N° d'ordre : 4477

THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

Par Tracey MARSHALL

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : CHIMIE ORGANIQUE

Chimie dynamique: reconnaissance de nucléobases par des récepteurs synthétiques et isomérie cis-trans d'hydrazones acylées

---

Directeur de thèse : Dr. Ivan Huc

Soutenue le 27 janvier 2012 après avis de :

Prof. Jerome LACOUR	Professor, University of Geneva	**Rapporteur**
Dr. Mihaï BARBOIU	Directeur de recherche CNRS, Université de Montpellier 2	**Rapporteur**

Devant la commission d'examen formée de :

Dr. Michel LAGUERRE	Directeur de recherche CNRS, Université de Bordeaux 1	**Président du jury**
Prof. Jerome LACOUR	Professor, University of Geneva	**Rapporteur**
Dr. Mihaï BARBOIU	Directeur de recherche CNRS, Université de Montpellier 2	**Rapporteur**
Dr. Ivan HUC	Directeur de recherche CNRS, Université de Bordeaux 1	**Directeur de thèse**
Dr. Frédéric GODDE	Maître de conférence, Université de Bordeaux 1	**Membre invité**

Université Bordeaux 1
Les Sciences et les Technologies au service de l'Homme et de l'environnement

To my parents,
Thomas and Thomasina
Marshall

For all of their support

Introductory remarks

The context of my PhD studies was within a Dynamic Combinatorial Chemistry (DCC) European Union Marie Curie Research Training Network [MRTN-CT-2006-035614]. This network brought together other European University groups with an active interest in developing the DCC concept to its full potential while also sharing knowledge and skills between its members. The aim was to train early stage researchers such as myself to develop chemistry and build on the theory behind these highly interesting reversible equilibrium systems. The role of the network enabled a transfer of information to provide efficiency in delivering new functional molecules and expanding the DCC toolbox. The network also provided a training platform bringing together different DCC disciplines such as biocatalysis and nanotechnology. My training was implemented at an institutional and network-wide level. My research was carried out at the Institut Européen de Chimie et Biologie (IECB), Université Bordeaux 1, Pessac, 33600, FRANCE under the guidance of Dr. Ivan Huc and Dr. Frédéric Godde.

Network Members *http://www-dcc.ch.cam.ac.uk/*

Sijbren Otto (coordinator), **University of Groningen (Netherlands)**

Saleh Hamieh, Lebanon

Jerome Peyralans, France

Jeremy K. M. Sanders, **University of Cambridge (UK)**

Jörg Klein, Germany

Aldrik Velders, **University of Twente (Netherlands)**

Ignacio Pérez-Victoria, Spain

Sabine Flitsch, **University of Manchester (UK)**

Roberto Castangia, Italy

Stefan Kubik, **Technische Universität Kaiserslautern (Germany)**

Eugenia Petkova Eugenieva, Bulgaria

Jerome Lacour, **University of Geneva (Switzerland)**

Franck Torricelli, France

Ulrich Lüning, **Christian Albrechts Universität Kiel (Germany)**

Vittorio Saggiomo, Italy

Olof Ramström, **Royal Institute of Technology Stockholm (Sweden)**

Morakot Sakulsombat, Thailand

Kay Severin, **EPFL (Lausanne, Switzerland)**

Vera Janowski, Germany

Abstract

The focus of this thesis was on novel molecular systems which can adapt upon the addition of substances that act as templates. This approach enables one major species to be identified from a mixture of compounds through the use of dynamic combinatorial chemistry (DCC). The first part of my PhD included the use of a single stranded DNA (ssDNA) as a template for information transfer with the self-assembly of receptors without the need for enzymes. The second part of my PhD focused on the self-sorting of acylhydrazone pyridine motifs and the interesting configurations they adopt.

First Project: The objective was to develop new compounds (A and G clamps) designed to recognise specifically purine on a DNA template (*Fig. 1*). Molecular recognition of the purine's adenine and guanine was envisaged. Both clamps were molecularly modelled with consideration to the number and nature of hydrogen bonding sites. The clamps also contained solubilizing groups, which while protected were hydrophobic and unprotected were hydrophilic. The binding studies were carried out first in organic solvent (with no competing hydrogen bonding effects) as proof of specific and efficient recognition. Once this was completed the next step was to assess the clamps in an aqueous environment with the presence of competing hydrogen bonding. This drawback was expected to be compensated by the possibility of induced cooperativity via *π-π* stacking interactions upon addition of the ssDNA template (the complementary strands were 3, 5, 8 and 13 mers).

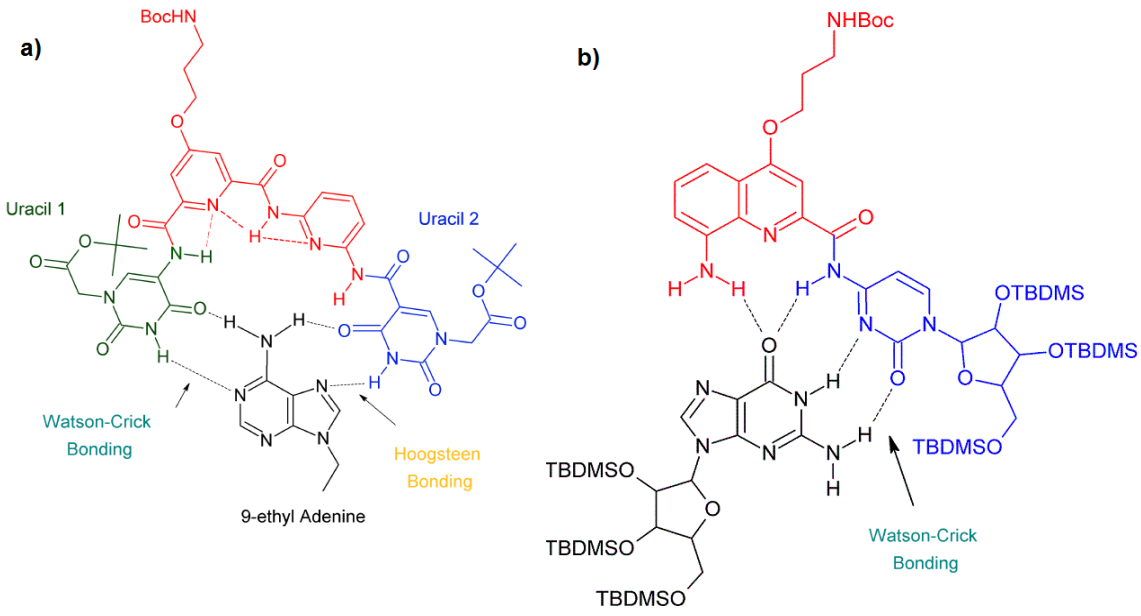


Figure 1: a) 9-ethyl adenine base (black) hydrogen bonded via Watson Crick and Hoogsteen Bonds to the A clamp containing two uracil groups (green, blue) for recognition joined by a linker (red); b) G Quinoline clamp hydrogen bonded to Guanosine (black) with protecting groups. Contains a cytidine group for recognition (blue).

Abstract

The first assessment in water involved Surface Plasmon Resonance (SPR); Circular Dichroism (CD) and UV melting techniques. CD was used to investigate the chromophoric interaction in detail. Optical activity was expected upon the self-assembly between the clamps and the ssDNA template resulting in a helical structure which might possess a preference for left or right-handed helicity. UV melting studies will determine the melting behaviour (T_m) upon heating between the clamps and a ssDNA duplex with the reorganisation of the coil back into a helix upon cooling. This study will give information about the strength of the cooperative process. Expectations are that there will be no organisation of stacking in organic solution. However the A clamp contains a large planar surface and is sp^2 hybridised. These features are expected to enable the possibility of *π-π* stacking interactions between the clamps and ssDNA template in water. This should induce a cooperative effect and result in the rapid zipping-up of the duplex. Future studies will include a poly A ssDNA template containing a guanine mismatch. It was expected that the clamps will organise in accordance to the sequence of the template.

It is envisaged that these receptors can be used in association with DNA as a template for information transfer with non-enzymatic replication using non-covalent interactions. Proof-reading is expected due to the reversible interactions to form one preferential stable product. Thus this method could also be applied to resolve a mixture of compounds for sequence selection. The initial investigations, synthesis and characterisations of these nucleobase clamps will enable us to progress towards this future goal.

Second Project: The next project was to explore the equilibrium of short dynamic systems derived of acylhydrazone pyridine motifs. The feasibility to synthesise these acylhydrazone pyridine motifs (dimer, trimers and pentamers) was investigated (*Fig. 2*). Behavioural observations monitored configurational changes of the cis/trans equilibrium of these dynamic systems via ^{1}H NMR studies (*Fig. 3*). It was expected that the equilibrium could be biased in such a way that is unusual. This was based on initial studies in which cis configuration was observed.

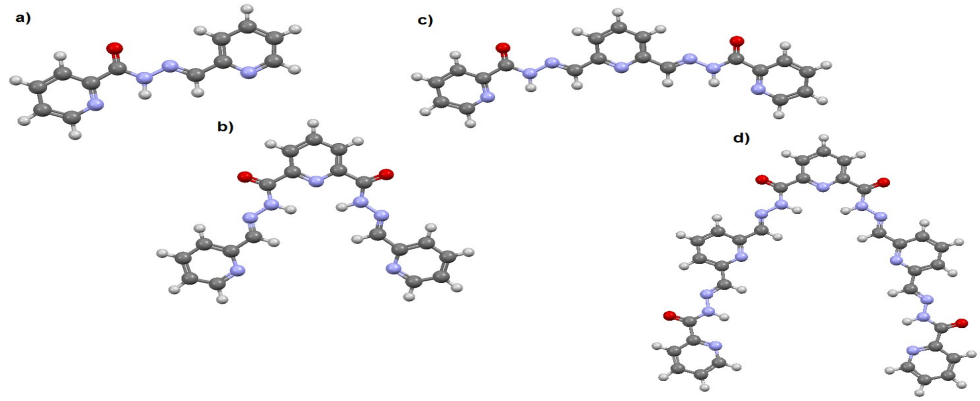


Figure 2: Molecular design of acylhydrazone pyridine motifs (side chains excluded); (a) dimer; (b) AHA trimer; (c) HAH trimer; (d) HAHAH pentamer.

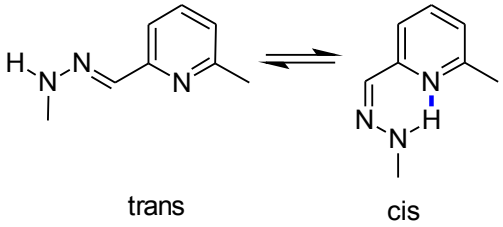


Figure 3: Intramolecular hydrogen bond (blue) of the acylhydrazone pyridine motif, between trans and cis

The intramolecular hydrogen bonding with pyridine-NH, induces a downfield chemical shift from 11.1 ppm (trans), to 15.9 ppm (cis) in $CDCl_3$ solvent as observed by ^1H NMR analysis. It was particularly interesting to investigate this system in depth as these motifs exhibit intramolecular hydrogen bonding and a configurational change. X-ray diffraction was also utilised to investigate cis configuration. It is important to note that with solution studies many species are in equilibrium. Whereas in solid state one isolated species was obtained via crystallisation from a mixture of compounds.

After experiments characterised the equilibrium between trans and cis configuration the next objective was to investigate how this phenomenon translates into longer oligomers. Oligomer length (such as trimers) (*Fig. 2.(b)(c)*) and environmental effects, (such as solvent and temperature) was investigated for its influence towards the configuration transformation and adaptation of the acylhydrazone pyridine motifs.

It is envisaged that this system can be useful and incorporated into a model which utilises folding as a driving force to enhance a bias in equilibrium and therefore isolate one species from a mixture of products (*Fig. 4*). The initial investigations, synthesis and characterisations of these acylhydrazone pyridine motifs will enable us to progress towards this future goal.

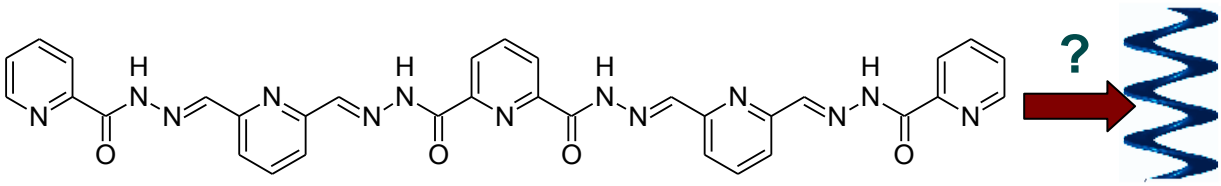


Figure 4: Can configurational change from trans to cis induce a folding of these systems? Can this driving force enable the isolation of one species from a mixture of products?

Acknowledgements

I would like to thank Dr. Ivan Huc for giving me this opportunity to study my PhD at the IECB, University of Bordeaux 1, France, within the EU Marie Curie RTN program. It has been a fantastic life changing experience. Ivan has been a great mentor throughout my studies and has been very supportive, helpful and inspiring. His patience while explaining scientific concepts helped me a lot to achieve my full potential. I am very grateful for the time he spent discussing my project and enabling me to have a greater insight and understanding into chemistry research. I would also like to acknowledge his contributions towards planning the various Marie Curie training and presentation network meetings around Europe. Overall, his help has contributed to a very rewarding and motivating experience. I would also like to thank Fred Godde for his support throughout my PhD.

I am also grateful to the EU Marie Curie Training Network for funding. There were many benefits and the funding enabled me to further my education, not only in chemistry but to also gain the experience of living and studying in another EU country with the opportunity to interact with different research groups across Europe. I also want to thank all the members of the network for the enriching discussions and for the enjoyable time I had at the network meetings. It was great to have the support from students in a similar situation.

I would also like to thank the rest of the Ivan Huc Group at IECB, for being helpful whenever I needed help. A PhD in chemistry is a great challenge and combined with moving to a new country and making new friends. However, meeting and working with everyone at IECB made my stay in France even easier and more enjoyable than expected. It was great to have people to connect with at IECB and to make the work day more fun. Whenever I needed support and time to relax there was always someone to enjoy a glass of Bordeaux wine/cheese plate with, to visit the French food market, to do the Irish pub quiz, to have a picnic along the quai, to do some sport, to watch the Ireland-France Rugby/Soccer or Irish Gaelic GAA matches, or to have a cooked English/Irish breakfast with. These moments helped to prevent me from feeling too homesick. Overall, I met some great friends at IECB during my time here. I would also like to thank the interesting friends I met via the Talence Cyclisme et Cyclotourisme Ufolep club, the Club Alpin Français FFCAM Bordeaux Escalade club, the Bordeaux surfing club, the “Blablabylone”, “After-work Anglais” and “Maison de l'Europe Bordeaux-Aquitaine” language exchange and culture clubs, the “On va sortir” club, Couchsurfing, Network Irlande and more recently the “Foreign Guests of the University of Bordeaux” club.

Finally, I would like to thank my parents for their support and for believing in me throughout my studies. I am very grateful to the rest of my family and friends in Ireland for staying in touch despite the long distances and for coming to visit me during my five years in Bordeaux.

Abbreviations

General

α	alpha
Å	angstrom(s)
°C	degree(s) celsius
δ	chemical shift
Δ	change, heating to reflux
Δδ	change in chemical shift
ΔG	spontaneous process, Gibbs free energy. $\Delta G = -RT (\ln K_{eq})$
Δ*H*	enthalpy
Δ*S*	entropy
μ	micro
AA	acceptor–acceptor
AD	acceptor–donor
calcd	calculated
CD	Circular dichroism
cm^{-1}	wavenumber(s)
^{13}C NMR	Carbon-13 nuclear magnetic resonance
d	doublet (spectral)
DCC/*CCD*	Dynamic combinatorial chemistry
DCL	Dynamic combinatorial libraries
dd	double doublet (spectral)
DD	donor-donor
DDA	donor-donor-acceptor
DNA/*ADN*	Deoxyribosenucleic acid
ESI	Electrospray ionization
eq, equiv	equivalent(s)
g	gram(s)
^{1}H NMR	Proton nuclear magnetic resonance
HMBC	Heteronuclear multiple bond correlation
HMQC	Heteronuclear multiple quantum coherence
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
Hz	Hertz
J	Coupling constant
k	Rate constant

K_a	Acid dissociation constant
K_{assoc}	Association constant. A + B◀⟶▶ AB, K_{assoc} = [AB]/[A]•[B]; K_{assoc} = k_{on}/k_{off} (M^{-1})
K_d	Dissociation constant. $K_d = 1/K_{assoc}$ (M, mM, µM, nM)
L	Litre(s)
LCMS	Liquid chromatography mass spectrometry
LRMS	Low resolution mass spectrometry
m	multiplet (spectral)
M	Molar (moles per litre)
M+	Parent molecular ion
MF	Molecular formula
mg	milligram
MHz	Megahertz
mins	minutes
mL	millilitre(s)
mmol	millimoles
mol	moles
MW	Molecular weight
m/z	mass-to-charge ratio
nm	nanometre
NMR/*RMN*	Nuclear magnetic resonance
NOESY	Nuclear overhauser enhancement spectroscopy
pK_a	the negative logarithmic of the acid dissociation constant
ppm	parts per million
q	quartet (spectral)
R_f	Retention factor (for TLC)
RT	Room temperature
RNA	Ribonucleic acid
ROESY	Rotating-frame overhauser spectroscopy
s	singlet (spectral)
SPR	Surface plasmon resonance
ssDNA	single stranded deoxyribosenucleic acid
SM	Starting material
t	triplet (spectral)
TLC	Thin layer chromatography
T_m	Melting temperature
UV	ultraviolet
v/v	volume per unit volume (volume-to-volume ratio)

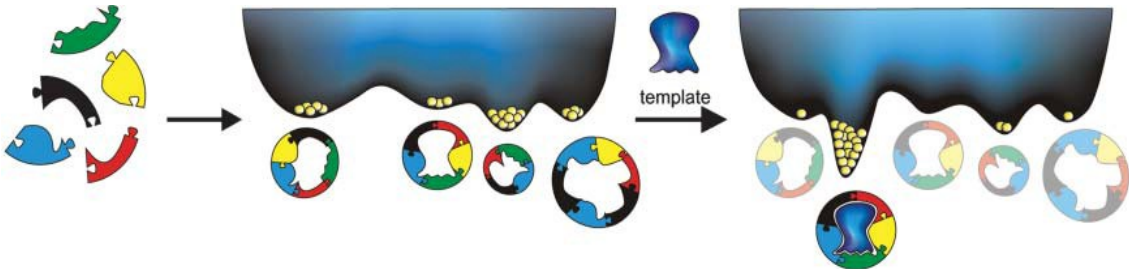
Reagents and solvents

A	Adenine
AcOH	Acetic acid
C	Cytosine
DCC	*N,N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DIAD	Diisopropylazadicarboxylate
DIEA	Diisopropylethylamine
DMA	Dimethylacetamide
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
EDCI	*N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride
EtOAc	Ethyl acetate
EtOH	Ethanol
Et_2O	Diethyl ether
G	Guanine
HBTU	*o*-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate
HOBt	1-Hydroxybenzotriazole
MeOH	Methanol
MeOD	Deuterated methanol
PBS	Phosphate buffered saline
Pd/C	Palladium on carbon
PPA	Poly(phosphoric) acid
PPh_3	Triphenylphosphine
PyBop	(Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophophate
Pet. Ether	Petroleum ether
SCBS	Sodium cacodylate buffer
T	Thymine
TBAF	Tetra-*n*-butylammonium fluoride
TBDMS Cl	*tert*-butyldimethylsilyl chloride
TEA/Et_3N	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
U	Uracil

Substituents and protecting groups

Ac	Acetyl
Bn	Benzyl
Boc	*tert*-Butyloxycarbonyl
Bu	Butyl
Et	Ethyl
*i*Bu	*iso*butoxy
Me	Methyl
t	*tert*
TBDMS	*tert*-butyldimethylsilyloxy

Table of Contents



Chapter 1: Dynamic Combinatorial Chemistry

1 Introduction

1.1 Dynamic Combinatorial Chemistry (DCC)

The term DCC, was first postulated in the early 1990's simultaneously by Lehn[1] and Sanders.[2,3] Lehn observed that the major product in a dynamic mixture of helicates was determined by the nature of the counter ion that binds in the centre of the helicate.[4] Huc and Lehn described the inhibition of carbonic anhydrase by a library of imines in situ.[5] Sanders concept of DCC was based on the mammalian immune system which exhibits combinatorial, selection, and amplification elements and decided to implement this general approach to become more efficient at developing synthetic receptors. DCC also has similarities to Emil Fischer's "Lock and Key" metaphor. In fact DCC is a 21st century technique utilising 19th century chemistry, with the help of modern analytical techniques, computational and statistical analysis.[6] A detailed account into the background and various applications of DCC is outlined in a review by Corbett et al.[7] and in two recently published books by B. Miller,[8] and J. Reek/S. Otto.[9]

DCC is defined as a library of building blocks that undergo reversible reactions via covalent or non-covalent linkages which give rise to equilibrating mixtures under thermodynamic control.[10] Addition of a template induces self-organisation which provides an in situ screening to a selection of the best binder. As the concentration of the stabilised species increases, a shift in equilibration towards that target is amplified (*Fig. 1*).[11] Not only can a chemical template stabilise a DCC system but also variation in environmental conditions can influence library distribution. By inducing a phase change, changing of temperature, pH or use of light and electric fields, it is possible to probe various species formed in the library.

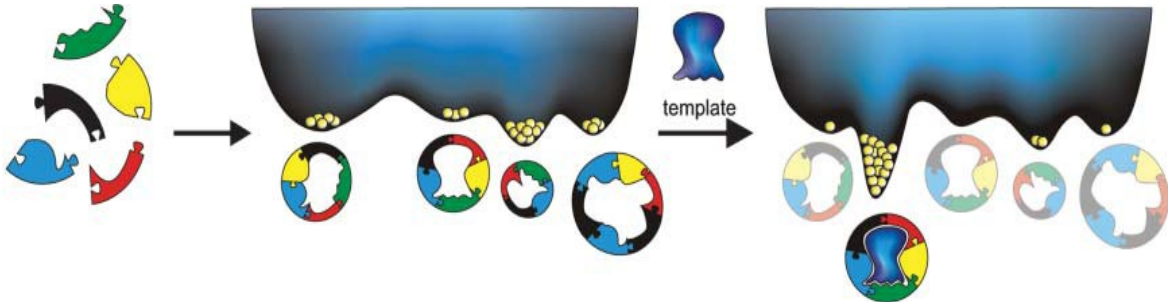


Figure 1: Bi-functionalised building blocks combine to form a dynamic combinatorial library of macrocycles that are potential receptors. The amounts of the individual macrocycles depend on their thermodynamic stabilities as represented by the depth of the corresponding well in the free energy landscape. Addition of a guest or template that selectively binds to one of the macrocycles induces a shift in the equilibrium towards this species at the expense of all other library members.[7]

The energetically most favourable situation of the entire system determines the outcome of an amplification experiment. In situations where the concentration of the receptor depends on the composition of the library, the best receptor is not likely to be amplified at the expense of a lower total concentration of receptor. Since the different receptors are built from the same building blocks, an increase in the concentration of certain receptors causes a decrease in the concentration of others.[12] Misleading amplification can be minimized by careful library design.[13] The desirable reaction properties for DCC, consist of a stoichiometric amount of each building block at low concentration (mM) in mild conditions compatible with templating. The system can be turned on, followed by proof reading and editing with connections between building blocks in flux, continuously being made and broken. Composition in the system is responsive with the addition of a template that selectively binds one member and induces a bias in the equilibrium with the amplification of a single thermodynamically stable product.

Traditional chemistry involving covalent bonds has allowed chemists to achieve overwhelming results, however the limits have been reached and the synthesis of even larger and more complicated molecules is likely to be impossible.[14] Traditional Combinatorial Chemistry involves one-to-one reactions [5] between many substrate's and a core library monomer joined by a robust bond, leading to expensive and time consuming methodology with little diversity and a necessity to analyse, purify and screen every single product against the desired target.[15] In contrast, DCC methodology enables *n*-to-one possibilities of adaptive correspondence between a set of components and the target site joined by dynamic covalent or non-covalent bonds, resulting in a hit-to-lead target molecule with diversity and minimal effort.[5] Dynamic Combinatorial Libraries (DCLs), seek to provide optimum targets mimicking nature's approach with new molecules waiting to be explored and discovered (*Fig. 2*).[15] In nature, templating is key in the replication of DNA, first observed by Watson and Crick in 1953 which inspired chemists to postulate that one day, organic templates might be utilised in organic synthesis.[16]

Figure 2: Chemist's (left) vs Nature's (right) Approach to discovering functional molecules.[15]

The differences between supramolecular synthesis and traditional covalent synthesis is compared *(Fig.3)*. The main advantages of DCC is the ability for the system to "proof-read" and the access to large structures. DCC noncovalent and covalent synthesis exhibit the limitation of disfavoured structures which is compensated with the addition of a template and thus offers substantial capabilities to that of traditional chemistry. This approach offers a powerful methodology for the discovery of substrates, inhibitors, receptors, catalysts and carriers for a variety of processes.[17]

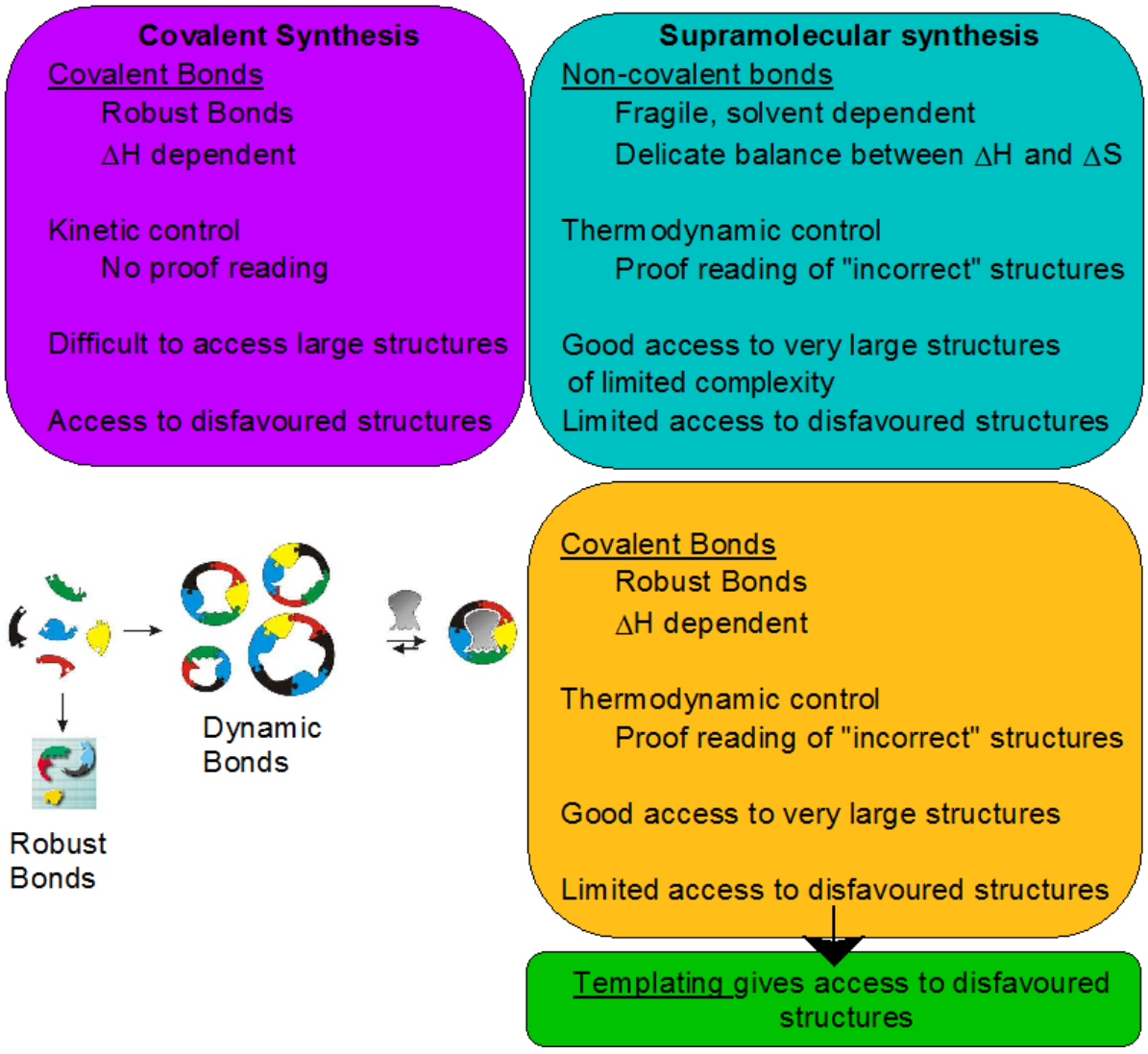


Figure 3: Differences between Covalent Synthesis and Supramolecular (and DCC) Synthesis.[6]

The advantage of utilising DCC is the ability for a "best-binder" to be obtained from a mixture of library building blocks with the addition of a template.[9] One can use an external template such as molding or casting techniques or internal, "self-template" in the case of self folding (intramolecular) or self aggregation (intermolecular) to influence species of library distribution *(Fig. 4)*. Different templates amplify different receptors from the same library. Immobilised templates can simultaneously amplify, isolate and purify receptors.

Hence, it is possible to turn off equilibrium to fix, isolate and identify the successful receptor which is otherwise inaccessible, thus the ability to create solutions that evolve in response to external stimuli. This type of synthesis under thermodynamic control, is extremely efficient and simulations can guide experimental design. The building blocks can be decided upon depending on which templating approach will be implemented.

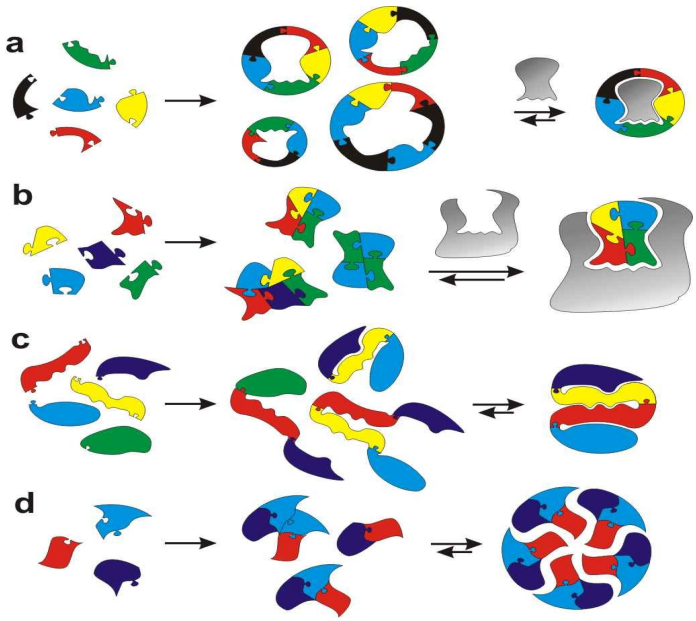


Figure 4: Various Approaches to DCC and templating; a) use a ligand to make an ideal receptor – molding; b) use a receptor to make an ideal ligand – casting; c) optimise a self folding structure – driven by internal non-covalent interactions; d) generate self assembled aggregates – driven by external non-covalent interactions.[7]

To incorporate DCC towards designing potential libraries of structural diversity and unique functionality, one must be knowledgeable with the various interactions that can be utilised in templating.[18] These include:

1. Hydrogen bonding
2. Aromatic π-π interactions
3. Metal ligand coordination
4. Cation- π interactions
5. Aromatic donor-acceptor interactions
6. Hydrophobic effects.

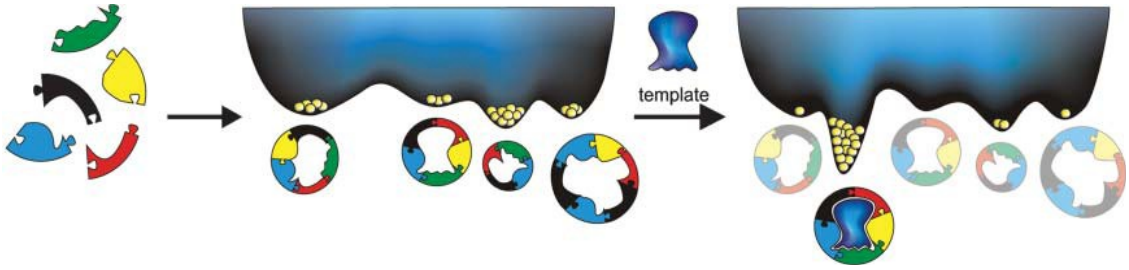
A reversible reaction pathway such as those listed can be utilised in DCC:

- Base-catalysed transesterification
- Transacetalisation
- Michael addition of thiols
- Biocatalysis
- Nitro-aldol (Henry) reaction
- Alkene metathesis
- Reversible Diels–Alder

- Folding
- Thioester exchange
- C=N bond exchange (imines, oximes, hydrazones…)
- Disulfide exchange
- Photochemical isomerisation

It is significant, that equilibration is carried out in conditions which are compatible with the desired target. The rate of exchange needs to be fast enough for equilibration distribution to occur within a reasonable time frame, but slow enough for binding to the target to occur. A fundamental question is, what is the correlation between amplified molecules and host-guest binding strength? Is the best species amplified, similar to supramolecular "Darwinism"?[5] At low template concentrations, most binding energy is released by amplifying the best receptor. This can also be interpreted that with more building blocks, and therefore larger libraries, the higher probability of obtaining the strongest binder. At high template concentrations more binding energy may be released by creating a larger number of weaker receptors. With regards to selectivity, the library constituents, must be stereorandom and nonsubstrate selective. It is critical that the library members react exclusively on the desired functional groups attached on the building block units and not the target functionality which would result in irreversible products and disrupt the equilibrium. As mentioned previously, not only can the DCL be directed by a "chemical template", it can also constitute selection by environmental changes and directed by external physical stimuli. Various ways to "freeze" the equilibration include removal of a catalyst or reactive functionality, turning off the light source, changes in temperature or pH. A method for halting this equilibration requires one which doesn't interfere or decompose the amplified product before analysis.[8]

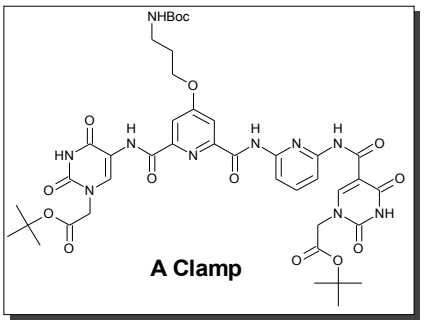
After halting the exchange, the mixture is then usually analysed via LCMS, in which a change in the distribution curves before and after templating in observed. Albeit small DCL which can be analysed by NMR. This is a single one step screening and analytical approach to DCC. The time taken to thoughtfully think through the process of DCC design and possible difficulties, rewards itself by being an efficient and effective method to develop new diverse complex receptors previously inaccessible. With this in mind, it was postulated to apply this technique towards the first project involving nucleobase recognition and the second project consisting of acylhydrazone pyridine isomerism. Libraries consisting of various ssDNA templates consisting of purines, adenine and guanine with the addition of the receptors A and G Clamps was envisaged. Whereas the self-folding of acylhydrazone pyridine oligomers was of interest for the second project. These two projects utilising DCC as a tool, in Chapter 2 and Chapter 3 will investigate whether these goals were attainable.

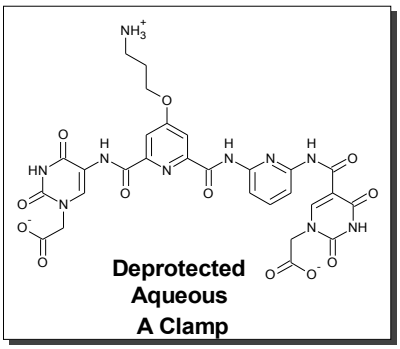


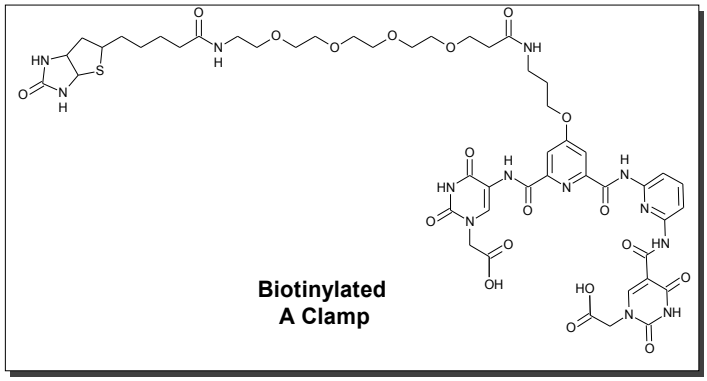
Chapter 2: Nucleobase Recognition and Non Enzymatic DNA Replication

Nomenclature

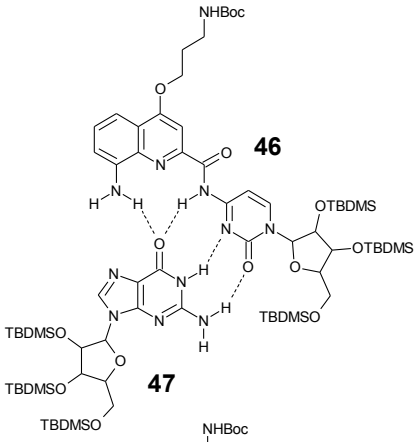
A clamp series

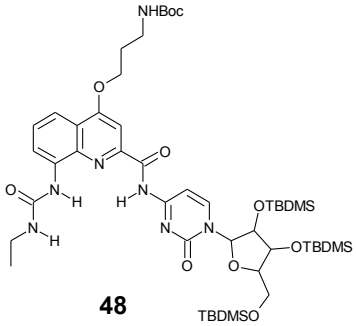


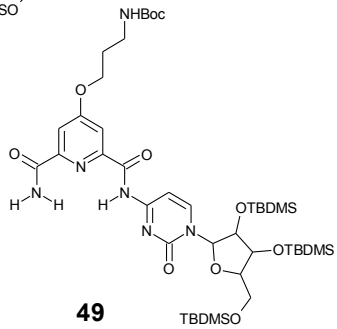


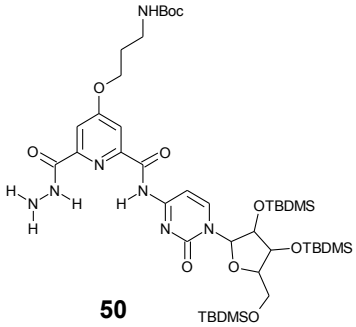


G clamp series



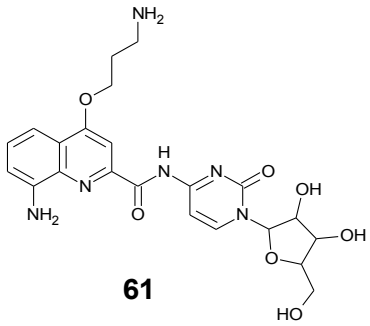






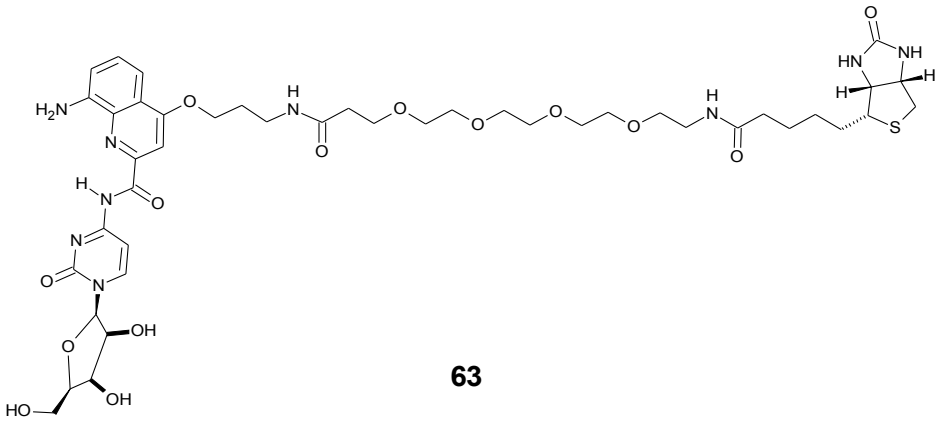
Protected G clamps (hydrophobic) : 46, 48 = quinoline G clamps; **49, 50** = pyridine G clamps and TBDMS protected guanosine **47.** Note: The above numbering is based upon the numerical sequence within this chapter. A clamp based molecules are **1** – **45** whereas G clamp based molecules are **46** – **70**.

G clamp series



61

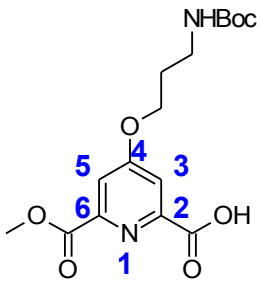
Deprotected quinoline G clamp (hydrophilic, aqueous soluble) : **61**

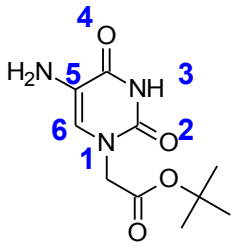


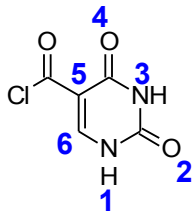
63

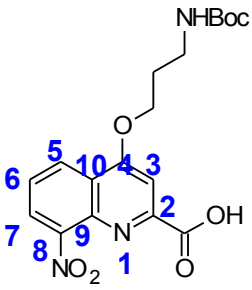
Biotinylated G Clamp: **63**

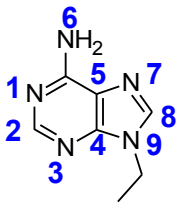
Numbering

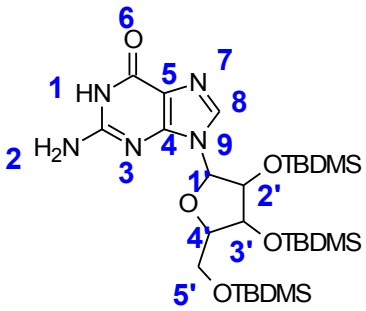


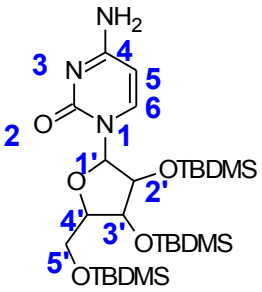












1 Introduction

1.1 Short Introduction to DNA

Nucleic acids are fundamental building blocks in life and essential to biological functions, from DNA information storage to gene expression (RNA transcription and translation). This makes for a desirable target for the development of nucleic acid ligands.[8] DNA is a long polymer of simple units called nucleotides *(Fig. 1)*, which are held together by a backbone made of sugars and phosphate groups.

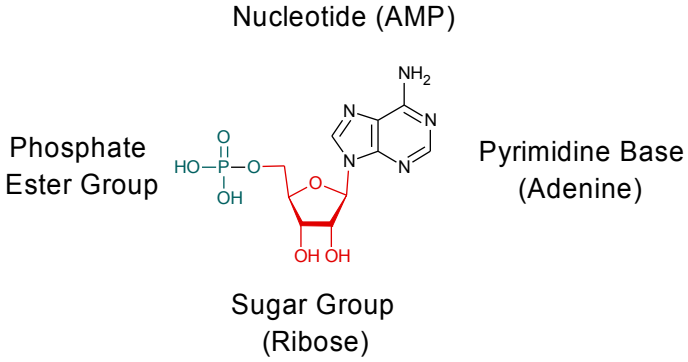


Figure 1: Nucleotide, AMP – adenosine monophosphate.

In nucleic acids there are five natural bases, two sugars and one phosphate group possible. There are two purine bases, adenine (A) and guanine (G) and three pyrimidine bases: uracil (U), thymine (T), and cytosine (C) *(Fig. 2)*. Adenine, guanine and cytosine are found in DNA and RNA, uracil in RNA only, and thymine in DNA only. Each base pairs up specifically with its complementary base – adenine with thymine (A-T); guanine with cytosine (G-C) and interacts via two or three hydrogen bonds, respectively *(Fig. 2)*.[19] While the majority of these motifs are by Watson-Crick bonding, it is not exclusive and other possibilities exist, including reverse Watson–Crick, Hoogsteen, 'wobble' (or mismatched) base pairs and base triplets. The purines, A and G, contain two major hydrogen bonding sites, the Watson–Crick and the Hoogsteen sites. Thus, in addition to being able to form 1:1 complexes through either of these two sites, both of these nucleobases can form 2:1 aggregates (or base-triplets) with an appropriate partner (two homo pyrimidines and one homo purine). The NH groups of the bases are good hydrogen bond donors **D**, while the sp^2-hybridised electron pairs on the oxygen (electronegative atom) of the base C=O groups are much better hydrogen bond acceptors **A** than the oxygen of either phosphate or pentose groups.[20]

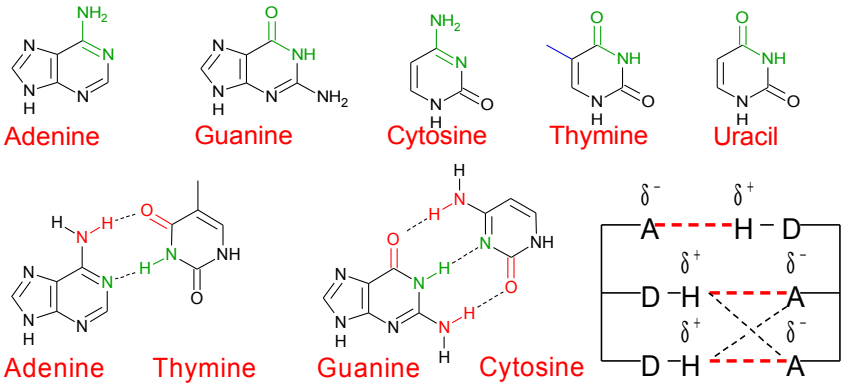


Figure 2: Five "natural" bases in nucleic acids, Base Pairing T-A and C-G (primary hydrogen bond interactions).

The major function of DNA is to encode the sequence of amino acid residues in proteins, using the genetic code. The fidelity of transmission of the genetic code rests on the specific pairings of AT and CG base pairs. If changes in the shape result from base mismatches, such as AG, they must be recognised and repaired by enzymes with high efficiency. In a double helix, the two strands are also held together by forces generated by the hydrophobic effect and π-π stacking, adding to the overall stability of the molecular structure.[21] The use of noncovalent bonds is essential to replication because they allow the strands to be separated and used to template new double stranded DNA. DNA is copied into RNA by RNA polymerase enzymes.[22]

1.2 Specific Nucleobase Recognition

Nature has excelled at information storage, transcription and replication with a limited number of building blocks. For this reason, research into understanding how these molecules interact is of great interest. These complex architectures are achieved via non-covalent interactions, the nature of the molecular components, and the type of interactions which they are involved in a subtle, yet tunable and responsive way for a desired and specific purpose. These range from hydrogen bonding, *π-π* interactions, electrostatic bonds, hydrophobic effects, and phase segregation to metal–ligand chemistry. Nucleobase recognition is possible not only by utilising hydrogen bonding (Watson-Crick and Hoogsteen bonding), but is also complemented by additional binding forces.[23] Previous investigations focused on complementary hydrogen bonding in relatively non-polar organic solvents. Kool has published an interesting review on the design and methods for nucleic acid recognition.[24] These investigations have now been extended to complexation in polar solvents or water by utilising:

1 Co-ordination
2 Ion pairing
3 π- π stacking interactions
4 Multiple hydrogen bonds.

1.2.1 Adenine

Since there are two hydrogen bonds between adenine and thymine (K_{AT} *ca.* 100 M^{-1} in $CDCl_3$), the degree of interaction is significantly lower than that of guanine – cytosine (K_{CG} *ca.* 10^4–10^5 M^{-1} in $CDCl_3$), which has a triple hydrogen bond complex. Hamilton, [25] Rebek, [26] Zimmerman, [27] Gokel [28] and Wilcox [29] are the pioneers in molecular receptors for adenine/thymine recognition *(Fig. 3)*. By utilising multiple nucleobase binding sites, significant binding constants could be obtained, via hydrogen bonding sites in conjunction with π-π stacking interactions. As with the case of larger supramolecular assemblies, substituents attached to the nucleobase can play a role in controlling the nature of the assembly.

The Hamilton group reported in the late 1980s a ***multi-site*** approach for the recognition of nucleotide bases in which hydrogen bonding and aromatic stacking groups within a macrocyclic receptor bind simultaneously to the substrate. By varying the hydrogen bonding, selective receptors for nucleobase recognition could be achieved *(Fig. 3)*. Changing the electronic characteristics of the stacking group results in different geometries for the aromatic-aromatic interaction.[25,30]

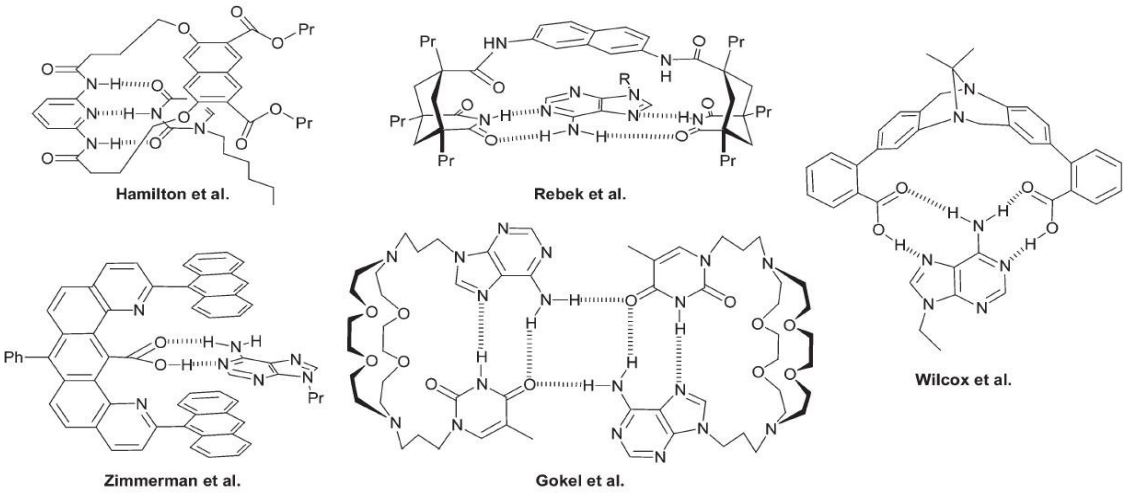


Figure 3: Examples of Adenine and thymine receptors.[23]

More recently, Odashima and co-workers have utilised complementary hydrogen bonding towards synthetic hosts for adenine *(Fig. 4,A)*. The planar receptor containing five-six-five-membered heteroaromatic rings and two carbamoyl NH sites exploits both Watson–Crick and Hoogsteen-type interactions. Furthermore, they have modified the receptor to incorporate π-π stacking by attachment of a 2,7-dialkoxynaphthalene unit tethered by the appropriate length of alkyl side chains *(Fig. 4,B)*.[31] As shown in their initial publication,[32] the receptor was successful in selectively recognising a lipophilised adenosine derivative in $CHCl_3$, in a highly selective manner (K_{assoc} = 12000 M^{-1}). The group is currently continuing these studies to adapt receptors for adenine nucleotides on a membrane/water interface.

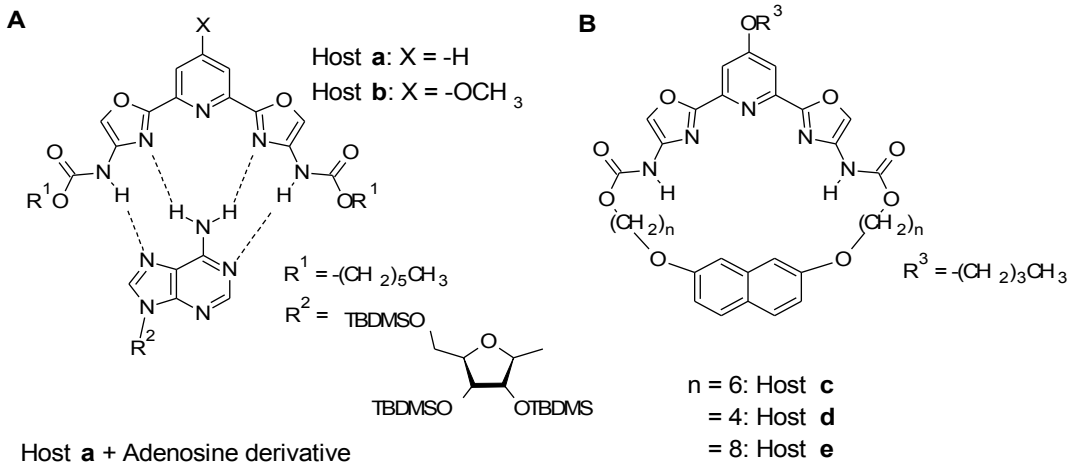


Figure 4: a) Complexation between host **a** and adenosine by multiple hydrogen bonding. b) Molecular structure of host **c-e** [31]

The level of development of receptors with high selectivity for adenine (and in particular, those which work in aqueous media) still remains insufficient. Based on these findings, the design of our adenine receptors was formulated to integrate multiple hydrogen bonding sites to exploit Watson-Crick and Hoogsteen type hydrogen bonds and an aromatic planar surface was expected to enhance *π-π* stacking interactions. An interchangeable side chain was incorporated to enable host-guest binding studies in organic and aqueous solvents.

1.2.2 Guanine

The bonding between guanine and cytosine is an attractive motif to base receptors upon due to its strong degree of interaction. The triple bond complex is of importance and various calculations were obtained for interaction energies as reported by Jorgensen.[33,34] Guanine receptors have been limited due to the reduced solubility of guanine derivatives in synthetic systems. In 1995, Sessler utilised Watson-Crick nucleobase-pairing interactions between guanosine and cytosine bearing porphyrins to form rigid hydrogen-bonded ensembles.[35] These were used to examine the effect of non-covalent interactions between donor and acceptor units on a photoinduced energy transfer process. They further developed this concept towards a novel guanosine-cytidine dinucleoside which aggregates via Watson-Crick hydrogen bonds to form a trimeric cyclic supramolecule. Such structures are potentially useful in the construction of self-assembled dendrimers and other nano-structures.[36] *(Fig. 5)*

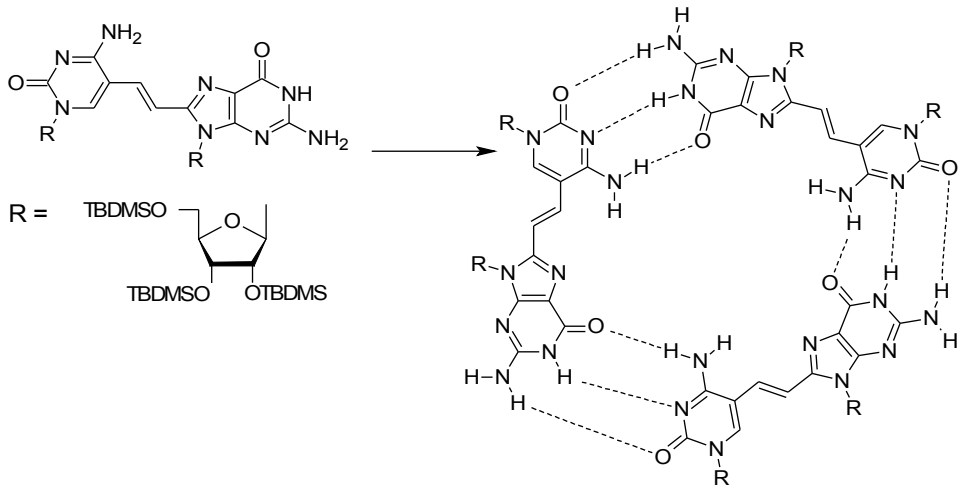


Figure 5: A guanosine–cytidine dinucleoside that self-assembles into a cyclic trimer.[36]

Zimmerman published in 2000, an example of a non-natural hydrogen bonding system for biological interest.[37] The major groove side of a cytosine-guanine (CG) base pair presents a non-contiguous donor-acceptor-acceptor (DAA) hydrogen bonding array. In addition to strong binding, the three hydrogen bonds formed between such a receptor and the CG base pair should provide selective binding because the thymine-adenine, guanine-cytosine, and adenine-thymine base pairs present different arrays of hydrogen-bond donor and acceptor sites *(Fig. 6)*. The strong complexation of the CG base pair by Zimmerman receptors bodes well for the development of non-natural nucleosides capable of recognising CG base pairs in the pyrimidine-purine-pyrimidine DNA triple helix motif.

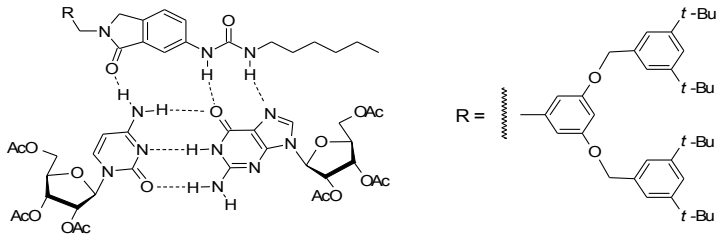


Figure 6: Hydrogen-bond-mediated complexation of a CG base pair by a hexylureido phthalimide and a hexylureido isoindolin-1-one.[37]

Zimmerman developed this idea further, resulting in an artificial receptor which simultaneously binds both the Watson-Crick and Hoogsteen sites of guanosine.[38] By forming multiple hydrogen-bonding contacts, receptors are able to complex nucleobases in competitive media, including some in water. Their research concentrated on naphthalene and quinoline derivatives. These motifs were based on the analogues by Umezawa et al. who reported a cytosine analogue with an 8-ureido-2-naphthoyl group attached at *N*-4 and complexes to a guanosine derivative **G** via five hydrogen bonds.[39] In addition to the Watson–Crick type hydrogen bonding site on the *N*-1 side of guanosine (two hydrogen bond donor atoms and an acceptor in a DDA array), natural guanosine derivatives have an acceptor–donor (AD) and an acceptor–acceptor (AA) hydrogen bonding site on the side of *N*-3 and *N*-7 which could also be utilised. Umezawa therefore linked a urea or thiourea group as a donor-donor (DD) hydrogen bonding site to a rigid naphthalene spacer which had a cytosine derivative attached *(Fig. 7)*. During ^{1}H NMR titration experiments in $CDCl_3$/DMSO-d_6 (4:1 v/v), complexation-induced downfield shifts for all NH hydrogens of both the receptor *(Fig. 7* ***a****)* (1.0 mM) and guanosine *(Fig. 7* ***G****)* (0-25 mM) were found, confirming the formation of hydrogen bonds. The association constant (K_{assoc}) of the urea containing receptor was found to be 170 M^{-1}.[39] Zimmerman's group modified the Umezawa analogues to be more stable to hydrolysis and investigated the importance of various design features in specific nucleoside recognition.*(Fig.8)*.[38] They wanted to exploit these receptors in duplex DNA. As a result, it is not sufficient to rely on hydrogen bonding alone, however multiple bonds do enhance selectivity. There was also considerable interest in the methylene linker because of its expected stability during oligonucleotide synthesis. Overall, the ^{1}H NMR chemical shift changes observed in both guanosine *(Fig. 8* ***G'****)* and Zimmerman's receptors *(Fig. 8* ***a-f****)* indicate Watson-Crick hydrogen bonding and minimally, one or two additional hydrogen bonds to the Hoogsteen edge of **G'** from the pendant amino or ureido group, respectively. The association constant (K_{assoc}) of the receptor *(Fig. 8* ***b****)* in $CDCl_3$/5% DMSO-d_6 had a value of 2530 M^{-1} for the naphthalene derivative whereas the quinoline derivative *(Fig. 8* ***a****)* was just 230 M^{-1}. This difference in binding constants suggests an unfavourable electrostatic interaction between the quinoline nitrogen lone pair and the guanosine carbonyl group. The addition of a tethered ureido group *(Fig. 8* ***c****)* maximised the Hoogsteen bonding to increase binding with an association constant (K_{assoc}) of 4070 M^{-1}.

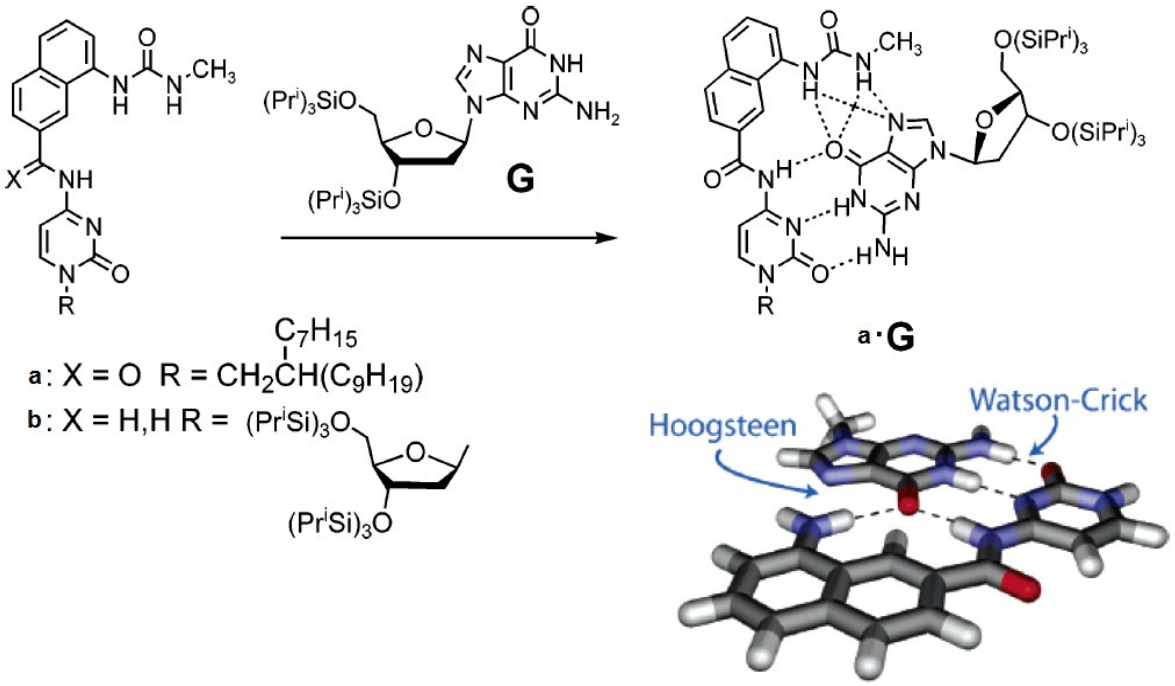


Figure 7: Umezawa guanosine (**G**) receptors (**a, b**).

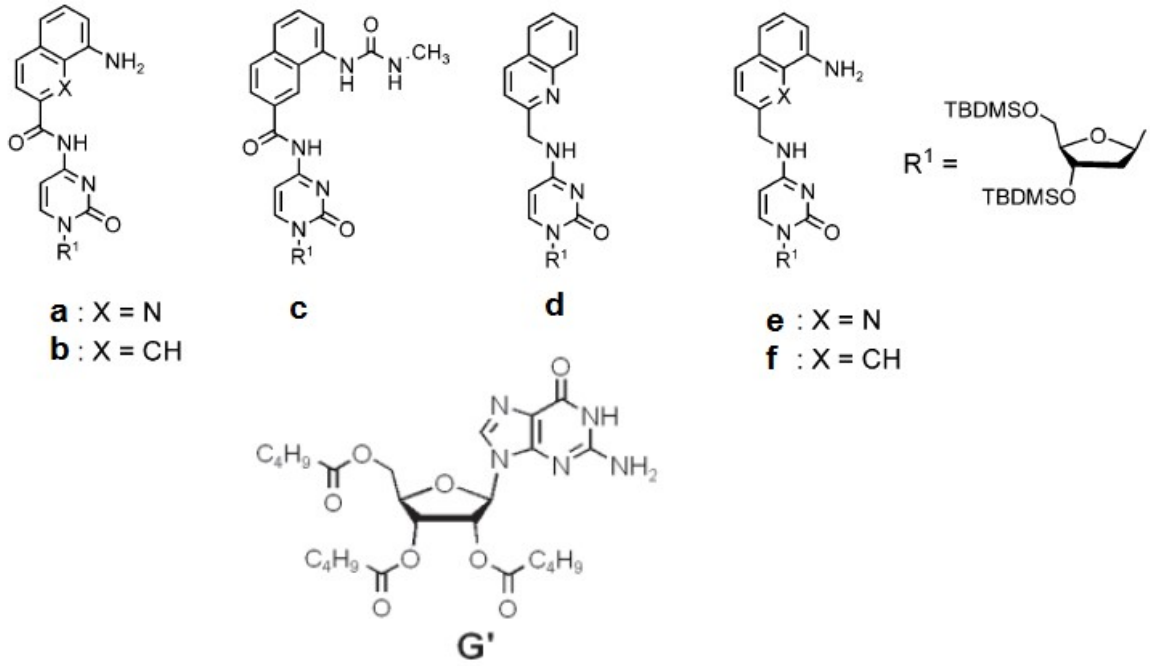


Figure 8: Zimmerman's guanosine (**G'**) receptors (**a-f**).

Some design guidelines set out by Zimmerman to develop novel nucleosides enhancing recognition of guanosine in duplex DNA include:

- Multiple hydrogen bonding for selectivity and base recognition
- Aromatic stacking for stabilisation
- Tethered receptor to maximise interaction to ligand
- Amide linker group to enable the N-H, to hydrogen bond with the "rabbit ear" lone pairs for the carbonyl group

Our guanosine receptors were based on these findings and were used as a comparison.

1.3 DNA Induced Assemblies

1.3.1 Targeting a nucleobase in water

Hydrogen bonding motifs have been widely demonstrated in organic media, however due to competitive interference from hydration in aqueous media, the number of receptors developed has not been as widespread. Some previous research in aqueous media for adenine receptors has been conducted by the Lehn,[40] Rebek,[41] Hong,[42] and Anslyn[43] research groups and for guanosine receptors by Yoon et al.[44]

Schmuck has recently published work regarding a water soluble tweezer receptor *(Fig. 9)* which contains two symmetric peptidic arms, connected by an aromatic scaffold and contains lysine (which allows additional charge–charge interactions with the substrate), phenylalanine (*π*-*π* interactions with the nucleobase), and a guanidinium-based oxoanion-binding site as the head group.[45] Their research showed the receptor to be capable of binding to phosphates and nucleotides with binding constants between 10^4 and 10^5 M^{-1} in buffered water at pH 7 via 1:1 complexes. Strong interactions between the highest charged primary phosphate group with the cationic groups of the receptor plus the hydrophobic interactions between the nucleobase and aromatic template lock the complex conformation. The efficient binding is based on noncovalent electrostatic and hydrophobic and/or *π*-*π*-stacking interactions with the nucleobase within the cavity-shaped receptor. Mono-nucleotides were bound much stronger than phosphates. The overall sequence of binding constants was UMP>GMP≈AMP>cAMP≈CMP>ADP>ATP≈PPi≈P. Hence this is the first artificial receptor that prefers the less-charged monophosphate nucleotide (AMP, K_{assoc} = 76000 M^{-1}) over the higher-charged diphosphate (ADP, K_{assoc} = 32500 M^{-1}) and triphosphate nucleotides (ATP, K_{assoc} = 8300 M^{-1}) in water. Thus the formation of a well-defined complex with specific interactions in contrast to more unspecific electrostatic interactions explains the unusual binding selectivity for this tweezer receptor. Their findings underline the importance of combining different types of interactions to achieve stable and selective binding of biological substrates in aqueous solutions.

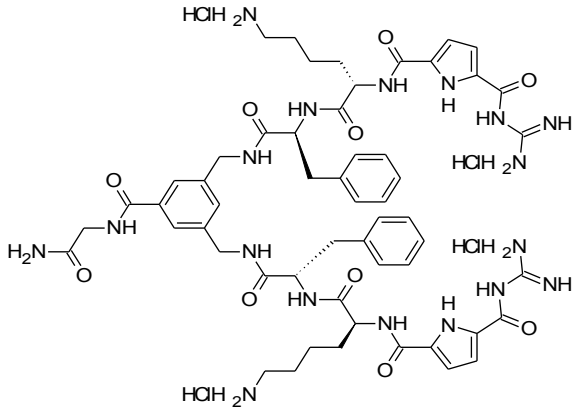


Figure 9: Guanidinium-based artificial tweezer receptor.[45]

Teramae and co-workers reported a method for nucleobase recognition in water which involved the construction of an abasic site (AP site, apurinic/apyrimidinic) in DNA duplexes. As in *(Fig. 10)*, normal DNA is hybridised with an AP site-containing DNA strand, with the AP site opposite from, but facing toward, a target nucleobase. This provides an opportunity for a ligand to recognise its corresponding target nucleobase through hydrogen bonding.[46] DNA damage is commonly due to naturally occurring AP sites. The Teramae group incorporated a tetrahydrofuranyl residue (**X** = dSpacer) *(Fig. 11)* which lacks a nucleobase moiety into a 11-mer oligodeoxynucleotide (5'-TCCAG**X**GCAAC-3', X = dSpacer) for the design of the AP site. The complementary DNA strand contained **Y** = A (adenine), C (cytosine), G (guanine), and T (thymine). The ligands were naphthyridine (**ND**) and quinoline (**AQ**) and 2-amino-7-methylnaphthyridine (**AMND**) derivatives *(Fig. 11)* which could utilise two or three hydrogen bonds for nucleobase recognition. **ND** has two aza-nitrogen atoms which work as acceptors for hydrogen bonds, whereas **AQ** has an acceptor and a donor sites for hydrogen bonding. As for **AMND**, it has three hydrogen bonding sites, that is, two acceptor and one donor sites.

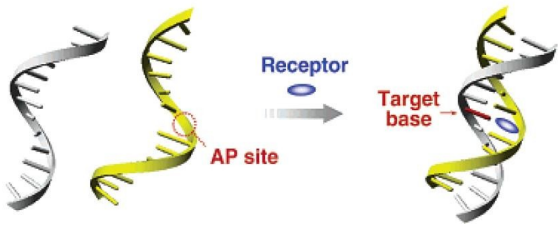


Figure 10: Nucleobase recognition in water using an abasic site (AP site)-containing DNA strand.[46]

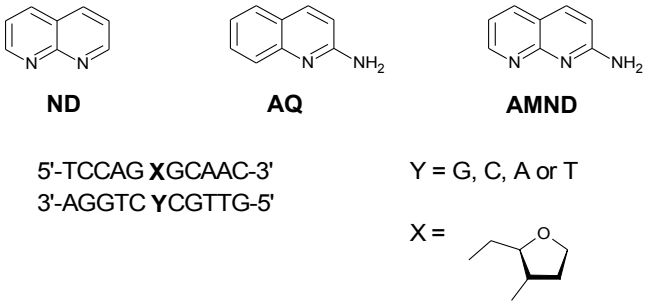


Figure 11: AP site duplexes, ligands **ND, AQ, AMND**

They used melting temperature measurements to determine the best ligand and **AMND** was found to strongly recognise target nucleobases (C > T > G > A) at the AP site via the formation of hydrogen bonds. **AMND** shows the selectivity for **C** (ΔT_m = 13.7) over G (ΔT_m = 4.9) at the AP site. The K_{assoc} between **AMND** and **C** was calculated to be greater than 10^6 M^{-1} in water. A stacking interaction with nucleobases flanking the AP site significantly controls the binding properties of the ligand (**AMND**). Therefore, **AMND** should bind to **C** in a cooperative fashion via hydrogen bonding with **C** and stacking with nucleobases flanking the AP site contributing to binding stability and selectivity. Circular dichroism (CD) and fluorescence experiments were also investigated. Their method can help develop a ligand-based detection method for single-nucleotide polymorphisms (SNPs) in aqueous media.

1.3.2 Self-assembly via complementary oligonucleotide templates

Biomolecules can be used as a template to control the size and sequence of non-natural polymers in a bottom up approach. In biological systems, enzymes help bring reagents in close proximity and thus accelerate intermolecular reactions which would otherwise not occur under physiological conditions. Without enzymes or proofreading, substrates need an alternative mode in replication and self-assembly. Bioorganic chemists are interested in progressing non-enzymatic synthetic approaches towards artificial biosystems. Developing these new biosystems can make way for numerous applications, but they might also help uncover the mystery of the origins of life. The first article reporting “non-enzymatic”, template directed synthesis of oligonucleotides was by Naylor and Gilham in 1966.[47]

Orgel pioneered the reasoning that DNA was too complex to have been the first repository of genetic information, Orgel and others speculated that RNA could have preceded it, simplifying the evolutionary process.[48] Since even RNA is very complicated, he synthesised simpler molecules that could have arisen even before it. Orgel suggested that RNA replaced a simpler precursor, a peptide nucleic acid, which could be replicated in a test tube. Orgel "looked at it like a chemist: how does it work and why does it work? How does a machinery like this self-synthesise?" Orgel showed that RNA and DNA could be used as templates to mimic their replication non-enzymatically and to create non-natural polymers with a predefined sequence.[49] A discussion on current research which implements templates towards self-assembly with a complementary oligonucleotide will be described herein.
Meijer and co-workers have utilised molecular recognition and non-covalent interactions to develop nano-sized functional objects using DNA as a template.[50] They synthesised well defined structures of π-conjugated oligo(*p*-phenylenevinylene) (**OPV**) nucleotide-appended (thymidylic acid) to form homo and binary self-assemblies with a complementary single-stranded 20-meric oligodeoxy-adenylic acid in aqueous conditions *(Fig. 12)*.

The templating process strongly depended on the concentration of complementary oligoadenylic acid 20-mer while the stability of the binary self-assembly is temperature dependent. Complementary T-A base pairs formed and induced a helical stack of the **OPV**s in the binary self-assembly. CD experiments revealed right-handed stacks self-assembled depending on the stoichiometry of A:T. Transformation from a helical to non-helical molecular assembly was found to occur non-cooperatively. Therefore the helical formation process of binary self-assembly, differs from DNA molecules where formation of double-stranded helical structures takes place cooperatively.

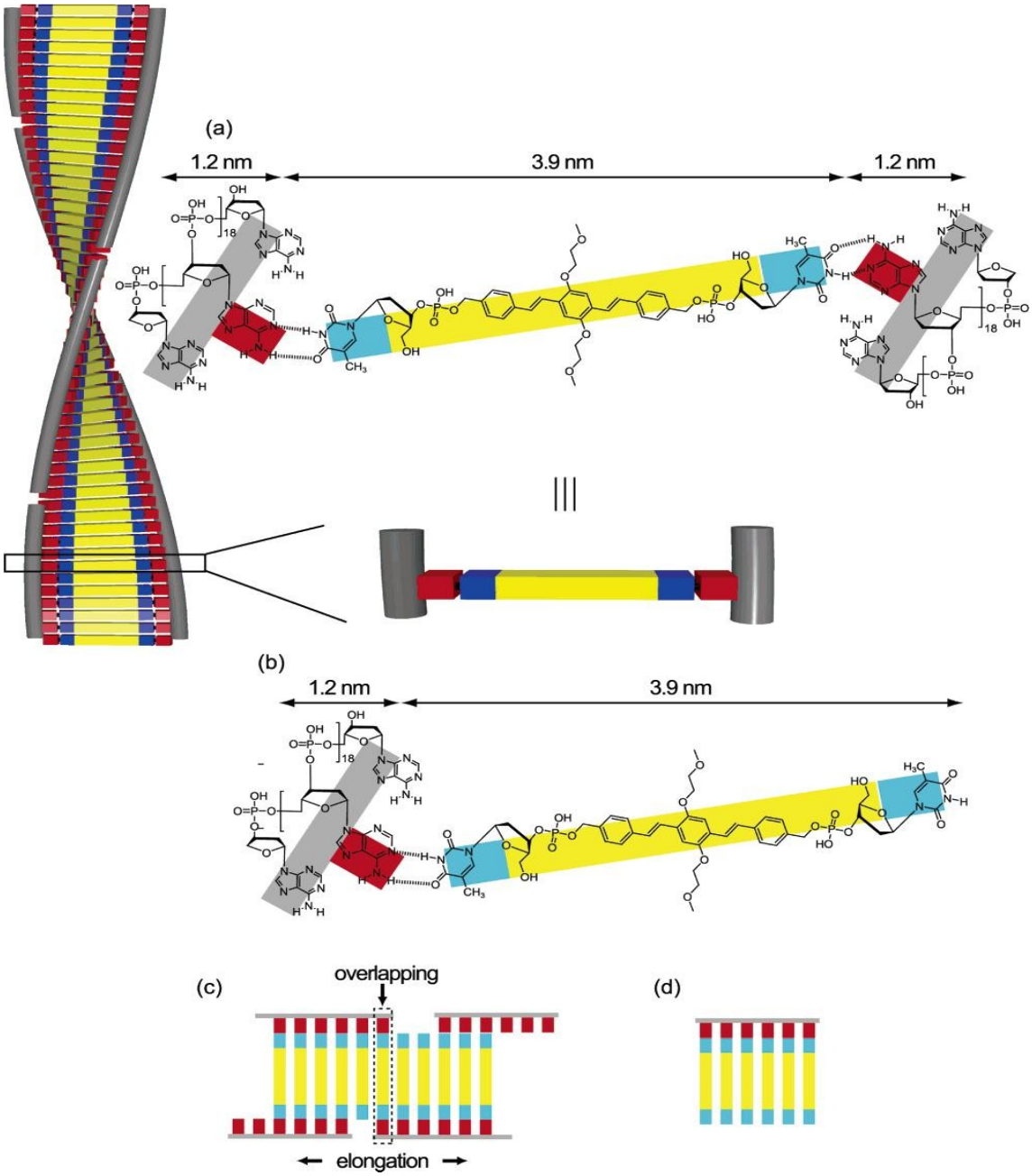


Figure 12: Proposed structure for the binary self-assembly of (a) (T:A) 1:1 and (b) (T:A) 2:1. Possible elongation mechanism for the helical stacks self-assembled from (c) (T:A) 1:1 and (d) (T:A) 2:1. (red/grey = complementary 20-meric adenylic acid, blue = thymine, yellow = OPV).[50]

Meijer and Janssen further developed their concept by binding an oligothymine (**dT*n,*** *n* = number of thymines) single stranded DNA (ssDNA) template via hydrogen bonding to diamino triazine chromophore units, stabilised by π-π interactions *(Fig. 13)*.[51] The number of chromophores was found to be controlled by the template. The first generation of diamino triazine units was the naphthalene derivative **NT**. Thus for a **dT40** ssDNA template, 40 **NT** could be bound in a fully reversible manner. ESI-MS could not detect **NT-dT40** complexes due to difficulties in deconvolution whereas **NT-dT10** could be detected. When **dT40** was replaced by the non-complementary strand **dA40,** no cotton effect was observed via CD studies, thus **NT** binds specifically to **dT40**. **NT** could be replaced by **dA40**, to form **dA40-**

dT40, suggesting that hydrogen bonding in **NT-dT40** is less stable than **dA40-dT40**. The stability could be extended by using a π-conjugated oligo(*p*-phenylene)vinylene (**OPVT**) which has a larger aromatic surface area and thus increased *π-π* stacking interactions. It was found that, at high concentration of guests resulted in non templated self-assembly interfering with the templated self-assembly. From here Meijers group modified the methyl terminated ethylene oxide of the naphthalene guest molecule **NT** with a hydroxyl terminated ethylene oxide group to form a new second generation naphthalene derivative **NT-OH** *(Fig. 13)*.[52] The new naphthalene derivative **NT-OH** could suppress both the non-templated self-assembly of the naphthalene guest molecules and the further aggregation of the entire **dT40-NT** DNA hybrid complex. The **NT-OH** guest was self-assembled on oligothymine templates of various lengths (**dT*n,*** *n* = 5, 10, 20, 40), yielding objects whose size was controlled by the length of the template. CD and UV-vis temperature dependent studies implies binding almost immediately gives rise to a helical arrangement of the complex formed by ssDNA and bound guest molecules. They discovered that short templates were filled one by one whereas large templates first have alternating sequences of filled and empty sections, after which at a large fraction of bound sites, virtually all of the binding sites for all template lengths were filled. Correlation length not only depends on the fraction of bound sites but also on the interaction between guests on the template. Results showed a cooperative templated self-assembly process. Guest molecules bind to the ssDNA via hydrogen bonding and were held together by *π-π* and hydrophobic interactions. Smaller template sizes meant less efficient binding of the guest molecules resulting in a decrease in CD intensity. However decreasing template size allows the size of the aggregates to be controlled, with lower temperatures needed to get high coverage. Temperature-dependent UV measurements showed a stabilising effect of larger templates of the hybrid complex.

This finding is known to be true for helix-coil transition in biopolymers and for the formation of double stranded DNA in which at least four bases were needed for favourable hybridisation. However chain ends were more statistically disordered than the central parts due to fewer interactions with neighbours. Meijer developed a mathematical model to fit experimental data in order to determine how guests were distributed on the different template lengths. The model does not account for additional processes such as two or more guest-guest or host-guest interaction energies, self-aggregation or higher-order aggregation. Each binding site, could either be empty or occupied by one guest molecule. Together with the enthalpy of binding, a prediction could be made for the transition temperature for any given concentration of guest and template. Ideally, the correlation length could be further improved by increasing the attractive interactions between the guest molecules. Overall in these constructs the chiral ssDNA host templates a supramolecular strand of achiral chromophores to yield a right-handed helical organisation.

To enhance the efficiency of the self-assembly process, Jassen and Meijer reported a third generation achiral naphthalene guest derivative (**P**) *(Fig. 13)*.[53] This contained a diaminopurine hydrogen bonding unit and a larger π surface than diaminotriazine. Therefore **P** bonded to the **T40** through Watson-Crick type hydrogen bonding (similar to the previous studies with diaminotriazine derivatives) (40:1 ratio by CD studies, at higher temperatures binding is lower, previous for UV/Vis mixed 80:1). Unexpectedly the helicity of this DNA-templated assembly could be switched by changing the pH value as a result of protonation of the guest. 2,6-Diaminopurine guest molecules could bind to oligothymine strands and construct assemblies with a pH-switchable stability and supramolecular helicity. Helicity could be controlled by a variety of factors ranging from temperature, pH value, solvent, light or chiral stimuli. At low pH values, other forces come into play that may cause the stabilisation of a left-handed structure. Changing the pH from 9 to 2, the Cotton effect is reversed in which a left-handed hybrid DNA complex is formed at low pH.

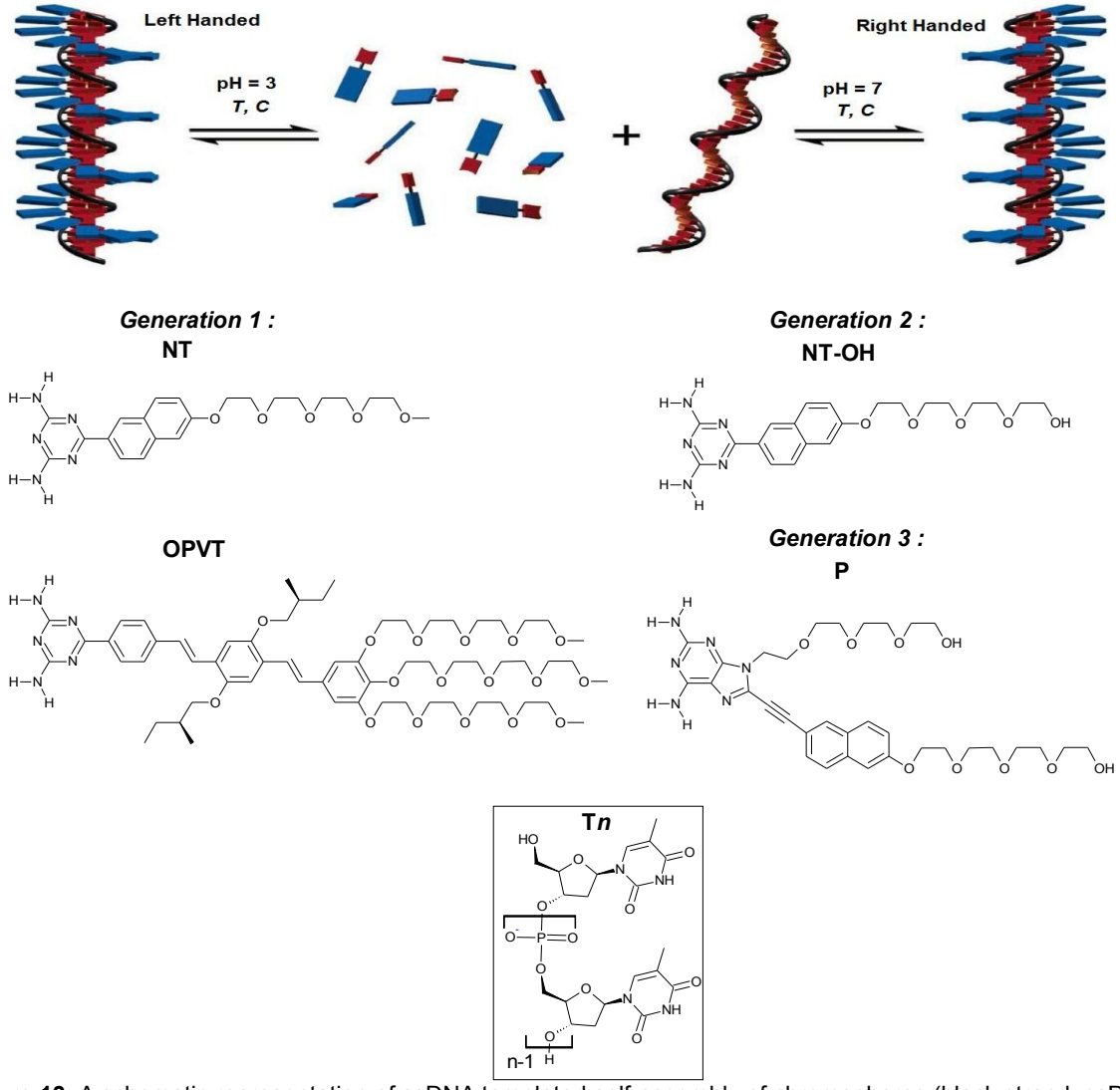


Figure 13: A schematic representation of ssDNA templated self-assembly of chromophores (black strand, ssDNA; blue bar, chromophores; red bar, hydrogen bonding unit) and molecular structures of **dTn**, first generation: **NT** and **OPVT**, second generation: **NT-OH** and third generation: **P** diamino chromophore units.

1.4 DNA Templating – Non-Enzymatic Replication

DNA transcribes information to form RNA molecules which allows for protein synthesis. In replicating systems, templates can be used to transfer the information. Complementary DNA strands are an example of excellent templates. These enable recognition, bind in a specific and predictable manner and bring reagents to predefined positions.

1.4.1 Modification of nucleic acids to enhance affinity

One concept to stabilise DNA/RNA complexes is to modify nucleic acids, as well as creating novel supramolecular polymers and hydrogels. Rayner and co-workers in 2004, performed studies aimed at modifying an existing nucleic acid to enhance its affinity for a complementary nucleic acid.[54] The oligonucleotide ligand *(Fig. 14* ***1****)* bearing a reactive amino group was allowed to react reversibly with a set of three aromatic aldehydes *(Fig. 14* ***2a-c****)* in an aqueous media. The aldehydes could stabilise the duplex structure by stacking on the outer base pairs while forming imine bonds with the amine functions of the terminal U'. The resulting unstable imines were trapped by reduction by $NaBH_3CN$ *(Fig. 14* ***3a-c****)*. One main product was obtained, and a control experiment with a non complementary DNA sequence (i.e., one that would not form a duplex) and measurements of the corresponding amines confirmed that the amplified product was indeed selected on the basis of its stabilising effect on the hairpin structure.

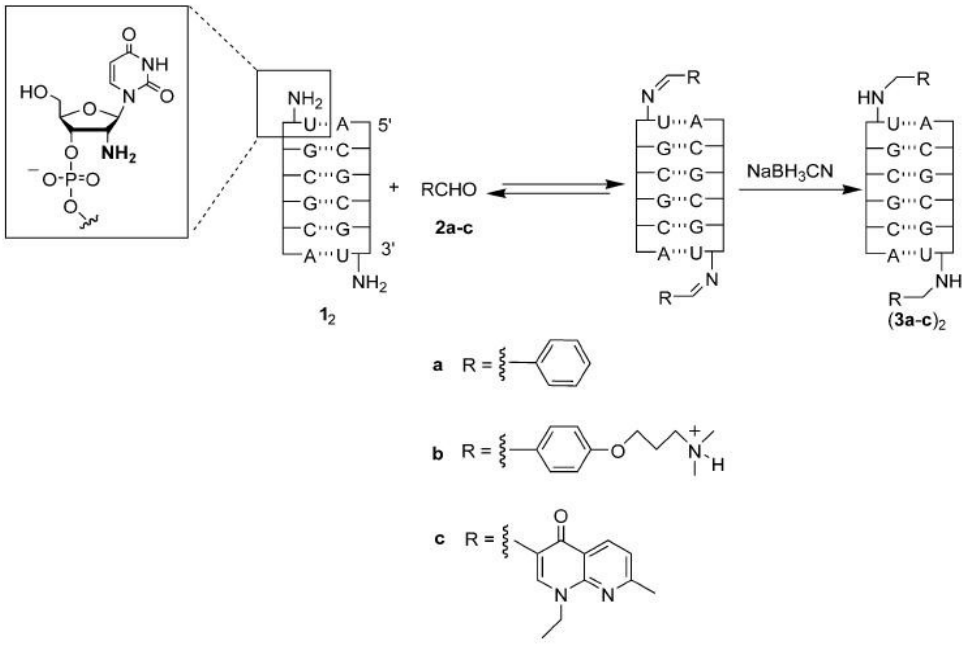


Figure 14: Selection by DCC of 3'-appended residues that stabilise a DNA duplex. Self-complementary oligonucleotide 1 bearing a 2'-amino-2'-deoxyuridine at its 3'-terminus is allowed to reversibly react with a set of aldehydes (benzaldehyde (2a), 4-[3-(dimethylamino)propoxy]benzaldehyde hydrochloride (2b), and nalidixic aldehyde (2c)) in an aqueous medium. This reaction provides a dynamic mixture of conjugated duplexes whose proportions were dictated by their relative thermodynamic stabilities. Subsequently, imines were reduced by NaBH3CN to chemically stable secondary amines, thus allowing HPLC analysis of the composition of the mixture. [54]

The same group used the same approach in a second publication.[55] Whereas the previous study had demonstrated that a functionalised complex could choose the building block that gives the best stabilisation of the complex, it was now shown that, in the presence of multiple sites for fixation of the building block on the complex, the most stabilising adduct was selectively formed and conventional synthesis of some of the individual adducts and melting point measurements confirmed that the most stabilising adduct had indeed been formed.

1.4.2 Replication

The spontaneous assembly of nucleotides on nucleic acid templates and the origins-of-life has inspired DNA nanostructures, [56] the dynamic assembly of DNA primers and the formation of PNA quadruplexes and metallocalixarenes. Templated directed synthesis based on DNA replication has gathered much interest with non-enzymatic self-replicating systems as the biological strategy for information storage and a current objective in bioorganic chemistry.[57–62] Studies carried out by Orgel, [49,63,64] von Kiedrowski, *[65]* and Lynn *[66,67]* have reported interesting findings with oligonucleotides to uncover synthetic molecular evolution mimicking Darwinian behaviour.

Lynn described a self-replicating monomer polymerisation method by using a poly-adenine DNA template with the addition of a thymine deoxyribose analogue equipped with a 5'-amino and a 3'-aldehyde group via reductive amination conditions ($NaBH_3CN$) instead of a phosphodiester reaction *(Fig. 15)*. The reaction followed step-growth kinetics exploiting pre-equilibrium imine formation on the template. As polymerization proceeds, the number of binding sites falls, but product/template affinity increases to enhance the later coupling steps. Balancing the rates of these steps is expected to be critical for controlling reaction fidelity along longer templates. This replication monomer method showed the ability to translate biopolymer-encoded information stereoselectively into sequence- and chain-length specific synthetic polymers using a DNA template.[68]

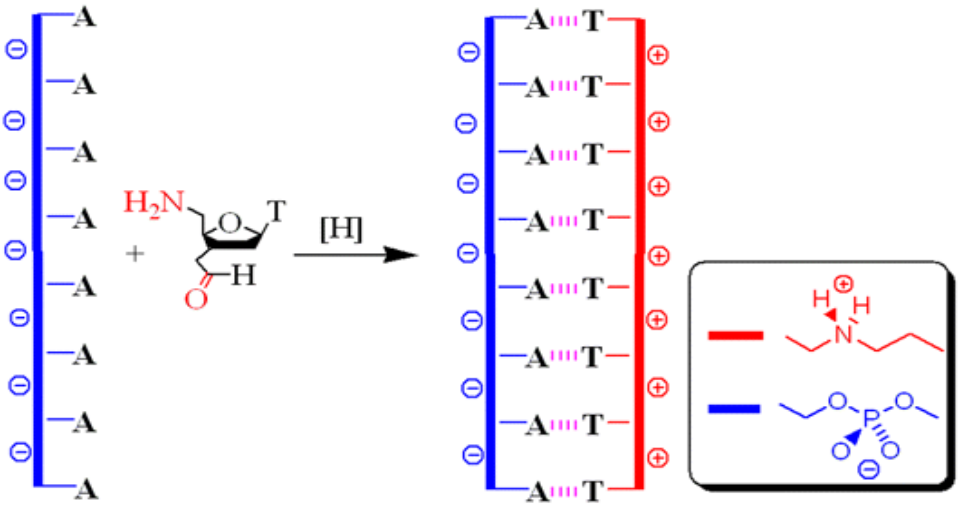


Figure 15: Template-directed polymerization of monomer (**T**) by **$(dA_p)_8$**

1.4.3 Nucleic acid templating

Benner [69] in 2006, was inspired by Lynn's research, [66,67,70,71] and utilised nucleic acid's with reversible ligation and thermodynamic stable complexes. This is in contrast to the irreversible template ligation suggested by Kool, [72] von Kiedrowski, [59,65] and Ellington, [73] as these result in kinetic traps such as partial mismatches, ligated with nearly the same frequency as full matches. Benner demonstrated the use of dynamic equlibria to assemble in situ composite DNA polymerase primers, having lengths of fourteen or sixteen nucleotides, from DNA fragments that are six or eight nucleotides in length. The fragments were joined together by an imine linker *(Fig. 16)*. The rationale behind the dynamic primer assembly was to increase primer specificity. They used template-directed assembly of dynamic primers to detect nucleic acids with single mismatch discrimination, but with a specificity equivalent to that displayed by a long oligonucleotide primer. In just one implementation, employing a 8-mer as the 5'-fragment carrying a 3'-aldehyde, a 6-mer as the 3'-fragment carrying a 5'-amino group, the 9ºN DNA polymerase, and a temperature of 70 °C, high discriminatory power could be achieved across the entire length of the composite primer. This primer formation approach is an interesting example of dynamic assembly applied to templating nucleic acids.

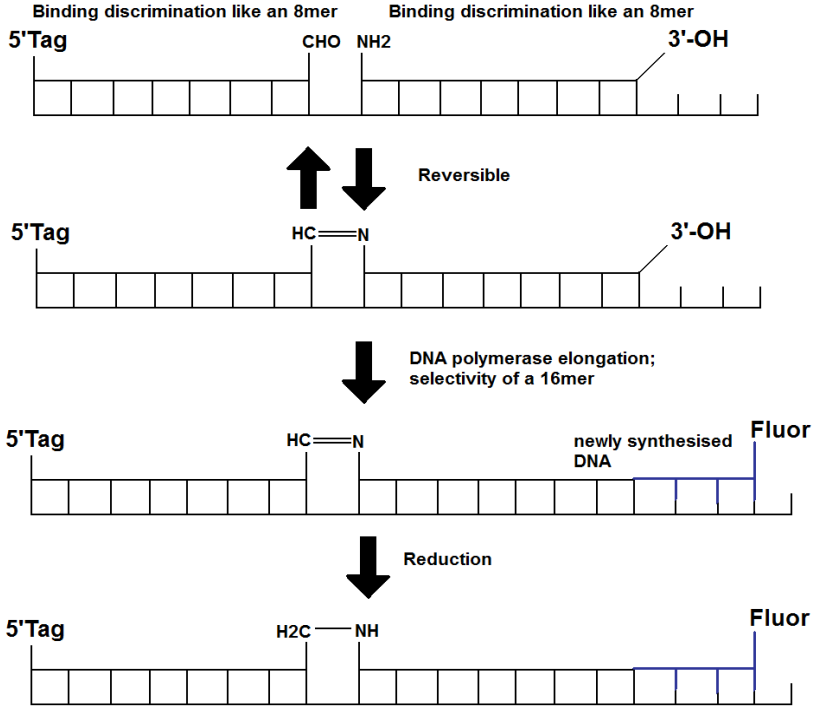


Figure 16: Architecture of an assay for a DNA template that dynamically assembles its own composite primer from two fragments. Two DNA fragments, each 8 nucleotides in length, terminated with 3'-CH_2CHO (on the 5'-DNA fragment) and 5'-NH_2 (on the 3'-DNA fragment) should reversibly form a composite, joined via an imine linker under conditions of dynamic equilibrium. Imine formation is easily reversible in aqueous conditions, and this reversibility ensures that the tightest binding complement perfectly matched to the template is formed. Should this composite be able to prime the synthesis of DNA using a DNA polymerase, the specificity of priming should be characteristic of a 16-mer, as both sequences must be adjacent on the template for priming to occur. The discrimination against mismatches, however, should be that characteristic of an 8-mer, and therefore be very high.

1.4.4 Self-assembling sequence-adaptive peptide nucleic acids

Ghadiri et al utilised reversible reactions involving nucleobase recognition units which self-assemble onto oligodipeptide backbones consisting of thioester peptide nucleic acids (tPNAs).[74] The oligomers self pair with tPNAs strands and cross-pair with RNA and DNA in a Watson-Crick fashion. Cysteine (Cys) residues at alternating amino acid positions enabled anchoring for reversible tethering of nucleobase thioester monomers from solution via transthioesterification reactions *(Fig. 17)*. This allowed the nucleobases to attach and disassemble on their own without enzymes, so that a given peptide strand could hold a shifting array of nucleobases. A variety of amino acid residues were used at the flanking positions of the backbone (Glu, Asp, Arg, and Gly) to provide spacing between Cys residues, improve aqueous solubility, increase backbone flexibility and enhance electrostatic properties. These features contributed towards improved interactions with nucleic acids.

The nucleobase thioester was allowed to equilibrate with tPNA strands. As such, when adenine thioesters were employed, base pair selectivity was observed with the addition of a **dT20** DNA strand. However the same experiment with noncomplentary **dA20** did not result in cooperative melting transition. With no DNA template the ratio of bases was 1:1, with no preference for either nucleobase. In the presence of a DNA template, various nucleobase thioesters were then enriched depending on the nature of the template. For example, the adenine species was selectively enriched in an approximate 3:1 ratio in the presence of **dT10** or **dT20** strands. Similar observations were made with 7-deazaguanine moieties elevated to an approximate 3:1 ratio in the presence of a **dC10** or **dC20** template. This demonstrates amplification of a specific complementary species that binds to the oligonucleotide via Watson-Crick pairing interactions upon the addition of a template, resulting in re-equilibration.

They also investigated templated-assembly reactions involving more complex templates by preparing a series of biotinylated mixed-sequence DNA templates containing different ratios of cytosine and thymine nucleotides. From this mixture they observed a linear relation between the percentage of cytosine nucleotides present in the DNA template and the percentage of 7-deazaguanine thioester anchored to the peptide at equilibrium. This finding supports the potential of the oligomers to assemble in a sequence-specific manner in response to mixed-sequence templates. In another experiment the peptide [Aba-$(ECGC)_5$-$CONH_2$] was incubated with the **adenine** and **7-deazaguanine** thioesters, and **dT20** template. This resulted in predominantly adenine nucleobases anchored to the peptide.

An excess of **dA20** and **dC20** DNA strands was then added to the reaction. The greater length DNA strands relative to the peptide meant **dA20** displaced any assembled oligomer that was bound to the **dT20** template, thereby preventing **dT20** from further templating

oligomer assembly thus leaving only the **dC20** strand available in solution to act as a template. Exchange involving the **7-deazaguanine** monomers in solution would eventually make possible nucleation of productive interactions between **dC20** and the oligomer. HPLC showed that four hours after the addition of the **dA20** and **dC20** DNA strands, a substantial enrichment of anchored **7-deazaguanine** moieties relative to the analysis before the addition of **dA20** and **dC20.** This is evidence that oligomers are capable of a dynamic response to changing external conditions.

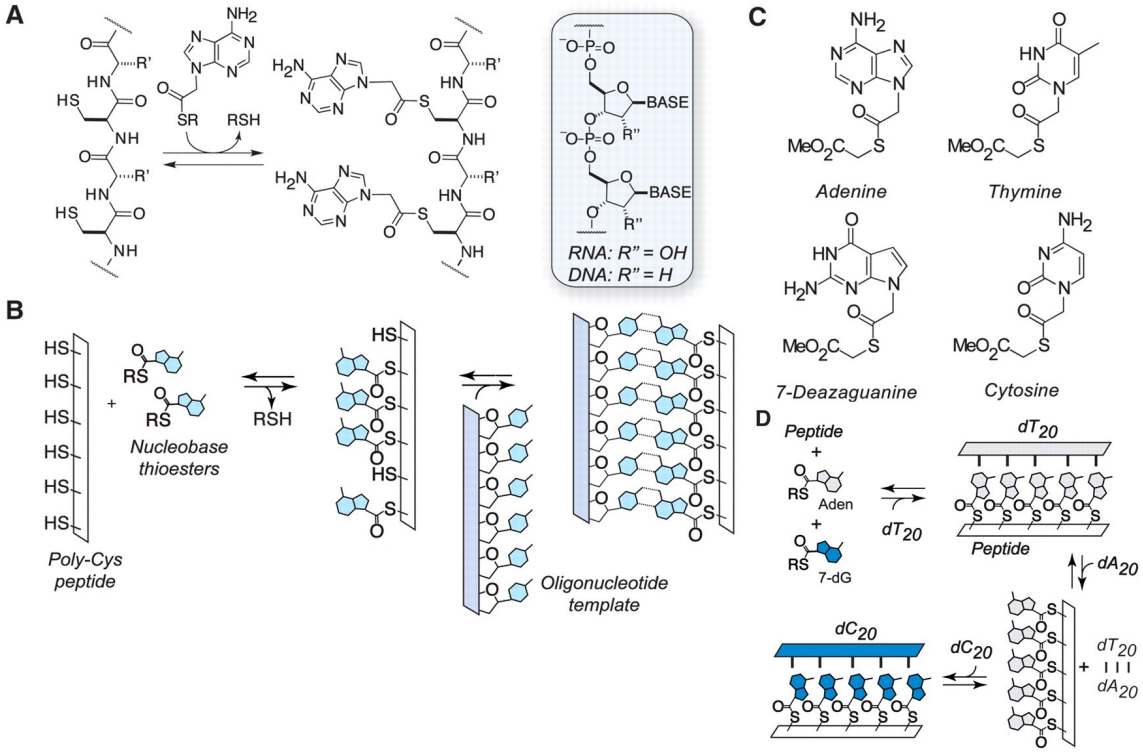


Figure 17: Chemical structure and assembly mechanism of the dynamic oligomers. **(A)** Mechanism for reversible covalent assembly of the oligomers via anchoring of thioester-derived recognition units (adenine is shown) onto an oligopeptide backbone. RNA and DNA structures are drawn for comparison (in the box at right). **(B)** Illustration of oligomer assembly and binding to a complementary oligonucleotide template. **(C)** Structures of the nucleobase thioester monomers used in the present study. Me, methyl. **(D)** Outcome of incubation of peptide [Aba-(ECGC)5-$CONH_2$] with adenine and 7-deazaguanine thioesters, initially in the presence of a dT20 template for 2 hours, and after 4 additional hours in the presence of dA20 and dC20 DNA strands, leaving only the dC20 strand available to serve as template because of duplex formation between dT20 and dA20.[74]

Ghardiri's work showed DNA analogues which could assemble and dissemble themselves without the use of enzymes. This research may show the way for possible materials that repair themselves or transform in response to their environment and the inspiration for this thesis is developed upon this work. It has provided the background knowledge to design and develop highly selective nucleobase receptors for applications in polar media of biological interest and a better understanding of nature's architecture of biological systems. The main inspiration for nucleobase receptors has been derived from Zimmerman's [38] and Odashima's [31,32] work.

2 Objectives

The complexation selectivity, particularly to adenine over all other nucleobases, is still insufficient and remains challenging. Primarily this is due to the adenine nucleobase not being furnished with three neighbouring hydrogen-bonding/coordinating sites in contrast with other nucleobases.

The key objective of the project described in this thesis was the use of DNA as a template for information transfer with non-enzymatic replication, using non-covalent interactions while incorporating dynamic combinatorial chemistry (DCC). Another aspect to this project is the use of an oligonucleotide template to resolve a mixture of compounds for sequence selection. In order to achieve this goal, focus towards nucleobase recognition was essential *(Fig. 18)*. A new planar adenine and guanine receptor was designed and synthesised for this purpose.

To measure the effectiveness of receptor binding, a titration study was carried out in organic media and analysed by NMR. It was also planned to test for an “abasic site” in which one purine was removed from an oligonucleotide duplex. A CD experiment in water, enabled an investigation to see if the clamp could bind specifically to its corresponding base. From here, nucleobase assembly was envisaged. The use of a single stranded oligonucleotide was expected to enhance nucleobase hybridisation due to noncovalent interactions and *π-π* stacking leading to an overall cooperative effect. Initially a poly-A template was used, followed by insertion of a mismatch. This evaluation was expected to give more information about whether a mismatch could enhance stability or disrupt the system. From here, prospective tests could determine if it is possible to displace the A Clamp aggregates (dA^*_n) from the poly-A strand (dA_8) with that of a poly-T strand (dT_8)? Thus testing which strand is the strongest bound, dA_8-dA^*_8 or dA_8-dT_8.

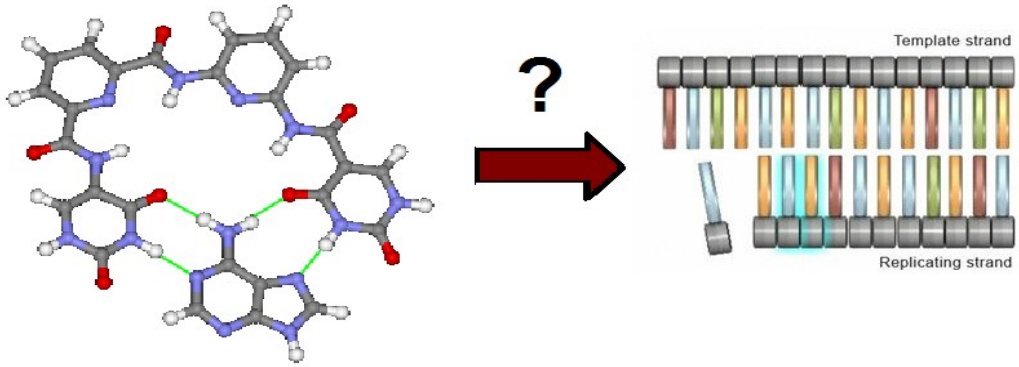


Figure 18: Molecular design of adenine receptor bound to its corresponding base and testing its ability to bind to a template strand in the presence of a mixture of receptors to form a replicating strand.

3 Design Strategy

The aim of this project was to develop receptors which could bind to nucleobases. It was envisaged that these receptors could self-assemble upon addition of an oligonucleotide template to resolve a mixture of compounds for sequence selection. It was decided to incorporate into the receptors the ability to simultaneously bind via four hydrogen bonding sites. The use of non-covalent Watson-Crick and Hoogsteen sites would enable recognition and selectivity. Some other design concepts incorporated into the receptors included planar aromatic surfaces to maximise the possibility of *π-π* stacking and thus increasing the overall stability via a cooperative effect. These electrostatic interactions and conjugation allowed for a predefined host molecule. The receptors contained solubilising groups, which when protected were hydrophobic, and when unprotected were hydrophilic, this feature enabled a functional molecule in both organic and aqueous solvents. Zimmerman [38] and Odashima [31,32] have studied molecular recognition of guanine and adenine derivative receptors respectively, in organic solvents which had medium affinity. However, a limited amount of research has investigated nucleobase receptors in water and this was envisaged for studies regarding the adenine and guanine receptors (A and G clamp). With regards to the design of an A clamp, several possibilities were conceived via molecular modelling studies *(Fig. 19, 20)*. It was not certain if the synthetic routes were feasible for each clamp. Attention was given to a short flat foldamer crescent due to previous expertise towards synthesis of similar molecules within the Huc group. *(Fig. 19 e))*.

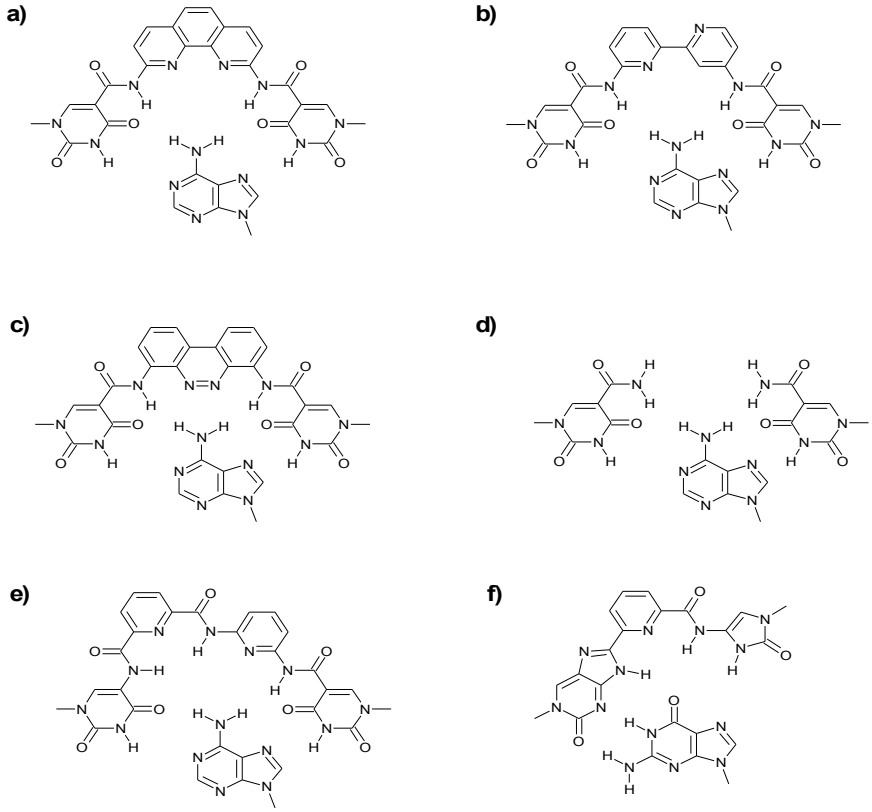


Figure 19: Suggested A clamp receptors for synthesis.

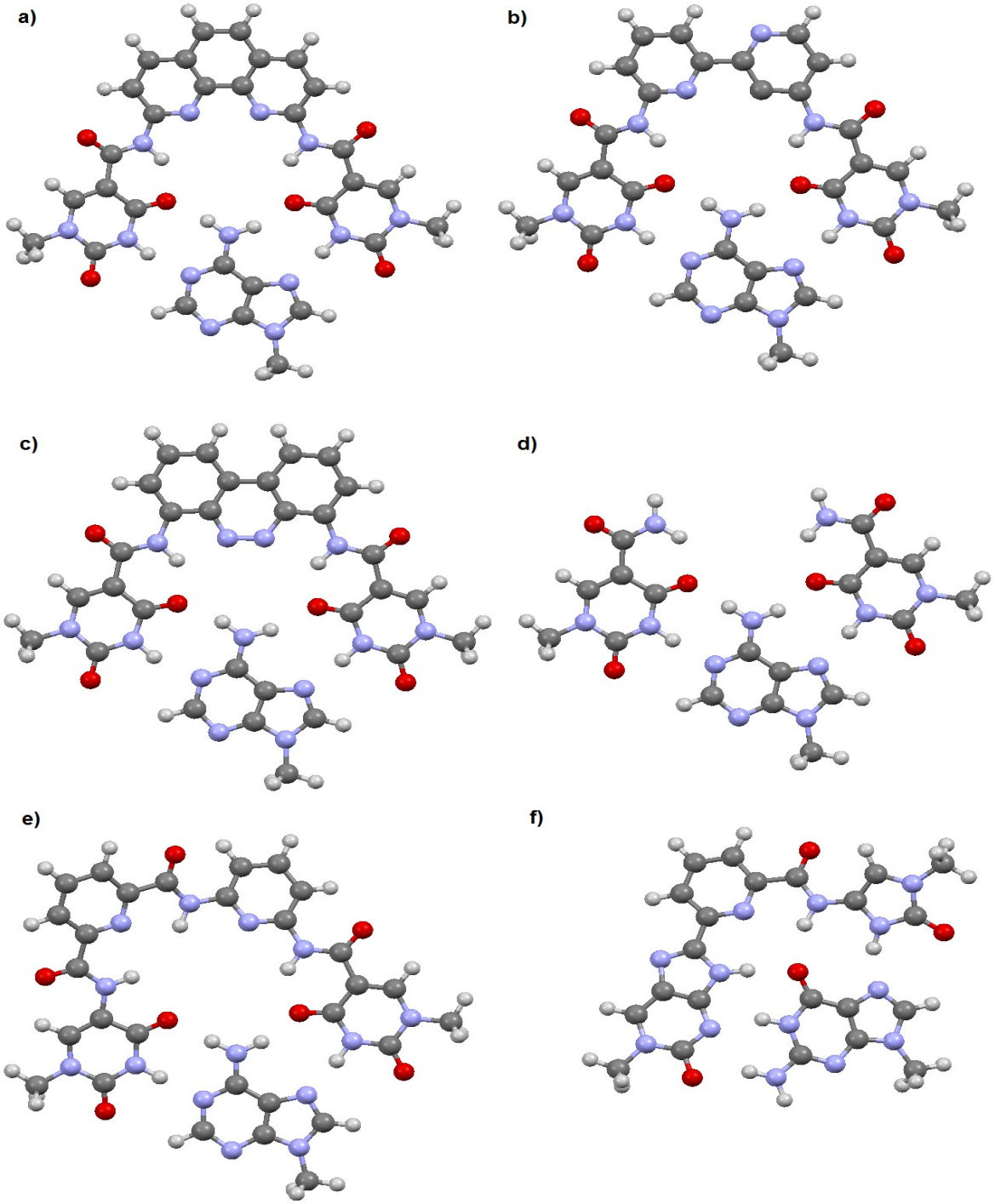


Figure 20: Suggested A clamp receptors for synthesis (a-e) and G clamp (f). Energy minimised models. { e) was chosen to synthesise.}

This short foldamer crescent *(Fig. 21)* contained a linker which has similarities to previous structures within the Huc group. The two terminal uracil groups were required for recognition and were bridged by a linker with recognition being pH dependant. It has an ADDA (acceptor, donor, donor acceptor) array to complement the adenines DAAD (donor, acceptor, acceptor, donor) hydrogen bonding array. Intramolecular hydrogen bonding within the linker group, has pitched the receptor at a 39˚ angle *(Fig. 21)*. The cinched receptor provides optimum orientation for Watson-Crick (2.918 & 2.866 Å) and Hoogsteen bonds (2.844 & 2.876 Å).

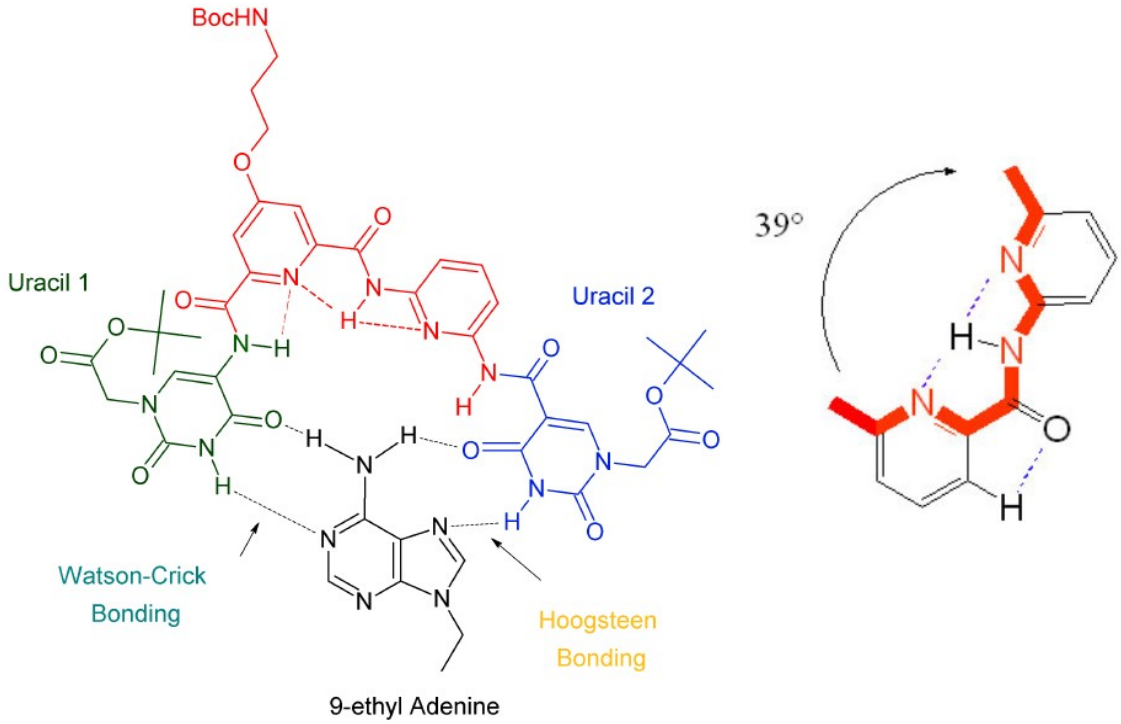


Figure 21: (Left) Adenine receptor binding to 9-ethyl adenine via Watson-Crick and Hoogsteen bonding. It contains two uracil derivatives for recognition joined by a central linker (red) (Right) Intramolecular hydrogen bonding of the linker, has pitched the receptor at an angle of 39°.

The guanosine receptors chosen were similar to those published by Umezawa and Zimmerman respectively. Additionally a water solubilising side chain was incorporated. Due to the smaller planar aromatic surface, it was expected that hybridisation would be weaker than that of the A clamp. The main purpose of synthesising these G clamp molecules was to compare them to published experimental data. Subsequently, these G clamps could be utilised later on in more critical experiments to determine selectivity between the A and G clamps.

Synthesis and methodology, and the challenges encountered with implementing such desired design features, will be elucidated in the following section (Section 4). The main obstacles included poor solubility and nucleophilicity issues of manipulating an amphiphilic receptor in organic solvents. Consequently these complications inhibited a quick and efficient progress to develop and analyse the receptors in aqueous solution. Initially the A and G clamps were to be tested for their ability to form non-covalent Watson-Crick and Hoogsteen hydrogen bonds with their respective bases in organic solvent (chloroform), and this interaction was to be monitored by NMR.

Aqueous studies were planned to be carried out using surface plasmon resonance (SPR) and CD. Preferentially SPR, would enable the targeting of an abasic site in aqueous buffer. The prerequisite, was to test the clamp **(A*)** for its ability to restore specifically *π-π* stacking in a DNA duplex in which one nucleotide was deleted *(Fig. 22)*.

This test was required due to the competitive effects in aqueous media and to establish how effectively the receptor could bind. This test was expected to verify the ability for the clamp to bind specifically to its corresponding base in aqueous solvent. Due to the relatively small nature of the receptors, a minuscule change in the SPR sensitivity would be compensated by an overall change in mass of the entire oligonucleotide-receptor complex. Thus SPR would detect the overall effect of the receptor on the oligonucleotide. *(Fig. 22)*

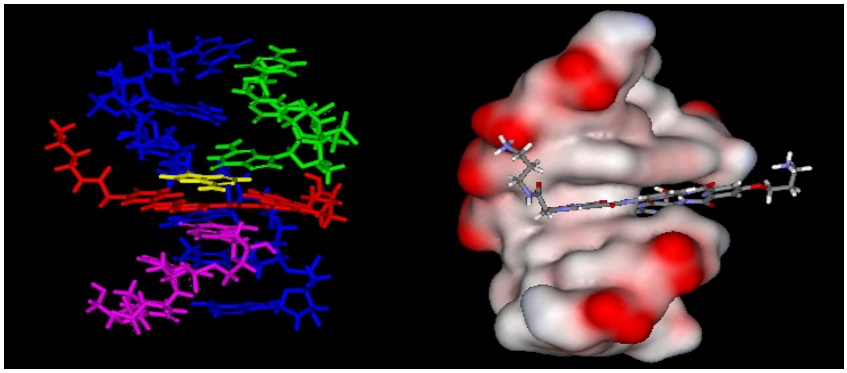


Figure 22: Targeting an abasic site. Oligonucleotide 10 mer with adenine base. In the corresponding duplex a purine is missing. Adenine clamp **(A*)** (red) is introduced to restore binding.

The initial dynamic combinatorial library (DCL) was expected to consist of A and G clamps, to investigate their non-covalent interaction's upon poly A and poly G oligonucleotides (template 10 mer) *(Fig. 23)*. Sequentially, it was thought that hybridisation could occur between the clamps specifically corresponding to their respective nucleobase contained the oligonucleotide template. This interaction was suspected to be aided by a driving force of *π-π* stacking interactions and hydrophobic effects of water thus increasing affinity to the strand and the formation of a self-replicating complementary strand via hybridisation. The initial binding was expected to be critical. Followed by a stabilisation, due to a cooperative effect when no other impeding forces exist. This experiment was expected to elicit some fundamental questions. Could the clamp bind efficiently to its respective base and did a strong cooperative effect exist? What was the shortest template length required for self-assembly? Was there be a significant difference in binding strength between the various template lengths? The introduction of a variety of template lengths: (T-A_n-T) with n = 3, 5, 8 and 13 would address these questions. Was it possible to disrupt the poly A-A clamp duplex (dA_8-dA^*_8) with a poly T strand (dT_8) resulting in a poly A-poly T duplex (dA_8-dT_8)? Subsequently, answering these questions would lead to further experiments involving the insertion of a mismatch base into the oligonucleotide strand *(Fig. 23)*. This would test the clamps ability to discriminate between bases. Thus the crucial overall question would be if the clamps could exhibit selectivity.

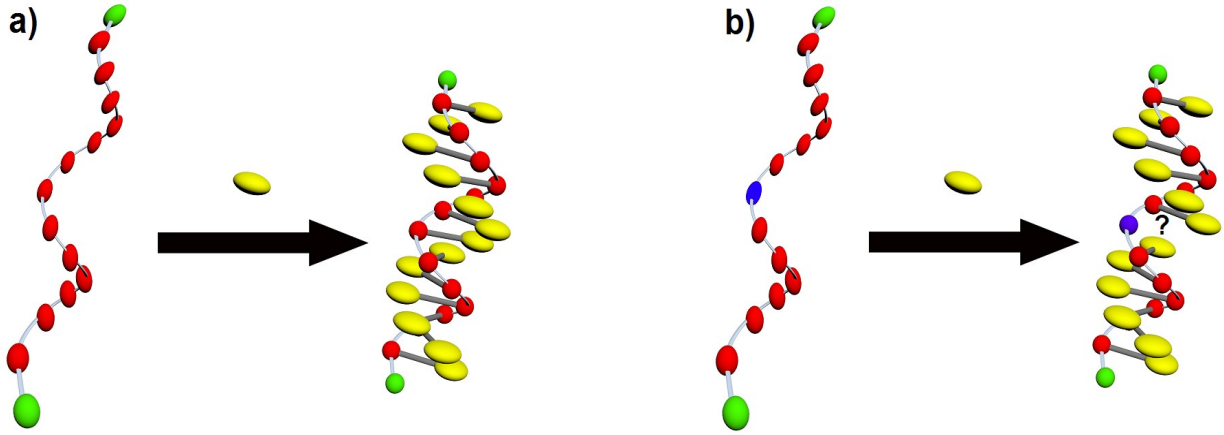


Figure 23: Step 1 (a): Poly A (or poly G) oligonucleotide strand (red) to be attached to a sensor chip, and a solution of A* (or G*) clamp receptors (yellow) is to be passed through the flow channel, binding non-covalently (grey) upon the oligonucleotide template forming a helix. This experiment measures the efficiency of binding and recognition. Step 2 (b): Insertion of a guanine mismatch (blue) is intercalated onto the template sequence of poly A oligonucleotide, to evaluate if it enhances stability or disrupts the system.[75]

Consequently, after favourable interactions in the preliminary homogeneous DCL has been achieved, more complex DCC assemblies could be generated to deduce the self replication of nucleobase receptors upon an oligonucleotide template. A heterogeneous A and G oligonucleotide template attached to the sensor chip would decipher if the A and G clamps could discriminate between their respective bases *(Fig. 24)*. It was expected that upon addition of four equivalents of A clamp, followed by addition of four equivalents of G clamp, in the absence of a template no self-assembled duplex would form. To this, the introduction of an oligonucleotide template (8-mer, containing four adenine and four guanines) is expected to induce binding resulting in the formation of a duplex. Additional equivalents of A or G clamp to the system is expected to not have any impact as the oligonucleotide template (8-mer) is expected to be completely filled. Overall it was expected that answering many of these questions, would achieve the final goal of sequence selection and translation of information using oligonucleotides as a template.

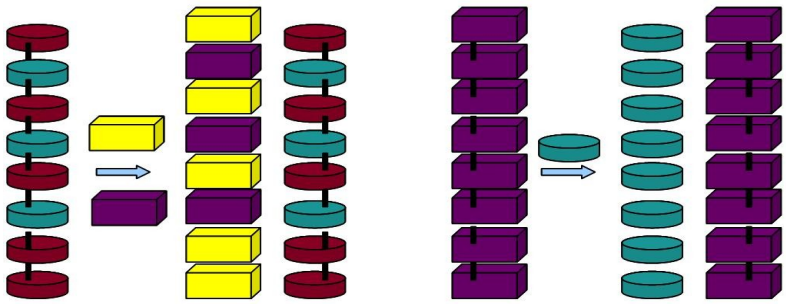


Figure 24: Step 3 (Left): Clamps discriminate between their respective bases. A variety of templates containing different arrangements of adenine and guanine nucleobases increase the possibilities of the combinatorial library. Specificity and self-assembling of the receptors was envisaged. Step 4 (Right): The clamps could also be arranged in a template fashion and the adenine and guanine nucleobases hybridised onto this system. This further explores the dynamic chemistry of the receptors.

4 Synthesis

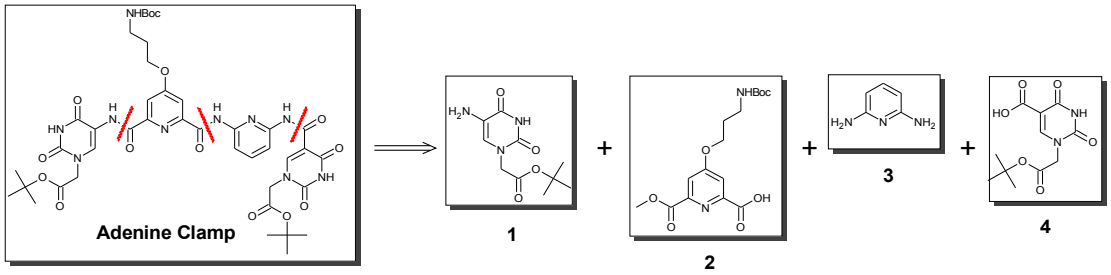
Synthesis towards the adenine and guanine receptors initially seemed straightforward, however, a few complications such as poor solubility and nucleophilicity issues were encountered. The strategic route and methods to overcome these challenges is described in the following sections (Section 4.1 and 4.2).

4.1 A Clamp

The synthesis of the chosen A clamp involved steps which had previously been investigated in our group and for that reason was preferred over the other molecular designs.

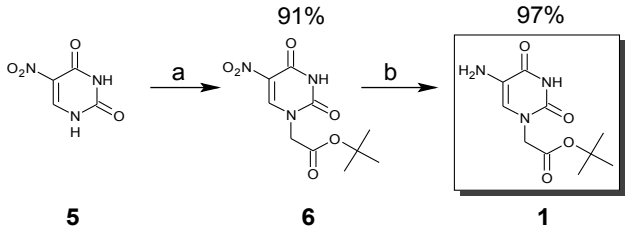
4.1.1 Proposed route 1

Looking at the A clamp from a retrosynthetic perspective *(Scheme 1)* three main building blocks, **1, 2,** and **4** seemed straightforward to construct and assemble together, with the fourth monomer 2,6-diaminopyridine **3** being commercially available. The first proposed synthetic route involved coupling **2** and **3** together, based on previous protocols published by our group. Previous work also established that coupling **1** and **2** presented no difficulties. Based on this knowledge the initial strategy was to couple the precursors **2** to **3**, then to **4** (which was thought to be a more problematic step) followed by **1** to form the final A clamp. This approach and the complications encountered (leading to this route being revised) are now described.



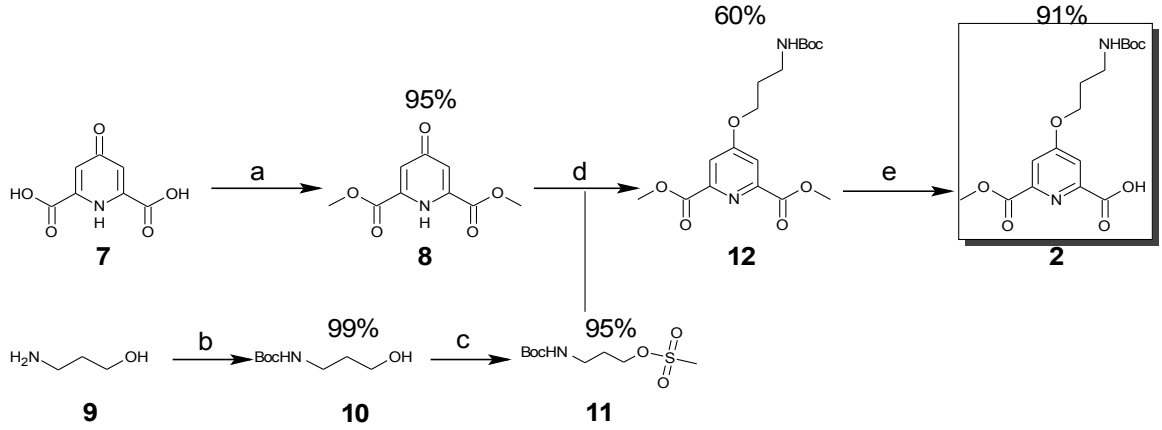
Scheme 1: Retrosynthetic analysis of the **adenine clamp.** Four building blocks (**1, 2, 3, 4**) were suitable for assembling the desired A clamp.

Synthesis of the 5-amino uracil derivative **1** proceeded smoothly *(Scheme 2)*. For the first step to form **6**, regioselective alkylation of **5** was carried out with *t*-butyl bromoacetate in the presence of TEA, and not DBU as previously described. According to the literature,[76] when TEA (pK_a = 10.8) is used as a deprotonating agent, exclusively N-1 adducts were obtained in a Michael-type addition reaction. However, in the presence of DBU (pK_a = 12) in DMF N-3 adducts were isolated as the major products. HMBC NMR analysis confirmed the correct product **6** was formed. For the reduction of the nitro group to form product **1**, hydrogenation conditions were employed. It was thought that the double bond within the uracil might be reduced so careful TLC monitoring ensured this was minimised. Column chromatography was required for purification. Note that the amino group could degrade and this step was carried out only when product **1** was required, to avoid oxidation over time.



Scheme 2: Synthesis of 5-amino uracil derivative 1: a) *t*-butyl bromoacetate, TEA, DMF, 0 °C to give **6** (91% yield); b) Pd/C 10%, H_2, MeOH/EtOAc, RT to give **1** (97% yield).

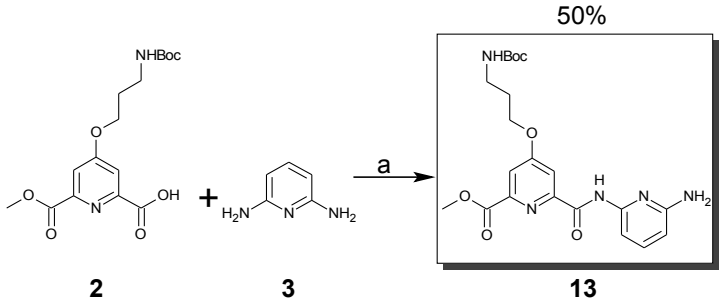
The synthetic route to obtain monomer **2,** involved esterification of commercially available chelidamic acid **7** to afford the dimethyl chelidamic ester **8** *(Scheme 3)*.[77] Simultaneously the Boc protected mesylated side chain **11** was prepared. To the 3-aminopropanol **9** was added Boc-anhydride to give 3-Boc-aminopropanol **10** in quantitative yield. The mesylate chloride was added to afford **11**.[78] It is important to point out that standard Mitsunobu conditions were not possible for the addition of side chain **11** to the diester **8** in forming product **12** *(Scheme 3)*. Instead nucleophilic substitution was employed to form **12,** which is similar to other side chain additions to chelidamic diesters developed within the group.[79] To compound **8** was added potassium carbonate in anhydrous DMF and heated to reflux for two hours. This was followed by cooling to RT and addition of the mesylated side chain **11**, continued by heating to 100 °C for two hours. Work up was then carried out and the product was purified by column chromatography to give monomer **12** in 60% yield.[79] The saponification step was carefully monitored by TLC so that it could be quenched when the mono carboxylic acid formed, to avoid formation of the bis-substituted carboxylic acid.



Scheme 3: Synthetic route for monomer 2: a) 2,2-dimethoxypropane, MeOH, HCl, 60 °C to give **8** (95% yield); b) Boc_2O, THF to give **10** (99% yield); c) CH_3SO_2Cl, THF, TEA to give **11** (95% yield); d) K_2CO_3, DMF, 100 °C to give **12** (60% yield); e) NaOH, MeOH/1,4-dioxane to give **2** (91% yield).

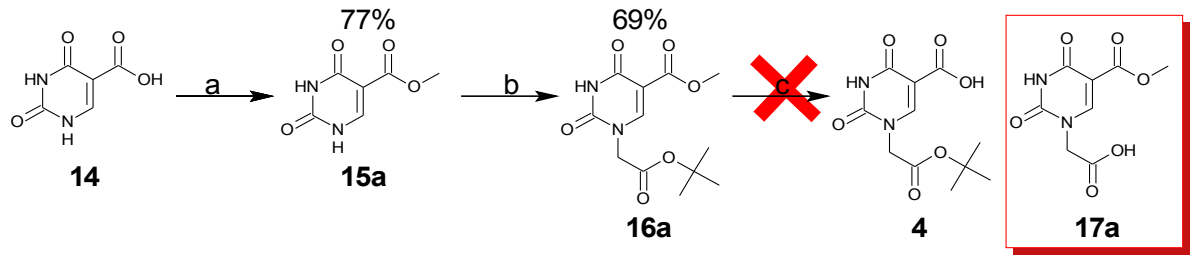
Due to previous knowledge involving coupling of pyridine carboxylic acid derivatives similar to **2** with 2,6-diaminopyridine **3**, it seemed like a natural progression to synthesise the central linker unit **13** *(Scheme 4).* The synthesis for **2** has been discussed previously. Regarding the linker unit **13** synthesis, it was important to monitor the rate of addition of the **2** PyBOP/HOBt ester derivative to that of 2,6-diaminopyridine **3** during the coupling step in dilute conditions.

This is because it could potentially form the bis-substituted compound if the rate of addition was too fast. Excess 2,6-diaminopyridine **3** was easily removed in the work-up extraction as it was soluble in the aqueous layer. Note, that the mono Boc protected aminopyridine could be used, but this would be an additional two extra steps onto the procedure (involving protection and deprotection). This coupling step was validated with no major difficulties encountered to form the linker **13**.

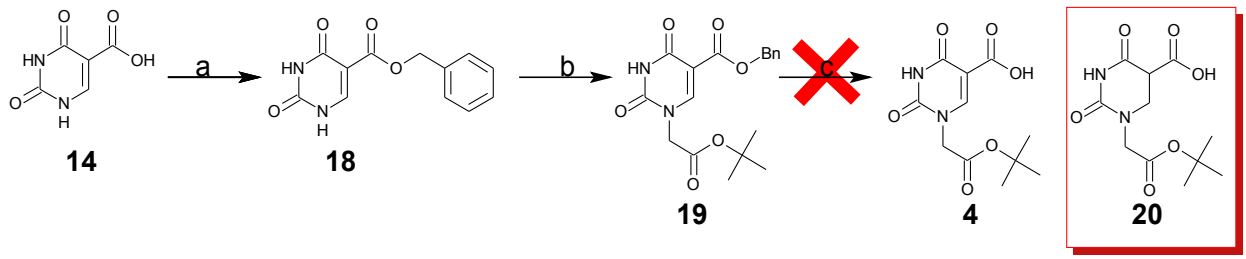


Scheme 4: Synthesis of the central linker by coupling **2** and **3** a) PyBOP, DIEA, DCM to give **13** (50% yield).

The synthesis of the iso-orotic acid derivative **4** initially proved troublesome *(Scheme 5)*. Unexpectedly saponification of **16a** did not give the desired product and the *t*-butyl group was cleaved giving **17a**.[80] The next strategy was to add a benzyl ester to **14** giving compound **18** followed by addition of *t*-butyl bromoacetate to produce **19** *(Scheme 6)*. However during hydrogenation to remove the benzyl group, the iso-orotic double bond was destroyed yielding product **20** and in future a selective catalyst should be used. Since selectivity of the ester to give the desired carboxylic acid **4** was not possible, an alternative route was implemented. It was thought to add a Boc aminopropyl **23** side chain onto compound **17a** *(Scheme 7)*. It was then envisaged to produce **25** and to couple this to the linker **13**.

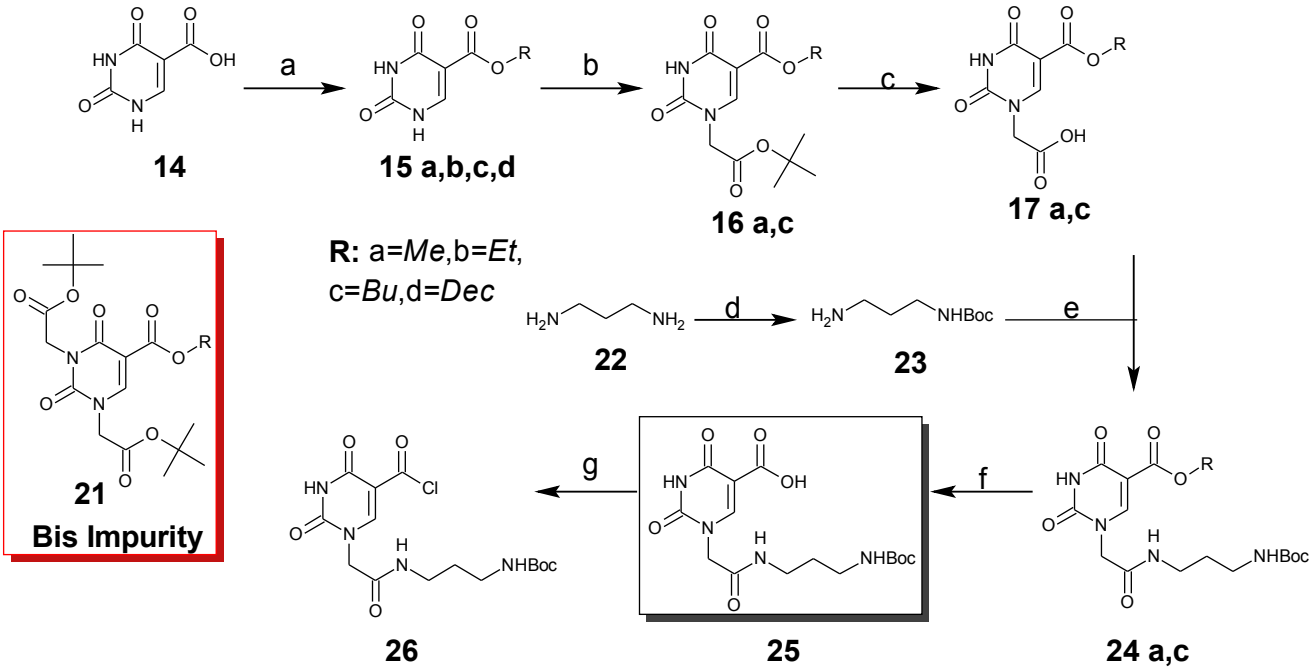


Scheme 5: Iso-orotic derivative synthesis **4:** a) MeOH, 6 eq, $SOCl_2$, Reflux 5 nights to give **15a** (77% yield); b) K_2CO_3 0.5 eq, *t*-butyl bromoacetate, DMF to give **16a** (69% yield); c) NaOH, H_2O/1,4-dioxane.



Scheme 6: Iso-orotic derivative synthesis 4: a) BnOH, 6 eq, $SOCl_2$, Reflux 5 nights; b) K_2CO_3 0.5 eq, *t*-butyl bromoacetate, DMF; c) Pd/C 10%, H_2, MeOH/EtOAc.

Amide coupling of **25** to the linker **13** was unsuccessful on the first attempt to produce the trimer **27** *(Scheme 8)*. More material was needed to be synthesised due to loss of product **25** through its poor solubility in organic solvents. To overcome these poor solubility issues, it was decided that addition of a long fatty side chain could improve the solubility in organic solvents. For this reason decyl alcohol was used for the esterification of **15d** *(Scheme 7)*. However, the actual decyl alcohol itself was difficult to remove even with use of a Kugelrohr, *in vacuo* and at 220 °C. Thus it was a challenge to isolate pure **15d**. Ethanol was used for the first esterification step to form compound **15b** (18% yield), however its properties were similar to that of the methyl iso-orotic ester derivative **15a**. It was decided to continue the synthesis with **15a** instead. It was finally decided that the use of butanol for esterification to form **15c**, would enhance solubility and improve ease of synthesis. Major solubility issues were still encountered, but it was hoped that forming the butyl ester would improve synthesis and somewhat resolve this problem.

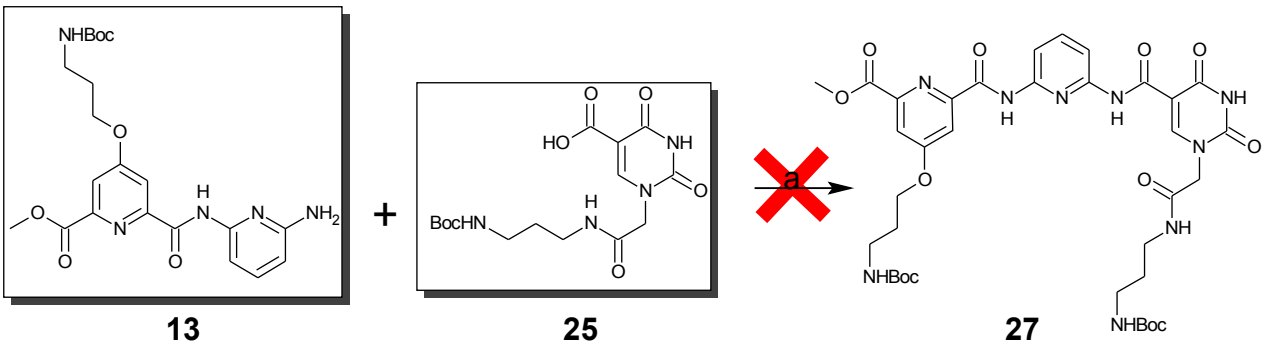


Scheme 7: Iso-orotic derivative synthesis: a) ROH, 6 eq, $SOCl_2$, Reflux 5 nights to give **15a** (77% yield), **15b** (18% yield), **15c** (98% yield), **15d** (not isolated); b) K_2CO_3 0.5 eq, *t*-butyl bromoacetate, DMF to give **16a** (69% yield), **16c** (82% yield); c) TFA, DCM to give **17a** (95% yield), **17c** (90% yield); d) Boc_2O, THF; e) PyBOP, DIEA, DMF to give **24a** (67% yield), DMAP also added to **17c** to give **24c** (77% yield); f) NaOH, H_2O/1,4-dioxane of **24a** gave **25** (54% yield), of **24c** gave **25** (78% yield); g) Chloroenamine, DCM.

The next challenge, was addition of *t*-butyl bromoacetate to form product **16c**. NMR analysis showed the bis-substituted compound **21** was obtained. Careful consideration of the rate of addition and using fewer equivalents of bromo *t*-butyl acetate, prevented this occurring with 82% yield obtained for **16c**. TFA removed the *t*-butyl group to give the carboxylic acid **17c**. Addition of the mono Boc aminopropyl side chain **23** to **17c** was possible with PyBOP coupling and addition of DMAP.

Some product was found to be in the aqueous layer and numerous organic extractions were required with a addition of some methanol to give **24c**. PyBOP was the major impurity in this step, however any attempt to purify by column chromatography, resulted in the poorly soluble desired compound precipitating on the silica. Thus the by-products needed to be removed during the aqueous extraction process. In the case of **24c** this compound was obtained pure (77% yield) after aqueous extraction. Regardless of which ester is used, the saponification step results in a common carboxylic acid, **25** despite its poor solubility. Continuous organic extractions from the aqueous layer were required to obtain **25** via **24a** [80] (54% yield of **25**) and **24c** (78% yield of **25**), with TLC (DCM/MeOH, 9:1) showing a new product **25** formed on the baseline. This gave a reddish spot when stained with ninhydrin and was very fluorescent under UV light. This compound was only soluble in DMSO-d_6 for NMR analysis. A large batch of **25** was produced again to compensate for possible low yields. Not enough material being readily available hindered synthesis to investigate different strategies, and work was planned efficiently to overcome this. Once the carboxylic acid **25** was obtained, it was decided to use standard amide coupling or to form an acid chloride via chloroenamine to give **26** in order to react this with the linker derivative **13** in an attempt to form product **27** *(Scheme 8)*. Chloroenamine was chosen to form the acid chloride **26** as it does not effect the Boc group nor result in Boc cleavage like other acids might do.[81]

Numerous attempts were made to react the iso-orotic acid chloride **26** with the linker **13**, however these resulted only in starting material being isolated. This is despite the fact that anhydrous conditions for the chloroenamine [81] step with iso-orotic acid **25** were employed and new product formation was indicated by TLC. Amide couplings with PyBOP, DIEA were envisaged to couple the iso-orotic derivative **25** with the linker **13** to give the trimer compound **27**. The PyBOP coupling synthesis failed, with only starting material being isolated. Other coupling reactions to form compound **27** in anhydrous conditions were generated: 1) EDCI/HOBt and TEA in DMF; 2) PyBOP/HBTU and DIEA in DMF and 3) DCC/LiCl in THF.



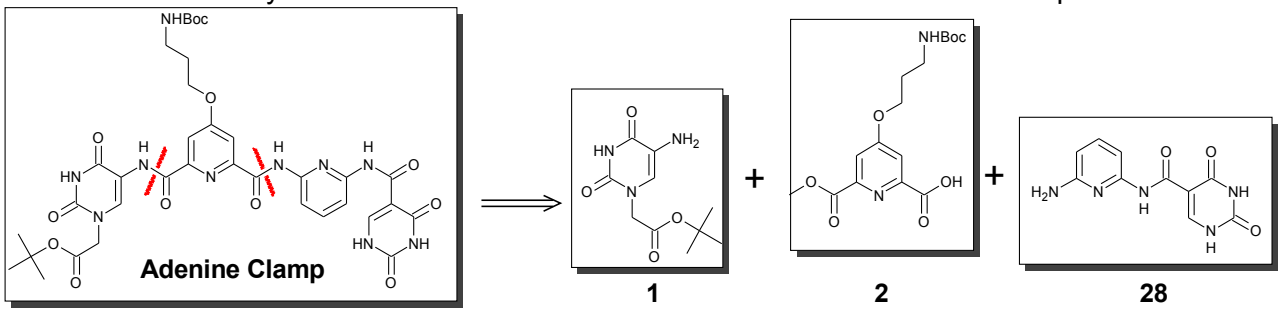
13 **25** **27**

Scheme 8: a) Coupling was unsuccessful using PyBOP, DIEA, DCM; or EDCI/HOBt and TEA in DMF; or PyBOP/HBTU and DIEA in DMF; or DCC/LiCl in THF; or by chloroenamine conditions.

The last conditions, containing lithium chloride salt, was hoped to disrupt any aggregation between the aromatic molecules and enhance coupling.[82] None of the mentioned couplings resulted in the desired product **27**. As a result it was not possible to continue the adenine clamp synthesis and carry out the final coupling between **27** and **1**. So far successful synthesis included making the precursors **1**, **2** and **13**.

4.1.2 Proposed route 2

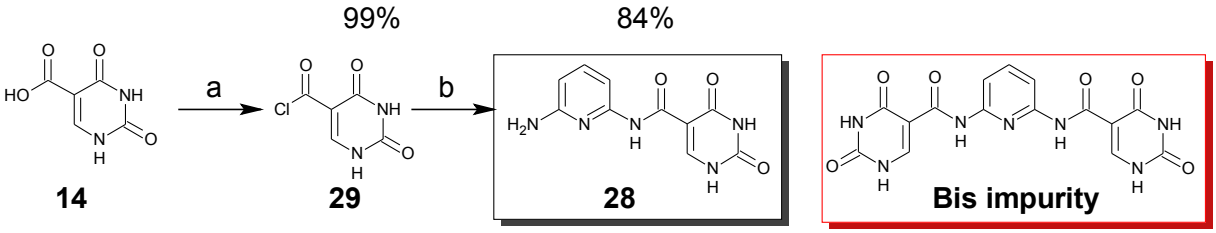
Due to the problems in forming the iso-orotic acid derivatives **4/25** and the coupling of these to the linker **13**, the synthetic route was revised to obtain the final adenine clamp *(Scheme 9)*. One major limitation with compound **25** was the choice of reaction conditions available to form a reactive acid chloride **26**. Previously these conditions were limited to chloroenamine in chlorinated solvents due to the presence of a Boc protecting group on compound **25**. If the acid labile protecting group was replaced or eliminated, it could be possible to investigate an alternative synthetic route. Other problems included obtaining the bis-substituted impurity **21** and ester selectivity during saponification of **16c**, as well as the poor solubility throughout each of the iso-orotic derivative steps. The new proposed route consisted of making the difficult amide first via the iso-orotic acid **14** and thus avoid the use of the Boc amino propyl side chain *(Scheme 10)*. Therefore the plan was to synthesise compound **28**, then couple this to **2** and finally to the 5-amino uracil derivative **1** to make the adenine clamp.



Scheme 9: Second retrosynthetic analysis of the **adenine clamp.** Three main building blocks: 5-amino uracil derivative **1**; pyridine carboxylic acid derivative **2**; amino pyridine iso-orotic derivative **28** to assemble and join together.

To form compound **28** the following synthesis was carried out. Commercially available 2,4-dihydroxypyrimidine-5-carboxylic acid **14** was reacted with thionyl chloride ($SOCl_2$), heating at 90 °C for three nights *(Scheme 10)*. A benefit of this, was that $SOCl_2$ could be used as the solvent for this reaction, thus avoiding the solubility issues which were encountered before. Since no Boc groups were present at this stage, the wider selection of compounds available to make an acid chloride was an advantage. These harsh conditions successfully formed the acid chloride **29** and thionyl chloride was removed *in vacuo* and 2,6-diaminopyridine **3** as a solution in DCM was added to **29** resulting in the iso-orotic pyridine derivative **28**. It is not possible to do this in the reverse order, due to the poor solubility of the iso-orotic derivative **29**.

It was important to ensure that the 2,6-diaminopyridine **3** was of good quality (not black in colour due to oxidation), and if necessary it was recrystallised. The resulting insoluble product 28 was isolated by filtration and analysed by NMR. This showed a new singlet at 10.9 ppm and a doublet at 6.2 ppm suggesting amide formation with an 84% yield and the compound was thus used for the next steps without further purification. Formation of **28** was further confirmed by mass spectra. Depending on the reaction conditions, the bis-substituted impurity could form.



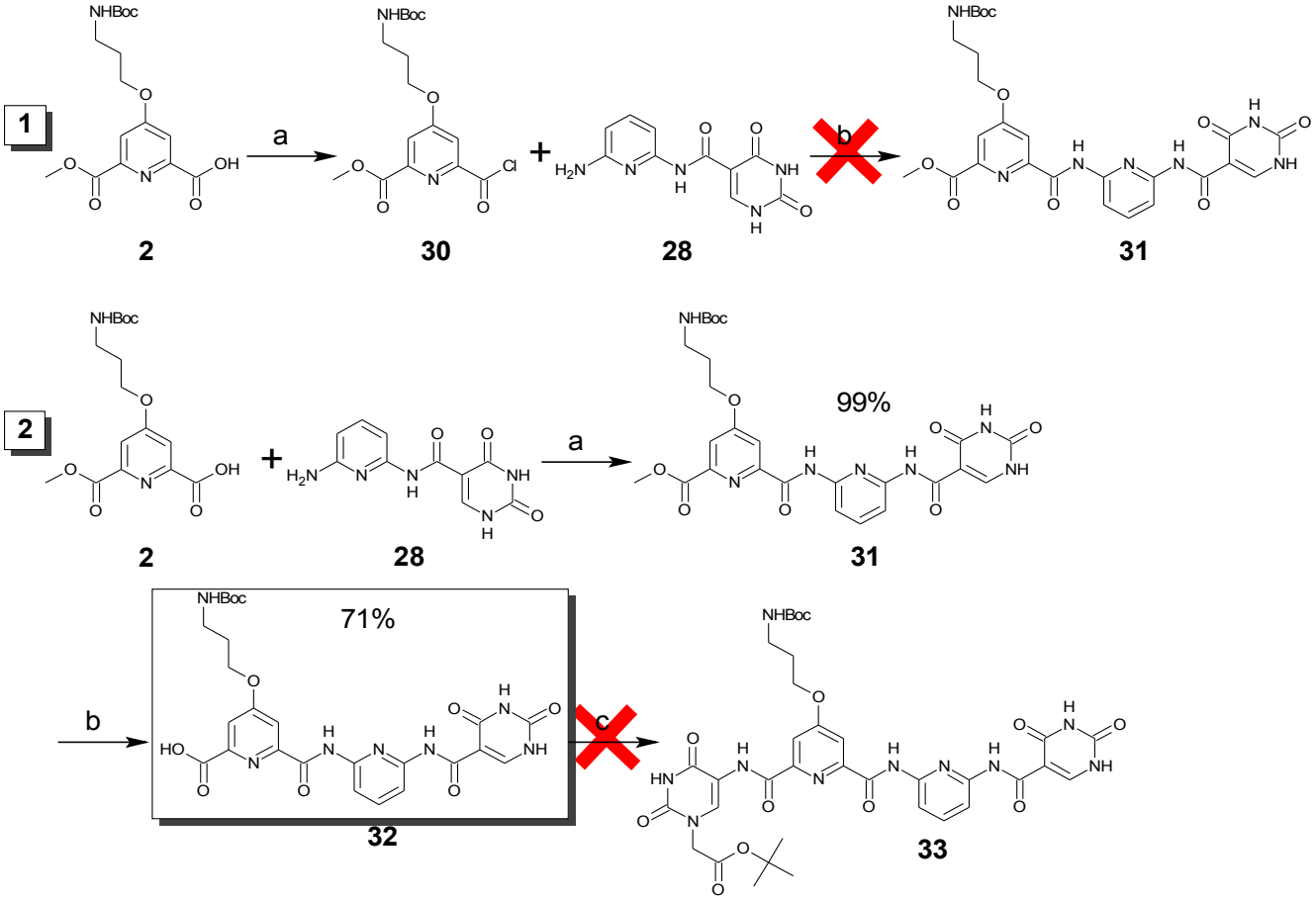
Scheme 10: Iso-orotic acid chloride derivative **29** coupled to 2,6-diaminopyridine **3** to give dimer product **28**. a) $SOCl_2$, 90 °C, 3 nights; b) 2,6-diaminopyridine **3**, DCM, TEA, RT.

Further reactions showed the product **28** also contained the bis-substituted iso-orotic pyridine derivative by NMR analysis. This occurred when 3.6 equivalents of 2,6-diaminopyridine **3** was used compared to a preferred 1.3 equivalent. This can be avoided by using the mono Boc protected amino pyridine derivative; however as previously mentioned, this would add a further two steps of protection and deprotection to the synthesis. Also due to solubility issues of compound **28**, column chromatography was not viable for purification. Instead, the crude mixture was dissolved in 10% $NaHCO_3$ and filtered to remove any unreacted 2,4-dihydroxypyrimidine-5-carboxylic acid **14** starting material. The solid recovered was dissolved in 1 M HCl. At this stage, the insoluble solid isolated from the acidic aqueous solution was shown to be the bis-substituted impurity by NMR. Therefore, the insoluble iso-orotic acid **14** and the bis-substituted compound had been successfully removed by work up conditions. The acidic aqueous solution was filtered off, was then neutralised with NaOH and resulted in a solid precipitating and this proved to be the pure desired product **28**. Any 2,6-diaminopyridine **3** still present was removed by trituration with diethyl ether and DCM. The pure product **28** was obtained after the acid/base aqueous work up, however this came at the expense of a poor yield of 38% compared to the 84% previously obtained.

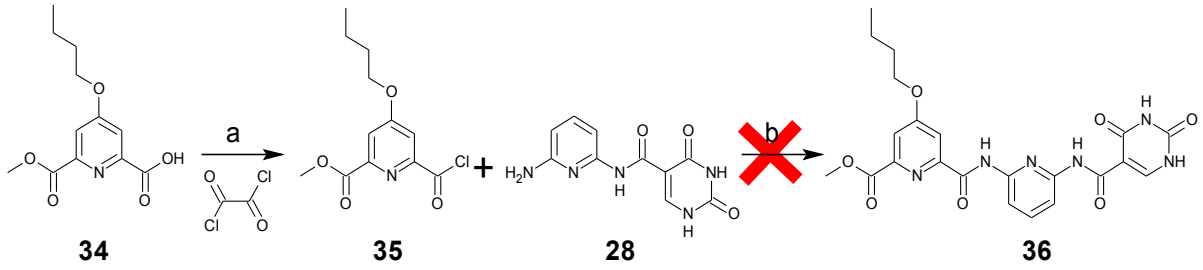
The pyridine mono acid derivative **2** was initially reacted with chloroenamine to form the acid chloride **30** and reacted with **28** and DIEA in anhydrous DCM; however this proved unsuccessful in forming **31** *(Scheme 11, Synthesis 1)*. This was mainly due to the poor solubility of **28** in DCM. Rather than lose the valuable compound **2**, it was decided to react **35** with **28** to test if amide coupling was possible *(Scheme 12)*. Since there were no acid sensitive groups (such as a Boc protecting group) on **34**, it was possible to react the carboxylic acid with oxalyl chloride to yield the acid chloride **35**.

However, upon reaction of **35** to **28**, the compound **36** was not observed by NMR analysis *(Scheme 12)*. It was thought that the lack of a solubilising chain may have hindered the reaction. Heat may have also been required. To alleviate these limiting factors optimised conditions such as increased temperatures and polar solvents were implemented.

Coupling reactions between **2** and **28** in the presence of standard coupling reagents were investigated in a variety of conditions, such as PyBOP, EDCI/HOBt, DCC/LiCl, with DIEA or TEA as base in anhydrous DMF or DMA as solvent. Despite using DMF or DMA, the starting materials were poorly soluble and at room temperature the product **31** was not formed. Coupling **2** and **28** in the presence of PyBOP, DIEA, pyridine, DMF at 60 ˚C produced the desired product **31** (99% yield) *(Scheme 11, Synthesis 2)*. The addition of anhydrous pyridine and heating the reaction resulted in increased the solubility of the reaction mixture. Product **31** precipitated from the DMF solution upon cooling and was was easily filtered from the reaction mixture. Drying reagents such as $MgSO_4$ and purification by column chromatography were avoided, due to poor solubility.



Scheme 11: Synthesis 1. a) Initial formation of the acid chloride **30** using chloroenamine in DCM b) Reacting acid chloride **30** with **28**, DIEA in DCM was unsuccessful in obtaining **31**; **Synthesis 2.** a) The carboxylic acid **2** proved successful in coupling with **28** to give **31** (99% yield) in the following conditions: PyBOP, DIEA, pyridine, DMF, 60 °C; b) NaOH, DMSO/MeOH gave **32** (71% yield); c) Unsuccessful coupling between **1** and **32**: PyBOP, DIEA, DMF.

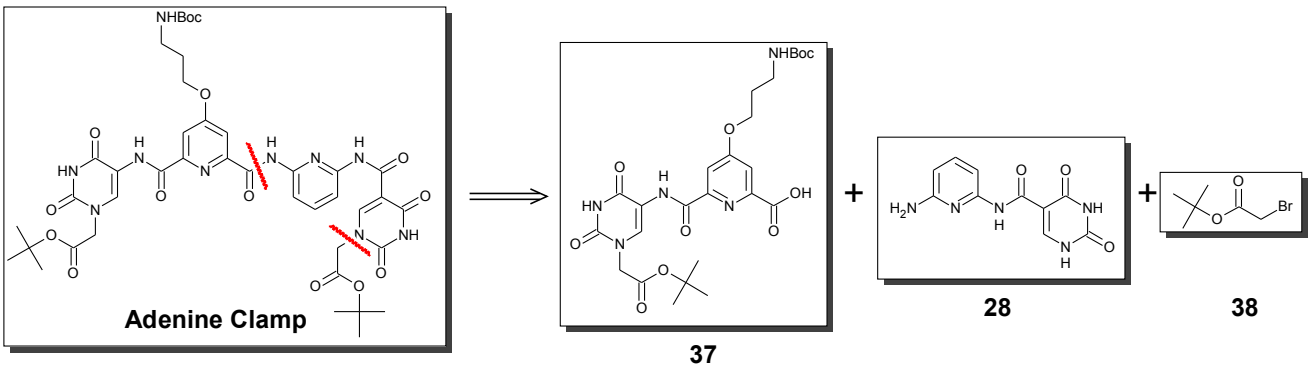


Scheme 12: Testing reaction conditions to form product 36: a) Oxalyl chloride, DCM; b) Reacting 35 and 28 proved unsuccessful in forming compound 36 in the following conditions: TEA, pyridine, DCM, RT.

Saponification of **31** was carried out in NaOH, DMSO/MeOH to form **32** (71% yield) and coupled to 5-amino uracil **1**. However this coupling did not work and compound **33** was not obtained *(Scheme 11, Synthesis 2)*. This reaction may need high temperature conditions; in addition perhaps introduction of the solubilising chain onto the iso-orotic pyridine derivative **28** might overcome the solubility issues. Nevertheless, this route successfully provided the precursors **28** and **32**. The ability to form the amide bond and obtain **28** was a critical turning point in being able to make progress in the remainder of the adenine clamp synthesis. This route also reduces the number of synthetic steps previously required to make the iso-orotic derivative **25**. The dimer **28** could be formed from the relatively cheap commercially available compounds 2,4-dihydroxypyrimidine-5-carboxylic acid **14** (via **29**) and 2,6-diaminopyridine **3**. In addition, the need for ester selectivity or a saponification step to obtain **28** was eliminated. The ability to form **32** was also considerably useful.

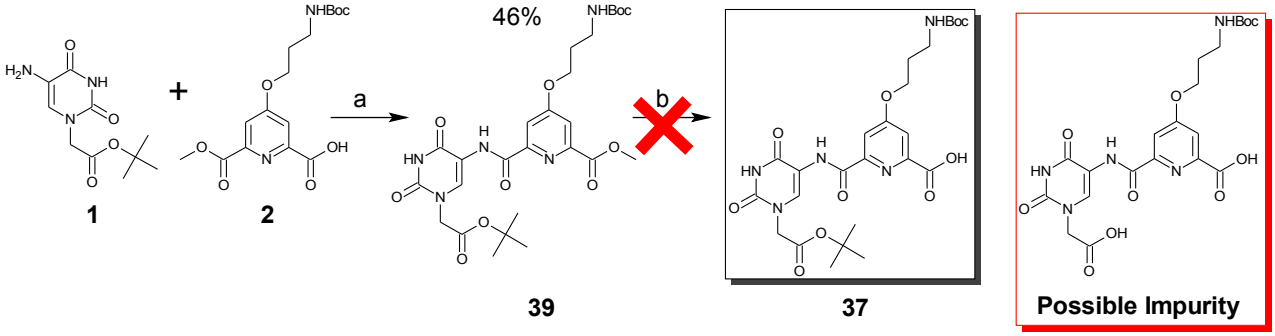
4.1.3 Proposed route 3

Despite previous failures, a new synthetic approach was implemented based on the knowledge gathered from the previous two synthetic routes. The logical progression towards the adenine clamp was to reintroduce a solubilising side chain **38** *(Scheme 13)*. A convergent strategy was selected rather than a linear one. Thus it was envisaged to add a *t*-butyl bromoacetate side chain **38** to the iso-orotic pyridine derivative **28** and couple this to compound **37** forming the adenine receptor *(Scheme 13)*.

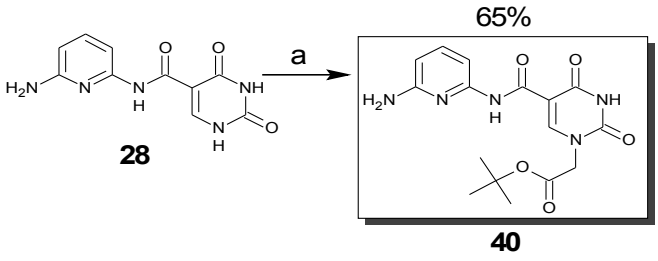


Scheme 13: Third retrosynthetic analysis of the adenine clamp. Building blocks consisting of the uracil pyridine carboxylic acid derivative **37** and amino pyridine iso-orotic derivative **28** and *t*-butyl bromoacetate side chain **38**.

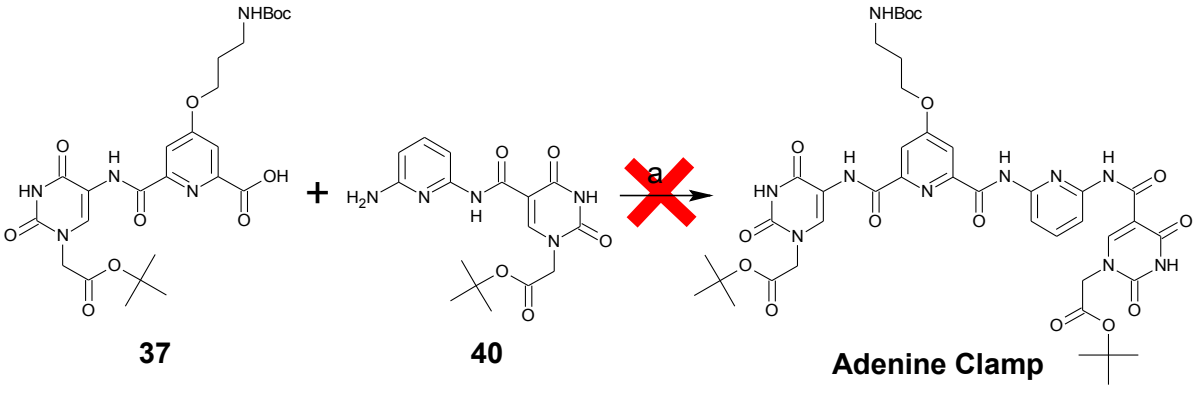
Compound **39** was obtained in 46% yield, without any issues by coupling **1** and **2** together in PyBOP conditions *(Scheme 14)*. Saponification of **39** was expected to give the mono carboxylic acid **37** however, it was thought that the *t*-butyl group might have been cleaved resulting in a bis-substituted carboxylic acid version of **37** as confirmed by NMR analysis *(Scheme 14)*. The third synthetic route did successfully result in compound **40** (65% yield) being obtained in minimal steps despite the other challenges. To increase the solubility of **28**, a *t*-butyl acetate side chain was added to form **40** *(Scheme 15)*. Even though there were several nucleophilic sites on compound **28** available to react with *t*-butyl bromoacetate, compound **40** was obtained in good yield. The number of synthetic steps required to produce **40** was therefore reduced, since there was no apparent requirement for protecting groups. Since **37** was not successfully obtained, coupling this to **40** was not possible in obtaining the final adenine clamp *(Scheme 16).* Thus a new synthetic route strategy had to be devised.



Scheme 14: a) Coupling of 5-amino uracil **1** with pyridine carboxylic acid derivative **2** to form the uracil pyridine ester derivative **39** (46% yield) in the following conditions: PyBOP, DIEA, DMF, RT; b) Saponification of **39** did not result in **37** and it is suspected the bis-substituted carboxylic acid impurity was obtained in the following conditions: NaOH, MeOH/1,4-dioxane/H_2O.



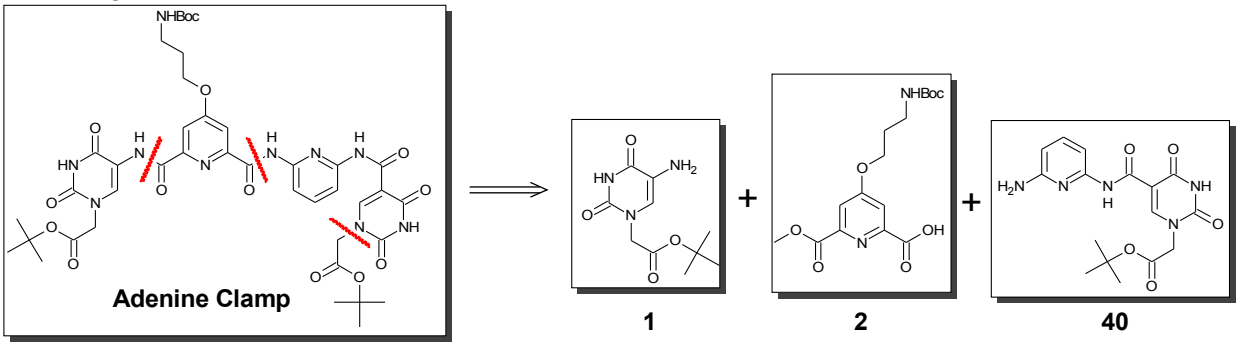
Scheme 15: Side chain addition: a) *t*-butyl bromoacetate, TEA, DMF, RT to give **40** (65% yield).



Scheme 16: Compound **37** was not obtained and therefore could not be coupled to **40** to form the adenine clamp.

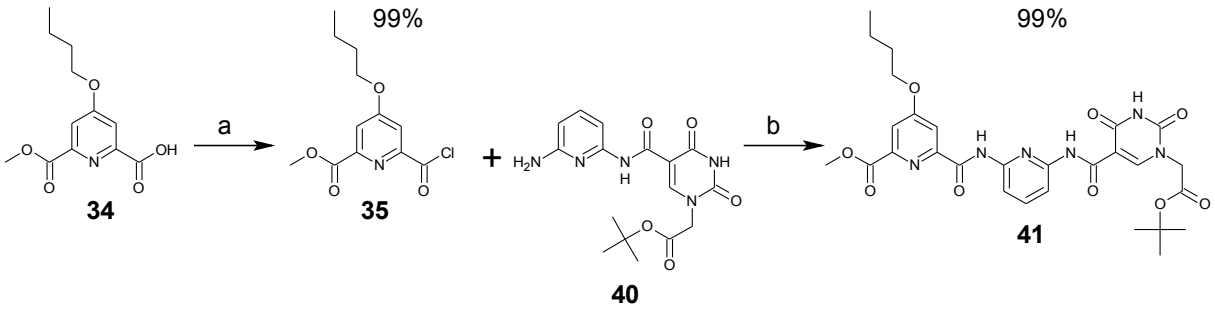
4.1.4 Proposed route 4

The fourth synthetic strategy was an opportunity to concentrate on forming an amide bond between **2** and **40** in an attempt to obtain **42** *(Scheme 17 and 19)*. This would be followed by coupling of **1** and **42** to obtain the desired adenine clamp.



Scheme 17: The fourth retrosynthetic analysis of the adenine clamp: Building blocks consisting of the 5-amino uracil derivative **1**; pyridine carboxylic acid derivative **2**; amino pyridine iso-orotic side chain derivative **40** to be assembled and joined together.

Initial attempts at the coupling reaction between **2** and **40** did not go to completion and mainly starting amine **40** (35%) and carboxylic acid **2** (44%) were recovered, with just 16% yield of the amide product **42** recovered on a 550 mg scale. The reaction was also monitored by TLC and analytical HPLC to investigate the synthesis. To further investigate the reactivity of amine **40**, it was decided to synthesise the acid chloride **35** of the pyridine acid derivative **34** with no acid sensitive protecting groups present *(Scheme 18)*. On a 250 mg scale, the carboxylic acid **34** and oxalyl chloride in anhydrous DCM were reacted to give the acid chloride **35** which was subsequently reacted with amine **40** in anhydrous pyrdine, TEA, DCM under N_2. The addition of pyridine greatly increased the reaction mixture solubility and the product **41** was successfully obtained in quantitative yield. It was decided to add anhydrous pyridine when coupling **2** to **40** in order to overcome the previous solubility issue.

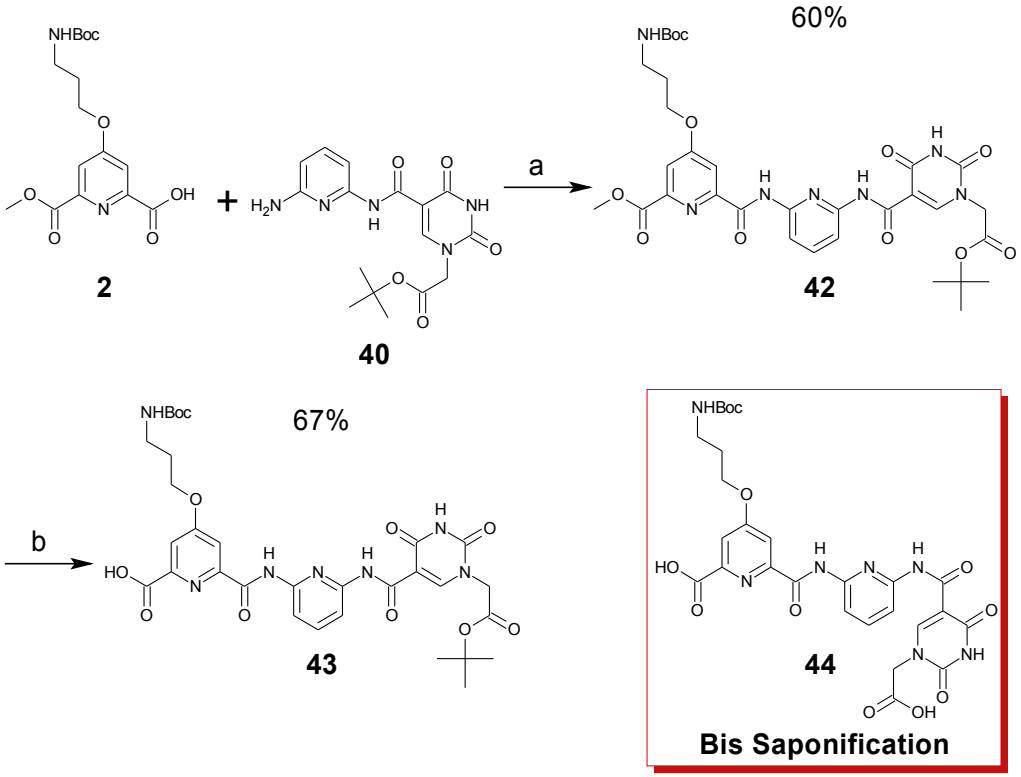


Scheme 18: Testing the feasibility to form the amide: a) Oxalyl chloride, DCM to give **35** (99% yield); b) Reacting **35** and **40** in pyridine, TEA, DCM to give **41** (99% yield).

However, during saponification of **41**, the *t*-butyl group was suspected to have been removed resulting in **44**, as it was difficult to solubilise the compound in a variety of deuterated solvents for NMR analysis *(Scheme 19)*. This was the case with compound **37** *(Scheme 14)* in which the *t*-butyl group was cleaved during saponification.

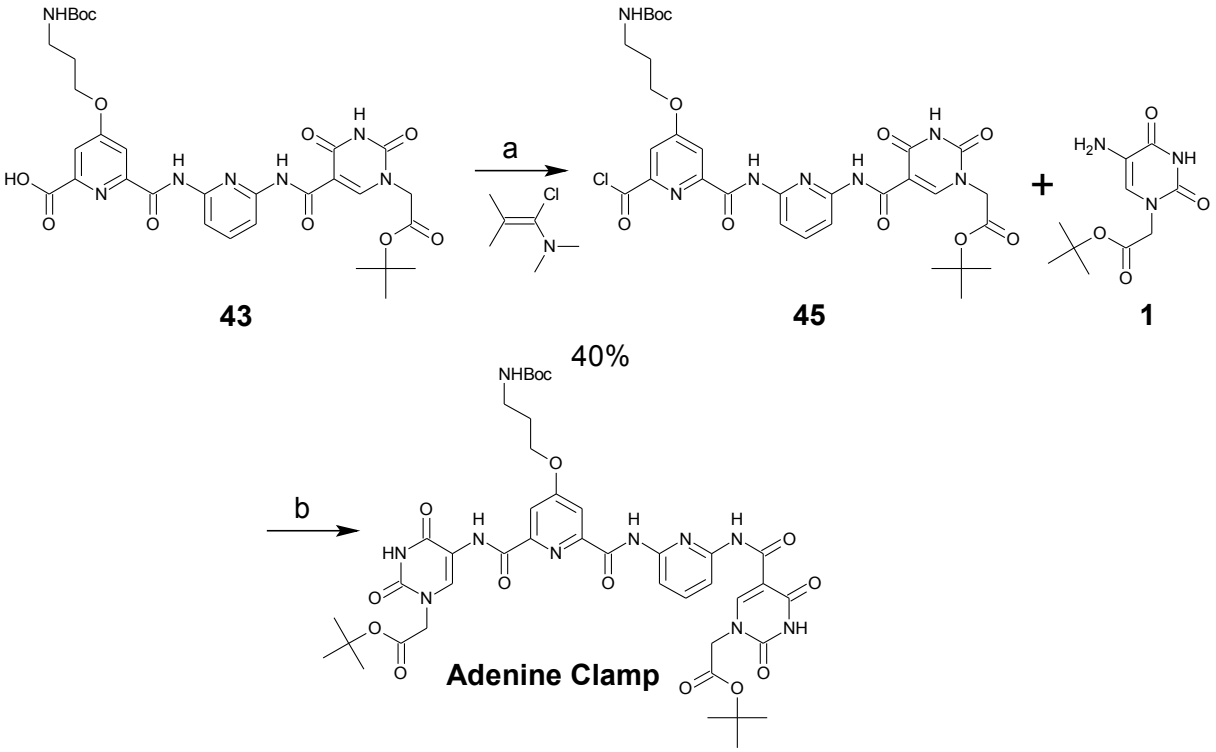
PyBOP coupling between **2** and **40** under argon in anhydrous DMF/TEA/pyridine and heated to 60 °C overnight gave product **42**, which was isolated in 60% yield *(Scheme 19)*. Saponification of **42** also provided a problem, as the *t*-butyl group on the solubilising side chain of **42** may have been removed during saponification, thus resulting in compound **44** with two carboxylic acids *(Scheme 19)*.

A variety of saponification conditions were investigated. Both sodium hydroxide and lithium hydroxide had the ability to cleave the *t*-butyl group and resulted in compound **44**. Lithium iodide refluxed in 1,4-dioxane resulted in the starting ester **42** and did not result in a carboxylic acid. Potassium hydroxide (3 equiv) in DMSO/H_2O was reacted with **42** and consequently **43** was obtained (67% yield) in these conditions, without cleavage of the *t*-butyl group. It was necessary to remove the DMSO/H_2O by lyophilisation, and the subsequent difficult removal of DMSO may have contributed to the low yield. This step has since been optimised by using a trimethylsilyl ethoxy carbonyl group on **2** which is more readily cleaved to that of the *t*-butyl group.[83]



Scheme 19: a) PyBOP, TEA, pyridine, DMF, 60 °C, overnight resulted in **42** (60% yield); b) KOH, H_2O/DMSO, 60 °C, 1 hour resulted in **43** (67% yield).

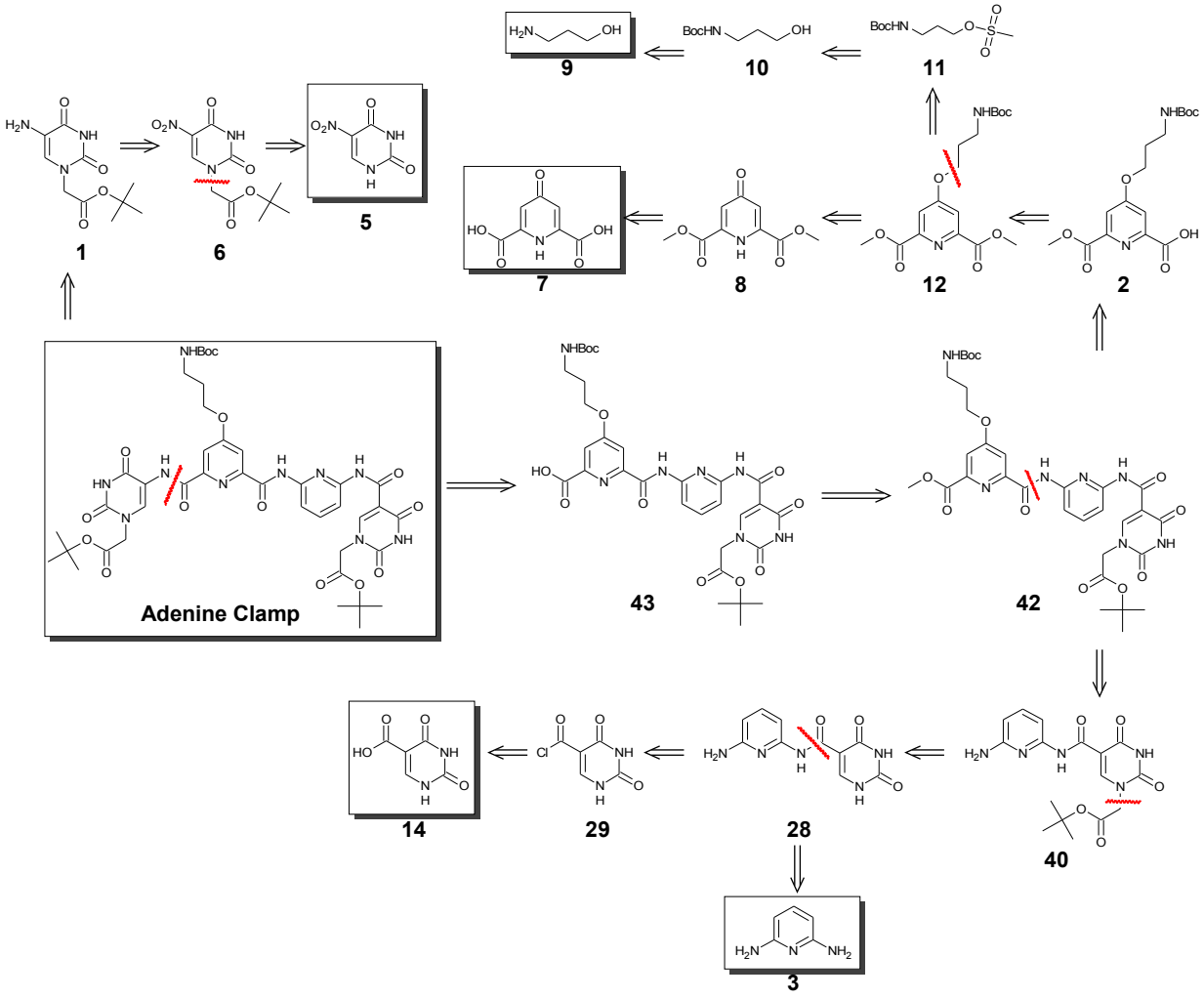
The next step using chloroenamine to form the acid chloride **45** proved unsuccessful *(Scheme 20)*. Instead PyBOP coupling between **43** and **1** gave the desired adenine clamp. To date the yield of the final coupling step is low (40% yield) and there was not enough material to investigate and optimise this step at the time. The adenine clamp obtained was used for ^{1}H NMR titration binding studies in chloroform with 9-ethyl adenine and for various conditions to grow crystals for X-ray crystallography analysis.



Scheme 20: a) Initial reaction of **43** with chloroenamine in DCM proved unsuccessful in the following step between **45** and **1** and did not result in the adenine clamp. b) A successful coupling between **43** and **1** resulted in the adenine clamp (40% yield) in the following conditions: PyBOP, TEA, pyridine, DMF at RT.

4.1.5 Synthesis summary

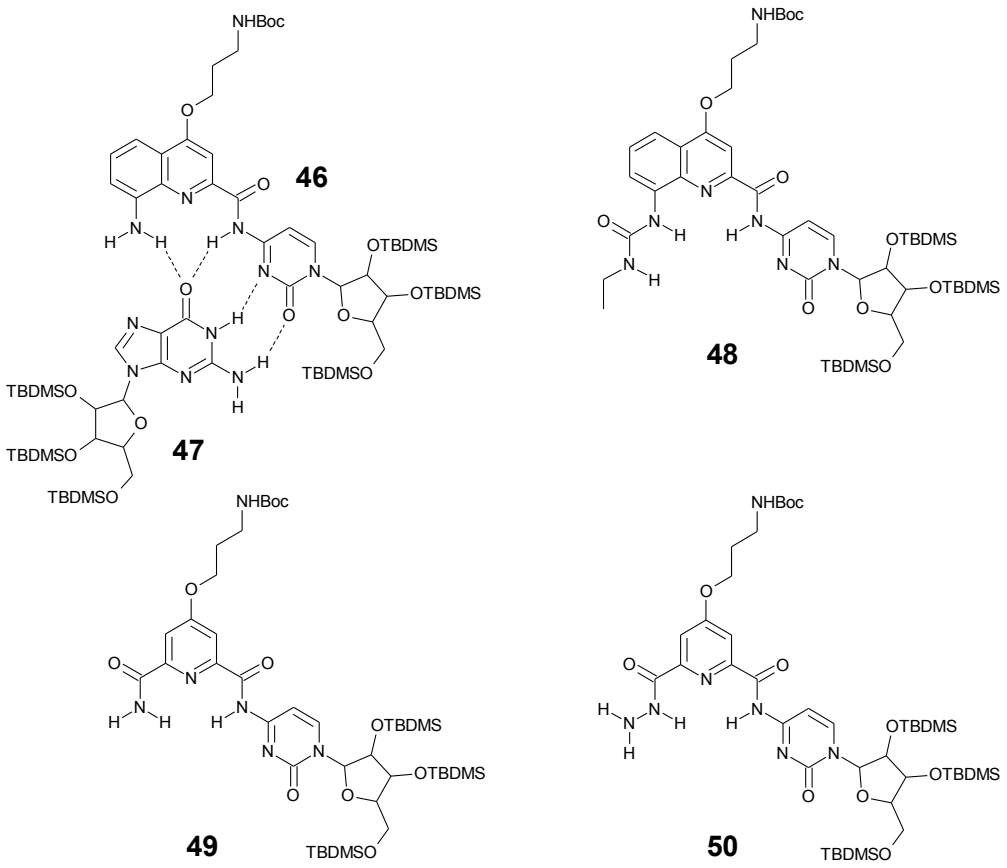
The initial adenine clamp synthesis consisted of coupling **2** and **3** to form **13**. This was followed by coupling to **4** to give **27** and coupling of this to **1**; however this approach was not as trivial as initially thought. In summary, the main challenges encountered throughout the adenine clamp synthesis consisted of the poor solubility of the iso-orotic derivatives **4** and **25**, and the poor ester selectivity during saponification. To overcome these issues it was thought to form the difficult amide coupling in the first step, i.e. coupling between **14** and **3** to give **28**. The next challenge was to form the second amide bond. It was possible to couple **28** with **2**, using PyBOP in DIEA, DMF and pyridine at 60 °C overnight to give **31**. However another setback came during the final step of coupling **1** with **32**. Again this could be attributed to poor solubility. For this reason it was decided to reintroduce a solubilising side chain. After some investigations it was possible to obtain compound **40**, despite the possibility to alkylate at several places. The second amide bond was made based on the previous coupling conditions of PyBOP in TEA, DMF and pyridine at 60 °C overnight to give compound **42**. A final complication involved ester selectivity during saponification of **42**. potassium hydroxide in DMSO/H_2O gave the desired carboxylic acid **43**. The third and final amide bond was made by coupling **1** and **43** to give the adenine clamp *(Scheme 21)*.



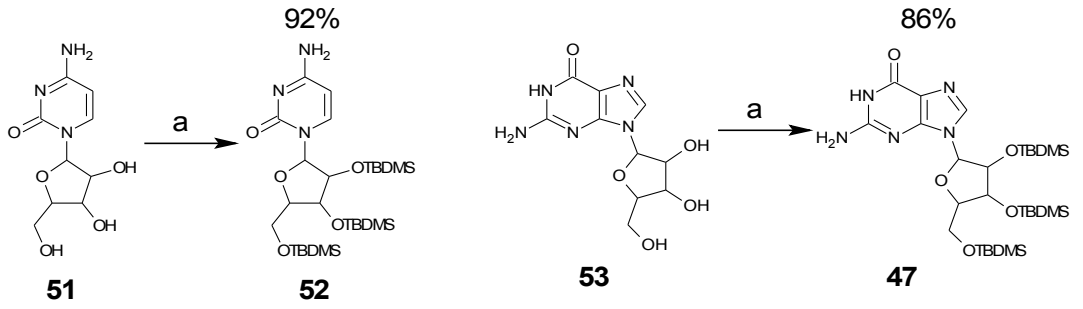
Scheme 21: A retrosynthetic overview of the final route used to obtain the adenine clamp

4.2 G Clamps

Guanine receptors similar to those published by Zimmerman's lab were synthesised for comparative reasons.[38] The receptors synthesised, contained a solubilising side chain which enabled studies in both organic and aqueous solvents. Thus in an organic solvent it was possible to compare the binding constants to those published by Zimmerman's group.[38] Synthesis to produce compounds quinoline **46, 48** and pyridine **49, 50** G clamp derivatives was investigated *(Scheme 22)*. Some design features of the quinoline G clamp motifs such as π-π stacking was expected to stabilise binding to a template during aqueous studies. The smaller aromatic surface area of the pyridine G clamp series is expected to be less stable than those of the larger quinoline G clamp series. Urea or hydrazide extension onto the free amino group was desired to utilise specific Hoogsteen bonding due to the $NHNH_2$ donor group. Protection of the cytidine **51** with a TBDMS group gave **52** *(Scheme 23).* The purpose was to increase solubility in organic solvents. The cytidine derivative could be deprotected later on to provide a hydrophilic compound for aqueous studies. The guanosine **53** was also protected with TBDMS protecting groups to give compound **47**.[84]



Scheme 22: The various G clamps **46**, **48**, **49**, **50**, and guanosine **47** synthesis feasibility was investigated.

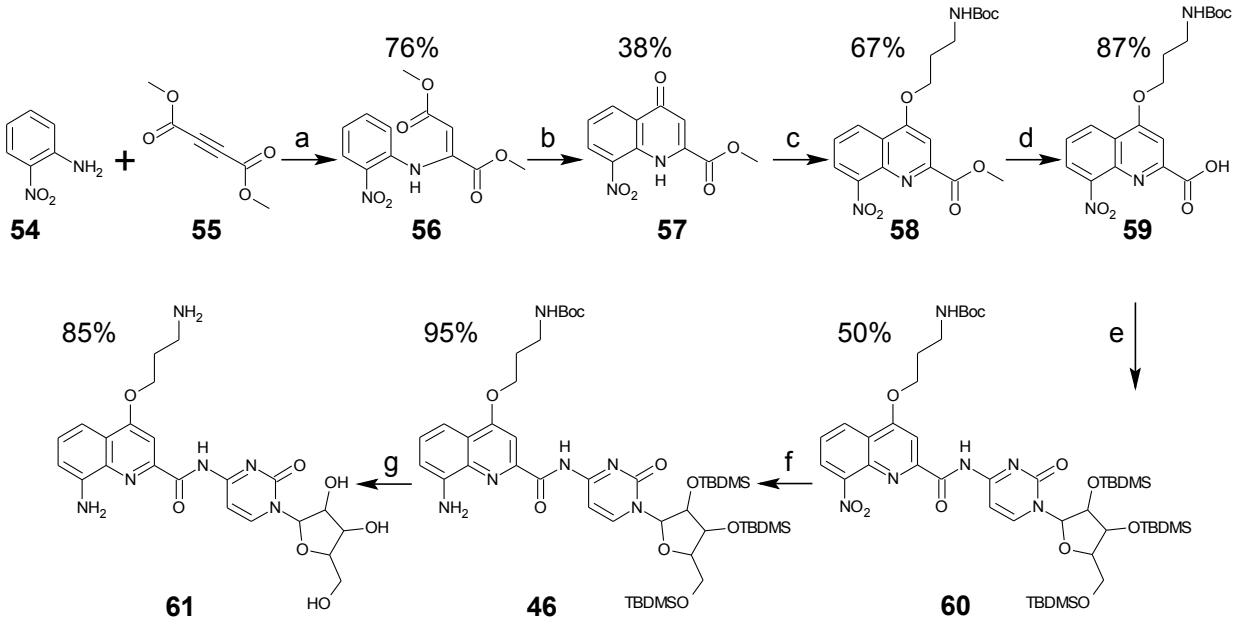


Scheme 23: Cytidine **51** and guanosine **53** TBDMS protection to give **52** (98% yield) and **47** (86% yield) respectively. a) TBDMS-Cl, imidazole, THF, RT, 2 nights.

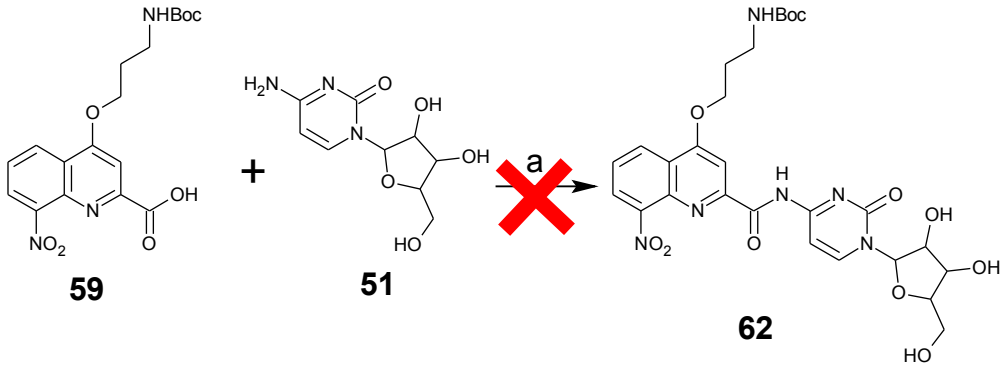
A standard Huc group protocol,[85] was used for the quinoline monomer synthesis **57** with addition of the 1,3-Boc aminopropanol side chain **10** via Mitsunobu conditions to give compound **58** *(Scheme 24)*. The saponification of **58** was a particular difficulty, since on each attempt a gel was formed, this is thought to be due to the presence of the potassium salt, since the literature reports the formation of nanogels with potassium salts.[86] Saponification conditions reported by Gillies involved reacting compound **58** with KOH in THF/MeOH to give 96% of product **59**.[85] Compound **60** was obtained by PyBOP coupling, between the quinoline derivative **59** and TBDMS cytidine derivative **52** in DMF. Hydrogenation of **60** with 10% Pd/C, H_2 in MeOH/EtOAc gave the amino derivative **46** without reducing the cytidine double bonds. The quinoline G clamp **46** was obtained and used to test its ability to bind with guanosine **47** in $CDCl_3$/5% DMSO-d_6 via NMR titration studies.

To obtain the water soluble compound, deprotection of **46** with TBAF initially proved troublesome as not all of the TBDMS groups were cleaved. It was planned, to couple the unprotected cytidine **51** with the quinoline acid monomer **59**. However due to poor solubility of cytidine **51**, this route proved a non-effective alternative in forming **62** *(Scheme 25)*. The TBAF remained in both aqueous and organic layers. Literature showed ways to remove the TBAF salt by use of an Amberlyst resin and HCl.[85–87] As an alternative straightforward method, 15% HCl in methanol deprotected the TBDMS and Boc groups in one step.

The pure deprotected quinoline G clamp **61** was obtained after purification by preparative HPLC. Analytical HPLC was carried out to observe the purity of **61**. Unfortunately, **61** decomposed and many impurities were observed by analytical HPLC. Possible oxidation of the free amine is thought to have occurred.

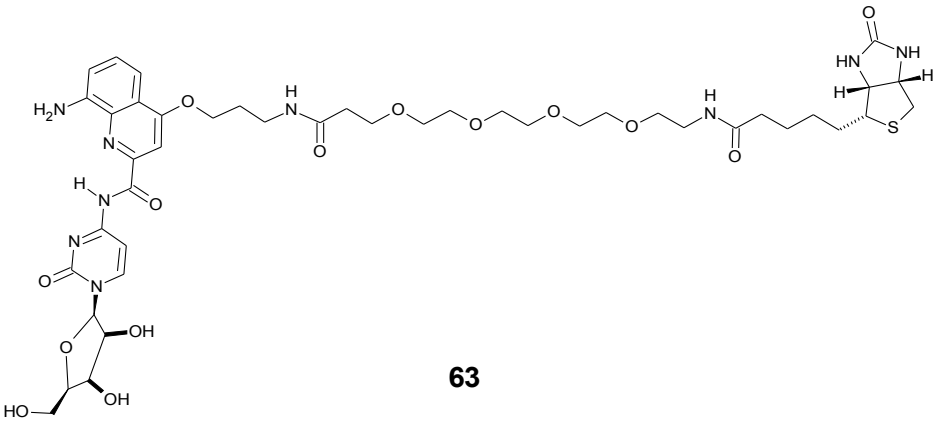


Scheme 24: Quinoline monomer **57** synthesis to give the G clamps **46** and **61** using the following conditions: a) MeOH, Reflux; b) PPA, 90 °C; c) 1,3-Boc aminopropanol **10**, PPh_3, DIAD, THF, 0°C; d) NaOH, MeOH/THF, 0°C; e) TBDMS cytidine **52**, PyBOP, TEA, DMF, RT; f) Pd/C 10%, H_2, MeOH/EtOAc; g) 15% HCl, MeOH.



Scheme 25: Alternative coupling which was unsuccessful due to solubility issues. a) PyBOP, DIEA, DMF, RT.

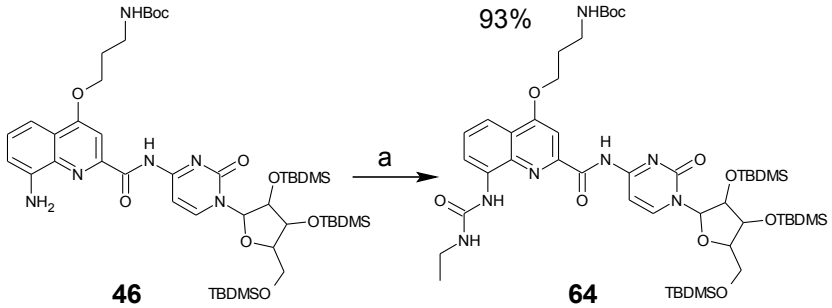
More material was synthesised to obtain **61**, which was subsequently coupled to Biotin to form **63** *(Scheme 26)*.[83] The purpose was to attach the quinoline G clamp **63** onto the surface of a SPR chip via streptavidin.



63

Scheme 26: Biotinylated G clamp **63**.

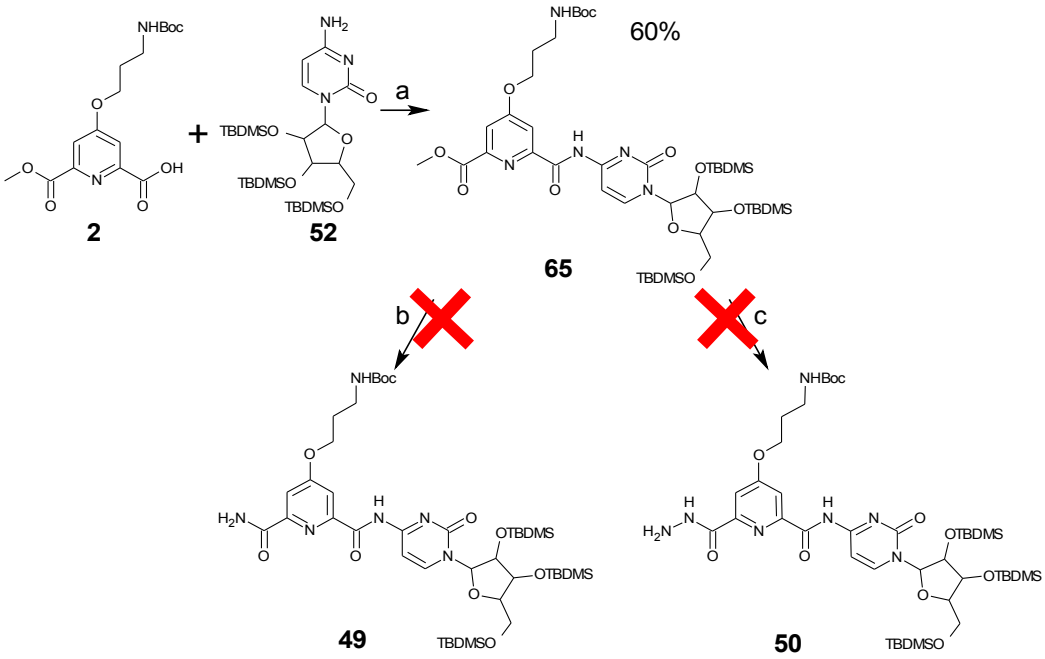
The amino quinoline G clamp **46** could easily be modified to incorporate Hoogsteen style hydrogen bonding and increase the binding to guanosine, by addition of a urea group *(Scheme 27)*. Compound **46** was reacted with ethyl isocyanate in THF resulting in quantitative formation **64** after column chromatography. It was expected that **64** would provide an enhanced binding to guanosine due to the utilisation of Hoogsteen bonding between the urea group and the N-7 of guanosine.



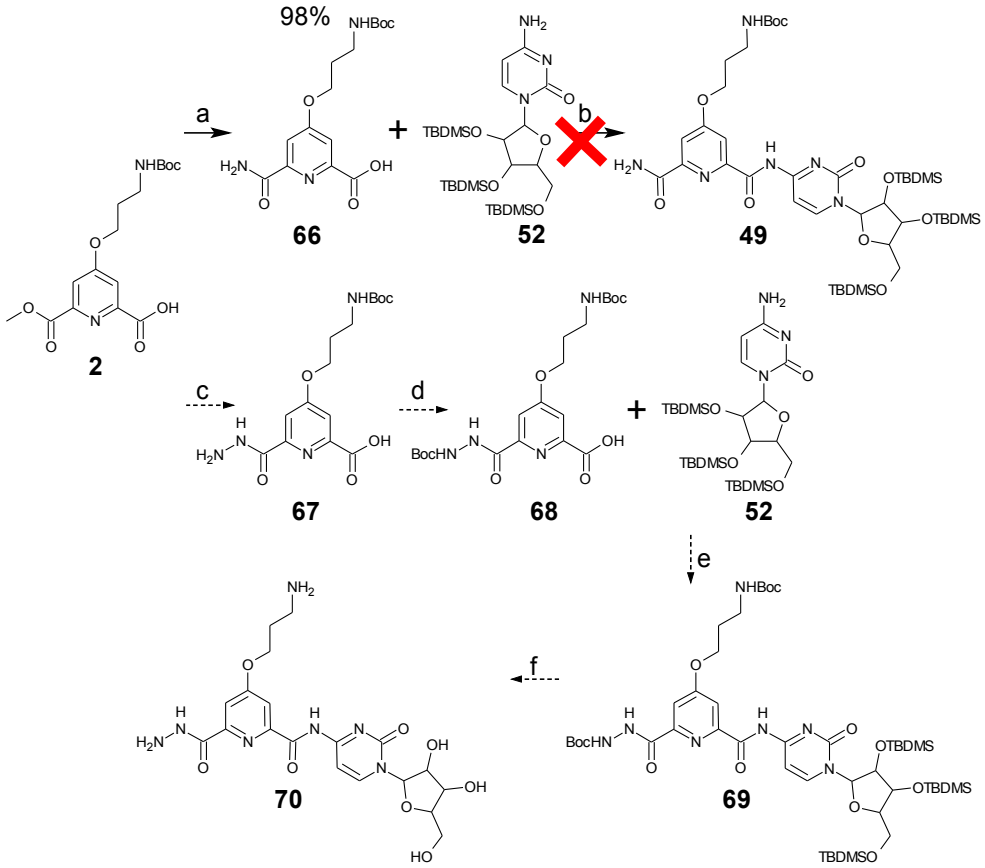
Scheme 27: Synthesis of urea derived quinoline G clamp **64** (93% yield): (a) Ethyl isocyanate, THF, RT.

The pyridine carboxylic acid monomer **2** previously used in the A clamp synthesis was coupled with TBDMS protected cytidine **52** via PyBOP coupling to obtain **65** in 60% yield *(Scheme 28)*. The pyridine ester cytidine derivative **65** and ammonium hydroxide (1 equiv) in MeOH/1,4-dioxane for ten minutes, was used for aminolysis in an attempt to obtain the carboxamide derivative **49**. However amide cleavage occurred giving starting materials **2** and **52**. Thus ammonia gas was bubbled through methanol for 15 minutes in another attempt of nucleophilic acyl substitution but this also resulted in amide cleavage of **65**. The reaction with the pyridine ester derivative **65** and hydrazine hydrate (1 equiv) in solution also resulted in amide bond cleavage with the cytidine and did not give the desired pyridine carbohydrazide derivative **50**. A new strategy reacting ammonia gas with pyridine derivative **2** formed **66** in quantitative yield (Scheme 29). Compound **66** was then coupled to the TBDMS cytidine derivative **52** in an attempt to form **49** as an alternative strategy to avoid amide cleavage. However **49** was not obtained on the initial attempt. Synthesis to **70** was not carried out and could be investigated in the future. The compound **67** would need to be Boc protected to give **68** and then coupled to cytidine **52** to give **69** followed by compound **70**.

Due to time restrictions, it was thought to concentrate on the quinoline G clamp synthesis and binding studies, as the larger aromaticity is advantageous for DCC studies.



Scheme 28: Initial synthesis for the pyridine G clamp derivatives **49** and **50**. a) PyBOP, DIEA, DMAP, DMF gave **65** (60% yield); b) 1st attempt: 1 equiv ammonium hydroxide, MeOH/1,4-dioxane; 2nd attempt: ammonia gas, MeOH; c) 1 equiv 1 M solution of hydrazine hydrate in THF.



Scheme 29: An alternative pyridine G clamp strategy to avoid amide bond cleavage: a) Ammonia gas, MeOH; b) PyBOP, DIEA, DMF, RT proved unsuccessful in obtaining **49**. Proposed synthesis: c) hydrazine hydrate, THF; d) Boc_2O, THF; e) PyBOP, DIEA, DMF, RT; f) 15% HCl, MeOH.

5 Analysis and Discussion

The objective this project, was to synthesise and analyse novel nucleobase receptors. It was desired to monitor the interaction and affinity of these receptors with their corresponding nucleobases, followed by investigations upon complexation with a single stranded DNA (ssDNA) template. Some questions which were to be addressed include:

1) Can the clamps strongly bind and complex with a single corresponding nucleobase molecule? ^{1}H NMR titration studies in $CDCl_3$/5% DMSO-d_6 was carried out to investigate the complexation between the receptor and nucleobase. Upon binding, non-covalent interactions between the host and guest molecule via simultaneous Watson-Crick and Hoogsteen hydrogen bonding was expected. This effect induced a change in chemical shift which was observed in the NMR spectra. From this, it was possible to plot a graph and calculate a binding constant to reveal the strength and affinity of the complex. The A clamp receptor has two possibilities for binding with adenine and thus varied temperature experiments were used to investigate if there was a preference. Attempts to obtain crystals of the receptor and nucleobase to observe the solid state complexation via X-ray crystallography will be discussed. An experiment via surface plasmon resonance (SPR) targeting an abasic site was also envisaged to test the ability for complexation in water. Such a competitive media for hydrogen bonding was expected to be compensated by π-π stacking and a cooperative effect.

2) Can more than one clamp self-assemble and form a duplex with a poly A and poly G ssDNA in water? This experiment was designed to measure the efficiency of binding and self-assembly. Evaluation was via Circular dichroism (CD) and UV melting temperature studies. Stability studies in aqueous buffer and increasing temperature was also to be addressed.

3) Investigation into the ability of the A clamp to distinguish between a guanine mismatch intercalated on a poly A ssDNA. Would the mismatch enhance stability or disrupt the system?

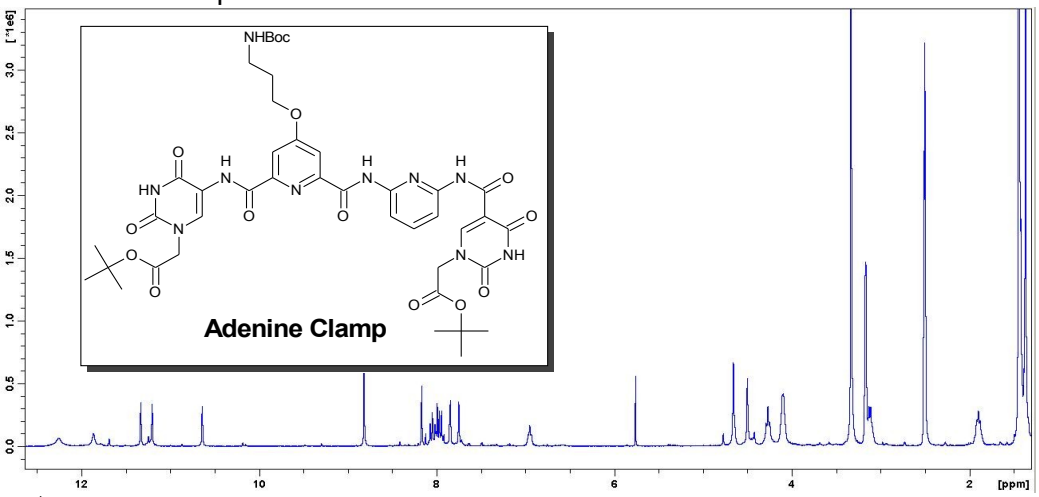
4) What would happen to a ssDNA containing both adenines and guanines when a mixture of clamps is introduced? This was designed to test the receptors specificity and recognition ability to discriminate between bases.

5) Would it be possible to introduce the A clamp onto a DNA template and induce transfer of information for self replication without the need for enzymes?

The above questions 3-5 were evaluated via circular dichroism (CD). A discussion on the various experiments carried out will be described in the following section (Section 5.3), with regards to these set goals. Synthesis completion of the adenine and guanine clamp receptors **46**, **48** gave the required products to carry out binding studies to examine the receptors ability to specifically bind and to determine how efficiently they could bind towards their respective nucleobases in organic and aqueous media.

5.1 A Clamp Characterisation

The adenine clamp (A clamp) was successfully synthesised. To confirm that the final A clamp was produced, ^{1}H NMR *(Spectra 1)*, ^{13}C NMR and HRMS analysis were obtained to characterise the A clamp.



Spectra 1: ^{1}H NMR organic soluble A clamp in DMSO-d_6

The A clamp was chosen from a design pool of several others as it contained a central linker motif **13** similar to motifs previously synthesised within the Huc group. Features such as two uracils for recognition, a planar foldamer crescent and solubilising side chains were discussed in detail in the design section (Section 3). An X-ray crystal structure was obtained for the trimer **42** in DMSO solvent *(Fig. 25)*. This crystal structure confirms that the *t*-butyl acetate side chain is attached at the N-1 position of the uracil derivative. A co-crystal of the A clamp and 9-ethyl adenine has yet to be obtained.

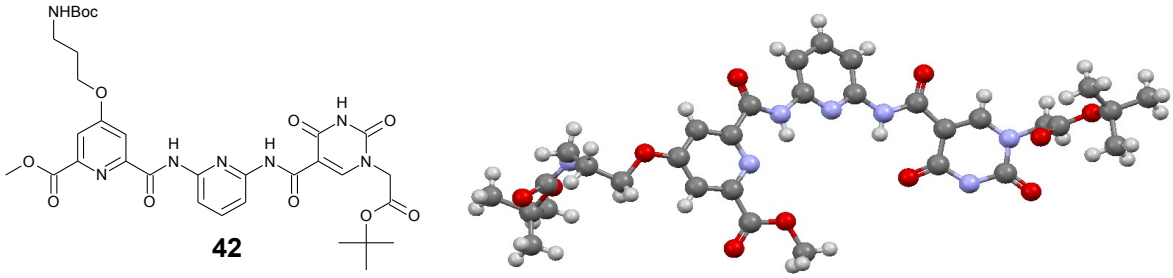


Figure 25: (Left) Structure of **42** (Right) X-ray crystal structure of trimer **42**.

5.2 G Clamp Characterisation

To confirm that the final quinoline G clamps **46** and **48** were produced, ^{1}H, ^{13}C NMR and HRMS analysis were obtained to characterise the G clamps. The design of the G clamp receptor was based on those reported by Zimmerman with an extra solubilising side chain attached. Synthesis of both the quinoline **59** and pyridine **2** monomers was similar to those previously synthesised within our group.

5.3 Complexation Studies in Solution

5.3.1 Clamp complexation to corresponding nucleobase in organic solvent

5.3.1.1 ^{1}H NMR titration: A clamp and 9-ethyl adenine.

The initial step was to evaluate the A clamp binding affinity towards 9-ethyl adenine in organic media $CDCl_3$ via NMR titration studies. Purity is an important factor in NMR binding studies and the A clamp was purified by column chromatography three times in $CHCl_3$/MeOH 99:1 to ensure any impurities were removed. Although a substantial amount of product was lost on the column in these conditions due to its poor solubility. The A clamp was obtained pure and was utilised for NMR titration studies. Previous work by Zimmerman [90] and Rebek [91] was used as a guideline for setting up the titration studies and co-crystallisation experiments respectively.

General procedure for ^{1}H NMR binding studies: $CDCl_3$ was stored over 4 Å molecular sieves and passed through a column of activated (flame dried) alumina (basic) plug prior to use, to remove any acidic residues. NMR tubes and vials were dried in a 110 °C oven and cooled in a desiccator prior to sample preparation. The nucleobase guest and the receptors were weighed on a precision milligram calibrated balance. The empty vial was weighed before and after the experiment to weigh any remaining residual compound. This was to ensure accuracy regarding the amount of compound addition and calculation of the correct concentration. Samples were prepared by adding aliquots of stock solutions of nucleobase and clamp via syringe directly into NMR tubes and diluting to 0.5 mL. After optimisation it was found that $CDCl_3$ with additional 5% DMSO-d_6 increased solubility for both the A and G clamps.

A 400 MHz NMR spectrometer was utilised for the binding studies. The NMR tube and vial containing the products in solution were sealed with a septum to ensure a minimal amount of $CDCl_3$ evaporated over the duration of the experiment. The increments of solvent was between 10-50 µL using a 50 µL micro syringe. The sample tube was shaken carefully after each addition.

Initial ^{1}H NMR analysis of the A clamp in deuterated chloroform was actually very broad and it was difficult to recognise the presence of the A clamp in solution. This could be due to self aggregation in solution of the highly aromatic planar molecule. Hence poor solubility in $CDCl_3$ was a concern when carrying out the “reverse titration” and “forward titration” binding studies in an attempt to observe a 1-1 complex between the A clamp and 9-ethyl adenine.

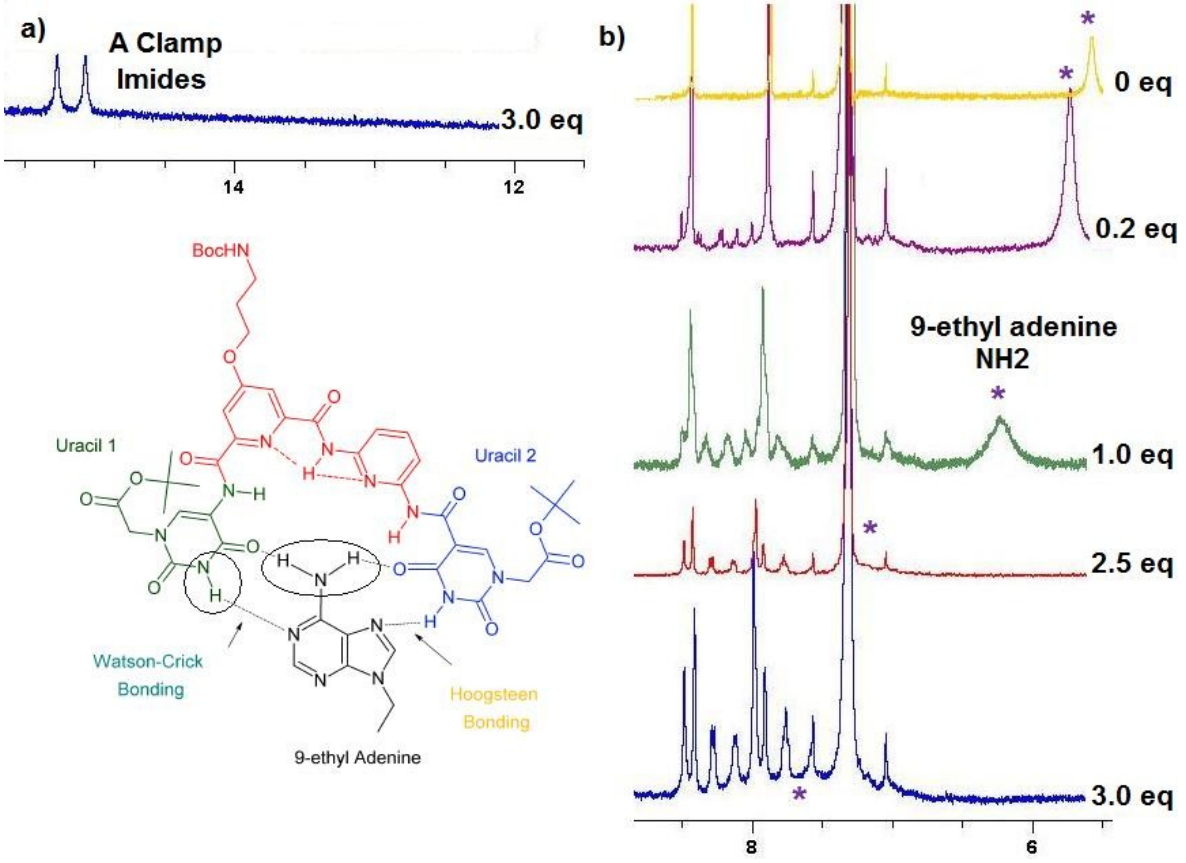
In contrast, the ^{1}H NMR spectra in DMSO-d_6 was sharp and clear due to the denaturing effect of this solvent. Based on Rebek's work on adenine receptors, it was thought that upon addition of 9-ethyl adenine this could contribute to reorganisation and therefore help resolve the spectra quality. It was decided to replicate conditions similar to those conducted by Rebek's group.[91] The “forward titration” was initially preferred due to the poor solubility of the A clamp. This enabled to fix the concentration of the A clamp with addition of 9-ethyl adenine aliquots to this.

For the first titration experiment, a forward titration was conducted on a 4 mM scale of A clamp (receptor). Due to the poor solubility of the A clamp in $CDCl_3$, the A clamp was used as a receptor and therefore the very soluble 9-ethyl adenine (guest) was added to it by micro syringe accordingly. Note that the poor solubility of the A clamp in $CDCl_3$ resulted in poor shimming and signal-to-noise ratio. Also any impurities which might not have been observed by TLC, HPLC or NMR analysis could be more prominent and apparent within the study due to the impurities being more soluble than the A clamp itself in $CDCl_3$. Thus only a small proportion of the overall compound was actually fully soluble in the NMR tube.

This was further confirmed when 0.2 equivalents of 9-ethyl adenine was added to the A clamp. According to ^{1}H NMR integration of the signals at 0.2 equivalent addition, it corresponds more so to that of one equivalent of 9-ethyl adenine. Hence the poor solubility issue obscured the initial titration results, since the equivalents added was not a true representation of what was happening within the system. Not only was solubility issues a challenge for determining the binding constant but also following the 9-ethyl adenine NH_2 signal was difficult as it shifted downfield and disappeared underneath the aromatic signals. The imides before binding were also not well resolved and became more sharp upon addition of 9-ethyl adenine. Thus there was no starting point to focus on, in order to monitor the signal throughout the titration study from start to finish. Any imide signals or even A clamp signals which were observed were those which were involved in binding. The unbound A clamp was not observed at all, due to solubility issues in $CDCl_3$. This meant it was difficult to calculate an accurate binding constant after the initial titration experiment.

A “reverse titration” experiment was carried out, using a fixed concentration of 9-ethyl adenine (receptor) at a 2 mM scale with the A clamp (guest) added to this during the titration. It was desired to monitor the change in chemical shift for the 9-ethyl adenine NH_2 and the H-C2 and H-C8 protons upon binding. The benefit of this is that there is a clear starting point. However, poor solubility of the A clamp was still an issue. The NMR titration conditions were optimised to obtain an accurate binding constant value. What was interesting though at this stage, was that the affinity seemed to be extremely high. This high affinity is evident from the significant change in chemical shift *(Spectra 2)*.

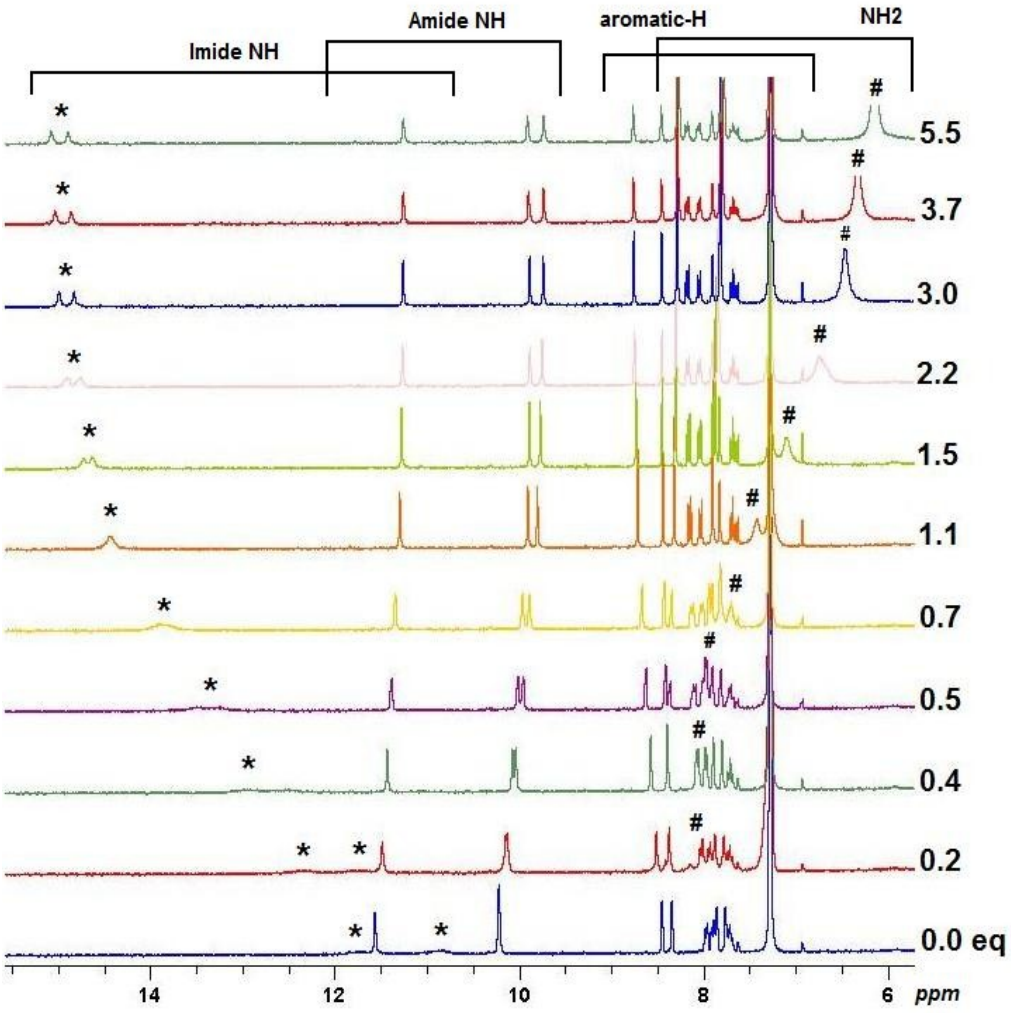
The A clamp imide protons involved in hydrogen bonding which normally reside around 10.9 ppm and 11.9 ppm were shifted downfield towards 15.2 ppm *(Spectra 2)*. The NH_2 from 9-ethyl adenine has also shifted downfield from 5.7 ppm towards ~7.9 ppm. This significant change in chemical shift suggests that the A clamp is binding very well via simultaneous Watson-Crick and Hoogsteen hydrogen bonding to the 9-ethyl adenine. For this to occur after the addition of 0.2 equivalents 9-ethyl adenine is indicative that any of the soluble A clamp is engaged in binding to 9-ethyl adenine.



Spectra 2: ^{1}H NMR Reverse titration in $CDCl_3$: 9-ethyl adenine is kept at a fix concentration (receptor, 2 mM) and aliquots of A clamp (guest) is added. a) A clamp imides exhibit a significant downfield chemical shift towards 14.5-15.2 ppm when bound to 9-ethyl adenine upon addition of 3 equivalents of A clamp b) 9-ethyl adenine NH_2 has a chemical shift downfield from 5.7 ppm towards 7.9 ppm when complexed to increasing aliquots of A clamp.

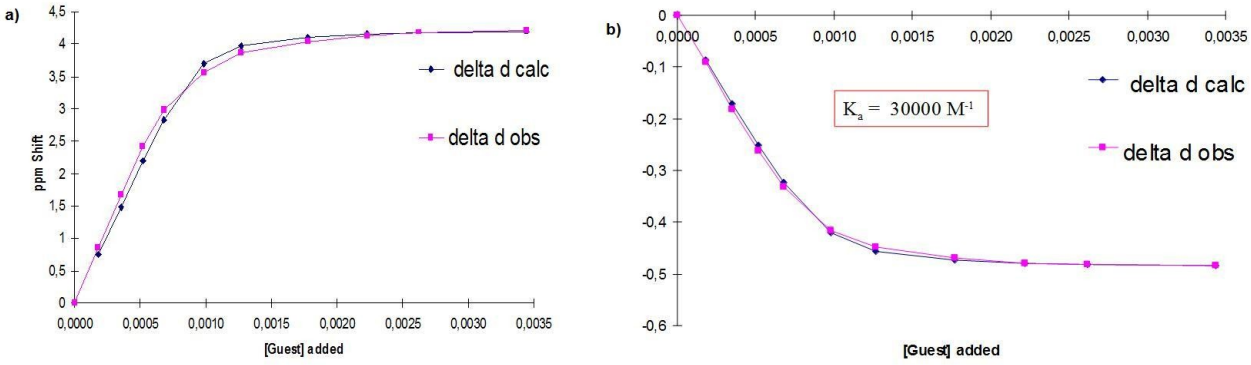
After these very interesting initial A clamp ^{1}H NMR studies, further titration experiments were optimised based on these findings and found to have a $K_{assoc} \geq 10^6$ M^{-1} in $CDCl_3$.[83] Due to solubility issues these results were not accurate. To overcome the previous solubility problem, 5% DMSO-d_6 was added to $CDCl_3$ increasing the A clamps solubility. This of course would have a slight impact on the hydrogen bonding and slightly reduce binding affinity. However the A clamps increased solubility would compensate for this drawback. It was also important to conduct the NMR studies at low concentration to prevent aggregation or precipitation, thus the A clamp concentration was fixed at 1 mM. This concentration was determined to be suitable after conducting NMR experiments with the A clamp at a variety of concentrations to investigate the formation of self aggregates.

A "forward titration" was therefore carried out using a ^{1}H NMR 400 MHz spectrophotometer, with the A clamp (receptor) concentration fixed at 1 mM in $CDCl_3$/5% DMSO-d_6 and aliquots of 9-ethyl adenine (guest) added to this *(Spectra 3)*. Similar observations to the previous titrations were made regarding a significant change in chemical shift for the A clamp imides N-H and the 9-ethyl adenine NH_2 upon increased binding. The spectra signals were sharp and the solubility issue has been addressed.



Spectra 3: ^{1}H NMR Forward titration in $CDCl_3$/5% DMSO-d_6: A clamp (receptor) concentration fixed at 1 mM. Addition of 9-ethyl adenine (guest) using a 400 MHz NMR in $CDCl_3$/5% DMSO-d_6

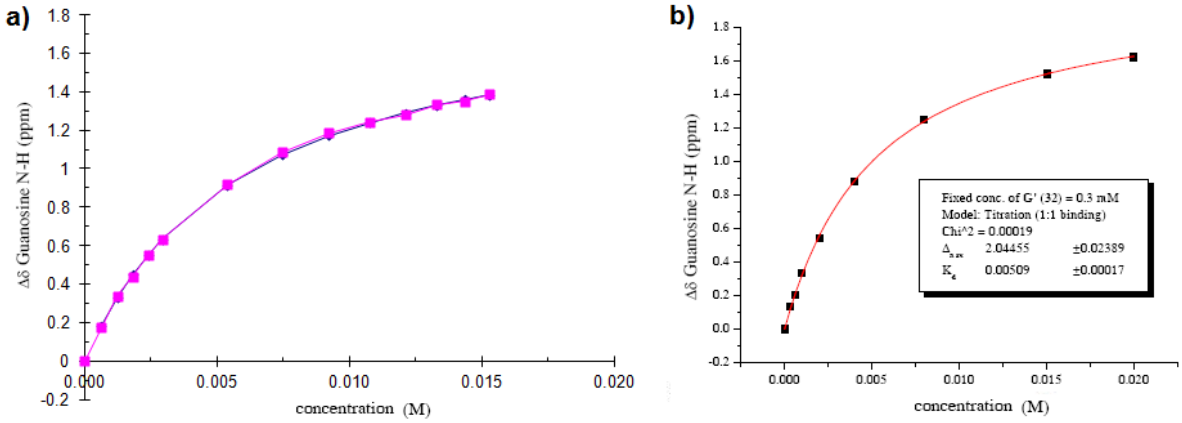
A graph was plotted "Δδ change in chemical shift of the A clamp imide/amide N-H (receptor)" versus "concentration of 9-ethyl adenine (guest) addition". The resulting binding curve (experimental vs calculated data) is shown *(Graph 1)*. The A clamp *(Graph 1a)* Δδ of the imides N-H and *(Graph 1b)* Δδ of the amides N-H were monitored upon complexation. A binding constant of K_{assoc} = 30000 M^{-1} was obtained, using a non-linear least squares curve-fitting program implemented within Excel. The programme yields binding constants K_{assoc} and limiting Δδ as output. This result shows a high affinity between the A clamp and adenine. It is higher than previous binding constants reported by Hamilton K_{assoc} = 3200 M^{-1},[29] Zimmerman K_{assoc}= 15500 M^{-1} [90] in $CDCl_3$.



Graph 1: ^{1}H NMR Forward titration in $CDCl_3$/5% DMSO-d_6: "Change in A clamp (receptor fixed at 1 mM) chemical shift Δδ (ppm)" versus "concentration (M) of 9-ethyl adenine (guest)" to obtain a binding constant between the A clamp and 9-ethyl adenine complex (13 °C). a) A clamp (receptor fixed at 1 mM) imide N-H chemical shift was monitored upon complexation with 9-ethyl adenine (guest). b) A clamp (receptor fixed at 1 mM) amide N-H chemical shift was monitored upon complexation with 9-ethyl adenine (guest). Both graphs give a binding constant K_{assoc} = 30000 M^{-1}.

5.3.1.2 ^{1}H NMR titration: G clamp and guanosine

The 1:1 complexation between the G clamps **46** and **48** and guanosine **47** was also investigated. The structures of these molecules were similar to those reported by Zimmerman [38] with an extra solubilising side chain attached. The titration conditions were optimised and validated. A "reverse titration" was carried out and the concentration of guanosine **47** (receptor) was fixed to 0.3 mM. Similar to the A clamp titration, 5% DMSO-d_6 was added to both of the $CDCl_3$ stock solutions of the receptor and guest to increase compound solubility. These parameters provided a clear ^{1}H NMR spectra in which to monitor changes in chemical shift and calculate a binding constant. The binding study for G clamp **46** and guanosine **47** spectra showed a significant chemical shift downfield for the G clamp **46** amide N-H from 10.61 ppm with no guanosine **47**, to 11.87 ppm when 0.5 equivalents of G clamp **46** was added to guanosine **47**. This is indicative of complexation and an affinity between the G clamp **46** and guanosine **47**. The binding is at its optimum at around 0.5 equivalents of G clamp **47**. As the concentration increases, the binding is not so tight, due to a slight increase in solvent and competition from self aggregation. The amide signal starts to move upfield to its original position due to the availability of more unbound G clamp **46** in solution. A graph was plotted "Δδ change in chemical shift of the guanosine N1 (N-H) **47** (receptor)" versus "concentration of G clamp **46** (guest) addition". The resulting binding curve (experimental vs calculated data) is shown *(Graph 2)*. A binding constant of K_{assoc} = 236 M^{-1} was obtained, using a non-linear least squares curve-fitting program implemented within Excel. The programme yields binding constants K_{assoc} and limiting Δδ as output. This result shows a low affinity in $CDCl_3$/ 5%-DMSO-d_6 between the G clamp **46** and guanosine **47**. It is similar to the binding constant K_{assoc} = 230 M^{-1} reported by Zimmerman for the amino G clamp derivative *(Fig. 26a)*.



Graph 2: ^{1}H NMR Reverse titration in $CDCl_3$/5% DMSO-d_6: a) "Change in guanosine N-H **47** (receptor fixed at 0.3 mM) chemical shift Δδ (ppm)" versus "concentration (M) of G Clamp **46** (guest)" resulting in a binding constant K_{assoc} = 236 M^{-1}. b) "Change in guanosine N-H (receptor fixed at 0.3 mM, *Fig. 26a*) chemical shift Δδ (ppm)" versus "concentration (M) of Zimmermans amino G clamp derivative (guest, *Fig. 26a*)" resulting in a binding constant K_{assoc} = 230 M^{-1}.

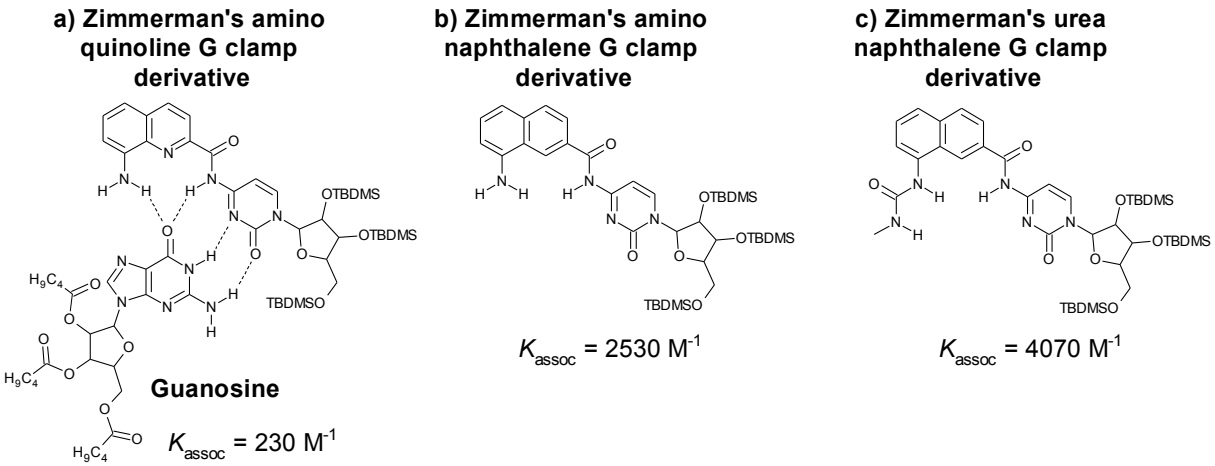
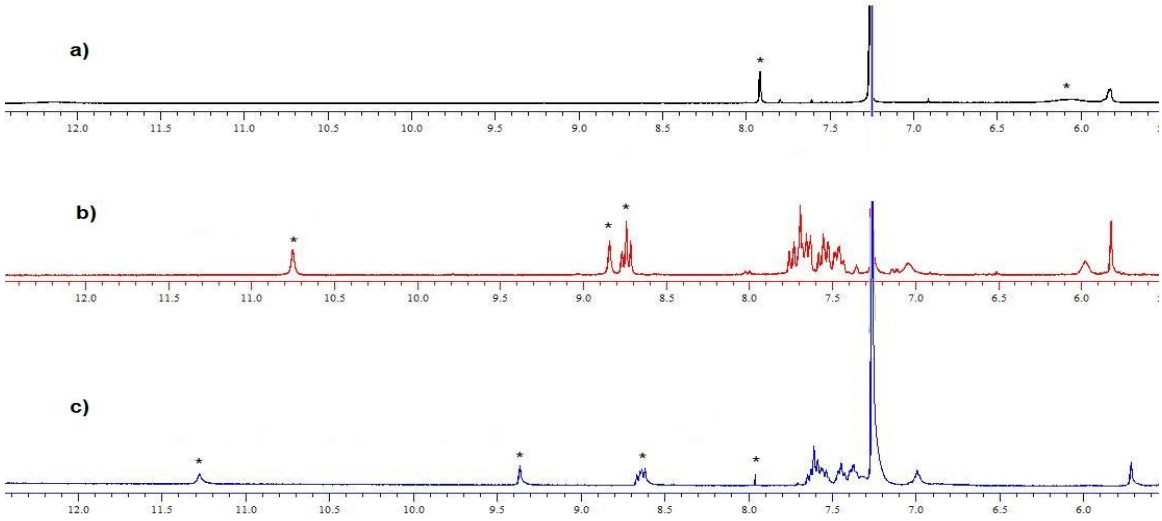


Figure 26: Zimmerman's G clamp derivatives. a) Zimmerman's amino quinoline G clamp derivative complexed to guanosine. The binding constant K_{assoc} = 230 M^{-1}. b) Zimmerman's amino naphthalene G clamp derivative with a binding constant K_{assoc} = 2530 M^{-1} upon complexation with guanosine. c) Zimmerman's urea naphthalene G clamp derivative with a binding constant K_{assoc} = 4070 M^{-1} upon complexation with guanosine.[38]

A similar titration study was carried out for quinoline urea G clamp **48**. Despite using similar NMR conditions, concentrations and addition of 5% DMSO-d_6 the solubility of the quinoline urea G clamp **48** was poor. As a result, the equivalents of **48** added to the guanosine **47** do not correspond to the calculated integration values. Also the guanosine N1 signal broadens and disappears upon addition of G clamp **48**. There was not enough material to optimise these conditions. However there is a noticeable chemical shift difference for the quinoline urea G clamp **48** before and after addition to the guanosine **47**. The urea amide signals move from 10.75 and 8.84 ppm to 11.25 and 9.38 ppm respectively (Spectra 4). Zimmerman's urea naphthalene G clamp derivative *(Fig. 26c)*, was reported to have a binding constant K_{assoc} = 4070 M^{-1}.[38] However their amino quinoline G clamp derivatives had

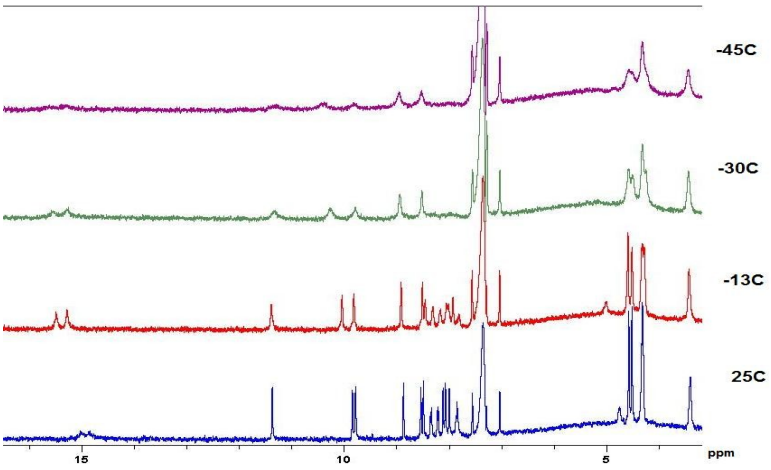
lower binding constant values K_{assoc} = 230 M^{-1} *(Fig. 26a)* and K_{assoc} = 2530 M^{-1} *(Fig. 26b).* This difference in binding constants suggests an unfavourable electrostatic interaction between the quinoline nitrogen lone pair and the guanosine carbonyl group. The addition of a tethered ureido group to the G clamp **48** was expected to maximise the Hoogsteen bonding and increase binding. However, the presence of the quinoline backbone in **48** as oppose to the naphthalene unit, was expected to create unfavourable electrostatic interactions. For this reason and due to time constraints the urea G clamp **48** compound was not investigated further.



Spectra 4: ^{1}H NMR Reverse titration in $CDCl_3$/5% DMSO-d_6: a) Guanosine **47** $CDCl_3$/5% DMSO-d_6; b) Urea G clamp **48**; c) Guanosine **47** complexed to urea G clamp **48**.

5.3.1.3 ^{1}H NMR variable temperature: assessment of induced binding geometry between A clamp and 9-ethyl adenine.[83]

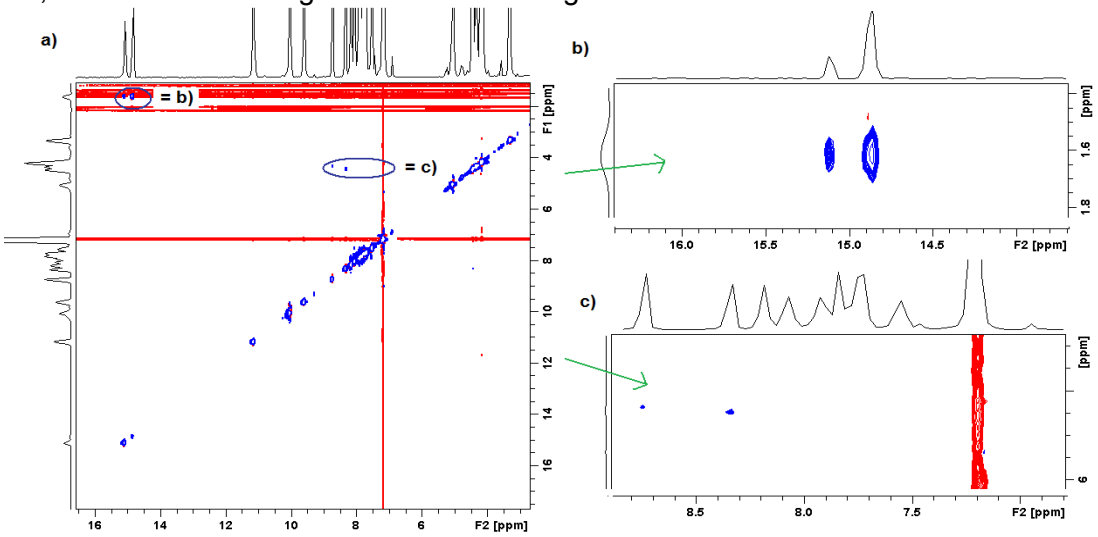
Addition of 9-ethyl adenine to a solution of A clamp (0.2 mM) in a 1:1 complex in $CDCl_3$ at low temperature (-45 °C, 228 K; -30 °C, 243 K; -13 °C, 260 K) result in the NMR signals becoming quite broad. *(Spectra 5).* However one species is observed and this suggests that the A clamp binds specifically with one preference for binding according to low temperature ^{1}H NMR studies. At room temperature, there is a rapid interconversion between the N^1 and N^7 positions of 9-ethyl adenine in a fast exchange, but at lower temperatures the exchange is slowed down. If exchange was occurring, it was expected that the two imide signals would split into four imide signals at 15.2 ppm at low temperatures. This phenomenon would be caused by the 9-ethyl adenine binding one way with the A clamp, exiting and re-binding in the opposite orientation via a reorganisation. As a result different chemical environments of the Watson-Crick and Hoogsteen modes would exist. However, splitting of the two imide signals is not observed, suggesting a preference for hydrogen bonding of the 9-ethyl adenine binding to the A clamp. This was an inconclusive result since exchange might be too fast to measure on an NMR time scale.



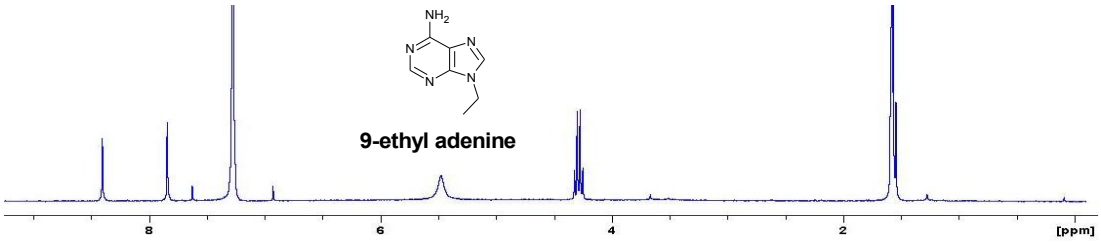
Spectra 5: Varied temperature experiment ^{1}H NMR of 1:1 clamp-Adenine Complex (0.2 mM), 25 °C, 298 K; -13 °C, 260 K; -30 °C, 243 K; -45 °C, 228 K.

5.3.1.4 ^{1}H NMR assessment of solution structural studies between A clamp and 9-ethyl adenine.[83]

Two-dimensional ^{1}H NMR studies in $CDCl_3$ were used to obtain information on the geometry of the one-to-one complexes between 9-ethyl adenine and the A clamp. A NOESY experiment was conducted to further access the intermolecular contacts. The A clamp imides bind to the NH_2 of the 9-ethyl adenine. The signal-to-noise ratio was high due to the *t*-butyl signals, however if these signals are reduced signal information would also be reduced.



Spectra 6: a) NOESY 1:1 Complex between 9-ethyl adenine and the organic soluble A clamp in $CDCl_3$ b) Expansion of the A clamp imides and their correlation with 9-ethyl adenine; c) Expansion of 9-ethyl adenine Ar C-H and their correlation with the A clamp.



Spectra 7: 9-ethyl adenine in $CDCl_3$

Solid state complexation studies focused on co-crystallisation of the 9-ethyl adenine and A clamp. Various solvents were investigated ranging from DMSO, DMF, chloroform, toluene with diethyl ether, *n*-hexane combinations in gas phase and solution phase studies did not give a desired co-crystal. It was thought that the binding of the A clamp to 9-ethyl adenine in solution might increase solubility and thus produce co-crystals. Rebek's work suggested slow diffusion of *n*-hexane into chloroform of the two solutions.[93] However all proved unsuccessful and a co-crystal of the final A clamp was not obtained. Using a tertiary solvent to create an interface did not provide the desired co-crystals. A few crystals were obtained but each time they consisted of 9-ethyl adenine. In future studies, solid state NMR could be investigated in between the 1:1 complex.

5.3.2 Clamp complexation to corresponding nucleobase in aqueous solvent

5.3.2.1 SPR studies: A clamp binding to an abasic site.[94]

The objective was to investigate binding effectiveness of the aqueous soluble A clamp towards an abasic site in water. For this study an oligonucleotide containing a free abasic site for binding should enable the A clamp to complex with its corresponding adenine base and restore *π-π* stacking *(Fig. 27)*.Thus the nucleobases flanking the abasic site were expected to play an important role in the complexation. The overall effect was expected to be detected by a change in MW on the sensor chip before and after complexation via surface plasmon resonance. The characterisation of the complexation could access the specificity of interaction, the kinetics via on and off rates ($K_d = k_{off}/k_{on}$), the affinity and binding strength, and the thermodynamic properties underlying association and dissociation. A high binding capacity would give a high response.

The aqueous A clamp was modified by attaching a biotinylated side chain *(Fig. 28)*. This enabled the biotinylated A clamp to be immobilised onto the sensor chip surface via streptavidin. A buffer solution containing the oligonucleotide was flown over the A clamp sensor chip surface. A real time output in the form of a sensorgram should be observed upon association *(Fig. 29)*. Thus enabling kinetics to be measured and the chip regeneration upon washing with EDTA. However this was not the case and the experiment did not go as accordingly. Two explanations could be: the adenine nucleobase on the oligonucleotide is inverted outwards causing the duplex to fold in upon itself preventing complexation; or the biotinylated side chain is too long and therefore the A clamp is not positioned in the desired upright position, perhaps residing on the surface and unable to bind with the oligonucleotide. Thus this experiment using these conditions proved inconclusive in determining the ability of the aqueous A clamp to bind to an abasic site in physiological conditions.

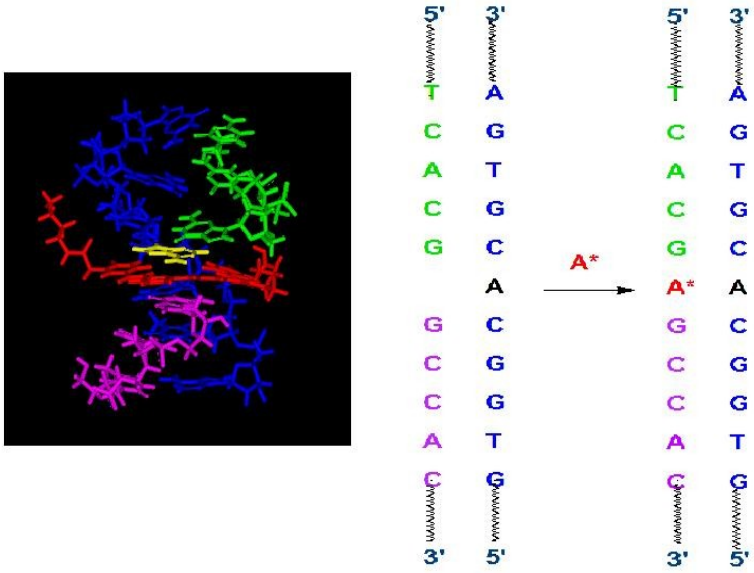


Figure 27: (Left) Molecular Model of oligonucleotide duplex and A clamp (red); (Right) Schematic representation of the oligonucleotide duplex with one nucleotide removed and insertion of A clamp within the abasic site.

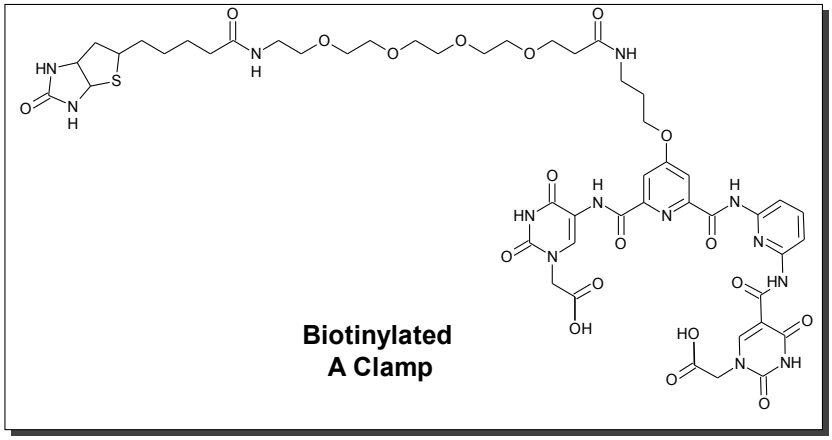


Figure 28: The modified aqueous soluble A clamp with a biotinylated chain attached.

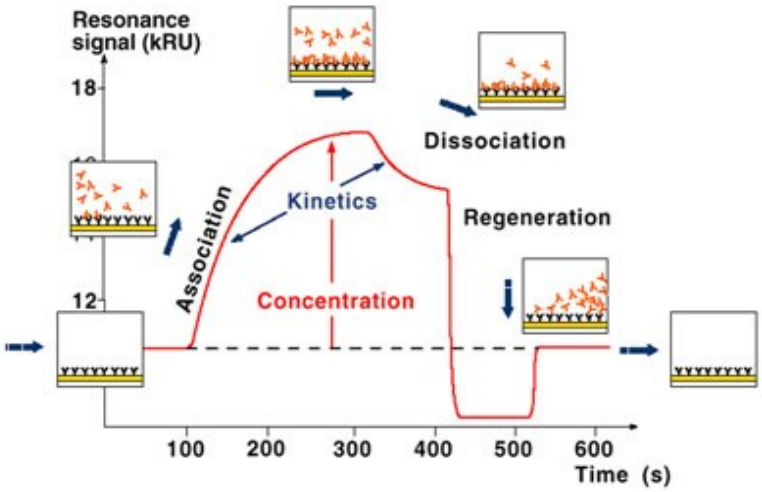


Figure 29: SPR sensorgram (Resonance signal vs Time). As molecules bind to and dissociate from the sensor chip surface, the resulting changes in the resonance signal create a sensorgram. The shape and amplitude of the measurements could then be used to determine overall concentration and the kinetics of the interaction. (Biacore)

5.3.2.2 CD studies: A clamp binding to a ssDNA.[83]

In preparation for CD and UV melting temperature studies the UV/Vis parameters of the organic soluble A clamp and deprotected aqueous A clamp were investigated before conducting binding experiments. *(Fig. 30)*.

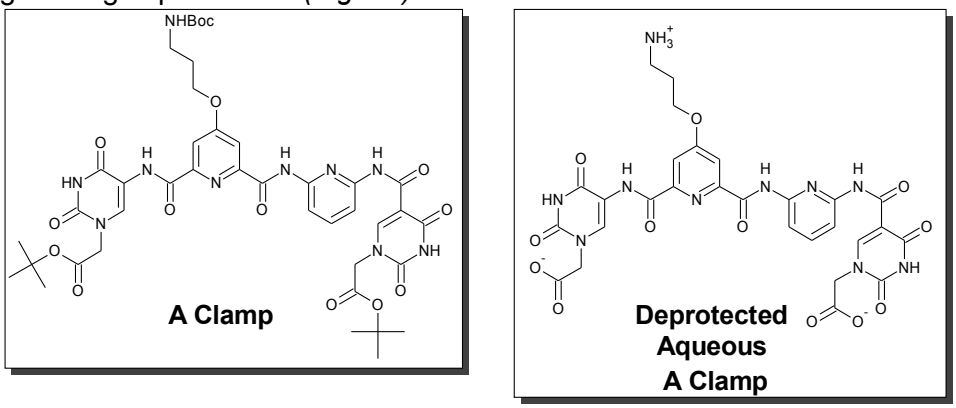
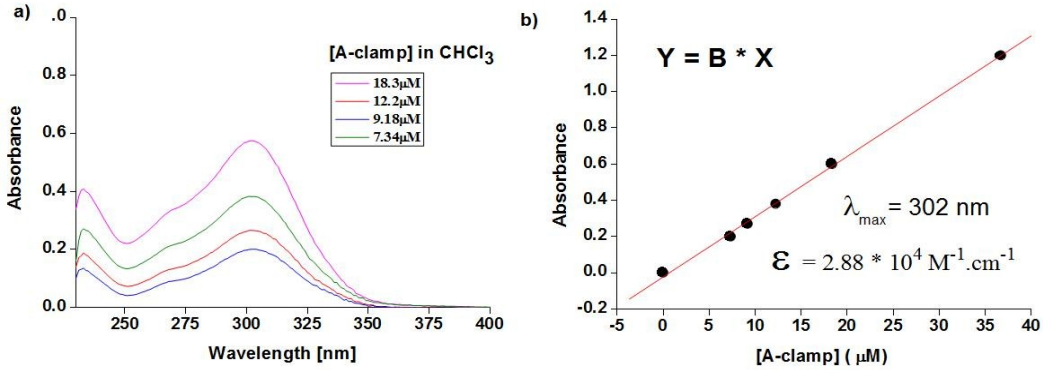
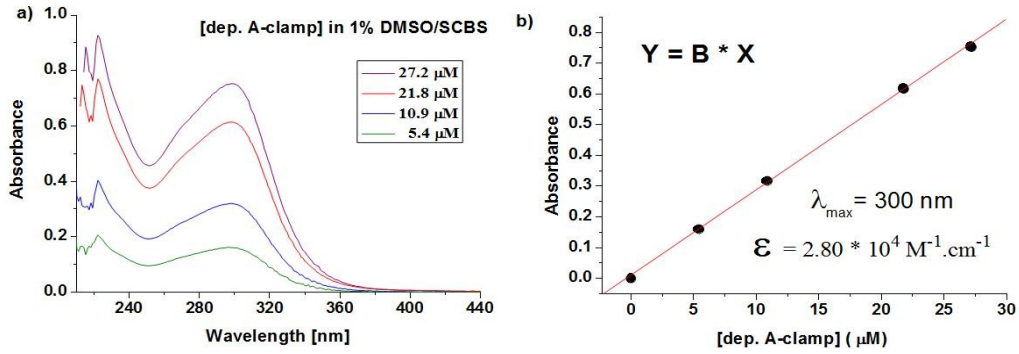


Figure 30: (Left) Organic soluble A clamp; (Right) deprotected aqueous soluble A clamp

The absorbance of different concentrations of the organic A clamp were investigated in $CHCl_3$. Using the Beer-Lambert law A = εcl, the ε value could be calculated *(Spectra 8)*. This was found to be $\varepsilon = 2.88*10^4 M^{-1}.cm^{-1}$. A similar scenario was carried out for the deprotected A clamp in sodium cacodylate buffer solution with 1% of DMSO added *(Spectra 9)*. The λ_{max} = 300 nm and $\varepsilon = 2.80*10^4 M^{-1}.cm^{-1}$.



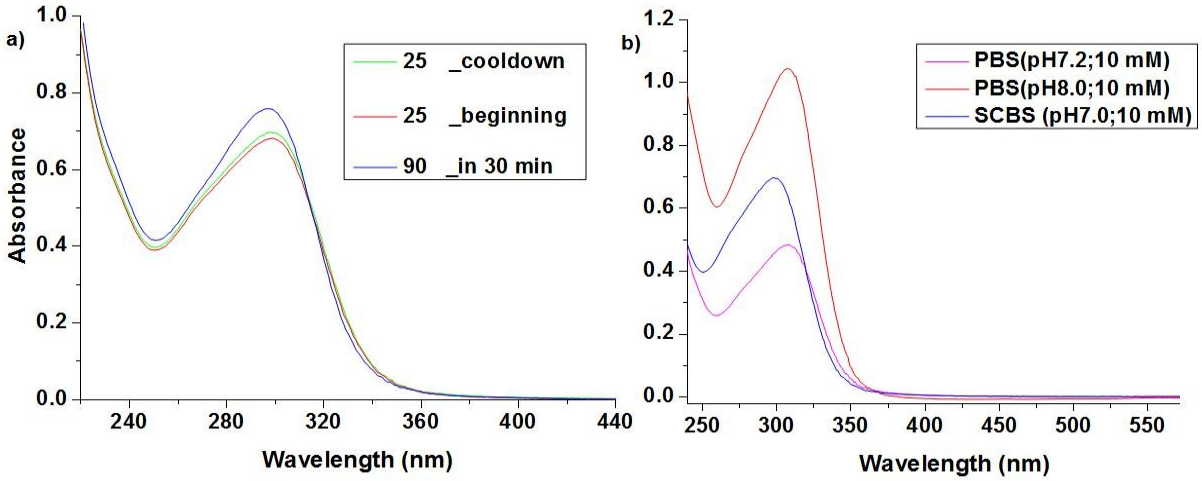
Spectra 8: a) UV Spectra: Absorbance vs Wavelength (nm) at various different [A clamp] in $CHCl_3$. b) Graph of A clamp Absorbance versus [A clamp] (µM), λ_{max} = 302 nm and $\varepsilon = 2.88*10^4 M^{-1}.cm^{-1}$



Spectra 9: a) UV Spectra: Absorbance vs Wavelength (nm) at various different [deprotected A clamp] in SCBS: sodium cacodylate buffer solution and 1% DMSO. b) Graph of deprotected A clamp Absorbance versus [deprotected A clamp] (µM), λmax = 300 nm and $\varepsilon = 2.80*10^4 M^{-1}.cm^{-1}$

The stability of the deprotected A clamp was also assessed by heating it to 90 °C for thirty minutes to observe if any degradation occurs upon heating in sodium cacodylate buffer solution with 1% of DMSO added *(Spectra 10)*. Little change was observed in the absorbance before and after heating from a wavelength of 200 to 440 nm. This suggests the A clamp is stable upon heating and cooling.

Buffer effects were also investigated by comparing phosphate buffered saline buffer (PBS) against sodium cacodylate buffer (SCBS) *(Spectra 10)*. There is a slight variation observed in λ_{max} from 300 nm to 308 nm. It was also noticeable that adenosine highly improves the solubility of deprotected A clamp. 1% of DMSO was also found to improve the deprotected A clamp solubility reducing the formation of precipitate.



Spectra 10: a) Stability of deprotected A-clamp, dep. A-clamp is stable while cooling and heating in buffer solution, b) Changes of λ_{max} from 300 to 308 nm in different buffer solutions as observed by UV/Vis. SCBS: sodium cacodylate buffer solution (10 mM, pH 7.0)

CD was used to investigate the chromophoric interaction in detail. A Jasco J-815 Circular Dichroism Spectrometer was used for the following experiments. This instrument simultaneously measured the UV/Vis absorbance *(see Appendix)*. The CD spectra of the ssDNA in the absence of the A clamp did not result in a Cotton effect at 320 nm. It was expected that despite the competition for hydrogen bonding in water that the aromatic A clamp surface could compensate for this via *π-π* stacking.

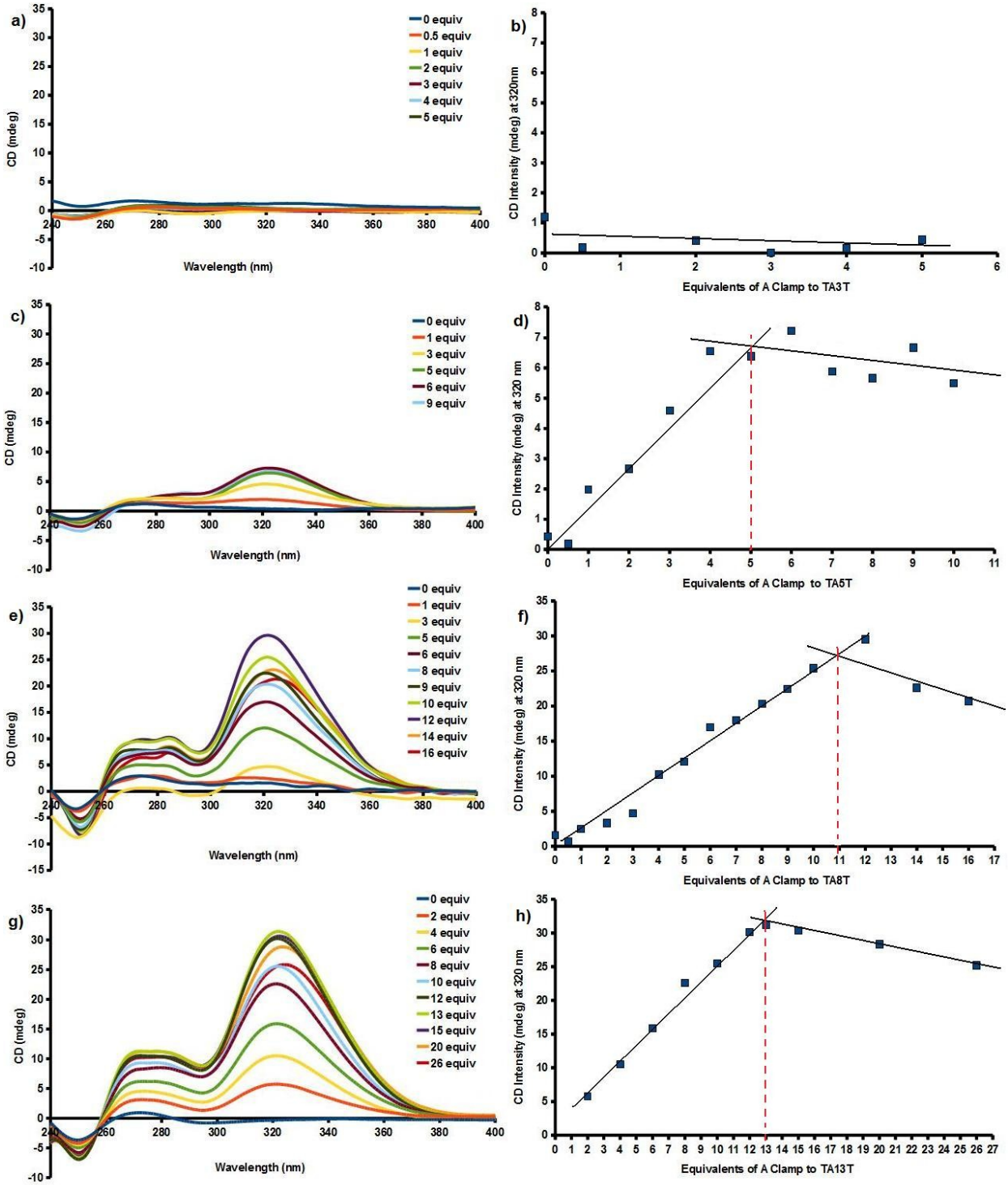
For CD titration optimisation, TA_8T was used as a template to validate investigations by varying the parameters such as template concentration and equilibration time *(see Appendix)*. Various equivalents of deprotected water-soluble A clamp were added to the ssDNA template at different concentrations. 2 μM of ssDNA was found to be the preferred concentration as higher concentrations gave irreproducible results. The phosphate buffered saline (PBS) buffer at 10 mM, pH 8, 150 mM NaCl also had an additional 1% DMSO to aid solubility.

Optical activity was evaluated upon the addition of A clamp to various template lengths, TA_nT (n = 5, 8, 13) resulting in a positive Cotton effect *(Spectra 11 c, e, g)*. Thus a preferred right-handed helical arrangement at a given pH 8.0. The template TA_nT (n = 3) upon addition of 0, 1, 3 and 5 equivalents of A clamp showed little capability to self-assembly as evident from the flat CD spectra between the wavelength 250-450 nm and therefore lacks a point of saturation *(Spectra/Graph 11 a, b)*. Thus the available nucleobase sites were disfavoured for binding on short templates.

Template TA_nT (n = 5) and addition of 0, 1, 3, 5 and 8 equivalents of A clamp showed a slightly improved ability to self-assemble. Five equivalents of deprotected A clamp proved optimum with a maximum at 323 nm and 7.08 mdeg CD intensity. More equivalents of A clamp had little influence towards this template as observed by the addition of eight equivalents of A clamp which had a maximum of 5.4 mdeg CD intensity at the same wavelength. Suggesting that all available nucleobase binding sites were filled *(Spectra/Graph 11 c, d).*

Template TA_nT (n = 8) and addition of 0, 2, 4, 6, 8, 12 and 16 equivalents of A clamp proved more favourable for self-assembly and inducing a cotton effect in the spectra at 322 nm, whereas this template alone has a maximum at 273 nm and CD intensity of 4.23 mdeg. Upon increasing A clamp equivalents, there is a progressive increase in CD intensity at 322 nm. After the addition of eight equivalents the intensity continues to increase upon adding twelve and sixteen equivalents of A clamp but the change in CD intensity is decreasing. The λ_{max} shifts a little, perhaps due to not enough time given for equilibration *(Spectra 11 e)*. Thus the experiment was repeated with time given for the A clamp and TA_8T template to equilibrate *(see Appendix)*. The graph CD (mdeg) vs wavelength (nm) was extrapolated and eleven equivalents of A clamp was found to be the point of saturation for TA_nT (n = 8) *(Spectra/Graph 11 e, f).*

Template TA_nT (n = 13) and addition of 0, 2, 4, 6, 8, 13, 15, 20 and 26 equivalents of A clamp exhibited excellent capabilities to self-assemble. Again a trend of increasing CD intensity at 321 nm upon increased addition of A clamp to the template suggested duplex formation. A CD maximum of 31.06 mdeg is obtained with addition of thirteen equivalents of A clamp, corresponding to complete binding with the complementary ssDNA. The graph CD (mdeg) vs wavelength (nm) was extrapolated and thirteen equivalents of A clamp was found to be the point of saturation for TA_nT (n = 13) *(Spectra/Graph 11 g, h).* Upon additional A clamp, the CD intensity did not increase in amplification.



Spectra (and Graphs) 11: CD Titration spectra of various oligonucleotides TA_nT (n = 3, 5, 8, 13) [2 μM] + addition of A clamp in PBS buffer, 10 mM, pH 8, 150 mM NaCl and 1% DMSO at 293 K. a) CD (mdeg) vs wavelength (nm), TA_3T template and 0, 0.5, 1, 2, 3, 4 and 5 equivalents of A clamp; b) Graph of CD intensity (mdeg) at 320 nm vs equivalents of A clamp added to TA_3T in order to extrapolate the point of saturation; c) CD (mdeg) vs wavelength (nm), TA_5T template and 0, 1, 3, 5, 6 and 9 equivalents of A clamp; d) Graph of CD intensity (mdeg) at 320 nm vs equivalents of A clamp added to TA_5T; e) CD (mdeg) vs wavelength (nm), TA_8T template and 0, 1, 3, 5, 6, 8, 9, 10, 12, 14 and 16 equivalents of A clamp; f) Graph of CD intensity (mdeg) at 320 nm vs equivalents of A clamp added to TA_8T; g) CD (mdeg) vs wavelength (nm), $TA_{13}T$ template and 0, 2, 4, 6, 8, 10, 13, 15, 20 and 26 equivalents of A clamp; h) Graph of CD intensity (mdeg) at 320 nm vs equivalents of A clamp added to $TA_{13}T$.

Upon addition of the A clamp to the template, it was suspected that the ssDNA would be initially filled while containing some empty sections before the complete formation of the duplex *(Fig. 31)*.[52] Thus, it was important to allow for equilibrium to be reached during the CD titration study.

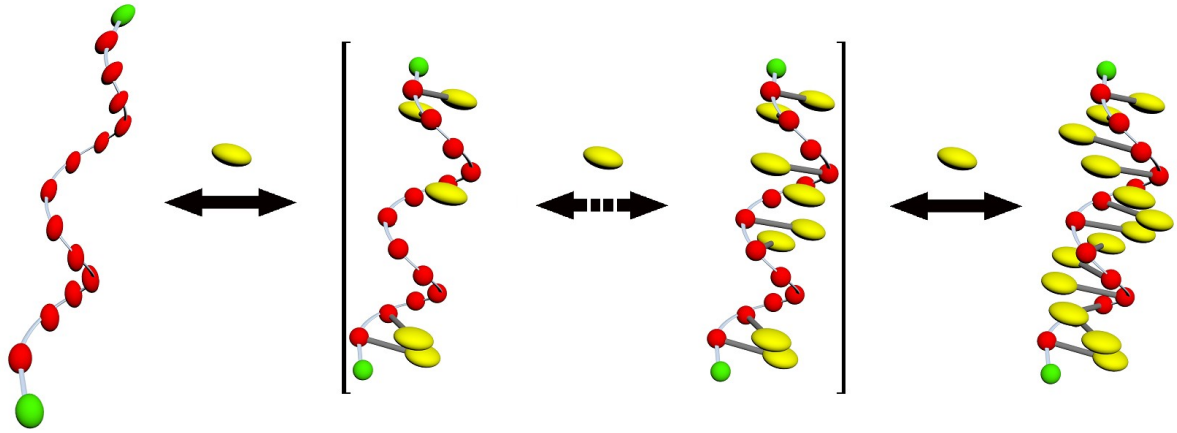
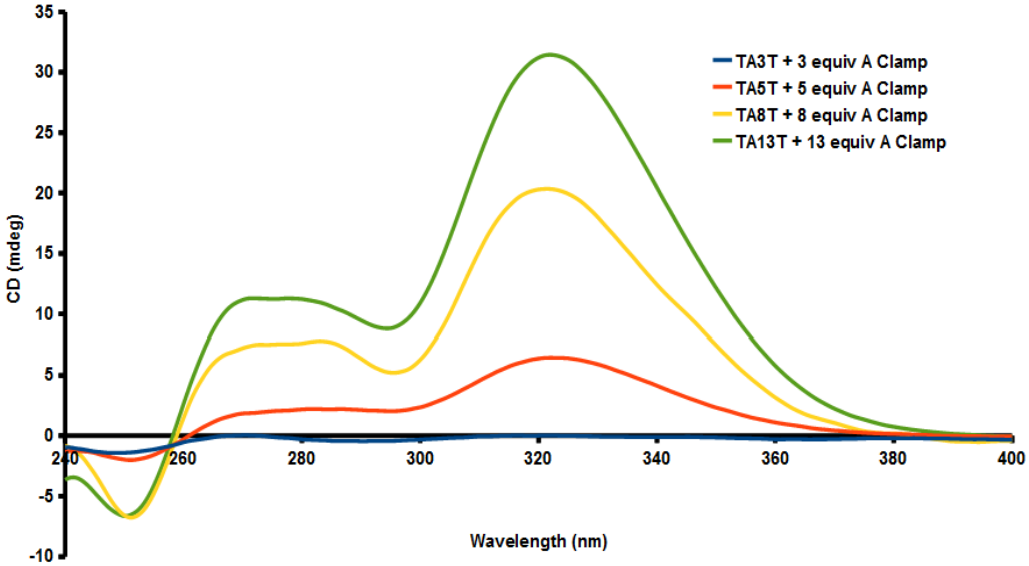


Figure 31: Schematic representation of the ssDNA, $TA_{13}T$ template (left) with addition of A clamp. The template is filled and contains some empty sections (centre) in equilibrium before the formation of the duplex (right).[75]

Overall this CD study compared the effect of template size upon self-assembly of A clamp in aqueous conditions, TA_3T < TA_5T < TA_8T < $TA_{13}T$ *(Spectra 12)(Fig. 32)*. The longer templates enhance duplex stability possibly due to a cooperative effect via host-guest interactions such as *π-π* stacking. The next objective will be to assess intercalation of a guanine mismatch onto a poly A ssDNA, however this was not carried out and should be investigated to test if the A clamp could distinguish between this. Would the mismatch enhance stability or disrupt the system and what would happen to a ssDNA containing both adenines and guanines when a mixture of clamps was introduced? This experiment would test the receptors specificity and recognition ability to discriminate between bases.



Spectra 12: CD spectra of various oligonucleotides TA_nT (n = 3, 5, 8, 13) [2 μM] + n equivalents of A clamp added in PBS buffer, 10 mM, pH 8, 150 mM NaCl and 1% DMSO at 293 K.

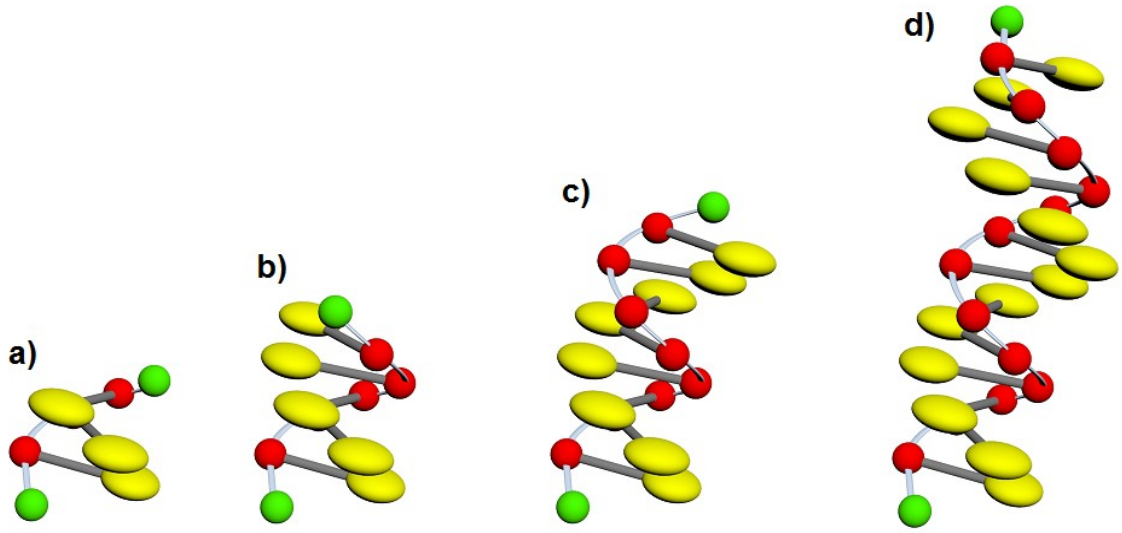


Figure 32: Schematic representation of the various DNA:A clamp duplexes: a) TA_3T:A $clamp_3$; b) TA_5T:A $clamp_5$; c) TA_8T:A $clamp_8$; d) $TA_{13}T$:A $clamp_{13}$.[75]

5.3.2.3 UV melting studies: A clamp and ssDNA

Experimental parameters: Experiment carried out on Cary 300 Scan UV-Vis Spectrophotometer. Buffer solution was sodium cacodylate; 10 mM; pH 7.0. Start temperature = 15 °C; return temperature = 95 °C; rate = 1 °C per minute. The advantage of using a cacodylate buffer is that there was only a slight variation in pH with increase in temperature.

UV melting studies were used to characterise the formation and melting behaviour (T_m) of an intramolecular DNA duplex between the oligonucleotide strands TA_nT (n = 3, 5, 8, 13) and the aqueous soluble deprotected A clamp analogue *(Fig. 33)*. The experiments showed that A clamp alone or the ssDNA do not exhibit a melting temperature curve or denaturation. After annealing the A clamp with the ssDNA TA_nT (n = 3, 5, 8, 13) to form a duplex in sodium cacodylate buffer, denaturation behaviour is observed and indicative of unfolding upon increasing temperature and reorganisation upon cooling. Thus base stacking is a relatively important stabilising interaction for a duplex structure.

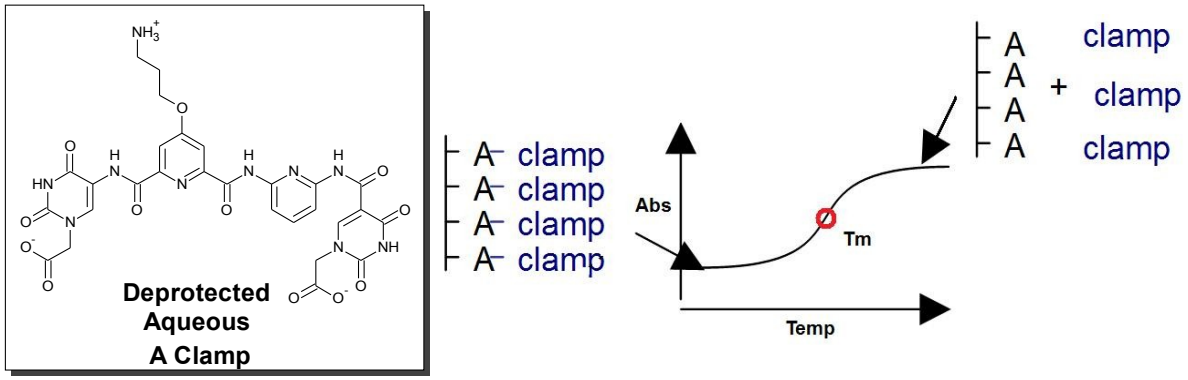


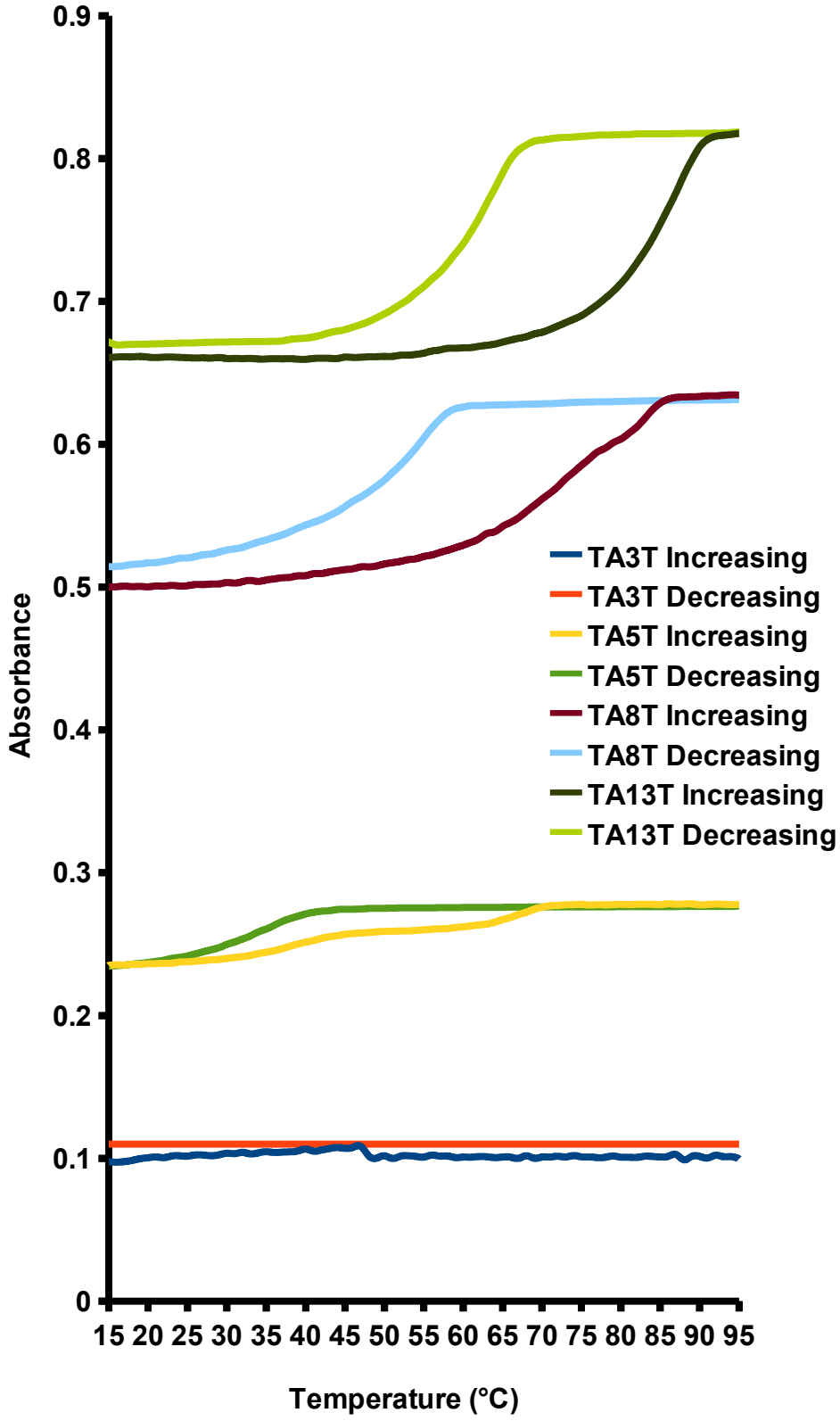
Figure 33: (Left) Structure of deprotected aqueous A clamp. (Right) Diagram depicting an expected UV melting temperature profile going from a poly A:A clamp duplex and unfolding upon heating with a T_m value.

Within the data set a difference between the various oligonucleotide lengths at 260 nm wavelength upon heating from 15 °C to 95 °C can be observed *(Spectra 13)*. Spectra were also evaluated at 320 nm in which there is an amplitude difference between the two different wavelengths but the same T_m value. No change in absorbance was apparent (ΔAbs = 0.006) for template TA_3T (2 μM) with three equivalents of A clamp indicating no significant transition and thus suggesting host-guest binding interaction was poor. However a T_m value of 48 °C was still calculated *(Table 1)*. Template TA_5T (2 μM) with five equivalents of A clamp was not well behaved showing a two-state transition with T_m values at 35 °C and 65 °C. This was perhaps because of characteristics it has between that of template TA_3T and TA_8T oligonucleotides. Template TA_8T (2 μM) with nine equivalents of A clamp showed a melting profile consistent with a cooperative helix-to-coil transition and an observed melting temperature T_m value of 72 °C. Template $TA_{13}T$ (2 μM) with thirteen equivalents of A clamp also showed a behaviour characteristic of duplex unfolding and reorganisation upon melting and cooling respectively. In addition, the $TA_{13}T$:A clamp duplex takes more time to reorganise in solution after unfolding. Little change is observed upon cooling from 95 °C to 65 °C, after which it forms a duplex again.

UV-monitored thermal denaturation experiments indicated cooperative interactions via expected *π-π* stacking upon increasing oligonucleotide length. As a result there is a significant increase in the T_m value. This suggests the A clamp is more proficient at stacking towards longer oligonucleotides with an increased stability. It demonstrates a cooperative self-assembly of unpaired bases in aqueous solution towards a ssDNA. The A clamp was designed with a flat aromatic surface area to enable stacking and the results were indicative of this phenomenon. This is in agreement with and compliments previously discussed CD studies. Although stacking in DNA is a favourable interaction, the thermodynamic data alone does not guarantee a face-to-face stacked geometry for the A clamp. For this reason NMR studies were conducted between the oligonucleotide and A clamp duplex for structural confirmation of stacking. However due to solubility issues the spectra were not well defined.

Oligonucleotide	Equiv A clamp	ΔAbs	Increasing T_m	Decreasing T_m
TA_3T (2 μM)	3	0.006	48 °C	0 °C
TA_5T (2 μM)	5	0.042	35 °C, 65 °C	32 °C
TA_8T (2 μM)	9	0.134	72 °C	46 °C
$TA_{13}T$ (2 μM)	13	0.157	83 °C	61 °C

Table 1: Melting temperatures Tm of various oligonucleotides TA_nT (n = 3, 5, 8, 13) [2 μM] + n equivalents of deprotected A clamp added in sodium cacodylate; 10 mM; pH 7.0. Heating from 15 °C to 95 °C to unfold the duplex and cooling to enable reorganisation of assemblies.



Spectra 13: UV temperature melting spectra of oligonucleotides TA_nT (n = 3, 5, 8, 13) [2 μM] + n equivalents of deprotected A clamp added in sodium cacodylate; 10 mM; pH 7.0. Heating from 15 °C to 95 °C to unfold the duplex and cooling to enable reorganisation of assemblies. a) TA_nT (n = 3) and 3 equivalents of A clamp; (b) TA_nT (n = 5) and 5 equivalents of A clamp; c) TA_nT (n = 8) and 9 equivalents of A clamp; d) TA_nT (n = 13) and 13 equivalents of A clamp.

5.3.2.4 CD and UV melting studies: G clamp and ssDNA

It was desired to also carry out aqueous binding tests involving the aqueous quinoline G clamp derivative **61** and oligonucleotides TG_nT, (n = 3, 4, 5, 6) in water. CD and UV-Vis melting temperature experiments were used to investigate binding between the aqueous G clamp and poly G oligonucleotides. Initially, the ssDNA TG_nT, (n = 3, 4, 5, 6) were tested in conditions which favour G-quadruplex formation in potassium cacodylate buffer and monitored by UV/Vis. A G-quadruplex consists of hydrogen bonding formation between four guanines comprising of a square co-planar array held together by eight Hoogsteen hydrogen bonds and stabilised by a Na^+ or K^+ ion.[95] *(Fig. 34).* The guanine oligomers were annealed by heating to 90 °C in the K^+ buffer, pH 7, 22.5 mM for ten minutes and allowed to cool down slowly overnight. In contrast, triethylamine cacodylate buffer was used to disfavour G-quadruplex formation and UV melting temperature studies were carried out.

All of the TG_nT strands (n = 3, 4, 5) apart from TG_6T showed little change upon heating in TEA cacodylate buffer. Template TG_6T in TEA cacodylate buffer showed a melting profile consistent with a helix-to-coil transition, suggesting unfolding of a G-quadruplex. This is despite the fact that TEA cacodylate buffer should disfavour G-quadruplex formation as it does not contain a potassium or sodium ion. G-quadruplex formation means it is impossible to bind the G clamps to the oligonucleotide and therefore accurate results could not be obtained. ssDNA comes as an ammonium salt, and therefore the slight presence of potassium or sodium could contribute towards G-quadruplex formation. Ammonium buffer could be used instead, but it is volatile at 90 °C, and evidently would affect the T_m results.

Another problem encountered, observed by the UV/Vis of the G clamp **61** (absorbance vs wavelength) showed that upon heating at 75 °C there is degradation of the compound *(Spectra 14)*. LCMS and HPLC analysis confirmed this degradation. The expected mass for G clamp **61** is 487.48 g/mol. However a mass of 286.8 g/mol was detected as the major compound by mass spec analysis after heating, with a difference of 200.68 g/mol. Investigation into the degradation of G clamp **61** consisted of testing the compound stored in various conditions. Followed by HPLC analysis on the compound using a mobile phase of 100% water to 100% acetonitrile with 1% TFA over fourteen minutes on a C18 column. The initial conditions involved storing the G clamp **61** in milliQ water, stored at -20 °C for a week. HPLC analysis showed that the compound was stable and only one peak was observed at 3.2 min. This sample was heated at 95 °C for fifteen minutes to recreate the annealing conditions. HPLC analysis was different to that before heating the G clamp **61**, thus degradation occurred. The next test was to store the G clamp **61** in TEA cacodylate buffer (pH 7.0), at 4 °C for one week.

The final test was to heat this sample in TEA cacodylate buffer. A difference in HPLC spectra before and after heating the sample further confirmed the G clamp **61** degrades upon heating. Therefore, the G clamp **61** is not stable during heating to 75 °C or above and experiments involving heat should be avoided. Experiments involving poly G ssDNA were inconclusive as upon increasing template length, the formation of G-quadruplexes in buffer solution was observed. These factors meant the G clamp **61** and poly G ssDNA were not practical for conducting experiments in aqueous conditions.

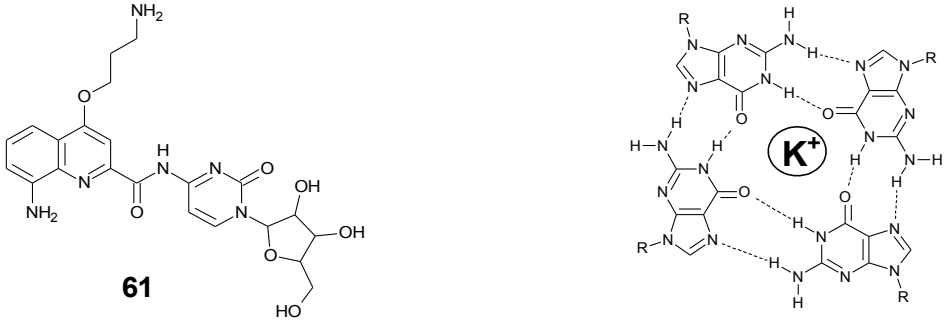
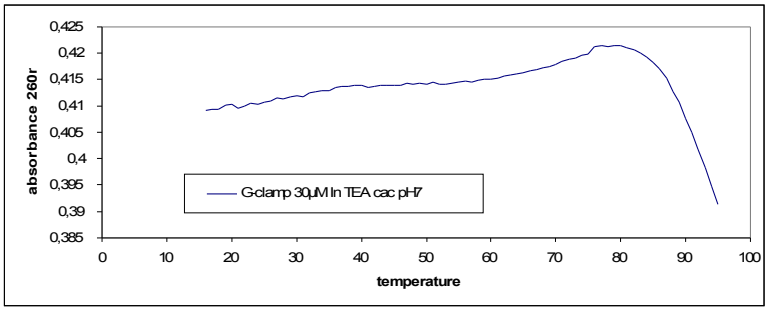


Figure 34: (Left) Water soluble G clamp **61**; (Right) G-Quartet stabilised by K^+. Dotted lines represent the Hoogsteen hydrogen bonding. R represent the connecting loops



Spectra 14: G clamp **61** 30 µM in TEA cacodylate Buffer, pH 7, degradation monitored by UV upon increasing temperature.

6 Conclusions

6.1 A clamp

After some challenges the A clamp was successfully synthesised. It proved to have a high affinity towards 9-ethyl adenine via Watson-Crick and Hoogsteen hydrogen bonding as observed by the significant downfield chemical shift of the imides (> 4 ppm) via ^{1}H NMR titration binding studies in $CDCl_3$/5% DMSO-d_6. The binding constant was found to be in the region of $K_{assoc} \geq 10^6$ M^{-1} in $CDCl_3$ and $K_{assoc} = 30000$ M^{-1} in $CDCl_3$/5% DMSO-d_6. Temperature variant experiments using ^{1}H NMR analysis proved inconclusive in determining if there is a preference for binding out of the two possibilities with 9-ethyl adenine. This could be due to the need for a lower temperature study (example: -53 °C, 220 K) or the exchange is just too fast to measure. It could also suggest that there is indeed only one species and preference for binding. Two-dimensional ^{1}H NMR studies also did not provide further information into the selectivity if any for binding. Co-crystallisation between 9-ethyl adenine and A clamp could be revisited in an attempt to study the complexation via X-ray diffraction. Investigations into binding effectiveness of the aqueous soluble A clamp towards an abasic site in which one purine has been removed from an oligonucleotide duplex in water did not give the desired

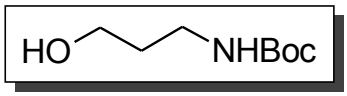
results, thus determining the efficiency of binding one clamp to its corresponding nucleobase proved ineffective. There were two possibilities for this: the adenine nucleobase on the oligonucleotide is inverted outwards causing the duplex to fold in upon itself preventing complexation; or the biotinylated side chain is too long and therefore the A clamp is not positioned in the desired upright position, perhaps residing on the surface and unable to bind with the oligonucleotide. ^{1}H NMR solution studies in 10% D_2O/H_2O of the water soluble A clamp (0.26 mM) and TA_8T also proved inconclusive. Future studies could involve investigations via solid state NMR to investigate the A clamp:ssDNA assembly. The aqueous A clamp proved highly successful for the self-assembly onto poly A ssDNA templates as monitored by CD and UV melting temperature studies in aqueous conditions. Template length was found to be an important factor and duplex stability improved upon increasing length with favourable binding for TA_nT (n = 8, 13). This could be attributed to enhanced by cooperativity and *π-π* stacking. Final experiments have yet to be investigated which would determine a vital question. If a nucleobase mismatch(es) is introduced onto the poly A ssDNA would this enhance stability or disrupt the system? Could the clamps discriminate between bases? There is still scope to evaluate if it is possible to use a mixture of A and G clamps to resolve ssDNA templates for sequence selection by incorporating dynamic combinatorial chemistry. Thus demonstrating the final objective of using ssDNA as a template for information transfer with non-enzymatic replication using non-covalent interactions.

6.2 G Clamps

The quinoline G clamps **46** and **48** were successfully synthesised and complexed with guanosine **47** and analysed by ^{1}H NMR binding studies in $CDCl_3$/5% DMSO-d_6. They were found to have low affinity, (K_{assoc} = 230 M^{-1}) and less than that of the A clamp (K_{assoc} = 30000 M^{-1}) as expected since the A clamp design was specifically designed with two sites of recognition and not just one as is the case with the G clamps. The G clamps utilised primarily Watson-Crick bonding. The A clamp was tailored to also incorporate Hoogsteen hydrogen bonding. Optimisation of the ^{1}H NMR binding study between G clamp **48** and guanosine **47** should be conducted. Co-crystals should be obtained for the G clamp and guanosine complex to study via X-ray crystallography. CD and UV temperature melting studies between the G clamp **46** and various oligonucleotide strands TG_nT strands (n = 3, 4, 5, 6) proved unsuccessful. This was due to two reasons: 1. The degradation of the G clamp **46** upon heating and **2.** The formation of G-quadruplexes within poly G ssDNA templates. However the G clamps could be used for future CD studies with ssDNA which contain a mixture of adenine and guanine nucleobases. It might also be interesting to further investigate the pyridine G clamp synthesis for **49** and **69** as outlined in the synthesis section (Section 4.2, *Scheme 29)*. Overall there is scope to continue this project and investigate the A clamp further, due to its high affinity and recognition towards adenine.

7 Experimental

tert-butyl (3-hydroxypropyl)carbamate **10**

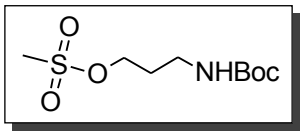


MF: $C_8H_{17}NO_3$

MW: 175.225 g.mol^{-1}

Ref: tcm001, tcm041, tcm083, tcm089, tcm180, tcm2-068
To a solution of 1,3-aminopropanol **9** (3.0 mL, 40 mmol) in anhydrous THF (30 mL) was added Boc-anhydride (10 g, 44 mmol) and stirred for 2 hours at RT under N_2. TLC (EtOAc/Pet. Ether, 1:1) showed a new product had formed R_f = 0.24. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (30 mL) and washed with 5% citric acid solution (30 mL), water and brine (organics extracted x 3). The combined organics were dried (Na_2SO_4), filtered and evaporated to dryness *in vacuo*. No further purification was carried out and the compound **10** was used in the following step as is. Yield is (7.0 g, 99%) of a colourless oil. **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 4.72 (1H, bs); 3.69-3.63 (2H, q); 3.32-3.26 (2H, q); 1.70-1.62 (2H, m); 1.44 (9H, s).

3-[(Boc)amino]propyl methanesulfonate **11**

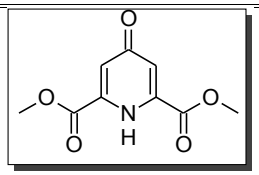


MF: $C_9H_{19}NO_5S$

MW: 253.22 g.mol^{-1}

Ref: [78] tcm003, tcm046, tcm087, tcm091, tcm173, tcm2-059, tcm2-070
To a solution of 1-3 Boc amino propanol **10** (7.0 g, 40 mmol) in anhydrous THF (70 mL) was added anhydrous TEA (17.5 mL, 126 mmol) and stirred under N_2. The solution was cooled to 0 °C and mesylate chloride (6.5 mL, 84 mmol) was added drop wise. The reaction mixture turned yellow and a precipitate formed. The mixture was stirred for 1 hour and allowed to reach RT. TLC (DCM/MeOH, 99:1) showed a new product had formed. Rf: 0.29. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (70 mL) and washed with 5% citric acid solution (70 mL), water and brine (organics extracted x 3). The combined organics were dried (Na_2SO_4), filtered and evaporated to dryness *in vacuo*. No further purification was carried out and the compound **11** was used in the following step as is. Yield is (9.66 g, 95%) of a clear crystalline product.**^{1}H NMR** ($CDCl_3$, 300 MHz): *δ* 4.74 (1H, bs); 4.28 (2H, t); 3.25 (2H, bt); 3.02 (3H, s); 1.97-1.88 (2H, m); 1.43 (9H, s).

Chelidamic Acid Dimethyl Ester 8



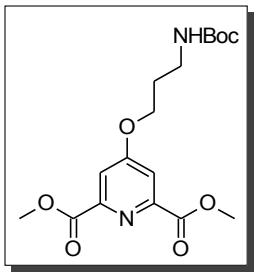
MF: $C_9H_9NO_5$ (HCl Salt)

MW: 211.17 g.mol^{-1} (247.63 salt)

Ref: [77] tcm084, tcm2-061, tcm3-038, tcm3-073

To a suspension of chelidamic acid **7** (40 g, 199 mmol) in absolute MeOH (600 mL) was added 2,2-dimethoxypropane (280 mL) and concentrated HCl (30 mL). The mixture was heated to reflux for 8 hours and then at RT for 3 nights under a $CaCl_2$ drying tube. The volatiles were removed *in vacuo* and the residual triturated with Et_2O (500 mL). The white solid was filtered and rinsed further with Et_2O. No further purification was carried out and the compound **8** was used in the following step as is. Yield is (44 g, 95%) of a white solid. **^{1}H NMR** (DMSO-d_6, 300 MHz): *δ* 7.61 (2H, s); 3.87 (6H, s). **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_9H_{10}NO_5$) 212.0559; found 212.0564

Dimethyl 4-{3-[Boc amino]propoxy}pyridine-2,6-dicarboxylate 12



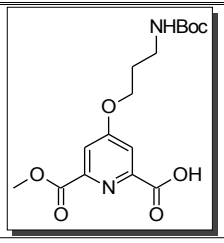
MF: $C_{17}H_{24}N_2O_7$

MW: 368.38 g.mol^{-1}

Ref: [79] tcm004, tcm037, tcm050, tcm092, tcm174, tcm2-060, tcm2-071, tcm3-017

Chelidamic diester **8** and K_2CO_3 was co-evaporated twice with anhydrous toluene before use. To a solution of chelidamic diester **8** (3.0 g, 14 mmol) in anhydrous DMF (50 mL) was added K_2CO_3 (3.92 g, 28 mmol) under N_2. The reaction mixture was heated to 105 ˚C for 2 hours. The reaction mixture was then heated to 60 ˚C and the mesylated propyl amino Boc derivative **11** (3.96 g, 15.6 mmol) was added drop wise in a solution of (30 mL) anhydrous DMF. The reaction mixture was heated to 100 ˚C for a further 2 hours, a precipitate had formed. TLC (EtOAc/Pet. Ether, 1:1) showed a new product had formed. Rf: 0.23. The volatiles were removed *in vacuo,* the residual was diluted with DCM (50 mL) and was washed with 5% citric acid (50 mL), water and brine (organics extracted x 3). The combined organics were dried (Na_2SO_4), filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (EtOAc/Pet. Ether, 8:2) to afford a white solid **12** (2.67 g, 60%). **^{1}H NMR** ($CDCl_3$, 300 MHz): *δ* 7.74 (2H, s); 4.77 (1H, bs); 4.15 (2H, t); 3.95 (6H, s); 3.32-3.25 (2H, q); 2.00 (2H, m); 1.37 (9H, s).**^{13}C NMR** ($CDCl_3$, 75 MHz): *δ* 166.8, 165.0, 155.9, 149.8, 114.4, 114.0, 66.5, 63.8, 59.2, 55.9, 53.1, 37.4, 32.9, 29.3, 28.3. **HRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{17}H_{24}N_2O_7Na$) 391.1481; found 391.1476.

4-{3-[Boc amino]propoxy}-6-(methoxycarbonyl)pyridine-2-carboxylic acid **2**



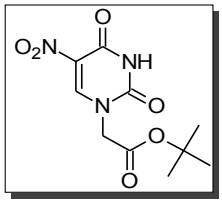
MF: $C_{16}H_{22}N_2O_7$

MW: 354.35 g.mol^{-1}

Ref: tcm006, tcm025, tcm040, tcm066, tcm096, tcm142, tcm183, tcm2-076, tcm3-029, tcm3-031

To a solution of the diester derivative **12** (3.11 g, 8.4 mmol) in MeOH (30 mL) and 1,4-dioxane (30 mL) was added NaOH (337 mg, 8.4 mmol) and stirred at RT for 1 hour. The volatiles were removed *in vacuo* and the crude mixture diluted with Et_2O (30 mL). The solid was filtered to remove starting material. The solid was diluted with DCM (60 mL) and washed with 5% citric acid (60 mL), water and brine (organics extracted x 3). The combined organics were dried (Na_2SO_4), filtered and evaporated to dryness *in vacuo*. The crude product was purified by trituration with cyclohexane, filtered and evaporated to dryness *in vacuo,* to afford a clear crystalline solid **2** (2.92 g, 91%). **^{1}H NMR** ($CDCl_3$, 300 MHz): *δ* 7.84-7.80 (2H, d, *J* = 11.14 Hz); 4.69 (1H, bs); 4.21 (2H, t); 3.99 (3H, s); 3.36-3.30 (2H, q); 2.05 (2H, m); 1.43 (9H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): *δ* 169.0, 166.6, 165.2, 156.0, 154.7, 146.6, 113.1, 112.6, 66.4, 53.1, 37.0, 29.0, 28.3.

***tert*-butyl (5-nitro-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetate** **6**



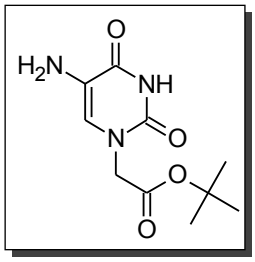
MF: $C_{10}H_{13}N_3O_6$

MW: 271.22 g.mol^{-1}

Ref: [76] tcm189, tcm2-028

To a solution of 5-nitro uracil **5** (15 g, 95 mmol) in anhydrous DMF (200 mL) was added anhydrous TEA (14.5 mL, 105 mmol) under N_2 and stirred for 20 minutes at RT. The solution was cooled to 0 °C and *t*-butyl bromoacetate (15.5 mL, 105 mmol) was added drop wise. The reaction mixture was allowed to reach RT and stirred for a further 4 hours. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (200 mL) and washed with 5% citric acid solution (200 mL), water and brine (organics extracted x 3). The combined organics were dried (Na_2SO_4), filtered and evaporated to dryness *in vacuo*. The crude product was purified by trituration with Et_2O, filtered and evaporated to dryness *in vacuo,* to afford a white solid **6** (24.5 g, 91%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 9.32 (1H, s); 4.58 (2H, s); 1.42 (9H, s). **^{13}C NMR** ($CDCl_3$/DMSO-d_6, 75 MHz): δ 165.0, 154.2, 149.6, 148.2, 124.4, 82.2, 49.0, 26.9. **LRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{10}H_{13}N_3O_6Na$) 294.07, found 294.09.**HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{10}H_{13}N_3O_6Na$) 294.0702, found 294.0715

t-butyl (5-amino-2,4-dioxo-3,4-dihydropyrimidin-1(*2H*)-yl)acetate 1



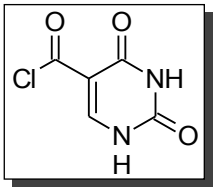
MF: $C_{10}H_{15}N_3O_4$

MW: 241.24 g.mol^{-1}

Ref: tcm156, tcm167, tcm2-30, tcm2-064

To a solution of 5-nitro uracil *t*-butyl acetate (7.8 g, 28.7 mmol) was added MeOH (125 mL) and EtOAc (125 mL) followed by 10% Pd/C (1.6 g) under N_2. The reaction was purged with N_2 three times, then H_2 and stirred at RT for 16 hours. The reaction was filtered on a bed of celite and washed with EtOAc (125 mL x 3). The filtrate evaporated to dryness *in vacuo,* to afford a crude product. The solid was purified by trituration with Et_2O, filtered and evaporated to dryness *in vacuo,* to afford a white solid **1** (6.79 g, 97%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 9.32 (1H, s); 4.58 (2H, s); 1.42 (9H, s). **^{13}C NMR** ($CDCl_3$/DMSO-d_6, 75 MHz): δ 165.0, 154.2, 149.6, 148.2, 124.4, 82.2, 49.0, 26.9. **HRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{10}H_{15}N_3O_4Na$) 264.0960, found 264.1024.

2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carbonyl chloride 29



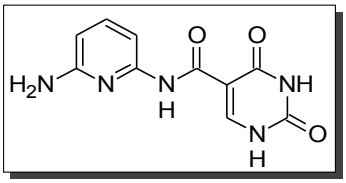
MF: $C_5H_3ClN_2O_3$

MW: 174.54 g.mol^{-1}

Ref: tcm152, tcm2-002, tcm2-039

To 2,4-dihydroxypyrimidine-5-carboxylic acid **14** (8.0 g, 51 mmol) was added $SOCl_2$ (80 mL) and stirred at 90 °C for 3 nights. A colour change from white solid to a pale yellow solid was observed. The volatiles were removed *in vacuo* to afford a pale yellow solid **29** (7.8 g, 99%). No further purification was carried out and the compound **29** was used in the following step as is. **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 13.32 (1H, s); 12.09 (1H, bd); 11.96 (1H, s); 8.24 (1H, d *J* = 6.23 Hz).

N-(6-aminopyridin-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide



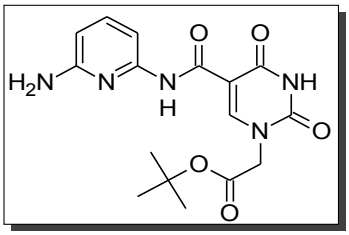
MF: $C_{10}H_9N_5O_3$

MW: 247.21 g.mol^{-1}

Ref: tcm153, tcm2-005, tcm2-040

To a solution of 2,6-diaminopyridine **3** (1.2 g, 10.8 mmol) in anhydrous DCM (20 mL) was added anhydrous TEA (3.0 mL, 21.7 mmol) and stirred for 20 mins under N_2. In a separate flask, contained the iso-orotic acid chloride **29** (1.5 g, 8.59 mmol) in anhydrous DCM (20 mL) and cooled to 0 °C. The 2,6-diaminopyridine **3** solution was cannulated onto the acid chloride mixture. The yellow reaction mixture was allowed to reach RT and stirred overnight. The volatiles were removed *in vacuo* and the solid filtered to afford a white solid **28** (1.77 g, 84%). If starting acid and bis-substituted compound is present as an impurity the following purification method was carried out in addition to the above. The crude solid is dissolved in 1 M HCl (20 mL) and the insoluble solid filtered off (bis-substituted adduct). The acidic aqueous solution is then neutralised with NaOH until a pH 7 is reached. The white precipitate is filtered and evaporated to dryness *in vacuo,* to afford a white solid **28** (34% yield if bis-substituted adduct is present and requires removal). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 11.09 (1H, s); 8.35 (1H, s); 7.34 (2H, m); 6.12 (1H, d, *J* = 7.36 Hz); 5.84 (2H, s). **^{13}C NMR** (DMSO-d_6, 75 MHz): δ 164.6, 160.7, 158.7, 152.1, 151.5, 149.8, 138.8, 103.6, 102.9, 100.7. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{10}H_{10}N_5O_3$) 248.0784, found 248.0775.

t-butyl {5-[(6-aminopyridin-2-yl)carbamoyl]-2,4-dioxo-3,4-dihydropyrimidin-1 (*2H)* -yl}acetate **40**

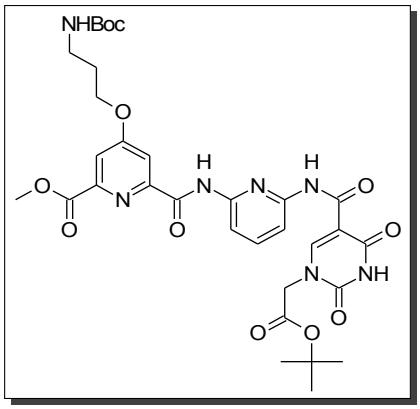


MF: $C_{16}H_{19}N_5O_5$

MW: 361.35 g.mol^{-1}

Ref: tcm2-038, tcm2-044, tcm3-018
To a solution of amino pyridine iso-orotic derivative **28** (2.5 g, 10.0 mmol) in anhydrous DMF (25 mL) was added anhydrous TEA (2.0 mL, 14.0 mmol) and stirred at RT for 10 mins under N_2. The solution was cooled to 0 °C and *t*-butyl bromoacetate (1.8 mL, 12.0 mmol) was added drop wise. The reaction mixture was allowed to reach RT and stirred overnight. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (50 mL) and washed with 5% citric acid solution (200 mL), water and brine (organics extracted x 3). The combined organics were directly co-evaporated with toluene to dryness *in vacuo*. The crude product was purified by trituration with Et_2O, filtered and evaporated to dryness *in vacuo,* to afford a pale yellow solid **40** (2.35 g, 65%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 12.21 (1H, s); 10.85 (1H, s); 8.72 (1H, s); 7.36 (2H, m); 6.19 (1H, d, *J* = 7.36 Hz); 5.98 (2H, s); 4.62 (2H, s); 1.42 (9H, s). **^{13}C NMR** (DMSO-d_6, 75 MHz): δ 167.2, 164.3, 160.2, 159.3, 153.1, 150.3, 150.0, 139.5, 104.8, 104.5, 101.4, 82.7, 50.5, 28.1. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{16}H_{20}N_5O_5$) 362.1464, found 362.1464.

***t*-butyl {5-[(6-aminopyridin-2-yl)carbamoyl]-2,4-dioxo-3,4-dihydropyrimidin-1 (*2H)*-yl}acetate 4-{3-Boc-aminopropoxy}-6-pyridine ester 42**



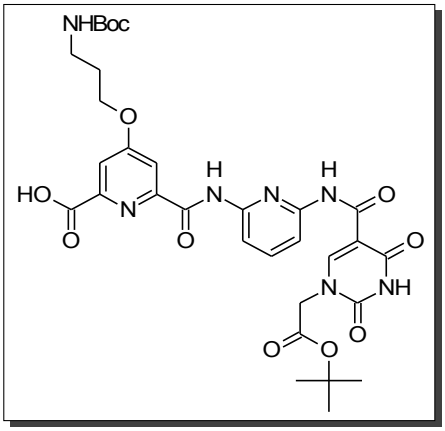
MF: $C_{32}H_{39}N_7O_{11}$

MW: 697.69 g.mol^{-1}

Ref: tcm2-045 x, tcm2-047, tcm2-049, tcm3-019, tcm3-034, tcm3-041

All starting materials were co-evaporated with anhydrous toluene (x 3) and dried in vacuo before use. To a solution of 4-{3-Boc aminopropoxy}-6-pyridine ester-2-carboxylic acid **2** (1.2 g, 3.4 mmol) in anhydrous DMF (6 mL) was added PyBOP (2.56 g, 4.9 mmol) and anhydrous TEA (0.5 mL, 3.7 mmol) and stirred at RT for 15 mins under Ar. The reaction mixture was cannulated into a separate flask containing the amino pyridine iso-orotic derivative **40** (0.9 g, 2.4 mmol), anhydrous pyridine (2 mL) and anhydrous DMF (3 mL). The reaction was stirred at 60 °C overnight. The solution was cooled to RT and a precipitate (product) was filtered off. The crude product was purified by trituration with Et_2O, filtered and evaporated to dryness in vacuo, to afford a white solid **42** (0.95 g, 60%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 12.30 (1H, bs); 11.27 (1H, s); 10.15 (1H, s); 8.79 (1H, s); 7.99 (2H, dd, *J* = 3.21 Hz); 7.89 (1H, t); 7.81 (1H, d, J = 2.45 Hz); 7.72 (1H, d, J = 2.45 Hz); 6.93 (1H, t); 4.65 (2H, s); 4.25 (2H, t); 3.97 (3H, s); 3.11 (2H, q); 1.88 (2H, m); 1.43 (9H, s); 1.36 (9H, s). **^{13}C NMR** (DMSO-d_6/CD_3OD, 75 MHz): δ 167.7, 167.0, 164.6, 164.1, 161.4, 160.6, 156.2, 153.3, 151.0, 150.2, 149.3, 148.5, 141.2, 114.7, 111.5, 110.2, 109.2, 104.4, 82.6, 78.0, 67.1, 53.1, 50.5, 36.9, 29.1, 28.4, 27.8. **LRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{32}H_{39}N_7O_{11}Na$) 720.6, found 720.4

t-butyl {5-[(6-aminopyridin-2-yl)carbamoyl]-2,4-dioxo-3,4-dihydropyrimidin-1 (*2H)*-yl}acetate 4-{3-Boc-aminopropoxy}-6-pyridine carboxylic acid 43



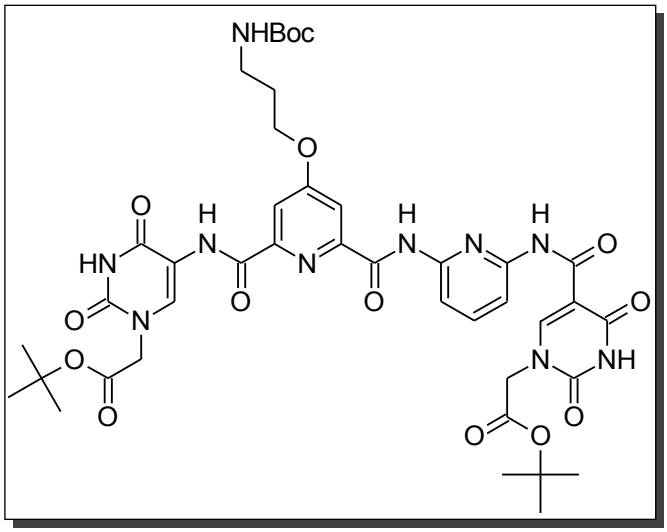
MF: $C_{31}H_{37}N_7O_{11}$

MW: 683.66 g.mol^{-1}

Ref: tcm3-088, tcm3-097

To a solution of the A clamp ester derivative **42** (260 mg, 0.37 mmol) in H_2O (6.5 mL) and DMSO (25 mL) was added KOH (62 mg, 1.10 mmol) and stirred at 60 °C for 1 hour. The reaction was allowed to reach RT and stirred overnight. Acetic acid (5 equiv) was added and the volatiles were removed *in vacuo*. Some DMSO remained and the crude mixture was purified by trituration with a combination of $H_2O/Et_2O/MeOH/CHCl_3$. The precipitated solid was filtered and evaporated to dryness *in vacuo,* to afford a white solid **43** (171 mg, 67%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 12.23 (1H, s); 11.17 (2H, s); 8.76 (1H, s); 7.98 (2H, d, *J* = 6.61 Hz); 7.87 (1H, t); 7.78 (1H, d *J* = 1.51 Hz); 7.70 (1H, d *J* = 1.69 Hz); 6.89 (1H, t); 4.61 (1H, s); 4.21 (2H, t); 3.07 (2H, m); 1.84 (2H, m); 1.39 (9H, s); 1.32 (9H, s). **^{13}C NMR** (DMSO-d_6, 75 MHz): δ 167.3, 166.6, 164.7, 163.7, 161.6, 160.3, 155.5, 152.9, 150.2, 149.7, 148.5, 140.7, 113.2, 111.6, 109.8, 103.9, 82.1, 77.5, 66.6, 50.0, 40.3, 36.5, 28.6, 28.1, 27.5. **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{31}H_{38}N_7O_{11}$) 684.26, found 684.27. **HRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{31}H_{37}N_7O_{11}Na$) 706.2449, found 706.2469.

A clamp



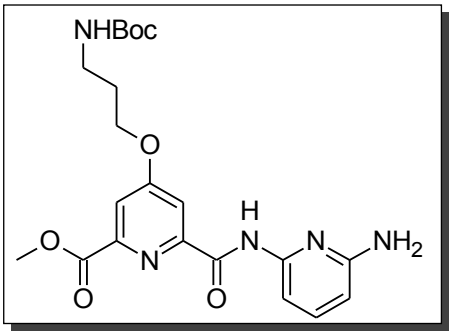
MF: $C_{41}H_{50}N_{10}O_{14}$

MW: 906.89 g.mol^{-1}

Ref: tcm3-096, tcm3-110, tcm3-112

All starting materials were co-evaporated with anhydrous toluene (x 3) and dried *in vacuo* before use. To a solution of carboxylic acid **43** (85 mg, 0.12 mmol) in anhydrous DMF (5 mL) was added PyBOP (161 mg, 0.3 mmol) and anhydrous TEA (0.04 mL, 0.3 mmol) and stirred at RT for 15 mins under Ar. The reaction mixture was cannulated into a separate flask containing the 5-amino uracil derivative **1** (120 mg, 0.5 mmol), anhydrous pyridine (2 mL) and anhydrous DMF (5 mL). The reaction was stirred at RT for 3 nights. TLC ($CHCl_3$/MeOH, 9:1) showed a new product had formed, white under UV 365 nm. Rf: 0.62. The volatiles were removed *in vacuo,* the residual diluted with $CHCl_3$ (50 mL) and washed with 5% citric acid solution (20 mL), water and brine (organics extracted x 3). The combined organics were directly co-evaporated with toluene to dryness *in vacuo*. The crude product was purified by column chromatography ($CHCl_3$/MeOH, 99:1) to afford an off white solid **A clamp** (45 mg, 40%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 12.24 (1H, bs); 11.86 (1H, bs); 11.32 (2H, s); 10.63 (1H, s); 8.78 (1H, s); 8.16 (1H, s); 8.07-7.90 (3H, m); 7.84 (1H, dd, *J* = 2.45 Hz); 7.74 (1H, dd, *J* = 2.26 Hz); 6.94 (1H, t); 4.64 (2H, s); 4.49 (2H, s); 4.26 (2H, t); 3.11 (2H, q); 1.89 (2H, m), 1.43 (18H, d *J* = 2.83 Hz); 1.36 (9H, s).**^{13}C NMR** (DMSO-d_6, 75 MHz): δ 167.7, 167.6, 167.2, 164.2, 163.2, 162.9, 161.5, 160.9, 156.1, 153.5, 151.2, 151.0, 150.5, 150.4, 150.3, 150.2, 141.7, 141.0, 112.0, 111.6, 110.6, 104.5, 82.7, 82.4, 78.0, 67.1, 50.0, 49.1, 46.3, 46.2, 37.1, 29.3, 28.7, 28.2, 28.1. **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{41}H_{51}N_{10}O_{14}$) 907.9, found 907.4. **HRMS (ESI):** *m/z* calcd $[M+Na]^+$ ($C_{41}H_{50}N_{10}O_{14}Na$) 929.3406, found 929.3751.

N-(6-aminopyridin-2-yl)-4-{3-Boc-aminopropoxy}-6-pyridine ester **13**



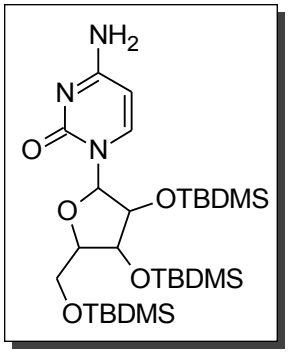
MF: $C_{21}H_{27}N_5O_6$

MW: 445.46 $g.mol^{-1}$

Ref: tcm026, tcm030, tcm039, tcm115,

To a solution of 4-{3-Boc aminopropoxy}-6-pyridine ester-2-carboxylic acid **2** (422 mg, 1.19 mmol) in anhydrous THF (10 mL) was added PyBOP (620 mg, 1.19 mmol) and anhydrous DIEA (0.4 mL, 2.38 mmol) and stirred at RT for 30 mins under N_2. The reaction mixture was cannulated into a separate flask containing the 2,6-diaminopyridine **3** (900 mg, 2.4 mmol) and anhydrous THF (60 mL). The reaction was stirred at RT overnight. The volatiles were removed *in vacuo,* the residual diluted with DCM (70 mL) and washed with 5% citric acid solution (70 mL), water and brine (organics extracted x 3). The combined organics were dried ($MgSO_4$), filtered and evaporated to dryness *in vacuo*. The crude product was purified by trituration with Et_2O, filtered and evaporated to dryness *in vacuo,* to afford a white solid **13** (264 mg, 50%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 9.92 (1H, s); 7.84-7.74 (2H, d, *J* = 25.87 Hz); 7.49-7.39 (2H, m); 6.93 (1H, t); 6.27 (1H, d, *J* = 7.55 Hz); 6.12 (2H, s); 4.26 (2H, t); 3.95 (3H, s); 3.11 (2H, m); 1.88 (2H, m); 1.36 (9H, s) **^{13}C NMR** (75 MHz, DMSO-d_6): δ 167.3, 164.1, 160.4, 158.9, 150.8, 148.5, 147.9, 139.2, 114.3, 110.9, 104.5, 99.9, 77.5, 66.7, 52.8, 45.7, 36.4, 28.6, 28.1. **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{21}H_{28}N_5O_6$) 446.4, found 446.0. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{21}H_{28}N_5O_6$) 446.2040, found 446.2034

TBDMS-cytidine **52**



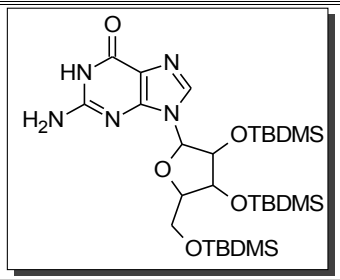
MF: $C_{27}H_{55}N_3O_5Si_3$

MW: 585.99 g.mol^{-1}

Ref:[32] tcm053

Cytidine (500 mg, 2.0 mmol) was co-evaporated twice with anhydrous toluene (10 mL) and suspended in anhydrous THF (15 mL). To this was added imidazole (700 mg, 10.3 mmol) and then TBDMS chloride (1.7 g, 11 mmol). The reaction was stirred at RT under N_2 for two nights. TLC (EtOAc/MeOH 9:1) showed a new product had formed Rf: 0.51. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (80 mL) and washed with 10% $NaHCO_3$ solution (70 mL), water and brine (organics extracted x 3). The combined organics were dried ($MgSO_4$), filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (EtOAc/MeOH, 99:1) to afford a white solid **52** (1.11 g, 92%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 8.19 (1H, d, *J* = 7.36 Hz); 5.79 (1H, s); 5.57 (1H, d, *J* = 7.17 Hz); 4.09-4.01 (4H, m); 3.76 (1H, d, *J* = 10.77 Hz); 0.95 (9H, s); 0.90 (9H, s); 0.88 (9H, s); 0.23 (3H, s); 0.13 (3H, s); 0.11 (6H, s) 0.04 (6H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 165.7, 155.8, 141.4, 93.6, 90.1, 82.8, 76.2, 69.4, 61.0, 26.0, 25.8, 18.5, 18.0, 18.0, -4.1, -5.0, -5.0, -5.3, -5.5. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{27}H_{56}N_3O_5Si_3$) 586.3528, found 586.3527.

TBDMS-guanosine 47



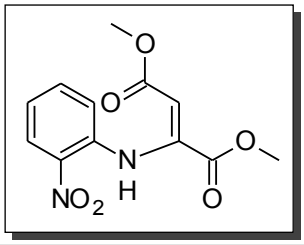
MF: $C_{28}H_{55}N_5O_5Si_3$

MW: 626.02 g.mol^{-1}

Ref:[32,84] tcm059

Guanosine (2.0 g, 7.1 mmol) was co-evaporated twice with anhydrous toluene (20 mL) and suspended in anhydrous THF (25 mL). To this was added imidazole (2.4 g, 39 mmol) and TBDMS chloride (7.2 mL, 35 mmol). The reaction is stirred at RT under N_2 overnight. The volatiles were removed *in vacuo,* the residual diluted with EtOAc (100 mL) and washed with 10% $NaHCO_3$ solution (50 mL), water and brine (organics extracted x 3). The combined organics were dried ($MgSO_4$), filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (Pet. Ether/EtOAc 9:1) gave a white solid **47** (3.8 g, 86%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 12.03 (1H, bs); 7.89 (1H, s); 6.22 (1H, bs); 5.81 (1H, bs); 4.42 (1H, bs); 4.27 (1H, bs); 4.09 (1H, bs); 3.99 (1H, d, *J* = 11.52 Hz); 3.78 (1H, d, *J* = 9.44 Hz); 0.95 (9H, s); 0.92 (9H, s); 0.86 (9H, s); 0.13 (6H, s); 0.09 (6H, s); 0.02 (3H, s)-0.04 (3H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 159.1, 153.4, 151.4, 88.2, 84.5, 76.1, 71.1, 62.0, 29.6, 26.1, 25.8, 25.7, 18.5, 18.0, 17.9, -4.2, -4.6, -4.8, -5.3.

Dimethyl-2-[(2-nitrophenyl)amino]fumarate 56



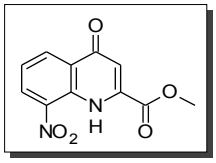
MF: $C_{12}H_{12}N_2O_6$

MW: 280.23 g.mol^{-1}

Ref: tcm069

To a solution of 2-nitroaniline (60 g, 434 mmol) in (650 mL) of methanol was added dimethylbut-2-ynedioate (53 mL, 434 mmol). The resulting mixture was heated at reflux for 24 hr. The solution was then cooled and the solvent was evaporated up to ca. 350 mL. The flask was then placed at -18 ˚C overnight for precipitation. The resulting yellow crystalline solid was collected by filtration, and washed 4 times with cold methanol and dried under reduced pressure to yield a yellow solid **56** (92.2 g, 76%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 11.11 (1H, bs); 8.14 (1H, dd, *J* = 1.4, 8.3 Hz); 7.46 (1H, t, *J* = 8.8Hz); 7.08 (1H, t, *J* = 8.4 Hz); 6.78 (1H, dd, *J* = 1.3, 8.4 Hz); 5.04 (1H, s); 3.81 (3H, s); 3.76 (3H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 167.9, 164.3, 143.4, 136.8, 134.1, 126.1, 121.4, 120.3, 102.9, 52.9, 51.7s.

Methyl 8-nitro-4-oxo-1,4-dihydroquinoline-2-carboxylate 57



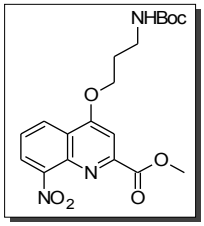
MF: $C_{11}H_8N_2O_5$

MW: 248.19 g.mol^{-1}

Ref:[96] tcm071

To polyphosphoric acid (PPA) (421 g) was added 2-(2-nitrophenylamino)fumarate **56** (83.4 g, 297 mmol) and the reaction heated to 90 °C for 3 hours. The solution was cooled down and poured slowly into a bucket containing 1.5 L sat. $NaHCO_3$ at 0 °C for neutralisation. The resulting slurry was filtered using a sintered glass funnel and washed several times with distilled water. The obtained brown solid was triturated in 70 mL MeOH and stirred for 1 hour at RT. The flask was subsequently placed at –20 °C for precipitation, which was subsequently filtered, washed with cold methanol and dried under reduced pressure to yield a yellow solid **57** (28.5 g, 38%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 11.8 (1H, bs); 8.74 (2H, m); 7.47 (1H, t); 7.08 (1H, s); 4.09 (3H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 170.2, 146.9, 145.4, 143.2, 139.1, 136.2, 130.8, 121.4, 62.1. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{11}H_9N_2O_5$) 249.0511, found 249.0506.

Methyl 4-propyl amino boc-8-nitroquinoline-2-carboxylate 58



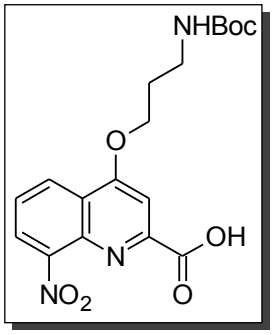
MF: $C_{19}H_{23}N_3O_7$

MW: 405.40 g.mol^{-1}

Ref:[97] tcm049, tcm169, tcm181, tcm2-020

To methyl-1,4-dihydro-8-nitro-4-oxoquinoline-2-carboxylate **57** (5.0 g, 20.15 mmol) was added 1,3-Boc aminopropanol **10** (3.9 g, 22.16 mmol) and triphenylphosphine (7.9 g, 30.23 mmol) in anhydrous THF (250 mL) under N_2. The reaction mixture was cooled to 0 °C and diisopropyl azodicarboxylate (6 mL, 30.23 mmol) was added and stirred for 30 mins. The reaction mixture was allowed to reach RT and stirred overnight. The volatiles were removed *in vacuo* and the residual was purified by recrystallisation (MeOH) to afford an off yellow solid **58** (5.46 g, 67%). Starting material **57** was recovered (3.12 g). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 8.48 (1H, dd, *J* = 1.3, 8.5 Hz); 8.13 (1H, dd, *J* = 1.3, 7.5 Hz); 7.78-7.63 (2H, m); 4.81-4.67 (1H, bs); 4.41 (2H, t, *J* = 6.3 Hz); 4.05 (3H, s); 2.34 (2H, q, *J* = 6.3 Hz); 2.21 (2H, quintet, *J* = 6.3 Hz); 1.45 (9 H, s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 165.7, 162.6, 156.1, 151.4, 148.5, 140.1, 126.5, 126.2, 125.3, 123.3, 102.3, 79.7, 67.4, 53.5, 37.8, 29.5, 28.5. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{19}H_{23}N_3O_7Na$) 428.1434, found 428.1447.

4-{3-[Boc amino]propoxy} 8-nitroquinoline-2-carboxylic acid 59



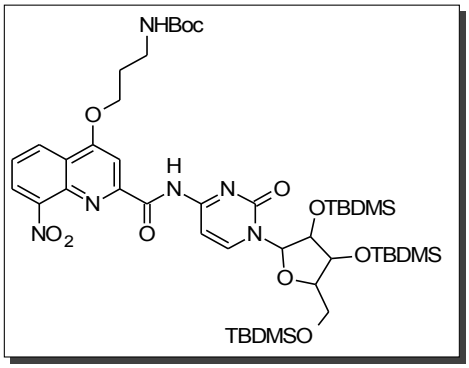
MF: $C_{18}H_{21}N_3O_7$

MW: 391.37 g.mol^{-1}

Ref:[97] tcm062, tcm086, tcm088, tcm093, tcm145, tcm178, tcm182, tcm190, tcm192

To the quinoline ester derivative **58** (0.8 g, 1.97 mmol) was added 1:1 mixture of (MeOH/THF, 24, 48 mL) and cooled to 0 °C. NaOH (173 mg, 4.34 mmol) added and reaction stirred at RT overnight. The volatiles were removed *in vacuo* and the residual was dissolved in DCM and washed with 5% citric acid. The combined organics were dried ($MgSO_4$), filtered and evaporated to dryness *in vacuo*. The product **59** (671 mg, 87%) was not purified further and used as is in the following step. Note this method was not reproducible due to gel formation. A new method using LiOH has been validated since within the group **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 11.85-10.05 (1H, bs); 8.56 (1H, dd, *J* = 1.1, 8.5 Hz); 8.25 (1H, dd , *J* = 1.3, 7.5 Hz); 7.79-7.72 (2H, m); 4.78-4.63 (1H, bs); 4.45 (2H, t, *J* = 6.3 Hz); 3.46 (2H, q, *J* = 6.3 Hz); 2.23 (2H, quintet, *J* = 6.2 Hz); 1.45 (9H, s).

4-{3-[Boc-amino]propoxy-8-nitroquinoline-2-TBDMS cytidine 60



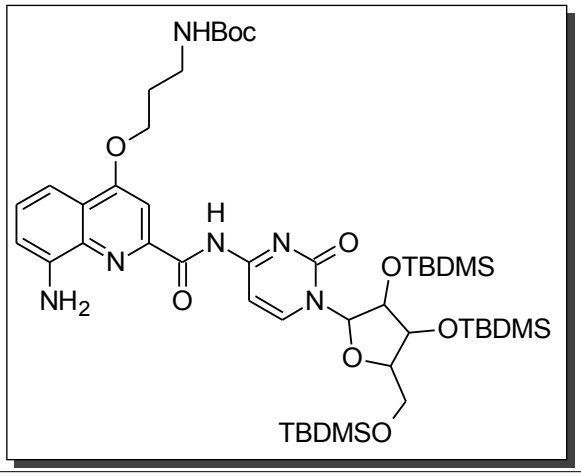
MF: $C_{45}H_{74}N_6O_{11}Si_3$

MW: 959.35 g.mol^{-1}

Ref: tcm090, tcm097, tcm2-058

To the quinoline acid derivative **59** (160 mg, 0.408 mmol) in anhydrous DMF (2 mL) was added PyBOP (318 mg, 0.612 mmol), and TEA (0.23 mL, 1.632 mmol) and the reaction mixture stirred at RT for 30 mins. To this was added TBDMS cytidine **52** (263 mg, 0.45 mmol) and the reaction mixture was stirred under N_2 at RT overnight. The volatiles were removed *in vacuo* and the residual was dissolved in DCM and washed with 5% citric acid, brine and the combined organics were dried ($MgSO_4$), filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (Pet. Ether/EtOAc 7:3) and this gave a yellow crystalline solid **60** (392 mg, 50%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 10.41 (1H, s); 8.67 (1H, d, *J* = 7.36 Hz); 8.48 (1H, d, *J* = 9.44 Hz); 8.13 (1H, d, *J* = 7.36 Hz); 7.81 (1H, s); 7.68 (1H, t); 7.54 (1H, d, *J* = 7.55 Hz); 5.80 (1H, s); 4.79 (1H, bs); 4.42 (2H, t); 4.21-4.04 (4H, m); 3.81 (1H, d, *J* = 10.95); 3.43 (2H, q); 2.20 (2H, quintet); 1.43 (9H,s); 0.99 (9H,s); 0.92 (9H,s); 0.87 (9H,s); 0.29 (3H,s); 0.18 (3H,s); 0.16 (6H,s); 0.04 (6H,s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 163.3, 163.1, 161.4, 155.9, 155.3, 151.3, 147.8, 145.2, 138.8, 126.3, 125.4, 123.3, 99.7, 94.4, 91.2, 82.8, 76.1, 68.5, 67.4, 60.4, 37.5, 29.6, 29.3, 28.3, 26.1, 25.9, 25.8, 18.5, 18.1, 18.0, -4.0, -4.1, -5.0, -5.2, -5.2, -5.5. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{45}H_{75}N_6O_{11}Si_3$) 959.4802, found 959.4791.

4-{3-[Boc-amino]propoxy}-8-aminoquinoline-2-TBDMS cytidine **46**



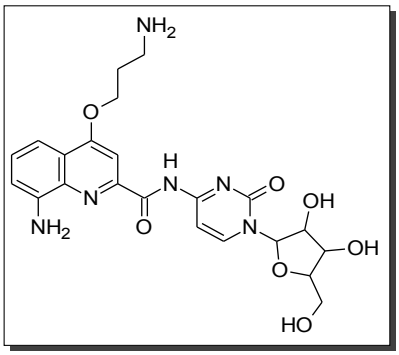
MF: $C_{45}H_{76}N_6O_9Si_3$

MW: 929.37 g.mol^{-1}

Ref:[38] tcm105, tcm2-061

To the nitro-quinoline TBDMS cytidine derivative **60** (195 mg, 0.203 mmol) in a mixture of MeOH/EtOAc (6 mL) was added Pd/C 10% (20 mg), under N_2 and flushed 3 times. Hydrogen gas was introduced and the reaction stirred at RT for 7 hours. The reaction mixture was filtered on a bed of celite and rinsed with EtOAc (100 mL). The volatiles were removed *in vacuo* to give an amber solid **46** (179 mg, 95%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 10.58 (1H, s); 8.65 (1H, d, *J* = 7.55 Hz); 7.64 (1H, s); 7.60 (1H, d, *J* = 7.36 Hz); 7.50 (1H, d, *J* = 7.93 Hz); 7.40 (1H, t); 6.97 (1H, d, *J* = 7.55 Hz); 5.86 (1H, s); 5.04 (2H, bs); 4.76 (1H, bs); 4.35 (2H, t); 4.19-4.06 (4H, m); 3.82 (1H, d, *J* = 11.33); 3.42 (2H, q); 2.16 (2H, quintet); 1.44 (9H,s); 1.01 (9H,s); 0.93 (9H,s); 0.89 (9H,s); 0.27 (3H,s); 0.19 (3H,s); 0.17 (3H,s); 0.14 (3H,s); 0.05 (6H,s). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ 164.0, 163.0, 161.5, 155.9, 155.3, 146.1, 145.2, 144.4, 136.9, 129.1, 123.1, 111.5, 109.5, 98.1, 95.4, 90.9, 83.0, 76.3, 68.8, 66.8, 60.6, 37.8, 29.6, 29.1, 28.3, 26.0, 25.8, 25.8, 18.5, 18.0, 0.97 **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{45}H_{77}N_6O_9Si_3$) 929.5060, found 929.5050.

4-{3-amino propoxy}-8-aminoquinoline-2 cytidine 61



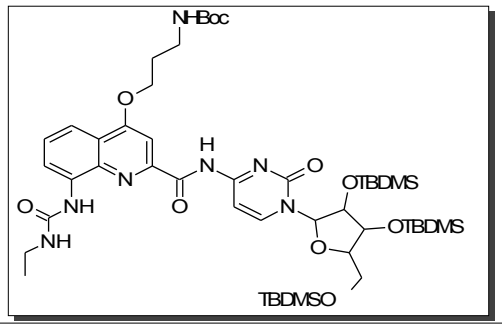
MF: $C_{22}H_{26}N_6O_7$

MW: 486.47 g.mol^{-1}

Ref: tcm2-069

To the amino-quinoline TBDMS cytidine derivative **46** (44 mg, 0.053 mmol) in MeOH (5 mL) was added conc. HCl (2 mL) and the reaction was stirred at RT overnight. The volatiles were removed *in vacuo* and the residual was co-evaporated with toluene. The residual was purified by trituration with Pet. Ether to give the expected product **61** (22 mg, 85%). The product was purified further by Prep. HPLC. **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{22}H_{27}N_6O_7$) 487.48, found 487.19. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{22}H_{27}N_6O_7$) 487.1941, found 487.1933.

8-[(ethylcarbamoyl)amino]-4-{3-[Boc-amino]propoxy}-quinoline-2-TBDMS cytidine 64

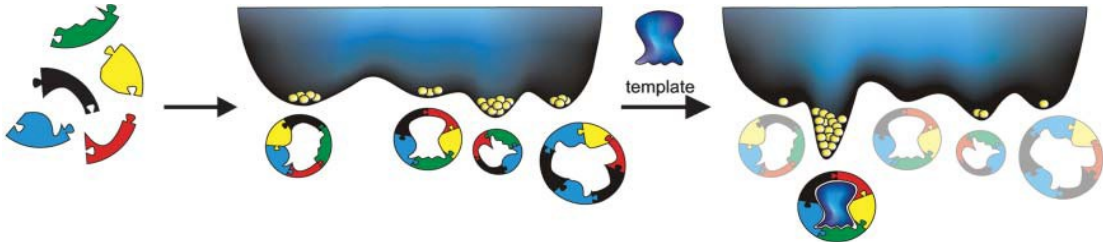


MF: $C_{48}H_{81}N_7O_{10}Si_3$

MW: 1000.45 g.mol^{-1}

Ref: tcm116

To the amino-quinoline TBDMS cytidine derivative **46** (30 mg, 0.032 mmol) in anhydrous THF (1 mL) under N_2 was added ethyl isocyanate (0.1 mL, 1.26 mmol) and the reaction was stirred at RT overnight. The volatiles were removed *in vacuo* and the residual was purified by column chromatography (Pet. Ether/EtOAc 7:3) which gave a yellow solid **64** (30 mg, 93%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 10.75 (1H, s); 8.84 (1H, s); 8.74 (1H, t); 7.75-7.43 (4H, m); 7.04 (1H, bs); 5.97 (1H, bs); 5.81 (1H, s); 4.35 (2H, t); 4.18-4.06 (4H, m); 3.84-3.70 (2H, m); 3.43-3.26 (7H, m); 2.16 (2H, quintet); 1.43 (9H,s); 1.01 (9H,s); 0.93 (9H,s); 0.89 (9H,s); 0.27 (3H,s); 0.19 (3H,s); 0.17 (3H,s); 0.14 (3H,s); 0.05 (6H,s). **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{48}H_{82}N_7O_{10}Si_3$) 1000.45, found 1000.43.

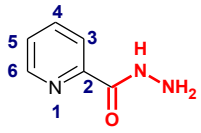


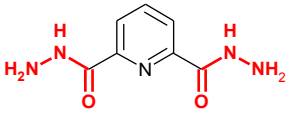
Chapter 3: Acylhydrazone Pyridine Motifs

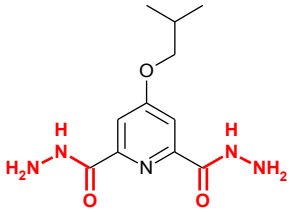
Nomenclature

Monomers

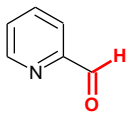
H (Hydrazide)

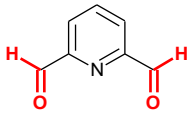




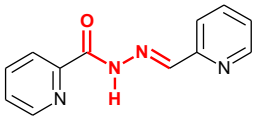


A (Aldehyde)

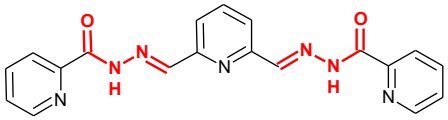




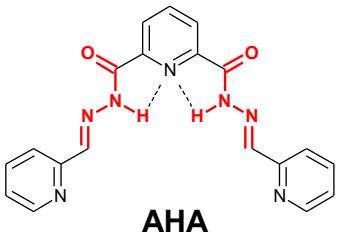
Dimer and Trimers



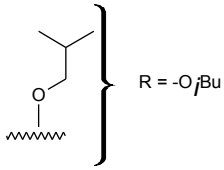
AH

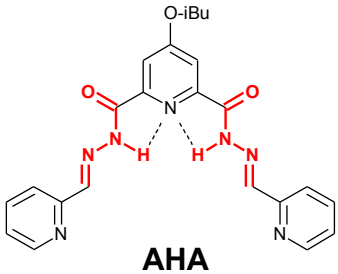


HAH



AHA





AHA

The naming of these motif's is based on the initial monomers used to produce the molecules. The dimer AH is therefore named due to the condensation of an aldehyde (A) and a hydrazide (H) to form a dimer hydrazone (AH). This nomenclature continues onto the trimers such as AHA and HAH. With regards to AHA, the central core monomer of a dihydrazide is reacted with two molar equivalents of aldehyde, thus resulting in the AHA trimer. As for HAH, the central core monomer of dialdehyde is reacted with two equivalents of hydrazide, to produce the HAH trimer. This form of nomenclature works well for symmetrical acylhydrazone motifs.

1 Introduction

1.1 Short Introduction to Foldamers

This is a brief introduction into a class of supramolecular structures consisting of long folded synthetic oligomers and collectively known as foldamers. These fascinating architectures mimic the folding of natural existing biopolymers such as DNA *(Fig. 1)*, peptides and proteins, which are held together by non-covalent forces (hydrogen bonding, stacking, electrostatic interactions and metal-ion coordination). Synthetic foldamers also fold into well-defined conformations in solution. They form folded structures mainly due to their backbone rigidity which is determined locally at the molecules rotatable bonds in the absence of other strong factors.[98]

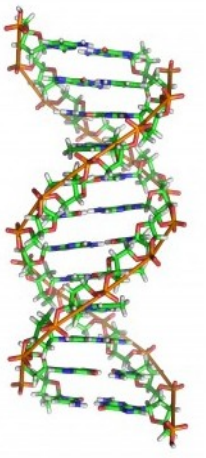


Figure 1: DNA, a natural helical structure

There are four contributing factors involved in foldamer folding which have varying dominance depending on the particular molecule. These factors include: local rotational restrictions, interactions between sites remote in the sequence, solvophobic effects and assembly/hybridization. Electrostatic and steric repulsions play a less visible but still an important role in folding. Examples of artificial folded structures include synthetic α-peptide sequences, artificial proteins, nucleic acids, and helical polymers, among others. The ability to increase the size and complexity of these synthetic foldamers, and use them in a variety of potential applications has lead to our interest in studying such molecules. An added benefit is the capability to mimic elegant processes of nature. Foldamers are also adaptable and responsive, thus increasing their contribution towards efficient recognition and the creation of functional molecules.

The Huc group has developed this world renowned research within the field of biomimetics and bioorganic chemistry.[99] The laboratory has focused their attention on a subfamily of foldamers called aromatic oligoamides (AOA).[100] These non-natural aromatic helical structures adopt exceptionally stable and specific secondary or tertiary folded conformations. This important property is due to π-π stacking from the aromatic surface and intra-molecular hydrogen bonds between the amide hydrogen and the endocyclic nitrogen.[101] *(Fig. 2)*

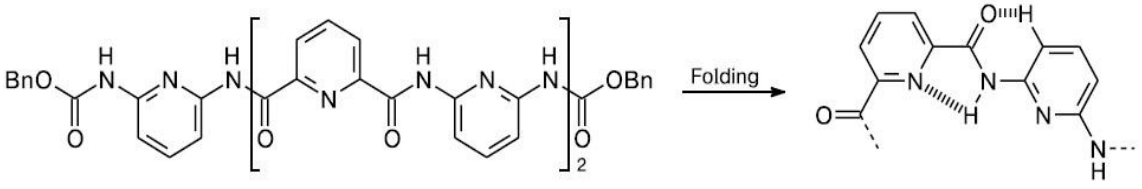


Figure 2: Structure of an aromatic oligoamide strand derived from 2,6-diaminopyridine and pyridine-2,6-dicarboxylic acid. Folding occurs via intramolecular hydrogen bonding about the rotational aryl-CONH bond.

The Huc group has established that these foldamers have a degree of structural predictability and tunability. The structure can be designed in such a way that complete folding is the final outcome. Local conformational preference of the aryl-amide bond can selectively be changed in order to induce conformational rearrangements such as helical-linear strand transitions. Some desired features of the AOA foldamers include their remarkable ability to adopt stable helical conformations even in polar solvents or at high temperatures. Other features include their ease of synthesis required to access large molecular weights, and their demonstrated resistance to hydrolysis. It has been observed that the resulting stability can vary with chemical nature and length of the foldamer. For example, by attaching various side chains the molecules solubility in water or their versatility in protein recognition can be increased.

Dynamic capsules can be engineered thanks to the ability to tailor subunits and thus adjust the diameter of the interior hollow of the foldamer.[98,99] The group has also previously demonstrated the potential biological applications of these aromatic oligomers, for example their ability to transport cargo molecules across the cell membrane with concomitant protection towards enzymatic degradation. The structure of foldamers can be characterised in the solid state by X-ray crystallography *(Fig. 3)* and in solution by NMR studies *(Fig. 4)*. [101,102]

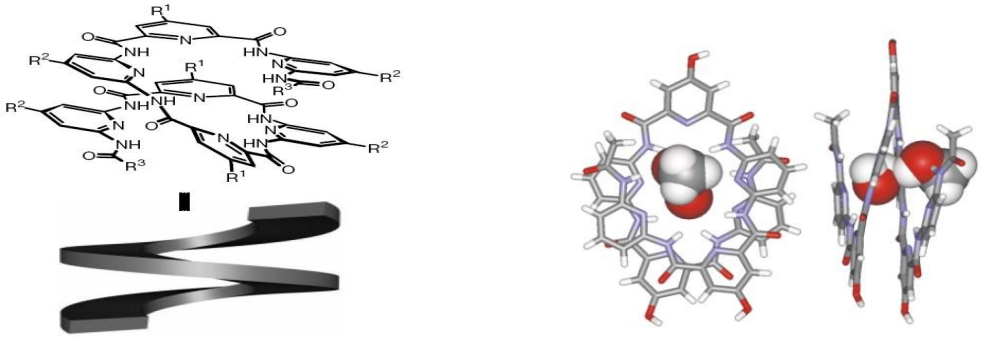


Figure 3: (Left) Example of a folded AOA. Folded conformation of the 2'-pyridyl-2-pyridinecarboxamide motif. R1 and R2 represent side chains.[103] (Right) Example of a solid state characterisation, aromatic oligoamide strand derived from 2,6-diaminopyridine and pyridine-2,6-dicarboxylic acid heptamer (contains seven pyridine units) crystal grown from DMSO-EtOH shows helical conformation with one and a half turn containing one water and one methanol molecule.[101]

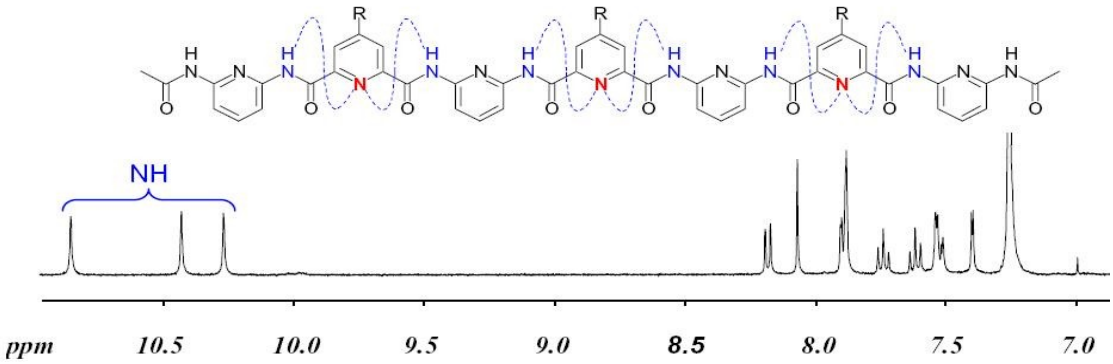


Figure 4: ^{1}H NMR spectra of a pyridinecarboxamide heptamer.[104]

The full potential of foldamers and their applications is still largely unexplored and more opportunities are constantly being unveiled. We are at the beginning of an exciting field, in which the past achievements can lead to new developments. The Huc group has thus provided a broad basis for the creation of future foldamer design by their established ability to synthesise these oligomers with an array of physiochemical properties, such as size, shape and higher order arrangements.

Despite the progress, many options remained to be explored and further understood. In practise, aromatic oligoamides have mostly been limited to structural studies, leaving the door open to more applied techniques. The construction of foldamers from various monomers and their role as small guest carriers across the cell membrane can be more aggressively pursued, and it is noted that previous oligomers within the group have had overall limited degrees of freedom. The development of new molecules could incorporate increased flexibility and utility. One such hypothesis to explore is the use of folding for the deconvolution of a library of oligomers. The question is can a mixture of compounds be resolved by utilising folding? This process of using a spontaneous thermodynamic driving force due to formation a helical structure could be used to produce a single dominant product. This process might also be enhanced by the use of a chemical effector such as metal ion complexation/protonation or by template recognition. With these questions and goals in mind, a new type of system containing acylhydrazone pyridine motifs was investigated *(Fig. 5)*.

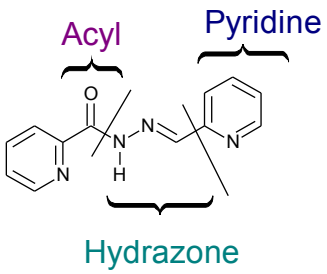


Figure 5: Structure of acylhydrazone pyridine motif. ***AH*** (Derived from **A** = Aldehyde and **H** = Hydrazine condensation.)

These hydrazone derivatives are interesting because they contain an acyl group and the conjugated system makes rotation about the hydrazone double bond possible, resulting in predictable configurational changes that can contribute a new feature to our foldamer libraries. This has not been reported in previous investigations involving hydrazone oligomer studies. Moreover this avenue provides a possible additional system within the field of dynamic combinational chemistry.

1.2 Dynamics and Conformational Control of Folding

Studies published by the groups of Lehn and Moore have inspired the idea's, objectives and design strategy of this thesis section (Chapter 3).[105,106] Lehn's group have demonstrated the control of self-organisation and folding by manipulating synthetic molecular strands into well-defined and ordered shapes, achieved by an appropriate choice of constituting subunits.[105,107,108] They observed the interconversion between two states, specifically of linear and helical molecular shapes triggered by metal-ion binding and also demonstrated the alternate use of chemical effectors to achieve similar effects. Our attention was drawn to molecular machine's fuelled by a mechanism of acid-base neutralisation that possesses on/off reversible motional processes. The devices contained units which can induce an extension/contraction (unfolding/folding) interconversion between a “free helical state” and a “bound linear form” through a reversible metal ion complex formation.[105] Polyheterocyclic strands containing pyridine-pyrimidine (py–pym) subunits are known to fold into extended helical structures.[109]

However, Lehn's group replaced the 2,6-disubstituted pyridine with its isomorphic analogue, a hydrazone group (hyz) while retaining the same structure-inducing features[108] *(Fig. 6)*. These compounds were more accessible via a facile synthesis based on simple condensation between hydrazine and carboxaldehyde derivatives of pyrimidines. A pym-hyz-pym sequence corresponds to a tridentate coordination site so that a hydrazone unit is associated with a coordination site *(Fig. 7)*. A molecular helix is possible when 4, 6, 8 or 10 hydrazone groups are present. In this case, the oligomers are folded in a helix without the presence of a metal ion. A linear strand is obtained upon coordination with Pb^{2+} or Zn^{2+} metal ions.[108]

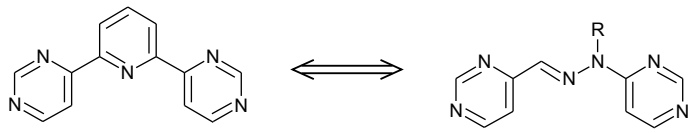


Figure 6: The isomorphic correspondence between the 2,6-disubstituted pyridine and its hydrazone counterpart.[108]

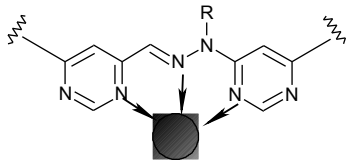


Figure 7: Tridentate coordination site with pym-hyz-pym sequence.[108]

In contrast, to the above work, Lehn et co-workers also published a system which is linear in the absence of metal ions, but folded upon ion complexation resulting in a bound helix *(Fig. 8)*.[105] As previously described, pyridine-pyrimidine (and hydrazone-pyrimidine) sequences undergo helical wrapping without an ion, but α,α'-linked oligopyridine chains adopt a linear

shape. It was shown that a pyridine group can be replaced by a hydrazone functionality while retaining the same features. Upon ion binding, these later strands are helical. This is another example of how chemical effectors can be put to use to inter-convert molecular shapes. The binding of this motif to a Pb^{2+} cation imposes a terpyridine-type coordination geometry of the subunits *(Fig. 9)*. For longer oligomers this results in a coiling of the strand as all transoid inter-heterocyclic connections are converted into cisoid one isomorphic s. isomorphic This process represents a self-induced molecular ion channel formation with “high ion selectivity”, as only the ion capable of inducing the folding is the subject of facilitated transfer.[105]

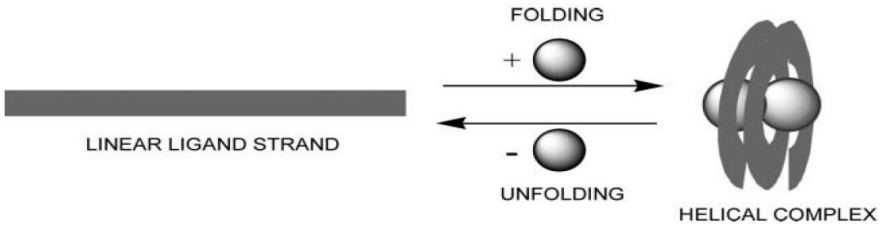


Figure 8: Coiling / uncoiling, contraction/extension molecular motion induced by ion binding and release.[105]

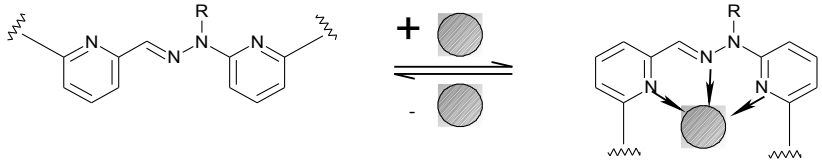


Figure 9: Conversion of a free linearity unit (left) into a bound helicity inducing unit, a terpy-like site (right), by coordination of a cation.

Attention is given to the fact that these hydrazone motifs lack an “acyl” group. Furthermore only the trans conformation was observed throughout investigations of folding and unfolding. The concept of molecular machines and use of chemical effectors, was successfully applied to studies within the Huc group. Protonation of an oligopyridine-dicarboxamide foldamer leads to a transition between two conformations via a linear intermediate.[110] Upon partial protonation, the foldamer unfolds into a linear oligomer and refolds into a helical structure when fully protonated. The basis of this can be assigned to the intramolecular hydrogen bonds stabilising the bending of the helices to form strongly preferred conformations. In the figure *(Fig. 10)*, one can see the effects of protonation upon the monomers dicarbonylaminopyridine and pyridinedicarboxamide. Depending on the location of the amide carbonyl group, a particular conformation will have preference upon protonation. Upon protonation, 2,6-dicarbonylaminopyridine carbonyl groups form hydrogen bonds with pyridinium $N^{+}H$ and not H^{3} aryl protons.[110]

In **A**, the C=O...H-C is weak, whereas in the case of **B** the N-H...N intramolecular bond is stronger. Electron-withdrawing effects of carbonyl groups and intramolecular hydrogen bonds results in a decreased basicity of the pyridine nitrogen. This property therefore plays a role in the different basicity of each of the monomers which in turn leads to a specificity and selective protonation via different strengths of acid. The various acids tested were: CF_3CO_2H (TFA, pK_a = 0), CH_3SO_3H (MsOH, pK_a = -2) and CF_3SO_3H (TfOH, pK_a = -16).[110] Selective control over the contraction and extension equilibrium of these two oligoamides was observed with a change in pH between an unprotonated helix, to a partially protonated linear strand towards a fully protonated helix distinct from the unprotonated helix.

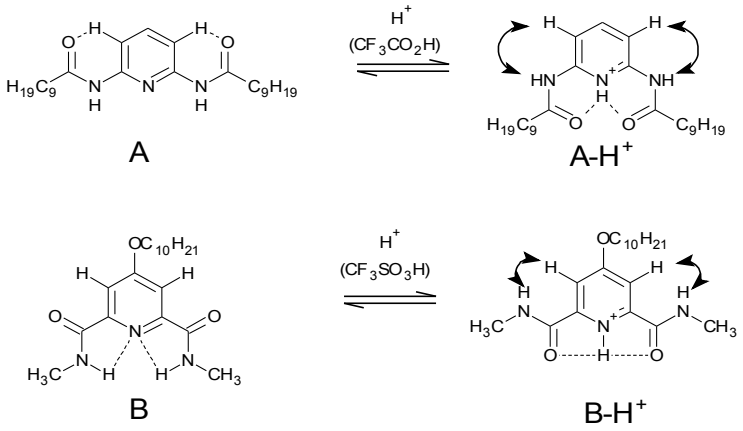


Figure 10: Changes in the conformations of **A** and **B** upon protonation to give **A-H$^+$** and **B-H$^+$** respectively. The double-headed arrows represent NOE correlations seen in the spectra of **A-H$^+$** and **B-H$^+$**.[110]

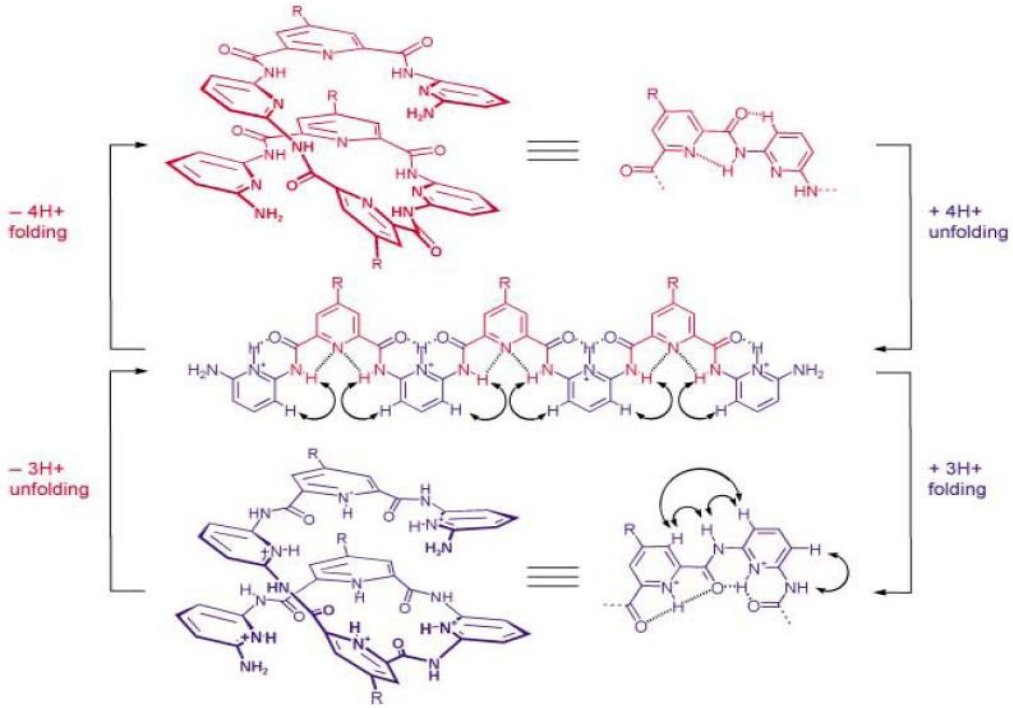


Figure 11: (Red) Helical conformation when unprotonated. (Centre) Linear conformation with partial protonation $(H^+)_4$,. (Blue) Fully protonated $(H^+)_7$ results in refolding into a helical conformation with inverted carbonyl groups. [104]

1.3 Combining Molecular Recognition with Orientations

The above examples from the Lehn and Huc group, show that the use of ions or protonation chemical effectors can contribute towards switching between two states selectively, from that of a folded state towards that of a linear form.[105,110] However both situations deal with folding which is consistently trans conformation and no ambiguity is involved. Our next quest is the consideration of libraries and how to identify and selectively form a preference for the expression of one specific member of a dynamic structural/conformational library.

For this we turn to work published by Lehn and Huc which examines libraries of molecular strands which generate supramolecular assemblies when a template is introduced.[107] Their receptor had thirty-six possible quasi-planar/conjugated rotamers due to rotations about the CO-aromatic single bond without the presence of a template *(Fig. 12)*. Some of the species have more than one occurrence depending on their symmetry. The 2,6-dicarbonylaminopyridine motif adopts a preference for *s-trans/s-trans* conformation. However the isophthalamide motif can have an equilibrium between three distinct different conformations *(Fig. 12)*.[107]

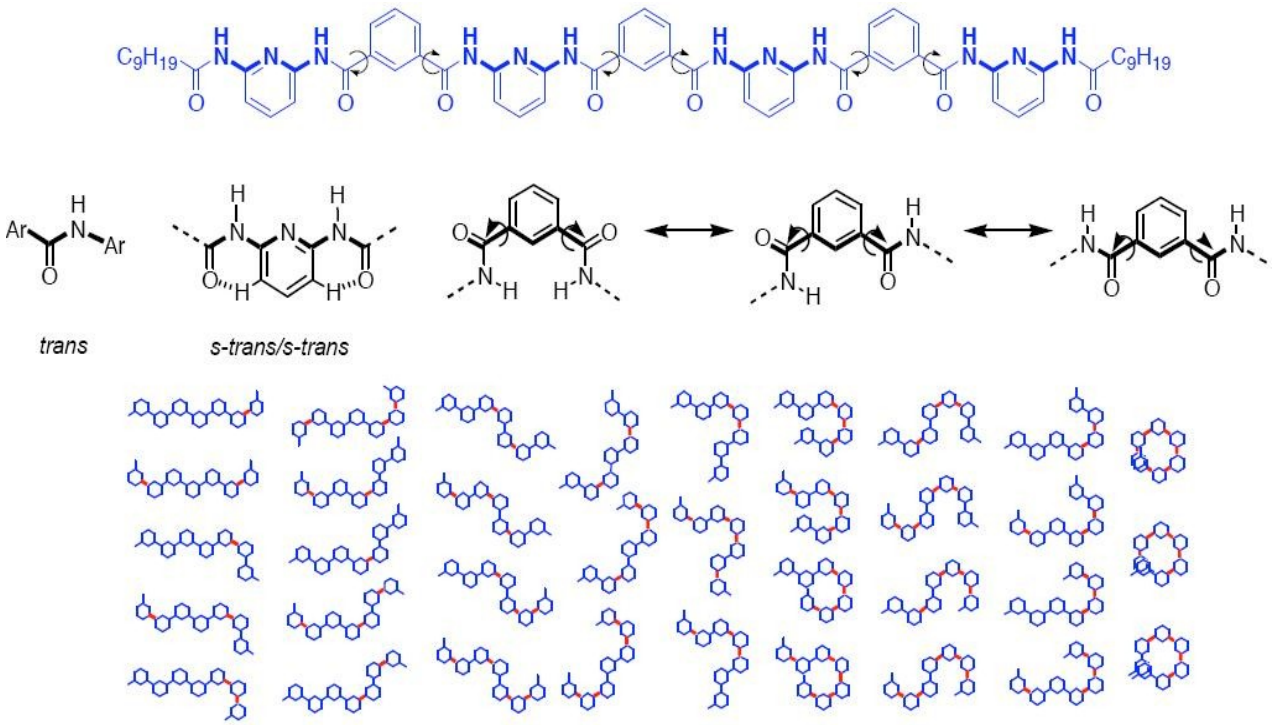


Figure 12: Oligo-isophthalamide receptor; the individual motif configurations; and the oligo-isophthalamide strands thirty-six possible rotameric forms. The linkages between the hexagons represent the CO-NH fragments. The thick linkages indicate those within which a 180° rotation was performed about the CO-aromatic ring bond.[111]

With many possible outcomes, Lehn and Huc, managed to isolate one specific library member *(Fig. 13, 14)*. To accomplish this, they utilised a template as an effector. The incorporation of H-bonding arrays into the components enabled molecular recognition between a linear oligo-isophthalamide strand and a cyanuric acid template. This resulted in a phenomenon of conformational reorganisation to obtain a helical disk-like object and can be considered as deconvolution from a library of conformers. This in turn drives a "second-level" self-assembly with increasing concentration resulting in aggregation due to solvophobic and stacking interactions to form columnar entity resulting in the formation of polymeric fibres *(Fig. 13)*. Additional means of control is represented by the medium, that is fibre formation depends critically on the solvent composition, temperature and concentration. This contributes to the receptor adaptability.[107]

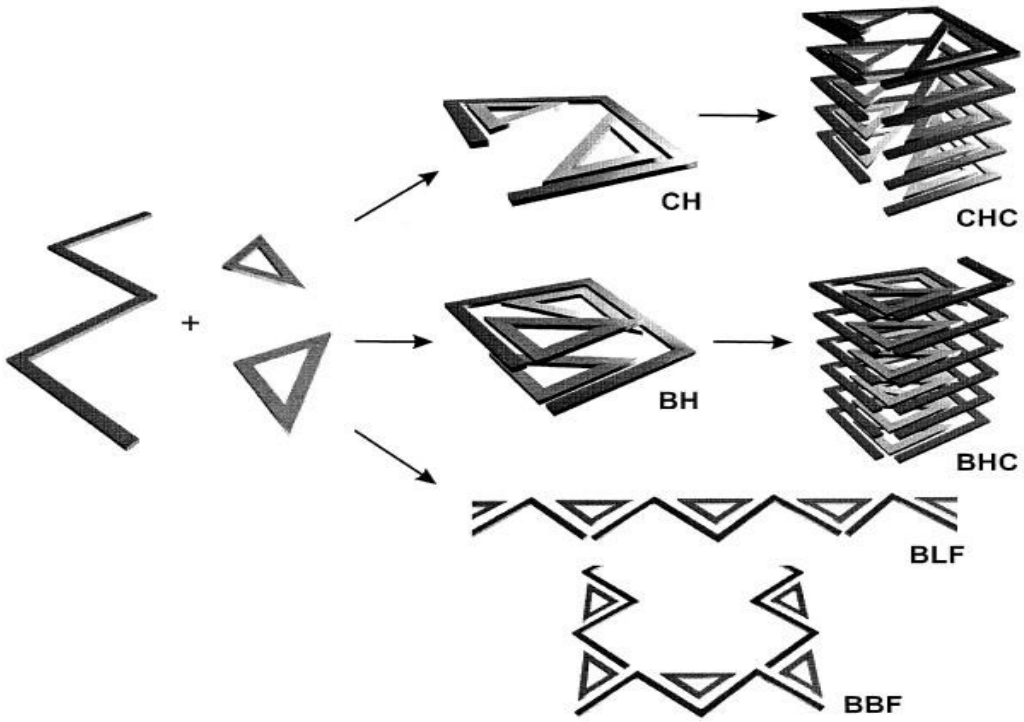


Figure 13: (Left) Receptor and two cyanuric templates can adopt various self-assemblies via three distinct pathways. 1) **CH:** Closed Helical form, internally bound 2) **BH:** Bridged helical form. **CH** and **BH** may assemble into a "2nd-level" assemblies **CHC** and **BHC** respectively. 3) Two types of non-helical bridge forms, linear **BLF** or bent **BBF**.[107]

The resulting selection and folding into a helical structure depends on the template employed due to different hydrogen bonding interactions. The receptor is linear with an ADA imide template (bottom), or may bind to a double-faced ADA/ADA cyanurate template (top), to form a disk-like object *(Fig. 14)*. What is very interesting, is that this represents programmed chemical systems and reading of molecular information to yield a defined structural output depending on the template. The addition of the cyanurate template in this instance meant the conformations were found to adapt consistently trans orientations despite the possibility of the oligo-isophthalamide strand being able to adapt a s-trans-cis conformation.[107]

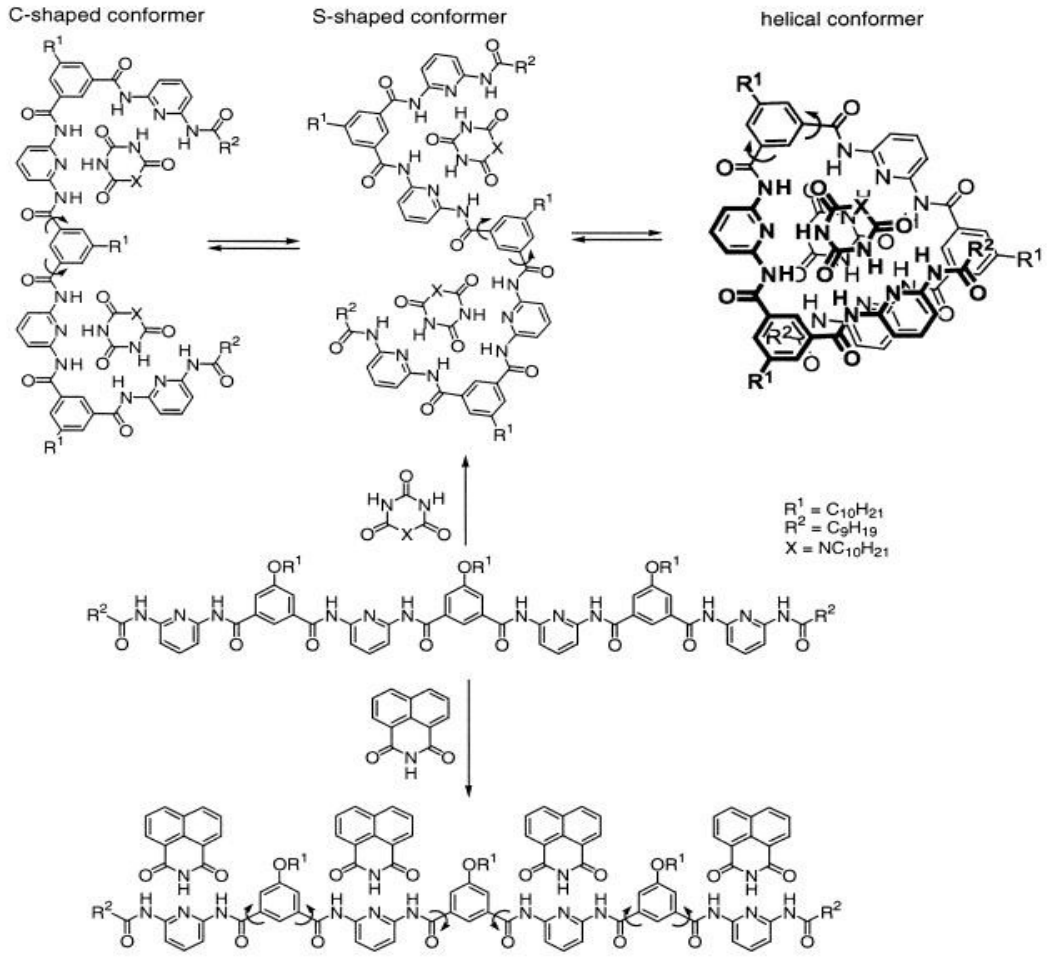


Figure 14: Template-dependent expression of the information stored in the oligo-isophthalamide receptor. Various conformer isomers of the receptor depending on the template employed (ADA/ADA cyanurate template, top or ADA imide template, bottom). This represents programmed chemical systems with proof reading resulting in a particular outcome.[107]

1.4 Dynamics & Selectivity from a Library of Oligomers

The above are examples of a folding process that exists only when a chemical effector is employed. However Moore's group showed that folding could also be used to select from a pool of oligomers, whose segments are most “prone” to fold.[106,112–114] They demonstrated that stability gained from folding can be used to control the synthesis of oligomers. They allowed two imines to compete for one equivalent of another imine. Short and long chain segments were reversibly ligated through an imine metathesis reaction *(Fig. 15)*. Therefore, folding shifts the ligation equilibrium in favour of conformationally ordered sequences, so that oligomers having the most stable solution structures form preferentially. However imine metathesis was slow in neutral organic solvents at room temperature and the presence of appropriate acid catalyst resulted in the reaction proceeding at a reasonable rate.[106] The folding equilibrium was found to depend on length of the ligated oligomer, solvent and temperature *(Fig. 15)*.

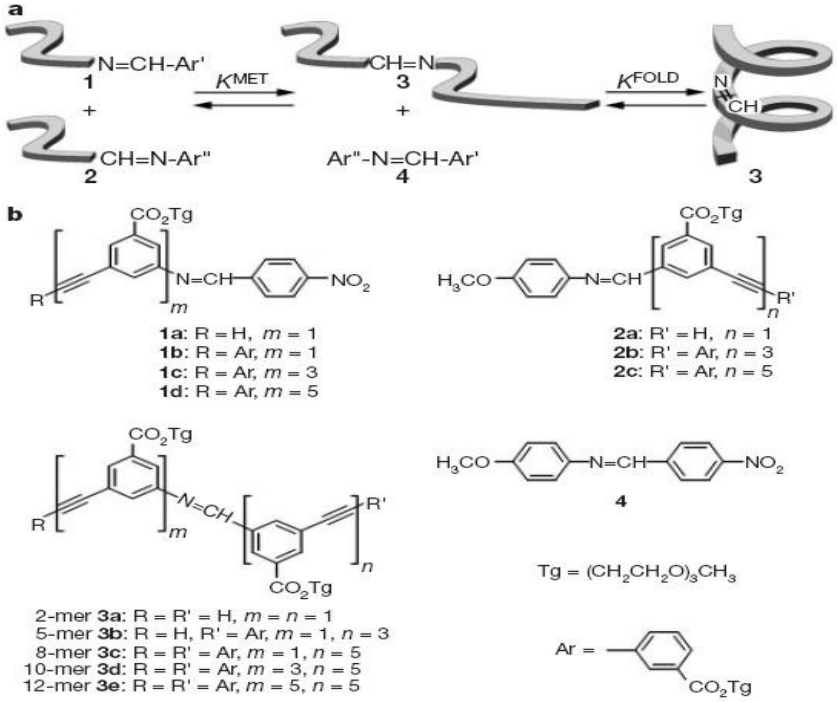


Figure 15: Imine metathesis of meta-connected phenylene ethynylene oligomers. (a) Representation of the ligation reaction of oligo (m-phenylene ethynylene) imines. The coupled equilibria correspond to imine metathesis and helical folding. The folding equilibrium depends on the length of the ligated oligomer, the solvent, and the temperature. Equilibrium constants for the metathesis of aromatic imines are 1.0 ± 0.2 regardless of substituents. (b) Chemical structure and corresponding compound numbers of various imines used in this study. Oligomer segments of different lengths (**1a-d** and **2a-c**) were synthesized by repeated palladium/copper-catalysed coupling of appropriate aryl halide and terminal acetylene precursors. *N*-terminal and *C*-terminal imine groups were added by condensation of the corresponding aldehydes and amines. K^{MET} = equilibrium constant for imine metathesis. K^{FOLD} = equilibrium constant for helix-coil transition.[106]

In the first experiment, Moore's group wanted to determine how the metathesis equilibrium depends on the length of the oligomer chain synthesised under conditions that favour folding. They also wanted to know “if” folding drives ligation and therefore expecting “critical” size is necessary for a folding-driven ligation to occur. In their second experiment, they monitored the metathesis equilibrium under solvent conditions that favoured folding and compared this to equilibrium state to metathesis reactions occurring in solvents which do not favour folding.

Folding was possible in acetonitrile but not in $CDCl_3$. The metathesis of longer oligomers in identical conditions exhibited considerable equilibrium shifting and therefore suggests equilibrium shifting which depends on folding. There is a linear correlation between $\Delta\Delta G$ and chain length above this critical value which is consistent with the notion that folding stability is directly proportional to the contact area between adjacent turns as expected for a compact helical conformation. Therefore the longer the oligomer, the greater stabilizing energy from folding, with this ligation product preferentially selected over shorter oligomers in competitive ligation reactions *(Fig. 16)*.[106] Moore's findings extend to size-selective synthesis of folded oligomers by dynamic templating.

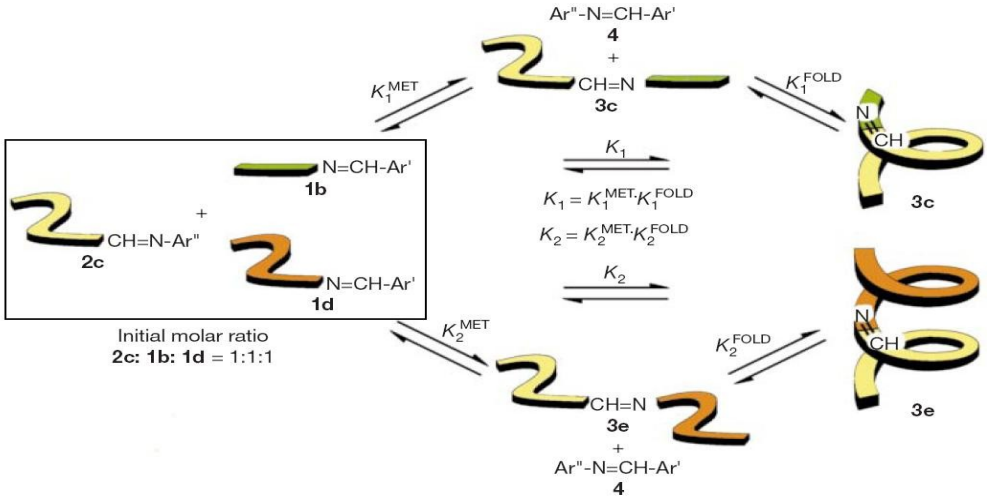


Figure 16: Segment selection in imine metathesis determined by foldability. Representation of the competitive ligation of oligo(m-phenylene ethynylene) imines. K_1^{MET} is the equilibrium constant for the imine metathesis between *N*-terminal dimer **1b** and *C*-terminal hexamer **2c**. K_2^{MET} is the equilibrium constant for the imine metathesis between *N*-terminal hexamer **1d** and *C*-terminal hexamer **2c**. K_1^{FOLD} and K_2^{FOLD} are equilibrium constants for the respective helix-coil transitions resulting in the overall equilibrium constants K_1 and K_2.[106]

That brings us to briefly summarize the main interesting points which have inspired this thesis:

1) A chemical effector such as a metal ion/protonation can induce folding and when removed, unfolding is possible and a linear strand is generated.[110]
2) The presence of a template was necessary to induce folding to obtain a single product from a mixture of thirty-six possible conformations. Depending on the template, a foldamer or a linear strand was obtained.[107]
3) Competition exists between two different lengths of imine metathesis oligomers with the longer oligomer having preference due to increased stability in a self-selection process.[106]

In all of the above, trans conformations were present and no ambiguity was involved. The objectives and the challenges due to the possibility of cis configurations will be discussed. It is expected that 100% cis configuration is possible via folding and cooperativity in driving the selection process forward. It is important to also note that these acylhydrazone pyridine motifs will not depend on a template.

2 Objectives

The initial objective was to explore the equilibrium of short dynamic systems derived from acylhydrazone pyridine motifs. Previous investigations of acylhydrazones, indicated the trans conformation is preferred. However, based on initial studies a system was devised which exhibits cis isomerisation due to the availability of hydrogen bonding in the acylhydrazone pyridine motif.[115] The resulting equilibrium is unusually biased such that cis configuration is

observed. This small system (AH) is shown below *(Fig. 17)* with the transition from trans to the cis configuration depicted. After experiments confirmed the equilibrium between the trans and cis configuration, the next objective was to investigate how this phenomenon translated into longer oligomers (foldamers). This involved the exploration of including multiple acylhydrazone pyridine motifs contained in a larger system. In principle, if oligomer length was increased, the complexity of the system was expected to be greatly increased. Therefore it was interesting to determine, if the outcome of these larger molecules might lead to a well-defined foldamer, as a co-operative effect from folding was expected *(Fig. 18)*. Some questions to be addressed included: Could folding bias the equilibrium and be the driving force to form a well defined helical structure? Would folding be the driving force to resolve mixtures? Was it possible to fine-tune the configuration?

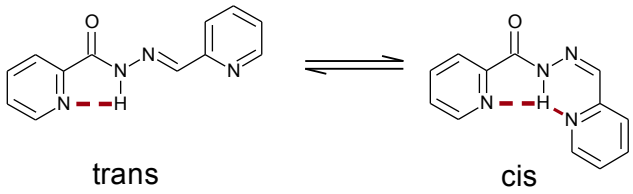


Figure 17: The acylhydrazone pyridine dimer trans-cis isomers. AH configuration between trans and cis.

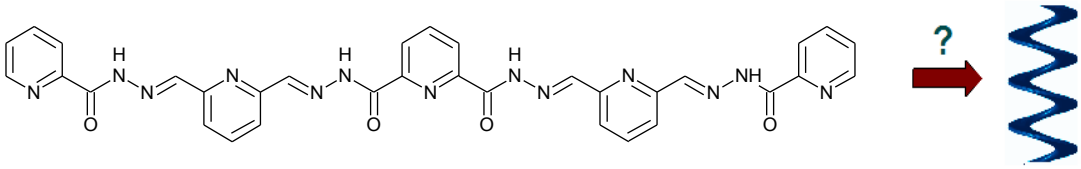


Figure 18: The acylhydrazone pyridine pentamer and determining if will folding occur.

3 Design Strategy

Pyridines are a widely exploited monomer in foldamers.[79,116] Incorporation of a novel acylhydrazone pyridine system that exhibits a bias for cis configuration could give access to a new family of foldamers. This in turn could lead to foldamers with new secondary structures and characteristics with less rigidity than their predecessors. The design utilised available intramolecular hydrogen bonding to induce an equilibrium biased towards the cis configuration. A short dynamic AH dimer system, with one hydrazone motif, was initially investigated for its trans-cis behaviour *(Fig. 19* ***1****)*. Initial studies have shown that only 50% cis configuration was achieved, and by increasing oligomer length it was expected folding and cooperativity would drive the system forward to obtain 100% cis configuration. Two symmetrical trimer systems, AHA and HAH, each containing two acylhydrazone pyridine motifs, were also tested. *(Fig. 19* ***2****)*. Intramolecular ring size and conjugation of the acylhydrazone pyridine motif may have an important effect on its ability to exist in the cis configuration form. AHA contains two central five-membered hydrogen bonding rings in a V shaped resulting in its trans-trans conformation. It was expected upon formation of the trans-

cis configuration in the AHA trimer a new six-membered intramolecular ring to form via hydrogen bonding at the terminal end. The HAH trimer contains two terminal five-membered hydrogen bonding rings in a linear form resulting in the trans-trans conformation. A new six-membered hydrogen bonding ring is expected upon formation of the trans-cis configuration for the HAH trimer with the central pyridine unit. For both trimers, upon formation of the cis-cis configuration, two five and six-membered intramolecular ring systems could exist, however for the purpose of clarity and explanation they have not been depicted in the diagram. These molecules were synthesised and their properties investigated.

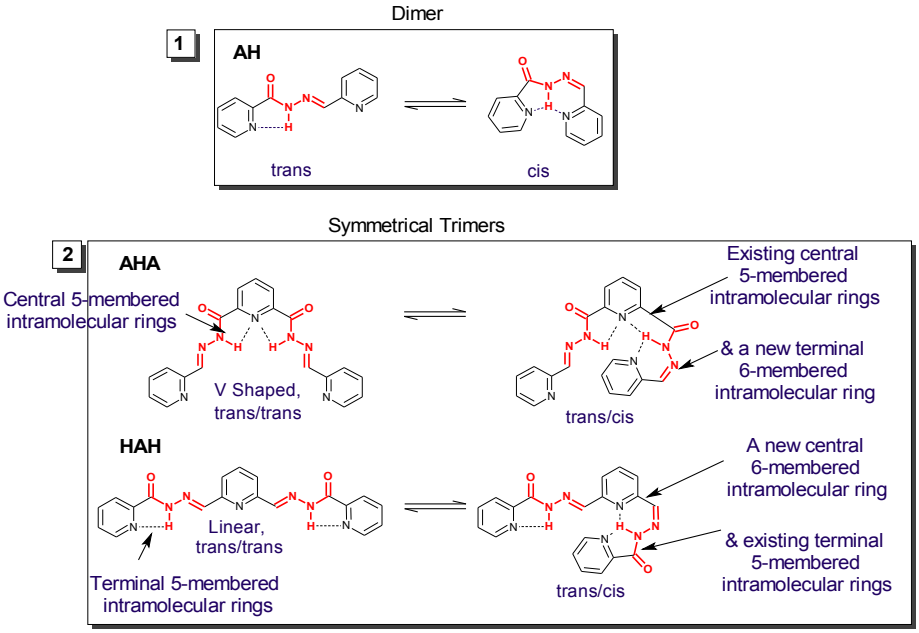


Figure 19: 1. The acylhydrazone pyridine dimer system (AH) equilibrium between the trans and cis configuration, exhibiting an unusual biased equilibria by utilisation of the available hydrogen bonding. **2.** The acylhydrazone pyridine trimers (AHA and HAH). Molecular modelling shows the differences in intramolecular hydrogen bonding ring size due to the order of acylhydrazone pyridine motif.

The previous trimers mentioned (AHA and HAH) are symmetrical and made with bi-functional central units. However, desymmetrisation and incorporating a central monomer unit containing masked functions was also desired. *(Fig. 20* ***1****)*. This was expected to allow for facile polymerisation of longer oligomers. It was not possible to have both an aldehyde and hydrazide functional group contained on a single monomer, as this would result in uncontrollable polymerisation. Masking of the hydrazide function as a methyl ester, or an aldehyde as a primary alcohol, was a solution to overcome this issue *(Fig. 20* ***2****)*. The introduction of a side chain to this monomer increases solubility of longer oligomers in organic solvents. The dis-symmetrical trimer contains both a terminal and central five-membered intramolecular ring system and its effect upon equilibria to the cis configuration was expected to provide interesting equilibria results. For clarity, the cis-cis configuration of this trimer is not depicted in the diagram. Hence the dis-symmetrical monomer is interesting due to both its versatility and desirable features.

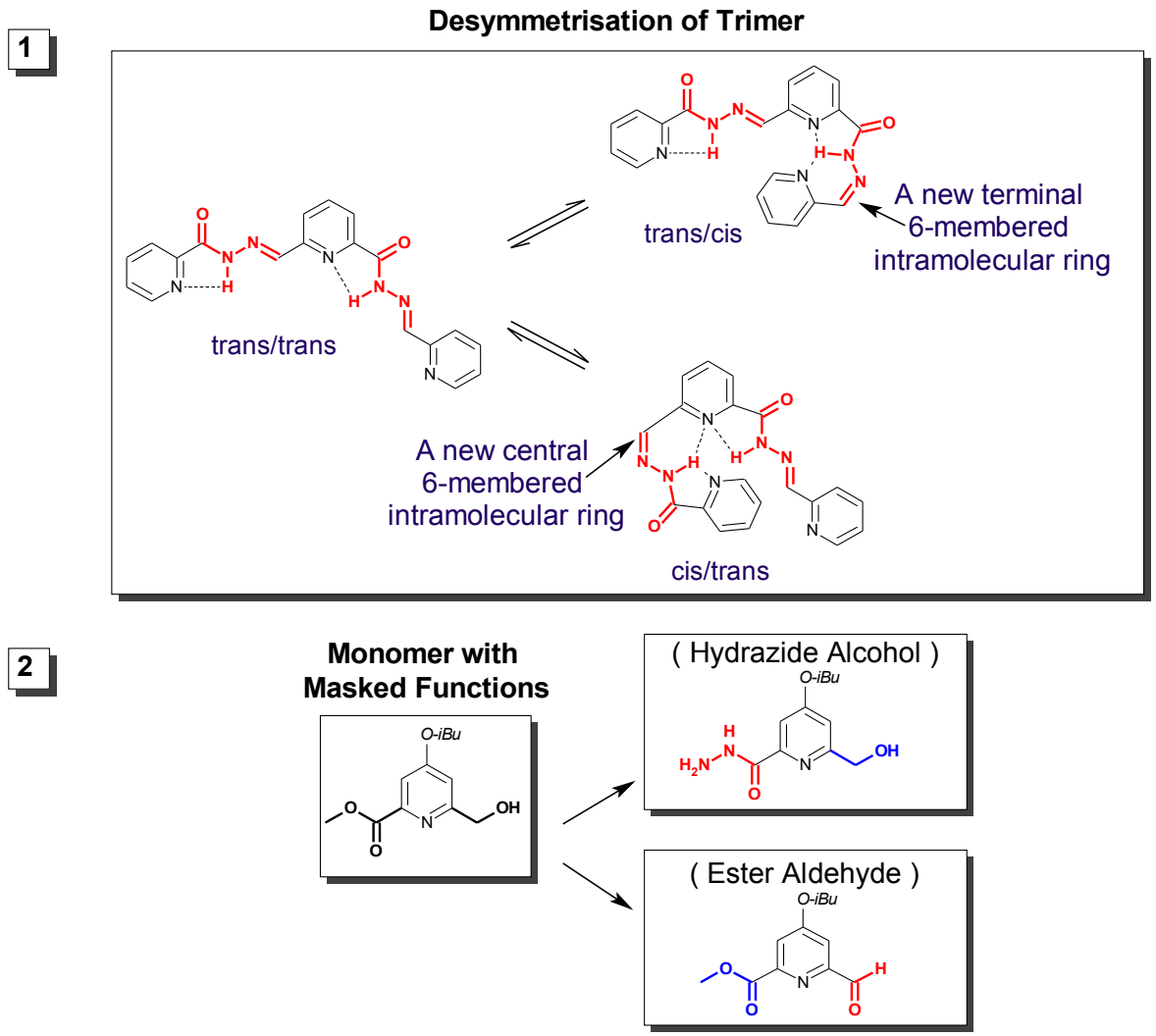


Figure 20: 1. Desymmetrisation of the trimer which contains both a terminal and central five-membered intramolecular rings to investigate their effects on the cis configuration. **2.** Monomer with two masked functions: (i) hydrazide masked as a methyl ester and (ii) an aldehyde masked as a primary alcohol. The monomer also contains an isobutoxy solubilising side chain.

Longer acylhydrazone pyridine oligomers were expected to have increased complexity. The increased number of acylhydrazone units allows for increased degrees of freedom and flexibility. The possibility of π-π stacking of the planar aromatic surface may further enhance folding stability and contribute to a co-operative effect of longer aromatic oligomers having a preference compared to that of aliphatic oligomers. Previous generations of foldamers exhibited rigidly locked conformations with an inability to unfold.[98] Whereas these dynamic systems could contain "*n*" possible library members and in the case of HAHAH pentamer, it contains four hydrazone motifs with sixteen possible configurations *(Fig. 21)*. However some configurations are degenerate, due to the occurrence of symmetrical derivatives. Thus for HAHAH pentamer, six configurations are degenerate, leaving ten unique configurations to be investigated. *(Fig. 21).* With the implementation of these design concepts, the question of "can folding influence equilibria" was investigated.

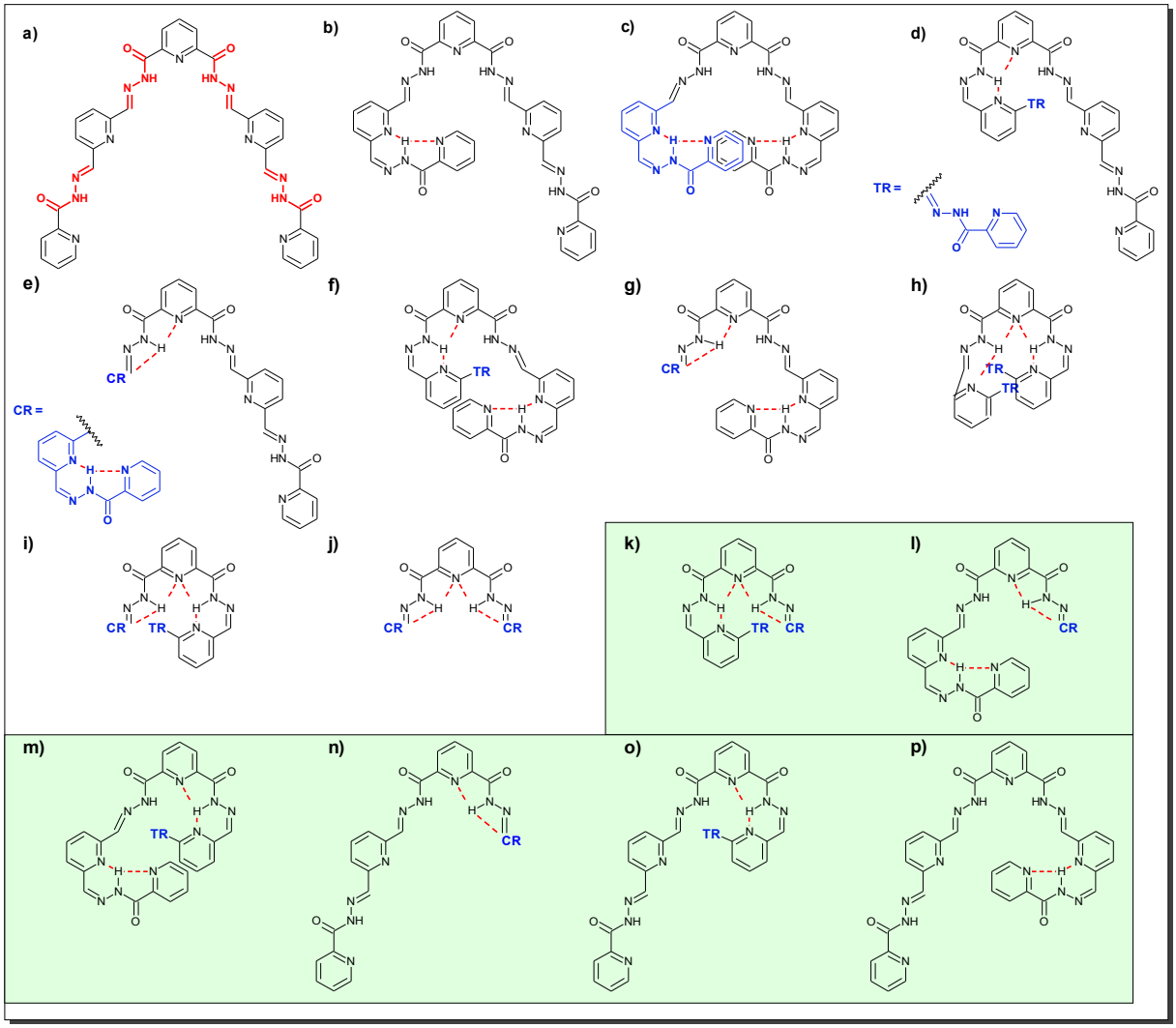


Figure 21: HAHAH configurations **a)** trans, trans, trans, trans; **b)** cis, trans, trans, trans; **c)** cis, trans, trans, cis; **d)** trans, cis, trans, trans; **e)** cis, cis, trans, trans; **f)** trans, cis, trans, cis; **g)** cis, cis, trans, cis; **h)** trans, cis, cis, trans; **i)** cis, cis, cis, trans; **j)** cis, cis, cis, cis; (the following in green are degenerate species) **k)** trans, cis, cis, cis; **l)** cis, trans, cis, cis; **m)** cis, trans, cis, trans; **n)** trans, trans, cis, cis; **o)** trans, trans, cis, trans; **p)** trans, trans, trans, cis.

Molecular modelling of HAHAH suggested the design features outlined above were expected to allow acylhydrazone pyridine motifs to adopt different geometries upon intramolecular hydrogen bonding *(Fig. 22)*. The environmental surroundings were expected to play a role in the equilibrium behaviour and contribute towards induced folding and for this reason studies in various solvents, acids and temperature were conducted. To achieve this goal, synthesis of the motifs was carried out to afford the various desired monomers, dimers and trimers. With these acylhydrazone pyridine derivatives in hand, the outcome and selection process was monitored. To test the deconvolution ability of these complex systems, length variation, protonation and solvent effects of the oligomers were investigated.

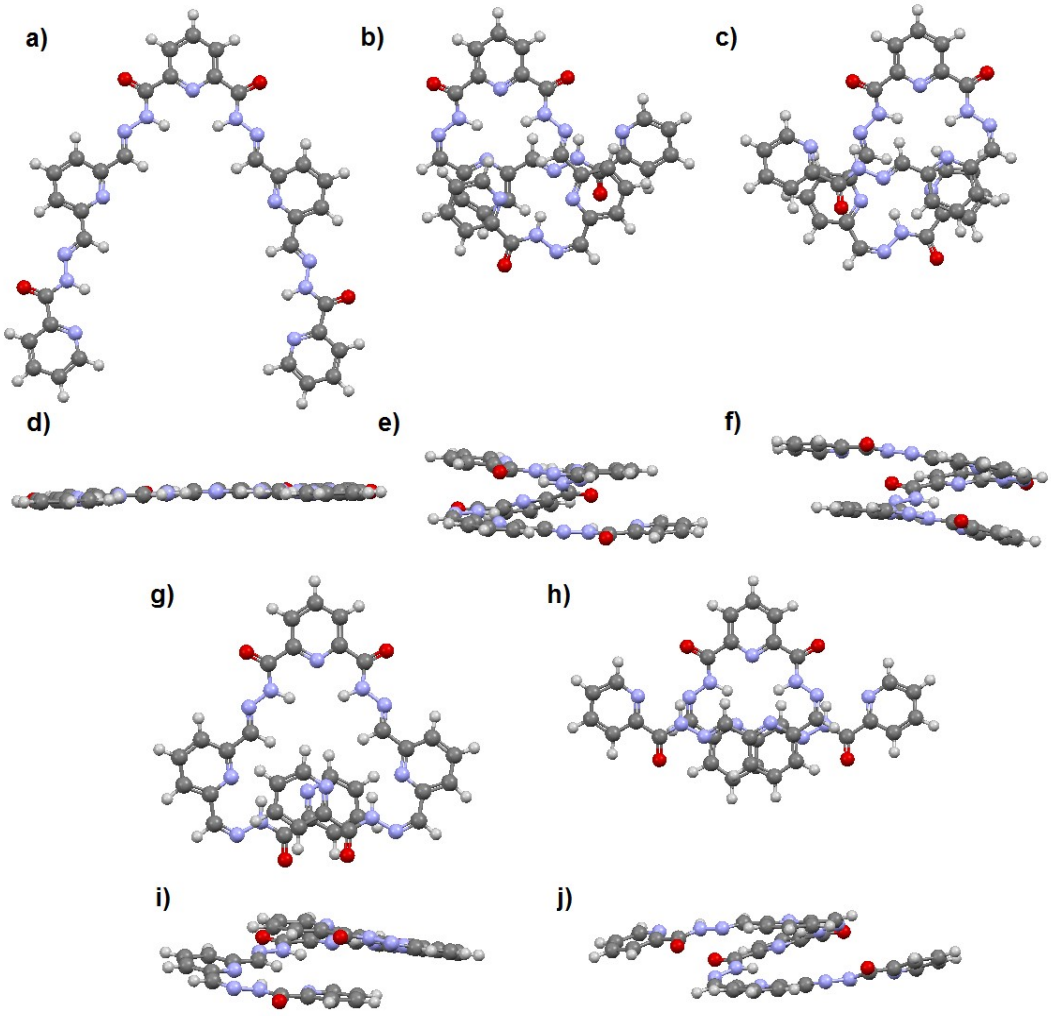


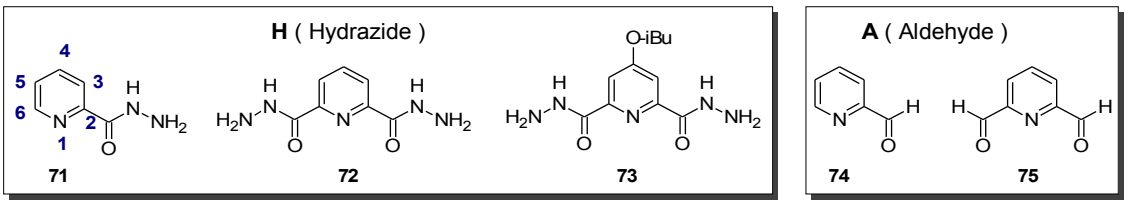
Figure 22: Pentamer HAHAH configurations molecular modelling **a)** trans, trans, trans, trans; **b)** trans, cis, trans, cis; **c)** cis, trans, cis, trans; **d)** trans, trans, trans, trans (side view); **e)** trans, cis, trans, cis (side view); **f)** cis, trans, cis, trans; (side view); **f)** cis, trans, trans, cis (side view); **g)** trans, cis, cis, trans (side view).

4 Synthesis

4.1 Symmetrical Units

4.1.1 Monomers

To date, the synthesis of some monomer building blocks required for this project had previously been established,[115] with further adjustments required to complete the desired structures, such as the addition of a solubilising side chain. As discussed previously, the monomers consist of bi-functional central units **72, 73** and **75**, and mono functional terminal units **71** and **74**. The ease of synthesis provided initial building blocks in good yield, **71** (99% yield), **72** (99% yield), **73** (99% yield), and **75** (40% yield) while one was commercially available **74** *(Scheme 1)*. From these monomers, it was evident that short dynamic systems could be efficiently made which did not require side chains. However upon oligomer elongation, it was envisaged that side chains would enhance organic solubility and thus ease the synthesis of larger constructions.

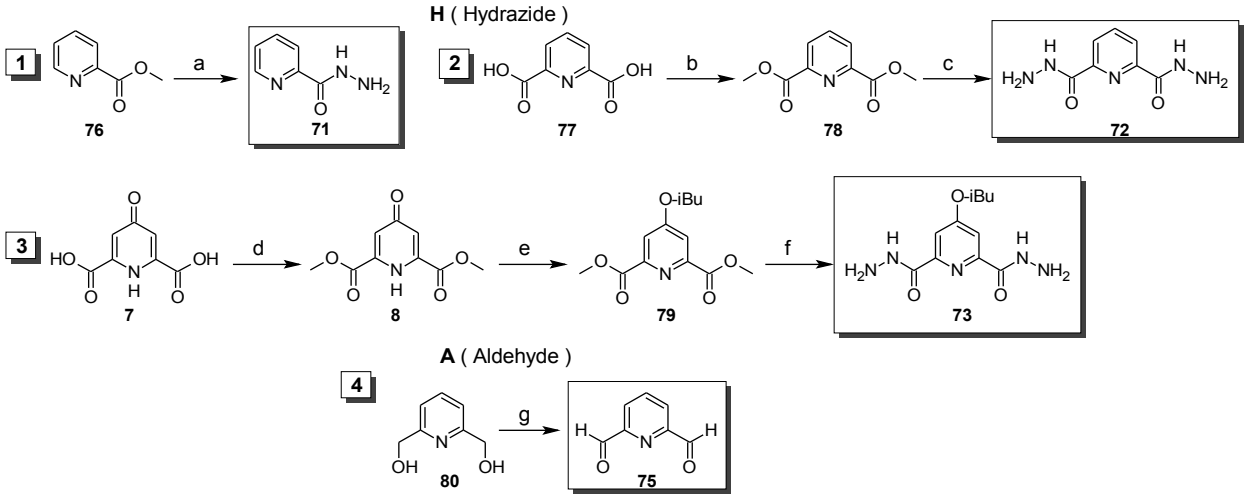


Scheme 1: Hydrazide and Aldehyde monomers (symmetrical central units **72, 73** and **75**).

The first aim in the synthesis was to establish a reliable method for reacting the pyridine esters with a hydrazine to form the hydrazide building blocks. To obtain 2-pyridine carbohydrazide **71**, commercially available methyl picolinate **76** was reacted with hydrazine monohydrate in methanol to give the product **71** (99% yield) *(Scheme 2, Synthesis 1)*. Dimethyl 2,6-pyridine dicarboxylate **78** was previously synthesised in our group by esterification from the commercially available 2,6-pyridine carboxylic acid **77** in the presence of methanol and a catalytic amount of sulphuric acid. The hydrazide synthesis for the pyridine-2,6-dicarbohydrazide **72** was similar to that of the pyridine-2-carbohydrazide **71**, carried out in the presence of hydrazine monohydrate and methanol at 0 °C to yield the dihydrazide product **72** (99% yield) *(Scheme 2, Synthesis 2)*.

Introduction of a side chain involved a familiar synthesis outlined in Chapter 2 (Section 4.1.1, *Scheme 3*, Compound **12**).[79] Esterification of chelidamic acid **7** afforded the dimethyl chelidamic ester **8** (95% yield) *(Scheme 2, Synthesis 3)*. To this was added potassium carbonate in anhydrous DMF and heated to reflux for two hours.[117] This was followed by cooling to room temperature and addition of isobutyl iodide, continued by heating to 70 °C for four hours. Dimethyl 4-isobutoxy-2,6-pyridine dicarboxylate **79** (84% yield) was obtained after purification by column chromatography, and then treated with hydrazine monohydrate to afford 4-(isobutoxy)pyridine-2,6-dicarbohydrazide **73** (99% yield), this was used in the next step without further purification.

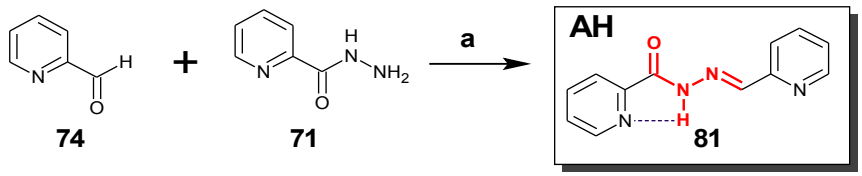
The next step in the building block monomer synthesis involved the formation of aldehyde derivatives. Pyridine-2,6-dicarbaldehyde **75** (40% yield) is commercially available, but can also be easily afforded by oxidation of 2,6-dipyridine methanol **80** was oxidised with activated manganese (IV) oxide under reflux in 1,4-dioxane *(Scheme 2, Synthesis 4)*. An alternative method to obtain **75** in higher yield (90%) is outlined by the Lüning group using selenium dioxide.[118]



Scheme 2: Hydrazide monomers **71, 72, 73:** *Synthesis 1:* (a) $NH_2NH_2.H_2O$, MeOH, 0 °C, (99% yield); *Synthesis 2:* (b) H_2SO_4, MeOH; (c) $NH_2NH_2.H_2O$, MeOH, 0 °C, (99% yield); *Synthesis 3:* (d) HCl, MeOH, Dimethoxypropane, 70 °C, (95% yield); (e) K_2CO_3, DMF, 120 °C 1.5 hr, then iBuI, 70 °C, (84% yield); (f) $NH_2NH_2.H_2O$, MeOH, 0 °C, (99% yield); Aldehyde monomer **75**: *Synthesis 4:* (g) MnO_2, 1,4-dioxane, 100 °C, (40% yield).

4.1.2 Dimers and Trimers

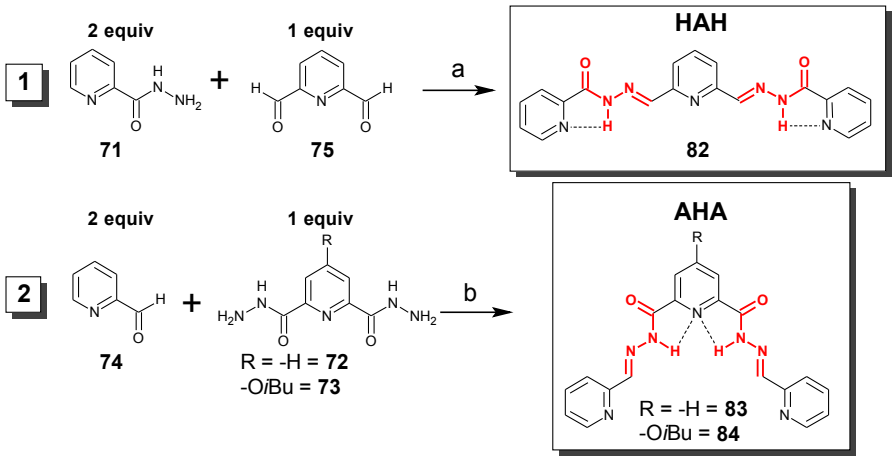
The synthesis of small oligomers (dimers and trimers) allowed the first objective, investigating the equilibria of trans-cis configuration, to be fulfilled. The AH dimer **81** contains one hydrazone unit and was thus useful for carrying out preliminary studies. For the synthesis of dimer **81** *(Scheme 3)*, the monomer 2-pyridine carbohydrazide **71** was reacted with commercial 2-pyridine carboxaldehyde **74** in ethanol under reflux to give the desired AH product **81**. However, the purification of this compound by column chromatography was found to be somewhat difficult. Despite this, the pure AH dimer **81** was obtained in 40% yield. This dimer containing one hydrazone unit was used for NMR studies to investigate the equilibria of the trans-cis configurations.



Scheme 3: AH dimer Synthesis: (a) EtOH, Reflux, (40% yield). Red: Hydrazone functionality.

AH dimer **81** equilibria studies were chosen to provide information about a system containing one acylhydrazone unit. To expand on this information, the trans-cis equilibria of a system containing two acylhydrazone units was investigated. For this reason, the synthesis of HAH **82** and AHA **83** and **84** trimers was carried out *(Scheme 4)*. The symmetrical trimers each contained two acylhydrazone motifs with a different configuration in each compound, for comparison of the intramolecular hydrogen bonding ring systems.

Synthesis proved relatively straightforward. The formation of the HAH trimer **82** (65% yield) consisted of reacting pyridine-2,6-dicarbaldehyde **75** with two equivalents of 2-pyridine carbohydrazide **71** in refluxing ethanol *(Scheme 4, Synthesis 1)*. The formation of the AHA trimer **83** (68% yield) consisted of reacting pyridine-2,6-dicarbohydrazide **72** with two equivalents of commercially available 2-pyridine carboxaldehyde **74** in refluxing ethanol. For NMR analysis, the compound had poor solubility in $CDCl_3$ and thus DMSO-d_6 was used instead. To produce a compound with improved solubility, the AHA trimer **84** containing an isobutoxy side chain was synthesised in good yield (87% yield) by reacting one equivalent of dihydrazide **73** with two equivalents of aldehyde **74**.



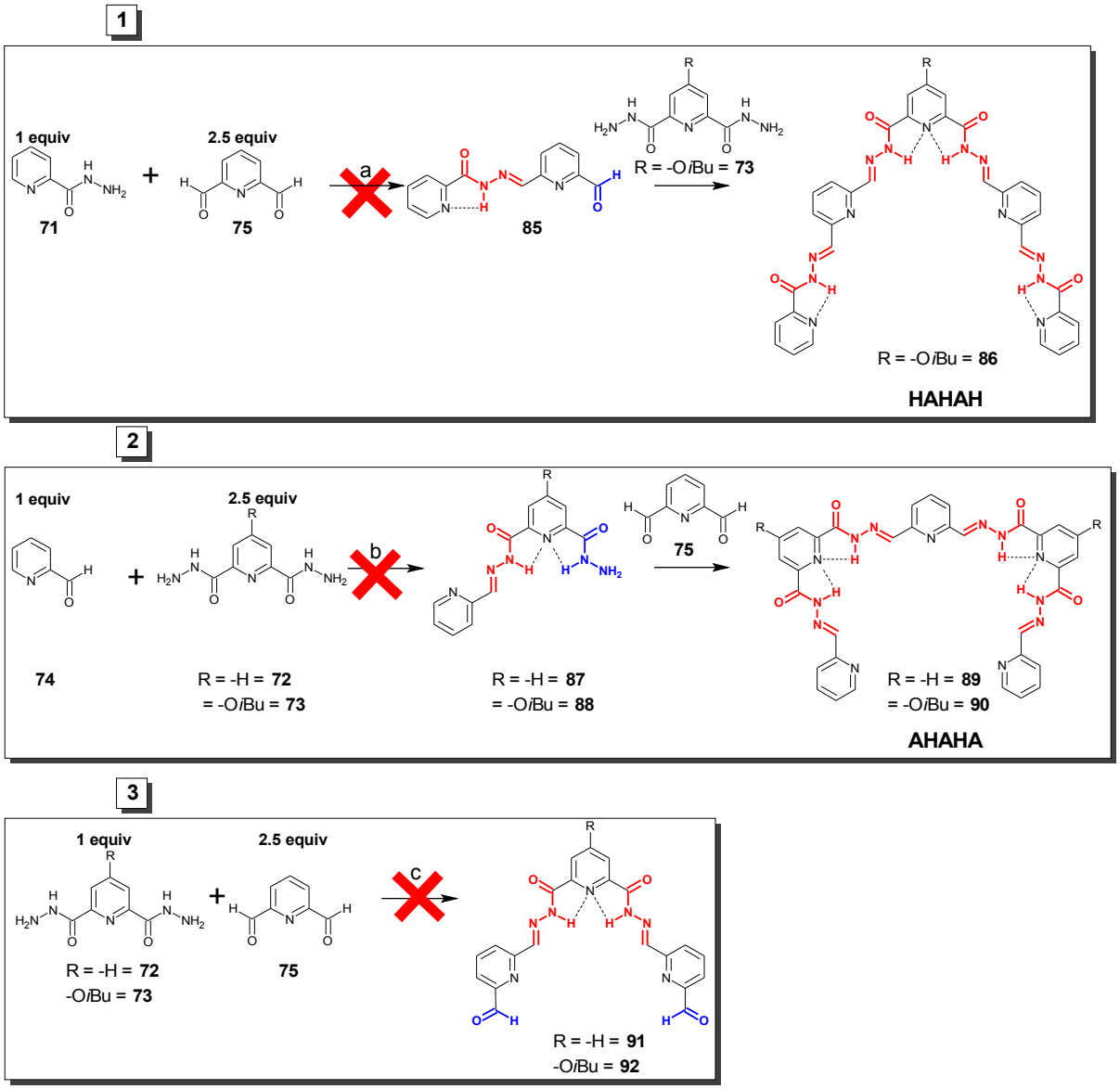
Scheme 4: Synthesis of trimers; HAH trimer **82**: *Synthesis 1* (a) EtOH, Reflux, (65% yield). AHA trimers **83** and **84**: *Synthesis 2*. (b) EtOH, Reflux, (68% and 87% yield respectively).

4.1.3 Elongation: Pentamers using bi-functional central units

It was intriguing to further observe this phenomenon by creating longer oligomers such as pentamers. The second objective was to investigate the effect of including multiple acylhydrazone pyridine motifs contained in an oligomer. It was therefore necessary to devise a synthetic route to achieve a variety of such pentamers. Initial investigations into the desymmetrisation of acylhydrazone pyridine pentamers proved unsuccessful using bi-functional central monomer units and resulted in trimers.

An initial attempt to synthesise the pentamer HAHAH **86** entailed reacting 2-pyridine carbohydrazide **71** (1 equivalent) and pyridine-2,6-dicarbaldehyde **75** (2.5 equivalents) in ethanol and heating to 65 °C. *(Scheme 5, Synthesis 1)*. It was expected that reacting excess dialdehyde with mono-hydrazide could easily form dimer **85**. However, the HAH trimer **82** was obtained instead. Similarly the same synthesis concept was applied to the pentamer AHAHA **89**. Compound 2-pyridine carboxaldehyde **74** (1 equivalent), was reacted with pyridine-2,6-dicarbohydrazide **72** (2.5 equivalents) in ethanol, and heated to 65 °C in an attempt to form dimer **87** *(Scheme 5, Synthesis 2)*.

As before, a trimer (AHA trimer **83**) was obtained with excess unreacted dihydrazide starting material **72**. To further confirm that this route was not viable, the above synthesis was repeated, with a solubilising isobutoxy side chain present on the dihydrazide **73**. It was hoped that this would increase solubility and reduce the possibility of producing trimers. The monomer, 2-pyridine carboxaldehyde **74** (1 equivalent), and 4-(isobutoxy)pyridine-2,6-dicarbohydrazide **73** (2.5 equivalents) in ethanol, were heated to 65 °C *(Scheme 5, Synthesis 2)*. However once again, the pentamer **90** was not obtained and a trimer formed resulting in product **84**. Reacting two bi-functional groups, the dihydrazide (**72** or **73**) and dialdehyde **75** in ethanol under reflux did not result in the desired dialdehyde trimer (**91** or **92**) but a multitude of polymeric products *(Scheme 5, Synthesis 3)*. Thus it was impossible to isolate a pure product from this reaction mixture.



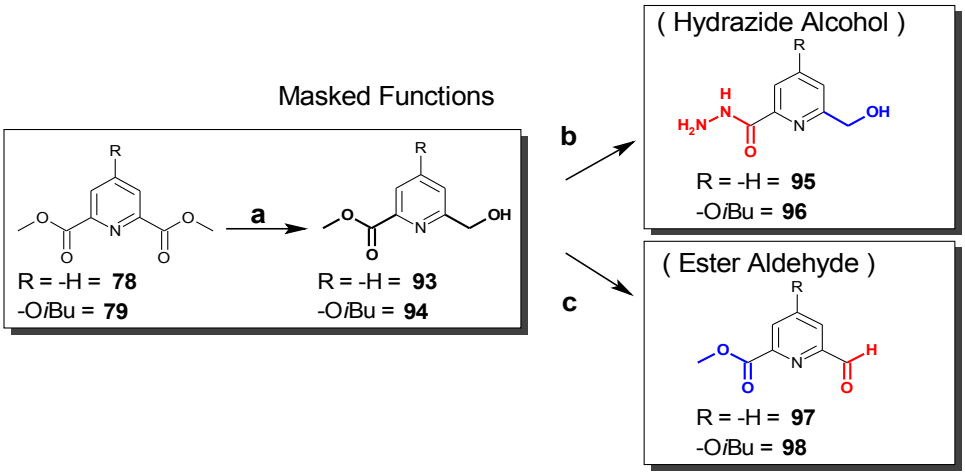
Scheme 5: Attempted pentamer synthesis of **86**, **89/90** and **91/92**. Bi-functional monomers and mono functional terminal units (Synthesis's 1 and 2) failed to give the desired functionalised dimers **85** or **87/88**. Reacting two different bi-functional monomers **72/73** with **75** failed to result in **91/92**, instead uncontrollable polymerisation resulted. (a); (b); (c) EtOH, 65 °C.

An explanation for this failure to form pentamers using the above method could be the existence of an equilibrium between starting materials and the dimer/trimer. The dimer may be formed and it might react with the terminal monomer to form the less soluble trimer compound which precipitates. It could be possible to address this issue by using DMSO or DMF as solvent instead of ethanol. If DMSO is used, it might also be possible to add water and obtain the desired product. These conditions remain to be investigated. However, the failure to obtain a pentamer from the above synthetic conditions suggested a need for a more sophisticated synthetic approach.

4.2 Dis-symmetrical Units

4.2.1 Monomers with masked functions

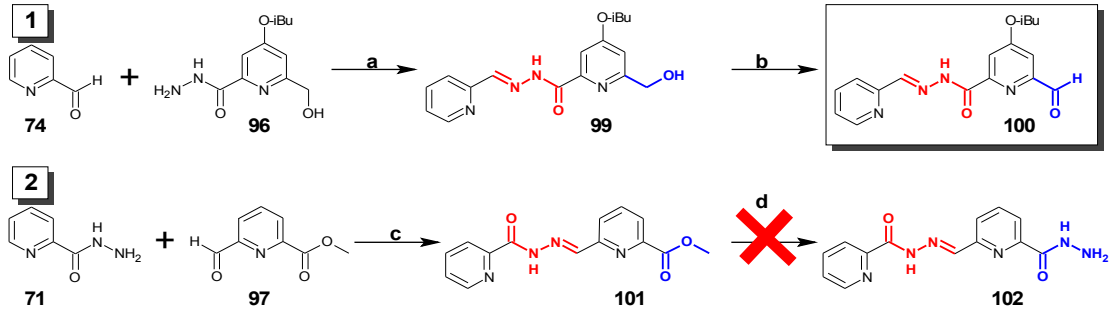
As discussed in the Design section (Section 3) it was decided to explore desymmetrisation of the pyridine unit to provide an alternative synthetic route in order to ease oligomer elongation and to investigate intramolecular hydrogen bonding *(Scheme 6)*. The initial diester **79** was mono reduced in the presence of sodium borohydride in anhydrous DCM and MeOH according to the published protocol of methyl 6-(hydroxymethyl)-4-isobutoxy-2-pyridine carboxylate **94**.[119,120] Advantageously, from here, two different desired building blocks could be obtained. The synthesis of the [mono hydrazide], 6-(hydroxymethyl)-4-(isobutoxy)pyridine-2-carbohydrazide **96**, and [mono aldehyde], methyl 6-formyl-4-(isobutoxy)pyridine-2-carboxylate **98**, were successful based on synthetic protocols established for the previous monomers **73** and **75**. Oxidation of the alcohol **94**, using MnO_2 in 1,4-dioxane resulted in the mono-aldehyde **98** (58% yield). Reacting monomer **94** with hydrazine monohydrate resulted in the desired mono-hydrazide **96** (99% yield). The synthesis was carried out for the monomer with and without an isobutoxy side chain. Overall the method to introduce mono-hydrazide (**95** or **96**) and mono-aldehyde (**97** or **98**) from a monomer (**93** or **94**) containing masked functions proved successful.



Scheme 6: Desymmetrisation of the diester **78/79** to form the monomer **93/94** containing masked functions. (a) $NaBH_4$, DCM, 0 °C, MeOH, (76% yield for **93**) and (78% yield for **94**). (b) $NH_2NH_2.H_2O$, MeOH, 0 °C, (99% yield for **95** and **96**); (c) MnO_2, 1,4-dioxane, 100 °C, (93% yield for **97**) and (58% yield for **98**).

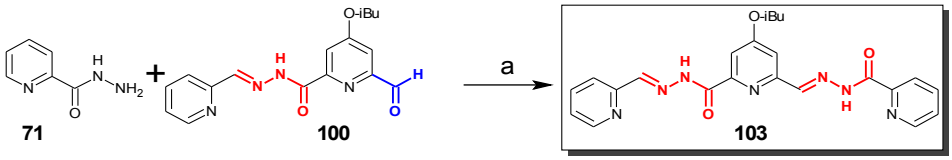
4.2.2 Dimers & Trimers: masked functionalised monomer

To overcome the previous challenges encountered for oligomer elongation, the dis-symmetrical monomer, methyl 6-(hydroxymethyl)-4-isobutoxy-2- pyridine carboxylate **94**, was employed. The addition of 2-pyridine carboxaldehyde **74** to mono hydrazide **96** proceeded well in methanol *(Scheme 7, Synthesis 1)* and the dimer product **99** (88% yield) was obtained after chromatographic purification. Subsequent oxidation of the dimer **99** was successful using manganese (IV) oxide in refluxing 1,4-dioxane, affording the desired dimer aldehyde **100** (75% yield). Coupling of 2-pyridine carbohydrazide **71** with mono aldehyde **97** in methanol also proved straightforward, forming the acylhydrazone dimer ester **101** (46% yield). *(Scheme 7, Synthesis 2)*. However, when the acylhydrazone dimer ester **101** was reacted with hydrazine monohydrate (0 °C in methanol) difficulties were encountered. This was thought to be due to hydrazone bond cleavage occurring, resulting in recovery of the starting monomer 2-pyridine carbohydrazide **71**. NMR integration showed the crude mixture contained approximately 16% of product **102**, which could not be obtained after column chromatography. To further confirm that bond cleavage occurred, the crude "dimer" mixture **102** was reacted with 2-pyridine carboxaldehyde **74** and the resulting product was found to correspond to the AH dimer **81** by ^{1}H NMR analysis. This is similar to the bond cleavage observed with the pyridine G clamps (discussed in Chapter 2, Section 4.2, *Scheme 28*) when hydrazine was used as a reactant. With this knowledge at hand, it was then possible to avoid reactions which involved hydrazine monohydrate in the later stages of the synthetic route. For improved synthetic ease, a solubilising side chain was incorporated.



Scheme 7: Dimer synthesis from the dis-symmetrical monomers. *Synthesis 1:* (a) MeOH, RT overnight, (88% yield); (b) MnO_2, 1,4-dioxane, 100 °C, (75% yield). *Synthesis 2:* (c) MeOH, RT overnight, (46% yield); (d) $NH_2NH_2.H_2O$, MeOH, 0 °C, (**102** not obtained due to hydrazone cleavage).

From the above synthesis, a dis-symmetrical trimer **103** could be synthesised *(Scheme 8)*. Dimer aldehyde **100** was reacted with 2-pyridine carbohydrazide **71** in DCM resulting in the desired dis-symmetrical trimer **103** (52% yield). Due to intramolecular hydrogen bonding, the ^{1}H NMR analysis contained many signals and the exact purity of the product could not be conclusively determined. Purification by column chromatography was carried out four times and despite isolation of one product, this soon had other components after analysis by TLC and NMR. It is thought that this is due to trans-cis isomerisation.



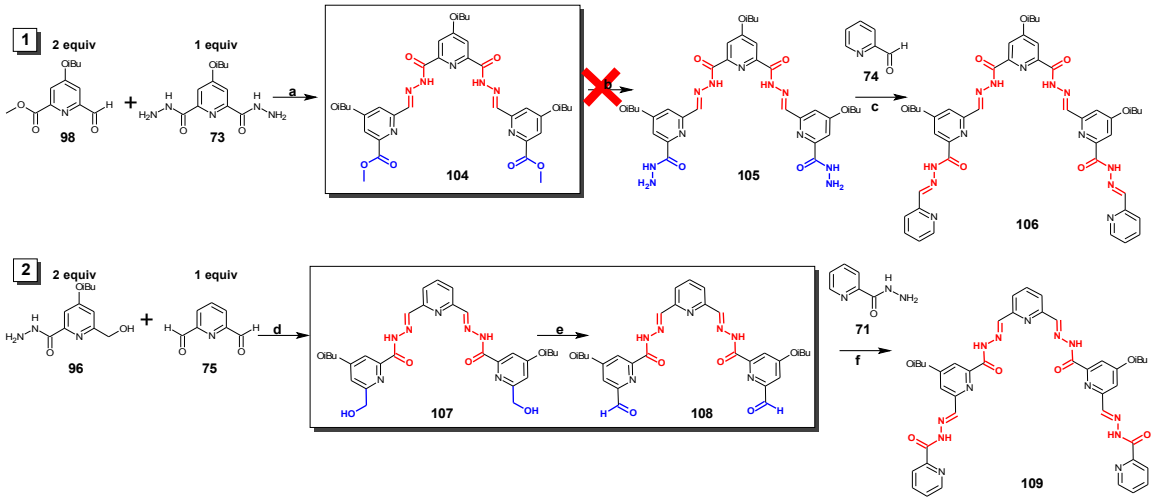
Scheme 8: Synthesis of the dis-symmetrical trimer. (a) DCM, RT overnight, (52% yield).

4.2.3 Elongation: Pentamers using masked functionalised monomers

The second objective was to investigate the effect of including multiple acylhydrazone pyridine motifs in a long oligomer. Preliminary studies showed it was not possible to obtain a pentamer by use of symmetrical bi-functional monomers (Section 4.1.3). Subsequently an alternative route utilising dis-symmetrical monomers for oligomer elongation was established.

4.2.3.1 Divergent pentamer synthesis

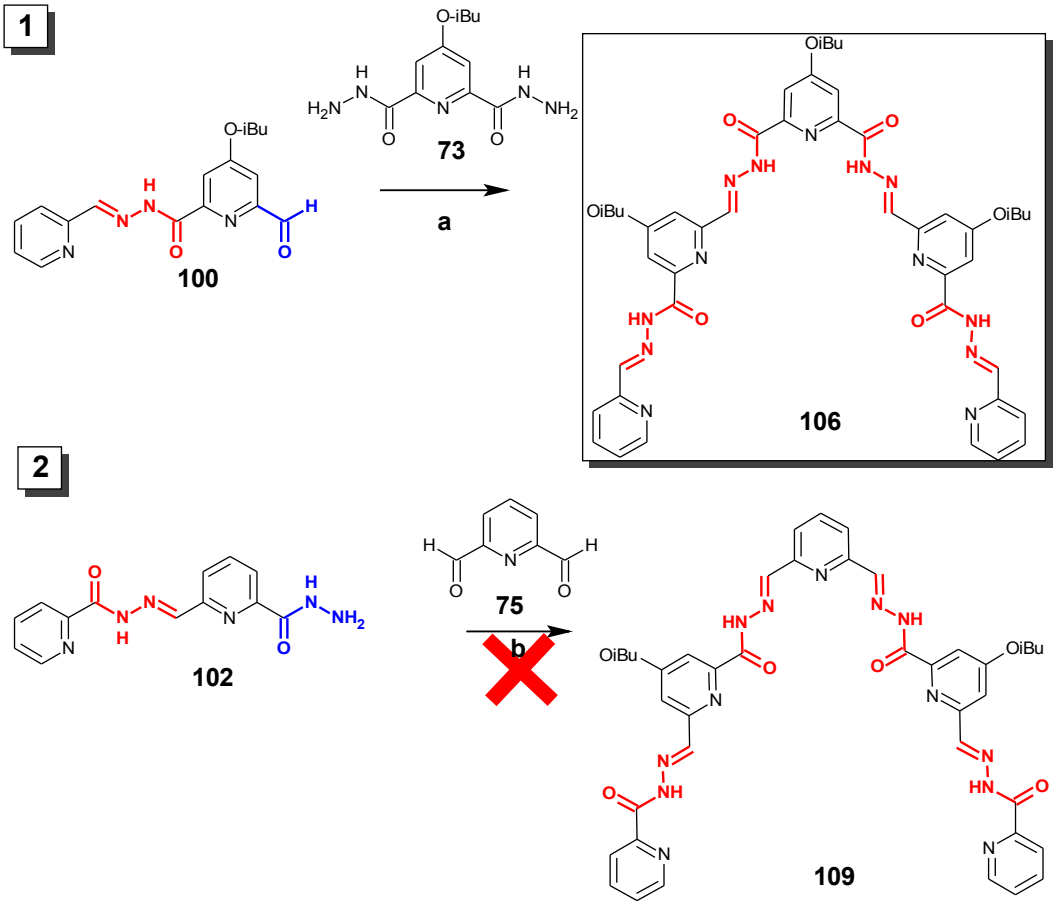
The initial synthetic route investigation involved a divergent route strategy. The mono aldehyde **98** (2 equivalents) was reacted with one equivalent of the dihydrazide **73** in MeOH/DCM to give the diester trimer **104** (98% yield). As mentioned for the dimer synthesis **102**, the use of hydrazine monohydrate with trimer **104** resulted in hydrazone cleavage and as a result, it was not possible to obtain the dihydrazide trimer **105** via this route *(Scheme 9, Synthesis 1)*. In contrast, the divergent route proved more viable for **107** despite some solubility issues. To the monomer **96** (2 equivalents) was added one equivalent of pyridine-2,6-dicarbaldehyde **75** to form the desired trimer **107** *(Scheme 9, Synthesis 2)*. The oxidation step with activated manganese (IV) oxide was carried out in refluxing 1,4-dioxane over five days, and NMR analysis of the crude material showed signals suggesting the presence of the aldehyde **108**. However a low yield was obtained after filtration on celite, perhaps suggesting poor solubility of the product. Thus, there was not enough material to react **108** with **71** to obtain **109**. This procedure therefore requires optimisation. SeO_2 might also be more effective instead of MnO_2.



Scheme 9: Attempts at elongation via a divergent synthetic route using a masked functionalised monomer: *Synthesis 1:* a) MeOH/DCM, (98% yield). b) $NH_2NH_2.H_2O$, MeOH, 0 °C; Suspected that hydrazone cleavage occurred. *Synthesis 2:* d) MeOH/DCM, (90% yield). e) MnO_2, 1,4-dioxane, 100 °C, Solubility poor. f) DCM. **109** was not obtained pure and it is suggested that optimisation of **108** using SeO_2 in step **e** is investigated.

4.2.3.2 Convergent pentamer synthesis

To improve the synthesis of **106**, a convergent route was investigated. The use of hydrazine monohydrate at early stages of synthesis is more desirable, for reasons which have been previously described. The bi-functional monomer **73** (1 equivalent) was stirred in DCM with the dimer aldehyde **100** (2 equivalents) overnight *(Scheme 10, Synthesis 1)*. Column chromatography was then carried out to afford the pentamer **106** in 60% yield. Due to configurational changes, the NMR showed the presence of cis products. However the pentamer **109** cannot be obtained via the convergent method *(Scheme 10, Synthesis 2)*. This is due to bond cleavage when using hydrazine monohydrate at later stages in synthesis.



Scheme 10: Attempts at elongation via a convergent synthetic route using a masked functionalised monomer: *Synthesis 1:* a) DCM, (60% yield). *Synthesis 2:* b) This route could not be carried out because **102** was not obtained due to hydrazone cleavage.

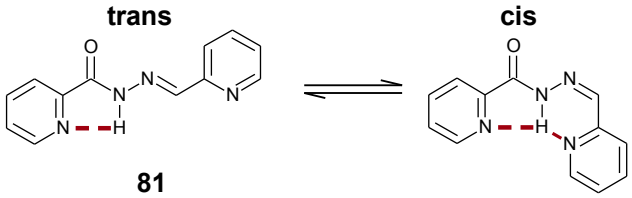
The synthesis of longer oligomers containing multiple hydrazones was found to be unexpectedly difficult. It was indeed possible to obtain both pentamers, however a different synthetic route needed to be implemented for both, in order to avoid hydrazone bond cleavage. For pentamer **106** a convergent synthetic route was preferred, whereas for pentamer **109**, a divergent synthetic route proved viable. The main problem in the pentamer synthesis is the purification of the final product. In addition, many configurations were observed by ^{1}H NMR analysis.

5 Analysis and Discussion

5.1 Cis/Trans Isomerism of Acylhydrazone Pyridine Motifs

5.1.1 Effects of intramolecular hydrogen bonding

The initial objective was to determine if a biased equilibria for cis configuration exists within a small system containing one or two acylhydrazone pyridine motifs *(Scheme 11)*. This cis configuration ability had not been reported for other hydrazone motifs which usually have a preference for trans configuration during the following investigations. Hydrogen bonding interactions was expected to contribute to this phenomenon. However since then Lehn et al. [121]and Dumitru et al.[122] published a similar observation of this phenomenon. What differs with the following investigations is that intramolecular hydrogen bonding is utilised compared to the published findings which incorporate metal ions or photochemical stimuli. The strength of a hydrogen bond depends on the immediate environment: the medium (solvent), chemical composition, pH and temperature. With these environmental factors in mind, ^{1}H NMR studies examined the cis configuration of the AH dimer **81** in $CDCl_3$ *(Scheme 11)*. This chosen solvent has no competition for hydrogen bonding.



Scheme 11: Equilibria of the AH dimer **81** between trans and cis configuration

The AH dimer in $CDCl_3$ (with 1.75 equivalents of acetic acid) was characterised by ^{1}H NMR analysis over a period of three weeks. Note that $CDCl_3$ was filtered on a bed of alumina to remove any acidic residue prior to this study. Initially the trans configuration is preferred, with a signal at 11.1 ppm corresponding to trans N-H acylhydrazone and supported by X-ray crystallography results. A dimer AH crystal was obtained in toluene/pentane conditions *(Fig. 23)*. X-ray crystallography of the AH dimer crystal displayed trans conformation. This corresponds to the observations of the initial ^{1}H NMR spectra.

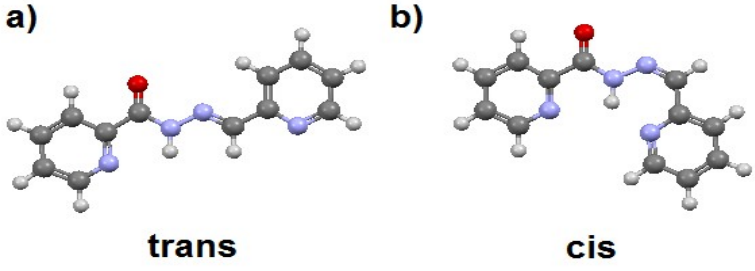
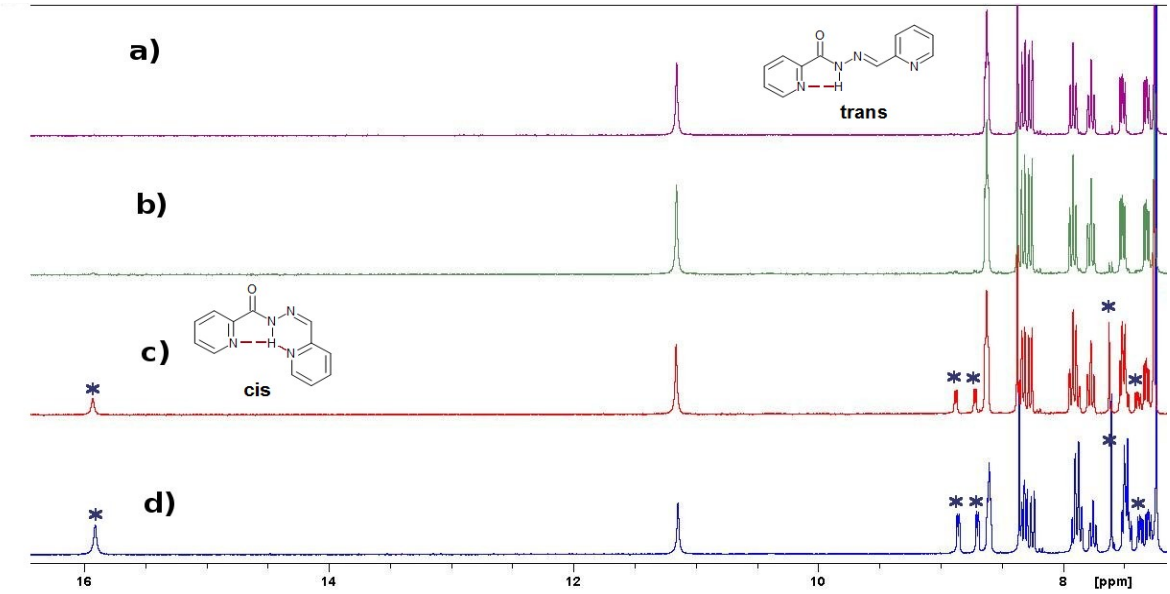


Figure 23: The AH dimer, a) trans, (X-ray crystal structure); b) cis, (molecular model).

There was no significant chemical shift change for the AH dimer over 24 hours *(Spectra 1)*. However the emergence of a new signal at 15.9 ppm on the second day *(Spectra 1,c)*, suggested the appearance of the cis configuration and this is likely due to intramolecular hydrogen bonding between the hydrazone and pyridine groups. Such a downfield shift is indicative of a change in configuration from trans to cis. There is a steady increase in the AH behaviour towards cis configuration proportional to a decrease in trans configuration. An equilibrium of 50:50 mixture trans:cis configuration is achieved after one week *(Spectra 1,d)*.



Spectra 1: ^{1}H NMR of AH in $CDCl_3$ with 1.75 eq AcOH; a) 30 mins, b) 3 hrs, c) 2 days, d) 1 week * represents cis configuration a) 0%,b) 7%, c) 23% and d) 45% respectively

This initial study lays the foundations for this hypothesis in which the equilibrium is biased towards the cis configuration within a system containing one acylhydrazone pyridine motif. The initial state of the AH dimer determines how the ^{1}H NMR results develop over time and coincides with the X-ray crystallography results. For instance, after the AH dimer purification by column chromatography in DCM/MeOH the trans configuration product was obtained. Perhaps with different purification conditions, the cis isomer might have been isolated. Therefore beginning with a cis isomer it is expected the X-ray crystallography and ^{1}H NMR analysis to show the cis configuration as the initial starting point.

The question now, is how will this phenomenon translate into a system containing two or more acylhydrazones? Not only that, but how does intramolecular ring size of the acylhydrazone motif influence this equilibrium? A review of AHA (**83, 84**) and HAH **82** set to distinguish the importance of these factors *(Fig. 24)*.

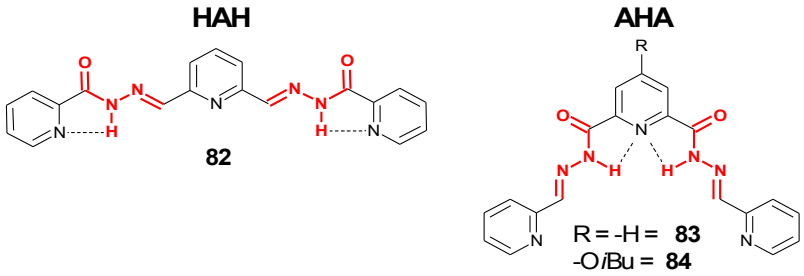


Figure 24: HAH trimer **82** and the AHA trimer **83/84**

The trimer samples in $CDCl_3$ solvent (with 1.75 equivalents of acetic acid) were monitored over a period of three weeks to observe if any configuration changes occurred. As the AHA ^{1}H NMR spectra shows *(Spectra 2,a)*, initially the two amide's, for trans-trans configuration are represented by one singlet at 11.73 ppm. This reflects the AHA trans-trans configuration observed for the X-ray crystal structure which was obtained in DMF/diethyl ether conditions *(Fig. 25)*.

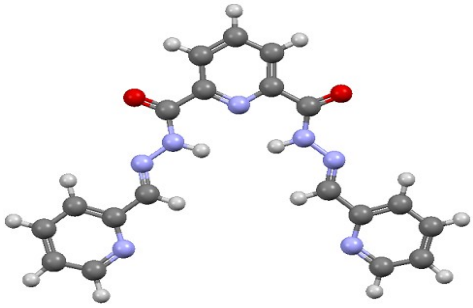
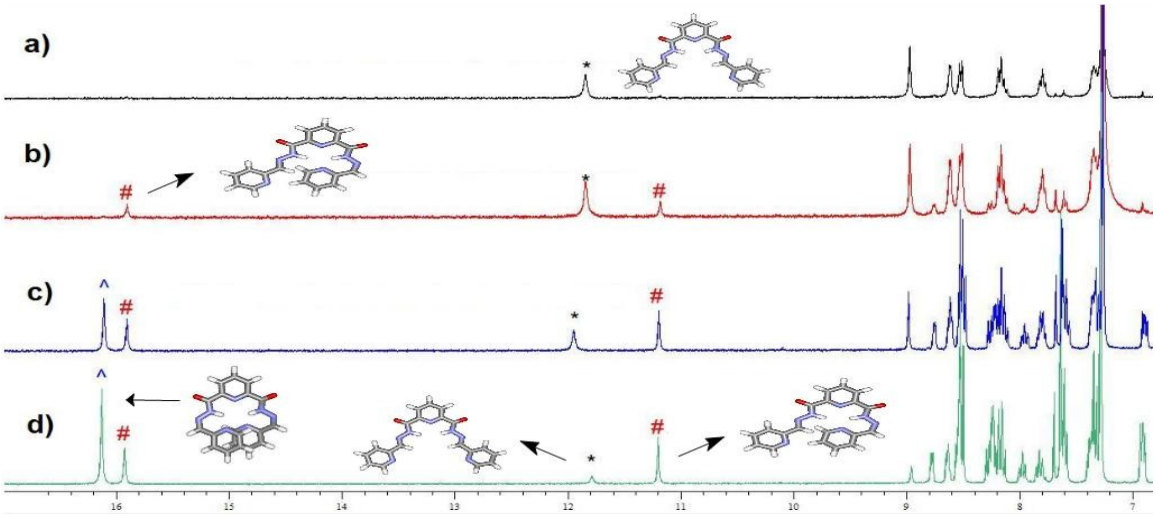


Figure 25: X-ray crystal structure of the AHA trimer in trans-trans configuration.

Continuing with the ^{1}H NMR analysis *(Spectra 2,b)*, new signals at 11.17 ppm and 15.89 ppm began to appear and these were expected to be representative of the trans-cis configuration *(Fig. 26,b,e)*. After three days of slow exchange, a signal at 16.10 ppm was present *(Spectra 2,c)* and it was expected that this corresponded to the cis-cis isomer *(Fig. 26,c,f)*. Thus AHA was initially 100% trans-trans configuration in $CDCl_3$ *(Fig. 26,a,d)* and after four weeks the equilibrium was 7% trans-trans, 37% trans-cis, 56% cis-cis *(Spectra 2,d)*. The increased stability of the cis-cis configuration could be considered due to interplanar π-π interactions of the overlapping aromatic rings with minimal steric hindrance or strain. The five-membered intramolecular rings at the central pyridine unit seem to contribute to the overall configuration and also enhance stability. Over time, the cis-cis isomer becomes increasingly predominant. Note that when calculating the ratios of trans-trans, trans-cis and cis-cis isomers, one must consider that the trans-trans (*) and cis-cis (^) singlet integration is representative of two amide signals whereas trans-cis (#) has two separate amide signals.



Spectra 2: ^{1}H NMR of the AHA trimer in $CDCl_3$ with 1.75 eq AcOH a) 30 mins, b) 6 hours, c) 3 days, d) 3 weeks. ***** trans-trans; **#** trans-cis; **^** cis-cis configurations.

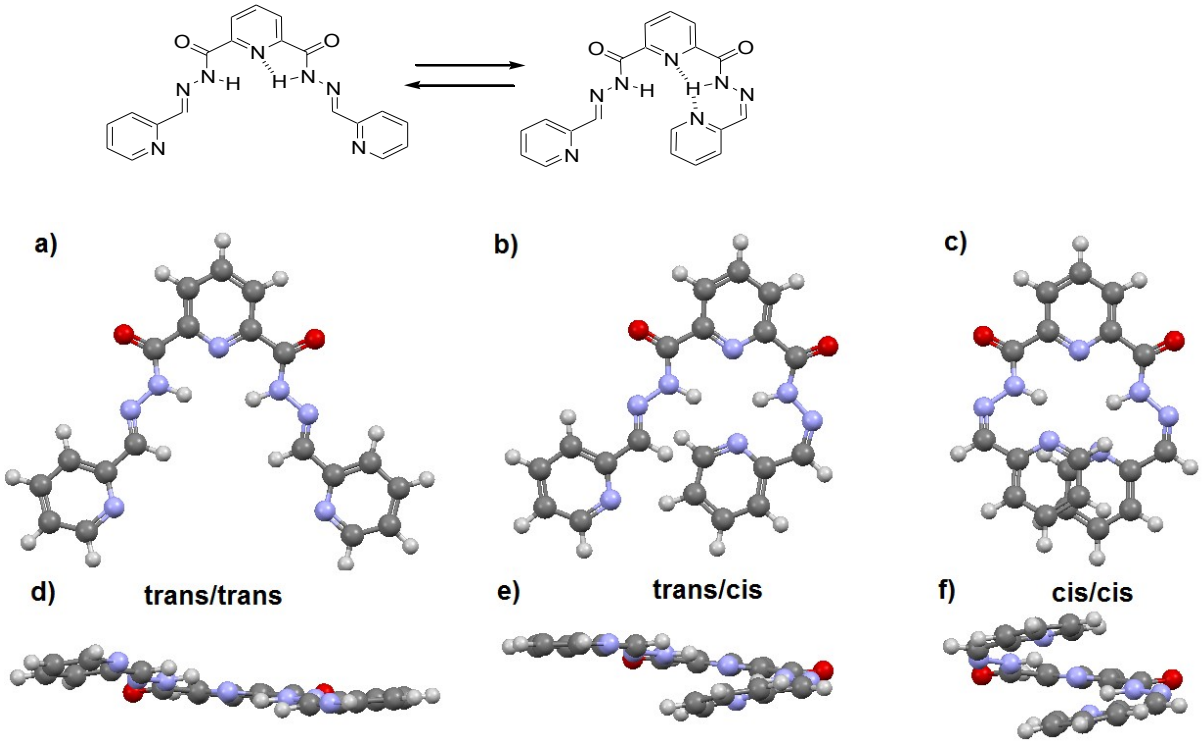
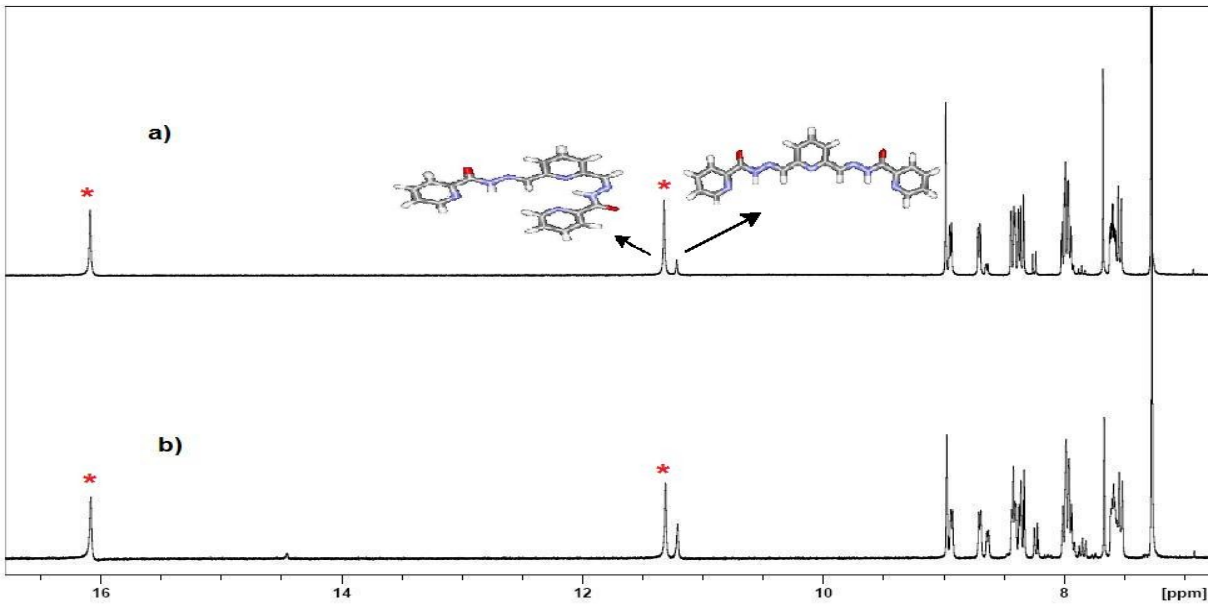


Figure 26: The AHA trimer configurations: **a)** trans-trans; **b)** trans-cis; **c)** cis-cis; **d)** trans-trans (side view); **e)** trans-cis (side view); **f)** cis-cis (side view).

In contrast, to AHA, the initial HAH ^{1}H NMR spectra in $CDCl_3$ *(Spectra 3,a)* showed a preference for the trans-cis configuration *(Fig. 27,b,e)*, 87%, represented by 2 separate singlets at 11.30 ppm and 16.06 ppm. This is in comparison to the diminutive trans-trans singlet, 13%, at 11.19 ppm *(Fig. 27,a,d)*. After three weeks, no new signals are observed, whereas the trans-trans signal increases slightly to 26% *(Spectra 3,b)*. The key difference between AHA and HAH, is the location of the five-membered and six-membered hydrogen bonding intramolecular rings. AHA is capable of having two central five-membered

intramolecular hydrogen bonding rings whereas HAH, is capable of having two terminal five-membered intramolecular hydrogen bonding rings. The possibility of a two central six-membered ring system might occur for HAH. However, from molecular modelling studies cis-cis HAH configuration *(Fig. 27,c,f)* could involve increased strain, steric hindrance and repulsive interactions between the close proximity of the two opposite N-H groups.



Spectra 3: ^{1}H NMR of HAH trimer in $CDCl_3$ with 1.75 eq AcOH; a) Initial; b) 3 weeks * trans-cis configuration.

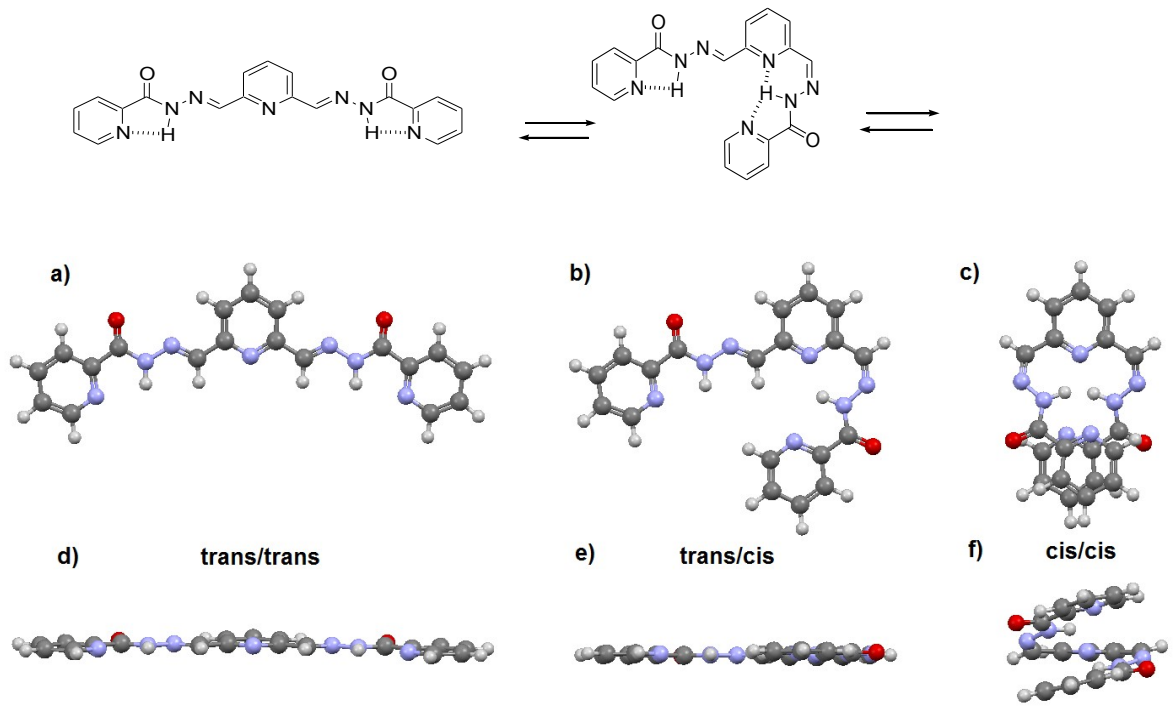


Figure 27: The HAH trimer configurations: **a)** trans-trans; **b)** trans-cis; **c)** cis-cis; **d)** trans-trans (side view); **e)** trans-cis (side view); **f)** cis-cis (side view).

The initial HAH trimer ^{1}H NMR spectra reflects the X-ray crystallography results. The HAH crystal was obtained in toluene/pentane solvent conditions and showed a preference for trans-cis configuration *(Fig. 28)*. This is the first crystal structure evidence to support that indeed the cis configuration exists. Note the AH crystal was obtained in the same toluene/pentane solvent conditions as used for the HAH trimer.

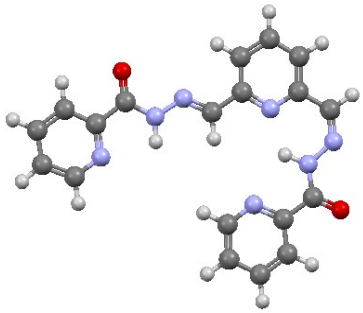
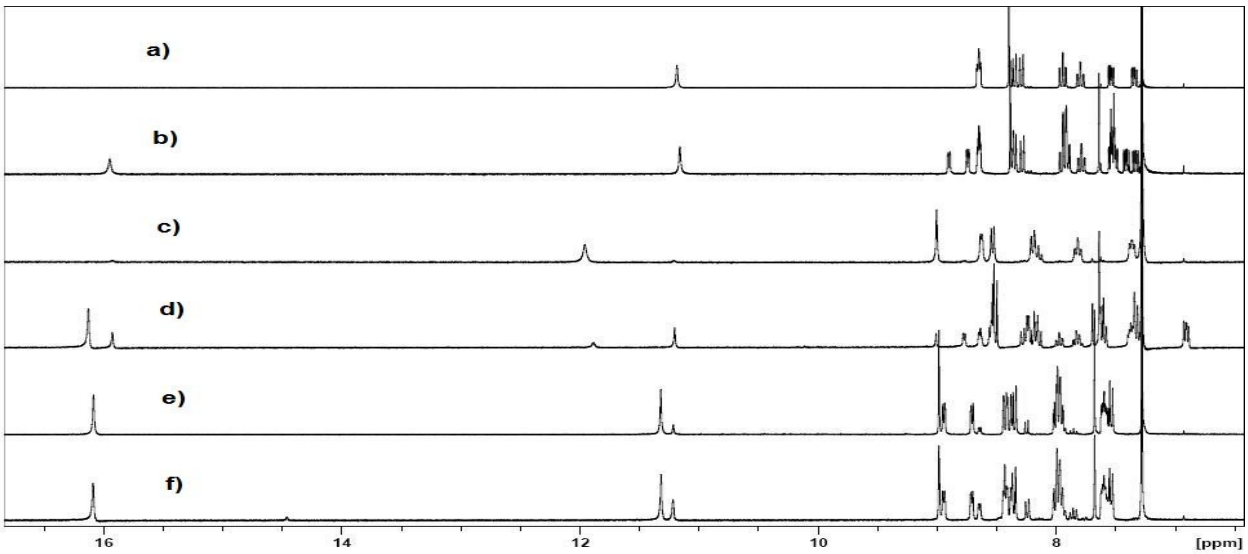


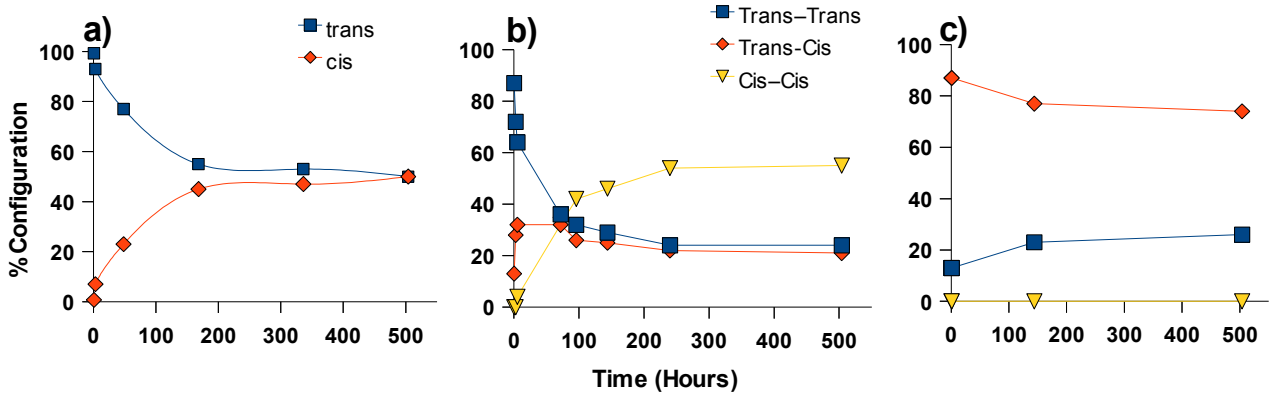
Figure 28: X-ray crystal structure of the HAH trans-cis configuration

The key observations between the acylhydrazone pyridine motifs are as follows. The AH dimer is initially 100% trans *(Spectra 4,a)* with the appearance of a second species of cis configuration (50%) after one week *(Spectra 4,b)*. The AHA trimer, is initially 100% trans-trans configuration in $CDCl_3$ *(Spectra 4,c)*. Despite the initial preference for trans-trans configuration exhibited by the AHA trimer, the NMR study revealed three species present, with a transitional shift in equilibria possible through to trans-cis and finally that of cis-cis configuration. After four weeks the AHA equilibrium is 7% trans-trans, 37% trans-cis, 56% cis-cis isomer *(Spectra 4,d)*. In contrast, the HAH trimer has an initial preference for trans-cis configuration (87% present *Spectra 4,e*) with only a slight change after one week to 77%. The second HAH species had a slight increase in trans-trans configuration to 23%, which remains stable with no change after four weeks *(Spectra 4,f)*. All of the initial ^{1}H NMR spectra is consistent with the crystal structures.



Spectra 4: ^{1}H NMR comparison between the dimer and trimer acylhydrazone pyridine motifs. 1.75 equivalents AcOH in $CDCl_3$: a) AH, Initial; b) AH, 6 weeks; c) AHA, Initial; d) AHA, 4 weeks; e) HAH Initial; f) HAH 4 weeks.

Graphs plotting the % increase of cis configuration versus time (hours) to compare all three motifs shows interesting observations derived from the ^{1}H NMR analysis *(Graph 1)*. The AH dimer shows a progressive transition from trans to cis configuration based on the integration of the N-H acylhydrazone at 15.9 ppm over three weeks *(Graph 1,a)*. The AHA trimer also shows a progressive decrease in the amount of trans-trans configuration present with a slight increase in trans-cis configuration, only to be dominated by a preference for cis-cis configuration *(Graph 1,b)*. As previously discussed, there is little variation with the HAH trimer which has a preference for trans-cis configuration *(Graph 1,c)*. In complete contrast to the AHA trimer, HAH does not exist at all in the form of cis-cis. These initial studies in deuterated chloroform, on the effects of available intramolecular hydrogen bonding within acylhydrazone pyridine motifs, suggest that indeed a cis configuration is biased within these systems.



Graph 1: 1.75 eq AcOH in $CDCl_3$, % Configuration change versus Time (Hours); a) AH; b) AHA; c) HAH

Despite the fact that both trimers are made up of similar building blocks, the way in which they are connected contributes to a difference in overall behaviour and properties. This fundamental characteristic gives rise to isomers which can have extraordinary complexity coupled with the potential ability to master or manipulate these combinations in a controlled environment. Obtaining a desired isolated isomer may be possible based on the advantage of this complexity, with folding as a driving force in proof reading.

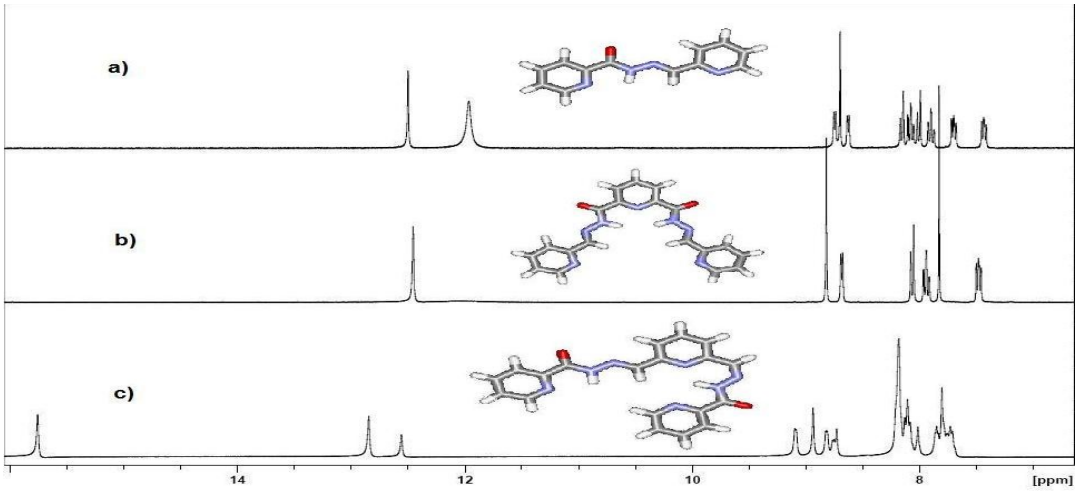
Our overall conclusions obtained from these initial ^{1}H NMR studies suggest that cis configuration exists for the AH dimer and the trimers AHA and HAH. The emergence of a new proton signal at 16 ppm is a feature of cis configuration. The NMR data shows an equilibrium for AH and AHA in a given solvent, $CDCl_3$. For trimer AHA, there are three different species that can exist. The trimer HAH, however, has an overall preference for one species, the trans-cis configuration, with trans-trans only as a minority. One isomer, the cis-cis configuration, is completely absent from the trimer HAH samples and appears to not be feasible. X-ray crystallography shows a trans-cis configuration for the HAH trimer, however the trans configuration is observed for AH and AHA. This crystallography data is substantial

in telling us that the solid state preference for configuration is coupled to the initial state as seen in the NMR analysis. The AH dimer was purified by column chromatography in DCM/MeOH whereas the AHA and HAH trimers were precipitated from EtOH. Further validation regarding purification techniques and solvent might result in different isomer crystals. In addition, using similar conditions as NMR analysis in which growth of various isomers was observed in these experiments could be applied to obtain different X-ray crystal structures. NMR experiments such as NOESY/ROESY or deuterium exchange could provide further substantial support towards the existence of cis configuration. With these extraordinary findings of favourable interactions and cis configuration, attention was give to the effects of solvent and surrounding environment on the equilibria of acylhydrazone pyridine motifs.

5.1.2 Solvent effects

Solvent is an important medium in which molecules can react with other molecules or encounter behavioural changes as a result of solvophobic effects. In non-polar solvents, hydrogen bonding interactions are much more significant whereas in polar solvents these interactions are destroyed within supramolecular constructions. Therefore it is interesting to investigate the strength of conformational stability in various solvents. It is often found that one conformer or rotamer is predominant in one medium but not in another. The dynamic system is in competition with solvation interactions and it is important to investigate since solvent is in a large molar excess. Thus it is of interest to discuss the configurational changes and an affinity for cis configuration across different solvent media. Differences in environment can have an impact on energetically favoured conformations. For this reason, studies in various solvents was conducted to investigate the behavioural effects on acylhydrazone pyridine motifs.

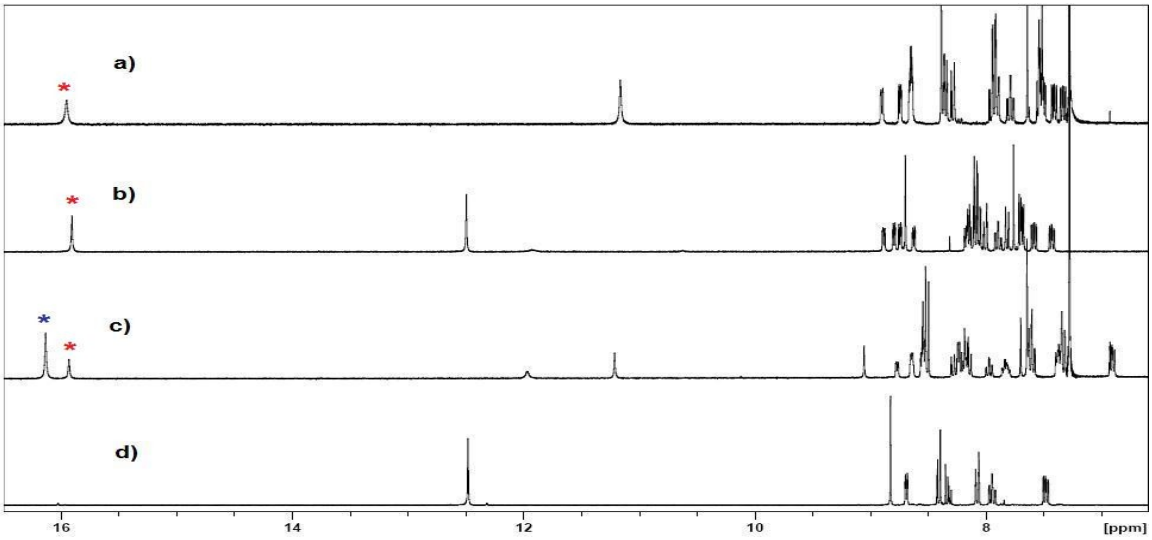
A comparison of the equilibrium studies discussed in the previous section *(Section 5.1.1)* in the non-polar solvent $CDCl_3$, was made to the same molecules in the polar aprotic solvent DMSO-d_6 *(Spectra 5).* The aim of this, was to recreate the configurational equilibrium changes from trans to cis. The AH dimer in DMSO-d_6 showed no cis configuration even after three weeks due to the denaturing effect and competition for hydrogen bonding in DMSO *(Spectra 5,a)*. After seven weeks, 94% trans and 6% cis was present. Thus to observe cis configuration, choice of solvent is important in order to enable the available intramolecular hydrogen bonding to form a cis biased configuration in equilibria. The same was observed for the AHA trimer in DMSO-d_6 with only trans configuration observed *(Spectra 5,b)*. In contrast, the HAH trimer still exhibits trans-cis configuration *(Spectra 5,c)* even in DMSO-d_6, possibly confirming the strong stability and closeness to equilibrium compared to that of AH or AHA.



Spectra 5: ^{1}H NMR, DMSO-d_6 with 1.75 equivalents of AcOH after 3 weeks (11.9 ppm = AcOH) a) AH dimer, 100% trans; b) AHA trimer, 100% trans; c) HAH, trans-cis configuration dominant.

Based on the solvent-dependence revelations, further interesting behaviours may be observed when $CDCl_3$ is removed from the AH dimer sample previously exhibiting a 50:50 mixture of trans-cis configuration, and replaced by the denaturing solvent DMSO-d_6 *(Spectra 6)*. One might expect the intramolecular hydrogen bonding to be diminished. In fact, the cis AH configuration remains after three weeks and is merely reduced from 50% to 41% *(Spectra 6,a,b)*. This suggests that equilibrium was not reached and must lie between 100% and 60% trans. In contrast, a majority (94%) of trans configuration was observed for the AH dimer in DMSO-d_6 solvent *(Spectra 5,a)* and not much change was observed after seven weeks. However, the mixture of trans-cis still exists in DMSO-d_6 solvent replacement *(Spectra 6,b)*, and it was speculated that the equilibrium was closer to 100% trans in this case.

Like the AH dimer, it was of interest to investigate this phenomenon in DMSO-d_6 denaturing solvent to observe its effect on this AHA trimer dynamic system *(Spectra 6,c,d)*. Similar to the previous experiment, $CDCl_3$ was removed and replaced with DMSO-d_6 to the three isomer-mixture. Surprisingly, unlike the AH dimer in which intramolecular hydrogen bonding and cis configuration remained, here the trans-cis and cis-cis configurations were completely destroyed and only trans-trans AHA trimer is observed *(Spectra 6,d)*. After 3 weeks, 94% trans-trans AHA trimer is present with 6% cis-trans AHA trimer. Thus the system appears to be reversible. This suggests that equilibrium was reached within the initial $CDCl_3$ solvent *(Spectra 6,c)*. Since there was no change with the HAH trimer in either the $CDCl_3$ or DMSO-d_6 solvent, this experiment was not carried out. As expected, the HAH trimer is close to thermodynamic equilibrium. After a suitably long wait a difference might be observed, however this would require a time period that is currently unknown, and may in fact range from months to years in order to see a change for the HAH trimer.



Spectra 6: ^{1}H NMR experiment; Deuterated chloroform removed and replaced by DMSO-d_6 (11.9 ppm = AcOH) a) AH in $CDCl_3$, 50% cis configuration b) AH sample a, replaced with DMSO-d_6, 3 weeks, 41% cis (* red); c) AHA in $CDCl_3$, after equilibrium was reached; (* blue = cis-cis; * red = trans-cis) d) AHA sample c, replaced with DMSO-d_6, 3 weeks, 94% trans configuration, 6% cis-trans configuration.

Other NMR solvents were investigated for configuration effects of the AHA trimer *(Table 1)*.

NMR Solvent	Type of Solvent	Property
$CDCl_3$ - Chloroform	Non polar	Inert solvent
DMSO-d_6 - Dimethyl sulfoxide	Dipolar aprotic	Do not act as H-bond donors since their C-H bonds are not strongly enough polarized. Large dielectric constant. (εr > 15)
CD_3COCD_3 - Acetone-d_6		
C_5D_5N - Pyridine-d_5	Aromatic Dipolar aprotic	
$C_6D_5OCD_3$ - Toluene-d_8	Aromatic Apolar aprotic	Inability to act as H-bond donor. Low dielectric constant. (εr < 15)

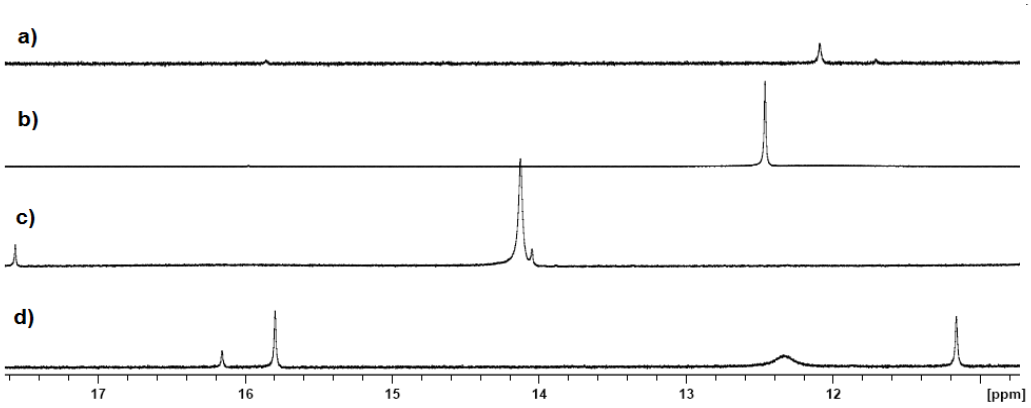
Table 1: Details and properties of the various NMR solvents investigated for their effects upon the AHA trimer

It is important to note that the various solvents investigated, included the dipolar aprotic solvent acetone-d_6, pyridine-d_5 and the apolar aprotic solvent toluene-d_8 *(Spectra 7,a,c,d)*. Intramolecular hydrogen bonding is lacking in these solvents, thus the trans-trans AHA configuration is the dominant species *(Spectra 7)*. Pyridine-d_5 and toluene-d_8 are aromatic and have *π-π* stacking influences. These results suggest the cis configuration of acylhydrazone pyridine motifs is possible in inert solvents, but in dipolar aprotic solvents heating is required to speed up the process.

In the polar aprotic solvent of acetone-d_6, only a minuscule amide signal was observed at 12 ppm and is thought to correspond to the trans-trans configuration (92%) with trans-cis configuration (8%) *(Spectra 7,a)*. Polar aprotic solvents such as DMSO-d_6 and pyridine-d_5 were thus investigated *(Spectra 7,b,c)*. Hydrogen bonding is lacking in these solvents, with an expected majority of the trans-trans AHA configuration. New signals for trans-cis (3%) and

cis-cis (5%) AHA configuration were obtained after heating the DMSO-d_6 sample to 80 °C for one day. However, these signals were minimal and little change was observed after further heating for three weeks.

Pyridine-d_5 is aromatic with *π-π* stacking potential. Within one hour at 80 °C in pyridine-d_5 a new trans-cis (38%) AHA signal at 17.6 ppm is observed *(Spectra 7,c)*. The emergence after one day in pyridine-d_5 at 80 °C of a signal at 17.8 ppm is thought to correspond to the cis-cis AHA configuration (12%) with trans–cis (50%) and trans–trans (38%). Little change is observed in this ratio of configuration after three weeks. Toluene-d_8 was shown to give interesting results for the AHA trimer. At room temperature, the trans-trans configuration was solely present. Within one hour of heating to 50 °C, the trans-trans signal (44%) had greatly diminished and the trans-cis signal (53%) was dominantly present with some cis-cis configuration (3%) *(Spectra 7,d)*. Little change was observed after heating to 80 °C for one week. However, the AHA seemed to be insoluble after this time in solution and a precipitate formed. In hindsight, more tests could have been carried out on this dominantly “trans-cis” AHA compound which might have been the pure precipitate that crashed out of solution. From these results it would be desirable to grow AHA crystals in toluene solvent in an effort to obtain a crystal structure of the trans-cis AHA configuration.



Spectra 7: ^{1}H NMR AHA trimer at room temperature, after heating the NMR tube to 50 °C 1 hour in an oil bath a) Acetone-d_6, 1.75 eq AcOH; b) DMSO-d_6, 1.75 eq AcOH; c) Pyridine-d_5, 1.75 eq AcOH; d) Toluene-d_8, 1.75 eq AcOH

The solvents such as pyridine-d_5 or toluene-d_8 are not completely inert to the equilibrium process. With increased time it could be possible to reach equilibrium, but currently it's unknown how long this would take. These experiments so far showed a “proof-of-concept” towards available intramolecular hydrogen bonding contributing towards a biased cis configuration in non-polar solvents with no competition for hydrogen bonding. However this bias in configuration is disrupted in polar aprotic solvents. The next experiments to be discussed deal with the chemical equilibria of this unique system. Attention will be given towards the various proportions between the trans and cis configurations.

5.2 Study of Equilibria

One prerequisite for understanding the study of equilibria is the evaluation of association free energies and the respective enthalpic and entropic contributions which are important to the overall free energy (ΔG) *(Fig. 29)*. This term refers to the portion of the system's energy which can perform work. The overall free energy equation is: $\Delta G = \Sigma(\Delta H) - T\Sigma(\Delta S)$.

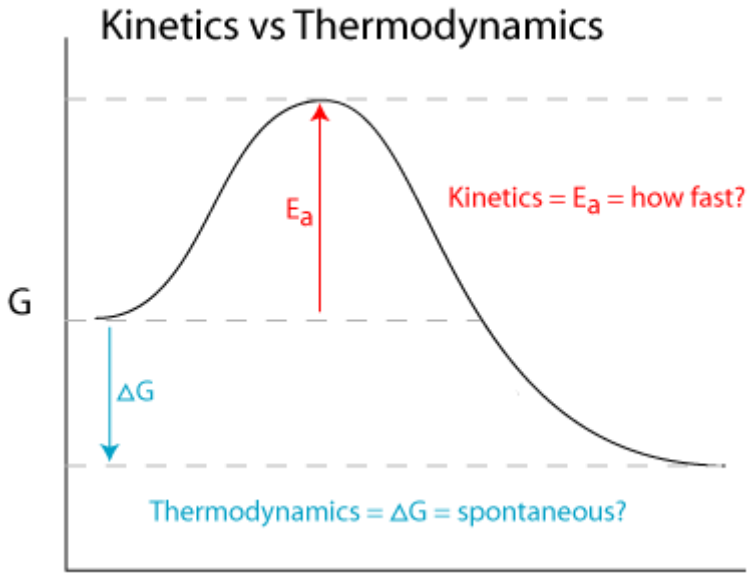


Figure 29: Kinetics vs Thermodynamics.

Kinetics is the energy of motion, whereas thermodynamics is the study of energy transformation *(Fig. 29)*. It is possible to manipulate experimental conditions to favour a preferred product: be it kinetically controlled or thermodynamically preferred. The use of a catalyst, such as an acid or an enzyme, changes the state of the reaction without being consumed and lowers the activation energy. Selectivity and stability are key to successful supramolecular systems.

In the case of folding, the structures assume the conformation of greatest thermodynamic stability. It was therefore important to evaluate and distinguish the acylhydrazone pyridine motifs by screening a variety of kinetic and thermodynamic conditions. The key focus was on the intramolecular dynamics of the system and the configurational changes which occurred. Understanding this system will enable progression towards future studies involving induced folding of larger oligomers and complexation mechanisms involving metal ions. In order for a configurational change to occur spontaneously, a negative ΔG must be present and for example the free energy of the cis product must be lower than the free energy of the trans configuration for a given environment. The configurational changes do not occur instantaneously and this rate of change can depend on the variation of conditions such as concentration of the species in the system, temperature, the presence of solvents or the

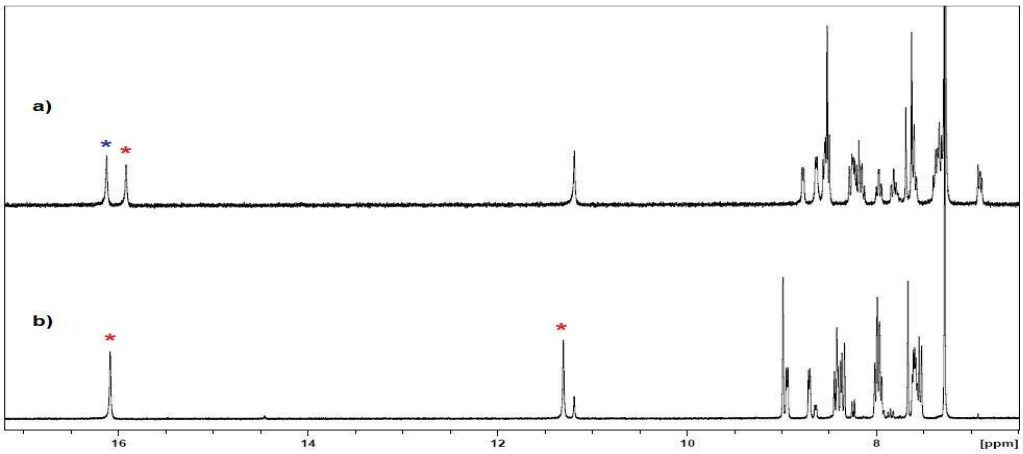
presence of catalysts, all of which influence the thermodynamics and kinetics of the system. This suggests that regardless of which combinations of acylhydrazone pyridine motifs are present, the final configuration is totally dependent upon free energy considerations. The transition between the trans and cis configuration relates to the system being in equilibrium. In the case of the AH dimer, there is one transition from trans to cis, however for the AHA trimer there is an intermediate trans-cis product from trans-trans to cis-cis in a two-step process. In the case of the HAH trimer, the trans-cis configuration is preferred and thus this intermediate is so stable that a deep energy well would be present in this case. The conformation assumed by the acylhydrazone pyridine motif need not be the one of absolute lowest free energy but that of the lowest free energy state kinetically available. Thus with time, and via the trans-cis configuration, the lowest free energy state of cis-cis is expected to be available. The rate of this change can be slow with no assistance and the aid of a catalyst can significantly speed up the process. The overall conclusion is that the acylhydrazone pyridine motif will spontaneously possess a configuration of the lowest available free energy in any given environment and under normal circumstances this will be the native one.

5.2.1 Kinetics

The ability to bring reactants into close proximity results in species that may not ordinarily form. This is the kinetic effect, and relevant studies are based on reaction rates, concentration and temporal coordinates. The fundamental aspects of kinetics in chemistry is the study of reaction rates as the possible association and dissociation of components. Also the intramolecular dynamics which contribute towards conformational changes and/or intra-complex mobility of bound guests is an interesting area with possible use for applications. For kinetic reactions, the rate is proportional to activation energy, with lower activation barriers relating to faster reactions.

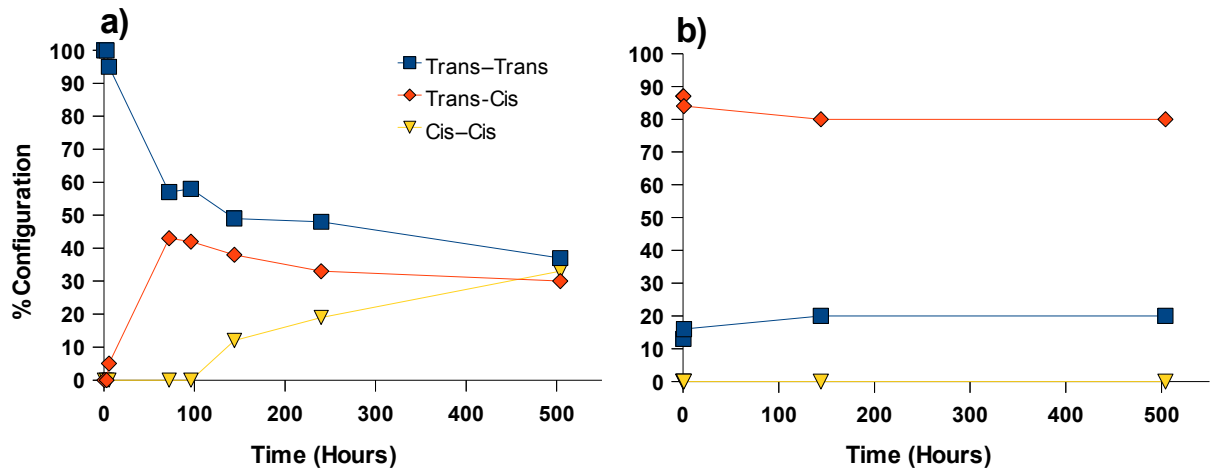
5.2.1.1 Kinetics: Reaction time

In the case of the AHA trimer kinetics, the trans-trans configuration under a given set of conditions can undergo competing reactions to give different proportions of trans-cis or cis-cis products. In terms of the HAH trimer, the trans-cis configuration was formed faster and the system displayed no evidence of being reversible and the system appears to be kinetically controlled. First consideration was with regards to the role of time on the rate of configurational change with fix values of concentration, temperature and solvent. With no additional acid catalyst, each trimer was monitored in $CDCl_3$ over a time period of three weeks. The AHA trimer displayed three species but the HAH trimer had only one isomer available as the major species and a slight amount of the trans-trans product present *(Spectra 8)*. This experiment provided an idea to the proportions of the different species present over time without an additional source of acid, catalyst or increased temperature.



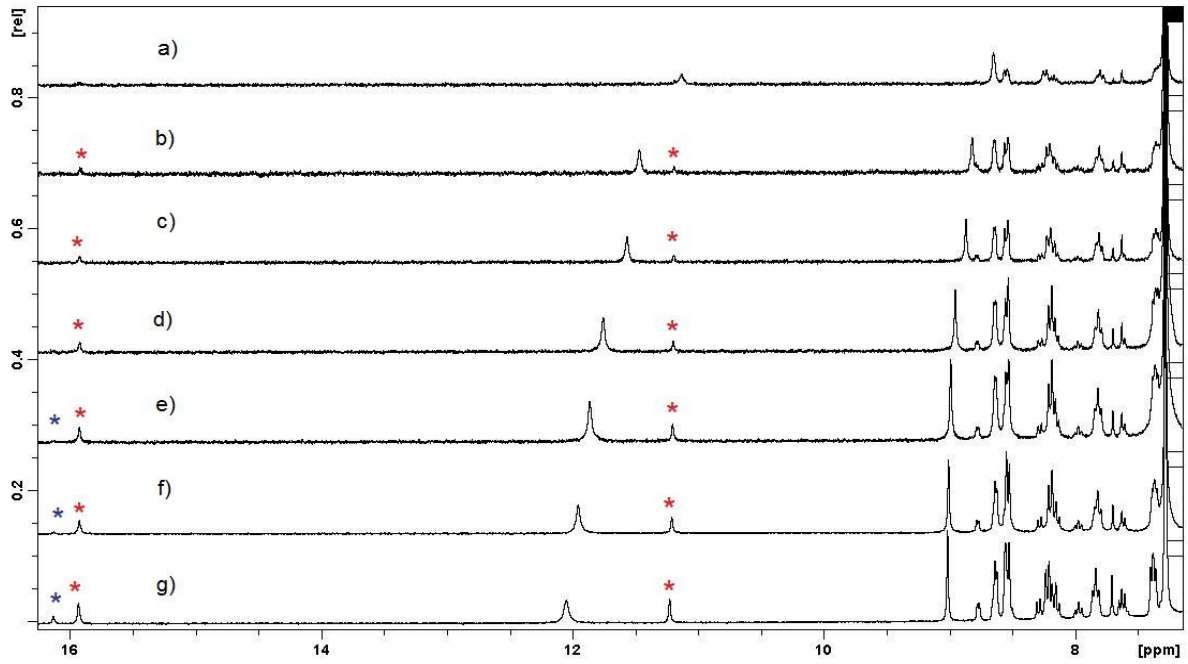
Spectra 8: ^{1}H NMR with 0 equivalents of Acid $CDCl_3$ 3 weeks a) AHA; b) HAH (* blue = cis-cis; * red = trans-cis).

In the case of the AHA trimer in $CDCl_3$, it is three days before a new signal is observed which corresponds to the trans-cis configuration. A total of six days is required to observe the cis-cis configuration, which continues to increase over time. By three weeks no significant change is observed for the AHA trimer. Little change is also observed for the HAH trimer over this time via ^{1}H NMR analysis. For clarity, the proportions of the trans/cis configurations was plotted in a graph for both the AHA and HAH trimer *(Graph 2)*. Note that previous studies show that if left longer, the trimer can revert back to trans-trans configuration due to a possible build up of acid occurring in the $CDCl_3$ solvent.



Graph 2: ^{1}H NMR spectra used to derive% Configuration versus Time (Hours) with no additional acid, $CDCl_3$ at room temperature; a) AHA; b) HAH

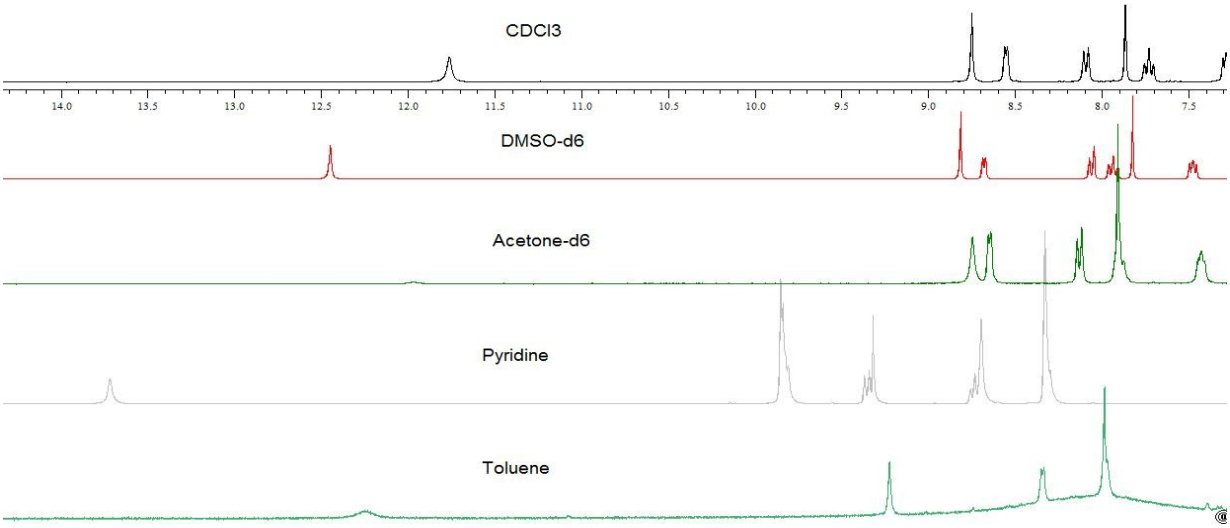
As a side note, to further investigate this phenomenon and the effects of time on the AHA trimer configuration, various equivalents of acetic acid was added to the trimer in $CDCl_3$ and monitored by NMR *(Spectra 9)*. For clarity, the spectra at six hours is shown, but results were obtained for a period of four weeks for each of the samples below. It was observed that the transitions between various configurations proceeded faster with an increase in equivalents of acetic acid.



Spectra 9: ^{1}H NMR of AHA trimer with various equivalents of AcOH in $CDCl_3$ after 6 hours; a) 0 eq; b) 0.1 eq; c) 0.2 eq; d) 0.5 eq; e) 1 eq; f) 2 eq; g) 5 eq. (*) trans-cis; (*) cis-cis configuration.

5.2.1.2 Kinetics: Increasing temperature

The next parameter was to investigate the effect of increasing temperature upon the configuration of the trimers. It was expected that an increase in the reaction rate would result from the increasing temperature. Kinetic studies were carried out by heating the NMR tubes in order to investigate if this process has an effect on configurational change *(Spectra 10)*. Note that the actual NMR experiment was obtained at room temperature, due to the "slow exchange" between different states of configuration it was deemed negligible.

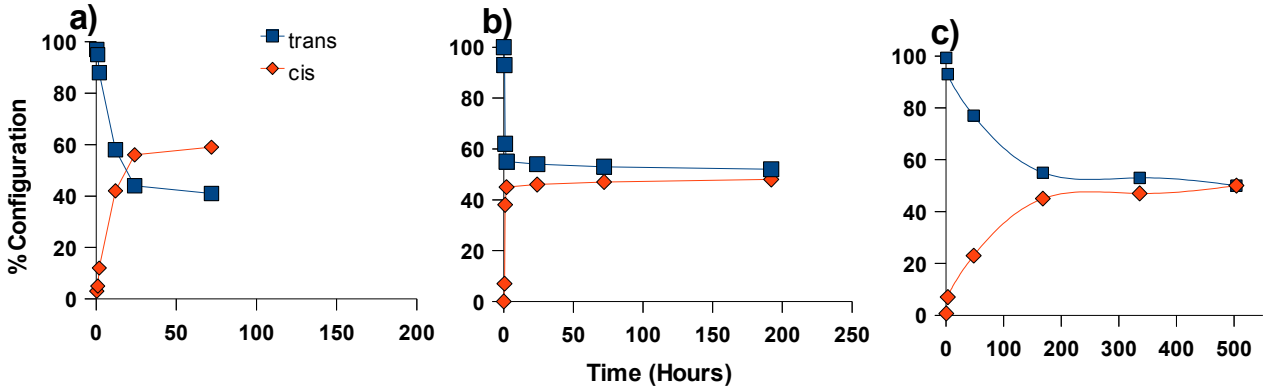


Spectra 10: ^{1}H NMR of AHA trimer with 1.75 equivalents of AcOH in various deuterated solvents at room temperature in which only trans-trans configuration is observed initially.

At room temperature, the AHA trimer sample with 1.75 equivalent of acetic acid in $CDCl_3$ had 100% trans-trans present before heating compared to 91% trans-cis and 9% cis-cis after heating to 50 °C for one hour. Note that due the low boiling point of $CDCl_3$, the kinetic studies were not continued for this sample. It is also important to point out, that upon heating, the equilibrium changed. Therefore when heated at 50 °C or 80 °C, the equilibrium will be different to that at room temperature. Overall, these results show that indeed the reaction rate is increased upon increasing the temperature for the sample in the solvent $CDCl_3$, with three species present for the AHA trimer. In pyridine-d_5 and toluene-d_8 aromatic aprotic solvents heated to 80 °C, it was also possible to observe an increased reaction rate *(Spectra 7)*. Three species were observed for the AHA trimer, with a preference for trans-cis configuration. This differs to $CDCl_3$ solvent, which has a preference for AHA cis-cis configuration upon heating.

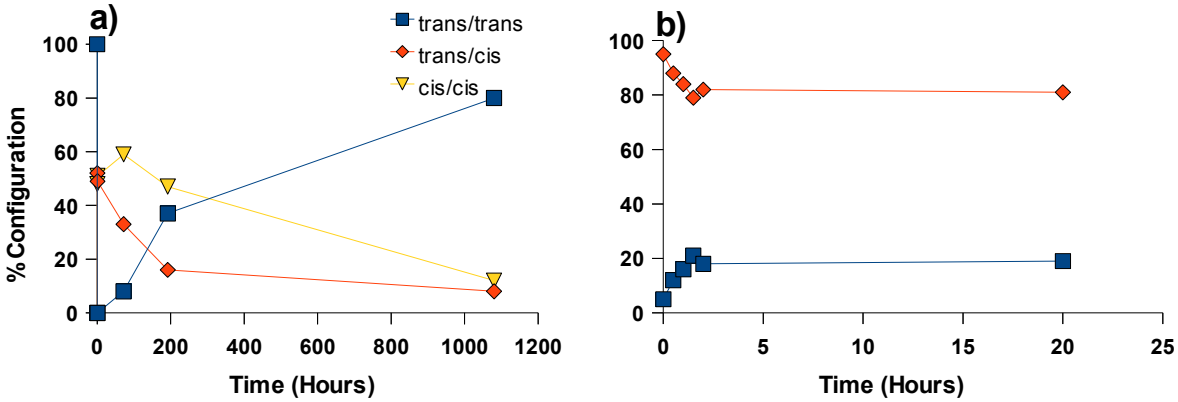
5.2.1.3 Kinetics: Addition of catalytic acid

The acylhydrazone pyridine motifs exhibit cis configuration spontaneously over time without the addition of a catalyst. However, the addition of catalytic acid to speed up this change in configuration process was also investigated and the additional catalytic acid was expected to bring about rate acceleration. It was expected to reduce entropic barriers and orientate the molecule in such a geometry to favour the cis configuration. For this experiment, 0.01 equivalents of trifluoroacetic acid (TFA) was used. This catalytic amount of acid allowed for the system to be monitored effectively by protonation in quantitative terms *(Graph 3)*. Upon addition of 0.01 equivalents of acid it was expected that the level of protonation will be 1% of the acylhydrazone pyridine molecules and this will shift the equilibria by 1%. The percent configuration of the cis AH dimer with 0.01 equivalent of TFA acid in $CDCl_3$ solvent reached equilibrium in twenty four hours *(Graph 3,a, Spectra 11,a),* whereas the cis configuration was 45% for the AH dimer after two hours with 0.05 equivalents of TFA acid in $CDCl_3$ solvent *(Graph 3,b)*. This faster rate to reach equilibrium is in contrast to the AH dimer with 1.75 equivalents of AcOH, which reached equilibrium after one week *(Graph 3,c)*.

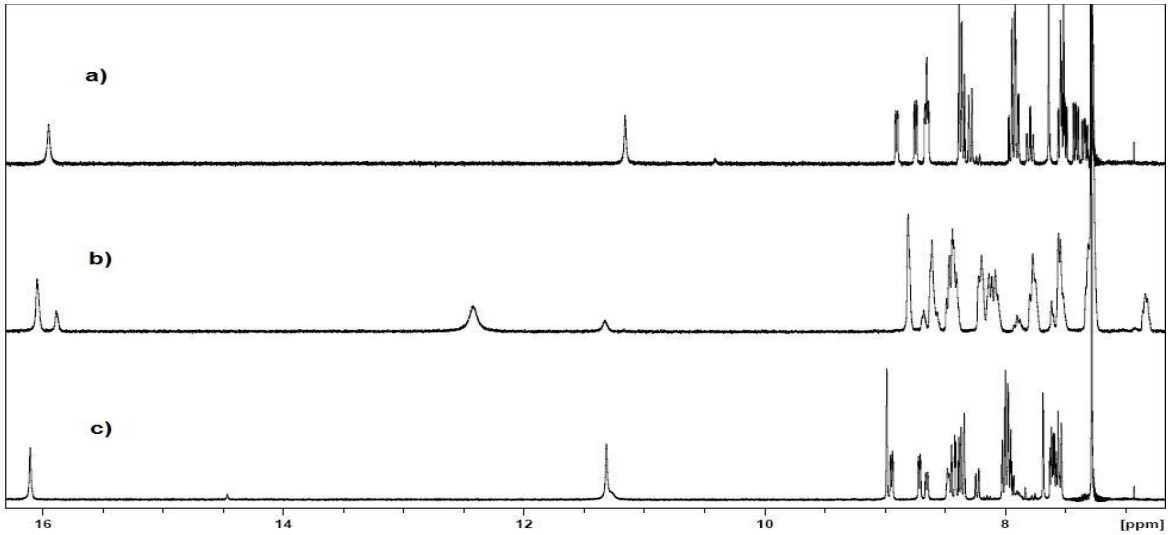


Graph 3: ^{1}H NMR of AH in $CDCl_3$ solvent, room temperature;% configuration versus Time (hours); a) 0.01 equivalents of TFA; b) 0.05 equivalents of TFA; c) 1.75 equivalents of AcOH.

Similar conditions were applied to the AHA trimer with addition of 0.01 equivalents of TFA acid in $CDCl_3$ solvent and monitored by ^1H NMR at room temperature *(Graph 4)*. After thirty minutes, the trans-trans configuration was present and little change was observed between thirty minutes and twelve hours *(Graph 4,a)*. After one day the cis-cis configuration is prominent and the spectra remains the same one week later *(Spectra 11,b)*. This is consistent with the AcOH experiments, but equilibrium is reached faster due to the TFA acid catalyst. Note that by week seven, the trans-trans AHA isomer is dominant, which could be attributed to the increase of acid within $CDCl_3$ over time. This phenomenon was further confirmed by monitoring the NMR spectra of the AHA trimer samples and will be discussed shortly. The same conditions were applied to the HAH trimer. In this case, the trans-cis configuration remained, 95% initially with a levelling off to 81% after 24 hours *(Graph 4,b, Spectra 11,c)*. The trans-trans configuration appeared to slightly increase over this time from from 5% to 19%. However its chemical shift of 11.16 ppm is very close to the trans-cis chemical shift at 11.28 ppm. The calculations based on peak integration were shown to be a challenge.

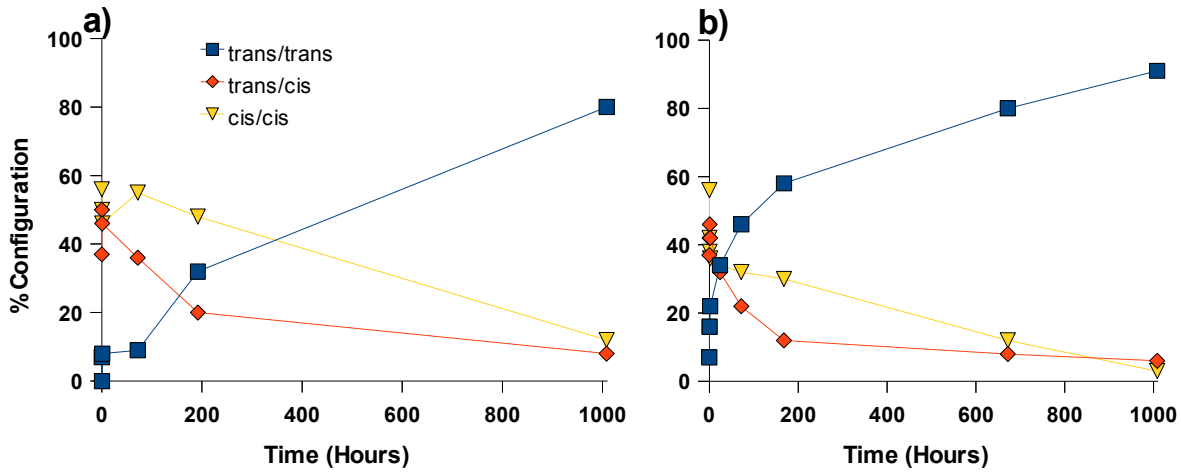


Graph 4:% configuration versus Time (Hours) at RT, $CDCl_3$ with 0.01 equivalents of TFA a) AHA; b) HAH.

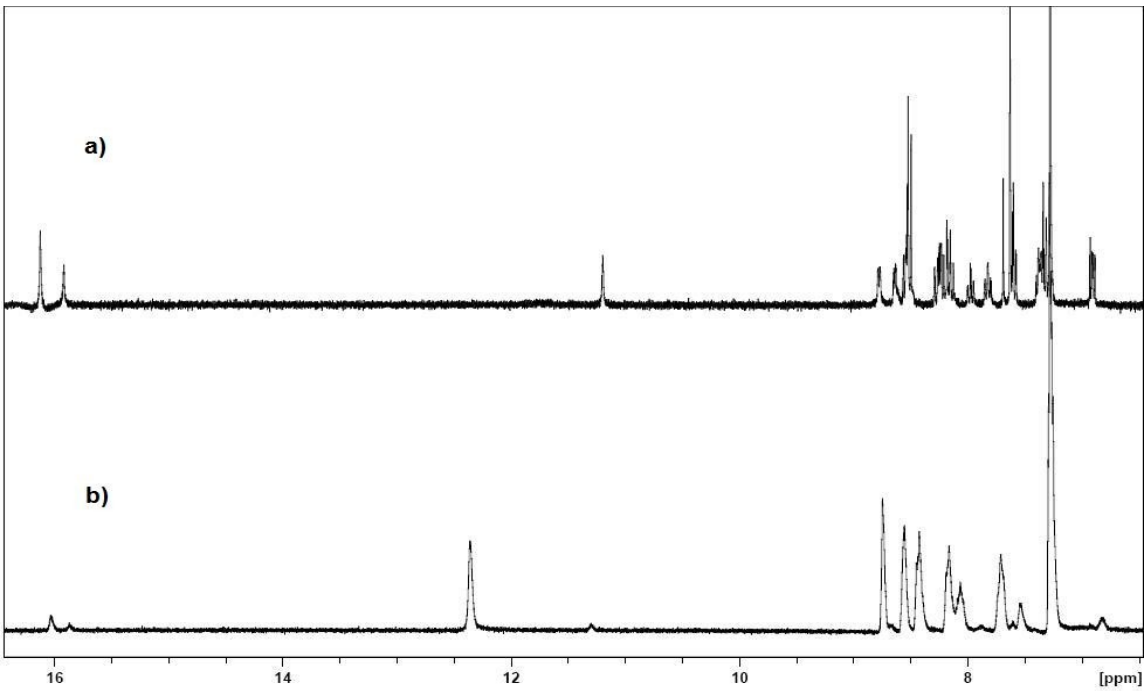


Spectra 11: All ^1H NMR spectra 0.01 eq TFA in $CDCl_3$ after 1 day all show cis isomers present; a) AH; b) AHA; c) HAH.

As mentioned previously, analysis was carried out on an AHA trimer sample which showed that too much acid build-up within the chloroform solvent can result in reversing the equilibrium. The AHA sample had previously reached equilibrium in conditions with five equivalents of AcOH in $CDCl_3$. The sample was dried and redissolved in chloroform and 0.01 equivalents of TFA acid added (Graph *5,a)*. The ^1H NMR spectra the changed from a predominantly cis-cis AHA isomer to that of trans-trans configuration over a period of six weeks *(Spectra 12)*. Ultimately, these results indicate the danger of acidic HCl build up within $CDCl_3$ solvent over time. As a result it is important to be aware of this issue which could effect results.



Graph 5: % configuration versus Time (Hours) at room temperature,Reverse Experiment, $CDCl_3$ with TFA a) 0.01 equivalents of TFA and the AHA trimer; b) 0.05 equivalents of TFA and the AHA trimer



Spectra 12: ^1H NMR Reverse Experiment, AHA with 0.01 eq TFA in $CDCl_3$ a) Initial spectra with cis-cis AHA isomer as dominant b) 6 weeks later, trans-trans configuration AHA trimer is dominant.

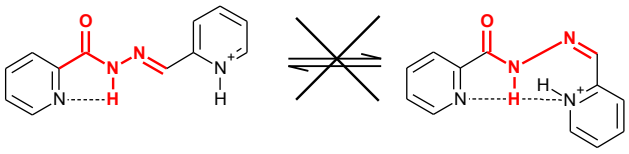
5.2.2 Thermodynamics

A characteristic of many metabolic processes is a dynamic state between forces, known as dynamic equilibrium, which can be influenced by changing the environment, temperature, solvent, pH and salt concentration. This reversible process will ultimately drive the process in favour of the thermodynamic product. As mentioned previously, when ΔG is negative the process proceeds spontaneously. The creation of a well-defined structure results in loss of entropy, which must be compensated by either an increase in favourable enthalpic interactions or a gain in entropy from the environment. The activation energy ($\Delta G^{\ddagger}$) required to overcome the entropic penalty of the initial step is overcome by the use of a catalyst.

It Is expected that after the initial transformation to the cis configuration in a long acylhydrazone pyridine oligomer, the induced effect will result in a helical structure. As more units are added, it is expected, based on biological processes, a momentum until the oligomer is fully assembled in a acylhydrazone pyridine helical foldamer. In such cases, the initial threshold is overcome and the thermodynamic conditions which existed at the beginning of the process have changed in favour of increased assembly. Thus the initial entropic loss encountered by folding will be compensated by favourable hydrogen bonding, electrostatic, hydrophobic and steric interactions in the secondary structure. One can obtain well-defined helices under thermodynamic control when the free energy of the most stable conformation is sufficiently lower than that of competing conformations (high ΔG_{AB}).

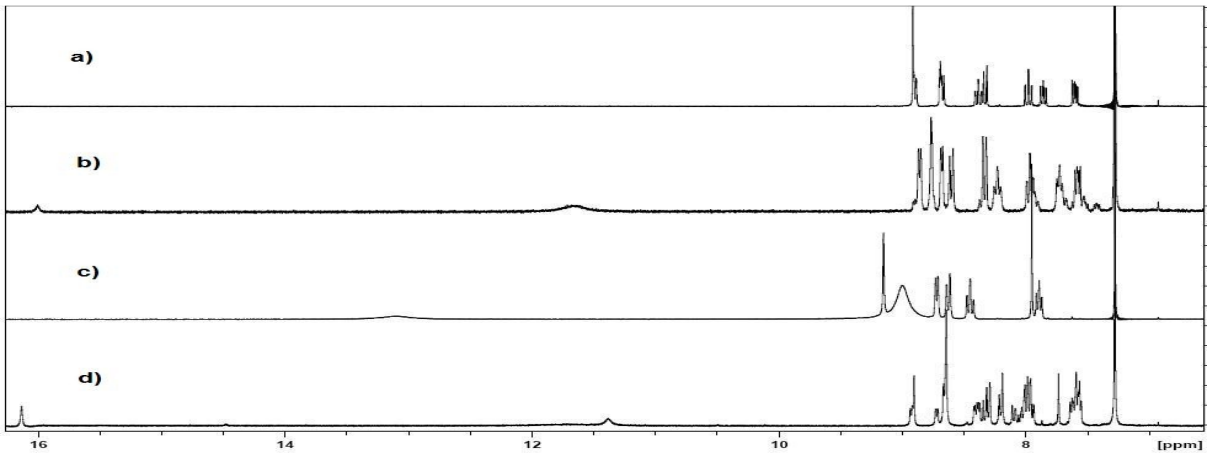
5.2.2.1 Thermodynamics: Addition of stoichiometric acid

In contrast, to the kinetic studies, a stoichiometric amount of acid was added to the acylhydrazone pyridine molecules. This ranged from 0.1, 0.5, 1 to 1.75 equiv. of TFA acid in $CDCl_3$. Notably with TFA, it was possible to fully protonate *(Fig. 30)* and thus restrict the configuration to 100% trans for the AH and AHA molecules *(Spectra 13,a,b,c).* With regards to the HAH trimer, it was difficult to determine an exact amount of trans-trans configuration from NMR integration *(Spectra 13,d)*. Upon complete protonation, the broad signal at 11.7 ppm corresponding to the trans-trans configuration increases the challenge to determine the integration accurately. Taking into account the integration of the trans-cis signal at 16.11 ppm and the combined trans-trans, trans-cis configuration at 11.7 ppm and 11.3 ppm respectively, the percentage isomers was determined and a graph plotted *(Graph 6,b).* This showed a preference for the trans-trans HAH configuration with 58% for the first twenty four hours. This preference was not observed before for HAH, although it shifts towards the trans-cis configuration. With this technique, it's possible to shift cis to trans configuration by protonation. Thus changing the dynamics from a kinetically driven equilibrium towards a thermodynamic one by fully protonating the system. This is similar to a concept previously reported within the Huc Group.[110] With AcOH in $CDCl_3$, one or five equiv. was not enough to fully protonate and thus cis configuration was observed (*Graph 6,a Spectra 14)*.

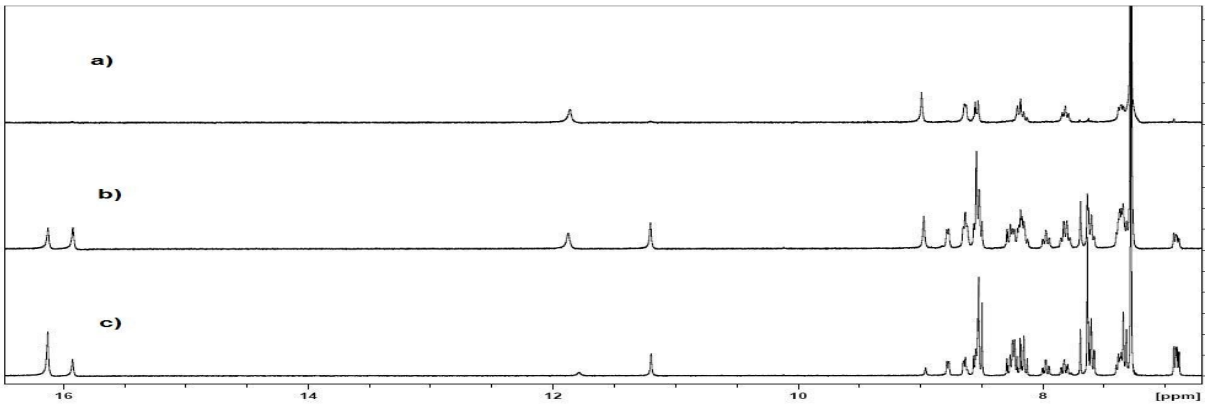


AH trans AH cis

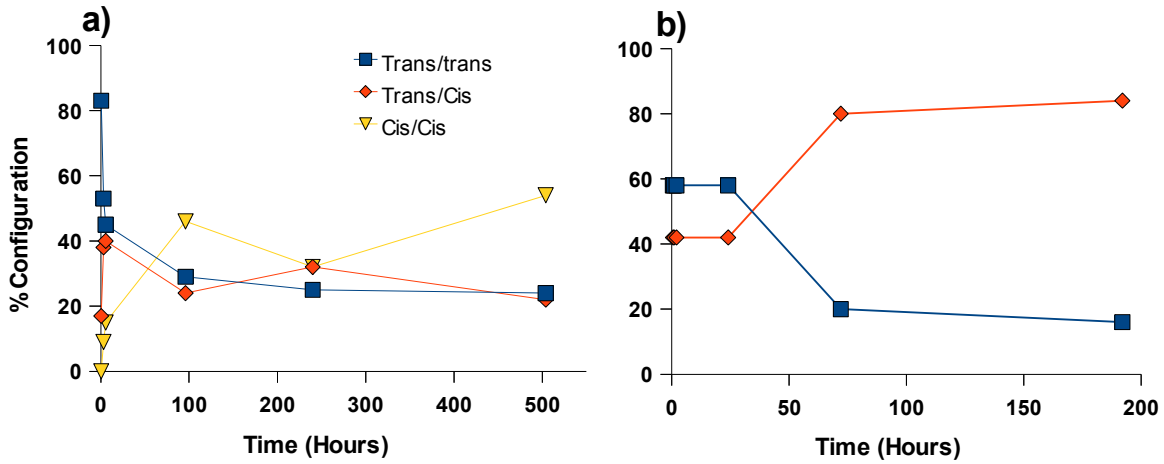
Figure 30: The AH dimer fully protonated and restricted to the trans configuration



Spectra 13: ^{1}H 1 equivalent of TFA in $CDCl_3$, spectra a) AH 1 hour b) AH 4 wks; c) AHA 1 hour; d) HAH 1 hour



Spectra 14: AHA with 100% AcOH in $CDCl_3$ a) 30 mins; b) 3 days; c) 4 weeks



Graph 6: % configuration versus Time (Hours) at room temperature, $CDCl_3$ a) The AHA trimer with 5 equivalents of AcOH; b) HAH trimer with 1 equivalent of TFA

6 Conclusions

This research showed that equilibrium is biased for acylhydrazone pyridine motifs due to available intramolecular hydrogen bonding and that the cis configuration is observed by ^{1}H NMR and X-ray crystallography analysis. At the time of when these studies were conducted, this phenomenon had not been previously reported for other acylhydrazone motifs. Since then, the Lehn [121] and Barboiu [122] groups have reported this concept with the aid of photochemical stimuli or metal ions. The results reported here, are still interesting as it shows the equilibrium is biased without the need for external stimuli.

The main focus of this study was on the dynamic chemistry of AH dimer, AHA trimer and HAH trimer systems. Synthesis using symmetrical units H (hydrazide) and A (aldehyde) monomers was suitable to obtain the AH dimer and AHA, HAH trimers. Upon attempted elongation to obtain HAHAH **86**, AHAHA **89**, **90**, and **91**, **92**, the use of symmetrical monomers proved insufficient. Desymmetrisation of the monomer units with masked functional groups provided an alternative synthetic route to ease oligomer elongation. The dis-symmetrical trimer **103** was obtained by reacting **71** with **100** in DCM. However despite four attempts at purification via column chromatography, the trimer **103** appeared to not be isolated in a pure form. Purification methods of the trimer **103**, perhaps by recrystallisation need to be re-investigated. The dynamic configurational changes of this dis-symmetrical compound would of interest to study in future.

The synthesis of longer oligomers containing multiple hydrazones was found to be unexpectedly difficult, as elongation to pentamer motifs exhibited increased complexity. For pentamer **106** a convergent synthetic route proved viable, whereas for pentamer **109**, a divergent synthetic route was preferred. The purification of pentamer **106** needs to be re-visited. Whereas the synthesis for **109** needs to be continued by reacting **108** with **71**, as previously there was not enough material to complete the final step. The main challenge in the pentamer synthesis was the purification of the final product to obtain a single pure species. Thus until now, it was not possible to obtain pentamers in a well defined form. In terms of overall synthetic challenges, oxidation was carried out using activated manganese (IV) oxide under reflux in 1,4-dioxane resulting in poor yields. For future experiments, conditions outlined by the Lüning group using selenium dioxide should be investigated.[118] Use of hydrazine monohydrate to form hydrazides, should be restricted to monomer units as hydrazone cleavage occurs for longer oligomers. In future, the constitutional interconversion method implemented by the Lehn group,[121] utilising reversible chemical reactions such as hydrazone exchange could be investigated to access longer oligomers.

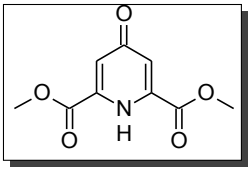
The overall conclusions obtained from the initial ^{1}H NMR studies suggest that cis configuration exists for the AH dimer and the trimers AHA and HAH. The emergence of a new proton signal at 16 ppm is a feature of cis configuration. The NMR data showed an equilibrium for AH and AHA in a given solvent, $CDCl_3$. For trimer AHA, there are three different species that can exist. The trimer HAH, however, has an overall preference for one species, the trans-cis configuration, with trans-trans only as a minority. One isomer, the cis-cis configuration, was completely absent from the trimer HAH samples and appears to not be feasible. X-ray crystallography shows a trans-cis configuration for the HAH trimer, however the trans configuration was observed for AH and AHA. In hindsight, more tests could have been carried out to further support our interpretation of the NMR results. The crystallography data for AH, AHA and HAH showed that the solid state preference for configuration is coupled to and coincides with the initial state observed in the NMR analysis. Perhaps with different purification conditions, different isomers crystals might have been isolated. For example, it would be desirable to grow AHA crystals in toluene solvent in an effort to obtain a crystal structure of trans-cis AHA configuration. NMR experiments such as NOESY/ROESY NMR or deuterium exchange could provide further substantial support towards the existence of cis configuration. Similar investigations need to be carried out for the dis-symmetrical trimer.

Various environmental conditions were tested to explore the equilibrium of each system. Due to the denaturing effect and competition for hydrogen bonding in DMSO-d_6 no cis configuration was observed for the AH dimer and AHA trimer. In contrast, the HAH trimer still exhibits trans-cis configuration even in DMSO-d_6, possibly confirming the strong stability and closeness to equilibrium compared to that of the AH or AHA. Thus to observe cis configuration, choice of solvent is important in order to enable the available intramolecular hydrogen bonding to form a cis biased configuration in equilibria. $CDCl_3$ was removed from the previous AH dimer and AHA, HAH trimer samples and replaced with DMSO-d_6 to the isomer-mixtures. The mixture of trans-cis still exists for the AH dimer suggesting equilibrium was not reached and must lie between 100% and 60% trans. Unlike the AH dimer, the isomer mixtures were destroyed for the AHA trimer and only trans-trans configuration remains, speculating that the equilibrium is closer to 100% trans in this case. After 3 weeks, 94% trans-trans AHA trimer was present with 6% cis-trans AHA trimer. Since there was no change with the HAH trimer in either the $CDCl_3$ or DMSO-d_6 solvent, this experiment was not carried out, as it's expected that the HAH trimer is close to thermodynamic equilibrium. The solvents pyridine-d_5 or toluene-d_8 were not completely inert to the equilibrium process. Toluene could be a good solvent to obtain cis-cis isomers of the AHA trimer.

In terms of kinetics, addition of catalytic acid and increasing temperatures was found to increase the rate of equilibrium and thus the configurational change from the trans-trans species to that of the cis-cis AHA trimer. In contrast, to the kinetic studies, a stoichiometric amount of acid was added to the AH, AHA and HAH acylhydrazone pyridine molecules. Notably with TFA, it was possible to fully protonate the molecule and thus restrict the configuration to 100% trans for the AH, and AHA molecules. HAH also exhibited trans-trans isomer in these conditions. Overall these studies show that despite the fact that these motifs are made up of similar building blocks, the way in which they are connected contributes to a difference in overall behaviour and properties. This fundamental characteristic gives rise to isomers which can have extraordinary complexity coupled with the potential ability to master or manipulate these combinations in a controlled environment. The results obtained from this thesis lay the foundations for further investigations including incorporating acylhydrazone pyridine motifs into foldamer backbones to expand a new generation of foldamers. The ability for self-folding and the possibility of information storage of these motifs contributes towards an interest in future development. Further emphasis on incorporating dynamic combinatorial chemistry via the use of templates and hydrazone exchange as discussed in the introduction and demonstrated by the Lehn [121] and Moore [106] groups could provide a pathway to access longer oligomers with interesting dynamics and features.

7 Experimental

Chelidamic acid dimethyl ester hydrochloride 8



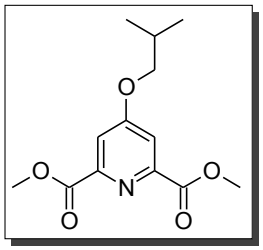
MF: $C_9H_9NO_5$ (HCl Salt)

MW: 211.17 g.mol^{-1} (247.63 salt)

Ref: [77] tcm084, tcm2-061, tcm3-038, tcm3-073

To a suspension of chelidamic acid **7** (40 g, 199 mmol) in absolute MeOH (600 mL) was added 2,2-dimethoxypropane (280 mL) and concentrated HCl (30 mL). The mixture was heated to reflux for 8 hours and then at RT for 3 nights under a $CaCl_2$ drying tube. The volatiles were removed *in vacuo* and the residue was triturated with diethyl ether (500 mL). The white solid was filtered and rinsed further with diethyl ether. No further purification was carried out and the compound **8** was used in the following step as is. Yield is (44 g, 95%) of a white solid. **^{1}H NMR** (DMSO-d_6, 300 MHz): *δ* 7.61 (2H, s); 3.87 (6H, s).**^{13}C NMR** ($CDCl_3$, 75 MHz): *δ* 174.5, 163.3, 142.8, 118.5, 53.6

Dimethyl 4-isobutoxy-2,6-pyridine dicarboxylate 79



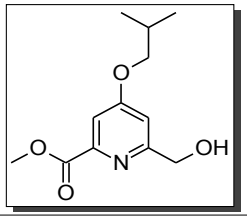
MF: $C_{13}H_{17}NO_5$

MW: 267.28 g.mol^{-1}

Ref: [119,120] tcm3-039, tcm3-060, tcm3-133

Dimethyl 4-hydroxy-2,6-pyridinedicarboxylate and K_2CO_3 was co-evaporated twice with anhydrous toluene before use. To a solution of dimethyl 4-hydroxy-2,6-pyridinedicarboxylate (19.6 g, 79.5 mmol, 1.0 equiv) in anhydrous DMF (300 mL) was added K_2CO_3 (33 g, 238 mmol, 3 equiv) under nitrogen. The reaction mixture was stirred at 120 °C for 2 h, then cooled to 70 °C and *i*BuI (27 mL, 238 mmol, 3 equiv) was added and the reaction mixture was stirred at 70 °C for 3 hours. The reaction mixture was cooled to room temperature, filtered, and rinsed with toluene (x 3). The filtrate was washed with H_2O (2 x 300 mL). The combined organic phases were washed with brine (300 mL), dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (EtOAc/Pet. Ether, 8:2) to afford a white solid (17.7 g, 84% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ 7.72 (2H, s), 4.01 (2H, d, *J* = 6.4 Hz), 3.90 (6H, s), 2.06 (1H, m), 1.00, 0.98 (6H, d, *J*= 6.1 Hz)

Methyl 6-(hydroxymethyl)-4-isobutoxy-2- pyridine carboxylate 94



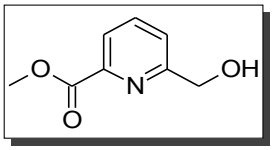
MF: $C_{12}H_{17}NO_4$

MW: 239.27 g.mol^{-1}

Ref: [119,120] tcm3-053, tcm3-069, tcm3-134

To the diester (7.2 g, 26.9 mmol, 1.0 equiv) was added $NaBH_4$ (1.53 g, 40.4 mmol, 1.5 equiv) in anhydrous DCM (70 mL) under nitrogen and cooled to 0 °C. MeOH (200 mL) was added via a syringe and allowed to reach ambient temperature. After 3 h, TLC showed starting material was still present and the reaction was stirred at RT overnight. The reaction mixture was neutralized with 1 M HCl solution (35 mL), and the volatiles removed *in vacuo*, water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (EtOAc/Pet. Ether, 1:1) to afford a white solid (5.02 g, 78% yield). However, 1.5 g of starting material was recovered, thus 98% recovered yield was obtained. **^{1}H NMR** ($CDCl_3$, 300 MHz): δ 7.55 (1H, d, *J* = 2.0 Hz), 7.02 (1H,d, *J* = 2.0 Hz), 4.79 (2H, s), 3.98 (3H, s), 3.83 (2H, d, *J* = 6.8 Hz), 2.16 (1H, m), 1.04 (6H, d, *J* = 6.4 Hz)

Methyl 6-(hydroxymethyl)pyridine-2-carboxylate 93



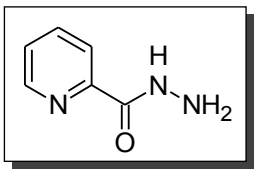
MF: $C_8H_9NO_3$

MW: 167.16 g.mol^{-1}

Ref: tcm3-125, SC044

To the diester (350 mg, 1.79 mmol, 1.0 equiv) was added $NaBH_4$ (102 mg, 2.7 mmol, 1.5 equiv) in anhydrous DCM (10 mL) under nitrogen and cooled to 0 °C. MeOH (10 mL) added via a syringe and allowed to reach ambient temperature. After 3 h, TLC showed starting material was still present and the reaction was stirred at RT overnight. The reaction mixture was neutralized with 1 M HCl solution (5 mL) and the volatiles removed *in vacuo*, water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The crude product was purified by column chromatography (EtOAc/Pet. Ether, 1:1) to afford a white solid (227 mg, 76% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 8.03-8.01 (1H, d, *J* = 7.8 Hz), 7.84 (1H, t), 7.54-7.51 (1H, d, *J* = 7.6 Hz), 4.85 (2H, s), 3.98 (3H, s)

Pyridine-2-carbohydrazide 71



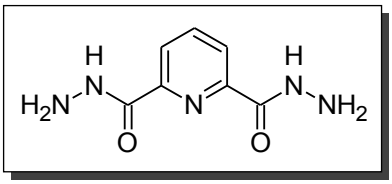
MF: $C_6H_7N_3O$

MW: 137.14 g.mol^{-1}

Ref: SC050a

To methyl picolinate (2.23 g, 16.26 mmol, 1 equiv) in MeOH (10 mL) was added hydrazine monohydrate (0.8 mL, 16.5 mmol, 1.01 equiv) at 0 °C under nitrogen. The reaction mixture was allowed to reach ambient temperate and was stirred overnight. The volatiles were removed *in vacuo*, water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a yellow solid (2.23 g, 99% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 9.89 (1H, bs), 8.6-8.58 (1H, dt, *J* = 2.0 Hz), 7.98 (2H, m), 7.57-7.52 (1H, m), 4.68 (2H, bs) **^{13}C NMR** (DMSO-d_6, 75 MHz) δ 162.5, 149.7, 148.4, 137.5, 126.1, 121.6. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_6H_8N_3O$) 138.0667; found 138.0664.

Pyridine-2,6-dicarbohydrazide 72



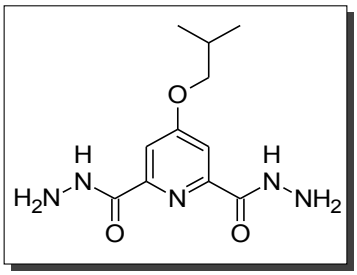
MF: $C_7H_9N_5O_2$

MW: 195.17 g.mol^{-1}

Ref: SC045, tcm3-072

To the diester (100 mg, 0.51 mmol, 1 equiv) in MeOH (5 mL) was added (0.25 mL, 5.1 mmol, 10 equiv) of hydrazine monohydrate at 0 °C under nitrogen. The reaction mixture was allowed to reach ambient temperate and was stirred overnight. The volatiles were removed *in vacuo*, water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a white solid (100 mg, 99% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 10.62 (2H, s), 8.12 (3H, s), 4.62 (4H, s) **^{13}C NMR** (DMSO-d_6, 75 MHz) δ 161.8, 148.3, 139.2, 123.5. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_7H_{10}N_5O_2$) 196.0834; found 196.0834.

4-(isobutoxy)pyridine-2,6-dicarbohydrazide 73



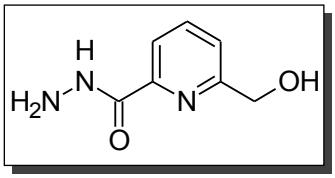
MF: $C_{11}H_{17}N_5O_3$

MW: 267.28 g.mol^{-1}

Ref: tcm3-052

To the diester (500 mg, 1.87 mmol, 1 equiv) in MeOH (10 mL) was added (1.0 mL, 18.7 mmol, 10 equiv) of hydrazine monohydrate at 0 °C under nitrogen. The reaction mixture was allowed to reach ambient temperature and was stirred overnight. The volatiles were removed *in vacuo,* water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a white solid (500 mg, 99% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 10.57 (2H, bs); 7.56 (2H, s); 4.64 (4H, bs); 3.96-3.93 (2H, d *J* = 6.4 Hz); 2.03 (1H, m); 0.98-0.96 (6H, d, *J* = 6.6 Hz) **^{13}C NMR** (DMSO-d_6, 75 MHz) δ 167.0, 161.6, 150.4, 109.4, 74.2, 27.4, 18.7

6-(hydroxymethyl)pyridine-2-carbohydrazide 95



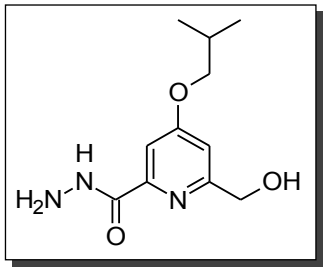
MF: $C_7H_9N_3O_2$

MW: 167 g.mol^{-1}

Ref: tcm3-130

To a solution of 6-hydroxymethyl pyridine-2-carboxylate (169 mg, 0.70 mmol, 1.0 equiv) in MeOH (5 mL) was added hydrazine monohydrate (0.2 mL, 4.2 mmol, 6 equiv) under nitrogen at 0 °C for 30 minutes and then at room temperature for 4 hours. The volatiles were removed *in vacuo,* water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a white solid (169 mg, 99% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 9.92 (1H, bs); 7.95 (1H, t); 7.84 (1H, d, *J* = 7.33 Hz); 7.56 (1H, d, *J* = 7.58 Hz); 4.62 (2H, d, *J* = 5.13 Hz); 4.57 (2H, bs)

6-(hydroxymethyl)-4-(isobutoxy)pyridine-2-carbohydrazide 96



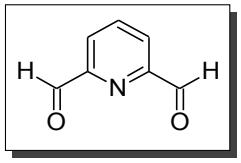
MF: $C_{11}H_{17}N_3O_3$

MW: 239.27 g.mol^{-1}

Ref: tcm3-081, tcm3-100, tcm3-135

To a solution of 6-hydroxymethyl 4-isobutoxypyridine-2-carboxylate (5.0 g, 20.89 mmol, 1.0 equiv) in MeOH (100 mL) was added hydrazine monohydrate (3 mL, 62.7 mmol, 3 equiv) under nitrogen at 0 °C for 30 minutes and then stirred at room temperature overnight. The volatiles were removed *in vacuo,* water was added to the residue and extracted with DCM (x 3). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a white solid (4.9 g, 99% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 9.26 (1H, bs); 7.59-7.58 (1H, d, *J* = 2.2 Hz); 6.90-6.89 (1H, d, *J* = 2.0 Hz); 4.69 (2H, s); 3.95 (2H, bs); 3.83-3.80 (2H, d, *J* = 6.6 Hz); 2.09 (1H, m); 1.02-0.99 (6H, d, *J* = 6.6 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 167.1, 164.6, 160.3, 149.8, 109.4, 107.6, 74.7, 64.3, 27.9, 19.0. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{11}H_{18}N_3O_3$) 240.1348; found 240.1339.

Pyridine-2,6-dicarbaldehyde 75



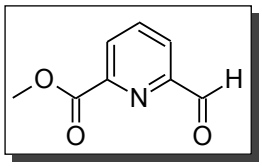
MF: $C_7H_5NO_2$

MW: 135.12 g.mol^{-1}

Ref: tcm3-047, tcm3-061

To a solution of 2,6-pyridinedimethanol (500 mg, 3.6 mmol, 1 equiv) was added activated MnO_2 (1.3 g, 15.1 mmol, 4.2 equiv) in 1,4-dioxane (4.5 mL) and heated to 100 °C for three nights. The reaction mixture was cooled down and filtered on a bed of celite, washed with chloroform and the filtrate evaporated to dryness *in vacuo*. The crude compound was purified by column chromatography (DCM/MeOH, 99:1 → 95:5) to afford a white solid (194 mg, 40% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 10.17 (2H, s); 8.18 (2H, d, *J* = 6.98 Hz); 8.08 (1H, t)

Methyl 6-formylpyridine-2-carboxylate 97



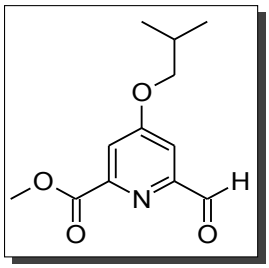
MF: $C_8H_7NO_3$

MW: 165.14 g.mol^{-1}

Ref: tcm3-048

To a solution of 6-hydroxymethyl pyridine-2-carboxylate (180 mg, 0.107 mmol, 1 equiv) was added activated MnO_2 (393 mg, 4.5 mmol, 4.2 equiv) in 1,4-dioxane (4 mL) and heated to 100 °C overnight. The reaction mixture was cooled down and filtered on a bed of celite, washed with DCM and the filtrate evaporated to dryness *in vacuo* to afford a white solid (165 mg, 93% yield). No further purification was carried out and the product was used in the following step as is. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 10.19 (1H, s); 8.35 (1H, d, *J* = 8.68 Hz); 8.16 (1H, d, *J* = 8.12 Hz); 8.05 (1H, t); 4.06 (3H, s)

Methyl 6-formyl-4-(isobutoxy)pyridine-2-carboxylate 98



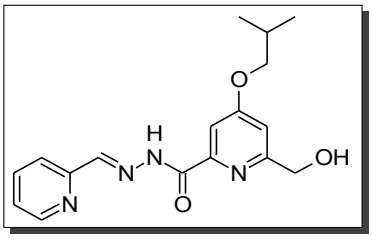
MF: $C_{12}H_{15}NO_4$

MW: 237.25 g.mol^{-1}

Ref: tcm3-082

To a solution of 6-hydroxymethyl 4-isobutoxypyridine-2-carboxylate (350 mg, 1.46 mmol, 1 equiv) was added activated MnO_2 (533 mg, 6.14 mmol, 4.2 equiv) in 1,4-dioxane (8 mL) and heated to 100 ºC overnight. The reaction mixture was cooled down and filtered on a bed of celite, washed with DCM and the filtrate evaporated to dryness *in vacuo*. The crude compound was purified by column chromatography (DCM, 100%) to afford a white solid (200 mg, 58% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 10.09 (1H, s); 7.82-7.81 (1H, d, *J* = 2.6 Hz); 7.57-7.56 (1H, d, J = 2.6 Hz); 4.02 (3H, s); 3.88-3.86 (2H, d, *J* = 6.4 Hz); 2.19-2.06 (1H, m); 1.03-1.01 (6H, d, *J* = 6.8 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 192.6, 167.1, 164.9, 154.5, 149.9, 116.2, 109.7, 75.3, 53.2, 27.9, 18.9.

6-(hydroxymethyl)-4-isobutoxy-*N'*-[(*E*)-pyridin-2-ylmethylidiene]pyridine-2-carbohydrazide **99**



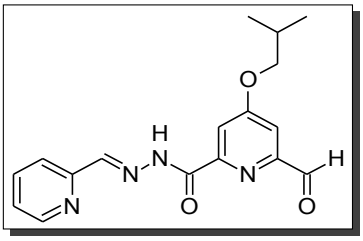
MF: $C_{17}H_{20}N_4O_3$

MW: 328 g.mol^{-1}

Ref: tcm3-136

To 6-(hydroxymethyl)-4-(2-methylpropoxy)pyridine-2-carbohydrazide (5 g, 20.9 mmol, 1 equiv) in MeOH (100 mL) was added 2-pyridine carboxyaldehyde (4 mL, 42.1 mmol, 2 equiv) and the reaction mixture was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (EtOAc/Pet. Ether 1:99 → 95:5) to afford a white solid (6.04 g, 88% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 11.53 (1H, s); 8.48 (1H, d, *J* = 4.91 Hz); 8.29 (1H, s); 8.11 (1H, d, *J* = 7.93 Hz); 7.64 (1H, t); 7.60 (1H, d, *J* = 2.45 Hz); 7.20 (1H, t); 6.96 (1H, d, *J* = 2.26 Hz); 4.89 (1H, t); 4.73 (2H, d, *J* = 4.72 Hz); 3.72 (2H, d, *J* =6.42 Hz); 2.01 (1H, m); 0.95 (6H, d, *J* = 6.79 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 167.1, 161.1, 160.8, 152.8, 149.7, 149.0, 148.9, 136.4, 124.2, 121.1, 110.0, 108.0, 74.6, 64.5, 27.8, 18.9. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{17}H_{21}N_4O_3$) 329.1614; found 329.1620.

6-formyl-4-isobutoxy-*N'*-[(*E*)-pyridin-2-ylmethylidiene]pyridine-2-carbohydrazide **100**



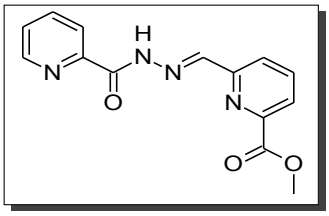
MF: $C_{17}H_{18}N_4O_3$

MW: 326 g.mol^{-1}

Ref: tcm3-137

To 6-(hydroxymethyl)-4-isobutoxy-*N'*-[(*E*)-pyridin-2-ylmethylidiene]pyridine-2-carbohydrazide (1g, 3.04 mmol, 1 equiv) was added activated MnO_2 (2.65 g, 30.5 mmol, 10 equiv) in 1,4-dioxane (25 mL) and heated to 100 ºC overnight. The reaction mixture was cooled down and filtered on a bed of celite, washed with DCM and the filtrate evaporated to dryness *in vacuo*. The residue was triturated with diethyl ether and the compound was filtered and dried to afford a white solid (740 mg, 75% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 12.3 (1H, bs); 10.0 (1H, s); 8.62 (1H, s); 8.06-7.42 (5H, m); 3.98 (2H, s); 3.37 (1H, s); 2.04 (1H, s); 0.99 (6H, d, *J* = 5.85 Hz). **LRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{17}H_{19}N_4O_3$) 327.3; found 327.3.

Methyl 6-{(*E*)-[2-(pyridin-2-ylcarbonyl)hydrazinylidene]methyl}pyridine-2-carboxylate **101**



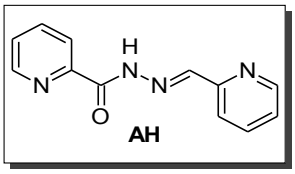
MF: $C_{14}H_{12}N_4O_3$

MW: 284 g.mol^{-1}

Ref: tcm3-118

To methyl 6-formylpyridine-2-carboxylate (130 mg, 0.78 mmol, 1 equiv) in MeOH (5 mL) was added pyridine-2-carbohydrazide (160 mg, 1.17 mmol, 1.5 equiv) and the reaction mixture was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (DCM/MeOH 98:2) to afford an off-white solid (100 mg, 46% yield).**^{1}H NMR** ($CDCl_3$, 300 MHz) δ 11.18 (1H, s), 8.59-8.57 (1H, d, *J* = 2.2 Hz); 8.46 (1H, s); 8.44-8.41 (1H, d, *J* = 5.8 Hz); 8.29-8.27 (1H, d, *J* = 5.8 Hz); 8.12-8.09 (1H, d, *J* = 6.6 Hz); 7.92-7.86 (2H, t); 7.51-7.46 (1H, t); 3.99 (3H, s) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 165.2, 160.2, 153.4, 148.6, 148.2, 147.8, 147.6, 137.6, 137.4, 126.9, 125.7, 124.3, 122.9, 52.9.

***N*-[(*E*)-pyridin-2-ylmethylidene]pyridine-2-carbohydrazide** **81**



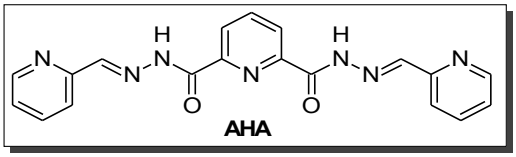
MF: $C_{12}H_{10}N_4O$

MW: 226 g.mol^{-1}

Ref: tcm3-076

To pyridine-2-carbohydrazide (300 mg, 2.19 mmol, 1 equiv) was added 2-pyridine carboxaldehyde (0.23 mL, 2.41 mmol, 1.1 equiv) in EtOH (3 mL) and refluxed overnight. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (DCM/MeOH 9:1) to afford a pale yellow solid (200 mg, 40% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 11.13 (1H, s); 8.60 (2H, t); 8.35 (1H, s); 8.32-8.29 (1H, d, *J* = 7.9 Hz); 8.25-8.22 (1H, d, *J* = 7.9 Hz); 7.92-7.87 (1H, t); 7.76-7.70 (1H, t); 7.51-7.47 (1H, q); 7.31-7.26 (1H, q) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 160.2, 152.9, 149.4, 148.9, 148.1, 137.6, 136.4, 126.9, 124.4, 123.0, 121.3. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{12}H_{10}N_4O$) 227.0933; found 227.0933.

N'2, N'6 -bis[(*E*)-pyridin-2-ylmethylidene]pyridine-2,6-dicarbohydrazide 83



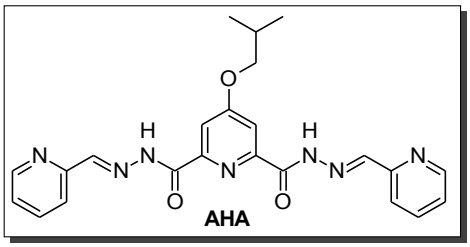
MF: $C_{19}H_{15}N_7O_2$

MW: 373.36 g.mol^{-1}

Ref: SC037, SC046, tcm3-065

Pyridine-2,6-dihydrazide (100 mg, 0.51 mmol, 1 equiv) was dissolved in EtOH (5 mL) and heated to reflux for 45 mins. To this was added 2-pyridinecarboxylate (123 µL, 1.13 mmol, 2.5 equiv). The reaction was stirred overnight at reflux and then allowed to cool to ambient temperature. A white precipitate was formed, which was filtered and washed with cold EtOH to yield a white solid (130 mg, 68% yield). No further purification was carried out and the product was used in the following step as is. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 12.31 (2H, s); 8.74 (2H, s); 8.55 (2H, d, *J* = 4.9 Hz); 8.42 (2H, d, *J* = 7.9 Hz); 8.15 (2H, d, J = 7.9 Hz); 8.04 (1H, t); 7.68 (2H, t); 7.23 (2H, t) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 160.0, 153.0, 150.6, 149.2, 148.0, 138.8, 136.2, 125.8, 124.1, 120.8. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{19}H_{16}N_7O_2$) 374.1365; found 374.1371.

4-isobutoxy-*N'2, N'6* -bis[(*E*)-pyridin-2-ylmethylidene]pyridine-2,6-dicarbohydrazide 84



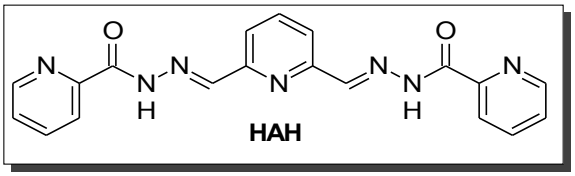
MF: $C_{23}H_{23}N_7O_3$

MW: 445.47 g.mol^{-1}

Ref: tcm3-055

2-pyridinecarboxaldehyde (22 µL, 0.23 mmol, 0.4 equiv) in EtOH (4 mL) was added drop wise to a round bottom flask, containing 4-(2-methylpropoxy)pyridine-2,6-dicarbohydrazide (150 mg, 0.56 mmol, 1 equiv) dissolved in EtOH (6 mL). The reaction was stirred overnight at reflux and then allowed to cool to ambient temperature. A white precipitate was formed, which was filtered and washed with cold EtOH to yield a white solid (145 mg, 87% yield). No further purification was carried out and the product was used in the following step as is. Note this procedure was initially carried out with the intention to achieve compound, 4-isobutoxy-*N'2*[(*E*)-pyridin-2-ylmethylidene]pyridine-2,6-dicarbohydrazide. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 12.27 (2H, s); 8.73 (2H, s); 8.57 (2H, d, *J* = 3.9 Hz); 8.17 (2H, d, *J* = 7.9 Hz); 7.90 (1H, s); 7.69 (2H, t); 7.24 (2H, t); 3.89 (2H, d, *J* = 6.4 Hz); 2.10 (1H, m); 1.00 (6H, d, *J* = 6.6 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 167.8, 160.1, 153.1, 150.5, 149.9, 149.3, 136.2, 124.1, 120.8, 112.1, 75.1, 27.8, 18.8

***N', N"*-[pyridine-2,6-diyldi(*E*)methylylidene]dipyridine-2-carbohydrazide 82**



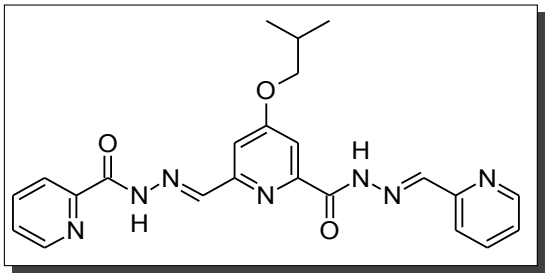
MF: $C_{19}H_{15}N_7O_2$

MW: 373.37 g.mol^{-1}

Ref: SC041

Pyridine-2,6-dicarbaldehyde (27 mg, 0.19 mmol, 1 equiv) was dissolved in EtOH (1 mL). To this was added pyridine-2-carbohydrazide (60 mg, 0.43 mmol, 2.2 equiv). The reaction was stirred overnight at reflux and then allowed to cool to ambient temperature. A white precipitate was formed, which was filtered and washed with cold EtOH to yield a white solid (48 mg, 65% yield). No further purification was carried out and the product was used in the following step as is. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 12.27 (2H, s); 8.73 (2H, s); 8.57 (2H, d, *J* = 3.9 Hz); 8.17 (2H, d, *J* = 7.9 Hz); 7.90 (1H, s); 7.69 (2H, t); 7.24 (2H, t); 3.89 (2H, d, *J* = 6.4 Hz); 2.10 (1H, m); 1.00 (6H, d, *J* = 6.6 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 162.4, 160.3, 152.3, 151.7, 150.0, 148.7, 148.4, 148.3, 138.9, 138.1, 137.8, 137.4, 127.1, 126.7, 126.0, 123.7, 123.1, 121.1. **HRMS (ESI):** *m/z* calcd $[M+H]^+$ ($C_{19}H_{16}N_7O_2$) 374.1365; found 374.1347.

4-isobutoxy-6-{(*E*)-[2-(pyridin-2-ylcarbonyl)hydrazinylidene]methyl}-*N'*-[(*E*)-pyridin-2-ylmethylidene]pyridine-2-carbohydrazide 103



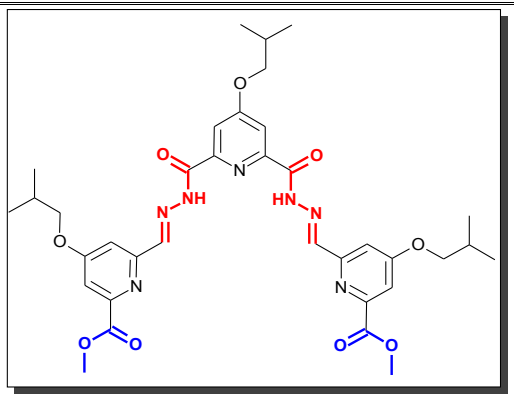
MF: $C_{23}H_{23}N_7O_3$

MW: 445.47 g.mol^{-1}

Ref: tcm3-139

To 6-formyl-4-isobutoxy-*N'*-[(*E*)-pyridin-2-ylmethylidiene]pyridine-2-carbohydrazide (150 mg, 0.46 mmol, 1 equiv) in anhydrous DCM (10 mL) was added Pyridine-2-carbohydrazide (63 mg, 0.46 mmol, 1 equiv) and was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (Chloroform/MeOH 99:1) to afford an off-white solid (106 mg, 52% yield). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 16.16 (1H, s, cis-cis species); 15.95 (1H, s, trans-cis species); 11.38 (1H, s, trans-cis species), 11.28 (1H, s, trans-cis species); 11.21 (1H, s, trans-trans species); 10.13 (1H, s); 10.02 (1H, s); 9.01-7.26 (8H, m); 3.90 (2H, d, *J* = 6.3 Hz); 2.12 (1H, m); 1.02 (6H, d, *J* = 6.6 Hz).

4-isobutoxy-6,6'-{pyridine-2,6-diylbis[carbonyl(1*E*)hydrazin-2-yl-1ylidene(E) methylyidene]} dipyridine-4-isobutoxy-2-carboxylate **104**



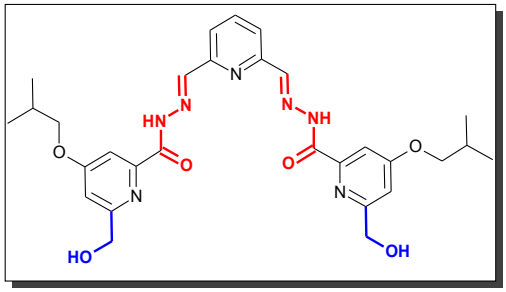
MF: $C_{35}H_{43}N_7O_9$

MW: 705.75 g.mol^{-1}

Ref: tcm3-086

To methyl 6-formyl-4-(isobutoxy)pyridine-2-carboxylate (200 mg, 0.84 mmol, 2 equiv) in MeOH (30 mL) and DCM (20 mL) was added 4-(isobutoxy)pyridine-2,6-dicarbohydrazide and was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was triturated with diethyl ether. The compound was filtered and dried to afford a white solid (289 mg, 98% yield). **^{1}H NMR** ($CDCl_3$ 300 MHz) δ 12.36 (1H, s); 10.52 (1H, s); 8.76 (1H, s); 7.72-7.58 (5H, m); 4.73 (2H, bs) 4.03 (6H, t), 3.90 (6H, s); 2.08 (3H, m); 1.01 (18H, d, *J* = 5.8 Hz) **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 168.0, 166.5, 165.3, 160.3, 154.8, 150.0, 149.6, 148.9, 113.9, 112.5, 109.0, 75.0, 52.8, 28.0, 19.0.

N',N''-[pyridine-2,6-diyldi(*E*)methylylidene]bis(6-hydroxymethyl-4-isobutoxypyridine-2-carbohydrazide) **107**



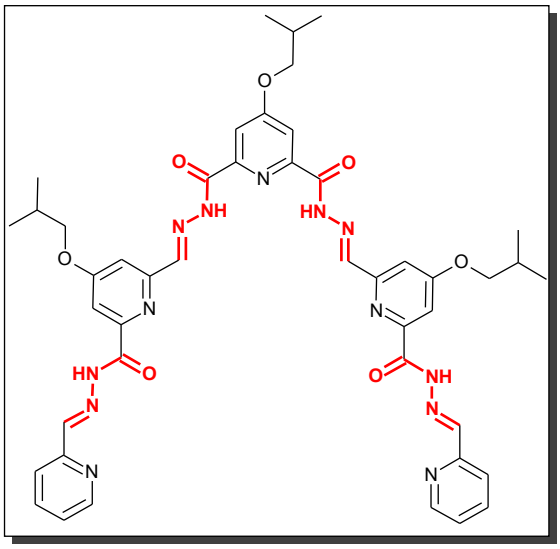
MF: $C_{29}H_{35}N_7O_6$

MW: 577.63 g.mol^{-1}

Ref: tcm3-085

To 6-(hydroxymethyl)-4-(isobutoxy)pyridine-2-carbohydrazide (147 mg, 0.61 mmol, 2 equiv) in MeOH (30 mL) and DCM (20 mL) was added pyridine-2,6-dicarbaldehyde (41 mg, 0.31 mmol, 1 equiv) and was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was triturated with diethyl ether. The compound was filtered and dried to afford a white solid (319 mg, 90% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 12.21 (2H, s); 8.72 (1H, s); 8.01 (2H, s); 7.48 (2H, d *J* = 2.2 Hz); 7.21 (2H, d, *J* = 2.1 Hz); 5.5 (1H, t); 4.74 (1H, t); 4.65 (4H, d *J* = 5.6 Hz) 3.94 (4H, bd, *J* = 6.4 Hz); 2.07 (2H, m); 1.02-1.00 (12H, d, *J* = 6.8 Hz).

2,6-{(*E*)-[4-isobutoxy-2-(pyridin-2-ylcarbonyl)hydrazinylidene]methyl}-N'-bis[4-isobutoxy(*E*)-pyridin-2ylmethylidene]pyridine-2-carbohydrazide **106**



MF: $C_{45}H_{49}N_{13}O_7$

MW: 883.95 g.mol^{-1}

Ref: tcm3-138

To 6-formyl-4-isobutoxy-*N'*-[(*E*)-pyridin-2-ylmethylidiene]pyridine-2-carbohydrazide (182 mg, 0.56 mmol, 2 equiv) in DCM (10 mL) was added 4-(isobutoxy)pyridine-2,6-dicarbohydrazide (75 mg, 0.28 mmol, 1 equiv) and the reaction was stirred at room temperature overnight. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (EtOAc/MeOH, 8:2) to afford a pale yellow solid (150 mg, 60% yield). **^{1}H NMR** (DMSO-d_6, 300 MHz) δ 12.65 (2H, s);12.11 (1H, s); 8.99 (2H, s); 8.61 (1H, s) 8.48 (1H, s); 8.13-7.34 (15 H, m); 3.98 (6H, s); 2.07 (3H, m), 1.02 (18H, d, *J* = 5.8 Hz).

Bibliography

[1] J.-M. Lehn, *Chemistry - A European Journal* **1999**, *5*, 2455–2463.

[2] P. A. Brady, R. P. Bonar-Law, S. J. Rowan, C. J. Suckling, J. K. M. Sanders, *Chem. Commun.* **1996**, 319–320.

[3] P. A. Brady, J. K. M. Sanders, *J. Chem. Soc., Perkin Trans. 1* **1997**, 3237–3254.

[4] B. Hasenknopf, J.-M. Lehn, N. Boumediene, A. Dupont-Gervais, A. Van Dorsselaer, B. Kneisel, D. Fenske, *Journal of the American Chemical Society* **1997**, *119*, 10956–10962.

[5] I. Huc, J.-M. Lehn, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 2106–2110.

[6] J. K. M. Sanders, Dynamic Combinatorial Chemistry - An Introduction, Marie Curie RTN Presentation, KTH, Stockholm, **2007**.

[7] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chemical Reviews* **2006**, *106*, 3652–3711.

[8] B. L. Miller, *Dynamic Combinatorial Chemistry: In Drug Discovery, Bioorganic Chemistry, and Materials Science*, Wiley-blackwell, **2010**.

[9] J. N. H. Reek, S. Otto, *Dynamic Combinatorial Chemistry*, Wiley Vch, **2010**.

[10] J. K. M. Sanders, *Drug Discovery Today* **2002**, *7*, 117–125.

[11] Huc I., Nguyen R., *Combinatorial Chemistry & High Throughput Screening* **2001**, *4*, 53–74.

[12] B. de Bruin, P. Hauwert, J. N. H. Reek, *Angewandte Chemie International Edition* **2006**, *45*, 2660–2663.

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

[14] U. Diederichsen, T. K. Lindhorst, B. Westermann, L. A. Wessjohann, *Bioorganic Chemistry: Highlights and New Aspects*, Wiley-vch, **1999**.

[15] S. Ladame, *Org. Biomol. Chem.* **2008**, *6*, 219–226.

[16] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.

[17] S. Otto, *Curr Opin Drug Discov Devel* **2003**, *6*, 509–520.

[18] G. L. Hornyak, H. F. Tibbals, J. Dutta, J. J. Moore, *Introduction to Nanoscience and Nanotechnology*, Crc Press, **2008**.

[19] J. Clayden, N. Greeves, S. Warren, P. Wothers, *Organic Chemistry*, Oup Oxford, **2001**.

[20] M. Egli, A. Flavell, A. M. Pyle, W. D. Wilson, S. I. Haq, B. Luisi, G. M. Blackburn, M. J. Gait, D. Loakes, D. M. Williams, *Nucleic Acids in Chemistry and Biology*, Royal Society Of Chemistry, **2005**.

[21] D. A. Braasch, D. R. Corey, *Biochemistry* **2002**, *41*, 4503–4510.

[22] J. Summerton, D. Weller, *Antisense and Nucleic Acid Drug Development* **1997**, *7*, 187 – 195.

[23] S. Sivakova, S. J. Rowan, *Chem. Soc. Rev.* **2005**, *34*, 9–21.

[24] E. T. Kool, *Chemical Reviews* **1997**, *97*, 1473–1488.

[25] A. D. Hamilton, D. Van Engen, *Journal of the American Chemical Society* **1987**, *109*, 5035–5036.

[26] B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, K. Parris, K. Williams, J. Rebek, *Journal of the American Chemical Society* **1989**, *111*, 1082–1090.

[27] S. C. Zimmerman, W. Wu, Z. Zeng, *Journal of the American Chemical Society* **1991**, *113*, 196–201.

[28] O. F. Schall, G. W. Gokel, *Journal of the American Chemical Society* **1994**, *116*, 6089–6100.

[29] J. C. Adrian, C. S. Wilcox, *Journal of the American Chemical Society* **1989**, *111*, 8055–8057.

[30] S. Goswami, D. Van Engen, A. D. Hamilton, *J. Am. Chem. Soc.* **1989**, *111*, 3425–3426.

[31] Y. Hisamatsu, H. Takami, N. Shirai, S. Ikeda, K. Odashima, *Tetrahedron Letters* **2007**, *48*, 617–621.

[32] Y. Hisamatsu, K. Hasada, F. Amano, Y. Tsubota, Y. Wasada-Tsutsui, N. Shirai, S. Ikeda, K. Odashima, *Chemistry - A European Journal* **2006**, *12*, 7733–7741.

[33] W. L. Jorgensen, J. Pranata, *Journal of the American Chemical Society* **1990**, *112*, 2008–2010.

[34] J. Pranata, S. G. Wierschke, W. L. Jorgensen, *Journal of the American Chemical Society* **1991**, *113*, 2810–2819.

[35] J. L. Sessler, B. Wang, A. Harriman, *Journal of the American Chemical Society* **1995**, *117*, 704–714.

[36] J. L. Sessler, J. Jayawickramarajah, M. Sathiosatham, C. L. Sherman, J. S. Brodbelt, *Organic Letters* **2003**, *5*, 2627–2630.

[37] E. Mertz, S. Mattei, S. C. Zimmerman, *Organic Letters* **2000**, *2*, 2931–2934.

[38] J. R. Quinn, S. C. Zimmerman, *The Journal of Organic Chemistry* **2005**, *70*, 7459–7467.

[39] S. Amemiya, P. Buhlmann, Y. Umezawa, *Chem. Commun.* **1997**, 1027–1028.

[40] M. W. Hosseini, A. J. Blacker, J. M. Lehn, *Journal of the American Chemical Society* **1990**, *112*, 3896–3904.

[41] G. Deslongchamps, A. Galán, J. de Mendoza, J. R. Jr, *Angewandte Chemie International Edition in English* **1992**, *31*, 61–63.

[42] W.-S. Yeo, J.-I. Hong, *Tetrahedron Letters* **1998**, *39*, 3769–3772.

[43] S. E. Schneider, S. N. O'Nei, E. V. Anslyn, *Journal of the American Chemical Society* **2000**, *122*, 542–543.

[44] J. Y. Kwon, N. J. Singh, H. N. Kim, S. K. Kim, K. S. Kim, J. Yoon, *Journal of the American Chemical Society* **2004**, *126*, 8892–8893.

[45] H. Y. Kuchelmeister, C. Schmuck, *Chemistry - A European Journal* **2011**, *17*, 5311–5318.

[46] K. Yoshimoto, S. Nishizawa, M. Minagawa, N. Teramae, *Journal of the American*

Chemical Society **2003**, *125*, 8982–8983.

[47] R. Naylor, P. T. Gilham, *Biochemistry* **1966**, *5*, 2722–2728.

[48] O. Leslie E., *Critical Reviews in Biochemistry and Molecular Biology* **2004**, *39*, 99–123.

[49] T. Inoue, L. Orgel, *Science* **1983**, *219*, 859–862.

[50] R. Iwaura, F. J. M. Hoeben, M. Masuda, A. P. H. J. Schenning, E. W. Meijer, T. Shimizu, *Journal of the American Chemical Society* **2006**, *128*, 13298–13304.

[51] P. G. A. Janssen, J. Vandenbergh, J. L. J. van Dongen, E. W. Meijer, A. P. H. J. Schenning, *Journal of the American Chemical Society* **2007**, *129*, 6078–6079.

[52] P. G. A. Janssen, S. Jabbari-Farouji, M. Surin, X. Vila, J. C. Gielen, T. F. A. de Greef, M. R. J. Vos, P. H. H. Bomans, N. A. J. M. Sommerdijk, P. C. M. Christianen, et al., *Journal of the American Chemical Society* **2009**, *131*, 1222–1231.

[53] P. G. A. Janssen, A. Ruiz-Carretero, D. González-Rodríguez, E. W. Meijer, A. P. H. J. Schenning, *Angewandte Chemie International Edition* **2009**, *48*, 8103–8106.

[54] A. Bugaut, J.-J. Toulmé, B. Rayner, *Angewandte Chemie International Edition* **2004**, *43*, 3144–3147.

[55] A. Bugaut, K. Bathany, J.-M. Schmitter, B. Rayner, *Tetrahedron Letters* **2005**, *46*, 687–690.

[56] N. C. Seeman, *Mol Biotechnol* **2007**, *37*, 246–257.

[57] S. Hoffmann, *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1013–1016.

[58] A. Eschenmoser, E. Loewenthal, *Chem. Soc. Rev.* **1992**, *21*, 1.

[59] D. Sievers, G. von Kiedrowski, *Nature* **1994**, *369*, 221–224.

[60] P. A. Bachmann, P. L. Luisi, J. Lang, *Nature* **1992**, *357*, 57–59.

[61] J. Visscher, C. Bakker, R. van der Woerd, A. Schwartz, *Science* **1989**, *244*, 329–331.

[62] G. R. Fleischaker, S. Colonna, P. L. Luisi, *Self-Production of Supramolecular Structures: From Synthetic Structures to Models of Minimal Living Systems (NATO Science Series C:*, Springer, **1994**.

[63] L. E. Orgel, *Nature* **1992**, *358*, 203–209.

[64] T. Wu, L. E. Orgel, *Journal of the American Chemical Society* **1992**, *114*, 317–322.

[65] B. G. Bag, G. von Kiedrowski, *Pure Appl. Chem.* **1996**, *68*, 2145–2152.

[66] J. T. Goodwin, D. G. Lynn, *Journal of the American Chemical Society* **1992**, *114*, 9197–9198.

[67] Z.-Y. J. Zhan, D. G. Lynn, *Journal of the American Chemical Society* **1997**, *119*, 12420–12421.

[68] X. Li, Z.-Y. J. Zhan, R. Knipe, D. G. Lynn, *Journal of the American Chemical Society* **2002**, *124*, 746–747.

[69] N. A. Leal, M. Sukeda, S. A. Benner, *Nucleic Acids Research* **n.d.**, *34*, 4702 –4710.

[70] M. Q. Zhang, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 565 –568.

[71] Zhan Z-y.J., Ye J., Li X., Lynn D.G., *Current Organic Chemistry* **2001**, *5*, 885–902.

[72] S. Sando, H. Abe, E. T. Kool, *Journal of the American Chemical Society* **2004**, *126*, 1081–1087.

[73] K. D. James, A. D. Ellington, *Origins of Life and Evolution of Biospheres* **1999**, *29*, 375–390.

[74] Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel, M. R. Ghadiri, *Science* **2009**, *325*, 73 –77.

[75] A. Carruthers, Marshall, Gary, Marshall, Tracey, *Graphics Design*, **2012**.

[76] S. Boncel, K. Walczak, *Letters in Organic Chemistry* **2006**, *3*, 534–538.

[77] R. Chenevert, M. Dickman, *The Journal of Organic Chemistry* **1996**, *61*, 3332–3341.

[78] A. J. Brouwer, R. M. J. Liskamp, *European Journal of Organic Chemistry* **2005**, *2005*, 487–495.

[79] B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *The Journal of Organic Chemistry* **2010**, *75*, 7175–7185.

[80] R. H. Hudson, F. Wojciechowski, *Canadian Journal of Chemistry* **2005**, *83*, 1731–1740.

[81] J. Marchand-Brynaert, L. Ghosez, *Journal of the American Chemical Society* **1972**, *94*, 2869–2870.

[82] Y. Kuroda, J. M. Lintuluoto, H. Ogoshi, *J. Chem. Soc., Perkin Trans. 2* **1997**, 333–340.

[83] Zeyuan Dong, *Post-doc Continued with the Research and Carried Out These Experiments*, **2010**.

[84] H. Hayakawa, K. Haraguchi, H. Tanaka, T. Miyasaka, *Chem. Pharm. Bull.* **1987**, *35*, 72–79.

[85] E. R. Gillies, F. Deiss, C. Staedel, J.-M. Schmitter, I. Huc, *Angewandte Chemie International Edition* **2007**, *46*, 4081–4084.

[86] H.-Y. Hu, Y. Yang, J.-F. Xiang, C.-F. Chen, *Chinese Journal of Chemistry* **2007**, *25*, 1389–1393.

[87] Y.-S. Liu, C. Zhao, D. E. Bergbreiter, D. Romo, *The Journal of Organic Chemistry* **1998**, *63*, 3471–3473.

[88] D. Amantini, R. Beleggia, F. Fringuelli, F. Pizzo, L. Vaccaro, *The Journal of Organic Chemistry* **2004**, *69*, 2896–2898.

[89] Y. Kaburagi, Y. Kishi, *Organic Letters* **2007**, *9*, 723–726.

[90] J. R. Quinn, S. C. Zimmerman, *The Journal of Organic Chemistry* **2005**, *70*, 7459–7467.

[91] M. M. Conn, G. Deslongchamps, J. de Mendoza, J. Rebek, *Journal of the American Chemical Society* **1993**, *115*, 3548–3557.

[92] S. C. Zimmerman, W. Wu, Z. Zeng, *Journal of the American Chemical Society* **1991**, *113*, 196–201.

[93] R. K. Castellano, V. Gramlich, F. Diederich, *Chemistry - A European Journal* **2002**, *8*, 118–129.

[94] Frédéric Godde, *Carried Out the SPR Experiments*, **2011**.

[95] S. Neidle, S. Balasubramanian, *Quadruplex Nucleic Acids*, Royal Society Of Chemistry, **2006**.

[96] H. Jiang, J.-M. Léger, C. Dolain, P. Guionneau, I. Huc, *Tetrahedron* **2003**, *59*, 8365–8374.

[97] P. V. Jena, P. S. Shirude, B. Okumus, K. Laxmi-Reddy, F. Godde, I. Huc, S. Balasubramanian, T. Ha, *Journal of the American Chemical Society* **2009**, *131*, 12522–12523.

[98] S. Hecht, I. Huc, *Foldamers: Structure, Properties and Applications*, Wiley Vch, **2007**.

[99] G. Guichard, I. Huc, *Chem. Commun.* **2011**, *47*, DOI 10.1039/C1CC11137J.

[100] I. Huc, *European Journal of Organic Chemistry* **2004**, *2004*, 17–29.

[101] I. Huc, V. Maurizot, H. Gornitzka, J.-M. Leger, *Chem. Commun.* **2002**, 578–579.

[102] V. Berl, I. Huc, R. G. Khoury, J.-M. Lehn, *Chemistry - A European Journal* **2001**, *7*, 2798–2809.

[103] V. Berl, I. Huc, R. G. Khoury, M. J. Krische, J.-M. Lehn, *Nature* **2000**, *407*, 720–723.

[104] V. Maurizot, Hélices Simples Et Hélices Doubles D'oligoamides Aromatiques De Synthèse, PhD Thesis, IECB, University of Bordeaux 1, **2003**.

[105] A.-M. Stadler, N. Kyritsakas, J.-M. Lehn, *Chem. Commun.* **2004**, 2024–2025.

[106] K. Oh, K.-S. Jeong, J. S. Moore, *Nature* **2001**, *414*, 889–893.

[107] V. Berl, M. J. Krische, I. Huc, J.-M. Lehn, M. Schmutz, *Chemistry - A European Journal* **2000**, *6*, 1938–1946.

[108] J. Schmitt, A. Stadler, N. Kyritsakas, J. Lehn, *Helvetica Chimica Acta* **2003**, *86*, 1598–1624.

[109] M. Ohkita, J. Lehn, G. Baum, D. Fenske, *Chemistry - A European Journal* **1999**, *5*, 3471–3481.

[110] C. Dolain, V. Maurizot, I. Huc, *Angewandte Chemie International Edition* **2003**, *42*, 2738–2740.

[111] I. Huc, Dynamic Equilibria and Folding Processes, Marie Curie RTN Presentation, KTH, Stockholm, **2007**.

[112] D. Zhao, J. S. Moore, *Macromolecules* **2003**, *36*, 2712–2720.

[113] D. Zhao, J. S. Moore, *Journal of the American Chemical Society* **2002**, *124*, 9996–9997.

[114] D. Zhao, J. S. Moore, *Org. Biomol. Chem.* **2003**, *1*, 3471.

[115] S. Chauffaille, Synthèse Et Analyse Conformationnelle De Motifs Pyridine Acylhydrazone, Stage de Licence, IECB, Universite Bordeaux 1, **2005**.

[116] C. Bao, Q. Gan, B. Kauffmann, H. Jiang, I. Huc, *Chemistry - A European Journal* **2009**, *15*, 11530–11536.

[117] J. Garric, Conception, Synthèse Et Étude De Nouvelles Capsules Hélicoïdales :des Pelures De Pomme Moléculaires., PhD Thesis, IECB, Universite Bordeaux 1, **2005**.

[118] U. Lüning, R. Baumstark, K. Peters, H. G. von Schnering, *Liebigs Ann. Chem.* **1990**, *1990*, 129–143.

[119] N. Delsuc, Structures Tertiaires Proteomimetiques Issues D'oligoamides Aromatiques., PhD Thesis, IECB, University of Bordeaux 1, **2007**.

[120] N. Delsuc, F. Godde, B. Kauffmann, J.-M. Léger, I. Huc, *Journal of the American Chemical Society* **2007**, *129*, 11348–11349.

[121] M. N. Chaur, D. Collado, J. Lehn, *Chemistry - A European Journal* **2011**, *17*, 248–258.

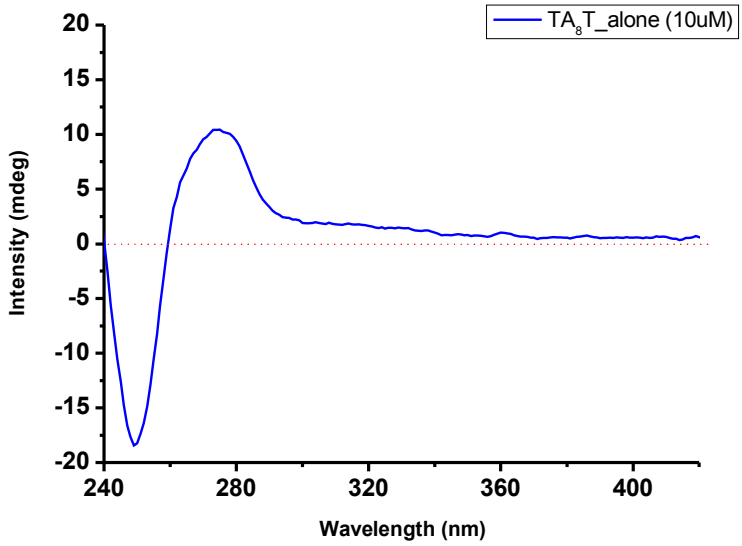
[122] F. Dumitru, Y.-M. Legrand, M. Barboiu, E. Petit, A. van der Lee, *Crystal Growth & Design* **2009**, *9*, 2917–2921.

DNA Sequence to Test:

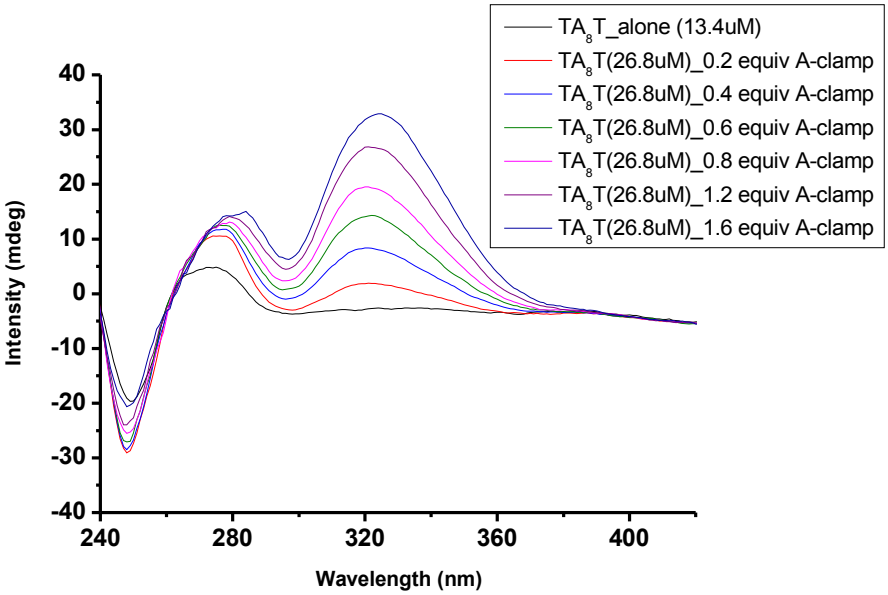
Name	Sequence	M_W	ε
TA3T	TAA-AT	1486	54800
TA5T	TAA-AAA-T	2112	78800
TA8T	TAA-AAA-AAA-T	3052	114800
TA13T	TAA-AAA-AAA-AAA-AAT	4618	174800
TA4GA3T	TAA-AAG-AAA-T	3068	114100
TA3G2A3T	TAA-AGG-AAA-T	3084	112200
TA2G3A3T	TAA-GGG-AAA-T	3100	110300
TA2GAGA3T	TAA-GAG-AAA-T	3084	113400
Biot-TA8T	Biot-TEG- TAA-AAA-AAA-T		114800
Biot-TG8T	Biot-TEG- TGG-GGG-GGG-T		98200
A-clamp		695	33000

Raw CD and UV/Vis Data

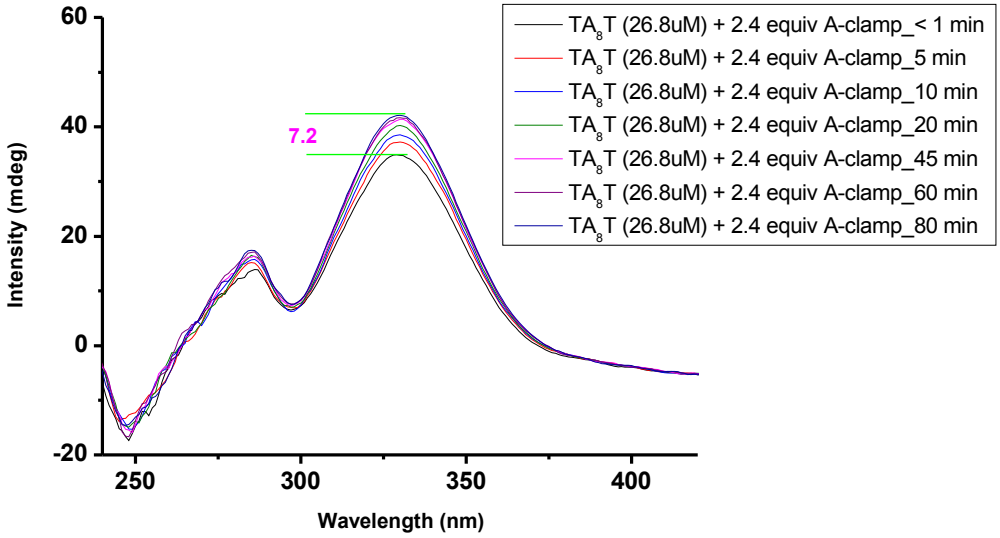
CD Spectra: TA_8T (10 μM): Template only, no A clamp present. PBS buffer. pH 8.0, 10 mM, 150 mM NaCl, 1% DMSO.



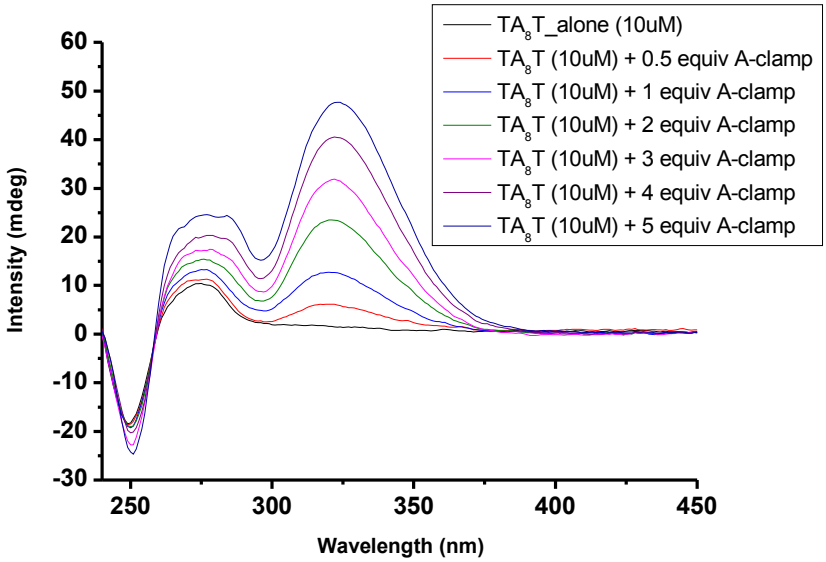
CD Spectra Titration Optimisation: TA_8T (26.8 μM): Testing the effects of template concentration at 26.8 μM. Addition of A clamp (0.2 equiv to 1.6 equiv). Change in λ_{max} from 320 nm to 323 nm.



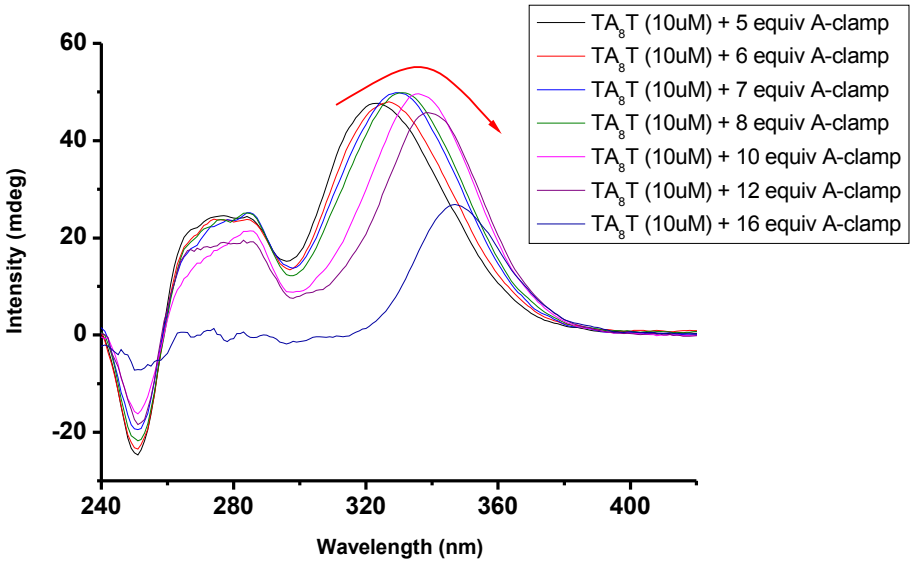
CD Spectra Titration Optimisation: TA_8T (26.8 μM): Monitoring the dynamic processes between TA_8T and 2.4 equiv of A clamp. Equilibrium reached after 60 minutes. Thus time is require in order to reach equilibrium. λ_{max} = 330 nm



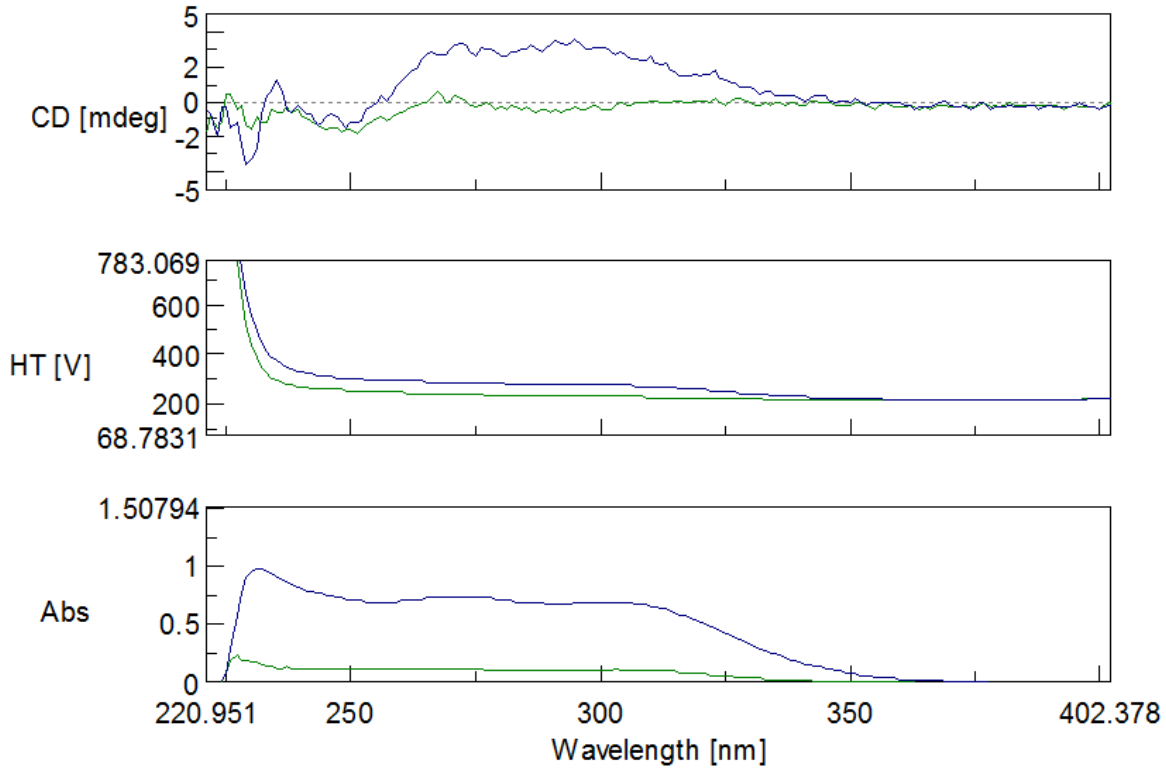
CD Spectra Titration Optimisation: TA_8T (10 μM): Testing the effects of template concentration at 10 μM. Addition of A clamp (0.5 equiv to 5 equiv). λ_{max} 323 nm. PBS buffer. pH 8.0, 10 mM, 150 mM NaCl, 1% DMSO.



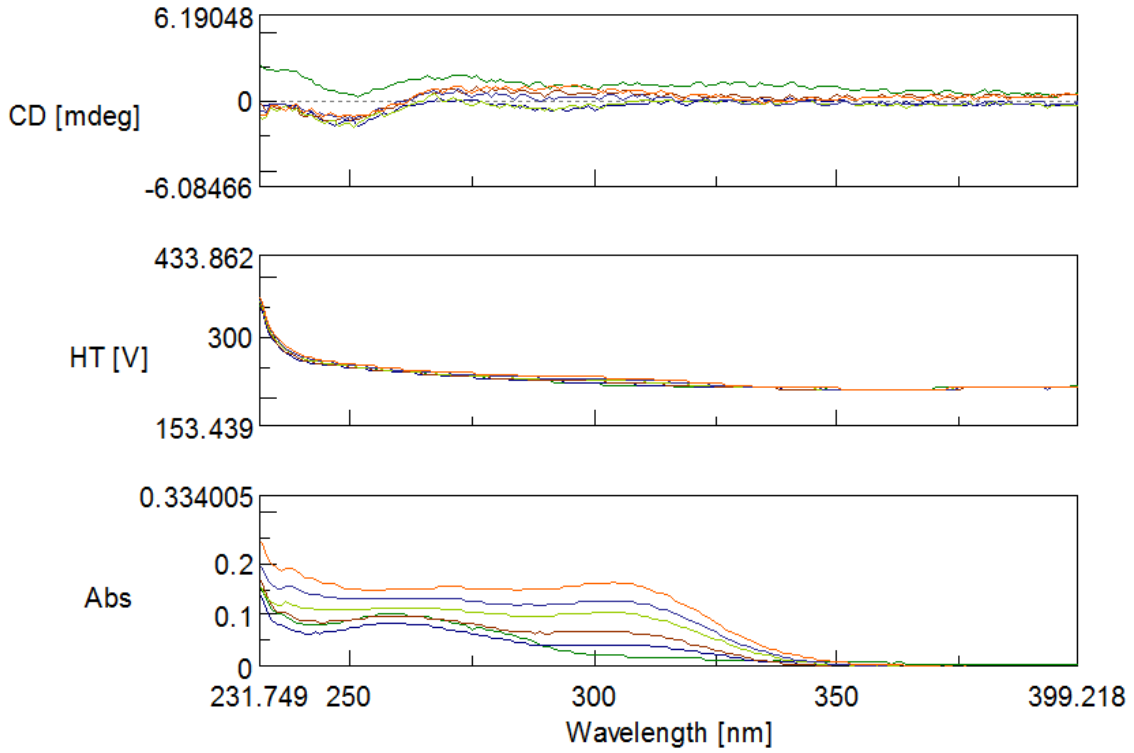
CD Spectra Titration Optimisation: **TA_8T (10 μM):** Testing the effects of template concentration at 10 μM. Addition of A clamp (5 equiv to 16 equiv). λ_{max} between 323 nm and 350 nm. Thus concentration too high. PBS buffer. pH 8.0, 10 mM, 150 mM NaCl, 1% DMSO.



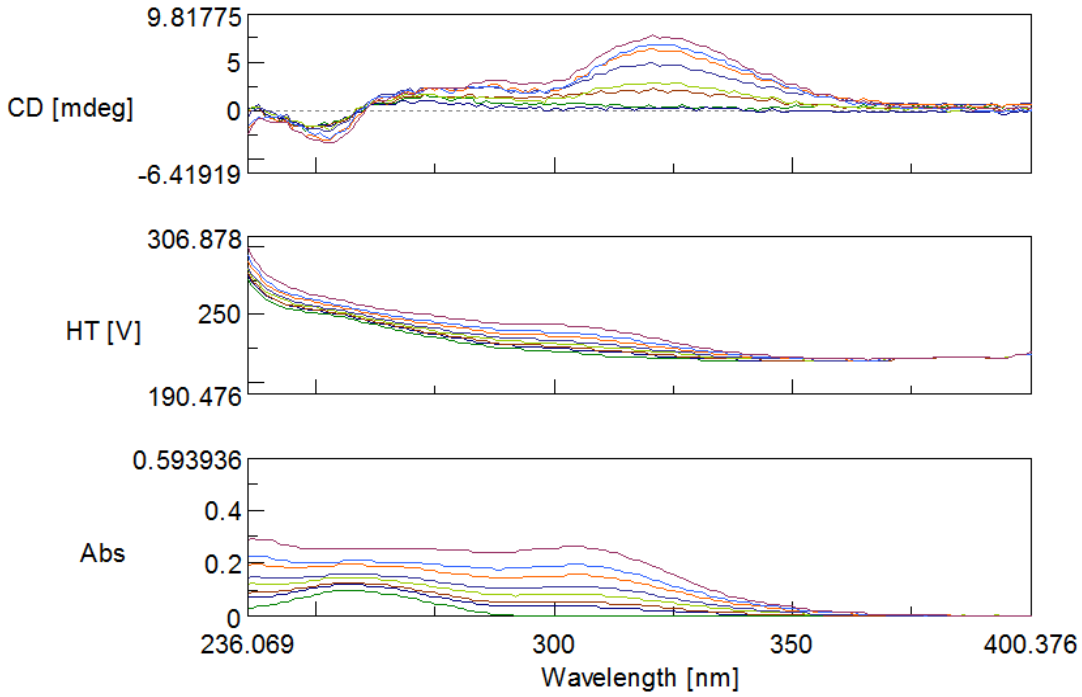
CD Spectra: TA_3T (2 μM) and TA_3T (10 μM): Testing the effects of increasing concentration.



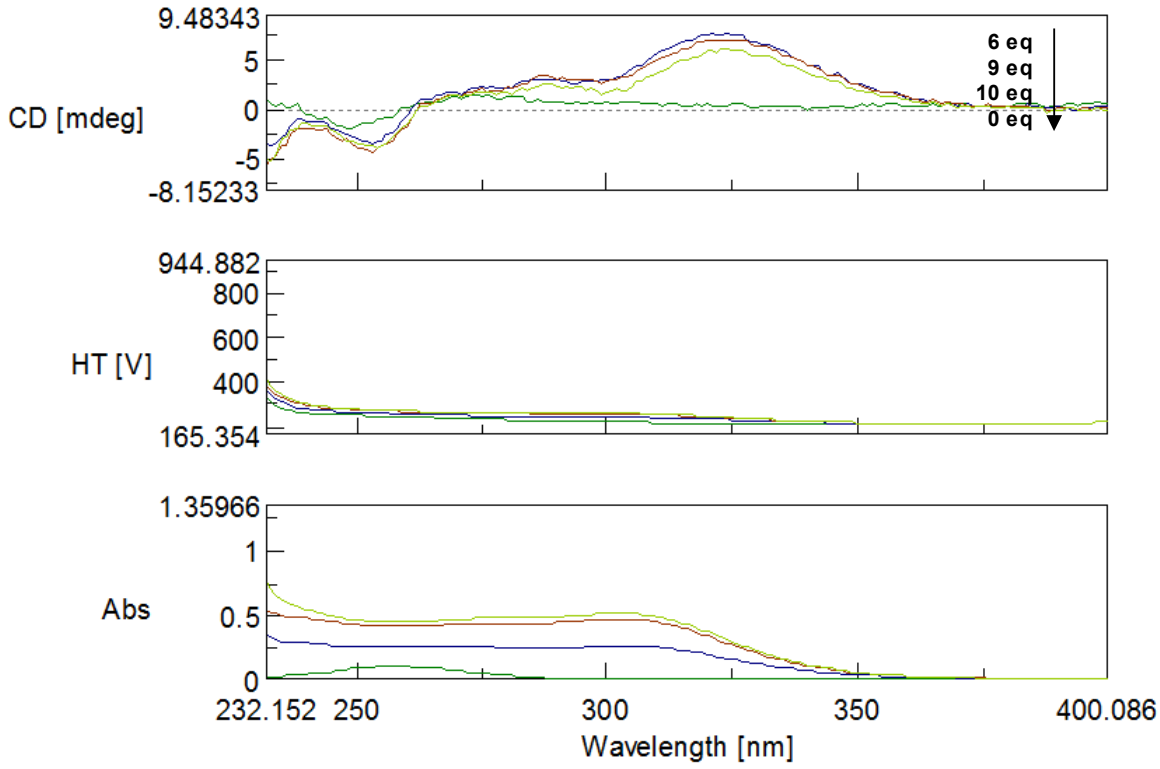
CD Spectra: TA_3T (2 μM) and addition of A Clamp (0 ~ 5 equiv):



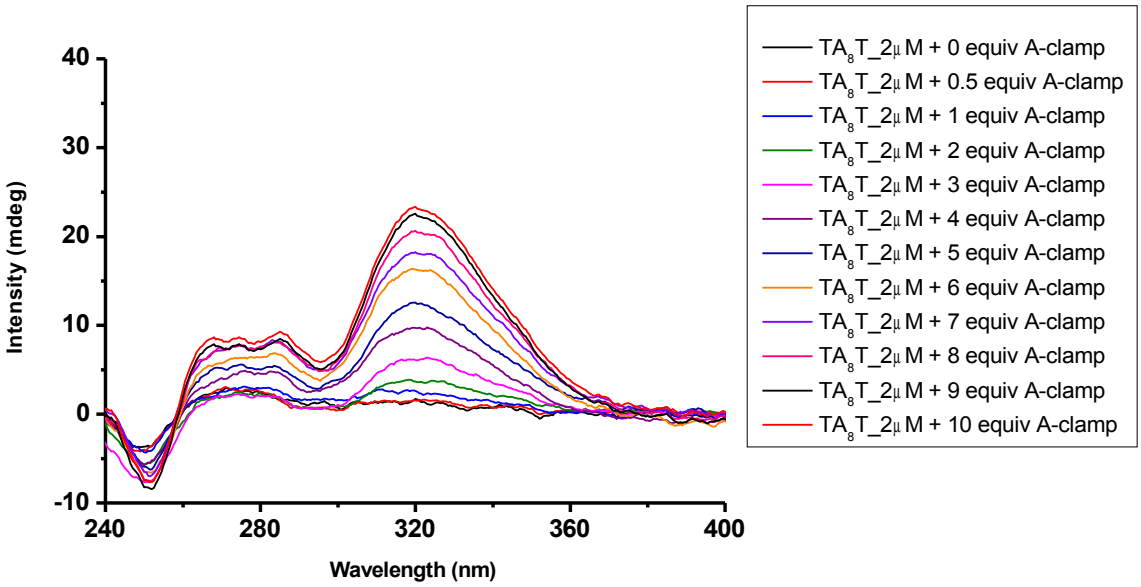
CD Spectra: TA_5T (2 μM) and addition of A Clamp (0 ~ 6 equiv):



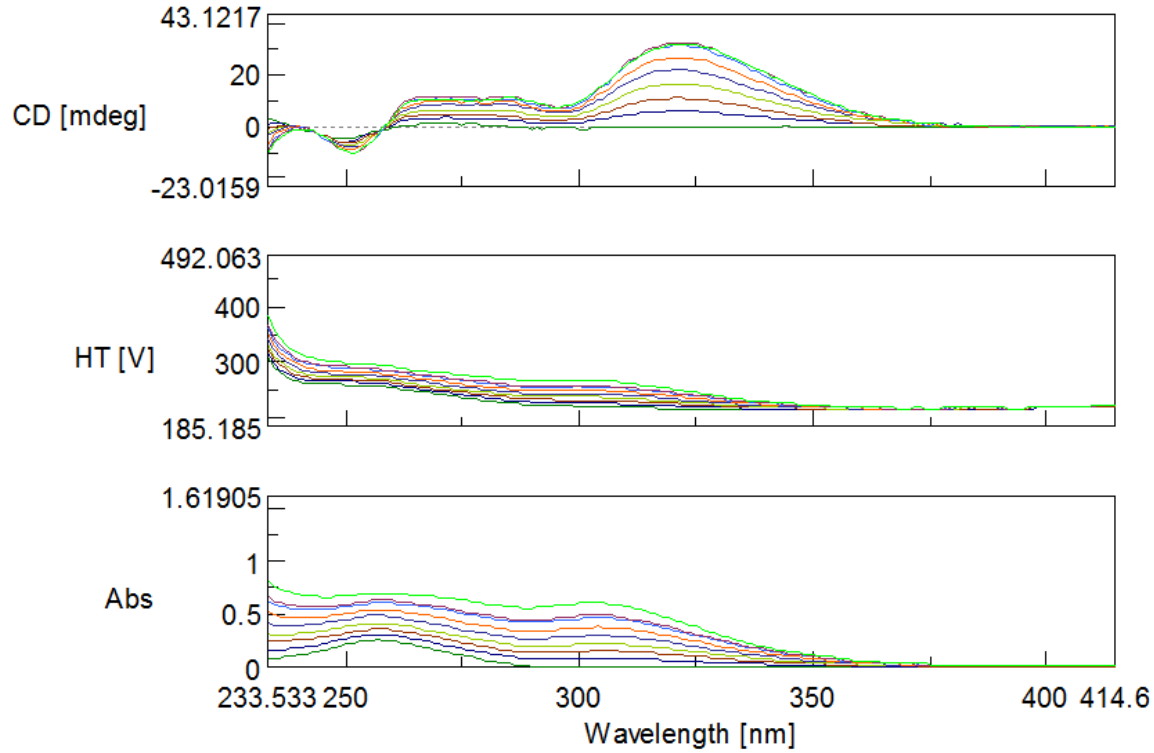
CD Spectra: TA_5T (2 μM) and addition of A Clamp (0, 10, 9, 6 equiv):



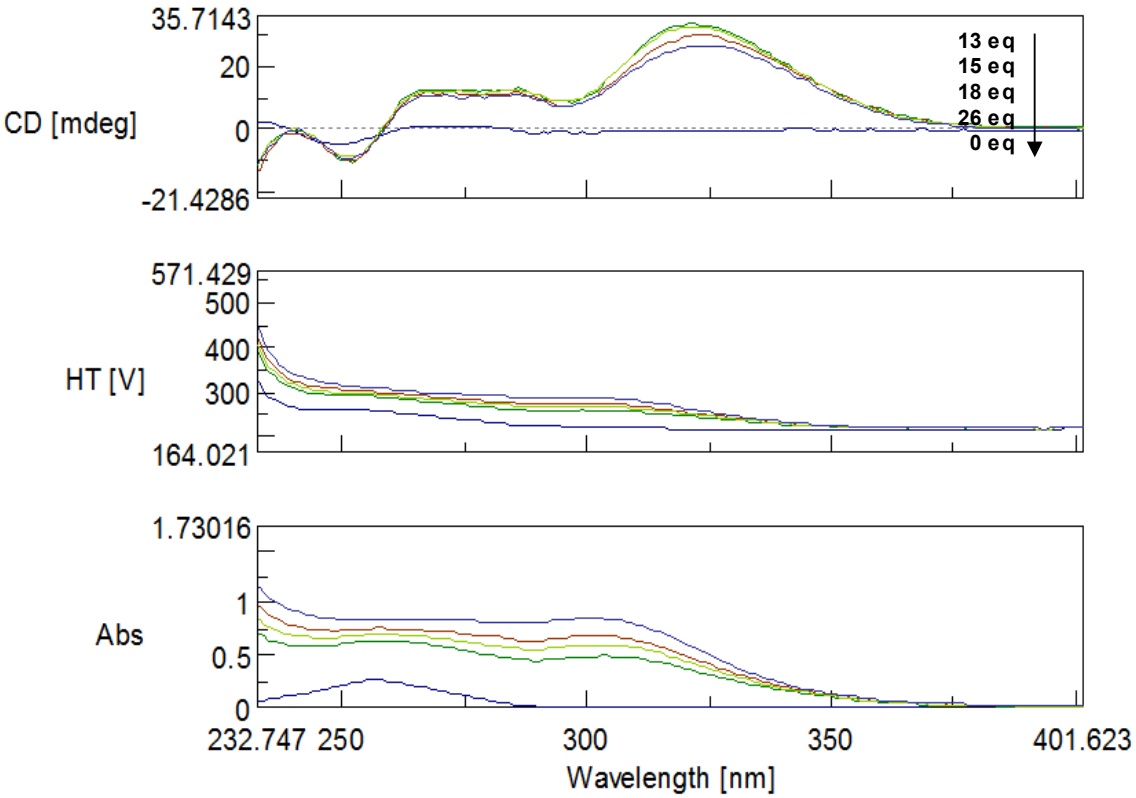
CD Spectra: TA_8T (2 μM) and addition of A Clamp (0 ~ 10 equiv):



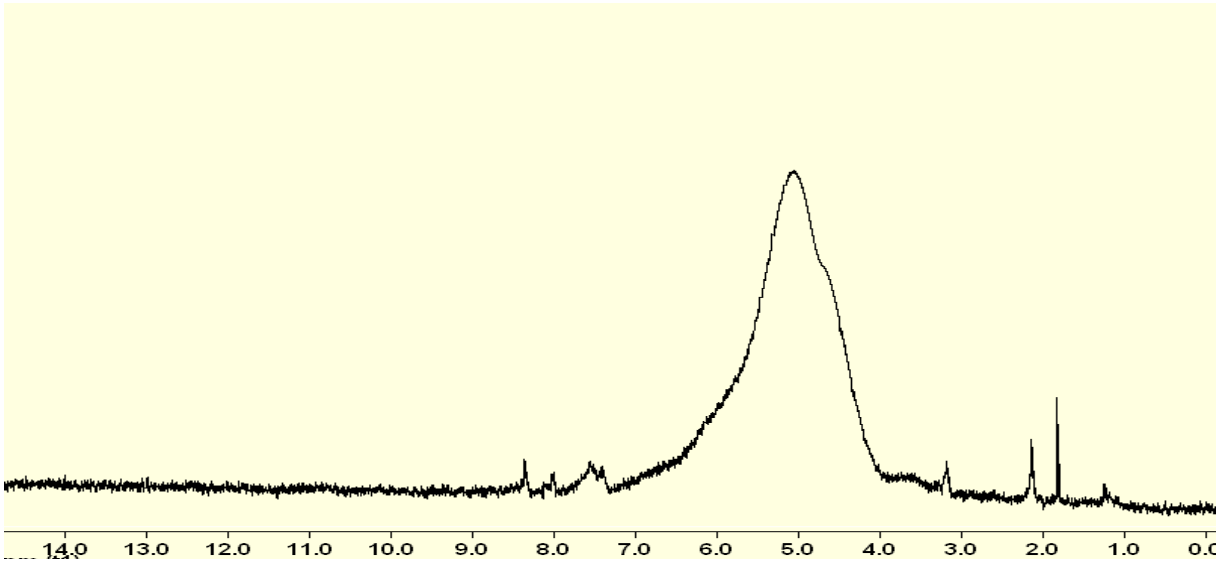
CD Spectra: $TA_{13}T$ (2 μM) and addition of A Clamp (0 ~ 15 equiv):



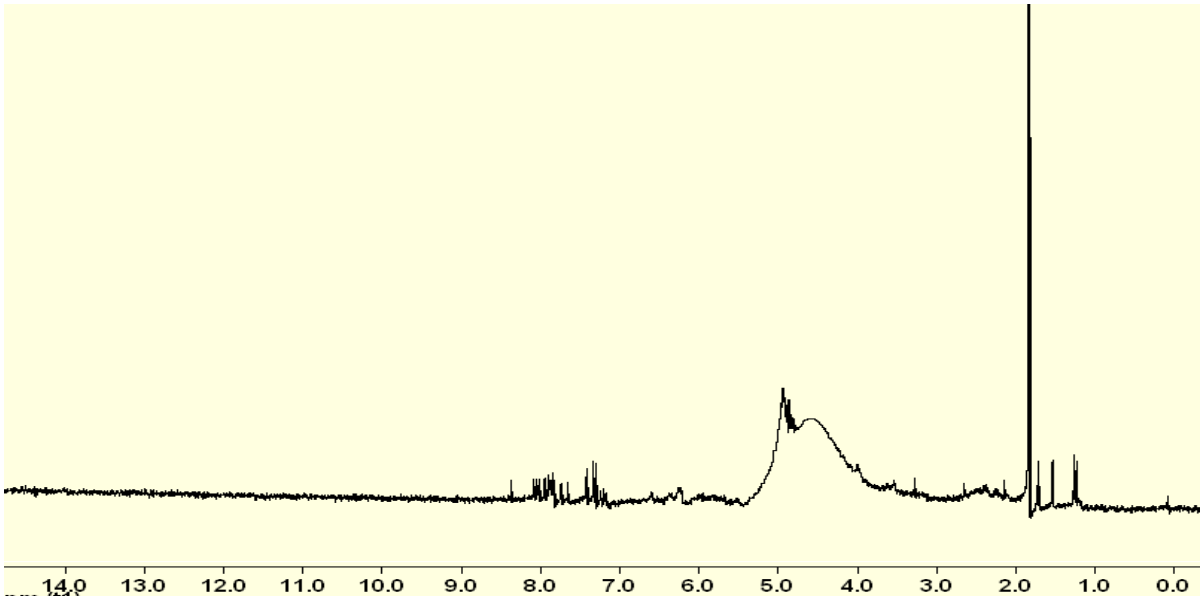
CD Spectra: $TA_{13}T$ (2 μM) and addition of A Clamp (0, 26, 18, 15, 13 equiv):



^{1}H NMR Spectra: A clamp (~0.26 mM in 10% D_2O/H_2O)
Initial suspension, followed by precipitate due to poor solubility.

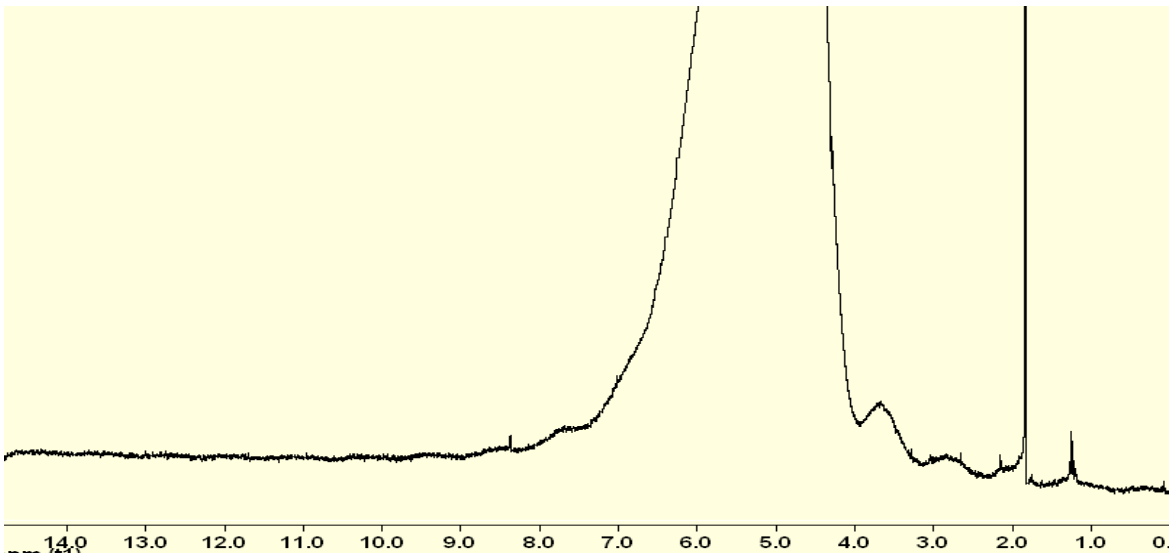


^{1}H NMR Spectra: A clamp (~0.26 mM in 10% D_2O/H_2O) and TA_8T template

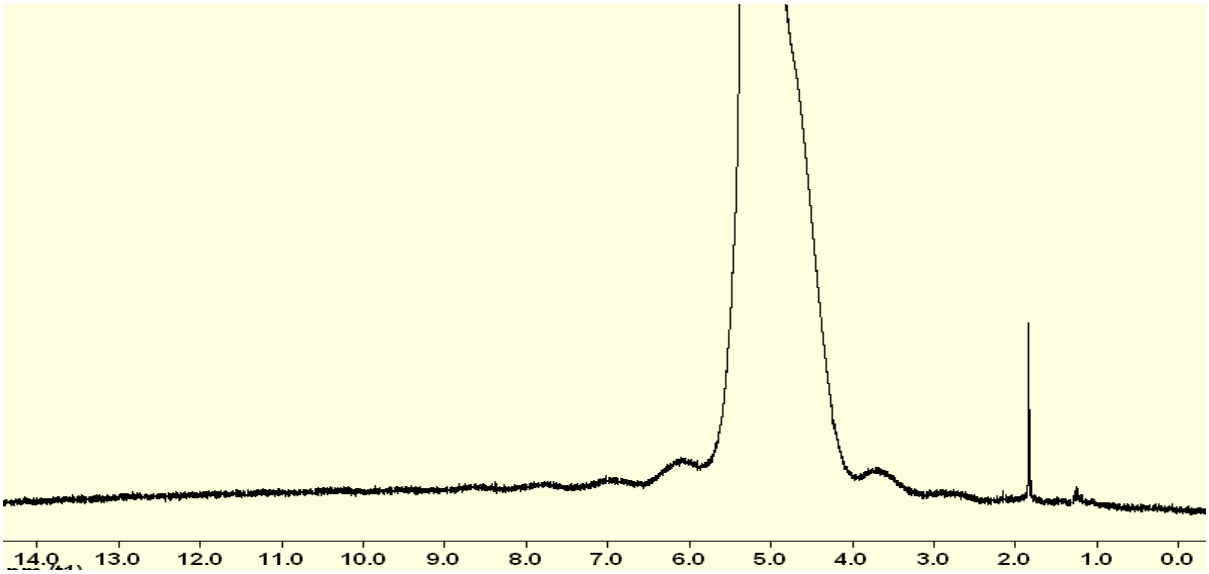


Inconclusive NMR results due to precipitation, thus future suggestion is to try Solid State NMR between the A clamp and TA_8T template assembly.

^{1}H NMR Spectra: A clamp (~0.5 mM in 10% D_2O/H_2O) and TA_8T template
Sample after 2 hours – precipitate observed.



^{1}H NMR Spectra: A clamp (~0.5 mM in 10% D_2O/H_2O) and TA_8T template
Sample after 2 hours – precipitate observed, thus possible aggregates.



Scientific Training

Oral Presentations

"Dynamic chemistry: nucleobase recognition by synthetic receptors and cis-trans acylhydrazone isomerism"

EU Marie Curie Research Training Network in Dynamic Combinatorial Chemistry

1. **July 2010** Las Palmas, Gran Canaria, Spain
2. **November 2009** University of Twente, Netherlands
3. **June 2009** Christian Albrechts Universität Kiel, Germany
4. **January 2009** University of Cambridge, UK
5. **July 2008** University of Geneva, Switzerland
6. **January 2008** ICIQ, Terragona, Spain
7. **June 2007** KTH, Royal Institute of Technology, Stockholm, Sweden

Presentations (total of 40) were also given on a triweekly basis from January 2007-July 2009 to fellow colleagues to update the Ivan Huc group on the research progress.

Poster Presentations

"Designing Aromatic Amide Foldamer Ligands for Nucleic Acid Recognition"

- **November 2008** 1st Chemistry and Biology of Membranes and Nano-objects Congress; IECB, Bordeaux, France
- **July 2008** 20th International Symposium on Chirality, (University of Geneva), Switzerland

"A New Adenine Receptor: Design and Synthesis"

- **May 2008** 1st Journées Jeunes Chercheurs, IECB, Bordeaux, France
- **October 2007** 1st Aquitaine Conference on Polymers, Arcachon, France
- **July 2007** 21st International Congress Heterocyclic Chemistry, (UNSW), Sydney, Australia

Attended Conferences

2007

April 9th Annual Meeting of the Chemistry PhD School of Bordeaux, France.

July 20th International Congress for Heterocyclic Chemistry, Sydney, Australia.

October 1st Aquitaine Conference on Polymers, Arcachon, France.

2008

April 10th Annual Meeting of the Chemistry PhD School of Bordeaux, France.

May 1st Journée Jeunes Chercheurs, (IECB), *"Various lectures by Post Graduate*

Researchers in Université Bordeaux"

July 20th International Symposium on Chirality, (University of Geneva), Switzerland.

October 2nd Aquitaine Conference on Polymers, Arcachon, France.

November 1st Chemistry & Biology of Membranes & Nanoobjects (CBMN) Workshop.

2009

April 11th Annual Meeting of the Chemistry PhD School of Bordeaux, France.

May 2nd Journée Jeunes Chercheurs, (IECB), *"Various lectures by Post Graduate Researchers in Université Bordeaux"*

December 2nd Chemistry & Biology of Membranes & Nanoobjects (CBMN) Workshop.

2010

January 1st Foldamer Symposium, (IECB), Bordeaux, France.

May 3rd Journée Jeunes Chercheurs, (IECB), *"Various lectures by Post Graduate Researchers in Université Bordeaux"*

Workshops/Training

Over 440 hours of scientific lectures, workshops and training was attended during my PhD.

EU Marie Curie Research Training Network in Dyanmic Combinatorial Chemistry

1. **July 2010** Las Palmas, Gran Canaria, Spain

Workshop on Careers and CV writing (7 hours)

2. **November 2009** University of Twente, Netherlands

Course on Nanotechnology (30 hours)

- Nanofabrication and Nanomaterials (8 hours)
- Nano-optics (7 hours)
- Nanoelectronics (4 hours)
- Bionanotechnology and Nanofluidics (7 hours)
- Nanotechnology and Society (4 hours)

3. **June 2009** Christian Albrechts Universität Kiel, Germany

Workshop Intellectual Property, Patents and how to found a company (7 hours)

- Dr. Heino Steentoft, (CAU Kiel), *"Technology Transfer"*
- Dr. Christian Manthey, (IP Bewertungs-AG), *"Creating Value from Intellectual Property"*
- Dr. Jeremy Carmichael, (Astex Therapeutics), *"IP, University-Industry Interaction and Agreements Commercialization of a Process"*
- Prof. Ralf Thiericke, (IZET), *"How to Found a Company"*

4. **April 2009** University of Groningen

Workshop on Systems Chemistry (14 hours)

Various guest speakers including: Prof. Bert Meijer, (Eindhoven University of Technology); Prof. Stefan Matile, (University of Geneva); Prof. Günter von Kiedrowski, (Ruhr University of Bochum); Prof. Alan Rowan, (Radboud University Nijmegen); Prof. Ben Feringa, (University of Groningen); Prof. Douglas Philp, (University of St.Andrew's); Dr. Sijbren Otto, (University of Groningen); Prof. Gonen Ashkenasy, (University of the Negev); Prof. Piet Herdewijn, (University of Leuven).

5. **February 2009** Brussels, Belgium

Career Development and Grant Proposal Writing

- Dr. Christina Schütte, (ProSciencia GmbH & Co. KG)
- Dr. Susan Kentner, (Helmholtz Association)
- Dr. Pierre Van Antwerpen, (Université Libre de Bruxelles)

6. **January 2009** University of Cambridge, UK

Workshop Biocatalysis

- Prof. Sabine Flitsch, (University of Manchester), *Biocatalysis,* (8 hours)

7. **July 2008** University of Geneva, Switzerland

Workshop on Chirality

- Th. Wenzel, (Bates College), *"NMR Spectroscopy for Determining Molecular Stereochemistry and Enantiomeric Purity"*
- N. Berova, (Columbia University), P. Polavarapu, (Vanderbilt University), *"Circular Dichroism, Vibrational Circular Dichroism VCD "*

8. **January 2008** ICIQ, Terragona, Spain

Workshop Supramolecular Chemistry: Self Assembly & Metal Co-ordination.

- Stefan Kubik, (Technische Universität Kaiserslautern), *"Strategies for Anion Coordination"*
- Kay Severin, (EPFL Lausanne), *"Novel Building Blocks for Supramolecular Chemistry"*
- Aldrik Velders, (University of Twente), *"Self Assembly from Solution to Surface"*
- Javier de Mendoza, (ICIQ), *"Recognition and Catalysis Based on Self-Assembly and Metal-Coordination"*
- Feliu Maseras, (ICIQ), *"A Computational Approach to Host Guest Catalysis"*
- *Anton Vidal* (ICIQ) *"Self-Assembled Chiral Biaryl Derivatives via Hydrogen*

Bonding and Metal-Ligand Interactions"

- Arjan Kleij, (ICIQ), "*Self-Assembling Heteromultimetallic Metallosalens: Towards New Bimetallic Catalysis"*
- Pablo Ballester, (ICIQ), *"Calix[4]pyrroles with Extended Aromatic Cavities. New tools for the quantification of the anion-pi interaction."*

9. **June 2007** KTH, Royal Institute of Technology, Stockholm, Sweden

Workshop on Dynamic Combinatorial Chemistry: An Introduction.

- Sijbren Otto, (University of Cambridge), "*Approaches to Dynamic Library Analysis"*
- Jeremy Sanders, (University of Cambridge), *"Dynamic Combinatorial Chemistry, An Introduction"*
- Ulrich Lüning, (Christian-Albrechts-Universität, Kiel), *"Dynamic Combinatorial Chemistry, Chemistry for Receptor Discovery Principles of Macrocycle Synthesis"*
- Ivan Huc, (IECB), *"Dynamic Equilibria and folding processes"*
- Fredrik Rahm, (AstraZeneca), *"Combinatorial Chemistry in the Drug Discovery Process"*
- Sabine Flitsch, (University of Manchester), *"The Application of Biocatalysis to Dynamic Combinatorial Chemistry"*
- Miles Congreve, (Astex), *"Industrial Partner Presentation on DCC"*

Universite Michel de Montaigne Bordeaux 3 DEFLE

April 2007–May 2008 French Language Course (120 hours)

IECB, Bordeaux

May 2007 Biacore Workshop, *"Surface Plasmon Resonance"*.

September 2007 Workshop Cancéropôle Grand Sud-Ouest, *"Approaches combinatoires pour la validation de cibles et de molécules anti-tumorales"*

February 2008 X-ray crystallography Workshop (3 hours).
Brice Kauffmann (IECB) *1. "X-ray crystallography, in direct space"* 2. *"X-ray crystallography in terms of X-ray diffraction, NMR – Nuclear spin transitions, electron diffraction"*

May 2008 Biacore Workshop, *"Surface Plasmon Resonance"*.

Attended Lectures

Lectures were attended on a weekly basis from January 2007-July 2010 at IECB from international scientific researchers including:

- Prof.Jonathan Sessler, (University of Texas, Austin), *"Biomedical Applications of Expanded Porphyrins"*
- Dr. Gilles Guichard, (Institut de Biologie Moleculaire et Cellulaire, Strasbourg; now at IECB, Bordeaux), *"Helical foldamers and cyclic oligomers as templates for the design of functional peptide / protein mimetics"*
- Prof. E. W. Meijer, (Eindhoven University), *"Dendrimers and supramolecular architectures for biomedical applications." "Chiral amplification in supramolecular architectures". "Supramolecular polymers at work"*
- Prof. Eiji Yashima, (Nagoya University), "*Double-Stranded Helical Polymers and Oligomers: Synthesis, Structures and Function"*
- Prof. Shankar Balasubramanian, (University of Cambridge), *"Do G-quadruplex nucleic acids play a role?"*
- Dr. Maité Paternostre, (CEA Saclay), *"Peptide self-assembly: mechanism and organization"*
- Prof. Klaus Muellen, (MPI Mainz, Germany), *"The Fascination and Function of Colour: from Cell Staining to Quantum Cryptography", "Organic Functional Nanoparticles", "Graphenes and Carbon Materials"*
- Prof. J Clayden, (University of Manchester), *"Conformational Stereocontrol"*
- Dr David Monchaud, (Centre Universitaire Paris XI Orsay), *"Stabilizing G-quadruplex DNA by small molecules : A novel avenue for cancer therapy"*
- Julius Rebek, Jr, (The Scripps Research Institute), "*The Inner space of molecules"*
- Debora Berti, (University of Florence), *"Hybrid Lipid/DNA Self- Assemblies"*
- Prof. Makoto Fujita, (The University of Tokyo), *"Bio-encapsulation by Self-assembled Hosts"*

Chimie dynamique: reconnaissance de nucléobases par des récepteurs synthétiques et isomérie cis-trans d'hydrazones acylées

Ce travail traite du développement des systèmes moléculaires qui peuvent s'adapter à l'addition de substances qui agissent comme un gabarit. Cette approche permet d'isoler une espèce majeure à partir d'un mélange de composés par le biais de la chimie combinatoire dynamique (CCD). La première partie de ma thèse de doctorat inclus l'utilisation d'un ADN simple brin (ADNsb) comme un gabarit pour le transfert d'information par auto-assemblage de récepteurs sans avoir besoin d'enzyme. De nouveaux récepteurs de l'adénine et de la guanine (pinces A et G) solubles dans l'eau ont été conçues dans ce but. Une approche utilisant la résonance magnétique nucléaire (RMN) a été utilisée pour déterminer l'affinité de liaison comme preuve d'une reconnaissance spécifique et efficace. Une évaluation dans l'eau par dichroïsme circulaire (CD) et mesure de la température de fusion par UV (T_m) a été réalisée. Cela a permis de tester respectivement la capacité d'auto-assemblage entre les pinces et un modèle ADNsb, et la force du processus de coopérativité. La deuxième partie de ce travail est axée sur le tri spontanné de motifs pyridine acylhydrazone et sur les configurations intéressantes qu'ils adoptent. Nous avons étudié la synthèse d'une série de motifs pyridine acylhydrazone: dimère, trimères et pentamères. Des études RMN ont permis d'évaluer les changements dans l'équilibre configurationnel cis / trans de ces systèmes dynamiques. Les études ont montré que l'équilibre attendu est biaise la cis acylhydrazone pyridine isomère a été observée par diffraction des rayons X. ***Mots clés:*** chimie dynamique, CCD, l'auto-assemblage, les templates, le transfert d'informations, récepteurs synthétiques, solubles dans l'eau, la reconnaissance nucléobase, acide nucléique, ADN, simple brin, acylhydrazone pyridine, dimère, trimère, pentamère, isomérie cis-trans, RMN, CD UV, des études de température de fusion.

Dynamic chemistry: nucleobase recognition by synthetic receptors and cis-trans acylhydrazone isomerism

This work deals with the development of molecular systems which can adapt upon the addition of substances that act as templates. This approach enables one major species to be identified from a mixture of compounds through the use of dynamic combinatorial chemistry (DCC). The first part of my PhD included the use of a single stranded DNA (ssDNA) as a template for information transfer via the self-assembly of receptors without the need for enzymes. New water soluble adenine and guanine receptors (A and G clamps) were designed and synthesised for this purpose. Nuclear magnetic resonance (NMR) titration studies were carried out to calculate the binding affinity and as a proof of specific and efficient recognition. An assessment in water via circular dichroism (CD) and UV temperature melting (T_m) studies was carried out. This tested the ability for self-assembly between the clamps and a ssDNA template and the strength of the cooperative process respectively. The second part of my PhD focused on the self-sorting of acylhydrazone pyridine motifs and the interesting configurations they adopt. The feasibility to synthesise these acylhydrazone pyridine motifs (dimer, trimers and pentamers) was investigated. X-ray and NMR studies showed that the equilibrium was found to be biased in an unusual way, and the cis acylhydrazone pyridine isomer was observed.

Key words: Dynamic chemistry, DCC, self-assembly, templating, information transfer, synthetic receptors, water soluble, nucleobase recognition, nucleic acid, DNA, ssDNA, acylhydrazone pyridine, dimer, trimer, pentamer, cis-trans isomerism, NMR, CD, UV melting temperature studies.