University of
Sheffield
&
University of Leeds

PhD Thesis

Synthesis and Applications of Solubilised Self-Assembling Cubic Cages M_8L_{12}

Author:

Christopher Graham Paul TAYLOR

Supervisor:

Prof. Michael D. WARD

November 2016

A thesis submitted to The University of Sheffield in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

Department of Chemistry, University of Sheffield, S3 7HF

Declaration

Except where specific reference have been made to other sources, the work within this thesis is the original work of the author. It has not been submitted, in whole or part, for any other degree.

Christopher Graham Paul TAYLOR

January 2017

Abstract

After the initial chapter (a general introduction to the relevant field) each following chapter contains an introductory section that gives relevant background information. The thesis of a whole consists of two general parts: firstly the synthetic development of functionalised cubic cage complexes; secondly the development of the host-guest chemistry of the Ward group's cubic cage, specifically concerned with the trapping, catalysed hydrolysis, and predicted binding of chemical warfare agent simulants.

Chapter 1: Introduction

This chapter consists of a brief history and general introduction to the supramolecular chemistry field. This starts with discussion of self-assembly and molecular interactions used. Principles behind the design of coordination-chemistry driven assemblies and key examples are introduced. Finally the various Ward group cages are discussed along with the recently published host-guest applications for the cubic cages ($\mathbf{H^A}$ and $\mathbf{H^W}$).

Chapter 2: Synthesis

The chapter focuses on the synthetic development of ligands to yield a cage complex soluble in non-polar solvents such as dichloromethane, using a range of different routes, similar to those for $\mathbf{H^A}$ and $\mathbf{H^W}$.

Chapter 3: Binding of Chemical Warfare Agent Simulants

This chapter introduces chemical warfare agents. The history, usage and related working involving G-series nerve agent simulants is discussed before $\mathbf{H^A}$ and $\mathbf{H^W}$ are used to bind a series of alkyl phosphonates (size and shape related simulants).

Chapter 4: Catalysed Destruction of Dichlorvos

Dichlorvos, a pesticide with similar chemical properties to G-series nerve agents, is bound in the cubic cage and then the cage catalysed hydrolysis is investigated. After finding some complications with $\mathbf{H^W}$, a different cubic cage host ($\mathbf{H^D}$) is used instead and gives a 500-fold rate enhancement for the hydrolysis of dichlorvos.

Chapter 5: Predicting Guest Binding

This final chapter delves into using a protein docking program (GOLD) to predict the binding affinities of a range of aliphatic ketones. While previous work has successfully predicted binding constants for rigid guests, improvement to the scoring function was required to predict the binding of flexible ketones. After improving the prediction of binding, the experimental binding affinities for the newly bound ketones and the series of alkyl phosphonates from chapter 3 closely matched those calculated.

Acknowledgements

Contents

Publication List

Published

J. R. Piper, L. Cletheroe, **C. G. P. Taylor**, A. Metherell, J. A. Weinstein, I. V. Sazanovich and M. D. Ward, Photoinduced energy- and electron-transfer from a photoactive coordination cage to bound guests, *Chem. Commun.*, 2017, **53**, 408-411.

C. G. P. Taylor, W. Cullen, O. Collier and M. D. Ward, A quantitative study of the effects of guest flexibility on binding inside a coordination cage host, *Chem Eur. J.*, 2017, **23**, 206-231.

C. G. P. Taylor, J. R. Piper and M. D. Ward, Binding of chemical warfare agent simulants as guests in a coordination cage: contributions to binding and a fluorescence-based response, *Chem. Commun.*, 2016, **52**, 6225-6228. Featured in two related news articles in Chemistry World and Scientific American.

In Preparation

C. G. P. Taylor, A. Metherell, M. D. Ward, Catalytic Hydrolysis of Organophosphate Toxin within the Cavity of a Coordination Cage.

Abbreviations

⊖			DMAE	2-Dimethylaminoethanol
A	Acceptor		DMF	Dimethylformamide
acac	acetylacetone		DMMP	Dimethyl methylphosphonate
AFIX	Proton fix (X-ray)		DOSY	Diffusion-Ordered Spectroscopy
AIBN	Azobisisobutyronitrile		DNA	Deoxyribonucleic acid
ATP	Adenosine triphosphate		Et	ethyl
$aq.$	aqueous enviroment		EtOAc	ethyl acetate
β_2	macroscopic binding constant		fac	facial geometry
BiPy	Biphenyl		F	Force
BuLi-LiDMAE	See Scheme 2.19		f_{HG}	molar fraction of H·G
C_7	7 carbon ketone		G	Guest
C_9	9carbon ketone		[G]	guest concentration
cat	catalysed		[G]$_o$	
CWA(s)	Chemical warfare agent(s)		GOLD	Genetic Optimisation of Ligand Docking
δ_H	chemical shift for H		h	hour(s)
δ_{HG}	chemical shift for H·G		H	Host
δ_{calc}	calculated chemical shift		[H]	host concentration
δ_{obs}	observed chemical shift		[H]$_o$	host concentration at time zero
ΔG	Gibb's free energy		HD	mustard gas
ΔH	enthalpy		H·G	host-guest complex
ΔS	entropy		[H·G]	host-guest complex concentration
ΔY	change in physical property		HPLC	High Performance Liquid Chromatography
D	Donor		$in\ vacuo$	in a vacuum
DCM	dichloromethane		$in\ vitro$	in cells (not whole organism)
DEEP	Diethyl ethylphosphonate		$in\ vivo$	in a living organism
DEMP	Diethyl methylphosphonate		IR	infra-red
DFIX	Distance fix (X-ray)		J	coupling constant
DIMP	Diisopropyl methylphosphonate		K_a or K	Binding constant
DMA	dimethylacetamide		K_{off}	rate of binding
DMA-DMF	N,N-Dimethylformamide dimethyl acetal		K_{on}	rate of unbinding

K_e	Coulomb's constant	pH	potential of hydrogen
K_{cat}	rate of catalysed reaction	pD	potential of deuterium
K_{uncat}	rate of background reaction	PXRD	Powder X-Ray Diffraction
L	ligand	R	gas constant
LDA	Lithium diisopropylamide	R^2	R-squared (stat)
M	metal	RBF	Round bottom flask
M	Molar	RMSD	Root-mean-square deviation
mCPBA	meta-Chloroperoxybenzoic acid-	RNA	Ribonucleic acid
Me	methyl	*sat.*	saturated
MeOH	methanol	s	second
mer	meridional	θ	angle
$MgSO_4$	magnesium sulfate	T	temperature
min	minute(s)	TBAF	Tetra-n-butylammonium fluoride
mol	moles	TBDMSCl	tert-Butyldimethylsilyl chloride
MS	Mass spectroscopy	TFA	Trifluoroacetic acid
NaH	sodium hydride	THF	tetrahydrofuran
NaOH	sodium hydroxide	TLC	thin layer chromatography
NBS	N-bromosuccinimide	TMSCN	Trimethylsilyl cyanide
n-BuLi	n-butyllithium	TMV	Tobacco Mosaic Virus
N_c	Number of Carbon atoms	uncat	uncatalysed
NMR	Nuclear Magnetic Resonance	UV	Ultraviolet
NRB	number of rotatable bonds	X	anion
OTBDMS	oxy-tert-Butyldimethylsilyl		

Ligands

$\mathbf{L^A}$	See Fig. 2.1	$\mathbf{L^{ter}}$	See Scheme 2.22
$\mathbf{L^B}$	See Fig. 2.1	$\mathbf{L^{Et}}$	See Scheme 2.8
$\mathbf{L^C}$	See Scheme 2.21	$\mathbf{L^{Me}}$	See Scheme 2.17
$\mathbf{L^W}$	See Fig. 2.1	$\mathbf{L^{??}}$	All other defined ligands See Fig. 1.41
$\mathbf{L^W_{mod}}$	See Scheme 2.4		

Hosts

$\mathbf{H^A}$	$[Co_8(L^A)_{12}](BF_4)_{16}$	$\mathbf{H^{Et}}$	$[Co_8(L^{Et})_{12}](BF_4)_{16}$
$\mathbf{H^B}$	$[Co_8(L^B)_{12}](BF_4)_{16}$	$\mathbf{H^W}$	$[Co_8(L^W)_{12}](BF_4)_{16}$
$\mathbf{H^D}$	$[Co_8(L^A)_{12}]Cl_{16}$		

Chapter 1

Introduction
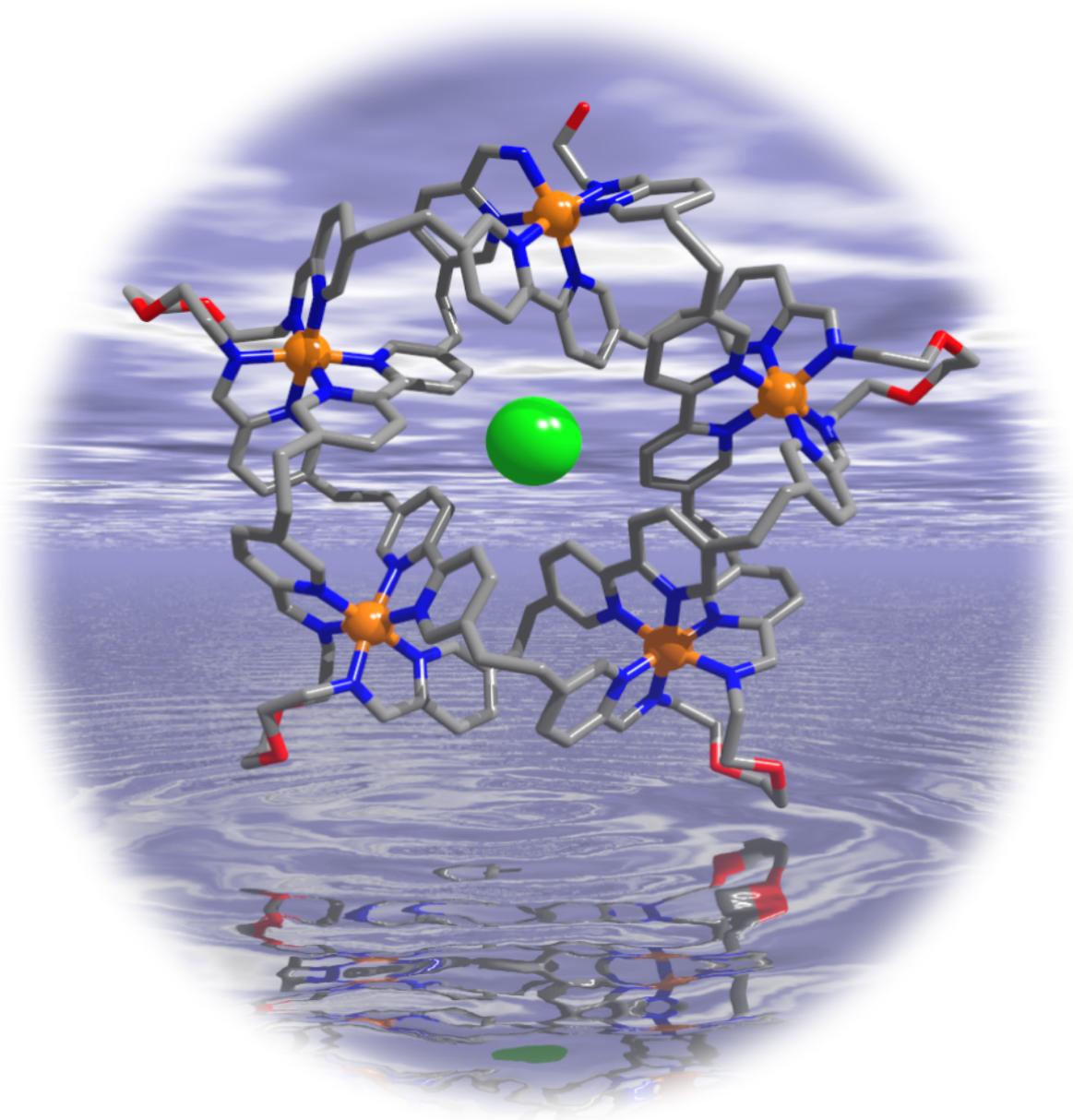
Rendering of a molecular knot with bound chloride ion by Prof D. Leigh[1]

1.1 Supramolecular Chemistry

Supramolecular chemistry, the "chemistry beyond the molecule"[2], dates back to the late 1960s and is the study of systems involving aggregates of molecules/ions held together by non-covalent interactions to form ordered structures, generally through self-assembly[3]. There is a wide range of examples from knots[1,4] (see coverpage of chapter 1) or cages[5] through to proteins, DNA origami[6,7] or nano-machines[8], all formed through this self-assembly method. Supramolecular chemistry came to the forefront of chemistry in the late 80's where it has remained ever since[9]. Although still in relative infancy, self-assembled structures prepared using supramolecular chemistry are starting to rival the complexity and functionality of some biological constructs. This rapidly expanding field of research impinges on many large research areas including organic, inorganic, biological, physical, and materials chemistry[10].

Many of the first supramolecular assemblies were reliant on serendipity for their discovery but as the understanding of the underlying principles that dictate these assemblies improves, rational design is being utilised more and more[11]. Biology has many examples of self-assembled complexes, evolved over millennia, which are highly complex and highly functional. Well known examples include photosynthetic reaction centres, capable of capturing and storing the energy from light, enzymes that can convert a proton gradient into cell energy currency (ATP molecules), and DNA which stores and encodes information for every single molecule, enzyme, metabolic reaction which occurs in living cells. As our ability to produce supramolecular assemblies improves, the complexity and functionality increases promising huge potential benefits. The growing wealth and importance of this field was first highlighted with a Nobel Prize awarded to Cram, Lehn and Pedersen in 1987[12] along with the recent Nobel Prize awarded to Sauvage, Stoddart and Feringa in 2016[8].

1.1.1 Self-Assembly

What is self-assembly? If you take various blocks of a castle (Fig. 1.1) and these are placed within a bag and shaken, will the blocks assemble into a completed castle? (Fig. 1.1) This illustrates the task that self-assembly aims towards, using non-covalent interactions at the molecular-scale to achieve[13]. Within recent history, self-assembly has transformed the interface between many of the different sciences including nanotechnology with 10 % of nanotechnology-related articles addressing the concept of self-assembly[13].
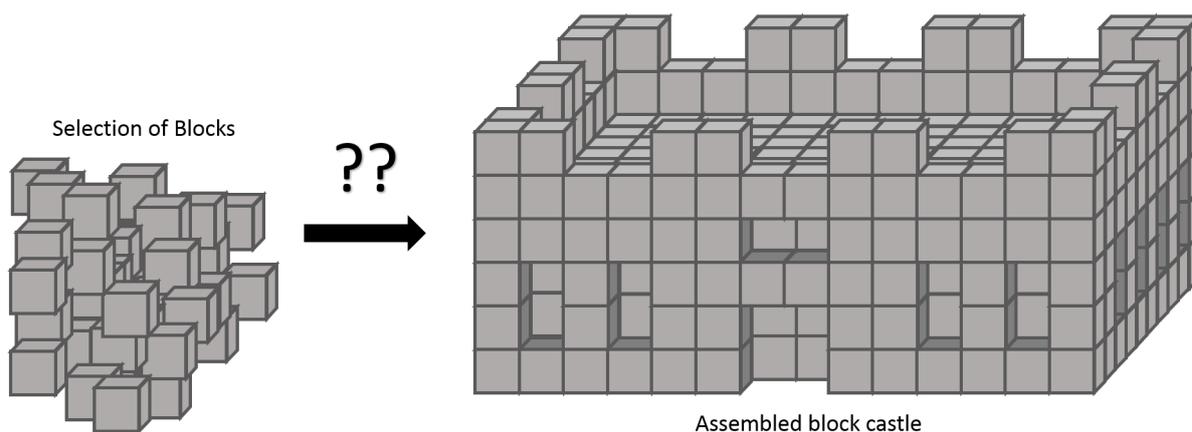
Fig. 1.1 Illustration (left) a random selection of blocks; (right) a "castle" made from blocks.

Self-assembly is the spontaneous and reversible association of two or more different components to form a larger, non-covalently bound aggregate[3]. These assemblies are self-driven and require no external input, often forming in relation to thermodynamic minima. It is the use of non-covalent interactions, typically weaker than covalent interactions, which allow the components to reach the lowest thermodynamic minimum by allowing incorrect systems to break apart and reform[3,14].

Self-Assembly:

"a process where pre-designed components assemble in a determined structure without the intervention of human operators"

— George Whitesides[15]

Self-assembly is prominent within nature, with the tobacco mosaic virus being one of the most commonly known examples which has been studied for over 60 years. The simple

structure is made from a core RNA strand to which a regular helical array of identical protein subunits binds to give the rod shaped virus. The helical array contains 16.5 subunits per turn with three individual nucleotides making up each subunit[16]. Upon changes in temperature or pH the subunits can dissociate while reverting back to the original conditions allows the re-assembly of the active virus. This reversibility is the key to self-assembly, allowing access to the structures at their thermodynamic minimum rather than the kinetic product.
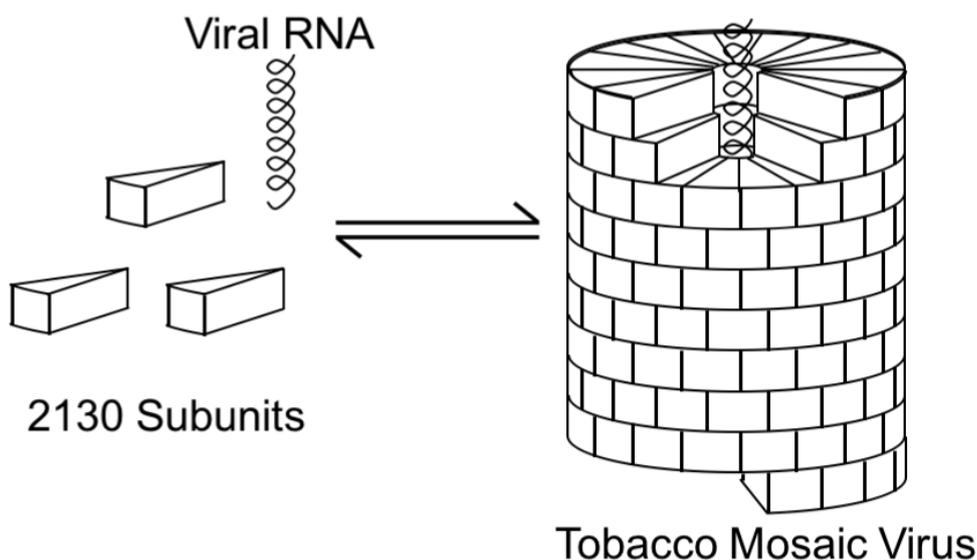
Fig. 1.2 A cartoon showing the assembly of a tobacco mosaic virus from viral RNA and subunits[16].

1.1.1.1 Non-Covalent Interactions

There are numerous further examples of self-assembly in nature based on functional building blocks which self-assemble in a deliberate manner. Protein assemblies, nucleic acid structures, phospholipid membrane mosaics, and ribosomes are all essential for living organisms and all reliant on self-assembly. Nature uses a range of weak, non-covalent bonds to achieve highly complex and often symmetrical architectures. Interactions used include charge-charge electrostatic interactions, dipole-dipole interactions, hydrogen bonding, $\pi - \pi$ and Van der Waals interactions[17]. Solvent interactions are also important in self-assembly and discussed in more detail later in this chapter (Section 1.4.1).
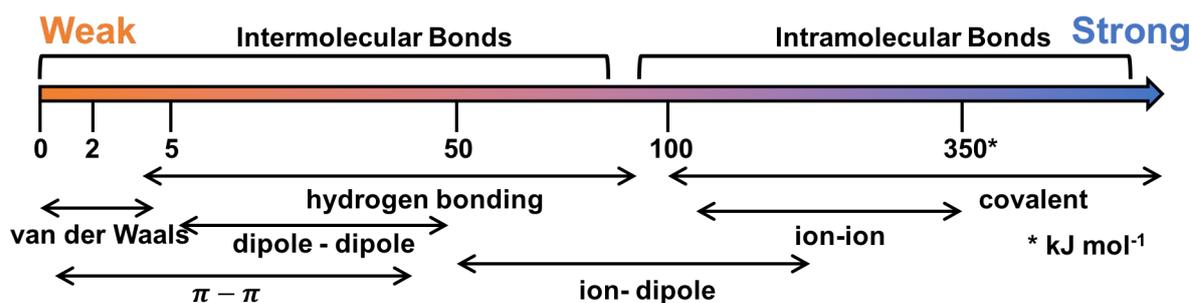
Fig. 1.3 Diagram depicting the relative strength of different interactions[17]

Electrostatic Interactions These are the strongest non-covalent interactions commonly seen in nature. Interactions include, from strongest to weakest, point charge to charge interactions, dipole-dipole interactions and higher multipole moment types (e.g. quadrupoles or octapoles). The strength of these interactions is dependent on distance (r^{-2} dependence)[17].

$$F = K_e \frac{q_1 q_2}{r^2} \qquad \text{Where } K_e \text{ is Coulomb's Constant} \tag{1.1}$$

Hydrogen bonding is a special case of a dipole-dipole interaction which occurs between an electron deficient hydrogen bound to an electronegative atom (hydrogen-bond donor) and an electron rich atom (hydrogen-bond acceptor). Hydrogen bond interactions are typically stronger than other dipole-dipole interactions. In contrast to typical electrostatic interactions, hydrogen bond strengths are highly dependent on the bonding angle. Uniquely a hydrogen atom in a donor group uses its core electronic orbital (1s) in bonding and so when bound to an electronegative atom the proton core of the hydrogen becomes exposed. As a result the optimal binding angle D-H●A is 180°[17,18].
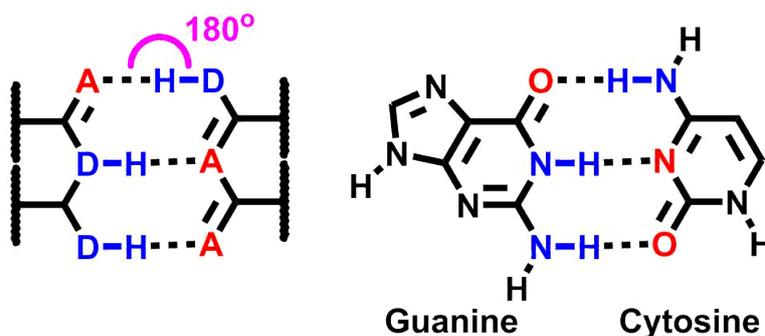
Fig. 1.4 Hydrogen bonding between two nucleotides, A = hydrogen bonding acceptor, D-H = hydrogen bond donor

$\pi - \pi$ interactions are quadrupole-quadrupole interactions. Although they are much weaker than hydrogen bonds, they are also dependent upon the bonding angle and offset between the aromatic systems. Typically these interactions occur between two aromatics, one being relatively electron rich whilst the other is electron poor. Most of the interactions are either face-to-face or edge-to-face stacking though a variety of other intermediate geometries known[17,19].
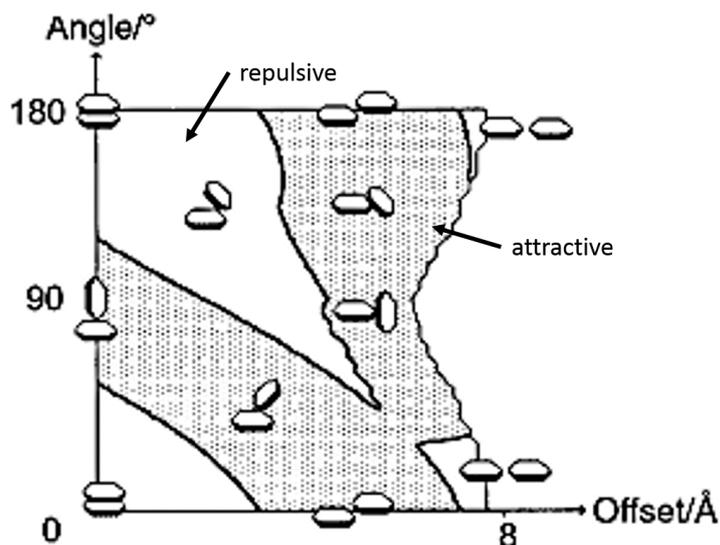
Fig. 1.5   The electrostatic interaction between two benzene rings as a function of offset and orientation (shaded = attractive, unshaded = repulsive). Reproduced (in part) from ref[19] with permission of The Royal Society of Chemistry

Induced Interactions   The easiest example of an induced interaction can be seen in polar solvents. The polar molecules all have dipoles but these are randomly arranged in solution with no long distance order. One molecule can cause the neighbouring molecule to line up with it momentarily leading to short-range order and temporary dipole-dipole interactions before disorder returns. These events only occur locally and temporarily while the bulk retains disorder.
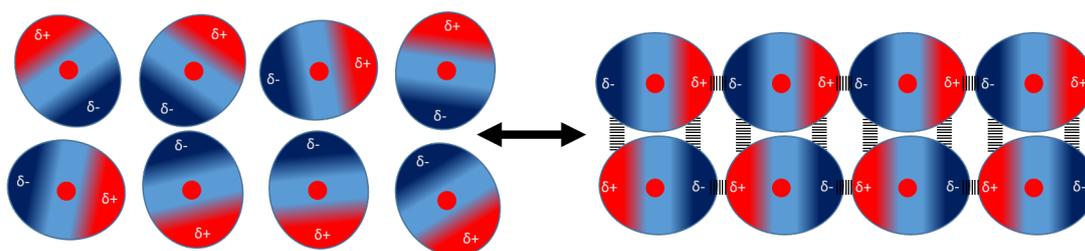
Fig. 1.6   Cartoon showing the temporary induction of long range interactions between polar molecules

Van der Waals' interactions, also known as London Dispersion Forces, are similar to the induced interactions seen in polar solvent except that they occur in non-polar materials which have no fixed dipole. Due to random electron density fluctuations even non-polar molecules can briefly gain an instantaneous dipole. Once formed the dipole is induced into the neighbouring molecules, building to form the weakest interaction commonly seen with the attraction force decreases rapidly with distance (r^{-6} dependency). Though weak when this interaction occurs over large surfaces the additive effect of a large number of interactions explains (for example) why large alkanes are often solids[17].
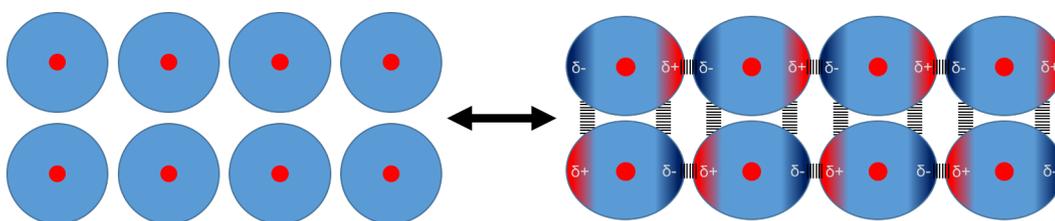
Fig. 1.7 Neutral atoms (left) can have temporary dipoles which induce further dipoles in neighbouring molecules leading to short-range attraction (right)

Combinations of these interactions work cooperatively to allow the formation of robust nano-sized assemblies. Self-assembly is an example of bottom-up construction (Fig. 1.8). In contrast nanotechnology has typically employed the top down approach for manufacturing nano-sized and larger structures. Micro-fabrication, for example, of integrated circuit transistors, found within all our modern technology, uses a range of techniques collectively called lithography. Combinations of light sensitive coatings, selective light exposures and chemical etching are used to produce the small features required from a large bulk starting material[20].

Organic synthesis is the most common form of a bottom up construction and has a wide range of powerful techniques for creating molecules. Control of synthesis within organic chemistry has been heavily developed and some nano-sized molecules can be accessed. These generally are "natural product" synthesis, requiring many synthetic steps, with purification between each, and give poor overall yields. The size and complexity of nano-scaled assemblies is such that generally organic synthesis becomes too difficult. This technique relies on the correct combination of kinetically inert covalent bonds and as a result any mis-formed bonds are often very difficult to remove. This
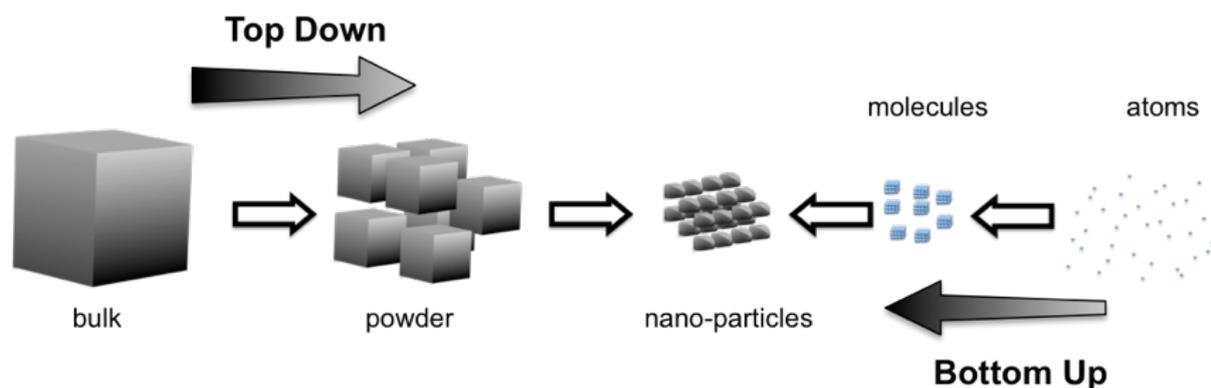
Fig. 1.8 Cartoon depicting the top down vs bottom up approach to reach nano-sized structures

means that once one mistake is made, that material is wasted and needs to be removed or recycled before any further synthetic steps can be completed[5].

The supramolecular field, which aims to construct nano-scaled and larger scaled structures formed using self-assembly, provides an alternative bottom-up approach to synthesis. The area covers a wide breadth of different structures in both size, type and function, with many of the structures being inaccessible using conventional synthetic methods. When divided according to the interactions utilised for controlling assembly, supramolecular chemistry can be broadly divided into three main branches[11]:

- Hydrogen bonded assemblies

- Assemblies based on combinations of other non-covalent interactions

- **Assemblies based on metal ligand coordinative bonding**

Herein the focus will remain within supramolecular coordination chemistry, the self-assembly driven by metal ligand bonds. The coordination bond offers an 'easy' access to self-assembly. The interactions are relatively strong, but still labile and reversible with clear a directional nature that is predictable, which is the ideal combination of properties for self-assembly to occur[14].

1.2 Supramolecular Coordination Chemistry

The supramolecular coordination chemistry field started in the 1960's with the discovery of crown ethers, cryptands and spherands by Pedersen[21] , Lehn[22] and Cram[23],

respectively (Fig. 1.9). This initial work was later recognised by the Nobel Prize awarded in 1987[12].

The work showed that small molecules could be "programmed" to recognize each other through weak non-covalent interaction such as those seen in nature's self-assembled structures. Typically the self-assembly process is kinetically reversible allowing for numerous association and dissociation steps on the path to the product at thermodynamic minimum.

The field of supramolecular chemistry quickly developed beyond its starting point in small molecule recognition of metal ions. Use of coordination bonds allowed access to a stable but reversible bonding system and therefore a route into self-assembly.

Crown Ethers　　　　**Cryptands**　　　　**Spherands**
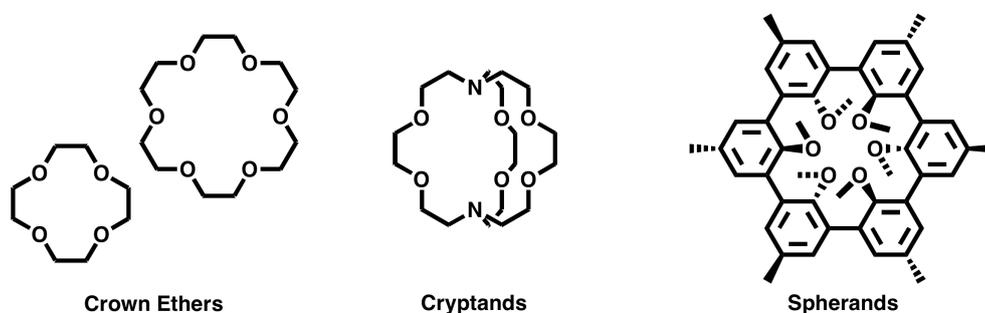
Fig. 1.9　Structures of the molecules key within the early supramolecular field

Lehn and co-workers used oligo-bipyridine ligands with Cu^{I} cations to "program" the assembly of double stranded helicates[24]. These helicates, like those found in biological systems, form through positive cooperation (formation of an initial interaction modifies the binding affinity for proceeding interactions allowing them to form more readily). They also showed that by substituting the 4-coordinate Cu^{I} ions with the 6-coordinate Ni^{II} ions that a triple stranded helicate would form instead[25]. "Programming" a steric preference into the oligo-bipyridine strands via positioned terminal methyl groups allowed for selection of the metal ion type. Mixing a pot of the two oligomer strands along with both types of metal ion resulted in ordered selection of double helicates made from the 2-methyl pyridine terminated ligand with Cu^{I} ions and the 3-substituted-pyridine selecting for the Ni^{II} ions forming the triple helix (Fig. 1.10). Lehn extended this work to create 2-dimensional structures including squares, grids and racks[26].

Beyond the early examples of coordination driven assemblies by Lehn[2] and Sauvage[27], there have been examples of grids, ladders, knots, rings, catenanes, rotaxanes and more
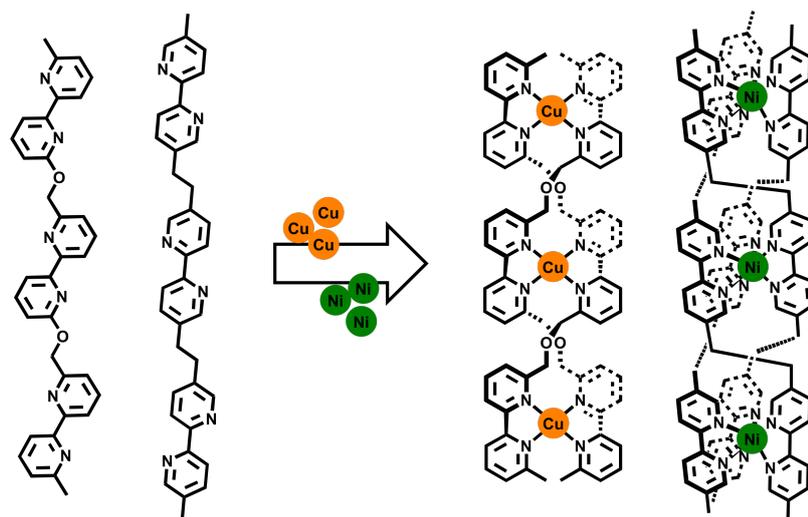
Fig. 1.10 Two oligomers (left) selectively form a double or triple helicate with Cu^{II} or Ni^{II} metal ions[25]

by several groups including those of Stang[28], Raymond[29], Fujita[30] and others[11,31].

1.2.1 Design Principles

Self-assembly can easily and often lead to infinite structures such as self-assembled polymers or metal-organic frameworks[32]. The difference between infinite and discrete structures is down to a careful balance between the entropic and enthalpic driving forces. The assembly forms through multiple intermediates to give the final product under thermodynamic control, and a single product is only formed when there is sufficient energetic stability of one species over other possibilities. There are many examples of two or more species being present during an assembly, and when several products are of similar energy then a combination of different products is formed[33].
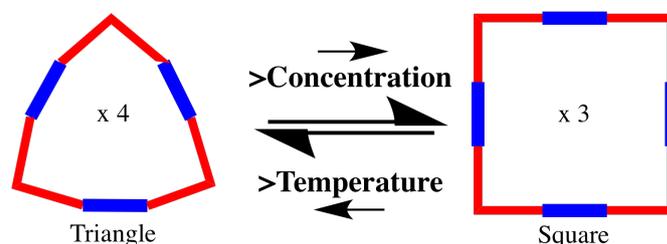
Fig. 1.11 Equilibrium between a molecular triangle and molecular square[33]

A simple example consists of an equilibrium between a molecular triangle and a molecular square (Fig. 1.11). A 90° angle between two coordination sites should, in combination with linear ligands, lead to a molecular square. Instead an equilibrium

10

between a square and triangular structure is often obtained[33]. Enthalpic driving forces favour the square structure as there is a negative effect from introducing bond strain to form the triangle. Entropic effects however favour the smaller, triangular structure as it contains fewer components and therefore a larger number of assemblies form (four trimers *vs.* only three tetramers).

$$\Delta G^{\ominus} = \Delta H^{\ominus} - T\Delta S^{\ominus} \tag{1.2}$$

The balance between the two driving forces in self-assembly is described in equation (1.2). It is clear that the temperature is important in determining Gibb's free energy ($\Delta G^{\ominus}$) and therefore the equilibrium point between a square and a triangle. As the temperature increases, the entropic term dominates leading to a smaller molecular triangle while at lower temperature the larger square will be favoured. Additionally, according to Le Chatelier's principle, the equilibrium is concentration dependent. As the concentration of building blocks is increased the equilibrium shifts to favour the structure containing more components and so favours the square structure. The ligand can also be tuned to control the product outcome. By increasing the flexibility of the ligand, the enthalpic cost of bond strain is reduced and the molecular triangle is favoured[11,33].

1.2.1.1 Directional-Bonding Approach

The directional-bonding approach, coined the "molecular library" approach by Stang and co-workers[5], involves the assembly of large structures through highly directional bonding between pre-defined rigid subunits. Through careful selection of the subunits, for example a cis-capped Pt^{II} that provides a 90° binding angle, formation of the resulting structure can be controlled.

This approach was first applied by Verkade[34] and later refined by Stang[35] and Fujita[36,37]. Fujita and co-workers created one of the first examples of a self-assembled molecular square. By combining linear bis-mono-dentate ligands, such as 4,4'-bipyridine, with cis-capped square-planar metals, such as Pt^{II} and Pd^{II}, the 90° corners select for discrete squares (Fig. 1.13).

The Pd^{II} ions are labile enough to allow the system to reach thermodynamic equilibrium. The different intermediates can be distinguished within the ^{1}H-NMR spectra
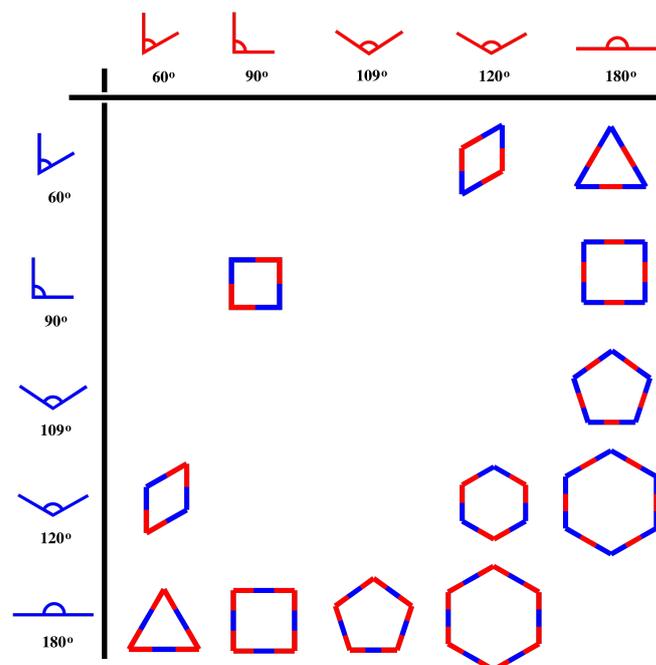
Fig. 1.12 Molecular library of potential two-dimensional structures[5]

by varying the ratio of ligand to metal. By mixing the components together in the correct 1 to 1 ratio the desired complex is formed quantitatively in less than 10 minutes[33]. The high level of lability has the disadvantage of the complex not being stable enough to be used in applications. The use of the more inert Pt^{II} ions in place of Pd^{II}, and heating at over 100 °C for a week, gives the more stable analogous square structure[37] and on cooling this becomes kinetically inert and the structure is "locked".
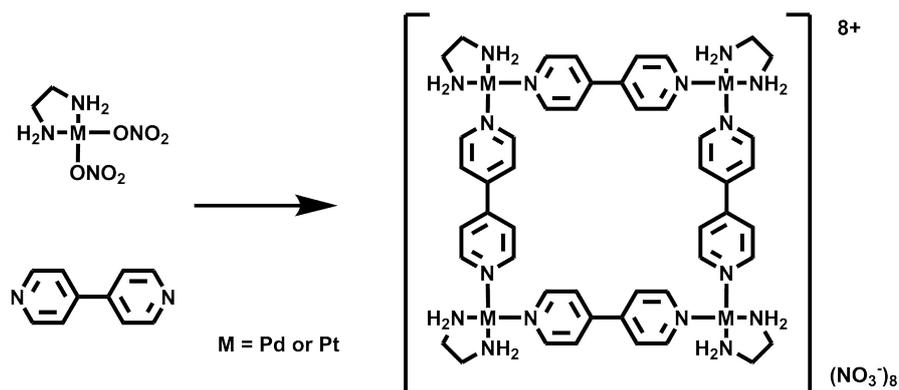
Fig. 1.13 Fujita's molecular square M_4L_4 complex[37]

There are many examples of dinuclear macrocycles, molecular triangles, squares and beyond in the literature[4,5,11,14,31]. This methodology has been extended into three-dimensions simply through the use of subunits with three or more connections. The expansion of this molecular library into polyhedral structures demonstrates the power of

this strategy and is still heavily utilised within the supramolecular chemistry field[31].
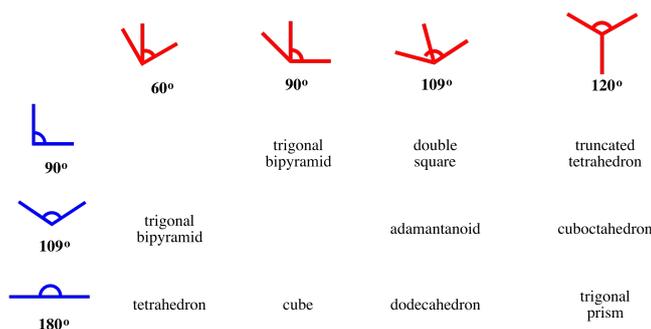
Fig. 1.14 Molecular components using to form three-dimensional assemblies[5]

1.2.1.2 Symmetry-Interaction Approach

The use of multibranching chelating ligands in combination with octahedral coordination transition metals allows access to a range of elegant shape and structures. This so called "symmetry approach" uses the high symmetry of the components to drive the formation of the supramolecular architectures[29,38]. An additional result of using multidentate ligands is the stronger associated binding energies, arising from the chelate effect, which lead to improved organisation within the assemblies. Although described as a "design principle", the majority of the structures formed using this methodology were discovered serendipitously. With growing understanding of the symmetry principles, learning from both nature and previous supramolecular structures, many groups are positively applying the design method towards new systems[31].
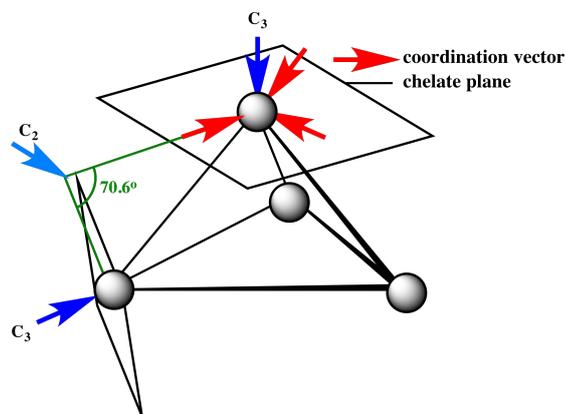
Fig. 1.15 Sketch of a tetrahedral structure depicting the components described in Raymond's symmetry-interaction design approach[29,38]

Raymond and co-workers developed the design principle further by defining the geometric relationship between the ligands and metals (Fig. 1.15). The *coordinate vector*

describes the direction from which the ligand binds to the metal; in the case of a bidentate ligand the vector bisects the two connections pointing towards the metal centre. The *chelate plane* is the plane orthogonal to the major symmetry axis of the metal ion.
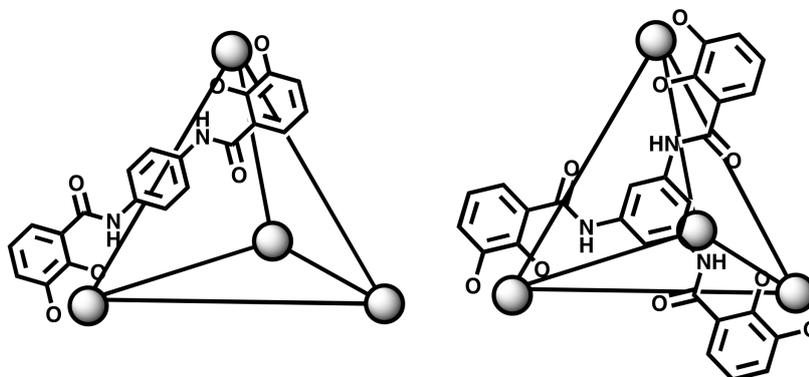
Fig. 1.16   Raymond's tetrahedral complexes (left) M_4L_6, (right) M_4L_4[29]

For the design of an M_4L_6 tetrahedron, there must be C_3 symmetry perpendicular to the chelate plane and also C_2 symmetry through the tetrahedral edge and therefore within the ligand (Fig. 1.15). The ideal angle between these vectors can also be calculated and is ideally 70.6°. Applying the same principles to an M_4L_4 tetrahedron requires two C_3 axis, one perpendicular to the chelate plane and the second within the ligand. Examples of both of these complexes can be seen in Fig. 1.16[29].

1.2.1.3   "Panelling" Approach
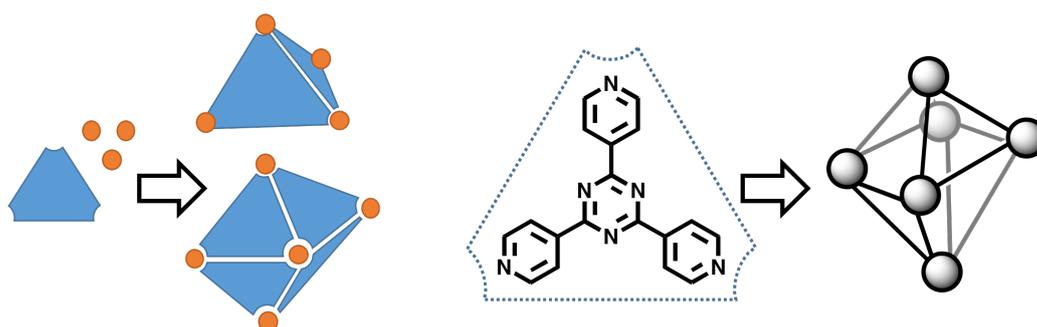
Fig. 1.17   Schematic showing Fujita's "panelling approach" to form an octahedral M_6L_8[30,37]

The "panelling approach", pioneered by Fujita and co-workers[30,37], uses the idea of connecting pre-formed shapes to form the desired platonic solid. For example, to form a tetrahedron one needs to connect four triangles together. To get a cube one must connect eight squares. This allows for the design of complex and elegant three-dimensional

structure through the easier design of two-dimensional building blocks. Typically the flat panels are constructed from an appropriate organic ligand and stuck together using metal ions as the 'glue' via metal-ligand coordination interactions.

1.3 Supramolecular Cages
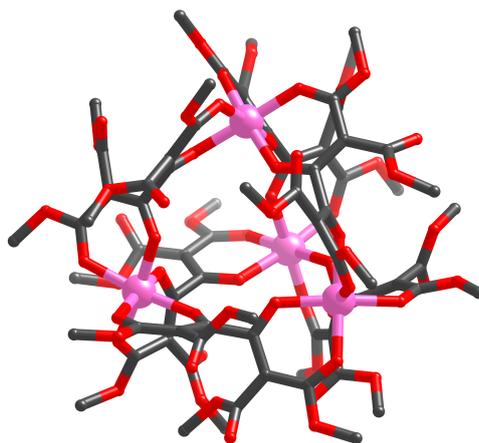
Fig. 1.18   Render of the crystal structure for Saalfrank's cage[39]

Access to three-dimensional structures, especially through rational design, remains challenging and many of the architectures observed are still a result of serendipity. Development of the above design principles, especially within the realm of high-symmetry coordination cages, has led to a wide range of Platonic solid structures as well as some of the more complex Archimedean solids. The polygons that occur within Platonic and Archimedean solids, defined below, normally consist of equilateral triangle, square, pentagon, and hexagon.

Platonic Solid:

A structure constructed with faces of the same regular polygons.

Archimedean Solid:

A structure constructed with faces of two or more different regular polygons which meet through identical vertices.

1.3.1 Tetrahedral Systems

A tetrahedron is the simplest form of a Platonic solid: even so, when thinking about the design of such a structure, there are more than a few synthetic approaches. From the molecular library approach, **(Fig. 1.14)**, we can see a combination of six linear edge-binding components with four tritopic angular components (vertices) can form the tetrahedron **(Fig. 1.19a)**. A symmetry approach uses four octahedral metal centres with a 70.6° angle between the three coordination vectors in combination with six bis-bi-dentate ligands **(Fig. 1.15)**. Both of these approaches can yield a M_4L_6 complex (Fig. 1.16 and Fig. 1.18).
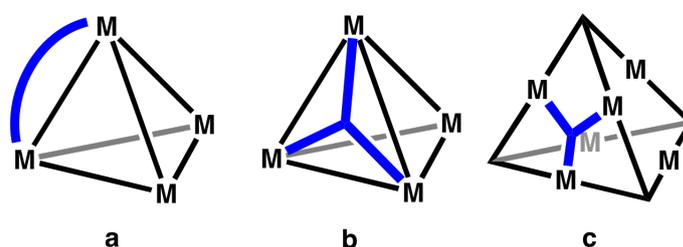
Fig. 1.19 Three simple topologies for a tetrahedral coordination complex. Only one ligand is highlighted in blue for each for clarity

Using the alternative "panelling" approach gives access to a different topology, an M_4L_4 complex **(Fig. 1.19b)**. Here the tritopic ligands (the panels) span the triangular face of the tetrahedron. The same approach using ditopic metal binders in middle of each edge with the ligand spanning the space between the edges gives an M_6L_4 complex **(Fig. 1.19c)**. This can either be described as a tetrahedron (based on the panel approach) or a truncated-tetrahedron if defined using the metal centres.
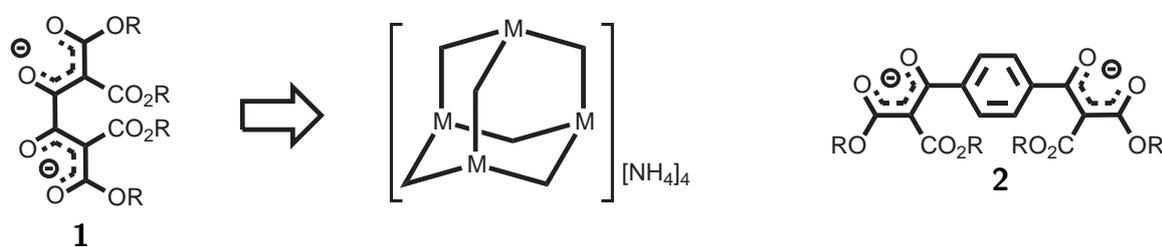
Fig. 1.20 (left) Saalfrank's ligand which forms an M_4L_6 complex shown in the render above (Fig. 1.18)[39], (right) the extended ligand[40]

One of the first three-dimensional cage structures was Saalfranks adamantanoid style structure (Fig. 1.18)[39], an M_4L_6 tetrahedron with ethyl malonate as the bridging ligand **1**. The structure can be seen as utilising the symmetry approach and was made in many

variations using different octahedral coordinating metal ions including Ni^{II}, Co^{II} and Fe^{III}. The ammonium counter-ions formed hydrogen bonds with the oxygen donors present along the triangle edges. The structure was found to be homo-chiral and kinetically stable on the NMR timescale. Later a phenylene spacer group was added into the bridging ligand to give **2**, allowing the formation of a tetrahedral cage with an expanded central cavity (Fig. 1.20)[40].

These tetrahedra are nearly all composed of achiral building blocks, but due to the ligands twisting around the bridging edges between metal ions most of the tetrahedra end up being chiral. This helical twist is introduced partly to attain the 70.9° angle required between chelate planes. Raymond and co-workers followed on from Saalfrank's design to produce a perfectly symmetrical tetrahedral cage from direct application of their developed theory[29]. They designed ligand **3** to have a perfect 70° angle between the two coordination vectors. While still containing the desired $\mathbf{C_2}$ symmetry, it allowed the bound metal ions to have $\mathbf{C_3}$ rotational symmetry.
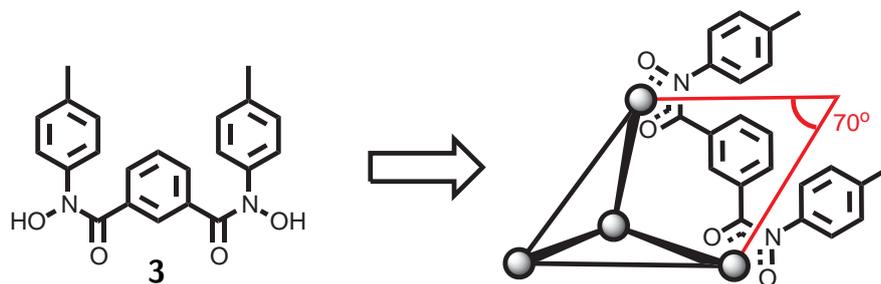
Fig. 1.21   Raymond's designed ligand which forms a symmetrical tetrahedral complex[29]

In this particular example, **Fig. 1.21**, either $Fe(acac)_3$ or $Ga(acac)_3$ were used to give $[M_4L_6]^{12-}$ complexes[29]. Exchanging the 1,4-phenyl spacer with a 1,5-naphthalene-diyl spacer results in the formation of a racemic mixture of the homochiral $[M_4L_6]$ species due to breaking the perfect 70° angle. Both Saalfrank's cage and Raymond's complexes are negatively charged as a result of the negatively charged ligands. The majority of examples from other workers are actually cationic cage complexes formed from the combination of neutral ligands with metal cations[11,31].

Fujita and co-workers also extended the principles used in their molecular square complex (Fig. 1.13) by making the ligand tritopic to give a three-dimensional cage complex $[M_3L_2]^{6+}$ (Fig. 1.23). The flexible ligand **4**, used in the $[M_3L_2]^{6+}$ cage,
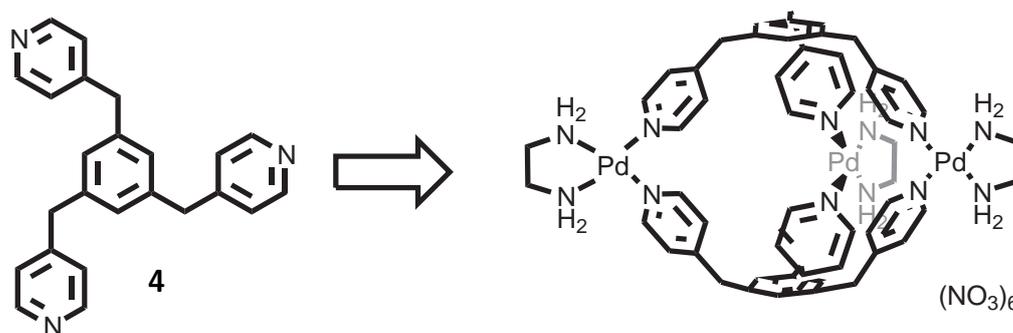
Fig. 1.22 Fujita's M_3L_2 "pancake" complex[41]

has pyridines each attached to a central benzene ring through a methyl connector (Fig. 1.22)[41]. Removing the flexibility within the ligand gives **5** which cannot bend and leads to a larger cage structure (Fig. 1.23). The ligand with three pyridines attached directly to a central triazine unit along with the same Pd^{II} complexes resulted in a $[M_6L_4]^{12+}$ cage (Fig. 1.23)[42].
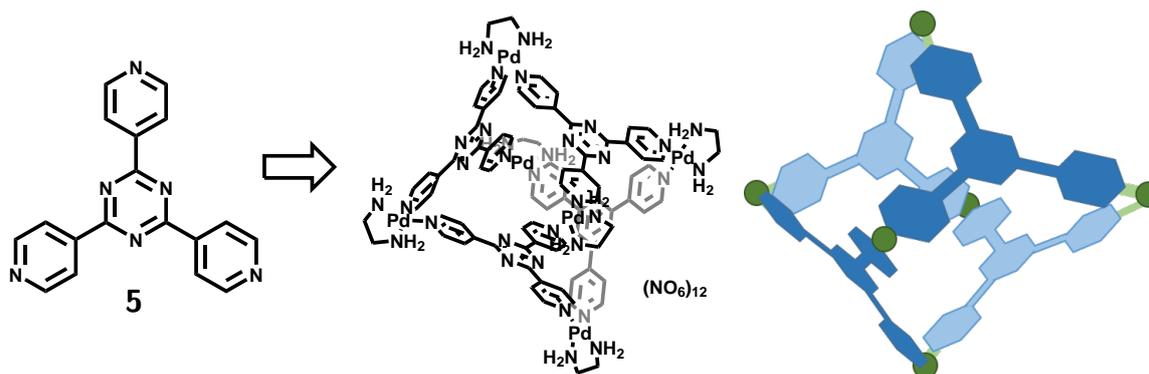
Fig. 1.23 Fujita's tetra-caped tetrahedron based on the "panelling design" approach[42]

1.3.2 Cubic Systems

From a design point of view cubic cages can be constructed from two approaches, edge directed or face-directed. The edge directed method allows use of the molecular library or symmetry approaches whereas the face-directed synthesis favours the panel approach. Many of the same design principles behind formation of tetrahedral cages also apply to cubic cages. One point of note, whereas many tetrahedral cage structures intrinsically have chirality, cubic cages more usually form achiral complexes. Thomas and co workers[43] prepared an M_8L_{12} complex from a *fac* capped tritopic Ru^{II} vertex and 4,4' bi-pyridine bridging ligand, which gave an achiral edge-directed cube through the molecular library approach[11,31].
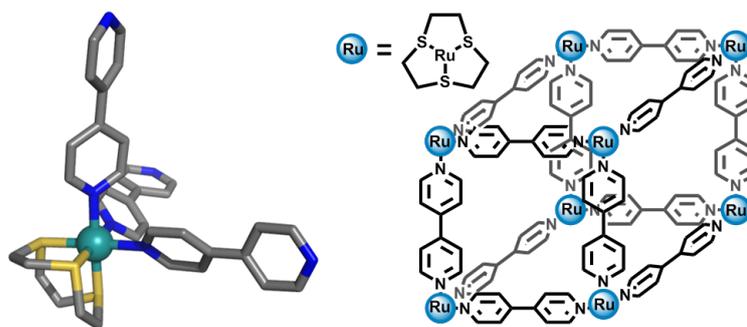
Fig. 1.24   Edge-directed cube (left) crystal structure of ML_3 vertex, (right) structure of cube[43]

Fujita and co workers[44] created a face-directed M_6L_{12} cube with a 4-coordinate square planar Pd^{II} ion in each face. Each metal is connected by one linker containing a 90° bend to give a large cage with the dimensions of 3 x 3 x 3 nm and volume of $27\,nm^3$. There are numerous further examples of cubic cages formed from various different approaches including Brisbois[45] face-directed $M_{12}L_6$ cube, an inverse to Fujita's structure, and Ward's[46] edge-directed M_8L_{12} cages created through the symmetry approach and described in more detail later in this chapter.
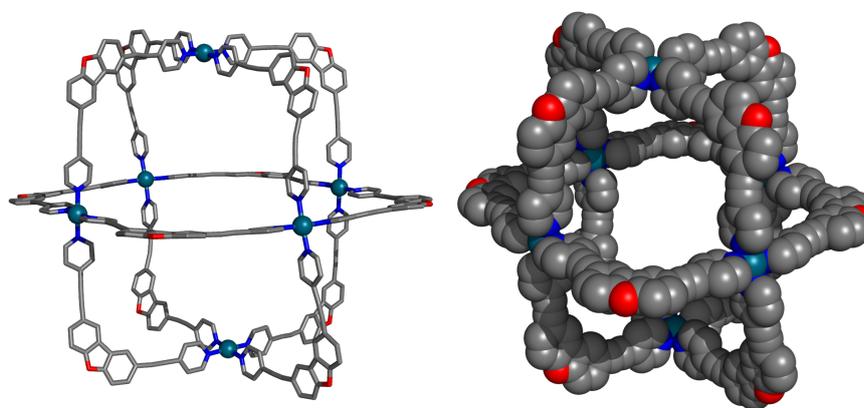
Fig. 1.25   Fujita face-directed $M_{12}L_6$ cube[44]

1.3.3   Larger Systems

With the development of design principles within supramolecular chemistry over the last few decades, the structures of simple polyhedral complexes can, in theory, be predicted and designed. In practice however ligands and the metal-ligand bonds are often too flexible for predictability to be possible. This is ever more apparent as the structures get larger with those containing more than 50 subunits becoming increasingly complex and their formation nearly impossible to predict. Fujita and co-workers have made some

of the largest supramolecular cage complexes known, with the current largest example being an $M_{24}L_{48}$ cage containing 72 components[47].
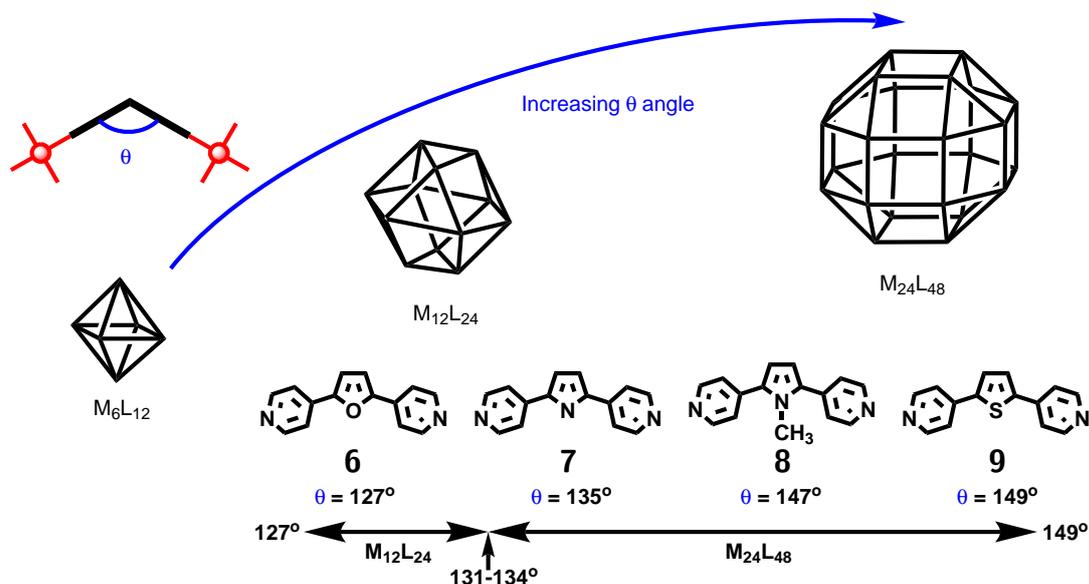
Fig. 1.26 Example highlighting that structure size can be controlled through ligand bend angle[48]

The size of the cage complex formed was found to depend on the curvature of the ligand, with the flatter ligands giving the larger, pseudo-spherical assemblies[41,47]. A range of ligands with different curvatures were investigated. Ligand **8** has a similar bite angle to **9** and so unsurprisingly forms the same larger $M_{24}L_{48}$ complex. Ligand **7** has a bite angle of 135°, around the angle where the smaller complex is expected to be favoured, surprisingly ligand **7** formed exclusively the $M_{24}L_{48}$ complex rather than the smaller $M_{12}L_{24}$ or a mixture of both. This shows that careful manipulation of a single parameter can result in predictably different structures but as the size and complexity of these complexes continues to increase, it is apparent that even very small changes can result in amplified differences within resultant structures[47].

1.4   Host-Guest Chemistry

Many of the self-assembled complexes prevalent throughout nature have particular functions or applications which the subunits alone could not perform. Only the assembled final structure has the desired functional behaviour, and in most of these cases host-guest interactions are vital[29]. Host-guest chemistry is in essence the interaction between a host

(in our case a large self-assembled structure with a central cavity) and a guest (normally a small organic molecule). A good example of host-guest chemistry in action can be found in the starch test for iodine. Starch is a polysaccharide consisting of a helical α-1,4 linked chain of glucose sugars (Fig. 1.27). The helicate encapsulates a number of linear iodine molecules to form a complex with a characteristic blue colour.
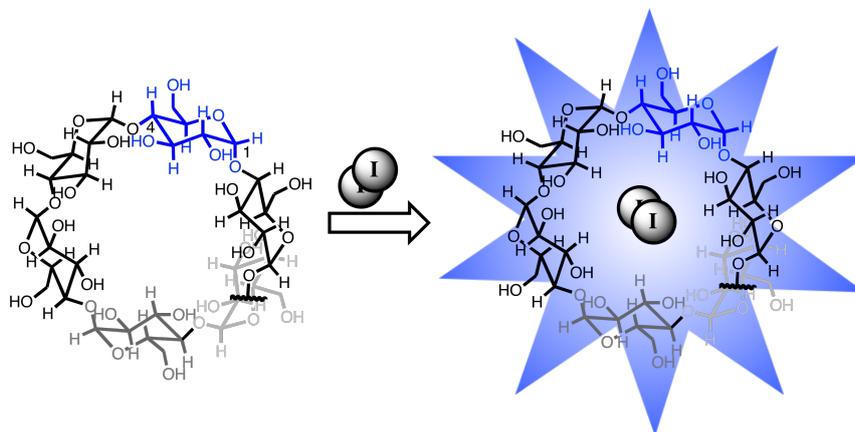
Fig. 1.27 α-1,4 glucose polymer can bind several I_2 guests to give a blue coloured complex

One of the first artificial host-guest complexes, which helped to establish this supramolecular chemistry field, was between 18-crown-6 (host) and a K^+ ion (guest). Six oxygen donors arranged in a circular scaffold provide a cavity with a matching size for the potassium ion, leading to a high binding affinity. On binding of the K^+ ion the crown ether requires little rearrangement, resulting in binding being supported by both the pre-organisation of host and the chelate effect.
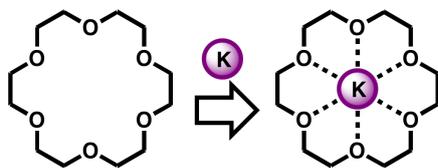
Fig. 1.28 18-crown-6 ether with potassium bound

The shape and size of the guest in relation to the host binding site are crucial for binding selectivity, and are utilized greatly throughout biology, particularly in enzymes. The "Lock and Key" principle is the most well-known model for understanding binding and, as the name suggests, the guest (the key) should match the binding site of the host (the lock) (Fig. 1.29). The model works well as an approximation but does not fully account for all known behaviour.

A better approach, based on the same underlying principle, is called the "induced fit" model in which the binding site does not necessarily match the shape of the guest initially, but on binding, adapts to the guests shape (Fig. 1.29). This mechanism allows for the host to be more selective, not only by size and shape, but also the positioning of key functional groups needed for inducing the required shape change in the binding site.
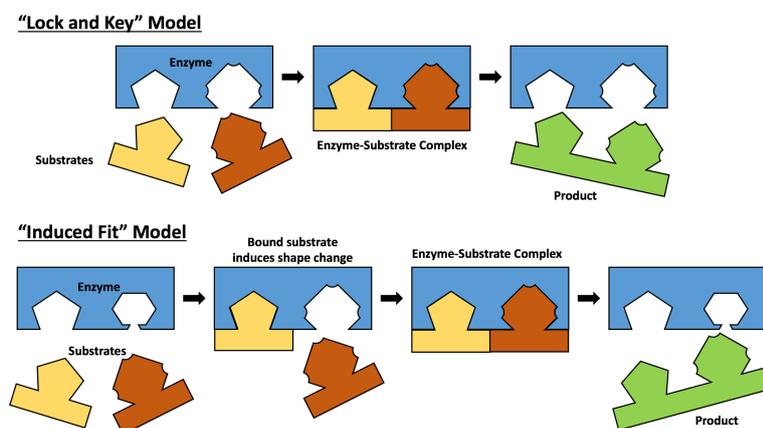
Fig. 1.29   Lock Key and Induced models of guest binding

1.4.1   Driving forces involved in guest binding

The interactions between different functional groups with the host binding site or cavity are key in understanding and quantifying host-guest interactions. In general these interactions are the same as the non-covalent interactions seen in self-assembly with contributions from; Van der Waals, charge-charge interactions, dipole-dipole interactions, hydrogen bonding, and $\pi-\pi$ interactions. The interactions involved in host-guest binding as well as self-assembly can be classed into three general groups; electrostatic interactions, induced interactions or solvent effects[49]. The former two interactions were discussed earlier (Chapter 1.1.1.1).

1.4.1.1   Solvent Effect

Solvent interactions play a vital role in self-assembly and host-guest chemistry and their importance is often overlooked. A clear example of the solvent's importance can be seen if one tries to model a single pair of DNA-double helix strands. Computer modelling in solvent gives a stable double helix as expected with the hydrogen bonding motifs stabilising the ordered structure. The formation of a DNA duplex also relies on the

stabilising contribution from the burying of the hydrophobic base pairs from surrounding solvent along with the stabilising interaction of solvent with the negatively charged sugar phosphate backbone. However removing the solvent, and therefore the stabilising contributions from the solvent, and modelling the duplex in a vacuum quickly results in the two DNA strands simply floating apart as the hydrogen bonding between nucleotides is not strong enough to overcome the electrostatic repulsion of the anionic backbones. The interactions of solvent with the host, and guest are both important in guest binding.

Hydrophobic Effect Hydrophobic, meaning to "dislike water", is a property that dominates intermolecular interactions within life and other aqueous media. The hydrophobic effect is what prevents oils from mixing readily with water and even causes surfactants to form membranes and vesicles. Despite being one of the most widespread non-covalent effects found in nature and one of the most studied solution effects, it remains one of the least understood. There have been many attempts to explain the cause of this effect however there is still an amount of controversy as to the cause[50–52].
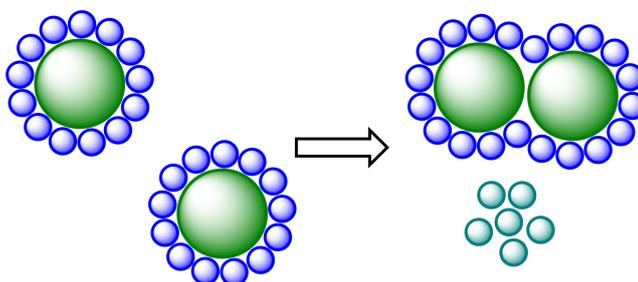
Fig. 1.30 Cartoon of the iceberg model showing the energy gain from aggregating two hydrophobic molecules (green) together allowing a lower number of ordered water molecules (blue)[53,54]

The first theory, and the most widely understood, is the "Iceberg Model" which explains this effect in terms of entropy. In water the individual molecules are relatively disordered. To form a solid the water molecules must order into a less dense structure; evident from the fact that ice floats on water. When a hydrophobic object is introduced into water, the molecules cannot bond with the object and instead form a crystalline structure at the interface, in other words forms a so called "iceberg". Removal of the object allows the "iceberg" to melt, an outcome which is favoured as the system becomes more disordered. This model fits well with explaining aggregation. If you take two hydrophobic objects, they both have an "iceberg" shell surrounding them. By aggregating

the two objects together the water can form a single "iceberg" around the two objects which has a smaller surface area compared to the objects having individual shell casings, with more water molecules freed, increasing disorder[53,54].

An alternative theory is the "high energy water" model. This model is based on enthalpy losses, born from supramolecular examples. Curcubiturils typically have small bowl shaped cavities capable of hosting a few "high energy water" molecules (molecules that cannot form the desired hydrogen bonds). Unlike in the bulk, these "high energy waters" are heavily restricted to which direction that hydrogen bonds can be formed, if at all. Upon displacement by a guest into the bulk, the "high energy water" can form the normal hydrogen bonds and therefore form a more stable system. Examples of this model are limited and generally occur in small cavities[51,52].

For both the described theories, and others, there is theoretical and practical data that both support and contradict the models with some examples simultaneously proving and disproving them[50–52]. When considering a larger hydrophobic cavity, within which there is enough space for small groups of water to form hydrogen bonding networks, the "iceberg" theory appears to better explain the hydrophobic effect. There is even crystallographic evidence of an ice-like ordered network of water molecules within a cage cavity[55].
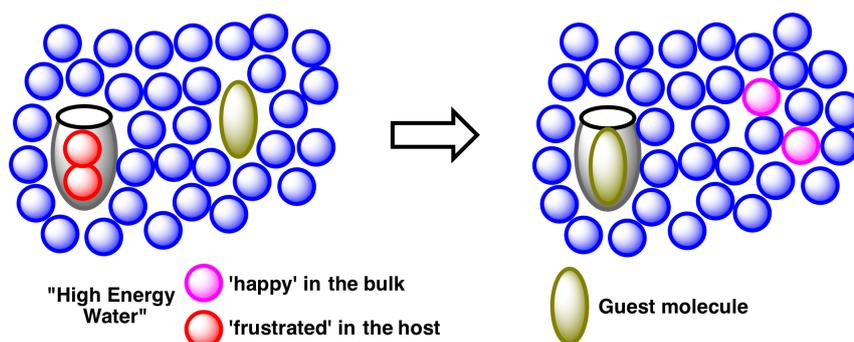
Fig. 1.31 Cartoon of the "high energy water" model showing how displacing frustrated water allows bonds to be formed with the bulk water[51,52]

In reality it is likely that the hydrophobic effect arises from a balance of both entropic and enthalpic contributions, with each term's contribution depending on the specific system being studied. Either way, while the cause is still disputed, the significance of the effect in aqueous solution is indisputable and often the most significant factor in both binding of guests and self-assembly itself.
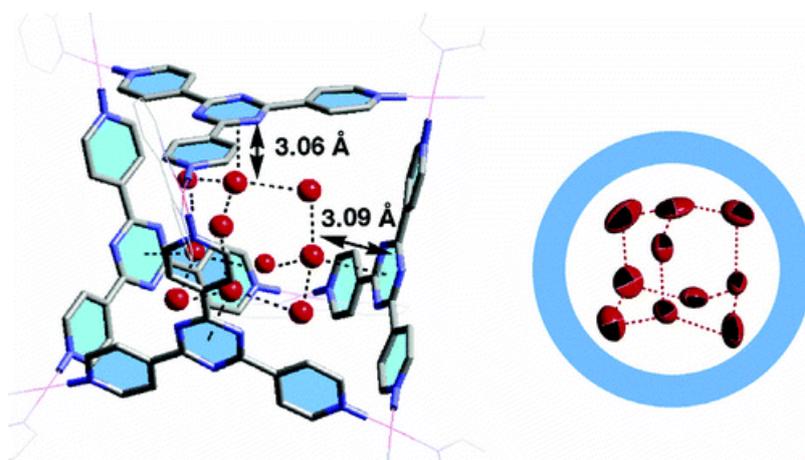
Fig. 1.32 X-ray structure of a cage complex containing an "iceberg". Reprinted with permission from ref[55]. Copyright 2017 American Chemical Society

1.4.2 Determining Binding Constants

The binding constant, K, is a value that describes the equilibrium associated with formation of a host-guest system, as described in equation (1.3), and is related to thermodynamic parameters, such as Gibbs' free energy, entropy and enthalpy. The relationship between these three thermodynamic parameters was described earlier in equation (1.2). The balance between the entropy and enthalpy determines the free energy of the system and therefore the binding strength of the guest. Though the enthalpy and entropy can be determined through a Van't Hoff plot of the change in equilibrium constant at various temperatures, any error in one of the terms (entropy or enthalpy) is magnified and corrected for in the other term. This can lead to the phenomenon of entropy-enthalpy compensation. As a result it is better to only look at the overall free energy of the system therefore only determine ΔG and K_a as these can be measured with a higher certainty.

$$\text{Host [H]} + \text{ Guest [G]} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{ Host-Guest Complex [HG]} \tag{1.3}$$

$$K_a = \frac{k_{on}}{k_{off}} = \frac{[\text{HG}]}{[\text{H}][\text{G}]} \qquad \Delta G = RT \ln K_a \tag{1.4}$$

Equation (1.4) defines the equilibrium of a 1:1 binding system. Though the equation can be adapted for other binding modes, 1:1 guest:host is the simplest system to solve as well as being the most common binding mode studied. Defining the host and guest is arbitrary but for simplicity the host will be defined as the supramolecular assembly. From equation (1.3) the k_{on} binding rate and the k_{off} un-binding rates can be seen. By

putting these terms together in equation (1.4) we can relate these kinetic parameters to a single constant that describes the equilibrium. At equilibrium the concentrations of free host, guest and complex can also be related to the equilibrium constant.

$$[H]_0 = [H] + [HG] \qquad [G]_0 = [G] + [HG] \tag{1.5}$$

$$f_{HG} = \frac{[HG]}{[H]_0} = \frac{K_a[G]}{1 + K_a[G]} \qquad [HG] = \frac{[H]_0 K_a[G]}{1 + K_a[G]} \tag{1.6}$$

Though the starting concentration of host $[H]_0$ and guest $[G]_0$ in a system is known, it is not normally possible to measure the concentration of free host $[H]$, free guest $[G]$ or host-guest complex $[HG]$ directly. As matter is being conserved throughout we can describe the starting concentration of host as equal to the sum of free host and host-guest complex at equilibrium to give equations (1.5). Combining the mass balance equations (1.5) with (1.4) gives equation (1.6). Inserting $[HG]$ into the second part of (1.5) and rearranging for $[G]$ gives the quadratic equation (1.7)

$$[G]^2 - [G]\left([G]_0 - [H]_0 - \frac{1}{K_a}\right) - \frac{[G]_0}{K_a} = 0 \tag{1.7}$$

$$[HG]^2 - [HG]\left([G]_0 + [H]_0 + \frac{1}{K_a}\right) + [H]_0[G]_0 = 0 \tag{1.8}$$

$$ax^2 + bx + c = 0 \qquad x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{1.9}$$

A similar insertion and rearrangement can be made to give the equation (1.8). With these two equations it is possible to solve for $[HG]$ and $[G]$ simply by knowing the initial concentrations of the system and the equilibrium constant K_a. Equally if $[HG]$ or $[G]$ can be determined then K_a can be calculated. Therefore any experiment that allows a physical property to be related to one of these concentrations found at equilibrium or the change in physical property to be related to the change in, for example, the host-guest complex concentration, can be used to determine the binding constant (equation (1.10)).

$$\Delta Y = Y_{\Delta HG}\left(\frac{[HG]}{[H]_0}\right) \qquad Y = \text{physical property being observed} \tag{1.10}$$

Typically the measurement of binding constant K_a is carried out on a system with a fixed host concentration and a varied guest concentration. The resulting change in physical property can be monitored through one of many spectroscopic techniques:

commonly NMR, UV-vis or Fluorescence spectroscopy. The observed change is plotted against change in guest concentration to give a binding isotherm which, once fitted, gives the binding constant. The different techniques, used to monitor different physical properties, have a number of different advantages and disadvantages. NMR spectroscopy, for example, is a slower technique and requires higher concentrations than UV-vis or fluorescence spectroscopy. The pros and cons of different techniques have been discussed at length in various reviews[56,57].
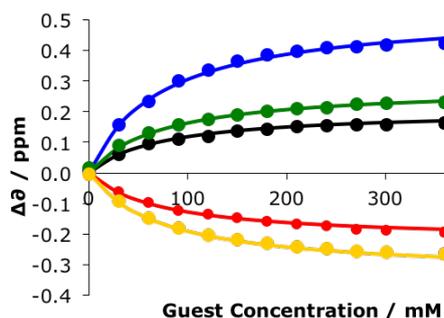
Fig. 1.33 An example of a binding isotherm

NMR spectroscopy is able to provide more information from a single titration than just the binding constant. The way the guest binds to the host can often be seen by observing which NMR peaks shift upon binding. It can, depending on the rate of guest exchange, be possible to directly observe [H] and [HG]. Depending on the timescale of the host-guest exchange, NMR titrations can be grouped into two categories: Fast exchange or slow exchange. In fast exchange the guest is exchanging in and out of the host faster than can be observed on the NMR timescale. In a similar way to a photo of a fast moving object, you observe an average of both the guest being free in solution and bound within the host. As a result, with increasing guest concentration, the observed peak for the host shifts steadily towards the complex peak (Fig. 1.34). Slow exchange is where the guest moves in and out of the host at a rate slower than the NMR timescale. In this case the host peaks and the host-guest complex peak are observed as individual peaks and separately quantified by integration (Fig. 1.34).

Once the host-guest concentration is determined the guest concentration can be determined from equation (1.5). These values can then be used directly in equation (1.4) to give the binding constant. Fast exchange results must be plotted to give a binding isotherm which is then fitted to indirectly determine the binding constant.
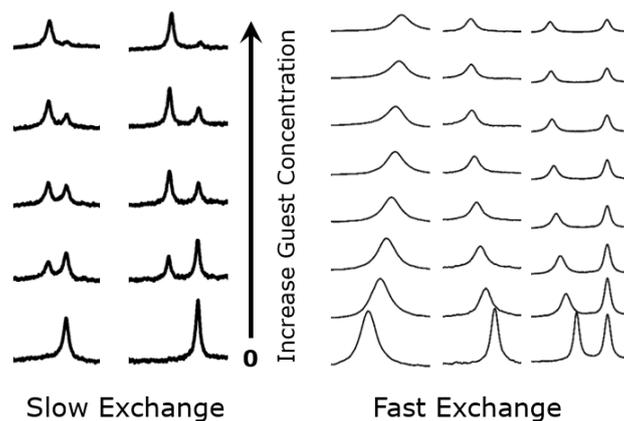
Fig. 1.34 Examples of changes in host NMR signals with increasing guest concentration when in: (left) slow exchange, (right) fast exchange

1.5 Functionality and Applications

The last couple of decades have seen a dramatic increase in the number of applications for discrete three-dimensional molecular structures (cages), from binding and sensing guests to catalysis and drug delivery. Even so the field has a long way further to go to truly start matching the breath of applications and uses that nature has developed for biological assemblies. Establishing and understanding the host-guest chemistry of cages is key to the future development of these applications.

1.5.1 Molecular Flask
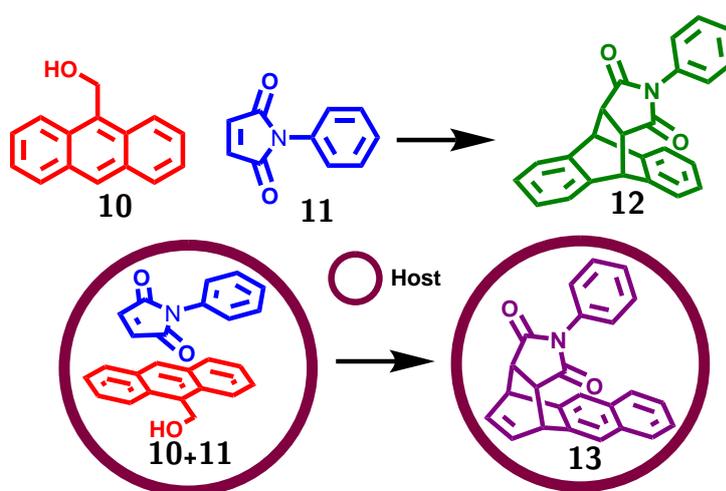
Scheme 1.1. Rebek's cage catalysed Diels-Alder reaction[58]

Research into molecular flasks has the eventual goal of mimicking the activity of enzymes but through new non-natural reactions[59]. The non-covalent interactions used

in self-assembly and guest binding can be used to effect catalysis of reactions inside the host cavity. Isolation of reactants from the bulk solvent leads to an increased local concentration and as a consequence an increase rate of reaction, whilst preorganization can be used to support and stabilize unusual conformations of bound guests in transition states[60].

Aside from preorganization, the size and shape of the internal cavity can be used to generate stereoisomers of products that are otherwise unobtainable. As a result of the spatial positioning within the host, a particular stereoisomer, known as a social isomer, can be made from two guest co-bound within the host in the desired arrangement[61].
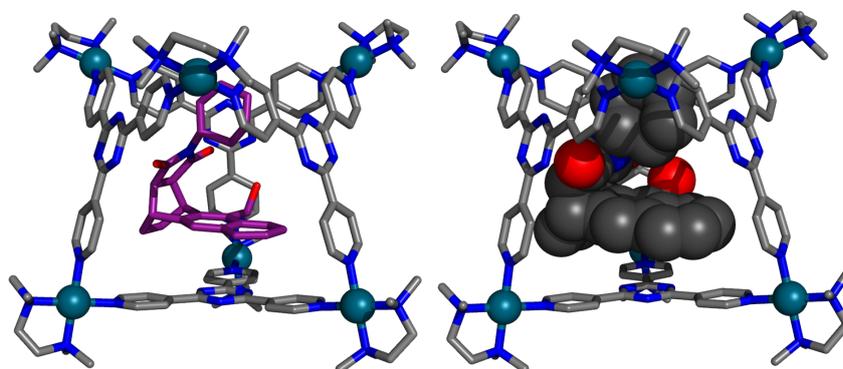
Fig. 1.35 Crystal structure of Fujita's cage containing product **13**[58]

Rebek and co-workers produced one of the first examples of a self-assembled complex that could function as a molecular flask, producing a 200 fold increase in the rate of a Diels-Alder reaction between a pair of diene and dienophile guest at room temperature[62]. Fujita and co-workers likewise showed that the sterics within a molecular flask can lead to unusual regio- and stereo-selective Diels-Alder reactions (Scheme 1.1). Normally **10** and **11** will reaction to produce **12** however when within the sterically hindering environment of the cage cavity the different isomer **13** is produced instead[58]. Binding of the two different guests within the same host cavity is required, in both of the examples, for the Diels-Alder reaction to occur.

1.5.2 Encapsulation

There are many compounds that are reactive, explosive, or toxic compounds when in a free state. Typically white phosphorus, P_4, rapidly decomposes in air to phosphoric acid. When trapped within the cage the compound is completely stable as the decomposition

intermediate is too large to fit in the cavity. Addition of a competing guest, in this case benzene, expels the P_4 molecule where it quickly decomposes (Fig. 1.36)[63].
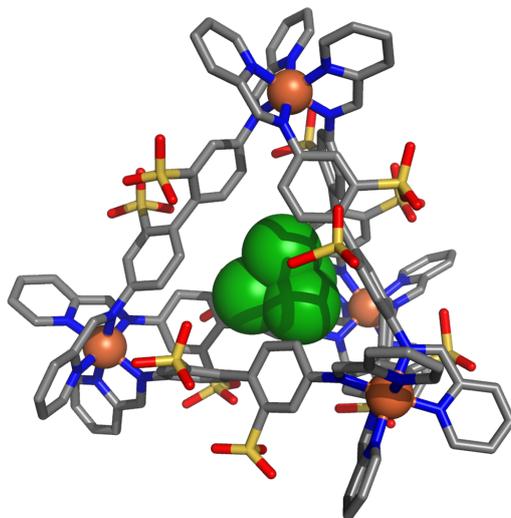
Fig. 1.36 Crystal structure of Nitschke's tetrahedral host cage showing P_4 encapsulation[63]

1.5.3 Drug Delivery

An obvious potential application of a cage-guest systems is to use the cage as a capsule to transport and deliver drug molecules. Targeted transport could be achieved through external functionalisation of the cage with protein tags, while delivery could be achieved by applying a stimulus for guest realise at the delivery site. Clear advantages to using cages as delivery devices are the ability to administer hydrophobic guests (which are normally insoluble) and keeping the cytotoxic guests trapped safely until they are released at the desired site.

The Crowley group have demonstrated a drug delivery capable system using an M_2L_4 Pd^{II} cage that is able to bind cis-platin[64]. This is a drug that is widely used to treat a variety of different cancers from ovarian cancer to lymphomas. The addition of a competing ligands, which can coordinate to the Pd^{II} ions, causes the cage to disassemble and release the guest. Though observable in an ^1H-NMR spectra this system is a long way from functioning *in vivo*.

Ruthenium based compounds are being developed as alternatives to some platinum based cancer drugs[65]; RAPTA-C is one such example[66]. While RAPTA-C is only weakly cytotoxic *in vitro*, the compound is selective and effective against metastasis at high
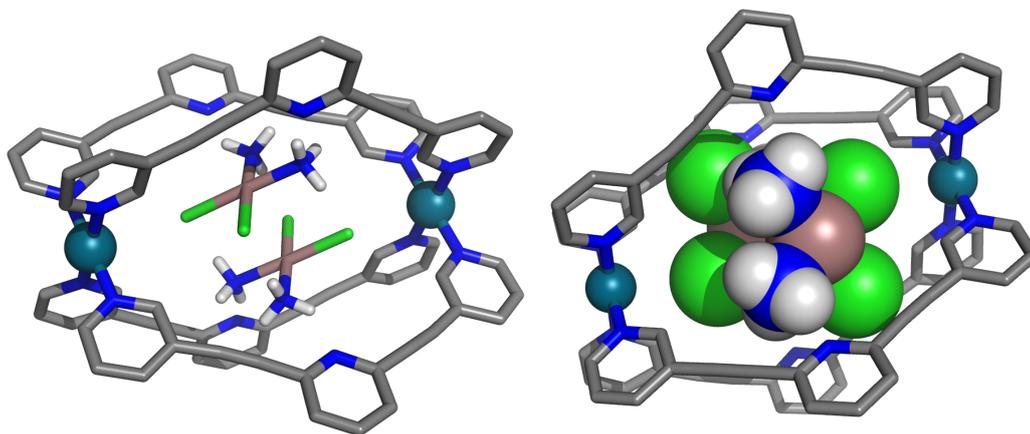
Fig. 1.37 Crystal structure of Crowley M_2L_4 complex with encapsulated cis-platin[64]

dosages *in vivo*. Therrien and co-workers showed that the use of water soluble cages as drug delivery vectors encapsulating RAPTA-C gave a great than 10 fold increase in cytotoxicily *in vitro*[67].

1.6 Ward Group Cages
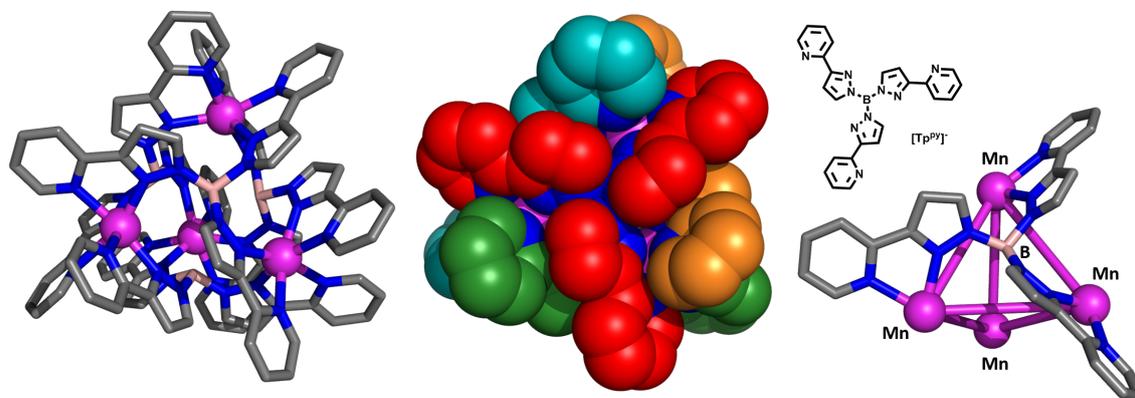
Fig. 1.38 Crystal Structure of tetrahedral complex $[Mn_4(Tp^{Py})_4](PF_6)_4$ (left) stick model, (middle) spacefill model, (right) cut-away stick model highlighting one ligand[68]

The family of cages of particular interest within the Ward Group are those based on ligands with terminal pyridine-pyrazole chelating groups. These complexes can be described as being constructed from the symmetry-approach, however the original cage complexes were discovered serendipitously. One of the first significant complexes within this family to be constructed used a tris(pyrazolyl)borate ligand $[Tp^{py}]^-$ shown in Fig. 1.38 (left)[68]. The ligand can bind either to a single metal ion, $[Co(Tp^{py})]^+$ or

span three metal centres to form a tetrahedral M_4L_4 cage complex, with each ligand binding three metals in a face-capping mode and each metal binding three ligands.

1.6.1 Tetrahedral Cages

The free-rotation around the pyrazole-borate bond gives the complex the flexibility to form the small M_4L_4 tetrahedral complex (Fig. 1.38) which follows the design rules described earlier with C_3 symmetry within both the ligand, and at the metal binding site[68,69]. Reducing the symmetry of the ligand to C_2 and introducing a spacer group, giving ligand $\mathbf{L}^{o\mathbf{Ph}}$, leads to a M_4L_6 tetrahedron (Fig. 1.39) with the metals in the vertices and bridging ligand along the edges of the cage, in keeping with the design rules.
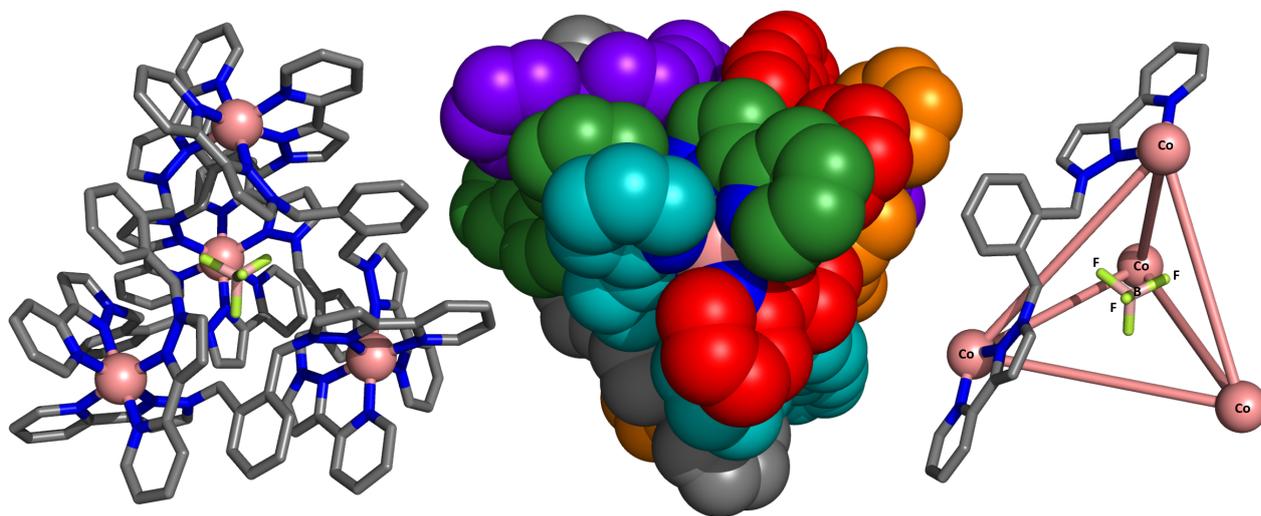
Fig. 1.39 Crystal structure of tetrahedral complex $[Co_4(L^{oPh})_6(BF_4)](BF_4)_7$ (left) stick model, (middle) spacefill model, (right) cut-away stick model highlighting one ligand and trapped anion[70]

Additionally the flexibility of the ligands allows helical twisting along these edges to give the required angle between the ligands coordination vectors. As a result each complexes is chiral, but they are produced as a racemic mix of the homochiral cages. Additionally this M_4L_6 tetrahedron was found to have a cavity the perfect size, shape and charge to bind a tetra-fluoroborate (BF_4^-) counter ion with the terminal fluorides interacting via CH•F interactions with the saturated methylene linkages within the ligand. A similar encapsulation is seen with a perchlorate counter-ion. The complex was found to not allow exchange of the incorporated guest on the NMR timescale, as the guest anion is completely encapsulated.

Variation of the spacer group between the chelating units quickly resulted in a series of different ligands. Collectively the ligands are all inherently highly flexible due to the saturated methylene linker between the pyridine-pyrazole terminals and the spacer group. Initially introduced for easy of synthesis, the flexibility that arises is key to allowing the formation of the resulting coordination cages. The disadvantage of the flexibility however is the inability to rationally design the structures.
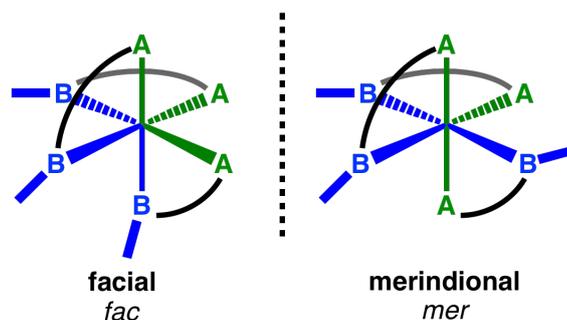
Fig. 1.40 Diagram showing the two possible octahedral metal geometries: (left) facial *fac* (right) meridional *mer*

Metal ions used in complexation were chosen to have six-coordination octahedral geometry. In combination with bidentate ligands the metals can bind to three ligands in either a facial (*fac*) or meridional (*mer*) geometry. As a result the complexes require a two-metal to three-ligand ratio (2M:3L) to be able to satisfy the "maximum site occupancy principle". This principle proposes that the highest stability is obtained when all ligand binding sites and metal coordination sites are saturated. The specific metal ion is selected based on additional properties that it provides to the structure, for example many of the complexes are made with Co^{II} which allows for study of the complexes using paramagnetic ^{1}H-NMR spectroscopy. This spreads the resulting peaks over a 200 ppm range, -100 to 100 ppm, allowing individual peaks to be discerned more easily.

Ligands $\mathbf{L}^{2,3naph}$ and $\mathbf{L}^{2,3anth}$ were both prepared as analogues of ligand $\mathbf{L}^{o-Ph}$. Of the ligands shown in Fig. 1.41 with C_2 symmetry, these three analogs are the only ones that can form mononuclear tetrahedral complexes because the two pyridine pyrazole units are close enough together to chelate to a single metal ion. In the M_4L_6 cages, all three structures form as a racemic mixture of homochiral cages with all 4 vertices having *fac* tris-chelate geometry and are homochiral (either $\Delta\Delta\Delta\Delta$ or $\Lambda\Lambda\Lambda\Lambda$). The structures all contain a single trapped anion at the centre of the cage structure as well as extensive
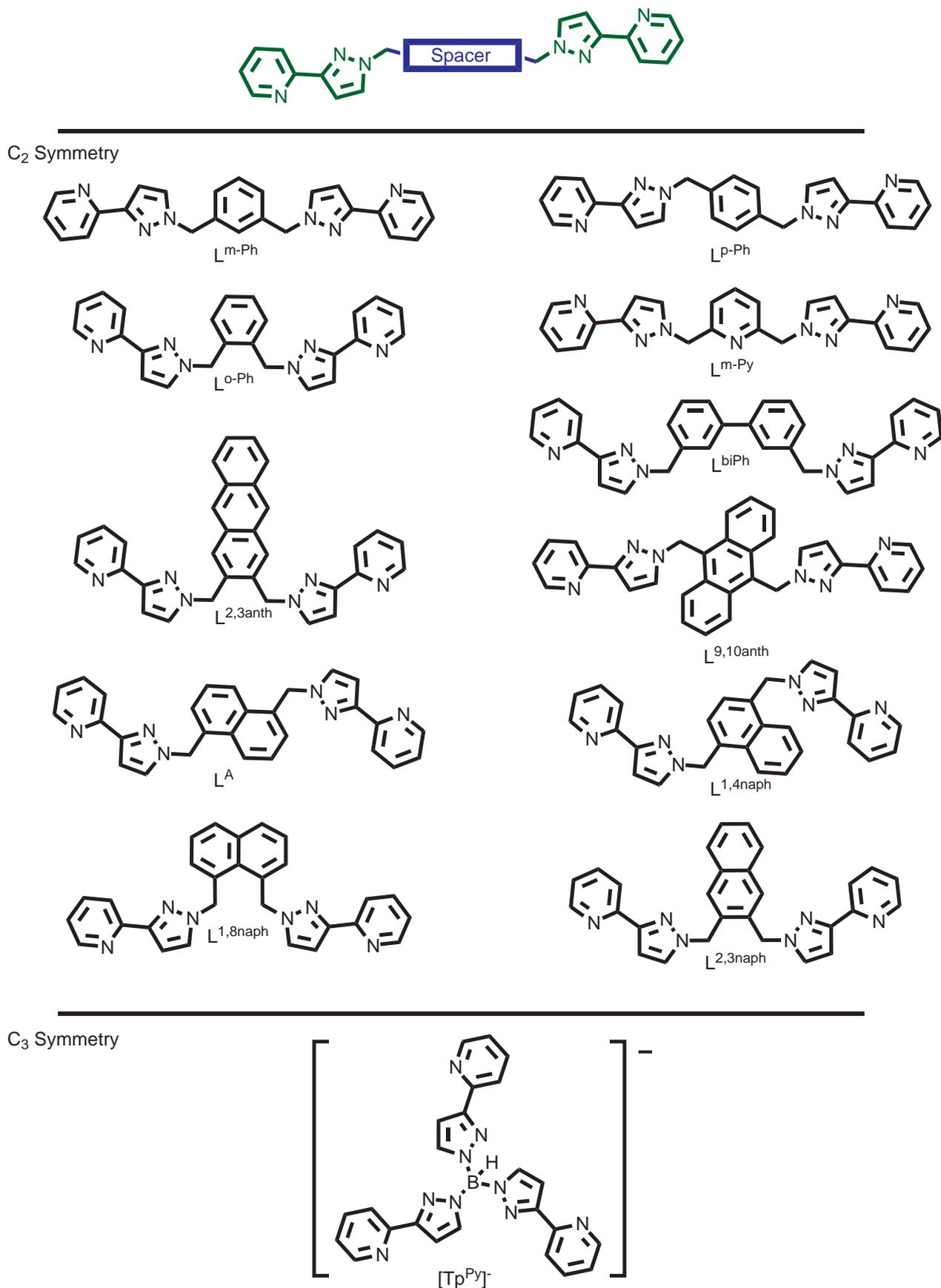
Fig. 1.41 Various ligand used within the Ward Group to prepare coordination polyhedra
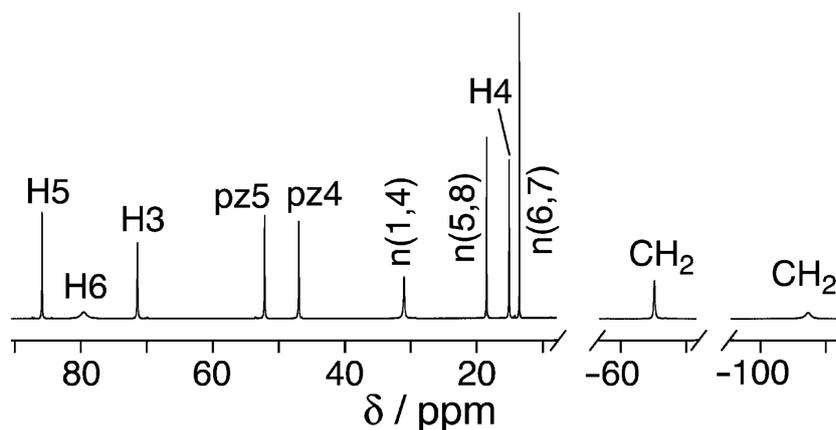
Fig. 1.42 Paramagentically shifted ^{1}H-NMR of tetrahedron $[Co_4(L^{2,3naph})_6(BF_4)](BF_4)_7$. Reproduced from ref[46] with permission of The Royal Society of Chemistry

aromatic π stacking between the ligands around the periphery of the cage (Fig. 1.43)[71]. On each of the cage faces a pyridine-pyrazole / spacer / pyridine-pyrazole interaction can be seen as a three-component stack. Further, the luminescence of the naphthalene group in the complex containing ligand $\mathbf{L^{2,3naph}}$ can be seen to have been red-shifted due to the π stacking interactions[72].
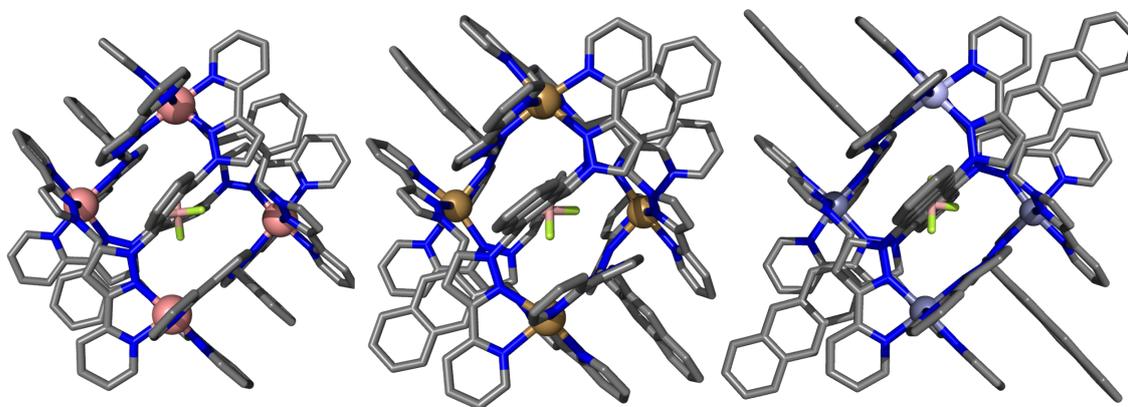
Fig. 1.43 Crystal structure of tetrahedral complexes (left)[70] $[Co_4(L^{oPh})_6(BF_4)](BF_4)_7$ (middle)[71] $[Cd_4(L^{2,3naph})_6(BF_4)](BF_4)_7$ (right)[71] $[Zn_4(L^{2,3anth})_6(BF_4)](BF_4)_7$ showing the π-π interactions between ligands, around the faces

Ligand $\mathbf{L^{BiPh}}$ is similar to those discussed above that form the tetrahedral cages, but is extended in length and contains an additional rotatable bond. When complexed with a metal ion a M_4L_6 tetrahedral cage complex[73] is again formed but this is interestingly different from the previous examples. The metal coordination geometries differ with three-*mer* and one-*fac* tris-chelate vertices (Fig. 1.44). As a result the structure is a "stretched" tetrahedron with a single C_3 axis through the *fac* vertex. The structure,
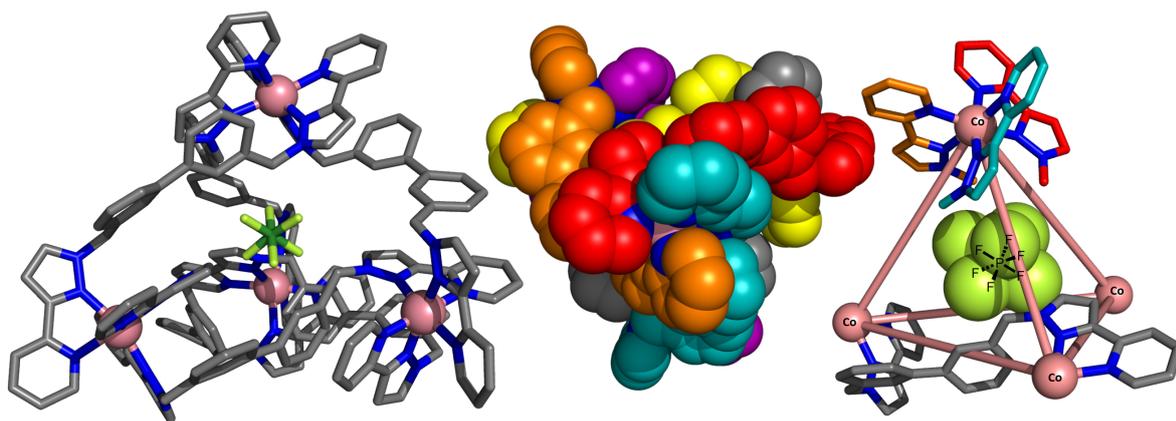
Fig. 1.44 Crystal structure of tetrahedra complex[73] $[Co_4(L^{BiPh})_6(PF_6)](PF_6)_7$ (left) stick model, (middle) spacefilled model looking at *fac* vertex, (right) cut away model highlighting a single ligand and *fac* vertex

containing a larger cavity than the previous tetrahedron, can bind a range of larger anions such as hexafluorophosphate or iodide. As the cavity is larger than the small cages formed by L^{o-Ph}, $L^{2,3naph}$, $L^{2,3anth}$, the bound anions can now exchange with the external ones on the NMR timescale.

1.6.2 Cubic Cage
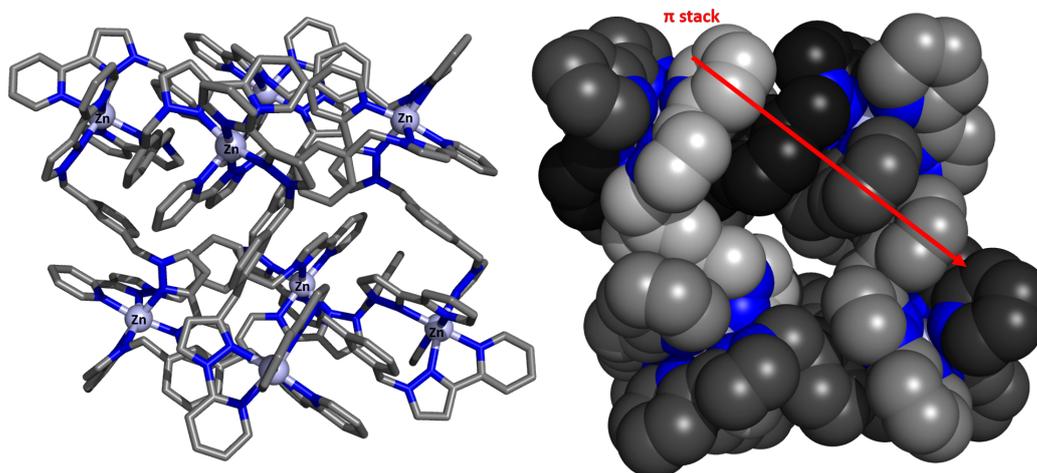
Fig. 1.45 Crystal structure[74] of $[Zn_8(L^{m-Ph})_{12}](BF_4)_{16}$

Ligands L^{m-Ph} and L^{m-Py} are near identical and both have a widened angle between the substituents on the aromatic spacer group due to the meta-substitution pattern. The nitrogen atom of the pyridine spacer is not involved in binding, resulting in both ligands assembling in the same manner with octahedral metal ions to form a distorted cubic

cage[74]. Like the tetrahedra, these cubes have extensive aromatic π stacking around the outer surface (Fig. 1.44).
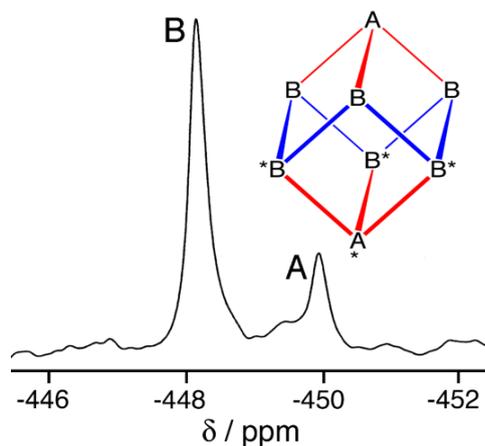
Fig. 1.46 $^{113}Cd-NMR$ of $[Cd_8(L^A)_{12}](BF_4)_{16}$ showing ratio of *fac* to *mer* vertices. Reproduced from ref[46] with permission of The Royal Society of Chemistry

The phenyl spacer group was replaced by a 1,5-dimethyl naphthalene spacer to give ligand L^A that assembles with octahedral metal ions to form a similar cubic structure. The cubic structure formed by L^A is less distorted than the two previous cubic structures while containing the same extensive π stacking (Fig. 1.47) with a single five component stack across each of the six faces. The five-component stacks each involve three pyridine-pyrazole units (electron deficient due to coordination of the metal) alternating with two naphthalene spacers (electron rich) to give A/D/A/D/A stacks (A = acceptor, D = donor).
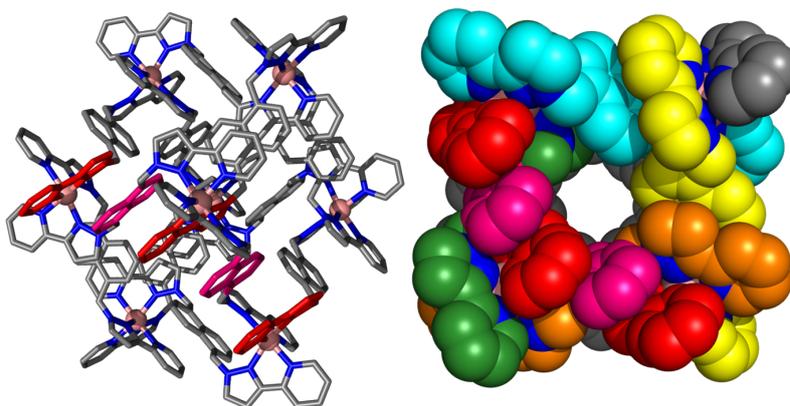
Fig. 1.47 Crystal structure of $[Co_8(L^A)_{12}](BF_4)_{16}$ highlighting π stacking (red/pink)

All three cubic structures contain two *fac* tris-chelate vertices at opposite corners of the cage complex and are connected through a mesh of six *mer* tris-chelate vertices. The 1:3, *fac* to *mer* ratio is evident in the ^{113}Cd-NMR spectrum where the differing

configurations give individual peaks (Fig. 1.46). The cage has an inversion centre: One *fac* vertex has a Λ optical configuration while the opposite *fac* vertex has Δ optical configuration and each form an $M_{fac}L_3(M_{mer})_3$ array (Fig. 1.48). The left-handed and right handed arrays come together, connected by the remaining ligands, to form the completed cubic structure which is achiral. The cubic cage contains an S_6 improper rotation axis with the two *fac* vertices being related through an inversion centre.
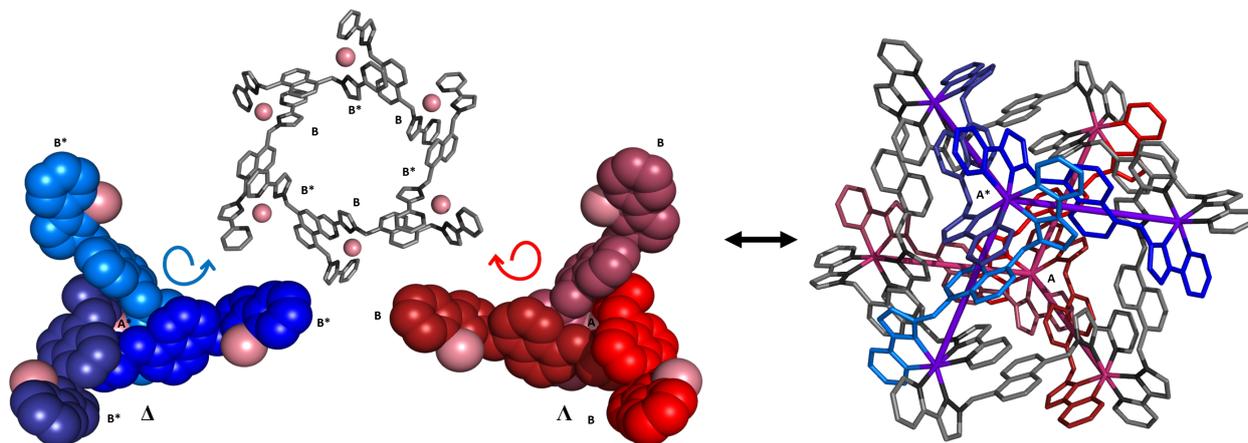
Fig. 1.48 (left) Each $M_{fac}L_3(M_{mer})_3$ array with the inter-connecting spiral of ligands, (right) completed cube highlighting the two arrays (shades of red vs shades of blue) with the inter-connecting ligands (grey)

Substituting the napthtyl spacer group with an anthracenyl group gives ligand $\mathbf{L^{9,10anth}}$ which (with an octahedral metal) also forms an M_8L_{12} cubic cage complex[75], but this differs significantly from the previous cubic complexes. In contrast to earlier the cubic complex formed from ligand $\mathbf{L^{9,10anth}}$ does not have any stabilization from $\pi - \pi$ stacking interactions which results in a different, less compact structure (Fig. 1.49). The cage consists of two M_4L_4 square helical arrays joined together by four "pillars". Unusually all eight metals have a *mer* tris-chelate coordination geometry. Though this ligand forms an elegant structure in the solid state with a comparatively large cavity, the exclusion of $\pi - \pi$ stacking results in the structure not being stable in solution.

1.6.3   Other Polyhedral Cages

Widening the angle between substituents again on the phenyl spacer by using para-substitution to giving ligand $\mathbf{L^{p-Ph}}$, allows larger cage structures to be formed depending on the metal ion used. The use of Cu^{II} gives the smallest architecture of an
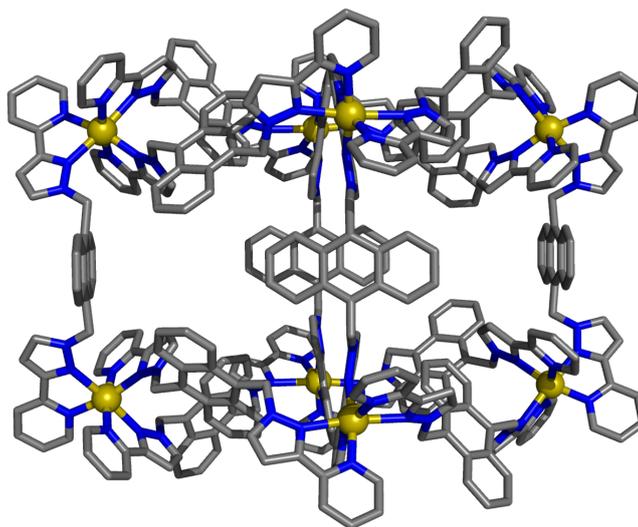
Fig. 1.49 Crystal structure[75] of $[Cu_8(L^{9,10anth})_{12}](BF_4)_{16}$

M_6L_9 triangular prism with two cyclic M_3L_3 arrays being held together with three ligand pillars. Each array consists of three *fac* vertices with the same optical configuration as well as a network of $\pi - \pi$ stacks, one along each edge. The top and bottom arrays have the opposite optical configurations resulting in the complex being achiral.
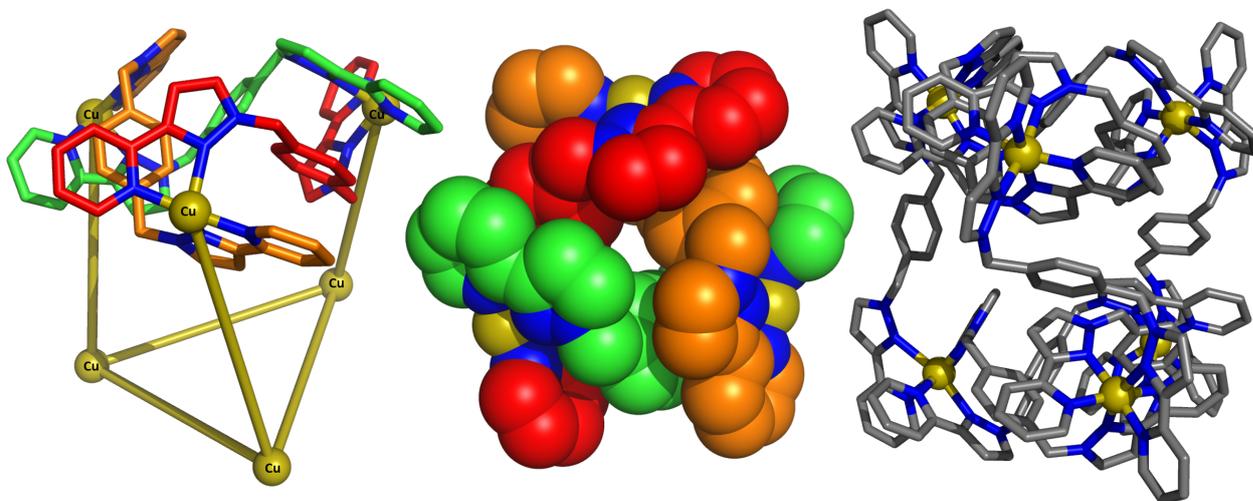
Fig. 1.50 Crystal structure[76] of $[Cu_6(L^{p-Ph})_9](BF_4)_{12}$ (left) stick model cut away structure highlighting ligands within the M_3L_3 array, (middle) spacefilling model of the same M_3L_3 array, (right) stick model complete structure

The use of Ni^{II} with ligand $\mathbf{L}^{p-\mathbf{Ph}}$ gives a cubic M_8L_{12} structure with identical features to those described above for $[Co_8(L^A)_{12}](BF_4)_{16}$ complex. This includes the $\pi - \pi$ stacks across each face of the cube (Fig. 1.51) as seen in most of the earlier cubic cage structures.

In contrast the use of Zn^{II} forms the $M_{16}L_{24}$ cage complex, the largest known cage within the Ward group. The tetra-capped truncated-tetrahedron consists of two
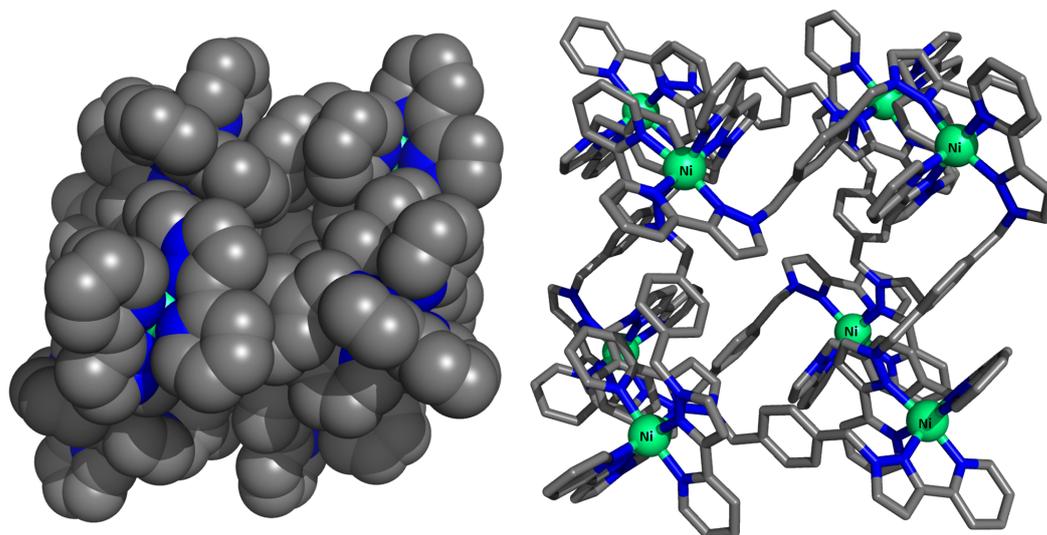
Fig. 1.51 Crystal structure[76] of $[\text{Ni}_8(\text{L}^{p-\text{Ph}})_{12}](\text{BF}_4)_{16}$ (left) spacefill model, (right) stick model

sub-assemblies. The first is the same M_3L_3 cyclic array as seen with the Cu^{II} complex above. Four of these arrays are connected through four ML_3 "caps". The 24 ligands are all entangled together in an extensive network of π stacks in the familiar pyridine-pyrazole and spacer alternating sandwich.
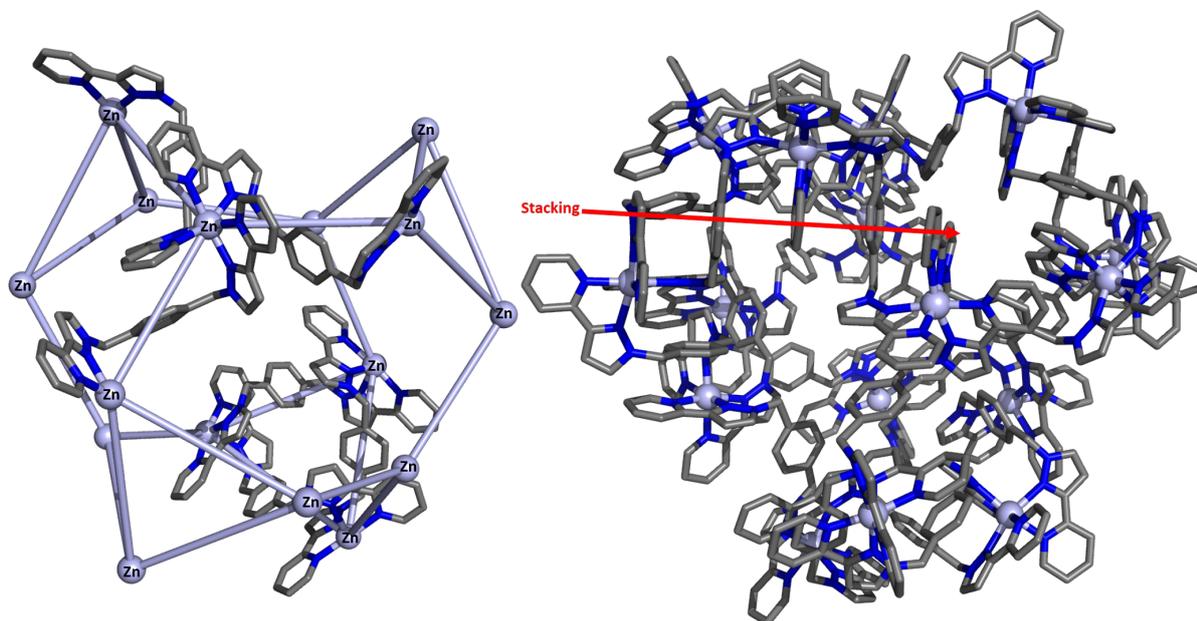
Fig. 1.52 Crystal structure[77] of $[\text{Zn}_{12}(\text{L}^{p-\text{Ph}})_{24}](\text{BF}_4)_{32}$ (left) cut-away model showing stick model of ligands in the ML_3 "caps" at the top and M_3L_3 cyclic array at the bottom, (right) stick model of the complex highlighting one of the π stacks

Though the ligand $\mathbf{L}^{p\text{-}\mathbf{Ph}}$ forms the three described structures with Cu^{II}, Ni^{II} and Zn^{II} in the solid state, none of these complexes appear to be present in solution. The use of Cd^{II} produces the same tetra-capped truncated-tetrahedron as Zn^{II} but the complex

is more stable and is evident in solution ^{1}H-NMR spectra. Even so the complex is not stable for a long period and within a few weeks in solution, the complex rearranges into smaller $[M_6L_9]^{12+}$ units.

Ligand $\mathbf{L^{1,8naph}}$ simply formed a single $M_{12}L_{18}$ complex with a range of metal ions $[Cd^{II}, Co^{II}, Cu^{II}, Zn^{II}]$[78]. The formed truncated tetrahedron shares many structural features with the other complexes in this family. Yet again we see the incorporation of M_3L_3 triangular cyclic helical arrays, this time with bridging ligands joining the four arrays together. $\pi - \pi$ stacking between ligands is again present this time in seven-component sandwiches of alternating electron rich and electron poor units. All 12 metal centres have a *mer* geometry and are all of the same optical configuration resulting in a homo-chiral cage and again, as before, the bulk material is racemic. The cage complex contains a large open cavity which, in the solid state, contains four anions (either BF_4^- or ClO_4^-).
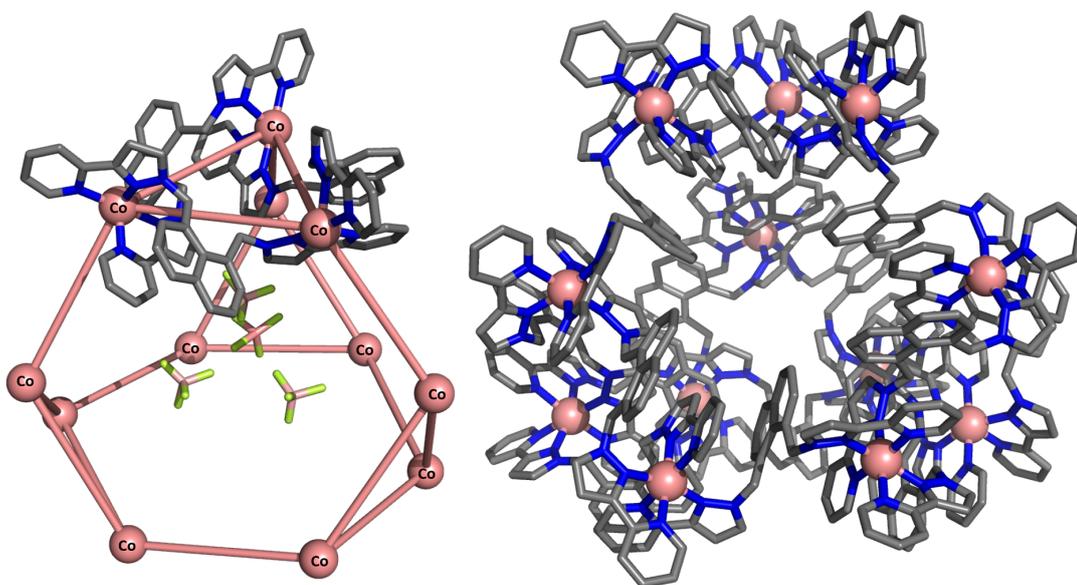
Fig. 1.53 Crystal structure[78] of $[Co_{12}(L^{1,8naph})_{18}](BF_4)_{24}$ (left) stick model, (right) cut away model showing the four counter ions and one of the ML_3 sub-assemblies

The final $\mathbf{C_2}$ symmetric ligand present in Fig. 1.41, ligand $\mathbf{L^{1,4naph}}$ is a clear analogue of $\mathbf{L^{p-Ph}}$. When complexed with Cd^{II} the resulting complex is a similar $M_{16}L_{24}$ tetra-capped truncated-tetrahedron. This structure contains the same M_3L_3 triangular array and ML_3 "cap" components as seen with the combination of Cd^{II} and $\mathbf{L^{p-Ph}}$, but is more stable in solution due to the improved $\pi - \pi$ stacking.
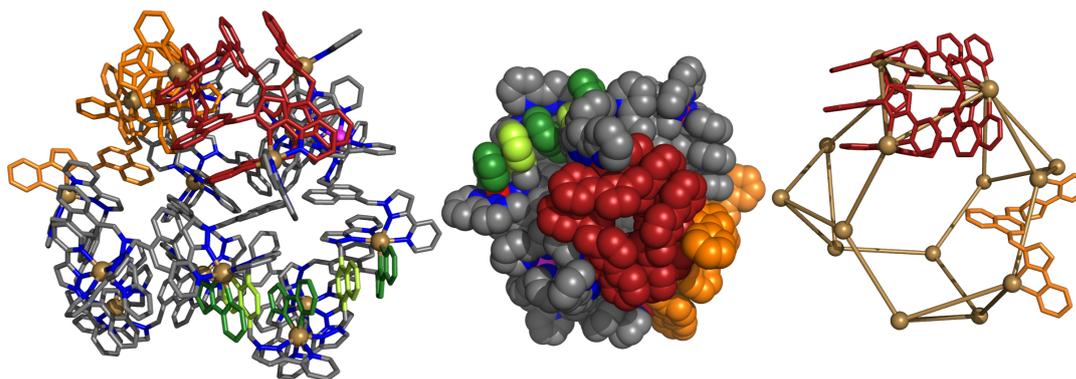
Fig. 1.54 Crystal structure[79] of $[Cd_{16}(L^{1,4naph})_{24}](BF_4)_{32}$ with highlighted M_3L_3 array (red) and ML_3 "cap" (orange) and π stack (green)

1.6.4   Host-Guest Binding in the Cubic Cage

A large majority of recent work, within the Ward group, as well as within this thesis, is based on the cubic cage formed with the ligand $\mathbf{L^A}$ as $[Co_8(L^A)_{12}](BF_4)_{16}$ and referred to hereafter as $\mathbf{H^A}$. Out of the structures within the Ward group family of cages, this cube is not only stable in solution, but possess a cavity where a guest can be readily exchanged and bound. Within each face, as can be seen in the space filling model (Fig. 1.58), there is a portal (4 Å wide) which leads to the internal cavity which has a calculated volume of 407 Å[80].
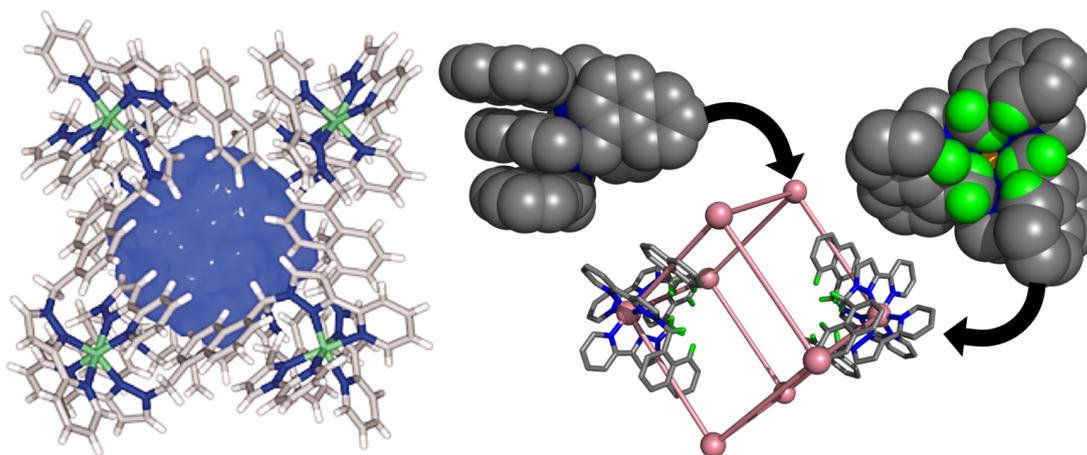
Fig. 1.55   (left)[80] volume of the cage cavity in blue, (right) frame of the cubic cage with the two *fac* and six *mer* vertices highlighted. The spacefilling models of the vertices is orientated from within the cavity

For host-guest titrations, $\mathbf{H^A}$ is normally used due to the inherent paramagnetic properties gained from Co^{II}. The ^{1}H-NMR spectrum of the cage shows 44 individual

peaks arising from two independent ligand environments nicely spread over a chemical shift range of 100 and -100 ppm. As a guest binds into the cage cavity, the shift in the peaks is enhanced by the paramagnetic environment, resulting in larger changes that are easier to observe.
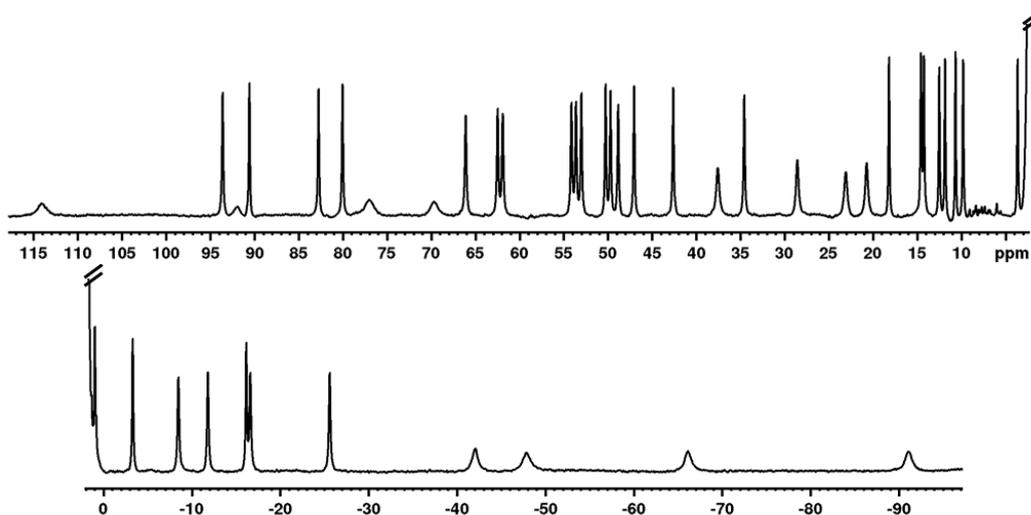
Fig. 1.56 ^{1}H-NMR of cage complex $\mathbf{H^A}$

The cage, soluble in acetonitrile, was investigated for host-guest binding using a series of small aromatic guests (Fig. 1.57). In organic solvents guest binding is driven by weak hydrogen bonding to the convergent set of CH protons present in the *fac* site (Fig. 1.55). The magnitude of guest binding was found to relate to the β (hydrogen bond accepting) parameters for the guests.
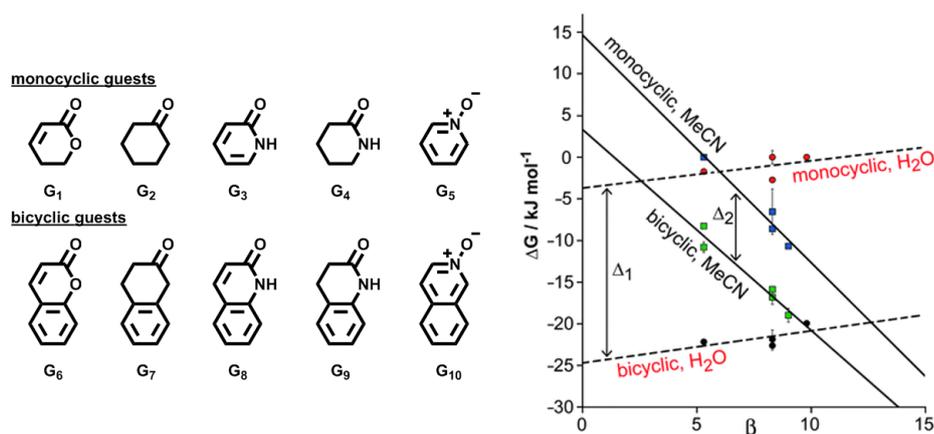
Fig. 1.57 Binding of cyclic guest with varying β values. (right portion) Reproduced (in-part) from ref[81] with permission of The Royal Society of Chemistry

In water, the hydrophobic effect dominates guest binding so a modified cage, soluble in water, was developed. This isostructural cage, $\mathbf{H^W}$, was solubilised into water through

the addition of twenty-four CH_2OH groups on to the exterior surface. The same series of guests were investigated and it was found that the changing β value had little impact on the binding, whereas the hydrophobic surface area of the guest contributed significantly.
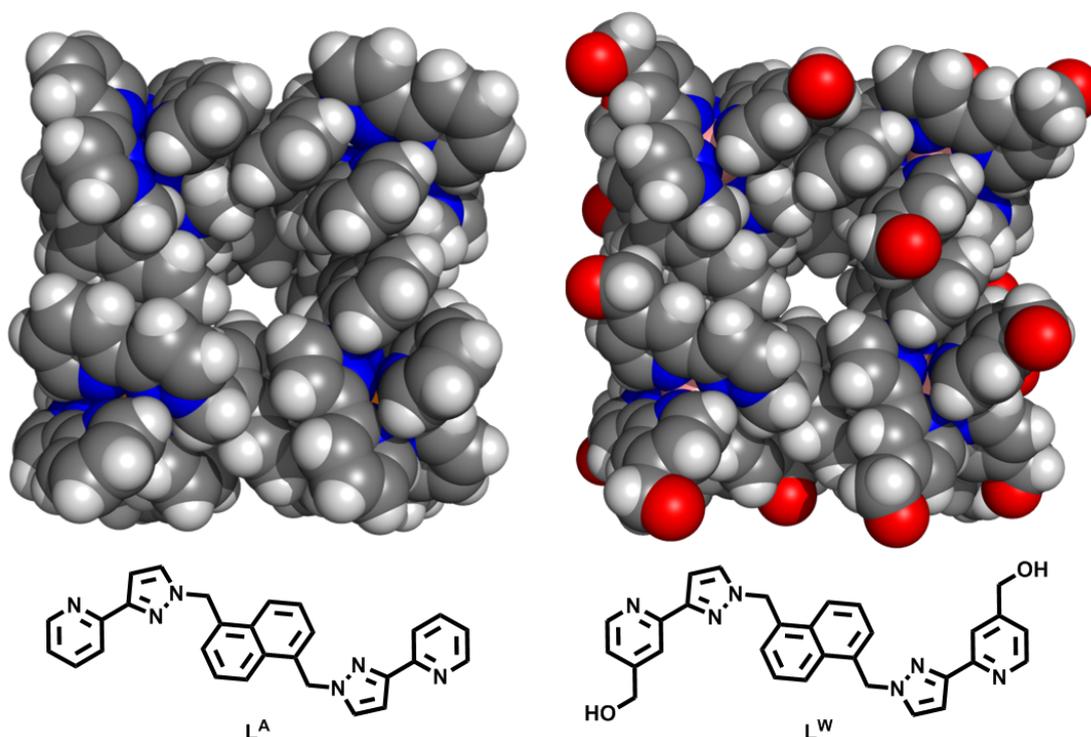
Fig. 1.58   Spacefilling model of the face of (left) $\mathbf{H^A}$ and (right) $\mathbf{H^W}$

In water a series of cyclic ketones showed that increasing the ring size of the guest resulted in a linear increase in binding strength. This increase is consistent with a ca. $4.7\,\mathrm{kJ\,mol^{-1}}$ with each additional CH_2 group[82]. This pattern held until the guest was too large to fit within the cage cavity. Cycloundecanone was found to be the strongest binding and also fill $50\,\%$ of the cavity, close to the optimal guest volume size of 55 ± 9 % as noted by Rebek[83–85].

Further work and examples are given within the introductory section of each relevant chapter.
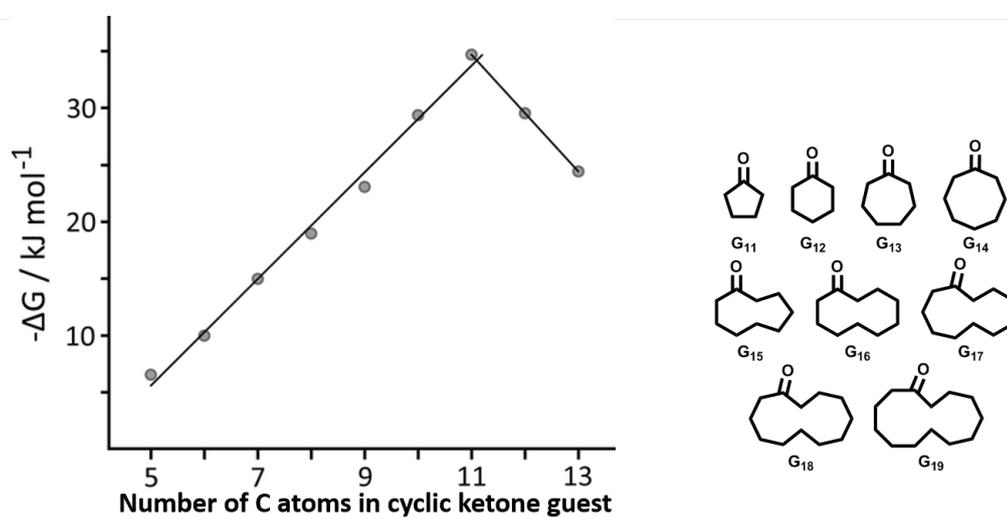
Fig. 1.59 Binding of a series of different sized cyclic ketones. (left portion) Reproduced (adapted) from ref[82] under Creative Commons CC-BY agreement

45

Chapter 2

Synthesis
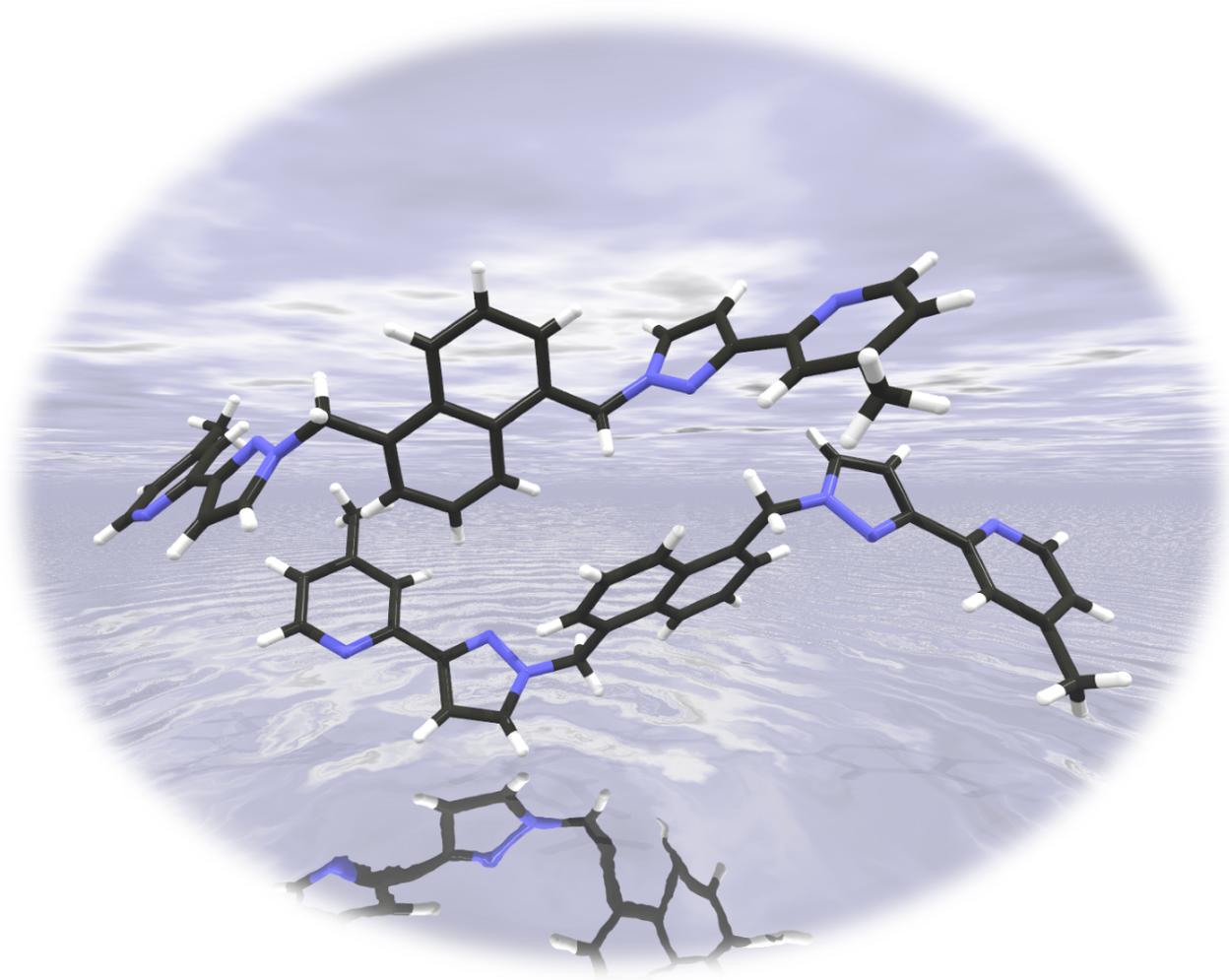
Rendering of the crystal structure for ligand $\mathbf{L^{Me}}$

2.1 Introduction

As discussed in Chapter 1.6 the cubic cage continues to be the most interesting of the Ward group supramolecular complexes. This cage is constructed by the self-assembly of twelve bis-bidentate bridging ligands with eight octahedral metal ions. As with all of these cage complexes the ligand on which it is based has two terminal pyridine-pyrazole units that provides the bidentate binding site for the chosen octahedral metal ions. The two units are connected together by a spacer group: naphthalene-1,5-diyl with saturated methylene linkages. The linkages help with the formation of the complex as well as helping with the synthesis for this ligand $\mathbf{L^A}$.
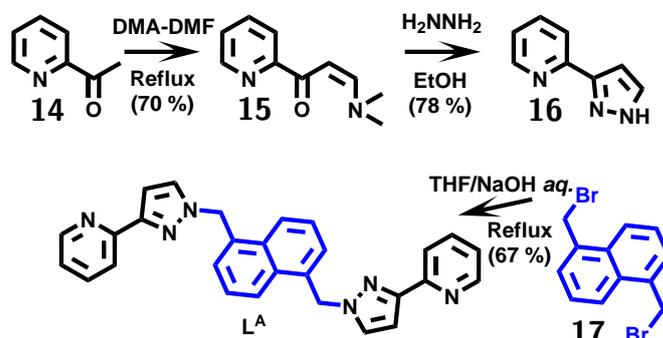
Scheme 2.1. Synthesis of ligand $\mathbf{L^A}$

The synthesis of the pyridine-pyrazole unit **16** starts from the commercially available pyridin-2-yl-ethan-1-one **14**. Through two efficient steps, via an keto-enamine intermediate **15**, the desired 3-(pyridin-2-yl)-pyrazole **16** is easily obtained. Two of these units are then coupled to the bis-brominated spacer group **17** to give the ligand $\mathbf{L^A}$.

The resulting cubic M_8L_{12} cage, from combining $Co(BF_4)_2$ and ligand $\mathbf{L^A}$, gives $\mathbf{H^A}$ which is soluble in acetonitrile. Initial investigations of host-guest chemistry of $\mathbf{H^A}$ were carried out in this solvent (Chapter 1.6.4). However to access more useful applications the cage needed to be modified to be soluble in water while maintaining the same structure and cavity properties. This was achieved through the addition of 24 hydroxymethyl

groups (CH$_2$OH) to the external surface of the cage[81]. A synthetic route to the modified ligand **L**$^{\mathbf{W}}$ was established to allow the creation of this water soluble cage starting from the commercially available hydroxymethyl-functionalised pyridine **18**.
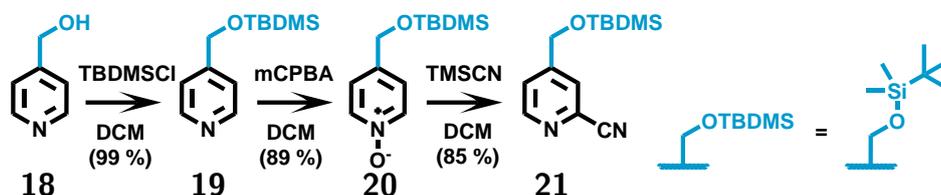
Scheme 2.2. Protection and conversion of starting pyridine **18**

A nitrile group was added in the 2-position, via the N-oxide **20**, before being modified to give the acetyl-pyridine **22**. From compound **22** the remaining two steps to **24** follow the same route as for ligand **L**$^{\mathbf{A}}$. The pyridine-pyrazole **24** is then coupled to the spacer group **17** before de-protection to give the CH$_2$OH functionalised ligand **L**$^{\mathbf{W}}$.
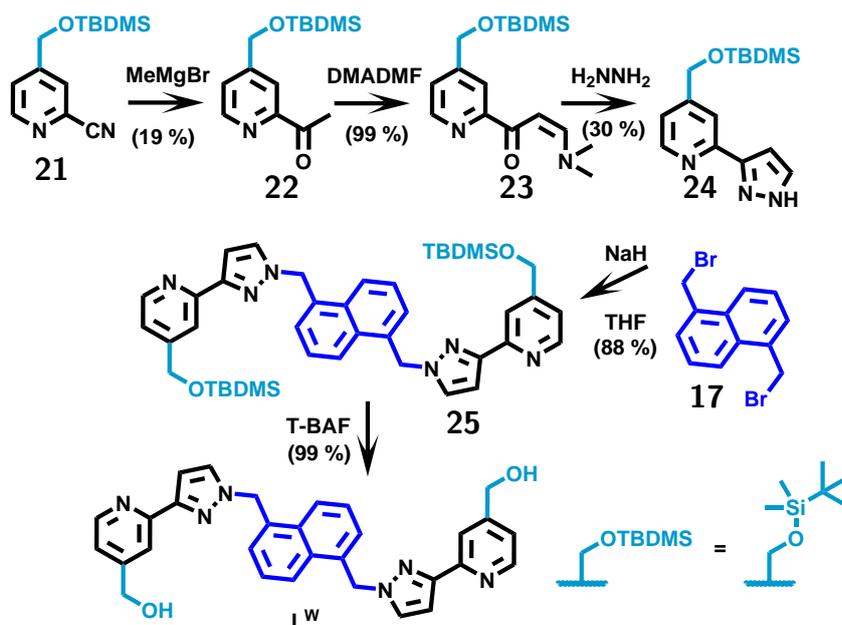
Scheme 2.3. Route through to ligand **L**$^{\mathbf{W}}$

Both of these ligands and their synthetic routes have been well established within the Ward group and appear in the literature[75,81].

2.2 Aim

The structure of the cage and some the principles for guest binding have been discussed within Chapter 1.4. Understanding of the individual contributions to guest binding,

particularly the hydrophobic effect, is key to the design of functional supramolecular systems.

Often useful information can be extracted from a supramolecular system via a thermodynamic cycle: a comparison between two solvent environments is required. To date these thermodynamic cycles have been used data on guest binding in water and acetonitrile[81]. In water the binding into the cavity is driven largely through the hydrophobic effect (Chapter 1.4.1), whereas in acetonitrile the binding occurs predominantly though a hydrogen bonding interaction between the guest and cage surface. Acetonitrile, though less polar than water, still favourably binds to polar guests. To study the hydrogen bonding strength of our guests within the cage in a less competitive environment we needed to be able to bind the guests in a non-competitive solvent. Therefore the initial aim of the work in this chapter was to develop a new isostructural cage which is soluble in a non-polar solvent such as dichloromethane or chloroform.
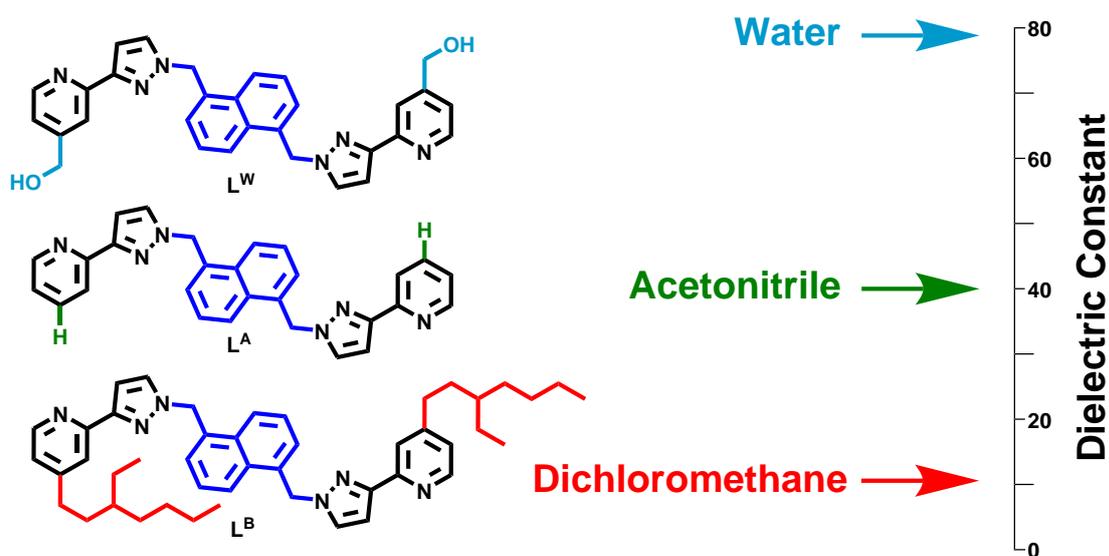
Fig. 2.1 Target ligands related to ideal solvent system

In a similar way to the synthesis of the water-solubilised cage, the external surface can be functionalised with alkyl chains to encourage solubility, this time in non-polar solvents. In a similar way that the hydrophobic contribution to guest binding can be determined, the host/guest hydrogen bonding interaction within the cage cavity should be able to be probed effectively through thermodynamic cycles involving comparison of guest binding in water and, for example, dichloromethane.

2.3 Discussion

2.3.1 Spacer Synthesis

The ligand used for the cubic cage contains a 1,5-naphthalene spacer group with saturated methylene linkages and the synthesis of the spacer starts with 1,5-dimethylnaphthalene **27**. Previously **27** was commercially available which allowed the brominated analogue **17** to be obtained through a simple bromination reaction. However the starting material **27** has since been discontinued by the commercial source: Sigma-Aldrich®. Other commercial sources are still available but at a much greater expense so instead alternative routes to **27** were investigated. The literature suggested that the best route is to prepare **27** from the dimethyl tetralin **26** which is readily available and still cheap.
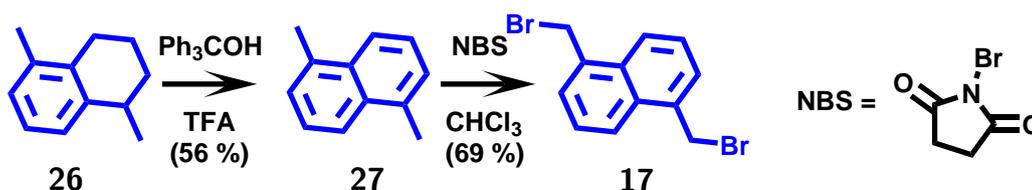
Fig. 2.2 Conversion of tetralin **26** to 1,5 dimethyl naphthalene

A successful route, which uses TFA with tri-phenylmethanol, was able to oxidise the tetralin **26** to **27**. The purified product was obtained with a yield of 56 %. Purification was achieved through silica chromatography.

The bromination of 1,5-dimethylnaphthalene to **27** was achieved by reaction with N-bromo-succinimide (NBS) through a photo-induced radical reaction. Originally the reaction was carried out in carbon tetrachloride solvent and purified by crystallisation to give the product **17**. However it has been found recently, within the Ward group, that it is possible to complete the bromination in other more readily available halogenated solvents including dichloromethane and chloroform. The use of chromatographic purification allowed for a greater yield to be obtained in a reduced time compared with recrystallisation. The general procedure now uses chloroform, which has a higher boiling point than dichloromethane, to give the bis-bromomethyl-naphthalene **17** in 69 % yield.

2.3.2 Post-Modification of Ligand $\mathbf{L^W}$

With a re-established route to obtain the spacer group the first attempt to create a non-polar solvent soluble cage was to modify the existing ligand $\mathbf{L^W}$. Using some simple chemistry allows addition of aliphatic chains through an ether linkage. The reaction of $\mathbf{L^W}$ with 1-bromo-2-methyl-hexane adds a branched chain to either end of the ligand to give the modified ligand $\mathbf{L^W_{mod}}$. Sodium hydride, 60 % dispersion, was used to deprotonate the alcohol and aid the addition of the alkyl group. Even under these conditions the reaction occurs slowly over two days. It was possible through thin layer chromatography (TLC) and mass spectrometry (MS) to observe the mono-substituted intermediate **28** being formed before the di-substituted product $\mathbf{L^W_{mod}}$ was made.
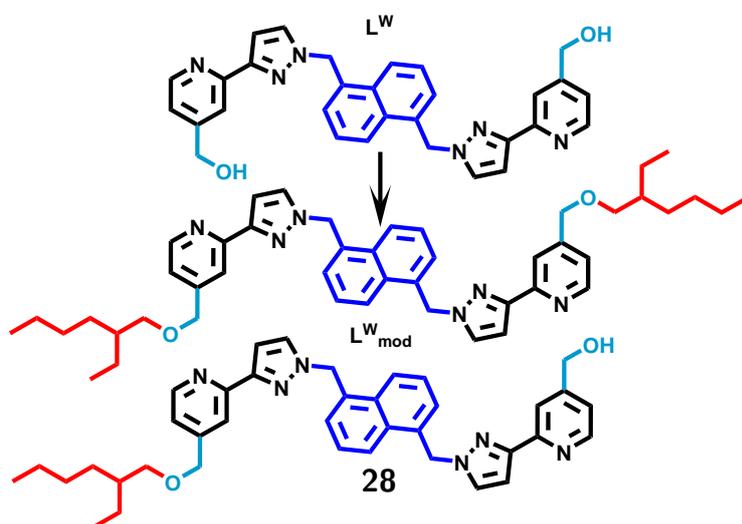
Scheme 2.4. Modification of ligand $\mathbf{L^W}$

The modified ligand $\mathbf{L^W_{mod}}$ was combined in a 3:2 ratio with $[\mathrm{Co(BF_4)_2}]$ in acetonitrile and refluxed overnight. After removal of solvent the remaining solid was collected and washed with water, methanol, dichloromethane, and then diethyl-ether sequentially. The $^1\mathrm{H-NMR}$ spectrum obtained showed several broad peaks spread-over a 200 ppm range as usual due to the paramagnetism. The peaks within the paramagnetic region that sharpen at higher temperature suggest that a metal-ligand complex was formed but it was not clear if this was cage complex or not. The spread of peaks across the large range (Fig. 2.3) was a result of the interaction between the ligands with the $\mathrm{Co^{II}}$ ions. Generally the cubic cage structure will give a characteristic pattern of mass peaks in an electrospray MS associated with sequential loss of anions from an intact $\mathrm{M_8L_{12}X_{16}}$ complex. However

these distinctive peaks could not be found therefore it was inconclusive to whether the self-assembly between Co^{II} ions and $\mathbf{L_{mod}^{W}}$ reaction had occurred.
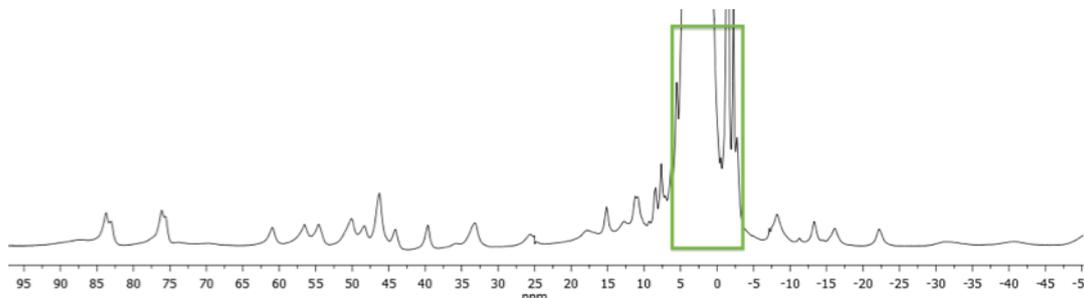
Fig. 2.3 ^{1}H-NMR spectrum of cage complex $[Co_8L_{12}^{mod}](BF_4)_{16}$ with solvent region highlighted in green

2.3.3 Synthesis of Ethyl-Functionalised Ligand $\mathbf{L^{Et}}$

The synthetic route used for the ligand $\mathbf{L^W}$ was applied to allow preparation of an ethyl functionalised-ligand starting from the commercially available ethyl pyridine **29**.
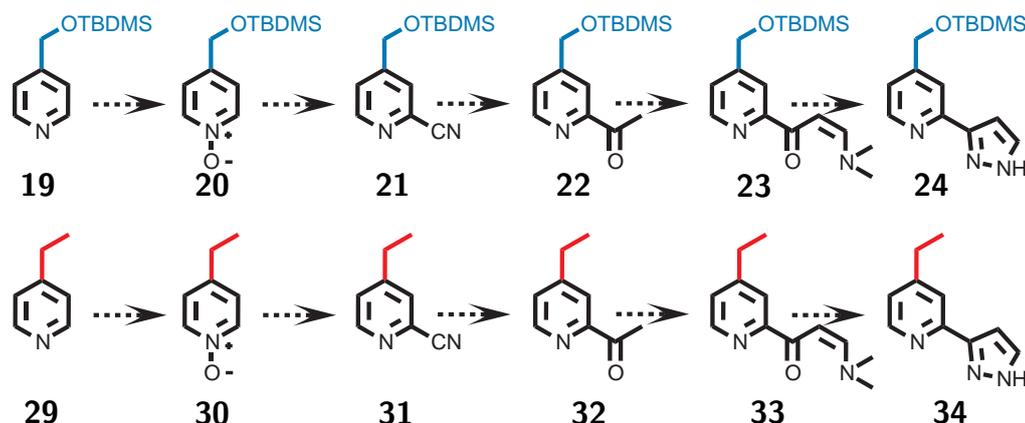
Scheme 2.5. Synthetic Route: (top) water soluble cage ligand synthesis route; (bottom) equivalent route for the ethyl pyridine pyrazole

The first step started with the commercially available 4-ethyl pyridine **29**. Initially the same conditions as previously used were able to afford the N-oxide **30**. The ^{1}H-NMR spectrum showed that the product was impure and had a low yield. Monitoring by TLC revealed that the reaction was going to completion in only 2.5 h. Shortening of the reaction time resulted in a three-fold increase in yield (23 % to 79 %).

The ^{1}H-NMR spectrum of N-oxide **30** has four signals, three of which are identical to those seen for the starting pyridine **29**. The fourth peak for pyridine **29** was at 8.43 ppm , shifted from the equivalent signal for the N-oxide **30** at 8.03 ppm. This is expected due

to the addition of the electron-withdrawing atom to the nitrogen resulting in a change to the electronic environment of the ortho-positioned protons.
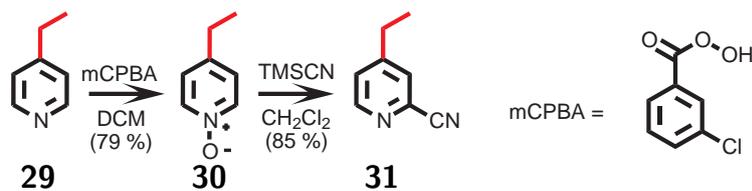
Scheme 2.6. Synthesis of (4-ethylpyridin-2-yl)2-nitrile **31**

The nitrile group is then added to the activated ortho position of the N-oxide **30** to give 2-cyano-4-ethyl pyridine **31** with an 85 % yield. The ^{1}H-NMR spectrum of **31** has similar peaks for the ethyl group as seen previously. The addition of the nitrile group breaks the symmetry of the pyridine so the ^{1}H-NMR contains individual peaks for the three remaining aromatic protons in **31**.

Conversion of **31** to 2-acetyl-4-ethyl-pyridine **32** using the Grignard reagent MeMgBr was achieved with a 59 % yield. The yield is greatly improved compared to the same step in the ligand $\mathbf{L^W}$ synthesis of 19 %. The ^{1}H-NMR spectrum has a new signal for the acetyl-methyl group at 2.73 ppm. The electrospray MS and ^{13}C-NMR spectrum also confirm that the product **32** has been formed.
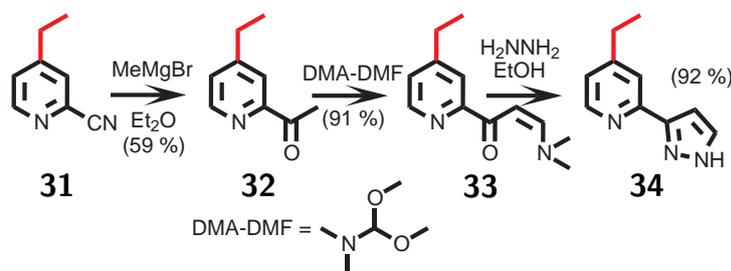
Scheme 2.7. Synthesis of ethyl functionalised pyridine-pyrazole

The next two steps, as with the synthesis of ligand $\mathbf{L^W}$, involves the reaction of the pyridine **32** with DMF-DMA, to give the intermediate **33**, followed by the cyclization using hydrazine monohydrate to give the pyridine pyrazole **34**. The two steps proceed with high yields and gave a high purity product **34**. The final step to synthesise ligand $\mathbf{L^{Et}}$ was the addition of the spacer, and was achieved by the combination of deprotonated **34** and reagent **17**, in a 2:1 ratio.

The ligand $\mathbf{L^{Et}}$ was combined in a 3:2 ratio with $Co(BF_4)_2$ in acetonitrile in a solvothermal reaction. The mixture was heated to 120 °C in a sealed autoclave reaction
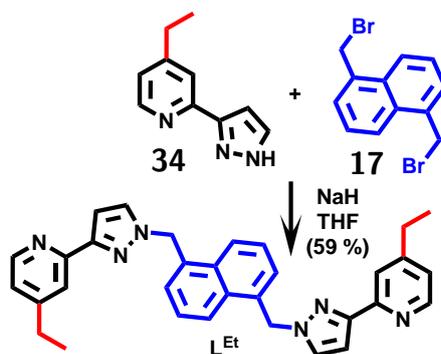
Scheme 2.8. Coupling to spacer group to give completed ligand $\mathbf{L^{Et}}$

vessel. A crude paramagnetic $^1\text{H}-\text{NMR}$ spectrum of the pale salmon-coloured product gave a distribution of 50 discernible proton peaks in acetonitrile. For a cubic cage complex there are 2 ligand environments and so for complex $[\text{Co}_8\text{L}^{Et}_{12}](\text{BF}_4)_{16}$ $\mathbf{H^{Et}}$ 60 individual peaks are expected in the ^{1}H-NMR spectra. The presences of 50 discernible ^{1}H-NMR signals for the complex means that there are greater than 1.5 ligand environments (45 signals) and likely that a few of the 60 expected signals for a cubic M_8L_{12} complex are overlapping. The product was not found to dissolve in other common solvents. The electrospray MS confirmed the formation an $[\text{M}_8\text{L}_{12}]^{16+}$ complex with a distinctive pattern of peaks related to the sequential loss of anions, for example $[M_8L_{12}X_{16-x}]^{x+}$, $x = 3$ to 8.

As the $\mathbf{H^{Et}}$ complex was only found to be soluble in acetonitrile an alternative anion was also used with the aim of making $[\text{Co}_8(\text{L}^{Et})_{12}](\text{BPh}_4)_{16}$, to improve solubility in low-polarity solvents. Though a complex was formed and was soluble in dichloromethane the ^{1}H-NMR spectrum only showed twenty-six peaks. One ligand environment should give a ^{1}H-NMR spectrum with 30 signals. The three protons in the terminal CH_3 groups are likely equivalent and we therefore would expect give 26 signals for a single ligand environment. The ^{1}H-NMR spectrum therefore suggested that the formed complex was not the M_8L_{12} complex which has two independent ligand environments giving 60 ^{1}H-NMR signals. The MS for this new complex with the BPh_4^- anion did not give a spectrum with the distinctive series of sequential peaks as seen for the complex with the BF_4^- anions.

Of the two complexes attempted with $\mathbf{L^{Et}}$, one formed a cubic structure and one did not. The cubic M_8L_{12} complex formed, $\mathbf{H^{Et}}$, was only soluble in acetonitrile and
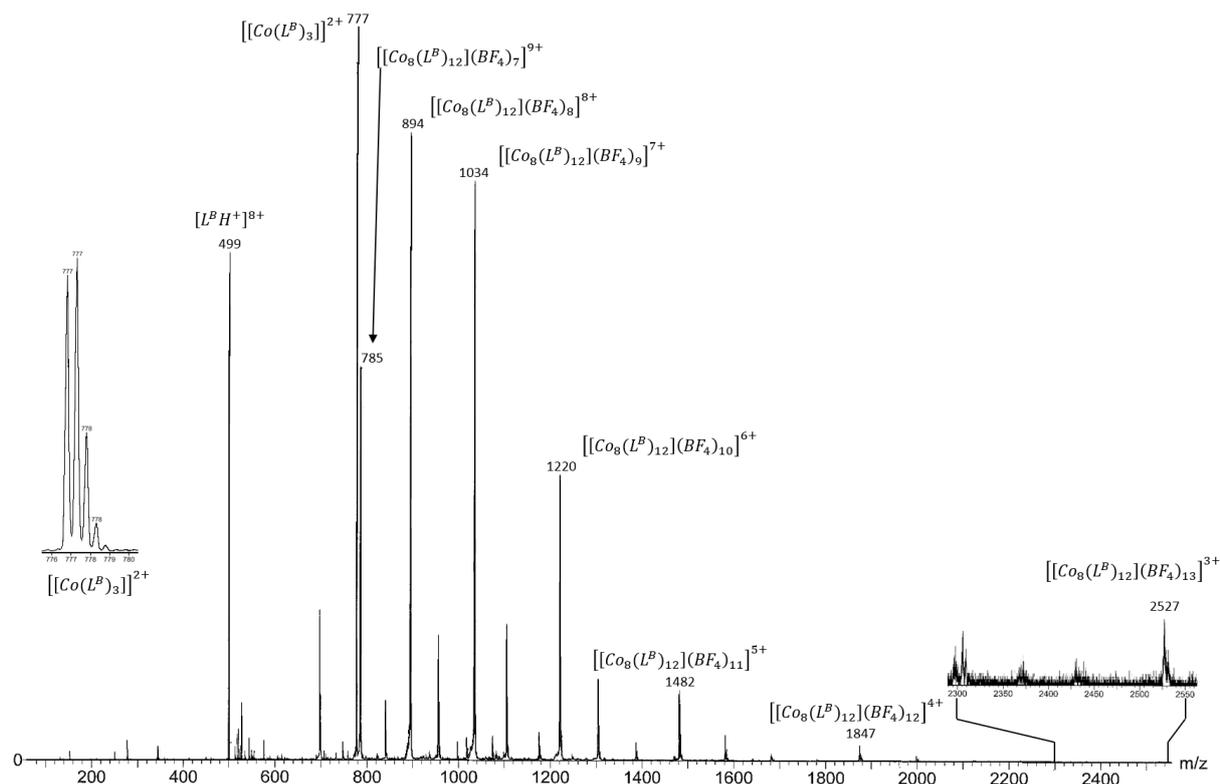
Fig. 2.4 Mass spectra of cage complex $[Co_8(L^{Et})_{12}](BF_4)_{16}$ **H^{Et}**

therefore, while involving a more challenging synthesis, offered no benefit in comparison to **H^A**. Therefore, the cage using **L^{Et}** was not taken any further and alternative methods to develop the desired low-polarity solvent soluble cage complex were looked into.

2.3.4 Synthesis of New Alkylated Pyridines

Alternative ligands to produce low-polarity solvent soluble cage complexes could have a range of different functional groups on the 4-position of the pyridine ring and these can simply be accessed using the synthetic route described above by starting with different alkyl-substituted pyridines. A great improvement in the overall yield of ligand synthesis was seen, with the removal of the OTBDMS group, in the ligand **L^{Et}** synthesis (Scheme 2.5). However the ethyl groups do not enable the cage to be solubilised sufficiently in non-polar solvents. Functionalisation in the 4-position of the pyridine stills allows cage formation so addition of longer aliphatic chains could allow the cage to dissolve more effectively in non-polar solvents.

It is possible to functionalise 4-methyl pyridine through addition to the methyl group by using simple lithium chemistry. Lithium diisopropylamide (LDA) can deprotonate

the acidic proton on the methyl group, which then allows the addition of an appropriate bromo-alkyl substrate to add the desired alkyl chain.
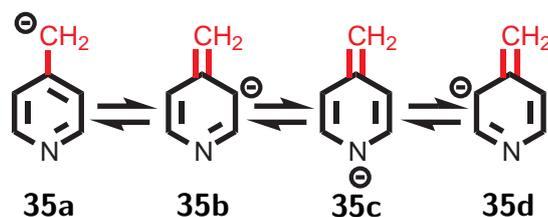
Scheme 2.9. Resonance structures of deprotonated 4-methyl-pyridine

The lithiated intermediate is stabilised through conjugation in the pyridine ring with a range of different possible resonance structures (**35a**, **35b**, **35c** and **35d**) as shown in Scheme 2.9. There are many examples of these pyridine lithiation reactions found in literature including the most relevant literature reaction where lithiated methylpyridine is reacted with 1-bromo hexane to give 4-heptyl pyridine **37** (Scheme 2.10)[86].

The same chemistry was applied to the synthesis of the C_{13}-substituted pyridine **38** by reacting 4-methylpyridine (deprotonated) with a bromo-dodecane chain and was achieved with 85 % yield. The ^{1}H-NMR data confirmed that the correct product was isolated. This reaction allows a variety of different alkylated pyridines to be easily obtained.
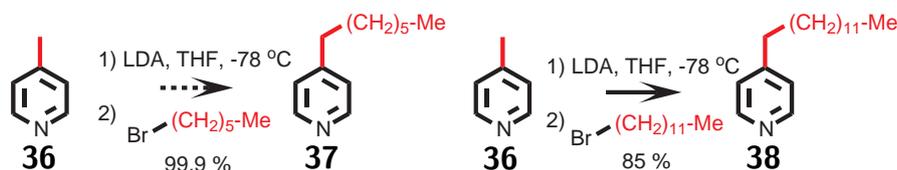
Scheme 2.10. Functionalisation of pyridine, (left) from literature[86], (right) test reaction

Initially two different alkyl substituents were decided upon to hopefully produce the desired functionalised ligand. The first of these side groups, a branched chain alkyl group, should allow for solubilisation of a formed cage into a non-polar solvents. A linear chain could cause unwanted aggregation whereas a branched chain helps reduce aggregations between the alkyl chains and improve solubility in non-polar solvents. The second side group decided upon was a cyclo-hexyl group attached through a methylene linker. This latter group should also provide similar solubilisation. Furthermore with the increased pre-organisation of the hexyl ring, as opposed to a floppy chain, this may allow for crystallisation.
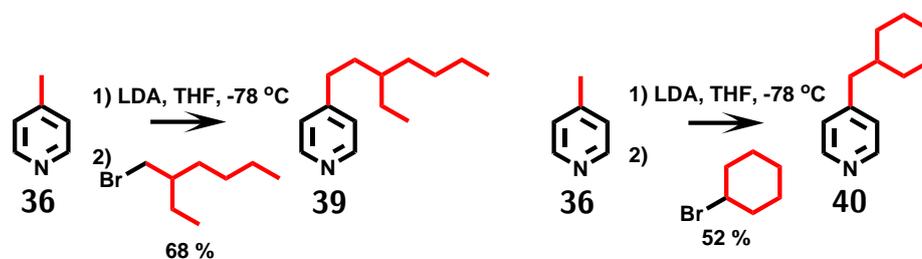
Scheme 2.11. Synthesis of two alkyl functionalised pyridines

These side groups were added through the same lithiation reaction starting from **36** to give the functionalised pyridine starting materials, **39** and **40**. Both of these pyridines have distinctive ^{1}H-NMR spectra. There are two signals in the aromatic region, corresponding to the pyridine protons, matching those seen previously for the synthesis of ligand $\mathbf{L^W}$ and $\mathbf{L^{Et}}$. Additionally a 'fingerprint' pattern that is indicative of the alkyl side chain is present between 0 and 3 ppm.
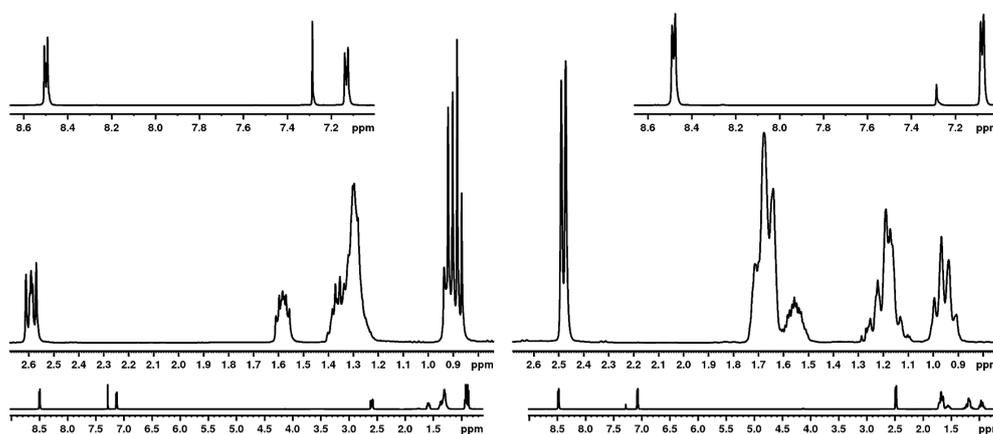
Fig. 2.5 'Fingerprint region of the alkyl pyridines: (left) branched chain pyridine **39**. (right) cyclo-hexyl pyridine **40**

For both pyridines **39** and **40**, the N-oxide was obtained using the same conditions as seen previously for ligand $\mathbf{L^W}$ and $\mathbf{L^{Et}}$. These products were easily isolated in high yield of 94 % and 90 % for the N-oxide **41** and **43** respectively. The addition of a nitrile group also proceeded with high yield to give **42** and **44**. The ^{1}H-NMR spectra show the same distinctive breaking of symmetry in the pyridine as seen with ligand $\mathbf{L^W}$ and $\mathbf{L^{Et}}$, with the aromatic region having three signals.

The next step in the synthesis was the Grignard reaction of the nitrile group with MeMgBr to give products **45** and **46**. Unlike the reaction of **31** to give **32** (Scheme 2.7), the reaction yields were poor with yields of 15 % and 13 % for **45** and **46** respectively. The reactions showed a high level of precipitate under the same conditions that had
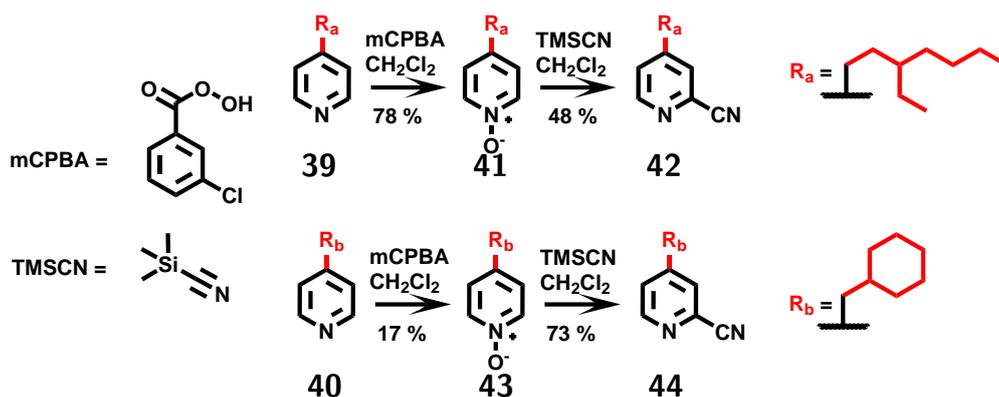
Scheme 2.12. Conversion of functionalised pyridines

previously worked for $\mathbf{L}^{\mathbf{Et}}$. The reduced solubility in di-ethyl ether, used in the Grignard reaction, is likely a result of the alkyl substituents. An ether containing solvent is required to stabilise the magnesium reagent used in the Grignard reaction and so the THF was tested as an alternative solvent for the Grignard reaction with **42**. Initially when using THF, at the previously used temperature of $0\,°C$, the reaction occurred in under a minute producing a mixture of undesired products. The reaction was instead cooled to $-78\,°C$ and, while still going to completion in under $5\,min$, gave a slight improvement in yield of **45** (15%). The reaction at $0\,°C$ in diethyl ether typically takes around $3\,h$ to go to completion. Even after changing the solvent the yield for **45** was still worse than the previous routes to either **22** or **32**.
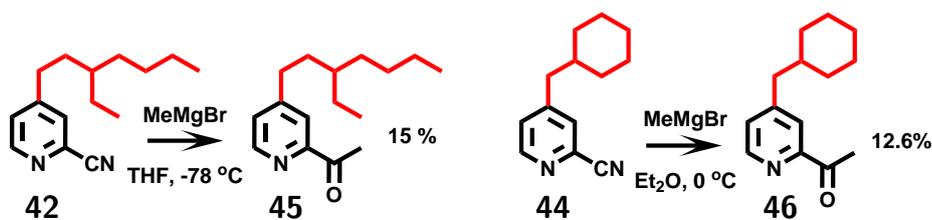
Scheme 2.13. Grignard reaction to give alkyl-functionalised acetyl-pyridines

2.3.5 Synthesis of Methyl-functionalised Ligand $\mathbf{L}^{\mathbf{Me}}$

The very low yields for the preparation of acetyl-pyridines **45** and **46** meant that the target ligands $\mathbf{L}^{\mathbf{B}}$ and $\mathbf{L}^{\mathbf{C}}$ could not be easily accessed through the standard synthetic routes. Instead a different route was attempted by first creating $\mathbf{L}^{\mathbf{Me}}$ with the aim to then use the lithiation chemistry to alkylate the pre-formed $\mathbf{L}^{\mathbf{Me}}$ and give the desired ligands (Scheme 2.14).
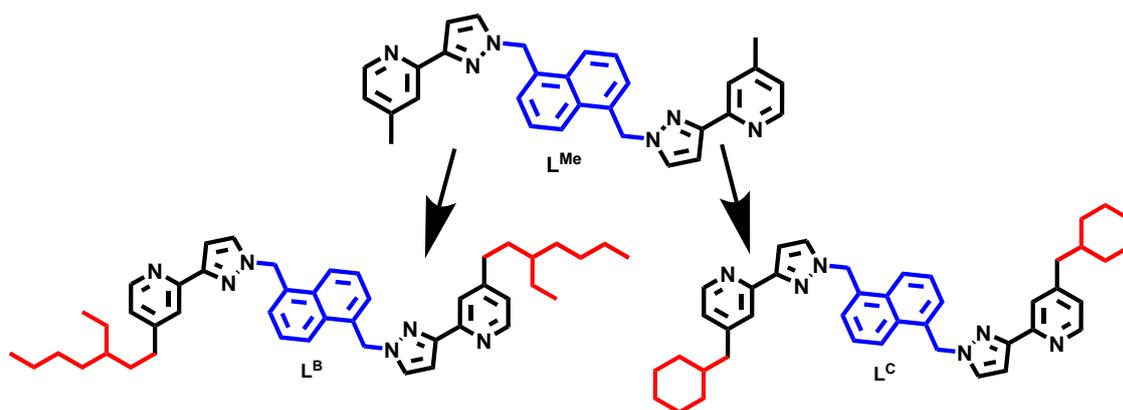
Scheme 2.14. Alternative route to obtain $\mathbf{L}^{\mathbf{B}}$ and $\mathbf{L}^{\mathbf{C}}$ through alkylation of $\mathbf{L}^{\mathbf{Me}}$

The first hurdle in creating a methyl substituted ligand was obtaining 4-methyl pyridine N-oxide **47**. The previously described conditions to give a pyridine N-oxide from a pyridine (for example **29** to **30**) were used for the synthesis of **47** but resulted in a surprisingly low yield. This was a result of the N-oxide **47** being highly soluble in the water layer of the reaction workup. Repeated extraction would continually give small extra amounts of product but, even with multiple separations, a very poor yield was gained. A method was found that did not require a separation during the workup. Boiling the pyridine **36** in a mixture of hydrogen-peroxide and glacial acetic acid produced the pyridine N-oxide by precipitation of the product, and the excess acetic acid was removed *in vacuo*. The addition of the nitrile to the ortho-position of pyridine ring then proceeded without issue and gave **48**.
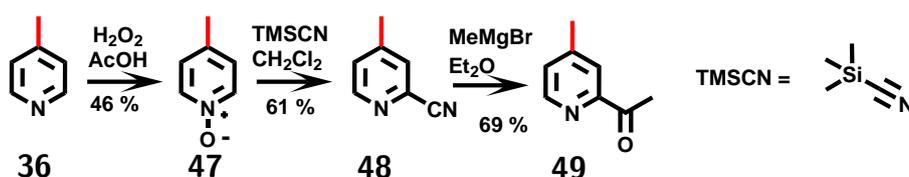
Scheme 2.15. Synthesis of (4-methylpyridin-2-yl)ethan-2-one **49**

The Grignard reaction to give **49**, performed using the previously described conditions (0 °C in diethyl ether), gave the acetyl product **49** in 69 % yield. The transformation of signals in the ^{1}H-NMR was clear with the addition of a new signal at 2.69 ppm for the methyl of the acetyl group.

4-methyl-2-acetyl pyridine **49** was then refluxed with DMF-DMA overnight to give intermediate **50**. Once separated the intermediate was used without purification and was reacted with hydrazine monohydrate, in a ring closing reaction, giving an 85 % yield of
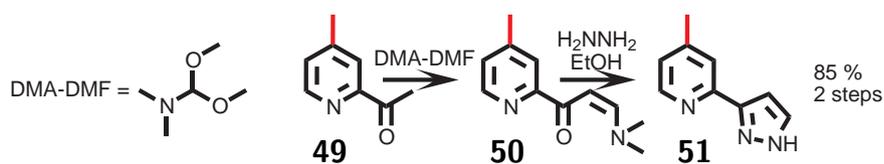
Scheme 2.16. Synthesis of 4-methyl(2-pyrazyl)-pyridine **51**

51 across the two steps.
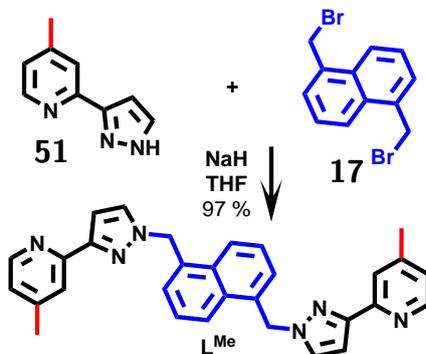
Scheme 2.17. Coupling to spacer group to give ligand **L**^{Me}

Pure **51** was isolated as a golden brown solid. A 2:1 mixture of **51** and **17** was then refluxed in a THF (175 mL) and 5.5 M NaOH (20 mL) biphasic solution. A crystal structure of **L**^{Me} was obtained verifying the expected structure.
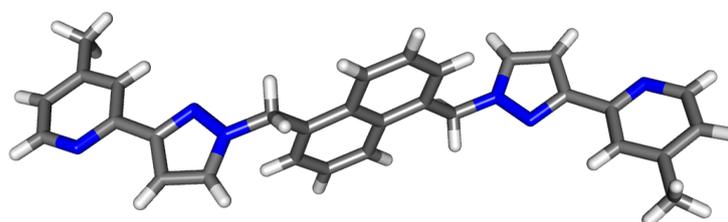
Fig. 2.6 Crystal structure of methyl functionalised ligand **L**^{Me}

After multiple attempts it was found that the lithiation chemistry, used successfully to alkylate **36**, did not work on the pre-formed ligand **L**^{Me} (Scheme 2.14). Using the more powerful deprotonating reagent (butyl lithium) in place of LDA also did not give the desired alkylated ligands **L**^B or **L**^C with the crude mass spectrum showing no evidence of any product.

2.3.6 Alkylated Ligands from Pre-alkylated Pyridines

With neither of the previous routes to the alkylated ligands performing as hoped, the literature was again searched to find a new procedure. The particular focus was on

replacing the low-yielding Grignard reaction as a route to the acetyl pyridines **45** and **46**. While looking at reactions related to the lithiation of 4-methylpyridine, which normally results in substitution of an acidic methyl proton, it was found that the deprotonation could instead be directed to the ortho-position of the pyridine ring. The directing agent, di-methylaminoethanol (DMAE), works with pyridine by coordinating to the pyridine's nitrogen and the lithiating species, bringing the lithium reagent close to the pyridine's proton in the 2-position.
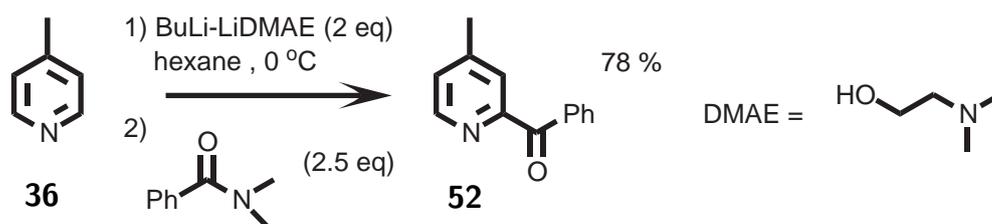
Scheme 2.18. Literature example of a directed butyl-lithium based reaction[87]

The conditions described in the literature[87] (Scheme 2.18) used two equivalents of butyl-lithium along with the directing agent DMAE. The first equivalent of butyl-lithium deprotonates DMAE and forms a complex with the second butyl-lithium equivalent. The formed complex, BuLi.LiDMAE, associates with the pyridine nitrogen atom and deprotonates in the 2-position of the pyridine to give the intermediate shown in Scheme 2.19. The literature makes use of N,N-dimethylbenzamide as an electrophile to add a benzo-ketone group in the 2-position of the pyridine core (Scheme 2.18). The use of DMA as an electrophile should, by analogy, allow the addition of the desired acetyl group in a single step to a pyridine nucleus at the carbon in the 2-position.

4-ethylpyridine **29** was used to test the reaction's viability as it was readily available and has less acidic protons than in the analogous methyl pyridine. After proceeding with the reaction, using the conditions as described in the literature[87], the crude mass spectrum showed peaks for the expected 4-ethyl-2-acetyl pyridine **32** while the crude ^{1}H-NMR spectrum showed a mixture of starting material and product signals. These product signals matched the ^{1}H-NMR spectrum previously obtained for pyridine **32**.

The new synthetic route to **32** was applied to the alkylated pyridines **39** and **40**. Both reactions yield a mixture of starting material, (**39** and **40**), and the corresponding 2-acetyl substituted products (**45** and **46**). The best previous combined yield to obtain
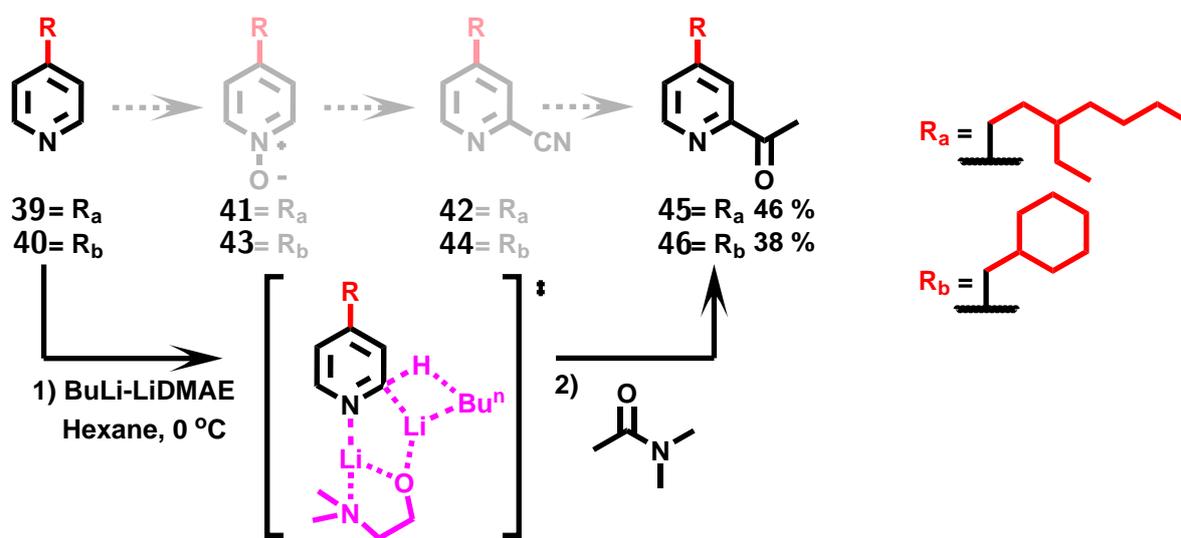
Scheme 2.19. Shortened synthesis to pyridine **46**

each **45** and **46** was 15 % and 13 % respectively and had two additional synthetic steps. The new reaction using BuLi.LiDMAE gave the 2-acetyl-pyridines in only one step with yields of 46 % and 38 % respectively.
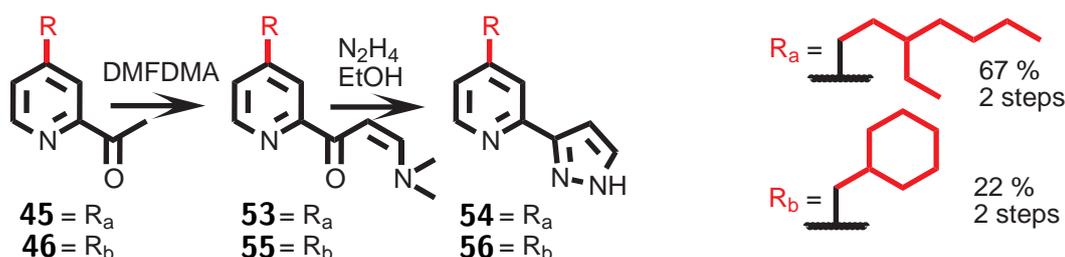
Scheme 2.20. Synthesis of functionalised pyridine-pyrazole

After some final refinements to the reaction conditions a reasonable quantity of each of **45** and **46** was obtained and the two remaining steps to create the pyridine pyrazoles **54** and **56** were carried out. The reaction conditions for these two steps did not need to be modified and both the targets **54** and **56** were finally obtained.

The final steps to create the ligands $\mathbf{L^B}$ and $\mathbf{L^C}$ involved coupling to the spacer group. The same conditions as used for ligand $\mathbf{L^A}$ (NaOH in THF) the used and yielded the completed ligands in a 54 % and 52 % for $\mathbf{L^B}$ and $\mathbf{L^C}$ respectively.

Ligand $\mathbf{L^B}$ was combined in a 3:2 ratio with $Co(BF_4)_2$ in methanol through a solvothermal reaction to give the standard M_8L_{12} cubic complex. The mass spectrum showed the typical series of signals for the completed cage due to the sequential loss of anions (Fig. 2.7). The ^{1}H-NMR spectrum gave a pattern of peaks similar to that for the previous cubic cages with two independent ligand environments.
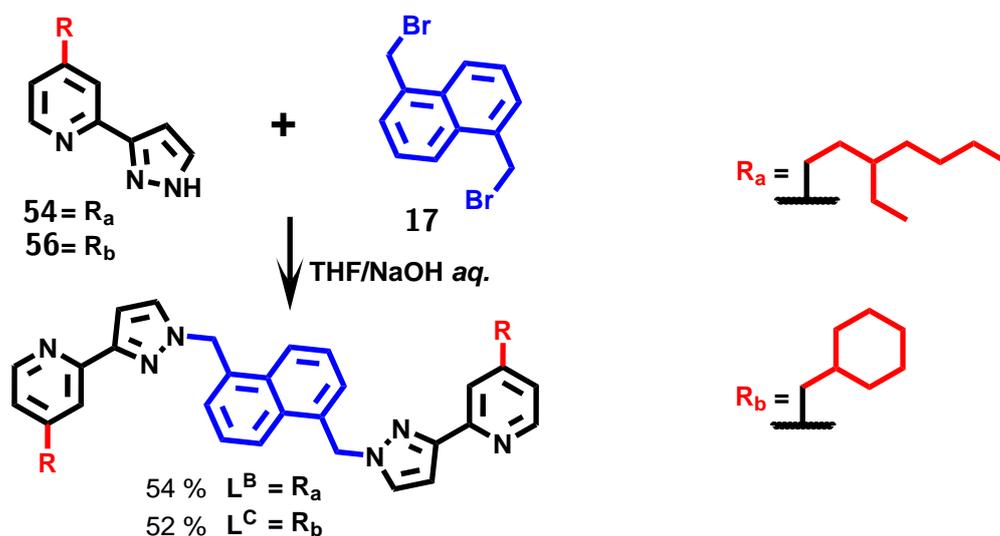
Scheme 2.21. Synthesis of alkylated ligands $\mathbf{L^B}$ and $\mathbf{L^C}$
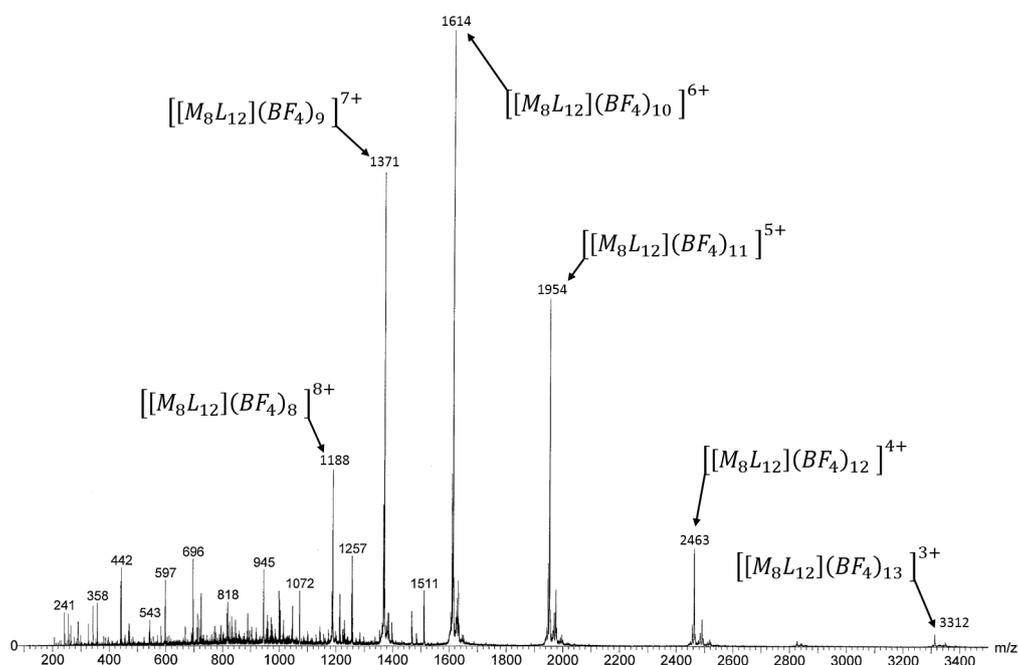
Fig. 2.7 $[Co_8(L^B)_{12}](BF_4)_{16}$ cage complexes mass spectrum

Ligand $\mathbf{L^B}$ was also combined in a 3:2 ratio with $Cd(NO_3)_2$ to yield another cubic complex. The ^{1}H-NMR spectrum (in chloroform) showed a range of peaks within the 0 to 10 ppm region. A DOSY-NMR spectrum of the formed complex was also measured. Disregarding the solvent peaks the complex showed the same diffusion coefficient across all of the signals for the complex.

Ligand $\mathbf{L^C}$ was combined in a 3:2 ratio with $Co(BF_4)_2$ in methanol through a solvothermal reaction to give a pink solid. The mass spectrum gave the expected series of peaks for sequential anion loss from an M_8L_{12} complex. However it was not
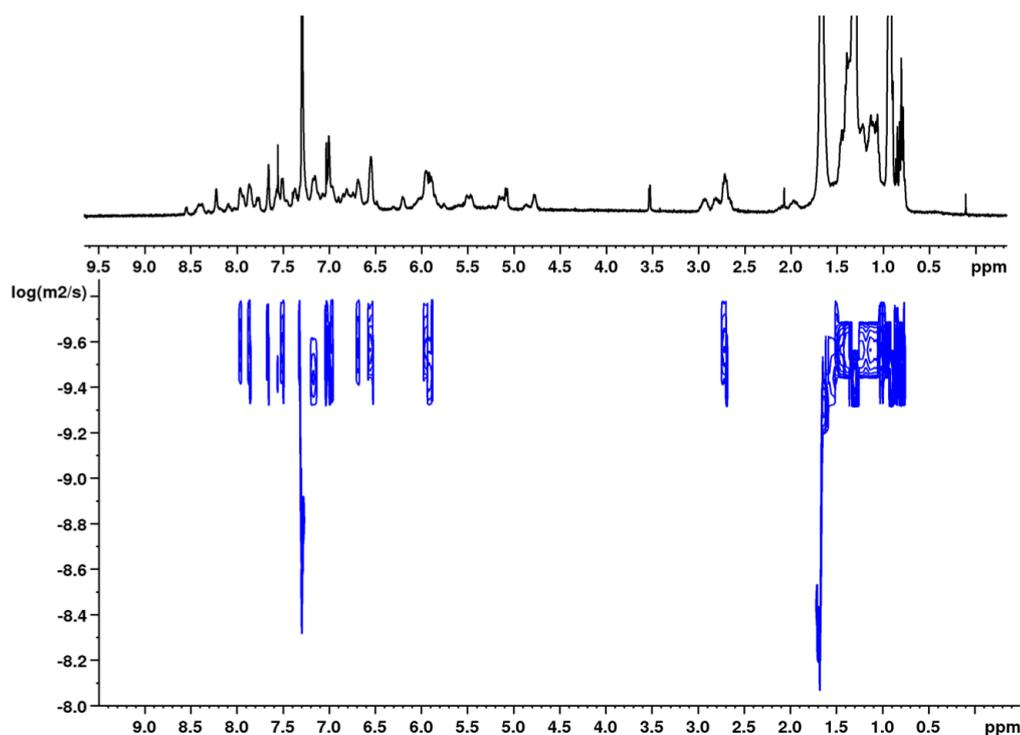
Fig. 2.8 DOSY-NMR spectrum of complex $[Cd_8(L^B)_{12}](NO_3)_{16}$

sufficiently soluble enough in acetonitrile, dichloromethane or chloroform to obtain an NMR spectrum.

Both ligands L^B and L^C formed were combined in a 3:2 ratio with $Co(BPh_4)_2$ and each formed a complex that was soluble in chloroform. However, the ^{1}H-NMR spectrum of each only showed signals equivalent to one ligand environment instead of the two ligand environments that are required for the M_8L_{12} complex. This difference is also apparent in that there were fewer paramagnetically shifted ^{1}H-NMR peaks and the signals did not correspond to those expected for the intact M_8L_{12} cubic cages. Neither of the MS spectra for either complex showed the expected series of signals for sequential loss of anions and therefore it is most likely the cubic cage complexes were not formed when BPh_4^- was the anion. This is the same result as seen when the ligand L^{Et} complex with $Co(BPh_4)_2$ was previously attempted (Chapter 2.3.3).

Both of the new ligands, L^B and L^C, can be complexed with $Co(BF_4)_2$ to give $M_8L_{12}X_{16}$ cubic cages. L^B also formed a cubic complex with $Cd(NO_3)_2$. Only the $[Co_8L_{12}^B](BF_4)_{16}$ complex, H^B, is sufficiently soluble in a desired solvent (in this case dichloromethane) to allow investigation of host-guest chemistry in a non-polar solvent (2.4). The use of BPh_4^- as an alternative anion proved to be unsuccessful. With H^B

providing the desired complex, the further development of alternative ligands or cage complexes for solubilisation into non-polar solvents was not continued.

2.3.7   Design and Synthesis of a Terpyridine Based Ligand

Many mixed metal complexes have previously been made[88,89] including some within the Ward group[90,91]. The design and synthesis of these complexes is often non-trivial. Easier access to a mixed metal cages containing large central cavities would provide new and exciting possibilities for host/guest chemistry.
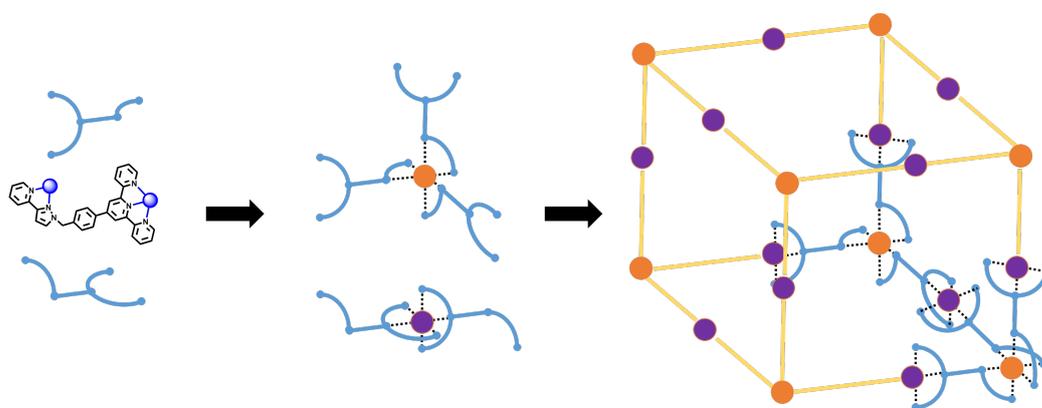
Fig. 2.9   Cartoon of a design idea to create a mixed-metal cage complex

Using a ligand that contains one bidentate binding site and one tridentate site should allow the two ligand termini to each select for a different metal ion. Octahedral coordination can be provided to metal ions by either two tridentate ligand termini or three bidentate termini. A ligand with both bidentate and tridentate termini can be made by connecting a terpyridine unit to a bidentate pyridine pyrazole unit through a phenyl spacer (Scheme 2.22).
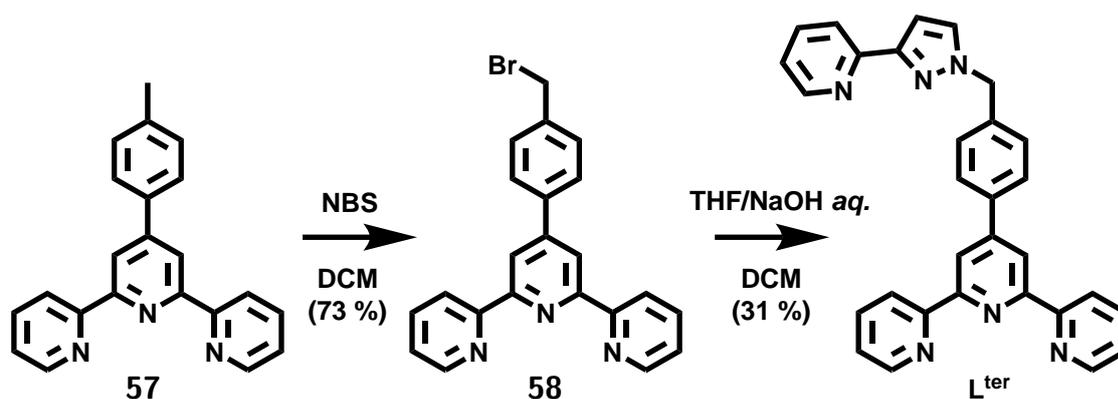
Scheme 2.22. Synthesis of new ter-pyridine based ligand

The simple bromination of the commercially available 4-tolyl-terpyridine **57** and coupling to pyridyl pyrazole yields the desired ligand $\mathbf{L^{ter}}$. The MS and ^{1}H-NMR spectrum show the signals expected for the ligand.
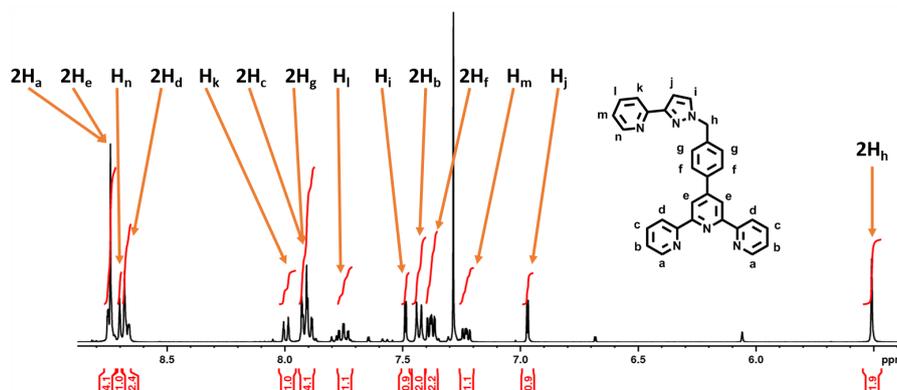
Fig. 2.10 ^{1}H-NMR spectra of $\mathbf{L^{ter}}$

Complexes of an unfunctionalised terpyridine (terpy), for example $[Fe(terpy)_2]^{2+}$, are known in the literature[92]. Therefore the formation of a complex between the new ligand $\mathbf{L^{ter}}$ and Fe^{II} were investigated. $FeCl_2$ was refluxed with ligand $\mathbf{L^{ter}}$ in a 1:2 ratio to assemble the larger ligand complex $[Fe(L^{ter})_2]Cl_2$. The ^{1}H-NMR spectrum signals for the formed complex $[Fe(L^{ter})_2]Cl_2$ were similar to those described for a $[Fe(terpy)_2]^{2+}$ complex in the literature[92] (See experimental for more details). The mass spectrum of $[Fe(L^{ter})_2]Cl_2$ also shows a signal at 494 (m/z) with an M^{2+} pattern typical of Fe^{II} complexes and which matches the calculated spectrum(Fig. 2.12).
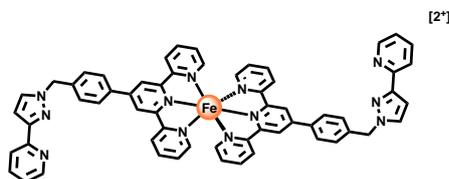
Fig. 2.11 Sketch of the assembled complex cation $[Fe(L^{ter})_2]^{2+}$

The $[Fe(L^{ter})_2](Cl)_2$ complex was combined with $Co(BF_4)_2$ in a 3:2 ratio in an attempt to form a larger complex such as the cube (sketched in Fig. 2.9). There was a mixture of different signals in the mass spectrum but it was not possible to isolate any single products. As it was not clear what had formed the work has not been progressed any further at this point.
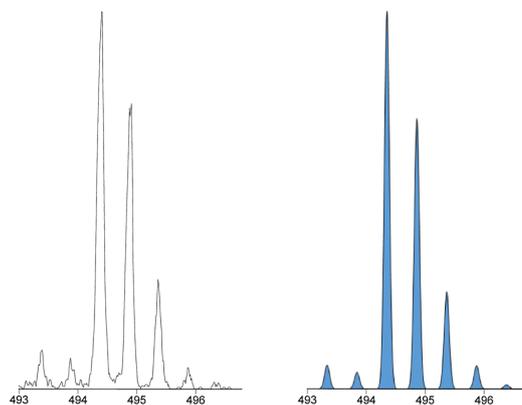
Fig. 2.12 Expansion of MS signal at 494 (m/z) for $[Fe(L^{ter})_2]^{2+}$ (left) measured, (right) calculated

2.4 Guest binding in the DCM soluble cage

The $[Co_8(L^B)_{12}](BF_4)_{16}$ complex $\mathbf{H^B}$, created through a solvothermal reaction described earlier, is soluble in dichloromethane and produces a similar pattern of signals in ^{1}H-NMR spectrum as seen for complexes $\mathbf{H^A}$ and $\mathbf{H^W}$ with two ligand environments. While the crystal structure of $\mathbf{H^B}$ has not been obtained, both the MS and ^{1}H-NMR spectrum match what is expected for the cubic cage structure. Therefore host-guest chemistry of the cage complex in dichloromethane can be investigated with confidence which was the initial aim of this part of the project.
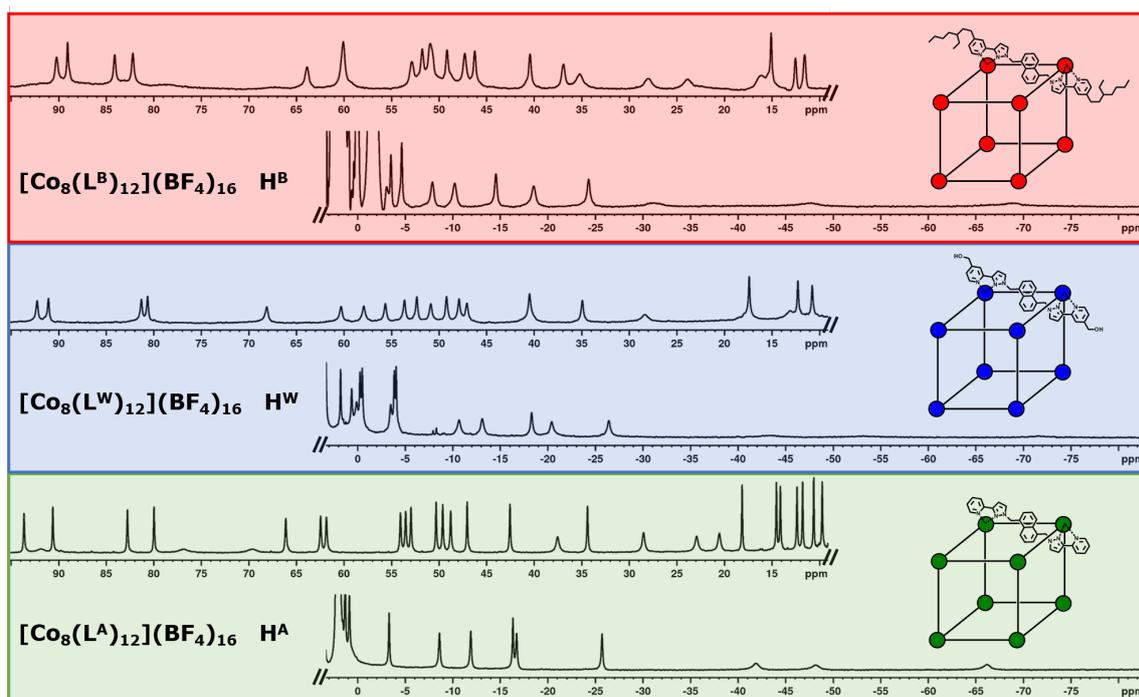
Fig. 2.13 ^{1}H-NMR of three cubic cage complexes, (bottom) $\mathbf{H^A}$, (middle) $\mathbf{H^W}$, (top) $\mathbf{H^B}$ each in their respective solvent (CD$_3$CN/D$_2$O/CD$_2$Cl$_2$)

It is clear that the ^{1}H-NMR spectrum for $\mathbf{H^A}$ has the sharpest signals while the spectrum for $\mathbf{H^B}$ has broader signals. This is as a result of the larger groups on the external surface of the cage complex that make the complex tumble more slowly in solution and lead to the broadening of signals.
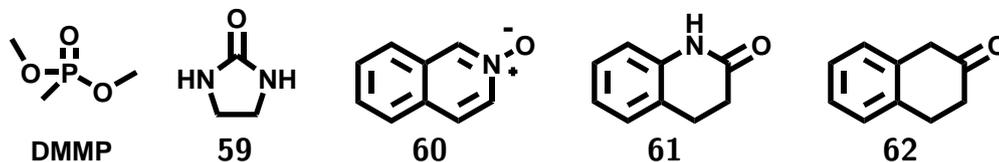
Fig. 2.14 Guests bound in $\mathbf{H^B}$ in dichloromethane

Host-guest chemistry of $\mathbf{H^W}$ and how binding constants are determined have been discussed briefly in Chapter 1.4.2 and are discussed in more detail in Chapter 3. Here a selection of guests have been bound with $\mathbf{H^B}$ and binding constants in dichloromethane have been determined. Two of these guests, DMMP and **59**, were found to bind in fast exchange on the NMR timescale.
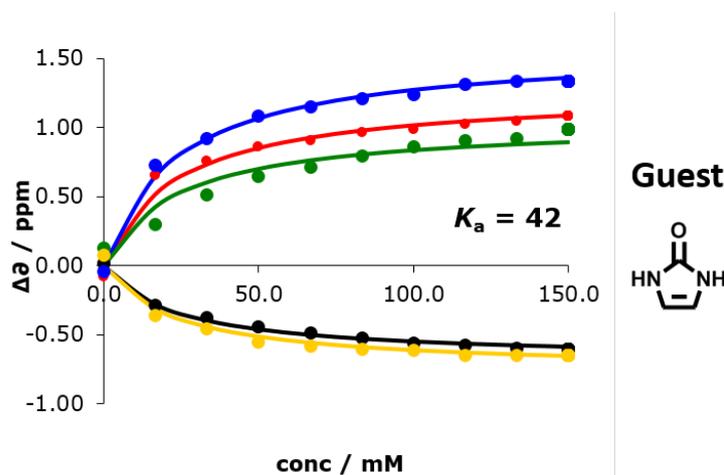
Fig. 2.15

Binding of the urea derivative **59** with $\mathbf{H^B}$ fits a 1:1 guest:host binding isotherm with a binding constant of $K_a = 42 \, \text{M}^{-1}$. Binding of DMMP with $\mathbf{H^B}$ instead fits a 2:1 guest:host isotherm which is the same as seen for binding of DMMP with $\mathbf{H^W}$ (Chapter 3.2.1.3)

The binding constants are displayed in Table 2.1. An assumption is made that only the change in solvent causes the changes the guest binding strength. The determined binding constants have one repeat (DMMP) or no repeat (**59**) and as such should not yet be considered definitive. While it appears that the binding of these two guests is similar

Table 2.1 Data for guest binding. [a] Data from reference[93] [b] Data from reference[81]

Solvent	Water	Acetonitrile	Dichloromethane
	$(K_a \; [/\text{M}]) \; \Delta G^o \; [\text{kJ mol}^{-1}]$		
DMMP	$(7(2))^a$ **-4.5(6)**^{a}	$(4(1))^a$ **-2.3(5)**^{a}	$(7.1(3))$ **-4.8(1)**
59	$(30(7))^b$ **8.4(6)**^{b}	$(1(1))^b$ **0(2)**^{b}	(42) **-9.3**

in water and dichloromethane it is apparent that the binding is significantly weaker for each when in acetonitrile. In both cases, the "middle" solvent (acetonitrile) gives weaker binding than either of the "extreme" solvents (water or dichloromethane).
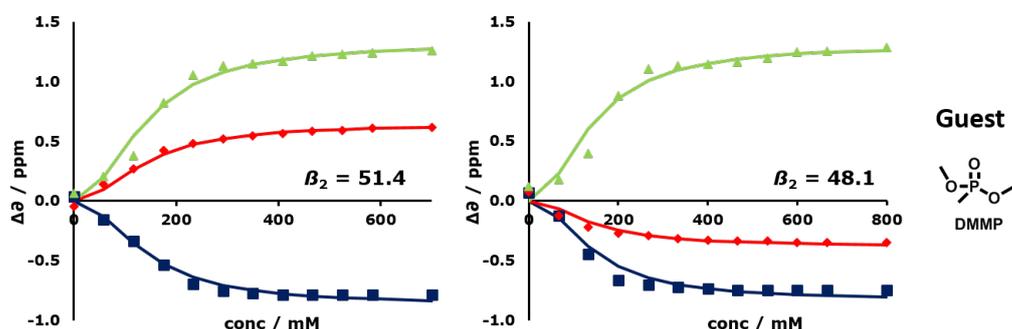
Fig. 2.16 Fitting of a 2:1 binding isotherm to the binding of DMMP with $\mathbf{H^B}$ in dichloromethane

The binding in water is driven by the hydrophobic effect and has a penalty for taking a polar group into a hydrophobic environment. Binding in dichloromethane will have no contribution from the hydrophobic effect and there will be no significant penalty from removing a polar group from the solvent. However there will be a favourable contribution from the interaction of the polar group with the converging protons in the *fac* vertices within the cage cavity. Binding in acetonitrile will not have the same significant contribution from the hydrophobic effect, as seen in water, but will still have a penalty for removing polar groups from the bulk solvent. As a result the binding in acetonitrile has the worst of both worlds (neither hydrogen bonding nor solvophobic effect are optimal) and so the weaker binding is to be expected.

The three guests, **60**, **61** and **62**, were all found to bind in slow exchange on the NMR timescale and therefore separate signals for the host and the host-guest complex are observed in the ^{1}H-NMR titrations. However upon binding of these guest the symmetry of the cage can be seen to be broken; likely as a result of the guests interacting with the *fac* vertices and not being able to freely rotate within the cage cavity on the NMR timescale.
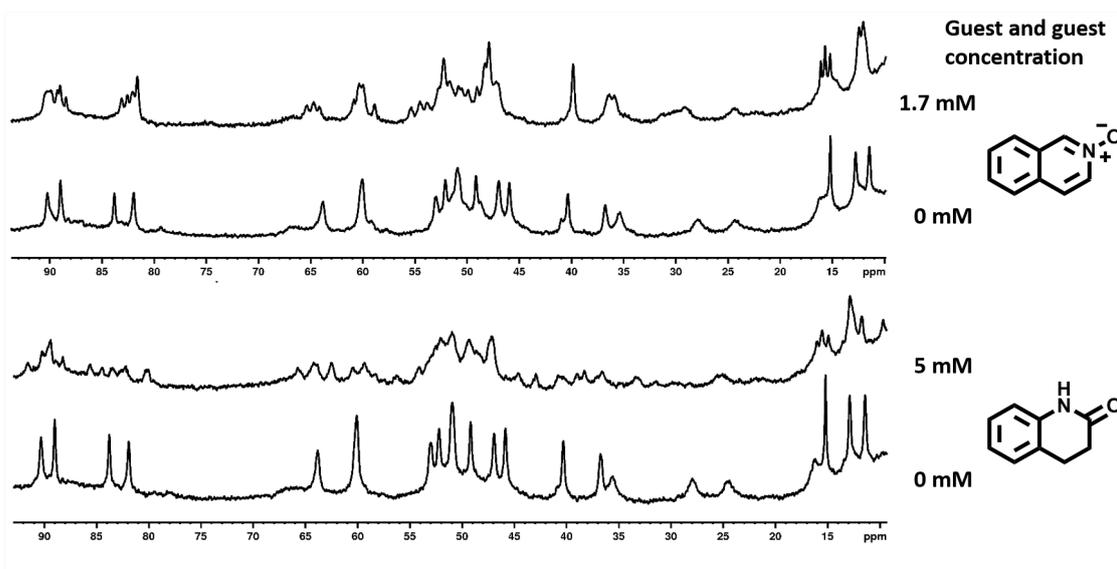
Fig. 2.17 Example spectrum of guests binding in slow exchange and splitting the symmetry of the host which precluded binding constants from being measured

The result of this is that each host signal, with increase of guest concentration, turns into multiple host-guest complex signals. Coupled with the broadness of the peaks this splitting of the host-guest complex peaks means that the spectra cannot be deconvoluted with any confidence and the binding constants cannot be determined using ^{1}H-NMR spectroscopy.
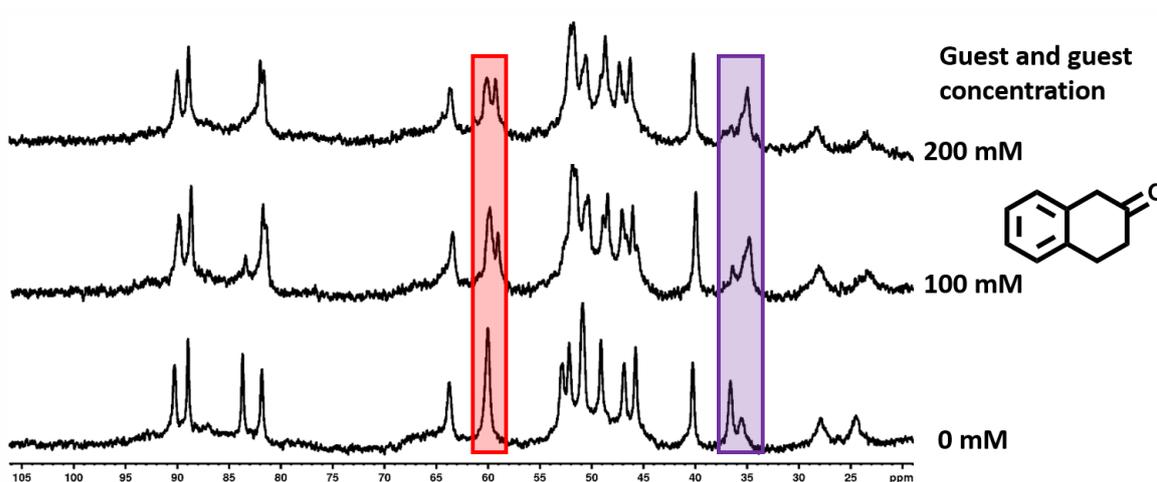
Fig. 2.18 Best example of guest binding with $\mathbf{H^{B}}$ in slow exchange and splitting the symmetry of the host. One peak changes completely to the host-guest signal (purple) while the second host peak goes to two signals (red)

2.5 Conclusion

Initially $\mathbf{L^{Et}}$ was successfully synthesised using an analogous route to the synthesis[81] of $\mathbf{L^{W}}$. Combining $\mathbf{L^{Et}}$ and $Co(BF_4)_2$ in a 3:2 ratio produced an M_8L_{12} cubic cage which was soluble in acetonitrile but not soluble in either dichloromethane or chloroform as had been desired.

Alkylation of 4-methyl-pyridine **36** gave 4-(3-ethyl-heptyl)pyridine **39** and 4-(cyclohexylmethyl)pyridine **40**. An analogous synthetic route[81] to $\mathbf{L^{W}}$, to synthesise the corresponding 2-acetyl-pyridines **45** and **46**. However the Grignard step to give **45** and **46** gave a very low yield so this route was clearly not viable.

The analogous route was used to successfully create $\mathbf{L^{Me}}$ with the intention to alkylate the 4-pyridyl units *after* forming the ligand backbone. The crystal structure of the ligand was obtained but the desired conversion of $\mathbf{L^{Me}}$ to give $\mathbf{L^{B}}$ did not work. Instead an alternative synthetic route from the alkylated pyridines **39** and **40** was found along with a one-step reaction to give **45** and **46**. The new one-step acylation reaction replaces three steps from the original synthetic route while also improving the overall yield. The ligands $\mathbf{L^{B}}$ and $\mathbf{L^{C}}$ were obtained and both complexed with $Co(BF_4)_2$ to give the respective M_8L_{12} cubic cage complexes ($\mathbf{H^{B}}$ and $\mathbf{H^{C}}$). The $\mathbf{H^{C}}$ complex was not soluble in the desired solvents, dichloromethane or chloroform, but $\mathbf{H^{B}}$ was soluble and its host-guest chemistry was investigated.

Initial results gave binding constants for DMMP and **59** with both having a higher binding free energy compared to the same guest binding in $\mathbf{H^{A}}$ acetonitrile. The binding constant for the remaining guests, which were in slow exchange, could not be determined. Three changes can be made to improve future titrations and possibly allow the determination of guest binding constants when in slow exchange: 1) an increased cage concentration; 2) increased temperature; 3) optimising the NMR pulse sequence to use spin-echo (Chapter 3.4.1.2). The combination of these changes should give sharper and better defined signals, with a increased signal to noise ratio, that can be deconvoluted satisfactorily. The directed lithiation reaction was applied to the synthesis of $\mathbf{L^{W}}$ reducing the overall eight step synthesis to just six steps while also significantly improving the overall yield.

A new ligand $\mathbf{L^{ter}}$ with one pyridine-pyrazole bidentate termini and one terpyridine tridentate termini was successfully synthesised. The ligand was complexed with $FeCl_2$ to give $[Fe(L^{ter})]Cl_2$. Further coordination with Co^{II} ions gave an unknown mixture. Future development of the ligand and investigation of complexes could lead to new interesting mixed-metal cage complexes.

2.6 Experimental

Petroleum refers to the fraction of Petroleum Ether which boil between $40\,°C$ to $60\,°C$. All solvents used were of HPLC quality and were purchased from Fisher Scientific® or Sigma-Aldrich®. All dry solvents were obtained from the in-house Grubbs solvent purification system.

Chromatography was performed under gravity on silica gel 60 unless otherwise stated. All chromatography columns were monitored by TLC using pre coated silica plates unless otherwise stated. TLC were visualised with UV light unless otherwise stated.

^{1}H-NMR spectra were recorded on a Bruker AV-400 spectrometer at room temperature unless otherwise stated. High-resolution mass spectra were recorded on a MicroMass LCT operating in electrospray mode unless otherwise stated. Sodium hydride $60\,\%$ dispersion or NaH $60\,\%$ dispersion both refer to dispersion in mineral oil. This dispersion was used without removal of the dispersion oil.

2.6.1 Synthesis of parent ligand L_A

26 → **27** → **17**

14 → **15** → **16**

17

L^A
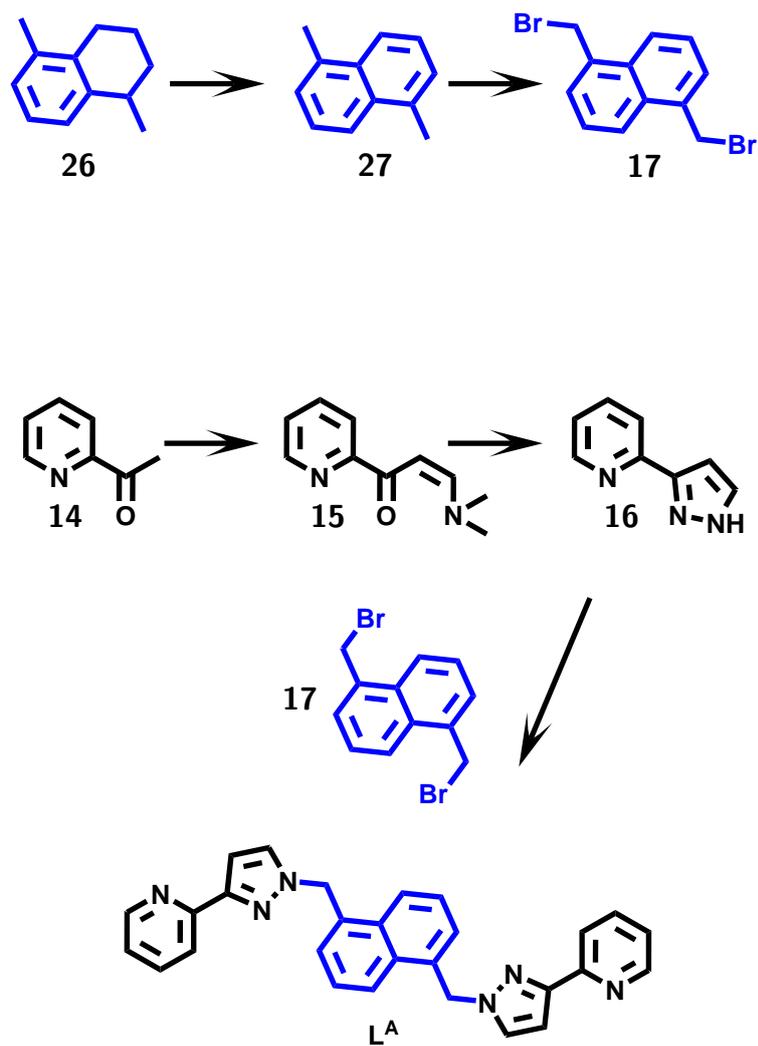
Scheme 2.23

2.6.1.1 1,5 dimethyl naphthalene 27

26   **27**
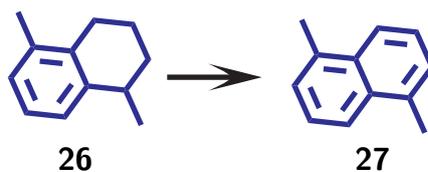
Triphenyl methanol (21.5 g, 82.8 mmol) was dissolved in TFA (30 mL) and stirred until the reaction has finished. 1,2,3,4-tetrahydro-1,5-dimethylnaphthalene **26** (6.0 g, 37.5 mmol) was then added to the mixture and the resulting solution was refluxed at 75 °C for 1 h. The reaction was quenched with *sat. aq.* potassium carbonate (50 mL) and then extracted with diethyl ether (4 x 70 mL). Additional water was added before extraction if a solid persisted after the first addition of diethyl ether. The organic layers were collected, dried ($MgSO_4$) and then solvent was removed *in vacuo* to yield the crude product. Chromatography (100 % hexane) yields the product.

Yield = 3.32 g, 21.0 mmol, 56 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 7.91 (d, J = 7.3 Hz, 2H), 7.44 (dd, J = 8.5 Hz, J = 6.9 Hz, 2H), 7.36 (d, J = 6.8 Hz, 2H), 2.73 (s, 6H)

MS-EI = m/z 156.1[M-H]$^+$ 141.1[M-CH$_3^+$]

2.6.1.2   1,5 bis-bromomethyl naphthalene 17
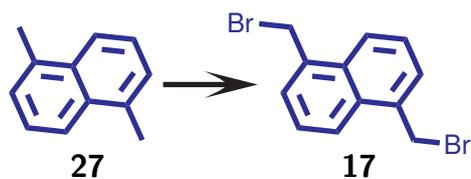
1,5-dimethyl-naphthalene **27** (3.3 g, 21.0 mmol) was dissolved in chloroform (150 mL) and stirred. N-bromo-succinimide (NBS) (8.23 g, 46.2 mmol) was added followed by azobisisobutyronitrile (AIBN) (approx. 5 mg) and the reaction was heated to reflux. The stirred solution was then exposed to a halogen lamp in a cycle of 15 min on then 15 min off. After two hours the solvent was removed *in vacuo* and the resulting solid was washed with water to yield product as an off-white solid.

Yield = 4.54 g, 14.5 mmol, 69 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.23 (dd, $J = 16.5$ Hz $J = 8.6$ Hz, 1H), 7.60 (ddd, $J = 15.2$ Hz $J = 7.5$ Hz $J = 4.1$ Hz, 2H), 4.98 (s, 2H)

MS-ES = (%) m/z 314.2 (100) [MH]$^+$, 312.2 (43), 316.2 (44)

2.6.1.3 1-(pyridin-2-yl)-(dimethylamino)prop-2-en-1-one 15[75]

14   **15**
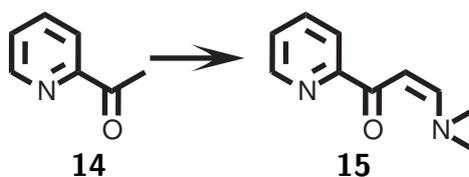
2-acetyl-pyridine **14** (115 mL, 1025.3 mmol) was added to DMA-DMF (280 mL) and refluxed overnight (90 °C) resulting in a colour change yellow to orange to brown. The mixture was transferred with methanol to into a 1 L RBF and a majority of the solvent is removed. Chloroform (200 mL) was added to the solution and then hexane (500 mL) was layered in before being left in the freezer overnight. Solvent was decanted, excess solvent removed then the chloroform/hexane layering was repeated. The resulting yellow brown crystals were washed with hexane and collected as product.

Yield = 127.14 g, 721.5 mmol, 70 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.65 (d, J = 4.7 Hz, 1H), 8.17 (d, J = 7.6 Hz, 1H), 7.94 (d, J = 12.7 Hz, 1H), 7.82 (td, J = 7.6 Hz J = 1.6 Hz, 1H), 7.38 (dd, J = 7.4 Hz J = 4.8 Hz, 1H), 6.47 (d, J = 12.4 Hz, 1H), 3.20 (s, 3H), 3.02 (s, 3H)

MS-ES = m/z 177.1 [M$^+$]

2.6.1.4 2-(pyrazol-3-yl) pyridine 16[75]

15   **16**
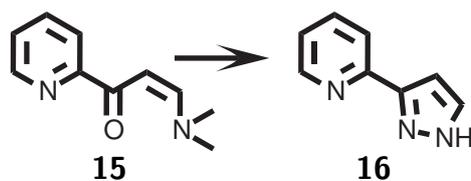
Compound **15** (70 g, 397 mmol) was dissolved in ethanol (100 mL). Hydrazine monohydrate (130 mL) was added and the solution was heated to 60 °C for 1 h. The mixture was poured into a large conical flask pre-filled with water (1 L) and refrigerated overnight. The resulting crystals, filtered and washed with water, were collected as product.

Yield = 44.7 g, 308 mmol, 78 %

^{1}H-NMR (**400 MHz, CDCl$_3$**): δ (ppm) = 11.31 (s, 1H), 8.66 (dt, J = 4.9 Hz J = 1.4 Hz, 1H) 7.80–7.73 (m, 2H), 7.69 (d, J = 2.0 Hz, 1H), 7.28–7.23 (m, 1H), 6.82 (d, J = 2.0 Hz, 1H)

MS-ES = m/z 146.1 [M$^+$]

Accurate Mass = calc. (C$_8$H$_8$N$_3$)[MH]$^+$ 146.0718, acqu. 146.0720

2.6.1.5 1,5 bis-(pyridin-2-yl-pyrazol-1-yl)methylnaphthalene (Ligand L^A)[75]
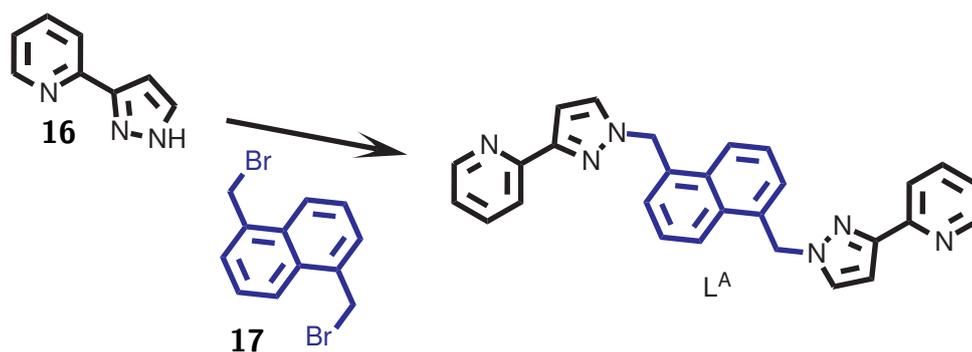
Compound **16** (6 g, 19.1 mmol) and spacer **17** (5.82 g, 40.1 mmol) were dissolved in THF (200 mL). Separately a solution of NaOH (4 g, 100 mmol) in water (50 mL) was prepared. The two solutions were mixed together and refluxed at 70 °C overnight. Teflon tap was used on all joints to avoid glass joints "welding" together. The aqueous layer was separated from the precipitated containing organic layer. After a further water wash the organic layer was filtered to collect the solid. The solid was washed with ice-cold THF and collected as product.

Yield = 5.65 g, 12.8 mmol, 67 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.67 (d, J = 4.7 Hz, 2H), 8.08 (d, J = 8.5 Hz, 2H), 7.99 (d, J = 7.9 Hz, 2H), 7.75 (td, J = 7.8 Hz J = 1.7 Hz, 2H), 7.52 (dd, J = 13.7 Hz J = 6.5 Hz, 2H), 7.36 (d, J = 7.0 Hz, 2H), 7.28 (s, 2H), 7.23 (dd, J = 6.9 Hz J = 5.4 Hz, 2H), 6.88 (d, J = 2.3 Hz, 2H), 5.90 (s, 4H)

^{13}C-NMR (101 MHz**, CDCl$_3$):** δ (ppm) = 152.27 (C), 151.72 (C), 149.50 (CH), 136.58 (CH), 132.38 (C), 131.78 (C), 130.84 (CH), 127.40 (CH) 126.48 (CH), 124.39 (CH), 122.40 (CH, 120.16 (CH), 104.88 (CH), 54.71(CH$_2$)

MS-ES = (%) m/z 465.3 (6) [MNa$^+$], 443.2 (24) [MH]$^+$, 222.1 (100) [MH$_2$]$^{2+}$

Accurate Mass = calc. (C$_{28}$H$_{23}$N$_6$) [MH]$^+$ 443.1984, acqu. 443.1675

2.6.2 Synthesis of 4-hydroxymethylpyridin-2-yl-pyrazole

Scheme 2.24

2.6.2.1 4-tert-butyldimethylsilyloxymethylpyridine 19[81]
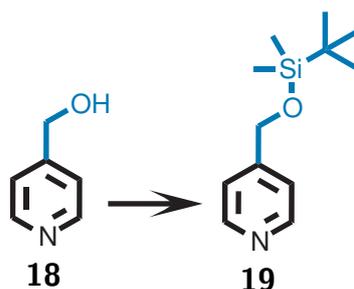
Method 1 (Original) Imidazole (14 g, 205.6 mmol) was dissolved in DMF/DCM 50:50 mix (100 mL) and stirred under N_2. *tert*-Butyldimethylsilyl chloride (TBDMSCl) (24.9 g, 165.2 mmol) was added slowly over 20 min while vigorously stirring. 4-methanol pyridine **18** (15 g, 135.7 mmol) was then added slowly over 5 min and resulting solution was stirred overnight. Solvent was removed *in vacuo* and then water (150 mL) was added. The product was extracted with EtOAc/hexane 50:50 mix (4 x 100 mL). Organic layers were combined, dried (MgSO$_4$) and evaporated *in vacuo* to give crude product.

Yield = 30.1 g, 134.6 mmol, 99 %

Method 2 (New) Imidazole (36.5 g, 536 mmol) was suspended in DCM (300 mL). TBDMSCl (64.2 g, 426 mmol) was added slowly over 10 min and solution was left to stir for another 10 min giving a white suspension. 4-methanol pyridine **18** (38.6 g, 354 mmol) was added and the reaction was stirred overnight to yield a light pink suspension. The solvent was remove *in vacuo*. Water (300 mL) was added to the resulting residue and shaken vigorously before the product was extracted with EtOAc (3 x 100 mL). Extracted layers were combined, dried (MgSO$_4$) and solvent removed *in vacuo* to give the pure product as a yellow oil.

Yield = 78.6 g, 352 mmol, 99 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.57 (d, $J = 5.9$ Hz, 2H), 7.27 (d, $J = 5.8$ Hz, 2H), 4.76 (s, 2H), 0.97 (s, 9H), 0.13 (s, 6H)

MS-ES = m/z 224.1 [MH]$^+$

2.6.2.2 (4-tert-butyldimethylsilyloxymethylpyridine)-N-oxide 20[81]

19   **20**
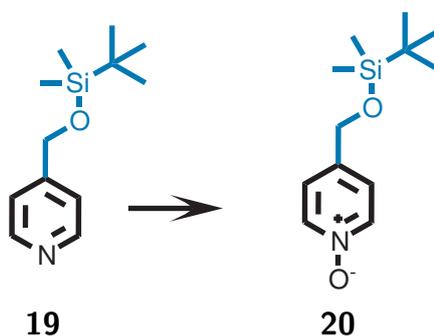
19 (30.41 g, 136.3 mmol) was dissolved in DCM (300 mL). *meta*-chloroperoxybenzoic acid (mCPBA) (40.32 g, 163.5 mmol) was added slowly and resulting mixture was stirred overnight. Product was washed into NaOH (1 M, 400 mL) and then extracted with DCM (3 x 100 mL). Organic extracts were dried (MgSO$_4$) and solvent removed *in vacuo* to yield crude product.

Yield = 29.32 g, 122.5 mmol, 89 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.19 (d, $J = 6.8$ Hz, 2H), 7.24 (d, $J = 6.8$ Hz, 2H), 4.70 (s, 2H), 0.94 (s, 9H), 0.11 (s, 6H)

2.6.2.3 (4-tert-butyldimethylsilyloxymethylpyridin-2-yl)-2-nitrile 21 [81]

20

21
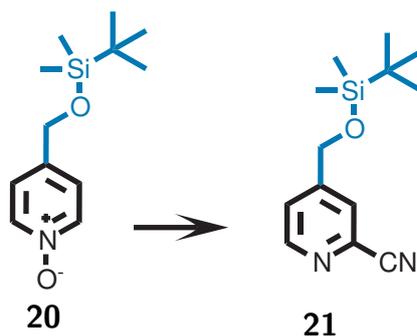
20 (29.31 g, 123 mmol) was dissolved in DCM (300 mL). Trimethylsilyl cyanide (TMSCN) (20.4 mL, 153 mmol) was slowly added to the solution and stirred for 10 min at room temperature. Dimethylcarboamoyl chloride (14.1 mL, 153 mmol) was added and reaction was stirred overnight. The reaction was quenched with the addition of *sat. aq.* K_2CO_3 (200 mL) while stirring for 10 min. The reaction changes from a clear orange to a purple colour during quenching. Product was extracted with DCM (3 x 250 mL) and solvent removed from the combined organic layers. Water was added to the residue and stirred for 1 h. The product was then re-extracted with DCM (3 x 250 mL) and solvent was removed *in vacuo* from the combined organic layers to give a brown oil. Chromatography (10 % EtOAc/Hexane) yields a yellow oil.

Yield = 25.8 g, 104.0 mmol, 85 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.66 (d, $J = 5.0$ Hz, 1H), 7.70 (d, $J = 0.6$ Hz, 1H), 7.48 (dd, $J = 5.0$ Hz, $J = 1.4$ Hz, 1H), 4.80 (s, 2H), 0.98 (s, 9H), 0.15 (s, 6H)

2.6.2.4 (4-tert-butyldimethylsilyloxymethylpyridin-2-yl)ethan-2-one 22
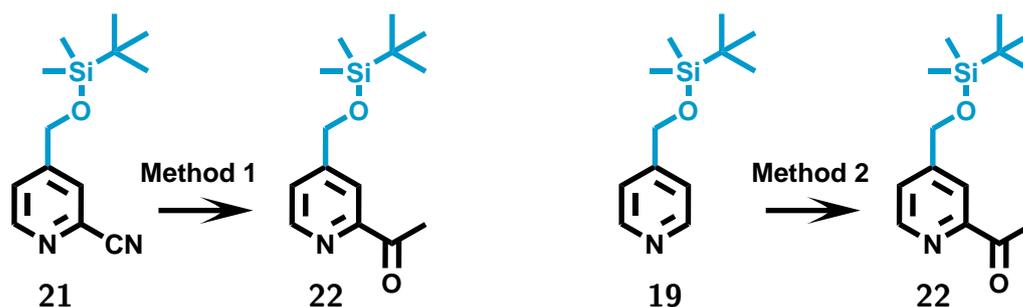
Method 1 (Original)

Product has previously been prepared with a 19 % yield following literature[81].

Method 2 (New)

Dimethylethanolamine (DMAE) (12.5 mL, 125 mmol) was added to toluene (100 mL) within dried glassware under N_2 and cooled to below 0 °C using rock-salt/ice bath. n-Butyllithium n-BuLi (100 mL, 2.5 mol dm^{-3}, 250 mmol) was added dropwise while stirring vigorously. A colour change should be noted half way through the addition. The solution was stirred for 30 min. **19** (22.4 g, 100 mmol) was added dropwise and then the solution was stirred for a further 30 min. Dimethylacetamide (DMA) (27.9 mL, 300 mmol) was added and the solution was stirred for a final 30 min. The reaction was quenched with water (200 mL) resulting precipitate re-dissolves. The layers were separated and product extracted from the aqueous layer. All the organic layers were combined, dried (MgSO$_4$), and solvent removed. The above was repeated a further two times then all three crudes were combined and chromatography (20 % EtOAc:petroleum ether) yielded product. (11.824 g) and recover starting material (23.132 g).

Yield = 11.824 g, 44.6 mmol, 25 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.66 (dd, $J = 5.0$ Hz $J = 0.6$ Hz, 1H), 7.97 (dd, $J = 1.7$ Hz $J = 0.8$ Hz, 1H), 7.52 (ddd, $J = 4.9$ Hz $J = 1.7$ Hz $J = 0.9$ Hz, 1H), 4.81 (s, 2H), 2.75 (s, 3H), 0.98 (s, 9H), 0.14 (s, 6H)

MS-ES = m/z 266.2 [MH]$^+$

2.6.2.5 1-(4-tert-butyldimethylsilyloxymethylpyridin-2-yl)-(dimethylamino)prop-2-en-1-one 23[81]

22   **23**
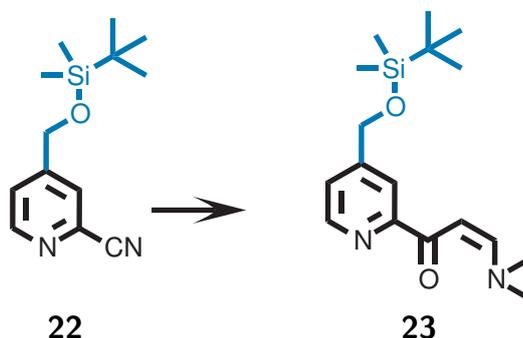
Compound **22** (3.78 g, 14.3 mmol) was added to excess DMF-DMA (5 mL) with solvent DMF (15 mL) and heated to 110 °C overnight. The solvent was removed *in vacuo* with aid of toluene (3 x 100 mL) washes. This yielded crude product as a brown solid which was taken forward without purification.

Yield = 4.55 g, 14.2 mmol, 99 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.57 (d, J = 5.1 Hz, 1H), 7.70 (s, 1H), 7.68 (d, J = 2.0 Hz, 1H), 7.23 (dd, J = 5.1 Hz J = 0.7 Hz, 1H), 6.79 (s, 1H), 4.81 (s, 2H), 1.65 (s, 6H), 1.00 (s, 9H), 0.15 (d, J = 2.6 Hz, 6H)

MS-EI = m/z 321.2 [MH]$^+$

2.6.2.6 3-(4-tert-butyldimethylsilyloxymethylpyridin-2-yl)-pyrazole 24[81]
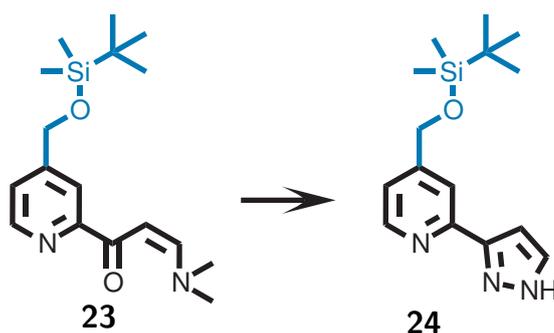
Compound **23** (4.55 g, 14.2 mmol) was dissolved in ethanol (50 mL). Hydrazine monohydrate (20 mL, 400 mmol) was added and the mixture was heated at 50 °C for 30 min. The solvent was removed *in vacuo*. Water was added to quench the reaction and then product was extracted with DCM (3 x 100 mL. The combined organic layers were dried (MgSO$_4$) and solvent remove *in vacuo*. Product was yielded from chromatography (5 % MeOH/DCM) to give an orange solid.

Yield = 1.24 g, 4.3 mmol, 30 %

^{1}H-NMR (400 MHz, **CDCl$_3$**): δ (ppm) = 8.58 (d, J = 5.1 Hz, 1H), 7.70 (s, 1H), 7.68 (d, J = 2.0 Hz, 1H), 7.25 7.21 (m, 1H), 6.79 (s, 1H), 5.32 (s, 1H), 4.81 (s, 2H), 1.00 (s, 9H), 0.16 (s, 6H)

MS-ES = m/z 290.2 [MH]$^+$

Accurate Mass = calc. (C$_{15}$H$_{24}$N$_3$OSi)[MH]$^+$ 290.1689, acqu. 290.1697

2.6.3 Synthesis of oxy-methyl ligand (L^W)
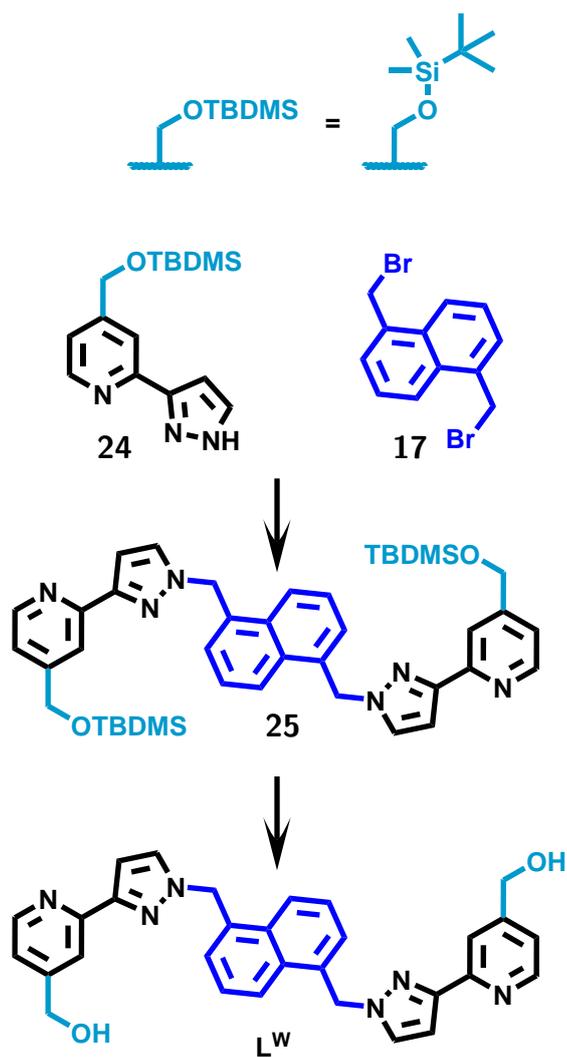
Scheme 2.25

2.6.3.1   1,5 bis-(4-tert-butyldimethylsilyloxymethylpyridin-2-yl-pyrazol-1-yl) methylnaphthalene 25[81]
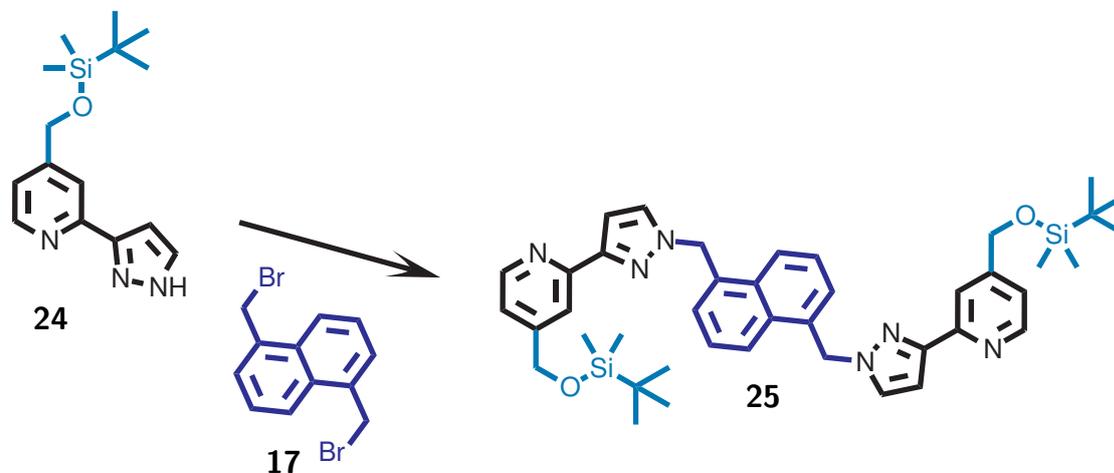
Compound **24** (2.90 g, 10.1 mmol) was added to the reaction vessel under N_2. The mixture was dissolved, while stirring, in dry THF (30 mL) followed by an excess of NaH (60 % suspension). 1,5-dibromo-naphthalene **17** (1.43 g, 4.6 mmol) was added when effervescing ends (10 min) and the solution was heated to 70 °C. The reaction was monitored by thin layer chromatography (TLC) through to complete conversion (22 h). The reaction was first cooled, then quench with MeOH, and finally evaporate to dryness *in vacuo*. Chromatography (5 % MeOH:DCM) yields product as a yellow oil.

Yield = 2.94 g, 4.0 mmol, 88 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.57 (d, J = 5.1 Hz, 1H), 8.03 (d, J = 8.5 Hz, 1H), 7.91 (d, J = 0.5 Hz, 1H), 7.45 (dd, J = 8.5 Hz J = 7.0 Hz, 1H), 7.30 (d, J = 7.1 Hz, 1H), 7.24 (s, 1H), 6.88 (d, J = 2.3 Hz, 1H), 5.82 (s, 2H), 4.78 (s, 2H), 0.99 (s, 9H), 0.15 (s, 6H)

MS-ES = m/z (%) 731.4 (8) [MH]$^+$, 366.2 (100) [MH$_2$]$^{2+}$

2.6.3.2 1,5 bis-(4-hydroxymethylpyridin-2-yl-pyrazol-1-yl) methylnaphthalene (L^W)[81]
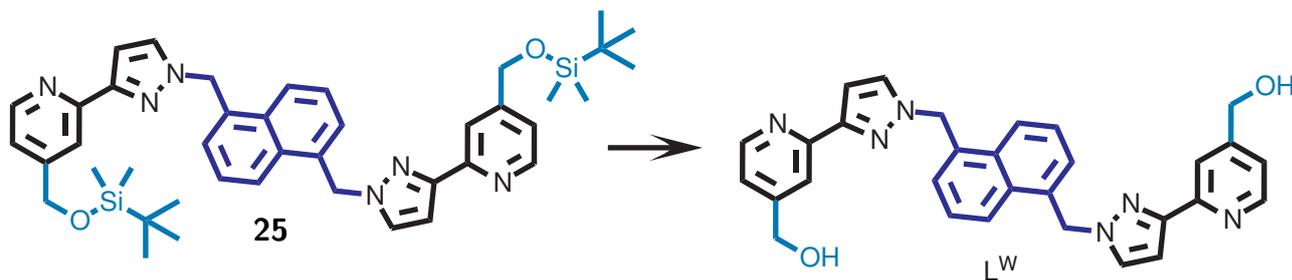
The protected ligand **25** (0.567 g, 0.77 mmol) was dissolved into THF (30 mL). Tetra-*n*-butylammonium fluoride (TBAF) (0.480 g, 1.52 mmol) was added and the mixture was stirred overnight. The precipitate was filtered off, washed with the mother liquor, and collect as a white solid product.

Yield = 0.39 g, 0.77 /millimol, 99 %

^{1}H-NMR (400 MHz, **CDCl$_3$**): δ (ppm) = 8.47 (d, J = 4.9 Hz, 2H), 8.22 (d, J = 8.3 Hz, 2H), 7.89 (s, 2H), 7.58 (t, J = 7.7 Hz, 2H), 7.27 (d, J = 7.0 Hz, 2H), 7.20 (d, J = 4.9 Hz, 2H), 6.84 (s, 2H), 5.94 (s, 4H), 5.45 (t, J = 5.1 Hz, 2H), 4.56 (d, J = 5.4 Hz, 4H)

MS-ES = m/z (%) 503.2 (44) [MH]$^+$, 252.1 (100) [MH$_2$]$^{2+}$

2.6.4 Modification of the oxy-methyl ligand and cage synthesis
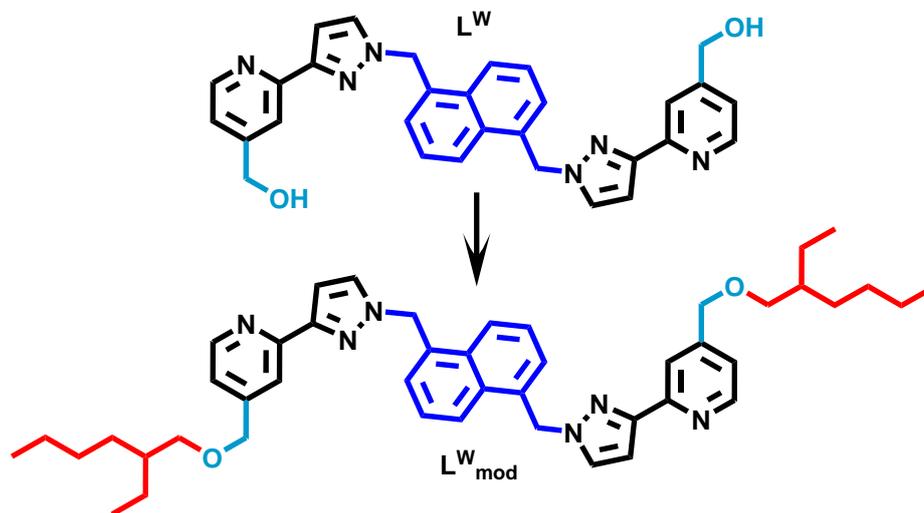
Scheme 2.26

2.6.4.1   1,5 bis-(4-[2-ethyl-hexyl]oxymethylpyridin-2-yl-pyrazol-1-yl) methylnaphthalene
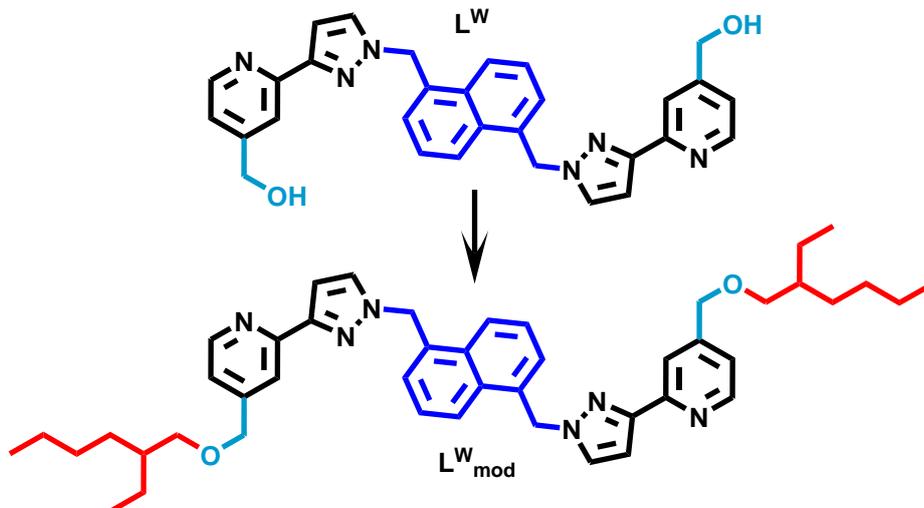
Ligand $\mathbf{L^W}$ (0.2 g, 0.4 mmol) and excess NaH (60 % dispersion) was added to the reaction vessel under N_2. The mixture was dissolved in dry THF (20 mL) and stirred until effervescing ends (10 minute). 2-ethyl-hexyl-1-bromide (0.23 g, 1.19 mmol) and Bu_4NI (10 % eq.) was added and the solution was heated to 70 °C. The reaction was monitored by TLC (10 % MeOH:DCM) and after 24 h additional 2-ethyl-hexyl-1-bromide (0.3 mL) and Bu_4NI (5 % eq.) was added. After full conversion (48 h) the reaction was cooled MeOH was added to quench the reaction. The solvent was removed *in vacuo*. Chromatography (20 % EtOAc:petroleum ether) yields product as a sticky brown solid.

Yield = 81 mg, 0.11 mmol, 28 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.52 (d, $J = 5.0$ Hz, 2H), 8.27 (d, $J = 8.6$ Hz, 2H), 8.03 (d, $J = 0.7$ Hz, 2H), 7.71 (d, $J = 2.3$ Hz, 2H), 7.56 (dd, $J = 8.4$ Hz $J = 7.1$ Hz, 2H), 7.37 (d, $J = 7.0$ Hz, 2H), 7.25 to 7.20 (m, 2H), 6.93 (d, $J = 2.3$ Hz, 2H), 5.95 (s, 4H), 4.57 (s, 4H), 3.44 (d, $J = 5.7$ Hz, 4H), 1.66 to 1.52 (m, 2H), 1.52 to 1.17 (m, 16H), 0.88 (ddd, $J = 15.2$ Hz $J = 11.2$ Hz $J = 7.4$ Hz, 12H)

ES-MS = m/z 727 [MH]$^+$, 749 [MNa]$^+$

2.6.5 Synthesis of the ethyl ligand ($\mathbf{L^{Et}}$) and cage
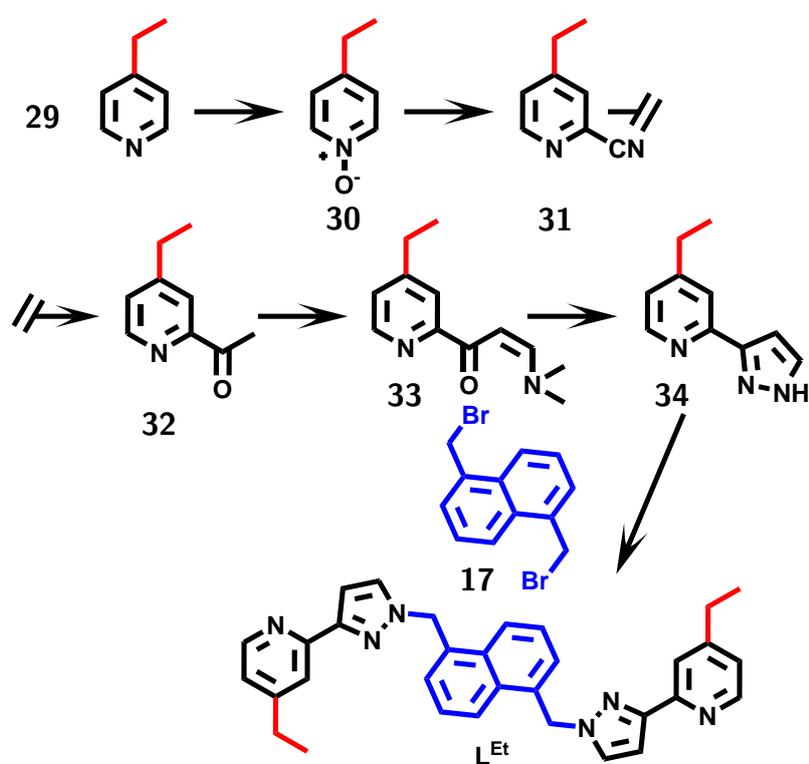
Scheme 2.27

2.6.5.1 (4-ethylpyridine)-N-oxide 30

29   **30**
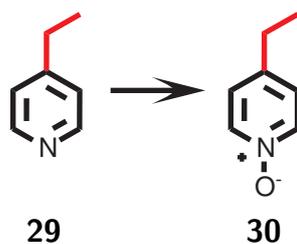
4-ethyl pyridine **29** (10 mL, 87.9 mmol) was dissolved in DCM (200 mL) at room temperature. mCPBA (26.0 g, 105.5 mmol) was added slowly over ten minutes to the stirring solution. The reaction was monitored by TLC (5 % MeOH in DCM) and after 2.5 h the reaction was quenched with aqueous NaOH (1 M 200 mL) and stirred for a further 10 min. The product was extracted with DCM (5 x 200 mL). Organic fractions were combined, dried ($MgSO_4$), and solvent removed under vacuum to give a white solid product.

Yield = 8.5 g, 69.0 mmol, 79 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.05 (d, J = 6.8 Hz, 2H), 7.04 (d, J = 6.7 Hz, 2H), 2.57 (q, J = 7.6 Hz, 2H), 1.16 (t, J = 7.6 Hz, 3H)

ES-MS = m/z 124 [MH]$^+$

Accurate Mass = calc. ($C_7H_{10}NO$)[MH]$^+$ 124.0762, acqu. 124.0758

2.6.5.2 (4-ethylpyridin-2-yl)-2-nitrile 31

30 **31**
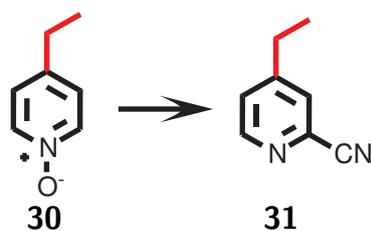
30 (7.0 g, 56.8 mmol) was dissolved in DCM (350 ml). TMSCN (9.5 ml, 71.1 mmol) was slowly added to the solution and stirred for 10 min at room temperature. Dimethylcarbamyl chloride (6.5 mL, 71.1 mmol) was added and reaction was stirred overnight. After 1 h a bung was added to prevent solvent evaporation. The reaction was quenched with the addition of *sat. aq.* K_2CO_3 (300 mL) while stirring for 10 min. The reaction changes from a clear cream to a strong magenta colour during quenching. Product was extracted with DCM (5 x 200 ml) and solvent removed from the combined organic layers. Water (250 mL) was added to the residue and stirred for 1 h. The product was then re-extracted with DCM (5 x 200 mL) and solvent was removed *in vacuo* from the combined organic layers to give a brown oil. Chromatography (10 % EtOAc/Hexane) yields a yellow oil.

Yield = 25.8 g, 104.0 mmol, 85 %

^{1}H-NMR **(**400 MHz**, CDCl$_3$):** δ (ppm) = 8.61 (d, J = 5.0 Hz, 1H), 7.57 (s, 1H), 7.37 (d, J = 5.0 Hz, 1H), 2.75 (q, J = 7.6 Hz, 2H), 1.31 (t, J = 7.6 Hz, 3H)

ES-MS = 133 [MH]$^+$

Accurate Mass = calc. $(C_8H_9N_2)$[MH]$^+$ 133.0766, acqu 122.0760

2.6.5.3 (4-ethylpyridin-2-yl)ethan-2-one 32
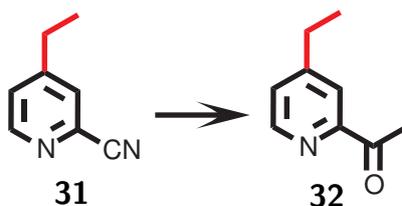
A solution of pyridine **31** (6.36 g, 48.1 mmol) in dry diethyl ether (120 mL) was split evenly into three round bottom flasks. To each flask, methyl magnesium bromide (6.41 mL, 19.2 mmol) was added dropwise and stirred for 3 h. A saturated aqueous solution of ammonium chloride (30 mL each) was added to each flask to quench the reaction. The solution for all three flasks was combined and the product was extracted with DCM. The combined organic layers were dried ($MgSO_4$) and solvent was removed under vacuum to give a dark brown oil. Chromatography (2 % MeOH/DCM) yields an orange/yellow oil.

Yield = 4.15 g, 28.5 mmol, 59 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.59 (d, J = 4.9 Hz, 1H), 7.92 (s, 1H), 7.33 (d, J = 5.0 Hz, 1H), 2.75 (s, 3H), 2.75 (q, J = 7.6 Hz, 2H), 1.30 (t, J = 7.6 Hz, 3H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 200.50 (C), 154.15 (C), 153.56 (C), 148.91 (CH), 126.80 (CH), 121.28 (CH), 28.26 (CH$_2$), 25.97 (CH$_3$), 14.27 (CH$_3$)

ES-MS = m/z 150 [MH]$^+$

Accurate Mass = calc. (C$_9$H$_{12}$NO)[MH]$^+$ 150.0919, acqu. 150.0912

2.6.5.4 1-(4-ethylpyridin-2-yl)-(dimethylamino)prop-2-en-1-one 33
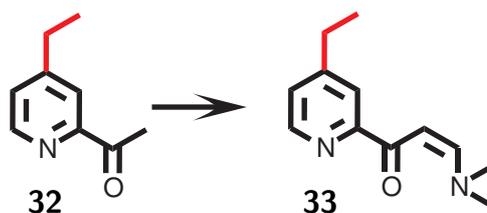
Pyridine **32** (2.7 g, 18.1 mmol) was dissolved in DMF-DMA (5.0 mL, 36.2 mmol) as solvent and reactant. The mixture was refluxed (110 °C) overnight. The solvent was then removed under reduced pressure to yield crude product. The crude product was taken onto the next step without further purification.

Yield = 3.38 g, 16.5 mmol, 91 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.51 (d, J = 4.9 Hz, 1H), 8.00 (s, 1H), 7.91 (d, J = 12.7 Hz, 1H), 7.20 (d, J = 5.0 Hz, 1H), 6.44 (d, J = 12.4 Hz, 1H), 3.18 (s, 3H), 2.99 (s, 3H), 2.71 (q, J = 7.6 Hz, 2H), 1.27 (t, J = 7.6 Hz, 3H)

^{13}C-NMR (400 MHz, CDCl$_3$): δ (ppm) = 156.15 (C), 154.72 (CH), 153.89 (C), 148.24 (CH), 125.10 (CH), 121.70 (CH), 100.26 (C), 91.35 (CH), 45.17 (CH$_3$), 37.50 (CH$_3$), 28.34 (CH$_2$), 14.34 (CH$_3$)

ES-MS = m/z 205.0 [MH]$^+$

Accurate Mass = calc. (C$_{12}$H$_{17}$N$_2$O)[MH]$^+$ 205.1341, acqu. 205.1335

2.6.5.5   3-(4-ethylpyridin-2-yl)-pyrazole 34
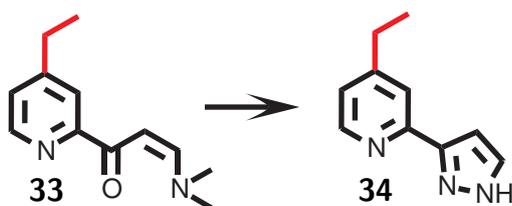
Keto enamine **33** (0.49 g, 2.41 mmol) was dissolved in ethanol (15 mL). Hydrazine monohydrate (2.34 mL, 48 mmol) was added and the solution was refluxed (60 °C) for 30 min. The solution changed from a dark brown to golden colour. The solvent was removed *in vacuo* to give an oil. Water (30 mL) was added to the oil and the product was extracted with DCM (3 x 30 mL) to give the crude product.

Yield = 0.39 g, 2.23 mmol, 92 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.55 (d, J = 5.0 Hz, 1H), 7.68 (d, J = 2.0 Hz, 1H), 7.60 (s, 1H), 7.11 (d, J = 5.1 Hz, 1H), 6.82 (d, J = 2.0 Hz, 1H), 2.73 (q, J = 7.6 Hz, 2H), 1.32 (t, J = 7.6 Hz, 3H)

ES-MS = m/z 174.1 [MH]$^+$

Accurate Mass = calc. (C$_{10}$H$_{12}$N$_3$) 174.1031, acqu. 174.1027

2.6.5.6 1,5 bis-(4-ethylpyridin-2-yl-pyrazol-1-yl)methylnaphthalene L ^{Et}
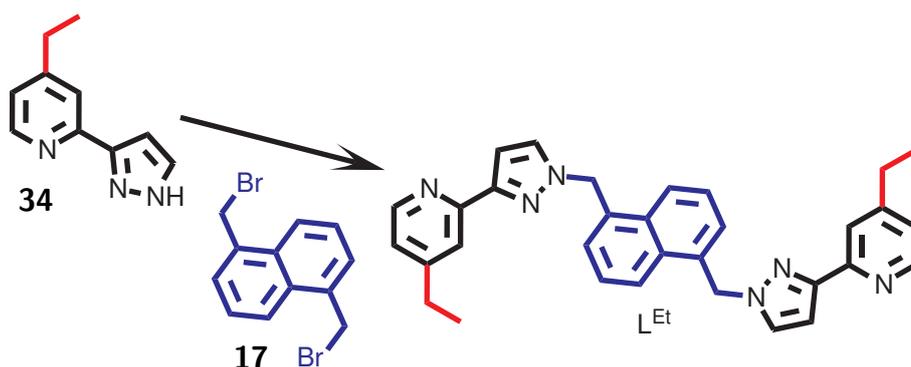
Pyridine pyrazole **34** (0.386 g, 2.23 mmol) was added to dry THF (25 mL) under nitrogen. NaH 60 % dispersion (2.2 eq, 0.098 g, 2.45 mmol) was added and reaction stirred vigorously for 10 min until bubbling stopped. 1,5 bis-bromomethyl naphtalene **17** (0.349 g, 1.11 mmol) was added and the solution was heated to 70 °C overnight. Once the reaction had gone to completion an aqueous solution of NaOH (1 mol dm^{-3}, 30 mL) was added and stirred for 10 min. Solvent was removed under vacuum and then water (40 mL) was added. The product was extracted using DCM (4 x 50 mL). Chromatography (5 % MeOH/DCM) yielded a cream coloured solid.

Yield = 0.326 g, 0.65 mmol, 59 %

^{1}H-NMR (400 MHz, **CDCl$_3$):** δ (ppm) = 8.53 (d, J = 5.1 Hz, 1H), 8.04 (d, J = 8.5 Hz, 1H), 7.85 (s, 1H), 7.49 (dd, J = 8.3 Hz J = 7.2 Hz, 1H), 7.32 (d, J = 7.0 Hz, 1H), 7.26 (d, J = 2.3 Hz, 1H), 7.07 (dd, J = 5.0 Hz J = 1.5 Hz, 1H), 6.88 (d, J = 2.3 Hz, 1H), 5.88 (s, 2H), 2.71 (q, J = 7.6 Hz, 2H), 1.30 (t, J = 7.6 Hz, 3H)

^{13}C-NMR (101 MHz, **CDCl$_3$):** δ (ppm) = 153.70 (C), 152.07 (C), 151.92 (C), 149.35 (CH), 132.41 (C), 131.71 (C), 130.86 (CH), 127.28 (CH), 126.47 (CH), 124.30 (CH), 122.25 (CH), 119.70 (CH), 104.96 (CH), 54.66 (CH$_2$), 28.36 (CH$_2$), 14.45 (CH$_3$)

ES-MS = m/z (%) 250 (100) [MH$_2$]$^{2+}$, 499 (18) [MH]$^+$, 521 (4) [MNa]$^+$

Accurate Mass = calc. (C$_{32}$H$_{22}$N$_6$)[MH]$^+$ 449.2610, acqu. 499.2612

2.6.6   Functionalization of pyridine

36 → **38**

36 → **39**

36 → **40**
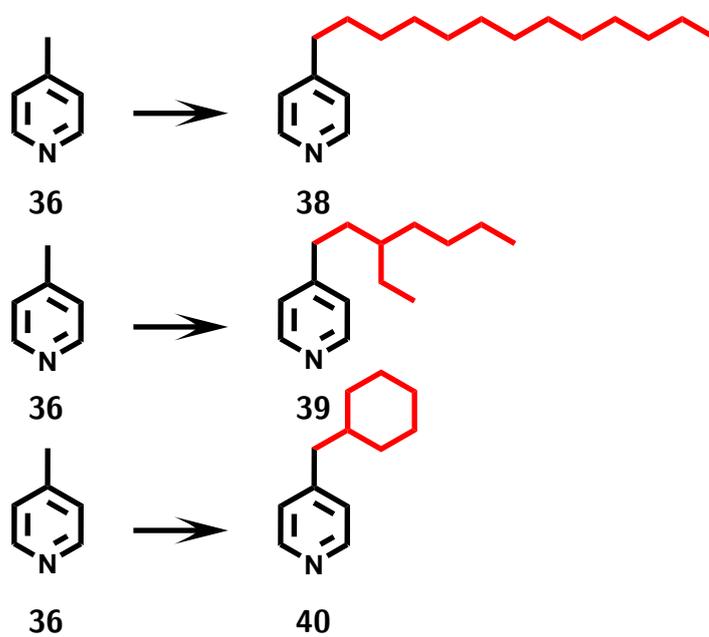
Scheme 2.28

2.6.6.1 4-tridecyl-pyridine 38

36   **38**
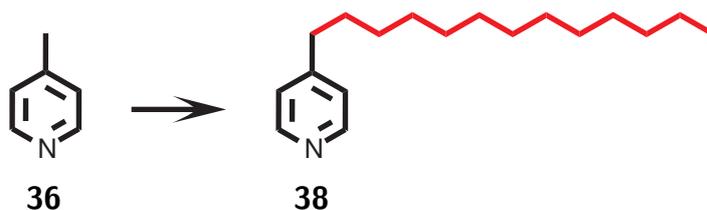
All glassware was either oven dried for 48 h or flame-dried. LDA (4.0 mL of 1.8 mol dm^{-3} in THF/heptane/ethylbenzene, 7.2 mmol) was added dropwise over 10 min to a stirred solution of pyridine **36** (0.5 mL, 5.34 mmol) and dry THF (5 mL) under nitrogen at $-78\,°$C. After stirring at $-78\,°$C for 30 min, a solution of 1-bromo-dodecane (0.83 mL, 3.6 mmol) in dry THF (5 mL) was added dropwise over 5 min. The mixture was allowed to warm to room temperature and stirred for 20 h. The reaction was quenched by the addition of *sat. aq.* NH$_4$Cl (10 mL) followed by water (10 mL). The product was extracted using EtOAc (3 x 20 mL). All organic layers were combined, dried (MgSO$_4$), and concentrated *in vacuo* to give a yellow-orange oil. Chromatography (20 % EtOAc/Petroleum) yielded 4-product as a pale yellow oil.

Yield = 800 mg, 3.06 mmol, 85 %

^{1}H-NMR (250 MHz, CDCl$_3$): δ (ppm) = 8.44 (td, J = 4.6 Hz J = 1.6 Hz, 2H), 7.07 (dd, J = 3.3 Hz J = 2.7 Hz, 2H), 2.72 to 2.45 (m, 2H), 1.71 to 1.43 (m, 2H), 1.39 to 1.05 (m, 20H), 0.96 to 0.72 (m, 3H)

EI-MS = m/z (%) 246.3 (3) [M$-$CH$_3$]$^+$, 232.3 (6) [M$-$C$_2$H$_5$]$^+$, 218.3 (8) [M$-$C$_3$H$_7$]$^+$, 204.3 (5) [M$-$C$_4$H$_9$]$^+$, 190.3 (4) [M$-$C$_5$H$_{11}$]$^+$, 176.2 (3) [M$-$C$_6$H$_{13}$]$^+$, 162.2 (6) [M$-$C$_7$H$_{15}$]$^+$, 148.2 (4) [M$-$C$_8$H$_{17}$]$^+$, 106.2 (100) [M$-$C$_{11}$H$_{23}$]$^+$

2.6.6.2 4-[3-ethyl-heptyl]pyridine 39

36   **63**
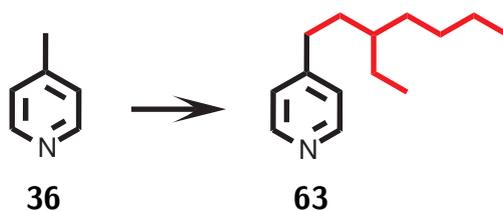
Lithium diisopropylamide (LDA) (60 mL, 1.8 M in THF/heptane/ethylbenzene, 108.0 mmol) was added dropwise over 10 min to a stirred solution of pyridine **36** (9.0 mL, 92.5 mmol) and dry THF (200 mL) under nitrogen. After stirring at $-78\,°C$ for 30 min a solution of 1-ethylhexyl bromide (16.0 mL, 89.9 mmol) in dry THF (200 mL) was added dropwise over 5 min. The mixture was allowed to warm to room temperature over night. The reaction was quenched by the addition of *sat. aq.* NH_4Cl (60 mL) followed by water (10 mL). The product was extracted using EtOAc (4 x 60 mL). All organic layers were combined, dried ($MgSO_4$), and concentrated *in vacuo* to give the crude product as orange oil. Chromatography (40 % EtOAc/Petroleum) gave a yellow oil.

Yield = 12.57 g, 61.2 mmol, 68 %

^{1}H-NMR (400 MHz, **CDCl$_3$**): δ (ppm) = 8.50 (dd, $J = 4.4$ Hz $J = 1.6$ Hz, 2H), 7.13 (dd, $J = 4.4$ Hz $J = 1.5$ Hz, 2H), 2.62 to 2.56 (m, 2H), 1.62 to 1.55 (m, 2H), 1.42 to 1.23 (m, 9H), 0.95 to 0.84 (m, 6H)

^{13}C-NMR (101 MHz, **CDCl$_3$**): δ (ppm) = 152.19 (C), 149.63 (CH), 123.87 (CH), 38.45 (CH), 34.00 (CH_2), 32.59 (CH_2), 32.50 (CH_2), 28.83 (CH_2), 25.63 (CH_2), 23.11 (CH_2), 14.16 (CH_3), 10.74 (CH_3)

ES-MS = m/z 206 $[MH]^+$

Accurate Mass = calc. $(C_{14}H_{24}N)[MH]^+$

Elemental = calc. $[(C_{14}H_{24}N) + 0 \cdot 3\,H_2O]$ C 79.79, H 11.29, N 6.65, acqu. C 80.01, H 11.37, N 6.82

2.6.6.3 4-(cyclohexylmethyl)pyridine 40
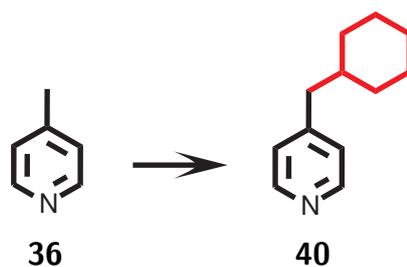
LDA (50 mL, 2 M in THF/heptane/ethylbenzene, 108.0 mmol) was added dropwise over 10 minute to a stirred solution of pyridine **36** (9.0 mL, 92.5 mmol) and dry THF (200 mL) under nitrogen. After stirring at $-78\,^{\circ}$C for 30 min a solution of 1-ethylhexyl bromide (11.0 mL, 89.7 mmol) in dry THF (200 mL) was added dropwise over 5 minute. The mixture was allowed to warm to room temperature over night. The reaction was quenched by the addition of *sat. aq.* NH$_4$Cl (60 mL) followed by water (10 mL). The product was extracted using EtOAc (4 x 60 mL). All organic layers were combined, dried (MgSO$_4$), and concentrated *in vacuo* to give the crude product as yellow oil. Chromatography (40 % EtOAc/Petroleum) gave a pale yellow oil.

Yield = 8.05 g, 45.9 mmol, 52 %

^{1}H-NMR (400 MHz, **CDCl$_3$**): δ (ppm) = 8.48 (d, $J = 5.7$ Hz, 2H), 7.08 (d, $J = 5.3$ Hz, 2H), 2.48 (d, $J = 7.1$ Hz, 2H), 1.77 to 1.62 (m, 5H), 1.61 to 1.49 (m, 1H), 1.30 to 1.07 (m, 3H), 1.03 to 0.87 (m, 2H)

ES-MS = m/z 176 [MH]$^+$

2.6.7 Synthesis of methyl ligand (L^{Me}) and functionalization
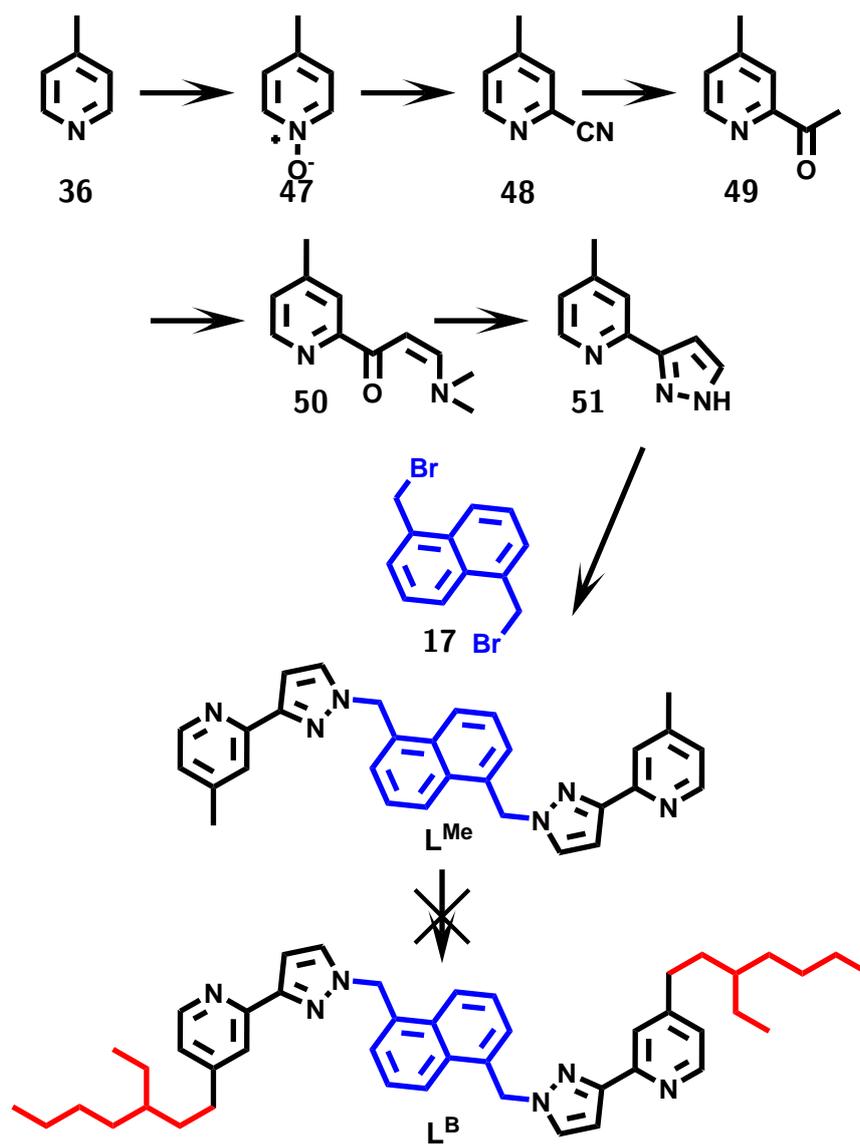
Scheme 2.29

2.6.7.1   (4-methylpyridine)-N-oxide 47

36   **47**
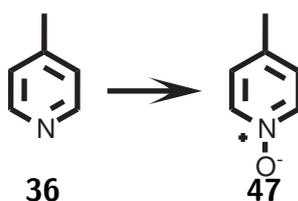
Pyridine **36** (32.1 g, 340 mmol) was dissolved in glacial acetic acid (200 mL). Hydrogen peroxide (34 mL, 300 mmol) was added to the reaction and the solution was refluxed (140 °C) for 24 h. The majority of the acetic acid was removed under vacuum. Excess di-ethyl ether was added and the product was collected as a white solid. The solvent was removed from the filtrate and then the product was again crashed out using diethyl ether. This was repeated a total of 5 times until no further product was crashed out. All solid was collected and washed with di-ethyl ether to give the desired product.

Yield = 15.2 g, 139 mmol, 46 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.04 (d, J = 6.9 Hz, 1H), 7.03 (d, J = 6.5 Hz, 1H), 2.29 (s, 2H)

ES-MS = m/z (%) 110 (98) [MH]$^+$

Accurate Mass = calc. (C$_6$H$_7$NONa)[MNa]$^+$ 132.0425, acqu. 132.0429

2.6.7.2 (4-methylpyridine)-2-nitrile 48

47   **48**
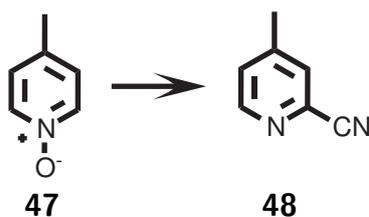
Pyridine N-oxide **47** (15.2 g, 139 mmol) was dissolved in DCM (300 mL). TMSCN (23.2 mL, 174 mmol) was added drop-wise to the reaction solution followed by the drop-wise addition of dimethyl carbamoyl chloride (16 mL, 174 mmol). The resulting mixture was stirred at room temperature overnight then quenched with *aq. sat.* K_2CO_3 (150 mL). The crude product was extracted with DCM (3 x 50 mL). The organics were collected and solvent removed under vacuum. Water (100 mL) was added to the crude and stirred for 1 h. The crude was again extracted with DCM (3 x 100 mL), dried with $MgSO_4$ and solvent removed under vacuum to give crude product. The crude was titurated with hexane to give a white solid product.

Yield = 10.0 /gram, 85 mmol, 61 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.38 (d, J = 5.0 Hz, 1H), 7.39 (s, 1H), 7.23 (d, J = 4.9 Hz, 1H), 2.28 (s, 1H)

2.6.7.3 (4-methylpyridin-2-yl)ethan-2-one 49

48 → **49**
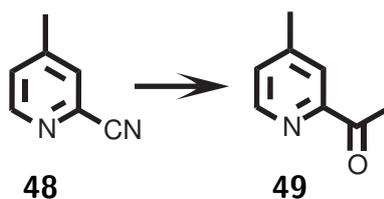
Methylmagnesium bromide $(3\,\mathrm{mol\,dm^{-3}}, 12.4\,\mathrm{mL}, 37.3\,\mathrm{mmol})$ was added drop-wise to a solution of pyridine **48** $(4\,\mathrm{g}, 34\,\mathrm{mmol})$ in di-ethyl ether $(125\,\mathrm{mL})$ at $0\,°\mathrm{C}$. After $3\,\mathrm{h}$ the reaction was quenched with *sat. aq.* $\mathrm{NH_4Cl}$ $(100\,\mathrm{mL})$. The organic layer was removed and then the product was further extracted from the *aq.* layer with DCM. All the organic layers were combined, dried with $\mathrm{MgSO_4}$ and solvent removed *in vacuo*. Product was purified through chromatography $(25\,\%\ \mathrm{EtOAc} / \mathrm{Pet\text{-}Ether})$

Yield $= 3.17\,\mathrm{g}, 23.5\,\mathrm{mmol}, 69\,\%$

^{1}H-NMR $(400\,\mathrm{MHz}, \mathbf{CDCl_3})$: $\delta\,(\mathrm{ppm}) = 8.51\ (\mathrm{d},\ J = 4.9\,\mathrm{Hz}, 1\mathrm{H}),\ 7.84\ (\mathrm{s},\ J = 0.8\,\mathrm{Hz}, 1\mathrm{H}),\ 7.26\ (\mathrm{d},\ J = 4.9\,\mathrm{Hz}, 1\mathrm{H}),\ 2.69\ (\mathrm{s}, 3\mathrm{H}),\ 2.39\ (\mathrm{s}, 3\mathrm{H})$

^{13}C-NMR $(101\,\mathrm{MHz}, \mathbf{CDCl_3})$: $\delta\,(\mathrm{ppm}) = 200.36\ (\mathrm{C}),\ 153.37\ (\mathrm{C}),\ 148.76\ (\mathrm{C}),\ 148.21\ (\mathrm{CH}),\ 127.94\ (\mathrm{CH}),\ 122.47\ (\mathrm{CH}),\ 25.90\ (\mathrm{CH_3}),\ 21.05\ (\mathrm{CH_3})$

2.6.7.4   1-(4-methylpyridin-2-yl)-(dimethylamino)prop-2-en-1-one 50
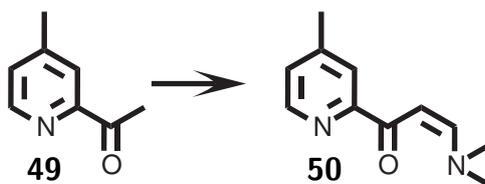
Pyridine **49** (3.43 g, 25.4 mmol) was dissolved in excess DMF-DMA (15 mL) and refluxed at 110 °C overnight. Solvent was then removed to give the crude product (4.682 g). This was used in the next step without purification.

^{1}**H-NMR** (400 MHz, **CDCl$_3$**): δ (ppm) = 8.49 (d, J = 4.9 Hz, 1H), 7.98 (s, 1H), 7.91 (d, J = 12.7 Hz, 1H), 7.18 (d, J = 4.9 Hz, 1H), 6.44 (d, J = 12.5 Hz, 1H), 3.18 (s, 3H), 2.99 (s, 3H), 2.41 (s, 3H)

^{13}**C-NMR** (101 MHz, **CDCl$_3$**): δ (ppm) = 195.02 (C), 156.01 (C), 154.73 (CH), 149.60 (CH), 147.97 (C), 126.32 (CH), 122.91 (CH), 90.30 (CH), 45.16 (CH$_3$), 37.50 (CH$_3$), 21.14 (CH$_3$)

2.6.7.5 3-(4-methylpyridin-2-yl)-pyrazole 51
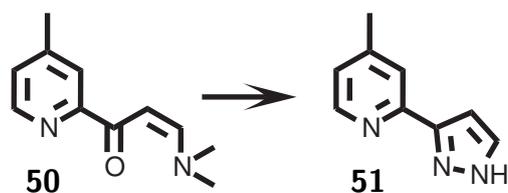
Keto enamine **50** (4.68 g crude) was dissolved in ethanol (125 mL). Hydrazine monohydrate was added and the solution was refluxed at 65 °C for 30 min. The dark brown solution turned to a golden colour. Water (125 mL) was added and the product was extracted with DCM (4 x 100 mL). Chromatography (5 % MeOH/DCM) yielded product.

Yield = 3.42 g, 21.5 mmol, 85 %

^{1}H-NMR (400 MHz, **CDCl$_3$**): δ (ppm) = 8.58 (d, J = 5.1 Hz, 1H), 7.68 (d, J = 2.0 Hz, 1H), 7.60 (s, 1H), 7.07 (d, J = 5.1 Hz, 1H), 6.81 (d, J = 2.0 Hz, 1H), 2.40 (s, 3H)

^{13}C-NMR (101 MHz, **CDCl$_3$**): δ (ppm) = 149.09, 148.39, 123.87, 121.12, 103.36, 21.16

2.6.7.6   1,5 bis-(4-methylpyridin-2-yl-pyrazol-1-yl)methylnaphthalene $\mathbf{L^{Me}}$
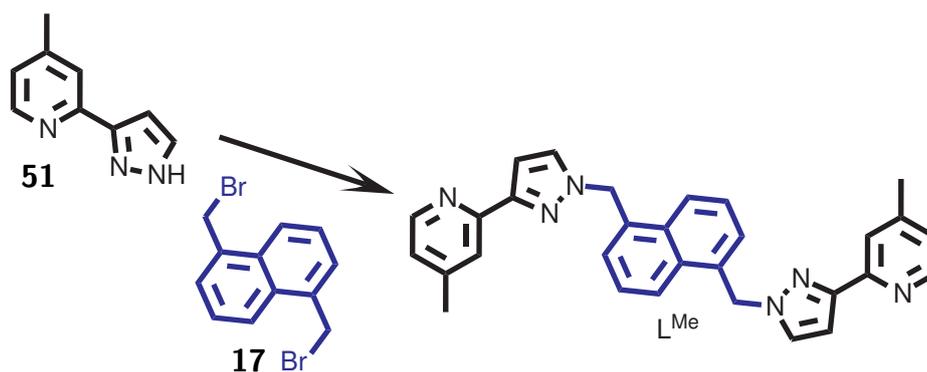
Pyridine pyrazole **51** (2.07 g, 13.0 mmol) and spacer **17** (2.00 g, 6.37 mmol) was dissolved in THF (175 mL). A solution of NaOH *aq.* (4.4 g, 20 mL, 5.5 M) was added and the reaction was refluxed at 60 % for 24 h. Upon cooling the reaction to room temperature the product sometimes precipitated out of solution. Excess aqueous layer was removed and then solvent was removed under vacuum. The residue was dissolved in DCM (50 mL) and water (5 mL). The organic layer was collected, dried using $MgSO_4$, and solvent removed under vacuum. The solid was washed with cold THF and solid was collected as the product.

Yield = 2.90 g, 6.17 mmol, 97 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.54 (d, $J = 5.2$ Hz, 2H), 8.05 (d, $J = 8.5$ Hz, 2H), 7.90 (s, 2H), 7.51 (dd, $J = 8.5$ Hz $J = 7.1$ Hz, 2H), 7.34 (d, $J = 7.0$ Hz, 2H), 7.29 (d, $J = 2.4$ Hz, 3H), 7.12 (d, $J = 5.1$ Hz, 2H), 7.03 (d, $J = 2.2$ Hz, 2H), 2.45 (s, 6H)

ES-MS = m/z (%) 236 (100) $[MH_2]^{2+}$, 471 $[MH]^+$

2.6.8 Synthesis of branched-chain ligand (L^B)
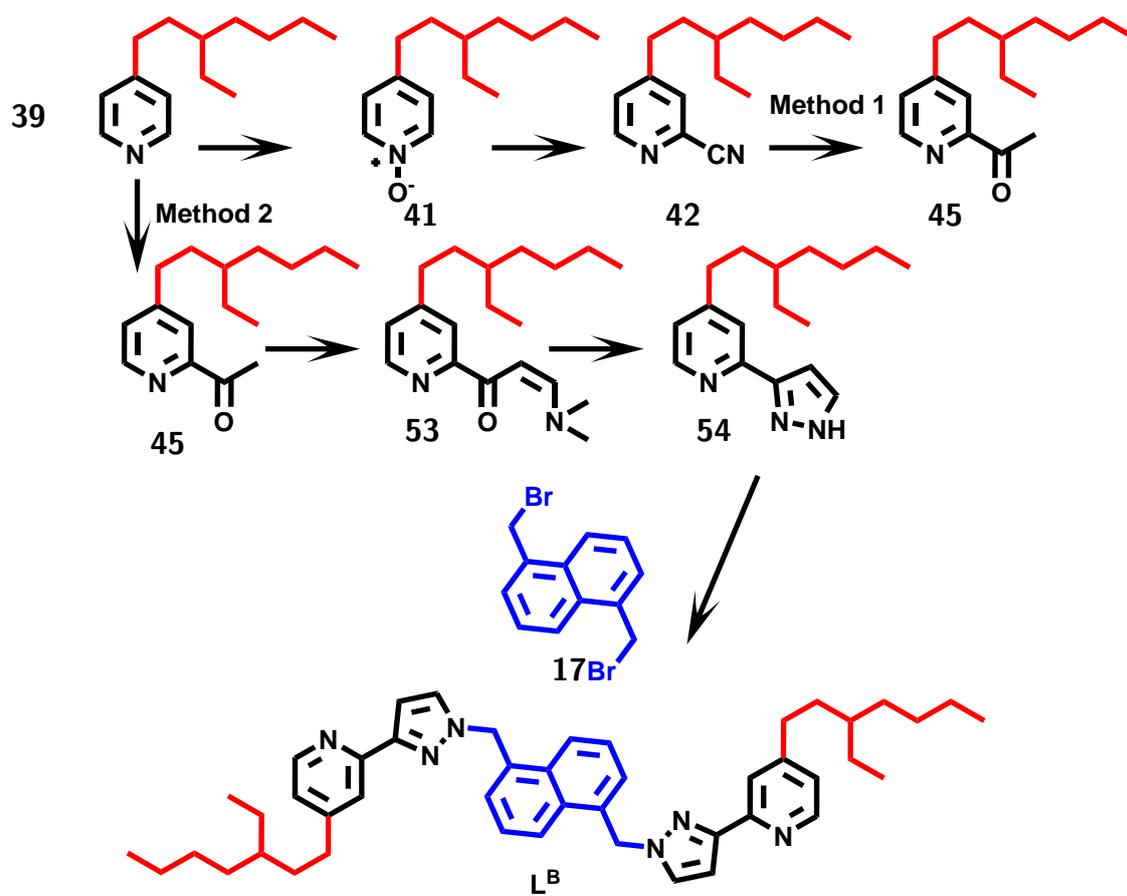
Scheme 2.30

2.6.8.1   (4-[3-ethyl-heptyl]pyridine)-N-oxide 41

39   **41**
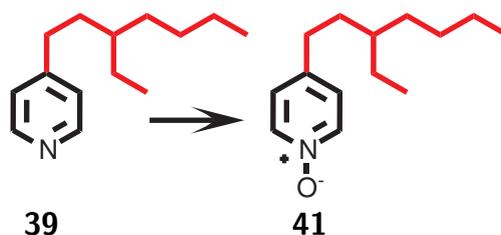
Pyridine **39** (6 g, 29.2 mmol) was dissolved in DCM (70 mL). mCPBA was added slowly while stirring. The solution was stirred for a further 2.5 h. Solution was quenched with *aq.* NaOH (70 mL, 1 mol dm^{-3}). Water (50 mL) was added and then the product was extracted with DCM (4 x 50 mL). The organic layers were collected, dried (MgSO$_4$), and reduced *in vacuo* to yield product as a clear oil.

Yield = 5.06 g, 22.9 mmol, 78 %

^{13}C-NMR (101 MHz**, CDCl$_3$):** δ (ppm) = 142.87 (C), 138.77 (CH), 125.88 (CH), 38.30 (CH), 33.81 (CH$_2$), 32.51 (CH$_2$), 31.66 (CH$_2$), 28.79 (CH$_2$), 25.56 (CH$_2$), 23.05 (CH$_2$), 14.12 (CH$_3$), 10.71 (CH$_3$)

ES-MS = m/z 222 [MH]$^+$

Accurate Mass = calc. (C$_{14}$H$_{24}$NO)[MH]$^+$ 222.1858, acqu. 222.1856

2.6.8.2 (4-[3-ethyl-heptyl]pyridine)-2-nitrile 42

41   **42**
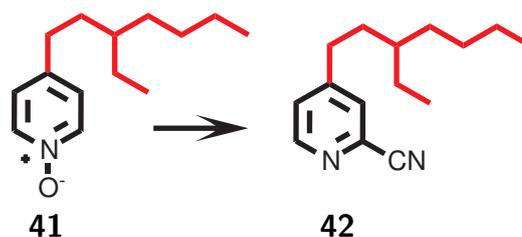
Pyridine N-oxide **41** (4.93 g, 22.3 mmol) was dissolved in DCM (70 mL) and TMSCN (3.7 mL, 27.9 mmol) was added dropwise. Di-methyl carbamyl was then added dropwise and the resulting mixture was stirred at room temperature overnight then quenched with aq. sat. potassium carbonate (150 mL). The crude product was extracted with DCM (3 x 50 mL). The organics were collected and solvent removed under vacuum. Water (100 mL) was added to the crude and stirred for 1 h. The crude was again extracted with DCM (3 x 100 mL), dried ($MgSO_4$) and solvent removed under vacuum to give crude product. Chromatography (10 % EtOAc/Petroleum) gave a yellow oil.

Yield = 2.46 g, 10.7 mmol, 48 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.60 (d, J = 5.0 Hz, 1H), 7.55 (s, 1H), 7.35 (d, J = 5.0 Hz, 1H), 2.71 to 2.60 (m, 2H), 1.66 to 1.52 (m, 2H), 1.45 to 1.16 (m, 9H), 0.96 to 0.81 (m, 6H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 153.92 (C), 150.88 (CH), 128.71 (CH), 127.04 (CH), 117.47 (C), 38.45 (CH), 33.78 (CH_2), 32.51 (CH_2), 32.31 (CH_2), 28.79 (CH_2), 25.57 (CH_2), 23.07 (CH_2), 14.14 (CH_3), 10.71 (CH_3)

ES-MS = m/z 231 [MH]$^+$

Accurate Mass = calc. $(C_{15}H_{23}N_2)$[MH]$^+$ 231.1861, acqu 231.1869

Elemental = calc. $[(C_{15}H_{22}N_2) + 0 \cdot 1\,H_2O + 0 \cdot 1\,EtOAc]$ C 76.76, H 9.62, N 11.63, acqu. C 76.93, H 9.78, N 11.49

2.6.8.3 (4-[2-ethylheptyl]pyridin-2-yl)ethan-2-one 45
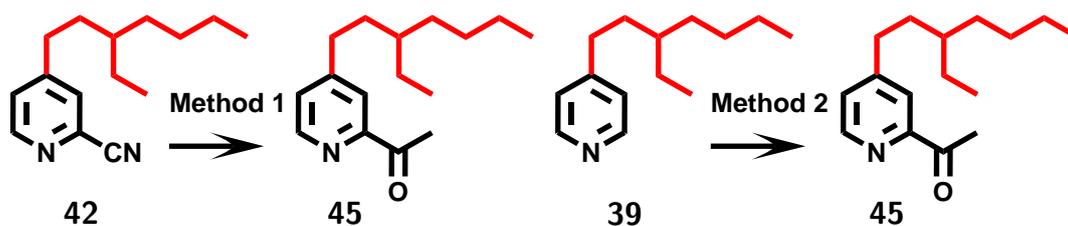
Method 1 Pyridine **42** (0.5 g, 2.17 mmol) was dissolved in dry THF (20 mL) at $-78\,^\circ$C. Methylmagnesium bromide (0.9 mL, 2.60 mmol) was added dropwise and stirred for 1 h. A *sat. aq.* ammonium chloride (10 mL each) was used to quench the reaction and the product was extracted with DCM (3 x 20 mL). The combined organic layers were dried (MgSO$_4$) and solvent was removed *in vacuo* to give a dark orange oil. Chromatography (10 % EtOAc/Pet-Ether) yields a pale yellow oil.

Yield = 80 mg, 0.32 mmol, 15 %

Method 2 At 0 $^\circ$C under N$_2$ DMAE (12.5 mL, 125 mmol) was added to hexane (250 mL). BuLi (100 mL, 250 mmol) was added dropwise over 30 min. Pyridine **42** (17.12 g, 83.4 mmol) in hexane (30 mL) was added dropwise to the reaction followed by the electrophile DMA (30 mL, 300 mmol) in hexane (30 mL). After a further 30 min stirring the reaction was quenched with water (200 mL) and product was extracted with DCM (5 x 100 mL). All organic layers were collected, dried (MgSO$_4$) and solvent removed *in vacuo* to give crude product. Chromatography (alumina 10 % to 20 % EtOAc/petroleum ether) yields product and (30 % to 40 % EtOAc/petroleum ether) yields unreacted starting material.

Yield = 9.46 g, 38.3 mmol, 46 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.58 (d, J = 5.0 Hz, 1H), 7.91 (s, 1H), 7.33 (d, J = 4.9 Hz, 1H), 2.76 (s, 3H), 2.71 to 2.63 (m, 2H), 1.65 to 1.56 (m, 2H), 1.41 to 1.21 (m, 9H), 0.98 to 0.84 (m, 6H)

ES-MS = m/z 248 [MH]$^+$

Accurate Mass = calc. (C$_{16}$H$_{26}$NO)[MH]$^+$ 248.2014, acqu. 248.2024

2.6.8.4   1-(4-[2-ethylheptyl]pyridin-2-yl)-(dimethylamino)prop-2-en-1-one 53
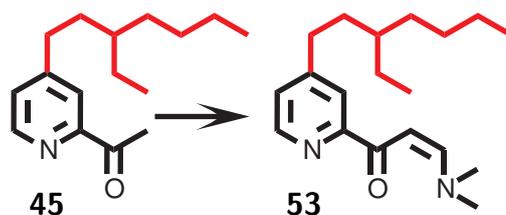
Pyridine **45** (0.77 g, 3.11 mmol) was dissolved in excess DMF-DMA (15 mL) and refluxed at 110 °C overnight. Solvent was then removed to give the crude product (1.20 g). This was used in the next step without purification.

¹H-NMR (400 MHz, **CDCl₃**): δ (ppm) = 8.52 (d, $J = 4.9$ Hz, 1H), 8.01 (s, 1H), 7.93 (d, $J = 12.7$ Hz, 1H), 7.21 (dt, $J = 4.9$ Hz $J = 2.6$ Hz, 1H), 6.46 (d, $J = 12.7$ Hz, 1H), 3.20 (s, 3H), 3.01 (s, 3H), 2.73 to 2.55 (m, 2H), 1.64 to 1.55 (m, 2H), 1.43 to 1.20 (m, 9H), 0.90 (dt, $J = 14.4$ Hz $J = 6.2$ Hz, 6H)

2.6.8.5   3-(4-[2-ethylheptyl]pyridin-2-yl)-pyrazole 54
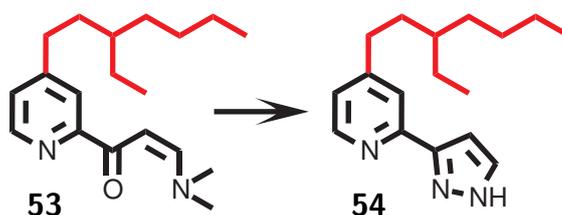
Keto enamine **53** (1.20 g crude) was dissolved in ethanol (50 mL). Hydrazine monohydrate (4 mL, 80 mmol) was added and the solution was refluxed at 65 °C for 30 min. The dark brown solution turned to a golden colour. Water (50 mL) was added and the product was extracted with DCM (3 x 50 mL). The organics where combined, dried with MgSO₄, and solvent removed under vacuum. Chromatography (2 % MeOH/DCM, Alumina Brockman III) yielded product.

Yield = 0.57 g, 2.08 mmol, 67 %

¹H-NMR (400 MHz, **CDCl₃**): δ (ppm) = 8.53 (d, $J = 5.1$ Hz, 1H), 7.68 (d, $J = 2.0$ Hz, 1H), 7.60 (s, 1H), 7.11 (dd, $J = 5.1$ Hz $J = 1.5$ Hz, 1H), 6.85 (d, $J = 1.9$ Hz, 1H), 2.73 to 2.58 (m, 2H), 1.73 to 1.54 (m, 3H), 1.43 to 1.21 (m, 9H), 0.96 to 0.84 (m, 6H)

ES-MS = m/z 272.0 [MH]$^+$

2.6.8.6 1,5 bis-(4-[2-ethylheptyl]pyridin-2-yl-pyrazol-1-yl) methylnaphthalene
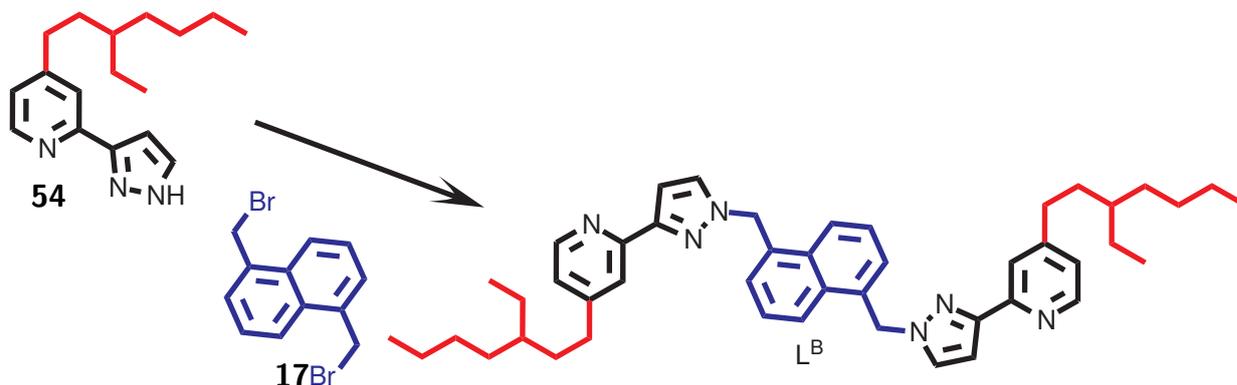
Pyridine pyrazole **54** (565 mg, 2.06 mmol) was dissolved in THF (20 mL). The spacer **17** (308 mg, 0.98 mmol) and NaOH *aq.* (5 mL 5.5 M) was added and the solution was refluxed for 24 h. After cooling the aqueous layer was removed and product extracted with DCM (3 x 15 mL. The organic layers were recombined, dried (MgSO$_4$) and solvent removed *in vacuo* to give product.

Yield = 365 mg, 0.52 mmol, 54 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.53 (d, J = 5.0 Hz, 2H), 8.06 (d, J = 8.5 Hz, 2H), 7.84 (s, 2H), 7.53 to 7.47 (m, 2H), 7.33 (d, J = 7.0 Hz, 2H), 7.26 (d, J = 2.3 Hz, 2H), 7.06 (dd, J = 5.0 Hz J = 1.5 Hz, 2H), 6.88 (d, J = 2.3 Hz, 2H), 5.89 (s, 4H), 2.97 to 2.59 (m, 4H), 1.68–1.57 (m, 4H), 1.45 to 1.16 (m, 18H), 1.00 to 0.80 (m, 12H)

^{13}C-NMR (101 MHz**, CDCl$_3$):** δ (ppm) = 152.91, 152.06, 151.99, 149.33, 132.41, 131.72, 130.83, 127.28, 126.47, 124.32, 122.74, 120.11, 104.95, 54.68, 38.67, 34.15, 32.73, 32.64, 28.88, 25.70, 23.12, 14.17, 10.80

ES-MS = m/z (%) 717.5 (6) [MNa]$^+$, 717.5 (6) [MH]$^+$, 717.5 (6) [MH$_2$]$^{2+}$

Accurate Mass = calc. (C$_{46}$H$_{59}$N$_6$)[MH]$^+$ 695.4801, acqu. 695.4803

2.6.9 Synthesis of cyclo-hexyl ligand (L^C)
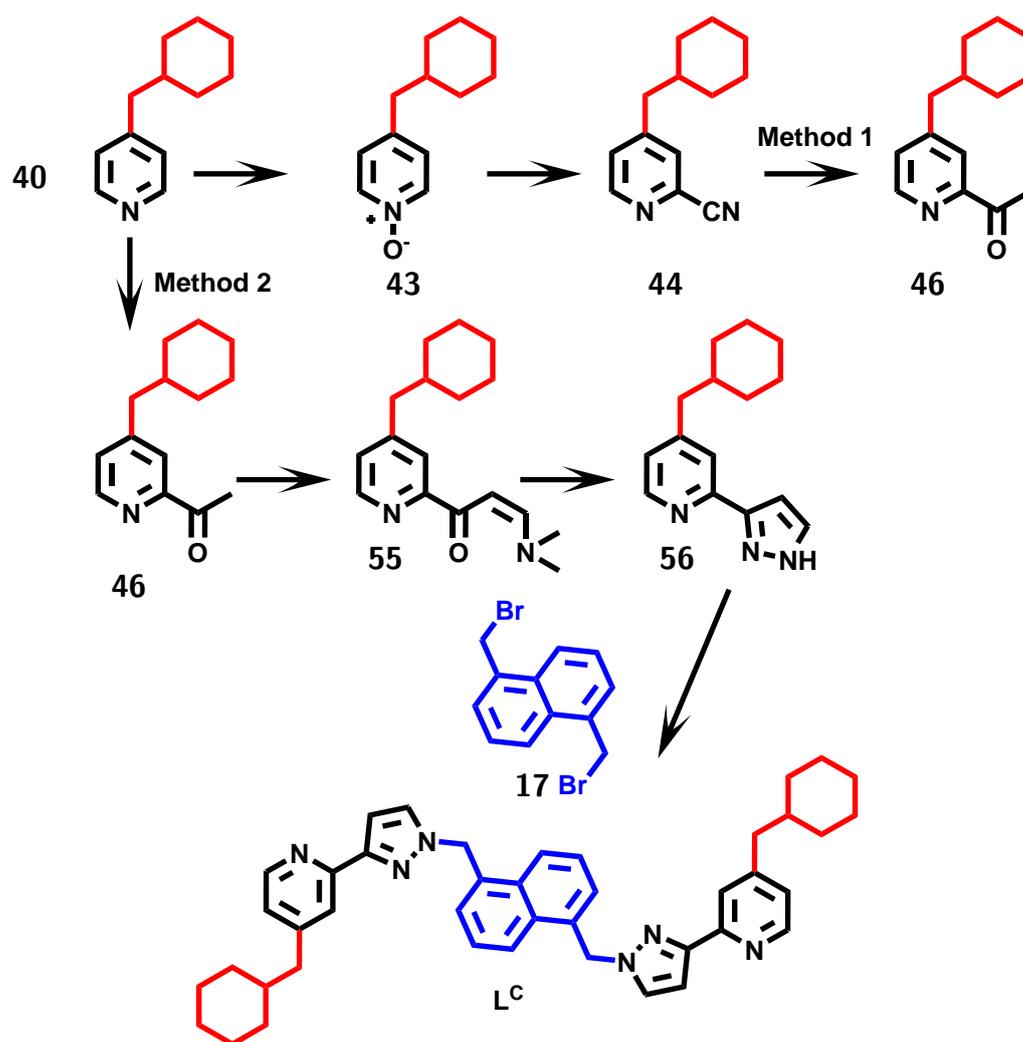
Scheme 2.31

2.6.9.1 (4-cyclohexylmethylpyridine) N-oxide 43

40   **43**
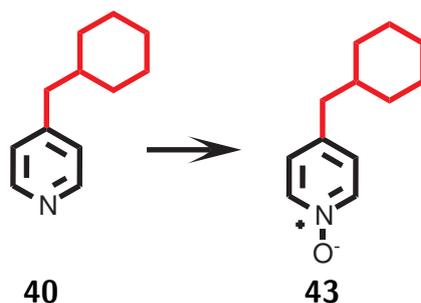
Pyridine **40** (4 g, 22.8 mmol) was dissolved in DCM (70 mL). mCPBA was added slowly while stirring. The solution was stirred for a further 2.5 h. Solution was quenched with *aq.* NaOH (70 mL, 1 M). Water (50 mL) was added and then the product was extracted with DCM (4 x 50 mL). The organic layers were collected, dried (MgSO$_4$), and reduced *in vacuo* to yield product as a white solid.

Yield = 3.11 g, 16.3 mmol, 17 %

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.11 (d, J = 7.0 Hz, 2H), 7.04 (d, J = 7.0 Hz, 2H), 2.46 (d, J = 7.1 Hz, 2H), 1.75 to 1.58 (m, 5H), 1.58 to 1.41 (m, 1H), 1.28 to 1.09 (m, 3H), 1.00 to 0.85 (m, 2H)

^{13}C-NMR (101 MHz**, CDCl$_3$):** δ (ppm) = 141.04 (C), 138.58 (CH), 126.64 (CH), 42.44 (CH2), 39.04 (CH), 32.84 (CH$_2$), 26.23 (CH$_2$), 26.05 (CH$_2$)

ES-MS = m/z 192 [MH]$^+$

Accurate Mass = calc. (C$_{12}$H$_{18}$NO)[MH]$^+$ 192.1388, acqu. 192.1396

Elemental = calc. [(C$_{12}$H$_{17}$NO) + 0 $\cdot$ 4 MeOH] C 72.98, H 9.19, N 6.86; acqu. C 72.83, H 9.15, N 7.03

2.6.9.2 (4-cyclohexylmethylpyridine)-2-nitrile 44
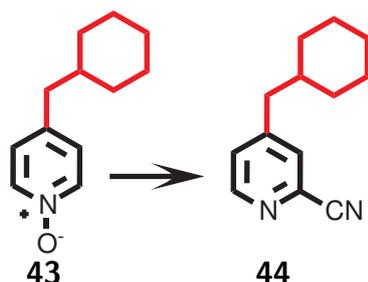
Pyridine N-oxide **43** (4.93 g, 22.3 mmol) was dissolved in DCM (70 mL) and TMSCN (3.7 mL, 27.9 mmol) was added dropwise. Di-methyl carbamyl was then added dropwise and the resulting mixture was stirred at room temperature overnight then quenched with aq. sat. potassium carbonate (150 mL). The crude product was extracted with DCM (3 x 50 mL). The organics were collected and solvent removed under vacuum. Water (100 mL) was added to the crude and stirred for 1 h. The crude was again extracted with DCM (3 x 100 mL), dried ($MgSO_4$) and solvent removed under vacuum to give crude product. Chromatography (10 % EtOAc/petroleum ether) gave a white solid.

Yield = 2.50 g, 11.5 mmol, 73 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.60 (dd, J = 5.0 Hz, J = 0.5 Hz, 1H), 7.51 (d, J = 0.8 Hz, 1H), 7.31 (dd, J = 5.0 Hz, J = 1.7 Hz, 1H), 2.56 (d, J = 7.0 Hz, 2H), 1.77 to 1.50 (m, 6H), 1.32 to 1.09 (m, 3H), 1.04 to 0.91 (m, 2H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 152.04 (C), 150.73 (CH), 129.38 (CH), 127.80 (CH), 117.50 (C), 43.04 (CH$_2$), 38.95 (CH), 32.89 (CH$_2$), 26.18 (CH$_2$), 26.03 (CH$_2$)

ES-MS = m/z 201 [MH]$^+$

Accurate MAss = calc. (C$_{13}$H$_{17}$N$_2$)[MH]$^+$

2.6.9.3 (4-cyclohexylmethylpyridin-2-yl)ethan-2-one 46
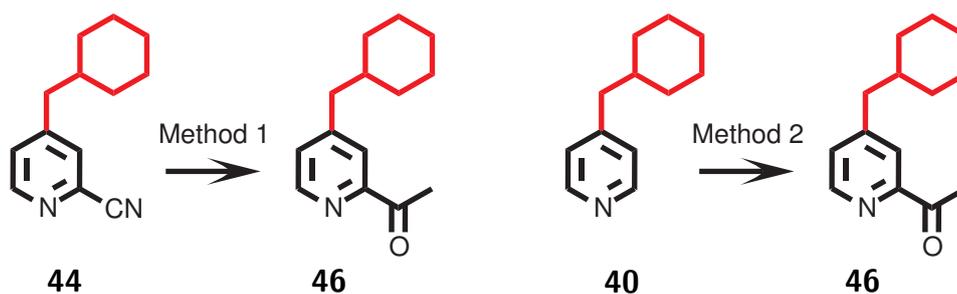
Method 1 Pyridine **44** (2.84 g, 14.18 mmol) was dissolved in dry diethyl ether (50 mL) at 0 °C. Methyl magnesium bromide (5.67 mL, 17.02 mmol) was added dropwise and stirred for 3 hrs. A *sat. aq.* ammonium chloride (50 mL each) was added to quench the reaction. The product was extracted with DCM (3 x 50 mL). The combined organic layers were dried ($MgSO_4$) and solvent removed *in vacuo* to give a yellow oil. Chromatography (10 % EtOAc/petroleum ether) yields a pale straw-coloured oil.

Yield = 389 mg, 1.79 mmol, 13 %

Method 2 At 0 °C under N_2 DMAE (12.5 mL, 125 mmol) is added to hexane (200 mL). BuLi (100 mL, 250 mmol) is added dropwise over 30 min. Pyridine **44** (14.0 g, 78.9 mmol) in hexane (15 mL) is added dropwise followed by the electrophile DMA (30 mL, 300 mmol). After 30 min stirring the reaction is quenched with water (100 mL). Product is extracted with DCM (3 x 100 mL). All organic layers are collect, dried ($MgSO_4$) and solvent removed *in vacuo* to give crude. Chromatography (alumina 10 % to 20 % EtOAc/petroleum ether) yields product.

Yield = 5.36 g, 30.5 mmol, 38 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.61 to 8.54 (m, 1H), 7.88 (s, 1H), 7.33 to 7.29 (m, 1H), 2.78 (s, J = 6.2 Hz, 3H), 2.61 to 2.50 (m, 2H), 1.77 to 1.53 (m, 6H), 1.29 to 1.08 (m, 3H), 1.05 0.87 (m, 2H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 200.30 (C), 153.14 (C), 151.92 (C), 148.51 (CH), 128.09 (CH), 122.56 (CH), 43.43 (CH$_2$), 39.05 (CH), 32.99 (CH$_2$), 26.31 (CH$_2$), 26.12 (2x CH$_2$), 26.05 (CH$_2$)

ES-MS = m/z 218 [MH]$^+$

Accurate Mass = calc. (C$_{14}$H$_{20}$NO)[MH]$^+$ 218.1545, acqu. 218.1552

2.6.9.4 1-(4-cyclohexylmethylpyridin-2-yl)-(dimethylamino)

prop-2-en-1-one 55
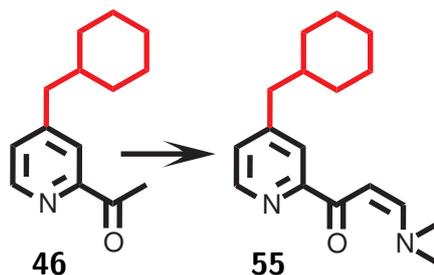
Pyridine **46** (0.76 g, 3.50 mmol) was dissolved in excess DMF-DMA (15 mL) and refluxed at $110\,°C$ overnight. Solvent was then removed to give the crude product (0.932 g). This was used in the next step without purification.

¹H-NMR (400 MHz, **CDCl₃**): δ (ppm) = 8.52 (d, $J = 4.9$ Hz, 1H), 8.33 (s, 1H), 7.93 (d, $J = 12.6$ Hz, 1H), 7.26 (s, $J = 2.5$ Hz, 1H), 6.48 (d, $J = 12.0$ Hz, 1H), 3.36 (s, 2H), 2.98 (s, 3H), 2.90 (d, $J = 0.5$ Hz, 3H), 1.48 to 1.43 (m, 6H), 1.32 (m, 3H), 1.17 (m, 2H)

ES-MS = m/z 273 [MH]$^+$

2.6.9.5   3-(4-cyclohexylmethylpyridin-2-yl)-pyrazole 56
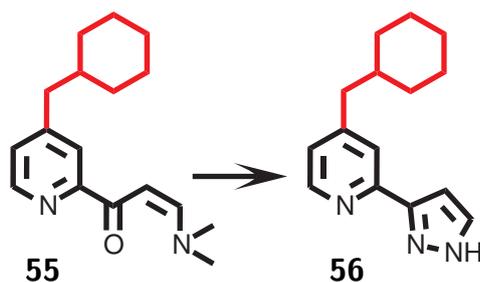
Keto enamine **55** (4.68 g crude) was dissolved in ethanol (125 mL). Hydrazine monohydrate (4 mL, 80 mmol) was added and the solution was refluxed at 65 °C for 30 min. The dark brown solution turned to a golden colour. Water (50 mL) was added and the product was extracted with DCM (3 x 50 mL). The organics where combined, dried with MgSO$_4$, and solvent removed under vacuum. Chromatography (5 % MeOH/DCM) yielded product.

Yield = 0.18 g, 0.78 mmol, 22 % over two steps

^{1}H-NMR (400 MHz**, CDCl$_3$):** δ (ppm) = 8.50 (dd, $J = 5.0$ Hz, $J = 0.4$ Hz, 1H), 7.67 (d, $J = 2.0$ Hz, 1H), 7.54 (s, 1H), 7.04 (dt, $J = 6.9$ Hz, $J = 3.4$ Hz, 1H), 6.81 (s, 1H), 2.53 (t, $J = 9.1$ Hz, 2H), 1.79 to 1.55 (m, 6H), 1.35 to 1.10 (m, 3H), 1.07 to 0.82 (m, 2H)

ES-MS = m/z 242 [MH]$^+$

2.6.9.6 1,5 bis-(4-cyclohexylmethylpyridin-2-yl-pyrazol-1-yl)

methylnaphthalene
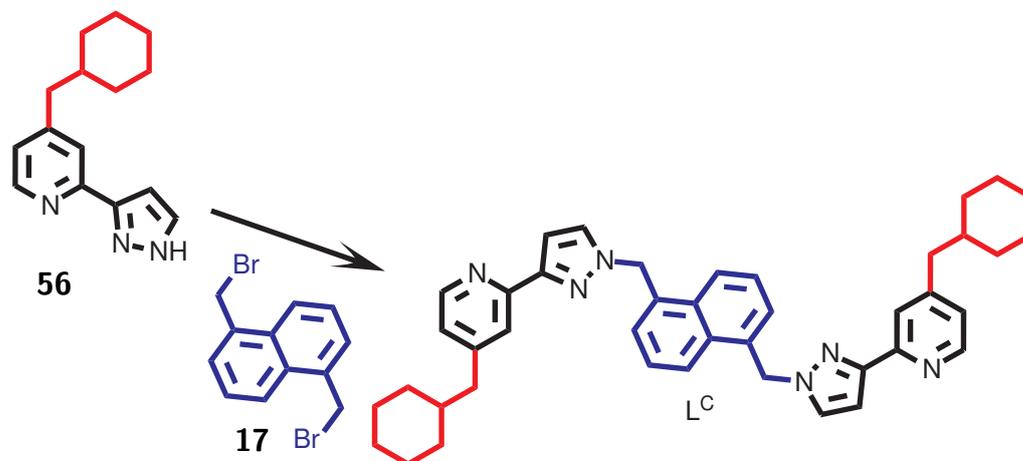
Pyridine pyrazole **56** (8.12 g, 33.7 mmol) was dissolved in THF (200 mL). The spacer **17** (5.04 g, 16.0 mmol) and NaOH *aq.* (50 mL 5.5 M) was added and the solution was refluxed for 24 h. After cooling the aqueous layer was removed and product extracted with DCM (3 x 200 mL. The organic layers were recombined, dried (MgSO$_4$) and solvent removed *in vacuo* to give product.

Yield = 5.27 g, 8.30 mmol, 52 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.53 (dd, J = 5.0 Hz, J = 0.6 Hz, 1H), 8.06 (d, J = 8.6 Hz, 1H), 7.80 (d, J = 0.8 Hz, 1H), 7.56 to 7.47 (m, 1H), 7.34 (d, J = 7.0 Hz, 1H), 7.26 (d, J = 2.4 Hz, 1H), 7.02 (dd, J = 5.0 Hz, J = 1.6 Hz, 1H), 6.88 (d, J = 2.4 Hz, 1H), 5.90 (s, 2H), 2.55 (d, J = 7.0 Hz, 2H), 1.68 (dd, J = 28.8 Hz, J = 12.8 Hz, 6H), 1.33 to 1.09 (m, 3H), 1.08 to 0.81 (m, 2H).

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 152.01, 151.94, 151.01, 149.15, 132.41, 131.74, 130.82, 127.32, 126.49, 124.35, 123.55, 120.84, 104.96, 54.69, 43.54, 39.09, 33.12, 26.42, 26.21.

ES-MS = (%) m/z 635.4 (13) [MH]$^+$, 318.2 (100) [MH$_2$]$^{2+}$.

Accurate Mass = calc. (C$_{42}$H$_{47}$N$_6$)[MH]$^+$ 635.3862, acqu. 635.3873

2.6.10   Synthesis of terpyridine-based complex
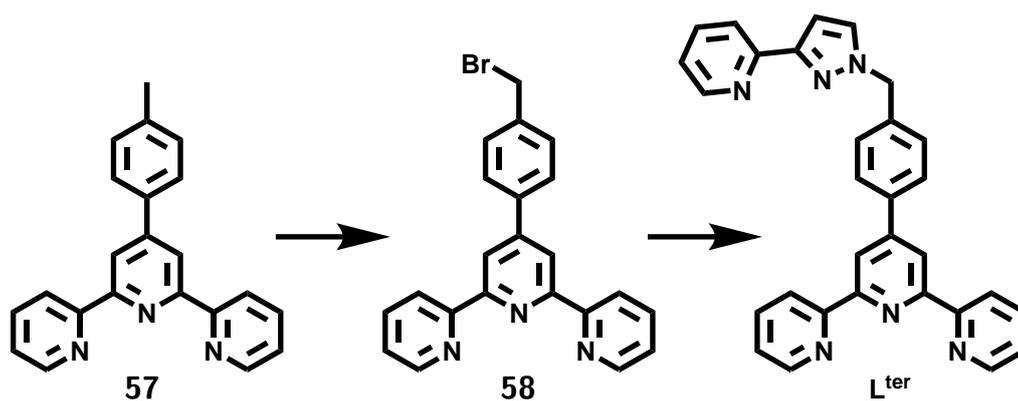
Scheme 2.32

2.6.10.1 4-[4-(bromomethyl)phenyl]-6-(pyridin-2-yl)-2,2'-bipyridine 58
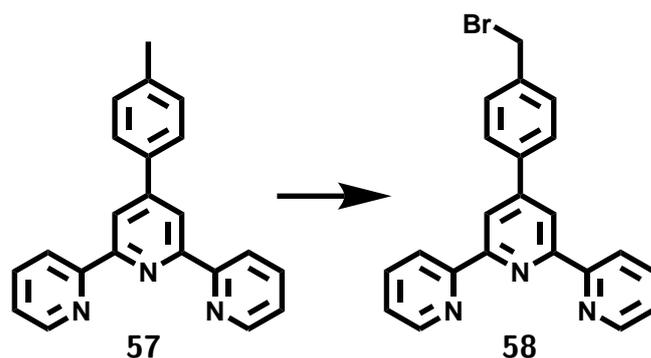
Phenyl ter-pyridine **57** (1.0 g, 3.09 mmol) and NBS (0.6 g, 3.37 mmol) were stirred in DCM (60 mL). A spatular of AIBN was added and the solution refluxed under a tungsten lamp (15 min on, 15 min off). The reaction was monitored by TLC and after 2.5 h was removed from heat and allowed to cool. Water (100 mL) was added and product extracted with DCM (3 x 50 mL). The solvent was dried ($MgSO_4$) and removed *in vacuo* to give product.

Yield = 900 mg, 2.24 mmol, 73 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.77 (m, 2H), 8.76 (m, 2H), 8.71 (d, J = 7.9 Hz, 2H), 7.93 (m, 4H), 7.58 (d, J = 8.3 Hz, 2H), 7.40 (m, 2H), 4.60 (s, 2H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ (ppm) = 156.17 (C), 156.05 (C), 149.54 (C), 149.17 (CH), 138.65 (C), 136.94 (CH), 129.67 (CH), 127.80 (CH), 127.18 (C), 123.92 (CH), 121.41 (CH), 118.91 (CH), 32.99 (CH$_2$)

2.6.10.2 6-(pyridin-2-yl)-4-(4-[3-(pyridin-2-yl)pyrazol-1-yl]methylphenyl)-2,2'-bipyridine L^{ter}
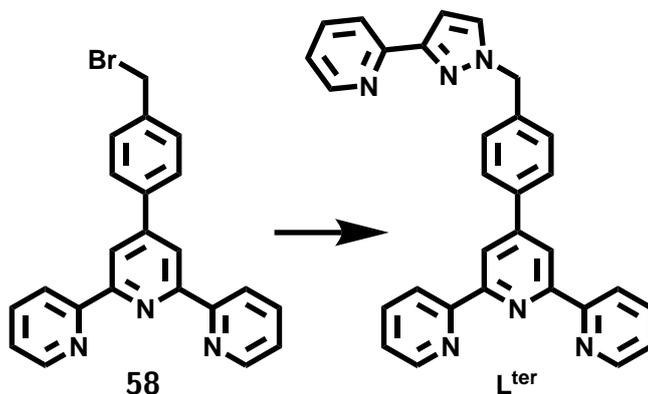
Brominated phenyl terpyridine **58** (900 mg, 2.24 mmol) and pyridine pyrazole **16** (325 mg, 2.24 mmol) were dissolved in THF (20 mL). NaOH *aq.*(1 M, 5 mL) was added and the solution refluxed for 6 days. After cooling the aqueous layer was removed. Water (50 mL) was added and the product extracted with DCM (3 x 30 mL). The solvent was dried (MgSO$_4$) and removed *in vacuo* to give product.

Yield = 320 mg, 0.69 mmol, 31 %

^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 8.76 to 8.72 (m, 4H), 8.70 to 8.65 (m, 3H), 7.99 (dt, J = 1.0 Hz J = 8.1 Hz, 1H), 7.93 to 7.87 (m, 4H), 7.77 (dt, J = 1.9 Hz J = 7.8 Hz, 1H), 7.49 (d, J = 2.3 Hz, 1H), 7.43 (m, 2H), 7.38 (ddd, J = 1.2 Hz J = 4.8 Hz J = 7.5 Hz, 2H), 7.23 (m, 1H), 6.97 (d, J = 2.3 Hz, 1H), 5.51 (s, 2H)

MS-ESI = m/z (%) 467.1 (100) [MH]$^+$ 489.2 (3) [MNa]$^+$

2.6.10.3   Co(BF$_4$)$_2$ complex with Ligand L^{ter} complex

Ligand **L**^{ter} was mixed with Co(BF$_4$)$_2$ in methanol and refluxed overnight. The solvent was removed *in vacuo* to give product.

* protons highlight the paramagneticly shifted peaks H_a through to H_e. See Fig. 2.10

^{1}H-NMR (400 MHz, **CD$_3$CN**): δ (ppm) = 93.16 (m, 2H*), 54.54 (m, 2H*), 43.37 (m, 2H*), 32.65 (m, 2H*), 13.68 (m, 2H*), 9.57 to 8.54 (m, 6H), 8.53 to 7.97 (m, 1H), 7.96 to 7.22 (m, 2H), 6.07 (s, 1H), 3.31 (s, 2H)

MS-ES = m/z (%) 1079 (22) [CoL$_2^{ter}$]$^{2+}$, 1079 (22) $\left[[CoL^{ter}2](BF4)\right]^{1+}$

2.6.10.4   FeCl$_2$ complex with Ligand L^{ter} complex

Ligand **L**^{ter} was mixed with FeCl$_2$ in methanol and refluxed overnight. The solvent was removed *in vacuo* to give product.
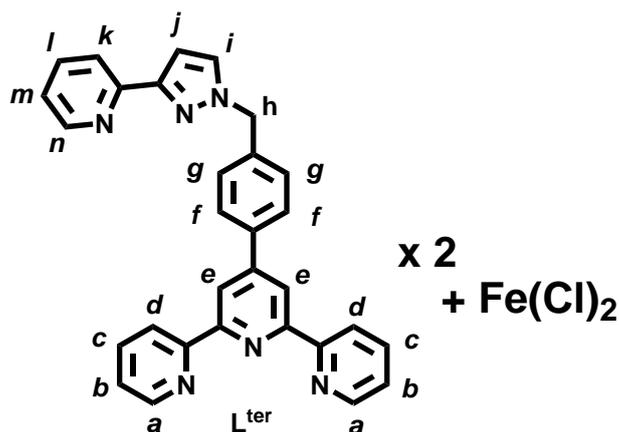
assignment is only an estimation based on literature[92,94] and the integration of peaks.

^{1}H-NMR (400 MHz, **CD$_3$CN**): δ (ppm) = 9.22 (s, H_e, 2H), 8.73 (m, H_j, 1H), 8.64 (m, H_d, 2H), 8.32 (m, H_f, 2H), 7.99 (m, H_l, 1H), 7.90 (m, H_c, 2H), 7.82 (m, H_g, 2H), 7.72 (m, H_k H_n, 2H), 7.32 (m, H_m, 1H), 7.18 (m, H_b, 2H), 7.09 (m, H_a, 2H), 7.00 (m, H_i, 1H), 5.63 (s, H_h, 2H)

MS-ES = m/z 494 [FeL$_2^{ter}$]$^{2+}$

2.6.11 Cage Complex Synthesis

General Procedure Ligand and metal-salt were placed into a teflon lined autoclave reaction vessel with methanol solvent. The sealed reaction pot was heated to 120 °C for 12 h followed by a slow gradual cooling of 0.1 ° min^{-1} to help crystallisation of product. The resulting solids were collected as product and washed with the reaction methanol, ice-cold fresh methanol, DCM, chloroform and finally diethyl ether before being air dried.

It is also, on a larger scale, to produce the cage though a simple reflux overnight. This method does not produce crystalline material and so the product is not as pure.

The synthesis with $Co(BPh_4)_2$ involved a different route. The structure is first formed using $Co(CH_3CO_2)_2$. The ligand and metal salt was dissolved into methanol and refluxed for 1 h. NaBPh$_4$ is added as a saturated aqueous solution. The resulting precipitate was then collected as product and washed with water, methanol and diethyl ether.

2.6.11.1 Ligand: L$^{\mathbf{A}}$, Metal salt: Co(BF$_4$)$_2$

^{1}H$-$NMR$_{\mathbf{paramagnetic}}$ (400 MHz, CD$_3$CN): δ (ppm) = 114.44, 93.65, 91.81, 90.61, 82.77, 79.97, 76.98, 69.62, 66.12, 62.46, 61.84, 54.06, 53.51, 52.95, 50.31, 49.61, 48.78, 47.04, 42.54, 37.53, 34.39, 28.50, 22.91, 20.52, 18.13, 14.52, 14.11, 12.36, 11.76, 10.60, 9.71, 3.43, 1.30, 0.87, -3.30, -8.59, -11.89, -16.32, -16.71, -25.69, -41.82, -48.13, -66.15, -91.33

ES-MS = m/z 2302 $\left[[M_8L_{12}]_\bullet(BF_4)_{13}\right]^{3+}$, 1706 $\left[[M_8L_{12}]_\bullet(BF_4)_{12}\right]^{4+}$, 1347 $\left[[M_8L_{12}]_\bullet(BF_4)_{11}\right]^{5+}$, 1108 $\left[[M_8L_{12}]_\bullet(BF_4)_{10}\right]^{6+}$, 938 $\left[[M_8L_{12}]_\bullet(BF_4)_9\right]^{7+}$, 810 $\left[[M_8L_{12}]_\bullet(BF_4)_8\right]^{8+}$

2.6.11.2 Ligand: L$^{\mathbf{W}}$, Metal salt: Co(BF$_4$)$_2$

^{1}H$-$NMR$_{\mathbf{paramagnetic}}$ (400 MHz, CD$_3$CN): δ (ppm) = 122.46, 92.32, 91.12, 87.96, 81.40, 79.82, 71.36, 67.86, 60.54, 57.78, 54.88, 53.44, 52.76, 50.94, 49.56, 47.58, 47.22, 40.61, 39.94, 33.91, 28.23, 17.35, 13.14, 12.42, 10.58, 5.09, 3.76, 3.35, 2.42, 2.12, 0.26, 0.02, -0.37, -0.60, -3.03, -3.68, -3.89, -10.63, -12.72, -18.68, -19.79, -26.63, -40.53, -54.33, -69.63, -93.78

2.6.11.3 Ligand: L_{mod}^{W}, Metal salt: $Co(BF_4)_2$

$^{1}H-NMR_{paramagnetic}$ (400 MHz, CD_3CN): δ (ppm) = 113.58, 86.98, 83.67, 82.98, 76.12, 75.43, 60.83, 56.54, 54.42, 49.94, 48.26, 446.14, 44.07, 39.54, 35.79, 33.13, 25.53, 17.69, 15.08, 12.62, 11.14, 10.79, 9.26, 8.38, 7.49, 7.05, 6.26, 5.47, 4.58, 4.08, 3.79, 3.35, 2.12, 1.96, 1.66, 1.37, 0.97, -1.43, -2.24 -2.84, -7.30, -8.21, -11.12, -13.35, -14.49, -15.98, -22.14, -31.27, -40.63, -51.82, -54.45, -59.92, -78.53

2.6.11.4 Ligand: L^{Et}, Metal salt: $Co(BF_4)_2$

$^{1}H-NMR_{paramagnetic}$ (400 MHz, CD_3CN): δ (ppm) = -92.0, -65.9, -51.3, -33.5, -26.5, -19.4, -16.4, -12.1, -9.8, -4.4, -1.9, -1.3, -0.9, -0.6, -0.4, 0.6, 0.9, 1.3, 2.0, 2.7, 2.9, 4.0, 5.0, 6.7, 10.1, 11.8, 12.6, 16.5, 18.3, 29.1, 35.2, 36.5, 41.5, 48.3, 48.9, 49.2, 49.9, 51.1, 53.9, 56.1, 58.7, 61.7, 66.4, 76, 77, 83, 89, 90, 92, 129

ES-MS = m/z 2527 $\left[[M_8L_{12}]_\bullet(BF_4)_{13}\right]^{3+}$, 1874 $\left[[M_8L_{12}]_\bullet(BF_4)_{12}\right]^{4+}$, 1482 $\left[[M_8L_{12}]_\bullet(BF_4)_{11}\right]^{5+}$, 1220 $\left[[M_8L_{12}]_\bullet(BF_4)_{10}\right]^{6+}$, 1034 $\left[[M_8L_{12}]_\bullet(BF_4)_9\right]^{7+}$, 894 $\left[[M_8L_{12}]_\bullet(BF_4)_8\right]^{8+}$, 785 $\left[[M_8L_{12}]_\bullet(BF_4)_7\right]^{9+}$, 697 $\left[[M_8L_{12}]_\bullet(BF_4)_6\right]^{10+}$, 499 $[LH]^+$

2.6.11.5 Ligand: L^{Et}, Metal salt: $Co(BPh_4)_2$

$^{1}H-NMR_{paramagnetic}$ (400 MHz, CD_2Cl_2): δ (ppm) = 81.33, 66.01, 64.02, 57.57, 50.28, 37.44, 26.90, 15.46, 15.14, 9.83, 8.38, 7.74, 7.38, 6.93, 5.35, 4.08, 2.13, 1.28, 0.91, -6.95, -7.62, -9.77, -15.97, -17.37, -17.67, -19.43, -29.09, -41.00

2.6.11.6 Ligand: L^{B}, Metal salt: $Co(BF_4)_2$

$^{1}H-NMR_{paramagnetic}$ (400 MHz, $CDCl_3$, 50 °C): δ (ppm) = -28.5, -21.4, -11.9, 9.7, 13.2, 20.3, 22.6, 32.9, 35.9, 40.9, 44.4, 46.5, 49.0, 55.87, 57,7, 74.9, 82.5, 84.5, 86.4

$^{1}H-NMR_{paramagnetic}$ (400 MHz, CD_3CN, 50 °C): δ (ppm) = -40.6, -31.5, -22.0, -16.1, -13.3, -8.2, 10.9, 12.8, 15.1, 17.7, 25.7, 33.1, 39.6, 44.1, 46.2, 48.3, 50.0, 54.6, 56.6, 69.9, 75.5, 76.1, 83.0, 83.7

$^{1}H-NMR_{paramagnetic}$ (400 MHz, CD_2Cl_2): δ (ppm) = 126.79, 90.30, 88.99, 87.45, 83.77, 81.91, 66.28, 63.81, 60.07, 53.01, 52.18, 50.92, 49.18, 46.92, 45.82, 40.28, 36.73, 35.57, 27.9, 24.38, 16.14, 15.18, 12.87, 11.40, 7.46, 7.33, 6.06, 5.35, 5.10, 4.89, 4.56, 4.46,

4.34, 4.15, 3.87, 3.47, 3.33, 2.69, 2.27, 2.18, 2.01, 1.57, 0.91, 0.38, 0.16, -1.51, -1.75, -1.82, -2.01, -3.55, -4.59, -7.39, -9.86, -14.89, -18.72, -24.23, -30.51, -47.04, -70.77

ES-MS = m/z 3312 $\left[[M_8L_{12}]_\bullet(BF_4)_{13}\right]^{3+}$, 2463 $\left[[M_8L_{12}]_\bullet(BF_4)_{12}\right]^{4+}$, 1954 $\left[[M_8L_{12}]_\bullet(BF_4)_{11}\right]^{5+}$, 1614 $\left[[M_8L_{12}]_\bullet(BF_4)_{10}\right]^{6+}$, 1371 $\left[[M_8L_{12}]_\bullet(BF_4)_9\right]^{7+}$, 1188 $\left[[M_8L_{12}]_\bullet(BF_4)_8\right]^{8+}$

2.6.11.7 Ligand: L^B, Metal salt: $Co(BPh_4)_2$

$^1H-NMR_{paramagnetic}$ (400 MHz, **CDCl$_3$**, 50 °C): δ **(ppm)** =113.19, 87.75, 72.27, 71.47, 65.76, 60.32, 53.32, 49.17, 42.92, 28.86, 26.12, 19.92, 19.55, 18.58, 17.82, 15.45, 12.71, 12.01, 11.36, 8.56, 8.24, 7.01, 4.32, 3.47, 2.96, 2.76, 2.19, 1.73, 1.33, 0.91, 0.10, -1.51, -1.74, -2.00, -2.84, -3.63, -3.79, -4.22, -4.33, -4.66, -4.89, -5.61, -6.81

2.6.11.8 Ligand: L^B, Metal salt: $Cd(NO_3)_2$

$^1H-NMR_{paramagnetic}$ (400 MHz, **CDCl$_3$**, 50 °C): δ **(ppm)** = 8.54, 8.38, 8.22, 8.09, 7.96, 7.92, 7.86, 7.78, 7.75, 7.65, 7.55, 7.51, 7.50, 7.46, 7.45, 7.36, 7.34, 7.29, 7.17, 7.15, 7.07, 7.06, 7.03, 7.00, 6.89, 6.80, 6.73, 6.68, 6.54, 6.45, 6.31, 6.19, 5.94, 5.91, 5.89, 5.75, 5.51, 5.45, 5.15, 5.11, 5.08, 5.07, 4.85, 4.77, 3.53, 3.52, 2.94, 2.92, 2.82, 2.80, 2.71, 2.09, 2.06, 1.96, 1.67, 1.44, 1.39, 1.32, 1.28, 1.22, 1.13, 1.11, 1.07, 1.06, 0.93, 0.91, 0.84, 0.82, 0.80, 0.78, 0.10

2.6.11.9 Ligand: L^C, Metal salt: $Co(BF_4)_2$

ES-MS = m/z 3072 $\left[[M_8L_{12}]_\bullet(BF_4)_{13}\right]^{3+}$, 2283 $\left[[M_8L_{12}]_\bullet(BF_4)_{12}\right]^{4+}$, 1809 $\left[[M_8L_{12}]_\bullet(BF_4)_{11}\right]^{5+}$, 1493 $\left[[M_8L_{12}]_\bullet(BF_4)_{10}\right]^{6+}$, 1267 $\left[[M_8L_{12}]_\bullet(BF_4)_9\right]^{7+}$, 1098 $\left[[M_8L_{12}]_\bullet(BF_4)_8\right]^{8+}$

2.6.11.10 Ligand: L^C, Metal salt: $Co(BPh_4)_2$

$^1H-NMR_{paramagnetic}$ (400 MHz, **CDCl$_3$**, 50 °C): δ **(ppm)** = 73.14, 71.19, 71.00, 66.81, 51.88, 48.63, 47.53, 44.98, 41.87, 39.89, 39.72, 30.50, 26.97, 19.82, 18.54, 18.29, 17.81, .17.07, 16.56, 16.31, 14.16, 12.86, 12.24, 12.10, 10.71, 10.01, 3.87, 2.12, 1.55, 1.18, 0.82, 0.00, -1.21, -1.95, -2.34, -2.57, -3.05, -3.30, -3.56, -4.24, -4.38, -6.54, -7.39, -7.87, 8-.55, -8.65, -8.90, -9.81, -9.96, -13.54, -17.21, -17.82, -17.99, -18.43, -18.69, -35.71, -41.83

2.6.12 X-ray Crystallography

The crystal structure data collection was performed using a Bruker APEX-2 CCD difractometer with Mo-$K\alpha$ radiation from a sealed tube source. Data was corrected for absorption using empirical methods (SADABS) based upon symmetry-equivalent reflections combined with measurements at different azimuthal angles. The structures were solved and refined using a combination of different software including Olex$_2$, WinGX, Apex3, PLATON and primarily the SHELX suite version 6.14.

Complex	L$^{\mathrm{M}}$e
Formula[a]	$C_{30}H_{26}N_6$
Molecular Weight	467.54
T, K	100(2)
Crystal System	Triclinic
Space Group	P-1
a, Å	8.681(3)
b, Å	8.956(3)
c, Å	16.001(4)
α, °	99.80(2)
β, °	102.135(19)
γ, °	95.91(2)
V, Å^3	1186.0(6)
Z	2
ρ, $\mathrm{g\,cm^{-1}}$	1.257
Crystal Size, mm^3	0.20 x 0.18 x 0.05
Data, restraints, parameters	5375, 0, 328
Final R1, wR2[b]	0.0726, 0.1484

a These formulae (and consequently the crystal densities) are necessarily approximate given that large amounts of diffuse electron density in solvent-accessible voids were removed from the refinements using the 'SQUEEZE' function in PLATON. See CIFs and comments below for more details.

b The value of R1 is based on 'observed' data with $I > 2\sigma(I)$; the value of wR2 is based on all data.

Chapter 3

Binding of Chemical Warfare Agent Simulants
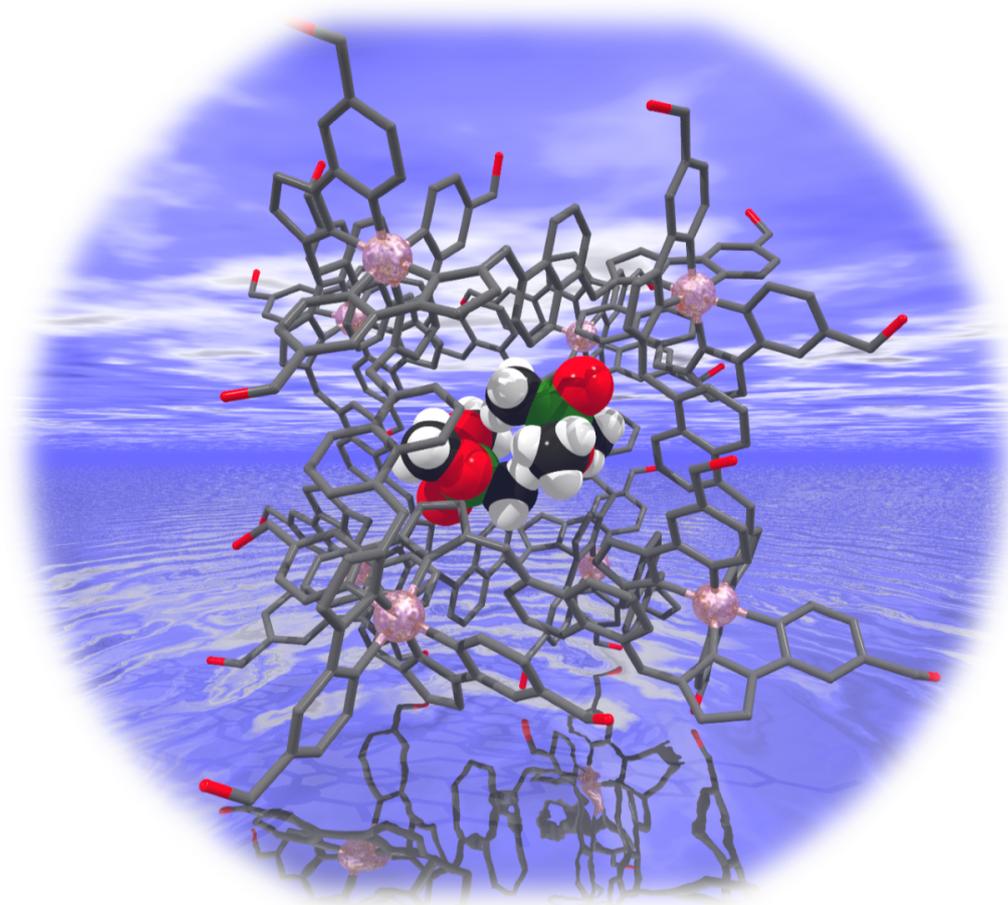
Render of $[\text{Co}_8(\text{L}^\text{W})_{12}](\text{BF}_4)_{16}$ with two bound DMMP guests

3.1 Introduction

3.1.1 Brief History of Chemical Warfare Agents

Chemical warfare agents (CWAs) were first used extensively during the First World War in the form of choking gases such as chlorine or phosgene. The development lead to the invention and usage of mustard gas (HD, or bis(2-chloroethyl)sulphide), named so due to the colour and smell of the impure form used in weapons.

During the Second World War and immediately after a new class of CWAs were develop based on a series of organophosphorus compounds. This G-series of nerve agents were further developed to give the related but different V-series of nerve agents. Unlike their predecessors these agents are colourless, odourless, tasteless, fast acting and deadly at low concentrations. Fortunately, these agents were never deployed during the Second World War. The use of these agents, and other chemical and biological weapons, were banned from modern warfare in the Geneva Protocol.
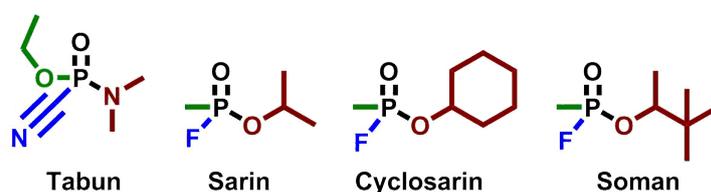
Fig. 3.1 The structures of G-series nerve agents (one family of CWA's

The first recorded use of nerve agents is believed to be during the Iraq civil war, in 1988, against a single Kurdish village. The most widely publicised usage of one of these agents, Sarin, was in the 1995 terrorist attack on the Tokyo subway system where five coordinated attacks killed twelve people while injuring around five thousand others. With the quantity of sarin involved the death count could easily have been far higher[95].

Since the growing threat of terrorist attacks there has been a growing fear of the further use of these horrific agents. More recently the use of Sarin was confirmed during the Syrian Civil War[96,97] in 2013, which killed hundreds. With this new threat it is vital more than ever that new materials aimed at detection, capture and destruction of these agents are developed. As such the detection of the CWAs and the hydrolysis products has significant worth. An ideal sensor would use soil and ground water samples to establish information about when an attack took place, verify which CWA was used (even when it has been destroyed), and establish the extent of exposure to aid treatment.
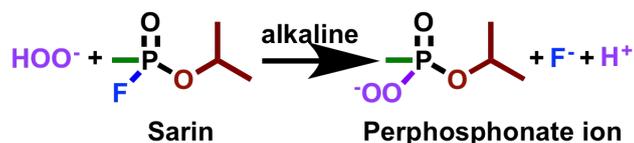
Fig. 3.2 Schoenemann Chemistry: Reaction of Sarin, under basic conditions, with hydrogen peroxide to give a perphosphonate salt.

Methods for detection and isolation of the agents are still in relative infancy of development. Early work towards this goal started in the 1950s[98,99] with the use of the Schoenemann chemistry (a reaction with hydrogen peroxide to give a perphosphonate ion, See Fig. 3.2) for the detection of anticholinesterase agents similar to Sarin. Recent work has continued to use selective reactions[100,101] with the active agents to aid detection through determining loss of a known starting reagent. More recently systems have used reversible interactions for sensing such as hydrogen bonding receptors[102], hydrophobic pore receptors[103], supramolecular gels[104–106], and metal-organic frameworks[107,108]. A range of different detection methods have been used including[101,109,110], surface plasmon resonance[111], luminescence[112], and NMR spectroscopies[102].

3.1.2 G-Series Nerve Agents

The acute toxicity of G-series nerve agents comes from their ability to irreversibly bind to, and inhibit, the enzyme acetylcholinesterase. The nerve agents all contain a good leaving group marked in blue Fig. 3.1, which makes the phosphorus centre vulnerable to nucleophilic attack by a residue within the active site of the enzyme. As well being driven to encapsulation within the enzyme through the hydrophobic effect, the CH_3 groups (or

equivalent) form favourable interactions with a pocket within the binding site of the enzyme.
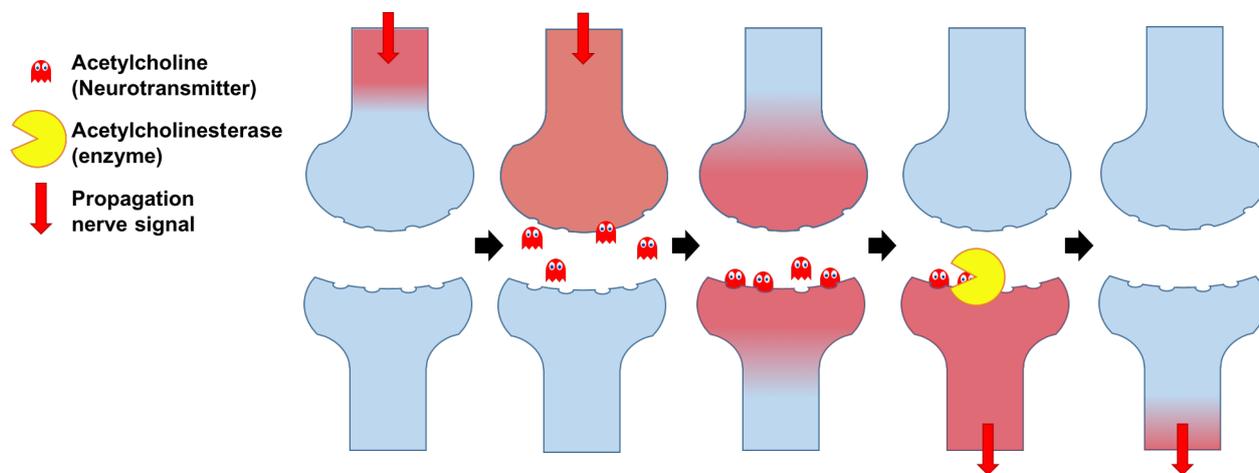
Fig. 3.3   How nerves normally transmit signals

For nerves to function normally, neurotransmitters are released over a synaptic junction to propagate the signal onwards. The acetylcholinesterase enzyme then digests the neurotransmitters to reset the synaptic junction (Fig. 3.3). The nerve agents bind to and deactivate this enzyme. The resulting build-up of neurotransmitter in the synapse means the nerve continuously signals which leads to muscle spasms, the inability to breathe, along with other unpleasant symptoms. Ultimately the combination of these affects can lead to asphyxiation and death. Skin contact with just a pin-head sized amount of Sarin can lead to death.
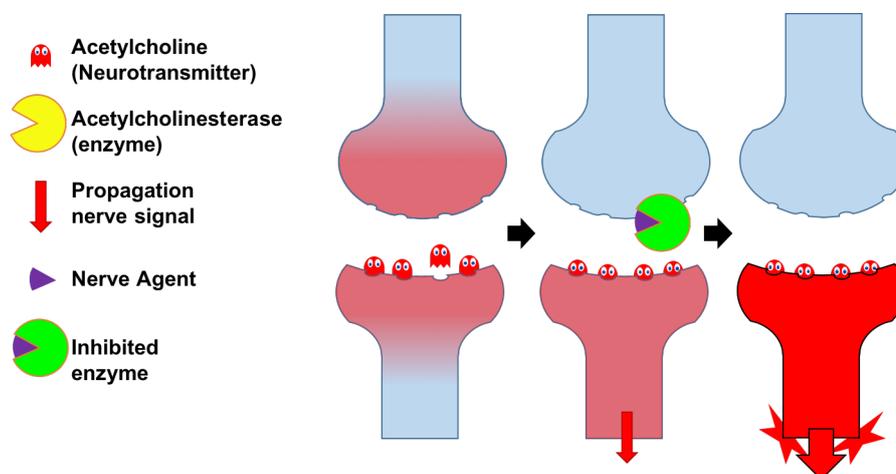
Fig. 3.4   How nerves signal transmission is affected by nerve agent

3.1.3 CWAs and Supramolecular Chemistry

Recent interest has turned to using supramolecular interactions for molecular recognition of the nerve agents; an area that so far has been explored relatively little[113]. While the physical properties and chemical reactivity of these agents have been explored, and are reported in comprehensive reviews[114,115], the use of supramolecular interactions to trap them still requires greater understanding. The interactions available from the CWAs for binding can involve the alkyl side chain or the phosphoryl P=O bond. The former requires a binding strategy involving hydrophobically driven encapsulation. In addition short-range interactions, such as π-CH interactions, can be used to enhance encapsulation and improve binding selectivity. To bind the P=O group the electronegative oxygen atom can be utilised as an effective hydrogen bond acceptor. The size and functionality of the CWAs differ across each series, and significantly between different series, meaning that hosts require tuning for the effective binding of particular agents. A second consideration that must be taken into account to investigate the interactions of these CWAs is their acute toxicity. Consequently the use and handling of CWAs is heavily restricted and they can only be used inside specialist facilities. Therefore most work is done on 'simulants' of the real agents.
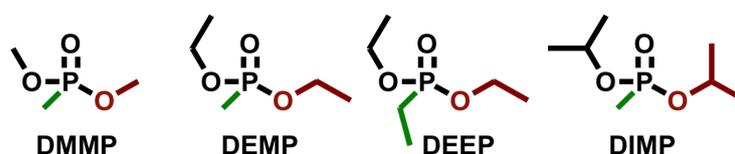
Fig. 3.5 CWA shape and size simulant

The choice of simulant is not straightforward. In the past many simulants have been chosen based on their physical properties and are not be appropriate for the investigation of supramolecular interactions. Typically most of the simulants found in the literature are pesticides or insecticides and are size/shape appropriate. The series of alkyl phosphonates (Fig. 3.1) contain the key functionality that need to be preserved in the chosen simulants. The active P-F group in the agents can be replaced by an alkyl ester, which results in the loss of chirality normally seen in the real CWAs.

Gale and co-workers have investigated the interactions of CWA simulants with supramolecular hydrogen-bonded gels[104–106,116]. Binding of the simulants in the gel

can result in a change to the gel's hydrogen bonding network. Small changes are then amplified through the extended gel network and a physical response is observable. Therefore the gels offer both a medium for trapping the CWAs while also providing a mechanism for optical sensing. The same group has also demonstrated that hydrogen-bonding based receptor molecules can lead to increased hydrolytic destruction of related substrates[102,117,118]. Recently it has been shown that metal-organic frameworks can also successfully be used for the catalytic destruction of chemical warfare agents[107,108].

There is still only a single example in literature where a coordination cage host has been used to bind a CWA simulant. Nitschke and co-workers showed that an organophosphate insecticide (dichlorvos) can be bound within a tetrahedral cage and undergo accelerated hydrolysis[119].

3.2 Discussion

3.2.1 Binding of CWA simulants

The guest binding of a series of alkyl phosphonates was investigated in two of the isostructural cages, $\mathbf{H^A}$ and $\mathbf{H^W}$ (described earlier in Chapter 1.6). Combination of ligand $\mathbf{L^A}$ or $\mathbf{L^W}$ with $Co(BF_4)_2$ provides cubic cages that are soluble in acetonitrile, $\mathbf{H^A}$, or water, $\mathbf{H^W}$, respectively. The alkyl phosphonates can be used as CWA simulants of real agents with a similar shape and size. The volumes of the phosphonate guests were calculated (see Table 3.1 later in this chapter) using Spartan-06. Based on the Rebek's 55 % rule[83–85] all the chosen simulants are below the optimal guest volume for the cage host[80] of 220 $\mathrm{\mathring{A}}^3$, and so there should be no steric barrier to binding.

The binding constants for the series of simulants were determined by ^{1}H-NMR spectroscopy. As a result of the inclusion of Co^{II} ions in both complexes ($\mathbf{H^A}$ and $\mathbf{H^W}$) the ^{1}H-NMR signals for the hosts are spread across 200 ppm, allowing the change in chemical shift caused by guest binding to be easily observed (Fig. 3.6). The two largest guests, DIMP and DEMP, were found to bind in slow guest exchange between the bound and free state on the NMR timescale in both water and acetonitrile whereas the smaller guests showed fast exchange behaviour.

3.2.1.1 Example of a guest in slow exchange on the NMR timescale
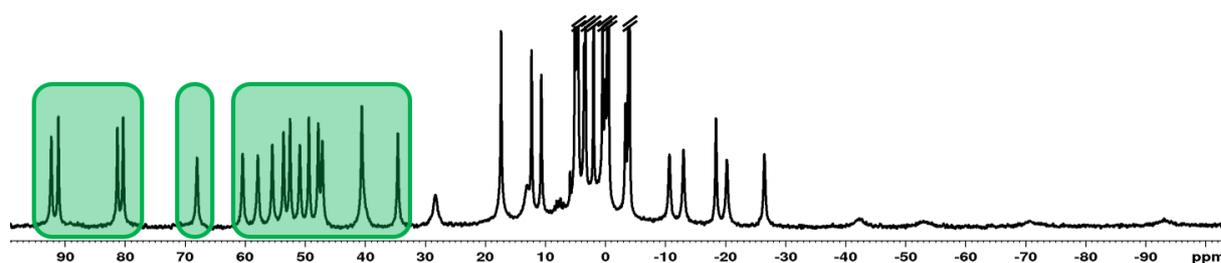
Fig. 3.6 Paramagnetic ^{1}H-NMR spectrum of H^W with expansions shown in Fig. 3.7 highlighted

The paramagnetic ^{1}H-NMR spectrum for **H^W** is shown in Fig. 3.6. The areas highlighted are expanded in Fig. 3.7 with the free host signals indicated with green arrows. The bottom spectrum shows the host by itself with the remaining spectra shifting with increasing guest concentration of DEEP . The host signals can be seen to reduce in intensity before disappearing with increasing guest concentration, while the host-guest complex signal can be seen to increase in intensity. The occurrence of separate signals for empty host and bound host is consistent with slow guest exchange on the NMR timescale.
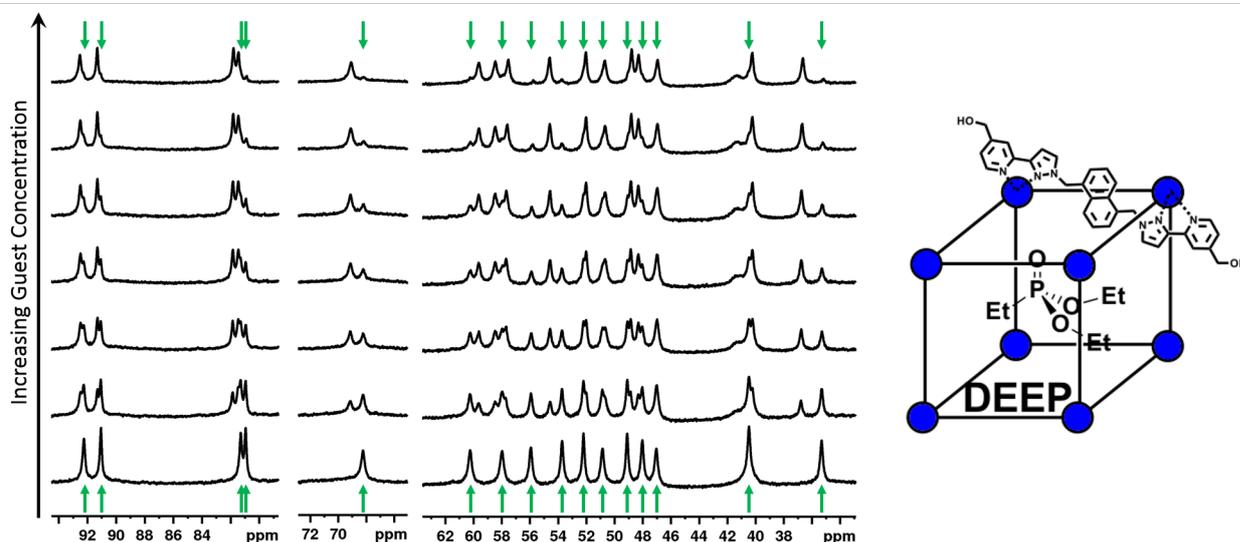
Fig. 3.7 Selections from ^{1}H-NMR spectra showing the change in peaks for the empty host and host/guest with increasing concentration of guest when the guest is in slow exchange on the NMR timescale: green arrows point at the peaks for empty host

When the guest is in slow exchange, each spectrum can be deconvoluted and integrated directly to get the ratio of H (host) to H·G (host-guest complex). In combination with the known [H]$_o$ (starting host concentration), the parameters [H] (concentration of free host), [G] (concentration of guest) and [HG] (concentration of H·G) can be calculated

137

and then equation (1.3) (see 1.4.2) can be used to obtain the desired binding constants. To obtain the spectra over a wide range (200 ppm) an intense and short pulse is required for excitation. Unfortunately the excitation is not completely uniform across the whole range so any integrations can only be compared with the immediate neighbouring peak.
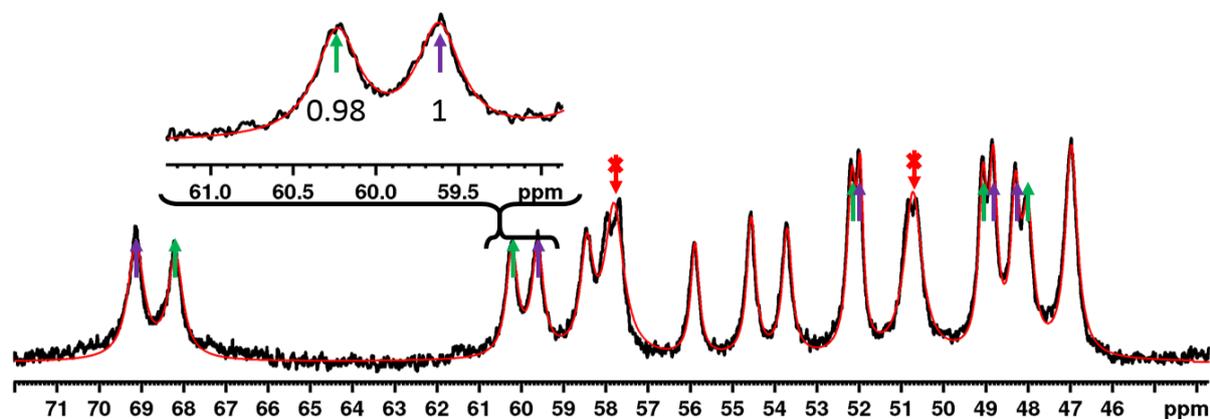
Fig. 3.8 Deconvolution of free host and host/guest signals for a guest in slow exchange: arrows highlight peaks for (green) empty host, (purple) host/guest, (red) areas that are not deconvoluted

3.2.1.2 Example of a guest in fast exchange on the NMR timescale

The two smaller guests, DMMP and DEMP, were found to be in fast exchange for binding in the host in both solvent systems. Fig. 3.9 shows the paramagnetic ^{1}H-NMR spectrum for $\mathbf{H^A}$. While some of the host NMR signals are consistent between the two isostructural cages, most of the peaks are slightly different. An expansion for each peak highlighted by a blue arrow (Fig. 3.9) is shown in Fig. 3.10. The shift of the host peak induced by guest can be seen binding in fast exchange with increasing guest concentration of DEMP.

Unlike what is seen for slow exchange, for guest binding in fast exchange the signals remain at a comparable integration throughout the titration. At low or zero concentration of guest, the host is fully unbound and only the free host peak is observed, while at high [G], all of H is bound as the $H \cdot G$ complex and only those signals are present. The peaks observed in the intermediate spectra are an averaging of the signal for H and $H \cdot G$ with a shift position are dependent on the ratio between [H] and [HG].

For the titration experiment, the shift of a signal (change in ppm) is tracked for several different peaks before being plotted against guest concentration to give a binding
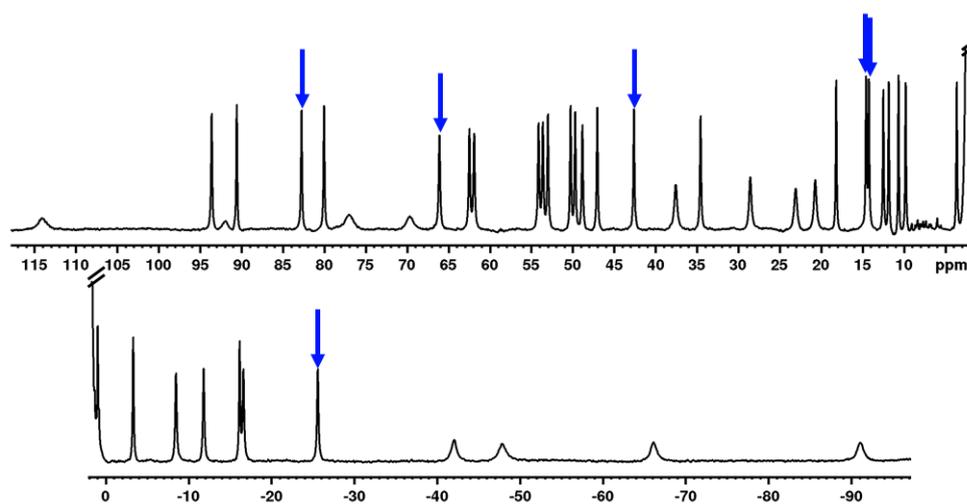
Fig. 3.9 Paramagnetic ^{1}H-NMR spectrum of $\mathbf{H^A}$ with selected peaks highlighted with blue arrows for Fig. 3.10
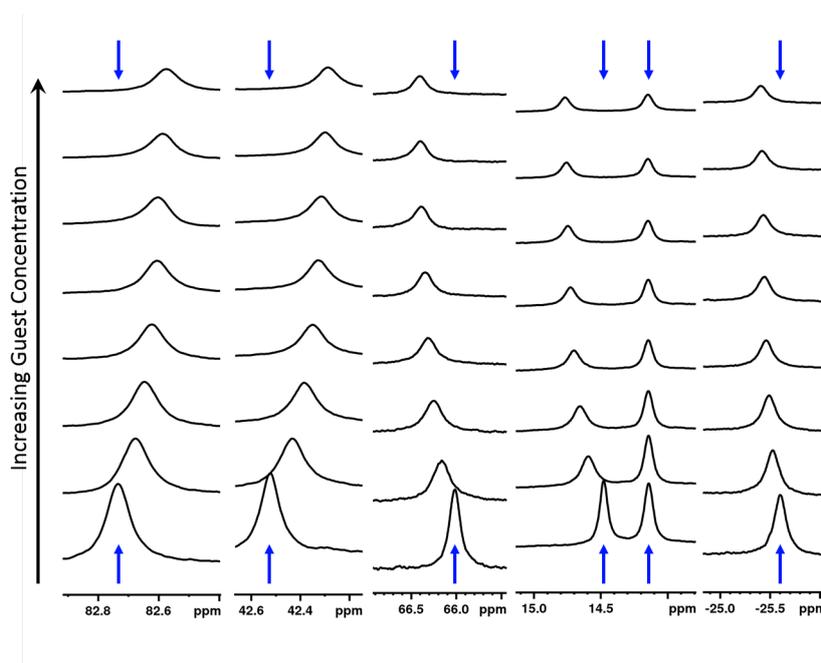
Fig. 3.10 Selection of ^{1}H-NMR spectra showing the changes in peak positions with increased guest concentration: blue arrows highlight the position of the peak for empty host

isotherm (circles in Fig. 3.11). A calculated binding isotherm is then fitted to each data set through a non-linear regression to give the binding constant.

3.2.1.3 Example of 2:1 guest:host binding

The binding of DMMP in the water soluble cage was also in fast exchange but unlike the others, fitted a 2:1 system (two guests bound to one host). The difference in the shape of the isotherm is visible in Fig. 3.12 where the shape at the start of the curve is notably
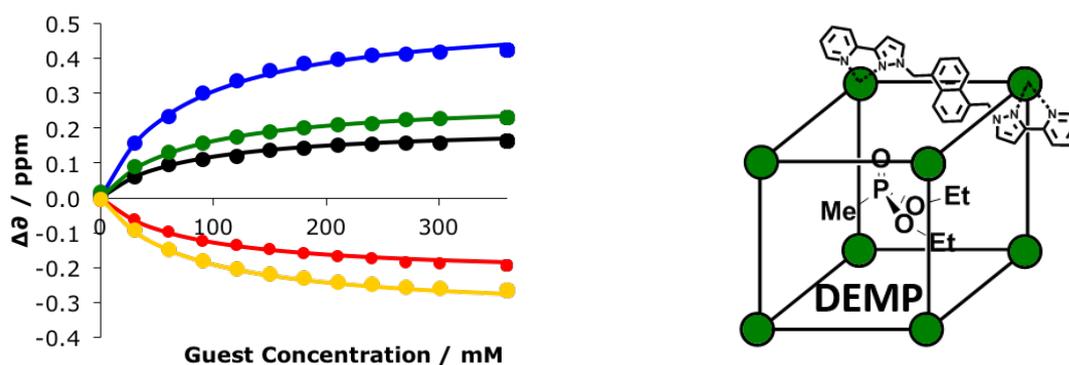
Fig. 3.11   (left) 1:1 binding isotherm for DEMP with $\mathbf{H^A}$ (right) cartoon of $\mathbf{H^A}$ and guest DEMP

different. The microscopic binding constants were determined to be the same for both of the DMMP guests. It is noted that the binding of two guests may be either cooperative or competitive (binding of the first guest can either improve or hinder the binding affinity of the second guest) so values of K_a for the first and second bound guests could differ. We did not observe any significant difference between the binding of the two DMMP guests which suggests that the two binding events are independent of each other. The averaged binding for each is sufficient for the comparison with the other phosphonate guests.
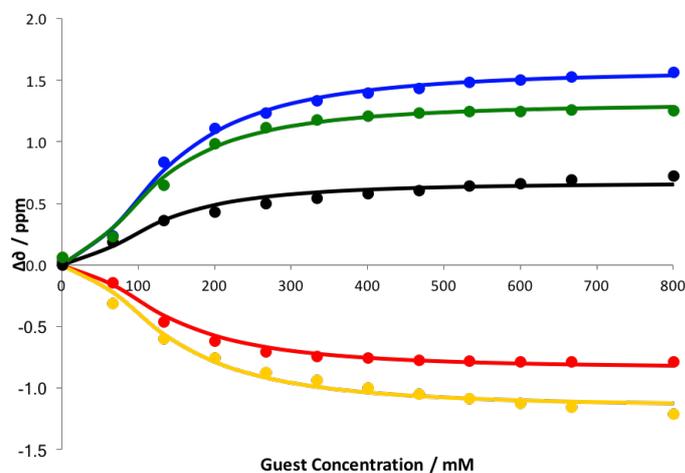
Fig. 3.12   2:1 G:H binding isotherm for guest DMMP with $\mathbf{H^W}$

The individual bindings of the two DMMP guests, averaged to $7\,\mathrm{M}^{-1}$ each, from the global binding constant of $49\,\mathrm{M}^{-1}$ for the pair. The combined volume $(238\,\mathrm{\mathring{A}}^3)$ for the two guest molecules is close to the optimum guest size based on Rebek 55% rule[83–85]. The binding constants in Table 3.1 are the average of at least three independent measurements with the error calculated as twice the standard deviation. In acetonitrile the binding constants were all found to be relatively small with only marginal differences between

the different guests.

Table 3.1 Guest data; volume of guests (calculated using Spartan-06) and binding constants

	DMMP	DEMP	DEEP	DIMP
Guest Volume (Å^3)	119	157	175	193
$\mathbf{Co_8L_{12}^W}$/D$_2$O (M^{-1})	7(2)	26(23)	160(30)	390(80)
$\mathbf{Co_8L_{12}^W}$/CD$_3$CN (M^{-1})	4(1)	14(3)	14(3)	9(1)

In water the binding constants differ across the series of alkyl phosphonate guests and showed an obvious progression from 7(2) M^{-1} for each bound DMMP, to 390(80) M^{-1} for DIMP: equal to an increase in the magnitude of ΔG from $-4.8(6)\,\text{kJ}\,\text{mol}^{-1}$ to $-15.0(5)\,\text{kJ}\,\text{mol}^{-1}$ respectively. The guests have an increasing number of methylene groups across the series with a corresponding increase in volume. This increase in size relates to an increase in the hydrophobic contribution for guest binding within the hydrophobic cage cavity[82].

A previous example demonstrating a hydrophobic contribution to cage/guest binding used a series of cyclic ketones[82]. It was demonstrated that the addition of each CH$_2$ group added $4.7\,\text{kJ}\,\text{mol}^{-1}$ to guest binding in water up to a point where the guest is to large for the cavity[120]. Within the series of phosphonate guests in Table 3.1 the increase in ΔG per CH$_2$ averages to $2.5\,\text{kJ}\,\text{mol}^{-1}$. This smaller contribution to binding energy compared to the ketones may be due to an increased flexibility in the phosphonates which results in a larger entropic penalty for binding. The cyclic ketones are held in a 'pre-organised' configuration while the flexible alkyl chains of the phosphonates have to reorganise to fit into the cavity. Another difference in contribution could be due to the guest not being fully desolvated upon binding with supporting evidence seen in the crystal structures (discussed below 3.2.2).

3.2.2 Crystal structures of host-guest complexes

To probe the specific interactions and positioning of the bound guests within the cage cavity, the crystal structures for host-guest complexes were determined for a series of guest bound within the two isostructural cages ($\mathbf{H^A}$ and $\mathbf{H^W}$). Only the $\mathbf{H^A}\bullet$DMMP complex gave good quality crystals through conventional solvent-diffusion methods. The
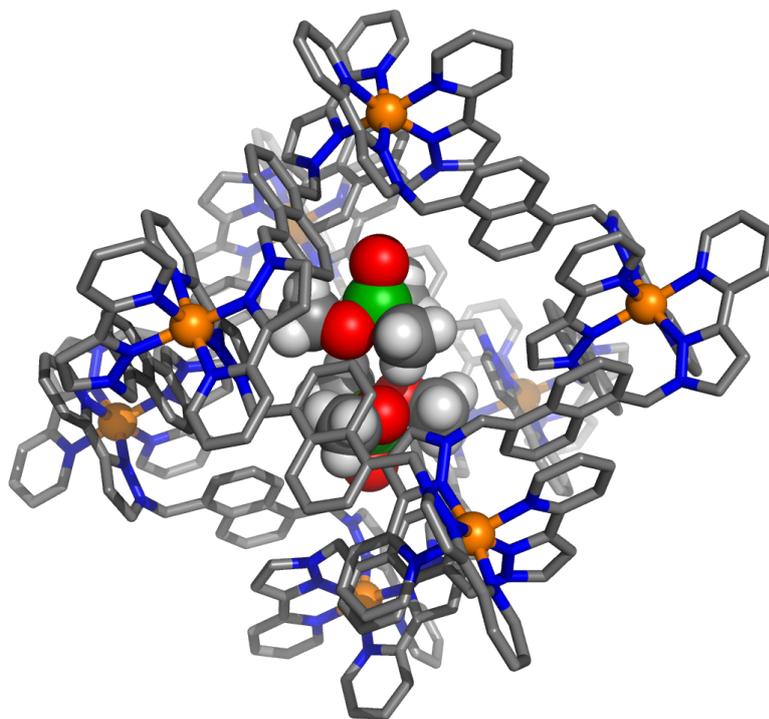
Fig. 3.13 Crystal structure of $[Co_8(L^A)_{12}](BF_4)_{16} \bullet 2(DMMP)$

framework of the cage itself in the $\mathbf{H^A} \bullet 2x(DMMP)$ complex did not differ from previously cage structures obtained (Fig. 3.13). Unusually two guest molecules were bound within the cavity with the two P=O bonds pointing to the opposite *fac* vertices where the hydrogen bond donor sites are located. We observed a 2:1 guest binding of DMMP with $\mathbf{H^W}$ in solution, and the crystal structure of $\mathbf{H^A} \bullet 2x(DMMP)$ also showed the same guest configuration. The two guests together appear to perfectly fill the cavity. Attempts to crystallise other host-guest complexes of this phosphonate series from a mixture of host and guest in solution failed by only yielding crystals of the cages with unoccupied cavities.

In contrast to the above $\mathbf{H^A} \bullet 2x(DMMP)$ complex, the remaining cage/guest structures were obtained through a crystal soaking method. Crystals of host cage obtained through solvothermal reactions were then soaked in a concentrated methanol-guest solution over the weekend before the structures were determined using Fujita's "crystalline sponge" method[121].

As with the DMMP-cage structure above, all the other host-guest complex crystal structures, with both isostructural cages, were found to have the guest P=O bond directed into one of the two hydrogen-bonding donor sites that arise from the converging C-H protons located around the electron deficient *fac* tris-chelate Co^{II} vertices. The protons within the binding pocket (highlighted in green Fig. 3.17) can collectively act
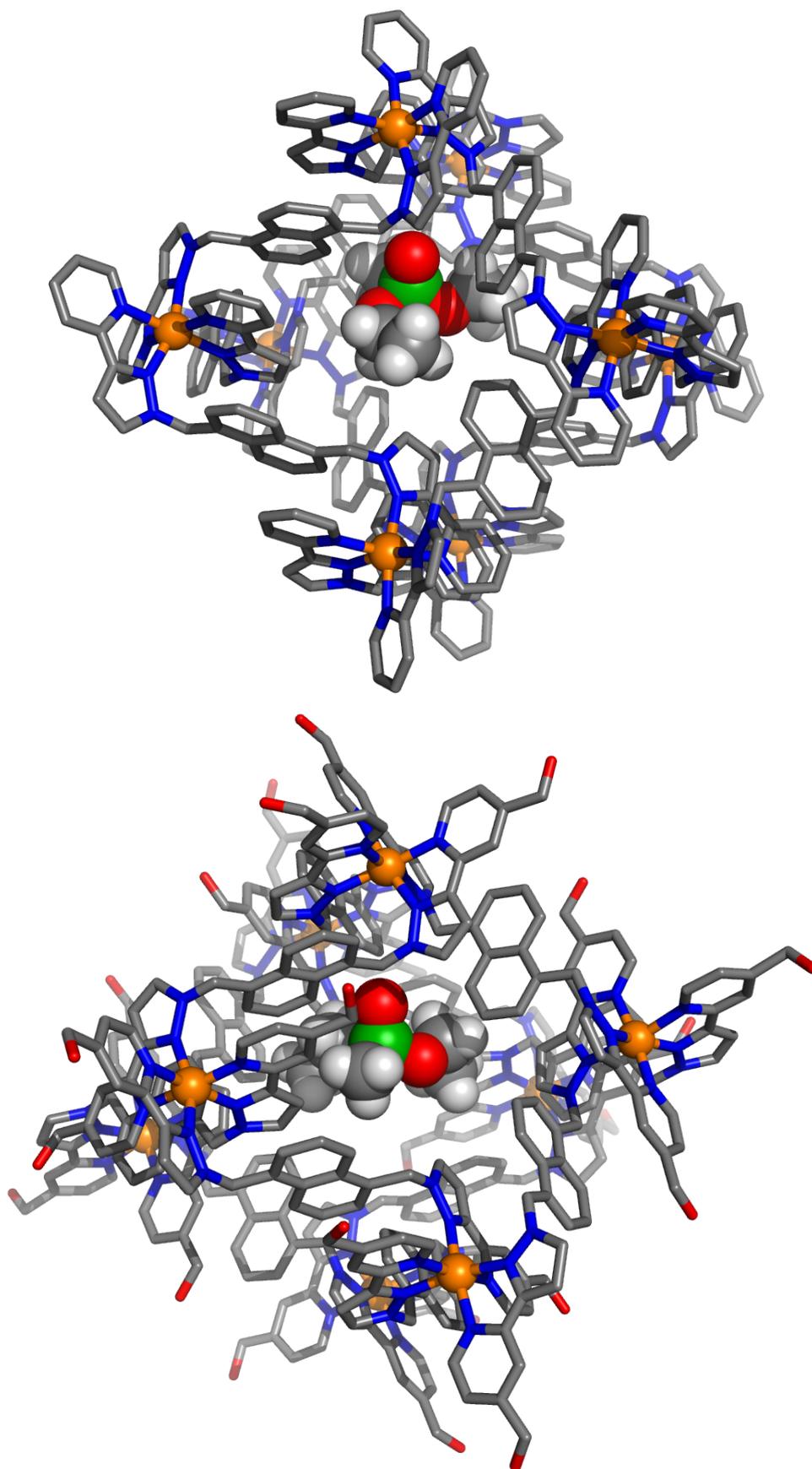
Fig. 3.14 Crystal structure of host-guest complex: (top) $[Co_8(L^A)_{12}](BF_4)_{16} \bullet$ DEMP (bottom) $[Co_8(L^W)_{12}](BF_4)_{16} \bullet$ DEMP
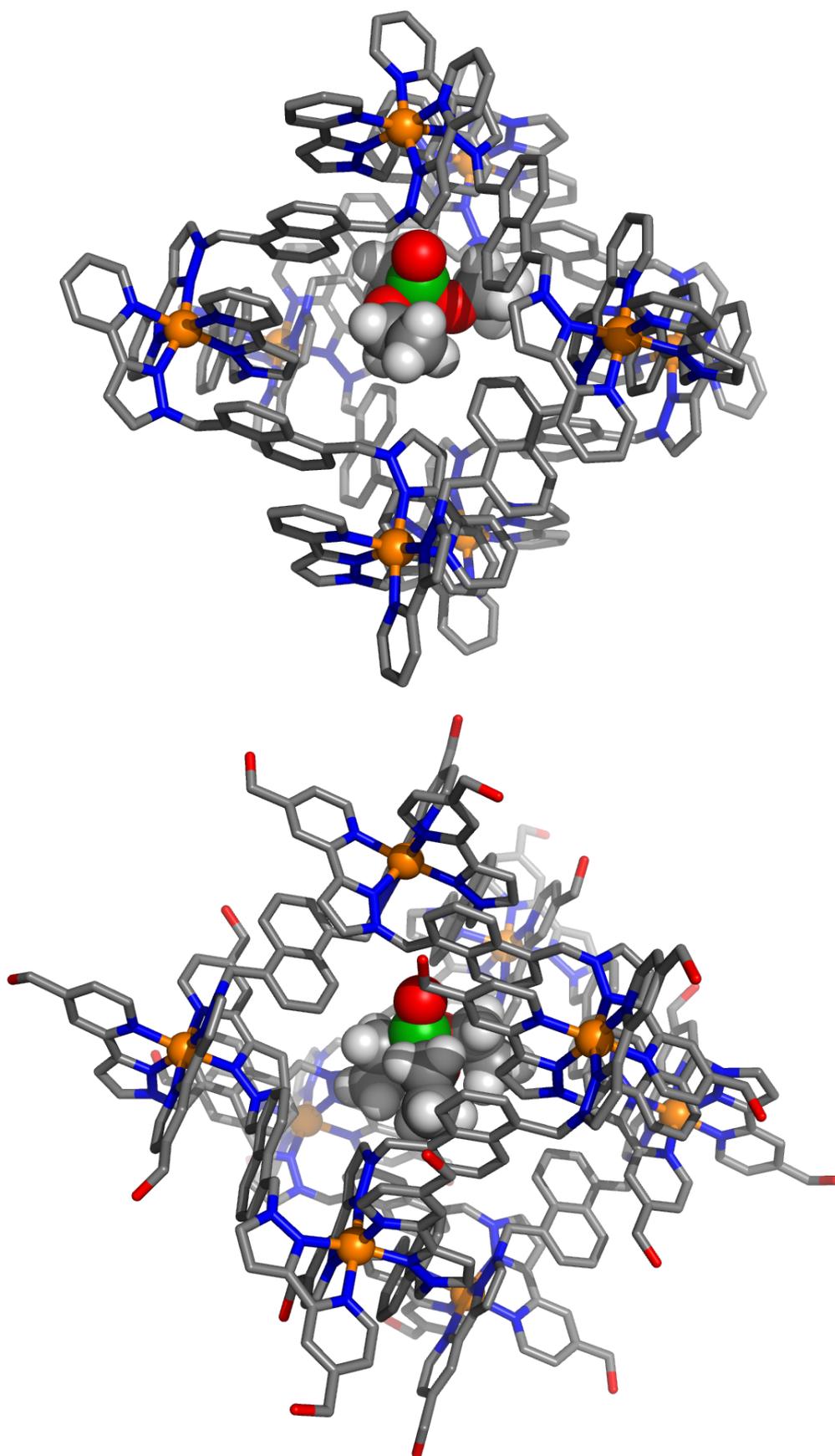
Fig. 3.15 Crystal structure of host-guest complexes: (top) $[Co_8(L^A)_{12}](BF_4)_{16} \bullet$ DEEP (bottom) $[Co_8(L^W)_{12}](BF_4)_{16} \bullet$ DEEP
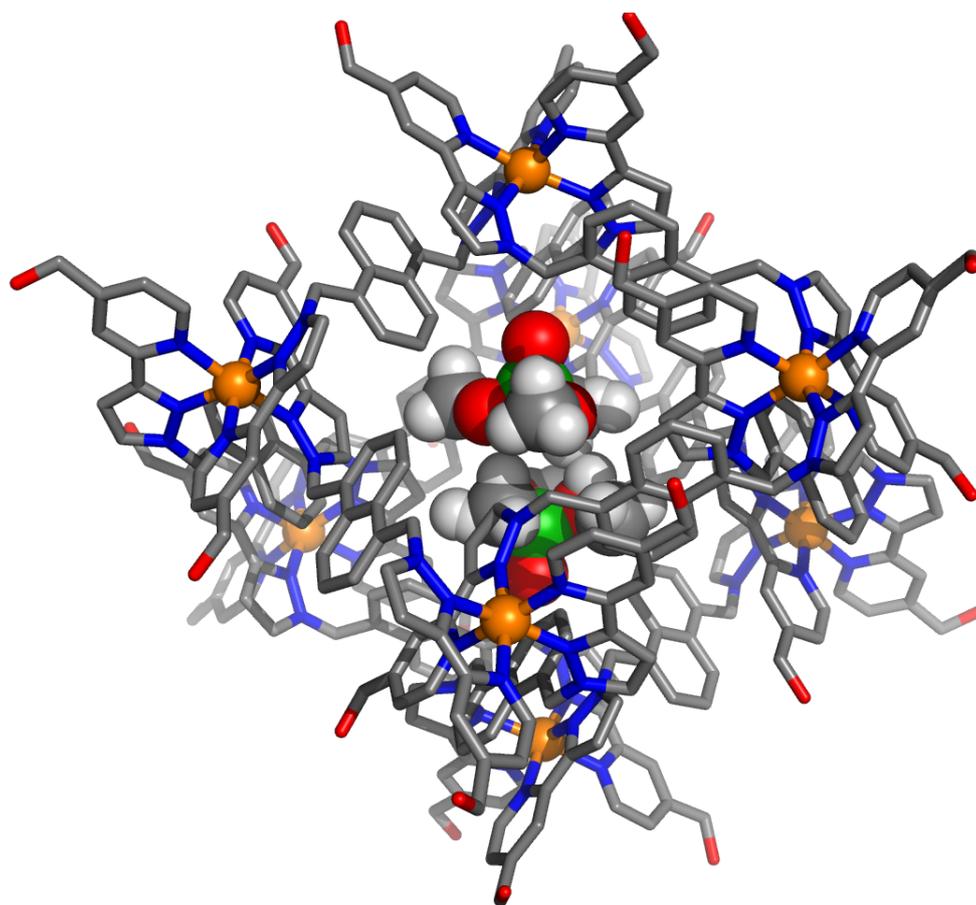
Fig. 3.16 Crystal structure of host-guest complexes: $[Co_8(L^W)_{12}](BF_4)_{16} \bullet 2(DMMP)$
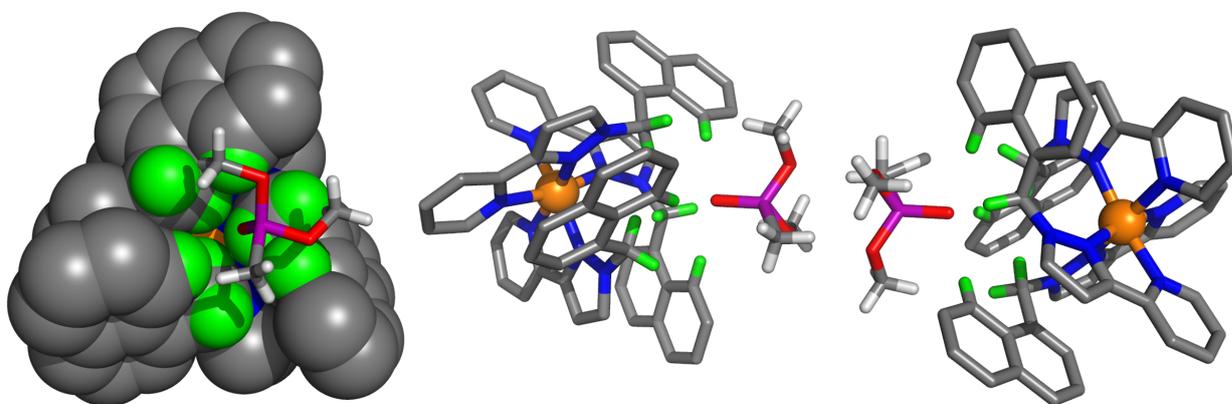
Fig. 3.17 Stick model of DMMP bound into the spacefilling *fac* corner

as a hydrogen bond donor site due to the proximity of the positively charged metal ion. In Fig. 3.18, the central image of the *fac* binding site's calculated electrostatic surface potential showing an area of electron deficiency around the converging protons. As a result the electron rich oxygen of the guest is oriented into the *fac* tris-chelate vertex. Often, when there is only space for a single guest, a solvent molecule can be found sat within the opposite *fac* vertex.

H^W•2x(DMMP) is nearly identical to the previous **H^A**•2x(DMMP) structure with
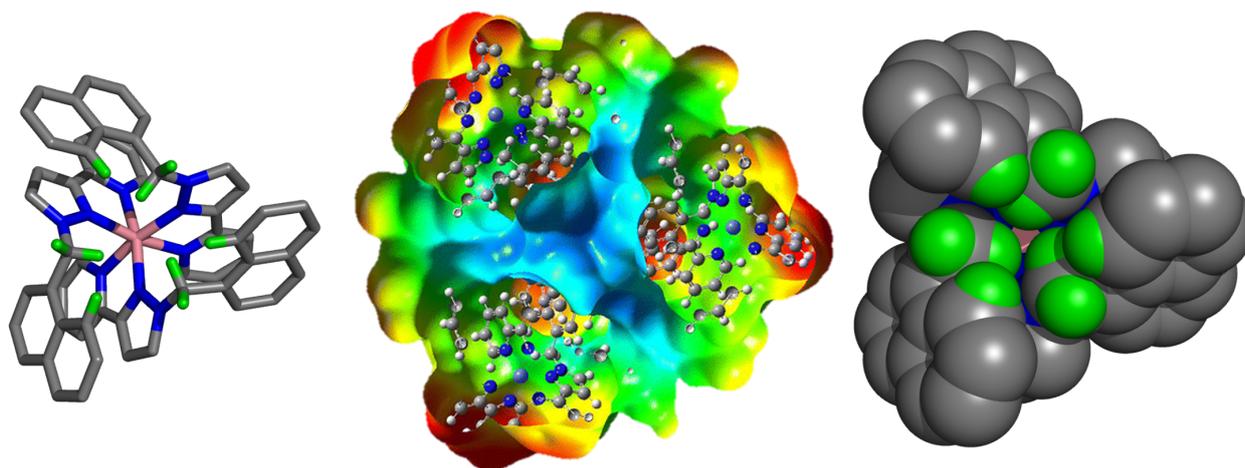
Fig. 3.18 The hydrogen bonding *fac* binding site (left) stick model, (middle)[120] Calculated electron potential surface map; values range from +0.81 atomic units (blue) to +0.5 (red), (right) spacefill model. Protons are highlighted in green. (central portion) Reproduced (in-part) from ref[120]. Copyright 2013 American Chemical Society

only the substituents on the external surface of the cage differing. The two DMMP guests are close together in the cavity, both have their P=O bond along an axis that runs through both *fac* vertices, with a rotational offset of 60° such that the two guests are arranged in a staggered confirmation with the alkyl groups interlocking between the two guests. The alkyl groups can be seen to pointing towards the cage portals in a arrangement that produces the least steric interference (Fig. 3.17).

3.3 Conclusion

The binding constants of a series of alkyl phosphonates, commonly used as chemical warfare agent simulants, have been determined within the host complexes $\mathbf{H^A}$ and $\mathbf{H^W}$ in acetonitrile and water solvents respectively. Determining the crystal structures of the host-guest complexes revealed that the phosphonates bound in the cavity of both hosts and were orientated toward the electron-deficient set of converging protons within one of the two *fac* vertices through a set of hydrogen bonding interactions.

It was found that two molecules of the smallest guest, DMMP, occupied the cavity in the solid state and accordingly was also found to have a 2:1 guest:host binding in water. It was observed that there was little difference in binding strength between the different guest across the series with in acetonitrile $\mathbf{H^A}$. Binding with $\mathbf{H^W}$ in water showed an

increase in binding strength across the series in proportion to increasing hydrophobic surface area of guest.

The successful binding of the series of CWA simulants is strong evidence that the active nerve agents would also bind within the cage host and therefore the cage could be used for entrapment of the agents. Additionally, other work within the Ward group has demonstrated that a Cd^{II} based cage ($[Cd_8L^W_{12}](BF_4)_{16}$) can produce partial quenching of the cage's fluorescence upon binding of the simulants and therefore could form the basis in a luminescent sensor which can signal CWA binding[93].

3.4 Experimental

3.4.1 Measurements and Calculations

3.4.1.1 Host and Guests

All guests were purchased from commercial sources and used as supplied unless otherwise stated.

$\mathbf{H^A}$, $[Co_8(L^A)_{12}](BF_4)_{16}$, and $\mathbf{H^W}$, $[Co_8(L^W)_{12}](BF_4)_{16}$ were both prepared as stated in Chapter 2.

3.4.1.2 Nuclear Magnetic Resonance Spectra

All NMR data was collected using a Bruker Avance III 400 MHz NMR spectrometer at 298 K. A majority of data for the binding constants within this chapter were collected using a standard `zg30` (3072 scans, 284 ppm sweep width, `O1p` value at 0 ppm, `D1` 0.4 s `AQ` 0.6 s) pulse program but some of the later data was collected with the improved `zgse` pulse program (same as before except `O1p` value at 5.29 ppm). The NMR data were all processed using Bruker Topspin 3.2.

The spectrum shown in Fig. 3.19 are both for $\mathbf{H^W}$ in D_2O using the two different pulse programs described above. The top spectra was obtained using a standard proton pulse sequence with a widened sweep width. Peaks are individually obtainable with a good signal-to-noise ratio by the end of collection. However it is apparent that there is a singular and very broad hump within the baseline of the spectra centred at around 5
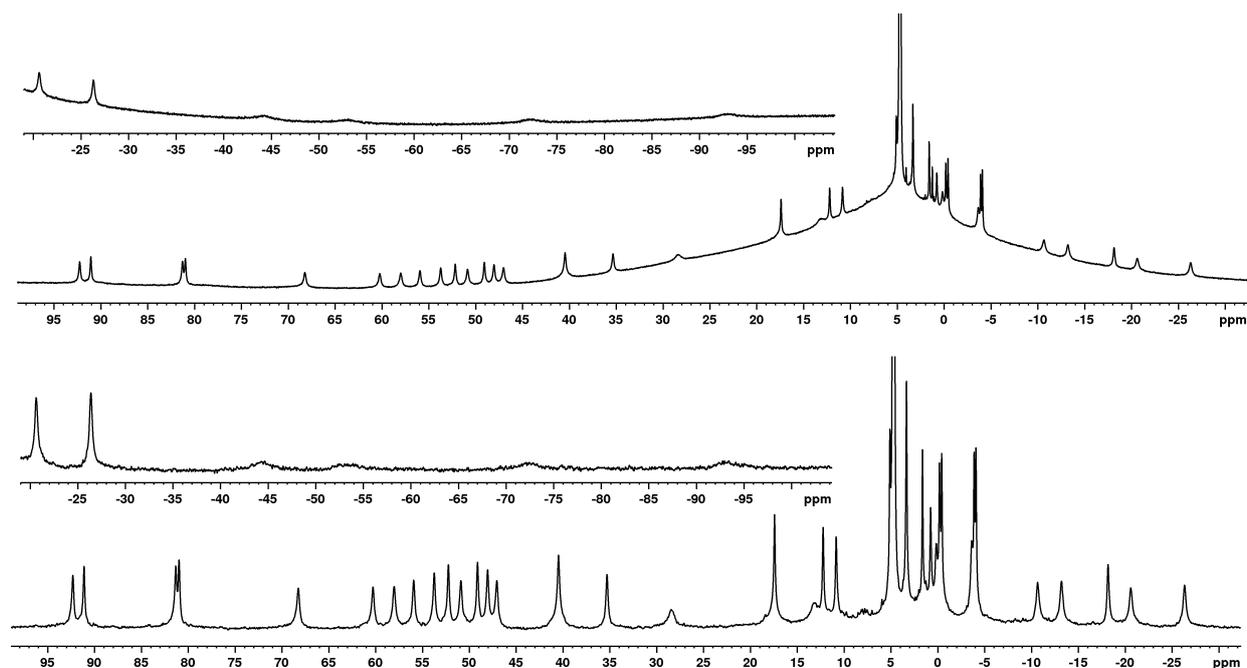
Fig. 3.19 Paramagnetic NMR

ppm. As a result it is not possible to use an auto-phasing and time is required for manual phasing of spectrum. It is only possible to correct the baseline for small selected regions and this is again time-consuming.

While developing a solvent suppression program used for binding of guests in host $\mathbf{H^B}$ in dichloromethane it was discovered that the broad peak, seen to decay quickly, originates from the instrument itself. Normally for standard proton NMR spectra the acquisition of the spectra occurs after a short delay and the broad peak would have disappeared. However due to the fast relation of the paramagnetic peaks the spectra must be collected before the broad hump has decayed. The presence of this hump is also re-inforced over the three-thousand scans required to obtain a spectra.

A spin-echo sequence was investigated to eliminate the broad hump. To obtain a spectrum for the paramagnetic complex a very fast and strong pulse is required. The strength of the pulse lies at the limits of the instrument. The resulting spectra, seen in the bottom half of Fig. 3.19 shows a significantly flatter baseline without a broad hump. The peaks are generally of a consistent size across the range. Due to the obvious improvement yielded from the spin-echo program the sequence has been adopted as the new standard for cage-based ^{1}H-NMR spectroscopy within the Ward group.

3.4.1.3 Determining Binding Constants

As described within the discussion, guest binding can be either slow or fast exchange on the NMR timescale. While processing of data differs, a standard setup is followed to obtain the titration spectra for both. Determining a binding constant with $\mathbf{H^W}$ used a host solution (0.2 mM, 10 mL) in D_2O. A guest solution was then made by dissolving the required concentration of guest into the host solution (5 mL resulting in two equally sized solutions). Both 5 mL solutions have identical host concentration but differing guest concentrations. From the two solutions a set of 12 tubes, each containing 600 µL, were created according to Table 3.2 to give a series of differing guest concentrations.

Table 3.2   Solution volumes used to setup standard titration

Tube Number	1	2	3	4	5	6	7	8	9	10	11	12
Host Solution (µL)	600	550	500	450	400	350	300	250	200	150	100	0
Guest Solution (µL)	0	50	100	150	200	250	300	350	400	450	500	600

For slow exchange guests individual pairs of peaks are manually phased and baselined before using the deconvolution command included within topspin. Once the ratio of host concentration [H] to host guest complex [HG] is determined the binding constant and free energy of binding can be determined from (1.4) as described in Chapter 1.4.2. Each tube is used as an independent experiment and multiple peaks are collected from each.

The process is performed twice and then the results used to give an averaged binding constant. Twice the standard deviation between the different experiments is determined and used as the given error for each binding constant. For the example data in Table 3.3, the binding constants for $\mathbf{H^W}$•DEEP averages to 170(36) M^{-1}. In combination with the repeated experiment the final average binding constant for DEEP with $\mathbf{H^W}$ in D_2O was 160(30) M^{-1}.

For fast exchange guests the spectra cannot be directly deconvoluted. Instead the change in chemical shift for different signals is tracked against changing guest concentration.

Example data for DEMP binding with $\mathbf{H^A}$ in fast exchange (Table 3.4) was fitted using 14allMaster.xls: a macro-based Excel fitting program written by Christopher

Table 3.3 Data extracted from the slow exchange spectra titration for DEEP in the water soluble cage

[H]/[HG]	$[G]_0$	[H]	[HG]	[G]	K	ΔG
2.2	3.33	0.14	0.06	3.27	140.2	-12.2
1.9	3.33	0.13	0.07	3.26	164.8	-12.6
1.5	3.33	0.12	0.08	3.25	204.1	-13.2
1.7	3.33	0.13	0.07	3.26	177.2	-12.8
0.3	20.00	0.05	0.15	19.80	167.4	-12.7
0.3	20.00	0.05	0.15	19.80	168.5	-12.7
0.3	20.00	0.05	0.16	19.80	172	-12.8
0.2	20.00	0.04	0.16	19.80	209.9	-13.2
0.4	13.30	0.06	0.15	13.20	199.3	-13.1
0.5	13.30	0.06	0.14	13.20	165.5	-12.7
0.5	13.30	0.06	0.14	13.20	160.6	-12.6
0.4	13.30	0.06	0.14	13.20	174.4	-12.8
0.8	6.67	0.09	0.11	6.56	182.4	-12.9
0.9	6.67	0.09	0.11	6.56	169.8	-12.7
0.9	6.67	0.09	0.11	6.56	172.2	-12.8
1.1	6.67	0.10	0.10	6.57	144.1	-12.3
0.6	10.00	0.07	0.13	9.87	169.2	-12.7
0.6	10.00	0.07	0.13	9.87	174.4	-12.8

A. Hunter (University of Cambridge). Once the binding constant is determined, the experiment is repeated a further two times and the three determined values are averaged.

3.4.1.4 Fitting of 2:1 Binding Constant

The standard equation for a binding constant (1.4) is only valid in a 1:1 binding system. Though support for other binding modes are available within the 14allMaster program, it was found that the program was unstable in the current version of Excel (from the Microsoft Office Professional Plus 2013 suite). Instead a custom spreadsheet was written for the purpose of fitting the alternative binding mode displayed by DMMP in water. The non-simplified version of (1.4) is defined below in

Table 3.4 Data extracted from the fast exchange spectra titration for DEMP in the acetonitrile soluble cage

$[\text{G}]_0$ mM^{-1}	δH 1	δH 2	δH 3	δH 4	δH 5	δH 6	δH 7
0	93.598	82.736	66.02	-25.601	42.524	52.92	-16.673
30	93.652	82.684	66.169	-25.529	42.437	53.082	-16.844
60	93.686	82.646	66.247	-25.487	42.383	53.184	-16.953
90	93.701	82.622	66.312	-25.46	42.35	53.247	-17.022
120	93.712	82.609	66.348	-25.443	42.326	53.295	-17.057
150	93.727	82.597	66.377	-25.428	42.31	53.327	-17.082
180	93.735	82.585	66.397	-25.417	42.299	53.351	-17.094
210	93.742	82.578	66.411	-25.409	42.289	53.374	-17.103
240	93.746	82.570	66.423	-25.404	42.281	53.391	-17.106
270	93.748	82.560	66.426	-25.395	42.272	53.402	-17.106
300	93.750	82.556	66.431	-25.39	42.269	53.414	-17.106
360	93.755	82.551	66.436	-25.387	42.264	53.435	-17.106

$$\text{M}[H] + \text{N}[G] \xrightleftharpoons{\beta_{MN}} [\text{H}_\text{M}\text{G}_\text{N}] \tag{3.1}$$
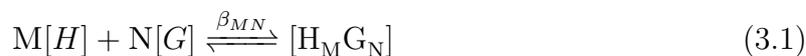
$$\beta_{MN} = \frac{[\text{H}_\text{M}\text{G}_\text{N}]}{[\text{H}]^\text{M}[\text{G}]^\text{N}} \qquad \Rightarrow \qquad \beta_2 = \frac{[\text{HG}_2]}{[\text{H}][\text{G}]^2} \tag{3.2}$$

For the titration of DMMP with $\mathbf{H^W}$ the concentration of guest $(800\,\text{mM})$ was four thousand times the concentration of host $(0.2\,\text{mM})$ allowing for the assumption that $[\text{G}]_0 = [\text{G}]$. Initially a crudely estimated binding constant is added to the spreadsheet. This value is then used in (3.3) to give δ_{calc}. The `SUMXMY2` function from excel is used to give a R^2 value; the sum of the squares of the difference between δ_{calc} and δ_{obs} across all of the isotherms being fitted. This R^2 value is used to describe the agreement between the calculated and observed values.

$$\delta_{calc} = \delta_H + (\delta_{HG} - \delta_H)f_{HG} \qquad f_{\,\text{HG}} = \frac{\beta_2[\text{G}]^2}{1 + \beta_2[\text{G}]^2} \tag{3.3}$$

δ_{obs} - observed chemical shift of peak
δ_{calc} - observed chemical shift of peak
δ_H - chemical shift for host
δ_{HG} - chemical shift for host guest complex
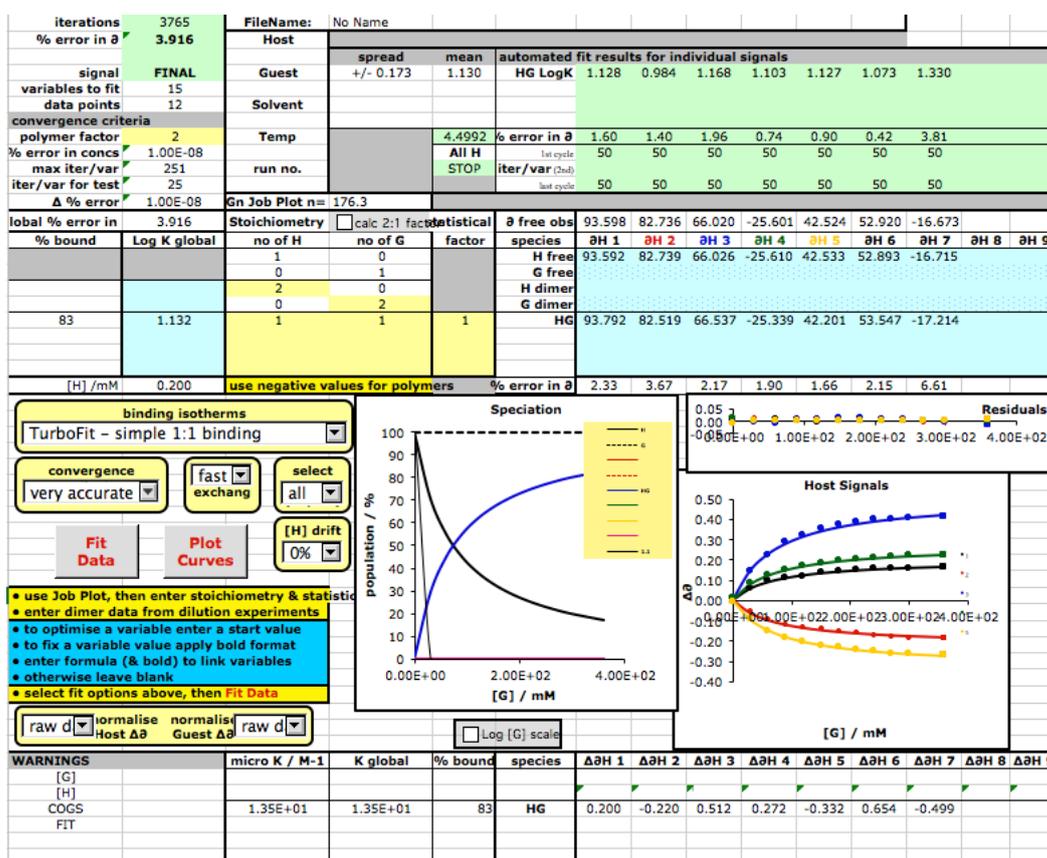
Fig. 3.20 Screen shot of 14AllMaster Excel program used to fit most binding constants

The Excel plug-in 'Solver' is then used with the objective set to minimise R^2 while varying β_2. δ_H and δ_{HG} values for each isotherm are set to the observed start and end values for each titration and are allowed to vary. The non-linear regression model, `GRG Nonlinear` is run to convergence to give a refined β_2. The change in chemical shift is plotted against guest concentration to check the final fit is sensible. It was assumed that the binding constant for each DMMP molecule was the same and therefore $(K_a)^2 = \beta_2$. The fitted constants for the the binding of DMMP with $\mathbf{H^W}$ giving $\beta_2 = 52$ and corresponding to $K_a = 7.2$ for each guest molecule.
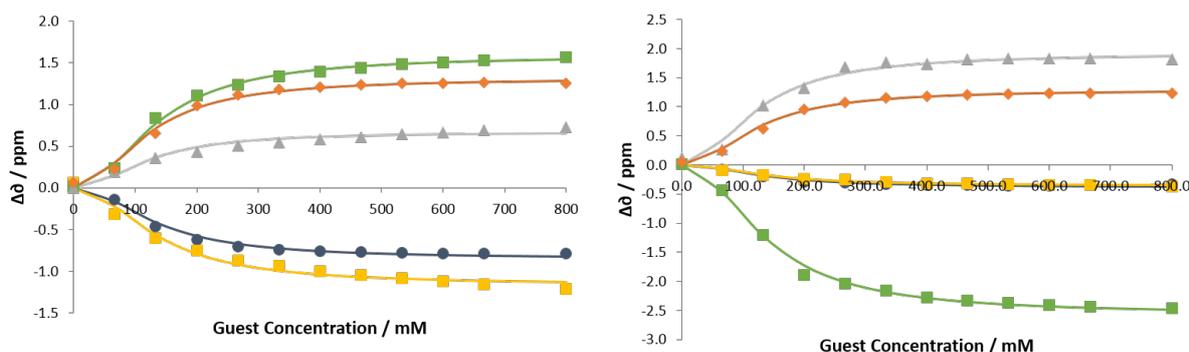
Fig. 3.21 Fitted binding isotherms for the 2:1 system ($\mathbf{H^W}\bullet2$DMMP)

3.4.2 X-ray Crystallography

The crystal structure data collections were performed using a Bruker APEX-2 CCD diffractometer with Mo-Kα radiation from a sealed tube source. Data were corrected for absorption effects using empirical methods (SADABS) based upon symmetry-equivalent reflections combined with measurements at different azimuthal angles. The structures were solved and refined using a combination of different software packages including Olex$_2$, WinGX, Apex3, PLATON and primarily the SHELX suite version 6.14.

AFIX restraints were applied to all the rings within the cage structure. Various weak global restraints were applied to the located anions and any disorder, where possible, was modelled. The guest was then located within the cavity and DFIX restrains used to optimise the guest parameters that are chemically sensible. Other solvent molecules that could be located were also modelled before the hydrogens were calculated for cage and guest. Due to a combination of weak diffraction data, strongly absorbing crystals and general disorder some hydrogen atoms, typically on the termini of the CH$_2$OH groups, were not retained. Once the model had converged there were large regions of diffuse electron density which could not be modelled accounting for any missing anions and solvent molecules. This electron density was removed from the final refinement using the 'SQUEEZE' function in the PLATON software package. The determined structure is of poor quality compared to conventional small-molecule standards though this is typical for these types of self-assembled coordination cage complexes. The overall structure of the cage, the presence of guest, and the relative position and orientation within the cavity are clear. While reference to relative occupancy of encapsulated guests and solvent molecule *vs.* the host complex are made it should be noted that these are approximate and therefore only used to support other data. No further claims for any specific structural details are used.

3.4.2.1 Crystallography Data Tables

Table 3.5

Complex	$[Co_8(L^A)_{12}](BF_4)_{16}$ • 2(DMMP)	$[Co_8(L^A)_{12}](BF_4)_{16}$ • DEMP
Formula[a]	$C_{339}H_{273}B_{16}Co_8F_{64}N_{72}O_3P$	$C_{341}H_{277}B_{16}Co_8F_{64}N_{72}O_3P$
Molecular Weight	7294.65	7322.70
T, K	100(2)	100(2)
Crystal System	Monoclinic	Monoclinic
Space Group	R_3	C2/c
a, Å	28.713(5)	32.788(12)
b, Å	28.713(5)	29.860(10)
c, Å	51.792(9)	40.029(14)
α, °	90	90
β, °	90	96.04(2)
γ, °	120	90
V, Å^3	36978(14)	38972(23)
Z	3	4
ρ, $g\,cm^{-1}$	0.983	1.248
Crystal Size, mm^3	0.16 x 0.14 x 0.12	0.20 x 0.10 x 0.10
Data, restraints, parameters	189974, 726, 510	44753, 2052, 1609
Final R1, wR2[b]	0.1728, 0.4310	0.2032, 0.4332

Table 3.6

Complex	$[Co_8(L^A)_{12}](BF_4)_{16}$ • DEEP • 0.5H$_2$O	$[Co_8(L^A)_{12}](BF_4)_{16}$ • DIMP
Formula[a]	$C_{342}H_{280}B_{16}Co_8F_{64}N_{72}O_{3.5}P$	$C_{343}H_{281}B_{16}Co_8F_{64}N_{72}O_3P$
Molecular Weight	7345.73	7350.75
T, K	100(2)	100(2)
Crystal System	Monoclinic	Monoclinic
Space Group	C2/c	C2/c
a, Å	32.8826(6)	32.7148(10)
b, Å	30.1786(6)	29.7919(8)
c, Å	40.0732(8)	39.8554(11)
α, °	90	90
β, °	96.4980(10)	96.733(2)
γ, °	90	90
V, Å^3	39511.2(13)	38576.6(19)
Z	4	4
ρ, $g\,cm^{-1}$	1.235	1.266
Crystal Size, mm^3	0.43 x 0.28 x 0.28	0.20 x 0.10 x 0.10
Data, restraints, parameters	30236, 2478, 1681	43994, 2142, 1674
Final R1, wR2[b]	0.1143, 0.3583	0.1335, 0.3770

Table 3.7

Complex	$[Co_8(L^W)_{12}](BF_4)_{16}$ • $2(DMMP)$	$[Co_8(L^W)_{12}](BF_4)_{16}$ • DEMP
Formula[a]	$C_{366}H_{330}B_{16}Co_8F_{64}N_{72}O_{30}P_2$	$C_{364}H_{323}B_{16}Co_8F_{64}N_{72}O_{27}P$
Molecular Weight	8139.34	8029.29
T, K	100(2)	100(2)
Crystal System	Monoclinic	Monoclinic
Space Group	C2/c	C2/c
a, Å	27.477(4)	27.475(3)
b, Å	39.201(6)	39.204(4)
c, Å	41.992(7)	42.208(5)
α, °	90	90
β, °	107.133(5)	107.517(6)
γ, °	90	90
V, Å^3	43224(12)	43354(9)
Z	4	4
ρ, $g\,cm^{-1}$	1.251	1.230
Crystal Size, mm^3	0.20 x 0.10 x 0.10	0.20 x 0.10 x 0.10
Data, restraints, parameters	38209, 2545, 1779	50015, 2499, 1786
Final R1, wR2[b]	0.1481, 0.4181	0.1851, 0.4716

Table 3.8

Complex	$[Co_8(L^W)_{12}](BF_4)_{16}$ • DEEP
Formula[a]	$C_{366}H_{327}B_{16}Co_8F_{64}N_{72}O_{27}P$
Molecular Weight	8057.35
T, K	100(2)
Crystal System	Monoclinic
Space Group	C2/c
a, Å	27.341(10)
b, Å	38.931(11)
c, Å	41.873(15)
α, °	90
β, °	107.25(3)
γ, °	90
V, Å^3	42565(26)
Z	4
ρ, $g\,cm^{-1}$	1.257
Crystal Size, mm^3	0.20 x 0.10 x 0.10
Data, restraints, parameters	49075, 2538, 1789
Final R1, wR2[b]	0.2040, 0.4612

a   These formulae (and consequently the crystal densities) are necessarily approximate given that large amounts of diffuse electron density in solvent-accessible voids were removed from the refinements using the 'SQUEEZE' function in PLATON. See CIFs and comments below for more details.

b   The value of `R1` is based on 'observed' data with $I > 2\sigma(I)$; the value of `wR2` is based on all data.

Chapter 4

Catalysed Hydrolysis of an Organophosphate Guest
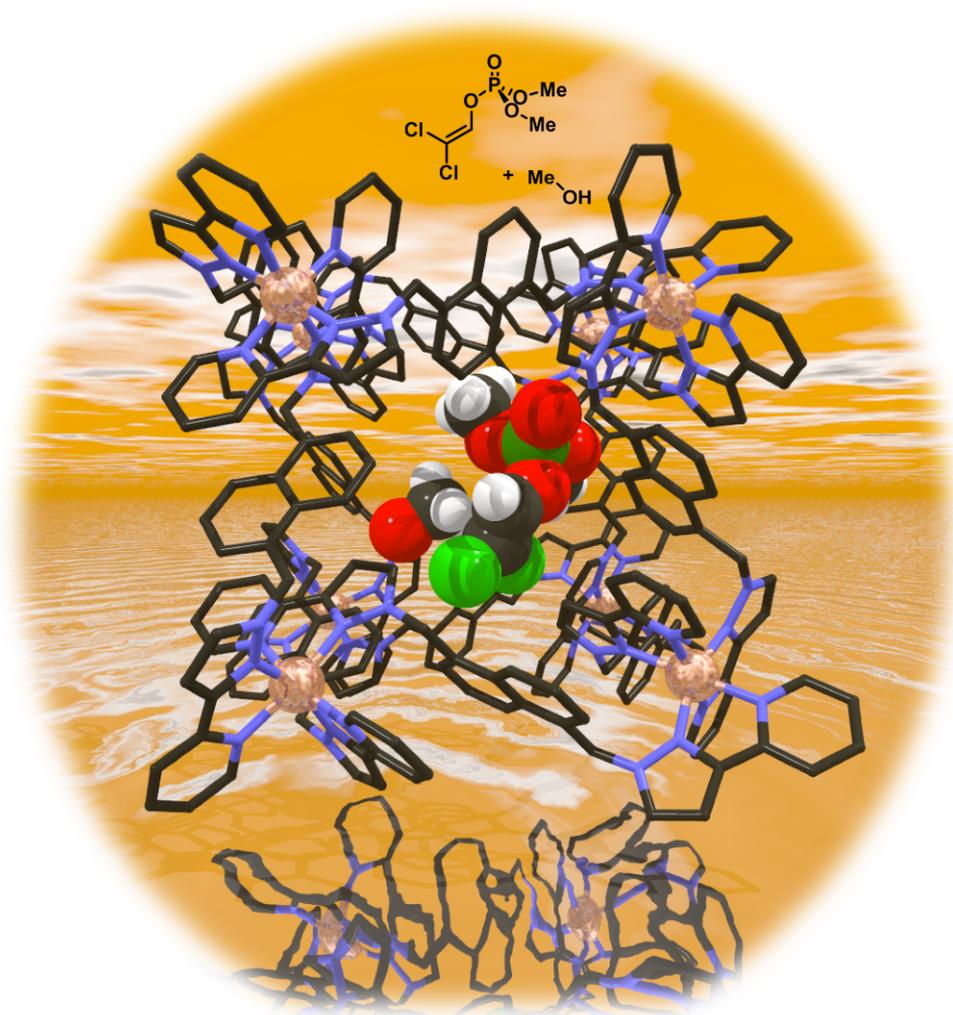
Rendering of the crystal structure of $\mathbf{H^A}$ containing guest; dichlorvos and methanol guests

4.1   Introduction

The stabilisation of bound guests within the cavity of coordination cage complexes, and modification of their reactivity, has been of great interest within the supramolecular chemistry field. Well known examples include Cram's stabilization of highly reactive cyclobutadiene[122] through to catalysed Diels-Alder reactions[37,58] (Chapter 1.5.1) with the ultimate aim of mimicking the ability of enzymes to catalyse reactions; both in selectivity and in rate enhancement[60,123]. For example, the enzyme β-amylase is able to catalyse the breakdown of starch with a rate enhancement of more than 10^{17} (k_{cat}/k_{non})[124].

To act as an efficient catalyst a molecular container needs to; recognise and bind guest(s), accelerate the reaction by either increasing local concentration and/or stabilisation of transition states, and expel the product to allow turnover. Guest binding within cage cavities have been well studied[9]. Many examples of accelerated reactions have also been reported[123]. However the factors leading to the reactions' acceleration often also result in tight binding of the reaction product leading to inefficient catalysis. Therefore the biggest challenge to establishing efficient catalysts is to achieve *both* reaction acceleration *and* product release.
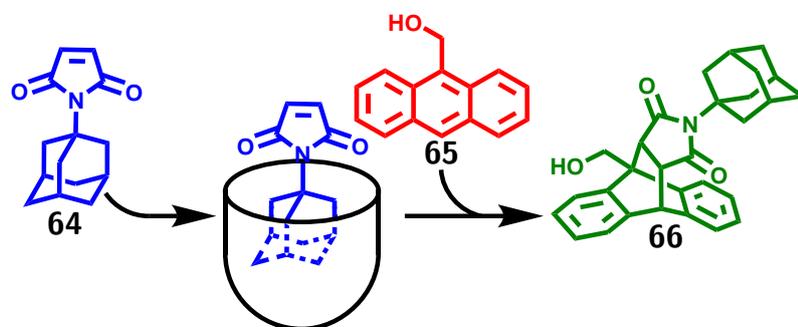
Fig. 4.1   Diels-Alder reaction within a cavitand with sterically-induced turnover[125]

Ensuring the release of product, and therefore turnover, can be achieved through simply introducing steric or electronic hindrance to product binding. Rebek and

co-workers achieved a Diels-Alder reaction, with turnover, in a cavitand with a rate enhancement of 30 fold by producing a product that is too sterically hindered to bind[125].

One of the largest known rate enhancements for a synthetic host is for the Nazarov cyclisation of pentamethyl cyclopentadienol **67**. Using an anionic cage, a rate enhancement of more than 10^6 was achieved for the reaction[126]. Recently the same cage assembly has been used to gain an improved rate enhancement of more than 10^7-fold for a catalysed reductive elimination of ethane from the dialkyl Au^{III} iodide complex **69**[127].
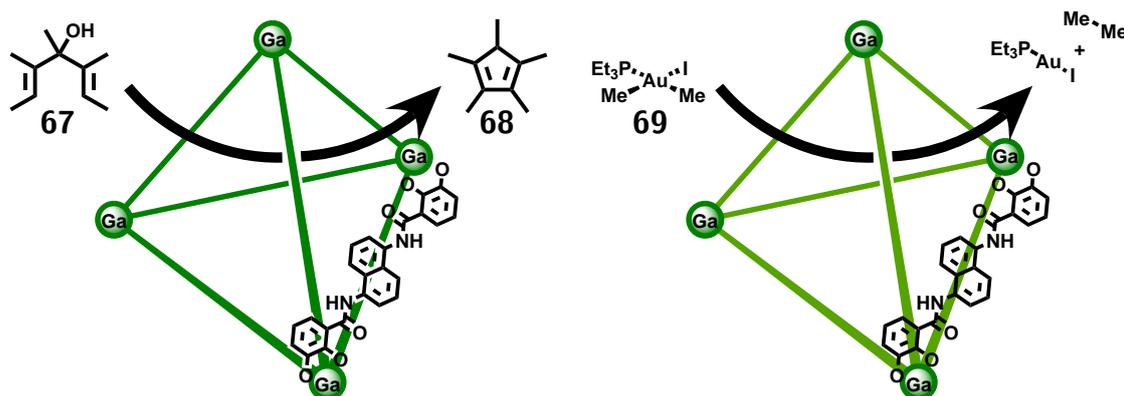
Fig. 4.2 (left) The Nazarov cyclisation within an $[Ga_4L_6]^{12-}$ complex[126], (right) catalysed elimination of ethane from a dimethyl Au^{III} complex[127]

4.1.1 Ward Group Catalysis

The cubic cage $[M_8L_{12}]$, $\mathbf{H^W}$, only binds neutral, hydrophobic guests strongly in water; cationic and anionic guests bind weakly as they are hydrophilic[128]. Protonated amines or deprotonated carboxylate anions, for example, bind weakly as they prefer to be solvated by the bulk water. As a result, guest uptake into and release from the cage can be controlled through changes to pH[128]. It has also been observed that, in nearly all crystal structures of the cubic cage, six of the anions are found to be located in the portals of the cage: one in each face.

The Ward group looked at the Kemp elimination[129], in which benzisoxazole undergoes a ring opening reaction with hydroxide to give 2-cyano-phenolate. Importantly the benzisoxazole has a binding constant of $K_a = 4 \times 10^3 M^{-1}$ and the background reaction is first order with respect to base down to pH 6.

The maximum rate enhancement of 2×10^5 was measured at pH 8.5. K_{cat} was observed to be independent of pH and therefore as the pH drops, K_{uncat} also drops leading
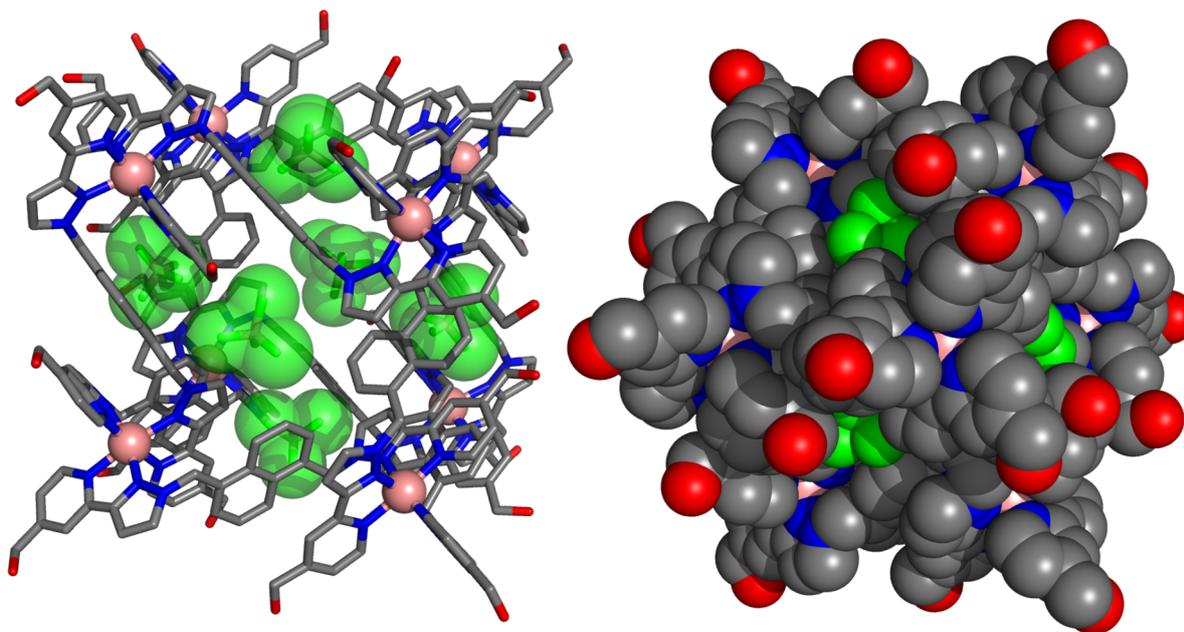
Fig. 4.3 Crystal structure of H$^{\text{W}}$: highlighting the six BF$_4$ anions located in the cage portals around the surface

to increased rate enhancements. If compared when both rates are pH independent the theoretical rate enhancement is $K_{cat}/K_{uncat} = 6 \times 10^{\,6}$. Both of these rate enhancements are greater than those previously observed by many other artificial systems for the Kemp elimination.
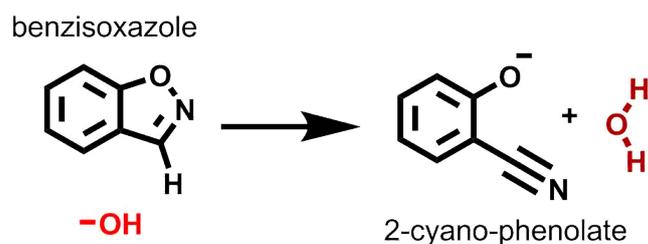
Scheme 4.1. The Kemp elimination[129]

This cubic coordination cage is effective at catalysis of the Kemp elimination on a bound benzisoxazole through a combination of two factors: a high concentration of hydroxide anions, partly desolvated, located in the portals by ion-pairing with the cage complex; and localisation of the hydrophobic substrate within the cavity. It is unusual but effective use of two orthogonal supramolecular interactions to bring together the two components required for the bimolecular reaction.
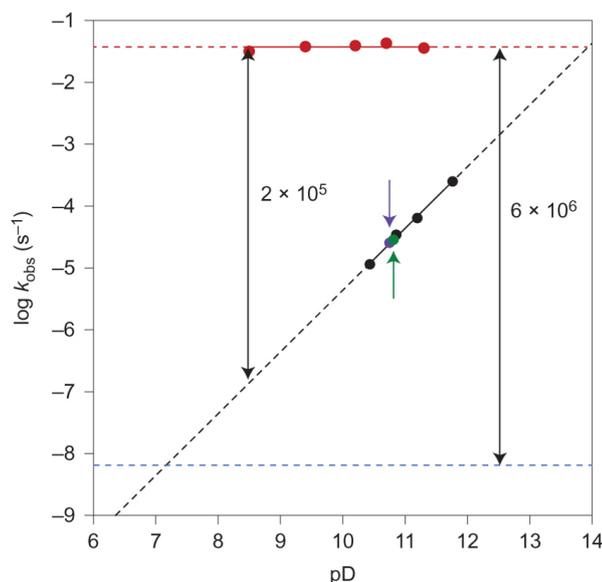
Fig. 4.4 Graph of reaction rates for the cage-catalysed Kemp elimination: (Black) K_{uncat}, (red) K_{cat}, (dashed blue) pD independent rate for K_{uncat}, (purple) suppressed cage catalysis by competing guest, (green) suppressed cage catalysis by anion competitor LiCl. Reprinted by permission from Macmillan Publishers Ltd: Nature Chemistry Ref[129], copyright 2017
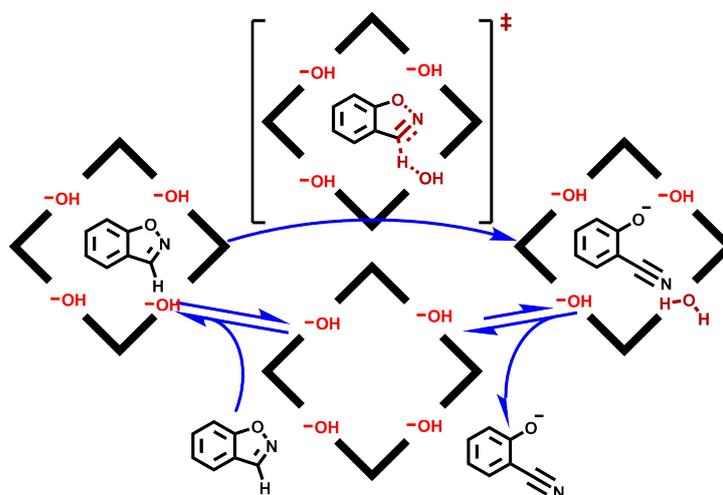
Fig. 4.5 Cartoon of the catalytic reaction cycle for the cage-catalysed Kemp elimination. Starting at the bottom left, benzisocazole binds in the cage with a proton next to a hydroxide ion in the adjacent cage window. The elimination reaction occurs to produce the negatively charged product which is ejected from the cage because it favours being solvent in the bulk[129]

4.1.2   CWA simulants hydrolysis

As discussed in Chapter 3, new methods for detection and/or trapping of CWAs remains important along with the development of strategies for their destruction[113]. While the agents can be hydrolysed in water, the reaction is either slow (at neutral pH) or needs to be very basic (high pH is hazardous to the environment).
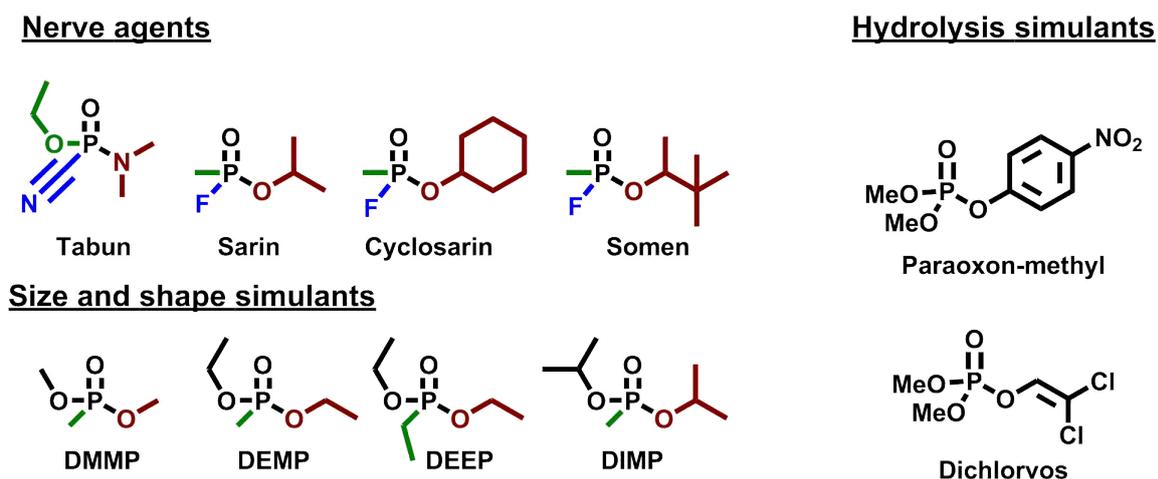
Fig. 4.6 CWAs and their related simulants

Cyclodextrins are small, simple, and easily available organic host molecules which are known to bind small nonpolar guests[103]. By encapsulation inside a cyclodextrin host, the hydrolysis of cyclosarin was observed to be enhanced in terms of both selectivity and rate[103].
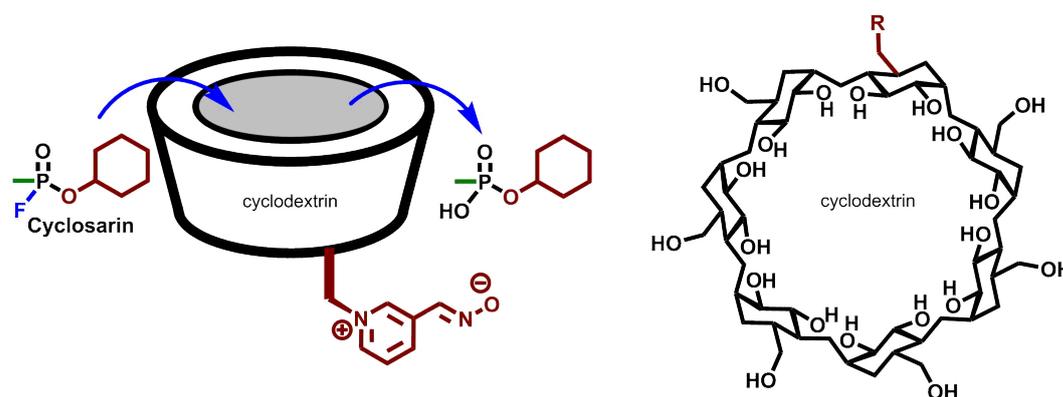
Fig. 4.7 Functionalised cyclodextrin that can hydrolysis cyclosarin[103]

Hupp and coworkers[107,108] have also shown that metal organic frameworks can both trap and catalysis the destruction of paraoxon-methyl (Pestanal), used here as a simulant of CWAs. The hydrolysis yields a p-nitrophenoxide anion that strongly absorbs light around 400 nm (paraoxon-methyl absorbs at around 260 nm) which allows for easy monitoring of the hydrolysis reaction by UV/Vis spectroscopy.

Nitschke and co-workers used an $[Fe_4L_6]$ cage complex to catalysis the organophosphate pesticide dichlorvos[119]; one of the only known such examples using a self-assembled coordination cage as the host. Monitoring using $^{31}P-NMR$ spectroscopy showed the accelerated conversion of dichlorvos to dimethyl phosphate acid **70** (major

product), and dichlorovinylmethyl phosphate acid **71** (minor product)[119].
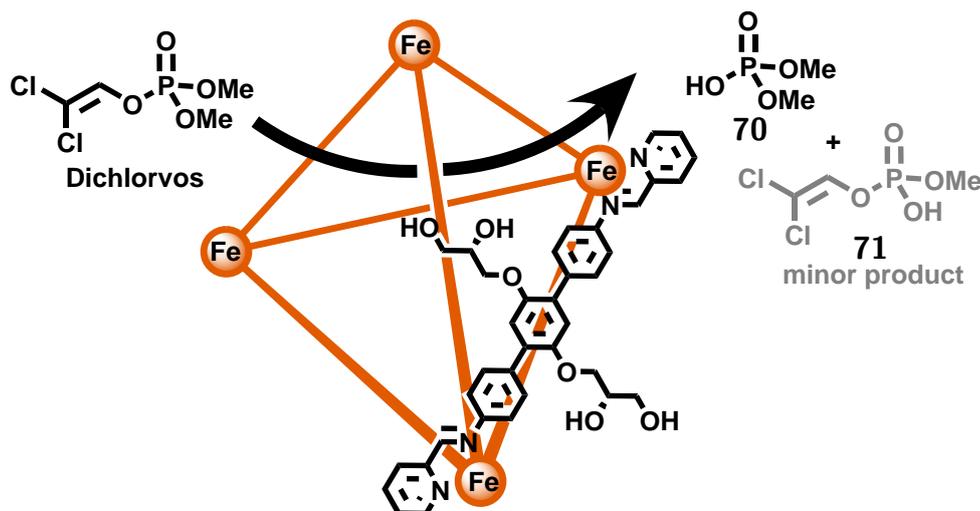
Fig. 4.8   Nitschke's [Fe$_4$L$_6$] cage complex can hydrolysis dichlorvos[119]

4.2   Discussion

While Paraoxon-methyl was initially investigated for binding in our cubic host cage, it was apparent that the guest did not bind within $\mathbf{H^W}$ due a size/shape misfit as a result of a rigid and bulky side group. An alternative substrate of interest was dichlorvos which has a smaller rigid side group (a dichloro-alkene) compared to Paraoxon-methyl (a nitro-benzyl). The binding constant of dichlorvos in $\mathbf{H^W}$ was initially measured under conditions previously described (Chapter 3.4) giving a value of $K = 62\,\text{M}^{-1}$. Dichlorvos was found to be in fast exchange on the NMR timescale and therefore the experimentally determined binding isotherm was fitted against a calculated isotherm.
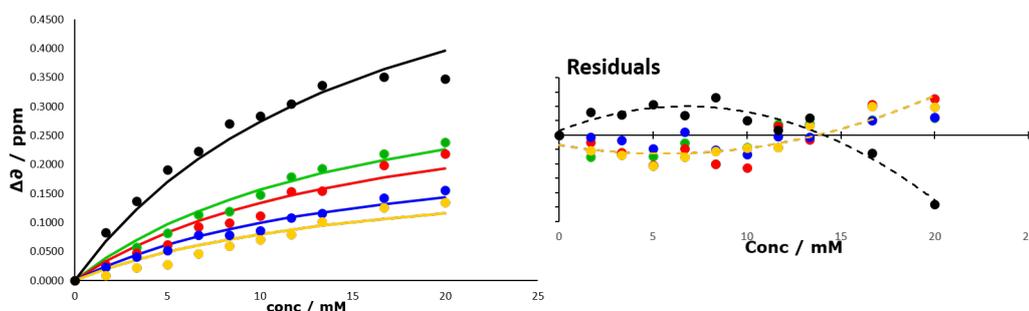
Fig. 4.9   Fitted binding isotherms for binding of guest dichlorvos in $\mathbf{H^W}$

There was an overall poor fit between the calculated and experimental binding isotherms with a clear systematic error between the values as seen in the residuals plot

in Fig. 4.9. The discrepancy between the experimentally determined chemical shifts of the H·G complex and the calculated values is likely down to a change in $[G]^o$. The experiment was performed using the previously described method (Chapter 3.4), with each spectrum at a different concentration taking around an hour to measure. Therefore the last spectrum to be obtained for the experiment corresponding to the largest guest concentration is measured around 11 h after the start of the experiment: during which time some of the dichlorvos guest will have hydrolysed. We return to this point later.
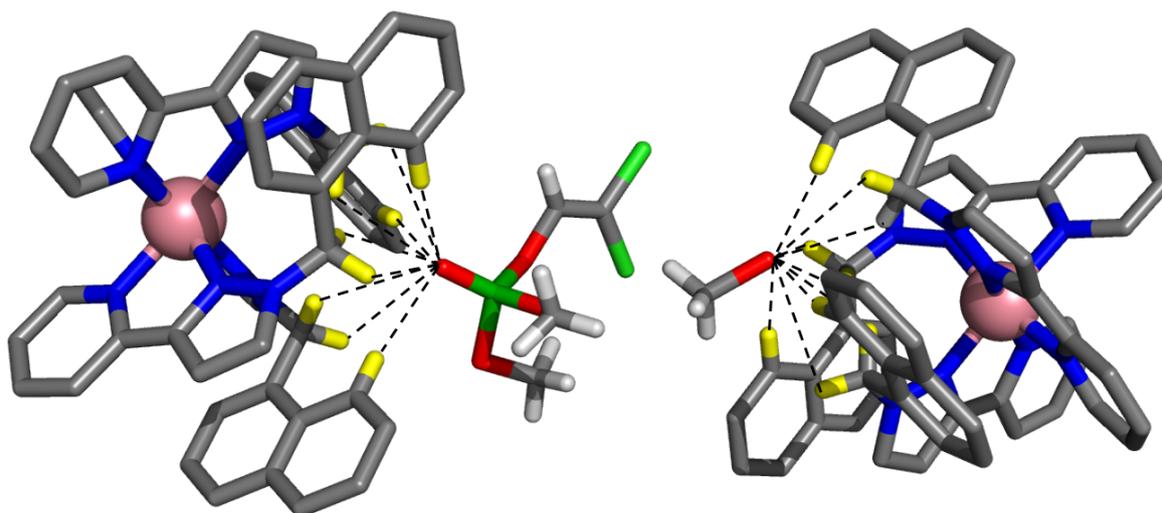
Fig. 4.10 A cut-away stick model of $\mathbf{H^A}$ containing a dichlorvos and a methanol guest in each *fac* vertex

The partial hydrolysis of dichorvos, even at the relative acidity of the unaltered cage solution (pH 3-4), clearly created issues for obtaining an accurate binding constant and equally caused difficulty in obtaining a crystal structure of the host-guest complex. Pre-grown crystals of $\mathbf{H^A}$ were soaked in a methanol solution of guest but after half a day the crystallinity was lost. Reducing the soaking period to just under 2.5 h, before picking and transferring a crystal directly into the cryostream (at 100 K), allowed successful collection of the diffraction data.

As expected the P=O double bond was found to be oriented into the hydrogen-bonding pocket of the cage. The two OMe groups of the dichlorvos guest point towards two of the nearest cage portals and appear to block access to the cavity. The dichloroalkene side group also points to a cage portal, but this is on the opposite side of the cage with the two chlorine atoms blocking the access to the cavity. The third
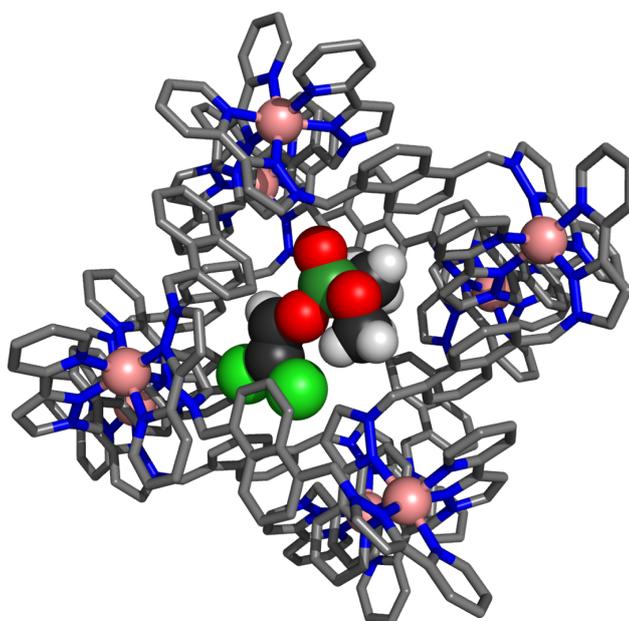
Fig. 4.11 A completed structure of $\mathbf{H^A}$ containing the guest molecule dichlorvos as a spacefilled model. The MeOH and anion are not included for clarity (Table 4.3)

portal, in close proximity to the binding site of the guest, is also blocked by the side group leaving just two unblocked portals. Within the crystal structure, the phosphorus atom is clearly exposed to potential nucleophilic attack by hydroxide ions through the remaining two portals. A methanol solvent molecule is also bound in the second hydrogen bonding pocket, which is likely to be replaced by a water molecule when under catalytic conditions.
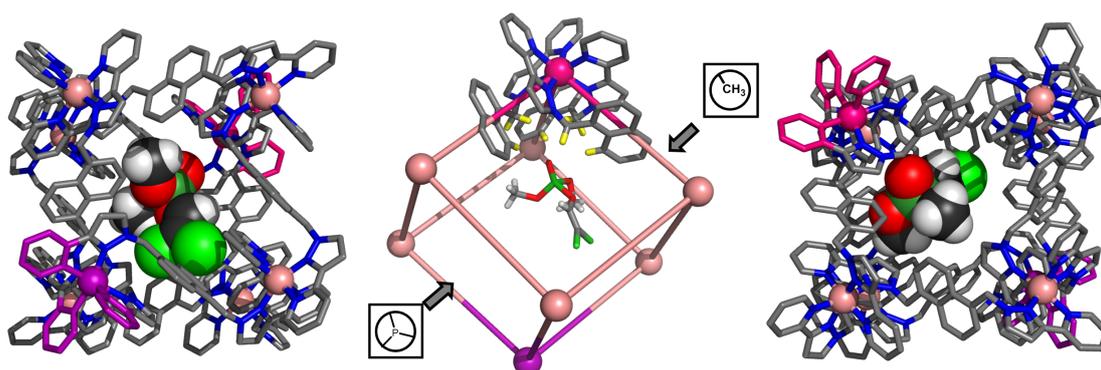
Fig. 4.12 View through the cage windows to the bound guest inside. Four of the windows are sterically blocked by the bound guest with the two remaining windows allowing a clear path of attach on the phosphorus atom of the bound guest. The two *fac* vertices are highlighted in pink (containing bound guest) and purple

The combination of the proposed mechanism for catalysis of the Kemp elimination, with an hydroxide anion located in each of the six portals, and the arrangement of the

dichlorvos in the solid state suggest that the cage complex might act as a catalyst for the hydrolysis of dichlorvos.
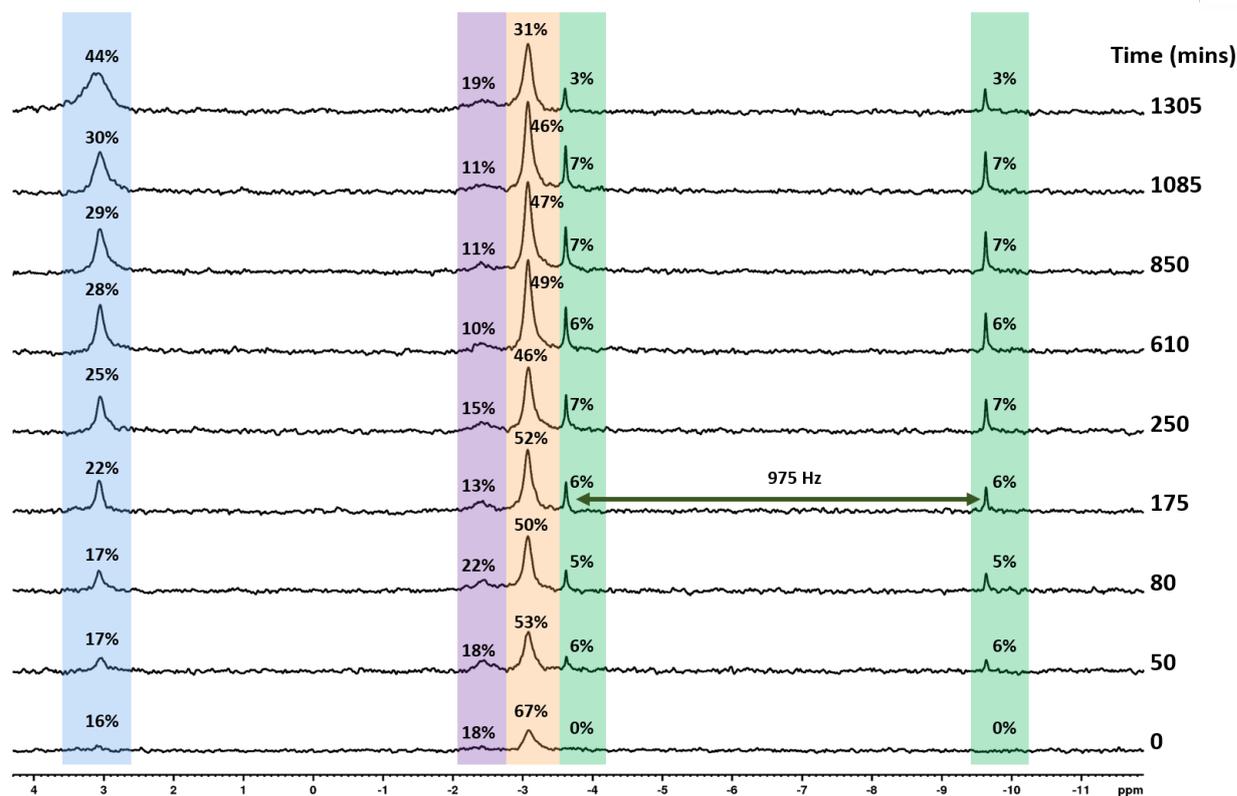
Fig. 4.13 Changes in ^{31}P$-$NMR spectra over time during the $\mathbf{H^W}$-catalysised hydrolysis of dichlorvos with peaks for (orange) dichlorvos, (blue) DMP **70**, (purple) DVMP **71**, (green) unknown species containing P-F bond.

In the first experiment, a solution of 25 mM dichlorvos guest (a concentration that allows collection of a ^{31}P$-$NMR spectra within 3 min) and $\mathbf{H^W}$ (1 mM) was prepared at pH 10. The hydrolysis of the dichlorvos and production of two products (**70** and **71**, Scheme 4.2) was monitored over a period of one day (Fig. 4.13). The signals for dichlorvos in the ^{31}P$-$NMR spectrum decrease in intensity over time synchronous with an increase in the signal for the hydrolysis product **70**. Additionally, the two peaks (highlighted in green Fig. 4.13) can be seen to initially increase before plateauing during a majority of the monitored time. After around 22 hours, these signals, and concentration of unknown species, decrease in intensity before disappearing altogether. These two peaks, at around -3.5 ppm and -9.5 ppm , both remain a comparable intensity to each other throughout and therefore are likely to be a doublet arising from a ^{31}P-^{19}F coupling with a splitting constant of $J = 975$ Hz (consistent with a single P-F bond quoted as $J = 955$ Hz in

literature for di-isopropyl fluorophosphate)[130]. This indicates an undesired reaction of dichlorvos with fluoride ions that were not expected to be present.

Hydrolysis Reaction

Possible Side Reaction

Scheme 4.2. Hydrolysis reaction of dichlorvos and possible side reaction of dichlorvos to give a P-F containing product

A ^{19}F$-$NMR spectrum of the system at the end of the catalysis experiment was obtained showing multiple signals, in contrast the single peak for the anion BF_4^- that had been expected. In particular a doublet was present at around -86 ppm with a coupling constant of 975 Hz, matching the signal seen in the ^{31}P$-$NMR spectrum (Fig. 4.14).

Fig. 4.14 ^{19}F$-$NMR of $\mathbf{H^W}$ during trial hydrolysis of dichlorvos

Attempts to fit kinetic data, using SimFit2008, were unsuccessful. The pH of the reaction solution, starting around pH 10, dropped significantly to below pH 7 by the end of the experiment, which is likely one cause of the misfitting of the experiment. Another cause of misfitting data was the unwanted P-F species. If the pH drifts during the experiment then the rate will change. Although catalysed hydrolysis of dichlorvos is clearly occurring the experiment needs further refinement.

As part of the process to make the aqueous cage solution, $\mathbf{H^W}$ was dissolved in D_2O by heating at 90° overnight. The BF_4^- anion is the only possible source of fluoride ions present in the reaction mixture which could react with the dichlorvos to form the unknown and undesired species containing the single P-F bond. The ^{11}B$-$NMR spectra and ^{19}F$-$NMR

spectra for a solution of just the $Co(BF_4)_2$ metal salt, at the same relative concentration of Co^{II} (1.6 mM) as used in the binding study, were obtained with each spectrum consisting of a single peak (in pink Fig. 4.15). Adding dichlorvos to the solutions gave no change to the resulting spectra. Instead the $Co(BF_4)_2$ was heated overnight at 90° before the $^{19}F-NMR$ spectrum were again measured (in blue Fig. 4.15). In both spectra there were additional peaks indicating that the BF_4^- anion had partly hydrolysed in the aqueous solution, and this change is not caused by the presence of the guest dichlorvos. This is the source of unwanted fluoride ions that interfered with the catalysis experiment.
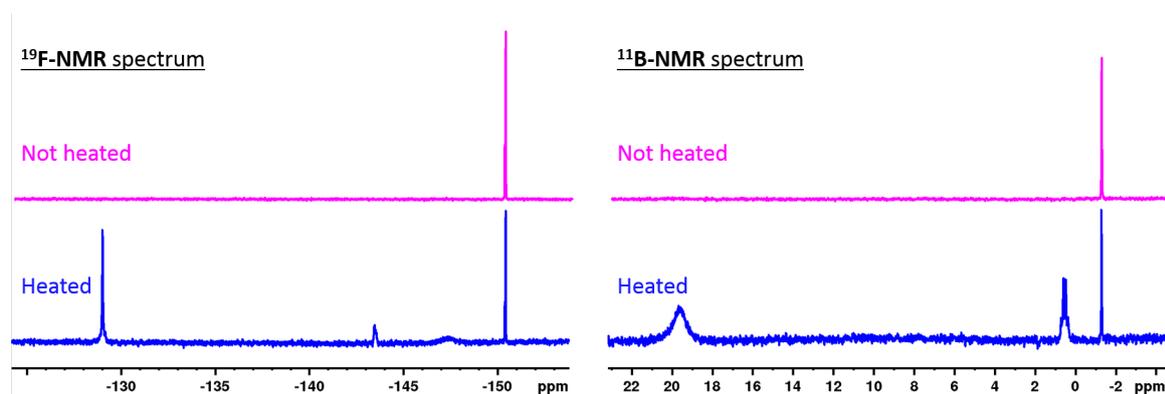
Fig. 4.15 (left) $^{19}F-NMR$ spectrum, (right) $^{11}B-NMR$ spectrum, (pink) $Co(BF_4)_2$ 1.6 mM solution, (blue) same solution heated at 90° overnight

Going forwards with the catalysis there are three issues that need resolving: 1) using a buffer that is compatible with the cage complex and can prevent the pH from drifting during the experiment as ⁻OH is consumed; 2) reducing the instrument time required for the experiment to minimise slow side-reactions; and 3) utilising a cage complex without the BF_4^- anion present to avoid the generation of F^- ions which can react with dichlorvos directly.

The cage was found to be stable in a borate-based buffer which provides the necessary weakly basic pH range appropriate for the hydrolysis of dichlorvos.

To reduce the time required on the ^{1}H-NMR spectrometer it was decided to perform the experiments at 40 °C, which is a high enough temperature to significantly accelerate the reaction.

Other recent work within the Ward group has shown that the cage $[Co_8(L^A)_{12}](Cl)_{16}$, $\mathbf{H^D}$ is water soluble and is made simply by anion exchange with Dowex ion exchange resin and $\mathbf{H^A}$ complex to replace BF_4^- anions with Cl^- ions. This allows the cage

to be water-solubilised without having to add CH_2OH groups to the external surface and removes the potential source of fluoride ions which interfered with the hydrolysis experiments.
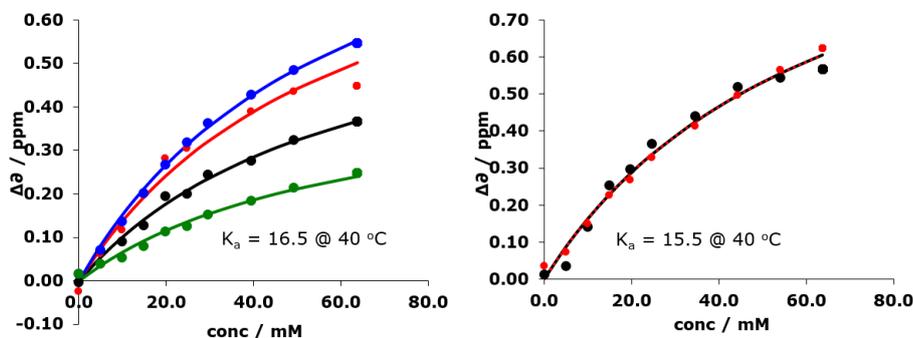
Fig. 4.16   Fitted binding isotherms from titration of dichlorvos with $\mathbf{H^D}$

The binding constant of dichlorvos with cage complex $\mathbf{H^D}$ was determined at 40 °C by ^{1}H-NMR titration at the native pH of the cage solution (pH 3 - 4). The resulting binding constant of $K_a = 16\,\mathrm{M}^{-1}$ is weaker than the previously determined binding constant due to the increased temperature and the presence of chloride ions. It has been demonstrated that the presence of Cl^- as the counter ion reduces guest binding affinity[129]. While the binding constant is small the data from NMR titration is now more reliable and fits well to a 1:1 binding isotherm.
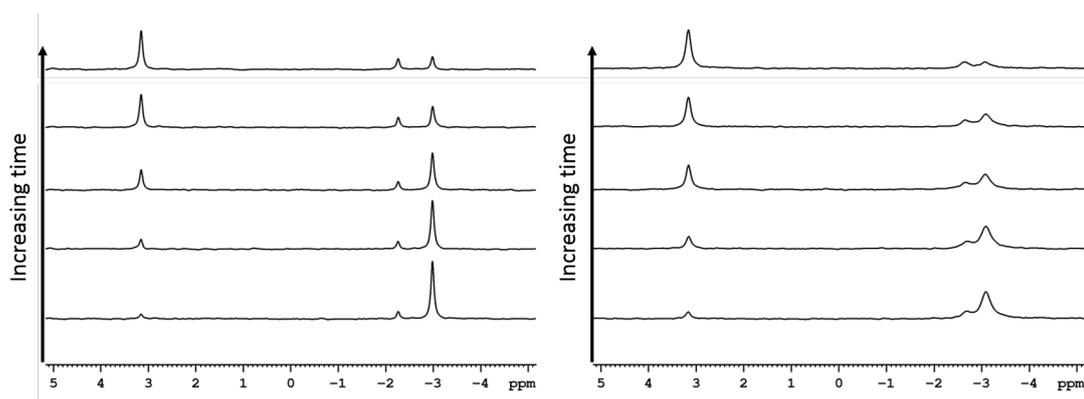
Fig. 4.17   (left) $^{31}P-NMR$ spectrum of background hydrolysis of dichlorvos, (right) of $\mathbf{H^D}$ catalysis hydrolysis

Buffered in 800 mM boric acid (hydrogen borate), both the background-hydrolysis of dichlorvos and cage-catalysed hydrolysis were carried out. The first noteworthy observation was the doublet, indicating formation of a P-F bond, was no longer present in the $^{31}P-NMR$ spectrum. For the background reaction the $^{31}P-NMR$ spectra contain

three peaks. The deconvoluted area of the peak at -3 ppm (dichlorvos) reduced over time with a corresponding increase in the intensity of the signals at +3 ppm and -2 ppm (for the major and minor products see Scheme 4.2). For the cage-catalysed hydrolysis experiment, the spectra contained the same three peaks and the signals evolved in the same way with time. However the presence of the paramagnetic cage resulted in broadening of the peaks and created difficultly in separating the deconvoluted areas of the two overlapping peaks for dichlorvos and the minor product **73** (see Fig. 4.17, right).

Table 4.1 Summary of k_{uncat} and k_{cat} at various pD

pD	7.54	8.7	10.0
k_{uncat}	5.19×10^{-6}	3.69×10^{-5}	2.63×10^{-4}
k_{cat}	2.11×10^{-3}	3.70×10^{-3}	1.29×10^{-2}

Formation of the major product was monitored and fitted at various pD values. Once fitted it was apparent that at pD 10, the cage-catalysed production of the major product was 50 times faster than the corresponding background reaction rate. This enhancement increased to 100 times at pD 8.7 and 400 times at pD 7.5 demonstrating the cage can catalyse the hydrolysis of dichlorvos efficiently. As dichlorvos was in a large excess, compared to the cage concentration, the reaction must be proceeding with effective turning over for the observed rate enhancement to occur. This is because all hydrolysis products are anionic and therefore hydrophilic and will be expelled from the hydrophobic cavity.
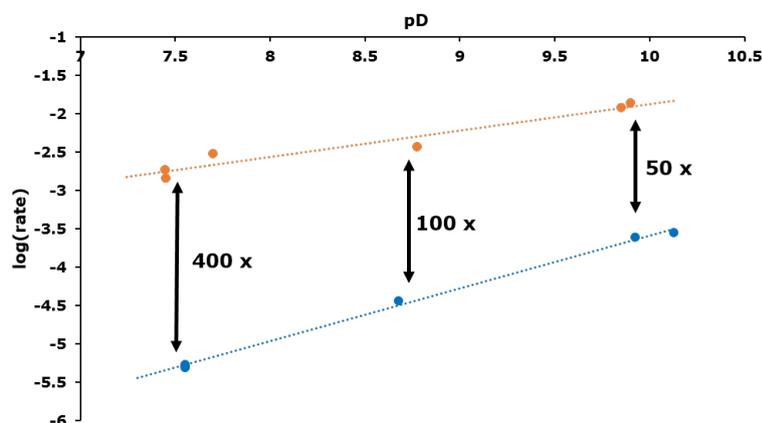
Fig. 4.18 Plot of k_{uncat} and k_{cat} at various pD values for the cage-catalysed hydrolysis of dichlorvos

The pD values used were selected to utilise the limited effective range of the borate buffer. Over this range the cage-catalysed reaction rate is only affected slightly by the

change in DO$^-$ concentration. This change is a result of competition of DO$^-$ with the Cl$^-$ anions for binding to the pockets in each of the cage windows (Fig. 4.3). The rate of hydrolysis for the background reaction drops significantly with decreasing concentration of base and therefore the difference between k_{uncat} and k_{cat} increases with lowering pD, to a factor of 400 at pD 7.5.

4.3 Conclusion

After showing that the cage is capable of catalysing the hydrolysis of dichlorvos, the conditions were optimised and cage complex modified to allow accurate measurement of the rate enhancement. At pD 7.5 the best achieved rate enhancement was $k_{uncat}/k_{cat} = 400$.

For the Kemp elimination reaction the presence of chloride ions was shown to inhibit catalysis by displacing $^-$OH from the surface sites and therefore achievable rate enhancement. Cage $\mathbf{H^D}$ was used to catalyse dichlorvos in place of $\mathbf{H^W}$ to prevent the formation of a unknown P-F species forming by the reaction of dichlorvos with traces of fluoride ions. It is likely that a cage with a non-competing anion would show a greater rate enhancement for the hydrolysis of dichlorvos. The determined rate enhancement was for the production of the major product **70**.

The powder X-ray diffraction pattern of $\mathbf{H^A}$ was obtained to see if binding of CWA simulants from the vapour phase afforded detectable changes in the PXRD pattern. An initial XRD powder pattern was obtained from dried and ground cage ($\mathbf{H^A}$) crystals but after exposure to just methanol vapour for a few hours the crystallinity of the powder was lost. Possible future studies can examine if the powder is able to bind CWA simulants from the vapour phase, by using IR spectroscopy and tracking any changes caused by binding to the host to the signal for the P=O double bond. The same IR-based method could also be used to obtain new data for the catalysed hydrolysis of the dichlorvos simulant or other related compounds.

Therefore, while the hydrolysis of dichlorvos has been shown to be significantly enhanced by the cage complex, a more comprehensive study using a different cubic cage complex with a non-inhibiting anion could lead to larger rate enhancements.

4.4 Experimental

Chemicals All chemicals were purchased from commercial sources and used as supplied unless otherwise stated.

Instruments Instruments used for spectroscopic analysis were: 1) Cary Bio for UV/Vis spectrophotometry; 2) Bruker AV3-400 for NMR spectroscopy; 3) Hamilton Spintrode pH combination electrode calibrated with standards at pH 4.01, 7.00, 10.01. $\mathbf{H^A}$ and $\mathbf{H^W}$ were prepared as previously described in Chapter 2.

Binding Constant Dichlorvos binding to $\mathbf{H^W}$ and $\mathbf{H^D}$ cages was in fast exchange by ^{1}H-NMR and so binding data could be collect using a shortened acquisition program. The data for guest binding was fitted as described before in Chapter 3.

4.4.1 Synthesis of cage complex $\mathbf{H^D}$

$\mathbf{H^A}$ cage (71.7 mg) was added into a centrifuge tube with D_2O (5 mL). Dowex 1x2 chloride form (around 10 mg) was added and stirred at room temperature overnight while ensuring that the Dowex and cage remained as an even suspension. After 12-18 h of stirring, the suspension was spun down in the centrifuge and the solution was filtered into a 10 mL volumetric flash using a membrane filter. A few mL of D_2O was added to the centrifuge tube and remaining solid was re-suspend the solids before the suspension was spinning down again and the solution was filtered into the same the volumetric flask. This was repeated until the solution in the volumetric flask is made up to 10 mL.

4.4.2 Collecting Catalysis Data

An aqueous solution of $\mathbf{H^D}$ (1 mM) was prepared in D_2O in 10 mL batches. Before the final dilution to gain the accurate 10 mL volume, the appropriate borate buffer, boric acid (375 mg, 800 mM), was added and dissolved. The accurate concentration of host cage, [H], was determined for each batch by UV spectroscopy at room temperature. A 3 mL solution of just buffer was measured as the background. 3 µL of cage solution was added, mixed

and then the absorption at 292 nm recorded. The cage solution was then calculated using the predetermined determined absorbance coefficient ($3.0766 \times 10^5 \, \text{mol}^{-1} \, \text{dm}^3 \, \text{cm}^{-1}$).

500 µL of host solution was added to an NMR tube and left, along with the pH probe, in a water bath at 40 °C. After the pH probe has been calibrated the pH of the host solution was recorded. The NMR tube was transferred to a conical flask filled with water at 40 °C for transport. Finally an addition of dichorvos (1.69 µL) was made before the tube was shaken and manually added to the NMR spectrometer (pre-set to 40°). A timer, started at the time of the addition, is referenced to the instrument's clock time such that when the timer hits 1 min 46 s, the experiment time is set to zero. Each spectrum has a time stamp at the end of the acquisition time so the adjustment time (half of the time taken to acquire a spectra) sets the start time at the midpoint of the acquisition.

4.4.3 Fitting of Catalysis Data

For the background hydrolysis of dichlorvos the three signals in the $^{31}\text{P}-$NMR spectrum were deconvoluted to give a ratio between dichlorvos, **72**, and **73** of each time. A $^{31}\text{P}-$NMR spectrum of dichlorvos in CDCl$_3$ is used for correcting initial concentrations (a small quantity of both products are present in the dichlorvos sample). The NMR ratios are converted into concentrations and then fitted using the SimFit2008 program. The coding used to fit the background reaction can be seen in Fig. 4.20 and gives rate constants for the production of **72** and **73**. The fits for the production of **72** are displayed (in blue) in Fig. 4.19 with extracted rate values in Table 4.2.

Table 4.2 Experimental data of the determined hydrolysis rate of dichlorvos

Exp Name	[H] (mM)	Background rate	[G] (mM)	pD	Rate
TC04	0.83	5.19×10^{-6}	17.8	7.4	1.87×10^{-3}
TC05	0.83	5.19×10^{-6}	17.6	7.7	3.03×10^{-3}
TC06	-	-	17.6	7.6	5.10×10^{-6}
TC07	-	-	17.6	7.6	4.99×10^{-6}
TC08	-	-	17.6	7.6	5.46×10^{-6}
TC10	0.791	1.97×10^{-5}	15.6	8.8	3.70×10^{-3}
TC17	-	-	14.0	8.7	3.69×10^{-5}
TC18	0.864	1.09×10^{-4}	13.5	9.9	1.19×10^{-2}
TC19	0.864	5.19×10^{-6}	14.9	7.5	1.44×10^{-3}
TC20	-	-	16.1	9.9	2.44×10^{-4}
TC21	-	-	16.1	10.1	2.83×10^{-4}
TC22	0.864	1.18×10^{-4}	16.1	9.9	1.39×10^{-2}

The ratio between signals in the cage catalysed reactions was also obtained through deconvolution of the $^{31}P-NMR$ spectrum. Due to the overlapping of the signals for dichlorvos and **73**, the ratio between the two could not be accurately determined. The ratio of (dichlorvos plus **73**) verse **72** was determined without issue and, after correction for starting concentrations, used to give concentration of **72** against time. This data was fitted in SimFit2008 using the coding in Fig. 4.21 along with the background rate of hydrolysis at the same corresponding pD. The fits are displayed (in orange) in Fig. 4.19 with extracted rate values in Table 4.2.

4.4.4 X-ray Crystallography

The crystal structure data collection was performed using a Bruker APEX-2 CCD diffractometer with Mo-Kα radiation from a sealed tube source. Data was corrected for absorption using empirical methods (SADABS) based upon symmetry-equivalent reflections combined with measurements at different azimuthal angles. The structure was solved and refined using a combination of different software including Olex$_2$, WinGX, Apex3, PLATON and primarily the SHELX suite version 6.14.

AFIX restraints were applied to all the rings within the cage structure. Various weak global restraints were applied to the located anions and any disorder, where possible, was modelled. The guest was then located within the cavity and DFIX restrains used to optimise the guest parameters. Other solvent molecules that could be located were also modelled before the hydrogens were calculated for cage and guest. Due to a combination of weak diffraction data, strongly absorbing crystals and general disorder some hydrogen atoms, typically on the terminal of the CH$_2$OH groups were not retained. Once the model had converged there were large regions of diffuse electron density which could not be modelled, accounting for any missing anions and solvent molecules. This electron density was removed from the final refinement using 'SQUEEZE' function in the PLATON software package. The determined structure is of poor quality compared to conventional small-molecule standards though this is typical for these types of self-assembled coordination cage complexes. The overall structure of the cage, the presence of guest, and the relative position and orientation within the cavity are clear.

While reference to relative occupancy of encapsulated guests and solvent molecule *vs.* the host complex are made it should be noted that these are approximate and therefore only used to support other data. No further claims for any specific structural details are used.

4.4.4.1 Crystallography Data Tables

Table 4.3

Complex	$[Co_8(L^A)_{12}](BF_4)_{16}$ ● Dichlorvos ● MeOH
Formula[a]	$C_{346}H_{289}B_{16}Cl_2Co_8F_{64}N_{72}O_{10}P$
Molecular Weight	7577.75
T, K	100(2)
Crystal System	Monoclinic
Space Group	C2/c
a, Å	32.8775(5)
b, Å	30.0232(5)
c, Å	39.7579(7)
α, °	90
β, °	96.3230(10)
γ, °	90
V, Å^3	39005.8(11)
Z	4
ρ, $g\,cm^{-1}$	1.290
Crystal Size, mm^3	0.20 x 0.20 x 0.10
Data, restraints, parameters	23750, 1774, 1891
Final R1, wR2[b]	0.1919, 0.4740

a   These formulae (and consequently the crystal densities) are necessarily approximate given that large amounts of diffuse electron density in solvent-accessible voids were removed from the refinements using the 'SQUEEZE' function in PLATON. See CIFs and comments below for more details.

b   The value of R1 is based on 'observed' data with $I > 2\sigma(I)$; the value of wR2 is based on all data.
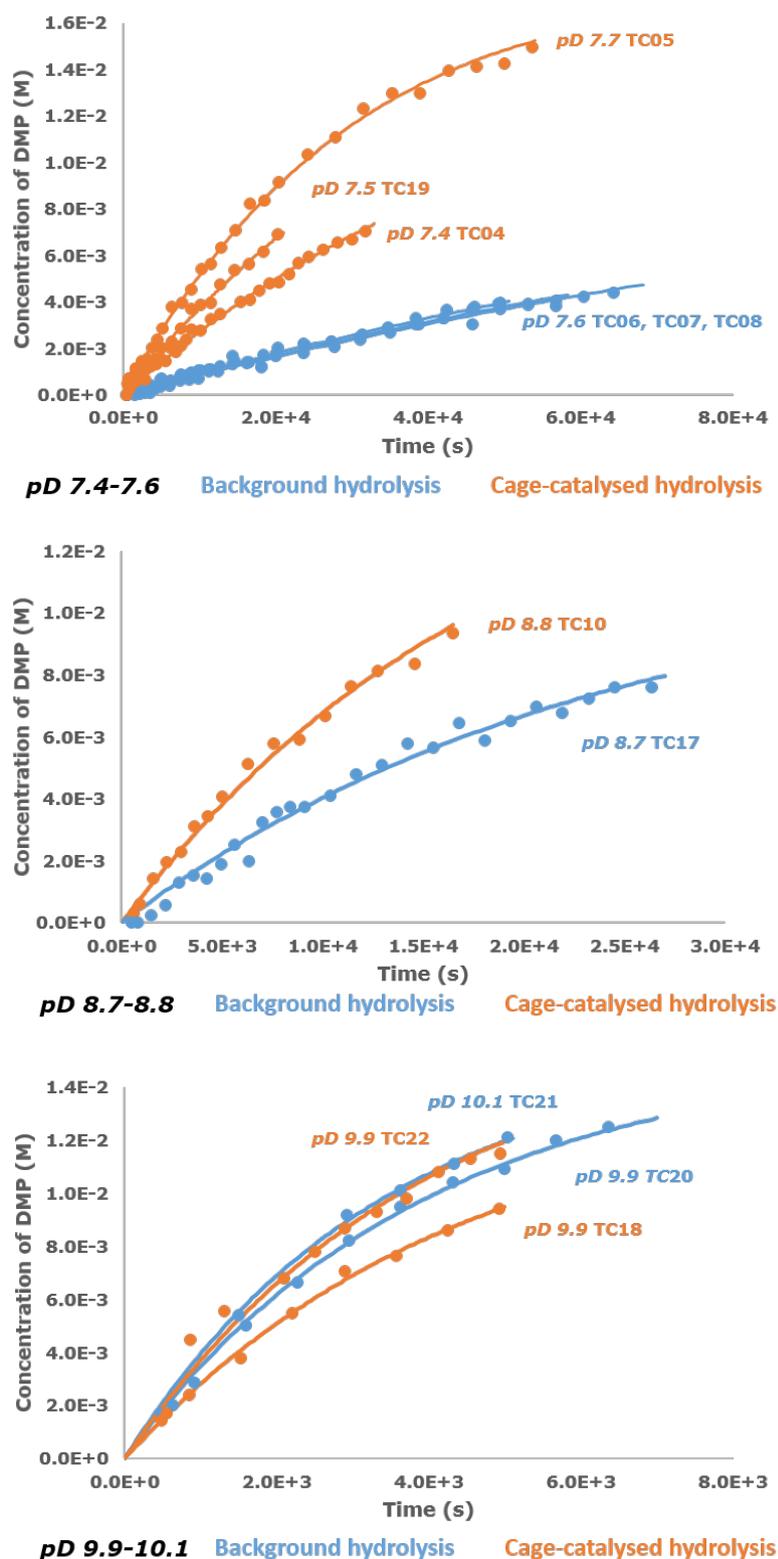
Fig. 4.19 Hydrolysis data (circles) and SimFit2008 fitted curve (line). (blue) background reaction, (orange) catalysed reaction

```
*===========================================================
* Chris Taylor
*
* A going to B and D (Background Fit)
* Background reaction   K1 = 1.956E-4
*
*===========================================================

*Background reaction

REACTION (A              --> B) CONSTANT ( 1, 9E-4, 1, 1, 10000)

REACTION (A              --> D) CONSTANT ( 2, 4E-7, 2, 1, 10000)


REACTION (COMPILE)

REACTION (SHOW)
CONSTANT (SHOW)

DEFINE (1, B, P, 1) SCALE (3,1)
DEFINE (2, D, I, 2) SCALE (3,1)

SELECT (B, D)

READ (tc06)
REACTION (DOC)
CONSTANT (DOC)

TIME (SEC)
CONC (M)
*Scales need adjusting WIN(0,Max Time, Maj, Min, 0, Max Conc, Maj, Min)
WIN (0, 110000, 3600, 360, 0, 1.8e-2, 0.5e-3, 1e-4)

DIM (2)
ASSIGN (OBS, B   = B )
ASSIGN (OBS, D   = D )
*Conc of dicholvos total
ASSIGN (SPEC, A = #1.761e-2 )

CHOOSE (EXP1)

INTEG (STIFF, 1E-9, 100, 0.075, 100, 50)

PLOT (OBS, RES)

* Optimise rate constants using simplex

SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)

PLOT (OBS, RES)
PLOT (FILE)
```

Fig. 4.20 Coding used in SimFit2008 for fitting of the background hydrolysis reaction of dichlorvos

```
*============================================================
* Chris Taylor
* Boric Buffered Cage CAT of dichlorvos
*
* Update Bkgd Rate, Koff, Scale, File Name, Conc Dichlorvos, Conc Cage
*============================================================

*Background reaction

REACTION (A              --> B) CONSTANT ( 1, 4.37E-6, 0)

* Assocation of Cage and A (Ka = Kon/Koff, Kon = 1E9) calc Koff (Ka = 16 from two exp)

REACTION (A + C          ==> AC) CONSTANT ( 2, 1E9, 0) CONSTANT (3, 6.25E7, 0)

*REACTION OF A IN CAGE

REACTION (AC             --> BC) CONSTANT ( 4, 1.956E-2, 1, 1, 10000)

* B CAGE EQUILIBRIUM

REACTION (B + C          ==> BC) CONSTANT ( 5, 1E9, 0) CONSTANT (6, 5E12, 0)


REACTION (COMPILE)

REACTION (SHOW)
CONSTANT (SHOW)

DEFINE (1, B, P, 1) SCALE (3,1)


SELECT (B)

READ (File Name)
REACTION (DOC)
CONSTANT (DOC)

TIME (SEC)
CONC (M)
* Scale (0, Max Time, mj, mi, 0, Max Conc, mj, mi)
WIN (0, 54000, 100, 20, 0, 1.6E-2, 0.1E-2, 1e-4)

DIM (1)

ASSIGN (OBS, B   = B + BC )
* Conc Dichlorvos
ASSIGN (SPEC, A = #1.7E-2 )
* Conc Cage
ASSIGN (SPEC, C = #0.76E-3 )

CHOOSE (EXP1)

INTEG (STIFF, 1E-9, 100, 0.075, 100, 50)

PLOT (OBS, RES)

* Optimise rate constants using simplex

SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)

PLOT (OBS, RES)
PLOT (FILE)
```

Fig. 4.21   Coding used in SimFit2008 for fitting of the cage catalysed hydrolysis reaction of dichlorvos

Chapter 5

Predicting Guest Binding
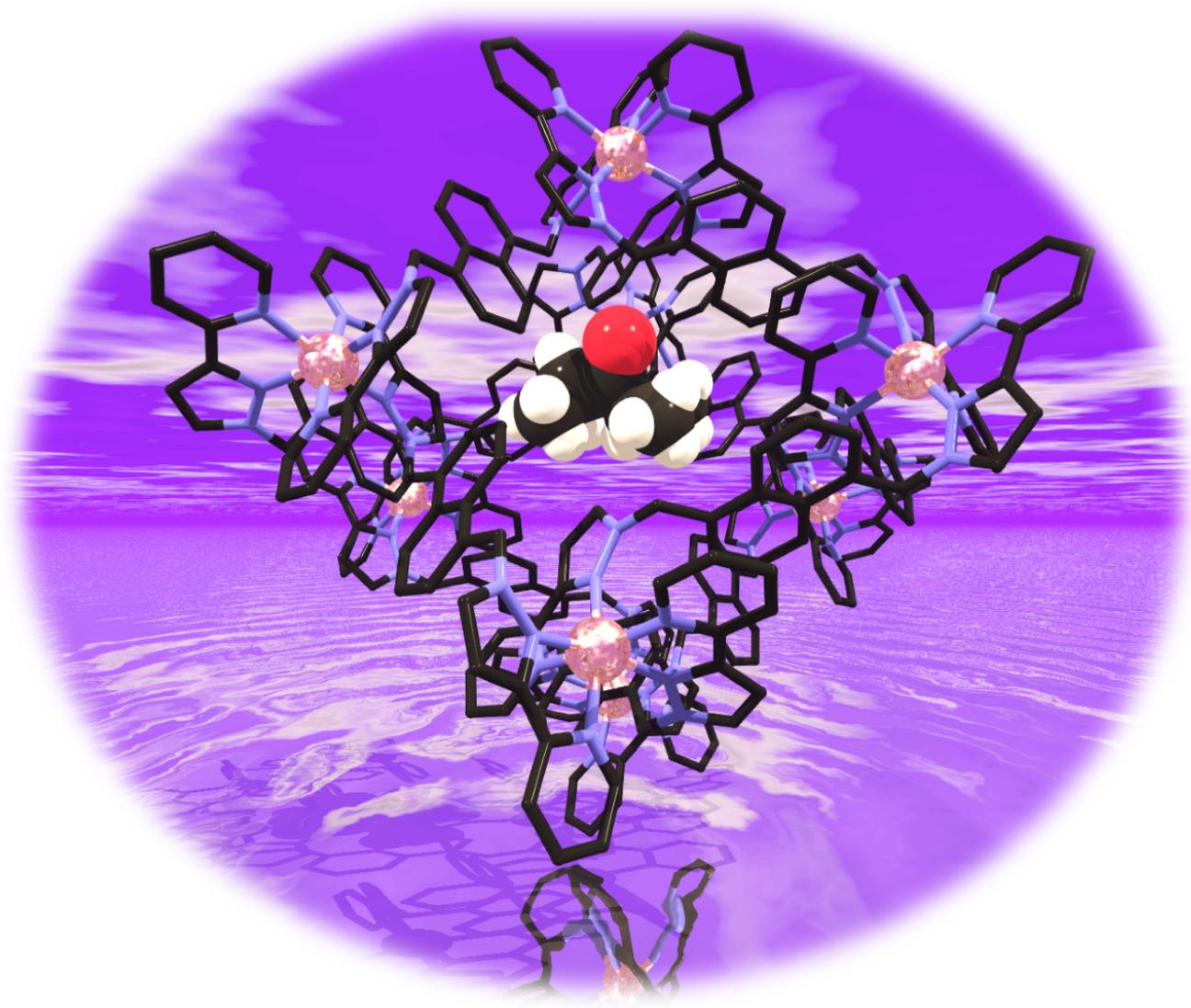
Rendering of crystal structure for **H**$^{\textbf{A}}$ containing 2,4-dimethyl-3-pentanone

5.1 Introduction

Applications of cage complexes continue to develop by exploiting the ability of these hollow containers to selectively bind guests of different shapes and sizes. Once a guest is bound within the central cavity of a cage complex its reactivity may change; it may trigger a response for sensing purposes; or it may be susceptible to stimulus-responsive binding, for example it can be used for drug delivery. All of these applications require a good understanding of which guest will bind in which cage cavity, why they will bind and what the strength of binding would be.

Beyond the obvious effect of guest shape and size, which need to match the cavity parameters, there are more subtle contributions to binding including electronic interactions between the cage and guests, desolvation of both the cage and guests, and conformational changes associated with the folding of flexible guests to allow them to fit with the cage cavity. In nearly all reported cases of guest binding in synthetic containers, empirical approaches were used to identify appropriate guests, often just through a trial and error approach.

A more systematic approach to this problem, attempting to quantify the various intermolecular interactions that contribute to guest binding, will allow a degree of prediction in identifying new guests. While the quantitative prediction of interactions is well established within the pharmaceutical industry[131] for identifying new drugs of interest, there are very limited examples of quantitative prediction of guest binding in synthetic cage complexes[82,128,132,133].

The Ward group has previously reported a virtual screening method for identifying new guests for binding within the cavity of $\mathbf{H^W}$[128]. The protein docking program GOLD (Genetic Optimisation of Ligand Docking) was used with an initial training set of 54 guests (black molecules Fig. 5.1) that had previously had their binding constants determined experimentally. This initial set of data allowed a scoring function to be

Training Set Guest (Binding)

Training Set Guests (Non-Binding)

Newly Identified Guests Based on Prediction from GOLD

CWA Simulants and CWAs

DMMP   DEMP   DEEP   DIMP   Tabun   Sarin   Cyclosarin   Somen
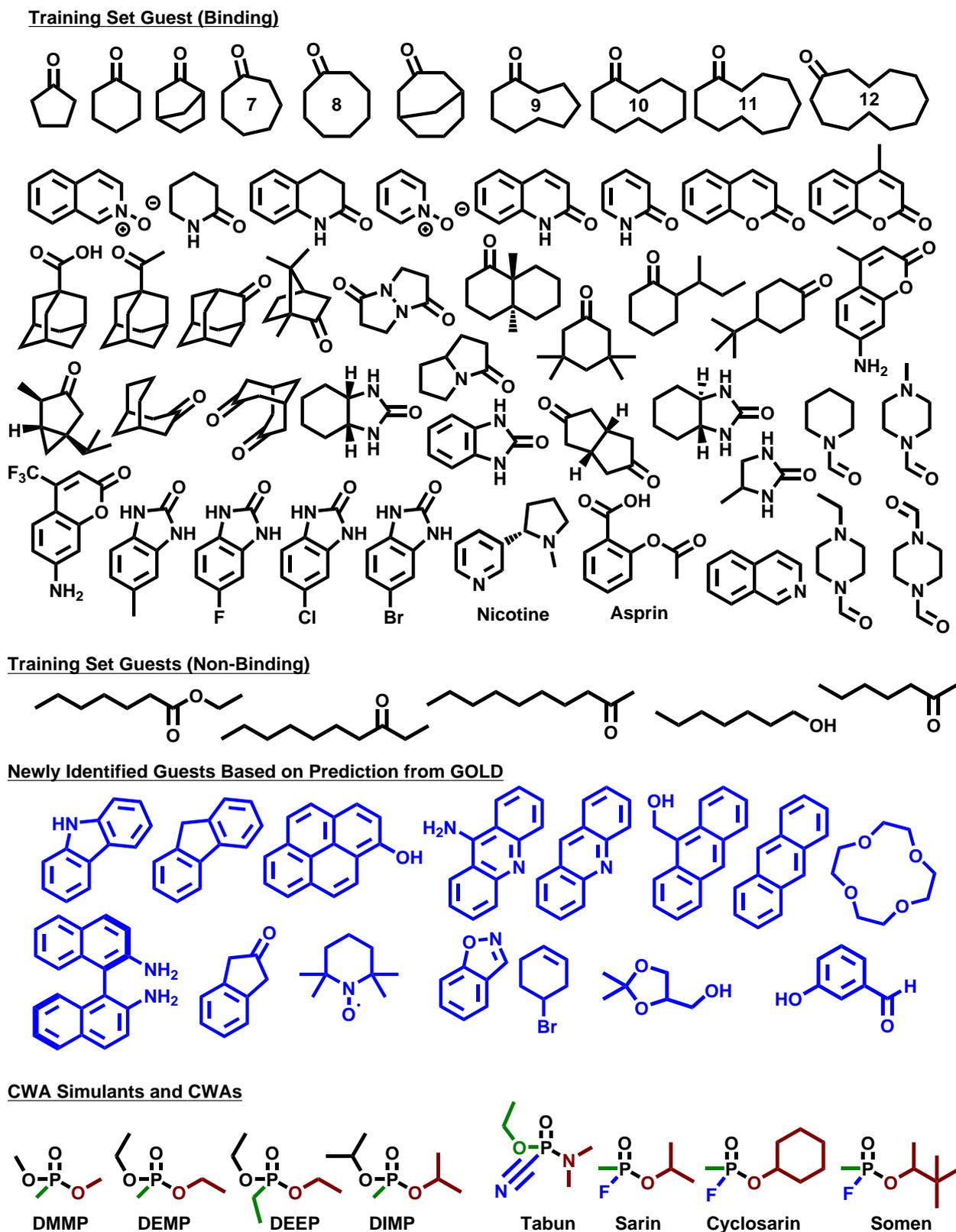
Fig. 5.1 (black)[128] Training set of guest, (blue) new guests identified from screening a virtual library, (bottom-left) alkylphosphonates, (bottom-right) chemical warfare agents

determined that predict binding of rigid guests. Applying the scoring function to predict binding properties of a virtual library of guests, 15 new potential guests (blue molecules in Fig. 5.1) were identified.

$$\log K_{calc} = - (4.8 \times f_{ligandclash}) + (0.20 \times f_{partburied}) - (0.10 \times f_{nonpolar})$$
$$+ (0.90 \times f_{ligandtorsion}) - (0.93 \times f_{ligandflexibility}) \tag{5.1}$$

GOLD has an inbuilt scoring function called CHEMPLP[134] from which the first four terms above are taken. The fifth term, $f_{ligandflexibility}$, is a custom term added to take account of the entropic cost of folding guest to fit within the cavity and is based on the calculated number of rotamers within the guest. Any term can be added to the scoring function, including something random and unconnected (for example the phase of the moon), but when refined against real data the weighting of irrelevant terms tends to zero as these do not correlate with guest binding. More details about each term and refinement to give the scoring function (5.1) can be found in the original publication[128].

- $f_{ligandclash}$ Steric clashes between the host and guest
- $f_{partburied}$ The burial of a polar group in a non-polar environment
- $f_{nonpolar}$ Hydrophobic interactions
- $f_{ligandtorsion}$ The torsional strain induced in the ligand (guest) on binding (enthalpy)
- $f_{ligandflexibility}$ - Conformational changes to guest to allow binding (entropy)

5.2 Discussion

The recent work, discussed in the two preceding chapters, focused on binding, detection and catalysed hydrolysis of CWA simulants or a chemically related pesticide. Therefore being able to accurately predict the binding of CWA simulants and other related guests should also allow for a prediction of the active nerve agents' binding affinity. When the alkylphosphonate guests were first screened through GOLD the calculated binding strengths were found to disagree significantly with the experimental values obtained despite the fact that the scoring function in equation (5.1) could be used to predict binding of rigid guest very well. The disagreement between predicted and calculated binding

constants increases as the alkyl chain length increased. Unsurprisingly the flexibility of the guests is a key parameter in the strength of guest binding, with relatively rigid cyclic aliphatic ketones binding strongly compared to the open-chain analogues whose binding was very weak or was not detectable, suggesting that the initial scoring function (5.1) does not properly account for the unfavourable entropy cost associated with the loss of guest flexibility on binding.
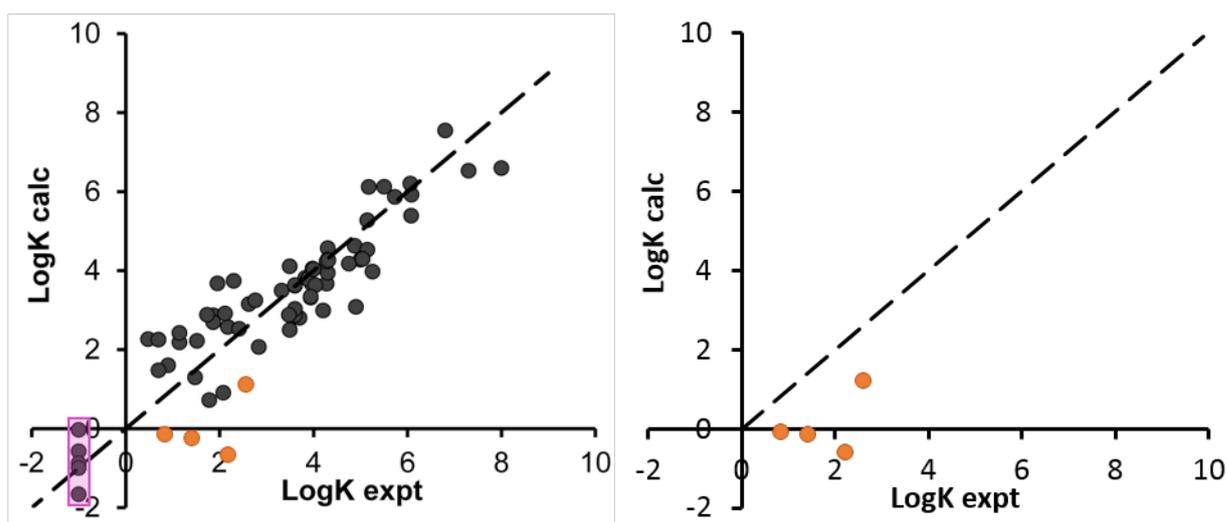
Fig. 5.2 Plot showing calculated *vs.* measured binding constants for the original training set and identified guests (black) along with the new flexible alkyl phosphonates (orange), the non-binding guests from the training set are highlighted by a pink box, (right) same plot showing just the new guests

All of the strongly binding guests previously found through both the empirical studies and those found through the virtual screening are relatively rigid and typically contain at least one ring system. The only highly flexible guests within the training set (Training Set Guests (Non-Binding) Fig. 5.1) were found to be very weakly binding and so were set to $\log K_{calc} = -1$ (highlighted in pink box Fig. 5.2).

5.2.1 Choice of Guests

A series of simple alkyl ketones were evaluated to determine the effect that guest flexibility has on guest binding. A set of heptanone isomers, and a set of nonanone isomers, were chosen for several reasons. Firstly, aliphatic ketones are known to bind strongly in the cage cavity in water, with the guest hydrophobic surface area providing the major contribution to binding, as demonstrated with cyclic ketones[82]. Secondly the polar carbonyl group means the ketones are sufficiently water soluble to allow the binding
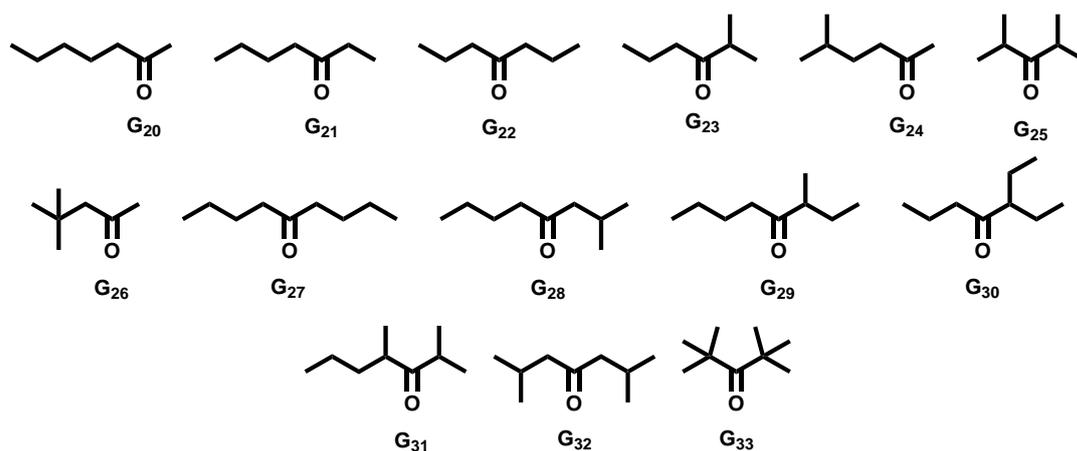
Fig. 5.3 Selection of flexible and branched C_7 and C_9 ketones

constants to be measured, while also providing a hydrogen bond acceptor such that the guest will anchor into the hydrogen bonding pocket of the cage cavity. Thirdly, these $\mathbf{C_7}$ and $\mathbf{C_9}$ ketones are readily available as a number of different isomers with a variation of branching of the alkyl chain, and so provide the desired structural variation while keeping similar molecular volumes, electronic properties, and hydrophobic surface areas. Lastly, chosen guests have some similarity to the alkylphosphonate guests with a selection of branched alkyl substituents around a polar oxygen double bond.

The surface areas, volumes, and the experimentally determined binding constants of these guests with $\mathbf{H^W}$ are displayed in Table 5.1. All of the new ketone guests are below the optimal guest volume of $220\,\text{Å}^3$, as described by Rebek's 55 % rule[85], and this has been confirmed by previous work within the Ward group[82]. The table also includes the number of rotatable bonds (NRB) for each guest; this is used as part of the prediction of scoring within the introduced term $f_{ligandflexibility}$ and is defined as the number of bonds which allow free rotation around themselves (single bonds that are not part of a ring, do not involve hydrogens, and have a non-terminal atom at each end)[131,135,136].

5.2.2 Crystal Structures of Host-Guest Complexes

Some crystal structures of the cages containing different guests were investigated to give insight into the binding mode of the ketones. Four structures containing different heptanone isomers were obtained. Good quality data for the nonanone guests could not be obtained; this is likely due to the increased length of the alkyl chains within the guest and resulting disorder within the structure. While the central section of the guest around

Table 5.1 Structural data, and binding data for investigated guests including the aliphatic ketone series and alkylphosphonates guest series

Guest	NRB [a]	Surface area [Å^2]	Volume [Å^3]	K [M^{-1}]	$-\Delta G^o$ [b] [kJ mol^{-1}]
G_{20}	4	174.5	146.0	59(36)	10.1(13)
G_{21}	4	174.2	145.9	72(37)	10.6(14)
G_{22}	4	174.5	146.0	210(160)	13.2(17)
G_{23}	3	173.4	145.8	275(160)	13.9(17)
G_{24}	3	174.5	146.4	204(93)	13.2(12)
G_{25}	2	172.8	145.5	625(300)	16.0(12)
G_{26}	2	169.7	145.6	699(300)	16.2(10)
G_{27}	6	213.8	182.5	863(400)	16.7(10)
G_{28}	5	215.0	183.2	2040(670)	18.9(9)
G_{29}	5	212.8	183.0	1590(540)	18.3(9)
G_{30}	5	213.4	183.2	2620(860)	19.5(9)
G_{31}	4	207.6	182.3	2440(730)	19.3(8)
G_{32}	4	210.2	182.3	3890(67)	20.5(1)
G_{33}	2	199.4	180.8	12100(2000)	23.3(5)
DMMP	2	151.0	118.6	7(2)	4.8(6)
DEMP	4	192.4	155.5	26(23)	8.1(17)
DEEP	5	212.2	174.0	160(45)	12.6(7)
DIMP	4	230.2	192.1	390(15)	14.8(1)

[a] NRB = number of rotatable bonds. [b] Each titration was repeated at least three times and the experimental error is quoted as twice the standard deviation

the carbonyl group is ordered and could be found within the structures, the remainder of the guest molecule for the nonanone isomers could not be resolved satisfactorily due to disorder of the alkyl chains.

For each of the structures obtained, the crystalline complexes were prepared by soaking preformed crystals of $\mathbf{H^A}$ in a methanol solution of the guest. $\mathbf{H^A}$ was used in place of the $\mathbf{H^W}$ as the crystals are easier to prepare, and the crystals give better diffraction. The two cage types are identical except for the hydroxymethyl substituents on the exterior of $\mathbf{H^W}$. Previous structures (described in Chapter 3.2.2) have showed that the guest binding does not differ in the solid-state between the two cages and so the obtained structures of guest with $\mathbf{H^A}$ can be used in confidence as models for how guests bind in the cavity of $\mathbf{H^W}$.

For each X-ray crystal structure of $\mathbf{H^A \bullet G_{21}}$, $\mathbf{H^A \bullet G_{22}}$, $\mathbf{H^A \bullet G_{23}}$, and $\mathbf{H^A \bullet G_{24}}$ it is clear that the guests are oriented within the cavity of the cage via a hydrogen bonding between the electron rich carbonyl group of the guest, and the electron deficient set of
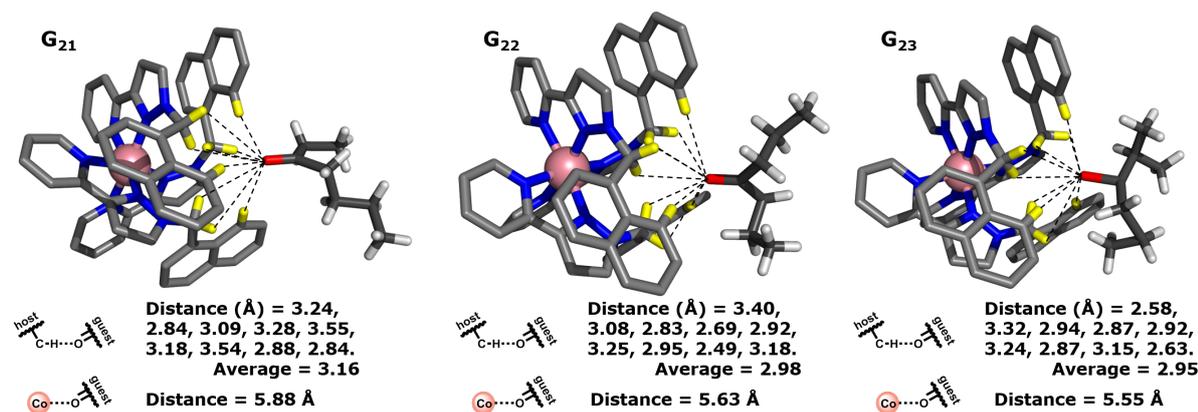
Fig. 5.4 Cutaway structure showing the *fac* vertex, bound guests and relative interactions for (left) **G₂₁**, (middle) **G₂₂**, (right) **G₂₃**

convergent C-H protons in each of the two *fac* tris-chelate vertices. These nine contacts (highlighted in yellow Fig. 5.4 and Fig. 5.5) have an average CH···O distance of around 3 Å and non-bonded Co···O distance of 5.6 Å. Along with one guest molecule, each cavity also contains a methanol solvent which is located at the opposite *fac* vertex. For two guests to occupying the cavity they must both be small [for example DMMP (Chapter 3.2.1.3)] but the ketone molecules are too large for two to bind in the cavity allowing the second *fac* vertex to be occupied by a solvent molecule.
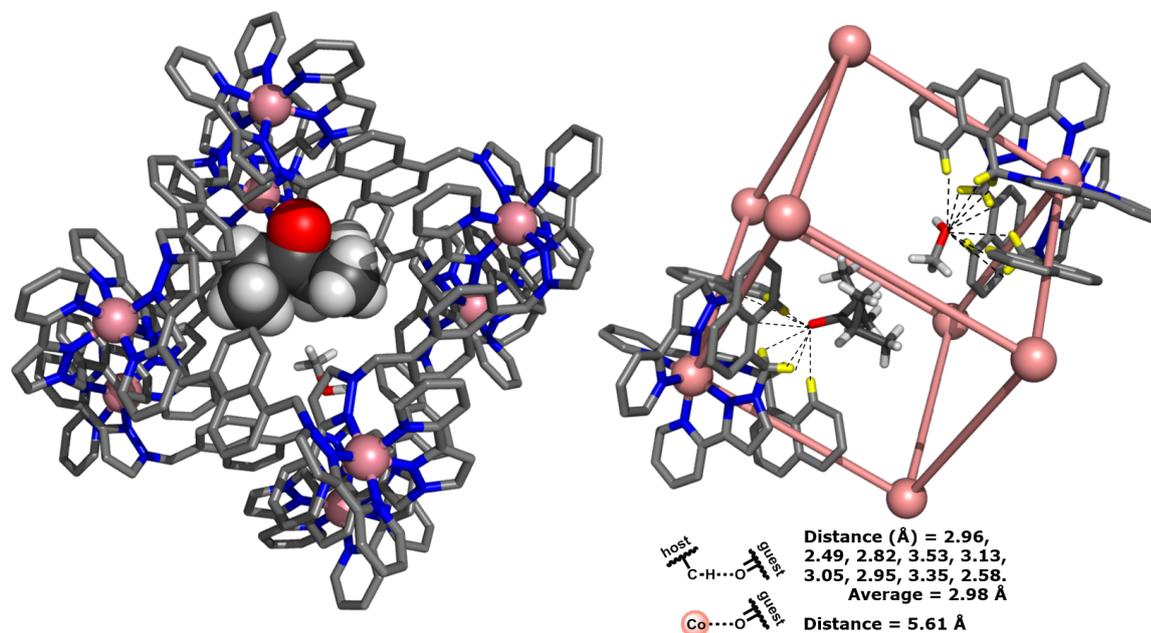
Fig. 5.5 Structure of H^A containing a **G₂₅** guest molecule and a methanol solvent molecule; (left) stick model, (right) cut away model showing the bound guest and solvent interactions with the *fac* vertex
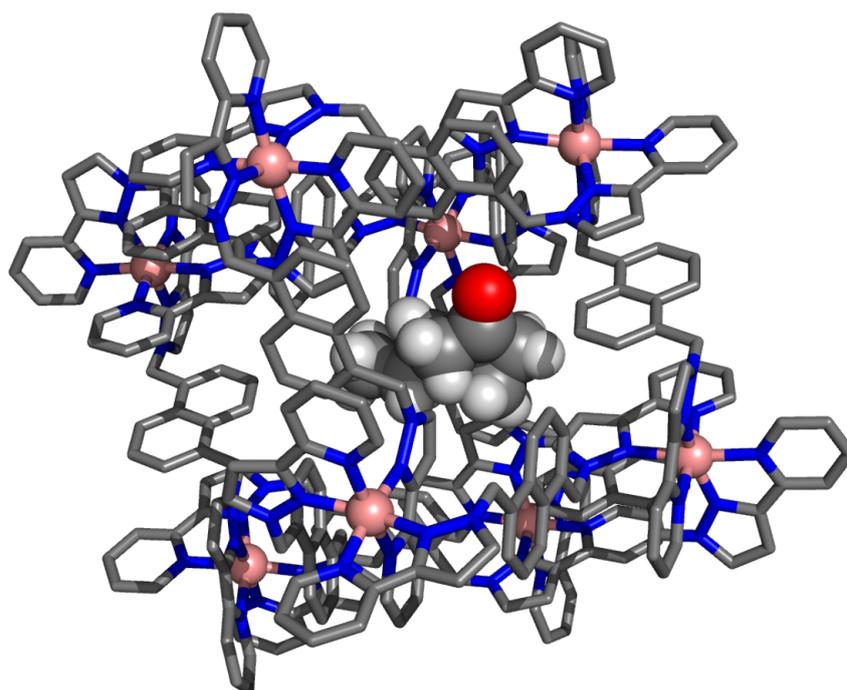
Fig. 5.6 Structure of $\mathbf{H^A}$ containing a $\mathbf{G_{21}}$ guest molecule
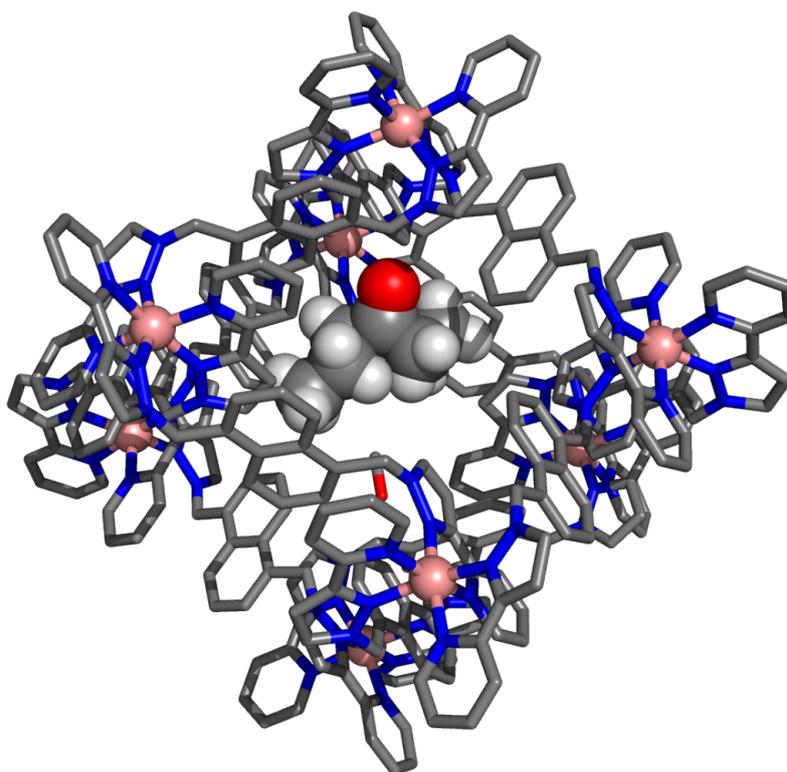
Fig. 5.7 Structure of $\mathbf{H^A}$ containing a $\mathbf{G_{22}}$ guest molecule
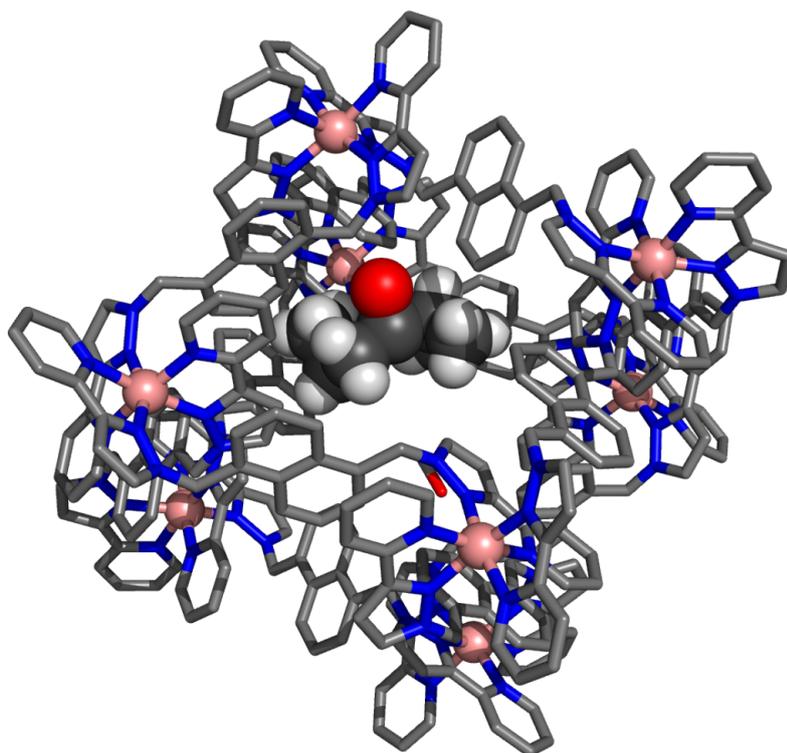
Fig. 5.8 Structure of $\mathbf{H^A}$ containing a $\mathbf{G_{23}}$ guest molecule

5.2.3   Solution Binding Properties

The binding constants for all of the ketone guests were measured in D_2O following the previously described NMR titration method; results are shown in Table 5.1. Unsurprisingly there is a clear difference between the two series of guests ($\mathbf{C_7}$ and $\mathbf{C_9}$ series), with the $\mathbf{C_9}$ guests generally binding more strongly than the $\mathbf{C_7}$ ketones. This is simply due to the greater hydrophobic surface area (two additional CH_2 groups).

In previous work by the Ward group (Fig. 1.59) looking at a series of cyclic ketones it was found that each additional CH_2 group contributed an additional 4–$5\,\mathrm{kJ\,mol^{-1}}$ of binding free energy when binding in $\mathbf{H^W}$ in water[82]. The binding constants for cycloheptanone and cyclononanone were $4.2(4) \times 10^2\,\mathrm{M^{-1}}$ and $1.2(3) \times 10^4\,\mathrm{M^{-1}}$ respectively, affording binding free energies of $15\,\mathrm{kJ\,mol^{-1}}$ and $23\,\mathrm{kJ\,mol^{-1}}$ respectively[82]. There is a spread of values within each of the open chain ketone series associated with the different substitution patterns and hence degrees of flexibility. If we compare between the most rigid $\mathbf{C_7}$ guest 2,4-dimethyl pentan-3-one $\mathbf{G_{25}}$, and $\mathbf{C_9}$ guest 1,1,1,3,3,3-hexamethylacetone $\mathbf{G_{33}}$, both of which have C_2 symmetry and the same degree of flexibility with NRB = 2, their respective binding parameters ($K = 6.3 \times 10^2\,\mathrm{M^{-1}}$ and

$1.2 \times 10^4 \, \text{M}^{-1}$ respectively; $-\Delta G$ 16 kJ mol^{-1} and 23 kJ mol^{-1} respectively) match well with the values found for their cyclic analogues.

Table 5.2 Average binding free energies of isomers with the same 'number of rotatable bonds' (NRB), with the step changes between them

Guest	Number of C atoms	NRB	$-\Delta G$ (average) [kJ mol^{-1}]	$-\Delta G$ increment per NRB [kJ mol^{-1}]
G_{25} G_{26}	7	2	16.1	-
G_{23} G_{24}	7	3	13.5	2.6
G_{20} G_{21} G_{22}	7	4	11.3	2.2
G_{33}	9	2	23.3	-
G_{31} G_{32}	9	4	19.9	1.7
G_{28} G_{29}	9	5	18.6	1.3

There is a clear inverse correlation between the NRB and the strength of guest binding. For example it is clear that, within the C_7 series, the most highly substituted guests (G_{25} G_{26}), which are therefore the most rigid with the smallest NRB values, have the strongest binding. In contrast, the most open-chain ketones within the series (G_{20} G_{21} G_{22}) have the weakest binding. The average of the binding free energy for a set of isomer guests with the same NRB value are given in Table 5.2. There is an incremental increase of 2.6 kJ mol^{-1} and 2.2 kJ mol^{-1} between NRB values of 2 to 3 and 3 to 4 respectively for the C_7 series. The C_9 series shows the same relationship between $-\Delta G$ and NRB, with increments of 1.7 kJ mol^{-1} and 1.3 kJ mol^{-1} for each increase of one in the NRB value.

While there is a given experimental uncertainty with each determined $-\Delta G$, the average incremental decrease for each additional NRB is 2 kJ mol^{-1}. This is nicely comparable to a penalty of 1.5–4 kJ mol^{-1} per rotatable bond that has been calculated for binding of many different substrates to proteins[131,135,136].

Comparison between the most rigid C_7 guest G_{25} with the most flexible C_9 guest G_{27} reveals that the binding free energies are similar. Guest G_{27} has two additional CH$_2$ groups but 4 additional NRB. Interestingly, the addition of 9.4 kJ mol^{-1} from the extra hydrophobic contribution due to the two extra CH$_2$ groups[82] is cancelled out by the increased entropic penalty of 4 x -2 kJ mol^{-1} associated with increased flexibility. This agrees with the generalisation made in a recent review that the loss of configurational entropy upon binding of substrates to proteins is nearly as large as the gain in favourable binding energy from extra hydrophobic surface area and the two approximately equal
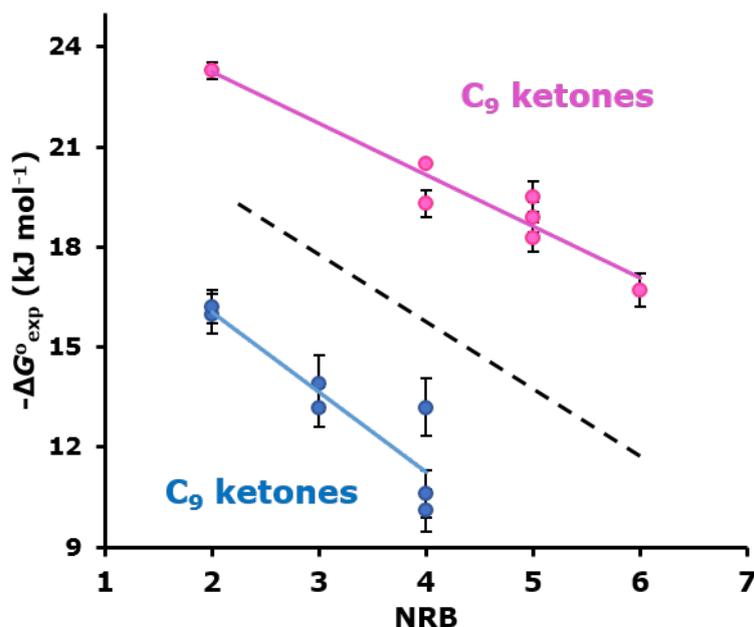
Fig. 5.9 Plot of the binding free energy of the guests against the 'number of rotatable bonds' (NRB). (blue) C_7 guests, (pink) C_9 guests, (black dotted) a line with a gradient equivalent to a decrease of $2\,\mathrm{kJ\,mol^{-1}}$ per NRB

and opposite contributions tend to correlate[131].

$$-\Delta G_{crude} = -13 + (4.7 \times N_c) - (2 \times \mathrm{NRB}) \tag{5.2}$$

Crudely, the binding constants can be calculated using just three terms in equation 5.2. The penalty for removing a carbonyl from the bulk water, and the entropic price for the formation of a bimolecular complexation, gives the unfavourable contribution of $-13\,\mathrm{kJ\,mol^{-1}}$ ($-7\,\mathrm{kJ\,mol^{-1}}$ and $-6\,\mathrm{kJ\,mol^{-1}}$ respectively)[82]. Starting from NRB = 2, each increment adds a further unfavourable contribution of $-2\,\mathrm{kJ\,mol^{-1}}$. Each CH_2 group contributes $4.7\,\mathrm{kJ\,mol^{-1}}$ due to the increase in hydrophobic surface area. Here the assumption was made that the number of CH_2 is equal to the number of C (N_c) to simplify the crude calculation.

A final observation within each set of unbranched ketones is that the most symmetrical guests (4-heptanone **G$_{22}$**, 5-nonanone **G$_{27}$**) bind significantly stronger than others despite having the same NRB.

Table 5.3 Crude calculation of binding constants with $\mathbf{H^W}$ cage complex: all numbers have units of $\mathrm{kJ\,mol^{-1}}$

Guest	G_{20}	G_{21}	G_{22}	G_{23}	G_{24}	G_{25}	G_{26}	G_{27}	G_{28}	G_{29}	G_{30}	G_{31}	G_{32}	G_{33}
	-13	-13	-13	-13	-13	-13	-13	-13	-13	-13	-13	-13	-13	-13
2 x (NRB - 2)	-8	-8	-8	-6	-6	-4	-4	-12	-10	-10	-10	-8	-8	-4
4.7 x N_c	32.9	32.9	32.9	32.9	32.9	32.9	32.9	42.3	42.3	42.3	42.3	42.3	42.3	42.3
$-\Delta G^o_{crude}$	11.9	11.9	11.9	13.9	13.9	15.9	15.9	17.3	19.3	19.3	19.3	21.3	21.3	25.3
$-\Delta G^o_{exp}$	10.1	10.6	13.2	13.9	13.2	16	16.2	16.7	18.9	18.3	19.5	19.3	20.5	23.3
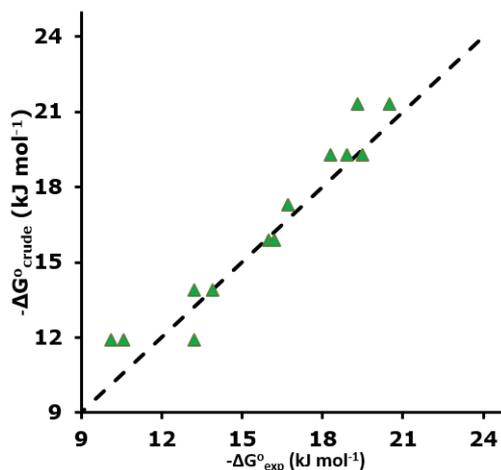
Fig. 5.10 Plot of crudely calculated (Equation 5.2) *vs.* measured binding constants for the series of aliphatic ketones

5.2.4 Improving Guest Binding Prediction

Of the original training set of 54 guests and the additional 15 identified guests, only five guests were flexible and had been defined as non-binding. Initially these five non-binding guests (pink box Fig. 5.2) were removed from the analysis and then the remaining 64 guests, from the original publication[128], were used to recalculate the weighting constants in the scoring function giving equation 5.3. Neither the organo-phosphonates nor ketones were included for the refinement of this scoring function. Following a small improvement in the prediction of binding for the organo-phosphonates, along with no significant change to the prediction of binding for the guests used to generate the scoring function, it was evident that the five non-binding guests were detrimental to the calculation so they were removed from the preceding training sets.

Fig. 5.11 shows the correlation between the predicted (using equation 5.3) and measured binding constants. The 64 training guests (black) on which 5.3 is based present an even spread along the line of agreement. The new flexible ketone guests ($\mathbf{C_7}$ in blue,
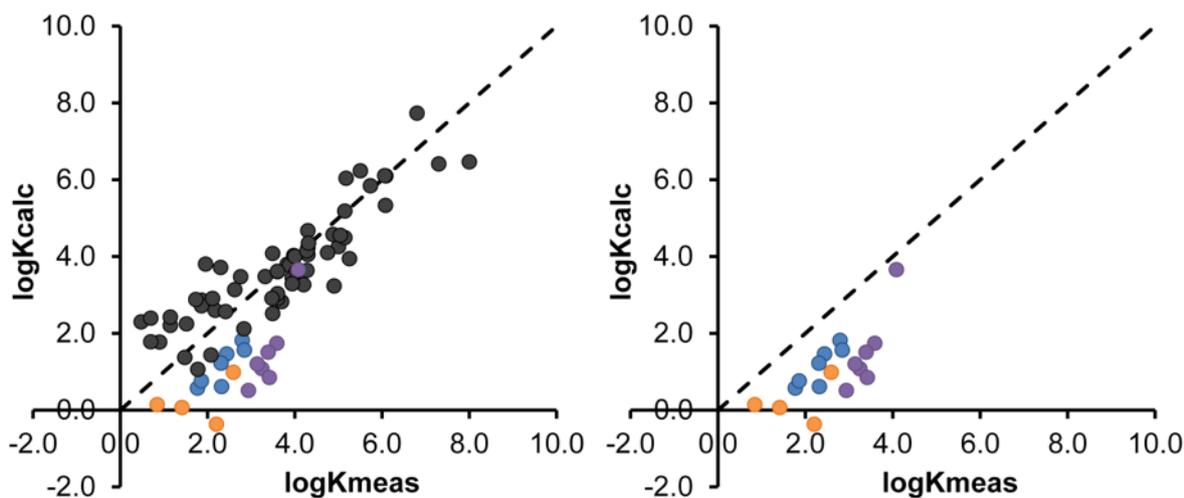
Fig. 5.11 Plot showing calculated (Equation 5.3) *vs.* measured binding constants for the original training set and identified guests (black) along with the new flexible guests; alkyl phosphonates (orange), **C$_7$** (blue), **C$_9$** (purple), (right) same plot showing just the new guests

C$_9$ in purple) have also been added. It is apparent that the calculated binding constants for all of the flexible guests have been significantly and systematically underestimated by both the original scoring function (5.1) and the revised scoring function (5.3).

$$\log K_{calc} = - (4.8 \times f_{ligandclash}) + (0.22 \times f_{partburied}) - (0.10 \times f_{nonpolar})$$
$$+ (0.32 \times f_{ligandtorsion}) - (0.76 \times f_{ligandflexibility}) \tag{5.3}$$

The enthalpic and entropic effects associated with the loss of conformational flexibility upon binding of the guest are not properly accounted in the current scoring function (Equation 5.3). Interestingly the more rigid the guest the closer the calculated/experimental values match with 1,1,1,3,3,3-hexamethylacetone **G$_{33}$** sitting closest to the line of agreement (y = x line).

Adding the new flexible guests (the two series of ketones and series of phosphonates) to the training set and refining the weighting terms in the scoring function leads to a new revised scoring function (Equation 5.4). Individual weightings for each term within the scoring function are allowed to vary freely to minimise the root-mean-square deviation between the calculated and measured binding constants.

$$\log K_{calc} = - (4.48 \times f_{ligandclash}) + (0.22 \times f_{partburied}) - (0.10 \times f_{nonpolar})$$
$$+ (0.022 \times f_{ligandtorsion}) - (0.36 \times f_{ligandflexibility}) \qquad (5.4)$$

The comparison between the new revised scoring function (Equation 5.4) and the original scoring function (Equation 5.1) shows little difference between the first three terms: $f_{ligandclash}$, $f_{partburied}$, and $f_{nonpolar}$. These terms are for unfavourable steric interactions between host/guest, unfavourable contribution from the burial of a polar group in a non-polar environment, and the contribution from matching hydrophobic surfaces, respectively. The remaining two functions, $f_{ligandtorsion}$ and $f_{ligandflexibility}$, have had a substantial change to their weightings in the scoring function and both are considerably diminished. Significantly, these are the two terms that take account of changes in guest conformations upon binding. The $f_{ligandtorsion}$ term accounts for the enthalpic penalty for the required conformational changes and resulting strained bonds that are needed to allow guest binding. The $f_{ligandflexibility}$ term is precisely the NRB, which therefore takes account of the entropic cost of restricting the number of possible conformations a flexible guest can adopt. The fact that both of the terms associated with the flexibility of the guests have diminished shows that the original scoring function (Equation 5.1) was over-estimating the consequence of flexibility, and therefore was systematically under-estimating the calculated binding constant for flexible guests.

Table 5.4   Comparison of weightings for the different scoring functions

Scoring Function	$f_{ligandclash}$	$f_{partburied}$	$f_{nonpolar}$	$f_{ligandtorsion}$	$f_{ligandflexibility}$
Original (5.1)	-4.80	0.20	-0.10	0.90	-0.93
Removed non-binders (5.3)	-4.80	0.22	-0.10	0.32	-0.76
New revised (5.4)	-4.48	0.22	-0.10	0.02	-0.36

Using the new revised scoring function (Equation 5.4) for the plot of calculated *vs.* measured binding, and using the same colours for each of the series of guests as before, showed there is a significant improvement in the agreement with most of the flexible guests sitting close to the agreement line (y = x). Importantly this improvement to the flexible guests is achieved without compromising the prediction of binding constants for the original set of more rigid guests. The original scoring function (5.1) gave a RMSD

value of 0.79 for the corresponding guests. Adding the new ketone and phosphonate guests into the calculation while still using the original scoring function (5.1) resulted in the RMSD increasing to 1.26. The final refinement to give the new scoring function (5.4) resulted in a RMSD value of 0.77; an overall small improvement in agreement.
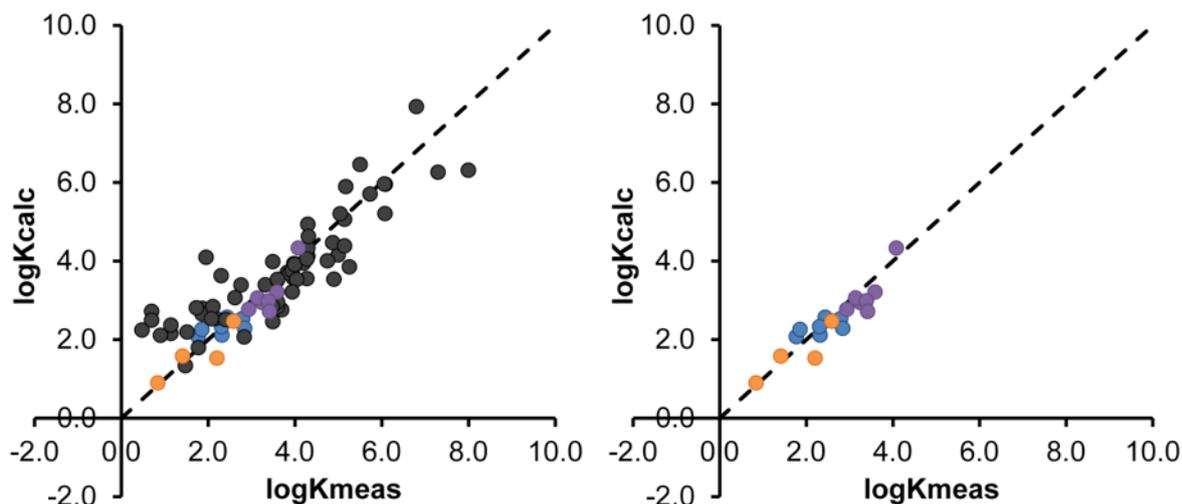
Fig. 5.12 Plot showing calculated (Equation 5.4) *vs.* measured binding constants for the original training set and identified guests (black) along with the new flexible guests; alkyl phosphonates (orange), **C$_7$** (blue), **C$_9$** (purple), (right) same plot showing just the new guests

The individual contributions from each of the five terms (including their associated weightings) are displayed in Table 5.5. The $f_{ligandclash}$ parameter is zero for all of these guests; all the guests have a molecular volume considerably below the Rebek 55% limit[83,85]. If the guests were too large, the weighting for this term would be very large leading to a large negative contribution to the calculated binding constant and the guests simply would not bind. The $f_{ligandtorsion}$ weighted parameter contributes very little to the overall calculated binding constant and has a correspondingly small weighting.

The $f_{partburied}$ weighted parameter is fairly consistent among the ketone and phosphonate guest sets. Interestingly the average $\log K[f_{partburied}]$ for the ketones is -0.44 which equates to a negative contribution of $2.5\,\mathrm{kJ\,mol^{-1}}$. The Ward group[82] has previously shown that there is a penalty of around $7\,\mathrm{kJ\,mol^{-1}}$ to remove a carbonyl group to a non-polar environment and that each additional CH_2 group drives binding by an additional contribution of $4.7\,\mathrm{kJ\,mol^{-1}}$. Earlier the equation 5.2, empirically derived, assumed that each C in the ketone was equivalent to a CH_2 and therefore a C=O group would contribute $4.7\,\mathrm{kJ\,mol^{-1}}$ for the C and $-7\,\mathrm{kJ\,mol^{-1}}$ for removing the group to a

non-polar environment totalling a negative contribution of $2.3\,\mathrm{kJ\,mol^{-1}}$; a value that is within error of the contribution from the $f_{partburied}$ weighted parameter.

The enthalpic contribution to binding described by $f_{nonpolar}$ gives a weighted parameter that is consistent within each series with the $\mathbf{C_7}$ and $\mathbf{C_9}$ sets giving the average $\log K[f_{nonpolar}]$ of 3.88 and 5.08 respectively, corresponding to $-\Delta G$ of $22\,\mathrm{kJ\,mol^{-1}}$ and $29\,\mathrm{kJ\,mol^{-1}}$ respectively. Bringing two separate molecules together has a entropic penalty of $6\,\mathrm{kJ\,mol^{-1}}$[49]. For the $\mathbf{C_7}$ series, one C has already been taken into account, leaving 6 further C atoms and therefore a hydrophobic contribution of $4.7\,\mathrm{kJ\,mol^{-1}}$ x 6 and the penalty of $6\,\mathrm{kJ\,mol^{-1}}$ gives an overall contribution of $22.2\,\mathrm{kJ\,mol^{-1}}$; a value that is within error of the contribution from the $f_{nonpolar}$ weighted parameter. Therefore, as expected, the $f_{nonpolar}$ term describes a majority of the hydrophobic contribution to binding.

Table 5.5 Weighted terms from the refined scoring function (Equation 5.4)

Guest	$-4.48 \times$ $f_{ligandclash}$	$0.22 \times$ $f_{partburied}$	$-0.10 \times$ $f_{nonpolar}$	$0.02 \times$ $f_{ligandtorsion}$	$-0.36 \times$ $f_{ligandflexibility}$	$\log K_{calc}$ $[\mathrm{M}^{-1}]$
$\mathbf{G_{20}}$	0	-0.48	3.99	0.00	-1.44	2.07
$\mathbf{G_{21}}$	0	-0.47	4.16	0.00	-1.44	2.25
$\mathbf{G_{22}}$	0	-0.46	4.01	0.00	-1.44	2.11
$\mathbf{G_{23}}$	0	-0.44	4.08	0.00	-1.08	2.56
$\mathbf{G_{24}}$	0	-0.38	3.79	0.00	-1.08	2.33
$\mathbf{G_{25}}$	0	-0.44	3.68	0.00	-0.72	2.53
$\mathbf{G_{26}}$	0	-0.47	3.47	0.00	-0.72	2.28
$\mathbf{G_{27}}$	0	-0.37	5.29	0.01	-2.16	2.77
$\mathbf{G_{28}}$	0	-0.39	5.12	0.01	-1.80	2.94
$\mathbf{G_{29}}$	0	-0.42	5.27	0.02	-1.80	3.06
$\mathbf{G_{30}}$	0	-0.44	4.93	0.02	-1.80	2.71
$\mathbf{G_{31}}$	0	-0.44	4.85	0.01	-1.44	2.98
$\mathbf{G_{32}}$	0	-0.46	5.10	0.01	-1.44	3.21
$\mathbf{G_{33}}$	0	0.06	4.98	0.01	-0.72	4.33
DMMP	0	-0.77	2.39	0.00	-0.72	0.89
DEMP	0	-0.70	3.71	0.01	-1.44	1.58
DEEP	0	-0.76	4.07	0.02	-1.80	1.53
DIMP	0	-0.82	4.70	0.03	-1.44	2.46
Tabun	0	-0.54	3.97	0.02	-1.08	2.37
Sarin	0	-0.62	3.51	0.00	-0.72	2.18
Cyclosarin	0	-0.60	4.45	0.01	-0.72	3.15
Soman	0	-0.48	4.91	0.00	-1.08	3.35

The final contribution, largely entropic, is described by $f_{ligandflexibility}$ which is effectively the NRB. The weighting for this function gives a value for $\log K[f_{ligandflexibility}]$

of -0.36 per NRB, which equates to a penalty of $2.05\,\mathrm{kJ\,mol^{-1}}$ per NRB. This matches the value calculated empirically earlier in this chapter. Overall, for this series of guests when binding to the host $\mathbf{H^W}$, the contributing parameters to binding (derived by GOLD) are principally: a strongly favourably hydrophobic contribution that scales with the guests surface area, and an unfavourable entropy effect arising from the number of rotatable bonds. Other factors are much less significant. With a relatively large penalty associated with each additional NRB compared to overall binding strength, it is unsurprising that large open chain guests bind weakly or not at all.

One of the initial purposes in accurately calculating the binding constants for the phosphonates was to be able to calculate those for the chemical warfare agents with the G-series of nerve-agents. These values are included in Table 5.5 with Soman predicted to bind the strongest with $K = 2.2 \times 10^3\,\mathrm{M^{-1}}$. While experimental determined binding constants for these agents with the $\mathbf{H^W}$ complex will hopefully be determined at some point in the future, for now it is only possible to compare the calculated binding constants of these nerve agents with the real binding constants of the simulant phosphonates. Typically the active agents have been found to bind more strongly than the simulants and this agrees with the calculated binding constants[113] . Looking at the terms within the calculation this general observation for the active agents becomes more apparent: the nerve agents have an increased $f_{nonpolar}$ along with a reduced penalty for $f_{ligandflexibility}$ which means that both terms should contribute to an increased binding affinity for the nerve agents compared to the corresponding simulants.

5.3   Conclusion

By using two different sets of ketone isomers ($\mathbf{C_7}$ and $\mathbf{C_9}$), along with some simple alkyl phosphonates, the effect of guest flexibility on the ability of the guest to bind within the host complex $\mathbf{H^W}$ has been investigated. The largest difference between the binding strength of the two ketone sets arises from the greater hydrophobic surface area of the $\mathbf{C_9}$ guests. Within each series there was a clear difference between the isomers' binding strength and this was associated with changes in the guests' flexibility. The flexibility of guests was described by the NRB parameter, with the most flexible and weakest binding

guests having the largest NRB in comparison to the least flexible and strongest-binding isomers which have the smallest NRB. An approximate incremental penalty of $2 \, \mathrm{kJ \, mol^{-1}}$ per NRB for the binding free energy was observed, exactly in the range observed for protein/ligand binding.

It was found that the binding constants could be crudely calculated using this incremental penalty, in conjunction with some parameters previously ascertained by the Ward group ($-13 \, \mathrm{kJ \, mol^{-1}}$ for binding a carbonyl group in water, $4.7 \, \mathrm{kJ \, mol^{-1}}$ per CH_2)[82]. Fig. 5.10 showed that this crude calculation actually gave a good agreement with the experimental values.

Refinement of the scoring function used with GOLD led to a significant improvement in the agreement between calculated and experimental binding constants for flexible guests. Importantly the improvement was made without compromising the prediction of binding for the more rigid guests in the original training set[82]. The parameters within the scoring function showed that there is generally one large positive contribution, linked to enthalpy, and one large negative contribution, linked to entropy, that partially cancel out each other. This is in agreement with observations made for binding of ligands to proteins[131].

Finally, it was possible to calculate the binding constants for the four G-series nerve agents using the optimised scoring function for GOLD. While these values can not be confirmed experimentally outside a specialist facility, the good agreement between the calculated and experimentally determined binding constants for the chemical warfare agent simulants and the flexible ketones means that the calculated binding constants can be used with confidence.

5.4 Experimental

Host cage $\mathbf{H^W}$ and $\mathbf{H^A}$ were prepared as described in Chapter 2. All guests were obtained from commercial suppliers and used as received.

5.4.1   Measurements of Binding Constants

Binding constants of guests were measured by ^{1}H-NMR spectroscopy in D_2O at 298 K using a Bruker AV3-400 spectrometer. Concentration of host cage $\mathbf{H^W}$ was always 0.2 mM. Guest binding could be in either fast or slow exchange depending on guest size; illustrations of each type of behaviour, with data analysis, are shown below.

5.4.2   Data from GOLD

Table 5.6   The unweighted terms for each guest from GOLD

	$f_{ligandclash}$	$f_{partburied}$	$f_{nonpolar}$	$f_{ligandtorsion}$	$f_{ligandflexibility}$	$\log K_{calc}$ [M^{-1}]	$\log K_{exp}$ [M^{-1}]
Weightings	-4.4802581	0.2162392	-0.1024202	0.0219217	-0.3599014	-	-
$\mathbf{G_{20}}$	0	-2.2374	-38.9919	0.0851	4	2.07	1.77
$\mathbf{G_{21}}$	0	-2.1715	-40.6066	0.1791	4	2.25	1.86
$\mathbf{G_{22}}$	0	-2.1433	-39.1524	0.2177	4	2.11	2.32
$\mathbf{G_{23}}$	0	-2.0210	-39.8382	0.0631	3	2.56	2.44
$\mathbf{G_{24}}$	0	-1.7773	-37.0008	0.1133	3	2.33	2.31
$\mathbf{G_{25}}$	0	-2.0297	-35.9659	0.1860	2	2.53	2.80
$\mathbf{G_{26}}$	0	-2.1523	-33.8366	0.0476	2	2.28	2.84
$\mathbf{G_{27}}$	0	-1.7020	-51.6439	0.2905	6	2.77	2.94
$\mathbf{G_{28}}$	0	-1.8014	-49.9531	0.5866	5	2.94	3.25
$\mathbf{G_{29}}$	0	-1.9556	-51.4204	0.6880	5	3.06	3.14
$\mathbf{G_{30}}$	0	-2.0342	-48.1200	1.0428	5	2.71	3.42
$\mathbf{G_{31}}$	0	-2.0171	-47.3961	0.2786	4	2.98	3.39
$\mathbf{G_{32}}$	0	-2.1282	-49.7833	0.4193	4	3.21	3.59
$\mathbf{G_{33}}$	0	0.2586	-48.6459	0.3353	2	4.33	4.08
DMMP	0	-3.5734	-23.2924	0.0519	2	0.89	0.85
DEMP	0	-3.2509	-36.1947	0.4886	4	1.58	1.41
DEEP	0	-3.5193	-39.7379	0.7325	5	1.53	2.20
DIMP	0	-3.7765	-45.8432	1.1596	4	2.46	2.59
Tabun	0	-2.5192	-38.7838	1.0845	3	2.37	-
Sarin	0	-2.8536	-34.3182	0.1064	2	2.18	-
Cyclosarin	0	-2.7754	-43.4854	0.6239	2	3.15	-
Soman	0	-2.2300	-47.8987	0.1589	3	3.35	-

5.4.2.1   Example Guest in Slow Exchange

For guests in slow exchange, separate signals for free host (H) and host/guest complex (HG) could be seen during a titration (Fig. 5.13). An expansion and example fit is shown in Fig. 5.14.
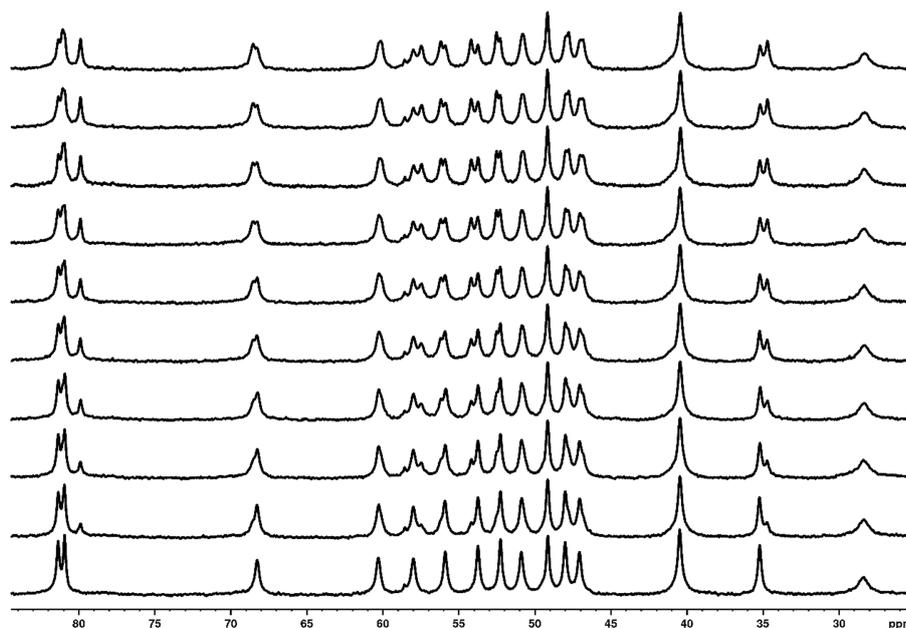
Fig. 5.13 Section of NMR peaks in slow exchange for the titration of H with 4,4-dimethyl-2-pentanone. The fitted data based on peak integrations are shown in Table 5.7 below.
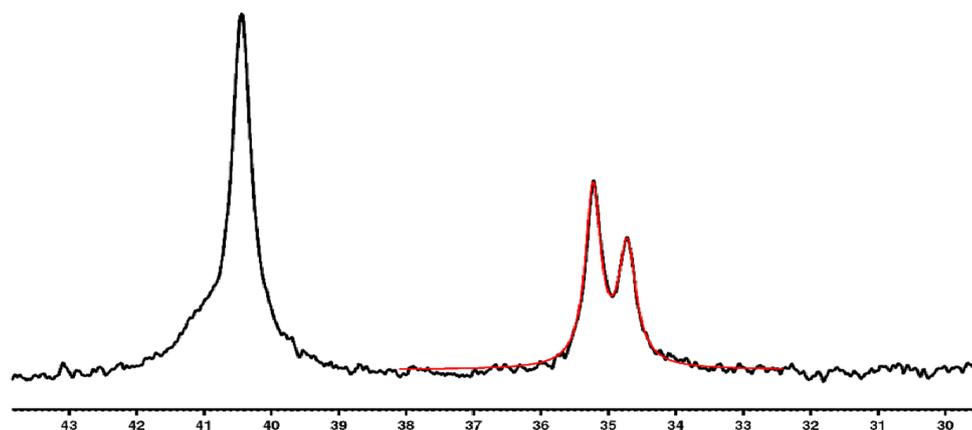
Fig. 5.14 Example of the fitted curve from deconvolution of the overlapping signals for H and HG, used to give the [H]/[HG] ratio (integral ratio 4.582 vs. 4.037, giving ratio of 1.1 : 1, see line 5 in Table 5.7).

5.4.2.2 Example Guest in Fast Exchange

Changes in the ^{1}H-NMR spectra recorded during titration of the host cage (0.2 mM) with 5-methyl-2- hexanone (0 to 10 mM) in D_2O are shown as an example (Fig. 5.15). As this guest is in fast exchange between free / bound states, we see a steady shift in some signals for the host cage, as shown in Fig. 5.16; these curves could be fitted to a 1:1 binding isotherm, and the quoted binding constant is taken from the average of the individual curve fits from several repeat titration experiments.

Table 5.7   The [H]/[HG] ratios for a range of starting guest concentrations, determined by integration of signals in the slow exchange NMR spectra. The free guest, host and host-guest complex concentrations can then be determined and converted to obtain the equilibrium constant and Gibbs free energy. Several repeats were averaged to give the final binding constant quoted in the main text.

[H]/[HG]	$[G]_o$	[H]	[HG]	[G]	K_a	ΔG
2.2	7.5×10^{-4}	1.38×10^{-4}	6.24×10^{-5}	6.88×10^{-4}	6.59×10^2	-16.1
1.7	1×10^{-3}	1.25×10^{-4}	7.45×10^{-5}	9.25×10^{-4}	6.42×10^2	-16
1.4	1.25×10^{-3}	1.15×10^{-4}	8.5×10^{-5}	1.17×10^{-3}	6.34×10^2	-16
1.3	1.5×10^{-3}	1.15×10^{-4}	8.55×10^{-5}	1.41×10^{-3}	5.28×10^2	-15.5
1.1	1.75×10^{-3}	1.06×10^{-4}	9.37×10^{-5}	1.66×10^{-3}	5.32×10^2	-15.6
0.7	2.25×10^{-3}	8.54×10^{-5}	1.15×10^{-4}	2.14×10^{-3}	6.29×10^2	-16
0.8	2.5×10^{-3}	8.64×10^{-5}	1.14×10^{-4}	2.39×10^{-3}	5.51×10^2	-15.6
2.3	7.5×10^{-4}	1.39×10^{-4}	6.14×10^{-5}	6.89×10^{-4}	6.43×10^2	-16
1.2	1.5×10^{-3}	1.11×10^{-4}	8.92×10^{-5}	1.41×10^{-3}	5.71×10^2	-15.7
0.7	2.5×10^{-3}	8.5×10^{-5}	1.15×10^{-4}	2.39×10^{-3}	5.67×10^2	-15.7
1.1	2×10^{-3}	1.04×10^{-4}	9.62×10^{-5}	1.9×10^{-3}	4.87×10^2	-15.3
0.7	2×10^{-3}	8.55×10^{-5}	1.15×10^{-4}	1.89×10^{-3}	7.11×10^2	-16.3
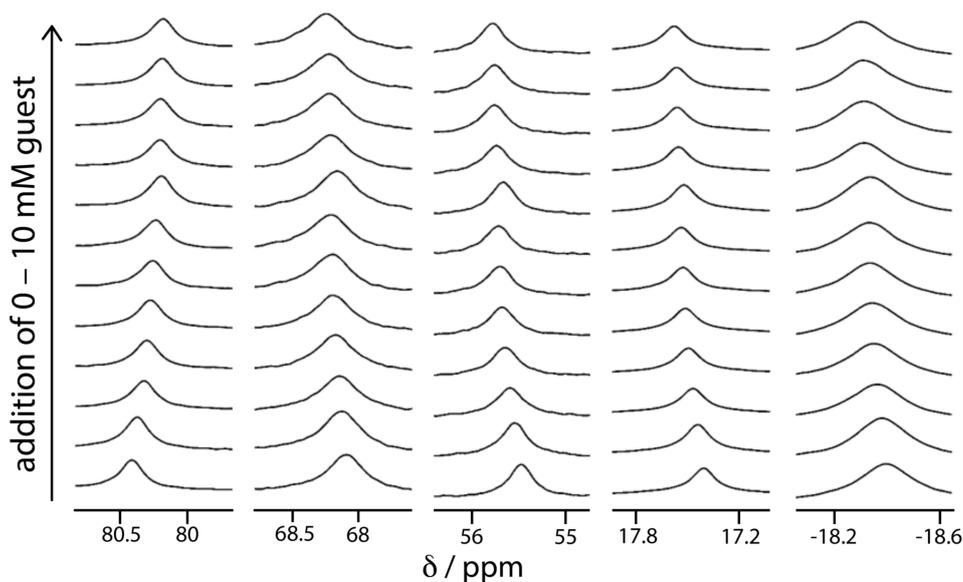
Fig. 5.15 Variation in some ^{1}H-NMR signals of H during titration with 5-methyl-2-hexanone

5.4.3   X-ray Crystallography

The crystal structure data collections were performed using a Bruker APEX-2 CCD diffractometer with Mo-Kα radiation from a sealed tube source unless otherwise stated (a Bruker D8 Venture diffractometer using Cu-Kα radiation was used for complex $[Co_8(L^A)_{12}](BF_4)_{16} \bullet G_{25}$). Data were corrected for absorption using empirical methods (SADABS) based upon symmetry-equivalent reflections combined with measurements at different azimuthal angles. The structures were solved and refined using a combination of
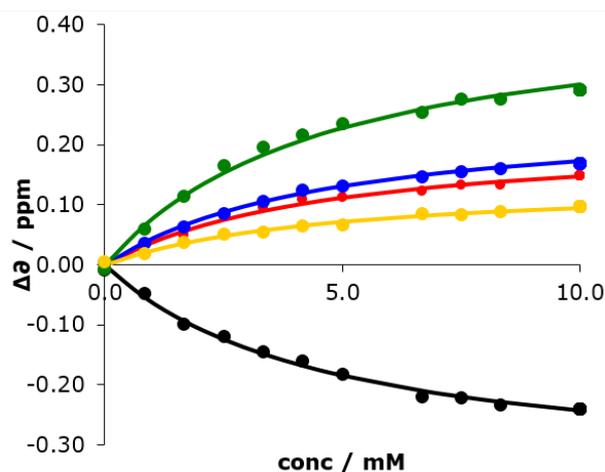
Fig. 5.16 Plot showing the steady shifts in ^{1}H-NMR signals based on the data in Fig. 5.15. All titrations were repeated three times, and binding constants calculated based on shifts of several individual signals in each titration

different software including Olex^2, WinGX, Apex3, PLATON and primarily the SHELX suite version 6.14.

AFIX restraints were applied to all the rings within the cage structure. Various weak global restraints were applied to the located anions and any disorder, where possible, was modelled. The guest was then located within the cavity and DFIX restrains used to optimise the guest parameters. Other solvent molecules that could be located were also modelled before the hydrogens were calculated for cage and guest. Due to a combination of poor diffraction data, strongly absorbing crystals, and disorder of anions/solvents some hydrogens, typically on solvent molecules, were not retained. Once the model was converged there were large regions of diffuse electron density which could not be modelled, accounting for any missing anions and solvent molecules. This electron density was removed from the final refinement using 'SQUEEZE' function in the PLATON software package. The determined structures are of relatively poor quality compared to conventional small-molecule standards though this is typical for these types of self-assembled coordination cage complexes. The overall structure of the cage, the presence of guest, and the relative position and orientation within the cavity are clear. While reference to relative occupancy of encapsulated guests and solvent molecule *vs.* the host complex are made, it should be noted that these are approximate and therefore only used to support other data. No further claims for any specific structural details are used.

5.4.3.1   Crystallography Data Tables

<div align="center">

Table 5.8

</div>

Complex	$[Co_8(L^A)_{12}](BF_4)_{16} \bullet \mathbf{G_{21}}$	$[Co_8(L^A)_{12}](BF_4)_{16} \bullet \mathbf{G_{22}}$
Formula[a]	$C_{344}H_{282}B_{16}Co_8F_{64}N_{72}O_2$	$C_{344}H_{282}B_{16}Co_8F_{64}N_{72}O_2$
Molecular Weight	7316.80	7316.80
$\textbf{\textit{T}}$**, K**	100(2)	100(2)
Crystal System	Monoclinic	Monoclinic
Space Group	C2/c	C2/c
a, Å	32.9486(14)	32.9184(17)
b, Å	29.7415(13)	29.9311(13)
c, Å	39.7282(18)	39.817(2)
α, °	90	90
β, °	96.281(2)	96.111(3)
γ, °	90	90
V, Å^3	38698(3)	39008(3)
Z	4	4
ρ, $g\,cm^{-1}$	1.256	1.246
Crystal Size, mm^3	0.12 x 0.12 x 0.10	0.20 x 0.20 x 0.20
Data, restraints, parameters	44394, 2572, 1974	44822, 2409, 1954
Final R1, wR2[b]	0.1027, 0.3295	0.1230, 0.3493

Table 5.9

Complex	$[Co_8(L^A)_{12}](BF_4)_{16} \bullet \mathbf{G_{23}}$	$[Co_8(L^A)_{12}](BF_4)_{16} \bullet \mathbf{G_{25}}$
Formula[a]	$C_{344}H_{282}B_{16}Co_8F_{64}N_{72}O_2$	$C_{344}H_{282}B_{16}Co_8F_{64}N_{72}O_2$
Molecular Weight	7316.80	7316.80
T, **K**	100(2)	100(2)
Crystal System	Monoclinic	Monoclinic
Space Group	`C2/c`	`C2/c`
a, Å	30.2476(5)	32.9008(15)
b, Å	39.201(6)	30.1251(13)
c, Å	39.9571(7)	39.7960(16)
α, °	90	90
β, °	96.2220(10)	96.327(2)
γ, °	90	90
V, Å^3	39592.5(11)	39203(3)
Z	4	4
ρ, $g\,cm^{-1}$	1.227	1.240
Crystal Size, mm^3	0.24 x 0.24 x 0.24	0.15 x 0.15 x 0.12
Data, restraints, parameters	45473, 2397, 1679	34484, 1442, 1818
Final `R1`, `wR2`[b]	0.1124, 0.2940	0.1133, 0.2933
μ, mm^{-1}	0.421 (Mo-Kα)	3.346 (Cu-Kα)

References

[1] J.-F. Ayme, J. E. Beves, D. A. Leigh, R. T. McBurney, K. Rissanen and D. Schultz, *Nat. Chem.*, 2012, **4**, 15–20.

[2] J. M. Lehn, *Supramolecular Chemistry: Concepts and Persepctives*, Wiley, 2006.

[3] J. W. Steed, D. R. Turner and K. Wallace, *Core Concepts in Supramolecular Chemistry and Nanochemistry*, Wiley, 2007.

[4] D. B. Amabilino and J. F. Stoddart, *Chem. Rev.*, 1995, **95**, 2725–2828.

[5] S. Leininger, B. Olenyuk and P. J. Stang, *Chem. Rev.*, 2000, **100**, 853–908.

[6] P. W. K. Rothemund, *Nature*, 2006, **440**, 297–302.

[7] N. C. Seeman, H. Wang, X. Yang, F. Liu, C. Mao, W. Sun, L. Wenzler, Z. Shen, R. Sha, H. Yan, M. H. Wong, P. Sa-Ardyen, B. Liu, H. Qiu, X. Li, J. Qi, S. M. Du, Y. Zhang, J. E. Mueller, T.-J. Fu, Y. Wang and J. Chen, *Nanotechnology*, 1998, **9**, 257–273.

[8] The Nobel Prize in Chemistry 2016, http://nobelprize.org/nobel_prize/chemistry/laureater/2016/, (accessed 17 Oct 2016).

[9] M. D. Ward and P. R. Raithby, *Chem. Soc. Rev.*, 2013, **42**, 1619–1636.

[10] Z. R. Bell, L. P. Harding and M. D. Ward, *Chem. Commun.*, 2003, **8**, 2432–2433.

[11] R. Chakrabarty, P. S. Mukherjee and P. J. Stang, *Chem. Rev.*, 2011, **111**, 6810–6918.

[12] The Nobel Prize in Chemistry 1987, http://nobelprize.org/nobel_prize/chemistry/laureater/1987/, (accessed 17 Oct 2016).

REFERENCES

[13]   B. Bensaude-Vincent, *NanoEthics*, 2009, **3**, 31–42.

[14]   B. J. Holliday and C. a. Mirkin, *Angew. Chem., Int. Ed. Engl.*, 2001, **40**, 2022–2043.

[15]   G. M. Whitesides and B. Grzybowski, *Science*, 2002, **295**, 2418–2421.

[16]   A. Klug, *Phil. Trans. R. Soc. Lond. B*, 1999, **354**, 531–535.

[17]   J. W. Steed and J. L. Atwood, *Supramolecular Chemistry*, John Wiley and Sons, Ltd, 2nd edn, 2009.

[18]   C. E. Housecroft and A. G. Sharpe, *Inorganic Chemistry*, Pearson, 3rd edn, 2008.

[19]   C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, *J. Chem. Soc., Perkin Trans. 2*, 2001, 651–669.

[20]   E. L. Wolf, *Nanophysics and Nanotechnology: An Introduction to Modern Concepts in Nanoscience, 2nd Edition*, Wiley, 2006, p. 308.

[21]   C. J. Pedersen, *Angew. Chem. Int. Ed. Engl.*, 1988, **27**, 1021–1027.

[22]   J.-M. Lehn, *Angew. Chem. Int. Ed. Engl.*, 1988, **27**, 89–112.

[23]   D. J. Cram, *Angew. Chem. Int. Ed. Engl.*, 1988, **27**, 1009–1020.

[24]   J. M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier and D. Moras, *Proc. Natl. Acad. Sci.*, 1987, **84**, 2565–2569.

[25]   R. Kramer, J. M. Lehn and A. Marquis-Rigault, *Proc. Natl. Acad. Sci.*, 1993, **90**, 5394–5398.

[26]   J. M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151–160.

[27]   C. Dietrich-Buchecker, B. X. Colasson and J.-P. Sauvage, *Top. Curr. Chem.*, 2005, **249**, 261.

[28]   S. R. Seidel and P. J. Stang, *Acc. Chem. Res.*, 2002, **35**, 972–983.

[29]   D. L. Caulder and K. N. Raymond, *Acc. Chem. Res.*, 1999, **32**, 975.

[30]   M. Fujita, M. Tominaga, A. Hori and B. Therrien, *Acc. Chem. Res.*, 2005, **38**, 369–378.

[31] T. R. Cook and P. J. Stang, *Chem. Rev.*, 2015, **115**, 7001–7045.

[32] J. J. Perry IV, J. A. Perman and M. J. Zaworotko, *Chem. Soc. Rev*, 2009, **38**, 1400–1417.

[33] M. Fujita, O. Sasaki, T. Mitsuhashi, T. Fujita, J. Yazaki, K. Yamaguchi and K. Ogura, *Chem. Commun.*, 1996, 1535–1536.

[34] P. Stricklen and J. Verkade, *J. Am. Chem. Soc.*, 1983, **105**, 2494–2495.

[35] P. J. Stang and B. Olenyuk, *Acc. Chem. Res.*, 1997, **30**, 502–518.

[36] M. Fujita and K. Ogura, *Coord. Chem. Rev.*, 1996, **148**, 249–264.

[37] M. Fujita, *Chem. Soc. Rev.*, 1998, **27**, 417.

[38] D. L. Caulder and K. N. Raymond, *J. Chem. Soc., Dalton Trans.*, 1999, 1185–1200.

[39] R. W. Saalfrank, A. Stark, M. Bremer and H.-U. Hummel, *Angew. Chem., Int. Ed. Engl.*, 1990, **29**, 311–314.

[40] R. W. Saalfrank, B. Horner, D. Stalke and J. Salbeck, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 1179.

[41] Q.-F. Sun, J. Iwasa, D. Ogawa, Y. Ishido, S. Sato, T. Ozeki, Y. Sei, K. Yamaguchi and M. Fujita, *Science*, 2010, **328**, 1144–1147.

[42] M. Fujita, D. Oguro, M. Miyazawa, H. Oka, K. Yamaguchi and K. Ogura, *Nature*, 1995, **378**, 469–471.

[43] S. Roche, C. Haslam, H. Adams, S. L. Heath and J. A. Thomas, *Chem. Commun.*, 1998, 1681.

[44] K. Suzuki, M. Tominaga, M. Kawano and M. Fujita, *Chem. Commun.*, 2009, 1638–1640.

[45] S. C. Johannessen, R. G. Brisbois, J. P. Fischer, P. A. Grieco, A. E. Counterman and D. E. Clemmer, *J. Am. Chem. Soc*, 2001, **123**, 3818–3819.

[46] M. D. Ward, *Chem. Commun.*, 2009, 4487–4499.

[47] J. Bunzen, J. Iwasa, P. Bonakdarzadeh, E. Numata, K. Rissanen, S. Sato and M. Fujita, *Angew. Chem. Int. Ed.*, 2012, **51**, 3161–3163.

[48] K. Harris, D. Fujita and M. Fujita, *Chem. Commun.*, 2013, **49**, 6703–6712.

[49] C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310–5324.

[50] N. Galamba, *J. Phys. Chem. B*, 2013, **117**, 2153–2159.

[51] G. Graziano, *J. of Phys. Chem. B*, 2014, **118**, 2598–2599.

[52] N. Galamba, *J. of Phys. Chem. B*, 2014, **118**, 2600–2603.

[53] F. Biedermann, W. M. Nau and H. J. Schneider, *J. Am. Chem. Soc.*, 2012, **134**, 15318–15323.

[54] F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau and A. De Simone, *J. Am. Chem. Soc.*, 2014, **53**, 11158–11171.

[55] M. Yoshizawa, T. Kusukawa, M. Kawano, T. Ohhara, I. Tanaka, K. Kurihara, N. Niimura and M. Fujita, *J. Am. Chem. Soc.*, 2005, **127**, 2798–2799.

[56] P. Thordarson, *Chem. Soc. Rev.*, 2011, **40**, 1305–1323.

[57] S. A. Kadam, K. Haav, L. Toom, T. Haljasorg and I. Leito, *J. Org. Chem.*, 2014, **79**, 2501–2513.

[58] M. Yoshizawa, M. Tamura and M. Fujita, *Science*, 2006, **312**, 251–254.

[59] A. J. Kirby, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 706–724.

[60] M. Yoshizawa, J. K. Klosterman and M. Fujita, *Angew. Chem., Int. Ed. Engl.*, 2009, **48**, 3418–3438.

[61] A. Shivanyuk and J. Rebek, *J. Am. Chem. Soc.*, 2002, **124**, 12074–12075.

[62] J. Kang and J. Rebek, *Nature*, 1997, **385**, 50–52.

[63] P. Mal, B. Breiner, K. Rissanen and J. R. Nitschke, *Science*, 2009, **324**, 1697–1699.

[64] J. E. M. Lewis, E. L. Gavey, S. A. Cameron and J. D. Crowley, *Chem. Sci.*, 2012, **3**, 778–784.

[65] C. G. Hartinger, S. Zorbas-Seifried, M. A. Jakupec, H. Kynast, Bernd Zorbas and B. K. Keppler, *J. Inorg. Biochem.*, 2006, **100**, 891–904.

[66] C. S. Allardyce, P. J. Dyson, D. J. Ellis and S. L. Heath, *Chem. Commun.*, 2001, 1396–1397.

[67] M. A. Furrer, F. Schmitt, M. Wiederkehr, L. Juillerat-Jeanneret and B. Therrien, *Dalton Trans.*, 2012, **41**, 7201–7211.

[68] R. L. Paul, A. J. Amoroso, P. L. Jones, S. M. Couchman, Z. R. Reeves, L. H. Rees, J. C. Jeffery, J. A. Mccleverty and M. D. Ward, *J. Chem. Soc., Dalton Trans.*, 1999, 1563–1568.

[69] A. J. Amoroso, J. C. Jeffery, P. L. Jones, J. A. McCleverty, P. Thornton and M. D. Ward, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 1443–1446.

[70] J. S. Fleming, K. L. V. Mann, C.-A. Carraz, E. Psillakis, J. C. Jeffery, J. A. McCleverty and M. D. Ward, *Angew. Chem. Int. Ed.*, 1998, **37**, 1279–1281.

[71] I. S. Tidmarsh, B. F. Taylor, M. J. Hardie, L. Russo, W. Clegg and M. D. Ward, *New J. Chem.*, 2009, **33**, 366–375.

[72] N. K. Al-Rasbi, C. Sabatini, F. Barigelletti and M. D. Ward, *Dalton Trans.*, 2006, 4769–4772.

[73] R. L. Paul, S. P. Argent, J. C. Jeffery, L. P. Harding, J. M. Lynam and M. D. Ward, *Dalton Trans.*, 2004, 3453–3458.

[74] S. P. Argent, H. Adams, L. P. Harding and M. D. Ward, *Dalton Trans.*, 2006, 542–544.

[75] I. S. Tidmarsh, T. B. Faust, H. Adams, L. P. Harding, L. Russo, W. Clegg and M. D. Ward, *J. Am. Chem. Soc.*, 2008, **130**, 15167–15175.

[76] A. Stephenson, S. P. Argent, T. Riis-Johannessen, I. S. Tidmarsh and M. D. Ward, *J. Am. Chem. Soc.*, 2011, **133**, 858–870.

[77] S. P. Argent, H. Adams, T. Riis-Johannessen, J. C. Jeffery, L. P. Harding and M. D. Ward, *J. Am. Chem. Soc.*, 2006, **128**, 72–73.

[78] Z. R. Bell, J. C. Jeffery, J. A. McCleverty and M. D. Ward, *Angew. Chem. Int. Ed.*, 2002, **41**, 2515–2518.

[79] A. Stephenson, D. Sykes and M. D. Ward, *Dalton Trans.*, 2013, **42**, 6756–6767.

[80] S. Turega, M. Whitehead, B. R. Hall, M. F. Haddow, C. A. Hunter and M. D. Ward, *Chem. Commun.*, 2012, **48**, 2752–2754.

[81] M. Whitehead, S. Turega, A. Stephenson, C. A. Hunter and M. D. Ward, *Chem. Sci.*, 2013, **4**, 2744–2751.

[82] S. Turega, W. Cullen, M. Whitehead, C. A. Hunter and M. D. Ward, *J. Am. Chem. Soc.*, 2014, **136**, 8475–8483.

[83] M. R. Ama, D. Ajami, S. L. Craig, J. S. Yang and J. Rebek, *J. Am. Chem. Soc.*, 2009, **131**, 13190–13191.

[84] J. Rebek, *Acc. Chem. Res.*, 2009, **42**, 1660–1668.

[85] S. Mecozzi and J. Rebek, *Chem. Eur. J.*, 1998, **4**, 1016–1022.

[86] L. Robertson and R. C. Hartley, *Tetrahedron*, 2009, **65**, 5284–5292.

[87] T. Kaminski, P. Gros and Y. Fort, *Eur. J. Org. Chem.*, 2003, **2003**, 3855–3860.

[88] V. C. M. Smith and J.-M. Lehn, *Chem. Commun.*, 1996, 2733–2734.

[89] W. J. Ramsay, T. K. Ronson, J. K. Clegg and J. R. Nitschke, *Angew. Chem. Int. Ed.*, 2013, **52**, 13439–13443.

[90] A. J. Metherell and M. D. Ward, *Chem. Commun.*, 2014, **50**, 10979–10982.

[91] A. J. Metherell and M. D. Ward, *Chem. Commun.*, 2014, **50**, 6330–6332.

[92] L. Pazderski, T. Pawlak, J. Sitkowski, L. Kozerski and E. Szlyk, *Magn. Reson. Chem.*, 2011, **49**, 237–241.

[93] C. G. P. Taylor, J. R. Piper and M. D. Ward, *Chem. Commun.*, 2016, **52**, 6225.

[94] C. S. Hawes, C. M. Fitchett and P. E. Kruger, *Supramol. Chem.*, 2012, **24**, 553–562.

[95] N. Yanagisawa, H. Morita, T. Nakajima, H. Okudera, M. Shimizu, H. Hirabayashi, M. Nohara, Y. Midorikawa and S. Mimura, *Lancet*, 1995, **346**, 290–293.

[96] E. Dolgin, *Nat. Med.*, 2013, **19**, 1194–1195.

[97] C. B. Caputo, L. J. Hounjet, R. Dobrovetsky and D. W. Stephan, *Science*, 2013, **341**, 1374–1377.

[98] B. Gehauf, J. Epstein, G. B. Wilson, B. Witten, S. Sass, V. E. Bauer and W. H. C. Rueggeberg, *Anal. Chem.*, 1957, **29**, 278–281.

[99] G. A. Grant, R. Blanchfield and D. M. Smith, *Can. J. Chem.*, 1957, **35**, 42–49.

[100] T. Dale and J. Rebek, *Angew. Chem. Int. Ed.*, 2009, **48**, 7850–7852.

[101] A. Wild, A. Winter, M. D. Hager and U. S. Schubert, *Chem. Commun.*, 2012, **48**, 964–966.

[102] M. R. Sambrook, J. R. Hiscock, A. Cook, A. C. Green, I. Holden, J. C. Vincent and P. A. Gale, *Chem. Commun.*, 2012, **48**, 5605–5607.

[103] M. Zengerle, F. Brandhuber, C. Schneider, F. Worek, G. Reiter and S. Kubik, *Beilstein J. Org. Chem.*, 2011, **7**, 1543–1554.

[104] J. R. Hiscock, I. L. Kirby, J. Herniman, G. John Langley, A. J. Clark and P. A. Gale, *RSC Adv.*, 2014, **4**, 45517–45521.

[105] J. R. Hiscock, M. R. Sambrook, J. A. Ede, N. J. Wells and P. A. Gale, *J. Mater. Chem.*, 2015, **3**, 1230–1234.

[106] F. Piana, M. Facciotti, G. Pileio, J. R. Hiscock, W. Van Rossom, R. C. D. Brown and P. A. Gale, *RSC Adv.*, 2015, **5**, 12287–12292.

[107] J. E. Mondloch, M. J. Katz, W. C. Isley Iii, P. Ghosh, P. Liao, W. Bury, G. W. Wagner, M. G. Hall, J. B. DeCoste, G. W. Peterson, R. Q. Snurr, C. J. Cramer, J. T. Hupp and O. K. Farha, *Nat. Mater.*, 2015, **14**, 512–516.

[108] S.-Y. Moon, Y. Liu, J. T. Hupp and O. K. Farha, *Angew. Chem. Int. Ed.*, 2015, **54**, 6795–6799.

[109] M. Burnworth, S. J. Rowan and C. Weder, *Chem. Eur. J.*, 2007, **13**, 7828–7836.

[110] S.-W. Zhang and T. M. Swager, *J. Am. Chem. Soc.*, 2003, **125**, 3420–3421.

[111] S. M. Daly, M. Grassi, D. K. Shenoy, F. Ugozzoli and E. Dalcanale, *J. Mater. Chem.*, 2007, **17**, 1809–1818.

[112] G. E. Southard, K. A. Van Houten, E. W. Ott Jr and G. M. Murray, *Anal. Chim. Acta*, 2007, **581**, 202–207.

[113] M. R. Sambrook and S. Notman, *Chem. Soc. Rev.*, 2013, **42**, 9251–9267.

[114] S. L. Bartelt-Hunt, D. R. U. Knappe and M. A. Barlaz, *Crit. Rev. Env. Sci. Technol.*, 2008, **38**, 112–136.

[115] K. Kim, O. G. Tsay, D. A. Atwood and D. G. Churchill, *Chem. Rev.*, 2011, **111**, 5345–5403.

[116] J. R. Hiscock, F. Piana, M. R. Sambrook, N. J. Wells, A. J. Clark, J. C. Vincent, N. Busschaert, R. C. D. Brown and P. A. Gale, *Chem. Commun.*, 2013, **49**, 9119–9121.

[117] A. Barba-Bon, A. M. Costero, M. Parra, S. Gil, R. Martnez-Mez, F. Sancenn, P. A. Gale and J. R. Hiscock, *Chem. Eur. J.*, 2013, **19**, 1586–1590.

[118] J. R. Hiscock, M. R. Sambrook, P. B. Cranwell, P. Watts, J. C. Vincent, D. J. Xuereb, N. J. Wells, R. Raja and P. A. Gale, *Chem. Commun.*, 2014, **50**, 6217–6220.

[119] J. L. Bolliger, A. M. Berlenguer and J. R. Nitschke, *Angew. Chem., Int. Ed.*, 2013, **52**, 7958.

[120] S. Turega, M. Whitehead, B. R. Hall, A. J. H. M. Meijer, C. A. Hunter and M. D. Ward, *Inorg. Chem.*, 2013, **52**, 1122–1132.

[121] Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen and M. Fujita, *Nature*, 2013, **495**, 461–466.

[122] D. J. Cram, M. E. Tanner and R. Thomas, *Angew. Chem. Int. Ed.*, 1991, **30**, 1024–1027.

[123] C. J. Brown, F. D. Toste, R. G. Bergman and K. N. Raymond, *Chem. Rev.*, 2015, **115**, 3012–3035.

[124] R. Wolfenden, X. Lu and G. Young, *J. Am. Chem. Soc.*, 1998, **120**, 6814–6815.

[125] R. J. Hooley and J. Rebek Jr., *Org. Biomol. Chem.*, 2007, **5**, 3631–3636.

[126] C. J. Hastings, M. D. Pluth, R. G. Bergman and K. N. Raymond, *J. Am. Chem. Soc.*, 2010, **132**, 6938–6940.

[127] D. M. Kaphan, M. D. Levin, R. G. Bergman, K. N. Raymond and F. D. Toste, *Science*, 2015, **350**, 1235–1238.

[128] W. Cullen, K. A. Thomas, C. A. Hunter and M. D. Ward, *Chem. Sci.*, 2015, **6**, 4025–4028.

[129] W. Cullen, M. C. Misuraca, C. A. Hunter, N. H. Williams and M. D. Ward, *Nat. Chem.*, 2016, **6**, 231–236.

[130] R. R. Dean and W. McFarlane, *Chem. Commun.*, 1967, 840–841.

[131] M. K. Gilson and H.-X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.*, 2007, **36**, 21–42.

[132] M. M. J. Smulders, S. Zarra and J. R. Nitschke, *J. Am. Chem. Soc.*, 2013, **135**, 7039–7046.

[133] J. L. Brumaghim, M. Michels and K. N. Raymond, *Eur. J. Org. Chem.*, 2004, **2004**, 4552–4559.

[134] O. Korb, T. Sttzle and T. E. Exner, *Journal of Chemical Information and Modeling*, 2009, **49**, 84–96.

[135] H. J. Bohm, *J. Computer-Aided Mol. Des.*, 1994, **8**, 243.

[136] M. D. Eldridge, C. W. Murry, T. R. Auton, R. P. Paolini and J. Mee, *J. Computer-Aided Mol. Des.*, 1997, **11**, 425.

Appendix A

Publication Reprints