

**Development of Building Blocks
Exhibiting Self-Sorting Molecular
Recognition Properties: Towards Coded
Self-Assembly Processes**

Maria Louise Pellizzaro

Submitted in accordance with the requirements for the degree of Doctor of
Philosophy

The University of Leeds
School of Chemistry
January 2012

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The candidate was responsible for the majority of the data collection and analysis, however gratefully acknowledges the assistance of the following individuals:

- Colin Kilner and Marc Little (X-ray – Chapter 2, Chapter 3);
- Dr. Julie Fisher (DOSY – Chapter 2, Chapter 4);
- Dr. Bruce Turnbull (ITC data analysis – Chapter 2);
- Dr. Mike Nix (Molecular modelling – Chapter 2);
- Simon Barrett (^1H - ^1H ROESY and VT NMR experiments – Chapter 2, Chapter 4)
- Dr. Adam Gooch (^1H NMR titrations – Chapter 3);
- Nedolisa Chinemelum (DSC measurements – Chapter 3);
- Tanya Marinko-Covell (MS).

Publications:

- Adam Gooch, Andrea M. McGhee, Maria L. Pellizzaro, Christopher I. Lindsay and Andrew J. Wilson, Substituent Control over Dimerization Affinity of Triply Hydrogen Bonded Heterodimers, *Org. Lett.*, **2011**, *13*, 240–243.
- Maria L. Pellizzaro, Andrea M. McGhee, Lisa C. Renton, Michael G. Nix, Julie Fisher, W. Bruce Turnbull, and Andrew J. Wilson, Conformer-Independent Ureidoimidazole Motifs – Tools to Probe Conformational and Tautomeric Effects on the Molecular Recognition of Triply Hydrogen Bonded Heterodimers, *Chem. Eur. J.*, **2011**, In press.

- Maria L. Pellizzaro, Simon A. Barrett, Julie Fisher and Andrew J. Wilson, Design, Synthesis and Binding Studies of a Novel Quadruple ADDA Hydrogen-Bond Array, *Org. Biomol. Chem.*, **2011**, Submitted.

This copy has been supplied on the understanding that it is copyright material and that no quotation from this thesis may be published without prior acknowledgement.

© 2012, The University of Leeds and Maria Louise Pellizzaro

Acknowledgements

Firstly I would like to thank my PhD supervisor Dr. Andrew Wilson, not only for giving me the opportunity to be involved in this project and helping me throughout but also for all the free food and drink. I would also like to thank Dr. Andrea McGhee for carrying out the initial work in the project, ensuring that there was a funded position for me to take and Adam Gooch for teaching me everything he knew, although it didn't take very long.

I have had the opportunity to collaborate with many other academic researchers throughout my PhD and I would like to thank all of them. Dr. Julie Fisher, who was very trusting and let me use the NMR machine alone, Simon Barrett, who didn't have a heart attack, even though the NMR machine broke many times, Dr. Mike Nix, for teaching someone who was computer illiterate to carry out molecular modelling, Prof. Jim Guthrie, for all the help with viscometry and Dr. Bruce Turnbull, for helping with the ITC data. Thanks must also go to Colin Kilner and Marc Little for determining my crystal structures.

A big thank you goes to the entire Wilson Group, past and present for inviting me on every night out, even though I hardly ever went. Thanks also to Tasha, your dreams were always an 'interesting' listen, but cake Wednesdays were a genius idea; Valeria, for persevering with the coffee offers, but sorry, I just don't like it; Kerya, for being the most sarcastic person I've ever met; Dave, for football chat; Maria, for being tall and leaving me with the nickname 'Little Maria'; George B, for falling asleep and having your hair cut off, amused me for days; Hannah, bringing hockey gossip to the lab ... thanks ...; Panchami, for always helping whenever I had a question, no matter how ridiculous you thought it was; Kelly, for taking over the project and bringing cake and George P, is this sentence long enough?!

Massive thanks has to go to Cathryn for booking me a trip to Australia, giving me a deadline so that I could see the light at the end of the never ending PhD tunnel.

And finally. Nana. Yes I did wear a white coat.

Abstract

Biology can achieve phenomenal information storage/processing capabilities from just four hydrogen bonding molecules. The orthogonal self-sorting of these four hydrogen bonding motifs is achieved alongside the hydrophobic collapse of the DNA backbone. Orthogonal self-sorting without the aid of a polymer backbone is difficult to achieve, as shown by the minimal examples of self-sorting hydrogen bonding motifs available in the literature. This thesis describes the design, synthesis and binding studies of molecules that are capable of orthogonal self-sorting, without a preorganising backbone and begins to use them in a signalling cascade.

In Chapter 2 ureidoimidazole, a conformer independent triple hydrogen-bond array, is introduced. Studies were carried out to investigate what effect, if any, preorganisation using intramolecular hydrogen bonding had on the binding affinity of triply hydrogen bonded complexes.

Chapter 3 investigates another factor that can effect the binding affinity of complexes, the remote substituent effect. Two series of complementary molecules were synthesised so that they contained a variety of electron donating/withdrawing groups and the effect that these had on the binding affinity of the complex was measured.

Chapter 4 describes a novel quadruple hydrogen-bond array, which was designed to interact strongly with its complementary partner. It was found that a combination of effects (differences in geometry and undesired conformers being favoured) lead to a low binding affinity being observed.

Chapter 5 begins to investigate non-linear arrays, however none of those proposed were able to form heterodimers. Therefore a self-sorting system was assembled using a triple and quadruple hydrogen-bonded array. High fidelity interactions were achieved, even though it was possible to form undesired complexes. These undesired molecular interactions were exploited in Chapter 5, where a signalling cascade is described. Careful planning of the order in which to add the molecules can give different routes in which to achieve the self-sorting system, each with it's own fidelity trace. A photolabile tag was introduced to one of the molecules so that a photosensitive system could be achieved.

Contents

Acknowledgements	v
Abstract	vi
Abbreviations	xi
1 Molecular Recognition and Self-Assembly of Hydrogen Bonding Motifs	1
1.1 Hydrogen Bonds in Supramolecular Assemblies	3
1.1.1 Strength of hydrogen bonds	4
1.1.2 Secondary Electrostatic Interactions	6
1.1.3 Conformer Considerations	8
1.1.4 Tautomer Considerations	12
1.2 Self-Assembly into Designed Architectures	16
1.2.1 Defined Two- and Three-Dimensional Architectures	16
1.2.2 Polymer Assembly	21
1.2.3 Self-Assembling Capsules	23
1.3 Self-Sorting in Hydrogen bonded Systems	25
1.4 Project Aims	32
2 Effects of Preorganisation	33
2.1 Conformer independent array	38
2.1.1 Synthesis of Molecules	40
2.1.2 Binding Studies	42
2.1.3 Measurement of Dimerisation	44
2.1.4 Confirmation of Conformer Independence	47
2.2 Effects of preorganisation in AAD·DDA arrays	54
2.2.1 Synthesis of Molecules	54
2.2.2 ¹ H NMR Titrations	56
2.2.3 NMR Studies of Compounds and Complexes	57
2.2.4 Molecular Modelling	58
2.3 Conclusions	62
3 Remote Substituent Effect and Applications in Supramolecular Polymer Assembly	65

3.1	Substituent Control over the Binding Affinity of DDA · AAD Arrays	68
3.1.1	Synthesis of the Molecules	69
3.1.2	¹ H NMR Titration Experiments	71
3.1.3	Molecular Modelling	74
3.2	Substituent Effect in Supramolecular Polymers	78
3.2.1	Synthesis of polymers	80
3.2.2	Self-Assembly into Polymers	85
3.3	Conclusions	88
4	A Novel ADDA Quadruple Hydrogen-Bond Array	89
4.1	Design and Synthesis of UDIM and DAN	92
4.2	Conformational Studies of UDIM	94
4.3	Heterodimerisation Studies of UDIM · DAN	102
4.4	Comparison of the Binding Properties of a Triple Hydrogen-Bond Array .	106
4.5	Conclusions	112
5	Novel Arrays for Self-Sorting Assemblies	113
5.1	Enol Array	115
5.1.1	Synthesis and Binding Studies	115
5.2	Bifurcated Array	119
5.2.1	Synthesis and Binding Studies	119
5.3	Three Dimensional Array	123
5.3.1	Synthesis and Binding Studies	123
5.4	A Self-Sorting System	127
5.5	Signalling Cascade	135
5.6	Photosensitive signalling cascade	142
5.7	Conclusions	152
6	Thesis Summary	155
6.1	Future Directions	158
7	Experimental	161
7.1	Experimental for Chapter 2: Effects of Preorganisation	162
7.2	Experimental for Chapter 3: The Substituent Effect	167
7.3	Experimental for Chapter 4: A Novel ADDA array	179
7.4	Experimental for Chapter 5: Self-Sorting of Hydrogen-Bonding Motifs . .	183
7.5	NMR Titration/Dilution Experiments	196
7.6	Molecular Modelling	196
7.7	DOSY data acquisition	197
7.8	NOESY data acquisition	197
7.9	ROESY data acquisition	197
7.10	ITC experiments	198
7.11	DSC experiments	199
7.12	Viscometry	199
7.13	HySS	199
	References	200

A	Crystal Structures	213
B	¹H NMR Titrations	217
B.1	Preorganisation	218
B.2	Substituent Effects	220
B.3	UDIM	226
B.4	Self-Sorting	227
C	Molecular Modelling	229
C.1	Molecular Modelling for Chapter 2: Effects of Preorganisation	230
C.2	Molecular Modelling for Chapter 3: The Substituent Effect	251
C.3	Molecular Modelling for Chapter 4: A Novel ADDA array	268

Abbreviations

Abbreviation/Symbol	Meaning
2-D	Two dimensional
3-D	Three dimensional
A	Hydrogen bond acceptor
Å	Angstrom
AIC	Amidoisocytosine
AIC ⁺	Benzylamidoisocytosine
AIC*	2-Nitrobenzylamidoisocytosine
ANT	Benzamidonaphthrydine
B3LYP	Becke, three-parameter, Lee-Yang-Parr
Boc	<i>t</i> -Butyloxycarbonyl
br	Broad (NMR, IR)
BSSE	Basis set superposition error
CDI	1,1-Carbonyldiimidazole
COSY	¹ H- ¹ H correlation NMR spectroscopy
d	Doublet (NMR)
D	Hydrogen bond donor/Diffusion coefficient
DAN	2,7-Diamido-1,8-naphthyridine
DAP	Diamidopyridine
DeAP	Deazapterin
DEB	5,5-diethylbarbituric acid
DFT	Density functional theory
DMAP	<i>N,N</i> -Dimethyl-4-aminopyridine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulphoxide

DNA	Deoxyribonucleic acid
DOSY	Diffusion-ordered 2D NMR spectroscopy
DP	Degree of polymerisation
DPU	Diphenylurea
DTU	Ditriazoleurea
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDG	Electron donating group
ES	Electrospray
ESI	Electrospray ionisation
EWG	Electron withdrawing group
HAC	Hexylamidocytosine
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HMQC	Heteronuclear multiple-quantum correlation spectroscopy
HOBT	Hydroxybenzotriazole
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
Hz	Hertz
IR	Infrared
ITC	Isothermal titration calorimetry
J	Coupling constant in Hz
K	Temperature (Kelvin)
K_a	Association constant
K_{dim}	Dimerisation constant
m	multiplet (NMR)
MS	mass spectrometry
MW	Molecular weight
m.p.	Melting point
m/z	Mass-to-charge ratio
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
ppm	Parts per million
PUPY	Phenylureidopyrimidine
Py	pyridine

R	Alkyl group/Gas constant ($8.314 \text{ J K}^{-1} \text{ mol}^{-1}$)
R_f	Retention factor
RNA	Ribonucleic acid
ROESY	Rotating-frame Overhauser effect spectroscopy
s	Singlet (NMR)
t	Triplet (NMR)
T	Temperature
TBD	Triazabicyclodecene
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
UCyt	Ureidocytosine
UDIM	Ureidodiimidazole
UG	Ureidoguanosine
UIM	phenylureidoimidazole
UPy	Ureidopyrimidinone
UPy*	Benzylureidopyrimidinone
UV	ultraviolet
Vis	Visible
VPO	Vapour pressure osmometry
δ	Chemical shift
ΔG°	Standard Gibbs free energy
ΔH°	Standard enthalpy
ΔS°	Standard entropy
σ	Sigma
χ	Mole fraction

Chapter 1

Molecular Recognition and Self-Assembly of Hydrogen Bonding Motifs

Supramolecular chemistry is the chemistry of the intermolecular bond,¹ allowing small molecules to assemble into much larger, more complicated structures² that would be difficult to synthesise using covalent bond formations. Supramolecular chemistry has been inspired by nature, where extensive examples of self-assembly using intermolecular interactions can be found,³⁻⁵ for example protein folding⁶ and DNA duplex formation^{7,8} (Figure 1.1). In these examples, many weak, non-covalent interactions are used in concert in order to obtain large, well-defined architectures.⁹

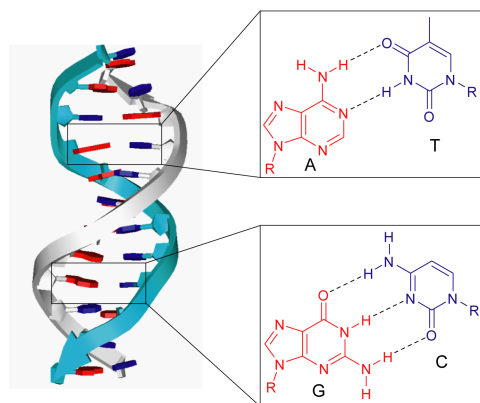


Figure 1.1: The DNA duplex is assembled using a combination of hydrophobic, π - π stacking and hydrogen bond non-covalent interactions.

Since intermolecular interactions are normally weaker than covalent bonds they can be reversible. This allows supramolecular assemblies to be under equilibrium control, resulting in the thermodynamic, not the kinetic, product being obtained (Figure 1.2). For example, consider a system containing A and B; three species can exist (A, B, and A·B), but the most prevalent will be the component that is thermodynamically favourable.

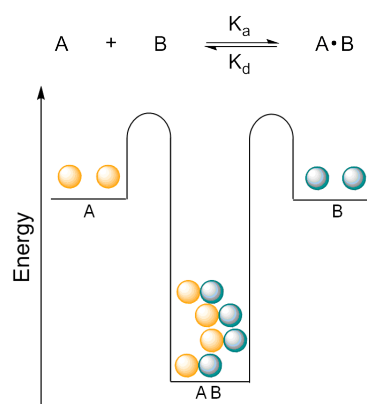


Figure 1.2: AB will be the major component of the mixture because it is thermodynamically favoured.

Another advantage of using reversible intermolecular interactions is that they allow supramolecular assemblies to be self-healing. If an assembly is formed that is higher in energy than the thermodynamic product, it can dissociate to allow assembly of the most stable, thermodynamic product. This is one of the reasons why nature uses lots of weak interactions rather than using a few strong interactions. Since each individual interaction is too weak to maintain the desired structure they must all be used in concert to ensure that only the desired assemblies are formed.

The reversible nature of the non-covalent intermolecular interactions also allow supramolecular assemblies to be stimuli responsive. Since the association constant (K_a) is proportional to the concentration of bound material compared to unbound molecules (Equation 1.1), changing the position of equilibrium results in a change in K_a and can permit previously unfavoured assemblies to be the most prevalent.

$$K_a \propto \frac{[A \cdot B]}{[A][B]} \quad (1.1)$$

The equilibrium position of supramolecular systems can be changed by altering the pH,¹⁰ chemical environment,¹¹ redox potential^{12,13} and by using light^{14–17} as a stimulus. K_a

is related to free energy, ΔG° by the Equation 1.2 and ΔG° can also be expressed by Equation 1.3. These two equations show that as well as changes in temperature,¹⁸ enthalpy and entropy can also affect the binding affinity of supramolecular assemblies. This indicates that alterations to the individual components, along with changes to the external environment can all influence interactions within supramolecular assemblies.

$$\Delta G^\circ = -RT\ln K_a \quad (1.2)$$

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (1.3)$$

1.1 Hydrogen Bonds in Supramolecular Assemblies

There are many different types of non-covalent interaction (Table 1.1) available for use by supramolecular chemists.¹⁹ Hydrogen bonds are commonly chosen because they can be strong yet reversible, selective and directional^{20,21} and in addition, it is relatively easy to manipulate the strength of binding. In principle, this makes it quite easy to design complementary motifs and allows the strength of interaction to be manipulated, according to the application that is required.

Selectivity arises because, for a hydrogen bond to form, an acceptor (A) must interact with a donor (D). A's are normally electronegative heteroatoms (e.g. O or N) and D's are normally hydrogens that are attached to an electronegative atom (Figure 1.3). The electronegative atom removes electron density from the hydrogen atom, resulting in

Table 1.1: Examples of non-covalent interactions and their associated typical energies.¹⁹

Non-covalent interaction	Typical Energy Range (kJ mol ⁻¹)
Ion-Ion	50–400
Ion-dipole	50–200
Dipole-dipole	4–40
Hydrogen Bonding	4–120
Halogen Bonding	5–180
Cation- π	5–80
π - π	4–20
Solvation	4–40

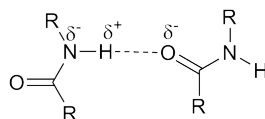


Figure 1.3: A typical hydrogen bond between amide functionalities. NH acts as the hydrogen bond donor and the carbonyl acts as the hydrogen bond acceptor.

a positive charge, which has attractive interactions towards the negatively charged A atom. Directionality arises because hydrogen bonds are formed between dipoles. The electronegative atoms position themselves as far apart as possible, resulting in a linear alignment of the atoms involved in the hydrogen bond.

1.1.1 Strength of hydrogen bonds

The strength of hydrogen bonds can be altered by manipulating the acidity/basicity of the hydrogen bonding atoms.²² The pK_a equalisation principle²³ shows that the strongest hydrogen bonds are obtained when ΔpK_a is zero (Equation 1.4). When ΔpK_a between A and D is zero the hydrogen bond can be considered as a three-centre-four-electron covalent bond, so it becomes irreversible. In medium strength hydrogen bonds, which are ideal for reversible interactions, ΔpK_a should be ~ 11 – 21 .²⁴ This is normally achieved by using molecules where the pK_a of the D is 11–21 and the pK_a of the A is -7 – 3 . It has also been suggested that changing the pK_a of the D can have a larger effect on the strength of the hydrogen bond than changing the pK_a of the A.^{25–27} Addition of electron withdrawing acyl groups to diaminopyridine **1** gave the diamidopyridine **2**.²⁸ Stronger K_a 's were obtained between **2** and the uracil derivative **3** ($K_a = 800 \text{ M}^{-1}$) compared to **1** ($K_a = 84 \text{ M}^{-1}$), presumably because the NH's were more acidic (Figure 1.4).

$$\Delta pK_a = pK_a(\text{D-H}) - pK_a(\text{A-H}^+) \quad (1.4)$$

Ordinarily, a single hydrogen bond does not have a strong enough K_a when used alone, so multiple hydrogen bonds are used simultaneously as part of a hydrogen-bonding array. It has been shown that large differences in binding affinity can be achieved by using different numbers of hydrogen bonds.^{29,30} For example, three orders of magnitude difference in K_a was measured between **4·5** and **4·6** where the only difference was the addition of an extra hydrogen bond (Figure 1.5). Supramolecular homo- and heterodimers that

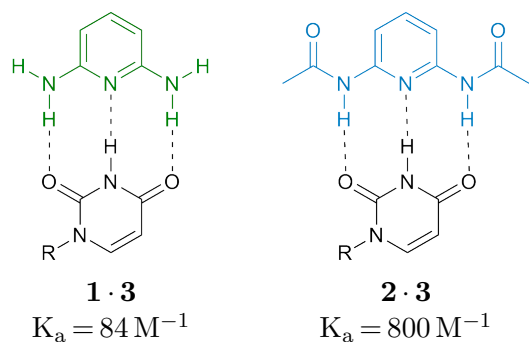


Figure 1.4: The association constant of diaminopyridine **1** can be increased by adding electron withdrawing groups to give diamidopyridine **2**. Where R denotes generic alkyl chain.

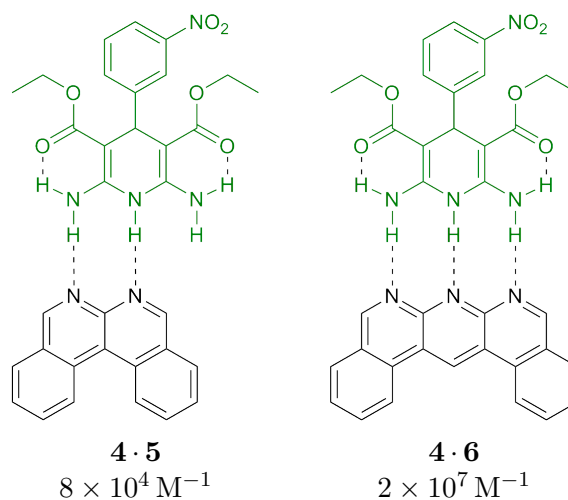


Figure 1.5: Large differences in binding affinity are seen between AA · DDD and AAA · DDD hydrogen bond arrays.

interact using two,³¹ three,^{30,32,33} four³⁴⁻⁴⁰ and six^{41,42} hydrogen bonds have all been described in the literature. It is generally accepted that using three or four intermolecular hydrogen bonds gives the best properties for most supramolecular applications because they assemble into strong, yet reversible complexes. Using less than three hydrogen bonds does not give strong enough interactions, but using more than four intermolecular hydrogen bonds can result in strong K_a 's that give essentially irreversible complexes.

1.1.2 Secondary Electrostatic Interactions

In addition to the number of hydrogen bonds used in the array, an important consideration is the order in which the D's and A's are arranged. Large differences in binding affinity have been obtained, simply by changing the order of the D's and A's in the array,^{32,43,44} due to secondary electrostatic interactions.⁴⁵ Secondary electrostatic interactions can be constructive or destructive in nature, depending on the arrangement of D's and A's in the array (Figure 1.6). Where triple hydrogen-bond arrays are concerned, DDD · AAA arrays normally have the strongest binding affinities ($K_a \approx 10^5 \text{ M}^{-1}$) because they contain four attractive secondary electrostatic interactions. Conversely, DAD · ADA arrays normally have the weakest binding affinities ($K_a \approx 10^2 \text{ M}^{-1}$) because they contain four repulsive secondary electrostatic interactions. Extremely strong binding affinities were measured in the AAAA · DDDD⁺ array **7·8** ($K_a > 3 \times 10^{12} \text{ M}^{-1}$) introduced by Leigh *et al.* (Figure 1.7) because there are four primary hydrogen bonds and six attractive secondary electrostatic interactions.³⁵ The molecules in this complex were also

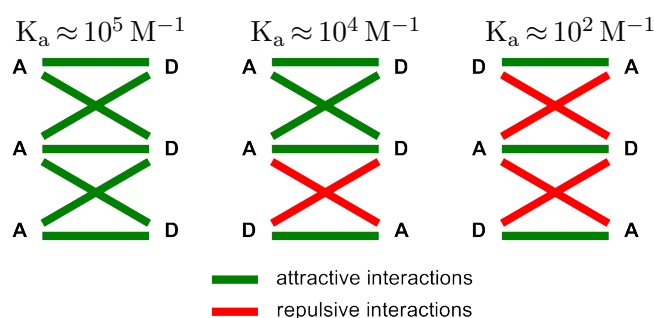


Figure 1.6: Adjacent A and D can contribute positively or negatively to the binding affinity of the complex allowing different association constants to be obtained.

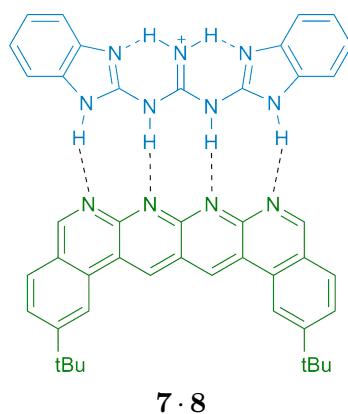


Figure 1.7: Extremely strong binding affinities were obtained in this AAAA · DDDD⁺ **7·8** array because there are no repulsive secondary electrostatic interactions.

designed so that they could be synthesised in just two steps, allowing complexes that have exceptionally strong binding affinities to be readily accessible.

Since hydrogen bonds are additive, it is possible to predict the association constants of dimers without having to synthesise and test each molecule.⁴⁶ It was calculated that primary hydrogen bonds increased the binding affinity by -7.9 kJ mol^{-1} and secondary electrostatic interactions added $\pm 2.9 \text{ kJ mol}^{-1}$ depending on whether they were constructive or destructive in nature.⁴⁷ By adding spacers between the atoms involved in hydrogen bonding the effects of secondary electrostatic interactions can be eliminated.⁴⁸⁻⁵³ Gong *et al.* have designed a molecule (**9**) that is capable of self-recognition ($K_{\text{dim}} = 6 \times 10^4 \text{ M}^{-1}$), but contains no secondary electrostatic interactions (Figure 1.8).⁵⁴ This allows the binding affinity of the complex to be calculated based solely on the number of primary hydrogen bonds that are present, without interference from secondary electrostatic interactions. Therefore the same binding affinities should be achieved, no matter what order the hydrogen bond array is in.

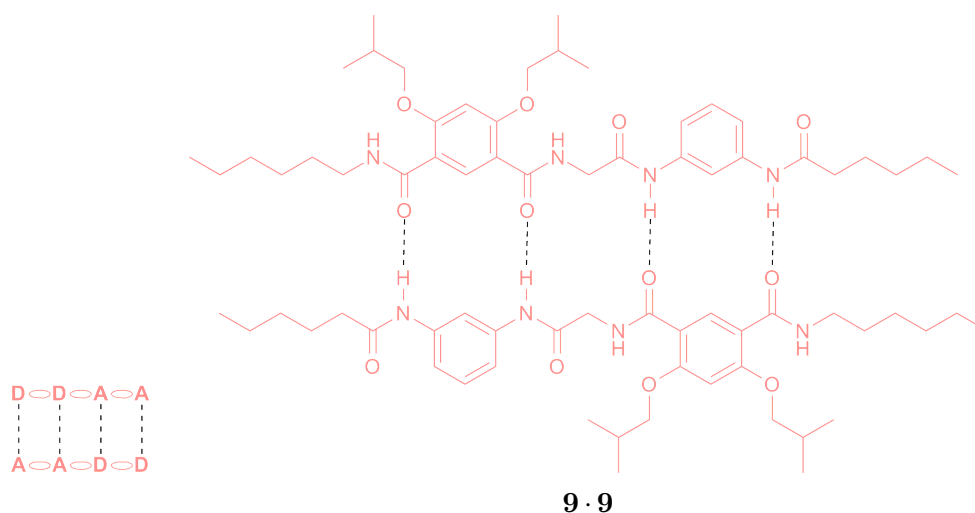


Figure 1.8: Spacers in molecules allow the binding strength of dimers to remain the same, no matter what order the array is in, because there are no secondary electrostatic interactions present.

Spacers can not only negate secondary electrostatic interactions, allowing any combination of hydrogen-bond array to be used, but can also be used to prevent homodimerisation occurring.⁵⁵ Where three hydrogen bonds are used, none of the arrays are self-complementary (Figure 1.9a), but when four hydrogen bonds are present there are two combinations (AADD and ADAD) that result in homocomplexes (Figure 1.9b). By

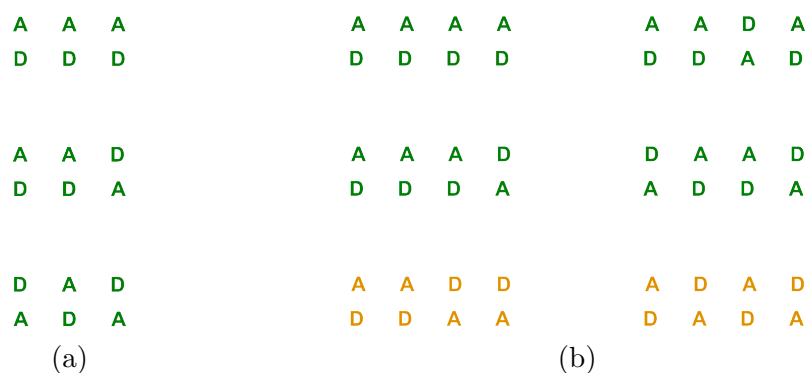


Figure 1.9: (a) Where three intermolecular hydrogen bonds are present six different arrays are possible, none of them are self-complementary. (b) When four intermolecular hydrogen bonds are used there are ten possible arrays. Because two are self-complementary six different complexes can be assembled.

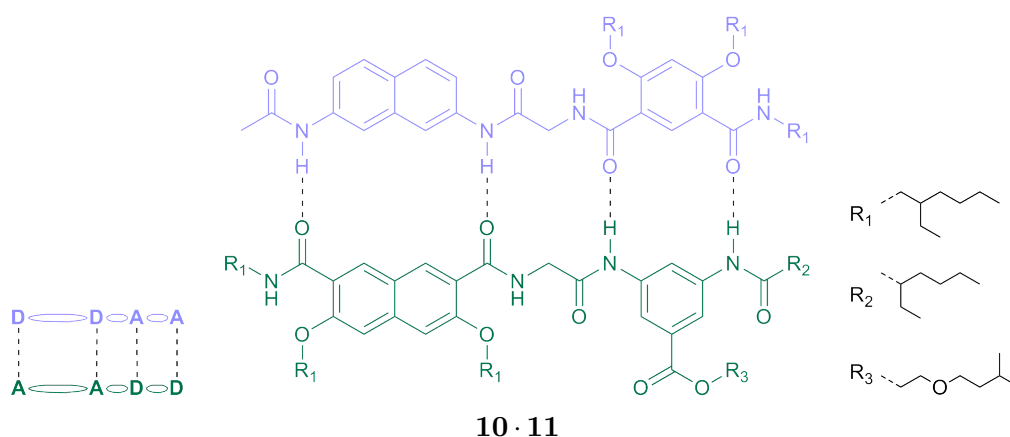


Figure 1.10: Homodimerisation of AADD arrays can be prevented by using different sized spacers.

altering the size of the spacer (**10** and **11**) AADD arrays ($K_a = 2 \times 10^4 \text{ M}^{-1}$) can be synthesised that are not self-complementary (Figure 1.10).

1.1.3 Conformer Considerations

Amide/urea functionalities are a common feature of many hydrogen-bonding motifs because they can act as both hydrogen bond D's and A's and they can normally be synthesised relatively easily. A detrimental feature of these types of functional groups however, is that they can adopt many alternative conformations, which can result in a lower than expected binding affinity because of the loss of entropy that occurs when the molecule adopts the correct binding conformation. One strategy to prevent molecules from adopting alternative conformations, ensuring that the desired hydrogen-bond array is always presented, has been to use covalent linkers.^{56,57} By forming a macrocycle, bonds

are prevented from rotating, so undesired conformations cannot be adopted. Chang and Hamilton⁵⁸ showed the importance of preorganisation by measuring the differences in K_a between a molecule that was free to adopt different conformers (**12**) and one that was locked in a macrocyclic ring (**13**) with 5,5-diethylbarbital (DEB) **14**. They found that the K_a was increased by 100-fold, from 10^4 M^{-1} to 10^6 M^{-1} respectively (Figure 1.11), simply by forcing all of the hydrogen bonding sites to be positioned in the desired orientation.

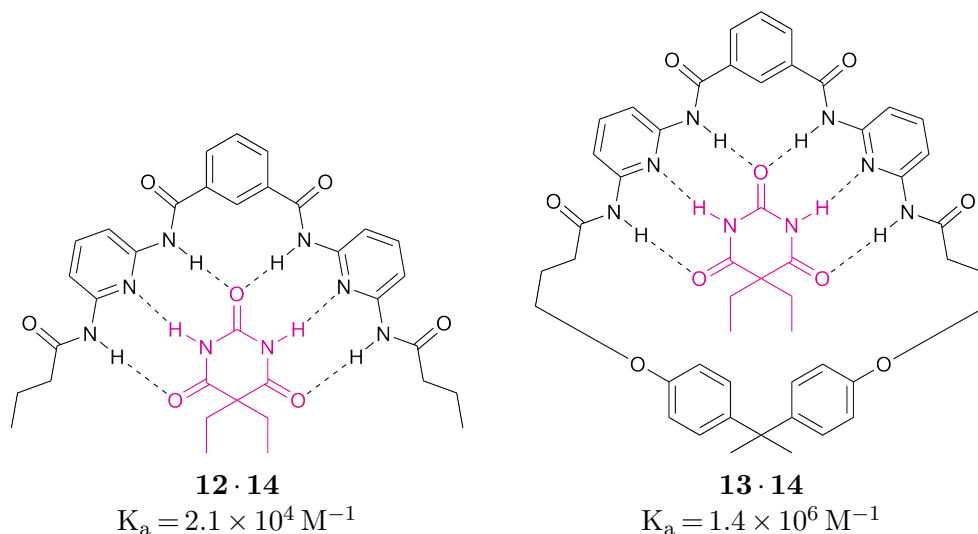


Figure 1.11: Adding a covalent linker prevents different conformers being present, allowing stronger interactions to be obtained.

In order to simplify the design and synthesis, hydrogen-bonding motifs often contain a heteroaromatic ring, as well as amide/urea functionalities. Because heteroaromatic rings are rigid and planar the position of the hydrogen-bonding atoms can be predicted, allowing complementary partners to be designed. In addition to making the design and synthesis easier, rigid molecules, if designed correctly, will have stronger binding affinities than non-rigid molecules because the entropic penalties on binding are reduced (Equation 1.3).

Although there are many advantages to using heteroaromatics, they can sometimes form intramolecular hydrogen bonds with adjacent amide/urea functionalities. This can often lead to undesired hydrogen-bond arrays being presented by the molecule, leading to the assembly of undesired complexes. The most likely conformation of the molecule can be predicted using Etter's rules,⁵⁹ which state that:

1. all good proton donors and acceptors are used in hydrogen bonding;

2. six-membered-ring intramolecular hydrogen bonds form in preference to intermolecular hydrogen bonds;
3. the best proton donors and acceptors remaining after intramolecular hydrogen bond formation form intermolecular hydrogen bonds to one another.

Phenylureidopyridine **15** was originally designed to present an ADD array **15i**. However, Etter's rules predicted that an intramolecular hydrogen bond would form first and a DA array **15ii** would then be presented (Figure 1.12). This explains the lower than expected binding affinity ($K_a = 84 \text{ M}^{-1}$)⁶⁰ obtained between **15** and complementary arrays.^{61,62}

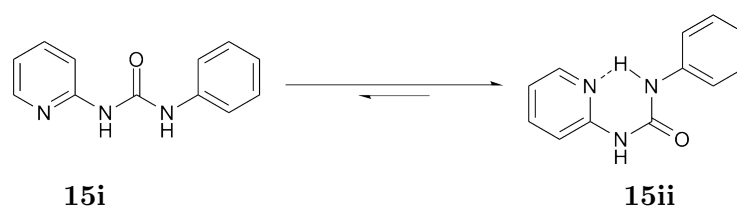


Figure 1.12: Intramolecular hydrogen bonding preorganises the molecule to present an undesired array, so intermolecular hydrogen bonding is not possible.

Careful design of hydrogen-bonding motifs can allow intramolecular hydrogen bonds to preorganise the motif for intermolecular interactions, rather than hinder them. The Wilson group have designed two molecules that present the desired hydrogen-bond array even when intramolecular hydrogen bonds are present.^{60,63} The first molecule exchanged the pyridine ring in **15** for a pyrimidine to give phenylureidopyrimidine (PUPY) **16**.⁶³ The presence of two nitrogens in the ring allowed the desired ADA hydrogen-bond array to be presented even when an intramolecular hydrogen bond was formed. Since either nitrogen of the pyrimidine can form the intramolecular hydrogen bond (Figure 1.13), PUPY **16** is a conformer independent array. A conformer independent array is one that presents the desired hydrogen-bond array, no matter which conformer is adopted. The second molecule designed by the Wilson group, where the pyridine ring in **15** was

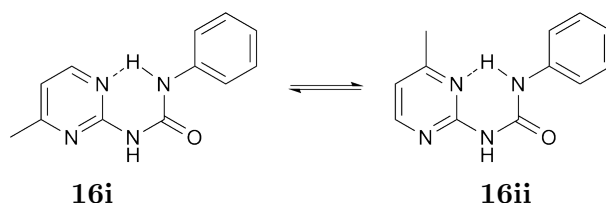


Figure 1.13: Exchanging pyridine ring with a pyrimidine allows an ADA array to be permanently presented.

replaced with an imidazole to give phenylureidoimidazole (UIM) **17**,⁶⁰ is also a conformer independent array (Figure 1.14). The K_a of UIM **17** with complementary arrays is greatly increased when compared to that of phenylureidopyridine **15** ($K_a = 8400 \text{ M}^{-1}$ and 84 M^{-1} respectively) because it presents the desired DDA array in both conformations.

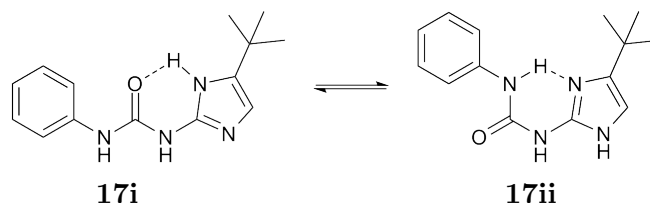


Figure 1.14: Exchanging pyrimidine for an imidazole gives UIM **17** and allows a DDA array to be presented in both conformations.

One exception to Etter's rule is ureidocytosine (UCyt) **18** designed by Hailes *et al.*⁶⁴ According to Etter's rules, intramolecular hydrogen bonds should be formed first, before intermolecular hydrogen bonds can form. However, it was found that UCyt **18** assembled through the self-complementary AADD array, which did not contain any intramolecular hydrogen bonds. It was proposed that even though both homodimers **18i** and **18ii** contained four hydrogen bonds, the unfolded conformation **18i** was preferred because it formed four hydrogen bonds, plus two weak $\text{CH} \cdots \text{O}$ interactions (Figure 1.15).

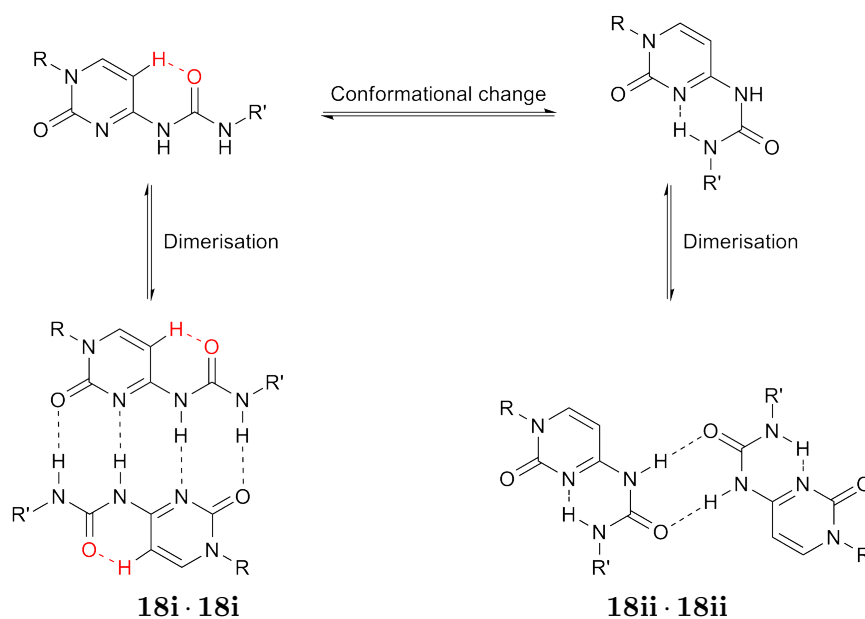


Figure 1.15: The two conformations that UCyt **18** can adopt. Both are self-complementary, but the AADD array is favoured, probably due to the additional weak $\text{CH} \cdots \text{O}$ interaction that forms. R denotes generic alkyl chain.

1.1.4 Tautomer Considerations

A further complication that can arise from using heteroaromatics is their ability to adopt different tautomers, possibly preventing the desired hydrogen-bond array from being presented. Ureidopyrimidine (UPy) **19**³⁴ was designed by Meijer *et al.* to self-associate through an AADD array. However, it was found to adopt many other tautomeric states, some of which did not present self-complementary arrays and one which self-associated through an ADAD array (Figure 1.16). The AADD array was normally more prevalent than the ADAD array, probably because it contained fewer destructive secondary electrostatic interactions than the ADAD array.³⁴ It was possible to control which tautomer was the most prevalent, simply by changing the electron donating/withdrawing nature of the substituents (Figure 1.17).³⁴ Careful choice of electron donating/withdrawing groups made it possible to change the electron density of the atoms involved in hydrogen bonding and so control which array was preferred.

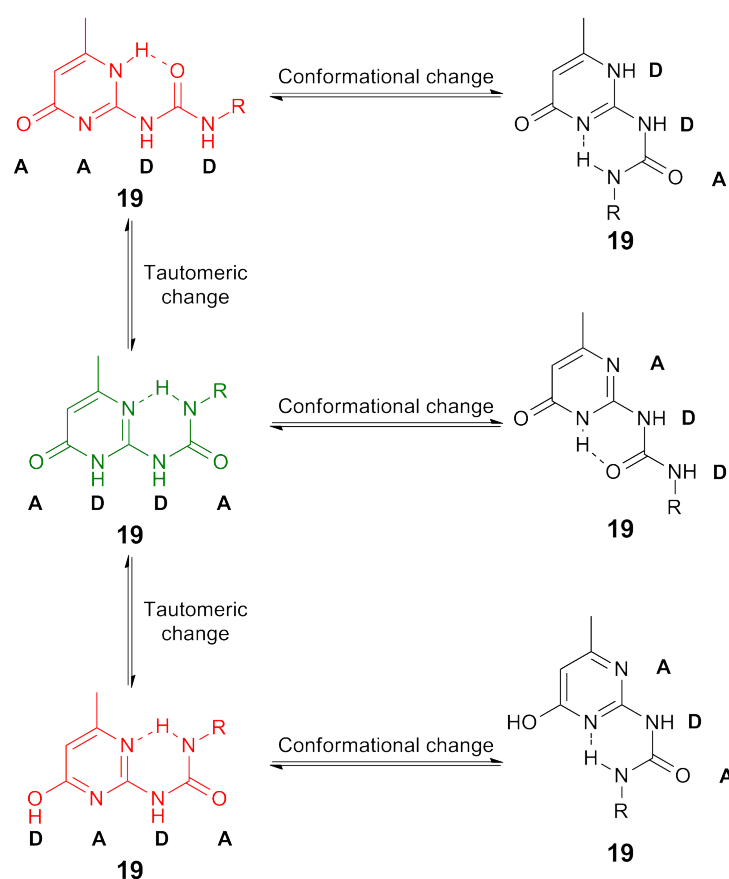


Figure 1.16: All the possible tautomers and conformers that UPy **19** can adopt whilst still maintaining an intramolecular hydrogen bond. Red molecules are self-complementary and green present an ADAD array. R denotes generic alkyl chain.

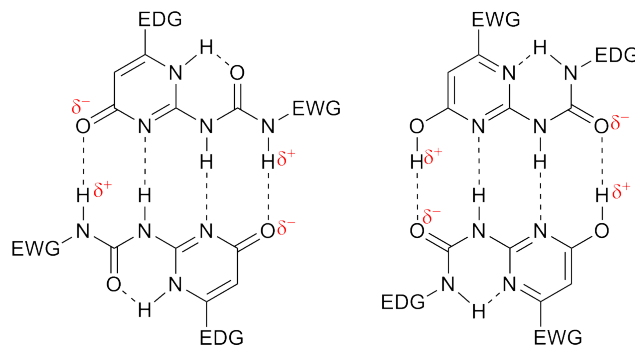


Figure 1.17: Changing the properties of the substituents can alter the preference of the tautomer.

Although UPy **19** has a strong dimerisation constant ($K_{\text{dim}} \approx 6 \times 10^7 \text{ M}^{-1}$),^{65,66} it was thought that it could be possible to increase it by ensuring that all possible tautomers presented self-complementary arrays. Zimmerman *et al.* designed deazapterin (DeAP) **20** ($K_{\text{dim}} = 9 \times 10^7 \text{ M}^{-1}$) in an attempt to test this hypothesis,⁶⁷ which does indeed have a slightly stronger K_{dim} than UPy **19**. Although more of the possible tautomer/conformers presented self-complementary arrays in DeAP **20** than UPy **19**, some tautomers still presented undesired DAAD and AADA arrays (Figure 1.18) and these can be exploited in order to form heterodimers.

Diamidonaphthalene (DAN) **21** was designed to form ADDA · DAAD complexes with both UPy **19**⁶⁸ and DeAP **20**⁶⁷ (Figure 1.19). It was found that, even though the heterodimers contained more destructive secondary electrostatic interactions, they were formed in preference to the homocomplexes. The rationale for this was that, even though the strength of interaction was lower in the heterocomplexes compared to the homocomplexes, more hydrogen bonds could be formed when the heterocomplexes were assembled, making it thermodynamically favourable. For example, in the UPy **19** homodimer there were six hydrogen bonds. When the UPy **19** dimer dissociated it could assemble into two complexes with DAN **21**, which each contained five hydrogen bonds. So the system went from six hydrogen bonds in the homodimer, to ten hydrogen bonds in the heterodimer.

Even though both UPy **19** and DeAP **20** interacted with high affinity to DAN **21**, there was a lack of fidelity^{69,70} due to the possibility of the formation of homodimers. In order to make the heterocomplex with DAN **21** as favourable as possible, Park *et al.* sought to design a molecule that presented the ADDA array through a tautomer that was low in energy and the self-complementary AADD array through a tautomer that was

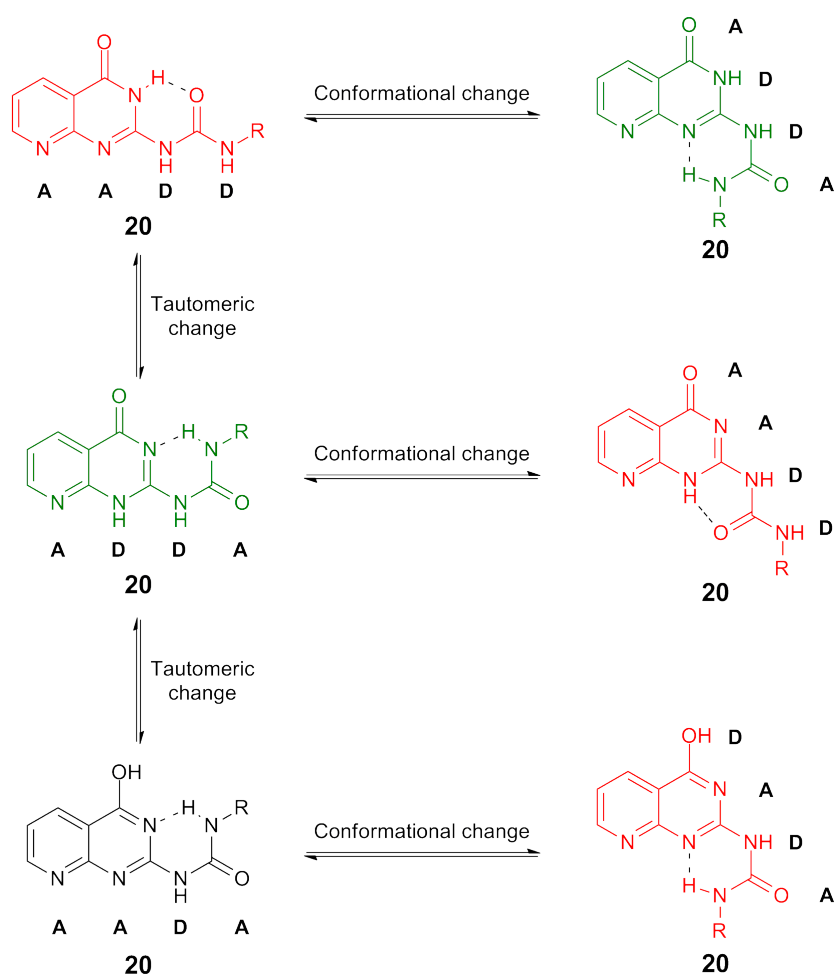


Figure 1.18: All the possible tautomers and conformers that DeAP **20** can adopt whilst still maintaining an intramolecular hydrogen bond. Molecules that are coloured red are self-complementary and green coloured molecules present an ADDA array. R denotes generic alkyl chain.

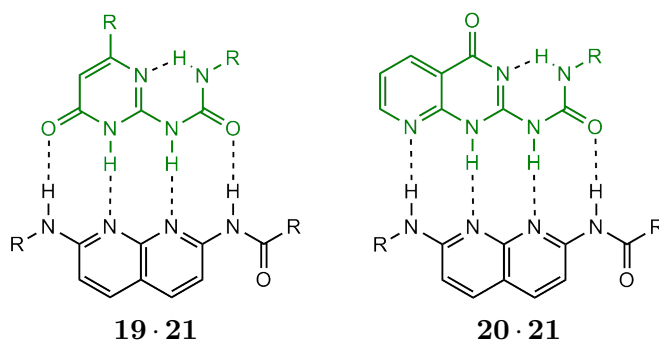


Figure 1.19: DAN **21** can form heterodimers with UPy **19** and DeAP **20**. R denotes generic alkyl chain.

higher in energy.³⁷ Using molecular modelling they designed ureidoguanine (UG) **22**, which presented the ADDA array in a tautomer that was 15 kcal mol^{-1} lower in energy than the tautomer that presented an AADD array (Figure 1.20). Indeed, UG **22** has a much lower dimerisation constant than UPy **19** ($K_{\text{dim}} \approx 200 \text{ M}^{-1}$ and $K_{\text{dim}} \approx 6 \times 10^7 \text{ M}^{-1}$ respectively). It was found that strong heterodimers ($K_{\text{a}} \approx 5 \times 10^7 \text{ M}^{-1}$) were formed between UG **22** and DAN **21**.

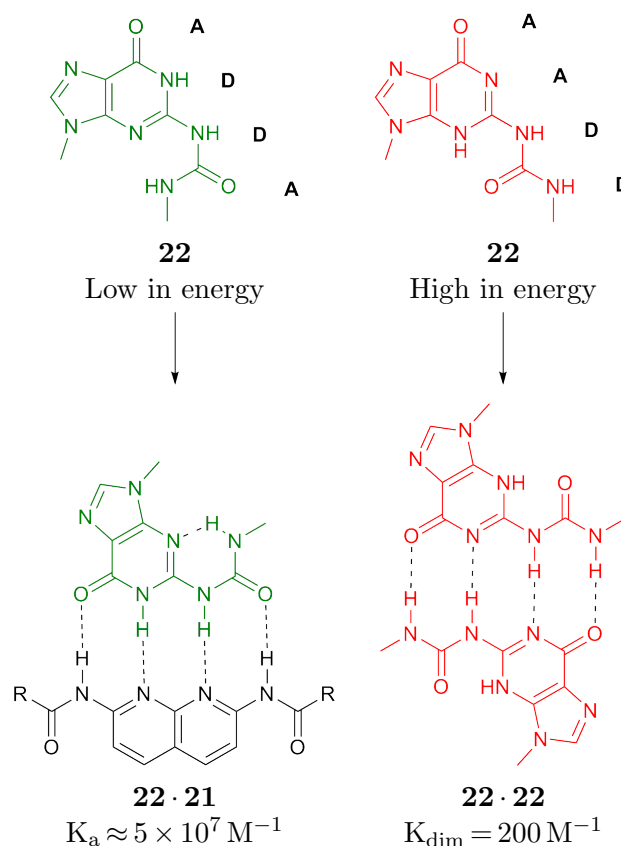


Figure 1.20: UG **22** presents an ADDA array through a low energy tautomer and the AADD array in a tautomer that is high in energy. R denotes generic alkyl chain.

1.2 Self-Assembly into Designed Architectures

The assembly of dimers is only part of the challenge. In order to obtain functionality, molecules must be programmed so that higher order architectures can be assembled. This can lead to applications such as; purification,⁷¹ detection of contaminants,⁷² catalysts^{73–76} and in the recognition of DNA mismatches.^{77–83} Careful design of molecules can allow them to assemble into two- and three-dimensional structures.^{84,85} The literature on hydrogen bond mediated assembly is already vast,⁸⁶ hence only selected examples that have relevance to the goals of this project will be discussed here.

1.2.1 Defined Two- and Three-Dimensional Architectures

It has been shown that guanosine **23** is templated by the addition of ions⁸⁷ to form macrocycles.^{88,89} Guanosine **23** alters its binding mode, switching from traditional Watson and Crick interactions to Hoogsteen interactions, which results in the formation of a 2-D disk (Figure 1.21a). These 2-D disks then assemble, using hydrophobic interactions, into 3-D stacks (Figure 1.21b). If hydrogen-bonding motifs can be designed so that they have two hydrogen bonding faces then the ability to assemble into 2-D disks and possibly 3-D stacks should be possible. For 2-D disks to be formed the molecules must be the correct shape and the correct hydrogen-bond array must be chosen so that only the desired macrocycle is obtained. For example, when symmetrical DAD · ADA arrays are

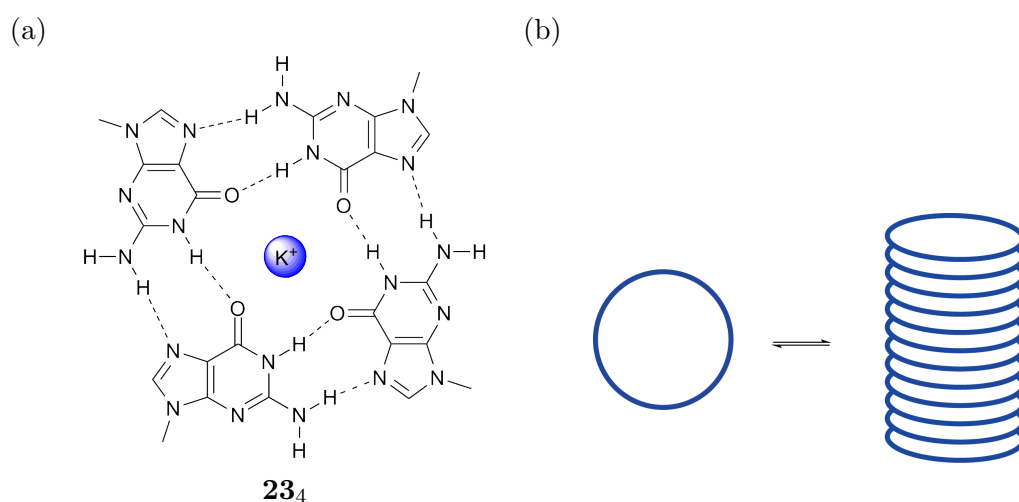


Figure 1.21: (a) Guanosine **23** can assemble into macrocycles in the presence of ions. (b) These two-dimensional disks can then form three-dimensional stacks using hydrophobic interactions.

used there is a possibility of extended tapes being formed (Figure 1.22),^{90,91} but by using unsymmetrical AAD · DDA arrays, 2-D disks are the only product that can be assembled (Figure 1.23).⁹² When the two arrays in a single molecule were complementary (**24**), 2-D disks were assembled that only contained one type of molecule (Figure 1.24a).⁹²⁻⁹⁴ However, when the two hydrogen-bond arrays were not complementary (**25** and **26**), a second molecule had to be added, which allowed the formation of 2-D disks that contained a mixture of molecules (Figure 1.24b).⁹⁰ As in discrete linear arrays, the binding affinity of 2-D disks has been increased by including additional secondary electrostatic interactions **27** (Figure 1.25),⁹⁵⁻⁹⁷ or by increasing the number of intermolecular hydrogen bonds.^{98,99} Subtle changes in the shape of the motifs can change the assembly that is obtained. Petersen *et al.* have shown that by using a mixture of six-membered (**28**) and five-membered (**29**) rings, the assembly no longer formed 2-D disks, but instead assembled into helical structures (Figure 1.26).¹⁰⁰

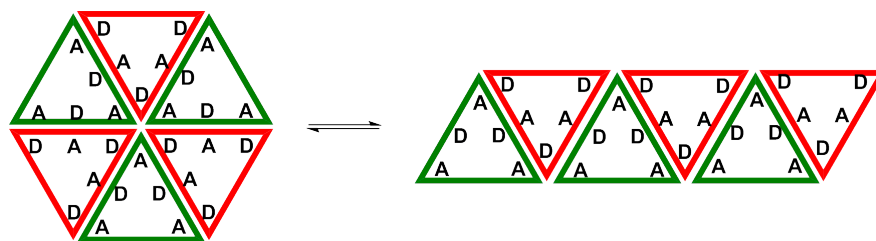


Figure 1.22: An equilibrium between rings and tapes can be established if symmetrical hydrogen-bond arrays are used.

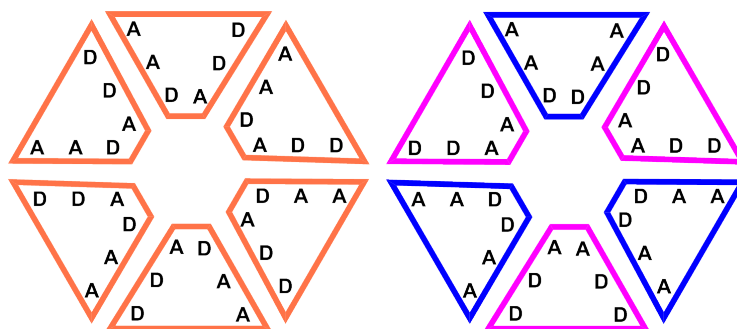
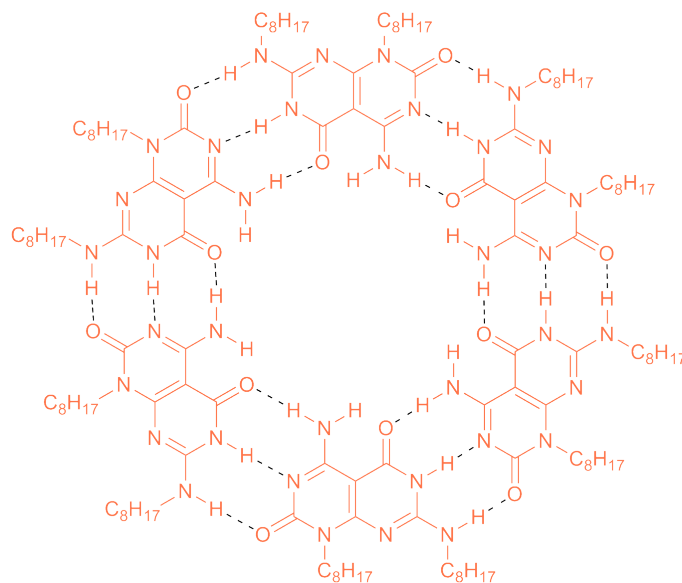


Figure 1.23: Homo- and heterocomplementary hydrogen-bonding motifs can be designed so that they can only form macrocycles.

(a)

 $(24)_6$

(b)

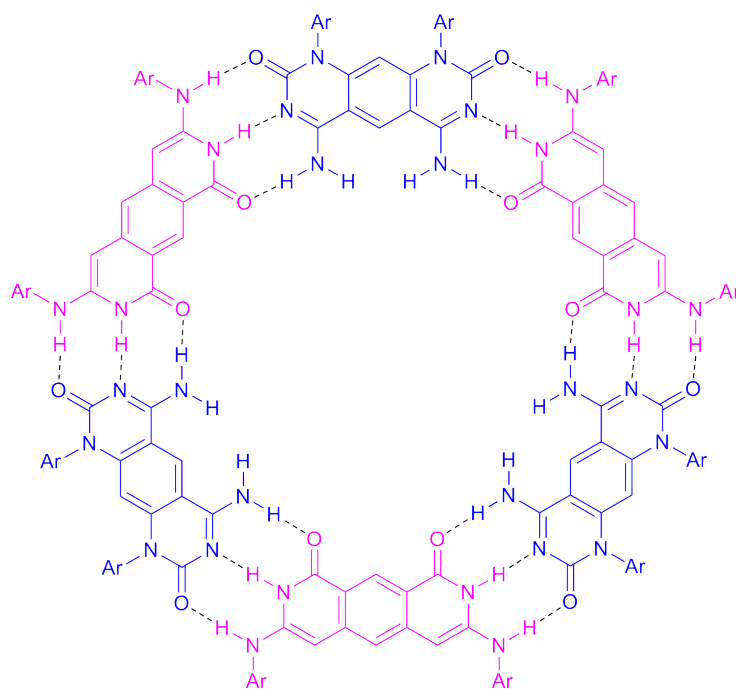
 $(25 \cdot 26)_3$

Figure 1.24: Macrocycles can be assembled using (a) homo or (b) hetero arrays.

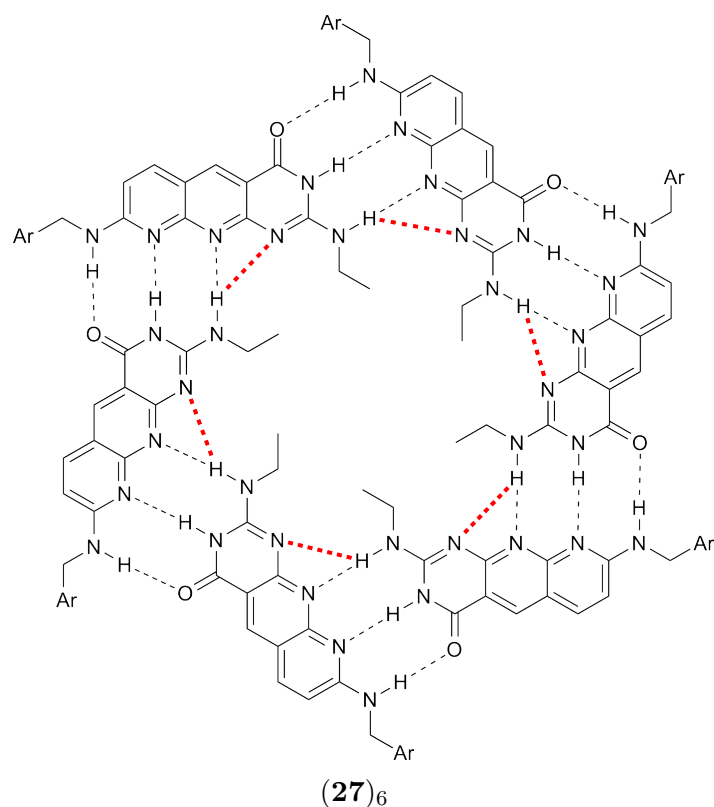


Figure 1.25: The binding affinity of macrocycles can be increased by using additional secondary electrostatic interactions.

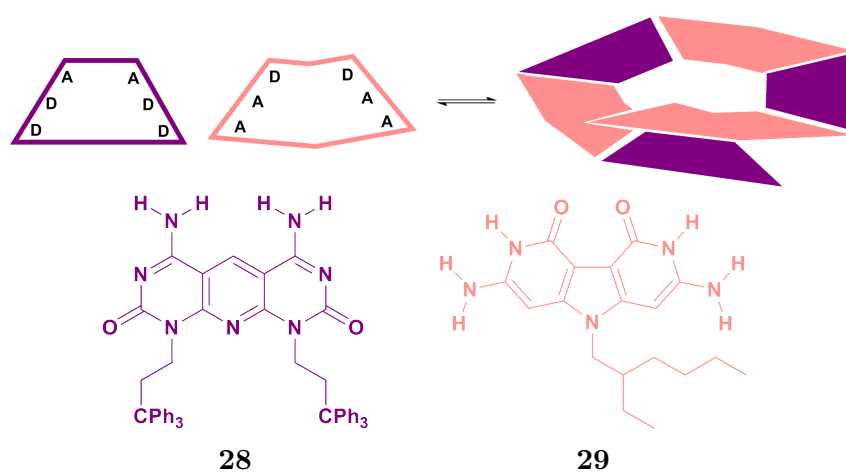


Figure 1.26: Slightly changing the shape of the supramolecular motifs can result in a dihelical assembly being obtained.

In the previous examples, intermolecular hydrogen bonds were used to form the 2-D disks, which could assemble into 3-D stacks using π - π interactions. By designing 2-D molecules, 3-D stacks can be assembled where hydrogen bonds are formed between the disks. These molecules must orientate hydrogen bond D's and A's above and below the plane of the molecule. This can result in a self-complementary molecule that forms 3-D stacks, which can then self-assemble into larger architectures (Figure 1.27a). The first examples of three-dimensional stacks had cyclohexane rings at their core **30** (Figure 1.27b),¹⁰¹ but these molecules required aromatic side chains to allow π - π stacking to stabilise the assembly. More recently, phenyl rings have been used in the core **31** (Figure 1.27c) because they increase stack stability through π - π stacking in the core and allow the side chains to be varied.¹⁰² By altering the side chains many new properties can be obtained, for example different gelation properties,^{103,104} enhanced polymer formation^{105,106} and the ability to control the chirality of the self-assembled 3-D stack.¹⁰⁷⁻¹¹⁴

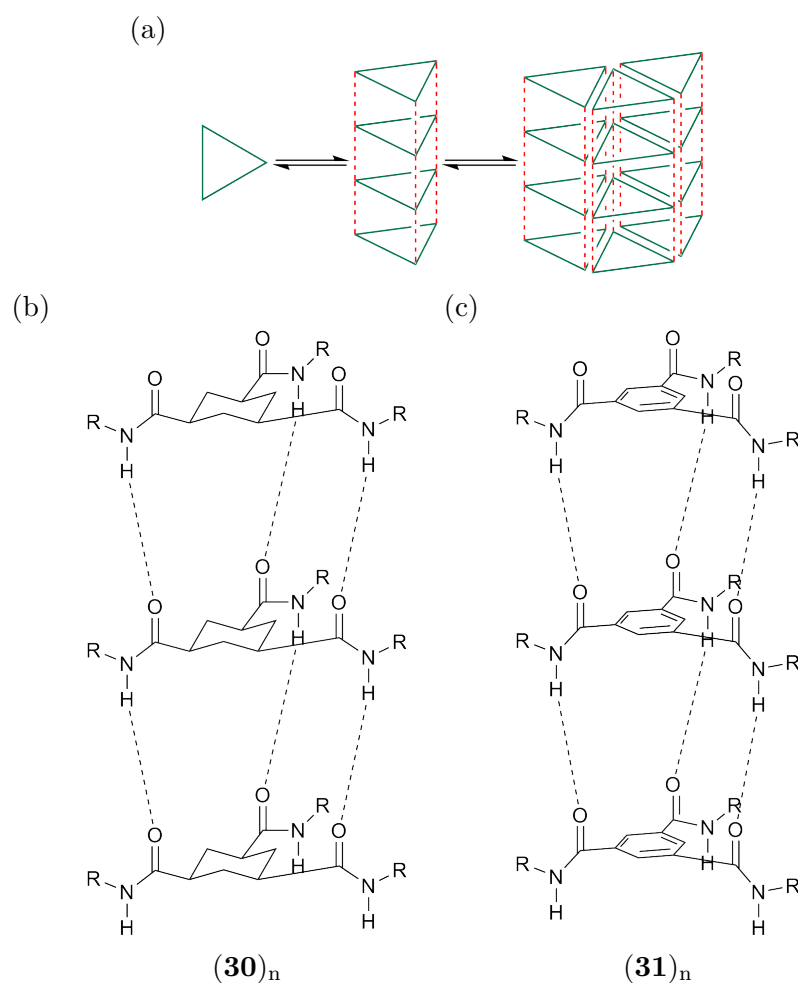


Figure 1.27: (a) A cartoon representation of a 3D stack which can be built from (b) a cyclohexane or (c) a benzene core.

1.2.2 Polymer Assembly

If molecules are designed with two hydrogen bonding faces, which are not oriented to allow the assembly of 2-D disks, extended polymeric structures can be obtained. An early example was from the Lehn group, where the two hydrogen bonding arrays were separated by a rigid, aromatic core.^{115,116} Mixtures of two different molecules (**32** and **33**) were made by ensuring that the individual molecules were not self-complementary (Figure 1.28). More recently flexible linkers have been used between the hydrogen bonding motifs.⁸⁴ Again, these arrays can have self-complementary hydrogen-bonding motifs at each end of the ditopic array (Figure 1.29a),¹¹⁷ or they can be designed so that a complementary component must be added for polymerisation to occur (Figure 1.29b).^{118,119} Because the linker is flexible, a concentration dependent equilibrium between ring and chain

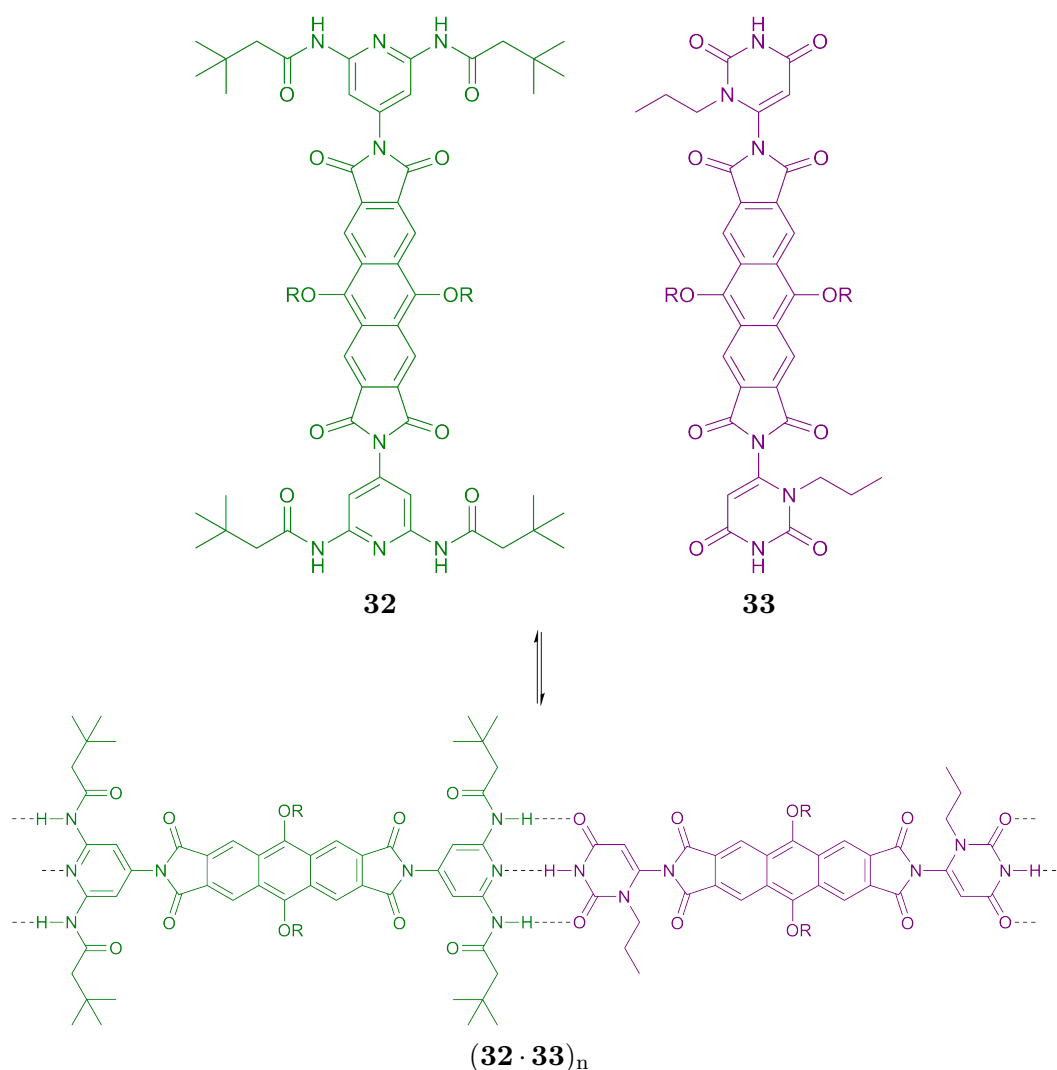


Figure 1.28: Rigid rods designed by the Lehn group present two hydrogen-bond arrays at either end of a rigid molecule.

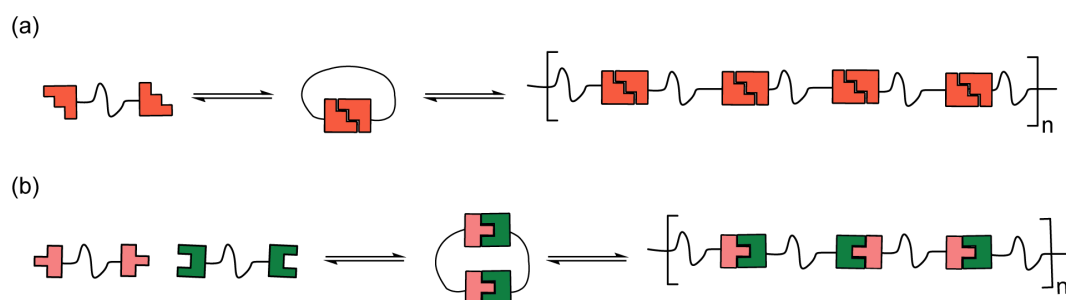


Figure 1.29: Cartoon representation of ditopic (a) homo and (b) hetero hydrogen-bonding motifs.

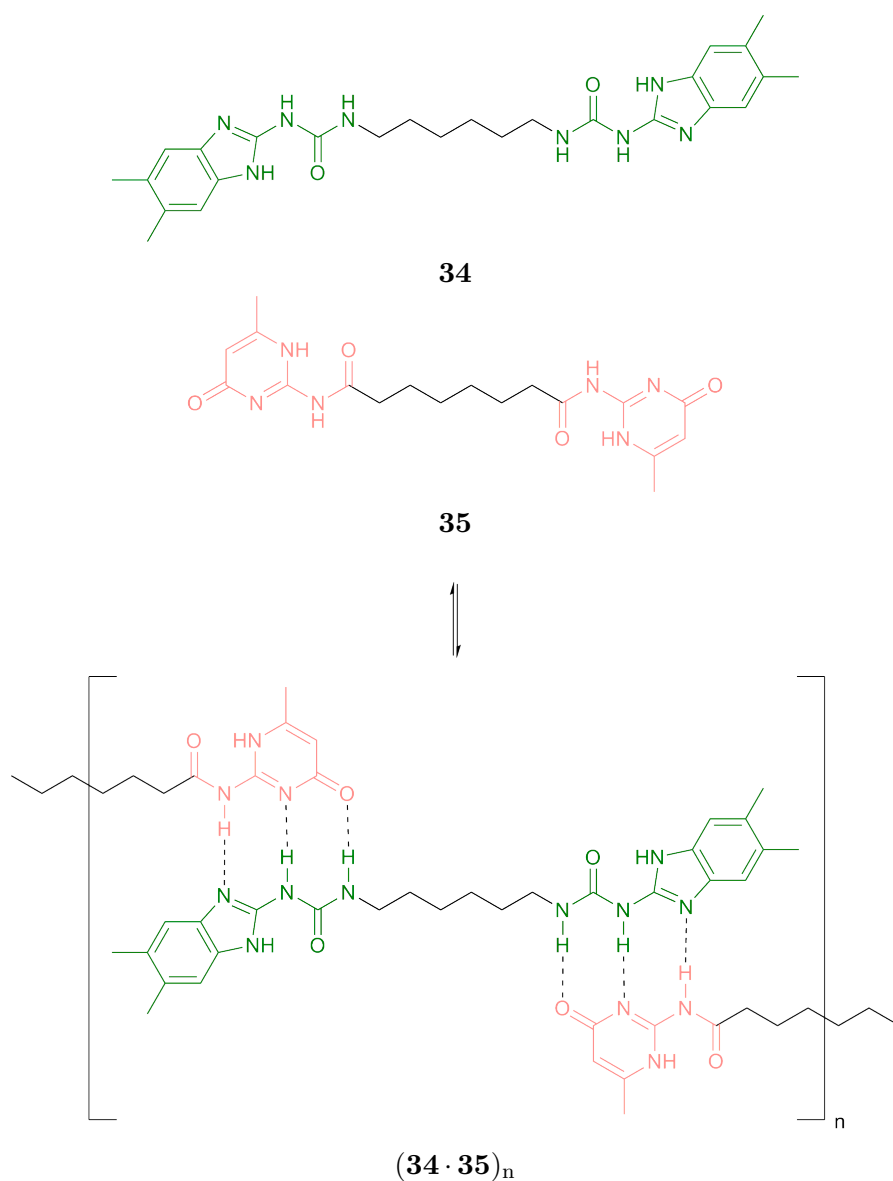


Figure 1.30: The ditopic triply hydrogen bonded heterodimers designed by the Wilson group.

formation is commonly established.^{120,121} The Wilson group have designed molecules **34** and **35** which can assemble into cyclic dimers and polymeric structures (Figure 1.30).¹²² Neither molecule was self-complementary, so they both had to be present in order to form extended assemblies.

Hydrogen-bonding motifs have also been attached to the side chains of polymers (Figure 1.31).¹²³ Appending the polymerising group to the hydrogen-bonding motif allows it to be integrated into the polymer during polymerisation.^{124–128} The addition of supramolecular motifs to the backbone of polymers can allow materials to have new properties. Supramolecular polymers have self-healing properties,¹²⁹ are stimuli responsive,¹³⁰ and often have lower melting temperatures than traditional covalent polymers,^{131–133} which can be useful for some applications.

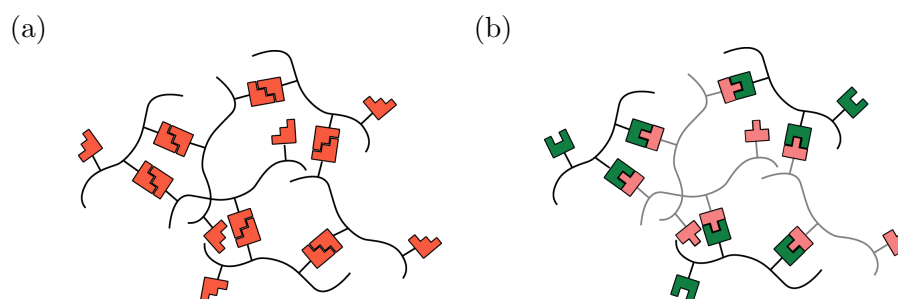


Figure 1.31: Supramolecular polymers that are (a) stimuli responsive and (b) able to form blends of polymers that are not normally miscible.

1.2.3 Self-Assembling Capsules

Careful integration of hydrogen-bonding motifs into rigid, 2D structures can allow the assembly of 3D capsules.¹³⁴ Rebek introduced the “tennis ball” capsule in 1993,¹³⁵ which was assembled from a self-complementary motif **36** (Figure 1.32). The existence of the capsule was confirmed by vapour pressure osmometry (VPO), which calculated the mass as being a dimer plus a solvent molecule, which was presumably trapped in the capsule.

A second capsule that can reversibly encapsulate guests¹³⁶ is resorcinarene **37** (Figure 1.33),¹³⁷ which assembles into a large capsule using 60 hydrogen bonds. Because the resorcinarene capsules are quite large, they can accommodate large guest molecules.¹³⁸ Another capsule, designed by Reinhoudt, appends hydrogen-bonding motifs to a calixarene to give **38**.¹³⁹ Because the hydrogen-bonding motif is not self-complementary, capsules are only assembled when two equivalents of 5,5-diethylbarbituric acid (DEB) **14**

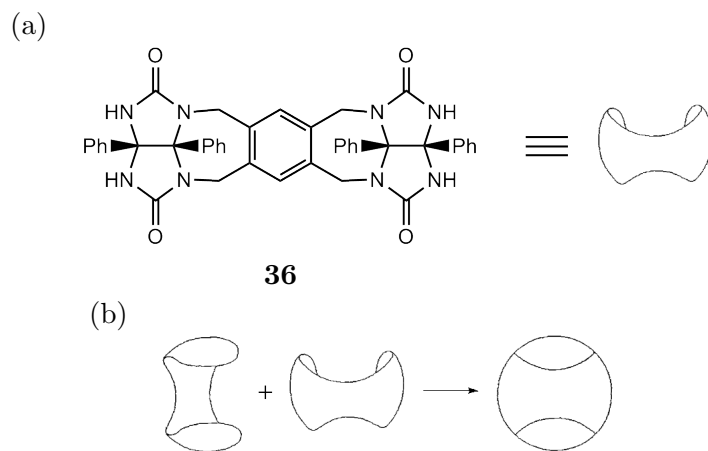


Figure 1.32: (a) The self-complementary hydrogen-bond motif **36** that assembles into (b) the “tennis ball”.

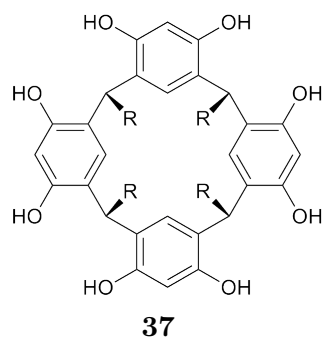


Figure 1.33: The resorcinarene **37** which has been shown to assemble into large capsules that can accommodate guest molecules.

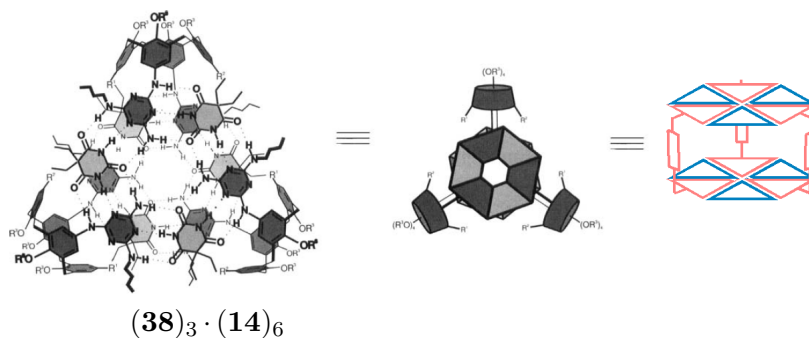


Figure 1.34: The capsule designed by Reinhoudt *et al.* that requires two complementary partners to both be present in order for capsules to assemble.

are added (Figure 1.34). The result is a capsule that is assembled from three calixarene **38** molecules and six DEB **14** molecules.

1.3 Self-Sorting in Hydrogen bonded Systems

Currently, only supramolecular assemblies constructed using one or two molecules have been described. Although these small assemblies can enable materials to have new characteristics, in order to obtain further new properties larger numbers of supramolecular motifs must be used. Many different types of self-sorting systems have been described in the literature,¹⁴⁰ but this thesis will focus on nonintegrative,¹⁴¹ social self-sorting systems,¹⁴² which use hydrogen bonds as the primary interaction. Nonintegrative systems are those that assemble into more than one architecture in solution, integrative systems assemble all of the motifs into a single structure (Figure 1.35). A social self-sorting system is one where molecules assemble into heterodimers and a narcissistic self-sorting system is one where all molecules are self-complementary, so only homodimers are present.

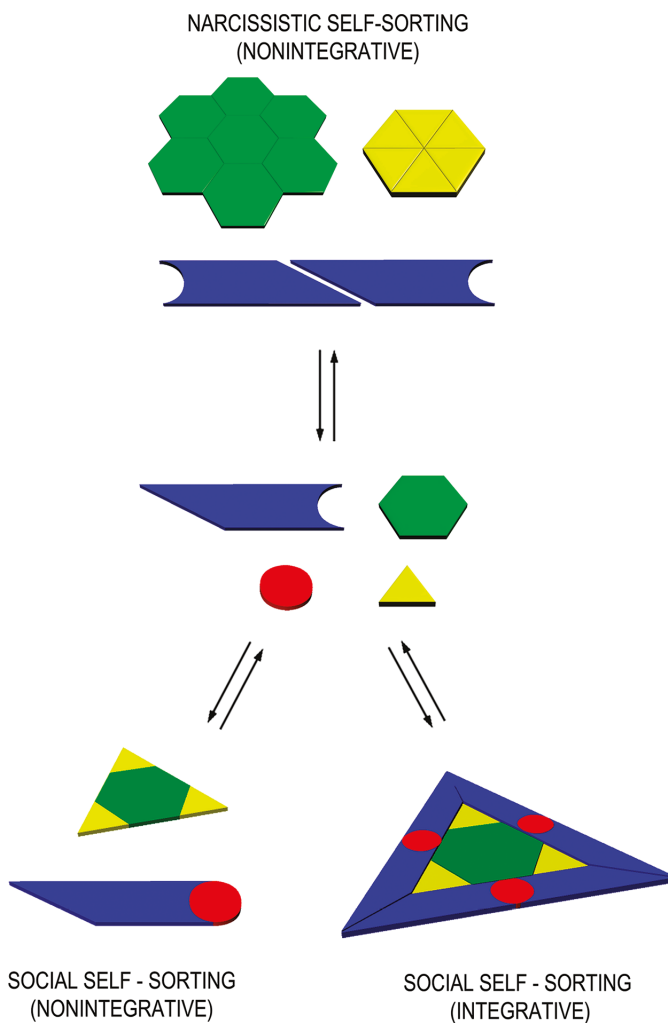


Figure 1.35: A cartoon representation of integrative and nonintegrative, social and narcissistic self-sorting systems.

The previously described calixarene **38** was extended to give a double calixarene **39** architecture, which could also interact with DEB **14** to give the $(\mathbf{39})_3 \cdot (\mathbf{14})_{12}$ complex.¹⁴³ On mixing of **38**, **39** and DEB **14** it could be possible to obtain a number of different homo- and heterocomplexes (Figure 1.36). However, it was found that no heterocomplexes were assembled, only the homocomplexes $(\mathbf{38})_3 \cdot (\mathbf{14})_6$ and $(\mathbf{39})_3 \cdot (\mathbf{14})_{12}$.¹⁴³

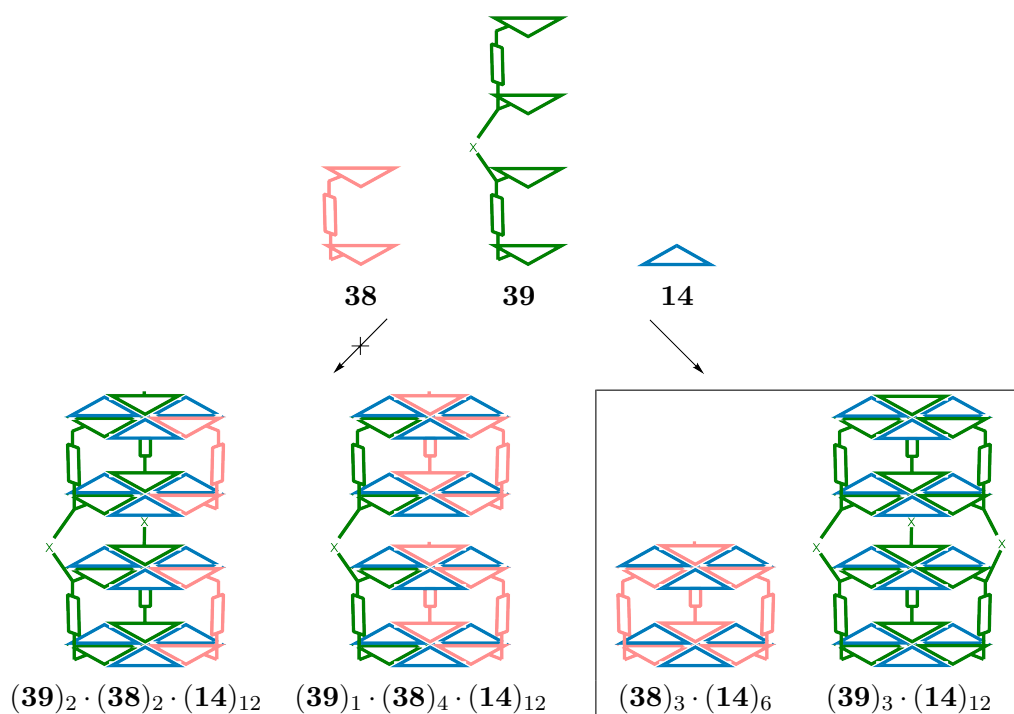


Figure 1.36: Heterocomplexes containing both **38** and **39** are not assembled when **38**, **39** and DEB **14** are mixed. Only the desired homocomplexes $(\mathbf{38})_3 \cdot (\mathbf{14})_6$ and $(\mathbf{39})_3 \cdot (\mathbf{14})_{12}$ are assembled.

Isaacs *et al.* have been able to obtain only homodimers from a mixture of two very similar molecular clips (**40** and **41**). They found that along with differences in hydrogen-bond arrays, shape complementarity and steric hindrances also prevented the assembly of the undesired heterodimers.¹⁴⁴ An extra molecule **42** could be added to this mixture, leading to the narcissitic self-sorting¹⁴⁵ of three very similar molecules (Figure 1.37).¹⁴⁶

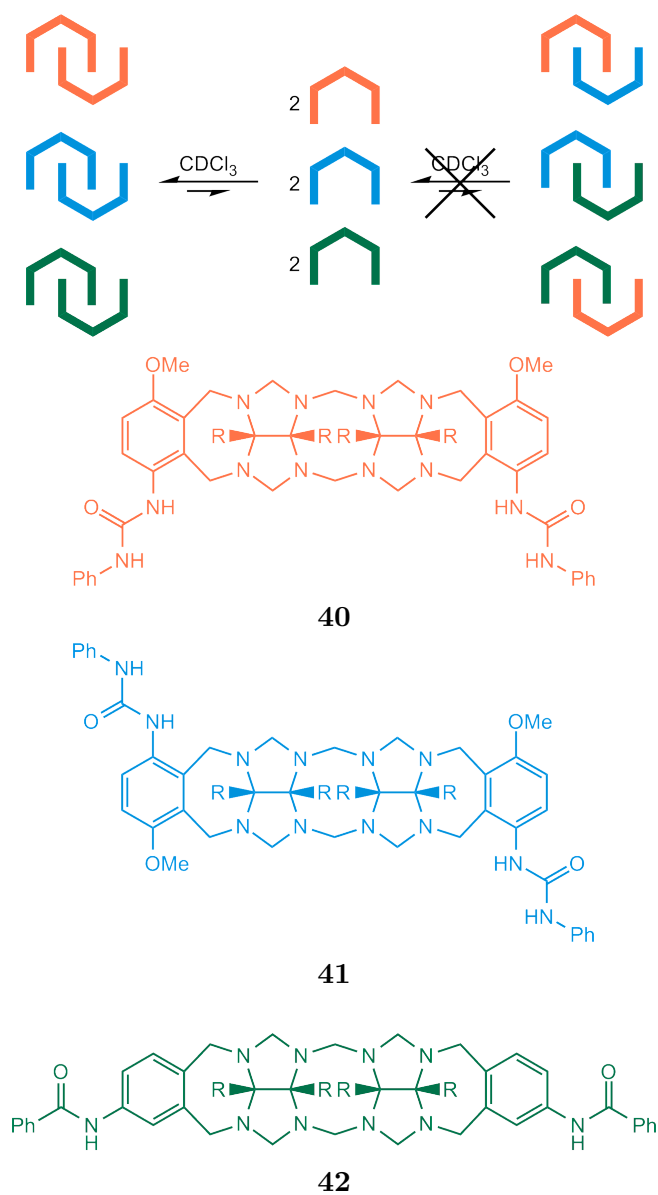


Figure 1.37: The self-sorting mixture proposed by Isaacs *et al.* which assembles into three homocomplexes. The assembly of heterocomplexes was not observed. $\text{R} = \text{CO}_2\text{Et}$.

The self-sorting of polymers is essential in biology for the replication and translation of genetic material.¹⁴⁷ To investigate whether self-sorting in synthetic polymers was possible Burd and Weck incorporated two different recognition motifs into the polymer backbone (Figure 1.38).¹⁴⁷ The two molecules that were chosen were a thymine derivative **3**, which interacts using a linear ADA hydrogen-bond array and cyanuric acid **14**, which interacts using a wedge-like ADAADA hydrogen-bond array. These two molecules have very similar hydrogen-bond arrays, so could both interact with the same molecules, preventing self-sorting processes from occurring. Nevertheless, it was found that on addition of diamidopyridine **2** and isophthalic diamidopyridine **12**, the self-sorting of molecules attached to a polymer backbone was achieved.

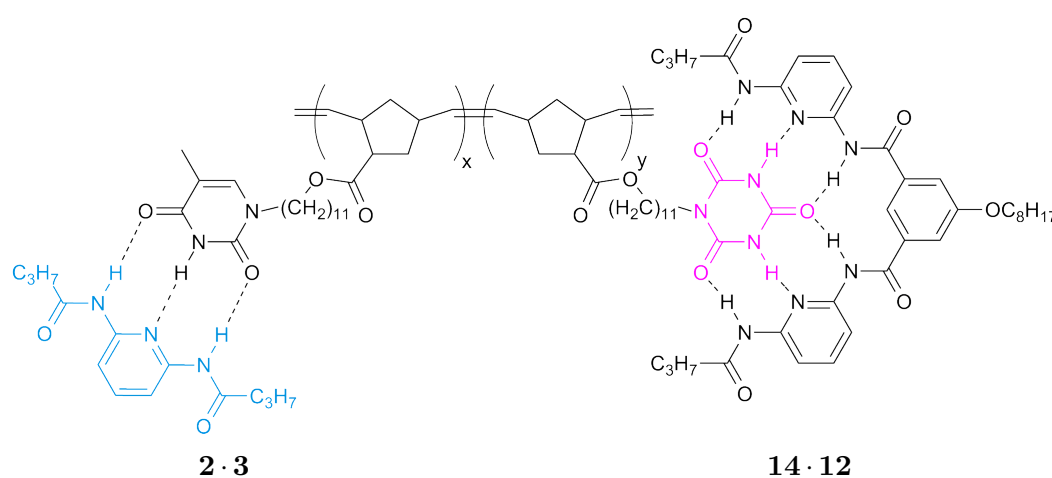


Figure 1.38: Self-sorting of molecules attached to a polymer backbone is possible, even when similar motifs are used.

Since all these examples appear to self-sort without requiring significantly different hydrogen-bond arrays, Isaacs *et al.* were interested in investigating whether these represented special examples, or whether self-sorting was inherent in hydrogen bonded systems. It was found that nine molecules could be mixed, some of which had the potential to assemble into undesired heterocomplexes, but only the desired homocomplexes were obtained (Figure 1.39).⁷⁰ All interactions could be broken and reformed by the addition/removal of DMSO, indicating that the thermodynamic, not the kinetic, product was obtained each time. Investigation into how various stimuli (e.g. temperature, concentration, K_a) could affect the system was carried out. It was found that narcissitic self-sorting can be robust to these external parameters, as long as the K_a of undesired complexes is a factor of ten lower than the K_{dim} of the desired homodimers and equimolar amounts of all components are present. It is expected that these general rules can

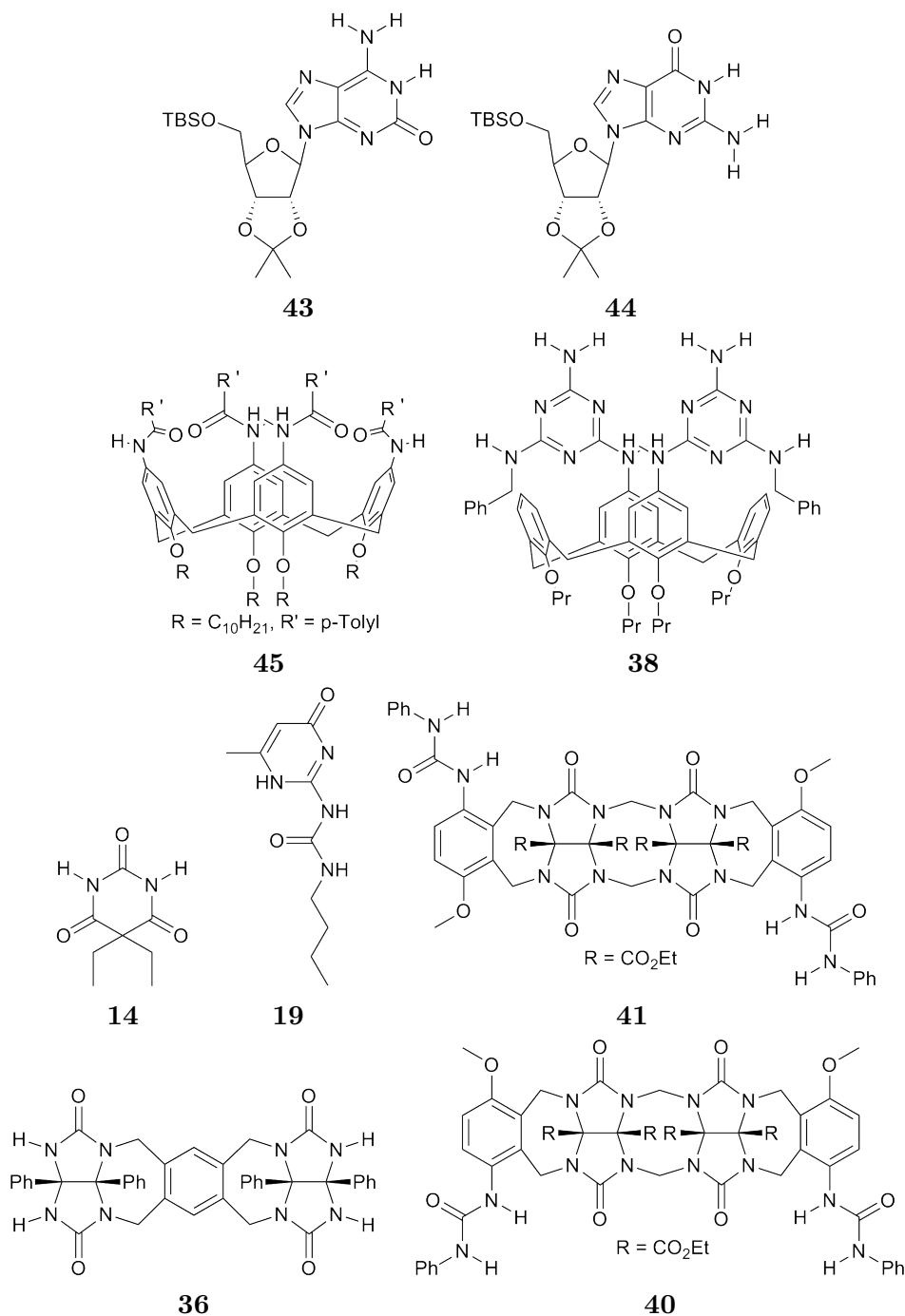


Figure 1.39: The nine molecules that were used to assemble a self-sorting mixture. $(43)_{10} \cdot Ba^{2+} + 2Pic^{-}$, $(44)_{16} \cdot 2Ba^{2+} + 4Pic^{-}$, $45 \cdot 45$, $(38)_3 \cdot (14)_6$, $19 \cdot 19$, $41 \cdot 41$, $40 \cdot 40$ and $36 \cdot 36$.

be reversed when heterocomplexes are desired over the formation of homocomplexes.

These self-sorted systems could lead to many polymer blends being made, allowing many new materials properties to be realised. However, if it could be possible to manipulate

the self-sorting system, allowing components to be altered using external stimuli, applications in information processing and molecular machines could be developed.^{148,149} An example of a system that uses a controlled self-sorting system to carry out the desired function has been described by Pérez *et al.*¹⁵⁰ In this system the ring **46** can be shuttled to the desired position along the rod **47**, using light as the stimulus (Figure 1.40). When the olefin is in the *Z*-configuration the ring **46** selectively interacts at this position; the peptide (red) preferentially adopts an alternative conformation, allowing the formation of an intramolecular hydrogen bond, preventing the interaction of the ring **46**. Converting the olefin to the *E*-configuration prevents the ring **46** from interacting at this position, so it is shuttled along the rod **47** to interact with the peptide. The isomerisation of the olefin is reversible, so the ring **46** can be shuttled along the rod **47** many times. Modification of the rod **47** into a macrocycle gives **48**, which can act as a molecular motor.¹⁵¹ In this system the direction that the ring **46** moves around the macrocycle **48** can be controlled by using a combination of light and chemical stimuli (Figure 1.41).

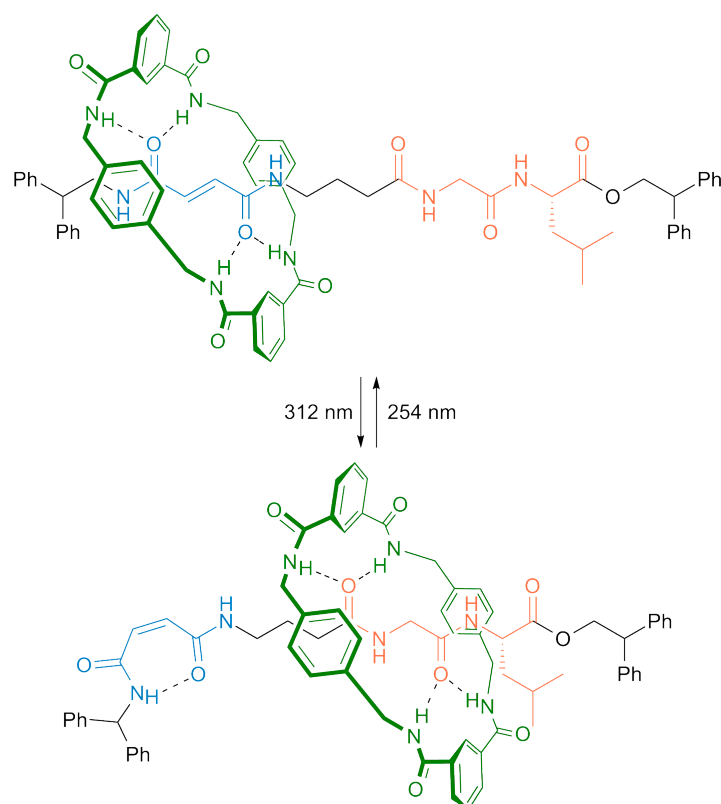


Figure 1.40: The ring **46** can be reversibly shuttled along the rod **47** using light as the stimulus.

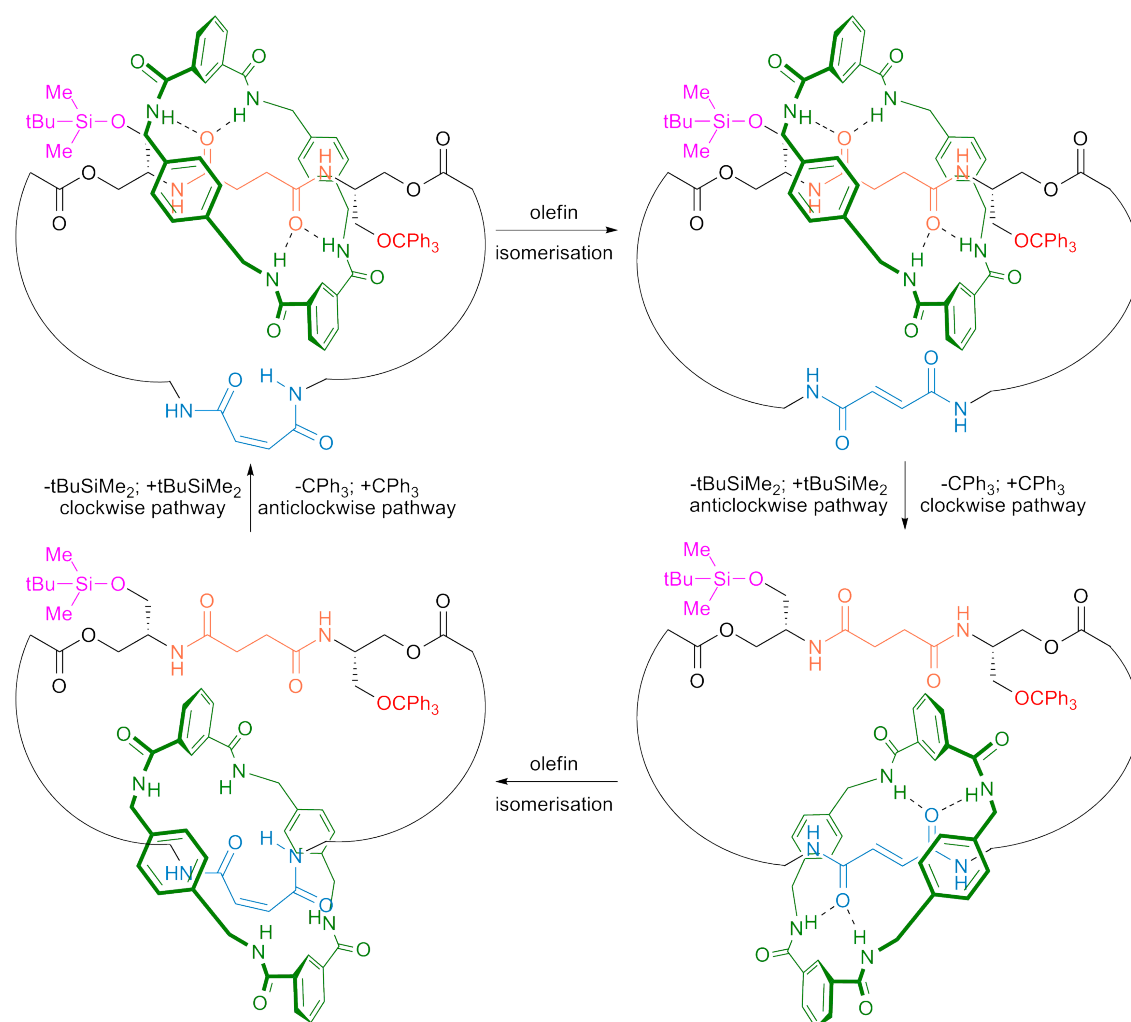


Figure 1.41: A molecular motor **48** can be assembled by using a combination of light and chemical stimuli.

1.4 Project Aims

The above examples illustrate that synthetic molecules can be programmed so that they only assemble into the desired structures, even in the presence of competing motifs. However, the examples that are present in the literature tend to use shape complementarity, alongside different hydrogen-bonding arrays to aid self-sorting, and these molecules can be difficult to synthesise. Ideally, self-sorting systems should be obtained from simple, easy-to-synthesise molecules, allowing complexity to be achieved rapidly. Generally, motifs with triple or quadruple linear hydrogen-bond arrays give strong K_a 's and are easy-to-synthesise. However, when only linear arrays are present in a self-sorting system there is the potential for many undesired complexes to be formed. In order to use linear hydrogen-bond arrays in self-sorting systems greater understanding of the fundamental features of hydrogen-bonding motifs is required. This thesis will focus on the investigation of some of the fundamental features that can confirm the formation of heterodimers; this increased understanding will enable a self-sorting system using only linear hydrogen-bond arrays to be assembled. It has also been shown that external stimuli, for example light, can be used to switch on and off the hydrogen bonding capabilities of some molecules. Investigation into appending a photolabile group onto one of the molecules, allowing a stimuli responsive self-sorting system to be obtained, is also described.

Chapter 2

Effects of Preorganisation

An objective for supramolecular chemists is to obtain complex systems from simple building blocks. Simple and easy-to-synthesise building blocks are desirable and therefore, all factors that can increase the binding affinity whilst keeping the molecules small must be considered. Binding affinity, specifically the association constant (K_a) (Equation 1.2), is dependant on the contributions of enthalpy and entropy (Equation 1.3). Therefore, in order to increase the binding affinity of supramolecular complexes, it is not always necessary to add extra intermolecular hydrogen bonding groups, but simply to rigidify the molecules so that they always present the desired hydrogen-bond array. This should maximise affinities when low numbers of intermolecular hydrogen bonds are used, which could also ease synthesis.

In complexes that interact using intermolecular hydrogen bonds there are many examples of how molecules have been preorganised.^{29,32,34,35,48,67,152–158} One strategy is to use fused rings, forcing the molecules to present the desired array because they cannot adopt alternative conformer states. The fused pyridine molecules (**6** and **8**), developed by Leigh,^{29,35} participate in high affinity interactions (Figure 2.1). This strong binding affinity is due to two reasons; there are no repulsive secondary electrostatic interactions⁴⁵ since the motifs present AAA **6** and AAAA **8** arrays and the molecules are fully preorganised, minimising entropic loss on complex formation.

It is also possible to use intramolecular hydrogen bonds to preorganise molecules for complex formation. There are many examples of molecules that are preorganised through a combination of both rigid aromatic rings and intramolecular hydrogen bonds,^{32,34,48,67,152–157}

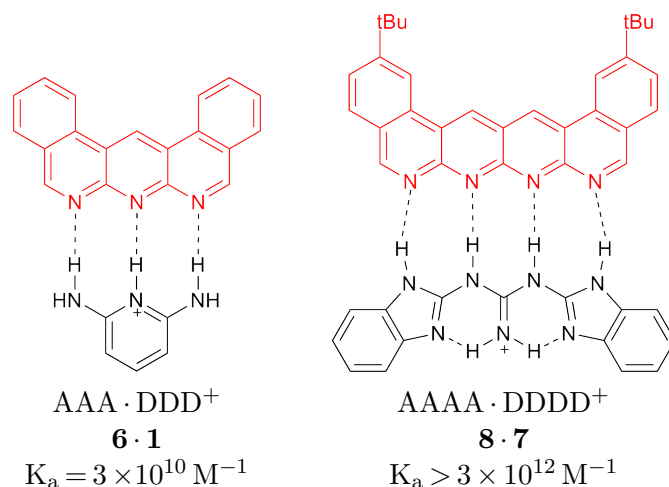


Figure 2.1: Molecules that are preorganised by fused rings (red). No undesired hydrogen-bond arrays can be presented because no alternative tautomer/conformer states can be adopted. All association constants were determined in CH_2Cl_2 .

some examples (**20**, **49**, **50**, **51** and **52**) are shown in Figure 2.2. Care needs to be taken when designing such molecules because they may have the ability to adopt undesired conformer/tautomer states due to the nature of the heteroaromatic and amide/urea functional groups that are used. There have also been examples of hydrogen-bonding motifs that do not contain any aromatic rings. These molecules are preorganised by using only intramolecular hydrogen bonds. This approach was used by Sanjayan *et al.* in order to obtain self-complementary motifs **53 · 53** and **54 · 54** (Figure 2.3).¹⁵⁸

There have been some previous studies on the effect that preorganisation using intramolecular hydrogen bonds has on the binding affinity of complexes. Meijer *et al.* have shown that the binding affinity of DADA · ADAD arrays can be dramatically increased by the addition of an intramolecular hydrogen bond (Figure 2.4).¹⁵⁶ Diamidopyrimidine **55** does not contain any intramolecular hydrogen bonds and has a dimerisation constant (K_{dim}) of 170 M^{-1} . Switching one of the amide groups for urea functionality to give 6-amido-2-ureidopyrimidine **56** allows for the formation of an intramolecular hydrogen bond. The self-association of **56** is dramatically increased ($K_{\text{dim}} = 2 \times 10^5 \text{ M}^{-1}$), which is much higher than predicted by the rules set by Schneider.⁴⁷ This strong binding affinity is probably because the intramolecular hydrogen bond not only preorganises the pyrimidine **56** for interactions, but also because it alters the pK_a of the hydrogen bonding atoms through inductive effects. Electron density is removed from the NH's that form intermolecular hydrogen bonds, making them more acidic, allowing them to form stronger interactions (Figure 2.4).

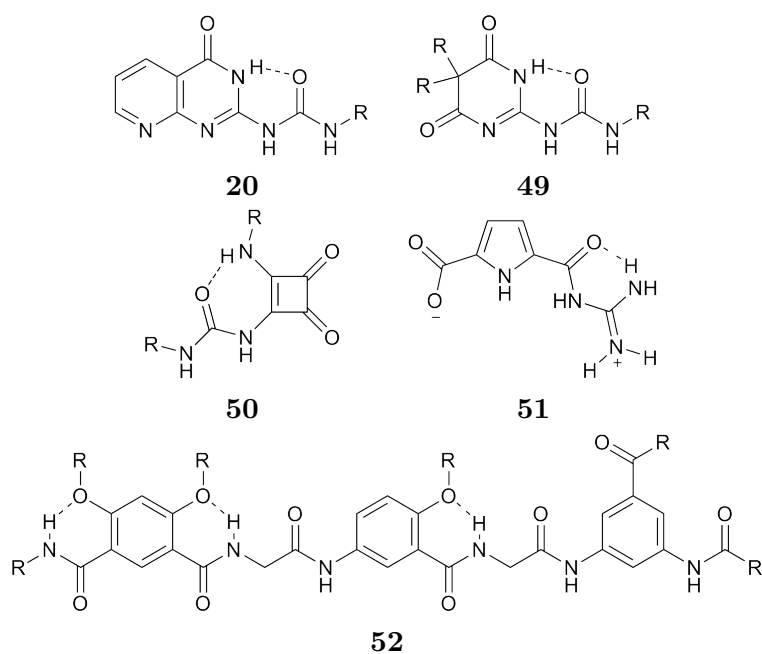


Figure 2.2: Examples of self-complementary molecules that are preorganised using a mixture of heteroaromatic rings and intramolecular hydrogen bonds. R denotes generic alkyl group.

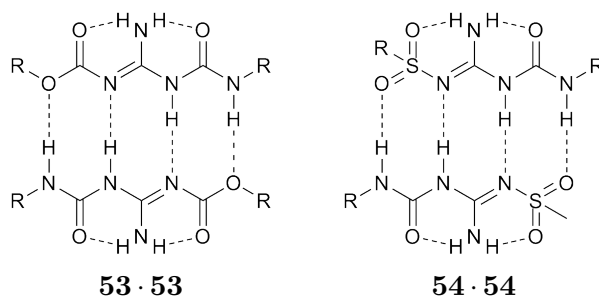


Figure 2.3: Self-complementary molecules that were introduced by Sanjayan *et al.* These molecules are preorganised solely by the use of intramolecular hydrogen bonds. R denotes generic alkyl group.

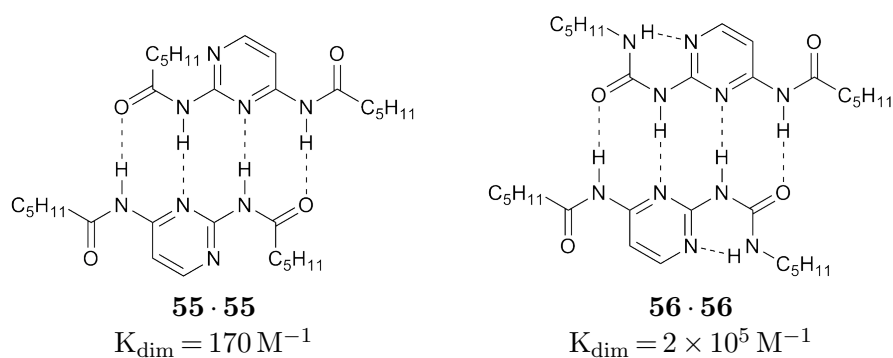


Figure 2.4: Addition of an intramolecular hydrogen bond dramatically increases the self-association constant of the pyrimidines proposed by Meijer *et al.*

Meijer *et al.* also introduced ureidopyrimidine (UPy) **19** as a self-complementary array that can form homocomplexes with a strong dimerisation constant ($K_{\text{dim}} = 2 \times 10^7 \text{ M}^{-1}$).⁶⁵ One issue is that UPy **19** can adopt different conformer/tautomer states (Figure 2.5), of which some do not present self-complementary hydrogen-bond arrays, but are all stabilised with an intramolecular hydrogen bond.³⁴ Hailes *et al.* then introduced ureidocytosine (UCyt) **18**,⁶⁴ which can also form homodimers through the self-complementary AADD array. One advantage of UCyt **18** over UPy **19** is that no alternative tautomers can be adopted, reducing the number of available conformations. However, UCyt **18** does not contain an intramolecular hydrogen bond when the desired AADD array (**18i**) is presented, but it can adopt a folded conformation (**18ii**), which does allow the formation of an intramolecular hydrogen bond (Figure 2.6). According to Etter's rules,⁵⁹ the folded conformation that contains the intramolecular hydrogen bond should be favoured, giving homodimers that interact through only two intermolecular hydrogen bonds, resulting in a reduced binding affinity compared to UPy **19**. However, because UCyt **18** forms strong homocomplexes ($K_{\text{dim}} > 2 \times 10^5 \text{ M}^{-1}$) it is most likely interacting through the desired AADD array. Even though both complexes (**18i** and **18ii**) contain a total of four hydrogen bonds, it is possible that the unfolded AADD array is favoured because it contains additional weak C-H \cdots O interactions to further stabilise the complex.³⁶

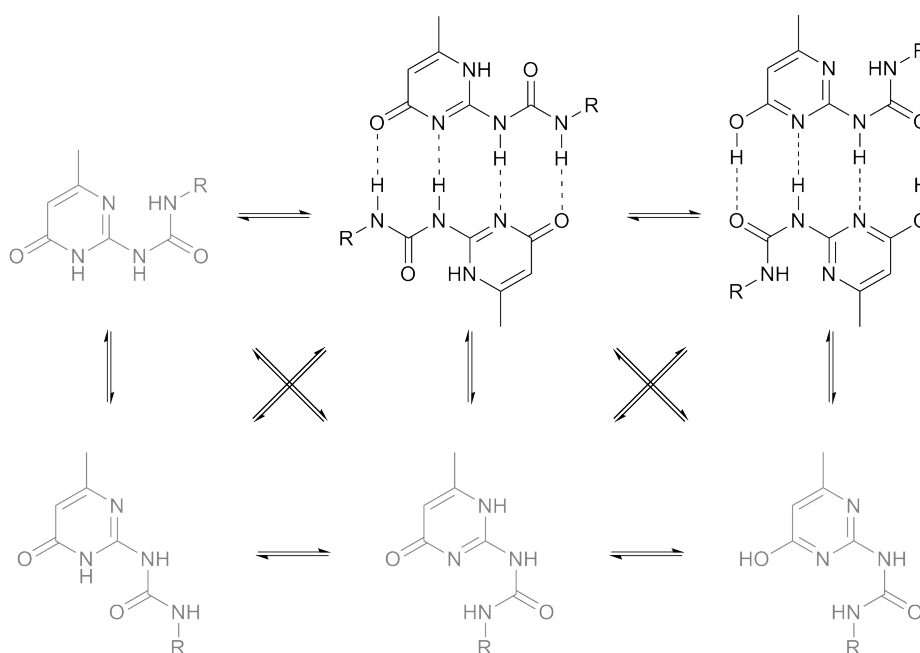


Figure 2.5: The possible tautomer/conformers of UPy **19** that are all stabilised with an intramolecular hydrogen bond. Only two (black) present self-complementary quadruple hydrogen-bond arrays, the others (grey) do not self-associate.

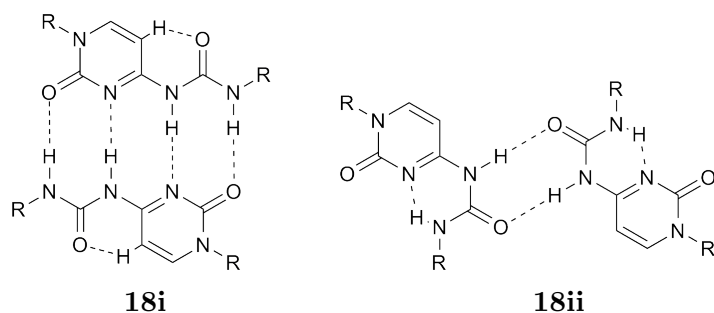


Figure 2.6: Two possible conformers of UCyt **18**. Although complexes with intramolecular hydrogen bonds are normally considered favourable, here the additional weak C–H ···O interactions favour self-association through the AADD array.

Currently, investigation of the effects of using intramolecular hydrogen bonding on binding affinity has only been carried out on self-complementary quadruple arrays. Furthermore, these effects are usually difficult to distinguish from other features; for example, changes in pK_a of the hydrogen bonding atoms due to inductive effects and preferential tautomeric alternatives. The objective of this work was to investigate the effects of preorganisation using triply hydrogen bonded heterocomplexes and has been published in *Chem. Eur. J.*¹⁵⁹ The phenylureidoimidazole (UIM) **17**, previously reported by the Wilson group,⁶⁰ was chosen as the DDA array, which remained constant throughout the study. It has been proposed that UIM **17** is a conformer independent DDA array (Figure 2.7),⁶⁰ which is important in this study because it remains constant, so must have the same binding capabilities through all possible conformations. Further studies were carried out in order to confirm that UIM **17** was conformer independent before

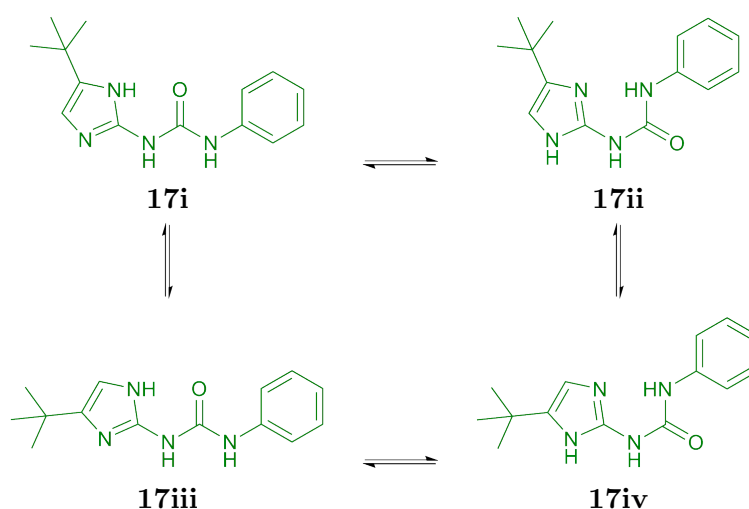


Figure 2.7: Conformers/tautomers of UIM **17** that contain an intramolecular hydrogen bond all present the desired hydrogen-bond array.

investigation into the effects of preorganisation could be completed. The complementary AAD array was then systematically changed in order to investigate the effect that intramolecular hydrogen bonds have on the binding affinity of AAD·DDA heterocomplexes.

2.1 Conformer independent array

UIM **17** can only be truly conformer independent if all possible conformations can form heterodimers with comparable binding affinities. In order to test this, a molecule that presents the complementary AAD array had to be designed. Amido*isocytosine* (AIC) **57** was chosen because of its similarities to UPy.⁶⁵ AIC **57i** contained an intramolecular hydrogen bond which preorganised the molecule for binding, but AIC **57** could also adopt an alternative tautomer which did not present the desired array **57ii**, so it was not a tautomer/conformer independent molecule (Figure 2.8a). Other tautomers (e.g. pyrimidinol **57iii** and **57iv**) were not considered, even though they would normally be favoured due to aromaticity, because they did not contain an intramolecular hydrogen bond. There were also additional steric and electrostatic repulsion between the pyrimidine nitrogens and the amide group (Figure 2.8b) disfavouring formation of the pyrimidinol tautomer. It was proposed that both conformers of UIM (**17i** and **17ii**) are able to interact with AIC **57** (Figure 2.9).

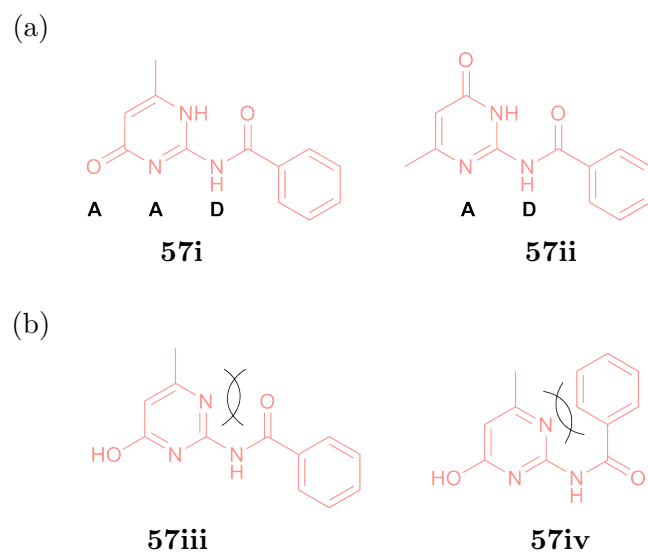


Figure 2.8: (a) There are two possible tautomers of AIC **57** that both contain an intramolecular hydrogen bond, but only one presents the desired AAD array. (b) Other possible tautomers do not allow the formation of an intramolecular hydrogen bond and have steric/electrostatic interactions.

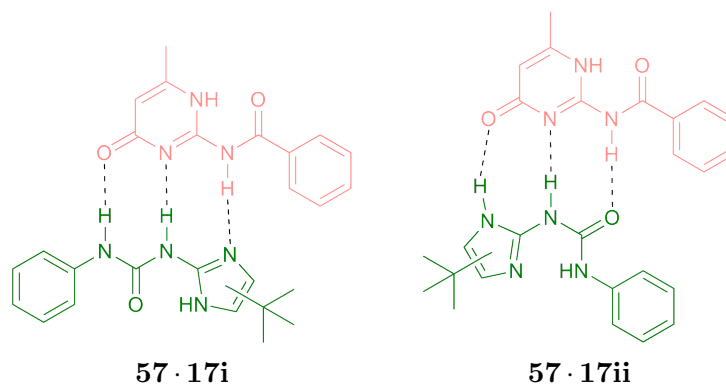
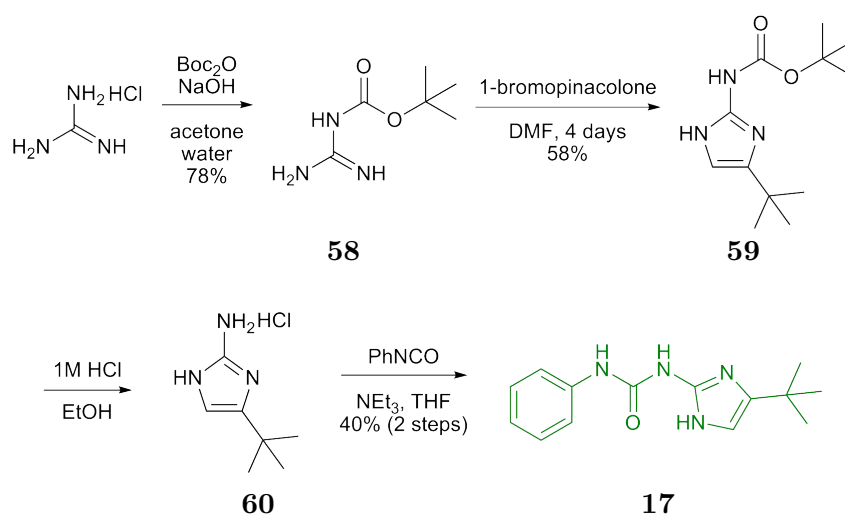


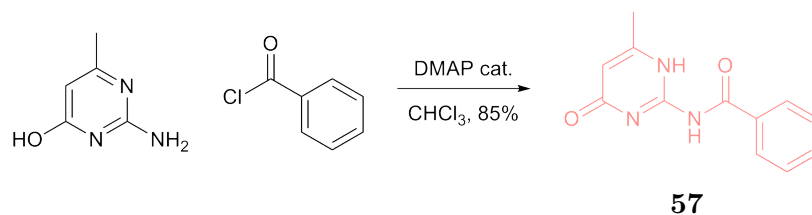
Figure 2.9: Heterocomplexes that can be formed between AIC **57** and UIM **17**.

2.1.1 Synthesis of Molecules

Both UIM **17** and AIC **57** were synthesised following procedures developed within the Wilson group. Boc protection of guanidine to give the boc-guanidine **58** followed by cyclisation with 1-bromopinacolone gave the imidazole **59**. Deprotection in acidic conditions to obtain the free amine **60**, followed by reaction with phenylisocyanate gave the UIM **17** (Scheme 2.1). Separation of UIM **17** from the side product, diphenylurea, was difficult and required multiple crystallisations. In the original paper⁶⁰ these crystallisations were not carried out and the measured association constants were lower than reported here as a result of the small amounts of impurities that were present.¹⁶⁰ AIC **57** was synthesised in a one step amide coupling between *isocytosine* and benzoyl chloride (Scheme 2.2). A catalytic amount of DMAP was required to further activate the acyl chloride because the amine is unreactive.



Scheme 2.1: Synthesis of UIM **17**.



Scheme 2.2: Synthesis of AIC **57**.

Crystal structures of both AIC **57** and UIM **17** were obtained (Figure 2.10).¹⁶¹ The crystal structure of AIC **57** does not present the desired hydrogen-bond array **57i**, but instead adopts the alternative tautomer **57ii**. There are two possible reasons for this; the tautomer found in the crystal structure could be the lowest energy conformation, or it could be because this tautomer is able to pack better in the solid state. The desired AAD array cannot self-associate, but the alternative tautomer presents a self-complementary AD array. Conversely, the crystal structure of UIM **17** does present one of the desired hydrogen-bond arrays. The required DDA array can be self-complementary through bifurcated hydrogen bonds, and this is the case here; molecules are linked through bridging methanol molecules (Figure 2.10c). Attempts were made to grow co-crystals of AIC **57** and UIM **17**, but these were all unsuccessful.

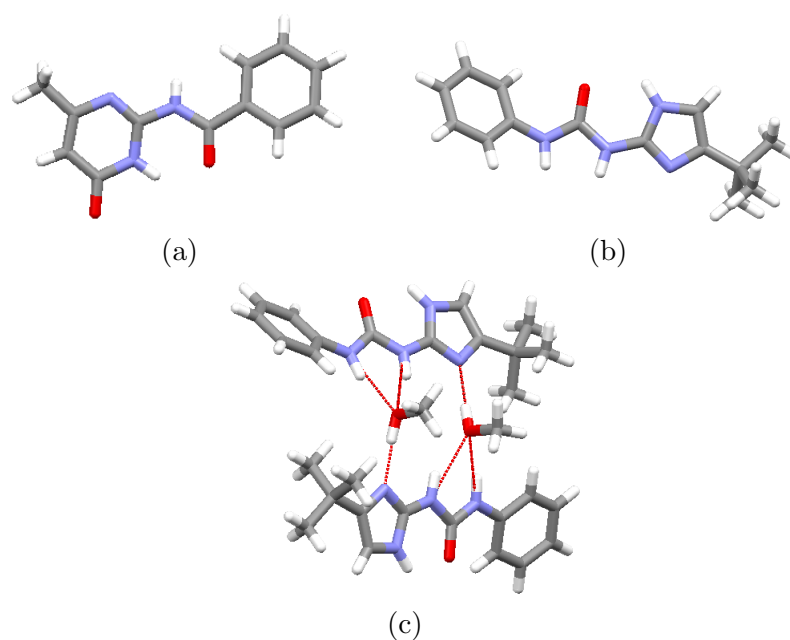


Figure 2.10: Crystal structures of (a) AIC **57**, (b) UIM **17** and (c) the packing structure of UIM **17** showing the bridging methanol molecules that form bifurcated hydrogen bonds between two UIM **17** molecules.

2.1.2 Binding Studies

Before investigations could be carried out to establish whether UIM **17** is a conformer independent array it was confirmed that the UIM · AIC **17** · **57** heterodimer can be formed in solution. When molecules interact they become close in space and this change in chemical environment can be followed by ^1H NMR.¹⁶² Comparison of ^1H NMR spectra of the individual molecules with the 1:1 mixture shows that there are significant complexation induced shifts (Figure 2.11). Any change in shift indicates that there are interactions between the molecules because the chemical environment of some protons has changed.

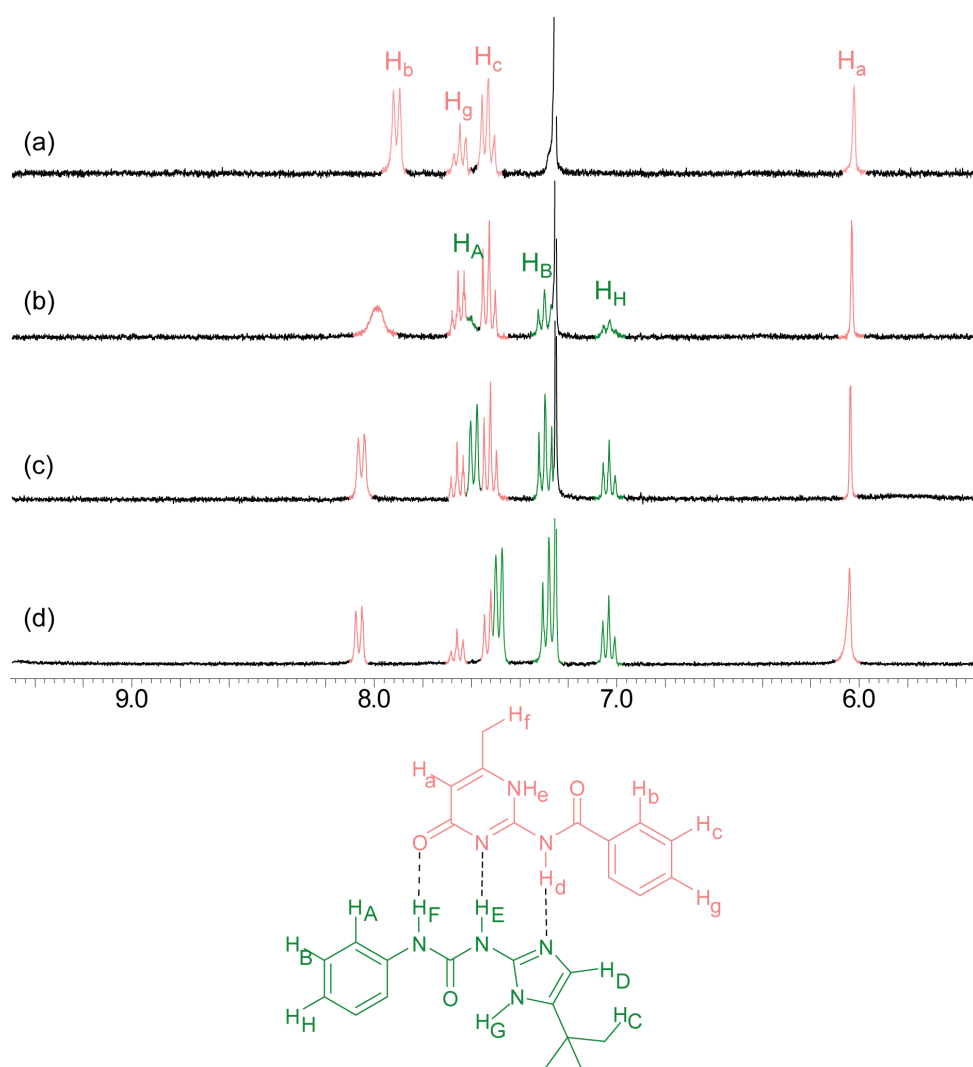


Figure 2.11: Typical ^1H NMR spectra from the titration of UIM **17** into AIC **57** (CDCl_3 , 300 MHz, 6 mM). Ratio of AIC:UIM **57**:**17** (a) 1:0, (b) 1:0.5, (c) 1:1, (d) 1:2.

To further confirm the formation of the complex in solution, diffusion ordered two-dimensional NMR spectroscopy (DOSY) studies were carried out. DOSY can be used to measure the size of components in solution. The speed with which they diffuse through the solution is related to their size. Small components tumble much quicker than larger ones, so formation of the complex can be observed by measuring the diffusion coefficient of a 1:1 mixture compared to the individual molecules. Because the diffusion coefficient is related to the size of the molecule, the size of the assembly that is formed in solution can be calculated using Equation 2.1.¹⁶³ Dimerisation of the individual molecules was established by measuring the diffusion coefficient at two extremes. For UIM **17**, diffusion coefficients of $D = 7.092 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ and $D = 5.247 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ at 10 and 100 mM respectively were measured. Using Equation 2.1 and assuming that dimeric species (517 Da) is formed at high concentrations, it was calculated (210 Da) that the species present at low concentrations is a monomer (actual mass 259 Da) of UIM **17**.^a Similarly, AIC **57** gave $D = 9.575 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ and $D = 7.058 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ for 10 and 100 mM respectively. Assuming AIC **57** is only present as the homodimer (458 Da) at the higher concentration, a mass of 182 Da is calculated for the low concentration, corresponding to monomer (actual mass = 229 Da). Since UIM **17** and AIC **57** are monomeric at 10 mM, this was the concentration of choice to study formation of the heterocomplex. The methyl group of AIC **57** was followed, giving a diffusion coefficient of $D = 7.721 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$. This corresponds to a mass of 493 Da, which suggests that a heterodimer (actual mass = 488 Da) is present.

$$\frac{D_{[\text{low}]}}{D_{[\text{high}]}} = \frac{\sqrt[3]{MW_{[\text{high}]}}}{\sqrt[3]{MW_{[\text{low}]}}} \quad (2.1)$$

^aThe calculated mass is lower than than the actual mass because the DOSY technique includes a layer of solvent around the molecule when calculating the radius, which was not taken into account for the $MW_{[\text{high}]}$.

2.1.3 Measurement of Dimerisation

Having demonstrated that the complex does indeed form in solution it was necessary to determine the binding affinity. There are many techniques available for measuring the association constant (K_a) of complexes, including by ^1H NMR titration. The change in proton shift is proportional to the strength of the interaction, making it possible to use this to calculate the strength of interaction in the complex. Firstly, dilution studies were carried out in order to ascertain whether the individual components can self-associate. If self-association is high it can impact upon the equilibrium position (Equation 2.2),¹⁶⁴ resulting in a lower concentration of the desired complex being present. This in turn impacts upon the K_a (Equation 2.3) because K_a is equal to the concentration of the desired complex over the concentration of all other species that are present. If the concentration of undesired species is high then K_a is lower as a result.



$$K_a = \frac{[\text{UIM} \cdot \text{AIC}]}{[\text{UIM} \cdot \text{UIM}][\text{AIC} \cdot \text{AIC}][\text{UIM}][\text{AIC}]} \quad (2.3)$$

Dimerisation constants (K_{dim}) for both UIM **17** and AIC **57** were found to be negligible ($10 \pm 2 \text{ M}^{-1}$ and $4 \pm 1 \text{ M}^{-1}$ respectively) by ^1H NMR dilutions and were therefore, not included in any of the calculations for K_a . The titrations to obtain K_a were then carried out; a solution of UIM **17** containing some AIC **57** was titrated into a solution of AIC **57**. After each aliquot of UIM **17** was added a ^1H NMR spectrum was recorded and the change in shift was plotted against the equivalents of UIM **17** (Figure 2.12). HypNMR¹⁶⁵ was used to calculate the K_a of the complex, which was found to be $33000 \pm 16000 \text{ M}^{-1}$. The titration was carried out in triplicate and the mean average taken as the association constant and the reported error is the standard deviation of the three measurements. Since $\text{error}(K_a)$ and $e^{\text{error}(\ln(K_a))}$ are not comparable, the average and error of K_a could not simply be converted to ΔG .^b ΔG was then calculated for each individual titration using Equation 1.2, where R is the Boltzmann gas constant ($8.314 \text{ J K}^{-1} \text{ mol}^{-1}$) and T is temperature (298 K), and the mean value taken as the ΔG and the error was the

^bGuidance from Dr. Matthew Greenwood-Nimmo, Lecturer in Applied Econometrics at the University of Leeds, was taken for calculation of average ΔG 's.

standard deviation of the converted ΔG 's. This gave $\Delta G = -25.6 \pm 1.2 \text{ kJ mol}^{-1}$. It was not possible to carry out the reverse titration (where AIC **57** was titrated into UIM **17**) because of solubility restrictions. The results obtained here were much higher than reported in the original paper,⁶⁰ probably due to the more rigorous purification of the UIM **17** to remove any trace impurities which could interfere with interactions.

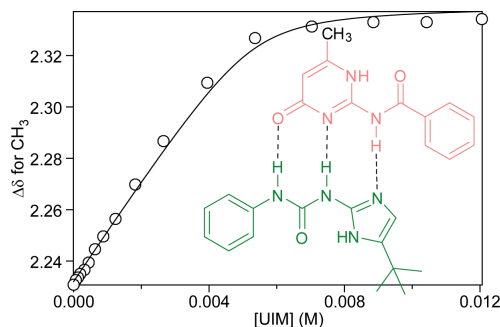


Figure 2.12: Typical titration graph obtained from HypNMR for the titration of UIM **17** into AIC **57** (0.006 M).

When molecules interact there is normally a change in enthalpy and entropy, which can result in a release of energy. This release of energy can be measured using isothermal titration calorimetry (ITC).¹⁶⁶ ITC measures any changes in temperature by maintaining a standard cell at the same temperature as the experimental cell. As molecules are titrated into the experimental cell, energy must be added or removed from the standard cell, ensuring that it remains at the same temperature as the experimental cell. This energy correlates to the energy released by the molecules when they interact. Because ITC can be carried out at lower concentrations than ^1H NMR, it was possible to perform the titration in both directions (i.e. UIM **17** into AIC **57** and AIC **57** into UIM **17**) (Figure 2.13). Titrations between UIM **17** and AIC **57** provided usable endotherms for curve fitting; however, since ITC measures a combination of hetero and homodimerisation, dilution studies were carried out as well as the normal titrations. It was shown that AIC **57** did not self-associate because there was no change in heat release throughout the dilution experiment. UIM **17** has a different dilution curve which shows that as it becomes more concentrated it releases more heat, indicating that self-association is occurring. Figure 2.13 shows the titration and dilution curves for the titration of UIM **17** into AIC **57** alongside the reverse experiment (i.e. AIC **57** into UIM **17**). Standard analysis of the data by subtraction of dilution curves from titration curves gave data that could not be fitted with standard MicroCal models available in the Origin software. Therefore, a global analysis of all titration and UIM **17** dilution data, using a mixed

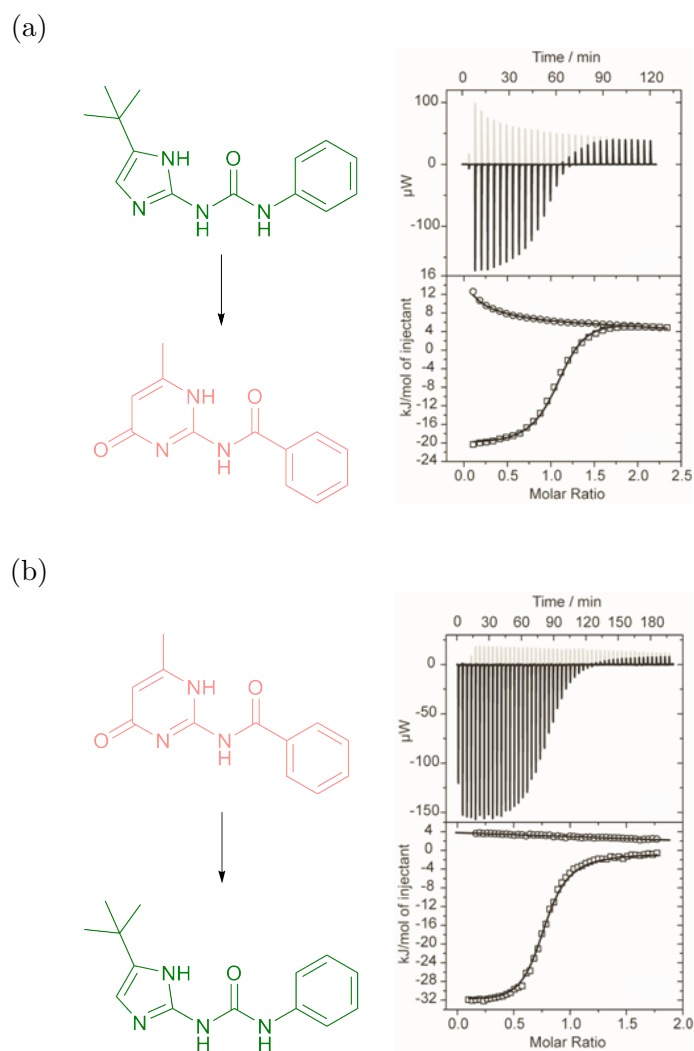


Figure 2.13: ITC titrations (black) and dilutions (grey) for (a) AIC **57** in the cell and (b) UIM **17** in the cell.

association model, that allows for UIM **17** dimerisation coupled to AIC·UIM **57**·**17** binding was carried out^c using the SEDPHAT software programme.¹⁶⁷ This data analysis provided a rigorous determination of the binding parameters for the interaction of UIM **17** with AIC **57**, giving values of $K_a = 24700 \pm 3800 \text{ M}^{-1}$, $\Delta H_a = -41.9 \pm 2.1 \text{ kJ mol}^{-1}$ and $K_{\text{dim}} = 520 \pm 70 \text{ M}^{-1}$, $\Delta H_{\text{dim}} = -28.3 \pm 1.3 \text{ kJ mol}^{-1}$. The association constant corresponds to $\Delta G_a = -24.6 \text{ kJ mol}^{-1}$, $\Delta S_a = 58.9 \text{ J K}^{-1} \text{ mol}^{-1}$. The analysis reveals, as expected, an enthalpy driven process. Furthermore, the association constant obtained using ITC is comparable to that determined by $^1\text{H NMR}$ spectroscopy. The self-association constant that is extracted for UIM **17** however, is one order of magnitude different between ITC and $^1\text{H NMR}$ ($K_{\text{dim}} = 520 \text{ M}^{-1}$ and $K_{\text{dim}} = 11 \text{ M}^{-1}$ respectively). This may

^cAnalysis was carried out with the guidance of Dr. Bruce Turnbull.

reflect the small signal changes observed and limited concentration range accessible in the ^1H NMR experiment due to solubility limitations.

2.1.4 Confirmation of Conformer Independence

Although extensive studies on the dimerisation of AIC · UIM **57** · **17** have been carried out, it has still not been confirmed whether UIM **17** is conformer independent. The ^1H NMR spectra illustrate that UIM **17** interacts with AIC **57** because of the large changes in shift seen between the individual components and the mixture. However, there is only one signal present for each proton of the UIM **17**, which can mean two things. Firstly, it could be that only one conformer/tautomer is present in solution and this is the one that interacts. Secondly, it could be that all conformers/tautomers interconvert and the signals that are seen are an average of all of tautomer/conformer states that are present in solution. The second option is most likely because each conformer/tautomer is stabilised by only one intramolecular hydrogen bond, so should be easily broken. ^1H - ^1H NOESY gives evidence that both conformations of UIM **17** can interact with AIC **57** (Figure 2.14) however, it is difficult to differentiate between the tautomers because the resonances corresponding to these are not resolved. It is likely that because of the bulk of the ^tBu group only two of the tautomers (**17i** and **17ii**) can interact because the other two (**17iii** and **17iv**) are sterically hindered.

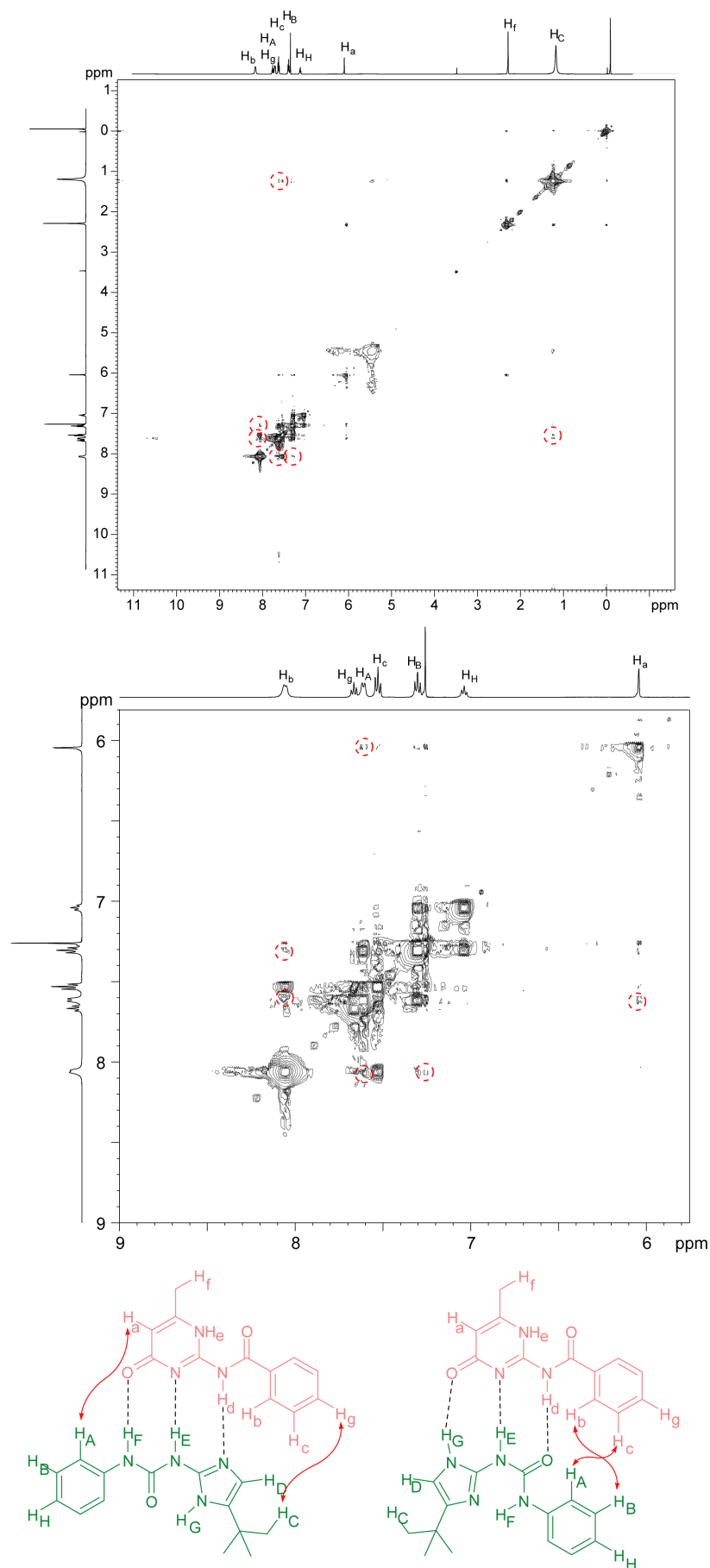


Figure 2.14: ^1H - ^1H NOESY of AIC · UIM 57 · 17 shows that two conformers of UIM 17 can interact.

A low temperature ^1H NMR spectrum (400 MHz, CDCl_3 , 213 K, 40 mM) of a 1:1 mixture of $\text{UIM} \cdot \text{AIC } \mathbf{17} \cdot \mathbf{57}$ shows that more than one conformer/tautomer of $\text{UIM } \mathbf{17}$ is present (Figure 2.15). As the sample is cooled, exchange between the different conformer/tautomer states is slowed. Since each conformer/tautomer places protons in different chemical environments and exchange is sufficiently slow at low temperatures, they can be distinguished by ^1H NMR. There are three visible ^tBu peaks, with one more prevalent than the others. However, it was not possible to assign the peaks to particular conformers because the sample precipitated when it was cooled down for extended periods of time required to perform ^1H - ^1H NOESY experiments.

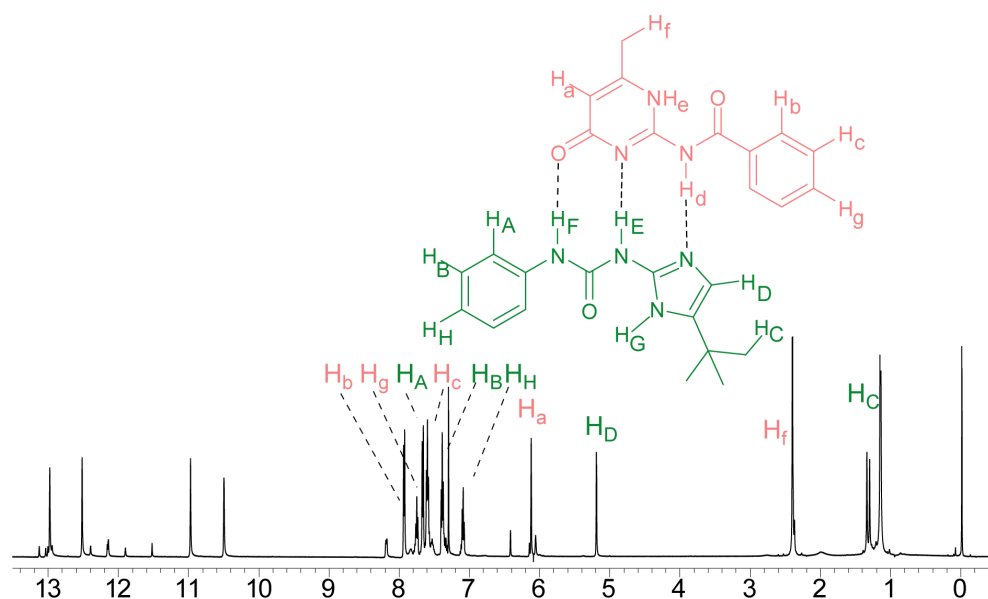


Figure 2.15: As the sample is cooled, exchange between the different tautomer/conformer states slows, meaning that the peaks resolve, allowing the different tautomers to be observed. (400 MHz, CDCl_3 , 213 K, 40 mM)

Although ^1H - ^1H NOESY and low temperature ^1H NMR spectra confirmed that different conformations of $\text{UIM } \mathbf{17}$ were present in the $\text{AIC} \cdot \text{UIM } \mathbf{57} \cdot \mathbf{17}$ complex, neither could conclude what ratios each conformer were in. It was necessary to turn to molecular modelling in order to ascertain the binding affinity of each conformer of $\text{UIM } \mathbf{17}$. If the two conformers were found to have similar binding affinities by molecular modelling then it could be assumed that $\text{UIM } \mathbf{17}$ is a conformer independent array. DFT calculations were carried out using Gaussian 03,¹⁶⁸ which is the program of choice because it can accurately predict hydrogen bonds. However, the molecules and complexes are large, which led to long calculation times when the desired, global basis set (6-311++G**) was used. This required new basis sets to be written so that hydrogen bonds could be

accurately predicted using a large basis set, whilst the rest of the molecule was calculated using a smaller basis set, decreasing the calculation time. All calculations were carried out using the B3LYP level of theory in the gas phase. The basis set that was chosen for calculations was C, 6-31G, N/O, 6-31G*, H, 6-31++G**; this includes polarisation¹⁶⁹ and diffusion¹⁷⁰ functions on hydrogen-bonding atoms, but keeps the calculation of non-hydrogen-bonding atoms simple. It is important to include polarisation and diffusion functions on the hydrogen-bonding atoms because they share electron density, so their orbitals become distorted as the hydrogen bond is formed. Results using this basis set were compared to those of a larger basis set (C/N/O, 6-311G*; H, 6-311++G**) and were found to be quite similar (Table 2.1). The individual molecules (**17** and **57**) and the complex **17**·**57** was calculated using both basis sets. Subtraction of the energy of the individual molecules from that of the complex gave the energy from intermolecular interactions. A difference of only 10 kJ mol⁻¹ was obtained between the two basis sets, which is a relatively small difference in energy. Since the purpose of molecular modelling in this study was not to obtain absolute binding affinities, but simply to observe trends, the smaller basis set was able to calculate the binding affinities to the required accuracy.

Table 2.1: Comparison of the basis sets. Small; C, 6-31G, N/O, 6-31G*, H, 6-31++G**. Large; C/N/O, 6-311G*, H, 6-311++G**.

	Small (Hartrees)	Big (Hartrees)
UIM 17	-681.243	-681.524
AIC 57	-778.527	-778.849
AIC · UIM 57 · 17	-1459.822	-1460.421
ΔG (kJ mol ⁻¹)	-136.9	-126.8

Firstly, the energies of the possible conformers and tautomers of UIM **17** and AIC **57** were calculated. It was found that for UIM **17**,^d both conformers that contain an intramolecular hydrogen bond have very similar energies (Figure 2.16). A difference of only ~ 2 kJ mol⁻¹ was calculated. AIC **57** was different, the tautomer that presents the desired hydrogen-bond array (**57i**) is ~ 30 kJ mol⁻¹ higher in energy than the alternative tautomer (**57ii**), found in the crystal structure (Figure 2.17).

^dThe ^tBu group was removed for two reasons. 1, to simplify the calculations and 2, to reduce the number of conformer/tautomer possibilities.

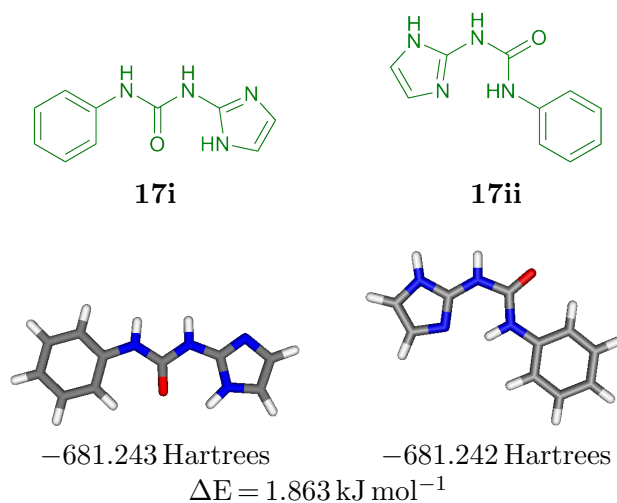


Figure 2.16: The energies of UIM **17i** and **17ii** that were obtained using B3LYP/C, 6-31G; N/O, 6-31G*; H, 6-31++G**.

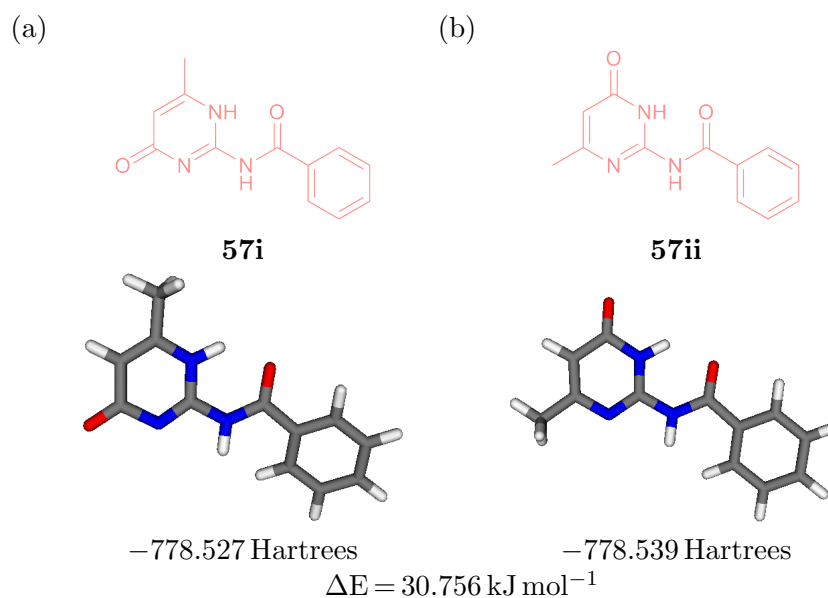


Figure 2.17: The energies of AIC **57i** and **57ii** that were obtained using B3LYP/C, 6-31G; N/O, 6-31G*; H, 6-31++G**.

Even though both conformers of UIM **17** are similar in energy, it still does not necessarily mean that they both interact with similar binding affinities. The energies of the two possible complexes (**57 · 17i** and **57 · 17ii**) were calculated using the same basis set as for the individual motifs (Figure 2.18), but with counterpoise^{171,172} included. Counterpoise allows the basis set superposition error (BSSE) to be calculated. As molecules approach one another their basis functions overlap, resulting in each molecule increasing its basis set. This results in a higher energy being calculated and using counterpoise allows this error to be removed. Counterpoise runs five calculations for each complex; the energy of each molecule in their own basis set, the energy of each molecule using the global basis set and the energy of the complex. The BSSE can be calculated by addition of the errors in calculating the individual molecules, which is the difference in energy for the molecules calculated using the individual and global basis set. The strength of the intermolecular interactions can be calculated by subtracting the energies of the molecules calculated in their own basis sets from the counterpoise corrected complex. This energy is higher than the energy measured experimentally because it does not include the entropic loss that occurs when the molecules twist out of their favoured, planar geometry,

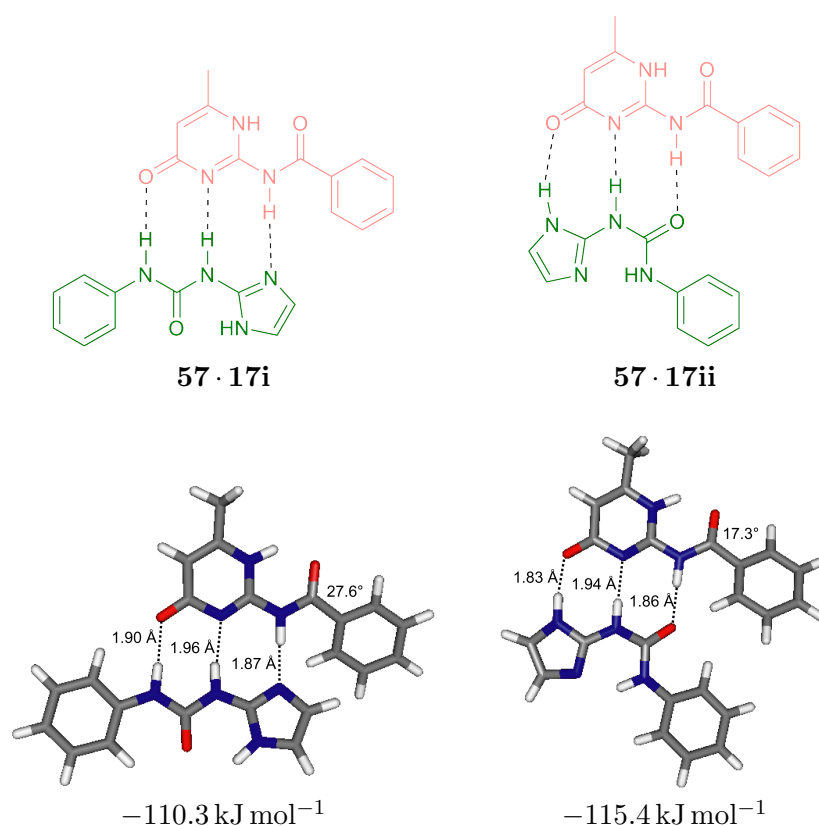


Figure 2.18: Molecular modelling of the AIC · UIM **57 · 17** complex.

Table 2.2: Molecular modelling of the complexes shows that both complexes have similar binding affinities.

	AIC · UIM 57 · 17i (Hartrees)	AIC · UIM 57 · 17ii (Hartrees)
AIC · UIM 57 · 17 ^a	-1459.822	-1459.824
AIC · UIM 57 · 17 ^b	-1459.812	-1459.813
AIC 57 ^c	-778.529	-778.530
AIC 57 ^d	-778.522	-778.523
AIC 57 ^e	-778.527	-778.527
UIM 17 ^c	-681.244	-681.244
UIM 17 ^d	-681.241	-681.239
UIM 17 ^e	-681.243	-681.242
Intermolecular interactions	-127.2 ^f	-133.0 ^f
Measured intermolecular interactions	-110.3 ^f	-115.4 ^f

^a Without counterpoise correction.

^b With counterpoise correction

^c Using the global basis set.

^d Using the basis set for the individual molecule.

^e The planar molecule calculated alone.

^f kJ mol⁻¹.

in order to reduce steric hindrance in the complex. To include this change in geometry to the calculation of the binding affinity, the energy of the molecules calculated in their planar geometry is subtracted from the counterpoise corrected complex. A summary of the results obtained for the AIC · UIM **57 · 17** complex can be found in Table 2.2. Both conformers of UIM **17** gave similar results, with UIM **17ii** having slightly stronger intermolecular interactions (-115.4 kJ mol⁻¹) than **17i** (-110.3 kJ mol⁻¹).^e This is probably because, in the unfolded UIM **17i**, steric interactions between the phenyl ring of AIC **57** and the imidazole of UIM **17** mean that the phenyl ring has to twist further out of the plane (27.6°), resulting in a loss of conjugation.

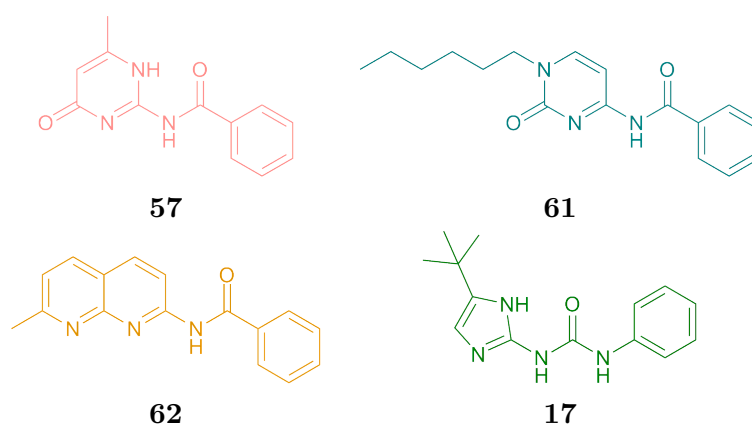
In summary, it has been shown that the UIM · AIC **17 · 57** complex forms in solution and the measured K_a using different methods, provides consistent results. ¹H NMR studies confirmed that UIM **17** can interact through two possible conformers and molecular

^eThe binding affinities obtained from molecular modelling are much higher than the experimental results because the calculations are carried out in the gas phase, so solvation is not taken into account.

modelling has shown that the binding affinity of these two conformers is similar. These results confirm that UIM **17** is a conformer independent array.

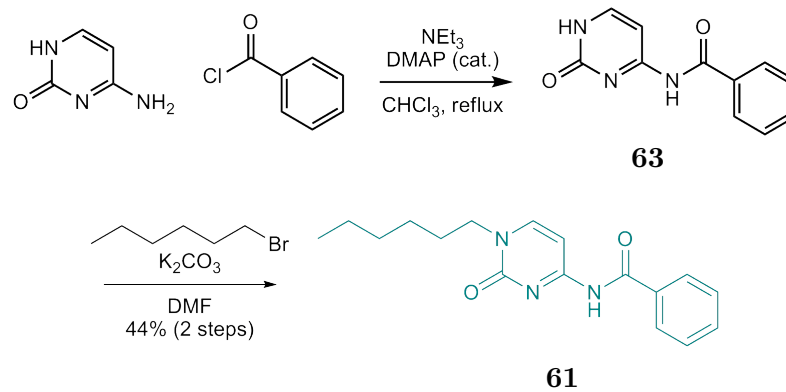
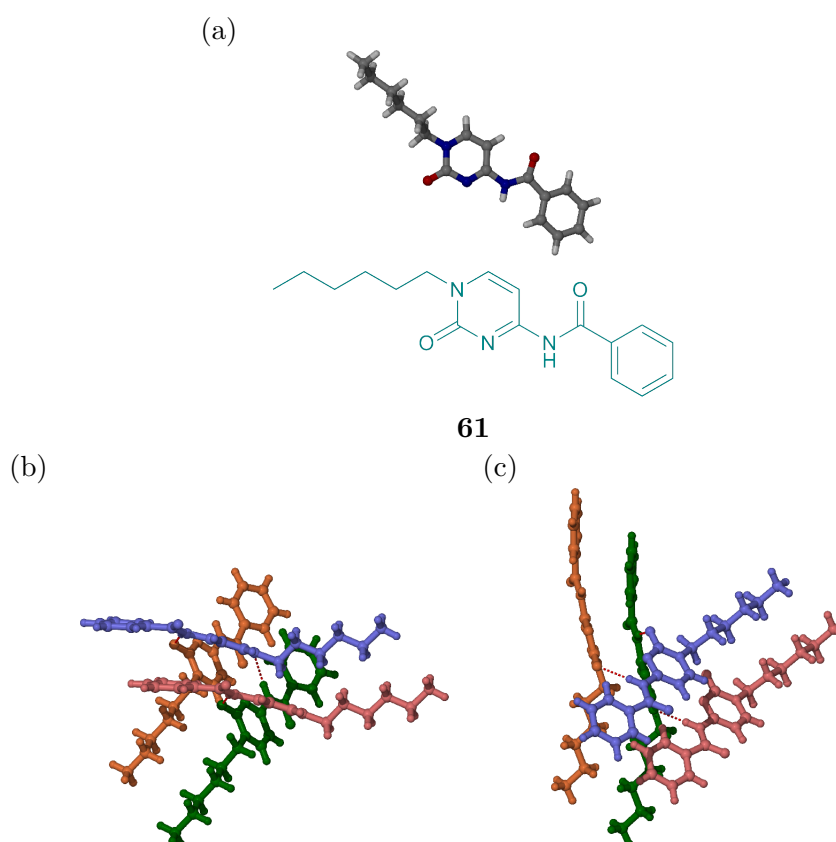
2.2 Effects of preorganisation in AAD·DDA arrays

Throughout the study described below, UIM **17** remained constant and the complementary partner was changed to different molecules that present an AAD array, but with different degrees of preorganisation and steric hindrance. Two new molecules were designed (**61** and **62**) that both present an AAD array, but do not contain a preorganising intramolecular hydrogen bond. *n*-Hexylbenzamidocytosine (HAC) **61** shows similarities to the UCyt **18** that was introduced by Hailes.⁶⁴ The benzamidonaphthyridine (ANT) **62** is similar to the naphthyridine described by Corbin *et al.*⁶¹ however, it has an extra methyl group which could interfere with the binding affinity because of steric interactions.

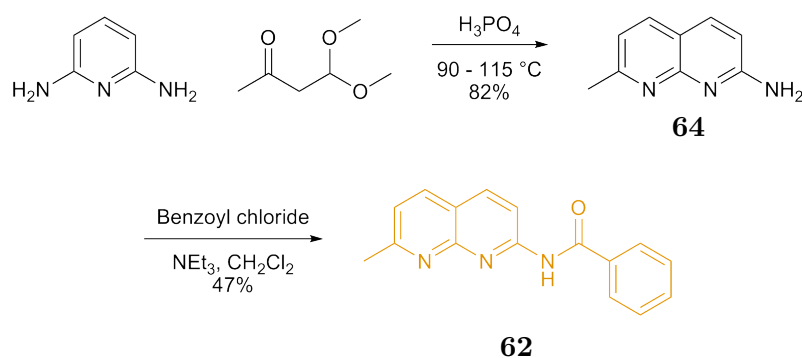


2.2.1 Synthesis of Molecules

For the synthesis of UIM **17** and AIC **57** see Section 2.1.1. HAC **61** was synthesised in two steps (Scheme 2.3). Firstly, amide coupling between cytosine and benzoyl chloride in the presence of base and catalytic DMAP gave the insoluble amidocytosine **63**. To aid solubility an alkyl chain was added *via* an S_N2 reaction to give the HAC **61**. The alkylation was low yielding, because there are other possible alkylation sites in the molecule and a mixture of all these products was obtained.¹⁷³ A crystal structure of HAC **61** was obtained from the slow addition of hexane into a CHCl₃ solution of HAC **61** (Figure 2.19). In the crystal structure the desired AAD hydrogen-bond array is presented

Scheme 2.3: Synthesis of HAC **61**.Figure 2.19: The crystal structure of (a) HAC **61** with (b) the packing structure from (c) different views.

and the molecules interact through two of these hydrogen bonds in the extended packing structure. The molecules are assembled into rows however, it does not appear that they interact through π - π stacking because the molecules are not aligned correctly. ANT **62** was synthesised (Scheme 2.4) following minor modifications to a literature procedure.⁴³ Diamidopyridine was dissolved in hot phosphoric acid, before dimethoxy-2-butanone was added dropwise, to give the aminonaphthyridine **64**. Reaction with benzoyl chloride in the presence of NEt_3 gave the ANT **62**. Multiple crystallisations were necessary in order to obtain material that was pure enough to acquire accurate results from the ^1H NMR titration experiments.



Scheme 2.4: Synthesis of ANT **62**.

2.2.2 ^1H NMR Titrations

It was expected that AIC **57** would have the strongest binding affinity towards UIM **17** because it is preorganised for interactions by the intramolecular hydrogen bond, which may also increase binding affinity through inductive effects. ANT **62** was predicted to have the lowest binding affinity because it does not contain a preorganising intramolecular hydrogen bond and also has a sterically hindering methyl group. HAC **61** was expected to have an association constant inbetween that of the other two complexes. ^1H NMR titrations were carried out for each of the complexes and data was processed using HypNMR¹⁶⁵ to give K_a and ΔG . It was found that ANT **62** had the lowest binding affinity for UIM **17** ($\Delta G = -19.3 \pm 0.4\text{ kJ mol}^{-1}$), but the binding affinities of AIC **57** and HAC **61** with UIM **17** ($\Delta G = -25.6 \pm 1.2$ and $-24.2 \pm 0.3\text{ kJ mol}^{-1}$ respectively) were both similar. This result shows that preorganisation of AIC **57** with an intramolecular hydrogen bond does not dramatically increase its binding affinity towards UIM **17**. The reasons for this were not immediately obvious, so conformational analysis of the individual components and their complexes was carried out.

2.2.3 NMR Studies of Compounds and Complexes

^1H - ^1H NOESY spectra of the individual components UIM **17** and AIC **57**, did not give good structural information because all of the NH protons were broadened. This is presumably because they are in rapid exchange, so are not visible by ^1H NMR. For the other molecules, HAC **61** and ANT **62**, the NH protons were visible in the ^1H NMR spectra, making it possible to gain more information. It was found that no correlations between the NH and other protons were observed in HAC **61**, but only one conformation was visible in the ^1H - ^1H NOESY spectrum for ANT **62** (Figure 2.20). Only correlations between the NH and H_e were observed. Correlations between the NH and H_d , which

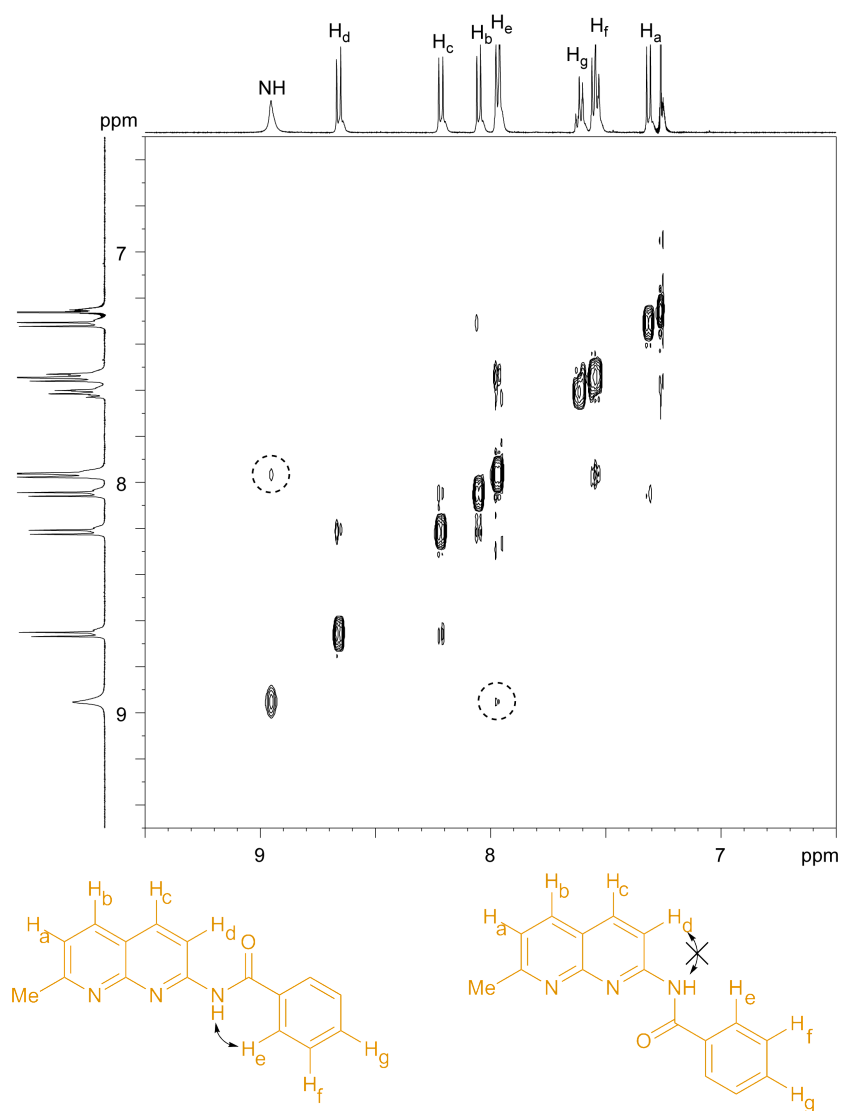


Figure 2.20: ^1H - ^1H NOESY of ANT **62** shows that the desired conformation is the only one that is present in solution.

could arise from the molecule being in an alternative conformation, were not seen. This suggests that ANT **62** presents the desired hydrogen-bond array in solution. ^1H - ^1H NOESY spectra were also obtained for the mixtures of HAC **61** and ANT **62** with UIM **17**. These however, only showed evidence that one conformer of UIM **17** was interacting. This could either mean that the second conformer does not form complexes or that it does not show up in the ^1H - ^1H NOESY spectra because non-covalently bound species can give weak nOe's. Since the binding array in AIC **57** and HAC **61** is very similar it is unlikely that both conformations of UIM **17** will interact with AIC **57**, but not HAC **61**. For ANT **62** the binding face is more hindered because of the methyl group, which could block the approach of one of the conformers of UIM **17**.

2.2.4 Molecular Modelling

Since no useful structural information was obtained for the individual molecules and complexes using ^1H NMR studies, molecular modelling was carried out. Firstly, the energy of four possible rotamers of HAC **61** (Figure 2.21) and ANT **62** (Figure 2.22) were calculated, using the same level of theory and basis set as described in Section 2.1.4. These four rotamers were chosen because they are likely to be the most stable. They can adopt a planar structure, allowing the sp^2 orbitals to align, so that delocalisation of the p-electrons can occur throughout the entire molecule. It was found that the rotamers that present the desired hydrogen-bond array were the lowest in energy for both molecules, which is presumably because of unfavourable steric/electronic clashes in the alternative rotamers. This is the opposite to AIC **57**, where the lowest energy tautomer presented an undesired hydrogen-bond array.

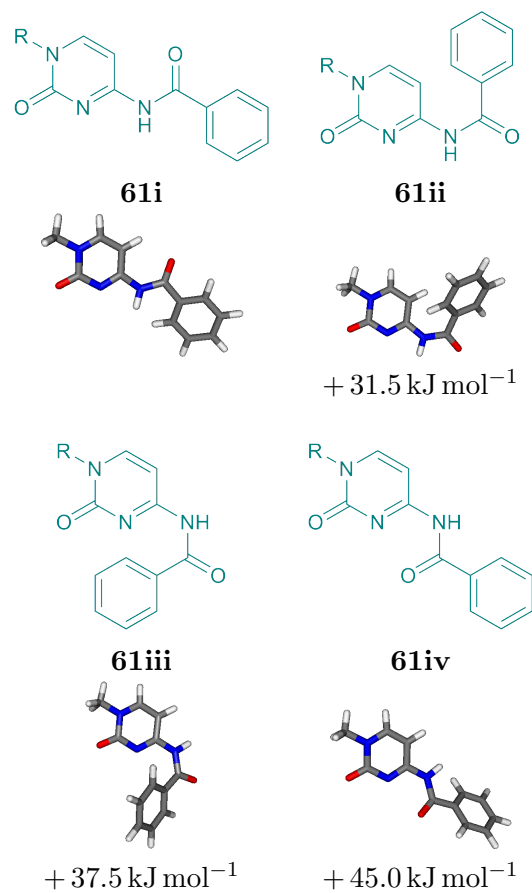


Figure 2.21: Molecular modelling of HAC **61** revealed that the conformer that presents the desired hydrogen-bond array (**61i**), is lowest in energy. Energies show how much higher in energy each conformer (**61ii**, **61iii** and **61iv**) is than the ground state molecule (**61i**).

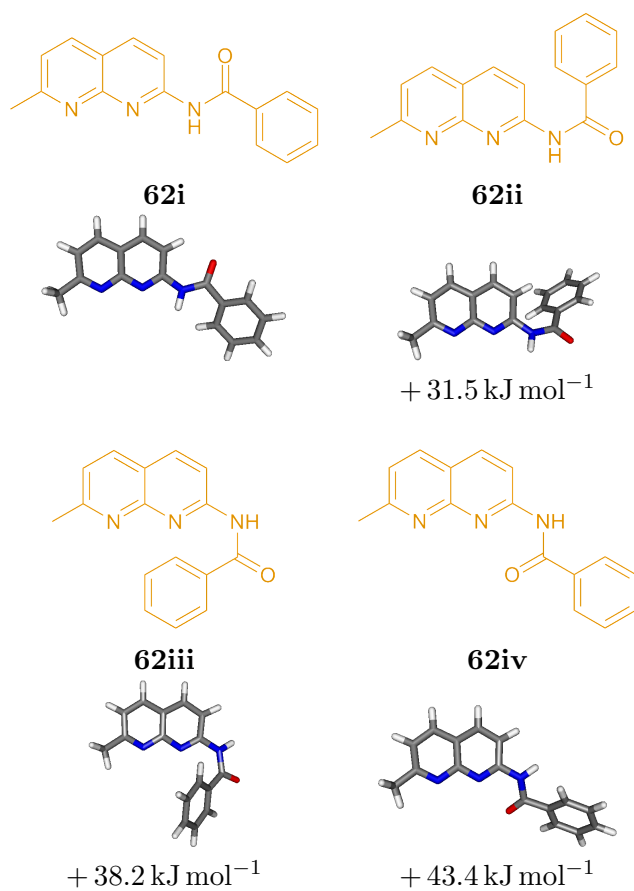


Figure 2.22: Molecular modelling of ANT **62** revealed that the conformer that presents the desired hydrogen-bond array (**62i**), is lowest in energy. Energies show how much higher in energy each conformer (**62ii**, **62iii** and **62iv**) is than the ground state molecule (**62i**).

The energy of the conformer/tautomer that presents the desired hydrogen-bond array of AIC **57** (Figure 2.18), HAC **61** (Figure 2.23) and ANT **62** (Figure 2.24a) with both conformers of UIM **17** were then calculated, using the same level of theory and basis set as described previously. It was shown that the AIC · UIM **57** · **17** complex had the strongest binding affinity ($\Delta G = -110.3$ and $-115.4 \text{ kJ mol}^{-1}$) and HAC · UIM **61** · **17** had a lower binding affinity ($\Delta G = -92.5$ and $-95.9 \text{ kJ mol}^{-1}$). This does not agree with the experimental values, where AIC **57** and HAC **61** had similar binding affinities to UIM **17** ($\Delta G = -25.6 \pm 1.2$ and $-24.2 \pm 0.3 \text{ kJ mol}^{-1}$ respectively). A likely explanation for this is that the lowest energy conformer of HAC **61** presents the desired hydrogen-bond array, so it does not have to change conformer in order to form complexes. AIC **57** on the other hand, possibly presents an undesired hydrogen-bond array. Therefore, in order to interact with UIM **17**, AIC **57** must change tautomer/conformer, losing energy, before it can form the heterocomplex and this is not accounted for in the calculations. This loss in energy means that, although molecular modelling indicated that the intermolecular hydrogen bonds were stronger in the AIC · UIM **57** · **17** complex, the measured binding affinities of AIC **57** and HAC **61** were very similar. ANT · UIM **62** · **17** had the lowest binding affinity ($\Delta G = -70.9$ and $-76.5 \text{ kJ mol}^{-1}$), which was because of steric hindrance from the methyl group, which can be seen when viewed from the side (Figure 2.24b). The methyl group of ANT **62** blocks the approach of both conformations of UIM **17**,

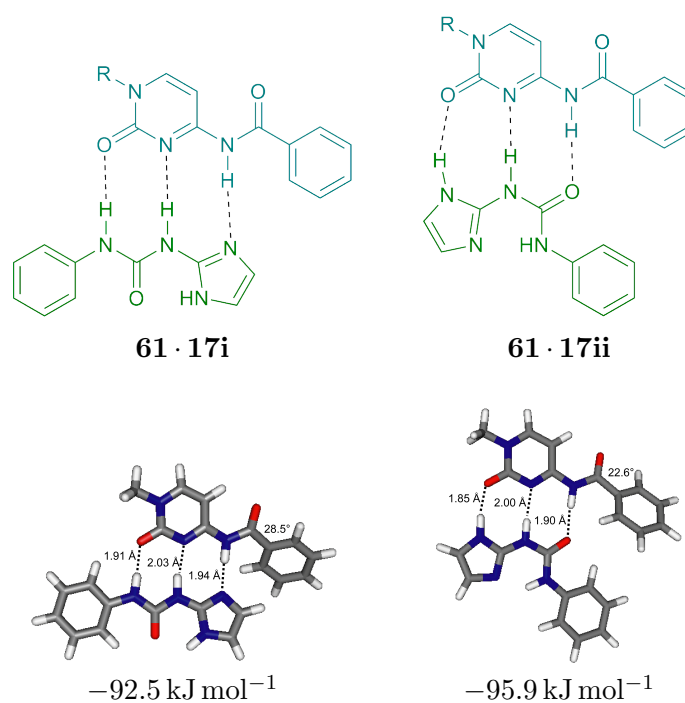


Figure 2.23: Molecular modelling for the HAC · UIM **61** · **17** complex.

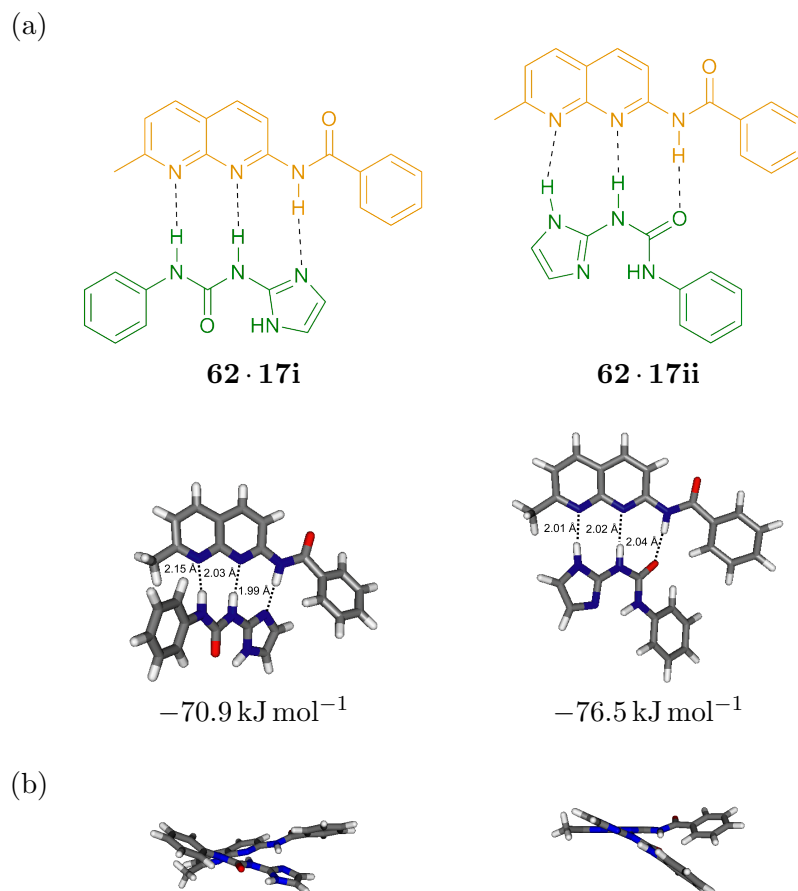


Figure 2.24: (a) Molecular modelling for the ANT · UIM **62 · 17** complex and (b) the side views showing how the methyl group of ANT **62** sterically hinders the approach of UIM **17** in both conformations, preventing strong intermolecular interactions from forming.

meaning that all the hydrogen bonds are longer and therefore weaker as a result.

2.3 Conclusions

It has been shown that UIM **17** is a conformer independent array that can interact with similar energies through both conformations. This makes it ideal for use in studies that compare the binding affinities of AAD arrays because it does not matter which conformer of UIM **17** interacts, they both give the same results. Therefore, sterics should not affect the results because both conformers have similar binding affinities. Intramolecular hydrogen bonding, tautomeric preferences and electrostatic and steric interactions all play important roles in determining the preferred conformation of the individual molecules. It is important to ensure that the lowest energy conformation presents the desired hydrogen-bond array in order to obtain the highest possible binding

affinities. For the molecules used in this study it was found that preorganisation using hydrogen bonds did not dramatically increase the binding affinity of the complexes, as had been predicted. However, the molecular modelling results suggest that, if molecules containing intramolecular hydrogen bonds could be designed so that they present the desired AAD array in the lowest energy tautomer/conformer, then binding affinities could be increased.

Chapter 3

Remote Substituent Effect and Applications in Supramolecular Polymer Assembly

It was shown in Chapter 2 that the binding affinity of hydrogen-bonding motifs can be modulated by changing the extent to which molecules are preorganised using intramolecular hydrogen bonds.¹⁵⁹ There are many other factors that can also affect the stability of supramolecular complexes including the conformer/tautomer preference,^{60,63} the number of hydrogen bonds,¹⁷⁴ the arrangement of the hydrogen bonds⁴⁵ and the strength of the individual hydrogen bonds.²⁴ The strength of the individual hydrogen bonds is controlled by the pK_a of the atoms involved, which can be modulated by changing the electron donating/withdrawing nature of adjacent substituents. The substituent effect could then be used in supramolecular polymers as a simple way to access different binding properties, without having to redesign the hydrogen-bonding motif in order to obtain the desired binding affinity for the chosen application.

There have been attempts to quantify the effect that electron donating/withdrawing substituents have on the binding affinity of hydrogen-bonding arrays.^{28,175,176} In order to study the binding affinity of the DNA bases in non-aqueous solvent, Kyogoku *et al.* made some modifications to the bases (**65** and **66**) using electron donating/withdrawing substituents (Figure 3.1).¹⁷⁵ They found that, in general, binding affinities increased with electron withdrawing substituents. This is most likely because electron withdrawing groups remove electron density from the NH of uracil, making it more acidic and therefore better at forming strong hydrogen bonds (Table 3.1) However, the measured differences are small and no errors were reported, indicating that the measured differences represent an insignificant change in the binding affinity.

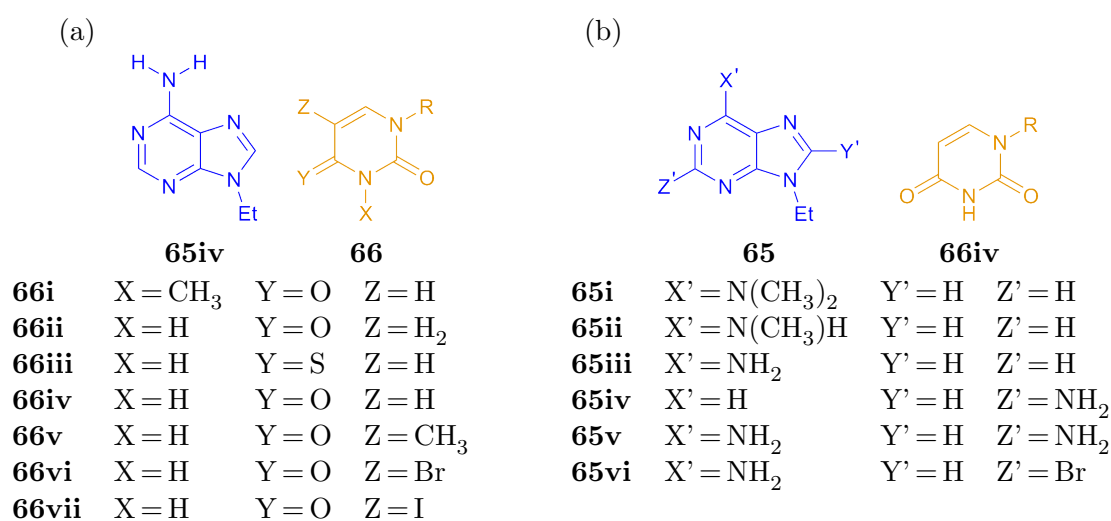
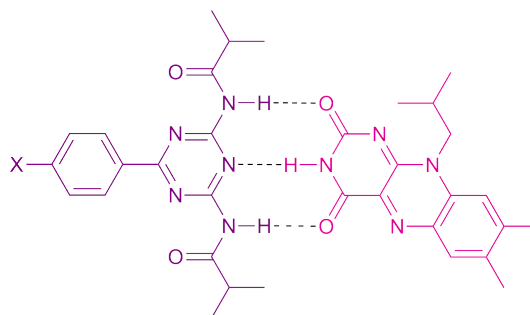


Figure 3.1: DNA base derivatives **65** and **66** designed by Kyogoku.

Table 3.1: Table to show the binding affinities of complexes in the study by Kyogoku.

Complex	K _a (M ⁻¹)	Complex	K _a (M ⁻¹)
65iii · 66i	<1	65i · 66iv	1.5
65iii · 66ii	30	65ii · 66iv	50
65iii · 66iii	90	65iii · 66iv	100
65iii · 66iv	100	65iv · 66iv	45
65iii · 66v	130	65v · 66iv	170
65iii · 66vi	240	65vi · 66iv	140
65iii · 66vii	220		

Another study of the effect that substituents can have on the binding affinity of complexes was carried out by Deans *et al.*,¹⁷⁶ who designed a receptor **67** for flavin **68**. Again, a correlation between binding affinity and the substituent was found (Figure 3.2), but the differences were small and there were some anomalies that did not fit the trend. The strongest interactions were obtained when electron withdrawing groups were present, presumably because they decreased the electron density of the two NH protons, affording stronger hydrogen-bonding atoms. Weaker interactions were obtained when electron donating groups were present because, although they increased the electron density of the triazine nitrogen, the binding capabilities of both NH's was reduced, making the overall binding affinity weaker as a result. It would normally be expected that, when hydrogen is the substituent, a binding affinity between that of electron donating and withdrawing groups would be obtained. In this example however, the hydrogen substituent had the lowest binding affinity, probably because it did not enhance or reduce the binding affinity of any of the hydrogen-bonding atoms. Therefore, although this study showed that there was a correlation between electron donating/withdrawing groups and the binding affinity of the complexes, it was difficult to predict the binding affinity of substituents that were not included in the experiment.

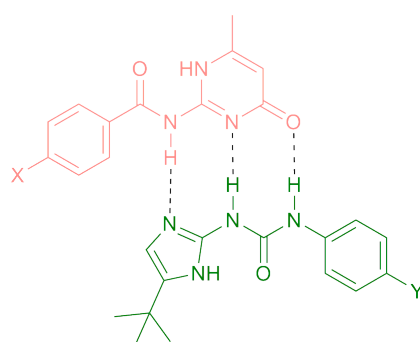
**67 · 68**

Molecule	Substituent (X)	Binding affinity (K_a)
67a	OMe	44 M^{-1}
67b	Me	53 M^{-1}
67c	H	24 M^{-1}
67d	CF_3	63 M^{-1}
67e	NO_2	97 M^{-1}

Figure 3.2: Binding affinity of **67** to **68** could be modulated by changing the electron donating/withdrawing nature of the substituent.

3.1 Substituent Control over the Binding Affinity of DDA · AAD Arrays

Although the studies described above showed that differences in binding affinity can be obtained by changing the electron donating/withdrawing nature of the substituent, the differences were small and it was difficult to predict what binding affinity untested molecules would have. For the substituent effect to be quantified, molecules with stronger binding affinities should be used, allowing any effects to be magnified. In order to ensure that a trend without anomalies can be observed, the substituent should be appended to the terminus of the array, which should allow each atom to be affected equally. This should also prevent steric interactions and the conformer preference of the molecules from changing, no matter which substituent was present. UIM **17** and AIC **57**,⁶⁰ described in Chapter 2, were chosen for this study because they interacted through an ADD · DAA array and both contained a phenyl ring appended to the end of the hydrogen-bond array, which was easy to functionalise (**57a-f** and **17a-i**). Because remote substituents were used a difference in binding affinity between electron donating and electron withdrawing groups was observed without interference from sterics and undesired conformations. This work was published in *Org. Lett.*¹⁷⁷

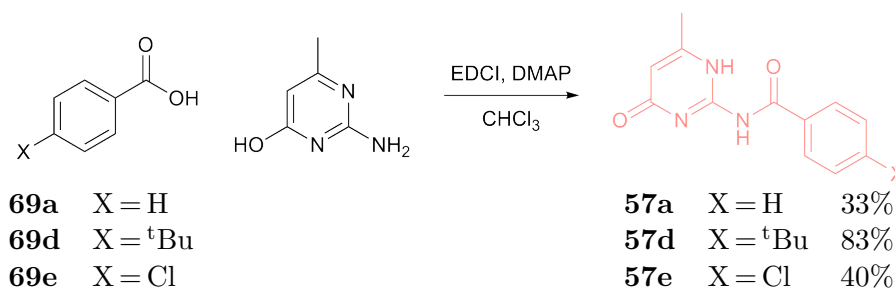


57 · 17

57a	X = H	17a	Y = H	17e	Y = Me
57b	X = MeO	17b	Y = CF ₃	17f	Y = Cl
57c	X = CO ₂ Me	17c	Y = MeO	17g	Y = F
57d	X = ^t Bu	17d	Y = CO ₂ Et	17h	Y = ^t Bu
57e	X = Cl				

3.1.1 Synthesis of the Molecules

Both series of molecules were synthesised^a using methods that had previously been developed within the Wilson group.⁶⁰ Compounds in the AIC series **57a-e** were synthesised in one step, which was amide coupling of the chosen benzoic acid **69** with *isocytosine*, using EDCI and DMAP as the activating agents (Scheme 3.1). A crystal structure of AIC **57d** was obtained by the slow evaporation of MeOH (Figure 3.3) and it was found to be in the same, undesired conformation as the crystal structure of AIC **57a** (Figure 2.10a). The second series (UIM **17a-h**) was synthesised in four steps (Scheme 3.2). Guanidine hydrochloride was protected to give Boc-guanidine **58** before cyclisation using 1-bromopinacolone was carried out to give the imidazole **59**. The next step was deprotection of the imidazole **59** to give the free amine **60**, which was reacted with the chosen *isocyanate* in the presence of NEt_3 , to give the desired UIM **17**. The final step was low yielding because it was difficult to separate the desired UIM **17** from the biproduct, diphenyl urea, leading to significant loss of product during purification.



Scheme 3.1: Synthesis of the AIC **57** series. Molecules shown were synthesised by myself, others were synthesised by Andrea McGhee and Adam Gooch.

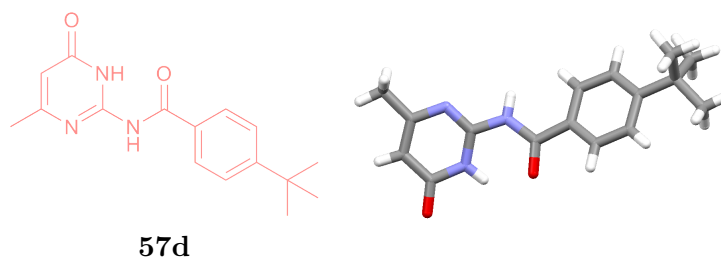
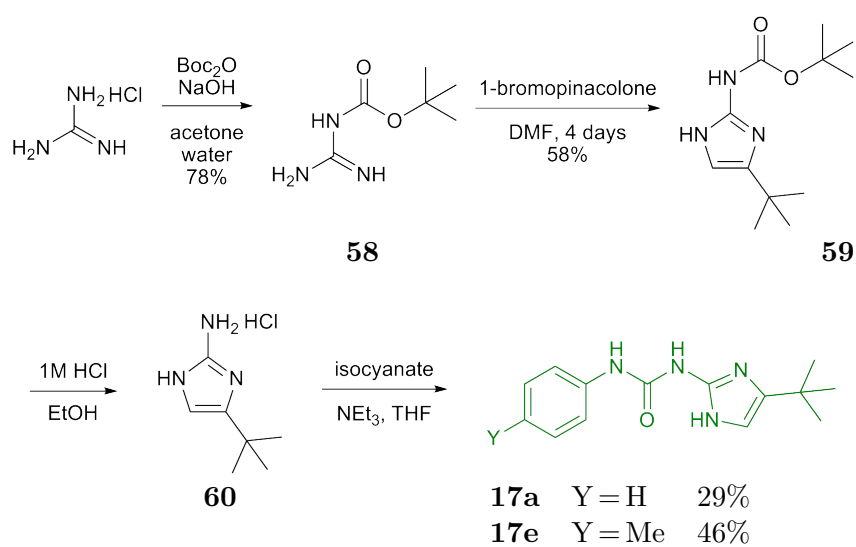


Figure 3.3: The crystal structure of AIC **57d** shows that an alternative conformation, which does not present the desired AAD array, is presented in the solid state.

^aAndrea McGhee, Adam Gooch and myself were all involved in the synthesis of the compounds.



Scheme 3.2: Synthesis of the UIM **17** series. Molecules shown were synthesised by myself, others were synthesised by Andrea McGhee and Adam Gooch.

3.1.2 ^1H NMR Titration Experiments

The self-association of UIM **17** and AIC **57** had already been studied⁶⁰ but dimerisation constants (K_{dim}) were required for the remaining molecules in each series to ensure that they would not impact on the association constants of the heterocomplexes. It was found that K_{dim} was negligible for all of the molecules (Table 3.2), so was not included in any of the calculations for association constants of the heterodimers. The K_a 's were measured^b for AIC **57a** with all of the UIM's **17a-h** and for UIM **17a** with all of the AIC's **57a-e** (Table 3.3). Each titration was carried out in triplicate in order to obtain more accurate results and attempts were made to remove human error from the results by using more than one person. K_a 's were then converted to ΔG 's and Hammett plots (Figure 3.4) were generated by plotting the sigma (σ) value¹⁷⁸ of the substituent against the measured ΔG for the given complex over ΔG for **57a** · **17a** ($\Delta G_{X \text{ or } Y} / \Delta G_H$). Errors were calculated (Equation 3.1) to show the standard deviation of the measurements.

Table 3.2: Table to show the dimerisation constants (K_{dim}) measured for each molecule.

Molecule	K_{dim} (M^{-1})	Molecule	K_{dim} (M^{-1})
57a	3.6 ± 0.3	17a	10.6 ± 2.2
57b	1.5 ± 0.3	17c	25.2 ± 1.0
57d	$< 1^*$	17e	2.0 ± 0.1
57f	$< 1^*$	17h	$< 1^*$
57c	4.7 ± 1.1	17f	$< 1^*$
		17g	$< 1^*$
		17d	12.5 ± 2.4
		17b	10.9 ± 1.8

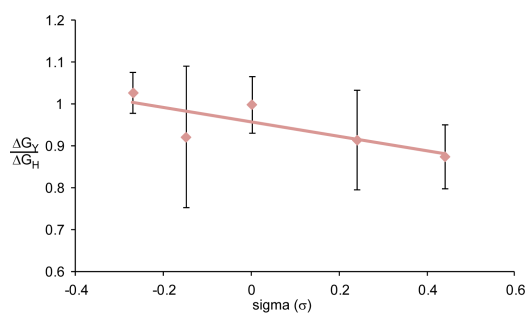
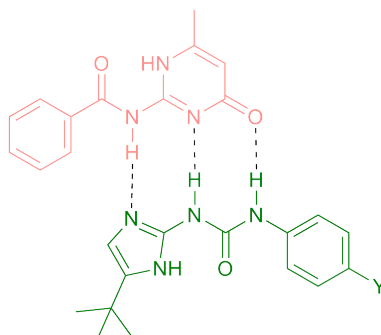
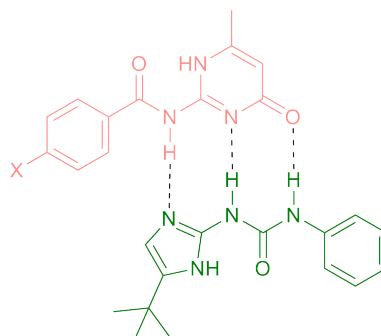
* Negligible change in shift was observed for these molecules, this equates to negligible dimerisation constants.

^bAll ^1H NMR titrations were carried out by myself and Adam Gooch, a former PhD student in the Wilson group.

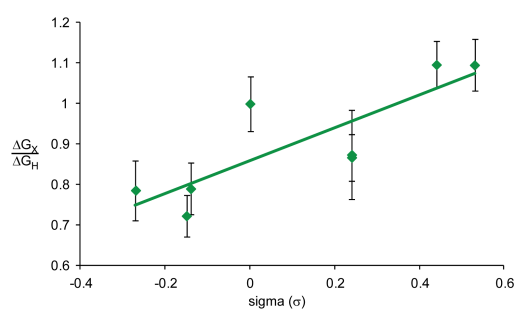
Table 3.3: Table to show the binding constants for substituents throughout both series and how these relate to their sigma values.

57-X	17-Y	σ	K_a ($\times 10^3 \text{ M}^{-1}$)	ΔG (kJ mol^{-1})
CO ₂ Me	H	0.44	10 ± 7.9	-22.4 ± 1.6
Cl	H	0.24	18 ± 13	-23.4 ± 2.8
H	H	0.00	33 ± 16	-25.6 ± 1.2
^t Bu	H	-0.15	25 ± 20	-23.6 ± 4.2
OMe	H	-0.27	41 ± 3.9	-26.3 ± 0.2

H	CF ₃	0.53	86 ± 34	-28.0 ± 0.9
H	CO ₂ Et	0.44	84 ± 22	-28.0 ± 0.6
H	Cl	0.24	11 ± 8.4	-22.4 ± 2.6
H	F	0.24	8.1 ± 2.8	-22.2 ± 1.0
H	H	0.00	33 ± 16	-25.6 ± 1.2
H	^t Bu	-0.15	1.9 ± 0.8	-18.5 ± 1.0
H	Me	-0.14	3.7 ± 2.2	-20.2 ± 1.3
H	OMe	-0.27	3.8 ± 2.2	-20.1 ± 1.6



(a)



(b)

Figure 3.4: Hammett plots to show the correlation between the association constants and substituent effect for (a) the AIC **57** series and (b) the UIM **17** series.

$$\frac{A_e^2}{A^2} = \frac{B_e^2}{B^2} + \frac{C_e^2}{C^2} \quad (3.1)$$

$$A = \frac{\Delta G_{X \text{ or } Y}}{\Delta G_H}, A_e = SD_{\frac{X \text{ or } Y}{H}}$$

$$B = \Delta G_{X \text{ or } Y}, B_e = SD_{X \text{ or } Y}$$

$$C = \Delta G_H, C_e = SD_H$$

In the AIC **57** series, a four fold difference in K_a was measured between the strongest and weakest interactions (**57c** $10100 \pm 7900 \text{ M}^{-1} \rightarrow$ **57b** $40900 \pm 3900 \text{ M}^{-1}$), which does not represent a significant effect. This was not surprising because, as has previously been suggested,¹⁷⁹ the carbonyl group can insulate against any effects that the substituent might have on the hydrogen-bonding atoms. It seems that, in this case, electron-withdrawing groups stabilise negative charge development on the NH nitrogen (Figure 3.5a), making it a better hydrogen bond donor. At the same time positive charge on the carbonyl carbon is destabilised, furnishing a weaker intramolecular hydrogen bond. The opposite was observed for electron donating groups; positive charge build-up on the carbonyl is stabilised, allowing stronger intramolecular hydrogen bonds to form. This gave a slightly stronger association constant when electron donating groups were present, even though the electron density of the NH increased. It was proposed that the increased binding affinity due to the stronger intramolecular hydrogen bond was cancelled out by the weakening of the NH hydrogen bond, leading to only a small change in the K_a as substituents were changed from electron-withdrawing to electron-donating. Conversely, there was nearly two orders of magnitude difference in the UIM **17** series (**17h** $1900 \pm 800 \text{ M}^{-1} \rightarrow$ **17b** $86200 \pm 33800 \text{ M}^{-1}$). There were no insulating groups present

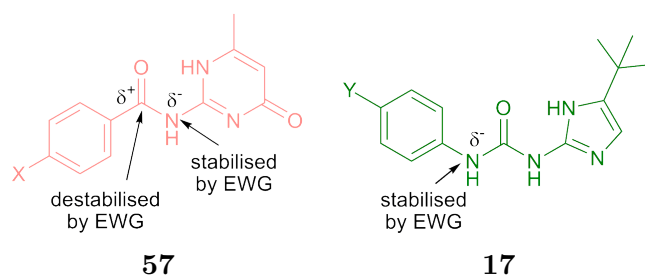


Figure 3.5: The proposed reasons for the observed differences in binding affinities for the AIC **57** and the UIM **17** series.

in this series, allowing the substituent to have a direct effect on the hydrogen-bonding atoms. Electron withdrawing groups could stabilise negative charge on the nitrogen bonds (Figure 3.5b), making the hydrogen more acidic, allowing it to form stronger hydrogen bonds. The opposite was true for electron donating groups, electron density increased at the NH's, affording weaker hydrogen bonds.

3.1.3 Molecular Modelling

Since the NH's did not appear in the ^1H NMR spectra of either series of compounds it was difficult to obtain structural information on the individual molecules and complexes using ^1H NMR techniques. Because of this, molecular modelling was used to provide further insight into the differences in binding affinities that had been measured. Following protocol from Jorgensen and Pranata,⁴⁵ DFT calculations at the B3LYP/6-31G* level of theory were carried out to obtain the geometry of each individual molecule using Gaussian03.¹⁶⁸ For the AIC **57** series, when the desired DAA array was presented, the phenyl ring was slightly twisted to prevent steric interactions between the ArCH and NH protons (Figure 3.6a). In contrast, the UIM **17** series could take on a planar geometry (Figure 3.6b) when the desired DDA binding array was presented, whichever conformation was adopted.

In order to compare the hydrogen bonding abilities of each molecule, the Mulliken values¹⁸⁰ were obtained for each atom involved in intermolecular hydrogen bonding (Table 3.4). Mulliken values are a measure of the electronegativity of atoms; stronger

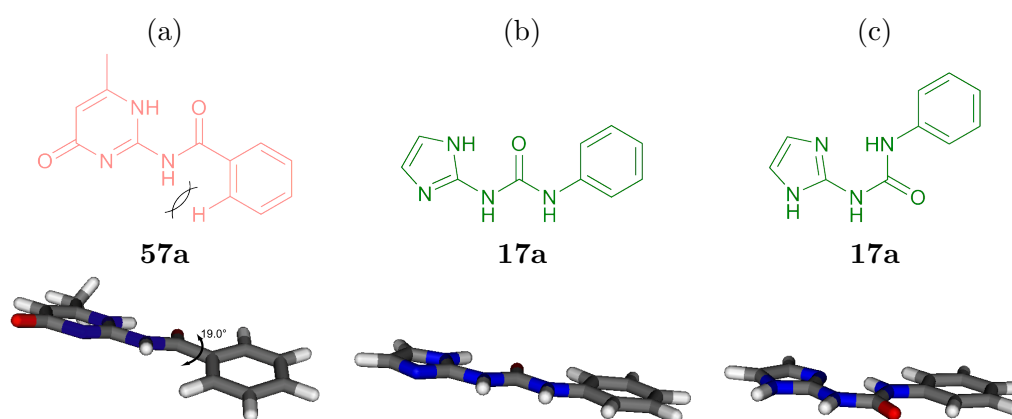
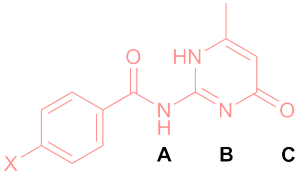


Figure 3.6: Figure showing (a) the steric interactions in AIC **57** that force the phenyl ring to twist and lose conjugation and (b) the planar structure of UIM **17** in (c) both conformations.

Table 3.4: Table to show how the Mulliken Values of the atoms involved in hydrogen bonding change as the strength of the interaction changes.

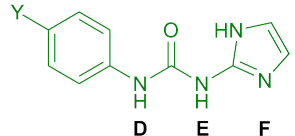
57-X	A	B	C
OMe	0.368025	-0.556871	-0.526338
Me	0.369315	-0.556096	-0.525138
H	0.369957	-0.555439	-0.524308
Cl	0.363866	-0.556925	-0.511179
CO ₂ Me	0.370544	-0.554845	-0.523051

↑ Increasing binding affinity



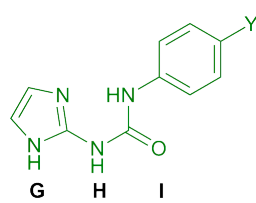
17-Y	D	E	F
CF ₃	0.328203	0.341659	-0.537103
CO ₂ Me	0.328262	0.341585	-0.537178
Cl	0.327507	0.341435	-0.357374
H	0.324916	0.339929	-0.538526
Me	0.323865	0.339470	-0.538972
OMe	0.324071	0.339681	-0.538843

↑ Increasing binding affinity



17-Y	G	H	I
CF ₃	0.337946	0.346312	-0.588508
CO ₂ Me	0.338111	0.346750	-0.586609
Cl	0.337779	0.345982	-0.590288
H	0.336564	0.343734	-0.590287
Me	0.336272	0.343162	-0.591346
OMe	0.336860	0.344270	-0.591015

↑ Increasing binding affinity



hydrogen bonds are formed when hydrogens are more acidic, with low electron density. Therefore, a more positive Mulliken value indicates that strong interactions should occur. For hydrogen bond acceptors (e.g. O and N) a more negative value leads to stronger interactions because this indicates that the atom has a higher electron density. ^1H NMR titration experiments showed that, for the AIC **57** series, electron donating groups gave the strongest binding affinities and the same trend was observed from the Mulliken values (Table 3.4). When electron donating groups were present, the hydrogen bond acceptors B and C became more negative. The hydrogen bond donor A however, became less positive, indicating that weaker hydrogen bonds could be formed. So, even though one hydrogen bond had weaker interactions, two increased their hydrogen bonding ability, which gave an overall gain in binding affinity. The UIM **17** series gave a similar result for both conformations. Two of the atoms (D,E and G,H) became stronger hydrogen bond donors, whilst the third (F and I) became weaker hydrogen-bonding atoms as the binding affinity increased.

This same effect could be visualised using electron density maps (Figure 3.7). Red shows areas of high electron density, allowing hydrogen-bond acceptors to form stronger intermolecular interactions. Blue shows areas of low electron density, allowing hydrogen-bond donors to form stronger intermolecular interactions. For the UIM **17** series the NH's became more blue (Figure 3.7a) when an electron withdrawing group was present. For the AIC **57** series the O and N became more red (Figure 3.7b) when an electron donating group was present. This result was in agreement with the ^1H NMR titration results.

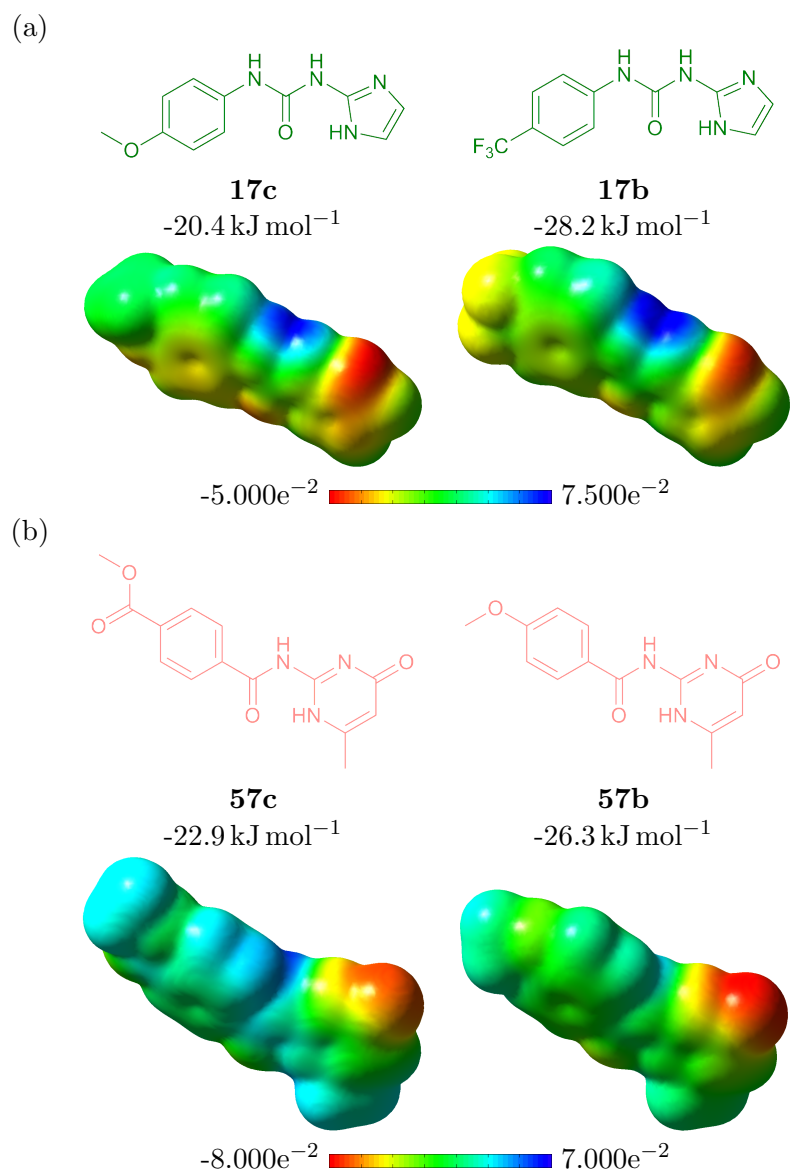


Figure 3.7: Electron density maps of the molecules that form the strongest and weakest interactions for (a) UIM **17** and (b) AIC **57**.

3.2 Substituent Effect in Supramolecular Polymers

It was proposed that the substituent effect described above can be applicable to many hydrogen-bonding motifs. The substituents are appended to the end of the hydrogen-bond array, so should not interfere sterically or conformationally. Currently, a popular area of research is in supramolecular polymers⁸⁴ because of their ease-of-synthesis, facile handling and new materials properties when compared to covalent polymers.¹⁸¹ A common supramolecular polymer monomer is the ditopic UPy **19** unit (Figure 3.8a),¹⁵⁶ which can assemble into polymers (Figure 3.8b) because UPy **19** is self-complementary.

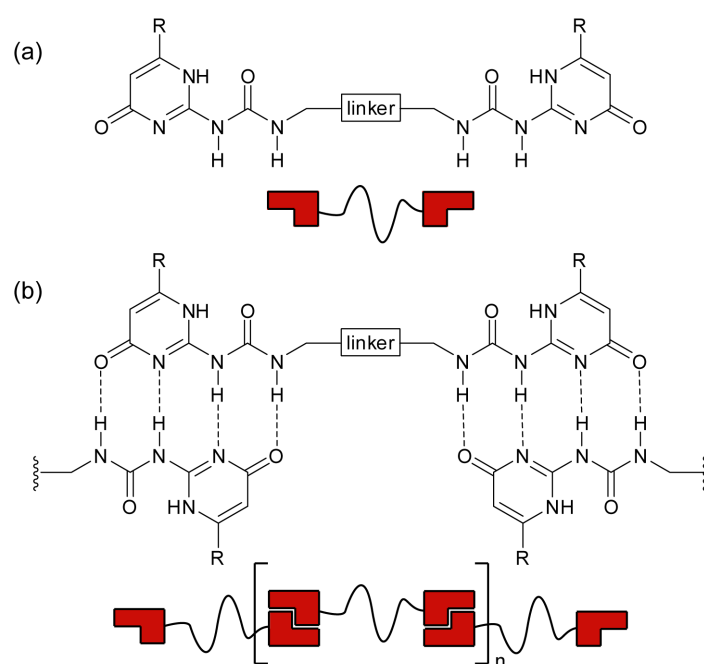


Figure 3.8: (a) The monomer building block. Self-complementary UPy can form (b) supramolecular polymers when linked together.

There are many examples of UPy **19** polymers in the literature. Short alkyl linkers^{65,117,182,183} have been used in order to study the ring/chain equilibrium of supramolecular polymers.^{120,184–186} In order to more easily obtain the desired materials properties the linker was extended so that it was a UPy **19** capped polymer chain.^{187–189} It can also be made so that the UPy **19** moiety is placed within the polymer chain,¹⁹⁰ allowing many contacts to be made between chains, making the assembly much stronger. The fact that UPy **19** can form strong interactions to DAN **21** can allow two different types of polymers to be mixed,^{191,192} enabling new materials properties to be made.

All of these examples use alkyl linkers between the UPy **19** units, but in order to study the remote substituent effect in supramolecular polymers, the UPy **19** units must be linked together through a phenyl ring (Figure 3.9). Using *isoelectronic* esters allows the desired electron donating/withdrawing properties to be obtained, whilst keeping the linkers the same, so that interactions between the linking groups should remain the same in both cases.¹⁹³ Although there it could be possible for the two *isoelectronic* supramolecular polymers to have different crystallisation/packing properties, they are as similar as possible, so any differences in the material properties can be assumed to be due to differences in their hydrogen bonding capabilities. By changing the electron donating/withdrawing nature of the linker it should be possible to change the strength of the intramolecular interactions and hence change the degree of polymerisation (DP). Since the remote substituent in the UPy molecules is attached to the hydrogen-bond array through the urea functionality, its binding properties should be analogous to the UIM **17** series. Electron withdrawing groups should afford more acidic NH's, making them better at forming hydrogen bonds, allowing large assemblies to accumulate. An electron donating group should have the opposite effect, increasing electron density at the urea NH's, affording a lower association constant, which should be visible by the formation of smaller assemblies in solution (Figure 3.9).

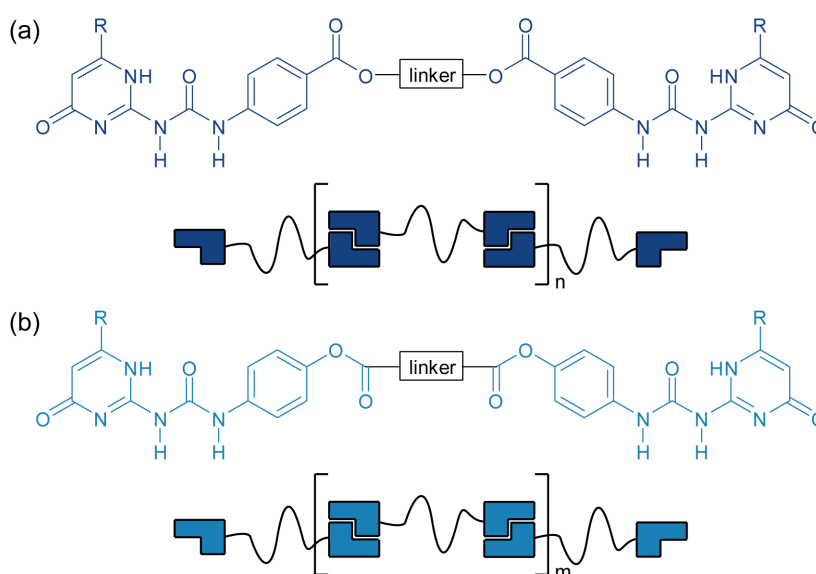
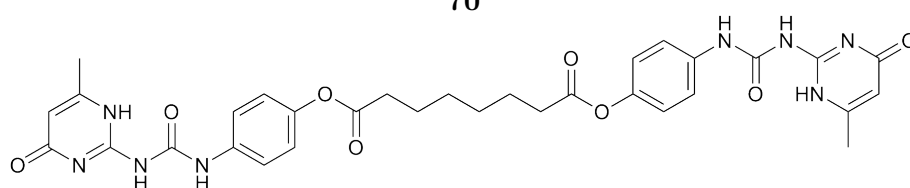
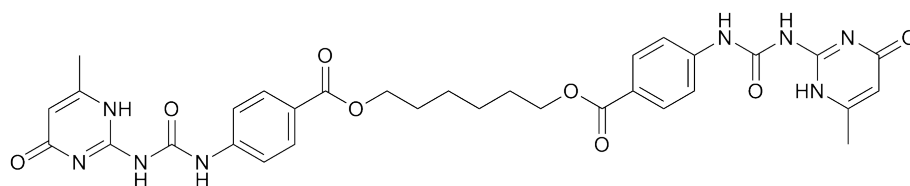


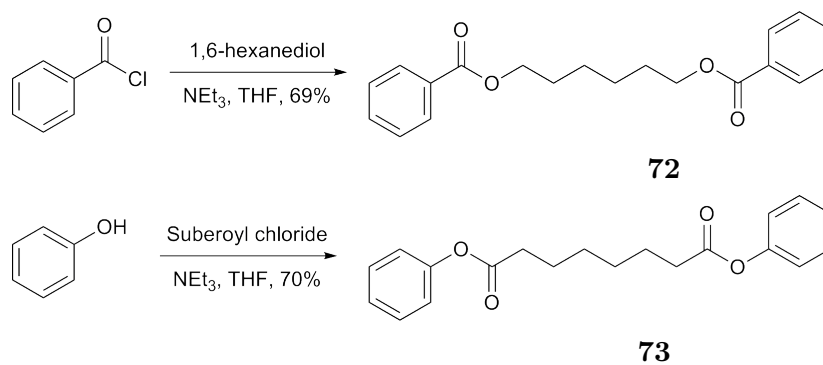
Figure 3.9: The two self-complementary polymers that were designed, with (a) electron withdrawing and (b) electron donating substituents. According to the substituent effect described in Section 3.1, when applied to the hydrogen-bonding motifs here, $n > m$ is predicted.

3.2.1 Synthesis of polymers

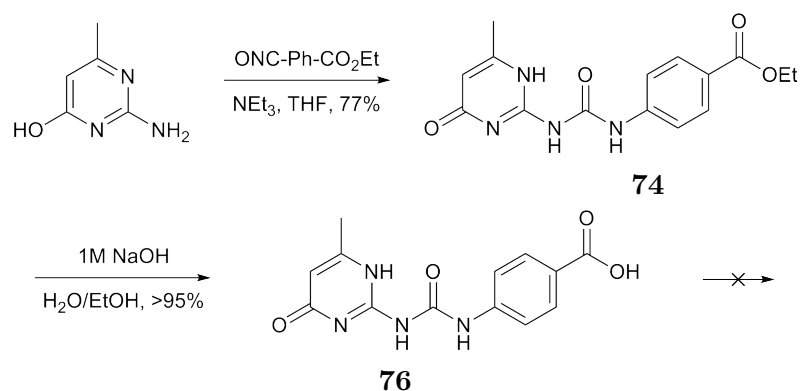
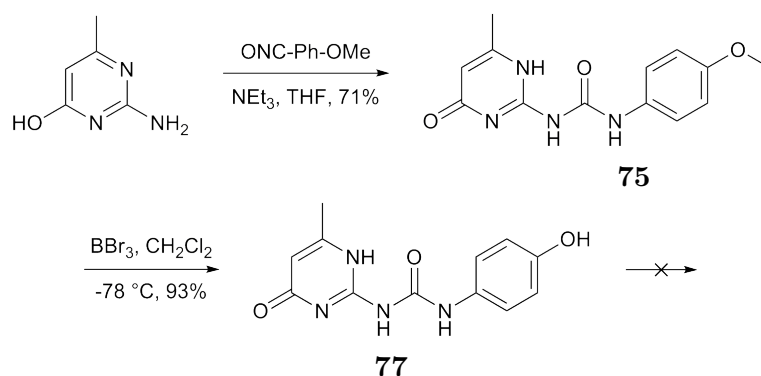
The monomer building block that was first proposed for the synthesis of both electron withdrawing and electron donating supramolecular polymers had a methyl group appended to the *isocytosine* functionality and a short, six-carbon linker (EWG **70** and EDG **71**).



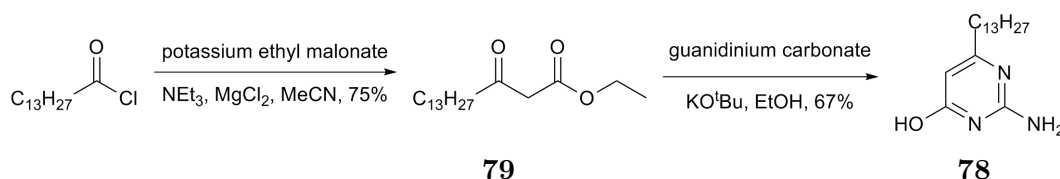
Investigation into the addition of the linker was tested using benzoyl chloride and phenol as mimics for the UPy functionality to give molecules **72** and **73** (Scheme 3.3). Because these reactions were successful, synthesis of the ditopic UPy's was attempted. Reaction of 2-amino-6-methylpyrimidin-4-ol with the desired *isocyanate* gave UPy-CO₂Et **74** and UPy-OMe **75**. UPy-CO₂Et **74** was then hydrolysed¹⁹⁴ using NaOH to give the UPy-COOH **76** (Scheme 3.4), which was insoluble, so further reactions could not be carried out. Demethylation of UPy-OMe **75** with boron tribromide,¹⁹⁵ to give the UPy-OH **77** (Scheme 3.5), also gave an insoluble material, which would not undergo further reactions. A different route was required so that the desired monomers could be synthesised.



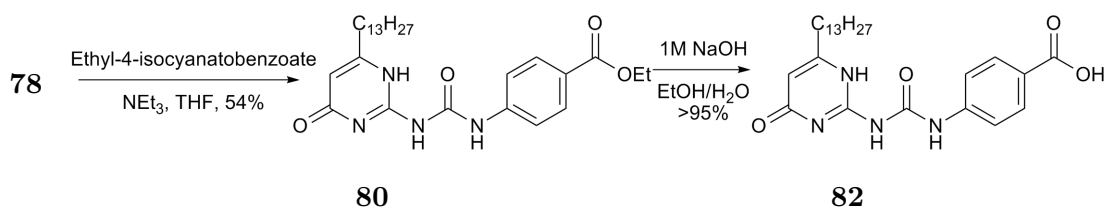
Scheme 3.3: Synthesis of the short linkers using benzoyl chloride and phenol.

Scheme 3.4: Synthesis of the insoluble UPy derivative **76**.Scheme 3.5: Synthesis of the insoluble UPy derivative **77**.

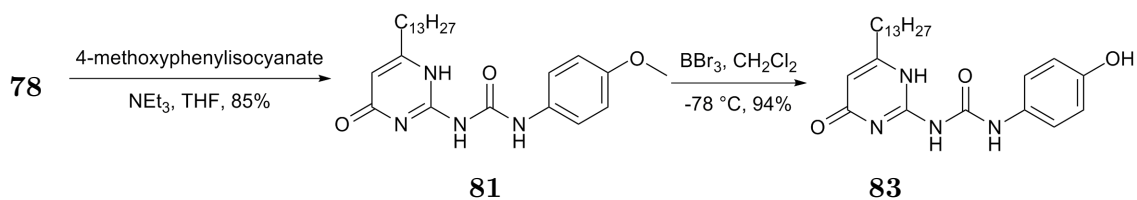
It was reasoned that addition of a long, solubilising alkyl chain to the UPy derivatives could make the precursor molecules more soluble, allowing further reactions to be carried out. Firstly *isocytosine* with a long alkyl chain **78** was synthesised (Scheme 3.6) by reaction of myristoyl chloride with potassium malonate to give ethyl 3-oxohexadecanoate **79**, which was cyclised by reaction with guanidinium carbonate to give the *isocytosine* **78**.¹⁹³ The *isocytosine* **78** was reacted with the chosen *isocyanate* to give the desired molecules, UPy-CO₂Et **80** (Scheme 3.7) and UPy-OMe **81** (Scheme 3.8). Hydrolysis of UPy-CO₂Et **80** was carried out in the presence of NaOH¹⁹⁴ to give UPy-COOH **82** and demethylation of UPy-OMe **81** was carried out in the presence of BBr₃ to give UPy-OH **83**.¹⁹⁵ Several attempts were made to synthesise the ditopic units with a short linker however, the desired products were never isolated.



Scheme 3.6: Synthesis of the *isocytosine* **78** with a long alkyl chain.

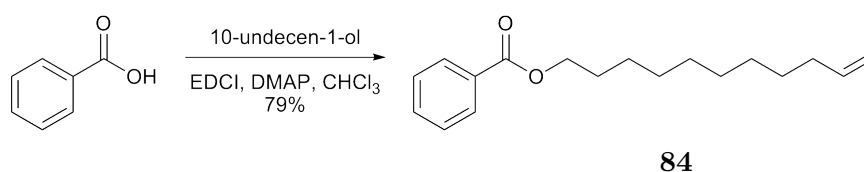


Scheme 3.7: Synthesis of electron withdrawing UPy.



Scheme 3.8: Synthesis of electron donating UPy.

It could be that the ditopic UPy's with a 6-carbon linker were not isolated because the direct difunctionalisation was not possible. Addition of a longer linker which contained a terminal alkene was first tested on benzoic acid (Scheme 3.9) to give **84**. The terminal alkene was present so that two molecules could be joined using metathesis to give the desired difunctionalised molecule. The mechanism¹⁶⁴ for the cross metathesis (Figure 3.10) is firstly a [2+2] cycloaddition between the carbene and the alkene to give a metallocyclobutane. This reaction then occurs in reverse, leaving the alkene attached to the Ru. A further [2+2] cycloaddition between another alkene in solution and the Ru carbene gives another 4-membered ring containing both functional groups. The reverse reaction can release the product, regenerating the catalyst.



Scheme 3.9: Synthesis of benzoic acid derivative **84** functionalised with a long alkyl chain with a terminal alkene.

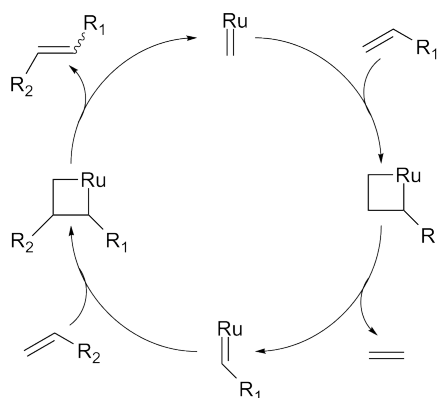
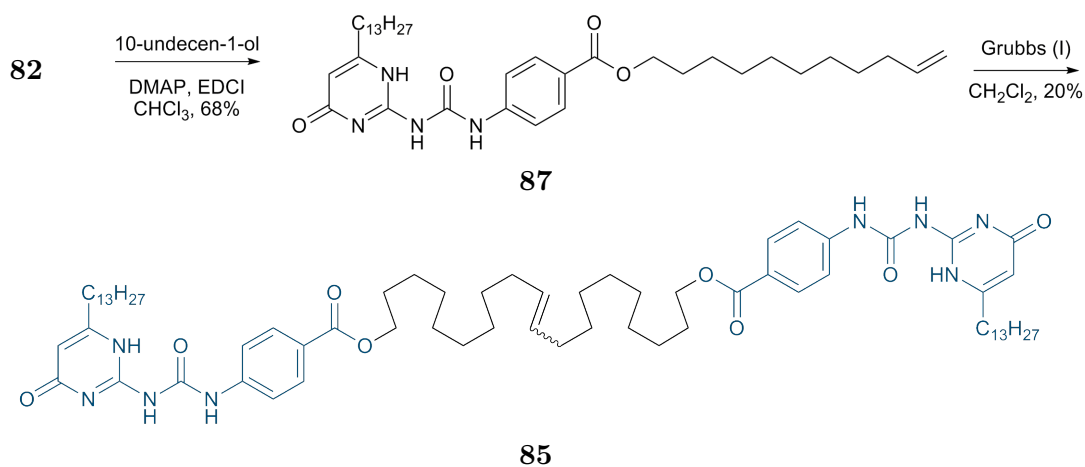
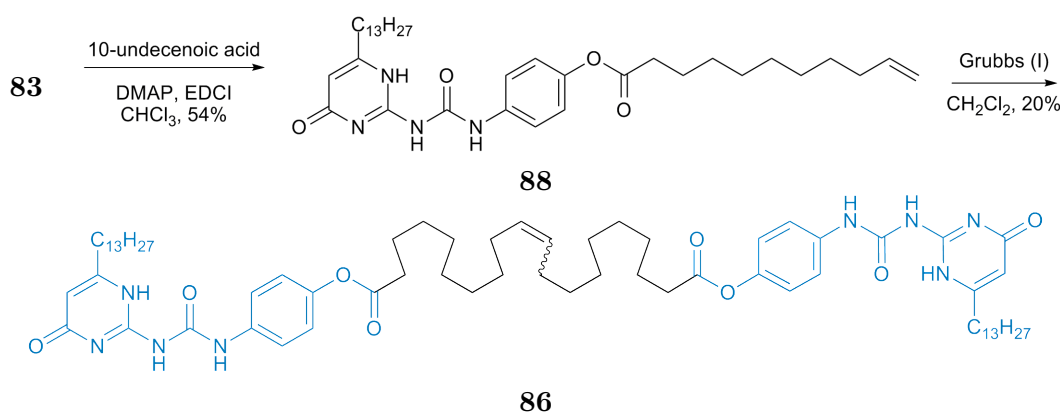


Figure 3.10: Cross metathesis mechanism.

The strategy of esterification followed by metathesis was then used in the synthesis of both UPy-EWG **85** and UPy-EDG **86** (Scheme 3.10 and Scheme 3.11). Esterification was carried out between the acid and alcohol precursors using DMAP and EDCI as the coupling reagents to give the UPy esters **87** and **88**. The metathesis reaction could then be carried out using the first generation of Grubbs catalyst.¹⁹⁶ The solvent was degassed for 1 hr before the alkene precursor was charged to the flask. The catalyst was then added dropwise over 10 hrs before the reactions mixture was left to stir for a further 12 hrs. The ditopic UPy's **85** and **86** were both obtained as a mixture of *cis*- and *trans*-alkene, but since they were probably both obtained in the same ratio (difficult to ascertain from the broad ¹H NMR spectra) in both polymers they were taken on as a mixture. Since such a long flexible linker was used it should not cause problems having both *cis*- and *trans*-alkene present because they should both have the same binding properties.

Scheme 3.10: Synthesis of UPy-EWG **85**.Scheme 3.11: Synthesis of UPy-EDG **86**.

3.2.2 Self-Assembly into Polymers

There are many different materials properties that can be measured, but to measure the differences between **85** and **86** the size of the assemblies and melting temperatures are important factors. Differential scanning calorimetry (DSC) is a technique that measures melting transitions of polymers and it was thought that if the strength of the intermolecular hydrogen bonds in **85** and **86** was different, then different melting temperatures would be obtained. Although a transition was observed for UPy-EWG **85** at about 70 °C, no transition was observed for UPy-EDG **86** (Figure 3.11). This could be because there are no transitions, but it could also be because the transition is very broad, so cannot be detected. These initial results indicated that DSC was not a suitable technique for the study of these types of supramolecular polymers, so no further experiments were carried out.

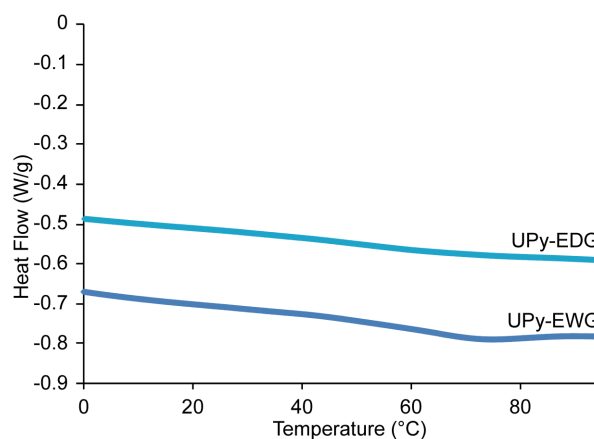


Figure 3.11: DSC experiment for both UPy-EWG **85** and UPy-EDG **86** indicates that there is a transition at ~ 70 °C for **85**, but that there is no transition for **86**.

Because interactions between these ditopic molecules can result in polymer networks being assembled, the viscosity of the solution can be changed when different concentrations of monomer are present and also when different types of monomer are present. It is possible to measure the viscosity of the solution using a viscometer (Figure 3.12); the time taken for the solution to travel between two lines is measured and compared to that of the solvent alone. Differences between chloroform and the supramolecular polymer solutions at different concentrations were measured (Table 3.5) however, the differences between the two polymers were found to be negligible. There are possibly two reasons for this: (1) the travel time was too short, making it difficult to know if differences are real, or measurement errors and (2) the solubility of the polymers was poor, resulting

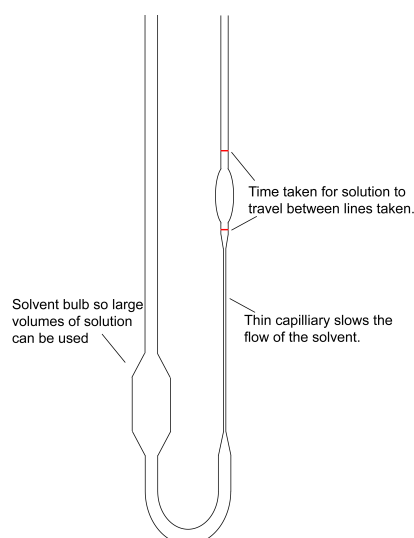


Figure 3.12: A cartoon representation of the viscometer that was used in these experiments. The solution was sucked up the right hand tube before the time taken for it to travel between the lines under gravity control was measured.

Table 3.5: Table to show the measured values for different concentrations of both supramolecular polymers.

Polymer	(Conc mM)	Run 1 (s)	Run 2 (s)	Run 3 (s)	Average (s)
	(0)	45.30	45.50	45.94	45.58 ± 0.33
85	(10.42)	48.25	48.40	48.33	48.33 ± 0.08
85	(8.34)	47.79	48.08	47.86	47.91 ± 0.15
85	(6.95)	47.61	48.15	47.82	47.86 ± 0.27
86	(12.59)	51.89	51.94	51.95	51.93 ± 0.03
86	(10.07)	50.31	50.52	50.35	50.39 ± 0.11
86	(8.40)	50.05	49.90	50.42	50.12 ± 0.27
86	(7.20)	49.40	49.81	50.01	49.74 ± 0.31
86	(6.30)	49.57	50.66	50.16	50.13 ± 0.55

in the capillary tube getting blocked which made measurements unreliable. These results indicated that viscometry was not sensitive enough to measure differences in the assembly properties of these similar supramolecular polymers.

Since the polymers did not stay in solution during the viscometry experiment, gelation was also attempted in order to detect differences between these two supramolecular polymers. It was thought that UPy-EWG **85** would form gels at lower concentrations than UPy-EDG **86** because of its predicted stronger hydrogen bond interactions, allowing larger polymer networks to assemble. It was found that the only solvent that the polymers would dissolve in was chloroform and gels were formed by **85** at 40 mM after 40 days, whereas no gelation was observed in the 40 mM solution of **86**. Although this does represent a difference in the two supramolecular polymers, the long time taken for

the gel to form is not viable as a form of testing the differences in these polymers. It could be that gelation only occurred because some chloroform evaporated resulting in a higher concentration of polymer, or that the concentration of H₂O increased over time, resulting in a solvent conformation that the supramolecular polymer would not dissolve in.

DOSY can be used to measure the size of the assemblies that are present in solution, so could be used to see if there are differences in the DP of the **85** and **86** supramolecular polymers at different concentrations. As the concentration was decreased it was found that the ¹H NMR peaks sharpened for both polymers, which was indicative of smaller assemblies being present in solution. The smaller the assemblies, the quicker they tumble through solution and so faster diffusion coefficients (D) are measured. D was measured for both UPy polymers (**85** and **86**) at 30, 20, 10 and 1 mM concentrations (Table 3.6). Because D was very similar for 10 and 1 mM in both polymers it was reasoned that this was the smallest assembly that could be formed in solution; probably representing a single molecule. For UPy-EWG **85**, D for 1 and 10 mM were similar, but when higher concentrations were tested D began to decrease, indicating that larger assemblies were formed. For UPy-EDG **86**, the 20 mM solution had a similar D to the 1 and 10 mM solutions and D only began to decrease at 30 mM concentrations. Assuming that the quickest D represented a single monomer at low concentrations, the DP could be calculated. Firstly the molecular weights (MW) of the assemblies at each concentration were calculated (Equation 3.2) where D_[low] was the D for 1 mM solutions and MW_[low] was assumed to be the monomer. The DP was then calculated by dividing the weight of the polymer assembly by the weight of the monomer units (Equation 3.3). Plotting the DP results against concentration (Figure 3.13) showed that the EWG **85** began to polymerise at lower concentrations than the EDG **86**. This showed that higher concentrations of UPy-EDG **86** were required to initiate polymer assembly, indicating

Table 3.6: Diffusion coefficients and the DP that these correlate to for the polymers at different concentrations.

Conc. (mM)	UPy-EWG 85		UPy-EDG 86	
	D ($\times 10^{-10} \text{ m}^2 \text{ s}^{-1}$)	DP	D ($\times 10^{-10} \text{ m}^2 \text{ s}^{-1}$)	DP
30	2.315 ± 0.005	16.58 ± 0.48	2.089 ± 0.010	12.57 ± 0.29
20	3.652 ± 0.016	4.22 ± 0.09	4.470 ± 0.120	1.28 ± 0.06
10	5.817 ± 0.007	1.05 ± 0.03	5.040 ± 0.025	0.90 ± 0.02
1	5.903 ± 0.070	1.00	4.858 ± 0.060	1.00

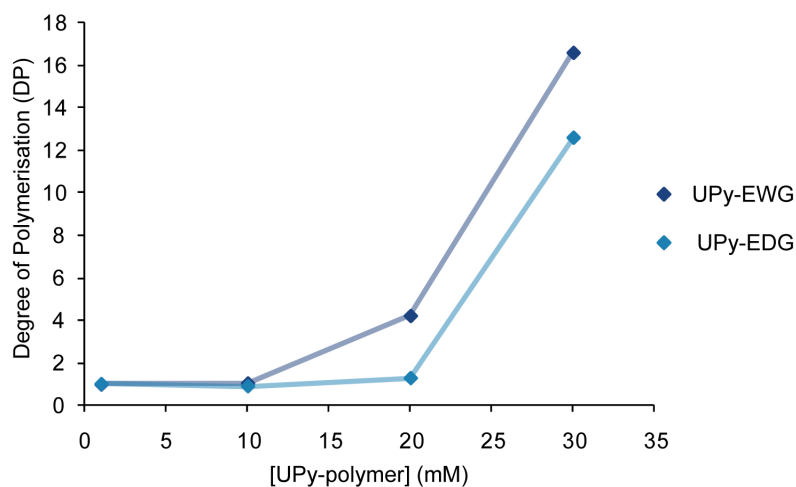


Figure 3.13: Graph to show the correlation between the concentration and DP for both polymers, UPy-EWG **85** and UPy-EDG **86**. Lines serve to guide the eye.

that weaker interactions were present compared to UPy-EWG **85**, which is in agreement with the predictions from earlier experiments (Section 3.1).

$$\frac{D_{[\text{low}]}}{D_{[\text{high}]}} = \frac{\sqrt[3]{MW_{[\text{high}]}}}{\sqrt[3]{MW_{[\text{low}]}}} \quad (3.2)$$

$$DP = \frac{MW_{\text{polymer}}}{MW_{\text{monomer}}} \quad (3.3)$$

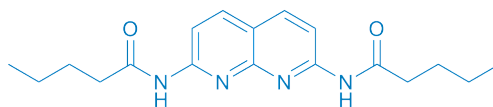
3.3 Conclusions

It has been shown that the binding affinity of supramolecular complexes can be modulated, simply by changing the electron donating/withdrawing nature of the substituent. By using remote substituents, connected to the end of the linear hydrogen-bond array through conjugation, it was possible to alter the binding affinities without favouring alternative conformations or sterically hindering the binding face of the molecules. The substituent effect was then applied to other hydrogen-bonding motifs that are commonly used in the assembly of supramolecular polymers. Although differences in the size of the assembly were observed using DOSY, the other techniques were not sensitive enough to measure differences between the two supramolecular polymers. This could indicate that using the remote substituent effect to modulate the properties of supramolecular polymers should only be used to fine tune the characteristics, not to make large changes.

Chapter 4

A Novel ADDA Quadruple Hydrogen-Bond Array

As shown in Chapters 2 and 3, the binding affinity of molecules can be manipulated by altering molecular features (e.g. conformation and electronics) that are not directly involved in hydrogen bonding. However, simply changing the strength of interactions will probably not allow the assembly of a high fidelity self-sorting system. New hydrogen-bonding motifs which interact using different arrays must be designed in order to achieve the desired high affinity and high fidelity systems. There are examples of hydrogen-bonding motifs that present ADDA arrays, allowing them to interact with the complementary DAAD array diamidonaphthalene (DAN) **21**; however there is an ongoing requirement for molecules that can achieve high affinity interactions to DAN **21** without competition from homodimerisation.



21

Diphenylurea (DPU) **89** dimerises with DAN **21** with a low association constant ($K_a = 10^3 \text{ M}^{-1}$).^{152,197} The reason for this low association constant has been attributed to DPU **89** being preferentially in the folded state,⁵⁹ to allow the formation of an intramolecular hydrogen bond (Figure 4.1). This results in a self-complementary AD array being presented, allowing DPU **89** to weakly self-associate ($K_{\text{dim}} = 5 \text{ M}^{-1}$ at 50°C).⁶¹ Ditrizoleurea (DTU) **90** has recently been introduced by Hisamatsu *et al.* (Figure 4.1).

Although DTU **90**, like DPU **89**, can fold so that an intramolecular hydrogen bond is formed allowing self-association to occur ($K_{\text{dim}} = 10^2 \text{ M}^{-1}$), it forms strong interactions with DAN **21** ($K_{\text{a}} = 10^5 \text{ M}^{-1}$).⁴⁰ The difference in binding affinities between DPU **89** and DTU **90** with DAN **21** has been attributed to steric hindrance between the carbonyl oxygen and aromatic protons in the unfolded conformers. In DPU **89** there is a short distance between the oxygen and proton, disfavouring the unfolded conformer. For DTU **90** the distance between the oxygen and proton is greater because of the size and shape of the five-membered-ring, allowing the unfolded conformer to be adopted with less steric hindrance.

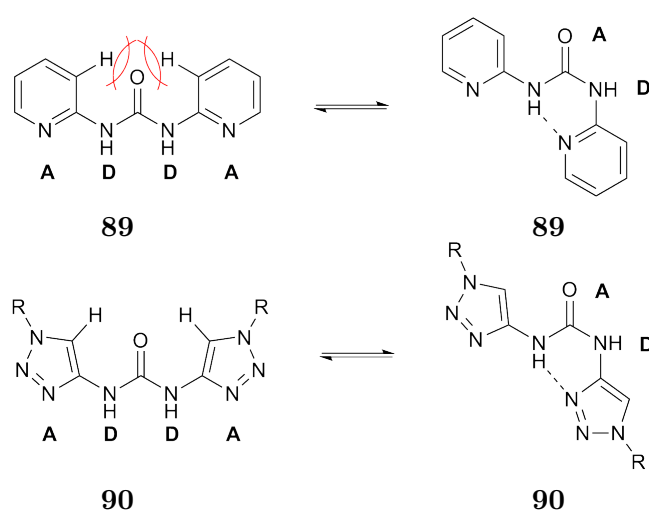


Figure 4.1: The two conformers of DPU **89** and DTU **90** which present the desired ADDA and self-complementary AD arrays. Steric hindrance in DPU **89** favours the folded conformation, preventing formation of the heterodimer.

Other hydrogen-bonding motifs that have been shown to interact with DAN **21** can also adopt alternative tautomeric states, allowing them to present self-complementary hydrogen-bonding arrays. Ureidopyrimidine (UPy) **19**, introduced by Meijer, forms a strong homodimer⁶⁶ ($K_{\text{dim}} > 10^5 \text{ M}^{-1}$) but on addition of DAN **21** this homodimer dissociates in favour of formation of the heterodimer ($K_{\text{a}} > 10^5 \text{ M}^{-1}$) (Figure 4.2).^{68,198,199} Although both the homo- (**19**·**19**) and heterocomplex (**21**·**19**) have strong association constants, formation of the heterocomplex is preferred, presumably because it allows the largest number of hydrogen bonds to be formed. There are six hydrogen bonds in the UPy homodimer, but this dissociates in order to form two UPy·DAN **19**·**21** complexes, which contain ten hydrogen bonds altogether. Zimmerman has also designed molecules that can self-associate and interact with DAN **21**. Deazapterin (DeAP) **20**, like UPy **19**, self-associates strongly ($K_{\text{dim}} > 10^5 \text{ M}^{-1}$) but can also interact with DAN

21 ($K_a > 10^5 \text{ M}^{-1}$). In an attempt to increase the fidelity of molecular association with DAN **21**, Zimmerman *et al.* designed ureidoguanosine (UG) **22**, which presents the desired ADDA array when the lowest energy tautomer is adopted.³⁷ The association constant of UG · DAN **22** · **21** complex is still high ($K_a = 10^8 \text{ M}^{-1}$) but the self-association of UG **22** ($K_{\text{dim}} = 10^2 \text{ M}^{-1}$) is much lower than that of UPy **19** and DeAP **20**.²⁰⁰ This means that the fidelity of the UG · DAN **22** · **21** system is high, but there still remains a requirement for molecules that form strong interactions to DAN **21** with even lower dimerisation constants than UG **22**. The design, synthesis and molecular recognition properties of one possible molecule is described below; this work has been submitted to *Org. Biomol. Chem.*²⁰¹

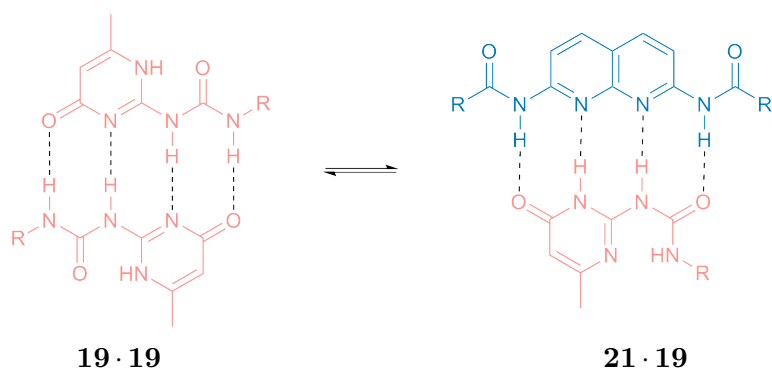


Figure 4.2: The UPy dimer **19** · **19** dissociates on addition of DAN **21** to preferentially form the heterodimer **21** · **19**.

4.1 Design and Synthesis of UDIM and DAN

Extension of the previously studied UIM **17** ADD array with a second imidazole gave the ureidodiimidazole (UDIM) **91**, which presents an ADDA array when it is in the unfolded conformation. As in UIM **17**, UDIM **91** can adopt two different conformers (Figure 4.3), along with several different tautomers. It was reasoned that the desired conformer would be easily adopted because it contained two stabilising intramolecular hydrogen bonds, the same as the undesired folded conformer, and none of the alternative conformer/tautomer states presented self-complementary hydrogen-bond arrays.

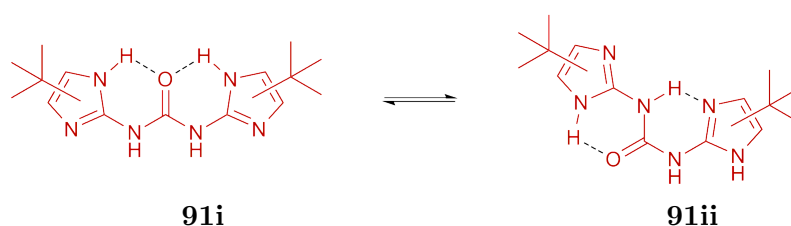
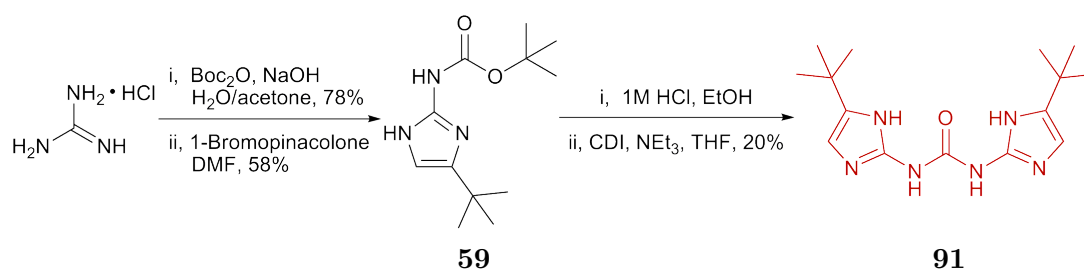


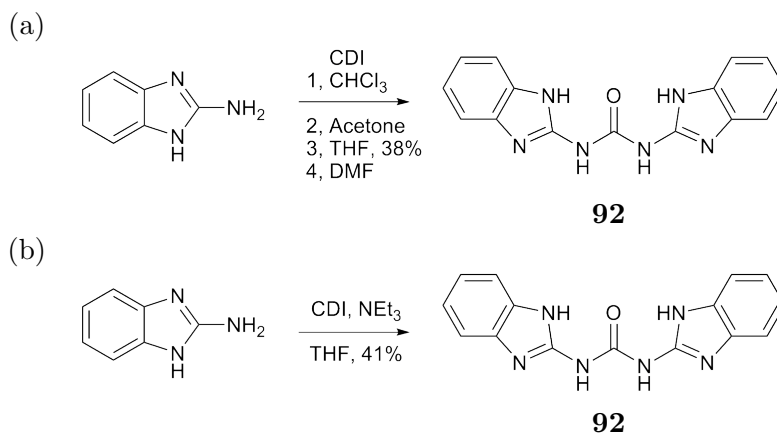
Figure 4.3: The two possible conformations that UDIM **91** could adopt whilst retaining two intramolecular hydrogen bonds.

UDIM **91** was synthesised (Scheme 4.1) following a similar procedure to that of UIM **17** (see Section 2.1.1). Protection of guanidine to give the boc-guanidine **58**, followed by cyclisation with 1-bromopinacolone gave the boc-imidazole **59** which was deprotected using 1 M HCl in EtOH to give the aminoimidazole **60**. Initial reactions to form the UDIM **91** were all unsuccessful, so the synthesis was optimised using 2-aminobenzimidazole to give the ureidodibenzimidazole **92**. 1,1-Carbonyldiimidazole (CDI) was reacted with 2-aminobenzimidazole and a range of solvents were tested (Scheme 4.2). It was found that THF was the solvent of choice for the reaction because it gave the largest conversion to the product which could be collected simply by filtration of the reaction mixture. Because the aminoimidazole **60** is present as the hydrochloride salt a base must be added to the reaction mixture in order to reveal the free amine. The addition of NEt_3 was tested



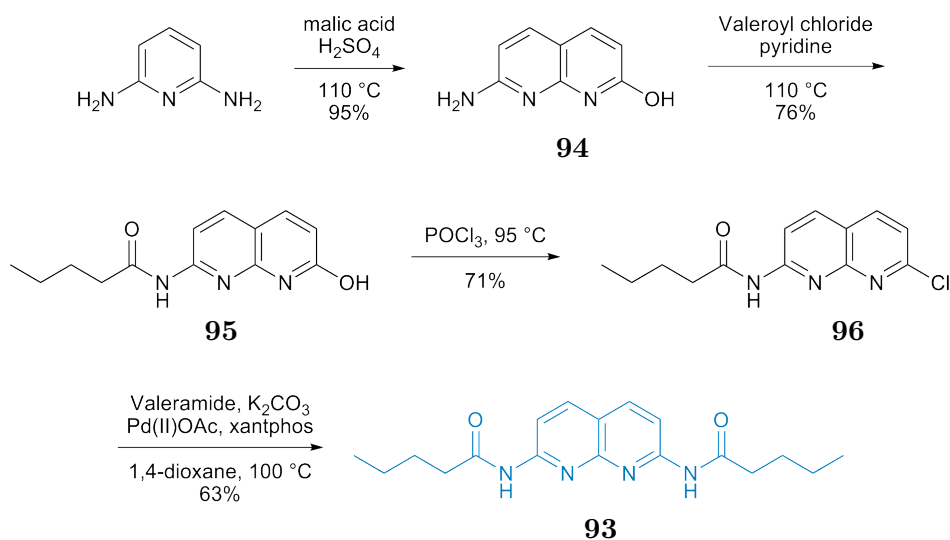
Scheme 4.1: The synthesis of UDIM **91**.

in the reaction of 2-aminobenzimidazole with CDI in THF to investigate whether the addition of NEt_3 could be tolerated. Similar conversion rates were observed with and without NEt_3 , indicating that it was not a problem for it to be present. The aminoimidazole **60** was then reacted with 1,1-carbonyldiimidazole in the presence of NEt_3 to give the UDIM **91**. Although this final step was low yielding no attempts were made to optimise the reaction because enough material was obtained in order to carry out the desired experiments.



Scheme 4.2: The synthesis of ureidodibenzimidazole **92** was used to test the formation of the urea. Firstly (a) the ideal solvent was found and secondly (b) the addition of NEt_3 to the reaction was tested.

DAN **93** was synthesised (Scheme 4.3) using a modified literature route.^{191,202} Reaction of 2,6-diaminopyridine and malic acid gave the aminohydroxy naphthalene **94**. The aminohydroxy naphthalene **94** was then reacted with valeroyl chloride in pyridine to give the naphthalamide **95** which was converted to the chloride **96** using POCl_3 . Addition of valeramide was then carried out in the presence of catalytic palladium to give DAN **93**.

Scheme 4.3: The synthesis of DAN **93**.

4.2 Conformational Studies of UDIM

Since UDIM **91** is symmetrical it should have four (broad) signals in the ^1H NMR spectrum, if rapid exchange between the different conformer/tautomer states occurs. However, the ^1H NMR spectrum of UDIM **91** at room temperature contained more than four signals (Figure 4.4a), which indicated that slow exchange between the different tautomer/conformer states occurred on the NMR timescale. The ^{13}C NMR spectrum also indicated that slow exchange was occurring because there were many more signals than would normally be expected (Figure 4.4b). Cold temperature ^1H NMR (Figure 4.4c) confirmed that many conformer/tautomer states were present in solution, with some being more prevalent than others. On heating (Figure 4.4d and e) the signals began to converge into four broad NH signals, which indicated that exchange between conformers/tautomers had become more rapid, but since they had not completely converged into the expected two NH signals this exchange was still quite slow.

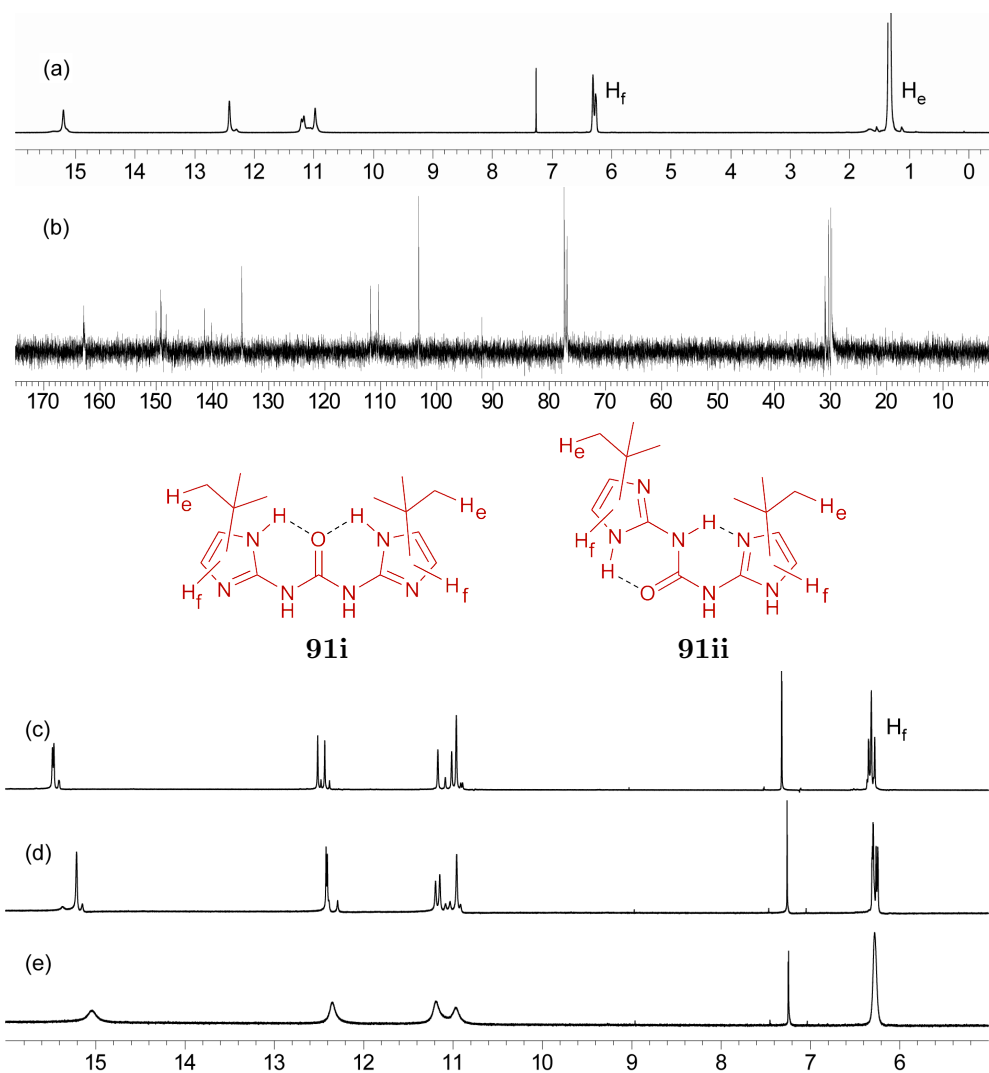
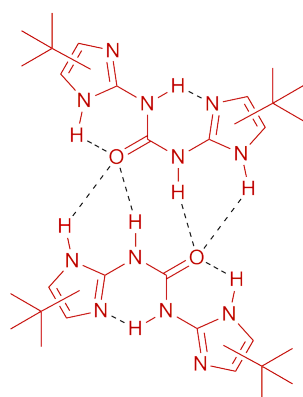


Figure 4.4: (a) ^1H NMR spectra of UDIM **91** (40 mM, 300 MHz, CDCl_3) and (b) ^{13}C NMR spectra of UDIM **91** confirm that slow exchange between tautomer/conformer states occur allowing many states to be observed (40 mM, 125 MHz, CDCl_3). Variable temperature ^1H NMR shows that at high temperatures the signals converge to give a more symmetrical spectra. (c) 213 K, (d) 293 K (e) 333 K (40 mM, 500 MHz, CDCl_3).

Ideally, in order for the strongest possible interactions to be formed the unfolded conformer of UDIM **91** must be the most prevalent in solution. However, the folded conformer has the potential to assemble into heterodimers, through bifurcated hydrogen bonds (Figure 4.5), so it could be possible that the folded conformer is favoured over the desired unfolded conformer. 2D NMR experiments were carried out in order to investigate which conformer of UDIM **91** was the most prevalent in solution. HMQC (Figure 4.6) showed that H_f was in two different environments within the molecule. There were correlations to two different carbon atoms from the ArCH proton, which confirmed that the molecule was not symmetrical in solution. However, this could arise from either conformation being present (Figure 4.6), so could not establish the structure of UDIM **91**.



91ii · 91ii

Figure 4.5: The folded conformer of UDIM **91ii** has the potential to assemble into homodimers through a bifurcated hydrogen-bond array.

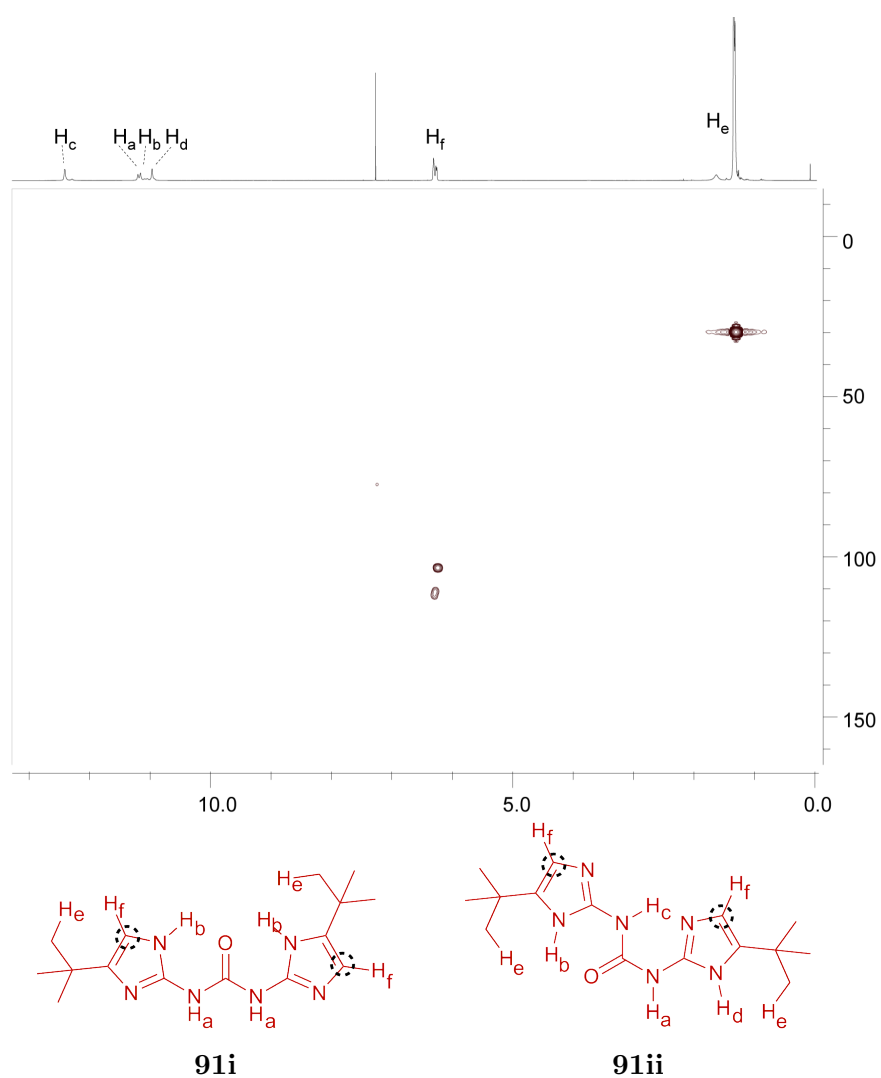


Figure 4.6: HMQC of UDIM **91** shows that H_f is in two different environments.

^1H - ^1H NOESY (Figure 4.7) showed that every proton was correlated to all of the other protons. However, this could have been obtained from either conformer state, because they can both place all protons within close proximity of all the other protons. One problem with this experiment is that it was not possible to distinguish between through-space correlations and proton exchange peaks, so many of the observed interactions may simply be exchange peaks.

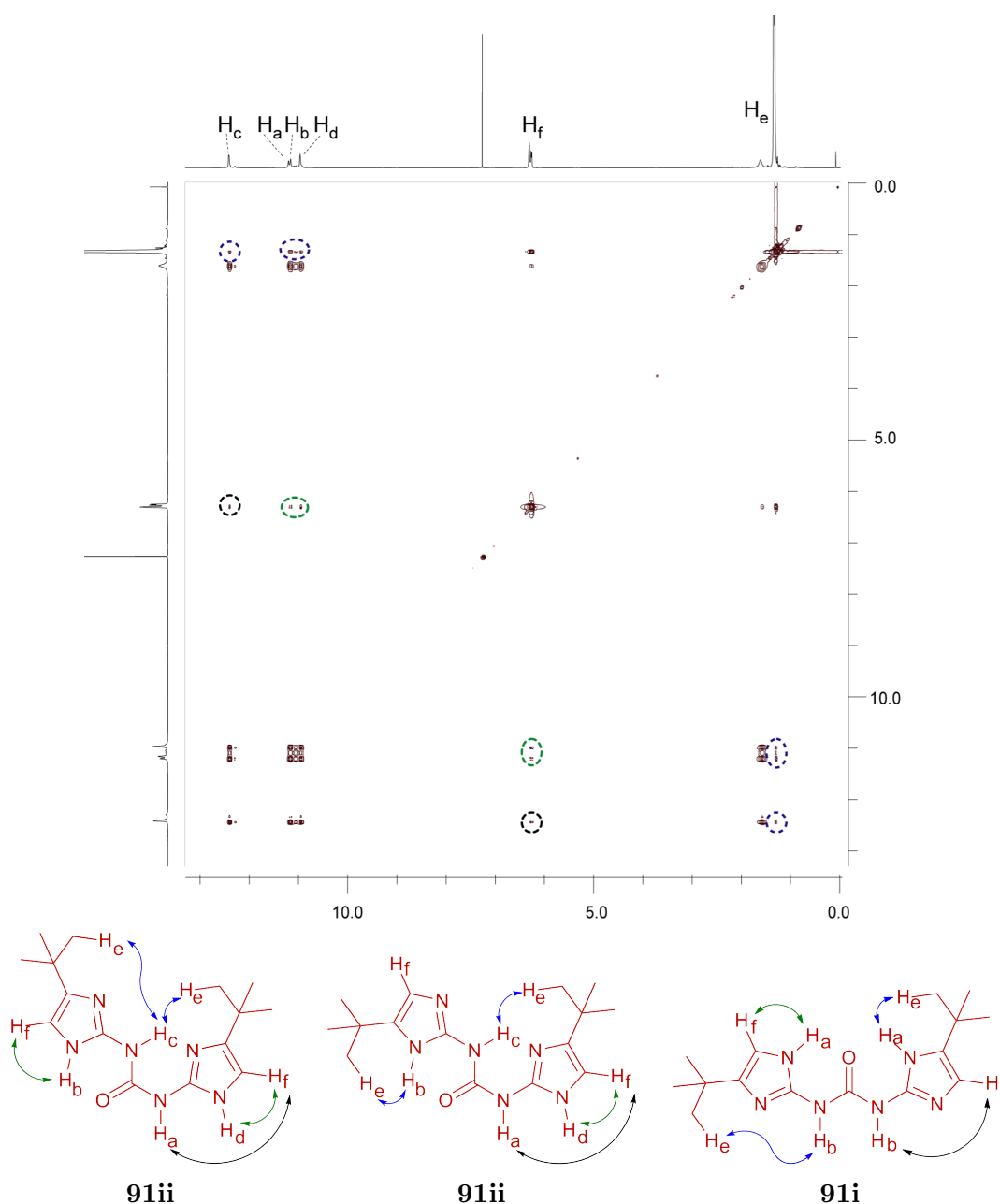


Figure 4.7: ^1H - ^1H NOESY of UDIM **91** shows that each proton is coupled to all of the other protons, but this could arise from either conformation.

Since the 2D NMR experiments could not confirm which conformer UDIM **91** adopted in solution it was not possible to predict whether it was monomeric, or present in larger assemblies. In order to measure the self-dimerisation of UDIM **91**, some dilution and co-solvent experiments were carried out. Firstly, a solution of UDIM **91** in CDCl_3 was diluted from 113 mM to 12 mM (Figure 4.8). No change in shift of any of the proton signals was observed, indicating that either UDIM **91** was monomeric, even at very high concentrations, or that it was bound strongly in assemblies, even when the concentration was reduced. To confirm whether UDIM **91** was present in larger assemblies, DMSO- d_6 was added to a solution of UDIM **91** in CDCl_3 (Figure 4.9). Even at very low concentrations of DMSO- d_6 the NH's began to converge and broaden, which indicated that either exchange between conformer/tautomer states occurred more rapidly with an increasing concentration of DMSO- d_6 , or that the equilibrium distribution of conformers had changed. At 4% DMSO the proton at 6 ppm had become sharp, indicating that exchange between the conformer/tautomer states was increasing. Although some of the proton peaks shifted as the concentration of DMSO- d_6 increased, this was not necessarily due to a complex induced shift, but because the composition of the solvent was different and protons can have different chemical shifts in different solvents. A solution of UDIM **91** in 5% DMSO- d_6 - CDCl_3 was then diluted from 100 mM to 0.5 mM (Figure 4.10). Again, there was no change in the chemical shift of any of the protons, which indicated

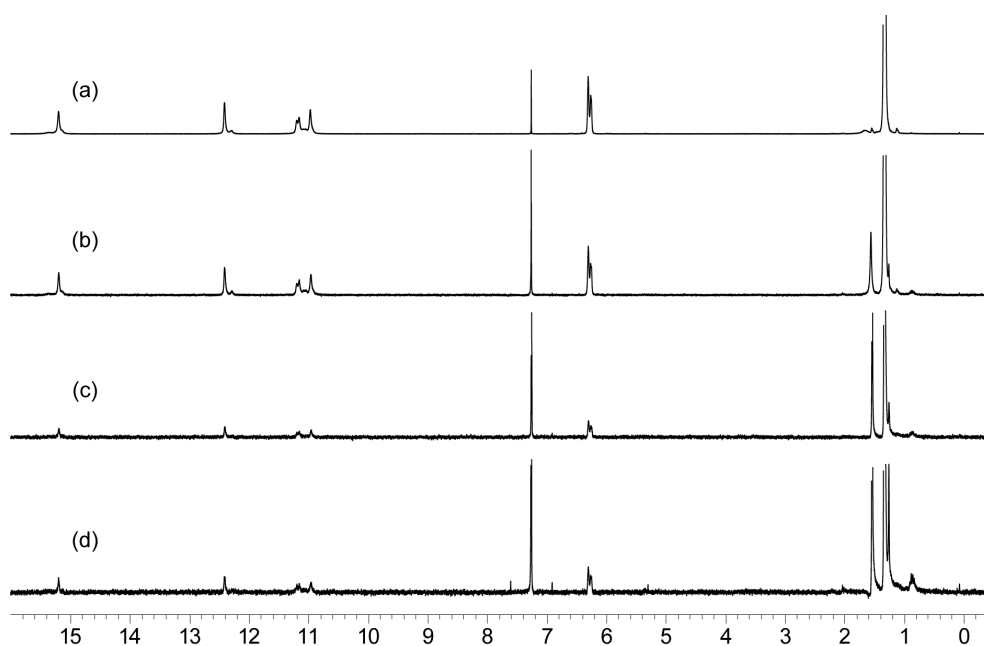


Figure 4.8: A solution of UDIM **91** was diluted (300 MHz, CDCl_3). (a) 113 mM, (b) 46 mM, (c) 19 mM, (d) 12 mM.

that UDIM **91** was monomeric because DMSO- d_6 is a very strong hydrogen bond acceptor and normally interferes with the assembly of hydrogen-bonded complexes as their concentration is decreased. This indicated that UDIM **91** was monomeric in CDCl_3 solutions, but it was not possible to distinguish what the most prevalent conformer was.

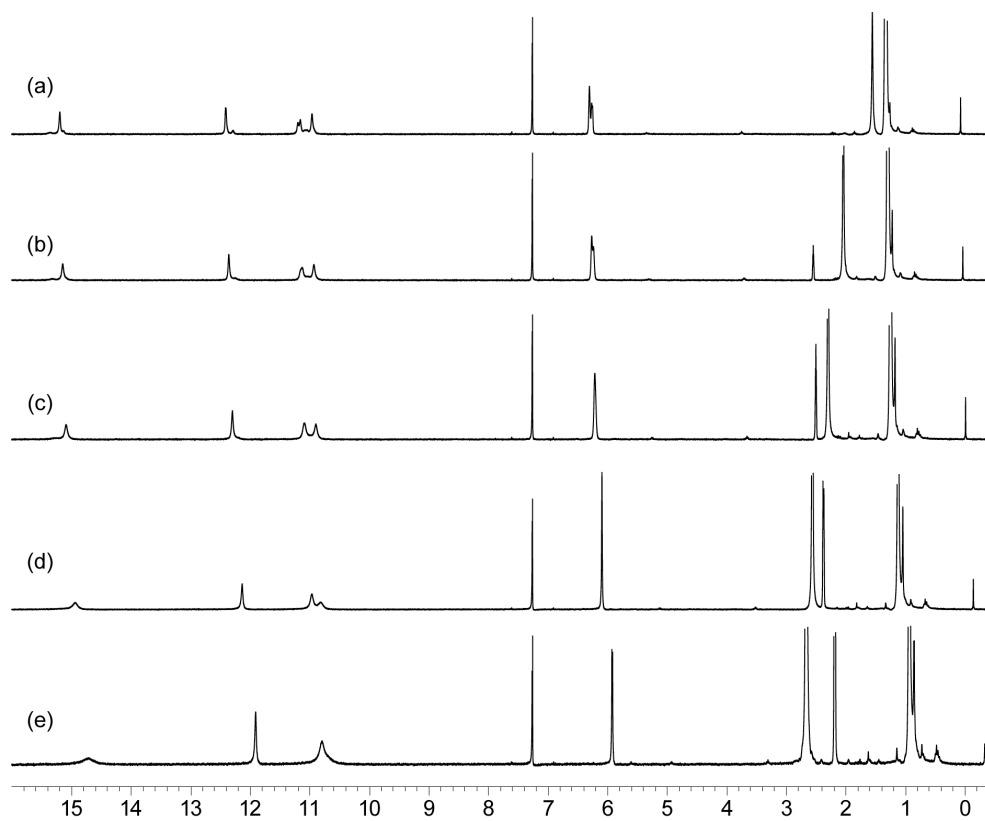


Figure 4.9: DMSO was added to a solution of UDIM **91** (10 mM, 300 MHz, CDCl_3). (a) 0.00%, (b) 1.64%, (c) 4.00%, (d) 10.45%, (e) 20.00%.

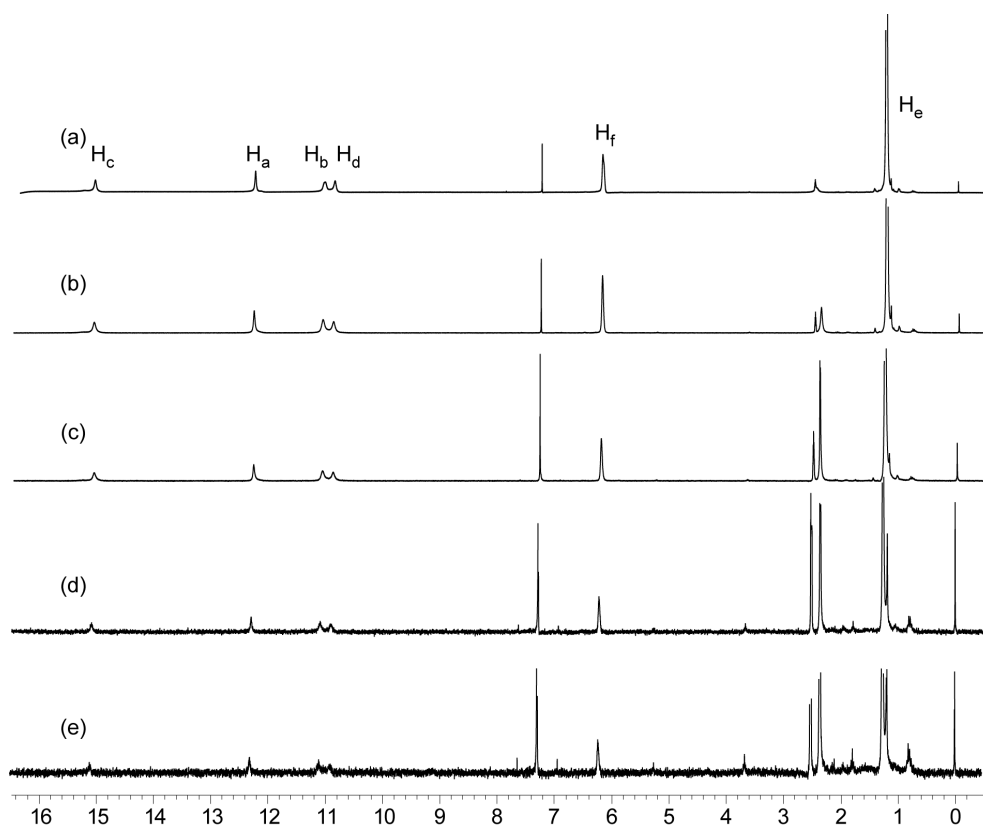
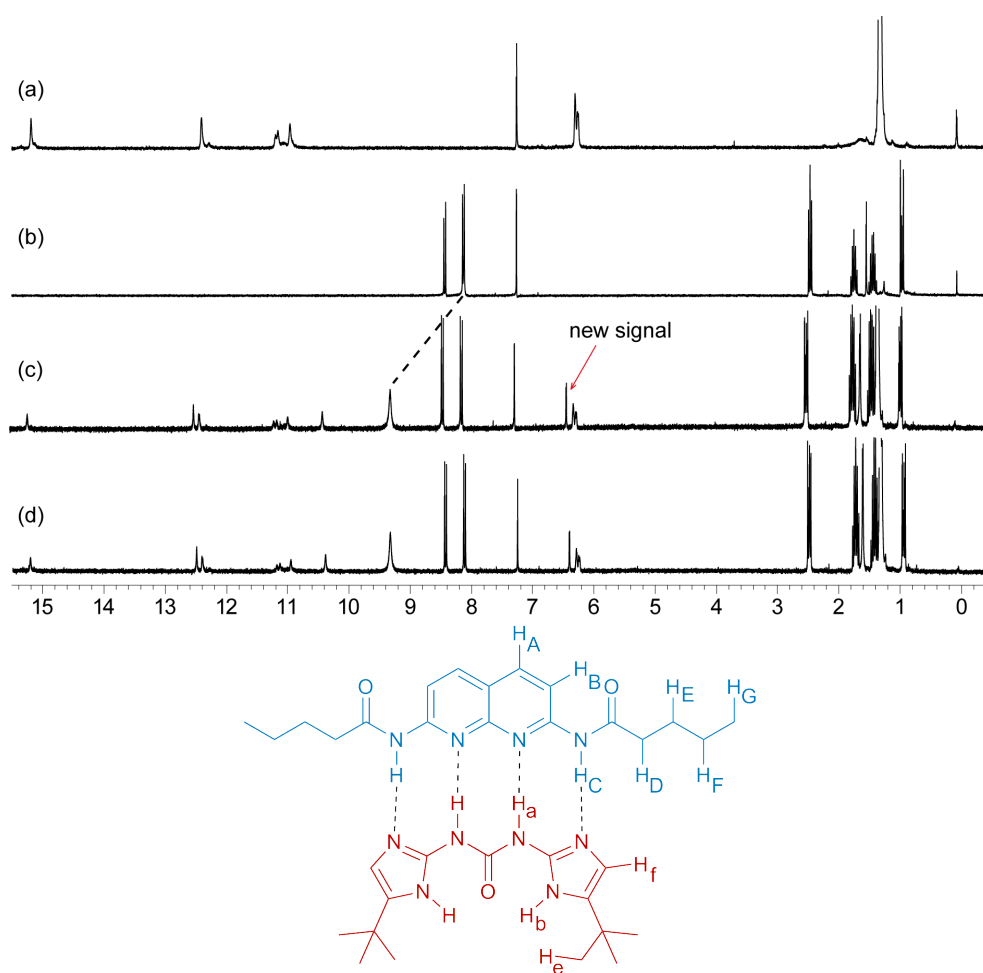


Figure 4.10: A solution of UDIM **91** in 5% DMSO: CDCl_3 was diluted (a) 96.9 mM, (b) 48.5 mM, (c) 9.7 mM, (d) 1.0 mM, (e) 0.5 mM.

4.3 Heterodimerisation Studies of UDIM · DAN

Since it had been shown that UDIM **91** was monomeric in solution, its interactions with DAN **93** could be studied. Comparison of the ^1H NMR spectra of DAN **93** and UDIM **91** with the 1:1 mixture indicated that there was some complex formation (Figure 4.11). The NH of DAN **93** shifts from ~ 8 ppm to ~ 9 ppm and a new signal appears just downfield of the UDIM **91** ArCH proton. This new signal is most likely from the unfolded conformation of UDIM **91i** interacting with DAN **93**. Strongly bound quadruple hydrogen-bond motifs should be nearly 100% bound at 10 mM, so the fact that it appeared that only about a third of UDIM **91** was in the complex at 10 mM gave an early indication that this motif did not behave as expected. Confirmation that the new



93 · 91

Figure 4.11: The NH_C of DAN **93** shifts downfield by ~ 1 ppm (10 mM, 300 MHz, CDCl_3). (a) UDIM **91**, (b) DAN **93**, (c) UDIM · DAN **91** · **93**, (d) UDIM · DAN **91** · **93** after 3 days. No extra complexation was observed. New signal integrates to $\frac{1}{3}$ and the original ArCH signal integrates to $\frac{2}{3}$.

signal at ~ 6 ppm was attributed to bound UDIM **91** was found from ^1H - ^1H NOESY and ROESY experiments. The ^1H - ^1H NOESY (Figure 4.12) showed that all NH's correlated with all other NH's but, as was the case in UDIM **91**, differentiating between exchange and through-space correlations was not possible. For this reason a ^1H - ^1H ROESY (Figure 4.13) was carried out which showed that only the new UDIM **91** signal interacted with DAN **93**. This indicated that it was only the unfolded conformer that formed complexes because the new signal was a singlet, which should only be obtained from a symmetrical molecule.

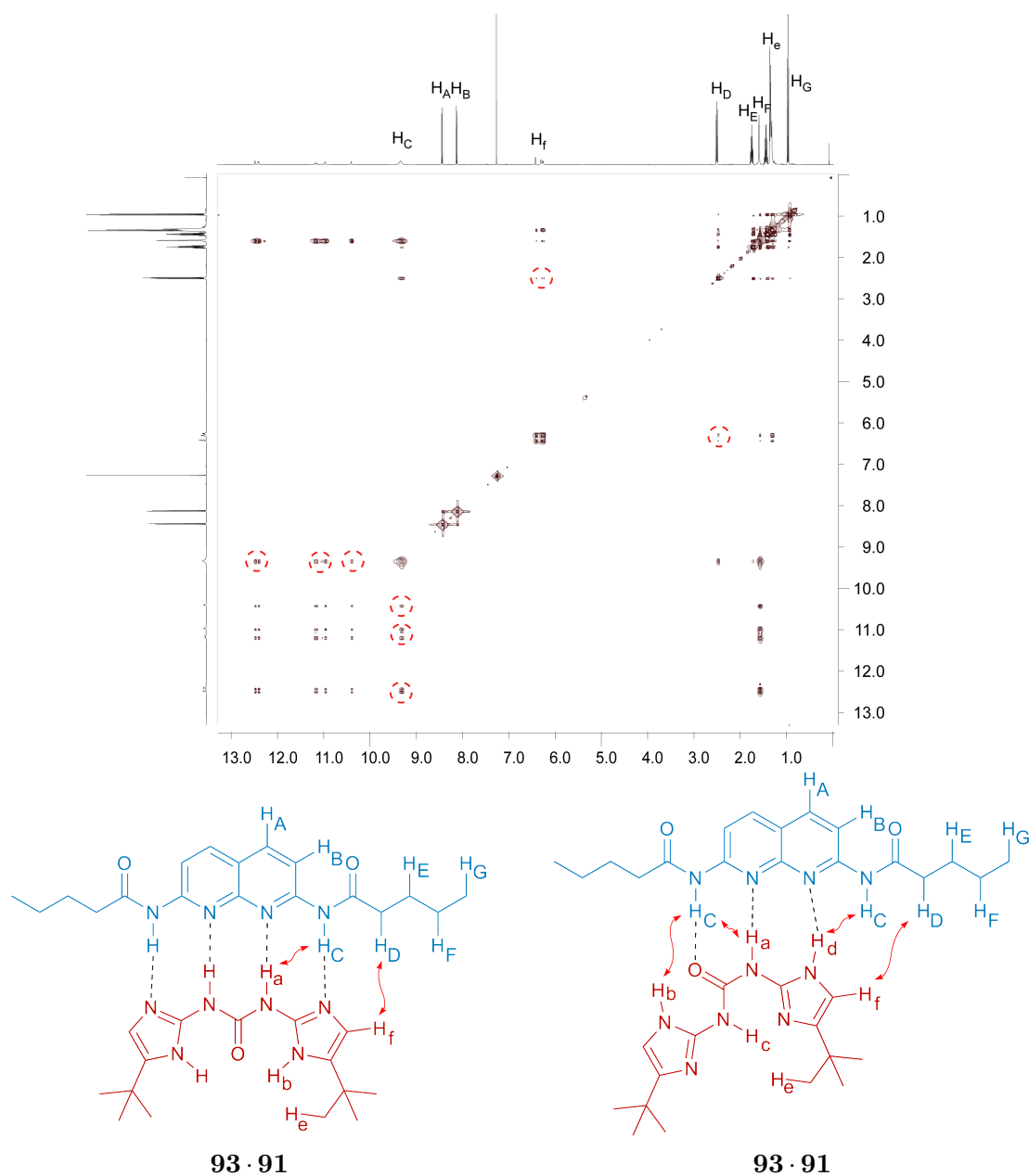
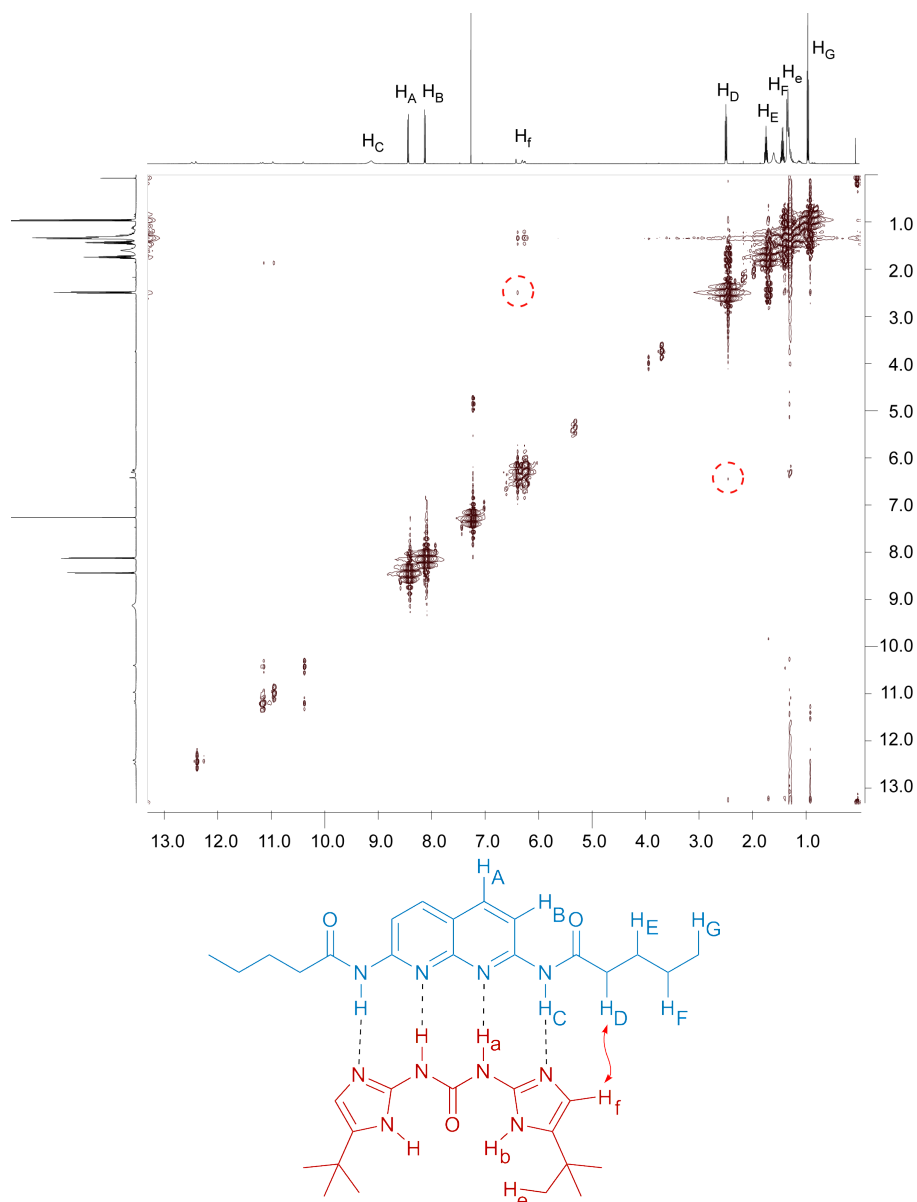


Figure 4.12: ^1H - ^1H NOESY of the DAN · UDIM **93 · 91** complex (10 mM, 500 MHz, CDCl_3).



93 · 91

Figure 4.13: ^1H - ^1H ROESY of the DAN · UDIM **93 · 91** complex (10 mM, 500 MHz, CDCl_3).

Since the DAN · UDIM **93 · 91** complex had been shown to exist in solution a ^1H NMR titration was carried out in order to quantify the strength of interaction. Aliquots of UDIM **91**, containing a small amount of DAN **93**, were added to a solution of DAN **93**. After each addition of UDIM **91** a ^1H NMR spectrum was taken and the change in shift of the NH of DAN **93** was plotted against the concentration of UDIM **91**. Because neither DAN **93** or UDIM **91** strongly self-associate, a simple 1:1 association

model in HypNMR¹⁶⁵ was used to fit the data, giving a K_a of $825 \pm 16 \text{ M}^{-1}$. This association constant was much lower than expected for quadruple hydrogen-bonded complexes, so DOSY NMR experiments were carried out in order to verify the ^1H NMR titration results. Diffusion coefficients were measured on 3.6 mM solutions at 20 °C in CDCl_3 . The diffusion coefficients for DAN **93** and UDIM **91** alone were 10.613 and $9.332 \text{ m}^2 \text{ s}^{-1} \times 10^{-10}$ respectively. The diffusion coefficients of the 1:1 mixture were different to those of the individual samples (10.057 and $8.671 \text{ m}^2 \text{ s}^{-1} \times 10^{-10}$ respectively), showing that larger assemblies were present in the mixture compared to the individual samples. In strongly bound species it is normally expected that the diffusion coefficients are coincident, so obtaining two sets of diffusion coefficients from the mixture indicated that a weakly bound complex was present. In order to calculate the association constant, the mole fraction (χ) must first be obtained. Calculation of χ can be carried out using Equation 4.1 and this was used to calculate the association constant using Equation 4.2. An association constant of 834 M^{-1} was calculated, which is in agreement with the ^1H NMR titrations ($825 \pm 16 \text{ M}^{-1}$).

$$\chi_{\text{DAN} \cdot \text{UDIM}} = \frac{D_{\text{obs}} - D_{\text{UDIM}}}{D_{\text{DAN}} - D_{\text{UDIM}}} \quad (4.1)$$

$$\chi_{\text{DAN} \cdot \text{UDIM}} = \frac{10.057 - 9.332}{10.613 - 9.332}$$

$$\chi_{\text{DAN} \cdot \text{UDIM}} = 0.566$$

$$K_a = \frac{\chi_{\text{DAN} \cdot \text{UDIM}}[\text{DAN}]}{((1 - \chi_{\text{DAN} \cdot \text{UDIM}})[\text{DAN}])^2} \quad (4.2)$$

$$K_a = \frac{0.566 \times 3.6 \times 10^{-3} \text{ M}^{-1}}{(0.434 \times 3.6 \times 10^{-3} \text{ M}^{-1})^2}$$

$$K_a = 834 \text{ M}^{-1}$$

4.4 Comparison of the Binding Properties of a Triple Hydrogen-Bond Array

It has been suggested that sometimes triply hydrogen-bonding motifs can form stronger interactions than quadruply hydrogen-bonding motifs with the same complementary partner.²⁰³ Because UDIM **91** was an extended version of UIM **17**, which interacts through an ADD triple hydrogen-bond array, it was hypothesised that UIM **17** might form stronger interactions with DAN **93** than UDIM **91** did. A 1:1 ^1H NMR spectra of UIM·DAN **17**·**93** confirmed that they interact because of the complexation induced shifts that were observed (Figure 4.14). Indeed, the NH of DAN **93** shifted further downfield when UIM **17** was present than when UDIM **91** was. ^1H - ^1H NOESY (Figure 4.15) confirmed that the molecules were close in space and formed a complex in solution.

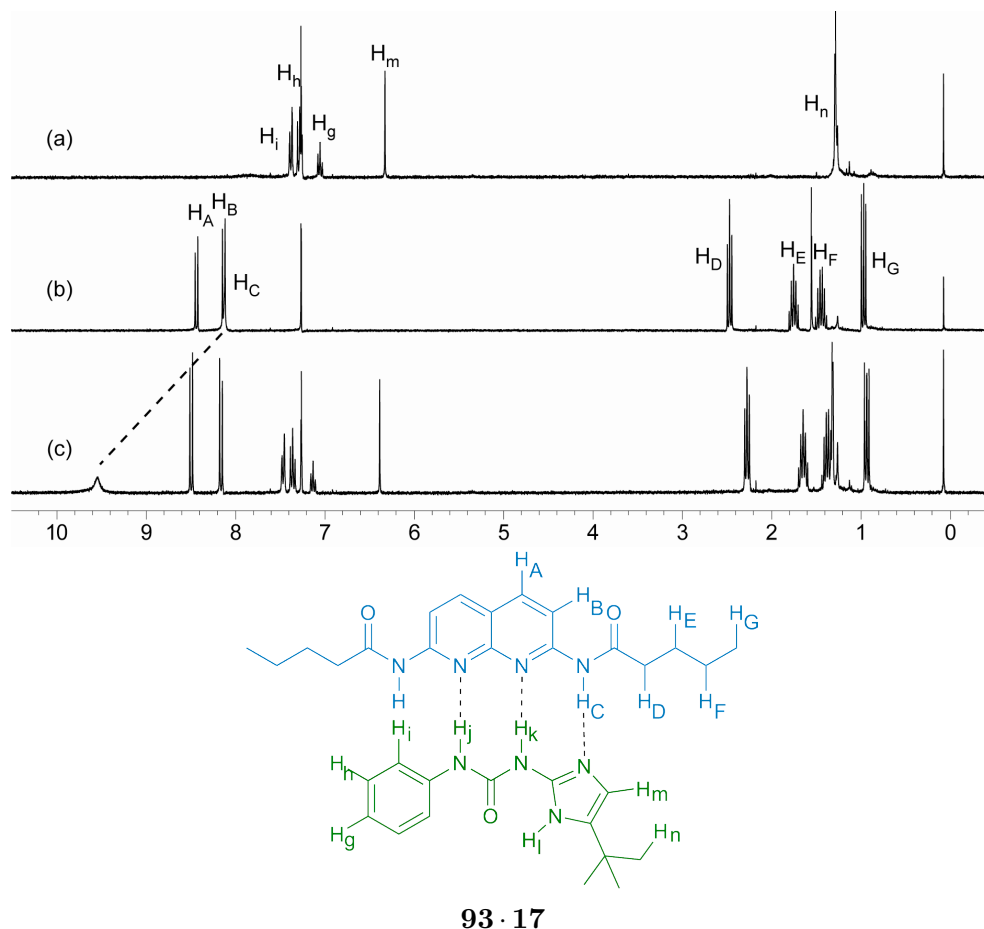
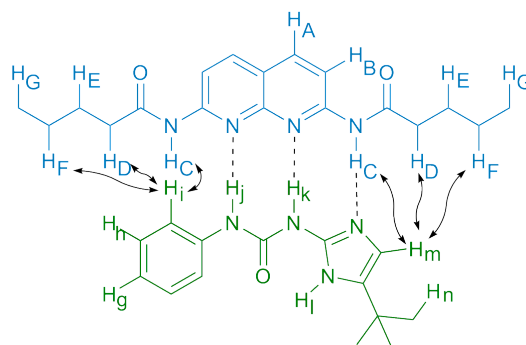


Figure 4.14: The NH_C of DAN **93** shifts downfield by ~ 1.5 ppm (10 mM, 300 MHz, CDCl_3). (a) UIM **17**, (b) DAN **93**, (c) UIM·DAN **17**·**93**.

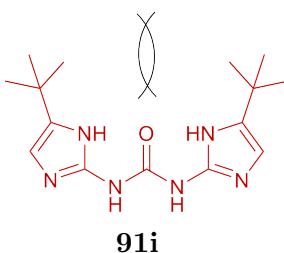


93 · 17

Figure 4.15: Correlations observed between DAN **93** and UIM **17** in the ^1H - ^1H NOESY of a 1:1 mixture (10 mM, 500 MHz, CDCl_3).

^1H NMR titrations were then performed, again following the shift of the NH of DAN **93**. It was found that the K_a of the UIM · DAN **17** · **93** complex was $2140 \pm 40 \text{ M}^{-1}$, which is significantly stronger than the DAN · UDIM **93** · **91** complex ($825 \pm 16 \text{ M}^{-1}$). In order to rationalise the differences in binding affinity observed between the triple and quadruple hydrogen-bond arrays, molecular modelling studies were carried out.

It could be possible that the tautomer required for UDIM **91** to interact with DAN **93** cannot be easily adopted due to steric hindrance (Figure 4.16). To investigate this the energies of all possible conformer/tautomer combinations were calculated (Figure 4.17).^a It was found that there was only $\sim 9 \text{ kJ mol}^{-1}$ energy difference between the lowest and highest energy conformer/tautomers, so it should be possible for all of them to be presented in solution. The conformer/tautomer that presents the desired hydrogen-bonding array with the ^tBu groups pointing away from the interaction face is 5.4 kJ mol^{-1} higher in energy than the lowest energy tautomer/conformer. This should not be a large enough energy difference to prevent the desired conformer being present in solution, especially when the complementary binding partner is present. However, solvation was



91i

Figure 4.16: Steric hindrance between the two ^tBu groups could prevent the desired hydrogen-bond array from being adopted.

^aThe B3LYP/C, 6-31G; N/O, 6-31G*; H, 6-31++G** basis set was used for all the calculations shown in this chapter.

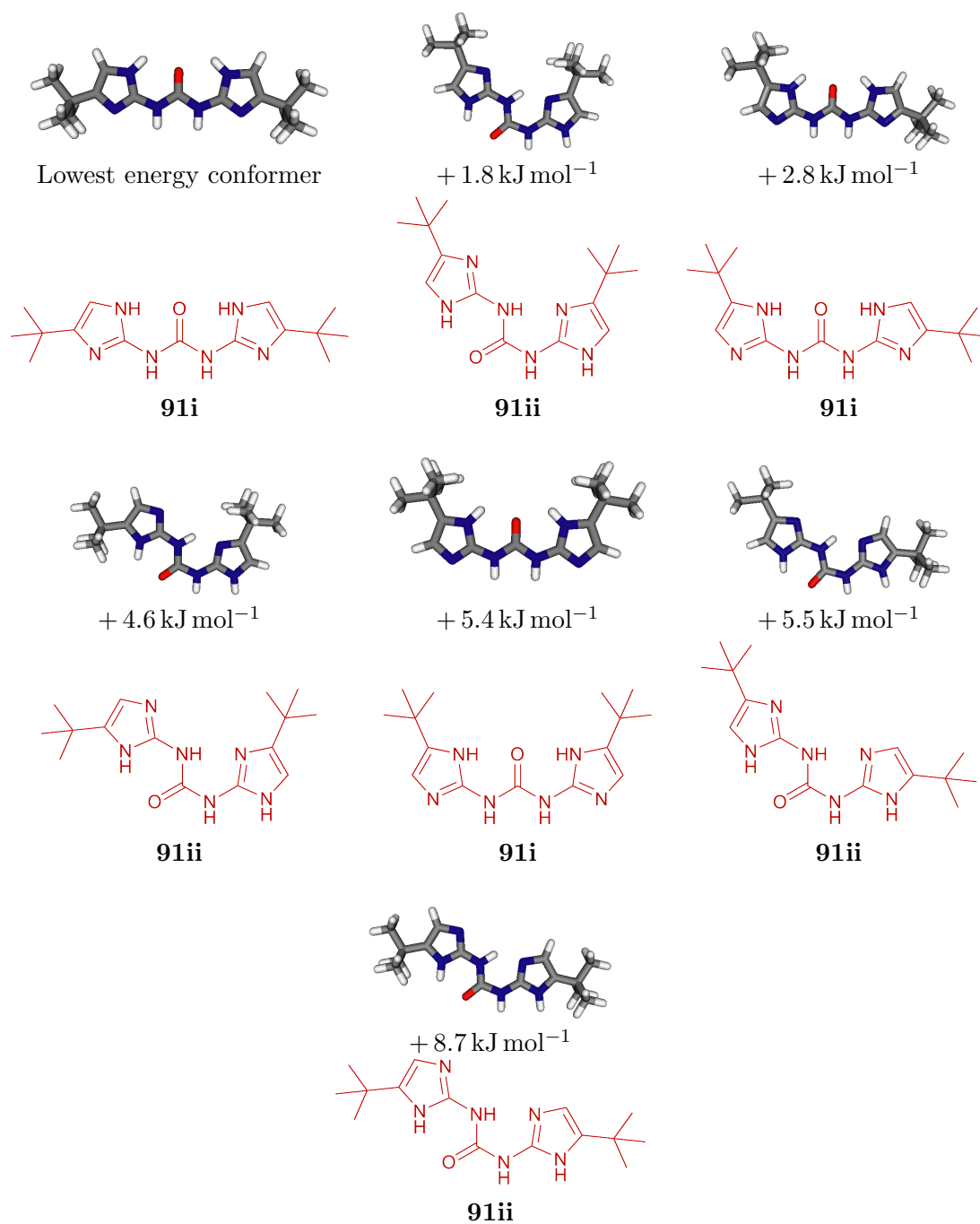


Figure 4.17: The energies of the lowest and highest energy conformer/tautomer combinations and the energy of the desired conformer/tautomer. All other combinations have energies inbetween the lowest and highest energy conformations. Values show how much higher in energy each conformer/tautomer is than the lowest energy conformer.

not taken into account in these gas phase calculations and in addition the energy of the transition states were unknown. It could take a lot of energy to change transition and desolvate the new hydrogen-bonding atoms.

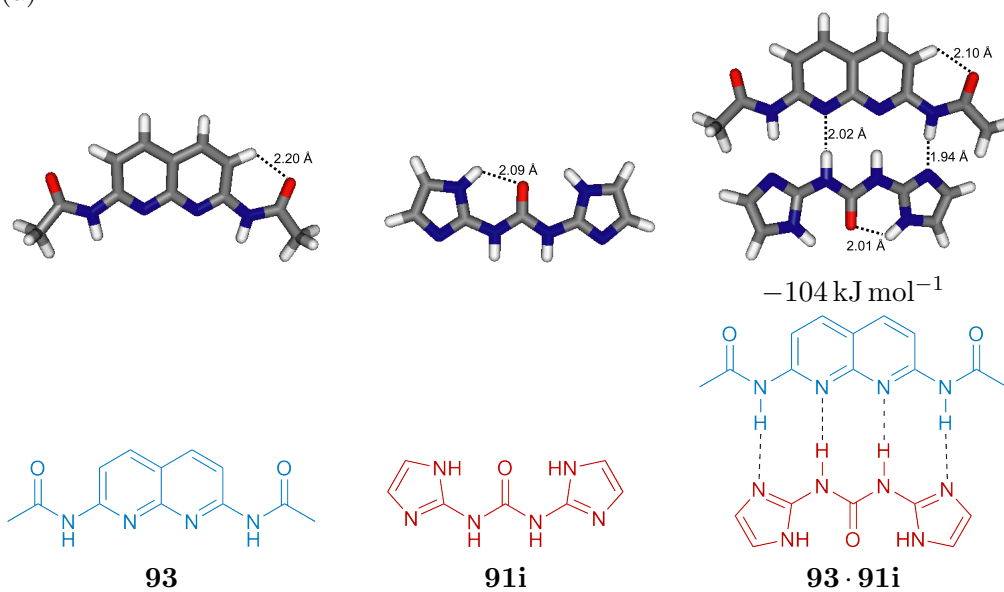
Since the desired conformer of UDIM **91i** can probably exist in solution there must be other factors that prevent it from interacting with DAN **93**. The binding affinities of DAN **93** were calculated with both conformers of UDIM **91** (Figure 4.18).^b It was found that the conformer that interacts with an ADDA array (**91i**) had a binding affinity of -104 kJ mol^{-1} and the folded conformer (**91ii**), which only interacts through three hydrogen bonds, had a lower binding affinity of -76 kJ mol^{-1} . Although the triply hydrogen-bonded complex had a lower binding affinity, it was found that the molecules did not have to alter their geometry so much in order to assemble, so there was less entropic loss. For the quadruple UDIM·DAN **91i·93** complex the ArCH···O=C distance has to become shorter (2.20–2.10 Å) to form interactions. The intramolecular hydrogen bonds in UDIM **91i** also have to become shorter (2.09–2.01 Å) to allow the intermolecular hydrogen-bonding atoms to align. For the triply hydrogen-bonded UDIM·DAN **91ii·93** complex only one of the ArCH···O=C has to become shorter (2.20–2.11 Å and 2.20–2.18 Å) and the intramolecular hydrogen bonds both remain to about the same length (2.05–2.06 Å and 1.93–1.91 Å).

The binding affinities of the two possible conformers of UIM **17**, which can both interact with DAN **93** using three hydrogen bonds, were also calculated (Figure 4.19). It was found that they both had higher calculated binding affinities to DAN **93** (-78 and -79 kJ mol^{-1}) than the folded UDIM **91ii** (-76 kJ mol^{-1}), even though they all interact with the same hydrogen-bonding array. This higher binding affinity was attributed to the fact that UDIM **91ii** had an extra intramolecular hydrogen bond which removed electron density from the carbonyl oxygen, rendering it a weaker hydrogen bond acceptor. It was again noted that, in order for the complex to form, in both cases neither of the intramolecular hydrogen bonds in UIM **17** had to change length (2.07–2.01 Å and 1.91–1.89 Å) as much as they did in UDIM **91** and also only one of the ArCH···O=C had to change length (2.20–2.12 Å and 2.20–2.11 Å).

It appears that the subtle geometry mismatch between the unfolded UDIM **91i** and DAN **93** molecules could be the main reason for the low binding affinities that were

^bFor these calculations the ^tBu groups were removed in order to reduce the time required for the calculation to be completed.

(a)



(b)

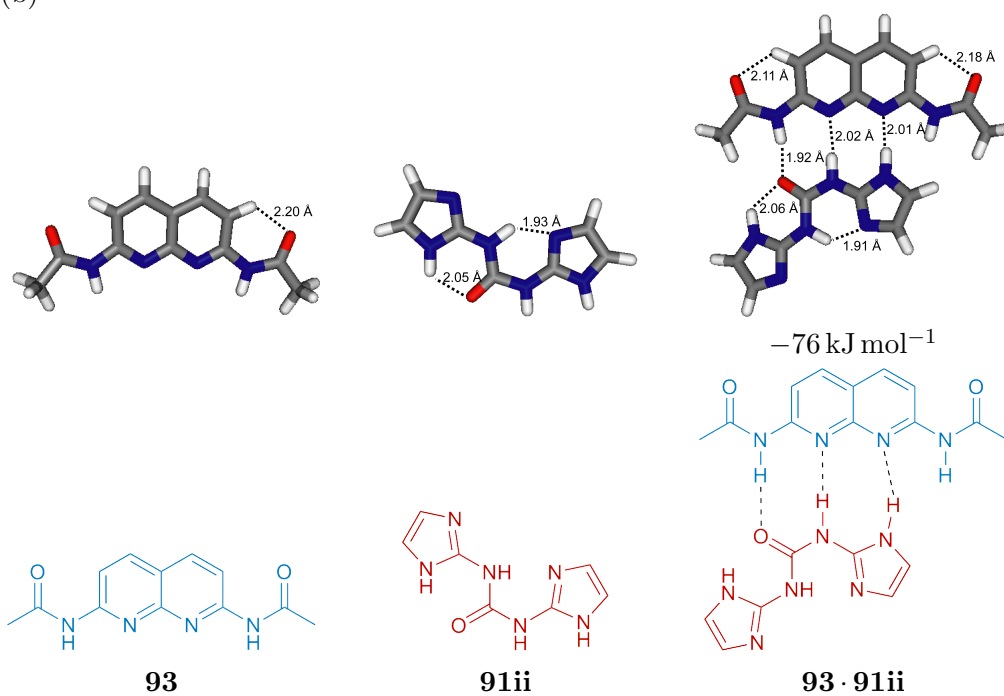


Figure 4.18: The structures and binding energies of the two possible DAN · UDIM $93 \cdot 91$ complexes.

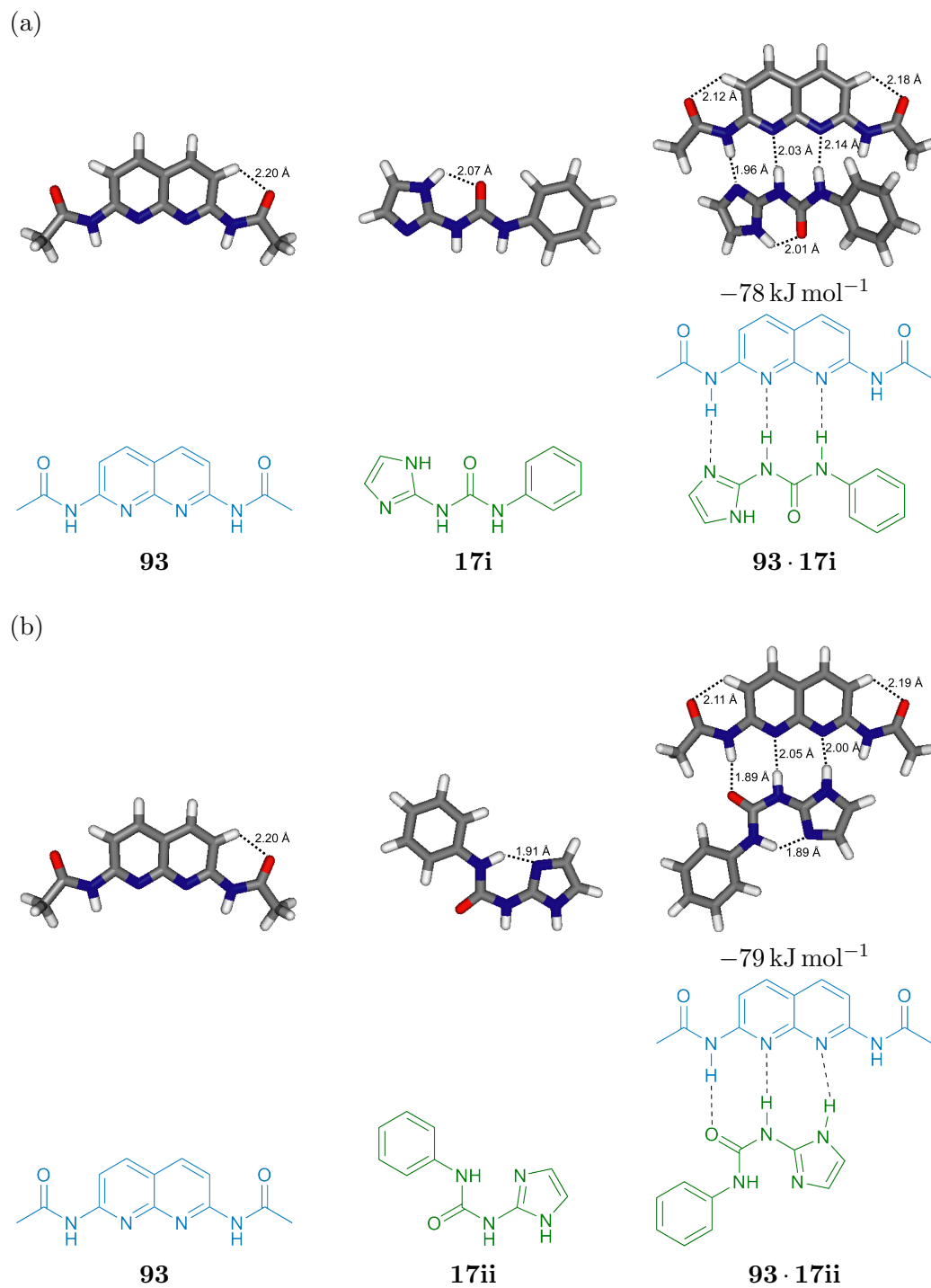


Figure 4.19: The structures and binding energies of the two possible DAN · UIM 93 · 17 complexes.

measured for this complex. Both DAN **93** and UDIM **91** have to change geometry in order to interact using four hydrogen bonds, leading to a loss of entropy, reducing the binding affinity. When the folded conformer is present there is less loss of entropy, which helps to keep the binding affinity as high as possible. Solvation can also play a part in controlling which conformer interacts and it could be that desolvation of the fourth hydrogen bonding atom requires more energy than is gained from it interacting, which would result in a thermodynamically unfavoured process.

4.5 Conclusions

The design and synthesis of a novel quadruple ADDA hydrogen-bonding motif (UDIM **91**) has been described. Lower than expected binding affinities with the complementary partner, DAN **93**, were observed and extensive studies to probe the reasons for this were performed. It was found that the reason for the low association constant was most likely due to a combination of subtle effects. These included the desired conformer of UDIM **91** not being the lowest energy conformer, possible problems with high solvation energies and entropic loss on binding due to differences in shape complementarity. The differences in shape complementarity arose because DAN **93** is not a linear array, but slightly curved, because C-N bonds are shorter than C-C bonds.²⁰³ This study has highlighted that highly preorganised hydrogen-bonding motifs, which present the desired array and do not self-associate, will not necessarily have high association constants with complementary motifs. All of the factors discussed above must be taken into consideration in order to obtain motifs that can interact with high binding affinities.

Chapter 5

Novel Arrays for Self-Sorting Assemblies

Since many of the fundamental features of hydrogen-bonding motifs have now been investigated it was possible to look towards the study of self-sorting systems.¹⁴⁰ Ideally, self-sorting systems are assembled using high affinity and high fidelity interactions. A high fidelity system is defined as one where only the desired complexes, and no undesired complexes, are assembled. The assembly of a self-sorting system using orthogonal hydrogen-bonding motifs could lead to many applications, for example supramolecular polymers and supramolecular switches/devices.²⁰⁴ There are many examples of self-sorting in biological systems;⁶ for example, substrates interacting with proteins, proteins interacting with proteins and the assembly of the DNA double helix. Unfortunately, biological molecules cannot simply be mimicked in synthetic systems because they are too large. Many of the atoms that are present in proteins and DNA are not involved in intramolecular interactions, but are used to preorganise the molecule. Ideal molecules for synthetic systems are small and easy-to-synthesise, allowing complexity to be obtained relatively quickly and cheaply.

Synthetic hydrogen-bonding motifs normally contain heteroaromatic and amide/urea functionalities, which have the potential to interact indiscriminately when there are many molecules present, leading to the formation of undesired complexes. If different types of linear arrays of hydrogen bonds are used, there is the possibility of undesired arrays being formed; for example, in Chapter 4 it was shown that DAN **93**, a quadruple

hydrogen-bond array, can form interactions with UIM **17**, a triply hydrogen-bonded motif. It was therefore proposed that, if hydrogen-bonding motifs could be designed so that they interact through different types of non-linear arrays, the chance of undesired complexes being formed would be reduced allowing high fidelity systems to be obtained.

Three new pairs of motifs were proposed (Figure 5.1) for the synthetic self-sorting system. The first (Figure 5.1a) was between diamidopyridine (DAP) **97** and oxazolidinone **98**, which were proposed to interact through a bifurcated enol array. The second (Figure 5.1b) was between the pyrrole **99** and its complementary piperidinone **100**, which were proposed to interact through a bifurcated hydrogen-bond array. The final array (Figure 5.1c) that was proposed for the synthetic self-sorting system was a 3D array, which was assembled using two 2D motifs **101** and **102**. Because all these complexes were designed to interact through different types of hydrogen-bond array, the formation of undesired complexes should be prevented when they are mixed, leading to the assembly of a high fidelity²⁰⁴ self-sorting system.

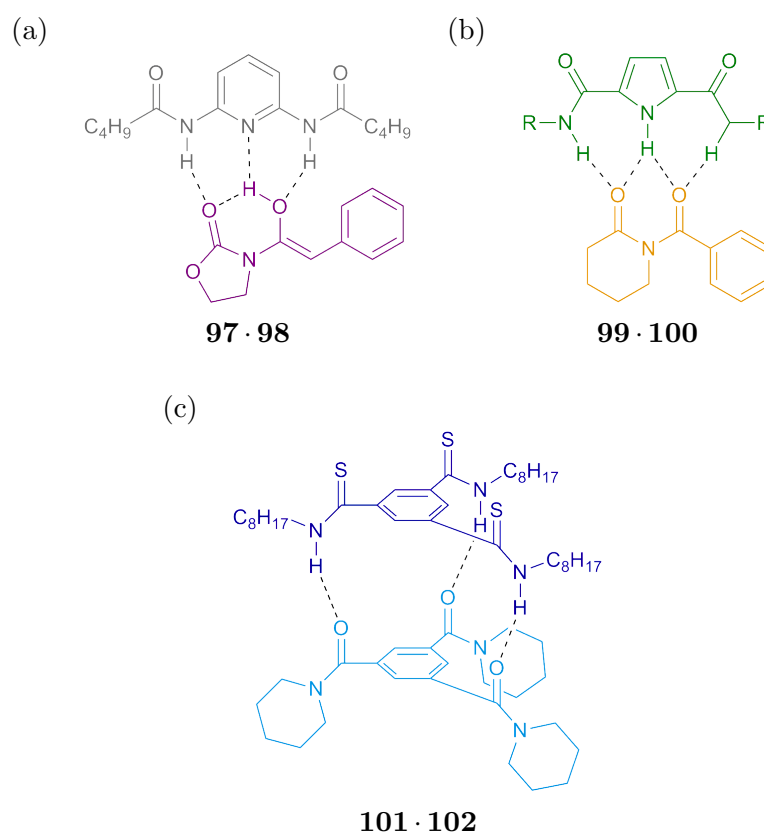


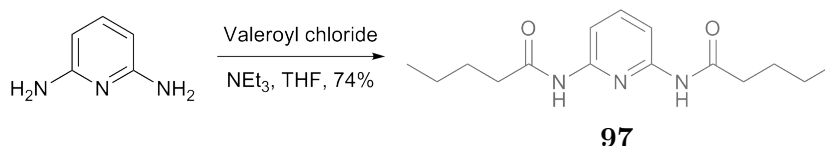
Figure 5.1: Three new motifs that were proposed for the synthetic self-sorting system (a) enol, (b) bifurcated and (c) 3D arrays.

5.1 Enol Array

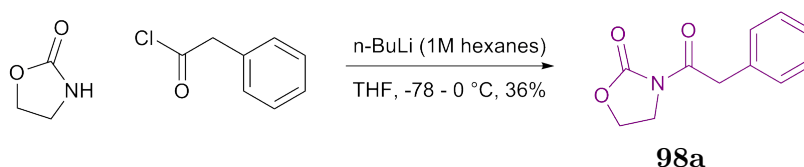
When carbonyls are placed at the one and three positions along an alkyl chain there is the possibility for a six-membered intramolecular hydrogen bond ring to form when an enol is present. This results in an ADA array being presented, which should interact with a complementary DAD array. Diamidopyridine (DAP) **97** has been previously studied by the Wilson group,⁶³ so it was chosen as the DAD array for this study.

5.1.1 Synthesis and Binding Studies

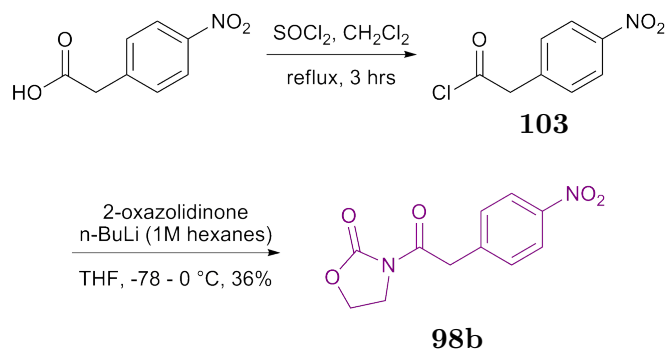
Synthesis of DAP **97** (Scheme 5.1) was carried out in one step, reaction of valeroyl chloride with 2,6-diaminopyridine gave DAP **97**.⁶³ The complementary oxazolidinones were then synthesised. Reaction of the commercially available phenylacetyl chloride with deprotonated 2-oxazolidinone gave the phenyl oxazolidinone **98a** (Scheme 5.2).²⁰⁵ It was reasoned that an electron withdrawing group on the phenyl ring could make the benzyl protons more acidic, enabling the enol to form more easily, so a second oxazolidinone was synthesised (Scheme 5.3) with an electron withdrawing nitro group. Formation of the acid chloride **103** was carried out in the presence of thionyl chloride²⁰⁶ before addition of 2-oxazolidinone gave the nitrophenyl oxazolidinone **98b**.²⁰⁵



Scheme 5.1: The synthesis of DAP **97**.



Scheme 5.2: The synthesis of the enol array **98a**.



Scheme 5.3: Synthesis of the enol array **98b** with an electron withdrawing group appended to it.

A crystal structure of the enol array with a nitro group **98b** was obtained by the slow evaporation of a solution of **98b** in CH_2Cl_2 - Et_2O . For interactions to be possible between **98b** and complementary motifs the carbonyl groups must be pointing in the same direction however, in the crystal structure of **98b** the carbonyl groups are orientated away from each other. This is probably because there are electrostatic interactions between the two oxygen atoms, preventing them from being on the same face when the enol form is not present.

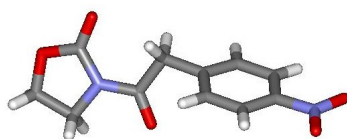


Figure 5.2: A crystal structure of the enol array **98b** was obtained.

Interactions between DAP **97** and the enol arrays **98a** and **98b** were tested by comparing the individual ^1H NMR spectra to that of a 1:1 mixture (Figure 5.3 and Figure 5.4). No complexation induced shifts were observed and the benzyl protons integrated to two throughout, indicating no conversion to enol. This showed that there were no interactions between **97** and **98a** or **98b** and the enol was never present.

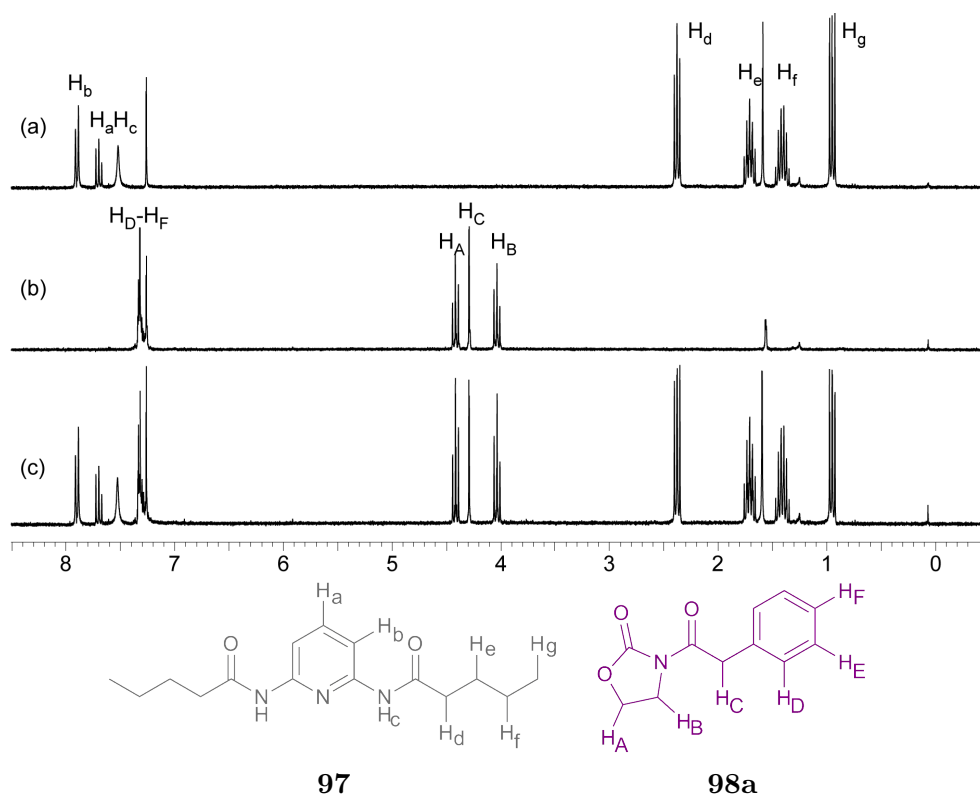


Figure 5.3: ^1H NMR spectra of (a) DAP **97**, (b) **98a** and (c) a 1:1 mixture of DAP **97** and **98a** (10 mM, 300 MHz, CDCl_3).

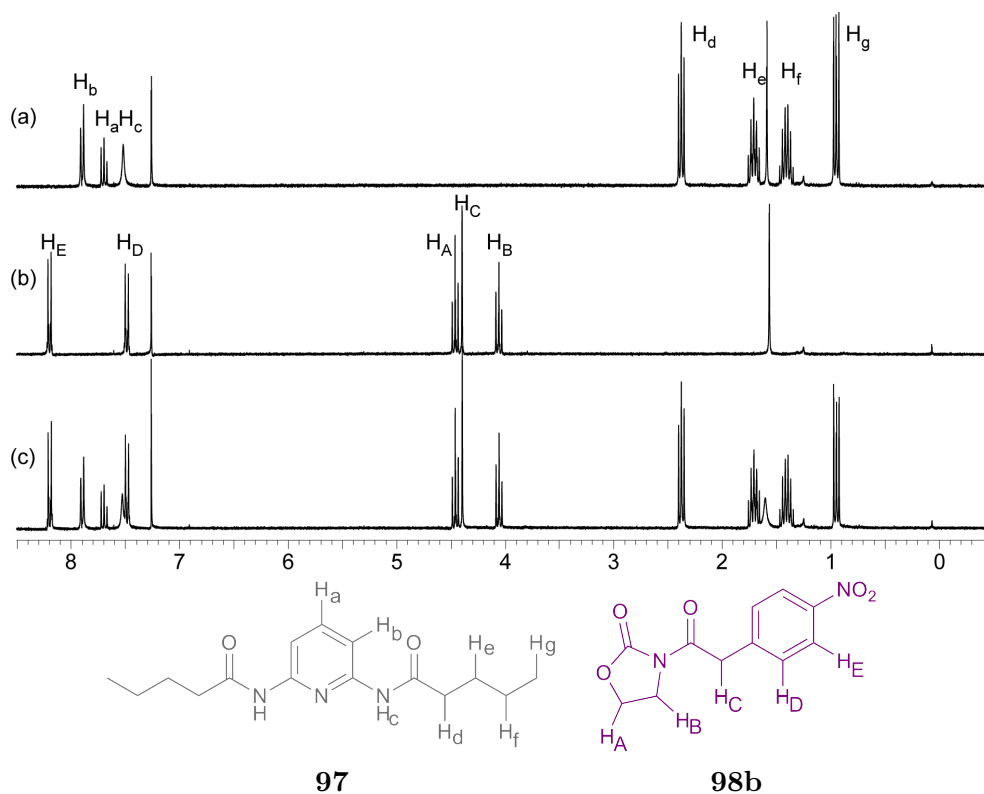


Figure 5.4: ^1H NMR spectra of (a) DAP **97**, (b) **98b** and (c) a 1:1 mixture of DAP **97** and **98b** (10 mM, 300 MHz, CDCl_3).

Since neither **98a** or **98b** could interact with DAP **97**, ethylacetoacetate **104** was tested because it has been shown that it can exist in the enol form.²⁰⁷ Unfortunately, no interactions between **104** and DAP **97** were observed (Figure 5.5), indicating that it is not possible to form complexes between DAP **97** and enols. This could be because protons can only form one hydrogen bond at a time and intramolecular hydrogen bonds are normally formed in favour of intermolecular hydrogen bonds.⁵⁹ So, even if the enol is formed, it is unlikely that the desired heterocomplex will be formed.

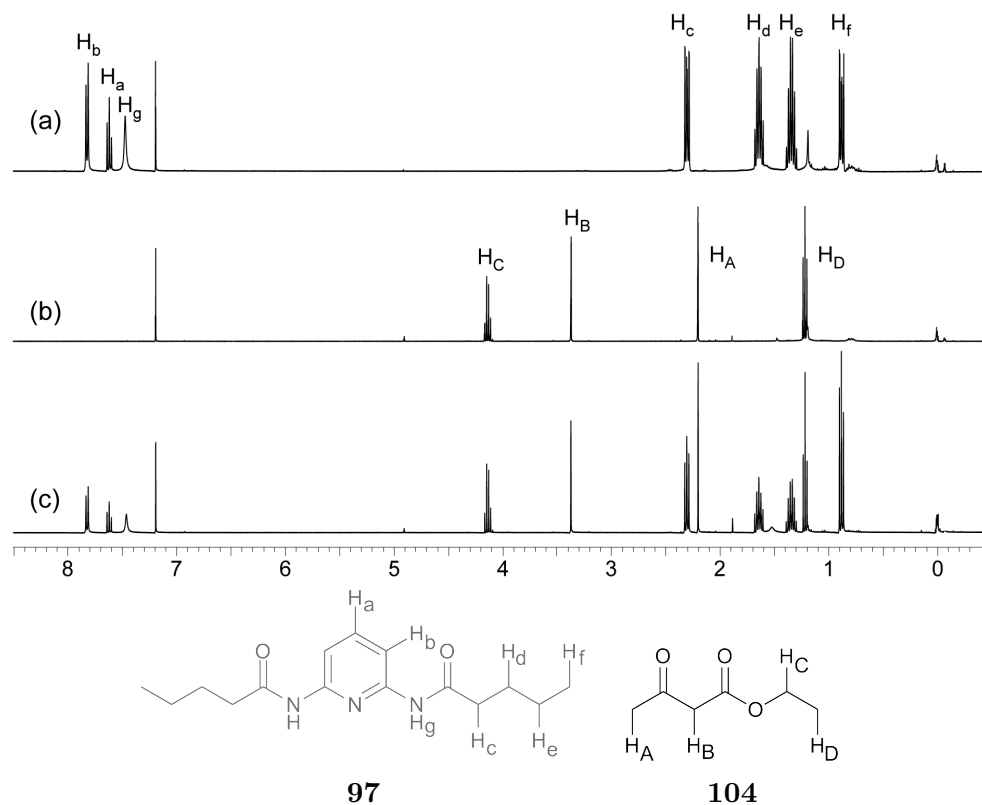


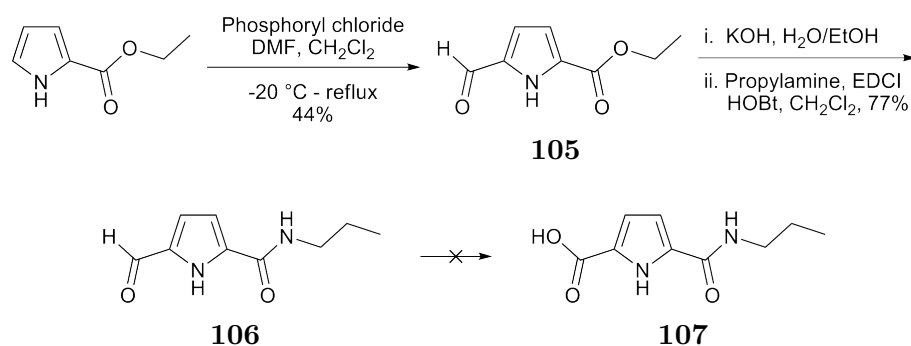
Figure 5.5: ¹H NMR spectra of (a) **97**, (b) **104** and (c) a 1:1 mixture of **97** and **104** (400 MHz, CDCl₃, 10 mM)

5.2 Bifurcated Array

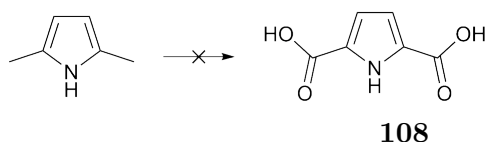
Since carbonyl oxygens have two lone pairs of electrons they can act as the acceptor of two hydrogen bonds and this is known as a bifurcated array. Using bifurcated hydrogen bonds can allow small numbers of hydrogen bonding atoms to form many interactions.²² This can increase the binding affinity whilst keeping the molecule small, which normally makes the synthesis easier. Two carbonyls can form up to four hydrogen bonds and it was proposed that the 2,5-diamidopyrrole **99** would present the complementary DDD array (Figure 5.1b).

5.2.1 Synthesis and Binding Studies

Many different routes were attempted towards the synthesis of the 2,5-diamidopyrrole **99**. Firstly, to enable an unsymmetrical molecule to be obtained (Scheme 5.4), functionality was added to ethyl pyrrole-2-carboxylate *via* a Vilsmeier reaction to give the aldehyde **105**.²⁰⁸ Deprotection of the ester in the presence of potassium hydroxide, followed by amide coupling using EDCI and HOBt gave the amide pyrrole **106**. Oxidation of the aldehyde in the presence of potassium permanganate did not give the desired acid pyrrole **107** and this route was abandoned in favour of the synthesis of symmetrical pyrroles. The second route that was attempted was the oxidation of 2,5-dimethyl pyrrole, to obtain the diacid pyrrole **108** (Scheme 5.5) however, this reaction was not

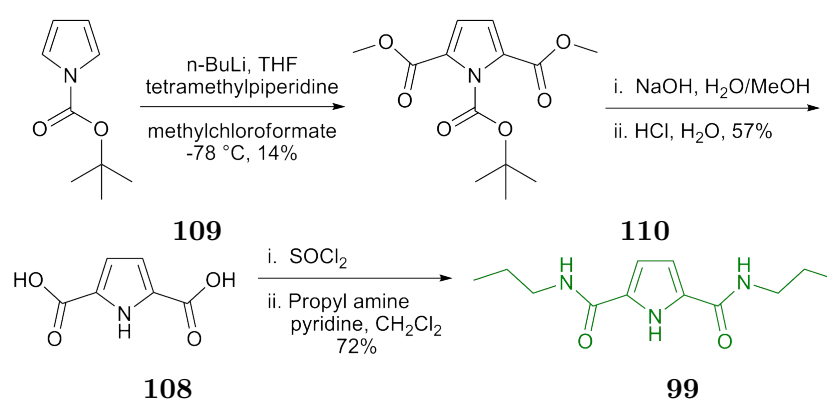


Scheme 5.4: The first route that was attempted for the synthesis of pyrroles.



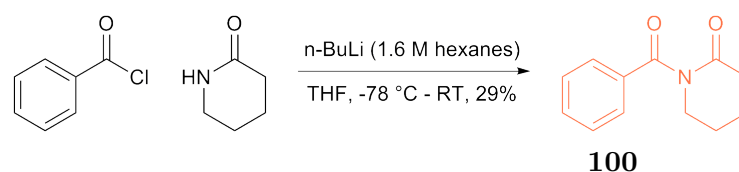
Scheme 5.5: Attempts to synthesise the diacid pyrrole **108** *via* oxidation of the 2,5-dimethyl pyrrole were all unsuccessful.

successful. The reason for this could be because the 3- and 4-positions were not substituted and there were no electron withdrawing groups present.²⁰⁹ The third route that was attempted (Scheme 5.6) involved functionalisation of the boc-pyrrole **109** *via* deprotonation of the 2- and 5-positions using *n*-BuLi before addition of methyl chloroformate gave the 2,5-diester pyrrole **110**.²¹⁰ Hydrolysis in the presence of sodium hydroxide, followed by removal of the protecting group with hydrochloric acid gave the 2,5-diacid pyrrole **108**. Amide coupling was carried out by preparation of the acid chloride with thionyl chloride followed by addition of propyl amine in the presence of pyridine to give the 2,5-diamidopyrrole **99**.

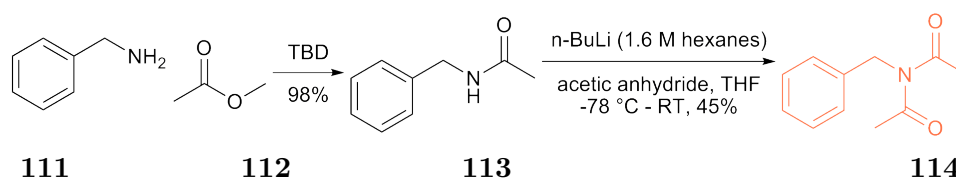


Scheme 5.6: Synthesis of the diamide pyrrole **99**.

Two complementary partners for the 2,5-diamidopyrrole **99** were synthesised. The first was the benzoylpiperidinone **100** which was synthesised (Scheme 5.7) by deprotonation of piperidinone using *n*-BuLi followed by addition of benzoyl chloride. Synthesis of the benzyl diacetamide **114** was first attempted by the reaction of diacetamide and benzyl bromide in the presence of both NaH and K₂CO₃²¹¹ however, no product was isolated. The second route (Scheme 5.8) involved amide coupling between benzyl amine and

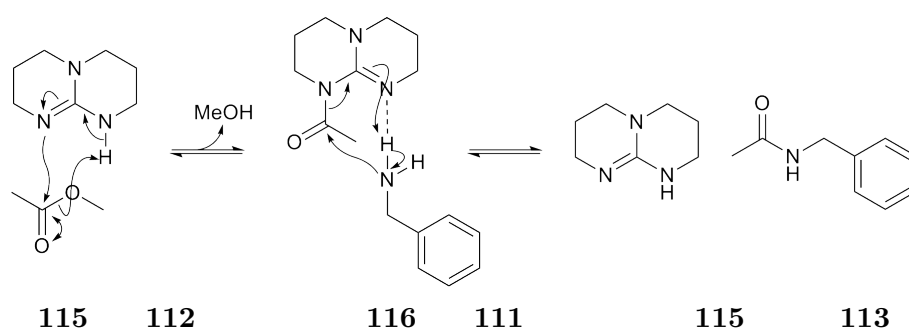


Scheme 5.7: Synthesis of the benzoylpiperidinone **100**.



Scheme 5.8: Synthesis of the acetylbenzylacetamide **114**.

methyl acetate using a catalytic amount of triazabicyclodecene (TBD) **115** in solvent free conditions to obtain the benzyl acetamide **113**.²¹² TBD **115** can catalyse the amide coupling of esters **112** and amines **111** (Scheme 5.9)²¹³ by first reacting with the ester **112** to give the TBD amide **116** and alcohol. The amine **111** can then interact with the TBD-amide **116** through an intermolecular hydrogen bond, which places the nitrogen in the desired position for amide formation to occur. TBD **115** is then regenerated along with formation of the desired amide **113**. Deprotonation of the benzyl acetamide **113** using *n*-BuLi followed by the addition of acetic anhydride gave the benzyl diacetamide **114**.



Scheme 5.9: The proposed reaction mechanism of TBD **115** with the ester **112** and amine **111** to give the amide **113**.

¹H NMR was used to investigate the binding properties of the 2,5-diamidopyrrole **99** to the piperidinone **100** (Figure 5.6) and the diacetamide **114** (Figure 5.7). No interactions were observed and this was probably due to a combination of reasons. Firstly, it is likely that there were unfavourable electrostatic interactions between the two carbonyl oxygens. This would hinder interactions between the pyrrole **99** and piperidinone **100**/diacetamide **114** because it would prevent the piperidinone **100** and the diacetamide **114** from adopting the desired conformation. Secondly, a derivative of the pyrrole **117** has previously been used to interact with just one oxygen atom²¹⁴ (Figure 5.8) and this could indicate that there is not enough space within the array for two oxygens to interact.

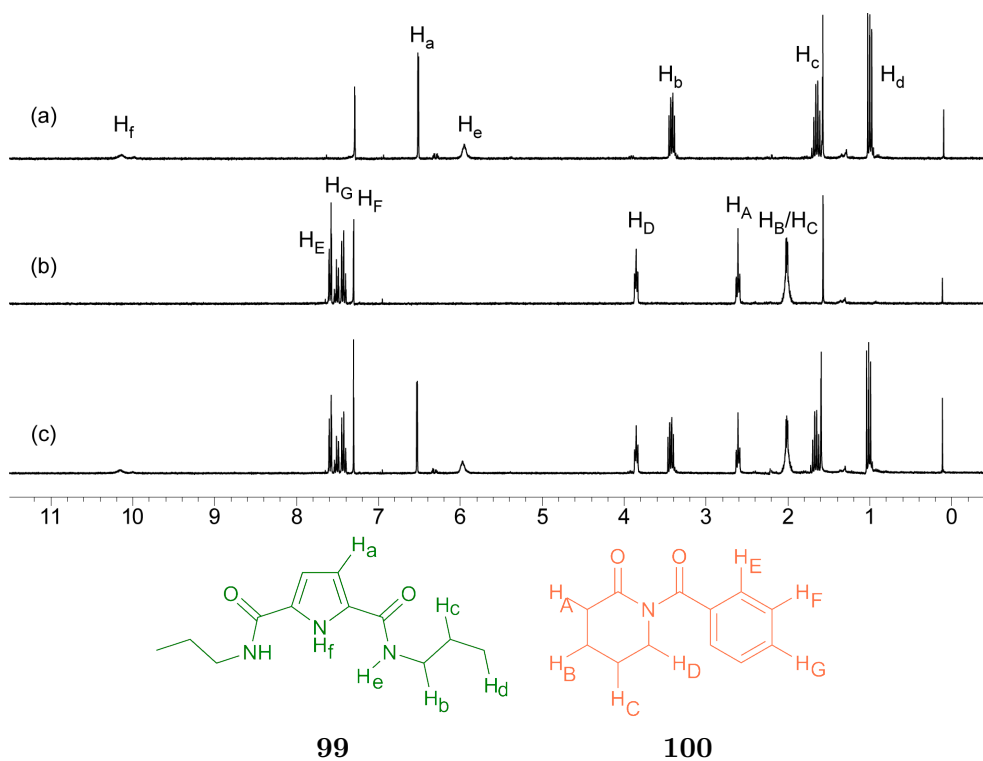


Figure 5.6: ^1H NMR spectra of (a) **99**, (b) **100** and (c) a 1:1 mixture of **99** and **100** (300 MHz, CDCl_3 , 10 mM).

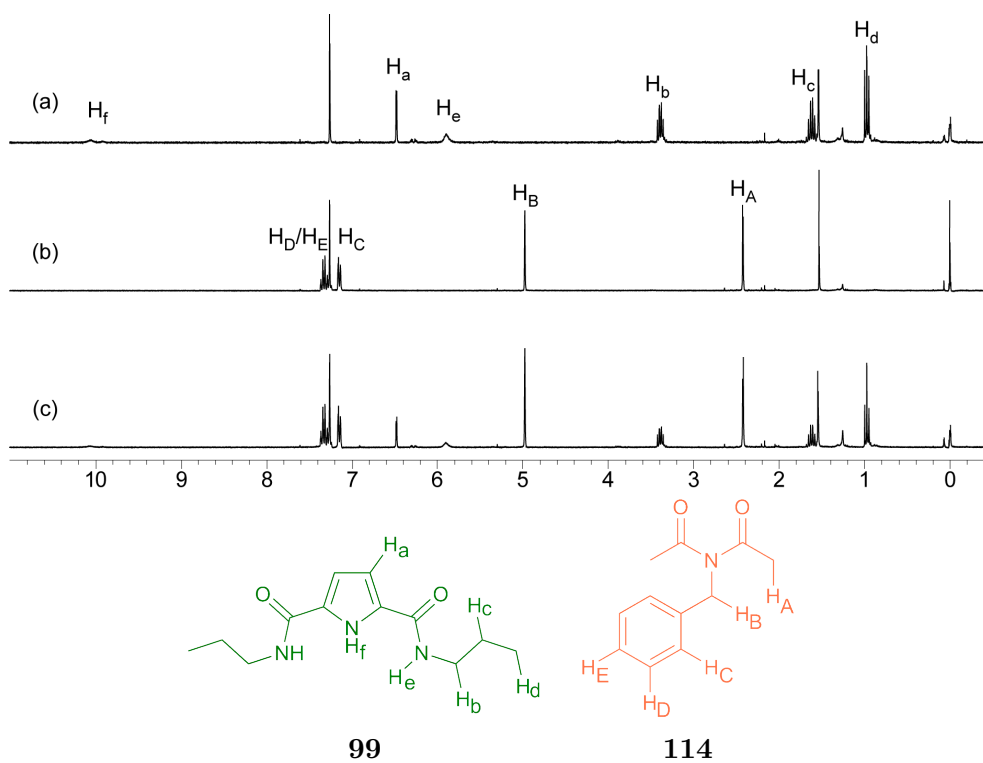


Figure 5.7: ^1H NMR spectra of (a) **99**, (b) **114** and (c) a 1:1 mixture of **99** and **114** (300 MHz, CDCl_3 , 10 mM).

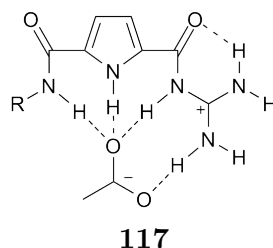


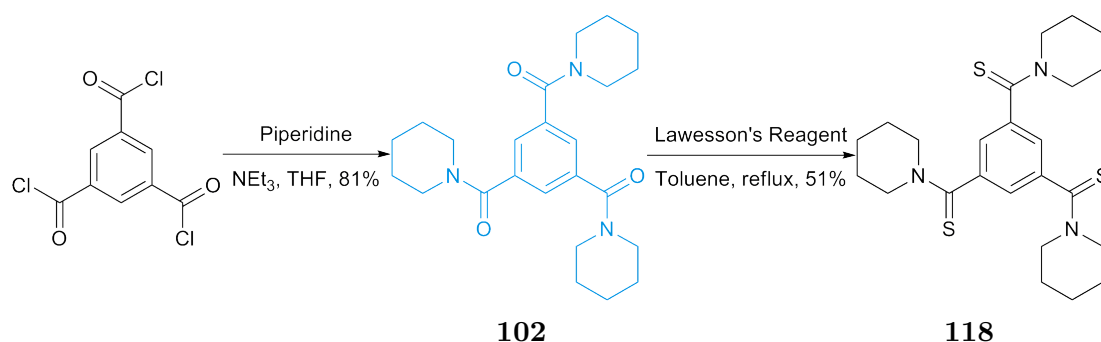
Figure 5.8: The pyrrole **117** that interacts with just one oxygen.

5.3 Three Dimensional Array

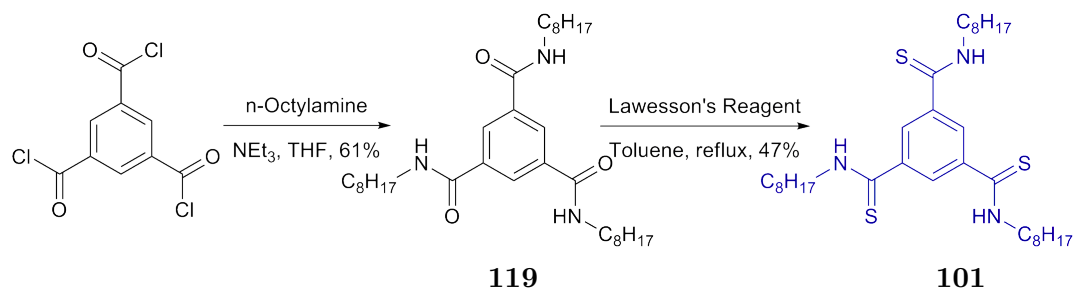
Benzene-1,3,5-triamides have been used to assemble three-dimensional (3-D) stacks;^{102,105,106} the amide functional group allows two-dimensional (2-D) molecules to be self-complementary, leading to the formation of 3-D stacks. By removing either the donor or acceptor from the molecule it was proposed that these molecules would no longer be self-complementary. This would prevent the assembly of 3-D stacks, allowing discrete complexes to be formed, but only when both complementary molecules were present.

5.3.1 Synthesis and Binding Studies

Benzene-1,3,5-tricarbonyltrichloride was reacted with the chosen amine to give the tripiperidine amide **102** (Scheme 5.10) and trioctyl amide **119** (Scheme 5.11).²¹⁵ Because the tripiperidine amide **102** is a tertiary amide it does not contain any protons that can act as hydrogen bond donors, so should not form 3-D stacks.^{107,108} The trioctyl amide **119** however, is self-complementary and the hydrogen bond acceptors must be removed to prevent it from assembling into 3-D stacks. The oxygen atoms were exchanged for sulphur atoms because soft sulphur atoms are not as strong hydrogen bond acceptors as

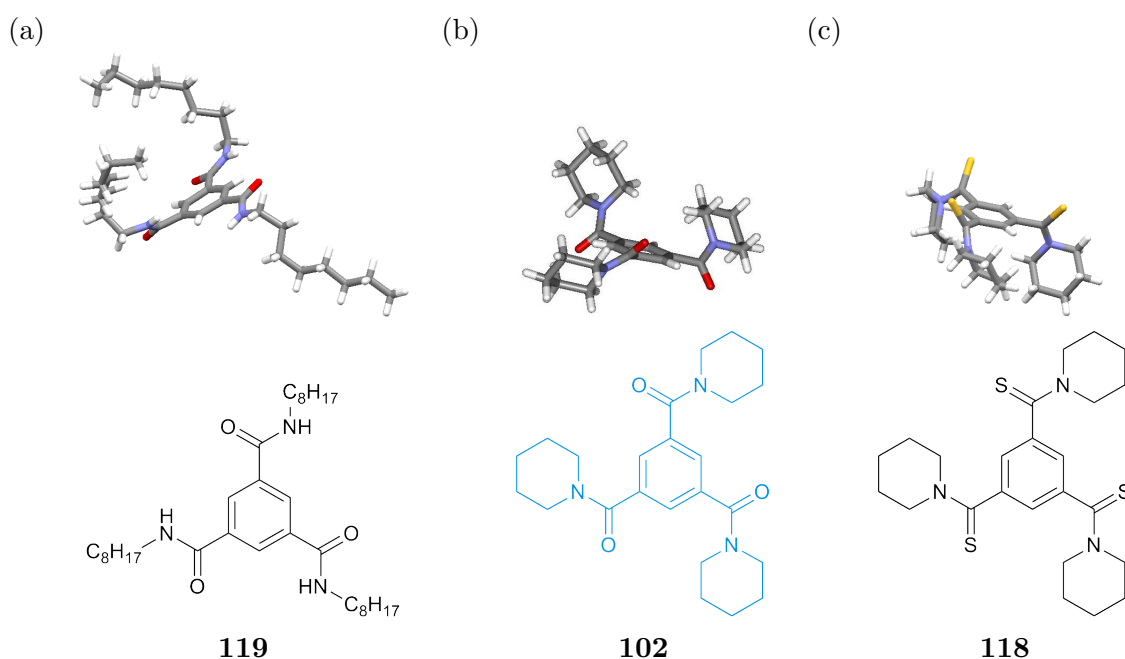


Scheme 5.10: Synthesis of the tripiperidine amide **102** and the trithioamide **118**.

Scheme 5.11: Synthesis of the trioctyl thioamide **101**.

hard oxygen atoms.²¹⁶ Reaction of the trioctyl amide **119** with Lawesson's Reagent²¹⁷ gave the trioctyl thioamide **101**

Crystals of **119** (Figure 5.9a), the precursor of the 1,3,5-trioctyl thioamide **101**, were obtained by the slow evaporation of a DMF–Et₂O solution. The phenyl ring appears to be in conjugation with the amide bonds because they are planar. For intermolecular interactions to form the conjugation would have to be broken and the amide bonds twisted out of the plane of the phenyl ring. The structure of **102** (Figure 5.9b), obtained by the slow evaporation of MeOH, is very different. The large piperidine groups sterically hinder the conjugated conformation, leading to a structure where the carbonyl groups are orientated away from the plane of the phenyl ring. This is ideal for intramolecular hydrogen bonds to form because the carbonyl groups are already in the correct orientation, so there will be less loss of entropy on binding. For heterodimers to be formed

Figure 5.9: Crystal structures of **119**, **102** and **118**.

between **102** and **101** the hydrogen bonding atoms must all be on the same face of the phenyl ring. In the crystal structure of the tripiperidine **102** only two are on the same face, with the third being on the opposite face. This could indicate that there is steric hindrance between the piperidine rings, preventing them from all being on the same face, which would give lower than expected binding affinities between **102** and complementary motifs. The crystal structure of 1,3,5-tripiperidine trithioamide **118** was also obtained (Figure 5.9c) and in this structure the piperidines are all on the same face. This indicates that there is enough space for all of the piperidines to be on the same face of the molecule and presumably **102** can also adopt this conformation.

^1H NMR of the 1:1 mixture of **102** and **101** gives a very broad spectrum (Figure 5.10) which sharpens as the temperature is increased. It was found that at low temperatures a broad spectrum was obtained, which could indicate that weak interactions were occurring and an equilibrium between bound and unbound **102** and **101** had been established. As the temperature was increased the spectrum became more defined, which showed that the molecules were most likely present as monomers. Although it was shown that the **102**·**101** complex was probably forming in solution the association constant could not be obtained using ^1H NMR titrations because the spectra were too broad. Other techniques (e.g. ITC or DOSY) could have been used to obtain the binding affinity however, ^1H NMR was the technique of choice with which to study the self-sorting system, so it was decided that the **102**·**101** heterodimer was not a suitable candidate for the self-sorting system.

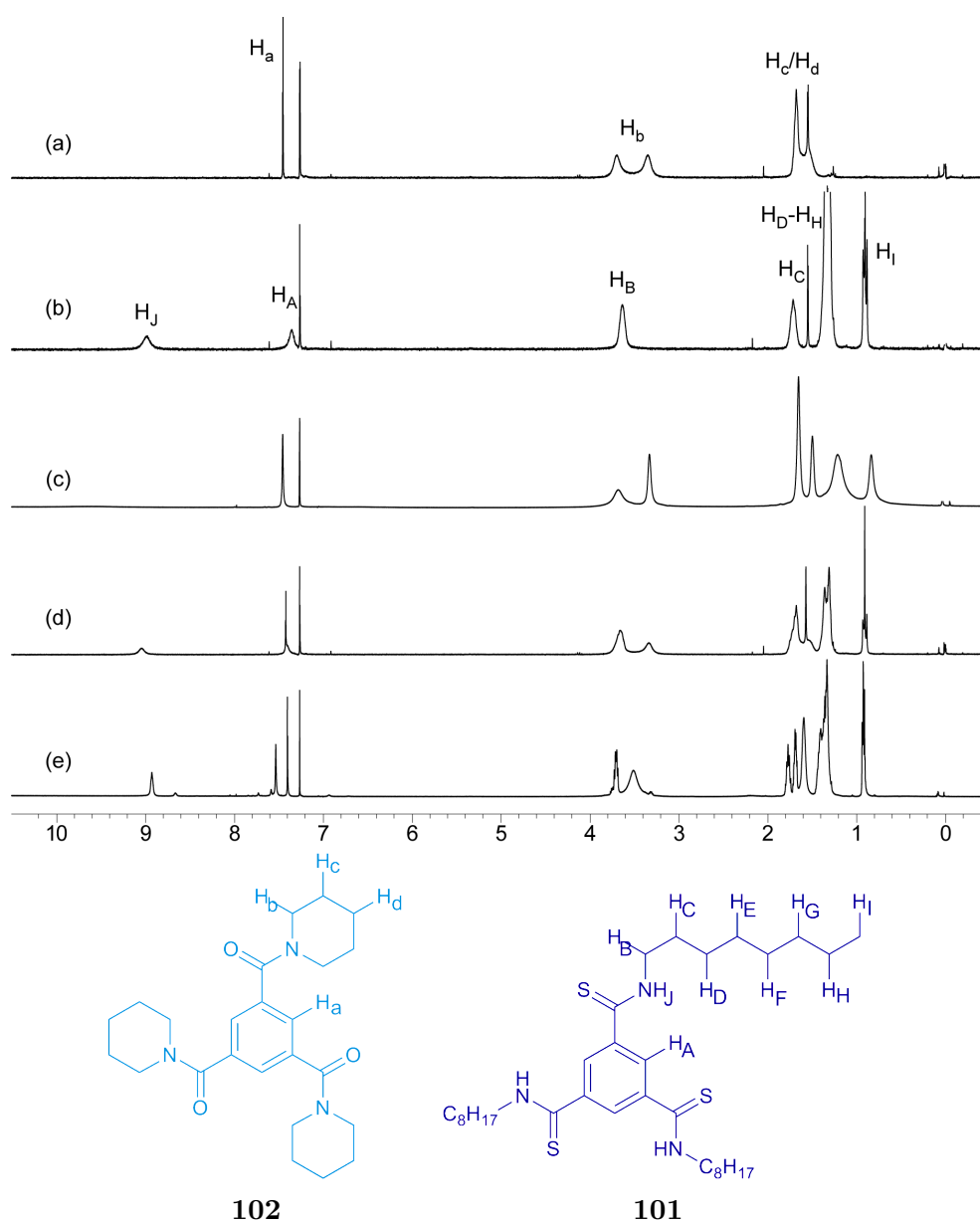


Figure 5.10: ^1H NMR spectra (300 MHz, CDCl_3 , 10 mM) of (a) **102** (293 K), (b) **101** (293 K) and the 1:1 mixtures of **102** and **101** at (c) 223 K, (d) 293 K, and (e) 333 K.

5.4 A Self-Sorting System

Because it was not possible to form complexes using novel hydrogen-bonding arrays a different approach was required to assemble an orthogonal self-sorting system. The DAN·UPy **21**·**19** complex has been used in a self-sorting system along with two hydrazide derivatives **120** and **121**.²¹⁸ The desired outcome of this system was the **21**·**19** and **120**·**121** pairs (Figure 5.11) however, by ¹H NMR it was possible to see that UPy **19** was present as both the homo- and heterodimer (at a ratio of 4:6 respectively). This suggested that the system had a low fidelity, possibly because both desired complexes interact using the same ADDA·ADDA array, so there was no chemical discrimination between the pairs.

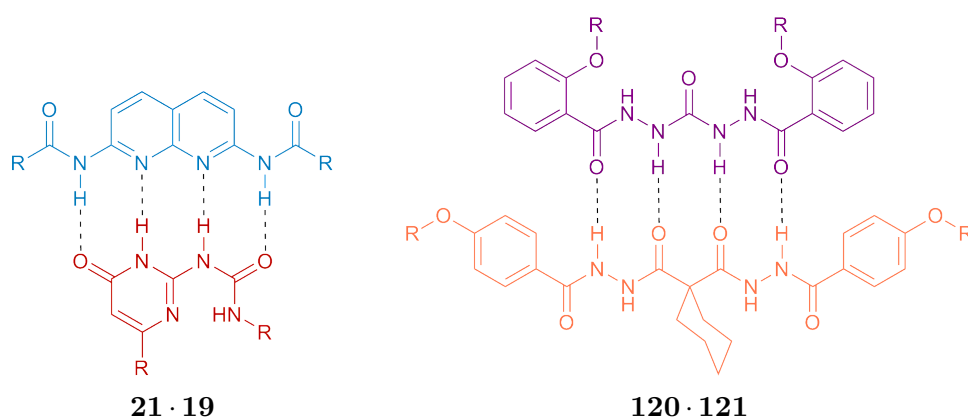


Figure 5.11: The self-sorting system proposed by Zhao *et al.* that includes DAN **21** and UPy **19**. R denotes generic alkyl chain.

By using different types of linear array the binding affinity of the desired complexes should be much higher than the undesired complexes, which should allow a high affinity, high fidelity system to be obtained. The previously studied AIC·UIM **57**·**17** triply hydrogen-bonded array was chosen along with the DAN·UPy **93**·**80** quadruple hydrogen-bond array (Figure 5.12). Since hydrogen bonded systems assemble into the thermodynamic product, using different types of array should give a high fidelity system because the maximum number of hydrogen bonds can only be present when the desired complexes are assembled.

Before this self-sorting system could be assembled the association constants of all possible dimers had to be measured. The self-association of all four molecules was already known; the homodimerisation of DAN **93** was found to be negligible ($K_{\text{dim}} < 5 \text{ M}^{-1}$) by Corbin *et al.*⁶¹ and the homodimerisation of UPy **80** is known to be high ($K_{\text{dim}} = 6 \times 10^7 \text{ M}^{-1}$).¹⁵⁶

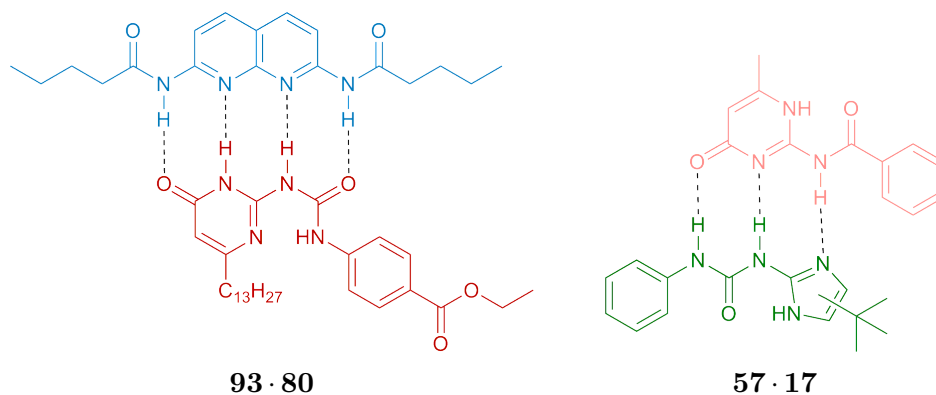


Figure 5.12: The proposed self-sorting system that uses different types of linear hydrogen-bond array.

Homodimerisation constants of UIM **17** and AIC **57** were measured in Section 2.1.3 and were found to be $K_{\text{dim}} = 10.6 \pm 2.2$ and $3.6 \pm 0.3 \text{ M}^{-1}$ respectively. The association constant of the DAN · UPy **93 · 80** heterodimer has been found to be¹⁹¹ $4 \times 10^5 \text{ M}^{-1}$ and the association constant of the AIC · UIM **57 · 17** was measured in Chapter 2 to be $3 \times 10^4 \text{ M}^{-1}$.

It was found in Chapter 4 that undesired interactions are formed between DAN **93** and UIM **17** ($K_a = 2140 \pm 40 \text{ M}^{-1}$). To investigate whether there were interactions between DAN **93** and AIC **57** the ^1H NMR spectra of the individual molecules was compared to that of the 1:1 mixture (Figure 5.13). There were no complexation induced shifts, indicating that there were no intermolecular interactions between the molecules.

A 1:1 mixture of UPy **80** and UIM **17** (Figure 5.14) showed that interactions were formed between these two molecules. ^1H NMR titrations were carried out in order to obtain the association constant of the UPy · UIM **80 · 17** complex. By fixing the homodimerisation of UPy **80** at $6 \times 10^7 \text{ M}^{-1}$ in HypNMR¹⁶⁵ an association constant of $68000 \pm 9600 \text{ M}^{-1}$ was obtained for the UPy · UIM **80 · 17** complex, which is much stronger than the DAN · UIM **93 · 17** complex. This was thought to be due to secondary electrostatic interactions,⁴⁵ DAN · UIM **93 · 17** has an additional destructive interaction, whereas UPy · UIM **80 · 17** has an additional constructive interaction (Figure 5.15).

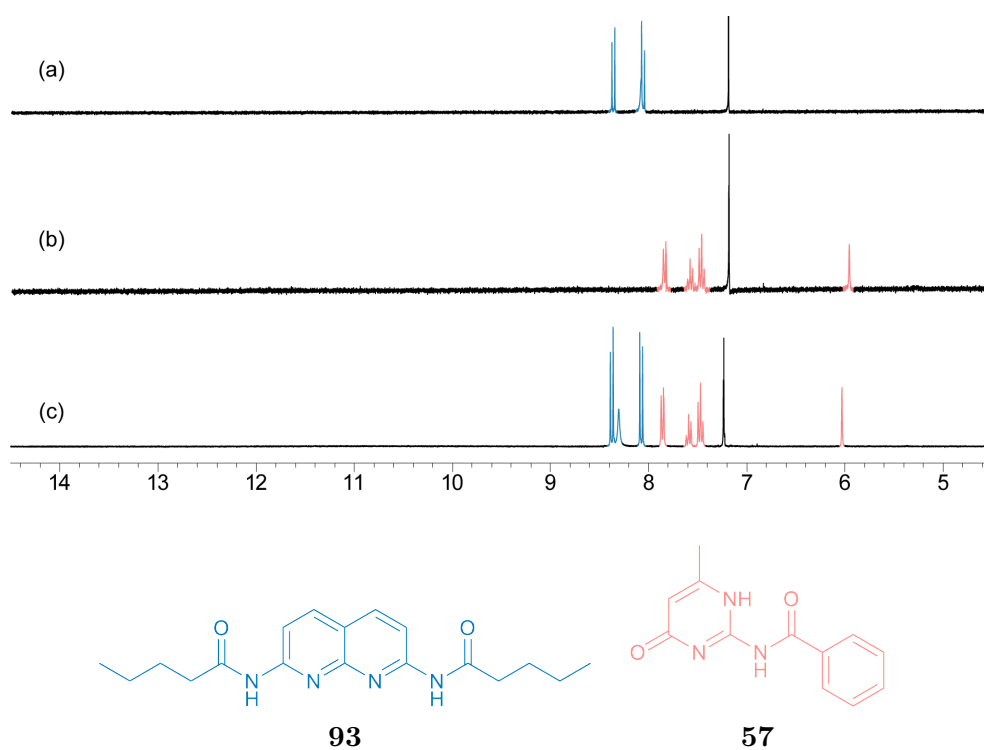


Figure 5.13: ^1H NMR spectra (10 mM, 300 MHz, CDCl_3) of the DAN **93** and AIC **57** molecules (a) DAN **93**, (b) AIC **57** and (c) a 1:1 mixture of DAN **93** and AIC **57**.

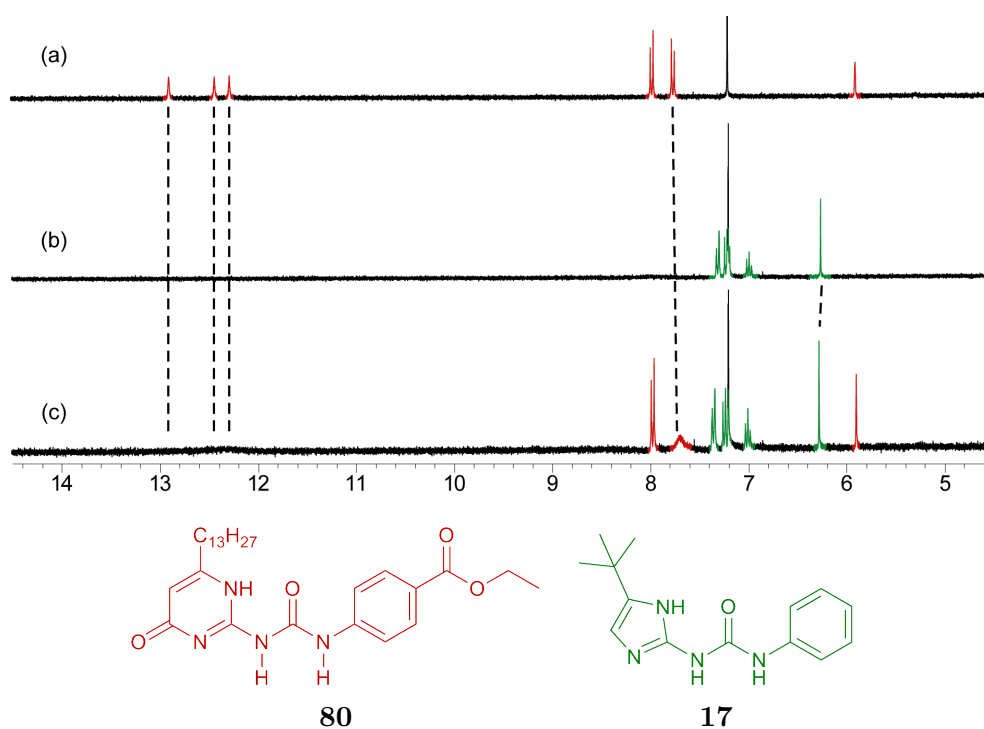


Figure 5.14: ^1H NMR spectra (10 mM, 300 MHz, CDCl_3) of the UPy **80** and UIM **17** molecules (a) UPy **80**, (b) UIM **17** and (c) a 1:1 mixture of UPy **80** and UIM **17**.

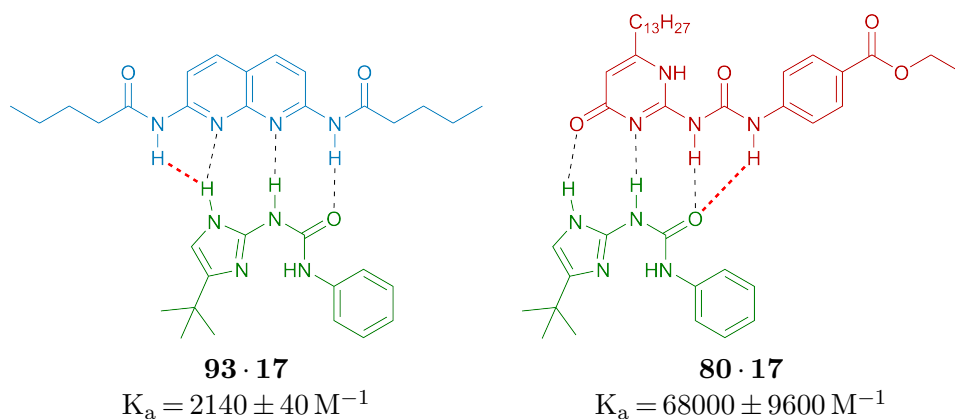


Figure 5.15: The additional secondary electrostatic interactions can have a dramatic effect over the stability of the complex.

The final possible dimer was between the UPy **80** and AIC **57**. Comparison of the ^1H NMR spectra of the individual molecules to a 1:1 mixture (Figure 5.16) indicated that there were no interactions between these molecules. All of the association constant results have been collated into Table 5.1. Although some undesired complexes can be formed it should be noted that AIC **57** only forms complexes with UIM **17** and it was predicted that this would drive self-sorting towards the desired complexes, along with the fact that formation of the desired complexes gives the maximum number of hydrogen

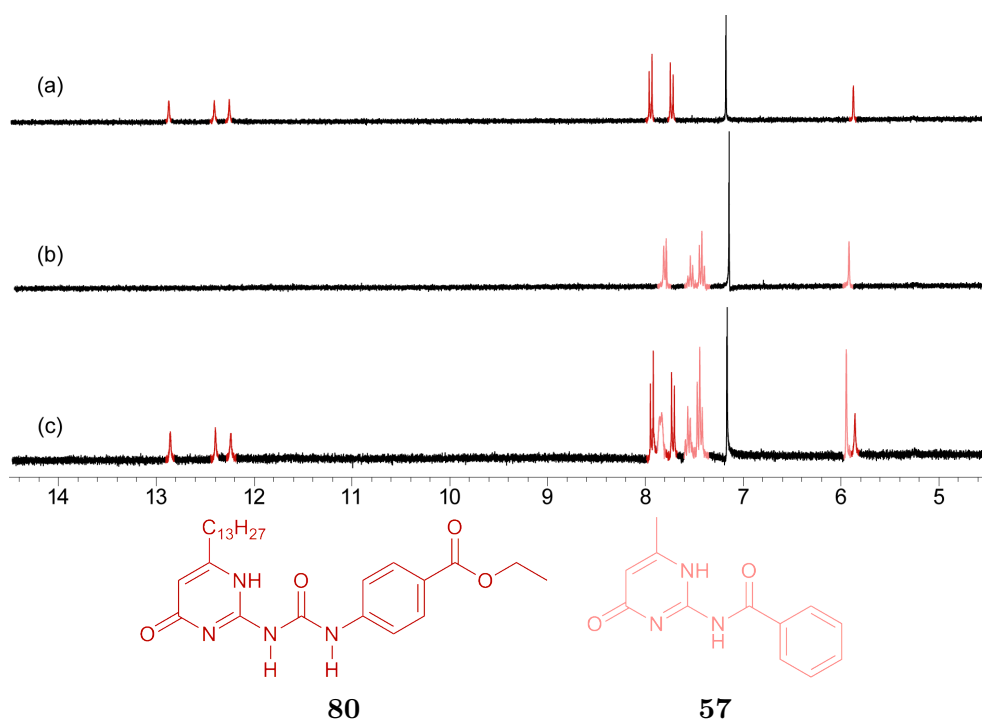


Figure 5.16: ^1H NMR spectra (10 mM, 300 MHz, CDCl_3) of the UPy **80** and AIC **57** molecules (a) UPy **80**, (b) AIC **57** and (c) a 1:1 mixture of UPy **80** and AIC **57**.

Table 5.1: The association constants of all possible homo- and heterodimers that can be formed between the four molecules.

Complex		$K_a/K_{\text{dim}} (\text{M}^{-1})$
UPy · UPy	80 · 80	6×10^7
DAN · DAN	93 · 93	5
AIC · AIC	57 · 57	3.6 ± 0.3
UIM · UIM	17 · 17	10 ± 2.2
UPy · DAN	80 · 93	4×10^5
AIC · UIM	57 · 17	33400 ± 16100
AIC · UPy	57 · 80	< 1
AIC · DAN	57 · 93	< 1
UIM · UPy	17 · 80	68000 ± 9600
UIM · DAN	17 · 93	2140 ± 40

bonds, so should therefore be thermodynamically favourable.

The ^1H NMR spectra of the 1:1:1:1 mixture at 10, 5 and 1 mM confirmed that a high fidelity system had been established; the ^1H NMR spectrum of the mixture was an overlay of the desired pairs (Figure 5.17) at all concentrations. There were some small differences in the NH region which could indicate that there was not 100% formation of the desired complexes. One of the UPy **80** NH's does not appear in the self-sorting mixture and the DAN **93** NH moved as the concentration was decreased. At 10 mM there was some evidence of a small amount of an undesired component being present because there were some broad peaks at ~ 10.5 , 7.8 and 6.4 ppm which were not present in the individual heterodimers. At lower concentrations these undesired components are not visible, even though the fidelity of the desired complexes is probably lower. This could be because, even though the relative concentration of the undesired components is higher compared to the desired complexes, it is too low to be detectable by ^1H NMR.

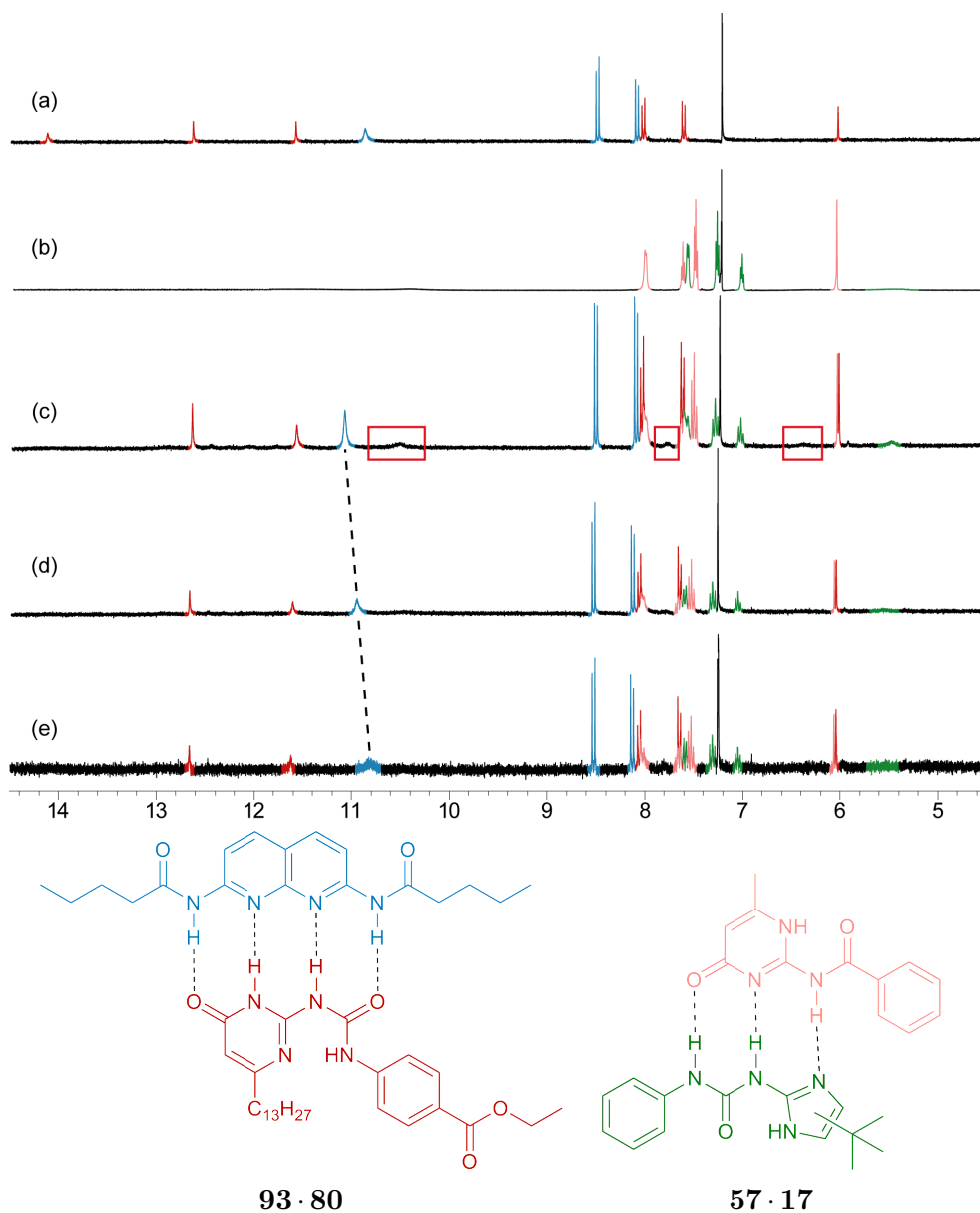


Figure 5.17: ^1H NMR spectra of the self-sorting mixture UPy·DAN **80·93** and AIC·UIM **57·17**. (a) UPy·DAN **80·93** (5 mM, 300 MHz, CDCl_3), (b) AIC·UIM **57·17** (5 mM, 300 MHz, CDCl_3), (c) UPy·DAN **80·93** and AIC·UIM **57·17** (10 mM, 300 MHz, CDCl_3), (d) UPy·DAN **80·93** and AIC·UIM **57·17** (5 mM, 300 MHz, CDCl_3) and (e) UPy·DAN **80·93** and AIC·UIM **57·17** (1 mM, 300 MHz, CDCl_3). Red boxes indicate where there is evidence of undesired components being present.

The fidelity of self-sorting systems can be calculated by dividing the sum of the concentration of the desired complexes by the sum of the concentration of all other components (Equation 5.1).²⁰⁴ In the proposed DAN · UPy **93 · 80**, AIC · UIM **57 · 17** self-sorting system there is the possibility for all four monomers and homodimers to be present along with the six possible heterodimers (DAN · UPy, DAN · UIM, DAN · AIC, UPy · UIM, UPy · AIC and UIM · AIC).

$$\text{Fidelity} = \frac{\text{total concentration of desired complexes}}{\text{total concentration of all complexes}} \quad (5.1)$$

$$\text{Fidelity} = \frac{[\text{DAN} \cdot \text{UPy}] + [\text{AIC} \cdot \text{UIM}]}{[\text{DAN}] + [\text{UPy}] + [\text{UIM}] + [\text{AIC}] + [\text{UPy} \cdot \text{UPy}] + [\text{DAN} \cdot \text{DAN}] + [\text{AIC} \cdot \text{AIC}] + [\text{UIM} \cdot \text{UIM}] + [\text{DAN} \cdot \text{UPy}] + [\text{DAN} \cdot \text{UIM}] + [\text{DAN} \cdot \text{AIC}] + [\text{UPy} \cdot \text{UIM}] + [\text{UPy} \cdot \text{AIC}] + [\text{UIM} \cdot \text{AIC}]}$$

Since the association constants of all possible dimers was known it was possible to estimate the concentration of each component in solution and hence the fidelity of the system. The concentration of all possible components in a 1:1:1:1 mixture of DAN **93**, UPy **80**, AIC **57** and UIM **17** was calculated as the concentration of all components was increased using the HySS program.²¹⁹ It was possible to see that the concentration of the desired components (AIC · UIM **57 · 17** and UPy · DAN **80 · 93**) was significantly higher than the other components throughout (Figure 5.18a). AIC · UIM **57 · 17** had a higher concentration than UPy · DAN **80 · 93**, presumably because AIC **57** and UIM **17** have low dimerisation constants, so there is less competition for assembly of the heterodimer. The fidelity of the system at different concentrations was then calculated (Figure 5.18b)

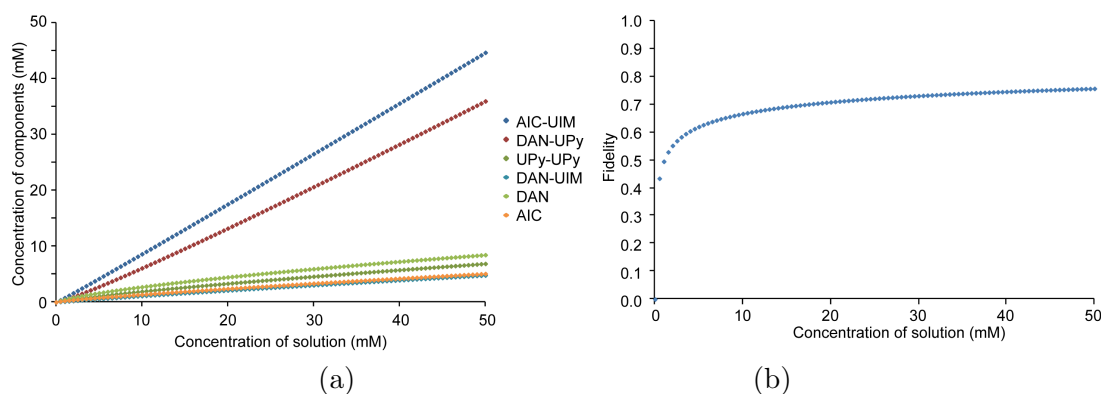


Figure 5.18: The fidelity of the self-sorting DAN · UPy **93 · 91**, AIC · UIM **57 · 17** system changes as the concentration is altered. High concentrations lead to higher fidelities being obtained.

and it was found that as the concentration increased the fidelity increased, to a maximum of ~ 0.7 at 50 mM. The low fidelity at low concentrations was presumably because at low concentrations there was a comparatively high concentration of the individual monomers, increasing the concentration of the undesired components. The fidelity curve begins to plateau at ~ 10 mM, which is the concentration that was chosen for these experiments to be carried out at.

5.5 Signalling Cascade

Biology uses signalling cascades in order to carry out a desired process.²²⁰ The binding of one molecule to a protein can alter the tertiary structure of the protein, revealing a second binding site where other molecules can interact, resulting in the desired output signal being released. Alternatively, the output obtained from the formation of the first complex can be stopped by the addition of a competing molecule, allowing the biological processes to be switched on and off.

In the self-sorting system that was described in the previous section many undesired complexes can be formed. These undesired interactions can be exploited in a signalling cascade if the order of addition of the molecules is carefully considered. Because UIM **17** can interact with both UPy **80** and DAN **93**, as well as AIC **57**, three signalling cascades could be envisioned (Figure 5.19) where UPy·UPy **80·80** was the starting point.

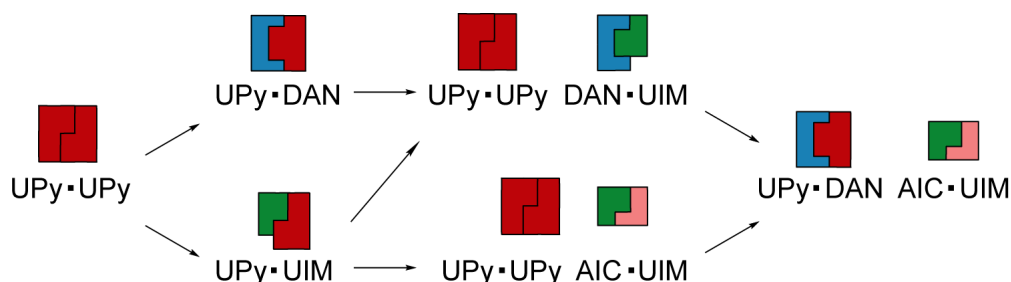


Figure 5.19: Some of the different signalling cascades that are possible using AIC **57**, UIM **17**, UPy **80** and DAN **93**.

The first that was investigated began with the UPy·UPy **80·80** homodimer, which was broken by the addition of UIM **17**. Addition of AIC **57** regenerated the UPy **80** homodimer along with the UIM·AIC **17·57** heterodimer and finally, addition of DAN **93** resulted in the self-sorting system being assembled. The ¹H NMR spectra (Figure 5.20) confirmed that the presence of the UPy **80** homodimer could be switched on and off by the presence (Figure 5.20a and c) and disappearance (Figure 5.20b and d) of the UPy **80** NH signals.

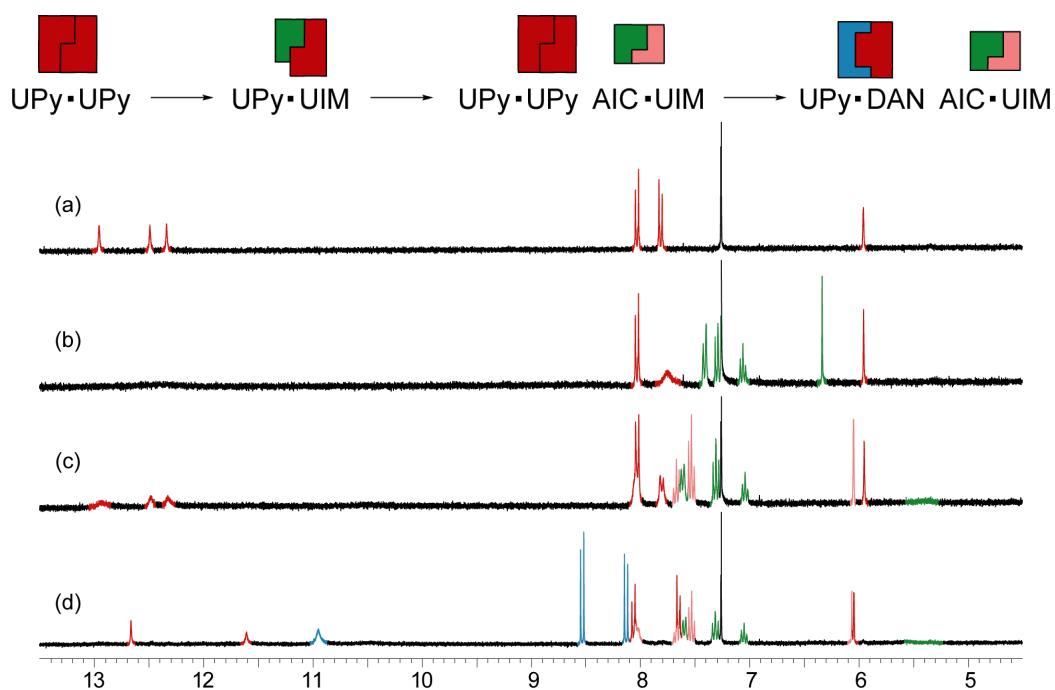


Figure 5.20: The signalling system (a) UPy·UPy **80·80**, (b) UPy·UIM **80·17**, (c) UPy·UPy **80·80** + UIM·AIC **17·57**, (d) UPy·DAN **80·93** + UIM·AIC **17·57** (10 mM, 300 MHz, CDCl₃).

The composition of the mixture at each stage of the signalling cascade was calculated using HySS.²¹⁹ The concentration of all molecules already present was set to 10 mM and the concentration of the molecule being added was increased from 0–50 mM (Figure 5.21). The fidelity of each step of the signalling cascade during addition of the new molecule was calculated using Equation 5.1. It was found that in all cases the fidelity was at a maximum when an equimolar amount of all molecules were present, which was not surprising because if there is an excess of one of the components present then the concentration of undesired components will begin to increase. The fidelity of the UPy·UIM **80·17** complex (Figure 5.21a) is quite low (0.24), probably because the UPy **80** dimer has a much stronger association constant than the UPy·UIM **80·17** heterodimer, so UIM **17** cannot compete and dissociate the UPy **80** homodimer. The next step, addition of AIC **57** (Figure 5.21b), had a very high fidelity (0.91), which is most likely because AIC **57** only interacts with UIM **17**, leaving UPy **80** free to assemble into strong homodimers. The final step was addition of DAN **93** (Figure 5.21c), which was found to have a moderate fidelity of 0.67 at equimolar concentrations.

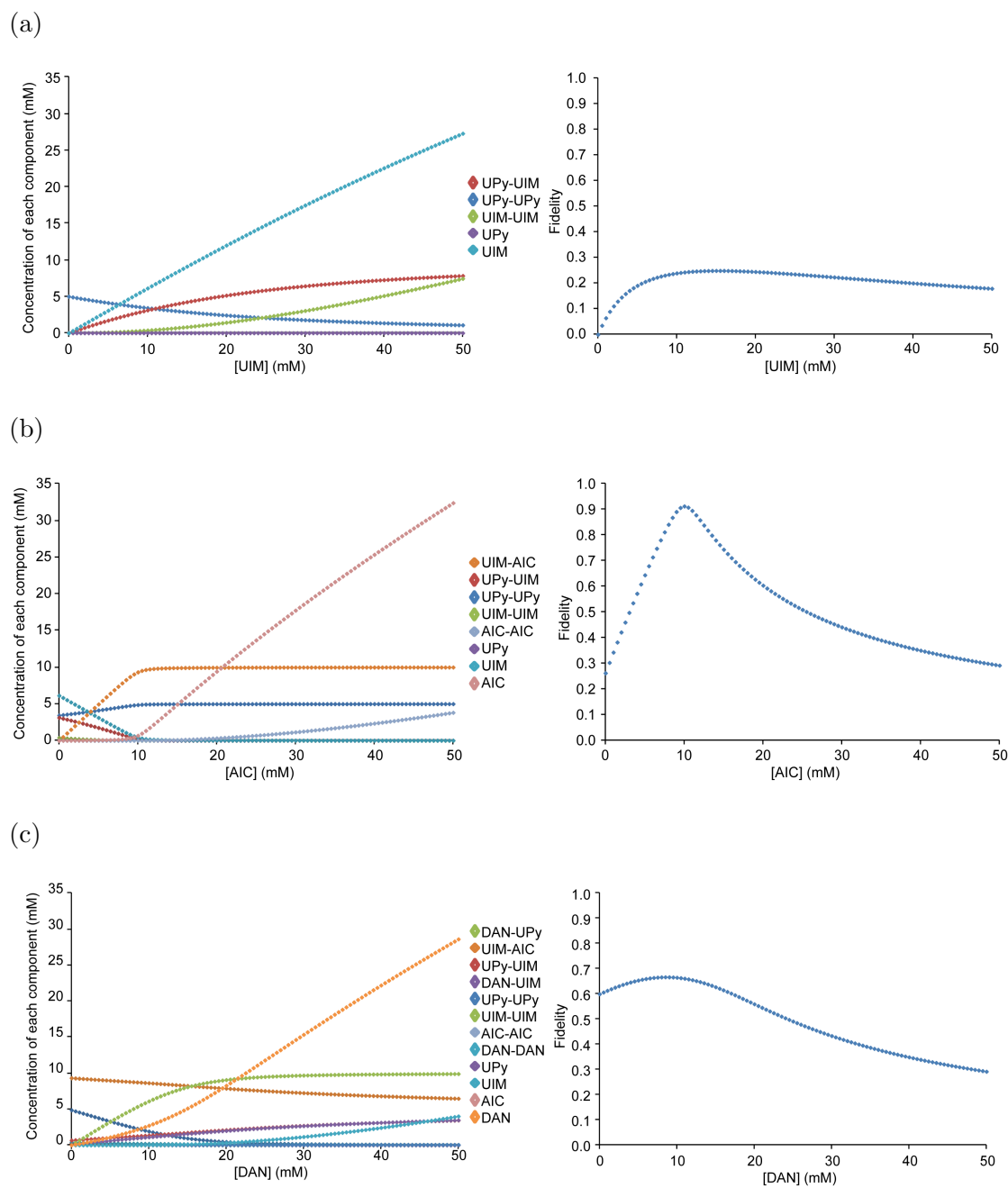


Figure 5.21: The concentration of all components of the mixture was calculated as one molecule was added. (a) $UPy \cdot UPy \rightarrow UPy \cdot UIM$, (b) $UPy \cdot UIM \rightarrow UPy \cdot UPy + UIM \cdot AIC$ and (c) $UPy \cdot UPy + UIM \cdot AIC \rightarrow UPy \cdot DAN + UIM \cdot AIC$. All mixtures began at 10 mM.

The second signalling cascade began the same as the first, by the formation of the UPy·UIM **80**·**17** complex (fidelity = 0.24). The next step was the addition of DAN **93** with the aim of reforming the UPy **80** homodimer along with the assembly of the DAN·UIM **93**·**17** heterodimer. The final step was addition of AIC **57** to give the original self-sorting system once more. The ^1H NMR spectra (Figure 5.22) for this signalling cascade indicated that there was a low fidelity for some of the steps. In the third step (Figure 5.22c) there was evidence for the presence of the undesired UPy·DAN **80**·**93** complex. This was most likely because this undesired heterocomplex had a higher association constant than the desired DAN·UIM **93**·**17** complex. Calculation of the fidelity of each of the steps (Figure 5.23), using the same protocol as described above, confirmed that the fidelity was low for this signalling cascade. Addition of DAN **93** to the UPy·UIM **80**·**17** mixture only had a fidelity of 0.53.

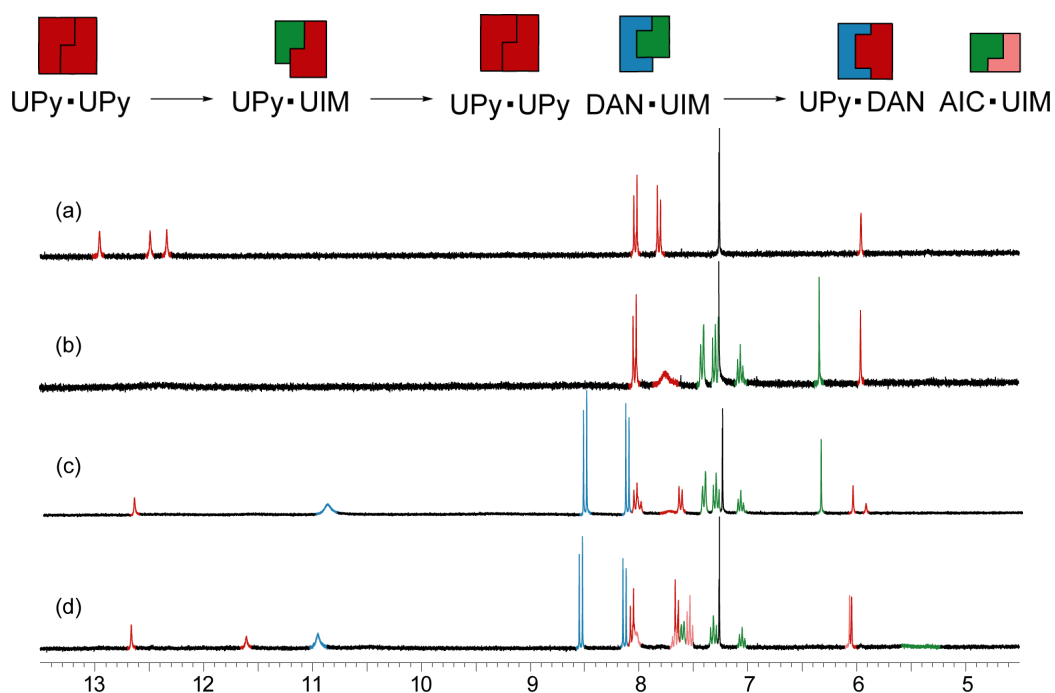


Figure 5.22: The signalling system (a) UPy·UPy **80**·**80**, (b) UPy·UIM **80**·**17**, (c) UPy·UPy **80**·**80** + UIM·DAN **17**·**93**, (d) UPy·DAN **80**·**93** + UIM·AIC **17**·**57** (10 mM, 300 MHz, CDCl_3).

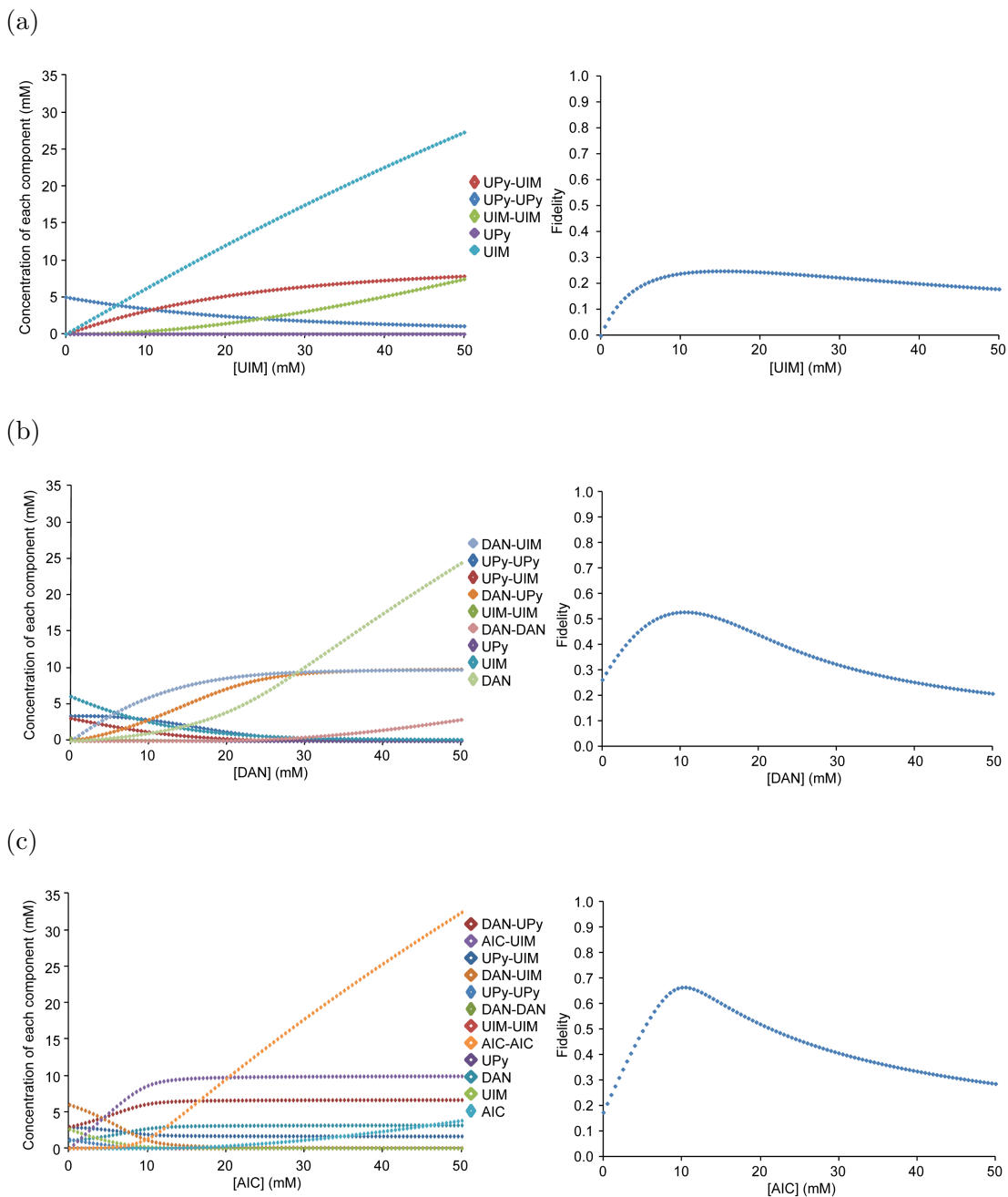


Figure 5.23: The concentration of all components of the mixture was calculated as one molecule was added. (a) $UPy \cdot UPy \rightarrow UPy \cdot UIM$, (b) $UPy \cdot UIM \rightarrow UPy \cdot UPy + UIM \cdot DAN$ and (c) $UPy \cdot UPy + UIM \cdot DAN \rightarrow UPy \cdot DAN + UIM \cdot AIC$. All mixtures began at 10 mM.

The third possible signalling cascade began with the addition of DAN **93** to a solution of UPy **80**, followed by addition of UIM **17** to give UPy·UPy **80**·**80** and DAN·UIM **93**·**17** and finally addition of AIC **57** to return to the self-sorting system. The ^1H NMR spectra (Figure 5.24) were obtained for each stage of this cascade and again the third step (UPy·UPy **80**·**80** and DAN·UIM **93**·**17**) appeared to have a low fidelity due to the presence of two species of UPy **80**. The fidelity of each step was then calculated and it was found that the UPy·DAN **80**·**93** heterodimer had a fidelity of 0.58. The addition of UIM **17** gave the same, low fidelity (0.53) mixture that was obtained between UPy **80**, DAN **93** and UIM **17** in the previous cascade. Finally, addition of AIC **57** resulted in the self-sorting system being obtained with the same fidelity (0.67) that had been previously obtained.

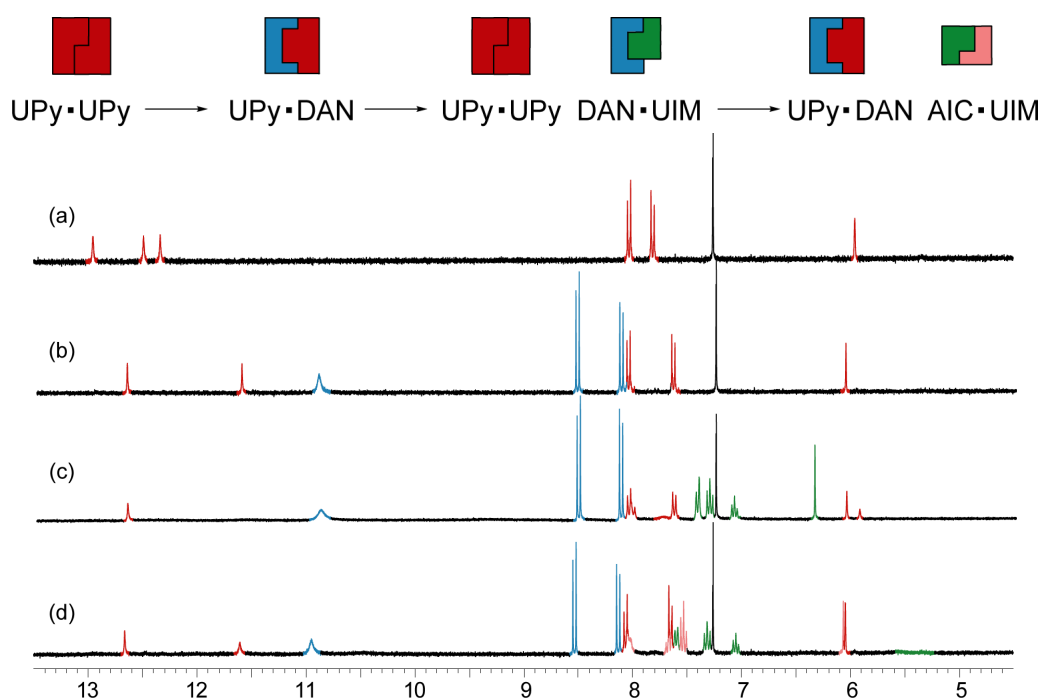


Figure 5.24: The signalling system (a) UPy·UPy **80**·**80**, (b) UPy·DAN **80**·**93**, (c) UPy·UPy **80**·**80** + UIM·DAN **17**·**93**, (d) UPy·DAN **80**·**93** + UIM·AIC **17**·**57** (10 mM, 300 MHz, CDCl_3).

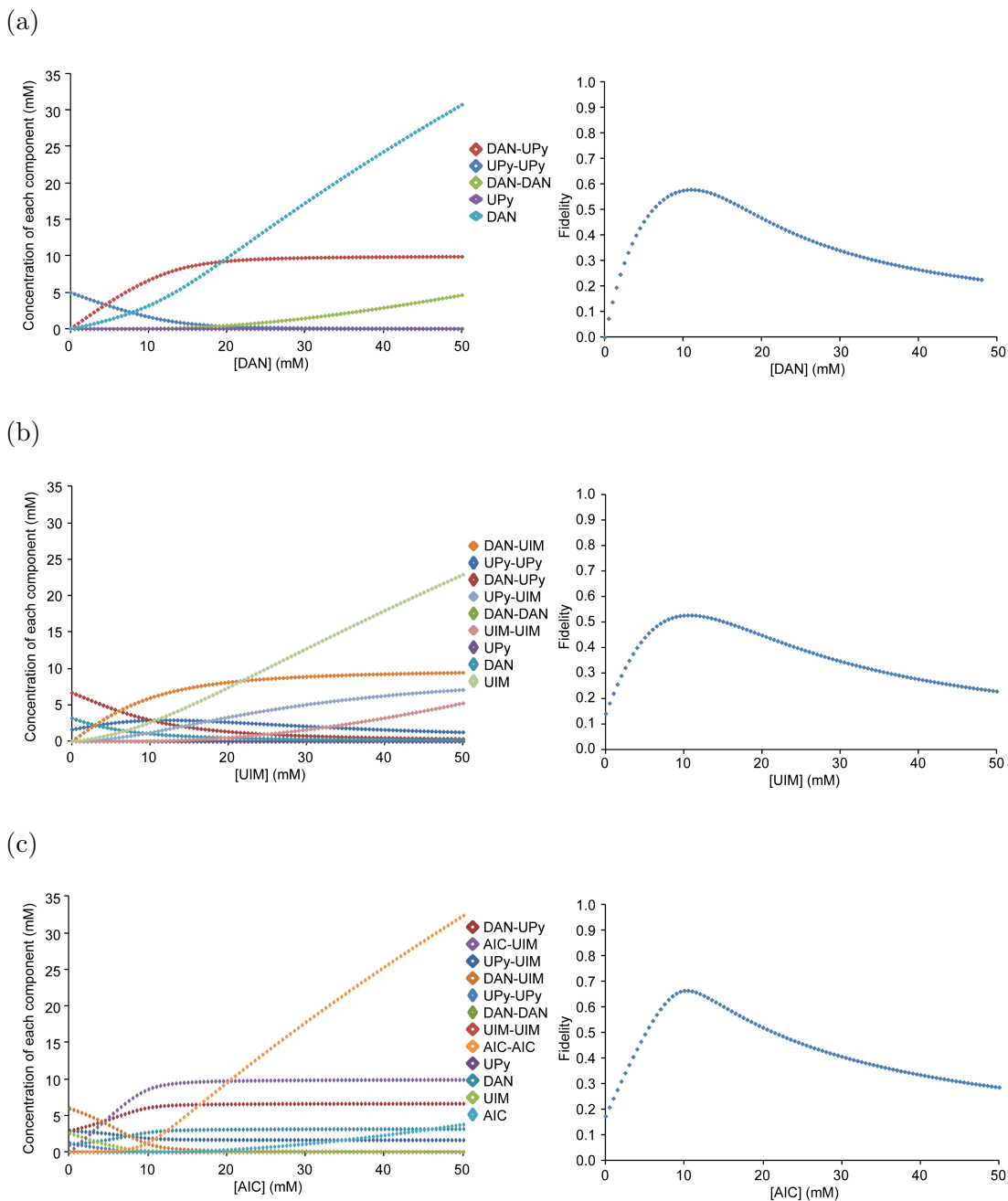


Figure 5.25: The concentration of all components of the mixture was calculated as one molecule was added. (a) $\text{UPy} \cdot \text{UPy} \rightarrow \text{UPy} \cdot \text{DAN}$, (b) $\text{UPy} \cdot \text{DAN} \rightarrow \text{UPy} \cdot \text{UPy} + \text{UIM} \cdot \text{DAN}$ and (c) $\text{UPy} \cdot \text{UPy} + \text{UIM} \cdot \text{DAN} \rightarrow \text{UPy} \cdot \text{DAN} + \text{UIM} \cdot \text{AIC}$. All mixtures begin at 10 mM.

5.6 Photosensitive signalling cascade

Attachment of a photolabile group to one of the molecules would result in a photosensitive signalling cascade, allowing interactions to be switched on and off by light. Photolabile protecting groups are often favoured over chemically labile protecting groups, especially in supramolecular systems, because they can be removed without having to add extra reagents and can often have 100% cleavage yields. There are vast examples of photolabile protecting groups in the literature,²²¹ many of which have been used in supramolecular systems.^{15,222} Some molecules can use photoswitchable groups, which change the conformation of the molecule, often using steric hindrance to block the hydrogen-bonding array from making interactions.²²³⁻²²⁷ Azobenzene can reversibly switch conformation using UV and Vis light. If large groups are appended to the azobenzene functionality then the hydrogen-bonding atoms can be sterically hindered to interact with complementary molecules (Figure 5.26). Other systems have used ring formation to turn interactions on and off.^{228,229} In these systems it is the modulation of the pK_a of the hydrogen-bonding atoms that alters the binding affinity of the system (Figure 5.27).²³⁰

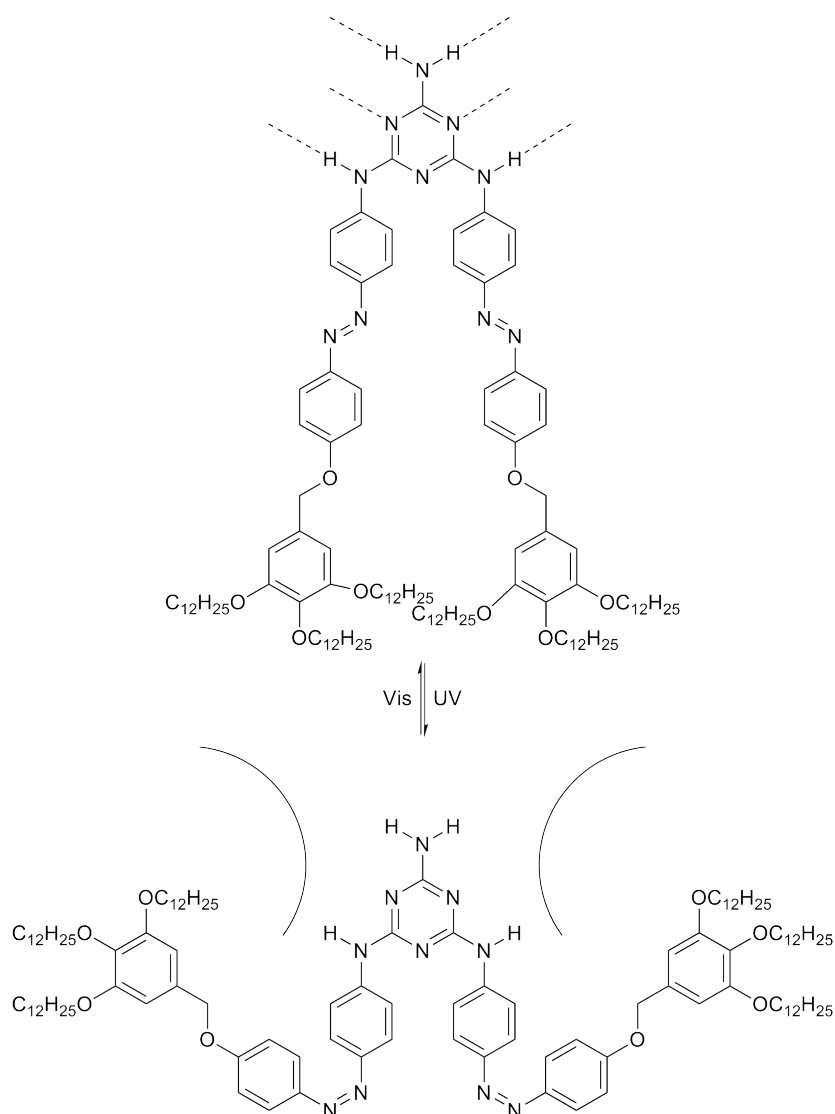


Figure 5.26: The binding site can be sterically hindered by changing the conformation of the azobenzene moieties, allowing interactions to be switched on and off.

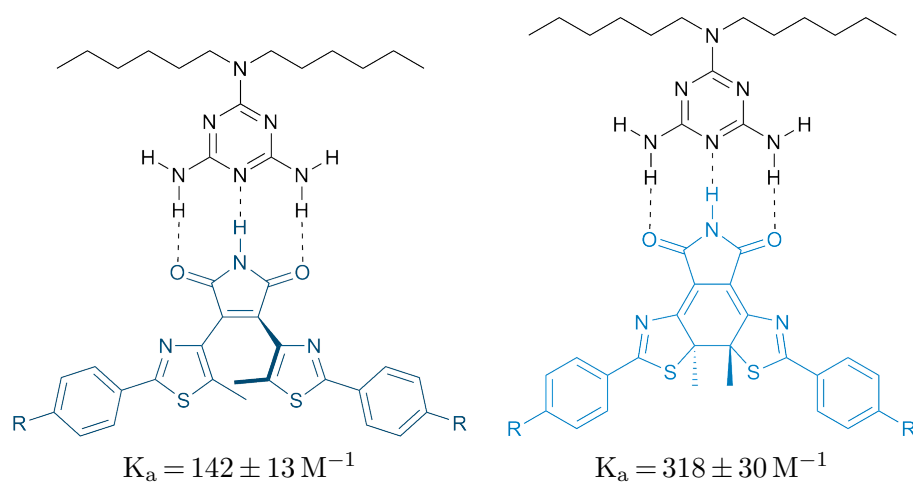


Figure 5.27: The binding affinity of supramolecular heterodimers can be modulated by changing the $\text{p}K_a$ of the hydrogen-bonding atoms using ring formation ($\text{R} = \text{CO}_2\text{C}_6\text{H}_{13}$).

The photolabile 2-nitrobenzyl group has been added to the UPy **19** moiety before to give UPy* **122** (Figure 5.28).^{16,231–233} This group is removed using UV light (Scheme 5.12),²²¹ firstly the nitro group forms a radical, which removes a hydrogen from the benzyl carbon and the resultant transition state can rearrange so that there is no longer a radical present. An intramolecular cyclisation then gives a fused ring system, which reopens to give the desired hydrogen-bonding motif along with the 2-nitrosobenzaldehyde side product. Since the 2-nitrosobenzaldehyde remains in the mixture once it has been removed it is possible that it could compete with the desired interactions however, it does not contain any hydrogen-bond arrays, so it is unlikely that it will be able to interact strongly.

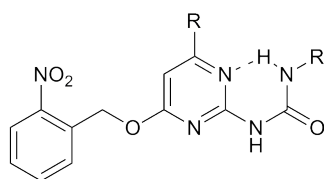
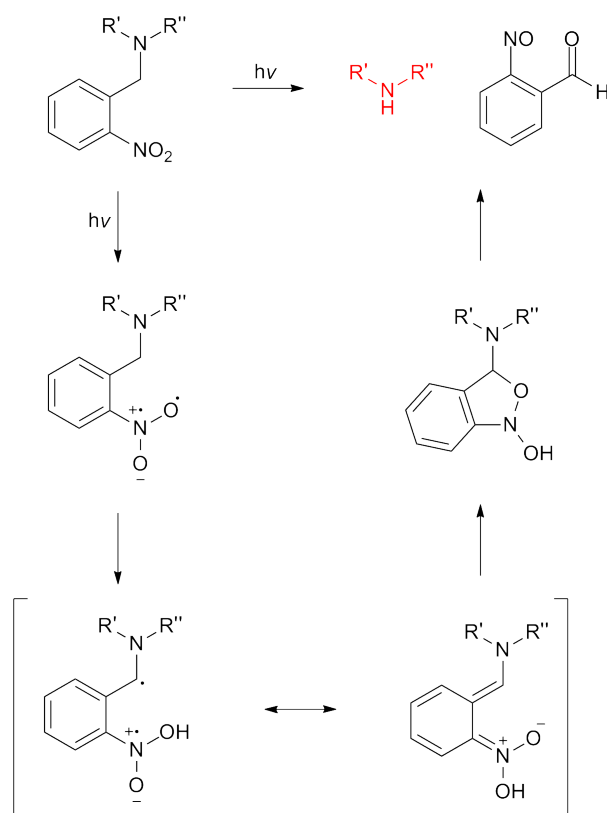
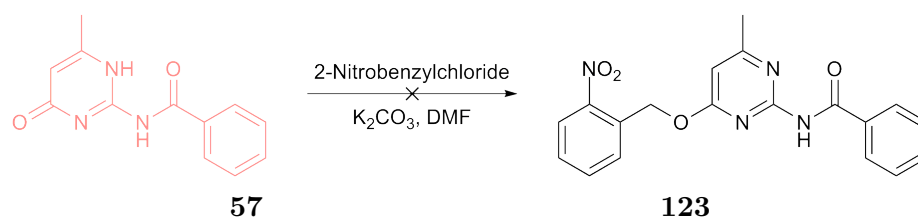
**122**

Figure 5.28: The 2-nitrobenzyl photolabile group has been appended to the end of the quadruple hydrogen-bond array of UPy **19** to give UPy* **122**.

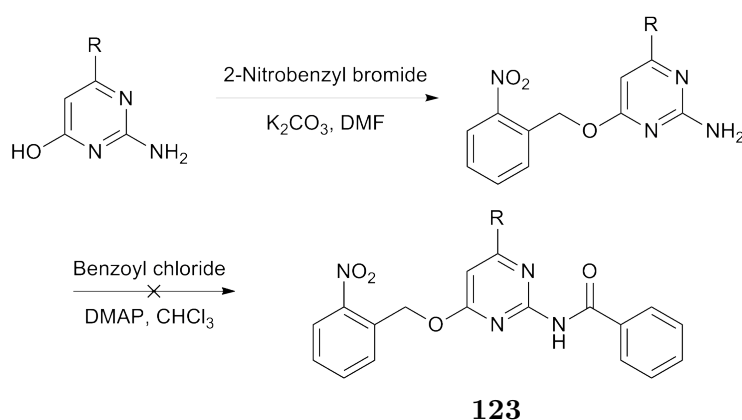


Scheme 5.12: The mechanism for the photocleavage of the 2-nitrobenzyl protecting group to reveal AIC **57**, along with the 2-nitrosobenzaldehyde.

Because the AIC **57** contains the same *isocytosine* functional group as UPy **19** it was thought that a photolabile group could be appended to the AIC **57** molecule. Attempts to append the 2-nitrophenyl group to the oxygen of AIC **57** to give AIC* **123** were all unsuccessful (Scheme 5.13). An alternative synthesis was also attempted; alkylation of 2-nitrobenzyl to the *isocytosine*, followed by acylation however, no product was isolated (Scheme 5.14). This was probably because, unlike UPy* **122**, AIC* **123** cannot form a stabilising intramolecular hydrogen bond and instead contains destabilising electrostatic and steric interactions (Figure 5.29). For an intramolecular hydrogen bond to be present



Scheme 5.13: Synthesis of AIC* **123** was unsuccessful when *O*-alkylation of AIC **57** was attempted.



Scheme 5.14: An alternative synthesis of AIC* **123** was attempted, but this was unsuccessful.

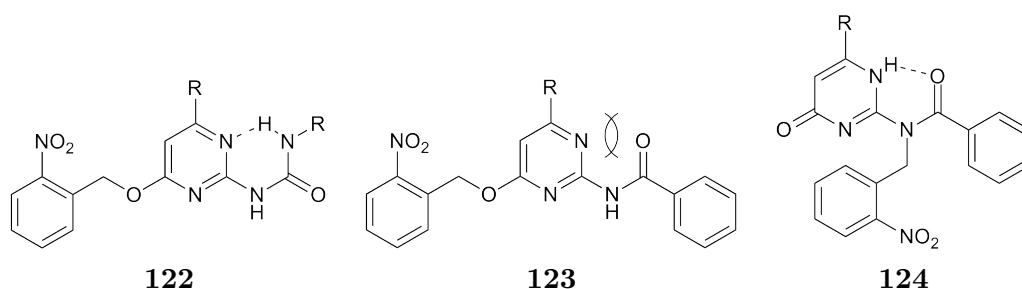
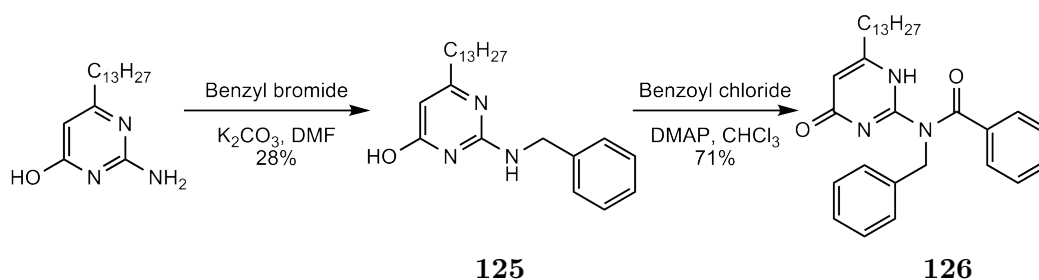


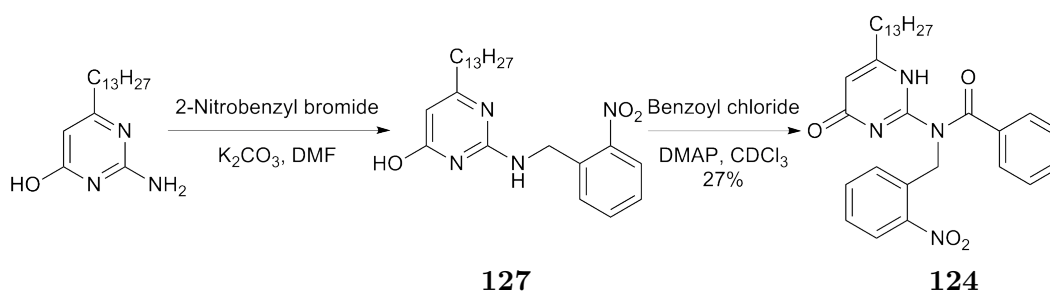
Figure 5.29: UPy* **122** is stabilised by an intramolecular hydrogen bond, but AIC* **123** is destabilised by electrostatic interactions, preventing it from being synthesised. Appending the 2-nitrobenzyl group to the amide allows a stabilising intramolecular hydrogen bond to be formed.

the 2-nitrophenyl group must be attached to the amide functional group to give AIC* **124** (Figure 5.29).

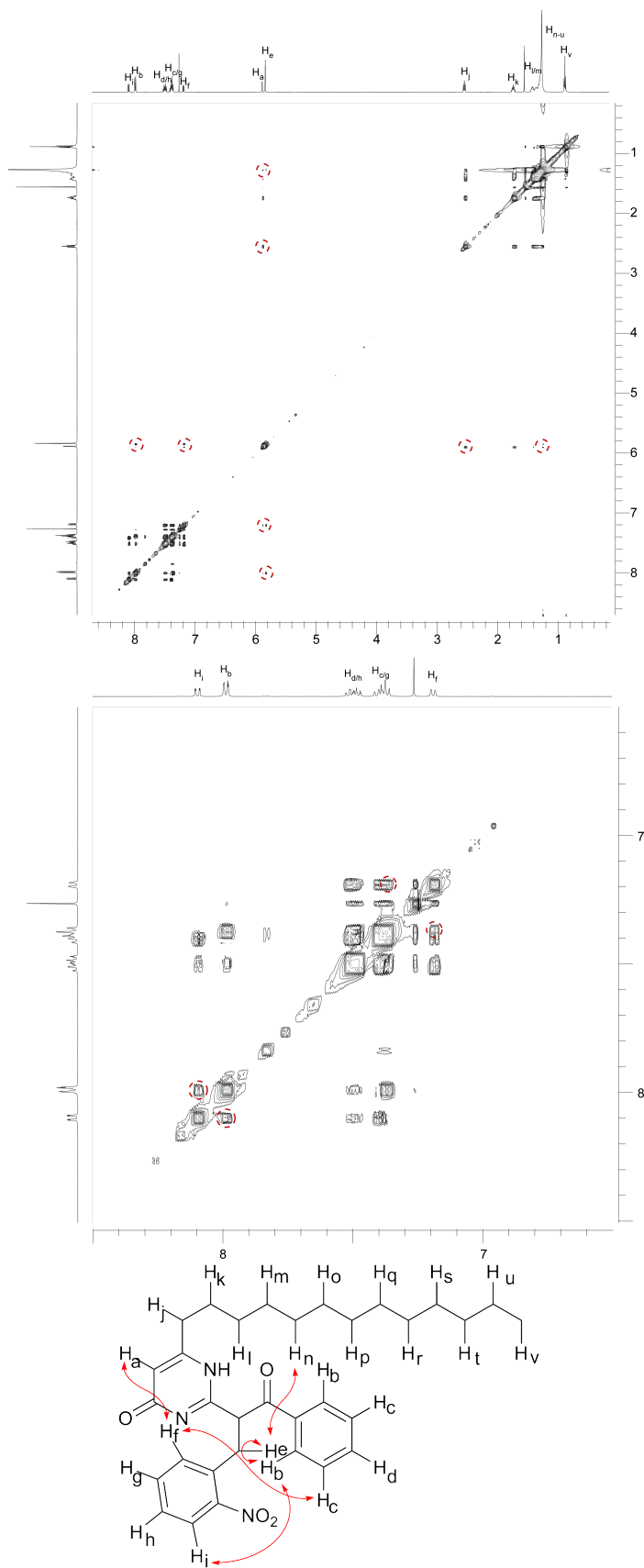
Attachment of the 2-nitrophenyl group to the amine group was first attempted *via* reductive amination. This approach was never successful which was thought to be due to the low reactivity of the amine, preventing the imine from being formed. The second approach was *via* *N*-alkylation; in the presence of K_2CO_3 in DMF. This reaction was first attempted with benzylbromide (Scheme 5.15) which resulted in both the *N*- and *O*-alkylation products being obtained, which explains the low yield (28%) for this step. These were separated by trituration (hexane) and 1H - 1H NOESY confirmed that the resultant solid was the desired *N*-alkylation product **125**. Reaction with benzoyl chloride in the presence of a catalytic amount of DMAP gave the AIC⁺ **126**. These reactions were then carried out with the 2-nitrophenyl group (Scheme 5.16) to give the 2-nitrophenyl *isocytosine* **127**, but in this case separation of the *N*- and *O*-alkylation products was not possible. Both *O*- and *N*-alkylation products were taken onto the acylation step to give the AIC* **124** in 27% yield over both steps. Confirmation that the desired product had been synthesised was obtained by 1H - 1H NOESY NMR (Figure 5.30). No correlations were observed between the isocytosine proton and the CH_2 of the 2-nitro benzyl group, indicating that they are not positioned next to each other.



Scheme 5.15: Synthesis of AIC⁺ **126** *via* *N*-alkylation.



Scheme 5.16: Synthesis of AIC* **124** *via* *N*-alkylation.

**124**Figure 5.30: The ^1H - ^1H NOESY spectra of AIC* **124** (25 mM, 500 MHz, CDCl_3).

To investigate the molecular interaction of substituted AIC molecules with UIM **17**, UPy **80** and DAN **93**, AIC⁺ **126** was used due to its ease of handling. Comparison of the ¹H NMR spectra of the individual molecules with a 1:1 mixture was carried out to see if there were any complexation induced shifts, which are indicative of interactions between molecules. It was found that, as expected, there were negligible interactions between AIC⁺ **126** and UIM **17** (Figure 5.31), UPy **80** (Figure 5.32) and DAN **93** (Figure 5.33) because the hydrogen-bond array had been blocked. It was presumed that AIC* **124** would have the same recognition properties as AIC⁺ **126** because they both contain the same hydrogen-bond array.

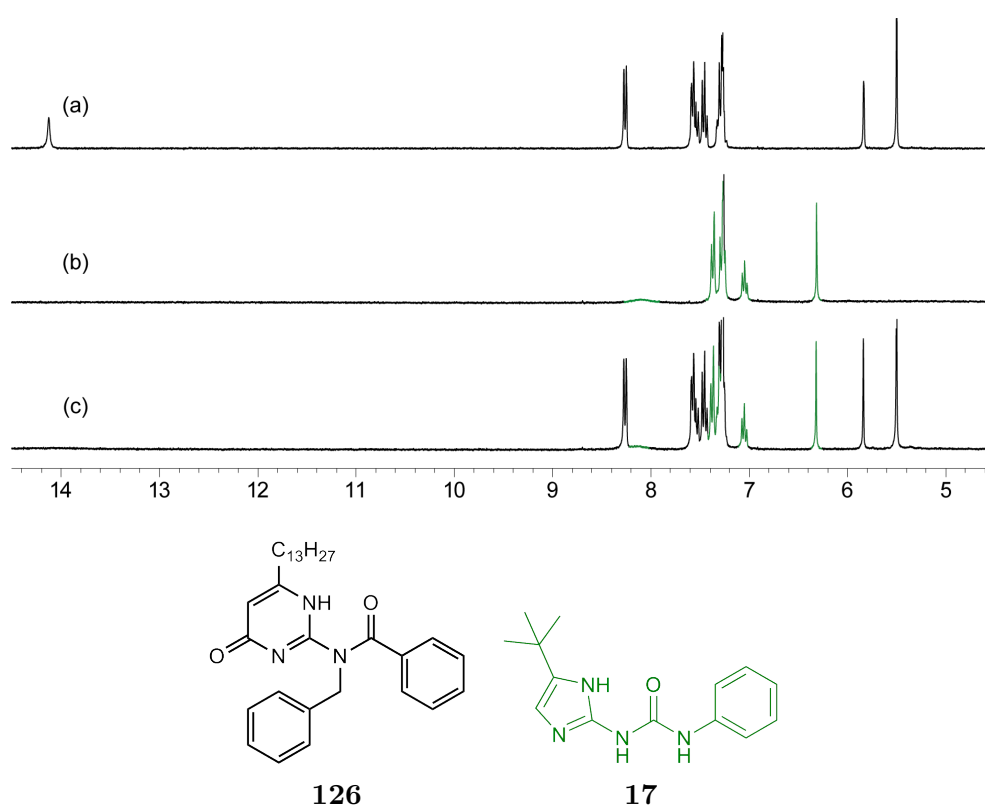


Figure 5.31: Comparison of the ¹H NMR spectra of (a) AIC⁺ **126** and (b) UIM **17** with (c) the 1:1 mixture of AIC⁺ **126** and UIM **17** shows that there are no complexation induced interactions (300 MHz, CDCl₃, 10 mM).

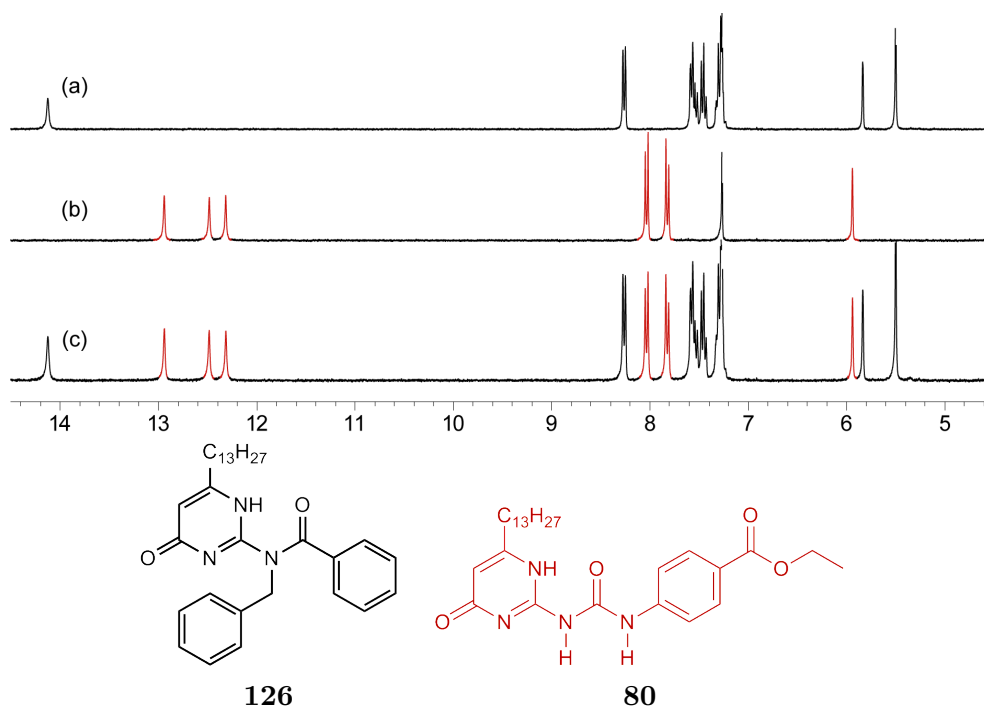


Figure 5.32: Comparison of the ^1H NMR spectra of (a) AIC^+ **126** and (b) UPy **80** with (c) the 1:1 mixture of AIC^+ **126** and UPy **80** shows that there are no complexation induced interactions (300 MHz, CDCl_3 , 10 mM).

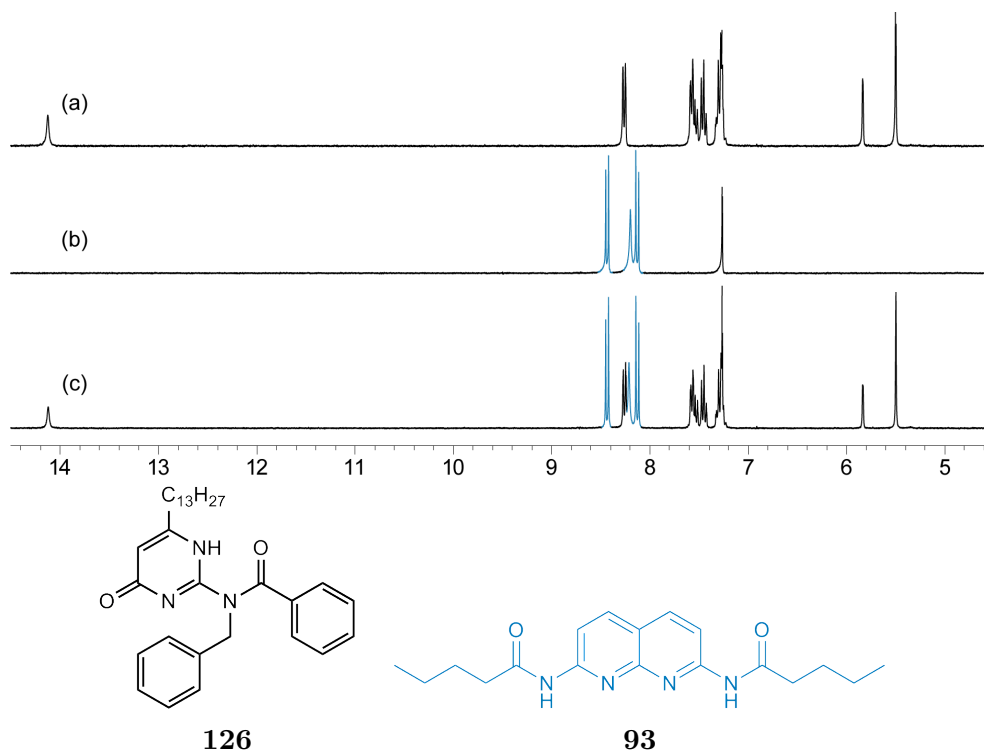


Figure 5.33: Comparison of the ^1H NMR spectra of (a) AIC^+ **126** and (b) DAN **93** with (c) the 1:1 mixture of AIC^+ **126** and DAN **93** shows that there are no complexation induced interactions (300 MHz, CDCl_3 , 10 mM).

Removal of the 2-nitrobenzyl protecting group can be achieved using UV light. The UV spectra of AIC* **124** showed that its maximum absorbance was 260-340 nm (Figure 5.34), so a lamp that emitted at 300 nm was used in the photocleavage experiments. A 10 mM solution of AIC* **124** was irradiated for 4 hr and this gave evidence for the removal of the protecting group (Figure 5.35). However, this reaction did not occur efficiently which could easily be seen in the ^1H NMR spectra, which was messy in the 8-6 ppm region, the important region for studying the self-sorting mixture. It was thought that using wet, degassed CDCl_3 could aid the photocleavage by making protons more easily available and reducing the amount of undesired radicals that were present in solution, but it was found that neither of these changes made any difference to the cleavage yield.

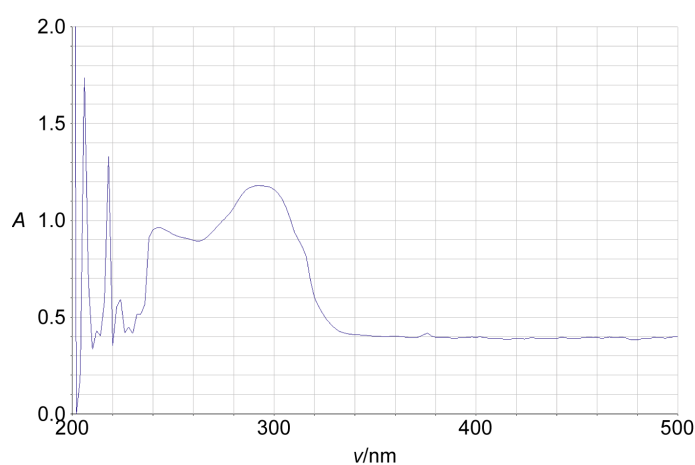


Figure 5.34: The UV spectrum of AIC* **124** ($1\ \mu\text{M}$, CDCl_3).

Currently only examples of self-complementary hydrogen bonding groups using the 2-nitrobenzyl protecting group have appeared in the literature. It could be that the cleavage yield could be increased by the presence of the complementary partner, UIM **17**, allowing a dimer to form when the protecting group is removed. A 1:1 mixture of AIC* **124** and UIM **17** was irradiated (Figure 5.36) and although this showed evidence that UIM **17** was interacting, presumably with AIC **57**, the cleavage yield was not increased and the ^1H NMR became increasingly more messy as the experiment was continued.

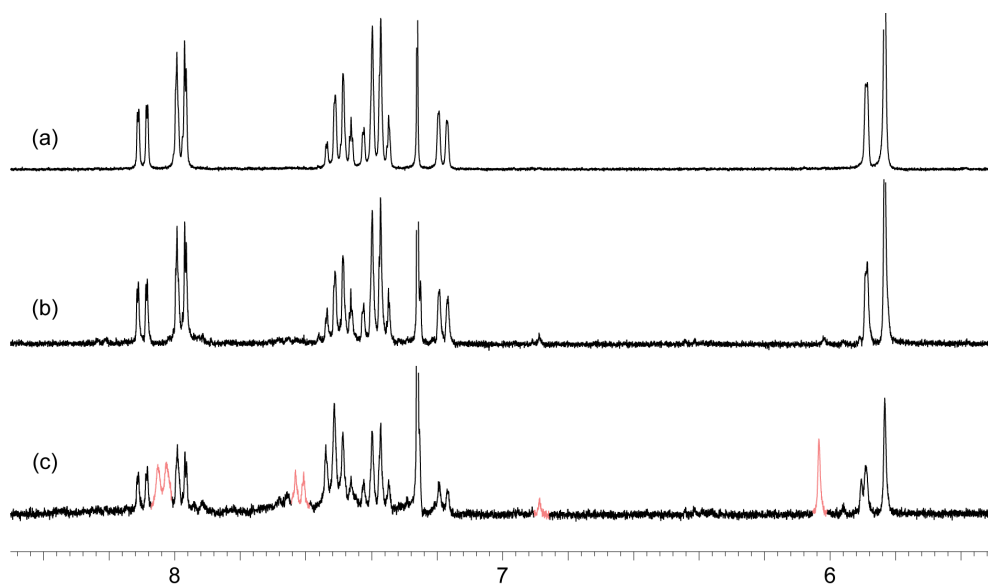


Figure 5.35: Irradiation of AIC* **124** with 300 nm for (a) 0 hr, (b) 0.5 hr, (c) 4 hr (10 mM, 300 MHz, CDCl_3). There is evidence of a new species being present as shown in pink.

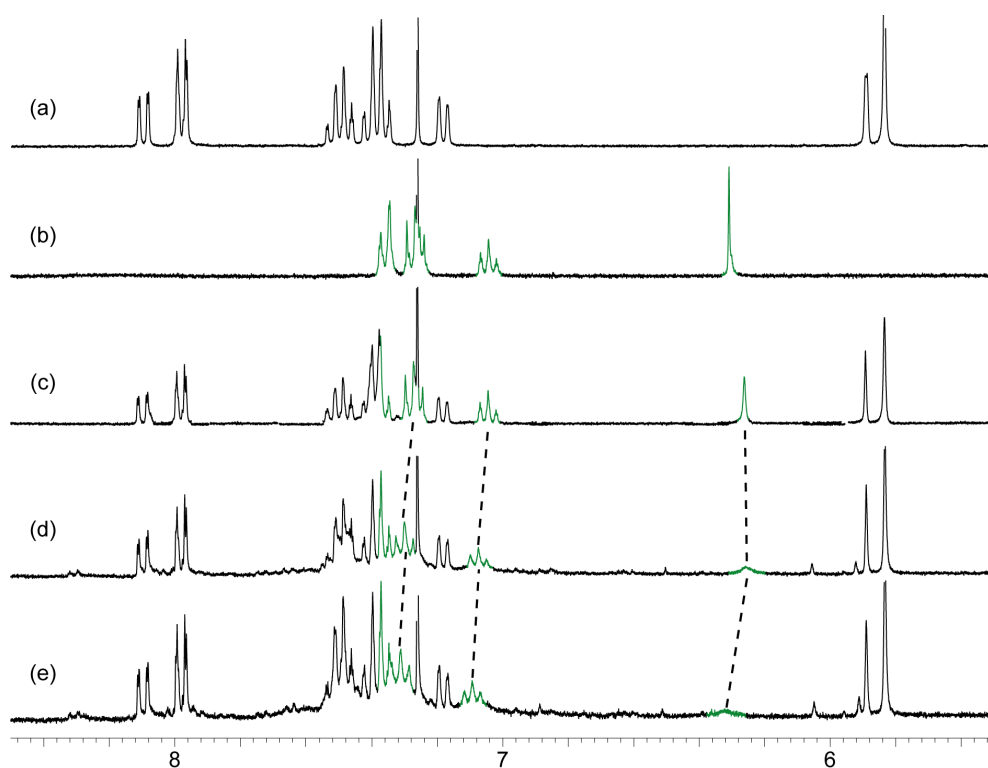


Figure 5.36: There is evidence of interactions being formed between UIM and AIC as the photolabile group is removed. (a) AIC* **124**, (b) UIM **17**, (c) UIM · AIC* **17** · **124**, (d) after irradiation for 1 hr, (e) after irradiation for 2 hrs.

Although there are examples of the 2-nitrobenzyl group being appended to a tertiary amide, there are no examples where there are two aromatic rings in conjugation with the amide, as well as the 2-nitrobenzyl group. This however is the case in the AIC* **124**, which could lead to steric interactions in the transition state structure (Figure 5.37). It seems that in order for the 2-nitrobenzyl group to be efficiently removed it must have more space than it has been given in this last example.

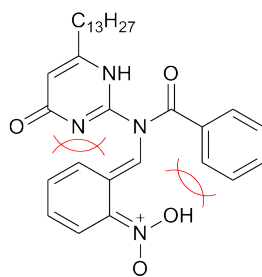


Figure 5.37: The transition state is very sterically hindered, which could be the reason for the poor cleavage yield.

5.7 Conclusions

Complementary hydrogen-bond arrays must be carefully designed in order to allow interactions to occur. Subtle differences in shape complementarity or electrostatic interactions can prevent the assembly of heterodimers. The dynamic nature of the reversibility of hydrogen bonds can lead to broad ^1H NMR spectra being obtained. Although a broad spectrum can indicate that interactions are occurring it makes it difficult to interpret the data, so association constants cannot be obtained. This can prevent some complexes from being used in self-sorting systems even though they probably assemble into heterodimers.

A self-sorting system was assembled using two different types of linear hydrogen-bond arrays and even though many undesired complexes could be formed this system was found to have high fidelity. As with all supramolecular systems these self-sorting systems are under thermodynamic control and the thermodynamically favoured state is generally the one that contains the most hydrogen bonds. In the case of the DAN **93**, UPy **80**, AIC **57**, UIM **17** system this was achieved by pairing the quadruple (DAN · UPy **93** · **80**) and triple (AIC · UIM **57** · **17**) hydrogen-bond arrays together. Another factor that aided the orthogonal binding in this system was the fact that AIC **57** only forms complexes with

UIM **17**, removing UIM **17** from the system so that it cannot form undesired complexes with DAN **93** or UPy **80**. Undesired interactions could be exploited by the addition of molecules in different orders to obtain a signalling cascade. The formation of the UPy dimer could be switched on and off using three different cascades, all of which had different fidelity profiles.

The signalling cascade could be made photosensitive by the addition of a photolabile group to the AIC **124** molecule. Although it was shown that partial removal of the 2-nitrobenzyl group could be achieved, illustrated by the change in shift of the UIM **17** molecule, complete removal of the protecting group was never accomplished. Due to the complexity of the resultant ^1H NMR after partial cleavage of the protecting group the photosensitive signalling cascade was never followed. In order to assemble a signalling cascade using a photolabile protecting group it seems that it must be in a less sterically hindered position, allowing the molecule to adopt any conformation that it requires for the cleavage to be possible.

Chapter 6

Thesis Summary

This thesis began by investigating UIM **17** as a conformer independent array. Thorough analysis, using both experimental and computational techniques, of UIM **17** alone and bound to its complementary partner AIC **57** showed that UIM **17** could interact through two possible conformations.¹⁵⁹ This made it an ideal tool with which to probe some of the factors that can affect the hydrogen bonding capabilities of molecules, including preorganisation using intramolecular hydrogen bonding and the remote substituent effect.

Firstly the effect that preorganisation using intramolecular hydrogen bonds was studied.¹⁵⁹ The design and synthesis of two novel molecules (HAC **61** and ANT **62**) that both present AAD array without an intramolecular hydrogen bond was described. Conformational studies of the individual molecules indicated that both HAC **61** and ANT **62** preferentially presented the desired hydrogen-bonding array in solution, whereas AIC **57** could adopt an alternative conformation which did not present the desired AAD array. It was found that although the binding affinity of AIC **57** was slightly stronger than that of HAC **61**, this was most likely due to inductive effects and not because of the increase in preorganisation. The methyl group of ANT **62**, which was appended to the end of the hydrogen-bond array, was found that have a detrimental effect on binding. These results indicated that preorganisation using intramolecular hydrogen bonds was not essential in order to obtain high binding affinities, but consideration of the lowest energy conformation and any detrimental steric interactions must be assessed.

Secondly, the UIM·AIC **17·57** heterodimer⁶⁰ was used to quantify the effects that remote substituents can have on the binding affinity of hydrogen bonded complexes.¹⁷⁷

Because both of these molecules contained phenyl rings appended to the end of the hydrogen-bond array it was possible to add electron withdrawing/donating groups as required. This allowed the effects that changing the pK_a of the hydrogen-bonding atoms had on the binding affinity of the complex, without interference from sterics or different conformer preferences. A small change in binding affinity was measured in the AIC **57** series, where a carbonyl group ‘blocked’ the hydrogen-bond array from the direct effects of the substituent. Much larger differences were observed for the UIM **17** where the substituent could directly affect the pK_a of the hydrogen-bonding atoms. It was proposed that this effect could be broadly applicable to many hydrogen-bonding motifs, making it possible to modulate the strength of interaction without having to redesign the motif.

It is important to be able to control the binding affinity of supramolecular polymers in order to obtain the desired materials properties. Ditopic UPy molecules, one with an electron donating linker **85** and the other with an electron withdrawing linker **86**, were designed in order to investigate whether differences in DP could be obtained simply by changing the nature of the substituent. It was found that changes in the size of the assembly could be measured by DOSY, showing that it is possible to alter the DP by changing the electron donating/withdrawing nature of the substituent. This could be important for applications where control over the DP is important, allowing subtle changes in the materials properties to be obtained.

Although there are examples in the literature of molecules that present ADDA arrays,⁴⁰ they either have low affinity with DAN **21** due to an undesired conformer being adopted^{152,197} or they can also assemble into strong homodimers,^{37,68,198,199} which can result in a low fidelity being obtained. UDIM **91** was designed so that it contained two intramolecular hydrogen bonds when the desired hydrogen-bond array was presented and so that none of the possible conformers were self-complementary. However, a lower than expected binding affinity of the UDIM · DAN **91** · **93** complex ($825 \pm 16 \text{ M}^{-1}$) was obtained.²⁰¹ Thorough experimental and computational techniques were used in order to explain this lack of binding affinity. It was found that, even though UDIM **91** did not self-associate, it was most likely in an undesired conformation in solution which could prevent it from interacting strongly with DAN **93**. The low binding affinity could also be due to subtle differences in shape complementarity between UDIM **91** and DAN **93**, preventing interactions using the desired quadruple hydrogen-bond array. Indeed, it was

found that UIM **17**, which can only interact through a triple hydrogen-bond array, had a stronger association constant towards DAN **93** than UDIM **91** because it did not contain an intramolecular hydrogen bond to the carbonyl oxygen, so could form stronger interactions. In order to design molecules that present ADDA arrays that interact strongly to DAN **21**, shape complementarity, steric interactions and the preferred conformation in solution must be carefully considered.

With the increased understanding of how molecules behaved in solution and the effect that some factors had on binding affinity, a self-sorting system could be assembled. Initially, molecules that interacted using entirely different hydrogen-bond arrays were designed and synthesised. It was found that many of the desired heterodimers did not form in solution, or the ^1H NMR spectra were too broad to be used in the studies. The assembly of a high fidelity self-sorting system using two different types of linear array was then investigated. It was found that when all molecules were present in equimolar amounts a high fidelity system (0.66 at 10 mM) was established, even though it was possible for many undesired complexes to be present. The high fidelity is achieved because the maximum number of hydrogen bonds can be formed only when the desired complexes are present, making it the thermodynamically favoured position of equilibrium. Another reason for the interactions being so selective is that AIC **57** only interacts with UIM **17**, removing them both from the mixture, leaving only UPy **80** to interact with DAN **93**.

The undesired complexes could be exploited in a signalling cascade. It was possible to assemble and disassemble the UPy **80** by carefully choosing the order of addition of the molecules. The cascade could be carried out in many different ways, resulting in different fidelities being established at each step. It was thought that addition of a photolabile protecting group to one of the molecules could allow it to be present in the self-sorting mixture without forming interactions. The hydrogen-bonding array would only be revealed by removal of the photolabile protecting group, allowing control over which complexes could be formed to be achieved. A photolabile protecting group was added to AIC **57**, to give AIC* **124**. Although it was shown that some of the 2-nitro benzyl group could be removed, revealing AIC **57** so that it could interact with UIM **17**, this cleavage was never achieved in quantitative yields, preventing it from being used in a photolabile signalling cascade.

In conclusion, the results represent a significant advancement in the understanding of the behaviour of hydrogen-bonding motifs. A high fidelity self-sorting system was assembled using extremely similar molecules, which had the potential to form strong interactions with undesired motifs. These undesired interactions were exploited in a signalling cascade, where the formation of the UPy **80** homodimer is switched on and off.

6.1 Future Directions

Several aspects of the research described in this thesis are ongoing within the Wilson group. For example, assembly of the ditopic supramolecular polymers that use the remote substituent effect (described in Chapter 3) to control polymerisation will be investigated further using other techniques (e.g. rheology).

The self-sorting system described in Chapter 5 presents an area for significant future research. The four molecules described could be appended to polymers, allowing a ABC triple polymer blend to be made (Figure 6.1). This could enable the development of many new materials, allowing new properties to be discovered.

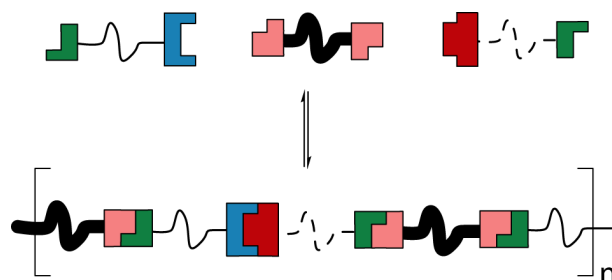


Figure 6.1: A supramolecular polymer that blends three different polymers could be synthesised using a four component self-sorting system.

Another area for future research is into the use of the signalling cascade, where the formation of the UPy **80** homodimer could be switched on and off. If the UPy molecule could be developed so that an output signal is released on binding then this could be switched on and off by the addition of a new molecule (Figure 6.2).

Further research is required into different photolabile groups that could be used in order to mask the hydrogen-bond array of AIC **57**. Once a group is found that can be more easily removed then a photosensitive signalling cascade can be realised.

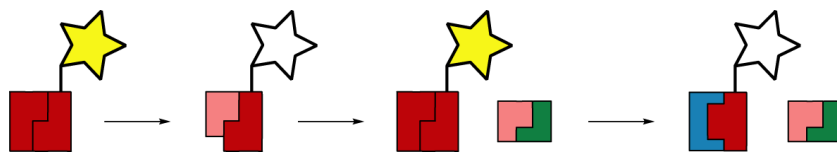


Figure 6.2: If UPy **80** were designed so that a signal was released when the homodimer was present then the signal could be turned on and off as the signalling cascade was carried out.

Chapter 7

Experimental

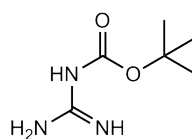
General

All reagents were purchased from Aldrich or Alfa-Aesar and used without further purification unless otherwise stated. Where anhydrous solvents were required, THF was freshly distilled from sodium benzophenone ketyl radical and CH_2Cl_2 was freshly distilled from calcium hydride, or they were obtained using a solvent purification system from Innovative Technology Inc. PureSolv[®]. CHCl_3 was freshly distilled from calcium chloride under a nitrogen atmosphere. Anhydrous DMF was obtained “sure-sealed” from Sigma-Aldrich. Triethylamine was distilled from calcium hydride and stored, under nitrogen, over potassium hydroxide pellets. All non-aqueous reactions were carried out under a nitrogen atmosphere. Analytical thin layer chromatography (TLC) was conducted using Merck Kieslegel 0.25 mm silica gel pre-coated aluminium plates with fluorescent indicator active at UV_{245} . Purification by column chromatography was carried out using Merck Kieselgel 60 silica gel. NMR spectra were obtained using Bruker DMX500 or Bruker AMD300 spectrometers operating at 500 MHz or 300 MHz for ^1H spectra and 100 MHz or 75 MHz for ^{13}C spectra as stated. Proton spectra are referenced to TMS at 0.00 ppm, and carbon spectra to CDCl_3 at 77.4 ppm, unless otherwise stated. Melting points were determined using a Griffin D5 variable temperature apparatus and are uncorrected. IR spectra were obtained using Perkin-Elmer FTIR spectrometer. Microanalysis was carried out on a Carlo Erba Elemental Analyser MOD 1106 instrument.

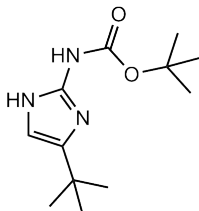
High Resolution Mass Spectra (HRMS) were recorded on a Micromass GCT Premier using electron impact ionisation (EI) or a Bruker Daltonics micrOTOF using electrospray ionisation (ESI).

7.1 Experimental for Chapter 2: Effects of Preorganisation

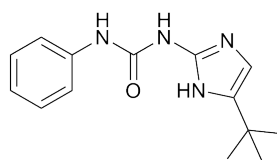
N-*tert*-Butoxycarbonylguanidine (58)



Synthesised following literature procedure.⁶⁰ Di-*tert*-butyl dicarbonate (23.7 g, 109 mmol) in acetone (250 mL) was added to a solution of guanidine hydrochloride (31.0 g, 326 mmol) and sodium hydroxide (26.0 g, 652 mmol) in water (200 mL) stirring at 0 °C. The reaction mixture was then allowed to warm to room temperature before being stirred for 20 hrs. The organics were evaporated, giving an aqueous suspension which was extracted into EtOAc (2 × 150 mL). The combined organics were dried (Na₂SO₄) and concentrated *in vacuo* before being triturated (1:2 hexane–EtOAc, 200 mL). Filtration gave *N*-*tert*-butoxycarbonylguanidine **58** (9.04 g, 52%) as a colourless powder; m.p. 165–168 °C [Lit.⁶⁰ decomposes > 350 °C] All other data is comparable to literature values (Found: C, 45.0; H, 8.20; N, 26.7; C₆H₁₃N₃O₂ requires C, 45.3; H, 8.23; N, 26.4%); *R*_f 0.00 (EtOAc); δ_H (300 MHz, DMSO-*d*₆); 1.36 (9H, s, *t*Bu); δ_C (75 MHz, DMSO-*d*₆); 163.7, 163.0, 75.8, 28.58; ν_{max}/cm⁻¹ (neat); 2971–3402, 1666, 1600; ESI-HRMS found *m/z* 160.1078 [M + H]⁺, C₆H₁₄N₃O₂ requires 160.1081.

***tert*-Butyl 4-*tert*-butyl-1*H*-imidazol-2-ylcarbamate (59)**

Synthesised following literature procedure.⁶⁰ 1-Bromopinacolone (0.56 mL, 4.19 mmol) was added to a solution of *N-tert*-butoxy carbonylguanidine **58** (2.00 g, 12.56 mmol) in dry DMF (14 mL) and was stirred at room temperature for 4 days. The reaction mixture was then filtered to give crude product. Water (30 mL) was added to the filtrate which was extracted into EtOAc (3 × 30 mL). The organics were washed with water (30 mL) and dried (Na₂SO₄) before being removed *in vacuo* to give further crude product. The combined crude material was purified by column chromatography (gradient elution: 0:1-1:4 EtOAc-CH₂Cl₂ followed by EtOAc) to give *tert*-Butyl 4-*tert*-butyl-1*H*-imidazol-2-ylcarbamate **59** (577 mg, 58%) as a colourless powder; m.p. decomposes > 173 °C [Lit.⁶⁰ decomposes > 167 °C] (Found: C, 60.0; H, 8.85; N, 17.4; C₁₂H₂₁N₃O₂ requires C, 60.2; H, 8.84; N, 17.6%); *R*_f 0.44 (1:9 MeOH-CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 6.40 (1H, s, ArCH), 1.53 (9H, s, ^{*t*}Bu), 1.12 (9H, s, ^{*t*}Bu); δ_C (75 MHz, DMSO-*d*₆); 150.2, 149.4, 148.3, 103.2, 84.4, 31.6, 29.4, 27.9; ν_{max}/cm⁻¹ (neat); 2850-3500, 1720, 1630; ESI-HRMS found *m/z* 240.1698 [M + H]⁺, C₁₂H₂₂N₃O₂ requires 240.1707.

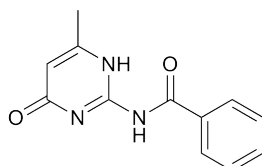
1-(5-*tert*-Butyl-1*H*-imidazol-2-yl)-3-phenylurea (17a)

Synthesised following literature procedure.⁶⁰ 2-*tert*-Butoxyamido-4-*tert*-butylimidazole (**59**) (1.40 g, 5.71 mmol) was dissolved in 1M HCl-EtOH (100 mL) before being stirred at reflux for 17 hrs. It was then evaporated to dryness *in vacuo* to give the aminoimidazole · hydrochloride **60** as a yellow oil.

Phenyl isocyanate (0.68 mL, 6.28 mmol) in THF (20 mL) was added dropwise to a solution of the aminoimidazole · hydrochloride **60** (1.00 g, 5.71 mmol) and triethylamine

(1.19 mL, 8.57 mmol) in THF (60 mL) and the reaction mixture was heated at reflux for 24 hrs before the organics were evaporated *in vacuo*. The resultant solid was dissolved in EtOAc (200 mL) and washed with water (2 × 50 mL), 10% aqueous hydrochloric acid (2 × 50 mL), saturated aqueous sodium bicarbonate (2 × 50 mL) and saturated aqueous sodium chloride (2 × 50 mL) before the organics were dried (Na₂SO₄) and removed *in vacuo*. The resultant solid was triturated (CHCl₃) and the solid removed by filtration. The filtrate was evaporated *in vacuo* and crystallised (1:1 MeOH–MeCN) to give 1-(5-*tert*-butyl-1*H*-imidazol-2-yl)-3-phenylurea **17a** (188 mg, 16%) as colourless needles; m.p. 89–93 °C [Lit.⁶⁰ decomposes > 147 °C] All other data is comparable to literature values; *R*_f 0.17 (1:19 MeOH–CH₂Cl₂); δ_H (300 MHz, CDCl₃); 7.36 (2H, d, *J* = 9.9 Hz, 2 × ArCH), 7.26 (2H, t, *J* = 7.9 Hz, 2 × ArCH), 7.04 (1H, t, *J* = 7.2 Hz, ArCH), 6.31 (1H, s, ArCH), 1.27 (9H, s, ^tBu); δ_C (75 MHz, CDCl₃); 155.1, 114.8, 138.3, 129.0, 123.3, 119.7, 30.6, 29.7; ν_{max}/cm⁻¹ (neat); 3292 (br), 2962, 1671, 1595, 1555; ESI-HRMS found *m/z* 259.1543 [M + H]⁺, C₁₄H₁₉N₄O₁ requires 259.1553.

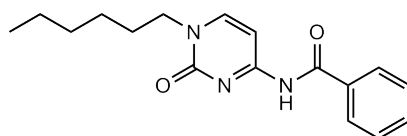
N-(6-Methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide (**57a**)



Synthesised following literature procedure.⁶⁰ Benzoic acid (500 mg, 4.1 mmol) was added to a solution of 6-methylisocytosine (427 mg, 3.4 mmol) in CHCl₃ before DMAP (625 mg, 5.1 mmol) and EDCI (1.20 g, 4.1 mmol) were added. The reaction mixture was heated at reflux for 19 hrs before being cooled and washed with 10% aqueous hydrochloric acid (2 × 50 mL), saturated aqueous sodium bicarbonate (2 × 50 mL) and saturated aqueous sodium chloride (2 × 50 mL). The organics were dried (Na₂SO₄) and evaporated *in vacuo* to give the crude material which was purified by column chromatography (gradient elution: 0:1-1:19 MeOH–CH₂Cl₂) and crystallised (CH₂Cl₂–hexane) to give *N*-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide **57a** (260 mg, 33%) as a colourless powder; m.p. 184–186 °C [Lit.⁶⁰ 189–192 °C] All other data is comparable to literature values (Found: C, 62.7; H, 4.80; N, 18.1; C₁₂H₁₁N₃O₂ requires C, 62.9; H, 4.80; N, 18.3%); *R*_f 0.35 (1:9 MeOH–CH₂Cl₂); δ_H (400 MHz, CHCl₃); 7.92 (2H, d, *J* = 7.2 Hz, 2 × ArCH), 7.65 (1H, t, *J* = 7.2 Hz, ArCH), 7.53 (2H, t, *J* = 7.2 Hz

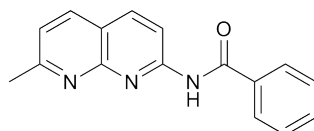
ArCH), 6.02 (1H, s, ArCH), 2.22 (3H, s, CH₃); δ_{C} (125 MHz, CDCl₃); 169.8, 162.3, 151.0, 133.6, 132.3, 128.9, 127.9 ($\times 2$), 108.4, 22.9; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3186, 1737, 1640; ESI-HRMS found m/z 230.0934 [M + H]⁺, C₁₂H₁₂N₃O₂ requires 230.0930.

N-(1-Hexyl-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (61)



Triethylamine (9.40 mL, 67.5 mmol) and benzoyl chloride (6.27 mL, 54.0 mmol) were added to a solution of cytosine (5.00 g, 45.0 mmol) and DMAP (0.55 g, 4.50 mmol) in CHCl₃ and the reaction mixture was heated to reflux for 16 hrs before being allowed to cool and filtered. The resultant solid was washed with CHCl₃ (100 mL) and water (100 mL) before being dried *in vacuo* to give benzamidocytosine **63** (10.3 g) as the crude product.

Potassium carbonate (9.60 g, 69.7 mmol) and 1-bromohexane (4.90 μL , 34.9 mmol) were added to a suspension of benzamidocytosine **63** (5 g, 23.3 mmol) in DMF (400 mL) and the mixture was heated to 80 °C for 20 hrs. The reaction mixture was then cooled to room temperature before the solid was filtered. The filtrate was diluted with EtOAc (200 mL) and washed with 10% aqueous hydrochloric acid (100 mL), water (100 mL) and saturated aqueous sodium chloride (100 mL). The organics were dried (Na₂SO₄) and evaporated *in vacuo* to give crude product which was purified by column chromatography (EtOAc) to give N-(1-hexyl-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide **61** (377 mg, 5%) as a colourless powder; m.p. 125–127 °C (Found: C, 67.9; H, 7.10; N, 14.2; C₁₇H₂₁N₃O₂ requires C, 68.2; H, 7.07; N, 14.0%); R_{f} 0.58 (EtOAc); δ_{H} (500 MHz, DMSO-*d*₆); 11.17 (1H, s, NH), 8.19 (1H, d, $J = 7.3$ Hz, ArCH), 8.01 (2H, d, $J = 8.1$ Hz, 2 \times ArCH), 7.68 (1H, t, $J = 8.1$ Hz, ArCH), 7.57 (2H, t, $J = 8.1$ Hz, 2 \times ArCH), 7.32 (1H, d, $J = 7.3$ Hz, ArCH), 3.85 (2H, t, $J = 7.3$ Hz, CH₂), 1.66 (2H, m, CH₂), 1.29 (6H, m, 3 \times CH₂), 0.91 (3H, t, $J = 7.3$ Hz, CH₃); δ_{C} (75 MHz, CDCl₃); 162.0, 148.8, 133.2, 129.1, 127.5, 96.3, 51.1, 31.3, 28.9, 26.2, 22.5, 14.0; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3381 (br), 2954, 1671, 1625; ESI-HRMS found m/z 300.1695 [M + H]⁺, C₁₇H₂₂N₃O₂ requires 300.1707.

2-Benzamido-7-methylnaphthyridine (62)

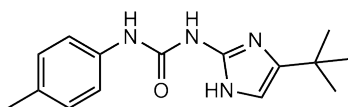
Synthesised following literature procedure.⁴³ 2,6-Diaminopyridine (2.99 g, 27.3 mmol) and phosphoric acid (35.0 g) were heated to 90 °C until melted. 4,4-Dimethoxy-2-butanone (3.75 mL, 28.1 mmol) was then added dropwise over 30 mins and the reaction mixture was then heated to 115 °C for 3 hrs. After cooling to room temperature, ammonium hydroxide was added dropwise until pH > 10. The resultant mixture was extracted with CHCl₃ (3 × 100 mL) and the combined organics were washed with saturated aqueous sodium chloride (2 × 100 mL), dried (Na₂SO₄) and concentrated *in vacuo* to give 2-amino-7-methylnaphthyridine **64** (3.32 g, 82%) as a red-brown solid.

Benzoyl chloride (2.00 mL, 17.2 mmol) was added dropwise to a solution of 2-amino-7-methylnaphthyridine **64** (2.10 g, 14.1 mmol) and NEt₃ (2.96 mL, 21.1 mmol) in CH₂Cl₂ (50 mL) and stirred at room temperature for 16 hrs. The reaction mixture was then diluted with CH₂Cl₂ (50 mL) and washed with saturated aqueous ammonium chloride (3 × 50 mL) and saturated aqueous sodium chloride (2 × 50 mL). The organics were dried (Na₂SO₄) and evaporated *in vacuo*. The resultant solid was purified by column chromatography (0:1–1:19 MeOH–EtOAc) and then crystallised (CH₂Cl₂–hexane) to give 2-benzamido-7-methylnaphthyridine **62** (1.66 g, 47%) as a colourless powder; m.p. 160–162 °C; *R*_f 0.68 (1:9 MeOH–CH₂Cl₂); δ_H (500 MHz, CDCl₃); 8.98 (1H, s, NH), 8.66 (1H, d, *J* = 8.8 Hz, ArCH), 8.21 (1H, d, *J* = 8.8 Hz, ArCH), 8.05 (1H, d, *J* = 8.2 Hz, ArCH), 7.97 (2H, d, *J* = 7.0 Hz, 2 × ArCH), 7.51–7.61 (3H, m, 3 × ArCH), 7.31 (1H, d, *J* = 8.2 Hz, ArCH), 2.79 (3H, s, CH₃); δ_C (75 MHz, DMSO-*d*₆); 166.7, 162.5, 154.5, 154.3, 139.1, 136.9, 133.9, 132.3, 128.5, 128.3, 121.6, 118.3, 115.0, 25.3; ν_{max}/cm⁻¹ (neat); 3225, 1677, 1610, 1579, 1505, 1437, 1391, 1322, 1272; ESI-HRMS found *m/z* 264.1131 [M + H]⁺, C₁₇H₁₃N₃O requires 264.1129.

7.2 Experimental for Chapter 3: The Substituent Effect

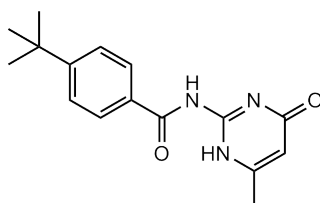
Synthesis of **17a** and **57a** described in Section 7.1.

1-(5-*tert*-Butyl-1*H*-imidazol-2-yl)-3-*p*-tolylurea (**17e**)

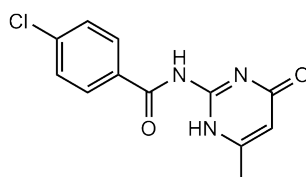


Synthesised following literature procedure.⁶⁰ 2-*tert*-Butoxyamido-4-*tert*-butylimidazole **59** (685 mg, 2.86 mmol) was dissolved in 1M HCl–EtOH (10 mL) before being stirred at reflux for 17 hrs. It was then evaporated to dryness *in vacuo* to give the aminoimidazole · hydrochloride **60** as a yellow oil.

4-Methylphenyl isocyanate (0.40 mL, 3.14 mmol) was added dropwise to a solution of the aminoimidazole · hydrochloride **60** (500 mg, 2.86 mmol) and triethylamine (0.79 mL, 5.71 mmol) in THF (50 mL) and the reaction mixture was heated at reflux for 15 hrs before the organics were evaporated *in vacuo*. The resultant solid was dissolved in EtOAc (50 mL) and washed with water (2 × 50 mL), 10% aqueous hydrochloric acid (2 × 50 mL), saturated aqueous sodium bicarbonate (2 × 50 mL) and saturated aqueous sodium chloride (2 × 50 mL) before the organics were dried (Na₂SO₄) and removed *in vacuo*. The crude product was purified by column chromatography (gradient elution: 0:1-3:7 EtOAc–hexane followed by 1:4 MeOH–EtOAc) and crystallised (1:1 MeOH/MeCN–H₂O) to give 1-(5-*tert*-Butyl-1*H*-imidazol-2-yl)-3-*p*-tolylurea **17e** (361 mg, 46%) as colourless needles; m.p. 185–188 °C; *R*_f 0.21 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, CDCl₃); 7.23 (2H, d, *J* = 8.4 Hz, 2 × ArCH), 7.07 (2H, d, *J* = 8.4 Hz, 2 × ArCH), 6.29 (1H, s, ArCH), 2.30 (3H, s, CH₃), 1.26 (9H, s, ^tBu); δ_C (75 MHz, CDCl₃); 155.5, 146.9, 144.9, 135.6, 133.2, 129.5, 121.7, 120.4, 30.6, 29.8, 20.8; ν_{max}/cm⁻¹ (neat); 3369 (br), 2963, 1667, 1598, 1552, 1515; ESI-HRMS found *m/z* 273.1713 [M + H]⁺, C₁₅H₂₁N₄O requires 273.1710.

4-*tert*-Butyl-*N*-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide (57d)

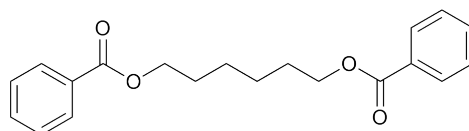
Synthesised following literature procedure.⁶⁰ 2-Amino-6-methylpyrimidin-4-ol (0.29 g, 2.34 mmol), 4-*tert*-butylbenzoic acid (0.50 g, 2.81 mmol), DMAP (0.43 g, 3.51 mmol) and EDCI (0.83 g, 2.81 mmol) were suspended in CHCl₃ (100 mL) and heated to reflux for 18 hrs. The reaction mixture was cooled to room temperature and washed with 10% aqueous hydrochloric acid (2 × 50 mL), saturated aqueous sodium bicarbonate (2 × 50 mL) and saturated aqueous sodium chloride (2 × 50 mL). The organic layer was then dried (Na₂SO₄) and evaporated *in vacuo* and the resultant solid was purified by column chromatography (1:1 EtOAc–hexane) followed by crystallisation (CH₂Cl₂–hexane) to give 4-*tert*-butyl-*N*-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide **57d** (552 mg, 83%) as a colourless flocculent solid; m.p. 238–240 °C (Found: C, 67.6; H, 6.65; N, 14.7; C₁₆H₁₉N₃O₂ requires C, 67.4; H, 6.71; N, 14.7%); *R*_f 0.33 (EtOAc); δ_H (400 MHz, CDCl₃); 7.84 (2H, d, *J* = 8.6 Hz, 2 × ArCH), 7.54 (2H, d, *J* = 8.6 Hz, 2 × ArCH), 6.02 (1H, s, ArCH), 2.22 (3H, s, CH₃), 1.36 (9H, s, ^tBu); δ_C (75 MHz, CDCl₃); 162.3, 158.1, 151.0, 129.5, 128.1, 126.4, 108.9, 35.6, 31.4, 23.6; ν_{max}/cm⁻¹ (neat); 3435 (br), 2079, 1650; ESI-HRMS found *m/z* 286.1546 [M + H]⁺, C₁₆H₂₀N₃O₂ requires 286.1550.

4-Chloro-*N*-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide (57e)

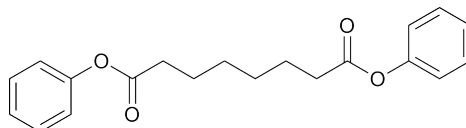
Synthesised following literature procedure.⁶⁰ 2-Amino-6-methylpyrimidin-4-ol (0.33 g, 2.65 mmol), 4-chlorobenzoic acid (0.50 g, 3.18 mmol), DMAP (0.49 g, 3.98 mmol) and EDCI (0.95 g, 3.18 mmol) were suspended in CHCl₃ (70 mL) and heated to reflux for 16 hrs. The reaction mixture was then allowed to cool to room temperature before being washed with 10% aqueous hydrochloric acid (2 × 50 mL), saturated aqueous sodium

bicarbonate (2×50 mL) and saturated aqueous sodium chloride (2×50 mL). The organic layer was then dried (Na_2SO_4) and evaporated *in vacuo*. The crude product was purified by column chromatography (1:9 MeOH– CH_2Cl_2) followed by crystallisation (CH_2Cl_2 –MeOH) to give *4-chloro-N-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)benzamide* **57e** (278 mg, 40%) as colourless plates; m.p. 245–248 °C (Found: C, 54.9; H, 3.75; N, 16.0; Cl, 12.3 $\text{C}_{12}\text{H}_{10}\text{ClN}_3\text{O}_3$ requires C, 54.7; H, 3.82; N, 16.0; Cl, 12.1%); R_f 0.61 (1:9 MeOH– CH_2Cl_2); δ_{H} (300 MHz, DMSO- d_6); 8.08 (2H, d, $J=8.6$ Hz, ArCH), 7.59 (2H, d, $J=8.6$ Hz, ArCH), 5.92 (1H, s, ArCH), 2.20 (3H, s, CH_3); δ_{C} (75 MHz, DMSO- d_6); 161.1, 153.0, 137.9, 133.4, 130.9, 128.8, 106.1, 31.0, 21.4; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3435 (br), 2078, 1638; ESI-HRMS found m/z 264.0524 $[\text{M} + \text{H}]^+$, $\text{C}_{12}\text{H}_{11}\text{ClN}_3\text{O}_2$ requires 264.0534. Found m/z 286.0354 $[\text{M} + \text{Na}]^+$, $\text{C}_{12}\text{H}_{10}\text{ClN}_3\text{NaO}_2$ requires 286.0354.

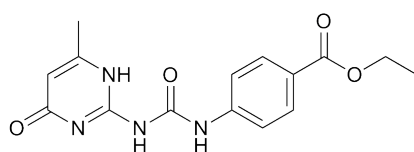
Hexane-1,6-diyl dibenzoate (**72**)



Benzoyl chloride (500 μL , 4.29 mmol) was added dropwise to a solution of 1,6-hexanediol (230 mg, 1.95 mmol) and triethylamine (543 μL , 3.90 mmol) in THF (10 mL) and the reaction mixture was stirred at room temperature for 18 hrs. The solvents were then evaporated and the crude product was purified by column chromatography (gradient elution: 1:0-0:1 hexane– CH_2Cl_2) to give hexane-1,6-diyl dibenzoate **72** (438 mg, 69%) as a colourless powder; m.p. 44–46 °C (Found: C, 73.5; H, 6.80; $\text{C}_{20}\text{H}_{22}\text{O}_4$ requires C, 73.6; H, 6.79%); R_f 0.34 (CH_2Cl_2); δ_{H} (300 MHz, CDCl_3); 8.04 (4H, d, $J=7.9$ Hz, $4 \times$ ArCH), 7.55 (4H, t, $J=7.9$ Hz, $2 \times$ ArCH), 7.43 (4H, t, $J=7.9$ Hz, $4 \times$ ArCH), 4.34 (4H, t, $J=6.6$ Hz, $2 \times$ CH_2), 1.82 (4H, m, $2 \times$ CH_2), 1.53 (4H, m, $2 \times$ CH_2); δ_{C} (75 MHz, CDCl_3); 167.1, 138.2, 130.9, 129.9, 128.7, 65.3, 29.1, 26.2; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3064, 2979, 2936, 2868, 1715; ESI-HRMS found m/z 327.1602 $[\text{M} + \text{H}]^+$, $\text{C}_{20}\text{H}_{23}\text{O}_4$ requires 327.1591.

Diphenyl octanedioate (73)

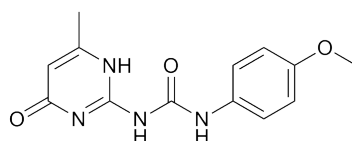
Suberoyl chloride (435 μL , 2.40 mmol) was added to a refluxing solution of phenol (500 mg, 5.30 mmol) and triethylamine (673 μL , 4.84 mmol) in THF (10 mL). The reaction mixture was stirred at reflux for 15 hrs before being cooled and concentrated *in vacuo*. The resultant solid was triturated (CH_2Cl_2), filtered and the filtrate concentrated to give the crude product which was purified by column chromatography (CH_2Cl_2) to give diphenyl octanedioate **73** (545 mg, 70%) as a colourless powder; m.p. 51–54 °C [Lit.²³⁴ 73 °C] All other data is comparable to literature values (Found: C, 73.4; H, 6.80; $\text{C}_{20}\text{H}_{22}\text{O}_4$ requires C, 73.6; H, 6.79%); R_f 0.61 (CH_2Cl_2); δ_{H} (500 MHz, CDCl_3); 7.30 (4H, t, $J=7.9$ Hz, $4 \times \text{ArCH}$), 7.14 (2H, t, $J=7.9$ Hz, $2 \times \text{ArCH}$), 7.00 (4H, d, $J=7.9$ Hz, $4 \times \text{ArCH}$), 2.50 (4H, t, $J=7.5$ Hz, $2 \times \text{CH}_2$), 1.72 (4H, m, $2 \times \text{CH}_2$), 1.42 (4H, m, $2 \times \text{CH}_2$); δ_{C} (75 MHz, CDCl_3); 172.2, 150.7, 129.5, 125.8, 121.6, 34.3, 28.8, 24.8; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 2937, 2913, 2970, 2862, 1758; ESI-HRMS found m/z 349.1407 $[\text{M} + \text{Na}]^+$, $\text{C}_{20}\text{H}_{25}\text{NaO}_4$ requires 349.1410.

Ethyl 4-(3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)ureido)benzoate (74)

Ethyl 4-isocyanatobenzoate (1.68 g, 8.8 mmol) was added to a solution of 2-amino-4-hydroxy-6-methylpyrimidine (1.00 g, 8.0 mmol) and NEt_3 (1.11 mL, 8.0 mmol) in THF (50 mL) and the reaction mixture was heated to reflux for 17 hrs. After cooling to room temperature the reaction mixture was filtered and the solid triturated with EtOAc. The suspension was filtered to give *ethyl 4-(3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)ureido)benzoate* **74** (1.94 g, 77%) as a colourless solid; m.p. 274–276 °C; R_f 0.00 (1:4 MeOH– CH_2Cl_2); δ_{H} (300 MHz, $\text{DMSO}-d_6$); 7.92 (2H, d, $J=9.0$ Hz, $2 \times \text{ArCH}$), 7.70 (2H, d, $J=9.0$ Hz, $2 \times \text{ArCH}$), 5.76 (1H, s, ArCH), 4.29 (2H, q, $J=6.5$ Hz, CH_2), 2.20 (3H, s,

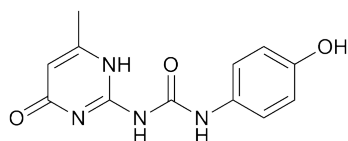
CH₃), 1.31 (3H, t, $J=6.5$ Hz, CH₃); Molecule was too insoluble to obtain δ_C ; $\nu_{\max}/\text{cm}^{-1}$ (neat); 3410 (br), 1698, 1660; Molecule was too insoluble to obtain MS.

1-(4-Methoxyphenyl)-3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)urea (75)



2-Amino-6-methylpyrimidin-4-ol (1.00 g, 8.0 mmol) was added to a solution of triethylamine (1.11 mL, 8.0 mmol) in THF (50 mL). The reaction mixture was heated to reflux before 4-methoxy phenylisocyanate (1.14 mL, 8.8 mmol) was added and the reaction mixture was heated at reflux for a further 18 hrs before being cooled to room temperature and filtered. The resultant crude solid was purified by repeated sonication in MeOH (200 mL) followed by filtration to give the 1-(4-methoxyphenyl)-3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)urea **75** (1.56 g, 71%) as a cream coloured solid; m.p. 275–276 °C (Found: C, 56.7; H, 5.05; N, 20.4; C₁₃H₁₄N₄O₃ requires C, 56.9; H, 5.14; N, 20.4%); R_f 0.57 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 9.82 (1H, s, NH), 7.41 (2H, d, $J=8.8$ Hz, 2 × ArCH), 6.91 (2H, d, $J=8.8$ Hz, 2 × ArCH), 5.82 (1H, s, ArCH), 3.73 (3H, s, OCH₃), 2.17 (3H, s, CH₃); Molecule was too insoluble to obtain δ_C ; $\nu_{\max}/\text{cm}^{-1}$ (neat); 2987 (br), 1702, 1652, 1261; ESI-HRMS found m/z 275.1142 [M + H]⁺, C₁₃H₁₅N₄O₃ requires 275.1139.

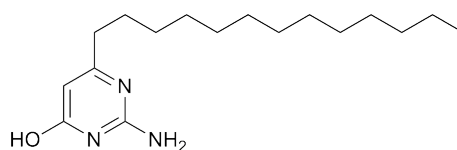
1-(4-Hydroxyphenyl)-3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)urea (77)



Synthesised following literature procedure.¹⁹⁵ Boron tribromide (1M in CH₂Cl₂, 2.19 mL, 2.19 mmol) was added dropwise to a suspension of the methoxy ureidopyrimidine **75** (200 mg, 0.73 mmol) in CH₂Cl₂ (10 mL) stirring at -78 °C. The reaction mixture was then allowed to warm to room temperature and stirred for 19 hrs before being quenched

with H₂O (20 mL) and stirred for a further 2 hrs. The reaction mixture was filtered to give 1-(4-hydroxyphenyl)-3-(6-methyl-4-oxo-1,4-dihydropyrimidin-2-yl)urea **77** (176 mg, 93%) as a colourless powder; m.p. 112–115 °C; *R*_f 0.36 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 9.67 (1H, s, OH), 7.27 (2H, d, *J* = 8.8 Hz, 2 × ArCH), 6.73 (2H, d, *J* = 8.8 Hz, 2 × ArCH), 5.86 (1H, s, ArCH), 2.18 (3H, s, CH₃); Molecule was too insoluble to obtain δ_C; ν_{max}/cm⁻¹ (neat); 3655-2337, 1710, 1646, 1621, 1514, 1227; Molecule was too insoluble to obtain MS.

2-Amino-6-tridecylpyrimidin-4-ol (**78**)

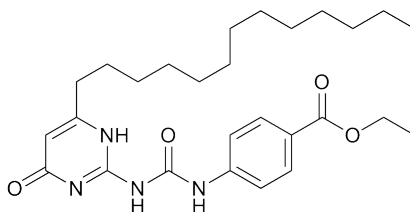


Synthesised following literature procedure.¹⁹³ Triethylamine (50.0 mL, 360 mmol) was added to a suspension of potassium ethyl malonate (40.8 g, 240 mmol) in MeCN (250 mL) that was stirring at 0 °C. The reaction mixture was stirred for 15 mins before magnesium chloride (28.5 g, 300 mmol) in MeCN (80 mL) was added and allowed to warm to 10 °C. After stirring for 2 hrs the reaction mixture was cooled to 0 °C before myristoyl chloride (32.6 mL, 120 mmol) was added dropwise. The reaction mixture was warmed to room temperature and stirred for 16 hrs before the solvents were evaporated *in vacuo*. The resultant solid was dissolved in Et₂O (350 mL) and washed with 30% aqueous hydrochloric acid (250 mL) and saturated aqueous sodium bicarbonate (100 mL) before the organics were dried (MgSO₄) and evaporated *in vacuo*. The resultant solid was purified by column chromatography (CHCl₃) to give ethyl 3-oxohexadecanoate **79** (26.9 g, 75%) as a colourless powder.

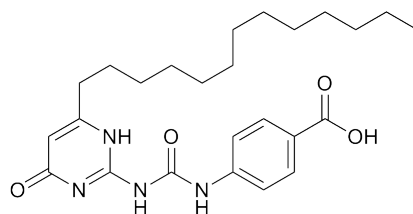
Guanidinium carbonate (12.8 g, 142.6 mmol) was added to a solution of **79** (25.0 g, 83.9 mmol) and potassium *tert*-butoxide (9.4 g, 83.9 mmol) in EtOH (300 mL) and the reaction mixture was heated to reflux for 3 days. After cooling the reaction mixture was filtered and the filtrate was evaporated *in vacuo*. The resultant solid was dissolved in H₂O (500 mL) and the solution was acidified to pH 6 (acetic acid). The resultant suspension was filtered and the solid washed with acetone and Et₂O before being crystallised (propan-2-ol) to give the 2-amino-6-tridecylpyrimidin-4-ol **78** as a cream coloured powder; m.p. 162–165 °C; *R*_f 0.19 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 6.58 (2H,

s, NH₂), 5.33 (1H, s, ArCH), 2.20 (2H, t, $J=6.0$ Hz, CH₂), 1.51 (2H, m, CH₂), 1.23 (20H, m, 10 × CH₂), 0.85 (3H, t, $J=6.0$ Hz, CH₃); Molecule was too insoluble to obtain δ_{C} ; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3360, 3144, 2919, 2850, 2679 (br), 1637, 1468, 1401; ESI-HRMS found m/z 294.2529 [M + H]⁺, C₁₇H₃₂N₃O requires 294.2540.

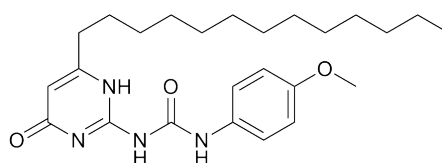
Ethyl 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoate (80)



Ethyl-4-isocyanatobenzoate (0.71 g, 3.75 mmol) was added to a refluxing solution of the amine **78** (1.00 g, 3.41 mmol) and triethylamine (0.48 mL, 3.75 mmol) in THF (50 mL). The reaction mixture was then stirred at reflux for 18 hrs before being allowed to cool and the volatiles evaporated *in vacuo*. The resultant solid was suspended in EtOAc (50 mL) and H₂O (50 mL) was added. The suspension was filtered and sonicated in EtOAc (50 mL) before being filtered and washed (Et₂O) to give ethyl 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoate **80** (1.22 g, 74%) as a colourless powder; m.p. 218–220 °C; R_f 0.38 (1:99 MeOH–CH₂Cl₂); δ_{H} (300 MHz, CDCl₃); 12.95 (1H, s, NH), 12.52 (1H, s, NH), 12.32 (1H, s, NH), 8.07 (2H, d, $J=8.7$ Hz, ArCH), 7.86 (2H, d, $J=8.7$ Hz, ArCH), 5.94 (1H, s, ArCH), 4.41 (2H, q, $J=7.2$ Hz, CH₂), 2.47 (2H, t, $J=7.7$ Hz, CH₂), 1.65 (4H, m, 2 × CH₂), 1.44 (3H, t, $J=7.2$ Hz, CH₃), 1.32 (18H, m, 9 × CH₂), 0.92 (3H, t, $J=6.6$ Hz, CH₃); δ_{C} (75 MHz, CDCl₃); 172.8, 166.2, 154.4, 154.3, 152.9, 142.8, 130.5, 125.3, 119.3, 105.9, 60.8, 32.4, 31.9, 29.6–29.2 (× 7), 28.9, 26.3, 22.7, 14.4, 14.1; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3466–2850, 1698, 1650; Molecule was too insoluble to obtain MS.

4-(3-(4-Oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoic acid (**82**)

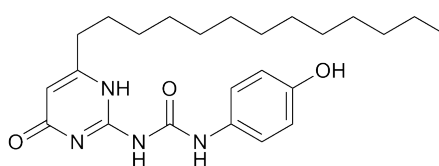
Synthesised following literature procedure.¹⁹⁴ 1 M Aqueous sodium hydroxide (4.12 mL, 4.12 mmol) was added to a suspension of the ester ureidopyrimidine **80** (500 mg, 1.03 mmol) in EtOH (50 mL) and the reaction mixture was heated to 70 °C for 18 hrs. The solvent was then evaporated *in vacuo* and the resultant solid was dissolved in H₂O (25 mL) and acidified with 10% aqueous hydrochloric acid. The reaction mixture was then filtered to give 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoic acid **82** (457 mg, 97%) as a colourless solid; m.p. 256–257 °C; *R*_f 0.39 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 7.91 (2H, d, *J* = 9.0 Hz, 2 × ArCH), 7.64 (2H, d, *J* = 9.0 Hz, 2 × ArCH), 5.84 (1H, s, ArCH), 2.44 (2H, t, *J* = 7.5 Hz, CH₂), 1.59 (2H, m, CH₂), 1.21 (20H, m, 10 × CH₂), 0.83 (3H, t, *J* = 6.0 Hz, CH₃); Molecule was too insoluble to obtain δ_C; ν_{max}/cm⁻¹ (neat); 3200–2343, 1689, 1656; ESI-HRMS found *m/z* 438.2747 [M–H₂O]⁻, C₂₅H₃₄N₄O₃ requires 438.2631.

1-(4-Methoxyphenyl)-3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)urea (**81**)

4-Methoxyphenylisocyanate (0.24 mL, 1.90 mmol) was added to a refluxing mixture of the amine **78** (0.50 g, 1.71 mmol) and triethylamine (0.24 mL, 1.71 mmol) in THF (20 mL). The reaction mixture was then stirred at reflux for 18 hrs before being allowed to cool and the volatiles evaporated *in vacuo*. The resultant solid was triturated in hot CHCl₃ (50 mL), filtered and the filtrate evaporated to give the crude product which was crystallised (CHCl₃–MeOH) to give 1-(4-methoxyphenyl)-3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)urea **81** (0.55 g, 73%) as colourless needles; m.p. 121–123 °C (Found: C, 67.7; H, 8.65; N, 12.7; C₂₅H₃₈N₄O₃ requires C, 67.8; H, 8.65; N, 12.7%);

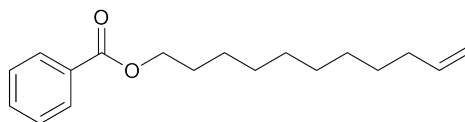
R_f 0.57 (1:1 EtOAc–hexane); δ_H (300 MHz, $CDCl_3$); 13.12 (1H, s, NH), 12.27 (1H, s, NH), 12.10 (1H, s, NH), 7.61 (2H, d, $J=9.1$ Hz, ArCH), 6.93 (2H, d, $J=9.1$ Hz, ArCH), 5.94 (1H, s, ArCH), 3.84 (3H, s, OCH_3), 2.50 (2H, t, $J=7.5$ Hz, CH_2), 1.68 (2H, m, CH_2), 1.60 (4H, m, $2 \times CH_2$), 1.32 (16H, m, $8 \times CH_2$), 0.92 (3H, t, $J=6.8$ Hz, CH_3); δ_C (75 MHz, $CDCl_3$); 173.1, 156.2, 154.7, 154.6, 152.7, 131.2, 122.4, 114.1, 105.9, 55.5, 32.6, 31.9, 29.7–29.2 ($\times 7$), 28.9, 26.7, 22.7, 14.2; ν_{max}/cm^{-1} (neat); 2916 (br), 1707, 1651, 1574, 1509; ESI-HRMS found m/z 443.3017 $[M + H]^+$, $C_{25}H_{39}N_4O_3$ requires 443.3017.

1-(4-Hydroxyphenyl)-3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)urea (83)



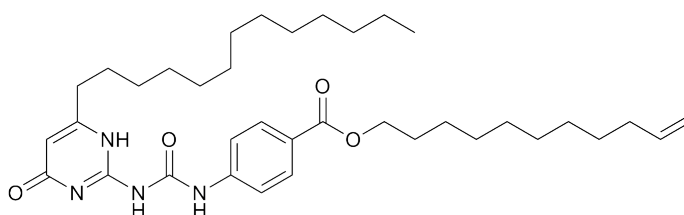
Synthesised following literature procedure.¹⁹⁵ BBr_3 (3.4 mL, 1 M in CH_2Cl_2 , 3.39 mmol) was added dropwise to a suspension of **81** (500 mg, 1.13 mmol) in CH_2Cl_2 (10 mL) cooled to $-78^\circ C$. The reaction mixture was then allowed to warm to room temperature and was stirred for 20 hrs. H_2O (10 mL) was then added dropwise before the reaction mixture was filtered and the resultant solid was triturated with $CHCl_3$ to give 1-(4-hydroxyphenyl)-3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)urea **83** (448 mg, 93%) as a colourless solid; m.p. 175 – $176^\circ C$; R_f 0.39 (1:9 MeOH– CH_2Cl_2); δ_H (300 MHz, $DMSO-d_6$); 9.80 (1H, s, OH), 7.25 (2H, d, $J=8.7$ Hz, $2 \times$ ArCH), 6.73 (2H, d, $J=8.7$ Hz, $2 \times$ ArCH), 5.84 (1H, s, ArCH), 2.43 (2H, t, $J=7.4$ Hz, CH_2), 1.59 (2H, m, CH_2), 1.22 (20H, m, $10 \times CH_2$), 0.85 (3H, t, $J=6.6$ Hz, CH_3); δ_C (75 MHz, $DMSO-d_6$); 163.6, 163.4, 154.1, 153.5, 152.1, 129.5, 121.7, 115.7, 104.2, 36.1, 31.6, 29.4–29.1 ($\times 7$), 28.7, 27.6, 22.4, 14.3; ν_{max}/cm^{-1} (neat); 3234–2853 (br), 1695, 1646, 1627, 1562, 1505, 1447, 1222, 834; Molecule was too insoluble to obtain MS.

Undec-10-enyl benzoate (**84**)



Benzoic acid (1.00 g, 8.20 mmol), DMAP (1.36 g, 11.18 mmol), 10-undecen-1-ol (1.49 mL, 7.45 mmol) and EDCI (1.72 g, 8.94 mmol) were all dissolved in CHCl_3 (50 mL) and the reaction mixture was heated to reflux for 16 hrs. After being allowed to cool to room temperature the solvent was evaporated *in vacuo* and the crude product was purified by column chromatography (CH_2Cl_2) to give undec-10-enyl benzoate **84** (1.88 g, 92%) as a yellow coloured oil; (Found: C, 78.9; H, 9.80; $\text{C}_{18}\text{H}_{26}\text{O}_2$ requires C, 78.8; H, 9.60%); R_f 0.41 (1:1 CH_2Cl_2 -hexane); δ_{H} (300 MHz, CDCl_3); 7.98 (2H, d, $J = 7.5$ Hz, $2 \times \text{ArCH}$), 7.49 (1H, t, $J = 7.5$ Hz, ArCH), 7.37 (2H, t, $J = 7.5$ Hz, $2 \times \text{ArCH}$), 5.75 (1H, m, alkene-CH), 4.90 (2H, m, $2 \times$ alkene-CH), 4.24 (2H, t, $J = 7.5$ Hz, CH_2), 1.96 (2H, m, CH_2), 1.70 (2H, m, CH_2), 1.37-1.18 (12H, m, $6 \times \text{CH}_2$); δ_{C} (75 MHz, CDCl_3); 166.7, 139.2, 132.8, 130.5, 129.5, 128.3, 114.1, 65.2, 33.8, 29.5, 29.4, 29.3, 29.1, 28.9, 28.7, 26.0; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3074, 2927, 2855, 1727, 1274, 1114, 711; ESI-HRMS found m/z 297.1823 $[\text{M} + \text{Na}]^+$, $\text{C}_{18}\text{H}_{26}\text{NaO}_2$ requires 297.1825.

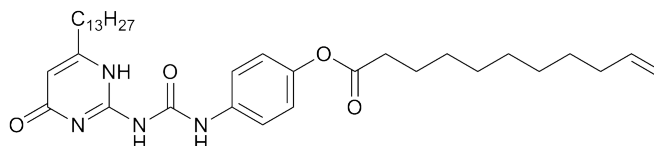
Undec-10-enyl 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido) benzoate (**87**)



10-Undecen-1-ol (358 μL , 1.79 mmol) was added to a solution of acid **82** (900 mg, 1.97 mmol), DMAP (328 mg, 2.69 mmol) and EDCI (412 mg, 2.15 mmol) in CHCl_3 (50 mL) and the reaction mixture was heated to reflux for 3 days. It was then allowed to cool to room temperature before the solvent was evaporated *in vacuo* and the resultant solid was purified by column chromatography (gradient elution: 0:1-1:19 MeOH- CH_2Cl_2) to give the undec-10-enyl 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido) benzoate **87** (740 mg, 68%) as a colourless powder; m.p. 122-125 $^\circ\text{C}$; R_f 0.51 (1:19 Et_2O - CHCl_3);

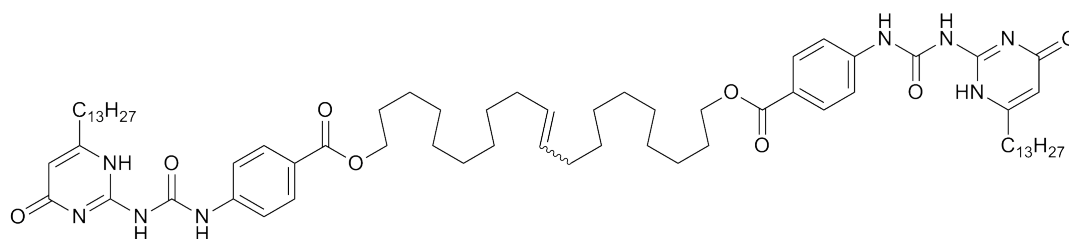
δ_{H} (500 MHz, CHCl_3); 12.96 (1H, s, NH), 12.48 (1H, s, NH), 12.33 (1H, s, NH), 8.03 (2H, d, $J=8.3$ Hz, ArCH), 7.82 (2H, d, $J=8.3$ Hz, ArCH), 5.96 (1H, s, ArCH), 5.82 (1H, ddt, $J=11.0, 17.0, 6.6$ Hz, CH), 4.97 (2H, ddd, $J=11.0, 17.0, 32.3$ Hz, CH_2), 4.30 (2H, t, $J=6.6$ Hz, CH_2), 2.52 (2H, t, $J=7.2$ Hz, CH_2), 2.05 (2H, q, $J=7.2$ Hz, CH_2), 1.79–1.26 (36H, m, alkyl), 0.87 (3H, t, $J=6.9$ Hz, CH_3); δ_{C} (75 MHz, CDCl_3); 172.9, 166.3, 154.4, 154.3, 153.0, 142.8, 139.2, 130.6, 125.4, 119.4, 114.2, 106.0, 65.0, 33.9, 32.5, 32.0, 29.7–28.8 ($\times 14$), 26.4, 26.1, 22.7, 14.2; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3454, 2922, 2850, 1714, 1696, 1660; ESI-MS found m/z 609.4 $[\text{M} + \text{H}]^+$, $\text{C}_{36}\text{H}_{56}\text{H}_4\text{O}_4$ requires 609.4.

4-(3-(4-Oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)phenyl undec-10-enoate (88)



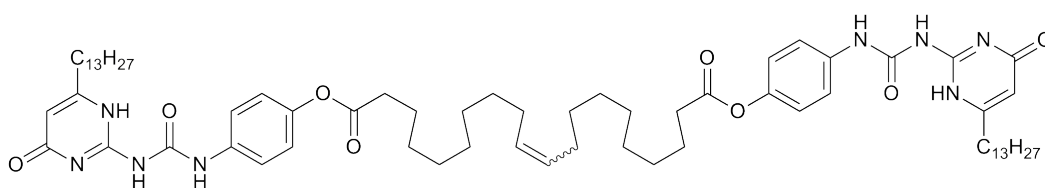
10-Undecenoic acid (473 mg, 2.57 mmol) was added to a solution of **83** (1.00 g, 2.34 mmol), DMAP (428 mg, 3.50 mmol) and EDCI (538 mg, 2.80 mmol) in CHCl_3 (50 mL) and the reaction mixture was heated to reflux for 15 hrs before being allowed to cool to room temperature. The volatiles were then removed *in vacuo* and the resultant solid was purified by column chromatography (gradient elution: 0:1–1:9 MeOH– CH_2Cl_2) and crystallisation (CHCl_3 –MeOH) to give the 4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)phenyl undec-10-enoate **88** (758 mg, 54%) as a colourless powder; m.p. 128–130 °C (Found: C, 70.5; H, 9.15; N, 9.5; $\text{C}_{35}\text{H}_{54}\text{N}_4\text{O}_4$ requires C, 70.7; H, 9.15; N, 9.4%); R_f 0.28 (1:4 EtOAc–hexane); δ_{H} (300 MHz, CDCl_3); 12.87 (1H, s, NH), 12.22 (1H, s, NH), 12.11 (1H, s, NH), 7.66 (2H, d, $J=9.0$ Hz, $2 \times$ ArCH), 6.97 (2H, d, $J=9.0$ Hz, $2 \times$ ArCH), 5.75 (1H, m, CH), 5.73 (1H, s, ArCH), 4.91 (2H, dd, $J=15.3, 22.2$ Hz, CH_2), 2.47 (2H, t, $J=7.4$ Hz, CH_2), 2.24 (2H, t, $J=7.5$ Hz, CH_2), 1.98 (2H, q, $J=6.9$ Hz, CH_2), 1.68 (2H, m, CH_2), 1.48 (2H, m, CH_2), 1.32–1.19 (30H, m, $15 \times \text{CH}_2$), 0.81 (3H, t, $J=6.6$ Hz, CH_3); δ_{C} (75 MHz, CDCl_3); 173.0, 172.2, 154.5 ($\times 2$), 152.9, 146.7, 139.2, 135.9, 121.8, 121.2, 114.2, 106.0, 34.4, 33.8, 32.5, 31.9, 29.7–28.8 ($\times 13$), 26.7, 24.9, 22.7, 14.1; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3143, 2924, 281, 1749, 1697, 1580, 1505, 1201; ESI-HRMS found m/z 617.4018 $[\text{M} + \text{Na}]^+$, $\text{C}_{35}\text{H}_{54}\text{N}_4\text{NaO}_4$ requires 617.4037.

(*E,Z*)-Icos-10-ene-1,20-diyl bis(4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoate) (85)



Grubbs (I) (10 mg, 0.01 mmol) in degassed CH_2Cl_2 (2 mL) was added dropwise to a solution of **87** (150 mg, 0.25 mmol) in degassed CH_2Cl_2 (10 mL) and the reaction mixture was stirred at room temperature for 2 days. It was then filtered through celite before the solvent was evaporated and the resultant solid was crystallised (CHCl_3 -MeOH) to give (*E,Z*)-icos-10-ene-1,20-diyl bis(4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)benzoate) **85** (125 mg, 84%) as an off white powder; δ_{H} (300 MHz, CDCl_3); 12.59 (2H, br, NH) 12.40 (2H, br, NH), 11.90 (2H, br, NH), 7.91 (4H, br, ArCH), 7.75 (4H, br, ArCH) 5.60 (0.5H, br, CH), 5.45 (0.5H, br, CH), 5.35 (1H, br, CH), 4.18 (4H, br, CH_2), 1.92 (4H, br, CH_2), 1.68 (4H, br, CH_2), 1.20 (72H, br, CH_2), 0.85 (6H, br, CH_3); Molecule too insoluble to carry out MS.

(*E,Z*)-Bis(4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)phenyl) icos-10-enedioate (86)



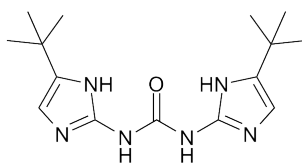
Grubbs (I) (10 mg, 0.01 mmol) in degassed CH_2Cl_2 (1 mL) was added dropwise to a solution of **87** (150 mg, 0.25 mmol) in degassed CH_2Cl_2 (10 mL) and the reaction mixture was stirred at room temperature for 2 days. It was then filtered through celite before the solvent was evaporated and the resultant solid was crystallised (CHCl_3 -MeOH) to give (*E,Z*)-bis(4-(3-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)ureido)phenyl) icos-10-enedioate **86** (140 mg, 93%) as an off white powder; δ_{H} (300 MHz, CDCl_3); 12.80 (2H, br, NH), 12.30 (2H, br, NH), 11.95 (2H, br, NH), 7.70 (4H, br, ArCH), 6.95 (4H, br, ArCH), 5.60

(2H, br, CH), 5.42 (2H, br, CH), 2.60 (4H, br, CH₂), 1.90 (4H, br, CH₂), 1.70 (4H, br, CH₂), 1.20 (68H, br, CH₂), 0.75 (6H, br, CH₃); Molecule too insoluble to carry out MS.

7.3 Experimental for Chapter 4: A Novel ADDA array

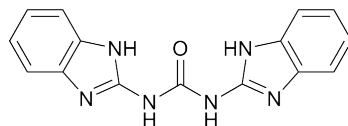
Synthesis of **17** described in Section 7.1.

1,3-Bis(5-*tert*-butyl-1*H*-imidazol-2-yl)urea (**91**)

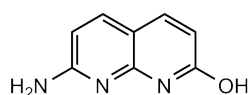


2-*tert*-Butoxyamido-4-*tert*-butylimidazole **59** (1.36 g, 5.68 mmol) was dissolved in 1M HCl–EtOH (100 mL) before being stirred at reflux for 17 hrs. It was then evaporated to dryness *in vacuo* to give the aminoimidazole · hydrochloride **60** as a yellow oil.

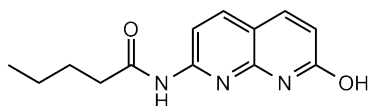
Triethylamine (1.98 mL, 14.20 mmol) was added to a solution of the aminoimidazole · hydrochloride **60** (1.00 g, 5.68 mmol) in THF (100 mL) and the reaction mixture was stirred at room temperature for 2 hrs. It was then heated to reflux and carbonyl diimidazole (0.55 g, 3.41 mmol) was added. The reaction mixture was then heated at reflux for a further 20 hrs before being cooled to room temperature and filtered. The filtrate was evaporated *in vacuo* and the resultant solid was purified by column chromatography (gradient elution: 0:1-1:19 MeOH–CH₂Cl₂) followed by trituration (MeOH) to give the 1,3-bis(5-*tert*-butyl-1*H*-imidazol-2-yl)urea **91** (150 mg, 17%) as a colourless solid; m.p. decomposes > 280 °C (Found: C, 59.1; H, 7.95; N, 27.7; C₁₅H₂₄N₆O requires C, 59.2; H, 7.95; N, 27.6%); *R*_f 0.45 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, CDCl₃); 15.20 (1H, a, NH), 12.40 (1H, s, NH), 11.16 (1H, s, NH), 10.96 (1H, s, NH), 6.30 and 6.26 (2H, 2 × s, 2 × ArCH), 1.33 (18H, s, ^tBu); δ_C (125 MHz, CDCl₃); 162.9, 162.8 (× 2), 162.7 (× 2), 149.9, 149.1 (× 2), 149.0, 148.1, 141.3 (× 2), 140.1 (× 2), 136.7, 134.7, 111.8, 110.4 (× 2), 103.2, 30.9 (× 3), 30.8, 30.2 (× 2), 29.8 (× 2), 29.7; ν_{max}/cm⁻¹ (neat); 3372-2251 (br), 2140, 1916, 1674, 1538, 1480, 1446, 1397, 1173, 1073, 1037, 851, 808; ESI-HRMS found *m/z* 305.2087 [M + H]⁺, C₁₅H₂₅N₆O requires 305.2084.

1,3-Di(1*H*-benzo[*d*]imidazol-2-yl)urea (128)

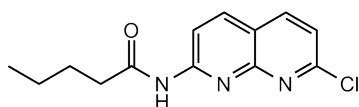
2-Aminobenzimidazole (100 mg, 0.75 mmol) was added to a solution of triethylamine (262 μ L, 1.88 mmol) in THF (10 mL). The reaction mixture was then heated to reflux and carbonyl diimidazole (73 mg, 0.45 mmol) was added. The reaction mixture was then heated at reflux for a further 20 hrs, before being cooled to 0 °C and filtered. The resultant solid was washed (Et₂O) to give 1,3-di(1*H*-benzo[*d*]imidazol-2-yl)urea **128** (51 mg, 45%) as a colourless solid; m.p. > 350 °C *R*_f 0.00 (1:9 MeOH–EtOAc); δ_{H} (300 MHz, DMSO-*d*₆); 12.15 (2H, s, 2 \times NH), 7.47–7.44 (4H, m, 4 \times ArCH), 7.14–7.11 (4H, m, 4 \times ArCH); Molecule was too insoluble to obtain δ_{C} ; ν_{max} /cm⁻¹ (neat); 3369–2681 (br), 1633, 1593, 1564, 1274, 752; ESI-HRMS found *m/z* 293.1142 [M + H]⁺, C₁₅H₁₃N₆O requires 293.1145.

7-Amino-1,8-naphthyridin-2-ol (94)

Synthesised following literature procedure.²³⁵ Concentrated sulphuric acid (10 mL) was added dropwise to a ground mixture of 2,6-diaminopyridine (2.2 g, 20.0 mmol) and malic acid (3.0 g, 22 mmol) cooled to 0 °C. The reaction mixture was then heated to 110 °C for 3 hrs before being cooled to 0 °C. Ammonium hydroxide solution was then added dropwise to pH 9 and the reaction mixture was filtered and washed with water and Et₂O to give 7-amino-1,8-naphthyridin-2-ol **94** (3.04 g, 95%) as an orange solid; m.p. decomposes > 350 °C [Lit.²³⁵ decomposes > 350 °C]; *R*_f 0.00 (EtOAc); δ_{H} (300 MHz, DMSO-*d*₆); 11.91 (1H, s, OH), 7.66 (1H, d, *J* = 9.5 Hz, ArCH), 7.65 (1H, d, *J* = 8.4 Hz, ArCH), 7.03 (2H, s, NH₂), 6.36 (1H, d, *J* = 8.4 Hz, ArCH), 6.12 (1H, d, *J* = 9.5 Hz, ArCH); δ_{C} (75 MHz, DMSO-*d*₆); 164.1, 160.9, 150.7, 140.1, 137.7, 115.2, 105.6, 105.4; ν_{max} /cm⁻¹ (neat); 3600–2500 (br), 1616, 1511, 1370, 1140, 833; ESI found *m/z* 162.1 [M + H]⁺, C₈H₈N₃O requires 162.2.

N-(7-Hydroxy-1,8-naphthyridin-2-yl)pentanamide (95)

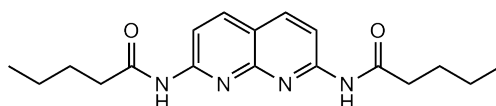
Synthesised following literature procedure.¹⁹¹ Valeroyl chloride (3.70 mL, 31.7 mmol) was added dropwise to a solution of 2-amino-7-hydroxy-1,8-naphthyridine **94** (3.00 g, 18.6 mmol) in pyridine (20 mL) and heated to 110 °C for 20 hrs before the reaction mixture was allowed to cool to room temperature. The pyridine was evaporated and then co-evaporated with toluene and the resultant black residue was dissolved in hot CHCl₃ (20 mL) and hexane (200 mL) was added. The solid was filtered and triturated (MeOH) to give N-(7-hydroxy-1,8-naphthyridin-2-yl)pentanamide **95** (3.47 g, 76%) as a pale yellow powder; m.p. 316–319 °C; *R*_f 0.39 (1:19 MeOH–CH₂Cl₂); δ_H (300 MHz, DMSO-*d*₆); 11.84 (1H, s, OH), 10.48 (1H, s, NH), 8.04 (1H, d, *J* = 8.5 Hz, ArCH), 7.94 (1H, d, *J* = 8.5 Hz, ArCH), 7.84 (1H, d, *J* = 9.4 Hz, ArCH), 6.42 (1H, d, *J* = 9.4 Hz, ArCH), 2.45 (2H, t, *J* = 7.4 Hz, CH₂), 1.55 (2H, m, CH₂), 1.32 (2H, m, CH₂), 0.89 (3H, t, *J* = 7.3 Hz, CH₃); ν_{max}/cm⁻¹ (neat); 3170–2873, 1673, 1585; ESI-HRMS found *m/z* 246.1239 [M + H]⁺, C₁₃H₁₆N₃O₂ requires 246.1237.

N-(7-Chloro-1,8-naphthyridin-2-yl)pentanamide (96)

Synthesised following literature procedure.¹⁹¹ 2-Pentanoylamino-7-hydroxy-1,8-naphthyridine **95** (500 mg, 2.04 mmol) was dissolved in POCl₃ (10 mL) and heated to 95 °C for 3 hrs. The reaction mixture was then cooled and POCl₃ was evaporated *in vacuo*. The residue was poured onto iced water (25 mL) and neutralised with concentrated ammonium hydroxide solution under vigorous stirring. The reaction mixture was extracted into CH₂Cl₂ (3 × 20 mL) and then the organic phase was washed with saturated aqueous sodium bicarbonate (3 × 20 mL), water (3 × 20 mL) and saturated aqueous sodium chloride (20 mL). The organic phase was dried (Na₂SO₄) and evaporated *in vacuo* before being crystallised (1:1 MeOH/MeCN–H₂O) to give N-(7-chloro-1,8-naphthyridin-2-yl)pentanamide **96** (380 mg, 71%) as yellow plates; m.p. 159–162 °C; *R*_f 0.60 (EtOAc); δ_H (300 MHz, CDCl₃); 8.57 (1H, d, *J* = 8.7 Hz, ArCH), 8.35 (1H, s, NH), 8.19 (1H, d,

$J=8.7$ Hz, ArCH), 8.07 (1H, d, $J=8.3$ Hz, ArCH), 7.40 (1H, d, $J=8.3$ Hz, ArCH), 2.48 (2H, t, $J=7.7$ Hz, CH₂), 1.75 (2H, m, CH₂), 1.44 (2H, m, CH₂), 0.96 (3H, t, $J=7.4$ Hz, CH₃); δ_{C} (75 MHz, CDCl₃); 172.7, 154.5, 154.4, 139.1, 138.8, 122.0, 119.2, 115.5, 37.7, 27.3, 22.3, 13.8; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3183, 3129, 3056, 3015, 2951, 2870, 1711, 1490; ESI-HRMS found m/z 264.0896 [M + H]⁺, C₁₃H₁₅ClN₃O requires 264.0898. Found m/z 286.0713 [M + Na]⁺, C₁₃H₁₄ClN₃NaO requires 286.0718.

N,N'-(1,8-Naphthyridine-2,7-diyl)dipentanamide (**93**)

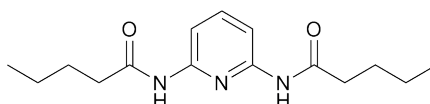


Synthesised following literature procedure.²⁰² 2-Pentanoylamino-7-chloro-1,8-naphthyridine **96** (100 mg, 0.38 mmol), valeramide (46 mg, 0.46 mmol), potassium carbonate (73 mg, 0.53 mmol), palladium (II) acetate (4 mg, 0.02 mmol) and xantphos (18 mg, 0.04 mmol) were suspended in 1,4-dioxane (10 mL) in a schlenk tube. The reaction mixture was heated to 100 °C for 23 hrs before being cooled to room temperature and filtered through celite. The solvent was evaporated *in vacuo* and the crude product was purified by column chromatography (gradient elution: 0:1-1:19 MeOH-CH₂Cl₂) followed by crystallisation (MeOH-H₂O) to give *N,N'*-(1,8-naphthyridine-2,7-diyl)dipentanamide **93** as colourless needles; m.p. 214–216 °C [Lit.⁶¹ 216–217 °C] (Found: C, 65.6; H, 7.45; N, 17.0; C₁₈H₂₄N₄O₂ requires C, 65.8; H, 7.37; N, 17.1%); R_{f} 0.66 (1:9 MeOH-CH₂Cl₂); δ_{H} (300 MHz, CDCl₃); 8.44 (2H, d, $J=8.8$ Hz, 2 × ArCH), 8.35 (2H, s, 2 × NH), 8.12 (2H, d, $J=8.8$ Hz, 2 × ArCH), 2.46 (4H, t, $J=7.5$ Hz, 2 × CH₂), 1.73 (4H, m, 2 × CH₂), 1.42 (4H, m, 2 × CH₂), 0.95 (6H, t, $J=7.3$ Hz, 2 × CH₃); δ_{C} (75 MHz, CDCl₃); 172.3, 153.9, 153.7, 139.0, 118.4, 113.4, 37.8, 27.3, 22.3, 13.8; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3592, 3421, 3297, 2932, 2869, 1674, 1504; ESI-HRMS found m/z 329.1968 [M + H]⁺, C₁₈H₂₅N₄O₂ requires 329.1972.

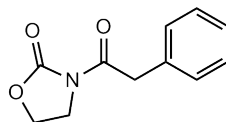
7.4 Experimental for Chapter 5: Self-Sorting of Hydrogen-Bonding Motifs

Synthesis of **17** and **57** described in Section 7.1. Synthesis of **80** described in Section 7.2. Synthesis of **93** described in Section 7.3.

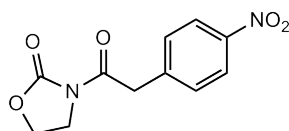
N,N'-(pyridine-2,6-diyl)dipentanamide (**97**)



Synthesised following literature procedure.⁶³ Valeroyl chloride (4.0 mL, 33.7 mmol) was added dropwise to a solution of 2,6-diaminopyridine (1.84 g, 9.5 mmol) and NEt₃ (5.0 mL, 33.7 mmol) in THF (50 mL) stirring at 0 °C and the reaction mixture was then stirred for a further 30 mins before the solvent was removed *in vacuo*. The resultant oil was dissolved in CHCl₃ (30 mL) and washed with H₂O (5 × 100 mL) before the organics were dried (MgSO₄) and removed *in vacuo*. Purification by column chromatography (3:97 MeOH–CHCl₃) followed by crystallisation (CHCl₃–hexane) gave *N,N'*-(pyridine-2,6-diyl)dipentanamide **97** (3.47 g, 74%) as a colourless powder; m.p. 123–124 °C (Found: C, 64.8; H, 8.55; N, 15.2; C₁₅H₂₃N₃O₂ requires C, 65.0; H, 8.36; N, 15.2%); *R*_f 0.37 (3:97 MeOH–CHCl₃); δ_H (300 MHz, CDCl₃); 7.90 (2H, d, *J* = 7.8 Hz, 2 × ArCH), 7.69 (1H, t, *J* = 7.8 Hz, ArCH), 7.55 (2H, s, 2 × NH), 2.38 (4H, t, *J* = 7.5 Hz, 2 × CH₂), 1.71 (4H, m, 2 × CH₂), 1.41 (4H, m, 2 × CH₂), 0.95 (6H, t, *J* = 7.3 Hz, 2 × CH₃); δ_C (75 MHz, CDCl₃); 172.0, 149.9, 141.3, 109.8, 38.0, 27.8, 22.7, 14.2; ν_{max}/cm⁻¹ (neat); 3400, 3255, 2958, 2874, 1981, 1754, 1662, 1591, 1463; ESI-HRMS found *m/z* 300.1682 [M + Na]⁺, C₁₅H₂₃N₃NaO₂ requires 300.1688.

3-(2-Phenylacetyl)oxazolidin-2-one (98a)

Synthesised following literature procedure.²⁰⁵ *n*-Butyl lithium (1.6 M in hexanes, 3.59 mL, 5.74 mmol) was added dropwise to a solution of 2-oxazolidinone (500 mg, 5.74 mmol) in THF (25 mL) stirring at -78°C . The reaction mixture was stirred for 10 mins before phenylacetyl chloride (837 μL , 6.32 mmol) was added and it was then stirred for a further 1 hr before being allowed to warm to room temperature. The reaction mixture was quenched with saturated aqueous ammonium chloride (5 mL) and the organics were evaporated *in vacuo*. The resultant aqueous solution was then extracted into CH_2Cl_2 (3×20 mL) which was dried (Na_2SO_4) and removed *in vacuo* to give the crude product which was purified by column chromatography (CH_2Cl_2) to give 3-(2-phenylacetyl)oxazolidin-2-one **98a** (422 mg, 36%) as a colourless powder; m.p. $54\text{--}56^{\circ}\text{C}$ [Lit.²⁰⁵ $67\text{--}68^{\circ}\text{C}$] All other data is comparable to literature values R_f 0.38 (CH_2Cl_2); δ_{H} (300 MHz, $\text{DMSO-}d_6$); 7.37-7.21 (5H, m, $5 \times \text{ArCH}$), 4.38 (2H, t, $J=7.8$ Hz, CH_2), 4.19 (2H, s, CH_2), 3.91 (2H, t, $J=7.8$ Hz, CH_2); δ_{C} (75 MHz, CDCl_3); 170.3, 152.5, 132.6, 128.7, 127.5, 126.2, 61.0, 41.7, 40.0; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3396 (br), 1775, 1697; ESI-HRMS found m/z 206.0819 $[\text{M} + \text{H}]^+$, $\text{C}_{11}\text{H}_{12}\text{NO}_3$ requires 206.0812. Found m/z 228.0623 $[\text{M} + \text{Na}]^+$, $\text{C}_{11}\text{H}_{11}\text{NNaO}_3$ requires 228.0631.

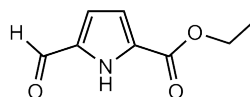
3-(2-(4-Nitrophenyl)acetyl)oxazolidin-2-one (98b)

Synthesised following literature procedure.^{205,206} Thionyl chloride (20.1 mL, 276 mmol) was added to a suspension of 4-nitrophenylacetic acid (5.00 g, 27.6 mmol) in CH_2Cl_2 (200 mL) and refluxed for 3 hrs. The reaction mixture was then allowed to cool to room temperature before the organics were evaporated *in vacuo* to give the acid chloride **103** as the crude material.

n-Butyl lithium (1.6 M in hexanes, 15.7 mL, 25.1 mmol) was added dropwise to a solution of 2-oxazolidinone (2.18 g, 25.1 mmol) in THF (100 mL) stirring at -78°C . The

reaction mixture was stirred for 15 mins before acid chloride **103** (5.51 g, 27.6 mmol) in THF (50 mL) was added and then stirred for a further 1 hr before being warmed to room temperature. Saturated aqueous ammonium chloride (200 mL) was added and the organics were removed *in vacuo* to give an aqueous solution which was extracted into CH₂Cl₂ (3 × 100 mL). The combined organics were dried (Na₂SO₄), removed *in vacuo* and the crude material was purified by column chromatography (gradient elution: 3:7-2:3 EtOAc–hexane) and crystallisation (CH₂Cl₂–hexane) to give 3-(2-(4-nitrophenyl)acetyl)oxazolidin-2-one **98b** (505 mg, 8%) as a pale yellow powder; m.p. 139–142 °C (Found: C, 52.6; H, 4.00; N, 10.8; C₁₁H₁₀N₂O₅ requires C, 52.8; H, 4.03; N, 11.2%); *R*_f 0.44 (CH₂Cl₂); δ_H (300 MHz, CDCl₃); 8.20 (2H, d, *J* = 8.9 Hz, ArCH), 7.49 (2H, d, *J* = 8.9 Hz, ArCH), 4.46 (2H, t, *J* = 8.1 Hz, CH₂), 4.40 (2H, s, CH₂), 4.06 (2H, t, *J* = 8.1 Hz, CH₂); δ_C (75 MHz, CDCl₃); 169.8, 153.5, 147.3, 140.9, 130.8, 123.7, 62.2, 42.7, 41.0; ν_{max}/cm⁻¹ (neat); 3424 (br), 1776, 1696; ESI-HRMS found *m/z* 273.0476 [M + Na]⁺, C₁₁H₁₀N₂NaO₅ requires 273.0482.

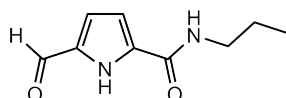
Ethyl 5-formyl-1*H*-pyrrole-2-carboxylate (**105**)



Synthesised following literature procedure.²⁰⁸ Phosphoryl chloride (201 μL, 2.16 mmol) was added dropwise to DMF (167 μL, 2.16 mmol) at 5 °C and stirred for 45 mins. The Vilsmeier reagent was then added to ethyl-1*H*-pyrrole-2-carboxylate (200 mg, 1.44 mmol) in CH₂Cl₂ (5 mL) at -20 °C over 15 mins. The reaction mixture was stirred at -20 °C for 1 hr, 0 °C for 1 hr, room temperature for 1 hr and finally 40 °C for 30 mins before being neutralised with saturated aqueous sodium bicarbonate and heated to reflux for 15 mins. The reaction mixture was then allowed to cool before being diluted with CH₂Cl₂ (20 mL). The layers were separated and the organic phase was washed with saturated aqueous sodium chloride (2 × 20 mL) and dried (Na₂SO₄) before being evaporated *in vacuo*. The crude material was purified by column chromatography (3.5:6.5 EtOAc–hexane) to give ethyl 5-formyl-1*H*-pyrrole-2-carboxylate **105** (107 mg, 44%) as a yellow powder; m.p. 63–65 °C [Lit.²³⁶ 72–73 °C] All other data is comparable to literature values; *R*_f 0.51 (3.5:6.5 EtOAc–hexane); δ_H (400 MHz, CDCl₃); 10.03 (1H, s, NH) 9.67 (1H, s, CHO) 6.95 (1H, d, *J* = 6.8 Hz, ArCH) 6.94 (1H, d, *J* = 6.8 Hz, ArCH) 4.39 (2H, q, *J* = 7.2 Hz,

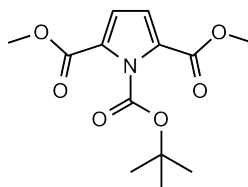
CH_2) 1.39 (3H, t, $J=7.2$ Hz, CH_3); δ_{C} (100 MHz, CDCl_3); 180.3, 160.3, 134.3, 128.6, 119.7, 115.6, 61.4, 14.3; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3435 (br), 2083, 1702, 1666; ES-HRMS found m/z 167.0582 $[\text{M} + \text{H}]^+$, $\text{C}_8\text{H}_9\text{NO}_3$ requires 167.0582.

5-Formyl-*N*-propyl-1*H*-pyrrole-2-carboxamide (106)

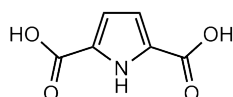


Synthesised following literature procedure.^{208,237} Potassium hydroxide (219 mg, 3.91 mmol) in H_2O (1.2 mL) was added to a solution of ethyl-5-formylpyrrole-2-carboxylate **105** (594 mg, 3.56 mmol) in ethanol (4.8 mL) and the reaction mixture was heated to reflux for 3 hrs. The reaction mixture was then allowed to cool to room temperature before the solid was filtered. The yellow solid was dissolved in H_2O (5 mL) and acidified with conc. HCl. The resultant solid was filtered and washed with H_2O to give the pyrrole acid **129**.

Propylamine (370 μL , 4.51 mmol) was added to a solution of the pyrrole acid **129** (570 mg, 4.10 mmol) in CH_2Cl_2 (60 mL) and EDCI (1.18 g, 6.15 mmol) and HOBT (740 mg, 4.77 mmol) were added. The reaction mixture was stirred at room temperature for 20 hrs before being diluted with CH_2Cl_2 (250 mL) and washed with saturated aqueous sodium chloride (3×100 mL). The combined organics were dried (Na_2SO_4) and removed *in vacuo* and the crude product was purified by column chromatography (EtOAc) followed by cold trituration (CH_2Cl_2) to give 5-formyl-*N*-propyl-1*H*-pyrrole-2-carboxamide **106** (565 mg, 77%) as a yellow solid; m.p. 162–164 °C; R_f 0.39 (1:19 MeOH– CH_2Cl_2); δ_{H} (300 MHz, $\text{DMSO}-d_6$); 12.48 (1H, s, ArNH), 9.64 (1H, s, CHO), 8.36 (1H, s, NH), 6.95 (1H, d, $J=3.9$ Hz, ArCH), 6.87 (1H, d, $J=3.9$ Hz, ArCH), 3.21 (2H, t, $J=6.0$ Hz, CH_2), 1.53 (2H, m, CH_2), 0.90 (3H, t, $J=7.5$ Hz, CH_3); δ_{C} (100 MHz, MeOD); 184.0, 181.5, 119.9, 116.5, 112.2 ($\times 2$), 41.7, 23.1, 11.1; ESI-HRMS found m/z 181.0979 $[\text{M} + \text{H}]^+$, $\text{C}_9\text{H}_{13}\text{N}_2\text{O}_2$ requires 181.0972.

1-*tert*-Butyl-2,5-dimethyl-1*H*-pyrrole-1,2,5-tri-carboxylate (110)

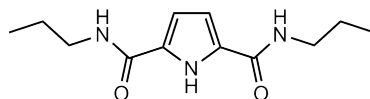
Synthesised following literature procedure.²¹⁰ *n*-Butyl lithium (1.6 M in hexanes, 9.40 mL, 14.97 mmol) was added to a solution of 2,2,6,6-tetramethylpiperidine (2.50 mL, 14.97 mmol) in THF (30 mL) stirring at -78 °C. *N*-Boc pyrrole (1.00 mL, 5.99 mmol) was then added and the reaction mixture was stirred at -78 °C for 3 hrs. Methyl chloroformate (1.40 mL, 17.96 mmol) was then added and the reaction mixture was stirred for a further 1 h before saturated aqueous ammonium chloride (5 mL) was added and allowed to warm to room temperature. The reaction mixture was then diluted with ether (10 mL) and washed with 10% aqueous hydrogen chloride (2 × 20 mL) and saturated aqueous sodium chloride (2 × 20 mL). The organics were dried (Na₂SO₄) and evaporated *in vacuo* to give the crude product which was purified by column chromatography (1:9 EtOAc–hexane) to give 1-*tert*-butyl-2,5-dimethyl-1*H*-pyrrole-1,2,5-tricarboxylate **110** (240 mg, 14%) as a cream coloured powder; m.p. 111–113 °C [Lit.²³⁸ 112–114 °C]; *R*_f 0.35 (1:4 EtOAc–hexane); δ_H (500 MHz, CDCl₃); 6.84 (2H, s, 2 × ArCH), 3.87 (6H, s, 2 × Me), 1.66 (9H, s, ^tBu); δ_C (75 MHz, CDCl₃); 162.1, 160.3, 127.1, 116.2, 86.7, 52.4, 27.7; ν_{max}/cm⁻¹ (neat); 2996, 1774, 1732, 1704, 1373, 1260, 745; ESI-HRMS found *m/z* 306.0957 [M + Na]⁺, C₁₃H₁₇NNaO₆ requires 306.0948.

1*H*-Pyrrole-2,5-dicarboxylic acid (108)

Synthesised following literature procedure.²¹⁰ Sodium hydroxide (171 mg, 4.26 mmol) was added to a suspension of the pyrrole **110** in H₂O–MeOH (5:1, 7.2mL) and the reaction mixture was heated to reflux for 22 hrs. Once cooled the reaction mixture was acidified with 30% aqueous hydrochloric acid before being cooled to 0 °C. The resultant solid was filtered and triturated (Et₂O) and evaporation of the organics gave 1*H*-pyrrole-2,5-dicarboxylic acid **108** (56 mg, 57%) as a white powder; m.p. decomposes > 250 °C

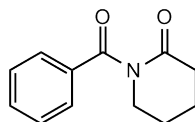
[Lit.²³⁹ decomposes $> 250\text{ }^{\circ}\text{C}$]; R_f 0.00 (1:4 MeOH–EtOAc); δ_H (300 MHz, DMSO- d_6); 12.70 (2H, s, CO₂H), 12.15 (1H, s, NH), 6.75 (2H, d, $J = 3.0\text{ Hz}$, $2 \times \text{ArCH}$); δ_C (75 MHz, DMSO- d_6); 161.6, 127.7, 115.3; $\nu_{\max}/\text{cm}^{-1}$ (neat); 3432, 3127–2571, 1679; ESI-HRMS found m/z 154.0166 $[\text{M} - \text{H}]^-$, C₆H₄NO₄ requires 154.0146.

***N*²,*N*⁵-Dipropyl-1*H*-pyrrole-2,5-dicarboxamide (99)**

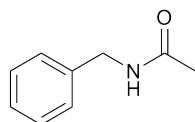


Synthesised following literature procedure.²¹⁰ Pyrrole-2,5-dicarboxylic acid **108** (20 mg, 0.14 mmol) was dissolved in SOCl₂ (7 mL) and the reaction mixture was heated at reflux for 20 hrs. SOCl₂ was then evaporated *in vacuo* to give the pyrrole-2,5-diacid chloride **130**.

Pyridine (47 μL , 0.59 mmol) and propyl amine (43 μL , 0.53 mmol) were added to a solution of pyrrole-2,5-diacid chloride **130** (27 mg, 0.14 mmol) in CH₂Cl₂ (5 mL) and the reaction mixture was stirred at room temperature for 2 days. The reaction mixture was then diluted with H₂O (20 mL), the layers were separated and the aqueous phase was extracted into CH₂Cl₂ ($2 \times 10\text{ mL}$). The organic phase was combined and dried (Na₂SO₄) before being evaporated *in vacuo*. The crude product was purified by column chromatography (gradient elution: 0:1–1:49 MeOH–CH₂Cl₂) and crystallised (EtOAc–hexane) to give the *N*²,*N*⁵-dipropyl-1*H*-pyrrole-2,5-dicarboxamide **99** (24 mg, 72%) as a colourless solid; m.p. 193–195 $^{\circ}\text{C}$; R_f 0.49 (1:9 MeOH–CH₂Cl₂); δ_H (300 MHz, CDCl₃); 6.47 (2H, d, $J = 2.6\text{ Hz}$, $2 \times \text{ArCH}$), 3.39 (4H, t, $J = 6.4\text{ Hz}$, $2 \times \text{CH}_2$), 1.62 (4H, m, $2 \times \text{CH}_2$), 0.97 (6H, t, $J = 7.4\text{ Hz}$, $2 \times \text{CH}_3$); δ_C (125 MHz, CDCl₃); 160.3, 128.6, 109.4, 41.3, 23.0, 11.4; $\nu_{\max}/\text{cm}^{-1}$ (neat); 3241 (br), 2963, 2876, 1595, 1527; ESI-HRMS found m/z 238.1543 $[\text{M} + \text{H}]^+$, C₁₂H₂₀N₃O₂ requires 238.1550.

1-Benzoylpiperidin-2-one (100)

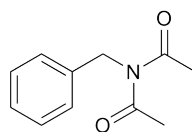
Synthesised following literature procedure.²⁴⁰ 2-Piperidinone (2.00 g, 20.2 mmol) in THF (20 mL) was added *via* cannula to a solution of *n*-butyl lithium (1.6 M in hexanes, 14.5 mL, 23.5 mmol) in THF (50 mL) and the solution was stirred at -78 °C for 20 mins. Benzoyl chloride (2.35 mL, 20.2 mmol) was then added to the reaction mixture before it was warmed to room temperature and stirred for 20 hrs. The reaction mixture was then quenched with saturated aqueous ammonium chloride (30 mL) and extracted into Et₂O (3 × 50 mL). The organics were combined and evaporated *in vacuo* and the crude product was crystallised (toluene) to give 1-benzoylpiperidin-2-one **100** (1.20 g, 29%) as a colourless powder; m.p. 98–101 °C [Lit.²⁴¹ 108–110 °C] All other data is comparable to literature values; *R*_f 0.44 (CH₂Cl₂); δ_H (300 MHz, CDCl₃); 7.54 (2H, d, *J* = 7.2 Hz, 2 × ArCH), 7.46 (1H, d, *J* = 7.2 Hz, ArCH), 7.38 (2H, t, *J* = 7.2 Hz, 2 × ArCH), 3.81 (2H, t, *J* = 6.3 Hz, CH₂), 2.57 (2H, t, *J* = 6.8 Hz, CH₂), 1.97 (4H, m, 2 × CH₂); δ_C (75 MHz, CDCl₃); 175.1, 173.9, 136.6, 131.9, 128.5, 128.3, 46.5, 35.0, 23.2, 21.9; ν_{max}/cm⁻¹ (neat); 2965, 1860, 1694, 1672; ESI-HRMS found *m/z* 204.1026 [M + H]⁺, C₁₂H₁₄NO₂ requires 204.1019.

***N*-Benzylacetamide (113)**

Benzylamine (3.00 mL, 27.50 mmol), methyl acetate (2.18 mL, 27.50 mmol) and TBD (0.38 g, 2.75 mmol) were all charged in a schlenk tube and the reaction mixture was stirred at room temperature for 17 hrs. The crude product was then purified by column chromatography (gradient elution: 0:1-1:9 MeOH-CH₂Cl₂) to give *N*-benzylacetamide **113** (4.01 g, 98%) as a colourless powder; m.p. 53–54 °C [Lit.²⁴² 61 °C]; *R*_f 0.46 (1:9 MeOH-CH₂Cl₂); δ_H (300 MHz, CDCl₃); 7.36–7.26 (5H, m, 5 × ArCH), 5.79 (1H, s, NH), 4.43 (2H, d, *J* = 5.7 Hz, CH₂), 2.02 (3H, s, CH₃); δ_C (75 MHz, CDCl₃); 169.8, 138.3, 128.7, 127.9, 127.6, 43.8, 23.3; ν_{max}/cm⁻¹ (neat); 3291, 3030, 2923, 2875, 1627, 1553;

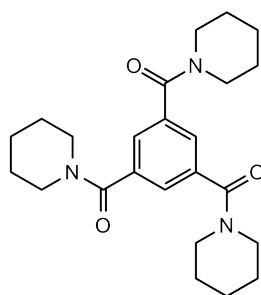
ESI-HRMS found m/z 150.0938 $[M + H]^+$, $C_9H_{12}NO$ requires 150.0913. Found m/z 172.0742 $[M + Na]^+$, $C_9H_{11}NNaO$ requires 172.0733.

N-Acetyl-*N*-benzylacetamide (**114**)



Synthesised following literature procedure.²⁴⁰ *n*-Butyl lithium (1.6 M in hexanes, 3.50 mL, 5.59 mmol) was added to a solution of **113** (1.00 g, 6.71 mL) in THF (30 mL) cooled to -78°C and the reaction mixture was stirred for 20 mins before being allowed to warm to room temperature. Acetic anhydride (0.63 mL, 6.71 mmol) was then added and the reaction mixture was stirred for a further 18 hrs. The reaction mixture was then quenched with saturated aqueous ammonium chloride (30 mL) and extracted into Et_2O (3×50 mL) before being dried (Na_2SO_4) and evaporated *in vacuo*. The resultant gel was purified by column chromatography (CH_2Cl_2) to give *N*-acetyl-*N*-benzylacetamide **114** (480 mg, 45%) as a yellow liquid; R_f 0.41 (1:1 EtOAc-hexane); δ_{H} (300 MHz, CDCl_3); 7.36-7.28 (3H, m, $3 \times \text{ArCH}$), 7.15 (2H, d, $J=8.2$ Hz, $2 \times \text{ArCH}$), 4.97 (2H, s, CH_2), 2.42 (6H, s, $2 \times \text{CH}_3$); δ_{C} (75 MHz, CDCl_3); 173.6, 136.9, 128.9, 127.5, 126.1, 47.7, 26.6; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3110-2941, 1704; ESI-HRMS found m/z 192.1027 $[M + H]^+$, $C_{11}H_{14}NO_2$ requires 192.1019.

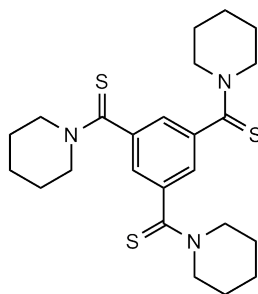
Benzene-1,3,5-triyltris(piperidin-1-ylmethanone) (**102**)



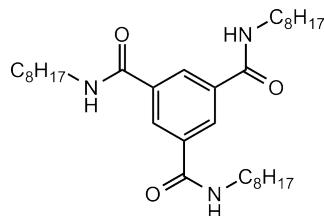
Synthesised following literature procedure.²¹⁵ Piperidine (3.30 mL, 33.8 mmol) in THF (100 mL) was added dropwise to a solution of benzene-1,3,5-tricarbonylchloride (2.00 mL, 11.3 mmol) and triethylamine (5.60 mL, 40.6 mmol) in THF (50 mL) and the reaction

mixture was stirred for 22 hrs at room temperature. The volatiles were then evaporated *in vacuo* and the resultant solid was dissolved in water (150 mL) and extracted into CH₂Cl₂ (2 × 100 mL). The organics were combined and dried (Na₂SO₄) before being evaporated *in vacuo*. The resultant crude product was purified by crystallisation (EtOAc) to give *benzene-1,3,5-triyltris(piperidin-1-ylmethanone* **102** (3.74 g, 81%) as yellow dendrites; m.p. 151–155 °C (Found: C, 69.8; H, 8.15; N, 10.3; C₂₄H₃₃N₃O₃ requires C, 70.0; H, 8.08; N, 10.2%); *R*_f 0.59 (EtOH); δ_H (300 MHz, DMSO-*d*₆); 7.38 (3H, s, 3 × ArCH), 3.58 (6H, br, 6 × NCH), 3.26 (6H, br, 6 × NCH), 1.99 (12H, br, 6 × CH₂), 1.47 (6H, br, 3 × CH₂); δ_C (75 MHz, CDCl₃); 168.6, 137.1, 126.3, 48.9, 43.3, 26.6, 25.7, 24.5; ν_{max}/cm⁻¹ (neat); 3479 (br), 2937, 2856, 1626; ESI-HRMS found *m/z* 412.2598 [M + H]⁺, C₂₄H₃₄N₃O₃ requires 412.2595.

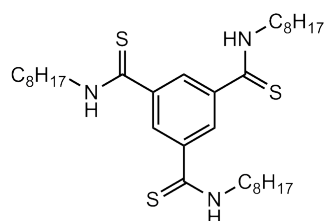
Benzene-1,3,5-triyltris(piperidin-1-ylmethanethione) (**118**)



Synthesised following literature procedure.²¹⁷ Lawesson's reagent (2.95 g, 7.30 mmol) was added to a solution of **102** (1.00 g, 2.43 mmol) in toluene (100 mL) and the reaction mixture was stirred at reflux for 18 hrs before being cooled. The volatiles were then evaporated *in vacuo* and the resultant solid was triturated (MeOH) to give *benzene-1,3,5-triyltris(piperidin-1-ylmethanethione)* **118** (574 mg, 51%) as a yellow powder; m.p. 241–243 °C; *R*_f 0.71 (EtOH); δ_H (500 MHz, CDCl₃); 7.22 (3H, br, 3 × ArCH), 4.34 (6H, br, 6 × NCH), 3.62 (6H, br, 6 × NCH), 1.83 (6H, br, 3 × CH₂), 1.75 (6H, br, 3 × CH₂) 1.59 (6H, br, 3 × CH₂); δ_C (75 MHz, CDCl₃); 197.8, 143.4, 123.9, 54.0, 51.3, 27.3, 25.9, 24.5; ν_{max}/cm⁻¹ (neat); 2938, 2856, 2214, 1488; ESI-HRMS found *m/z* 460.1907 [M + H]⁺, C₂₄H₃₄N₃S₃ requires 460.1909.

***N*¹,*N*³,*N*⁵-Trioctylbenzene-1,3,5-tricarboxamide (119)**

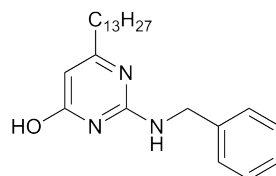
Synthesised following literature procedure.²¹⁵ Octylamine (5.60 mL, 33.8 mmol) in THF (80 mL) was added dropwise to a solution of benzene-1,3,5-tricarbonylchloride (2.00 mL, 11.3 mmol) and triethylamine (5.60 mL, 40.6 mmol) in THF (50 mL) and the reaction mixture was allowed to stir for 3 days at room temperature. The volatiles were then evaporated *in vacuo* and the resultant solid was dissolved in CH₂Cl₂ (100 mL) which was washed with water (3 × 75 mL). The organics were then combined and dried (Na₂SO₄) before being evaporated *in vacuo* to give the crude product which was purified by crystallisation (EtOAc) to give *N*¹,*N*³,*N*⁵-trioctylbenzene-1,3,5-tricarboxamide **119** (3.72 g, 61%) as brown plates; m.p. 178–182 °C [Lit.²⁴³ 196 °C] All other data is comparable to literature values (Found: C, 72.6; H, 10.55; N, 7.6; C₃₃H₅₇N₃O₃ requires C, 72.9; H, 10.60; N, 7.7%); *R*_f 0.68 (EtOAc); δ_H (400 MHz, CDCl₃); 8.32 (3H, s, 3 × ArCH), 6.65 (3H, s, 3 × NH), 3.44 (6H, t, *J* = 7.2 Hz, 3 × CH₂), 1.61 (6H, m, 3 × CH₂), 1.32 (30H, m, 15 × CH₂), 0.88 (9H, t, *J* = 6.8 Hz, 3 × CH₃); δ_C (75 MHz, CDCl₃); 166.1, 135.7, 128.3, 40.8, 32.2, 29.9, 29.7, 29.6, 27.4, 23.0, 14.4; ν_{max}/cm⁻¹ (neat); 3240, 3074, 2926, 2855, 1630; ESI-HRMS found *m/z* 544.4488 [M + H]⁺, C₃₃H₅₈N₃O₃ requires 544.4473. Found *m/z* 566.4303 [M + Na]⁺, C₃₃H₅₇N₃NaO₃ requires 566.4292.

***N*¹,*N*³,*N*⁵-Trioctylbenzene-1,3,5-tris(carbothioamide) (101)**

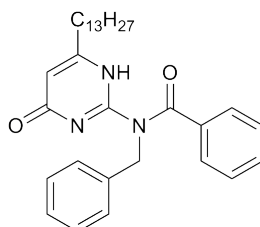
Synthesised following literature procedure.²¹⁷ Lawesson's reagent (2.23 g, 5.52 mmol) was added to a solution of **119** (1.00 g, 1.84 mmol) in toluene (100 mL) and the reaction

mixture was stirred at reflux for 19 hrs before being cooled. The solvents were then evaporated *in vacuo* and the resultant solid was triturated from MeOH to give N¹,N³,N⁵-trioctylbenzene-1,3,5-tris(carbothioamide) **101** (517 mg, 47%) as a sticky yellow solid; (Found: C, 66.7; H, 9.70; N, 7.3; S, 16.4; C₃₃H₅₇N₃S₃ requires C, 67.0; H, 9.70; N, 7.1; S, 16.3%); *R*_f 0.82 (EtOAc); δ_H (300 MHz, CDCl₃); 9.01 (3H, br, 3 × NH), 7.37 (3H, br, 3 × ArCH), 3.64 (6H, br, 3 × CH₂), 1.72 (6H, br, 3 × CH₂), 1.34 (30H, br, 15 × CH₂), 0.91 (9H, t, *J* = 8.1 Hz, 3 × CH₃); δ_C (75 MHz, CDCl₃); 195.4, 137.6, 130.2, 47.4, 32.3, 29.9, 27.9 (× 3), 23.2, 14.6; ν_{max}/cm⁻¹ (neat); 3176, 3040, 2925, 2854, 1542; ESI-HRMS found *m/z* 592.3800 [M + H]⁺, C₃₃H₅₈N₃S₃ requires 295.3787.

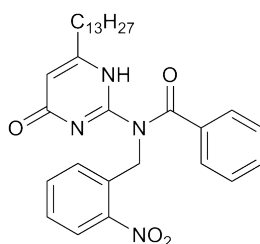
2-(Benzylamino)-6-tridecylpyrimidin-4-ol (**125**)



Benzyl bromide (200 μL, 1.7 mmol) was added dropwise to a solution of **78** (500 mg, 1.7 mmol) and K₂CO₃ (470 mg, 3.4 mmol) in DMF (20 mL) and the reaction mixture was heated to 50 °C for 18 hrs. Once cooled the reaction mixture was filtered before the DMF was removed by vacuum distillation. The resultant solid was triturated (hexane) to give 2-(benzylamino)-6-tridecylpyrimidin-4-ol **125** (184 mg, 28%) as a colourless powder; m.p. 108–110 °C; *R*_f 0.47 (1:19 MeOH–CH₂Cl₂); δ_H (500 MHz, CDCl₃); 7.38 (2H, t, *J* = 7.8 Hz, 2 × ArCH), 7.31 (1H, t, *J* = 7.8 Hz, ArCH), 7.27 (2H, d, *J* = 7.8 Hz, 2 × ArCH), 5.92 (1H, s, ArCH), 5.25 (2H, s, CH₂), 4.74 (2H, s, OH/NH), 2.36 (2H, t, *J* = 7.8 Hz, CH₂), 1.59 (4H, m, 2 × CH₂), 1.32 (18H, m, 9 × CH₂), 0.89 (3H, t, *J* = 6.8 Hz, CH₃); δ_C (75 MHz, CDCl₃); 168.3, 163.0, 154.9, 134.6, 129.4, 128.3, 126.7, 101.8, 44.7, 37.5, 31.9, 29.7–29.4 (× 8), 28.0, 22.7, 14.2; ν_{max}/cm⁻¹ (neat); 3350, 3081, 2917, 2850, 1638; ESI-HRMS found *m/z* 384.3017 [M + H]⁺, C₂₄H₃₈N₃O requires 384.3009.

***N*-Benzyl-*N*-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)benzamide (126)**

Benzoyl chloride (45 μ L, 0.39 mmol) was added to a solution of **125** (100 mg, 0.26 mmol) and DMAP (3 mg, 0.03 mmol) in CHCl_3 (10 mL) and the reaction mixture was heated to reflux for 20 hrs. It was then allowed to cool before the solvent was removed *in vacuo* and the resultant solid was purified by column chromatography (gradient elution: 1:0-0:1 hexane- CH_2Cl_2) to give *N*-benzyl-*N*-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)benzamide **126** (90 mg, 71%) as a colourless powder; m.p. 65–68 °C; R_f 0.37 (CH_2Cl_2); δ_{H} (300 MHz, CDCl_3); 14.06 (1H, s, NH), 8.20 (2H, d, $J=7.5$ Hz, $2 \times \text{ArCH}$), 7.50 (3H, m, $3 \times \text{ArCH}$), 7.39 (2H, t, $J=7.5$ Hz, $2 \times \text{ArCH}$), 7.20 (3H, m, $3 \times \text{ArCH}$), 5.76 (1H, s, ArCH), 5.43 (2H, s, CH_2), 2.41 (2H, t, $J=7.5$ Hz, CH_2), 1.62 (2H, m, CH_2), 1.23 (20H, m, $10 \times \text{CH}_2$), 0.81 (3H, t, $J=7.5$ Hz, CH_3); δ_{C} (75 MHz, CDCl_3); 178.0, 160.5, 154.4, 151.6, 136.0, 135.6, 131.3, 128.5, 127.8, 127.4, 127.2, 126.7, 102.5, 43.7, 32.0, 30.9, 28.6-27.9 ($\times 8$), 25.8, 21.7, 13.1; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3088, 2920, 2853, 1697; ESI-HRMS found m/z 488.3281 $[\text{M} + \text{H}]^+$, $\text{C}_{31}\text{H}_{42}\text{N}_3\text{O}_2$ requires 488.3272.

***N*-(2-Nitrobenzyl)-*N*-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl) benzamide (124)**

2-Nitrobenzylbromide (369 mg, 1.71 mmol) was added to a solution of **78** (500 mg, 1.71 mmol) and K_2CO_3 (471 mg, 3.41 mmol) in DMF (20 mL) and the reaction mixture was heated to 50 °C for 18 hrs. Once cooled the reaction mixture was filtered before the DMF was removed by vacuum distillation. The resultant solid was purified by column chromatography (gradient elution: 0:1-1:9 MeOH- CH_2Cl_2) to give a mixture of

the *O*- and *N*-alkylation products (506 mg, 77%). These could not be separated so were both taken onto the next step.

Benzoyl chloride (140 μ L, 1.17 mmol) was added to a solution of the *O*- and *N*-alkylation products (500 mg, 1.17 mmol) and DMAP (14 mg, 0.12 mmol) in CHCl_3 (50 mL) and the reaction mixture was heated to reflux for 20 hrs. It was then allowed to cool before the solvent was removed *in vacuo* and the resultant solid was purified by column chromatography (gradient elution: hexane– CH_2Cl_2) to give *N*-(2-Nitrobenzyl)-*N*-(4-oxo-6-tridecyl-1,4-dihydropyrimidin-2-yl)benzamide **124** (170 mg, 27%) as a colourless sticky solid; R_f 0.42 (CH_2Cl_2); δ_{H} (300 MHz, CDCl_3); 13.94 (1H, s, NH), 8.04 (1H, d, $J=9.0$ Hz, ArCH), 7.92 (2H, d, $J=9.0$ Hz, $2 \times$ ArCH), 7.44 (2H, m, $2 \times$ ArCH), 7.32 (3H, m, $3 \times$ ArCH), 7.12 (1H, d, $J=9.0$ Hz, ArCH), 5.82 (1H, s, ArCH), 5.77 (2H, s, CH_2), 2.48 (2H, t, $J=7.5$ Hz, CH_2), 1.65 (2H, m, CH_2), 1.25 (20H, m, $10 \times$ CH_2), 0.81 (3H, t, $J=6.0$ Hz, CH_3); δ_{C} (75 MHz, CDCl_3); 179.2, 161.5, 155.0, 153.4, 148.5, 136.4, 133.7, 132.6, 132.4, 129.4, 128.2, 128.0, 127.3, 125.1, 103.1, 42.3, 33.2, 31.9, 29.7×2 , 29.6×2 , 29.4×2 , 29.2, 29.0, 26.9, 22.7, 14.2; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat); 3054, 2940, 2856, 1751, 1699, 1448, 1176, 917; ESI-HRMS found m/z 555.2936 $[\text{M} + \text{Na}]^+$, $\text{C}_{31}\text{H}_{40}\text{N}_4\text{NaO}_4$ requires 555.2942.

7.5 NMR Titration/Dilution Experiments

Anhydrous CDCl_3 was purchased from Aldrich and was distilled over CaCl_2 , to remove residual H_2O and other impurities, before being used immediately. For ^1H NMR titration experiments in Chapters 4 and 5 the CDCl_3 was washed with K_2CO_3 prior to distillation in order to remove trace amounts of HCl . For titrations the ^1H NMR spectrum was recorded for a solution of host (2–10 mM) in CDCl_3 upon sequential additions of aliquots of a solution of guest (20–120 mM) containing host (2–10 mM) in CDCl_3 . The change in chemical shift of key proton resonances was recorded at each titration point. The data was analysed using the HypNMR program using the appropriate model.¹⁶⁵ HypNMR uses data from multiple resonances for curve fitting. Representative data for individual resonances is shown in Appendix B. Each titration was performed in triplicate and the mean binding constant given with the standard deviation given as the error. Dimerisation constants were calculated by taking ^1H NMR spectra of a solution of the desired molecule as it was diluted from ~ 20 –0.1 mM. The change in chemical shift of the key resonances was recorded at each dilution point and analysed in HypNMR using the appropriate model.¹⁶⁵ The self-dimerisation constants were sufficiently small for most molecules that they were not included in the determination of the association constants for heterodimer formation. Self-dimerisation of UPy **80** was assumed to be 1×10^7 in the titration of UIM **17** into UPy **80** (Chapter 5) and was included in the calculation for heterodimerisation.

7.6 Molecular Modelling

Ab initio density-functional calculations were carried out using the Gaussian 03 Package and performed at the level of theory stated in the text. Where B3LYP/6-31G* was used a standard desktop computer was used to carry out the calculation. Where the larger basis set (B3LYP/C, 6-31G, N/O, 6-31G*, H, 6-31++G**) was used calculations were carried out in parallel using eight nodes on a supercomputer.

7.7 DOSY data acquisition

DOSY NMR measurements were made on a Varian Inova 500 MHz spectrometer. All experiments were conducted at 20 °C on CDCl₃ solutions and used a 5 mm ID probe. The bipolar pulse pair simulated echo (BPPSTE) sequence²⁴⁴ was employed operating in ONESHOT mode.²⁴⁵ Additional parameters: number of different gradient strengths, 20; gradient stabilisation delay, 0.002 s; gradient length, 0.002 s; diffusion delay, 0.03 s; relaxation delay, 2.5 s (following measurement of T₁); acquisition time, 2 s; kappa (unbalancing factor), 0.2. Data were processed using a 3 Hz line broadening and exponential multiplication. Data were zero-filled once. Spectra were phased and baseline-corrected prior to production of the pseudo 2D DOSY plots. For measurement of diffusion coefficients, a calibration curve was plotted (using the Stokes-Einstein relationship) for diffusion coefficient ($\times 10^6 \text{ cm}^2 \text{ s}^{-1}$) versus the reciprocal cube root of the molecular mass ($1/(\text{molecular mass})^{\frac{1}{3}}$). The compounds used for this calibration exercise were 4-aminopyrimidine (95.11 Da, $17.281 \times 10^6 \text{ cm}^2 \text{ s}^{-1}$), dipyridyl (184.24 Da, 13.573), 4-(2',6',2''-terpyridinyl)benzyl (323.4 Da, 10.197), benzene-1,3,5-tricarboxylic acid tris(6-methyl-2-pyridinyl)amide (480.0 Da, 11.503), hexakis(2-pyridylmethyl)CTV (913.11 Da, 8.703), tris[4-(2-pyridyl)benzyloxy]CTG (910.14 Da, 9.037). The straight line determined was $y = 73.275x + 1.1263$ ($R^2 = 0.93$).

7.8 NOESY data acquisition

¹H-¹H NOESY data were recorded on a Bruker Avance500 instrument operating at 300 K at a frequency of 500 MHz. A pulse sequence of 8.2 μs pulse, 2.0 s delay, 16.4 μs pulse, 1.2 s delay was used.

7.9 ROESY data acquisition

¹H-¹H ROESY data were recorded on a Bruker Avance DRX500 instrument operating at 300 K at a frequency of 500 MHz. A mix time of 350 ms was applied using a constant wave spin lock mixing method. A pulse sequence repetition of 2.2 s was used.

7.10 ITC experiments

Titration experiments were performed at 25 °C in a MicroCal VP-ITC with a cell volume of 1.4109 mL. The concentration of receptor in the cell was 0.2 mM and the concentration of titrant in the syringe was 2.5 mM. For data fitting, the standard MicroCal Origin software was used for initial data processing to calculate host and guest concentrations and to integrate heat changes for each step in the titrations. For reverse titrations (**57** in syringe and **17** in cell), the heat associated with diluting the host was subtracted from the titration data prior to curve fitting. Dilution experiments with **57** in the syringe gave non-linear curves that are indicative of significant self-association at the concentrations used for the titrations. Therefore, rather than subtracting the heat of **17** dilution on a point-by-point basis, several direct and reverse titrations, along with **17** dilution experiments, were subjected to a global analysis in the SEDPHAT software¹⁶⁷ using a mixed association model that allows for **17** dimerisation coupled to **17**·**57** binding. The model was constrained to exclude the possibility of **57** dimerisation or the formation of any ternary or quaternary complexes. For, **57** into **17** titration was fit to a simple 1:1 binding model to obtain an initial estimate for K_a and ΔH_a for the **17**·**57** interaction. These values were fixed while three dilution experiments for **17** were simultaneously fit to the mixed association model to provide initial estimates for K_{dim} and ΔH_{dim} . A wide range of initial parameter values and repetitive switching between simplex and Marquardt-Levenberg fitting methods were required to avoid local minima at this stage in the fitting process. Then five titration experiments (three **17** into **57** titrations, and two **57** into **17** titrations) were fit to the global model, while keeping K_{dim} and ΔH_{dim} fixed. Any apparent deviations from 1:1 binding were accommodated by using the “incompetent fraction” parameter for each experiment to estimate any errors in concentrations of the **17** or **57**.¹⁶⁷ Finally, K_a , ΔH_a , K_{dim} and ΔH_{dim} were all allowed to float freely in a global fit of eight datasets. Parameter errors (assuming Gaussian-distributed errors) were estimated using Monte Carlo simulations²⁴⁶ as implemented in the SEDPHAT software.¹⁶⁷

7.11 DSC experiments

Samples were run on a DSC Q200 V24.9 build 121 instrument and processed with TA instruments Universal Analysis 2000 software. TZero pans and lids were used for all samples. Analysis was performed over a temperature range -120–150 °C in a cyclic manner (two cycles per sample) with an isothermal stage of 5 mins. The instrument was modulated to FE type, with a cell constant of 0.9708 without correction.

7.12 Viscometry

Viscometry was performed in CHCl₃ at 24.9 °C using a water bath. A Cannon-Fenske viscometer was used and all solutions allowed to equilibrate to the correct temperature before experiments were run. The most concentrated samples were run first before being diluted, mixed and allowed to equilibrate to the desired temperature.

7.13 HySS

All known binding constants were inputted to the HySS²¹⁹ program operating in the simulation mode. The concentration of each component was set to the appropriate amount, 10 mM throughout if it was present from the beginning or 0-50 mM if it was being added. The table of data was then accessed and the concentration of each possible component at each possible concentration was obtained and converted to fidelity using the appropriate calculation.

References

- [1] J. M. Lehn, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89–112.
- [2] M. C. T. Fyfe, J. F. Stoddart, *Acc. Chem. Res.* **1997**, *30*, 393–401.
- [3] C. Britton, M. Poe, G. D. V. van Rossum, *Eng. Princ. Physiol.* **1973**, *2*, 373–391.
- [4] E. S. Yeung, *Anal. Chem.* **1999**, *71*, 522A–529A.
- [5] N. Klonas, M. Rug, I. Harper, M. Wickham, A. Cowman, L. Tilley, *Eur. Biophys. J.* **2002**, *31*, 36–51.
- [6] L. Stryer, *Biochemistry*, W. H. Freeman and Company, New York, USA, 3rd ed., **1988**.
- [7] G. J. Quigley, G. Ughetto, G. A. van der Marel, J. H. van Boom, A. H.-J. Wang, A. Rich, *Science* **1986**, *232*, 1255–1258.
- [8] Y. He, Y. Tian, Y. Chen, A. E. Ribbe, C. Mao, *Chem. Commun.* **2007**, 165–167.
- [9] A. T. Krueger, E. T. Kool, *Curr. Opin. Chem. Biol.* **2007**, *11*, 588–594.
- [10] J. L. Wood, *Biochem. J.* **1974**, *143*, 775–777.
- [11] A. P. Bisson, F. J. Carver, D. S. Egglestone, R. C. Haltiwanger, C. A. Hunter, D. L. Livingstone, J. F. McCabe, C. Rotger, A. E. Rowan, *J. Am. Chem. Soc.* **2000**, *122*, 8856–8868.
- [12] M. Gray, A. O. Cuello, G. Cooke, V. M. Rotello, *J. Am. Chem. Soc.* **2003**, *125*, 7882–7888.
- [13] Y. Li, P. T. J. K. Quansah, S. C. Zimmerman, *J. Am. Chem. Soc.* **2011**, *133*, 17118–17121.
- [14] J. Berna, D. A. Leigh, M. Lubomska, S. M. Mendoza, E. M. Perez, P. Rudolf, G. Teobaldi, F. Zerbetto, *Nat. Mater.* **2005**, *4*, 704–710.
- [15] S. Yagai, A. Kitamura, *Chem. Soc. Rev.* **2008**, *37*, 1520–1529.
- [16] B. J. B. Folmer, E. Cavini, R. P. Sijbesma, E. W. Meijer, *Chem. Commun.* **1998**, 1847–1848.
- [17] S. Masiero, S. Lena, S. Pieraccini, G. P. Spada, *Angew. Chem. Int. Ed.* **2008**, *47*, 3184–3187.
- [18] L. Sun, C. D. Wick, J. I. Siepmann, M. R. Schure, *J. Phys. Chem. B.* **2005**, *109*, 15118–15125.

- [19] A. J. Goshe, I. M. Steele, C. Ceccarelli, A. L. Rheingold, B. Bosnich, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4823–4829.
- [20] J. C. Speakman, *The Hydrogen Bond and other Intermolecular Forces*, The Chemical Society, London, England, **1975**.
- [21] G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, England, **1997**.
- [22] X. Li, Y. Fang, P. Deng, J. Hu, T. Li, W. Feng, L. Yuan, *Org. Lett.* **2011**, *13*, 4628–4631.
- [23] P. Gilli, L. Pretto, G. Gilli, *J. Mol. Struct.* **2007**, *844–845*, 328–339.
- [24] P. Gilli, L. Pretto, V. Bertolasi, G. Gilli, *Acc. Chem. Res.* **2009**, *42*, 33–44.
- [25] A. M. Kelly-Rowley, V. M. Lynch, E. V. Anslyn, *J. Am. Chem. Soc.* **1995**, *117*, 3438–3447.
- [26] L. Bian, *J. Phys. Chem.* **2003**, *107*, 11517–11524.
- [27] J. Taubitz, U. Lüning, *Eur. J. Org. Chem.* **2008**, 5922–5927.
- [28] F. H. Beijer, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, H. Kooijman, A. L. Spek, *J. Org. Chem.* **1996**, *61*, 6371–6380.
- [29] B. A. Blight, A. Camara-Campos, S. Djurdjevic, M. Kaller, D. A. Leigh, F. M. McMillan, H. McNab, A. M. Z. Slawin, *J. Am. Chem. Soc.* **2009**, *131*, 14116–14122.
- [30] S. Djurdjevic, D. A. Leigh, H. McNab, S. Parsons, G. Teobaldi, F. Zerbetto, *J. Am. Chem. Soc.* **2007**, *129*, 476–477.
- [31] J. L. Sessler, R. Wang, *J. Org. Chem.* **1998**, *63*, 4079–4091.
- [32] T. J. Murray, S. C. Zimmerman, *J. Am. Chem. Soc.* **1992**, *114*, 4010–4011.
- [33] Y. Hisamatsu, Y. Fukumi, N. Shirai, S. I. Ikeda, K. Odashima, *Tetrahedron Lett.* **2008**, *49*, 2005–2009.
- [34] F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek, E. W. Meijer, *J. Am. Chem. Soc.* **1998**, *120*, 6761–6769.
- [35] B. A. Blight, C. A. Hunter, D. A. Leigh, H. McNab, P. I. T. Thomson, *Nature Chem.* **2011**, *3*, 244–248.
- [36] E. Greco, A. E. Aliev, V. G. H. Lafitte, K. Bala, D. Duncan, L. Pilon, P. Golding, H. C. Hailes, *New J. Chem.* **2010**, *34*, 2634–2642.
- [37] T. Park, E. M. Todd, S. Nakashima, S. C. Zimmerman, *J. Am. Chem. Soc.* **2005**, *127*, 18133–18142.
- [38] J. Taubitz, U. Lüning, *Aust. J. Chem.* **2009**, *62*, 1550–1555.
- [39] Y. Hisamatsu, N. Shirai, K. Odashima, *Org. Lett.* **2009**, *11*, 4342–4345.
- [40] Y. Hisamatsu, N. Shirai, S. Ikeda, K. Odashima, *Org. Lett.* **2010**, *12*, 1776–1779.

- [41] T. W. Bell, Z. Hou, S. C. Zimmerman, P. A. Thiessen, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2163–2165.
- [42] P. S. Corbin, S. C. Zimmerman, *J. Am. Chem. Soc.* **2000**, *122*, 3779–3780.
- [43] E. E. Fenlon, T. J. Murray, M. H. Baloga, S. C. Zimmerman, *J. Org. Chem.* **1993**, *58*, 6625–6628.
- [44] T. J. Murray, S. C. Zimmerman, S. V. Kolotuchin, *Tetrahedron* **1995**, *51*, 635–648.
- [45] W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.
- [46] H. J. Schneider, R. K. Junega, S. Simova, *Chem. Ber.* **1989**, *122*, 1211–1213.
- [47] J. Sartorius, H. J. Schneider, *Chem. Eur. J* **1996**, *2*, 1446–1452.
- [48] H. Zeng, X. Yang, A. L. Brown, S. Martinovic, R. D. Smith, B. Gong, *Chem. Commun.* **2003**, 1556–1557.
- [49] H. Zeng, R. S. Miller, R. A. I. Flowers, B. Gong, *J. Am. Chem. Soc.* **2000**, *122*, 2635–2644.
- [50] H. Zeng, X. Yang, R. A. I. Flowers, B. Gong, *J. Am. Chem. Soc.* **2002**, *124*, 2903–2910.
- [51] X. Yang, S. Martinovic, R. D. Smith, B. Gong, *J. Am. Chem. Soc.* **2003**, *125*, 9932–9933.
- [52] J. B. Bialecki, L. H. Yuan, B. Gong, *Tetrahedron* **2007**, *63*, 5460–5469.
- [53] R. Cao, J. Zhou, W. Wang, W. Feng, X. Li, P. Zhang, P. Deng, L. Yuan, B. Gong, *Org. Lett.* **2010**, *12*, 2958–2961.
- [54] B. Gong, Y. Yan, H. Zeng, E. Skrzypczak-Jankunn, Y. W. Kim, J. Zhu, H. Ickes, *J. Am. Chem. Soc.* **1999**, *121*, 5607–5608.
- [55] P. Zhang, H. Chu, X. Li, W. Feng, P. Deng, L. Yuan, B. Gong, *Org. Lett.* **2011**, *13*, 54–57.
- [56] A. D. Hamilton, D. Van Engen, *J. Am. Chem. Soc.* **1987**, *109*, 5035–5036.
- [57] A. D. Hamilton, D. Little, *J. Chem. Soc. Chem. Commun.* **1990**, 297–300.
- [58] S.-K. Chang, A. D. Hamilton, *J. Am. Chem. Soc.* **1988**, *110*, 1318–1319.
- [59] M. C. Etter, *Acc. Chem. Res.* **1990**, *23*, 120–126.
- [60] A. M. McGhee, C. Kilner, A. J. Wilson, *Chem. Commun.* **2008**, 344–346.
- [61] P. S. Corbin, S. C. Zimmerman, P. A. Thiessen, N. A. Hawryluk, M. T. J., *J. Am. Chem. Soc.* **2001**, *123*, 10475–10488.
- [62] C. H. Chien, M. K. Leung, J. K. Su, G. H. Li, Y. H. Liu, Y. Wang, *J. Org. Chem.* **2004**, *69*, 1866–1871.
- [63] A. Gooch, A. M. McGhee, L. Renton, J. Plante, C. Lindsay, A. J. Wilson, *Supramol. Chem.* **2009**, *21*, 12–17.

- [64] V. G. H. Lafitte, A. E. Aliev, P. N. Horton, M. B. Hursthouse, K. Bala, P. Golding, H. C. Hailes, *J. Am. Chem. Soc.* **2006**, *128*, 6544–6545.
- [65] R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. H. K. K. Hirschberg, R. F. M. Lange, J. K. L. Lowe, E. W. Meijer, *Science* **1997**, *278*, 1601–1604.
- [66] S. H. M. Sontjens, R. P. Sijbesma, M. H. P. van Genderen, E. W. Meijer, *J. Am. Chem. Soc.* **2000**, *122*, 7487–7493.
- [67] P. S. Corbin, S. C. Zimmerman, *J. Am. Chem. Soc.* **1998**, *120*, 9710–9711.
- [68] T. F. A. de Greef, G. B. W. L. Ligthart, M. Lutz, A. L. Spek, E. W. Meijer, R. P. Sijbesma, *J. Am. Chem. Soc.* **2008**, *130*, 5479–5486.
- [69] P. Mukhopadhyay, P. Y. Zavalij, L. Isaacs, *J. Am. Chem. Soc.* **2006**, *128*, 14093–14102.
- [70] A. Wu, L. Isaacs, *J. Am. Chem. Soc.* **2003**, *125*, 4831–4835.
- [71] B. Feibush, M. Saha, K. Onan, B. Karger, R. Giese, *J. Am. Chem. Soc.* **1987**, *109*, 7531–7533.
- [72] J. Jo, D. Lee, *J. Am. Chem. Soc.* **2009**, *131*, 16283–16291.
- [73] R. R. Knowles, S. Lin, E. N. Jacobsen, *J. Am. Chem. Soc.* **2010**, *132*, 5030–5032.
- [74] L. Diab, T. Šmejkal, J. Geier, B. Breit, *Angew. Chem. Int. Ed.* **2009**, *48*, 8022–8026.
- [75] W. Seiche, A. Schuschkowski, B. Breit, *Adv. Synth. Catal.* **2005**, *347*, 1488–1494.
- [76] A. C. Laungani, B. Breit, *Chem. Commun.* **2008**, 844–846.
- [77] A. Kobori, T. Murase, H. Suda, I. Saito, K. Nakatani, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3431–3433.
- [78] H. Suda, A. Kobori, J. Zhang, G. Hayashi, K. Nakatani, *Bioorg. Med. Chem.* **2005**, *13*, 4507–4512.
- [79] F. Takei, H. Suda, M. Hagihara, J. Zhang, A. Kobori, K. Nakatani, *Chem. Eur. J.* **2007**, *13*, 4452–4457.
- [80] K. Nakatani, S. Sando, I. Saito, *J. Am. Chem. Soc.* **2000**, *122*, 2172–2177.
- [81] P. Cekan, S. T. Sigurdsson, *J. Am. Chem. Soc.* **2009**, *131*, 18054–18056.
- [82] H. C. Ong, J. F. Arambula, S. R. Ramisetty, A. M. Baranger, S. C. Zimmerman, *Chem. Commun.* **2009**, 668–670.
- [83] J. F. Arambula, S. R. Ramisetty, A. M. Baranger, S. C. Zimmerman, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16068–16073.
- [84] A. J. Wilson, *Soft Matter* **2007**, *3*, 409–425.
- [85] J. M. Lehn, *Polym. Int.* **2002**, *51*, 825–839.
- [86] D. N. Reinhoudt, M. Crego-Calama, *Science* **2002**, *295*, 2403–2407.

- [87] A. Ciesielski, S. Lena, S. Masiero, G. P. Spada, P. Samori, *Angew. Chem. Int. Ed.* **2010**, *49*, 1963–1966.
- [88] G. Gottarelli, S. Masiero, G. P. Spada, *J. Chem. Soc. Chem. Commun.* **1995**, 2555–2557.
- [89] D. González-Rodríguez, J. L. J. van Dongen, M. Lutz, A. L. Spek, A. P. H. J. Schenning, E. W. Meijer, *Nature Chem.* **2009**, *1*, 151–155.
- [90] A. Marsh, E. G. Nolen, K. M. Gardinier, J. M. Lehn, *Tetrahedron Lett.* **1994**, *35*, 397–400.
- [91] N. Kimizuka, S. Fujikawa, H. Kuwahara, T. Kunitake, A. Marsh, J. M. Lehn, *J. Chem. Soc. Chem. Commun.* **1995**, 2103–2104.
- [92] A. Marsh, M. Silvestri, J. M. Lehn, *Chem. Commun.* **1996**, 1527–1528.
- [93] H. Fenniri, P. Mathivanan, K. L. Vidale, D. M. Sherman, K. Hallenga, K. V. Wood, J. G. Stowell, *J. Am. Chem. Soc.* **2001**, *123*, 3854–3855.
- [94] H. Fenniri, B. L. Deng, A. E. Ribbe, *J. Am. Chem. Soc.* **2002**, *124*, 11064–11072.
- [95] S. V. Kolotuchin, S. C. Zimmerman, *J. Am. Chem. Soc.* **1998**, *120*, 9092–9093.
- [96] Y. Ma, V. Kolotuchin, S. C. Zimmerman, *J. Am. Chem. Soc.* **2002**, *124*, 13757–13769.
- [97] Y. Yang, M. Xue, J. F. Xiang, C. F. Chen, *J. Am. Chem. Soc.* **2009**, *131*, 12657–12663.
- [98] H. M. Keizer, J. J. González, M. Segura, P. Prados, R. P. Sijbesma, E. W. Meijer, J. de Mendoza, *Chem. Eur. J.* **2005**, *11*, 4602–4608.
- [99] E. M. Todd, S. C. Zimmerman, *J. Am. Chem. Soc.* **2007**, *129*, 14534–14535.
- [100] P. M. Petersen, W. Wu, E. E. Fenlon, S. Kim, S. C. Zimmerman, *Bioorg. Med. Chem.* **1996**, *4*, 1107–1112.
- [101] E. Fan, J. Yang, S. J. Geib, T. C. Stoner, M. D. Hopkins, A. D. Hamilton, *J. Chem. Soc. Chem. Commun.* **1995**, 1251–1252.
- [102] M. P. Lightfoot, F. S. Mair, R. G. Pritchard, J. E. Warren, *Chem. Commun.* **1999**, 1945–1946.
- [103] Y. Yasuda, E. Iishi, H. Inada, Y. Shirota, *Chem. Lett.* **1996**, 575–576.
- [104] K. Hanabusa, C. Koto, M. Kimura, H. Shirai, A. Kakehi, *Chem. Lett.* **1997**, 429–430.
- [105] I. Paraschiv, K. de Lange, M. Giesbers, B. van Lagen, F. C. Grozema, R. D. Abellon, L. D. A. Siebbeles, E. J. R. Sudhölter, H. Zuilhof, T. M. Marcelis, *J. Mater. Chem.* **2008**, *18*, 5475–5481.
- [106] M. K. Müller, L. Brunsveld, *Angew. Chem. Int. Ed.* **2009**, *48*, 2921–2924.
- [107] A. J. Wilson, J. van Gestel, R. P. Sijbesma, E. W. Meijer, *Chem. Commun.* **2006**, 4404–4406.

- [108] A. R. A. Palmans, E. W. Meijer, *Angew. Chem. Int. Ed.* **2007**, *46*, 8948–8968.
- [109] M. M. J. Smulders, T. Buffeteau, D. Cavagnat, M. Wolffs, A. P. H. J. Schenning, E. W. Meijer, *Chirality* **2008**, *20*, 1016–1022.
- [110] M. M. J. Smulders, A. P. H. J. Schenning, E. W. Meijer, *J. Am. Chem. Soc.* **2008**, *130*, 606–611.
- [111] P. J. M. Stals, M. M. J. Smulders, R. Martín-Rapún, A. R. A. Palmans, E. W. Meijer, *Chem. Eur. J.* **2009**, *15*, 2071–2080.
- [112] M. M. J. Smulders, P. J. M. Stals, T. Mes, T. F. E. Paffen, A. P. H. J. Schenning, A. R. A. Palmans, E. W. Meijer, *J. Am. Chem. Soc.* **2010**, *132*, 620–626.
- [113] M. M. J. Smulders, I. A. W. Filot, J. M. A. Leenders, P. van der Schoot, A. R. A. Palmans, A. P. H. J. Schenning, E. W. Meijer, *J. Am. Chem. Soc.* **2010**, *132*, 611–619.
- [114] S. Cantekin, D. W. R. Balkenende, M. M. J. Smulders, A. R. A. Palmans, E. W. Meijer, *Nature Chem.* **2011**, *3*, 42–46.
- [115] M. Kotera, J. M. Lehn, J. P. Vigneron, *J. Chem. Soc. Chem. Commun.* **1994**, 197–199.
- [116] M. Kotera, J. M. Lehn, J. P. Vigneron, *Tetrahedron* **1995**, *51*, 1953–1972.
- [117] R. F. M. Lange, M. van Gulp, E. W. Meijer, *J. Polym. Sci. A: Polym. Chem.* **1999**, *37*, 3657–3670.
- [118] C. Fouquey, J. M. Lehn, A. M. Levelut, *Adv. Mater.* **1990**, *2*, 254–257.
- [119] A. Gooch, S. Barrett, J. Fisher, C. I. Lindsay, A. J. Wilson, *Org. Biomol. Chem.* **2011**, *9*, 5938–5940.
- [120] A. T. ten Cate, H. Kooijman, A. L. Spek, R. P. Sijbesma, E. W. Meijer, *J. Am. Chem. Soc.* **2004**, *126*, 3801–3808.
- [121] S. L. Li, Y. Xiao, W. Xia, X. Ding, Y. Yu, J. Jiang, L. Wang, *Chem. Eur. J.* **2011**, *17*, 10716–10723.
- [122] A. Gooch, PhD thesis, School of Chemistry, University of Leeds, **2011**.
- [123] T. F. A. de Greef, M. M. J. Smulders, M. Wolffs, A. P. H. J. Schenning, R. P. Sijbesma, E. W. Meijer, *Chem. Rev.* **2009**, *109*, 5687–5754.
- [124] A. Khan, D. M. Haddleton, M. J. Hannon, D. Kukulj, A. Marsh, *Macromolecules* **1999**, *32*, 6560–6564.
- [125] O. Uzun, H. Xu, E. Jeoung, R. J. Thibault, V. M. Rotello, *Chem. Eur. J.* **2005**, *11*, 6916–6920.
- [126] T. Park, S. C. Zimmerman, *J. Am. Chem. Soc.* **2006**, *128*, 11582–11590.
- [127] T. Park, S. C. Zimmerman, *J. Am. Chem. Soc.* **2006**, *128*, 13986–13987.
- [128] T. Park, S. C. Zimmerman, *J. Am. Chem. Soc.* **2006**, *128*, 14236–14237.
- [129] P. Cordier, F. Tournilhac, C. S. Ziakovic, L. Leibler, *Nature* **2008**, *451*, 977–980.

- [130] A. W. Bosman, R. P. Sijbesma, E. W. Meijer, *Mater. Today* **2004**, *7*, 34–39.
- [131] P. Woodward, D. H. Merino, I. W. Hamley, A. T. Slark, W. Hayes, *Aust. J. Chem.* **2009**, *62*, 790–793.
- [132] S. Sivakova, D. A. Bohnsack, M. E. Mackay, P. Suwanmala, S. J. Rowan, *J. Am. Chem. Soc.* **2005**, *127*, 18202–18211.
- [133] S. Kiyonaka, K. Sugiyasu, S. Shinkai, I. Hamachi, *J. Am. Chem. Soc.* **2002**, *124*, 10954–10955.
- [134] M. M. Conn, J. Rebek Jr., *Chem. Rev.* **1997**, *97*, 1647–1668.
- [135] R. Wyler, J. de Mendoza, J. Rebek Jr., *Angew. Chem. Int. Ed.* **1993**, *32*, 1699–1701.
- [136] A. Shivanyuk, J. Rebek Jr., *Chem. Commun.* **2001**, 2424–2425.
- [137] L. R. MacGillivray, J. L. Atwood, *Nature* **1997**, *389*, 469–472.
- [138] A. Shivanyuk, J. Rebek Jr., *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7662–7665.
- [139] P. Timmerman, R. H. Vreekamp, R. Hulst, W. Verboom, D. N. Reinhoudt, K. Rissanen, K. A. Udachin, J. Ripmeester, *Chem. Eur. J.* **1997**, *3*, 1823–1832.
- [140] M. M. Safont-Sempere, G. Fernández, F. Würthner, *Chem. Rev.* **2011**, *111*, 5784–5814.
- [141] W. Jiang, H. D. F. Winkler, C. A. Schalley, *J. Am. Chem. Soc.* **2008**, *130*, 13852–13853.
- [142] A. Shivanyuk, J. Rebek Jr., *J. Am. Chem. Soc.* **2002**, *124*, 12074–12075.
- [143] K. A. Jolliffe, P. Timmerman, D. N. Reinhoudt, *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 933–937.
- [144] A. Wu, A. Chakraborty, J. C. Fettingner, R. A. Flowers II, L. Isaacs, *Angew. Chem. Int. Ed.* **2002**, *41*, 4028–4031.
- [145] P. N. Taylor, H. L. Anderson, *J. Am. Chem. Soc.* **1999**, *121*, 11538–11545.
- [146] S. Ghosh, A. Wu, J. C. Fettingner, P. Y. Zavalij, L. Isaacs, *J. Org. Chem.* **2008**, *73*, 5915–5925.
- [147] C. Burd, M. Weck, *Macromolecules* **2005**, *38*, 7225–7230.
- [148] E. R. Kay, D. A. Leigh, F. Zerbetto, *Angew. Chem. Int. Ed.* **2007**, *46*, 72–191.
- [149] K. Konstas, S. J. Langford, M. J. Latter, *Int. J. Mol. Sci.* **2010**, *11*, 2453–2472.
- [150] E. M. Pérez, D. T. F. Dryden, D. A. Leigh, G. Teobaldi, F. Zerbetto, *J. Am. Chem. Soc.* **2004**, *126*, 12210–12211.
- [151] J. V. Hernández, E. R. Kay, D. A. Leigh, *Science* **2004**, *306*, 1532–1537.
- [152] U. Lüning, C. Köhl, A. Uphoff, *Eur. J. Org. Chem.* **2002**, *2002*, 4063–4070.
- [153] C. Schmuck, *Eur. J. Org. Chem.* **1999**, 2397–2403.

- [154] P. K. Baruah, R. Gonnade, U. D. Phalgune, G. J. Sanjayan, *J. Org. Chem.* **2005**, *70*, 6461–6467.
- [155] A. P. Davis, S. M. Draper, G. Dunne, P. Ashton, *Chem. Commun.* **1999**, 2265–2266.
- [156] F. H. Beijer, H. Kooijman, A. L. Spek, R. P. Sijbesma, E. W. Meijer, *Angew. Chem. Int. Ed.* **1998**, *37*, 75–78.
- [157] C. Schmuck, W. Wienand, *J. Am. Chem. Soc.* **2003**, *125*, 452–459.
- [158] P. Prabakharan, V. G. Puranik, G. J. Sanjayan, *J. Org. Chem.* **2005**, *70*, 10067–10072.
- [159] M. L. Pellizzaro, A. M. McGhee, L. C. Renton, M. G. Nix, J. Fisher, W. B. Turnbull, A. J. Wilson, *Chem. Eur. J.* **2011**, *17*, 14508–14517.
- [160] J. L. Cook, C. A. Hunter, C. M. R. Low, A. Perez-Velasco, J. G. Vinter, *Angew. Chem. Int. Ed.* **2007**, *46*, 3706–3709.
- [161] A. M. McGhee, J. P. Plante, C. A. Kilner, A. J. Wilson, *Supramol. Chem.* **2011**, *23*, 470–479.
- [162] F. Bovey, L. Jelinski, P. Mirau, A. Telephone, T. Company, *Nuclear magnetic resonance spectroscopy*, Academic Press New York, **1969**.
- [163] K. S. Cameron, L. Fielding, *J. Org. Chem.* **2001**, *66*, 6891–6895.
- [164] J. Clayden, N. Greeves, S. Warren, P. Wothers, *Organic Chemistry*, Oxford University Press, New York, **2001**.
- [165] C. Frassinetti, S. Ghelli, P. Gans, A. Sabatini, M. S. Moruzzi, A. Vacca, *Anal. Biochem.* **1995**, *231*, 374–382.
- [166] M. M. Pierce, C. S. Raman, B. T. Nall, *Methods* **1999**, *19*, 213–221.
- [167] J. C. D. Houtman, P. H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L. E. Samelson, *Protein Sci.* **2007**, *16*, 30–42.
- [168] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, *Gaussian 03, Revision E.01*, Gaussian, Inc., Wallingford, CT, 2004.
- [169] M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265–3269.

- [170] T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. v. R. Schleyer, *J. Comp. Chem.* **1983**, *4*, 294–301.
- [171] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553–566.
- [172] M. Simon, S Duran, J. J. Dannenberg, *J. Chem. Phys.* **1996**, *105*, 11024–11031.
- [173] V. G. H. Lafitte, A. E. Aliev, E. Greco, K. Bala, P. Golding, H. C. Hailes, *New J. Chem.* **2011**, *35*, 1522–1527.
- [174] S. C. Zimmerman, P. S. Corbin, *Struct. Bonding* **2000**, 63–94.
- [175] Y. Kyogoku, R. C. Lord, A. Rich, *Proc. Natl. Acad. Sci. USA* **1967**, *57*, 250–257.
- [176] R. Deans, G. Cooke, V. M. Rotello, *J. Org. Chem.* **1997**, *62*, 836–839.
- [177] A. Gooch, A. M. McGhee, M. L. Pellizzaro, C. I. Lindsay, A. J. Wilson, *Org. Lett.* **2011**, *13*, 240–243.
- [178] E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, Sausalito, CA, USA, **2004**.
- [179] J. A. Donohue, R. M. Scott, F. M. Menger, *J. Org. Chem.* **1970**, *35*, 2035–2036.
- [180] R. S. Mulliken, *J. Chem. Phys.* **1935**, *3*, 573–585.
- [181] L. S. Shimizu, *Polym. Int.* **2007**, *56*, 444–452.
- [182] M. Wübbenhorst, J. van Turnhout, B. J. B. Folmer, R. P. Sijbesma, E. W. Meijer, *IEEE T. Dielect. E. In.* **2001**, *8*, 365–372.
- [183] H. M. Keizer, R. P. Sijbesma, E. W. Meijer, *Eur. J. Org. Chem.* **2004**, 2553–2555.
- [184] S. H. M. Sontjens, R. P. Sijbesma, M. H. P. van Genderen, E. W. Meijer, *Macromolecules* **2001**, *34*, 3815–3818.
- [185] A. T. ten Cate, P. Y. W. Dankers, H. Kooijman, A. L. Spek, R. P. Sijbesma, E. W. Meijer, *J. Am. Chem. Soc.* **2003**, *125*, 6860–6861.
- [186] J. H. K. K. Hirschberg, R. A. Koevoets, R. P. Sijbesma, E. W. Meijer, *Chem. Eur. J.* **2003**, *9*, 4222–4231.
- [187] J. H. K. K. Hirschberg, F. H. Beijer, H. S. van Aert, P. C. M. M. Magusin, R. P. Sijbesma, E. W. Meijer, *Macromolecules* **1999**, *32*, 2696–2705.
- [188] B. J. B. Folmer, R. P. Sijbesma, R. M. Versteegen, J. A. J. van der Rijt, E. W. Meijer, *Adv. Mater.* **2000**, *12*, 874–878.
- [189] H. M. Keizer, R. P. Sijbesma, J. F. G. A. Jansen, G. Pasternack, E. W. Meijer, *Macromolecules* **2003**, *36*, 5602–5606.
- [190] K. E. Feldman, M. J. Kade, E. W. Meijer, C. J. Hawker, E. J. Kramer, *Macromolecules* **2009**, *42*, 9072–9081.
- [191] G. B. W. L. Ligthart, H. Ohkawa, R. P. Sijbesma, E. W. Meijer, *J. Am. Chem. Soc.* **2005**, *127*, 810–811.

- [192] O. A. Scherman, G. B. W. L. Ligthart, H. Ohkawa, R. P. Sijbesma, E. W. Meijer, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11850–11855.
- [193] T. F. A. de Greef, M. M. L. Nieuwenhuizen, P. J. M. Stals, C. F. C. Fitié, A. R. A. Palmans, R. P. Sijbesma, E. W. Meijer, *Chem. Commun.* **2008**, 4306–4308.
- [194] Z. Chen, A. M. Venkatesan, C. M. Dehnhardt, S. Ayrál-Kaloustian, N. Brooijmans, R. Mallon, I. Feldberg, L. Hollander, J. Lucas, K. Yu, F. Kong, T. S. Mansour, *J. Med. Chem.* **2010**, *53*, 3169–3182.
- [195] A. M. Salaheldin, A. M. F. Oliveira-Campos, P. Parpot, L. M. Rodrigues, M. M. Oliveira, F. P. Feixoto, *Helv. Chim. Acta.* **2010**, *93*, 242–248.
- [196] T. M. Trnka, R. H. Grubbs, *Acc. Chem. Res.* **2001**, *34*, 18–29.
- [197] U. Lüning, C. Köhl, *Tetrahedron Lett.* **1998**, *39*, 5735–5738.
- [198] X. Z. Wang, X. Q. Li, X. B. Shao, X. Zhao, P. Deng, X. K. Jiang, Z. T. Li, Y. Q. Chen, *Chem. Eur. J.* **2003**, *9*, 2904–2913.
- [199] T. F. A. de Greef, G. Ercolani, G. B. W. L. Ligthart, E. W. Meijer, *J. Am. Chem. Soc.* **2008**, *130*, 13755–13764.
- [200] T. Park, S. C. Zimmerman, S. Nakashima, *J. Am. Chem. Soc.* **2005**, *127*, 6520–6521.
- [201] M. L. Pellizzaro, S. A. Barrett, J. Fisher, A. J. Wilson, *Org. Biomol. Chem.* **2011**, submitted.
- [202] G. B. W. L. Ligthart, H. Ohkawa, R. P. Sijbesma, E. W. Meijer, *J. Org. Chem.* **2006**, *71*, 375–378.
- [203] J. R. Quinn, S. C. Zimmerman, *Org. Lett.* **2004**, *6*, 1649–1652.
- [204] E. M. Todd, J. R. Quinn, S. C. Zimmerman, *Isr. J. Chem.* **2005**, *45*, 381–389.
- [205] D. A. Evans, S. G. Nelson, *J. Am. Chem. Soc.* **1997**, *119*, 6452–6453.
- [206] T. Terai, K. Kikuchi, S. Y. Iwasawa, T. Kawabe, Y. Hirata, Y. Urano, T. Nagano, *J. Am. Chem. Soc.* **2006**, *128*, 6938–6946.
- [207] P. H. Beloso, P. Roy, D. L. H. Williams, *J. Chem. Soc. Perkin Trans. 2* **1991**, 17–21.
- [208] C. Schmuck, V. Bickert, M. Merschky, L. Geiger, D. Rupperecht, J. Dudaczek, P. Wich, T. Rehm, U. Machon, *Eur. J. Org. Chem.* **2008**, 324–329.
- [209] R. Voloshchuk, M. Gałęzowski, D. T. Gryko, *Synthesis* **2009**, *7*, 1147–1152.
- [210] D. Yoon, D. Gross, V. Lynch, J. Sessler, B. Hay, C. Lee, *Angew. Chem* **2008**, *47*, 5038–5042.
- [211] R. C. Clevenger, K. D. Turnbull, *Synth. Comm.* **2000**, *30*, 1379–1388.
- [212] C. Sabot, K. A. Kumar, S. Meunier, C. Mioskowski, *Tetrahedron Lett.* **2007**, *48*, 3863–3866.
- [213] L. Simón, J. M. Goodman, *J. Org. Chem.* **2007**, *72*, 9656–9662.

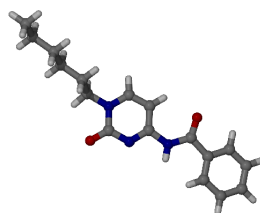
- [214] C. Schmuck, *Chem. Eur. J.* **2000**, *6*, 709–718.
- [215] R. van Heerbeek, P. C. J. Kamer, P. N. M. van Leeuwen, J. N. H. Reek, *Org. Biomol. Chem.* **2006**, *4*, 211–223.
- [216] J. Valdés-Martínez, S. Hernández-Ortega, M. Rubio, D. T. Li, J. K. Swearingen, W. Kaminsky, D. R. Kelman, D. X. West, *J. Chem. Crystallogr.* **2004**, *34*, 533–540.
- [217] T. Olszewska, M. Gdaniec, Połoiński, *J. Org. Chem.* **2008**, *73*, 4859–4864.
- [218] X. Zhao, X. Z. Wang, X. K. Jiang, Y. Q. Chen, Z. T. Li, G. J. Chen, *J. Am. Chem. Soc.* **2003**, *125*, 15128–15139.
- [219] L. Alderighi, P. Gans, A. Ienco, D. Peters, A. Sabatini, A. Vacca, *Coord. Chem. Rev.* **1999**, *184*, 311–318.
- [220] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell*, 4th Ed., Garland Science, Taylor and Francis Group, New York, NY, USA, **2002**.
- [221] C. G. Bochet, *J. Chem. Soc. Perkin Trans. 1* **2002**, 125–142.
- [222] S. Yagai, T. Karatsu, A. Kitamura, *Chem. Eur. J.* **2005**, *11*, 4054–4063.
- [223] M. S. Vollmer, T. D. Clark, C. Steinem, M. R. Ghadiri, *Angew. Chem. Int. Ed.* **1999**, *38*, 1598–1601.
- [224] F. G. Gatti, D. A. Leigh, S. A. Nepogodiev, A. M. Z. Slawin, S. J. Teat, J. K. Y. Wong, *J. Am. Chem. Soc.* **2001**, *123*, 5983–5989.
- [225] A. Goodman, E. Breinlinger, M. Ober, V. M. Rotello, *J. Am. Chem. Soc.* **2001**, *123*, 6213–6214.
- [226] F. Rakotondradany, M. A. Whitehead, A. M. Lebuis, H. F. Sleiman, *Chem. Eur. J.* **2003**, *9*, 4771–4780.
- [227] S. N. Uno, C. Dohno, H. Bitterman, V. L. Malinovskii, R. Häner, K. Nakatani, *Angew. Chem. Int. Ed.* **2009**, *48*, 7362–7365.
- [228] T. Okuyama, Y. Yokoyama, Y. Yokoyama, *Bull. Chem. Soc. Jpn.* **2001**, *74*, 2181–2187.
- [229] M. Takeshita, M. Hayashi, S. Kadota, K. H. Mohammed, Y. Yamato, *Chem. Commun.* **2005**, 761–763.
- [230] M. Herder, M. Pätzelt, L. Grubert, S. Hecht, *Chem. Commun.* **2011**, *47*, 460–462.
- [231] E. J. Foster, E. B. Berda, E. W. Meijer, *J. Am. Chem. Soc.* **2009**, *131*, 6964–6966.
- [232] E. B. Berda, E. J. Foster, E. W. Meijer, *Macromolecules* **2010**, *43*, 1430–1437.
- [233] E. J. Foster, E. B. Berda, E. W. Meijer, *J. Polym. Sci.* **2011**, *49*, 118–126.
- [234] S. Hvilsted, F. Andruzzi, P. Cerrai, M. Tricoli, *Polymer* **1991**, *32*, 127–133.
- [235] G. R. Newkome, S. J. Garbis, V. K. Majestic, F. R. Fronczek, G. Chiari, *J. Org. Chem.* **1981**, *46*, 833–839.

- [236] D. M. Wallace, S. H. Leung, M. O. Senge, K. M. Smith, *J. Org. Chem.* **1993**, *58*, 7245–7257.
- [237] Z. Q. Tian, B. B. Brown, D. P. Mack, C. A. Hutton, P. A. Bartlett, *J. Org. Chem.* **1997**, *62*, 514–522.
- [238] T. J. Donohoe, C. E. Headley, R. P. C. Cousins, A. Cowley, *Org. Lett.* **2003**, *5*, 999–1002.
- [239] T. Zieliński, J. Jurczak, *Tetrahedron* **2005**, *61*, 4081–4089.
- [240] A. de Filippis, D. G. Pardo, J. Cossy, *Tetrahedron* **2004**, *60*, 9757–9767.
- [241] C. J. Foti, D. L. Comins, *J. Org. Chem.* **1995**, *60*, 2656–2657.
- [242] H. O. Nicholas, Erickson, *J. Am. Chem. Soc.* **1926**, *48*, 2174–2176.
- [243] K. Rehse, C. Woyke, A. Rodloff, H. Hahn, *Arch. Pharm. Pharm. Med. Chem.* **1996**, *329*, 155–160.
- [244] M. D. Pelta, H. Barjat, G. A. Morris, A. L. Davis, S. J. Hammond, *Magn. Reson. Chem.* **1998**, *36*, 706–714.
- [245] M. D. Pelta, G. A. Morris, M. J. Stchedroff, S. J. Hammond, *Magn. Reson. Chem.* **2002**, *40*, S147–S152.
- [246] L. M. Schwartz, *Anal. Chem.* **1975**, *47*, 963–964.

Appendix A

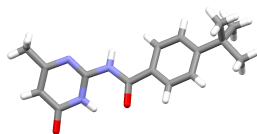
Crystal Structures

Crystal Structure Determination for **61**



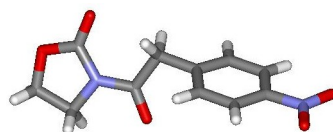
Single crystals were grown by the slow evaporation of a solution of HAC **61** in CDCl_3 -hexane. X-ray diffraction data were collected at the University of Leeds. Crystal data: $\text{C}_{17}\text{H}_{21}\text{N}_3\text{O}_2$, $M = 299.37$, crystal size $0.34 \times 0.32 \times 0.08$ mm, monoclinic, $a = 34.9139(19)$, $b = 5.7312(3)$, $c = 18.4423(10)$ Å, $\alpha = \gamma = 90^\circ$ $\beta = 120.660(1)^\circ$, $U = 3174.4(3)$ Å³, $T = 150$ K, $Z = 8 = 0.084$ mm⁻¹, $\lambda = 0.71073$ Å [Mo- K_α], 17580 reflections measured, 3278 unique [$R(\text{int}) = 0.0446$], 2764 observed [$I > 2\sigma(I)$]. The final R_1 was 0.0349 (observed reflections 0.0439) and wR (F2) was 0.0930 (all data 0.3278) for 200 parameters.

Crystal Structure Determination for 57e



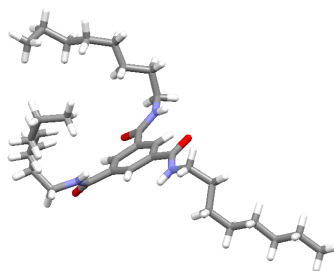
Single crystals were grown by the slow evaporation of a solution of **57e** in methanol. X-ray diffraction data were collected at the University of Leeds. Crystal data: $C_{16}H_{19}N_3O_2$, $M = 285.34$, crystal size $0.36 \times 0.24 \times 0.18$ mm, orthorhombic, $a = 17.2425(18)$, $b = 10.2252(9)$, $c = 18.2845(19)$ Å, $\alpha = \gamma = \beta = 90^\circ$, $U = 3223.7(6)$ Å³, $T = 150$ K, $Z = 8 = 0.079$ mm⁻¹, $\lambda = 0.71073$ Å [Mo- K_α], 48811 reflections measured, 4382 unique [$R(\text{int}) = 0.0569$], 3264 observed [$I > 2\sigma(I)$]. The final R_1 was 0.0610 (observed reflections 0.0808) and wR (F_2) was 0.1814 (all data 0.2023) for 194 parameters.

Crystal Structure Determination for 98b



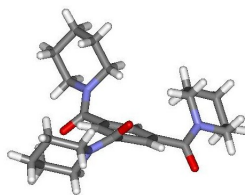
Single crystals were grown by the slow evaporation of a solution of the nitro oxazolidinone **131** in CH_2Cl_2 - Et_2O . X-ray diffraction data were collected at the University of Leeds. Crystal data: $C_{11}H_{10}N_2O_5$, $M = 250.21$, crystal size $0.31 \times 0.19 \times 0.15$ mm, monoclinic, $a = 12.1095(15)$, $b = 5.5934(6)$, $c = 16.9885(19)$ Å, $\alpha = \gamma = 90^\circ$, $\beta = 104.914(5)^\circ$, $U = 1111.9(2)$ Å³, $T = 150$ K, $Z = 4 = 0.120$ mm⁻¹, $\lambda = 0.71073$ Å [Mo- K_α], 19416 reflections measured, 2563 unique [$R(\text{int}) = 0.0363$], 2123 observed [$I > 2\sigma(I)$]. The final R_1 was 0.0362 (observed reflections 0.0450) and wR (F_2) was 0.0921 (all data 0.0992) for 163 parameters.

Crystal Structure Determination for 119



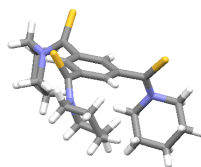
Single crystals were grown by the slow evaporation of a solution of the benzene-1,3,5-tricarboxamide **119** in DMF–Et₂O. X-ray diffraction data were collected at the University of Leeds. Crystal data: C₃₃H₅₇N₃O₃, M = 543.82, crystal size 0.29 × 0.26 × 0.04 mm, monoclinic, a = 11.9302(5), b = 7.6632(4), c = 18.1674(10) Å, α = γ = 90° β = 97.232(3)°, U = 1647.71(14) Å³, T = 150 K, Z = 2 = 0.069 mm⁻¹, λ = 0.71073 Å [Mo-K_α], 24679 reflections measured, 3143 unique [*R*(int) = 0.0933], 2442 observed [*I* > 2σ(*I*)]. The final *R*1 was 0.0729 (observed reflections 0.0954) and w*R* (F₂) was 0.1900 (all data 0.2097) for 355 parameters.

Crystal Structure Determination for 102



Single crystals were grown by the slow evaporation of a solution of the benzene-1,3,5-tricarboxamide **102** in MeOH. X-ray diffraction data were collected at the University of Leeds. Crystal data: C₂₄H₃₃N₃O₃, M = 411.53, crystal size 0.31 × 0.13 × 0.11 mm, monoclinic, a = 6.6907(7), b = 11.4515(13), c = 29.715(3) Å, α = γ = 90° β = 95.640(6)°, U = 2265.7(4) Å³, T = 150 K, Z = 4 = 0.08 mm⁻¹, λ = 0.71073 Å [Mo-K_α], 42958 reflections measured, 5179 unique [*R*(int) = 0.0376], 4269 observed [*I* > 2σ(*I*)]. The final *R*1 was 0.0470 (observed reflections 0.0594) and w*R* (F₂) was 0.1113 (all data 0.1192) for 271 parameters.

Crystal Structure Determination for 118



Single crystals were grown by the slow evaporation of a solution of the benzene-1,3,5-trithioamide **118** in DMSO–hexane. X-ray diffraction data were collected at the University of Leeds. Crystal data: $C_{24}H_{33}N_3S_3$, $M = 459.71$, crystal size $0.22 \times 0.15 \times 0.10$ mm, orthorhombic, $a = 6.9235(5)$, $b = 16.934(11)$, $c = 22.2411(15)$ Å, $\alpha = \gamma = \beta = 90^\circ$, $U = 2478.2(3)$ Å³, $T = 150$ K, $Z = 4 = 0.315$ mm⁻¹, $\lambda = 0.71073$ Å [Mo- K_α], 33002 reflections measured, 6190 unique [$R(\text{int}) = 0.0612$], 5498 observed [$I > 2\sigma(I)$]. The final R_1 was 0.0320 (observed reflections 0.0398) and wR (F2) was 0.0714 (all data 0.0758) for 271 parameters.

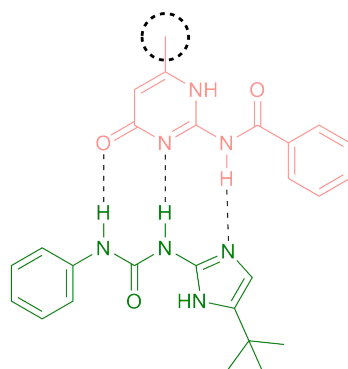
Appendix B

¹H NMR Titrations

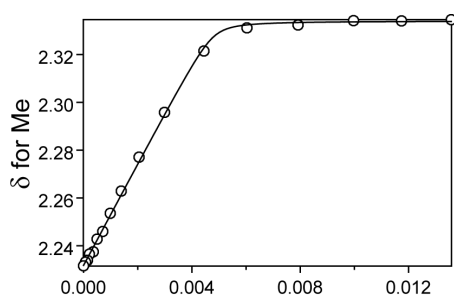
Representative ¹H NMR titration curves are included. The curves for just one resonance is shown, however multiple resonances were followed in order to calculate the association constant of each complex. Each titration was also carried out in triplicate, but just one graph for each titration is shown. The black circle on the molecular structures indicate which proton resonance is being followed.

B.1 Preorganisation

UIM into AIC



57 · 17

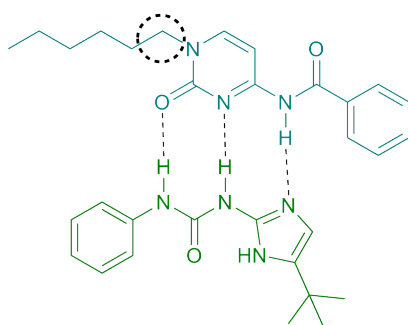


4 iterations
 Refinement converged successfully
 Chi-squared = 22.25
 sigma = 0.00215011245
 RMS weighted residual = 0.00199711202

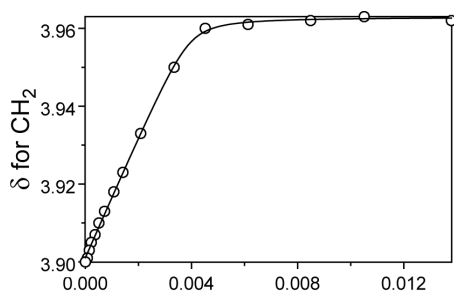
K	rel. std. dev.	logK	std. dev.
2.2490E+004	0.3401	4.3520	0.1477

[17]

UIM into HAC



61 · 17

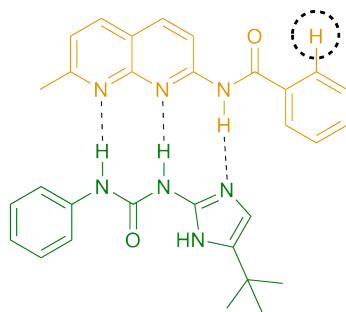
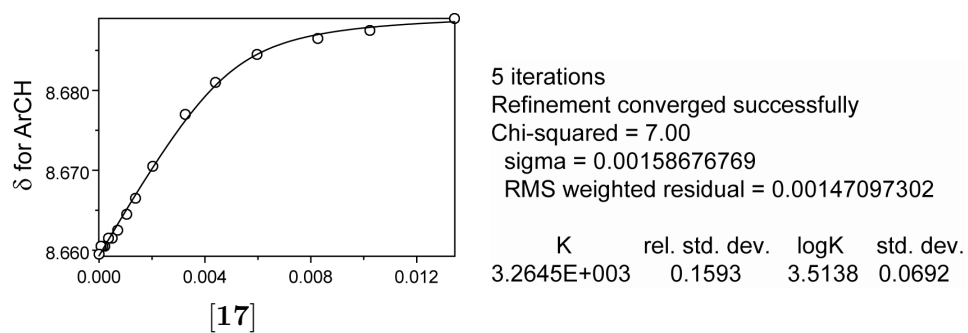


2 iterations
 Refinement converged successfully
 Chi-squared = 1.33
 sigma = 0.00151716576
 RMS weighted residual = 0.00140218185

K	rel. std. dev.	logK	std. dev.
1.9823E+004	0.2204	4.2972	0.0957

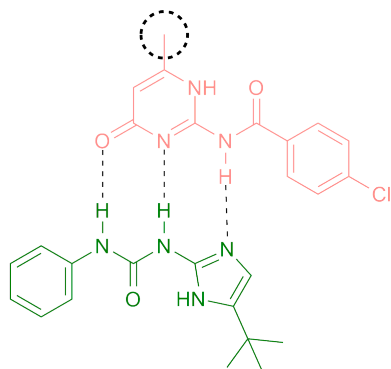
[17]

UIM into ANT

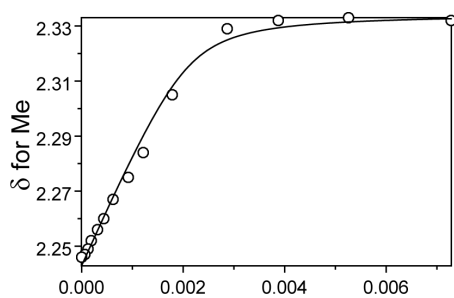
**62 · 17**

B.2 Substituent Effects

UIM into AIC-Cl



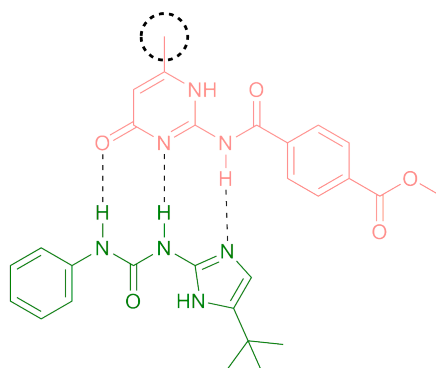
57e · 17a



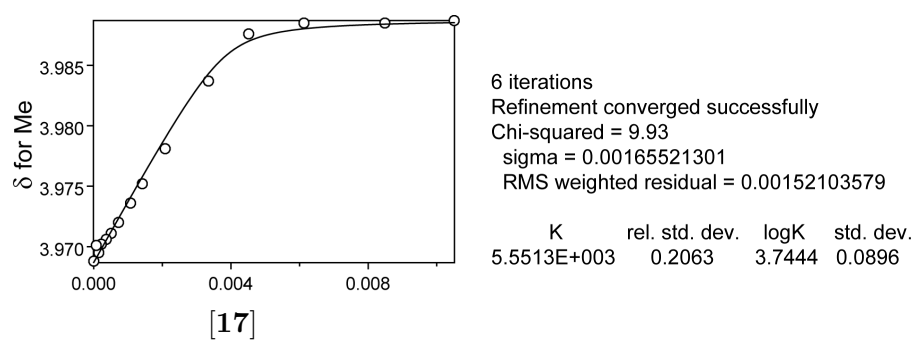
5 iterations
Refinement converged successfully
Chi-squared = 13.62
sigma = 0.00527678668
RMS weighted residual = 0.00481702516

K	rel. std. dev.	logK	std. dev.
3.4372E+003	0.3180	3.5362	0.1381

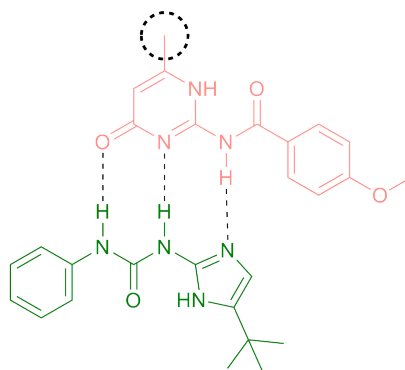
[17]

UIM into AIC-CO₂Me

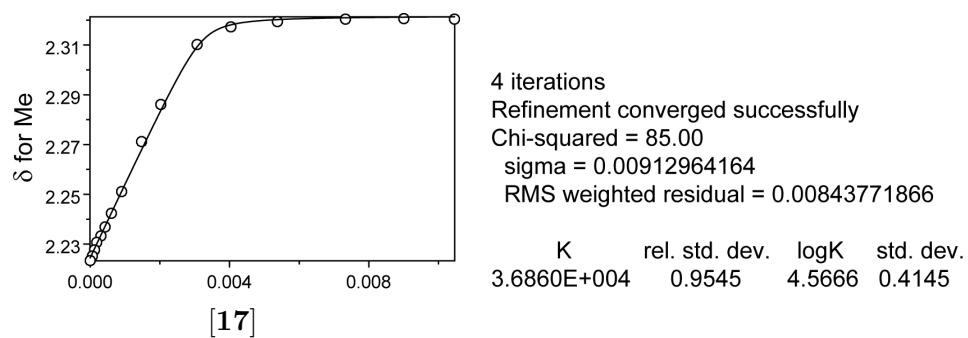
57c · 17a



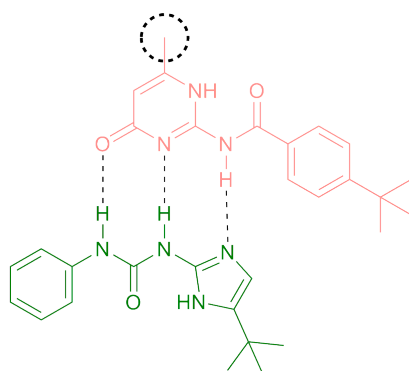
UIM into AIC-OMe



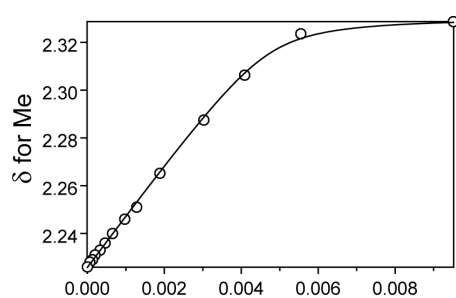
57b · 17a



UIM into AIC-tBu



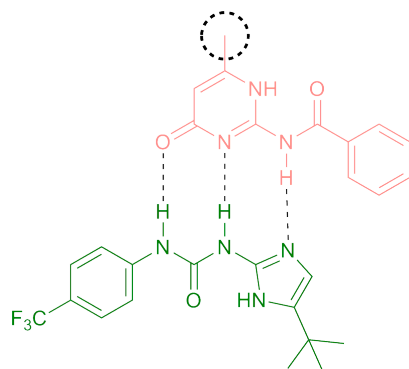
57d · 17a



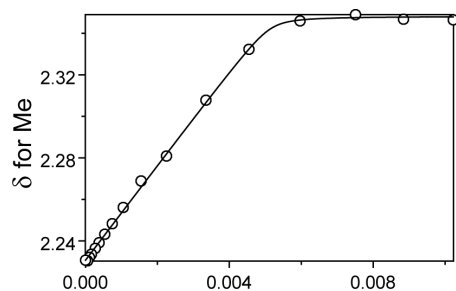
[17]

7 iterations
 Refinement converged successfully
 Chi-squared = 52.48
 sigma = 0.00511400689
 RMS weighted residual = 0.00466842822

K	rel. std. dev.	logK	std. dev.
3.8025E+004	0.3880	4.5801	0.1685

UIM-CF₃ into AIC

57a · 17b

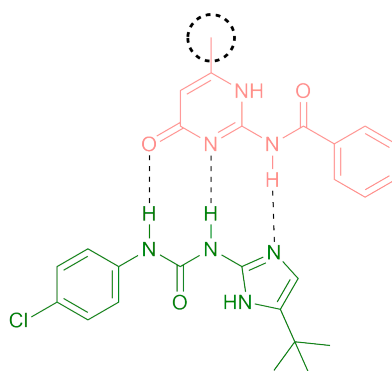


[17b]

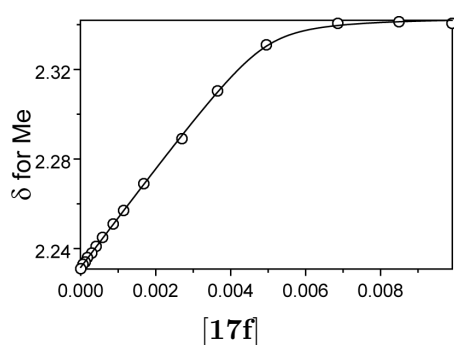
1 iteration
 Refinement converged successfully
 Chi-squared = 18.18
 sigma = 0.00199425569
 RMS weighted residual = 0.00185234591

K	rel. std. dev.	logK	std. dev.
6.9403E+004	0.8975	4.8414	0.3898

UIM-Cl into AIC

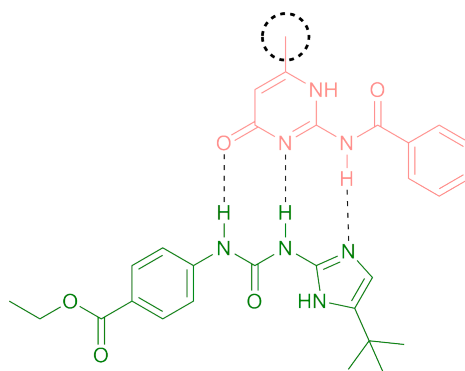


57a · 17f

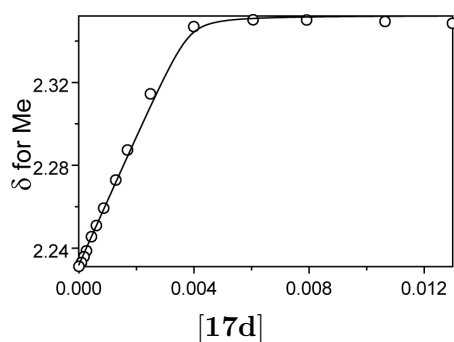


4 iterations
 Refinement converged successfully
 Chi-squared = 32.67
 sigma = 0.00240879873
 RMS weighted residual = 0.00222623918

K	rel. std. dev.	logK	std. dev.
1.1827E+004	0.3914	4.0729	0.1700

UIM-CO₂Et into AIC

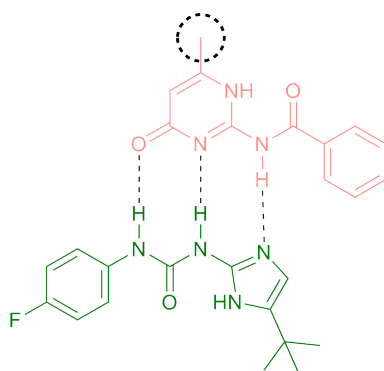
57a · 17d



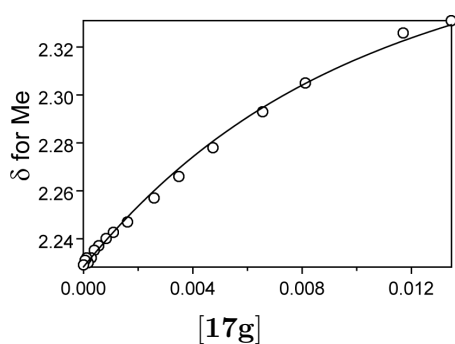
4 iterations
 Refinement converged successfully
 Chi-squared = 52.24
 sigma = 0.00314289134
 RMS weighted residual = 0.00288811783

K	rel. std. dev.	logK	std. dev.
1.0901E+005	0.9782	5.0375	0.4248

UIM-F into AIC



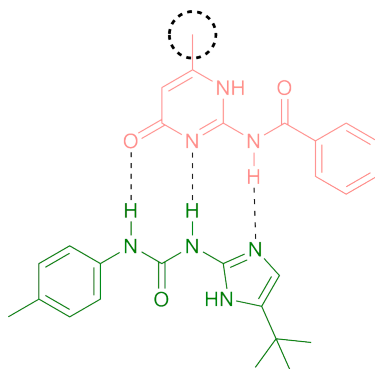
57a · 17g



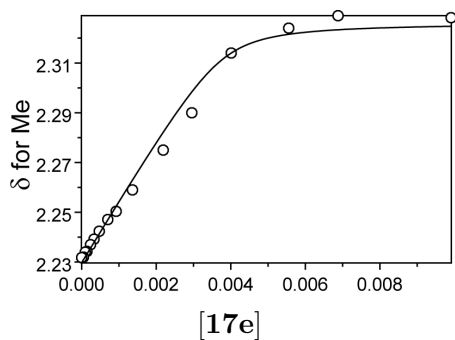
3 iterations
 Refinement converged successfully
 Chi-squared = 2.80
 sigma = 0.00200338069
 RMS weighted residual = 0.00186082158

K	rel. std. dev.	logK	std. dev.
1.3110E+002	0.0011	2.1176	0.0005

UIM-Me into AIC



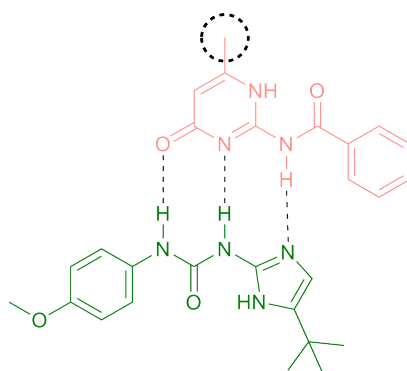
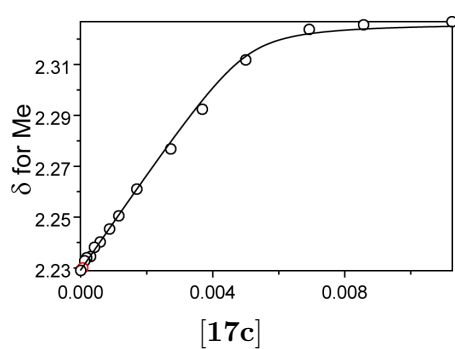
57a · 17e



3 iterations
 Refinement converged successfully
 Chi-squared = 17.33
 sigma = 0.00357849727
 RMS weighted residual = 0.00330728789

K	rel. std. dev.	logK	std. dev.
3.1409E+003	0.2628	3.4971	0.1141

UIM-OMe into AIC

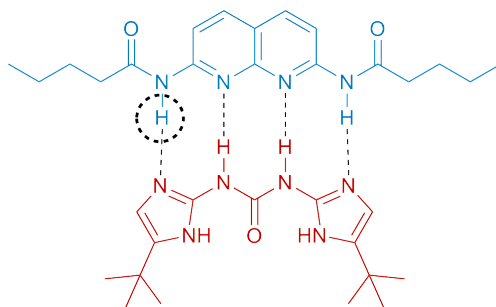
**57a · 17c**

6 iterations
Refinement converged successfully
Chi-squared = 16.33
sigma = 0.00268850968
RMS weighted residual = 0.00247056990

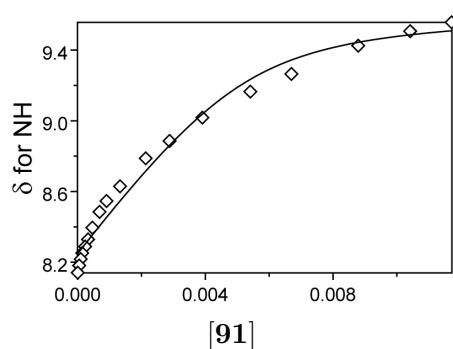
K	rel. std. dev.	logK	std. dev.
5.9786E+003	0.3134	3.7766	0.1361

B.3 UDIM

UDIM into DAN



21 · 91

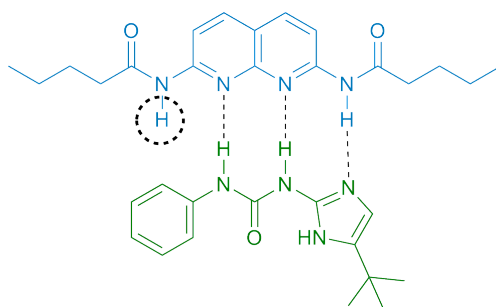


5 iterations
 Refinement converged successfully
 Chi-squared = 4.22
 sigma = 0.06311379850
 RMS weighted residual = 0.05761475188

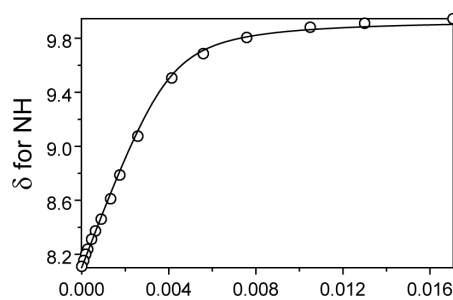
K	rel. std. dev.	logK	std. dev.
8.1955E+002	0.1059	2.9136	0.0460

[91]

UIM into DAN



21 · 17



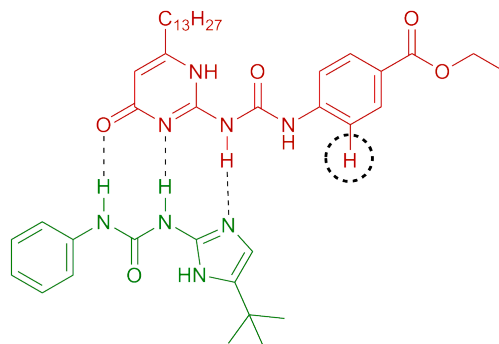
6 iterations
 Refinement converged successfully
 Chi-squared = 4.00
 sigma = 0.01631721812
 RMS weighted residual = 0.01470814165

K	rel. std. dev.	logK	std. dev.
2.1513E+003	0.0822	3.3327	0.0357

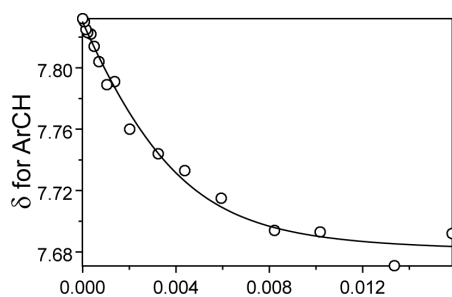
[17]

B.4 Self-Sorting

UIM into UPy



19 · 17



3 iterations
Refinement converged successfully
Chi-squared = 53.31
sigma = 0.00394602473
RMS weighted residual = 0.00353807180

K	rel. std. dev.	logK	std. dev.
5.8271E+004	0.4912	4.7655	0.2133

[17]

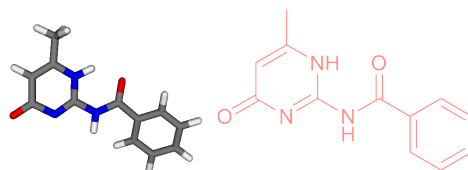
Appendix C

Molecular Modelling

Coordinates that were obtained for all of the calculations that were carried out. *Ab initio* density-functional calculations were carried out using the Gaussian 03 Package and performed at the B3LYP level of theory. The basis set that was used is stated for each molecule. ^tBu groups have been removed from all of the UIM **17** calculations, further simplifying the calculation and reducing the number of possible tautomers and conformers that can be present in each complex. ^tBu groups were also been removed from some of the UDIM **91** calculations for the same reasons.

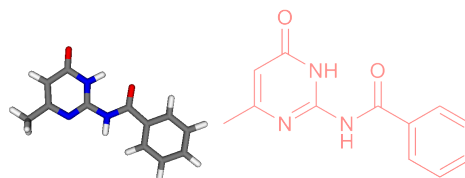
C.1 Molecular Modelling for Chapter 2: Effects of Preorganisation

AIC 57i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



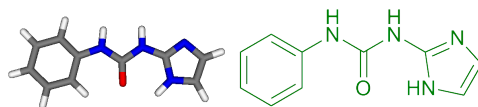
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	1.231291	1.275269	-4.214699
2	6	1.959403	1.355094	-3.020861
3	6	1.335652	1.083446	-1.799400
4	6	-0.025753	0.729808	-1.764594
5	6	-0.745766	0.633658	-2.968221
6	6	-0.121429	0.910332	-4.186130
7	1	1.716461	1.488972	-5.159813
8	1	3.009606	1.620056	-3.040139
9	1	1.927810	1.114091	-0.892060
10	1	-1.787337	0.343054	-2.925695
11	1	-0.685349	0.841731	-5.108596
12	6	-0.759260	0.423742	-0.500393
13	7	-0.201247	0.911899	0.663324
14	1	0.609255	1.516171	0.628613
15	8	-1.820813	-0.223513	-0.502518
16	6	-0.678607	0.756756	1.978507
17	6	-2.312773	-0.173829	3.432754
18	1	-2.248178	-0.361558	1.301770
19	6	-0.439686	1.208673	4.256485
20	6	-1.653329	0.405206	4.469473
21	1	-2.004370	0.293790	5.485581
22	7	-1.809141	0.004672	2.148763
23	7	0.004922	1.341627	2.914333
24	8	0.175871	1.747224	5.175577
25	6	-3.558269	-1.006840	3.545804
26	1	-4.381184	-0.569198	2.969245
27	1	-3.391658	-2.020927	3.165172
28	1	-3.872517	-1.080000	4.587445

AIC 57ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



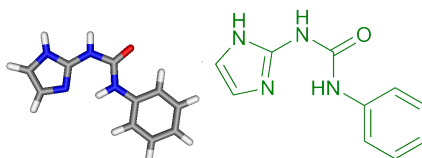
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.408827	-2.840416	-0.389442
2	6	-0.338885	-2.012817	-0.025999
3	6	-0.571117	-0.678161	0.320492
4	6	-1.879599	-0.161354	0.304017
5	6	-2.951024	-1.002960	-0.042220
6	6	-2.715515	-2.333976	-0.392532
7	1	-1.226169	-3.873745	-0.659971
8	1	0.670379	-2.405938	-0.003817
9	1	0.265676	-0.066988	0.639296
10	1	-3.953331	-0.594876	-0.029865
11	1	-3.545748	-2.973954	-0.665850
12	6	-2.207639	1.250949	0.663234
13	8	-3.345878	1.592102	1.010877
14	7	-1.164253	2.158543	0.566740
15	1	-0.276612	1.868599	0.178763
16	6	-1.194948	3.527720	0.857641
17	6	-2.456104	5.457064	1.648682
18	1	-3.160777	3.464299	1.419532
19	6	-0.102454	5.537141	0.945993
20	6	-1.218969	6.175987	1.424541
21	1	-1.217582	7.233390	1.646646
22	7	-2.346153	4.073423	1.324357
23	7	-0.076595	4.184211	0.653605
24	8	-3.523273	5.904685	2.064559
25	6	1.195236	6.259227	0.698517
26	1	1.485842	6.166705	-0.353196
27	1	1.120198	7.318225	0.951239
28	1	1.997888	5.809722	1.292275

UIM 17i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



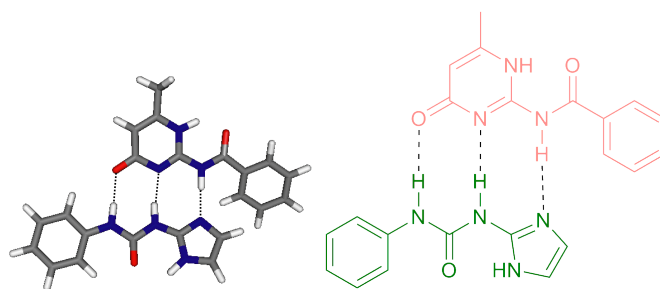
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.114703	-3.794048	0.442554
2	6	0.765834	-3.014201	-0.521368
3	6	0.411720	-1.675988	-0.698807
4	6	-0.602140	-1.095040	0.087331
5	6	-1.259319	-1.874118	1.057016
6	6	-0.892359	-3.214216	1.222314
7	1	0.387671	-4.833093	0.582313
8	1	1.548558	-3.443896	-1.135773
9	1	0.920877	-1.077307	-1.448123
10	1	-2.035306	-1.430175	1.658989
11	1	-1.404342	-3.806623	1.971968
12	7	-0.897747	0.268003	-0.158707
13	1	-0.343399	0.697081	-0.885847
14	6	-1.837752	1.065424	0.467126
15	7	-1.865004	2.362154	-0.021811
16	1	-1.250123	2.670677	-0.761698
17	8	-2.592297	0.687587	1.374769
18	6	-2.723432	3.362610	0.441321
19	6	-4.288247	4.410930	1.594273
20	1	-3.769221	2.305751	1.911575
21	6	-3.705638	5.252830	0.674418
22	1	-5.068572	4.555795	2.319671
23	1	-3.928426	6.289235	0.485359
24	7	-2.723223	4.590226	-0.047654
25	7	-3.642574	3.188900	1.432864

UIM 17ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.282345	-2.708766	-1.836345
2	6	-2.127152	-1.998722	-2.188457
3	6	-1.441112	-1.249670	-1.231368
4	6	-1.901791	-1.198657	0.099457
5	6	-3.062009	-1.910576	0.457426
6	6	-3.738109	-2.656742	-0.514174
7	1	-3.815698	-3.290670	-2.578862
8	1	-1.758420	-2.026640	-3.207663
9	1	-0.546950	-0.701110	-1.507622
10	1	-3.413460	-1.870725	1.476035
11	1	-4.630793	-3.201417	-0.227766
12	7	-1.146390	-0.416193	1.002555
13	1	-0.317922	0.061078	0.636222
14	6	-1.400889	-0.221443	2.331125
15	7	-0.470336	0.616617	2.995392
16	1	-0.701880	0.728539	3.971664
17	8	-2.331434	-0.687416	2.992347
18	6	0.647988	1.256585	2.485611
19	6	2.480201	2.528085	2.444835
20	1	1.341998	2.237422	4.248560
21	6	2.225343	2.001826	1.205814
22	1	3.255068	3.174610	2.815142
23	1	2.782759	2.136433	0.294887
24	7	1.082860	1.210024	1.234985
25	7	1.458308	2.043033	3.265944

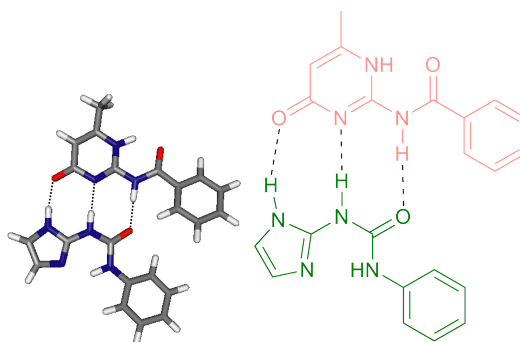
AIC · UIM 57 · 17i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	6.987070	-1.341844	0.295437
2	6	5.600755	-1.495538	0.182107
3	6	4.772677	-0.356808	0.203252
4	6	5.356394	0.920240	0.338131
5	6	6.740854	1.054407	0.450356
6	6	7.568364	-0.076117	0.429986
7	1	7.614510	-2.226291	0.277187
8	1	5.156546	-2.471632	0.074390
9	1	4.717009	1.795166	0.351429
10	1	7.171787	2.044264	0.552977
11	1	8.643716	0.028697	0.516408
12	7	3.363105	-0.399587	0.108027
13	1	2.889279	0.509730	0.135369
14	6	2.577480	-1.501496	-0.103978
15	7	1.215735	-1.196935	-0.117621
16	1	0.871546	-0.235287	-0.014811
17	8	2.988994	-2.666590	-0.264954
18	6	0.237841	-2.126357	-0.418085
19	6	-1.675886	-3.050852	-0.823004
20	6	-0.735188	-4.035995	-0.991360
21	1	1.437263	-3.788609	-0.721997
22	1	-2.743588	-3.121436	-0.936495
23	1	-0.815758	-5.072731	-1.263914
24	7	-1.068009	-1.852832	-0.455961
25	7	0.487415	-3.428327	-0.727597
26	8	2.107931	2.235280	0.038696
27	6	0.897585	2.517817	-0.032621
28	6	0.416782	3.889618	-0.149467
29	6	-0.916477	4.148914	-0.226770
30	1	1.148074	4.684175	-0.174891
31	6	-1.331158	1.801781	-0.068717
32	1	-2.821788	3.167256	-0.256540

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
33	7	-0.058109	1.491315	0.005006
34	7	-1.798908	3.082024	-0.187174
35	6	-1.529525	5.514289	-0.355915
36	1	-2.115640	5.597129	-1.277881
37	1	-2.202418	5.725005	0.482550
38	1	-0.753389	6.279875	-0.373760
39	7	-2.266136	0.767921	-0.000977
40	1	-1.854946	-0.194280	-0.085244
41	6	-3.639602	0.949330	0.002177
42	6	-4.483531	-0.247832	0.272485
43	6	-5.814316	-0.223050	-0.184569
44	6	-4.026987	-1.344583	1.024129
45	6	-6.669716	-1.292239	0.084870
46	1	-6.156773	0.639222	-0.741751
47	6	-4.891824	-2.408303	1.302776
48	1	-3.011346	-1.368687	1.395289
49	6	-6.209742	-2.387208	0.830497
50	1	-7.690084	-1.272223	-0.278936
51	1	-4.536498	-3.247533	1.888527
52	1	-6.875894	-3.214597	1.046162
53	8	-4.170471	2.062396	-0.190029

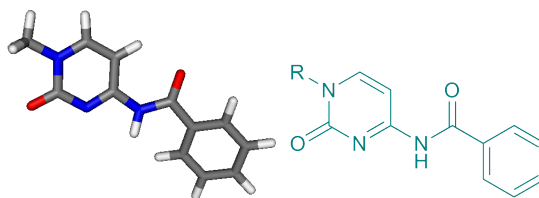
AIC · UIM 57 · 17ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-0.613640	5.766778	0.038603
2	6	-1.531960	4.732072	-0.138330
3	6	-1.181216	3.406080	0.184209
4	6	0.098127	3.133969	0.700719
5	6	1.012756	4.176914	0.881622
6	6	0.662396	5.490943	0.550336

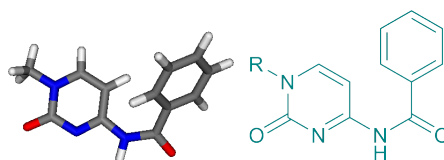
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
7	1	-0.889153	6.782560	-0.218893
8	1	-2.524716	4.926966	-0.521723
9	1	0.397048	2.126404	0.951641
10	1	1.996032	3.957986	1.279848
11	1	1.376100	6.294612	0.690908
12	6	-2.234923	2.370473	-0.000598
13	7	-1.855191	1.033566	0.010007
14	1	-0.856013	0.759111	-0.005030
15	6	-2.739128	-0.045793	-0.056689
16	6	-4.997454	-0.802550	-0.220262
17	1	-4.306581	1.229862	-0.212224
18	6	-3.115000	-2.358387	-0.052524
19	6	-4.542512	-2.083332	-0.165491
20	1	-5.219212	-2.924307	-0.204307
21	8	-3.427484	2.709168	-0.150827
22	7	-4.073787	0.227461	-0.164524
23	7	-2.241322	-1.260505	-0.002463
24	6	-6.438874	-0.396611	-0.339140
25	1	-6.747443	0.215153	0.515707
26	1	-6.607440	0.195023	-1.245704
27	1	-7.080582	-1.277138	-0.380273
28	8	-2.666306	-3.519311	0.002701
29	6	3.672194	1.167145	-0.558172
30	6	3.851750	-0.182673	-0.200478
31	6	5.157025	-0.668801	0.018264
32	6	6.259771	0.175681	-0.117171
33	6	6.083404	1.520162	-0.469546
34	6	4.787622	2.002131	-0.687076
35	1	2.675552	1.539033	-0.731985
36	1	5.296977	-1.707696	0.296337
37	1	7.255452	-0.217655	0.054245
38	1	6.939152	2.176798	-0.574372
39	1	4.636385	3.038696	-0.967999
40	7	2.799314	-1.113342	-0.064658
41	1	3.050390	-2.103874	0.050812
42	6	1.456186	-0.862998	-0.055678
43	7	0.635606	-1.967513	0.071331
44	1	-0.367487	-1.747528	0.050684
45	8	0.932698	0.274589	-0.149254
46	6	1.009637	-3.303487	0.167338
47	6	2.087475	-5.164436	0.315801
48	6	0.752810	-5.492967	0.307640
49	1	-0.945972	-4.120777	0.156317
50	1	2.939619	-5.819561	0.378668
51	1	0.242881	-6.438220	0.359308
52	7	2.246127	-3.786567	0.228439
53	7	0.067693	-4.287799	0.212843

HAC 61i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



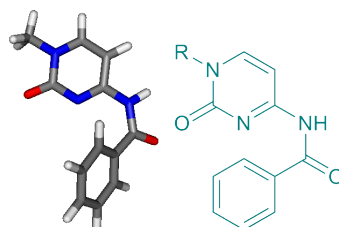
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.893744	-3.153882	-0.899827
2	6	-1.384050	-2.276872	0.662109
3	6	-2.250001	-1.205796	1.036906
4	6	-3.455435	-1.161072	0.388142
5	1	-1.965548	-0.492463	1.789350
6	1	-4.189170	-0.392467	0.594077
7	7	-1.680349	-3.193080	-0.245051
8	7	-3.791460	-2.085595	-0.550084
9	7	-0.114462	-2.452564	1.221010
10	1	0.348522	-3.273936	0.851781
11	6	0.500394	-1.737989	2.241861
12	6	-5.079989	-2.055249	-1.255476
13	1	-4.908991	-1.979412	-2.330362
14	1	-5.623162	-2.983330	-1.070725
15	1	-5.658794	-1.201421	-0.903147
16	8	-0.053954	-0.820974	2.850962
17	8	-3.252192	-3.962963	-1.753294
18	6	1.896431	-2.174834	2.572329
19	6	2.725851	-2.866601	1.671020
20	6	2.390802	-1.838346	3.844333
21	6	4.022674	-3.232295	2.047165
22	1	2.389043	-3.091105	0.665395
23	6	3.683523	-2.209550	4.219921
24	1	1.746921	-1.287085	4.517428
25	6	4.501170	-2.910837	3.323336
26	1	4.657414	-3.758578	1.344251
27	1	4.053110	-1.953235	5.205700
28	1	5.505102	-3.197923	3.613698

HAC 61ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



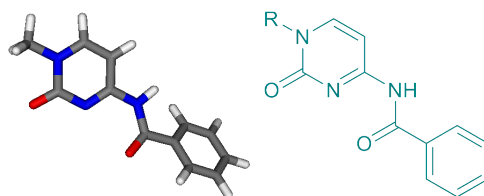
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-0.800254	-3.080016	-2.282303
2	6	-0.699786	-3.267274	-0.897232
3	6	-0.870606	-2.184336	-0.028645
4	6	-1.151956	-0.906090	-0.544079
5	6	-1.280861	-0.729718	-1.932469
6	6	-1.091500	-1.809605	-2.798794
7	1	-0.662966	-3.919781	-2.953665
8	1	-0.495486	-4.252517	-0.495116
9	1	-0.803793	-2.331269	1.041989
10	1	-1.537001	0.251785	-2.310850
11	1	-1.180185	-1.664952	-3.868915
12	6	-1.449930	0.260739	0.341338
13	7	-0.799030	0.379660	1.577625
14	1	-1.274920	1.006508	2.219998
15	6	0.494162	0.065544	2.005585
16	6	1.511109	-0.416610	1.120998
17	6	2.740634	-0.643102	1.679565
18	1	1.342503	-0.577282	0.070852
19	6	1.933661	0.092411	3.857098
20	1	3.578293	-1.000186	1.094330
21	8	-2.318299	1.081061	0.049574
22	7	0.697506	0.315465	3.287828
23	7	2.971431	-0.418363	2.998591
24	8	2.203771	0.293685	5.038458
25	6	4.282200	-0.654112	3.619385
26	1	4.647339	0.271375	4.066763
27	1	4.185781	-1.392816	4.416682
28	1	4.978580	-1.008312	2.859068

HAC 61iii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



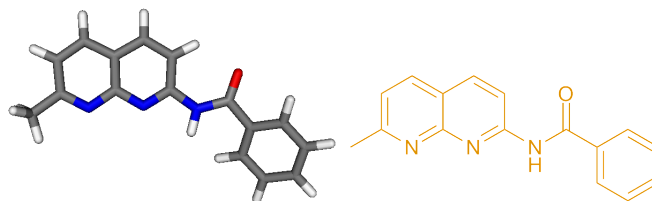
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.266700	3.356918	5.885846
2	6	-1.454902	2.223273	5.748426
3	6	-1.435813	1.516193	4.544060
4	6	-2.239827	1.933260	3.470255
5	6	-3.070644	3.057384	3.620985
6	6	-3.075155	3.773218	4.819821
7	1	-2.269919	3.910727	6.817746
8	1	-0.832898	1.896898	6.573306
9	1	-0.801461	0.647555	4.438357
10	1	-3.701568	3.351998	2.792386
11	1	-3.706048	4.648009	4.923963
12	6	-2.342998	1.175279	2.196800
13	7	-1.294938	0.319110	1.769591
14	1	-1.639142	-0.279052	1.026040
15	6	0.088206	0.520358	1.749429
16	6	0.861185	-0.464067	1.045815
17	6	2.213472	-0.267739	1.008072
18	1	0.404687	-1.323151	0.576914
19	6	1.976027	1.765939	2.332781
20	1	2.884544	-0.949083	0.500811
21	8	-3.366383	1.178890	1.513154
22	7	0.608771	1.571433	2.343958
23	7	2.779800	0.806170	1.619549
24	8	2.547598	2.703824	2.882686
25	6	4.229327	1.045798	1.601821
26	1	4.435829	2.018978	1.154017
27	1	4.613347	1.062087	2.622985
28	1	4.714290	0.255710	1.028069

HAC 61iv C, 6-31G; N/O, 6-31G*; H, 6-31++G**



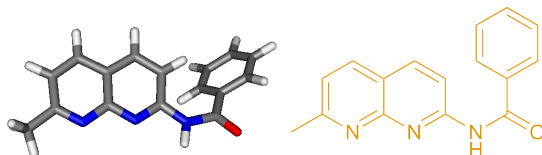
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-6.766831	1.003679	1.812065
2	6	-5.803122	1.009895	0.796265
3	6	-4.443987	1.092618	1.119880
4	6	-4.040074	1.168184	2.465056
5	6	-5.014055	1.183825	3.477663
6	6	-6.369581	1.095964	3.153408
7	1	-7.819068	0.938208	1.560405
8	1	-6.107485	0.962407	-0.242751
9	1	-3.710057	1.143855	0.323175
10	1	-4.685049	1.270889	4.505435
11	1	-7.113905	1.101207	3.940856
12	6	-2.601000	1.274355	2.888276
13	7	-1.696136	0.724672	1.958578
14	1	-2.099364	0.088668	1.285919
15	6	-0.297780	0.774352	1.976059
16	6	0.382850	-0.185516	1.151515
17	6	1.748446	-0.124118	1.141498
18	1	-0.147388	-0.930882	0.576601
19	6	1.689149	1.758744	2.693031
20	1	2.356865	-0.803677	0.558245
21	8	-2.268800	1.760233	3.957420
22	7	0.308938	1.691935	2.692345
23	7	2.407089	0.805291	1.881933
24	8	2.344585	2.578209	3.330398
25	6	3.872620	0.901693	1.897719
26	1	4.180394	1.891757	1.557757
27	1	4.238772	0.773090	2.917468
28	1	4.288305	0.132067	1.246645

ANT 62i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



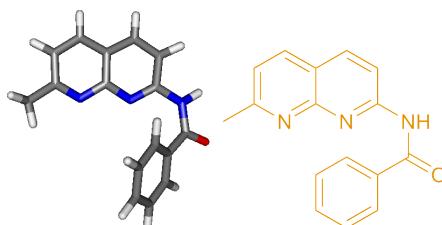
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-4.549589	-2.389591	-0.994584
2	6	-2.959022	-1.113396	0.099405
3	6	-3.938243	-0.151057	0.508854
4	6	-5.284248	-0.368025	0.117586
5	6	-5.591624	-1.481456	-0.631207
6	6	-3.497161	0.955680	1.280752
7	1	-6.049752	0.342407	0.411751
8	1	-6.607886	-1.680730	-0.948850
9	6	-2.167657	1.080566	1.611813
10	6	-1.274984	0.066136	1.150205
11	1	-4.215094	1.699891	1.608123
12	1	-1.783088	1.895476	2.202575
13	6	-4.874271	-3.616258	-1.815666
14	1	-5.592596	-4.262118	-1.297770
15	1	-5.320518	-3.343182	-2.778745
16	1	-3.960883	-4.182233	-1.998623
17	7	0.105775	0.089194	1.410395
18	1	0.573736	-0.721384	1.025064
19	6	0.836362	0.973578	2.184862
20	6	2.314340	0.712591	2.227380
21	6	3.005403	-0.001004	1.231536
22	6	3.029075	1.242347	3.315245
23	6	4.386343	-0.198904	1.336879
24	1	2.487542	-0.372369	0.354580
25	6	4.406730	1.039582	3.421479
26	1	2.485138	1.808272	4.060485
27	6	5.088127	0.314490	2.434473
28	1	4.911492	-0.744163	0.561560
29	1	4.947655	1.444937	4.268438
30	1	6.157525	0.157708	2.515594
31	8	0.326825	1.904957	2.814949
32	7	-1.643754	-0.985009	0.428451
33	7	-3.282059	-2.214955	-0.644445

ANT 62ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



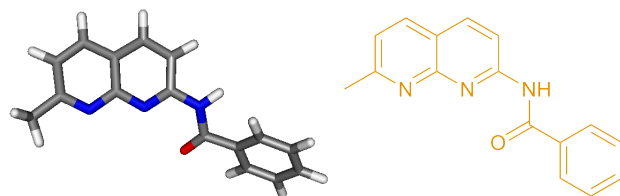
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-5.093818	-0.350378	0.136919
2	6	-3.277291	0.772179	1.029734
3	6	-4.081254	1.352340	2.063419
4	6	-5.462394	1.032652	2.089689
5	6	-5.969947	0.185568	1.130341
6	6	-3.435860	2.195351	3.006199
7	1	-6.097605	1.455454	2.861109
8	1	-7.019361	-0.082357	1.117145
9	6	-2.087287	2.439332	2.897093
10	6	-1.380732	1.839904	1.807532
11	1	-4.008242	2.627988	3.819569
12	1	-1.569601	3.051083	3.619146
13	7	-3.797551	-0.070410	0.087139
14	7	-1.943373	1.030796	0.920369
15	7	0.005037	1.997498	1.625392
16	1	0.426028	1.233193	1.106879
17	6	-5.636784	-1.283693	-0.920294
18	1	-4.827706	-1.589388	-1.583483
19	1	-6.082384	-2.177844	-0.469623
20	1	-6.418206	-0.797779	-1.515720
21	6	0.941823	2.864649	2.186725
22	6	0.517815	4.235020	2.615169
23	6	-0.439274	4.979431	1.902471
24	6	1.193633	4.827588	3.695228
25	6	-0.730092	6.293668	2.282814
26	1	-0.938583	4.540321	1.047770
27	6	0.888405	6.134783	4.084456
28	1	1.956865	4.255696	4.207723
29	6	-0.074757	6.869596	3.379291
30	1	-1.460214	6.866634	1.723405
31	1	1.404799	6.581605	4.925821
32	1	-0.305407	7.886383	3.675716
33	8	2.121240	2.514662	2.249020

ANT 62iii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



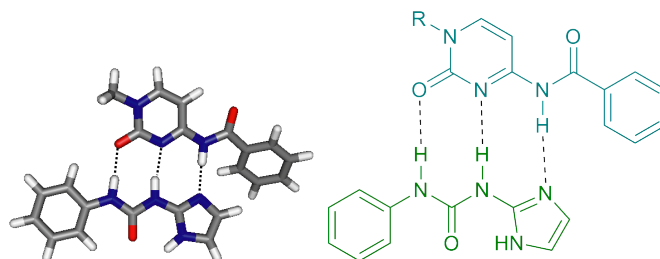
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-4.885527	0.253676	-0.859818
2	6	-3.005969	1.316628	-0.024592
3	6	-3.802435	2.209435	0.763316
4	6	-5.212050	2.075706	0.699323
5	6	-5.755133	1.102157	-0.109379
6	6	-3.126225	3.173936	1.559377
7	1	-5.842093	2.737554	1.284539
8	1	-6.827495	0.969478	-0.184777
9	6	-1.754252	3.212861	1.549819
10	6	-1.050489	2.279169	0.720744
11	1	-3.698144	3.863535	2.170842
12	1	-1.202371	3.919308	2.157540
13	7	-3.562411	0.355019	-0.819821
14	7	-1.641848	1.378647	-0.041270
15	7	0.354334	2.366971	0.729189
16	1	0.717211	3.295825	0.914921
17	6	-5.464564	-0.818186	-1.753918
18	1	-4.653552	-1.393032	-2.200940
19	1	-6.115679	-1.497510	-1.192445
20	1	-6.068945	-0.379928	-2.556647
21	6	1.308876	1.622403	0.006769
22	6	1.166292	0.146806	-0.132976
23	6	0.483352	-0.643622	0.806080
24	6	1.839904	-0.471485	-1.199577
25	6	0.465248	-2.033940	0.670950
26	1	-0.025564	-0.175335	1.637679
27	6	1.806588	-1.860425	-1.343284
28	1	2.382033	0.149551	-1.900900
29	6	1.118883	-2.644397	-0.407786
30	1	-0.060136	-2.638673	1.400791
31	1	2.316383	-2.329195	-2.176813
32	1	1.094580	-3.722839	-0.516067
33	8	2.303443	2.220799	-0.406692

ANT 62iv C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-4.657492	-0.164523	-0.103801
2	6	-2.659928	0.979458	0.150020
3	6	-3.333416	2.244472	0.168505
4	6	-4.745318	2.250292	0.035648
5	6	-5.407924	1.051647	-0.101569
6	6	-2.542732	3.414265	0.321750
7	1	-5.282233	3.193355	0.043315
8	1	-6.485445	1.014847	-0.206727
9	6	-1.179407	3.292438	0.433267
10	6	-0.605903	1.980527	0.400464
11	1	-3.019261	4.388579	0.340023
12	1	-0.542024	4.163227	0.531528
13	7	-3.337229	-0.201100	0.017463
14	7	-1.305878	0.869042	0.278052
15	7	0.790613	1.888694	0.557458
16	1	1.248301	2.723679	0.893573
17	6	-5.370703	-1.489031	-0.249204
18	1	-4.638558	-2.296230	-0.234621
19	1	-6.087750	-1.645960	0.564858
20	1	-5.932204	-1.536891	-1.189251
21	6	1.640100	0.905995	0.026627
22	6	3.070863	1.036164	0.471966
23	6	3.445675	1.645205	1.682955
24	6	4.061645	0.488763	-0.360163
25	6	4.796076	1.726402	2.042119
26	1	2.688662	2.015498	2.365349
27	6	5.409078	0.574546	-0.002733
28	1	3.750706	-0.001020	-1.274276
29	6	5.779699	1.198416	1.196876
30	1	5.076822	2.188900	2.981209
31	1	6.167639	0.156196	-0.653822
32	1	6.825319	1.262751	1.475334
33	8	1.275905	0.039312	-0.755233

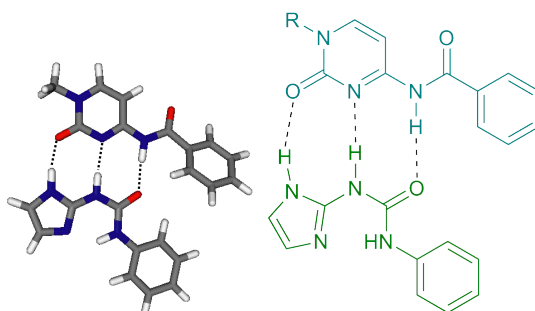
HAC · UIM 61 · 17i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	6.736628	-1.826606	0.245908
2	6	5.341996	-1.872739	0.141008
3	6	4.603605	-0.673928	0.172974
4	6	5.286036	0.552978	0.310761
5	6	6.677395	0.580172	0.414360
6	6	7.415198	-0.610426	0.382603
7	1	7.293697	-2.756716	0.219458
8	1	4.823036	-2.811068	0.032371
9	1	4.716088	1.474576	0.334452
10	1	7.184138	1.533226	0.519697
11	1	8.495870	-0.589471	0.462524
12	7	3.195379	-0.607359	0.084634
13	1	2.792521	0.332174	0.122521
14	6	2.322792	-1.646608	-0.117085
15	7	0.990790	-1.238597	-0.134467
16	1	0.717287	-0.255948	-0.036320
17	8	2.645645	-2.840128	-0.267361
18	6	-0.061071	-2.102461	-0.391089
19	6	-2.045680	-2.887799	-0.724809
20	6	-1.186829	-3.946682	-0.888068
21	1	1.002303	-3.861438	-0.676308
22	1	-3.118365	-2.879256	-0.808269
23	1	-1.352282	-4.981284	-1.128752
24	7	-1.340300	-1.731266	-0.406877
25	7	0.084707	-3.427254	-0.670776
26	8	2.104703	2.112107	0.073587
27	6	0.914829	2.466978	-0.008550
28	6	-0.652471	4.291609	-0.184974
29	6	-1.392476	2.024945	-0.106120
30	7	-0.133117	1.585508	-0.018420
31	7	-2.361372	1.038372	-0.059739
32	1	-1.999209	0.061272	-0.079482
33	6	-3.743267	1.238113	-0.172195

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
34	6	-4.602549	0.072055	0.199271
35	6	-5.897989	0.020333	-0.347068
36	6	-4.201661	-0.905742	1.125677
37	6	-6.771221	-1.009032	0.007860
38	1	-6.199022	0.796075	-1.039355
39	6	-5.084516	-1.928383	1.490421
40	1	-3.213852	-0.868436	1.564859
41	6	-6.365833	-1.985715	0.929156
42	1	-7.763640	-1.049086	-0.425464
43	1	-4.771223	-2.675536	2.209853
44	1	-7.045993	-2.781503	1.210638
45	8	-4.246062	2.303727	-0.536846
46	1	-0.793018	5.362168	-0.250908
47	6	-1.704774	3.420212	-0.198734
48	1	-2.722192	3.751944	-0.295136
49	7	0.635031	3.854108	-0.091463
50	6	1.781397	4.775779	-0.078917
51	1	2.347722	4.647126	0.844465
52	1	2.444255	4.550413	-0.915339
53	1	1.412546	5.798232	-0.155129

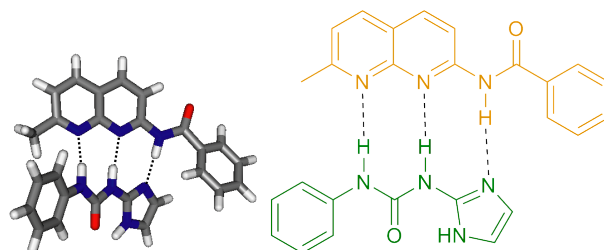
HAC · UIM 61 · 17ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	2.269140	5.322613	-0.114903
2	6	0.976219	4.880540	-0.396064
3	6	0.516306	3.652436	0.116553
4	6	1.364104	2.882421	0.931140
5	6	2.657150	3.333339	1.217984
6	6	3.113115	4.548845	0.694773
7	1	2.618060	6.264741	-0.521086
8	1	0.303523	5.469949	-1.005419
9	1	1.033732	1.936441	1.335090
10	1	3.304416	2.733147	1.846000

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
11	1	4.116908	4.892585	0.917252
12	6	-0.899372	3.282484	-0.187517
13	7	-1.241829	1.927980	-0.043021
14	1	-0.486993	1.224888	0.005664
15	6	-2.521346	1.402129	-0.092410
16	6	-4.898008	1.568961	-0.205632
17	1	-3.616842	3.281912	-0.312159
18	6	-3.802304	-0.569918	0.006275
19	8	-1.715534	4.138169	-0.538390
20	7	-2.590770	0.069699	0.008159
21	8	-3.933267	-1.804912	0.095101
22	6	3.825464	-0.390700	-0.421741
23	6	3.382824	-1.707719	-0.194205
24	6	4.332147	-2.745533	-0.092951
25	6	5.695216	-2.473704	-0.216785
26	6	6.138585	-1.163789	-0.441745
27	6	5.195833	-0.134358	-0.542086
28	1	3.101181	0.403499	-0.501634
29	1	3.992867	-3.760428	0.083938
30	1	6.409314	-3.285671	-0.136744
31	1	7.197025	-0.951648	-0.538019
32	1	5.523598	0.883893	-0.720485
33	7	2.026797	-2.079558	-0.071551
34	1	1.814454	-3.082248	0.002597
35	6	0.933957	-1.259742	-0.031296
36	7	-0.294038	-1.896464	0.078042
37	1	-1.093035	-1.254335	0.077043
38	8	0.962944	-0.009670	-0.080963
39	6	-0.545182	-3.262191	0.138138
40	6	-0.392536	-5.412568	0.220741
41	6	-1.735397	-5.123406	0.248363
42	1	-2.658785	-3.144116	0.180399
43	1	0.087124	-6.376206	0.245716
44	1	-2.607277	-5.750670	0.299277
45	7	0.352802	-4.241772	0.152300
46	7	-1.822688	-3.736761	0.195587
47	7	-4.979583	0.213002	-0.099562
48	6	-3.694313	2.215499	-0.209179
49	1	-5.835421	2.102354	-0.288489
50	6	-6.266384	-0.499960	-0.098044
51	1	-6.301187	-1.200195	-0.933684
52	1	-6.369421	-1.071755	0.825061
53	1	-7.072526	0.228208	-0.182425

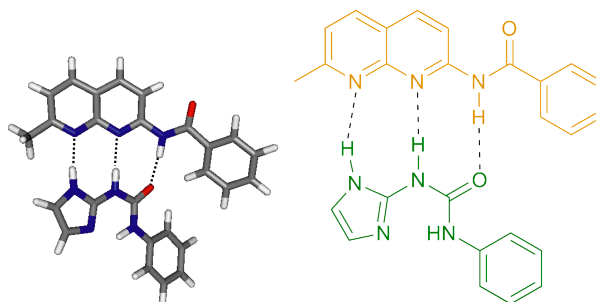
ANT · UIM 62 · 17i C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	3.089877	2.667847	0.860199
2	6	0.838101	2.439887	0.327816
3	6	0.679305	3.854229	0.199315
4	6	1.810717	4.675384	0.424265
5	6	3.008428	4.088122	0.767841
6	6	-0.612264	4.339461	-0.133727
7	1	1.717277	5.752485	0.335664
8	1	3.890116	4.684312	0.966428
9	6	-1.658226	3.464408	-0.288745
10	6	-1.413309	2.063429	-0.103778
11	1	-0.769629	5.405980	-0.251293
12	1	-2.657686	3.787533	-0.523946
13	6	4.400823	2.020605	1.240939
14	1	5.141485	2.115203	0.439102
15	1	4.825072	2.496967	2.130727
16	1	4.256448	0.960889	1.448084
17	7	-2.425235	1.105151	-0.202387
18	1	-2.190057	0.187694	0.215149
19	6	-3.751627	1.330447	-0.563388
20	6	-4.591124	0.103610	-0.753005
21	6	-4.059909	-1.136649	-1.145316
22	6	-5.982697	0.243011	-0.602463
23	6	-4.910289	-2.225156	-1.369330
24	1	-2.994036	-1.252442	-1.287712
25	6	-6.828055	-0.847520	-0.814325
26	1	-6.377717	1.212469	-0.327226
27	6	-6.292676	-2.085297	-1.198534
28	1	-4.493154	-3.176737	-1.676819
29	1	-7.898264	-0.734190	-0.686296
30	1	-6.948773	-2.931234	-1.369588
31	8	-4.236255	2.453368	-0.740209
32	7	-0.200868	1.571362	0.172560
33	7	2.049332	1.870227	0.625844
34	6	-2.621654	-2.384528	1.888402
35	6	-0.600674	-1.980158	1.250589

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	6	-1.909214	-3.532239	2.138945
37	1	-3.665736	-2.193734	2.067534
38	1	0.208902	-3.852060	1.657045
39	1	-2.191312	-4.482562	2.555068
40	7	-1.801264	-1.414707	1.323732
41	7	-0.607696	-3.255638	1.730376
42	7	0.519678	-1.330731	0.756816
43	1	0.372987	-0.340696	0.525022
44	6	1.706212	-1.987532	0.463290
45	7	2.661088	-1.143099	-0.059735
46	1	2.453376	-0.144087	-0.006945
47	8	1.857238	-3.206373	0.659330
48	6	3.900715	-1.509707	-0.629895
49	6	4.495310	-2.775758	-0.464263
50	6	4.563206	-0.535364	-1.404086
51	6	5.729557	-3.042868	-1.066506
52	1	3.985784	-3.525679	0.118852
53	6	5.799053	-0.813851	-1.990452
54	1	4.094391	0.430731	-1.554671
55	6	6.392350	-2.072040	-1.826313
56	1	6.175423	-4.022258	-0.933238
57	1	6.291381	-0.051268	-2.583744
58	1	7.348991	-2.291873	-2.285475

ANT · UIM 62 · 17ii C, 6-31G; N/O, 6-31G*; H, 6-31++G**

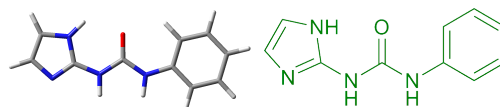


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	3.849086	-1.443542	-1.434568
2	6	3.283920	-2.384799	-0.553142
3	6	4.077853	-3.449274	-0.080638
4	6	5.410141	-3.570184	-0.478741
5	6	5.976500	-2.632048	-1.351537
6	6	5.187171	-1.576105	-1.821629
7	1	3.241775	-0.631952	-1.802132
8	1	3.644481	-4.174091	0.599975
9	1	6.004480	-4.396238	-0.104718

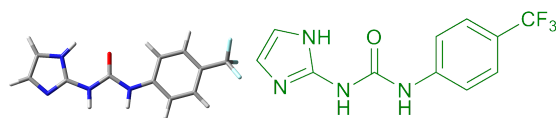
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
10	1	7.011316	-2.724241	-1.659842
11	1	5.610033	-0.846429	-2.503385
12	7	1.940062	-2.353116	-0.120608
13	1	1.586904	-3.166074	0.396776
14	6	1.047352	-1.326772	-0.260767
15	7	-0.227173	-1.560823	0.235700
16	1	-0.877391	-0.777506	0.104736
17	8	1.285705	-0.219586	-0.788385
18	6	-0.708920	-2.714603	0.836572
19	6	-0.951607	-4.653368	1.747642
20	6	-2.205701	-4.093771	1.704621
21	1	-2.743315	-2.164883	0.810700
22	1	-0.659807	-5.614190	2.136300
23	1	-3.164364	-4.451666	2.034928
24	7	-0.012910	-3.784811	1.203517
25	7	-2.037055	-2.841246	1.120764
26	6	-1.656594	2.078436	-0.284643
27	6	-3.409643	0.553582	-0.345263
28	6	-4.340218	1.565479	-0.740098
29	6	-3.846237	2.887869	-0.896145
30	6	-2.521869	3.157124	-0.666913
31	6	-5.687204	1.191298	-0.952386
32	1	-4.522258	3.678828	-1.201143
33	1	-2.112501	4.147906	-0.755491
34	6	-6.057650	-0.124116	-0.768329
35	6	-5.079550	-1.068697	-0.352211
36	1	-6.409042	1.940209	-1.259818
37	1	-7.077661	-0.448352	-0.928968
38	6	-5.476323	-2.506688	-0.103849
39	1	-6.388137	-2.768980	-0.645148
40	1	-4.679798	-3.185504	-0.415121
41	1	-5.666196	-2.677523	0.962848
42	7	-3.803538	-0.743133	-0.143293
43	7	-2.086367	0.822266	-0.137483
44	7	-0.287405	2.277717	-0.076221
45	1	0.280935	1.428146	-0.139320
46	6	0.300297	3.495143	0.278739
47	6	1.781105	3.515540	0.512621
48	6	2.302689	4.696226	1.076556
49	6	2.654229	2.459920	0.198257
50	6	3.669013	4.816539	1.331618
51	1	1.616906	5.501529	1.304228
52	6	4.024840	2.585705	0.455115
53	1	2.294446	1.542249	-0.246315
54	6	4.535443	3.758644	1.022286
55	1	4.057668	5.728587	1.769635
56	1	4.687333	1.763790	0.210556
57	1	5.597489	3.849147	1.220850
58	8	-0.370870	4.522590	0.425163

C.2 Molecular Modelling for Chapter 3: The Substituent Effect

UIM-H 17ai 6-31G*

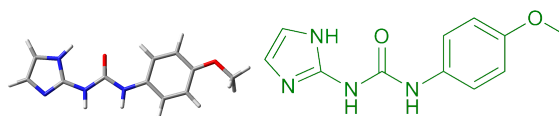


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.843967	1.399328	-0.000002
2	6	-2.496944	1.036434	-0.000002
3	6	-2.153102	-0.323883	0.000000
4	6	-3.169331	-1.292443	0.000002
5	6	-4.508502	-0.912450	0.000002
6	6	-4.856367	0.439077	0.000000
7	1	-4.099506	2.455657	-0.000003
8	1	-1.718191	1.785307	-0.000003
9	1	-2.907574	-2.349296	0.000003
10	1	-5.279007	-1.678560	0.000003
11	1	-5.900205	0.738708	0.000000
12	7	-0.821992	-0.796922	0.000000
13	1	-0.736748	-1.804167	0.000002
14	6	0.349552	-0.067665	-0.000002
15	7	1.483621	-0.859298	0.000000
16	1	1.434465	-1.869151	0.000002
17	8	0.408582	1.161311	-0.000004
18	6	2.792764	-0.376237	0.000000
19	1	2.421812	1.680631	0.000010
20	6	4.501273	1.004398	0.000005
21	6	4.920438	-0.299203	-0.000007
22	1	5.026304	1.946989	0.000009
23	1	5.935084	-0.672688	-0.000014
24	7	3.114527	0.942411	0.000006
25	7	3.843850	-1.167340	0.000000

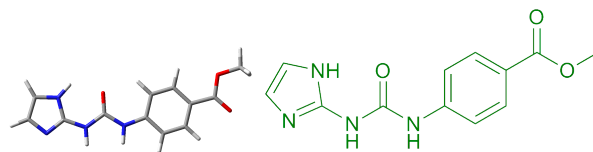
UIM-CF₃ 17bi 6-31G*

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	3.298507	0.029197	-0.038341
2	6	2.858026	1.354854	-0.025518
3	6	1.498383	1.636416	-0.016716
4	6	0.551107	0.598484	-0.018382
5	6	0.992258	-0.734650	-0.032737
6	6	2.357949	-1.003346	-0.041783
7	1	3.577489	2.166731	-0.030821
8	1	1.165981	2.672507	-0.011810
9	1	0.270140	-1.537857	-0.039639
10	1	2.693177	-2.035392	-0.059791
11	7	-0.804739	0.973718	-0.009380
12	1	-0.963596	1.972161	-0.000754
13	6	-1.921980	0.156284	-0.008093
14	7	-3.108929	0.858967	0.003495
15	1	-3.136729	1.869866	0.010169
16	8	-1.882687	-1.072751	-0.016216
17	6	-4.379434	0.279440	0.008027
18	6	-6.495272	0.045967	0.020324
19	6	-5.980267	-1.223214	0.008962
20	1	-3.859568	-1.746598	-0.008177
21	1	-7.534353	0.344345	0.028998
22	1	-6.436587	-2.200937	0.005760
23	7	-5.485065	0.990437	0.019763
24	7	-4.602883	-1.059769	0.000972
25	6	4.766432	-0.281439	0.012110
26	9	5.508812	0.701127	-0.548357
27	9	5.060978	-1.432466	-0.633455
28	9	5.210499	-0.425800	1.284559

UIM-OMe 17ci 6-31G*

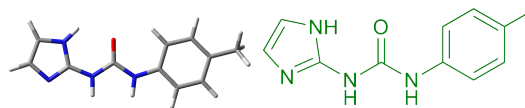


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.967629	1.299157	-0.000196
2	6	-1.622690	0.952583	-0.000231
3	6	-1.251675	-0.402951	-0.000082
4	6	-2.255141	-1.376883	0.000092
5	6	-3.606349	-1.027409	0.000119
6	6	-3.971934	0.322106	-0.000019
7	1	-3.263708	2.343597	-0.000305
8	1	-0.858196	1.716173	-0.000363
9	1	-1.985858	-2.431817	0.000198
10	1	-4.352771	-1.813347	0.000248
11	7	0.090605	-0.853106	-0.000142
12	1	0.193742	-1.858545	0.000003
13	6	1.244620	-0.102029	-0.000087
14	8	1.278124	1.129109	-0.000067
15	7	2.396977	-0.868820	-0.000065
16	1	2.369631	-1.879473	-0.000154
17	6	3.694809	-0.357759	0.000020
18	6	5.820482	-0.230261	-0.000009
19	6	5.370110	1.063074	0.000187
20	1	3.274079	1.688512	0.000481
21	1	6.843477	-0.580360	-0.000088
22	1	5.875046	2.016647	0.000326
23	7	4.764276	-1.123554	-0.000148
24	7	3.985955	0.968547	0.000286
25	8	-5.258780	0.784311	0.000003
26	6	-6.308405	-0.167743	0.000192
27	1	-7.236157	0.407476	0.000245
28	1	-6.278050	-0.804575	-0.894484
29	1	-6.277865	-0.804425	0.894966

UIM-CO₂Me 17di 6-31G*

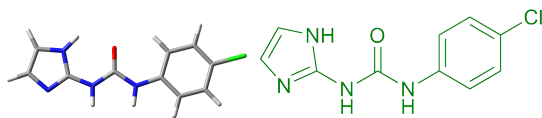
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	2.374321	-0.815155	-0.000072
2	6	0.998947	-0.604394	-0.000056
3	6	0.501486	0.709374	0.000013
4	6	1.403386	1.789283	0.000066
5	6	2.770920	1.564904	0.000059
6	6	3.275776	0.257428	-0.000015
7	1	2.754530	-1.830629	-0.000121
8	1	0.310729	-1.436980	-0.000092
9	1	1.024576	2.809608	0.000123
10	1	3.469569	2.394886	0.000110
11	7	-0.868304	1.027492	0.000016
12	1	-1.067342	2.018866	0.000083
13	6	-1.952603	0.167095	0.000077
14	7	-3.166275	0.824074	0.000040
15	1	-3.232694	1.833177	-0.000069
16	8	-1.867171	-1.059518	0.000063
17	6	-4.413036	0.195740	0.000000
18	6	-6.518150	-0.121284	-0.000104
19	6	-5.953344	-1.369151	0.000002
20	1	-3.813084	-1.807720	0.000182
21	1	-7.568299	0.135643	-0.000178
22	1	-6.370523	-2.364224	0.000047
23	7	-5.546163	0.862459	-0.000129
24	7	-4.583573	-1.151351	0.000100
25	6	4.749941	0.077706	-0.000018
26	8	5.113538	-1.227247	0.000046
27	8	5.556358	0.988862	-0.000078
28	6	6.528237	-1.466934	-0.000022
29	1	6.992413	-1.032598	0.889574
30	1	6.640725	-2.551656	-0.000644
31	1	6.992520	-1.031531	-0.889030

UIM-Me 17ei 6-31G*



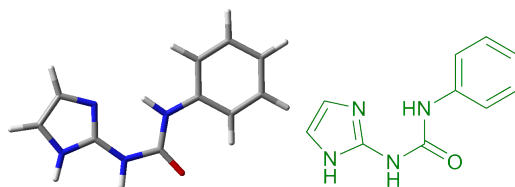
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	4.427495	-0.235706	-0.006687
2	6	4.015367	1.102314	-0.003123
3	6	2.666248	1.442623	-0.000421
4	6	1.678411	0.447094	-0.003638
5	6	2.070381	-0.899004	-0.011807
6	6	3.427969	-1.215881	-0.014220
7	1	4.759644	1.895392	-0.004028
8	1	2.376887	2.492419	-0.000574
9	1	1.319681	-1.676070	-0.019990
10	1	3.713255	-2.265705	-0.024063
11	7	0.332156	0.877452	-0.004398
12	1	0.215116	1.881511	-0.002307
13	6	-0.813300	0.110218	-0.002005
14	7	-1.974359	0.862428	-0.001148
15	1	-1.959650	1.873386	-0.002558
16	8	-0.830405	-1.120638	-0.000967
17	6	-3.265858	0.334846	0.001302
18	6	-4.922627	-1.107230	0.004899
19	1	-2.818898	-1.705894	0.002402
20	6	-5.389567	0.180301	0.004741
21	1	-5.415379	-2.067147	0.006349
22	1	-6.416942	0.517253	0.006085
23	7	-3.539946	-0.995132	0.002857
24	7	-4.344849	1.086895	0.002386
25	6	5.891664	-0.606794	0.022135
26	1	6.248828	-0.755605	1.050291
27	1	6.079654	-1.538559	-0.522417
28	1	6.512939	0.175927	-0.426387

UIM-Cl 17gi 6-31G*

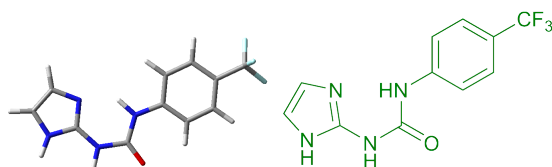


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.035936	-1.126357	-0.000014
2	6	-1.675884	-0.820147	-0.000012
3	6	-1.263813	0.521114	0.000003
4	6	-2.234470	1.535168	0.000016
5	6	-3.591079	1.227379	0.000014
6	6	-3.986995	-0.108682	-0.000001
7	1	-3.353003	-2.163889	-0.000026
8	1	-0.936876	-1.608243	-0.000023
9	1	-1.928430	2.579530	0.000028
10	1	-4.332400	2.019113	0.000024
11	7	0.086734	0.926833	0.000005
12	1	0.224008	1.928322	0.000017
13	6	1.218394	0.134018	-0.000006
14	7	2.392750	0.860026	0.000001
15	1	2.400504	1.871215	0.000009
16	8	1.204996	-1.096374	-0.000021
17	6	3.673913	0.305489	-0.000001
18	6	5.303049	-1.166889	0.000033
19	1	3.191459	-1.729124	0.000079
20	6	5.793968	0.111752	-0.000053
21	1	5.777801	-2.135818	0.000059
22	1	6.827298	0.429700	-0.000108
23	7	4.766195	1.037253	-0.000005
24	7	3.922701	-1.029281	0.000042
25	17	-5.699757	-0.509453	-0.000003

UIM-H 17aii 6-31G*

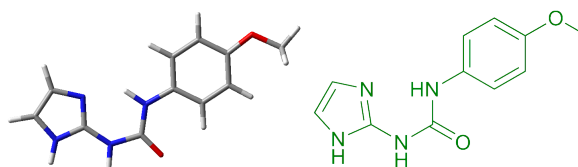


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-4.054940	0.565695	0.000008
2	6	-4.448697	-0.772775	0.000012
3	6	-3.469168	-1.767902	0.000009
4	6	-2.119841	-1.427052	0.000003
5	6	-1.726112	-0.078506	-0.000001
6	6	-2.706866	0.925508	0.000002
7	1	-4.805983	1.351680	0.000010
8	1	-3.753451	-2.817013	0.000012
9	1	-1.360622	-2.205921	0.000000
10	1	-2.405587	1.963623	-0.000002
11	7	-0.339537	0.174891	-0.000009
12	1	0.281121	-0.639032	-0.000010
13	6	0.280836	1.389750	-0.000014
14	7	1.693782	1.327469	-0.000017
15	1	2.098880	2.252905	-0.000015
16	8	-0.246864	2.494574	-0.000010
17	6	2.532777	0.228171	-0.000017
18	6	4.422554	-0.934357	0.000045
19	1	4.421550	1.215170	0.000068
20	6	3.341414	-1.765235	-0.000054
21	1	5.486453	-1.112225	0.000081
22	1	3.326036	-2.845722	-0.000108
23	7	3.893390	0.355092	0.000036
24	7	2.163364	-1.037744	-0.000001
25	1	-5.502228	-1.037792	0.000018

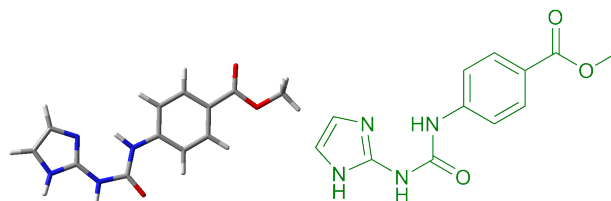
UIM-CF₃ 17bii 6-31G*

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.467335	1.145635	-0.029763
2	6	-3.007931	-0.142242	-0.041440
3	6	-2.148446	-1.244829	-0.043587
4	6	-0.773221	-1.056782	-0.034857
5	6	-0.227119	0.239082	-0.021674
6	6	-1.090070	1.347651	-0.021045
7	1	-3.128983	2.005926	-0.035378
8	1	-2.556002	-2.250274	-0.060234
9	1	-0.109626	-1.917798	-0.041508
10	1	-0.676133	2.345808	-0.016684
11	7	1.173055	0.333439	-0.013438
12	1	1.701769	-0.544537	-0.014935
13	6	1.926309	1.475513	-0.001064
14	7	3.320228	1.255225	0.007716
15	1	3.829041	2.128061	0.017213
16	8	1.518954	2.629025	0.003023
17	6	4.027033	0.064996	0.007674
18	6	5.765417	-1.311829	0.016658
19	1	6.018563	0.822296	0.031474
20	6	4.594359	-2.010024	0.001400
21	1	6.800931	-1.613653	0.024885
22	1	4.451756	-3.081045	-0.006018
23	7	5.392079	0.030901	0.020622
24	7	3.510715	-1.148253	-0.004243
25	6	-4.493244	-0.344761	0.011527
26	9	-4.869377	-1.491219	-0.603066
27	9	-5.166131	0.669937	-0.578285
28	9	-4.952270	-0.423232	1.285731

UIM-OMe 17cii 6-31G*

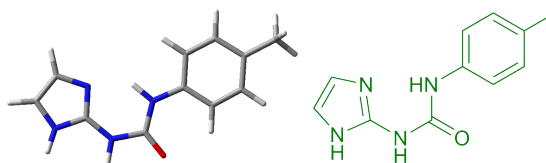


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.194134	0.783473	0.000073
2	6	-3.640738	-0.540474	-0.000065
3	6	-2.698041	-1.577701	-0.000901
4	6	-1.341720	-1.289751	-0.001637
5	6	-0.885221	0.040152	-0.001494
6	6	-1.826386	1.075102	-0.000595
7	1	-3.897590	1.608570	0.000781
8	1	-3.050693	-2.604335	-0.000846
9	1	-0.621310	-2.104567	-0.002208
10	1	-1.488780	2.102176	-0.000353
11	7	0.513476	0.235040	-0.002472
12	1	1.099586	-0.603669	-0.003135
13	6	1.179895	1.421984	-0.000785
14	7	2.590942	1.305843	-0.002874
15	1	3.031246	2.214931	0.001996
16	8	0.695572	2.547632	0.001883
17	6	3.385632	0.174863	-0.000431
18	6	5.227931	-1.062319	0.002385
19	1	5.312188	1.085719	0.000215
20	6	4.114575	-1.849363	0.001187
21	1	6.283974	-1.282104	0.004174
22	1	4.056394	-2.928426	0.001612
23	7	4.750306	0.247337	0.001500
24	7	2.966228	-1.075761	-0.000550
25	8	-4.954555	-0.927904	0.000451
26	6	-5.945052	0.084453	0.002440
27	1	-6.905901	-0.434278	0.003274
28	1	-5.875992	0.718596	0.896935
29	1	-5.878366	0.719833	-0.891396

UIM-CO₂Me 17dii 6-31G*

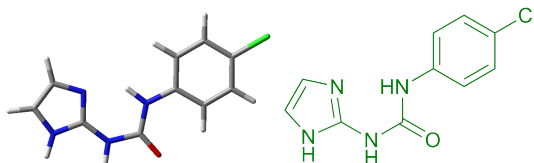
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.446113	1.005108	0.000076
2	6	-2.952748	-0.301687	-0.000062
3	6	-2.051465	-1.376307	-0.000197
4	6	-0.684872	-1.147713	-0.000193
5	6	-0.179646	0.166513	-0.000050
6	6	-1.076523	1.248917	0.000092
7	1	-3.134665	1.842827	0.000162
8	1	-2.444638	-2.387571	-0.000310
9	1	0.007086	-1.986467	-0.000312
10	1	-0.692673	2.259103	0.000202
11	7	1.216274	0.302971	-0.000066
12	1	1.769333	-0.559758	-0.000256
13	6	1.937895	1.465733	-0.000032
14	7	3.337634	1.283916	-0.000088
15	1	3.821663	2.170776	-0.000166
16	8	1.499962	2.607828	-0.000014
17	6	4.078314	0.114564	-0.000038
18	6	5.856314	-1.210817	-0.000061
19	1	6.047099	0.929875	-0.000520
20	6	4.705963	-1.942905	0.000356
21	1	6.900225	-1.482338	-0.000229
22	1	4.594477	-3.017639	0.000617
23	7	5.444028	0.120454	-0.000218
24	7	3.597651	-1.113128	0.000510
25	6	-4.404546	-0.605771	-0.000127
26	8	-4.879229	-1.726679	-0.000268
27	8	-5.170996	0.513899	0.000108
28	6	-6.585810	0.283444	0.000289
29	1	-7.043799	1.273409	0.000730
30	1	-6.885988	-0.277909	-0.888913
31	1	-6.885707	-0.278651	0.889096

UIM-Me 17eii 6-31G*



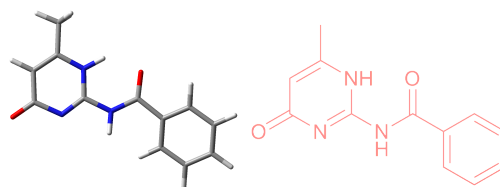
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.583174	0.841583	-0.006802
2	6	-4.071314	-0.469995	-0.007658
3	6	-3.129525	-1.507024	-0.011888
4	6	-1.763797	-1.243887	-0.008447
5	6	-1.291123	0.077831	-0.003099
6	6	-2.218578	1.129658	-0.003512
7	1	-4.287815	1.670790	-0.010442
8	1	-3.468014	-2.541000	-0.019630
9	1	-1.053325	-2.067607	-0.014841
10	1	-1.866993	2.151947	-0.005571
11	7	0.107949	0.253799	-0.002875
12	1	0.682262	-0.593285	-0.003440
13	6	0.793385	1.431982	-0.000286
14	7	2.201739	1.292687	0.000256
15	1	2.656700	2.194606	0.002798
16	8	0.327282	2.564394	0.001335
17	6	2.978860	0.149171	0.001338
18	6	4.801692	-1.116111	0.004551
19	1	4.919207	1.030262	0.006202
20	6	3.676330	-1.885954	0.001584
21	1	5.854206	-1.352221	0.006720
22	1	3.601451	-2.963975	0.000779
23	7	4.344379	0.200672	0.004358
24	7	2.540160	-1.094664	-0.000466
25	6	-5.554282	-0.758339	0.020999
26	1	-5.922915	-0.876712	1.049320
27	1	-6.128928	0.054250	-0.436641
28	1	-5.794512	-1.683149	-0.515528

UIM-Cl 17gii 6-31G*



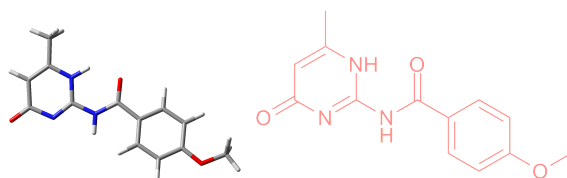
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	3.169337	0.999477	-0.000003
2	6	3.657958	-0.304556	-0.000002
3	6	2.782085	-1.388686	0.000002
4	6	1.410754	-1.155943	0.000005
5	6	0.900512	0.152791	0.000003
6	6	1.794974	1.234640	-0.000001
7	1	3.859316	1.836792	-0.000007
8	1	3.166195	-2.403124	0.000003
9	1	0.725503	-2.000082	0.000008
10	1	1.410856	2.244954	-0.000001
11	7	-0.499516	0.288356	0.000008
12	1	-1.053195	-0.573427	0.000011
13	6	-1.216452	1.450926	0.000009
14	7	-2.617621	1.275054	0.000008
15	1	-3.098062	2.163803	0.000006
16	8	-0.774790	2.592615	0.000004
17	6	-3.362014	0.108689	0.000008
18	6	-5.145140	-1.210381	-0.000022
19	1	-5.327878	0.931055	-0.000033
20	6	-3.997457	-1.946600	0.000023
21	1	-6.189985	-1.478305	-0.000039
22	1	-3.890123	-3.021788	0.000048
23	7	-4.727932	0.119330	-0.000018
24	7	-2.885877	-1.121030	-0.000004
25	17	5.396763	-0.589112	-0.000006

AIC-H 57a 6-31G*

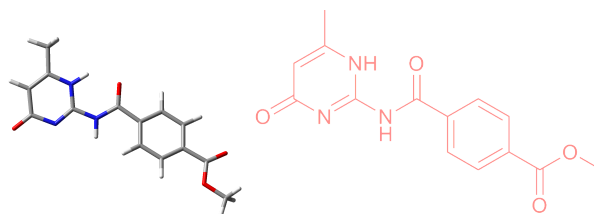


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	1.316013	-0.377024	-0.042106
2	6	3.236517	1.008522	0.086737
3	6	3.993986	-0.109156	-0.005953
4	6	3.385390	-1.442725	-0.127862
5	1	1.207304	1.656049	0.110404
6	1	5.075980	-0.049440	0.004980
7	7	1.971540	-1.487238	-0.135913
8	7	1.856770	0.870036	0.068755
9	7	-0.084429	-0.466962	-0.052381
10	1	-0.392171	-1.422185	-0.181414
11	8	4.047056	-2.467190	-0.217509
12	6	-1.008109	0.553005	0.022853
13	8	-0.673994	1.739836	0.073111
14	6	-2.445373	0.142233	0.023286
15	6	-2.884234	-1.144283	0.371569
16	6	-3.388567	1.119855	-0.325509
17	6	-4.244217	-1.449414	0.354224
18	1	-2.181506	-1.907739	0.694524
19	6	-4.745311	0.809652	-0.350014
20	1	-3.035963	2.114738	-0.575252
21	6	-5.175549	-0.476098	-0.012643
22	1	-4.575536	-2.445175	0.633963
23	1	-5.468251	1.570356	-0.630386
24	1	-6.234713	-0.717567	-0.028871
25	6	3.752776	2.411589	0.211029
26	1	4.844729	2.417573	0.211560
27	1	3.404513	3.035530	-0.621843
28	1	3.404025	2.878899	1.140623

AIC-OMe 57c 6-31G*

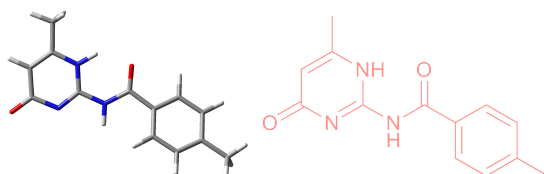


Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-4.356402	-0.370776	-0.047075
2	6	-3.426097	-1.390374	-0.312278
3	6	-2.069072	-1.112944	-0.312324
4	6	-1.599277	0.188006	-0.050003
5	6	-2.537956	1.198048	0.196678
6	6	-3.903566	0.932478	0.206196
7	1	-3.796526	-2.388259	-0.522472
8	1	-1.384682	-1.921890	-0.553161
9	1	-2.175611	2.203028	0.384471
10	1	-4.601255	1.736573	0.408526
11	6	-0.163321	0.576137	-0.034390
12	7	0.752257	-0.454948	0.032982
13	1	0.437637	-1.410264	0.139879
14	8	0.190936	1.759691	-0.067806
15	6	2.151948	-0.378067	0.034608
16	6	4.211091	-1.464047	0.119022
17	6	4.086428	0.991366	-0.055847
18	1	2.061736	1.656637	-0.086967
19	6	4.833275	-0.134406	0.024289
20	1	5.915845	-0.084509	0.022952
21	7	2.705880	0.865406	-0.051170
22	7	2.797734	-1.495926	0.115245
23	8	4.863696	-2.495775	0.196745
24	6	4.616409	2.391610	-0.152234
25	1	5.708394	2.387451	-0.147184
26	1	4.276649	2.878930	-1.074864
27	1	4.269480	3.003677	0.689981
28	8	-5.658944	-0.746754	-0.065046
29	6	-6.657006	0.236358	0.184854
30	1	-7.611500	-0.287403	0.117228
31	1	-6.549320	0.669950	1.186886
32	1	-6.626160	1.036752	-0.564793

AIC-CO₂Me 57d 6-31G*

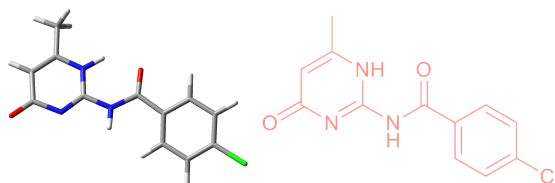
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.853410	0.884730	0.308525
2	6	-1.470868	0.724552	0.335325
3	6	-0.895065	-0.516080	0.021760
4	6	-1.725801	-1.598597	-0.303154
5	6	-3.105021	-1.437262	-0.338628
6	6	-3.677125	-0.194593	-0.035410
7	1	-3.296973	1.841236	0.559167
8	1	-0.855989	1.567513	0.638611
9	1	-1.269063	-2.556759	-0.525398
10	1	-3.759527	-2.262575	-0.597742
11	6	0.580228	-0.769659	0.040177
12	7	1.386895	0.339885	-0.074108
13	1	0.978240	1.250613	-0.240548
14	6	-5.166378	-0.080249	-0.087432
15	8	-5.597395	1.159036	0.228921
16	8	-5.909695	-0.994784	-0.380681
17	8	1.035988	-1.911194	0.139871
18	6	2.789780	0.403577	-0.057008
19	6	4.729969	1.686811	-0.177462
20	6	4.849774	-0.756530	0.128434
21	1	2.905738	-1.622988	0.174412
22	6	5.480199	0.433658	-0.003967
23	1	6.562089	0.493659	0.012100
24	8	5.274594	2.773993	-0.303621
25	7	3.462895	-0.771662	0.101640
26	7	3.319288	1.574625	-0.190950
27	6	5.516298	-2.088740	0.305668
28	1	6.602095	-1.973504	0.315109
29	1	5.210412	-2.560175	1.248131
30	1	5.249011	-2.774974	-0.507900
31	6	-7.023521	1.340409	0.198191
32	1	-7.509367	0.667111	0.908987
33	1	-7.190458	2.380953	0.476801
34	1	-7.412038	1.141391	-0.803841

AIC-Me 57e 6-31G*



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	4.287139	1.006494	-0.302555
2	6	2.922771	1.271238	-0.287470
3	6	2.004882	0.258696	0.027252
4	6	2.491042	-1.016892	0.350021
5	6	3.859947	-1.271192	0.340146
6	6	4.781501	-0.269891	0.004927
7	1	4.983424	1.802909	-0.553898
8	1	2.544154	2.261685	-0.516168
9	1	1.818400	-1.815679	0.651461
10	1	4.218968	-2.262398	0.605265
11	6	0.557824	0.622456	0.022754
12	7	-0.334505	-0.426748	-0.046556
13	1	0.002886	-1.373029	-0.166084
14	6	-1.736285	-0.380551	-0.038624
15	6	-3.699195	0.945880	0.075790
16	1	-1.690629	1.655776	0.098060
17	6	-3.771744	-1.510317	-0.118603
18	6	-4.421752	-0.195319	-0.008731
19	1	-5.505096	-0.169000	0.000156
20	8	0.184190	1.798341	0.065015
21	8	-4.401748	-2.555426	-0.200630
22	7	-2.357575	-1.511353	-0.124224
23	7	-2.316029	0.850001	0.061211
24	6	-4.258815	2.333350	0.187211
25	1	-5.350429	2.305123	0.190505
26	1	-3.922649	2.820884	1.111014
27	1	-3.932111	2.959290	-0.652896
28	6	6.261816	-0.560352	-0.040893
29	1	6.530519	-1.384742	0.627257
30	1	6.573592	-0.846197	-1.054312
31	1	6.851415	0.317145	0.244179

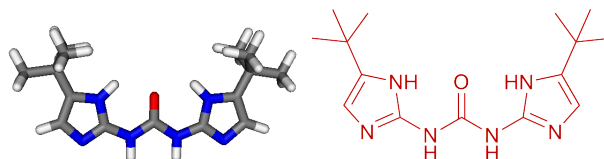
AIC-Cl 57f 6-31G*



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-3.884566	1.119602	-0.317872
2	6	-2.513060	1.347712	-0.299497
3	6	-1.617755	0.317688	0.022541
4	6	-2.124672	-0.948092	0.352105
5	6	-3.496624	-1.187140	0.345267
6	6	-4.366158	-0.150881	0.003358
7	1	-4.576124	1.914134	-0.575908
8	1	-2.114167	2.328247	-0.535353
9	1	-1.466133	-1.757902	0.653614
10	1	-3.889017	-2.162988	0.609012
11	17	-6.092480	-0.447978	-0.013580
12	6	-0.162178	0.657318	0.018001
13	7	0.711542	-0.404282	-0.053565
14	1	0.362060	-1.345507	-0.179298
15	8	0.224693	1.827926	0.063923
16	6	2.115747	-0.380469	-0.042441
17	6	4.130591	-1.546799	-0.120473
18	6	4.101523	0.909962	0.081950
19	1	2.108340	1.657468	0.101146
20	6	4.803277	-0.243990	-0.003853
21	1	5.886837	-0.237195	0.008752
22	8	4.740378	-2.602982	-0.204576
23	7	2.715589	-1.521641	-0.129749
24	7	2.716202	0.839320	0.062205
25	6	4.685694	2.286646	0.199518
26	1	5.776509	2.238278	0.208456
27	1	4.375175	2.920274	-0.640956
28	1	4.353856	2.778159	1.122761

C.3 Molecular Modelling for Chapter 4: A Novel ADDA array

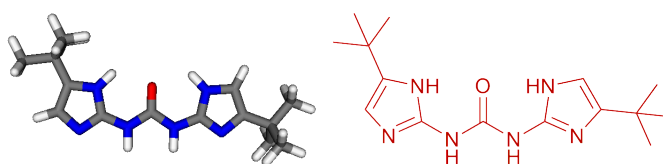
UDIM 1 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.970823	5.312410	1.843401
2	6	-2.813798	3.396679	1.101217
3	6	-2.997127	5.522188	0.943444
4	1	-1.241815	3.343749	2.479987
5	1	-3.393443	6.460652	0.597036
6	7	-3.024543	2.026079	0.914424
7	1	-3.767650	1.813853	0.262963
8	6	-2.309749	1.029969	1.539797
9	7	-2.693768	-0.247636	1.201424
10	1	-3.442006	-0.417647	0.543481
11	8	-1.388181	1.266112	2.348476
12	6	-2.113109	-1.415603	1.707924
13	6	-0.793123	-2.777178	2.863433
14	1	-0.669511	-0.589917	2.976492
15	6	-1.685833	-3.478913	2.077379
16	1	-1.789725	-4.545742	1.983238
17	7	-2.508599	-2.622533	1.357277
18	7	-1.087845	-1.434907	2.608636
19	7	-1.866934	3.921443	1.932420
20	7	-3.520443	4.320983	0.482872
21	6	-1.079085	6.250398	2.626759
22	6	-1.261330	6.010541	4.152113
23	1	-0.998609	4.984577	4.431114
24	1	-0.617762	6.687912	4.723658
25	1	-2.299110	6.185641	4.451040
26	6	-1.458218	7.712532	2.295971
27	1	-0.819261	8.401175	2.857210
28	1	-1.329413	7.922511	1.229751
29	1	-2.498799	7.920890	2.562954
30	6	0.297320	-3.207329	3.819753
31	6	0.355786	-4.751778	3.867981
32	1	0.576223	-5.169761	2.880977
33	1	1.141848	-5.075381	4.557087
34	1	-0.593340	-5.172428	4.214010
35	6	0.000526	-2.664922	5.246355

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	1	-0.045434	-1.570751	5.257936
37	1	-0.956452	-3.046153	5.614958
38	1	0.787387	-2.974205	5.942673
39	6	1.672341	-2.661147	3.341814
40	1	1.914599	-3.039671	2.344311
41	1	1.675799	-1.566846	3.297274
42	1	2.465400	-2.970385	4.031121
43	6	0.409804	6.012794	2.246953
44	1	0.721743	4.986839	2.469616
45	1	0.570702	6.189543	1.179305
46	1	1.059616	6.690108	2.811469

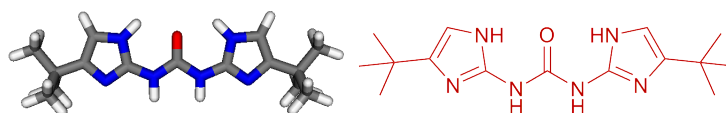
UDIM 2 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.740102	5.179481	1.727844
2	6	-2.369143	3.226485	0.878466
3	6	-2.484640	5.340939	0.576009
4	1	-1.226344	3.247269	2.630266
5	1	-2.764963	6.259803	0.091394
6	7	-2.534133	1.847523	0.707046
7	1	-3.075666	1.601235	-0.110119
8	6	-2.022605	0.886723	1.548732
9	7	-2.312738	-0.408667	1.180539
10	1	-2.859963	-0.612902	0.355800
11	8	-1.350865	1.164875	2.562708
12	6	-1.900459	-1.546280	1.881789
13	6	-0.962879	-2.845782	3.397465
14	1	-0.839508	-0.661200	3.454131
15	6	-1.613902	-3.606278	2.446709
16	7	-2.200574	-2.772424	1.497074
17	7	-1.155633	-1.518915	3.020681
18	7	-1.676329	3.794941	1.907819
19	7	-2.875165	4.116626	0.049418
20	6	-1.083973	6.157713	2.676893
21	6	-1.672424	5.990305	4.106283
22	1	-1.499447	4.981325	4.495609
23	1	-1.204649	6.699865	4.797276
24	1	-2.751648	6.170100	4.104842
25	6	-1.347487	7.600744	2.187186

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
26	1	-0.877351	8.318618	2.866157
27	1	-0.934750	7.759413	1.186208
28	1	-2.419757	7.816636	2.153931
29	6	0.450047	5.908542	2.723239
30	1	0.681717	4.897284	3.074278
31	1	0.892765	6.029749	1.730080
32	1	0.931606	6.617538	3.405283
33	6	-1.743238	-5.112961	2.340553
34	6	-3.246162	-5.501896	2.375339
35	1	-3.788124	-4.999906	1.569921
36	1	-3.364450	-6.584732	2.256645
37	1	-3.702093	-5.208538	3.326827
38	6	-1.007363	-5.801767	3.512015
39	1	0.060763	-5.558957	3.509125
40	1	-1.426338	-5.499831	4.478013
41	1	-1.103888	-6.889024	3.430191
42	6	-1.128476	-5.586396	0.995412
43	1	-1.618157	-5.086247	0.156086
44	1	-0.059127	-5.354109	0.953258
45	1	-1.250185	-6.669060	0.878631
46	1	-0.401051	-3.110154	4.274355

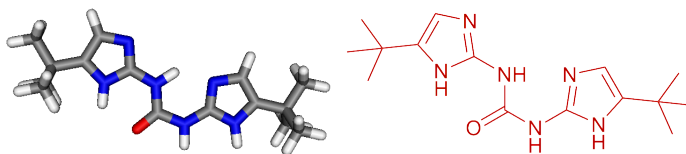
UDIM 3 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.327559	5.351656	2.345681
2	6	-2.438553	3.464441	1.209702
3	6	-2.632698	5.602575	1.022717
4	1	-1.990208	3.388389	3.252833
5	7	-2.409543	2.102496	0.891782
6	1	-2.606628	1.905339	-0.079633
7	6	-2.143195	1.094793	1.792099
8	7	-2.168797	-0.173137	1.254811
9	1	-2.370227	-0.329124	0.276826
10	8	-1.898428	1.312591	2.995045
11	6	-1.929562	-1.346874	1.977246
12	6	-1.472853	-2.728398	3.634702
13	1	-1.573206	-0.555873	3.882111
14	6	-1.676920	-3.431027	2.463817
15	7	-1.964169	-2.545703	1.426995

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
16	7	-1.638698	-1.385169	3.306346
17	7	-2.204912	3.968476	2.452253
18	7	-2.699197	4.402100	0.318622
19	6	-1.625230	-4.925485	2.215077
20	6	-3.003514	-5.404680	1.683582
21	1	-3.273797	-4.859037	0.775966
22	1	-2.973603	-6.475549	1.453204
23	1	-3.788089	-5.235292	2.428633
24	6	-1.288974	-5.678456	3.521776
25	1	-0.314177	-5.372810	3.916654
26	1	-2.045959	-5.495149	4.291807
27	1	-1.252920	-6.756566	3.336629
28	6	-0.537404	-5.230440	1.149442
29	1	-0.740886	-4.680075	0.227362
30	1	0.453771	-4.935548	1.509927
31	1	-0.515365	-6.301871	0.920796
32	1	-1.234071	-3.042671	4.634078
33	1	-2.190559	6.000041	3.191473
34	6	-2.879230	6.925365	0.324582
35	6	-2.746502	8.095912	1.324432
36	1	-1.743202	8.132447	1.762191
37	1	-2.925358	9.048174	0.815405
38	1	-3.474743	8.007988	2.137766
39	6	-4.308299	6.925663	-0.283320
40	1	-4.431797	6.084079	-0.969637
41	1	-5.065553	6.837343	0.502885
42	1	-4.490230	7.856020	-0.832927
43	6	-1.842540	7.102927	-0.818083
44	1	-1.899247	6.266145	-1.518882
45	1	-2.032344	8.032733	-1.365956
46	1	-0.824290	7.142224	-0.416884

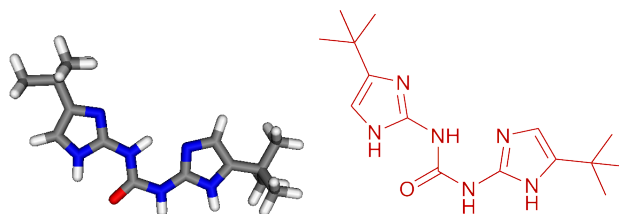
UDIM alt 1 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.924840	4.739878	1.691793
2	6	-2.663226	2.709269	1.168283
3	6	-2.839701	4.780039	0.656953
4	1	-1.265741	2.909628	2.709691
5	1	-3.198120	5.643329	0.123423

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
6	7	-2.829431	1.322563	1.190823
7	1	-3.483360	0.906626	0.518683
8	6	-2.170035	0.502797	2.056648
9	7	-2.455525	-0.866053	1.941992
10	1	-1.940736	-1.423937	2.607322
11	8	-1.354405	0.873796	2.918188
12	6	-3.328730	-1.484489	1.055231
13	6	-4.455023	-3.134887	0.040060
14	1	-3.071876	-3.515261	1.651434
15	6	-4.767207	-1.908014	-0.496754
16	1	-5.450180	-1.686416	-1.297676
17	7	-4.065464	-0.887405	0.136920
18	7	-3.519628	-2.843797	1.046864
19	7	-1.823754	3.385092	2.010188
20	7	-3.297402	3.510626	0.334029
21	6	-1.138289	5.810613	2.414692
22	6	-1.497610	5.807760	3.927443
23	1	-1.265339	4.844065	4.392826
24	1	-0.929924	6.581156	4.456576
25	1	-2.564644	6.002093	4.072029
26	6	-1.484476	7.193745	1.816231
27	1	-0.920815	7.977572	2.331763
28	1	-1.232367	7.236938	0.752230
29	1	-2.550891	7.414130	1.924458
30	6	-4.916219	-4.543170	-0.264550
31	6	-5.941308	-4.504097	-1.422207
32	1	-5.498104	-4.081393	-2.328831
33	1	-6.282789	-5.517716	-1.652189
34	1	-6.816161	-3.903527	-1.155708
35	6	-5.591062	-5.168669	0.988934
36	1	-4.896995	-5.229320	1.834657
37	1	-6.453493	-4.573905	1.303403
38	1	-5.933644	-6.185165	0.769185
39	6	-3.706298	-5.423026	-0.688611
40	1	-3.218142	-5.010549	-1.576300
41	1	-2.956429	-5.491152	0.107443
42	1	-4.038051	-6.440988	-0.917969
43	6	0.386700	5.554584	2.252296
44	1	0.676465	4.583176	2.666594
45	1	0.670820	5.567372	1.195723
46	1	0.961711	6.327001	2.774911

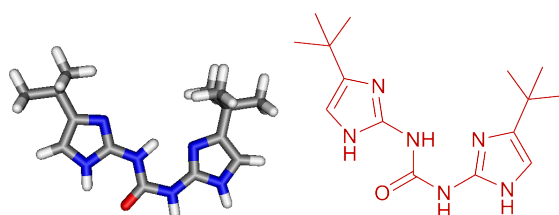
UDIM alt 2 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.192658	4.760225	2.183841
2	6	-2.824129	2.748817	1.524627
3	6	-3.069301	4.852118	1.120361
4	1	-1.489889	2.912602	3.127428
5	7	-2.941156	1.359412	1.454638
6	1	-3.554479	0.963167	0.733960
7	6	-2.282669	0.509451	2.293621
8	7	-2.514719	-0.858163	2.083715
9	1	-2.002954	-1.439446	2.731151
10	8	-1.513068	0.853698	3.205368
11	6	-3.332046	-1.448583	1.127303
12	6	-4.360529	-3.068762	-0.030120
13	1	-3.024898	-3.502847	1.607348
14	6	-4.697098	-1.821766	-0.502044
15	7	-4.055954	-0.819964	0.219853
16	7	-3.473391	-2.810198	1.027768
17	7	-2.043660	3.400229	2.433820
18	7	-3.458531	3.578905	0.715711
19	6	-3.601013	6.084141	0.415062
20	6	-3.018856	7.365885	1.052265
21	1	-3.293050	7.443323	2.109976
22	1	-1.926135	7.385168	0.978047
23	1	-3.406339	8.251839	0.538882
24	6	-5.149709	6.109902	0.526930
25	1	-5.559831	6.969031	-0.016226
26	1	-5.575464	5.193885	0.109893
27	1	-5.462075	6.183759	1.574092
28	6	-3.203347	6.027197	-1.084993
29	1	-2.114099	6.041662	-1.198658
30	1	-3.579836	5.109020	-1.542673
31	1	-3.617378	6.886409	-1.625042
32	1	-1.680001	5.504676	2.764986
33	1	-5.358002	-1.573366	-1.313518
34	6	-4.759469	-4.469786	-0.438582
35	6	-3.504250	-5.280391	-0.869235

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	1	-2.782290	-5.373992	-0.050284
37	1	-3.791340	-6.292198	-1.173623
38	1	-2.998674	-4.796898	-1.710120
39	6	-5.456145	-5.195804	0.747038
40	1	-6.350058	-4.651577	1.065160
41	1	-5.753808	-6.207096	0.451230
42	1	-4.790892	-5.287129	1.612932
43	6	-5.742617	-4.391548	-1.629948
44	1	-6.647298	-3.838481	-1.360158
45	1	-5.281947	-3.897697	-2.490767
46	1	-6.040037	-5.398863	-1.936364

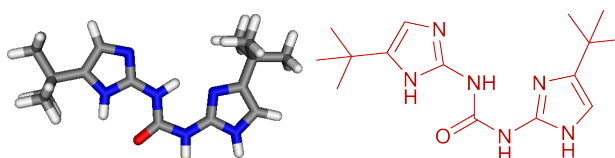
UDIM alt 3 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.191932	4.718608	2.105145
2	6	-2.818757	2.688171	1.503364
3	6	-3.142962	4.782048	1.105368
4	1	-1.380351	2.894728	3.007556
5	7	-2.906842	1.296013	1.449398
6	1	-3.559200	0.878776	0.776583
7	6	-2.171903	0.466595	2.244272
8	7	-2.387678	-0.907276	2.055912
9	1	-1.819445	-1.474059	2.668304
10	8	-1.349062	0.833748	3.098066
11	6	-3.256151	-1.520797	1.161649
12	6	-4.326875	-3.145871	0.083227
13	1	-2.875610	-3.569120	1.617419
14	6	-4.736693	-1.918359	-0.381066
15	7	-4.056458	-0.909852	0.304744
16	7	-3.375194	-2.881598	1.074702
17	7	-1.992601	3.364285	2.351535
18	7	-3.528448	3.496605	0.736273
19	1	-4.605367	-4.149177	-0.180984
20	6	-5.750401	-1.578149	-1.454400
21	6	-6.352297	-2.871057	-2.050113
22	1	-6.864597	-3.462263	-1.283384

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
23	1	-7.083844	-2.622250	-2.824832
24	1	-5.578626	-3.496383	-2.508642
25	6	-6.885247	-0.717719	-0.834539
26	1	-7.414574	-1.274201	-0.054061
27	1	-6.479978	0.193569	-0.387677
28	1	-7.609673	-0.433147	-1.605192
29	6	-5.053252	-0.766285	-2.580149
30	1	-4.601171	0.143766	-2.177900
31	1	-4.264083	-1.357719	-3.055985
32	1	-5.780137	-0.481655	-3.348462
33	6	-3.753223	5.995457	0.431670
34	6	-3.159372	7.295792	1.018594
35	1	-3.360162	7.374678	2.092534
36	1	-2.075366	7.340815	0.867586
37	1	-3.603109	8.168664	0.529026
38	6	-5.290237	5.983813	0.652398
39	1	-5.759940	6.827934	0.134547
40	1	-5.720949	5.054362	0.271671
41	1	-5.530018	6.057553	1.718478
42	6	-3.460419	5.935572	-1.092123
43	1	-2.382560	5.974523	-1.282271
44	1	-3.845893	5.004952	-1.516035
45	1	-3.931925	6.779767	-1.608217
46	1	-1.657602	5.479045	2.644539

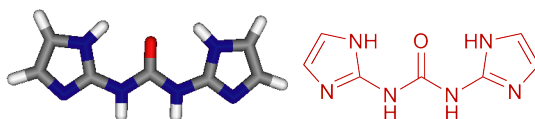
UDIM alt 4 C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.636149	-1.975271	0.387535
2	8	0.998252	-2.756814	-0.507789
3	7	0.724022	-2.213636	1.726101
4	1	0.391849	-1.491418	2.374857
5	7	0.086678	-0.735473	0.024425
6	1	0.044376	-0.609609	-0.976279
7	6	1.243532	-3.391524	2.266896
8	6	1.895926	-4.892028	3.646914
9	6	2.148529	-5.418009	2.394475
10	1	1.718419	-4.398108	0.498010
11	1	2.085043	-5.344143	4.605116

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
12	6	-0.374821	0.270389	0.864388
13	6	-1.252652	2.214072	1.499114
14	1	-0.998285	1.710793	-0.579720
15	6	-0.929822	1.457198	2.600571
16	1	-1.687750	3.190686	1.394499
17	7	-0.892754	1.444119	0.387008
18	7	-0.380855	0.242310	2.186237
19	7	1.717435	-4.427404	1.510937
20	7	1.331073	-3.627137	3.562204
21	6	-1.090588	1.770399	4.074019
22	6	-1.718928	3.170470	4.257904
23	1	-2.710449	3.226052	3.795695
24	1	-1.833255	3.392660	5.323139
25	1	-1.088389	3.951013	3.818570
26	6	-2.008087	0.702067	4.729301
27	1	-3.006802	0.716877	4.280534
28	1	-1.590476	-0.299421	4.598841
29	1	-2.110234	0.897402	5.802251
30	6	0.301543	1.734315	4.761913
31	1	0.966960	2.492875	4.336648
32	1	0.198894	1.929426	5.834856
33	1	0.772202	0.756536	4.632208
34	6	2.741362	-6.728646	1.926512
35	6	1.709709	-7.499237	1.055179
36	1	1.425241	-6.921735	0.169369
37	1	2.132192	-8.450865	0.714446
38	1	0.800404	-7.708953	1.626429
39	6	3.112070	-7.592579	3.154386
40	1	3.542467	-8.544481	2.827800
41	1	3.848208	-7.085655	3.785916
42	1	2.230101	-7.810042	3.764751
43	6	4.021932	-6.464156	1.085373
44	1	4.771294	-5.931052	1.678165
45	1	4.457841	-7.409863	0.745075
46	1	3.799941	-5.858954	0.200189

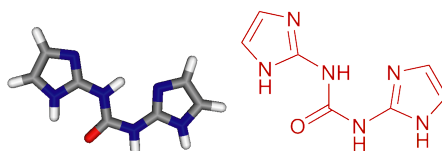
UDIM no tbu C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.412563	-0.691433	-0.514208
2	7	-0.676646	-0.675362	0.649083

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
3	1	-0.998902	-1.140678	1.486421
4	7	-2.599313	-1.384015	-0.431018
5	1	-2.889232	-1.837413	0.424489
6	6	0.558029	-0.035152	0.798638
7	6	2.379883	0.692060	1.658382
8	6	2.377411	1.134244	0.354876
9	1	0.806851	0.757923	-1.122696
10	1	3.128542	0.853628	2.415064
11	1	3.076044	1.719445	-0.215661
12	6	-3.507006	-1.533523	-1.484733
13	6	-4.444811	-1.380370	-3.477508
14	1	-2.516560	-0.467032	-2.989737
15	6	-5.230295	-2.112927	-2.616568
16	1	-4.565157	-1.097181	-4.507750
17	1	-6.181041	-2.577689	-2.814548
18	7	1.190559	0.655945	-0.191836
19	7	1.234934	-0.042007	1.932249
20	7	-4.636750	-2.206244	-1.366103
21	7	-3.330323	-1.010377	-2.731531
22	8	-1.034559	-0.124301	-1.559039

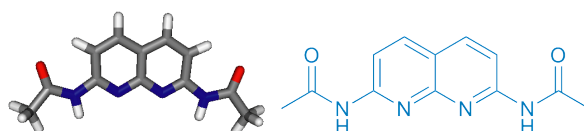
UDIM alt no tbu C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.142493	-0.554832	-0.471677
2	7	-0.238164	-0.410921	0.593230
3	1	0.573776	0.130957	0.336201
4	7	-2.267404	-1.275427	-0.204425
5	1	-2.393328	-1.665683	0.735360
6	6	-0.350792	-0.912774	1.883276
7	6	-1.027609	-1.902959	3.680886
8	6	0.178332	-1.338401	4.005028
9	1	1.467944	-0.182088	2.727290
10	1	-1.684731	-2.483069	4.305307
11	1	0.755291	-1.324021	4.911905
12	6	-3.267434	-1.513502	-1.150006
13	6	-4.384649	-1.489957	-3.057146
14	1	-2.450073	-0.498740	-2.787151
15	6	-5.065139	-2.199270	-2.092396

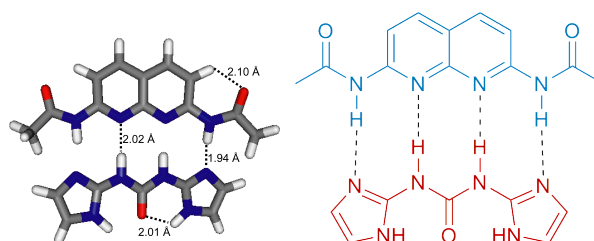
Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
16	1	-4.606608	-1.264467	-4.084913
17	1	-6.014112	-2.700596	-2.180505
18	7	0.609885	-0.700082	2.838477
19	7	-1.353317	-1.634853	2.356836
20	7	-4.361093	-2.211403	-0.898334
21	7	-3.222592	-1.051175	-2.434760
22	8	-0.868069	-0.029970	-1.563083

DAN C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.164319	5.205780	-5.629096
2	6	-0.573289	3.743762	-4.820588
3	6	0.385321	4.195073	-5.786591
4	6	-0.017017	5.208797	-6.693382
5	6	-1.290459	5.727457	-6.629820
6	6	1.671517	3.597390	-5.770611
7	1	0.685711	5.571166	-7.435614
8	1	-1.646188	6.497866	-7.294061
9	6	1.967447	2.618362	-4.849399
10	6	0.938905	2.244302	-3.933207
11	1	2.418342	3.917677	-6.488744
12	1	2.927877	2.132731	-4.794369
13	7	-1.832708	4.258663	-4.759736
14	7	-0.277937	2.774910	-3.910065
15	7	-3.486142	5.657086	-5.467391
16	1	-3.961741	5.179629	-4.711638
17	7	1.123514	1.257987	-2.948249
18	1	0.292103	1.120089	-2.386944
19	6	-4.181281	6.614777	-6.185202
20	6	2.240067	0.486739	-2.675983
21	8	3.301800	0.571187	-3.290408
22	8	-3.703756	7.256728	-7.118902
23	6	2.040754	-0.518573	-1.549790
24	1	1.193783	-0.281920	-0.901639
25	1	1.883982	-1.513729	-1.978284
26	1	2.952401	-0.556402	-0.951851
27	6	-5.625776	6.797745	-5.739503
28	1	-6.288340	6.285282	-6.444422
29	1	-5.822251	6.413608	-4.735860
30	1	-5.867871	7.860944	-5.772073

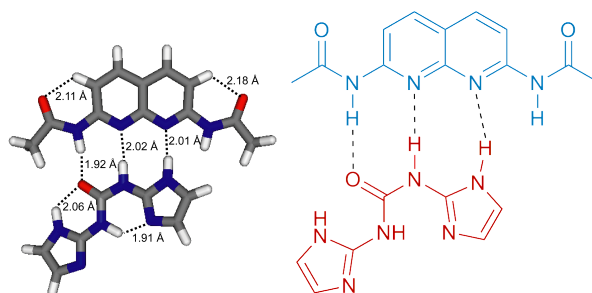
DAN · UDIM C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-0.000013	2.484629	-0.000274
2	7	1.124098	1.736466	0.251648
3	1	1.106954	0.708195	0.155492
4	7	-1.124090	1.736429	-0.252259
5	1	-1.106889	0.708178	-0.155974
6	6	2.352753	2.309510	0.556009
7	6	4.445903	2.545296	1.015994
8	6	3.905007	3.807480	1.059996
9	1	1.815421	4.321165	0.645510
10	1	5.465996	2.250737	1.192850
11	1	4.334062	4.771457	1.265709
12	6	-2.352885	2.309388	-0.556174
13	6	-3.905375	3.807250	-1.059759
14	1	-1.815761	4.321126	-0.645799
15	6	-4.446194	2.545044	-1.015519
16	1	-4.334560	4.771186	-1.265404
17	1	-5.466330	2.250422	-1.192003
18	7	2.556706	3.641029	0.766033
19	7	3.470198	1.609410	0.696369
20	7	-3.470321	1.609221	-0.696188
21	7	-2.556957	3.640889	-0.766263
22	8	-0.000067	3.737715	-0.000169
23	6	-2.315458	-1.968304	0.241566
24	6	0.000049	-2.014295	-0.000091
25	6	0.000080	-3.445136	-0.000103
26	6	-1.234574	-4.114823	0.171182
27	6	-2.392960	-3.394263	0.308902
28	6	1.234759	-4.114782	-0.171435
29	1	-1.254540	-5.198821	0.189619
30	1	-3.353557	-3.855739	0.460146
31	6	2.393115	-3.394179	-0.309102
32	6	2.315554	-1.968214	-0.241641
33	1	1.254730	-5.198779	-0.189939
34	1	3.353755	-3.855551	-0.460397
35	7	-1.163442	-1.296022	0.098000

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	7	1.163523	-1.295974	-0.098150
37	7	-3.473842	-1.187588	0.296612
38	1	-3.392807	-0.209297	-0.019251
39	7	3.473938	-1.187466	-0.296539
40	1	3.392862	-0.209241	0.019431
41	6	-4.753222	-1.605850	0.647121
42	6	4.753392	-1.605705	-0.646832
43	8	5.045640	-2.744792	-1.009679
44	8	-5.045391	-2.744949	1.009998
45	6	5.807510	-0.514456	-0.533888
46	1	5.450200	0.442702	-0.917855
47	1	6.081752	-0.370087	0.516739
48	1	6.690210	-0.836022	-1.084927
49	6	-5.807397	-0.514625	0.534415
50	1	-5.449982	0.442594	0.918128
51	1	-6.082011	-0.370420	-0.516136
52	1	-6.689890	-0.836158	1.085806

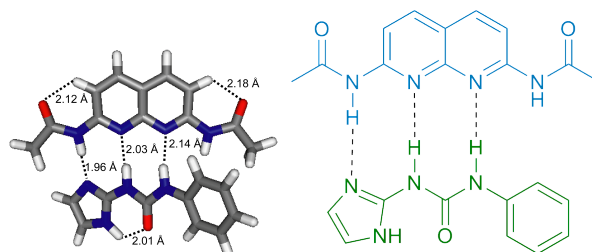
DAN · UDIM alt C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-2.432729	-0.634982	-0.163251
2	7	-1.100665	-0.884334	0.104532
3	1	-0.473043	-0.077763	-0.002506
4	7	-3.294245	-1.680056	-0.017028
5	1	-2.918655	-2.572591	0.328093
6	6	-0.555238	-2.066727	0.588500
7	6	-0.217761	-4.063537	1.325462
8	6	1.013065	-3.452540	1.309865
9	1	1.459187	-1.428234	0.602459
10	1	-0.466236	-5.063763	1.636431
11	1	1.988905	-3.793246	1.607313
12	6	-4.664529	-1.606850	-0.270916
13	6	-6.637957	-0.816778	-0.884617
14	1	-4.797025	0.366925	-0.949086
15	6	-6.735085	-2.125900	-0.467915

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
16	1	-7.366703	-0.113283	-1.245941
17	1	-7.614943	-2.744869	-0.418824
18	7	0.783744	-2.162035	0.836116
19	7	-1.198514	-3.191532	0.871764
20	7	-5.497036	-2.616785	-0.084929
21	7	-5.293037	-0.492724	-0.752486
22	8	-2.783421	0.516613	-0.529360
23	6	0.275750	2.765188	0.112040
24	6	2.110168	1.350439	-0.044716
25	6	3.023190	2.445957	0.062228
26	6	2.474153	3.744712	0.220386
27	6	1.112577	3.919674	0.248008
28	6	4.409793	2.164625	0.000765
29	1	3.137685	4.597181	0.312524
30	1	0.648986	4.884740	0.364942
31	6	4.850750	0.871627	-0.175627
32	6	3.865590	-0.151746	-0.275112
33	1	5.122339	2.977463	0.084699
34	1	5.893779	0.609716	-0.237300
35	7	0.757668	1.522636	-0.014044
36	7	2.552194	0.064865	-0.187192
37	7	-1.114722	2.856891	0.086133
38	1	-1.615839	1.984515	-0.124758
39	7	4.196215	-1.499934	-0.489421
40	1	3.391516	-2.089441	-0.659678
41	6	-1.897321	3.994555	0.271657
42	6	5.452898	-2.088074	-0.588902
43	8	6.506289	-1.477384	-0.436587
44	8	-1.447380	5.111887	0.518915
45	6	5.415546	-3.569437	-0.933000
46	1	4.459217	-4.044918	-0.703141
47	1	5.617223	-3.696971	-2.001578
48	1	6.211090	-4.074608	-0.384080
49	6	-3.388060	3.744311	0.115592
50	1	-3.677987	3.930262	-0.924670
51	1	-3.668562	2.721205	0.367958
52	1	-3.924848	4.452698	0.746598

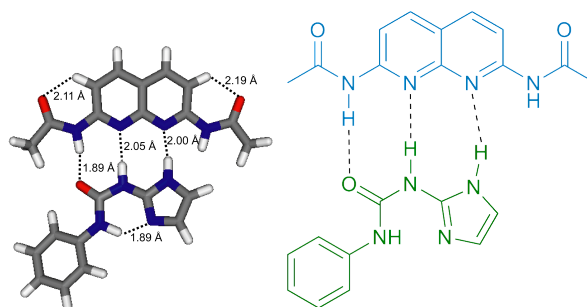
DAN · UIM C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.809594	-2.241089	-0.596593
2	6	0.438838	-2.040333	-0.063112
3	6	0.553930	-3.452613	0.128236
4	6	-0.602884	-4.245408	-0.070985
5	6	-1.790362	-3.658843	-0.446728
6	6	1.824103	-3.962125	0.500751
7	1	-0.543168	-5.319376	0.065967
8	1	-2.693159	-4.219378	-0.623234
9	6	2.895293	-3.113255	0.634007
10	6	2.695257	-1.718126	0.377795
11	1	1.943965	-5.026424	0.670075
12	1	3.877654	-3.451968	0.916796
13	7	3.745799	-0.802005	0.409481
14	1	3.544513	0.126729	0.001856
15	7	-2.962881	-1.543863	-0.980732
16	1	-2.820993	-0.545099	-1.067489
17	6	5.048857	-1.028357	0.837837
18	6	-4.243337	-2.028396	-1.230971
19	6	-5.228952	-0.941265	-1.630500
20	1	-5.268128	-0.147754	-0.877843
21	1	-4.941010	-0.487871	-2.585020
22	1	-6.215009	-1.390482	-1.735551
23	6	5.968602	0.169362	0.660020
24	1	6.374191	0.171308	-0.357904
25	1	5.450841	1.116881	0.816795
26	1	6.798571	0.071848	1.359144
27	8	5.454451	-2.091424	1.306958
28	8	-4.554089	-3.213047	-1.145394
29	7	-0.750994	-1.454979	-0.390896
30	7	1.502809	-1.196785	0.065041
31	6	-5.109878	2.413739	1.776122
32	6	-4.503098	1.171425	2.001201
33	6	-3.216085	0.917601	1.521089
34	6	-2.516071	1.902146	0.793479
35	6	-3.124465	3.152510	0.567249

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	6	-4.409449	3.395131	1.064299
37	1	-6.104696	2.614926	2.155468
38	1	-5.022565	0.403141	2.562812
39	1	-2.737812	-0.033616	1.725723
40	1	-2.584783	3.911732	0.024612
41	1	-4.864580	4.363177	0.887749
42	7	-1.230965	1.557629	0.314871
43	1	-1.008853	0.561771	0.272193
44	6	-0.240866	2.426178	-0.093784
45	8	-0.382361	3.655069	-0.219388
46	7	0.959173	1.783303	-0.357349
47	1	1.077803	0.772180	-0.209124
48	6	2.109016	2.462355	-0.732942
49	6	3.472947	4.085806	-1.377132
50	1	1.337393	4.380165	-0.960396
51	6	4.158627	2.903028	-1.242071
52	1	3.784483	5.073738	-1.664570
53	1	5.205204	2.716807	-1.413202
54	7	3.302950	1.886852	-0.832204
55	7	2.154617	3.786413	-1.051908

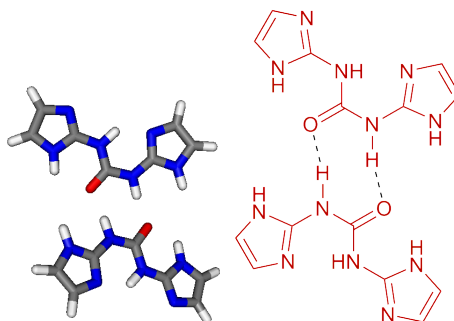
DAN · UIM alt C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	4.072755	-0.356424	-0.253839
2	6	2.424384	1.264589	-0.031278
3	6	3.412250	2.289226	0.110486
4	6	4.775641	1.909715	0.071360
5	6	5.126575	0.590876	-0.117733
6	6	2.954913	3.622118	0.278128
7	1	5.542745	2.667862	0.182498
8	1	6.149179	0.255632	-0.163830
9	6	1.609485	3.894185	0.282305
10	6	0.693376	2.805080	0.113275

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
11	1	3.676227	4.422988	0.396027
12	1	1.213247	4.888125	0.405102
13	7	-0.684580	2.998235	0.063548
14	1	-1.246404	2.163504	-0.153860
15	7	4.309663	-1.722069	-0.486355
16	1	3.467994	-2.248254	-0.682857
17	6	-1.385408	4.190544	0.238001
18	6	5.521658	-2.397092	-0.576386
19	6	5.384883	-3.865805	-0.949924
20	1	4.386883	-4.270650	-0.766812
21	1	5.619615	-3.990450	-2.011983
22	1	6.117204	-4.439757	-0.380726
23	6	-2.885701	4.054102	0.043741
24	1	-3.139110	4.332013	-0.985437
25	1	-3.243922	3.039332	0.217114
26	1	-3.385820	4.756198	0.711615
27	8	-0.858288	5.269942	0.503756
28	8	6.614039	-1.868077	-0.394363
29	7	2.776871	-0.047273	-0.187669
30	7	1.087723	1.532450	-0.023013
31	6	-7.418745	-1.210154	-0.488932
32	6	-6.773194	-2.336663	0.037281
33	6	-5.384646	-2.353029	0.175959
34	6	-4.614182	-1.238767	-0.213361
35	6	-5.259601	-0.106879	-0.745571
36	6	-6.652760	-0.104670	-0.875670
37	1	-8.497069	-1.195717	-0.595296
38	1	-7.348963	-3.203244	0.342119
39	1	-4.888699	-3.225109	0.588237
40	1	-4.670798	0.744515	-1.045874
41	1	-7.138979	0.772824	-1.287192
42	7	-3.214557	-1.350764	-0.053155
43	1	-2.846998	-2.253402	0.267873
44	6	-2.277949	-0.369298	-0.207925
45	8	-2.510284	0.814229	-0.533942
46	7	-0.957947	-0.752541	0.013853
47	1	-0.273891	0.007452	-0.073163
48	6	-0.496659	-1.968811	0.495779
49	6	0.977731	-3.458750	1.213878
50	1	1.555675	-1.460799	0.527322
51	6	-0.289477	-3.988836	1.225160
52	1	1.929057	-3.863281	1.510886
53	1	-0.601150	-4.972924	1.531264
54	7	-1.211732	-3.053198	0.772959
55	7	0.832883	-2.154436	0.744827

UDIM · UDIM C, 6-31G; N/O, 6-31G*; H, 6-31++G**



Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	-1.684491	-0.174824	-1.386715
2	7	-1.141975	-0.456679	-0.147674
3	1	-1.653179	-1.143576	0.425253
4	7	-0.969000	0.649758	-2.197318
5	1	-0.076861	1.019666	-1.842965
6	6	0.041002	0.048419	0.374592
7	6	1.909139	1.077493	0.709330
8	6	1.644499	0.400068	1.873059
9	1	-0.078329	-0.835851	2.298316
10	1	2.750644	1.707059	0.477517
11	1	2.176526	0.323832	2.804023
12	6	-1.356136	1.018545	-3.486910
13	6	-1.386226	1.934905	-5.424435
14	6	-2.484891	1.105135	-5.388184
15	1	-3.071978	-0.159017	-3.701618
16	1	-1.058062	2.569255	-6.230270
17	1	-3.254128	0.886724	-6.107280
18	8	-2.793954	-0.683950	-1.708655
19	7	-0.682974	1.877287	-4.232114
20	7	-2.454791	0.518312	-4.128588
21	7	0.437899	-0.261849	1.647640
22	7	0.904559	0.853569	-0.224396
23	6	-3.720626	-2.311242	1.524828
24	7	-4.346975	-1.763625	0.421749
25	1	-3.727953	-1.416838	-0.325263
26	7	-4.521247	-2.868508	2.472161
27	1	-5.538244	-2.844975	2.319139
28	6	-5.714801	-1.686982	0.196986
29	6	-7.865898	-1.825552	0.317399
30	6	-7.601997	-1.144520	-0.844392
31	1	-5.659327	-0.599647	-1.622741
32	1	-8.821083	-2.097805	0.731803
33	1	-8.241611	-0.728795	-1.601774
34	6	-4.055722	-3.485559	3.634691
35	6	-3.995089	-4.500268	5.521841

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
36	6	-2.680950	-4.348744	5.138454
37	1	-1.993835	-3.397820	3.291586
38	1	-4.378140	-4.962706	6.415646
39	1	-1.752268	-4.633078	5.600222
40	8	-2.461272	-2.274635	1.605070
41	7	-4.851759	-3.958350	4.577412
42	7	-2.733503	-3.689785	3.915810
43	7	-6.211710	-1.060622	-0.914412
44	7	-6.683094	-2.163268	0.964008
