Metal-Peptidic Cages – Helical Oligoprolines Generate Highly Anisotropic Nanospaces with Emergent Isomer Control

Ben E. Barber,[1,2] Ellen M. G. Jamieson,[1,2] Leah E. M. White,[1,2] Charlie T. McTernan[1,*]

[1]Artificial Molecular Machinery Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK and Department of Chemistry, Britannia House, 7 Trinity Street, King's College London, SE1 1DB, UK. [2]Authorship is alphabetical. *Correspondence: charlie.mcternan@crick.ac.uk

Keywords: Supramolecular Chemistry; Metal-Organic Cage; Metal-Organic Capsule; Self-Assembly; Proline; PPII; Tetra-Pyridyl Palladium; Isomer control; Regioselectivity; Peptide

Abstract:

The self-assembly of metal-organic cages enables the rapid creation of atomically defined, three-dimensional, nanoscale architectures reminiscent of proteins. However, existing metal-organic cages are almost exclusively built from rigid and flat aromatic panels, limiting binding selectivity and, often, water solubility. Herein, we disclose a new class of cages - metal-peptidic cages - which utilise water-soluble, chiral and helical oligoproline strands of varying length to generate highly anisotropic nanospaces. Further, we find formation of the *cis* isomer of the cage is strongly favoured, and is an emergent property of using complex and chiral building blocks in the formation of defined nanospaces. We demonstrate that the use of peptidic building blocks allows us to rapidly tune the size of the nanospace formed, from c. 1 - 4 nm, and that the use of biologically relevant components enables targeted binding of therapeutic molecules, highlighting the potential of these systems for selective drug delivery.

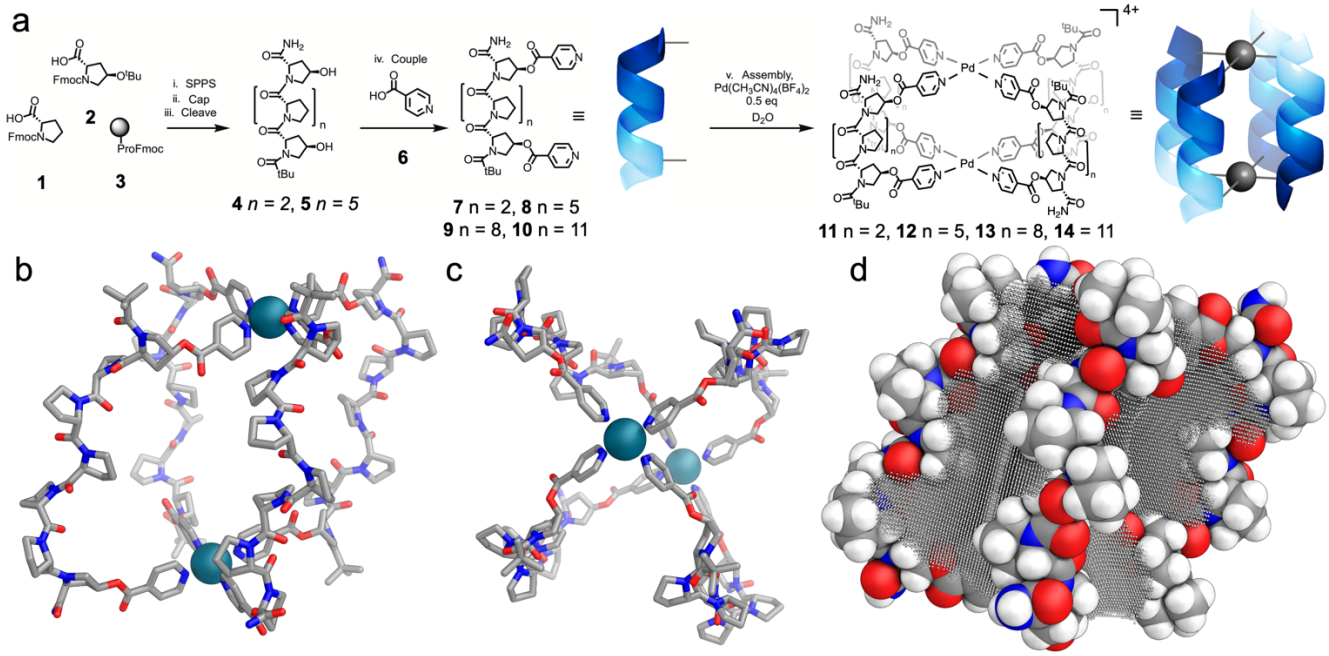

Figure 1. a. Synthesis of Metal-Peptidic Cages. i. Solid phase peptide synthesis (see SI Section 2 for protocol). ii. Pivalic Anhydride:CH_2Cl_2:DMF 1:4.5:4.5, r.t., 45 min. iii. Trifluoroacetic acid:Triisopropyl Silane:H_2O 38:1:1, r.t, 2 hr. iv. EDCI (4 eq), DMAP (2 eq), Isonicotinic acid **6** (4 eq), CH_2Cl_2, r.t., 16 hr. Ligands **9** and **10** synthesised by submitting crude cleaved peptide directly to coupling conditions. v. $Pd(CH_3CN)_4(BF_4)_2$ (0.5 eq), D_2O, r.t., <5 min. b. Side view and c. top view of molecular model of cage **12**, generated from the crystal structure of a hexaproline,[50] using molecular mechanics simulations to find energy minima. d. Cavity of cage **12** calculated with Molovol, volume = 4138 Å^3.

Advances in self-assembly have enabled the construction of increasingly complex architectures from simple building blocks.[1] Recent work has focused on the construction of simple metal-organic polyhedra, due to their reliable and predictable assembly from rigid, flat and conjugated aromatic systems.[2-4] These metal-organic cages contain defined internal cavities, and their functions often stem from guest binding in this space, the properties of which can significantly diverge from the general solution.[5-12] Metal-organic cages been applied in sensing,[13] separation,[14] catalysis[15] and drug delivery,[16] and many researchers have explored the rules that control their assembly, thus enabling rational design.[17]

Despite these advances, the ubiquitous use of rigid, flat, conjugated building blocks limits the wider applicability of metal-organic cages.[18] The extended planar surfaces of ligands often make their assembly in water difficult or impossible, limiting biomedical applications.[19-21] Further, the flat ligands present a uniform surface to the interior cavity, and so the selectivity of guest binding is limited, as cavities often approximate to spheres.[1,22,23] This contrasts strikingly with biology, where evolution has furnished a vast array of well-defined and exquisitely selective binding architectures, which operate in a crowded milieu, built from the sequence polymers of life - peptides, D/RNA, and glycans.[24-27] The current generation of metal-organic cages are unable to match the achievements and specificity of biological binders due to the limited repertoire of building blocks used in their construction.[28-33]

Inspired by both the limitations of current methods, and the advantages of biological systems, we considered alternative ways to generate reliable coordination vectors, without using extended aromatic building blocks. Peptides present a particularly versatile platform, as they are intrinsically chiral, often highly water soluble, and easy to synthesise from commercially available building blocks in a modular and automated way by solid phase peptide synthesis.[34] However, due to their unconjugated backbones, they have an extremely high degree of conformational freedom, compared to a biphenyl axis.[35] In metal-organic cages, this would typically be expected to lead to the formation of poorly defined polymers rather than cages. Di-, tri- and tetrapeptides have been used to form gels and metal-organic frameworks, and longer peptides have been used to generate porous extended frameworks, but reports of discrete structures are extremely limited.[36-40] Flexible peptides have been used by Sawada and Fujita to great effect, enabling generation of extraordinarily complex, woven molecular knots and links, but at the cost of designability.[41-45] Longer peptides, particularly those with defined secondary structures, have been shown to create non-covalent aggregates of micrometre scale,[46-48] and metal-organic or non-covalent extended peptide frameworks showing novel isomerism and tuneable guest binding.[37,38,49]

Oligoprolines with six or more repeat units reliably form polyproline II (PPII) structures in aqueous solution with a repeat length of c. 9 Å, where every third residue aligns.[50] This structure is relatively rigid, tolerates substitution, and is often used as a 'molecular ruler' due to these properties.[51] Inspired by prior work using PPII structures to generate triaxial weaves,[47] porous extended frameworks,[39] length controlled oligomerisation,[46] and metal-organic frameworks,[49] we decided to use the rigidity and stability of the PPII-fold to design and synthesise discrete metal-organic cages.

We decided to use a proline rod with seven residues -(*trans*-4-hydroxyproline)-(proline)$_5$-(*trans*-4-hydroxyproline (HPPPPPH, **5**) – as previous reports had suggested that six residues were the minimum required to ensure stable helix formation in aqueous solution.[52] The two *trans*-4-hydroxyproline (Hyp/H) residues are aligned on the same face of an idealised PPII helix, enabling attachment of metal binding moieties (Figure S226-S227).

We synthesised HPPPPPH **5** using standard solid-state microwave assisted peptide synthesis techniques, *N*-capping with a *tert*-butyl carbonyl group, to prevent competitive binding of a free amine to palladium(II) and to act as a reporter signal in ^{1}H NMR (SI, Section 3). We protected the *C*-terminus as an amide to reduce the risk of the carboxylic acid interfering in the self-assembly process.[53] We cleaved and globally deprotected the peptide, to furnish free HPPPPPH **5**. We confirmed the stability of the polyproline II structure in HPPPPPH **5** by circular dichroism (CD), observing a characteristic negative peak at 205 nm and a positive peak at 225 nm (Figure S12). Initial modelling (SI Section 7), based on the single previously reported PPII oligoproline crystal structure,[50] suggested that isonicotinic acid (**6**) would provide suitable binding vectors with square planar metals to favour the formation of the targeted discrete molecular cages over polymeric species. Therefore, we coupled isonicotinic acid (**6**) to each free Hyp hydroxy group in **5**. Purification by preparatory HPLC furnished ligand **8** in 46% yield based on loaded resin. We again confirmed the folding of ligand **8** to the PPII structure by CD (Figure S22). As a population of the *cis* isomer of ligand **8** would be detrimental to self-assembly (as it would introduce a kink in the ligand, creating diverging metal coordination vectors), we probed the *cis*/*trans* propensity of ligand **8** by varying ^{1}H NMR solvent. In D_2O, a single set of peaks corresponding to the all *trans* isomer was seen (Figure 2a) whereas in d_3-MeCN a complex spectra was seen, corresponding to *cis*-*trans* isomerism (Figure S214). The contrasting case of d_3-MeCN supports the assignment of **8** as preferring all *trans* in D_2O.

With these results in hand, we screened self-assembly conditions. Assemblies in d_3-MeCN were unsuccessful, likely due to the significant population of *cis* peptide bonds observed in ligand **8** in d_3-MeCN (Figure S214) favouring polymeric species. However, in D_2O, we observed the formation of a single species when ligand **8** was assembled using $Pd(CH_3CN)_4(BF_4)_2$ in a

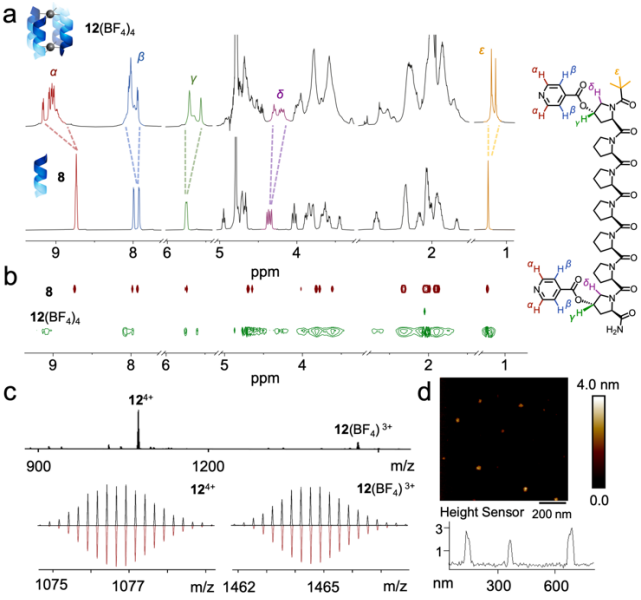

Figure 2. a. ^{1}H NMR (600 MHz, D_2O, 298 K) of cage **12** (top) and ligand **8** (bottom). b. ^{1}H DOSY NMR (600 MHz, D_2O, 298 K). Hydrodynamic radius of 13 Å for **8** and 21 Å for **12**, correlating well to molecular modelling. c. HRMS of cage **12**, showing 4+ and 3+ ions and their isotopic distributions (recorded, top; simulated, bottom). d. AFM of cage **13**, showing a height of 2-3 nm, in agreement with molecular modelling.

precisely 4:2 ligand:metal ratio (Figure 2a). Small deviations in stoichiometry led to partial assembly or a complex population of sub-species, which simplified when stoichiometries were corrected by internal standards, and highly purified ligand was used (total impurities <2% by LCMS, SI Section 3).

Significant downfield shifts were observed in the pyridyl protons (*α*, *β*, Figure 2a), corresponding to palladium(II) coordination, and desymmetrisation of these signals and the *tert*-butyl capping group was seen (*ε*, Figure 2a). DOSY ^{1}H NMR spectroscopy showed formation of a single species whose diffusion constant was approximately twice that of the free ligand **8**, with a hydrodynamic radius of 21 Å matching well with molecular modelling (Figure 2b). High resolution electrospray ionisation mass spectrometry (ESI-HRMS) showed +3 and +4 charge states of **12**, and the isotopic distributions matched well with simulated values (Figure 2c). A corresponding assembly could be formed with $Pd(NO_3)_2.2H_2O$ whose characterisation data matched well with those of cage **12**$(BF_4)_4$ (Figures S71 – S81). However, the assembly proved sensitive to anion identity – assemblies with $PdCl_2$ or $PdSO_4$ proved unsuccessful, generating complex mixtures or precipitation. We explored the sensitivity of the reaction to palladium(II) addition by titrating $Pd(CH_3CN)_4(BF_4)_2$ into ligand **8**. We saw the clear formation of cage **12** for sub-stoichiometric quantities, suggesting a level of cooperativity to cage assembly (Figure S2). For superstoichiometric quantities, a more complex regime was seen, with multiple species present (including cage **12** and Pd_2**8**$_2(BF_4)_4$ species). We confirmed cage stability down to 0.2 mM (below which partial disassembly was seen, Figure S4). CD confirmed that the PPII structure of ligand **8** was

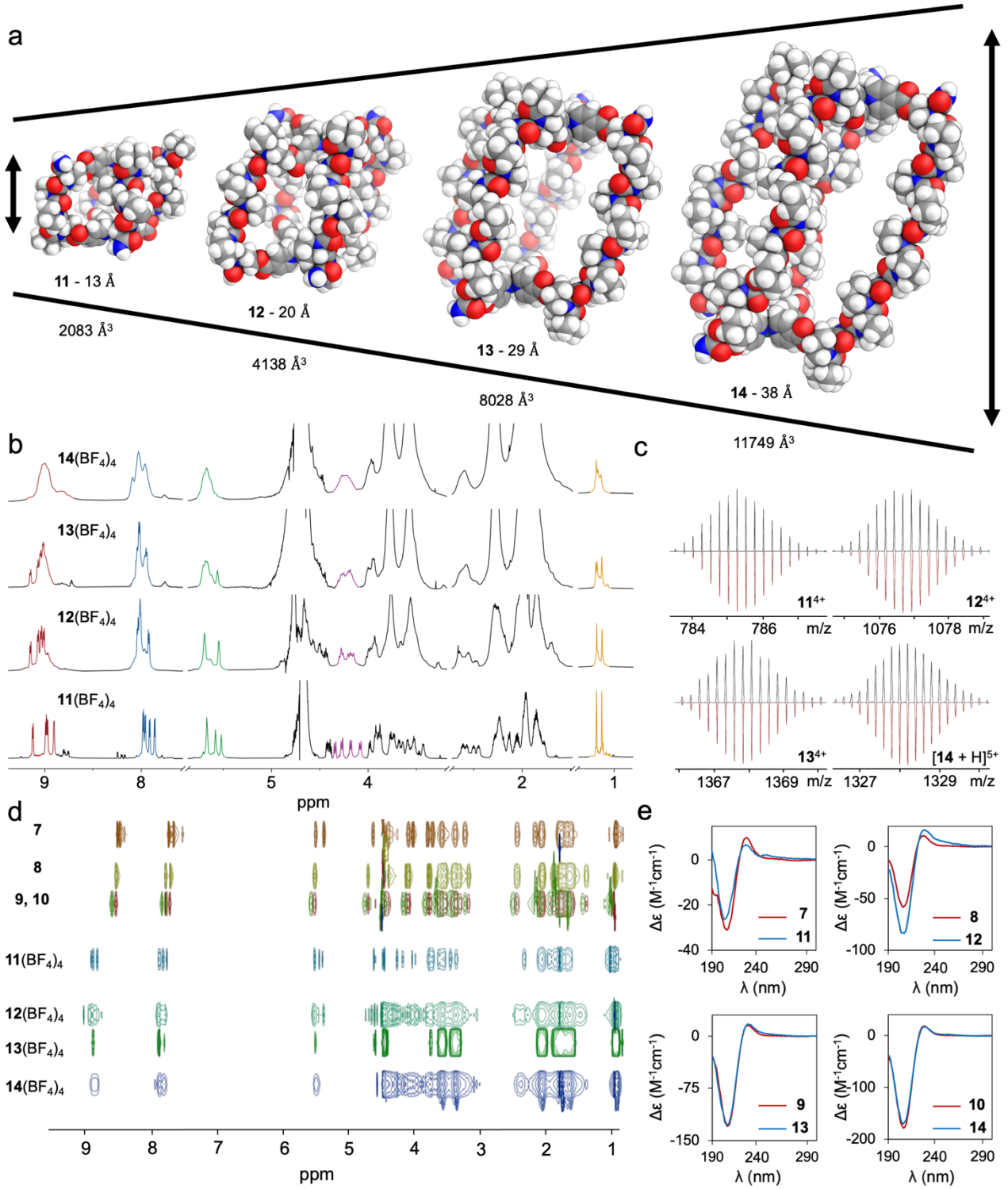

Figure 3. a. Molecular models of cages **11**-**14**, showing calculated lengths and volumes. b. ^{1}H NMR (600 MHz, D_2O, 298K) of cages **14** (top) to **11** (bottom). c. High-Resolution Mass Spectrometry data for **11**-**14**. Observed in black, simulated in red. d. ^{1}H DOSY NMR (600 MHz, D_2O, 298K). Hydrodynamic radius of: 11 Å for **7**, 13 Å for **8**, 14 Å for **9** and **10**, 17 Å for **11**, 21 Å for **12**, 24 Å for **13** and 28 Å for **14**, correlating well with molecular modelling. e. Circular Dichroism of **7**-**14** (273 K) with absorbances of cages **11**-**14** standardised to effective ligand concentration. Characteristic peaks at 205 nm and 225 nm are observed in all cases.

preserved in cage **12** (Figure 3e). ^{1}H NMR spectra (Figure 2a) showed a twofold desymmetrisation of the ligand (i.e. two peaks per proton environment), and an increased differentiation between *N*- and *C*-termini, which was particularly well resolved in the pyridine signals (*α*, *β*) and the *tert*-butyl reporter group (*ε*, Figure 2a, S69, S82). This desymmetrisation suggested that we had formed the *cis* isomer of cage (see SI Section 6), but the complexity of the spectra of cage prevented us from assigning the cause of desymmetrisation with certainty at this point (*vide infra*). Over 2000 attempts to grow crystals, using vapour diffusion, slow evaporation, layering, hanging drop, and protein crystallisation screening methodologies failed, underlining the difficulties in crystallising PPII structures, as others have reported.[50]

To explore the programmability of our system, and showcase the modularity of our approach, we

synthesised both short and longer analogues of ligand **8**, removing or adding three proline residues each time, targeting ligands with 4 (**7**), 10 (**9**), 13 (**10**) or 16 Pro/Hyp residues. This modularity is a key advantage of our system – extending typically used aromatic systems is often synthetically non-trivial - and the known, dependable, repeat length of oligoprolines means that we can reliably design and synthesise cages of varying sizes, in a similar way to the use of DNA origami in nanotechnology.[25,55] The targeted ligands were rapidly synthesised by solid-state peptide synthesis. We found that the 16mer irreversibly aggregated after cleavage from the resin, and so could not be explored further. We expected ligand **8**, containing only four Hyp and Pro residues, to lack a PPII structure in solution, given previous reports,[52] and so to fail to assemble a cage, acting as a negative control. However, we found that HPPH **4** and coupled ligand **7** both showed strong folding in solution.[39] As such, we assembled cages **11**, **13** and **14** in a similar manner to cage **12**, by addition of a carefully controlled stoichiometry of $Pd(CH_3CN)_4(BF_4)_2$. 13mer **10** was challenging to handle, showing a greater propensity to aggregation, and a lower solubility.

We collected ^{1}H NMR, DOSY, ESI-HRMS and CD on the successful assemblies, **11**, **13** and **14** (see SI Section 4). In each case, the data were consistent with cage assembly, with the cage^{4+} and cage^{3+} peaks seen by HRMS (along with matching isotopic distributions, Figure 3c), characteristic shifts in the ^{1}H NMR on palladium(II) binding (Figure 3b), and DOSY (Figure 3d) spectra reflecting a gradually increasing hydrodynamic radius consistent with that predicted by molecular models (Figure 3a and SI Section 7). AFM confirmed that nanostructures with heights consistent with predictions from molecular modelling formed and were preserved on surface deposition for cage **13** (Figure 2d and Figures S127 – S130, height c. 2.5 nm).[56] For cages **13** and **14**, scaled CD intensity and maxima matched that of free ligands **9** and **10**, as expected for PPII structures maintained between ligand and cage. Interestingly, in **11** we saw a decrease in CD intensity upon cage assembly, and conversely, in **12**, we saw an enhancement of CD intensity on assembly. We attribute this to the small size of **11** causing a distortion of the PPII structure of **7**, decreasing absorption intensity; whilst in **12**, assembly reinforces folding to the PPII structure, as folding becomes cooperative (PPI would not be accommodated in the cage architecture), enhancing absorption intensity. We were able to identify a minor (c. 2-9% by NMR) species in the ^{1}H NMRs of cages formed from $Pd(NO_3)_2.2H_2O$ as a $Pd_2\mathbf{L}_3$ species by HRMS (Figures S82 + S83), which agrees well with the ^{1}H NMR data. This species could form either due to steric clash between ligands,[18] or the electron-withdrawing ester *para* to the coordinating pyridine nitrogen making solvent coordination more competitive. This may also explain the very rapid cage assembly (<5 mins, 298 K)[3] we observe.

As such, the simplicity of solid-state peptide synthesis enabled us to rapidly prototype a family of metal-peptidic cages with lengths ranging from 13 - 38 Å and volumes of c. 2000 Å^3 to 11750 Å^3, whilst retaining the same basic architecture; showcasing a great advantage of building metal-organic cages from intrinsically tuneable sub-components using a repeating, modular system.

Furthermore, these extended and contracted systems allowed us to assign the cause of the twofold desymmetrisation of ligand signals we observe in the ^{1}H NMR spectra of our cages. Each showed twofold desymmetrisation compared to the corresponding ligand spectra, with the resolution of differentiation of the *N*-terminus and *C*-terminus residues increasing with decreasing ligand length. To maximise peak dispersion and aid our assignment, we undertook full characterisation of cage $\mathbf{11}(NO_3)_4$ on a 950 MHz ^{1}H NMR instrument.[57] To our delight, the significant increase in peak dispersion enabled us to confirm the cause of desymmetrisation across cages **11**, **12**, **13** and **14**.

Our ligands are enantiopure, helical, and have an intrinsic directionality (i.e. the *C*-terminus is distinct from the *N*-terminus). As such, while our cage is homoleptic,[58-65] there are four distinct cage isomers which can form (Figure 4a): where the *C*-termini are all aligned and sit on one end (i.e. bind one palladium(II) ion) of the cage (denoted *CCCC*), where three *C*-termini and one *N*-terminus lie at one end (denoted *CCCN*), and two cases where two *C*-termini and two *N*-termini are at each end, either with *C*-termini *cis* or *trans* to each other across the palladium(II) centre (*CCNN* and *CNCN* respectively). Along with this, further dynamic species could occur by *cis*/*trans* isomerism of individual proline amides[51] or of the terminal *tert*-butylcarbonyl groups (SI Section S6).[66]

We were particularly interested to see the resolution of the H_α and H_β pyridine peaks to form 8 doublets, with some partially overlapping. In the free ligand, the expected four doublets (again, partially overlapped) are seen, due to ^3J coupling to the neighbouring $CH_{\alpha/\beta}$, and the differentiation of the *C*-terminal and *N*-terminal Hyp residues. The eight equal intensity doublets seen for H_α and H_β in **11** provides a first hint to the isomerism present. We also saw two singlets for the *tert*-butyl protons H_ε, and both of these patterns are consistent across VT ^{1}H NMR experiments over the accessible range in D_2O (Figure S216).

These patterns show that cage **11** contains two separate and equal ligand environments. The only isomer consistent with this is the *CCNN* isomer, where the 'top' and 'bottom' metal binding faces are equivalent. Each of the four pyridines around a metal is in a distinct environment – two are *C*-termini, and two *N*-termini, providing a first differentiation. The helicity of the proline rods, and how they relate to one another through space, leads to diastereotopic splitting between the aligned (i.e. the *C*-termini) ligands on the 'left' and the 'right' of the pair, matching our experimental

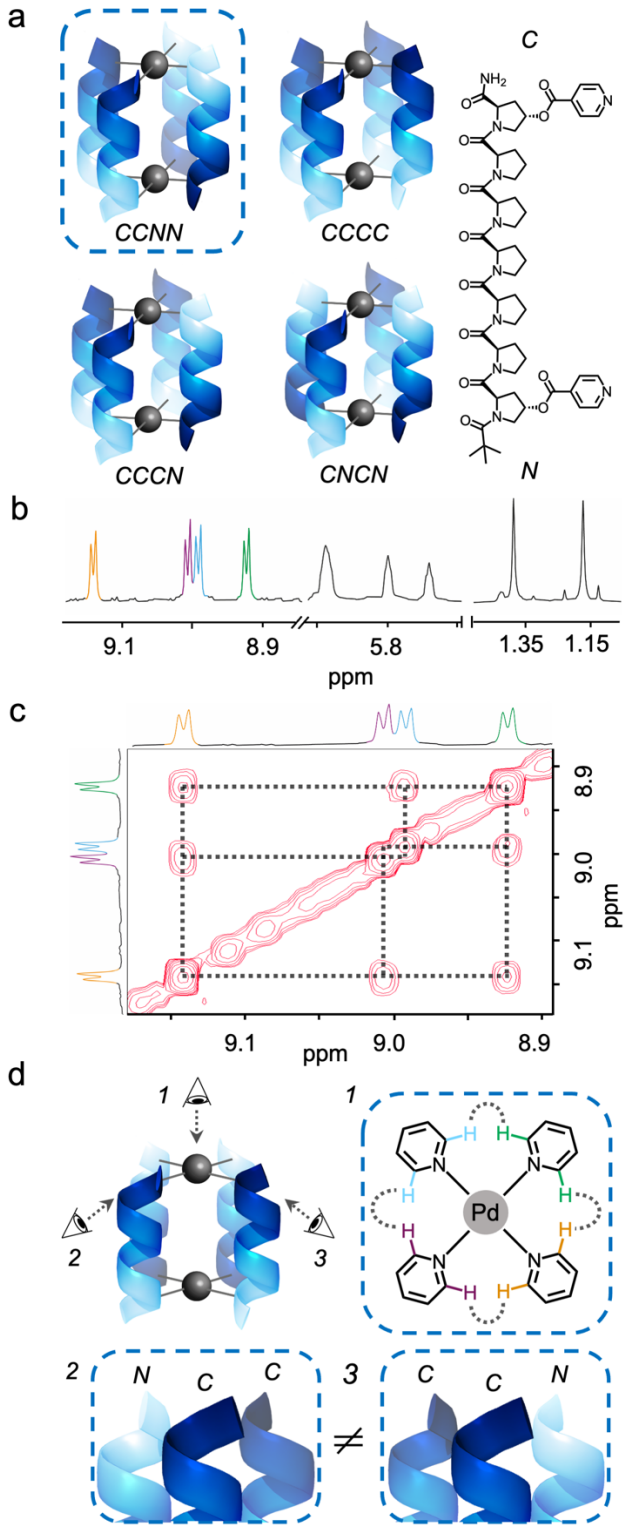

Figure 4. a. Representations of potential cage structural isomers, and ligand **8** with *C*- and *N*-termini highlighted. b. ^{1}H NMR (950 MHz, D_2O, 298 K) of cage **11**, highlighting desymmetrisation of cage peaks. Distinct H$_\alpha$ environments are shown in different colours. c. ^{1}H NOESY (950 MHz, D_2O, 298 K) correlations between H$_\alpha$ environments, with each H$_\alpha$ environment seeing two of the three remaining environments. d. Representation of the identified H$_\alpha$ environments (*1*), and spatial reasoning for distinction of the two aligned (i.e. *CC*) ligands (*2*,*3*).

observations (Figure 4b,c,d). Both the *CCCC* and *CNCN* contain a single equivalent ligand environment, which would lead to four aromatic doublets, while the *CCCN* would have four separate ligand environments and so 16 aromatic doublets (see SI Section 6 detailed discussion and reasoning). Each H$_\alpha$ environment shows two NOESY correlations with other, different, H$_\alpha$ environments, which can only be explained by a *CCNN cis* arrangement of the ligands (Figure 4d).

This is an extraordinary level of regiocontrol, with a single structural isomer of cage formed from a ligand with a high degree of conformational freedom. The system is capable of forming a vast array of possible species, considering the cage structural isomers, *tert*-butyl and proline *cis*/*trans* isomers (of which there are 28 separate, potentially isomerisable, bonds in cage **12**).

Furthermore, this is an emergent property of the system; we had not designed our ligands to form a single isomer, and indeed expected this to be a long term and challenging project.[67] By using complex, chiral, and helical building blocks in our self-assembly process, rather than observing an intractable morass, we instead see clean formation of a single species, underlining the need for, and advantages of, supramolecular chemists embracing complex and chiral building blocks.

To further support this assignment, we undertook extensive molecular modelling studies of cage **12**, based on the previously reported crystal structure of a hexaproline chain.[50] We had initially expected the *CCCC* aligned isomer to be lowest in energy, reasoning this would best align pyridine coordination vectors. However, when molecular models were made of each isomer, the *CCNN cis* isomer was found to be significantly lower in energy (SI Section 7) than any other isomer. Close examination of the molecular model provides potential reasons for this selective stabilisation. Firstly, in the *CCNN cis* isomer the cage tilts, forming an oblique rhombic prism rather than the expected cube. This geometry is not possible in any other isomer, enabling the *CCNN cis* to relieve torsional strain. Secondly, orientation of the carbonyl groups of the isonicotinic acid seems essential to enable this selectivity – the lowest energy system is obtained when the isonicotinic carbonyl of the *N*-terminus points into the cage cavity, and that of the *C*-terminus points away (and so the carbonyls of a single strand point in the same direction). Other permutations were found to be strongly disfavoured. We propose that steric clash between the carbonyl and the *tert*-butyl groups at the *N*-terminus favours these carbonyls facing inwards (Figure S234); which then dictates the orientation of the *C*-terminus carbonyl.[68] We probed this effect by synthesising an analogue of ligand **8** with an acetyl *N*-terminus protecting group rather than the *tert*-butyl carbonyl (Figures S141-S148), anticipating that decreasing the steric bulk at this position should reduce the preference for the internally facing carbonyl group, and so reduce control of the isomer distribution. Upon assembly, we saw a significantly decreased level of isomer selection, with more species present in the ^{1}H NMR, supporting our reasoning for the emergent isomer control seen in this system (Figures S224 + S225).

Having established that cage **12** contained an unusual and large cavity, we next investigated its binding behaviour towards a range of anionic and neutral

guests by ^{1}H NMR. We first found binding to a range of traditional metal-organic cage guests - small hydrocarbons (Figure S198 – S205), anions (SI Section 5.1) and dyes (Figures S206 + S207). Cage **12** showed little affinity for a range of sugars, which we attribute to competitive guest solvation in D_2O.

Having established that common classes of guests bind cage **12**, we focused our attention on unusual classes more suited to taking advantage of the peptidic framework of cage **12**. Nucleotide phosphates such as ATP and AMP caused cage precipitation. However, we found that amino acids and dipeptides such as Boc-Ala-Ala-OH **24** bound to cage **12** (Figure 5 and S189-S195). We were surprised to see binding to Boc-Ala-Ala-OH **24** in water, as previous reports have shown a preference for incorporation of two carboxylate and two pyridine ligands around palladium(II) centres.[53] Removal of the Boc group, and attempted binding of NH_2-Ala-OH was unsuccessful, causing cage disassembly, although Boc-Lys-OH **25** was tolerated (Figure S194), suggesting that a chelating interaction with unprotected amino acids is responsible for disassembly, rather than the presence of an amine. Boc-Phe-Phe-OH **22**, Boc-Ala-OH **23** and Boc-Hyp-OH **26** were also bound.

Having established that cage **12** could bind to amino acids and dipeptides, we explored whether therapeutic molecules could be bound, as this area has been poorly explored given the critical importance of drug delivery and release in modern biomedicine.[69] Cage **12** is uniquely suited to tackle this problem due to its unusual

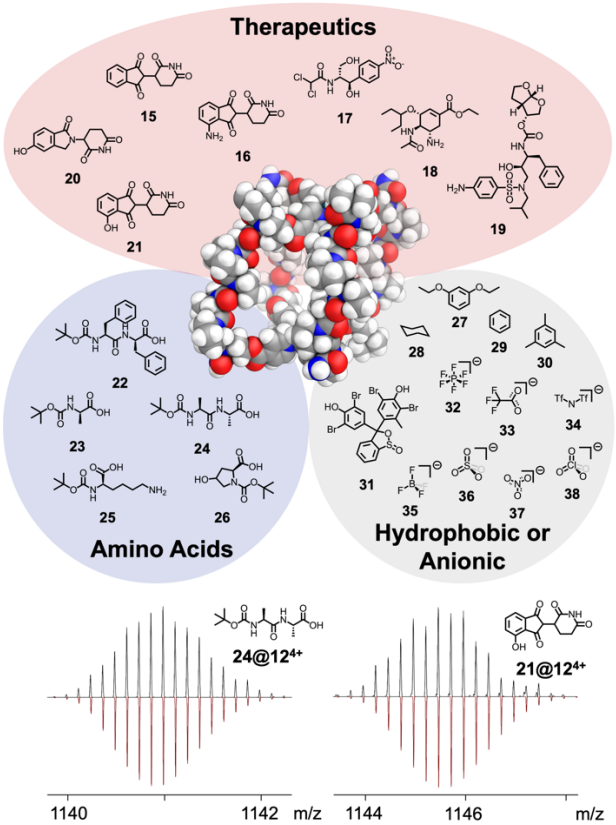

Figure 5. Host-Guest chemistry of **12**. A range of anionic and neutral molecules are bound, with some associations confirmed by HRMS showing 1:1 binding. Where guest solubility allowed, we obtained binding isotherms (SI Section S5).

constituent chemistry, biologically derived building blocks, and water solubility. Previous approaches have typically used binding induced aggregation, or metal-organic helices, often to great effect.[70-73] We were delighted to find that an array of FDA approved drugs bound cage **12** (Figure 5) including anti-retroviral Darunavir (**19**, Figure S175), antibiotic chloramphenicol (**17**, Figure S171), anti-viral Oseltamivir (**18**, Figure S173), and perhaps most excitingly, the molecular glue thalidomide and its analogues (**15**, **16**, **20**, **21**, Figures S165 – S170 + S177 – S182), which is both an anti-cancer drug and a ubiquitous subunit in CRBN ligase recruiting PROTACs. In certain cases, we were even able to observe association by HRMS (Figure 5 and S167 – S168 + S181 – S182), which is associated with strong binding. Given the known association of oligoprolines with specific cellular proteins,[74] the binding of therapeutics within self-assembled metal-peptidic cages provides a novel avenue to selective drug delivery which will be explored in future work. Consideration of the molecular model of cage **12** furnishes a justification for the binding of these unusual guests – the chiral patterning of hydrophobic and hydrophilic areas means that the internal cavity of **12** does not resemble the extended aromatic surfaces common in metal-organic cages, and instead delivers an open and textured surface for association, more reminiscent of the protein targets of approved drug molecules.[75]

Herein, we have shown that the rigid and flat aromatic panels typically used in metal-organic cages can be replaced by peptides, whose defined secondary structures in solution provide the requisite rigidity. We chose to demonstrate this principle using oligoprolines, given the rich literature on their use as molecular rulers. We were able to assemble a range of Pd_2L_4 metal-peptidic cages, using oligoproline rods of different lengths, with two Hyp residues modified with metal binding isonicotinic acid units. The complexity of our ligands led to emergent structural isomer control, with a single isomer of cage formed by emergent isomer selection, showcasing the power of self-assembly in amplifying individually weak interactions. We showed that these metal-peptidic cages bound to an array of different classes of guests, including biologically relevant peptides, drugs, and therapeutics, taking advantage of their patterned, chiral and helical internal surfaces. Future work will look to diversify the peptides used, enabling selective cavity modulation, and leverage the binding of therapeutics for targeted drug delivery applications.

Supplementary Information

The Supplementary Information is available free of charge online.

Author Information

Corresponding Author

* Charlie T. McTernan (charlie.mcternan@crick.ac.uk)

Author Contributions

BEB, EMGJ and LEMW conducted all experimental research and analysed the results with CTM. CTM conceived, designed and supervised the project, drafted the paper, and obtained funding. All authors edited the final manuscript.

Funding Sources

This work was supported by the Francis Crick Institute which receives its core funding from Cancer Research UK (FC0011221), the UK Medical Research Council (FC0011221), and the Wellcome Trust (FC0011221) and by the UKRI under Future Leaders Fellowship MR/W00657X/1.

Acknowledgements

CTM acknowledges that this work was supported by the Francis Crick Institute which receives its core funding from Cancer Research UK (FC0011221), the UK Medical Research Council (FC0011221), and the Wellcome Trust (FC0011221) and by the UKRI under Future Leaders Fellowship MR/W00657X/1. We thank Dr Richard Thorogate, UCL for recording AFM data, Dr James Jarvis, Dr Alain Oregioni and Dr Geoff Kelly for their advice and support with NMR spectroscopy. This work was supported by the Francis Crick Institute through provision of access to the MRC Biomedical NMR Centre. The Francis Crick Institute receives its core funding from Cancer Research UK (CC1078), the UK Medical Research Council (CC1078), and the Wellcome Trust (CC1078).

References

1. McTernan, C. T., Davies, J. A., & Nitschke, J. R. Beyond Platonic: How to Build Metal-Organic Polyhedra Capable of Binding Low-Symmetry, Information-Rich Molecular Cargoes. *Chem. Rev.* **122**, 10393–10437 (2022).

2. Seidel, S. R., & Stang, P. J. High-Symmetry Coordination Cages *via* Self-Assembly. *Acc. Chem. Res.* **35**, 972–983 (2002).

3. Han, M., Engelhard, D. M., & Clever, G. H. Self-assembled Coordination Cages Based on Banana-shaped Ligands. *Chem. Soc. Rev.* **43**, 1848–1860 (2014).

4. Fujita, M., & Ogura, K. Supramolecular Self-Assembly of Macrocycles, Catenanes, and Cages through Coordination of Pyridine-Based Ligands to Transition Metals. *Bull. Chem. Soc. Jpn.* **69**, 1471–1482 (1996).

5. Rizzuto, F. J., von Krbek, L. K. S., & Nitschke, J. R. Strategies for Binding Multiple Guests in Metal–Organic Cages. *Nat. Rev. Chem.* **3**, 204–222 (2019).

6. Pluth, M. D., Bergman, R. G., & Raymond, K. N. Acid Catalysis in Basic Solution: A Supramolecular Host Promotes Orthoformate Hydrolysis. *Science* **316**, 85–88 (2007).

7. Cullen, W., Misuraca, M. C., Hunter, C. A., Williams, N. H., & Ward, M. D. Highly Efficient Catalysis of the Kemp Elimination in the Cavity of a Cubic Coordination Cage. *Nat. Chem.* **8**, 231–236 (2016).

8. Ubasart, E. *et al.* A Three-shell Supramolecular Complex Enables the Symmetry-mismatched Chemo- and Regioselective Bis-functionalization of C60. *Nat. Chem.* **13**, 420–427 (2021).

9. Howlader, P., Das, P., Zangrando, E., & Mukherjee, P. S. Urea-Functionalized Self-Assembled Molecular Prism for Heterogeneous Catalysis in Water. *J. Am. Chem. Soc.* **138**, 1668–1676 (2016).

10. Ngai, C., Sanchez-Marsetti, C. M., Harman W. D., & Hooley, R. J. Supramolecular Catalysis of the Oxa-Pictet–Spengler Reaction with an Endohedrally Functionalized Self-Assembled Cage Complex. *Angew. Chem. Int. Ed.* **59**, 23505–23509 (2020).

11. Wang, Q. *et al.* Self-assembled Nanospheres with Multiple Endohedral Binding Sites Pre-organize catalysts and substrates for highly efficient reactions. *Nat. Chem.* **8**, 225–230 (2016).

12. Preston, D., Sutton, J. J., Gordon, K. C., & Crowley, J. D. A Nona-nuclear Heterometallic Pd_3Pt_6 'Donut'-Shaped Cage: Molecular Recognition and Photocatalysis. *Angew. Chem. Int. Ed.* **57**, 8659–8663 (2018).

13. Yamashina, M. *et al.* Preparation of Highly Fluorescent Host–Guest Complexes with Tunable Color upon Encapsulation. *J. Am. Chem. Soc.* **137**, 9266–9269 (2015).

14. Zhang, D., Ronson, T. K., Zou, Y.-Q. & Nitschke, J. R. Metal–organic Cages for Molecular Separations. *Nat. Rev. Chem.* **5**, 168–182 (2021).

15. Piskorz, T. K., Martí-Centelles, V., Spicer, R. L., Duarte, F., & Lusby, P. J. Picking the Lock of Coordination Cage Catalysis. *Chem. Sci.* (2023) Advance Article, doi: https://doi.org/10.1039/D3SC02586A.

16. Ahmad, N., Younus, H. A., Chughtai, A. H. & Verpoort, F. Metal–Organic Molecular Cages: Applications of Biochemical Implications. *Chem. Soc. Rev.* **44**, 9–25 (2015).

17. Chakrabarty, R., Mukherjee, P. S. & Stang, P. J. Supramolecular Coordination: Self-Assembly of Finite Two- and Three-Dimensional Ensembles. *Chem. Rev.* **111**, 6810–6918 (2011).

18. Pullen, S. & Clever, G. H. Mixed-Ligand Metal–Organic Frameworks and Heteroleptic Coordination Cages as Multifunctional Scaffolds - a Comparison. *Acc. Chem. Res.* **51**, 3052–3064 (2018).

19. Percástegui, E. G., Ronson, T. K. & Nitschke, J. R. Design and Applications of Water-Soluble Coordination Cages. *Chem. Rev.* **120**, 13480–13544 (2020).

20. Burke, B. P. *et al.* Visualizing Kinetically Robust $Co^{III}_4L_6$ Assemblies in Vivo: SPECT Imaging of the Encapsulated [^{99m}Tc]TcO_4^- Anion. *J. Am. Chem. Soc.* **140**, 16877–16881 (2018).
21. Cosialls, R. *et al.* PET Imaging of Self-Assembled ^{18}F-Labelled Pd_2L_4 Metallacages for Anticancer Drug Delivery. *Chem. Eur. J.* **29**, e202202604 (2022).
22. Zhang, W. *et al.* Controlling the Recognition and Reactivity of Alkyl Ammonium Guests Using an Anion Coordination-Based Tetrahedral Cage. *J. Am. Chem. Soc.* **140**, 5248–5256 (2018).
23. Zhong, W. *et al.* A Platinum(II)-Based Molecular Cage with Aggregation-Induced Emission for Enzymatic Photocyclization of Alkynylaniline. *Angew. Chem. Int. Ed.* **62**, e202214577 (2022).
24. Brudno, Y. & Liu, D. R. Recent Progress Toward the Templated Synthesis and Directed Evolution of Sequence-Defined Synthetic Polymers. *Chem. Biol.* **16**, 265–276 (2009).
25. Seeman, N. C. & Sleiman, H. F. DNA Nanotechnology. *Nat. Rev. Mat.* **3**, 17068 (2017).
26. Yang, H. *et al.* Metal–Nucleic Acid Cages. *Nat. Chem.* **1**, 390–396 (2009).
27. Edwardson, T. G. W., Tetter, S. & Hilvert, D. Two-tier Supramolecular Encapsulation of Small Molecules in a Protein Cage. *Nat. Commun.* **11**, 5410 (2020).
28. Xue, Y. *et al.* Catalysis within Coordination Cages. *Coord. Chem. Rev.* **430**, 213656 (2021).
29. Sen, S. K. & Natarajan, R. Influence of Conformational Change and Interligand Hydrogen Bonding in a Chiral Metal–Organic Cage. *Inorg. Chem.* **58**, 7180–7188 (2019).
30. Jurček, O. *et al.* Superchiral Pd_3L_6 Coordination Complex and its Reversible Structural Conversion into $Pd_3L_3Cl_6$ Metallocycles. *Angew. Chem. Int. Ed.* **54**, 15462–15467 (2015).
31. Samantray, S., Krishnaswamy, S. & Chand, D. K. Self-assembled Conjoined-Cages. *Nat. Commun.* **11**, 880 (2020).
32. Catti, L., Sumida, R. & Yoshizawa, M. Aqueous Polyaromatic Receptors for Biomolecules with High Selectivity. *Coord. Chem. Rev.* **460**, 214460 (2022).
33. Mesquita, L. M. *et al.* Palladium(II)-Mediated Assembly of a M_2L_2 Macrocycle and M_3L_6 Cage from a Cyclopeptide-Derived Ligand. *Org. Lett.* **21**, 6442–6446 (2019).
34. Behrendt, R., White, P. & Offer, J. Advances in Fmoc Solid-phase Peptide Synthesis. *J. Pept. Sci.* **22**, 4–27 (2016).
35. Huang, F. & Nau, W. M. A Conformational Flexibility Scale for Amino Acids in Peptides. *Angew. Chem. Int. Ed.* **42**, 2269–2272 (2003).
36. Mondal, S., Das, S. & Nandi, A. K. A Review on Recent Advances in Polymer and Peptide Hydrogels. *Soft Matter* **16**, 1404–1454 (2020).
37. Heinz-Kunert, S. *et al.* Assembly of π-Stacking Helical Peptides into a Porous and Multivariable Proteomimetic Framework. *J. Am. Chem. Soc.* **144**, 7001–7009 (2022).
38. Hess, S. T. *et al.* Noncovalent Peptide Assembly Enables Crystalline, Permutable, and Reactive Thiol Frameworks. *J. Am. Chem. Soc.* **145**, 19588–19600 (2023).
39. Brightwell, D. F. *et al.* A Reversibly Porous Supramolecular Peptide Framework. *Chem. Eur. J.***28**, e202202368 (2022).
40. Brightwell, D. F., Truccolo, G., Samanta, K., Shepherd, H. J. & Palma, A. Supramolecular Self-Assembly of Engineered Polyproline Helices. *ACS Macro Lett.* **12**, 908–914 (2023).
41. Inomata, Y., Sawada, T. & Fujita, M. Metal-Peptide Torus Knots from Flexible Short Peptides. *Chem* **6**, 294–303 (2020).
42. Inomata, Y., Sawada, T. & Fujita, M. Metal–Peptide Nonafoil Knots and Decafoil Supercoils. *J. Am. Chem. Soc.* **143**, 16734–16739 (2021).
43. Sawada, T. & Fujita, M. Folding and Assembly of Metal-Linked Peptidic Nanostructures. *Chem* **6**, 1861–1876 (2020).
44. Sawada, T., Inomata, Y., Shimokawa, K. & Fujita, M. A Metal–Peptide Capsule by Multiple Ring Threading. *Nat. Commun.* **10**, 5687 (2019).
45. Yamagami, M., Sawada, T. & Fujita, M. Synthetic *β*-Barrel by Metal-Induced Folding and Assembly. *J. Am. Chem. Soc.* **140**, 8644–8647 (2018).
46. Lewandowski, B. M., Schmid, D., Borrmann, R., Zetschok, D., Schnurr, M., Wennemers, H. Catalytic Length-Controlled Oligomerization with Synthetic Programmable Templates. *Nat. Synth.* **2**, 331–337 (2023).
47. Lewandowska, U. *et al.* A Triaxial Supramolecular Weave. *Nat. Chem.* **9**, 1068–1072 (2017).
48. Lewandowska, U. *et al.* Hierarchical Supramolecular Assembly of Sterically Demanding π-Systems by Conjugation with Oligoprolines. *Angew. Chem. Int. Ed.* **53**, 12537–12541 (2014).
49. Schnitzer, T., Paenurk, E., Trapp, N., Gershoni-Poranne, R. & Wennemers, H. Peptide–Metal Frameworks with Metal Strings Guided by Dispersion Interactions. *J. Am. Chem. Soc.* **143**, 644–648 (2021).
50. Wilhelm, P., Lewandowski, B., Trapp, N. & Wennemers, H. A Crystal Structure of an Oligoproline PPII-Helix, at Last. *J. Am. Chem. Soc.* **136**, 15829–15832 (2014).
51. Dobitz, S., Aronoff, M. R. & Wennemers, H. Oligoprolines as Molecular Entities for Controlling

Distance in Biological and Material Sciences. *Acc. Chem. Res.* **50**, 2420–2428 (2017).
52. Kakinoki, S., Hirano, Y. & Oka, M. On the Stability of Polyproline-I and II Structures of Proline Oligopeptides. *Polym. Bull.* **53**, 109–115 (2004).
53. Zheng, Y. *et al.* A Facile Approach toward Multicomponent Supramolecular Structures: Selective Self-Assembly *via* Charge Separation. *J. Am. Chem. Soc.* **132**, 16873–16882 (2010).
54. Yan, X., Cook, T. R., Wang, P., Huang, F. & Stang, P. J. Highly Emissive Platinum(II) Metallacages. *Nat. Chem.* **7**, 342–348 (2015).
55. Yamashina, M. *et al.* An Antiaromatic-Walled Nanospace. *Nature* **574**, 511–515 (2019).
56. Sun, B. *et al.* From Ring-in-Ring to Sphere-in-Sphere: Self-Assembly of Discrete 2D and 3D Architectures with Increasing Stability. *J. Am. Chem. Soc.* **137**, 1556–1564 (2015).
57. Comparing NMRs across with increasing magnet size, as coupling constants remain a constant hertz value, the splitting in ppm decreases and so dispersion between overlapping multiplets increases.
58. Lewis, J. E. M., Tarzia, A., White, A. J. P. & Jelfs, K. E. Conformational Control of Pd_2L_4 Assemblies with Unsymmetrical Ligands. *Chem. Sci.* **11**, 677–683 (2020).
59. Lewis, J. E. M. Pseudo-Heterolepticity in Low-Symmetry Metal-Organic Cages. *Angew. Chem. Int. Ed.* **61**, e202212392 (2022).
60. Lewis, J. E. M. Multi-functional, Low Symmetry Pd_2L_4 Nanocage Libraries. *Chem. Eur. J* .**27**, 4454–4460 (2021).
61. Tarzia, A., Lewis, J. E. .M. & Jelfs, K. E. High-Throughput Computational Evaluation of Low Symmetry Pd_2L_4 Cages to Aid in System Design. *Angew. Chem. Int. Ed.* **60**, 20879–20887 (2021).
62. Li, R.-J. *et al.* Orientational Self-Sorting in Cuboctahedral Pd Cages. *Chem. Sci.* **13**, 11912–11917 (2022).
63. Lisboa, L. S., Findlay, J. Y., Wright, L. J., Hartinger, C. G. & Crowley, J. L. A Reduced-Symmetry Heterobimetallic $[PdPtL_4]^{4+}$ Cage: Assembly, Guest Binding, and Stimulus-Induced Switching. *Angew. Chem. Int. Ed.* **59**, 11101–11107 (2020).
64. Wu, K., Benchimol, E., Baksi, A. & Clever, G. Evolution of Multicomponent $[Pd_2ABCD]$ Cages (2023) https://doi.org/10.26434/chemrxiv-2023-5gb4q.
65. Ogata, D. & Yuasa, J. Dynamic Open Coordination Cage from Nonsymmetrical Imidazole–Pyridine Ditopic Ligands for Turn-On/Off Anion Binding. *Angew. Chem. Int. Ed.* **58**, 18424–18428 (2019).
66. Dugave, C. & Demange, L. *Cis−Trans* Isomerization of Organic Molecules and Biomolecules: Implications and Applications. *Chemical Reviews* **103**, 2475–2532 (2003).
67. Wang, J., Li, H. & Xu, B. Biological Functions of Supramolecular Assemblies of Small Molecules in the Cellular Environment. *RSC Chem. Biol.* **2**, 289–305 (2021).
68. Morris, D. L. *et al.* A Molecular Communication Channel Consisting of a Single Reversible Chain of Hydrogen Bonds in a Conformationally Flexible Oligomer. *Chem* **7**, 2460–2472 (2021).
69. Moreno-Alcántar, G. & Casini, A. Bioinorganic Supramolecular Coordination Complexes and their Biomedical Applications. *FEBS Lett.* **597**, 191–202 (2022).
70. Melidis, L. *et al.* Supramolecular Cylinders Target Bulge Structures in the 5' UTR of the RNA Genome of SARS-CoV-2 and Inhibit Viral Replication. *Angew. Chem. Int. Ed.* **60**, 18144–18151 (2021).
71. Craig, J. S. *et al.* Organometallic Pillarplexes that Bind DNA 4-Way Holliday Junctions and Forks. *J. Am. Chem. Soc.* **145**, 13570–13580 (2023).
72. Howson, S. E. *et al.* Optically Pure, Water-Stable Metallo-Helical 'Flexicate' Assemblies with Antibiotic Activity. *Nat. Chem.* **4**, 31–36 (2012).
73. Faulkner, A. D. *et al.* Asymmetric Triplex Metallohelices with High and Selective Activity against Cancer Cells. *Nat. Chem.* **6**, 797–803 (2014).
74. Mompeán, M., Oroz, J. & Laurents, D. V. Do Polyproline II Helix Associations Modulate Biomolecular Condensates? *FEBS Open Bio* **11**, 2390–2399 (2021).
75. Anderson, A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* 10, 787–797 (2003).