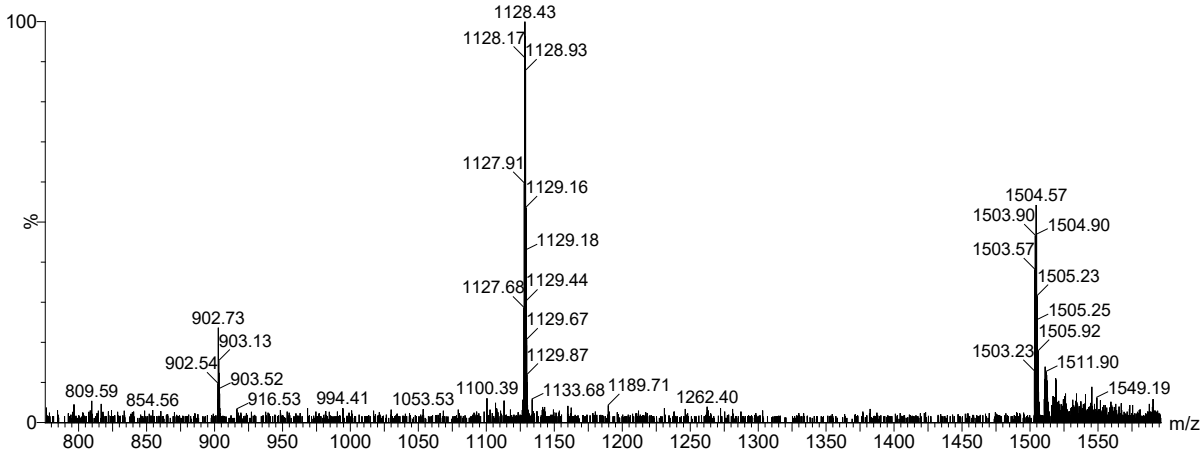


Figure S114. Mass spectrum of hexamer $\mathbf{7}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 902.63 $[M+5H]^{5+}$, 1128.04 $[M+4H]^{4+}$, 1503.72 $[M+3H]^{3+}$; observed m/z: 902.54 $[M+5H]^{5+}$, 1127.68 $[M+4H]^{4+}$, 1503.23 $[M+3H]^{3+}$

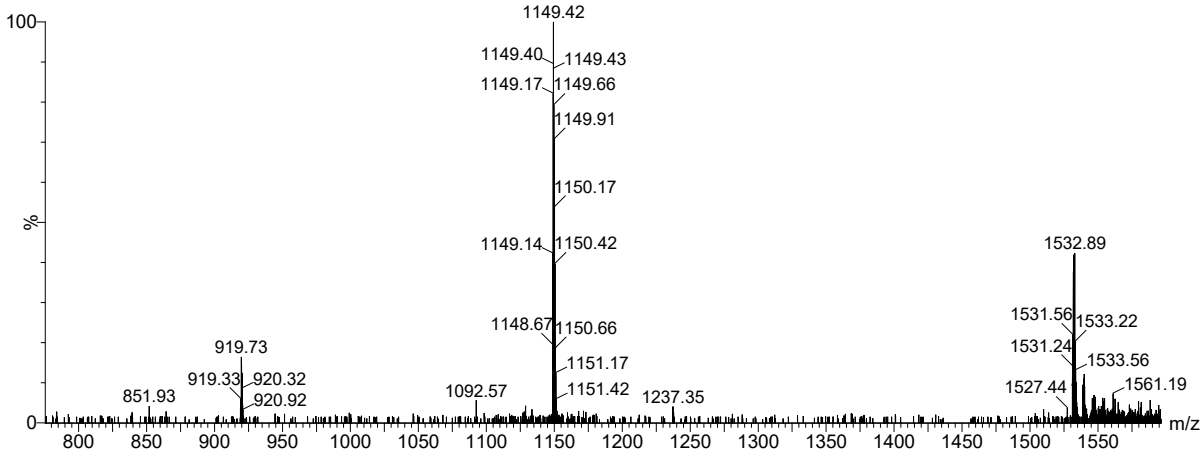


Figure S115. Mass spectrum of hexamer $\mathbf{7}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 919.43 $[M+5H]^{5+}$, 1149.04 $[M+4H]^{4+}$, 1531.72 $[M+3H]^{3+}$; observed m/z: 919.33 $[M+5H]^{5+}$, 1148.67 $[M+4H]^{4+}$, 1531.24 $[M+3H]^{3+}$

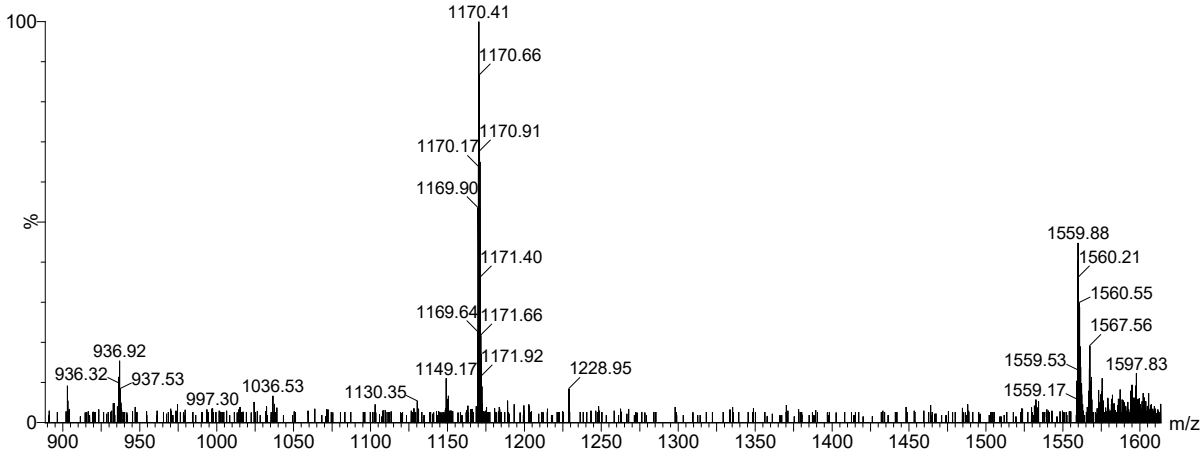


Figure S116. Mass spectrum of hexamer $\mathbf{7}_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 936.23 $[M+5H]^{5+}$, 1170.04 $[M+4H]^{4+}$, 1559.72 $[M+3H]^{3+}$; observed m/z: 936.32 $[M+5H]^{5+}$, 1169.64 $[M+4H]^{4+}$, 1559.53 $[M+3H]^{3+}$

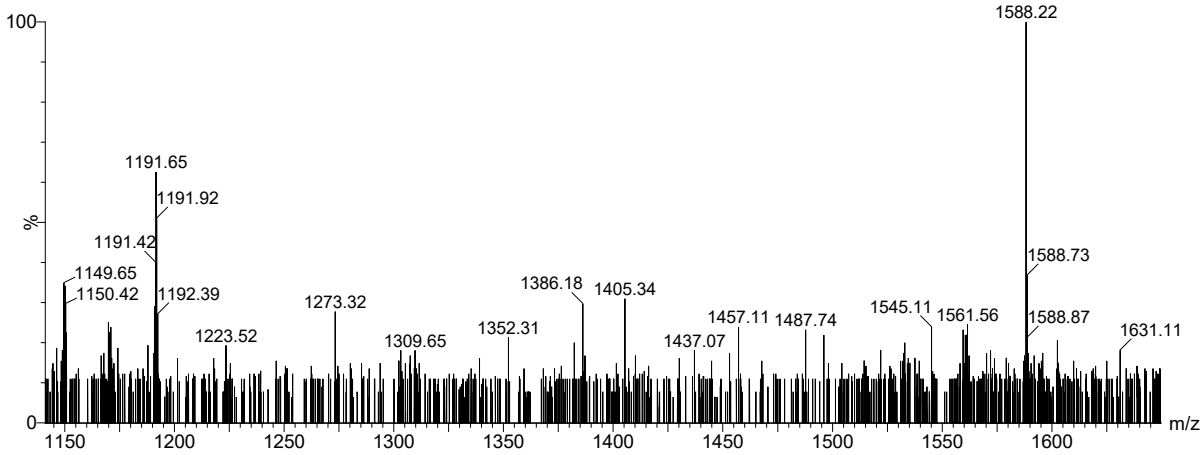


Figure S117. Mass spectrum of hexamer $\mathbf{7}_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **7** and **2** after 4 days (corresponding to Figure S110). Calculated m/z: 1191.04 $[M+4H]^{4+}$, 1587.72 $[M+3H]^{3+}$; observed m/z: 1191.42 $[M+4H]^{4+}$, 1588.22 $[M+3H]^{3+}$

X-GLKFK (1) + X-GLKLK (5)

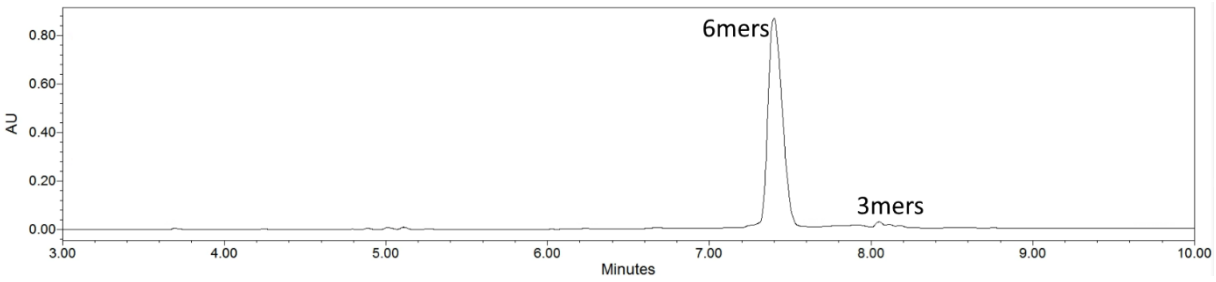


Figure S118. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **1** and **5** at 45 °C after 5 days (corresponding to **Figure S31**).

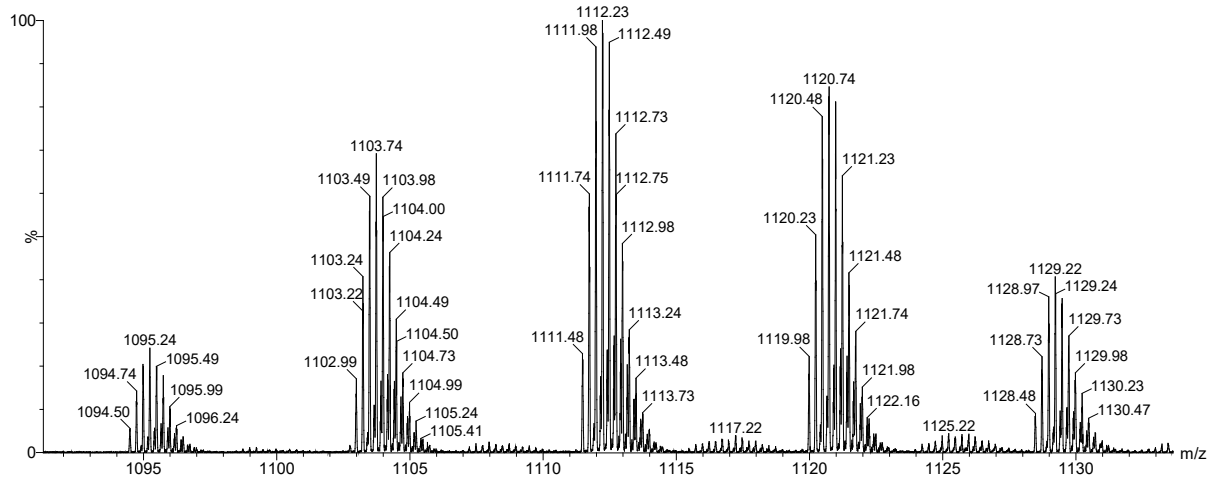


Figure S119. Partial mass spectrum of hexamers $\mathbf{1}_5\mathbf{5}_1$, $\mathbf{1}_4\mathbf{5}_2$, $\mathbf{1}_3\mathbf{5}_3$, $\mathbf{1}_2\mathbf{5}_4$ and $\mathbf{1}_1\mathbf{5}_5$ from the LC-MS analysis of a stirred library made from peptides **1** and **5** after 5 days (corresponding to **Figure S118**). These hexamers (together with $\mathbf{1}_6$ and $\mathbf{5}_6$) elute in a single peak in the LC-MS analysis. $\mathbf{1}_5\mathbf{5}_1$: Calculated m/z: 1128.52 $[M+4H]^{4+}$; observed m/z: 1128.48 $[M+4H]^{4+}$. $\mathbf{1}_4\mathbf{5}_2$: Calculated m/z: 1120.02 $[M+4H]^{4+}$; observed m/z: 1119.98 $[M+4H]^{4+}$. $\mathbf{1}_3\mathbf{5}_3$: Calculated m/z: 1111.53 $[M+4H]^{4+}$; observed m/z: 1111.48 $[M+4H]^{4+}$. $\mathbf{1}_2\mathbf{5}_4$: Calculated m/z: 1103.03 $[M+4H]^{4+}$; observed m/z: 1102.99 $[M+4H]^{4+}$. $\mathbf{1}_1\mathbf{5}_5$: Calculated m/z: 1094.54 $[M+4H]^{4+}$; observed m/z: 1094.50 $[M+4H]^{4+}$.

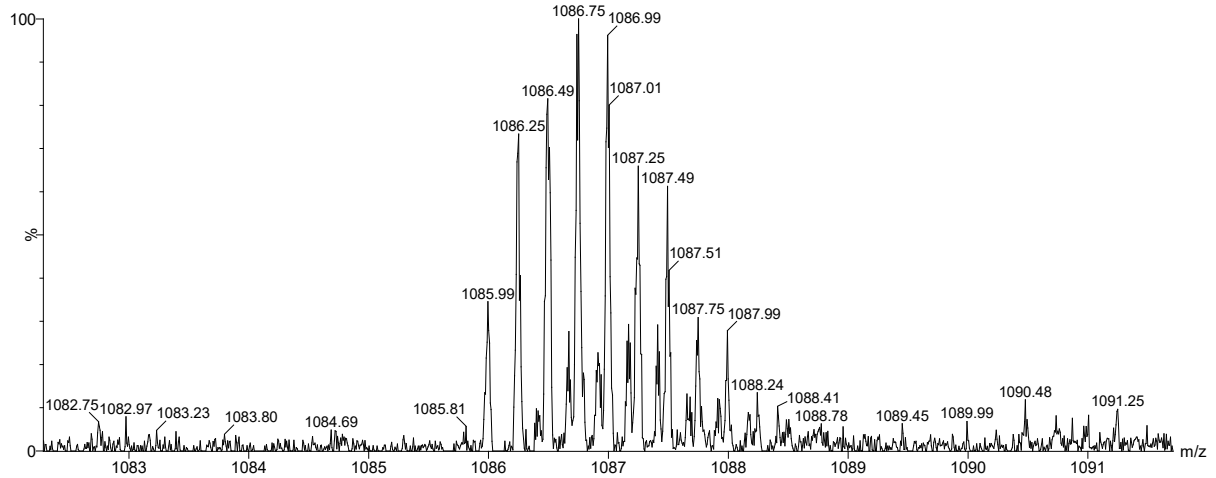


Figure S120. Partial mass spectrum of hexamer $\mathbf{5}_6$ from the LC-MS analysis of a stirred library made from peptides **1** and **5** after 5 days (corresponding to **Figure S118**). All hexamers elute in a single peak

in the LC-MS analysis.
5$_6$: Calculated m/z: 1086.04 $[M+4H]^{4+}$; observed m/z: 1085.99 $[M+4H]^{4+}$.

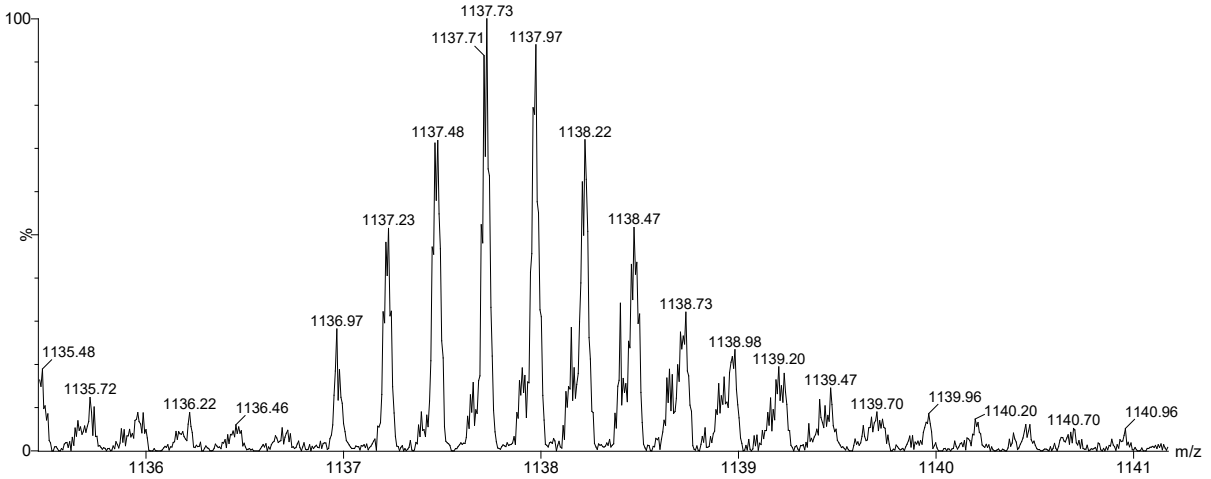


Figure S121. Partial mass spectrum of hexamer **1**$_6$ from the LC-MS analysis of a stirred library made from peptides **1** and **5** after 5 days (corresponding to **Figure S118**). All hexamers elute in a single peak in the LC-MS analysis.
1$_6$: Calculated m/z: 1137.01 $[M+4H]^{4+}$; observed m/z: 1136.97 $[M+4H]^{4+}$.

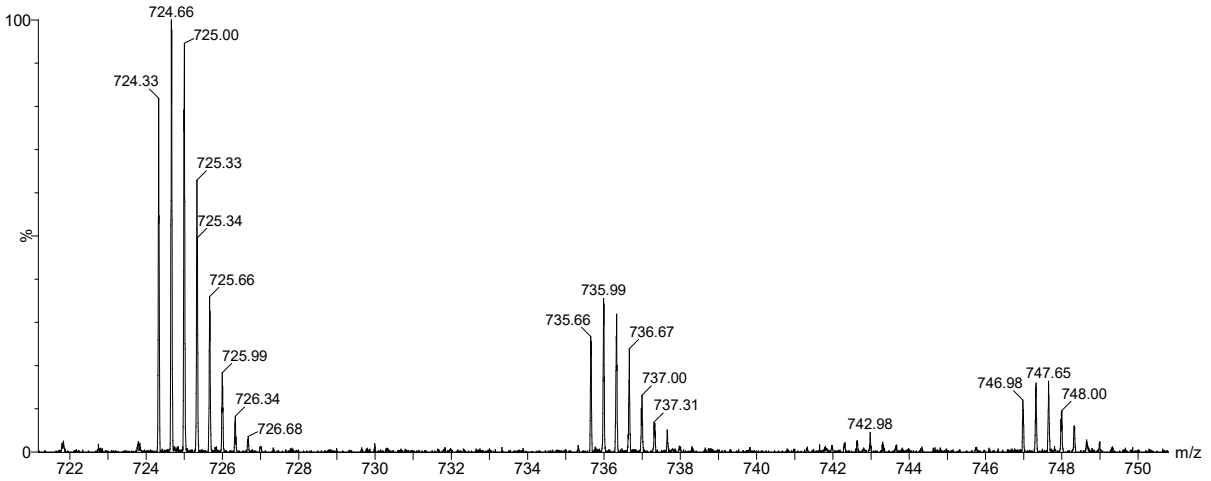


Figure S122. Mass spectrum of trimers **1**$_2$**5**$_1$, **1**$_1$**5**$_2$ and **5**$_3$ from the LC-MS analysis of a stirred library made from peptides **1** and **5** after 5 days (corresponding to **Figure S118**). These trimers elute in a single peak in the LC-MS analysis.
1$_2$**5**$_1$: Calculated m/z: 747.01 $[M+4H]^{4+}$; observed m/z: 746.98 $[M+3H]^{3+}$.
1$_1$**5**$_2$: Calculated m/z: 735.69 $[M+4H]^{4+}$; observed m/z: 735.66 $[M+3H]^{3+}$.
5$_3$: Calculated m/z: 724.36 $[M+3H]^{3+}$; observed m/z: 724.33 $[M+3H]^{3+}$.

X-GLKWK (9) + X-GKLKL (2)

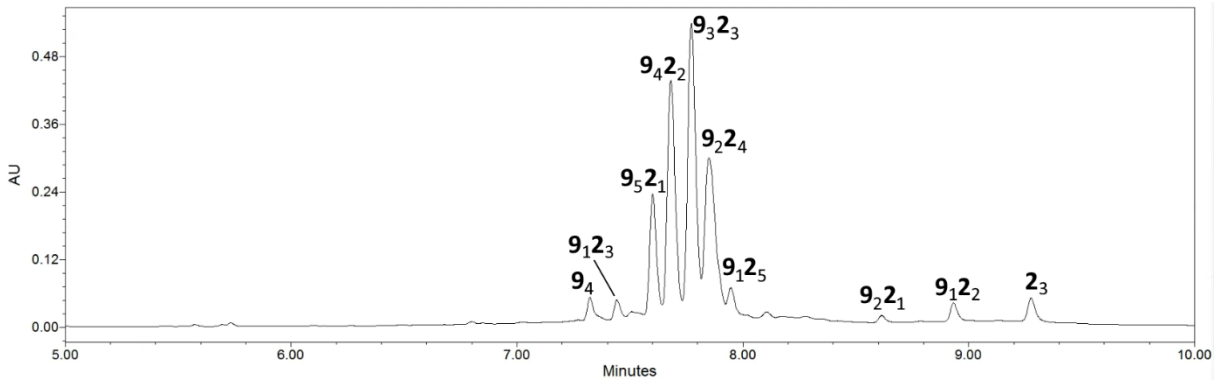


Figure S123. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **9** and **2** at 45 °C after 5 days (corresponding to **Figure S36**).

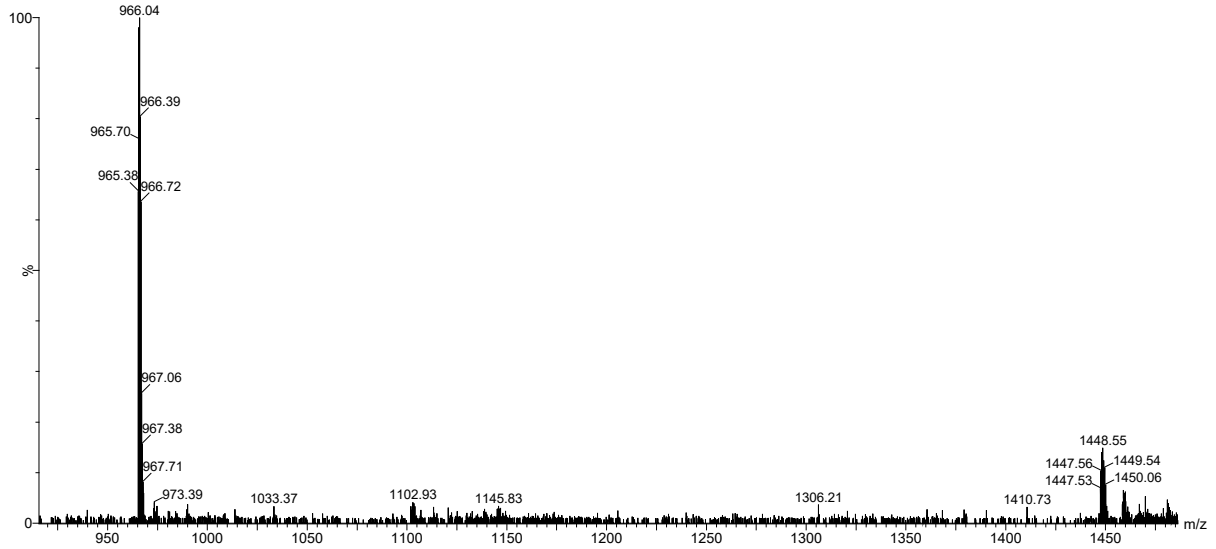


Figure S124. Mass spectrum of tetramer $\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 965.48 $[M+3H]^{3+}$, 1447.72 $[M+2H]^{2+}$; observed m/z: 965.38 $[M+3H]^{3+}$, 1447.53 $[M+2H]^{2+}$

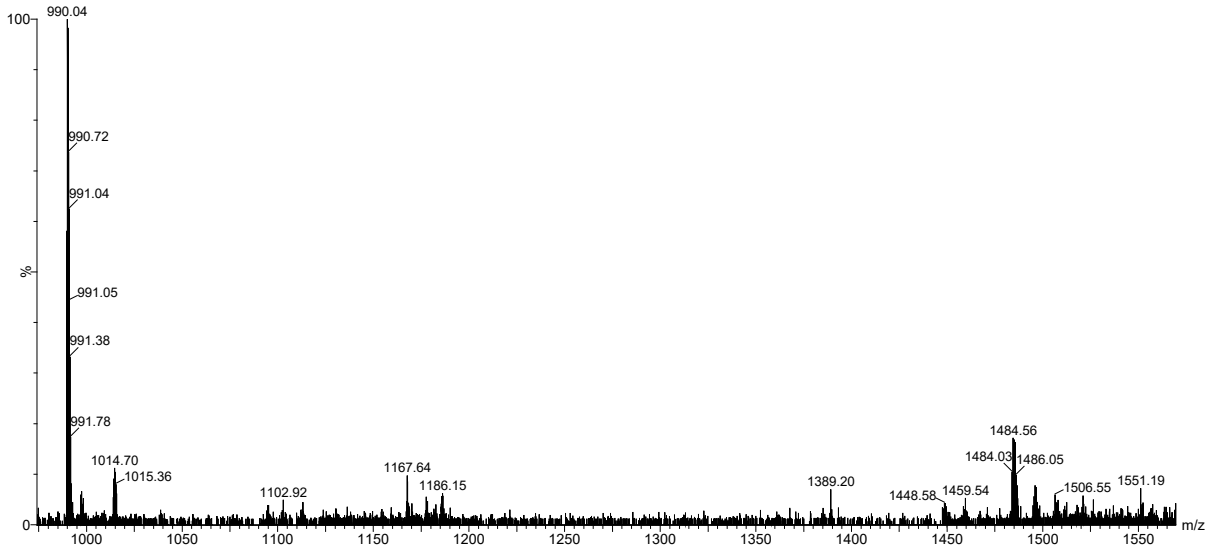


Figure S125. Mass spectrum of tetramer $\mathbf{9}_1\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 989.81 $[M+3H]^{3+}$, 1484.22 $[M+2H]^{2+}$; observed m/z: 990.04 $[M+3H]^{3+}$, 1484.03 $[M+2H]^{2+}$

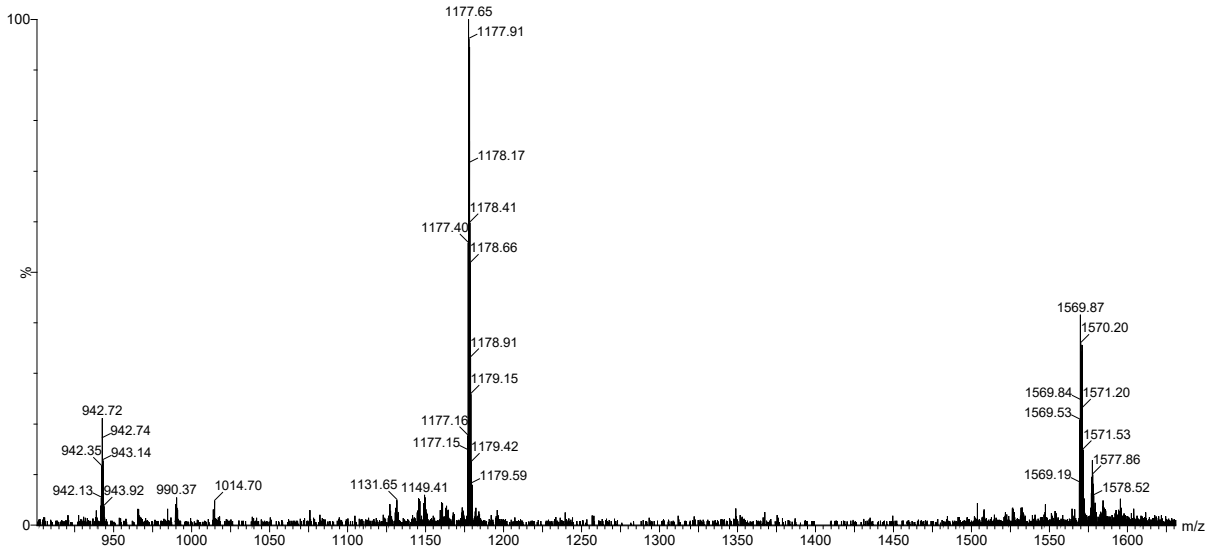


Figure S126. Mass spectrum of hexamer $\mathbf{9}_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 942.03 $[M+5H]^{5+}$, 1177.29 $[M+4H]^{4+}$, 1569.39 $[M+3H]^{3+}$; observed m/z: 942.13 $[M+5H]^{5+}$, 1177.15 $[M+4H]^{4+}$, 1569.19 $[M+3H]^{3+}$

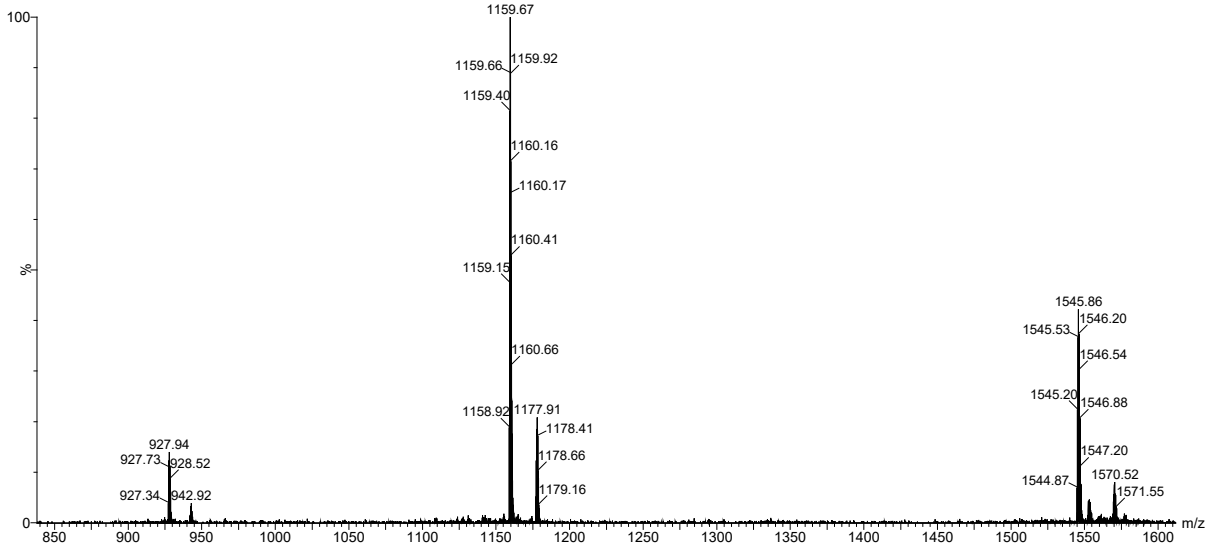


Figure S127. Mass spectrum of hexamer $\mathbf{9}_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 927.43 $[M+5H]^{5+}$, 1159.04 $[M+4H]^{4+}$, 1545.05 $[M+3H]^{3+}$; observed m/z: 927.34 $[M+5H]^{5+}$, 1158.92 $[M+4H]^{4+}$, 1544.87 $[M+3H]^{3+}$

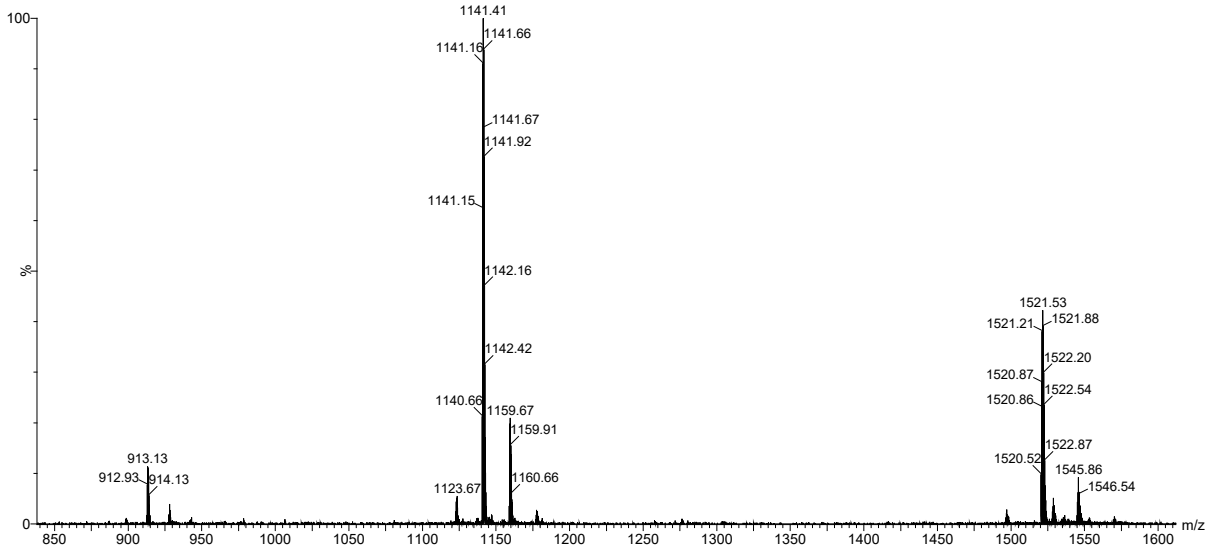


Figure S128. Mass spectrum of hexamer $\mathbf{9}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 912.83 $[M+5H]^{5+}$, 1140.79 $[M+4H]^{4+}$, 1520.72 $[M+3H]^{3+}$; observed m/z: 912.93 $[M+5H]^{5+}$, 1140.66 $[M+4H]^{4+}$, 1520.52 $[M+3H]^{3+}$

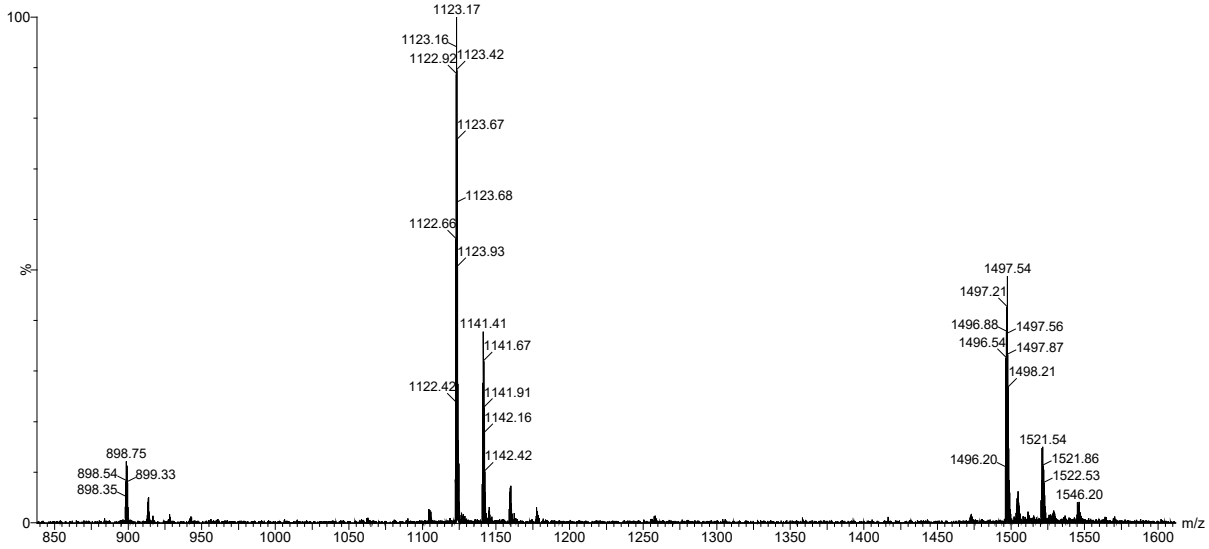


Figure S129. Mass spectrum of hexamer $\mathbf{9}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 898.23 $[M+5H]^{5+}$, 1122.54 $[M+4H]^{4+}$, 1496.39 $[M+3H]^{3+}$; observed m/z: 898.35 $[M+5H]^{5+}$, 1122.42 $[M+4H]^{4+}$, 1496.20 $[M+3H]^{3+}$

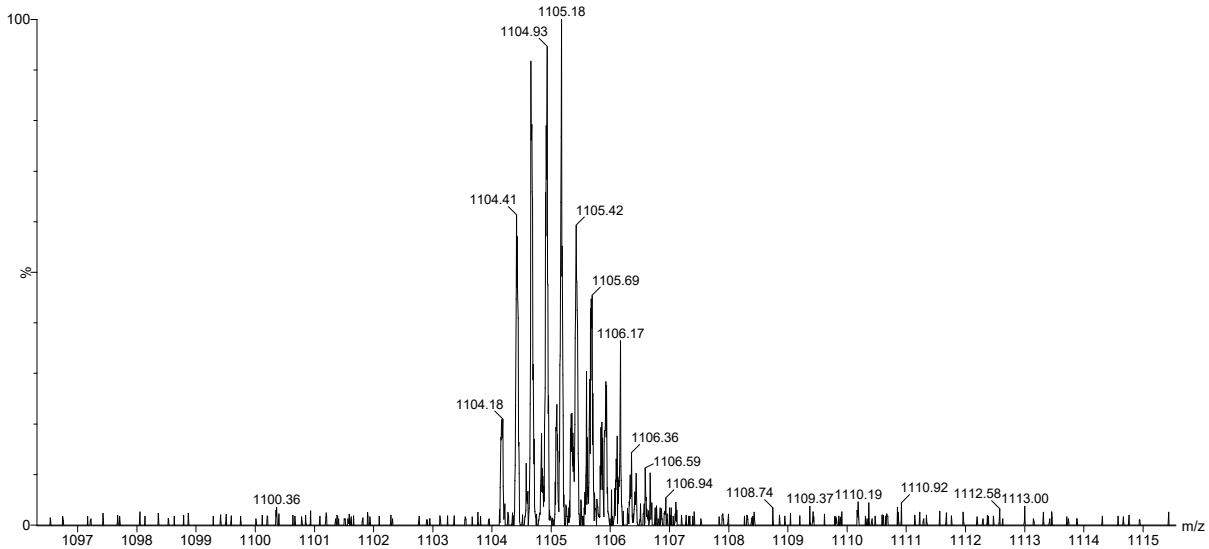


Figure S130. Mass spectrum of hexamer $\mathbf{9}_1\mathbf{2}_5$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 1104.29 $[M+4H]^{4+}$; observed m/z: 1104.18 $[M+4H]^{4+}$

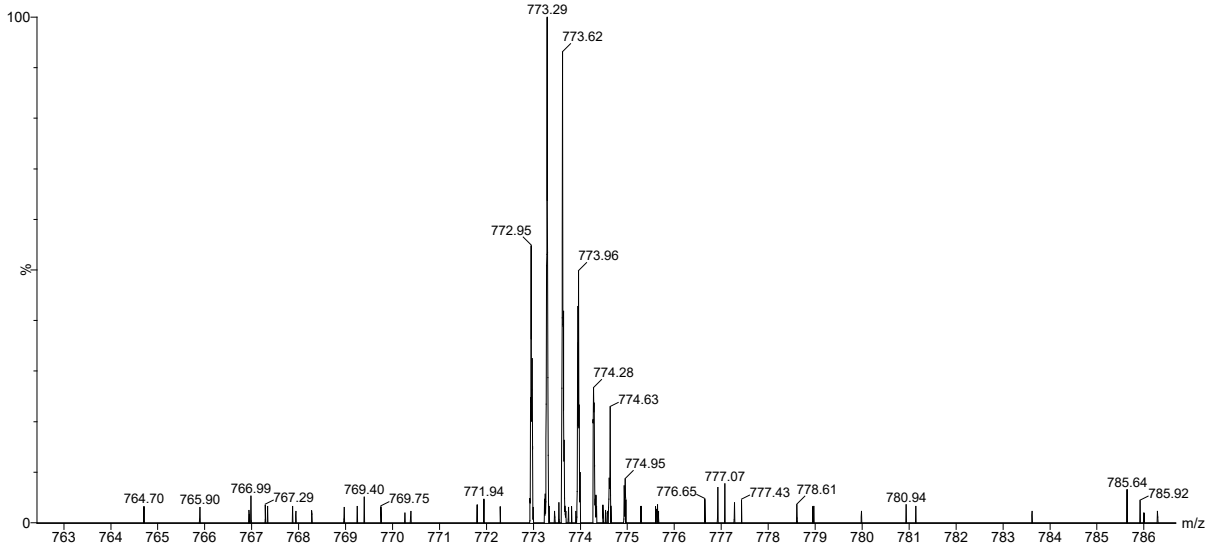


Figure S131. Mass spectrum of trimer $\mathbf{9}_2\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 773.03 $[M+3H]^{3+}$; observed m/z: 772.95 $[M+3H]^{3+}$

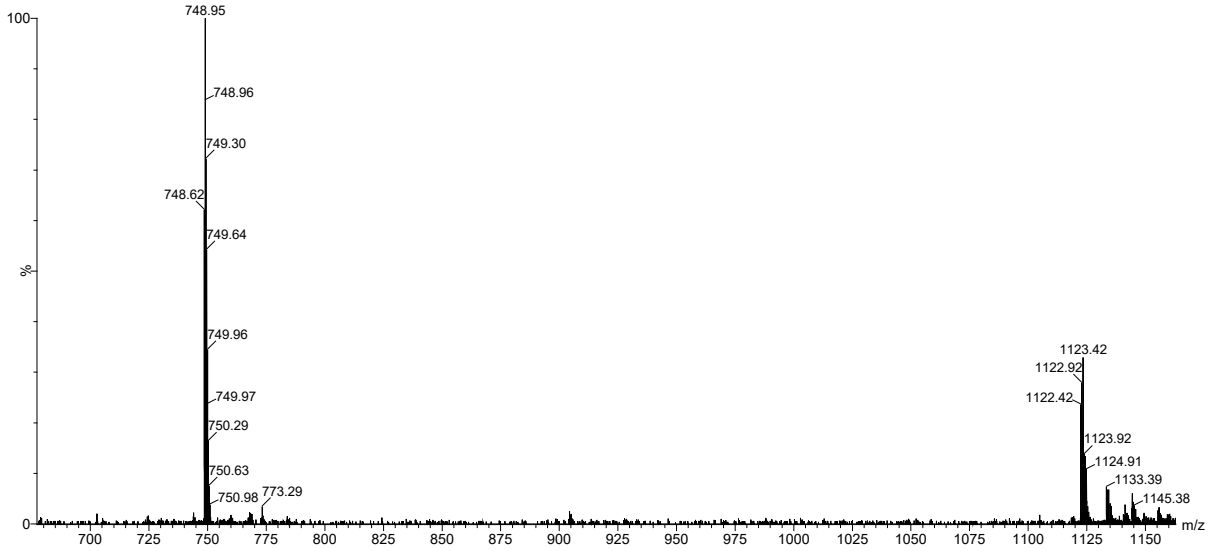


Figure S132. Mass spectrum of trimer $\mathbf{9}_1\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 748.69 $[M+3H]^{3+}$, 1122.54 $[M+2H]^{2+}$; observed m/z: 748.62 $[M+3H]^{3+}$, 1122.42 $[M+2H]^{2+}$

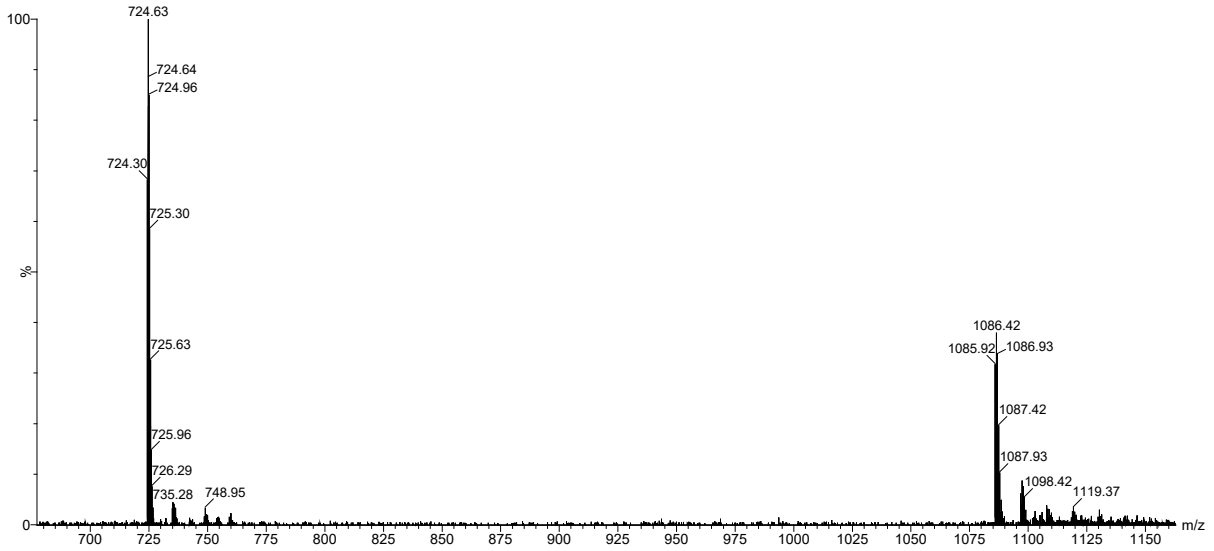


Figure S133. Mass spectrum of trimer $\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **9** and **2** after 5 days (corresponding to **Figure S123**). Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.30 $[M+3H]^{3+}$, 1085.92 $[M+2H]^{2+}$

X-GLKYK (10) + X-GKLKL (2)

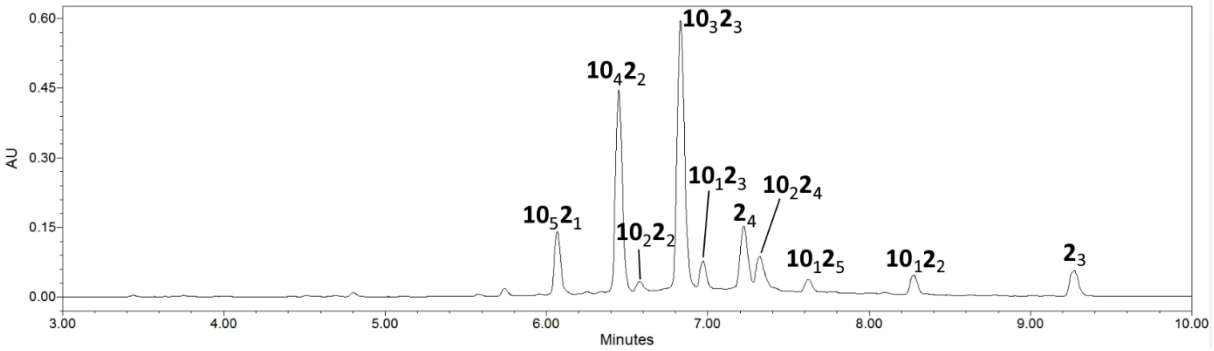


Figure S134. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **10** and **2** at 45 °C after 5 days (corresponding to Figure S30).

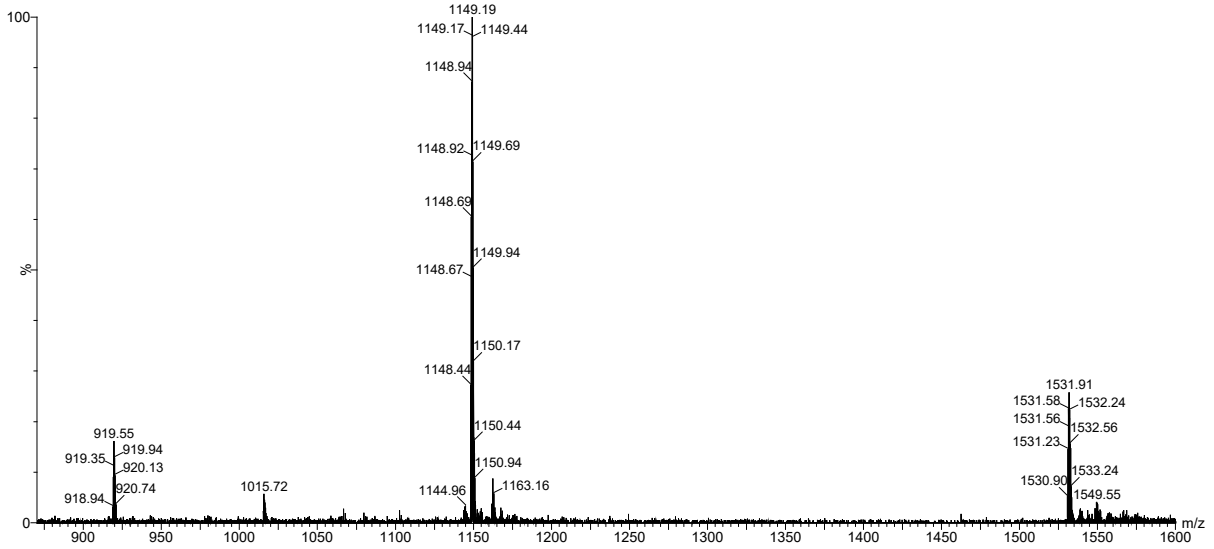


Figure S135. Mass spectrum of hexamer $\mathbf{10}_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). Calculated m/z: 919.01 $[M+5H]^{5+}$, 1148.52 $[M+4H]^{4+}$, 1531.02 $[M+3H]^{3+}$; observed m/z: 918.94 $[M+5H]^{5+}$, 1148.44 $[M+4H]^{4+}$, 1530.90 $[M+3H]^{3+}$

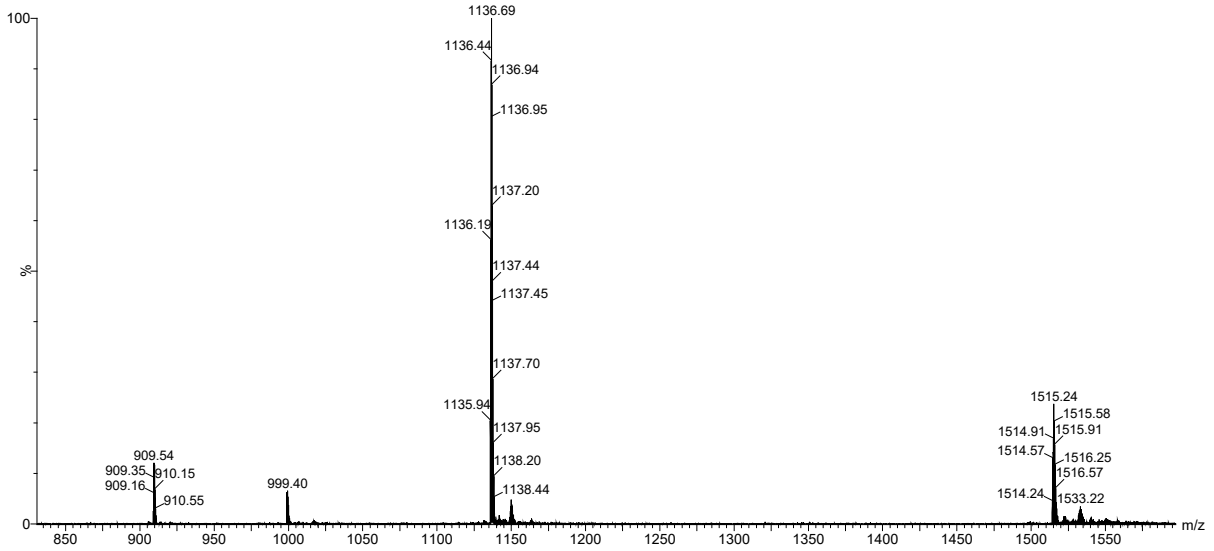


Figure S136. Mass spectrum of hexamer $\mathbf{10}_4\mathbf{2}_2$ and tetramer $\mathbf{10}_2\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). $\mathbf{10}_4\mathbf{2}_2$: calculated m/z: 909.02 $[M+5H]^{5+}$, 1136.02 $[M+4H]^{4+}$, 1514.36 $[M+3H]^{3+}$; observed m/z: 909.16 $[M+5H]^{5+}$, 1135.94 $[M+4H]^{4+}$, 1514.24 $[M+3H]^{3+}$
$\mathbf{10}_2\mathbf{2}_2$: calculated m/z: 999.13 $[M+1+3H]^{3+}$; observed m/z: 999.40 $[M+1+3H]^{3+}$

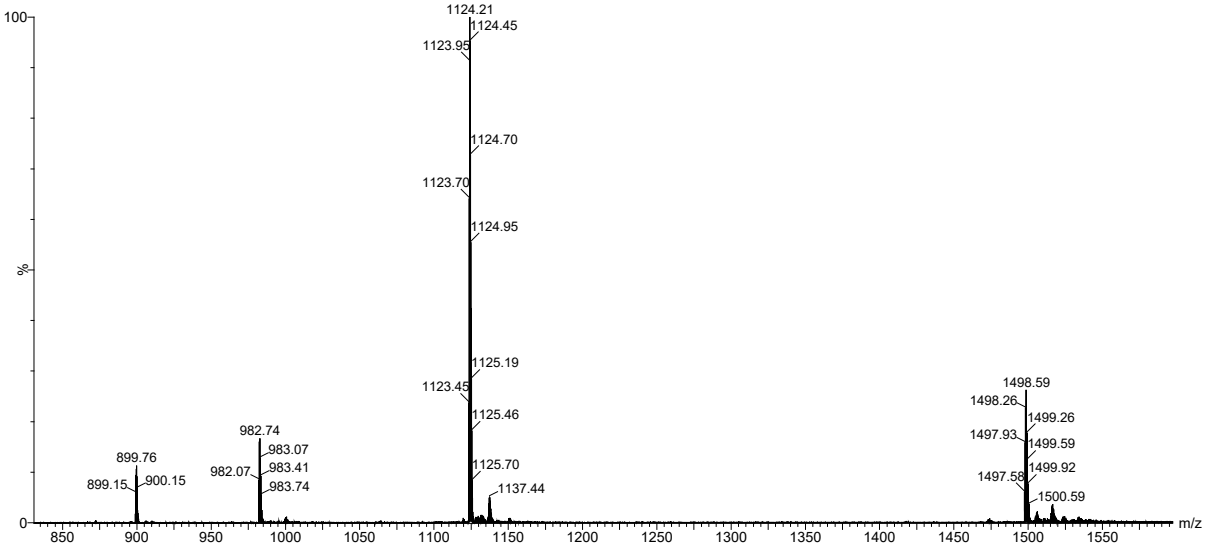


Figure S137. Mass spectrum of hexamer $\mathbf{10}_3\mathbf{2}_3$ and tetramer $\mathbf{10}_1\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). $\mathbf{10}_3\mathbf{2}_3$: calculated m/z: 899.02 $[M+5H]^{5+}$, 1123.53 $[M+4H]^{4+}$, 1497.70 $[M+3H]^{3+}$; observed m/z: 899.15 $[M+5H]^{5+}$, 1123.45 $[M+4H]^{4+}$, 1497.58 $[M+3H]^{3+}$ $\mathbf{10}_1\mathbf{2}_3$: calculated m/z: 982.14 $[M+3H]^{3+}$; observed m/z: 982.07 $[M+3H]^{3+}$

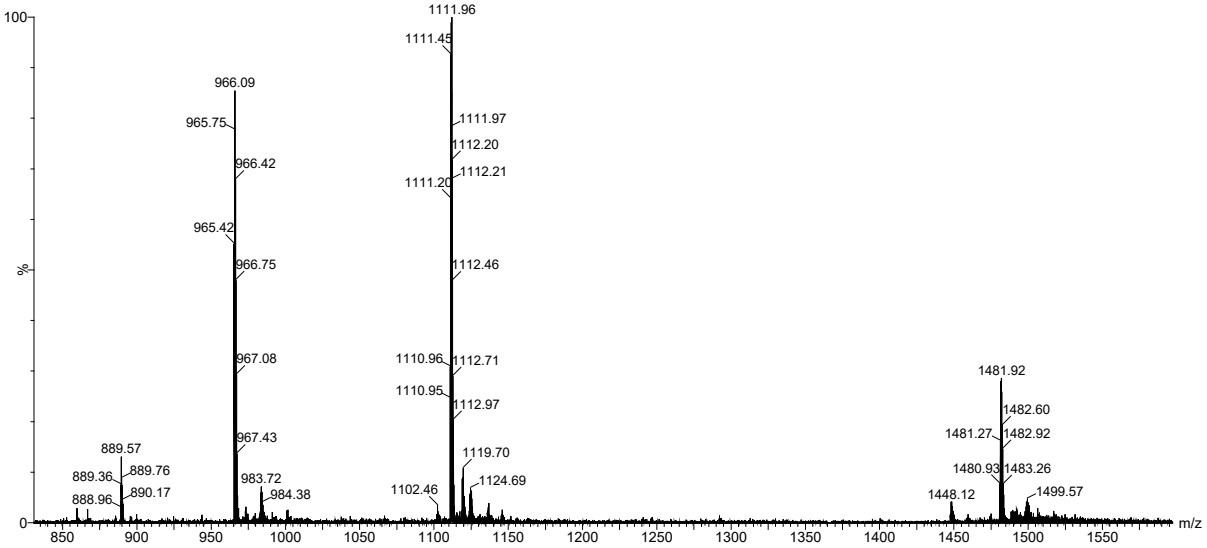


Figure S138. Mass spectrum of hexamer $\mathbf{10}_2\mathbf{2}_4$ and tetramer $\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). $\mathbf{10}_2\mathbf{2}_4$: Calculated m/z: 889.02 $[M+5H]^{5+}$, 1111.03 $[M+4H]^{4+}$, 1481.04 $[M+3H]^{3+}$; observed m/z: 888.96 $[M+5H]^{5+}$, 1110.95 $[M+4H]^{4+}$, 1480.93 $[M+3H]^{3+}$ $\mathbf{2}_4$: Calculated m/z: 965.48 $[M+3H]^{3+}$; observed m/z: 965.42 $[M+3H]^{3+}$

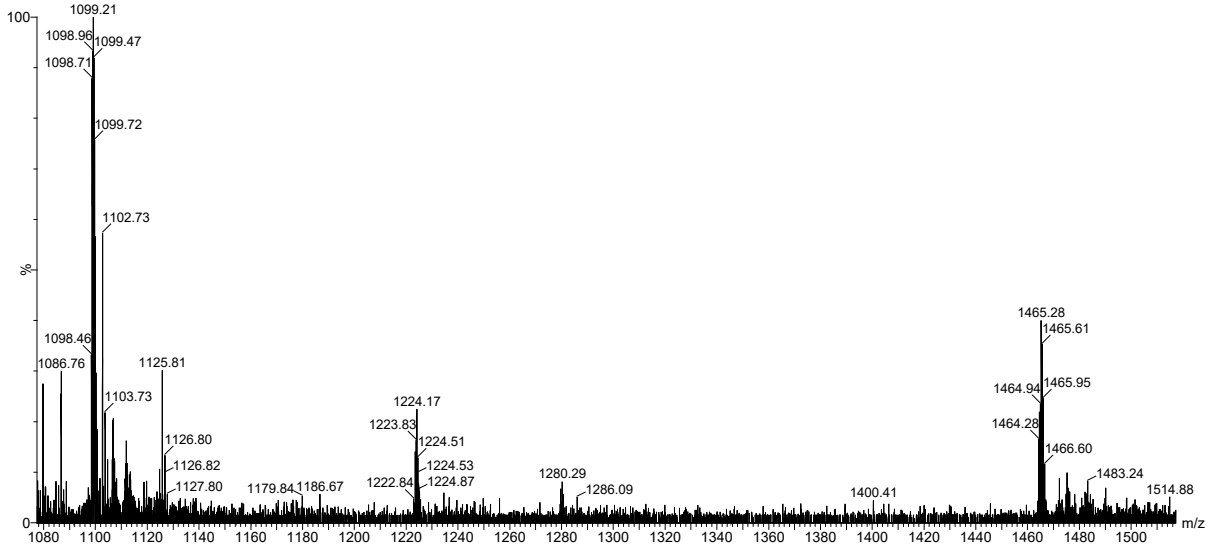


Figure S139. Mass spectrum of hexamer $\mathbf{10_1 2_5}$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). Calculated m/z: 1098.54 $[M+4H]^{4+}$, 1464.38 $[M+3H]^{3+}$; observed m/z: 1098.46 $[M+4H]^{4+}$, 1464.28 $[M+3H]^{3+}$

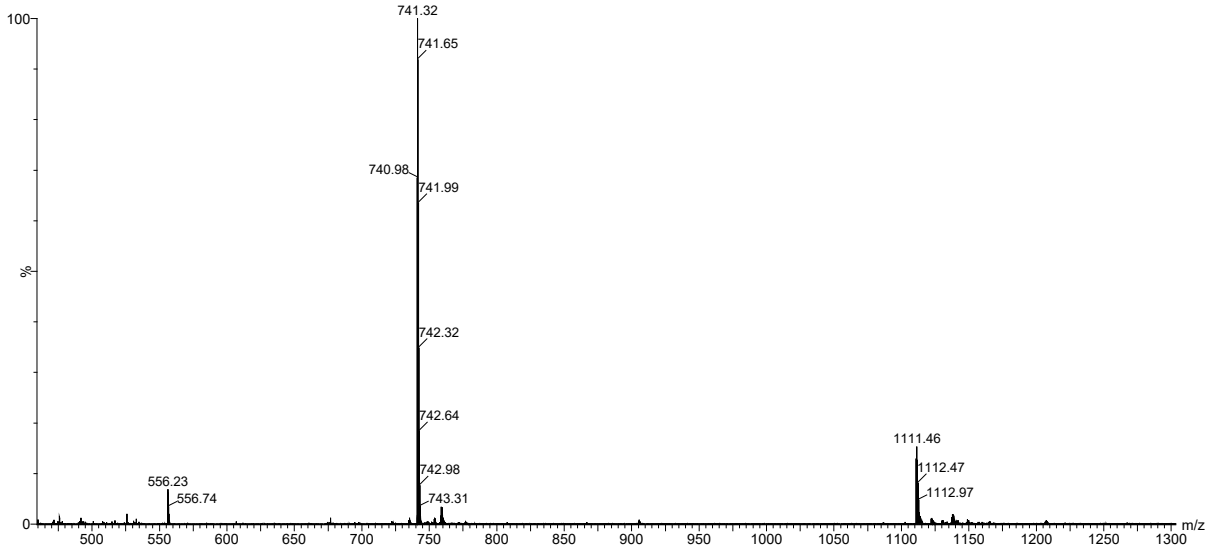


Figure S140. Mass spectrum of trimer $\mathbf{10_1 2_2}$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). Calculated m/z: 556.27 $[M+1+4H]^{4+}$, 741.02 $[M+3H]^{3+}$, 1111.53 $[M+1+2H]^{2+}$; observed m/z: 556.23 $[M+1+4H]^{4+}$, 740.98 $[M+3H]^{3+}$, 1111.46 $[M+1+2H]^{2+}$

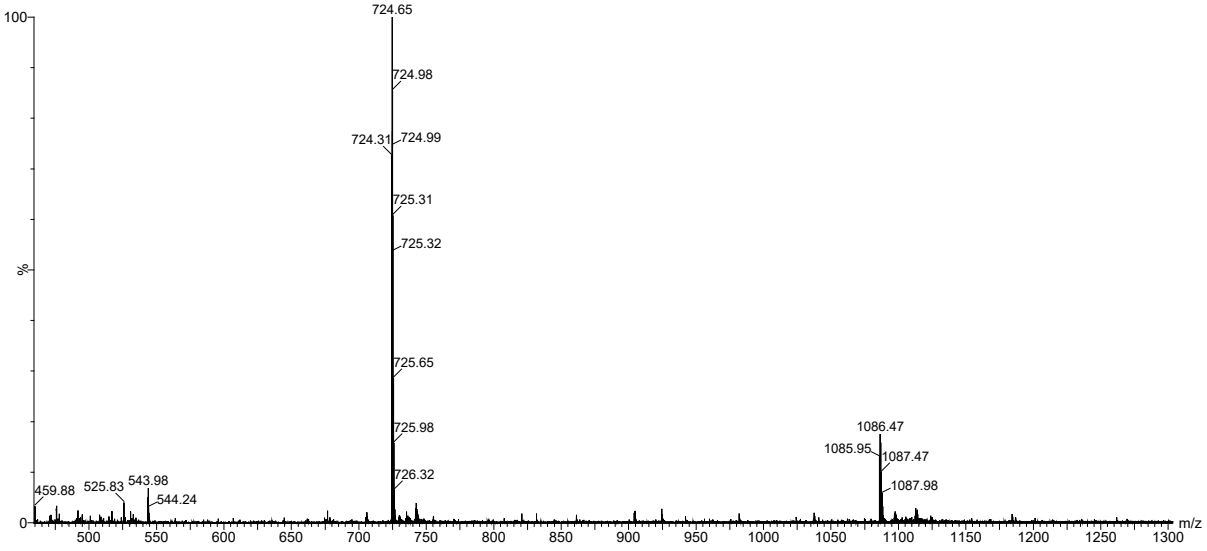


Figure S 141. Mass spectrum of trimer $\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **10** and **2** after 5 days (corresponding to **Figure S134**). Calculated m/z: 544.02 $[M+1+4H]^{4+}$, 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 543.98 $[M+1+4H]^{4+}$, 724.31 $[M+3H]^{3+}$, 1085.95 $[M+1+2H]^{2+}$

X-GLKFO (11) + X-GKLKL (2)

DCLs containing building block **11** show 3 different distributions of mixed hexamers. The second distribution in size (elution times between 8.0 and 8.5 minutes) has the same relative abundances as the first distribution in size (elution times between 7.4 and 8.9 minutes). In the second distribution all macrocycles contain one building block **11** in which the C-terminal ornithine has cyclized to form a six membered cyclic amide, losing one molecule of H_2O in the process. The third distribution (in which only $\mathbf{11}_3\mathbf{2}_3$ could be detected; elution time between 8.9 and 9.0 minutes) corresponds to hexamers in which two building blocks **11** have cyclized ornithines. The distribution shown in **Figure S39** results from the sum of integration of all three distributions.

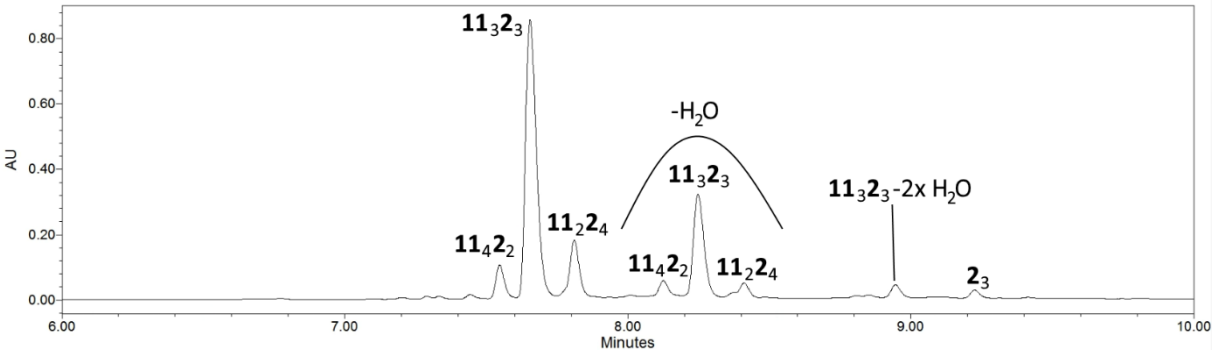


Figure S142. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **11** and **2** at 45 °C after 7 days (corresponding to **Figure S39**).

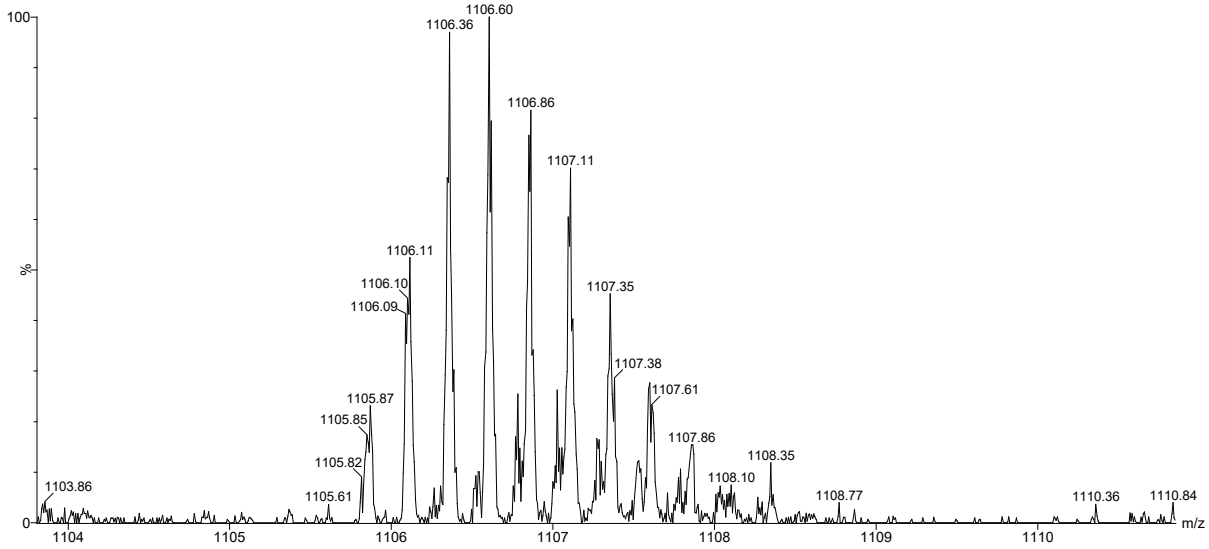


Figure S143. Mass spectrum of hexamer $\mathbf{11}_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 7 days (corresponding to **Figure S142**). Calculated m/z: 1106.01 $[M+4H]^{4+}$; observed m/z: 1105.82 $[M+4H]^{4+}$

Figure S144. Mass spectrum of hexamer $\mathbf{11}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 7 days (corresponding to **Figure S142**). Calculated m/z: 1101.02 $[M+4H]^{4+}$, 1467.69 $[M+3H]^{3+}$; observed m/z: 1101.11 $[M+4H]^{4+}$, 1467.48 $[M+3H]^{3+}$

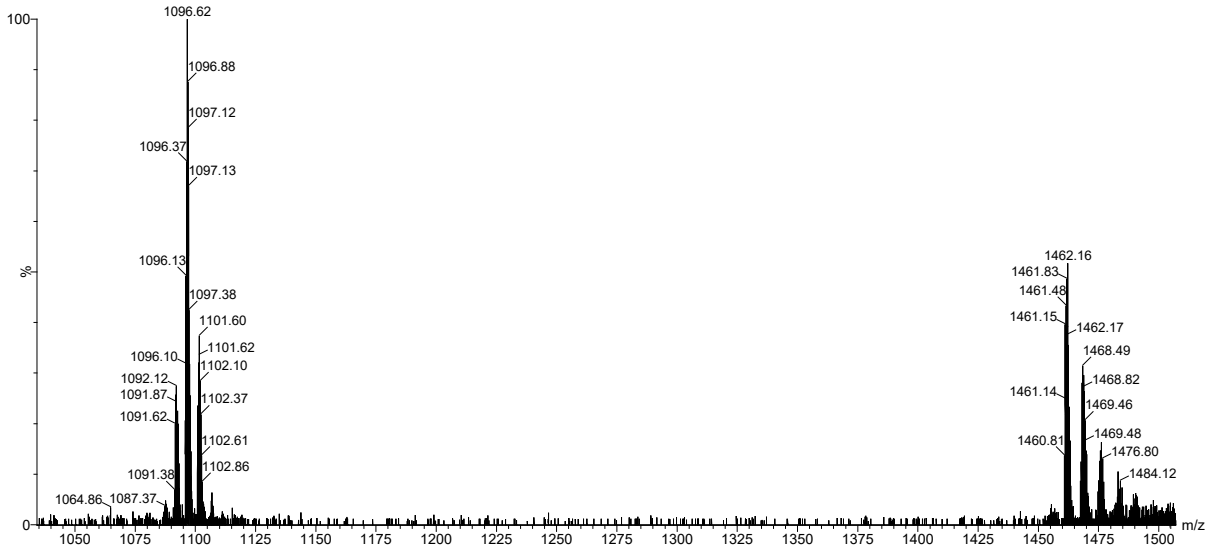


Figure S145. Mass spectrum of hexamer $\mathbf{11}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 7 days (corresponding to **Figure S142**). Calculated m/z: 1096.03 $[M+4H]^{4+}$, 1461.03 $[M+3H]^{3+}$; observed m/z: 1096.10 $[M+4H]^{4+}$, 1460.81 $[M+3H]^{3+}$

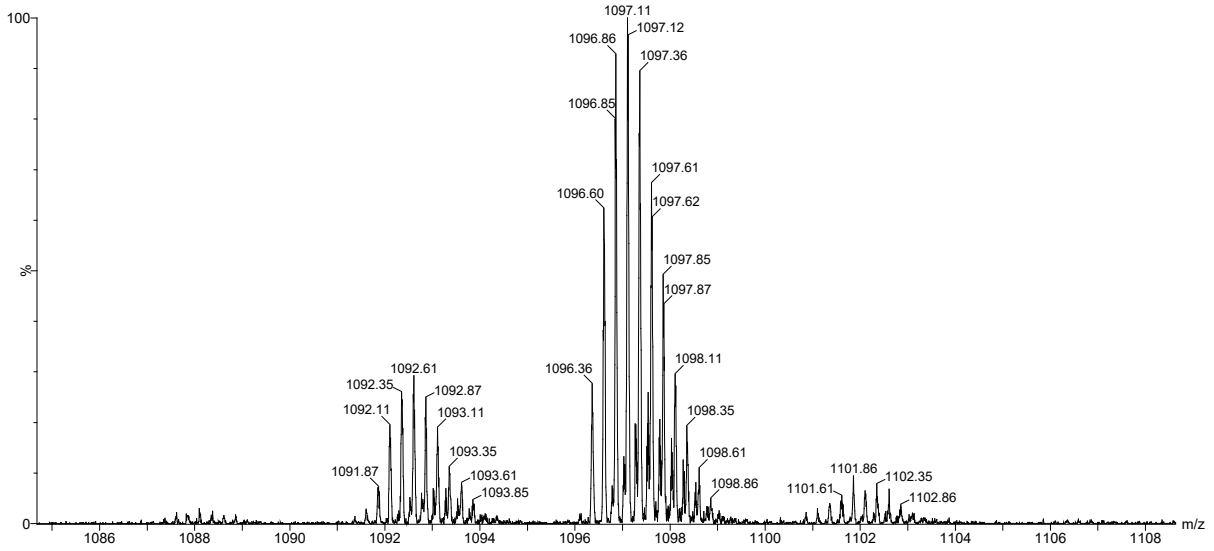


Figure S146. Mass spectrum of hexamer $\mathbf{11}_4\mathbf{2}_2$-$H_2O$, $\mathbf{11}_3\mathbf{2}_3$-$H_2O$ and $\mathbf{11}_2\mathbf{2}_4$-$H_2O$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 7 days (corresponding to **Figure S142**). $\mathbf{11}_4\mathbf{2}_2$-$H_2O$: Calculated m/z: 1101.50 $[M+4H]^{4+}$; observed m/z: 1101.61 $[M+4H]^{4+}$ $\mathbf{11}_3\mathbf{2}_3$-$H_2O$: Calculated m/z: 1096.51 $[M+4H]^{4+}$; observed m/z: 1096.36 $[M+4H]^{4+}$ $\mathbf{11}_2\mathbf{2}_4$-$H_2O$: Calculated m/z: 1091.52 $[M+4H]^{4+}$; observed m/z: 1091.87 $[M+4H]^{4+}$

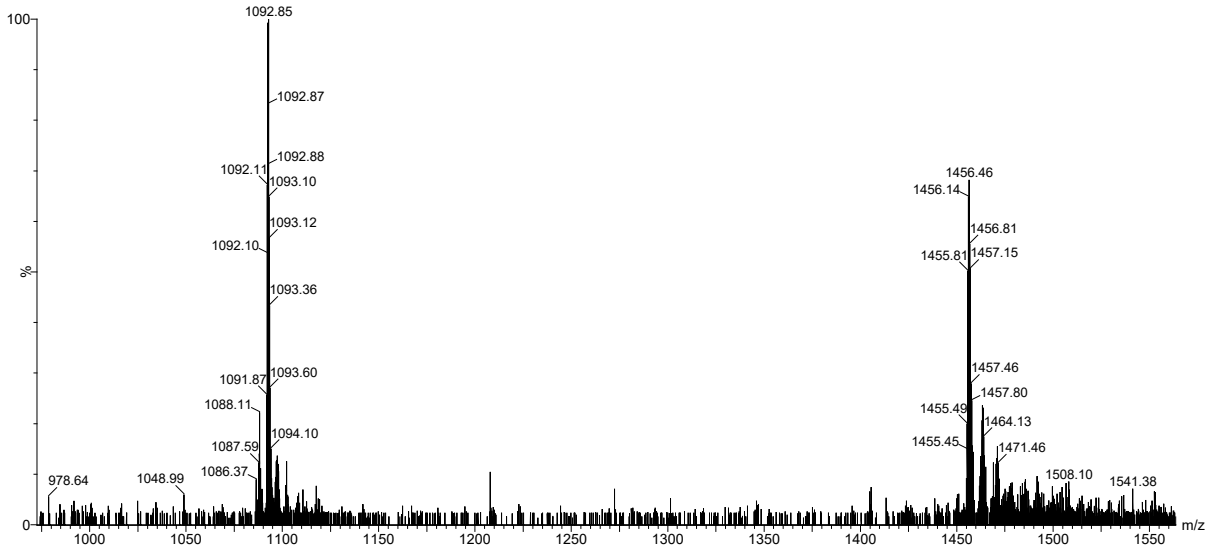


Figure S147. Mass spectrum of hexamer $\mathbf{11}_3\mathbf{2}_3$-$2H_2O$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 7 days (corresponding to **Figure S142**). Calculated m/z: 1092.01 $[M+4H]^{4+}$, 1455.68 $[M+3H]^{3+}$; observed m/z: 1091.87 $[M+4H]^{4+}$, 1455.45 $[M+3H]^{3+}$

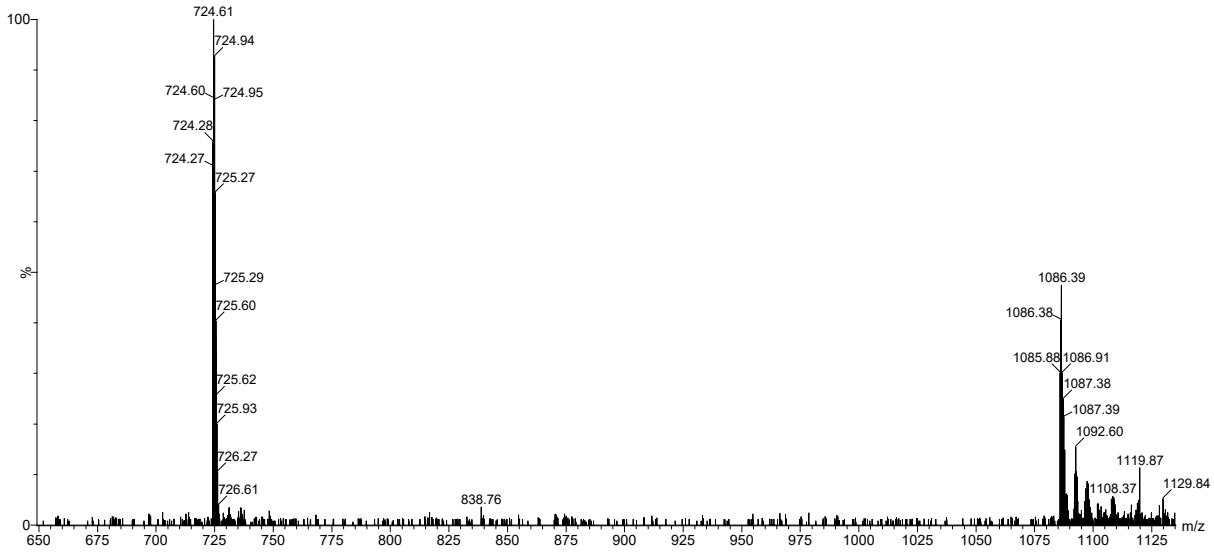


Figure S148. Mass spectrum of trimer $\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **11** and **2** after 5 days (corresponding to **Figure S142**). Calculated m/z: 724.36 $[M+3H]^{3+}$, 1086.04 $[M+2H]^{2+}$; observed m/z: 724.27 $[M+3H]^{3+}$, 1085.88 $[M+2H]^{2+}$

X-GLRFK (12) + X-GKLKL (2)

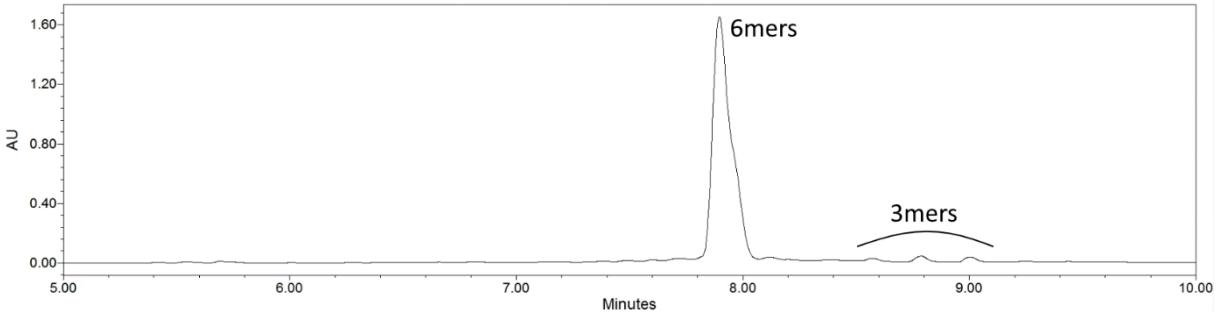


Figure S149. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **12** and **2** at 45 °C after 5 days (corresponding to **Figure S40**).

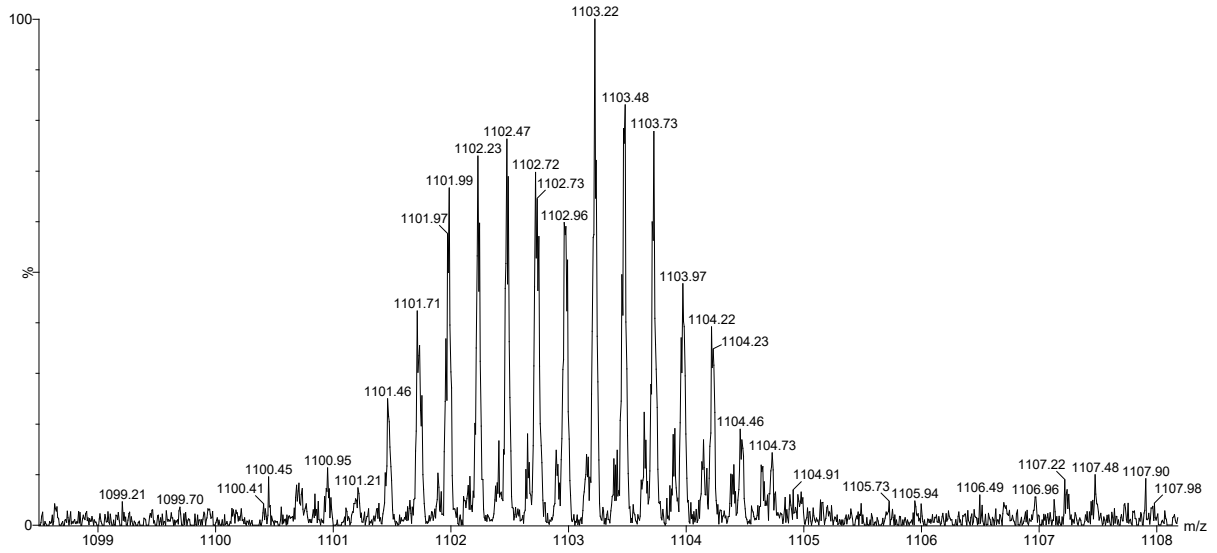


Figure S150. Mass spectrum of hexamer $\mathbf{12}_1\mathbf{2}_5$ from the LC-MS analysis of a stirred library made from peptides **12** and **2** after 5 days (corresponding to **Figure S149**). $\mathbf{12}_1\mathbf{2}_5$: Calculated m/z: 1101.54 $[M+4H]^{4+}$; observed m/z: 1101.46 $[M+4H]^{4+}$

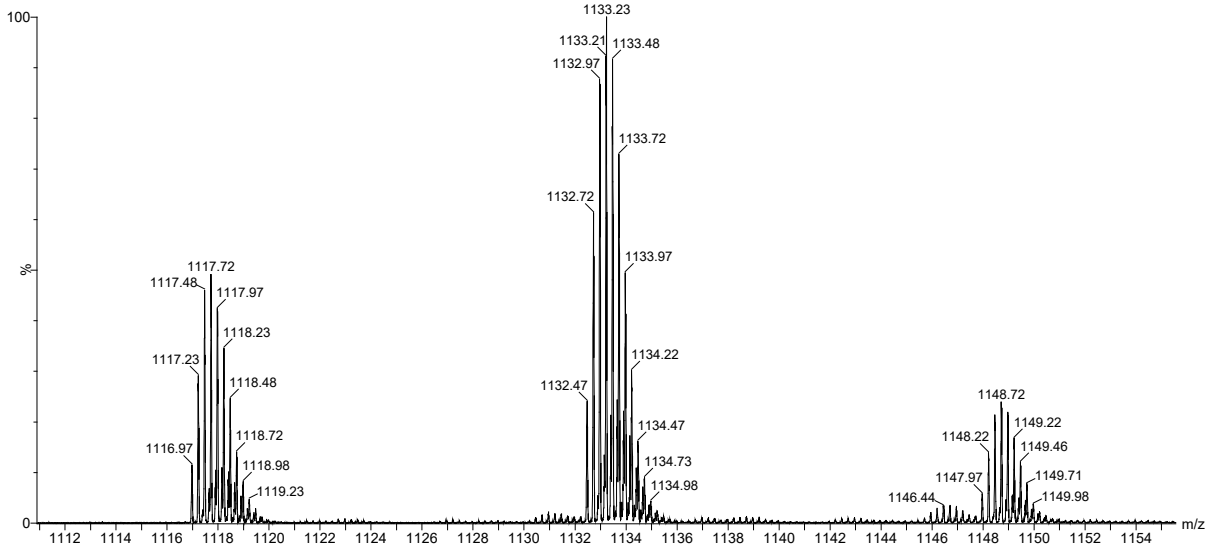


Figure S151. Mass spectrum of hexamers $\mathbf{12}_4\mathbf{2}_2$, $\mathbf{12}_3\mathbf{2}_3$ and $\mathbf{12}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **12** and **2** after 5 days (corresponding to **Figure S149**). $\mathbf{12}_4\mathbf{2}_2$: Calculated m/z: 1148.03 $[M+4H]^{4+}$; observed m/z: 1147.97 $[M+4H]^{4+}$ $\mathbf{12}_3\mathbf{2}_3$: Calculated m/z: 1132.53 $[M+4H]^{4+}$; observed m/z: 1132.47 $[M+4H]^{4+}$ $\mathbf{12}_2\mathbf{2}_4$: Calculated m/z: 1117.04 $[M+4H]^{4+}$; observed m/z: 1116.97 $[M+4H]^{4+}$

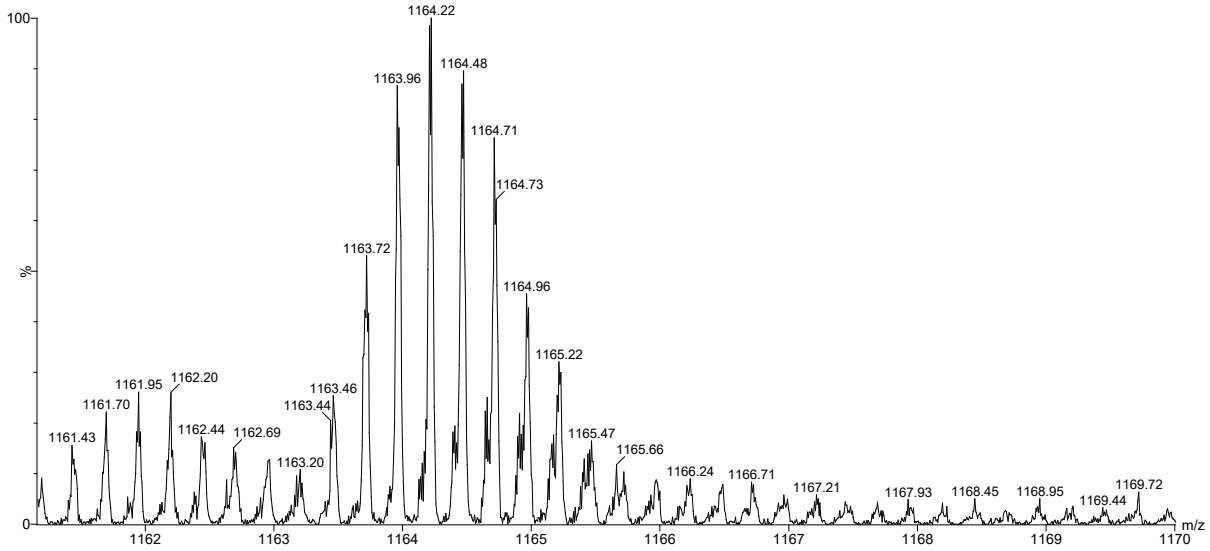


Figure S152. Mass spectrum of hexamer $\mathbf{12}_5\mathbf{2}_1$ from the LC-MS analysis of a stirred library made from peptides **12** and **2** after 5 days (corresponding to **Figure S149**). $\mathbf{12}_5\mathbf{2}_1$: Calculated m/z: 1163.53 $[M+4H]^{4+}$; observed m/z: 1163.46 $[M+4H]^{4+}$

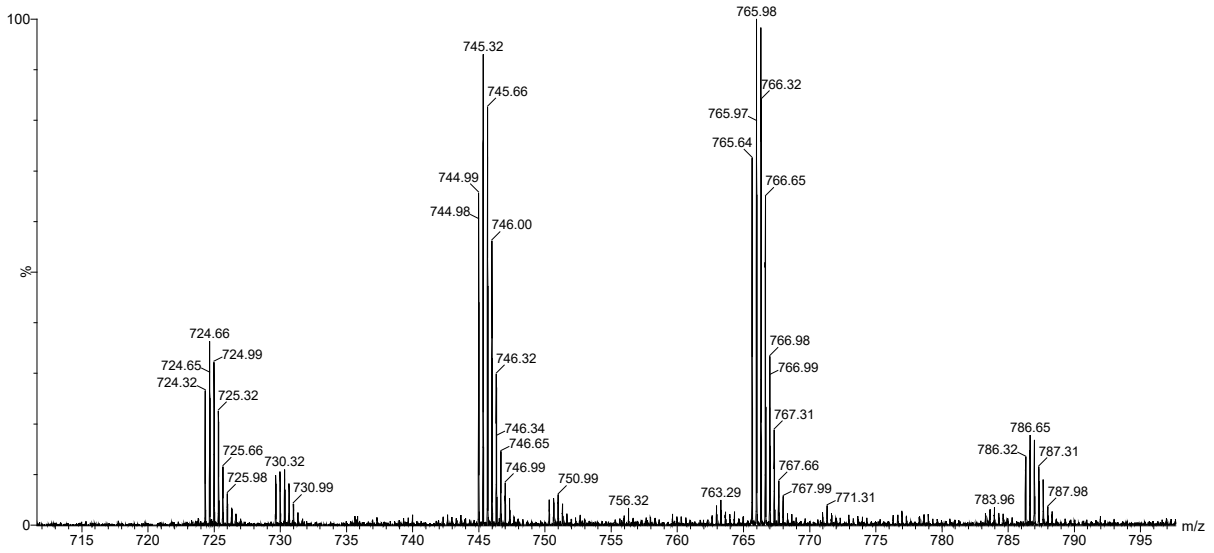


Figure S153. Mass spectrum of trimers $\mathbf{12}_3$, $\mathbf{12}_2\mathbf{2}_1$, $\mathbf{12}_1\mathbf{2}_2$ and $\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **12** and **2** after 5 days (corresponding to **Figure S149**). $\mathbf{12}_3$: Calculated m/z: 786.35 $[M+3H]^{3+}$; observed m/z: 786.32 $[M+3H]^{3+}$ $\mathbf{12}_2\mathbf{2}_1$: Calculated m/z: 765.69 $[M+3H]^{3+}$; observed m/z: 765.64 $[M+3H]^{3+}$ $\mathbf{12}_1\mathbf{2}_2$: Calculated m/z: 745.02 $[M+3H]^{3+}$; observed m/z: 744.98 $[M+3H]^{3+}$ $\mathbf{2}_3$: Calculated m/z: 724.36 $[M+3H]^{3+}$; observed m/z: 724.32 $[M+3H]^{3+}$

X-GLKFR (13) + X-GKLKL (2)

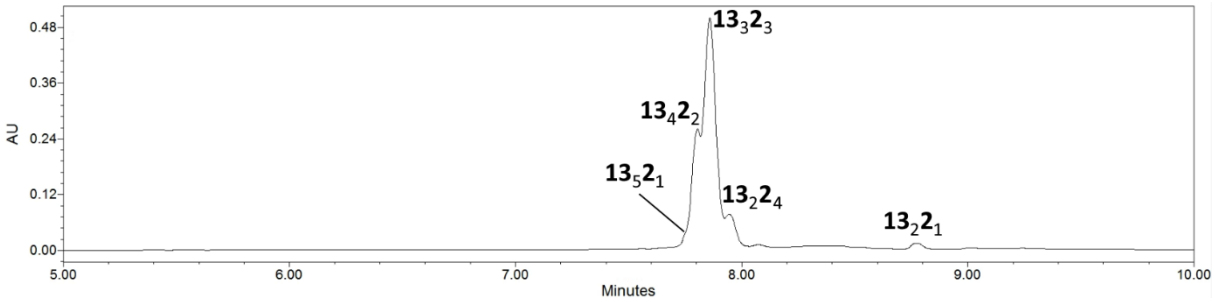


Figure S154. RP-UPLC trace (monitored at 254 nm) of the product mixture obtained from mixture of peptides **13** and **2** at 45 °C after 5 days (corresponding to **Figure S41**).

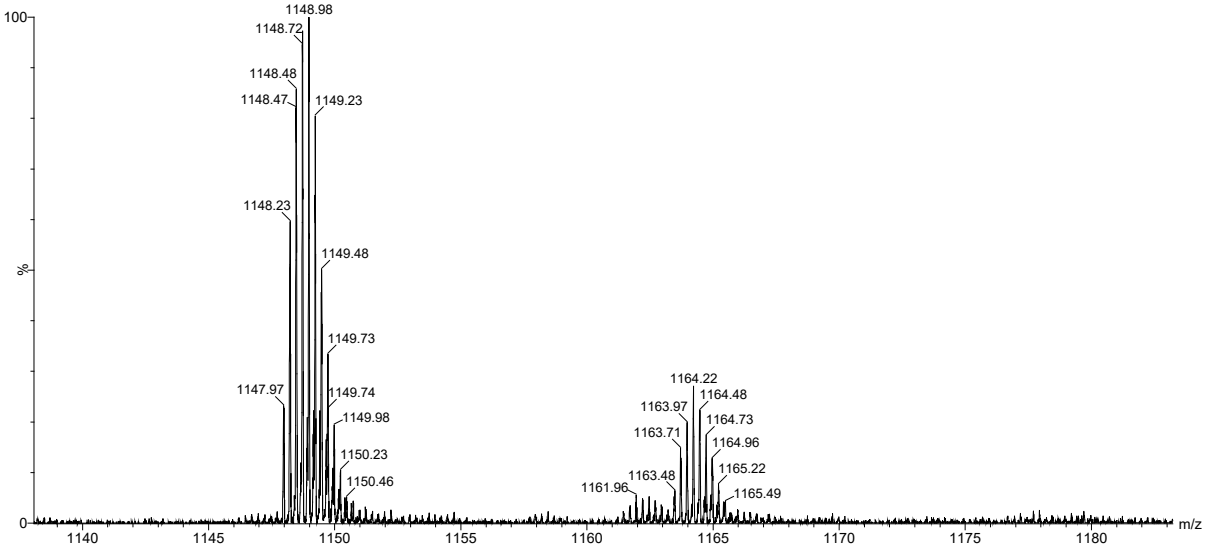


Figure S155. Mass spectrum of hexamers $\mathbf{13}_5\mathbf{2}_1$ and $\mathbf{13}_4\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **13** and **2** after 5 days (corresponding to **Figure S154**). $\mathbf{13}_5\mathbf{2}_1$: Calculated m/z: 1163.53 $[M+4H]^{4+}$; observed m/z: 1163.48 $[M+4H]^{4+}$ $\mathbf{13}_4\mathbf{2}_2$: Calculated m/z: 1148.03 $[M+4H]^{4+}$; observed m/z: 1147.97 $[M+4H]^{4+}$

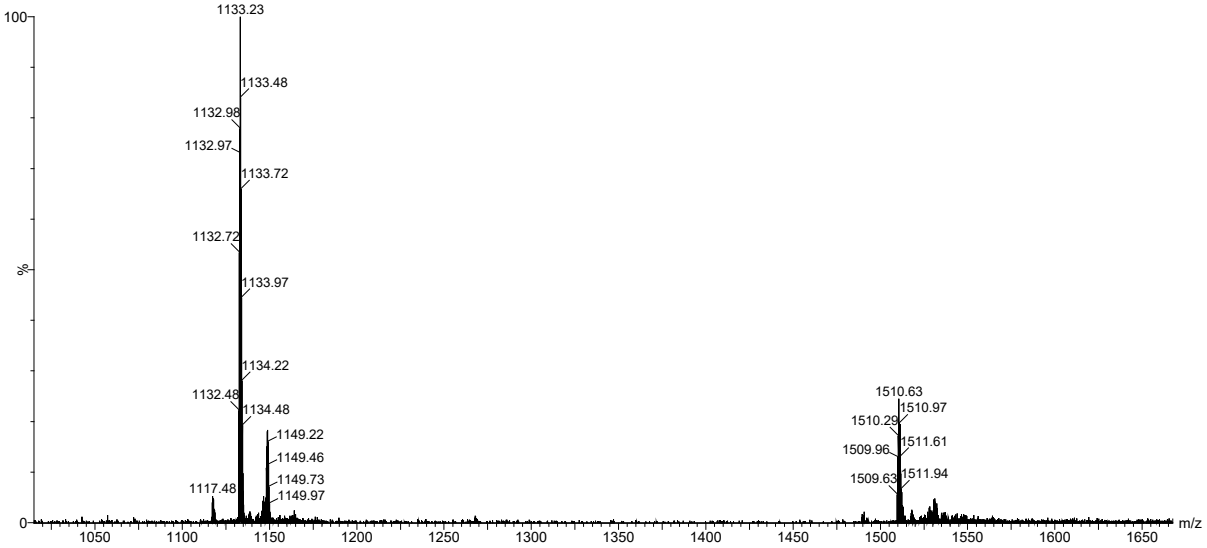


Figure S156. Mass spectrum of hexamer $\mathbf{13}_3\mathbf{2}_3$ from the LC-MS analysis of a stirred library made from peptides **13** and **2** after 5 days (corresponding to **Figure S154**). Calculated m/z: 1132.53 $[M+4H]^{4+}$, 1509.71 $[M+3H]^{3+}$; observed m/z: 1132.48 $[M+4H]^{4+}$, 1509.63 $[M+3H]^{3+}$

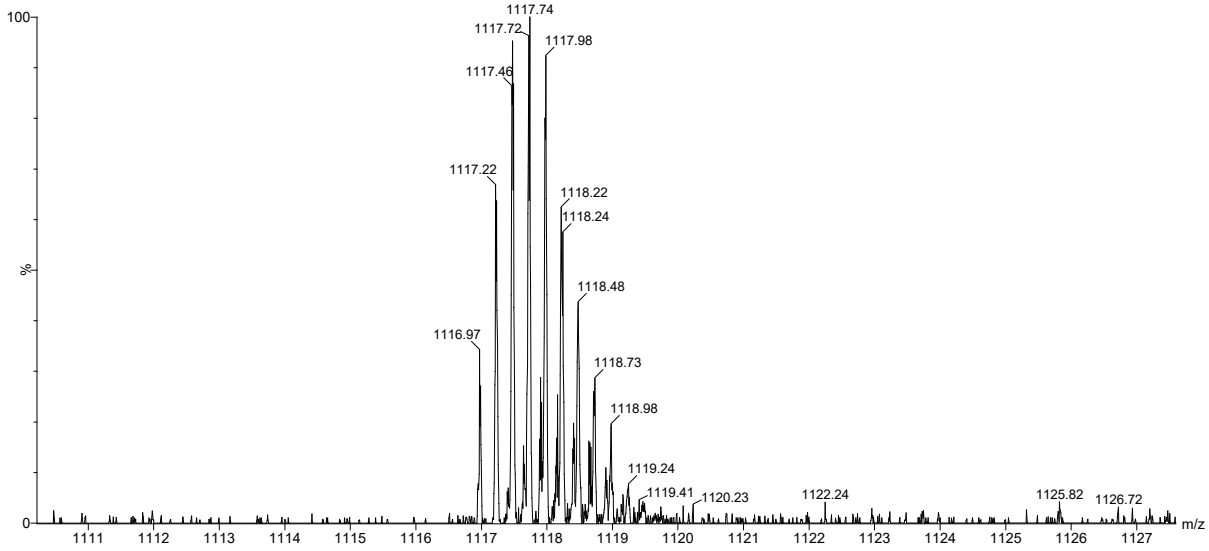


Figure S157. Mass spectrum of hexamer $\mathbf{13}_2\mathbf{2}_4$ from the LC-MS analysis of a stirred library made from peptides **13** and **2** after 5 days (corresponding to **Figure S154**). Calculated m/z: 1117.04 $[M+4H]^{4+}$; observed m/z: 1116.97 $[M+4H]^{4+}$

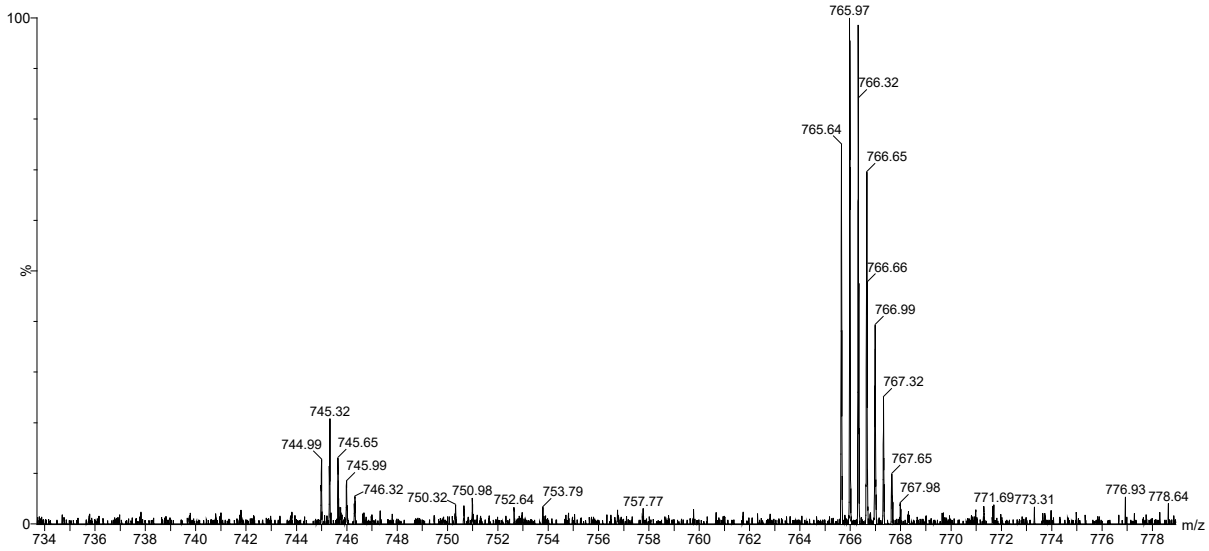


Figure S158. Mass spectrum of trimers $\mathbf{13}_2\mathbf{2}_1$ and $\mathbf{13}_1\mathbf{2}_2$ from the LC-MS analysis of a stirred library made from peptides **13** and **2** after 5 days (corresponding to **Figure S154**). $\mathbf{13}_2\mathbf{2}_1$: Calculated m/z: 765.69 $[M+3H]^{3+}$; observed m/z: 765.64 $[M+3H]^{3+}$ $\mathbf{13}_1\mathbf{2}_2$: Calculated m/z: 745.02 $[M+3H]^{3+}$; observed m/z: 744.99 $[M+3H]^{3+}$

S23. Mixed-building block replicators yielding statistical distributions

Libraries were prepared for a selection of building block combinations (100 µM per building block; 200 µM total). Library progression was monitored by UPLC at day zero, day four and day 14. LC-MS measurements were performed at day nine. All experiments were performed in borate buffer (100 mM B_2O_3, pH 8.12), were kept at 40 °C and stirred at 1200 rpm. Mixed hexamers were found to be the dominant products in all libraries. For building block structures, see Section S1.

X-GLKFK (1) + X-GLOFK (14)

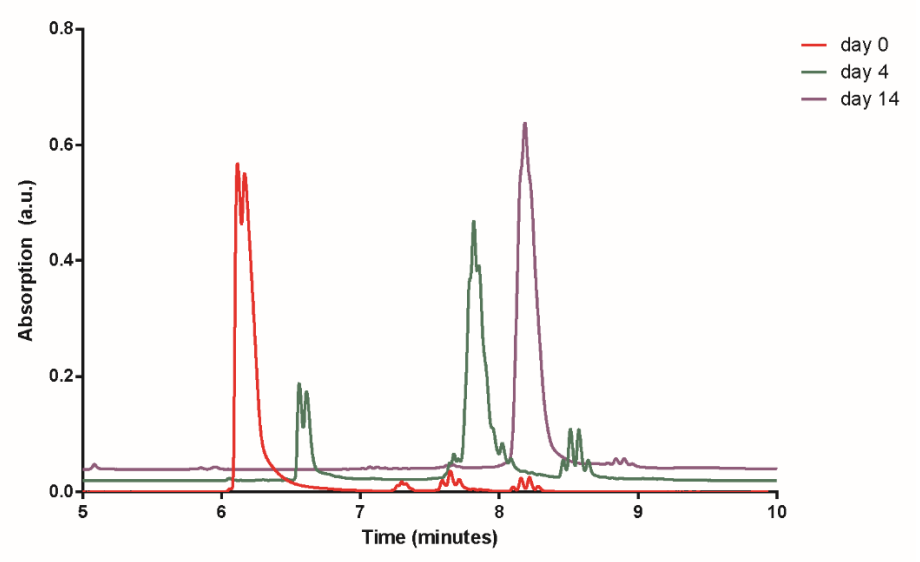


Figure S159. Progression of a library constructed from building blocks **1** and **14**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.

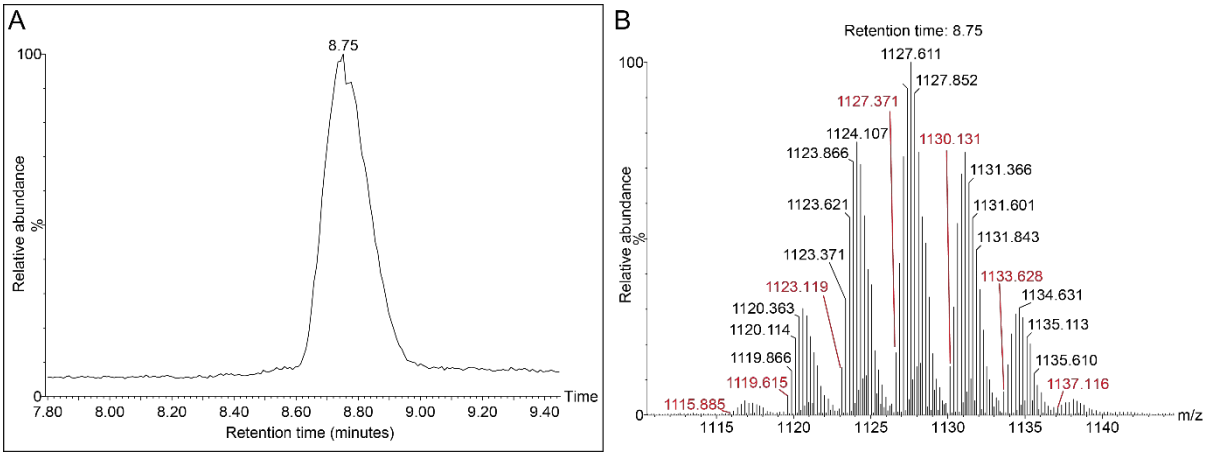


Figure S160. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **1** and **14** after nine days. A) Chromatogram showing the main product of the library. B) Mass spectrum identifying a statistical distribution of hexamers:

Observed mass	Calculated mass	Macrocycle
1115.89	1116.00	$[\mathbf{14}_6+4H]^{4+}$
1119.62	1119.50	$[\mathbf{1}_1\mathbf{14}_5+4H]^{4+}$
1123.12	1123.00	$[\mathbf{1}_2\mathbf{14}_4+4H]^{4+}$
1127.37	1126.50	$[\mathbf{1}_3\mathbf{14}_3+4H]^{4+}$
1130.13	1130.01	$[\mathbf{1}_4\mathbf{14}_2+4H]^{4+}$
1133.63	1133.51	$[\mathbf{1}_5\mathbf{14}_1+4H]^{4+}$
1137.12	1137.01	$[\mathbf{1}_6+4H]^{4+}$

X-GLKFK (1) + X-GLRFK (12)

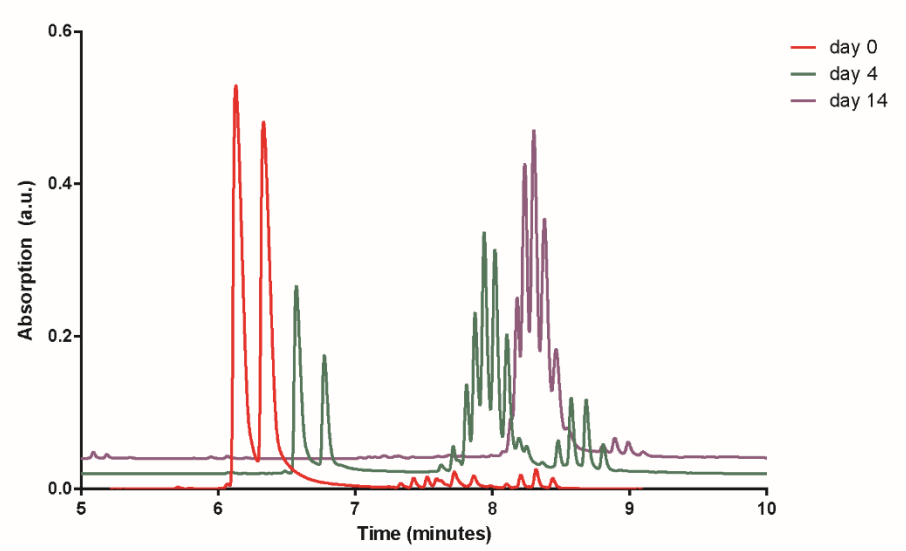
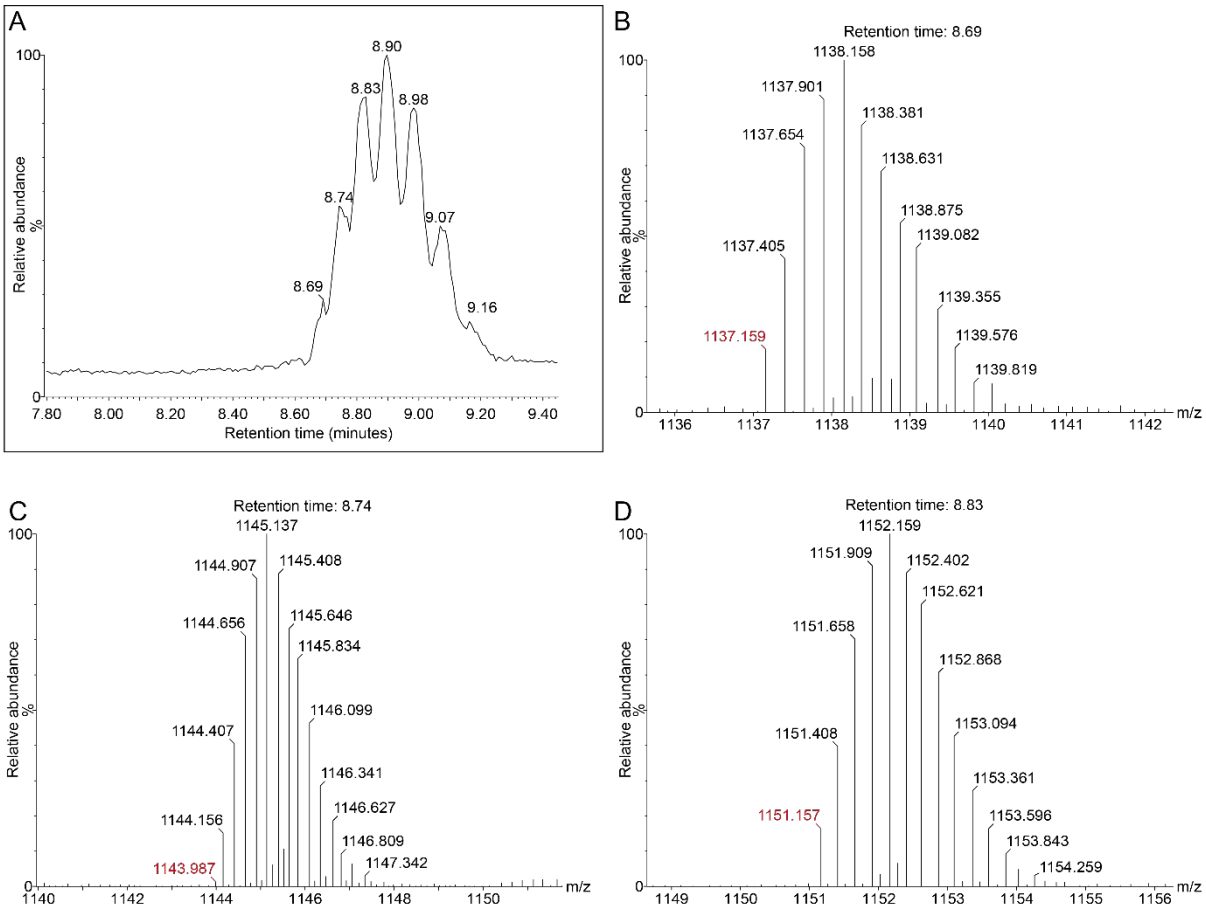


Figure S161. Progression of a library constructed from building blocks **1** and **12**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.


A
Relative abundance
%
100
0
8.90
8.83
8.98
8.74
9.07
8.69
9.16
7.80
8.00
8.20
8.40
8.60
8.80
9.00
9.20
9.40
Retention time (minutes)
B
Retention time: 8.69
1138.158
1137.901
1138.381
1137.654
1138.631
1138.875
1139.082
1137.405
1139.355
1137.159
1139.576
1139.819
1136
1137
1138
1139
1140
1141
1142
m/z
C
Retention time: 8.74
1145.137
1144.907
1145.408
1144.656
1145.646
1145.834
1146.099
1144.407
1146.341
1146.627
1144.156
1146.809
1147.342
1143.987
1140
1142
1144
1146
1148
1150
D
Retention time: 8.83
1152.159
1151.909
1152.402
1152.621
1151.658
1152.868
1153.094
1151.408
1153.361
1151.157
1153.596
1153.843
1154.259
1149
1150
1151
1152
1153
1154
1155
1156

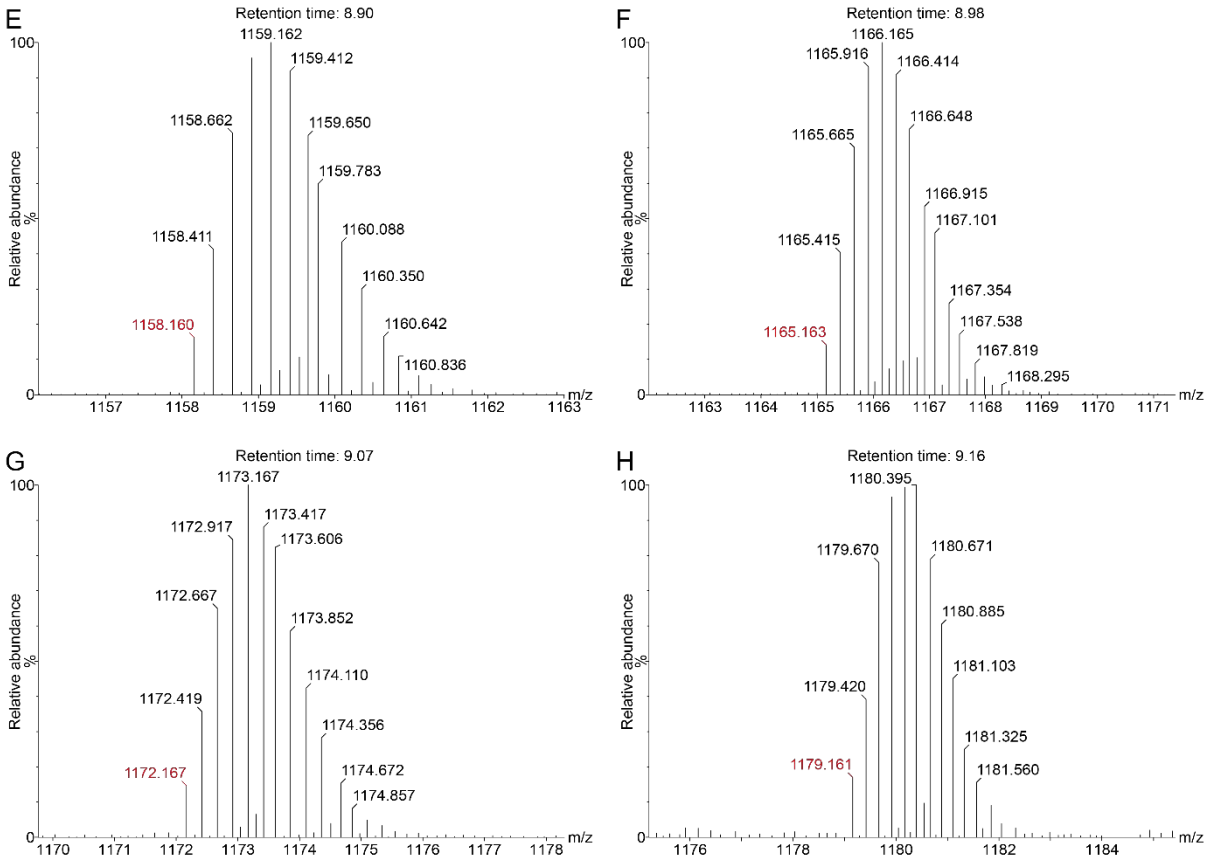


Figure S162. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **1** and **12** after nine days. A) Chromatogram showing the main product of the library. B-H) Mass spectrum identifying a statistical distribution of hexamers:

Observed mass	Calculated mass	Macrocycle
1137.16	1137.01	$[\mathbf{1}_6+4H]^{4+}$
1143.99	1144.01	$[\mathbf{1}_5\mathbf{12}_1+4H]^{4+}$
1151.16	1151.02	$[\mathbf{1}_4\mathbf{12}_2+4H]^{4+}$
1158.16	1158.02	$[\mathbf{1}_3\mathbf{12}_3+4H]^{4+}$
1165.16	1165.02	$[\mathbf{1}_2\mathbf{12}_4+4H]^{4+}$
1172.17	1172.02	$[\mathbf{1}_1\mathbf{12}_5+4H]^{4+}$
1179.16	1179.03	$[\mathbf{12}_6+4H]^{4+}$

X-GLKFK (1) + X-GLK(1-Nal)K (7)

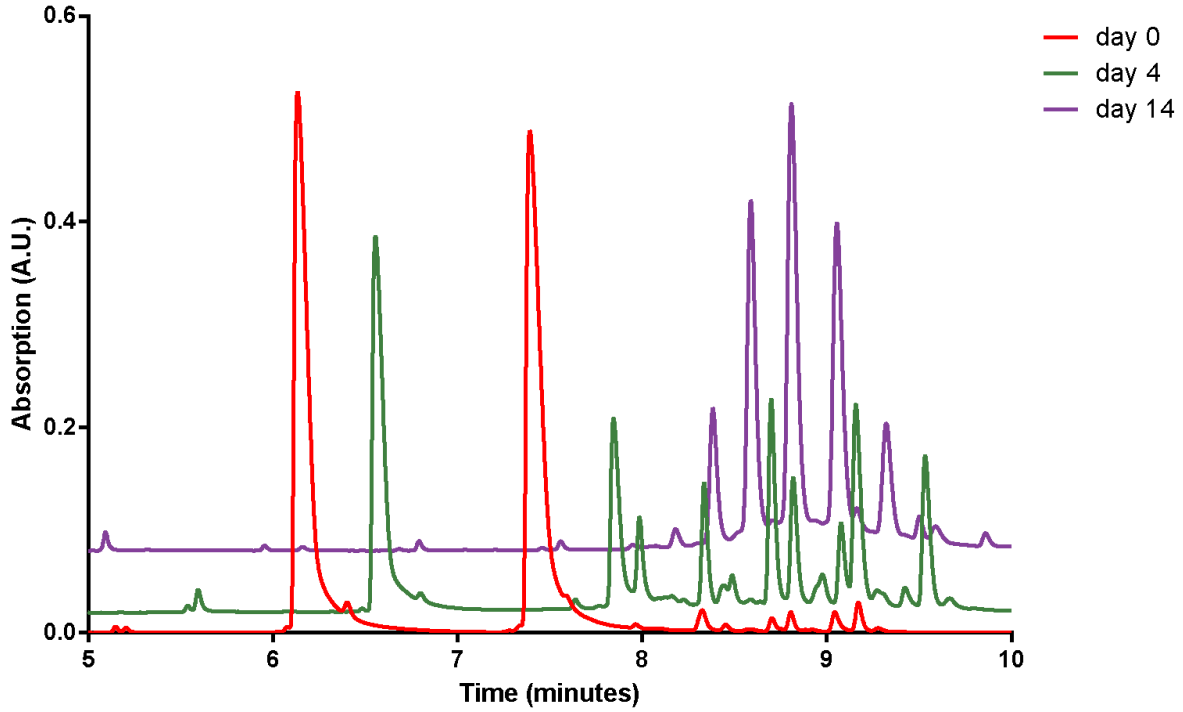
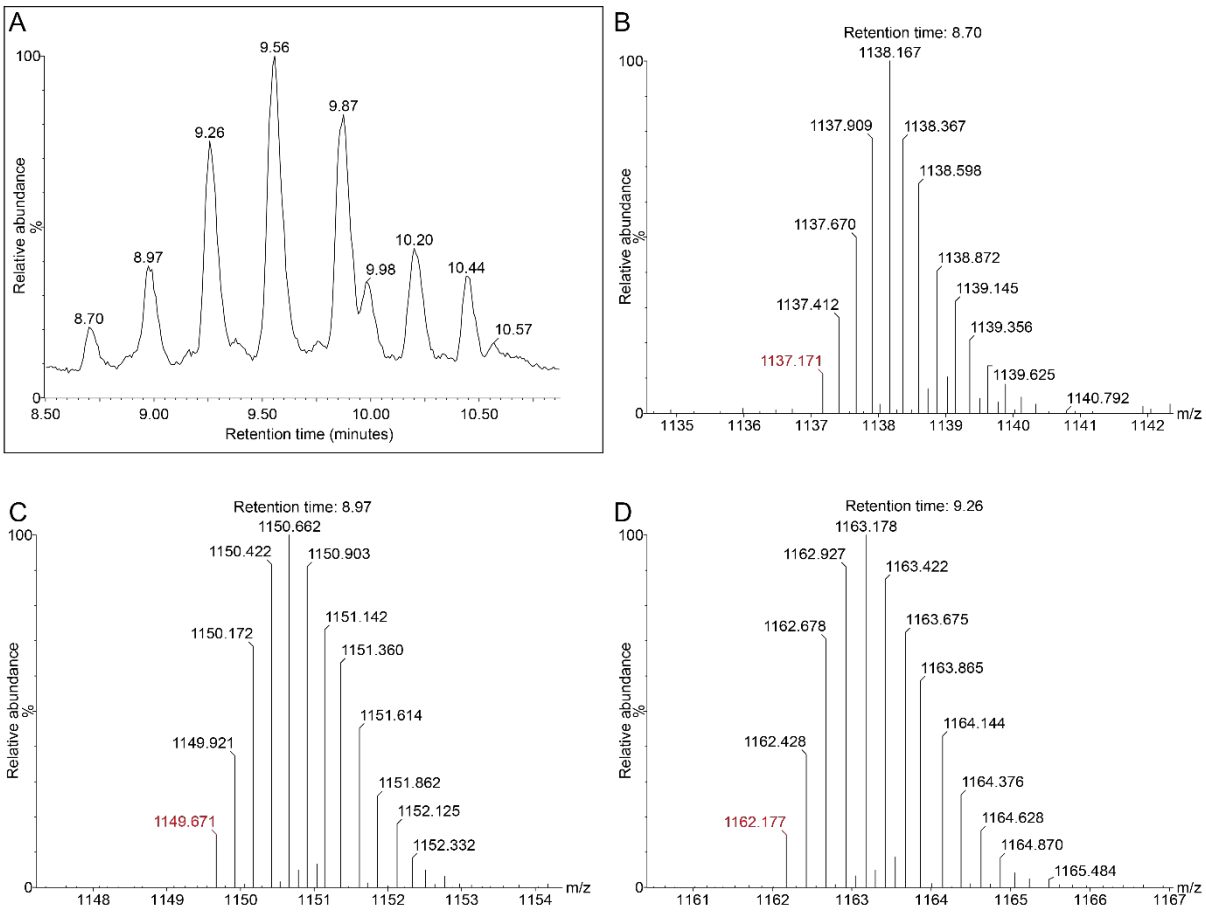


Figure S163. Progression of a library constructed from building blocks 1 and 7. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.



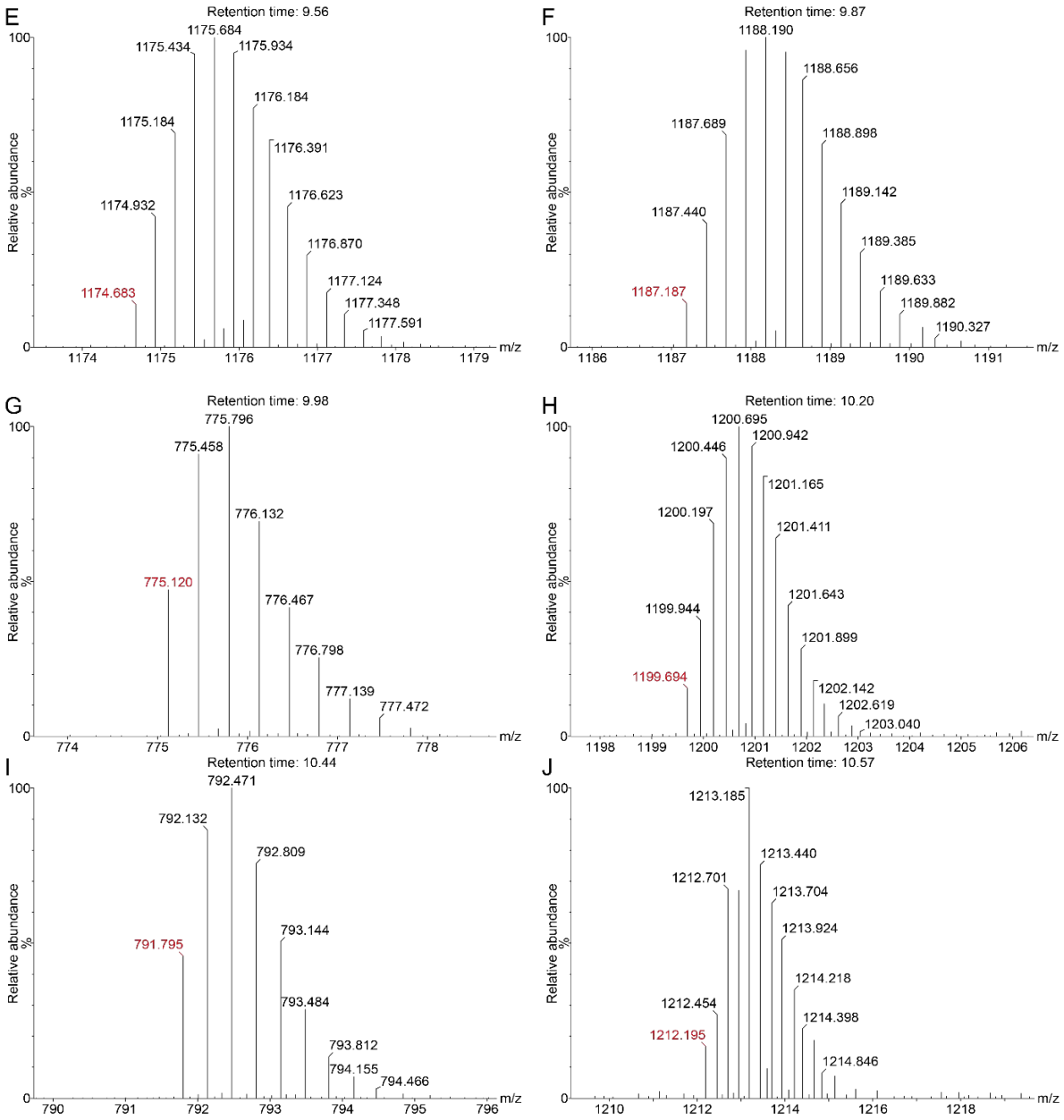


Figure S164. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **1** and **7** after nine days. A) Chromatogram showing the main product of the library. B-J) Mass spectrum identifying a statistical distribution of hexamers and other products**:**

Observed mass	Calculated mass	Macrocycle
1137.17	1137.01	$[\mathbf{1}_6+4H]^{4+}$
1149.67	1149.52	$[\mathbf{1}_5\mathbf{7}_1+4H]^{4+}$
1162.18	1162.02	$[\mathbf{1}_4\mathbf{7}_2+4H]^{4+}$
1174.68	1174.53	$[\mathbf{1}_3\mathbf{7}_3+4H]^{4+}$
1187.19	1187.03	$[\mathbf{1}_2\mathbf{7}_4+4H]^{4+}$
775.12	775.01	$[\mathbf{1}_2\mathbf{7}_1+3H]^{3+}$
1199.69	1199.54	$[\mathbf{1}_1\mathbf{7}_5+4H]^{4+}$
791.80	791.69	$[\mathbf{1}_1\mathbf{7}_2+3H]^{3+}$
1212.20	1212.04	$[\mathbf{7}_6+4H]^{4+}$

X-GLKFK (1) + X-GLKAK (15)

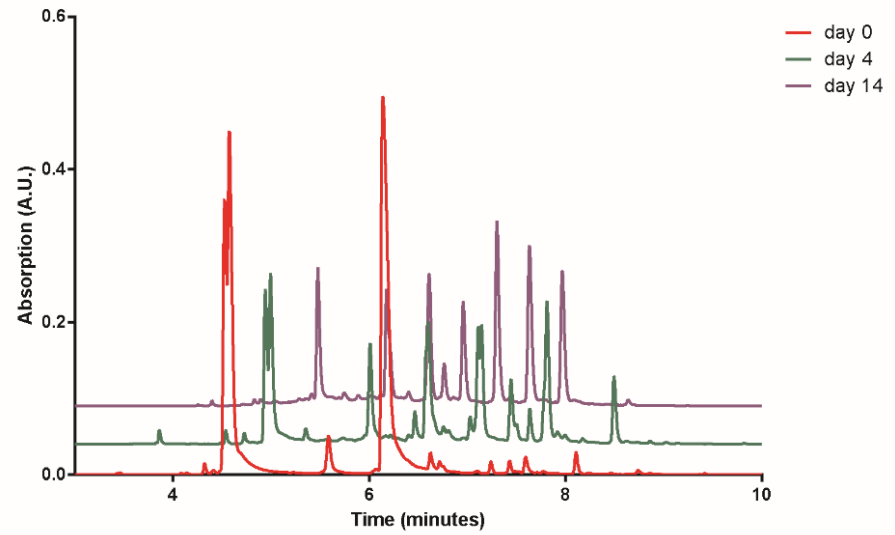
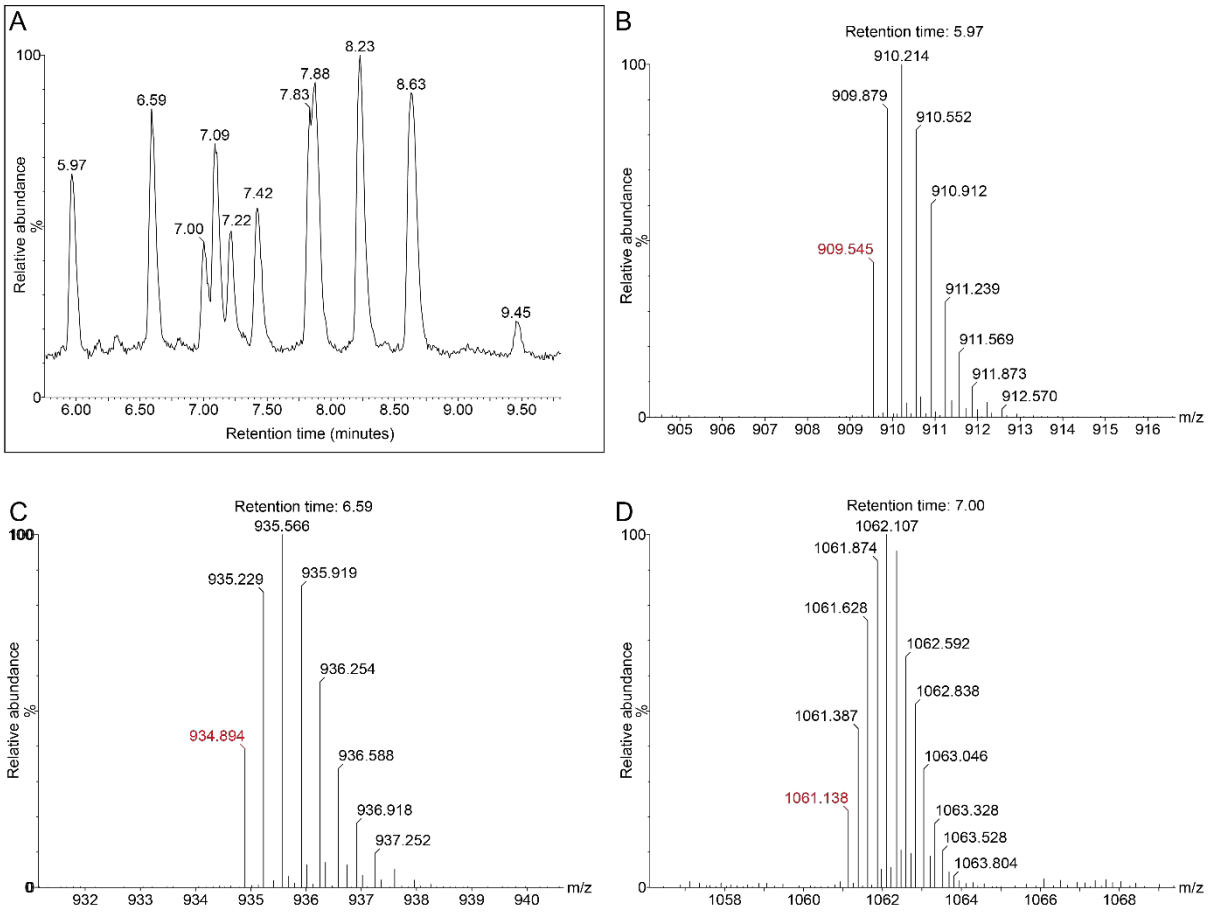


Figure S165. Progression of a library constructed from building blocks **1** and **15**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.



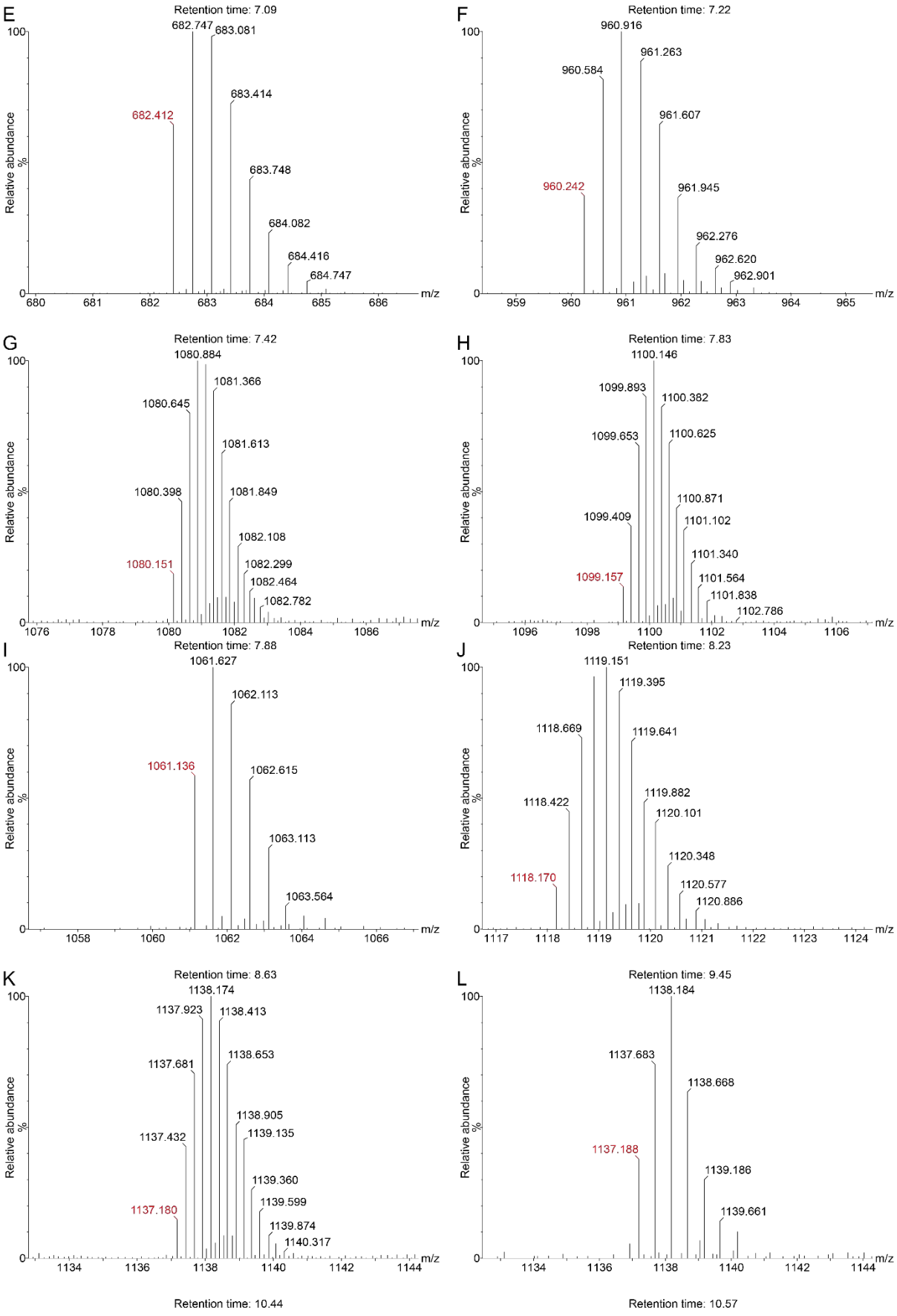

E
Retention time: 7.09
682.747
683.081
683.414
682.412
683.748
684.082
684.416
684.747
F
Retention time: 7.22
960.916
961.263
960.584
961.607
960.242
961.945
962.276
962.620
962.901
G
Retention time: 7.42
1080.884
1081.366
1080.645
1081.613
1080.398
1081.849
1082.108
1080.151
1082.299
1082.464
1082.782
H
Retention time: 7.83
1100.146
1099.893
1100.382
1099.653
1100.625
1100.871
1099.409
1101.102
1101.340
1099.157
1101.564
1101.838
1102.786
I
Retention time: 7.88
1061.627
1062.113
1061.136
1062.615
1063.113
1063.564
J
Retention time: 8.23
1119.151
1119.395
1118.669
1119.641
1119.882
1118.422
1120.101
1120.348
1118.170
1120.577
1120.886
K
Retention time: 8.63
1138.174
1137.923
1138.413
1137.681
1138.653
1138.905
1137.432
1139.135
1139.360
1137.180
1139.599
1139.874
1140.317
Retention time: 10.44
L
Retention time: 9.45
1138.184
1137.683
1138.668
1137.188
1139.186
1139.661
Retention time: 10.57
Relative abundance
%
m/z

Figure S166. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **1** and **15** after nine days. A) Chromatogram showing the main product of the library. B-L) Mass spectrum identifying the formation of hexamers and other products:

Observed mass	Calculated mass	Macrocycle
909.55	909.41	$[\mathbf{15}_4+3H]^{3+}$
934.89	934.76	$[\mathbf{1}_1\mathbf{15}_3+3H]^{3+}$
1061.14	1060.98	$[\mathbf{1}_2\mathbf{15}_4+4H]^{4+}$
682.41	682.31	$[\mathbf{15}_3+3H]^{3+}$
960.24	960.10	$[\mathbf{1}_2\mathbf{15}_2+3H]^{3+}$
1080.15	1079.99	$[\mathbf{1}_3\mathbf{15}_3+4H]^{4+}$
1099.16	1099.00	$[\mathbf{1}_4\mathbf{15}_2+4H]^{4+}$
1061.14	1060.98	$[\mathbf{1}_2\mathbf{15}_1+2H]^{2+}$
1118.17	1118.00	$[\mathbf{1}_5\mathbf{15}_1+4H]^{4+}$
1137.18	1137.01	$[\mathbf{1}_6+4H]^{4+}$
1137.18	1137.01	$[\mathbf{1}_3+2H]^{2+}$

X-GLOFK (14) + X-GLKFR (13)

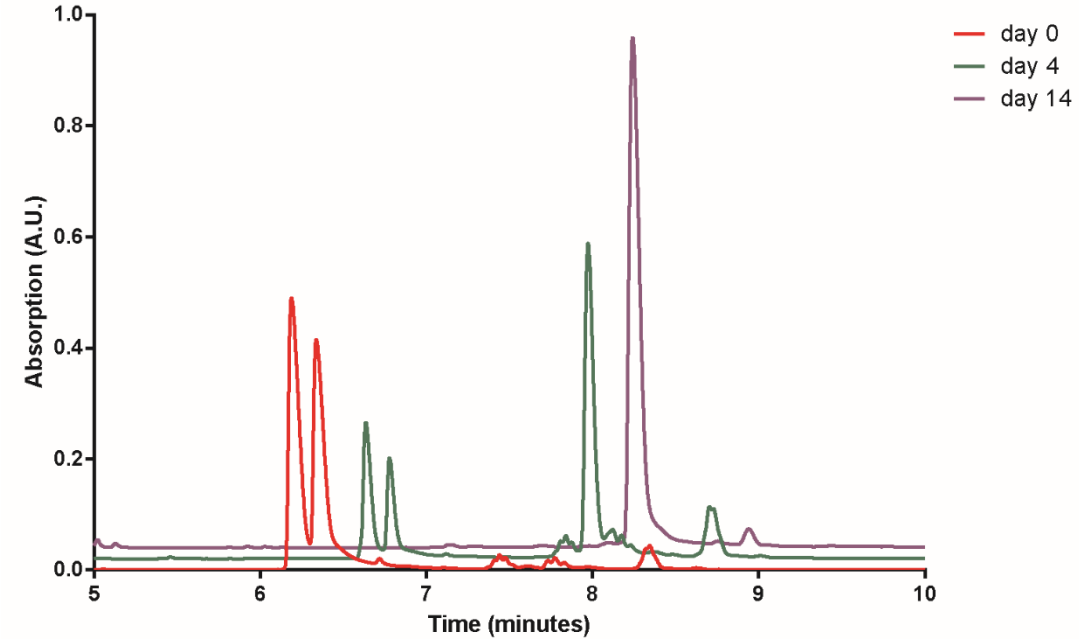


Figure S167. Progression of a library constructed from building blocks **14** and **13**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.

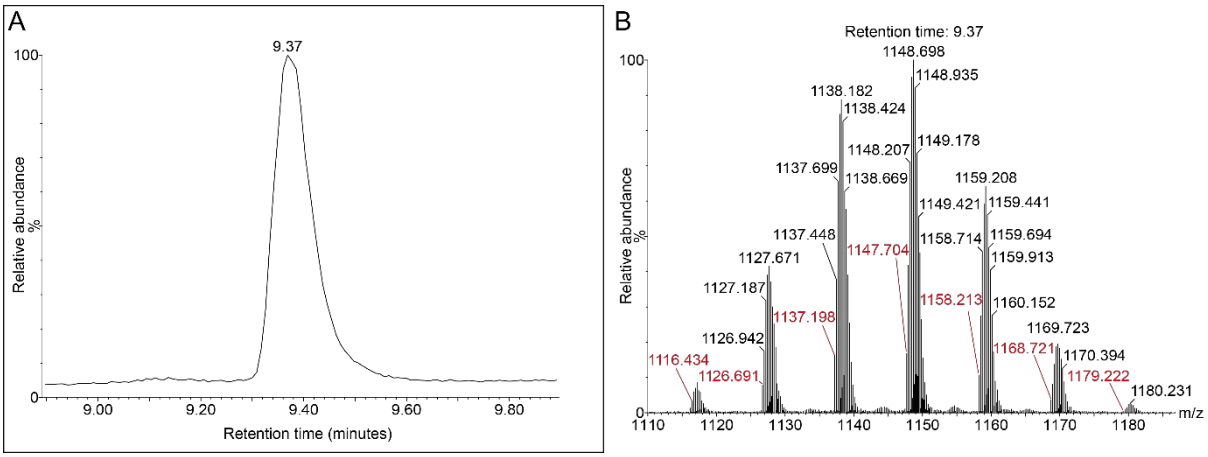


Figure S168. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **14** and **13** after nine days. A) Chromatogram showing the main product of the library. B) Mass spectrum identifying a statistical distribution of hexamers:

Observed mass	Calculated mass	Macrocycle
1116.17	1116.00	$[\mathbf{14}_6+4H]^{4+}$
1126.69	1126.50	$[\mathbf{13}_1\mathbf{14}_5+4H]^{4+}$
1137.20	1137.01	$[\mathbf{13}_2\mathbf{14}_4+4H]^{4+}$
1147.71	1147.51	$[\mathbf{13}_3\mathbf{14}_3+4H]^{4+}$
1158.46	1158.02	$[\mathbf{13}_4\mathbf{14}_2+4H]^{4+}$
1168.72	1168.52	$[\mathbf{13}_5\mathbf{14}_1+4H]^{4+}$
1179.23	1179.03	$[\mathbf{13}_6+4H]^{4+}$

X-GLOFK (14) + X-GLRFK (12)

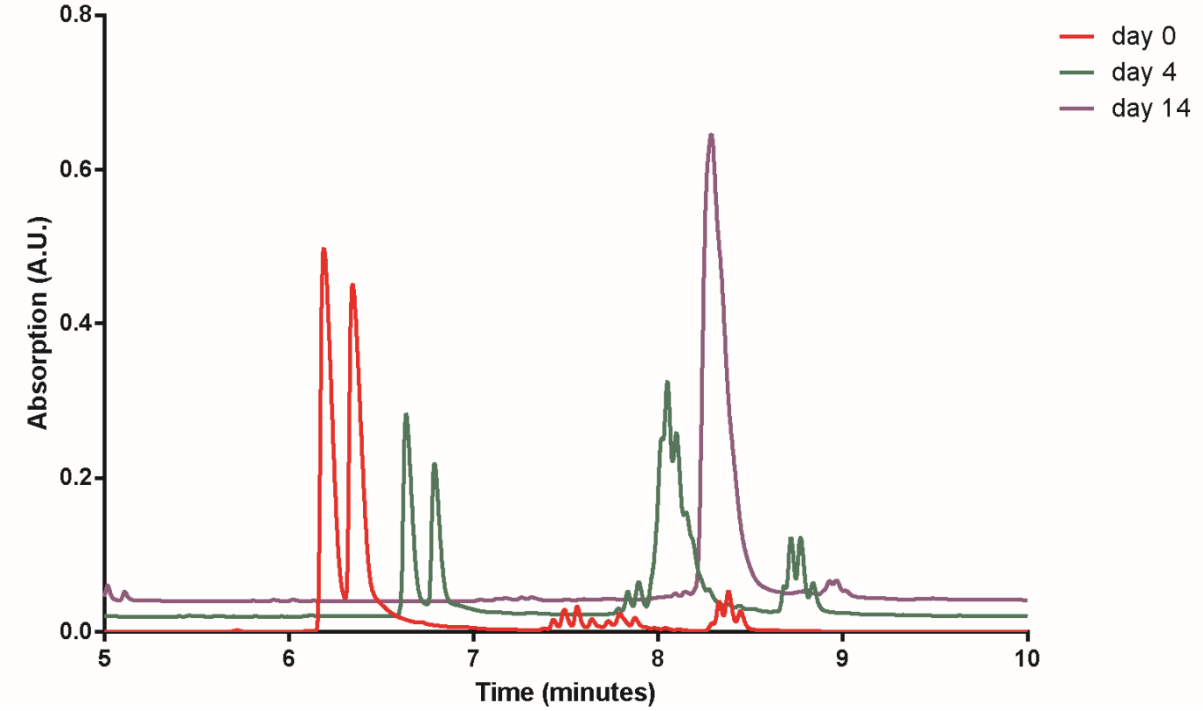


Figure S169. Progression of a library constructed from building blocks **14** and **12**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.

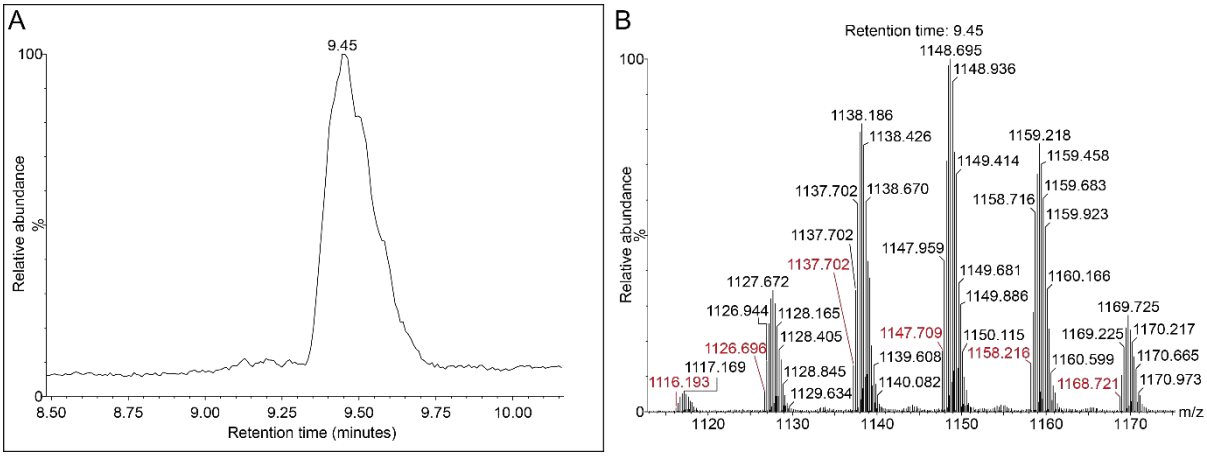


Figure S170. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **14** and **12** after nine days. A) Chromatogram showing the main product of the library. B) Mass spectrum identifying a statistical distribution of hexamers:

Observed mass	Calculated mass	Macrocycle
1116.19	1116.00	$[\mathbf{14}_6+4H]^{4+}$
1126.70	1126.70	$[\mathbf{12}_1\mathbf{14}_5+4H]^{4+}$
1137.70	1137.70	$[\mathbf{12}_2\mathbf{14}_4+4H]^{4+}$
1147.71	1147.71	$[\mathbf{12}_3\mathbf{14}_3+4H]^{4+}$
1158.22	1158.22	$[\mathbf{12}_4\mathbf{14}_2+4H]^{4+}$
1168.72	1168.72	$[\mathbf{12}_5\mathbf{14}_1+4H]^{4+}$

X-GLKWK (9) + X-GLKFR (13)

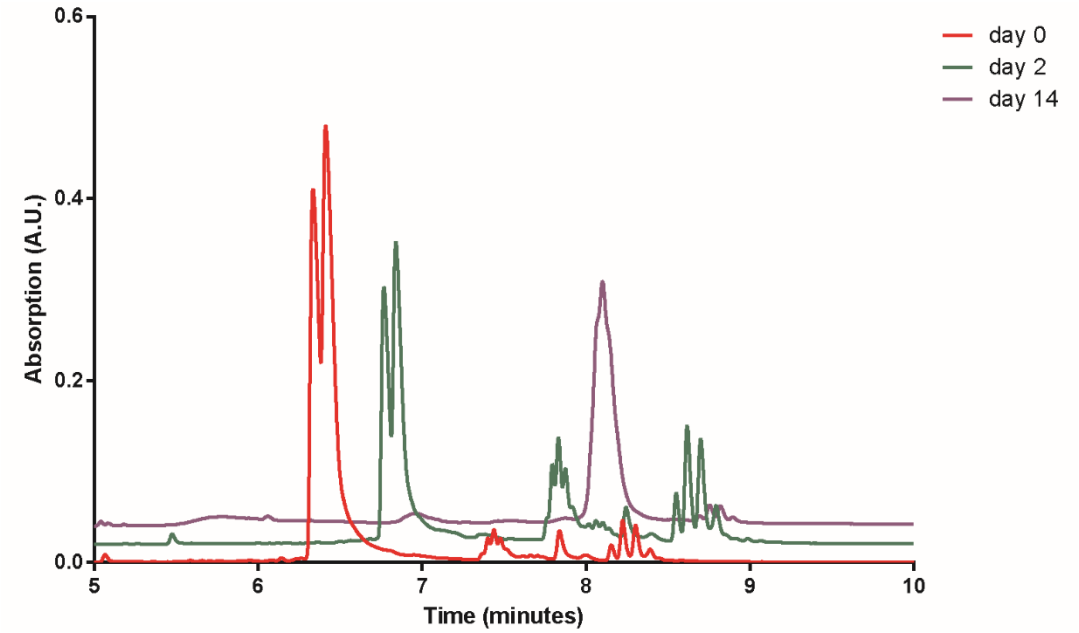


Figure S171. Progression of a library constructed from building blocks **9** and **13**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.

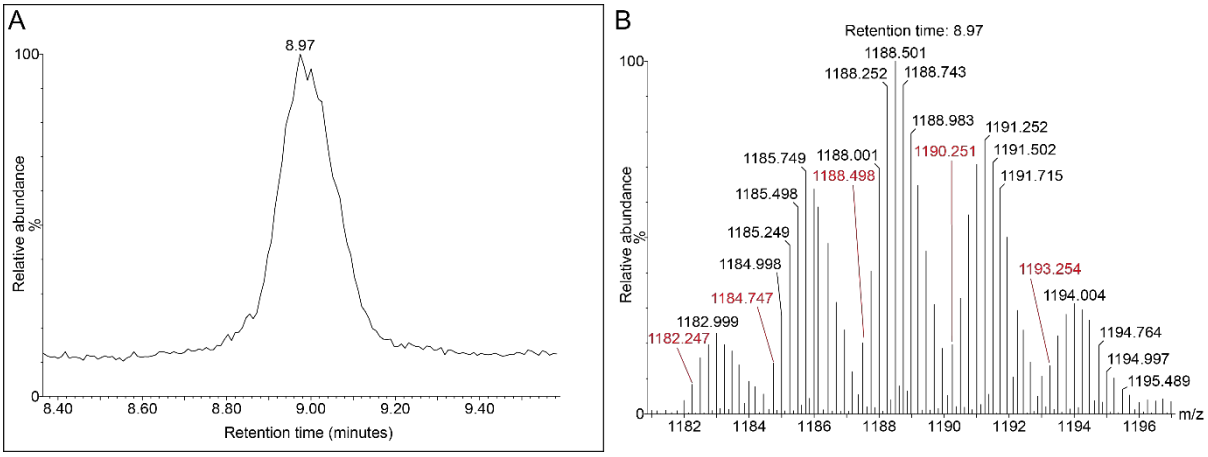


Figure S172. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **9** and **13** after nine days. A) Chromatogram showing the main product of the library. B) Mass spectrum identifying a statistical distribution of hexamers:

Observed mass	Calculated mass	Macrocycle
1182.25	1181.78	$[\mathbf{9}_1\mathbf{13}_5+4H]^{4+}$
1184.75	1184.53	$[\mathbf{9}_2\mathbf{13}_4+4H]^{4+}$
1188.50	1188.28	$[\mathbf{9}_3\mathbf{13}_3+4H]^{4+}$
1190.25	1190.04	$[\mathbf{9}_4\mathbf{13}_2+4H]^{4+}$
1193.25	1192.79	$[\mathbf{9}_5\mathbf{13}_1+4H]^{4+}$

X-GLKWK (9) + X-GLK(1-Nal)K (7)

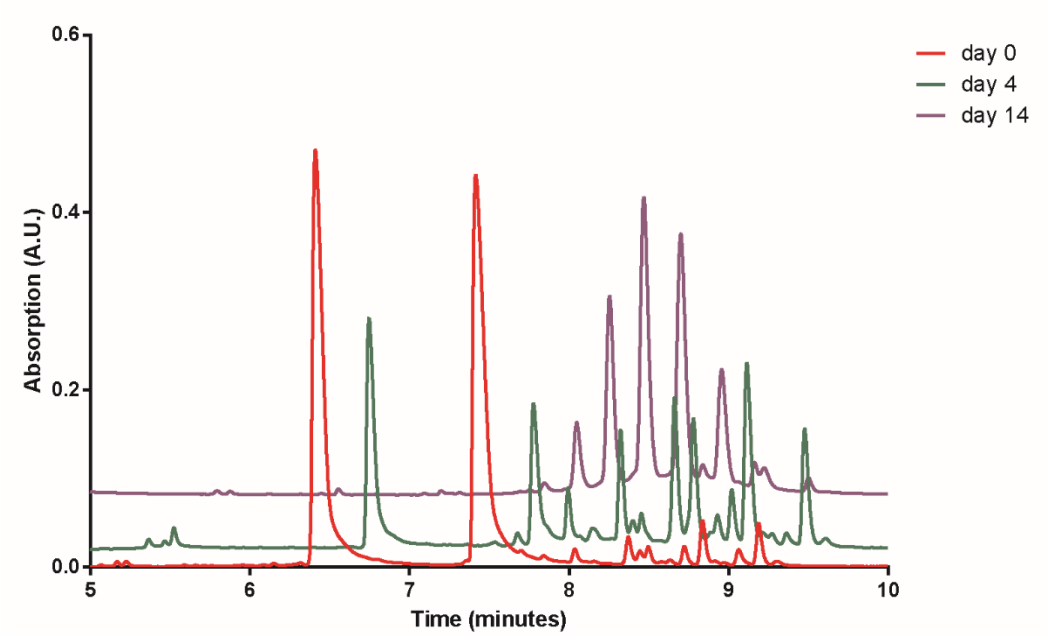
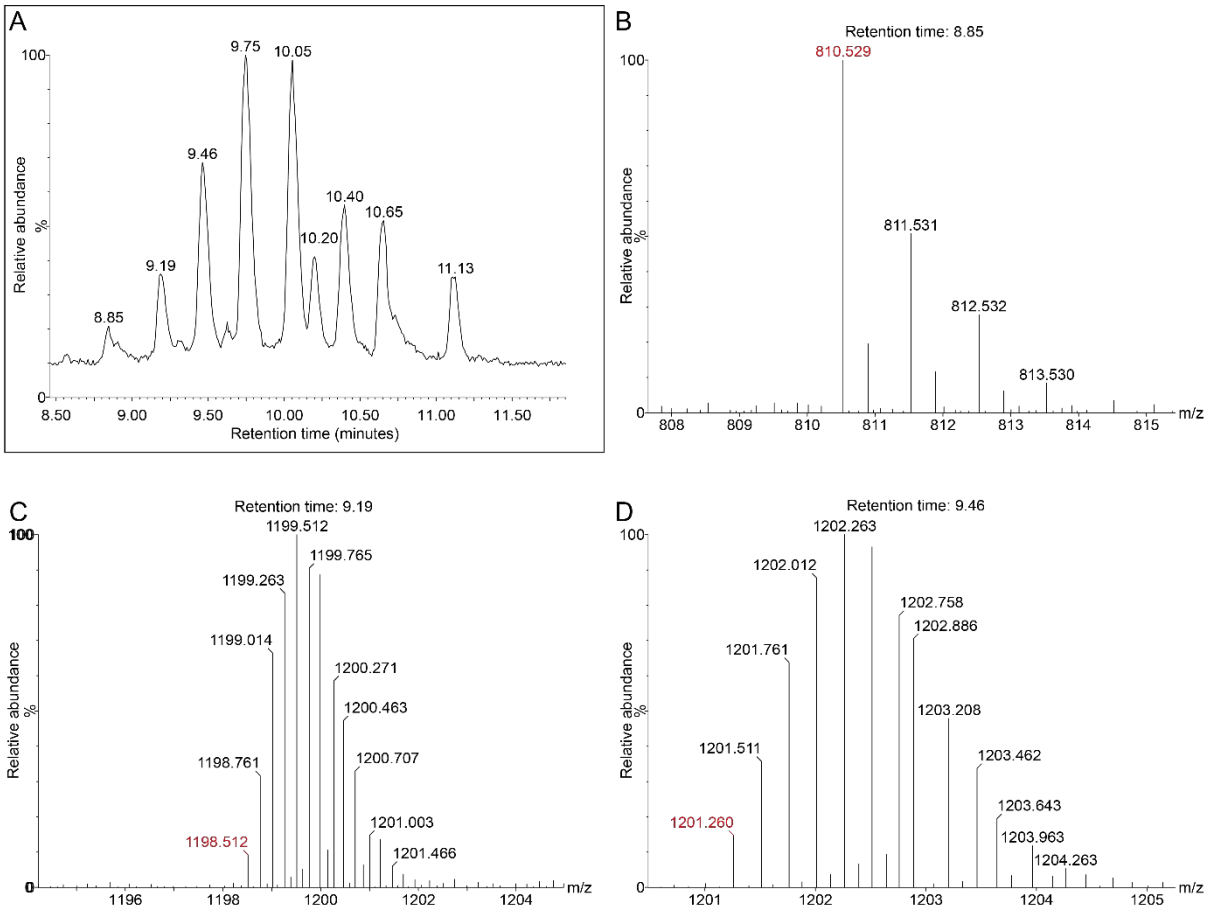


Figure S173. Progression of a library constructed from building blocks **9** and **7**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.



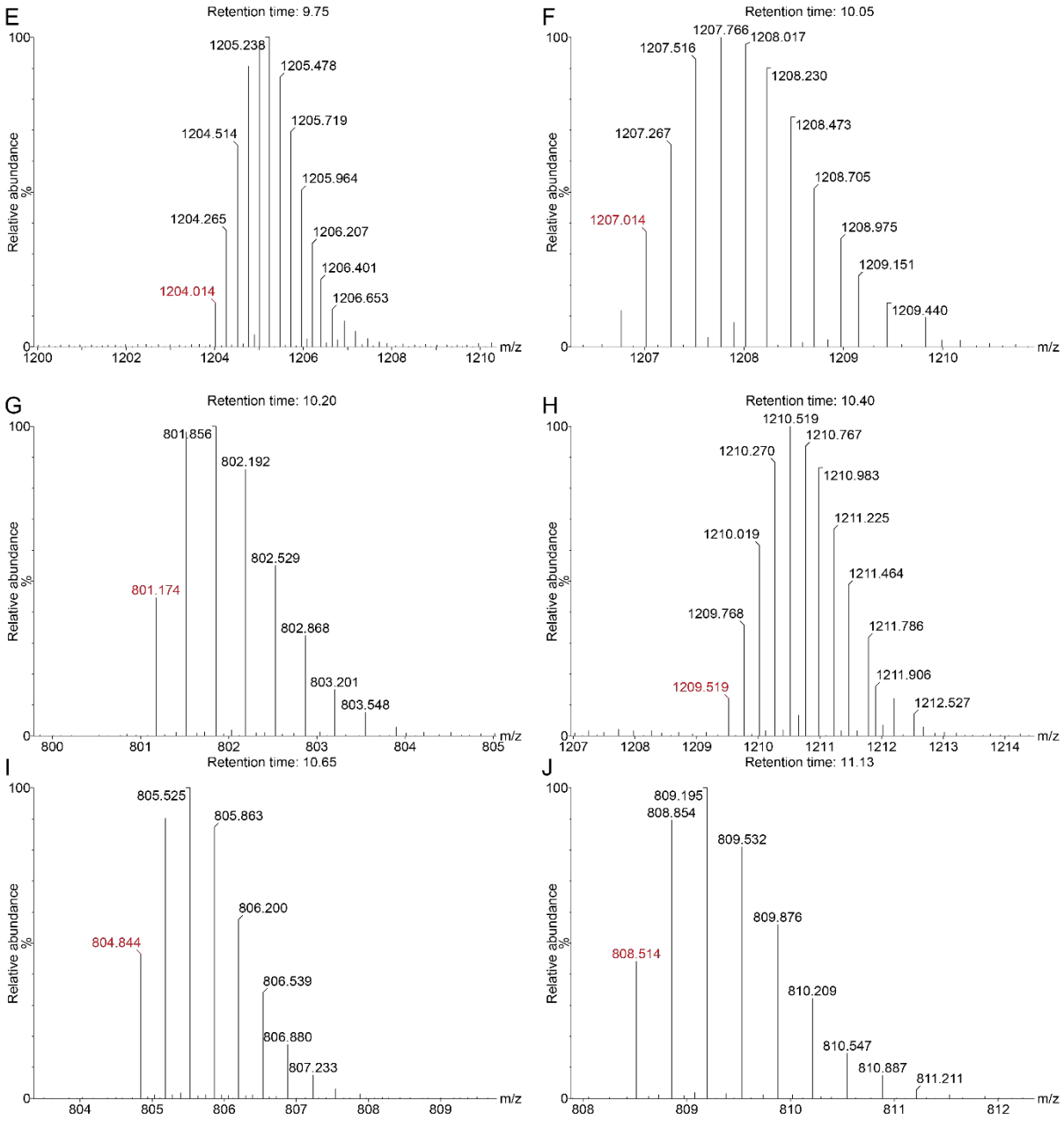


Figure S174. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **9** and **7** after nine days. A) Chromatogram showing the main product of the library. B-J) Mass spectrum identifying the formation of hexamers and other products:

Observed mass	Calculated mass	Macrocycle
810.53	810.36	$[\mathbf{7}_1+1H]^{1+}$
1198.51	1198.29	$[\mathbf{7}_1\mathbf{9}_5+4H]^{4+}$
1201.26	1201.04	$[\mathbf{7}_2\mathbf{9}_4+4H]^{4+}$
1204.01	1203.79	$[\mathbf{7}_3\mathbf{9}_3+4H]^{4+}$
1207.01	1206.54	$[\mathbf{7}_4\mathbf{9}_2+4H]^{4+}$
801.17	801.03	$[\mathbf{7}_1\mathbf{9}_2+3H]^{3+}$
1209.52	1209.29	$[\mathbf{7}_5\mathbf{9}_1+4H]^{4+}$
804.84	804.69	$[\mathbf{7}_2\mathbf{9}_1+4H]^{4+}$
808.51	808.36	$[\mathbf{7}_3+3H]^{3+}$

X-GLKFR (13)+ X-GLKSK (16)

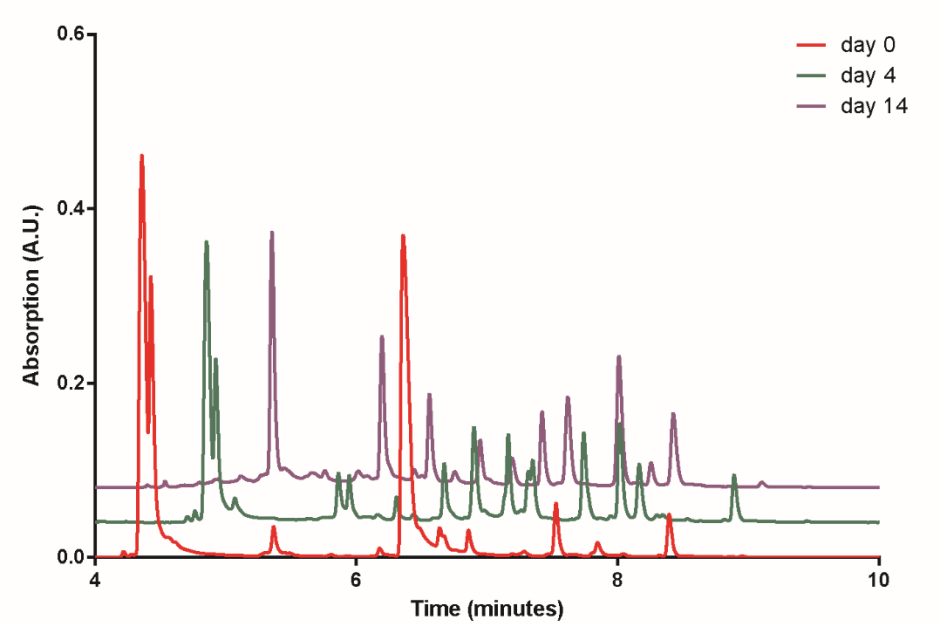
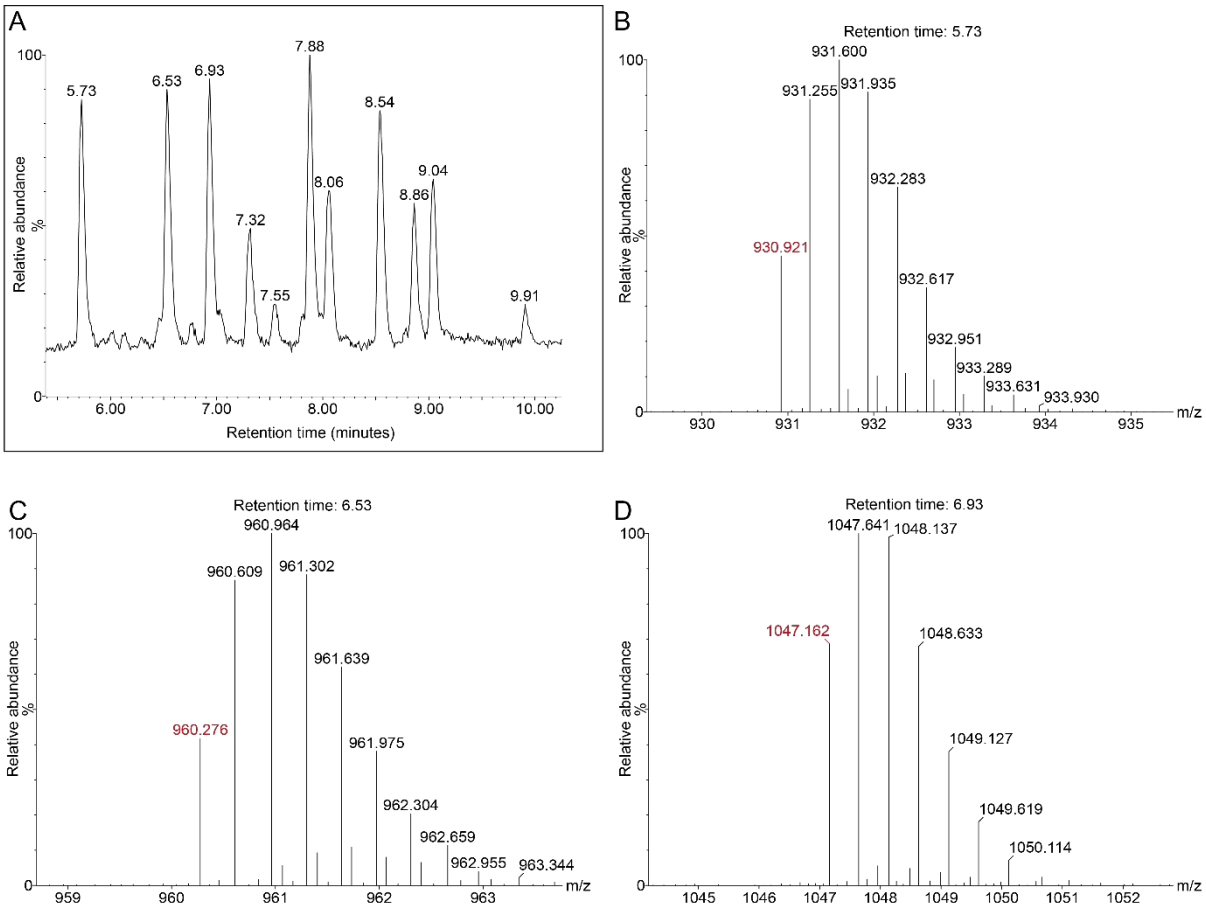


Figure S175. Progression of a library constructed from building blocks **13** and **16**. Red lines mark the library composition at day zero, green lines at day four and purple lines at day 14. Chromatograms are plotted with an offset (on both the X and Y axis) for clarity.



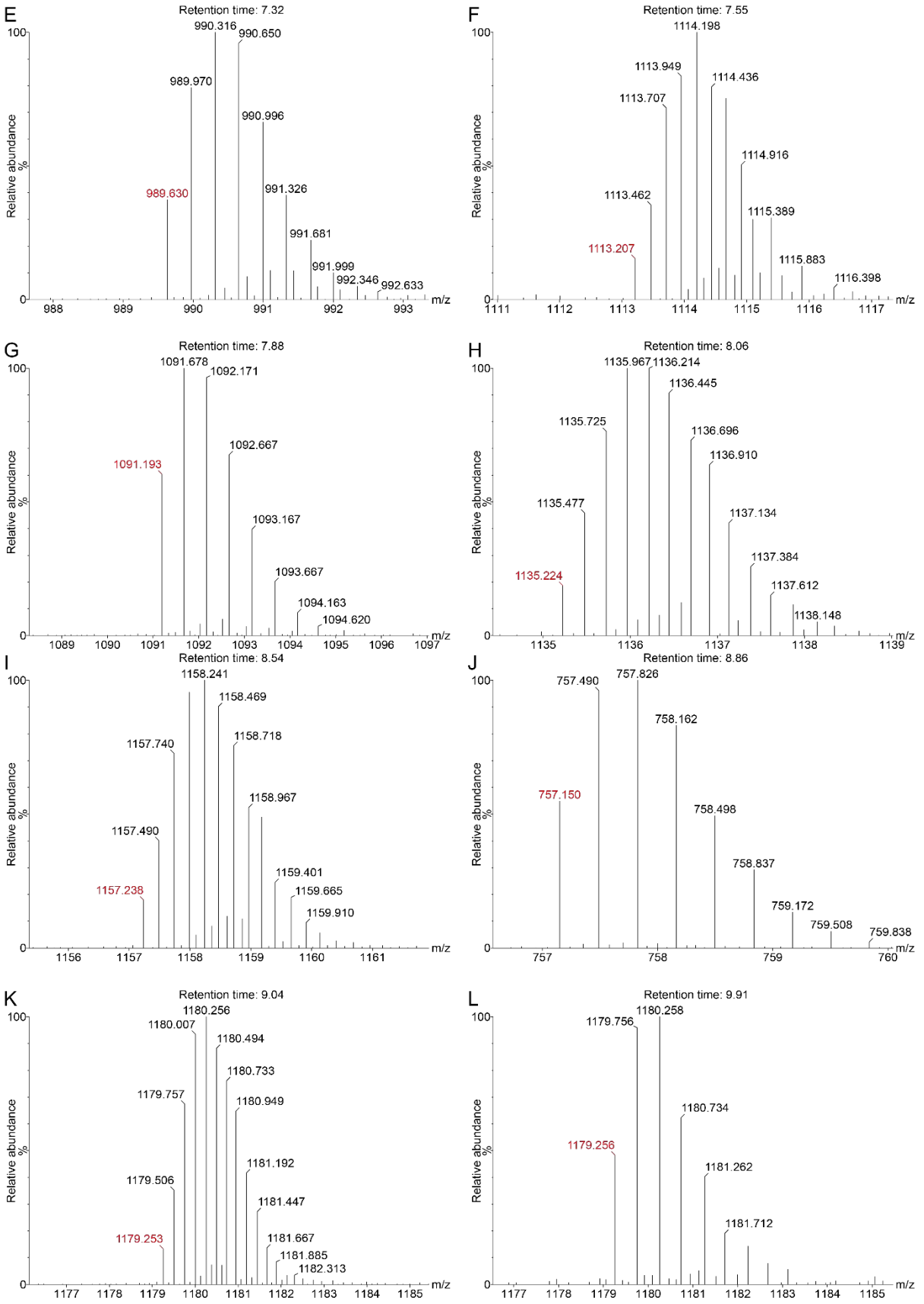


Figure S176. Chromatogram and mass spectrum obtained by LC-MS of a library constructed from building blocks **13** and **16** after nine days. A) Chromatogram showing the main product of the library. B-L) Mass spectrum identifying the formation of hexamers and other products:

Observed mass	Calculated mass	Macrocycle
930.92	930.75	$[\mathbf{16}_4+3H]^{3+}$
960.28	960.06	$[\mathbf{13}_1\mathbf{16}_3+3H]^{3+}$
1047.16	1046.97	$[\mathbf{16}_3+2H]^{2+}$
989.63	989.44	$[\mathbf{13}_2\mathbf{16}_2+3H]^{3+}$
1113.21	1113.00	$[\mathbf{13}_3\mathbf{16}_3+4H]^{4+}$
1091.19	1090.99	$[\mathbf{13}_1\mathbf{16}_2+2H]^{2+}$
1135.22	1135.01	$[\mathbf{13}_4\mathbf{16}_2+4H]^{4+}$
1157.24	1157.02	$[\mathbf{13}_5\mathbf{16}_1+4H]^{4+}$
757.15	757.00	$[\mathbf{13}_2\mathbf{16}_1+3H]^{3+}$
1179.25	1179.03	$[\mathbf{13}_6+4H]^{4+}$
1179.26	1179.03	$[\mathbf{13}_3+2H]^{2+}$

S24. References

1. Markovitch, O., Ottele, J., Veldman, O., and Otto, S. (2020). Automated device for continuous stirring while sampling in liquid chromatography systems. COMMUNICATIONS CHEMISTRY *3*, 180, 180. 10.1038/s42004-020-00427-5.
2. Mandal, A., Boatz, J.C., Wheeler, T.B., and van der Wel, P.C.A. (2017). On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR. Journal of Biomolecular NMR *67*, 165-178. 10.1007/s10858-017-0089-6.
3. Bennett, A.E., Rienstra, C.M., Auger, M., Lakshmi, K.V., and Griffin, R.G. (1995). Heteronuclear decoupling in rotating solids. The Journal of Chemical Physics *103*, 6951-6958. 10.1063/1.470372.
4. Takegoshi, K., Nakamura, S., and Terao, T. (2001). 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. Chemical Physics Letters *344*, 631-637. https://doi.org/10.1016/S0009-2614(01)00791-6.
5. Hohwy, M., Rienstra, C.M., Jaroniec, C.P., and Griffin, R.G. (1999). Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. The Journal of Chemical Physics *110*, 7983-7992. 10.1063/1.478702.
6. Kumashiro, K.K., Schmidt-Rohr, K., Murphy, O.J., Ouellette, K.L., Cramer, W.A., and Thompson, L.K. (1998). A Novel Tool for Probing Membrane Protein Structure: Solid-State NMR with Proton Spin Diffusion and X-Nucleus Detection. Journal of the American Chemical Society *120*, 5043-5051. 10.1021/ja972655e.
7. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293. 10.1007/bf00197809.
8. Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687-696. 10.1002/prot.20449.
9. Harris, R.K., Becker, E.D., De Menezes, S.M., Granger, P., Hoffman, R.E., and Zilm, K.W. (2008). Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008). Magn Reson Chem *46*, 582-598. 10.1002/mrc.2225.
10. Morcombe, C.R., and Zilm, K.W. (2003). Chemical shift referencing in MAS solid state NMR. J Magn Reson *162*, 479-486. 10.1016/s1090-7807(03)00082-x.

11. Tamiola, K., Acar, B., and Mulder, F.A.A. (2010). Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. Journal of the American Chemical Society *132*, 18000-18003. 10.1021/ja105656t.
12. Merg, A.D., Boatz, J.C., Mandal, A., Zhao, G., Mokashi-Punekar, S., Liu, C., Wang, X., Zhang, P., van der Wel, P.C.A., and Rosi, N.L. (2016). Peptide-Directed Assembly of Single-Helical Gold Nanoparticle Superstructures Exhibiting Intense Chiroptical Activity. J Am Chem Soc *138*, 13655-13663. 10.1021/jacs.6b07322.
13. Matlahov, I., and van der Wel, P.C.A. (2018). Hidden motions and motion-induced invisibility: Dynamics-based spectral editing in solid-state NMR. Methods *148*, 123-135. https://doi.org/10.1016/j.ymeth.2018.04.015.
14. Yang, S., Schaeffer, G., Mattia, E., Markovitch, O., Liu, K., Hussain, A.S., Ottelé, J., Sood, A., and Otto, S. (2021). Chemical fueling enables molecular complexification of self-replicators. Angew. Chem. Int. Ed. *60*, 11344-11349.