HHS Public Access

Author manuscript

Nat Chem. Author manuscript; available in PMC 2022 February 23.

Published in final edited form as:

Nat Chem. 2021 September ; 13(9): 850–857. doi:10.1038/s41557-021-00770-7.

Dynamic covalent self-assembly of mechanically interlocked molecules solely made from peptides

Hendrik V. Schröder[†], **Yi Zhang**[†], **A. James Link**[†,‡,§,*]

[†]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, United States

[‡]Department of Chemistry, Princeton University, Princeton, NJ 08544, United States

[§]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, United States

Abstract

Mechanically interlocked molecules (MIMs), such as rotaxanes and catenanes, have captured the attention of chemists both from a synthetic perspective and because of their role as simple prototypes of molecular machines. Although examples exist in nature, most synthetic MIMs are made from artificial building blocks and assembled in organic solvents. Synthesis of MIMs from natural biomolecules remains highly challenging. Here we report on a synthesis strategy for interlocked molecules solely made from peptides—mechanically interlocked peptides (MIPs). Fully peptidic, cysteine-decorated building blocks were self-assembled in water to generate disulfide-bonded dynamic combinatorial libraries consisting of multiple different rotaxanes, catenanes and daisy chains as well as more exotic structures. Detailed NMR spectroscopy and mass spectrometry characterization of a [2]catenane comprised of two peptide macrocycles revealed that this structure has rich conformational dynamics reminiscent of protein folding. Thus, MIPs can serve as a bridge between fully synthetic MIMs and those found in nature.

Graphical Abstract

[*]Corresponding author: ajlink@princeton.edu.

Author contributions

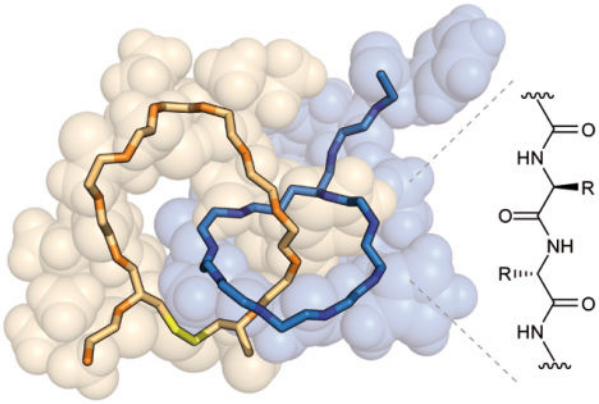
H.V.S. and A.J.L. conceived the project idea. H.V.S. and Y.Z. carried out the molecular cloning experiments and peptide purification. H.V.S. carried out mass spectrometry and NMR analyses. H.V.S. and A.J.L. analyzed data. H.V.S. and A.J.L. acquired funding. H.V.S wrote the first draft of the manuscript, and it was edited by A.J.L. All the authors contributed to and approved the final version of the manuscript.

Data availability statement

The coordinates for [2]catenane **3**-2H have been deposited in the Protein Data Bank (PDB) with accession code 7LL7. Coordinates have also been deposited to the Biological Magnetic Resonance Data Bank (BMRB), the accession number is 30852. All data, including mass spectra for all compounds, is present in the main text or the supplementary information. Due to the large data file size, raw mass spectrometry data underlying figures will be provided upon request.

Competing Interests Statement

The authors declare no competing interests.



As molecules progress from linear forms to entangled forms, a rich spectrum of properties emerges from an increase of structural complexity.[1] Prime examples in nature are the folds of proteins, which incorporate motifs of preorganized and entangled amino acid chains.[1–4] A dynamic yet precise structural alignment of functional units equips proteins and their complexes with a plethora of properties and functions. Chemists have recognized the power of spatial preorganization and have synthetically constructed novel types of intertwined architectures, in which molecules are held together by mechanical bonds[5–7] while often retaining mobility of subcomponents in an assembly. Mechanically interlocked molecules (MIMs), such as rotaxanes and catenanes, incorporating mobile parts are promising prototypes to mimic processes carried out by nature's biomolecular structures. Some of these smaller-sized artificial molecular machines with biologically inspired mechanisms[8,9] can, at least rudimentarily, resemble the by far more sophisticated operations of larger proteins. This has been shown by recent examples of artificial molecular motors[10], pumps[11], transporters[12] or assemblers[13]. To achieve the level of structural and functional complexity in biological systems, however, a logical next step is the construction of artificial molecular switches and machines from natural building blocks such that these machines can operate in aqueous conditions.

It is somewhat surprising that both domains, biomolecular machines and their synthetic mimics, have little overlap in their chemical structures. Chemical synthesis often involves template-directed[14,15] assembly of artificial components in organic solvents. Although various attempts[16–20] have been made, applicable synthesis strategies for wholly peptidic MIMs similar in size to synthetic MIMs,[21] have not been found yet. Elegant synthetic strategies that leverage the bisecting U motif of the p53 tetramerization domain have been developed to generate protein homocatenanes[22,23], heterocatenanes[24,25], and other entangled structures[26]. However, these mechanically interlocked proteins are much larger (~9 kDa and up) and bear only a passing resemblance to synthetic MIMs. While there is no hard size cutoff that delineates a peptide from a protein, we will use the term mechanically interlocked peptide (MIP) throughout the rest of this paper to refer to wholly peptidic systems comparable in size to synthetic MIMs. Syntheses of small hybrid systems made from biomolecules usually require synthetic support structures, such as crown ether/ammonium ion or metal coordination complexes, which remain in the structure. Rotaxanes,[27–31] catenanes,[32–37] and molecular knots[38,39] with hybrid structures, in which amino acids are

incorporated into such supports, have been reported. The challenge for synthetically derived MIPs without artificial supports arises from a lack of recognition motifs in peptides for template-directed synthesis in polar solvents.[29] This makes the construction of even simple MIPs, e.g., peptide [2]catenanes, synthetically challenging.

Here we report on a dynamic covalent synthesis for topologically complex MIPs using a simple one-pot protocol in water. Previously we have demonstrated that the lasso peptide microcin J25 (MccJ25), a naturally occurring peptide [1]rotaxane, could be genetically engineered to program the formation of [3]- and [4]catenanes via disuflide bonding.[40] Here we show that placement of Cys residues into different locations of the MccJ25 building block can be used as a precise tool to pre-determine the structure and topology of a self-assembled product. Structural selectivity is achieved without templating but instead by rational positioning of the Cys residues in the building block leading to a diversity of different complex structures. The synthesized set of MIPs includes [1–3]rotaxanes, molecular [c2]daisy chains, a lasso peptide with non-native chirality, [2–5]catenanes, and a double-lasso macrocycle. Detailed characterization of a peptide [2]catenane shows that MIPs can exist in stable conformations, which are accessible through co-conformational motion.

Results and discussion

Synthesis strategy and building block design.

A promising synthetic strategy for MIPs is the dynamic combinatorial chemistry approach, which can yield highly complex yet discrete assemblies as constituents in a dynamic combinatorial library (DCL).[41] Self-assembly of thiol-decorated building blocks has been extensively studied as strategy to synthesize sophisticated molecular architectures via disulfide bonding.[42–44] Prominent examples are DCLs studied by the groups of Sanders and Otto.[45–49] Reminiscent of synthetic DCLs, the biosyntheses of naturally occurring and cysteine-rich peptides with knotted structures, such as cyclotides, are known to proceed through efficient oxidative folding mechanisms in dynamic equilibria.[50]

Lasso peptides are a class of ribosomally synthesized and post-translationally modified peptides with a slipknot-like structure.[16,51] The entanglement is biosynthetically achieved by formation of an intramolecular isopeptide bond catalyzed by two enzymes. As a ‘pre-threaded’ molecule, lasso peptides are a well-suited supramolecular synthon for bottom-up self-assembly towards larger MIPs.[18] MccJ25 (**L1**) is a 21 amino acid (aa) lasso peptide with an 8 aa macrolactam ring, an 11 aa loop, and a 2 aa tail, in which the residues Phe-19 and Tyr-20 serve as sterically demanding residues (‘stoppers’) preventing the isopeptide-bonded ring from slipping off its thread (Fig. 1a).[52,53] Multiple amino acid substitutions in the MccJ25 scaffold are well-tolerated in its biosynthesis.[54–56] We hypothesized that rational substitutions by Cys residues in **L1**’s primary sequence and subsequent cleavage by a protease in its loop region would provide cysteine-rich [2]rotaxane building blocks, which can then selectively self-assemble into larger MIPs by disulfide bonding (Fig. 1b). We expect that the positions of Cys residues, in the otherwise structurally identical building blocks, dictate the resulting covalent connectivity and topology of the resulting MIPs.

The general strategy of generating [2]rotaxanes can be also applied to other lasso peptide scaffolds.[57]

MccJ25 variants **L2**–**L8** were engineered with multiple Cys substitutions at 4 different positions (Fig. 2b, see Methods and Supplementary Information). Trypsinization[40] cleaves variants with an additional substitution of Gly-12 with Arg (**L4**, **L5**, and **L7**) in the lasso peptide's loop region between Arg-12/Cys-13 yielding [2]rotaxanes building blocks with a 9 aa thread. In contrast, thermolysin[58,59] was used for variants **L2**, **L3**, **L6**, and **L8** and cleaves between Phe-10/Val-11 (Fig. 2a) yielding an 11 aa thread. Longer digestion times enable a second cleavage between Xaa-12/Ile-13. The selective Phe-10/Val-11 cleavage was confirmed by HPLC (Supplementary Fig. 2) and tandem mass spectrometry (MS^2) with collision induced dissociation (CID) experiments (Supplementary Figs 16 and 17).

Synthesis of mechanically interlocked peptides.

Synthesis of MIPs is carried out in a one-pot, two-step protocol in aqueous solution: (1) enzymatic digestion and (2) *in situ* self-assembly of the generated [2]rotaxane intermediates driven by air oxidation (see Methods). The general synthesis strategy is demonstrated by the example of variant **L5**, which bears Cys-13 (loop) and Cys-5 (ring) residues as shown in Figure 1b and 3. We envisioned that this positioning of Cys residues could generate homomeric oligomers with architectures of molecular [c*n*]daisy chains.

Trypsin cleavage (Arg-12/Cys-13) of lasso peptide **L5** quantitatively yields [2]rotaxane intermediate **5** as indicated by a mass increase of +18 Da and a retention time shift (Figs 3b and 3c). Air-driven oxidation generates DCL $(\mathbf{5}\text{-}2H)_n$ consisting of homomeric constituents (n = 1, 2, …) with identical *m/z* values (Supplementary Figs 55 and 56). Isotopic patterns identify the isolated library members as a monomer, two dimers, and traces of trimers, with dimer $(\mathbf{5}\text{-}2H)_2$ (15.8 min retention time) as the major constituent (Fig. 3c, see Supplementary Information). CID experiments reveal a characteristic gas phase fragmentation pattern (Supplementary Figs 15 and 19) indicative of connectivity between Cys-5 and Cys-13 within the dimer. Thus, the preferred dimer exhibits a sterically favorable 'head-to-tail' linkage with the overall structure of a molecular [c2]daisy chain. This is the first example of molecular daisy chains solely made from peptides. The [c2]daisy chain $(\mathbf{5}\text{-}2H)_2$ is formed in 72% yield after 1 day and 80% yield after 9 days, as determined by HPLC (Supplementary Fig. 3). We subsequently added dithiothreitol (DTT) to the DCL and quantitatively generated building block **5** in its reduced form (Fig. 3c, see Methods). Hence, the self-assembly process is reversible suggesting that multiple rounds of DCL formation can be carried out to drive the DCL towards a single desired product.

Kinetic profiles of DCL evolution (Supplementary Fig. 3) show that air-driven oxidation proceeds within ~2 h. Full oxidation of **5**'s thiols was ensured by addition of iodoacetamide (Supplementary Fig. 4, see Methods). After this initial oxidation, the kinetic distribution converts, by disulfide exchange, into a thermodynamic one, which energetically favors the [c2]daisy chain $(\mathbf{5}\text{-}2H)_2$, in a slow equilibration phase over several days. To exclude effects of trypsin in the reaction mix, we repeated the oxidation with isolated, DTT-reduced **5**

and found that the oxidation phase proceeded similarly. In accordance with the literature,[60] addition of DMSO (5 % (v/v)) slightly accelerates oxidation (Supplementary Fig. 3).

Steric and conformational constraints arising from an intramolecular disulfide bond, disfavor the monomer and amplify the [c2]daisy chain. Additionally, oligomerization may also be driven by supramolecular interactions of the building blocks in water (e.g. hydrophobic effects or hydrogen bonding).[40] Although a dimer can adopt two different connectivities (Supplementary Fig. 19), a head-to-tail connectivity is preferred (80% vs 4% yield as determined by HPLC). Overall, the positioning of Cys residues pre-determines the [c2]daisy chain as preferred library member.

We were interested in demonstrating the broad applicability of our synthesis strategy by synthesizing different MIPs. Variants **L2**–**4** and **L6**–**8** (0.25 mM) were digested in aqueous buffer (Fig. 3d and Supplementary Fig. 5, see Methods and Supplementary Information for HPLC and MS data). Building block **4** with only one Cys residue in the thread dimerizes by disulfide bonding between the N-terminal residues (Cys-13) into the symmetric [3]rotaxane (**4**-H)$_2$ (*m/z* 1107.0, 73% yield). The structure and disulfide connectivity was confirmed by MS^2/CID experiments (Supplementary Figs 22 and 23).

Since the macrolactam ring in a lasso peptide displays N-to-C sequence directionality, a particularly interesting stereochemical feature is that the lasso's tail can be, in theory, threaded from different sides into the ring resulting in two mechanically chiral species. Thus far, only right-handed lasso peptides like MccJ25 (**L1**) have been found in nature. Building block **L2** with spatially close Cys residues (Cys-5 and Cys-21) forms an intramolecular disulfide bond between tail and ring, which can only result in a non-native, left-handed lasso peptide after cleaving its loop. Indeed, digestion and oxidation generates [1]rotaxane **2**-2H (*m/z* 1068.0, 49% yield) accompanied by a minor fraction of the side product **2**-2H ΔN2 (*m/z* 990.0, 34% yield), which is the same [1]rotaxane after a second thermolysin cleavage and loss of the N-terminal Val-Gly dipeptide (Supplementary Figs 20 and 21).

Building block **3** incorporates two Cys residues in its thread (Cys-12 and Cys-21) with a Cys/Cys sequence distance of 10 aa and self-assembles into a mixture of [2–5] catenanes. The resulting DCL consists of [2]catenane **3**-2H (16.2 min, 11%), [3]catenane (**3**-2H)$_2$ (17.6 min, 52%) and minor amounts of [4]- and [5]catenanes (Figure 3d, Supplementary Figs 5 and 24–27, see below for a more detailed characterization of the [2]catenane). The [3]catenane has a head-to-tail structure consisting of two isopeptide-bonded 8 aa rings encircling a central, bis(disulfide)-bonded 20 aa macrocycle. The [4]- and [5]catenanes have different building block orientations and a central, disulfide-bonded macrocycle of 30 aa ([4]catenane) and 40 aa ([5]catenane).

Building block **6** (Cys-12, Cys-13, Cys-21) forms an intramolecular bond between Cys-12 and Cys-13 on its thread (Supplementary Fig. 30). A subsequent intermolecular dimerization of [2]rotaxane **6**-2H via Cys-21 proceeds slowly, because this residue is sterically more shielded. However, building block **7** (Cys-21, Cys-13, and Cys-5) shows a more complex self-assembly behavior (Supplementary Fig. 31). A dimer ((**7**-3H)$_2$) is the most abundant species along with minor amounts of a tetramer (Supplementary Fig. 33). In theory, the 6

cysteine residues in the homodimer can display 11 different connectivities (Supplementary Fig. 32). CID of the dimer revealed a Cys-13/Cys-13 and two Cys-5/Cys-21 connectivities indicating a daisy chain structure in which the ends are conjoined to form a closed loop, a double-lasso macrocycle (Fig. 3d, Supplementary Figs 33 and 34). Disulfide bonds within the dimer were further rationalized by a computational structure derived at the GFN2-xTB level of theory (Supplementary Fig. 6). These findings are reminiscent of reported chemical syntheses of double-lasso macrocycles from crown ether-based rotaxanes in organic solvents.[61]

Mixed DCLs and building block modifications.

Besides homooligomers, we were also interested in the synthesis of heteromeric constructs from mixed DCLs, which allows the exploration of a broader structural and topological space for MIPs. To exemplify this, dithiol building blocks **3** (catenane-forming) and **5** (daisy chain-forming) were mixed and self-assembled in equimolar amount (Fig. 4a). LC-MS analysis after 2 days (Supplementary Fig. 7) reveals a mixture of monomers, homo- and heterooligomers without significant preferences for social or narcissistic self-sorting.[62] Thus, the three possible head-to-tail dimers, (**3**-2H)$_2$, **3**•**5**-4H, and (**5**-2H)$_2$ are obtained in the statistically expected ratio of ~1:2:1, respectively (Fig. 4, Supplementary Fig. 7). Based on this, the head-to-tail heterodimer **3**•**5**-4H (*m*/*z* 1098.7, 27% yield) is the most abundant constituent. Dimer **3**•**5**-4H has a [2]catenane structure exhibiting two mechanical bonds, in which the larger macrocycle is composed of a bis(disulfide)-bonded lasso structure.

Pre-programming of building blocks by substitutions with other amino acids than cysteine can be utilized as an additional tool to control a peptide DCL. Library (**3**-2H)$_n$ consists of several homomeric species in thermodynamic equilibrium. We anticipated that the Ser-18 residues are spatially close in a head-to-tail dimer. We reasoned that substitution of Ser with a charged aa would lead to a shift in the DCL away from the homodimer. For building block **8** (Cys-12, Cys-21, Asp-18) with a S18D substitution (Supplementary Fig. 8), we observed a significant shift of the DCL (**8**-2H)$_n$ towards the monomer (62%) in comparison to building block **3** without S18D substitution (27% monomer) at a 100 μM concentration (Fig. 4b). Furthermore, in a mixed DCL, starting from equimolar amounts of building block **3** and **8**, this effect prevails over the formation of the heterodimer **3**•**8**-4H and drives the DCL towards monomer **8**-2H (Supplementary Fig. 9). We assume that besides charge repulsion further inter-building block interactions and solvent effects play a role. These results demonstrate that even simple side chain modifications can be utilized to control equilibria in a peptide DCL.

A [2]catenane made from macrocyclic peptides.

Since their first discovery,[63,64] synthesis of [2]catenanes, composed of two singly interlocked macrocycles, has been a major driving force for the field of mechanically interlocked molecules. To demonstrate that our DCL-approach is well-suited for a preparative-scale synthesis of MIPs, we were interested in the production and in-depth analysis of this fundamental molecular structure. Despite improved monomer formation in DCL (**8**-2H)$_n$, the higher expression level of lasso peptide **L3** (Supplementary Table 4) made the synthesis from DCL (**3**-2H)$_n$ more attractive. We performed its synthesis under high

dilution conditions (25 μM) and at optimized temperature (see Methods, Supplementary Fig. 10) and were pleased to obtain the [2]catenane **3**-2H in 47% yield. After isolation from the DCL, the [2]catenane was stable in aqueous solution (pH 6.0) without signs of oligomerization or degradation after more than a week (Supplementary Fig. 10).

The intramolecular disulfide bond in [2]catenane **3**-2H (*m/z* 1108) is reflected by a mass loss of 2 Da compared to building block **3** (Fig. 5b). Mass selection of the $[M+2H]^{2+}$ ion followed by CID confirms the interlocked structure of the rings and demonstrates a characteristic property of [2]catenanes: a covalent bond must break to separate both macrocycles (fragments *m/z* 1015.5 and 1200.5). Disassembly in the gas phase, without fragmentation of the disulfide-bonded ring, indicates that partial pirouetting of the isopeptide-bonded ring through the annulus of the disulfide-bonded ring is possible (Supplementary Fig. 28). It is likely that the sequence Glu-8–Phe-10 is sterically demanding enough[65] to prevent a complete 360° pirouetting of the isopeptide-bonded ring and that the observed dissociation in the gas phase results from partial co-conformational motion. Unthreading of [2]rotaxane intermediates in significant amounts was neither observed during CID nor in a heating experiment (95 °C, 2h) in solution (Supplementary Fig. 13).[59] Pirouetting of the isopeptide-bonded ring is presumably the favored degree of freedom on the otherwise restricted co-conformational energy landscape.

Additional insight into the [2]catenane structure came from a protease reaction using carboxypeptidase[65], which removes the Phe-10 residue of the isopeptide-bonded ring without re-opening the disulfide-bonded ring (Fig. 5c and Supplementary Fig. 14, see Methods). The trimmed version of the [2]catenane was generated in 90% yield. The precise removal of only 1 aa demonstrates that MIPs are a promising synthetic platform suitable for post-assembly modifications[66] including enzymatic transformations.

Evidence for a dynamic structure was provided by broadening and unusual fronting of the [2]catenane peak during HPLC purification (Supplementary Fig. 10). This fronting is consistent across different stationary and mobile phases (Supplementary Fig. 11 and Supplementary Table 6). CID experiments with the leading and trailing edges of the peak give similar fragmentation patterns (Supplementary Fig. 29) and re-injection of the HPLC-isolated peak's trailing edge gives a peak with the initially observed fronting (Supplementary Fig. 12). We hypothesized that this unusual chromatographic behavior may emerge from a dynamic structure of [2]catenane **3**-2H. *Cis*/*trans* isomerization of the Pro-16 residue in wild-type microcin J25 (**L1**) has been recently identified to enable conformational exchange between two defined conformations.[67] To test whether Pro-16 plays a role in the structural dynamics, we synthesized another [2]catenane made from the Ala-16 variant **3 P16A** (Cys-12, Cys-21, Ala-16). LC/MS analysis does not show the indicative peak broadening and fronting as observed for the similar [2]catenane **3**-2H with Pro-16 residue (Supplementary Fig. 12).

The presence of defined isomers was further confirmed by NMR experiments in H_2O/D_2O (95/5) at 295 K (see Supplementary Information). A ^1H NMR and a total correlation spectroscopy (TOCSY) spectrum display two related sets of sharp signals with an average ratio of peak intensities of 1:4 indicating the presence of the two species in a slow exchange

regime (Fig. 5d, Supplementary Figs 35 and 36). An atypical TOCSY pattern and chemical shifts helped us to identify the minor conformer of **3**-2H as *cis*-Pro-16 form (Supplementary Figs 37 and 38). However, dispersed chemical shift differences between the *cis* (minor) and *trans* (major) form of **3**-2H in the isopeptide-bonded and disulfide-bonded ring indicate that the change of chemical environment cannot only be localized[67] on Pro-16 but is rather distributed on both macrocycles (Supplementary Fig. 37). Hence, we propose a mechanism of co-conformational transfer in which the Pro-induced conformational change of one macrocycle triggers a structural response in the other due to its constriction in the crowded ensemble. This is reminiscent of the pivotal role of Pro *cis*/*trans* isomerization in the folding of proteins.[68]

^{1}H,^{1}H exchange spectroscopy (EXSY, t_m = 1.5 s, 295 K) shows no evidence for magnetization transfer and confirms a slow conformer interconversion on the NMR timescale (Supplementary Fig. 39). Although variable-temperature NMR (277 to 337 K) does not reveal peak coalescences, significant chemical shifts in the backbone amide region are observed while retaining sharp peaks in the spectrum overall (Supplementary Fig. 40). Thus, the co-conformational states remain well-defined at higher temperatures while the rings exhibit local structural mobility.[69] A subsequent cooling re-generated the initial ^{1}H NMR spectrum (Supplementary Fig. 41). The combined chromatography, MS, and NMR data suggests that [2]catenane **3**-2H shows local mobility while existing in a slow equilibrium of Pro-16 *cis/trans* isomerization.

TOCSY and a nuclear Overhauser effect spectroscopy (NOESY) NMR additionally reveal a wide chemical shift dispersion in the fingerprint region for both co-conformers, characteristic of highly structured peptides (Supplementary Figs 36 and 42 and Supplementary Table 7).[53] Despite partial overlapping of the co-conformers' peak sets in the TOCSY and NOESY spectra, we were able to generate a conformational ensemble for the major species (i.e. ground-state co-conformer)[7] of **3**-2H (Fig. 5e, Supplementary Figs 43–45; see Supplementary Information). The disulfide-bonded ring preferably encircles the sterically less crowded region on the isopeptide-bonded ring between residues Gly-1–Ala-3. The isopeptide-bonded macrocycle is conformationally constrained while the disulfide-bonded ring is more flexible (Supplementary Fig. 45), which can be potentially ascribed to the larger circumference of the latter. A similar behavior is well-known for MccJ25 (**L1**), which displays a rigid ring yet a more flexible loop region.[52,53] The residues excluded from the interlocked rings (Tyr-9, Phe-10, and Val-11) are less structured compared to the rings. These conformational properties were further supported by a 1 μs molecular dynamics (MD) simulation at 298 K and a *B*-factor plot (Supplementary Fig. 46) based on the root-mean-square deviation (RMSD) values of the MD trajectories (see Supplementary Information for details). The MD results indicate relatively defined conformations of both interlocked rings, with a slightly increased rigidity of the isopeptide-bonded ring over the disulfide-bonded ring, and high flexibility of residues 9–11.

MIP **3**-2H is the first [2]catenane solely made from two macrocyclic peptides. With ring circumferences of 32 and 26 atoms (10 aa and 8 aa), [2]catenane **3**-2H is similar in size to Sauvage's copper-based [2]catenanes[70] (30 × 30 atom rings) or Stoddart's cylophane-based [2]catenanes[71] (34 × 28 atom rings). The isopeptide-bonded ring (26 atoms)

resembles the dimensions of common synthetic macrocycles such as dibenzo-24-crown-8 (24 atoms) or cyclobis(paraquat-*p*-phenylene) (28 atoms). Another compelling feature is the [2]catenane's complex stereochemistry: both macrocycles have distinct faces characteristic for mechanically axially chiral[72] catenanes. Both interlocked rings also display sequence directionality (Fig. 5a). Thus, the [2]catenane **3**-2H is topologically chiral[72] whereby only one isomer can be generated from the right-handed lasso peptide **L3**.

Conclusions

Here we show that mechanically interlocked molecules, including rotaxanes, daisy chains, catenanes and double-lasso macrocycles, can be solely made from peptides without synthetic supports. Using dynamic covalent self-assembly, these mechanically interlocked peptides (MIPs) are efficiently synthesized by a simple, one-pot protocol in water. All constructs are built from genetically engineered peptide building blocks and structural selectivity is predetermined by different positioning of Cys residues. The water solubility and peptidic nature of MIPs enable precise post-assembly modifications by enzymes. As shown by the example of a [2]catenane made from 32 and 26 atom-sized macrocyclic peptides, a crowded but partially dynamic structure exhibits co-conformational isomerism, reminiscent of the defined folding pathways of proteins. A dynamic structural exchange between well-defined co-conformations equips MIPs with excellent prerequisites for the construction of molecular switches and machines. MIPs without synthetic supports also enable the route towards artificial constructs with increased biocompatibility, for example for drug delivery or biomaterial applications. As a next step, functionality and stimuli-responsiveness need to be engineered into MIPs to provide motion control using external stimuli, chemical fuels or even chaperones. Enzymes such as peptidyl prolyl isomerase could be used to induce conformational changes. Furthermore, we have recently shown that thermally induced, reversible wheel translation in lasso peptides is possible, allowing for the introduction of further motion into the next generation of MIPs.[57] Functional MIPs can provide the missing link to connect the worlds of synthetic interlocked molecules and nature's sophisticated protein machinery.

Methods

Plasmid construction.

A set of genes encoding the variants was generated using our previously established methods (see Supplementary Information). Briefly, the plasmids contain the genes for a swappable wild type MccJ25 precursor (McjA), as well as for the maturation enzymes McjB and McjC, and exporter McjD. The plasmids also contain an isopropyl *β*-D-1-thiogalactopyranoside (IPTG) inducible promoter and ampicillin resistance gene. Site-directed mutagenesis was performed on the precursor gene *mcjA*, which was subsequently amplified by PCR and then ligated into the vector pJP73 from which the wild-type gene was removed. The sequences of all constructs were confirmed by Sanger sequencing.

Expression and purification of MccJ25 variants.

Plasmids were transformed into electrocompetent BL21 cells and plated onto lysogeny broth (LB) agar plates containing 100 mg/L ampicillin. Starter cultures were created by picking colonies into LB (5 mL, 100 mg/L ampicillin) and grown overnight at 37 °C. These cells were sub-cultured into M9 minimal media (500 mL) at a starting OD_{600} of 0.02, and grown at 37 °C. Upon reaching an OD_{600} of 0.3–0.4, cells were induced by adding IPTG to 1 mM final concentration. The cells were grown for 16–20 h at room temperature after induction. Afterwards, the culture was centrifuged, the cells were discarded and the supernatant was ran through a 1 g / 6 mL Strata C8 column. The column was eluted by adding MeOH (6 mL) containing 5 mM DTT, the eluent was concentrated under reduced pressure, and the residue was taken up in 500 μL of a water/ACN (1:1) mix. The extracts were purified by semi-preparative HPLC (see Supplementary Information). Expression of all variants was confirmed by LC/MS and HPLC (Supplementary Fig. 1, Supplementary Tables 4 and 5). The variants were obtained in isolated yields between ~0.3–0.9 mg per liter of culture.

Preparation of peptide DCLs.

Peptides (0.25 mM) containing an Arg-12 substitution were trypsinized in $(NH_4)HCO_3$ buffer (50 mM, pH 7.8). Sequencing-grade trypsin was added (1:50 ratio of trypsin mass:peptide mass) and the solution was left at room temperature in air for typically 1 or 2 days. The DCLs were directly analyzed or purified via LC/MS and/or HPLC. Analogously, peptides (0.25 mM) without an Arg-12 substitution were thermolysin digested in buffered solution containing Tris (50 mM) and $CaCl_2$ (0.5 mM) at pH = 8.0. Thermolysin was added in a 1:50 ratio (thermolysin mass:peptide mass). The digestions were left at rt for 10 min and subsequently quenched by addition of formic acid (10 % in water) to a final concentration of 0.5 % (v/v) for a single cleavage or for 1 day at rt. Digests were directly analyzed or purified by HPLC. Isolated building blocks (0.25 or 0.10 mM) were dissolved in $(NH_4)HCO_3$ buffer (50 mM, pH 7.8) to self-assemble the corresponding DCLs via air oxidation.

Preparative [2]catenane synthesis.

Dithiol **3** (25 μM) was dissolved in $(NH_4)HCO_3$ buffer (50 mM, pH 7.8) and heated to 37 °C over 2 days. Afterwards, the reaction was quenched by addition of formic acid to a final concentration of 0.5 % (v/v). The mixture was frozen, lyophilized and the dry residue was dissolved in a minimal volume of water/acetonitrile (9:1) containing TFA (0.1 % (v/v)). The [2]catenane was purified by HPLC and obtained in 47% yield (determined by HPLC peak area). The rest of the DCL was isolated, reduced by DTT and the regenerated building block **3** was used for additional rounds of DCL formation.

DTT reduction and iodoacetamide reactions.

DTT (100 equiv.) was added to peptides (0.25 mM) in buffered solution and the mixture was heated for 30 min at 55 °C. The mixtures were directly analyzed via HPLC. Iodoacetamide (100 equiv.) was dissolved in a minimum amount of water and added to a DCL (0.25 mM). The mixture was left at rt for 2 h and was then analyzed by HPLC.

Carboxypeptidase assay.

[2]Catenane **3**-2H (6 μg) was digested with 1 U carboxypeptidase B (Sigma-Aldrich) and 1 U carboxypeptidase Y in carboxypeptidase digestion buffer (50 mM sodium acetate, pH 6.0) with a total volume of 20 μL for 1 day at room temperature. The mixture was analyzed by LC/MS after 7, 17 and 31 h.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank István Pelczer and Hader Elashal for support with the NMR experiments. This work was supported by NIH grant GM107036 to A.J.L. H.V.S. gratefully acknowledges support by the Deutsche Forschungsgemeinschaft (DFG Research Fellowship 427725459).

References

1. Siegel JSDriving the formation of molecular knots. Science338, 752–753 (2012). [PubMed: 23139321]
2. Liang C & Mislow K Knots in Proteins. J. Am. Chem. Soc 116, 11189–11190 (1994).
3. Mallam AL, Morris ER & Jackson SE Exploring knotting mechanisms in protein folding. Proc. Natl Acad. Sci. USA 105, 18740–18745 (2008). [PubMed: 19015517]
4. Dabrowski-Tumanski P & Sulkowska JI Topological knots and links in proteins. Proc. Natl Acad. Sci. USA 114, 3415–3420 (2017). [PubMed: 28280100]
5. Stoddart JFMechanically Interlocked Molecules (MIMs)-Molecular Shuttles, Switches, and Machines (Nobel Lecture). Angew. Chem. Int. Ed56, 11094–11125 (2017).
6. Sauvage J-PFrom Chemical Topology to Molecular Machines (Nobel Lecture). Angew. Chem. Int. Ed56, 11080–11093 (2017).
7. Stoddart JF & Bruns CJ The Nature of the Mechanical Bond: From Molecules to Machines, (John Wiley & Sons, Inc., Hoboken, New Jersey, 2016).
8. Browne WR & Feringa BL Making molecular machines work. Nat. Nanotechnol 1, 25–35 (2006). [PubMed: 18654138]
9. Zhang L, Marcos V & Leigh DA Molecular machines with bio-inspired mechanisms. Proc. Natl Acad. Sci. USA 115, 9397–9404 (2018). [PubMed: 29483259]
10. Erbas-Cakmak Set al.Rotary and linear molecular motors driven by pulses of a chemical fuel. Science358, 340–343 (2017). [PubMed: 29051374]
11. Cheng Cet al.An artificial molecular pump. Nat. Nanotechnol10, 547–553 (2015). [PubMed: 25984834]
12. Chen Set al.An Artificial Molecular Shuttle Operates in Lipid Bilayers for Ion Transport. J. Am. Chem. Soc140, 17992–17998 (2018). [PubMed: 30445811]
13. Lewandowski Bet al.Sequence-specific peptide synthesis by an artificial small-molecule machine. Science339, 189–193 (2013). [PubMed: 23307739]
14. Schalley CA, Vögtle F & Dötz KH Templates in Chemistry I, (Springer-Verlag Berlin Heidelberg, Heidelberg, 2004).
15. Crowley JD, Goldup SM, Lee AL, Leigh DA & McBurney RT Active metal template synthesis of rotaxanes, catenanes and molecular shuttles. Chem. Soc. Rev 38, 1530–1541 (2009). [PubMed: 19587949]
16. Ferguson ALet al.An experimental and computational investigation of spontaneous lasso formation in microcin J25. Biophys. J99, 3056–3065 (2010). [PubMed: 21044604]

17. Steemers L, Wanner MJ, Lutz M, Hiemstra H & van Maarseveen JH Synthesis of spiro quasi[1]catenanes and quasi[1]rotaxanes via a templated backfolding strategy. Nat. Commun 8, 15392 (2017). [PubMed: 28541349]
18. Martin-Gomez H & Tulla-Puche J Lasso peptides: chemical approaches and structural elucidation. Org. Biomol. Chem 16, 5065–5080 (2018). [PubMed: 29956707]
19. Waliczek Met al.Attempting to synthesize lasso peptides using high pressure. PLoS One15, e0234901 (2020). [PubMed: 32579565]
20. Soudy R, Wang L & Kaur K Synthetic peptides derived from the sequence of a lasso peptide microcin J25 show antibacterial activity. Bioorg. Med. Chem 20, 1794–1800 (2012). [PubMed: 22304849]
21. Moretto A, Crisma M, Formaggio F & Toniolo C Peptide-based rotaxanes and catenanes: an emerging class of supramolecular chemistry systems. Biomol. Concepts 3, 183–192 (2012). [PubMed: 25436531]
22. Yan LZ & Dawson PE Design and Synthesis of a Protein Catenane. Angew. Chem. Int. Ed 40, 3625 (2001).
23. Blankenship JW & Dawson PE Threading a peptide through a peptide: protein loops, rotaxanes, and knots. Protein Sci. 16, 1249–1256 (2007). [PubMed: 17567748]
24. Da XD & Zhang WB Active Template Synthesis of Protein Heterocatenanes. Angew. Chem. Int. Ed 58, 11097–11104 (2019).
25. Liu Yet al.Cellular Synthesis and X-ray Crystal Structure of a Designed Protein Heterocatenane. Angew. Chem. Int. Ed59, 16122–16127 (2020).
26. Liu Yet al.Lasso Proteins: Modular Design, Cellular Synthesis, and Topological Transformation. Angew. Chem. Int. Ed59, 19153–19161 (2020).
27. Zhai C, Schreiber CL, Padilla-Coley S, Oliver AG & Smith BD Fluorescent Self-Threaded Peptide Probes for Biological Imaging. Angew. Chem. Int. Ed 59, 23740–23747 (2020).
28. Saito F & Bode JW Synthesis and stabilities of peptide-based [1]rotaxanes: molecular grafting onto lasso peptide scaffolds. Chem. Sci 8, 2878–2884 (2017). [PubMed: 28553526]
29. Aucagne V, Leigh DA, Lock JS & Thomson AR Rotaxanes of cyclic peptides. J. Am. Chem. Soc 128, 1784–1785 (2006). [PubMed: 16464065]
30. Young MJ, Akien GR & Evans NH An amide hydrogen bond templated [1]rotaxane displaying a peptide motif - demonstrating an expedient route to synthetic mimics of lasso peptides. Org. Biomol. Chem 18, 5203–5209 (2020). [PubMed: 32597913]
31. Leigh DA, Murphy A, Smart JP & Slawin AMZ Glycylglycine Rotaxanes—The Hydrogen Bond Directed Assembly of Synthetic Peptide Rotaxanes. Angew. Chem. Int. Ed 36, 728–732 (1997).
32. Sawada T, Yamagami M, Ohara K, Yamaguchi K & Fujita M Peptide [4]Catenane by Folding and Assembly. Angew. Chem. Int. Ed 55, 4519–4522 (2016).
33. Sawada T, Inomata Y, Shimokawa K & Fujita M A metal-peptide capsule by multiple ring threading. Nat. Commun 10, 5687 (2019). [PubMed: 31831732]
34. Lam RTet al.Amplification of acetylcholine-binding catenanes from dynamic combinatorial libraries. Science308, 667–669 (2005). [PubMed: 15761119]
35. Schulte TRet al.A New Mechanically-Interlocked [Pd2 L4] Cage Motif by Dimerization of two Peptide-based Lemniscates. Angew. Chem. Int. Ed59, 22489–22493 (2020).
36. Chung MK, White PS, Lee SJ, Waters ML & Gagné MR Self-assembled multi-component catenanes: structural insights into an adaptable class of molecular receptors and [2]-catenanes. J. Am. Chem. Soc 134, 11415–11429 (2012). [PubMed: 22686480]
37. Chung MK, Lee SJ, Waters ML & Gagné MR Self-assembled multi-component catenanes: the effect of multivalency and cooperativity on structure and stability. J. Am. Chem. Soc 134, 11430–11443 (2012). [PubMed: 22686511]
38. Inomata Y, Sawada T & Fujita M Metal-Peptide Torus Knots from Flexible Short Peptides. Chem 6, 294–303 (2020).
39. Song YWet al.Effects of turn-structure on folding and entanglement in artificial molecular overhand knots. Chem. Sci12, 1826–1833 (2021).

40. Allen CD & Link AJ Self-Assembly of Catenanes from Lasso Peptides. J. Am. Chem. Soc 138, 14214–14217 (2016). [PubMed: 27768305]
41. Corbett PTet al.Dynamic combinatorial chemistry. Chem. Rev106, 3652–3711 (2006). [PubMed: 16967917]
42. Furusho Yet al.Dynamic covalent approach to [2]- and [3]roxtanes by utilizing a reversible thiol-disulfide interchange reaction. Chem. Eur. J9, 2895–2903 (2003). [PubMed: 12868421]
43. Lu Set al.Directed Disulfide Pairing and Folding of Peptides for the De Novo Development of Multicyclic Peptide Libraries. J. Am. Chem. Soc142, 16285–16291 (2020). [PubMed: 32914969]
44. Otto S & Kubik S Dynamic combinatorial optimization of a neutral receptor that binds inorganic anions in aqueous solution. J. Am. Chem. Soc 125, 7804–7805 (2003). [PubMed: 12822990]
45. Ponnuswamy N, Cougnon FB, Clough JM, Pantos GD & Sanders JKM Discovery of an organic trefoil knot. Science 338, 783–785 (2012). [PubMed: 23139329]
46. Otto S, Furlan RL & Sanders JKM Selection and amplification of hosts from dynamic combinatorial libraries of macrocyclic disulfides. Science 297, 590–593 (2002). [PubMed: 12142534]
47. Carnall JMet al.Mechanosensitive self-replication driven by self-organization. Science327, 1502–1506 (2010). [PubMed: 20299594]
48. Bartolec B, Altay M & Otto S Template-promoted self-replication in dynamic combinatorial libraries made from a simple building block. Chem. Commun 54, 13096–13098 (2018).
49. Liu Bet al.Complex Molecules That Fold Like Proteins Can Emerge Spontaneously. J. Am. Chem. Soc141, 1685–1689 (2019). [PubMed: 30562021]
50. de Veer SJ, Kan MW . & Craik DJ Cyclotides: From Structure to Function. Chem. Rev 119, 12375–12421 (2019). [PubMed: 31829013]
51. Hegemann JD, Zimmermann M, Xie X & Marahiel MA Lasso peptides: an intriguing class of bacterial natural products. Acc. Chem. Res 48, 1909–1919 (2015). [PubMed: 26079760]
52. Bayro MJet al.Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. J. Am. Chem. Soc125, 12382–12383 (2003). [PubMed: 14531661]
53. Wilson KAet al.Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. J. Am. Chem. Soc125, 12475–12483 (2003). [PubMed: 14531691]
54. Pan SJ & Link AJ Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. J. Am. Chem. Soc 133, 5016–5023 (2011). [PubMed: 21391585]
55. Pavlova O, Mukhopadhyay J, Sineva E, Ebright RH & Severinov K Systematic structure-activity analysis of microcin J25. J. Biol. Chem 283, 25589–25595 (2008). [PubMed: 18632663]
56. Ducasse Ret al.Sequence determinants governing the topology and biological activity of a lasso peptide, microcin J25. Chembiochem13, 371–380 (2012). [PubMed: 22287061]
57. Zong C, Wu MJ, Qin JZ & Link AJ Lasso Peptide Benenodin-1 Is a Thermally Actuated [1]Rotaxane Switch. J. Am. Chem. Soc 139, 10403–10409 (2017). [PubMed: 28696674]
58. Blond Aet al.Thermolysin-linearized microcin J25 retains the structured core of the native macrocyclic peptide and displays antimicrobial activity. Eur. J. Biochem269, 6212–6222 (2002). [PubMed: 12473117]
59. Rosengren KJet al.Structure of thermolysin cleaved microcin J25: extreme stability of a two-chain antimicrobial peptide devoid of covalent links. Biochemistry43, 4696–4702 (2004). [PubMed: 15096038]
60. Atcher J & Alfonso I The effect of DMSO in the aqueous thiol–disulphide dynamic covalent chemistry of model pseudopeptides. RSC Adv. 3, 25605–25608 (2013).
61. Romuald C, Cazals G, Enjalbal C & Coutrot F Straightforward synthesis of a double-lasso macrocycle from a nonsymmetrical [c2]daisy chain. Org. Lett 15, 184–187 (2013). [PubMed: 23256863]
62. Safont-Sempere MM, Fernández G & Würthner F Self-Sorting Phenomena in Complex Supramolecular Systems. Chem. Rev 111, 5784–5814 (2011). [PubMed: 21846150]
63. Wasserman EThe Preparation of Interlocking Rings: A Catenane. J. Am. Chem. Soc82, 4433–4434 (1960).

64. Schill G & Lüttringhaus A The Preparation of Catena Compounds by Directed Synthesis. Angew. Chem. Int. Ed. Engl 3, 546–547 (1964).

65. Allen CDet al.Thermal Unthreading of the Lasso Peptides Astexin-2 and Astexin-3. ACS Chem. Biol11, 3043–3051 (2016). [PubMed: 27588549]

66. Roberts DA, Pilgrim BS & Nitschke JR Covalent post-assembly modification in metallosupramolecular chemistry. Chem. Soc. Rev 47, 626–644 (2018). [PubMed: 29142998]

67. Romano Met al.Structural Basis for Natural Product Selection and Export by Bacterial ABC Transporters. ACS Chem. Biol13, 1598–1609 (2018). [PubMed: 29757605]

68. Wedemeyer WJ, Welker E & Scheraga HA Proline cis-trans isomerization and protein folding. Biochemistry 41, 14637–14644 (2002). [PubMed: 12475212]

69. Trainor K, Palumbo JA, MacKenzie DWS & Meiering EM Temperature dependence of NMR chemical shifts: Tracking and statistical analysis. Protein Sci. 29, 306–314 (2020). [PubMed: 31730280]

70. Dietrich-Buchecker CO, Sauvage J-P & Kintzinger J-P Une nouvelle famille de molecules : les metallo-catenanes. Tetrahedron Lett. 24, 5095–5098 (1983).

71. Ashton PRet al.A [2] Catenane Made to Order. Angew. Chem. Int. Ed28, 1396–1399 (1989).

72. Jamieson EMG, Modicom F & Goldup SM Chirality in rotaxanes and catenanes. Chem. Soc. Rev 47, 5266–5311 (2018). [PubMed: 29796501]

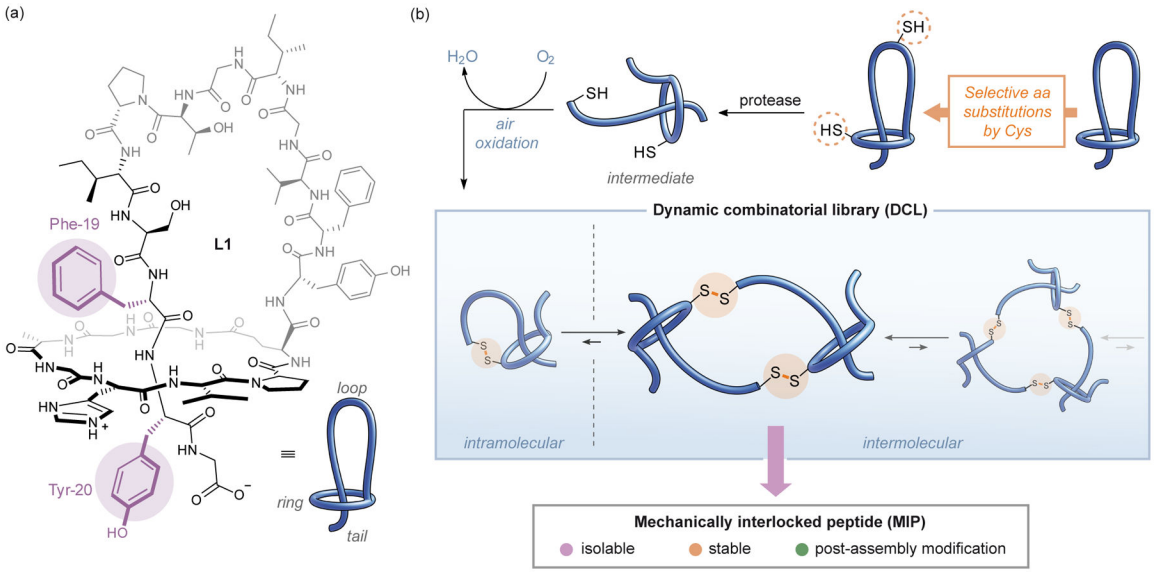


Figure 1: Strategy for a dynamic covalent self-assembly synthesis of MIPs.

(a) The mechanically interlocked structure of lasso peptide MccJ25 (**L1**) with highlighted stopper residues Phe-19 and Tyr-20 (*cis/trans* configuration of amide bonds is not taken into account). (b) Schematic representation of the synthesis protocol for MIPs made from engineered lasso peptides. After selective substitution of lasso peptide **L1** by Cys residues, enzymatic cleavage of a Cys-decorated variant in its loop region yields [2]rotaxane intermediates which undergo dynamic self-assembly into topologically complex MIPs via disulfide bonding. The positioning of Cys residues controls the pathway selection and thus the product distribution in the resulting DCL. A desired MIP can be isolated from the DCL, is stable and can be subsequently modified and functionalized in aqueous solution.

(a)

(b)

	ring	loop			tail
variant	5	12	13	18	21
L2	C	G	I	S	C
L3	H	C	I	S	C
L4	H	R	C	S	G
L5	C	R	C	S	G
L6	H	C	C	S	C
L7	C	R	C	S	C
L8	H	C	I	D	C

(c)

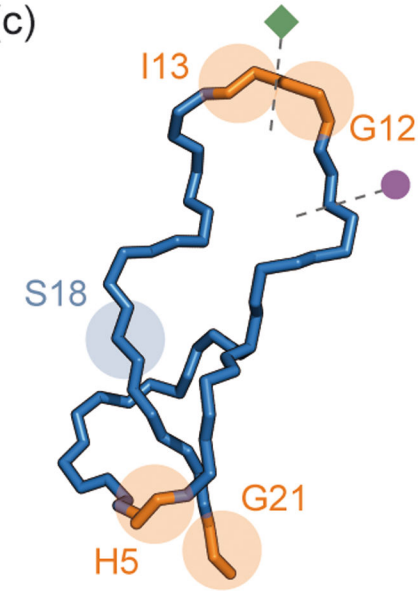


Figure 2: Synthesized set of Cys-decorated lasso peptide variants.

(a) Sequence of wild-type MccJ25 (**L1**) with highlighted substitution sites (orange) and predetermined cleavage sites (grey lines) in the loop region. Trypsin cleaves variants with a substitution by Arg-12 between Arg-12/Cys-13 (green rhombus) and thermolysin cleaves between Phe-10/Val-11 (purple circle). The link between Gly-1 and Glu-8 represents the isopeptide bond. (b) Set of engineered MccJ25 variants **L2**–**L8** used in this study. (c) Structure of wild-type MccJ25 (PDB code 1PP5) with highlighted substitution and cleavage sites.

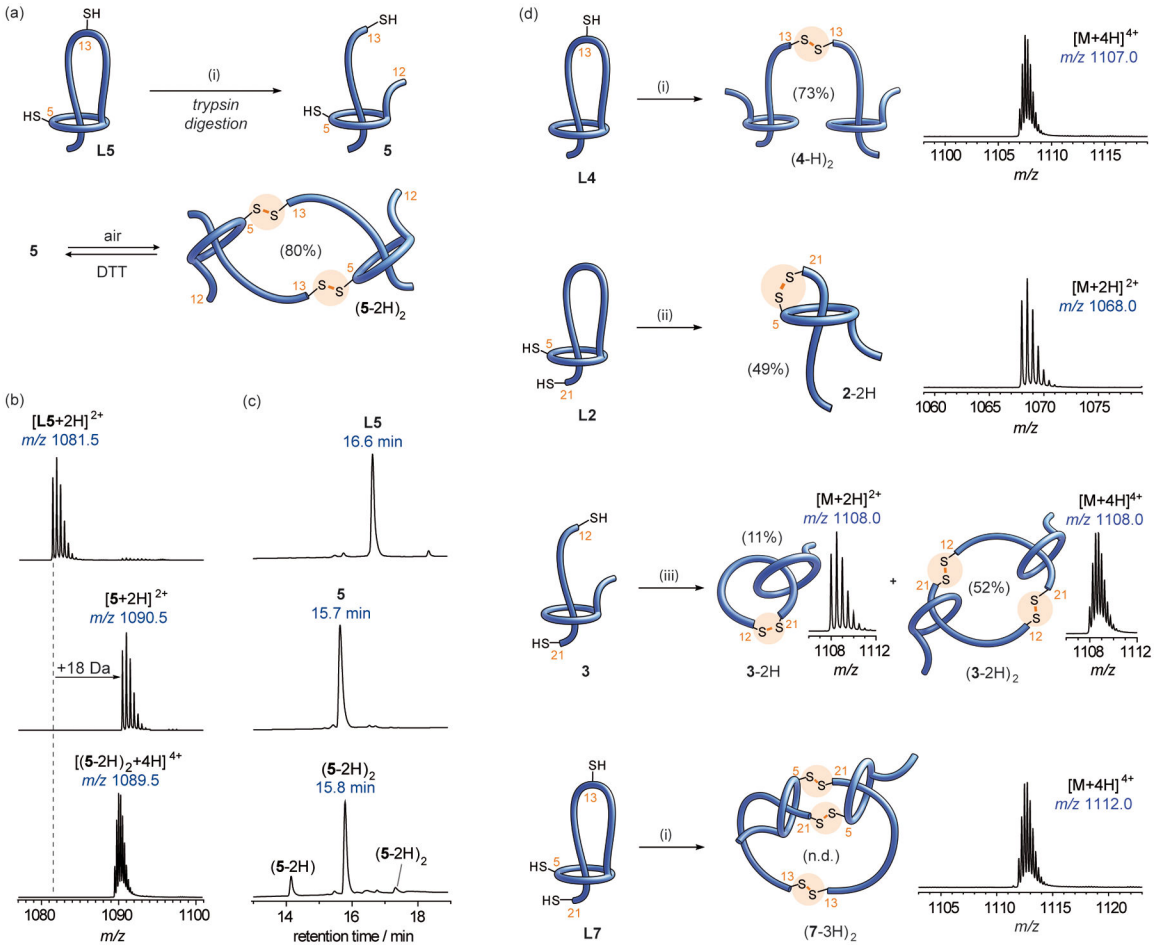


Figure 3: Dynamic covalent self-assembly synthesis of different MIPs.

(a) Schematic representation of lasso peptide **L5**, [2]rotaxane **5** and [c2]daisy chain (**5**-2H)$_2$. Addition of DTT quantitatively reduces dimer (**5**-2H)$_2$ back to building block **5**. (b) Electrospray ionization (ESI) mass spectra showing isotopic patterns of **L5** (top), **5** (middle) and (**5**-2H)$_2$ (bottom) and (c) the corresponding HPLC traces. The most abundant species after 9 days is the head-to-tail dimer (**5**-2H)$_2$ with the structure of a [c2]daisy chain (80% yield as determined by HPLC peak area). (d) Synthesis of [3]rotaxane (**4**-H)$_2$, left-handed lasso **2**-2H, [2]- and [3]catenane **3**-2H and (**3**-2H)$_2$ and double-lasso macrocycle (**7**-3H)$_2$. Yields were determined by integration of HPLC peak areas. HPLC traces can be found in the Supplementary Information (Supplementary Figs 5 and 33). Due to metastability of DCL (**7**-3H)$_n$, a yield could not be determined (n.d.). Reaction conditions: (i) trypsin, $(NH_4)HCO_3$ (50 mM), pH 7.8, air; (ii) thermolysin, Tris (50 mM), $CaCl_2$ (0.5 mM), pH 8.0, air; (iii) $(NH_4)HCO_3$ (50 mM), pH 7.8, air. Building blocks were assembled at a peptide concentration of 0.25 mM. DTT, dithiothreitol.

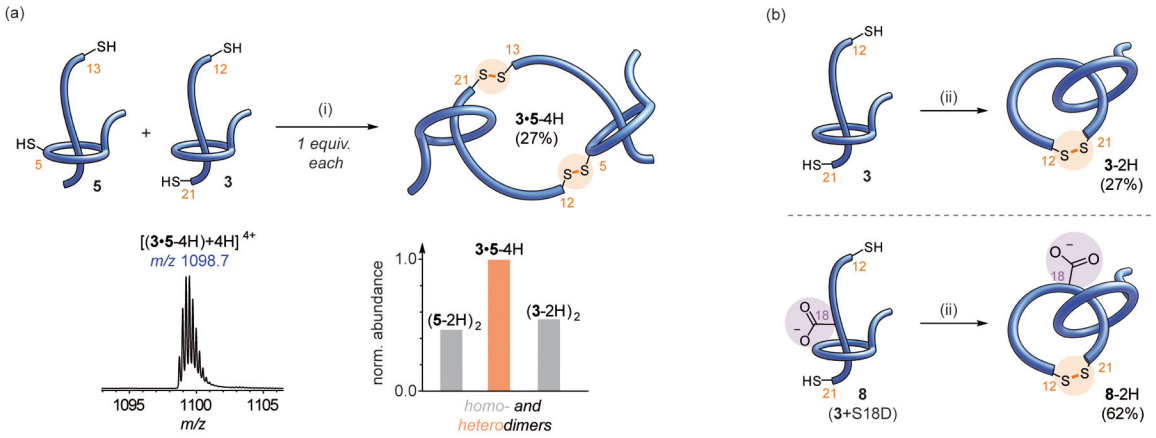


Figure 4: Mixed DCLs and building block modifications.

(a) Mixed DCL incorporating dithiol building blocks **3** and **5**.The hetereodimer **3•5**-4H (*m/z* 1098.7) is the most abundant constituent (see Supplementary Information for additional data). The bar chart displays the relative abundance of head-to-tail homo- and heterodimers in the DCL (**3**-2H)$_n$(**5**-2H)$_n$ based on LC/MS peak area (see Supplementary Fig. 7). (b) Pre-programming of catenane-forming building block **3** by a S18D substitution yields building block **8**. At equimolar concentrations, the DCL (**8**-2H)$_n$ is shifted towards the monomer compared to building block **3** without S18D substitution. Reaction conditions: $(NH_4)HCO_3$ (50 mM), pH 7.8, air, overall building block concentration of (i) 250 μM and (ii) 100 μM.

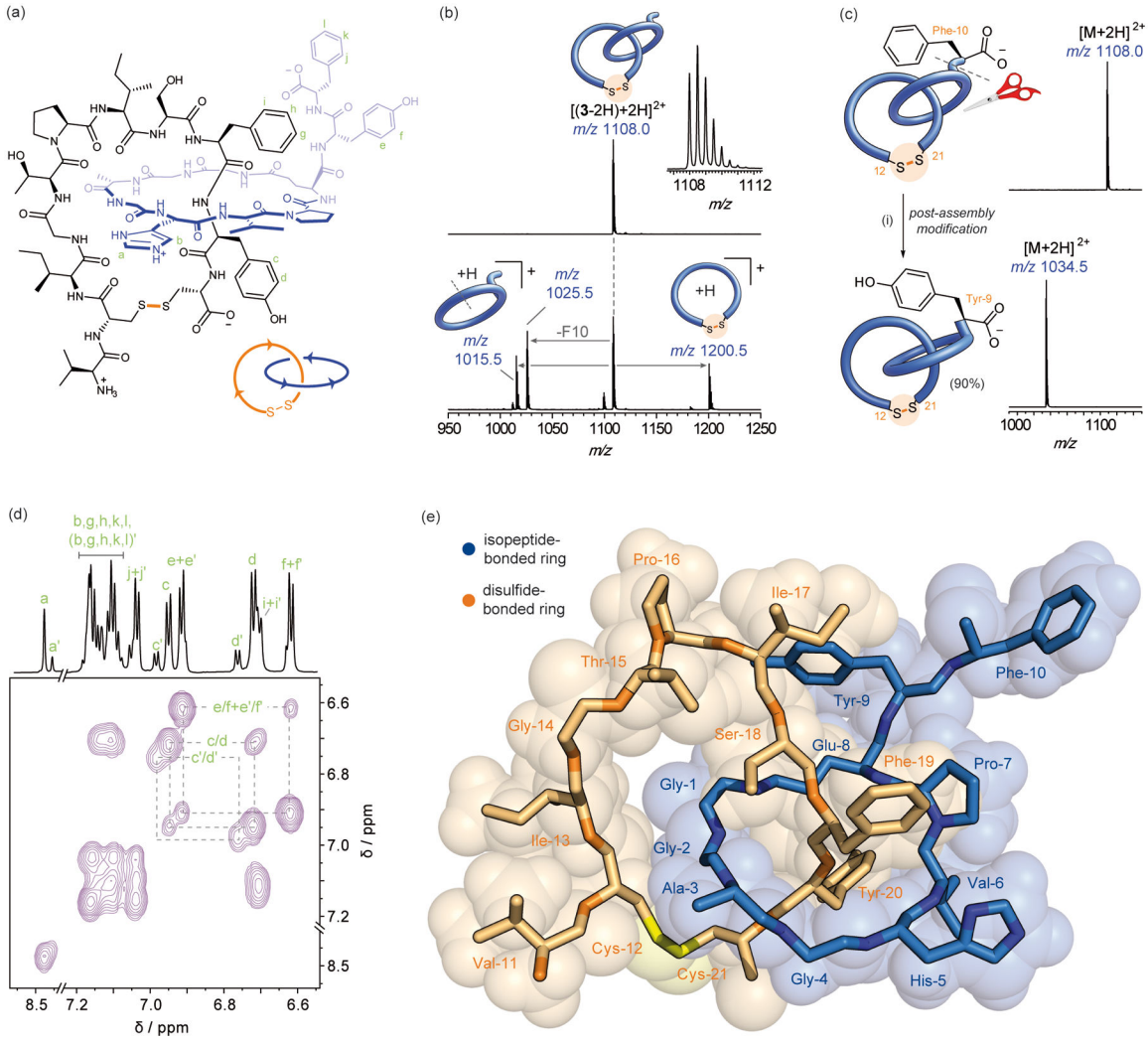


Figure 5: Structural characterization of peptide [2]Catenane 3-2H by NMR spectroscopy and mass spectrometry.

(a) Chemical structure of [2]catenane **3**-2H (*cis/trans* configuration of amide bonds is not taken into account) and schematic representation to indicate the directionality of the rings (arrows point from N- to C-terminus). (b) MS analysis of [2]catenane **3**-2H: (top) mass spectrum and (bottom) CID experiment (25 V collision voltage) with mass-selected $[(\mathbf{3}\text{-}2H)+2H]^{2+}$ ion. After cleavage of a covalent bond, the [2]catenane ion (*m/z* 1008.0) fragments into its ring subcomponents (*m/z* 1015.5 and 1200.5). (c) Schematic representation and MS spectra of post-assembly modification of **3**-2H by carboxypeptidase, which selectively removes the Phe-10 residue. Reaction conditions: (i) Carboxypeptidase B and carboxypeptidase Y, sodium acetate (50 mM, pH 6.0), 31 h, rt. (d) Partial ^{1}H NMR and $^{1}H,^{1}H$ TOCSY spectra (800 MHz, 295 K, H_2O/D_2O = 95:5) of [2]catenane **3**-2H (1.03 mM) show the overlap of two peak sets. Signals of the minor conformer are marked with an apostrophe. Correlations between the aromatic protons of Tyr-9 and Tyr-20 are highlighted in the TOCSY spectrum. (e) Solution NMR structure of the major co-conformation of [2]catenane **3**-2H in water. The backbone of all residues and selected side chains are

shown as sticks. Hydrogen and backbone oxygen atoms are removed for the sake of clarity, nitrogen atoms are shaded darker and the disulfide bond is colored yellow. The lowest energy structure of the conformational ensemble is shown. An overlay of the top 20 energy-minimized structures can be found in the Supplementary Information (Supplementary Fig. 44).