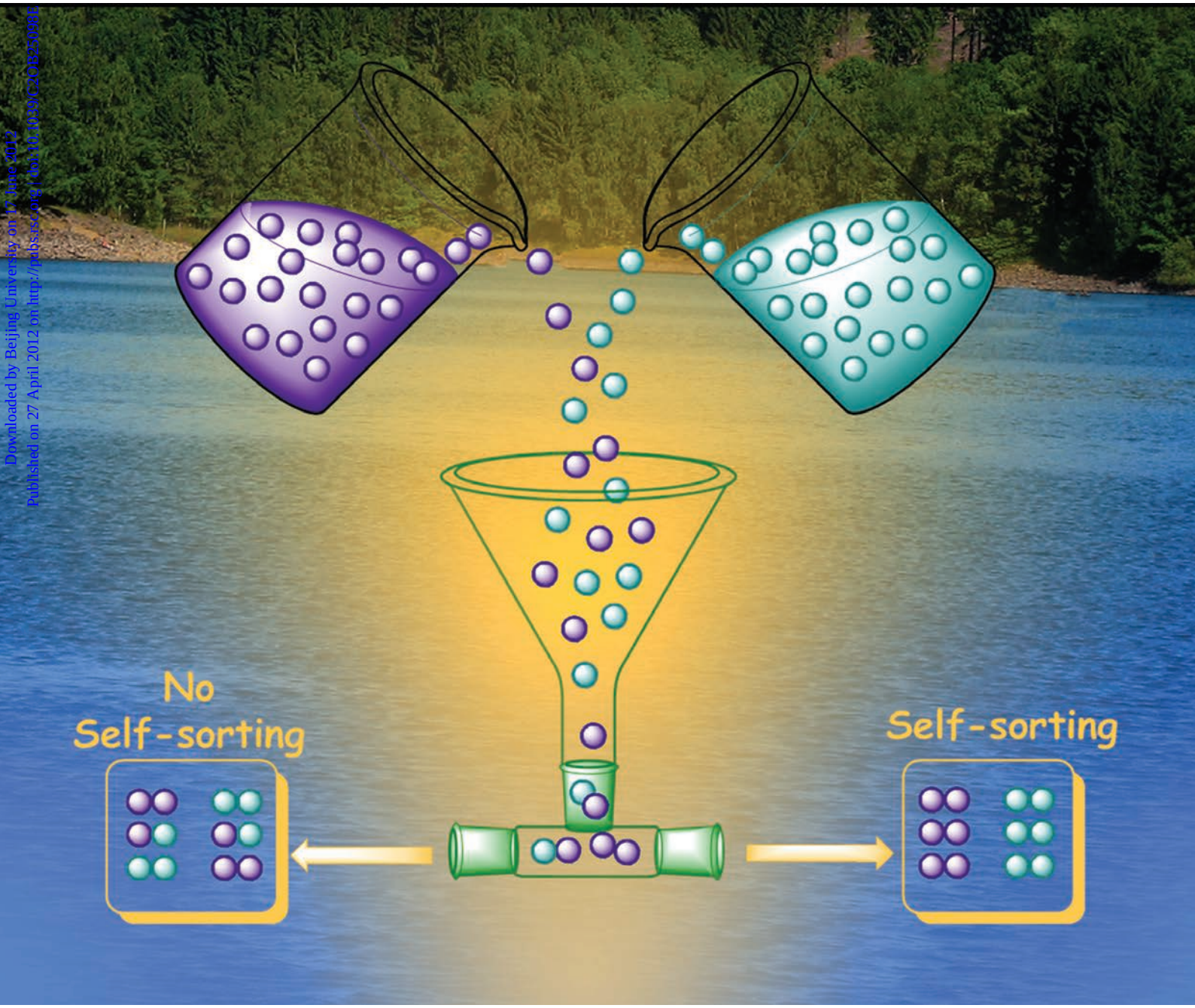
# Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 10 | Number 24 | 28 June 2012 | Pages 4629–4808






ISSN 1477-0520

RSC Publishing

**PERSPECTIVE**

M. Lal Saha and M. Schmittel

Degree of molecular self-sorting in multicomponent systems

Organic & Biomolecular Chemistry

Dynamic Article Links
View Online

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 4651

www.rsc.org/obc

PERSPECTIVE

# Degree of molecular self-sorting in multicomponent systems

**Manik Lal Saha and Michael Schmittel***

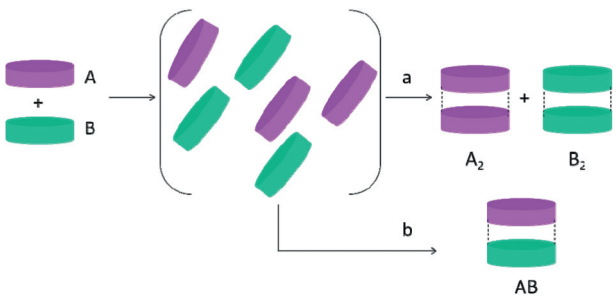
***Received 13th January 2012, Accepted 13th March 2012***
**DOI: 10.1039/c2ob25098e**



Self-sorting represents the spontaneous and high fidelity self and/or non-self-recognition of two or more related components within a complex mixture. While the effective management of self-sorting principles perceptibly requires some key expertise in *molecular programming*, at a higher stage of operation it is of supreme interest to guide the process to increasingly higher *degrees of self-sorting*. In this article, we present the emerging principles of how to guide several components toward formation of self-sorted multicomponent architectures. To provide further guidance in denominating such systems, we suggest to utilise a systematic classification as well as a formula to evaluate their *degree of self-sorting* (*M*).

## 1. Introduction

To guide an ensemble of species toward formation of a single well-defined, but fully dynamic aggregate requires independent (orthogonal) and/or competitive self-assembly protocols with a substantial degree of self-sorting. Self-sorting, in its original meaning, described the ability to distinguish "self" from "non-self" in a mixture of related components, *i.e.* the formation of well-defined homomeric assemblies at the expense of a random mixture of products (Scheme 1, path a).[1–3] However, later, this term has also been used for the mutual recognition of complementary components in artificial self-assembly (Scheme 1, path b).[4] In both scenarios, the sorting is guided by the correct read-out of specific information encoded in the




**Scheme 1** Self-sorting in a 2-component system depicting arbitrarily chosen 1 : 1 complexes.

*Center of Micro and Nanochemistry and Engineering, Department of Chemistry and Biology, Organische Chemie I, Universität Siegen, Adolf-Reichwein-Str. 2, D-57068 Siegen, Germany.*
*E-mail: schmittel@chemie.uni-siegen.de*

**Manik Lal Saha**

***Manik Lal Saha*** *was born in 1986 in West Bengal, India. He received his B.Sc. (Honours) in Chemistry from University of Calcutta (India) in 2006. After finishing his M.Sc. from Indian Institute of Technology Kanpur in 2008, he joined Prof. Schmittel's group at the University of Siegen (Germany). His current research is focused on multicomponent nano-assemblies and dynamers.*

**Michael Schmittel**

***Michael Schmittel*** *studied chemistry and French in Freiburg and Paris. After habilitation at the University of Freiburg he joined the University of Würzburg as an associate professor in 1993 and since 1999 holds a Chair in Organic Chemistry at the University of Siegen. His broad interests, published in* ca. *200 papers and one textbook, cover research areas as diverse as non-statistical dynamics of thermal diradical cyclizations, metallosupramolecular coordination chemistry, multichannel chemosensors for lab-on-microchip applications and, starting recently, molecular machines.*

molecules present in the mixture without any additional external input. To achieve effective error correction during the sorting process, any interaction between the building blocks has to be reversible in nature.

Self-sorting is a universal phenomenon that has manifold appearances in Nature. Actually, *order-out-of-chaos* processes are omnipresent in Nature, ranging from the formation of galaxies and weather patterns to biological self-organisation.[5] Examples of self-sorting are particularly well established in biological processes,[6] such as in DNA double helix formation: there, four bases (adenine (A), thymine (T), guanine (G), and cytosine (C)) automatically self-sort to form two base pairs (AT and GC)[7] allowing to store an immense amount of information in a specific sequence of the DNA. Furthermore, carbohydrates, peptides and fatty acids undergo self-sorting in the construction of a cell.[6] On a non-molecular level, self-sorting of white blood cells[8] and microtubules[9] has equally been well documented, and in many ways, we even encounter self-sorting in some of our daily life experiences. For instance, oil and water do not mix, but when mixed together they self-organise into two separate layers.

The above examples convincingly demonstrate that Nature efficiently uses the principles of self-assembly/self-organisation in combination with self-sorting protocols to construct intricate, functional architectures, which are capable of responding to chemical or physical stimuli of their environment, of displaying adaptive behaviour, and of undergoing evolutionary developments.[10] At present, the simulation or even reproduction of complex biological functions in man-made artificial devices and machines is still not in reach. It will demand a hitherto unprecedented error-free self-assembly/self-organisation level to manage and guide a crowd of different components in multiple orthogonal interactions.[11] As a consequence, artificial self-assembly amalgamated with self-sorting algorithms has received considerable attention over the past two decades, but is still far away from a mature state. Only in few cases, self-sorting has led to the production of a single aggregate as in biological self-assembly. In artificial self-sorting, the majority of processes rather leads to the formation of multiple assemblies or to the formation of a single assembly along with excess free ligand(s).

Both thermodynamic and kinetic control of self-sorting is known. Clearly, any thermodynamically controlled outcome will depend on the difference in Gibbs free energies ($\Delta\Delta G_{rxn}$) of all possible pathways, and only with $\Delta\Delta G_{rxn}$ being sufficiently large, a substantial difference in the products' population will be observed. On the other hand, kinetically controlled self-sorting is conditional and solvent dependent.[12] In some cases, even addition of external stimuli (*e.g.* oxidant) is necessary.[13,14] A well-known example is fractional crystallisation in a DCL (Dynamic Combinatorial Library): while many related molecules are present in the same solution, molecules of one kind recognise and interact with themselves at much faster rate to form well-ordered three-dimensional aggregates (crystals), thus ignoring all other molecules in the mixture (Fig. 1).[12b]

A large amount of self-sorting systems rely on either geometric match of their global shapes or match of their local interactions. Obviously, the difficulty of finding effective self-sorting algorithms increases with increasing similarity of the individual components, because structural differences, on which the discrimination has to be based, become smaller and smaller. Detailed comprehension of intrinsic factors, often described as "molecular programming",[15] and of external parameters is thus required to ultimately furnish an optimised blue print of highly selective self-sorting processes. For further progress, it is important to investigate the wide range of potential variables that may influence self-sorting and to identify the main player(s).

In a landmark paper, Isaacs and coworkers raised the question of whether self-sorting in appropriately instructed systems should be expected or not; thus, is it the "*exception or the rule*".[3a] To address this question, they investigated the sorting of nine components, *i.e.* **1–9**, in competition. Previously the individual components had been shown to give rise to well-defined supramolecular architectures. Experimentally, the competitive self-sorting of **1–9** in $CDCl_3$ affords eight distinct supramolecules in presence of barium picrate (Scheme 2). The fidelity of self-sorting is controlled by the pattern of H-bonding, spatial orientation, and presence of a closed network of hydrogen bonds. Equally, the temperature, concentration, association constants of complementary pairs, and presence of competitors influence the observed sorting phenomenon. After thoroughly investigating the effects at different conditions, the authors conclude that only small differences in the equilibrium constant (<10-fold) are necessary to promote interesting discrimination.[3a] When the equilibrium constant of a homomeric species is at least 100-fold higher than that of heteromeric assemblies, self-recognition products are present to more than 98% in the mixture. These results certainly demonstrate that self-sorting is neither the exception nor the rule. Based on the above results, Issacs summarises "*...the subset of known molecular aggregates that exceed the criteria required for thermodynamic self-sorting is larger than previously appreciated and potentially quite broad. This realization, in turn, offers a straightforward method for the preparation of complex, potentially functional, self-sorting systems.*"

Excellent reviews have recently become available by Würthner *et al.*[16] and Miljanić & Osowska[17] covering the topic in more detail. To our perspective, the level and advancement of sorting processes also depend on the process itself, thus suggesting to distinguish them on some quantitative scale. In the

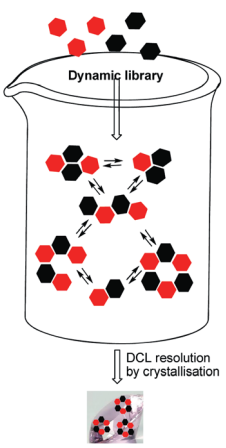
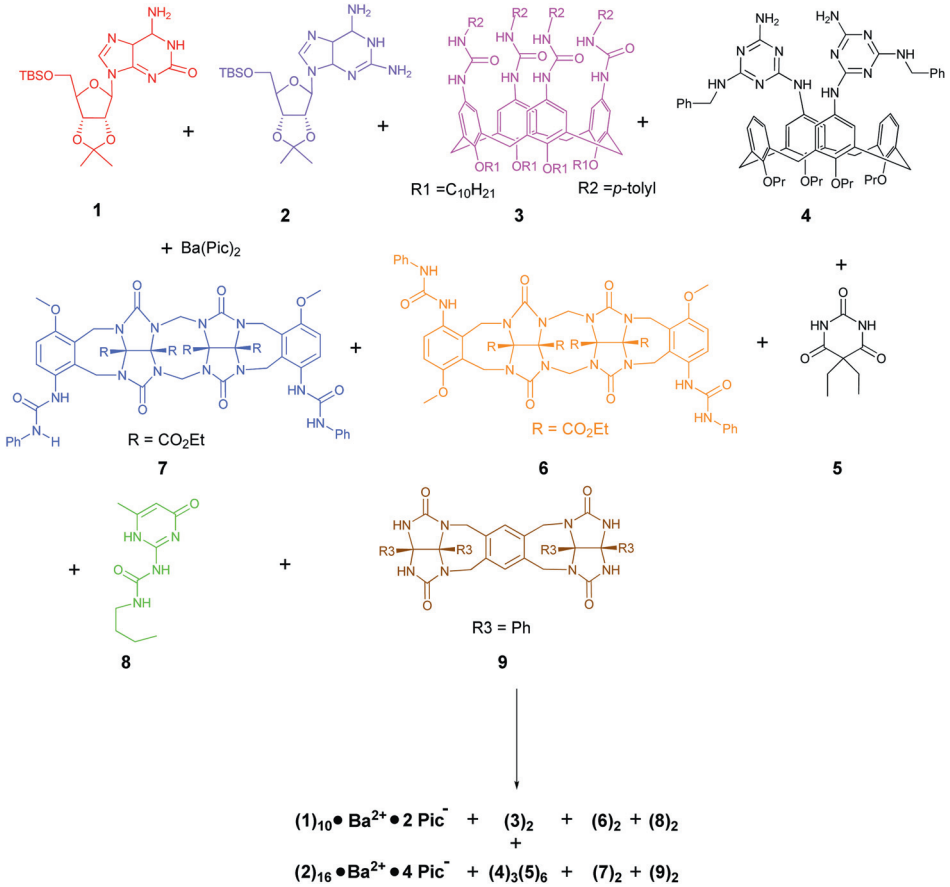



**Fig. 1** Crystallisation-induced selection from a Dynamic Combinatorial Library (DCL).



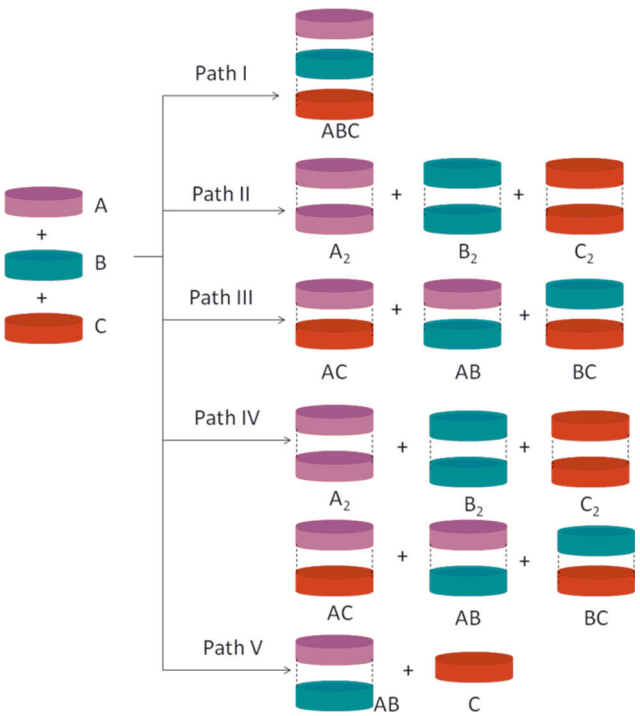

**Scheme 2** *Narcissistic self-sorting* of a complex mixture of hydrogen-bonding species. Barium picrate is added to aid the self-assembly of **1** and **2**.[3a]

present tutorial review, we focus on developing a systematic classification for the reported self-sorting systems and an evaluation of their *degree of self-sorting* (*M*).[18]

## 2. Classification of self-sorting systems

In 2003, Isaacs *et al.* classified artificial self-sorting into two categories; (a) *social self-sorting* and (b) *narcissistic self-sorting*.[3] Social self-sorting occurs between different species (Scheme 3, path I and III) whereas a narcissistic sorting process[1] takes place between the same species (Scheme 3, path II). In 2009, Stang and co-workers added two new terms in order to distinguish artificial self-sorting: (a) *absolute self-organisation* (Scheme 3, path I–III), and (b) *non-absolute self-organisation* (Scheme 3, path IV).[19] According to Stang's classification the exclusive formation of either homomeric or heteromeric aggregates from different components should be considered as absolute self-organisation, while formation of a mixture of homomeric and heteromeric aggregates would be designated as non-absolute sorting.

Although, both classifications help to discern between artificial self-sorting processes, they have obvious limitations due to the existence of only two categories. According to Isaacs's classification, the two paths I and III in Scheme 3 represent social self-sorting although they are chemically quite different. Path I utilises all components for a single assembly, whereas in path III each component ends up in all possible hetero-dimeric




**Scheme 3** Selected scenarios of self-sorting in a 3-component system.

assemblies. Furthermore, we cannot properly assess path IV and path V in this classification. Along Stang's categorisation, paths

I–III (Scheme 3) have to be described as *absolute self-sorting* while path IV constitutes a case of *non-absolute* self-sorting. Quite apparently, the classification is not instructed to grasp the differences in the sorting processes I–III.

Recently, Schalley and co-workers advocated to classify certain self-sorting systems as *integrative* or *nonintegrative*.[15,20] In *nonintegrative* systems, the components of the mixture combine to yield more than one final complex (Scheme 3, paths II–IV), whereas in *integrative* systems all sorted species present in the *nonintegrative* mixture are integrated into one single assembly through the use of multiple binding sites with positional control (Scheme 3, path I and chapter 4). Slightly later, Schmittel & Mahata coined the expressions *completive* and *incomplete self-sorting*.[18] *Incomplete self-sorting* leads to the formation of one or several assemblies along with unused components (Scheme 3, path V),[21] whereas *completive self-sorting* makes quantitative use of all members of the library to produce one (*i.e.* 1-fold completive, Scheme 3, path I) or several (*i.e.* *x*-fold completive ($x > 1$: Scheme 3, paths II–IV)) assemblies. In this review we would like to extend this nomenclature by denoting additionally the number of components in the mixture (in brackets) and the total number of particles accumulated in individual self-sorted assemblie(s) (as a superscript to 1 or *x*, being the number of self-sorted assemblies): For example in Scheme 3, path I represents $1^3$-fold(3) completive self-sorting as a single assembly arises out of 3 components. For the incomplete scenario we further address how many components remain unused. The number of components is now presented as a summation of components used in the self-sorting (1. number) and of components not participating in the self-sorting (2. number). For example, in Scheme 3, path V represents a $1^2$-fold(2 + 1) incomplete scenario.

To distinguish even further between the various processes, we herein list them in three subcategories depending on the nature of sorting: (i) *homomeric self-sorting, i.e.* formation of homoleptic assemblies (Scheme 3, path II); (ii) *heteromeric self-sorting, i.e.* molecules displaying a high affinity for others may form exclusively hetero aggregates (Scheme 3, path III); (iii) *mixed self-sorting* (Scheme 3, path IV) describes a cocktail of both homomeric and heteromeric assemblies. In principle, there exists a "window of possibilities" (Fig. 2) ranging from statistical mixtures to clean self-sorting. Accordingly, in Scheme 3, path I depicts a heteromeric $1^3$-fold(3) completive self-sorting, path II a homomeric $3^{2,2,2}$-fold(3) completive process, path III a heteromeric $3^{2,2,2}$-fold(3) completive scenario, and path IV a mixed $6^{2,2,2,2,2,2}$-fold(3) completive self-sorting and path V a heteromeric $1^2$-fold(2 + 1) incomplete scenario.

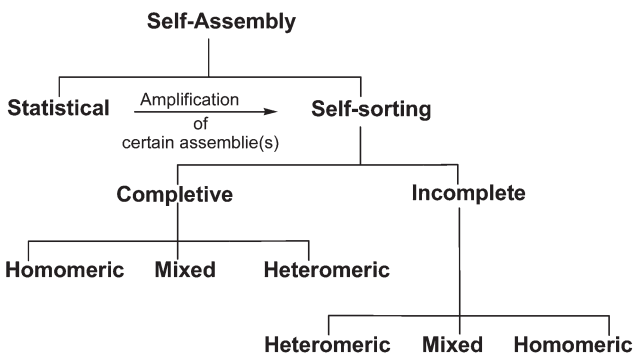



**Fig. 2** Classification of self-sorting in artificial self-assembly.

## 3. Degree of self-sorting (*M*)

Whilst self-sorting algorithms are differing in complexity the obvious question arises: *what is the level of sorting in a self-sorted system*? To address this question, Schmittel suggested in 2009 to define the *degree of self-sorting* (*M*) for mononuclear assemblies as $M = P_0/P$, with $P_0$ being the number of all possible aggregates[22] and *P* being the number of all observed assemblies in the experiment.[18] Accordingly, the maximum value for *M* is $P_0$ indicating the highest level of self-sorting, and on the other side, *M* will be 1 when $P = P_0$. The numerical value of *M* may allow differentiating quantitatively between the various sorting processes. Clearly, with increasing *M* value the system will show a smaller number of products. Therefore, the higher the *M* value should reflect a higher emergence of artificial self-sorting systems.

Although in many publications on self-sorting the authors have themselves provided $P_0$, the assessment of this number involves some arbitrary assumptions rendering it highly disputable, even among experts. For illustration, in repetitive cyclic assemblies, the unlimited number of potential oligomeric aggregates increases $P_0$ to infinity. In reality, though, discrete structures are always preferred over oligomeric ones due to entropic reasons. Thus, for a reliable and manageable determination of $P_0$ we assume for our analysis that any self-sorting will be intrinsically guided by (i) the maximum site occupancy rule[1] and (ii) low entropic costs. Thus, a self-sorting protocol will produce discrete products of the lowest entropic costs, in which all possible interactions (metal–ligand binding, hydrogen bonding, *etc.*) are realised.[23] Obviously, the use of *M* also has limitations. According to our definition, a self-sorted system will have an *M* value of 1 once producing all entities that are to be expected according to points (i) and (ii). Consequently *M* can only be used to identify those self-sorting aspects that go beyond the design criteria of (i) maximum site occupancy and (ii) low entropic costs. Despite this shortcoming, *M* still provides useful information about the degree of self-sorting, and we will apply this systematics for the present review.

There are further shortcomings concerning the informative value of *M*, for example, when all possible products are formed but not in a statistical product distribution.[24–27] This setting represents a special case of *mixed self-sorting*, for which *M* fails to provide any incisive information. Here it is worthwhile to determine the *amplification* value, *i.e.* the extent of experimental abundance that exceeds the statistical probability, according to a suggestion of Barboiu *et al.*[25] The percentage of amplification can be calculated using the difference between the experimental and the statistical abundances (*a*) as follows, *amplification* $A = [(a_{\text{experimental}} - a_{\text{statistical}})/a_{\text{statistical}}] \times 100$. Even the latter approach, however, has shortcomings. For example, it is not possible to describe quantitatively combinatorial libraries that show sometimes substantial self-sorting, as not only one but several products may form far beyond the statistically expected amounts.

To evaluate various sorting processes with respect to their *M* value let us scrutinise first the 2-component self-sorting system in Scheme 1. For the heteromeric $1^2$-fold(2) completive self-sorting system (path b) *M* is 3, whereas *M* is 3/2 for the homomeric $2^{2,2}$-fold(2) completive self-sorting (path a). This finding suggests that a 1-fold completive self-sorting, with formation of a single assembly, generates a higher genuine degree of self-sorting and is intrinsically heteromeric (*social*) in nature. Accordingly, an increase in number of components augments the complexity of any 1-fold completive self-sorting but at the same time the number of particles, whether small or large as in a cuboctahedron, does not find reflectance in *M*.[28] How is this possible? Apparently, *M* describes the complexity of the self-sorting process and not the complexity of the final assembly itself.[29]

While chemical assemblies composed of more and more dissimilar components are likely to reflect a large *M*, we realise that the same *M* value in, for example, various 1-fold completive processes may belong to conceptually quite different protocols with one being elegant and the other considered routine.

1-Fold completive self-sorted systems may be more attractive than *x*-fold completive or incomplete self-sorted ensembles because they are more apt for chemical analysis and for establishing a context between structure and function. However, incomplete self-sorting may become quite useful in many practical applications, in particular those that require repair or rejuvenation due to irreversible decomposition of one component. We may grasp the *M* value of incomplete self-sorting by the same formula as given above for completive recognition, except when the unused components were not true competitors but irrelevant spectators (Scheme 4).

In the ensuing selected examples, we will first discuss the guiding principles of self-sorting lined up according to their *degree of self-sorting* (*M*) value. We will also describe the *amplification factor* (*A*) for some special self-sorting systems. Due to its importance in synthetic multicomponent ($n \geq 3$) supramolecular chemistry, we have dedicated an independent section to *1-fold completive self-sorting*.

To limit the scope of this tutorial review, we have excluded self-sorting in polymers, surfaces, solid-state assemblies and gels thus only considering solution based discrete assemblies with high purity. Furthermore, we disregard heterochiral aggregation and enantiomeric self-recognition as they are special cases of self-sorting, allowing us to avoid their discussion in any great detail here. For interested readers, the recent review by Würthner *et al.*[16] covers chiral self-sorting in all facets.

### 3.1. *M* in *x*-fold completive ($x > 1$) self-sorting systems

In the following, we describe selected cases, in which self-sorting does not lead to a single product. Basically all of these thermodynamically controlled sorting processes are operating through maximum site occupancy and lowest possible entropic costs. However, when $M > 1$, additionally steric constraints, preferential geometry, length of the ligands, *etc.* are used to set up self-sorting.

**3.1.1. *M* = 1.0.** As pointed out above, the level of self-sorting as defined by *M* relies heavily on the number of possible aggregates $P_0$. Accepting our definition of $P_0$, self-sorting at the level $M = 1.0$ indicates that the sorting has exclusively been based on maximum site occupancy combined with generating assemblies at the lowest possible entropic costs. Along these principles, several classics of self-sorting have been conceived in the 1990s.

In 1993, Lehn *et al.* evaluated the spontaneous formation of the discrete supramolecular double stranded helicates **10–13** (Scheme 5) from a mixture of four oligo-bipyridines of different length. The homomeric helicates **10–13** assort in a selective manner, directed by the preferred tetrahedral coordination geometry of the copper(I) ions and by the number of binding sites available in each strand.[1] Any mixed ligand helicate would thus be destabilised by free coordination sites or by larger entropic costs. The exclusive formation of the discrete double helicates is therefore to be expected in light of the maximum site occupancy principle.

An aesthetically pleasing case, due to the difference in number of binding sites, has been described by Reinhoudt *et al.* with the spontaneous formation of the hydrogen-bonded rosettes $(\mathbf{4})_3(\mathbf{5})_6$ and $(\mathbf{14})_3(\mathbf{5})_{12}$ (Scheme 6).[31] A mixture[32] of **14** : **4** : **5** in a ratio of 1 : 1 : 6 led to exclusive formation of the above mentioned rosettes. The observed selectivity is mainly due to the molecular shape and hydrogen-bonding complementarity that are encoded into the di- or tetramelamine building blocks. Accordingly, maximum site occupancy dictates three melamine units of **4** and **14** to combine with six and twelve units of 5,5-diethylbarbituric acids (**5**) respectively, highlighting that structural information can efficiently drive the self-sorting processes.

In 1999, Taylor and Anderson reported on examples of *narcissistic self-sorting* in linear conjugated zinc(II)porphyrin oligomers, leading to the formation of dimeric oligoporphyrin ladder arrangements, all held together by DABCO molecules.[2] As an example, they described the clean formation of the dimeric ladder complexes **18** and **19** from mixtures of **15**, **16**, and DABCO (Scheme 7). Following guidance from maximum site occupancy and minimal entropic costs, only two homo-ladders **18** and **19** sort out in a selective manner. The selective formation of these ladders is directed by the number of porphyrin units available in each zinc(II)porphyrin oligomer. The alternative mixed ligand ladder would be destabilised by either free coordination sites or large entropic costs.

Utilising a clever design on the basis of highly stable and directionally well-defined amidinium-carboxylate salt bridges, such as depicted by formula **20** (Fig. 3), Furusho, Yashima *et al.* recently described the spontaneous formation of double helices consisting of two strands with the same number of complementary binding sites.[33] Notably, when six trimeric molecular strands (AAA, CCC, AAC, CCA, ACA, and CAC) were mixed in solution, the complementary strands were sequence-specifically hybridised to form the one-handed double-helical trimers AAA·CCC, AAC·CCA, and ACA·CAC through complementary amidinium–carboxylate salt bridges. Circular dichroism (CD) and $^1$H-NMR data support the assignment. Statistically, the above 6-component library is able to form 21 different double helices,[34] but if we account for maximum site occupancy effects and lowest entropic costs, only three structure are likely. Thus $M = 3/3 = 1$. Furthermore, when CCA was added to a mixture of AAA, AAC, and ACA, the AAC·CCA double helix was

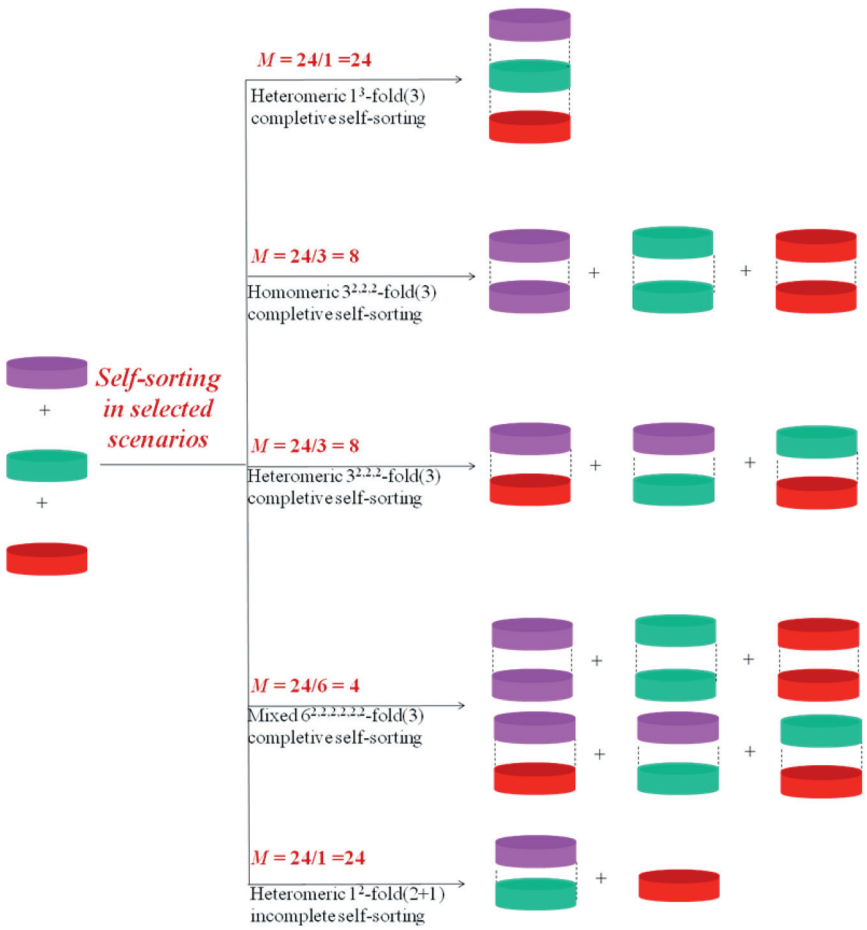
**(a) Homomeric/ heteromeric/ mixed $P^{y1,y2,..yp}$-fold(n) completive self-sorting, degree of self-sorting ($M = P_0 / P$)**

n = Total number of components, $P_0$ = The number of all possible aggregates, $P$ = The number of all observed aggregates, and y1, y2,...yp represent the total number of particles associated with found assemblies 1, 2...p in ascending order.

**Homomeric/ heteromeric/ mixed $P^{y1,y2,..yp}$-fold(n1+n2) incomplete self-sorting, degree of self-sorting ($M = P_0 / P$)**

n1 = Number of components used in the self-sorting and n2 = number of unused components.

**(b)**



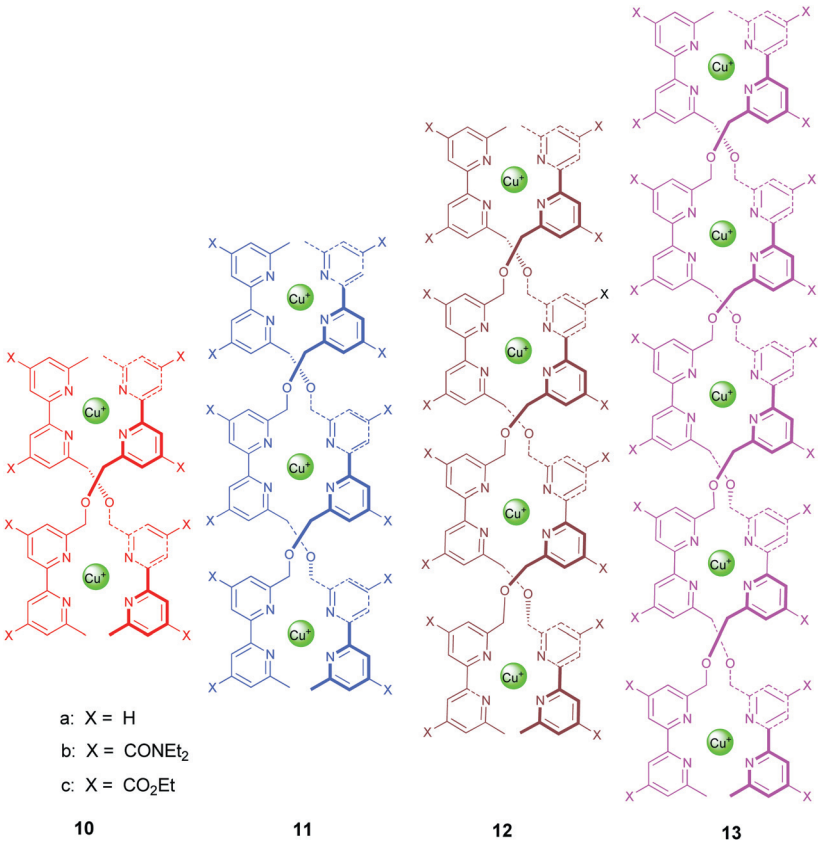

**Scheme 4** (a) Descriptor for *completive* and *incomplete* self-sorting. (b) Depiction of the various scenarios in a selected 3-component self-sorting.[30]

selectively formed. In a similar manner, the homooligomers composed of amidines or carboxylic acids (A, AA, AAAA, C, CC, and CCCC) assembled with a precise chain length specificity to form A·C, AA·CC, and AAAA·CCCC.

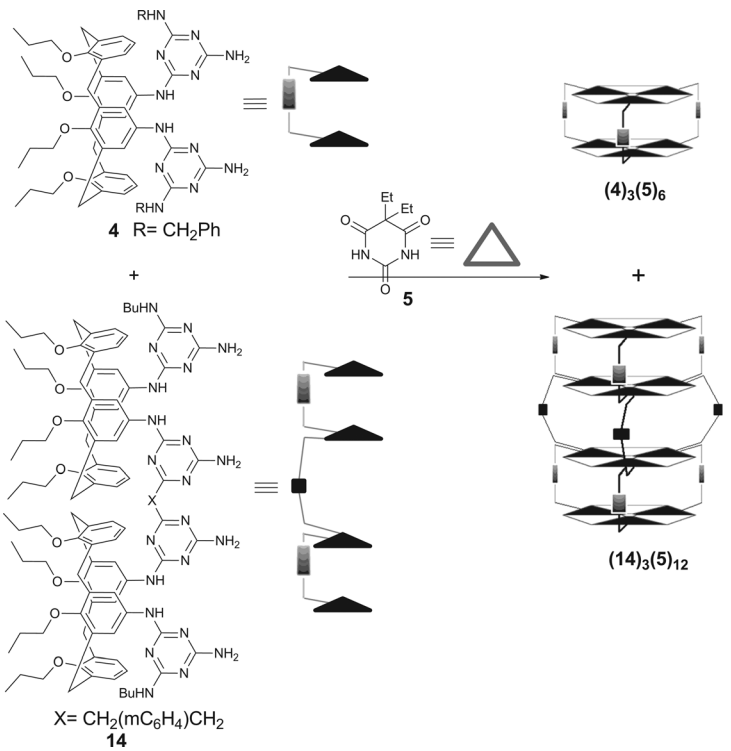
**3.1.2. 1.0 < $M$ ≤ 1.5.** Over last decade, Stang and coworkers have been successfully employing the "directional bonding model" with its precise control of geometric factors to prepare a variety of supramolecular structure(s).[35,36] In the ensuing recent example (Scheme 8), they described how the numbers of binding sites of donors and their directionality affect the self-selection in both 2D and 3D supramolecular polygons and polyhedra.[37] For example, the ditopic ligands **22** and **23** and the tritopic ligand **24** were successfully self-sorted into the differently sized 2D rectangles **25** and **26** as well as the 3D triangular prism **27** (Scheme 8). By considering the maximum site occupancy rule and lowest entropic costs, such system is still able to produce 4 discrete nanoassemblies, thus suggesting $M$ = 4/3 = 1.3.

***Model I.*** For a methodical presentation of self-sorting, we will first portray the ensuing systems according to the general model I. As denoted in Scheme 9, A and B are two molecules that, in

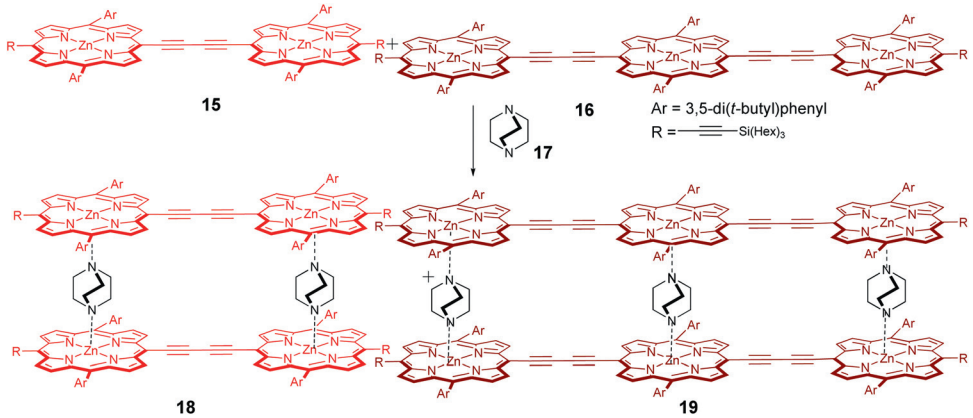
 




**Scheme 5** The four discrete supramolecular double-stranded helicates **10(a–c)**–**13(a–c)** formed in a homomeric $4^{4,5,6,7}$-fold(5) completive self-sorting.[1]




**Scheme 6** $2^{9,15}$-Fold(3) completive self-selection in a mixture of **4**, **5** and **14** leading to the formation of H-bonded rosettes $(\mathbf{4})_3(\mathbf{5})_6$ and $(\mathbf{14})_3(\mathbf{5})_{12}$.[31]




**Scheme 7** $2^{4,5}$-Fold(3) completive self-sorting of a mixture of **15** and **16** in presence of DABCO.[2]

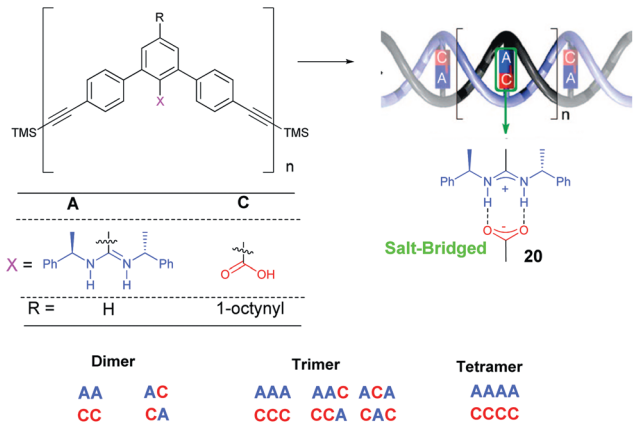



**Fig. 3** Structures of *m*-terphenyl-based molecular strands bearing amidine (A) and/or carboxyl (C) groups and an illustration of the formation of double-helical oligomers consisting of complementary molecular strands stabilised by amidinium–carboxylate salt bridges.[33] The helical artwork is reprinted with permission from ref. 33. Copyright 2008 American Chemical Society.

light of our general guiding principles, may combine to either homomeric ($A_2$ and $B_2$) or heteromeric (AB) aggregates. In the homomeric 2-fold completive self-sorting scenario only two assemblies $A_2$ and $B_2$ are observed. The *M* value for such system is always 3/2 = 1.5.

In 1999, Taylor and Anderson[2] successfully exploited the length difference in linear conjugated zinc(II)porphyrin dimers to set up self-recognition in the supramolecular ladders **18** and **29**. When DABCO is added to a 1 : 1 mixture of the butadiyne-linked dimer **15** and its 1,4-diethynylbenzene-linked analogue **28**, mixed complexes such as **15**·**28**·$(DABCO)_2$ do not form. Only observed are the two homoladders **18** and **29** (Scheme 10), ostensibly due to an effective self-selection based on the distance of the binding sites.

Recently, Stang *et al.* exploited the length difference of linear bipyridine ligands in the construction of two-dimensional rectangles, triangles and three-dimensional cages.[4] For example, they utilised the length difference in the ditopic donors **22** and **23** to select the discrete rectangles **25**, **26** (Scheme 11). Thus, here also the size of the linkers directs the self-selection, aside of maximum site occupancy and entropic factors.

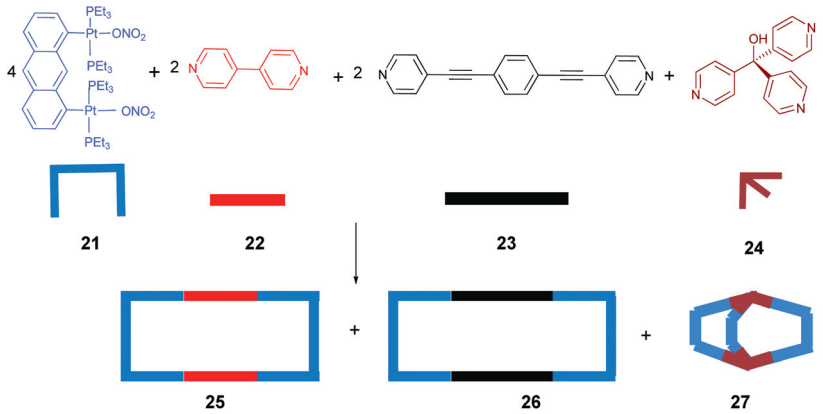
Dalcanale and co-workers have developed the self-recognition of tetradentate deep-cavity cavitand ligands during the formation of nanoscale coordination cages.[38] Upon addition of palladium(II) or platinum(II) building blocks, the cavitands **30** or **31** self-assemble into the dimeric capsules **32** and **33**, respectively, of varying depth, depending on the length of the pyridyl anchoring units. Likewise, competition experiment, in which the cavitands **30** and **31** were mixed with a stoichiometric amount of [Pd(en)]$(CF_3SO_3)_2$ in acetone-$d_6$ (en = ethylenediamine), showed the exclusive formation of homocages **32** and **33** (Scheme 12). Apparently, the geometrical mismatch between the biting angles of the two cavitands leads to complete self-sorting during the self-assembly process.

In a recent example, Maeda and his group reported on carboxylate-appended pyrrole based π-conjugated molecules that form self-assembled dimers and exhibit narcissistic self-sorting depending on the substituted positions of the anionic sites.[39] As an example, they monitored the narcissistic self-sorting behaviour of the dimers $(\mathbf{34})_2$ and $(\mathbf{35})_2$ in $CDCl_3$ by $^1$H-NMR spectroscopy. The data suggest that there are no signals from the mixed dimer **34**·**35** (Fig. 4b). According to theoretical calculations, the aforementioned homomeric sorting is mainly due to difference in the appropriate hydrogen-bonding directionality and less steric hindrance. As a result, $(\mathbf{35})_2$ is stabilised over both **34**·**35** (+2.26 kcal mol$^{-1}$) and $(\mathbf{34})_2$ (+2.36 kcal mol$^{-1}$).

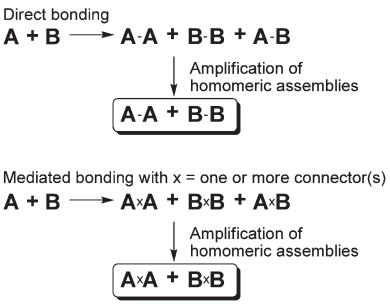
The factors influencing self-sorting discussed so far have been essentially limited to those with differences in molecular shape and size. Rather little attention has been paid to realise self-sorting driven by hydrophobic and hydrophilic interactions. In 2001, Kumar and coworkers successfully utilised the hyperstability of a fluorinated homodimeric coiled coil assembly to initiate the self-recognition of peptides **36** and **37**.[40] Both **36** and **37** are designed to form parallel homodimeric coiled coil assemblies. They have an identical sequence except that in **37** seven core leucines are replaced by fluorinated leucines to enhance the hydrophobic forces in the fluorinated homodimer $(\mathbf{37})_2$. Starting from the heterodimer **36**·**37**, the pure homodimers $(\mathbf{36})_2$ and $(\mathbf{37})_2$ formed by disproportionation under disulfide exchange conditions (Scheme 13). Hydrophobic interactions are the major driving force for such self-assembly. Later, the same group also used the same strategy for programming specific protein–protein interactions, equally based on the coiled coil motif with the






**Scheme 8** Homomeric $3^{4,4,5}$-fold(4) completive self-recognition of ligands **22–24** affording the 2D rectangles **25** and **26** and the 3D prism **27**.[37]




**Scheme 9** A homomeric 2-fold completive self-sorting scenario (model I).

hydrophobic core composed exclusively of phenylalanine residues. The resulting phenylalanine interface exclusively self-sorted from a related peptide assembly containing aliphatic residues in the core.[41]

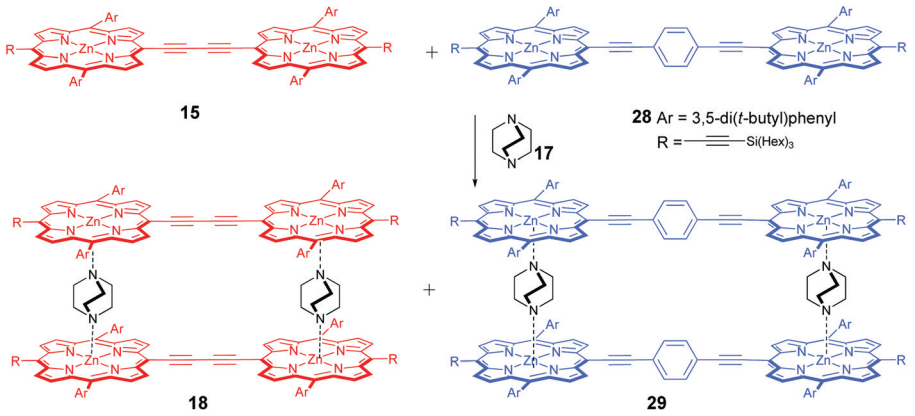
**3.1.3. *M* = 2.0.** Pei *et al.* probed the length difference of the aromatic diamines **40** and **41** in the context of their selective inclusion into the 'smart' macrocycles **38** and **39**.[42] As an illustration, they mixed equimolar amounts of **38–41** in $CDCl_3$ with excess (100 equiv.) of TFA. Under kinetically controlled conditions for the condensation they observed a reaction of **38** with **40**, and one of **39** with **41**. However, they equally observed some minor inclusion of **40** in **39**. The ratio of matched and mismatched pairs was estimated to be more than 90 : 10. After heating the same sample for 1 h, the matched pairs were finally obtained in nearly quantitative yield (Scheme 14). Out of 4 possible combinations (**42a**, **42b**, **43a**, **43b**) the two assemblies **42a** and **43a** formed exclusively by virtue of *heteromeric self-sorting*, furnishing $M = 2$.

The observed selectivity is explained by the conformational preference of the highly flexible oligo(ethylene glycol) unit. According to theoretical calculations, the macrocycle in **42a** adopts a *trans–gauche–trans* conformation as the most stable conformation. However, in the alternative product **42b**, the oligo (ethylene glycol) chains are forced to extend almost linearly, and as a result, the enthalpic contribution increases due to repulsion of the chains preventing the formation of **42b**. In contrast, **43a** and **43b** form in enthalpically comparable processes. Because **42b** is thermodynamically disfavoured, only **42a** is formed, while **39** is forced to transform into **43a** by virtue of maximum site occupancy.

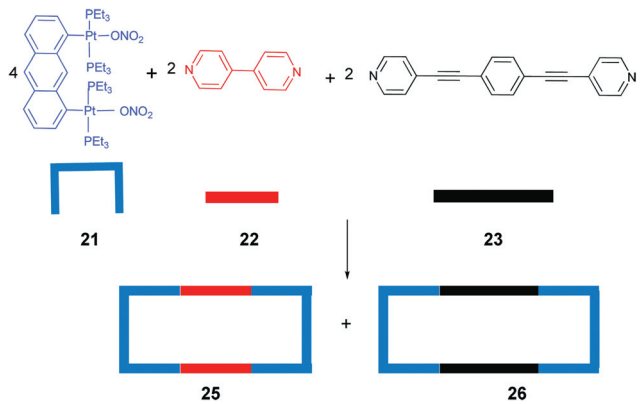
Based on the difference of the twist angle between two naphthalimide subunits, Li *et al.* recently reported on the self-selection behaviour of three bay-substituted perylene bisimide dyes **44–46** (Fig. 5).[43] The observed selectivity is due to the fact that bulky bay substituents twist the perylene unit dihedrally out of plane with angles ranging from 0 to 37, thereby influencing the available contact surface area for π–π stacking. To our appreciation, $M = 2$ because out of 6 homo- and hetero-dimers only 3 homodimers are formed.

In 1999, Albrecht *et al.* described the alkali metal-ion templated self-recognition of the alkyl-bridged bis-(catechol) ligands **47** and **48** in presence of $Ti^{4+}$ salts.[44] With $K^+$ or $Li^+$ as template, mixtures of homoleptic and heteroleptic helicates as well as polymers formed. In contrast, with $Na^+$ as countercation, exclusive formation of the dinuclear titanium(IV) helicates **49** and **50** is observed (Scheme 15). Thus, the choice of the counter cation is crucial in order to induce a specific selectivity. The large effect seems to be due to an electrostatic stabilisation or destabilisation between the anionic titanium fragments and the templating countercation. The system therefore ends up with two homohelicates out of 4 possible homo- and heterohelicates, suggesting an $M$ value of 4/2 = 2.

**3.1.4. 2.0 < *M* ≤ 3.0.** Within this category most of the reported examples exhibit $M = 3$, but the following one has a lower value. Based on steric demands, Hooley & Johnson developed the self-discrimination of the endohedrally functionalised bis(pyridine) **52** in presence of coordinating metal salts, such as $Pd(NO_3)_2$.[45] Tuning the size of the internal substituent allows selective heterocluster formation by noncovalent and space-filling interactions. Thus, when **51** was combined with trifluoroacetate ligand **52** and $Pd(NO_3)_2$, only two complexes were detected: the **53** = $(\mathbf{51})_4Pd_2$ cluster and the **54** = $(\mathbf{51})_3(\mathbf{52})Pd_2$ cluster (Scheme 16). Therefore, out of 5 possible homo- and hetero-combinations only one homo- and one hetero assembly are observed,[46] suggesting an $M$ value of 5/2 = 2.5. In a control experiment, the bulky trifluoroacetate ligand **52** does not show any self-assembly with $Pd(NO_3)_2$ and molecular modeling of the $(\mathbf{52})_4Pd_2$ complex proposes that the arrangement of four




**Scheme 10** $2^{4,4}$-Fold(3) completive self-sorting of a mixture of **15** and **28** in presence of DABCO.[2]



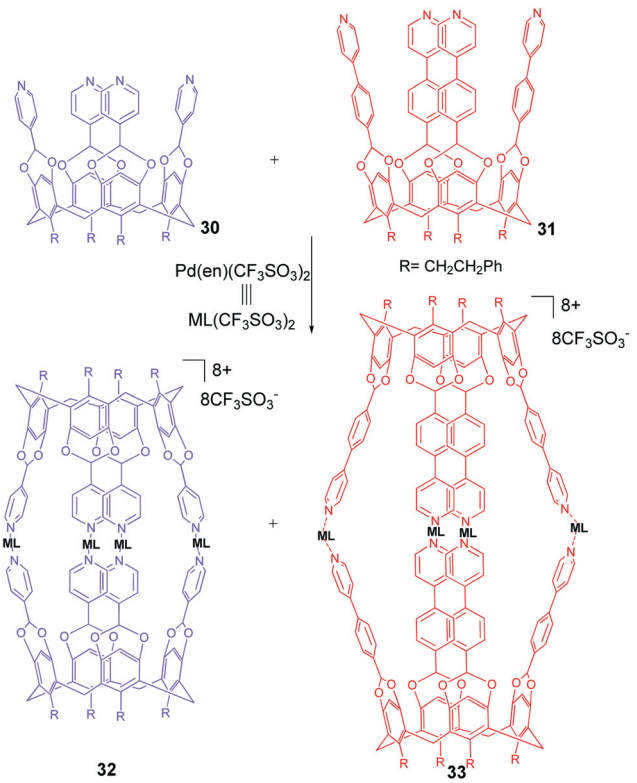

**Scheme 11** $2^{4,4}$-Fold(3) completive self-recognition of ditopic donors **22** and **23** in the formation of small rectangle **25** and large rectangle **26**.[4]

trifluoroacetamide (TFA) groups in the inside cavity is sterically impossible. It seems that one cannot place more than one TFA group in the small inside cavity.

Work by Severin *et al.* reveals the importance of geometric restraints in the self-selection of organometallic macrocycles. They describe the selective formation of 8 out of 24 possible trinuclear macrocycles ($M = 3$) from a mixture of the metal building blocks **55** and **56** as well as the bridging ligands **57** and **58**. The eight complexes contain either ligand **57** or **58**, but no combination of both. Self-sorting seems to originate from geometrical restraints: all complexes based on ligand **57** have a trigonal prismatic structure, whereas complexes based on ligand **58** exhibit a concave, domelike arrangement. Apparently, there is no adequate low-energy assembly for mixed complexes using both **57** and **58** (Scheme 17).[47]

Work by Barboiu *et al.* nicely demonstrates the potential of metal coordination in sorting processes.[48] Clean formation of the two homonuclear metallosupramolecular [2 × 2] grids **60** and **61** arises from a single ligand **59** in presence of both copper and silver ions (Scheme 18). The driving force for the observed selectivity is seen in the different coordination behaviour and ionic size of both metal ions that trigger a different conformational arrangement of the ligands in both grids. According to our analysis, for this self-sorting process $P_0 = 6$ and $M = 3$.

We have assigned the following example to $M = 3$, although our evaluation of $P_0$ may be disputed. In recent reports about




**Scheme 12** $2^{6,6}$-Fold(3) completive self-recognition of cavitands **30** and **31** upon addition of $[Pd(en)](CF_3SO_3)_2$ furnishing the homocages **32** and **33**.[38]

morphological switches,[49] Lehn *et al.* assessed how complementary pair selection can be altered by simply adding metal ions. In absence of metal ions, the authors observe a thermodynamically controlled mixture of the linear bisimine **66** formed by the reaction of *n*-octylamine (**65**) with the W-shaped dialdehyde **62** and of the cyclic diimine **67** after reaction of diamine **64** with the U-shaped dialdehyde **63** (Scheme 19). Using the maximum site occupancy principle but allowing for both [1 + 1] and [2 + 2] adducts with **64**, one would expect 6 possible structures. Lehn and coworkers only observe 2 products ($M = 3$).

Addition of lead triflate along with imidazole to a mixture of **66** and **67** resulted in an imine swapping to form products **68**



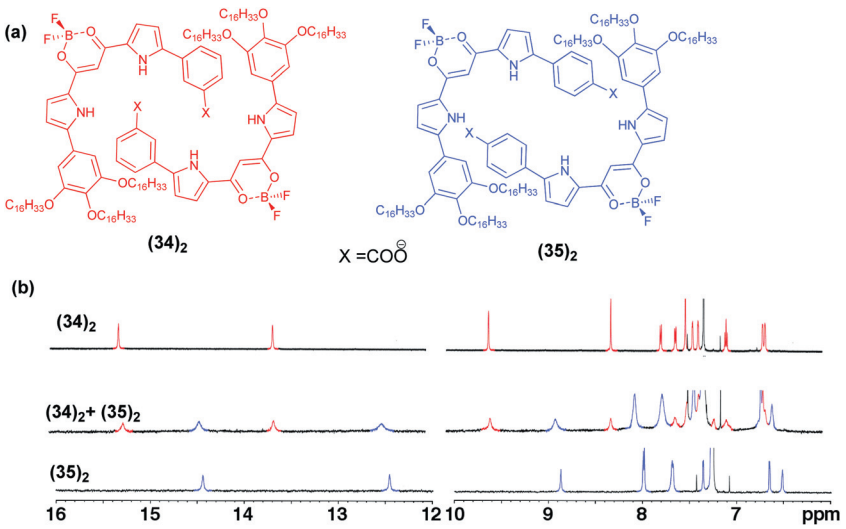
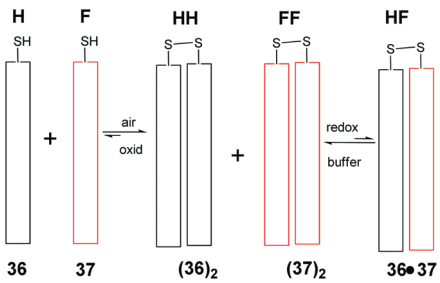



**Fig. 4** (a) Self-assembled dimers $(\mathbf{34})_2$ and $(\mathbf{35})_2$ from a homomeric $2^{2,2}$-fold(2) completive sorting. (b) $^1$H-NMR spectra of $(\mathbf{34})_2$ (red, top), $(\mathbf{35})_2$ (blue, bottom) and the mixture of equivalent amounts of $(\mathbf{34})_2$ and $(\mathbf{35})_2$ (middle) in $CDCl_3$.[39] Redrawn with support from the authors and reproduced from ref. 39 with permission from the Royal Society of Chemistry.




**Scheme 13** A homomeric $2^{2,2}$-fold(2) completive self-sorting of peptides **36** and **37** to afford coiled coil assemblies $(\mathbf{36})_2$ and $(\mathbf{37})_2$.[40]

and **69** supported by both metal ion coordination and morphological changes.[49a] Actually, the possibility of additional coordination through oxygen atom drives **64** to combine with **62** in complex **68**.

All other selected systems of this section will fit the two general models II and III. Therefore in the following, we will first describe the corresponding general model followed by some selected examples.

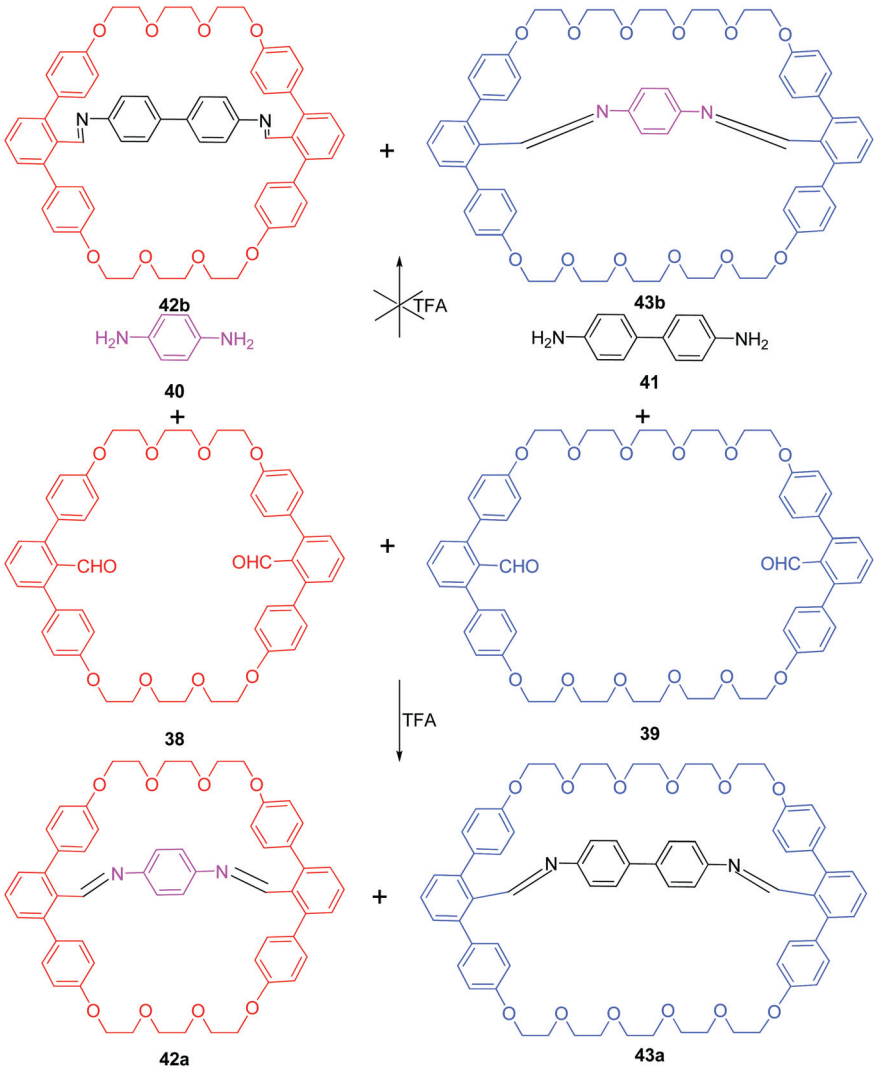
***Model II.*** The general setting for the 4-component model II is delineated in Scheme 20. For example, both singly- (C) and doubly-functionalised (D) guests are proficient to bind with hosts A, B. As a result of its two functionalities, D may generate two isomers upon binding. Thus, a mixture of A–D may combine to 6 possible inclusion complexes (AC, BC, $\overrightarrow{DA}$, $\overrightarrow{DB}$, $\overleftarrow{DA}$, and $\overleftarrow{DB}$). In a *heteromeric self-sorting* two assemblies, for example, AC and $\overleftarrow{DB}$ are forming exclusively, with the *M* value for such system being 6/2 = 3.

Isaacs *et al.*[50] and Masson *et al.*[51] probed the relationship between cavity sizes of cucurbit[*n*]urils and the steric demands of bulky guests. As a test, Isaacs and coworkers examined the complexation behaviour of **CB[6]** and **CB[7]** in presence the doubly-functionalised guest **70**, containing both a slim alkylammonium and a bulky adamantylammonium terminal, and the single-ended guest **71** (Scheme 21). Under kinetically controlled conditions (6 min after mixing) **70** inserts its slim alkylammonium binding epitope into **CB[6]** to form **77**, whereas **71** engages in the formation of **73**. After 56 days, however, a dramatically different selectivity was observed. After full thermodynamic equilibration and self-sorting, exclusive formation of the thermodynamically favoured state was observed as **72** and **75**, thus the degree of self-sorting amounts to *M* = 3. Now guest **70** has used its bulkier, but also much tighter adamantyl binding terminal for complexation.[50b]

In a related example with *M* = 3, Inoue *et al.* utilised cucurbit[*n*]urils of different cavity size to set up a sequence-specific recognition and self-sorting of dipeptides **78** and **79** that led to complexes **80** and **81** (Scheme 22). The observed selectivity may readily be rationalised: the smaller **CB[6]** forms only weak complexes with dipeptides having an aromatic residue at the *N*-terminus, but interacts more strongly with those having an aliphatic residue, particularly Lys, at the *N*-terminus. **CB[7]** forms a highly stable complex with aromatic residues at the *N*-terminus. Strong binding is achieved by a good fit of the guest's aromatic group in the CB cavity, by attractive electrostatic attractions between its ammonium group and the carbonyl oxygens of the CB portal, and additionally by extensive desolvation. The arrangement also allows avoiding electrostatic repulsions between the carbonyl oxygens and the guest's carboxylate.[52]

***Model III.*** Under the heading "model III" we denote systems of category *M* = 3 that follow the general schematics depicted in Scheme 23. For example, A, B and C are different molecules that bind to each other or oneself. Accordingly, they may combine to homomeric ($A_2$, $B_2$ and $C_2$) and/or heteromeric (AB, AC, BC) aggregates. In model III, only the two assemblies $A_2$ and BC do form by virtue of *mixed 2-fold completive self-sorting*, being characterised by an *M* value of 6/2 = 3.

Isaacs and others[53,54,55] performed extensive studies in order to determine the extent of self-sorting between various different




**Scheme 14** Heteromeric $2^{2,2}$-fold(2) completive self-sorting in the condensation reaction of macrocycles **38–39** and diamines **40–41**.[42]

complementary hydrogen bond donors and acceptors, as guided by the pattern of their H-bonding array. In one of the early reports, Isaacs *et al.* explained the self-selective dimerisation of molecular clips based on molecular shape and chirality (Scheme 24). The observed selectivity in the preferred self-sorting of racemic dimer $(\mathbf{83})_2$ is due to its extra stabilisation by an additional hydrogen bond.[53a]

Recently, Böhmer *et al.* successfully utilised appropriate steric programming in the ligand core to achieve self-sorting.[56] From a mixture of **84–86**, they observed selective formation of **87** and **88** (Scheme 25). The explanation for the observed selectivity is straightforward. The tetraloop compound **84** cannot form homodimers because the aliphatic chains connecting the adjacent urea units would have to overlap in a sterically very unfavourable arrangement. Furthermore, the bulkiness of the *tert*-butyl substituents attached to **86** strongly disfavours its intercalation in the small loops of **84**. This system thus can only evolve in one direction: **84** cannot homodimerise nor can it form heterodimers with **86**. Its only choice is to form a heterodimer with **85**. Left alone, **86** associates with itself, to form homodimer **88**.

**3.1.5. 3.0 < *M* ≤ 3.5.** In 1997, Raymond & Caulder successfully exploited the length difference in the rigid biscatechols **89–91** to set up self-recognition in the supramolecular triple helicates **92–94** (Scheme 26).[57] The observed self-selection is apparently governed by the thermodynamic preference for a finite system through maximum site occupancy and entropic constraints. Due to the possibility of forming 10 homo- and heteroleptic triple helicates, the exclusive formation of 3 discrete homomeric helicates represents a length-selective self-sorting process with $M = 3.3$. The rigidity and different distances of the coordinating catechol units in ligands **89–91** neatly direct the system towards exclusive formation of the homomeric helicates. Remarkably, when mixtures of any two or three of the ligands are reacted with gallium(III) ions, only complexes containing one type of ligand are formed and no trace of mixed-ligand species is observed in solution using $^1$H-NMR or electrospray ionisation (ESI) mass spectrometry.

Sanders and coworkers demonstrate the potential of "predisposition" in the self-selection of the two hydroxy esters **95** and **96**, a cinchonidine and a xanthene derivative, respectively

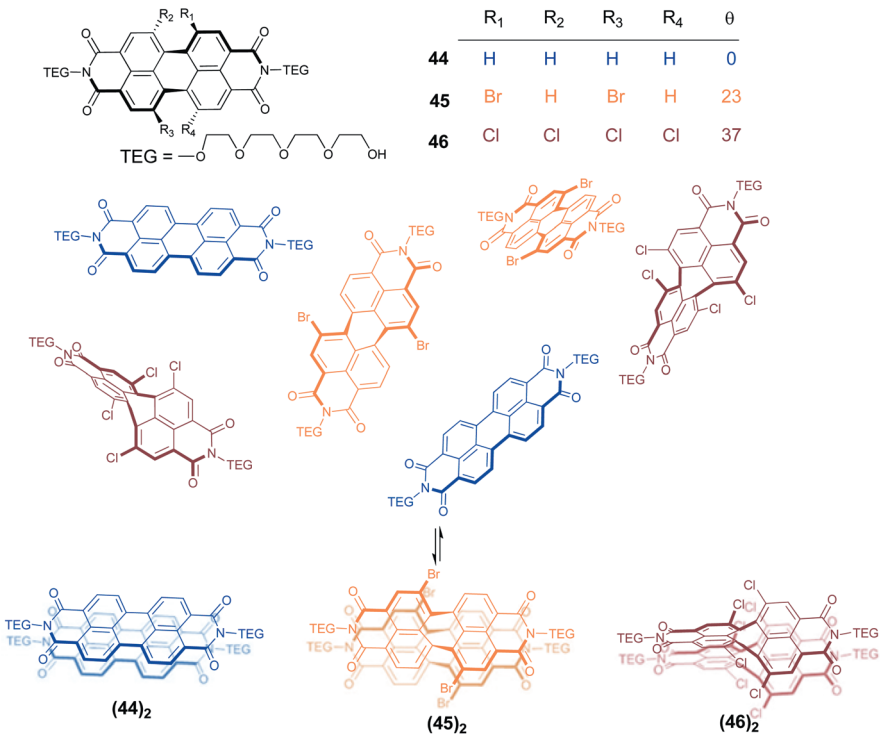
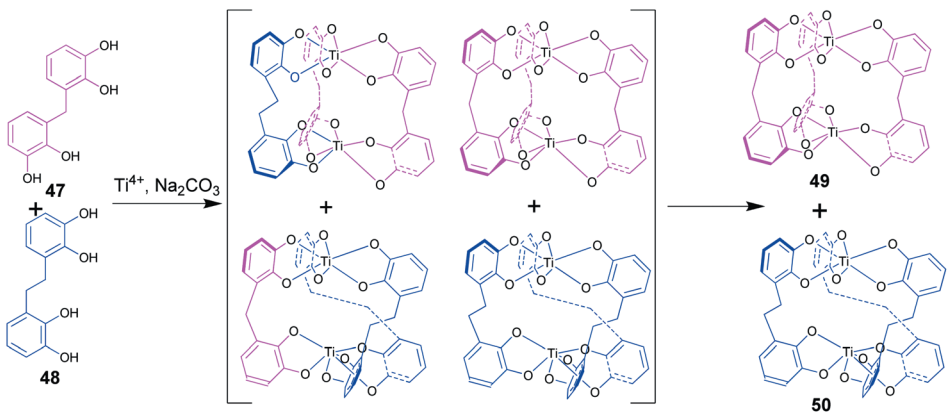
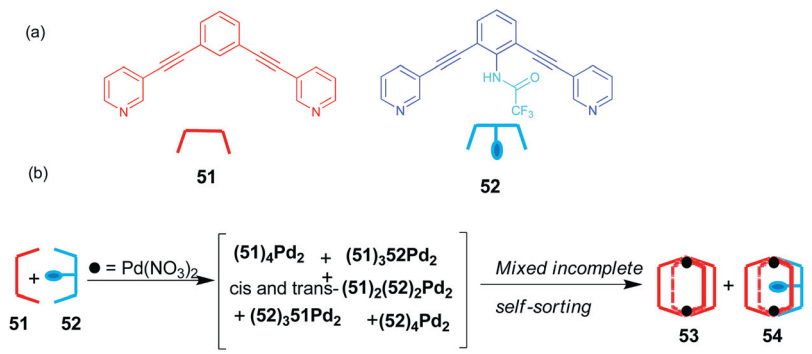



**Fig. 5** Homomeric $3^{2,2,2}$-fold(3) completive self-recognition of perylene bisimindes **44–46** by π–π stacking.[43]



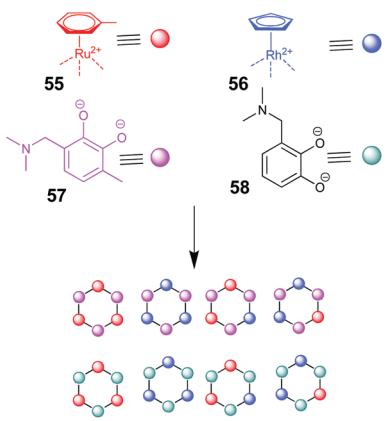

**Scheme 15** $2^{5,5}$-Fold(3) completive self-sorting of dinuclear helicates **49**, **50** from a 1 : 1 mixture of ligands **47** and **48** reacting with $Ti^{4+}$ in the presence of $Na_2CO_3$ as a base.[44]




**Scheme 16** (a) Chemical structures and cartoon representations of **51–52**. (b) Mixed $2^{6,6}$-fold(3) completive self-sorting of bis(pyridines) **51–52** through steric effects.[45]

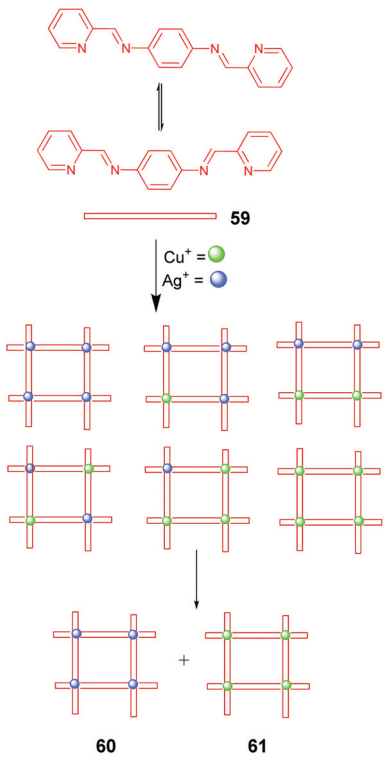
(Scheme 27). The homomeric selection is primarily driven by the specific structural, *i.e.* conformational preferences, of the corresponding building blocks in the macrocycles.[58] While our evaluation of $P_0$ may certainly be disputed, we derive a value of $P_0 = 7$ from constructing all possible dimerisation and trimerisation products that have the same entropic costs as **97** and **98**: *i.e.* **(95)**$_3$ (=**97**), **(95)**$_2$·**96**, **95**·**(96)**$_2$, **(96)**$_3$, **(95)**$_2$, **95**·**96**, **(96)**$_2$ (=**98**). As only two products are formed, this example is characterised by $M = 3.5$.

Lehn *et al.* described the self-assembly of the trinuclear triple helicate **101** from three oligobipyridine strands **100** and three octahedrally coordinated nickel(II) ions.[59] In a seminal contribution, Lehn *et al.*[1] combine some earlier results (see Scheme 5) in an elegant way with the formation of **101** to design the self-sorting of a double and triple helicate, *i.e.* **11a** and **101**, from a mixture of $Cu^+$ and $Ni^{2+}$ ions and the two different *tris*-bipyridine ligands **99–100** (Scheme 28). To our appreciation, out of 7 possible discrete homo- and heterohelicates as well as double and triple helicates, all designed along the maximum site occupancy rule[60] and lowest entropic costs, only one double and one triple helicate is found. This suggests $M = 3.5$.

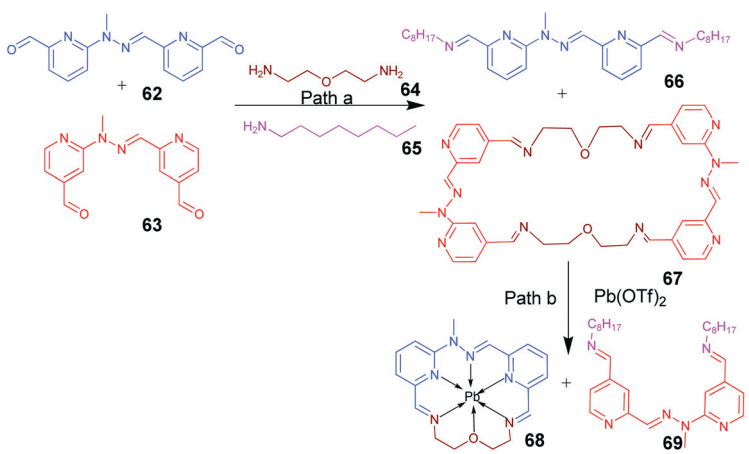



**Scheme 17** Mixed $8^{6,6,6,6,6,6,6,6}$-fold(4) completive self-assembly of the building blocks **55–58** leads to the formation of eight different macrocycles.[47]

The great majority of the examples of self-sorting systems described in literature are investigated in organic solvents. This fact stands in sharp contrast to biological systems, which operate in aqueous media. In recent years, Nitschke *et al.* successfully used subcomponent self-assembly techniques in water, marrying dynamic covalent (C=N) and metallo-supramolecular (M ← N) processes.[61–63] The chosen set always consists of "non-orthogonal" subcomponents (*e.g.* aldehydes and amines) that combine with each other in all possible ways to furnish a dynamic library




**Scheme 18** Formation of the homonuclear [2 × 2] grids **60**, **61** from a virtual dynamic library in a homomeric $2^{8,8}$-fold(3) completive self-sorting.[48]




**Scheme 19** Path a: A heteromeric $2^{3,4}$-fold(4) completive self-sorting process. Path b: An adaptation process induced by metal ion coordination leads to an imine swapping.[49a]

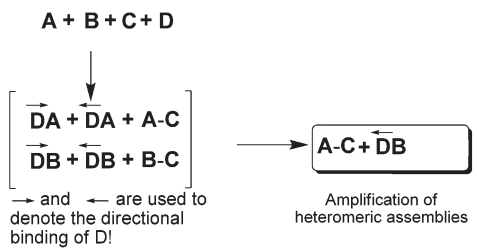
 

of several imines in the absence of metal ions but undergo a thermodynamic self-sorting process in the presence of the metal ion.

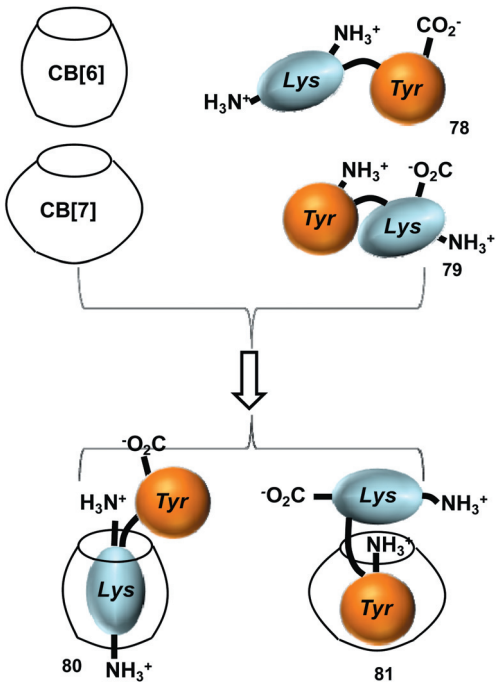
As a typical example, we depict herein the self-sorting of the two rather similar aldehydes **102** and **103** in the presence of amines **104–105** and $Cu^{+}/Fe^{3+}$ resulting in the formation of metal complexes **106** and **107**. The self-selection is guided by precise steric programming, optimal metal coordination and chelate effects. In absence of metal ions, a dynamic library of ~11 imines is present in equilibrium with all starting materials. However, addition of copper(I) tetrafluoroborate and iron(II) sulfate to the dynamic library generates **106** and **107** as the only products (Scheme 29). In our understanding, if one considers iron(II) to prefer a octahedral and copper(I) to favour a tetrahedral coordination geometry, then even in light of maximum site occupancy the system can end up with 7 possible metal complexes. Thus, self-sorting to 2 metal complexes suggests $M = 3.5$.[63]

**3.1.6.** ***M* > 5.0.** In 2004 Stang described the selective formation of the 2D polygons **25**, **110** and **111**, arising from 4,4′-dipyridyl (**22**) in the presence of various organoplatinum acceptors (Scheme 30).[64] In light of the "directional bonding model", self-sorting is clearly attributed to the different angles (0, 90 and 60 deg) among the acceptors **21**, **108** and **109**. To our understanding, aside of the found products **25** = $(\mathbf{21})_2(\mathbf{22})_2$, **110** = $(\mathbf{108})_4(\mathbf{22})_4$, and **111** = $(\mathbf{109})_3(\mathbf{22})_3$, there are many other options possible but do not form due to severe strain: further five [2 + 2], eight [3 + 3] and eleven [4 + 4] homo- and heterocomplexes. Thus, using the maximum site occupancy principle for [4 + 4], [3 + 3] and [2 + 2] complexes, altogether 27 discrete structures are possible. Because Stang and coworkers only observe 3 products, *M* will amount to 9.

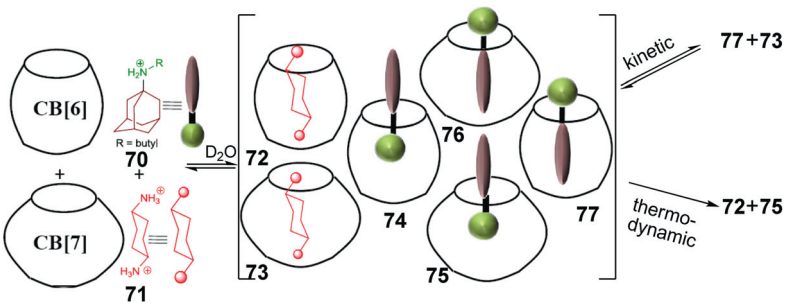
By using steric programming at the core of the ligand, Schalley and coworkers triggered self-sorting among eleven appropriately programmed tetra-urea calixarene derivatives.[65] All calixarenes **84** and **112–121** carrying four urea groups at their wide rim, are fixed in the cone conformation by four pentyl ether groups. Because they are based on the same scaffold they offer the same hydrogen-bond donor and acceptor pattern. Their only difference arises from the nature of the peripheral substituents (small or bulky) attached to the urea groups. Because all calixarenes are of analogous size, shape and hydrogen-bonding patterns, any bias is precluded on geometric or complementary grounds. In nonpolar aprotic solvents, such as benzene, cyclohexane or chloroform, the tetra-urea calixarenes **112–117** dimerise along a seam of interdigitating urea groups with altogether 16 hydrogen bonds. Statistically, the 11-component mixture **84** and **112–121** can combine to form 66 different dimers.[34] Experimentally, an equimolar mixture of all 11 components self-sorts into only six different dimers: **84**·**112**, **113**·**121**, **114**·**120**, **115**·**119**, **116**·**118**, and **117**·**117** (Scheme 31), suggesting $M = 11$. In reality, dimerisation is obstructed when adjacent urea residues




**Scheme 20** Amplification of two heteromeric assemblies from six possible combinations in a 4-component self-sorted system (model II).




**Scheme 22** Heteromeric $2^{2,2}$-fold(4) completive self-sorting of **48** and **49** in an aqueous solution containing **CB[6]** and **CB[7]**.[52]

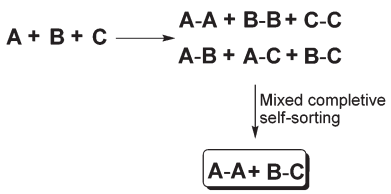



**Scheme 21** Kinetic *versus* thermodynamic heteromeric $2^{2,2}$-fold(4) completive self-sorting of **CB[6]** and **CB[7]** in presence of the doubly-functionalised guest **70** and guest **71**.[50b]

are covalently connected through one or more loops, as is the case of calixarenes **84** and **118–121**. Therefore, the latter are only capable to dimerise with a suitable partner (for example the "open chain" calixarenes **112–117**) equipped with nonbulky groups, whose urea substituents are small enough to penetrate the loops. Taking into account the above-mentioned constraints only 35 combinations are feasible.

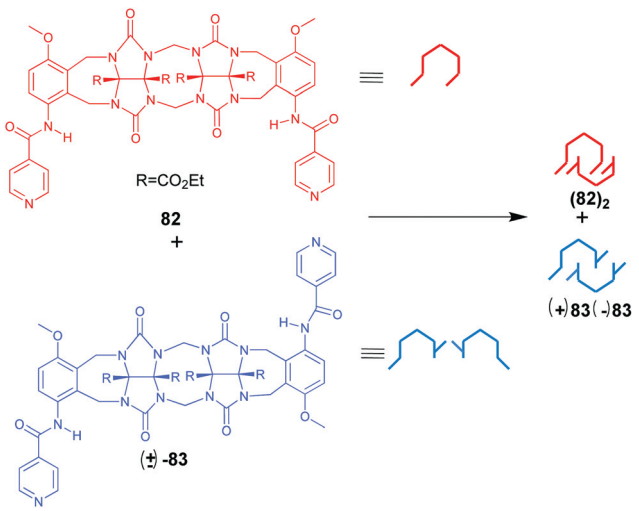
### 3.2. Degree of self-sorting in 1-fold completive systems

Nature is a grandmaster of selectivity. In biology, living organisms synthesise extremely pure intricate assemblies without any need for purification. Such delicate control is largely enabled by self-sorting, wherein the components of a metabolic pathway recognise each other within the complex cellular environment and sort out into a single assembly. The charm of self-sorting for a synthetic chemist is the possibility to self-assemble single species from a multicomponent library. In the following, we describe some selected abiological cases, in which self-sorting leads to a single product. Control in 1-fold completive sorting is operating mostly through geometry, appropriate steric constraints, *etc.* *Integrative self-sorting* as a way to 1-fold completive and integrative multicomponent systems is described in the next section.

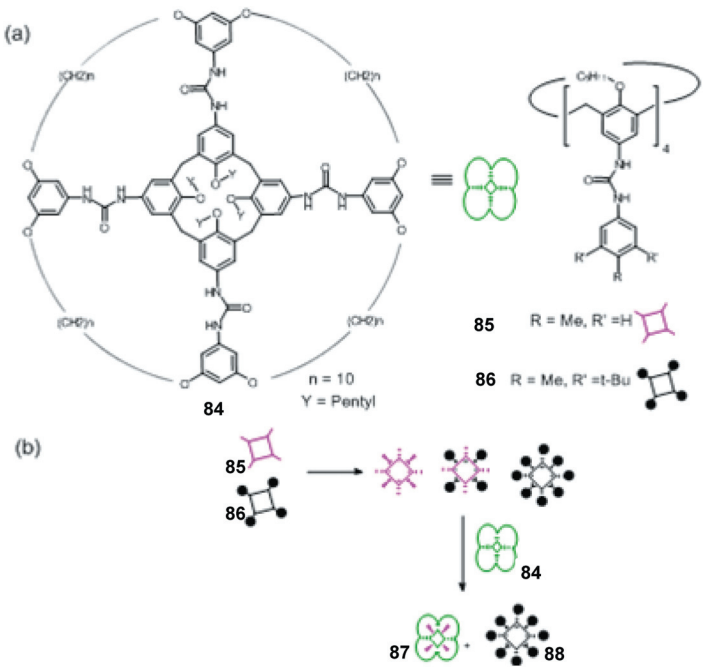



**Scheme 23** Amplification of two assemblies from six possible combinations in a 3-component self-sorted system (model III).

**3.2.1.** ***M* = 2.0.** Stang *et al.* describe the exclusive formation of discrete supramolecular rhomboids from the 90° monoplatinum acceptor **108** and the unsymmetrical bidentate ligands **122a–c**, the latter bearing two different binding sites: a pyridine and a carboxylate site. The ditopic ligands can potentially form two isomeric products by self-assembly: one with the bidentate ligands oriented in the same direction and another with the ligands oriented in opposite directions. Experimentally, only isomers **123a–c** with alternating ligand orientations are formed

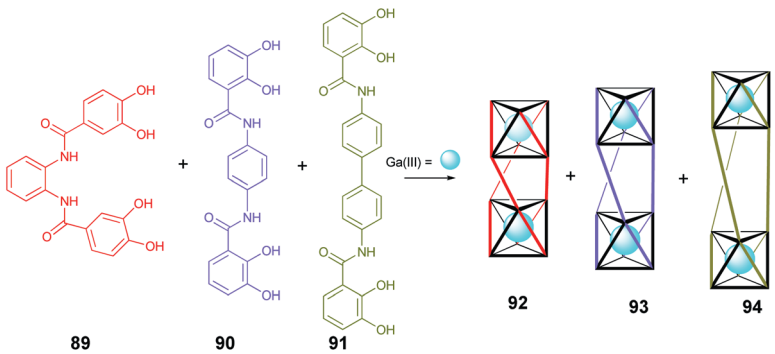



**Scheme 24** A mixed $2^{2,2}$-fold(3) completive self-sorting process of molecular clips (**82** and (±)-**83**).[53a]

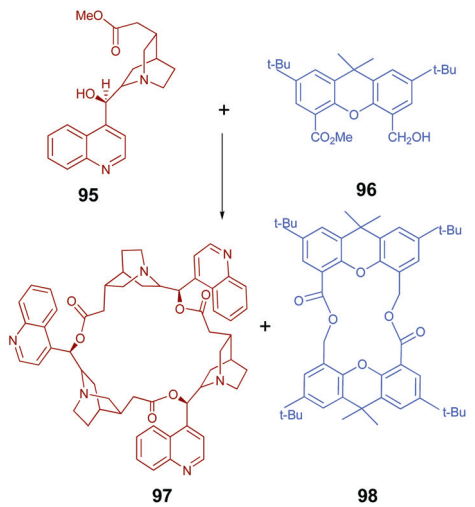



**Scheme 25** (a) Chemical structures and cartoon representations of **84–86**. (b) A mixed $2^{2,2}$-fold(3) completive self-sorting of a mixture of **85** and **86** by **84**.[56]




**Scheme 26** Schematic representation of the self-recognition of ligands **89–91** in triple helicates **92–94**. The coloured rods depict **92–94** and the spheres represents gallium(III) ions. This is a $3^{5,5,5}$-fold(4) completive self-sorting.[57]




**Scheme 27** Homomeric $2^{2,3}$-fold(2) completive self-recognition of cinchonidine hydroxy esters **95** and xanthene hydroxy esters **96** to afford macrocycles **97** and **98**, respectively.[58]

(Scheme 32), suggesting $M = 2$. The self-assembly is primarily driven by charge separation as well as minimisation of ring strain. The more favourable geometry can be attained if the ligands are arranged in an alternating fashion thus creating identical coordination environments at each Pt center.[66] A similar concept was employed in the construction of heterobimetallic triangles.[67]

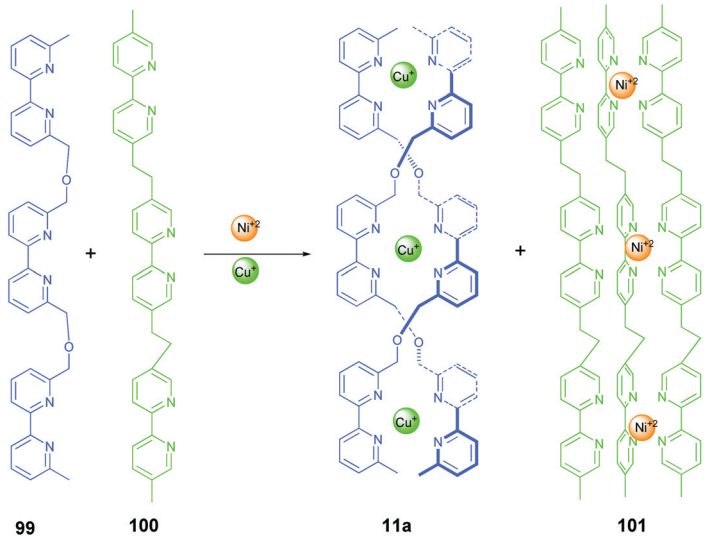
**3.2.2. *M* = 3.0. *Model IV***: For a methodical presentation, the ensuing systems may be portrayed according to the general model they follow (model IV). As described in Scheme 33, A and B are two related molecules that may form homo ($A_2$ and $B_2$) or heterodimers (AB). In a 1-fold completive self-sorting the heteromeric assembly AB is solely observed. Hence, the *M* value for such system is 3.

As an example for model IV, Rebek *et al.* illustrate the selective formation of **126**, a hydrogen-bonded heterodimeric capsule, from two very different monomeric units, which are both capable of forming their own homomeric capsules through hydrogen bonding (Scheme 34). The alternative homomeric capsules have either strong hydrogen bond donors (imide N-H's in $(\mathbf{124})_2$) or strong hydrogen bond acceptors (ureido carbonyls in $(\mathbf{125})_2$), but these complement each other optimally in the heterodimer **126** = **124**·**125**.[68]

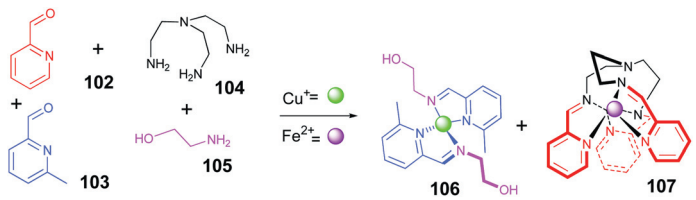
In a related contribution, Ballester and coworkers describe the solvent dependent self-sorting during the self-assembly of tetra-urea tolyl calix[4]arene **127** and tetra-urea benzyl calix[4]pyrrole **128**.[69] Dimerisation of **128** is induced by the pairwise encapsulation of trimethylamine-*N*-oxide (**129**) in both $CD_2Cl_2$ and $CDCl_3$. Importantly, $CDCl_3$ is also a competitive guest for the homodimeric capsule $(\mathbf{127})_2$ as it complements nicely its cavity size and lipophilicity, whereas $CD_2Cl_2$ disrupts the formation of $(\mathbf{127})_2$. Therefore, the authors investigated the self-sorting of the two tetra-ureas **127**, **128** and *N*-oxide **129** in $CDCl_3$. The corresponding $^1$H-NMR signals reveal that the percentages of assemblies are approximately: $CDCl_3$@$(\mathbf{127})_2$, **129**@**127**·**128**, $(\mathbf{129})_2$@$(\mathbf{128})_2$ as 12%, 76%, and 12% respectively. When the same reaction was performed in $CD_2Cl_2$, though, the heterocapsule **129**@**127**·**128** formed exclusively (Scheme 35). The obtained results demonstrate that the solvent plays an important role in the self-sorting of this system. Actually, the exclusive assembly of the hybrid capsule **127**·**128** in $CD_2Cl_2$, is not driven by filling its cavity with solvent molecules but by the maximisation of the available intermolecular interactions as a result of the disruption of the capsule $(\mathbf{127})_2$ in this solvent.

Based on the interplay of hydrogen bonding and dynamic covalent disulfide formation, Gong *et al.* realised the sequence-specific formation of molecular strands, containing multiple different complementary hydrogen bond donors and acceptors that organised into covalently cross-linked duplexes in aqueous media.[70] When a 1 : 1 mixture of **130** and **131** was treated with iodine in aqueous media, the MALDI-TOF results indicated that the heteroduplex **132** appeared as the overwhelmingly major product (Scheme 36).

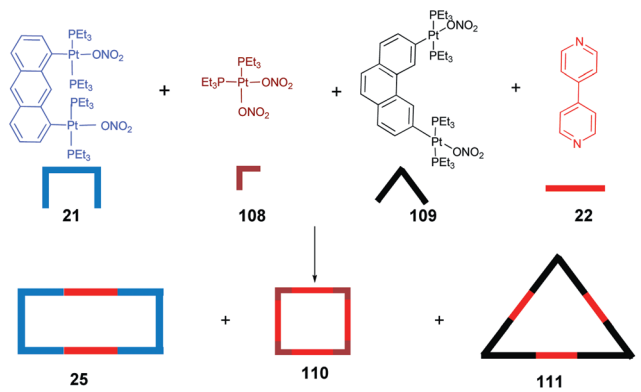
Kobayashi *et al.*[71] utilised steric benefits to favour the hetero-cavitand cage **136** (Scheme 37). A 1 : 1 : 4 mixture of pyridyl cavitand **133**, cyanophenyl cavitand **134**, and **135** = [Pd(dppp)]-$(OTf)_2$ (dppp = bis-1,3-(diphenylphosphanyl)-propane) in $CDCl_3$ at room temperature instantaneously self-sorted into **136** as shown in Scheme 37. The rotational motions of the pyridyl and cyanophenyl groups in **133** and **134** are highly restricted because both are placed between sterically demanding acetal moieties on the cavitand scaffold. Furthermore, the homoleptic cage of **134** is less stable than the hetero-cage **136** as the latter profits from the better donor ability of the pyridyl group.[71b]




**Scheme 28** Self-recognition in the $2^{5,6}$-fold(4) completive self-assembly of the double helicate **11a** and the triple helicate **101** from a mixture of **99** and **100** and of $Cu^{+}$ and $Ni^{2+}$ ions.[1]




**Scheme 29** $2^{5,5}$-Fold(6) completive self-sorting of metal complexes **106** and **107**.[63]



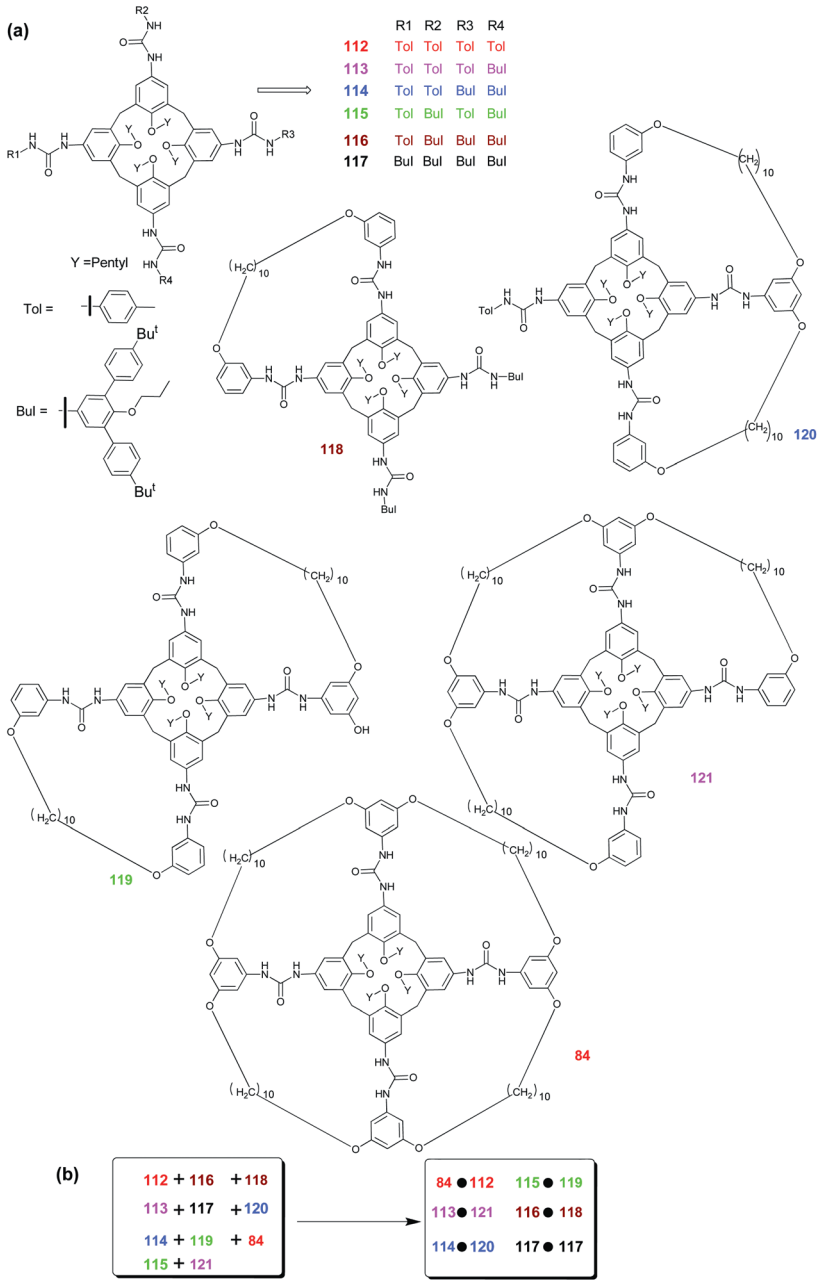

**Scheme 30** $3^{4,6,8}$-Fold(4) completive self-sorting during the formation of 2D polygons **25**, **110** and **111**.[64]

Hupp and coworkers introduced bulky ligands at the tin in tin(II)porphyrins **137–138** to prepare a highly ordered, rigid porphyrin box.[72] As illustrated in Scheme 38b, the two pyridine-derivatised porphyrin dimers **137**, when combined with four porphyrins trimers **139**, generate the symmetrical porphyrin box **140**. The steric crowdedness at the tin(II)porphyrin units forces **137** to coordinate selectively *via* the pyridine ligands to the first and third zinc(II)porphyrin terminals of **139** in $(\mathbf{137})_2(\mathbf{139})_4$ = **140**, leaving the central zinc site unoccupied. In a competition experiment, a mixture of **137–139** yielded only the hetero-loaded assembly **141** (Scheme 38c). PFG-NMR and SAXS measurements confirm that the solution contains a single assembly rather than a mixture of homo-loaded assemblies. The observed 1-fold completive self-sorting is presumably driven by elimination of the strain associated with the twisted assembly, $(\mathbf{138})_2(\mathbf{139})_4$. In a related paper, the authors furthermore describe the encapsulation of a slim manganese(II)porphyrin dimer into the large cavity of **140**.[73]

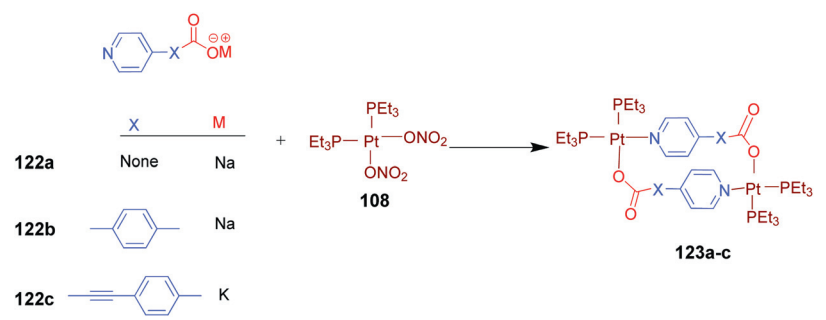
**3.2.3.** ***M* = 4.0.** Over the last decade, several groups[74,75] have successfully employed steric programming in the ligand core to influence the outcome of sorting processes. As a nice example, Stang and coworkers describe the selective formation of **144**, utilising the unsymmetric bipyridine **142** in combination with a 90° monoplatinum acceptor as corner unit **143** (Scheme 39).[74] Due to the lack of constitutional symmetry in **142**, it may theoretically self-assemble with **143** into four squares, but experimentally only the isomer **144** was found, suggesting $M = 4$. Most likely, the three other isomers were not able to compete due to the presence of at least one unfavourable steric interaction arising from two bulky pyridine terminals at a single metal center.

Originally, the same principle has been utilised by Fujita *et al.* in the selective multicomponent assembly of two- and three-dimensional polynuclear palladium(II) complexes, such as squares, as well as rectangular and trigonal prisms.[75a] Within the cavities of the pillared coordination prisms, the electrostatic

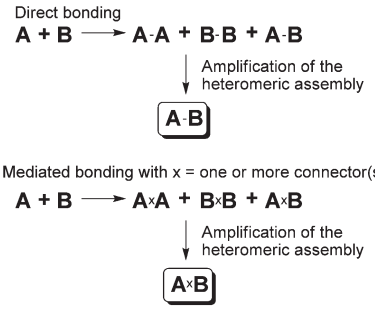
 



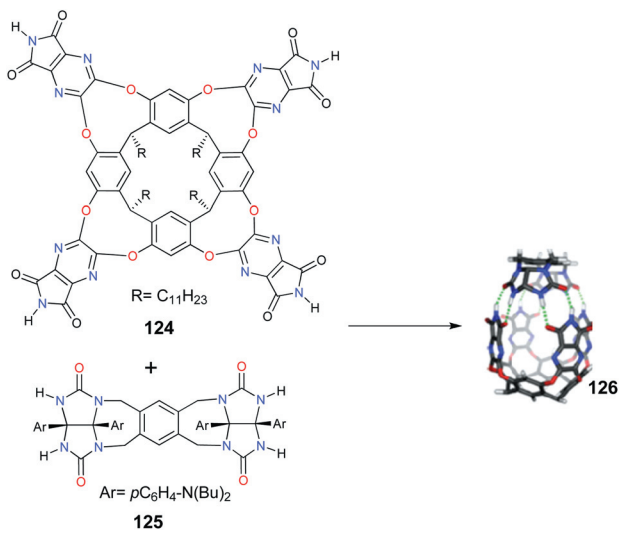

**Scheme 31** (a) Chemical structures of calixarenes **84** and **112–121** and (b) their self-sorting into six assemblies.[65]



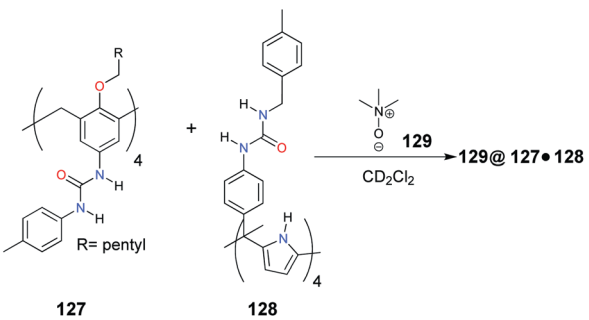

**Scheme 32** $1^2$-Fold(2) completive self-assembly of ligands **122a–c** with the 90° linker **108**.[66]




**Scheme 33** The heteromeric assembly AB formed in a 1-fold completive self-sorting process (model IV).



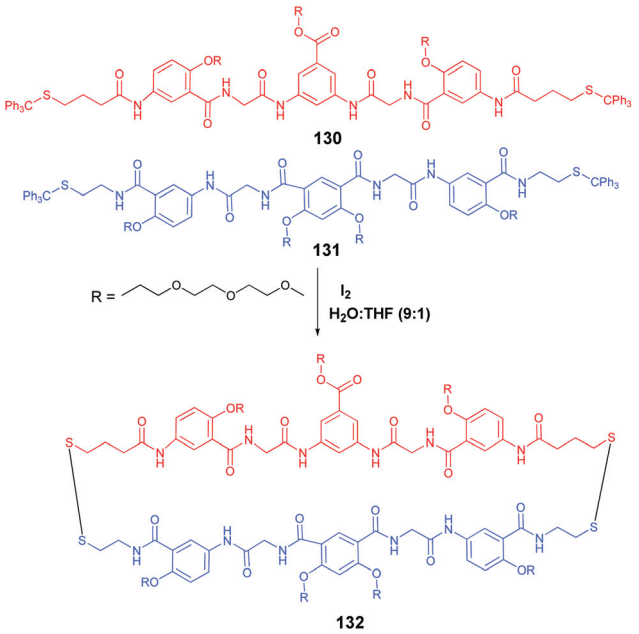

**Scheme 34** Heterodimeric capsule **124** formed from **125** and **126** in a $1^2$-fold(2) completive process.[68] **126** reprinted with permission from ref. 68. Copyright 2009 National Academy of Sciences.



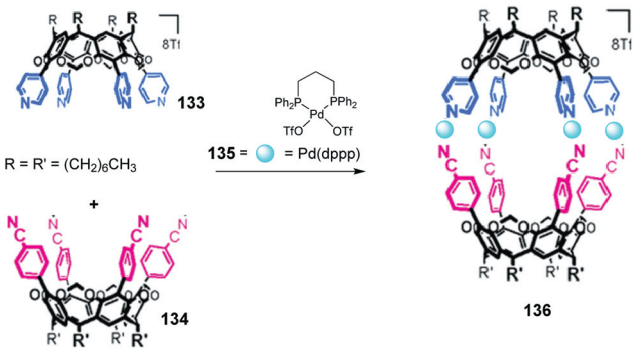

**Scheme 35** $1^3$-Fold(3) completive self-sorting toward the trimethylamine-*N*-oxide encapsulated heterocapsule **129@127·128**.[69]

interactions between electron-rich and electron-poor aromatics guests were further exploited for the formation of discrete hetero stacks.[76]

The proper choice of an unnatural amino acid residue, *e.g.* cyclohexylalanine, at the central 16th position of the amino acid sequence in a heptad repeat (*abcdefg*) triggers a heteromeric peptide selection during coiled-coil self-assembly formation, as

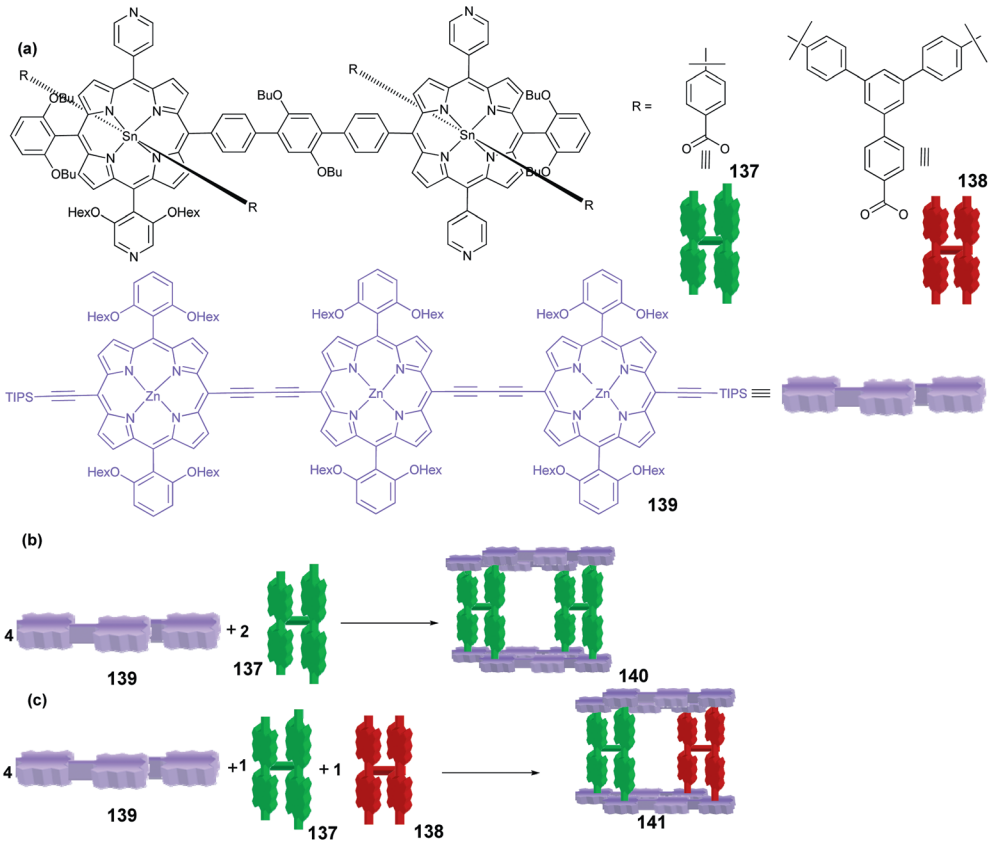



**Scheme 36** Synthesis of the heteroduplex **132** in a $1^2$-fold(2) completive process.[70]

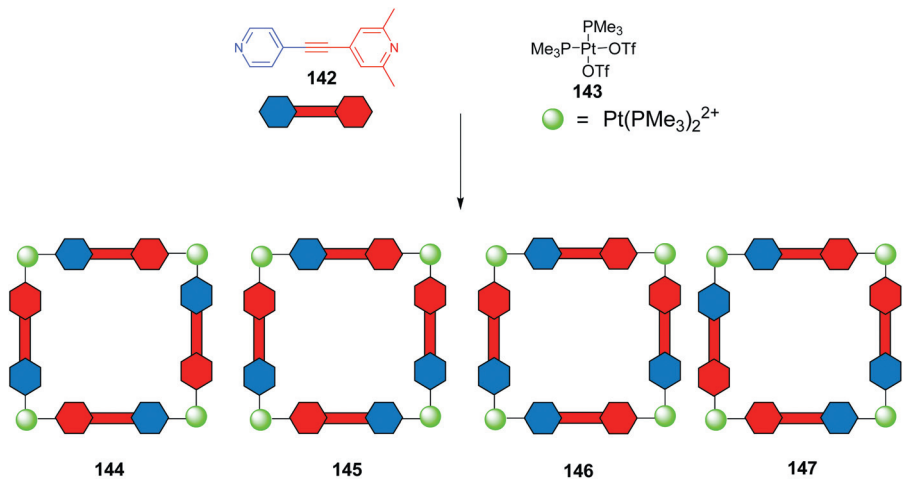



**Scheme 37** Sterically and electronically controlled hetero-assembly from a $1^6$-fold(3) completive process.[71b]

described by Kennan *et al.*[77] The amino acid sequence in both peptides **148** and **149** only differs in the 16th position, with **148** accommodating alanine and **149** cyclohexylalanine (Fig. 6). Notably, from a mixture of both peptides the heterotrimer **150** = $(\mathbf{148})_2$·**149** formed selectively. Considering 4 likely homo- and hetero coiled-coil self-assemblies ($(\mathbf{149})_3$, $(\mathbf{148})_2$·**149** (=**150**), **148**·$(\mathbf{149})_2$, $(\mathbf{148})_3$), the exclusive formation of **150** indicates $M$ = 4. The composition of the heterotripeptide assembly is rationalised mostly due to steric matching of the hydrophobic core side chains. Moreover, the geometry of **150** facilitates the packing of the small side chain positioned at one core position with the same residue on the opposing strands. Therefore, the resulting pocket is additionally stabilised by hydrophobic packing interactions *via* the introduction of a complementary peptide, substituted with a larger side chain, *i.e.* **149**. The proper choice of steric bulk is crucial. Indeed, substitution with naphthylalanine or cyclopropylalanine at the central a position failed to furnish similar heterotrimers with the alanine containing




**Scheme 38** (a) Chemical structures and cartoon representations of **137–139**. (b) Synthesis of **140**. (c) The heteromeric assembly **141** forms in a $1^6$-fold(3) completive self-sorting process.[72]




**Scheme 39** Steric interactions lead to the preferential formation of square **144** *via* a $1^8$-fold(2) completive process.[74]

peptide. Actually, in these systems core layers with steric voids are more damaging than those with steric repulsions. Thus, both the cyclohexylalanine and naphthylalanine peptides form complexes even more stable than **150**. The specificity for the 2 : 1 complex appears to arise mainly from poor interactions in the homotrimer of the alanine peptide. Later, the same group used the complementarity of cyclohexylalanine as a binding partner for the alanine peptide in the construction of a 1 : 1 : 1 heterotrimer.[78]

**3.2.4.** ***M* = 8.0.** Recently, Schalley and coworkers reported on the thermodynamically controlled self-sorting of supramolecular heterobimetallic and homometallic macrocycles.[79] Ligand **151** has two different binding sites, a bipyridine and two

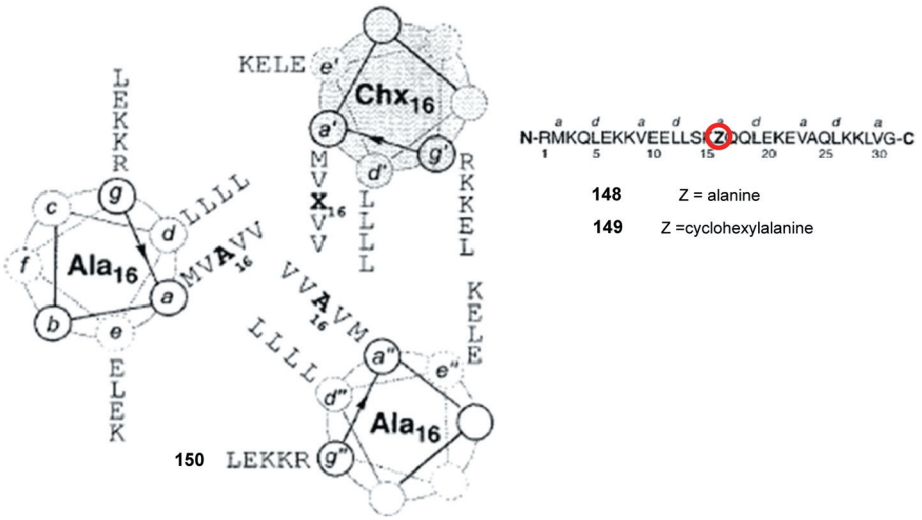



**Fig. 6** Schematic representation of **148–150**.[77] Reproduced with permission from ref. 77. Copyright 2001 American Chemical Society.

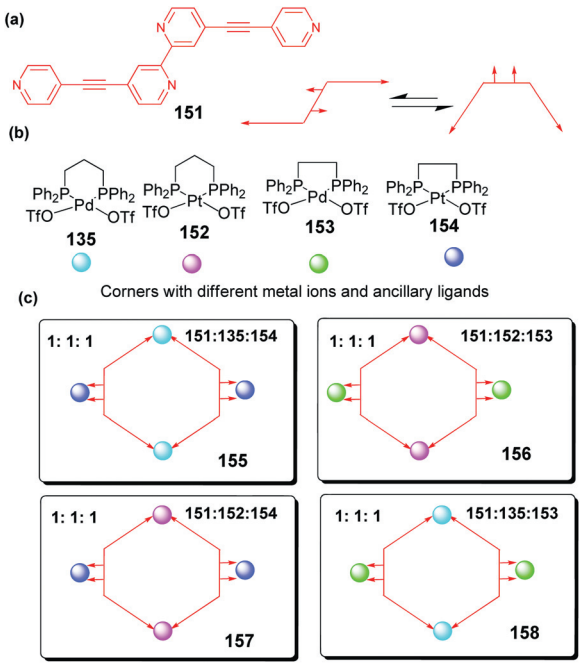



**Fig. 7** (a) Tetradentate ligand **151**; (b) metal–ligand units **135**, **152–154**, and (c) the four self-sorted metallo-supramolecular macrocycles **155–158**.[79]

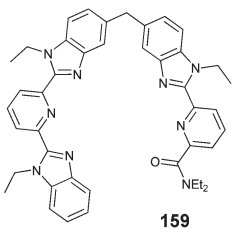



**Fig. 8** Chemical structure of **159**.[80]

individual pyridines. Interestingly, metal corners (Pd, Pt) with dppe [bis-1,2-(diphenylphosphanyl)-ethane] produce self-sorting in presence of dppp containing metal corners. For example, in a 1 : 1 : 1 mixture of **151**, **135** along with **154**, the selective binding of **135** at the pyridine sites triggers the dppe containing ligand to coordinate exclusively at the bipyridine site of **155** (Fig. 7). In light of maximum site occupancy, a mixture of both dppp and dppe appended metal corner in 1 : 1 ratio seems to generate 8 possible macrocycles, thus *M* should amount to 8. Apparently, the presence of an additional methylene group produces a huge effect in the orientation of the phenyl groups at the phosphorus surrounding the metal ions. Most likely, it is this steric effect that leads the dppp metal corners to coordinate exclusively at the pyridine site.

Based on the difference in ionic radii, Bunzil and coworkers portray the selective incorporation of two different lanthanide ions, *i.e.* $La^{3+}$ and $Lu^{3+}$, into a HHH-arranged (H stands for head) triple-stranded helicate generated from ligand **159**.[80] If HHH and HHT arrangements are considered along with 4 possible homo- and heterometallic triple-stranded helicates ($P_0$ = 8), then the formation of only one mixed metal helicate indicates $M$ = 8. Apparently, the benzimidazole–pyridine–benzimidazole coordination unit of **159** binds preferentially to the heavier $Ln^{3+}$ ion, while the benzimidazole–pyridine–carboxamide terminals of the triple helicate furnish a coordination cavity for the lighter $Ln^{3+}$ ion. The interstrand π–π and C–H···π interactions favour the HHH-helicate arrangement over the alternative HHT arrangement (T stands for tail). Actually, the size-discriminating effect along the lanthanide(III) series becomes effective only upon complexation of the third ligand strand, because the size of the self-assembled cavity becomes the key factor in the recognition process. In the case of HHT-bimetallic helicates, the three tridentate coordination units would not allow for such pronounced size discrimination (Fig. 8).

In a related approach, Hahn and coworkers utilise the different binding preferences of benzene–*o*-dithiol *vs.* catechol in the mixed ligand **160**. In the heterobimetallic triple-stranded helicate **162**, the three catecholates selectively choose to coordinate to



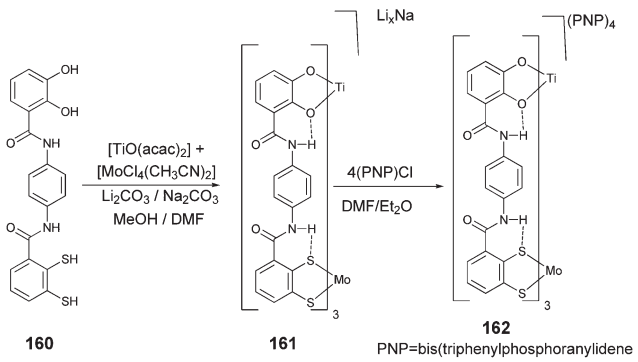
 

titanium(IV), while the benzene-*o*-dithiolate (bdt) prefers the molybdenum(IV) as a metal centre (Scheme 40). The observed selectivity is best explained by the planar geometry of the {Mo-(bdt)$_3$} moiety favouring formation of strong NH···S hydrogen bonds. In contrast, similar homodinuclear titanium(IV) helicates containing a {Ti(bdt)$_3$} unit are severely bent about the S···S vector, effectively preventing the formation of NH···S hydrogen bonds.[81] In principle, neglecting the chemoselectivity of catecholate and bdt, one may expect 8 possible homo- and heterometallic triple-stranded helicates. The formation of **162** as a single product indicates $M = 8$.

## 4. Integrative self-sorting

Recently, Schalley and Schmittel independently utilised *integrative self-sorting* to erect intricate multicomponent architectures. At the heart of such approach (Fig. 9), one needs first to optimise nonintegrative *x*-fold completive self-sorting of all required mononuclear cornerstones in a suitable library and second to implement these motifs in multiligand building blocks in a way that integrative and 1-fold completive self-sorting may be achieved. For this purpose, at least one component must now be designed in a way that it becomes part of two (or more) hitherto independent self-sorting protocols. Thus, two or more binding sites must be integrated into at least one of the components in order to transport the building algorithm and positional information for the final architecture. To our perception, the *integrative* strategy based on orthogonal binding motifs incorporated in the key building blocks represents a true advancement in the art of supramolecular synthesis, a field that has long been dominated by highly symmetric structures.

In 2008, Schalley *et al.* explored the sorting properties of two very similar macrocyclic crown ethers, *i.e.* **163–164**, in presence of the dialkylammonium ions **165–166**.[20] The observed sorting, as described in Scheme 41, is most likely guided by the distinct size of the various groups attached to the dialkylammonium unit: anthracenyl > phenyl > hydroxyalkyl. While secondary




**Scheme 40** Syntheses of the heterobimetallic helicates **162** = (PNP)$_4$[TiMo(**150**)$_3$] in a $1^5$-fold(3) completive process.[81]

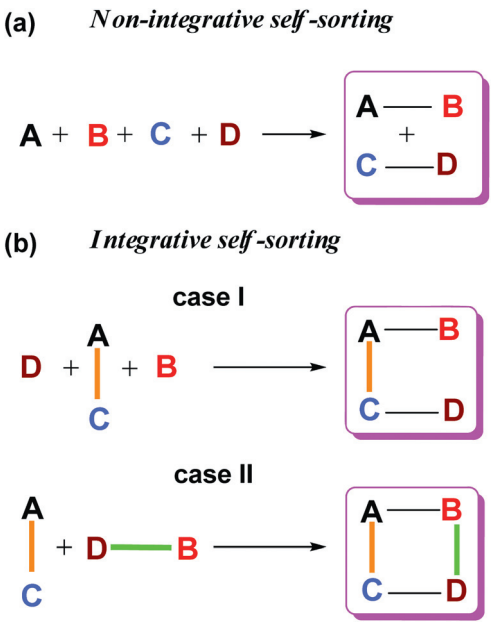
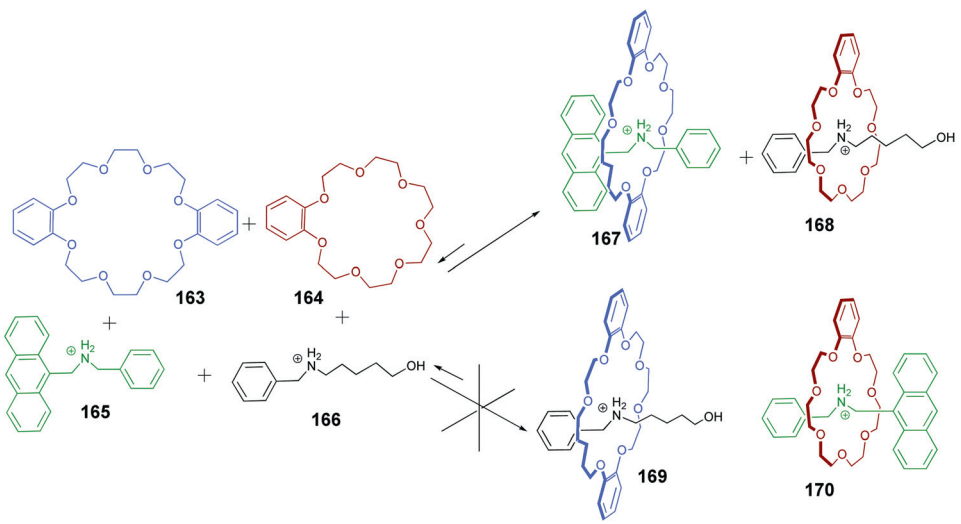
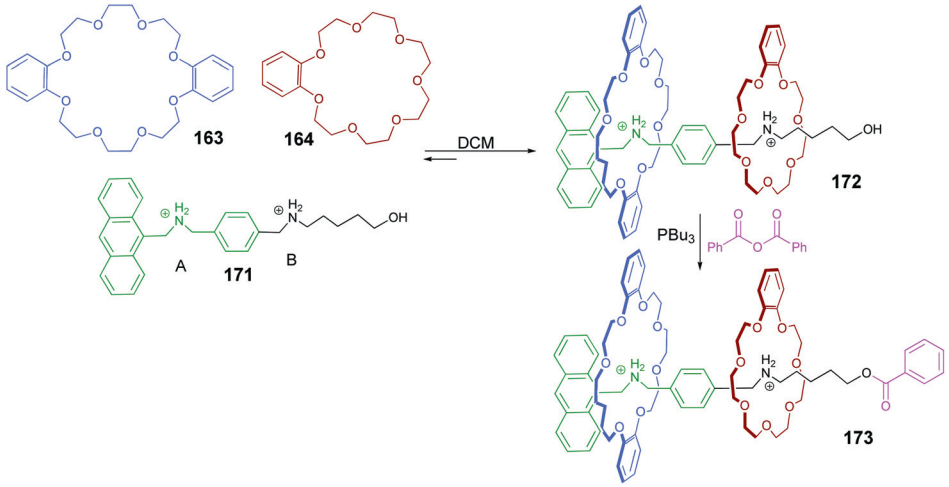



**Fig. 9** Illustration of the *integrative self-sorting* approach.




**Scheme 41** A $2^{2,2}$-fold(4) completive self-sorting system constructed from an equimolar mixture of **163–166**.[20]




**Scheme 42** Sequence-specific formation of hetero[3]pseudorotaxane **172** based on a $1^3$-fold(3) completive integrative self-sorting process and the follow-up synthesis of the "cascade-stoppered" hetero[3]rotaxane **173**.[20]



dialkylammonium ions are able to thread through the cavity of benzo-21-crown-7 (**164**) forming a pseudorotaxane, phenyl groups may effectively act as stoppers in these systems to trap **164** on the axle.[82] Indeed, the smaller macrocyclic crown ether **164** did not pass over a phenyl group under the conditions of the experiment, suggesting that formation of pseudorotaxane **170** is kinetically impeded. In light of the maximum site occupancy rule, the system thus can only evolve in one direction, *i.e.*, toward forming **167** and **168**. Out of four options, only two are realised, indicating $M = 2$. The corresponding association constants are additionally supportive of the observed self-sorting, as both **167** ($496 \pm 18\ M^{-1}$) and **168** ($615 \pm 36\ M^{-1}$) are thermodynamically more stable than **169** ($155 \pm 8\ M^{-1}$).

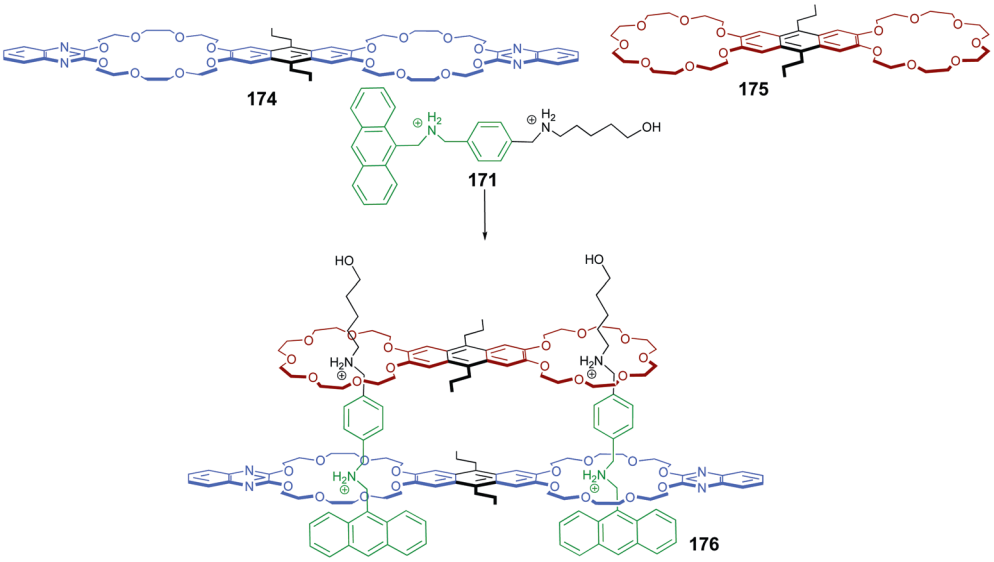
From this useful 4-component self-sorting system (Scheme 41), numerous integrative self-sorting systems may now be created by suitable *molecular programming*. For a first demonstration, Schalley *et al.* merged the two dialkylammonium motifs **165** and **166** in the key building block **171**.[20] As described in Scheme 42, axle **171** is equipped with the two binding sites A and B. Site A represents the binding site in **165** and site B that in **166**. Interestingly, upon mixing **171** and **164** (1 : 1), macrocycle **164** binds exclusively at site B of **171**, indicating that a kinetic barrier still prevents its coordination at site A in **171**. In contrast, the reaction of macrocycle **163** with **171** does not show any selectivity for sites A or B, as indicated by $^1$H-NMR spectroscopy. It is thus a combination of maximum site occupancy and of kinetically preventing one combination that leads, alike the 4-component self-sorting system in Scheme 41, to the sequence-specific hetero[3]-pseudorotaxane **172** (Scheme 42). **172** is the major product from equimolar amounts of the heterodivalent axle **171** and the two crown ethers **163** and **164**, as confirmed by $^1$H-NMR and ESI-FTICR mass spectra. This integrative 1-fold completive self-sorting shows $M = 4$. Finally, the received complex was subjected to an esterification reaction with benzoic anhydride to yield a "cascade stoppered" **173**.

Recently, the Schalley group utilised the same guiding idea to fabricate highly sophisticated rotaxane-based 3-component rectangles. For example, the heterodivalent axle **171** and homodivalent hosts **174–175** self-sort into the [4]pseudorotaxane **176** (Scheme 43) at $M = 4$.[15] In all of these examples, ESI-FTICR mass spectrometry was demonstrated to be the key tool to monitor intermediates, falsely assembled structures as well as the final thermodynamic products of the self-sorted mixtures.[83]

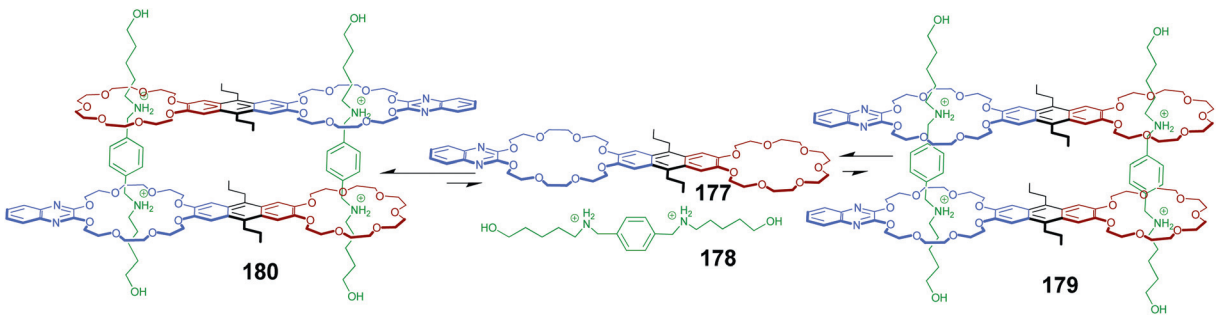
Very recently, preferential formation of [4]pseudorotaxane **180** over **179** (>49 : 1) was observed indicating a predilection for the antiparallel over the parallel alignment of crown ether building blocks **177** (Scheme 44).[84] By design, the hydroxypentyl-substituted ammonium units of **178** are able to coordinate to both crown ethers to form [4]pseudorotaxanes, and in principle the two constitutional isomers **179** and **180** may form. No significant stability difference is to be expected and both isomers may coexist in solution. However, the $^1$H-NMR spectrum of the 1 : 1 mixture of **177** and **178** suggests the predominance of the antiparallel isomer **180**. The selective formation of **180** may be due to the favourable alignment of dipoles in the antiparallel arrangement of the two crown ether dimers. Additionally, the two quinoxaline ring systems would be close to each other in **179**, leading to an unfavourable arrangement of local dipoles. In addition, differential geometric strain accounts at least for part of the *ca.* 9.6 kJ $mol^{-1}$ energy difference between the isomers. By comparing the statistical and experimental abundance of **180**, we conclude that the antiparallel assembly is increased by ≅96% compared to the theoretical statistical distribution.

In 2011, Schalley *et al.* elaborated an *integrative self-sorting* protocol with water soluble cucurbituril heteropseudorotaxanes.[85] Both, **CB[7]** and **CB[8]** are known to bind methyl viologen (**MV**) with almost identical binding affinities ($K = 2.0 \times 10^5\ M^{-1}$ and $1.1 \times 10^5\ M^{-1}$, respectively), which seems to preclude any self-sorting on thermodynamic grounds. In order to set up self-sorting, they explored the larger cavity size of **CB[8]** for accommodating a second electron-rich co-guest. Using 2,6-dihydroxynaphthalene (**DHN**) as a co-guest, the ternary complex **182** was formed with a charge-transfer guest pair inside the **CB[8]** cavity ($K_1 \times K_2 - 10^9\ M^{-2}$) (Scheme 45). In contrast, the cavity of **CB[7]** is not large enough to accommodate both guests.

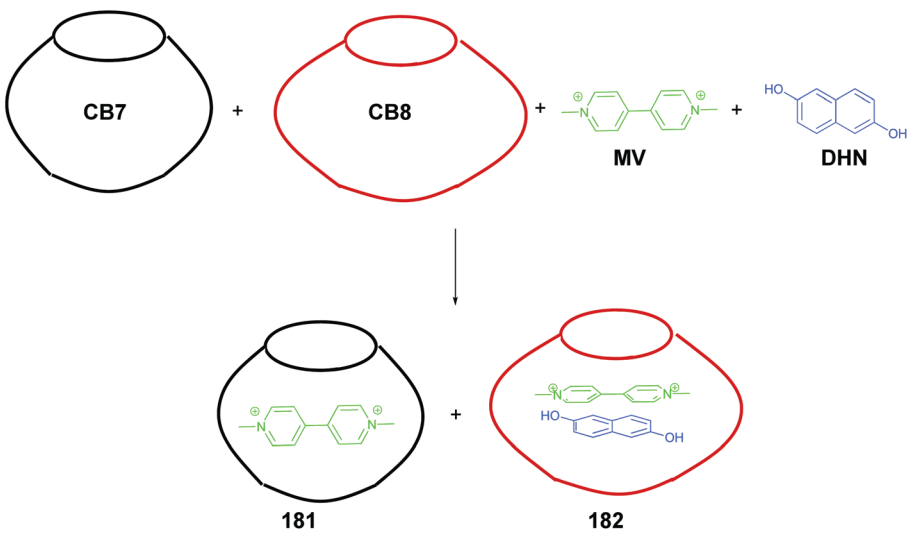
 




**Scheme 43** $1^4$-Fold(3) completive self-sorting toward [4]pseudorotaxane **176**.[83]




**Scheme 44** Self-assembly of **179–180** from **177** and **178**. The equilibrium of **179** and **180** is shifted far to the antiparallel crown ether arrangement (>49 : 1).[84]

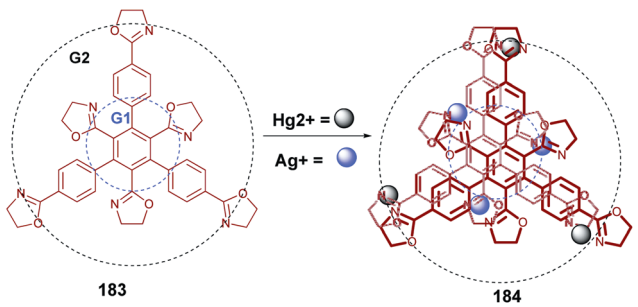



**Scheme 45** A $2^{2,3}$-Fold(4) completive self-sorting system constructed from an equimolar mixture of **CB[7]**, **CB[8]**, **MV** and **DHN**.[85]

Finally, implementation of **CB[7]** and **CB[8]** into hetero[*n*]pseudorotaxanes (*n* = 2 to 5) was achieved by an *integrative self-sorting* strategy.

At a time when the expression integrative self-sorting was still unknown, Shionoya and coworkers reported on the controlled arrangement of $Ag^+$ and $Hg^{2+}$ ions in the sandwich structure **184**

with its two different binding sites (G1 and G2). In the thermodynamically stable sandwich-shaped assembly three $Ag^{+}$ ions are arranged in the G1 and the three $Hg^{2+}$ ions in the G2 sites (Scheme 46).[86] The site-selective metal arrangement was best explained by minimisation of Coulombic repulsion between adjacent positively charged metal centers. While the silver ions can accommodate in both positions, $Hg^{2+}$ can only be placed in G2 positions due to the strong electrostatic repulsion. Thus, in an equimolar mixture of $Hg^{2+}$ and $Ag^{+}$, the strongly coordinating metal ion ($Hg^{2+}$) prefers to locate in G2, leaving no other option for the silver ion than the G1 sites. Simply guided by maximum site occupancy for a 1-fold completive system, one would expect 16 sandwich structures with two different metal ions arrangements in the G1 *vs.* G2 locations. The clean formation of **184** thus suggests $M = 16$. Later on, the same methodology was utilised for the electrostatically controlled arrangement of two

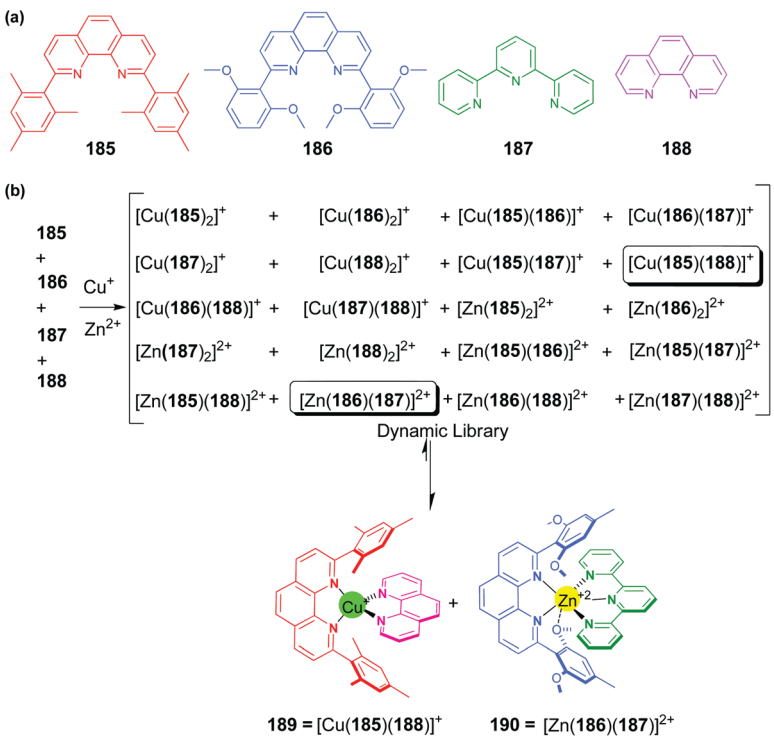



**Scheme 46** Coulomb repulsion-controlled arrangement of $Ag^{+}$ and $Hg^{2+}$ ions between two disk-shaped hexa-monodentate ligands **183** represents a $1^{8}$-fold(3) completive self-sorting.[86]

different metal ions on a cyclic array with six chemically equivalent oxazoline binding sites.[87]

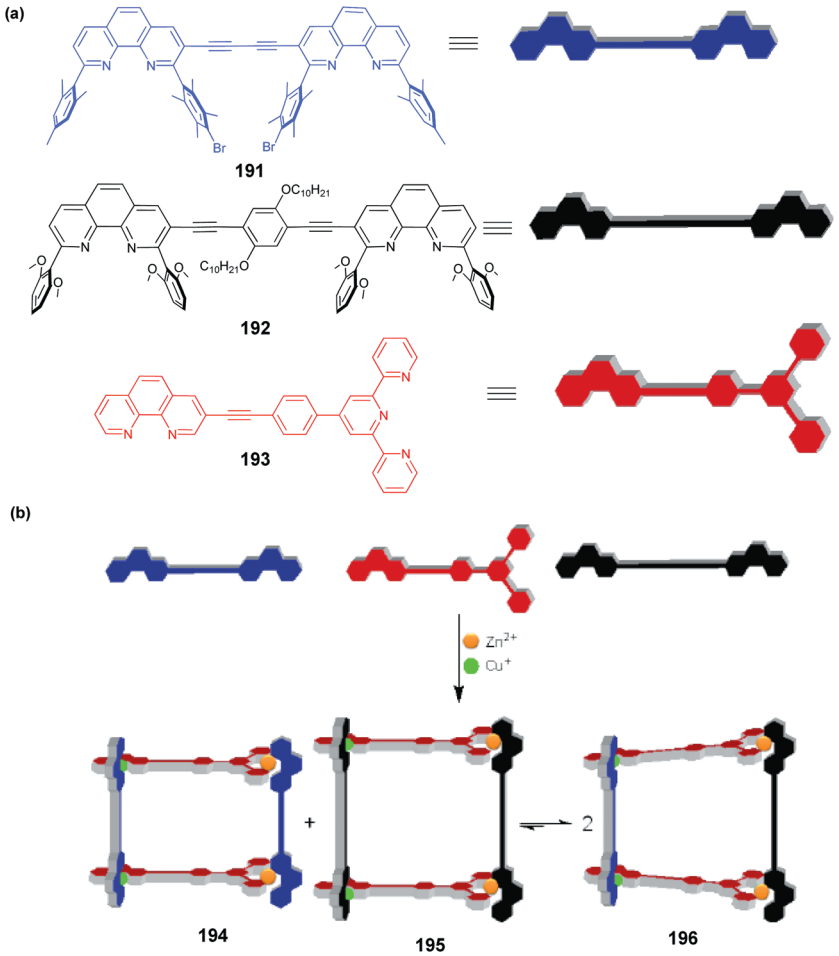
In 2009, Schmittel & Mahata utilised the preferred coordination geometries of zinc(II) and copper(I) ions to achieve a six-component $A^{1}A^{2}D^{1}D^{2}D^{3}D^{4}$ self-sorting system, with A representing an acceptor and D a donor (Scheme 47).[18] As described in Scheme 47, out of 20 possible combinations, the $2^{3,3}$-fold(6) completive self-sorting library ended up with only two metal complexes: **189** = [Cu(**185**)(**188**)]$^{+}$ and **190** = [Zn(**186**)(**187**)]$^{2+}$, indicating ($M = 10$). This high degree of self-sorting has been credited to precise tuning of steric and electronic effects, π–π interactions, and metal-ion coordination specifics. To our believe, the additional ion-dipole interaction(s) present in **211** provide(s) a suitable *pseudo*-octahedral geometry to the $Zn^{2+}$ ions thus guiding the observed sorting phenomenon in light of HETPHEN[88a] and HETTAP concept.[88b]

Based on this 2-fold completive self-sorting the authors integrated the binding motifs into the ditopic ligands **191–193** to fabricate the five-component supramolecular trapezoid **196**. Theoretically, ligands **191–193** along with zinc(II) and copper(I) metal ions may assemble to more than 300 discrete irregular tetragons. This self-sorting will thus have a degree of sorting $M$ of 300. However, using the information from Scheme 47, one even expects the formation of three architectures as depicted in Scheme 48: a trapezoid **196**, a small rectangle **194** and a large rectangle **195**. Both **196** and **195** have an ideal situation at both $Zn^{2+}$ and $Cu^{+}$ complexation sites. According to theoretical calculations, however, the geometry at the metal coordination centres in **196** is not perfect due to the difference in length of **191** and **192**. Thus, along a reasonable stability sequence the large rectangle **195** may be expected to constitute the most stable entity:




**Scheme 47** (a) Chemical structures of **185–190**. (b) A $2^{3,3}$-fold(6) completive self-sorting system using $Zn^{2+}$ and $Cu^{+}$ ions.[18]

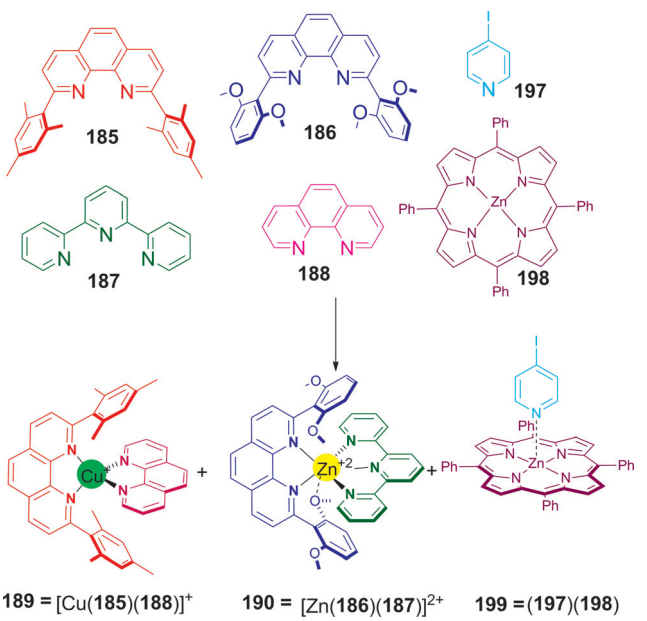



**Scheme 48** (a) Chemical structures and corresponding cartoon representation of ligands **191–193**. (b) $1^8$-Fold(5) completive self-sorting of the supramolecular trapezoid **196**.[18]

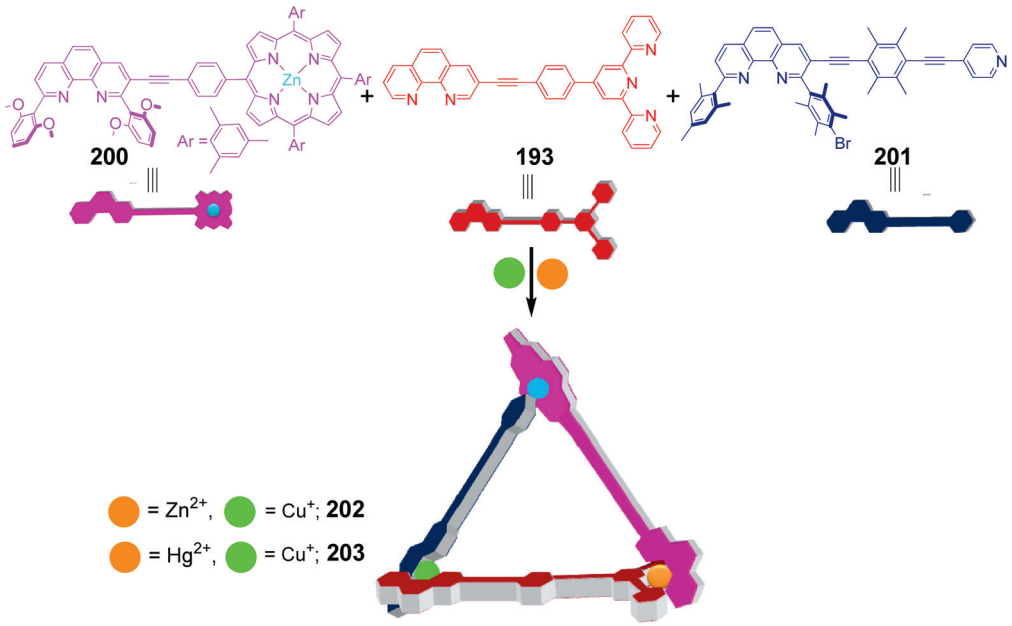


**195** > **196** > **194**. Yet for the global reaction outcome, the total energy of the ensemble is of importance, not the individual energy of a species. Thus, due to the initial choice of a 1 : 1 stoichiometry of **191** and **192**, the formation of rectangle **195** will be paralleled by an equal amount of **194**. Because the cornerstone [Zn(**185**)(**187**)](OTf)$_2$ is less stable than [Zn(**186**)(**187**)](OTf)$_2$, we expect **194** to be the least stable entity in this competition. Thus, one molecule of **194** and one molecule of **195** will dynamically reassemble into two molecules of **196** under the guidance of thermodynamics. Later, the same principle was also applied to fabricate a multicomponent supramolecular triangle.[89,90]

Recently, the Schmittel group has extended the conceptual insights of the above self-sorting system into an 8-component self-sorting system to prepare the scalene triangles **202**, **203**. In continuation to the aforementioned 6-component library, the authors describe a 8-component library. Out of 35 possible metallo-organic complexes that could result from such mixture (Scheme 49), only three were observed, proving a degree of self-sorting $M$ = 11.7. Here, the orthogonality of the porphyrin–pyridine interaction in **199** = (**197**)(**198**) with the other heteroleptic complexes **189** and **190** is mostly governed by the maximum site occupancy rule and the associated steric bulk of bi- or




**Scheme 49** $3^{2,3,3}$-Fold(8) completive self-sorting in a $A^1A^2A^3D^1D^2D^3D^4D^5$ library.[91]




**Scheme 50** Synthesis of the supramolecular scalene triangles **202**, **203**.[91]

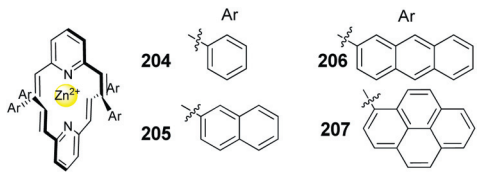



**Fig. 10** Chemical structures of homoduplex complexes **204**–**207**.[25]

tridentate ligands interacting with the zinc porphyrin **198**.[91] Using these components, they designed the three compounds **193**, **200** and **201**. Clearly, the compounds are devised in a way that the three complexes formed by self-sorting would make up the three corners of a triangle (Scheme 50). The assembly was characterised by ESI-MS, $^{1}$H-NMR, $^{1}$H-$^{1}$H COSY, diffusion ordered spectroscopy (DOSY) and differential pulse voltammetry (DPV). Additional studies allowed replacing one $Zn^{2+}$ centre by $Hg^{2+}$ to yield the trisheterometallic scalene triangle **203**, the first of its kind, as demonstrated by $^{1}$H-NMR and ESI-MS. Metal exchange in **202** by adding $Hg^{2+}$ entails a transformation of **202** to **203** within one day.

## 5. Evaluation of completive self-sorting using the *A* value

In the following we describe several selected cases, where all possible products are formed but their distribution violates the statistical expectation.

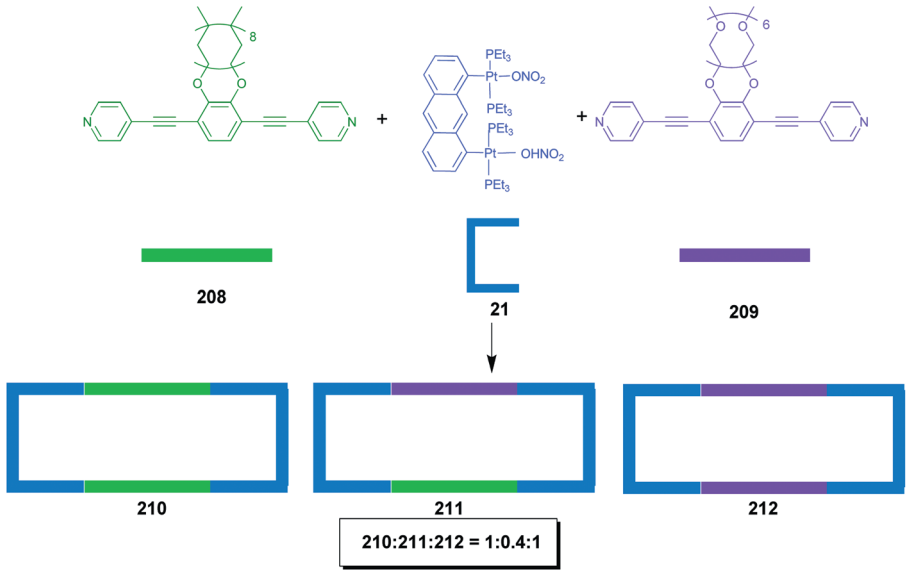
Recently, Barboiu *et al.* surveyed the ligand exchange of the constitutionally dynamic 2,6-bis(iminoarene)pyridine zinc(II) complexes **204**–**207** (Fig. 10).[25] When the pyrene-based homoduplex complex **207** was combined with the other homoduplex complexes **204**–**206**, a mixed incomplete self-sorting process occurred involving roughly a 50% amplification of the heteroduplex complexes over the statistical probability. The enriched self-sorting of the heteroduplex complexes is likely to be the result of geometric constraints about the pyrene-based ligand. On the other side, in the solid state structures, intramolecular C···π interactions between N=CH protons of the phenyl, naphthyl, and anthracenyl substituted ligands and the pyrenyl moieties of a second coordinated ligand are realised in heteroduplex complexes.

Stang *et al.* utilised solvophobic interactions to trigger second-order self-organisation, *i.e.* effects that are far remote from the principle determinants of self-assembly.[26] To make sure that the observed self-recognition in **208** and **209** (Scheme 51) arises solely from remote substituent effects, the geometry, length, and dimensionality of each pyridine was kept constant, while the long alkyl and ethylene glycol chains were kept away from the nitrogen coordination sites. When both donors **208** and **209** were mixed with the diplatinum(II) clip **21** as a 0° acceptor in a 1 : 1 : 2 ratio, the hydrophobic, amphiphilic, and hydrophilic supramolecular rectangles **210**–**212** were formed in a 1 : 0.4 : 1 ratio, in contrast to a statistical 1 : 2 : 1 ratio. The observed amplification in self-recognition is mostly triggered by the aggregation of like-philicity donors, thus biasing the formation of rectangles with identical donors. The driving force for aggregation increases with increasing chain length, accounting for the trend in amplified self-organisation.[26] By comparing the statistical and experimental abundance of **211**, we conclude that the mixed assembly is decreased by ≅80% compared to the theoretical statistical distribution.

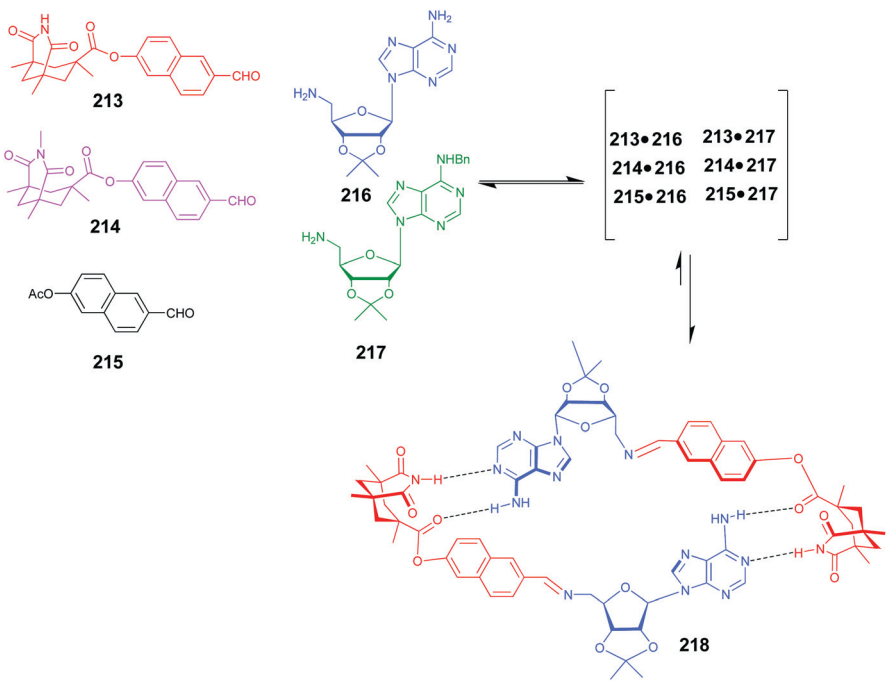
Giuseppone & Xu utilised complementary hydrogen bonded DA pairs (D: donor, A: acceptor) to achieve complementary pair selection from a pool of suitably programmed self-instructed aldehydes and amines.[27] For example, the key molecule **213**·**216**, synthesised by condensation of aldehyde **213** (derived from a Kemp's imide) and adenosine amine **216**, is able to strongly associate with itself into the dimeric complex **218**






**Scheme 51** Amplification in self-sorting of ditopic donors **208** and **209** in the formation of rectangles **210–212**.[26]




**Scheme 52** Amplification of **218** from a mixture of three aldehydes **213–215** and the two amines **216–217**.[27]

(Scheme 52). In a first set of experiment (DCL1), **214–217** (15 mM each at 22 °C in $CDCl_3$) were mixed, and the product distribution at equilibrium clearly showed statistical distribution. Notably, when **213–217** (15 mM each at 22 °C in $CDCl_3$) were mixed together, the production of homodimer **218** now strongly biased the expression of all constituents at equilibrium, far off the statistical distribution (DCL2). In a third round with DCL3, actually two experiments were performed: (i) a pre-equilibrated library was set up without **213** and then (ii) **213** (15 mM) was added to this library. In DCL3, it took five times longer to reach equilibrium (22 days) compared to DCL2 ($t$ = 109 h), but the competition produced an identical distribution of constituents in both DCL2 and DCL3. By comparing the "non-self-duplicating" DCL1 and the "self-duplicating" DCL2 and DCL3, one concludes that the self-duplicator **213·216** (9.03 mM) is increased by +83% compared to the theoretical statistical distribution (4.94 mM).

In 2012, Gibb & Gan examined the extent of guest-controlled amplification of the heterodimer **219·220** from mixtures of the deep-cavity cavitands **219** and **220**.[92] Fig. 11b shows the percentage of hetero-capsule formation as a function of the guest size. Clearly, the values irregularly range from 24%, as in the case of

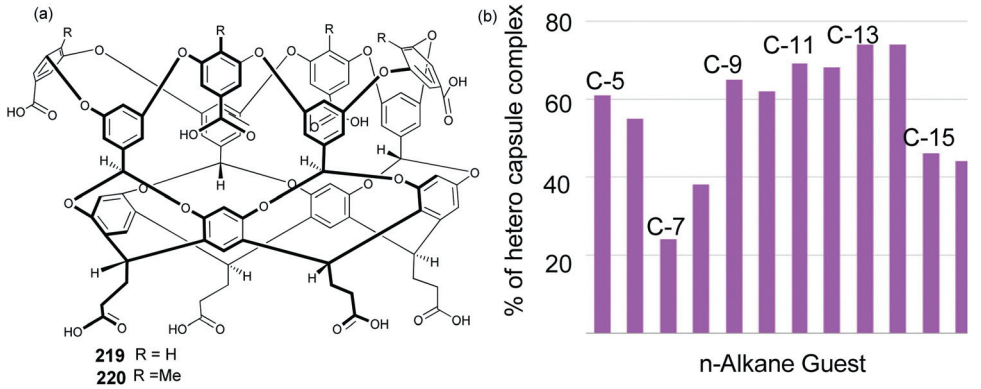



**Fig. 11** (a) Chemical structures of **219**, **220**. (b) Extent of hetero-capsule (**219**·**220**) formation *versus* the size of the guest(s).[92] Reproduced from ref. 92 with permission from the Royal Society of Chemistry.



*n*-heptane, to 74% for *n*-tridecane. Smaller guests such as *n*-pentane and *n*-hexane form 61% and 55% heterocomplexes. Thus, for these guests there is only a slight bias away from a statistical mixture (50%) toward formation of the hetero-complex. In contrast, *n*-heptane and *n*-octane as guests afford the least amount of hetero-capsules (24% and 38%, respectively). Actually, both guests are poor templates for **220**·**220** and **219**·**220**. As a result, self-sorting favours the stable **219**·**219**, driving the formation of **220**·**220** along with very little amounts of the hetero-complex **219**·**220**. However, larger guests, such as *n*-nonane to *n*-tetradecane, form the heterocomplex in yields of 62% and 74%. The key difference between the latter complexes and those involving smaller guests is that efficient packing of the capsule requires only one guest, although it is not yet known whether the effect is of enthalpic or entropic origin. Finally, for the largest guests examined, *n*-pentadecane and *n*-hexadecane, there is a relative drop in the amount of heterocomplex formed to a below statistical level. Why this occurs is unclear although these molecules are close to the maximum permissible guest size. Probably, the shift towards increased self-sorting indicates that one of the homocapsules is better suited for binding these larger guests than **219**·**220**.

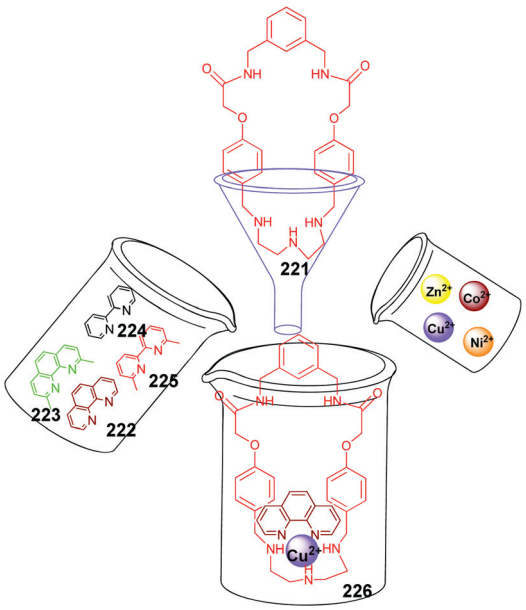



**Fig. 12** Selective formation of the ternary complex **226** in a heteromeric $1^3$-fold(3 + 6) incomplete self-sorting.[93]

## 6. Incomplete self-sorting

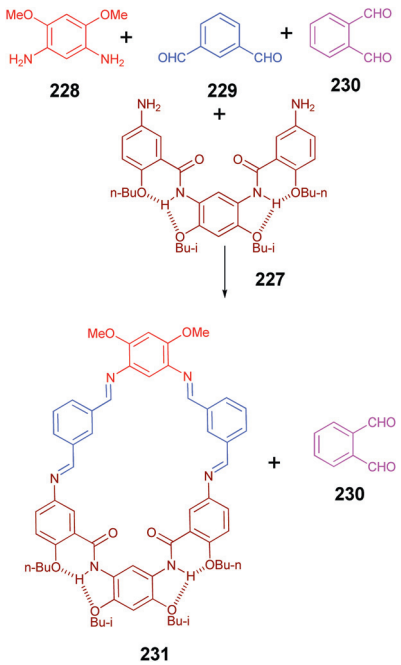
Incomplete self-sorting has several valuable aspects. Through the use of a self-sorting algorithm generating not only the self-sorted assemblies, but equally unused components, one can either demonstrate the robustness and high selectivity of the self-sorting ensemble in comparison to the left-over component(s) or one may even further utilise unused components for effecting various purposes, such as rejuvenation of components that were destroyed, input for chemical or electrical signalling, *etc*. In the ensuing sections we will first portray examples of incomplete self-sorting and then its application in rejuvenation.

### 6.1 Examples of incomplete self-sorting

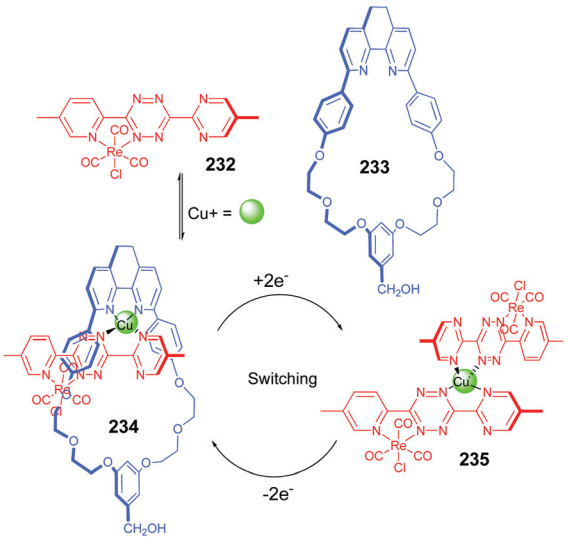
In a recent work, Ghosh *et al.* describe the copper(II)-templated formation of the pseudorotaxane **226** from macrocycle **221** and phenanthroline **222**.[93] **221** and $Cu^{2+}$ preferentially form a pseudorotaxane with the parent 1,10-phenanthroline (**222**) over those with 2,9-dimethyl-1,10-phenanthroline (**223**), 2,2′-bipyridine (**224**), and 6,6′-dimethylbipyridine (**225**). Interestingly, a 1 : 1 combination of **221** and **222** selectively binds to $Cu^{2+}$ and less to $Co^{2+}$, $Ni^{2+}$, $Zn^{2+}$. Furthermore, selective formation of the copper(II) pseudorotaxane **226** occurs from a mixture of **221**–**225** and $Cu^{2+}$, $Co^{2+}$, $Ni^{2+}$, $Zn^{2+}$ (Fig. 12). The selective formation of **226** may be attributed to the suitable coordination sites and geometry of $Cu^{2+}$ in the cavity of **221**, whereas preference of **222** over other bidentate ligands **223**–**225** could be due to steric effect as well as π–π stacking interactions. Out of 16 possible pseudorotaxanes, the mixture furnishes **226** along with free **223**–**225** and $Co^{2+}$, $Ni^{2+}$, $Zn^{2+}$ metal ions, suggesting $M = 16$.

As described in Scheme 53, a mixture of two dialdehydes **229**, **230** and two diamines **227**, **228** successfully self-sorts in a reversible macrocyclisation. To our understanding, the two diamines are structurally preorganised in a way that dialdehyde




**Scheme 53** Selective formation of macrocycle **231** from a 4-component reaction involving $1^4$-fold(3 + 1) incomplete self-sorting.[94]




**Scheme 54** The self-sorting mixture of two ligands (**232** and **233**) about a $Cu^+$ ion shows redox-driven interconversion between two distinct architectures: pseudorotaxane **234** and grid-corner complex **235**.[96]

**229**, due to its geometric match, exclusively selects a widely and a narrow spaced diamine, *i.e.* **227** and **228**.[94] Dialdehyde **230** will not be used as component. The intramolecular hydrogen bonding in **227** by may provide a tool to control the structural preorganisation by turning ON and OFF the hydrogen bonding.

### 6.2. Applications of incomplete self-sorting

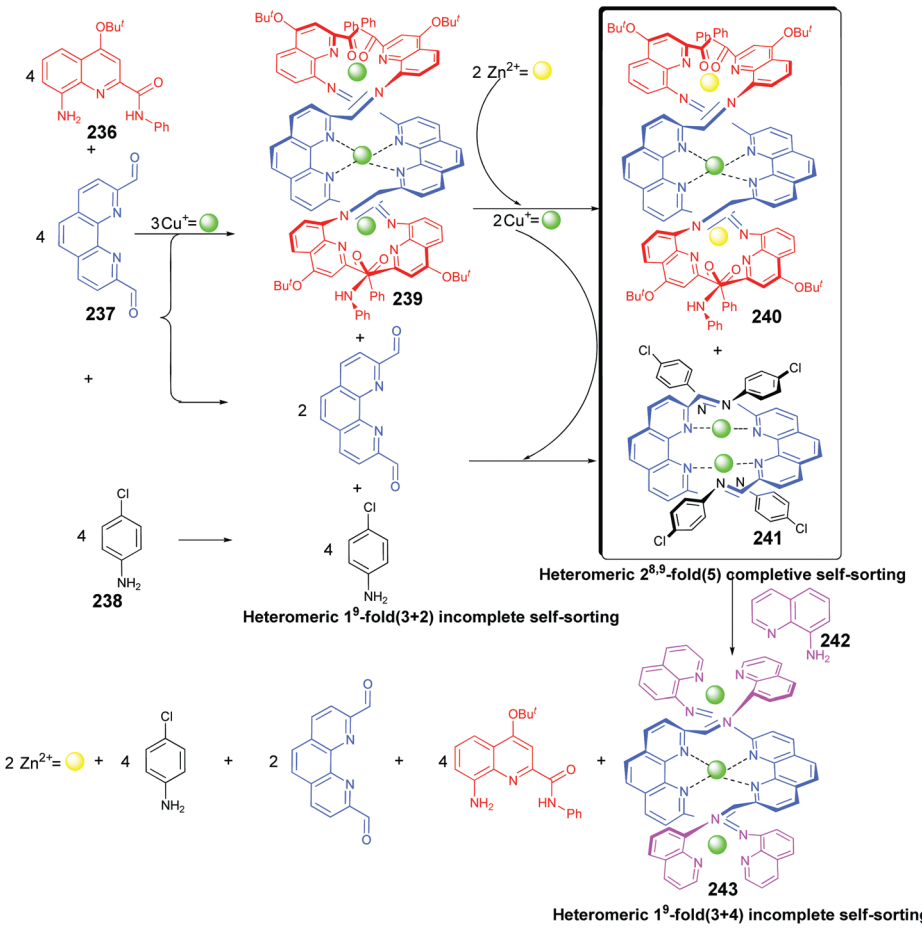
The detailed investigations on the self-sorting process presented herein may provide valuable insight into analogous biological self-sorting processes. For example, multiproteins complexes are astoundingly intricate assemblies. They form various types of molecular machinery and are responsible for a vast array of biological functions. For example, the voltage-gated potassium channel in a plasma membrane of a neuron is composed out of heteromultimeric proteins.[95]

No system, however, is truly perfect, and it is possible, even necessary at times, to rejuvenate elegant self-sorted systems. Therefore, in truly complex functioning systems that require repair, *incomplete self-sorting* (*completive* + extra ligand(s)) may become the true choice.

As a beautiful example of incomplete sorting, Flood *et al.* successfully demonstrated the redox controlled self-selection and self-discrimination behaviour of the tetrazine based ligand **232**.[96] When a mixture of **232** and **233** was exposed to copper(I) ion (2 : 1 : 1), [2]pseudorotaxane **234** was observed as exclusive, thermodynamically driven product. Due to topological constraints, the phenanthroline-based macrocycle fails to produce the $[Cu(\mathbf{233})_2]^+$ complex. $[Cu(\mathbf{232})_2]^+$ does not form due to the lack of σ donor quality of **232**. Whereas the pyridyl *N* is a strong σ donor and weak π acceptor, the tetrazyl *N* is a weak σ donor and strong π acceptor. Thus, the two nitrogen atoms in **232** have quite distinct electronic character. The electronic properties of the tetrazine unit can be inverted (to a strong σ donor and π donor) *via* reduction. Upon reduction of **232**, the electronic properties of the tetrazine core are inverted furnishing strong donor *N* atoms, while the redox-innocent pyridine unit remains almost unchanged. Therefore, two of the reduced ligands preferentially form the homoleptic complex **235**, which resembles a corner piece of a grid-type complex. The reversible switching between the different architectures was also achieved (Scheme 54).

Recently, Huc and Nitschke *et al.* developed a magnificent system, in which the application of different chemical signals through the addition of a metal ion or a small molecule – induces the system to reconstitute in intricate ways.[97] As shown in Scheme 55, helicate **239** is the thermodynamically controlled product in the presence of **236–238** (4 equiv.) and $Cu^+$ (3 equiv.). **239** is afforded in preference to the dicopper(I) double helicate **241**, when limited in $Cu^+$.[62b,98] In light of the maximum site occupancy rule, the mixture furnishes helicate **239** along with free **237–238** (2 equiv.), from the 2 possible homo helicates, suggesting an *M* value of 2. Addition of $Zn(OAc)_2$ (2 equiv.) to this system, however, gives rise to two distinct transformations: helicate **239** was converted into **240**, and the ejected $Cu^+$ (2 equiv.) reacted with the free **237–238** (2 equiv.) to generate the dicopper double helicate **241**. At this stage the system represents a heteromeric $2^{8,9}$-fold(5) completive self-sorting system with $M = 4.5$. Furthermore, the addition of 4 equiv. of unsubstituted 8-aminoquinoline (**242**) to a stable mixture of **240** and **241** resulted in the formation of tricopper helicate **243**. 24 h after the addition of **242**, the $^1$H-NMR spectra show resonances corresponding to **243** and free **236–238**, but no resonances corresponding to **240** and **241**. In this scenario, the system represents a heteromeric $1^9$-fold(3 + 4) incomplete self-sorting, with $M = 47$.

The greater thermodynamic stability of **243** compared to **240** and **241** is attributed to the more electron-rich character of the incorporated unit **242** than with either **236** or **238**. The preference of helicate **243** to incorporate $Cu^+$ rather than $Zn^{2+}$ is




**Scheme 55** Self-assembly in which the addition of one signal ($Zn^{2+}$ or 8-aminoquinoline) induced two distinct transformations.[97]



assigned to the lack of the carbonyl groups in unit **242**, while in **240** they are well positioned to stabilise a $Zn^{2+}$ centre. As each equivalent of **243** requires three $Cu^{+}$ ions, its formation triggers the disassembly of both **240** and **241**.

## 7. Conclusion

In one of his famous statements, Leonardo da Vinci declared[99] "*Where nature finishes producing its own species, man begins, using natural things and with the help of this nature, to create an infinity of species*". In this article, we describe selected examples of artificial supramolecular assemblies that form due to the sophisticated design of highly selective molecular recognition or discrimination. In principle, also experimental parameters, such as temperature, concentration of individual components, association constants of complementary pairs and presence of competitors, may also influence the observed sorting. Therefore in a broader perspective, under a given set of experimental conditions, self-sorting chemical systems must yield a product distribution that is different from the plausible compositions of aggregates estimated on the basis of statistical, chemical and geometrical arguments. As a tool for deeper understanding, we determined the degree of self-sorting ($M$) for many of those processes. To our appreciation, both the classification in *completive* and *incomplete self-sorting* and additionally the determination of $M$ values provide a good guideline to classify and quantify the basic sorting phenomenon.

Following our classification, most of the examples are $x$-fold completive ($x > 1$) giving rise to at least two self-sorted aggregates. However, there are large differences to be accounted for in these processes as the degree of self-sorting may vary from $M = 1$ to $M > 5$. Not unexpectedly, the degree of self-sorting in 1-fold completive is generally somewhat higher, whereas integrative self-sorting leads to numbers as high as $M \geq 300$. While the degree of self-sorting $M = P_0/P$ is not undisputable due to the vagueness to determine the number of all possible structures $P_0$, $M$ will still give a good orientation on the selectivity of the process.

The observed fidelity in self-sorting systems is basically always controlled through maximum site occupancy and keeping entropic costs as low as possible. Processes with $M > 1$ need to be complemented by the clever use of additional players, such as geometrical constraints, steric effects, coordination geometry of the metal ions, electrostatic interactions, *etc.* In spite of the presented richness of fascinating examples, self-assembly of multiple different components into a single supramolecular species is still a huge challenge to chemists. To our perspective, *completive self-sorting* in multicomponent system requires that each individual component is perfectly pre-programmed so that

multiple recognition events occur in a globally complementary manner between the pieces during the self-assembly process.

What is the future of self-sorting? Although *completive* and in particular *integrative self-sorting* seem to provide the largest degrees of self-sorting, we feel that equally *incomplete* self-sorting combined with its potential for functional use in molecular machinery will attract increasing attention in the near future. The examples given in Section 6.2 may serve as an illustration for the power of cleverly designed *incomplete* self-sorted systems.

## Acknowledgements

We are indebted to the Deutsche Forschungsgemeinschaft and the University of Siegen for financial support. We are grateful to Dr. Kingsuk Mahata from the University of Würzburg for his initial help and suggestions and to Mr. Anup Rana from the University of Siegen for assistance with the some drawings.

## References

1 R. Krämer, J.-M. Lehn and A. Marquis-Rigault, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 5394.
2 P. N. Taylor and H. L. Anderson, *J. Am. Chem. Soc.*, 1999, **121**, 11538.
3 (*a*) A. Wu and L. Isaacs, *J. Am. Chem. Soc.*, 2003, **125**, 4831; (*b*) P. Mukhopadhyay, A. Wu and L. Isaacs, *J. Org. Chem.*, 2004, **69**, 6157.
4 Y.-R. Zheng, H.-B. Yang, B. H. Northrop, K. Ghosh and P. J. Stang, *Inorg. Chem.*, 2008, **47**, 4706.
5 (*a*) J. Gleick, *Chaos: Making a New Science*, Penguin, New York, 1988; (*b*) S. Kauffman, *At Home in the Universe*, Oxford University Press, Oxford, U.K., 1995; (*c*) G. M. Whitesides and B. Grzybowski, *Science*, 2002, **295**, 2418.
6 D. Voet and J. G. Voet, *Biochemistry*, Wiley, New York, 1995.
7 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737.
8 R. H. Carlson, C. V. Gabel, S. S. Chan, R. H. Austin, J. P. Brody and J. W. Winkelman, *Phys. Rev. Lett.*, 1997, **79**, 2149.
9 (*a*) L. A. Amos and T. S. Baker, *Nature*, 1979, **279**, 607; (*b*) A. Desai and T. J. Mitchison, *Annu. Rev. Cell Dev. Biol.*, 1997, **13**, 83.
10 J. S. Wicken, *J. Theor. Biol.*, 1979, **77**, 349.
11 M. Schmittel and K. Mahata, *Angew. Chem., Int. Ed.*, 2008, **47**, 5284.
12 (*a*) M. Hutin, C. J. Cramer, L. Gagliardi, A. R. M. Shahi, G. Bernardinelli, R. Cerny and J. R. Nitschke, *J. Am. Chem. Soc.*, 2007, **129**, 8774; (*b*) M. Barboiu, F. Dumitru, Y.-M. Legrand, E. Petit and A. van der Lee, *Chem. Commun.*, 2009, 2192.
13 (*a*) K. Osowska and O. Š. Miljanić, *J. Am. Chem. Soc.*, 2011, **133**, 724; (*b*) K. Osowska and O. Š. Miljanić, *Angew. Chem., Int. Ed.*, 2011, **50**, 8345.
14 G. Celtek, M. Artar, O. A. Scherman and D. Tuncel, *Chem.–Eur. J.*, 2009, **15**, 10360.
15 W. Jiang and C. A. Schalley, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10425.
16 M. M. Safont-Sempere, G. Fernández and F. Würthner, *Chem. Rev.*, 2011, **111**, 5784.
17 K. Osowska and O. Š. Miljanić, *Synlett*, 2011, **12**, 1643.
18 K. Mahata and M. Schmittel, *J. Am. Chem. Soc.*, 2009, **131**, 16544.
19 B. H. Northrop, Y.-R. Zheng, K.-W. Chi and P. J. Stang, *Acc. Chem. Res.*, 2009, **42**, 1554.
20 W. Jiang, H. D. F. Winkler and C. A. Schalley, *J. Am. Chem. Soc.*, 2008, **130**, 13852.
21 Incomplete self-sorting is often observed for chemosensors where the receptor/sensor binds only one out of *x* analytes.
22 The evaluation of $P_0$ and $M$ is only based on the formation of discrete assemblies; thus, free/unused component(s) are not taken into account.
23 Even based on our boundary conditions, (i) maximum site occupancy rule, and (ii) lowest possible entropy, the evaluation of $P_0$ remains disputable, as we consider these rules only on a local but not global level. Analysis on a local level means that consideration of an individual plausible assembly based on (i) and (ii) may disregard that (i) and (ii) in principle needs to be fulfilled on the global level as well. For example, when we account for the binary complexes of A and B (1 : 1), our local analysis to $M$ uses $P_0 = 3$ for AA, AB and BB, while a global approach would only account for two options: AA + BB versus 2AB. While the latter approach is easily accomplished for small libraries, it delivers either unmanageable or even infinite numbers for larger self-sorting scenarios. Thus for practical reasons, we have utilised the local approach to determine $P_0$.
24 F. Dumitru, E. Petit, A. van der Lee and M. Barboiu, *Eur. J. Inorg. Chem.*, 2005, 4255.
25 Y.-M. Legrand, A. van der Lee and M. Barboiu, *Inorg. Chem.*, 2007, **46**, 9540.
26 B. H. Northrop, H.-B. Yang and P. J. Stang, *Inorg. Chem.*, 2008, **47**, 11257.
27 S. Xu and N. Giuseppone, *J. Am. Chem. Soc.*, 2008, **130**, 1826.
28 B. Olenyuk, J. A. Whiteford, A. Fechtenkötter and P. J. Stang, *Nature*, 1999, **398**, 796.
29 S. De, K. Mahata and M. Schmittel, *Chem. Soc. Rev.*, 2010, **39**, 1555.
30 In the evaluation of $P_0$ all possible homo-dimerisations (3), hetero-dimerisations (3), homo-trimerisations (3) and hetero-trimerisations (15) are considered.
31 K. A. Jolliffe, P. Timmerman and D. N. Reinhoudt, *Angew. Chem., Int. Ed.*, 1999, **38**, 933.
32 We have listed this self-assembly process under completive self-sorting although a slight, but undefined excess of **5** was used.
33 H. Ito, Y. Furusho, T. Hasegawa and E. Yashima, *J. Am. Chem. Soc.*, 2008, **130**, 14008.
34 In general, $N$ objects will combine to form $N(N + 1)/2$ different dimeric pairs.
35 P. J. Stang and B. Olenyuk, *Acc. Chem. Res.*, 1997, **30**, 502.
36 (*a*) R. Chakrabarty, P. S. Mukherjee and P. J. Stang, *Chem. Rev.*, 2011, **111**, 6810; (*b*) H.-B. Yang, K. Ghosh, B. H. Northrop and P. J. Stang, *Org. Lett.*, 2007, **9**, 1561.
37 Y.-R. Zheng, H.-B. Yang, K. Ghosh, L. Zhao and P. J. Stang, *Chem.–Eur. J.*, 2009, **15**, 7203.
38 R. Pinalli, V. Cristini, V. Sottili, S. Geremia, M. Campagnolo, A. Caneschi and E. Dalcanale, *J. Am. Chem. Soc.*, 2004, **126**, 6516.
39 H. Maeda, K. Kinoshita, K. Naritani and Y. Bando, *Chem. Commun.*, 2011, **47**, 8241.
40 B. Bilgiçer, X. Xing and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 11815.
41 N. C. Yoder and K. Kumar, *J. Am. Chem. Soc.*, 2006, **128**, 188.
42 J.-M. Han, J.-L. Pan, T. Lei, C. Liu and J. Pei, *Chem.–Eur. J.*, 2010, **16**, 13850.
43 A. D. Shaller, W. Wang, H. Gan and A. D. Q. Li, *Angew. Chem., Int. Ed.*, 2008, **47**, 7705.
44 M. Albrecht, M. Schneider and H. Röttele, *Angew. Chem., Int. Ed.*, 1999, **38**, 557.
45 A. M. Johnson and R. J. Hooley, *Inorg. Chem.*, 2011, **50**, 4671.
46 The present example furnishes actually six possible products, if we would additionally take into account stereoisomers. In that case, for the adduct $Pd_2(\mathbf{51})_2(\mathbf{52})_2$ there would be *cis*- and *trans*-isomers and $M$ would amount to 3.
47 I. Saur, R. Scopelliti and K. Severin, *Chem.–Eur. J.*, 2006, **12**, 1058.
48 M. Barboiu, E. Petit, A. van der Lee and G. Vaughan, *Inorg. Chem.*, 2006, **45**, 484.
49 (*a*) S. Ulrich and J.-M. Lehn, *J. Am. Chem. Soc.*, 2009, **131**, 5546; (*b*) S. Ulrich and J.-M. Lehn, *Chem.–Eur. J.*, 2009, **15**, 5640.
50 (*a*) S. Liu, C. Ruspic, P. Mukhopadhyay, S. Chakrabarti, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 15959; (*b*) P. Mukhopadhyay, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2006, **128**, 14093.
51 E. Masson, X. Lu, X. Ling and D. L. Patchell, *Org. Lett.*, 2009, **11**, 3798.
52 M. V. Rekharsky, H. Yamamura, Y. H. Ko, N. Selvapalam, K. Kim and Y. Inoue, *Chem. Commun.*, 2008, 2236.
53 (*a*) A. Wu, A. Chakraborty, J. C. Fettinger, R. A. Flowers II and L. Isaacs, *Angew. Chem., Int. Ed.*, 2002, **41**, 4028; (*b*) S. Ghosh, A. Wu, J. C. Fetinger, P. Y. Zavalij and L. Isaacs, *J. Org. Chem.*, 2008, **73**, 5915.
54 L.-P. Cao, J.-G. Wang, J.-Y. Ding, A.-X. Wu and L. Isaacs, *Chem. Commun.*, 2011, **47**, 8548.
55 (*a*) L. Avram and Y. Cohen, *J. Am. Chem. Soc.*, 2004, **126**, 11556; (*b*) E. S. Barrett, T. J. Dale and J. Rebek Jr., *J. Am. Chem. Soc.*, 2008, **130**, 2344.
56 D. Braekers, C. Peters, A. Bogdan, Y. Rudzevich, V. Böhmer and J. F. Desreux, *J. Org. Chem.*, 2008, **73**, 701.

57 D. L. Caulder and K. N. Raymond, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 1440.
58 S. J. Rowan, D. G. Hamilton, P. A. Brady and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1997, **119**, 2578.
59 R. Krämer, J.-M. Lehn, A. D. Cian and J. Fischer, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 703.
60 The use of maximum site occupancy for metal ion–ligand interactions requires to apply a reasonable coordination number. For $Ni^{2+}$ this generates a problem, as it can form octahedral, tetrahedral and square-planar complexes, even with N donor ligands. However, $NiN_4$ low-spin complexes are usually nickel(II)porphyrin complexes, so that for the present case we assume an octahedral coordination only.
61 J. R. Nitschke, *Acc. Chem. Res.*, 2007, **40**, 103.
62 (*a*) D. Schultz and J. R. Nitschke, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 11191; (*b*) M. Hutin, R. Frantz and J. R. Nitschke, *Chem.–Eur. J.*, 2006, **12**, 4077.
63 D. Schultz and J. R. Nitschke, *Angew. Chem., Int. Ed.*, 2006, **45**, 2453.
64 C. Addicott, N. Das and P. J. Stang, *Inorg. Chem.*, 2004, **43**, 5335.
65 Y. Rudzevich, V. Rudzevich, F. Klautzsch, C. A. Schalley and V. Böhmer, *Angew. Chem., Int. Ed.*, 2009, **48**, 3867.
66 K.-W. Chi, C. Addicott, A. M. Arif and P. J. Stang, *J. Am. Chem. Soc.*, 2004, **126**, 16569.
67 S. Ghosh, D. R. Turner, S. R. Batten and P. S. Mukherjee, *Dalton Trans.*, 2007, 1869.
68 D. Ajami, J.-L. Hou, T. J. Dale, E. Barrett and J. Rebek, Jr., *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10430.
69 M. Chas, G. G. Ramírez and P. Ballester, *Org. Lett.*, 2011, **13**, 3402.
70 M. Li, K. Yamato, J. S. Ferguson and B. Gong, *J. Am. Chem. Soc.*, 2006, **128**, 12628.
71 (*a*) K. Kobayashi, Y. Yamada, M. Yamanaka, Y. Sei and K. Yamaguchi, *J. Am. Chem. Soc.*, 2004, **126**, 13896; (*b*) M. Yamanaka, Y. Yamada, Y. Sei, K. Yamaguchi and K. Kobayashi, *J. Am. Chem. Soc.*, 2006, **128**, 1531.
72 S. J. Lee, K. L. Mulfort, X. Zuo, A. J. Goshe, P. J. Wesson, S. T. Nguyen, J. T. Hupp and D. M. Tiede, *J. Am. Chem. Soc.*, 2008, **130**, 836.
73 S. J. Lee, S.-H. Cho, K. L. Mulfort, D. M. Tiede, J. T. Hupp and S. T. Nguyen, *J. Am. Chem. Soc.*, 2008, **130**, 16828.
74 L. Zhao, B. H. Northrop, Y.-R. Zheng, H.-B. Yang, H. J. Lee, Y. M. Lee, J. Y. Park, K.-W. Chi and P. J. Stang, *J. Org. Chem.*, 2008, **73**, 6580.
75 (*a*) M. Yoshizawa, M. Nagao, K. Kumazawa and M. Fujita, *J. Organomet. Chem.*, 2005, **690**, 5383; (*b*) M. Yamanaka, Y. Yamada, Y. Sei, K. Yamaguchi and K. Kobayashi, *J. Am. Chem. Soc.*, 2006, **128**, 1531; (*c*) J. Zhang, P. W. Miller, M. Nieuwenhuyzen and S. L. James, *Chem.–Eur. J.*, 2006, **12**, 2448.
76 (*a*) Y. Yamauchi, M. Yoshizawa, M. Akita and M. Fujita, *J. Am. Chem. Soc.*, 2010, **132**, 960; (*b*) T. Murase, K. Otsuka and M. Fujita, *J. Am. Chem. Soc.*, 2010, **132**, 7864.
77 N. A. Schnarr and A. J. Kennan, *J. Am. Chem. Soc.*, 2001, **123**, 11081.
78 N. A. Schnarr and A. J. Kennan, *J. Am. Chem. Soc.*, 2002, **124**, 9779.
79 B. Brusilowskij, E. V. Dzyuba, R. W. Troff and C. A. Schalley, *Chem. Commun.*, 2011, **47**, 1830.
80 N. André, T. B. Jensen, R. Scopelliti, D. Imbert, M. Elhabiri, G. Hopfgartner, C. Piguet and J.-C. G. Bünzli, *Inorg. Chem.*, 2004, **43**, 515.
81 F. E. Hahn, M. Offermann, C. Schulze Isfort, T. Pape and R. Fröhlich, *Angew. Chem., Int. Ed.*, 2008, **47**, 6794.
82 C. Zhang, S. Li, J. Zhang, K. Zhu, N. Li and F. Huang, *Org. Lett.*, 2007, **9**, 5553.
83 W. Jiang, A. Schäfer, P. C. Mohr and C. A. Schalley, *J. Am. Chem. Soc.*, 2010, **132**, 2309.
84 W. Jiang, D. Sattler, K. Rissanen and C. A. Schalley, *Org. Lett.*, 2011, **13**, 4502.
85 W. Jiang, Q. Wang, I. Linder, F. Klautzsch and C. A. Schalley, *Chem.–Eur. J.*, 2011, **17**, 2344.
86 S. Hiraoka, T. Tanaka and M. Shionoya, *J. Am. Chem. Soc.*, 2006, **128**, 13038.
87 S. Hiraoka, M. Goda and M. Shionoya, *J. Am. Chem. Soc.*, 2009, **131**, 4592.
88 (*a*) M. Schmittel and A. Ganz, *Chem. Commun.*, 1997, 999; (*b*) M. Schmittel, V. Kalsani, R. S. K. Kishore, H. Cölfen and J. W. Bats, *J. Am. Chem. Soc.*, 2005, **127**, 11544.
89 M. Schmittel and K. Mahata, *Chem. Commun.*, 2010, **46**, 4163.
90 K. Mahata and M. Schmittel, *Beilstein J. Org. Chem.*, 2011, **7**, 1555.
91 K. Mahata, M. L. Saha and M. Schmittel, *J. Am. Chem. Soc.*, 2010, **132**, 15933.
92 H. Gan and B. C. Gibb, *Chem. Commun.*, 2012, **48**, 1656.
93 S. Saha, I. Ravikumar and P. Ghosh, *Chem. Commun.*, 2011, **47**, 6272.
94 J.-B. Lin, X.-N. Xu, X.-K. Jiang and Z.-T. Li, *J. Org. Chem.*, 2008, **73**, 9403.
95 S. B. Long, E. B. Campbell and R. MacKinnon, *Science*, 2005, **309**, 897.
96 K. Parimal, E. H. Witlicki and A. H. Flood, *Angew. Chem., Int. Ed.*, 2010, **49**, 4628.
97 V. E. Campbell, X. D. Hatten, N. Delsuc, B. Kauffmann, I. Huc and J. R. Nitschke, *Nat. Chem.*, 2010, **2**, 684.
98 In ref. 59b, the authors describe that both helicates may form when 5 equiv. of Cu+ are used.
99 J.-M. Lehn, *Supramolecular Chemistry*, VCH Publishers, New York, 1995.

