



Synthetic H-bonding information oligomers

By

Alexander Stross

Department of Chemistry

University of Sheffield

Submitted in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

March 2016

Summary

In Nature, the nucleic acids contain chemical information in the form of a sequence of bases. This information content is expressed through sequence selective duplex formation and template directed synthesis. To date, the only programmable artificial information molecules that can truly rival the nucleic acids, in terms of their function, are structurally very similar to the nucleic acids. This thesis describes a synthetic approach to duplex forming H-bonding oligomers that contain information in the form of a sequence of H-bond donor and acceptor groups, in the pursuit of a programmable material that is orthogonal to the nucleic acids.

Chapter 1 is a literature review of natural and synthetic information molecules and their applications in nanotechnology, including an overview of the structurally reengineered versions of the nucleic acids, foldamers, template polymerisation and synthetic duplexes. The review highlights the absence of totally synthetic information oligomers, that are orthogonal to the nucleic acids, and Chapter 2 sets out the aims of this thesis, which is to address this gap.

Chapter 3 describes the synthesis of oligomers equipped with phenol H-bond donors and phosphine oxide H-bond acceptors. Through cooperative H-bond formation, these oligomers form double stranded complexes, which were characterised by NMR titrations and thermal denaturation experiments. For each additional H-bond there is an order of magnitude increase in association constant.

In Chapter 4 it is demonstrated that the modular design for the oligomers in Chapter 3 represents a general strategy to synthetic information oligomers. Two new classes of H-bond acceptor oligomer were synthesised, bearing pyridine and pyridine

N-oxide groups. Both these systems also exhibit cooperative duplex formation with H-bond donor oligomers, which were characterised by NMR titration.

Chapter 5 examines the ability of mixed sequence 3-mers formed of H-bond donors (phenol) and H-bond acceptors (pyridine *N*-oxide) to form duplexes in a sequence selective manner. All 8 combinations of donor and acceptor were synthesised, and NMR titrations were used to measure the association constants for each pairwise combination of oligomers. Sequence matched duplexes generally have the highest association constants, but there are some anomalies.

Declaration

The work presented in this thesis was carried out at the University of Sheffield from October 2011 to September 2014 and at the University of Cambridge from September 2014 to March 2015. The work was carried out by me, except where otherwise stated.



Alexander Stross

Contents

| | | |
|----------------------|---|-----------|
| Summary | i | |
| Declaration | iii | |
| Abbreviations | vii | |
| | | |
| Chapter 1 | Natural and synthetic information molecules | 1 |
| 1.1 | Nucleic acids..... | 2 |
| 1.1.1 | Re-engineering of the nucleic acids..... | 4 |
| 1.1.2 | Designed assembly..... | 11 |
| 1.1.3 | Evolvable functional materials (aptamers) | 14 |
| 1.1.4 | Nucleic acids as templates | 17 |
| 1.2 | Synthetic information oligomers..... | 25 |
| 1.2.1 | Polymers with recognition groups (template polymerisation) | 25 |
| 1.2.2 | Sequence controlled polymerisation..... | 27 |
| 1.2.3 | Foldamers..... | 29 |
| 1.2.4 | Synthetic duplexes without sequence information..... | 35 |
| 1.2.5 | Synthetic duplexes with sequence information | 49 |
| 1.3 | Conclusions | 60 |
| 1.4 | References..... | 62 |
| | | |
| Chapter 2 | Aims | 73 |
| | | |
| Chapter 3 | Single sequence H-bonding information oligomers | 77 |

| | | |
|--------------------|---------------------------------------|------------|
| 4.5 | Experimental | 215 |
| 4.5.1 | Synthesis | 215 |
| 4.5.2 | Binding studies | 253 |
| 4.6 | Acknowledgements..... | 254 |
| 4.7 | References | 255 |
| Chapter 5 | Mixed sequence trimer duplexes | 257 |
| 5.1 | Introduction..... | 258 |
| 5.2 | Approach..... | 260 |
| 5.3 | Results and discussion..... | 263 |
| 5.3.1 | Synthesis | 263 |
| 5.3.2 | Binding..... | 265 |
| 5.4 | Conclusions | 273 |
| 5.5 | Experimental | 275 |
| 5.5.1 | Synthesis | 275 |
| 5.5.2 | Binding studies | 306 |
| 5.6 | Acknowledgements..... | 307 |
| 5.7 | References | 308 |
| Appendix I | Fitting analysis | 311 |
| I.1 | Fitting from Chapter 3..... | 312 |
| I.2 | Fitting from Chapter 4..... | 317 |
| Appendix II | Thermal denaturing data | 327 |

Abbreviations

| | |
|-------------|---|
| 3D | Three dimensional |
| A | Adenine |
| AIBN | Azobisisobutyronitrile |
| C | Cytosine |
| d | Doublet |
| DABCO | 1,4-Diazabicyclo[2.2.2]octane |
| DCE | 1,2-Dichloroethane |
| DCM | Dichloromethane |
| dd | Doublet of doublets |
| DMF | Dimethylformamide |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| EM | Effective molarity |
| EtOAc | Ethyl acetate |
| EtOH | Ethanol |
| FT-IR | Fourier transform infrared spectroscopy |
| G | Guanine |
| GNA | Glycol Nucleic Acid |
| H-bond(ing) | Hydrogen bond(ing) |
| HRMS | High Resolution Mass Spectrometry |
| <i>J</i> | J-coupling constant |

| | |
|---------------|---|
| LNA | Locked nucleic acid |
| m | Multiplet |
| MALDI | Matrix assisted laser desorption/ionisation |
| <i>m</i> CPBA | 3-Chloroperoxybenzoic acid |
| MS | Mass spectrometry |
| NMR | Nuclear magnetic resonance |
| Pd/C | Palladium on activated carbon |
| PCR | Polymerase chain reaction |
| Ph | Phenyl |
| PNA | Peptide nucleic acid |
| RAFT | Reversible addition fragmentation radical transfer |
| RNA | Ribonucleic acid |
| ROMP | Ring opening metathesis polymerisation |
| s | Singlet |
| SELEX | Systematic evolution of ligands by exponential enrichment |
| T | Thymine |
| TBAF | Tetra- <i>n</i> -butylammonium fluoride |
| TFA | Trifluoroacetic acid |
| THF | Tetrahydrofuran |
| TNA | Threose nucleic acid |

Chapter 1

Natural and synthetic information molecules

1.1 Nucleic acids

The remarkable properties of nucleic acids have not ceased to amaze chemists since the discovery of the structure of the double helix in 1953.¹ The different levels of structure of deoxyribonucleic acid (DNA) are shown in Figure 1.1. DNA is comprised of a defined sequence of nucleobases attached to deoxyribose sugars, which are joined into a polymer using phosphodiester chemistry (Figure 1.1(a)). There are four different nucleobases each with a different arrangement of H-bond donor and acceptor functional groups. Each base can form H-bonds with every other base and with itself, but the energetically favoured interactions are when adenine (A) binds with thymine (T) and guanine (G) binds with cytosine (C) (Figure 1.1(b)).²⁻⁶ A duplex formed of two strands of DNA adopts a double helix structure (Figure 1.1(c)). The main driving force for duplex formation in DNA, is the π -stacking interaction between nucleobases,^{7,8} but H-bonding between nucleobases is responsible for the sequence specific interaction between different oligonucleotides.⁹ In Nature, DNA exclusively operates in aqueous media and the presence of the negative charge on the phosphate groups is important for solubility and promotes an elongated structure.¹⁰ Watson and Crick proposed that Nature uses DNA to express and encode genetic information in two key ways: sequence selective duplex formation and template directed synthesis.¹¹ Sequence selective duplex formation means that two strands will selectively form a duplex, or one strand will selectively fold, such that complementary base pairs are formed (Figure 1.2(a)). Template directed synthesis means that one strand of DNA can act as a template for the synthesis of another strand of DNA with the complementary sequence of base pairs (Figure 1.2(b)).

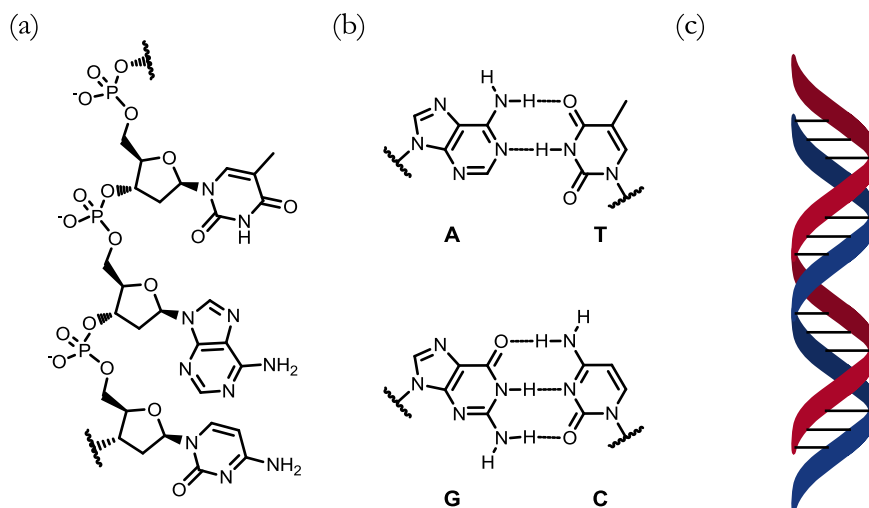


Figure 1.1. (a) Chemical structure of DNA; (b) Watson-Crick base pairing of adenine A with thymine T and guanine G with cytosine C; (c) DNA double helix.

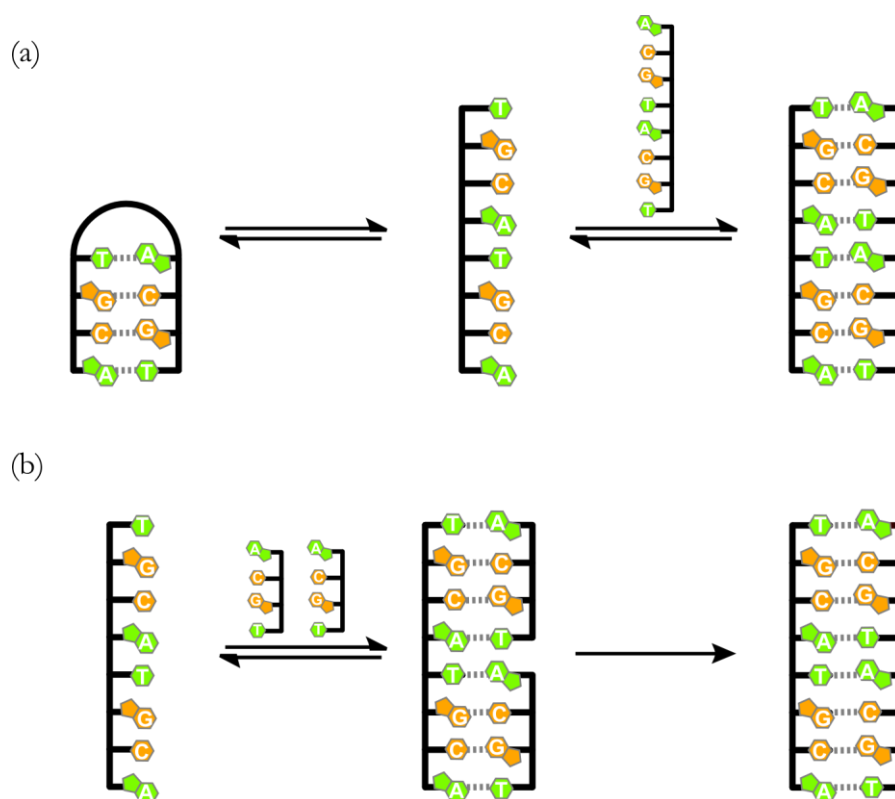


Figure 1.2. Schematic representation of (a) sequence selective duplex formation leading to duplex and folded structures; (b) template directed synthesis. Green represent A•T base pairs and orange represent C•G base pairs.

This Section begins by highlighting how robust the nucleic acids are to structural reengineering in terms of their ability to form sequence selective duplexes. Below we will discuss how the key properties displayed schematically in Figure 1.2

can be exploited, highlighting the many applications information oligomers can be used for.

1.1.1 Re-engineering of the nucleic acids

1.1.1.1 Modified bases

Numerous examples exist of artificially expanded genetic information systems where oligonucleotides are made, which include non-natural base pairs in addition to the natural G C and A T pairs (Figure 1.1(b)). Benner has contributed a great deal to this area reporting several artificial bases analogous to purines and pyrimidines, beginning with the successful enzymatic incorporation of isoG opposite isoC in a DNA template (Figure 1.3(a)).¹² Problems with hydrolysis of isoC and tautomers of isoG that made it pair with uracil in RNA, led to the design and successful incorporation of X into a template alongside its complement Kappa (Figure 1.3(b)).¹³ More recently the same group has reported the Z•P base pair. (Figure 1.3(c)),¹⁴ which has been used in an *in vitro* directed evolution experiment to generate six base alphabet aptamers (GCATZP) which bind liver cancer cells.¹⁵

The Kool group has successfully incorporated size-expanded A and T base pairs onto a native DNA backbone, where an extra benzene ring is incorporated into the natural nucleoside (Figure 1.3(d) and (e)).^{16,17} The configuration shown in Figure 1.3(d) led to base pairs which enhanced duplex stability, shown by DNA melting experiments. In a similar approach, the Inouye group have made a full set of size expanded nucleotides where an acetylene unit is inserted between the base and the native backbone (Figure 1.3(f)).¹⁸ Duplexes incorporating the four artificially expanded bases were found to have similar thermal stabilities compared to their native counterparts.

Minakawa and Matsuda *et al.* have incorporated two artificial four H-bond base pairs into native strands, which significantly increased the thermal stability of sequence matched duplexes (Figure 1.3(g)).¹⁹ The Kool group has demonstrated that H-bonding is not essential for duplex formation by incorporating non-polar isosteres of T and A into native DNA (F and D in Figure 1.3(h)).²⁰ They found from thermal stability experiments that the hydrophobic base pairs were more stable than a native polar base pairing with a non-polar analogue.

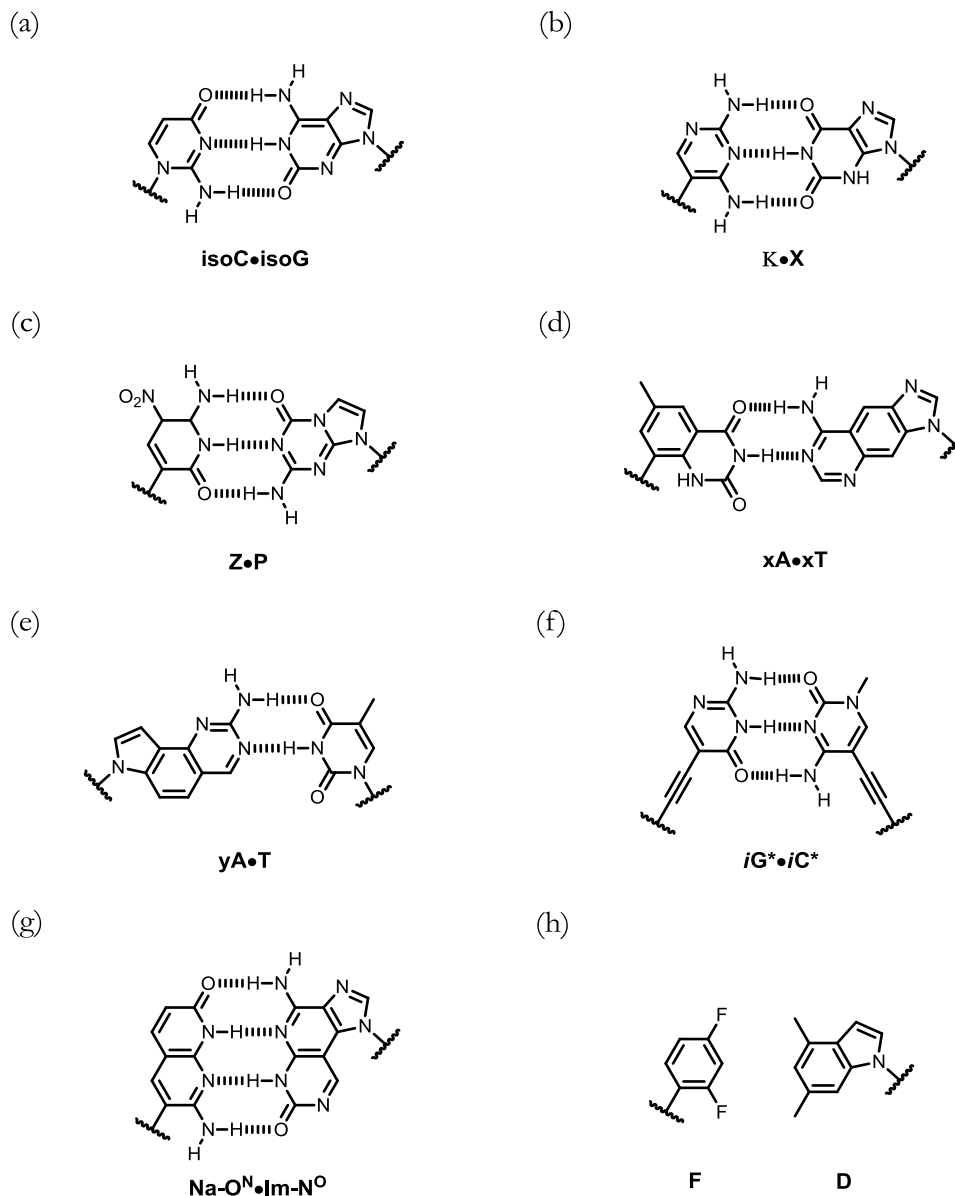


Figure 1.3. Artificial base pairs: (a) isoC•isoG; (b) X•Kappa; (c) Z•P; (d) xA and xT; (e) yA; (f) *iG**•*iC**; (g) NaO^N•ImN^O; (h) F and D.

1.1.1.2 Sugars

The furanose sugar in DNA has been replaced with a variety of alternative spacers, many of which lead to artificial oligonucleotides that can form stable duplexes (Figure 1.4). Some modifications have been simple, for instance just adding substituents onto the 2' OH of RNA,²¹ such as the 2-methoxyethyl-RNA shown in Figure 1.4(b) which, when fully incorporated into oligonucleotides, showed higher

thermal stability when binding to RNA, than RNA binding with itself.²² Other pentoses have been explored such as pyranosyl-RNA (Figure 1.4(c)).²³ Analogues not based on pentoses have also been made, such as Eschenmoser's four-carbon threose nucleic acid (TNA in Figure 1.4(d)),²⁴ and Aerschot's six-carbon 1,5-anhydrohexitol (Figure 1.4(e)).²⁵ Leumann and co-workers have replaced the sugar with bicyclic (Figure 1.4(f)),^{26,27} and tricyclic sugar rings (Figure 1.4(g)).²⁸ A prevalent example of bicyclic nucleic acids is "Locked" nucleic acids (LNA), where a 2'-O 4'-C methylene bridge locks RNA into the 3-C' *endo* conformation which is favourable for duplex formation and accounts for the increased thermal stability of LNA strands hybridising with DNA and RNA. (Figure 1.4(h)).^{29,30} Even completely acyclic sugar analogues have been made, such as glycol nucleic acid (GNA), which has a propylene glycol phosphodiester backbone (Figure 1.4(i)).^{31,32} These artificial oligonucleotides were made using enantiomerically pure (*S*)- or (*R*)-glycidol and interestingly (*R*)-GNA-(*R*)-GNA and (*S*)-GNA-(*S*)-GNA duplexes show similar thermal stability to DNA and RNA duplexes but (*R*)-GNA and (*S*)-GNA are incompatible with each other.

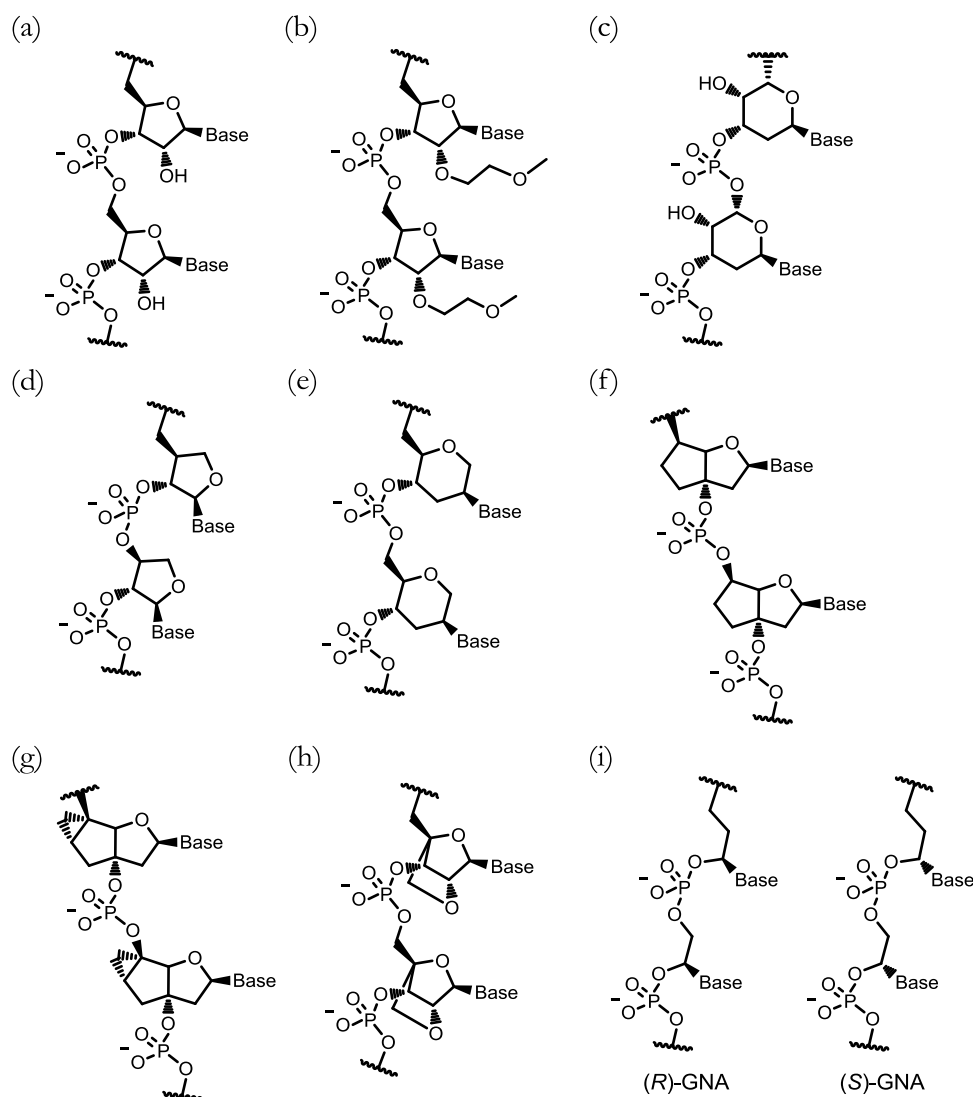


Figure 1.4. (a) native RNA; artificial nucleic acids where the furanose ring has been modified for: (b) furanose substituted on the 2'-O atom; (c) pyranose; (d) therosyl; (e) 1,5-anhydrohexitol; (f) a bicyclic ring; (g) a tricyclic ring; (h) “locked” furanose; (i) (*S*)- and (*R*)-glycidol.

1.1.1.3 Phosphates

There are many reported modifications of the phosphodiester chemistry in nucleic acids.³³ The phosphodiester has been replaced with phosphoramidate chemistry (Figure 1.5(a)),³⁴ and it was found that strands with only phosphoramidate linkages formed stable duplexes with native RNA and DNA. Uncharged linkages such as amide (Figure 1.5(b)),³⁵ 3'-thioformacetal (Figure 1.5(c)),³⁶ sulfone (Figure 1.5(d))^{37,38} and triazole (Figure 1.5(e))³⁹ have been reported. It has been suggested by

Benner that the charged backbone is essential for nucleic acid duplex formation.⁴⁰ For instance, oligonucleotides fully substituted with the sulfone linker shown in Figure 1.5(d) formed an intramolecular fold rather than forming duplexes. However, there are artificial nucleic acids with uncharged backbones that are capable of duplex formation. For instance the oligonucleotide 10-mer analogue with triazole linkers shown in Figure 1.5(e) formed stable duplexes with DNA. The most prevalent example of an information molecule, which is analogous to DNA and has an uncharged backbone is PNA, which will be discussed in the next Section.

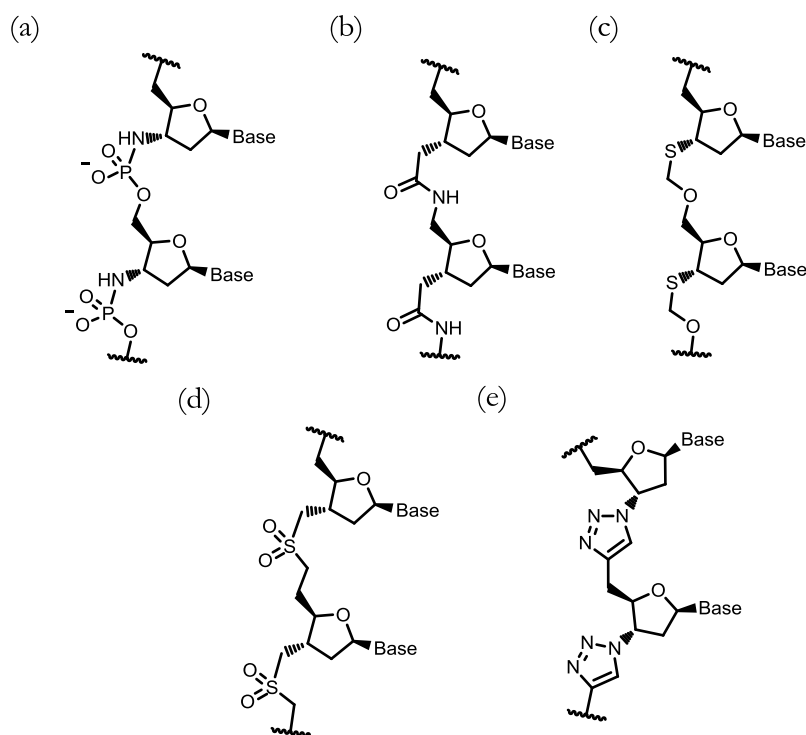


Figure 1.5. Artificial nucleic acids where the phosphodiester coupling has been replaced for other chemistries: (a) phosphoramidate; (b) amide; (c) 3'-thioformacetal; (d) sulfone; (e) triazole.

1.1.1.4 PNAs

Peptide or Polypeptide Nucleic Acids (PNAs) are arguably the most prevalent artificial nucleic acids. The sugar and phosphate backbone is entirely changed for uncharged pseudopeptide chains, but nucleobases still form the recognition motif

(Figure 1.6). The Nielsen group opened up this field when they reported a PNA based on a *N*-2-aminoethylglycine backbone (Figure 1.6(a)),⁴¹ which forms duplexes with strands of PNA, DNA and RNA.⁴²⁻⁴⁴ The crystal structure of a PNA-PNA duplex has shown that it adopts a similar base-pair stacking overlap motif as seen in DNA.^{45,46} As in DNA, stacking forces are responsible for PNA-PNA duplex formation in water, yet PNA duplexes show a greater stability than DNA in water-DMF mixes.^{47,48} The PNA architecture is quite robust in that modifications to the backbone, for instance at the α ,⁴⁹ and γ positions,⁵⁰ and exotic amino acids such as pyrrolidinyl PNA (Figure 1.6(b)) also lead to stable duplexes.⁵¹ A recent example of a PNA used dynamic covalent thioester linkages to join the nucleobases to the backbone (Figure 1.6(c)), so that the sequence of nucleobases could adapt to different DNA and RNA templates.⁵²

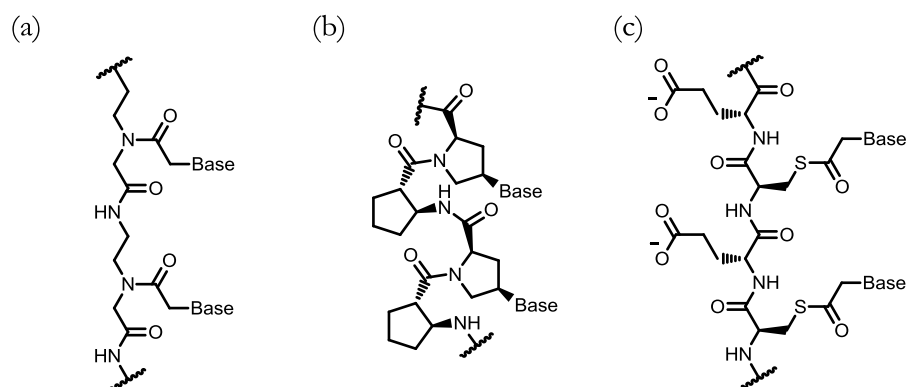


Figure 1.6. PNA chemical structures: (a) *N*-2-aminoethylglycine PNA (b) pyrrolidinyl PNA (c) thioester PNA.

1.1.2 Designed assembly

Nucleic acids are programmable in that a strand will fold or form a duplex in a predictable way, depending on the sequence of nucleobases. This feature combined with advances in nucleic acid synthesis, make DNA an attractive material for *designing* and constructing nanoscale structures.⁵³ In 1982, the basis for DNA nanotechnology was outlined by Seeman who described requirements to obtain sequences of oligonucleotides that form immobile junctions instead of linear duplexes, for instance the structure shown in (Figure 1.7(a)).⁵⁴ A junction occurs where multiple double stranded sections of DNA meet at a branching point and the junction is immobile when, as a consequence of having unique base pairing patterns, it cannot be formed at other sites along the DNA strands. A common assembly approach is called DNA origami,⁵⁵ which involves a long single strand of DNA interacting with several short staple strands, inducing it to fold into a defined shape (Figure 1.7(b)). Computer aided design now allows one to predict the sequence of strands that would lead to the desired structure.^{56,57} Static tiles and objects have been made using DNA,^{55,58-62} as well as nanodevices,⁶³⁻⁶⁷ ion channels⁶⁸ and DNA computers.⁶⁹⁻⁷¹ This approach has been used by the Church group to construct a DNA origami barrel that responds to external stimuli to expose a payload of antibody fragments (Figure 1.8).⁷² The barrel is made from a 7308-base scaffold strand that is held together using 196 oligonucleotide staple strands and formed in a one-pot fashion via rapid heating and then slow cooling. The barrel has two domains that are covalently joined at one end through the staple strand but non-covalently attached at the other through a hybridised double strand of DNA, which acts as a “lock”. One of the lock strands belongs to one half of the barrel and the other strand belongs to the other side of barrel. The two sets of hybridised strands fastening the barrel shut separate in

response to target proteins and the barrel opens and exposes the covalently attached antibody fragments, delivering them to their target.

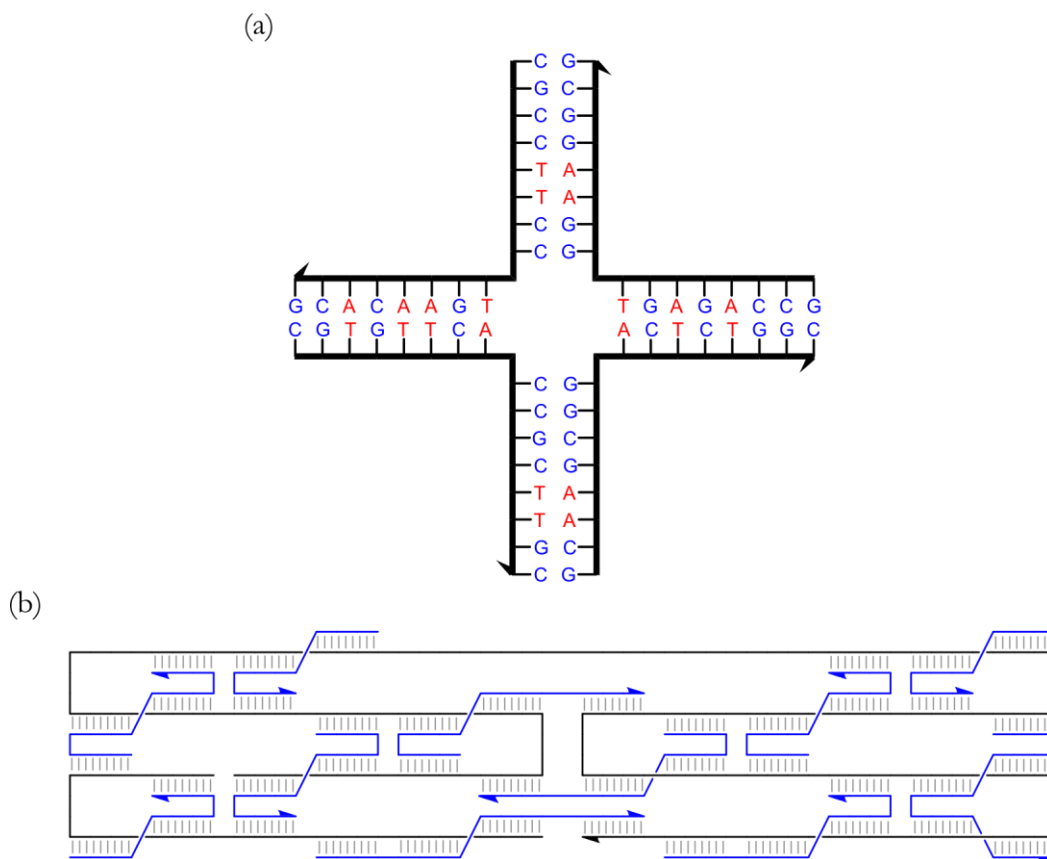


Figure 1.7. (a) A four-arm immobile (Holliday) junction where the sequence of base pairs are chosen to reduce the stability of competing structures; (b) DNA origami where a long strand (black) is induced to fold into a designed shape through interactions with short staple strands (blue).

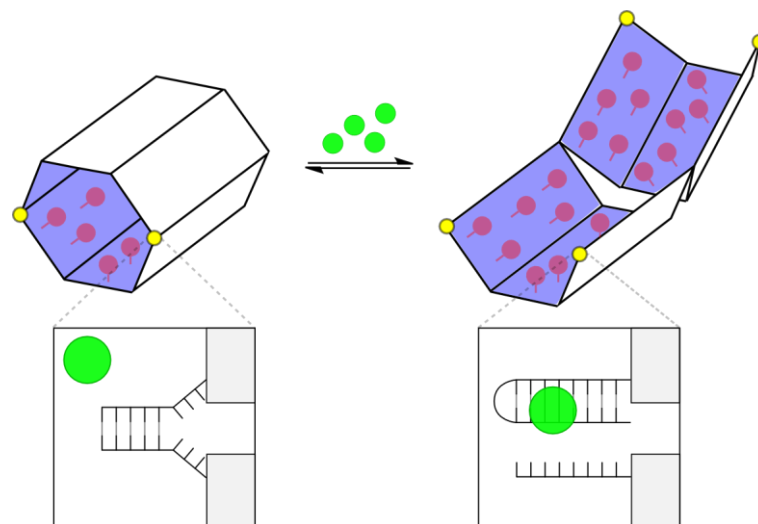


Figure 1.8. A molecular payload transporter based on a DNA origami barrel, which opens upon interaction of “lock” strands (yellow) with a molecular target (green). The barrel in its open form exposes covalently attached therapeutic sites (red).

Another elegant approach to construct 3D structures of DNA was reported by the Yin group, which they describe as DNA bricks.⁶⁰ Oligonucleotide strands, 32 base pairs in length, with 4×8 nucleotide binding domains hybridise selectively with four other strands owing to their unique sequence (Figure 1.9(a) and (b)). One thousand of these strands self-assemble to create a cube in a one pot fashion (Figure 1.9(c)). By selecting subsets of bricks, shapes with cavities and tunnels were made (Figure 1.9(d)) and characterised using transmission electron microscopy imaging.

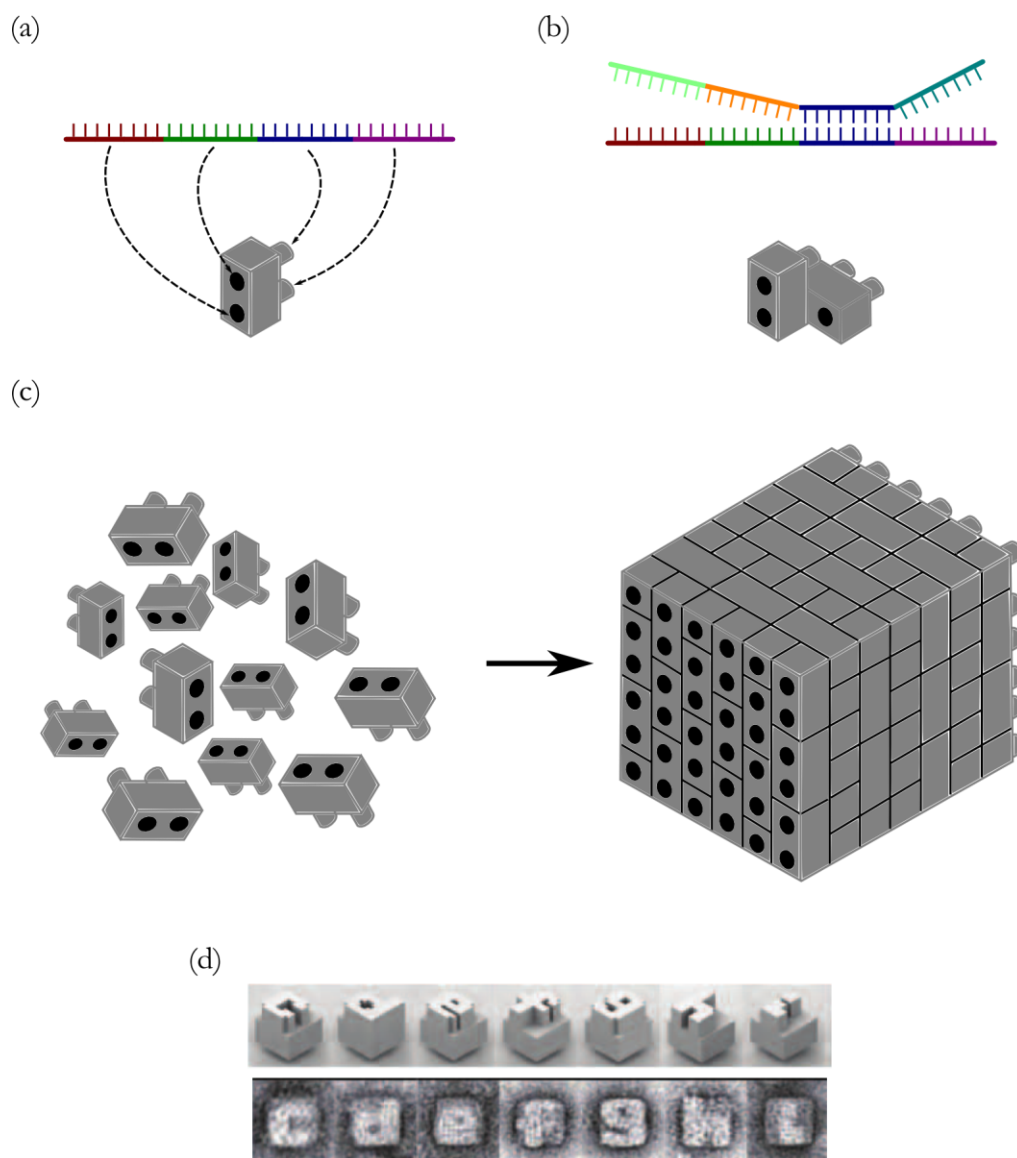


Figure 1.9. DNA brick (a) a brick consisting of a 32-mer DNA strand with a unique sequence comprised of four binding domains; (b) Each strand binds with up to 4 other strands in a sequence selective fashion; (c) the DNA bricks selectively self-assemble to give a 3D cube where each brick has a position in the cube determined by its sequence; (d) 3D structures with programmed shapes are designed using computer software (top) and imaged using transmission electron microscopy (bottom). The images in part (d) are taken directly from reference 60.

1.1.3 Evolvable functional materials (aptamers)

Functional strands of nucleic acids can be obtained by exploiting the ability of DNA and RNA to self-replicate and evolve through several rounds of template directed synthesis. When nucleic acids fold into a variety of sequence dependent

tertiary structures, such as G-quadruplexes or stem loops (Figure 1.10), the resulting structures can have functionality.^{73,74} Oligonucleotides with tertiary structures that perform functions are called aptamers, which is the term used by Ellington and Szostak in 1990 when they isolated sequences of RNA that selectively bound organic dye molecules.⁷⁵ Aptamers are isolated using a technique known as systematic evolution of ligands by exponential enrichment (SELEX). The process begins with selection of a library of $10^{13} - 10^{18}$ DNA strands with unique primary structures. These strands are either transcribed for RNA selection or strand separated for DNA selection, and then an assay takes place to select sequences which display the desired function. These sequences are isolated and a polymerase chain reaction (PCR) is used to create a new aptamer library. After typically 8 to 15 cycles of selection and PCR, a library of functional oligonucleotides is isolated.⁷⁶

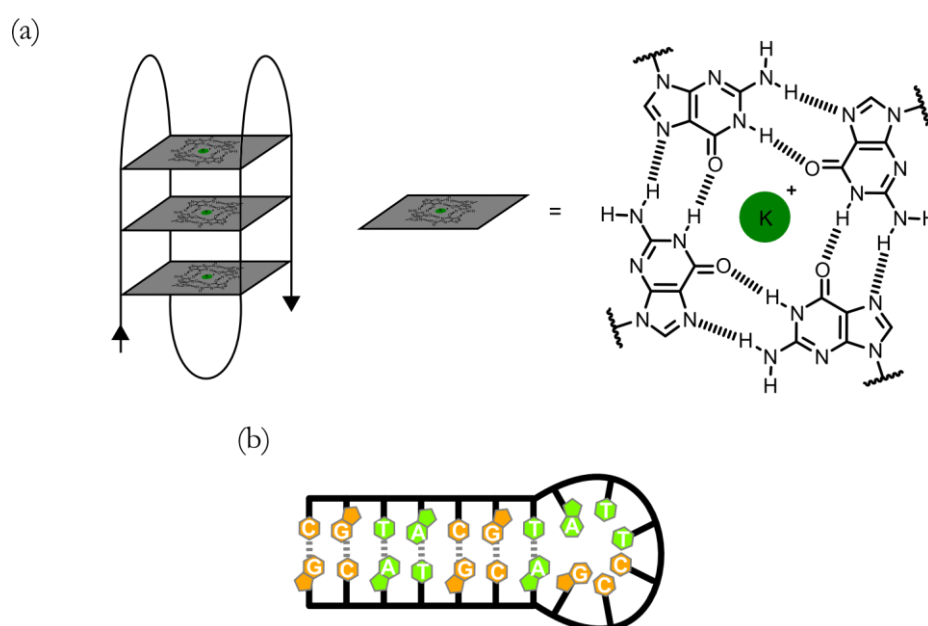


Figure 1.10. Typical structures of DNA that can perform functions (a) G-quadruplex; (b) stem-loop.

Potential functions of aptamers includes receptors for selective binding of target molecules,⁷⁷⁻⁷⁹ and catalysis.⁸⁰⁻⁸² For instance Worgall and Stojanovic have

reported an approach to develop aptamers that selectively bind low epitope molecules such as sugars (Figure 1.11).⁷⁹ The oligonucleotide library is attached to the stationary phase of a column via a biotinylated complementary oligonucleotide. The target sugar unselectively forms a ternary complex with a receptor molecule and the sugar-receptor mixture is eluted through the column and oligonucleotide sequences which bind the complex are eluted preferentially by stabilisation of a stem loop. The eluted oligonucleotide sequences are eluted and a new library generated by a PCR step and the process repeated (Figure 1.11(a)). The derived aptamers are transformed into sensors by functionalising the aptamer with a fluorescent label and the capture strand with a quencher. Upon complexation of the aptamer strand with the target complex the fluorescence label is no longer quenched and the signal increases (Figure 1.11(b)).

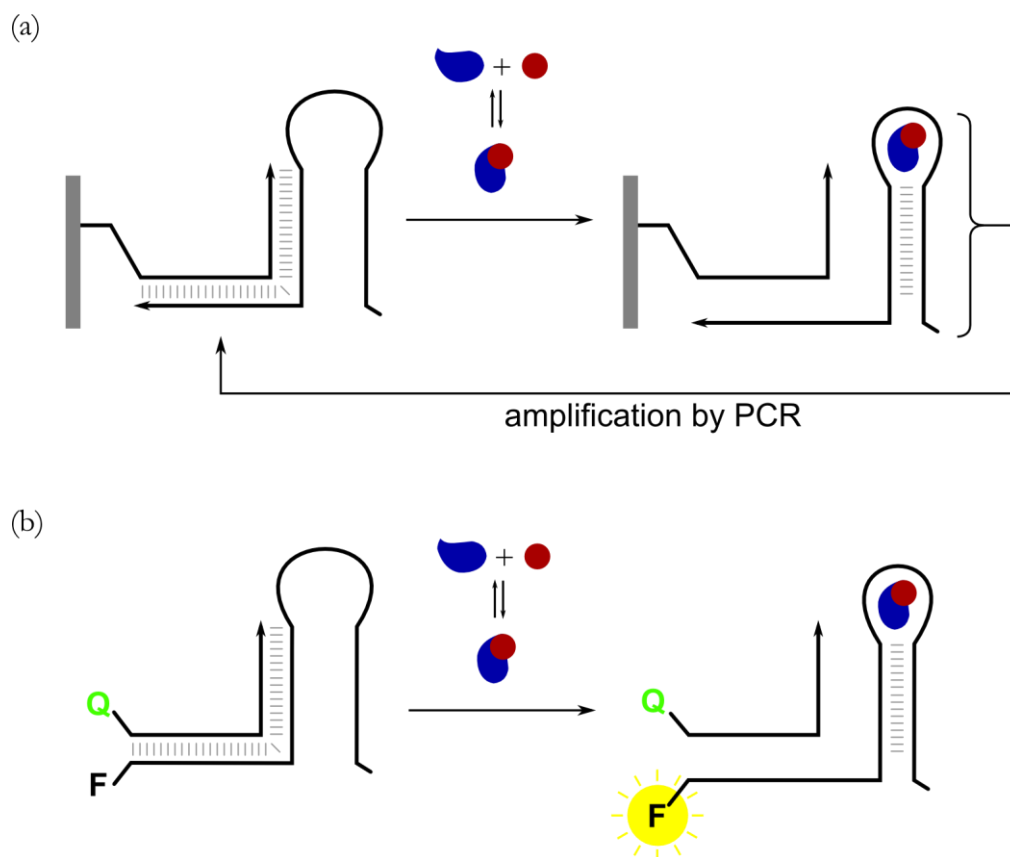


Figure 1.11. (a) An aptamer library is attached to the stationary phase of a column (grey box) through hybridisation with a covalently attached capture strand. Sugar-receptor target complex (blue and red) is eluted through the column which can form a ternary complex with the aptamer strands. Formation of the ternary complex induces stem loop formation in the aptamer strand and separates the complex from the capture strand so it can be washed out of the column. The eluted strands are amplified via PCR and the process repeated; (b) the derived aptamer strands are transformed into sensors when coupled with a fluorescent probe (F). Q on the capture strand quenches the fluorescence but upon aptamer binding the fluorescence is no longer quenched.

1.1.4 Nucleic acids as templates

In Nature, nucleic acids are exploited as sequence specific templates for enzymatic synthesis. Duplex DNA can be copied to form two new duplexes where each strand in the original duplex templates the formation of a daughter strand in the product duplexes (Figure 1.12(a)). DNA duplexes can template the formation of RNA strands through transcription (Figure 1.12(b)). Messenger RNA strands can template the formation of proteins in ribosomes by hybridisation of amino acid

bearing tRNA units to the messenger RNA, which allows the amino acids to react with the growing peptide chain (Figure 1.12(c)). Importantly, the sequence of amino acids depends on the primary structure of the RNA.

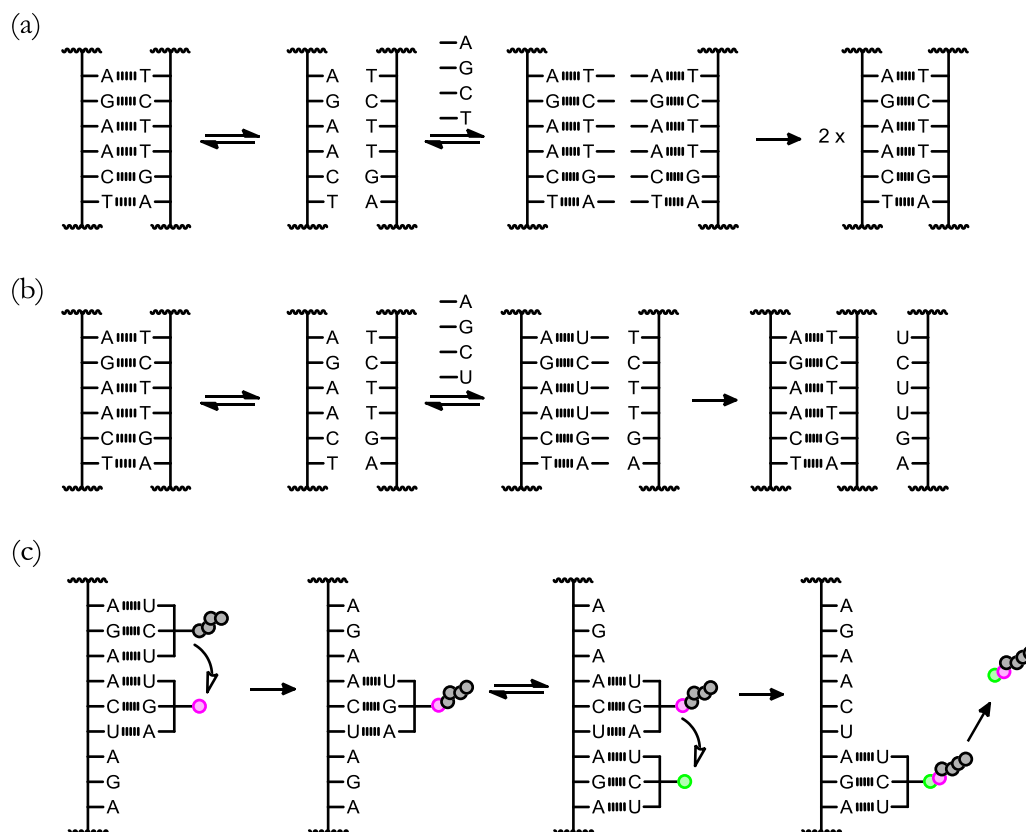


Figure 1.12. Schematic representation of DNA acting as a template where all steps are mediated by enzymes (a) DNA replication where both strands in a duplex can template the formation of their complementary daughter strand (b) DNA transcription where one strand in a duplex templates the formation of an RNA strand with a complementary sequence to the template strand (c) RNA translation where the sequence of amino acids in protein synthesis is determined by the sequence of bases in the template strand. A growing peptide (grey circles) attached to an RNA 3-mer sequences is hybridised with the template strand. RNA 3-mer fragments bearing amino acids bind to the template in a sequence specific manner and react with the growing peptide

Early examples of nucleic acid template directed synthesis typically involve DNA or RNA acting to mediate ligation reactions. In 1966 Naylor and Gilham used polyadenylic acid to template the formation of a native phosphodiester bond between two thymidine 6-mer strands (Figure 1.13(a)).⁸³ Since then many examples

of template ligation have been reported which use non-natural coupling chemistries and backbone structures. For instance Lynn has used DNA to template the ligation of modified DNA using reductive amination coupling chemistry (Figure 1.13(b)).⁸⁴ Eschenmoser has used TNA-TNA ligation reactions to demonstrate that replacing adenine with 2,6-diaminopurine, enhances the template effect (Figure 1.13(c)).⁸⁵ Nielson and Orgel have used an RNA 10-mer to ligate up to five PNA 2-mers (Figure 1.13(d)).⁸⁶

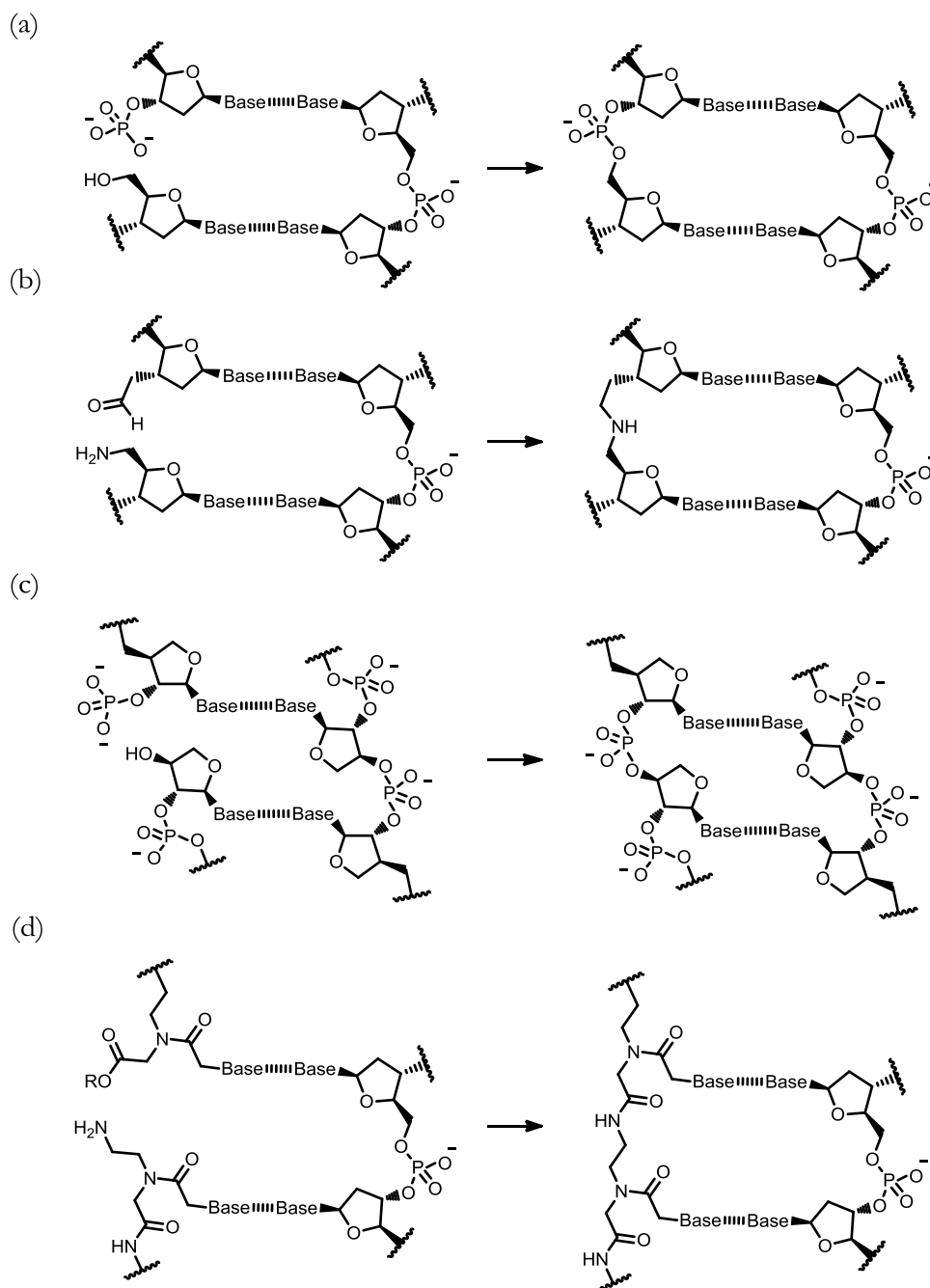


Figure 1.13. Nucleic acid ligation reactions (a) phosphodiester chemistry using native substrates and templates; (b) reductive amination chemistry using native template; (c) phosphodiester chemistry using TNA substrates and template; (d) RNA templating PNA ligation (amide chemistry).

Using nucleic acids and their close analogues as templates to mediate polymerisations from monomer units in the absence of enzymes has had some limited success.⁸⁷⁻⁹⁰ However there are some examples of polymers which use base

pairs as recognition groups to template polymerisation. For instance the Sleiman group have reported using a nucleobases containing copolymer made through a living ring-opening metathesis polymerisation (ROMP), to template adenine containing monomers in the formation of a conjugated polymer through Sonogashira coupling (Figure 1.14).⁹¹ The product polymers had chain lengths similar to the template and within a narrow molecular weight distribution.

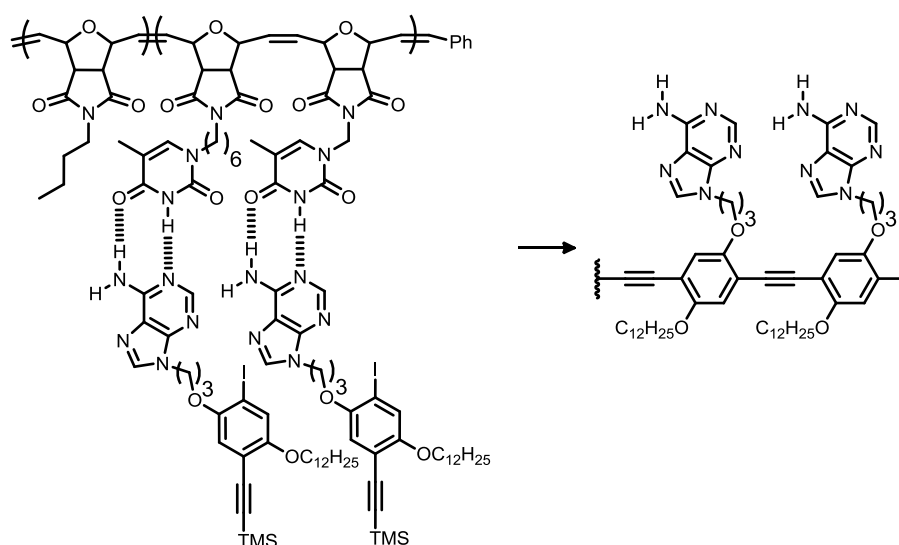


Figure 1.14. Template polymerisation: monomer units interact with a template through base pairing and couple together through Sonogashira chemistry.

The O'Reilly group have used a different approach to control polydispersity (Figure 1.15).⁹² They synthesised a thymine containing amphiphilic polyvinyl copolymer, which self-aggregates into micelles. When vinylbenzyl adenine monomer is added in CHCl₃, the monomer units induce the templates to dynamically enter and leave the micelles and base-pair with the thymine units. Upon addition of initiator (azobisisobutyronitrile, AIBN) dynamic template polymers are initiated outside the micelle before completing polymerisation within the micelle. This approach yields very narrow polydispersity products with molecular weights which do not vary as a function of reaction time.

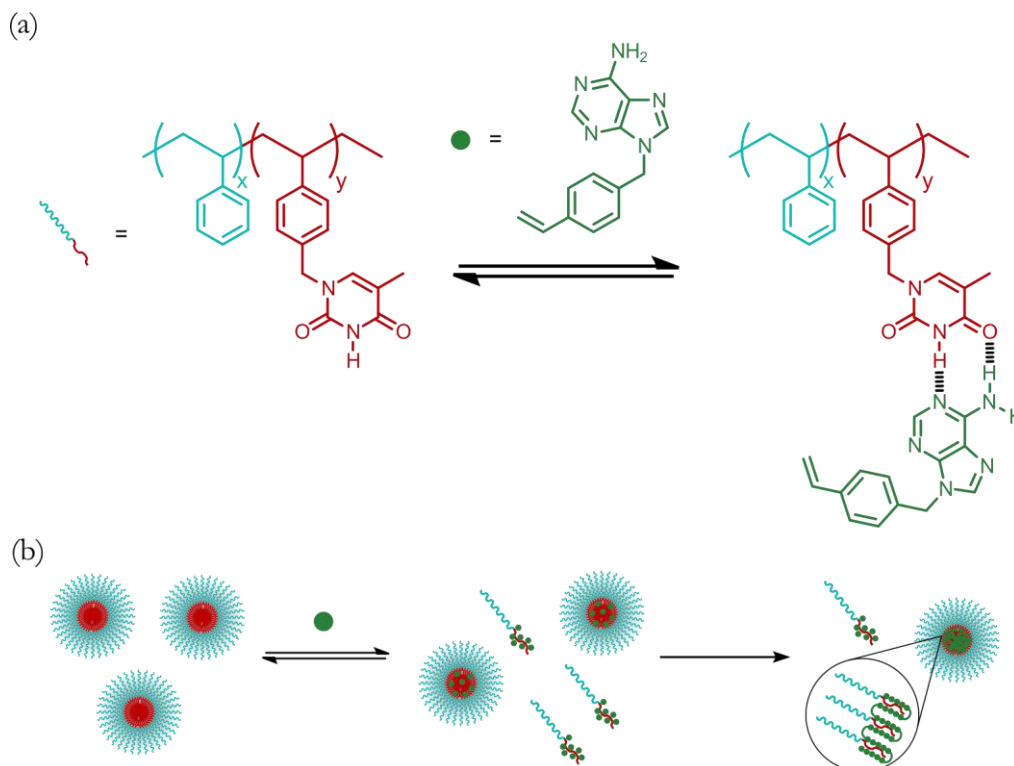


Figure 1.15. Template polymerisation: (a) vinylbenzyl adenine monomers base pair with amphiphilic thymine functionalised polyvinyl copolymers; (b) the copolymers self assemble into micelles and the monomer dynamically enters and leaves the micelle. Polymerisation initiation occurs outside the micelle and continues inside the micelle.

A separate line of research does not involve ligation or polymerisation but rather exploits the sequence selective duplex formation of nucleic acids to bring reactive groups together in more elaborate multi-step syntheses.⁹³⁻⁹⁵ Turberfield and O'Reilly have synthesised 4-mer oligomers via sequential Wittig reactions using DNA as a template.⁹⁶ In this approach a “toehold” displacement strategy is used, whereby a strand in a duplex, formed between two strands with only partially complementary sequences, can be displaced by a strand that has a fully complementary sequence (Figure 1.16(b)). The oligonucleotide strands were modified at one end with a phosphine ylide, an aldehyde group and an amino acid. In each synthetic cycle the first step is hybridisation whereby the ylide of one strand and the aldehyde on another are brought into close proximity for Wittig coupling (Figure 1.16(a)), which

transfers the growing oligomer from one strand to another. The strand that now has a phosphine oxide is removed by a “toehold” displacement and the strand bearing the oligomer is hybridised with a new strand with a reactive aldehyde group. A similar approach to this has been used by the Liu group to generate functional macrocycles coupled to DNA strands that are obtained using a protocol similar to SELEX.⁹⁷

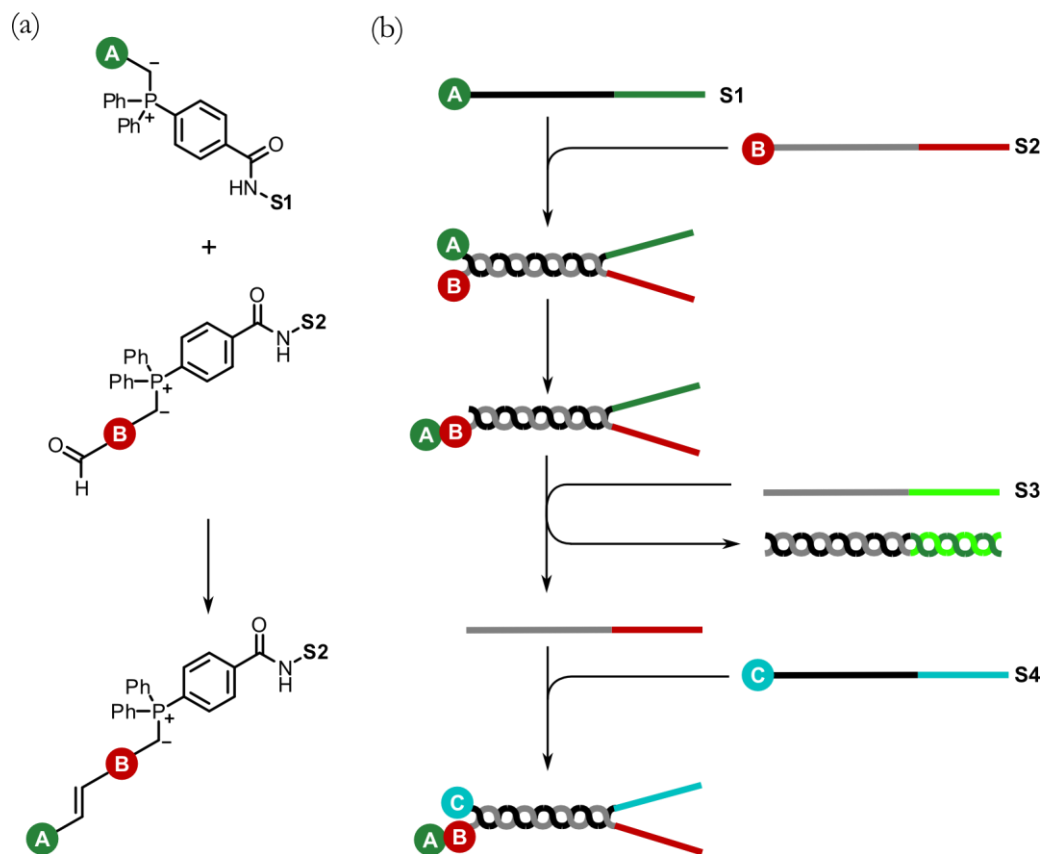
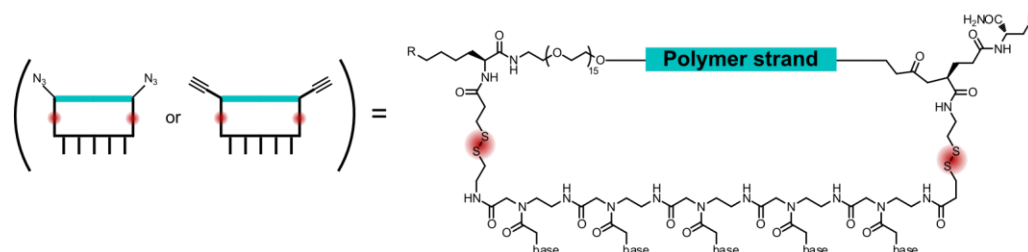


Figure 1.16. Schematic representation of DNA toehold displacement synthesis: DNA hybridisation between strands S1 and S2 in (a) promotes a Wittig reaction between the aldehyde of one strand and the phosphonium ylide on the other in (b). Toehold displacement of the spent ylide strand by a new aldehyde bearing strand templates another Wittig reaction. Repeated cycles leads to growth of a polymer chain.

In a different approach, Liu and co-workers translate the sequence information from a DNA template into the sequence of a non-nucleic acid polymer.⁹⁸ A DNA strand that folds into a step loop forms the template, where a “toehold” section is exposed. Sequences of PNA are joined to a polymer chain in a macrocyclic structure

through two cleavable disulphide linkers (Figure 1.17(a)). The PNA sequences base-pair with the template in a sequence specific manner, which aligns azide and alkyne groups at the end of the polymer chains. Through copper catalysed azide-alkyne coupling the polymer units covalently attach in a template directed synthesis. The product is a DNA strand with abiotic copolymer at one end (Figure 1.17(b)).

(a)



(b)

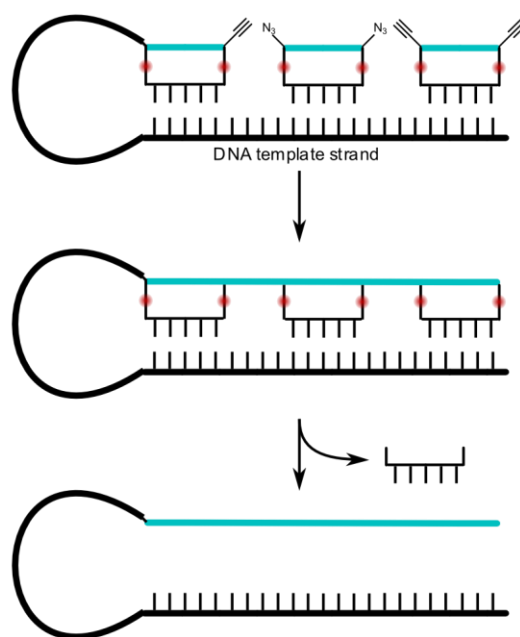


Figure 1.17. (a) PNA sequences coupled to a polymer strand via cleavable disulphide linkages. The polymer chain is terminated at either end with either azide or alkyne groups; (b) the PNA sequences hybridise with a template DNA stem loop. The azide and alkyne groups can react to join the polymer fragments together. Cleaving the disulphide linkers gives the template DNA stem loop covalently coupled to the polymer.

1.2 Synthetic information oligomers

1.2.1 Polymers with recognition groups (template polymerisation)

The concept of template polymerisation was postulated in 1954 by Szwarc who discussed the possibility that a polymer dissolved in a monomer might cause the monomer units to arrange into a pattern, ready for polymerisation.⁹⁹ Template polymerisations tend to affect the kinetics, degree of polymerisation and polydispersity of the product polymer. The monomer units can be organised through intermolecular interactions or covalent bonds. Most examples of template polymerisation mediated by intermolecular interactions involve radical reactions and some ring opening reactions.¹⁰⁰ A typical radical example is the polymerisation of acrylic acid in the presence of poly(*N*-vinylpyrrolidone) (Figure 1.18),¹⁰¹ where the template and product are held together by H-bonds.¹⁰² The polyacrylic acid products have a degree of polymerisation which is determined by the size of the template, and higher molecular weight templates lead to a larger rate enhancement.¹⁰³ Template polymerisations have also been combined with enzymes, for instance poly(ethylene glycol) acts as a template for the polymerisation of phenol by peroxidase.¹⁰⁴

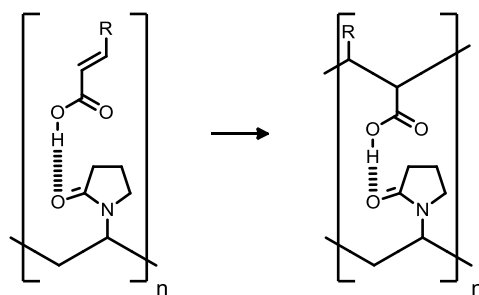


Figure 1.18. Template radical polymerisation of acrylic acid or methacrylic acid in the presence of a polyvinylpyrrolidone template, where R = H or CH₃.

The Luh group have used covalent bonds to template a ring-opening metathesis polymerisation (Figure 1.19).¹⁰⁵ Monomer units were coupled to a polynorbornene

derivative via ester linkers. The authors incorporated a ferrocene bridge to aid solubility and flexibility. Treatment with Grubbs I catalyst lead to a double stranded polymer and the ester groups were then hydrolysed to yield the daughter polymer which had a comparable degree of polymerisation to the template.

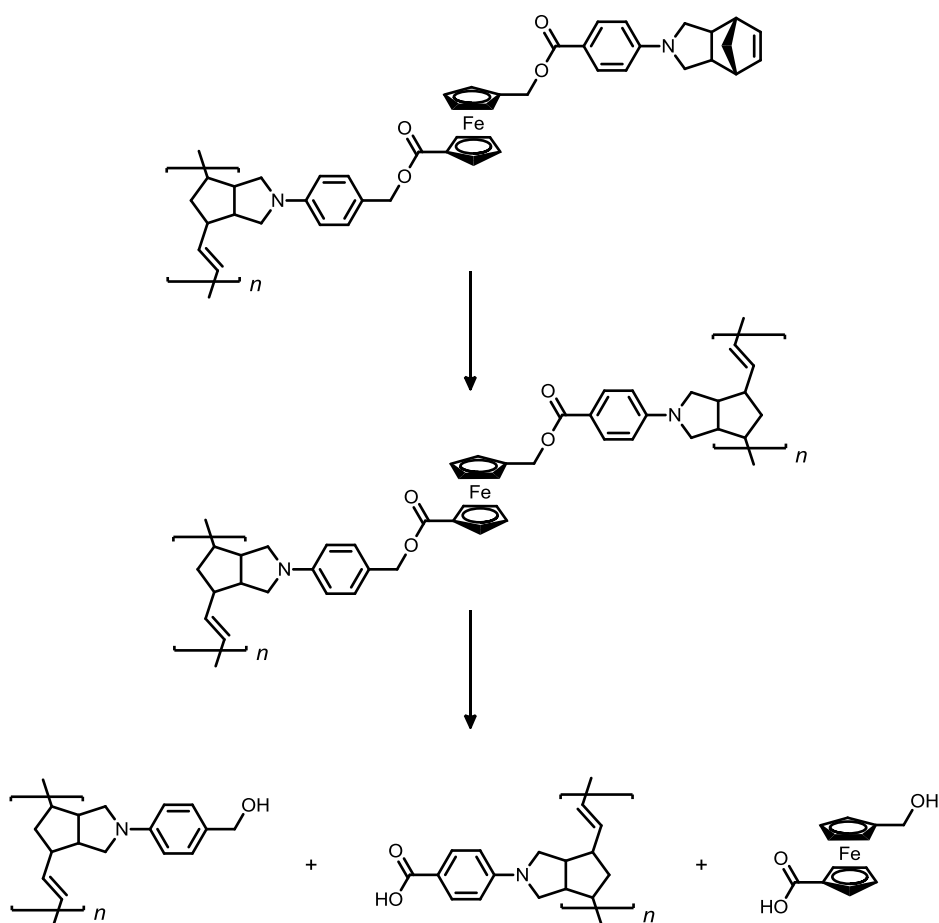


Figure 1.19. Template ring-opening metathesis polymerisation of norbornene derived monomers, covalently attached to a preformed template through ester groups. The daughter polymer was removed from the template via hydrolysis of the ester bonds.

Templates have been used to synthesise copolymers with periodic sequences but these templates are often not polymers themselves and can be viewed as more of an approach to control sequence rather than replicate a sequence. These systems will be discussed further in the following section.

1.2.2 Sequence controlled polymerisation

Currently, precise control of sequence when synthesising Nature's information oligomers in the laboratory is still achieved through solid support syntheses.^{106,107} In the absence of alternative methods, synthetic oligomers are often made using the same approach. Many contributors have achieved control of periodic sequence to create copolymers with repeating $-AB-$,¹⁰⁸⁻¹¹¹ or $-AAB-$ units,^{112,113} using syndiospecific catalysts,¹⁰⁹ templates,^{110,112} H-bonding,^{108,113} and using monomer units with sequence.^{111,114,115} However, these systems do not have the precise level of control that is necessary for an information molecule. For example, the Sawamoto group have reported a polymerisation where two (4-vinylphenyl)methanamine monomers and one 4-vinylpyridine monomer were preprogrammed as a template complex with palladium (Figure 1.20),¹¹² which after polymerisation and hydrolysis yielded an $-ABA-$ sequence polymer.

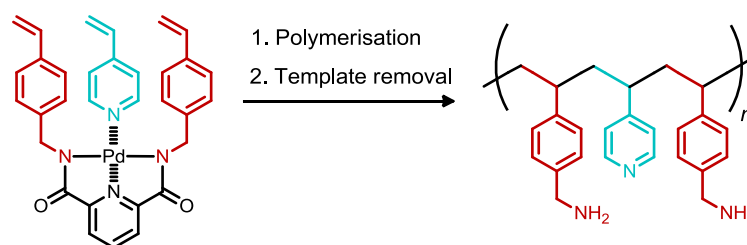


Figure 1.20. Sequence regulated radical polymerisation of where palladium acts a template, forming a complex with the monomer units.

A number of approaches have been reported that begin to show more complex control over monomer sequence in polymerisation.¹¹⁶⁻¹¹⁹ However, some of these approaches only allow regions of certain monomer units to be incorporated into a polymer (multiblock copolymers).^{116,118} For instance, Lutz has demonstrated how time-controlled additions of various *N*-substituted maleimides during the atom transfer radical polymerisation of styrene leads to well defined blocks of monomers

(Figure 1.21).¹¹⁶ The *N*-substituted maleimides have a much greater tendency for copolymerisation with styrene than homopolymerisation.

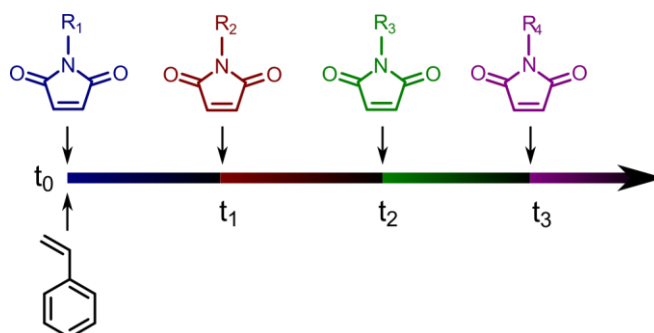


Figure 1.21. Atom transfer radical copolymerisation of styrene with time-controlled additions of various *N*-substituted maleimides where the regional distribution of monomers is determined by the time of comonomer addition. $R_1 - R_4$ represent various aryl and alkyl chains and $t_0 - t_3$ represent additions of *N*-substituted maleimides at different times.

Examples of monomer level sequence control in polymerisations do exist but each synthetic cycle is not quantitative and purification is still required.^{117,119} For instance the Junkers group have synthesised sequence controlled oligoacrylates using reversible addition fragmentation radical transfer (RAFT).¹¹⁹ The approach involves reacting an alkyl acrylate monomer with a trithiocarbonate bearing oligomer using AIBN initiator and then quenching the reaction after a short time (ca. 10 minutes). Size exclusion chromatography was used to separate oligomers which have reacted with just one monomer from oligomers that have reacted with multiple oligomers. The process was then repeated by reacting with another alkyl acrylate monomer (Figure 1.22). Essentially, this approach offers little or no advantages over conventional sequence controlled oligomer syntheses using the Merrifield approach and highlights the difficulties polymer chemists face in the pursuit of polymers with monomer level sequence control.

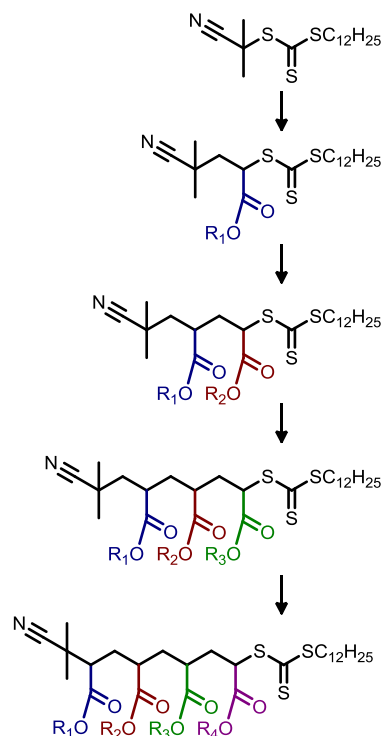


Figure 1.22. Schematic representation of a sequence controlled RAFT polymerisation. Each arrow represents addition of an alkyl acrylate monomer in the presence of initiator (AIBN), followed by purification using size exclusion chromatography. $R_1 - R_4$ represent various different alkyl chains.

1.2.3 Foldamers

Oligomers that adopt a specific folded conformation, stabilised by non-covalent interactions, are termed foldamers. The factors affecting the conformation can be rigidity in the backbone, interactions between functional groups along the oligomer, solvophobic effects or a combination thereof.¹²⁰ This area of chemistry began with the study of the conformations adopted by artificial amino acids.¹²¹⁻¹²³ Gellman coined the term foldamer in 1996 to describe the conformation adopted by β -peptide oligomers (Figure 1.23).¹²⁴ The principles dictating folding in these peptidomimetic compounds are the same as those of biopolymers such as proteins.

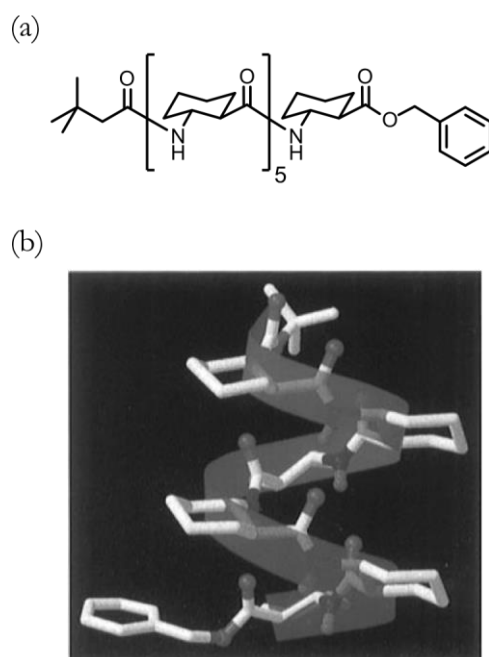


Figure 1.23. (a) Structure of foldamer β -peptide oligomer; (b) solid state conformation of the structure in (a), which adopt a well-defined helical conformation due to H-bonding between amide groups. The image in part (b) is taken directly from reference 124.

Foldamers with an aromatic oligoamide backbone have been extensively studied, beginning with work by Hamilton and co-workers who reported oligoanthranilamides that fold into helical bends.^{125,126} Since then, many similar designs have been reported including a foldamer based on 2'-pyridyl-2-pyridinecarboxamide units reported by Huc, Lehn and Krische, which adopts a helical curvature in chloroform or benzene (Figure 1.24).¹²⁷ The folded structure is the result of bifurcated intramolecular H-bonds between the amide protons and adjacent pyridine groups, and double-helical duplexes of two folded oligomers formed as a result of inter-strand aromatic stacking. Huc has extensively explored foldamers with structures similar to that in Figure 1.24 and has shown that the helix sense can be controlled by chiral groups at the end of the foldamer,¹²⁸⁻¹³⁰ or through the mutual steric repulsion of two covalently joined helices.¹³¹ In particular, Huc has shown the cavities created by a single or double helical folded aromatic oligoamides

make excellent hosts for guest binding.^{132,133} For instance, the oligomer with the structure shown in Figure 1.25 was able to fold and bind rod-like guests which shuttle within the foldamer.¹³⁴

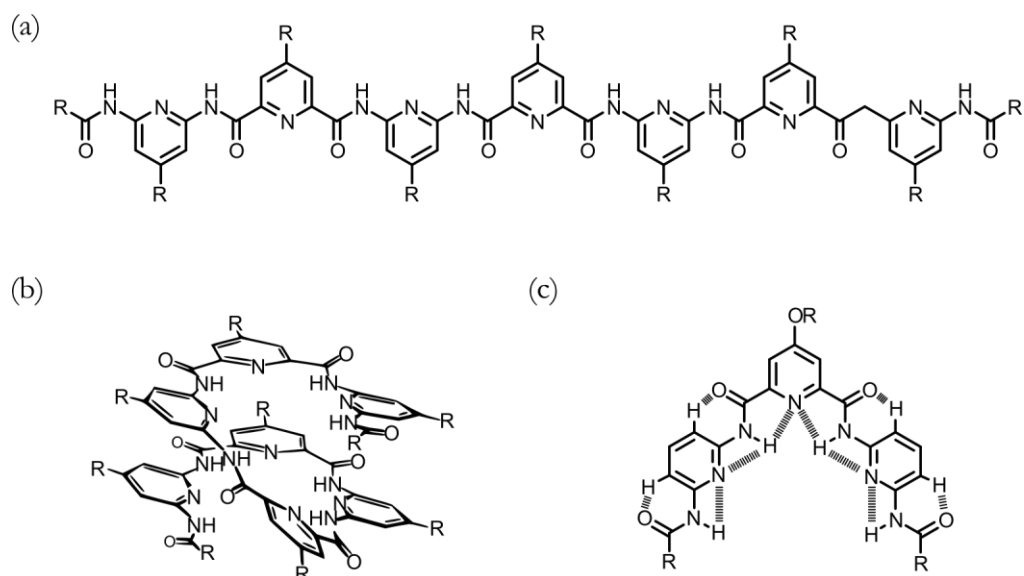


Figure 1.24. (a) Aromatic amide foldamer; (b) that adopts a helical conformation; (c) promoted by bifurcated intramolecular H-bonds. The R groups represent various different alkyl chains.

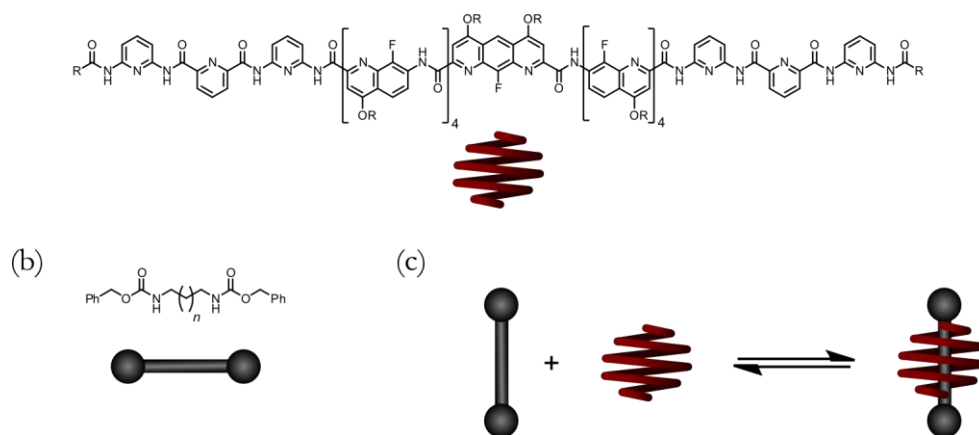


Figure 1.25. Helical foldamer which can bind a rod like guest: (a) structure of foldamer; (b) structure of rod guest; (c) schematic of helical foldamer binding guest. The R groups represent various different alkyl chains.

Gong has reported a number of aromatic oligoamides that adopt helical structures which rely on bifurcated intramolecular H-bonds between amide protons

and two aryl alkyl ethers.¹³⁵⁻¹³⁹ Crescent shaped oligomers based on this design have been shown to selectively bind guanadinium ions (Figure 1.26).¹⁴⁰

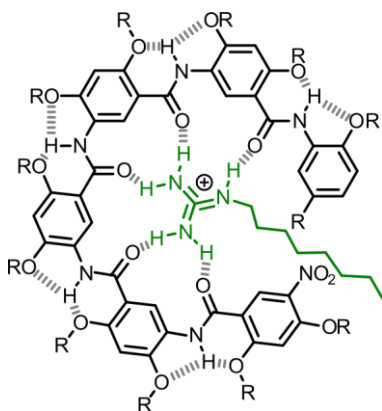


Figure 1.26. Foldamer, which can bind a guanadinium ion, where bifurcated H-bonds promote a helical conformation. The R groups represent various different alkyl chains.

Moore and co-workers have extensively studied a class of foldamers based on oligo-*m*-phenylacetylenes, which fold into a helical conformation stabilised by the solvophobic effect (Figure 1.27(a)).¹⁴¹ Upon addition of chloroform to a solution of foldamer in acetonitrile the well-defined structure is lost.¹⁴² Moore has shown that these foldamers can bind guests,¹⁴³⁻¹⁴⁵ and that foldamers strands can couple through dynamic imine bonds, the formation of which is promoted by the free energy released upon folding,¹⁴⁶ and binding guests (Figure 1.27(b)).¹⁴⁷

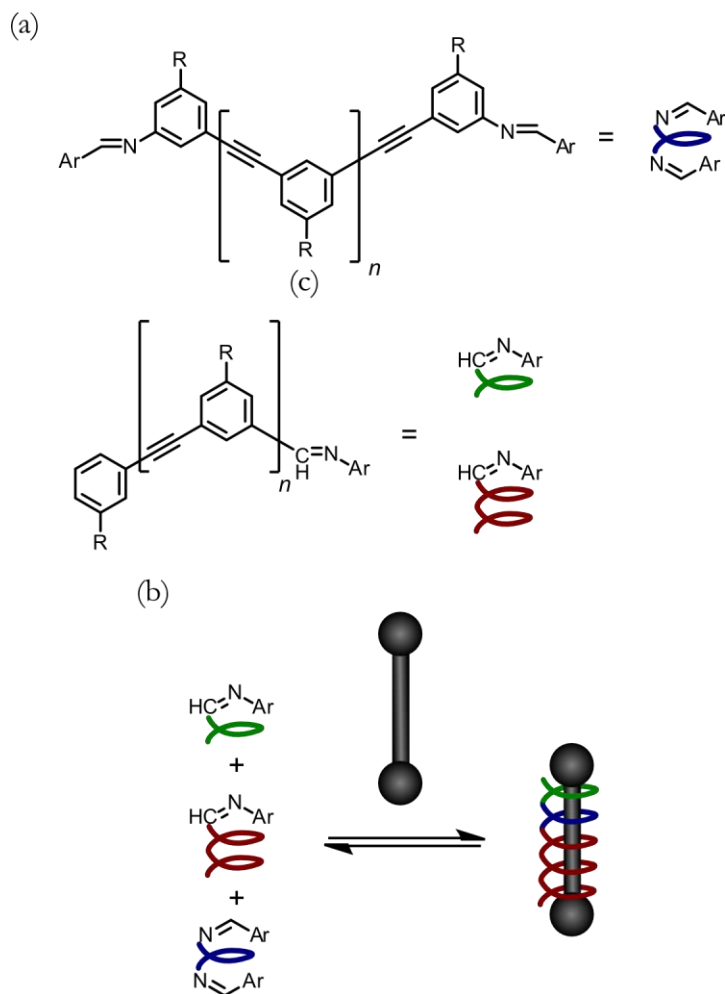


Figure 1.27. (a) foldamer based on oligo-*m*-phenylacetylenes; (b) foldamer formation is promoted by binding of a rod like guest where the foldamer is constructed of fragments coupled through dynamic imine chemistry. R = CO₂(CH₂CH₂O)₃CH₃.

Although many of the seminal examples of foldamers adopt helical structures, the term foldamer encompasses other conformations. For instance, Huc has reported an aromatic oligoamide that adopts a β -sheet type conformation through π - π interactions (Figure 1.28(a)),¹⁴⁸ and Gong has reported covalently tethered 1,3,5-benzenetricarboxamides that form a columnar conformation, promoted by H-bonds between the amide groups (Figure 1.28(b)).¹⁴⁹

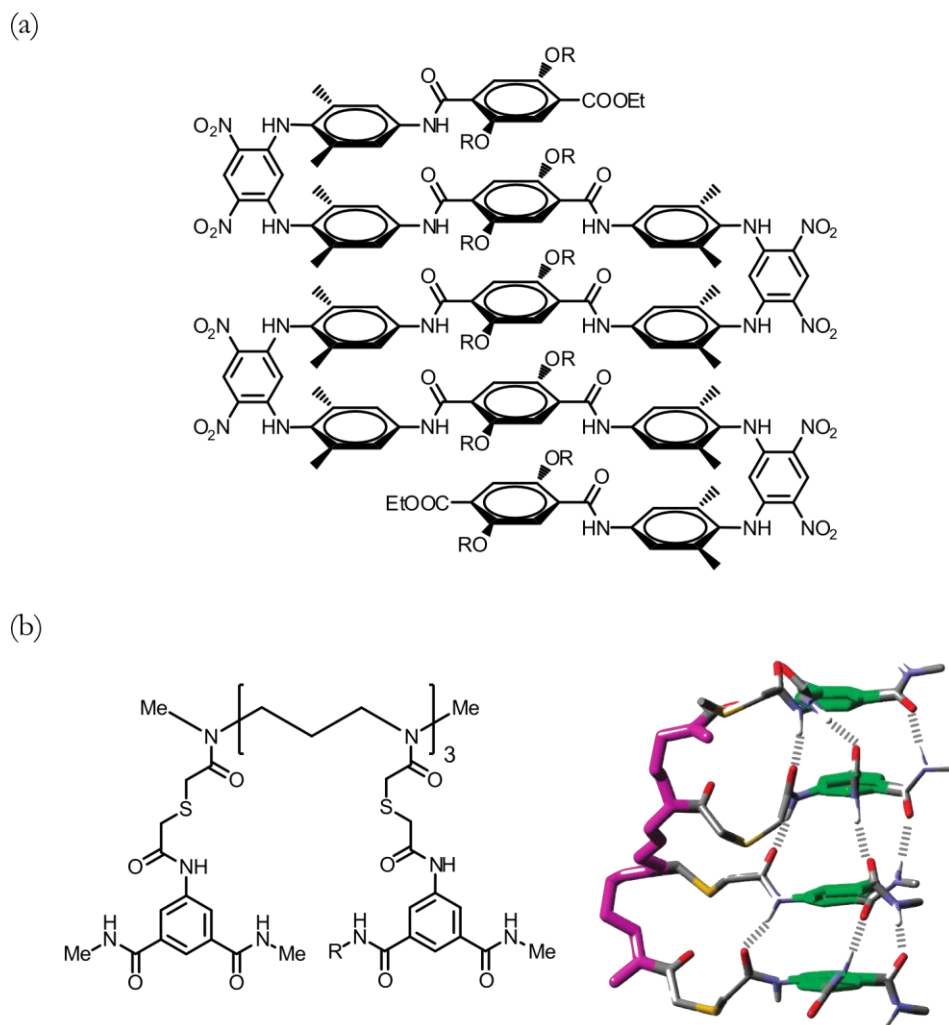


Figure 1.28. Foldamer examples that adopt conformations other than helical conformations: (a) a β -sheet type conformation; (b) the chemical structure (left) of an oligomer which adopts a columnar conformation (right) when modelled using density functional theory (the image on the right is taken from reference 149. R represents various different alkyl groups).

1.2.4 Synthetic duplexes without sequence information

Synthetic oligomers can also form duplexes via the interaction of several non-covalent interactions. However, most synthetic duplexes lack the ability to form in a sequence-specific manner, in similar way to DNA. This section will discuss systems which do not incorporate chemical sequence information, and focus on their synthesis and binding properties in the solution phase.

1.2.4.1 Metal helicates

Lehn *et al* have developed a series of metal helicates that display narcissistic self-sorting behaviour, with respect to ligand length. A series of oligo-bipyridines were synthesised via a series of alkoxylation reactions between mono-hydroxymethyl bipyridine derivatives and mono- or bis-bromomethyl bipyridines. Ligand oligomers from 2-mer up to 5-mer were made (**1.1**–**1.4** in Figure 1.29).

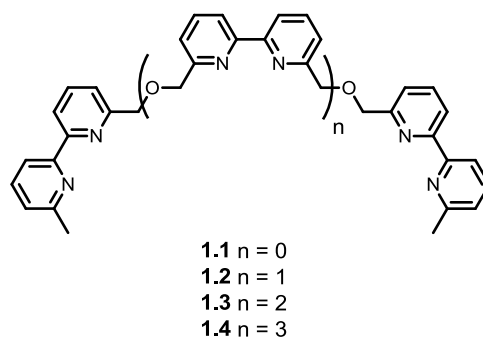


Figure 1.29. Structure of duplex helicates by Lehn *et al*.

Upon mixing equal equivalents of **1.1**, **1.2**, **1.3** and **1.4**, with a small excess of $[\text{Cu}(\text{CH}_3\text{CN})_4]\text{BF}_4$ (or PF_6) in CH_3CN ,¹⁵⁰ a complex mixture of species formed, which slowly over the course of days began to self-sort into duplex helicates formed of identical strands complexed with Cu(I) (Figure 1.30). A small amount of complexes formed between non-identical strands but the clear thermodynamic preference was for duplexes to form between strands of equal length.

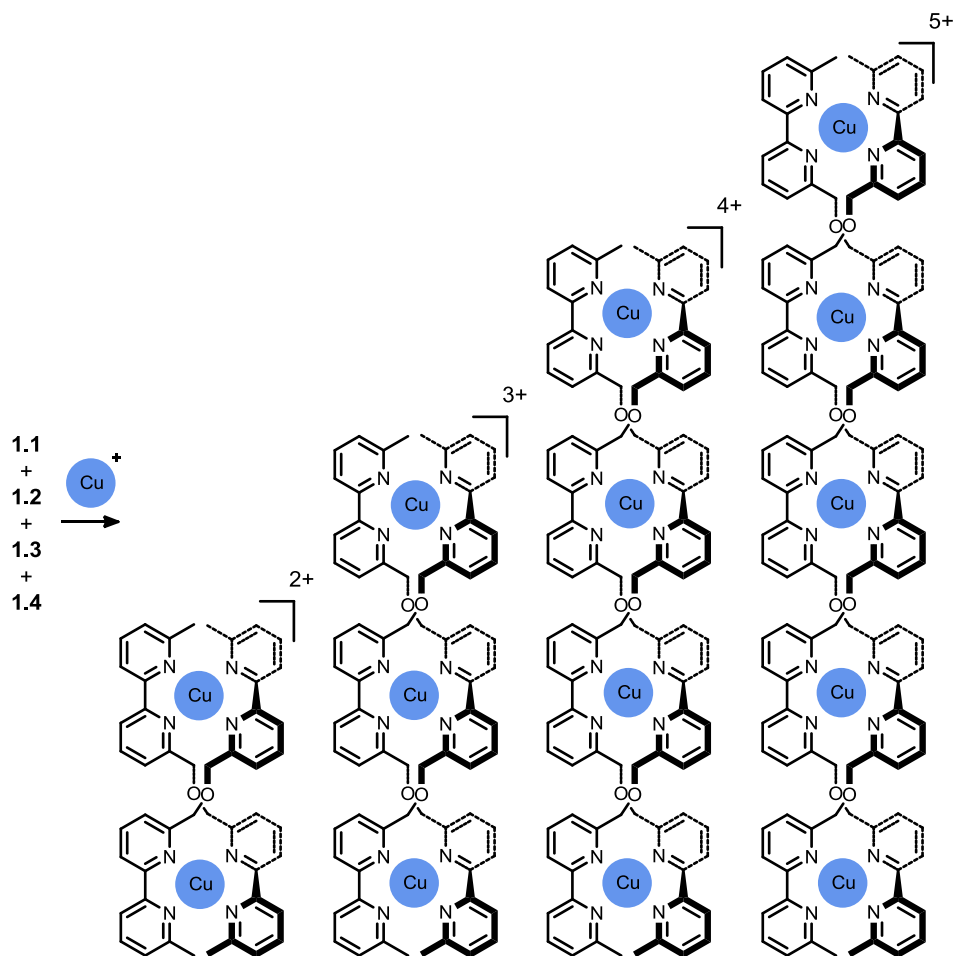


Figure 1.30. Duplex formation from a mixture of ligands **1.1–1.4** with Cu^+ into duplex helicates formed between two identical ligands.

1.2.4.2 H-bonded duplexes

Krische *et al* envisaged highly preorganised duplexes, based on the packing motif of 2,4-dichloro-6-aminotriazine in the solid state. Figure 1.31(a) shows an oligomer formed of covalently linked molecules of 2,4-dichloro-6-aminotriazine, and how a duplex of the oligomer could conceivably mimic the hydrogen bonding network seen in the crystal structure of 2,4-dichloro-6-aminotriazine.¹⁵¹

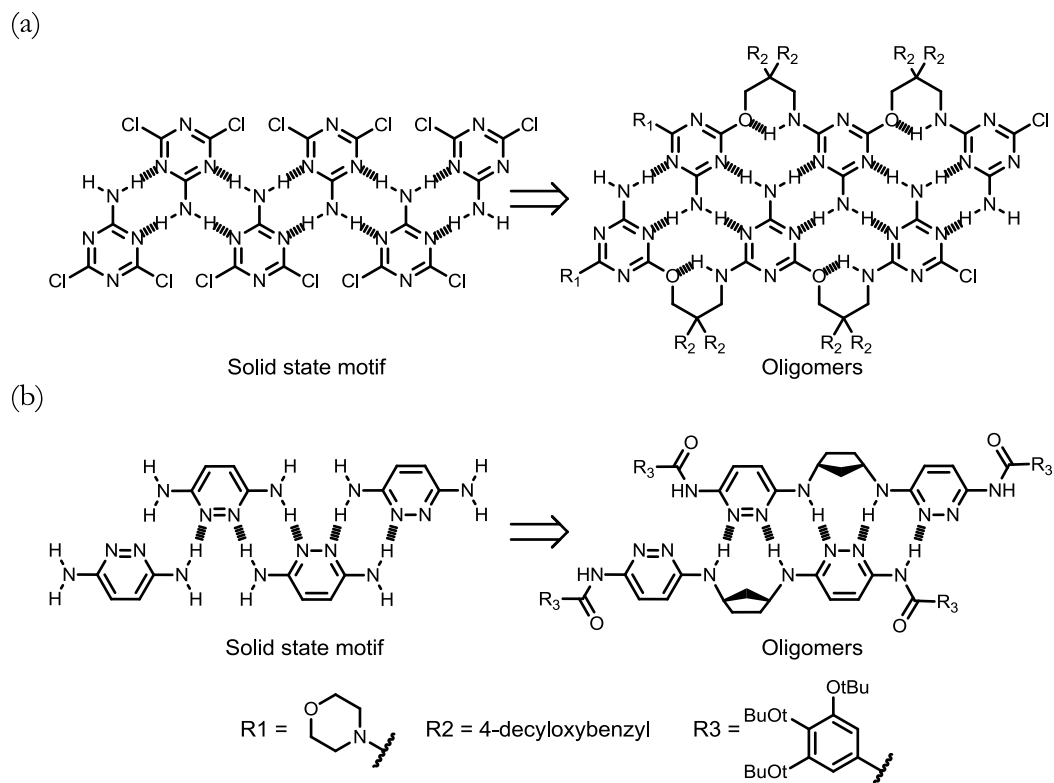


Figure 1.31. One dimensional H-bonding motifs in the solid state of (a) 2,4-dichloro-6-aminotriazine and (b) 3,6-diaminopyridazine, which acted as a starting point for the oligomer design.

Figure 1.32 shows the synthesis of the oligomers shown in Figure 1.31(a). The nucleophilic aromatic substitution reaction between amine **1.5** and monomer **1.6**, followed by removal of the *tert*-butyloxycarbonyl group gave the amine **1.7**. Further iterative coupling-deprotection steps propagated the oligomer, giving amines **1.8–1.9**, which are coupled with a capping group **1.10**. This approach generated oligomers **1.12–1.14** from 2-mer to 4-mer, and **1.11** was used as the smallest oligomer unit.

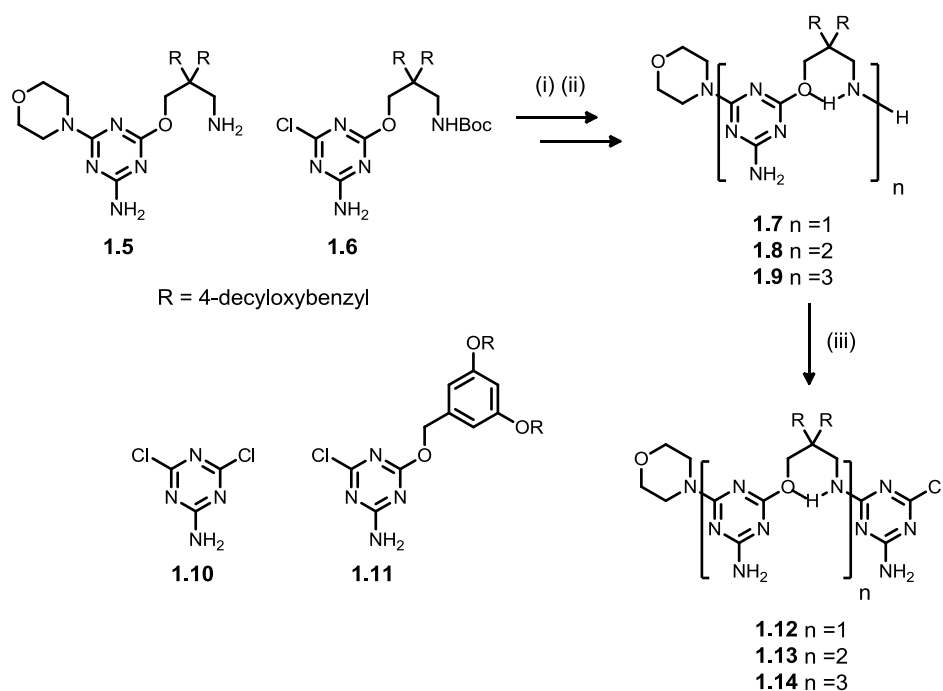


Figure 1.32. Synthesis of oligomers, via iterative coupling-deprotection reactions followed by capping with **1.10**.

Reagents and conditions: (i) $i\text{Pr}_2\text{NEt}$, CHCl_3 ; (ii) 10% TFA-DCM; (iii) **1.10** $i\text{Pr}_2\text{NEt}$, CHCl_3 .

Table 1.1 shows the dimerisation constants for oligomers **1.11–1.14**, measured via isothermal titration calorimetry. The association constant increases by 7 orders of magnitude from 1-mer to 3-mer and vapour pressure osmometry confirmed the stoichiometry of duplex formation. However, the association constant for the 4-mer **1.14** is lower than that of 3-mer **1.13**, which is explained by **1.14** being able to undergo intramolecular folding. Therefore, despite the presence of intermolecular $\text{N-H}\cdots\text{H}$ H-bonds between the secondary anilines and alkoxy oxygen atoms along the backbone, the oligomers are not rigid enough to prevent intramolecular folding.

Table 1.1. Dimerisation constants of oligomers **1.11–1.14**, determined by isothermal titration calorimetry in 1,2-dichloroethane at 20 °C.

| duplex | No. of HB ^a | K_a/M^{-1} |
|------------------|------------------------|-------------------|
| 1.11•1.11 | 2 | 1.9 |
| 1.12•1.12 | 6 | 2.3×10^4 |
| 1.13•1.13 | 10 | 6.9×10^8 |
| 1.14•1.14 | N/A | 1.1×10^3 |

^aNo. of HB = the number of H-bonds formed between two strands in a duplex.

Krische *et al* have also reported a set of duplexes based on oligomers of 3,6-diaminopyridazine units, using the same design approach as that used for oligomers based on 2,4-dichloro-6-aminotriazine (Figure 1.31(b)).¹⁵² The synthesis of oligomers from 1-mer to 3-mer involved the nucleophilic aromatic substitution reactions of cis-1,3-diaminocyclopentane **1.15** with aromatic halides **1.16** and **1.17**, yielding oligomers **1.18–1.20** (Figure 1.33). The oligomers all form homoduplexes with a large increase in duplex stability from 2-mer to 3-mer (Table 1.2), suggesting positive cooperativity between H-bonding interactions along the backbone.

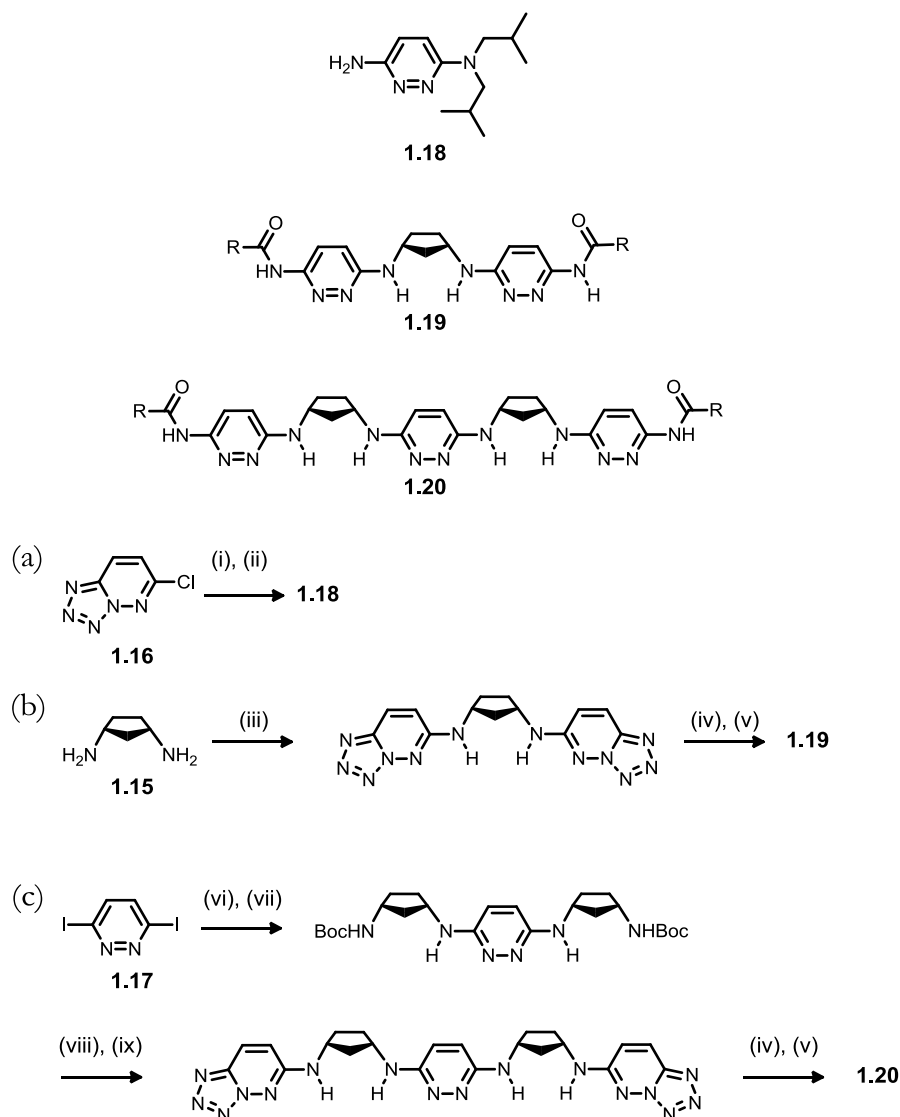


Figure 1.33. Synthesis and structures of oligomers (a) **1.18**; (b) **1.19** and (c) **1.20**.

Reagents and conditions: (i) $\text{HN}(\text{CH}_2\text{CH}(\text{CH}_3)_2)_2$; (ii) P^tBu_3 , CH_3CN , then $\text{HCl}(\text{aq})$; (iii) *cis*-1,3-diaminocyclopentane hydrochloride, K_2CO_3 , CH_3CN ; (iv) P^tBu_3 , CH_3CN ; (v) 3,4,5-tributoxybenzoic acid, toluene; (vi) CuI , **1.15**, ethylene glycol, K_3PO_4 , DMF ; (vii) di-*tert*-butyl dicarbonate (viii) TFA; (ix) **1.16**, K_2CO_3 .

Table 1.2. Dimerization constants of oligomers **1.18–1.20** measured via isothermal titration calorimetry in 1,2-dichloroethane at 20 °C.

| duplex | No. of HB ^a | K_d/M^{-1} |
|------------------|------------------------|---------------------|
| 1.18•1.18 | 2 | 5 |
| 1.19•1.19 | 6 | 870 |
| 1.20•1.20 | 10 | 8.0×10^5 |

^aNo. of HB = the number of H-bonds formed between two strands in a duplex.

Hunter *et al* have investigated the binding of a series of duplexes of aromatic amide oligomers.¹⁵³⁻¹⁵⁵ The reaction of diacid-chloride **1.21** and dianiline **1.22** generated oligomers **1.23** and **1.24**, which were separated from higher order species via column chromatography (Figure 1.34). Using a high ratio of dianiline **1.22**, limited the generation of higher order oligomers. The capping of dianilines **1.23** and **1.24** with two acid chlorides in different combinations, generated a series of aromatic amides capable of forming duplexes.

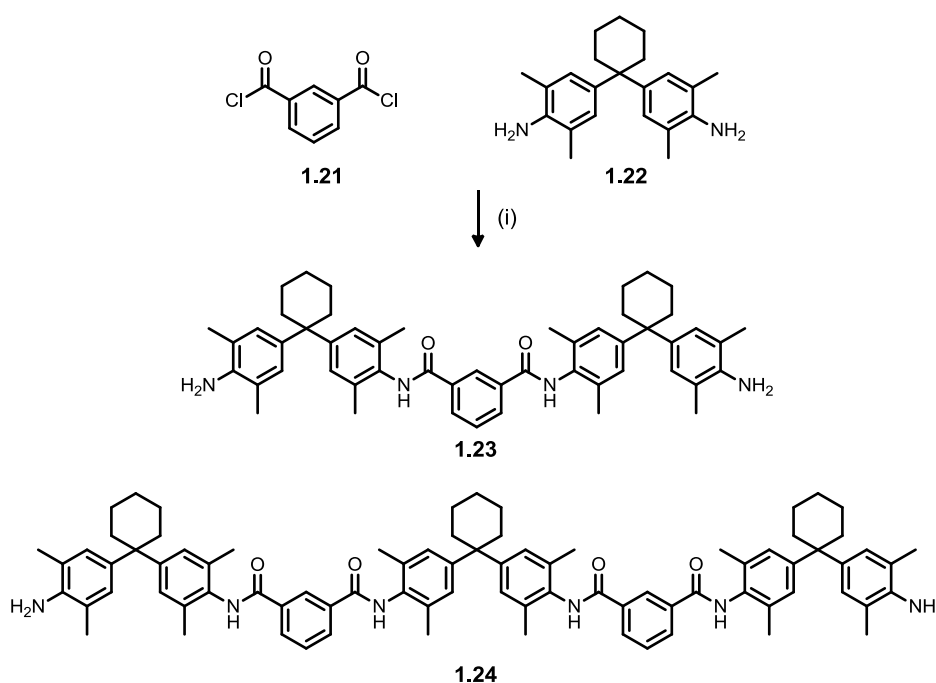


Figure 1.34. Synthesis of dianiline oligomers **1.23** and **1.24** via the acid chloride coupling reactions of **1.21** and **1.22**.
Reagents and conditions: (i) DCM, NEt_3 .

Figure 1.35 shows the structure of some of the homoduplexes measured. Longer oligomers had higher association constants (Table 1.3). Figure 1.36 shows the structure of the heteroduplexes formed, the association constants of which are at least an order of magnitude stronger than the self-association constants of the oligomers themselves (Table 1.4).

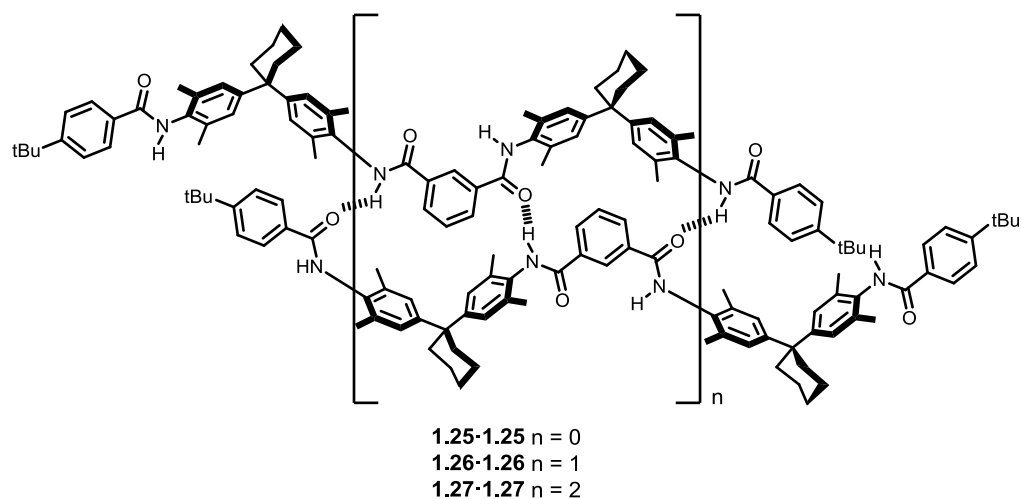


Figure 1.35. Duplexes of oligomers 1.25–1.27.

Table 1.3. Dimerisation constants of the oligomers 1.25–1.27, measured via ^1H NMR dilution in $\text{CDCl}_3/\text{CD}_3\text{OD}$ (95:5).

| duplex | m^a | K / M^{-1} |
|-----------|-------|---------------------|
| 1.25•1.25 | 2 | <1 |
| 1.26•1.26 | 4 | 12 ± 2 |
| 1.27•1.27 | 6 | 640 ± 170 |

^a m = the number of amide groups on each duplex.

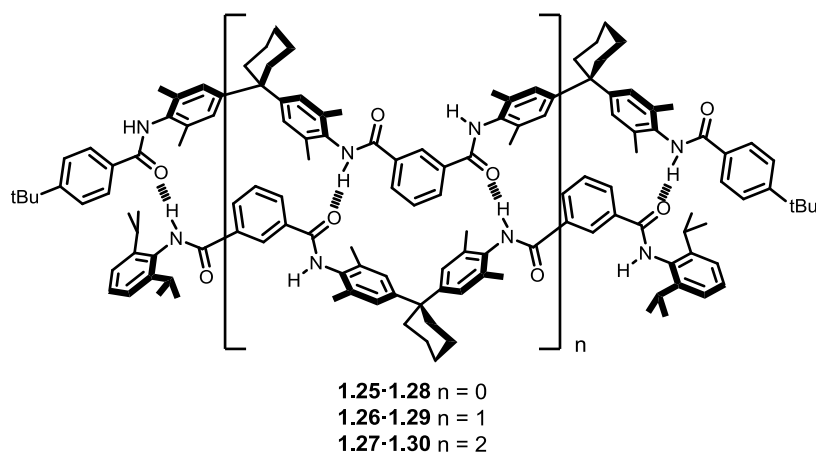


Figure 1.36. Heteroduplexes 1.25•1.28, 1.26•1.29 and 1.27•1.30.

Table 1.4. Heteroduplex association constants measured by ^1H NMR titration and dilution experiments in $\text{CDCl}_3/\text{CD}_3\text{OD}$ (95:5).

| duplex | K / M^{-1} |
|------------------|---------------------------|
| 1.25•1.28 | 18 ± 3 |
| 1.26•1.29 | 240 ± 11 |
| 1.27•1.30 | $5.5 \pm 3.5 \times 10^4$ |

Wisner *et al* have developed a series of H-bonded homo- and hetero-duplexes based on oligomers of nitrogen heterocycles. They used a convergent strategy to construct sequences of pyridine derivatives, as H-bond acceptors, and 1,3-thiazine-1,1-dioxide or pyrrole derivatives, as the H-bond donor. The key steps in the oligomer synthesis were the coupling of 2-(tributylstannyl)pyridine derivatives with 2-iodopyridines, or by forming thioethers, which were oxidised to the corresponding sulfone and cyclised with ammonium acetate to give the 1,4-thiazine-1,1-dioxide (Figure 1.37).

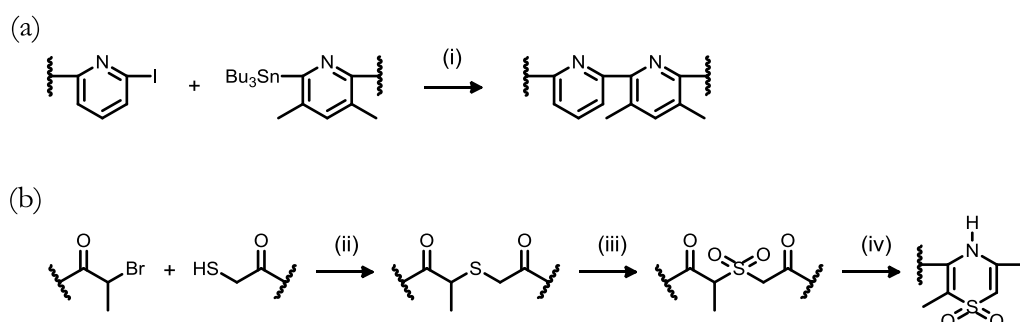


Figure 1.37. Key reaction steps in the synthesis of oligomers from Wisner *et al*.
 Reagents and conditions: (i) $\text{Pd}(\text{Ph}_3)_4$, toluene; (ii) Et_3N , CH_3CN ; (iii) *m*CPBA, DMF; (iv) NH_4OAc , AcOH .

In principle the oligomers are sequence programmable but alternating sequences of donor and acceptor groups such as **1.31•1.31** formed very weak complexes ($K_a = 4.5 \text{ M}^{-1}$ in CDCl_3).¹⁵⁶ The hydrogen bond acceptor/donor groups on adjacent sites are able to interact with each other in an either repulsive or attractive fashion.

These interactions are known as through secondary interactions and are represented by arrows in Figure 1.38.¹⁵⁷ The stabilities of the complexes are not independent of sequence, as duplex **1.32•1.33** has a dimerisation constant 4 orders of magnitude greater than **1.31•1.31**.¹⁵⁸ Alternating sequences have unfavourable interactions between adjacent H-bonding sites and lead to duplexes with lower association constants.¹⁵⁹

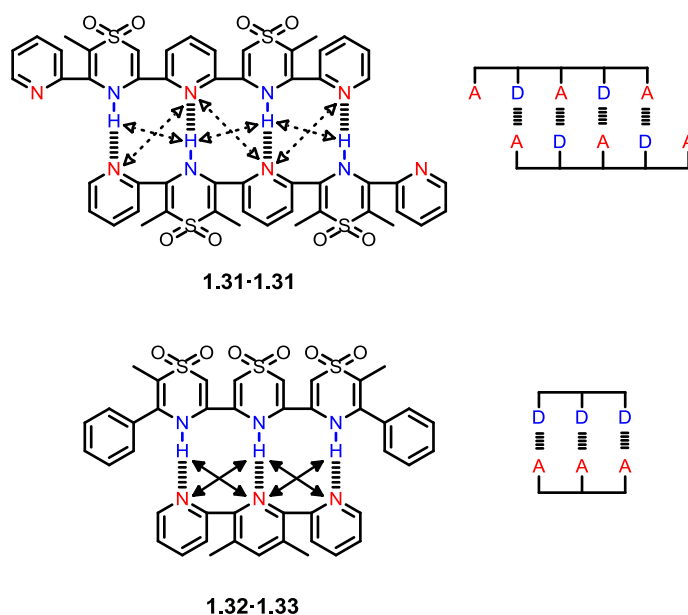


Figure 1.38. Examples of the duplexes developed by Wisner *et al.* Solid arrows show attractive secondary interactions and dashed arrows show repulsive secondary interactions. In the labels A is H-bond acceptor and D is H-bond donor.

1.2.4.3 Porphyrin ladders

Anderson *et al* have synthesised a series of oligomers based on linear chains of zinc porphyrins. The aim was to develop “ladder” complexes, where the binding of bidentate 1,4-diazabicyclo[2.2.2]octane (DABCO) ligands with the zinc metal sites on two adjacent porphyrin oligomers would lead to double stranded complexes.

Porphyrin oligomers can aggregate through interaction of the π -systems,¹⁶⁰ but in this example, bulky aromatic groups in the other two available *meso*-positions prevent aggregation by steric hindrance.¹⁶¹ Iterative Glaser type coupling of mono-

and di-acetylene bearing porphyrins, followed by deprotection of the silyl groups using tetra-*n*-butylammonium fluoride (TBAF), gave oligomers of 1–6 porphyrins in length.

Figure 1.39 shows ladder formation of oligomers with DABCO, given by process K_F . An excess of DABCO disrupts ladder formation, causing single stranded oligomer-ligand complex to form, given by process K_B in Figure 1.39. They observe an increase in ladder stability with oligomer length, with longer oligomers being less susceptible to disruption by an excess of DABCO.

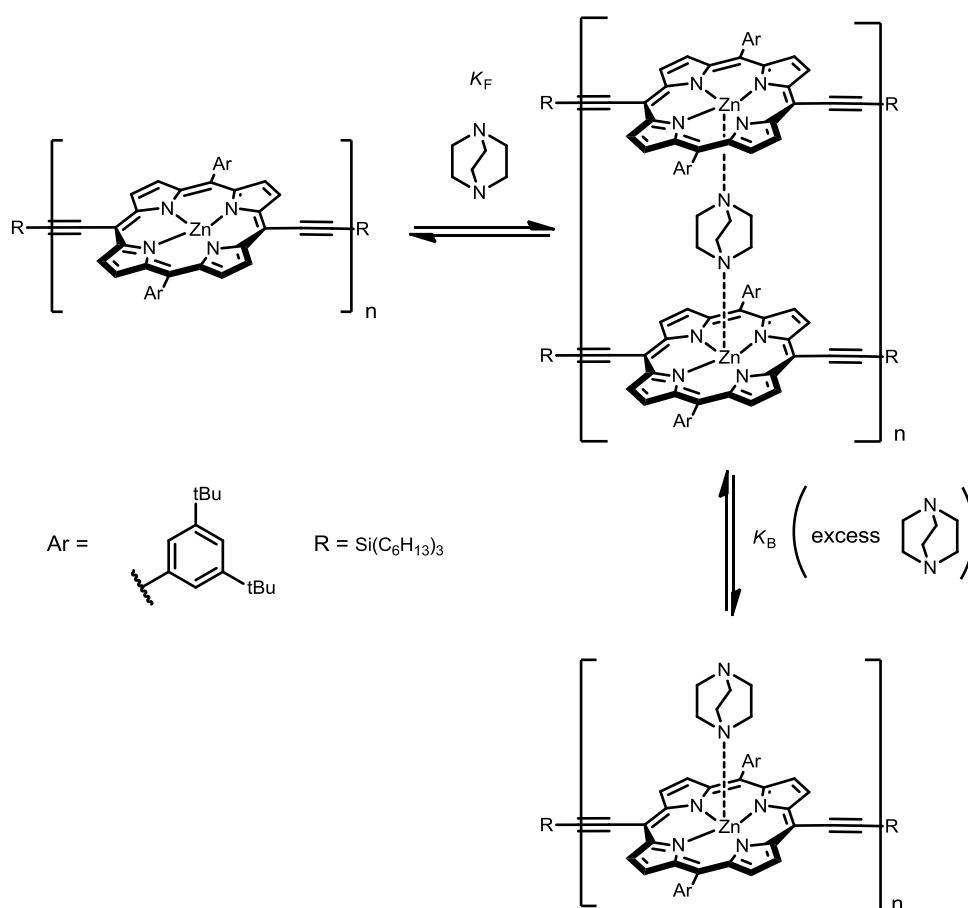


Figure 1.39. Ladder complexes form between porphyrin oligomers and DABCO (process K_F) but an excess of DABCO disrupts the ladders and leads to single stranded oligomer-ligand complex (process K_B).

The system displays all or nothing behaviour with respect to ladder formation and dissociation, with the only species present in solution being unbound oligomer,

ladder complex and single stranded oligomer-ligand complex. In other words, no frayed, partially bound species were observed. Only the addition of excess DABCO leads to the formation of single stranded oligomer-DABCO complex. The relationship between the Gibbs free energy of ladder formation ($\Delta G = -RT \ln K_F$) and ladder length is linear in both chloroform and toluene, where K_F can either be directly measured or calculated from K_B (Figure 1.40).

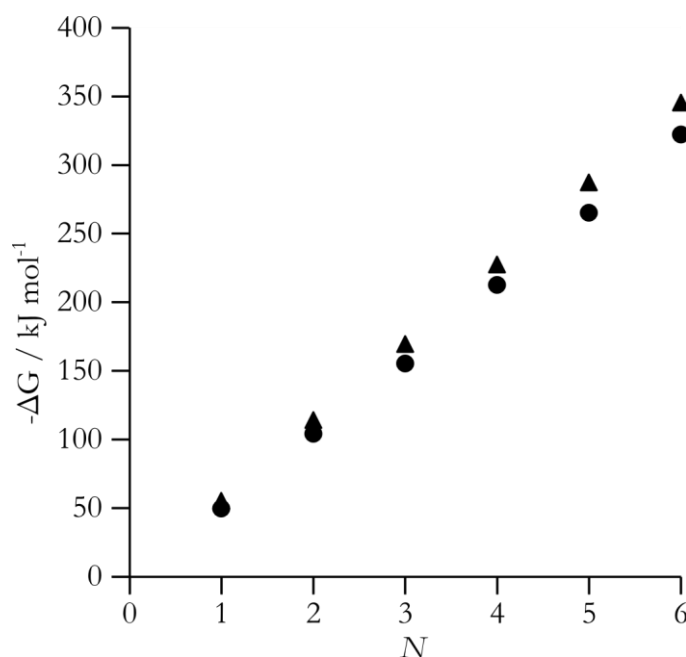


Figure 1.40. Gibbs free energy of ladder formation ($\Delta G = -RT \ln K_F$) against N (the number of porphyrins in the ladder complex) in toluene (triangles) and chloroform (circles).

A mixture of 2-mer and 3-mer oligomers, with DABCO, displays narcissistic self-sorting, with only homo-strand complexes being present. Furthermore, a mixture of two oligomers with the same number of porphyrins but one with a 1,4-diethynylbenzene linker rather than buta-1,3-diyne, also sorts into homoduplexes.

1.2.4.4 Duplexes assembled via aromatic stacking

Yamaguchi *et al* have prepared a number of ethynylhelicene oligomers, where the stability of homo- and hetero-duplexes depends on the sequence and type of

side-chains along the oligomer (group X in Figure 1.41).¹⁶²⁻¹⁶⁴ The synthesis involved coupling repeating units of (*P*)- or (*M*)-5,8-diethynyl-1,12-dimethylbenzo[*c*]phenanthrene with the aromatic triflate of the 1-mer unit via Sonogashira chemistry, followed by removal of the trimethylsilyl group with TBAF (Figure 1.41(a)). The yields of the coupling step diminish as the number of repeating units increases. This method allows the synthesis of homo- and alternating-sequences of oligomers, with respect to the nature of side chain group X.

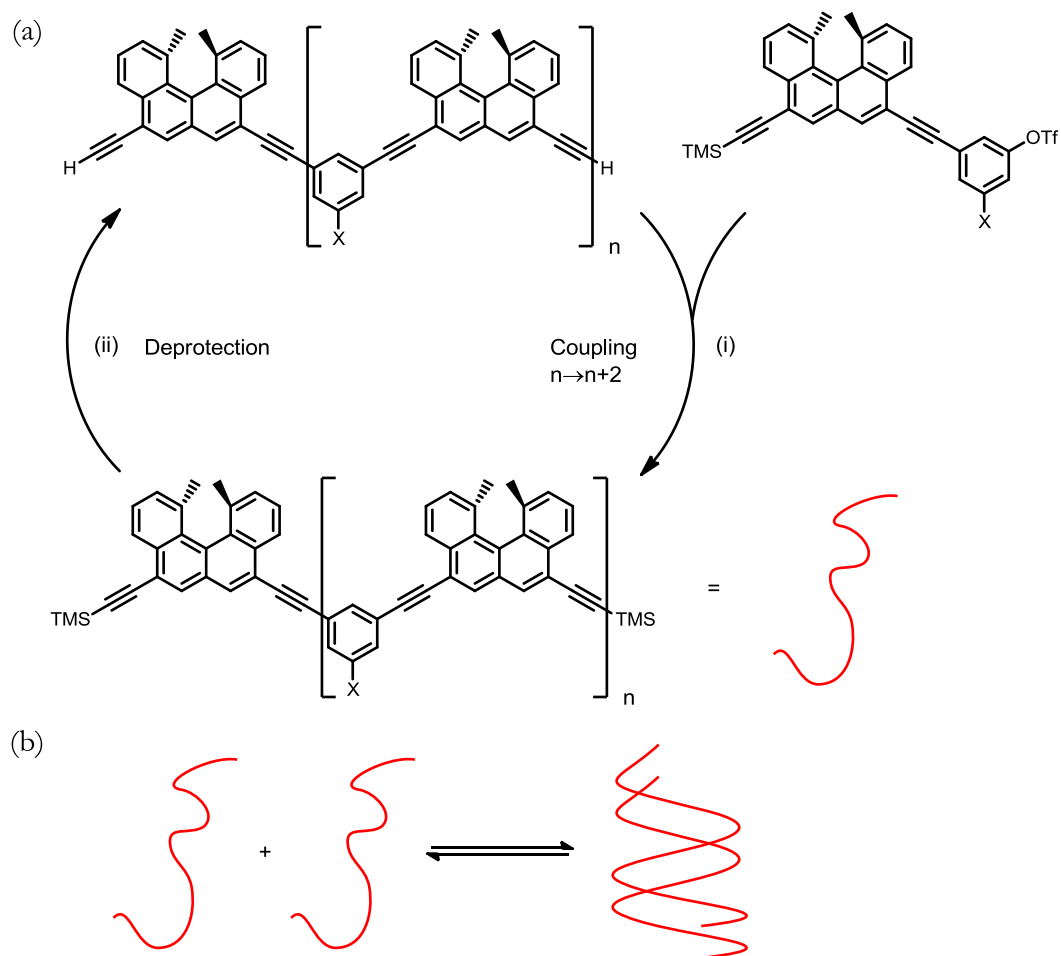


Figure 1.41. (a) General synthesis of linear or alternating sequences of oligo-helicenes where X represents a changeable side chain; (b) the helicenes, here represented by a red line, can reversibly form duplexes.

Reagents and conditions: (i) $\text{Pd}(\text{dba})_3$, CuI , PPh_3 , $\text{P}(\text{mesitylene})_3$; (ii) TBAF, NEt_3 .

Homooligomer duplexes form between identical chains with the same axial chirality and the same side-chain (Figure 1.41(b)), but the stability is dependent on

the nature of the side chain X in Figure 1.41(a). Changing the axial chirality of the 1,12-dimethylbenzo[c]phenanthrene unit and varying the “hardness/softness” of the side-chains, affects the stability of the duplexes formed between oligomers. Homoduplexes of oligomers with “harder” side-chains are more stable than those with “softer” properties, and they narcissistically sort into a mixture of “hard” and “soft” duplexes. The driving force of duplex formation is not clear. For instance, oligomers bearing alternating hard/soft side-chains do not form a stable duplex. Furthermore, a stable heteroduplex between a “hard” and “soft” oligomer forms when the “soft” oligomer have opposite axial chirality along the 1,12-dimethylbenzo[c]phenanthrene groups.

1.2.5 Synthetic duplexes with sequence information

There are few examples of synthetic oligomer systems which selectively form duplexes based on a specific sequence of non-covalent interactions. They are distinct from the examples described above, where either there is no sequence incorporated into the oligomer, or the stability of a sequence matched duplex is heavily dependent on the order of interaction sites. This section covers the synthesis and binding properties of synthetic duplexes which demonstrate sequence information.

1.2.5.1 Salt bridge duplexes

Yashima and co-workers have developed several examples of double helical duplex systems based on the H-bonding interaction between amidinium and carboxylate groups (Figure 1.42). The H-bonding sites form a sequence along a backbone comprised of *m*-terphenyl or *m*-phenylene derivatives.^{165,166} Several strategies have been used to couple together these units into oligomers and polymers, including Glaser,^{165,167,168} and Sonogashira type coupling,^{166,169} platinum coordination complexes,¹⁷⁰⁻¹⁷² and imine bond formation.¹⁷³⁻¹⁷⁵

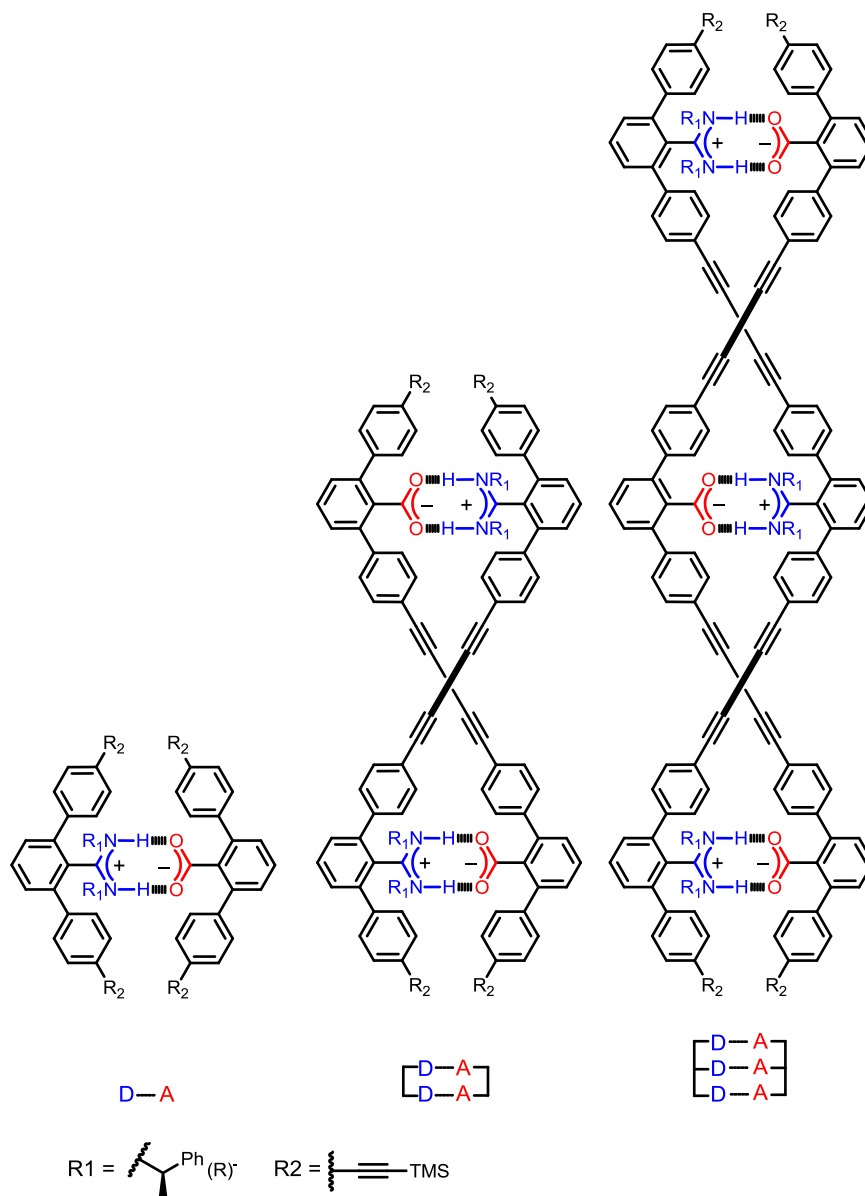


Figure 1.42. Examples of heteroduplexes formed from the oligomers developed by Yashima *et al.*

Trimethylsilane protected ethynyl groups act as protecting groups on the end of the monomers and oligomers. Deprotection of monomers and oligomers, with TBAF, gave doubly and singly deprotected species, which were isolated by flash chromatography (Figure 1.43(a)). All the oligomers are the products of Glaser type carbon–carbon bond formations between different combinations of deprotected species (Figure 1.43(a)). This strategy led to 2-mers (AA, DD, AD) 3-mers (AAA,

DDD, AAD, DDA, ADA, DAD) and 4-mers (AAAA, DDDD). They synthesised each oligomer with a complementary pair, so in the resulting duplex each amidine/carboxylate group would be able to form H-bonds with the paired oligomer. The pairs of oligomers which form duplexes give characteristic circular dichroism and ^1H NMR spectra. The duplexes are stable to analysis by high-performance liquid chromatography so characteristic retention times could be obtained for the duplexes formed.

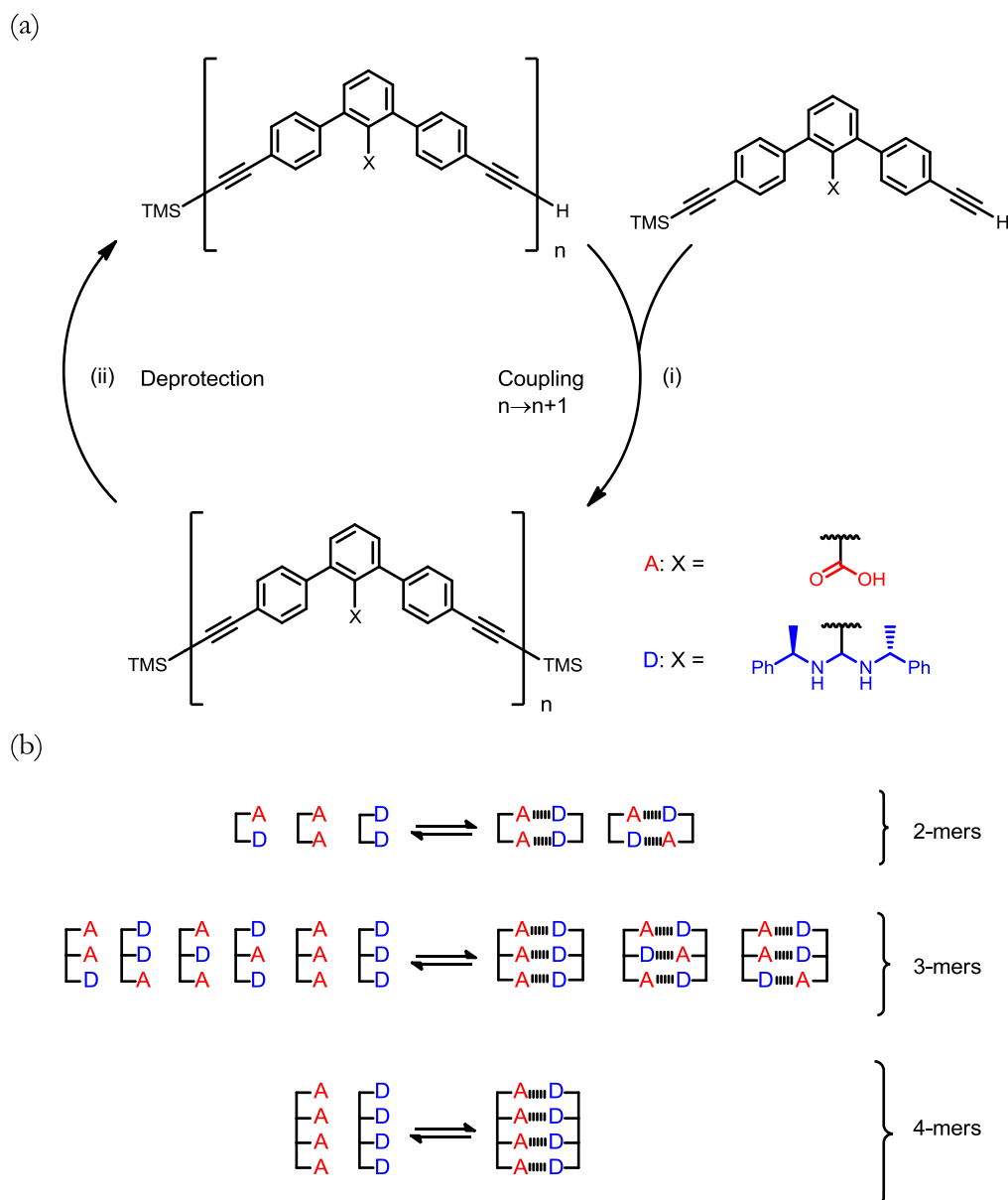


Figure 1.43. (a) Synthesis of oligomers involving deprotection and coupling steps to generate oligomers of any sequence and (b) All the different oligomers made form sequence selective duplexes, shown as a sequences of amidinium D and carboxylate A groups.

Reagents and conditions: (i) TBAF, THF; (ii) $\text{PdCl}_2(\text{PPh}_3)_2$, CuI, triethylamine/chloroform.

Despite not having matching sequences, some 3-mers form homoduplexes (AAD•AAD, ADA•ADA, DAD•DAD and DDA•DDA) and some pairs of 3-mers form heteroduplexes (AAA•DAD, AAA•DDA, AAD•ADA, AAD•DAD, AAD•DDD ADA•DDA, ADA•DDD and DAD•DDA). However, mixtures of

oligomers sort into sequence-matched pairs only, with no sequence mismatched duplexes being observable (Figure 1.43(b)). The circular dichroism and ^1H NMR spectra show a summation of the individual spectra corresponding to each sequence matched duplex. Furthermore, the high-performance liquid chromatography trace only shows peaks corresponding to sequence matched duplexes. This observation suggests the complexes formed between sequence matched partners are more stable than those formed between mismatched pairs.

A mixture of A, D, AA, DD, AAAA and DDDD in chloroform-*d* quickly self-sorts into complementary pairs, demonstrating the ability of the oligomers to also discriminate based on oligomer length.

Table 1.5 shows the binding constants in different solvents for the duplexes from 1-mer (A•D) to 2-mer (AA•DD). Even in polar solvent, duplex AA•DD binds appreciably. Binding constants for longer oligomers are not reported, but it is plausible that the binding of the duplexes increases dramatically with oligomer length.

Table 1.5. Association constants of duplexes measured at 25 °C.

| duplex | K_a (M^{-1}) | |
|--------|---------------------------|---------------------------------|
| | chloroform ^a | dimethyl sulfoxide ^b |
| A•D | 2.82×10^6 | ~ 0 |
| AA•DD | 6.43×10^{13} | 3.8×10^4 |

^a Measured by circular dichroism titration; ^b measured by ^1H NMR titration.

1.2.5.2 H-bonded duplexes

Gong *et al* have used a convergent strategy to synthesise fully programmable arrays of H-bond donors and acceptors, based on a backbone of aromatic amides (Figure 1.44(a)).^{176,177} The synthesis of the oligomers involves the amide coupling of

various previously synthesised components, an example of which is shown in Figure 1.45. Chemical sequence is incorporated by varying the configurations of the amide groups, whereby each benzene ring is part of a pair of aromatic amides, coding for DD, DA AD or AA sequences (Figure 1.44(b)). The oligomers have well-defined conformations dictated by steric hindrance and the intramolecular H-bonds between the N–H of the *N*-alkylbenzamide groups and the oxygen of the phenoxyalkyl groups.

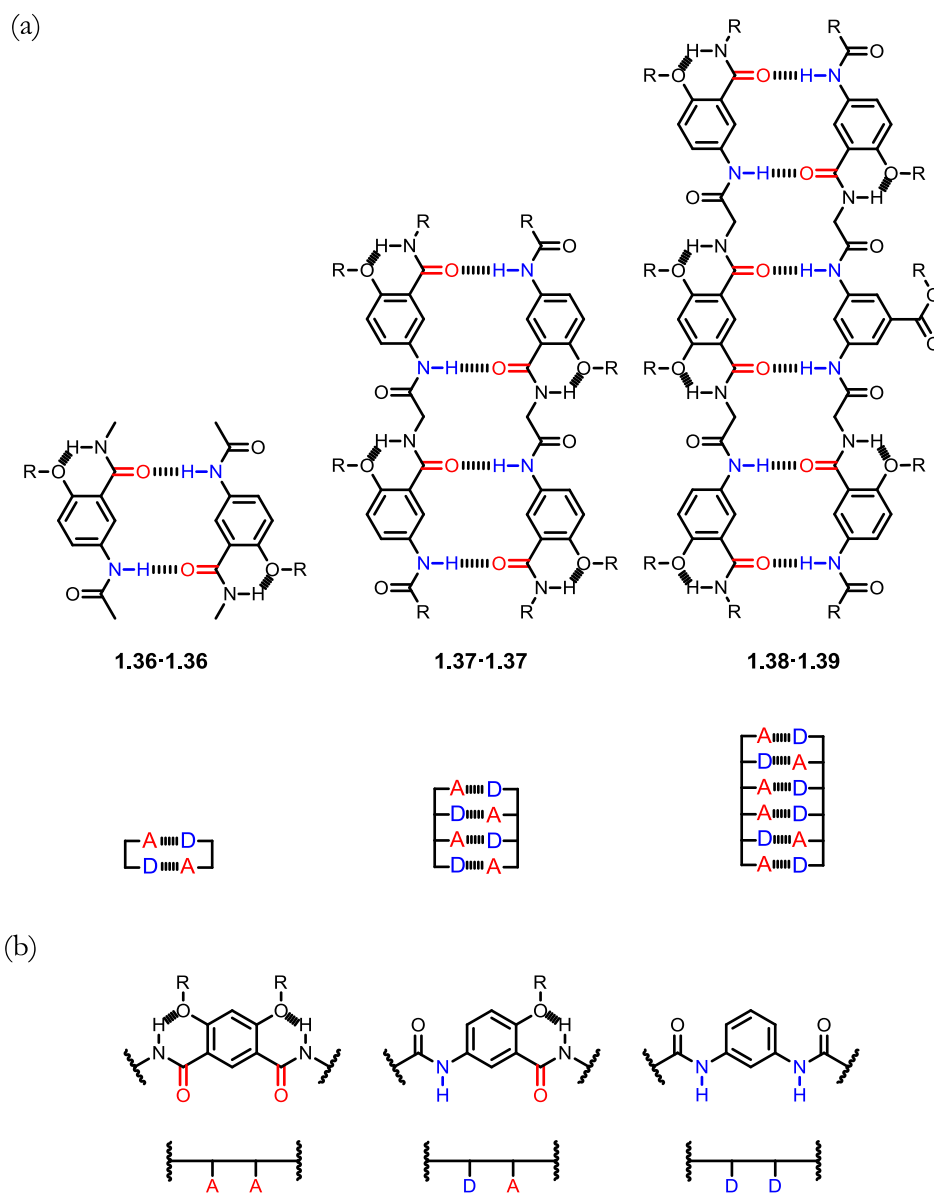


Figure 1.44. (a) Duplexes of oligomers based on aromatic amide H-bonded duplexes and (b) a representation of how different amide configurations leads to different sequences of H-bond donor and acceptor sites. The label A is H-bond acceptor and D is H-bond donor.

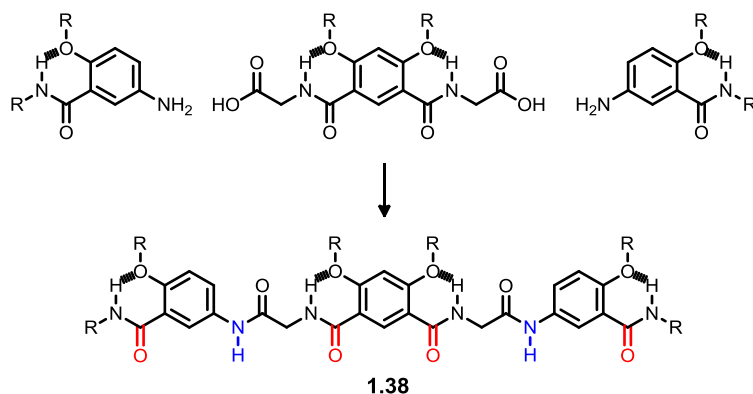


Figure 1.45. Final amide coupling step in the synthesis of aromatic oligoamides from Gong *et al.*

Reagents and conditions: 1-hydroxybenzotriazole, dimethylformamide and 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide hydrochloride.

Gong and co-workers have made different length aromatic oligoamides, capable of forming 2, 4 and 6 H-bonds (Figure 1.44(a)). The systems display positive cooperativity between H-bonds with an increase in association constant of 9 orders of magnitude between the duplex with 2 H-bonds **1.36•1.36** and the duplex with 6 H-bonds **1.38•1.39** (Table 1.6).¹⁷⁶ Importantly, duplex stability is essentially independent of the sequence of donor/acceptor interactions in a sequence matched duplex.

Table 1.6. Dimerisation constants measured by either ¹H NMR dilution experiments or isothermal titration calorimetry in chloroform.

| duplex | N^a | K (M^{-1}) |
|------------------|-------|--------------------------------|
| 1.36•1.36 | 2 | 25 ^b |
| 1.37•1.37 | 4 | 4.4×10^4 ^b |
| 1.38•1.39 | 6 | $>10^9$ ^c |

^a N = the number of amide groups per oligomer in the duplex; ^b measured by ¹H NMR dilution experiment in *d*-chloroform and ^c measured by isothermal titration calorimetry at 25 °C.

The binding constants of duplexes formed between pairs of sequence mismatched oligomers gave an insight into the thermodynamic cost of a single

H-bond mismatch, at different positions in the H-bond array (Figure 1.46).¹⁷⁸ Mismatched pairs still form duplexes (**1.38•1.40** and **1.38•1.41**), although competition experiments show that addition of sequence matched oligomer (**1.39**) fully displaces a mismatched sequence. No nuclear Overhauser effect contacts are seen between protons in the region around a mismatch, suggesting the chain twists to avoid unfavourable repulsion between two N–H or two C=O groups. The association constant for a complex with one mismatch is 40 times less stable than a sequence matched duplex with 6 H-bonds, and is in the same order of magnitude as the duplex with 4 H-bonds.

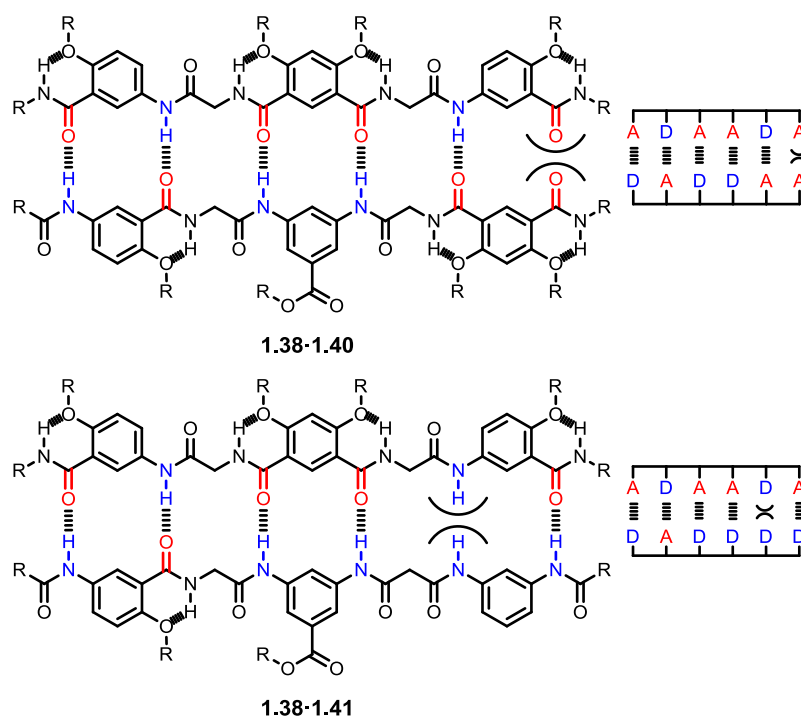


Figure 1.46. Sequence mismatched duplexes studied by Gong *et al*, where A is H-bond acceptor and D is H-bond donor.

A second strategy for incorporating sequence specificity in the duplexes is to incorporate a sequence of different sized aromatic spacers into the backbone (Figure 1.47). All the oligomers **1.42–1.46** have four aromatic amide groups and bear the same sequence of H-bond donor and acceptor groups (DDAA), which is conducive

to duplex formation between any pair of oligomers from the set, including homoduplex formation (Figure 1.47). However, each oligomer from **1.42–1.46** has a different sequence of aromatic spacers, which has an effect on the binding strength of each duplex (Table 1.7). The association constant of duplex **1.42•1.43** is three orders of magnitude greater than the homoduplexes **1.42•1.42** and **1.43•1.43**, due to the benzene and naphthalene linkers adding another element of sequence specificity to the duplex.¹⁷⁹ Oligomer **1.44** has the same H-bond sequence as **1.42** and **1.43** but is comprised only of benzene linkers. Addition of **1.44** to **1.42•1.43** failed to significantly disrupt the duplex. Likewise, addition of **1.44** to **1.45•1.55** also failed to disrupt the duplex. The binding constant of **1.44•1.44** was not reported, but the stability of the analogous **1.46•1.46** is in the same order of magnitude as **1.42•1.43**.

The use of different aromatic spacers leads to a small reduction in association constant at higher duplex lengths. For instance, the binding of **1.45•1.45** with two naphthalene spacers per oligomer is an order of magnitude weaker than **1.42•1.43** with just one naphthalene per oligomer. Furthermore, the binding of **1.48•1.48** with one naphthalene spacer is a factor of 10 times lower than the corresponding oligomer with just benzene linkers, **1.38•1.39** (Table 1.7).¹⁷⁶

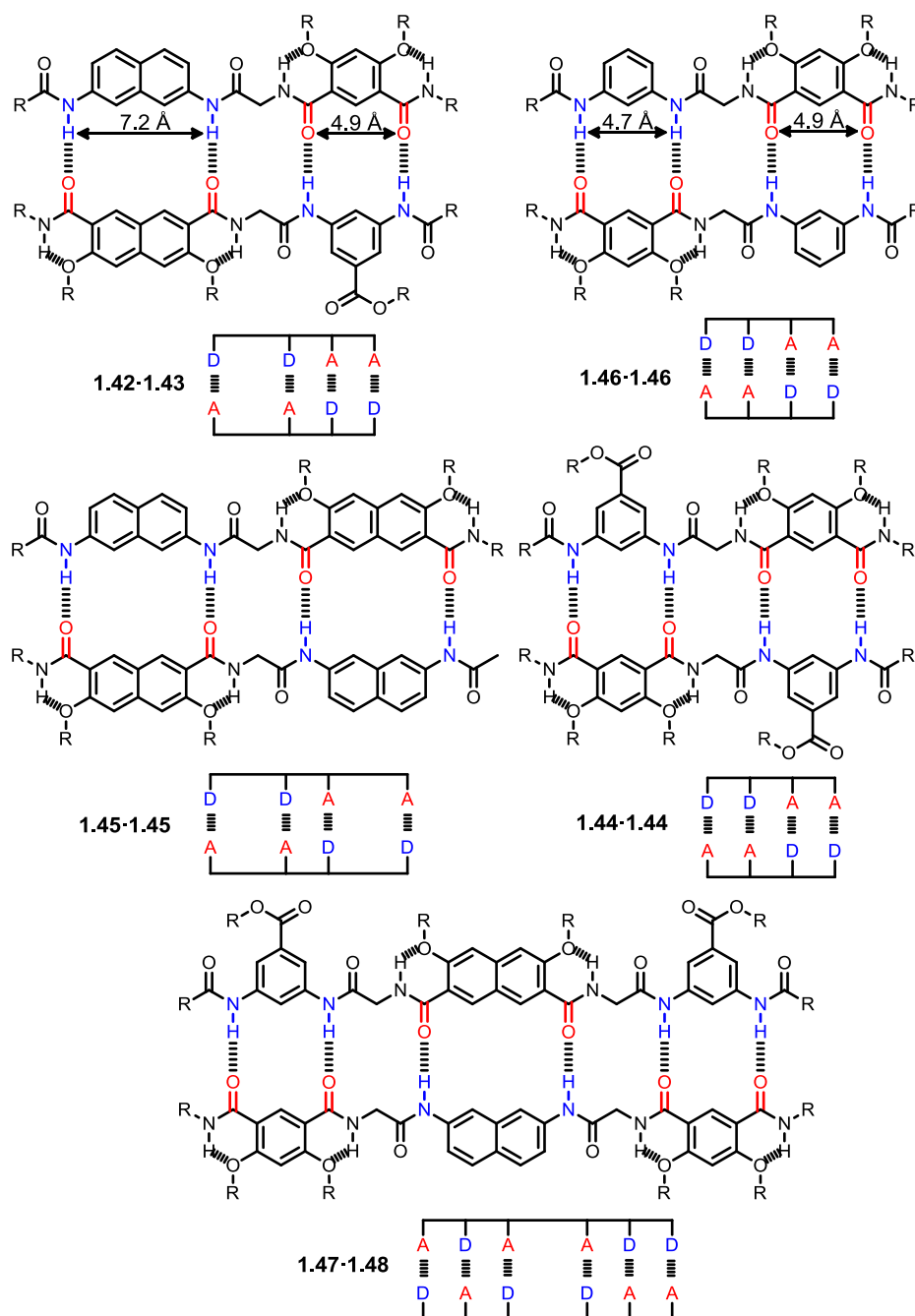


Figure 1.47. Structures and duplexes of oligomers, which show the use of both benzene and naphthalene spacers, where A is H-bond acceptor and D is H-bond donor. The distances indicated in the duplex of **1.42•1.43** were calculated from density functional theory calculations. The distances indicated in **1.44•1.44** were taken from single crystal X-ray diffraction data.

Table 1.7. Association constants measured by either isothermal titration calorimetry or ^1H NMR dilution and/or titration.

| duplex | N^a | K (M^{-1}) | technique |
|------------------|-------|-----------------------------|-----------|
| 1.42•1.42 | 4 | $(3.3 \pm 0.3) \times 10^1$ | b |
| 1.43•1.43 | 4 | $(1.7 \pm 0.2) \times 10^1$ | b |
| 1.42•1.43 | 4 | $(1.9 \pm 0.6) \times 10^4$ | b |
| 1.42•1.44 | 4 | $(9.6 \pm 0.8) \times 10^2$ | b |
| 1.43•1.44 | 4 | $(6.8 \pm 0.7) \times 10^2$ | b |
| 1.46•1.46 | 4 | $\sim 6.4 \times 10^4$ | b |
| 1.45•1.45 | 4 | $(1.6 \pm 0.3) \times 10^3$ | b |
| 1.47•1.47 | 6 | $(4.2 \pm 0.3) \times 10^1$ | b |
| 1.48•1.48 | 6 | $(5.2 \pm 0.4) \times 10^1$ | b |
| 1.47•1.48 | 6 | $(2.4 \pm 0.7) \times 10^5$ | c |
| 1.38•1.39 | 6 | $(3.5 \pm 1.6) \times 10^6$ | d |

^a N = number of amide groups in the each oligomer; ^b ^1H NMR dilution and titration experiments in chloroform- d ; ^c ^1H NMR titration experiment in 5% dimethyl sulfoxide- d_6 in chloroform- d ; ^d isothermal titration calorimetry in 5% dimethyl sulfoxide in chloroform.

1.3 Conclusions

Considering the many examples of chemically reengineered nucleic acids, the precise structures of nature's information molecules are not unique in their ability to perform sequence selective duplex formation and template directed synthesis. Many synthetic oligomer systems are capable of forming duplexes with adjustable stabilities and programmable sequence specificity, often with characteristics in parallel with the recognition properties of the nucleic acids. The binding constant of duplex formation increases with respect to the length of the oligomer. Oligomers capable of forming duplexes, which lack sequence information, will often selectively form

duplexes with an oligomer of the same length. H-bonding interactions are often an important intermolecular interaction in the construction of synthetic duplexes and are particularly suited to systems which incorporate sequence information. Oligomers incorporating a sequence of H-bond donor and acceptors have been shown to discriminate both on length and sequence, selectively forming duplexes with oligomers of matched sequence. Specificity has also been achieved by varying the distance between H-bonding sites, as another strategy to incorporate sequence information into an oligomer. Duplexes formed of sequence mismatched oligomers have significantly reduced stability and can be disrupted when the matched oligomer is added in a competition experiment. In certain cases, duplex formation is unaffected by the presence of competing H-bond donor and acceptor sites elsewhere in the molecule. Intramolecular folding of oligomers can be a competitive process to duplex formation in longer oligomers, but a rigid backbone seems to negate this. Some oligomers are preorganised in a certain conformation to facilitate duplex formation, via intramolecular H-bonds along the length of the oligomer.

A coupling-deprotection strategy is often advantageous in the synthesis of oligomers, as it gives control over oligomer length and sequence. An alternative strategy to oligomer synthesis is a convergent strategy, where sections of oligomer are combined, which may already be programmed with sequence information. Unsurprisingly, duplexes formed in non-polar solvents have larger binding constants, but often with lower solubility. A solution to this problem is to incorporate “greasy” side-chains in the oligomer design. Negative secondary electrostatic interactions are possible between H-bond donors and acceptor sites and these effects seemed to be minimised in oligomers where interaction sites are not in close proximity.

1.4 References

- 1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738.
- 2 Y. Kyogoku, R. C. Lord, and A. Rich, *Science*, 1966, **154**, 518–520.
- 3 R. M. Hamlin, R. C. Lord, and A. Rich, *Science*, 1965, **148**, 1734–1737.
- 4 R. A. Newmark and C. R. Cantor, *J. Am. Chem. Soc.*, 1968, **90**, 5010–5017.
- 5 Y. Kyogoku, R. C. Lord, and A. Rich, *Biochim. Biophys. Acta*, 1969, **179**, 10–17.
- 6 Y. Kyogoku, R. C. Lord, and A. Rich, *J. Am. Chem. Soc.*, 1967, **89**, 496–504.
- 7 D. Porschke, *Biopolymers*, 1971, 1989–2013.
- 8 S. I. Chan, M. P. Schweizer, P. O. P. Ts'o, and G. K. Helmkamp, *J. Am. Chem. Soc.*, 1964, **86**, 4182–4188.
- 9 W. Saenger, in *Principles of Nucleic Acid Structure*, Springer, New York, 1984, pp. 116–158.
- 10 W. Saenger, in *Principles of Nucleic Acid Structure*, Springer, New York, 1984, pp. 50–104.
- 11 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 964–967.
- 12 C. Switzer, S. E. Moroney, and S. A. Benner, *J. Am. Chem. Soc.*, 1989, **111**, 8322–8323.
- 13 J. A. Piccirilli, T. Krauch, S. Moroney, E., and S. A. Benner, *Nature*, 1990, **343**, 33–37.
- 14 Z. Yang, D. Hutter, P. Sheng, A. M. Sismour, and S. A. Benner, *Nucleic Acids Res.*, 2006, **34**, 6095–6101.
- 15 L. Zhang, Z. Yang, K. Sefah, K. M. Bradley, S. Hoshika, M.-J. Kim, H.-J. Kim, G. Zhu, E. Jiménez, S. Cansiz, I.-T. Teng, C. Champanhac, C. McLendon, C. Liu, W. Zhang, D. L. Gerloff, Z. Huang, W. Tan, and S. A. Benner, *J. Am. Chem. Soc.*, 2015, **137**, 6734–6737.
- 16 H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, and E. T. Kool, *Science*, 2003, **302**, 868–71.
- 17 H. Lu, K. He, and E. T. Kool, *Angew. Chemie Int. Ed.*, 2004, **43**, 5834–5836.
- 18 Y. Doi, J. Chiba, T. Morikawa, and M. Inouye, *J. Am. Chem. Soc.*, 2008, **130**, 8762–8768.

- 19 S. Hikishima, N. Minakawa, K. Kuramoto, Y. Fujisawa, M. Ogawa, and A. Matsuda, *Angew. Chemie Int. Ed.*, 2005, **44**, 596–598.
- 20 B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 1863–1872.
- 21 V. Tereshko, S. Portmann, E. C. Tay, P. Martin, F. Natt, K.-H. Altmann, and M. Egli, *Biochemistry*, 1998, **37**, 10626–10634.
- 22 M. Manoharan, M. Egli, M. Teplova, G. Minasov, V. Tereshko, G. B. Inamati, and P. D. Cook, *Nat. Struct. Biol.*, 1999, **6**, 535–539.
- 23 S. Pitsch, R. Krishnamurthy, M. Bolli, S. Wendeborn, A. Holzner, M. Minton, C. Lesueur, I. Schlönvogt, B. Jaun, and A. Eschenmoser, *Helv. Chim. Acta*, 1995, **78**, 1621–1635.
- 24 K.-U. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser, *Science (80-.)*, 2000, **290**, 1347–1351.
- 25 A. Aerschot Van, I. Verheggen, C. Hendrix, and P. Herdewijn, *Angew. Chemie Int. Ed.*, 1995, **34**, 1338–1339.
- 26 M. Tarköy, M. Bolli, and C. Leumann, *Helv. Chim. Acta*, 1994, **77**, 716–744.
- 27 M. Bolli, H. U. Tafelet, and C. Leumann, *Nucleic Acids Res.*, 1996, **24**, 4660–4667.
- 28 D. Renneberg and C. J. Leumann, *J. Am. Chem. Soc.*, 2002, **124**, 5993–6002.
- 29 S. K. Singh, A. A. Koshkin, J. Wengel, and P. Nielsen, *Chem. Commun.*, 1998, 455–456.
- 30 M. Petersen, C. B. Nielsen, K. E. Nielsen, G. A. Jensen, K. Bondensgaard, S. K. Singh, V. K. Rajwansi, A. A. Koshkin, B. M. Dahl, J. Wengel, and J. P. Jacobsen, *J. Mol. Recognit.*, 2000, **13**, 44–53.
- 31 L. Zhang, A. Peritz, and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174–4175.
- 32 E. Meggers and L. Zhang, *Acc. Chem. Res.*, 2010, **43**, 1092–102.
- 33 S. M. Freier and K.-H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429–4443.
- 34 S. M. Gryaznov, D. H. Lloyd, J. K. Chen, R. G. Schultz, L. A. DeDionisio, L. Ratmeyer, and W. D. Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 5798–5802.
- 35 A. De Mesmaeker, C. Lesueur, M.-O. Bévière, A. Waldner, V. Fritsch, and R. M. Wolf, *Angew. Chemie Int. Ed.*, 1996, **35**, 2790–2794.
- 36 R. J. Jones, K. Y. Lin, J. F. Milligan, S. Wadwani, and M. D. Matteucci, *J. Org. Chem.*, 1993, **58**, 2983–2991.

- 37 A. L. Roughton, S. Portmann, S. A. Benner, and M. Egli, *J. Am. Chem. Soc.*, 1995, **117**, 7249–7250.
- 38 C. Richert, A. L. Roughton, and S. A. Benner, *J. Am. Chem. Soc.*, 1996, **118**, 4518–4531.
- 39 H. Isobe, T. Fujino, N. Yamazaki, M. Guillot-Nieckowski, and E. Nakamura, *Org. Lett.*, 2008, **10**, 3729–32.
- 40 S. A. Benner, *Acc. Chem. Res.*, 2004, **37**, 784–797.
- 41 P. Nielsen, M. Egholm, R. Berg, and O. Buchardt, *Science (80-.)*, 1991, **254**, 1497–1500.
- 42 P. Wittung, P. E. Nielsen, O. Buchardt, M. Egholm, and B. Nordén, *Nature*, 1994, **368**, 561–3.
- 43 S. Brown, S. Thomson, J. Veal, and D. Davis, *Science (80-.)*, 1994, **265**, 777–780.
- 44 P. E. Nielsen and M. Egholm, *Curr. Issues Mol. Biol.*, 1999, 1, 89–104.
- 45 H. Rasmussen, J. S. Kastrup, J. N. Nielsen, J. M. Nielsen, and P. E. Nielsen, *Nat. Struct. Biol.*, 1997, **4**, 98–101.
- 46 D. J. Wilcock, A. Adams, C. J. Cardin, and L. P. G. Wakelin, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1996, **52**, 481–485.
- 47 A. Sen and P. E. Nielsen, *Nucleic Acids Res.*, 2007, **35**, 3367–3374.
- 48 A. Sen and P. E. Nielsen, *Biophys. Chem.*, 2009, **141**, 29–33.
- 49 P. E. Nielsen, G. Haaima, A. Lohse, and O. Buchardt, *Angew. Chemie Int. Ed.*, 1996, **35**, 1939–1942.
- 50 A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil, and D. H. Ly, *J. Am. Chem. Soc.*, 2006, **128**, 10258–10267.
- 51 T. Vilaivan and C. Srisuwannaket, *Org. Lett.*, 2006, **8**, 1897–1900.
- 52 Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel, and M. R. Ghadiri, *Science*, 2009, **325**, 73–7.
- 53 Y. Krishnan and F. C. Simmel, *Angew. Chem. Int. Ed.*, 2011, **50**, 3124–56.
- 54 N. C. Seeman, *J. Theor. Biol.*, 1982, **99**, 237–247.
- 55 P. W. K. Rothmund, *Nature*, 2006, **440**, 297–302.
- 56 N. C. Seeman, *J. Biomol. Struct. Dyn.*, 1990, **8**, 573–581.

- 57 S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih, *Nucleic Acids Res.*, 2009, **37**, 5001–5006.
- 58 H. Yan, T. H. LaBean, L. Feng, and J. H. Reif, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 8103–8108.
- 59 P. Yin, R. F. Hariadi, S. Sahu, H. M. T. Choi, S. H. Park, T. H. Labean, and J. H. Reif, *Science*, 2008, **321**, 824–826.
- 60 Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, *Science*, 2012, **338**, 1177–1183.
- 61 S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, *Nature*, 2009, **459**, 414–418.
- 62 W. M. Shih, J. D. Quispe, and G. F. Joyce, *Nature*, 2004, **427**, 618–621.
- 63 B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, and J. L. Neumann, *Nature*, 2000, **406**, 605–8.
- 64 Z.-G. Wang, J. Elbaz, and I. Willner, *Nano Lett.*, 2011, **11**, 304–9.
- 65 E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, and J. Kjems, *Nature*, 2009, **459**, 73–6.
- 66 R. Chhabra, J. Sharma, Y. Liu, and H. Yan, *Nano Lett.*, 2006, **6**, 978–83.
- 67 H. Gu, J. Chao, S.-J. Xiao, and N. C. Seeman, *Nature*, 2010, **465**, 202–205.
- 68 M. Langecker, V. Arnaut, T. G. Martin, J. List, S. Renner, M. Mayer, H. Dietz, and F. C. Simmel, *Science (80-.)*, 2012, **338**, 932–936.
- 69 N. C. Seeman, C. Mao, T. H. LaBean, and J. H. Reif, *Nature*, 2000, **407**, 493–496.
- 70 L. Adleman, *Science*, 1994, **266**, 1021–1024.
- 71 R. S. Braich, N. Chelyapov, C. Johnson, P. W. K. Rothmund, and L. Adleman, *Science*, 2002, **296**, 499–502.
- 72 S. M. Douglas, I. Bachelet, and G. M. Church, *Science (80-.)*, 2012, **335**, 831–834.
- 73 R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, and J. Feigon, *Proc. Natl. Acad. Sci.*, 1993, **90**, 3745–3749.
- 74 J. B.-H. Tok, J. Cho, and R. R. Rando, *Nucleic Acids Res.*, 2000, **28**, 2902–2910.
- 75 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.
- 76 M. Mascini, I. Palchetti, and S. Tombelli, *Angew. Chemie Int. Ed.*, 2012, **51**, 1316–1332.

- 77 A. D. Ellington and J. W. Szostak, *Nature*, 1992, **355**, 850–852.
- 78 C. Wilson and J. W. Szostak, *Chem. Biol.*, 1998, **5**, 609–617.
- 79 K.-A. Yang, M. Barbu, M. Halim, P. Pallavi, B. Kim, D. M. Kolpashchikov, S. Pecic, S. Taylor, T. S. Worgall, and M. N. Stojanovic, *Nat. Chem.*, 2014, **6**, 1003–1008.
- 80 B. Seelig, S. Keiper, F. Stuhlmann, and A. Jäschke, *Angew. Chemie Int. Ed.*, 2000, **39**, 4576–4579.
- 81 M. Chandra and S. K. Silverman, *J. Am. Chem. Soc.*, 2008, **130**, 2936–2937.
- 82 Y. Lu and J. Liu, *Curr. Opin. Biotechnol.*, 2006, **17**, 580–588.
- 83 R. Naylor and P. T. Gilham, *Biochemistry*, 1966, **5**, 2722–2728.
- 84 Z.-Y. J. Zhan and D. G. Lynn, *J. Am. Chem. Soc.*, 1997, **119**, 12420–12421.
- 85 X. Wu, G. Delgado, R. Krishnamurthy, and A. Eschenmoser, *Org. Lett.*, 2002, **4**, 1283–1286.
- 86 C. Böhrer, P. E. Nielsen, and L. E. Orgel, *Nature*, 1995, **376**, 578–581.
- 87 T. Inoue, G. F. Joyce, K. Grzeskowiak, L. E. Orgel, J. M. Brown, and C. B. Reese, *J. Mol. Biol.*, 1984, **178**, 669–676.
- 88 J. G. Schmidt, P. E. Nielsen, and L. E. Orgel, *J. Am. Chem. Soc.*, 1997, **119**, 1494–1495.
- 89 X. Li, Z.-Y. J. Zhan, R. Knipe, and D. G. Lynn, *J. Am. Chem. Soc.*, 2002, **124**, 746–747.
- 90 A. Khan, D. M. Haddleton, M. J. Hannon, D. Kukulj, and A. Marsh, *Macromolecules*, 1999, **32**, 6560–6564.
- 91 P. K. Lo and H. F. Sleiman, *J. Am. Chem. Soc.*, 2009, **131**, 4182–4183.
- 92 R. McHale, J. P. Patterson, P. B. Zetterlund, and R. K. O'Reilly, *Nat. Chem.*, 2012, **4**, 491–497.
- 93 Y. He and D. R. Liu, *Nat. Nanotechnol.*, 2010, **5**, 778–82.
- 94 S. Liao and N. C. Seeman, *Science*, 2004, **306**, 2072–4.
- 95 T. M. Snyder and D. R. Liu, *Angew. Chem. Int. Ed.*, 2005, **44**, 7379–82.
- 96 M. L. McKee, P. J. Milnes, J. Bath, E. Stulz, A. J. Turberfield, and R. K. O'Reilly, *Angew. Chemie Int. Ed.*, 2010, **49**, 7948–7951.

- 97 Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, and D. R. Liu, *Science*, 2004, **305**, 1601–5.
- 98 J. Niu, R. Hili, and D. R. Liu, *Nat. Chem.*, 2013, **5**, 282–292
- 99 M. Szwarc, *J. Polym. Sci.*, 1954, **13**, 317–318.
- 100 S. Połowiński, *Prog. Polym. Sci.*, 2002, **27**, 537–577.
- 101 J. Ferguson and S. A. O. Shah, *Eur. Polym. J.*, 1968, **4**, 343–354.
- 102 I. Rainaldi, C. Cristallini, G. Ciardelli, and P. Giusti, *Polym. Int.*, 2000, **49**, 63–73.
- 103 J. Ferguson, S. Al-Alawi, and R. Granmayeh, *Eur. Polym. J.*, 1983, **19**, 475–480.
- 104 Y.-J. Kim, H. Uyama, and S. Kobayashi, *Macromolecules*, 2003, **36**, 5058–5060.
- 105 N.-T. Lin, S.-Y. Lin, S.-L. Lee, C. Chen, C.-H. Hsu, L. P. Hwang, Z.-Y. Xie, C.-H. Chen, S.-L. Huang, and T.-Y. Luh, *Angew. Chem. Int. Ed.*, 2007, **46**, 4481–5.
- 106 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
- 107 R. L. Letsinger and K. K. Ogilvie, *J. Am. Chem. Soc.*, 1967, **89**, 4801–4803.
- 108 Y. Kang, A. Lu, A. Ellington, M. C. Jewett, and R. K. O'Reilly, *ACS Macro Lett.*, 2013, **2**, 581–586.
- 109 J. W. Kramer, D. S. Treitler, E. W. Dunn, P. M. Castro, T. Roisnel, C. M. Thomas, and G. W. Coates, *J. Am. Chem. Soc.*, 2009, **131**, 16042–16044.
- 110 Y. Hibi, S. Tokuoka, T. Terashima, M. Ouchi, and M. Sawamoto, *Polym. Chem.*, 2011, **2**, 341–347.
- 111 J. Li and J. He, *ACS Macro Lett.*, 2015, **4**, 372–376.
- 112 Y. Hibi, M. Ouchi, and M. Sawamoto, *Angew. Chemie Int. Ed.*, 2011, **50**, 7434–7437.
- 113 K. Satoh, M. Matsuda, K. Nagai, and M. Kamigaito, *J. Am. Chem. Soc.*, 2010, **132**, 10003–10005.
- 114 K. Satoh, S. Ozawa, M. Mizutani, K. Nagai, and M. Kamigaito, *Nat. Commun.*, 2010, **1**, 1–6.
- 115 J. Zhang, M. E. Matta, and M. A. Hillmyer, *ACS Macro Lett.*, 2012, **1**, 1383–1387.
- 116 S. Pfeifer and J.-F. Lutz, *J. Am. Chem. Soc.*, 2007, **129**, 9542–9543.
- 117 X. Tong, B. Guo, and Y. Huang, *Chem. Commun.*, 2011, **47**, 1455–1457.
- 118 G. Gody, T. Maschmeyer, P. B. Zetterlund, and S. Perrier, *Nat. Commun.*, 2013, **4**.

- 119 J. Vandenberg, G. Reekmans, P. Adriaenssens, and T. Junkers, *Chem. Commun.*, 2013, **49**, 10358.
- 120 S. Hecht and I. Huc, Eds., *Foldamers: Structure, Properties, and Applications*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2007, 1–108.
- 121 M. Goodman and I. Listowsky, *J. Am. Chem. Soc.*, 1962, **84**, 3770–3771.
- 122 J. Kovacs, R. Ballina, R. L. Rodin, D. Balasubramanian, and J. Applequist, *J. Am. Chem. Soc.*, 1965, **87**, 119–120.
- 123 J. M. Fernández-Santín, J. Aymamí, A. Rodríguez-Galán, S. Muñoz-Guerra, and J. A. Subirana, *Nature*, 1984, **311**, 53–54.
- 124 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071–13072.
- 125 Y. Hamuro, S. J. Geib, and A. D. Hamilton, *Angew. Chemie Int. Ed.*, 1994, **33**, 446–448.
- 126 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071–13072.
- 127 V. Berl, I. Huc, R. G. Khoury, M. J. Krische, and J. M. Lehn, *Nature*, 2000, **407**, 720–723.
- 128 H. Jiang, C. Dolain, J.-M. Léger, H. Gornitzka, and I. Huc, *J. Am. Chem. Soc.*, 2004, **126**, 1034–1035.
- 129 C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, and I. Huc, *J. Am. Chem. Soc.*, 2005, **127**, 12943–12951.
- 130 A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, and I. Huc, *J. Org. Chem.*, 2011, **76**, 195–200.
- 131 V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, and I. Huc, *J. Am. Chem. Soc.*, 2004, **126**, 10049–10052.
- 132 C. Bao, Q. Gan, B. Kauffmann, H. Jiang, and I. Huc, *Chem. - A Eur. J.*, 2009, **15**, 11530–11536.
- 133 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, and I. Huc, *Nat. Chem.*, 2015, **7**, 334–341.
- 134 Q. Gan, Y. Ferrand, C. Bao, B. Kauffmann, A. Grélard, H. Jiang, and I. Huc, *Science*, 2011, **331**, 1172–5.
- 135 L. Yuan, H. Zeng, K. Yamato, A. R. Sanford, W. Feng, H. S. Atreya, D. K. Sukumaran, T. Szyperski, and B. Gong, *J. Am. Chem. Soc.*, 2004, **126**, 16528–37.

- 136 R. D. Parra, H. Zeng, J. Zhu, C. Zheng, X. C. Zeng, and B. Gong, *Chem. - A Eur. J.*, 2001, **7**, 4352–4357.
- 137 X. Yang, S. Martinovic, R. D. Smith, and B. Gong, *J. Am. Chem. Soc.*, 2003, **125**, 9932–3.
- 138 B. Gong, *Acc. Chem. Res.*, 2008, **41**, 1376–86.
- 139 B. Gong, H. Zeng, J. Zhu, L. Yua, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, E. Skrzypczak-Jankun, S. Martinovic, R. D. Smith, C. Zheng, T. Szyperski, and X. C. Zeng, *Proc. Natl. Acad. Sci.*, 2002, **99**, 11583–11588.
- 140 K. Yamato, L. Yuan, W. Feng, A. J. Helsel, A. R. Sanford, J. Zhu, J. Deng, X. C. Zeng, and B. Gong, *Org. Biomol. Chem.*, 2009, **7**, 3643.
- 141 J. C. Nelson, J. G. Saven, J. S. Moore, and P. G. Wolynes, *Science (80-.)*, 1997, **277**, 1793–1796.
- 142 R. B. Prince, J. G. Saven, P. G. Wolynes, and J. S. Moore, *J. Am. Chem. Soc.*, 1999, **121**, 3114–3121.
- 143 R. B. Prince, S. A. Barnes, and J. S. Moore, *J. Am. Chem. Soc.*, 2000, **122**, 2758–2762.
- 144 A. Tanatani, M. J. Mio, and J. S. Moore, *J. Am. Chem. Soc.*, 2001, **123**, 1792–1793.
- 145 A. Tanatani, T. S. Hughes, and J. S. Moore, *Angew. Chemie*, 2002, **114**, 335–338.
- 146 K. Oh, K.-S. Jeong, and J. S. Moore, *Nature*, 2001, **414**, 889–893.
- 147 T. Nishinaga, A. Tanatani, K. Oh, and J. S. Moore, *J. Am. Chem. Soc.*, 2002, **124**, 5934–5935.
- 148 L. Sebaoun, V. Maurizot, T. Granier, B. Kauffmann, and I. Huc, *J. Am. Chem. Soc.*, 2014, **136**, 2168–2174.
- 149 Z. Chen, N. D. Urban, Y. Gao, W. Zhang, J. Deng, J. Zhu, X. C. Zeng, and B. Gong, *Org. Lett.*, 2011, **13**, 4008–4011.
- 150 R. Kramer, J. M. Lehn, and A. Marquis-Rigault, *Proc. Natl. Acad. Sci.*, 1993, **90**, 5394–5398.
- 151 E. A. Archer and M. J. Krische, *J. Am. Chem. Soc.*, 2002, **124**, 5074–5083.
- 152 H. Gong and M. J. Krische, *J. Am. Chem. Soc.*, 2005, **127**, 1719–1725.
- 153 A. P. Bisson, F. J. Carver, C. A. Hunter, and J. P. Waltho, *J. Am. Chem. Soc.*, 1994, **116**, 10292–10293.

- 154 A. P. Bisson, F. J. Carver, D. S. Eggleston, R. C. Haltiwanger, C. A. Hunter, D. L. Livingstone, J. F. McCabe, C. Rotger, and A. E. Rowan, *J. Am. Chem. Soc.*, 2000, 122, 8856–8868.
- 155 C. A. Hunter, P. S. Jones, P. M. N. Tiger, and S. Tomas, *Chem. Commun. (Camb)*., 2003, 1642–1643.
- 156 J. Li, J. a Wisner, and M. C. Jennings, *Org. Lett.*, 2007, 9, 3267–3269.
- 157 W. L. Jorgensen and J. Pranata, *J. Am. Chem. Soc.*, 1990, **112**, 2008–2010.
- 158 H.-B. Wang, B. P. Mudraboyina, J. Li, and J. A. Wisner, *Chem. Commun. (Camb)*., 2010, 46, 7343–7345.
- 159 J. Li, A. T. Pandelieva, C. N. Rowley, T. K. Woo, and J. A. Wisner, *Org. Lett.*, 2012, 14, 5772–5775.
- 160 H. L. Anderson, *Inorg. Chem.*, 1994, 33, 972–981.
- 161 P. N. Taylor and H. L. Anderson, *J. Am. Chem. Soc.*, 1999, 121, 11538–11545.
- 162 H. Sugiura, Y. Nigorikawa, Y. Saiki, K. Nakamura, and M. Yamaguchi, *J. Am. Chem. Soc.*, 2004, 126, 14858–14864.
- 163 R. Amemiya, N. Saito, and M. Yamaguchi, *J. Org. Chem.*, 2008, 73, 7137–7144.
- 164 N. Saito, R. Terakawa, M. Shigeno, R. Amemiya, and M. Yamaguchi, *J. Org. Chem.*, 2011, 76, 4841–4858.
- 165 Y. Tanaka, H. Katagiri, Y. Furusho, and E. Yashima, *Angew. Chem. Int. Ed.*, 2005, 44, 3867–3870.
- 166 Z.-Q. Wu, Y. Furusho, H. Yamada, and E. Yashima, *Chem. Commun. (Camb)*., 2010, 46, 8962–8964.
- 167 H. Ito, Y. Furusho, T. Hasegawa, and E. Yashima, *J. Am. Chem. Soc.*, 2008, 130, 14008–14015.
- 168 H. Iida, M. Shimoyama, Y. Furusho, and E. Yashima, *J. Org. Chem.*, 2010, 75, 417–423.
- 169 T. Maeda, Y. Furusho, S. Sakurai, J. Kumaki, K. Okoshi, and E. Yashima, *J. Am. Chem. Soc.*, 2008, 130, 7938–7945.
- 170 M. Ikeda, Y. Tanaka, T. Hasegawa, Y. Furusho, and E. Yashima, *J. Am. Chem. Soc.*, 2006, 128, 6806–6807.
- 171 T. Hasegawa, Y. Furusho, H. Katagiri, and E. Yashima, *Angew. Chem. Int. Ed.*, 2007, 46, 5885–5888.

-
- 172 H. Ito, M. Ikeda, T. Hasegawa, Y. Furusho, and E. Yashima, *J. Am. Chem. Soc.*, 2011, 133, 3419–3432.
- 173 H. Yamada, Y. Furusho, H. Ito, and E. Yashima, *Chem. Commun. (Camb)*, 2010, 46, 3487–3489.
- 174 H. Yamada, Y. Furusho, and E. Yashima, *J. Am. Chem. Soc.*, 2012, 134, 7250–7253.
- 175 J. Tanabe, D. Taura, H. Yamada, Y. Furusho, and E. Yashima, *Chem. Sci.*, 2013, 4, 2960–2966.
- 176 H. Zeng, R. S. Miller, R. A. Flowers, and B. Gong, *J. Am. Chem. Soc.*, 2000, 122, 2635–2644.
- 177 B. Gong, Kim, J. Zhu, and H. Ickes, *J. Am. Chem. Soc.*, 1999, 121, 5607–5608.
- 178 H. Zeng, H. Ickes, R. A. Flowers, and B. Gong, *J. Org. Chem.*, 2001, 66, 3574–3583.
- 179 P. H. Zhang, H. Z. Chu, X. H. Li, W. Feng, P. C. Deng, L. H. Yuan, and B. Gong, *Org. Lett.*, 2011, 13, 54–57

Chapter 2

Aims

The aim of this work is to develop a synthetic programmable information oligomer which rivals the nucleic acids in terms of function, but is orthogonal in terms of chemical structure, the solvent in which it operates and the information molecules it can interact with. The principle functions of nucleic acids are their ability to store and express information through sequence selective duplex formation and template directed synthesis. In the same manner as the nucleic acids, the information oligomer sought in this work also incorporates information content in a programmed array of recognition groups along an oligomer. However, duplex assembly between strands will be driven by the cooperative interaction of many high affinity single point H-bonds between donor and acceptor groups (red and blue bars in Figure 2.1).

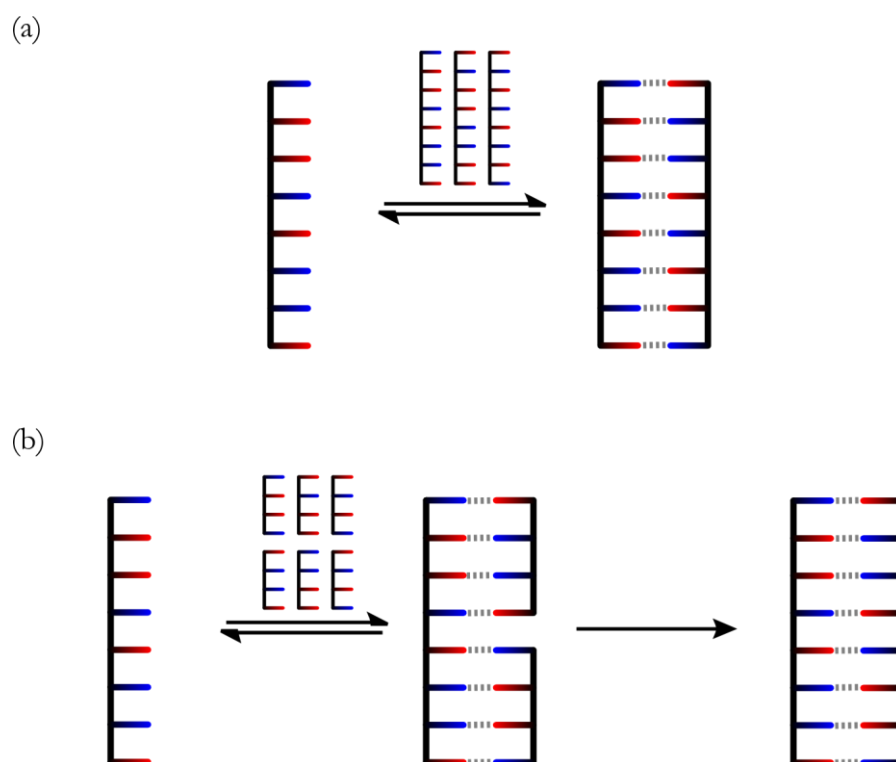


Figure 2.1. Schematic representations of the two key functions of idealised H-bonding synthetic information oligomers where the red and blue bars represent single H-bond donor and acceptor groups: (a) sequence selective duplex formation; (b) template directed synthesis.

The chemical architecture of the oligomer, and the solvents in which it operates, will be non-polar and organic, such that nothing will effectively compete with the H-bonding recognition.

The first aim is to identify a chemical architecture and a set of H-bonding recognition groups that lead to oligomers that form duplexes. This will involve synthesising a candidate system and measuring the extent of duplex formation by NMR titrations. To begin, no complex sequences will be studied, but rather oligomers which only have either donor or acceptor groups. Once cooperative duplex formation has been established, oligomers with simple sequences that bear both H-bond donor and acceptor groups will be made in order to investigate the ability of single oligomers to fold in an intramolecular fashion.

Once a system is identified, that forms cooperative duplexes and has a low propensity to fold, more complicated sequences will be synthesised and studied via NMR titrations in order to investigate the ability of the oligomers to discriminate based upon sequence when forming duplexes.

Chapter 3

Single sequence H-bonding information oligomers

3.1 Introduction

In the pursuit of a general solution to supramolecular problems, chemists often turn to programmable oligomers to form the basis of their approach to the construction of static nanoobjects,¹⁻⁷ dynamic devices,⁸⁻¹⁴ and in template directed synthesis of small molecules.¹⁵⁻¹⁹ Truly versatile programmable oligomers remain the domain of biology, in the form of polypeptides, polysaccharides and oligonucleotides. In nucleic acids, Nature expresses the chemical information contained within the sequence of nucleobases through sequence-specific duplex formation and template-directed synthesis.²⁰ It is fascinating to consider how Nature arrived at the precise structure of the nucleic acid, since duplex formation between oligonucleotides is surprisingly robust with respect to structural re-engineering: the integrity of the double helix is maintained when the furanose moieties are replaced by different sugars,²¹⁻²⁴ when the entire backbone is replaced by a peptide chain,²⁵⁻³⁰ and when the base-pairs are replaced by non-natural analogues.³¹⁻³³ The success of genetic mimetics demonstrates that other programmable materials, with very different chemical structures from nucleic acids, are possible. These synthetic information molecules would be orthogonal to Nature's information-containing oligomers and offer the exciting prospect of materials that resemble traditional plastics in chemical structure but can be programmed to develop functionality through forced evolution. As a first step towards the development of synthetic information molecules that form well-defined sequence-selective duplexes,³⁴ this chapter describes the synthesis and characterisation of a new class of oligomeric molecules, which form stable double-stranded complexes in solution.

The common characteristic of duplex-forming natural and artificial systems is cooperative non-covalent interaction between multiple sites along oligomer strands to form a closed complex (Figure 3.1). Such multivalent complexes display all-or-nothing behaviour, where partially bound states are not significantly populated, and the stability increases relative to the individual single point interactions.^{35,36} Macroscopically this leads to sharp transitions from mainly unbound complex to mainly bound states over a small range of conditions.^{37,38} More importantly, only duplexes formed of oligomers which are complementary in sequence and length predominate, which accounts for the special role the nucleic acids play in supramolecular chemistry. The synergistic behaviour displayed by multivalent systems is called chelate cooperativity and is quantified by the dimensionless product of σKEM , the equilibrium constant for the intramolecular binding event in Figure 3.1.³⁵ The value of K is the association constant for the corresponding single point intermolecular interaction, σ is a statistical factor that accounts for the difference in the degeneracies of the complexes, and EM is the effective molarity for the intramolecular interaction.³⁹ If $KEM \gg 1$ for all steps in multivalent complex formation, the fully bound species predominates. Therefore, KEM is a good parameter to judge whether oligomeric molecules are likely to form sequence-selective duplexes.

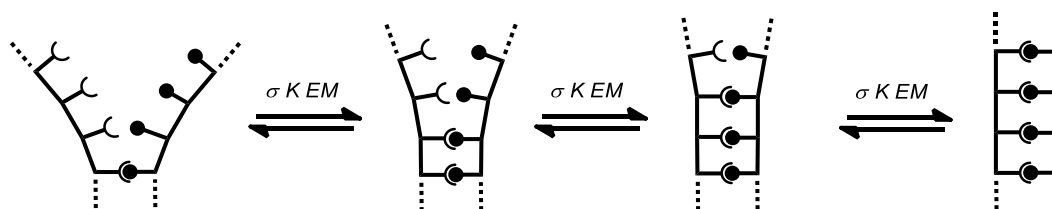


Figure 3.1. Schematic representation of multivalent duplex formation where each microscopic interaction has an association constant of σKEM .

In this chapter, we present the approach, design and synthesis of H-bonding oligomers orthogonal to nature's oligonucleotides. In order to quantify the oligomers potential as a supramolecular building block we measure the *KEM* of the H-bonding duplexes formed between them and describe their denaturation profiles. We also demonstrate that different sequences of H-bonding groups can be incorporated into the oligomers.

3.2 Approach

3.2.1 Functional design

Sequence selective duplex forming information oligomers all share the same functional design. As in DNA, there are four key functions to satisfy (Figure 3.2): non-covalent *recognition* between strands; a group to control *solubility*; efficient *coupling-chemistry* to synthesise the oligomers; and a *backbone* to connect all the sub-units together.

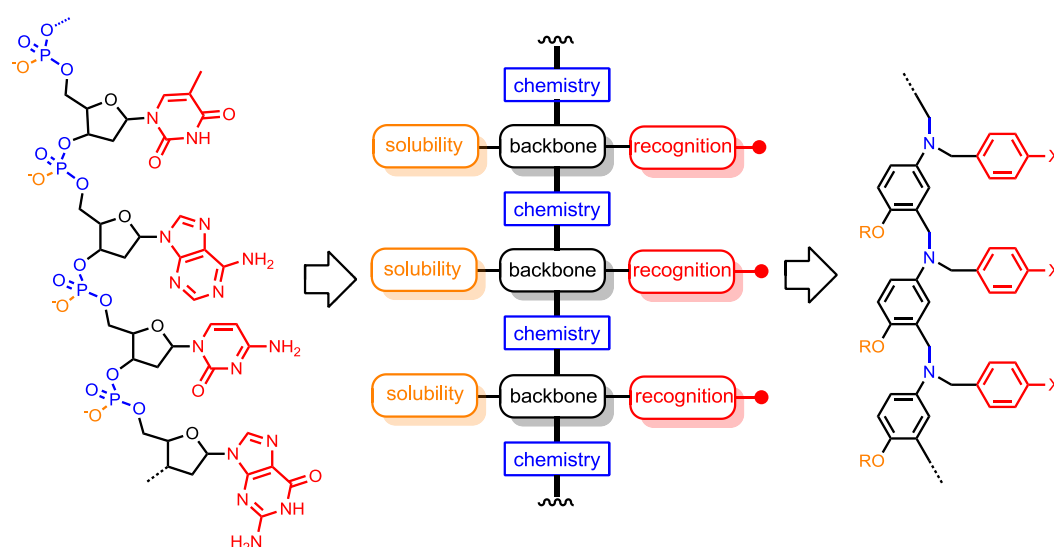


Figure 3.2. Schematic showing the structure and function of each part of a nucleic acid oligomer and how the functional design of an information oligomer can be derived using modules which perform each of the same four functions: (i) linking chemistry to join monomer units together (blue); (ii) recognition sites responsible for sequence selectivity in duplex formation (red); (iii) a control over solubility (orange) and (iv) a backbone for attachment of all the other functions.

Our design is comprised of an aniline oligomer, with branched alkoxy side-chains and recognition based on single point hydrogen bonds (given by X in Figures 3.2 and 3.3). DNA operates in water and forms duplexes through stacking interactions between the four hydrophobic bases.^{40,41} We have opted for duplexes formed through H-bonds in non-polar organic solvent. The backbone and coupling

chemistry are chosen to ensure no strong H-bond donor or acceptor sites are present anywhere on the oligomer other than on the recognition motif. The free energy gain of a H-bond interaction in solution can be accurately predicted using Equation 3.1,⁴² where α and β describe the H-bond donor and acceptor properties of two interacting functional groups, respectively. The H-bond properties of the solvent are given by α_s and β_s and the value 6 kJ mol^{-1} is the free energy cost of the bimolecular association.

$$-RT\ln K = \Delta G^\circ = -(\alpha - \alpha_s)(\beta - \beta_s) + 6 \text{ kJ mol}^{-1} \quad (3.1)$$

Table 3.1 shows the predicted pairwise association constants of interactions between functional groups, in a range of polar to non-polar solvents, derived using Equation 3.1. Phenol is considered an ideal candidate for the H-bond donor, since it forms strong interactions with a range of possible H-bond acceptors (Table 3.1, entries 1-3) but weak interactions with the functional groups elsewhere in the oligomer (Table 3.1, entries 4-7).

Table 3.1. Predicted association constants of potential H-bonding interactions between functional groups calculated using equation (3) in toluene ($\alpha_s = 1.0$, $\beta_s = 2.2$)^a, CHCl₃ ($\alpha_s = 2.2$, $\beta_s = 0.8$)^b and DMSO ($\alpha_s = 0.8$, $\beta_s = 8.9$)^b.

| Entry | Donor | α | Acceptor | β | K / M ⁻¹ | | |
|-------|------------------|------------------|-------------------------------------|-------------------|---------------------|-------------------|------|
| | | | | | toluene | CHCl ₃ | DMSO |
| 1 | <i>p</i> -cresol | 3.7 ^c | tri- <i>n</i> -butylphosphine oxide | 10.2 ^d | 540 | 26 | <1 |
| 2 | <i>p</i> -cresol | 3.7 ^c | 4-methylpyridine | 7.7 ^d | 36 | 5.8 | <1 |
| 3 | <i>p</i> -cresol | 3.7 ^c | pyridine- <i>N</i> -oxide | 9.0 ^d | 150 | 13 | <1 |
| 4 | <i>p</i> -cresol | 3.7 ^c | <i>N,N</i> -dimethylaniline | 4.2 ^d | <1 | <1 | <1 |
| 5 | <i>p</i> -cresol | 3.7 ^c | anisole | 3.3 ^d | <1 | <1 | <1 |
| 6 | <i>p</i> -cresol | 3.7 ^c | nitrobenzene | 3.7 ^d | <1 | <1 | <1 |
| 7 | <i>p</i> -cresol | 3.7 ^c | 2-methoxybenzaldehyde | 5.5 ^d | 3.2 | 1.5 | <1 |
| 8 | aldehyde | 1.6 ^a | tri- <i>n</i> -butylphosphine oxide | 10.2 ^d | <1 | <1 | <1 |
| 9 | aldehyde | 1.6 ^a | 4-methylpyridine | 7.7 ^d | <1 | <1 | <1 |
| 10 | aldehyde | 1.6 ^a | pyridine- <i>N</i> -oxide | 9.0 ^d | <1 | <1 | <1 |

^a values taken from reference 43; ^b values taken from reference 42; ^c value taken from reference 44; ^d values taken from reference 45.

The predicted association constants in Table 3.1 also show that the interaction between the functional groups associated with the recognition motif (Table 3.1, entries 1-3) are only formed to a significant extent in a non-polar solvent such as toluene. The racemic branched 2-ethyl hexyl alkoxy side chain is therefore important to promote solubility in toluene by disrupting formation of a stable lattice.

Figure 3.3 shows the retrosynthesis of the target oligomer. The secondary aniline and benzaldehyde groups, which react via reductive amination, can be protected using standard protecting group chemistry allowing the chain to propagate in either direction via coupling-deprotection steps. Reductive amination has been shown to produce relatively high yields, and when NaBH(OAc)₃ is used, the

iminium ion or imine intermediate can be selectively reduced without significant reduction of the aldehyde.^{46,47,†}

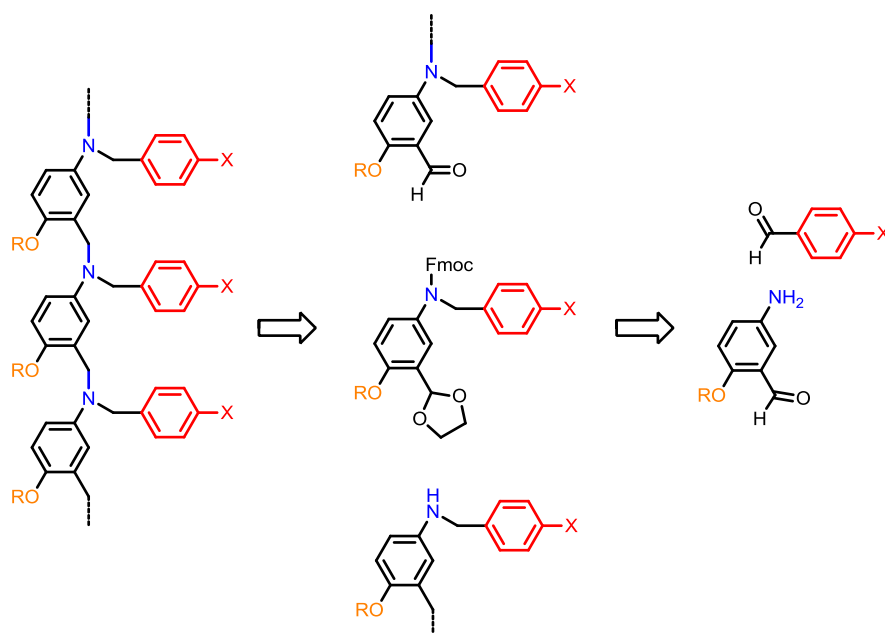


Figure 3.3. Retrosynthesis showing how reductive amination can be used to synthesise the oligomers from basic building blocks equipped with orthogonal protecting groups (R = 2-ethylhexyl, X = H-bond donor or H-bond acceptor).

[†] The system described in this Section is not the first design attempted within the Hunter group. The results of previous designs are summarised on pages 94 and 95.

3.2.2 Duplex formation and $K_{ref} \overline{EM}_N$

Figure 3.4(a) shows the equilibria of the stepwise duplex formation for two oligomers each bearing four recognition modules, represented as D for H-bond donors and A for H-bond acceptors. The first step is formation of a single intermolecular H-bond between the oligomers, with association constant $\sigma_1 K_{ref}$ where σ_1 is the statistical factor for forming the 1st H-bond and K_{ref} is the association constant of the reference single point interaction of A and D. Each subsequent step involves intramolecular H-bond formation, and the association constant for the N th H-bond formation is $\sigma_N K_{ref} EM_N$, where σ_N is the statistical factor and EM_N is the stepwise effective molarity. The observed association constant for the complex between two oligomers with N interaction sites K_N is the product of the association constants associated with the step-wise formation of the duplex and can be expressed in terms of $\sigma_{D,N}$, K_{ref} and \overline{EM}_N using Equation 3.2, where \overline{EM}_N is the average effective molarity for the formation of N H-bonds in the duplex and $\sigma_{D,N}$ is the product of all the stepwise statistical factors.

$$K_N = \sigma_{D,N} K_{ref}^N \overline{EM}_N^{(N-1)} \quad (3.2)$$

Rearranging Equation 3.2 allows the statistically corrected value of \overline{EM}_N to be expressed in terms of K_N , K_{ref} and $\sigma_{D,N}$, where $\sigma_{D,N} = 2$ for all values of N , which gives Equation 3.3.

$$\overline{EM}_N = \sqrt[N-1]{\frac{K_N}{2K_{ref}^N}} \quad (3.3)$$

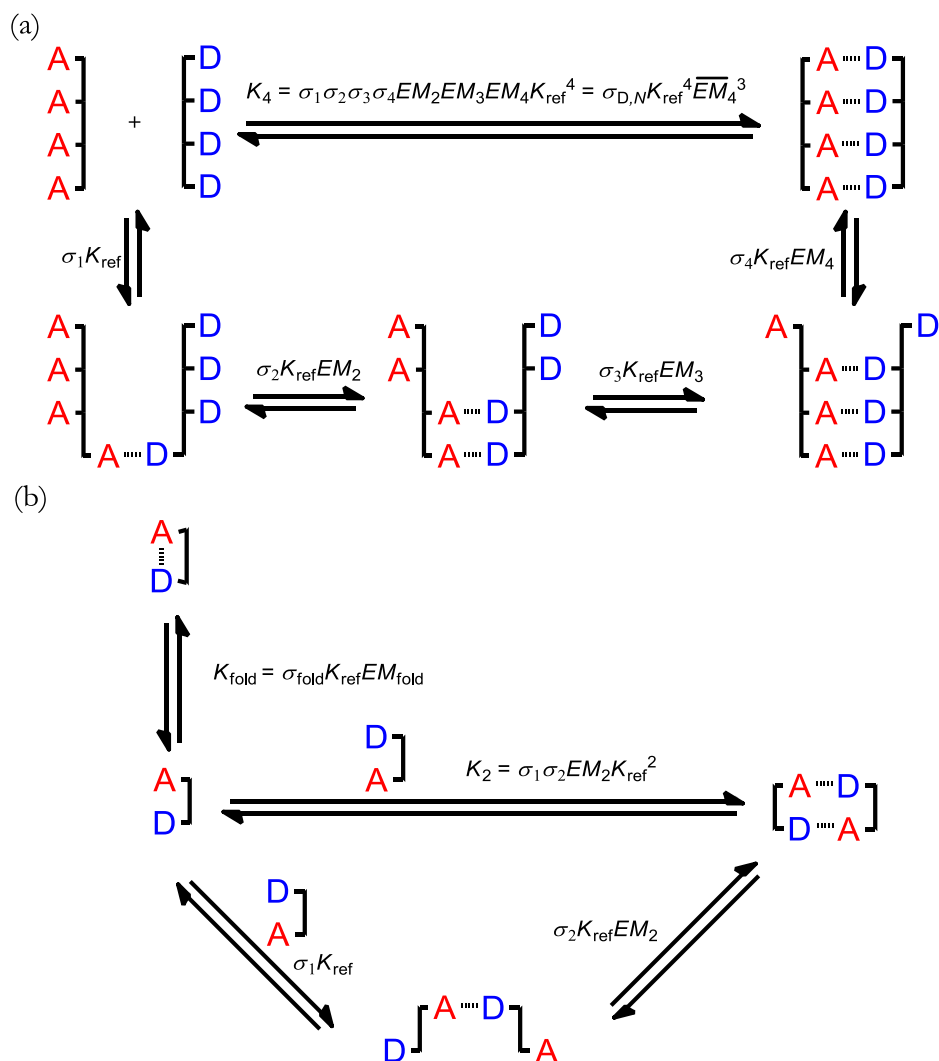


Figure 3.4. Equilibria involved in duplex formation: (a) two oligomers with four H-bonding sites each. The macroscopic association constant, K_4 , can be described as the product of all the microscopic association constants, which themselves can be described using K_{ref} , EM_N and σ_N ; (b) oligomers with a neighbouring H-bond acceptor (A) and donor (D) can form an intramolecular H-bond, which competes with duplex formation, where the equilibrium constant of folding K_{fold} can be expressed by the product of the statistical factors, K_{ref} and the effective molarity of folding EM_{fold} .

A molecule bearing an H-bond acceptor and an H-bond donor has the potential to form an intramolecular H-bond, the association constant of which is determined by the $K_{\text{ref}}EM$ of that interaction. Therefore, duplex formation between sequence programmable oligomers bearing H-bonding donor and acceptor groups will have competing unimolecular folding processes. Figure 3.4(b) shows the equilibria

involved with duplex formation and folding of a 2mer with both an H-bond donor and acceptor. Duplex formation competes with folding between the two H-bonding groups, where the association constant of folding K_{fold} , can be expressed in terms of K_{ref} and the effective molarity of folding EM_{fold} .

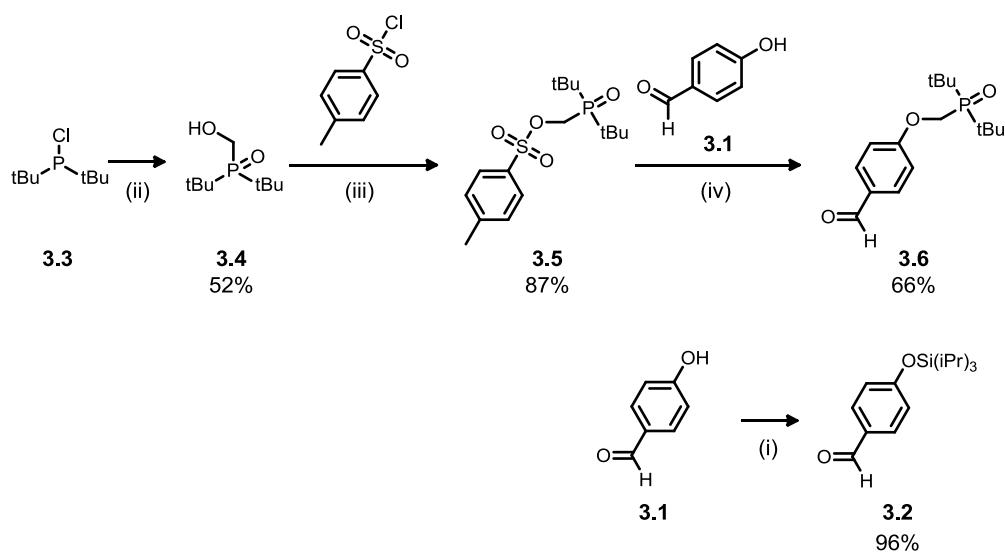
To ensure $K_{\text{ref}}EM > 1$ is satisfied for an intramolecular binding process one can attempt to increase EM or K_{ref} . Amongst other factors, the EM of a binding process is sensitive to the degree of preorganization and structural complementarity.⁴⁸ Designing a system with a high degree of preorganisation and structural complementarity can be far from trivial, as rigid systems can just as easily lead to very low values of EM which can hinder binding (i.e. $K_{\text{ref}}EM \ll 1$). Previous work in the group has shown that the decrease in EM resulting from an increase in conformational flexibility is not dramatic, and the EM for highly flexible systems can be as high as 0.1 M.⁴⁸ In comparison to EM , K_{ref} is much easier to predict and control, so assuming that $EM \approx 0.1$ M, then if $K_{\text{ref}} \geq 100 \text{ M}^{-1}$ $K_{\text{ref}}EM$ will be significantly greater than 1. Therefore, our design focuses on the appropriate choice of high affinity single point H-bonding recognition groups and a flexible backbone that is free from any strong H-bonding functional groups, rather than focusing on designing a system that is highly preorganised.

3.3 Results and discussion

3.3.1 Synthesis

3.3.1.1 Recognition units

All the recognition units are based on derivatives of 4-hydroxybenzaldehyde **3.1** (Scheme 3.1). The phenol H-bond acceptor was protected as silyl ether **3.2** for the synthesis of the oligomers.⁴⁹ Phosphine oxide **3.6** was made via the oxidation of **3.3** in the presence of formaldehyde to give alcohol **3.4**, which was reacted with *p*-toluenesulfonyl chloride in the presence of base to give the toluenesulfonate ester **3.5**. Finally, **3.1** was alkylated with **3.5** in the presence of base to yield **3.6**. Phosphine oxide **3.24** (shown in Figure 3.5, below) was synthesised in the same way as **3.6** using **3.5**.

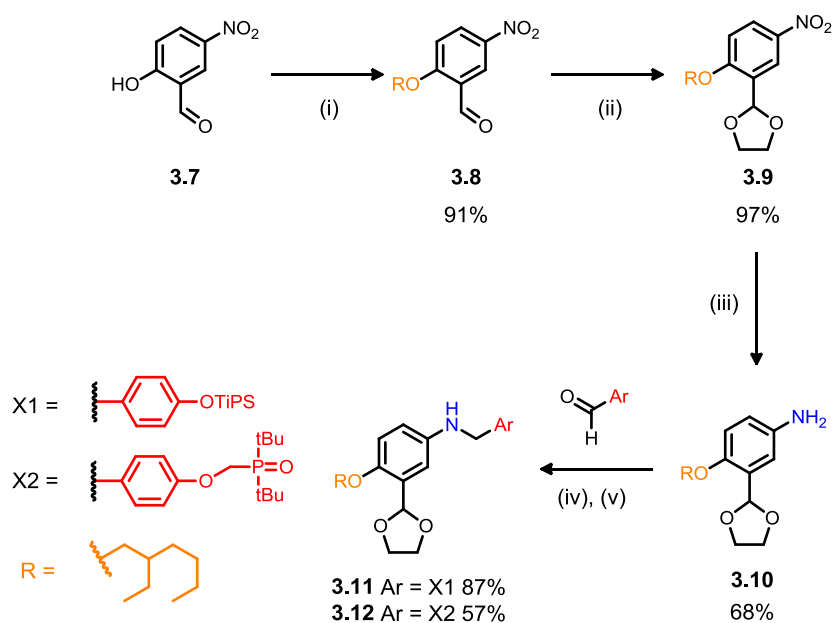


Scheme 3.1.

Reagents and conditions: (i) triisopropylsilyl chloride, imidazole, DMF; (ii) formaldehyde, aqueous HCl; (iii) *p*-toluenesulfonyl chloride, NEt₃, THF; (iv) Cs₂CO₃, DMF.

3.3.1.2 Monomers

Monomer units **3.11** and **3.12** were each prepared according to Scheme 3.2. Phenol **3.7** was alkylated with 2-ethylhexyl bromide, which gave **3.8**, and then the benzaldehyde was protected as the corresponding acetal by condensation with ethylene glycol **3.9**. Primary aniline **3.10** was formed by reducing **3.9** with H₂ in the presence of Pd/C catalyst. Monomer units **3.11** and **3.12** were each prepared stepwise by reduction of the preformed imines of primary aniline **3.10** with **3.2** and **3.6**, respectively.

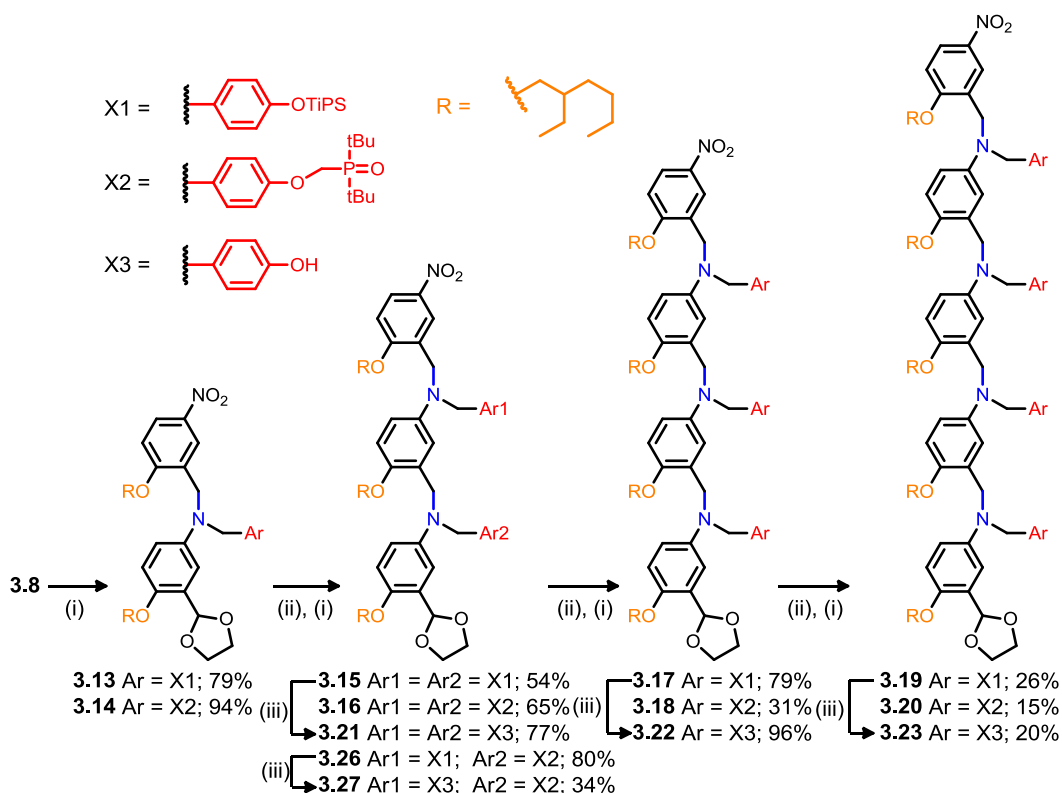


Scheme 3.2.

Reagents and conditions: (i) RBr, K₂CO₃, DMF; (ii) *p*-toluenesulfonic acid, ethylene glycol, toluene; (iii) H₂ Pd/C, EtOAc; (iv) CHCl₃ or toluene; (v) NaBH₄, MeOH.

3.3.1.3 Oligomers

Benzaldehyde **3.8** was coupled with one of each monomer unit via reductive amination using $\text{NaBH}(\text{OAc})_3$ as the reducing agent to give 1-mer chains **3.13** and **3.14**. The acetal group was then deprotected in the presence of aqueous HCl , and the product was coupled with a further monomer unit to give 2-mers **3.15** and **3.16**. In the same fashion 3-mers (**3.17** and **3.18**) and 4-mers (**3.19** and **3.20**) were synthesised. Purification by column chromatography was necessary after all reductive amination steps but not after acetal deprotection (Scheme 3.3). Phenol bearing oligomers **3.21**, **3.22** and **3.23** were prepared from the corresponding silyl ether oligomer by deprotection using tetra-*n*-butylammonium fluoride (TBAF).



Scheme 3.3.

Reagents and conditions: (i) Monomer (**3.11** or **3.12**), $\text{NaBH}(\text{OAc})_3$, CHCl_3 or DCE; (ii) aqueous HCl , CHCl_3 (iii) TBAF, THF.

3.3.2 Binding Studies

Complexation of length-complementary oligomers was measured using ^1H and ^{31}P NMR titration experiments in toluene- d_8 and all data were fitted to a 1:1 binding isotherm. The value of K_{ref} was determined by titration of **3.24** into **3.25** (Figure 3.5). Figures 3.6 – 3.9 show the spectra for titrations of duplexes **3.25•3.24**, **3.21•3.16**, **3.22•3.18** and **3.23•3.20** and the fitting of the changes in chemical shift to a 1:1 binding isotherm. The association constants derived are shown in Table 3.2, along with the corresponding values of $K_{\text{ref}}EM$, calculated using Equations 3.2 and 3.3. For **3.21•3.16**, **3.22•3.18** and **3.23•3.20** the phosphine oxide was used as host. ^{31}P NMR spectroscopy has the advantage that the spectra are much simpler and the limiting complexation-induced changes in chemical shift are over an order of magnitude greater than the changes observed in ^1H NMR spectra. For **3.21•3.16** the ^{31}P NMR signals overlap and are only resolved at the end of the titration, so the signals were fitted as one signal, recording the chemical shift at the highest point of the peak or the average of the two chemical shifts when two peaks could be seen. In the case of the **3.22•3.18** and **3.23•3.20** duplexes, all three ^{31}P signals began to broaden as guest was added, and then became sharp again towards the end of the titration, but all could be fit separately. For **3.23•3.20**, due to the low host concentration and the broadening of the signals, a very high line broadening was applied when processing the spectra (40 Hz), and three of the ^{31}P signals were fit as one signal, using the chemical shift recorded at the highest point of the peak.

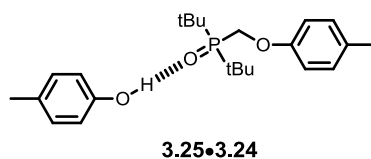


Figure 3.5. Complex studied in order to obtain a value of K_{ref} for the phenol-phosphine oxide interaction.

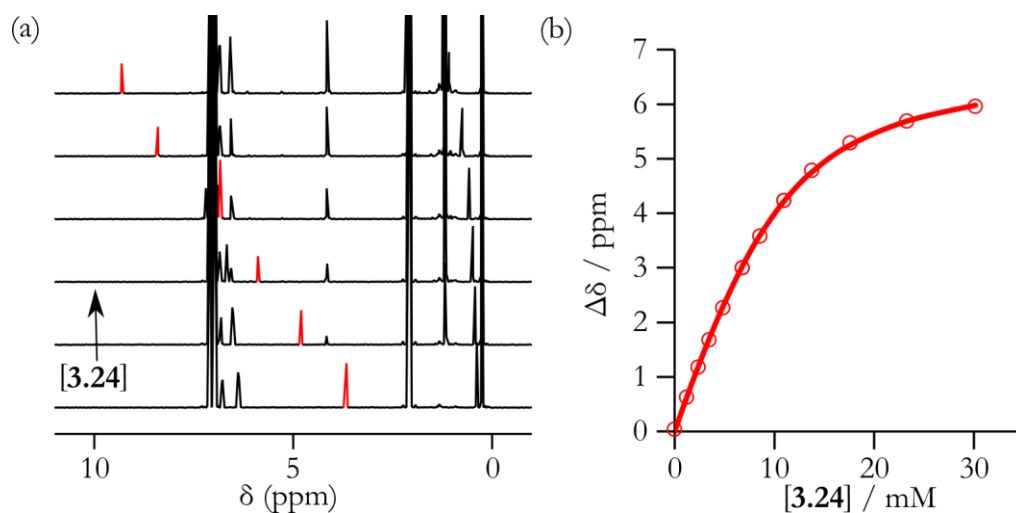


Figure 3.6. ^1H NMR data for titration of **3.24** into **3.25** (10 mM) at 298 K in toluene- d_8 . (a) Example 400 MHz ^1H NMR spectra (OH signal highlighted in red). (b) Plot of the change in chemical shift of the OH signal as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm).

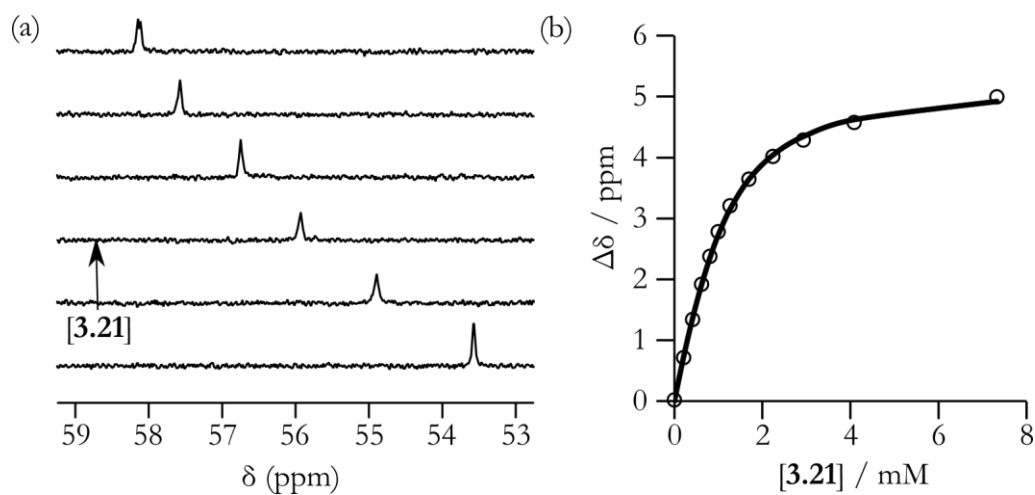


Figure 3.7. ^{31}P NMR data for titration of **3.21** into **3.16** (2.0 mM) at 298 K in toluene- d_8 . (a) Example 162 MHz ^{31}P NMR spectra. (b) Plot of the change in chemical shift of the ^{31}P signal as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm).

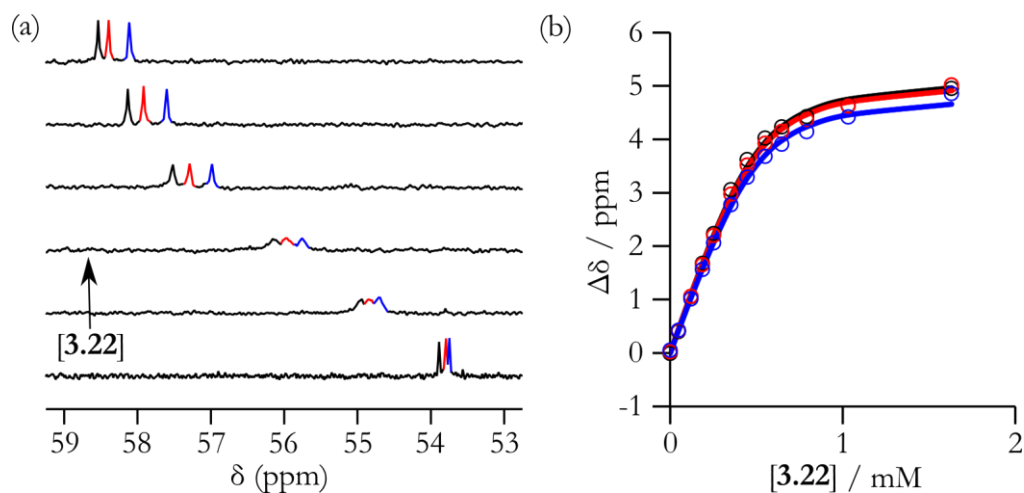


Figure 3.8. ^{31}P NMR data for titration of **3.22** into **3.18** (0.5 mM) at 298 K in toluene- d_8 . (a) Example 162 MHz ^{31}P NMR spectra. (b) Plot of the change in chemical shift of the ^{31}P signals as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm).

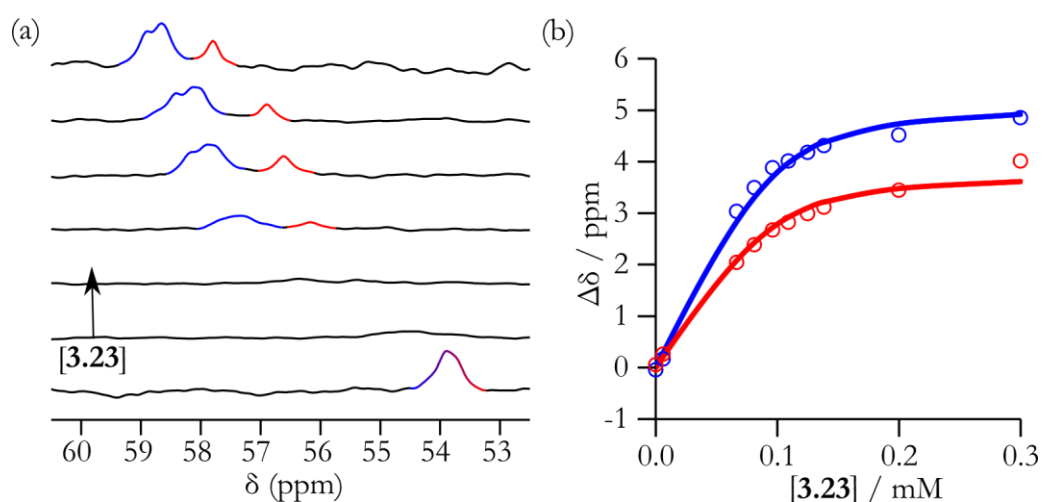


Figure 3.9. ^{31}P NMR data for titration of **3.23** into **3.20** (0.1 mM) at 298 K in toluene- d_8 . (a) Example 162 MHz ^{31}P NMR spectra. (b) Plot of the change in chemical shift of the ^{31}P signals as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm). Due to the low host concentration and peak broadening throughout the titration a high line broadening (40 Hz) was used to monitor three of the four ^{31}P chemical shifts as one broad peak.

Table 3.2. Association constants and values of EM_N and $K_{\text{ref}}EM_N$ for formation of 1:1 duplexes between phosphine oxide and phenol oligomers based on ^{31}P NMR titrations in toluene- d_8 at 298 K.^a

| Complex | K_N (M^{-1}) | EM_N / mM | $K_{\text{ref}} EM$ |
|---|---------------------------|-------------|---------------------|
| -A=D- 3.24•3.25 | 3.5×10^2 (1%) | - | |
| $\left[\begin{array}{c} \text{A=D} \\ \text{A=D} \end{array} \right]$ 3.16•3.21 | 1.9×10^3 (33%) | 8.1 (33%) | 2.8 (12%) |
| $\left[\begin{array}{c} \text{A=D} \\ \text{A=D} \\ \text{A=D} \end{array} \right]$ 3.18•3.22 | 1.7×10^4 (15%) | 14 (8%) | 5.0 (2%) |
| $\left[\begin{array}{c} \text{A=D} \\ \text{A=D} \\ \text{A=D} \\ \text{A=D} \end{array} \right]$ 3.20•3.23 | 2.6×10^5 (117%) | 21 (39%) | 7.2 (6%) |

^a Each titration was repeated twice and the average value is reported with errors at the 95% confidence limit (percentage error reported in brackets).

Tables 3.3-3.5 show the limiting complexation-induced changes in ^{31}P chemical shift ($\Delta\delta$) for oligomer duplexes **3.16•3.21**, **3.18•3.22** and **3.20•3.23**, and Figure 3.10 gives the proton labelling scheme. The ^{31}P NMR chemical shift of the free and bound signals for all duplexes are similar, as are the complexation-induced changes in chemical shift (4-5 ppm). This shows that all of the phosphine oxide groups are fully H-bonded in all of the duplexes. Although many ^1H NMR signals could not be followed during the titration of the oligomers, there was a clear pattern in the ^1H NMR complexation-induced chemical shift changes (Tables 3.6-3.8). Signals Aa6, Ab6 and Ac6 all display a small positive complexation-induced chemical shift changes (+0.2 - +0.1 ppm). The CH_2 protons on the backbone can be separated into two classes; those associated with the recognition motif (Aa5, Ab5, Ac5 and Ad5) and those associated with the backbone (X7, a7, b7 and c7). The former all have

small negative changes (-0.1 ppm) and the latter small positive changes (+0.1 ppm). In the case of **3.24•3.25**, a large positive complexation-induced change in chemical shift of the OH signal (+6.8 ppm) was observed, which is typical of H-bond formation.⁴⁸ During the titrations of **3.16•3.21**, **3.18•3.22** and **3.20•3.23**, the broad OH signals of the guest show a large positive complexation-induced chemical shift change as well.

Table 3.3. ^{31}P NMR chemical shifts of free host (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.

| $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ |
|--|--|--|
| 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| 53.6 | 53.9 | 53.9 |
| 53.6 | 53.8 | 53.9 |
| | 53.7 | 53.9 |
| | | 53.8 |

Table 3.4. Limiting complexation-induced ^{31}P NMR chemical shifts of fully bound host (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.^a

| $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \\ \text{A}\cdots\text{D} \end{bmatrix}$ |
|--|--|--|
| 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| 57.9 | 59.0 | 59.0 |
| 57.9 | 58.7 | 59.0 |
| | 58.5 | 59.0 |
| | | 57.6 |

Table 3.5. Limiting complexation-induced changes in ^{31}P NMR chemical shift (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.

| $\begin{bmatrix} \text{A}=\text{D} \\ \text{A}=\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}=\text{D} \\ \text{A}=\text{D} \\ \text{A}=\text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A}=\text{D} \\ \text{A}=\text{D} \\ \text{A}=\text{D} \\ \text{A}=\text{D} \end{bmatrix}$ |
|--|---|--|
| 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| 4.3 | 5.1 | 5.1 |
| 4.3 | 4.9 | 5.1 |
| | 4.8 | 5.1 |
| | | 3.8 |

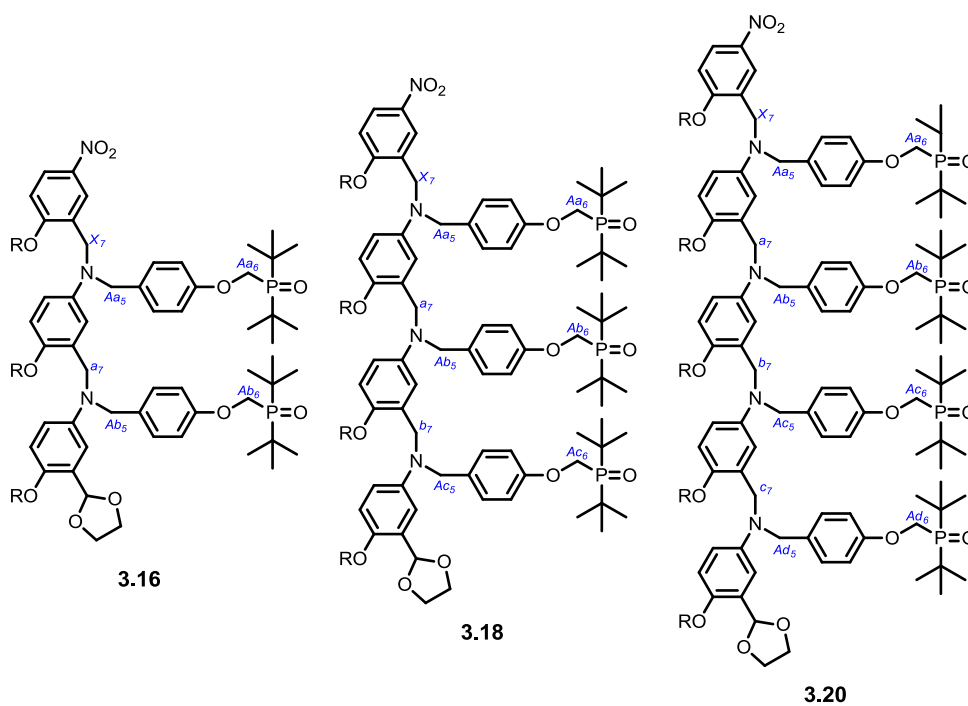


Figure 3.10. Proton labelling scheme.

Table 3.6. ^1H NMR chemical shifts of the free host (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.^a

| | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ |
|-----|--|--|--|
| | 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| Aa6 | 4.2 | 4.2 | 4.3 |
| Ab6 | 4.1 | 4.3 | 4.2 |
| Ac6 | | 4.2 | 4.4 |
| Ad6 | | | 4.2 |
| X7 | 4.5 | 4.6 | 4.5 |
| a7 | 4.6 | 4.7 | 4.7 |
| b7 | | 4.7 | 4.7 |
| c7 | | | 4.7 |
| Aa5 | 4.4 | 4.4 | 4.3 |
| Ab5 | 4.3 | 4.1 | 4.3 |
| Ac5 | | 4.1 | 4.1 |
| Ad5 | | | 4.1 |

^a refer to Figure 3.10 for proton labelling scheme

Table 3.7. Limiting complexation-induced ^1H NMR chemical shifts of the free host (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.^a

| | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ |
|-----|--|--|--|
| | 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| Aa6 | 4.3 | 4.4 | 4.5 |
| Ab6 | 4.2 | 4.5 | 4.3 |
| Ac6 | | 4.4 | 4.6 |
| Ad6 | | | ^b |
| X7 | 4.5 | 4.6 | 4.5 |
| a7 | 4.6 | 4.8 | ^b |
| b7 | | 4.8 | ^b |
| c7 | | | ^b |
| Aa5 | 4.3 | 4.3 | ^b |
| Ab5 | 4.3 | 4.0 | ^b |
| Ac5 | | 4.1 | 4.0 |
| Ad5 | | | 4.1 |

^a refer to Figure 3.10 for numbering scheme; ^b signal could not be monitored due to signal overlap.

Table 3.8. Limiting complexation-induced changes in ^1H NMR chemical shift changes (ppm) obtained by fitting titration data measured in toluene- d_8 at 298 K to a 1:1 binding isotherm.^a

| | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ |
|-----|--|--|--|
| | 3.16•3.21 | 3.18•3.22 | 3.20•3.23 |
| Aa6 | 0.1 | 0.2 | 0.2 |
| Ab6 | 0.1 | 0.2 | 0.1 |
| Ac6 | | 0.2 | 0.2 |
| Ad6 | | | |
| X7 | 0.0 | 0.1 | 0.0 |
| a7 | 0.0 | 0.1 | ^b |
| b7 | | 0.1 | ^b |
| c7 | | | ^b |
| Aa5 | -0.1 | 0.0 | ^b |
| Ab5 | 0.0 | -0.1 | ^b |
| Ac5 | | -0.1 | -0.1 |
| Ad5 | | | 0.0 |

^a refer to Figure 3.10 for numbering scheme; ^b signal could not be monitored due to signal overlap.

Figure 3.12 shows the association constants of the complexes plotted as a function of the number of H-bonding sites N (red data). An order of magnitude increase in association constant is observed for each additional H-bond formed. In all cases, $K_{\text{ref}} \overline{EM}_N$ is greater than 1 (Table 3.2), which indicates fully closed duplex is more stable than open frayed species. Although the average EM in each duplex is small, it is sufficient for duplex formation since K_{ref} is large. Previously in our group,

two other H-bonding oligomer systems have been synthesised and the association constants of the corresponding duplexes measured (the structures are shown in Figure 3.11).⁵⁵ Figure 3.12 compares the association constants of the duplexes measured here with previous the generations of multivalent H-bonding duplexes shown in Figure 3.11. The striking feature of Figure 3.12 is that the previous H-bonding duplexes show no increase in association constant for oligomers greater than the 2-mer but, encouragingly, this trend is not observed for the data presented here which show an order of magnitude increase in K_N for each extra H-bond.

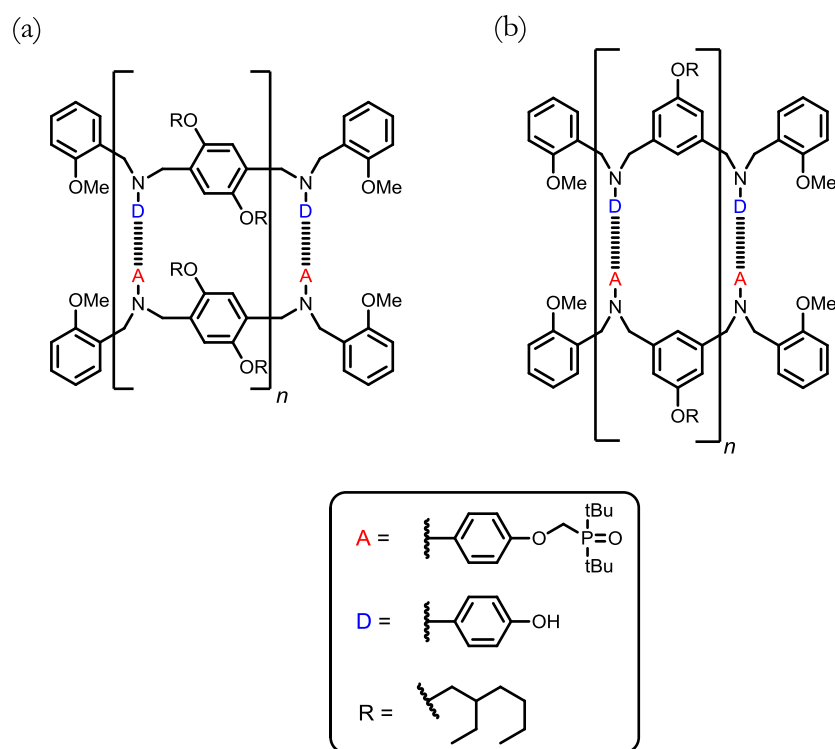


Figure 3.11. H-bonding duplexes synthesised previously in our group where $n = 0 - 3$.

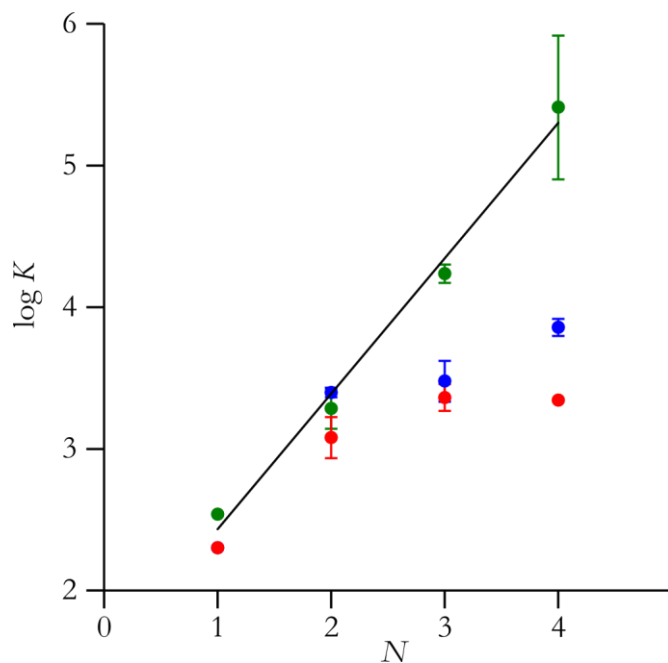


Figure 3.12. Log K of duplex formation as a function of the number of recognition modules in an oligomer (N), for duplexes measured in this Chapter (green), the duplexes shown in Figure 3.11(a) (blue) and the duplexes shown in Figure 3.11(b) (red). Error bars are shown to the 95% confidence limit. The line of best fit is shown for the duplexes measured in this chapter, $\log K_N = 0.96N + 1.48$.

3.3.3 Duplex denaturation through temperature change

Temperature denaturation data were measured for duplexes of **3.25•3.24**, **3.21•3.16**, **3.22•3.18** and **3.23•3.20**, by making 1:1 solutions of oligomers at 1 mM concentrations in toluene- d_8 and measuring the ^{31}P NMR spectra in temperature increments between -45 and 100 °C (Figure 3.13). The two ^{31}P signals in **3.21•3.16** gave a single overlapping peak that began to split into two signals at lower temperatures so for high temperatures, the chemical shift at the peak maximum was used, and at lower temperatures, the average of the chemical shifts of the two peaks was used (Figure 3.13(b)). Due to significant broadening of the ^{31}P NMR peaks for **3.22•3.18** 40 Hz line broadening was applied to all spectra. At high temperatures where one broad peak was observed, the chemical shift was recorded at the highest point of the peak. At lower temperatures, two of the three separate signals were resolved, so the average of the chemical shifts of the peaks was used (Figure 3.13(c)). Even after applying the same processing (40 Hz line broadening) to the spectra of **3.23•3.20** the peaks were too broad at lower temperatures to reliably distinguish from the background noise. An attempt to analyse these data is not included in this Section but is shown in Appendix II.

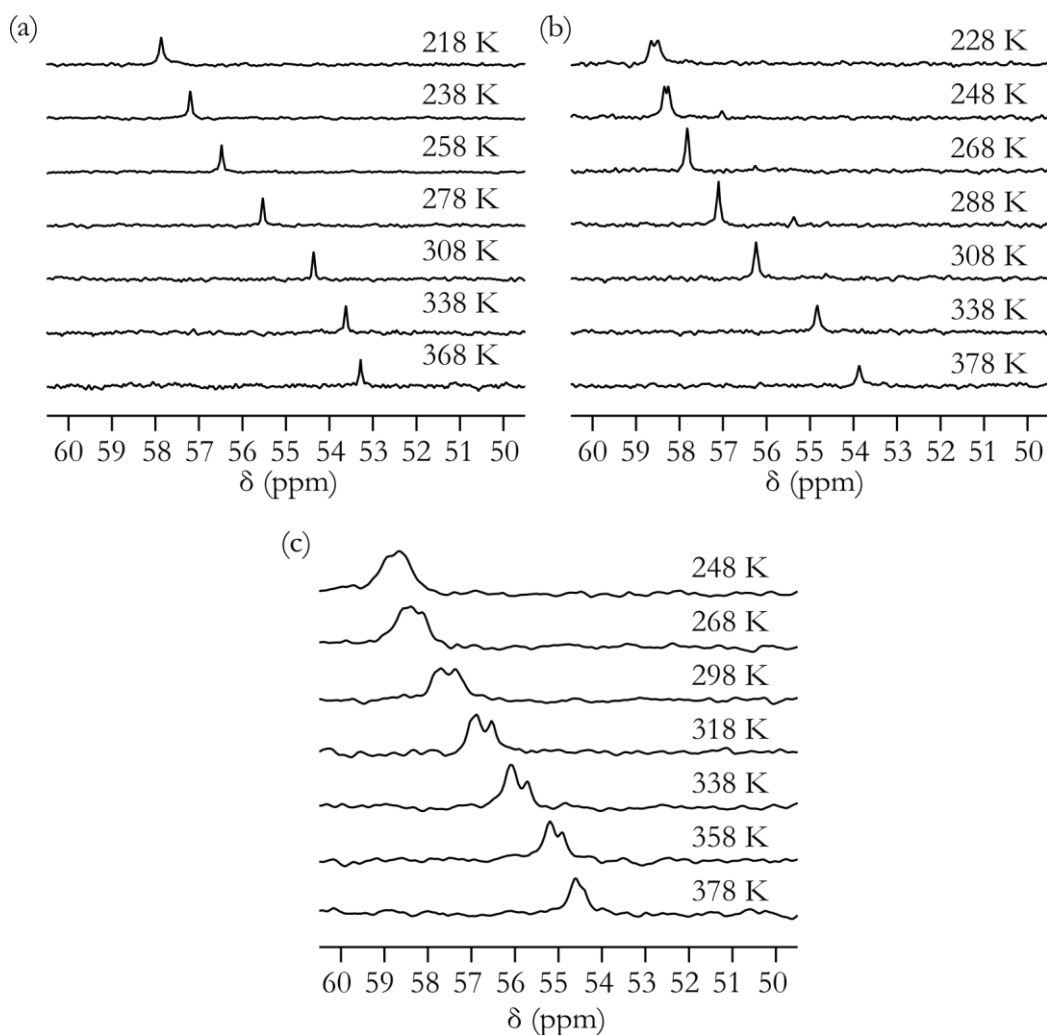


Figure 3.13. Variable temperature ^{31}P NMR spectra for 1 mM solutions of 1:1 mixtures of (a) **3.25•3.24**; (b) **3.21•3.16**; and (c) **3.22•3.18**, in toluene- d_8 . A high line broadening (40 Hz) was applied to the spectra in (c).

The thermodynamic parameters for duplex formation can be extracted by fitting the thermal denaturation data to a two-state model at equilibrium. A two state model considers only fully bound duplex and completely denatured single strands. The equilibrium constant for formation of a duplex between two oligomers with N binding sites that are not self-complementary, K_N , is given by Equation 3.4. If the two oligomers are present in equal concentrations, the total concentration of all free strands can be written as ϵ and the total fraction of all bound species as α .⁵⁰

$$K_N = \frac{[D_N \bullet A_N]}{[D_N][A_N]} = \frac{2\alpha}{c(1-\alpha)^2} \quad (3.4)$$

where $[D_N]$ and $[A_N]$ are the concentrations of single strand oligomers bearing donor groups and acceptor H-bonding groups respectively, $[D_N \bullet A_N]$ is concentration of the duplex.

Both the enthalpy and entropy of duplex formation are assumed to be temperature independent, and the change in heat capacity between free and bound states is assumed to be zero. Although it has been suggested that making these assumptions is an oversimplification,⁵¹ we are using these models to establish the duplex length dependence of the thermodynamic parameters and the transition melting temperature T_m , so we are more interested in relative changes in the parameters rather than the absolute values. The association constant at the transition melting temperature $T_{m,N}$ can be expressed as a function of ϵ by substituting $\alpha = 0.5$ into Equation 3.4. The value of $K_N(T_{m,N})$ can then be extrapolated to a different temperature T by using the integrated form of the van't Hoff equation (Equation 3.5)

$$\ln \left[\frac{K_N(T)}{K_N(T_{m,N})} \right] = -\frac{\Delta H_N}{R} \left(\frac{1}{T} - \frac{1}{T_{m,N}} \right) \quad (3.5)$$

By rearranging Equation 3.5 and substituting in the value of $K_N(T_{m,N})$ the temperature dependence of the association constant $K_N(T)$ can be expressed as Equation 3.6.⁵²

$$K_N(T) = \frac{4}{c} \exp \left\{ -\frac{\Delta H_N}{R} \left(\frac{1}{T} - \frac{1}{T_{m,N}} \right) \right\} = \frac{2\alpha}{c(1-\alpha)^2} \quad (3.6)$$

Equation 3.6 can be solved as a quadratic for α and, since the model assumes all-or-nothing binding, the observed chemical shift of a two state system in equilibrium δ can be expressed in terms of α by Equation 3.7

$$\delta = \alpha\delta_{\text{bound}} + (1 - \alpha)\delta_{\text{free}} \quad (3.7)$$

where δ_{bound} and δ_{free} are the chemical shifts of the duplex and single strand states.

Since ΔH_N is assumed not to change with T , the entropy and Gibbs free energy of the system can be calculated from $K_N(T)$ and ΔH_N using Equations 3.8 and 3.1.

$$\Delta G_N = \Delta H_N - T\Delta S_N \quad (3.8)$$

The thermal denaturation data of duplexes **3.25•3.24**, **3.21•3.16** and **3.22•3.18** were fit to Equations 3.6 and 3.7 where δ_{bound} , δ_{free} , ΔH_N and $T_{m,N}$ were allowed to be variable parameters. Strictly speaking both δ_{bound} and δ_{free} should be allowed to vary as a function of T and N . In order to assess the temperature dependence of the chemical shifts, ^{31}P spectra of a 1 mM toluene- d_8 solution of **3.24** were recorded between -45 and 100 °C. Although there was some variation in chemical shift as a function of temperature, there was no clear trend, and the variation ($\Delta\delta = \sim 1$ ppm) was small relative to the thermal denaturation data ($\Delta\delta > 4$ ppm) (see Appendix II). Therefore, we ignore variation in δ_{bound} and δ_{free} as a function of T . The δ_{free} and δ_{bound} ^{31}P NMR chemical shift values calculated from fitting the titration data recorded at 298 K in Section 3.2 are approximately independent of N (53.6 – 53.9 ppm for free oligomers and 57.9 – 59.0 ppm for bound duplexes, see Tables 3.3-3.5), so we fit all the thermal denaturation data assuming the same values of δ_{bound} and δ_{free} for each duplex. The results, $\delta_{\text{bound}} = 58.9$ ppm and $\delta_{\text{free}} = 52.8$ ppm, are similar to the values obtained from the titration data. Figure 3.14 shows the lines of best fit calculated

using Equations 3.6 and 3.7 (solid lines) for each thermal denaturation data set, and Table 3.9 shows the values of the fitted parameters along with ΔS_N and ΔG_N , calculated using Equations 3.1 and 3.8. The values of K_N at 298 K in Table 3.9 are similar to the corresponding values from the titration data (Table 3.2), which suggests this method of fitting thermal denaturation data is valid.

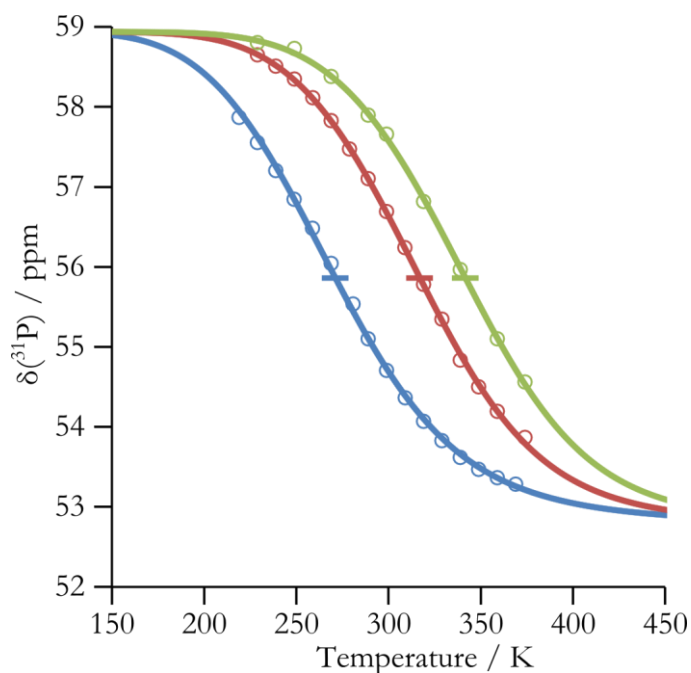


Figure 3.14. Experimental ^{31}P NMR chemical shift values as a function of temperature for 1 mM 1:1 mixtures in toluene- d_8 of **3.25•3.24** (blue), **3.21•3.16** (red), and **3.22•3.18** (green), compared to the calculated values from Equations 3.6 and 3.7 (solid lines). The bars in the middle of the curves show the transition temperature T_m . The total root mean square deviation for the fit is < 0.2 ppm.

Table 3.9. Thermodynamic parameters for the formation of duplexes in toluene- d_8 calculated from ^{31}P NMR thermal denaturation data.

| N | Duplex | $T_{m,N} / \text{K}$ | $T = 298 \text{ K}$ | | | |
|-----|------------------|----------------------|--|---|--|------------------------------|
| | | | ΔH_N (kJ mol^{-1}) | $T\Delta S_N$ (kJ mol^{-1}) | ΔG_N (kJ mol^{-1}) | K_N (M^{-1}) |
| 1 | 3.25•3.24 | 272 | -26 | -9 | -16 | 720 |
| 2 | 3.21•3.16 | 318 | -37 | -17 | -21 | 4900 |
| 3 | 3.22•3.18 | 342 | -42 | -18 | -24 | 18000 |

The experimental data in Figure 3.14 clearly follow a sigmoidal curve corresponding to the melting transition. For increasing values of N , the free-bound transitions are increasingly steeper and the point of inflection occurs at higher temperatures. These visual observations are confirmed by the calculated parameters shown in Table 3.9. The thermodynamic parameters in Table 3.9 show that the increased association constants for higher values of N are associated with an increase in the enthalpy change and transition melting temperature. A steeper melting curve and increasing $T_{m,N}$ indicates that duplex stability increases with N as a result of cooperativity between the H-bonding interactions along the duplex.⁵³

3.3.4 Duplex denaturation through competition

Complexes formed through highly cooperative interactions show “all or nothing” behaviour when broken apart using a competing ligand. As cooperativity increases the sigmoidal isotherm for binding of the denaturant to the complex becomes increasingly steep.⁵⁴ Therefore, in order to further probe the extent of cooperativity between the H-bonding interactions in the duplexes of **3.25•3.24**, **3.21•3.16**, **3.22•3.18** and **3.23•3.20**, aliquots of dimethyl sulfoxide- d_6 (DMSO- d_6) were added to 1 mM solutions of each duplex in toluene- d_8 until a large excess was reached (>1 M). The extent of denaturation was monitored using ^{31}P and ^1H NMR recorded at 298 K after each addition of DMSO- d_6 . The ^{31}P NMR spectra are shown in Figure 3.15 and the chemical shifts of a representative signal are plotted as a function of DMSO- d_6 concentration in Figure 3.16.

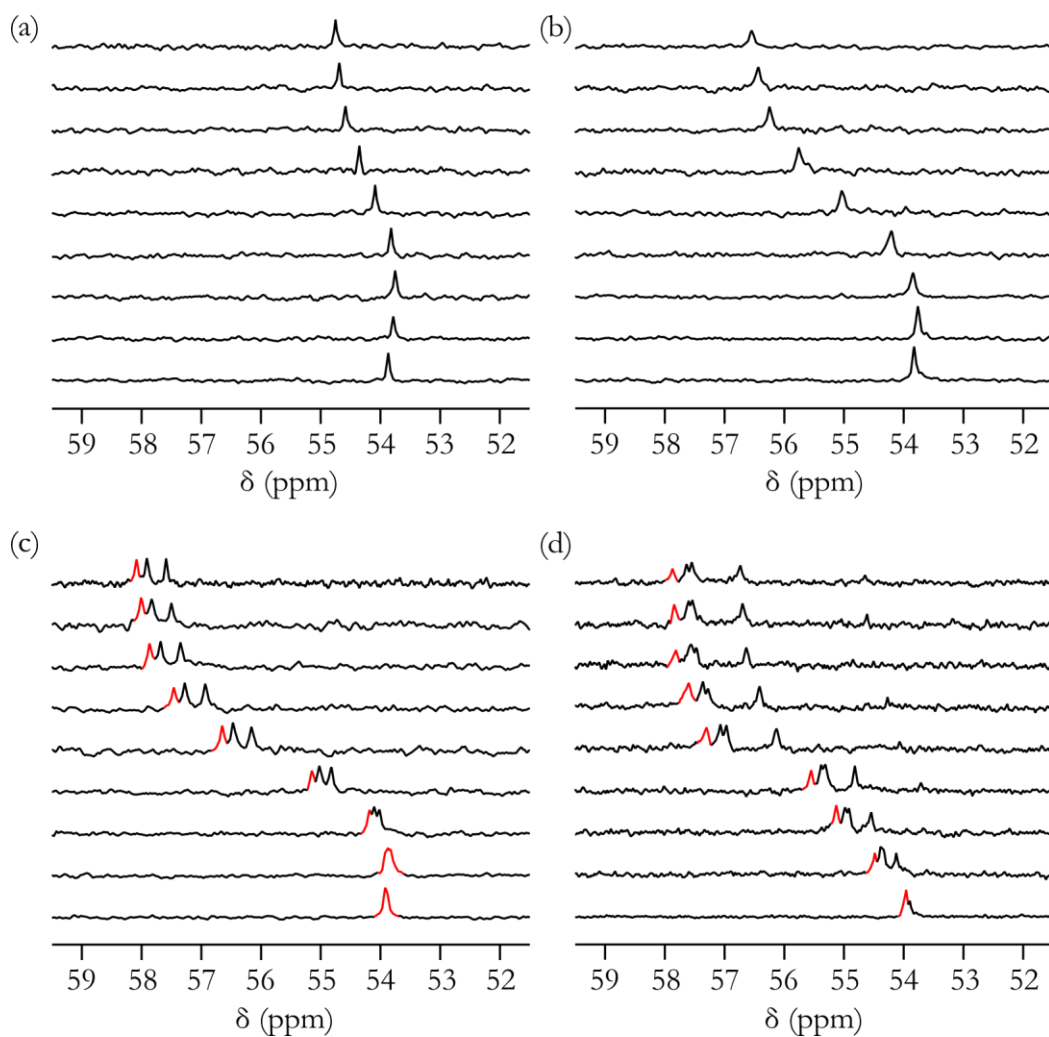


Figure 3.15. ^{31}P NMR titration spectra in toluene- d_8 at 298 K where the concentration of DMSO- d_6 increases from top to bottom: (a) **3.25•3.24**; (b) **3.21•3.16**; (c) **3.22•3.18**; and (d) **3.23•3.20**. The highlighted signals are plotted in Figures 3.16 and 3.18 (below). The concentration of oligomers starts at 1.0 mM and decreases to 0.7 mM from top to bottom.

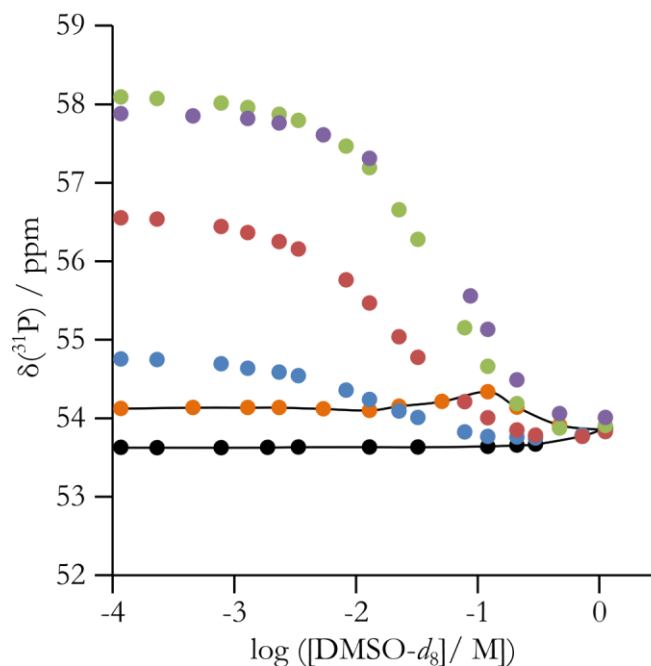


Figure 3.16. ^{31}P NMR chemical shifts as a function of the log concentration of $\text{DMSO-}d_8$ in $\text{toluene-}d_8$ at 298 K, for **3.25•3.24** (blue), **3.21•3.16** (red), **3.22•3.18** (green) and **3.23•3.20** (purple). The signals followed are highlighted in Figure 3.15. Data for **3.24** (black) and **3.18** (orange) are shown as references for unbound complex.

The denaturation curves are all sigmoidal and were analysed using three different isotherms:

- (i) *1:1 binding*, occurs in the case that no cooperativity is observed between each H-bonding interaction. For this analysis the data was fit to a simple 1:1 isotherm where the host concentration used is the concentration of phenol groups in the solution (i.e. $N \times [\text{D}_N]$), and the guest is $\text{DMSO-}d_6$.
- (ii) *Stepwise binding*, is shown schematically in Figure 3.17 and accounts for each partially bound state of duplex and denaturant complex, taking into account the higher association constants of oligomers with more H-bonds. This analysis assumes that the oligomers can still form a duplex after one or more of the phenol groups have formed H-bonds with the $\text{DMSO-}d_6$.

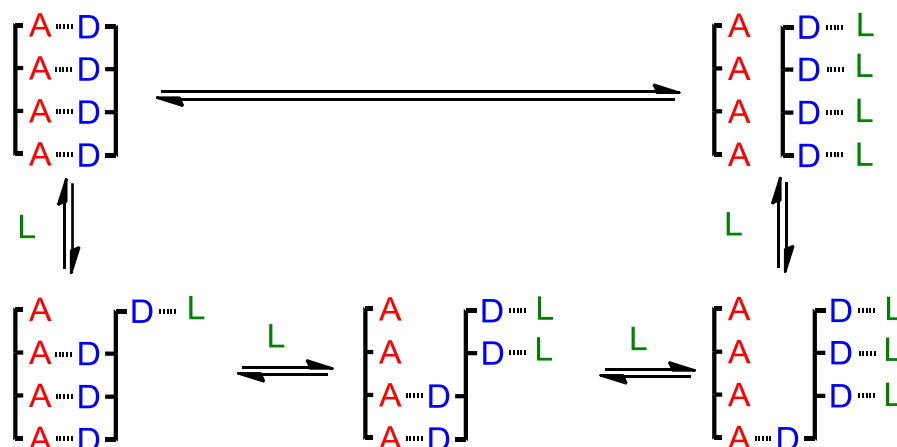


Figure 3.17. Equilibria involved in stepwise denaturation of a 4-mer duplex by ligand L.

In this case, a duplex ($D_N \bullet A_N$) which is a complex of two oligomers (D_N and A_N) formed of N H-bonds, can bind with up to M denaturant molecules of ligand L where $M \leq N$. The concentration of duplex of length N bound to M ligand molecules is given by Equation 3.9, where σ_{N-M} is the statistical factor, K_{N-M} is the experimentally determined association constant for oligomer of length $N - M$ and K_L is the association constant of **3.25** with dimethyl sulfoxide- d_6 (experimentally determined to be $K = 100 \pm 10 \text{ M}^{-1}$ in toluene- d_8 at 298 K).

$$[D_N \bullet A_N \bullet (L)_M] = \sigma_{N-M} K_{N-M} [D_N][A_N] \left((K_L [L]_{\text{free}})^M \right) \text{ for } M < N \quad (3.9)$$

The concentration of phenol oligomer fully bound to DMSO- d_6 is given by Equation 3.10 and the concentration of fully bound duplex is given by Equation 3.11.

$$[D_N \bullet (L)_N] = [D_N] \left((K_L [L]_{\text{free}})^N \right) \quad (3.10)$$

$$[D_N \bullet A_N] = \sigma_N K_N [D_N][A_N] \quad (3.11)$$

The binding isotherm is given by Equation 3.12

$$\frac{\delta - \delta_{\text{free}}}{\delta_{\text{bound}} - \delta_{\text{free}}} = \frac{N[\text{D}_N \bullet \text{A}_N] + \sum_{M=1}^{M=N-1} (N-M)[\text{D}_N \bullet \text{A}_N \bullet \text{L}_M]}{[\text{D}_N \bullet (\text{L})_N] + N[\text{D}_N \bullet \text{A}_N] + \sum_{M=1}^{M=N-1} (N-M)[\text{D}_N \bullet \text{A}_N \bullet \text{L}_M]} \quad (3.12)$$

where δ is the ^{31}P NMR chemical shift at a given point in the titration, δ_{bound} is the chemical shift of fully bound duplex and δ_{free} is the chemical shift of the denatured duplex.

(iii) *1:N binding*, is the extreme all or nothing case that applies to highly cooperative systems where only duplex and fully denatured oligomer are considered. The association constant for 1:N denaturation is given by Equation 3.13.

$$K_D = \frac{[\text{D}_N \bullet (\text{L})_N][\text{A}_N]}{[\text{D}_N \bullet \text{A}_N][\text{L}]^N} \quad (3.13)$$

Figure 3.18 shows a comparison of the three different binding isotherms considered. The data are best fit by a 1:1 isotherm in all cases. The implication of these data is that, when DMSO- d_6 denatures the duplex structure, there is no detectable cooperativity between the H-bonding events. It is difficult to rationalise the conclusions from these denaturation data with the binding data discussed in Table 3.2.

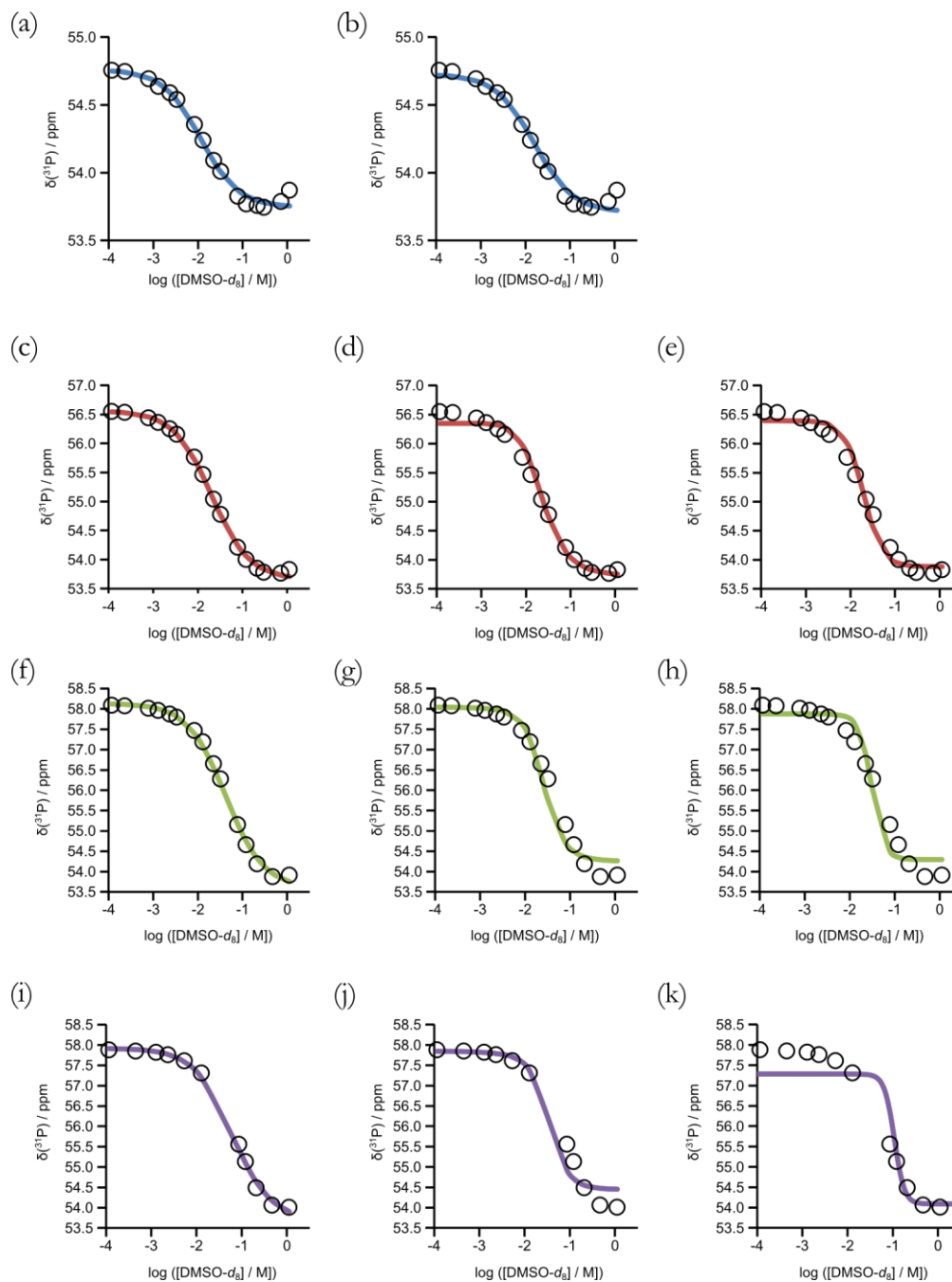


Figure 3.18. ^{31}P NMR denaturation titration data at 298 K in toluene- d_8 (circles) and calculated binding isotherms (lines). The root mean square of deviation of the fit is given in brackets. (a) **3.25•3.24**, 1:1, (0.04 ppm); (b) **3.25•3.24**, stepwise, (0.05 ppm); (c) **3.21•3.16**, 1:1, (0.04 ppm); (d) **3.21•3.16**, stepwise, (0.11 ppm); (e) **3.21•3.16**, 1: N , (0.14 ppm); (f) **3.22•3.18**, 1:1, (0.09 ppm); (g) **3.22•3.18**, stepwise, (0.25 ppm); (h) **3.22•3.18**, 1: N , (0.30 ppm); (i) **3.23•3.20**, 1:1, (0.08 ppm); (j) **3.23•3.20**, stepwise, (0.30 ppm); (k) **3.23•3.20**, 1: N , (0.42 ppm).

3.3.5 Intramolecular folding

In order to determine how facile intramolecular H-bond formation would be between adjacent sites in mixed sequence oligomers, we used **3.27**, which is a 2-mer oligomer with both an H-bond acceptor and donor (Figure 3.4(b)). Figure 3.19 shows the equilibria of **3.27** in solution, where an H-bond can form between the phenol and phosphine oxide within the same molecule, or a duplex is formed between two molecules of **3.27** in an anti-parallel configuration.

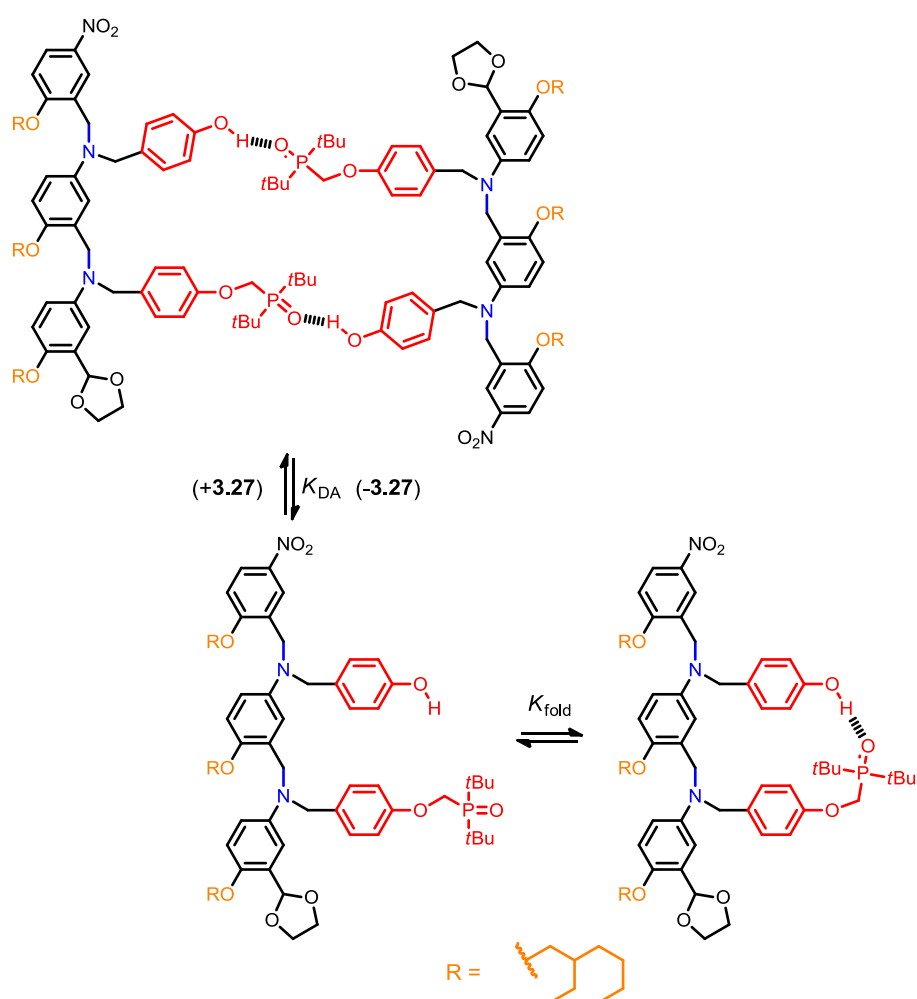


Figure 3.19. Competing equilibria of **3.27**: unimolecular folding and anti-parallel duplex formation.

If one assumes the association constant of duplex formation of **3.27** to form the self-complementary duplex **3.27•3.27** is the same as for the formation of **3.16•3.21**,

then the observed equilibrium association constant K_{DA} is given by $1/2K_{\text{ref}}^2EM_2$ where EM_2 is the effective molarity determined for **3.16•3.21**. The association constant of the unimolecular folding of **3.27** K_{fold} has an association constant of $K_{\text{ref}}EM_{\text{fold}}$. K_{fold} cannot be measured directly but can be estimated from the dimerisation association constant of **3.27•3.27** and the association constant for formation of **3.16•3.21**. The observed association constant obtained from dilution data of **3.27** would be lower than K_2 for **3.16•3.21** if the folded species is populated. Figure 3.20(b) shows the change in the ^{31}P NMR chemical shift of **3.27** in toluene- d_8 at 298 K as a function of the concentration of **3.27**. The data were fit to a dimerisation isotherm (Figure 3.20(b), solid line).

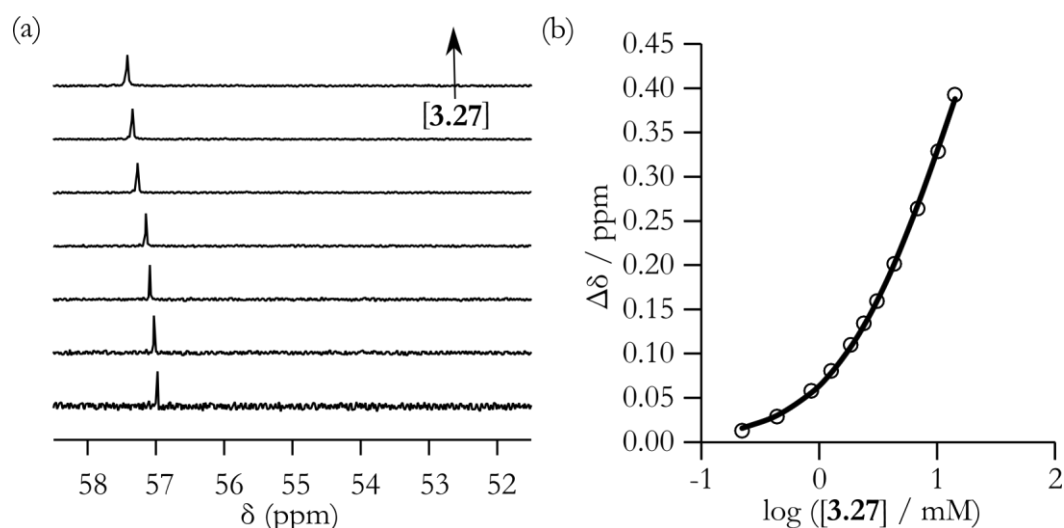


Figure 3.20. Dilution data (400 MHz, toluene- d_8 , 298 K): (a) ^{31}P NMR spectra of **3.27** at different concentrations (increasing bottom to top); (b) ^{31}P NMR chemical shift as a log function of the concentration of **3.27**. The black circles show experimental data and the line represents the best fit to a dimerisation isotherm. ($K = 40 \text{ M}^{-1}$, maximum %bound = 39%, %bound, minimum %bound = 2%, $\delta_{\text{free}} = 57.0 \text{ ppm}$; $\delta_{\text{bound}} = 58.1 \text{ ppm}$, $\Delta\delta = +1.1 \text{ ppm}$).

The striking result from the dilution of **3.27** is the difference in the limiting complexation-induced change in ^{31}P NMR chemical shift for the formation of **3.27•3.27** and the formation of **3.21•3.16** (compare Figure 3.20(a) and Figure 3.7(a)). The ^{31}P NMR chemical shift of free **3.27** is 57.0 ppm, which is much higher than the

corresponding value for **3.16** (53.6 ppm) and closer to the bound chemical shift for the **3.21•3.16** complex (57.9 ppm) suggesting that the phosphine oxide in **3.27** is forming an H-bond in the free state. The fit shown in Figure 3.20(b) gives an association constant of $K = 40 \text{ M}^{-1}$, which is significantly lower than the association constant for formation of the **3.21•3.16** duplex ($K_2 = 1900 \text{ M}^{-1}$).

The extent of folding and K_{fold} can be determined using Equations 3.19 and 3.20, which are derived through Equations 3.14 – 3.18. This derivation relies on the association constant for **3.21•3.16** (K_2) being the same as the dimerisation constant of **3.27•3.27**

$$K_2 = \frac{[\mathbf{3.21} \bullet \mathbf{3.16}]}{[\mathbf{3.21}][\mathbf{3.16}]} = \frac{[\mathbf{3.27} \bullet \mathbf{3.27}]}{[\mathbf{3.27}]_{\text{open}}^2} \quad (3.14)$$

where $[\mathbf{3.27}]_{\text{open}}$ is the concentrations of **3.27** in unfolded form. Equations 3.15 and 3.16 define K_{DA} and K_{fold} , respectively, where $[\mathbf{3.27}]_{\text{closed}}$ is the concentration of free DA in its folded form.

$$K_{\text{DA}} = \frac{[\mathbf{3.27} \bullet \mathbf{3.27}]}{[\mathbf{3.27}]^2} \quad (3.15)$$

$$K_{\text{fold}} = \frac{[\mathbf{3.27}]_{\text{closed}}}{[\mathbf{3.27}]_{\text{open}}} \quad (3.16)$$

The total concentration of **3.27** can be expressed in terms of $[\mathbf{3.27}]_{\text{open}}$ and K_{fold} (Equation 3.17) and substituted into Equation 3.15 to give Equation 3.18.

$$\begin{aligned} [\mathbf{3.27}] &= [\mathbf{3.27}]_{\text{open}} + [\mathbf{3.27}]_{\text{closed}} \\ &= [\mathbf{3.27}]_{\text{open}} + K_{\text{fold}}[\mathbf{3.27}]_{\text{open}} \\ &= (1 + K_{\text{fold}})[\mathbf{3.27}]_{\text{open}} \end{aligned} \quad (3.17)$$

$$\begin{aligned}
 K_{\text{DA}} &= \frac{[\mathbf{3.27} \bullet \mathbf{3.27}]}{[\mathbf{34}]_{\text{open}}^2} \cdot \frac{1}{(1 + K_{\text{fold}})^2} \\
 &= \frac{K_2}{(1 + K_{\text{fold}})^2}
 \end{aligned}
 \tag{3.18}$$

Equation 3.18 can be rearranged in terms of K_{fold} to give Equation 3.19.

$$K_{\text{fold}} = \sqrt{\frac{K_2}{K_{\text{DA}}}} - 1
 \tag{3.19}$$

Dividing K_{fold} by the intermolecular association K_{ref} gives the effective molarity of folding EM_{fold} (Equation 3.20).

$$EM_{\text{fold}} = \frac{K_{\text{fold}}}{K_{\text{ref}}}
 \tag{3.20}$$

Using Equation 3.19, K_{fold} is found to be 6 which gives an EM_{fold} of 20 mM, roughly twice that of EM_1 for **3.21•3.16** (8 mM). This analysis, along with the low dilution $\Delta\delta$ and high δ_{free} suggest that intramolecular H-bond folding competes strongly with duplex formation for **3.27**.

3.4 Conclusions

Herein a new class of H-bonding duplexes which show an order of magnitude increase in stability for each additional H-bond is presented. Despite a modest effective molarity (10 – 20 mM) duplex formation between the oligomers was cooperative ($K \overline{EM}_N = 3 - 7$) since a strong H-bonding interaction was chosen. Binding constants calculated from thermal denaturation experiments show a similar trend to those determined from titration experiments for 1-mer to 3-mer duplexes. Denaturation using dimethyl sulfoxide- d_6 suggests the duplexes break open with a non-cooperative mechanism where partially bound states contribute a significant

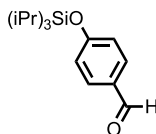
amount to the overall composition. It is difficult to make firm conclusions from these binding studies as the denaturation experiments suggest a non-cooperative model for binding.

The system presented in this chapter is not suitable for programming oligomer sequences which have both the phenol and phosphine oxide groups adjacent to each other, since H-bond formation between adjacent H-bonding groups effectively competes with duplex formation. One approach to deal with the problem of sequence would be to change the H-bond donor and acceptor to a pair that cannot readily H-bond to adjacent positions. The synthetic strategy we have chosen will allow us to make modifications to the recognition groups without a drastic redesign, and this approach will be explored in the next chapter.

3.5 Experimental section

3.5.1 Synthesis

All the reagents and materials used in the synthesis of the compounds described below were bought from commercial sources, without prior purification. Thin layer chromatography was carried out using with silica gel 60F (Merck) on aluminium. Flash chromatography was carried out on silica gel 40 – 60 μm (BDH) or on an automated system (Combiflash Companion) using pre-packed cartridges of silica (50 μm PuriFlash® Column). All NMR spectroscopy was carried out on either a Bruker AVI250, AVI400, DPX400, AVIII400 or DRX500 spectrometer using the residual solvent as the internal standard. All chemical shifts (δ) are quoted in ppm and coupling constants given in Hz. Splitting patterns are given as follows: s (singlet), d (doublet), t (triplet), m (multiplet). FT-IR spectra were measured on a PerkinElmer Spectrum 100 spectrometer. ES+ was carried out on a Micromass Platform spectrometer. Reactions were carried out at ambient temperature unless otherwise stated.

3.5.1.1 Synthesis of **3.2**⁴⁹

3.1 (5.01 g, 40.9 mmol, 1 equiv.) was dissolved in DMF (50 mL) and TIPS-Cl (10.1 mL, 47.1 mmol, 1.2 equiv.) and imidazole (5.58 g, 81.9 mmol, 2 equiv.) were added with stirring. After 12 hours stirring, the reaction mixture was poured onto water (100 mL), before extraction into hexane (5×20 mL). The organic fractions were washed several times with water (5×100 mL) then brine (1×100 mL), dried (MgSO_4) and the solvent was removed with a rotary evaporator to yield an oil (12.8 g). The crude product was purified by flash chromatography on silica eluting with 10% EtOAc in hexane to yield a pale yellow oil (11.0 g, 96%).

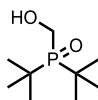
^1H NMR (250 MHz, CDCl_3): $\delta_{\text{H}} = 9.89$ (s, 1H), 7.79 (d, 2H, $J = 9.0$), 6.99 (d, 2H, $J = 9.0$), 1.20 – 1.38 (m, 3H), 1.07 – 1.16 (m, 18H);

^{13}C NMR (63 MHz, CDCl_3): $\delta_{\text{C}} = 190.8, 161.9, 131.9, 130.2, 120.3, 17.8, 12.7$;

MS (ES+): m/z (%) = 279 (100) [$\text{M} + \text{H}^+$];

HRMS (ES+): calculated for $\text{C}_{16}\text{H}_{27}\text{O}_2\text{Si}$ 279.1780, found 279.1791;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2945, 2893, 2867, 2730, 1698, 1596, 1575, 1508, 1463, 1275, 1211, 1155.

3.5.1.2 Synthesis of **3.4**⁵⁵

3.3 (20.0 g, 111 mmol) was dissolved in 37% aqueous formaldehyde (210 mL) and concentrated aqueous HCl (200 mL) was added and heated to 100 °C for 12 hours with stirring. The reaction mixture was cooled to room temperature before being neutralised using NaOH. The aqueous solution was washed with EtOAc (3 × 100 mL). The organic fractions were washed with brine (1 × 100 mL), dried (MgSO₄) and the solvent removed using a rotary evaporator. The crude product was recrystallized in boiling hexane to yield a white solid (11.1 g, 52%).

¹H NMR (250 MHz, CDCl₃): δ_H = 4.05 (s, 2H), 1.28 (d, 18H, *J* = 13);

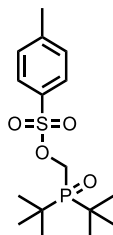
¹³C NMR (63 MHz, CDCl₃): δ_C = 54.8 (d, *J* = 62), 34.8 (d, *J* = 55), 26.4;

³¹P NMR (101 MHz, CDCl₃): δ_P = 154.0;

MS (ES+): *m/z* (%) = 193 (100) [M + H⁺];

HRMS (ES+): calculated for C₉H₂₂O₂P 193.1357, found 193.1367;

FT-IR (thin film): ν_{max}/cm⁻¹ 3142, 2957, 1474, 1129, 1120, 1064.

3.5.1.3 Synthesis of **3.5**⁵⁵

3.4 (9.2 g, 48 mmol, 1 equiv.) was dissolved in THF 130 mL and NEt₃ (10 mL) at 0 °C with stirring. 4-toluenesulfonyl chloride (11.4 g, 60 mmol, 1.3 equiv.) in THF (30 mL) was added slowly over 20 minutes. After 18 hours the mixture was poured onto water (100 mL) and extracted into EtOAc (3 × 20 mL). The organic extracts were washed with water (1 × 20 mL) and brine (1 × 2 mL) before drying (MgSO₄). The solvent was removed using a rotary evaporator and the crude product purified by flash chromatography on silica eluting with a gradient from 0% to 3% MeOH in DCM to yield a colourless oil, which after several days crystallised into a white amorphous solid (14.4 g, 87%).

¹H NMR (250 MHz, CDCl₃): δ_H = 7.81 (d, 2H, *J* = 8.0), 7.39 (d, 2H, *J* = 8.0), 4.27 (d, 2H, *J* = 7.0), 2.47 (s, 3H), 1.29 (d, 18H, *J* = 13.8);

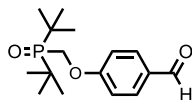
¹³C NMR (63 MHz, CDCl₃): δ_C = 145.7, 131.1, 130.0, 128.2, 61.4 (d, *J* = 60.5), 35.6 (d, *J* = 58.5), 26.2, 21.6;

³¹P NMR (101 MHz, CDCl₃): δ_P = 56.0;

MS (ES+): *m/z* (%) = 347 (100) [M + H⁺];

HRMS (ES+): calculated for C₁₆H₂₈O₄PS 347.1446, found 347.1456;

FT-IR (thin film): ν_{max}/cm⁻¹ 3431, 2968, 2874, 1597, 1478, 1369, 1190, 1176, 1146, 1095, 996.

3.5.1.4 Synthesis of **3.6**

3.1 (0.55 g, 4.5 mmol, 1 equiv.) was dissolved in DMF (20 mL) and **3.5** (1.70 g, 4.9 mmol, 1.1 equiv.) and Cs₂CO₃ (2.18 g, 6.69 mmol, 1.5 equiv.) were added and the mixture heated at 80 °C for 3 days with stirring. The reaction was then poured onto water (100 mL) and extracted into EtOAc (5 × 20 mL), before being washed with water (5 × 40 mL) and brine (1 × 40 mL) and then dried (MgSO₄). The solvent was removed using a rotary evaporator before the crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in diethyl ether to yield an amorphous white solid (0.88 g, 66%).

¹H NMR (250 MHz, CDCl₃): δ_H = 9.93 (s, 1H), 7.89 (d, 2H, *J* = 9.0), 7.07 (d, 2H, *J* = 9.0), 4.48 (d, 2H, *J* = 7.0), 1.40 (d, 18H, *J* = 14.0);

¹³C NMR (63 MHz, CDCl₃): δ_C = 190.2, 162.8 (d, *J* = 9.6), 131.8, 130.7, 114.33, 62.6 (d, *J* = 68.1), 35.2 (d, *J* = 57.6), 26.2;

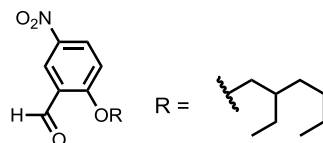
³¹P NMR (101 MHz, CDCl₃): δ_P = 56.5;

MS (ES+): *m/z* (%) = 297 (100) [M + H⁺], 338 (15) [M + CH₃CN], 360 (10);

HRMS (ES+): calculated for C₁₆H₂₆O₃P 297.1620, found 297.1619;

FT-IR (thin film): ν_{max}/cm⁻¹ 2967, 2906, 2871, 1694, 1603, 1582, 1509, 1306, 1242, 1163, 1133, 1041.

3.5.1.5 Synthesis of 3.8



3.7 (15 g, 90 mmol, 1 equiv.), K_2CO_3 (18.6 g, 135 mmol, 1.5 equiv.) and 2-ethylhexyl bromide (23.9 mL, 135 mmol, 1.5 equiv.) were dissolved in DMF (100 mL) and heated to 70 °C for 4 days with stirring. The mixture was then poured onto water (100 mL) and extracted into ethyl acetate (3×100 mL), before being washed with water (5×200 mL) and finally brine (1×100 mL). The organic fractions were dried (MgSO_4) and the solvent removed using a rotary evaporator to give the crude oil (38.8 g), which was purified by flash chromatography on silica eluting with 10% ethyl acetate in hexane to give product yellow oil (23.0 g, 91%).

^1H NMR (250 MHz, CDCl_3): $\delta_{\text{H}} = 10.48$ (s, 1H), 8.70 (d, 1H, $J = 3.0$), 8.42 (dd, 1H, $J = 9.0, 3.0$), 7.12 (d, 1H, $J = 9.0$), 4.11 (d, 2H, $J = 6.0$), 1.86 (s, 1H), 1.23 – 1.64 (m, 8H), 0.84 – 1.03 (m, 6H);

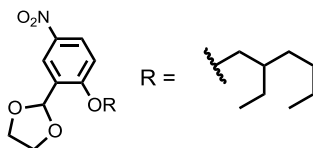
^{13}C NMR (63 MHz, CDCl_3): $\delta_{\text{C}} = 187.4, 165.4, 141.4, 124.7, 124.4, 112.9, 72.3, 39.3, 30.4, 29.0, 23.9, 22.9, 13.9, 11.0$;

MS (ES+): m/z (%) = 280 (100) [$\text{M} + \text{H}^+$], 291 (20).

HRMS (ES+): calculated for $\text{C}_{15}\text{H}_{22}\text{NO}_4$ 280.1549, found 280.1541;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2960, 2931, 2873, 1692, 1608, 1590, 1522, 1488, 1461, 1429, 1176, 1142, 1077, 1005.

3.5.1.6 Synthesis of 3.9



3.8 (14.1 g, 50.6 mmol, 1 equiv.) was dissolved in toluene (75 mL) and *p*-toluenesulfonic acid (0.19 g, 1.01 mmol, 0.02 equiv.) and ethylene glycol (9.43, 151 mmol, 3 equiv.) were added and the solution heated to reflux with stirring. The toluene/water condensate was removed using Dean-Stark apparatus until the condensate ran clear. The reaction was cooled to room temperature and poured onto water (100 mL) before being extracted into EtOAc (3 × 50 mL). The organic fractions were collected and washed with brine (1 × 100 mL), dried (MgSO₄), and concentrated on a rotary evaporator. The crude product required no further purification and was isolated as a golden viscous oil (15.9 g, 96% yield).

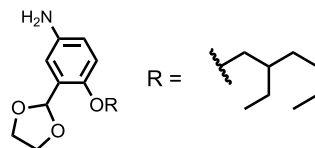
¹H NMR (250 MHz, CDCl₃): δ_H = 8.43 (d, 1H, *J* = 3.0), 8.24 (dd, 1H, *J* = 9.0, 3.0), 6.96 (d, 1H, *J* = 9.0), 6.14 (s, 1H), 4.03 – 4.22 (m, 4H), 4.01 (d, 2H, *J* = 5.5), 1.73 – 1.89 (m, 1H), 1.24 – 1.61 (m, 8H), 0.87 – 1.00 (m, 6H);

¹³C NMR (63 MHz, CDCl₃): δ_C = 162.3, 141.1, 127.5, 126.4, 123.3, 111.2, 98.3, 71.8, 65.4, 39.2, 30.4, 29.0, 14.0, 11.0;

MS (ES+): *m/z* (%) = 324 (100) [M + H⁺].

HRMS (ES+): calculated for C₁₇H₂₆NO₆ 324.1811, found 324.1820.

FT-IR (thin film): ν_{max}/cm⁻¹ 2959, 2931, 2875, 1613, 1596, 1518, 1494, 1463, 1340, 1270, 1105, 1067.

3.5.1.7 Synthesis of **3.10**

3.9 (30.56 g, 94.51 mmol, 1 equiv., 0.2 M) was dissolved in in degassed EtOAc (475 mL) and Pd/C (2.01 g, 18.9 mmol) was added with rigorous stirring. The reaction was stirred in an atmosphere of H₂ overnight. The reaction is then filtered through a plug of Celite® and the solvent was removed using a rotary evaporator to yield a dark red oil. The crude product was purified by flash chromatography on silica eluting with a gradient from 30% to 40% of EtOAc/hexane. The product was isolated as a yellow oil (23.4 g, 84%).

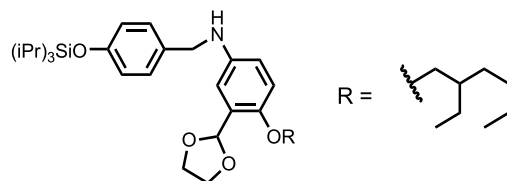
¹H NMR (250 MHz, CDCl₃): δ_{H} = 6.94 (d, 1H, J = 3.0), 6.76 (d, 1H, J = 8.5), 6.69 (dd, 1H, J = 8.50, 3.0), 6.12 (s, 1H), 3.98 – 4.18 (m, 4H), 3.82 (d, 2H, J = 5.5), 1.64 – 1.82 (m, 1H), 1.22 – 1.59 (m, 8H), 0.85 – 0.97 (m, 6H);

¹³C NMR (63 MHz, CDCl₃): δ_{C} = 150.9, 139.5, 127.0, 117.0, 114.2, 113.8, 99.2, 72.0, 65.2, 39.5, 30.5, 29.1, 23.9, 23.0, 14.0, 11.1.

MS (ES⁺): m/z (%) = 294 (100) [M + H⁺];

HRMS (ES⁺): calculated for C₁₇H₂₈NO₃ 294.2069, found 294.2055;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 3360, 2957, 2927, 2873, 1627, 1501, 1455, 1383, 1264, 1221, 1172, 1067.

3.5.1.8 Synthesis of **3.11**

3.10 (4.92 g, 16.8 mmol, 1 equiv.) and **3.2** (5.14 g, 18.5 mmol, 1.1 equiv.) were dissolved in toluene (30 mL) with stirring and heated under Dean-Stark apparatus until the condensate ran clear. The solvent was removed using a rotary evaporator and the crude mixture was dissolved in MeOH (60 mL) and diethyl ether (30 mL) and NaBH₄ (1.90 g, 50.3 mmol, 3 equiv.) was added slowly at 0 °C with stirring. This mixture was stirred for 2 hours before the solution was neutralised using concentrated aqueous HCl. This solution was washed with EtOAc (5 × 30 mL) and all the organic extracts were combined and washed with brine (1 × 20 mL) and then dried (MgSO₄). The crude oil was purified via flash chromatography on silica eluting with a gradient from 0% to 10% EtOAc in hexane to yield a yellow oil (8.07 g, 87%).

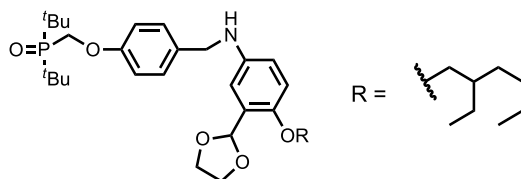
¹H NMR (250 MHz, CDCl₃): δ_H = 7.21 (d, 2H, *J* = 8.5), 6.89 (d, 1H, *J* = 3.0), 6.85 (d, 2H, *J* = 9.0), 6.78 (d, 1H, *J* = 9.0), 6.60 (dd, 1H, *J* = 9.0, 3.0), 6.14 (s, 1H), 4.20 (s, 2H), 3.97 – 4.14 (m, 4H), 3.82 (d, 2H, *J* = 6.0), 1.74 (s, 1H), 1.20 – 1.56 (m, 11H), 1.11 (d, 18H, *J* = 7.0), 0.87 – 0.98 (m, 6H);

¹³C NMR (63 MHz, CDCl₃): δ_C = 155.2, 150.2, 142.2, 131.9, 128.8, 127.1, 119.9, 114.4, 113.9, 112.1, 99.4, 72.0, 65.1, 48.9, 39.5, 30.5, 29.1, 23.9, 23.1, 17.9, 14.0, 12.7, 11.1;

MS (ES⁺): *m/z* (%) = 556 (100) [M + H⁺];

HRMS (ES⁺): calculated for C₃₃H₅₄NO₄Si 556.3822, found 556.3800;

FT-IR (thin film): ν_{max}/cm⁻¹ 2928, 2866, 1608, 1506, 1465, 1263, 1227, 1169, 1069, 913.

3.5.1.9 Synthesis of **14**

3.6 (0.79 g, 2.7 mmol, 1 equiv.) and **3.10** (1.02 g, 3.5 mmol, 1.3 equiv.) were dissolved in toluene (50 mL) with stirring and heated under Dean-Stark apparatus until the condensate ran clear. The solvent was removed with a rotary evaporator and the crude mixture was dissolved in MeOH (20 mL) and NaBH₄ (0.61 g, 16.0 mmol, 6 equiv.) was added slowly at 0 °C with stirring. After 2 hours the solution was neutralised using concentrated aqueous HCl. This solution was washed with EtOAc (5 × 20 mL) and all the organic extracts were combined and washed with brine (1 × 20 mL) and then dried (MgSO₄) and the solvent removed with a rotary evaporator. The crude oil was purified via flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in EtOAc to yield a yellow oil (0.87 g, 57%).

¹H NMR (250 MHz, CDCl₃): δ_H = 7.25 (d, 2H, *J* = 8.5), 6.82 – 6.85 (m, 3H), 6.70 (d, 1H, *J* = 9.0), 6.51 (dd, 1H, *J* = 9.0, 3.0), 6.07 (s, 1H), 4.32 (d, 2H, *J* = 7.5), 4.16 (s, 2H), 3.90 – 4.05 (m, 4H), 3.75 (d, 2H, *J* = 6.0), 1.63 – 1.71 (m, 1H), 1.20 – 1.50 (m, 26H), 0.83 – 0.91 (m, 6H);

¹³C NMR (63 MHz, CDCl₃): δ_C = 157.2 (d, *J* = 10.6), 149.7, 141.8, 132.9, 128.6, 126.7, 113.8, 113.8, 113.5, 111.8, 99.0, 71.6, 64.8, 62.3 (d, *J* = 71.0), 48.1, 39.2, 35.1 (d, *J* = 57.6), 30.2, 28.7, 26.1, 22.7, 13.8, 10.8;

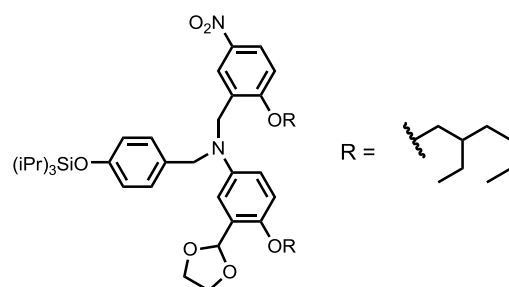
³¹P NMR (101 MHz, CDCl₃): δ_P = 56.6;

MS (ES⁺): *m/z* (%) = 574 (100) [M + H⁺], 1148 (15);

HRMS (ES⁺): calculated for C₃₃H₅₃NO₅P 574.3661, found 574.3642;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2956, 2927, 2871, 1691, 1601, 1506, 1468, 1392, 1226, 1136, 1069, 1040.

3.5.1.10 Synthesis of 3.13



3.11 (2.11 g, 3.8 mmol, 1 equiv.) and **3.8** (1.77 g, 6.3 mmol, 1.2 equiv.) were dissolved in DCE (20 mL) and NaBH(OAc)₃ (3.13 g, 2.8 mmol, 1.2 equiv.) was added with stirring. After 12 hours the reaction was quenched with saturated aqueous NaHCO₃ solution, and extracted into CHCl₃ (4 × 20 mL). All the organic fractions were washed with water (1 × 20 mL), brine (1 × 10 mL) and dried (MgSO₄) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% of EtOAc in hexane to yield a golden yellow oil (2.48 g, 79%).

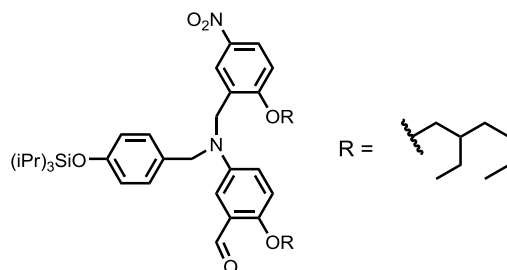
¹H NMR (400 MHz, CDCl₃): δ_H = 8.08 – 8.15 (m, 2H), 7.12 (d, 2H, *J* = 8.5), 6.97 (d, 1H, *J* = 3.0), 6.90 (d, 1H, *J* = 9.5), 6.81 (d, 2H, *J* = 8.5), 6.74 (d, 1H, *J* = 9.0), 6.60 (dd, 1H, *J* = 9.0, 3.0), 6.14 (s, 1H), 4.55 (s, 2H), 4.49 (s, 2H), 3.97 (d, 2H, *J* = 6.5), 3.94 (s, 4H), 3.80 (d, 2H, *J* = 6.0), 1.67 – 1.84 (m, 2H), 1.17 – 1.55 (m, 19H), 1.09 (d, 18H, *J* = 7.0); 0.86 – 0.97 (m, 12H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.8, 155.0, 149.9, 143.0, 141.40, 130.8, 128.80, 128.20, 126.90, 124.3, 123.5, 119.9, 114.8, 113.5, 112.3, 110.2, 99.6, 71.6, 71.3, 65.0, 55.1, 49.8, 39.4, 39.2, 30.5, 29.0, 23.9, 23.1, 22.9, 17.9, 14.1, 14.0, 12.6, 11.1;

MS (ES⁺): *m/z* (%) = 820 (100) [M + H⁺];

HRMS (ES⁺): calculated for C₄₈H₇₅N₂O₇Si 819.5344, found 819.5344;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2921, 2852, 1660, 1633, 1610, 1595, 1508, 1465, 1378, 1340, 1264, 1076.

3.5.1.11 Synthesis of **3.13'**

3.13 (1.39 g, 1.7 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 18 hours the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (4×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator to yield a yellow oil (1.30 g, 99%) requiring no further purification.

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 10.45$ (s, 1H), 8.14 (dd, 1H, $J = 9.0, 3.0$), 8.01 (d, 1H, $J = 2.5$), 7.17 (d, 1H, $J = 3.0$), 7.10 (d, 2H, $J = 8.5$), 6.89 – 6.94 (m, 2H), 6.80 – 6.87 (m, 3H), 4.57 (s, 2H), 4.54 (s, 2H), 3.99 (d, 2H, $J = 5.5$), 3.89 (d, 2H, $J = 5.5$), 1.69 – 1.84 (m, 2H), 1.17 – 1.59 (m, 19H), 1.09 (d, 18H, $J = 7.0$), 0.86 – 0.98 (m, 12H);

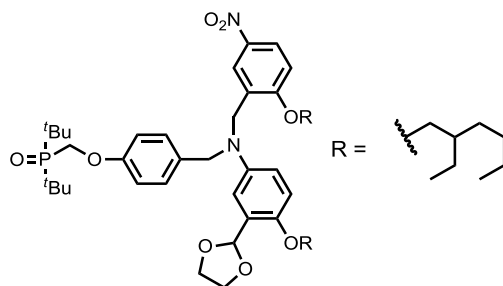
$^{13}\text{C NMR}$ (63 MHz, CDCl_3): $\delta_{\text{C}} = 190.0, 161.9, 155.1, 154.6, 142.9, 141.4, 130.2, 128.0, 127.9, 125.3, 124.5, 123.0, 121.2, 120.1, 114.2, 110.9, 110.4, 71.5, 71.3, 54.7, 49.6, 39.5, 39.2, 30.6, 30.5, 29.1, 29.0, 24.0, 23.9, 23.0, 22.9, 17.9, 14.0, 12.6, 11.2, 11.1$;

MS (ES+): m/z (%) = 263 (40), 775 (100) $[\text{M} + \text{H}^+]$;

HRMS (ES+): calculated for $\text{C}_{46}\text{H}_{71}\text{N}_2\text{O}_6\text{Si}$ 775.5081, found 775.5060;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2956, 2925, 2858, 1684, 1610, 1591, 1508, 1463, 1341, 1266.

3.5.1.12 Synthesis of 3.14



3.12 (0.160 g, 0.28 mmol, 1 equiv.) and **3.8** (0.156 g, 0.56 mmol, 2 equiv.) were dissolved in DCE (1 mL) and $\text{NaBH}(\text{OAc})_3$ (0.165 g, 0.78 mmol, 2.8 equiv.) was added with stirring. After 18 hours the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting from 0% to 3% of MeOH in EtOAc to yield a pale yellow oil (0.22 g, 94%).

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} = 8.12$ (dd, 1H, $J = 9.0, 3.0$), 8.10 (d, 1H, $J = 3.0$), 7.22 (d, 2H, $J = 9.0$), 6.97 (d, 1H, $J = 3.0$), 6.90 (d, 1H, $J = 9.0$), 6.86 (d, 2H, $J = 9.0$), 6.74 (d, 1H, $J = 9.0$), 6.61 (dd, 1H, $J = 9.0, 3.0$), 6.11 (s, 1H), 4.56 (s, 2H), 4.50 (s, 2H), (d, 2H, $J = 7.5$), 3.97 (d, 2H, $J = 6.0$), 3.93 – 3.99 (m, 4H), 3.79 (d, 1H, $J = 6.0$), 1.74 – 1.82 (m, 1H), 1.67 – 1.74 (m, 1H), 1.23 – 1.54 (m, 34H), 0.85 – 0.96 (m, 12H);

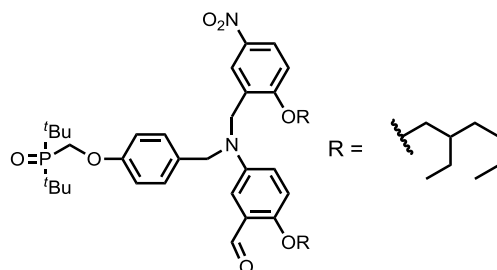
^{13}C NMR (126 MHz, CDCl_3): $\delta_{\text{C}} = 161.8, 157.4$, (d, $J = 10.6$), 150.1, 142.7, 141.4, 132.0, 128.6, 128.5, 126.9, 124.3, 123.5, 115.1, 114.1, 113.5, 112.6, 110.3, 99.6, 71.6, 71.4, 65.0, 62.6 (d, $J = 71.0$), 55.0, 50.2, 39.5, 39.2, 35.4 (d, $J = 57.6$), 30.5, 29.0, 26.4, 23.9, 23.0, 22.9, 14.0, 11.1;

^{31}P NMR (101 MHz, CDCl_3): $\delta_{\text{P}} = 56.8$;

MS (ES+): m/z (%) = 419 (10) [M + 2H⁺], 838 (100) [M + H⁺], 860 (45) [M + Na⁺], 883 (15), 938 (30);

HRMS (ES+): calculated for C₄₈H₇₄N₂O₈P 837.5183, found 837.5213;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2922, 2853, 1738, 1610, 1592, 1508, 1466, 1340, 1265, 1231, 1139, 1078.

3.5.1.13 Synthesis of **3.14'**

3.14 (0.22 g, 0.26 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 18 hours the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) and dried (MgSO_4) and the solvent removed with a rotary evaporator to yield a yellow oil (0.20 g, 94%) requiring no further purification.

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 10.43$ (s, 1H), 8.12 (dd, 1H, $J = 9.0, 3.0$), 7.97 (d, 1H, $J = 3.0$), 7.19 (d, 2H, $J = 8.5$), 7.14 (d, 1H, $J = 3.0$), 6.82 – 6.95 (m, 5H), 4.59 (s, 2H), 4.55 (s, 2H), 4.35 (d, 2H, $J = 7.5$), 3.99 (d, 2H, $J = 5.5$), 3.87 (d, 2H, $J = 5.5$), 1.66 – 1.83 (m, 2H), 1.21 – 1.56 (m, 34H), 0.83 – 0.96 (m, 12H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 189.9, 161.8, 157.4$ (d, $J = 11.6$), 154.6, 142.5, 141.2, 131.1, 128.1, 127.6, 125.2, 124.5, 122.9, 121.0, 114.2, 114.1, 110.8, 110.4, 71.4, 71.1, 62.5 (d, $J = 70.9$), 54.4, 49.8, 39.4, 39.1, 35.3 (d, $J = 57.0$), 30.5, 30.4, 29.0, 28.9, 26.4, 23.9, 23.8, 22.9, 22.8, 14.0, 13.9, 11.1, 11.0;

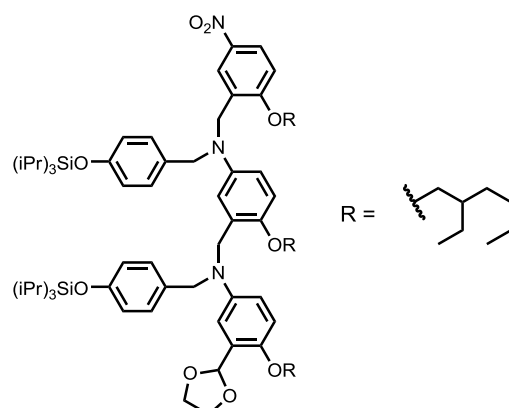
^{31}P NMR (162 MHz, CDCl_3): $\delta_{\text{P}} = 56.6$;

MS (ES⁺): m/z (%) = 397 (5) [$\text{M} + 2\text{H}^+$], 794 (100) [$\text{M} + \text{H}^+$], 816 (65) [$\text{M} + \text{Na}^+$], 839 (25), 864 (50);

HRMS (ES⁺): calculated for $\text{C}_{46}\text{H}_{70}\text{N}_2\text{O}_7\text{P}$ 793.4921, found 793.4921;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2927, 2871, 1610, 1591, 1507, 1465, 1338, 1265, 1227, 1137, 1047, 1017.

3.5.1.14 Synthesis of 3.15



3.13' (2.02 g, 2.61 mmol, 1 equiv.) and **3.11** (2.90 g, 5.22 mmol, 2 equiv.) were dissolved in DCE (9 mL) and $\text{NaBH}(\text{OAc})_3$ (1.55 g, 7.31 mmol, 2.8 equiv.) was added with stirring. After 1 day the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent was removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% of a 1:1 CHCl_3 /diethyl ether mix in hexane) to yield a golden yellow oil (1.86 g, 54%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 7.99$ (dd, 1H, $J = 9.0, 3.0$), 7.92 (d, 1H, $J = 3.0$), 7.07 (d, 2H, $J = 8.5$), 6.99 (d, 2H, $J = 8.5$), 6.81 (d, 2H, $J = 8.5$), 6.76 (d, 2H, $J = 8.5$), 6.72 (d, 1H, $J = 5.5$), 6.67 – 6.70 (m, 2H), 6.49 – 6.54 (m, 2H), 6.37 (d, 1H, $J = 3.0$), 6.30 (dd, 1H, $J = 9.0, 3.0$), 6.10 (s, 1H), 4.49 (s, 2H), 4.39 – 4.45 (m, 4H), 4.18 (s, 2H), 3.85 – 3.93 (m, 6H), 3.73 – 3.80 (m, 4H), 1.68 – 1.79 (m, 2H), 1.60 – 1.68 (m, 1H), 1.16 – 1.54 (m, 27H), 1.08 (dd, $J = 7.0, 2.0$, 31H), 0.81 – 1.00 (m, 16H);

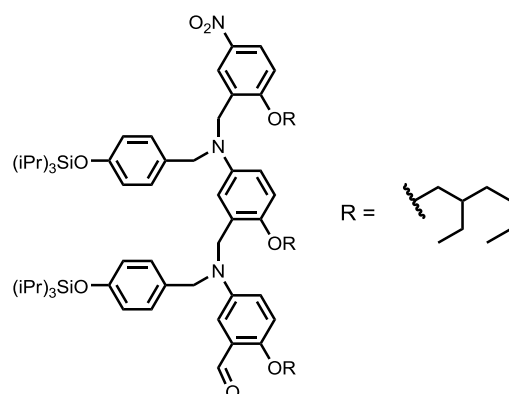
$^{13}\text{C NMR}$ (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.4, 154.9, 154.6, 149.1, 148.9, 143.1, 142.5, 141.2, 131.5, 131.3, 128.4, 128.1, 127.6, 127.2, 126.4, 124.2, 122.9, 119.9, 119.7, 113.4, 113.3, 112.9, 112.2, 111.3, 110.8, 110.4, 99.7, 71.7, 71.3, 70.7, 64.9, 55.6, 53.9, 50.1,$

49.6, 39.6, 39.5, 39.2, 30.7, 30.5, 30.6, 29.1, 29.0, 24.0, 23.9, 23.1, 23.0, 17.9, 14.1, 14.0, 12.7, 11.2;

MS (ES+): m/z (%) = 1315 (80) [M + H⁺], 1337 (100) [M + Na⁺];

HRMS (ES+): calculated for C₇₉H₁₂₄N₃O₉Si₂ 1314.8876, found 1314.8842;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2959, 2928, 2867, 1608, 1593, 1507, 1464, 1383, 1339, 1263, 1226, 1165, 1075, 1014.

3.5.1.15 Synthesis of **3.15'**

3.15 (0.68 g, 0.52 mmol, 1 equiv.) was dissolved in $CHCl_3$ (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 18 hours the mixture was neutralised using aqueous $NaHCO_3$ and the organic portion separated from the aqueous part. The aqueous layer was washed with $CHCl_3$ (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried ($MgSO_4$) and the solvent removed with a rotary evaporator to yield a yellow oil (0.65 g, 98%) requiring no further purification.

1H NMR (400 MHz, $CDCl_3$): $\delta_H = 10.43$ (s, 1H), 7.97 (dd, 1H, $J = 9.0, 2.5$), 7.87 (d, 1H, $J = 2.5$), 7.07 (d, 2H, $J = 8.5$), 6.98 (d, 2H, $J = 8.5$), 6.96 (d, 1H, $J = 3.0$), 6.82 (d, 2H, $J = 8.5$), 6.78 (d, 2H, $J = 8.5$), 6.73 (d, 1H, $J = 9.0$), 6.71 (d, 2H, $J = 9.0$), 6.61 (s, 1H), 6.54 – 6.60 (m, 1H), 6.28 (d, 1H, $J = 3.0$), 4.50 (s, 1H), 4.44 (s, 1H), 4.40 (s, 1H), 4.23 (s, 1H), 3.88 (d, 1H, $J = 5.5$), 3.86 (d, 1H, $J = 5.5$), 3.78 (d, 1H, $J = 5.5$), 1.62 – 1.82 (m, 3H), 1.16 – 1.58 (m, 27H), 1.02 – 1.15 (m, 36H), 0.82 – 1.01 (m, 18H);

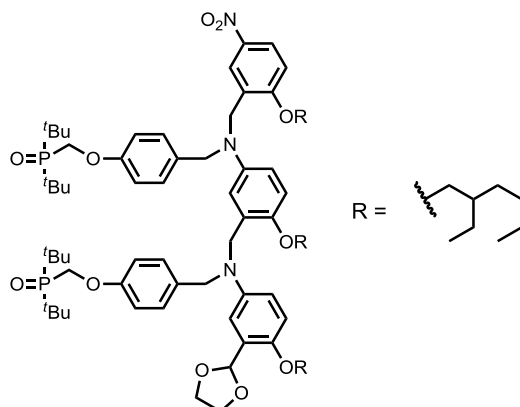
^{13}C NMR (101 MHz, $CDCl_3$): $\delta_C = 190.1, 161.4, 155.0, 154.8, 153.7, 149.1, 142.9, 142.5, 141.2, 131.0, 130.7, 128.4, 128.0, 127.6, 126.4, 124.9, 124.0, 122.9, 120.1, 119.9, 113.6, 112.5, 112.2, 111.3, 110.1, 109.4, 71.4, 71.3, 70.7, 55.4, 53.8, 50.2, 49.2, 39.6,$

39.5, 39.2, 30.7, 30.6, 30.5, 29.1, 29.0, 24.0, 23.9, 23.1, 23.0, 22.9, 17.9, 14.1, 12.7, 12.6, 11.2, 11.1;

MS (ES+): m/z (%) = 1271 (100) [M + H⁺], 1293 (15) [M + Na⁺];

HRMS (ES+): calculated for C₇₇H₁₂₀N₃O₈Si₂ 1270.8614, found 1270.8586;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2959, 2925, 2866, 1682, 1608, 1593, 1507, 1463, 1263, 1226, 1165, 1013.

3.5.1.16 Synthesis of **3.16**

3.14' (0.190 g, 0.24 mmol, 1 equiv.) and **3.12** (0.206 g, 0.36 mmol, 1.5 equiv.) were dissolved in DCE (900 μ L) and $\text{NaBH}(\text{OAc})_3$ (0.142 g, 0.67 mmol, 2.8 equiv.) was added with stirring. After 2 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting 0% to 10% MeOH in diethyl ether to yield a pale yellow oil (0.208 g, 65%).

$^1\text{H NMR}$ (400 MHz, CD_3CN): δ_{H} = 7.98 (dd, 1H, J = 9.0, 3.0), 7.75 (d, 1H, J = 3.0), 7.10 (d, 2H, J = 9.0), 7.06 (d, 2H, J = 8.0), 6.87 – 6.97 (m, 5H), 6.74 (d, 1H, J = 9.0), 6.64 (d, 1H, J = 3.0), 6.58 (d, 1H, J = 9.0), 6.54 (dd, 1H, J = 9.0, 3.0), 6.28 – 6.35 (m, 2H), 5.93 (s, 1H), 4.48 (s, 2H), 4.45 (s, 2H), 4.40 (s, 2H), 4.34 (d, 2H, J = 3.0), 4.33 (d, 2H, J = 3.0), 4.17 (s, 2H), 3.96 (d, 2H, J = 5.5), 3.83 (br. d., 4H, J = 2.0), 3.72 – 3.78 (m, 4H), 1.55 – 1.78 (m, 3H), 1.16 – 1.53 (m, 60H), 0.78 – 0.95 (m, 18H);

$^{13}\text{C NMR}$ (101 MHz, CD_3CN): δ_{C} = 162.8, 158.7 (d, J = 8.5), 158.6 (d, J = 8.5), 150.1, 143.6, 143.1, 142.0, 133.5, 133.4, 129.4, 129.2, 129.1, 128.2, 127.70, 125.2, 123.5, 115.4, 115.3, 114.7, 114.4, 114.1, 113.8, 112.7, 112.2, 100.0, 72.5, 72.4, 71.8,

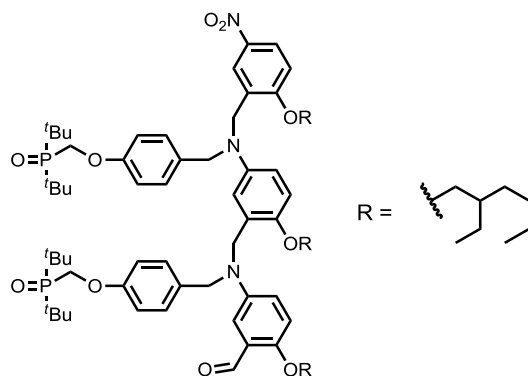
65.8, 63.6 (d, $J = 69.4$), 63.5 (d, $J = 70.1$), 56.1, 55.1, 51.2, 50.6, 40.6, 40., 40.2, 36.0 (d, $J = 57.8$), 31.5, 31.4, 30.8, 29.9, 26.9, 24.9, 24.8, 23.9, 23.8, 14.6, 14.5, 11.7;

^{31}P NMR (162 MHz, CD_3CN): $\delta_{\text{p}} = 55.3, 55.3$;

MS (ES+): m/z (%) = 676 (50) $[\text{M} + 2\text{H}^+]$, 1351 (100) $[\text{M} + \text{H}^+]$, 1373 (100) $[\text{M} + \text{Na}^+]$, 1396 (40);

HRMS (ES+): calculated for $\text{C}_{79}\text{H}_{122}\text{N}_3\text{O}_{11}\text{P}_2$ 1350.8555, found 1350.8549;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2927, 2871, 1680, 1610, 1592, 1505, 1465, 1432, 1392, 1264, 1225, 1174, 1135, 1044, 1017, 968.

3.5.1.17 Synthesis of **3.16**'

3.16 (0.21 g, 0.16 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 3 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator to yield a yellow oil (0.20 g, 98%) requiring no further purification.

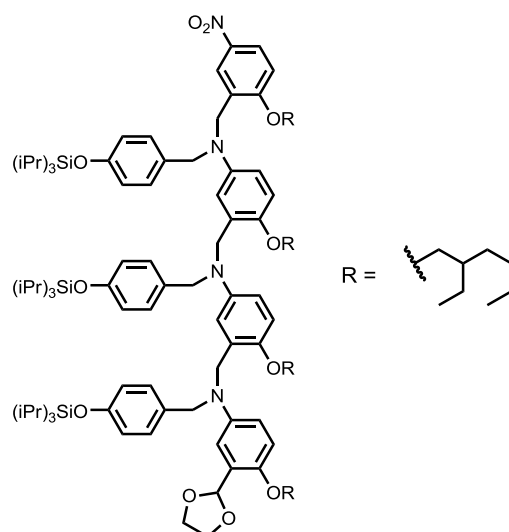
^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 10.42$ (s, 1H), 7.96 (d, 1H, $J = 9.0$), 7.84 (d, 1H, $J = 3.0$), 7.16 (d, 2H, $J = 9.0$), 7.06 (d, 2H, $J = 9.0$), 6.79 – 6.93 (m, 5H), 6.72 (dd, 2H, $J = 9.0, 3.5$), 6.64 (d, 1H, $J = 9.5$), 6.60 (dd, 1H, $J = 9.0, 3.0$), 6.51 – 6.57 (m, 1H), 6.26 (d, 1H, $J = 3.0$), 4.51 (s, 2H), 4.45 (s, 2H), 4.43 (s, 1H), 4.32 – 4.39 (m, 4H), 4.22 (s, 2H), 3.88 (d, 2H, $J = 5.5$), 3.86 (d, 2H, $J = 5.5$), 3.77 (d, 2H, $J = 5.5$), 1.75 (d, 2H, $J = 7.0$), 1.60 – 1.69 (m, 1H), 1.18 – 1.59 (m, 60H), 0.80 – 0.99 (m, 15H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 189.9, 161.3, 157.3$ (d, $J = 11.2$), 157.2 (d, $J = 10.8$), 153.80, 149.10, 142.6, 142.1, 141.1, 132.1, 131.6, 128.1, 128.0, 127.7, 126.2, 124.8, 124.0, 122.7, 120.0, 114.1, 113.6, 112.4, 112.2, 111.3, 110.1, 109.2, 71.3, 71.2, 70.5, 62.6 (d, $J = 70.1$), 62.5 (d, $J = 70.1$), 55.3, 53.6, 50.3, 49.2, 39.5, 39.4, 39.1, 35.3

(d, $J = 57.8$), 30.5, 30.4, 30.3, 30.2, 29.0, 28.9, 26.4, 23.9, 23.8, 22.9, 22.8, 14.0, 13.9, 11.1;

^{31}P NMR (101 MHz, CDCl_3): $\delta_{\text{p}} = 56.8$;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2928, 2872, 1683, 1592, 1508, 1466, 1339, 1265, 1229, 1176, 1138, 1081, 1047, 1014, 968.

3.5.1.18 Synthesis of **3.17**

3.15' (0.885 g, 0.70 mmol, 1 equiv.) and **3.11** (0.774 g, 1.4 mmol, 2 equiv.) were dissolved in DCE (0.5 mL) and $NaBH(OAc)_3$ (0.413 g, 2.0 mmol, 2.8 equiv.) was added with stirring. After 2 days the reaction was quenched with saturated aqueous $NaHCO_3$ solution, and extracted into $CHCl_3$ (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried ($MgSO_4$) before the solvent removed on a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 5% of diethyl ether in hexane to yield a golden yellow oil (1.00 g, 79%).

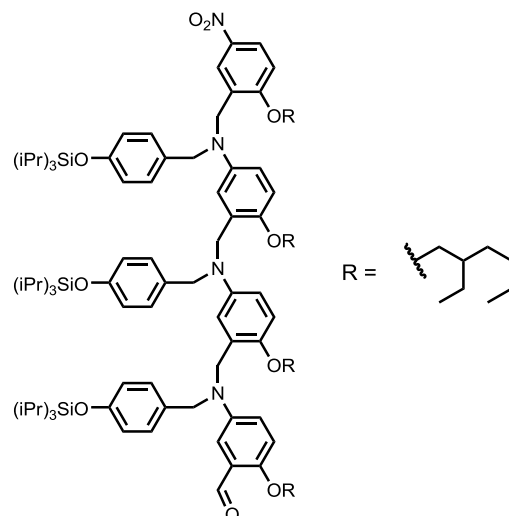
1H NMR (500 MHz, $CDCl_3$): $\delta_H = 7.98 - 8.02$ (m, 2H), 7.05 (d, 2H, $J = 8.5$), 7.00 (d, 2H, $J = 8.5$), 6.96 (d, 2H, $J = 8.5$), 6.89 (d, 1H, $J = 3.0$), 6.82 (d, 2H, $J = 8.5$), 6.79 (d, 2H, $J = 8.5$), 6.77 (d, 2H, $J = 8.5$), 6.72 (d, 1H, $J = 4.0$), 6.71 (d, 1H, $J = 4.0$), 6.69 (d, 1H, $J = 9.0$), 6.59 (d, 1H, $J = 9.0$), 6.51 – 6.55 (m, 2H), 6.49 (d, 1H, $J = 2.9$), 6.44 (d, 1H, $J = 3.0$), 6.27 – 6.30 (m, 1H), 6.18 (s, 1H), 4.50 (s, 2H), 4.48 (s, 2H), 4.46 (s, 2H), 4.44 (s, 2H), 4.10 (s, 2H), 4.07 (s, 2H), 3.92 (s, 4H), 3.79 – 3.83 (m, 6H), 3.77 (d, 2H, $J = 5.5$), 1.72 (d, 4H, $J = 6.2$), 1.21 – 1.57 (m, 35H), 1.04 – 1.18 (m, 54H), 0.86 – 0.99 (m, 24H);

¹³C NMR (63 MHz, CDCl₃): δ_c = 161.4, 154.9, 154.5, 149.0, 148.9, 148.4, 143.3, 143.1, 142.6, 141.3, 131.8, 131.6, 131.7, 128.7, 128.2, 127.7, 127.6, 127.0, 126.8, 124.1, 123.2, 119.8, 119.7, 119.6, 114.2, 113.6, 113.3, 112.5, 112.2, 112.1, 111.6, 111.4, 110.9, 110.3, 99.8, 71.9, 71.3, 70.8, 70.7, 64.9, 55.4, 54.3, 53.7, 50.5, 50.2, 49.7, 39.6, 39.5, 39.2, 30.7, 30.6, 30.5, 29.1, 24.0, 23.9, 23.1, 23.0, 17.9, 14.1, 14.0, 12.6, 11.2, 11.1;

MS (ES+): m/z (%) = 1810 (80) [M + H⁺], 1811 (100), 1812 (70), 1813 (40);

HRMS (ES+): calculated for C₁₁₀H₁₇₃N₄O₁₁Si₃ 1810.2409, found 1810.2484;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2959, 2929, 2867, 1608, 1508, 1465, 1340, 1264, 914.

3.5.1.19 Synthesis of **3.17'**

3.17' (0.89 g, 0.49 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed using a rotary evaporator to yield a yellow oil (0.78 g, 90%) requiring no further purification.

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} = 10.46$ (s, 1H), 7.99 – 8.03 (m, 2H), 7.14 (d, 1H, $J = 2.0$), 7.03 (d, 2H, $J = 8.5$), 6.97 (d, 2H, $J = 8.5$), 6.92 (d, 2H, $J = 8.5$), 6.72 – 6.80 (m, 9H), 6.67 (d, 1H, $J = 8.5$), 6.59 (d, 1H, $J = 8.5$), 6.50 (dd, 1H, $J = 8.5, 3.0$), 6.46 (d, 1H, $J = 3.0$), 6.34 (d, 1H, $J = 3.0$), 6.30 (dd, 1H, $J = 9.0, 3.0$), 4.48 (s, 2H), 4.45 (s, 2H), 4.42 (s, 4H), 4.10 (s, 2H), 4.07 (s, 2H), 3.85 (d, 4H, $J = 5.5$), 3.79 (d, 2H, $J = 6.0$), 3.72 (d, 2H, $J = 6.0$), 1.67 – 1.77 (m, 3H), 1.57 – 1.67 (m, 2H), 1.18 – 1.56 (m, 41H), 1.04 – 1.16 (m, 54H), 0.84 – 0.96 (m, 24H);

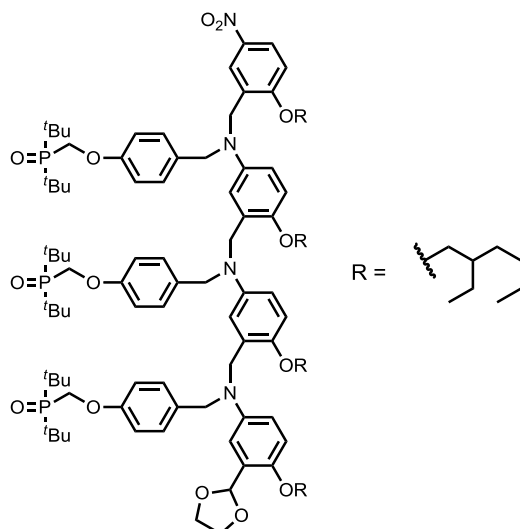
^{13}C NMR (63 MHz, CDCl_3): $\delta_{\text{C}} = 189.9, 161.4, 154.9, 154.7, 154.6, 153.7, 149.1, 148.5, 143.4, 143.0, 142.7, 141.3, 131.7, 131.1, 131.0, 128.8, 128.2, 127.7, 127.6, 127.5,$

126.2, 125.1, 124.0, 123.2, 120.7, 119.8, 119.7, 113.9, 113.2, 112.3, 112.2, 112.1, 111.4, 111.1, 110.3, 110.1, 71.4, 71.3, 70.8, 70.7, 55.4, 53.4, 50.5, 50.3, 49.6, 39.6, 39.5, 39.2, 30.7, 30.6, 30.5, 29.1, 29.0, 24.0, 23.9, 23.1, 23.0, 22.9, 17.9, 14.1, 14.0, 12.7, 12.6, 11.2, 11.1;

MS (ES+): m/z (%) = 1765 (40), 1766 (100) [M + H⁺], 1767 (95), 1768 (50), 1769 (20);

HRMS (ES+): calculated for C₁₀₈H₁₆₉N₄O₁₀Si₃ 1766.2147, found 1766.2076;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2959, 2929, 2867, 1683, 1608, 1507, 1464, 1340, 1264, 913.

3.5.1.20 Synthesis of **3.18**

3.16' (0.211 g, 0.16 mmol, 1 equiv.) and **3.12** (0.185 g, 0.32 mmol, 2.0 equiv.) were dissolved in DCE (1100 μ L) and $\text{NaBH}(\text{OAc})_3$ (0.096 g, 0.45 mmol, 2.8 equiv.) was added with stirring. After 4 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in EtOAc to yield a pale yellow oil (0.093 g, 31%).

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 7.98$ (dd, 1H, $J = 9.0, 3.0$), 7.94 (d, 1H, $J = 2.5$), 7.09 (d, 2H, $J = 8.5$), 7.02 (d, 2H, $J = 8.5$), 6.97 (d, 2H, $J = 8.5$), 6.87 (d, 1H, $J = 3.0$), 6.71 – 6.84 (m, 7H), 6.64 – 6.71 (m, 2H), (d, 1H, $J = 9.0$), 6.43 – 6.50 (m, 3H), 6.40 (d, 1H, $J = 2.5$), (dd, 1H, $J = 8.5, 2.5$), 6.10 (s, 1H), 4.42 (br. s., 8H), 4.30 – 4.37 (m, 6H), 4.03 (br. s., 4H), 3.92 (d, 4H, $J = 3.3$), 3.83 (d, 2H, $J = 5.5$), 3.70 – 3.78 (m, 6H), 1.57 – 1.75 (m, 4H), 1.16 – 1.53 (m, 86H), 0.77 – 0.94 (m, 24H);

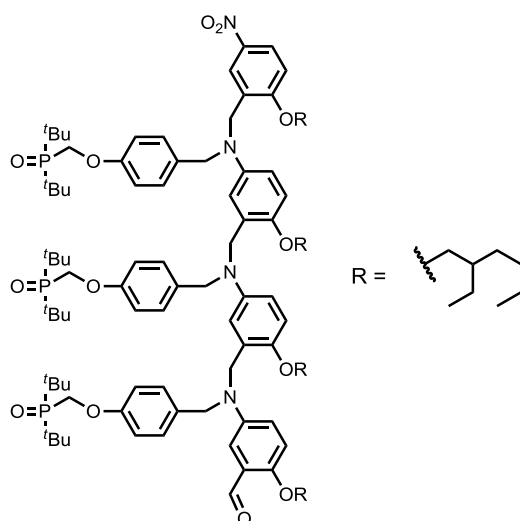
^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.3, 157.2, 157.1$, (d, $J = 10.8$), 149.1, 149.0, 148.5, 143.1, 142.8, 142.2, 141.1, 132.8, 132.5, 132.1, 128.4, 128.3, 128.0, 127.3, 126.7,

126.6, 124.0, 123.0, 114.0, 113.9, 113.8, 113.6, 113.3, 112.5, 112.1, 111.6, 111.4, 110.9, 110.2, 99.5, 71.8, 71.2, 70.7, 70.6, 64.8, 62.5 (d, $J = 71$), 55.1, 54.1, 53.3, 50.3, 49.4, 39.4, 39.3, 39.0, 35.2, 35.3 (d, $J = 58.0$), 30.5, 30.4, 30.3, 28.9, 23.8, 23.7, 22.9, 22.8, 14.0, 13.9, 11.0;

^{31}P NMR (162 MHz, CDCl_3): $\delta_{\text{p}} = 56.6, 56.6, 56.5$;

MS (ES+): m/z (%) = 1862 (50), 1863 (100) $[\text{M} + \text{H}^+]$, 1864 (100), 1865 (50), 1866 (20);

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2926, 2871, 1610, 1591, 1506, 1467, 1339, 1265, 1225, 1142, 965.

3.5.1.21 Synthesis of **3.18'**

3.18 (0.093 g, 0.051 mmol, 1 equiv.) was dissolved in CHCl_3 (5 mL) and concentrated aqueous HCl (5 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator to yield a yellow oil (0.082 g, 88%) requiring no further purification.

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 10.42$ (s, 1H), 8.01 (dd, 1H, $J = 9.0, 2.5$), 7.94 (d, 1H, $J = 2.5$), 7.10 (d, 2H, $J = 8.5$), 6.95 – 7.07 (m, 5H), 6.76 – 6.83 (m, 7H), 6.73 (d, 2H, $J = 1.0$), 6.65 (d, 1H, $J = 9.0$), 6.60 (d, 1H, $J = 9.0$), 6.43 – 6.50 (m, 2H), 6.28 – 6.34 (m, 2H), 4.41 – 4.47 (m, 6H), 4.30 – 4.40 (m, 9H), 4.08 (br. s., 4H), 3.88 (d, 2H, $J = 5.5$), 3.83 (d, 2H, $J = 5.5$), 3.77 (d, 2H, $J = 5.5$), 3.67 (d, 2H, $J = 5.5$), 1.63 – 1.78 (m, 3H), 1.53 – 1.63 (m, 1H), 1.17 – 1.52 (m, 86H), 0.78 – 0.95 (m, 24H);

$^{13}\text{C NMR}$ (101 MHz, CDCl_3): $\delta_{\text{C}} = 189.8, 161.4, 157.3$ (d, $J = 10.8$), 157.1 (d, $J = 11.5$), 157.0 (d, $J = 11.0$), 153.8, 149.2, 148.6, 143.2, 142.7, 142.3, 141.2, 132.7, 132.2, 131.8, 128.5, 128.4, 127.9, 127.8, 127.2, 126.0, 125.1, 124.1, 123.1, 120.4, 114.0,

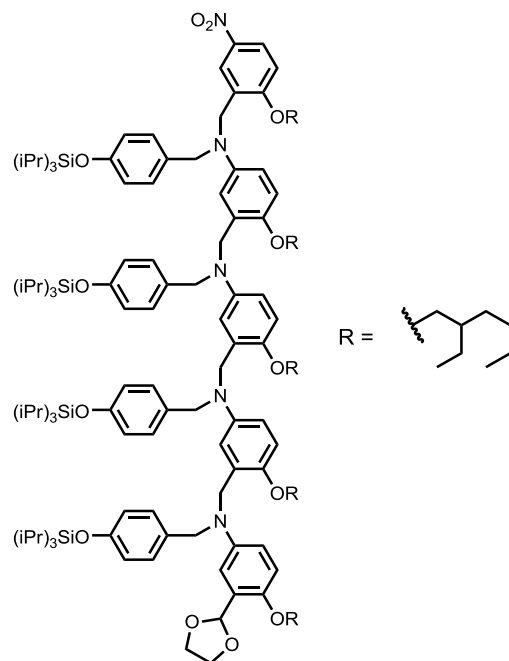
113.9, 113.3, 112.4, 112.2, 112.1, 111.5, 111.2, 110.2, 110.1, 71.4, 71.3, 70.8, 70.7, 62.6 (d, $J = 70.1$), 62.5 (d, $J = 70.1$), 55.3, 54.4, 53.2, 50.4, 50.3, 49.5, 39.5, 39.4, 39.1, 35.4 (d, $J = 57.8$), 30.6, 30.4, 29.6, 29.0, 28.9, 26.4, 23.9, 23.8, 23.0, 22.9, 22.8, 14.0, 11.1, 11.0;

^{31}P NMR (162 MHz, CDCl_3): $\delta_{\text{p}} = 56.8, 56.7, 56.6$;

MS (ES+): m/z (%) = 1820 (80) $[\text{M} + \text{H}^+]$, 1821 (100), 1842 (80) $[\text{M} + \text{Na}^+]$, 1843 (100);

HRMS (ES+): calculated for $\text{C}_{108}\text{H}_{166}\text{N}_4\text{O}_{13}\text{P}_3$ 1820.1664, found 1820.1619;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2921, 2852, 1683, 1508, 1466, 1233, 1138.

3.5.1.22 Synthesis of **3.19**

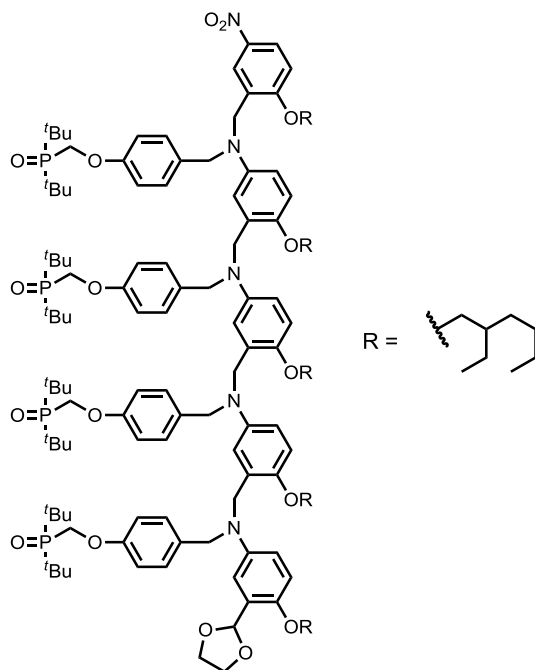
3.17' (0.815 g, 0.06 mmol, 1 equiv.) and **3.11** (0.713 g, 0.13 mmol, 2 equiv.) were dissolved in DCE (0.5 mL) and $\text{NaBH}(\text{OAc})_3$ (0.038 g, 0.13 mmol, 2.8 equiv.) was added with stirring. After 2 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 3% to 5% of diethyl ether in hexane to yield a golden yellow oil (0.036 g, 26%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 7.96 - 8.01$ (m, 2H), 7.03 (d, 2H, $J = 8.5$), 6.97 (d, 2H, $J = 8.5$), 6.89 (d, 4H, $J = 8.5$), 6.64 – 6.81 (m, 11H), 6.60 – 6.64 (m, 2H), 6.46 – 6.57 (m, 6H), 6.43 (dd, 1H, $J = 9.0, 2.5$), 6.30 (dd, 1H, $J = 9.0, 3.0$), 6.15 (s, 1H), 4.45 (s, 2H), 4.42 (s, 2H), 4.40 (s, 4H), 4.38 (s, 2H), 4.10 (s, 2H), 4.06 (s, 2H), 4.04 (s, 2H), 3.79 – 3.87 (m, 6H), 3.70 – 3.77 (m, 8H), 1.58 – 1.75 (m, 5H), 1.15 – 1.54 (m, 52H), 0.99 – 1.15 (m, 72H), 0.81 – 0.96 (m, 30H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.4, 154.9, 154.5, 154.4, 149.2, 148.9, 148.6, 148.5, 143.4, 143.3, 142.8, 141.3, 132.1, 131.7, 131.1, 128.8, 128.3, 127.9, 127.8, 127.7, 127.3, 126.9, 126.8, 125.5, 124.0, 123.3, 119.8, 119.7, 119.6, 114.0, 113.6, 113.5, 113.3, 113.1, 112.3, 112.2, 111.8, 111.6, 111.5, 111.4, 110.1, 99.9, 71.8, 71.3, 70.9, 70.7, 64.8, 55.4, 54.5, 54.3, 53.6, 50.9, 50.4, 50.0, 49.9, 39.6, 39.5, 39.1, 30.6, 30.5, 30.3, 29.2, 29.1, 29.0, 24.0, 24.9, 23.9, 23.1, 23.0, 17.9, 14.1, 14.0, 12.6, 11.2, 11.1;

MS (ES+): m/z (%) = 2305 (50), 2306 (100) [M + H⁺], 2307 (95), 2308 (55), 2309 (20);

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2927, 2859, 1680, 1610, 1592, 1500, 1464, 1394, 1338, 1265, 1249, 1231, 1170, 1081, 1013, 970.

3.5.1.23 Synthesis of **3.20**

3.18' (0.106 g, 0.06 mmol, 1 equiv.) and **3.12** (0.12 g, 0.21 mmol, 3.5 equiv.) were dissolved in DCE (0.5 mL) and $\text{NaBH}(\text{OAc})_3$ (0.035 g, 0.16 mmol, 2.8 equiv.) was added with stirring. After 1 week, more $\text{NaBH}(\text{OAc})_3$ (0.035 g, 0.16 mmol, 2.8 equiv.) was added to the reaction. After another week the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% of MeOH in a 25:75 mixture of $\text{CHCl}_3/\text{EtOAc}$ and then 10% MeOH in neat CHCl_3 to yield a golden yellow oil (0.021 g, 15%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 7.99$ (dd, 1H, $J = 9.0, 2.5$), 7.93 (d, 1H, $J = 2.5$), 7.11 (d, 1H, $J = 8.5$), 7.02 (d, 1H, $J = 8.5$), 6.96 (d, 1H, $J = 8.5$), 6.92 (d, 1H, $J = 8.5$), 6.70 – 6.85 (m, 9H), 6.64 – 6.69 (m, 2H), 6.57 – 6.63 (m, 2H), 6.55 (d, 1H, $J = 2.5$), 6.44 – 6.51 (m, 4H), 6.41 (dd, 1H, $J = 9.0, 3.0$), 6.33 (dd, 1H, $J = 9.0, 3.0$), 6.10 (s,

1H), 4.40 (br. s., 9H), 4.29 – 4.36 (m, 9H), 4.11 (s, 2H), 4.01 (d, 4H, $J = 5.0$), 3.90 (s, 4H), 3.85 (d, 2H, $J = 5.5$), 3.69 – 3.77 (m, 8H), 1.56 – 1.76 (m, 5H), 1.18 – 1.49 (m, 112H), 0.79 – 0.93 (m, 30H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.4, 157.3$ (d, $J = 11.0$), 157.0 (d, $J = 11.0$), 156.9 (d, $J = 11.0$), $149.4, 149.2, 148.7, 143.3, 143.2, 143.1, 142.5, 141.3, 133.1, 132.7, 132.6, 132.2, 128.5, 128.2, 128.1, 127.6, 127.1, 126.7, 124.1, 123.3, 114.1, 113.9, 113.7, 113.6, 113.3, 113.2, 112.4, 112.3, 111.8, 111.7, 111.6, 110.2, 99.7, 72.0, 71.4, 71.0, 70.9, 70.8, 65.0, 62.6$ (d, $J = 71.0$), $55.3, 54.4, 54.1, 53.2, 50.8, 50.3, 50.1, 49.5, 39.5, 39.4, 39.1, 35.4$ (d, $J = 58.0$), $30.6, 30.5, 30.4, 29.7, 29.1, 29.0, 26.4, 23.9, 23.8, 23.0, 22.9, 14.1, 14.0, 11.2, 11.1$;

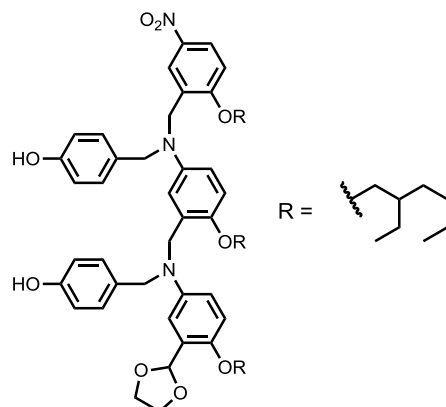
^{31}P NMR (162 MHz, CDCl_3): $\delta_{\text{P}} = 56.8, 56.7$;

MS (ES+): m/z (%) = 2379 (100), 2395 (60), 2378 (30) $[\text{M} + \text{H}^+]$, 2379 (50), 2395 (30);

HRMS (ES+): calculated for $\text{C}_{141}\text{H}_{217}\text{N}_5\text{O}_{17}\text{P}_4$ 2399.5112, found 2399.5032;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2920, 2851, 1659, 1632, 1506, 1470, 1427, 1243, 1137.

3.5.1.24 Synthesis of 3.21



3.15 (0.321 g, 0.24 mmol, 1 equiv.) was dissolved in THF (2 mL) at 0 °C and TBAF (720 μ L, 0.72 mmol, 3 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4×10 mL). All organic fractions were combined and washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 20% to 60% of diethyl ether in hexane to yield a light red oil (0.189 g, 77%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.09$ (dd, 1H, $J = 9.0, 2.5$), 8.01 (d, 1H, $J = 2.5$), 7.05 (d, 2H, $J = 8.5$), 6.93 (d, 2H, $J = 8.5$), 6.87 (d, 1H, $J = 9.0$), 6.61 – 6.77 (m, 4H), 6.51 – 6.60 (m, 3H), 6.37 – 6.46 (m, 2H), 6.28 (dd, 1H, $J = 9.0, 3.0$), 6.09 (s, 1H), 4.54 (s, 2H), 4.44 (s, 2H), 4.41 (s, 2H), 4.29 (s, 2H), 3.91 – 4.06 (m, 6H), 3.80 (d, 2H, $J = 6.0$), 3.75 (d, 2H, $J = 5.5$), 1.57 – 1.82 (m, 3H), 1.18 – 1.56 (m, 24H), 0.80 – 0.98 (m, 18H);

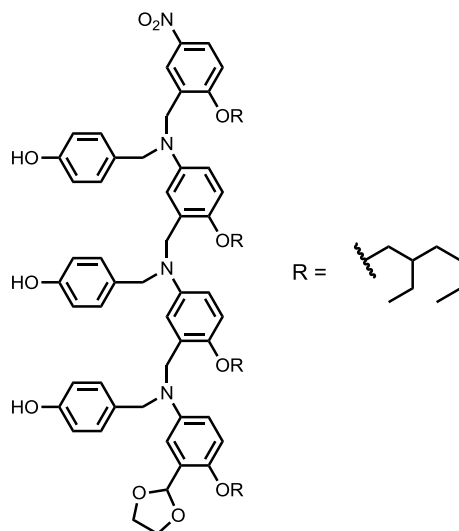
$^{13}\text{C NMR}$ (126 MHz, CDCl_3): $\delta_{\text{C}} = 161.7, 154.6, 154.4, 149.0, 143.1, 142.5, 141.3, 131.3, 130.5, 128.7, 128.1, 127.7, 127.4, 125.5, 124.3, 123.1, 115.8, 115.3, 113.7, 113.6, 113.5, 112.1, 111.3, 111.1, 110.4, 100.3, 71.8, 71.3, 70.7, 65.9, 65.0, 56.0, 54.5, 50.9, 49.7, 39.6, 39.5, 39.2, 30.6, 30.5, 29.1, 24.0, 23.9, 23.1, 23.0, 22.9, 14.1, 14.0, 11.1;$

MS (ES+): m/z (%) = 853 (10), 897 (35), 959 (10), 1003 (100) [$\text{M} + \text{H}^+$], 1025 (25);

HRMS (ES+): calculated for $C_{61}H_{84}N_3O_6$ 1002.6208, found 1002.6224;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2927, 2871, 1613, 1593, 1505, 1462, 1337, 1261, 1223, 1167, 1077, 1015, 964.

3.5.1.25 Synthesis of 3.22



3.17 (0.122 g, 0.07 mmol, 1 equiv.) was dissolved in THF (2 mL) at 0 °C and TBAF (210 μ L, 0.21 mmol, 3 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (3 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 5 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 20% to 70% of diethyl ether in hexane to yield a light red oil (0.087 g, 96%).

$^1\text{H NMR}$ (500 MHz, CDCl_3): δ_{H} = 8.11 (dd, 1H, J = 9.0, 3.0), 8.04 (d, 1H, J = 2.5), 6.98 (d, 2H, J = 8.5), 6.85 – 6.89 (m, 3H), 6.83 (d, 2H, J = 8.5), 6.81 (d, 1H, J = 3.5), 6.66 – 6.72 (m, 5H), 6.60 (d, 2H, J = 8.5), 6.47 – 6.51 (m, 2H), 6.37 – 6.45 (m, 4H), 6.34 (d, 1H, J = 3.0), 6.14 (s, 1H), 4.58 (s, 2H), 4.48 (s, 2H), 4.45 (s, 2H), 4.38 (s, 2H), 4.09 (s, 2H), 3.98 – 4.08 (m, 4H), 3.96 (d, 2H, J = 5.7), 3.87 (s, 2H), 3.81 (d, 2H, J = 5.7), 3.73 – 3.77 (m, 4H), 1.74 – 1.81 (m, 1H), 1.62 – 1.74 (m, 3H), 1.22 – 1.54 (m, 32H), 0.84 – 0.97 (m, 24H);

$^{13}\text{C NMR}$ (126 MHz, CDCl_3): δ_{C} = 161.7, 154.7, 154.2, 154.1, 149.2, 149.0, 148.1, 143.4, 143.2, 142.3, 141.4, 131.6, 131.4, 130.5, 129.0, 128.1, 127.9, 127.5, 127.4, 126.9,

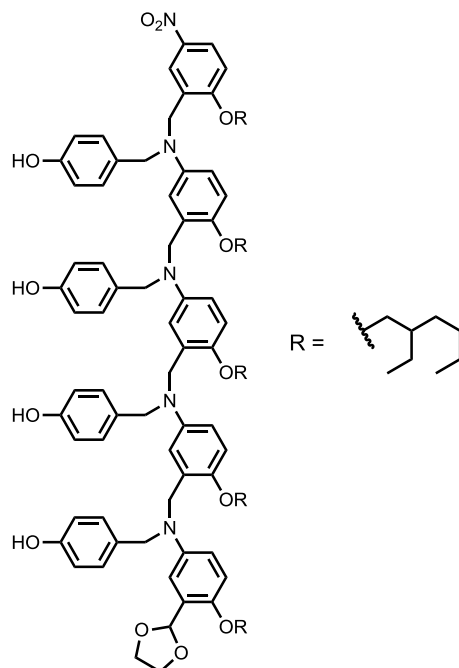
125.2, 124.3, 123.2, 115.6, 115.5, 115.3, 114.3, 113.8, 113.3, 112.6, 112.3, 112.0, 111.8, 110.8, 110.7, 110.3, 100.4, 71.7, 71.4, 71.0, 70.6, 65.0, 55.6, 54.9, 54.2, 51.3, 50.8, 49.5, 39.7, 39.5, 39.2, 30.7, 30.5, 29.1, 29.0, 24.0, 23.9, 23.8, 23.1, 23.0, 22.9, 14.1, 14.0, 11.2, 11.1, 11.0;

MS (ES+): m/z (%) = 565 (30), 618 (75), 671 (75), 1342 (100) [M + H⁺];

HRMS (ES+): calculated for C₈₃H₁₁₃N₄O₁₁ 1341.8406, found 1341.8342;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2956, 2918, 2850, 2161, 2030, 1616, 1504, 1338, 1261, 1022.

3.5.1.26 Synthesis of 3.23



3.19 (0.690 g, 0.30 mmol, 1 equiv.) was dissolved in THF (10 mL) at 0 °C and TBAF (2390 μ L, 2.4 mmol, 8 equiv.) was added with stirring. After 1 hour water (10 mL) was added and the aqueous mixture washed with diethyl ether (3×10 mL). All organic fractions were combined and washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 30% to 80% of diethyl ether in hexane to yield a light red oil (0.105 g, 20%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.10$ (dd, 1H, $J = 9.0, 2.5$), 8.04 (d, 1H, $J = 2.5$), $6.90 - 6.99$ (m, 4H), $6.80 - 6.89$ (m, 4H), 6.76 (d, 2H, $J = 8.5$), $6.66 - 6.73$ (m, 6H), $6.56 - 6.64$ (m, 5H), 6.53 (dd, 1H, $J = 9.0, 3.0$), $6.33 - 6.46$ (m, 7H), 6.17 (s, 1H), $4.45 - 4.56$ (m, 8H), 4.41 (s, 2H), $3.93 - 4.10$ (m, 10H), $3.75 - 3.84$ (m, 10H), $1.63 - 1.82$ (m, 5H), $1.23 - 1.56$ (m, 40H), $0.84 - 0.97$ (m, 30H);

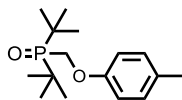
$^{13}\text{C NMR}$ (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.7, 154.5, 154.3, 154.2, 154.1, 149.2, 149.0, 148.1, 148.0, 143.6, 143.3, 143.1, 142.5, 141.4, 131.4, 131.2, 130.7, 128.8, 128.1, 128.0,$

127.7, 127.6, 127.2, 126.9, 125.2, 124.3, 123.2, 115.6, 115.5, 115.3, 114.4, 113.9, 113.5, 112.9, 112.6, 112.3, 112.1, 111.8, 111.1, 110.8, 110.7, 110.3, 100.3, 71.8, 71.4, 71.3, 71.0, 70.7, 65.0, 55.4, 54.8, 54.3, 51.3, 50.8, 50.6, 49.6, 39.6, 39.5, 39.2, 30.7, 30.6, 30.5, 30.4, 29.1, 29.0, 24.0, 23.9, 23.8, 23.0, 22.9, 14.0, 11.2, 11.1, 11.0;

MS (ES+): m/z (%) = 484 (100), 1575 (10), 1680 (10), 1681 (50) [M + H⁺], 1682 (50), 1683 (25), 1684 (10);

HRMS (ES+): calculated for C₁₀₅H₁₄₂N₅O₁₃ 1681.0604, found 1681.0530;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2927, 2858, 2163, 1613, 1593, 1505, 1464, 1339, 1264, 1225, 1168.

3.5.1.27 Synthesis of **3.25**

p-Cresol (0.21 g, 1.9 mmol, 1 equiv.) was dissolved in DMF (10 mL) and **3.5** (0.73 g, 2.1 mmol, 1.1 equiv.) and Cs₂CO₃ (0.93 g, 2.86 mmol, 1.5 equiv.) were added and stirred at 80 °C for 12 h. The reaction mixture was poured onto water (100 ml) and extracted into EtOAc (5 × 20 mL), before being washed with water (5 × 20 mL), brine (1 × 20 mL) and dried (MgSO₄). The solvent was removed using a rotary evaporator and the crude product was purified via flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in EtOAc to yield a white amorphous solid (0.21 g, 39%).

¹H NMR (250 MHz, CDCl₃): δ_H = 7.12 (d, 2H, *J* = 8.5), 6.83 (d, 2H, *J* = 8.5), 4.37 (d, 2H, *J* = 8.0), 2.30 (s, 3H), 1.39 (d, 18H, *J* = 13.5);

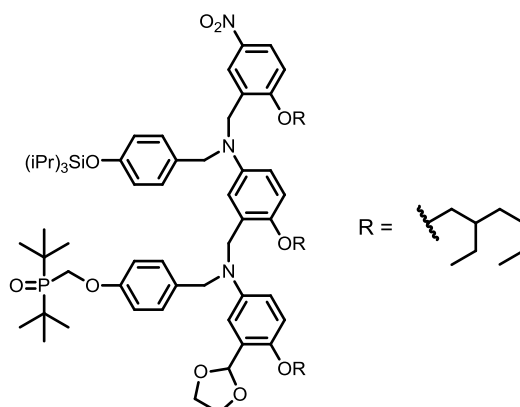
¹³C NMR (63 MHz, CDCl₃): δ_C = 156.3 (d, *J* = 10.5), 131.1, 130.1, 113.1, 62.6 (d, *J* = 70.5), 35.4 (d, *J* = 56.5), 26.4, 20.4;

³¹P NMR (101 MHz, CDCl₃): δ_P = 56.8;

MS (ES+): *m/z* (%) = 283 (100) [M + H⁺], 346 (50), 383 (15);

HRMS (ES+): calculated for C₁₆H₂₈O₂P 283.1827, found 283.1840;

FT-IR (thin film): ν_{max}/cm⁻¹ 3431, 2955, 2874, 1597, 1478, 1368, 1190, 1176, 1145, 996.

3.5.1.28 Synthesis of **3.26**

17' (0.158 g, 0.20 mmol, 1 equiv.) and **3.12** (0.175 g, 0.31 mmol, 1.5 equiv.) were dissolved in DCE (700 μ L) and $\text{NaBH}(\text{OAc})_3$ (0.121 g, 0.57 mmol, 2.8 equiv.) was added with stirring. After 2 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting 0% to 10% MeOH in diethyl ether to yield a pale yellow oil (0.217 g, 80%).

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.00$ (dd, 1H, $J = 9.0, 3.0$), 7.94 (d, 1H, $J = 3.0$), 7.11 (d, 2H, $J = 8.5$), 7.07 (d, 2H, $J = 8.5$), 6.79 – 6.85 (m, 4H), 6.76 (d, 1H, $J = 9.0$), 6.74 (d, 1H, $J = 3.0$) 6.72 (dd, 1H, $J = 9.0, 3.0$), 6.51 – 6.58 (m, 2H), 6.41 (d, 1H, $J = 3.0$), 6.34 (dd, 1H, $J = 9.0, 3.0$), 6.08 (s, 1H), 4.49 (s, 2H), 4.41 – 4.45 (m, 4H), 4.36 (d, 2H, $J = 7.5$), 4.21 (s, 2H), 3.90 – 3.93 (m, 6H), 3.79 (d, 2H, $J = 5.5$), 3.76 (d, 2H, $J = 5.5$), 1.60 – 1.79 (m, 3H), 1.18 – 1.58 (m, 45H), 1.09 (d, 18H, $J = 7.5$), 0.81 – 0.98 (m, 17H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.6, 157.4$ (d, $J = 11.0$), 155.1, 149.3, 149.3, 143.3, 142.8, 141.5, 132.9, 131.4, 128.6, 128.4, 128.3, 127.3, 126.6, 124.4, 123.2, 120.1, 114.2, 113.9, 113.5, 113.2, 112.5, 111.6, 111.2, 110.6, 99.9, 71.9, 71.5, 71.0, 65.2, 62.8

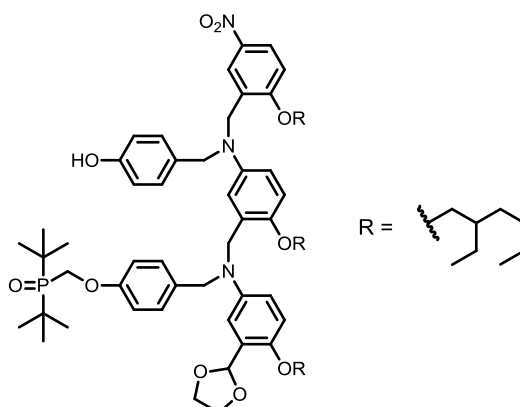
(d, $J = 70.5$), 55.8, 54.0, 50.4, 49.8, 39.9, 39.7, 39.5, 35.7 (d, $J = 57.5$), 30.9, 30.8, 30.8, 29.3, 29.3, 29.3, 26.7, 24.3, 24.2, 24.2, 23.3, 23.2, 18.2, 14.3, 14.3, 12.9, 11.4;

MS (MALDI+): m/z (%) = 1317.0 (50), 1332.2 (100) $[M + H^+]$;

HRMS (ES+): calculated for $C_{79}H_{123}N_3O_{10}P^{28}Si$ 1332.8710, found 1332.8678;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2953, 2929, 2867, 1609, 1507, 1464, 1339, 1263, 1226.

3.5.1.29 Synthesis of 3.27



3.26 (0.101 g, 0.08 mmol, 1 equiv.) was dissolved in THF (2 mL) at 0 °C and TBAF (80 μ L, 0.08 mmol, 1 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (3×10 mL). All organic fractions were combined and washed with brine (3×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 50% to 100% of ethyl acetate in hexane to yield a light yellow oil (0.030 g, 34%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} = 8.03 (1H, dd, $^3J = 9.0$, $^4J = 3.0$), 7.97 (1H, d, $^4J = 3.0$), 7.09 (2H, d, $^3J = 8.5$), 6.93 (2H, d, $^3J = 7.0$), 6.82 (1H, d, $^4J = 3.0$), 6.80, 1H, d, $^3J = 9.0$), 6.72 (2H, d, $^3J = 7.0$), 6.70 (2H, d, $^3J = 7.0$), 6.63 (1H, d, $^3J = 9.0$), 6.61 – 6.56 (2H, m), 6.44 – 6.35 (2H, m), 6.07 (1H, s), 4.44 (4H, s), 4.37 (2H, s), 4.33 (2H, d, $^2J_{\text{p}} = 6.5$), 4.28 (2H, s), 4.01 – 3.89 (4H, m), 3.91 (2H, d, $^3J = 6.0$), 3.78 (2H, d, $^3J = 5.5$), 3.72 (2H, d, $^3J = 5.5$), 1.79 – 1.57 (3H, m), 1.56 – 1.15 (42H, m), 0.97 – 0.77 (18H, m);

$^{13}\text{C NMR}$ (126 MHz, CDCl_3): δ_{C} = $^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ_{C} = 161.8, 157.3 (d, $J = 10.5$), 155.9, 149.6, 148.9, 143.7, 142.7, 141.5, 133.2, 129.7, 128.8, 128.5, 127.9, 127.8, 126.1, 124.4, 123.2, 116.1, 114.9, 114.2, 113.9, 113.5, 112.2, 111.5, 110.5,

100.1, 71.8, 71.5, 70.8, 65.2, 62.6 (d, $J = 70.0$), 55.7, 55.5, 50.9, 50.3, 39.7, 39.3, 35.5

(d, $J = 57.5$), 30.8, 30.7, 29.2, 26.6, 24.2, 24.1, 23.2, 23.1, 14.2, 11.3;

HRMS (ES+): calculated for $C_{70}H_{103}N_3O_{10}P$ 1176.7381, found 1176.7355;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2956, 2921, 2853, 1679, 1609, 1592, 1502, 1467, 1336,

1263, 1225, 1175, 1124.

3.5.2 Binding studies

All association constants were measured by ^1H NMR and additionally, if applicable, ^{31}P NMR titration. One species, labelled the host, is dissolved in toluene- d_8 to a known concentration. A second species, labelled the guest, is dissolved in the host solution and made to a known concentration such that the concentration of the host is the same in both solutions. A known volume of host is added to an NMR tube and the spectra measured. Known volumes of guest in host solution are added to the tube, and the spectra measured after each addition. The chemical shifts of the host spectra are recorded as a function of guest concentration and analysed using standard calculations in a fitting-program written in Microsoft Excel®. Errors were calculated as two times the standard deviation from the average value from repetitions.

3.5.3 Temperature denaturation

The temperature of the sample was changed using the internal thermostat of the NMR spectrometer and the sample was allowed to equilibrate in the probe until the probe thermometer gave a stable reading.

3.6 References

- 1 N. C. Seeman, *Annu. Rev. Biochem.*, 2010, **79**, 65–87.
- 2 H. Yan, T. H. LaBean, L. Feng, and J. H. Reif, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 8103–8.
- 3 P. W. K. Rothmund, *Nature*, 2006, **440**, 297–302.
- 4 W. M. Shih, J. D. Quispe, and G. F. Joyce, *Nature*, 2004, **427**, 618–21.
- 5 S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, *Nature*, 2009, **459**, 414–8.
- 6 Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, *Science*, 2012, **338**, 1177–83.
- 7 P. Yin, R. F. Hariadi, S. Sahu, H. M. T. Choi, S. H. Park, T. H. Labean, and J. H. Reif, *Science*, 2008, **321**, 824–826.
- 8 J. Bath and A. J. Turberfield, *Nat. Nanotechnol.*, 2007, **2**, 275–84.
- 9 Y. Krishnan and F. C. Simmel, *Angew. Chem. Int. Ed.*, 2011, **50**, 3124–56.
- 10 E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, and J. Kjems, *Nature*, 2009, **459**, 73–6.
- 11 Z.-G. Wang, J. Elbaz, and I. Willner, *Nano Lett.*, 2011, **11**, 304–9.
- 12 B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, and J. L. Neumann, *Nature*, 2000, **406**, 605–8.
- 13 R. Chhabra, J. Sharma, Y. Liu, and H. Yan, *Nano Lett.*, 2006, **6**, 978–83.
- 14 L. Adleman, *Science*, 1994, **266**, 1021–1024.
- 15 Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, and D. R. Liu, *Science*, 2004, **305**, 1601–5.
- 16 Y. Chen and C. Mao, *J. Am. Chem. Soc.*, 2004, **126**, 13240–1.
- 17 Y. He and D. R. Liu, *Nat. Nanotechnol.*, 2010, **5**, 778–82.
- 18 S. Liao and N. C. Seeman, *Science*, 2004, **306**, 2072–4.
- 19 T. M. Snyder and D. R. Liu, *Angew. Chem. Int. Ed.*, 2005, **44**, 7379–82.
- 20 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 964–967.
- 21 T. P. Prakash, *Chem. Biodivers.*, 2011, **8**, 1616–41.

- 22 R. N. Veedu and J. Wengel, *Chem. Biodivers.*, 2010, **7**, 536–42.
- 23 C. Leumann, *Bioorg. Med. Chem.*, 2002, **10**, 841–854.
- 24 A. Eschenmoser, *Science*, 1999, **284**, 2118–2124.
- 25 P. Wittung, P. E. Nielsen, O. Buchardt, M. Egholm, and B. Nordén, *Nature*, 1994, **368**, 561–3.
- 26 P. Nielsen, M. Egholm, R. Berg, and O. Buchardt, *Science.*, 1991, **254**, 1497–1500.
- 27 K. Murayama, Y. Tanaka, T. Toda, H. Kashida, and H. Asanuma, *Chem. - A Eur. J.*, 2013, **19**, 14151–14158.
- 28 E. Meggers and L. Zhang, *Acc. Chem. Res.*, 2010, **43**, 1092–102.
- 29 H. Asanuma, T. Toda, K. Murayama, X. Liang, and H. Kashida, *J. Am. Chem. Soc.*, 2010, **132**, 14702–3.
- 30 H. Kashida, K. Murayama, T. Toda, and H. Asanuma, *Angew. Chem. Int. Ed.*, 2011, **50**, 1285–8.
- 31 S. A. Benner, *Acc. Chem. Res.*, 2004, **37**, 784–97.
- 32 H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, and E. T. Kool, *Science*, 2003, **302**, 868–71.
- 33 E. M. Harcourt and E. T. Kool, in *Synthetic Biology*, ed M. Ryadnov, L. Brunsveld, H. Suga, Royal Society of Chemistry, Cambridge, 2014, vol. 1, pp 1–30.
- 34 H. Ito, Y. Furusho, T. Hasegawa, and E. Yashima, *J. Am. Chem. Soc.*, 2008, **130**, 14008–15.
- 35 C. A. Hunter and H. L. Anderson, *Angew. Chem. Int. Ed.*, 2009, **48**, 7488–99.
- 36 J. D. Badjić, A. Nelson, S. J. Cantrill, W. B. Turnbull, and J. F. Stoddart, *Acc. Chem. Res.*, 2005, **38**, 723–732.
- 37 D. Pörschke and M. Eigen, *J. Mol. Biol.*, 1971, **62**, 361–381.
- 38 M. E. Craig, D. M. Crothers, and P. Doty, *J. Mol. Biol.*, 1971, **62**, 383–401.
- 39 H. J. Hogben, J. K. Sprafke, M. Hoffmann, M. Pawlicki, and H. L. Anderson, *J. Am. Chem. Soc.*, 2011, **133**, 20962–20969.
- 40 D. Porschke, *Biopolymers*, 1971, 1989–2013.
- 41 S. I. Chan, M. P. Schweizer, P. O. P. Ts'o, and G. K. Helmkamp, *J. Am. Chem. Soc.*, 1964, **86**, 4182–4188.

- 42 C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310–24.
- 43 E. Chekmeneva, C. A. Hunter, M. C. Misuraca, and S. M. Turega, *Org. Biomol. Chem.*, 2012, **10**, 6022–31.
- 44 M. H. Abraham, H. S. Chadha, G. S. Whiting, and R. C. Mitchell, *J. Pharm. Sci.*, 1994, **83**, 1085–1100.
- 45 C. S. Calero, J. Farwer, E. J. Gardiner, C. a Hunter, M. Mackey, S. Scuderi, S. Thompson, and J. G. Vinter, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18262–73.
- 46 A. F. Abdel-Magid and S. J. Mehrman, *Org. Process Res. Dev.*, 2006, **10**, 971–1031.
- 47 A. F. Abdel-Magid, K. G. Carson, B. D. Harris, C. a. Maryanoff, and R. D. Shah, *J. Org. Chem.*, 1996, **61**, 3849–3862.
- 48 M. C. Misuraca, T. Grecu, Z. Freixa, V. Garavini, C. A. Hunter, P. W. N. M. van Leeuwen, M. D. Segarra-Maset, and S. M. Turega, *J. Org. Chem.*, 2011, **76**, 2723–32.
- 49 M. A. Ischay, Z. Lu, and T. P. Yoon, *J. Am. Chem. Soc.*, 2010, **132**, 8572–4.
- 50 K. J. Breslauer, *Methods Enzymol.*, 1995, 259, 221–242.
- 51 I. Rouzina and V. A. Bloomfield, *Biophys. J.*, 1999, **77**, 3242–3251.
- 52 T. Ratilainen, A. Holmén, E. Tuite, P. E. Nielsen, and B. Nordén, *Biochemistry*, 2000, **39**, 7781–7791.
- 53 D. Pörschke, *Biopolymers*, 1971, 1989–2013.
- 54 P. N. Taylor and H. L. Anderson, *J. Am. Chem. Soc.*, 1999, **121**, 11538–11545.
- 55 G. Iadevaia, Ph.D. Thesis, University of Sheffield, 2014

Chapter 4

**Recognition group
compatibility of
single sequence H-
bonding
information
oligomers**

4.1 Introduction

Highly versatile and tuneable supramolecular building blocks are advantageous in that they allow one material to be useful for a whole variety of applications. The obvious example of such a material is the programmable oligomers found in Nature, for instance the nucleic acids, polypeptides and oligosaccharides. The principle source of tuneability comes from the sequence of units. Single stranded forms of RNA and polypeptides can fold to form pockets, which selectively bind other molecules.¹⁻⁴ The role of oligosaccharides in Nature is varied and it is known different sequences of sugars play a part in recognition.^{5,6} However, chemical modifications of the types of components used allow information oligomers to be tuned in other ways. For instance RNA and DNA are structurally very similar but the lack of the 2'-hydroxyl group is enough to distinguish them such that DNA and RNA perform different functions in the body,⁷ and DNA is more stable than RNA.⁸ The stability of the DNA helix can be tuned by the proportion of weaker Adenine-Thymine stacking interactions and stronger Cytosine-Guanine interactions.⁹ Chemists have further altered the properties of each part of DNA through structural reengineering. For instance duplex stability can be enhanced by using expanded base pairs to enhance π -stacking,¹⁰ and using Locked Nucleic Acids which fix the backbone in a favourable position for hybridisation.¹¹

In the previous Chapter, a completely synthetic information oligomer was presented with cooperative H-bonding interactions forming the driving force for duplex formation. In this Chapter we demonstrate the generality of this design, by changing the phosphine oxide acceptor groups for both pyridine and pyridine *N*-oxide and show that duplex formation is still cooperative.

4.2 Approach

4.2.1 Design

Figure 4.1 shows the design of the information oligomer where there are four key components: H-bonding *recognition* groups; a group to control *solubility*; a type of *coupling chemistry* which allows sequences of polymers to be made and a *backbone* which joins all the other components together. In Chapter 3 we showed the applicability of this design based on an aniline backbone with branched alkoxy groups to promote solubility in non-polar solvents. The oligomers are constructed through reductive amination chemistry between benzaldehyde and aniline derivatives. Duplex formation occurs through cooperative H-bonding between the recognition groups. A key feature of this approach is the functions are separated into separate modules, each of which is exchangeable. For instance, the recognition unit is derived from a benzaldehyde, which can bear many different H-bonding groups with variation in properties such as conformational flexibility and H-bond strength. The alkoxy solubilising chains are formed through alkylating a phenol on the backbone. The degree of solubility and crystallinity can be manipulated by changing the alkoxy group for more or less branched alkyl chains.

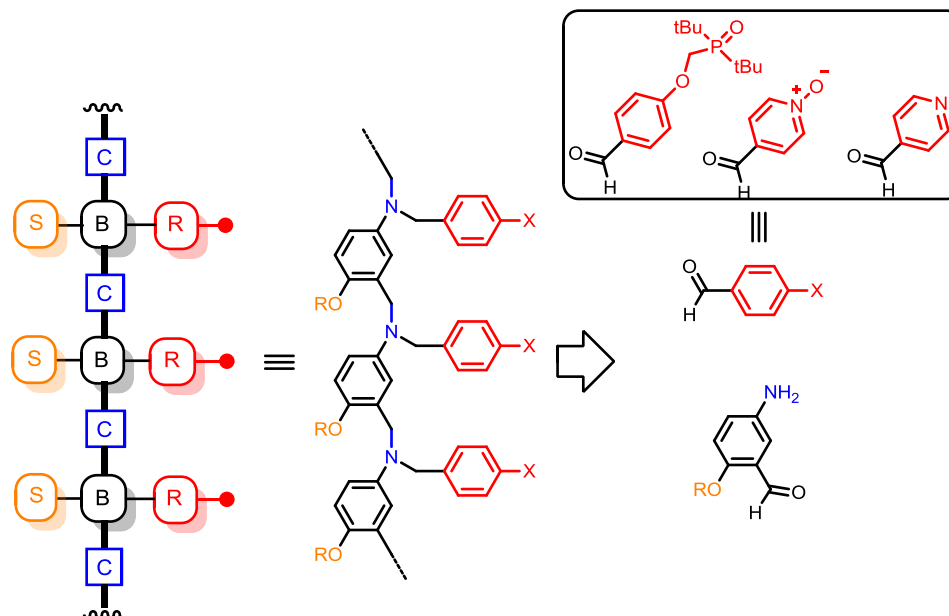


Figure 4.1. Information oligomers made from monomers with four key functions: linking chemistry to join monomer units together (C, blue); recognition sites responsible for sequence selectivity in duplex formation (R, red); a control over solubility (S, orange) and a backbone for attachment of all the other functions (B, black). A retrosynthesis is shown, which gives simple starting materials, of which there are a number of possible choices, each with different properties.

4.2.2 Duplex formation and $K_{\text{ref}} EM_N$

In Chapter 3 we showed that the dimensionless product $K_{\text{ref}} EM$ is an excellent measure of the degree of cooperativity of duplex formation between two strands. The effective molarity EM is the parameter that determines how readily the intramolecular H-bond forms relative to an intermolecular interaction. The value of K_{ref} is the association constant for the reference single point intermolecular H-bond between the H-bond acceptor and donor recognition group. Figure 4.2(a) shows the stepwise formation of a duplex from an H-bond donor oligomer and an H-bond acceptor oligomer. The nucleation step is the formation of a single H-bond, which has an equilibrium constant of $\sigma_1 K_{\text{ref}}$, where σ_1 is the statistical factor for forming the first H-bond. Each subsequent step in the duplex formation has an association constant corresponding to the product $\sigma_N K_{\text{ref}} EM_N$, where EM_N is the stepwise effective molarity, and σ_N is the statistical factor associated with forming the N th H-bond. It follows that the association constant for forming a duplex comprised of an oligomer with N H-bonds K_N is the product of all the stepwise equilibrium constants (Equation 4.1)

$$K_N = \sigma_{D,N} K_{\text{ref}}^N \overline{EM}_N^{(N-1)} \quad (4.1)$$

where \overline{EM}_N is the average effective molarity observed for all steps and σ_D is the statistical factor for forming the fully bound duplex. By rearranging Equation 4.1 the \overline{EM}_N for a duplex can be determined using Equation 4.2.

$$\overline{EM}_N = \sqrt[N-1]{\frac{K_N}{\sigma_{D,N} K_{\text{ref}}^N}} \quad (4.2)$$

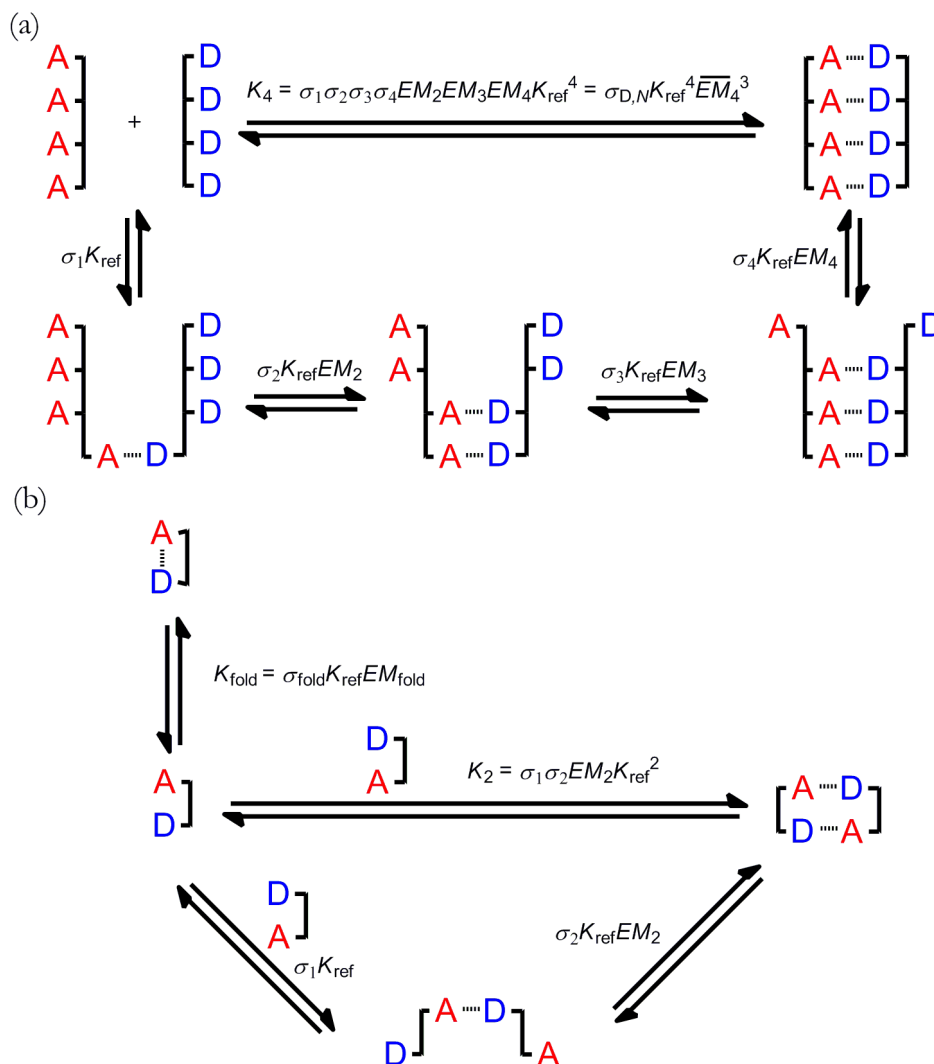


Figure 4.2. Equilibria involved in duplex formation (a) two oligomers with four H-bonding sites each bind together with a macroscopic association constant of K_4 which is the product of all the microscopic association constants. The stepwise equilibria have association constants described using K_{ref} , EM_N and σ_N ; (b) oligomers with a neighbouring H-bond acceptor (A) and donor (D) can form an intramolecular H-bond, which may compete with duplex formation. The equilibrium constant of folding K_{fold} can be expressed by the product of the statistical factor σ_{fold} , K_{ref} and the effective molarity of folding EM_{fold} .

The fully bound duplex will predominate and the partially bound species shown in Figure 4.2(a) will not be significantly populated if $K_{\text{ref}} EM \gg 1$. In Chapter 3 we showed that by choosing a sufficiently high K_{ref} (350 M^{-1}), the condition that $K_{\text{ref}} EM > 1$ can be achieved with low values of \overline{EM}_N ($8 - 20 \text{ mM}$). By changing the recognition motif in this design, both the K_{ref} and EM values change, since K_{ref} is

dependent on the H-bonding properties of a functional group and EM is dependent on, the degree of flexibility or preorganization of the system.¹² K_{ref} can be quantitatively predicted using Equation 4.3¹³

$$-RT\ln K_{\text{ref}} = \Delta G^{\circ} = -(\alpha - \alpha_s)(\beta - \beta_s) + 6 \text{ kJ mol}^{-1} \quad (4.3)$$

where α and β are the H-bond donor and acceptor properties of the two interacting functional groups, α_s and β_s describe the H-bond donating and accepting properties of the solvent respectively, R is the gas constant, T is the temperature and the 6 kJ mol^{-1} is the free energy cost of a bimolecular association in solution.

The phosphine oxide used in Chapter 3 is one the best H-bond acceptors ($\beta = 10.2$). In this Chapter we will investigate pyridine ($\beta = 7.7$) and pyridine *N*-oxide ($\beta = 9.0$) as H-bond acceptor groups. In toluene- d_8 ($\alpha_s = 1.0$, $\beta_s = 2.2$)¹⁴ these groups have predicted association constants of 36 M^{-1} and 150 M^{-1} respectively, which together with the phosphine oxide system from Chapter 3, span a wide range of H-bond strengths. The geometries and number of rotors is different in each system, so this is likely to affect the value of EM (and $K_{\text{ref}}EM$). Therefore, it will be interesting to investigate whether this approach to H-bonding oligomers is tolerant of variation in the magnitudes of single point H-bond association constants, flexibilities and H-bond geometries.

Figure 4.2(b) shows the simplest example of duplex formation between strands that contain a mixture of H-bond donors and acceptors. DA is able to form an antiparallel duplex with itself, but has the potential to form an intramolecular H-bond, between the donor and acceptor groups within the same molecule. The equilibrium constant for the unimolecular folding process K_{fold} is related to the effective molarity for folding EM_{fold} and statistical factor σ_{fold} (Equation 4.4).

$$K_{\text{fold}} = \sigma_{\text{fold}} K_{\text{ref}} EM_{\text{fold}} \quad (4.4)$$

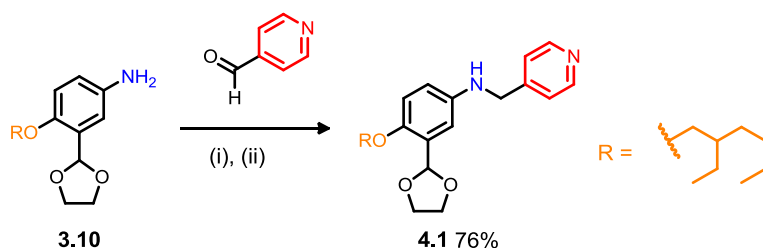
If folding between adjacent H-bonding sites in an oligomer is facile, duplex formation would be effectively disrupted and also prevent that oligomer acting as a template for template directed synthesis. By comparing the association constants for duplexes of AA•DD and AD•AD, EM_{fold} can be determined which gives a quantifiable measure of the potential for an information oligomer to express sequence information through duplex formation and template directed synthesis. Sequence selective duplexes will be further explored in Chapter 5.

4.3 Results and discussion

4.3.1 Synthesis

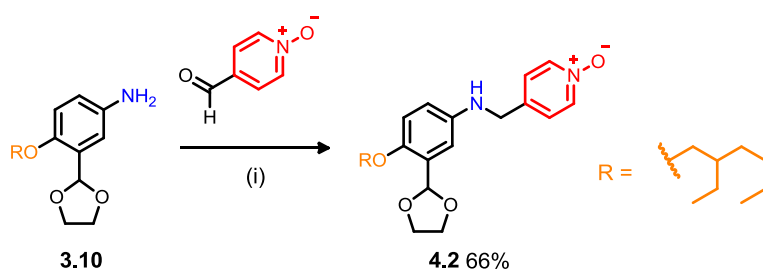
4.3.1.1 Monomers

The synthesis of primary aniline **3.10** is given in Chapter 3. All monomer units for oligomer synthesis were made by coupling a benzaldehyde derivative with **3.10** (Schemes 4.1 – 4.3). Pyridine **4.1** was made by reducing the imine formed between **3.10** and 4-nicotinaldehyde using NaBH_4 (Scheme 4.1). Pyridine *N*-oxide **4.2** was made via reductive amination of **3.10** and 4-formylpyridine-*N*-oxide using $\text{NaBH}(\text{OAc})_3$. The synthesis of **3.11** is given in Chapter 3 and phenol **4.3** was made by deprotection of the silyl-ether using tetra-*n*-butylammonium fluoride (TBAF).



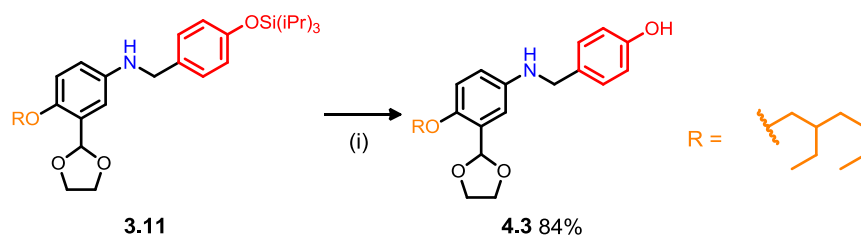
Scheme 4.1.

Reagents and conditions: (i) heat; (ii) NaBH_4 .



Scheme 4.2.

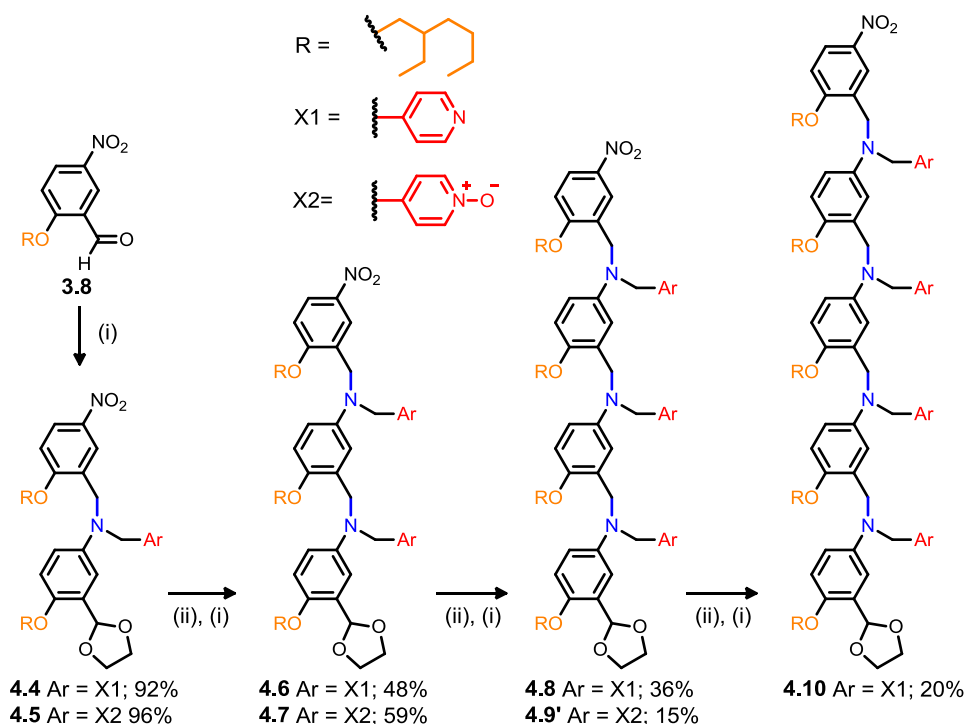
Reagents and conditions: (i) $\text{NaBH}(\text{OAc})_3$.

**Scheme 4.3.**

Reagents and conditions: (i) TBAF.

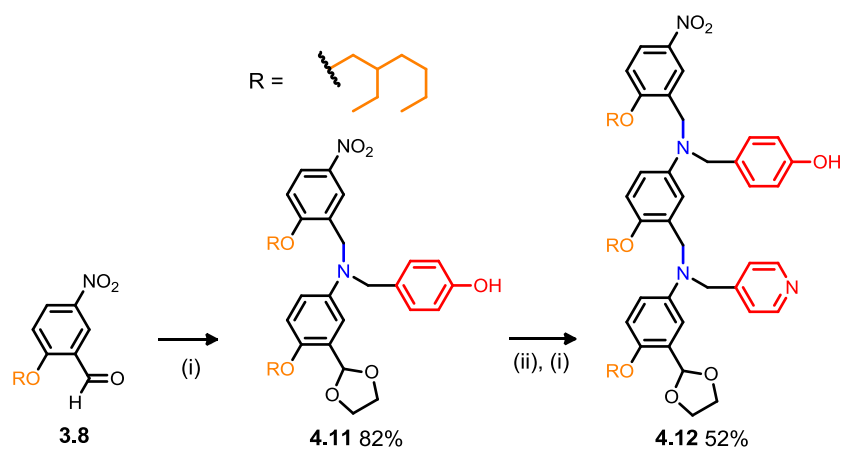
4.3.1.2 Oligomers

The synthesis of benzaldehyde **3.8** is given in Chapter 3 and this compound formed the capping group at the end of each oligomer. Through iterative reductive amination and acetal deprotection steps, all homo- and hetero-sequence oligomers shown in Schemes 4.4 – 4.6 were synthesised, requiring purification after each reductive amination step. The synthesis of H-bond donor phenol oligomers **3.21**, **3.22** and **3.23**, shown in Figure 4.3, is described in Chapter 3. The pyridine *N*-oxide 3-mer was isolated as the benzaldehyde of the structure shown in Scheme 4.4 and is designated with a prime (**4.9'**).



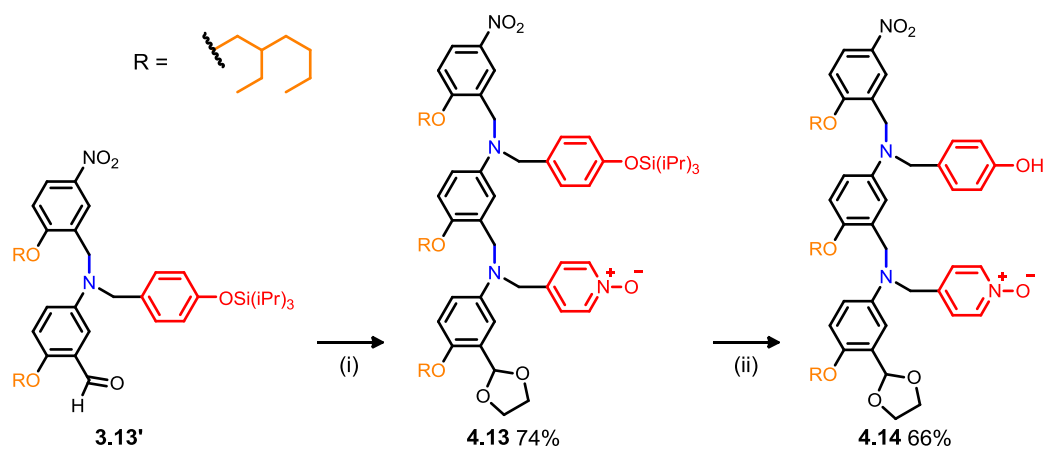
Scheme 4.4. (Compound **4.9'** corresponds to the benzaldehyde derivative of the structure shown).

Reagents and conditions: (i) **4.1** or **4.2**, NaBH(OAc)₃; (ii) HCl.



Scheme 4.5.

Reagents and conditions: (i) 4.3 or 4.1, NaBH(OAc)₃; (ii) HCl.



Scheme 4.6.

Reagents and conditions: (i) 4.2, NaBH(OAc)₃; (ii) TBAF, THF.

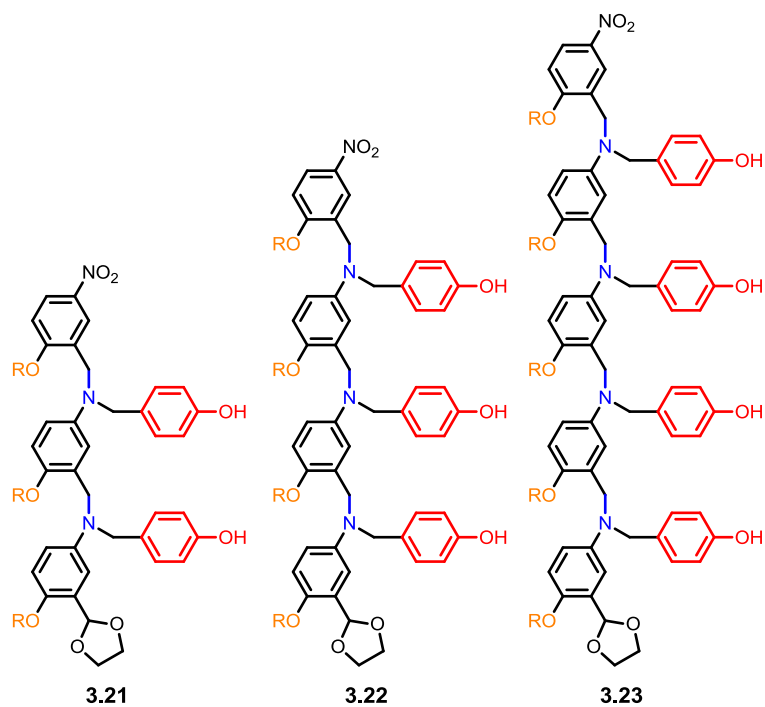


Figure 4.3. Structures of phenol H-bond donor oligomers synthesised in Chapter 3.

4.3.2 NMR Binding Studies

The association constants of length-complementary duplexes were measured by fitting ^1H NMR titration data in toluene- d_8 to a 1:1 isotherm. Values of K_{ref} for the phenol-pyridine and the phenol-pyridine N -oxide interactions were measured by titrating **4.15** or **4.16** into phenol **3.25** (Figure 4.4). Figures 4.6 – 4.12 show representative ^1H NMR spectra and plots of the changes in chemical shift $\Delta\delta$ as a function of guest concentration compared to the best fit 1:1 isotherm. The derived association constants are shown in Tables 4.1 and 4.2, along with values of \overline{EM}_N and $K_{\text{ref}} \overline{EM}_N$, calculated using Equation 4.2. In titrations of **3.25**•**4.15** and **3.25**•**4.16** (Figures 4.6 and 4.10, respectively), the phenol was used as the host as the large downfield chemical shift change of the phenol proton could be monitored (> 6.5 ppm in both cases), which is typical of H-bond formation).¹⁵

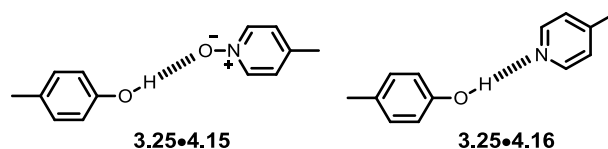


Figure 4.4. Complexes studied in order to obtain K_{ref} for the phenol pyridine and phenol pyridine N -oxide interactions.

For complexes where $N > 1$, the pyridine and pyridine N -oxide oligomers were used as the host. Table 4.5 shows the limiting complexation-induced changes in ^1H NMR chemical shift for the pyridine oligomer series (the proton labelling scheme is shown in Figure 4.5). The H1 signals due to protons on the pyridine unit show a similar upfield shift ($\Delta\delta \approx -0.2$) for complexes **4.6**•**3.21**, **4.8**•**3.22** and **4.10**•**3.23**. There are two types of CH_2 protons on the backbone: those benzylic to the pyridine ring (H2) and those benzylic to the rings bearing the alkoxy groups (H3). The H2 protons all show upfield shifts ($\Delta\delta = 0.0 - -0.3$ ppm), whereas the H3 protons all

show a small downfield shift (<0.05 ppm). In contrast, the H1 signals on the pyridine *N*-oxide system do not show a consistent trend. For instance for the titration of **4.9**•**3.22** two of the H1 protons give a downfield shift, and the third has an upfield shift. However, the H2 and H3 protons in the pyridine *N*-oxide series show the same trend as the pyridine series: the H2 protons move slightly downfield and H3 protons move slightly upfield on complexation. For the pyridine *N*-oxide, the H4 signal could also be followed and gave a consistent upfield shift ($\Delta\delta = -0.2 - -0.3$).

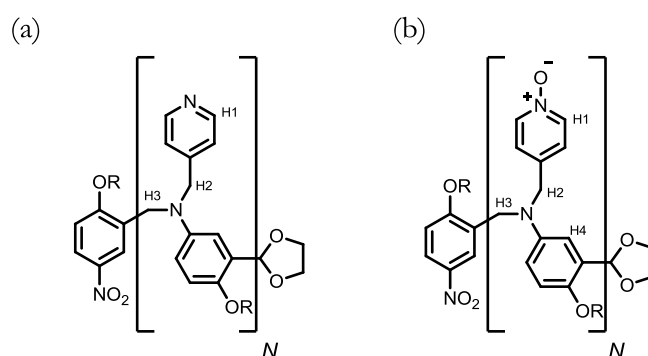


Figure 4.5. Proton labelling scheme for (a) pyridine oligomers and (b) pyridine *N*-oxide oligomers. R = 2-ethylhexyl

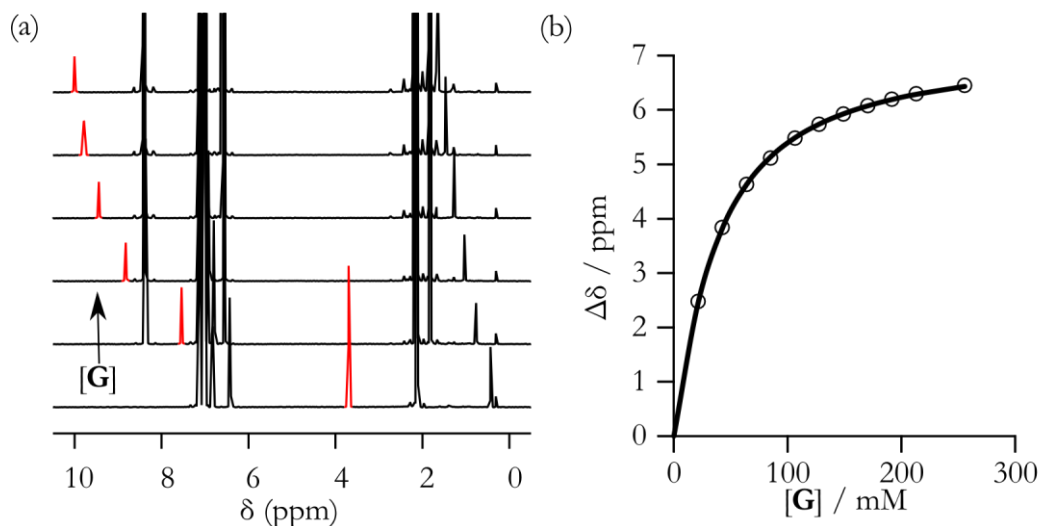


Figure 4.6. Titration of **4.16** into **3.25** (19 mM) at 298 K in toluene- d_8 . (a) 400 MHz ^1H NMR spectra (the OH signal is highlighted in red). (b) Chemical shift change as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm and **G** = **4.16**).

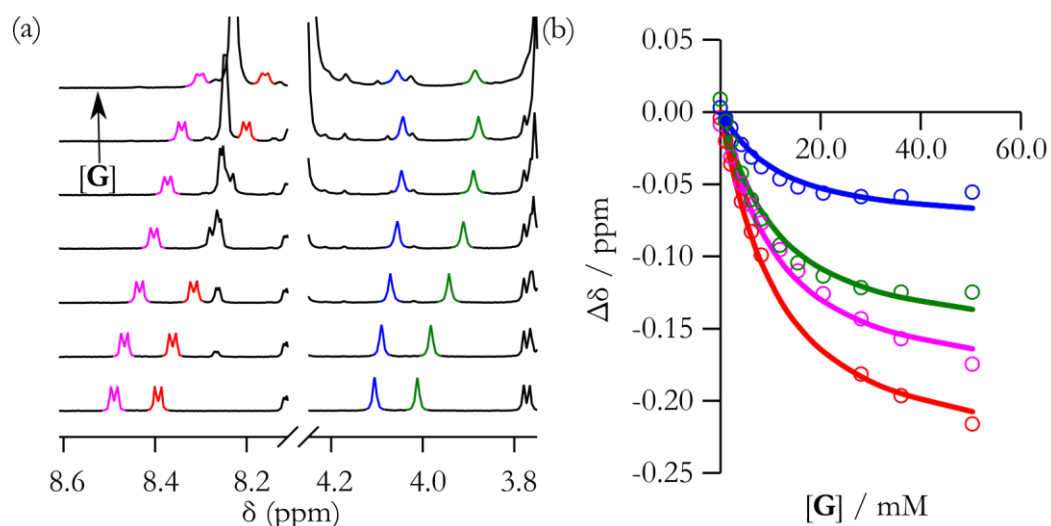


Figure 4.7. Titration of **3.21** into **4.6** (5.0 mM) at 298 K in toluene- d_8 . (a) 400 MHz ^1H NMR spectra. (b) Chemical shift change as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, $\mathbf{G} = \mathbf{3.21}$). The magenta and red signals correspond to H1; the blue and green signals correspond to H2 (see Figure 4.5(a) for labelling scheme).

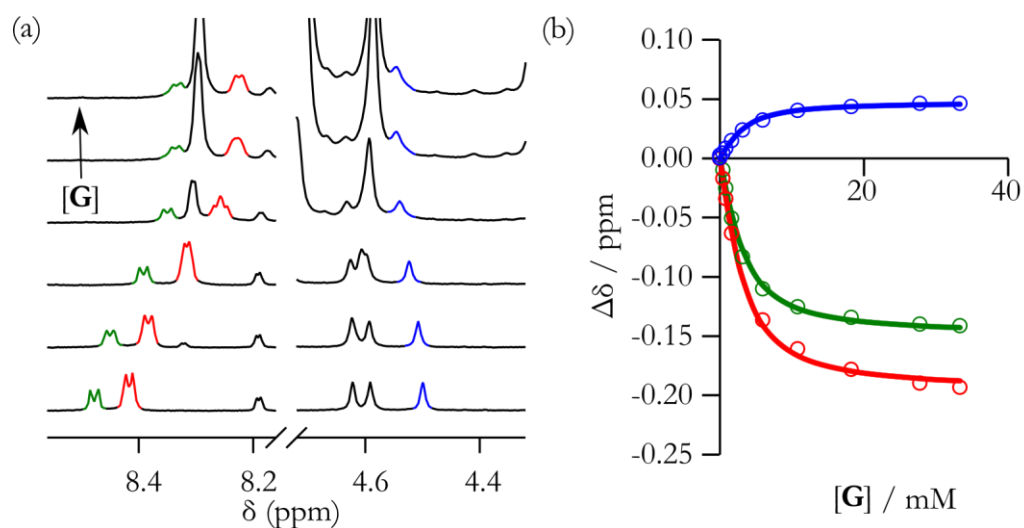


Figure 4.8. Titration of **3.22** into **4.8** (3.0 mM) at 298 K in toluene- d_8 . (a) 400 MHz ^1H NMR spectra. (b) Chemical shift change as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, $\mathbf{G} = \mathbf{3.22}$). The green and red signals correspond to H1 and blue corresponds to H3 (see Figure 4.5(a) for labelling scheme).

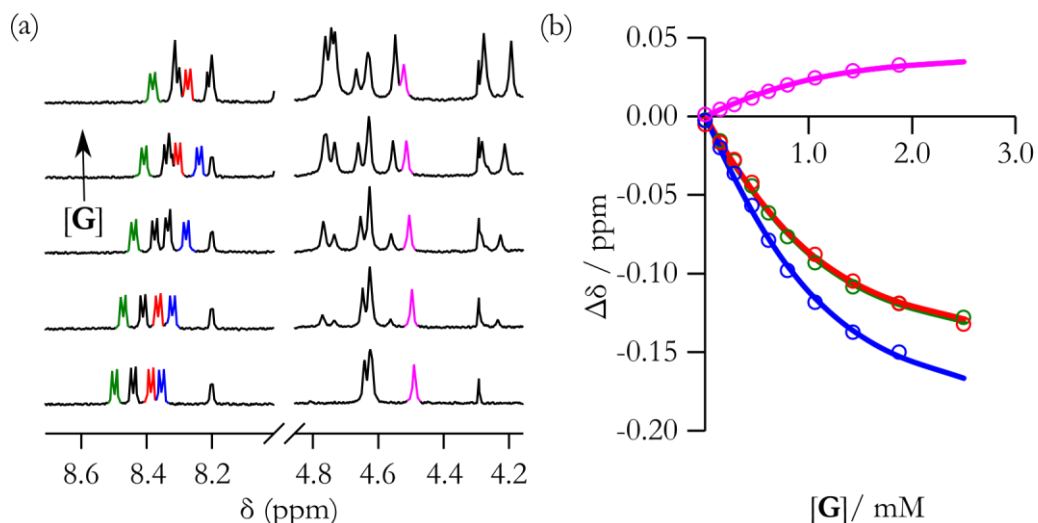


Figure 4.9. Titration of **3.23** into **4.10** (1.0 mM) at 298 K in toluene-*d*₈. (a) 400 MHz ¹H NMR spectra. (b) Chemical shift change as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, **G** = **3.23**). The green red and blue signals correspond to H1, magenta corresponds to H3 (see Figure 4.5(a) for labelling scheme).

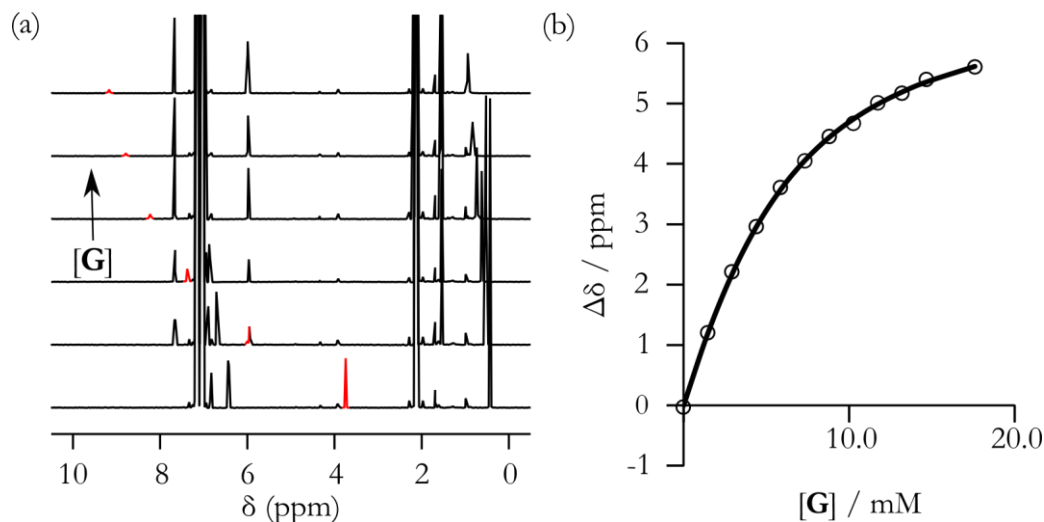


Figure 4.10. ¹H NMR data for titration of **4.15** into **3.25** (4.5 mM) at 298 K in toluene-*d*₈. (a) Example 400 MHz ¹H NMR spectra (OH signal highlighted in red). (b) Plot of the change in chemical shift of the OH signal as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, **G** = **4.15**).

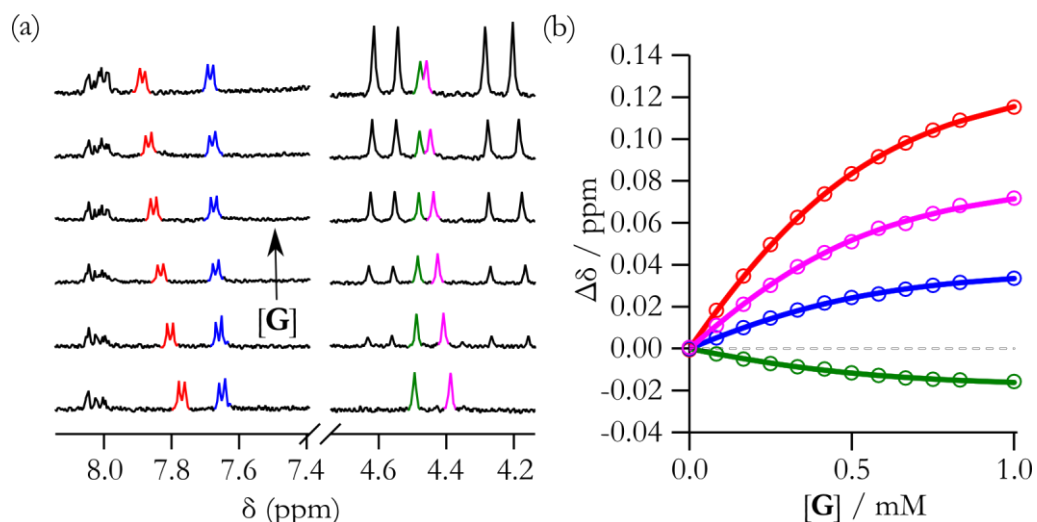


Figure 4.11. ^1H NMR data for titration of **3.21** into **4.7** (0.5 mM) at 298 K in toluene- d_8 . (a) Example 400 MHz ^1H NMR spectra. (b) Plot of the change in chemical shift as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, $G = \mathbf{3.21}$). The red and blue signals correspond to H1 and magenta and green signals correspond to H3 (see Figure 4.5(b) for labelling scheme).

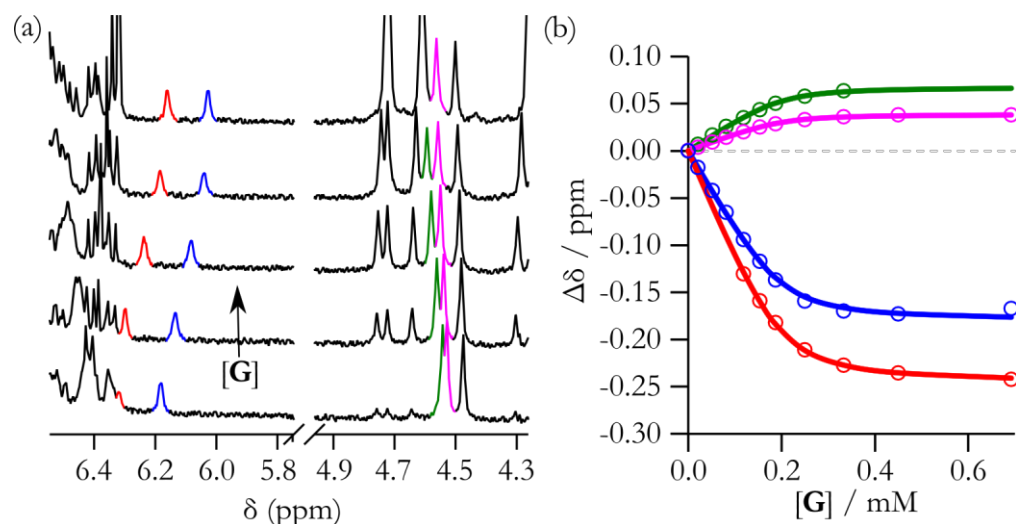


Figure 4.12. ^1H NMR data for titration of **3.22** into **4.9'** (0.2 mM) at 298 K in toluene- d_8 . (a) Example 400 MHz ^1H NMR spectra. (b) Plot of the change in chemical shift as a function of guest concentration (the line represents the best fit to a 1:1 binding isotherm, $G = \mathbf{3.22}$). The red and blue signals correspond to H4 and magenta and green signals correspond to H3 (see Figure 4.5(b) for labelling scheme).

Table 4.1. Association constants (K_N), average effective molarities (EM_N) and $K_{\text{ref}}EM_N$ for formation of 1:1 complexes between pyridine and phenol oligomers based on ^1H NMR titrations in toluene- d_8 at 298 K.^a

| Complex | K_N (M^{-1}) | EM_N / mM | $K_{\text{ref}}EM_N$ |
|---|---------------------------|-------------|----------------------|
| -A-D- 3.25•4.16 | 3.4×10^1 (4%) | - | - |
| [^A-^D] 4.6•3.21 | 1.3×10^2 (13%) | 57 (7%) | 1.9 (4%) |
| [^A-^D] 4.8•3.22 | 5.4×10^2 (13%) | 82 (5%) | 2.8 (3%) |
| [^A-^D] 4.10•3.23 | 4.0×10^3 (68%) | 110 (18%) | 3.9 (5%) |

^a Each titration was repeated twice and the average value is reported with relative standard errors (in brackets) at the 95% confidence limit.

Table 4.2. Association constants (K_N), average effective molarities (EM_N) and $K_{\text{ref}}EM_N$ for formation of 1:1 duplexes between pyridine *N*-oxide and phenol oligomers from ^1H NMR titrations in toluene- d_8 at 298 K.^a

| Complex | K (M^{-1}) | EM_N / mM | $K_{\text{ref}}EM_N$ |
|--|-------------------------|-------------|----------------------|
| -A-D- 3.25•4.15 | 3.3×10^2 (24%) | - | - |
| [^A-^D] 4.7•3.21 | 5.4×10^3 (30%) | 25 (16%) | 8.2 (4%) |
| [^A-^D] 4.9•3.22 | 1.0×10^5 (53%) | 37 (19%) | 12 (3%) |

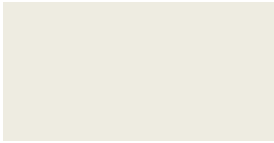

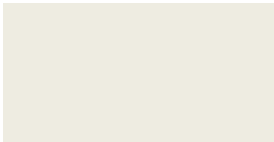
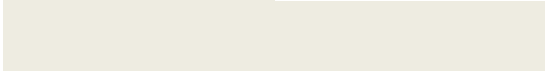
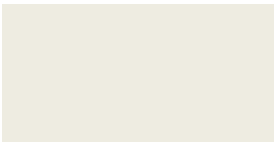
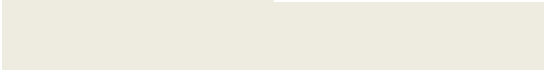
^a Each titration was repeated twice and the average value is reported with relative standard errors (in brackets) at the 95% confidence limit.

Table 4.3. ^1H NMR chemical shifts of the free pyridine host (ppm) obtained by fitting the pyridine-phenol titration data to a 1:1 binding isotherm.^a

| proton | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ |
|--------|--|--|--|
| | 4.6•3.21 | 4.8•3.22 | 4.10•3.23 |
| H1 | 8.5 | 8.5 | 8.5 |
| | 8.4 | 8.4 | 8.4 |
| | | 8.4 | 8.3 |
| | | | 8.3 |
| H2 | 4.1 | 4.1 | 4.1 |
| | 4.0 | 3.9 | 3.9 |
| | | 3.9 | 3.9 |
| | | | 3.9 |
| H3 | 4.6 | 4.6 | 4.6 |
| | 4.4 | 4.6 | 4.6 |
| | | 4.5 | 4.6 |
| | | | 4.5 |

^a refer to Figure 4.5(a) for proton labelling scheme

Table 4.4. Limiting ^1H NMR chemical shifts of the fully bound pyridine host (ppm) obtained by fitting the pyridine-phenol titration data to a 1:1 binding isotherm.^a

| proton | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ | $\begin{bmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{bmatrix}$ |
|--------|--|--|--|
| | 4.6•3.21 | 4.8•3.22 | 4.10•3.23 |
| H1 | 8.3 | 8.3 | 8.3 |
| | 8.2 | 8.2 | 8.3 |
| |  | 8.2 | 8.2 |
| |  | | 8.2 |
| H2 | 4.0 | 4.1 | 4.0 |
| | 3.8 | 3.7 | 3.8 |
| |  | 3.6 | 3.8 |
| |  | | 3.7 |
| H3 | 4.6 | 4.6 | 4.6 |
| | 4.5 | 4.6 | 4.6 |
| |  | 4.5 | 4.6 |
| |  | | 4.5 |

^a refer to Figure 4.5(a) for numbering scheme.

Table 4.5. Limiting complexation-induced changes in pyridine host ^1H NMR chemical shift (ppm) obtained by fitting pyridine-phenol titration data to a 1:1 binding isotherm.^a

| proton | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ | $\begin{bmatrix} \text{A-D} \\ \text{A-D} \\ \text{A-D} \\ \text{A-D} \end{bmatrix}$ |
|--------|--|--|--|
| | 4.6•3.21 | 4.8•3.22 | 4.10•3.23 |
| H1 | -0.2 | -0.2 | -0.1 |
| | -0.3 | -0.2 | -0.2 |
| | | -0.2 | -0.1 |
| | | | -0.2 |
| | | | |
| H2 | -0.1 | 0.0 | 0.0 |
| | -0.2 | -0.2 | -0.1 |
| | | -0.3 | -0.1 |
| | | | -0.2 |
| | | | |
| H3 | 0.0 | 0.0 | 0.0 |
| | 0.0 | 0.0 | 0.0 |
| | | 0.0 | 0.0 |
| | | | 0.0 |
| | | | |




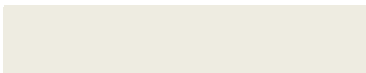
^a refer to Figure 4.5(a) for numbering scheme.

Table 4.6. ^1H NMR chemical shifts of the free pyridine *N*-oxide host (ppm) obtained by fitting the pyridine *N*-oxide titration data to a 1:1 binding isotherm.^a

| proton | $\left[\begin{array}{c} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{array} \right]$ | $\left[\begin{array}{c} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{array} \right]$ |
|--------|--|--|
| | 4.7•3.21 | 4.9•3.22 |
| | 7.8 | 7.8 |
| H1 | 7.7 | 7.7 |
| | | 7.6 |
| | | |
| H2 | ^b | 4.0 |
| | ^b | 3.9 |
| | | 3.8 |
| | | |
| H3 | 4.5 | 4.5 |
| | 4.4 | 4.5 |
| | | 4.5 |
| H4 | 6.3 | 6.2 |
| | | 6.4 |
| | | |

^a refer to Figure 4.5(b) for proton labelling scheme; ^b signal could not be followed due to signal overlap.

Table 4.7. Limiting ^1H NMR chemical shifts of the fully bound pyridine *N*-oxide host (ppm) obtained by fitting the pyridine *N*-oxide titration data to a 1:1 binding isotherm.^a

| proton | $\left[\begin{smallmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{smallmatrix} \right]$ | $\left[\begin{smallmatrix} \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \\ \text{A} \cdots \text{D} \end{smallmatrix} \right]$ |
|--------|---|---|
| | 4.7•3.21 | 4.9•3.22 |
| H1 | 7.9 | 7.8 |
| | 7.7 | 7.8 |
| |  | 7.6 |
| H2 | ^b | 3.9 |
| | ^b | 3.8 |
| |  | 3.5 |
| H3 | 4.5 | 4.6 |
| | 4.5 | 4.6 |
| |  | 4.5 |
| H4 | 6.1 | 6.0 |
| |  | 6.1 |

^a refer to Figure 4.5(b) for numbering scheme; ^b signal could not be followed due to signal overlap.

Table 4.8. Limiting complexation-induced changes in pyridine *N*-oxide ¹H NMR chemical shift (ppm) obtained by fitting the pyridine *N*-oxide titration data to a 1:1 binding isotherm.^a

| proton | $\begin{bmatrix} A-D \\ A-D \end{bmatrix}$ | $\begin{bmatrix} A-D \\ A-D \\ A-D \end{bmatrix}$ |
|--------|--|---|
| | 4.7•3.21 | 4.9•3.22 |
| H1 | 0.2 | 0.1 |
| | 0.0 | 0.2 |
| | | -0.1 |
| H2 | ^b | 0.0 |
| | ^b | 0.0 |
| | | -0.2 |
| H3 | 0.0 | 0.1 |
| | 0.1 | 0.0 |
| | | 0.0 |
| H4 | -0.2 | -0.2 |
| | | -0.3 |

^a refer to Figure 4.5(b) for numbering scheme; ^b signal could not be followed due to signal overlap.

In Figure 4.13, the association constants measured in this chapter are compared to the data from Chapter 3. For all three systems, the association constant of duplex formation increases significantly with *N*. The pyridine *N*-oxide system, shows an order of magnitude increase for each additional H-bond, and for the pyridine system there is half an order of magnitude increase for each additional H-bond. There is a linear relationship between $\log K_N$ and *N*. For all duplexes measured $K_{\text{ref}} \overline{EM}_N$ is greater than 1, indicating that cooperativity between the H-bonding sites promotes

the fully bound duplex (Tables 4.1 and 4.2). For the pyridine system the \overline{EM}_N increases slightly as the duplex gets longer but the differences are small. The \overline{EM}_N for the pyridine system is roughly twice as large as the \overline{EM}_N for the pyridine *N*-oxide system for duplexes of the same length, which somewhat compensates for the lower value of K_{ref} for the pyridine system. However, the $K_{\text{ref}} \overline{EM}_N$ for the pyridine *N*-oxide system is larger, which indicates there is a higher degree of cooperativity between the H-bonds in the pyridine *N*-oxide system relative to the pyridine system. For each value of *N*, the values of \overline{EM}_N for the pyridine and pyridine *N*-oxide systems are all larger than the corresponding values for the phosphine oxide system presented in Chapter 3. This result can be explained by the fact that the pyridine and pyridine *N*-oxide systems have fewer rotors between H-bonding sites, resulting in a more rigid structure in the duplex. Often more rigid and preorganised structures lead to higher values of *EM*.¹²

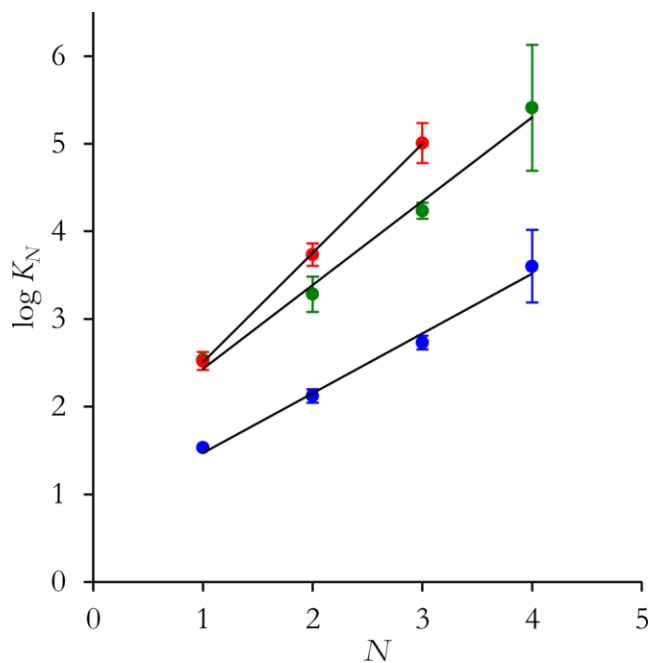


Figure 4.13. Log K_N of duplex formation as a function of the number of recognition modules in an oligomer (N), where the acceptor group is pyridine N -oxide (red), pyridine (blue) and phosphine oxide (green). Lines of best fit are shown: pyridine N -oxide, $\log K_N = 1.24N + 1.27$; pyridine, $\log K_N = 0.68N + 0.79$; phosphine oxide, $\log K_N = 0.96N + 1.48$.

4.3.3 Intramolecular folding

4.3.3.1 Dilution of DA 2-mers

As shown in Chapter 3, the effective molarity for forming an H-bond between two adjacent H-bonding sites EM_{fold} can be quantified by comparing the observed association constant of a duplex formed between two identical 2-mer strands bearing both an H-bond donor and acceptor (K_{DA} for DA•DA) and the association constant of the duplex formed between the H-bond donor and acceptor 2-mers (K_2 for DD•AA). Figures 4.14 and 4.15 show the equilibria for the pyridine AD 2-mer **4.12** and the pyridine *N*-oxide AD 2-mer **4.14**, respectively. In both cases, the duplexes resulting from the self-association of **4.12** and **4.14** are forced into an anti-parallel configuration.

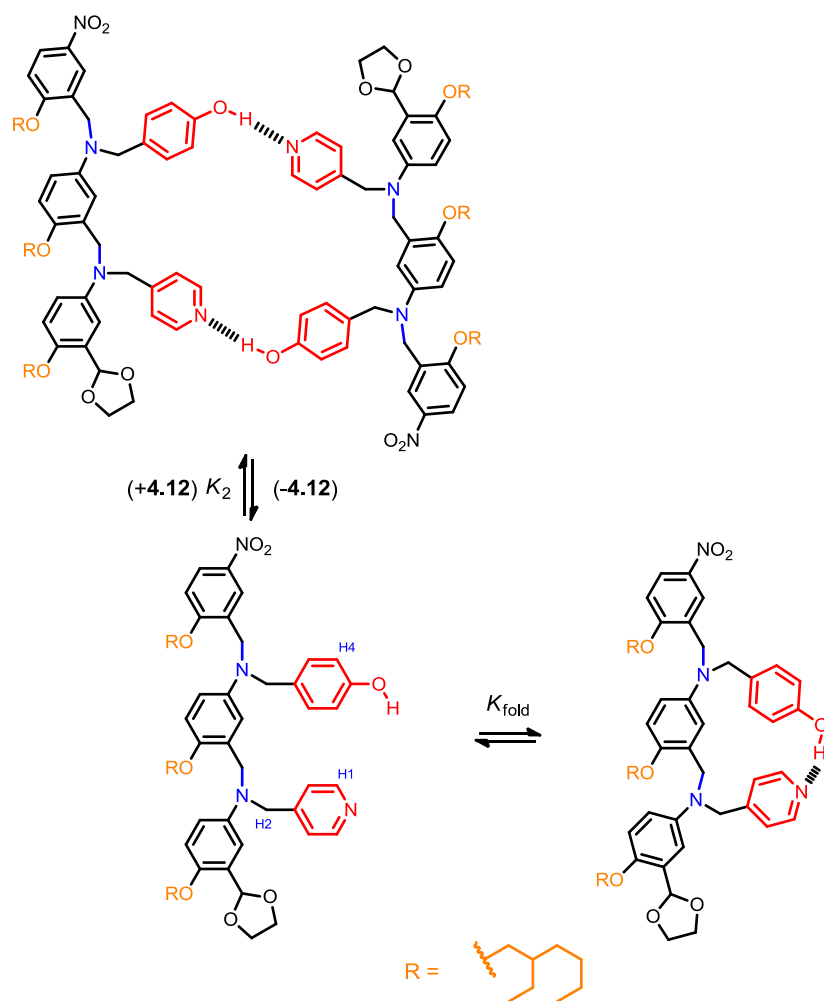


Figure 4.14. Competing equilibria of **4.12**: intramolecular folding and anti-parallel intermolecular duplex formation. Proton labelling is shown on the free unfolded structure.

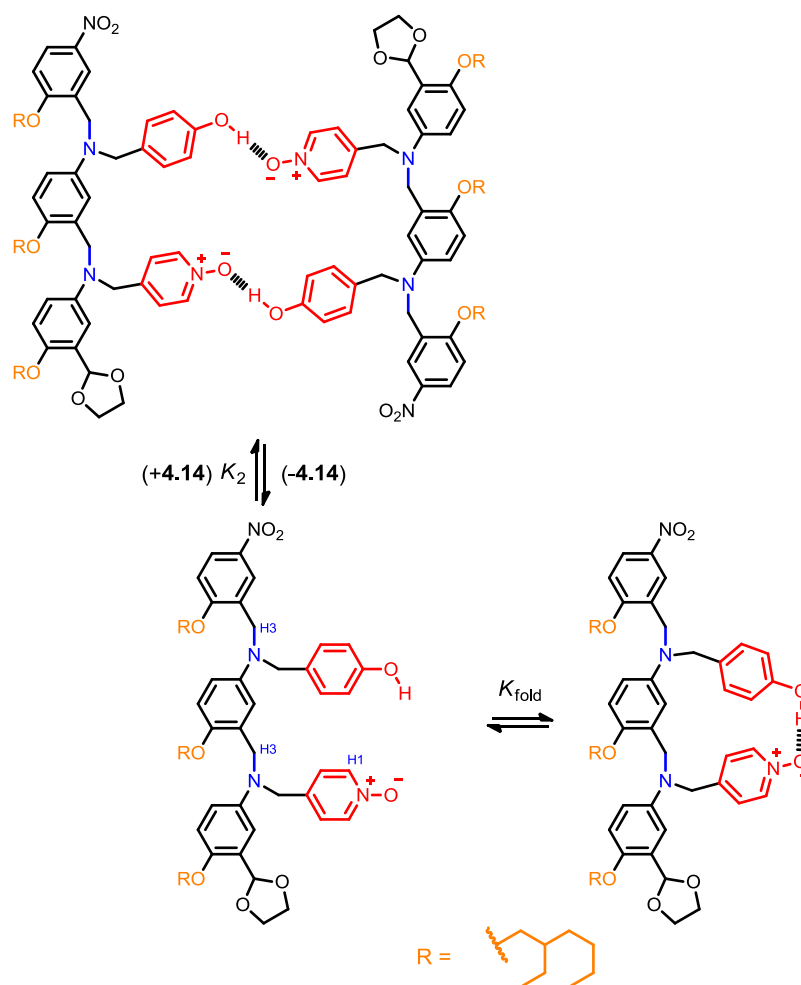


Figure 4.15. Competing equilibria of **4.14**: intramolecular folding and anti-parallel intermolecular duplex formation.

The value of K_{DA} for a DA•DA complex can be determined by fitting dilution data for DA to a dimerisation isotherm. K_{DA} will be significantly lower than K_2 if the unimolecular association constant of forming an H-bond between two adjacent H-bonding groups K_{fold} is appreciable. Figures 4.16 and 4.17 show how ^1H NMR chemical shift of selected protons for **4.12** and **4.14** vary as a function of concentration in toluene- d_6 .

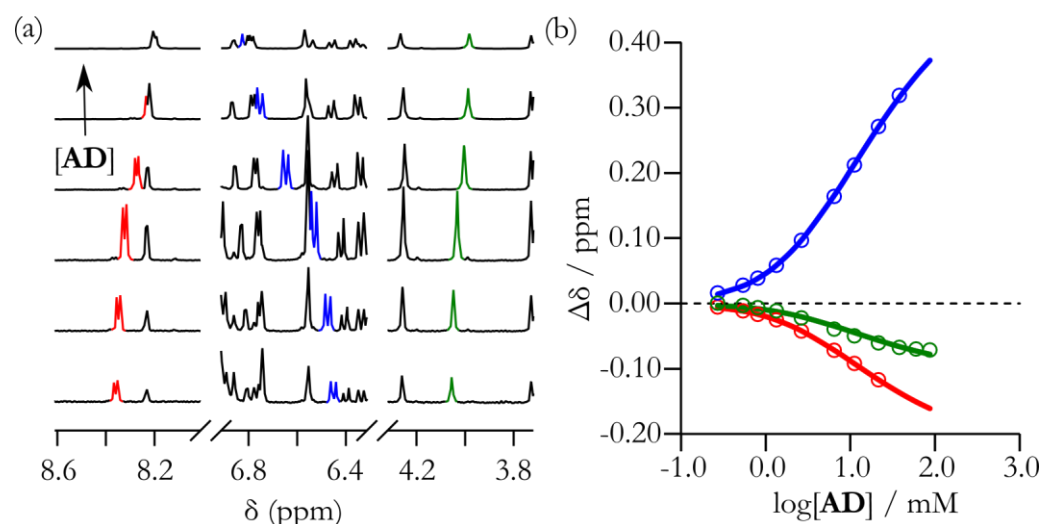


Figure 4.16. ¹H NMR data for dilution of **4.12** at 298 K in toluene-*d*₈. (a) Example 400 MHz ¹H NMR spectra. (b) Plot of the change in chemical shift as a function of guest concentration (the line represents the best fit to a dimerisation isotherm, where $K_{DA} = 54 \text{ M}^{-1}$; %bound (maximum) = 72%; %bound (minimum) = 3%). The red signal corresponds to H1; the blue signal corresponds to H4 and green corresponds to H2 (see Figure 4.14 for labelling screen). **AD = 4.12**

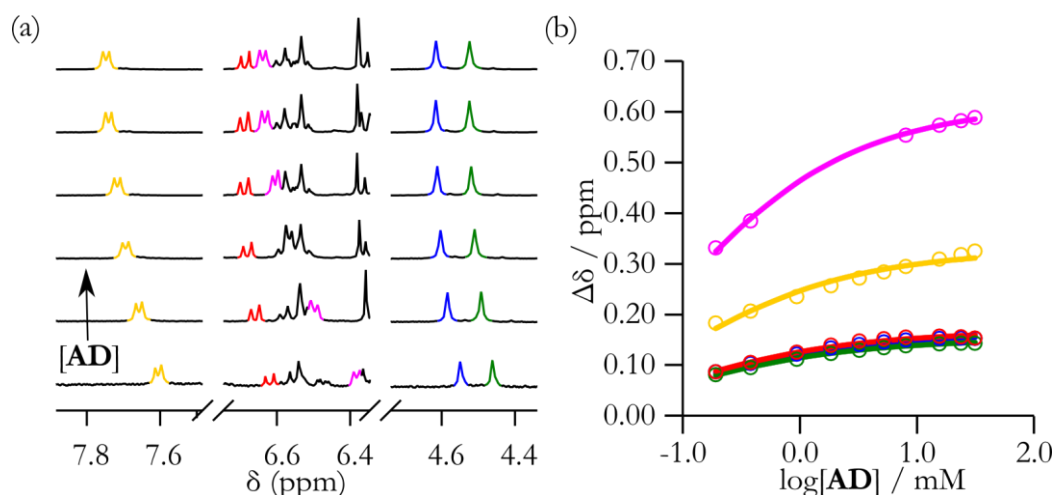


Figure 4.17. ¹H NMR data for dilution of **4.14** at 298 K in toluene-*d*₈. (a) Example 400 MHz ¹H NMR spectra. (b) Plot of the change in chemical shift as a function of guest concentration (the line represents the best fit to a dimerisation isotherm, where $K_{DA} = 6.1 \times 10^3 \text{ M}^{-1}$; maximum %bound = 95% and minimum %bound = 53%). The yellow signal is H1, and the blue and green are H3 (see Figure 4.15 for labelling scheme) **AD = 4.14**

K_{fold} cannot be measured directly but can be determined from K_{DA} and K_2 using Equation 4.10, which is derived through Equations 4.6 – 4.9. This derivation relies on the assumption given in Equation 4.5, that the equilibrium complex for forming

DA•DA from a DA 2-mer, which does not fold at all, is the same as the association constant for formation of DD•AA from DD and AA.

$$K_2 = \frac{[\text{DD}\bullet\text{AA}]}{[\text{DD}][\text{AA}]} = \frac{[\text{DA}\bullet\text{DA}]}{[\text{DA}]_{\text{open}}^2} \quad (4.5)$$

where [DD] and [AA] are the concentrations of the H-bond donor and H-bond acceptor 2-mers, respectively, [DD•AA] is the concentration of the duplex formed between them, [DA]_{open} is the concentration of the free 2-mer bearing an H-bond donor and H-bond acceptor in the unfolded open form, and [DA•DA] is the concentration of the self-complementary duplex formed between two 2-mer molecules each bearing an H-bond donor and H-bond acceptor.

Equations 4.6 and 4.7 give the definitions of K_{DA} and K_{fold} , respectively, where [DA]_{closed} is the concentration of free DA in its folded form.

$$K_{\text{DA}} = \frac{[\text{DA}\bullet\text{DA}]}{[\text{DA}]^2} \quad (4.6)$$

where [DA] is the total concentration of DA.

$$K_{\text{fold}} = \frac{[\text{DA}]_{\text{closed}}}{[\text{DA}]_{\text{open}}} \quad (4.7)$$

Using Equation 4.7, the total concentration of DA given by [DA] in Equation 4.6, can be expressed in terms of K_{fold} and [DA]_{open} (Equation 4.8) which, using Equation 4.5, allows Equation 4.6 to be written as Equation 4.9.

$$\begin{aligned} [\text{DA}] &= [\text{DA}]_{\text{open}} + [\text{DA}]_{\text{closed}} \\ &= [\text{DA}]_{\text{open}} + K_{\text{fold}} [\text{DA}]_{\text{open}} \\ &= (1 + K_{\text{fold}}) [\text{DA}]_{\text{open}} \end{aligned} \quad (4.8)$$

$$\begin{aligned}
 K_{\text{DA}} &= \frac{[\text{DA} \bullet \text{DA}]}{[\text{DA}]_{\text{open}}^2} \cdot \frac{1}{(1 + K_{\text{fold}})^2} \\
 &= \frac{K_2}{(1 + K_{\text{fold}})^2}
 \end{aligned}
 \tag{4.9}$$

Rearranging Equation 4.9 allows K_{fold} to be expressed in terms of experimentally determined values of K_2 and K_{DA} (Equation 4.10).

$$K_{\text{fold}} = \sqrt{\frac{K_2}{K_{\text{DA}}}} - 1
 \tag{4.10}$$

The effective molarity of folding EM_{fold} is given by the ratio of K_{fold} relative to the intermolecular association K_{ref} (Equation 4.11).

$$EM_{\text{fold}} = \frac{K_{\text{fold}}}{K_{\text{ref}}}
 \tag{4.11}$$

The dilution of the pyridine AD 2-mer **4.12** gives an observed association constant of $K_{\text{DA}} = 54 \text{ M}^{-1}$, which is lower than K_2 for the pyridine AA•DD complex ($K_2 = 130 \text{ M}^{-1}$). Using Equations 4.10 and 4.11, gives values of $K_{\text{fold}} = 0.6$ (dimensionless) and $EM_{\text{fold}} = 17 \text{ mM}$ for the pyridine DA oligomer **4.12**. This result means that the effective molarity of forming a bimolecular duplex ($EM_2 = 57 \text{ mM}$ in Table 4.1) is greater than for intramolecular folding. The chemical shift changes of the pyridine DA oligomer **4.12** have a similar pattern to the observed changes in the titration of **4.6•3.21** in Table 4.5: both the H1 and H2 protons have a downfield shift ($\Delta\delta = -0.2 \text{ ppm}$ and $\Delta\delta = -0.1 \text{ ppm}$, respectively, see Figure 4.14 for proton labelling scheme). For **4.12**, the free chemical shifts of H1 and H2 ($\delta_{\text{free}} = 8.4$ and ($\delta_{\text{free}} = 4.1$, respectively) and the limiting complexation induced bound chemical shifts ($\delta_{\text{bound}} = 8.2$ and $\delta_{\text{bound}} = 3.9$, respectively) are very similar to the corresponding chemical shifts for **4.6•3.21** (protons H1 and H2, respectively in Tables 4.3 and 4.4).

The dilution of the pyridine *N*-oxide DA oligomer **4.14** gives an observed association constant of $K_{\text{DA}} = 6.1 \times 10^3 \text{ M}^{-1}$ which is, within error, the same as the association constant of **4.7•3.21** ($K_2 = 5.4 \pm 1.6 \times 10^3 \text{ M}^{-1}$, Table 4.2), suggesting that **4.14** does not fold. Furthermore, the limiting complexation induced change in chemical shift of the H1 and H3 protons in **4.14** (see Figure 4.15 for proton labelling scheme) show a downfield shift ($\Delta\delta = 0.3 \text{ ppm}$ and $\Delta\delta = 0.2$, respectively) which is similar to the changes observed for the same protons in the titration of **4.7•3.21** (Table 4.8). Also the free chemical shifts of H1 and H3 in **4.12** ($\delta_{\text{free}} = 7.4$ and $\delta_{\text{free}} = 4.4$, respectively) and the limiting complexation induced bound chemical shifts ($\delta_{\text{bound}} = 7.8$ and $\delta_{\text{bound}} = 4.5$, respectively) are very similar to the corresponding chemical shifts for **4.7•3.21** (protons H1 and H3, respectively in Tables 4.6 and 4.7).

The implication from the dilutions of **4.12** and **4.14** is that neither of these systems show significant H-bond formation between adjacent H-bonding sites in solution.

4.3.3.2 Crystal structure of 4.17

Work by my colleague, Mike Jinks, provides further evidence that the pyridine *N*-oxide system does not fold. The analogue of **4.14** where the 2-ethylhexyloxy side chains have been replaced for methoxy groups (**4.17** in Figure 4.18) forms crystals from slow evaporation of a solution of **4.17** in chloroform. Figure 4.19 shows the single crystal X-ray structure of **4.17**.¹⁶

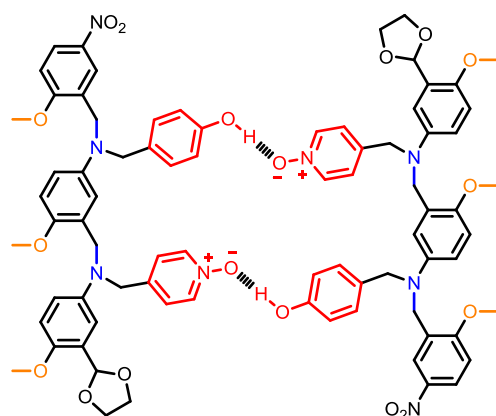


Figure 4.18. Self-complementary anti-parallel duplex of **4.17** which is an analogue of **4.14** where the solubilising groups are methoxy groups.

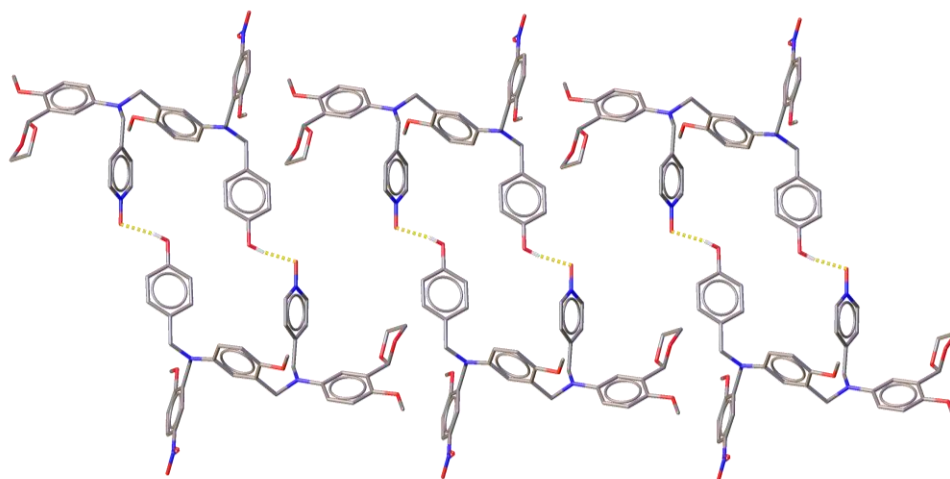


Figure 4.19. Single crystal X-ray structure of **4.17**. Three adjacent unit cells (joined along the *c*-axis edge of the unit cell) with a total of 6 molecules of **4.17** are shown. Hydrogen atoms have been omitted for clarity. Oxygen, nitrogen and carbon atoms are shown in red, blue and grey respectively. The intermolecular H-bond between the phenol and pyridine-*N*-oxide is shown in yellow.¹⁶

In the crystal structure of **4.17**, each molecule is arranged in an antiparallel conformation relative to another molecule of **4.17** and adopts an H-bonded duplex structure, where two H-bonds are formed with the same connectivity as shown in the two dimensional representation in Figure 4.18. The fact that **4.17** adopts the structure shown in Figure 4.19 adds further evidence that K_{fold} for H-bond formation between adjacent phenol and pyridine *N*-oxide groups in the same molecule is too low for folding to compete with duplex formation.

4.4 Conclusions

This Chapter has shown that the modular design presented in the previous Chapter is tuneable and versatile. The requirement that $K_{\text{ref}} \overline{EM}_N$ be greater than 1 is satisfied even when the relatively weak H-bond acceptor pyridine is used ($K_{\text{ref}} = 34 \text{ M}^{-1}$). The effect of fewer rotors between H-bonding sites increases the EM to compensate for the decrease in K_{ref} . Furthermore, the reduction in flexibility has dramatically reduced the impact of folding between adjacent H-bonding sites to the point that folding does not compete effectively with duplex formation. Duplex formation for the pyridine *N*-oxide system has been shown to occur in both solution and the solid state.

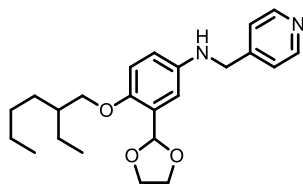
The obvious next step in the pursuit of an effective information oligomer is to establish how the system is able to distinguish between different H-bonding sequences, beyond a simple DA dimer. A system is required, which has a high $K_{\text{ref}} \overline{EM}_N$ for each additional H-bond formed and where intramolecular folding does not compete effectively with duplex formation. Based on the results in Chapters 3 and 4, the pyridine *N*-oxide system seems an ideal candidate for testing mixed sequence oligomers. This approach will be explored further in the next Chapter.

4.5 Experimental section

4.5.1 Synthesis

All the reagents and materials used in the synthesis of the compounds described below were bought from commercial sources, without prior purification. Thin layer chromatography was carried out using with silica gel 60F (Merck) on aluminium. Flash chromatography was carried out on silica gel 40 – 60 μm (BDH) or on an automated system (Combiflash Companion) using pre-packed cartridges of silica (50 μ PuriFlash® Column). All NMR spectroscopy was carried out on either a Bruker AVI250, AVI400, DPX400, AVIII400 or DRX500 spectrometer using the residual solvent as the internal standard. All chemical shifts (δ) are quoted in ppm and coupling constants given in Hz. Splitting patterns are given as follows: s (singlet), d (doublet), t (triplet), m (multiplet). FT-IR spectra were measured on a PerkinElmer Spectrum 100 spectrometer. ES+ was carried out on a Micromass Platform spectrometer. Reactions were carried out at ambient temperature unless otherwise stated.

4.5.1.1 Synthesis of 4.1



3.10 (11.5 g, 39.3 mmol, 1 equiv.) and 4-nicotinaldehyde (4.1 mL, 43.2 mmol, 1.1 equiv.) were dissolved in CHCl_3 (80 mL) in the presence of molecular sieves with stirring. After 6 hours the solution was filtered and the solvent removed by rotary evaporator. This crude mixture was dissolved in MeOH (150 mL) and then NaBH_4 (4.45 g, 118 mmol, 3 equiv.) was added slowly at 0 °C with stirring. This mixture was stirred for 2 hours before the solution was neutralized using concentrated aqueous HCl. This solution was washed with EtOAc (5×20 mL) and all the organic extracts were combined and washed with brine (1×20 mL) and then dried (MgSO_4). The crude oil was purified via flash chromatography on silica eluting with a gradient from 75% to 100% EtOAc in hexane to yield a golden oil (11.4 g, 76%).

$^1\text{H NMR}$ (250 MHz, CDCl_3): $\delta_{\text{H}} = 8.53$ (d, 2H, $J = 6.0$), 7.28 (d, 2H, $J = 6.0$), 6.85 (d, 1H, $J = 3.0$), 6.74 (d, 1H, $J = 9.0$), 6.48 (dd, 1H, $J = 9.0, J = 3.0$), 6.11 (s, 1H), 4.33 (s, 2H), 3.96 – 4.10 (m, 4H), 3.79 (d, 2H, $J = 5.5$), 1.67 – 1.77 (m, 1H), 1.24 – 1.54 (m, 8H), 0.86 – 0.95 (m, 6H);

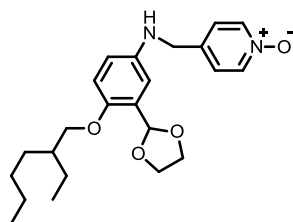
$^{13}\text{C NMR}$ (101 MHz, CDCl_3): $\delta_{\text{C}} = 150.4, 150.0, 149.2, 141.4, 127.2, 122.2, 114.1, 113.8, 112.2, 99.2, 71.9, 65.2, 47.9, 39.5, 30.5, 29.1, 23.9, 23.1, 14.1, 11.1$;

MS (ES+): m/z (%) = 385 (100) $[\text{M} + \text{H}^+]$, 426 (40);

HRMS (ES+): calculated for $\text{C}_{23}\text{H}_{33}\text{N}_2\text{O}_3$, 385.2491, found 385.2494;

FT-IR (ATR): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2926, 2859, 1680, 1603, 1496, 1465, 1388, 1253, 1224, 1174, 1068, 1030.

4.5.1.2 Synthesis of 4.2



3.10 (1.008 g, 3.44 mmol, 1 equiv.) and 4-formylpyridine-*N*-oxide (0.508 g, 4.12 mmol, 1.2 equiv.) were dissolved in CHCl_3 (12 mL) with stirring and $\text{NaBH}(\text{OAc})_3$ (2.04 g, 9.62 mmol, 2.8 equiv.) was added at room temperature. After 1 h the reaction was quenched with saturated aq. NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0 to 8% of MeOH in CHCl_3 to yield a yellow powder (0.91 g, 66%) Mp 126 – 134 °C.

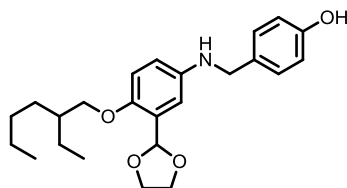
^1H NMR (400 MHz, CDCl_3): δ_{H} = 8.16 (d, 2H, $^3J = 7.0$, Ar-H), 7.29 (d, 2H, $^3J = 7.0$, Ar-H), 6.85 (d, 1H, $^4J = 3.0$, Ar-H), 6.75 (d, 1H, $^3J = 9.0$, Ar-H), 6.48 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 6.08 (s, 1H, OCHO), 4.32 (s, 2H, CH_2N), 4.13 – 3.96 (m, 4H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.80 (d, 2H, $^3J = 6$, CH_2O), 1.76 – 1.66 (m, 1H, Alk-CH), 1.55 – 1.23 (m, 8H, Alk- CH_2), 0.94 – 0.85 (m, 6H, Alk- CH_3);

^{13}C NMR (101 MHz, CDCl_3): δ_{C} = 150.7, 141.0, 140.1, 139.2, 127.3, 124.6, 114.3, 113.9, 112.3, 99.2, 72.0, 65.3, 47.2, 39.6, 30.6, 29.1, 24.0, 23.2, 14.2, 11.2;

HRMS (ES+): calculated for $\text{C}_{23}\text{H}_{33}\text{N}_2\text{O}_4$ 401.2440, found 401.2435;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 3055, 2987, 2686, 2411, 2306, 1505, 1483, 1422, 1265, 1168, 1071, 896.

4.5.1.3 Synthesis of 4.3



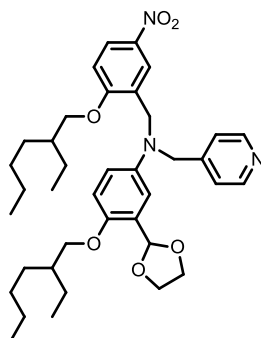
3.11 (3.66 g, 6.58 mmol, 1 equiv.) was dissolved in THF (20 mL) at 0 °C and TBAF (13.2 mL, 13.2 mmol, 2 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 × 10 mL). All organic fractions were combined and washed with brine (1 × 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator. The crude mixture was then purified via recrystallization from hot DCM and hexane to yield a yellow solid (1.80 g, 69%) Mp 110 – 114 °C.

¹H NMR (250 MHz, CDCl₃): δ_H = 7.23 (d, 2H, ³J = 8.5, Ar-H), 6.90 (d, 1H, ⁴J = 3.0, Ar-H), 6.81 – 6.73 (m, 3H, Ar-H), 6.61 (dd, 1H, ³J = 9.0, ⁴J = 3.0, Ar-H), 6.15 (s, 1H, OCHO), 4.20 (s, 2H, CH₂N), 4.17 – 3.98 (m, 4H, OCH₂CH₂O), 3.82 (d, 2H, CH₂O), 1.80 – 1.66 (m, 1H, Alk-CH), 1.62 – 1.22 (m, 8H, Alk-CH₂), 0.99 – 0.86 (m, 6H, Alk-CH₃);

¹³C NMR (101 MHz, CDCl₃): δ_C = 154.8 (Ar-C), 150.2 (Ar-C), 142.1 (Ar-C), 131.6 (Ar-C), 129.1 (Ar-C), 126.9 (Ar-C), 115.4 (Ar-C), 114.5 (Ar-C), 113.9 (Ar-C), 112.2 (Ar-C), 99.4 (OCHO), 72.0 (CH₂O), 65.2 (OCH₂CH₂O), 48.8 (CH₂N), 39.5 (Alk-CH), 30.5 (Alk-CH₂), 29.1 (Alk-CH₂), 23.9 (Alk-CH₂), 23.1 (Alk-CH₂), 14.1 (Alk-CH₃), 11.1 (Alk-CH₃);

HRMS (ES⁺): calculated for C₂₄H₃₄NO₄ 400.2488, found 400.2495;

FT-IR (thin film): ν_{max}/cm⁻¹ 3290, 2952, 2921, 2873, 2853, 1611, 1598, 1493, 1469, 1243, 1074

4.5.1.4 Synthesis of **4.4**

4.1 (0.913 g, 2.37 mmol, 1 equiv.) and **3.8** (1.061 g, 3.80 mmol, 1.6 equiv.) were dissolved in DCE (8.5 mL) and $\text{NaBH}(\text{OAc})_3$ (1.41 g, 6.6 mmol, 2.8 equiv.) was added with stirring. After 18 hours the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 50% of EtOAc in hexane to yield a golden yellow oil (1.4 g, 92%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} = 8.53 (d, 2H, $J = 6.0$), 8.15 (dd, 1H, $J = 9.0$, $J = 3.0$), 8.11 (d, 1H, $J = 3.0$), 7.21 (d, 2H, $J = 6.0$), 6.93 (d, 1H, $J = 9.0$), 6.91 (d, 1H, $J = 3.0$), 6.75 (d, 1H, $J = 9.0$), 6.57 (dd, 1H, $J = 9.0$), 6.10 (s, 1H), 4.62 (s, 2H), 4.55 (s, 2H), 3.98 (d, 2H, $J = 6.0$), 3.93 (s, 4H), 3.80 (d, 2H, $J = 6.0$), 1.66 – 1.83 (m, 2H), 1.23 – 1.53 (m, 16H), 0.83 – 0.96 (m, 12H);

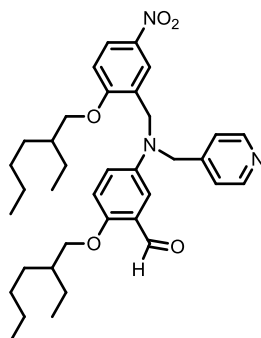
$^{13}\text{C NMR}$ (101 MHz, CDCl_3): δ_{C} = 161.8, 150.4, 149.9, 148.4, 142.1, 141.4, 128.0, 127.1, 124.7, 123.4, 122.1, 114.8, 113.5, 112.2, 110.5, 99.4, 71.5, 65.0, 54.6, 50.6, 39.4, 39.1, 30.5, 29.0, 23.9, 23.0, 22.9, 14.0, 11.1;

MS (ES⁺): m/z (%) = 648 (100) [$\text{M} + \text{H}^+$], 689 (30) [$\text{MH}^+ + \text{CH}_3\text{CN}$];

HRMS (ES⁺): calculated for $\text{C}_{38}\text{H}_{54}\text{N}_3\text{O}_6$ 648.4013, found 648.4016;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2927, 2859, 1681, 1592, 1501, 1463, 1338, 1264,
1230, 1179, 1074, 1014.

4.5.1.5 Synthesis of 4.4'



4.4 (1.51 g, 2.3 mmol, 1 equiv.) was dissolved in CHCl₃ (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 18 hours the mixture was neutralised using aqueous NaHCO₃ and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl₃ (3 × 10 mL) before all organic fractions were washed with brine (1 × 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator to yield an intense yellow oil (1.38 g, 98%) requiring no further purification.

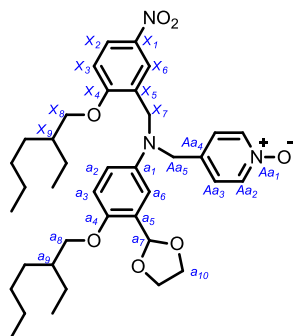
¹H NMR (500 MHz, CDCl₃): δ_H = 10.44 (s, 1H), 8.54 (d, 1H, *J* = 6.0), 8.15 (dd, 1H, *J* = 9.0), 8.00 (d, 1H, *J* = 3.0), 7.18 (d, 2H, *J* = 6.0), 7.13 (d, 1H, *J* = 2.5), 6.95 (d, 1H, *J* = 9.0), 6.87 (d, 1H, *J* = 9.0), 6.84 (dd, 1H, *J* = 9.0, *J* = 3.0), 4.64 (s, 2H), 4.59 (s, 2H), 4.00 (d, 2H, *J* = 5.5), 3.85 – 3.91 (m, 2H), 1.69 – 1.82 (m, 2H), 1.23 – 1.53 (m, 16H), 0.84 – 0.96 (m, 12H);

¹³C NMR (126 MHz, CDCl₃): δ_C = 189.7, 161.9, 155.1, 150.0, 147.6, 142.1, 141.3, 127.2, 125.3, 124.8, 122.9, 121.8, 120.8, 114.3, 110.9, 110.6, 71.6, 71.2, 54.2, 50.2, 39.5, 39.1, 30.6, 30.4, 29.0, 28.9, 23.9, 23.8, 22.9, 22.8, 14.0, 13.9, 11.1, 11.0;

MS (ES⁺): *m/z* (%) = 604 (100) [M + H⁺], 645 (50) [MH⁺ + CH₃CN];

HRMS (ES⁺): calculated for C₃₆H₅₀N₃O₅ 604.3750, found 604.3740;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2923, 2855, 1680, 1598, 1501, 1464, 1339, 1265,
1228, 1180, 1015.

4.5.1.6 Synthesis of **4.5**

4.2 (0.19 g, 0.48 mmol, 1 equiv.) and **3.8** (0.56 g, 1.99 mmol, 4 equiv.) were dissolved in CHCl_3 (1.7 mL) and $\text{NaBH}(\text{OAc})_3$ (0.28 g, 1.34 mmol, 2.8 equiv.) was added at room temperature with stirring. After 18 h the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed on a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0 to 4% of MeOH in DCM to yield a golden yellow oil (0.305 g, 96%).

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} =$ 8.17 (d, 2H, $^3J = 6.5$, H-Aa2), 8.16 (dd, 1H, $^3J = 6.0$, $^4J = 3.0$, H-X2) 8.08 (d, 1H, $^4J = 3.0$) 7.20 (d, 2H, $^3J = 6.5$, H-Aa3) 6.95 (d, 1H, $^4J = 3.0$ H-a6), 6.94 (d, 1H, $^3J = 9.0$, H-X3), 6.76 (d, 1H, $^3J = 9.0$, H-a3) 6.58 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-a2) 6.07 (s, 1H, H-a7) 4.57 (s, 2H, H-Aa5) 4.51 (s, 2H, H-X7) 4.03 – 3.93 (m, 6H, H-X8 and H-a10) 3.81 (d, 2H, $^3J = 5.5$, H-a8) 1.83 – 1.68 (m, 2H, H-X9 and H-a9) 1.54 – 1.24 (m, 16H, Alk- CH_2) 0.98 – 0.84 (m, 12H, Alk- CH_3);

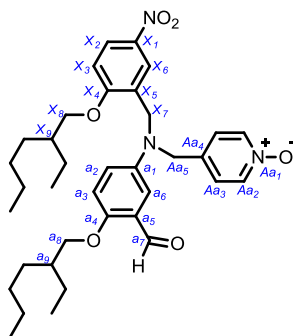
^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} =$ 162.1 (1C, Ar- C^{quat}), 151.2 (1C, Ar- C^{quat}), 141.8 (1C, Ar- C^{quat}), 141.6 (1C, Ar- C^{quat}), 139.4 (2C, C-Aa2), 127.8 (1C, Ar- C^{quat}), 127.3 (1C, Ar- C^{quat}), 125.1 (1C, C-X2), 124.8 (2C, C-Aa3), 123.7 (1C, C-X6), 115.8 (1C C-a2), 113.7 (1C C-a3), 113.1 (1C C-a6), 110.9 (1C, C-X3), 99.4 (1C C-a7), 71.8 (2C, C-X8 and C-a8), 65.4 (2C, C-a10), 54.3 (1C C-Aa5), 51.0 (1C C-X7), 39.6 (1C C-X9), 39.4

(1C C-a9), 30.7 (2C, Alk-CH₂), 29.2 (2C, Alk-CH₂) 24.1 (2C, Alk-CH₂), 23.3 (1C, Alk-CH₂), 23.2 (1C, Alk-CH₂), 14.3 (2C, Alk-CH₃), 11.3 (2C, Alk-CH₃);

HRMS (ES+): calculated for C₃₈H₅₄N₃O₇ 664.3962, found 664.3990;

FT-IR (thin film): ν_{max} /cm⁻¹ 3020, 2962, 2931, 2401, 2254, 1594, 1516, 1484, 1382, 1342, 1266, 1216, 1083, 909.

4.5.1.7 Synthesis of 4.5'



4.5 (0.255 g, 0.384 mmol, 1 equiv.) was dissolved in CHCl_3 (5 mL) and concentrated aqueous HCl (5 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed using a rotary evaporator to yield a yellow oil (0.240 g, 99%) requiring no further purification.

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 10.41$ (s, 1H, H-a7), 8.17 – 8.11 (m, 3H, H-Aa2 and H-X2), 7.94 (d, 1H, $^4J = 3.0$, H-X6), 7.16 (d, 2H, $^3J = 7.0$, H-Aa3), 7.11 (d, 1H, $^4J = 3.0$, H-a6), 6.94 (d, 1H, $^3J = 9.0$, H-X3), 6.86 (d, 1H, $^3J = 9.0$, H-a3), 6.82 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-a2) 4.58 (s, 2H, CH_2N), 4.54 (s, 2H, CH_2N), 3.98 (d, 2H, $^3J = 5.5$, CH_2O), 3.87 (d, 2H, $^3J = 5.5$, CH_2O), 1.83 – 1.66 (m, 2H, Alk-CH), 1.53 – 1.18 (m, 16H, Alk- CH_2), 0.96 – 0.80 (m, 12H, Alk- CH_3);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 189.8$ (Ar-CHO), 162.1 (Ar-C), 155.5 (Ar-C), 141.8 (Ar-C), 141.5 (Ar-C), 139.4 (Ar-C), 137.9 (Ar-C), 127.0 (Ar-C), 125.5 (Ar-C), 125.2 (Ar-C), 124.5 (Ar-C), 123.1 (Ar-C), 121.5 (Ar-C), 114.5 (Ar-C), 111.7 (Ar-C), 110.9 (Ar-C), 71.8 (CH_2O), 71.4 (CH_2O), 53.8 (CH_2N), 50.6 (CH_2N), 39.6 (Alk-CH), 39.3 (Alk-CH), 30.7 (Alk- CH_2), 30.6 (Alk- CH_2), 29.2 (Alk- CH_2), 29.1 (Alk- CH_2), 24.1

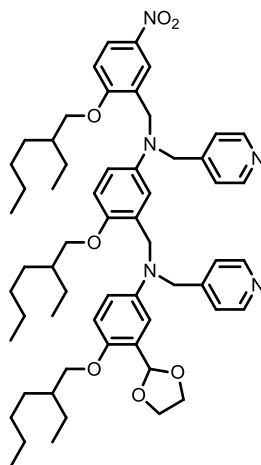
(Alk-CH₂), 24.1 (Alk-CH₂), 23.1 (Alk-CH₂), 23.1 (Alk-CH₂), 14.2 (Alk-CH₃), 14.2

(Alk-CH₃), 11.3 (Alk-CH₃), 11.2 (Alk-CH₃);

HRMS (ES+): calculated for C₃₆H₅₀N₃O₆ 620.3700, found 620.3704;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 3055, 2988, 2306, 1422, 1265, 896.

4.5.1.8 Synthesis of 4.6



4.4' (0.913 g, 2.37 mmol, 1 equiv.) and **4.1** (1.061 g, 3.80 mmol, 1.6 equiv.) were dissolved in DCE (8.5 mL) and NaBH(OAc)₃ (1.41 g, 6.6 mmol, 2.8 equiv.) was added with stirring. After 18 hours the reaction was quenched with saturated aqueous NaHCO₃ solution, and extracted into CHCl₃ (4 × 10 mL). All the organic fractions were washed with water (1 × 10 mL), brine (1 × 10 mL) and dried (MgSO₄) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 50% of EtOAc in hexane to yield a golden yellow oil (1.4 g, 92%).

¹H NMR (500 MHz, CDCl₃): δ_H = 8.51 (d, 2H, *J* = 6.0), 8.47 (d, 2H, *J* = 6.0), 8.07 (dd, 1H, *J* = 9.0, 3.0), 7.94 (d, 1H, *J* = 3.0), 7.12 – 7.08 (m, 4H), 6.84 (d, 1H, *J* = 9.0), 6.72 (d, 1H, *J* = 9.0), 6.68 (d, 1H, *J* = 3.0), 6.59 (d, 1H, *J* = 9.0), 6.46 (dd, 1H, *J* = 9.0, 3.0), 6.37 – 6.33 (m, 2H), 6.07 (s, 1H), 4.53 – 4.49 (m, 4H), 4.47 (s, 2H), 4.18 (s, 2H), 3.94 (d, 2H, *J* = 6.0), 3.93 – 3.90 (m, 4H), 3.80 (d, 2H, *J* = 6.0), 3.76 (d, 2H, *J* = 5.5), 1.80 – 1.70 (m, 2H), 1.68 – 1.61 (m, 1H), 1.57 – 1.19 (m, 24H), 0.98 – 0.81 (m, 18H);

¹³C NMR (126 MHz, CDCl₃): δ_C = 161.5, 149.8, 149.7, 149.6, 149.4, 148.7, 142.4, 141.6, 141.2, 127.6, 126.8, 126.6, 124.5, 122.8, 121.9, 113.9, 113.4, 113.0, 112.3, 111.6,

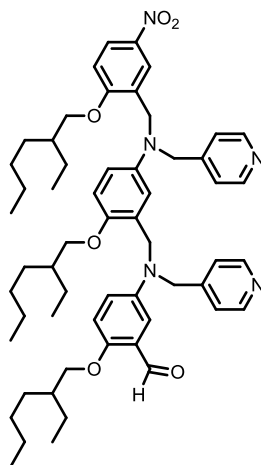
111.2, 110.6, 99.3, 71.7, 71.4, 70.7, 65.0, 55.1, 53.7, 50.8, 50.0, 39.5, 39.4, 39.2, 30.6,
30.5, 29.1, 29.0, 24.0, 23.9, 23.9, 23.0, 22.9, 22.9, 14.0, 14.0, 11.1, 11.1;;

MS (ES+): m/z (%) = 487 (20), 507 (80), 973 (100) [M + H⁺], 995 (30) [M + Na⁺];

HRMS (ES+): calculated for C₅₉H₈₂N₅O₇ 972.6214, found 972.6240;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2958, 2928, 2872, 1681, 1599, 1504, 1465, 1340, 1266,
1228, 1066, 1026.

4.5.1.9 Synthesis of 4.6'



4.6 (1.51 g, 2.3 mmol, 1 equiv.) was dissolved in CHCl₃ (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 18 hours the mixture was neutralised using aqueous NaHCO₃ and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl₃ (3 × 10 mL) before all organic fractions were washed with brine (1 × 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator to yield an intense yellow oil (1.38 g, 98%) requiring no further purification.

¹H NMR (400 MHz, CDCl₃): δ_H = 10.42 (s, 1H), 8.53 – 8.47 (m, 4H), 8.05 (dd, 1H, *J* = 9.0, 3.0), 7.90 (d, 1H, *J* = 3.0), 7.11 (d, 2H, *J* = 6.0) 7.07 (d, 2H, *J* = 6.0), 6.89 (d, 1H, *J* = 3.0), 6.80 (d, 1H, *J* = 9.0), 6.73 (d, 1H, *J* = 9.0), 6.67 (d, 1H, *J* = 9.0), 6.59 (dd, 1H, *J* = 9.0, 3.0), 6.47 (dd, 1H, *J* = 9.0, 3.0), 6.27 (d, 1H, *J* = 3.0), 4.54 (s, 2H), 4.48 (s, 4H), 4.27 (s, 2H), 3.91 (d, 2H, *J* = 6.0), 3.88 (d, 2H *J* = 6.0), 3.78 (d, 2H, *J* = 6.0), 1.82 – 1.71 (m, 2H), 1.71 – 1.61 (m, 1H), 1.59 – 1.19 (m, 24H), 1.01 – 0.81 (m, 18H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 189.9, 161.5, 154.3, 149.9, 149.7, 149.6, 148.4, 148.3, 142.2, 141.7, 141.2, 127.6, 126.1, 125.0, 124.5, 122.8, 121.8, 121.7, 120.2, 113.9, 112.6, 112.4, 111.7, 110.4, 109.6, 71.5, 71.3, 70.7, 55.0, 53.6, 50.9, 49.9, 39.6, 39.4,

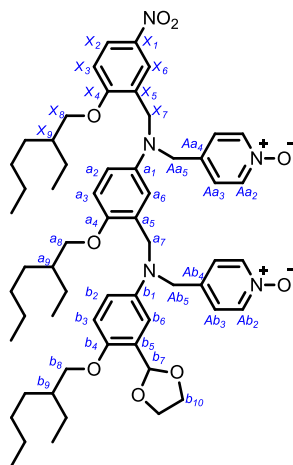
39.2, 30.6, 30.6, 30.5, 29.1, 29.0, 29.0, 24.0, 24.0, 23.9, 23.0, 23.0, 22.9, 14.1, 14.0,
11.2, 11.1;

MS (ES+): m/z (%) = 485 (85) , 506 (50), 929 (100) [M + H⁺], 951 (40) [M + Na⁺];

HRMS (ES+): calculated for C₅₇H₇₈N₅O₆ 928.5952, found 928.5937;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2926, 2858, 1680, 1598, 1500, 1463, 1414, 1338,
1264, 1226, 1179, 1015, 969.

4.5.1.10 Synthesis of 4.7



4.5' (0.50 g, 0.80 mmol, 1 equiv.) and **4.2** (0.48 g, 1.2 mmol, 1.5 equiv.) were dissolved in CHCl_3 (2.9 mL) with stirring and $\text{NaBH}(\text{OAc})_3$ (0.476 g, 2.3 mmol, 2.8 equiv.) was added at room temperature. After 18 h the reaction was quenched with saturated aq. NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed using a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient of 0 to 10% of MeOH in a 1:1 mixture of acetonitrile and CHCl_3 to yield a golden yellow oil (0.47 g, 59%).

^1H NMR (500 MHz, CDCl_3): δ_{H} = 8.08 (d, 4H, $^3J = 6.$ and $^3J = 6.5$, H-Aa2 and H-Ab2), 8.07 (dd, 1H, $^3J = 9.0$, $^2J = 3.0$), 7.88 (d, 1H, $^3J = 3.0$, H-X6) 7.09 (d, 2H, $^3J = 6.5$, H-Ab3), 7.02 (d, 2H, $^3J = 6.5$), 6.88 (d, 1H, $^3J = 9.0$, H-X3), 6.75 (d, 1H, $^4J = 3.0$, H-b6), 6.72 (d, 1H, $^3J = 9.0$, H-a3), 6.65 (d, 1H, $^3J = 9.0$, H-b3), 6.46 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-a2), 6.39 (dd, 1H, $^3J = 9.0$, $^3J = 3.0$, H-b2), 6.32 (d, 1H, $^4J = 3.0$, H-a6), 6.04 (s, 1H, H-b7), 4.48 (s, 2H, H-X7), 4.45 (s, 2H, H-a7), 4.41 (s, 2H, H-Aa5), 4.19 (s, 2H, H-Ab5) 4.03 – 3.93 (m, 4H, H-b10), 3.94 (d, 2H, $^3J = 5.5$, H-X8), 3.82 (d, 2H, $^3J = 5.5$, H-b8) 3.76 (d, 2H, $^3J = 5.5$, H-a8), 1.80 – 1.70 (2H, H-X9

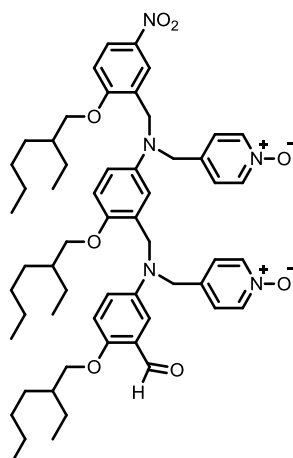
and H-b9), 1.69 – 1.61 (m, 1H, H-b9), 1.18 – 1.56 (m, 24H, Alk-CH₂), 0.97 – 0.80 (m, 18H, Alk-CH₃);

¹³C NMR (126 MHz, CDCl₃): δ_c = 161.6 (1C, s, C-X4), 150.1 (1C, s, C-b4), 150.0 (1C, s, C-a4), 142.0 (1C, s, C-b1), 141.2 (2C, s, C-X1 and C-a1), 139.1 (4C, s, C-Aa2 and C-Ab2), 138.5 (1C, s, C-Aa4), 127.3 (1C, s, C-X5), 126.6 (2C, s, C-a5 and C-b5), 124.7 (1C, s, C-X2), 124.3 (4C, s, C-Aa3, and C-Ab3), 122.9 (1C, s, C-X6), 114.4 (1C, s, C-b2), 113.5 (2C, s, C-b3 and C-a6), 112.5 (1C, s, C-a3), 112.3 (1C, s, C-a2), 111.6 (1C, s, C-b6), 110.7 (1C, s, C-X3), 99.1 (1C, s, C-b7), 71.7 (1C, s, C-b8), 71.5 (1C, s, C-X8), 70.7 (1C, s, C-a8), 65.1 (1C, s, C-b9), 54.6 (1C, s, C-Aa5), 53.1 (1C, s, C-Ab5), 50.8 (1C, s, C-X7), 49.9 (1C, s, C-a7), 39.5 (1C, s, C-X9, C-a9 or b9), 39.4 (1C, s, C-X9, C-a9 or b9), 39.2 (1C, s, C-X9, C-a9 or b9), 30.6 (1C, s, Alk-CH₂), 30.5 (2C, s, Alk-CH₂), 29.0 (3C, s, Alk-CH₂), 24.0 (1C, s, Alk-CH₂), 23.9 (2C, s, Alk-CH₂), 23.0 (1C, s, Alk-CH₂), 22.9 (2C, s, Alk-CH₂), (3C, s, Alk-CH₃), 11.1 (3C, s, Alk-CH₃);

HRMS (ES⁺): calculated for C₅₉H₈₁N₅O₉Na 1026.5932, found 1026.5944;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 3686, 3020, 2401, 1520, 1477, 1424, 1216, 1020, 929, 909.

4.5.1.11 Synthesis of 4.7'



4.7 (0.36 g, 0.36 mmol, 1 equiv.) was dissolved in CHCl₃ (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO₃ and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl₃ (3 × 10 mL) before all organic fractions were washed with brine (1 × 10 mL) and dried (MgSO₄) to yield a bright viscous yellow oil (0.34 g, 97%) requiring no further purification.

¹H NMR (500 MHz, CDCl₃): δ_H = 10.40 (s, 1H, Ar-CHO), 8.09 – 8.06 (m, 4H, Ar-H), 8.04 (dd, 1H, ³J = 9.0, ⁴J = 3.0, Ar-H), 7.82 (d, 1H, ⁴J = 3.0, Ar-H), 7.07 – 7.02 (m, 4H, Ar-H), 6.90 (d, 1H, ⁴J = 3.0, Ar-H), 6.81 (d, 1H, ³J = 9.0, Ar-H), 6.72 (d, 1H, ³J = 9.0, Ar-H), 6.69 (d, 1H, ³J = 9.0, Ar-H), 6.59 (dd, 1H, ³J = 9.0, ⁴J = 3.0, Ar-H), 6.47 (dd, 1H, ³J = 9.0, ⁴J = 3.0, Ar-H), 6.26 (d, 1H, ⁴J = 3.0, Ar-H), 4.46 (s, 2H, CH₂N), 4.45 (s, 2H, CH₂N), 4.42 (s, 2H, CH₂N), 4.26 (s, 2H, CH₂N), 3.93 – 3.84 (m, 4H, OCH₂CH₂O), 3.76 (d, 2H, ³J = 5.5, CH₂O), 3.73 – 3.69 (m, 4H, CH₂O), 1.79 – 1.61 (m, 3H, Alk-CH), 1.55 – 1.19 (24H, m Alk-CH₂), 0.97 – 0.82 (m, 18H, Alk-CH₃);

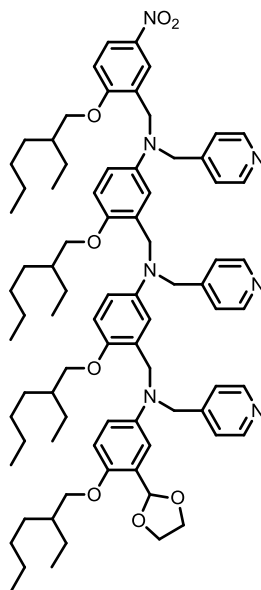
¹³C NMR (126 MHz, CDCl₃): δ_C = 189.7 (Ar-CHO), 161.5 (Ar-C), 154.6 (Ar-C), 150.0 (Ar-C), 141.7 (Ar-C), 141.2 (Ar-C), 141.2 (Ar-C), 139.2 (Ar-C), 139.1 (Ar-C),

138.1 (Ar-C), 127.2 (Ar-C), 126.1 (Ar-C), 125.0 (Ar-C), 124.6 (Ar-C), 124.2 (Ar-C), 124.1 (Ar-C), 122.8 (Ar-C), 120.5 (Ar-C), 114.0 (Ar-C), 113.2 (Ar-C), 112.5 (Ar-C), 112.5 (Ar-C), 110.6 (Ar-C), 110.0 (Ar-C), 71.6 (CH₂O), 71.4 (CH₂O), 70.7 (CH₂O), 54.4 (CH₂N), 53.0 (CH₂N), 50.8 (CH₂N), 50.0 (CH₂N), 39.5 (Alk-CH), 39.4 (Alk-CH), 39.1 (Alk-CH), 30.6 (Alk-CH₂), 30.6 (Alk-CH₂), 30.5 (Alk-CH₂), 29.1 (Alk-CH₂), 29.0 (Alk-CH₂), 29.0 (Alk-CH₂), 24.0 (Alk-CH₂), 23.9 (Alk-CH₂), 23.9 (Alk-CH₂), 23.0 (Alk-CH₂), 23.0 (Alk-CH₂), 22.9 (Alk-CH₂), 14.0 (Alk-CH₃), 14.0 (Alk-CH₃), 14.0 (Alk-CH₃), 11.2 (Alk-CH₃), 11.1 (Alk-CH₃), 11.1 (Alk-CH₃);

HRMS (ES⁺): calculated for C₅₇H₇₈N₅O₈ 960.5850, found 960.5817;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 3020, 2400, 1521, 1425, 1265, 1217, 909.

4.5.1.12 Synthesis of 4.8



4.1 (0.0519 g, 0.13 mmol, 1 equiv.) and **4.6'** (0.188 g, 0.200 mmol, 1.5 equiv.) were dissolved in DCE (0.5 mL) and $\text{NaBH}(\text{OAc})_3$ (0.080 g, 0.38 mmol, 2.8 equiv.) was added with stirring. After 3 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed on a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 80% to 100% of EtOAc (with 1% NEt_3) in hexane (with 1% NEt_3) to yield a golden yellow oil (0.062 g, 36%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.40 - 8.46$ (m, 6H), 8.08 (dd, 1H, $J = 9.0$, $J = 3.0$), 7.97 (d, 1H, $J = 3.0$), 7.07 (d, 2H, $J = 6.0$), 7.02 (d, 2H, $J = 6.0$), 6.97 (d, 2H, $J = 6.0$), 6.86 (d, 1H, $J = 9.0$), 6.78 (d, 1H, $J = 3.0$), 6.71 (d, 1H, $J = 8.5$), 6.68 (d, 1H, $J = 9.0$), 6.62 (d, 1H, $J = 9.0$), 6.39 – 6.47 (m, 3H), 6.34 (dd, 1H, $J = 9.0$, 2.9), 6.26 (d, 1H, $J = 3.0$), 6.09 (s, 1H), 4.51 (s, 2H), 4.42 – 4.48 (m, 6H), 4.02 – 4.07 (m, 4H), 3.89 – 3.95 (m, 6H), 3.77 – 3.81 (m, 4H), 3.75 (d, 2H, $J = 6.0$), 1.59 – 1.79 (m, 4H), 1.16 – 1.54 (m, 36H), 0.77 - 0.97 (m, 24H);

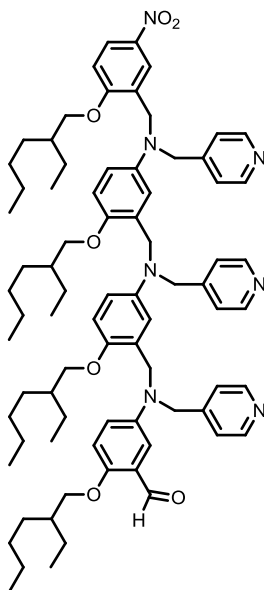
^{13}C NMR (63 MHz, CDCl_3): δ_{C} = 161.6, 149.8, 149.5, 149.3, 149.2, 149.1, 148.5, 142.5, 142.2, 141.7, 141.3, 127.9, 127.1, 126.8, 126.5, 124.6, 123.0, 122.0, 114.4, 113.6, 113.1, 112.8, 112.4, 112.3, 111.9, 111.8, 110.5, 99.4, 71.8, 71.5, 70.8, 70.7, 65.0, 54.9, 54.1, 53.8, 51.1, 50.9, 50.2, 39.5, 39.2, 30.7, 30.6, 30.5, 29.7, 29.1, 29.0, 24.0, 23.9, 23.1, 23.0, 22.9, 14.1, 14.0, 11.2, 11.1;

MS (ES+): m/z (%) = 274 (60), 649 (20), 1253 (10), 1297 (100) $[\text{M} + \text{H}^+]$;

HRMS (ES+): calculated for $\text{C}_{80}\text{H}_{110}\text{N}_7\text{O}_8$ 1296.8416, found 1296.8459;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2926, 2874, 2859, 1679, 1599, 1502, 1463, 1414, 1339, 1264, 1226, 1180, 1065, 1028, 969.

4.5.1.13 Synthesis of 4.8'



4.8 (0.25 g, 0.19 mmol, 1 equiv.) was dissolved in CHCl₃ (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO₃ and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl₃ (3 × 10 mL) before all organic fractions were washed with brine (1 × 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator to yield a yellow oil (0.22 g, 90%) requiring no further purification.

¹H NMR (400 MHz, CDCl₃): δ_H = 10.42 (s, 1H), 8.44 (d, 2H, *J* = 6.0), 8.40 (d, 4H, *J* = 6.0), 8.06 (dd, 1H, *J* = 9.0, *J* = 3.0), 7.97 (d, 1H, *J* = 2.5), 7.02 (d, 4H, *J* = 5.5), 6.94 (d, 2H, *J* = 6.0), 6.87 (d, 1H, *J* = 9.0), 6.71 – 6.76 (m, 2H), 6.68 (d, 1H, *J* = 9.0), 6.59 – 6.66 (m, 2H), 6.45 (dd, 1H, *J* = 9.0, *J* = 3.0), 6.39 (d, 1H, *J* = 2.5), 6.32 (dd, 1H, *J* = 9.0, *J* = 3.0), 6.17 (d, 1H, *J* = 3.0), 4.51 (s, 2H), 4.47 (s, 2H), 4.46 (s, 2H), 4.43 (s, 2H), 4.10 (s, 2H), 4.08 (s, 2H), 3.93 (d, 2H, *J* = 5.5), 3.86 (d, 2H, *J* = 5.5), 3.79 (d, 2H, *J* = 5.5), 3.71 (d, 2H, *J* = 5.5), 1.55 – 1.79 (m, 4H), 1.17 – 1.54 (m, 32H), 0.76 – 0.96 (m, 24H);

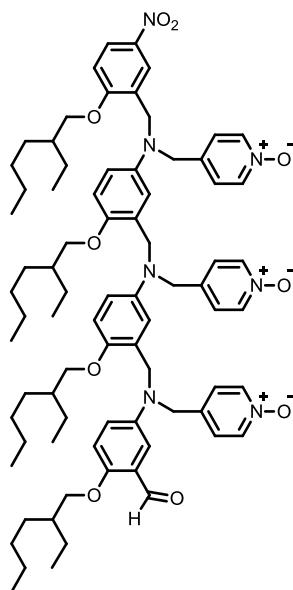
^{13}C NMR (63 MHz, CDCl_3): δ_{C} = 189.5, 161.4, 154.1, 149.7, 149.6, 149.3, 148.8, 148.7, 148.1, 148.0, 142.4, 141.9, 141.6, 141.1, 127.8, 126.8, 125.5, 124.9, 124.3, 122.8, 121.6, 121.5, 121.4, 120.3, 113.9, 112.8, 112.2, 112.1, 111.5, 110.3, 110.1, 71.3, 71.1, 70.5, 54.7, 54.1, 53.1, 50.8, 50.7, 49.9, 39.3, 39.2, 39.0, 30.5, 30.4, 30.3, 28.9, 28.8, 28.7, 23.8, 23.7, 22.8, 22.7, 13.9, 13.8, 13.7, 11.0, 10.9;

MS (ES+): m/z (%) = 1252 (50), 1253 (100) $[\text{M} + \text{H}^+]$, 1254 (60), 1255 (20);

HRMS (ES+): calculated for $\text{C}_{78}\text{H}_{106}\text{N}_7\text{O}_7$ 1252.8154, found 1252.8148;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2957, 2922, 2854, 1681, 1599, 1502, 1464, 1339, 1265, 1226, 1180.

4.5.1.14 Synthesis of 4.9'



4.7' (0.36 g, 0.37 mmol, 1 equiv.) and **4.2** (0.26 g, 0.65 mmol, 2.0 equiv.) were dissolved in CHCl₃ (2 mL) and NaBH(OAc)₃ (0.19 g, 0.91 mmol, 2.8 equiv.) was added with stirring. After 5 days the reaction was quenched with saturated aqueous NaHCO₃ solution, and extracted into CHCl₃ (4 × 10 mL). All the organic fractions were washed with water (1 × 10 mL), brine (1 × 10 mL) and dried (MgSO₄) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in a 2:8 mixture of acetonitrile and CHCl₃ to yield a mixture of the expected product and the benzaldehyde derivative. This mixture was dissolved in CHCl₃ (5 mL) and concentrated aqueous HCl (5 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO₃ and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl₃ (3 × 10 mL) before all organic fractions were washed with brine (1 × 10 mL) dried (MgSO₄) and the solvent removed using a rotary evaporator. The crude mixture was then purified via flash

chromatography on silica eluting with a gradient from 5% to 50% of MeOH in DCM to yield a viscous red oil (0.073 g, 15%).

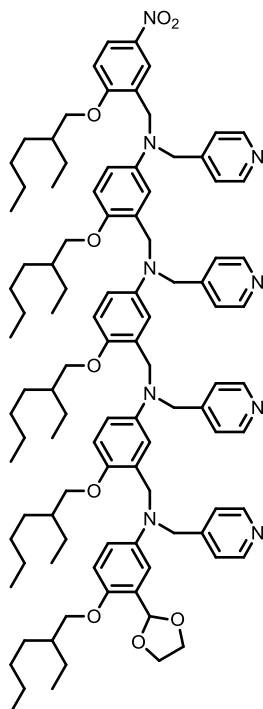
¹H NMR (400 MHz, CDCl₃): δ_{H} = 10.43 (s, 1H), 8.12 (dd, 1H, ³J = 9.0, ⁴J = 3.0), 8.09 (d, 2H, ³J = 7.0), 8.04 – 7.99 (m, 4H), 7.91 (d, 1H, ⁴J = 3.0), 7.10 – 7.04 (m, 3H), 7.00 (d, 2H, ³J = 7.0), 6.97 (d, 2H, ³J = 7.0), 6.93 (d, 1H, ³J = 9.0), 6.81 (d, 1H, ³J = 9.0), 6.73 – 6.66 (m, 3H), 6.47 (dd, 1H, ³J = 9.0, ⁴J = 3.0), 6.39 – 6.34 (m, 2H), 6.24 (d, 1H, ⁴J = 3.0), 4.52 – 4.47 (m, 4H), 4.43 (s, 2H), 4.41 (s, 2H), 4.24 (s, 2H), 4.16 (s, 2H), 3.99 (d, 2H, ³J = 5.5), 3.91 (d, 2H, ³J = 5.5), 3.85 (d, 2H, ³J = 5.5), 3.74 (d, 2H, ³J = 5.5), 1.82 – 1.60 (m, 4H), 1.55 – 1.22 (m, 32H), 0.98 – 0.83 (m, 24H);

¹³C NMR (126 MHz, CDCl₃): δ_{C} = 189.7, 161.7, 154.8, 149.9, 149.5, 142.1, 141.7, 141.3, 141.2, 139.2, 139.0, 138.9, 138.6, 138.4, 127.6, 126.8, 126.7, 125.9, 125.1, 124.7, 124.2, 124.2, 124.1, 123.0, 121.8, 120.8, 114.1, 113.5, 112.9, 112.5, 112.4, 112.3, 110.9, 110.6, 71.6, 71.4, 70.9, 70.8, 54.3, 53.5, 53.1, 50.9, 50.8, 50.6, 39.5, 39.4, 39.2, 30.6, 30.5, 29.7, 29.1, 29.0, 24.0, 23.9, 23.0, 14.1, 11.2;

HRMS (ES⁺): calculated for C₇₈H₁₀₅O₁₀N₇²³Na 1322.7815, found 1322.7782;

FT-IR (thin film): ν_{max} /cm⁻¹ 2959, 2925, 2872, 1682, 1502, 1483, 1340, 1264, 1245, 1229.

4.5.1.15 Synthesis of 4.10



4.8' (0.20 g, 0.16 mmol, 1 equiv.) and **4.1** (0.12 g, 0.31 mmol, 2 equiv.) were dissolved in DCE (1.2 mL) and $\text{NaBH}(\text{OAc})_3$ (0.093 g, 0.44 mmol, 2.8 equiv.) was added with stirring. After 3 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% of MeOH in CHCl_3 and then on reverse-phase silica eluting with a gradient from 70% to 100% of MeOH in acetonitrile) to yield a golden yellow oil (0.051 g, 20%).

^1H NMR (400 MHz, CDCl_3): δ_{H} = 8.44 (d, 2H, J = 6.0), 8.39 – 8.42 (m, 4H), 8.31 (d, 1H, J = 6.0), 8.09 (dd, 1H, J = 9.0, J = 3.0), 7.98 (d, 1H, J = 3.0), 7.03 (d, 2H, J = 6.0), 7.00 (d, 2H, J = 6.0), 6.88 – 6.92 (m, 4H), 6.86 (d, 1H, J = 9.0), 6.73 (d, 1H, J = 3.0), 6.66 – 6.72 (m, 2H), 6.64 (dd, 2H, J = 9.0, J = 3.0), 6.43 – 6.48 (m, 3H), 6.35 – 6.43 (m, 3H), 6.32 (d, 1H, J = 3.0), 6.26 (d, 1H, J = 3.0), 6.10 (s, 1H), 4.49 (s,

2H), 4.40 – 4.47 (m, 8H), 4.07 (s, 2H), 4.00 (s, 2H), 3.97 (s, 2H), 3.93 (d, 2H, $J = 6.0$), 3.87 – 3.91 (m, 4H), 3.72 – 3.81 (m, 8H), 1.58 – 1.82 (m, 5H), 1.18 – 1.55 (m, 40H), 0.79 – 0.95 (m, 30H);

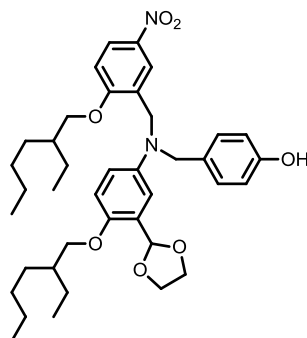
^{13}C NMR (63 MHz, CDCl_3): $\delta_{\text{C}} = 161.7, 150.1, 149.9, 149.8, 149.7, 149.2, 149.1, 149.0, 148.9, 148.3, 142.7, 142.6, 142.5, 142.0, 141.4, 128.1, 127.3, 126.9, 126.6, 124.7, 123.2, 122.0, 121.9, 114.3, 113.8, 113.3, 113.2, 113.0, 112.5, 112.4, 112.0, 111.9, 111.7, 110.6, 99.6, 71.9, 71.6, 70.9, 70.8, 55.1, 55.0, 54.3, 54.1, 53.6, 51.4, 51.2, 50.9, 50.2, 39.7, 39.6, 39.3, 30.7, 30.6, 29.2, 29.1, 24.1, 24.0, 23.2, 23.1, 14.2, 14.1, 11.3, 11.2;$

MS (ES+): m/z (%) = 1622 (20) $[\text{M} + \text{H}^+]$, 1644 (100) $[\text{M} + \text{Na}^+]$;

HRMS (ES+): calculated for $\text{C}_{101}\text{H}_{138}\text{N}_9\text{O}_9$ 1621.0618, found 1621.0594;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2927, 2859, 1599, 1503, 1464, 1414, 1340, 1264, 1225, 1065, 968.

4.5.1.16 Synthesis of 4.11



4.3 (1.30 g, 3.25 mmol, 1 equiv.) and **3.8** (1.82 g, 6.51 mmol, 2 equiv.) were dissolved in DCE (12 mL) and $\text{NaBH}(\text{OAc})_3$ (1.93 g, 9.11 mmol, 2.8 equiv.) was added at room temperature with stirring. After 18 h the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed on a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with 20% EtOAc in hexane to yield a waxy yellow solid (2.05 g, 95%).

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.17 - 8.10$ (m, 2H, Ar-H), 7.12 (d, 2H, $^3J = 8.0$, Ar-H) 7.01 (d, 1H, $^4J = 3.0$, Ar-H), 6.93 (d, 1H, $^3J = 10.0$, Ar-H), 6.77 (d, 1H, $^3J = 9.0$, Ar-H), 6.72 (d, 2H, $^3J = 8.0$, Ar-H), 6.55 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 6.20 (s, 1H, OCH₂O), 4.57 (s, 2H, CH₂N), 4.54 (s, 2H, CH₂N), 4.06 – 3.94 (m, 6H, OCH₂CH₂O and CH₂O), 3.81 (d, 2H, $^3J = 6.0$, CH₂O), 1.86 – 1.68 (m, 2H, Alk-CH), 1.58 – 1.25 (m, 16H, Alk-CH₂) 1.02 – 0.86 (m, 12H, Alk-CH₃);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.9$ (Ar-C), 154.9 (Ar-C), 150.0 (Ar-C), 143.0 (Ar-C), 141.3 (Ar-C), 130.2 (Ar-C), 128.8 (Ar-C), 128.4 (Ar-C), 126.6 (Ar-C), 124.4 (Ar-C), 123.5 (Ar-C), 115.4 (Ar-C), 115.2 (Ar-C), 113.6 (Ar-C), 112.5 (Ar-C), 110.4 (Ar-C), 99.7 (OCHO), 71.7 (CH₂O), 71.4 (CH₂O), 65.0 (OCH₂CH₂O), 55.2 (CH₂N), 50.0 (CH₂N), 39.5 (Alk-CH), 39.2 (Alk-CH), 30.5 (Alk-CH₂), 29.1 (Alk-CH₂), 29.0

(Alk-CH₂), 24.0 (Alk-CH₂), 23.9 (Alk-CH₂), 23.1 (Alk-CH₂), 23.0 (Alk-CH₂), 14.1

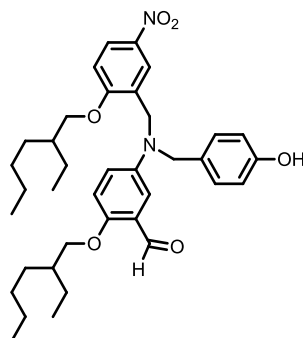
(Alk-CH₃), 14.1 (Alk-CH₃), 11.1. (Alk-CH₃);

HRMS (ES+): calculated for C₃₉H₅₅N₂O₇ 663.4009, found 663.4008;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 3420.63, 2956, 2929, 2869, 1677, 1609, 1592, 1510,

1494, 1336, 1267, 1177

4.5.1.17 Synthesis of 4.11'



4.11 (1.23 g, 1.85 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous HCl (10 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator to yield a yellow solid which was recrystallized from DCM and hexane (1.15 g, 95%) Mp 105 – 107 °C.

^1H NMR (500 MHz, CDCl_3): δ_{H} = 10.42 (s, 1H, Ar-CHO), 8.11 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 7.98 (d, 1H, $^4J = 3.0$, Ar-H), 7.17 (d, 1H, $^4J = 3.0$, Ar-H), 7.09 *d, 2H, $^3J = 8.5$, Ar-H), 6.93 – 6.88 (m, 2H, Ar-H), 6.84 (d, 1H, $^3J = 9.0$, Ar-H), 6.77 (d, 2H, $^3J = 8.5$, Ar-H), 4.55 (s, 2H, CH_2N), 4.54 (s, 2H, CH_2N), 3.98 (d, 2H, $^3J = 5.5$, CH_2O), 3.87 (dd, 2H, $^3J = 5.5$, $^4J = 1.5$, CH_2O), 1.83 – 1.68 (m, 2H, Alk-CH), 1.52 – 1.23 (m, 16H, Alk- CH_2), 0.96 – 0.82 (m, 12H, Alk- CH_3);

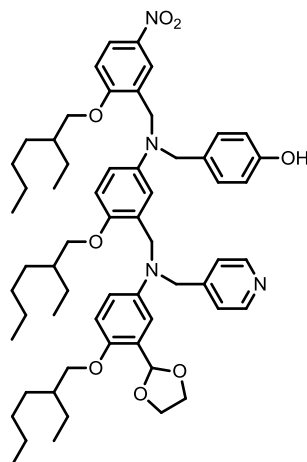
^{13}C NMR (126 MHz, CDCl_3): 190.4 (Ar-CHO), 161.9 (Ar-C), 155.0 (Ar-C), 154.7 (Ar-C), 142.7 (Ar-C), 141.3 (Ar-C), 129.7 (Ar-C), 128.2 (Ar-C), 127.8 (Ar-C), 125.1 (Ar-C), 124.5 (Ar-C), 123.0 (Ar-C), 121.4 (Ar-C), 115.5 (Ar-C), 114.2 (Ar-C), 110.9 (Ar-C), 110.5 (Ar-C), 71.5 (CH_2O), 71.2 (CH_2O), 54.6 (CH_2N), 49.6 (CH_2N), 39.5 (Alk-CH), 39.1 (Alk-CH), 30.6 (Alk- CH_2), 30.5 (Alk- CH_2), 29.0 (Alk- CH_2), 28.9 (Alk-

CH₂), 23.9 (Alk-CH₂), 23.9 (Alk-CH₂), 22.9 (Alk-CH₂), 22.9 (Alk-CH₂), 14.0 (Alk-CH₃), 14.0 (Alk-CH₃), 11.1 (Alk-CH₃), 11.1 (Alk-CH₃);

HRMS (ES+): calculated for C₃₇H₅₁N₂O₆ 619.3747, found 619.3738;

FT-IR (thin film): ν_{max} /cm⁻¹ 3433, 2956, 2925, 2861, 1677, 1610, 1592, 1510, 1493, 1465, 1439, 1338, 1269, 1245, 1207, 1177.

4.5.1.18 Synthesis of 4.12



4.1 (0.100 g, 0.260 mmol, 1 equiv.) and **4.11'** (0.242 g, 0.391 mmol, 1.5 equiv.) were dissolved in DCE (1 mL) and $\text{NaBH}(\text{OAc})_3$ (0.155 g, 0.731 mmol, 2.8 equiv.) was added at room temperature with stirring. After 18 h the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed on a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with 40% to 70% of EtOAc in hexane to yield an orange oil (0.134 g, 52%).

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} =$ 8.43 (d, 2H, $^3J = 5.5$, Ar-H), 8.06 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 7.98 (d, 1H, $^4J = 3.0$, Ar-H), 7.16 (d, 2H, $^3J = 5.5$, Ar-H), 6.96 (d, 2H, $^3J = 8.5$, Ar-H), 6.84 (d, 1H, $^3J = 9.0$, Ar-H), 6.71 (d, 2H, $^3J = 8.5$, Ar-H), 6.69 (d, 1H, $^3J = 9.0$, Ar-H), 6.67 (d, 1H, $^4J = 3.0$, Ar-H), 6.62 (d, 1H, $^3J = 9.0$, Ar-H), 6.47 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 6.43 (d, 1H, $^4J = 3.0$, Ar-H), 6.38 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, Ar-H), 6.06 (s, 1H, OCHO), 4.52 (s, 1H, CH_2N), 4.44 (m, 4H, CH_2N), 4.23 (s, 1H, CH_2N), 3.96 – 3.90 (m, 6H, CH_2O and $\text{OCH}_2\text{CH}_2\text{O}$), 3.80 (d, 2H, $^3J = 6.0$, CH_2O), 3.74 (d, 2H, $^3J = 6.0$, CH_2O), 1.80 – 1.69 (m, 2H, Alk-CH),

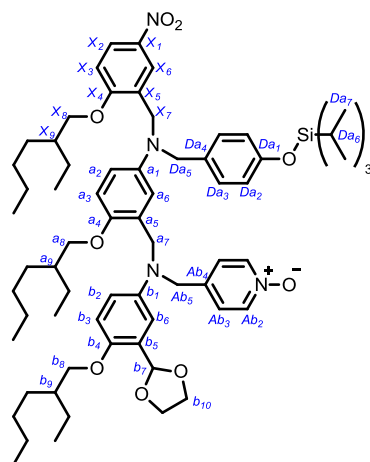
1.67 – 1.60 (m, 1H, Alk-CH), 1.56 – 1.19 (m, 24H, Alk-CH₂), 0.97 – 0.80 (m, 18H, Alk-CH₃);

¹³C NMR (126 MHz, CDCl₃): δ_c = 161.6 (Ar-C), 155.5 (Ar-C), 151.1 (Ar-C), 149.7 (Ar-C), 149.0 (Ar-C), 148.2 (Ar-C), 142.4 (Ar-C), 141.3 (Ar-C), 129.8 (Ar-C), 128.5 (Ar-C), 127.7 (Ar-C), 126.4 (Ar-C), 126.1 (Ar-C), 124.3 (Ar-C), 123.0 (Ar-C), 122.3 (Ar-C), 115.8 (Ar-C), 114.3 (Ar-C), 113.4 (Ar-C), 113.3 (Ar-C), 112.3 (Ar-C), 111.6 (Ar-C), 111.5 (Ar-C), 110.4 (Ar-C), 99.6 (OCHO), 71.6 (CH₂O), 71.3 (CH₂O), 70.6 (CH₂O), 65.0 (OCH₂CH₂O), 55.7 (CH₂N), 53.9 (CH₂N), 50.8 (CH₂N), 50.4 (CH₂N), 39.5 (Alk-CH), 39.4 (Alk-CH), 39.2 (Alk-CH), 30.6 (Alk-CH₂), 30.5 (Alk-CH₂), 29.0 (Alk-CH₂), 29.0 (Alk-CH₂), 24.0 (Alk-CH₂), 23.9 (Alk-CH₂), 23.9 (Alk-CH₂), 23.0 (Alk-CH₂), 23.0 (Alk-CH₂), 22.9 (Alk-CH₂), 14.0 (Alk-CH₃), 14.0 (Alk-CH₃), 11.1 (Alk-CH₃);

HRMS (ES+): calculated for C₆₀H₈₃N₄O₈ 987.6205, found 987.6201;

FT-IR (thin film): ν_{\max} /cm⁻¹ 3671, 2988, 2901, 1502, 1405, 1336, 1225.

4.5.1.19 Synthesis of 4.13



3.13' (2.02 g, 2.60 mmol, 1 equiv.) and **4.2** (2.09 g, 5.21 mmol, 2 equiv.) were dissolved in CHCl_3 (9 mL) and $\text{NaBH}(\text{OAc})_3$ (1.55 g, 7.29 mmol, 2.8 equiv.) was added with stirring. After 2 days more $\text{NaBH}(\text{OAc})_3$ was added (1.0 g) and then after 2 more days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 5% to 10% EtOAc in hexane and then 10% to 100% MeOH in a 1:1 mixture of acetonitrile and CHCl_3 to yield a pale yellow oil (2.24 g, 74%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.06 - 7.99$ (m, 3H, H-Ab2 and H-X2), 7.92 (d, 1H, $^4J = 3.0$, H-X6), 7.08 – 6.96 (m, 4H, H-Ab3 and H-Da3), 6.82 – 6.76 (m, 3H, H-Da2 and H-X3), 6.70 – 6.65 (m, 2H, H-b3 and H-b6), 6.60 (d, 1H, $^3J = 9.0$, H-a3), 6.49 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-b2), 6.36 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-a2), 6.34 (d, 1H, $^4J = 3.0$, H-a6), 6.02 (s, 1H, H-b7), 4.46 – 4.37 (m, 6H, H-X7, H-a7 and H-Da5), 4.05 (s, 2H, H-Ab5), 3.99 – 3.87 (m, 6H, H-b10 and H-X8), 3.77 (d, 2H, $^3J = 5.5$, CH_2O), 3.73 (d, 2H, $^3J = 5.5$, CH_2O), 1.77 – 1.57 (m, 3H, Alk-CH), 1.55 – 1.15 (m,

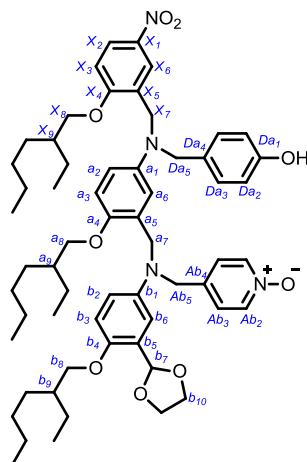
27H, Alk-CH₂ and H-Da6), 1.06 (d, 18H, ³J = 7.0, H-Da7), 0.94 – 0.79 (m, 18H, Alk-CH₃);

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.7 (C-X4), 155.1 (Ar-C), 150.2 (Ar-C), 149.4 (Ar-C), 142.6 (Ar-C), 142.5 (Ar-C), 141.5 (Ar-C), 140.0 (Ar-C), 139.1 (C-Ab2), 131.3 (Ar-C), 128.6 (Ar-C), 128.2 (Ar-C), 126.8 (Ar-C), 126.3 (Ar-C), 124.5 (Ar-C), 124.4 (Ar-C), 123.2 (Ar-C), 120.2 (Ar-C), 114.7 (Ar-C), 113.7 (Ar-C), 113.4 (Ar-C), 112.6 (Ar-C), 112.0 (Ar-C), 111.7 (Ar-C), 110.6 (Ar-C), 99.5 (C-b7), 71.9 (CH₂O), 71.6 (CH₂O), 70.9 (CH₂O), 65.3 (OCH₂CH₂O), 55.7 (CH₂N), 53.0 (C-Ab5), 50.6 (CH₂N), 50.4 (CH₂N), 39.7 (Alk-CH), 39.7 (Alk-CH), 39.4 (Alk-CH₂), 30.8 (Alk-CH₂), 30.7 (Alk-CH₂), 30.7 (Alk-CH₂), 29.3 (Alk-CH₂), 29.3 (Alk-CH₂), 29.3 (Alk-CH₂), 24.2 (Alk-CH₂), 24.1 (Alk-CH₂), 24.1 (Alk-CH₂), 23.3 (Alk-CH₂), 23.2 (Alk-CH₂), 23.2 (Alk-CH₂), 18.1 (TiPS-CH₃), 14.3 (Alk-CH₃), 14.3 (2 x Alk-CH₃), 12.8 (TiPS-CH), 11.4 (3 x Alk-CH₃);

HRMS (ES⁺): calculated for C₆₉H₁₀₃N₄O₉Si 1159.7494, found 1159.7439;

FT-IR (thin film): ν_{max}/cm⁻¹ 2958, 2926, 2863, 1681, 1609, 1592, 1504, 1463, 1338, 1262, 1227, 1166.

4.5.1.20 Synthesis of 4.14



4.13 (0.082 g, 0.07 mmol, 1 equiv.) was dissolved in THF (2 mL) at 0 °C and TBAF (140 μ L, 0.14 mmol, 2 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 5% of methanol in dichloromethane to yield a yellow oil (0.048 g, 66%).

^1H NMR (400 MHz, CDCl_3): δ_{H} = 8.08 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$, H-X2), 8.04 (d, 2H, $^3J = 6.5$, H-Ab2), 7.98 (d, 1H, $^3J = 9.0$, $^4J = 3.0$, H-X6), 7.05 (d, 2H, $^3J = 6.5$, H-Ab3), 6.92 (d, 2H, $^3J = 8.5$, H-Aa2), 6.86 (d, 1H, $^3J = 9.0$, H-X3), 6.77 – 6.72 (m, 3H, H-Aa2 and H-b6), 6.68 (d, 1H, $^3J = 9.0$, H-b3), 6.66 (d, 1H, $^3J = 9.0$, H-a3), 6.46 – 6.38 (m, 3H, H-b2, H-a2 and H-a6), 6.05 (s, 1H, H-b7), 4.54 (s, 2H, H-X7), 4.40 (s, 4H, H-Aa5 and H-a7), 4.04 (s, 2H, H-Ab5), 4.03 – 3.92 (m, 6H, H-X8 and H-b9) 3.80 (d, 2H, $^3J = 5.5$, H-b8), 3.72 (d, 2H, $^3J = 5.5$, H-a8), 1.81 – 1.57 (m, 3H, H-X9, H-a9 and H-b9), 1.55 – 1.18 (m, 24H, Alk- CH_2), 0.94 – 0.79 (m, 18H, Alk- CH_3);

^{13}C NMR (101 MHz, CDCl_3): δ_{C} = δ 161.9 (1C, s, C-X4), 156.0 (1C, s, C-Aa1), 150.5 (1C, s, C-b4), 149.2 (1C, s, C-a4), 142.5 (2C, s, C-Aa4 and C-Ab4), 142.4 (1C, s,

C-b1), 141.6 (1C, s, Ar-C^{quat}), 139.0 (2C, s, C-Ab2), 129.7 (1C, s, Ar-C^{quat}), 128.8 (1C, s, C-X5), 127.9 (2C, s, C-Aa3), 126.6 (1C, s, C-b5), 126.3 (1C, s, C-a5), 124.6 (2C, s, C-Ab3), 124.5 (1C, s, C-X2), 123.3 (1C, s, C-X6), 116.2 (2C, s, C-Aa2), 115.4 (1C, s, Ar-CH), 113.9 (1C, s, Ar-CH), 113.7 (1C, s, C-b3), 112.6 (1C, s, Ar-CH), 112.5 (1C, s, Ar-CH), 111.9 (1C, s, Ar-CH), 110.6 (1C, s, C-X3), 99.7 (1C, s, C-b7), 71.9 (1C, s, C-b8), 71.6 (1C, s, C-X8), 70.8 (1C, s, C-a8), 65.3 (2C, s, C-b10), 55.9 (1C, s, C-Aa5), 53.5 (1C, s, C-Ab5), 51.4 (1C, s, C-X7), 51.1 (1C, s, C-a7), 39.7 (2C, s, Ar-CH), 39.4 (1C, s, Ar-CH), 30.8 (1C, s, Ar-CH₂), 30.7 (2C, s, Ar-CH₂), 29.3 (2C, s, Ar-CH₂), 29.2 (1C, s, Ar-CH₂), 24.3 (1C, s, Ar-CH₂), 24.2 (1C, s, Ar-CH₂), 24.1 (1C, s, Ar-CH₂), 23.4 (1C, s, Ar-CH₂), 23.2 (1C, s, Ar-CH₂), 23.1 (1C, s, Ar-CH₂), 14.3, (3C, s, Ar-CH₃), 11.3 (3C, s, Ar-CH₂);

HRMS (ES+): calculated for C₆₀H₈₂N₄O₉²³Na 1025.5974, found 1025.5943;

FT-IR (thin film): ν_{max} /cm⁻¹ 2956, 2925, 2857, 1681, 1613, 1592, 1500, 1463, 1338, 1263, 1221, 1167.

4.5.2 Binding studies

All association constants were measured by ^1H NMR and additionally, if applicable, ^{31}P NMR titration. One species, labelled the host, is dissolved in toluene- d_8 to a known concentration. A second species, labelled the guest, is dissolved in the host solution and made to a known concentration such that the concentration of the host is the same in both solutions. A known volume of host is added to an NMR tube and the spectra measured. Known volumes of guest in host solution are added to the tube, and the spectra measured after each addition. The chemical shifts of the host spectra are recorded as a function of guest concentration and analysed using standard calculations in a fitting-program written in Microsoft Excel®. Errors were calculated as two times the standard deviation from the average value from repetitions.

4.6 Acknowledgements

I would like to thank Mike and Giulia who measured the association constant for **3.25•4.15** and Mike for allowing me to include the crystal structure of **2.9**.

4.7 References

- 1 M. Mascini, I. Palchetti, and S. Tombelli, *Angew. Chemie Int. Ed.*, 2012, **51**, 1316–1332.
- 2 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.
- 3 K.-A. Yang, M. Barbu, M. Halim, P. Pallavi, B. Kim, D. M. Kolpashchikov, S. Pecic, S. Taylor, T. S. Worgall, and M. N. Stojanovic, *Nat. Chem.*, 2014, **6**, 1003–1008.
- 4 P. Colas, B. Cohen, T. Jessen, I. Grishina, J. McCoy, and R. Brent, *Nature*, 1996, **380**, 548–550.
- 5 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683–720.
- 6 A. Varki, *Glycobiology*, 1993, **3**, 97–130.
- 7 W. Saenger, in *Principles of Nucleic Acid Structure*, Springer, New York, 1984, pp. 253–282.
- 8 H. Korhonen, T. Koivusalo, S. Toivola, and S. Mikkola, *Org. Biomol. Chem.*, 2013, **11**, 8324.
- 9 J. Marmur and P. Doty, *J. Mol. Biol.*, 1962, **5**, 109–118.
- 10 H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, and E. T. Kool, *Science*, 2003, **302**, 868–71.
- 11 R. N. Veedu and J. Wengel, *Chem. Biodivers.*, 2010, **7**, 536–42.
- 12 H. J. Hogben, J. K. Sprafke, M. Hoffmann, M. Pawlicki, and H. L. Anderson, *J. Am. Chem. Soc.*, 2011, **133**, 20962–20969.
- 13 C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310–24.
- 14 E. Chekmeneva, C. A. Hunter, M. C. Misuraca, and S. M. Turega, *Org. Biomol. Chem.*, 2012, **10**, 6022–31.
- 15 M. C. Misuraca, T. Grecu, Z. Freixa, V. Garavini, C. A. Hunter, P. W. N. M. van Leeuwen, M. D. Segarra-Maset, and S. M. Turega, *J. Org. Chem.*, 2011, **76**, 2723–32.
- 16 M. Jinks and C Hunter, Masters Thesis, University of Cambridge, 2014.

Chapter 5

Mixed sequence trimer duplexes

5.1 Introduction

In Nature, the ability of nucleic acids to form sequence selective duplexes is essential for the expression and replication of genetic information.¹ The expression of information is determined by the sequence of the nucleobases, which makes nucleic acids programmable. This feature has allowed nucleic acids to be the basis for a diverse range of applications,²⁻⁵ which can be reached through design,⁶ or controlled evolution.⁷ Efforts towards a synthetic alternative to the nucleic acids, seldom stray too far from Nature's design,⁸⁻¹² and require stringent structural complementarity. Although polymers bearing recognition groups have been made,¹³⁻¹⁵ the lack of control over sequence has meant they cannot challenge nucleic acids in many applications.¹⁶⁻¹⁹ Therefore stepwise synthesis is still the only reliable route to oligomers with well-defined sequences.^{20,21}

In Chapters 3 and 4, we have shown that it is possible to generate synthetic oligomers which form duplexes through H-bonding. The ability to form duplexes is maintained after changing the type of pairwise H-bonding interactions. Although synthetic H-bonding duplexes have been explored by others, the sequence of donor and acceptor groups is rarely programmable, due to being intrinsically dictated by the coupling chemistry.²²⁻²⁵ There is one example of a completely synthetic duplex which forms sequence selective duplexes via interaction of salt bridges off the side of the oligomer.²⁶ In this Chapter we investigate the binding properties of duplexes formed from eight 3-mers bearing all possible sequences of an H-bond donor or acceptor recognition group (Figure 5.1). By measuring and comparing the association constants of each pairwise combination of duplex, we directly assess the ability of this system to discriminate the sequence of H-bond donor and acceptor groups.

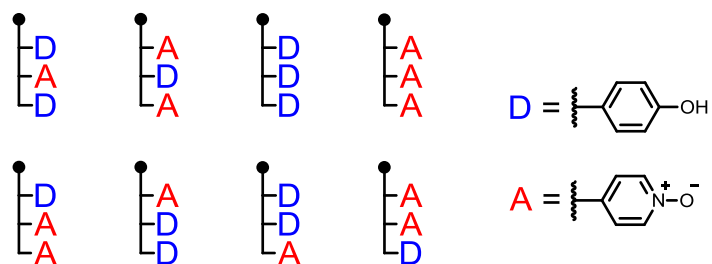


Figure 5.1. The eight possible sequences of 3-mers bearing three H-bond donor groups (D, phenol) or acceptor groups (A, pyridine *N*-oxide), for an oligomer, which is not symmetric with respect to the two ends (represented by the circle).

5.2 Approach

For constructing the mixed sequence oligomers we have used the same modular system described in Chapters 3 and 4. Oligomers bearing a sequence of H-bond donor and acceptor groups can be synthesised through a series of reductive amination and deprotection steps (Figure 5.2(a)). Using a divergent approach, each oligomer can be constructed by splitting supplies of the intermediates (Figure 5.2(b)). Three different H-bond acceptor groups have been investigated so far by measuring the association constants for duplex formation between oligomers bearing just acceptor groups and oligomers with just phenol donor groups (see Chapters 3 and 4). Each system can be compared through two key parameters (Table 5.1). The first is the dimensionless product of $K_{\text{ref}} \overline{EM}$, the association constant for each stepwise intramolecular H-bond. K_{ref} is the association constant for the single point H-bond interaction between the H-bond donor and acceptor groups, and \overline{EM} is the average effective molarity per H-bond in duplex formation, which quantifies the additional stability of forming an intramolecular H-bond in a duplex relative to forming an intermolecular H-bond. The second parameter is the equilibrium constant for forming an intramolecular H-bond between adjacent recognition sites on the same oligomer, $K_{\text{ref}} EM_{\text{fold}}$. If intramolecular H-bond formation is facile between adjacent H-bonding sites on the same oligomer, duplex formation will be disrupted at all concentrations by internal H-bonds. Therefore a system that satisfies $K_{\text{ref}} \overline{EM} \gg 1$ and $K_{\text{ref}} EM_{\text{fold}} \ll 1$ is going to be a good candidate for sequence selective duplex formation, since both of these criteria lead to high fidelity duplexes. Table 5.1 shows values of $K_{\text{ref}} \overline{EM}$ and $K_{\text{ref}} EM_{\text{fold}}$ for the phosphine oxide, pyridine and pyridine *N*-oxide systems measured so far. The pyridine *N*-oxide and system has

the largest $K_{\text{ref}}\overline{EM}$ and the lowest propensity to fold, which is why we have chosen this system to study the binding properties of mixed sequences.

Table 5.1. $K_{\text{ref}}\overline{EM}$ and $K_{\text{ref}}EM_{\text{fold}}$ values for the duplex systems presented in Chapters 3 and 4.

| H-bond acceptor | $K_{\text{ref}}\overline{EM}^a$ | $K_{\text{ref}}EM_{\text{fold}}$ |
|--------------------------|---------------------------------|----------------------------------|
| phosphine oxide | 5 | 7 |
| pyridine | 3 | 1 |
| pyridine <i>N</i> -oxide | 10 | ≈ 0 |

^a average of all duplex lengths measured

Comparing the association constant of each pairwise combination of 3-mer oligomers means we can investigate a number of interesting properties of the system. For instance,

- what is the impact of mismatches in sequence between two oligomer strands, for example DDD•AAA versus DDD•ADA?
- what is the difference in stability between two duplexes which are forced to bind in a parallel and antiparallel sense, for example AAD•DDA versus AAD•DDA?
- what is the effect of interactions between the two recognition groups at each terminus of the same oligomer (1,3-folding)?

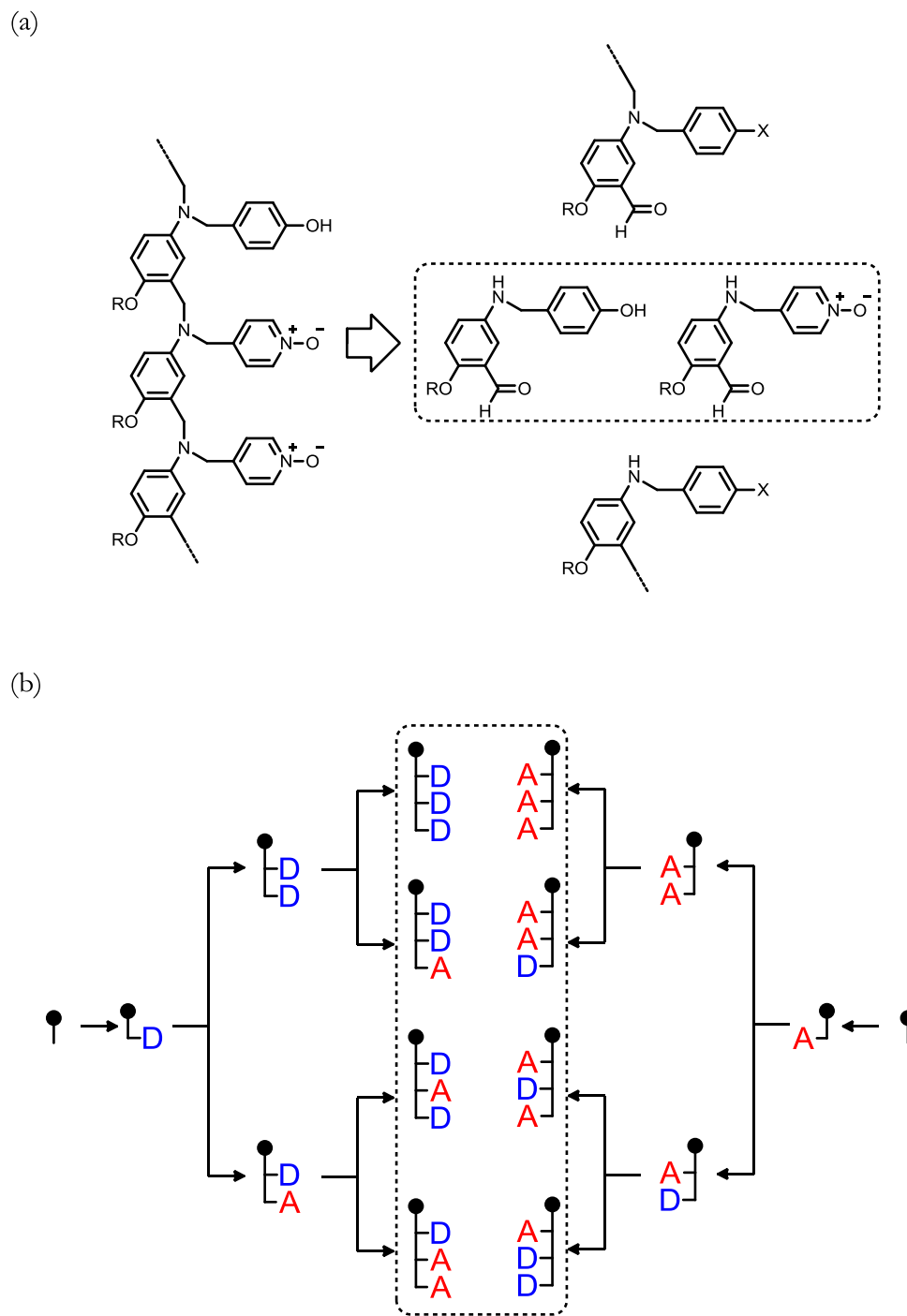


Figure 5.2. (a) Retrosynthesis of an oligomer showing how sequences are incorporated by choosing one of two monomers. (b) 3-mers synthesised by a divergent approach (circle is capping group).

5.3 Results and discussion

5.3.1 Synthesis

The syntheses of monomers **3.11** and **4.2** are described in Chapters 3 and 4 respectively (Figure 5.3). Scheme 5.1 summarises the synthesis of all oligomers in this Chapter, which have been made through sequential reductive amination and acetal deprotection steps. The phenol groups were all protected as the TIPS silyl ethers for oligomer synthesis and then the precursor 3-mers were deprotected using TBAF. Oligomers corresponding to sequences with DAD, DAA AAD and AAA were isolated as the corresponding benzaldehyde in Scheme 5.1 and are designated with a prime (**5.5'**, **5.7'**, **5.11'** and **4.9'** respectively).

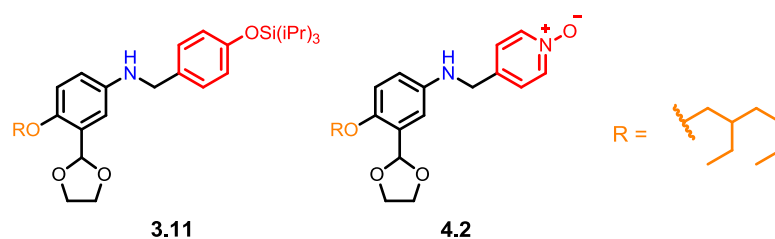
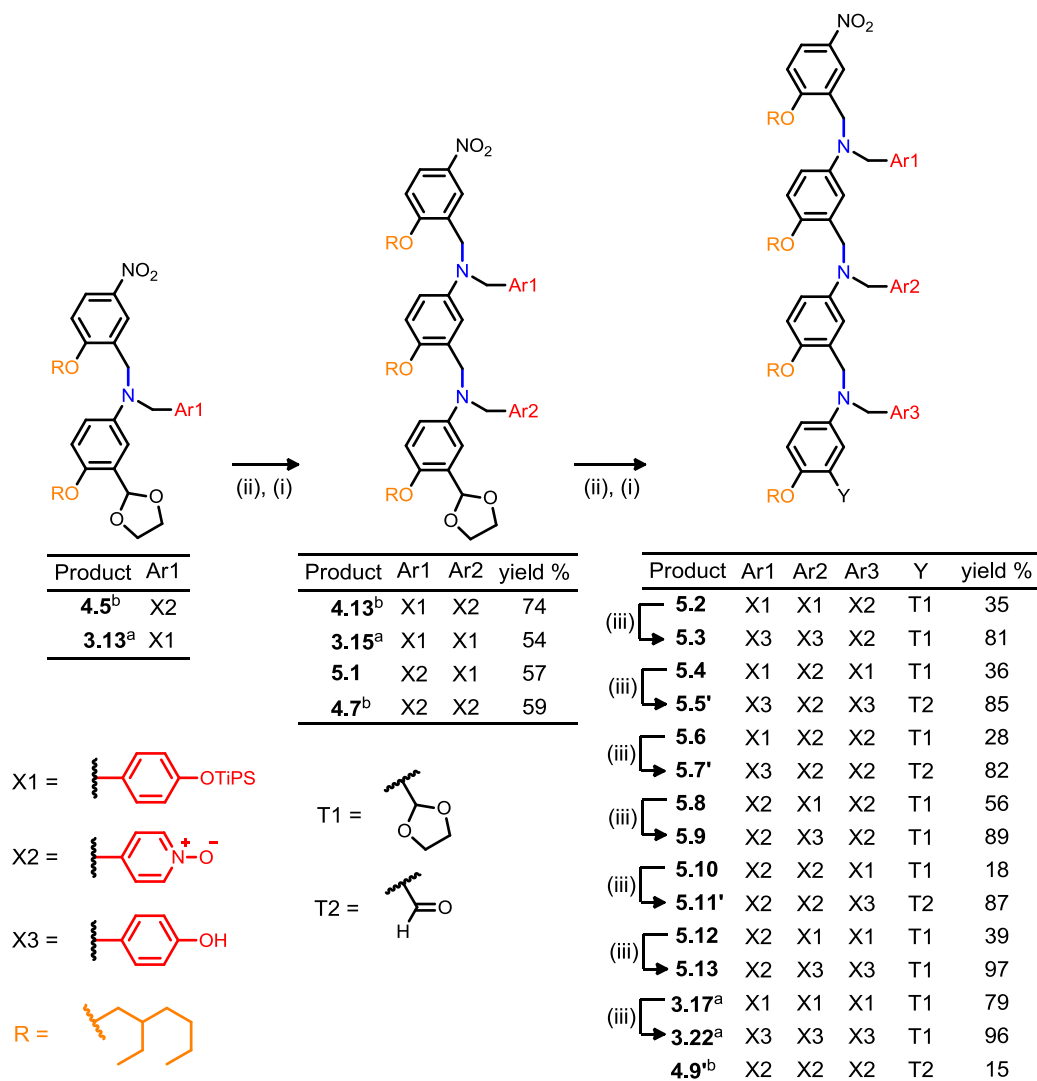


Figure 5.3. Monomer units used in the synthesis of oligomers shown in Scheme 5.1.



Scheme 5.1.^a Synthesis reported in Chapter 3; ^b synthesis reported in Chapter 4. The prime for compounds **5.5'**, **5.7'**, **5.11'** and **4.9'** designates that they correspond to the benzaldehyde in the Y position.

Reagents and conditions: (i) Monomer (**3.11** and **4.2**), NaBH(OAc)₃; (ii) aqueous HCl; (iii) TBAF.

5.3.2 Binding

The association constants for each pairwise combination of 3-mers were measured by ^1H NMR titrations in toluene- d_8 and the data fit well to a 1:1 binding isotherm (Table 5.2). Each oligomer is allocated a unique three letter code (shown in Table 5.2) which is used to describe the benzaldehyde to nitrobenzene sequence of donor (D) and acceptor (A) groups.

Table 5.2. Association constants for 1:1 complexes measured by ^1H NMR titrations in toluene- d_8 at 298 K.^a

| | DDD (3.22) | DDA (5.3) | DAD (5.5') | DAA (5.7') | ADD (5.13) | ADA (5.9) | AAD (5.11') | AAA (4.9') |
|-----|---------------|--------------|-----------------------------------|----------------------------|-----------------------------------|----------------------------|----------------------------|-----------------------------------|
| DDD | ^c | ^c | 1.1×10^3 _b | 1.1×10^4 (12%) | 2.2×10^2 _b | 1.5×10^4 (23%) | 1.5×10^4 (39%) | 1.0×10^5 (53%) |
| DDA | | ^c | 4.2×10^3 _b | 4.2×10^3 (16%) | 1.2×10^3 _b | 4.6×10^3 (4%) | 2.2×10^3 (5%) | 1.7×10^3 (52%) |
| DAD | | | ^c | 8.9×10^3 (6%) | 2.2×10^3 (10%) | 2.5×10^4 (40%) | 6.3×10^3 (35%) | 8.4×10^3 (6%) |
| DAA | | | | ^c | 1.0×10^3 (7%) | 2.6×10^3 (9%) | 8.0×10^2 (51%) | 4.5×10^2 _b |
| ADD | | | | | ^c | 2.3×10^3 (4%) | 8.0×10^3 (6%) | 1.6×10^3 (10%) |
| ADA | | | | | | ^c | 1.0×10^3 (48%) | ^c |
| AAD | | | | | | | ^c | ^c |
| AAA | | | | | | | | ^c |

^a Average values from repetitions and percentage error of the mean at the 95% confidence limit in brackets; ^b value based on one experiment; ^c experiment not conducted yet.

A number of interesting observations can be made from the data in Table 5.2 which will be discussed in more detail below.

5.3.2.1 Sequence matching/mismatching

Two oligomers are defined as having matched sequences if H-bonds in the duplex match the order in which the recognition groups are connected to the backbone. For example, ADA and AAD have two pyridine *N*-oxide groups and one phenol group, but only ADA has a sequence that is complementary to DAD (Figure 5.4). Three H-bonds could form in the mismatched duplex AAD•DAD but this requires recognition units to interact with adjacent groups within the duplex.

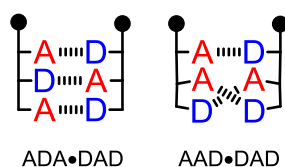


Figure 5.4. Comparison of two duplexes where H-bonds form between recognition units in the same order as they are placed on the backbone (matched ADA•DAD) and where they form in a different order (mismatched AAD•DAD)

The data in Table 5.2 shows that duplexes generally have higher association constants when they are sequence matched. For instance, the average association constant for duplexes that are sequence matched and can form three H-bonds ($2.4 \times 10^4 \text{ M}^{-1}$) is an order of magnitude higher than the average association constant for sequence mismatched duplexes that could form three H-bonds ($5.5 \times 10^3 \text{ M}^{-1}$). This suggests the association constant of a duplex is higher when the recognition groups on both oligomers are arranged in a complementary order.

5.3.2.2 Duplex symmetry: parallel versus antiparallel

It is important to note that the oligomers studied here are not symmetric along the axis of the backbone. For instance both DDA and ADD have recognition side chains in the order donor, donor, acceptor, but in DDA the pyridine *N*-oxide is at the benzaldehyde terminus of the oligomer whereas in ADD the pyridine *N*-oxide

group is at the nitrobenzene terminus. Therefore both DDA and ADD have sequences which are complementary to both AAD and DAA (Figure 5.5). Two of the four possible duplexes form in a parallel sense, where the backbone of each oligomer is aligned in the same direction; the other two duplexes are antiparallel and have backbones which are aligned in the opposite sense. Comparing the stability of the parallel duplexes (DAA•ADD and AAD•DDA) with the antiparallel duplexes (DAA•DDA, AAD•ADD) suggests that there is a preference for forming duplexes in the antiparallel direction. The average association constant for antiparallel duplexes ($6.1 \times 10^3 \text{ M}^{-1}$) is four times larger than for parallel duplexes ($1.6 \times 10^3 \text{ M}^{-1}$).

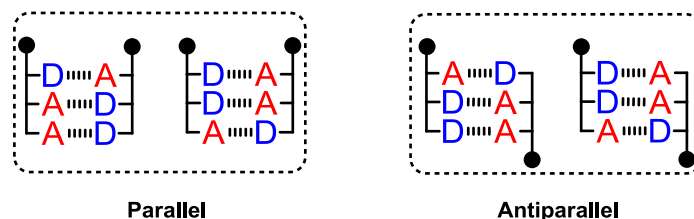


Figure 5.5. Representation of four duplexes which can make three H-bonds but are fixed in either parallel or antiparallel orientation by the sequence of recognition units. The black circle represents the nitrobenzene terminus of the oligomer.

5.3.2.3 1,3-folding

The data in Table 5.2 show that complementary duplexes involving oligomers which have a non-symmetric sequence (DDA, AAD, DAA, or ADD) have inherently lower association constants compared to all of the other duplexes. The average association constant for complementary duplexes without DDA, AAD, DAA, or ADD is $4.5 \times 10^4 \text{ M}^{-1}$ whereas the average association constant for duplexes that contain one or more of DDA, AAD, DAA, or ADD is an order of magnitude lower at $3.9 \times 10^3 \text{ M}^{-1}$. The average association constant for duplexes where both of the oligomers have a non-symmetric sequence is in the same order of magnitude ($2.9 \times 10^3 \text{ M}^{-1}$) as duplexes that have just one non-symmetric sequence. This suggests

that a non-symmetric sequences has a destabilising effect upon duplex formation. A possible explanation for this is that an intramolecular H-bond is formed between the recognition groups in the 1st and 3rd position of the oligomer (Figure 5.6). This form of 1,3-folding would compete with duplex formation and reduce the observed association constant.

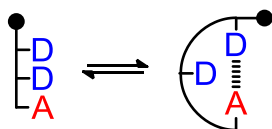


Figure 5.6. Schematic representation of 1,3-folding where an intramolecular H-bond can form between phenol (D) and pyridine N-oxide (A) on the same oligomer. The black circle represents the nitrobenzene terminus of the oligomer.

5.3.2.4 Double H-bond acceptors

If the association constants in Table 5.2 are all ranked in order of magnitude the five highest association constants in descending order correspond to duplexes of AAA•DDD ($1.0 \times 10^5 \text{ M}^{-1}$), ADA•DAD ($2.5 \times 10^4 \text{ M}^{-1}$), AAD•DDD ($1.5 \times 10^4 \text{ M}^{-1}$), ADA•DDD ($1.5 \times 10^4 \text{ M}^{-1}$) and DAA•DDD ($1.1 \times 10^4 \text{ M}^{-1}$). What is striking is that the association constant of DAD•ADA is significantly lower than DDD•AAA. Furthermore, each duplex of AAD•DDD, ADA•DDD and DAA•DDD is comprised of one oligomer with two *N*-oxide acceptors and a phenol paired with DDD. The models of 1,3-folding and the parallel and anti-parallel effect cannot be used to explain the difference between DAD•ADA and DDD•AAA, since none of the oligomers involved can fold and each duplex can be formed in both orientations. These duplexes all contain one D•D mismatch and do not have enough H-bond acceptors to form three H-bonds. However, there are multiple lone pairs on the oxygen atom of pyridine *N*-oxide groups, and these may be available to form additional H-bonds. Thus DDD could form two matched H-bonds and the third

donor group could form a second H-bond to one of the H-bonded acceptor groups (Figure 5.7). Further evidence for this effect was found by titrating **3.21** (Chapter 3) into the monofunctional 4-methylpyridine *N*-oxide **4.15** (Figure 5.8). The observed association constant for this duplex is $1.2 \times 10^3 \text{ M}^{-1}$, which is significantly larger than the association constant of *p*-cresol and 4-methylpyridine *N*-oxide ($3.3 \times 10^2 \text{ M}^{-1}$, Chapter 4). This double H-bond effect can also be used to explain the difference between DDD•AAA and DAD•ADA, since the double H-bond model means there are more ways for three H-bonds to be formed in DDD•AAA which would increase the binding constant for statistical reasons.

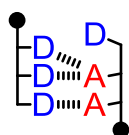


Figure 5.7. *N*-oxide acceptors can form more than one H-bond, as in this example (DDD•AAD). The black circle represents the nitrobenzene terminus of the oligomer.

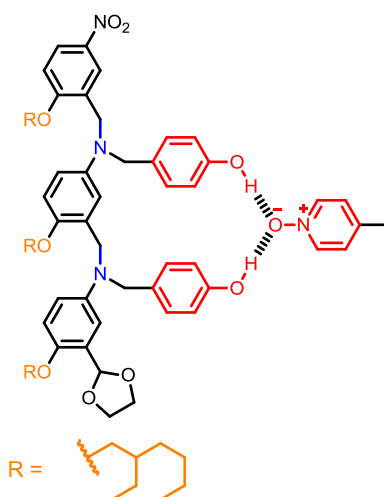


Figure 5.8. Two H-bond complex formed between **3.21** and **4.15**.

5.3.2.5 Sequence discrimination

Using the data in Table 5.2 and the trends discussed above we can begin to understand the ability of the oligomers to discriminate sequence when forming a duplex. Although the association constants in Table 5.2 range over three orders of magnitude from $1.0 \times 10^5 \text{ M}^{-1}$ (DDD•AAA) to $2.2 \times 10^2 \text{ M}^{-1}$ (ADD•DDD) not all of the oligomers take part in duplexes that span this entire range. In other words, the range of association constants for the duplexes that one oligomer forms with every other oligomer can occur over a much smaller range. For example the maximum association constant observed for duplexes involving DDA is just $4.6 \times 10^3 \text{ M}^{-1}$ (DDA•ADA) and the minimum is $1.2 \times 10^3 \text{ M}^{-1}$ (DDA•ADD). Therefore to examine selectivity, it is convenient to normalise the data relative to the highest association constant observed for each oligomer. These data are represented as a surface plot in Figure 5.9 and χ is defined in Equation 5.1,

$$\chi = \frac{K_{xy}}{K_{x,\max}} \quad (5.1)$$

where K_{xy} is the duplex association constant for the pair of oligomers on the x- and y-axes in Figure 5.9 and $K_{x,\max}$ is the maximum duplex association constant for the oligomer on the x-axis in Figure 5.9.

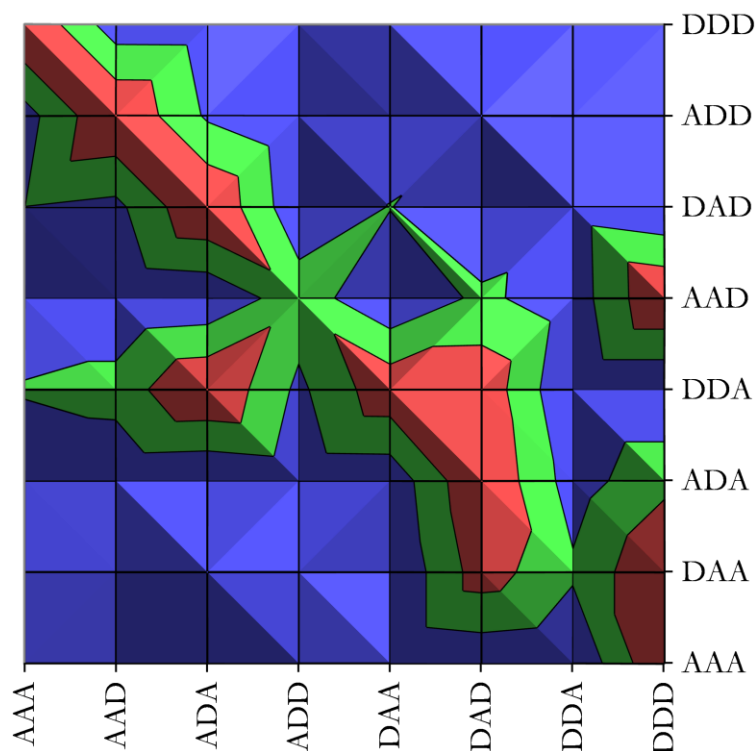


Figure 5.9. Normalised surface plot of χ (Equation 5.1) where blue is $\chi = 0.01 - 0.34$; green is $\chi = 0.35 - 0.67$ and red is $\chi = 0.68 - 1.00$.

In Figure 5.9, a diagonal line from the top left to the bottom right represents antiparallel sequence matched duplexes. Encouragingly the maximum relative association constant for most of the oligomers lies on the diagonal at the sequence matched complex. However, this is not the case for AAD, DDA and DAA, where sequence mismatched duplexes have the highest association constant. The unusually high association constants observed for the mismatched duplexes, AAD•DDD and DAA•DDD, are due to the formation of double H-bonds with an acceptor. The unusually low association constants observed for the corresponding matched duplexes, AAD•DDA and DAA•ADD, are due to 1,3-folding in both components of the duplex. The result is that non-symmetric sequences (AAD, ADD, DAA and DDA) discriminate poorly based on sequence. Even ADD which does have the highest association constant for the antiparallel sequence matched duplex formed

with AAD ($8.0 \times 10^3 \text{ M}^{-1}$) forms a mismatched duplex of similar stability with ADA ($2.3 \times 10^3 \text{ M}^{-1}$). Three of the four symmetric sequences (DDD, AAA, and DAD) are much better at discriminating based on sequence. The highest association constant for each of these oligomers is an order of magnitude higher than the next most stable duplex. However, symmetric sequence ADA is poor at discriminating based on sequence since the association constant of sequence matched ADA•DAD ($2.5 \times 10^4 \text{ M}^{-1}$) is not significantly greater than ADA•DDD ($1.5 \times 10^4 \text{ M}^{-1}$). This observation can be explained by the ability of double H-bond formation being possible in ADA•DDD.

5.4 Conclusions

Herein a sequence discrimination study of H-bonding duplexes is presented. A series of eight 3-mers has been synthesised using reductive amination chemistry to prepare all possible sequences of three H-bond donor (phenol) or acceptor (pyridine *N*-oxide) groups. The association constants for duplex formation for most pairwise combinations of the oligomers were measured by means of titration in toluene- d_8 . Collectively, the association constants show patterns which highlight a number of properties of the duplexes that impact upon their ability discriminate sequence:

- (i) oligomers that have a sequence of recognition groups that match the sequence on another oligomer tend to form duplexes with higher association constants;
- (ii) complementary sequences that form in an antiparallel sense have higher association constants than the same sequence duplex formed in a parallel sense;
- (iii) intramolecular H-bonds can form between the 1st and 3rd positions in the oligomer and this 1,3-folding decreases the association constant for duplexes formed from these oligomers;
- (iv) some sequence mismatched duplexes are stabilised by the ability of pyridine *N*-oxide groups to form more than one H-bond.

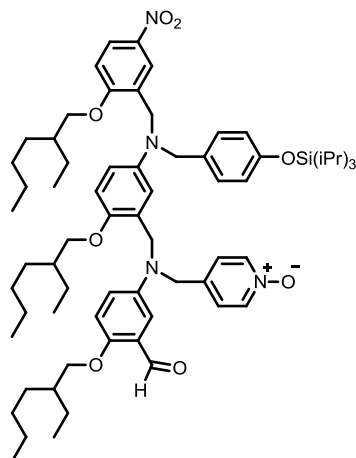
For many of the oligomers the highest association constant observed is for the duplex formed with its sequence complementary counterpart, but some sequences show poor discrimination. The ability of this system to discriminate based on recognition group sequence could be improved if points (iii) and (iv) above were to be addressed. For instance a slightly more rigid backbone might reduce 1,3-folding. Furthermore, if an H-bond acceptor that is only capable of forming a single H-bond were used, the fidelity of sequence recognition might be improved.

5.5 Experimental section

5.5.1 Synthesis

All the reagents and materials used in the synthesis of the compounds described below were bought from commercial sources, without prior purification. Thin layer chromatography was carried out using with silica gel 60F (Merck) on aluminium. Flash chromatography was carried out on silica gel 40 – 60 μm (BDH) or on an automated system (Combiflash Companion) using pre-packed cartridges of silica (50 μ PuriFlash® Column). All NMR spectroscopy was carried out on either a Bruker AVI250, AVI400, DPX400, AVIII400 or DRX500 spectrometer using the residual solvent as the internal standard. All chemical shifts (δ) are quoted in ppm and coupling constants given in Hz. Splitting patterns are given as follows: s (singlet), d (doublet), t (triplet), m (multiplet). FT-IR spectra were measured on a PerkinElmer Spectrum 100 spectrometer. ES+ was carried out on a Micromass Platform spectrometer. Reactions were carried out at ambient temperature unless otherwise stated.

5.5.1.1 Synthesis of 4.13'



4.13 (2.07 g, 1.79 mmol, 1 equiv.) was dissolved in CHCl_3 (10 mL) and concentrated aqueous acid (10 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed using a rotary evaporator to yield a bright yellow oil (2.10 g, 95%) requiring no further purification.

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} = 10.44$ (s, 1H), 8.10 (d, 2H, $^3J = 7.0$), 8.04 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 7.07 – 7.02 (m, 4H), 6.91 (d, 1H, $^4J = 3.0$), 6.84 (d, 2H $^3J = 8.5$), 6.79 (d, 1H, $^3J = 9.0$), 6.74 (d, 2H, $^3J = 9.0$), 6.68 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.56 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.26 (d, 1H, $^4J = 3.0$), 4.49 – 4.42 (m, 6H), 4.14 (s, 2H), 3.94 – 3.87 (m, 4H), 3.78 (d, 2H, $^3J = 5.5$), 1.82 – 1.72 (m, 2H), 1.71 – 1.63 (m, 1H), 1.60 – 1.19 (m, 27H), 1.10 (d, 18H, $^3J = 7.5$), 1.01 – 0.83 (m, 18H);

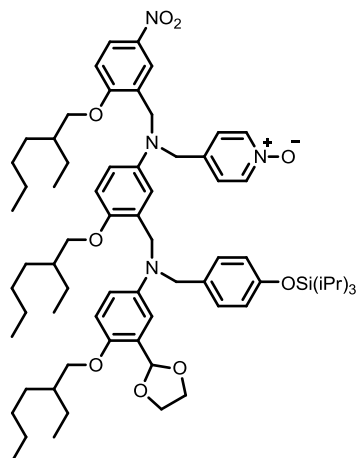
^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 189.7$, 161.5, 155.0, 154.6, 149.1, 142.5, 142.1, 141.3, 139.1, 138.5, 130.9, 128.4, 127.9, 125.5, 125.0, 124.2, 124.1, 123.0, 120.8, 120.0, 114.1, 113.0, 112.5, 112.0, 110.3, 110.2, 71.5, 71.4, 70.7, 55.4, 52.8, 50.4, 50.1, 39.6,

39.5, 39.2, 30.7, 30.6, 30.5, 29.1, 29.1, 29.0, 24.0, 24.0, 24.0, 23.0, 23.0, 17.9, 14.0, 12.6, 11.2, 11.2, 11.2;

HRMS (ES+): calculated for $C_{67}H_{99}N_4O_8Si$ 1115.7232, found 1115.7229;

FT-IR (thin film): ν_{max}/cm^{-1} 2957, 2927, 2865, 1682, 1609, 1592, 1505, 1464, 1338, 1262, 1226, 1166.

5.5.1.2 Synthesis of 5.1



4.7' (0.51 g, 0.82 mmol, 1 equiv.) and **4.2** (0.69 g, 1.2 mmol, 1.5 equiv.) were dissolved in CHCl_3 (3 mL) and $\text{NaBH}(\text{OAc})_3$ (0.49 g, 2.3 mmol, 2.8 equiv.) was added with stirring. After 4 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in a 1:1 mixture of EtOAc and hexane to yield a yellow oil (0.57 g, 59%).

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} = 8.12 - 8.03$ (m, 3H), 7.91 (d, 1H, $^3J = 3.0$), 7.07 – 6.98 (m, 4H), 6.84 (d, 1H, $^3J = 9.0$), 6.78 (d, 2H, $^3J = 8.5$), 6.75 (d, 1H, $^4J = 3.0$), 6.71 (d, 1H, $^3J = 9.0$), 6.59 (d, 1H, $^3J = 9.0$), 6.43 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.39 (d, 1H, $^4J = 3.0$), 6.34 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), (s, 1H), 4.47 (s, 2H), 4.45 (s, 2H), 4.44 (s, 2H), 4.23 (s, 2H), 3.96 (d, 2H, $^3J = 5.5$), 3.93 (s, 4H), 3.81 (d, 2H, $^3J = 5.5$), 3.77 (d, 2H, $^3J = 5.5$), 1.83 – 1.60 (m, 3H), 1.59 – 1.17 (m, 27H), 1.09 (d, 18H, $^3J = 7.5$), 0.99 – 0.84 (m, 18H);

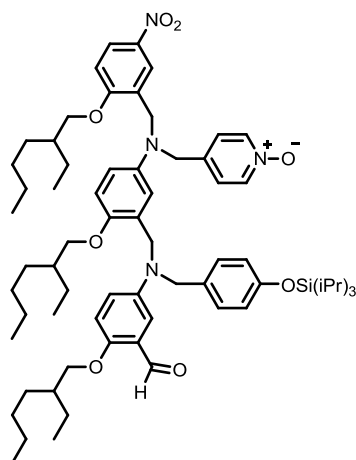
^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.5, 154.8, 150.0, 149.1, 143.0, 141.2, 139.1, 138.6, 131.2, 127.7, 127.7, 127.4, 126.4, 124.6, 124.3, 122.9, 119.8, 113.5, 113.5, 113.3,$

112.2, 111.8, 111.0, 110.7, 99.5, 71.7, 71.5, 70.7, 65.0, 54.6, 54.3, 50.8, 49.4, 39.6, 39.5, 39.2, 30.6, 30.5, 30.5, 29.1, 29.1, 29.0, 24.0, 23.9, 23.9, 23.1, 23.0, 23.0, 17.9, 14.1, 14.0, 12.6, 11.2, 11.1, 11.1;

HRMS (ES+): calculated for $C_{69}H_{103}N_4O_9Si$ 1159.7494, found 1159.7478;

FT-IR (thin film): ν_{max}/cm^{-1} 3055, 2987, 2686, 2306, 1422, 1266, 1156.

5.5.1.3 Synthesis of 5.1'



5.1 (0.57 g, 0.49 mmol, 1 equiv.) was dissolved in CHCl_3 (5 mL) and concentrated aqueous acid (5 mL) was added with stirring. After 2 days the mixture was neutralised using aqueous NaHCO_3 and the organic portion separated from the aqueous part. The aqueous layer was washed with CHCl_3 (3×10 mL) before all organic fractions were washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed using a rotary evaporator to yield a bright yellow oil (0.50 g, 91%) requiring no further purification.

^1H NMR (500 MHz, CDCl_3): $\delta_{\text{H}} = 10.42$, (s, 1H), 8.08 (d, 2H, $^3J = 7.0$), 8.03 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 7.85 (d, 1H, $^4J = 3.0$), 7.08 (d, 2H, $^3J = 7.0$), 6.99 (d, 2H, $^3J = 8.5$), 6.96 (d, 1H, $^4J = 3.0$), 6.82 – 6.76 (m, 3H), 6.73 (d, 1H, $^3J = 9.0$), 6.64 (d, 1H, $^3J = 9.0$), 6.56 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.45 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.30 (d, 1H, $^4J = 3.0$), 4.49 (s, 2H), 4.45 (s, 2H), 4.43 (s, 2H), 4.29 (s, 2H), 3.92 (d, 2H, $^3J = 5.5$), 3.88 (d, 2H, $^3J = 5.5$), 3.79 (d, 2H, $^3J = 5.5$), 1.81 – 1.63 (m, 3H), 1.59 – 1.17 (m, 27H), 1.08 (d, 18H, $^3J = 7.5$), 1.02 – 0.83 (m, 18H);

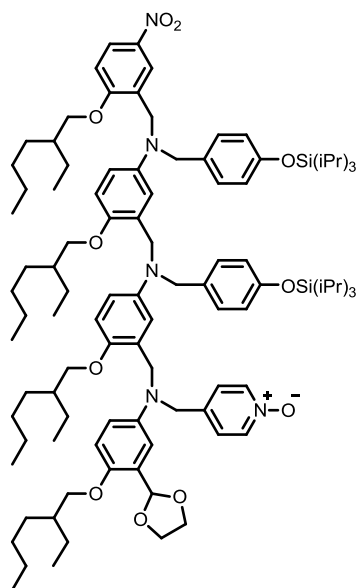
^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 190.2$, 161.7, 155.2, 154.1, 150.2, 142.9, 141.4, 141.4, 139.3, 138.6, 130.7, 127.9, 127.5, 127.3, 125.2, 124.7, 124.4, 123.0, 120.3, 120.2, 113.9, 113.3, 112.5, 112.1, 110.7, 109.7, 71.8, 71.5, 70.9, 54.7, 54.4, 51.0, 49.6, 39.8,

39.7, 39.4, 30.9, 30.8, 30.7, 29.3, 29.3, 29.2, 24.2, 24.2, 24.2, 23.2, 23.2, 23.2, 18.1,
14.3, 14.3, 14.2, 12.8, 11.4, 11.4, 11.4;

HRMS (ES+): calculated for $C_{67}H_{98}N_4O_8NaSi$ 1137.7052, found 1137.7024;

FT-IR (thin film): ν_{max}/cm^{-1} 3054, 2927, 2306, 1677, 1507, 1422, 1266.

5.5.1.4 Synthesis of 5.2



3.15' (0.70 g, 0.55 mmol, 1 equiv.) and **4.2** (0.33 g, 0.82 mmol, 1.5 equiv.) were dissolved in CHCl_3 (4 mL) and $\text{NaBH}(\text{OAc})_3$ (0.33 g, 1.5 mmol, 2.8 equiv.) was added with stirring. After 4 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 5% to 100% EtOAc in hexane and then 10% MeOH in CHCl_3 to yield a pale yellow oil (0.32 g, 35%).

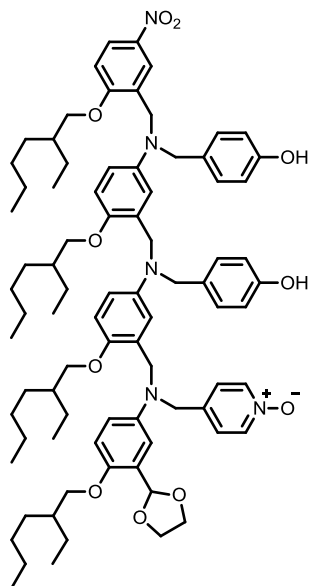
^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.04$ (d, 2H, $^3J = 7.0$), 8.00 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 7.96 (d, 1H, $^4J = 3.0$), 7.04 – 6.99 (m, 4H), 6.95 (d, 2H, $^3J = 8.5$), 6.85 (d, 1H, $^4J = 3.0$), 6.82 (d, 2H, $^3J = 8.5$), 6.79 – 6.76 (m, 3H), 6.75 – 6.69 (m, 3H), 6.61 (d, 1H, $^3J = 9.0$), 6.52 – 6.44 (m, 3H), 6.34 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.25 (d, 1H, $^4J = 3.0$), 6.11 (s, 1H), 4.49 – 4.43 (m, 6H), 4.06 (s, 2H), 4.02 – 3.93 (m, 4H), 3.90 (s, 2H), 3.85 (d, 2H, $^3J = 5.5$), 3.82 – 3.75 (m, 6H), 1.78 – 1.62 (m, 4H), 1.56 – 1.19 (m, 38H), 1.14 – 1.07 (m, 36H), 0.97 – 0.85 (m, 24H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.4, 154.9, 154.6, 150.0, 149.1, 148.4, 142.8, 142.7, 142.4, 141.3, 139.6, 138.8, 131.9, 131.1, 128.7, 128.2, 127.7, 127.5, 126.9, 125.8, 124.3, 124.1, 123.1, 119.9, 115.0, 113.6, 113.2, 112.8, 112.4, 112.3, 112.2, 111.4, 111.4, 110.3, 99.4, 71.8, 71.4, 70.8, 70.8, 65.1, 55.4, 54.5, 52.6, 50.8, 50.5, 50.2, 39.6, 39.6, 39.2, 30.7, 30.6, 30.5, 29.1, 29.1, 24.0, 23.9, 23.9, 23.1, 23.0, 23.0, 17.9, 17.9, 14.1, 12.7, 11.2, 11.1;$

HRMS (ES+): calculated for $\text{C}_{100}\text{H}_{152}\text{N}_5\text{O}_{11}\text{Si}_2$ 1655.1027, found 1655.1080;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2928, 2865, 1608, 1592, 1505, 1463, 1339, 1261, 1225.

5.5.1.5 Synthesis of 5.3



5.2 (0.097 g, 0.059 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (110 μ L, 0.11 mmol, 1.8 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4×10 mL). All organic fractions were combined and washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 5% of MeOH in CHCl_3 to yield a viscous pale yellow oil (0.064 g, 81%).

$^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} = 8.16 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 8.11 – 8.06 (m, 3H), 6.99 (d, 1H, $^3J = 7.0$), 6.94 (d, 1H, $^3J = 9.0$), 6.88 (d, 2H, $^3J = 8.5$), 6.83 – 6.72 (m, 5H), 6.70 (d, 1H, $^3J = 9.0$), 6.71 – 6.64 (m, 3H), 6.62 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.54 – 6.48 (m, 4H), 6.39 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.24 (d, 1H, $^4J = 3.0$), 6.14 (s, 1H), 4.66 (s, 2H), 4.50 – 4.46 (m, 4H), 4.41 (s, 2H), 4.24 – 4.03 (m, 6H), 4.02 (d, 2H, $^3J = 5.5$), 3.87 – 3.81 (m, 4H), 3.77 (d, 2H, $^3J = 5.5$), 3.74 (s, 2H), 1.87 – 1.60 (m, 4H), 1.58 – 1.24 (m, 32H), 1.01 – 0.85 (m, 18H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 162.0, 156.0, 155.3, 150.5, 149.1, 148.2, 143.6, 143.6, 142.4, 142.4, 141.7, 138.8, 130.7, 129.6, 129.3, 127.8, 127.7, 127.5, 125.8, 125.0, 124.5, 123.3, 116.0, 115.6, 115.4, 113.7, 113.3, 112.7, 112.7, 112.5, 112.1, 111.8, 110.7, 110.5, 100.1, 71.7, 71.6, 70.8, 70.7, 65.3, 55.6, 55.0, 54.0, 51.7, 51.5, 51.1, 39.9, 39.8, 39.7, 39.4, 31.0, 30.9, 30.8, 30.7, 29.4, 29.3, 29.3, 24.3, 24.2, 24.2, 24.1, 23.3, 23.3, 23.2, 23.2, 14.3, 14.3, 11.5, 11.4, 11.4, 11.3;

HRMS (ES+): calculated for C₈₂H₁₁₂N₅O₁₁ 1342.8358, found 1342.8372;

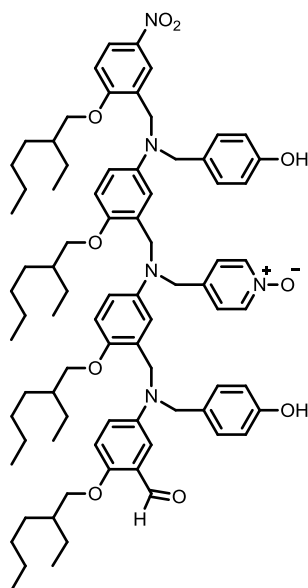
FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 3052, 3007, 2961, 2930, 2873, 1505, 1468, 1382, 1340, 1264, 1225.

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.7, 155.1, 154.9, 149.6, 149.4, 143.2, 142.8, 142.5, 141.5, 139.6, 139.0, 131.4, 131.1, 128.9, 128.3, 128.0, 127.5, 126.9, 126.9, 124.4, 123.4, 120.1, 120.0, 114.2, 113.9, 113.7, 113.4, 112.5, 112.5, 112.4, 112.2, 111.9, 110.5, 99.8, 72.0, 71.6, 71.1, 70.9, 65.2, 55.4, 54.6, 53.4, 51.3, 50.4, 49.8, 39.8, 39.7, 39.3, 30.9, 30.8, 30.7, 30.7, 29.3, 29.3, 29.2, 24.2, 24.2, 24.1, 24.1, 23.3, 23.2, 23.2, 18.1, 18.1, 14.3, 14.3, 14.3, 12.8, 11.4, 11.3, 11.3;

HRMS (ES+): calculated for C₁₀₀H₁₅₂N₅O₁₁Si₂ 1655.1027, found 1655.0964;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2927, 2866, 1609, 1592, 1506, 1464, 1339, 1261, 1225.

5.5.1.7 Synthesis of 5.5'



5.4 (0.10 g, 0.062 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (110 μ L, 0.11 mmol, 1.8 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 5% of MeOH in DCM to yield the benzaldehyde derivative of expected product which is a viscous yellow oil (0.071 g, 85%).

¹H NMR (500 MHz, CDCl₃): δ_{H} = 10.44 (s, 1H), 8.12 (dd, 1H, ³*J* = 9.0, ⁴*J* = 3.0), 8.01 (d, 1H, ⁴*J* = 3.0), 7.96 (d, 2H, *J* = 7.0), 7.25 (d, 1H, ⁴*J* = 3.0), 6.92 – 6.88 (m, 3H), 6.87 – 6.83 (m, 4H), 6.81 (d, 1H, ⁴*J* = 3.0), 6.79 (d, 1H, ³*J* = 9.0), 6.74 – 6.70 (m, 3H), 6.67 – 6.63 (m, 3H), 6.52 (dd, 1H, ³*J* = 9.0, ⁴*J* = 3.0), 6.43 (d, 1H, ⁴*J* = 3.0), 6.39 (dd, 1H, ³*J* = 9.0, ⁴*J* = 3.0), 6.33 (d, 1H, ⁴*J* = 3.0), 4.56 (s, 2H), 4.42 (s, 2H), 4.38 (s, 2H), 4.32 (s, 2H), 4.19 (s, 2H), 3.97 (d, 2H, ³*J* = 5.5), 3.90 (s, 2H), 3.86 (d, 2H, ³*J* = 5.5),

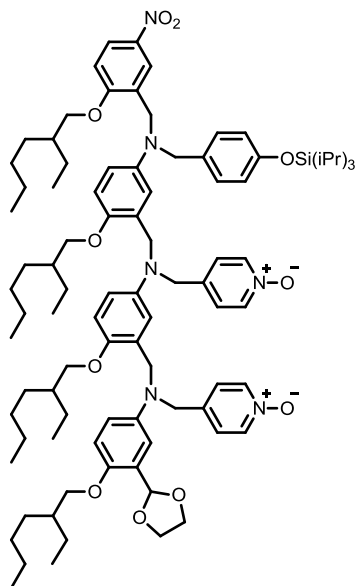
3.83 (d, 2H, $^3J = 5.5$), 3.71 (d, 2H, $^3J = 5.5$), 1.80 – 1.67 (m, 4H), 1.53 – 1.19 (m, 32H), 0.96 – 0.81 (m, 24H);

^{13}C NMR (126 MHz, CDCl_3): $\delta_{\text{C}} = 190.5, 161.8, 155.4, 154.6, 149.6, 149.0, 143.8, 143.1, 142.6, 142.2, 141.4, 138.7, 130.0, 129.7, 128.7, 128.2, 127.6, 127.2, 126.2, 124.9, 124.4, 124.2, 123.1, 122.4, 115.8, 115.6, 113.8, 113.7, 113.0, 112.4, 112.3, 112.1, 111.5, 110.4, 71.4, 71.3, 70.8, 70.6, 55.5, 55.2, 53.3, 51.1, 51.0, 50.4, 39.6, 39.5, 39.2, 30.7, 30.6, 30.6, 30.5, 29.7, 29.1, 29.1, 29.0, 24.0, 24.0, 23.9, 23.1, 23.0, 23.0, 14.1, 14.0, 11.2, 11.2, 11.1$;

HRMS (ES+): calculated for $\text{C}_{82}\text{H}_{112}\text{N}_5\text{O}_{11}$ 1342.8358, found 1342.8336;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2959, 2928, 2871, 2856, 1680, 1666, 1613, 1592, 1505, 1465, 1340, 1263, 1225, 1165.

5.5.1.8 Synthesis of 5.6



4.13' (0.77 g, 0.69 mmol, 1 equiv.) and **4.2** (0.41 g, 1.0 mmol, 1.5 equiv.) were dissolved in CHCl_3 (5 mL) and $\text{NaBH}(\text{OAc})_3$ (0.41 g, 1.9 mmol, 2.8 equiv.) was added with stirring. After 4 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% EtOH in diethyl ether and then 2% to 10% EtOH in CHCl_3 to yield a pale yellow oil (0.31 g, 28%).

$^1\text{H NMR}$ (500 MHz, CDCl_3): $\delta_{\text{H}} = 8.09 - 8.04$ (m, 3H), 8.02 (d, 2H, $^3J = 7.0$), 7.95 (d, 1H, $^4J = 3.0$), 7.05 (d, 2H, $^3J = 7.0$), 6.98 (d, 2H, $^3J = 8.5$), 6.91 (d, 2H, $^3J = 7.0$), 6.86 (d, 1H, $^3J = 9.0$), 6.80 (d, 1H, $^4J = 3.0$), 6.77 (d, 2H, $^3J = 8.5$), 6.73 – 6.66 (m, 3H), 6.51 – 6.39 (m, 4H), 6.26 (d, 1H, $^4J = 3.0$), 6.08 (s, 1H), 4.46 (s, 2H), 4.45 (s, 2H), 4.43 (s, 2H), 4.41 (s, 2H), 4.09 (s, 2H), 4.01 – 3.92 (m, 8H), 3.82 – 3.77 (m, 4H),

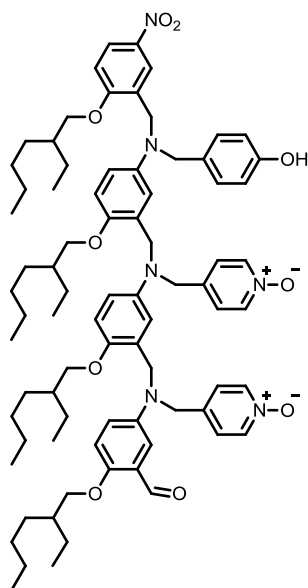
3.75 (d, 2H, $^3J = 5.5$), 1.81 – 1.60 (m, 4H), 1.56 – 1.17 (m, 35H), 1.08 (d, 18H, $^3J = 7.0$), 0.97 – 0.80 (m, 24H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} = 161.6, 155.0, 150.2, 149.4, 149.2, 142.6, 142.1, 142.1, 141.3, 139.3, 139.0, 138.8, 130.8, 128.6, 128.1, 126.8, 126.5, 126.3, 124.3, 124.2, 123.1, 119.9, 114.8, 113.6, 113.6, 113.4, 112.6, 112.4, 112.3, 112.0, 110.3, 99.3, 71.8, 71.5, 70.9, 70.8, 65.1, 55.2, 53.4, 53.1, 51.2, 50.3, 50.1, 39.5, 39.2, 30.6, 30.5, 29.1, 29.1, 29.0, 24.0, 24.0, 23.9, 23.0, 23.0, 23.0, 17.9, 14.1, 14.0, 12.6, 11.2, 11.1$;

HRMS (ES+): calculated for $\text{C}_{90}\text{H}_{131}\text{N}_6\text{O}_{11}\text{Si}$ 1499.9645, found 1499.9701;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 2958, 2927, 2868, 1592, 1506, 1339, 1263.

5.5.1.9 Synthesis of 5.7'



5.6 (0.078 g, 0.052 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (47 μ L, 0.047 mmol, 0.9 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 10 mL) dried (MgSO₄) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 10% of MeOH in DCM to yield the benzaldehyde derivative of expected product which is a viscous yellow oil (0.057 g, 82%).

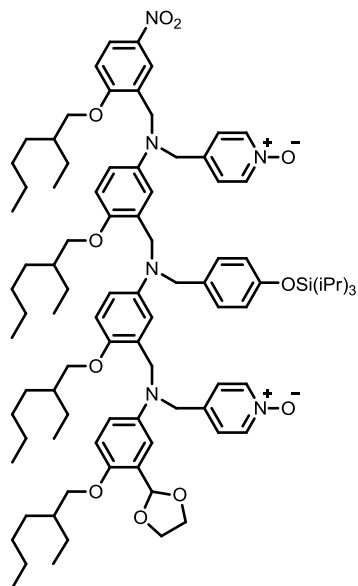
¹H NMR (500 MHz, CDCl₃): δ_{H} = 10.46 (s, 1H), 8.13 (dd, 1H, ³*J* = 9.0, ⁴*J* = 3.0), 8.08 (d, 2H, ³*J* = 6.5), 7.99 (d, 1H, ⁴*J* = 3.0), 7.90 (d, 2H, ³*J* = 6.5 Hz), 7.11 (d, 1H, ⁴*J* = 3.0), 7.01 (d, 2H, ³*J* = 6.5), 6.93 – 6.87 (m, 4H), 6.84 – 6.78 (m, 2H), 6.76 – 6.64 (m, 5H), 6.59 (dd, 1H ⁴*J* = 9.0, ⁴*J* = 3.0), 6.47 (d, 1H, ⁴*J* = 3.0), 6.42 – 6.38 (m, 1H), 6.24 (d, 1H, ⁴*J* = 3.0), 4.56 (s, 2H), 4.43 (s, 2H), 4.40 (s, 2H), 4.34 (s, 2H), 4.20 (s, 2H), 3.98 (d, 2H, ³*J* = 5.5), 3.93 – 3.88 (m, 4H), 3.80 (d, 2H, *J* = 5.5), 3.71 (d, 2H, ³*J* = 5.5), 1.83 – 1.64 (m, 4H), 1.52 – 1.17 (m, 32H), 0.98 – 0.80 (m, 24H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 190.0, 161.8, 155.9, 155.1, 149.5, 149.0, 142.6, 142.3, 141.4, 140.3, 139.9, 138.9, 138.8, 129.4, 128.6, 127.6, 126.2, 125.9, 125.0, 124.5, 123.9, 123.0, 121.7, 115.8, 114.1, 113.9, 113.7, 112.6, 112.3, 111.7, 111.5, 110.4, 71.4, 71.4, 70.8, 70.6, 55.3, 53.7, 52.7, 51.5, 51.0, 50.9, 39.5, 39.2, 30.6, 30.6, 30.5, 29.7, 29.1, 29.1, 29.0, 24.0, 23.0, 23.0, 14.1, 11.2, 11.1;

HRMS (ES+): calculated for C₈₁H₁₁₁N₆O₁₁ 1343.8311, found 1343.8279;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2925, 2855, 1682, 1612, 1503, 1339, 1226, 1167, 1019.

5.5.1.10 Synthesis of 5.8



5.1' (0.51 g, 0.46 mmol, 1 equiv.) and **4.2** (0.36 g, 0.91 mmol, 2 equiv.) were dissolved in CHCl_3 (3 mL) and $\text{NaBH}(\text{OAc})_3$ (0.27 g, 1.3 mmol, 2.8 equiv.) was added with stirring. After 3 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% EtOH in EtOAc to yield a pale yellow oil (0.38 g, 56%).

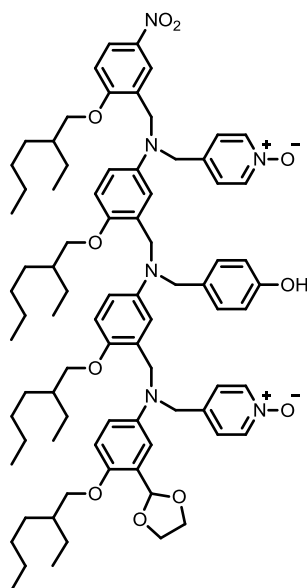
^1H NMR (400 MHz, CDCl_3): δ_{H} = 8.03 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 8.00 (d, 2H, $^3J = 7.0$), 7.92 (d, 2H, $^3J = 7.0$), 7.89 (d, 1H, $^4J = 3.0$), 6.99 (d, 2H, $^3J = 7.0$), 6.92 (d, 2H, $^3J = 8.5$), 6.88 (d, 2H, $^3J = 7.0$), 6.84 – 6.77 (m, 4H), 6.68 (d, 1H, $^3J = 9.0$), 6.67 (d, 1H, $^3J = 9.0$), 6.58 (d, 1H, $^3J = 9.0$), 6.44 – 6.37 (m, 2H), 6.32 (d, 1H, $^4J = 3.0$), 6.26 – 6.21 (m, 2H), 6.02 (s, 1H), 4.47 – 4.40 (m, 6H), 4.33 (s, 2H), 4.04 (s, 2H), 3.96 – 3.85 (m, 8H), 3.78 (d, 2H, $^3J = 5.5$), 3.75 – 3.71 (m, 4H), 1.75 – 1.59 (m, 4H), 1.51 – 1.16 (m, 35H), 1.06 (d, 18H, $^3J = 7.5$), 0.94 – 0.78 (m, 24H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.7, 154.9, 150.1, 150.0, 148.6, 142.6, 142.3, 141.4, 141.3, 139.8, 139.1, 139.0, 138.6, 131.6, 128.0, 127.9, 127.8, 127.0, 126.1, 124.8, 124.5, 124.4, 123.2, 120.2, 114.9, 113.8, 113.5, 112.6, 112.5, 112.3, 112.3, 112.2, 111.2, 110.8, 99.5, 72.0, 71.7, 70.9, 65.2, 54.8, 54.5, 53.2, 51.3, 50.8, 50.6, 39.8, 39.6, 39.3, 30.8, 30.8, 30.6, 29.3, 29.2, 29.2, 24.2, 24.1, 24.0, 23.2, 23.2, 23.2, 23.1, 18.1, 14.3, 14.2, 14.2, 12.8, 11.4, 11.3, 11.3, 11.2;

HRMS (ES+): calculated for C₉₀H₁₃₁N₆O₁₁Si 1499.9645, found 1499.9578;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2962, 2929, 2867, 1612, 1593, 1507, 1479, 1466, 1340, 1263.

5.5.1.11 Synthesis of 5.9



5.8 (0.11 g, 0.073 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (66 μ L, 0.66 mmol, 0.9 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 10% of MeOH in DCM to yield a viscous yellow oil (0.088 g, 89%).

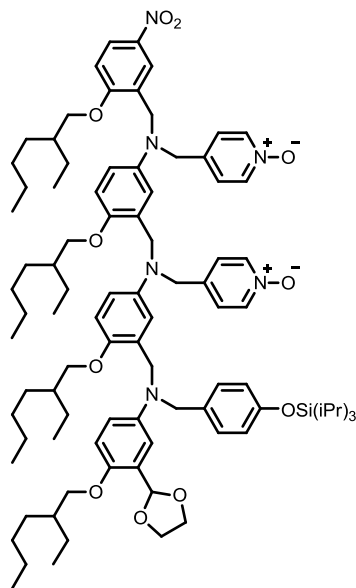
$^1\text{H NMR}$ (500 MHz, CDCl_3): δ_{H} = 8.11 – 8.06 (m, 3H), 8.04 (d, 2H, $^3J = 7.0$), 7.96 (d, 1H, $^4J = 3.0$), 7.09 (d, 2H, $^3J = 7.0$), 6.99 (d, 2H, $^3J = 7.0$), 6.93 – 6.88 (m, 3H), 6.86 (d, 1H, $^3J = 9.0$), 6.80 (d, 2H, $^3J = 8.5$), 6.77 (d, 1H, $^3J = 9.0$), 6.73 (d, 1H, $^3J = 9.0$), 6.65 (d, 1H, $^3J = 9.0$), 6.52 – 6.44 (m, 3H), 6.37 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.29 (d, 1H, $^4J = 3.0$), 6.10 (s, 1H), 4.51 (s, 2H), 4.49 (s, 2H), 4.47 (s, 2H), 4.43 (s, 2H), 4.10 (s, 2H), 4.07 – 3.91 (m, 8H), 3.85 – 3.78 (m, 6H), 1.82 – 1.65 (m, 4H), 1.56 – 1.23 (m, 32H), 0.99 – 0.84 (m, 24H).

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.6, 156.1, 150.2, 149.8, 148.6, 142.8, 142.2, 141.7, 141.3, 141.2, 139.6, 138.9, 138.8, 129.7, 128.3, 127.7, 126.7, 125.8, 124.7, 124.5, 124.5, 123.1, 115.9, 115.3, 113.6, 113.5, 113.2, 112.6, 112.2, 112.1, 112.0, 111.8, 110.6, 99.4, 71.8, 71.5, 70.7, 70.7, 65.1, 55.2, 54.3, 53.2, 51.7, 51.0, 50.6, 39.6, 39.5, 39.5, 39.2, 30.7, 30.7, 30.5, 29.1, 29.1, 29.1, 29.1, 24.1, 24.0, 23.9, 23.1, 23.1, 23.1, 23.0, 14.1, 14.1, 14.1, 11.2, 11.2, 11.2, 11.1;

HRMS (ES+): calculated for C₈₁H₁₁₁N₆O₁₁ 1343.8311, found 1343.8309;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2956, 2927, 2856, 1682, 1612, 1593, 1504, 1464, 1339, 1265, 1225, 1165,

5.5.1.12 Synthesis of 5.10



4.7' (0.37 g, 0.38 mmol, 1 equiv.) and **3.11** (0.31 g, 0.76 mmol, 2 equiv.) were dissolved in CHCl_3 (3 mL) and $\text{NaBH}(\text{OAc})_3$ (0.23 g, 1.1 mmol, 2.8 equiv.) was added with stirring. After 3 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4×10 mL). All the organic fractions were washed with water (1×10 mL), brine (1×10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% MeOH in a 1:1 mixture of EtOAc and CHCl_3 to yield a pale yellow oil (0.11 g, 18%).

^1H NMR (400 MHz, CDCl_3): $\delta_{\text{H}} = 8.10$ (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 8.03 (d, 2H, $^3J = 6.5$), 7.99 (d, 2H, $^3J = 6.5$), 7.93 (d, 1H, $^4J = 3.0$), 7.01 (d, 2H, $^3J = 8.5$), 6.89 – 6.94 (m, 4H), 6.88 (d, 1H, $^3J = 9.0$), 6.82 (d, 1H, $^4J = 3.0$), 6.77 (d, 2H, $^3J = 8.5$), 6.62 – 6.72 (m, 3H), 6.49 (dd, 1H, $^3J = 9$, $^4J = 3$), 6.45 (dd, 1H, $^3J = 9$, $^4J = 3$), 6.39 (d, 1H, $^4J = 3.0$), 6.35 (d, 1H, $^4J = 3.0$), 6.31 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.09 (s, 1H), 4.47 (s, 4H), 4.40 (s, 2H), 4.34 (s, 2H), 4.22 (s, 2H), 4.03 (s, 2H), 3.93 (d, 2H, $^3J = 5.5$), 3.90 (s,

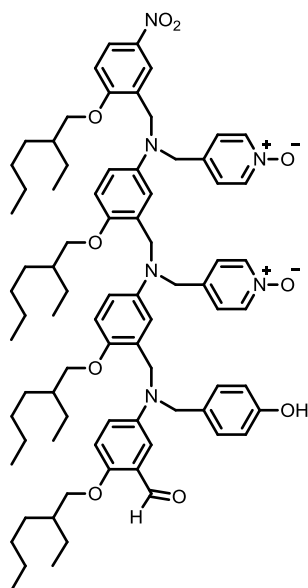
4H), 3.82 (d, 2H, $^3J = 5.5$), 3.73 – 3.79 (m, 4H), 1.61 – 1.80 (m, 4H), 1.17 – 1.53 (m, 32H), 1.04 – 1.11 (m, 21H), 0.82 – 0.95 (m, 24H);

^{13}C NMR (101 MHz, CDCl_3): $\delta_{\text{C}} =$ 161.81, 155.00, 150.14, 149.66, 149.45, 143.18, 142.21, 141.47, 141.37, 139.17, 139.10, 138.50, 131.32, 128.02, 127.81, 127.19, 126.88, 124.88, 124.49, 124.39, 123.29, 120.03, 114.41, 113.86, 113.78, 113.08, 112.76, 112.64, 112.45, 111.96, 110.81, 99.72, 72.04, 71.79, 71.09, 70.89, 68.14, 65.20, 55.08, 54.46, 53.51, 51.24, 50.90, 49.73, 39.74, 39.69, 39.62, 39.32, 30.83, 30.80, 30.66, 29.31, 29.24, 29.22, 29.17, 24.18, 24.09, 24.02, 23.22, 23.18, 23.14, 18.08, 14.27, 14.24, 12.80, 11.37, 11.34, 11.28;

HRMS (ES+): calculated for $\text{C}_{90}\text{H}_{131}\text{N}_6\text{O}_{11}\text{Si}$ 1499.9645, found 1499.9613;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 3020, 2400, 1521, 1425, 1265, 1217, 909.

5.5.1.13 Synthesis of 5.11'



5.10 (0.054 g, 0.036 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (32 μ L, 0.032 mmol, 0.9 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4 \times 10 mL). All organic fractions were combined and washed with brine (1 \times 10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 10% of MeOH in DCM to yield the benzaldehyde derivative of expected product which is a viscous yellow oil (0.041 g, 87%).

^1H NMR (400 MHz, CDCl_3): δ_{H} = 10.41 (s, 1H), 8.10 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 8.04 (d, 2H, $^3J = 7.0$), 7.97 (d, 2H, $^3J = 7.0$), 7.92 (d, 1H, $^4J = 3.0$), 7.11 (d, 1H, $^4J = 3.0$), 7.00 (d, 2H, $^3J = 7.0$), 6.93 (d, 2H, $^3J = 8.5$), 6.90 – 6.85 (m, 3H), 6.80 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.76 (d, 1H, $^3J = 9.0$), 6.71 – 6.66 (m, 3H), 6.64 (d, 1H, $^3J = 9.0$), 6.45 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.34 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.32 – 6.28 (m, 2H), 4.48 (s, 2H), 4.42 (s, 2H), 4.39 (s, 2H), 4.34 (s, 2H), 4.21 (s, 2H), 4.01 (s, 2H),

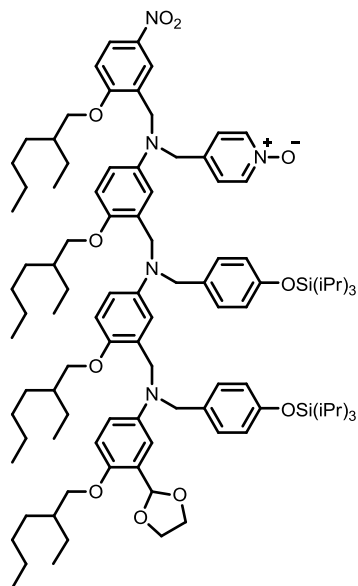
3.94 (d, 2H, $^3J = 5.5$), 3.87 – 3.82 (m, 4H), 3.72 (d, 2H, $^3J = 5.5$), 1.78 – 1.58 (m, 4H),
1.54 – 1.19 (m, 32H), 0.96 – 0.80 (m, 24H);

^{13}C NMR (126 MHz, CDCl_3): $\delta_{\text{C}} = 189.9, 161.7, 155.9, 154.5, 150.0, 149.3, 143.5,$
 $142.2, 141.3, 141.1, 139.6, 139.5, 138.9, 138.9, 130.1, 128.5, 127.5, 127.3, 126.8, 124.9,$
 $124.8, 124.4, 124.0, 123.0, 121.9, 115.7, 113.7, 113.2, 112.9, 112.5, 112.5, 112.3, 112.1,$
 $111.9, 110.6, 71.6, 71.3, 70.8, 70.6, 56.6, 54.2, 52.5, 51.1, 50.5, 50.4, 39.6, 39.5, 39.4,$
 $39.2, 30.7, 30.6, 30.5, 29.1, 29.1, 29.0, 29.0, 24.0, 24.0, 23.9, 23.1, 23.0, 23.0, 14.1,$
 $14.0, 11.2, 11.2;$

HRMS (ES+): calculated for $\text{C}_{79}\text{H}_{107}\text{N}_6\text{O}_{10}$ 1299.8049, found 1299.8041;

FT-IR (thin film): $\nu_{\text{max}}/\text{cm}^{-1}$ 3389, 2924, 2854, 1614, 1593, 1505, 1464, 1451, 1340,
1265, 1212, 1162, 1048.

5.5.1.14 Synthesis of 5.12



5.1' (0.33 g, 0.29 mmol, 1 equiv.) and **3.11** (0.32 g, 0.58 mmol, 2 equiv.) were dissolved in CHCl_3 (4 μL) and $\text{NaBH}(\text{OAc})_3$ (0.35 g, 1.64 mmol, 5.6 equiv.) was added with stirring. After 4 days the reaction was quenched with saturated aqueous NaHCO_3 solution, and extracted into CHCl_3 (4 \times 10 mL). All the organic fractions were washed with water (1 \times 10 mL), brine (1 \times 10 mL) and dried (MgSO_4) before the solvent removed with a rotary evaporator. The crude product was purified using flash chromatography on silica eluting with a gradient from 0% to 10% EtOH in a 1:1 mixture of EtOAc and hexane to yield a pale yellow oil (0.19 g, 0.48%).

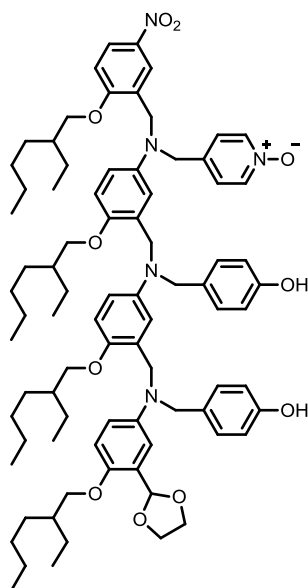
$^1\text{H NMR}$ (500 MHz, CDCl_3): δ_{H} = 8.04 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 7.96 – 7.92 (m, 3H), 6.98 – 6.92 (m, 4H), 6.88 – 6.83 (m, 3H), 6.79 (d, 2H, $^3J = 8.5$), 6.72 (d, 2H, $^3J = 8.5$), 6.67 (d, 2H, $^3J = 9.0$), 6.58 (d, 1H, $^3J = 9.0$), 6.50 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.42 – 6.37 (m, 2H), 6.33 (d, 1H, $^4J = 3.0$), 6.20 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 6.08 (s, 1H), 4.46 (s, 2H), 4.44 (s, 2H), 4.43 (s, 2H), 4.30 (s, 2H), 4.09 (s, 2H), 4.04 (s, 2H), 3.92 – 3.83 (m, 6H), 3.80 (d, 2H, $^3J = 5.5$), 3.76 – 3.70 (m, 4H), 1.76 – 1.60 (m, 4H), 1.56 – 1.14 (m, 38H), 1.12 – 1.01 (m, 36H), 0.96 – 0.78 (m, 24H);

¹³C NMR (101 MHz, CDCl₃): δ_C = 161.7, 154.9, 154.8, 150.0, 149.1, 148.7, 143.4, 142.9, 141.5, 141.3, 139.1, 138.9, 131.6, 131.6, 128.1, 128.0, 127.9, 127.9, 127.4, 126.9, 124.8, 124.5, 123.3, 120.1, 119.9, 114.4, 114.0, 113.7, 112.5, 112.4, 112.3, 112.1, 111.7, 110.9, 110.8, 99.8, 72.1, 71.7, 71.0, 70.9, 65.1, 54.7, 54.5, 54.4(CH₂N), 51.5, 50.5, 49.8, 39.8, 39.7, 39.7, 39.4, 30.9, 30.8, 30.7, 29.3, 29.3, 29.2, 24.2, 24.2, 24.1, 24.0, 23.3, 23.2, 23.2, 23.2, 18.1, 18.1, 14.3, 14.3, 14.2, 12.8, 12.8, 11.4, 11.4, 11.3, 11.3;

HRMS (ES+): calculated for C₁₀₀H₁₅₂N₅O₁₁Si₂ 1655.1027, found 1655.1057;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2926, 2955, 1608, 1591, 1504, 1463, 1339, 1260, 1225, 1165.

5.5.1.15 Synthesis of 5.13



5.12 (0.091 g, 0.055 mmol, 1 equiv.) was dissolved in THF (3 mL) at 0 °C and TBAF (100 μ L, 0.10 mmol, 1.8 equiv.) was added with stirring. After 1 hour water (5 mL) was added and the aqueous mixture washed with diethyl ether (4×10 mL). All organic fractions were combined and washed with brine (1×10 mL) dried (MgSO_4) and the solvent removed with a rotary evaporator. The crude mixture was then purified via flash chromatography on silica eluting with a gradient from 0% to 5% of MeOH in CHCl_3 to yield a viscous yellow oil (0.071 g, 97%).

$^1\text{H NMR}$ (500 MHz, CDCl_3): δ_{H} = 8.17 (dd, 1H, $^3J = 9.0$, $^4J = 3.0$), 8.12 (d, 2H, $^3J = 6.5$), 8.01 (d, 1H, $^4J = 3.0$), 7.00 (d, 2H, $^3J = 6.5$), 6.96 – 6.89 (m, 3H), 6.86 (d, 1H, $^4J = 3.0$), 6.80 – 6.66 (m, 7H), 6.58 (m, 3H), 6.49 – 6.39 (m, 3H), 6.36 (d, 1H, $^3J = 3.0$), 6.12 (s, 1H), 4.55 (s, 2H), 4.43 (s, 2H), 4.37 (s, 2H), 4.34 (s, 2H), 4.25 (s, 2H), 4.10 – 3.94 (m, 6H), 3.87 (s, 2H), 3.84 (d, 2H, $^3J = 5.5$), 3.82 – 3.75 (m, 4H), 1.85 – 1.61 (m, 4H), 1.59 – 1.21 (m, 32H), 1.02 – 0.80 (m, 24H);

$^{13}\text{C NMR}$ (126 MHz, CDCl_3): δ_{C} = 161.7, 155.6, 154.7, 149.9, 149.5, 148.2, 143.9, 143.5, 141.4, 141.3, 141.2, 138.7, 131.0, 129.5, 128.9, 128.4, 127.8, 127.7, 127.3, 125.5,

124.8, 124.7, 123.0, 116.2, 115.5, 114.9, 113.5, 113.3, 112.5, 112.3, 111.8, 111.5, 111.2, 110.6, 100.6, 71.7, 71.5, 70.7, 70.6, 65.1, 57.7, 54.1, 53.7, 51.1, 50.8, 49.6, 39.7, 39.5, 39.5, 39.2, 30.8, 30.6, 30.5, 29.2, 29.1, 29.1, 29.0, 24.1, 24.0, 23.9, 23.1, 23.0, 23.0, 14.2, 14.1, 14.1, 11.3, 11.1, 11.1;

HRMS (ES+): calculated for $C_{82}H_{112}N_5O_{11}$ 1342.8358, found 1342.8383;

FT-IR (thin film): $\nu_{\max}/\text{cm}^{-1}$ 2957, 2928, 2871, 2859, 1613, 1593, 1505, 1464, 1380, 1340, 1265, 1224, 1167, 1079, 1029.

5.5.2 Binding studies

All association constants were measured by ^1H NMR and additionally, if applicable, ^{31}P NMR titration. One species, labelled the host, is dissolved in toluene- d_8 to a known concentration. A second species, labelled the guest, is dissolved in the host solution and made to a known concentration such that the concentration of the host is the same in both solutions. A known volume of host is added to an NMR tube and the spectra measured. Known volumes of guest in host solution are added to the tube, and the spectra measured after each addition. The chemical shifts of the host spectra are recorded as a function of guest concentration and analysed using standard calculations in a fitting-program written in Microsoft Excel®. Errors were calculated as two times the standard deviation from the average value from repetitions.

5.6 Acknowledgements

I would like to thank Giulia who measured the association constants for all complexes except AAA•DDD and helped with characterising some of the 3-mers.

5.7 References

- 1 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 964–967.
- 2 W. M. Shih, J. D. Quispe, and G. F. Joyce, *Nature*, 2004, **427**, 618–21.
- 3 H. Gu, J. Chao, S.-J. Xiao, and N. C. Seeman, *Nature*, 2010, **465**, 202–205.
- 4 Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, and D. R. Liu, *Science*, 2004, **305**, 1601–5.
- 5 G. M. Church, Y. Gao, and S. Kosuri, *Science (80-.)*, 2012, **337**, 1628–1628.
- 6 M. Zuker, *Nucleic Acids Res.*, 2003, **31**, 3406–3415.
- 7 A. D. Ellington and J. W. Szostak, *Nature*, 1992, **355**, 850–852.
- 8 J. A. Piccirilli, S. A. Benner, T. Krauch, S. Moroney, E., and S. A. Benner, *Nature*, 1990, **343**, 33–37.
- 9 K.-U. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser, *Science (80-.)*, 2000, **290**, 1347–1351.
- 10 S. M. Gryaznov, D. H. Lloyd, J. K. Chen, R. G. Schultz, L. A. DeDionisio, L. Ratmeyer, and W. D. Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 5798–5802.
- 11 Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel, and M. R. Ghadiri, *Science*, 2009, **325**, 73–7.
- 12 S. M. Freier and K.-H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429–4443.
- 13 S. Polowiński, *Prog. Polym. Sci.*, 2002, **27**, 537–577.
- 14 R. McHale, J. P. Patterson, P. B. Zetterlund, and R. K. O'Reilly, *Nat. Chem.*, 2012, **4**, 491–497.
- 15 Y.-J. Kim, H. Uyama, and S. Kobayashi, *Macromolecules*, 2003, **36**, 5058–5060.
- 16 J.-F. Lutz, M. Ouchi, D. R. Liu, and M. Sawamoto, *Science*, 2013, **341**, 628.
- 17 G. Gody, T. Maschmeyer, P. B. Zetterlund, and S. Perrier, *Nat. Commun.*, 2013, **4**, 1-9.
- 18 K. Satoh, M. Matsuda, K. Nagai, and M. Kamigaito, *J. Am. Chem. Soc.*, 2010, **132**, 10003–10005.
- 19 J. Vandenberg, G. Reekmans, P. Adriaensens, and T. Junkers, *Chem. Commun.*, 2013, **49**, 10358.
- 20 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154

-
- 21 R. L. Letsinger and K. K. Ogilvie, *J. Am. Chem. Soc.*, 1969, **91**, 3350–3355.
 - 22 A. P. Bisson, F. J. Carver, D. S. Eggleston, R. C. Haltiwanger, C. A. Hunter, D. L. Livingstone, J. F. McCabe, C. Rotger, and A. E. Rowan, *J. Am. Chem. Soc.*, 2000, **122**, 8856–8868.
 - 23 H. Zeng, R. S. Miller, R. A. Flowers, and B. Gong, *J. Am. Chem. Soc.*, 2000, **122**, 2635–2644.
 - 24 B. GongKim, J. Zhu, and H. Ickes, *J. Am. Chem. Soc.*, 1999, **121**, 5607–5608.
 - 25 H. Zeng, H. Ickes, R. A. Flowers, and B. Gong, *J. Org. Chem.*, 2001, **66**, 3574–3583.
 - 26 H. Ito, Y. Furusho, T. Hasegawa, and E. Yashima, *J. Am. Chem. Soc.*, 2008, **130**, 14008–15.

Appendix I

Fitting analysis

I.1 Fitting from Chapter 3

I.1.1 Titration 1:1 fitting for complex 3.25•3.24 (toluene-*d*₈ 298 K)

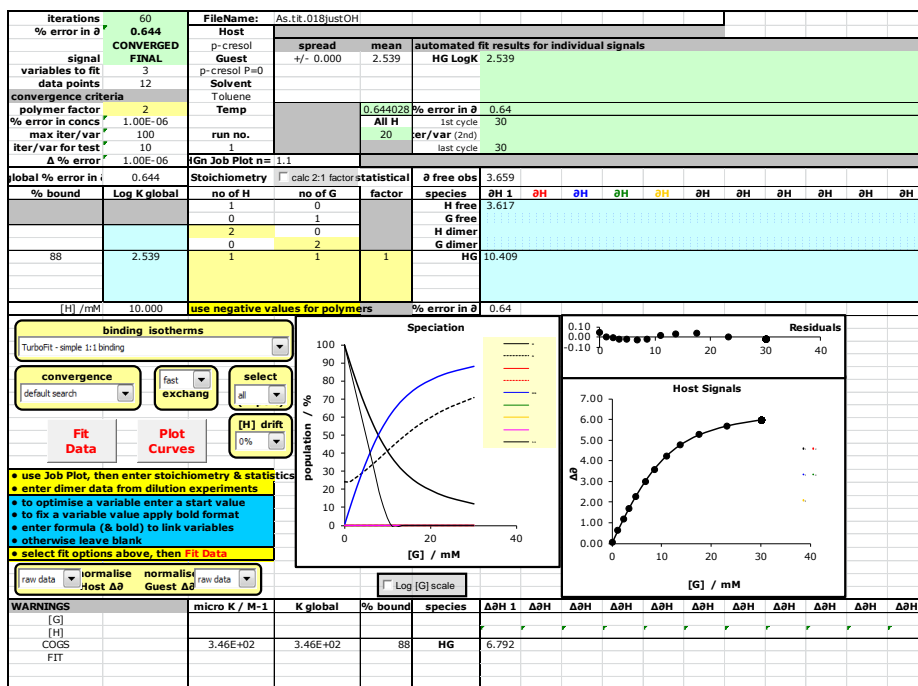


Figure I.1. H = 3.25; G = 3.24.

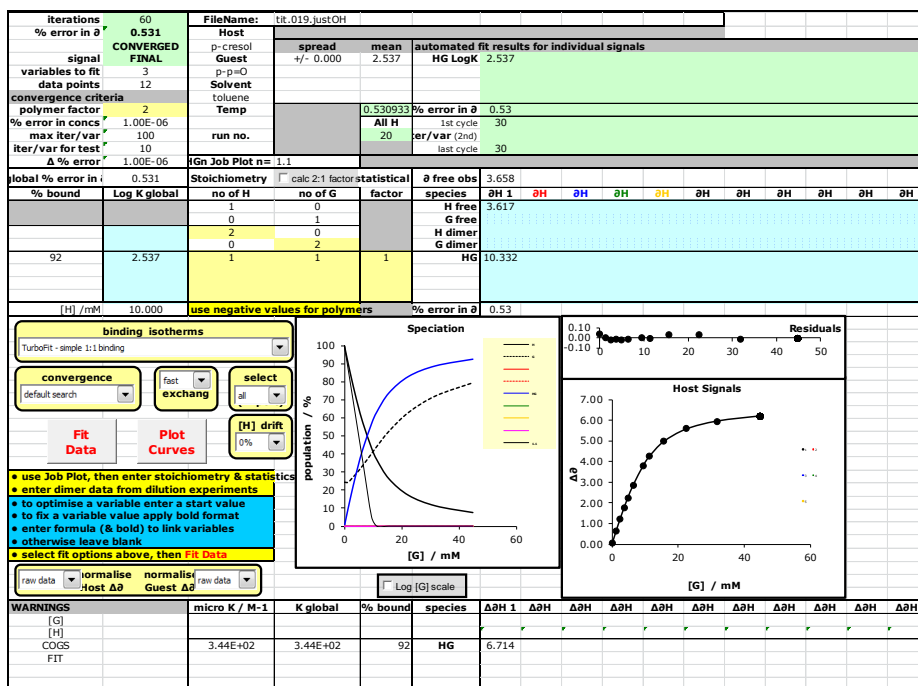


Figure I.2. H = 3.25; G = 3.24.

I.1.2 Titration 1:1 fitting for complex 3.21•3.16 (toluene-*d*₆ 298 K)

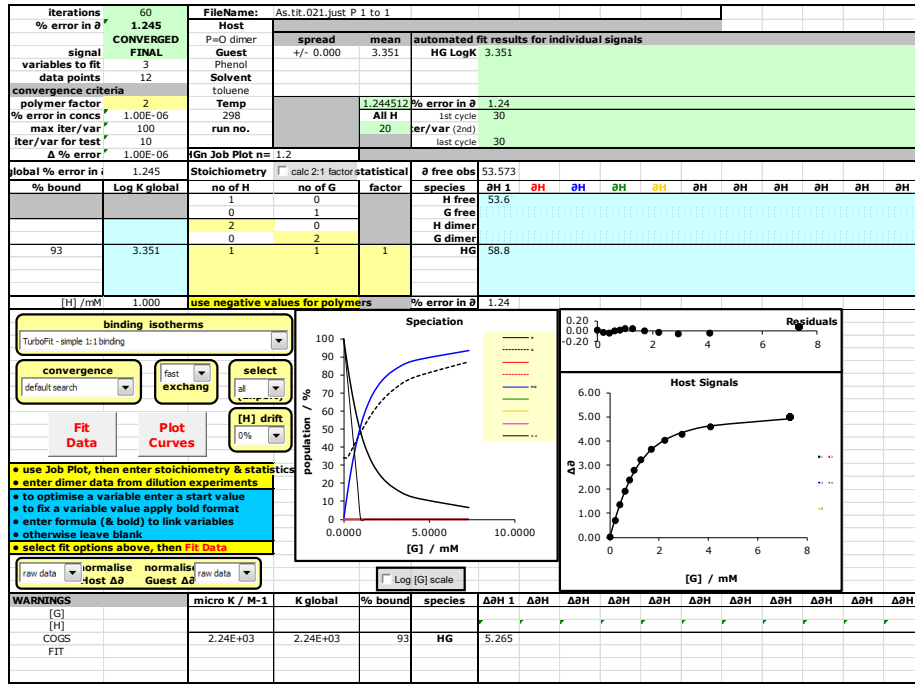


Figure I.3. H = 3.16; G = 3.21.

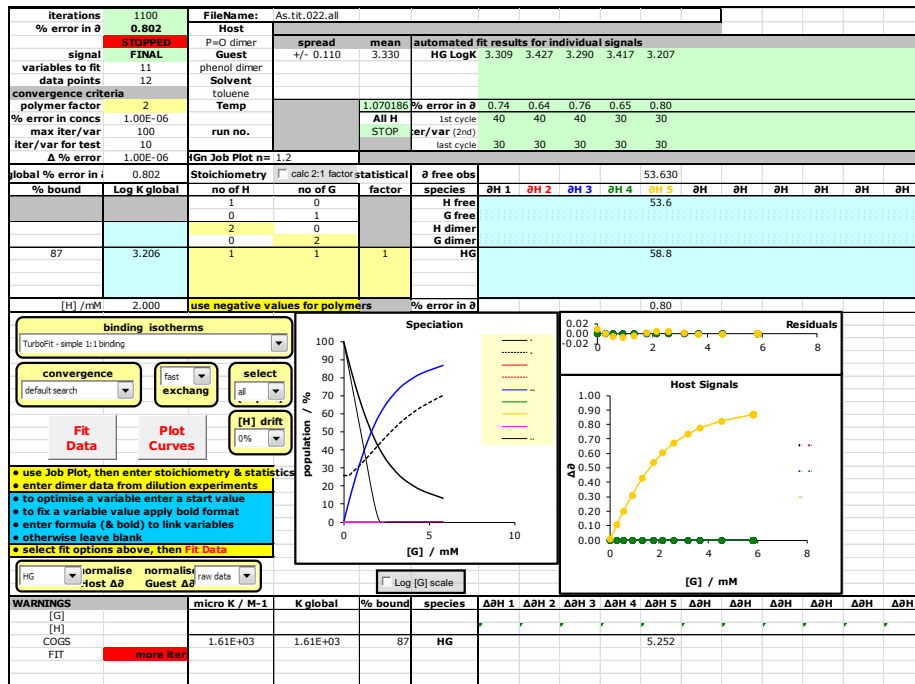


Figure I.4. H = 3.16; G = 3.21.

I.1.3 Titration 1:1 fitting for complex 3.22•3.18 (toluene-*d*₈ 298 K)

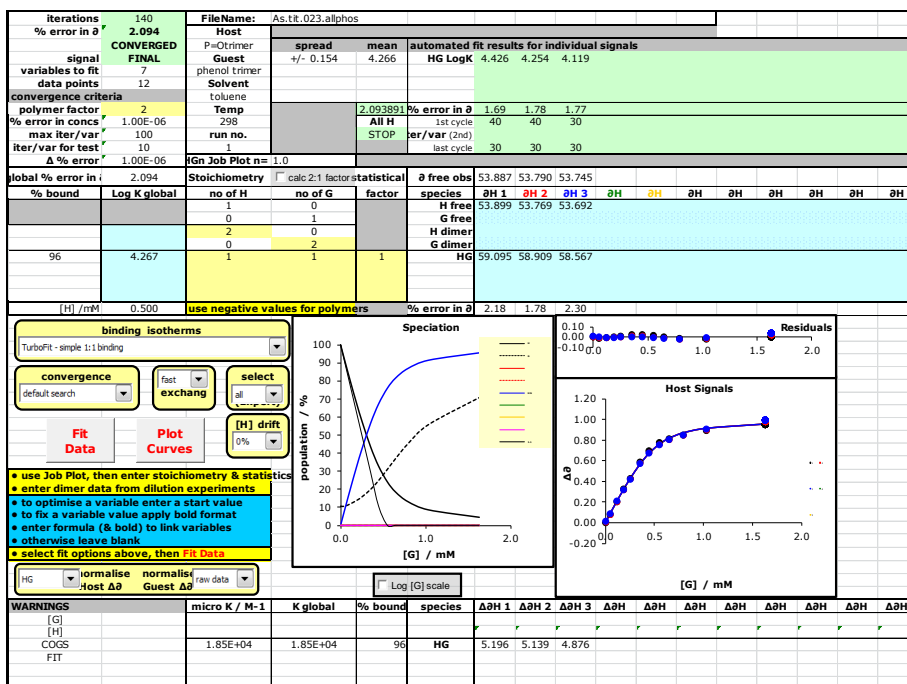


Figure I.5. H = 3.18; G = 3.22.

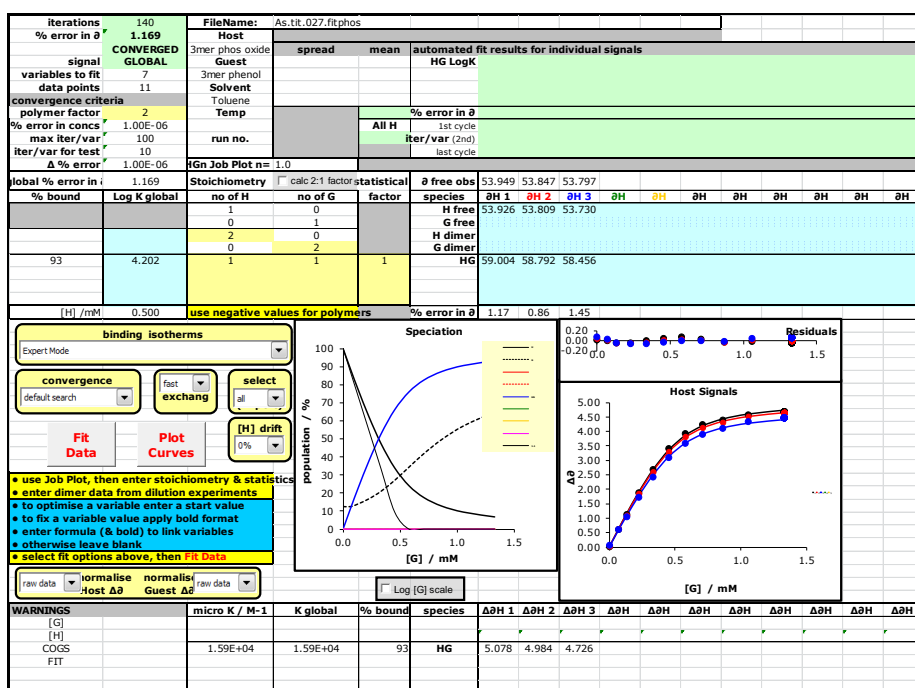


Figure I.6. H = 3.18; G = 3.22.

I.1.4 Titration 1:1 fitting for complex 3.23•3.20 (toluene-*d*₈ 298 K)

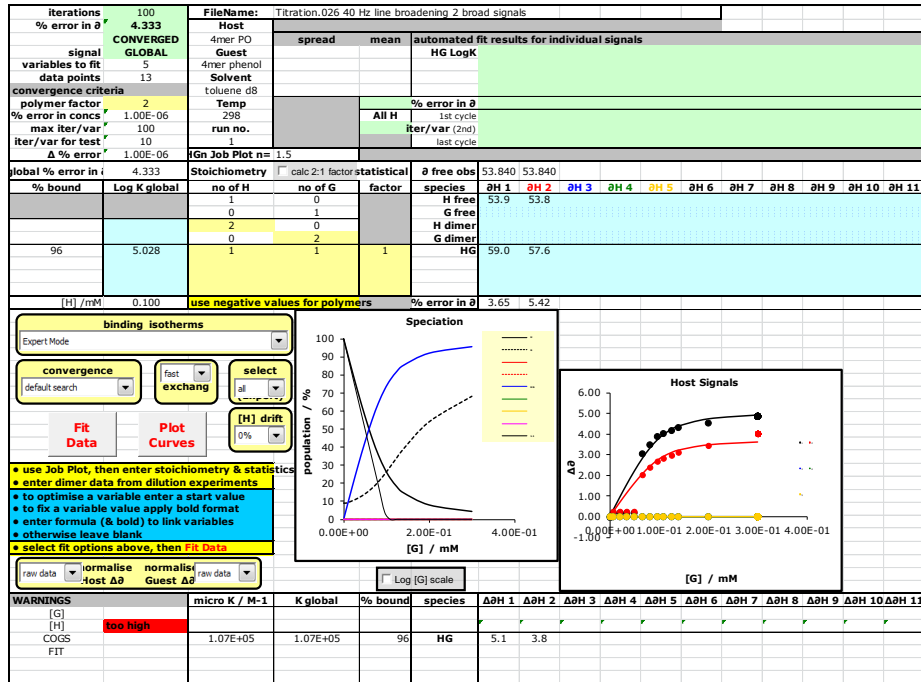


Figure I.7. H = 3.20; G = 3.23.

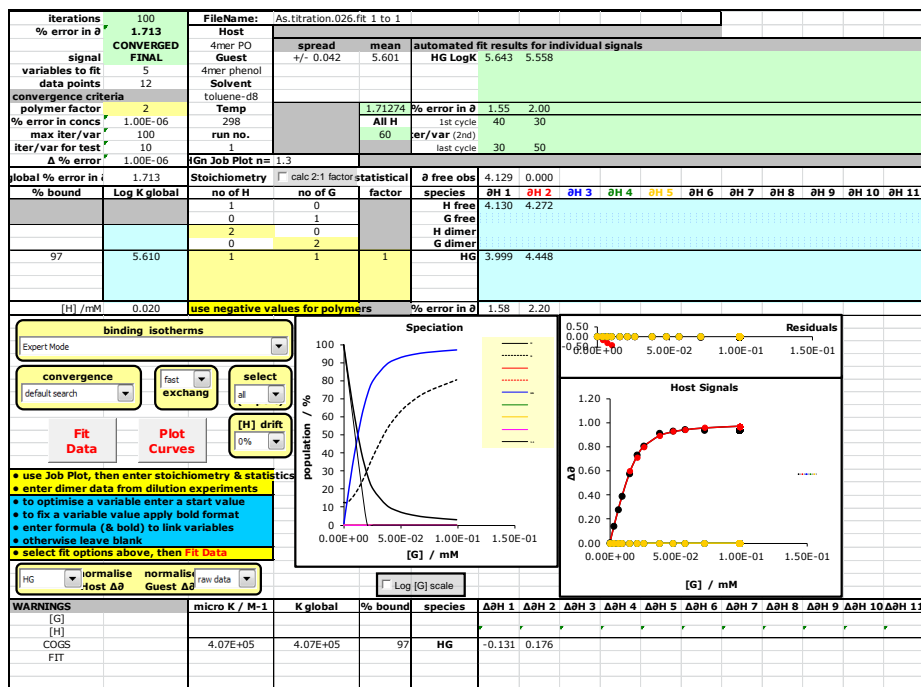


Figure I.8. H = 3.20; G = 3.23.

I.1.5 Dilution dimerisation fitting for 3.27•3.27 (toluene-*d*₈ 298 K)

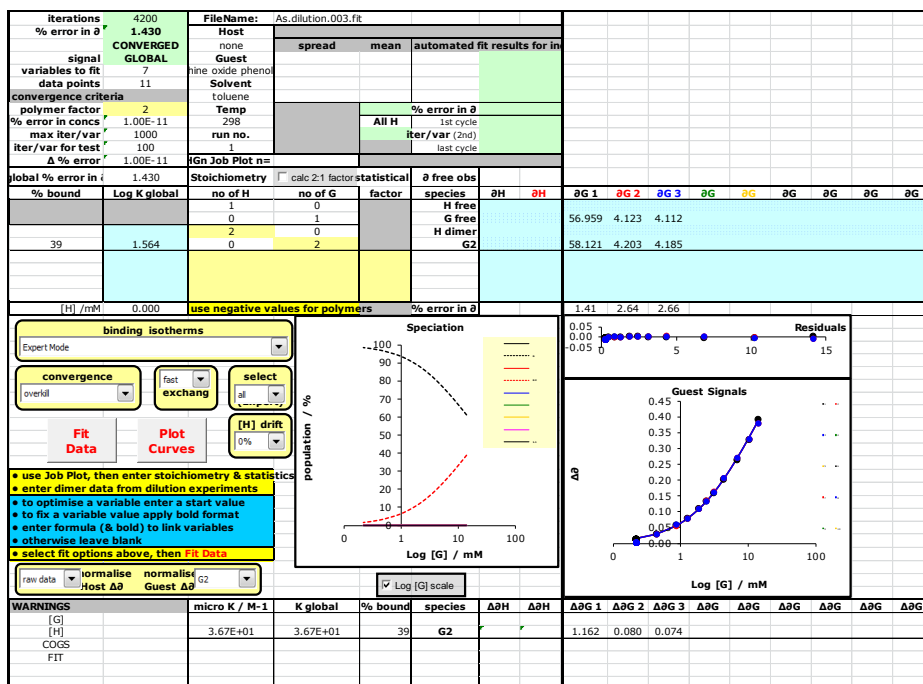


Figure I.9.

I.2 Fitting from Chapter 4

I.2.1 Titration 1:1 fitting for complex 3.25•4.16 (toluene-*d*₈ 298 K)

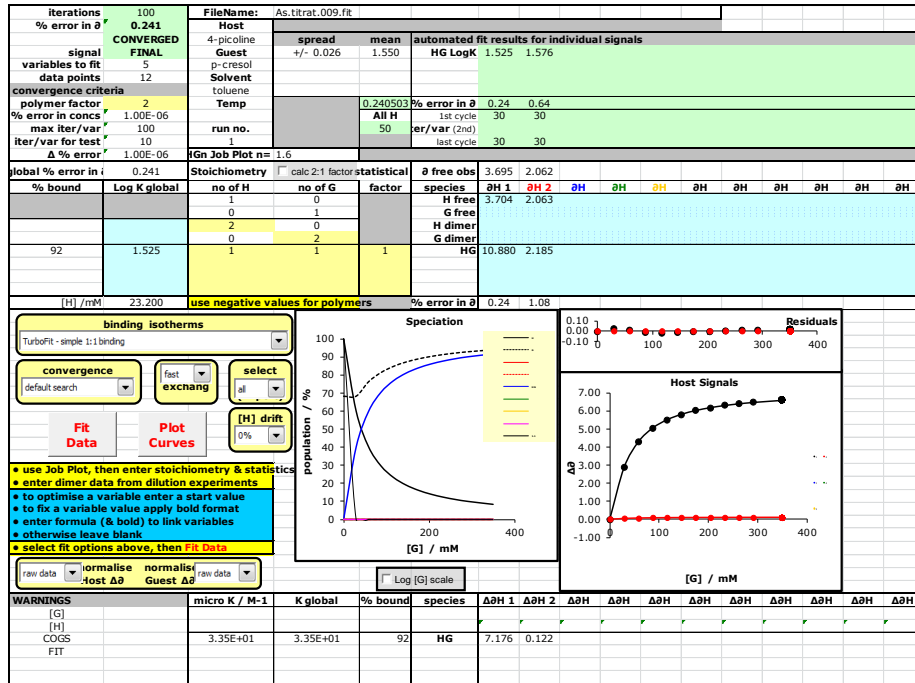


Figure I.10. H = 3.25; G = 4.16.

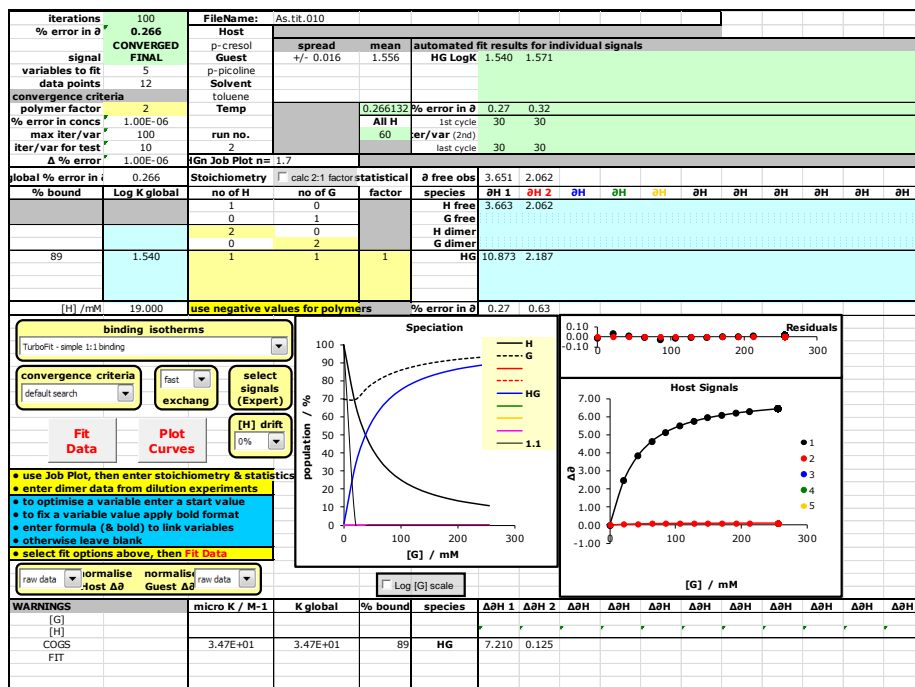


Figure I.11. H = 3.25; G = 4.16.

I.2.2 Titration 1:1 fitting for complex 3.21•4.6 (toluene-*d*₈ 298 K)

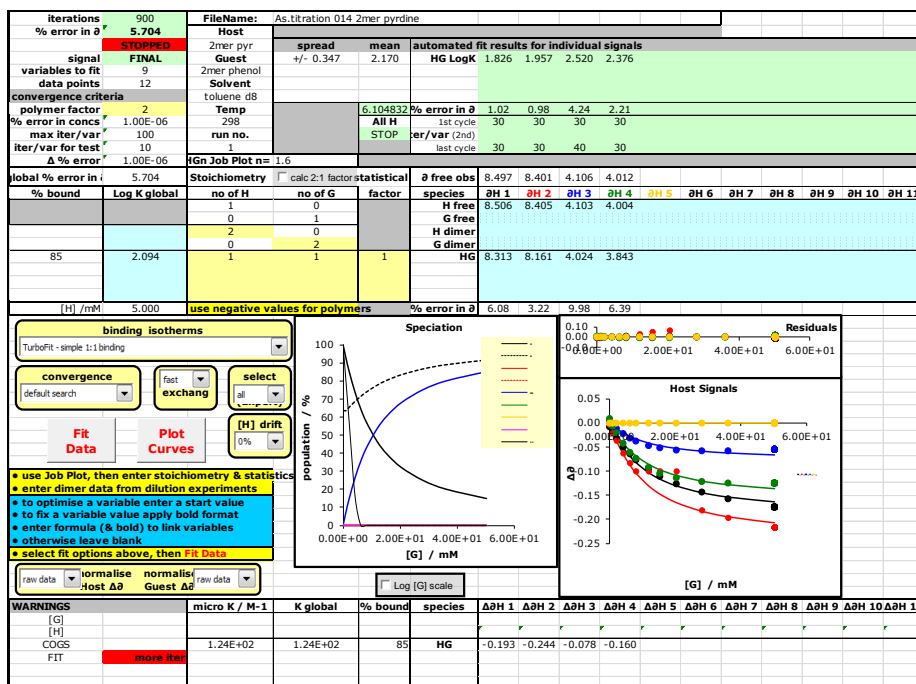


Figure I.12. H = 4.6; G = 3.21.

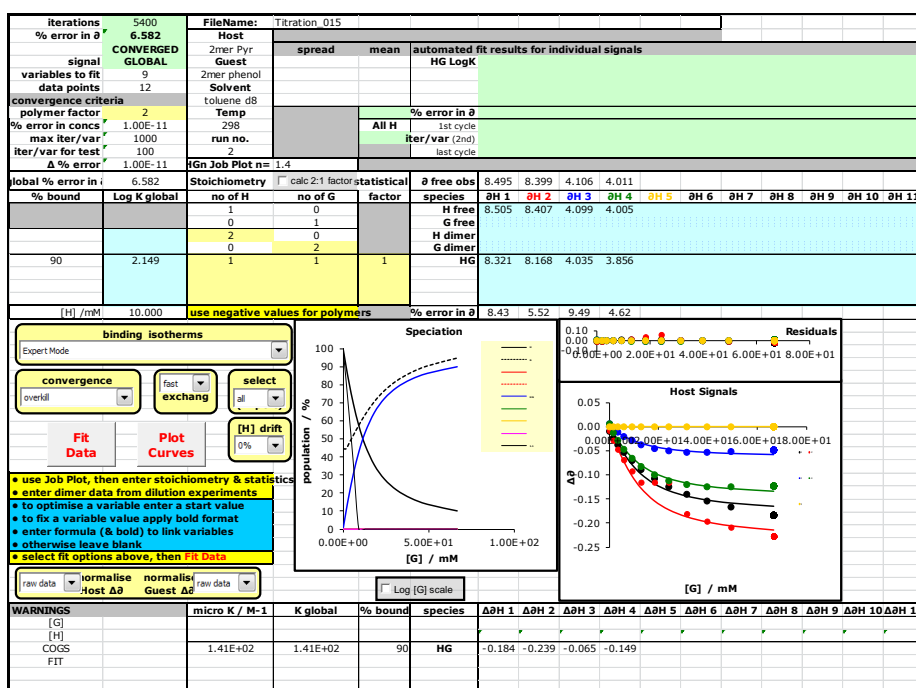


Figure I.13. H = 4.6; G = 3.21.

I.2.3 Titration 1:1 fitting for complex 3.22•4.8 (toluene-*d*₈ 298 K)

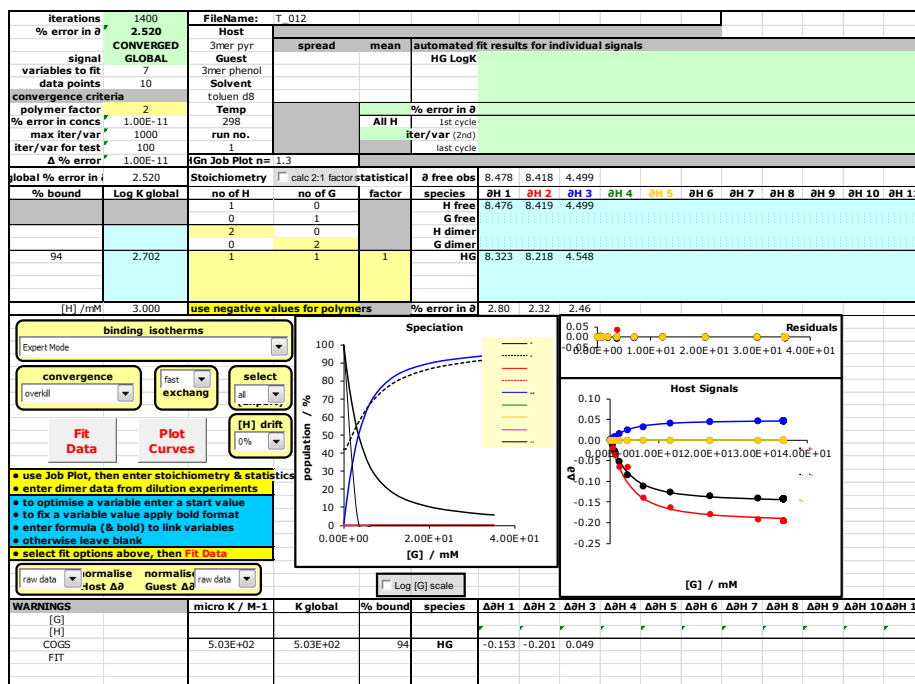


Figure I.14. H = 4.8; G = 3.22.

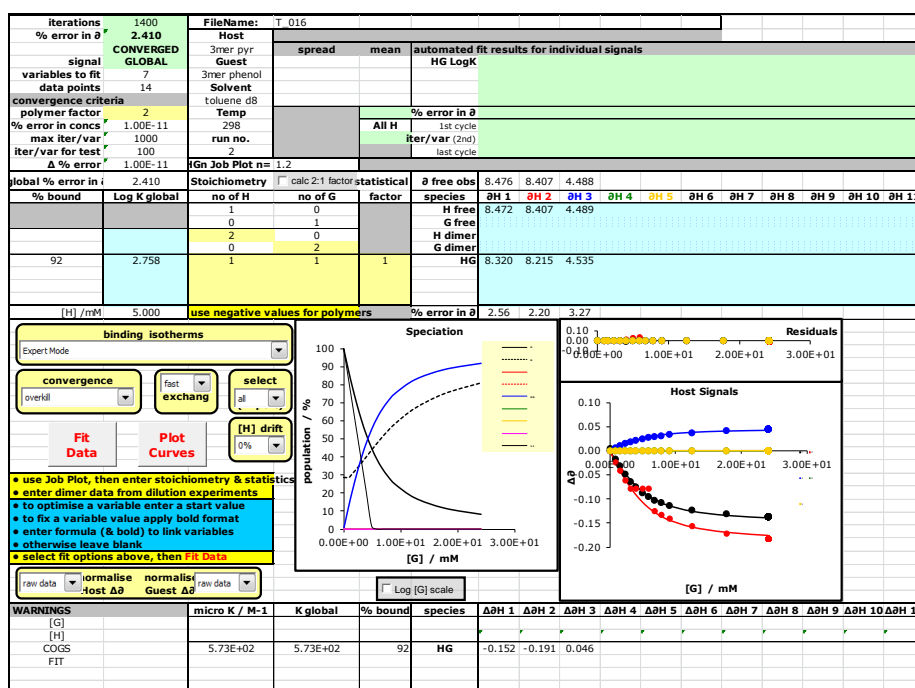


Figure I.15. H = 4.8; G = 3.22.

I.2.4 Titration 1:1 fitting for complex 3.23•4.10 (toluene-*d*₈ 298 K)

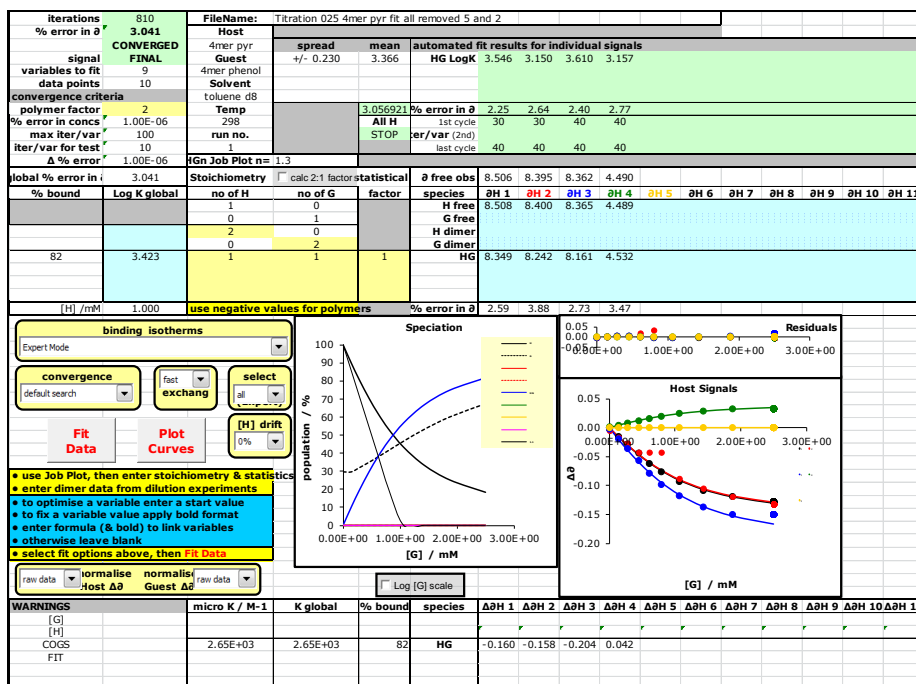


Figure I.16. H = 4.10; G = 3.23.

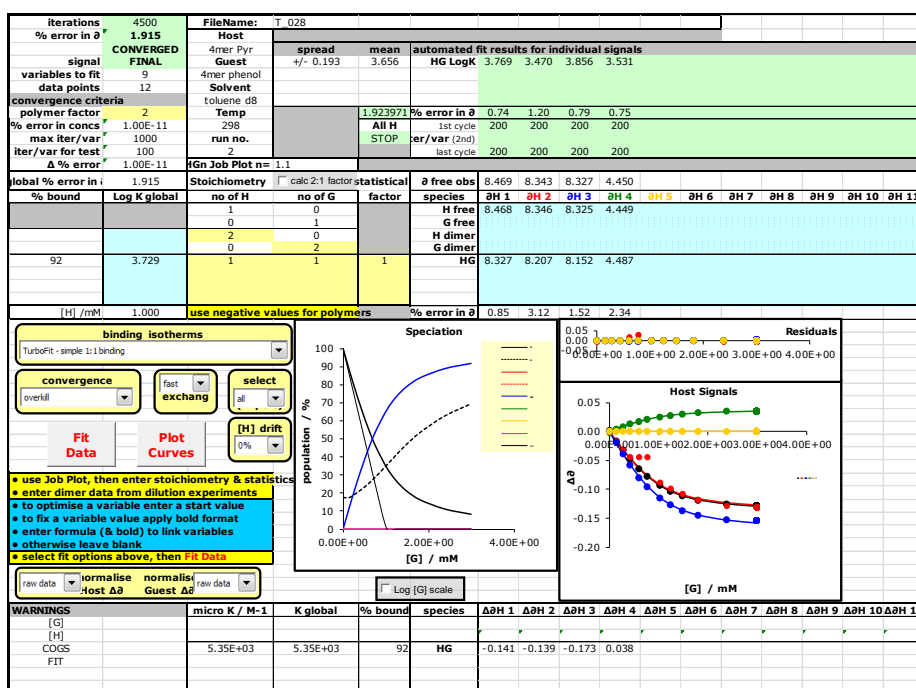


Figure I.17. H = 4.10; G = 3.23.

I.2.5 Titration 1:1 fitting for complex 3.25•4.15 (toluene-*d*₆ 298 K)

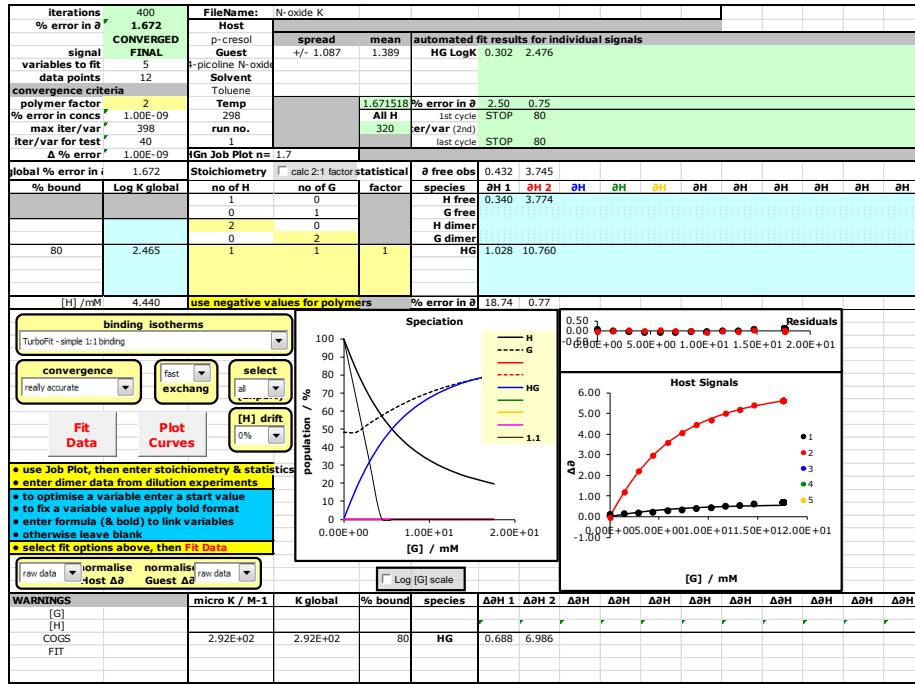


Figure I.18. H = 3.25; G = 4.15.

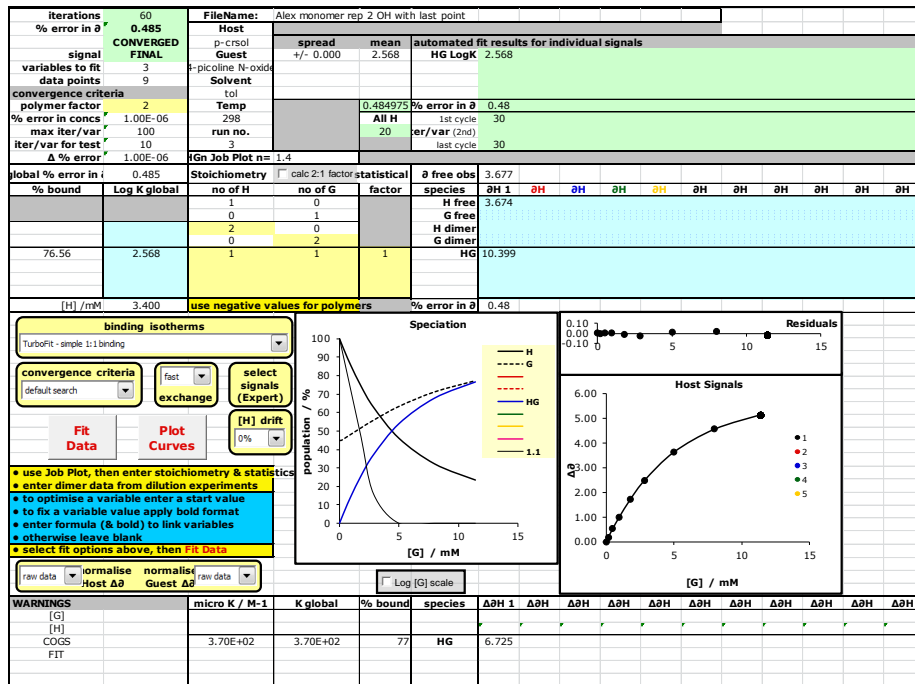


Figure I.19. H = 3.25; G = 4.15.

I.2.6 Titration 1:1 fitting for complex 3.21•4.7 (toluene-*d*₈ 298 K)

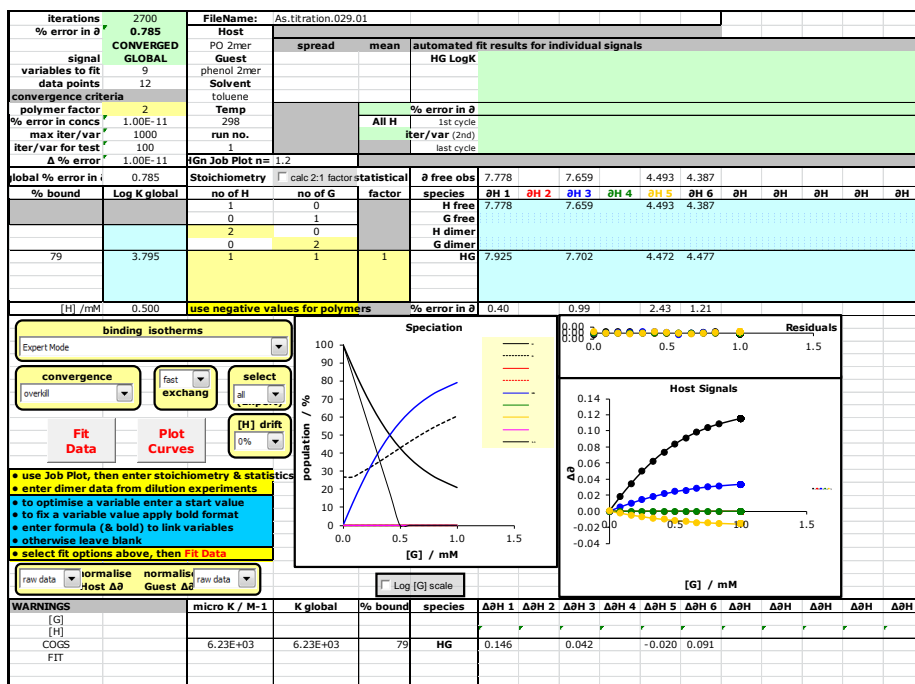


Figure I.20. H = 4.7; G = 3.21.

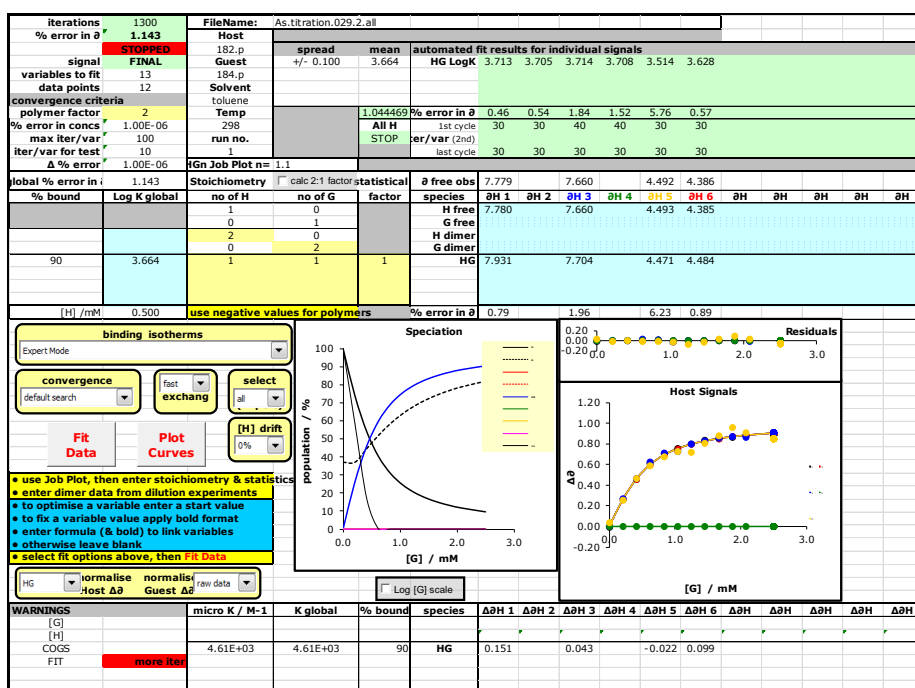


Figure I.21. H = 4.7; G = 3.21.

I.2.7 Titration 1:1 fitting for complex 3.22•4.9 (toluene-*d*₈ 298 K)

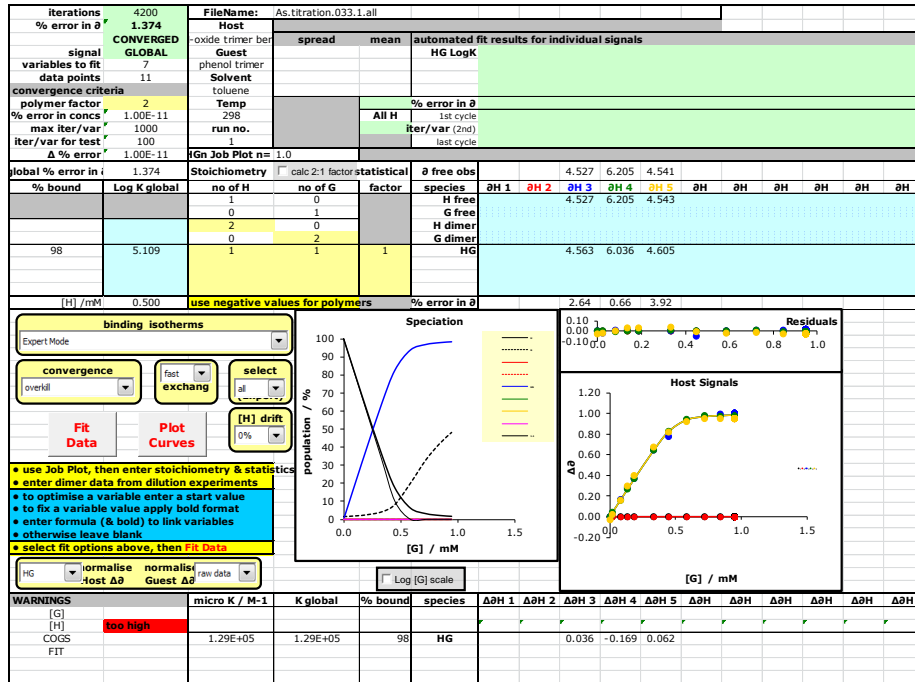


Figure I.22. H = 4.9; G = 3.22.

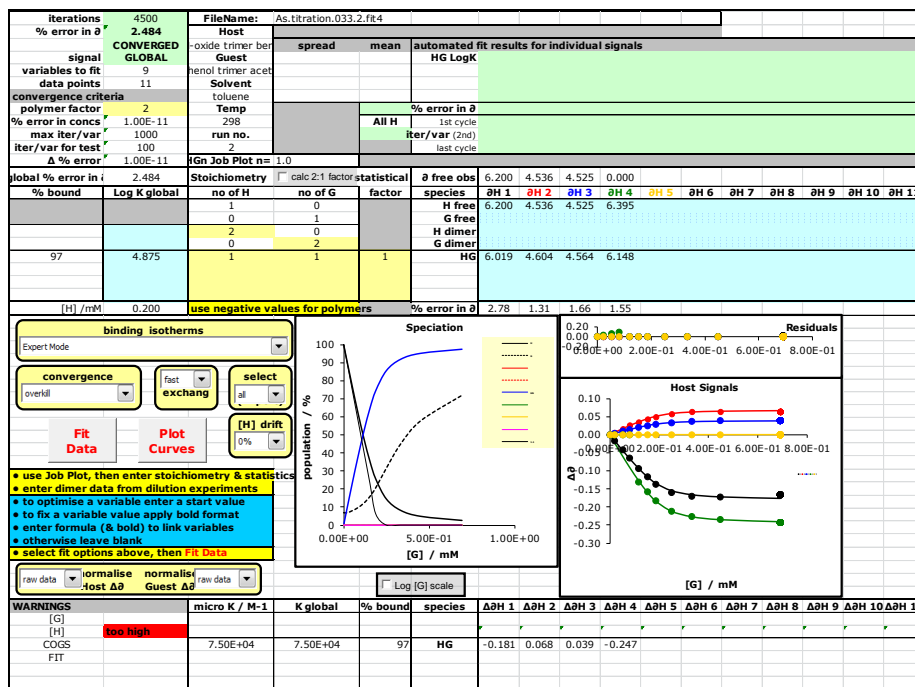


Figure I.23. H = 4.9; G = 3.22.

I.2.8 Dilution dimerisation fitting for 4.12•4.12 (toluene-*d*₈ 298 K)

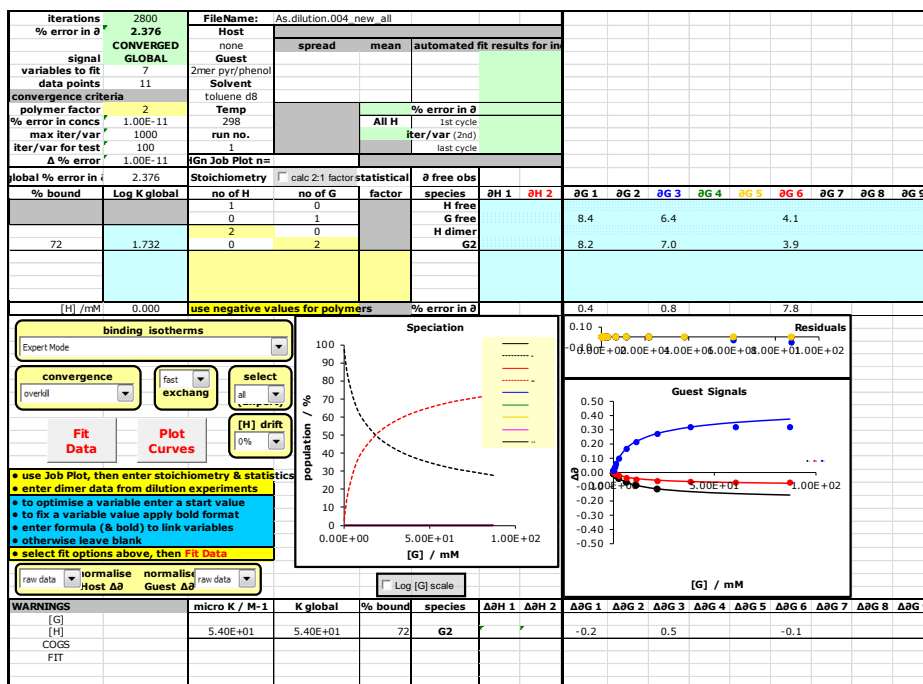


Figure I.24.

I.2.9 Dilution dimerisation fitting for 4.14•4.14 (toluene-*d*₈ 298 K)

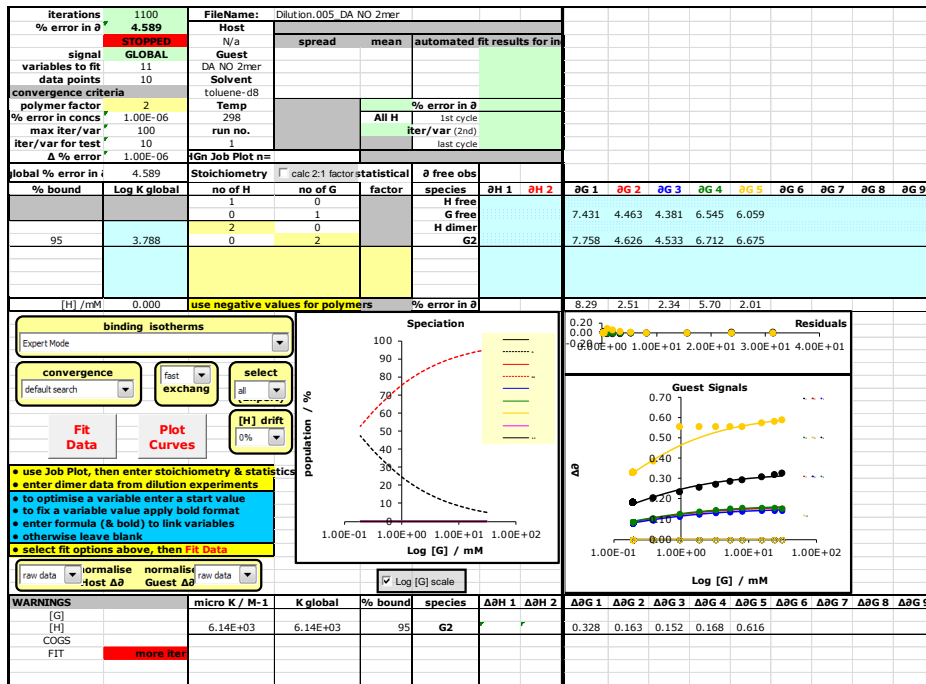


Figure I.25.

Appendix II

Thermal denaturing data

The temperature of the sample was changed using the internal thermostat of the NMR spectrometer, and the sample was allowed to equilibrate in the probe until the probe thermometer gave a stable temperature. The two ^{31}P signals in **3.21•3.16** gave a single overlapping peak that began to split into two signals at lower temperatures, so for high temperatures the chemical shift at the peak maximum was used, and at lower temperatures, the average of the chemical shifts of the two peaks was used (Figure II.1 (b)). Due to significant broadening of the ^{31}P NMR peaks for **3.22•3.18** 40 Hz line broadening was applied to all spectra. At high temperatures where one broad peak was observed, the chemical shift was recorded at the highest point of the peak. At lower temperatures, two of the three separate signals were resolved, so the average of the chemical shifts of the peaks was used (Figure II.1(c)). Even after applying the same processing (40 Hz line broadening) to the spectra of **3.23•3.20** the peaks were too broad at lower temperatures to reliably distinguish from the background noise. (Figure II.1(d)). Figure II.1(e) shows the ^{31}P NMR spectra of phosphine oxide 1-mer **3.24** only (1 mM) at various temperatures.

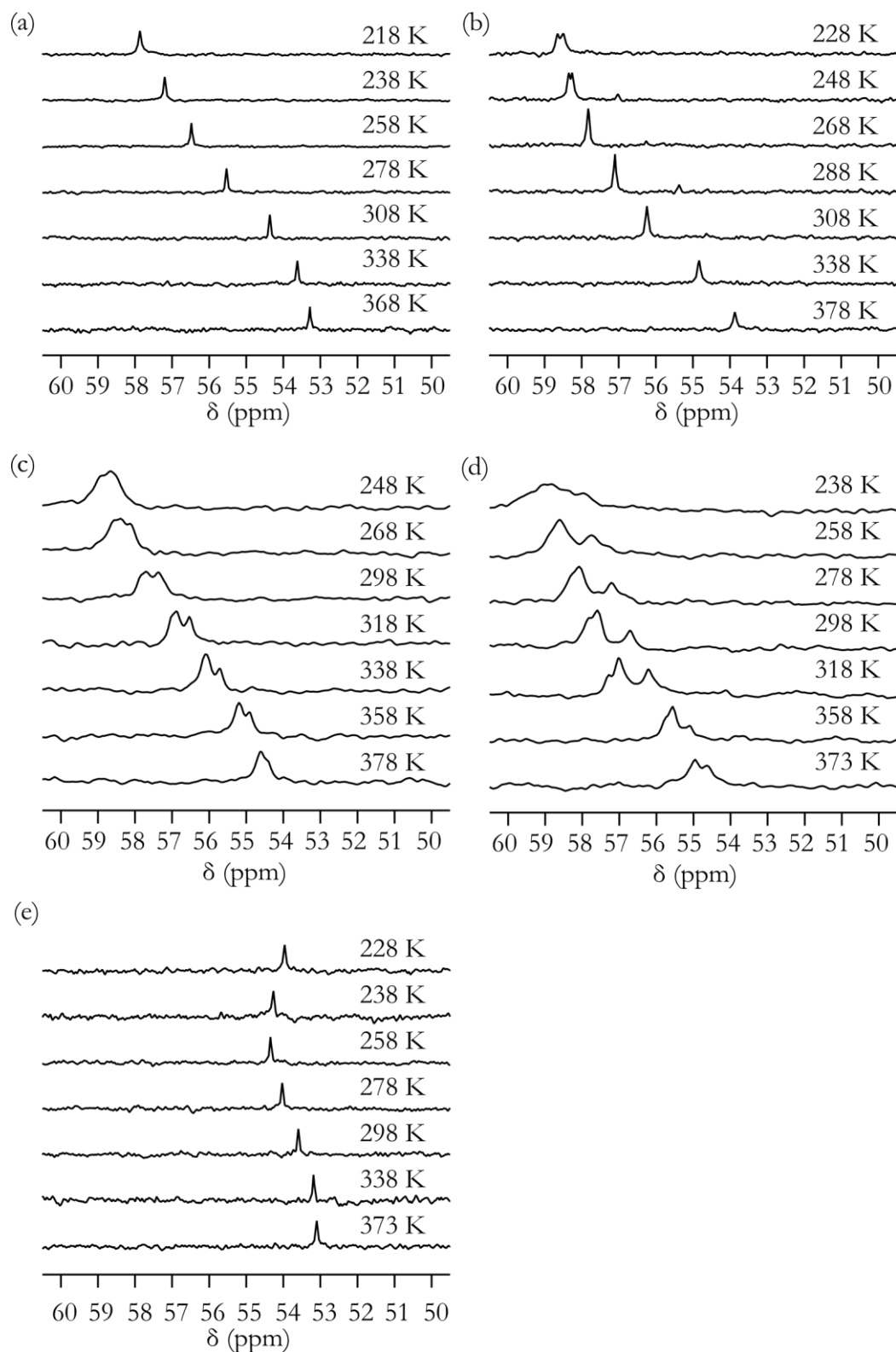


Figure II.1. Variable temperature 162 MHz ^{31}P NMR spectra in toluene- d_8 for 1 mM solutions of 1:1 mixtures of (a) **3.25•3.24**; (b) **3.21•3.16**; (c) **3.22•3.18**; (d) **3.23•3.20**; and (e) only **3.24**. A high line broadening (40 Hz) was applied to the spectra in (c) and (d).