Article

Controlling Supramolecular Plasticity through Stoichiometry

Dávid Komáromy,^a Theodora Tiemersma-Wegman,^a Johan Kemmink,^a Giuseppe Portale,^b Friso S. Aalbers,^c Ivana Marić,^a Guillermo Monreal Santiago,^a Jim Ottelé,^a Ankush Sood,^a Vittorio Saggiomo,^{a,d} Bin Liu,^a Pieter van der Meulen,^a and Sijbren Otto^a

SUMMARY

Self-assembly is a powerful method to access complex structures, but is traditionally confined to structures that represent global minima in the Gibbs energy landscape. We now report that it is possible to access structures other than those with the lowest individual Gibbs energy, while remaining under thermodynamic control. By tuning building block stoichiometry it may be more efficient for the system as a whole to produce many structures of moderate stability as opposed to a smaller number of high-stability assemblies. We prepared dynamic combinatorial libraries by mixing two building blocks designed to have complementary binding motifs. Depending on the building block stoichiometry one of three competing self-assembling macrocycles can be formed with remarkable selectivity. Thus, from the same two building blocks we could access a self-replicating octamer, self-assembling into fibers; a hexamer with a precise 4:2 building block stoichiometry, assembling into hexagonally packed fiber bundles; and a tetramer with 1:3 building block ratio affording a [c3]-daisy chain pseudorotaxane. These results show that systems chemistry approaches can enhance the versatility of self-assembly by allowing to navigate complex Gibbs energy landscapes to access structures that go beyond those that are individually the most stable.

INTRODUCTION

Chemistry involves navigating through complex Gibbs energy landscapes. In covalent chemistry (i.e. organic synthesis) the challenge resides in guiding starting material(s) across the right saddle points (activation barriers) to selectively end up in the energy well which corresponds to the desired product. In most cases the reactions are controlled by kinetics (i.e. heights of activation barriers). Great advances have been made in this area of covalent chemistry and, given enough time and resources, essentially any target molecule can now be accessed by synthesis.

Until recently supramolecular chemistry^{1,2,11–13,3–10} has followed a different paradigm. Constructing a specific supramolecular assembly has traditionally required it to be the thermodynamically most stable structure. Not barrier heights, but depths of the wells in the Gibbs energy landscape typically determine which assembly is formed. This approach to self-assembly is limiting, as for a given system only assemblies that correspond

The Bigger Picture

Self-assembly is a powerful approach to access complex assemblies in chemistry as well as in biology. As self-assembly relies on relatively weak non-covalent interactions it tends to produce assemblies that have the lowest Gibbs energy. This tendency restricts the diversity of structures that may be accessed. However, when considering systems in which assemblies compete for common building blocks, it is not necessarily the structure that is thermodynamically the most stable that wins the competition. This study shows that the ratio of components of the self-assembling structures can control which assembly is formed and that assemblies can be transformed into other assemblies upon adjusting the ratio of their components. These principles increase the versatility of thermodynamically controlled self-assembly.

^a Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

^b Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

^c Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

^d Present Address: Laboratory of BioNanoTechnology, Wageningen University, PO Box 8038, 6700 EK, Wageningen, The Netherlands

to the thermodynamic minimum can be obtained. More recently it has become appreciated that Gibbs energy landscapes pertaining to noncovalent interactions can feature energy barriers that cannot be transversed using only thermal energy, allowing specific supramolecular assemblies to be accessed by kinetic control.¹⁴⁻²⁰ In principle, such pathway control greatly expands the range of structures that may be accessed by self-assembly. In practice, the still limited mechanistic insights in self-assembly processes hampers our ability to make full use of the opportunities that are now opening up.

Dynamic combinatorial chemistry (DCC) has proven to be a powerful tool for probing complex Gibbs energy landscapes of systems featuring non-covalent interactions.^{21,22} In DCC, reactions between building blocks equipped with functional groups capable of reversible bond formation give rise to a dynamic combinatorial library (DCL), i.e. a mixture of interconverting oligomers. Reversibility implies that the DCL composition can be shifted towards one (or more) specific oligomer(s) by non-covalent interactions either among DCL members or between an external template and (a) DCL member(s). Most of the work on DCLs made use of external templates.²³⁻²⁶ Only more recently DCC has been used for exploring self-assembly processes, leading to foldamers,²⁷ interlocked structures²⁸⁻³² surfactant assemblies³³⁻³⁵ or assembly-driven self-replication.³⁶⁻³⁹ In most of these systems a single product was obtained. Yet DCC should allow for accessing multiple different structures, even when controlled by thermodynamics. The simplest, almost trivial, approach is to vary building block stoichiometry in a DCL prepared from more than one building block. In systems in which library members with different stoichiometries compete for common building blocks it is not necessarily the most stable product that dominates. Instead, the library member with a building block stoichiometry that matches the composition of the library may become the dominant product, as the system can produce more of this compound than of the competing species with different stoichiometries. Thus, the system may produce many species that, individually, feature weaker non-covalent interactions, as opposed to a smaller number of species that form individually stronger interactions. It is well established that such mass-balance related effects can override the individual host-guest binding strengths in DCLs of synthetic receptors.⁴⁰⁻⁴³ We now show that similar effects can be utilized to navigate multi-well self-assembly Gibbs energy landscapes. We introduce a systems chemistry approach by which the stoichiometry of the components controls, with remarkable selectivity, which well in the Gibbs energy landscape becomes populated, hence, which assembly is produced. We show that, in a DCL made from two building blocks with complementary binding sites, three markedly different types of self-assembling structures are formed, i.e. fibers, fiber bundles or molecular daisy chains **1** (Scheme 1). The system can be instructed to afford a specific assembly solely through building block stoichiometry, even when, individually, this assembly does not correspond to the deepest well in the Gibbs energy landscape.

RESULTS

Selective Formation of Three Different Molecules in a Two-Building-Block DCL

In order to achieve self-assembly in DCLs building block should be designed to be complementary to each other. This principle has proven potential in protein assembly. For example, inserting complementary amino acids at

more or less well-defined positions of known proteins can result in their assembly into non-natural types of aggregates.^{44,45}

In order to explore building block complementarity in DCC, we made use of already well-characterized systems arising from building blocks **1** or **4**. Peptide derived building block **1** features two lysine residues while the tetramer macrocycle derived from carboxylate functionalized building block **4** has a high affinity for protonated alkylamines, such as spermine (**Sp**), which are bound through a combination of hydrogen bonds, electrostatic interaction and hydrophobic interactions.⁴⁶ Thus the ingredients are in place for binding between (oligomers of) **4** and the lysine groups of **1**.

Oxidation of a stirred aqueous solution of dithiol building block **1** (Scheme 1) gives rise to a small DCL of macrocyclic disulfide oligomers of different ring sizes, among which octamer **1₈** is stabilized by self-assembly into non-discrete supramolecular polymer nanofibers⁴⁷ and is thus formed preferentially.

In order to probe whether interactions between **1** and **4** can indeed give rise to new self-assembling species, we prepared a series of DCLs containing the same total amount of building blocks ($[1] + [4] = 2.0 \text{ mM}$) with increasing molar ratios of **4**, ranging from 0 to 100 mol%. The libraries were stirred at 1200 rpm at room temperature and their composition was monitored with UPLC-MS (Supplemental Information, Sections 2.2.1 and 2.3.) until no further change was detected (typically 7-8 days). We anticipated that, in the absence of additional secondary interactions, a large number of small oligomers would form and the building block composition for a given oligomer size would follow a binomial distribution, reflecting the building block stoichiometry. Instead, we observed selective and preferential formation of three disulfide macrocycles with different ring sizes and stoichiometries (Figure 1A, traces (1)-(5)). In libraries containing 0-8 mol% of **4** the octamer **1₈** dominates, as expected. Between 8-20 mol% **4** we could no longer detect **1₈**. Instead a broad distribution of products was formed, differing in ring size and composition. At 20 mol% of **4** a mixed hexamer with stoichiometry **1₄4₂** appeared, becoming the dominant species (90 mol%) at 35 mol% **4**. Between 35-75 mol% **4**, the mixed hexamer persists, however, also tetramer **1₁4₃** appears as the sole other product. Oligomers **1₄4₂** and **1₁4₃** coexist up to 75 mol % of **4**. At 80 mol% **4** the mixed tetramer is the dominant species, whereas upon further increasing the relative amount of **4** to 95 mol%, the mixed tetramer is still present, eventually giving way to **4**-only oligomers when the mol% of **4** approached 100%. (See Figure S59 for detailed chromatograms).

Scheme 1. Dynamic Combinatorial Library Made From Building Blocks 1 and 4 Can Give Rise to Three Different Self-Assembling Species: 1_8 (fibers), 1_4_2 (fiber bundles) or $(1_14_3)_3$ ([c3]daisy-chain pseudorotaxane).

Figure 1. Selective Emergence of Three Different Macrocycles Can be Directed Solely by Stoichiometry.

(A) UPLC traces of DCLs prepared from **1** and **4** ($[1] + [4] = 2.0$ mM) in aqueous phosphate buffer (50 mM, pH = 8.15; for (1)-(5)) or aqueous acetate buffer (50 mM, pH = 3.2; for (6)) (1) 0 mol % (2) 10 mol % (3) 40 mol % (4) 60 mol %, (5) 80 mol %, (6) 33 mol % **4**.

(B) Relative amount of preferred species **18**, **1442** and **1433** as a function of mol % of **4** ($[1] + [4] = 2.0$ mM) in aqueous phosphate buffer (50 mM, pH = 8.15).

Importantly, each macrocycle forms in exceptionally high yields (corresponding to 90 % of the total UPLC peak area^a) if the initial building block ratio approximates the macrocycle stoichiometry, i.e. 0, 35 and 80 mol % of **4** for **18**, **1442** and **1433**, respectively (Figure 1B). The high yields indicate that, at the specified stoichiometries, the corresponding macrocycle is formed preferentially. Note that between 35-80 mol% **4** only **1442** and **1433** are formed, representing a novel type of stoichiometry-specific social self-sorting.⁴⁸

Structure and assembly of Octamer **18**

The systems chemistry of a **1**-only system has been extensively studied before.^{47,49} Upon oxidation of a stirred aqueous solution of **1**, initially mainly cyclic trimer and tetramer disulfides are formed, alongside a small amount of octamer **18**. The latter self-assembles into fibrous stacks, which further self-assemble into laterally assembled helical fibers (Figure S64).⁵⁰ Fiber growth is autocatalytic, but the activation barrier for octamer formation and assembly nucleation is relatively high. As the stacks grow beyond a certain length, they become susceptible to mechanical breakage (conferred by stirring), whereby each breakage event redoubles the number of catalytic fiber ends. Finally, due to reversible thiol-disulfide exchange, the DCL material is channeled almost exclusively into octamer stacks via a fiber growth-breakage mechanism, thereby realizing exponential self-replication.³⁸

^a Note that at the applied detection wavelength (254 nm) the extinction coefficients of the two building blocks differ by ca. 30% (Figure S1). Thus, peak areas do not accurately reflect concentrations. However, as our estimations show, the calculated relative molar percentage values obtained from the relative UPLC peak areas (by taking into consideration the difference in extinction coefficients) differ at most 10% from the actual relative UPLC peak areas (Supplemental Information, Section 1.18).

Figure 2. Elucidating the Molecular and Supramolecular Organization of the Mixed Hexamer **1442.**

(A) ESI-Orbitrap mass spectrum of **1442** after mass selection of the $[1442]^{3+}$ ($m/z = 1083$ Da) ion at collision voltages of 0 V (top) and 15 V (bottom). Inset: three possible positional isomers of **1442**. The appearance of $[14]^{3+}$ indicates the presence of mainly the *ortho* isomer. Asterisks indicate consecutive water loss from the indicated parent ions.

(B) Cryo-TEM shows the presence of a dense network of fibres. (see also Figure S67 for better magnification).

(C) AFM shows densely associated fibers forming linear, bundled (red arrow) and circular (blue arrow) assemblies, which are 5–15 μm in length.

(D) CD spectroscopy shows characteristic peptide β -sheet signatures (at 195 and 205 nm).

(E) ThT fluorescence assay indicates the presence of peptide β -sheets.

Structure and assembly of the Mixed Hexamer **1442**

Subsequently, we attempted to uncover the factors bringing about the highly specific stoichiometry and ring size observed for hexamer **1442**. At pH 8.2 each unit of **1** has a single positive charge (two protonated lysine side chains and one terminal carboxylate), whereas each unit of **4** carries two negative charges (two carboxylates). Consequently, in oligomers composed of **1** and **4** in a 2:1 molar ratio, charge neutrality is potentially satisfied. Moreover, a DCL with the same stoichiometry and concentration ($[1] + [4] = 2.0$ mM, $[1]:[4] = 2:1$) at considerably lower pH (50 mM acetate buffer, pH = 3.2) yielded only a minor amount of **1442**, alongside other oligomers (Figure 1A, trace (6)). Notably, at this pH, both lysines and the terminal carboxylates of **1** ($pK_{a1} = 2.2$) and the carboxylates of **4** are partially or fully protonated (for the structurally related thiosalicylic acid the $pK_a = 3.5$,⁵¹ whereas for terephthalic acid $pK_{a1} = 3.5$ and $pK_{a2} = 4.8$ ⁵²), thus possible oligomers are positively charged at this pH. We conclude that **1442** is formed preferentially under conditions where its net charge is zero and its self-assembly is not hampered by charge repulsion.

Next, we investigated the building block sequence of **1442**. Three regioisomers can potentially form: *ortho*, *meta* and *para*, similar to disubstituted benzene derivatives (Figure 2A, inset). UPLC analysis of **1442** only revealed a single peak under all chromatographic conditions tried, which suggests that a single regioisomer dominates (although we cannot rigorously exclude persistent co-elution of different regioisomers). Microscopy (TEM, AFM) indicated the presence of large nanoscale aggregates in DCLs containing the mixed hexamer (*vide infra*), precluding analysis by NMR spectrometry due to slow relaxation dynamics. ESI-

Orbitrap tandem mass spectrometry of the selected parent ion $[144_2]^{3+}$ ($m/z = 1083$ Da) at low collision voltage (15V) gave rise to several fragments (Figure 2A, bottom). Most importantly, tetramer fragment $[14]^{3+}$ ($m/z = 931$ Da) was observed, besides $[144_1]^{3+}$ ($m/z = 1007$ Da), and $[124_1]^{2+}$ ($m/z = 812$ Da). Whereas the two latter can form from any of the three regioisomers, $[14]^{3+}$ can only originate from the *ortho* isomer, featuring four contiguous 1-units (see Figure S60 for details). At higher collision voltages we observed water loss as well as the formation of b_5 fragments (i.e. loss of the C-terminal lysine) of one peptide unit within the macrocycle (Figure S61), both of which are well-described phenomena in gas-phase fragmentation of serine-containing peptide cations.⁵³ Notably, the presence of non-specific fragments does not allow us to exclude the initial presence of the other positional isomers. However, the ion intensity of the diagnostic fragment **14** is the highest compared to that of **1341** and **1242** fragments (Figure S62), suggesting that the *ortho* isomer is formed as the main product.

Next, we investigated the supramolecular organization of **1442**. Cryo-TEM and AFM analysis of a two days old sample (containing exclusively **1442**) indicated the presence of a network of nanoscale fibers (Figure 2B,C). The length of these fibers extended to the micrometer range (5 -15 μm) and they were several hundred nanometers wide, as indicated by TEM (Figure S65F-J). AFM and negative staining TEM (Figure S65) showed that the fibers further assembled into thick bundles of diverse morphologies (linear, curled and circular). In contrast, fibers assembled from peptide-only replicators (i.e. **18**, see Figure S64) are significantly shorter (0.5-5 μm) and only 2-5 nm wide.^{47,50,54,55} CD spectroscopy (Figure 2D) and thioflavin T (ThT) fluorescence assays (Figure 2E) both indicated a (partial) β -sheet structure, similar to those of peptide fibers.⁴⁷ A comparison of the spectroscopic signatures of octamer **18** and mixed hexamer **1442** reveals a noticeable difference: whereas the intensity of the CD spectra are comparable, that of the ThT assay is markedly lower for **1442** (Figure S63). This observation is in agreement with the lateral association of the mixed hexamer fibers: whereas the intercalation of ThT molecules between the macrocycles is more hindered for the dense fiber bundles of **1442** than for the more accessible fibers of **18**, no such difference holds for the interaction with circularly polarized light.

Figure 3. Comparison of the Nanoscale Structure and Evolution of Hexamer 1_{442} and Octamer 1_8 with SAXS.

(A) Temporal evolution of the SAXS profile of 1_{442} between 1-11 days.

(B) Temporal evolution of the SAXS profile of 1_8 between 1-17 days.

(C) Comparison of the SAXS profiles of 1_{442} and 1_8 after 1 day.

(D) Comparison of the SAXS profiles of 1_{442} and 1_8 after 9 days.

In order to obtain insight into the supramolecular structure of the two macrocycles in the native (solution) state, we monitored their self-assembly process with SAXS. Two DCLs were prepared from dithiol monomers under identical conditions (50 mM phosphate buffer, pH = 8.2, $[1] + [4] = 2.0$ mM, 1200 rpm stirring rate); the first composed of both building blocks ($[1]:[4] = 2:1$) whereas the second one was composed only of **1**. The temporal evolution of the two systems showed remarkable differences: The SAXS profile of the DCL leading to the formation of 1_{442} (Figure 3A), showed a characteristic q^{-1} dependence of the intensity in the $0.1 \text{ nm}^{-1} < q^{-1} < 1 \text{ nm}^{-1}$ region, indicative of the formation of fibers with a diameter of ca. 2.0 nm (close to the theoretically calculated value of 1.8 nm, Figure S69) already after 1 day (Figure 3C). In line with this observation, the UPLC peak area of hexamer (corresponding to the molecular composition) and the characteristic CD intensity (corresponding to the supramolecular chirality) increased simultaneously (Figure 4A) in this period, indicating that, as soon as hexamers are formed, they self-assemble into fibers. Moreover, after 7 days, three characteristic signals (Figure 3A, indicated with asterisks) were observed. The relative intensities of these signals increased in time, indicating progressive higher-order association of the fibers. The peaks were centered at $d = 5.8$, 3.4 and 2.9 nm, corresponding to a $1:\sqrt{3}:2$ ratio in the q -space, which indicates a hexagonal arrangement of the fibers (Figure 3D). In contrast, the second DCL resulting in the formation of 1_8 (Figure 3B) does not feature any fibrous aggregate after 1 day (Figure 3C). However, after 4 days, a broad peak centered at $d =$

4.8 nm is observed, most possibly corresponding to the pairwise association of fibers. The relative intensity of this signal increases over time until 11 days, concomitantly with the formation of **1₈**, as indicated by UPLC analysis. These results show that the self-assembly of **1₄4₂** occurs in two steps: fibers are formed first (concomitantly with the evolution of the macrocycle at the molecular level), which later self-assemble into a hexagonal bundle. In contrast, **1₈** self-assembles in one step (parallelly to the evolution of the macrocycle), giving rise to pairwise fibers.

Notably, monitoring the structural evolution of **1₄4₂** with negative staining TEM did not reveal a clear trend, as fiber bundles of various morphology were observed throughout an observation time of period of 2 hours up to 8 days (Figure S65), although the bundles became more regular after 3 days (Figure S65I). This discrepancy can be a consequence of drying effects during TEM sample preparation, which are often disregarded but could be consequential, especially for samples which are highly prone to secondary association.⁵⁶

Figure 4. Time evolution of the formation of the mixed hexamer **1₄4₂.**

(A) Time evolution of an agitated (stir rate set at 1200 rpm) DCL composed of **1** and **4** ([**1**]:[**4**] = 2:1, [**1**] + [**4**] = 2.0 mM), showing concomitant increase of the amount of **1₄4₂** molecules and the supramolecular CD signal.

(B) Concentration profile as a function of time in a non-agitated DCL composed of **1** and **4** ([**1**]:[**4**] = 2:1, [**1**] + [**4**] = 2.0 mM).

(C) Seeding experiment in a non-agitated DCL composed of **1** and **4** ([**1**]:[**4**] = 2:1, [**1**] + [**4**] = 50 μ M) comparing a non-seeded and a seeded (15 mol% pre-formed and sheared **1₄4₂**) sample.

Given the remarkably high self-assembly propensity of the mixed hexamer, we finally investigated the kinetics of its formation and compared it to that of the one-component peptide replicators (e.g. **1₈**). The latter are only formed under mechanical agitation due to the aforementioned breakage-growth mechanism and their evolution features an initial lag phase, followed by an exponential growth phase, both taking several days. In sharp contrast, the time evolution of the mixed hexamer **1₄4₂** from its building blocks upon agitation with the same stir rate used for the one-component peptide replicators (1200 rpm) featured a much shorter lag phase of ca. 1 hour (Figure 4A). Even more surprisingly, a similar kinetic profile was observed for an unagitated sample, showing that the formation of **1₄4₂** is not mechanosensitive (Figure 4B, C). Furthermore, also monomer oxidation is two orders of magnitude faster than for the peptide-only replicators (3-6 hours vs. 4-5 days, Figure 4B). These characteristics only change slightly by decreasing the stir rate to 400 rpm (Figure S70B) or the total building block concentration to 50 μ M (Figure 4C, black line). These results indicate that the formation of **1₄4₂** from its building blocks and its self-assembly proceeds with a low kinetic barrier, and most likely follows a different mechanism than that of **1₈** replicator.

Furthermore, **1****8** can form autocatalytically, whereas **1****4****2** shows only weak autocatalytic activity. In a non-agitated, partially oxidized DCL of **1** octamer formation is very slow, whereas seeding of this sample with a small amount of pre-formed **1****8** results in fast octamer formation.⁴⁷ In contrast, in a slightly oxidized, non-agitated mixture of **1** and **4** ($[1] + [4] = 2.0$ mM, $[1]:[4] = 2:1$), fast autonomous conversion (70 mol% hexamer in 50 minutes) into **1****4****2** was observed (Figure S72). This result showed that the non-catalyzed reaction is very fast leaving little room for autocatalysis. Additionally, TEM showed (Figure S66) that in sheared hexamer seeds mostly the degree of lateral association decreases, while fiber length is essentially retained. Thus, it proved difficult to produce a seed with a large number of fiber ends (which are the sites of potential autocatalysis). A somewhat increased fluorescence intensity of ThT was observed (Figure S68) for the sheared seeds, suggesting better accessibility of the hexamer molecules for dye intercalation. Decreasing the oxidation level, i.e. using a monomer mixture as the starting point for seeding experiment, led to partial reduction of the seed (Figure 4C, Figure S71). Only a weak autocatalytic effect was observed at very low concentrations (50 μ M), using an extensively sheared seed sample (Figure 4C); however, this effect was poorly reproducible (Figure S71). Thus, we conclude that the very fast formation of **1****4****2** and the resulting tightly bound and not readily fragmentable assemblies hamper autocatalysis, as most of the hexamers are buried within the fiber bundles and only a small fraction is available at the fiber ends.

Structure and Assembly of the Mixed Tetramer **1****1****4****3**

In contrast to the mixed hexamer, no nanoscale assemblies were detected in samples containing mixed tetramer **1****1****4****3**. However, its efficient (80 % by UPLC) and stoichiometry-specific formation suggests that non-covalent interactions stabilize this particular oligomer. As no nanoscale objects were detected by TEM (Figure S74) and kinetic analysis did not provide any indication for autocatalysis (Figure S73), it seemed more likely that a discrete assembly was formed. Thus, **1****1****4****3** was isolated and characterized using NMR, ESI-MS and SDS-PAGE. Nearly all protons of **1** and **1****1****4****3** (Figure 5B) were assigned following an extensive series of TOCSY, NOESY and HSQC experiments (Supplemental Information, Section 2.5.2). The most striking feature in the NMR spectrum of **1****1****4****3** (Figure 5D) compared to that of **1** (Figure 5C) was the large upfield shift of the lysine side chain protons K3 β -K3 ϵ ($\Delta\delta = -2.9$ ppm for K3 γ -K3 ϵ) and the smaller but also significant upfield shifts for K3 α and K3_{NH}. Furthermore, NOE cross-peaks were detected between the protons of the aromatic macrocycle and the aforementioned side chain protons (Figure 5E), indicating that these protons are in close spatial proximity. These observations indicate that the lysine side chain of K3 is threaded through cavity of the macrocycle. The threaded structure is further supported by the observed splitting of the K3 β -K3 γ protons, indicating that the two protons belonging to the same CH₂ group experience a different chemical environment, as their rotation is restricted in the confined space of the cavity (Figure 5D). Moreover, in contrast to **1**, the lysine ammonium protons (K3_{NH3}, K5_{NH3}) are detectable in the proton spectrum of **1****1****4****3**, suggesting that they are involved in strong hydrogen bonding (presumably with the carboxylates of **4**). Finally, the protons belonging to amino acids L2 and S4 show a significant downfield shift (up to $\Delta\delta = +1.5$ ppm for L2_{NH}), most possibly due to the aromatic ring current on the exterior of the neighboring threaded macrocycle. Similar features were observed in the NMR spectrum of pseudorotaxane **4****4**Sp,

which is stabilized by hydrogen bonding between ammonium protons and the macrocyclic carboxylates, as well as by the threading of nonpolar alkyl chains through the hydrophobic aromatic cavity.⁴⁶ Given the structural similarity between **Sp** and the lysine side chains, the proposed threaded structure seems plausible.

The presence of only two lysine K ϵ protons (a threaded and a non-threaded one) in **1143** suggests a (mixture of) cyclic [c η]daisy chain(s) (Figure 5A)⁵⁷ or lasso [1]pseudorotaxane⁵⁸ structure, rather than a (mixture of) acyclic (linear) [a η] daisy chain(s),⁵⁹ as the latter would give rise to multiple K ϵ proton signals. DOSY-NMR showed the presence of a single species with $D = 2.1 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ (Figure 5F).

Figure 5. Structure Elucidation and Self-Assembly of [c3]-(1₁4₃)₃.

- (A) Schematic structure of acyclic ([a_n]) and cyclic ([c_n]) daisy chains from (1₁4₃)₃.
(B) Proton assignment of [c3]-(1₁4₃)₃, highlighting the K3 lysine side chain threaded through the macrocycle
(C) ¹H-NMR spectrum of peptide building block **1** (500 MHz, H₂O/D₂O 9:1, phosphate buffer, pH = 7.2, 278 K). * = dioxane (internal standard); ** = TCEP and TCEP-oxide
(D) ¹H-NMR spectrum of [c3]-(1₁4₃)₃ (600 MHz, H₂O/D₂O 9:1, phosphate buffer, pH = 7.2, 278 K), highlighting large upfield shifts of the protons of the K3 side chain and considerable downfield shifts in the L2 and S4 amide protons.
(E) ¹H-¹H NOESY spectrum of [c3]-(1₁4₃)₃, highlighting NOE cross-peaks between the K3 side chain and the aromatic protons of the macrocycle.
(F) DOSY spectrum of [c3]-(1₁4₃)₃, confirming the presence of one species with a diffusion constant $D = 2.1 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$.
(G) SDS-PAGE analysis of [c3]-(1₁4₃)₃, showing the presence of one species, with a mass slightly higher than 3.5 kDa. Lanes: **1**: ultra-low weight molecular ladder, **2**: [c3]-(1₁4₃)₃ in phosphate buffer (50 mM, pH = 8.2), **3**: [c3]-(1₁4₃)₃ in ammonium acetate buffer (50 mM, pH = 7.0). BPB = Bromophenol Blue ($m = 669 \text{ Da}$)
(H) Negative mode ESI-Orbitrap spectrum of **1**₁4₃ in ammonium acetate buffer (50 mM, pH = 7.0), showing the presence of daisy chain [c3]-(1₁4₃)₃ as well as dimer fragment (1₁4₃)₂ and monomer 1₁4₃.

In order to determine the exact value of n (i.e. the number of macrocycles in one daisy chain unit), native SDS-PAGE experiments (both in phosphate and ammonium acetate buffer) were performed (Figure 5G). These indicated the presence of one species, with a molecular mass slightly bigger than 3.5 kDa, which most possibly corresponds to the [c3]- (4142 Da) rather than the [c2]- (2764 Da) or the [c4]- (5528 Da) species. Finally, native Orbitrap-ESI-MS in ammonium acetate buffer (Figure 5H) showed the presence of the [c3] species ($[(1_14_3)_3]^{4-}$, 1035 Da; $[(1_14_3)_3]^{3-}$, 690 Da) alongside with dimer ($[(1_14_3)_2]^{3-}$, 920 Da) and monomer ($[1_14_3]^-$, 690 Da) fragments, but no higher oligomers (i.e. $[(1_14_3)_4]^{5-}$, 1840 Da). The extensive fragmentation may be due to the absence of stabilization through hydrophobic interactions in the gas phase (i.e. in the absence of aqueous solvent). We conclude that **1**₁4₃ selectively assembles into a [c3] daisy chain as shown in Scheme 1.

Relative Stabilities of the Self-Assembling Species in the (1+4+Sp)-System

In view of the substantially different self-assembly modes of the three preferred macrocycles in the system, we were particularly interested in their relative stabilities and the kinetic barrier of their interconversion. We thus prepared samples of **1**, **1**₁4₂, **1**₁4₃ as well as of spermine complex **4**₄Sp and let them react with **1**, **4** or **Sp** (2.0 mM total building block concentration pH = 8.2, 50 mM phosphate buffer, stir rate 1200 rpm; reactions were conducted at 20 mol% thiol content in order to keep the systems dynamic) so that conversion to another preferred macrocycle could potentially occur (see overview on Scheme 2A). The building block stoichiometry in these experiments was chosen to correspond to that of a 1:1 mixture of the initial and the intended product macrocycles. Hence, any substantial deviation from a 1:1 molar ratio of products would indicate a high kinetic barrier and/or a large Gibbs free energy change of interconversion. The results of this series of experiments is shown in Figure 6 (see Figure S88 for the UPLC chromatograms).

Figure 6. Comparison of the Relative Stabilities of Self-Assembling Macrocycles 1_8 , 1_44_2 , 1_14_3 and 4_4Sp by Comparing Their Expected (red bars) and Observed (blue bars) Conversion in Various Interconversion Processes Starting from the Macrocycles Shown as White Bars.

- (A) The reaction between 1_8 and 4 (0.67 eq.) affords 1_44_2 in lower yield (26%) than expected (42 %).
- (B) The reaction between 1_8 and 4 (2.7 eq., 33 mol% thiolate) fails to produce 1_14_3 and affords 1_44_2 instead.
- (C) The reaction between 1_44_2 and 4 (2.0 eq) affords 1_14_3 in considerably lower yield (8 %) than expected (37 %).
- (D) The reaction between 1_44_2 and Sp (32 eq.) fails to produce 1_8 .
- (E) The reaction between 1_14_3 and Sp (0.5 eq) affords 1_44_2 and 4_4Sp .
- (F) The reaction between 1_14_3 and 1 (2.0 eq) affords 1_44_2 in slightly lower yield (46 %) than expected (63 %).
- (G) The reaction between 1_14_3 and Sp (0.75 eq., stoichiometric) fails to produce 1_8 and affords 1_44_2 instead.
- (H) The reaction between 4_4Sp and 1 (0.57 eq) fails to produce 1_14_3 and affords 1_44_2 instead.
- (I) The reaction between 4_4Sp and 1 (2.7 eq) affords 1_44_2 in the expected yield.

First, 1_44_2 can be obtained from any other macrocycle, i.e. in reactions from $1_8 + 4$ (Figure 6A), $1_14_3 + \text{Sp}$ (Figure 6E), $1_14_3 + 1$ (Figure 6F) and $4_4\text{Sp} + 1$ (Figure 6I), although in somewhat lower conversions than expected (due to side reactions, i.e. conversion to small oligomers). These results indicate a relatively low barrier of interconversion and a substantial negative Gibbs energy change for these reactions. Second, no reaction was

observed between **144₂** and even very high amounts of **Sp** (32 eq), although the dissociation constant of the putative product **44Sp** is in the nanomolar range (Figure 6D) and **Sp** has recently been reported to be able to sequester **4** from 2-component DCLs, in which secondary interactions between **4** and the other building block moiety are moderate.⁶⁰ We speculate that the multivalent nature of the fiber-fiber ion-pairing interaction may render **144₂** exceptionally stable, while the tight packing of the fibers might bring about a high activation barrier for the interconversion of **144₂** into **44Sp**.

Third, most intriguingly, even reactions which were predicted to give rise to other macrocycles, afforded **144₂** (to the extent stoichiometry allowed) at the cost of the expected product. For example, attempting to convert half an equivalent of **18** to **114₃** gives rise to **144₂** instead, alongside with unreacted **18** (Figure 6B) and the same trend was observed in the reverse reaction between **114₃** and **Sp** to give **18** (Figure 6G). Similarly, attempts to obtain **114₃** in the reaction between **44Sp** and **1** afforded **144₂** as well as unreacted **44Sp** (Figure 6H). Previously, simulations on thermodynamically governed DCLs have shown that if DCL members do not interact, then the formation of a large number of smaller macrocycles is preferred to the formation of a smaller number of larger ones.^{40,43,61} In this case, however, this trend does not hold due to the exceptionally high stability of **144₂**.

Based on these results, a free energy landscape shown on Scheme 2B is suggested. According to our results, **144₂** represents the deepest well in the Gibbs energy landscape. Macrocyclic **18** is stabilized to a lesser extent and can be converted to other species, although with lower yields, while **114₃** and **44Sp** are the least stable and can be easily converted to other species with multiple interaction sites. Interestingly, despite **144₂** having the lowest Gibbs energy, it is still possible for **114₃** to become the dominant library member, despite the fact that it is made from the same building blocks. This perhaps counterintuitive effect is a consequence of the fact that at high mole% of **4** the system prefers to use **1** for producing **114₃** instead of using it for **144₂**. For every four units of **1** (and with sufficient supply of **4**) it can harness four units of **114₃** against one unit of **144₂**. Even though individually **144₂** may be more efficient at forming non-covalent interactions than **114₃**, apparently it cannot compete against the interactions formed by four units of **114₃**.

Scheme 2. Systems Chemistry of 1, 4 and Sp

(A) Overview of the interconversion reactions probed in Figure 6.

(B) Potential energy surface of the system composed of 1, 4 and Sp (at 100 % oxidation level).

Substrate Scope

Finally, we investigated how structural modification of the building blocks affects the described systems behavior. For this purpose, we prepared DCLs in 2:1 and 1:3 molar ratios, either substituting **1** for peptide building blocks **2**, **3** or **5** (containing an alanine, threonine and phenylalanine moiety in place of serine, respectively) or **4** for regioisomer **6** (Scheme 1 and Figure 7). The selective formation of mixed hexamers (2_44_2 and 3_44_2) as well as of mixed tetramers (2_14_3 and 3_14_3) was observed for building blocks **2** (Figure 7A) and **3** (Figure 7B), respectively, showing that small structural modifications near to the guest binding site (lysine side chain) do not affect the self-assembly propensity (Figure S89-S91). In contrast, **5** and **4** form hexamer 5_44_2 with only poor selectivity, but are capable of forming 5_14_3 with high selectivity at the right stoichiometries (Figure 7C). This discrepancy suggests that the formation of the mixed hexamer is more susceptible to structural changes (phenylalanine instead of serine) than that of the mixed tetramer. We speculate that this difference reflects the fact that in the fiber bundles the building blocks are brought into close proximity, which is spatially more demanding than the much more open arrangement of building blocks in the daisy chains (whose formation is supposedly the driving force also in the case for the other peptide analogues). Upon substituting **4** for **6**, no selective formation of any specific macrocycle was observed upon mixing with **1** (Figure 7D), although building blocks **4** and **6** only differ in the position of one carboxylate group. Instead a complex mixture of different ring sizes and compositions was obtained. Apparently, the position of the carboxylate binding site has a pronounced effect on the self-assembly propensity. Notably, the regioisomers of several mixed macrocycles (1_26_n , $n = 3-6$) could be resolved (Figure 7D, in red) applying the same UPLC gradient as used for the **1**+**4** system. As the polarity of **4** and **6** is expected to be similar, this finding supports the notion that, in the case of 1_44_2 , indeed only one regioisomer was formed.

Figure 7. Effect of Building Block Structure on Systems Behavior.

(A) UPLC traces of DCLs prepared from **2** and **4** ($[2] + [4] = 2.0$ mM): $[2]:[4] = 2:1$ (top), $[2]:[4] = 1:3$ (bottom).
 (B) UPLC traces of DCLs prepared from **3** and **4** ($[3] + [4] = 2.0$ mM): $[3]:[4] = 2:1$ (top), $[3]:[4] = 1:3$ (bottom).
 (C) UPLC traces of DCLs prepared from **5** and **4** ($[5] + [4] = 2.0$ mM): $[5]:[4] = 2:1$ (top), $[5]:[4] = 1:3$ (bottom).
 (D) UPLC traces of DCLs prepared from **1** and **6** ($[1] + [6] = 2.0$ mM): $[1]:[6] = 2:1$ (top), $[1]:[6] = 1:3$ (bottom).

DISCUSSION

We showed that incorporating complementary molecular recognition units into separate building blocks **1** and **4** brings about the spontaneous and stoichiometry-selective formation of three oligomers. Each oligomer possesses a different ring size and self-assembles into markedly different nanoscale assemblies, while competing for common building blocks. The system can be directed to specific assemblies solely through building block stoichiometry, overriding the differences in thermodynamic stabilities of these assemblies. Specifically, **1₈** self-assembles into fibers via hydrophobic self-stacking. Mixed hexamer **1₄4₂** also self-assembles into fibers which subsequently arrange into hexagonally packed bundles. The remarkably high efficiency of this assembly process precludes significant autocatalysis, which essential to the formation of **1₈**. Finally, if **4** is in excess, daisy chain pseudorotaxane $[c\text{-}\mathbf{3}]\text{-(}\mathbf{1}_1\mathbf{4}_3)_3$ is obtained. Most importantly, both mixed macrocycles are stabilized by secondary interactions between the

complementary binding motifs (i.e. hydrophobic interactions as well as ion pairing between the lysine ammonium side chains from one building block and the carboxylates from the other).

The remarkably stoichiometry- and sequence- specific formation of **1₄4₂** can be attributed to several factors. First, the formation of a neutral species renders a 2:1 stoichiometry favorable as macrocycles with this building block ratio can potentially benefit maximally from attractive electrostatic interactions and allow for charge repulsion to be minimized. These characteristics make assembly by **1₄4₂** more facile than assembly of **1₈**, which is accompanied by charge repulsion. Second, we know from the structure of **4₄** that, if two or more carboxylates are in close proximity, they can engage in hydrogen bonds and ion pairing interactions with protons of a primary ammonium ion from **1**.⁴⁶ We speculate that this arrangement renders the formation of hexamer **1₄4₂** to be thermodynamically more favorable than that of the competing trimer **1₂4₁** (not observed in significant quantities, but also satisfying the 2:1 stoichiometry and entropically less demanding as it has a smaller ring size), as the latter has no vicinal **4**-moieties. Arrangements featuring **4**-dyads in **1₄4₂** in close proximity to each other along the fiber axis create a microenvironment with four carboxylates, which can bind lysine ammonium groups protruding from other fibers (Scheme 1, middle), resulting in strong lateral fiber association. Thus, in contrast to other nucleation-elongation processes^{62,63} where recognition sites required for self-assembly are only available at the fiber ends, **1₄4₂** can self-assemble along the fiber surface as well, further lowering the barrier for self-assembly. However, most of the macrocycles are buried in the tightly packed bundles of long fibers, thus only a relatively low number of catalytic fiber ends are available, explaining limited autocatalytic activity. Similar reduction of autocatalytic activity upon lateral association has been observed previously.⁶⁴

Using the appropriate building block ratio we can efficiently access a [c3]-daisy-chain assembly of **1₁4₃**. Molecular daisy chains^{57,65} are candidates for biomimicking motors, such as molecular muscles.^{66,67} However, their development lags behind those of other molecular motors, in part due to their demanding synthesis. Moreover, reports about daisy chains in water are scarce.⁶⁸ Our results show that a [c3] daisy chain can form spontaneously in high yield in water through a process of self-synthesis, where non-covalent interactions direct covalent bond formation in a subcomponent self-assembly process.⁶⁹ We attribute the high selectivity of formation of **1₁4₃** to the fact that this is the smallest macrocycle that provides a binding pocket for an alkyl ammonium group (requiring at least four building blocks) while offering a matching number of alkyl ammonium groups at the same time.

Most importantly, our results demonstrate that stoichiometry can be a powerful tool to navigate Gibbs energy landscapes of supramolecular systems, enabling the remarkably selective formation of specific assemblies even where these compete with other assemblies that are, when compared one-to-one, of higher thermodynamic stability. That stoichiometry can override thermodynamic stability is a systems effect: the system as a whole can achieve more non-covalent interactions by producing a larger number of individually less stable assemblies as opposed to a smaller number of individually more stable assemblies. Making use of such systems effects expands the scope of self-assembly beyond only populating

thermodynamic minima, but without having to resort to kinetically controlled steps.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 91 Supplementary Figures and 1 Supplementary Scheme and can be found online with this article.

ACKNOWLEDGMENTS

We are grateful for ERC (AdG ToDL 741774), the NWO and the Dutch Ministry of Education, Culture and Science (Gravitation program 024.001.035) for financial support. Bartosz Matysiak, Dr. Gaël Schaeffer and Dr. Charalampos G. Pappas are gratefully acknowledged for fruitful discussions.

AUTHOR CONTRIBUTIONS

D. K. and S. O. conceived the project. D. K., S. O., I. M., J. K. and G. P. designed the experiments. D.K. prepared the libraries, performed UPLC, LC-MS and spectroscopy analyses. T. T-W. performed Orbitrap ESI-MS analyses. J. K. and P. v. d. M. performed NMR measurements on **1** and **1143**. D. K., J.K. and B.L. analyzed and interpreted NMR data. D. K. and B. L. isolated **1143**. V. S. synthesized building block **6**. J. O. performed negative staining TEM measurements, G. M. S. performed cryo-TEM measurements, I. M. performed AFM measurements, F. S. A. performed SDS-PAGE analysis, G. P. and A. S. performed SAXS analysis, A. S. performed UV-VIS analysis. D. K. and S. O. wrote the manuscript, with inputs from J. K. and G. P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES AND NOTES

1. Steed, J.W., and Atwood, J.L. (2009). *Supramolecular chemistry* (Wiley).
2. Jain, A., and George, S.J. (2015). New directions in supramolecular electronics. *Mater. Today* **18**, 206–214.
3. Zarra, S., Wood, D.M., Roberts, D.A., and Nitschke, J.R. (2015). Molecular containers in complex chemical systems. *Chem. Soc. Rev.* **44**, 419–432.
4. Singh, N., Tena-Solsona, M., Miravet, J.F., and Escuder, B. (2015). Towards Supramolecular Catalysis with Small Self-assembled Peptides. *Isr. J. Chem.* **55**, 711–723.
5. Cook, T.R., and Stang, P.J. (2015). Recent Developments in the Preparation and Chemistry of Metallacycles and Metallacages via Coordination. *Chem. Rev.* **115**, 7001–7045.
6. Zhang, Q., and Tiefenbacher, K. (2015). Terpene cyclization catalysed inside a self-assembled cavity. *Nat. Chem.* **7**, 197–202.
7. Bruns, C.J., and Stoddart, J.F. (2016). *The Nature of the Mechanical Bond* (Wiley).
8. Thota, B.N.S., Urner, L.H., and Haag, R. (2016). Supramolecular Architectures of Dendritic Amphiphiles in Water. *Chem. Rev.* **116**, 2079–2102.
9. Wang, J., Liu, K., Xing, R., and Yan, X. (2016). Peptide self-assembly: thermodynamics and kinetics. *Chem. Soc. Rev.* **45**, 5589–5604.
10. Seeman, N.C., and Sleiman, H.F. (2017). DNA nanotechnology. *Nat. Rev. Mater.* **3**, 17068.
11. Lampel, A., Ulijn, R. V., and Tuttle, T. (2018). Guiding principles for peptide nanotechnology through directed discovery. *Chem. Soc. Rev.* **47**, 3737–3758.
12. Adelizzi, B., Van Zee, N.J., De Windt, L.N.J., Palmans, A.R.A., and Meijer, E.W. (2019). Future of Supramolecular Copolymers Unveiled by Reflecting on Covalent Copolymerization. *J. Am. Chem. Soc.*

- 147, 6110–6121.
13. Mako, T.L., Racicot, J.M., and Levine, M. (2019). Supramolecular Luminescent Sensors. *Chem. Rev.* **119**, 322–477.
14. Korevaar, P.A., George, S.J., Markvoort, A.J., Smulders, M.M.J., Hilbers, P.A.J., Schenning, A.P.H.J., De Greef, T.F.A., and Meijer, E.W. (2012). Pathway complexity in supramolecular polymerization. *Nature* **481**, 492–496.
15. Ogi, S., Sugiyasu, K., Manna, S., Samitsu, S., and Takeuchi, M. (2014). Living supramolecular polymerization realized through a biomimetic approach. *Nat. Chem.* **6**, 188–195.
16. Kang, J., Miyajima, D., Mori, T., Inoue, Y., Itoh, Y., and Aida, T. (2015). Noncovalent assembly. A rational strategy for the realization of chain-growth supramolecular polymerization. *Science* **347**, 646–51.
17. Cheng, C., McGonigal, P.R., Schneebeil, S.T., Li, H., Vermeulen, N.A., Ke, C., and Stoddart, J.F. (2015). An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553.
18. Tantakitti, F., Boekhoven, J., Wang, X., Kazantsev, R. V., Yu, T., Li, J., Zhuang, E., Zandi, R., Ortony, J.H., Newcomb, C.J., *et al.* (2016). Energy landscapes and functions of supramolecular systems. *Nat. Mater.* **15**, 469–476.
19. van Rossum, S.A.P., Tena-Solsona, M., van Esch, J.H., Eelkema, R., and Boekhoven, J. (2017). Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535.
20. Wehner, M., Röhr, M.I.S., Bühler, M., Stepanenko, V., Wagner, W., and Würthner, F. (2019). Supramolecular Polymorphism in One-Dimensional Self-Assembly by Kinetic Pathway Control. *J. Am. Chem. Soc.*
21. Jin, Y., Yu, C., Denman, R.J., and Zhang, W. (2013). Recent advances in dynamic covalent chemistry. *Chem. Soc. Rev.* **42**, 6634–6654.
22. Zhang, W., and Jin, Y. (2018). *Dynamic Covalent Chemistry: Principles, Reactions, and Applications*. W. Zhang and Y. Jin, eds. (Wiley).
23. Li, J., Nowak, P., and Otto, S. (2013). Dynamic combinatorial libraries: from exploring molecular recognition to systems chemistry. *J. Am. Chem. Soc.* **135**, 9222–9239.
24. Teichert, J.F., Mazunin, D., and Bode, J.W. (2013). Chemical sensing of polyols with shapeshifting boronic acids as a self-contained sensor array. *J. Am. Chem. Soc.* **135**, 11314–11321.
25. Lafuente, M., Solà, J., and Alfonso, I. (2018). A Dynamic Chemical Network for Cystinuria Diagnosis. *Angew. Chem. Int. Ed.* **57**, 8421–8424.
26. Komáromy, D., Nowak, P., and Otto, S. (2018). *Dynamic Combinatorial Libraries. In Dynamic Covalent Chemistry - Principles, Reactions and Applications* (Chichester, UK: John Wiley & Sons, Ltd), pp. 31–119.
27. Liu, B., Pappas, C.G., Zangrando, E., Demitri, N., Chmielewski, P.J., and Otto, S. (2019). Complex Molecules That Fold Like Proteins Can Emerge Spontaneously. *J. Am. Chem. Soc.* **141**, 1685–1689.
28. Ponnuswamy, N., Cougnon, F.B.L., Clough, J.M., Pantoş, G.D., and Sanders, J.K.M. (2012). Discovery of an organic trefoil knot. *Science* **338**, 783–785.
29. Ayme, J.-F., Beves, J.E., Leigh, D.A., McBurney, R.T., Rissanen, K., and Schultz, D. (2012). A synthetic molecular pentafoil knot. *Nat. Chem.* **4**, 15–20.
30. Ponnuswamy, N., Cougnon, F.B.L., Pantoş, G.D., and Sanders, J.K.M. (2014). Homochiral and meso figure eight knots and a Solomon link. *J. Am. Chem. Soc.* **136**, 8243–8251.
31. Li, J., Nowak, P., Fanlo-Virgós, H., and Otto, S. (2014). Catenanes from catenanes: quantitative assessment of cooperativity in dynamic combinatorial catenation. *Chem. Sci.* **5**, 4968–4974.
32. Wood, C.S., Ronson, T.K., Belenguer, A.M., Holstein, J.J., and Nitschke, J.R. (2015). Two-stage directed self-assembly of a cyclic [3]catenane. *Nat. Chem.* **7**, 354–358.
33. Minkenberg, C.B., Florusse, L., Eelkema, R., Koper, G.J.M., and van Esch, J.H. (2009). Triggered self-assembly of simple dynamic covalent surfactants. *J. Am. Chem. Soc.* **131**, 11274–11275.
34. Minkenberg, C.B., Li, F., van Rijn, P., Florusse, L., Boekhoven, J., Stuart, M.C.A., Koper, G.J.M., Eelkema, R., and van Esch, J.H. (2011). Responsive Vesicles from Dynamic Covalent Surfactants. *Angew. Chem. Int. Ed.* **50**, 3421–3424.
35. Wang, G., Wang, C., Wang, Z., and Zhang, X. (2012). H-Shaped Supra-Amphiphiles Based on a Dynamic Covalent Bond. *Langmuir* **28**, 14567–14572.
36. Carnall, J.M.A., Waudby, C.A., Belenguer, A.M., Stuart, M.C.A., Peyralans, J.J.-P., and Otto, S. (2010). Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502–1506.
37. Moulin, E., and Giuseppone, N. (2012). Dynamic combinatorial self-replicating systems. *Top. Curr. Chem.* **322**, 87–105.
38. Colomb-Delsuc, M., Mattia, E., Sadownik, J.W., and Otto, S. (2015). Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Commun.* **6**, 7427.
39. Kosikova, T., and Philp, D. (2017). Exploring the emergence of complexity using synthetic replicators. *Chem. Soc. Rev.* **46**, 7274–7305.
40. Severin, K. (2004). The Advantage of Being Virtual—Target-Induced Adaptation and Selection in Dynamic Combinatorial Libraries. *Chem. Eur. J.* **10**, 2565–2580.
41. Corbett, P.T., Otto, S., and Sanders, J.K.M. (2004). Correlation between host-guest binding and host amplification in simulated dynamic combinatorial libraries. *Chem. Eur. J.* **10**, 3139–3143.
42. Corbett, P.T., Tong, L.H., Sanders, J.K.M., and Otto, S. (2005). Diastereoselective amplification of an induced-fit receptor from a dynamic combinatorial library. *J. Am. Chem. Soc.* **127**, 8902–8903.
43. Saur, I., and Severin, K. (2005). Selection experiments with dynamic combinatorial libraries: the importance of the target concentration. *Chem. Commun.* **41**, 1471–1473.
44. Garcia-Seisdedos, H., Empereur-Mot, C., Elad, N., and Levy, E.D. (2017). Proteins evolve on the edge of supramolecular self-assembly. *Nature* **548**, 244–247.
45. Simon, A.J., Zhou, Y., Ramasubramani, V., Glaser, J., Pothukuchy, A., Gollihar, J., Gerberich, J.C., Leggere, J.C., Morrow, B.R., Jung, C., *et al.* (2019). Supercharging enables organized assembly of synthetic biomolecules. *Nat. Chem.* **11**, 204–212.
46. Vial, L., Ludlow, R.F., Leclaire, J., Pérez-Fernandez, R., and Otto, S. (2006). Controlling the biological effects of spermine using a synthetic receptor. *J. Am. Chem. Soc.* **128**, 10253–10257.

47. Malakoutikhah, M., Peyralans, J.J.-P., Colomb-Delsuc, M., Fanlo-Virgós, H., Stuart, M.C.A., and Otto, S. (2013). Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *J. Am. Chem. Soc.* **135**, 18406–18417.
48. Safont-Sempere, M.M., Fernández, G., and Würthner, F. (2011). Self-Sorting Phenomena in Complex Supramolecular Systems. *Chem. Rev.* **111**, 5784–5814.
49. Leonetti, G., and Otto, S. (2015). Solvent Composition Dictates Emergence in Dynamic Molecular Networks Containing Competing Replicators. *J. Am. Chem. Soc.* **137**, 2067–2072.
50. Frederix, P.W.J.M., Idé, J., Altay, Y., Schaeffer, G., Surin, M., Beljonne, D., Bondarenko, A.S., Jansen, T.L.C., Otto, S., and Marrink, S.J. (2017). Structural and Spectroscopic Properties of Assemblies of Self-Replicating Peptide Macrocycles. *ACS Nano* **11**, 7858–7868.
51. Zhang, G., Ma, N., Jiang, B., Shi, F., and Tu, S.J. (2010). Thiosalicylic acid catalyzed multicomponent reactions: Microwave-assisted synthesis of new extended angular fused azaheterocycles. *Synthesis (Stuttg.)*, 3993–3998.
52. Threeprom, J., Som-Aum, W., and Lin, J.-M. (2007). Separation of the positional isomers of phthalic acids in hydroorganic solvents using capillary electrophoresis. *J. Anal. Chem.* **62**, 1126–1130.
53. Harrison, A.G. (2012). Pathways for Water Loss from Doubly Protonated Peptides Containing Serine or Threonine. *J. Am. Soc. Mass Spectrom* **23**, 116–123.
54. Pal, A., Malakoutikhah, M., Leonetti, G., Tezcan, M., Colomb-Delsuc, M., Nguyen, V.D., van der Gucht, J., and Otto, S. (2015). Controlling the Structure and Length of Self-Synthesizing Supramolecular Polymers through Nucleated Growth and Disassembly. *Angew. Chem. Int. Ed.* **54**, 7852–7856.
55. Altay, M., Altay, Y., and Otto, S. (2018). Parasitic Behavior of Self-Replicating Molecules. *Angew. Chem. Int. Ed.* **57**, 10564–10568.
56. Franken, L.E., Boekema, E.J., and Stuart, M.C.A. (2017). Transmission Electron Microscopy as a Tool for the Characterization of Soft Materials: Application and Interpretation. *Adv. Sci.* **4**, 1600476.
57. Rotzler, J., and Mayor, M. (2013). Molecular daisy chains. *Chem. Soc. Rev.* **42**, 44–62.
58. Saito, F., and Bode, J.W. (2017). Synthesis and stabilities of peptide-based [1]rotaxanes: molecular grafting onto lasso peptide scaffolds. *Chem. Sci.* **8**, 2878–2884.
59. Iwaso, K., Takashima, Y., and Harada, A. (2016). Fast response dry-type artificial molecular muscles with [c2]daisy chains. *Nat. Chem.* **8**, 625–632.
60. Komáromy, D., Tezcan, M., Schaeffer, G., Marić, I., and Otto, S. (2017). Effector-Triggered Self-Replication in Coupled Subsystems. *Angew. Chem. Int. Ed.* **56**, 14658–14662.
61. Corbett, P.T., Sanders, J.K.M., and Otto, S. (2005). Competition between Receptors in Dynamic Combinatorial Libraries: Amplification of the Fittest? *J. Am. Chem. Soc.* **127**, 9390–9392.
62. Zhao, D., and Moore, J.S. (2003). Nucleation–elongation: a mechanism for cooperative supramolecular polymerization. *Org. Biomol. Chem.* **1**, 3471–3491.
63. Morris, A.M., Watzky, M.A., and Finke, R.G. (2009). Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**, 375–397.
64. Rubinov, B., Wagner, N., Matmor, M., Regev, O., Ashkenasy, N., and Ashkenasy, G. (2012). Transient Fibril Structures Facilitating Nonenzymatic Self-Replication. *ACS Nano* **6**, 7893–7901.
65. Ashton, P.R., Parsons, I.W., Raymo, F.M., Stoddart, J.F., White, A.J.P., Williams, D.J., and Wolf, R. (1998). Self-Assembling Supramolecular Daisy Chains. *Angew. Chem. Int. Ed.* **37**, 1913–1916.
66. Romuald, C., Ardá, A., Clavel, C., Jiménez-Barbero, J., and Coutrot, F. (2012). Tightening or loosening a pH-sensitive double-lasso molecular machine readily synthesized from an ends-activated [c2]daisy chain. *Chem. Sci.* **3**, 1851.
67. Goujon, A., Mariani, G., Lang, T., Moulin, E., Rawiso, M., Buhler, E., and Giuseppone, N. (2017). Controlled Sol-Gel Transitions by Actuating Molecular Machine Based Supramolecular Polymers. *J. Am. Chem. Soc.* **139**, 4923–4928.
68. Aeschi, Y., Drayss-Orth, S., Valášek, M., Häussinger, D., and Mayor, M. (2019). Aqueous Assembly of Zwitterionic Daisy Chains. *Chem. Eur. J.* **25**, 285–295.
69. Nitschke, J.R. (2007). Construction, Substitution, and Sorting of Metallo-organic Structures via Subcomponent Self-Assembly. *Acc. Chem. Res.* **40**, 103–112.

Contents

1. Supplemental Procedures	3
1.1. General Remarks	3
1.2. Buffer Preparation	3
1.3. UPLC Analysis	3
1.4. Mass Spectroscopy	4
1.4.1. UPLC-MS Analysis	4
1.4.2. (HPLC)-Orbitrap-MS Analysis	4
1.5. UV-VIS spectroscopy	4
1.6. Transmission Electron Microscopy	5
1.6.1. Negative-staining Transmission Electron Microscopy	5
1.6.2. Cryogenic Transmission Electron Microscopy (cryo-TEM)	5
1.7. Atomic Force Microscopy	5
1.8. Fluorescence Measurements	5
1.9. CD Spectroscopy	5
1.10. NMR Analysis	5
1.11. SDS-PAGE Analysis	6
1.12. SAXS Analysis	6
1.13. Synthesis of building block 6	7
1.14. Isolation of macrocycle 1 ₁₄ 3	8
1.15. Procedure for the seeding experiments with 1 ₄₄ 2	9
1.16. Sample Preparation Procedures	10
1.16.1. Sample Preparation for NMR Analysis of Building Block 1	10
1.16.2. Sample Preparation for NMR Analysis of [c3]-(1 ₁₄ 3) ₃	10
1.16.3. Sample Preparation for Direct Injection ESI-Orbitrap-MS analysis of [c3]-(1 ₁₄ 3) ₃	10
1.17. Mutual interconversion experiments in the 1+4 system.	11
1.18. Estimation of the difference between measured relative UPLC peak areas and molar percentages	12
2. Supplemental Data Items	14
2.1. NMR spectra	14
2.2. LC-MS analyses of DCLs	16

2.2.1.	DCLs prepared from 1 and 4	16
2.2.2.	DCLs prepared from 1 and 4 (chromatography with FA).....	23
2.2.3.	DCLs prepared from 2 and 4	24
2.2.4.	DCLs prepared from 3 and 4	26
2.2.5.	DCLs prepared from 5 and 4	28
2.2.6.	DCLs prepared from 1 and 6	30
2.3.	Detailed Stoichiometric Study of DCLs Composed of 1 and 4	35
2.4.	Supplemental characterization of the structure and assembly of 1 ₄ 4 ₂	36
2.4.1.	Tandem mass spectroscopy of 1 ₄ 4 ₂	36
2.4.2.	Analysis of the assemblies of 1 ₈ and 1 ₄ 4 ₂ with spectroscopy and TEM	39
2.4.3.	Kinetic Analyses of The Formation of 1 ₄ 4 ₂	42
2.5.	Supplemental Characterization of the Structure and Assembly of [c3]-(1 ₁ 4 ₃) ₃ ...	43
2.5.1.	UPLC and TEM analyses of the formation of [c3]-(1 ₁ 4 ₃) ₃	43
2.5.2.	Full ¹ H-NMR Assignment of Building Block 1 of [c3]-(1 ₁ 4 ₃) ₃	45
2.5.3.	Full ¹ H-NMR Assignment of Daisy Chain [c3]-(1 ₁ 4 ₃) ₃	49
2.6.	UPLC Chromatograms from the Interconversion Experiments in the 1:4:Sp System.....	55
2.7.	Structural characterization of the peptide analogues of 1 ₁ 4 ₃ and 1 ₄ 4 ₂	56
3.	Supplemental References	57

1. Supplemental Procedures

1.1. General Remarks

All chemicals, unless otherwise stated, were purchased from Sigma-Aldrich and used as received. Acetonitrile (UPLC-MS grade), water (UPLC-MS grade) and trifluoroacetic acid (HPLC grade) were purchased from Biosolve BV. Reagent grade solvents were purchased from Macron and used as received. Anhydrous solvents used in synthesis were freshly collected from a dry solvent purification system prior to use. Microwave synthesis was performed on a CEM DISCOVER hybrid 12 microwave reactor. Flash column chromatography was performed on a Reveleris® X2 Flash Chromatography System (Grace Davison Discovery Sciences, Deerfield IL) on reverse phase (C18) silica cartridges. Peptides **1**,¹ **2**,¹ **3**² and **5**¹ were purchased from Cambridge Peptides Ltd.

1.2. Buffer Preparation

Phosphate buffer (50 mM, pH = 8.16) was prepared by dissolving Na₂HPO₄·7H₂O (6.43 g, 24.0 mmol) and NaH₂PO₄·H₂O (0.166 g, 1.20 mmol) in 500 mL doubly distilled water. The pH was adjusted either with 1 M NaOH or with solid NaH₂PO₄·H₂O (ca. 90 mg).

Important: The phosphate buffer was filtered through a Whatman 0.2 µm PTFE syringe filter (GE Healthcare) just prior to library preparation to prevent growth of bacteria. If this step is missed, proteases secreted by eventual bacteria can hydrolyze peptide bonds. The reaction products (smaller peptides) appear in the UPLC chromatogram as minor peaks accompanying the major peaks, as well as a series of unresolved peaks in the wash (20-23 min).

Ammonium acetate buffer (50 mM, pH = 8.2) was prepared by dissolving ammonium acetate (185 mg, 2.4 mmol) in doubly distilled water (40 mL) and adjusting the pH to 8.2 with 1M NH₃ solution (210 µL). The solution was filled up to 50 mL and stored at 5 °C.

1.3. UPLC Analysis

Generally, samples were diluted to 200 µM total building block concentration with water (UPLC grade) and 10 µL of this diluted sample was injected for UPLC measurements. For the seeding experiments, undiluted samples were injected on the UPLC column.

UPLC measurements were performed on a Waters Acquity H-class system equipped with a PDA detector. UPLC analyses, unless otherwise stated, were performed on an Aeris 1.7 µm XB-C18 (150 × 2.1 mm) column, purchased from Phenomenex, using UPLC-MS grade water (eluent A) and UPLC-MS grade acetonitrile (eluent B), containing 0.1 V/V % TFA as a modifier. A flow rate of 0.3 mL/min and a column temperature of 35 °C were applied. Spectra were recorded at a detection wavelength of 254 nm.

Method for the analysis of DCLs prepared from **1+4**, **2+4**, **3+4**, **1+6**:

t / min	% A
0	90
1	90
2	80
8	72
16	71
17	50
18	5
19	5
20	90
23	90

Note: The very shallow gradient between 8 and 16 minutes was necessary to resolve **14**₂ from accompanying minor peaks of smaller oligomers. However, the pumps of the instrument could not always reproduce this minute gradient change, resulting in somewhat varying retention times (as shown for example in **Figure S 59**). In case of doubt, the corresponding peaks were always additionally identified with LC-MS.

Method for the analysis of DCLs prepared from **5+4**:

t / min	% A
0	90
1	90
1.3	75
3	72
11	60
13	30
17	5
18	5
19	90
22	90

1.4. Mass Spectroscopy

1.4.1. UPLC-MS Analysis

UPLC-MS measurements in positive mode were 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 with the following ionization parameters: capillary voltage: 3 kV, sampling cone voltage: 20 V, extraction cone voltage: 4 V, source gas temperature: 120°C, desolvation gas temperature: 450°C, cone gas flow (nitrogen): 1 L/h, desolvation gas flow (nitrogen): 800 L/h.

UPLC-MS measurements in negative mode were performed using a Waters Acquity UPLC H-class system coupled to a Waters Xevo TQD. The mass spectrometer was operated in the negative electrospray ionization mode with the following ionization parameters: capillary voltage: 2.4 kV, sampling cone voltage: 30 V, extraction cone voltage: 3 V, source gas temperature: 150°C, desolvation gas temperature: 500°C, cone gas flow (nitrogen): 5 L/h, desolvation gas flow (nitrogen): 800 L/h.

1.4.2. (HPLC)-Orbitrap-MS Analysis

Orbitrap-MS analyses were carried out on an Orbitrap XL (Thermo Fisher Scientific Instrument) with ESI ionization source. HPLC-ESI-MS/MS studies on **1₄2** were carried out on a Waters XBridge Phenyl 3.5 μ m (150 $\times$ 3.0 mm) column, with HPLC-MS grade water (eluent A) and HPLC-MS grade acetonitrile (eluent B), containing 0.1 V/V % formic acid (FA) as a modifier. A flow rate of 0.3 mL/min and a column temperature of 35 °C were applied, with the following gradient:

t / min	% A
0	90
2	85
10	50
22	5
24	5
30	90
31	90

ESI-MS/MS studies on **1₄3** were carried out in FIA (flow injection analysis) mode, applying ionization conditions as mild as possible (flow rate: 3 μ L/min, ion spray voltage: -2 kV, capillary voltage: -30 V, source gas temperature: 125°C).

1.5. UV-VIS spectroscopy

UV-Vis measurements were performed with a Jasco V-650 UV spectrophotometer, using a Hellma Analytics quartz high precision cell with a path length of 10 mm. UV-Vis data was analyzed with Jasco Spectra Manager (version 2) and OriginPro8. The measurements were performed at ambient temperature. For determination of molar absorption coefficients, 2.5 mM solutions of **1** and **4** were prepared in phosphate buffer (50 mM, pH = 8.2) and then

diluted to different concentrations with the buffer. 1 mL of these solutions was used for measurements. Phosphate buffer (50 mM, pH = 8.2) was used as a reference solution. Absorbance was recorded at 254 nm. Measurements were performed three times for each sample and the average of three measurements was used for the calibration curve.

1.6. Transmission Electron Microscopy

1.6.1. Negative-staining Transmission Electron Microscopy

Samples were diluted to 200 μ M total building block concentration with UPLC grade water. A small drop (5 μ L) of sample was then deposited on a 400 mesh copper grid covered with a thin carbon film (supplied by Agar Scientific). After 60 seconds, the droplet was blotted on filter paper. The sample was then stained with a solution of 2% uranyl acetate (5 μ L) deposited on the grid, subsequently washed and blotted on filter paper after 30 seconds. The staining procedure was repeated a second time. The grids were observed in a Philips CM12 electron microscope operating at 120 kV. Images were recorded on a slow scan CCD camera.

1.6.2. Cryogenic Transmission Electron Microscopy (cryo-TEM)

Samples were diluted to 200 μ M total building block concentration with UPLC grade water. A 10 μ L drop of the sample was placed on a Quantifoil 3.5/1 holey carbon coated grid. Blotting and vitrification in ethane was done in a Vitrobot (FEI, Eindhoven, the Netherlands). The grids were observed in a Philips CM20 cryo-electron microscope operating at 120 keV with a Gatan model 626 cryo-stage. Images were recorded under low-dose conditions with a slow-scan CCD camera.

1.7. Atomic Force Microscopy

AFM samples were prepared by depositing 100 μ L of the sample (diluted to 10 μ M building block concentration with UPLC grade water) onto a clean mica surface (Grade V1, Van Loenen Instruments). Subsequently, the solvent was evaporated in a gentle stream of air in ca. 10 minutes. The surface was then washed with 100 μ L of UPLC grade water and blotted on a piece of paper twice and finally air-dried. The AFM measurements have been performed using a Bruker Multimode 8 instrument in Scan Asyst-Air imaging mode. Measurements were performed in air at room temperature. As a probe, a ScanAsyst Air (Bruker) silicon tip on a nitride cantilever was used with the following parameters: length: 115 μ m, width: 25 μ m, resonance frequency: 70 kHz, force constant: 0.4 N/m. The images were recorded with frequencies between 0.5 and 1.5 Hz and analyzed with NanoScope Analysis 1.50 software (Bruker Corporation, 2015).

1.8. Fluorescence Measurements

Fluorescence measurements were performed on a JASCO FP 6200 fluorimeter using quartz cuvettes with 1 cm path length. For the ThT measurements, a freshly prepared solution of thioflavin T (dissolved to 2.2 mM in 50 mM phosphate buffer at pH = 8.2 and filtered with a 0.2 μ m syringe filter) was diluted to 22 μ M with the same buffer. 450 μ L of this diluted solution was transferred to the cuvette, followed by the addition of 80 μ L of sample (diluted with phosphate buffer to 80 μ M total building block concentration just prior to the measurement). Spectra were recorded after an incubation time of 2 minutes. The excitation wavelength was set at 440 nm and spectra were recorded in the range of 480-700 nm with a slit width of 5 nm, using a cutoff filter at 480 nm. The blank spectrum (i.e. fluorescence of the corresponding dye in solvent only) was subtracted from each spectrum.

1.9. CD Spectroscopy

Samples were diluted to 250 μ M with water (UPLC grade). Measurements were performed on a JASCO J-715 CD spectrophotometer using a HELMA quartz cuvette with a path length of 1 mm. Spectra were recorded in continuous mode between 190-400 nm at room temperature at a scan rate of 125 nm/min with 2 nm step-interval in millidegrees. All recorded spectra were an average of 3 repeats with solvent spectrum subtracted from each spectrum.

1.10. NMR Analysis

NMR spectra for routine purposes (analysis of synthetic intermediates in the synthesis of building block **6**) were recorded on a Varian VXR400 spectrometer.

NMR analyses on peptide **1** and [c3]-(**1,4**)₃ were carried out on a Varian Inova 500 and a Bruker Avance Neo 600 MHz instrument, respectively. For all spectra, the chemical shift of dioxane was used as an internal reference ($\delta(^1\text{H})$ = 3.75 ppm, $\delta(^{13}\text{C})$ = 69.3 ppm, with respect to DSS). Analyses were carried out at 278 K with water suppression. For the NOESY, TOCSY, dqfCOSY and HSQC experiments, 32 scans were recorded; for the DOSY experiments, 16 scans were recorded. For the NOESY experiment, a mixing time of 300 ms was applied.

1.11. SDS-PAGE Analysis

SDS-PAGE was performed using a stacking polyacrylamide gel (4 % cross-linking), a resolving polyacrylamide gel (20 % cross-linking) and a running buffer (1x TGS) consisting of 25 mM of Tris-HCl, 192 mM glycine, and 0.1% w/v of SDS. The standard ladder (M3546-1VL Molecular Weight Maker Ultra-Low Range) was obtained from Sigma Aldrich. The solutions for analysis were prepared by mixing 15 μ L sample ([c3]-1,4₃ at a total building block concentration of 2.0 mM in 50 mM phosphate or ammonium acetate buffer) with 5 μ L of loading dye (62.5 mM Tris-HCl, pH 6.8, 40% glycerol, 0.01% bromophenol blue). The total amount of the resulting mixture was injected into the gel wells. In total 6 μ L of the standard ladder was loaded in a separate well. The gel was run at 10 mA constant current, until the bromophenol blue reached the bottom of the gel, then the gel was removed and stained with InstantBlue™ Coomassie (Expedeon, US).

1.12. SAXS Analysis

A sample-to-detector distance of 6 m was used together with an X-ray photon wavelength of 1.24 Å. The SAXS profiles were recorded on two different detectors, the one measuring in an observed q range of $10^{-1} \text{ nm}^{-1} \leq q \leq 1 \text{ nm}^{-1}$, the other in an observed q range of $1 \text{ nm}^{-1} \leq q \leq 10 \text{ nm}^{-1}$, where q is the magnitude of the scattering vector ($q = (4\pi/\lambda) \sin\theta$), λ is the x-ray wavelength and θ is half of the scattering angle.

SAXS images were recorded using a 2D multiwire gas filled detector with 512x512 pixel dimension and with 260 μm^2 pixel size. The 2D images were radially averaged in order to obtain the intensity $I(q)$ vs q profiles. The liquid samples were contained in 2 mm borosilicate capillaries. Standard data reduction procedures, i.e. subtraction of the empty capillary contribution, correction for the sample absorption, were applied. Further information on data processing can be found elsewhere.³

1.13. Synthesis of building block 6

Scheme S 1. Synthesis of building block 6.

2,6-Bis(dimethylthiocarbamoyloxy)terephthalic Acid Dimethyl Ester (S1).

2,6-Dihydroxyterephthalic acid dimethyl ester (5.0 g, 22 mmol) and DABCO (9.8 g, 88 mmol) were dissolved in dry DMA (50 mL) under a nitrogen atmosphere and cooled to 0 °C in an ice bath. Dimethylthiocarbamoyl chloride (10.8 g, 88 mmol) dissolved in dry DMA (25 mL) was added dropwise under nitrogen. The mixture was stirred overnight at room temperature. Afterwards the reaction was cooled to 0 °C in an ice bath and 50 mL of water was added slowly. The yellowish precipitate was filtered and washed extensively with water and dried to yield 8.19 g (20 mmol, 93% yield) of the final product.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ 7.69 (s, 2H), 3.91 (s, 3H), 3.78 (s, 3H), 3.43 (s, 6H), 3.36 (s, 6H).

$^{13}\text{C-NMR}$ (100 MHz, CDCl_3): δ 186.26, 164.76, 162.57, 152.52, 132.91, 125.00, 122.76, 52.66, 52.50, 43.35, 38.96.

ESI-HRMS: $[\text{M}+\text{H}]^+$ found: 401.0827, 402.0860, 404.0825 Da (expected: 401.0836, 402.0869, 403.0794, 404.0827 Da); $[\text{M}+\text{Na}]^+$ found: 423.0647 Da (expected: 423.0660 Da).

2,6-Bis(dimethylthiocarbamoylsulfanyl)terephthalic Acid Dimethyl Ester (S2).

Compound **S1** (250 mg, 0.62 mmol) was suspended in 5 mL of diphenyl ether in a 10 mL microwave reactor. The reaction was performed in a CEM DISCOVER hybrid 12 microwave reactor (300 W, 240 °C, hold time 30 min, high stirring rate). The reaction mixture was passed through a silica plug eluting first with hexane to remove diphenyl ether, followed by chloroform : methanol 9 : 1. The solvent was removed under reduced pressure and the crude product was purified by column chromatography eluting with CHCl_3 (LAB-SCAN analytical reagent, stabilizer EtOH c.a. 1%). Product **S2** (r.f.=0.3) was obtained in 93% yield (232 mg, 0.58 mmol).

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ 8.26 (s, 2H), 3.91 (s, 3H), 3.90 (s, 3H), 3.06-3.02 (overlapping peaks, 12H).

$^{13}\text{C-NMR}$ (100 MHz, CDCl_3): δ 166.87, 165.20, 164.77, 148.31, 139.72, 131.81, 127.74, 52.83, 52.75, 37.28.

ESI-HRMS: $[\text{M}+\text{H}]^+$ found: 401.0827, 402.0861, 404.0825 Da (expected: 401.0836, 402.0869, 403.0794, 404.0827); $[\text{M}+\text{Na}]^+$ found: 423.0651 Da (expected: 423.0660 Da).

Building block 6

Compound **S2** (200 mg, 0.47 mmol) was refluxed (108 °C, oil bath) under inert atmosphere in 2 mL of iodotrimethylsilane for 12 h. After this time, the reaction was cooled down to 0 °C and slowly diluted with 20 mL of Et_2O . Subsequently, 10 mL of a 10% KOH solution (in water) was added dropwise. The aqueous phase was separated and the organic phase was extracted with 2x10 mL of 10% KOH water solution. The three combined aqueous phases were pooled and slowly acidified to pH = 2 with a concentrated solution of HCl. The resulting mixture was extracted with 4x20 mL of EtOAc. The combined organic phases were subsequently dried over MgSO_4 . The solvent was removed under vacuo to yield a dark yellow solid. After the solid was completely dried, it was reacted under strictly inert atmosphere with a 2 M KOH solution (degassed, in water). The reaction mixture was refluxed for 3 hours then cooled down to 0 °C. The solution was titrated with concentrated HCl under an inert atmosphere until a yellow precipitate appeared. At this point the flask was opened and the precipitate was collected through centrifugation. The collection/centrifugation was repeated 3 times with 5% HCl solution (degassed). The precipitate was then dried under reduced pressure yielding 92 mg of building block **6** (0.40 mmol, 69% yield).

$^1\text{H-NMR}$ (400 MHz, Degassed D_2O +10% KOH): δ 7.16 (s, 2H).

$^{13}\text{C-NMR}$ (100 MHz, Degassed D_2O +10% KOH, lock on D_2O): δ 179.37, 176.32, 149.85, 139.74, 132.70, 217.93 ppm.

ESI-HRMS: $[\text{M}-\text{H}]^-$ found: 228.9638, 229.9667, 230.9591 Da (expected: 228.9624, 229.9657, 230.9581 Da)

1.14. Isolation of macrocycle **1₁₄3**

1 (4.90 mg, 5.28 μ mol) was dissolved in phosphate buffer (50 mM, pH = 8.16, 1056 μ L) and was added to a solution of **4** (4.96 mg, 21.8 μ mol) in phosphate buffer (4313 μ L) in a 20 mL scintillation vial. The solution was stirred at 1200 rpm until UPLC analysis indicated full conversion of the monomers and a yield of ca. 70 % of **1₁₄3** (ca. 2 days). The reaction mixture was neutralized carefully (pH paper) with ca. 30 μ L TFA. This solution was purified directly on a Grace Reveleris X2 instrument on a 12 g C18 column, at a flow rate of 30 mL/min, with water + 0.1 % TFA (eluent A) and acetonitrile + 0.1 % TFA (eluent B). Simultaneous detection with UV(230 nm) and ELSD was applied. The UV detection range was set between 0 – 0.1 A.U. due to low signal intensity. The following gradient was applied:

t / min	% A
0	100
2	80
22	50
32	5
37	5
38	100
41	100

The product eluted at 13-15 minutes (ca. 62 % A). After elution the product-containing fractions were analyzed with UPLC and the pure fractions were pooled and immediately freeze-dried. The product was obtained as a white fluffy solid (4.1 mg, 2.5 μ mol, 48 % yield).

1.15. Procedure for the seeding experiments with **144₂**

1. A solution of **144₂** (150 μ L in phosphate buffer, 2.0 mM total building block concentration) was sheared in a Couette cell (described previously⁵) with the following program:

t / min	Ω / rpm
5	500
5	1000
5	1500
5	2000
5	2500
5	3000
5	3500
30	4000
5	3500
5	3000
5	2500
5	2000
5	1500
5	1000

2. The program above was immediately repeated on the same sample without opening the Couette cell.
3. The sheared solution of **144₂** was removed from the Couette cell and 10 μ L of the resulting sample was added to UPLC grade water (390 μ L) in a 4 mL vial to give a 50 μ M seed solution.
4. Building block **1** (1.04 mg) was weighed in a 2 mL UPLC vial and dissolved in phosphate buffer (560 μ L) to give a 2.0 mM solution.
5. Building block **4** (1.09 mg) was weighed in a 4 mL vial and dissolved in phosphate buffer (2370 μ L) to give a 2.0 mM solution.
6. A solution of **1** (2.0 mM, 30 μ L) was added to UPLC grade water (1170 μ L) in a 4 mL vial to give a 50 μ M solution.
7. A solution of **4** (2.0 mM, 30 μ L) was added to UPLC grade water (1170 μ L) in a 4 mL vial to give a 50 μ M solution.
8. A solution of **1** (50 μ M, 1000 μ L) was mixed with 500 μ L of a solution of **4** (50 μ M) in a 2 mL UPLC vial and 225 μ L of phosphate buffer was added to give a blank sample.
9. The vial was placed into the UPLC and 10 μ L was immediately injected on the column.
10. New dilutions of building blocks **1** and **4** were prepared (50 μ M, 1200 μ L) as described above.
11. At the end of the first measurement of the blank, 1000 μ L of a solution of **1** (50 μ M) was mixed with 500 μ L of a 50 μ M solution of **4** (50 μ M) in a 2 mL UPLC vial and 225 μ L of a solution of **144₂** seed (50 μ M) was added to give a blank sample.
12. Measurements were performed on the blank and on the seeded sample in an alternating manner, resulting in sampling periods of 48 minutes (23 minutes run time + 1 minute delay) per sample.

1.16. Sample Preparation Procedures

1.16.1. Sample Preparation for NMR Analysis of Building Block 1

A mixture of 2000 μL phosphate buffer (50 mM, pH = 8.16), 2500 μL UPLC grade water and 500 μL D_2O was purged with nitrogen for 2 hours in a 15 mL falcon tube. Subsequently, TCEP (1.43 mg, 1 mM) was added to inhibit oxidation of the thiol during the acquisition time. This was followed by the addition of dioxane (1.1 μL , 1 mM) as an internal reference. **1**·2 TFA (2.78 mg, 3 μmol) was weighed in a UPLC vial and dissolved in 600 μL of the solvent mixture prepared. The solution was transferred into an NMR tube and immediately subjected to measurement.

1.16.2. Sample Preparation for NMR Analysis of [c3]-(**1**₄)₃

In a 20 mL scintillation vial, 2000 μL phosphate buffer (50 mM, pH = 8.16) was mixed with 2500 μL UPLC grade water and 500 μL D_2O , followed by the addition of 8.8 μL of a 1 V/V % dioxane solution (in UPLC grade water). 600 μL of this mixture was added to a sample of purified **1**₄ (4.1 mg) in a 15 mL falcon tube and briefly sonicated if dissolution was not complete. The solution was transferred into an NMR tube and immediately subjected to measurement.

1.16.3. Sample Preparation for Direct Injection ESI-Orbitrap-MS analysis of [c3]-(**1**₄)₃

Building block **1** (1.17 mg, 0.126 μmol) was dissolved in the ammonium acetate buffer (630 μL) described above. Similarly, building block **4** (1.05 mg, 4.56 μmol) was dissolved in the same buffer (2283 μL). From the latter solution, 2205 μL was added to the solution of **4** in a 4 mL scintillation vial and the DCL was stirred at room temperature at 1220 rpm. After 36h, UPLC analysis indicated complete oxidation of the monomers and **1**₄ as the main product. The resulting slightly turbid solution was centrifuged (4000 rpm, 12 min) at 0 °C. UPLC analysis of the supernatant showed that **1**₄ corresponds to 78% of the total peak area, besides a few minor products. To a portion of the supernatant (300 μL) methanol (100 μL , UPLC grade) was added and this solution was used for ESI-Orbitrap analysis.

1.17. Mutual interconversion experiments in the 1+4 system.

The following samples were prepared from freshly dissolved monomers in phosphate buffer: **1₈**, **1₄4₂**, **1₁4₃**, **4₄Sp**, fully oxidized **1** (**1(ox)**), fully oxidized **4** (**4(ox)**). Samples **1(ox)** and **4(ox)** contained mainly trimers and tetramers. The overall building block concentration was 2.0 mM. The stirring rate was 1200 rpm.

	1₈	1₄4₂	1₁4₃	4₄Sp	1 (ox)	4 (ox)
V (1 , 2 mM)	600 µL	400 µL	200 µL		420 µL	300 µL
V (4 , 2 mM)		200 µL	700 µL	800 µL		
V (Sp , 32 mM)				12.5 µL		
Stirring	yes	yes	yes	yes	no	no

After 7 days, the DCLs reached their desired composition and no monomers were detected, as judged by UPLC. The following samples were prepared (experiment assignment corresponds to that shown on **Figure 6** of the main text):

	Reaction	Component 1	Component 2	V(TCEP, 20 mM)
A	$3 \times \mathbf{1}_8 + 2 \times \mathbf{4} \rightarrow 2 \times \mathbf{1}_4\mathbf{4}_2 + 2 \times \mathbf{1}_8$	1₈ (240 µL)	4 (ox) (40 µL)	5.6 µL
B	$9 \times \mathbf{1}_8 + 24 \times \mathbf{4} \rightarrow 8 \times \mathbf{1}_1\mathbf{4}_3 + 8 \times \mathbf{1}_8$	1₈ (240 µL)	4 (ox) (80 µL)	6.4 µL
C	$5 \times \mathbf{1}_4\mathbf{4}_2 + 10 \times \mathbf{4} \rightarrow 4 \times \mathbf{1}_1\mathbf{4}_3 + 4 \times \mathbf{1}_4\mathbf{4}_2$	1₄4₂ (240 µL)	4 (ox) (80 µL)	6.4 µL
D	$2 \times \mathbf{1}_4\mathbf{4}_2 + \mathbf{Sp} \rightarrow \mathbf{1}_8 + \mathbf{4}_4\mathbf{Sp}$	1₄4₂ (240 µL)	Sp (40 µL)	4.8 µL
E	$10 \times \mathbf{1}_1\mathbf{4}_3 + 5 \times \mathbf{Sp} \rightarrow 2 \times \mathbf{1}_4\mathbf{4}_2 + 2 \times \mathbf{1}_1\mathbf{4}_3 + 5 \times \mathbf{4}_4\mathbf{Sp}$	1₁4₃ (240 µL)	Sp (1.5 µL)	4.8 µL
F	$5 \times \mathbf{1}_1\mathbf{4}_3 + 10 \times \mathbf{1} \rightarrow 3 \times \mathbf{1}_4\mathbf{4}_2 + 3 \times \mathbf{1}_1\mathbf{4}_3$	1₁4₃ (240 µL)	1 (ox) (120 µL)	7.2 µL
G	$8 \times \mathbf{1}_1\mathbf{4}_3 + 6 \times \mathbf{Sp} \rightarrow \mathbf{1}_8 + 6 \times \mathbf{4}_4\mathbf{Sp}$	1₁4₃ (240 µL)	Sp (42 µL)	6.4 µL
H	$7 \times \mathbf{4}_4\mathbf{Sp} + 4 \times \mathbf{1} \rightarrow 4 \times \mathbf{1}_1\mathbf{4}_3 + 4 \times \mathbf{4}_4\mathbf{Sp} + 4 \times \mathbf{Sp}$	4₄Sp (280 µL)	1 (ox) (40 µL)	6.4 µL
I	$3 \times \mathbf{4}_4\mathbf{Sp} + 8 \times \mathbf{1} \rightarrow 2 \times \mathbf{1}_4\mathbf{4}_2 + 2 \times \mathbf{4}_4\mathbf{Sp} + \mathbf{Sp}$	4₄Sp (280 µL)	1 (ox) (160 µL)	8.0 µL

The reactions are written in an unusual way, i.e. assuming that the starting materials and the products are formed in a 1:1 molar ratio (except for equations **D** and **G**, where excess spermine was added). TCEP was added to the fully oxidized libraries to reach 20% reduction level in order to ensure thiol-disulfide exchange.

The expected relative UPLC intensities were calculated as follows: First, the ratio of the molar absorption coefficients of the two building blocks at 254 nm (wavelength of the UPLC-UV measurements) with UV-VIS spectroscopy. The molar absorption coefficients were determined to be $1.758 \times 10^4 \text{ L mol}^{-1} \text{ cm}^{-1}$ for **1** and $1.342 \times 10^4 \text{ L mol}^{-1} \text{ cm}^{-1}$ for **4** (Figure S 1), giving a molar absorption coefficient ratio of $\frac{\epsilon(1)}{\epsilon(4)} = 1.31$.

Figure S 1. UV-VIS calibration curves for building blocks A) **1**, B) **4** at 254 nm.

Second, we assume that the molar absorption coefficients are additive, i.e. for a species with a general formula $\mathbf{1}_k\mathbf{4}_n$:

$$\varepsilon_{kn} = \varepsilon(\mathbf{1}_k\mathbf{4}_n) = k \times \varepsilon(\mathbf{1}) + n \times \varepsilon(\mathbf{4}) \quad (1)$$

Substituting the experimentally determined ratio of the molar absorption coefficients, one gets:

$$\varepsilon_{kn} = (1.31 \times k + n) \times \varepsilon(\mathbf{4}) \quad (2)$$

Third, the expected UPLC intensity (absorbance) ratios can be calculated from the expected molar amounts (i.e. stoichiometric coefficients) of the corresponding species, weighted with the corresponding absorption coefficients. For example, in case of reaction **E**:

$$\begin{aligned} I_{rel}(\mathbf{1}_4\mathbf{4}_2) &= \frac{(4 \times 1.31 + 2) \times n_{rel}(\mathbf{1}_4\mathbf{4}_2)}{(4 \times 1.31 + 2) \times n_{rel}(\mathbf{1}_4\mathbf{4}_2) + (1.31 + 3) \times n_{rel}(\mathbf{1}_1\mathbf{4}_3) + 4 \times n_{rel}(\mathbf{4}_4 \bullet \mathbf{Sp})} \\ &= \frac{7.24 \times 2}{7.24 \times 2 + 4.31 \times 2 + 4 \times 5} = 33.6 \% \\ I_{rel}(\mathbf{1}_1\mathbf{4}_3) &= \frac{(1.31 + 3) \times n_{rel}(\mathbf{1}_1\mathbf{4}_3)}{(4 \times 1.31 + 2) \times n_{rel}(\mathbf{1}_4\mathbf{4}_2) + (1.31 + 3) \times n_{rel}(\mathbf{1}_1\mathbf{4}_3) + 4 \times n_{rel}(\mathbf{4}_4 \bullet \mathbf{Sp})} \\ &= \frac{4.31 \times 2}{7.24 \times 2 + 4.31 \times 2 + 4 \times 5} = 20.0 \% \\ I_{rel}(\mathbf{4}_4 \bullet \mathbf{Sp}) &= \frac{4 \times n_{rel}(\mathbf{4}_4 \bullet \mathbf{Sp})}{(4 \times 1.31 + 2) \times n_{rel}(\mathbf{1}_4\mathbf{4}_2) + (1.31 + 3) \times n_{rel}(\mathbf{1}_1\mathbf{4}_3) + 4 \times n_{rel}(\mathbf{4}_4 \bullet \mathbf{Sp})} \\ &= \frac{4 \times 5}{7.24 \times 2 + 4.31 \times 2 + 4 \times 5} = 46.4 \% \end{aligned}$$

For libraries involving **Sp**, (Figure S 88E-G, I) we used FA (formic acid) instead of TFA in the UPLC analysis, as we observed that the $\mathbf{4}_4\mathbf{Sp}$ complex decomposes upon longer exposure (> 7 minutes) to TFA under the conditions used for chromatography, giving rise to unknown species.

1.18. Estimation of the difference between measured relative UPLC peak areas and molar percentages

As noted above, the molar absorption coefficients of the building blocks differ by ca. 30 %. We estimated the resulting discrepancy between relative UPLC peak areas and relative molar concentrations as follows:

For the species $1_k 4_n$ in a given DCL, let us denote the total building block concentration with c_0 , the molar concentration of the species with c_{kn} , the corresponding relative molar concentration (equal to the molar concentration of building blocks contained in the species) with c_{kn}^r , the corresponding measured intensity (i.e. UPLC peak area) with I_{kn} , the relative intensity with I_{kn}^r and the absorption coefficient of the species with ε_{kn} . As shown above,

$$\varepsilon_{kn} = (1.31 \times k + n) \times \varepsilon(4) = (1.31k + n)\varepsilon. \quad (1')$$

From here, the Lambert-Beer law gives (taking into account that the path length is the same during the measurement, thus it can be regarded as unity):

$$I_{kn} = (1.31k + n)\varepsilon c_{kn} \quad (3)$$

and

$$I_{kn}^r = \frac{(1.31k + n)\varepsilon c_{kn}}{\sum_{(k,n)} (1.31k + n)\varepsilon c_{kn}} = \frac{(1.31k + n)c_{kn}}{\sum_{(k,n)} (1.31k + n)c_{kn}} \quad (4)$$

Similarly, the absolute and relative molar concentration of $1_k 4_n$:

$$c_{kn} = (k + n)c_{kn} \quad (5)$$

and

$$c_{kn}^r = \frac{(k + n)c_{kn}}{\sum_{(k,n)} (k + n)c_{kn}} = \frac{(k + n)c_{kn}}{c_0} \quad (6)$$

From equations (4) and (6) the absolute and relative molar percentages of all species can be calculated by solving the corresponding system of equations. As, however, direct quantification is not important to support the main statements of the paper and given the relatively small difference between the corresponding coefficients, we introduce the following correction factor (f_{kn}) by comparing equations (4) and (6):

$$f_{kn} = \frac{\sum_{(k,n)} (1.31k + n) \frac{k + n}{1.31k + n}}{\sum_{(k,n)} (k + n)} \quad (7)$$

So that

$$c_{kn}^r \approx f_{kn} I_{kn}^r \quad (8)$$

The f_{kn} values for the measurements described are typically in the range of 0.93-1.05, i.e. the discrepancy between the relative UPLC peak areas and relative concentrations is at most $\pm 10\%$.

2. Supplemental Data Items

2.1. NMR spectra

Figure S 2. ^1H -NMR (400 MHz, CDCl_3 , bottom) and ^{13}C -NMR (101 MHz, CDCl_3 , top) spectra of **S1**.

Figure S 3. ^1H -NMR (400 MHz, CDCl_3 , bottom) and ^{13}C -NMR (101 MHz, CDCl_3 , top) spectra of **S2**.

Figure S 4. ^1H NMR (400 MHz, degassed D_2O +10% KOH, bottom) and ^{13}C NMR (101 MHz, degassed D_2O +10% KOH, top) spectra of building block 6.

2.2. LC-MS analyses of DCLs

2.2.1. DCLs prepared from 1 and 4

Figure S 5. Mass spectrum of **4** (t_R = 6.60 min) from the LC-MS analysis of a DCL made from **4**.
m/z calculated: 228.96 [M-H]⁻; m/z observed: 229.00 [M-H]⁻.

Figure S 6. Mass spectrum of **1** (t_R = 7.10 min) from the LC-MS analysis of a DCL made from **1** and **4**.
m/z calculated: 700.31 [M+H]⁺, 350.66 [M+2H]²⁺; m/z observed: 700.24 [M+H]⁺, 350.79 [M+2H]²⁺.

Figure S 7. Mass spectrum of **11341** and **11342** (t_R = 7.66 min) from LC-MS analysis of a DCL made from **1** and **4**.
11341: m/z calculated: 1549.80 [M+6H]⁶⁺, 1859.56 [M+5H]⁵⁺; m/z observed: 1549.78 [M+6H]⁶⁺, 1859.47 [M+5H]⁵⁺.
11342: m/z calculated: 1587.79 [M+6H]⁶⁺, 1905.15 [M+5H]⁵⁺; m/z observed: 1587.76 [M+6H]⁶⁺, 1905.07 [M+5H]⁵⁺.

Figure S 11. Mass spectrum of **19**, **1941**, **1942**, **110** ($t_R=8.61$ min) from LC-MS analysis of a DCL made from **1** and **4**.
19: m/z calculated: 1256.13 [M+5H]⁵⁺, 1569.92 [M+4H]⁴⁺; m/z observed: 1256.05 [M+5H]⁵⁺, 1569.91 [M+4H]⁴⁺.
1941: m/z calculated: 1301.73 [M+5H]⁵⁺, 1626.90 [M+4H]⁴⁺; m/z observed: 1301.64 [M+5H]⁵⁺, 1626.86 [M+4H]⁴⁺.
1942: m/z calculated: 1347.32 [M+5H]⁵⁺, 1683.89 [M+4H]⁴⁺; m/z observed: 1347.17 [M+5H]⁵⁺, 1683.79 [M+4H]⁴⁺.
110: m/z calculated: 1395.59 [M+5H]⁵⁺, 1744.24 [M+4H]⁴⁺; m/z observed: 1395.62 [M+5H]⁵⁺, 1744.12 [M+4H]⁴⁺.

Figure S 12. Mass spectrum of **1841** ($t_R=8.83$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
m/z calculated: 1162.27 [M+5H]⁵⁺, 1452.58 [M+4H]⁴⁺, 1936.44 [M+3H]³⁺;
m/z observed: 1162.23 [M+5H]⁵⁺, 1452.42 [M+4H]⁴⁺, 1936.31 [M+3H]³⁺.

Figure S 13. Mass spectrum of **18** ($t_R=9.20$ min) from the LC-MS analysis of a DCL made from **1**.
m/z calculated: 930.73 [M+6H]⁶⁺, 1116.68 [M+5H]⁵⁺, 1395.59 [M+4H]⁴⁺, 1860.45 [M+3H]³⁺;
m/z observed: 930.64 [M+6H]⁶⁺, 1116.67 [M+5H]⁵⁺, 1395.68 [M+4H]⁴⁺, 1860.35 [M+3H]³⁺.

Figure S 14. Mass spectrum of **14** ($t_R=9.30$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 698.30 $[M+4H]^{4+}$, 930.73 $[M+3H]^{3+}$, 1395.59 $[M+2H]^{2+}$;
 m/z observed: 698.23 $[M+4H]^{4+}$, 930.83 $[M+3H]^{3+}$, 1395.46 $[M+2H]^{2+}$.

Figure S 15. Mass spectrum of **1541** and **17** ($t_R=9.67$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
1541: m/z calculated: 929.61 $[M+4H]^{4+}$, 1239.14 $[M+3H]^{3+}$; m/z observed: 929.62 $[M+4H]^{4+}$, 1238.98 $[M+3H]^{3+}$.
17: m/z calculated: 1221.37 $[M+4H]^{4+}$, 1628.02 $[M+3H]^{3+}$; m/z observed: 1221.27 $[M+4H]^{4+}$, 1628.18 $[M+3H]^{3+}$.

Figure S 16. Mass spectrum of **16** ($t_R=10.08$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 1046.94 $[M+4H]^{4+}$, 1395.59 $[M+3H]^{3+}$; m/z observed: 1046.85 $[M+4H]^{4+}$, 1395.62 $[M+3H]^{3+}$.

Figure S 17. Mass spectrum of **15** ($t_R = 10.21$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 872.61 $[M+4H]^{4+}$, 1163.15 $[M+3H]^{3+}$; m/z observed: 872.67 $[M+4H]^{4+}$, 1163.06 $[M+3H]^{3+}$.

Figure S 18. Mass spectrum of **1241** ($t_R = 9.96$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 541.85 $[M+3H]^{3+}$, 812.27 $[M+2H]^{2+}$; m/z observed: 541.93 $[M+3H]^{3+}$, 812.18 $[M+2H]^{2+}$.

Figure S 19. Mass spectrum of **1441** ($t_R = 10.18$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 755.28 $[M+4H]^{4+}$, 1006.71 $[M+3H]^{3+}$, 1509.56 $[M+2H]^{2+}$;
 m/z observed: 755.24 $[M+4H]^{4+}$, 1006.81 $[M+3H]^{3+}$, 1509.45 $[M+2H]^{2+}$.
 Peaks at 541.9 and 812.2 Da arise from coelution with **1241** and/or fragmentation.

Figure S 20. Mass spectrum of **144₂** ($t_R=10.56$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 650.01 $[M+5H]^5+$, 812.27 $[M+4H]^4+$, 1082.69 $[M+3H]^3+$, 1623.54 $[M+2H]^2+$;
 m/z observed: 649.99 $[M+5H]^5+$, 812.37 $[M+4H]^4+$, 1082.68 $[M+3H]^3+$, 1623.49 $[M+2H]^2+$.

Figure S 21. Mass spectrum of **134** ($t_R=10.97$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 774.28 $[M+3H]^3+$, 1160.92 $[M+2H]^2+$; m/z observed: 774.17 $[M+3H]^3+$, 1160.99 $[M+2H]^2+$.

Figure S 22. Mass spectrum of **124₂** ($t_R=12.29$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 617.83 $[M+3H]^3+$, 926.25 $[M+2H]^2+$; m/z observed: 617.79 $[M+3H]^3+$, 925.95 $[M+2H]^2+$.

Figure S 23. Mass spectrum of **13** ($t_R = 12.44$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 523.98 $[M+4H]^{4+}$, 698.30 $[M+3H]^{3+}$, 1046.94 $[M+2H]^{2+}$;
 m/z observed: 524.02 $[M+4H]^{4+}$, 698.24 $[M+3H]^{3+}$, 1047.13 $[M+2H]^{2+}$.

Figure S 24. Mass spectrum of **143** ($t_R = 16.13$ min) from the LC-MS analysis of a DCL made from **1** and **4**.
 m/z calculated: 691.58 $[M+2H]^{2+}$, 1382.16 $[M+H]^+$; m/z observed: 691.48 $[M+2H]^{2+}$, 1382.45 $[M+H]^+$.

2.2.2. DCLs prepared from 1 and 4 (chromatography with FA)

Figure S 25. Mass spectrum of **142** ($t_R = 9.03$ min) from the LC-MS analysis of a DCL made from **1** and **4**.

m/z calculated: 541.85 $[M+6H]^{6+}$; 650.01 $[M+5H]^{5+}$, 812.27 $[M+4H]^{4+}$, 1082.69 $[M+3H]^{3+}$, 1623.54 $[M+2H]^{2+}$;
 m/z observed: 541.99 $[M+6H]^{6+}$; 649.99 $[M+5H]^{5+}$, 812.37 $[M+4H]^{4+}$, 1082.68 $[M+3H]^{3+}$, 1623.49 $[M+2H]^{2+}$.

Figure S 26. Mass spectrum of **143** ($t_R = 12.02$ min) from the LC-MS analysis of a DCL made from **1** and **4**.

m/z calculated: 691.58 $[M+2H]^{2+}$, 1382.16 $[M+H]^+$; m/z observed: 691.62 $[M+2H]^{2+}$, 1381.74 $[M+H]^+$.

Figure S 27. Mass spectrum of **4** ($t_R = 18.5$ min) from the LC-MS analysis of a DCL made from **Sp** and **4**.

m/z calculated: 911.44 $[M-H]^-$, 682.88 (fragment); m/z observed: 911.05 $[M-H]^-$, 683.05 (fragment).

2.2.3. DCLs prepared from 2 and 4

Figure S 28. Mass spectrum of **24** ($t_R=10.08$ min) from the LC-MS analysis of a DCL made from **2** and **4**.

m/z calculated: 682.29 $[M+4H]^{4+}$, 909.40 $[M+3H]^{3+}$, 1363.60 $[M+2H]^{2+}$.

m/z observed: 682.24 $[M+4H]^{4+}$, 909.50 $[M+3H]^{3+}$, 1363.53 $[M+2H]^{2+}$.

Figure S 29. Mass spectrum of **2442** ($t_R=11.62$ min) from the LC-MS analysis of a DCL made from **2** and **4**.

m/z calculated: 637.22 $[M+5H]^{5+}$, 796.27 $[M+4H]^{4+}$, 1061.37 $[M+3H]^{3+}$, 1591.55 $[M+2H]^{2+}$;

m/z observed: 637.22 $[M+5H]^{5+}$, 796.14 $[M+4H]^{4+}$, 1061.38 $[M+3H]^{3+}$, 1591.28 $[M+2H]^{2+}$.

Figure S 30. Mass spectrum of **23** ($t_R=13.37$ min) from the LC-MS analysis of a DCL made from **2** and **4**.

m/z calculated: 511.97 $[M+4H]^{4+}$, 682.29 $[M+3H]^{3+}$, 1022.95 $[M+2H]^{2+}$.

m/z observed: 512.03 $[M+4H]^{4+}$, 682.25 $[M+3H]^{3+}$, 1022.65 $[M+2H]^{2+}$.

Figure S 31. Mass spectrum of **2143** (t_R = 17.63 min) from the LC-MS analysis of a DCL made from **2** and **4**.
 m/z calculated: 683.58 [M+2H]²⁺, 1365.56 [M+H]⁺; m/z observed: 683.54 [M+2H]²⁺, 1365.58 [M+H]⁺.

2.2.4. DCLs prepared from 3 and 4

Figure S 32. Mass spectrum of **34** (t_R = 9.91 min) from the LC-MS analysis of a DCL made from **3** and **4**.

m/z calculated: 712.30 [M+4H]⁴⁺, 949.41 [M+3H]³⁺, 1423.62 [M+2H]²⁺.

m/z observed: 712.23 [M+4H]⁴⁺, 949.49 [M+3H]³⁺, 1423.45 [M+2H]²⁺.

Figure S 33. Mass spectrum of **3541** and **3542** (t_R = 10.18 min) from the LC-MS analysis of a DCL made from **3** and **4**.

3541: m/z calculated: 947.13 [M+4H]⁴⁺, 1262.50 [M+3H]³⁺; m/z observed: 947.13 [M+4H]⁴⁺, 1262.33 [M+3H]³⁺. **3542**: m/z calculated: 1004.11 [M+4H]⁴⁺, 1338.48 [M+3H]³⁺; m/z observed: 1004.07 [M+4H]⁴⁺, 1338.24 [M+3H]³⁺.

Figure S 34. Mass spectrum of **3241** (t_R = 10.72 min) from the LC-MS analysis of a DCL made from **3** and **4**.

m/z calculated: 551.19 [M+3H]³⁺, 826.29 [M+2H]²⁺; m/z observed: 551.23 [M+3H]³⁺, 826.13 [M+2H]²⁺.

Figure S 35. Mass spectrum of **3442** ($t_R=11.62$ min) from the LC-MS analysis of a DCL made from **3** and **4**.
 m/z calculated: 661.23 $[M+5H]^5+$, 826.29 $[M+4H]^4+$, 1101.38 $[M+3H]^3+$, 1651.57 $[M+2H]^2+$;
 m/z observed: 661.20 $[M+5H]^5+$, 826.13 $[M+4H]^4+$, 1101.34 $[M+3H]^3+$, 1651.72 $[M+2H]^2+$.

Figure S 36. Mass spectrum of **33** ($t_R=13.36$ min) from the LC-MS analysis of a DCL made from **3** and **4**.
 m/z calculated: 534.42 $[M+4H]^4+$, 712.30 $[M+3H]^3+$, 1067.96 $[M+2H]^2+$.
 m/z observed: 534.52 $[M+4H]^4+$, 712.23 $[M+3H]^3+$, 1067.63 $[M+2H]^2+$.

Figure S 37. Mass spectrum of **3143** ($t_R=17.23$ min) from the LC-MS analysis of a DCL made from **3** and **4**.
 m/z calculated: 698.59 $[M+2H]^2+$, 1395.58 $[M+H]^+$; m/z observed: 698.53 $[M+2H]^2+$, 1395.59 $[M+H]^+$.

2.2.5. DCLs prepared from 5 and 4

Figure S 38. Mass spectrum of **5₂4₁**, **5₄4₁**, **5₆4₁**, **5₇4₁** ($t_R=7.25$ min) from the LC-MS analysis of a DCL made from **5** and **4**.

5₂4₁: m/z calculated: 581.87 [M+3H]³⁺, 872.31 [M+2H]²⁺; m/z observed: 581.94 [M+3H]³⁺, 872.19 [M+2H]²⁺.

5₄4₁: m/z calculated: 815.32 [M+4H]⁴⁺, 1086.76 [M+3H]³⁺; m/z observed: 815.25 [M+4H]⁴⁺, 1086.83 [M+3H]³⁺.

5₆4₁: m/z calculated: 955.39 [M+5H]⁵⁺, 1193.98 [M+4H]⁴⁺, 1591.64 [M+3H]³⁺;

m/z observed: 955.41 [M+5H]⁵⁺, 1193.88 [M+4H]⁴⁺, 1591.66 [M+3H]³⁺.

5₇4₁: m/z calculated: 1106.85 [M+5H]⁵⁺, 1383.32 [M+4H]⁴⁺; m/z observed: 1106.76 [M+5H]⁵⁺, 1383.33 [M+4H]⁴⁺.

Figure S 39. Mass spectrum of **5₄** and **5₅4₁** ($t_R=7.37$ min) from the LC-MS analysis of a DCL made from **5** and **4**.

5₄: m/z calculated: 758.33 [M+4H]⁴⁺, 1010.77 [M+3H]³⁺, 1515.66 [M+2H]²⁺;

m/z observed: 758.29 [M+4H]⁴⁺, 1010.56 [M+3H]³⁺, 1515.56 [M+2H]²⁺.

5₅4₁: m/z calculated: 803.92 [M+5H]⁵⁺, 1004.65 [M+4H]⁴⁺, 1339.20 [M+3H]³⁺;

m/z observed: 803.84 [M+5H]⁵⁺, 1004.69 [M+4H]⁴⁺, 1339.41 [M+3H]³⁺.

Figure S 40. Mass spectrum of **5₃4₁** ($t_R=7.88$ min) from the LC-MS analysis of a DCL made from **5** and **4**.

m/z calculated: 500.98 [M+5H]⁵⁺, 625.98 [M+4H]⁴⁺, 834.31 [M+3H]³⁺, 1250.97 [M+2H]²⁺;

m/z observed: 501.09 [M+5H]⁵⁺, 626.02 [M+4H]⁴⁺, 834.22 [M+3H]³⁺, 1251.08 [M+2H]²⁺.

Figure S 41. Mass spectrum of **53** ($t_R = 8.44$ min) from the LC-MS analysis of a DCL made from **5** and **4**.
 m/z calculated: 455.40 $[M+5H]^{5+}$, 569.00 $[M+4H]^{4+}$, 758.33 $[M+3H]^{3+}$, 1136.99 $[M+2H]^{2+}$;
 m/z observed: 455.53 $[M+5H]^{5+}$, 569.06 $[M+4H]^{4+}$, 758.29 $[M+3H]^{3+}$, 1136.69 $[M+2H]^{2+}$.

Figure S 42. Mass spectrum of **542** ($t_R = 8.61$ min) from the LC-MS analysis of a DCL made from **5** and **4**.
 m/z calculated: 698.05 $[M+5H]^{5+}$, 872.31 $[M+4H]^{4+}$, 1162.74 $[M+3H]^{3+}$, 1743.61 $[M+2H]^{2+}$;
 m/z observed: 698.04 $[M+5H]^{5+}$, 872.19 $[M+4H]^{4+}$, 1162.74 $[M+3H]^{3+}$, 1743.82 $[M+2H]^{2+}$.

Figure S 43. Mass spectrum of **5143** ($t_R = 9.81$ min) from the LC-MS analysis of a DCL made from **5** and **4**.
 m/z calculated: 721.61 $[M+2H]^{2+}$, 1442.19 $[M+H]^+$; m/z observed: 721.57 $[M+2H]^{2+}$, 1442.63 $[M+H]^+$.

2.2.6. DCLs prepared from 1 and 6

Figure S 44. Mass spectrum of **126₁** (t_R =9.76 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 541.85 $[M+3H]^3+$, 812.27 $[M+2H]^2+$; m/z observed: 541.93 $[M+3H]^3+$, 812.19 $[M+2H]^2+$.

Figure S 45. Mass spectrum of **116₂** (t_R =9.43 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 577.60 $[M+3H]^3+$, 1154.20 $[M+H]^+$; m/z observed: 577.66 $[M+3H]^3+$, 1153.86 $[M+H]^+$.

Figure S 46. Mass spectrum of **136₁** (t_R =10.58 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 774.28 $[M+3H]^3+$, 1161.08 $[M+2H]^2+$; m/z observed: 774.22 $[M+3H]^3+$, 1161.08 $[M+2H]^2+$.

Figure S 47. Mass spectrum of **1262** ($t_R=10.75$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 617.83 $[M+3H]^3+$, 926.25 $[M+2H]^2+$; m/z observed: 617.86 $[M+3H]^3+$, 926.09 $[M+2H]^2+$.

Figure S 48. Mass spectrum of **1263** ($t_R=11.06$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 693.82 $[M+3H]^3+$, 1040.23 $[M+2H]^2+$; m/z observed: 693.80 $[M+3H]^3+$, 1039.97 $[M+2H]^2+$.

Figure S 49. Mass spectrum of **1263** and **1364** ($t_R=11.23$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
1263: m/z calculated : 693.82 $[M+3H]^3+$, 1040.23 $[M+2H]^2+$; m/z observed: 693.80 $[M+3H]^3+$, 1039.97 $[M+2H]^2+$.
1364: m/z calculated : 1002.23 $[M+3H]^3+$, 1502.85 $[M+2H]^2+$; m/z observed: 1002.34 $[M+3H]^3+$, 1502.73 $[M+2H]^2+$.

Figure S 50. Mass spectrum of **13**, **1262** and **1365** ($t_R=11.50$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
13: m/z calculated : 698.30 $[M+3H]^3+$, 1046.94 $[M+2H]^2+$; m/z observed: 698.28 $[M+3H]^3+$, 1046.68 $[M+2H]^2+$.
1262: m/z calculated: 617.83 $[M+3H]^3+$, 926.25 $[M+2H]^2+$; m/z observed: 617.87 $[M+3H]^3+$, 926.08 $[M+2H]^2+$.
1365: m/z calculated: 1078.22 $[M+3H]^3+$, 1616.83 $[M+2H]^2+$; m/z observed: 1078.27 $[M+3H]^3+$, 1616.57 $[M+2H]^2+$.

Figure S 51. Mass spectrum of **1264** ($t_R=11.73$, 11.97min) from the LC-MS analysis of a DCL made from **1** and **6**.
m/z calculated: 769.80 $[M+3H]^3+$, 1154.20 $[M+2H]^2+$; m/z observed: 769.73 $[M+3H]^3+$, 1154.36 $[M+2H]^2+$.

Figure S 52. Mass spectrum of **1365** ($t_R=12.17$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
m/z calculated: 1078.22 $[M+3H]^3+$, 1616.83 $[M+2H]^2+$; m/z observed: 1078.28 $[M+3H]^3+$, 1616.61 $[M+2H]^2+$.

Figure S 53. Mass spectrum of **1265** (t_R = 12.77 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 845.79 $[M+3H]^3+$, 1268.18 $[M+2H]^2+$; m/z observed: 845.67 $[M+3H]^3+$, 1268.26 $[M+2H]^2+$.

Figure S 54. Mass spectrum of **1263** (t_R = 13.19 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 845.79 $[M+3H]^3+$, 1268.18 $[M+2H]^2+$; m/z observed: 845.67 $[M+3H]^3+$, 1268.26 $[M+2H]^2+$.

Figure S 55. Mass spectrum of **1264** (t_R = 13.85 min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 769.80 $[M+3H]^3+$, 1154.20 $[M+2H]^2+$; m/z observed: 769.74 $[M+3H]^3+$, 1154.36 $[M+2H]^2+$.

Figure S 56. Mass spectrum of **1265** ($t_R=14.61$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
 m/z calculated: 845.79 $[M+3H]^3+$, 1268.18 $[M+2H]^2+$; m/z observed: 845.68 $[M+3H]^3+$, 1268.25 $[M+2H]^2+$.

Figure S 57. Mass spectrum of **1265**, **1266** ($t_R=15.30$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
1265: m/z calculated: 845.79 $[M+3H]^3+$, 1268.18 $[M+2H]^2+$; m/z observed: 845.67 $[M+3H]^3+$, 1268.25 $[M+2H]^2+$.
1266: m/z calculated: 921.77 $[M+3H]^3+$, 1382.16 $[M+2H]^2+$; m/z observed: 921.60 $[M+3H]^3+$, 1382.13 $[M+2H]^2+$.

Figure S 58. Mass spectrum of **1266** ($t_R=16.11$ min) from the LC-MS analysis of a DCL made from **1** and **6**.
1266: m/z calculated: 921.77 $[M+3H]^3+$, 1382.16 $[M+2H]^2+$; m/z observed: 921.60 $[M+3H]^3+$, 1382.15 $[M+2H]^2+$.

2.3. Detailed Stoichiometric Study of DCLs Composed of 1 and 4

Figure S 59. UPLC chromatograms of DCLs obtained from building blocks **1** and **4** ($[1] + [4] = 2.0 \text{ mM}$) after 7 days (stirred, 50 mM phosphate buffer at pH = 8.15) containing A) 0, B) 2, C) 4, D) 6, E) 8, F) 10, G) 12, H) 15, I) 20, J) 25, K) 30, L) 35, M) 40, N) 45, O) 50, P) 55, Q) 60, R) 65, S) 70, T) 75, U) 80, V) 85, W) 90, X) 95 mol% of **4**.

2.4. Supplemental characterization of the structure and assembly of $1_4 4_2$

2.4.1. Tandem mass spectroscopy of $1_4 4_2$

Figure S 60. Schematic fragmentation pattern of the three different positional isomers (top) of the $1_4 4_2$ macrocycle, yielding different pentamers (middle) and tetramers (bottom). The only fragment specific for any of the isomers is tetramer 1_4 (namely, for the *ortho* isomer).

Figure S 61. ESI-Orbitrap-MS² spectra of $[1442]^{3+}$ at A) 5 V B) 15 V C) 20 V D) 25 V collisional energy in the 300-1600 Da mass range and E) at 15 V in the 1010-1080 Da mass range. Increasing the collisional energy does not lead to fragmentation of the macrocycle core, rather to water loss and fragmentation of the peptide chains. The $[1442-K]$ and $[1442-SK]$ species denote *b* type fragments of the peptide chain.

Figure S 62. Excerpts from the ESI-Orbitrap-MS² spectrum of $[1442]^{3+}$ at 15 V collision energy showing the presence or absence (denoted with dotted lines) of the corresponding oligomer cations.

2.4.2. Analysis of the assemblies of 1_8 and 1_44_2 with spectroscopy and TEM

Figure S 63. Comparison of the circular dichroism (A) and thioflavin T fluorescence emission spectra (B) of the macrocycles 1_44_2 and 1_8 at 2.0 mM total building block concentration (50 mM phosphate buffer, pH = 8.16).

Figure S 64. Negative staining TEM images taken from a DCL mainly composed of octamer 1_8 .

Figure S 65. Negative staining TEM images taken from a DCL prepared by mixing a 2.0 mM solution of **1** (pre-oxidized to 40 % with sodium perborate) and a non-oxidized 2.0 mM solution of **4** (both in 50 mM phosphate buffer, pH = 8.16) in a 2:1 ratio, after A) 0 B) 2 C) 3 D) 4 E) 4.5 F) 21 G) 23 H) 48 I) 72 J) 144 K) 192 hours of stirring at 1200 rpm, featuring morphologically diverse assemblies composed of laterally associated fibers. Inset L) shows the time evolution of $^{144}\text{I}_2$ in the initial phase of the experiment.

Figure S 66. Negative staining TEM images taken from a 29 days old (A-B) or a 3 days old (C-D) DCL containing $^{144}\text{I}_2$ ([**1**]+[**4**] = 2.0 mM, 50 mM phosphate buffer, pH = 8.16) before (A,C) and after applying a shear rate (maximum 33700 s⁻¹) for 100 minutes (B) or 200 minutes (D) in a Couette cell.

Figure S 67. Cryogenic TEM micrographs of **1442** ($[1] + [4] = 2.0$ mM, $[1]:[4] = 2.0$, phosphate buffer, pH = 8.2), highlighting a dense network of fibers. Micrograph A) is the same as shown on **Figure 2B** of the main text.

Figure S 68. Thioflavin T fluorescence emission spectra of **1442** before (red trace) and after (blue trace) applying a shear rate (maximum 33700 s^{-1} for 60 minutes) for 200 minutes in a Couette cell (2.0 mM total building block concentration, 50 mM phosphate buffer, pH = 8.16).

Figure S 69. MM simulation of the mixed hexamer **1442**, resulting in a diameter of 1.8 nm.

2.4.3. Kinetic Analyses of The Formation of 1_44_2

Figure S 70. Time evolution of the mixed hexamer 1_44_2 (50 mM phosphate buffer, pH = 8.16) from stirred DCLs containing building blocks **1** and **4** in a 2:1 molar ratio at a total building block concentration of A) 2 mM (stirred at 1200 rpm) B) 50 μ M (stirred at 400 rpm).

Figure S 71. Seeding experiments in a non-agitated DCL composed of **1** and **4** ($[1]:[4] = 2:1$, $[1] + [4] = 50 \mu$ M) comparing non-seeded and seeded (15 mol% pre-formed and sheared 1_44_2) sample. These experiments show no notable difference in the rate of hexamer formation between seeded and non-seeded samples.

Figure S 72. UPLC chromatogram of a DCL prepared by fully oxidizing a non-stirred solution of **1** (2.0 mM) A) before B) 50 minutes after the addition of 0.5 equivalent of **4** (2.0 mM).

2.5. Supplemental Characterization of the Structure and Assembly of $[c3]-(1_1 4_3)_3$

2.5.1. UPLC and TEM analyses of the formation of $[c3]-(1_1 4_3)_3$

Figure S 73. Time evolution of the mixed tetramer the $1_1 4_3$: (50 mM phosphate buffer, pH = 8.16) from DCLs containing building blocks **1** and **4** in a 1:3.5 molar ratio at a total building block concentration of A) 2 mM B) 50 μ M. C) 50 μ M, without (black trace) and with (red trace) of the addition of 15 mol % $1_1 4_3$ seed.

Figure S 74. Negative staining TEM image of a sample containing mainly **114₃** (2.0 mM total building block concentration) provides no evidence for the presence of well-defined supramolecular assemblies.

2.5.2. Full ^1H -NMR Assignment of Building Block 1 of $[\text{c}3]-(1,4)_3$

In this section the assignment of all protons of peptide building block **1** is described in a stepwise manner. An overview of the final proton assignment is shown on **Figure S 75**.

Figure S 75. A) Labeling of relevant protons in peptide **1**. B) Full-range ^1H -NMR spectrum of **1**. Excerpts from the full spectrum highlighting C) amide and aromatic protons D) α -protons E) protons of the peptide side chains

First, the aromatic protons were assigned based on their unique shifts and integrals. Thus, the peaks at 7.31 ppm and 7.50 ppm were assigned as R1 and R2, respectively.

The amide peaks of the peptide chain were first identified based on their number of cross-peaks in the protons and the ^1H - ^1H TOCSY spectrum. As the protons belonging to one peptide chain are expected to give pairwise cross-peaks with each other, and the number of protons is characteristic for each amino acid residue (except for the two lysines) all but one amide protons can be identified based on the number of their TOCSY cross-peaks (**Figure S 76A**). For glycine and serine, 1 and 2 cross-peaks are expected, respectively. That is, the amide protons at 8.90 and 8.38 can be assigned to G1 NH and S4 NH , respectively. Furthermore, the peak at 3.8 ppm (2 H) can be assigned to the S4 β protons.

For leucine, 4 cross-peaks are expected, including a cross-peak with the diastereotopic L2 δ methyl groups (doublet, 6H). The latter only holds for the proton at 8.55 ppm, which can be assigned to L2 NH ; however, only 3 cross-peaks

are observed, due to overlap of L2 β and L2 γ . Finally, the signals at 8.43 and 8.15 ppm can be assigned to K3 NH and K5 NH (5 cross-peaks).

Figure S 76. Excerpts of the ^1H - ^1H TOCSY spectrum of **1** highlighting cross-peaks in the A) amide region (8-9 ppm) B) α -proton region (3.8-4.5 ppm).

The α -protons can also be partially identified based on the number of their TOCSY cross-peaks (**Figure S 76B**). Based on the assignment of the amide protons, the signals at 4.35 and 4.05 ppm can be assigned to L2 α (cross-peak with the methyl signal at 0.9 ppm) and G1 α , respectively. Based on the number of cross-peaks, the signal at 4.15 ppm can be assigned to K3 α or K5 α . The other lysine (K3 α /K5 α) and S4 α cannot be distinguished at this stage.

Figure S 77. A) ^1H - ^1H NOESY cross-correlations between amide and α -protons in **1**. B) Spectrum shown in A) enlarged between 8.05-8.60 ppm. C) Excerpt of the ^1H - ^{13}C HSQC spectrum of **1**, highlighting the cross-peaks corresponding to the α -protons.

Next, we examined the NOESY cross peaks of the amide and α -protons. Most notably, at this stage we could identify K5 α as the only α -proton showing exactly one cross-peak (with K5 NH). Thus, we identified the protons at

4.15 and 8.13 ppm as K5 α and K5 NH (**Figure S 77A**). Consequently, we could identify the protons at 8.43 ppm as K3 NH .

The protons K3 α and S4 α overlap at 4.37-4.38 ppm. However, they can be distinguished based on their ^1H - ^{13}C HSQC cross-peaks (**Figure S 77C**). The ^{13}C signal of S4 α is expected to be more downfield than that of K3 α , due to the proximity of the hydroxyl group. Thus, the signal at 57.8 ppm (4.37 ppm) corresponds to S4 α and the signal at 56.1 ppm (4.38 ppm) corresponds to K3 α . The consecutive NOESY cross-correlations ("NOESY-walk") of the amide and α -protons were consistent with the proton assignments so far (**Figure S 77A-B**).

Next, the peaks of the lysine side chains were identified. Most importantly, the ϵ -CH $_2$ groups could be distinguished based on the TOCSY cross-peaks with the corresponding α -protons. Thus, the peaks at 2.91 and 2.97 ppm were identified as K3 ϵ and K5 ϵ , respectively (**Figure S 78A** and inset).

Based on their chemical environment, the γ -protons are expected to be the most upfield-shifted, i.e. the signals at 1.39 and 1.41 ppm can be assigned to K5 γ and K3 γ , respectively.

The β -protons are expected to be the most downfield-shifted (1.7-1.9 ppm). Moreover, their proton peaks are expected to split as the protons are diastereotopic. Correspondingly, in the HSQC spectrum, the split K3 β and K5 β signals are expected to give a cross-peak with the same carbon signal (**Figure S 78B**). Thus, the signals at 1.78 and 1.87 ppm can be assigned to K3 β , whereas those at 1.72 and 1.82 ppm can be assigned to K5 β .

Consequently, the signals at 1.65 and 1.66 ppm can be assigned to K3 δ and K5 δ , respectively.

Figure S 78. A) Excerpt of the ^1H - ^1H TOCSY spectrum of 1, highlighting cross-peaks between amide (K3 NH , K5 NH) and side chain protons of the lysine moieties. Excerpts of the ^1H - ^{13}C HSQC spectrum of 1, highlighting cross-peaks corresponding to B) K3 β and K5 β methylene groups C) L2 β , L2 γ , K3 δ and K5 δ protons.

Figure S 79. Excerpt of the ^1H - ^{13}C HSQC spectrum of **1**, displaying cross-peaks of the carbon-bound protons of the peptide chain.

The rest of the protons (L2 β , L2 γ , K3 δ , K5 δ), which overlap between 1.57-1.70 ppm (**Figure S 78C**), were identified based on their HSQC cross-correlations (**Figure S 79**). Among these four peaks, the most downfield shifted ^{13}C signal (42.1 ppm) can be assigned to the L2 β methylene group as it is the closest to the amide group. The corresponding proton signals appear at 1.62 and 1.65 ppm as a doublet as the protons are diastereotopic (**Figure S 78C**).

In contrast, the most upfield shift ^{13}C peak, i.e. at 27.0 ppm can be assigned to the L2 γ methyne group, as this is the farthest from the amide group; the corresponding proton signals appear at 1.63 ppm (**Figure S 78C**).

The remaining two protons with the ^{13}C signals at 29.2 ppm should correspond to K3 δ and K5 δ , with ^1H -shifts at 1.63-1.67 ppm. All protons of the peptide side (except for K3 δ -K5 δ) can be resolved by the HSQC spectrum (**Figure S 79**).

2.5.3. Full ^1H -NMR Assignment of Daisy Chain $[\text{c}3]\text{-(1,4)}_3$

Figure S 80. Excerpts of the ^1H - ^1H TOCSY spectrum of $[\text{c}3]\text{-(1,4)}_3$, highlighting cross-peaks in the A) amide region (8-9 ppm) B) α -proton region (3.8-4.5 ppm) and the C) ^1H - ^1H dqfCOSY spectrum of $[\text{c}3]\text{-(1,4)}_3$, highlighting cross-peaks in the amide region.

First, the amide peaks were identified based on the number of their cross-peaks in the ^1H - ^1H TOCSY spectrum (**Figure S 80A**). Following the similar logic applied to **1**, the G1_{NH} , L2_{NH} and S4_{NH} protons, having 1, 3 and 2 TOCSY cross-peaks, respectively, were assigned to the peaks at 9.26, 10.20 and 9.06 ppm, respectively.

The $\text{L2}\delta$ protons can also be identified based on their unique integral and chemical shift (6H, doublet of doublets at 0.92 ppm), as well as their TOCSY cross-peak with L2_{NH} .

In this region, both protons at 8.17 and 7.52 ppm give TOCSY cross-peaks with the proton at 3.00 ppm (which belongs either to $\text{K3}\epsilon$ or $\text{K5}\epsilon$, see **Figure S 78A** and **Figure S 82**), indicating that they belong to the same spin system. Consequently, they cannot be both amide protons, as in this case, one spin system corresponds to one amino acid residue, which features only one amide proton. However, the ϵ proton (3.00 ppm) shows only a weak cross-peak with peak at 8.17 ppm, whereas the cross-peak with the proton at 7.52 ppm is strong (**Figure S 80A**) and also appears in the dqfCOSY spectrum (**Figure S 80C**). Consequently, the latter one (7.52 ppm) must correspond to the ammonium protons of lysine, while the former one (8.17 ppm) is an amide proton. These ammonium and amide protons are also clearly distinguishable based on their different linewidths (**Figure S 84A**).

The peak at 6.93 ppm, which thus must correspond to the other lysine amide proton shows one cross-peak in the α region (3.73 ppm) and numerous cross-peaks with highly upfield shifted protons. The latter observation gives the first indication of the alkyl chain of this lysine unit being threaded through the macrocycle and thus experiencing a diamagnetic ring current.

The α -protons (**Figure S 80B**) can be assigned similarly, based on the number of their TOCSY cross-peaks. Thus, the signals at 4.06 and 4.62 ppm (overlapping) corresponds to $\text{G1}\alpha$ and $\text{L2}\alpha$, respectively, whereas the signals at 4.37 and 3.73 ppm correspond to the non-threaded (cross peak at 3.00 ppm) and threaded (cross-peaks in the upfield region) lysines, respectively. Finally, the peak at 4.66 ppm (overlapping with $\text{L2}\alpha$) can be assigned to $\text{S4}\alpha$, whereas the remaining doublet at 3.99 ppm corresponds to $\text{S4}\beta$.

Interestingly, the S4 α and S4 β protons are shifted +0.41 and +0.14 ppm downfield with regard to **1**. This is a second indication of the threading of one lysine unit through the macrocycle as the serine moiety is close to the threaded lysine unit, consequently close to the encircling macrocycle and its protons are thus deshielded by the outward aromatic ring current from the macrocycle.

Figure S 81. ^1H - ^1H NOESY cross-correlations between amide and α -protons in **1143**, assuming that the peak at 6.93 ppm corresponds to A) K3 NH , resulting in full assignment B) K5 NH , leading to contradictions.

At this stage, the α - and amide protons of G1, S2 and S4, as well as of the threaded and the unthreaded lysine units have been assigned. However, it is still unclear which lysine moiety (K3 or K5) is threaded. From the NOESY cross-correlations of the amide and alpha protons (**Figure S 81A**) the position of the two lysine residues could be distinguished. Similarly to the logic applied for **1**, we could identify K5 α as the only α -proton showing exactly one cross-peak with α -protons (K5 NH , both shown in blue). Thus, the peaks at 8.16 and 4.37 ppm can be assigned to K5 α and K5 NH , respectively. Consequently, the peaks at 6.97 and 3.73 ppm can be assigned to K3 NH and K3 α , respectively.

Figure S 82. Excerpt of the A) ^1H - ^1H dqfCOSY B) ^1H - ^1H TOCSY spectrum of **1143**, highlighting cross-peaks between the alkyl protons of L2 and those of K5.

The rest of the NOESY cross-peaks is in line with the assignments obtained from TOCSY and COSY. It should be noted that L2 α overlaps with S4 α and thus, an alternative cross-walk scenario can be imagined in which the peak at 6.93 ppm corresponds to K5 NH (**Figure S 81B**). However, this leads to contradictions, most notably to the absence of the S4 NH /S4 α cross-peaks and the presence of at least 3 NOESY cross-peaks for K5 α . Thus, the amide proton at 6.97, giving TOCSY cross-peaks with the threaded, upfield shifted protons can be identified as K3 NH and the same holds for the corresponding α -proton at 3.73 ppm. Thus, the alkyl chain of K3 is threaded through the macrocycle, whereas K5 is non-threaded.

The alkyl protons of L2 and K5 could also be partially identified (**Figure S 82A**). Thus, L2 δ (0.9 ppm) gives a COSY cross-peak at 1.76 ppm, which thus corresponds to L2 γ . Similarly, L2 α (4.62 ppm) gives a COSY cross-peak at 1.92 ppm, which consequently can be assigned to L2 β . Moreover, K5 ϵ (3.00 ppm) shows COSY correlation with the peak at 1.75 ppm, which is thus identified as K5 δ . Finally, K5 α (4.37 ppm) shows COSY cross-correlation with the peak at 1.91 ppm, identified as K5 β . The only remaining peak at 1.55 ppm, although not well-resolved presumably corresponds to K5 γ . The corresponding cross-peaks also appear in the TOCSY spectrum (**Figure S 82B**).

Figure S 83. A) Depiction of the threaded [c3] daisy chain trimer of **1,4**, highlighting the aromatic protons and the methylenes of the threaded lysine side chain. B) Expanded view of the upfield region of the ^1H -NMR spectrum. C) ^1H - ^1H NOESY spectrum of **1,4**, highlighting cross-correlations between protons of the threaded lysine side chain and upfield shifted protons. D) Expanded view of the spectrum in C), highlighting cross-correlations between protons of the threaded lysine side chain with amide and aromatic protons.

The 4 methylene groups of the threaded lysine side chain (**Figure S 83A**) give rise to 7 clearly distinguishable signals between 0.72 and -1.65 ppm (**Figure S 83B**). This splitting might be caused by restricted rotation of the alkyl chain within the confined cavity of the macrocycle. Moreover, these features were also observed in the NMR spectrum of the **4₄Sp** complex.⁶ These upfield-shifted protons give rise to NOESY cross-peaks with at least 6 protons between 6.9-10.5 ppm (**Figure S 83C**). Most notably, only 3 from these peaks can be assigned to amide or ammonium protons (L2 NH , K3 NH , S4 NH). Thus, the 3 other cross-peaks must correspond to through-space correlation with the aromatic protons (**Figure S 83D**).

Figure S 84. A) ^1H -NMR spectrum of **143** in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (top) and pure D_2O (bottom) to distinguish between amide/ammonium and aromatic protons (purple strips). B) ^1H - ^1H TOCSY spectrum of **143**, highlighting aromatic protons belonging to the same spin system (cross-peaks connected by dashed lines).

Although the signals of the aromatic protons are relatively broad, most possibly due to the partial rotation (tumbling) of the aromatic rings around the C-S bonds, they could be distinguished by comparing the spectra of **143** recorded in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (**Figure S 84A**, top) and pure D_2O (**Figure S 84A**, bottom). In this manner, an additional exchangeable proton can be identified at 7.31 ppm, possibly belonging to K3_{NH3} (vide infra). In this manner, the peaks at 8.53, 8.26, 7.98, 7.93, 7.61 and 7.41 ppm can be assigned to the 6 aromatic protons of **143**. The off-diagonal cross-peaks of the ^1H - ^1H TOCSY spectrum can distinguish between spin systems, i.e. protons belonging to the same benzene ring (**Figure S 84B**). According to the TOCSY spectrum, the protons at 8.53 and 7.98 ppm, at 8.26 and 7.61 ppm and at 7.93 and 7.41 ppm are attached to the same ring.

Figure S 85. A) Depiction of the structure of **143**, highlighting the aromatic protons and the methylenes of the threaded lysine side chain ($\text{R}=\text{COOH}$). ^1H - ^1H TOCSY cross-correlations of the threaded K3 side chain protons with B) K3_{NH} and K3_{NH3} C) each other, indicating that they belong to the same spin system. C) ^1H - ^{13}C HSQC cross-correlations of the threaded K3 side chain protons. D) ^1H - ^1H COSY cross-correlations between the threaded K3 side chain protons.

The upfield shifted K3 side chain protons can also be distinguished from each other to certain extent (despite the complicated spectra in the upfield region). Thus, ^1H - ^1H TOCSY spectrum shows that the upfield shifted protons of the K3 side chain are coupled with two downfield protons: K3_{NH} (6.93 ppm) and another proton at 7.31 ppm, which is thus assigned as K3_{NH_3} (**Figure S 85B**).

The sizes of the cross-peaks belonging to these two protons correlate with their distance from the respective methylene group. Most importantly, the proton at -0.05 ppm does not give any cross-peak with K3_{NH} , whereas a distinct cross-peak can be observed with K3_{NH_3} . Therefore, the peak at -0.05 ppm belongs to K3_ϵ (**Figure S 85B**). Conversely, the peak at 0.72 ppm gives a large cross-peak with K3_{NH} , whereby no cross-peak can be observed with K3_{NH_3} . Thus, it could be assigned to K3_β (**Figure S 85B**). Finally, K3_{NH_3} shows only one additional TOCSY cross-peak in the upfield range, i.e. with the proton at -1.41 ppm, which thus most possibly belong to K3_δ (**Figure S 85B**).

Figure S 86. ^1H - ^1H NOESY cross-correlations of the threaded K3 side chain protons with aromatic (pink square), amide and ammonium protons.

Intriguingly, the HSQC spectrum shows splitting of the threaded lysine side chain protons (**Figure S 85C**). More closely, the protons at 0.72 and 0.36 ppm give HSQC cross-peaks with the same carbon at 33.1 ppm. As the proton at 0.72 ppm was assigned to K3_β , that at 0.36 ppm can be assigned to the vicinal $\text{K3}_\beta'$. Similarly, the protons at 0.36 and -1.00 ppm give HSQC cross-peaks with the same carbon at 25.5 ppm. As the only proton that has not been assigned yet is K3_γ , these two protons can be assigned to K3_γ and $\text{K3}_\gamma'$. Consequently, in the peak at +0.36 ppm the signals of K3_β as well as K3_γ partially overlap, which is also evident from the relative integrals of the corresponding protons (**Figure S 83B**). The ^1H - ^1H COSY cross-correlations can be readily interpreted upon this assumption (**Figure S 85D**).

It must be noted that the signals at -0.3 ppm and -1.66 ppm cannot be assigned at this stage, due to the possibly too low intensities in the HSQC and TOCSY spectra. Moreover, due to overlaps between the aromatic protons as well as between the methylenes of the threaded lysine side chain, the aromatic protons R1-R6 cannot be assigned based on the spectral information collected (**Figure S 86**). However, all other carbon signals can be assigned based on the full HSQC spectrum (**Figure S 87**).

Figure S 87. Aliphatic part of the ^1H - ^{13}C HSQC spectrum of **143**, displaying cross-peaks of the carbon-bound protons of the peptide chain.

2.6. UPLC Chromatograms from the Interconversion Experiments in the 1:4:Sp System

Figure S 88. UPLC chromatograms showing the initial and final composition of the DCLs involved in the mutual interconversion experiments shown on **Figure 6** of the main text.

- A) **1₈** and **4** (0.67 eq.) affords **1₄4₂**.
 B) **1₈** and **4** (0.5 eq., 33 mol % thiolate) fails to produce **1₁4₃** and affords **1₄4₂** instead.
 C) **1₄4₂** and **4** (2.0 eq.) affords a 5.2:1 mixture of **1₁4₃** and **1₄4₂**.
 D) **1₄4₂** and **Sp** (32 eq.) fails to produce **1₈**.
 E) **1₁4₃** and **Sp** (0.5 eq.) results in the expected conversion to **1₄4₂** and **4₄Sp**.
 F) **1₁4₃** and **1** (2.0 eq.) results in somewhat lower than expected conversion to **1₄4₂**.
 G) **1₁4₃** and **1** (0.75 eq., stoichiometric amount) fails to give **1₈** and affords **1₄4₂** instead.
 H) **4₄Sp** and **1** (0.57 eq.) fails to produce **1₁4₃** and yields **1₄4₂** as a major product instead.
 I) **4₄Sp** and **1** (2.7 eq.) affords **1₄4₂** in the expected yield.

2.7. Structural characterization of the peptide analogues of 1_{14_3} and 1_{44_2}

Figure S 89. Comparison of the circular dichroism (A) and thioflavin T fluorescence emission spectra (B) of the macrocycles 1_{44_2} , 2_{44_2} and 3_{44_2} at 2.0 mM total building block concentration (50 mM phosphate buffer, pH = 8.16).

Figure S 90. Negative staining TEM image of 2.0 mM (total building block concentration) samples of A) 2_{44_2} and B) 3_{44_2} , both showing fibers similar to 1_{44_2} .

Figure S 91. Comparison of the circular dichroism spectra of the macrocycles 1_{44_3} , 2_{44_3} , 3_{44_3} and 5_{44_3} at 2.0 mM total building block concentration (50 mM phosphate buffer, pH = 8.16).

3. Supplemental References

1. Malakoutikhah, M., Peyralans, J.J.-P., Colomb-Delsuc, M., Fanlo-Virgós, H., Stuart, M.C.A., and Otto, S. (2013). Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *J. Am. Chem. Soc.* **135**, 18406–18417.
2. Altay, Y., Tezcan, M., and Otto, S. (2017). Emergence of a New Self-Replicator from a Dynamic Combinatorial Library Requires a Specific Pre-Existing Replicator. *J. Am. Chem. Soc.* **139**, 13612–13615.
3. Besenius, P., Portale, G., Bomans, P.H.H., Janssen, H.M., Palmans, A.R.A., and Meijer, E.W. (2010). Controlling the growth and shape of chiral supramolecular polymers in water. *Proc. Natl. Acad. Sci.* **107**, 17888–17893.
4. Frederix, P.W.J.M., Idé, J., Altay, Y., Schaeffer, G., Surin, M., Beljonne, D., Bondarenko, A.S., Jansen, T.L.C., Otto, S., and Marrink, S.J. (2017). Structural and Spectroscopic Properties of Assemblies of Self-Replicating Peptide Macrocycles. *ACS Nano* **11**, 7858–7868.
5. Pal, A., Malakoutikhah, M., Leonetti, G., Tezcan, M., Colomb-Delsuc, M., Nguyen, V.D., van der Gucht, J., and Otto, S. (2015). Controlling the Structure and Length of Self-Synthesizing Supramolecular Polymers through Nucleated Growth and Disassembly. *Angew. Chem. Int. Ed.* **54**, 7852–7856.
6. Vial, L., Ludlow, R.F., Leclaire, J., Pérez-Fernandez, R., and Otto, S. (2006). Controlling the biological effects of spermine using a synthetic receptor. *J. Am. Chem. Soc.* **128**, 10253–10257.