

# Encoding stereochemical molecular information on cyclophanes using non-directional interactions and achiral building blocks

Yuan Zhang, Benjamin Ourri, Pierre-Thomas Skowron, Emeric Jeamet, Ana M. Belenguer, Nicolas Vanthuyne, Olivier Cala, Pradeep K. Mandal, Christian Duchamp, Elise Dumont, Florent Perret, Laurent Vial, Ivan Huc, Julien Leclaire,

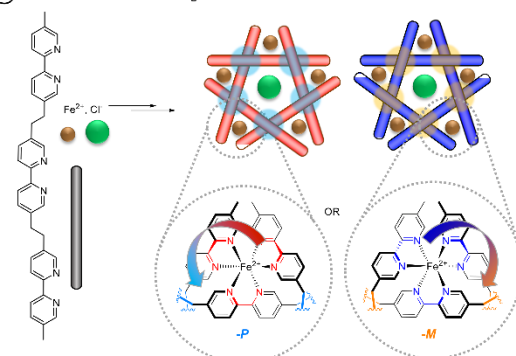
The stereoselective assembly of achiral constituents through a single spontaneous process into complex covalent architectures bearing multiple stereogenic elements currently seems out of reach to the synthetic chemist. It even seems beyond what nature itself has managed to attain through evolution. Here, we show that such an extreme level of control can be achieved by molecular programming, *i.e.* by implementing stereo-electronic information on synthetic organic building blocks and exploiting the features of the covalent reactions and interactions, whose interplay acts as a powerful assembling algorithm. Remarkably, we show that non-directional bonds and interactions can reliably transfer this information, delivering in near to physiological conditions, high-molecular weight macrocyclic species carrying up to 8-bits of conformational and configuration information. Beyond the field of supramolecular chemistry, this proof of concept should stimulate the on-demand production of highly structured polyfunctional architectures.

**Introduction.** The concept of *Molecular Information* emerged with the rise of supramolecular chemistry<sup>1–3</sup>. For instance, complementary molecular information is what underpins the host and guest binding in the lock and key paradigm. When combined to a reliable algorithm (the set of assembling covalent and non-covalent bonds), the molecular information borne by small building blocks can propagate at the nanoscopic level through information transfer, yielding highly ordered and organized architectures such as grids, MOFs or interlocked systems<sup>4–9</sup>. One of the key features of the algorithm lies in the reversibility of the linkages, which enables error correction<sup>10,11</sup>.

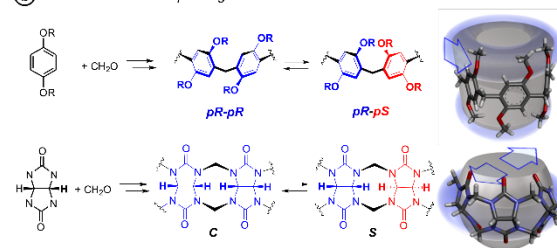
A seminal example of information transfer was brought by J.M. Lehn, with the self-assembling of circular metal helicates (figure 1a)<sup>12–14</sup>. The first level of information is compositional, with the chloride anion templating the formation of a hosting pentagonal cavity among all possible sizes and topologies (grids, helices) of organometallic architectures. The second level, which remains poorly commented, even in revisited versions of this work<sup>15</sup>, is stereochemical. The coordination requirement from the metal cation, the poly-denticity and the rigidity of the tris(bipyridine) strands are three inputs which are translated at the supramolecular level by the metal-ligand assembling chemistry (the algorithm). The first two inputs are encoded into chiral Werner-like complexes involving three different ligand strands on the edges of the cavity. Their  $\Delta$  (respectively  $\Lambda$ ) central configuration is stereoselectively interrelated (as a translation of the third input) to windings of  $M$  (respectively  $P$ ) neighbouring helices at the vertices of the pentagone. Remarkably, this process stereoselectively yields the  $(\Delta-M)_5/(\Lambda-P)_5$  mixture among all possible pentameric systems. Thus,

supramolecular chirality, defined as the *nonsymmetric arrangement of constituents in a non-covalent assembly*<sup>16</sup>, has spontaneously emerged as an information output. This seminal example allows for the formulation of the following take-home messages

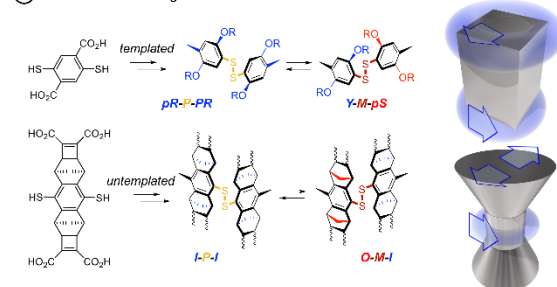
## (a) Seminal work : discovering molecular information



## (b) Previous work : deciphering molecular information



## (c) This work : encoding molecular information



**Figure 1: Information transfer during macrocyclization:** a) seminal example of anion-templated

pentahelicate self-assembling from achiral components with interdependent helical and central configurations (bicolour arrow); b) pillararenes and cucurbiturils stereoselectively self-assemble into isotactic species ( $pR$ )<sub>n</sub> and ( $pS$ )<sub>n</sub> and all-C facial configuration respectively); c) in the present work, stereoisomerism arises from monomeric faces positioning with respect to the cavity and disulfide bridge configuration. Blue arrows and highlights materialize intra- and intermolecular information transfer.

First, stereochemical information (in this example, the ( $A-P$ )<sub>5</sub> and ( $A-M$ )<sub>5</sub> configurational sequence) can be encoded at the supramolecular level, from simple achiral constituents. To reliably encode and propagate this stereochemical information, directional and reversible linkages (herein metal-ligand, the coding algorithm) enable error correction. In fact, nanoscopic self-organization of MOFs, helicates, nucleic acids, cages and knots is mostly involving directional reversible linkages such as hydrogen bonds, azomethine connections and/or metal-ligand interactions<sup>10,17,18</sup>. The last message is that, although the linear topography was more frequently chosen by nature and man (to engineer natural and synthetic foldamers), molecular information can also be efficiently encoded within (pseudo) macrocyclic species displaying a well-defined cavity<sup>19,20</sup>.

Despite their central role in molecular recognition, cavitands and cyclophanes have rarely been considered as substrates for the encryption, propagation and storage of molecular information. During their assembly, most attention is focused on the control of their size (the compositional level). Yet, two recent blockbusters, namely pillar[n]arenes and cucurbiturils, display striking examples of stereoconvergent assembly<sup>21–28</sup>. For pillar[n]arenes, although there are two equiprobable  $pR$  and  $pS$  configurations for each monomeric unit (figure 1b), potentially yielding  $n(n-1)/2$  stereoisomers, only the homochiral pair is obtained<sup>29</sup>. The information input (prochiral flat monomer with lateral hindrance in future positions 2 and 5) is encoded by the assembling chemistry (methylene bridging through reversible Friedel Craft reactions<sup>30–32</sup>) into a homochiral output, which minimizes the steric hindrance on both rims (figure 1b, blue arrow and highlight). Similarly, while the convex/concave glycosuril faces can potentially be oriented toward the cavity or toward the outer medium, the hindrance of the axial CH protons stands as the input, which is encoded by the same assembling algorithm (figure 1b, blue arrow and highlight) into the fully concave output, limiting inverted cucurbiturils to trace amounts<sup>33–35</sup>. In both cases, *a posteriori* analyses indicate that the stereoselectivity of the assembling process is guided by steric interactions which

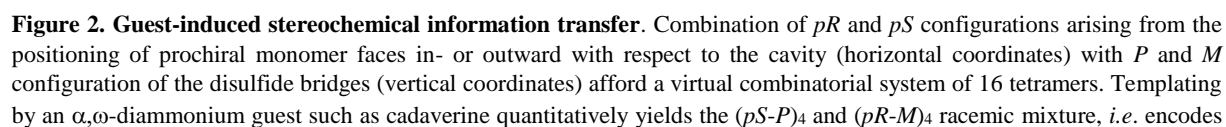
propagates (figure 1b, blue arrow and highlight) either during chain growth (cucurbiturils) or ring closure (pillararenes)<sup>36</sup>. In the field of disulfide based cavitands, which can be assembled by dynamic combinatorial strategies, examples of stereoselective assembling remain scarce and almost systematically involve chiral building blocks<sup>37–40</sup>. Template-induced (by casting and/or molding) selection and amplification of different facial stereoisomers of polydisulfide cyclotetramers was reported by Otto et al.<sup>38,41</sup> Yet, the rationale enabling to program on demand the configuration and conformation of a macrocycle by exploiting the stereochemical features of its building blocks and of the template used has remained out of reach so far. In particular, the role and relative configuration of the disulphide linkage between two stereogenic monomeric units have never been explored, even within protein architectures.<sup>42,43</sup> During the past decade, we have designed and studied the disulfide-based analogues of pillar[n]arenes named dyn[n]arenes (figure 1c up)<sup>43</sup>. Interestingly, thiol-to disulphide cyclooligomerization is accompanied by the creation of two classes of stereogenic elements: the monomeric faces and the disulphide bridges linking them. Once the covalent assembling is complete, the constitutional lability (disulphide exchange) is frozen and the only potential remaining exchange processes are of conformational nature. A central question is how to control these stereogenic elements formed i.e. how to program beforehand their configuration and how to preserve them from racemization over time.

We herein report the straightforward encryption and storage of up to 8-bits of stereochemical supramolecular information within such macrocyclic architectures. We show that this process can be directed either by a template or by the design of the building blocks themselves. Remarkably, we demonstrate that non-directional reversible linkages and steric interaction can indeed be powerful algorithm to reliably translate and propagate information from the molecular to the supramolecular level.

## Result and discussion

**Template-assisted stereochemical information encoding.** Recently, we showed that the diversity of the libraries generated from building block **A**, both in terms of macrocycle ring size and facial stereoisomers (figure 2), is largely underestimated by chromatographic analyses<sup>44</sup>. Although this technique is universally adopted for disulfide-based DCLs, it only enables to detect the trace amount of the homochiral tetramer **A4-a** (i.e., the strict

respectively **A4-a**, **A4-b** and **A4-c** (no (*pSpR*)<sub>2</sub> tetramer could be detected upon thiol-to-disulfide oxidation of **A**), were modelled by a combination of DFT and MD calculations (supporting section II.1.e). These two classes of simulations respectively yielded the populations of most stable conformers corresponding to each configuration as well as an evaluation of their conformational flexibility. The structures obtained and the repartition of the dihedral angles of the disulfide bridges of each species are gathered in tables S4-S9). The first main information which could be extracted from the calculations was that heterochiral species **A4-b** and **A4-c** have a marked tendency to adopt a flattened conformation wherein the cavity is too narrow to be accessed by any guest whatever its size (figure S10-S11).



8 bits of stereochemical information on the host (i). Top and side view of the crystallographic structure (with Van der Waals radii in contact) reveals contacts between substituents of the phenyl rings (ii), and crystalizes into a conglomerate (iii) wherein homochiral packing occurs (*i.e.* (*pR*)<sub>4</sub> and (*pS*)<sub>4</sub> species self-sort).

This fully agrees with the external binding mode observed experimentally with polyamine partners<sup>44</sup>. It also indicates that the formation of a pseudo-rotaxane by complex guest inclusion does not represent a sufficient energetic benefit to compensate the unfavourable flattened-to-open induced fit. The second main observation we made is that all stereoisomers **A4-x** (x= a-c), despite splitting each into several conformers **A4-x,n** (n = 1-4) of various stability, share a common configurational-conformational inter-dependence: *pR-P* and *pS-M* facial-dihedral combinations seem to be systematically the most stable for steric reasons. When conflicts arise (a disulfide bridge surrounded by a *pR* and a *pS* face), it forces the -S-S- linkage to adopt a *cis* conformation, which eventually explains the flattened shaped observed on heterochiral species **A4-b** and **A4-c**. While the virtual sub-library of cyclotetramers potentially encompasses 16 combinations of configurations and conformations, this stereochemical inter-dependence guides the real library toward the (x,n) = (a,1), (b,2), (c,3) diagonal on the gridded landscape in figure 2.

The interconnection between the two levels of stereochemical information borne by dyn[4]arenes was confirmed experimentally on the homochiral *pR*<sub>4</sub>/*pS*<sub>4</sub> species **A4-a**. As previously reported<sup>43</sup>, this cavitand can be quantitatively and stereoselectively self-assembled, at the expense of other ring sizes and stereochemistries, using biogenic polyamines such as spermine and spermidine as templates (figures S4-S5). Such an object combines the ideal cavity size and carboxylate group layout (at the vertices of its parallelepipedal structure) to optimize the binding of polyammoniums by salt bridges (figure 2i). Previously published optimized procedures<sup>43</sup> enable to obtain on demand either the pure polyamine **A4-a** pseudo-rotaxane complex or the free host upon guest departure. Both architectures were submitted to a crystallogenesis campaign, which eventually afforded samples eligible for structural elucidation by X-ray diffraction analyses. Guest-containing and free structures (figures S6-S7) support and validate *a posteriori* the modelling data (figure 2ii). As expected, diammonium guests transfer their stereochemical information (a cylindrical linear backbone terminated by two electron poor moieties) to the host building block, by templating (the algorithm) its assembly into a parallelepipedal structure which displays the optimal carboxylate groups layout (at the vertices of the parallelepiped, *i.e.* corresponding to the homochiral configuration).

From this intermolecular guest-to-host information transfer, which selects the homochiral **A4-a** species among all cyclotri- to penta-meric combinations, a second level of information is intramolecularly encrypted. In fact, the steric hindrance of the carboxylate groups propagates along the vertical edges of the parallelepipedic cavity, positioning each S-S linkage right in the axis formed by the two neighbouring carboxylates (figure 2ii). Eventually, only the *pR-P* and *pS-R* pairs of stereoisomers are observed, in full agreement with the computational model. Remarkably, not only a simple guest such as an  $\alpha,\omega$ -diammonium can transfer up to 8 bits of stereochemical information (coded as (*pR-P*)<sub>4</sub> or (*pS-M*)<sub>4</sub>) to the self-assembled host, but this information is recorded, *i.e.* remains implemented into the host upon guest departure, as confirmed by the X-ray structure of the free host (figures S6-S7). Fortuitously, the informational input initiated by the guest propagates beyond the individual host as conglomerate crystals are obtained (figure 2iii), *i.e.* crystals in which (*pR-P*)<sub>4</sub> and (*pS-M*)<sub>4</sub> segregate. This indeed means that a third stereoselective self-assembling process takes place consecutively to the templation step: the self-sorting of nearest neighbour cavitands of identical configuration. In other words,  $\alpha,\omega$ -polyamines encode the encryption of (*pR-P*)<sub>4</sub> and (*pS-M*)<sub>4</sub> 8-letter words. Upon, guest departure and supersaturation, these 8-letter words subsequently self-sort into 1-word sentences (the homochiral [(*pR-P*)<sub>4</sub>]<sub>n</sub> and [(*pS-M*)<sub>4</sub>]<sub>n</sub> aggregates).

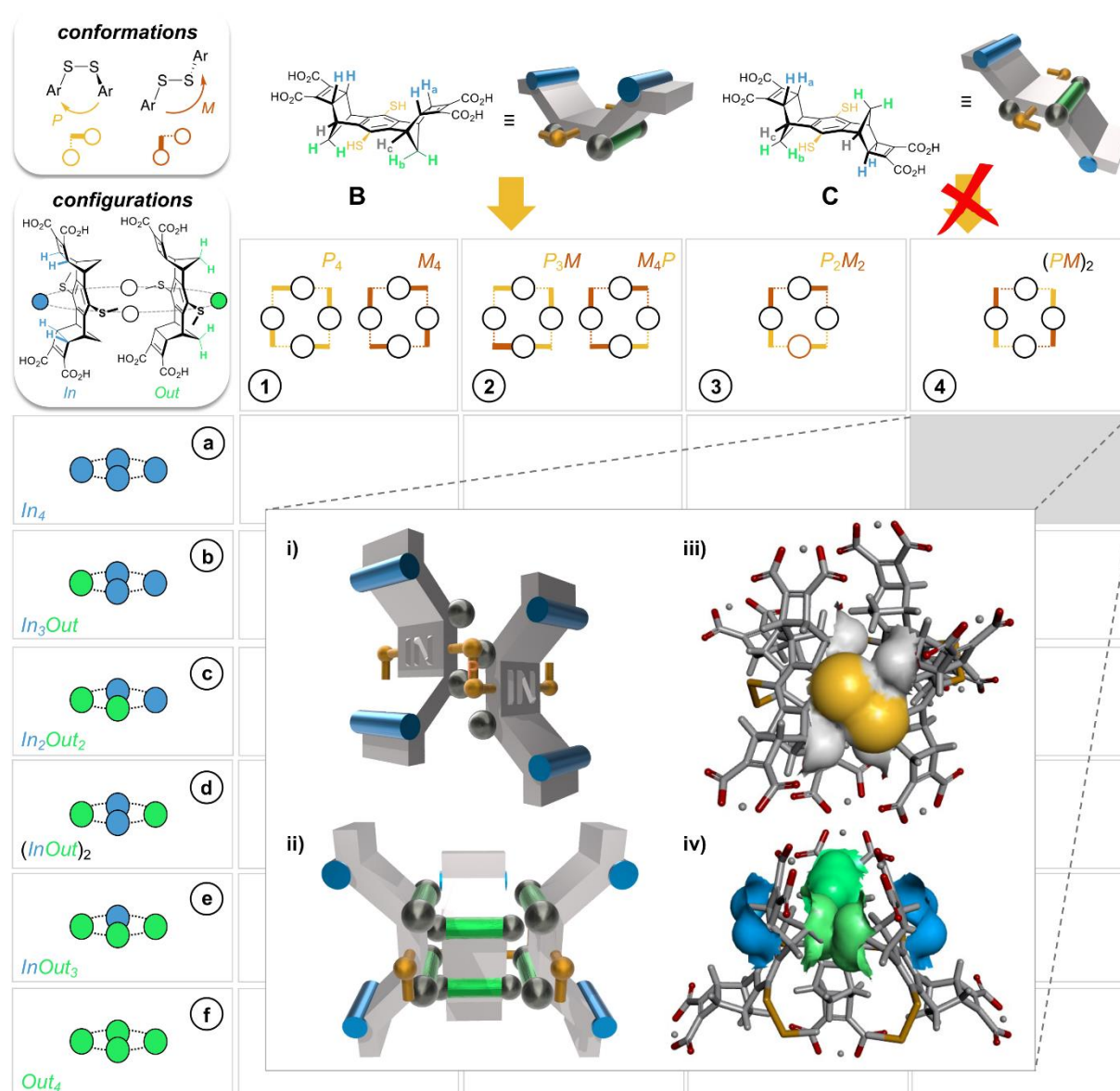
**Template-free stereochemical information encoding.** From this first example of guest-to-host transfer and storage of 8-bits of stereochemical information, we envisaged the possibility of extending the concept to guest-free systems. To do so, information should be entirely encoded into the building blocks themselves, through their stereochemical features, *i.e.* by synthetic design. In the first generation of dynarenes based on **A**, the stereochemical information was able to propagate along the edges of the monomeric faces through steric interactions (figure 1c, blue highlight).

In a second-generation system, the same class of challenging non-directional interaction was conserved but implemented into three moieties located at different positions of the building blocks (named **B** and **C**, figure 3). The profile of these building blocks was inspired from Schraders's molecular tweezers which are equipped with benzene walls and norbornadiene edges<sup>45</sup>. With this



design, three groups differing in bulkiness are borne by the cis and trans staircase-shaped monomers **B** and **C**: the pair of axial protons  $H_a$  found at the edge between the norbornene and cyclobutene rings is the bulkiest class of C-H protons (figure 3i-ii, cyan), the methylene  $CH_{2b}$  bridge of the norbornene is slightly less hindered and ranks second (figure 3i-ii, green) and finally, the lateral protons  $H_c$ , which are parallel to the C-S bond or the thiol group, should be sufficiently bulky to conformationally freeze the disulfide bridge on the NMR timescale. With this design, when primary and secondary bulky units are borne on different faces (such as in monomer **B**), thiol to disulfide cyclooligomerization should in principle stereoselectively deliver convex dyn[n]arenes (configuration noted  $(In)_n$ ). In this case,

the bulkiest  $C-H_a$  group prevent concave ring closure. In addition, disulfide bridges should adopt  $M/P$  alternate conformations (to avoid contact between protons  $H_b$  and  $H_c$  borne by nearest neighbours, see figure 3iii-iv). In contrast, cyclooligomerization, should be prohibited when both types of bulky units  $H_a$  and  $H_b$  are present on each face (such as in monomer **C**, figure 3ii). The synthesis of these new monomers **B** and **C** relied on a sequence of successive metal-free and metal-catalyzed cycloadditions, starting from simple hydroquinone. (See scheme S1 and supporting section S.II.2.a) Diastereoisomer separation was conducted by chiral chromatographic means at the middle of this sequence (see supporting section S.I.2)



**Figure 3. Principle of stereochemical information encoding by self-assembly of hindered dithiophenol monomers **B** and **C**** (i and ii: schematic representation). Axial protons (cyan: methine  $C-H_a$ ; green: methylene  $C-H_b$ ) hinder the concave face of **B** and both faces of **C**, preventing the cyclooligomerization of the latter. Axial  $C-H_b$  (green) and equatorial  $C-H_c$  (grey)

groups restrict disulfide dihedral rotation (ii and iv). Out of 24 possible tetrameric combinations of *In* and *Out* facial configurations (arising from the positioning of concave/convex monomer faces in- or outward with respect to the cavity, horizontal coordinate) and *P* or *M* disulfide bridges conformation (vertical coordinate), **B** is programmed to selectively deliver a *meso* alternate *MPMP* conformation and (*In*)<sub>4</sub> configuration. DFT modelling at the M06 level (v and vi: with Van der Waals surface) supports the design.

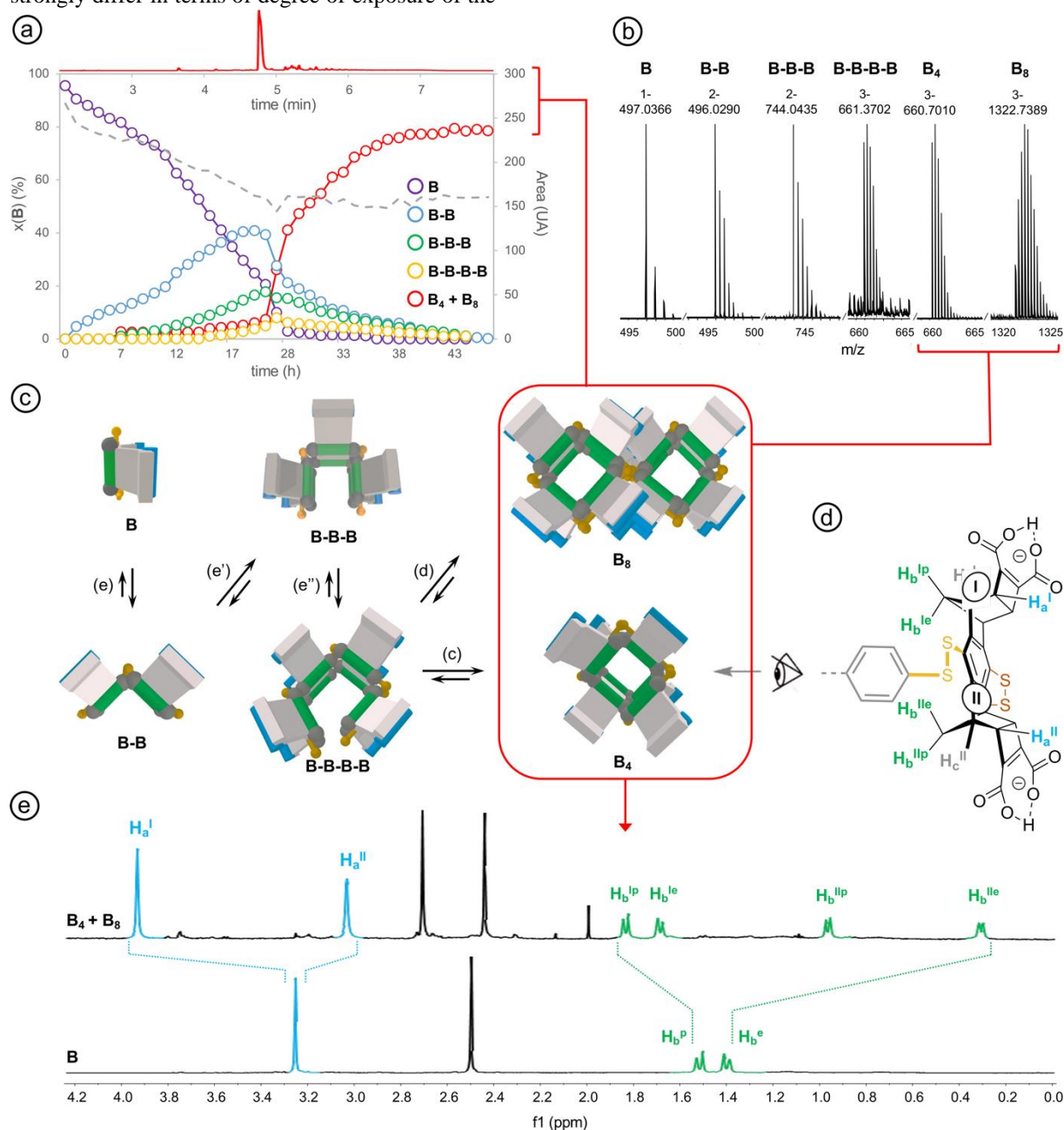
To enable the quantitative formation of discrete cyclic species rather than polymers<sup>30</sup>, libraries were set in the millimolar range from building block **B** or **C**. Slightly acidic pH (5.5) was chosen so that the monomeric units were mono(de)protonated at each dicarboxylic end (figure 4d), thereby yields soluble oligomers within which hydrogen bonds may potentially be shared between nearest monomeric neighbours. This intermediate protonation state was verified *a posteriori* by the potentiometric titration of the isolated cyclooligomers obtained (figure S120-121). Monitoring of the aerobic thiols-to-disulfide based cyclo-oligomerization by (-)-ESI-MS allowed us to qualitatively validate our design. In the case of building block **B**, disulfide-based linear species were detected during the first 4 hours, followed by the rise of their macrocyclic homologues and the concomitant decline of linear species (figure S97-100). A similar first regime of thiol-to-disulfide-based chain growth was observed on **C**, which then diverged during the second regime from the mild oxidation pathway. In fact, ring closure was replaced by overoxidation of the disulfides linkages (into thiolsulfonates) on the linear strands, which subsequently underwent cleavage into smaller linear fragments and further overoxidation (from sulfenic to sulfonic acids, figure S101-105). Remarkably, no trace of such overoxidized species could be detected by ESI-MS from libraries based on **B**. Beyond the qualitative validation of our design, this preliminary analysis seems to indicate that the disulfide oxidation state may be metastable in some circumstances and therefore require additional stabilization sources, which are herein provided by intramolecular non-covalent interactions within the macrocyclic species **B<sub>n</sub>**.

In-depth compositional and structural analyses of the architectures formed from monomer **B** were conducted by a combination of computational and experimental techniques. UPLC-UV/ESI-MS monitoring over time first provided a quantitative assessment of populations of linear and cyclic objects produced throughout the oxidation process (figure 4a, S106-109). It appeared that linear species of increasing chain length (from dimer **B-B** to tetramer **B-B-B-B**) are progressively generated from monomer **B** during the first 24 hours while the cyclic species, although detectable in ESI-MS by direct infusion (*vide infra*), remain limited to trace

amounts. In fact, the macrocyclization regime by ring closure or dimerization (eq. (c) and (d), figure 4a) effectively takes over the elongation phase (eq. (e)-(e''), figure 4a) after 24h, yielding quantitatively a cyclotetrameric product noted **B<sub>4</sub>**. Intriguingly, the corresponding mass spectra, indicated the presence of **B<sub>8</sub>-3H<sup>3-</sup>** and **B<sub>8</sub>-5H<sup>5-</sup>** ions (figure S116). Careful examination of the isotopic patterns corresponding to **B<sub>4</sub>-2H<sup>2-</sup>** and **B<sub>4</sub>-3H<sup>3-</sup>** also reveal that their isotopic pattern was respectively interfered with those of **B<sub>8</sub>-4H<sup>4-</sup>** and **B<sub>8</sub>-6H<sup>6-</sup>** (figure S115). Given the octaanionic nature of **B<sub>4</sub>** at pH 5.5, formation of a non-covalent dimeric **B<sub>4</sub>-B<sub>4</sub>** complex which may survive to the ionization process seemed highly unlikely<sup>1</sup>. Although a single and relatively sharp peak is obtained in UPLC after equilibration (and remains stable over more than two weeks), it was ascribed to the co-elution of the cyclic tetramer **B<sub>4</sub>** and octamer **B<sub>8</sub>** (figure 4a-c and S108). MS/MS fragmentation of the **B<sub>8</sub>-3H<sup>3-</sup>** ion systematically delivered the **B-H<sup>-</sup>** monomeric species, which is consistent with a covalent octamer (figure S117). Despite a thorough screening of all currently available chromatographic phases, any attempt to separate these two species, which are presumably obtained from the same tetrameric linear parent **B-B-B-B** by ring closure (figure 4a, eq. (c)) or cyclodimerization (figure 4a, eq. (d')), remained fruitless. Similarly, our generic procedure of dynarene isolation by controlled acidic precipitation proved to be ineffective to separate both species. The mixture obtained upon equilibration was thus analyzed by multi-nuclear and multi-dimensional NMR (supporting section S.II.2.f). Quantitative <sup>1</sup>H and <sup>13</sup>C spectra in the presence of an internal standard indicated the presence of a single well-resolved major species (approx. 80 % molar, figure S132), involving two sets of H<sub>a-c</sub> magnetically different protons (figure 4d and 4e, noted I and II). COSY and NOESY experiments confirmed that these two sets are geometrically isolated, while heteronuclear 2D experiments (HSQC, HMBC) enabled to assign the signals within each set to one of the H<sub>a-c</sub> class (figures S122-129). Finally, DOSY confirmed that these two sets belong to the same molecular entity, which displays a hydrodynamic radius of 9.08 +/- 0.75 Å (as a comparison, **A<sub>4</sub>-a**'s radius is 8.30 +/- 0.80 Å and the radius of monomer **B** is measured by the same technique around 5.71 +/- 0.07 Å in the same conditions, see figures S87 and S130 and table S10-11). These analyses enabled to

narrow the scope of **B**<sub>4</sub> cyclotetramer candidates (those displaying two geometrically distinct H<sub>a-c</sub> units) to the *meso*-(*In-M-In-P*)<sub>2</sub>, the (*Out-M-Out-P*)<sub>2</sub> species and the homochiral racemic (*In-P-Out-P*)<sub>2</sub> (*In-M-Out-M*)<sub>2</sub> mixture. These three stereochemical combinations are located on the gridded landscape of all possible configurations and conformations summarized in figure 3 at the (a,4), (f,4) and (d,1) coordinates respectively. Structurally speaking, they strongly differ in terms of degree of exposure of the

H<sub>a</sub> and H<sub>b</sub> protons to the outer solution. Experimentally, it appears that macrocyclization of **B** induces the shielding and deshielding of H<sub>a</sub> and H<sub>b</sub> protons, with a splitting within each class (figure 4e). This supports the hypothesis of a **B**<sub>4</sub> stereoisomer bearing H<sub>a</sub> (and H<sub>b</sub> protons) which all point outward (respectively inward) with respect to the macrocyclic cavity, wherein two distinct local environments can be found.



**Figure 4.** UPLC-MS analysis of the dynamic combinatorial library generated from monomer **B** (4 mM) in 50 mM AcONH<sub>4</sub> at pH 5.5. a) Schematic representation of the linear and cyclic species quantified and characterized by UPLC-UV/ESI-MS (b and c), d) monomeric unit displaying two sets of H<sub>a-c</sub> protons (I and II) and e) snapshot of the <sup>1</sup>H NMR spectra of starting material **B** and final product **B**<sub>4</sub>+**B**<sub>8</sub>.

Assessment of the stability of each member of the virtual combinatorial sub-library of cyclotetramer

stereoisomers gathered on the map displayed in figure 3 was performed by DFT calculations

(structure optimization with three alternative functionals, followed by single point (SP) energy calculation, see section S.I.8 and table S12). Comparison of the energy levels, which are displayed in figure S139 unambiguously reveal that, the *(In-M-In-P)*<sub>2</sub> stereoisomer is the most stable species and should experimentally dominate the sub-library of **B**<sub>4</sub> stereoisomers obtained from **B**. In addition to complementing and refining the NMR analyses, this modelling campaign also provided some insight on the main contributions to the stability of the species. In agreement with our design, it appears that orienting the H<sub>a</sub>-bearing concave faces toward the cavity (i.e. the *Out* configuration) destabilizes the corresponding architectures for steric reason. Stabilizing London dispersion forces follow the same evolution (increasing in strength with the number of *Out* facial configurations) but their amplitude is insufficient and fail to compensate the repulsion. To minimize steric clash between H<sub>b</sub> protons of facing monomeric units (which are in Van der Waals contact, see figure 3iii and iv), the structure adopts distorted octahedral shape. Finally, H<sub>b</sub> and H<sub>c</sub> protons play their ascribed role of disulfide bridge instructors. Energy calculation validates *a posteriori* the encoding process i.e. the effective translation of the stereochemical information borne by **B** by the assembling algorithm (the concomitant disulfide bond formation/exchange operating under the guidance of steric interactions). DFT modelling indeed confirm that the *MPMP* conformation is systematically favoured within each configurational series. On the most stable and selectively obtained **B**<sub>4</sub>, this conformation minimizes the hindrance between H<sub>b</sub> protons from neighbouring monomeric units. Finally, molecular dynamics confirm that H<sub>c</sub> provide sufficient hindrance in the vicinity of the S-S bridges to freeze and lock their conformation (see figure 5i and ii), at least on the millisecond NMR timescale.

At this stage, it should be mentioned that NMR spectra of equilibrated libraries either analysed directly (*in situ*) or after acid-induced precipitation and redissolution (*ex situ*) both display sharp signals corresponding to *meso*-(*In-M-In-P*)<sub>2</sub>-**B**<sub>4</sub> which account for 80% of the total material, while the remaining 20% imputed to a **B**<sub>8</sub> adduct, are dispatched into broad signals buried in the baseline (figure S). We first investigated the impact of the amphoteric salt used for pH regulation, by changing either the anion or cation while maintaining the pH of the library at 5.5). UPLC-analysis did not reveal any deviation from the scenario of **B**<sub>4</sub>/**B**<sub>8</sub> selective co-production (figure S106). To rule out the possibility of precipitation-induced interconversion

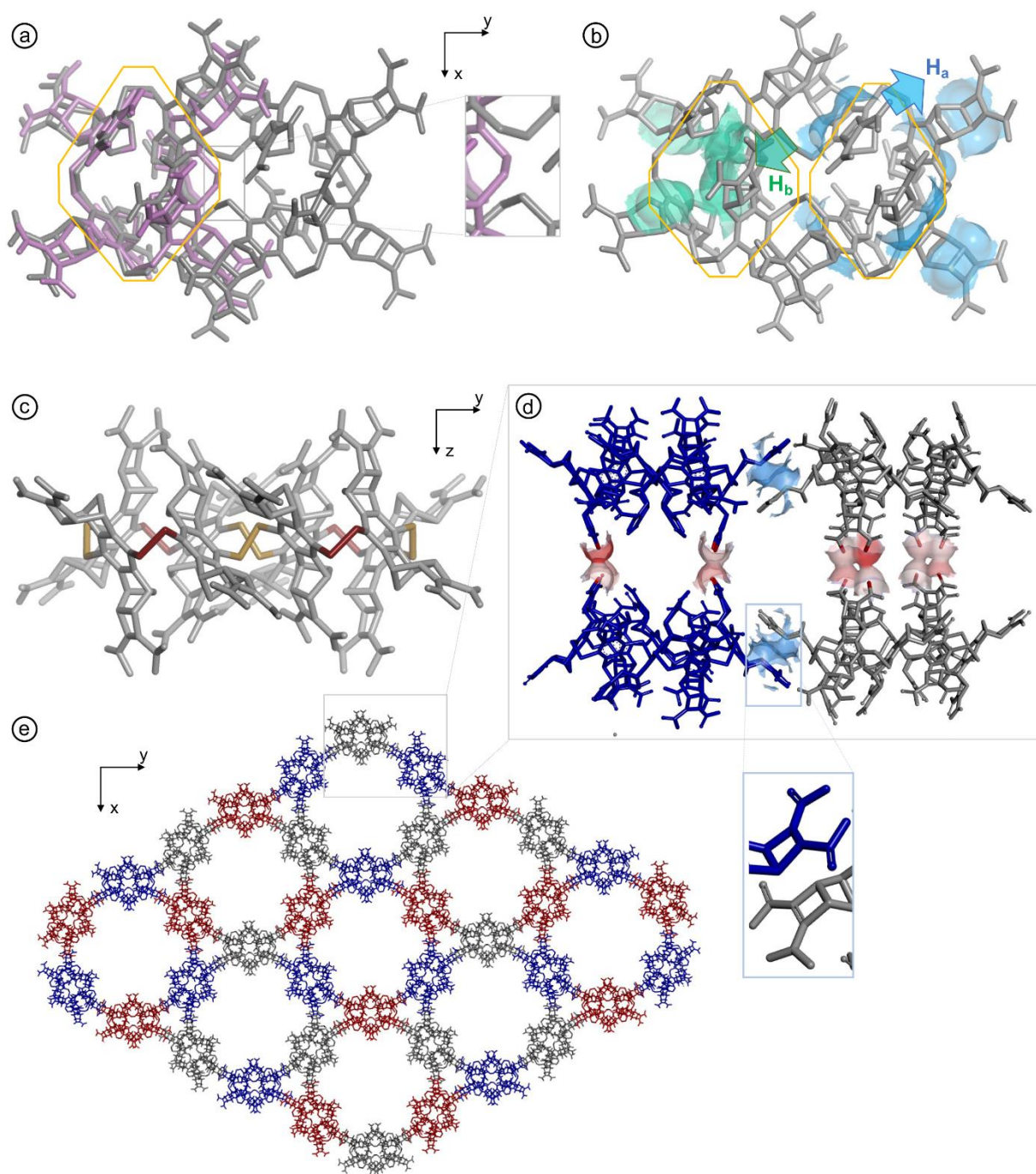
between **B**<sub>4</sub> and **B**<sub>8</sub>, which would remain undetectable upon redissolution, we then analysed the material collected by precipitation by NMR and MALDI-MS spectroscopy in the solid state (figures S118-119). Apart from the line broadening, solid and liquid state spectra can be superimposed, thereby ruling out this hypothesis. As **B**<sub>4</sub> and **B**<sub>8</sub> should result from the same **B-B-B-B** precursors through a monomolecular (ring closure) or bimolecular (dimerization) pathway, we reasoned that experimental parameters affecting the assembling process such as temperature, stirring rate or concentration may affect the final proportions of **B**<sub>4</sub> and **B**<sub>8</sub> within the libraries.

Increasing incubation temperature (from 25° to 75 °C) or stirring rate (from 0 to 1000 rpm) accelerated the elongation and cyclization regime, resulting in faster equilibration but final mixtures of identical compositions (see figure S109 for UPLC-MS analyses). Increasing the concentration of the starting material from 4 to 32 mM induced a similar accelerating effect but also resulted into the formation of solid material after several hours (above 15 mM, see figure S96). This solid, which was filtered and dried, accounted at best for 10 % of the total mass. It proved to be insoluble in all conventional solvents and no data could be extracted from ESI-MS or liquid-state NMR analyses. Solid state NMR delivered a <sup>13</sup>C CP-MAS spectra displaying broad peaks (figure S138), indicating that this sample differs from the **B**<sub>4</sub>+**B**<sub>8</sub> fraction collected upon acidification at least in terms of tridimensional organization.

As the analysis of the supernatant by ESI-MS and multinuclear NMR were similarly inconclusive (detection of the **B**<sub>8</sub>-3H<sup>3+</sup>, **B**<sub>8</sub>-5H<sup>5+</sup> and **B**<sub>4</sub>-3H<sup>3+</sup> ions for the latter and faint broad signals merged with the baseline accompanying the sharp peaks corresponding to **B**<sub>4</sub>, see figure S123 and S131), we performed some semi-quantitative ESI-MS sampling during the assembling process. Quantification of the intensities and areas of the isotopic patterns of identical m/z corresponding to **B**<sub>4</sub>-nH<sup>n+</sup> and **B**<sub>8</sub>-2nH<sup>2n+</sup> with time was performed after a preliminary validation by calibration (figure S133-136), i.e. injection of precise and increasing amounts of **B**<sub>4</sub>/**B**<sub>8</sub> mixtures whose proportions had been determined first by NMR analyses (figure S132). This kinetic monitoring revealed that the **B**<sub>8</sub>/**B**<sub>4</sub> proportion is increasing with time during the cyclization regime, reaching the 0.25 asymptotic value measured by NMR, for libraries whose concentration was below the 15 mM threshold value. In contrast, more concentrated sample lead to higher proportions of **B**<sub>8</sub> (up to 0.45), which abruptly



dropped to the same asymptote amount (0.25) upon precipitation (figure S137).



**Figure 5. Structural elucidation of B<sub>8</sub> octamer by X-ray diffraction crystallography.** a): comparison of B<sub>4</sub> and B<sub>8</sub> structures obtained *in silico* and experimentally by X-ray diffraction respectively (pink: B<sub>4</sub>, grey: B<sub>8</sub>, framed: zoom on the junction between tetrameric subunits), c): analysis of the facial configuration (green: H<sub>a</sub>; cyan H<sub>b</sub>) and d): disulphide conformation (yellow: P-SS, red: M-SS) and d) intermolecular packing between four B<sub>8</sub> molecules (blue and red: Van der Waals contact between cyclobutene and carboxylic moieties, with a framed zoom of the former) from two xy layers displayed in e)

These observations provided the framework for crystallogenes campaign based on supersaturated solution prepared by dissolution of the solid B<sub>4</sub>+B<sub>8</sub> material collected after thiol to disulfide assembling in dilute conditions and acid precipitation. To our delight, single crystals eligible for structural

elucidation by X-ray diffraction analysis could be collected after screening (figure S140). It should be noted that attempt to isolate them from the mother liquor lead to their collapse and prevented us from recording MS or NMR spectra in the liquid state. The tridimensional structure of the solid material

obtained from supersaturated mixtures of **B**<sub>4</sub> and **B**<sub>8</sub> is shown in figure 5. In agreement with ESI-MS monitoring or the assembling and CP-MAS NMR, this analysis confirms that the insoluble material corresponds to a covalent octamer **B**<sub>8</sub> (and rules out the unlikely hypothesis of a complex between two **B**<sub>4</sub> species). The shared filiation of **B**<sub>4</sub> and **B**<sub>8</sub> with **B-B-B-B** is confirmed by the eight-shaped profile of **B**<sub>8</sub> whose octahedral subunits (figure 5a, grey) almost perfectly superimpose with **B**<sub>4</sub> (figure 5a, pink). This solid-state structure also reveals how the molecular information we encoded into the monomer is handled during the dimerization of **B-B-B-B** into **B**<sub>8</sub>, which competes with the ring closure into **B**<sub>4</sub>, i.e. when two tetramers are to be fused. As for **B**<sub>4</sub>, the stereochemical information encoded into the monomeric unit is stereoselectively translated into the cyclooctamer both at the configurational and conformational levels. An all-*In* facial configuration, wherein the bulkiest H<sub>a</sub> protons point toward the outer medium (figure 5b, cyan) and the slightly less bulky H<sub>b</sub> (figure 5b, green) are oriented inward is indeed observed. Regarding disulfide bridge conformation, it is confirmed that the *M/P* alternance is the only combination which accommodates the hindrance borne by the building blocks, even on a larger ring size. (figure 5c and S141). As in **B**<sub>4</sub>, the unusual eight-shaped folding finds its origin in the interplay between intramolecular attractive and repulsive interactions. Yet, careful examination of the packing between (*M-In-P-In*)<sub>4</sub>-**B**<sub>8</sub> homologues reveal that intermolecular interactions in the solid lattice should act here as an additional folding driving force. In the equatorial plane of the macrocyclic ring, contacts between **B**<sub>8</sub> molecules occur at the tip of their long sides, i.e. between the monomeric units which are not engaged into the junctions between **B**<sub>4</sub> subunits, though unexpected stacking of their H<sub>a</sub>-bearing cyclobutene moieties (figure 5d). At each pole of the **B**<sub>8</sub> structure, 2 monomers are paired through this unusual interaction with a different **B**<sub>8</sub> neighbour (figure S143). Repetition of this pattern along the xy plane yields cyclotrimeric and hexameric superstructures displaying respectively triangular and circular cavities (figures 5e and S142). In the z-direction, stacking of the **B**<sub>8</sub>-layers occurs through H-bond pairing between the carboxylic moieties of homologous monomers (figure 5d, red). The carboxylic pairing concerns the four central building blocks of the 8-shaped **B**<sub>8</sub> ring, i.e. the zone resulting from the cyclodimerization of two linear **BBBB** tetramers. This analysis reveals that stacking in the solid phase, especially along the z axis, may be the driving force underpinning the unexpected formation and selection of a large ring size, which in

the absence of enthalpic driving force, is entropically unfavoured.<sup>47</sup> Despite this unexpected powerful driving force, 8 and 16-bits of (*In-M-In-P*)<sub>n</sub> stereochemical sequences could be successfully and simultaneously encoded into dyn[4]- and dyn[8]-arene, respectively in solution and in the solid state.

## Conclusion

We have brought here the proof of feasibility of multi-bit stereochemical information encoding into disulfide-based cyclophanes. More specifically we could show that the steric hindrance borne by simple methylene and methine moieties on achiral dithiophenol building blocks can serve as informational input. This input can be translated, by a non-directional assembling algorithm (disulfide bridge formation and steric hindrance minimization) with or without the assistance of a template, into a specific stereochemical output (the configurational and conformational sequence of the assembled product). Monomers and connecting linkages can be viewed as two channels wherein two alternative letters/values may be encoded, their inter-relation depending on the input i.e. on the nature of the building block. Encryption is therefore underpinned by design and of the starting material (input) and the choice of the assembling chemistry and interactions at work (the algorithm). The readout requires structural analysis (herein provided by NMR and X-ray crystallography), which currently limits the potential applicability of this innovative concept in molecular informatics. The scope of synthetic application of templating and folding, well known driving forces for compositional selection and amplification in DCC was herein extended to stereochemistry. Unexpectedly, the third driving force, namely aggregation and crystal packing can also influence the stereochemical selection by reinforcing (in the case of **A4-a**) or perturbing (in the case of **B**<sub>8</sub>) the encoding process. This conceptual toolbox should therefore be of great value, enabling the community to program on demand the stereoselective assembling of sophisticated architectures from very simple components.

**Acknowledgements** We thank the CCSM, CRMN and CCRMN for assistance on Mass Spectrometry, DOSY, High field liquid and solid-state NMR analyses respectively. We acknowledge the contribution of SFR Biosciences (UMS3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) facilities: Protein Science Facility - Frédéric DELOLME for performing MALDI-TOF-MS analysis. We are grateful to S. Vidal for providing access to his UPLC facility (ANR-17-CE07-0054). Financial support from the Chinese Scholarship council (Y. Z

scholarship) and UCB Biopharma (P.T.S and B.O scholarships) and the IR-RMN-THC Fr3050 CNRS is gratefully acknowledged. This work was supported by the LABEX iMUST (ANR-10-LABX-0064) of Université de Lyon, within the program "Investissements d'Avenir" operated by the French National Research Agency (ANR).

**Author contributions** J.L. designed the system, J.L., F. P., L. V. and I. H designed the experiments. Y. Z., P-T. S. and B. O. carried out the experimental work. A. B. designed the chromatographic methods and conducted the analyses. O. C. designed the NMR methods and conducted the analyses. N. V. designed the chiral chromatographic methods and conducted the analyses. P. M. conducted the X-ray diffraction analyses. C. D. designed and performed the MS experiments. E.D. designed the simulations. E.J. and B.O. performed the simulation. J.L. wrote the paper. All authors contributed to revising the paper.

**Competing interests.** The authors declare no competing interest

**Data availability statement** "Crystallographic data for the structures reported in this Article have been deposited at the Cambridge Crystallographic Data Centre, under deposition numbers CCDC 1554744, 2003290, 2001333, 2003509 for **A4-a**, **A4-a-Cadav**, **A4-a-Put** and **B8** respectively. Copies of the data can be obtained free of charge via <https://www.ccdc.cam.ac.uk/structures/>. All other data supporting the findings of this study, including synthetic and analytical procedures are available within the Article and its Supplementary Information, or from the corresponding author upon reasonable request."

## References

1. Lehn, J.-M. Supramolecular Chemistry—Scope and Perspectives Molecules, Supramolecules, and Molecular Devices (Nobel Lecture). *Angewandte Chemie International Edition in English* **27**, 89–112 (1988).
2. Lehn, J.-M. Perspectives in Supramolecular Chemistry—From Molecular Recognition towards Molecular Information Processing and Self-Organization. *Angewandte Chemie International Edition in English* **29**, 1304–1319 (1990).
3. Lehn, J.-M. Programmed Chemical Systems: Multiple Subprograms and Multiple Processing/Expression of Molecular Information. *Chemistry – A European Journal* **6**, 2097–2102 (2000).
4. Funeriu, D. P., Lehn, J.-M., Fromm, K. M. & Fenske, D. Multiple Expression of Molecular Information: Enforced Generation of Different Supramolecular Inorganic Architectures by Processing of the Same Ligand Information through Specific Coordination Algorithms. *Chemistry – A European Journal* **6**, 2103–2111 (2000).
5. Ayme, J.-F., Lehn, J.-M., Bailly, C. & Karmazin, L. Simultaneous Generation of a [2 × 2] Grid-Like Complex and a Linear Double Helicate: a Three-Level Self-Sorting Process. *J. Am. Chem. Soc.* **142**, 5819–5824 (2020).
6. Wang, J. & Ferguson, A. L. A Study of the Morphology, Dynamics, and Folding Pathways of Ring Polymers with Supramolecular Topological Constraints Using Molecular Simulation and Nonlinear Manifold Learning. *Macromolecules* **51**, 598–616 (2018).
7. Pease, A. R. *et al.* Switching Devices Based on Interlocked Molecules. *Acc. Chem. Res.* **34**, 433–444 (2001).
8. Dong, J., Liu, Y. & Cui, Y. Supramolecular Chirality in Metal–Organic Complexes. *Acc. Chem. Res.* **54**, 194–206 (2021).
9. Jędrzejewska, H. & Szumna, A. Making a Right or Left Choice: Chiral Self-Sorting as a Tool for the Formation of Discrete Complex Structures. *Chem. Rev.* **117**, 4863–4899 (2017).
10. Stewart, D. *et al.* Stable and ordered amide frameworks synthesised under reversible conditions which facilitate error checking. *Nat Commun* **8**, 1102 (2017).
11. Segura, J. L., Royuela, S. & Ramos, M. M. Post-synthetic modification of covalent organic frameworks. *Chem. Soc. Rev.* **48**, 3903–3945 (2019).
12. Hasenknopf, B., Lehn, J.-M., Boumediene, N., Leize, E. & Dorsselaer, A. V. Kinetic and Thermodynamic Control in Self-Assembly: Sequential Formation of Linear and Circular Helicates. *Angewandte Chemie International Edition* **37**, 3265–3268 (1998).
13. Hasenknopf, B. *et al.* Self-Assembly of Tetra- and Hexanuclear Circular Helicates. *J. Am. Chem. Soc.* **119**, 10956–10962 (1997).
14. Hasenknopf, B., Lehn, J.-M., Kneisel, B. O., Baum, G. & Fenske, D. Self-Assembly of a Circular Double Helicate. *Angewandte Chemie International Edition in English* **35**, 1838–1840 (1996).
15. Ayme, J.-F. *et al.* A synthetic molecular pentafoil knot. *Nature Chem* **4**, 15–20 (2012).
16. Suárez, M., Branda, N., Lehn, J.-M., Decian, A. & Fischer, J. Supramolecular Chirality: Chiral hydrogen-bonded supermolecules from achiral molecular components. *Helvetica Chimica Acta* **81**, 1–13 (1998).



17. Nitschke, J. R. Construction, Substitution, and Sorting of Metallo-organic Structures via Subcomponent Self-Assembly. *Acc. Chem. Res.* **40**, 103–112 (2007).
18. Langner, A. *et al.* Self-recognition and self-selection in multicomponent supramolecular coordination networks on surfaces. *PNAS* **104**, 17927–17930 (2007).
19. Rota Martir, D. *et al.* Homochiral Emissive  $\Delta$ 8- and  $\Delta$ 8-[Ir8Pd4]16+ Supramolecular Cages. *Chemistry – A European Journal* **23**, 14358–14366 (2017).
20. Nierengarten, I., Buffet, K., Holler, M., Vincent, S. P. & Nierengarten, J.-F. A mannosylated pillar[5]arene derivative: chiral information transfer and antiadhesive properties against uropathogenic bacteria. *Tetrahedron Letters* **54**, 2398–2402 (2013).
21. Pavlović, R. Z. *et al.* From Selection to Instruction and Back: Competing Conformational Selection and Induced Fit Pathways in Abiotic Hosts. *Angewandte Chemie* **133**, 20095–20101 (2021).
22. Han, C., Zhao, D., Li, T. & Sun, D. Stereoselective Synthesis of Pillar[4]arene[1]cis-diepoxy-p-dione and X-Ray Crystal Structure of Host–Guest System. *Chemistry – An Asian Journal* **12**, 2354–2358 (2017).
23. Ogoshi, T., Yamagishi, T. & Nakamoto, Y. Pillar-Shaped Macrocyclic Hosts Pillar[n]arenes: New Key Players for Supramolecular Chemistry. *Chem. Rev.* **116**, 7937–8002 (2016).
24. Xue, M., Yang, Y., Chi, X., Zhang, Z. & Huang, F. Pillararenes, A New Class of Macrocycles for Supramolecular Chemistry. *Acc. Chem. Res.* **45**, 1294–1308 (2012).
25. Vinciguerra, B. *et al.* Synthesis and Self-Assembly Processes of Monofunctionalized Cucurbit[7]uril. *J. Am. Chem. Soc.* **134**, 13133–13140 (2012).
26. Zhang, M., Sigwalt, D. & Isaacs, L. Differentially functionalized acyclic cucurbiturils: synthesis, self-assembly and CB[6]-induced allosteric guest binding. *Chem. Commun.* **51**, 14620–14623 (2015).
27. Day, A., Arnold, A. P., Blanch, R. J. & Snushall, B. Controlling Factors in the Synthesis of Cucurbituril and Its Homologues. *J. Org. Chem.* **66**, 8094–8100 (2001).
28. Wang, X., Jia, F., Yang, L.-P., Zhou, H. & Jiang, W. Conformationally adaptive macrocycles with flipping aromatic sidewalls. *Chem. Soc. Rev.* **49**, 4176–4188 (2020).
29. Fa, S., Kakuta, T., Yamagishi, T. & Ogoshi, T. Conformation and Planar Chirality of Pillar[n]arenes. *Chem. Lett.* **48**, 1278–1287 (2019).
30. Holler, M., Allenbach, N., Sonet, J. & Nierengarten, J.-F. The high yielding synthesis of pillar[5]arenes under Friedel–Crafts conditions explained by dynamic covalent bond formation. *Chem. Commun.* **48**, 2576–2578 (2012).
31. Li, J., Zhou, H.-Y., Han, Y. & Chen, C.-F. Saucer[n]arenes: Synthesis, Structure, Complexation, and Guest-Induced Circularly Polarized Luminescence Property. *Angewandte Chemie* **n/a**.
32. Sala, P. D. *et al.* An intramolecularly self-templated synthesis of macrocycles: self-filling effects on the formation of prismarenes. *Chem. Sci.* **12**, 9952–9961 (2021).
33. Isaacs, L. *et al.* The Inverted Cucurbit[n]uril Family. *J. Am. Chem. Soc.* **127**, 18000–18001 (2005).
34. Stancl, M., Gargulakova, Z. & Sindelar, V. Glycoluril Dimer Isomerization under Aqueous Acidic Conditions Related to Cucurbituril Formation. *J. Org. Chem.* **77**, 10945–10948 (2012).
35. Huang, W.-H., Zavalij, P. Y. & Isaacs, L. Chiral Recognition inside a Chiral Cucurbituril. *Angewandte Chemie International Edition* **46**, 7425–7427 (2007).
36. Pian, M. D., Schalley, C. A., Fabris, F. & Scarso, A. Insights into the synthesis of pillar[5]arene and its conversion into pillar[6]arene. *Org. Chem. Front.* **6**, 1044–1051 (2019).
37. Gianga, T.-M. & Pantoş, G. D. Structurally divergent dynamic combinatorial chemistry on racemic mixtures. *Nat Commun* **11**, 3528 (2020).
38. Corbett, P. T., Tong, L. H., Sanders, J. K. M. & Otto, S. Diastereoselective Amplification of an Induced-Fit Receptor from a Dynamic Combinatorial Library. *J. Am. Chem. Soc.* **127**, 8902–8903 (2005).
39. ten Cate, A. T., Dankers, P. Y. W., Sijbesma, R. P. & Meijer, E. W. Disulfide Exchange in Hydrogen-Bonded Cyclic Assemblies: Stereochemical Self-Selection by Double Dynamic Chemistry. *J. Org. Chem.* **70**, 5799–5803 (2005).
40. Lam, R. T. S. *et al.* Amplification of Acetylcholine-Binding Catenanes from Dynamic Combinatorial Libraries. *Science* **308**, 667–669 (2005).
41. Li, J., Nowak, P. & Otto, S. An Allosteric Receptor by Simultaneous “Casting” and “Molding” in a Dynamic Combinatorial Library. *Angewandte Chemie International Edition* **54**, 833–837 (2015).
42. Panijpan, B. Chirality of the disulfide bond in biomolecules. *J. Chem. Educ.* **54**, 670–672 (1977).
43. Carmack, M. Chirality of the Disulfide in the Prion Proteins. *J. Chem. Inf. Comput. Sci.* **44**, 286–288 (2004).
44. Skowron, P.-T. *et al.* On-Demand Cyclophanes: Substituent-Directed Self-Assembling, Folding, and Binding. *J. Org. Chem.* **81**, 654–661 (2016).
45. Dumartin, M. *et al.* The dark side of disulfide-



based dynamic combinatorial chemistry. *Chem. Sci.* **11**, 8151–8156 (2020).

46. Fokkens, M., Schrader, T. & Klärner, F.-G. A Molecular Tweezer for Lysine and Arginine. *J. Am. Chem. Soc.* **127**, 14415–14421 (2005).

47. Corbett, P. T., Sanders, J. K. M. & Otto, S. Exploring the Relation between Amplification and Binding in Dynamic Combinatorial Libraries of Macrocyclic Synthetic Receptors in Water. *Chem. Eur. J.* **14** (2008).

---