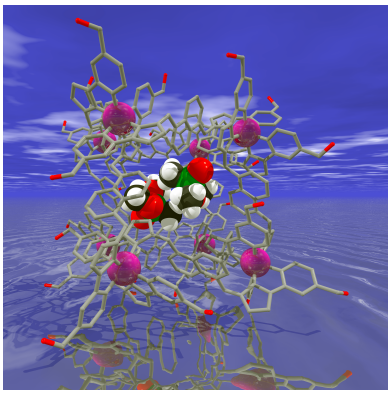
Coordination cages based on bis(pyrazolyl-pyridine) ligands: structures, dynamic behaviour, guest binding and catalysis

Michael D. Ward,[a,*] Christopher A. Hunter[b] and Nicholas H. Williams[c]

a *Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK*
Email: m.d.ward@warwick.ac.uk

b *University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK*

c *Department of Chemistry, University of Sheffield, S3 7HF, UK*



Biographical details

Mike Ward studied in Cambridge for his BA and Ph.D. before a post-doc position in Strasbourg with Jean-Pierre Sauvage. His independent academic career started with a lectureship in Bristol from 1990, and he remained there for 13 years before moving to Sheffield as a Professor of Inorganic Chemistry in 2003. His research interests cover many aspects of transition-metal and lanthanide ion coordination and supramolecular chemistry.

Chris Hunter is the Herchel Smith Professor of Organic Chemistry at the University of Cambridge. After a BA and then Ph.D. in Cambridge, he started his academic career as a lecturer at the University of Otago in New Zealand and then worked at the University of Sheffield for 23 years before taking up his current post in 2014. He has research interests in various aspects of supramolecular and physical organic chemistry.

Nick Williams did his BA, Ph.D., and early post-doctoral work in Cambridge. In 1994 he went to McGill University in Montréal as a RS/NSERC Research Fellow, working with Prof. J. Chin. In 1996 he was appointed to a lectureship at the University of Sheffield, where he has remained. His interests are in physical organic chemistry, working on understanding reactivity, mechanism, and catalysis in a range of contexts from organic surfaces to biological catalysis. These insights have underpinned the creation and understanding of a variety of biomimetic and supramolecular systems.

Conspectus

We describe here a family of coordination cages with interesting structural, guest-binding and catalytic properties. Flexible bridging ligands containing two bidentate pyrazolyl-pyridine termini assemble with transition metal dications to afford coordination cages containing a metal ion at each vertex, a bridging ligand spanning each edge, and a 2:3 metal:ligand ratio, expressed in structures from M_4L_6 tetrahedra to $M_{16}L_{24}$ tetra-capped truncated tetrahedra stabilised by formation of π-stacked arrays between electron-rich and electron-poor ligand segments which form around the cage periphery. In some cases concentration and/or temperature-dependent equilibria between multiple cage structures occur.

The cages are hollow and can accommodate guests (often anions or solvent molecules) in the central cavity. For one cage family, M_8L_{12} species with an approximately cubic structure and a *ca.* 400 $Å^3$ cavity, the guest binding properties have been studied extensively. This cage can accommodate a wide range of neutral organic guests, with binding in water being driven principally by the hydrophobic effect which leads to binding constants of up to 10^8 M^{-1}. The accumulation of a large amount of empirical data on guest binding in the M_8L_{12} cage in water provided the basis of a predictive tool for *in silico* screening of potential guests using the molecular docking programme GOLD; this methodology has allowed identification of numerous new guests with accurately predicted binding constants and provides a transformative new approach to exploring the host/guest chemistry of cages.

Binding of benzisoxazole inside the M_8L_{12} cage results in substantial rate enhancements – by a factor of up to 2 x 10^5 – of the Kemp elimination in which benzisoxazole reacts to give 2-cyanophenolate. Catalysis arises because the 16+ cage cation accumulates anions around the surface by ion-pairing, leading to a high effective concentration of hydroxide ions surrounding the guest even when the bulk pH is modest. Thus the catalysis relies on operation of two orthogonal interactions which bring the reaction partners together: hydrophobic guest binding in the cavity, and ion pairing around the cage surface. A consequence of this is that under some conditions the product of the cage-catalysed Kemp elimination (the 2-cyanophenolate anion) accumulates around the cage surface and deprotonates another benzisoxazole guest, perpetuating the reaction in an autocatalytic manner. Thus different anions accumulating around the cage can act as reaction partners, opening up the possibility of the M_8L_{12} cage being a general catalyst for reactions of electrophilic guests with surface-bound anions.

Introduction

The popularity of metal-organic polyhedral cages as a major research field has two distinct underpinnings.[1,2] The first is structural: such cages represent some of the most structurally elaborate and attractive edifices available using self-assembly methods, and our emerging understanding of the underlying principles of symmetry and structural control allow preparation of species that were, until recently, unimaginable.[1] The second is functional, and largely driven by host-guest chemistry. The central cavities of many coordination cages can accommodate guest species in an environment different from that in bulk solution, leading to potential applications in areas from transport across cell membranes to catalysis of reactions inside cage cavities.[2]

This review summarises work spanning 20 years on the synthesis and properties of a family of coordination cages; what started as an open-ended exploration of their structures has evolved into a directed exploration of their remarkable properties. In particular a detailed, quantitative understanding of the host-guest chemistry of one cage type has culminated in demonstration of catalytic and autocatalytic reactions controlled by the central cavity, and the discovery of a catalysis mechanism that may have a high degree of generality.

Syntheses and structures

The ligands (Scheme 1) contain two bidentate chelating pyrazolyl-pyridine units connected to an aromatic spacer *via* methylene units. Combination of these with labile transition-metal dications afforded a family of cages (Fig. 1, ranging from $[M_4(L^{12Ph})_6]^{8+}$ tetrahedra (the earliest examples) to $[M_{16}(L^{14Ph})_{24}]^{32+}$ tetra-capped truncated tetrahedra (the largest examples identified so far).[3] Simple changes to the structure or substitution pattern of the aromatic spacer have a dramatic effect on the structure of the resultant assembly, and there is no obvious way to predict which cage structure should arise from which ligand. However some underlying themes are clear.

(i) All cages have a 2M:3L ratio of components in with a metal ion at each vertex of a polyhedron, and a ditopic bridging ligand spanning every edge. This follows from simple considerations of coordination number: each ligand provides four donor atoms and each metal ion requires six, necessitating a 2:3 metal:ligand ratio for coordinative saturation. This is expressed in the set of cage architectures[3] which includes tetrahedra (4 vertices : 6 edges),[4,5] trigonal prisms (6:9),[6] cubes (8:12),[7] truncated tetrahedra (12:18)[8] and tetra-capped truncated tetrahedra (16:24).[9]

(ii) The unpredictability of the structures is a consequence of ligand flexibility. In contrast to many well-known cage types which are based on rigid edge-bridging ligands with predictable bend angles,[10] or rigid face-capping ligands which form triangular panels,[11] our inclusion of conformationally flexible methylene spacers means that predictability in the resulting cage structure is lost: but this is compensated for by allowing the ligands to determine the optimal assembly type. Rigid ligands are ideal from a rational design perspective, but the likelihood of a happy accident is reduced: each ligand type can adjust its conformation to optimise both metal-ligand coordinate bonding and inter-ligand stacking interactions.

(iii) The tris-chelate metal vertices can exist as either *fac* or *mer* geometric isomers, providing another degree of structural flexibility. In the smallest M_4L_6 tetrahedra (Fig. 1) all four vertices are *fac* tris-chelates;[4] conversely, in the $M_{12}L_{18}$ truncated tetrahedra, all twelve vertices are *mer* tris-chelates.[8] Other cages contain a mixture of both types of geometric isomer at different positions in the cage. Importantly the *fac* tris-chelate vertices provide an inwardly-directed, convergent set of C-H bonds in a region of high positive electrostatic potential close to a metal dication, generating a hydrogen-bond donor site to interact with electron-rich guests (Fig. 2).[12,13]

(iv) All of the cages incorporate π-stacked arrays around their periphery (Fig. 3), based on alternating stacks of electron-deficient (pyrazolyl-pyridine chelates coordinated to metal dications) and electron-rich units (aromatic spacers). The incorporation of both electron-deficient and electron-rich domains in the ligands arose from the synthetic methodology, but has turned out to be fundamentally important.[3] Ligand flexibility is also important in allowing the π-systems to fold and intertwine so as to maximise stacking interactions between donor and acceptor domains. The $[M_{16}L_{24}]^{32+}$ cage (Fig. 3) contains twelve five-component stacks around the periphery,[9b] affording a total of 48 charge-assisted pairwise interactions. The resulting stabilisation from the stacking interactions helps to overcome the entropy penalty associated with forming such large cages from flexible components. This behaviour is notably different to the assembly of cages based on Pd(II) and Pt(II) ions with rigid ligands, in which optimisation of metal/ligand bonding around d^8 ions with a strong stereoelectronic preference for square planar coordination provides the main driving force for cage assembly.[1]

The ligands' flexibility can allow multiple 'solutions' to a given self-assembly 'problem'. Combination of L^{14Ph} (Scheme 1) with Cd(II) ions afforded the 16-nuclear cage $[Cd_{16}(L^{14Ph})_{24}]^{32+}$. However redissolving crystals of this cage in MeCN gives slow (weeks)

rearrangement to the simpler trigonal prismatic $[Cd_6(L^{14Ph})_9]^{12+}$ cage (Fig. 4);[9a] the structures are in equilibrium (eq. 1) with the position being highly concentration-dependent. Clearly entropy considerations will always favour fragmentation into the Cd_6 species, with the opposing driving force that permits formation of the larger Cd_{16} cage being enthalpy effects associated with aromatic stacking, and the removal of a greater proportion of the ligands' solvophobic surface from solvent as aggregation occurs.

Similarly, combination of L^{18NapW} (Scheme 1) with Co(II) salts in water affords an equilibrium mixture of Co_2L_3 triple helicate, Co_4L_6 tetrahedral cage and $Co_{12}L_{18}$ truncated tetrahedral cage (Fig. 5).[14] Interconversion is slow (days); again the balance between components adjusts to changes in concentration and temperature, with high temperature or low concentration favouring the Co_2L_3 form, and low temperature or high concentration favouring the $Co_{12}L_{18}$ form.[14]

Host guest chemistry and the basis of guest binding.

We have studied in detail the guest binding properties of the octanuclear cubic $[Co_8L_{12}]^{16+}$ cages (Fig. 1, centre right), in which L is either L^{15naph},[7a] or the analogue $L^{15naphW}$ bearing hydroxymethyl substituents (Scheme 1).[15] In the former case the resultant $[Co_8(L^{15naph})_{12}]^{16+}$ cage (denoted **H** for host) is soluble in polar organic solvents.[7a] In the latter case, $[Co_8(L^{15naphW})_{12}]^{16+}$ ($\mathbf{H^w}$) bears twenty-four externally-directed hydroxyl substituents and is water-soluble.[15] We focussed on this cage type because of its stability in solution (there is no evidence for any structural rearrangement) and cavity size (large enough to accommodate a range of guests). Whilst many smaller cages are robust in solution, their cavities are too small to bind guests beyond counter-ions or solvents; conversely, many of the larger cages exist in solution in equilibrium with other forms, as described above, precluding any meaningful study of guest binding. **H** and $\mathbf{H^w}$ provide, from this family of cages, a sweet spot of stability and cavity size.

The two cages are essentially isostructural, with the central cavity volume being *ca.* 400 Å^3. On the basis of the Rebek '55% rule' [16] we expected optimally-sized guests to have molecular volumes of *ca.* 220 Å^3. In MeCN, **H** binds a range of bicyclic guests such as coumarin and isoquinoline-*N*-oxide (Fig. 6), driven in part by hydrogen-bonding interactions between the electron-rich O atom of the guest, and the hydrogen-bond donor site on the cage interior surface which provides a net H-bonding interaction comparable in strength to phenolic OH.[12] With guests of this nature we observed modest binding constants in MeCN in the range $10^1 - 10^3$ M^{-1} depending on the hydrogen-bond acceptor properties of the guest.[12]

In water however, using $\mathbf{H^w}$, we observed much stronger binding for the same guests[15,17] because the basis for binding is different. In water, weak cage/guest hydrogen bonds actually contribute negatively to guest binding, because the polar region of any guest will always form stronger interactions with water than with the cage interior surface. Instead, guest binding is driven by the hydrophobic effect: the tendency of hydrophobic regions of both guest and cage interior surface to associate and, in doing so, liberate water molecules into the bulk solution. The effect is substantial and proportional to the hydrophobic surface area of the bound guest.[18] For example K for coumarin binding inside **H** in MeCN is 80 M^{-1}; for binding inside $\mathbf{H^w}$ in water it is 7600 M^{-1}.[15]

The clearest demonstration of this is provided by a series of cyclic ketone guests from cyclopentanone to cycloundecanone.[17] All contain the same carbonyl group which forms hydrogen bonds to the cage interior surface, but the stepwise increase in the hydrophobic surface area associated with the increasing number of CH_2 groups afforded a linear increase in the ΔG value for guest binding in $\mathbf{H^w}$ of 5 kJ mol^{-1} per CH_2 group across the series of 7 guests in water (Fig. 7). This value is reassuringly consistent with previous correlations of the magnitude of the hydrophobic effect with burial of non-polar surface area obtained both from biological and artificial systems. As the series of guests expands further to cyclododecanone and cyclotridecanone the binding strength decreases as the guests are now above Rebek's 55% limit and steric limitations become apparent. In this series the strongest binding guest, cycloundecanone, has $K = 1.2 \times 10^6$ M^{-1} ($\Delta G_{binding} = -35$ kJ mol^{-1}). This value can be quantitatively dissected: (i) the ten non-polar CH_2 groups provide *ca.* –48 kJ mol^{-1} of binding energy arising from the hydrophobic effect; (ii) the cost of combining two species into one when a bimolecular complex forms is 6 kJ mol^{-1};[19] and (iii) replacement of strong hydrogen bonds of the polar carbonyl group to water by weaker hydrogen bonds to the cage interior surface costs *ca.* 7 kJ mol^{-1} (–35 = –48 + 6 + 7).[17]

We likewise observed substantial increases in binding constants of guests in water when an additional aromatic ring was added to the structure. Double mutant cycle experiments – in which other effects such as van der Waals' interactions and steric issues are factored out – showed that the magnitude of the additional contribution to guest binding precisely matched what would be expected the basis of the additional hydrophobic surface area.[15]

Importantly, the requirement for strongly-binding guests to be hydrophobic means that charged species – inherently more hydrophilic – bind more weakly, irrespective of the sign of the charge. This is important for ionisable guests such as amines, which can protonate to give

hydrophilic cations; or carboxylic acids, which can deprotonate to give hydrophilic anions. Guests of this nature therefore bind in their neutral / hydrophobic form, but a pH swing can be used to make the guest charged, at which point the value of K decreases by 2 – 3 orders of magnitude, and the guest exits the cavity; this was demonstrated with a range of guests having pK_a values from 3.5 to 11 (Fig. 8).[20] This issue becomes important in the context of cage-based catalysis (see later).

We probed further the different contributions to the hydrophobic effect in $\mathbf{H^w}$/guest complexes in water using NMR measurements of guest binding as a function of temperature, allowing us to separate (using Van't Hoff plots) the enthalpy and entropy contributions to binding using pairs of guests differing only in the presence / absence of a single CH_2 group (*eg.* cycloheptanone / cyclooctanone).[21] We consistently observed that, when several such pairs of guests were compared, the 5 kJ mol^{-1} contribution to the free energy change of complexation per CH_2 group[17] was almost completely enthalpy-based. This contradicts the traditional picture of the hydrophobic effect as being predominantly entropy-based, arising from liberation of water molecules in contact with hydrophobic surfaces. In the confined cavity of a synthetic host like $\mathbf{H^w}$, the curvature of the interior surface means that the array of bound water molecules is unable to form as many hydrogen-bonds to other water molecules as would normally happen in bulk solution. This can be seen in a crystal structure of hydrated $\mathbf{H^w}$ in which each water molecule in the cavity-bound $(H_2O)_{10}$ cluster forms (on average) only 3.2 hydrogen-bonds to other water molecules, compared to an average value of 3.7 in bulk solution (Fig. 9).[21] Thus the cluster of ten waters is 'frustrated' to the extent of five strong hydrogen-bonds which can be regained when the bound water molecules are liberated by guest binding, providing a substantial enthalpy boost.[22]

In silico screening of guest binding

Despite the large number of quantitative studies of guest binding inside cages, Rebek's work leading to the 55% rule is a unique example of a systematic analysis that has been predictively useful in other systems.[16] Given that, in the course of our work, we had accumulated binding constant data for many different types of guest inside $\mathbf{H^w}$ in water, we were interested to see if we could use these data to develop a quantitative prediction tool.

We used the software package GOLD, developed as a protein/ligand docking tool for drug discovery applications.[23] GOLD takes the three-dimensional structure of a protein, positions a potential ligand at the binding site and adjusts the position and conformation to optimise binding. Empirical parameters for contributions from hydrophobic contacts,

hydrogen-bonding, electrostatic interactions, the enthalpy and entropy costs of conformational changes in the guest, are summed to produce a binding score that is used to rank a list of potential ligands. We used the structure of the cage cation of $\mathbf{H^w}$ as the binding site in GOLD to score the set of guests whose binding constants were already known.[24]

Initially the correlation between predicted and observed binding was poor, but this was based on GOLD's default scoring function – a set of assumptions about the relative importance of the different contributions to binding. Whilst GOLD can estimate contributions from (for example) hydrophobic effects, restrictions of conformational flexibility, and hydrogen-bonding, the way in which these are combined – in particular their relative weightings in the scoring function – are at the discretion of the user. We therefore allowed the weightings of the different contributions to the scoring function to vary freely, and refined the whole set of weighting coefficients until the calculated score for our training set of known guests matched the observed binding constants as closely as possible.

The resultant scoring function worked well in predicting binding constants of new guests. A virtual screening experiment was used to rank a library of 3000 potential guests, and the top 13 were investigated experimentally. The agreement between calculated and measured binding constants for this set of new guests was excellent (Fig. 10) and underlines the value of this methodology in putting identification of new guests for synthetic hosts on a predictive footing.[24] The highest affinity guest for $\mathbf{H^w}$ that we had found previously (cycloundecanone)[17] came from two years of trial-and-error measurements; following the *in silico* screening process using GOLD, we identified three higher affinity guests (*K* values up to 10^8 M^{-1}) in a few days. A improved scoring function was developed using experimental data for flexible guests (aliphatic ketones with varying degree of conformational freedom) in an expanded training set to take account of the enthalpy and entropy penalties associated with conformational restriction of guests when they bind. This scoring function predicts well binding constants for any guest.[25]

Structures of cage / guest complexes

Crystallising either **H** or $\mathbf{H^w}$ from solvent in the presence of a guest usually resulted in crystals of the empty cage. However we could successfully exploit Fujita's 'crystalline sponge' method:[26] pre-formed crystals of either cage are immersed either in pure guest (if the desired guest is an oil), or in a concentrated methanolic solution of the guest, for a few hours. This often results in uptake of guest into the cage cavity in the intact crystals, allowing the complexes to be structurally characterised. Some examples are in Fig. 11 and include cyclic[17]

and open-chain aliphatic ketones,[25] alkyl phosphonates[27] (used as simulants for phosphonofluoridate chemical warfare agents), and rigid aromatic species such as benzisoxazole.[28] Note that we are not using the 'crystalline sponge' method to determine structures of unknown guests: typical *R*1 values for these systems are 10 - 15% due to the weak scattering associated with the large unit cell and disorder of anions and solvent molecules. The relatively simple guests we used have known structures which are highly restrained during refinement, with the sole aim being to identify the position and orientation of the guest in the cavity. In every case we observe the electron-rich regions of the guests (*e.g.* the C=O or P=O oxygen atoms) forming hydrogen bonds with a CH hydrogen-bond donor pocket on the cage interior surface. In some cases, when the guests are small, we can see two guest molecules binding (Fig. 11b).

We also note that in all crystal structures of **H** or $\mathbf{H^w}$, the portals in the cage faces are occupied by anions which participate in an array of CH•••X hydrogen-bonding interactions with the ligands (Fig. 12). It is well known that the cavity in the centre of a preorganised ligand array in a cyclic helicate can bind anions very strongly,[31] and this is an extension of that principle. The accumulation of anions around the cage surface, surrounding the cavity-bound guests, turns out to be of fundamental importance for the cage-based catalysis discussed next.

Catalysis of the Kemp elimination in the cage cavity

A guest identified by our *in silico* screening process was benzisoxazole,[24] which binds inside $\mathbf{H^w}$ in water with $K \approx 4000$ M^{-1}. Benzisoxazole has been known since the 1970s to react with base in a ring opening E2 elimination that generates 2-cyanophenolate (Scheme 2, commonly known as the Kemp elimination[32]); the reaction is first order in hydroxide over a wide pH range. The Kemp elimination has been widely studied to probe effects of medium on reactivity, and as such we were interested to investigate how this reaction is affected by binding of the substrate in the cavity of $\mathbf{H^w}$.

On comparing the rates of conversion of benzisoxazole to 2-cyanophenolate in NMR experiments in D_2O / DO^-, in the absence and presence of $\mathbf{H^w}$ and buffered at a range of pD values, two separate effects emerged.[28] Firstly the cage-catalysed reaction was substantially accelerated compared to the background reaction, with $k_{cat}/k_{uncat} = 4500$ at pD 10.2. Secondly, and in contrast to the uncatalysed reaction which is first order in hydroxide, the rate of the catalysed reaction was independent of [DO^-] over the pD range of 8.5 to 11.4 – nearly a thousandfold change in [DO^-]. These results are summarised in Fig. 13. It is clear from this

that the catalytic rate enhancement – the value of k_{cat}/k_{uncat} – steadily increases as the pD is reduced, with a maximum measured k_{cat}/k_{uncat} value of 2 x 10^5 at pD 8.5. The point at which the two lines intersect (at pD 13.8) is where the catalysed and uncatalysed reactions proceed with the same rate: *i.e.* the cage-catalysed reaction at pD 8.5 proceeds at the same rate as the background reaction at pD 13.8.

This effectively means that the benzisoxazole in the cavity, when the pD of the bulk solution is 8.5, reacts as if its local environment were at pD 13.8 (corresponding to a local [DO^-] of 0.1M, since the pK_a of D_2O is 14.88). This suggests the basis for the catalysis. We can see from crystal structures that the cationic cage accumulates anions in the surface portals, and we propose that something similar happens in solution in a manner analogous to formation of the Stern layer of counter-ions around the surface of charged micelles.[33] From crystal structures of $\mathbf{H^w}$ the volume of one complete cage (plus anions) is *ca.* 11000 Å^3;[15,28] the presence of one DO^- anion in each of the six portals would provide a local concentration of partly desolvated DO^- anions corresponding to *ca.* 1800 Å^3 per anion (about 1M). This is reasonably consistent with the observed catalysed reaction rate which suggests a local [DO^-] concentration around the guest of 0.1M. The modest discrepancy, with the catalysed reaction proceeding slightly more slowly than the high local concentration of DO^- suggests, is consistent with the fact that the hydrophobic cage interior provides an environment that is less good at stabilising the developing negative charge of the cyanophenolate anion than is water – which is the basis for ejection of the anionic product and catalytic turnover (see below). Thus the effect of the high local concentration of DO^- ions around the cavity (≈ 1M) is slightly offset by the poorer ability of the cavity to stabilise the developing negative charge as the reaction proceeds. This behaviour is similar to that observed for micelles, where catalysis of bimolecular reactions is explained by an increased local concentration of the reagents rather than greater reactivity of the reactants within the micelle environment.[34,35]

This accumulation of anions around the cage surface also explains the pH independence of the catalysed reaction: if the cage surface is saturated with DO^- anions at pD 8.5, increasing further the concentration of DO^- ions in the bulk solution cannot result in any significant increase in the local [DO^-] around the substrate.[28] A more quantitative description of this behaviour requires consideration of the local ion concentrations and might be achieved by considering ion exchange models or solving the Poisson-Boltzmann equation to determine the ionic surroundings of the cage.[35]

Two control experiments confirmed this mechanism (Fig. 13). Firstly, in the presence of cycloundecanone (which is inert but binds more tightly, blocking the cage cavity) the

catalysed reaction was reduced to background levels – proving that catalysis only occurs when benzisoxazole is inside the cage. Secondly, addition of chloride ions also reduced the reaction rate to background levels. Chloride is more easily desolvated than hydroxide and therefore accumulates more readily around cationic surfaces in water,[36] displacing hydroxide. Typically, the ion exchange constant between chloride and hydroxide is around 10 for cationic micelles,[36] illustrating the preference for ion pairing to occur with the less densely charged chloride ions. The presence of chloride ions does not affect the rate of the Kemp elimination in the absence of the cage, so their ability to switch off the cage-catalysed reaction confirms that the basis of the catalysis is the accumulation of hydroxide ions around the cage surface.

The overall catalytic cycle is shown in Fig. 14 and has two noteworthy features. Firstly, cage catalysis is based on two orthogonal interactions which bring the two reacting partners together: the substrate binds inside the cavity principally because of the hydrophobic effect (a non-polar interaction); hydroxide ions accumulate around the bound guest because of ion-pairing (a polar interaction). We know that the cage can bind a wide variety of different substrates, and the ability of chloride to displace hydroxide from the portals around the cage surface also suggests that we can also control which type of anion surrounds the substrate. The cage therefore provides a platform for surrounding any hydrophobic substrate in the cavity with a high local concentration of any anions occupying the surface sites, leading to the possibility of a versatile catalytic system. Further work towards this goal is described in the next section.

Secondly, efficient catalysis requires turnover – the product must be expelled from the cavity to make way for a new substrate – a challenge that has been a major stumbling block in the development of supramolecular catalysts and model enzymes. In $\mathbf{H^w}$, turnover is achieved on the basis of charge. Whilst benzisoxazole is neutral and hydrophobic enough to bind in the cage cavity, 2-cyanophenolate is anionic under the conditions used (pK$_a$ = 7) and therefore rapidly exits the cavity to be solvated in the aqueous phase.[20] This prevents any product inhibition and ensures that the reaction continues for at least 100 cycles with no loss of catalytic activity.[28]

Binding of different anions around the cage; onset of autocatalytic behaviour.

The observation that chloride ions could displace hydroxide ions from around the cage surface and thereby inhibit the cage-catalysed Kemp elimination reaction[28] suggested a set of experiments in which the concentration of chloride was increased steadily; we would expect

that this would result in the steady shut-down of the cage-catalysed reaction as the amount of hydroxide around the cage decreases. This turned out to be the case but the behaviour that emerged was more complex than expected.[30]

Fig. 15 shows the reaction progress profiles for the cage-catalysed reaction in the presence of increasing amounts of chloride. For these experiments we used the cage **H** (unsubstituted exterior surface) but as its chloride salt: this is an easier way of making the cage water-soluble than the lengthy synthesis of the hydroxymethyl-substituted ligands used to prepare $\mathbf{H}^{\mathbf{w}}$. As the charge on the cage is 16+, we necessarily started with 16 equivalents of chloride per $\mathbf{H}^{\mathbf{w}}$ cation and under these conditions the catalysed reaction is already substantially slowed down. As additional chloride was added, not only did the reaction slow down further, but the shape of the curve – accumulation of 2-cyanophenolate product with time – became sigmoidal. A similar effect was seen when fluoride was added, although more fluoride was needed compared to chloride to have a similar inhibiting effect, consistent with stronger solvation in water and therefore weaker interaction with the cage surface.

The emergence of a sigmoidal reaction profile in the presence of excess chloride or fluoride as the reaction slows down indicates the onset of autocatalytic behaviour, in which the product of the Kemp elimination (the 2-cyanophenolate anion) catalyses its own formation. The sigmoidal shape occurs because early in the reaction there is little product to catalyse further reaction so it starts slowly; as the product accumulates and catalysis increases the reaction accelerates, until substrate is consumed and the rate slows again. Crucially, in order to act as a base to deprotonate cage-bound benzisoxazole, the 2-cyanophenolate ions must themselves accumulate around the portals in the cage surface, where their basicity allows them to deprotonate a molecule of bound benzisoxazole in a way that halides cannot. The mechanism is schematically illustrated in Fig. 16. In the absence of cage the accumulation of phenolate anions has no effect on progress of the Kemp elimination, and the normal hydroxide-based reaction pathway dominates to give accurate first order behaviour at constant pH. We can therefore see how the progress of the cage-catalysed Kemp elimination depends on the nature of the anion that accumulates around the cage surface. In the presence of hydroxide we see rate accelerations of $>10^5$ fold because of the high local hydroxide concentration. In the presence of chloride or fluoride the reaction is slowed because the halide replaces hydroxide around the cage surface, and the halides are too poorly basic to react with benzisoxazole. In the domain where the normal hydroxide-based reaction has been almost completely stopped by the halide ions, an autocatalytic pathway emerges in which the cyanophenolate anions can in turn displace halide ions from around the cage, and they are

basic enough to propagate the reaction. Indeed, the induction period can be removed by adding the 2-cyanophenolate at the start of the reaction, and a similar effect can be observed by adding a range of phenolates with comparable pK_a values.[30]

Thus, different anions can be induced to be reaction partners with cavity-bound guests depending on the conditions. Given the range of guests that we have identified, this suggests the possibility of the cage acting as a general catalyst for bimolecular reactions of cavity-bound hydrophobic guests with surface-bound anions in water: studies to this end are in progress.

Acknowledgements

We thank EPSRC and the Leverhulme Trust for funding; and particularly thank our many talented co-workers whose names appear in the reference list.

References

1 (a) Chakrabarty, R.; Mukherjee, P. A.; Stang, P. J. *Chem. Rev.* **2011**, *111*, 6810. (b) Cook, T. R.; Stang, P. J. *Chem. Rev.* **2015**, *115*, 7001.

2 (a) Ward, M. D.; Raithby, P. R. *Chem. Soc. Rev.* **2013**, *42*, 1619. (b) Brown, C. J.; Toste, F. D.; Bergman, R. G.; Raymond, K. N. *Chem. Rev.* **2015**, *115*, 3012. (c) Yoshizawa, M.; Klosterman, J. K.; Fujita, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 3418. (d) Zarra, S.; Wood, D. M.; Roberts, D. A.; Nitschke, J. R. *Chem. Soc. Rev.* **2015**, *44*, 419.

3 Ward, M. D. *Chem. Commun.* **2009**, 4487.

4 Paul, R. L.; Bell, Z. R.; Jeffery, J. C.; McCleverty, J. A.; Ward, M. D. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4883.

5 Hall, B. R.; Manck, L. E.; Tidmarsh, I. S.; Stephenson, A.; Taylor, B. F.; Blaikie, E. J.; Vander Griend, D. A.; Ward, M. D. *Dalton Trans.* **2011**, *40*, 12132.

6 Najar, A. M.; Avci, C.; Ward, M. D. *Inorg. Chem. Commun.* **2012**, *15*, 126.

7 Tidmarsh, I. S.; Faust, T. B.; Adams, H.; Harding, L. P.; Russo, L.; Clegg, W.; Ward, M. D. *J. Am. Chem. Soc.* **2008**, *130*, 15167.

8 Argent, S. P.; Adams, H.; Riis-Johannessen, T.; Jeffery, J. C.; Harding, L. P.; Mamula, O.; Ward, M. D. *Inorg. Chem.* **2006**, *45*, 3905.

9 (a) Stephenson, A.; Argent, S. P.; Riis-Johannessen, T.; Tidmarsh, I. S.; Ward, M. D. *J. Am. Chem. Soc.* **2011**, *133*, 858. (b) Stephenson, A.; Sykes, D.; Ward, M. D. *Dalton Trans.* **2013**, *42*, 6756.

10 Sun, Q.-F.; Iwasa, J.; Ogawa, D.; Ishido, Y.; Sato, S.; Ozeki, T.; Sei, Y.; Yamaguchi, K.; Fujita, M. *Science* **2010**, *328*, 1144.

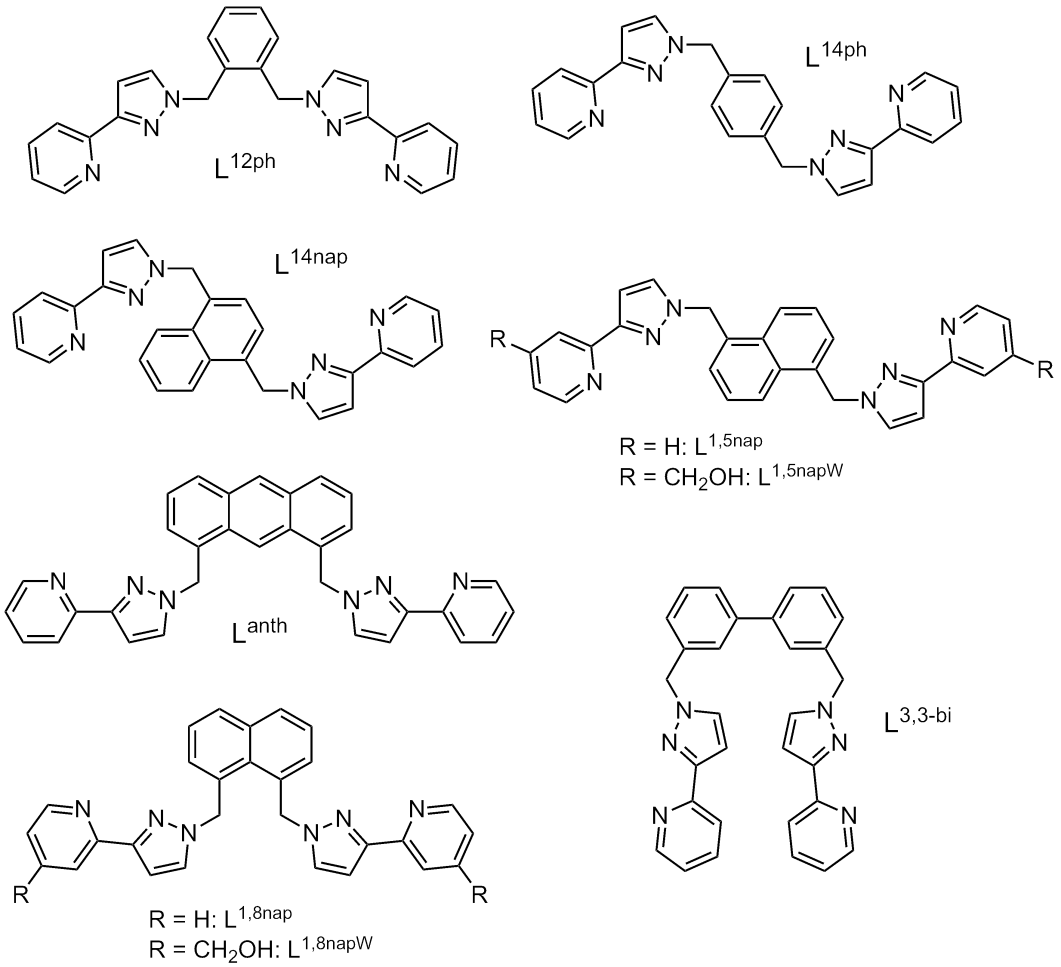
11 Fujita, M.; Umemoto, K.; Yoshizawa, M.; Fujita, N.; Kusukawa, T.; Biradha, K. *Chem. Commun.* **2001**, 509.

12 Turega, S.; Whitehead, M.; Hall, B. R.; Meijer, A. J. H. M.; Hunter, C. A.; Ward, M. D. *Inorg. Chem.* **2013**, *52*, 1122.

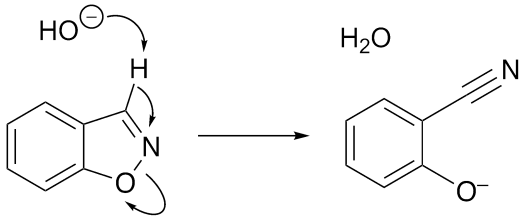
13 Metherell, A. J.; Ward, M. D. *Dalton Trans.* **2016**, *45*, 16096.

14 Cullen, W.; Hunter, C. A.; Ward, M. D. *Inorg. Chem.* **2015**, *54*, 2626.

15 Whitehead, M.; Turega, S.; Stephenson, A.; Hunter, C. A.; Ward, M. D. *Chem. Sci.* **2013**, *4*, 2744.
16 (a) Rebek, J. *Acc. Chem. Res.* **2009**, *42*, 1660. (b) Mecozzi, S.; Rebek, J. *Chem. Eur. J.* **1998**, *4*, 1016.
17 Turega, S.; Cullen, W.; Whitehead, M.; Hunter, C. A.; Ward, M. D. *J. Am. Chem. Soc.* **2014**, *136*, 8475.
18 (a) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210. (b) Hunter, C. A. *Chem. Sci.* **2013**, *4*, 834.
19 (a) Hunter, C. A. *Angew. Chem., Int. Ed.* **2004**, *43*, 5310. (b) Andrews, P. R.; Craik, D. J.; Martin, J. L. *J. Med. Chem.* **1984**, *27*, 1648.
20 (a) Cullen, W.; Turega, S.; Hunter, C. A.; Ward, M. D. *Chem. Sci.* **2015**, *6*, 625. (b) Cullen, W.; Thomas, K. A.; Hunter, C. A.; Ward, M. D. *Chem. Sci.* **2015**, *6*, 4025.
21 Metherell, A. J.; Cullen, W.; Williams, N. H.; Ward, M. D. *Chem. Eur. J.* **2018**, *24*, 1554.
22 (a) Biedermann, F.; Nau, W. M.; Schneider, H.-J. *Angew. Chem., Int. Ed.* **2014**, *53*, 11158. (b) Snyder, P. W.; Lockett, M. R.; Moustakas, D. T.; Whitesides, G. M. *Eur. Phys. J. Spec. Top.* **2014**, *223*, 853.
23 https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/
24 Cullen, W.; Turega, S.; Hunter, C. A.; Ward, M. D. *Chem. Sci.* **2015**, *6*, 2790.
25 Taylor, C. G. P.; Cullen, W.; Collier, O. M.; Ward, M. D. *Chem. Eur. J.* **2017**, *23*, 206.
26 (a) Hoshino, M.; Khutia, A.; Xing, H.; Inokuma, Y.; Fujita, M. *IUCrJ* **2016**, *3*, 139. (b) Inokuma, Y.; Yoshioka, S.; Ariyoshi, J.; Arai, T.; Hitora, Y.; Takada, K.; Matsunaga, S.; Rissanen, K.; Fujita, M. *Nature* **2013**, *495*, 461
27 Taylor, C. G. P.; Piper, J. R.; Ward, M. D. *Chem. Commun.* **2016**, *52,* 6225.
28 Cullen, W.; Misuraca, M. C.; Hunter, C. A.; Williams, N. H.; Ward, M. D. *Nat. Chem.* **2016**, *8*, 231.
29 Piper, J. R.; Cletheroe, L.; Taylor, C. G. P.; Metherell, A. J.; Weinstein, J. A.; Sazanovich, I. V.; Ward, M. D. *Chem. Commun.* **2017**, *53*, 408.
30 Cullen, W.; Metherell, A. J.; Wragg, A. B.; Taylor, C. G. P.; Williams, N. H.; Ward, M. D. *J. Am. Chem. Soc.* **2018**, *140*, 2821.
31 (a) Hasenknopf, B.; Lehn, J.-M.; Kneisel, B. O.; Baum, G.; Fenske, D. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1838. (b) Riddell, I. A.; Smulders, M. M. J.; Clegg, J. K.; Hristova, Y. R.; Breiner, B.; Thoburn, J. D.; Nitschke, J. R. *Nat. Chem.* **2012**, *4*, 751.
32 (a) Casey, M. L.; Kemp, D. S.; Paul, K. G.; Cox, D. D. *J. Org. Chem.* **1973**, *38*, 2294. (b) Kemp, D. S.; Casey, M. L. *J. Am. Chem. Soc.* **1973**, *95*, 6670.
33 Stigter, D. *J. Phys. Chem.* **1975**, *79*, 1008.
34 (a) Bunton, C. A.; Nome, F.; Quina, F. H.; Romsted, L. S. *Acc. Chem. Res.* **1991**, *24*, 357. (b) Quina, F. H. *J. Braz. Chem. Soc.* **2016**, *27*, 267.
35 (a) Bunton, C. A.; Moffatt, J. R. *J. Phys. Chem.* **1986**, *90*, 538. (b) Bunton, C. A.; Moffatt, J. R. *J. Phys. Chem.* **1988**, *92*, 2896.
36 Buurma, N. J. *Adv. Phys. Org. Chem.* **2009**, *43*, 1.



Scheme 1 Examples of the ligands used in this work.



Scheme 2 The Kemp elimination: reaction of benzisoxazole with base to give 2-cyanophenolate.

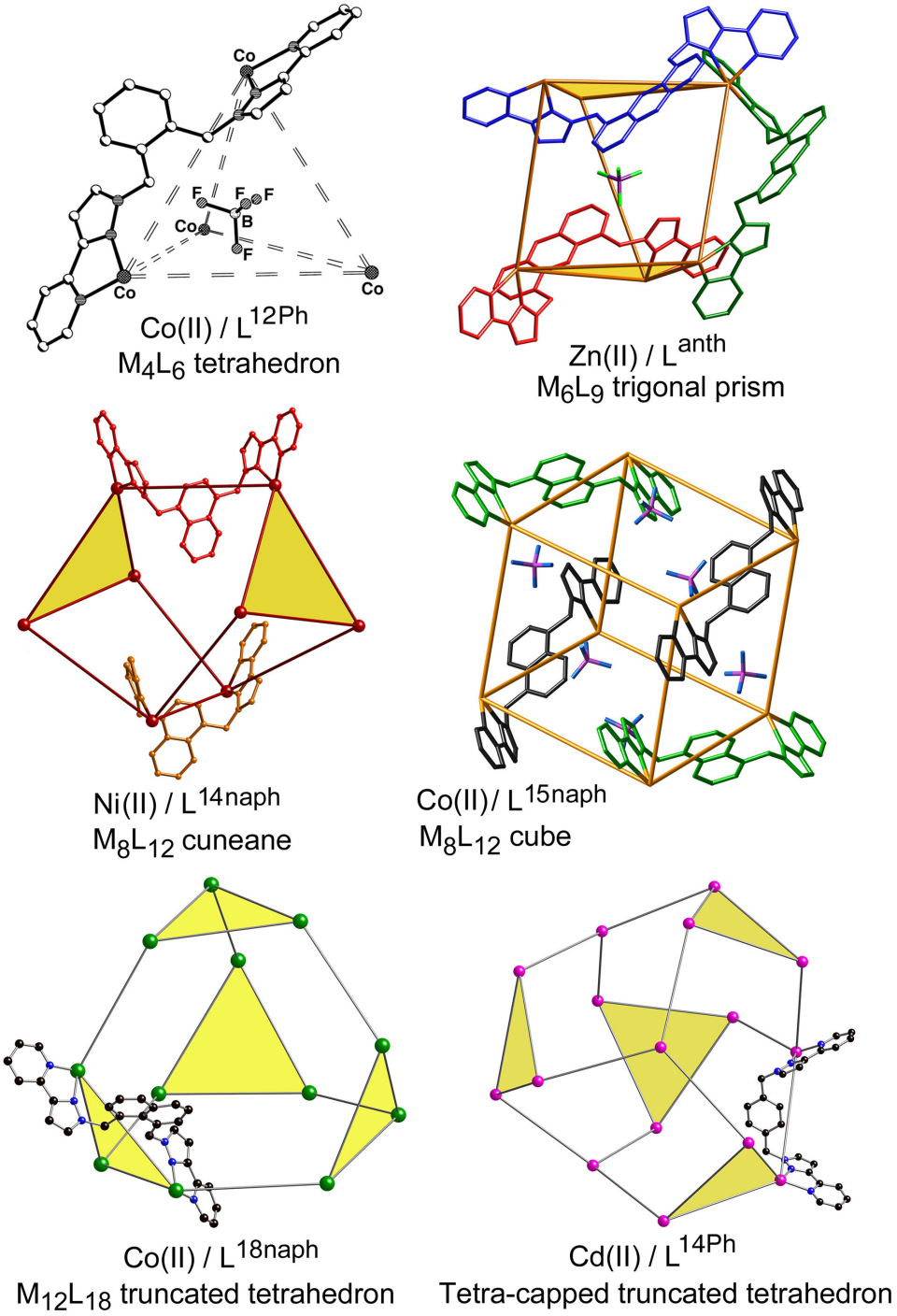


Figure 1 Examples of cage types arising from the different ligands assembling with M(II) ions.

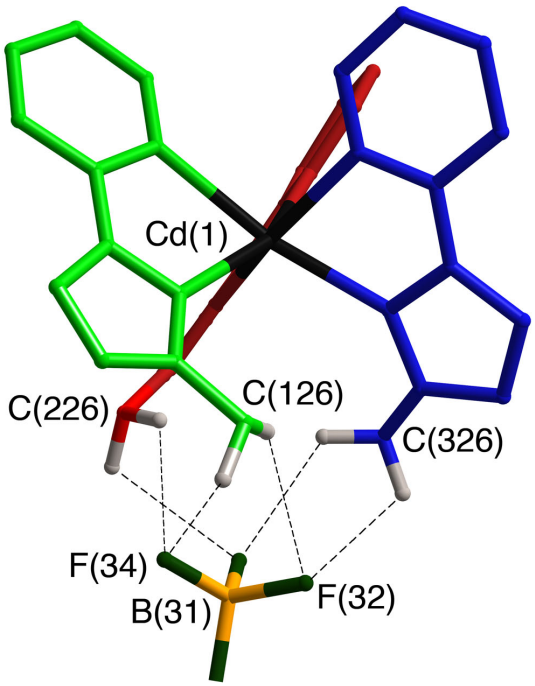


Figure 2 The *fac* tris-chelate Cd(II) vertex in the structure of $[Cd_4(L^{3,3\text{-biph}})_6](BF_4)_8$, showing how the convergent set of CH protons forms a hydrogen-bond donor pocket that interacts with electron-rich guests (ref. 13).

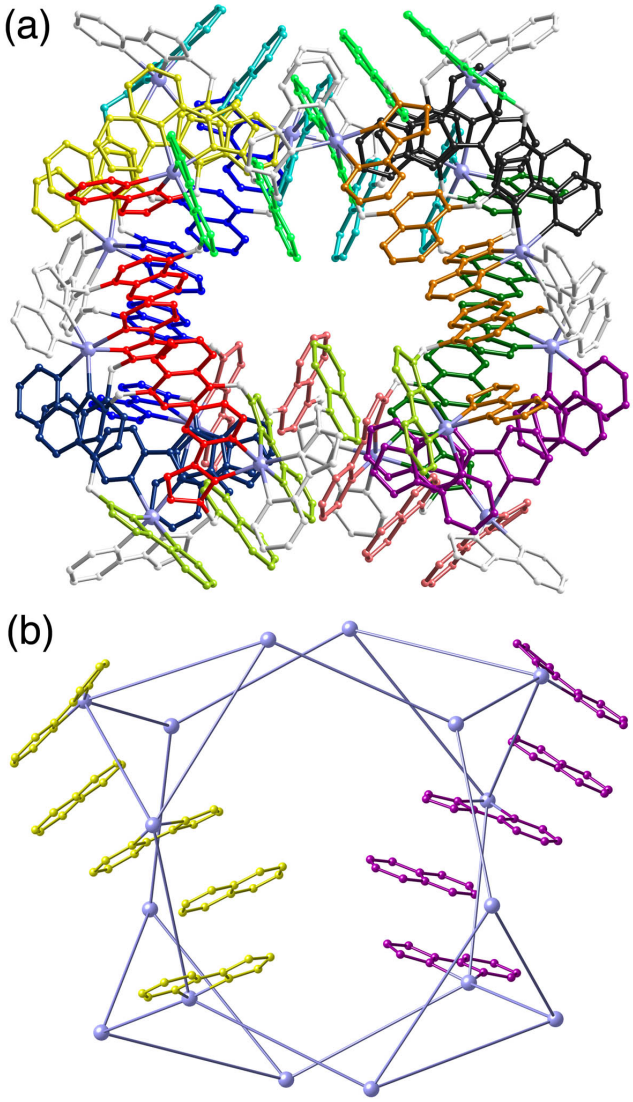


Figure 3 Two views of the cage cation of $[Cd_{16}(L^{14Nap})_{24}](BF_4)_{32}$ emphasising the inter-ligand π-stacking: (a) the complete cage with the ligand fragments involved in each five-layer stack shaded the same colour; (b) a simpler view showing two of the twelve stacks superimposed on the Cd_{16} core (ref. 9b).

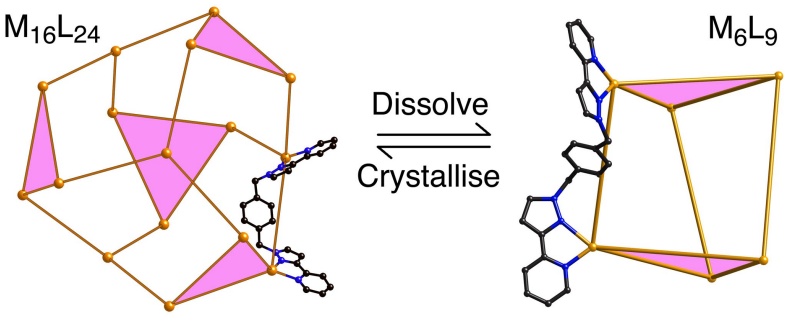


Figure 4 Slow rearrangement between $M_{16}L_{24}$ and M_6L_9 cages ($L = L^{14Ph}$) in MeCN (ref. 9a).

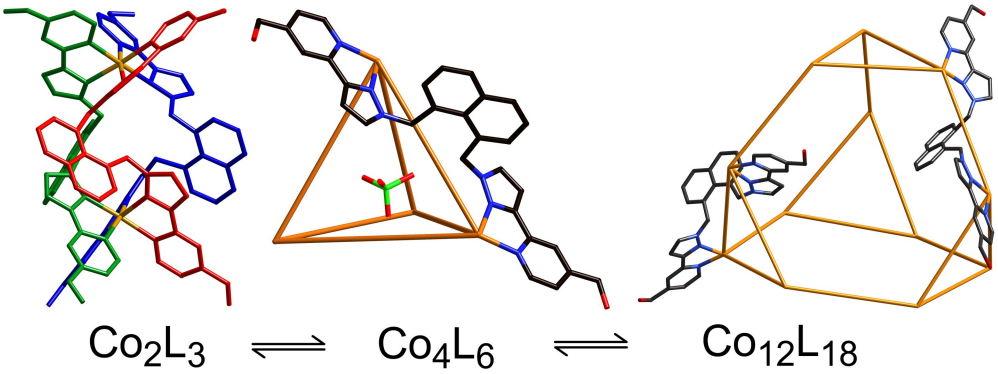


Figure 5 Equilibrium between M_2L_3, M_4L_6 and $M_{12}L_{18}$ species ($L = L^{18NaphW}$) in water (ref. 14).

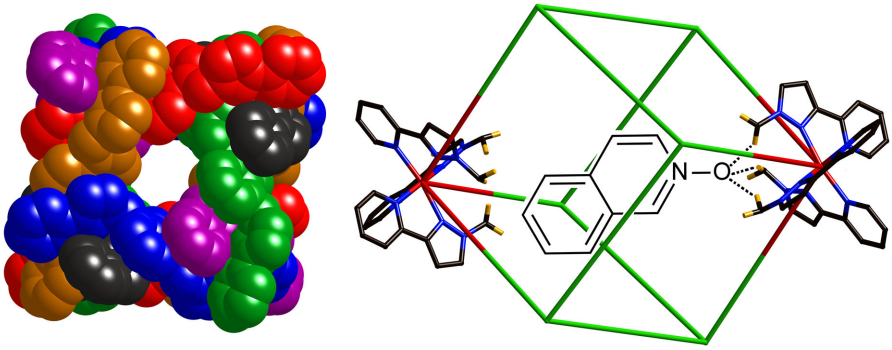


Figure 6 Binding of isoquinoline-*N*-oxide inside **H** in MeCN, illustrating the hydrogen-bonding interaction with the cage interior surface (ref. 12).

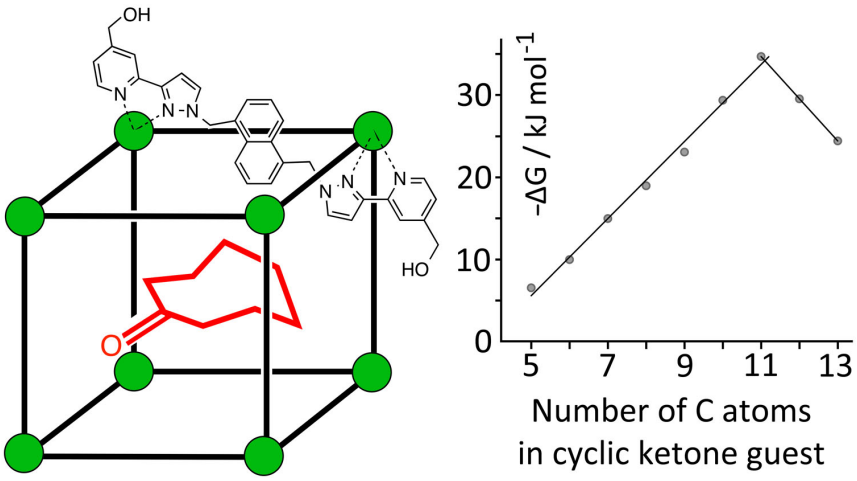


Figure 7 Binding free energies of a series of aliphatic cyclic ketones inside $\mathbf{H}^{w}$ in water (ref. 17).

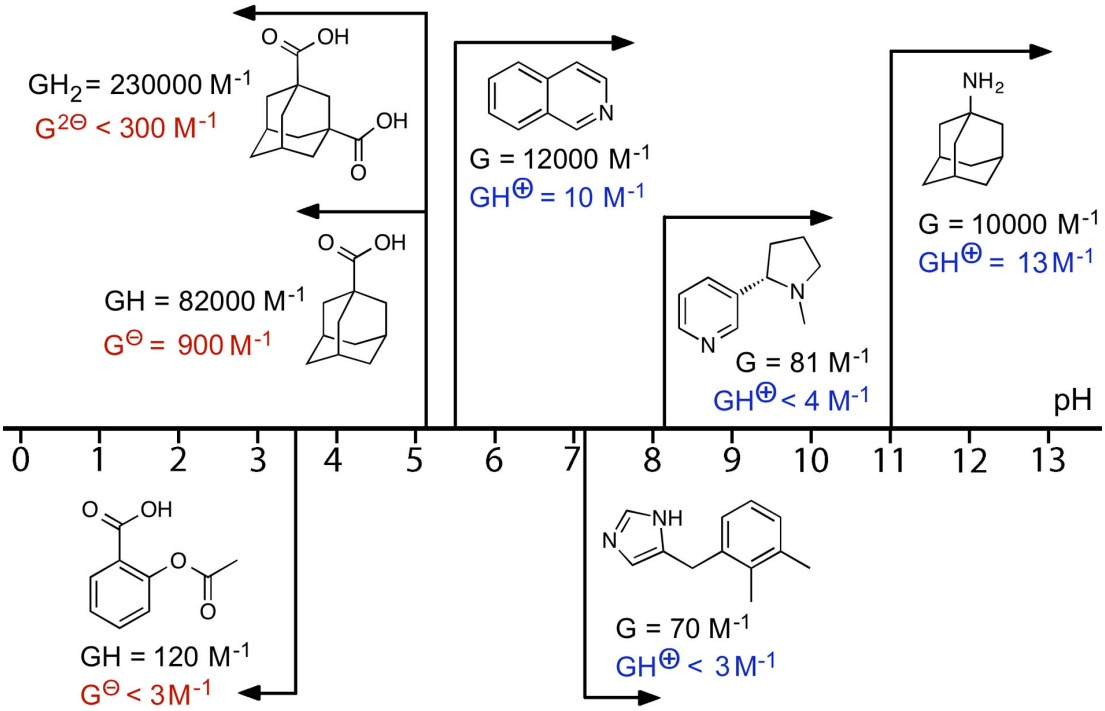


Figure 8 pH dependence of guest binding: summary of binding constants of guests in the cavity of $\mathbf{H}^{w}$ in water, their neutral and charged states; the direction of the arrow indicates the pH range in which the guest is neutral and stronger binding occurs (ref. 20a).

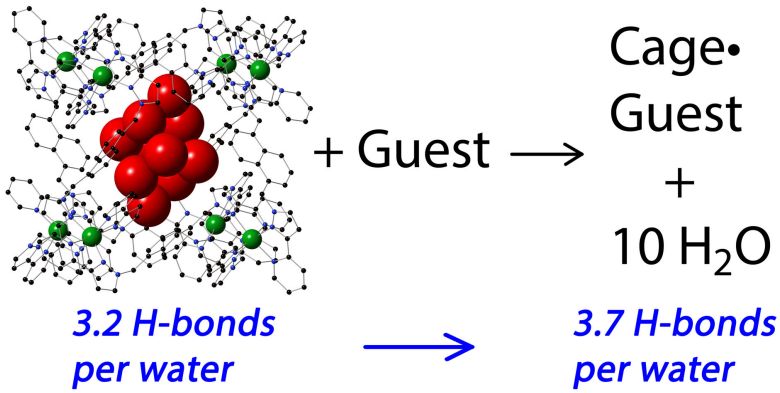


Figure 9 The origin of the substantial enthalpy contribution to the hydrophobic effect associated with binding of guests inside $\mathbf{H}^{\mathbf{w}}$ in water (ref. 21).

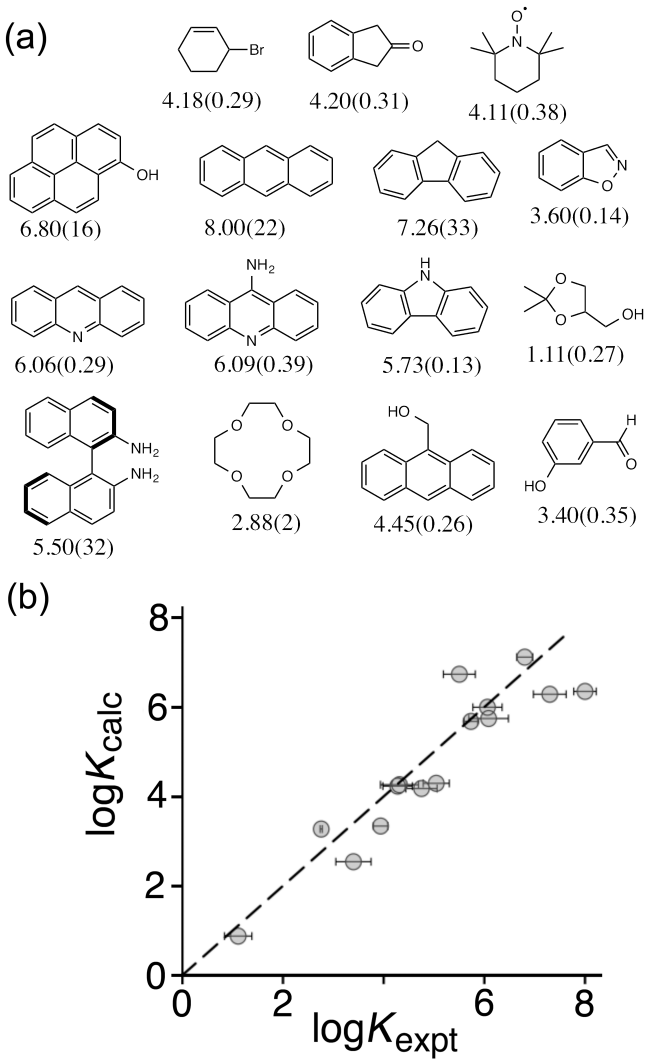


Figure 10 (a) A series of new guests identified by an *in silico* screen of a virtual library of thousands of compounds (numbers are log*K* values with esd's in parentheses); (b) the correlation of predicted (using GOLD) with observed binding constants for these guests in $\mathbf{H}^{\mathbf{w}}$ (ref. 24).

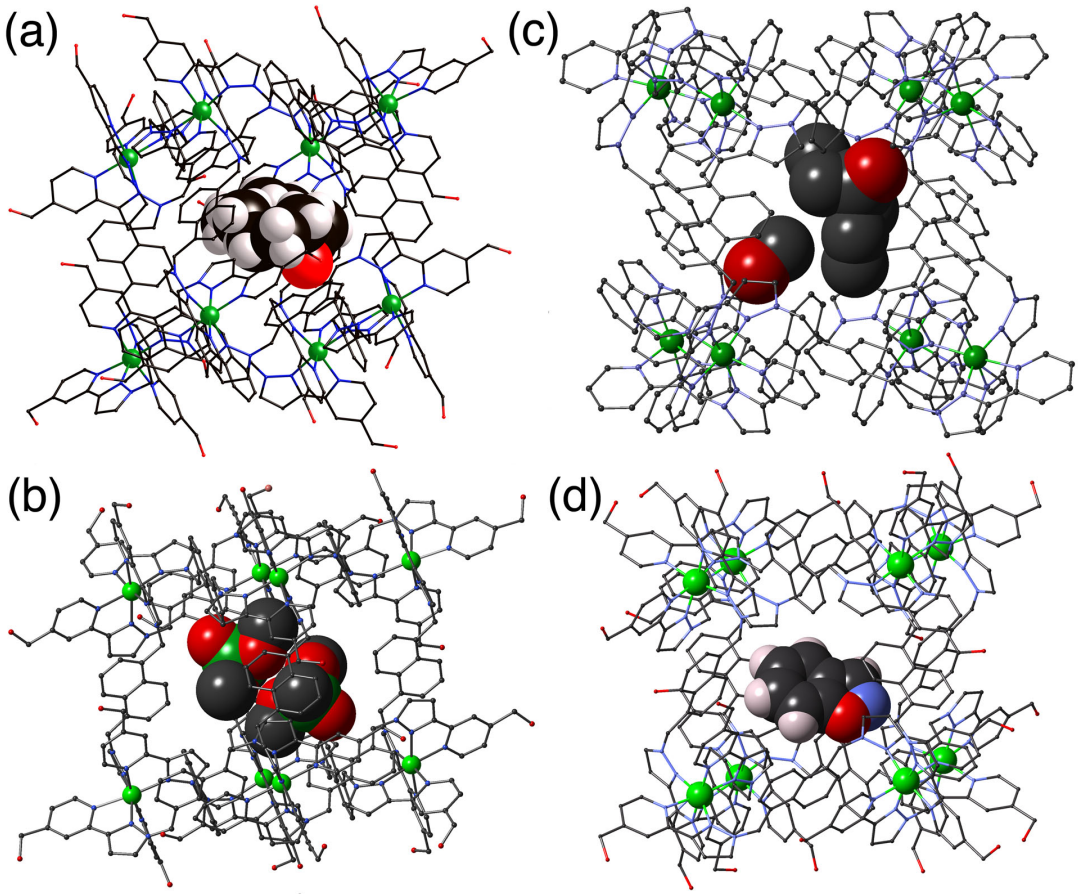


Figure 11 Crystal structures of hosts **H** or $\mathbf{H}^{w}$ containing as guests (a) cycloundecanone; (b) two molecules of $Me(MeO)_2P{=}O$; (c) heptan-4-one (and one MeOH molecule); (d) benzisoxazole.

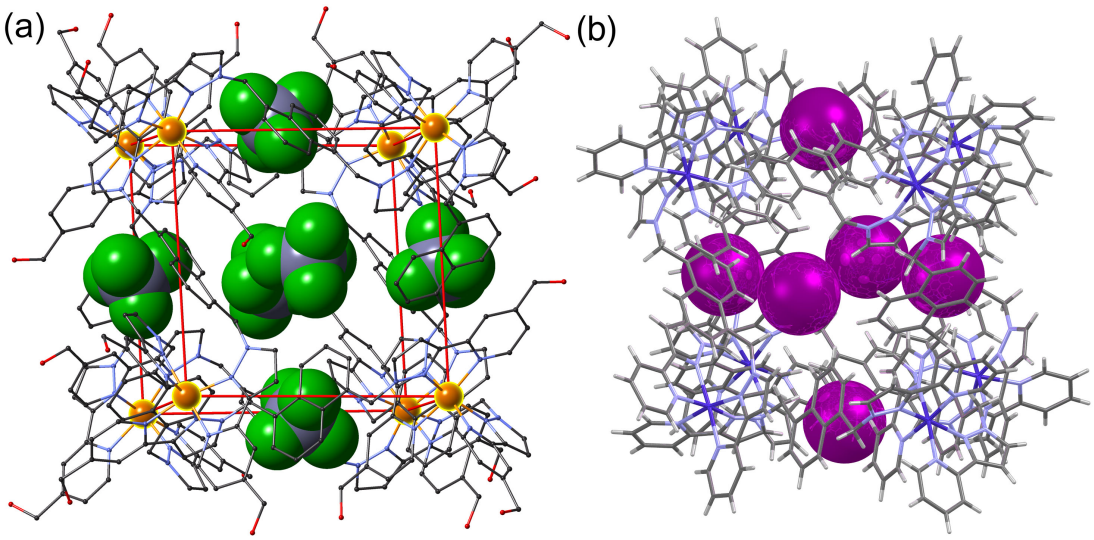


Figure 12 Crystal structures of hosts **H** or $\mathbf{H}^{w}$ illustrating the positions of anions in the portals around the cage surfaces: (a) tetrafluoroborate, and (b) iodide (refs. 7, 30).

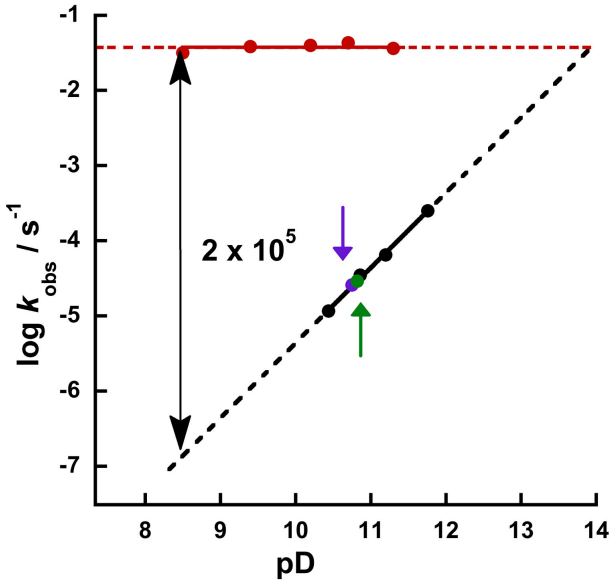


Figure 13 pD-dependence of rate constants of the uncatalysed Kemp elimination (black circles) and the $\mathbf{H}^{w}$-catalysed reaction (red circles) in D_2O at 298 K (for conditions see ref. 28). The purple circle shows the reaction rate in the presence of catalyst but with a competing guest (20 mM cycloundecanone) present as an inhibitor. The green circle shows the reaction rate in the presence of catalyst and 47 mM chloride (LiCl) as a competitor for the hydroxide binding sites (ref. 28).

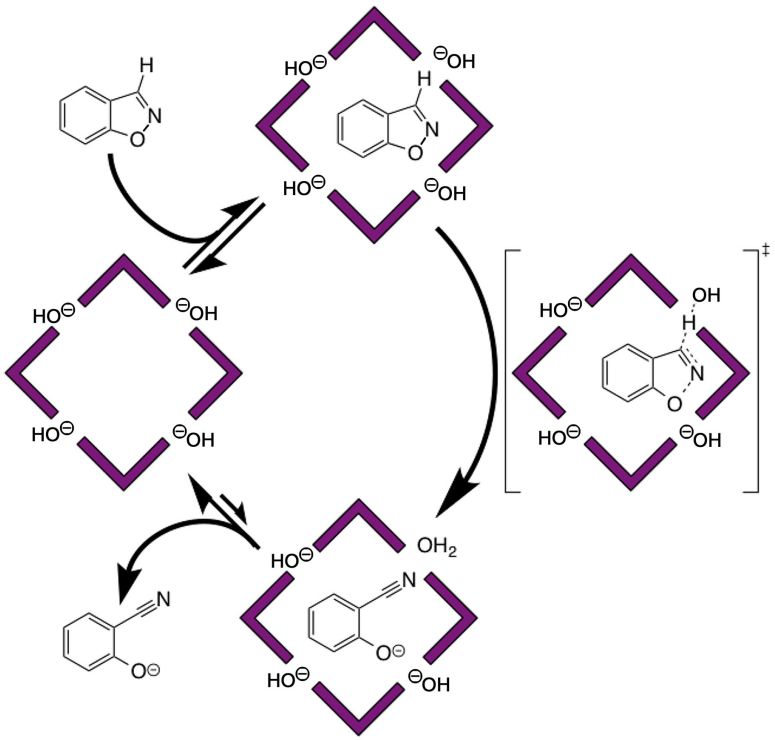


Figure 14 Cartoon of the catalytic reaction cycle, showing the role of the cage in bringing the benzisoxazole substrate (in the cavity) and the hydroxide ions (around the cage surface) into close proximity (ref. 28).

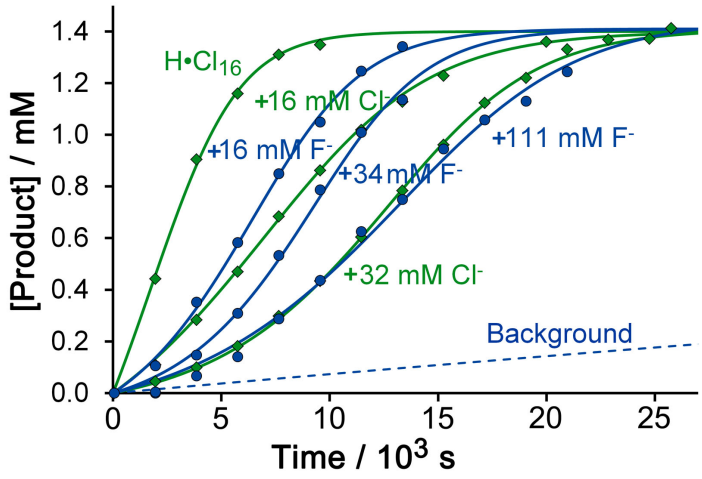


Figure 15 Comparison of the inhibiting effects on the **H**-catalysed Kemp elimination of added chloride (green curves) and fluoride (blue curves); the change in shape of the reaction progress curves to sigmoidal at high halide concentrations – indicating onset of autocatalysis – is clear (ref. 30).

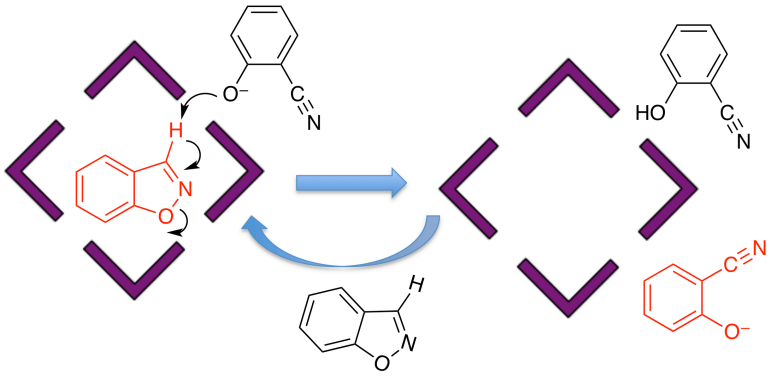


Figure 16 Autocatalysis of the Kemp elimination in cage **H** (ref. 30).