Targeting receptors and DNA secondary structures with small molecules and calix[4]arene conjugates

By

Qiran Sheng

A doctoral thesis submitted in accordance with the requirements of the

degree of Doctoral of Philosophy (PhD) from the University of East

Anglia

School of Pharmacy

December 2016

Declaration

This thesis is submitted to the University of East Anglia for the Degree of Doctor of Philosophy and has not been previously submitted at this or any university assessment or for any other degree. Except where stated, and reference and acknowledgment is given, this work is original and has been carried out by the author alone.

Qiran Sheng

To

My Grandma <u>Gu</u> Zhongxiu

(1926.8.11-2013.9.13)

Acknowledgement

Firstly, I would like to thank my first doctoral supervisor Dr. Zoë A. E. Waller and Dr Susan E. Matthews for their endless support in my research throughout my PhD, especially in my hardest time during the four years. Literally, I feel I am so lucky to have these two excellent personal investigators as my supervisors. In the past four years, they gave me all their patience, encouragement and academic advices in teaching me to be a qualified researcher.

I would like to thank to the other members in the Chemistry and Pharmacy School. Prof Ganesan and Dr Lesley Howell for their advice on my solid phase peptide synthesis; Dr Marco Cominetti and Dr Andrew Beekman for their support on the analytical and preparative-HPLC. Dr Myles Cheesman for the use of circular dichoism. Dr Peter McCormick and Quim for the use of plate reader; Colin Macdonald for the help with the NMR. I would also like to thank my collaborator at the John Innes Centre, Dr Clare Stevenson, in the surface plasmon resonance in *Chapter 2*.

I would like to thank to Waller group members for the past few years. Dr Henry A Day helped me a lot at the beginning of my PhD, teaching me about the chemical synthesis as well as other biophysical techniques; Dr Elise P Wright for her encouragement and support over the past two years. I would like to thank Tasnim Mahmoud and Joe Neavearson for their excellent work in *Chapter 2*. Rouven Becker for his help in preparing the Nile Red precursor in *Chapter 3*. Also I will never forget Mahmoud Abdelhamid, Phil Spence and Phil Gillespie for their support and company in the group. Thank you for appreciating my inappropriate jokes.

I would also like to thank to the members in Ganesan's group for sharing the desk room, CAP 1.01, with me for the past years. These people includes Dr Hanae Benelkebir, Dr Ke Liu, Dr Maria Teresa Bollero, Dr James Buttress, Dr Francesca Kinsey, Remy Narozny, Carys Thomas and Adam. And other members in the School of Pharmacy also

helped me a lot for the past few years. These people includes Oliver Cartwight, Emma Gould, Dr Ryan Tinson, Hui Hui, Noelia, Grace Walden, Vera, Sara and Xin Liao.

Additionally, I would like to thank Elise, Zoe, Mahmoud, Sue and Lavinia for proof-reading my thesis. All your comments and advices were very useful.

In the end, I am grateful to my parents for their support and encouragement during the past four years.

Abstract

This body of research is focused on developing calixarene conjugates targeting i-motif structures and integrin receptors. In *Chapter 1*, a general background of DNA secondary structures and i-motifs was described, mainly focused on the biological relevance, the experimental techniques and known interacting ligands in i-motif studies.

In *Chapter 2*, a high-throughput i-motif ligand screen method was established, based on fluorescent intercalator displacement. Thiazole orange was used as the fluorescent intercalator in the screen against human telomeric i-motif. Its binding was studied using several spectroscopic techniques. A compound library was used to evaluate the newly developed high-throughput screen method using a plate reader and tobramycin was found as the most valuable hit compound in this screen.

In *Chapter 3*, a family of water soluble, DNA-targeting calixarene conjugates were synthesis and characterised. They were functionalised with and DNA-binding moiety on the lower rim. It was found that two of the calixarene conjugates, **28** and **54**, were able to condense G-quadruplex and i-motif forming sequences from human telomere and c-MYC promoter. The calixarene induced condensation was stable under acidic pH, but behave reversible by heating at neutral pH.

Chapter 4 discussed the possibility to develop a calixarene based tumour recognising ligand. In order to achieve tumour targeting, a novel cyclic RGD peptide bearing alkyne was made and tested in the 'click' reaction. Later on, a route to conjugate the novel cyclic RGD peptide with a calixarene tethered with a fluorescent tag was established. It was found that a linker between calixarene and peptide moiety or the copper (I) catalyst was crucial in making this calixarene-peptide conjugate.

Chapter 5 described the experimental procedures used in *Chapter 2, 3* and *4*.

Chapter 6 summarised the key findings in *Chapter 2, 3* and *4*, as well as proposing the future work for all three chapters.

List of Abbervations

Å	Angstrom
Bcl-2	B-cell lymphoma 2
°C	Degree Celsius
CD	Circular Dichroism
c-kit	Mast/stem cell growth factor receptor, CD 117
c-MYC	Cellular myelocytomatosis oncogene
CuAAC	Copper-catalysed alkyne-azide cyclisation
CMC	Critical micelle concentration
DIPEA	*N, N*-Diisopropylethylamine
DCM	Dichloromethane
DMF	*N, N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EtOAc	Ethylacetate
EtOH	Ethanol
FRET	Förster resonance energy transfer
HCl	Hydrochloric acid
HSQC	Heteronuclear single quantum coherence spectroscopy
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HIF-1α	Hypoxia-inducible factor 1α
HIV	Human immunodeficiency virus
HOBt	Hydroxybenzotriazole
hTelo	human telomeric repeat sequence

HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluorophosphate
MALDI	Matrix-assisted laser desorption/ionization
MeOH	Methanol
MeCN(ACN)	Acetonitrile
NaCaco	Sodium cacodylate
NMR	Nuclear magnetic resonance
NRD	Nile red deritives
PDGF	Platelet-derived growth factor
RNA	Ribonucleic acid
SPPS	Solid phase peptide synthesis
SPR	Surface plasmon resonance
TLC	Thin layer chromatography
TFA	Trifluoroacetic acid
UV	Ultraviolet

List of Figures

Chapter 4 Synthesis of cRGD-Conjugated Calix[4]arenes

List of Schemes

Chapter 4 Synthesis of cRGD-Conjugated Calix[4]arenes

Chapter 5 Overview and Future work

List of Tables

Table of Contents

Chapter 1   General Introduction

1.1   Oligonucleotides and DNA secondary structures

Parents pass their heritable genetic information to their offspring through DNA. This biopolymer, along with its ribonucleic acid analogue is crucial to all living cellular organisms. They distinguish life from other chemical or physical processes. On a cellular level, they serve as a 'guide book' to cells for the synthesis of proteins, orchestrating the process of cell cycle and reproduction.[1] In 1869, Friedrich Miescher first discovered and isolated a mysterious phosphorous-rich material, later called nuclein, from the nuclei of white blood cells; this marked the beginning of oligonucleotide chemistry.[2, 3] In the late 1940s, structures of basic monomer nucleosides and their corresponding nucleobases in DNA & RNA (adenine, guanine, thymine, uracil and cytosine) were discovered.[1] In the year 1953, two Cambridge scientists, Francis Crick and James D. Waston, presented a milestone paper proposing the secondary structure of B-form DNA based on Rosalind Franklin's X-ray diffraction images.[4] This model showed two complementary DNA strands parallel paired *via* classical 'Watson-Crick' hydrogen bonding (adenine to thymine and cytosine to guanine, Figure 1.1 a) resulting in the iconic right-handed B-form DNA 'double helix' (Figure 1.1 b). [5]
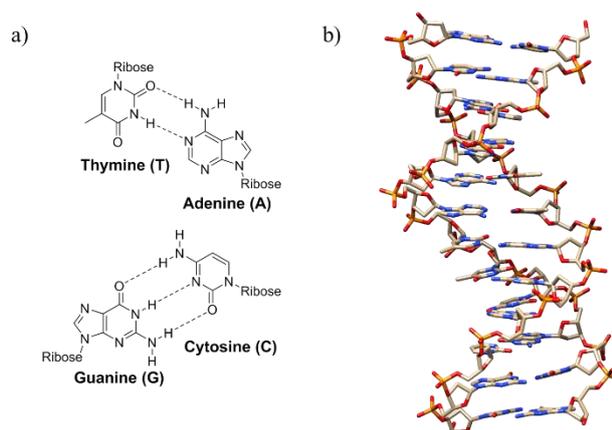
Figure 1.1   *a) Watson & Crick base paring (Adenine to Thymine & Guanine to Cytosine) and b) The structure of B-DNA (X- ray diffraction, PDB ID: 1BNA [6])*

However, the subsequent 60 years of research has revealed that the conventional duplex is not the only form of DNA structure. Depending on the sequences, non-Watson Crick base pairing, surrounding conditions and other factors, non-B DNA secondary structures are also possible as well.[7] Known DNA secondary structures include B-, A-and Z-form duplex, triplex, G-quadruplexes, i-motif, A-motif, cruciform and Holliday junctions (Figure 1.2).

G-quadruplexes **i-motif** **Hairpin**

Triplex DNA **Holliday junction**
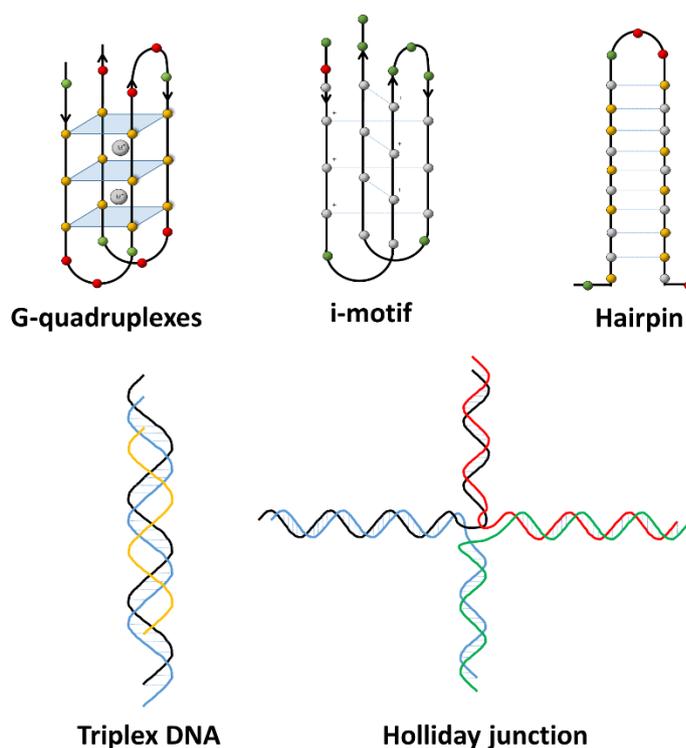
Figure 1.2   Examples of some alternative DNA secondary structures

During DNA transcription, recombination and replication, several alternative DNA secondary structures have been discovered with specific biological functions.[7] Evidence has also shown that genomic instability induced by non-canonical DNA structure formation not only contributes to the predisposition to disease,[8,9] but to rapid evolutionary changes.[10] Since scientists are interested in their molecular mechanisms in cells, non-canonical DNA secondary structures have been extensively studied. Moreover, scientists have found that these non-canonical DNA structures are ideal materials for applications in bio-related nanotechnology or drug

vectors due to their safety and robustness in physiological conditions.[11, 12] Thus the study of DNA secondary structures has become a popular branch of chemical biology, driving more and more researchers and resources into this field.

1.2   G-quadruplexes

G-quadruplexes are a family of well-characterized oligonucleotide secondary structures which have been intensively investigated over the past twenty five years.[13] G-quadruplexes arise from guanine-rich nucleic acid sequences, in both DNA and RNA，and have a 3D structure composed of at least two or three planar arrangements called G-tetrads. The first published example of a G-tetrad had 4 guanine bases joined *via* eight Hoogsteen hydrogen bonds in a clockwise or anti-clockwise orientation (N_1-H to N_7 & N_2-H to O_6 atoms, Figure 1.3 a).[14] Two or more tandem layers of G-tetrads can be assembled as a G-quadruplex structure stabilised by monovalent cations (Figure 1.3 b & c).[13] Thus the strong negative-charged electrostatics at the centre of the G-tetrad created by the lone pair of O_6 atoms of each guanine is partially neutralized, further enhancing the rigidity of the structure. As a result, G-quadruplexes formed from some sequences possess higher stability when compared to their conventional Watson-Crick double helical counterparts under physiological conditions.[13]
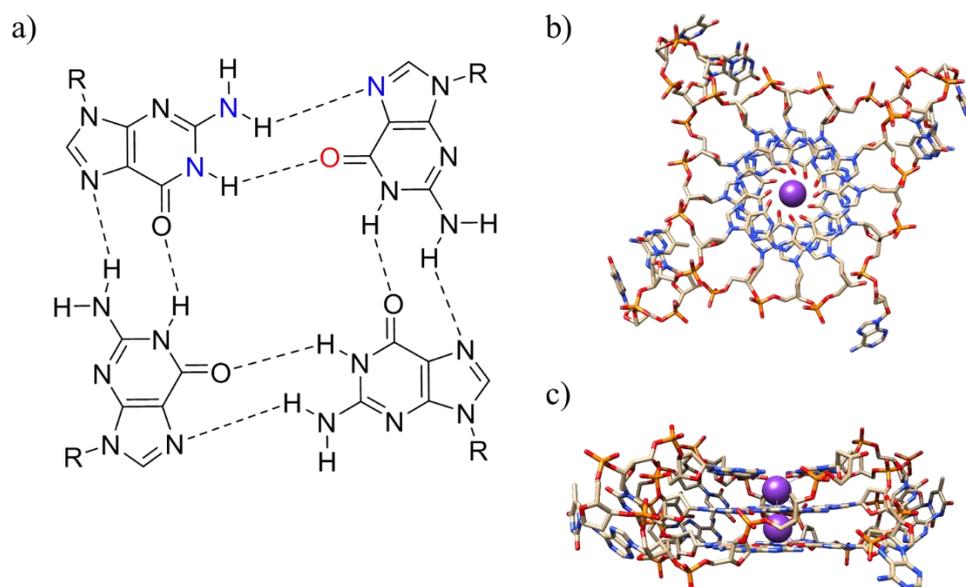
Figure 1.3   *a) Planar G-tetrad held together by Hoogsteen hydrogen bonds. b) A top view of human telomeric G-quadruplex stabilised by K^+ (X-ray diffraction, PDB ID: 3T5E). c) A side view of human telomeric G-quadruplex stabilised by K^+ (X-ray diffraction, PDB ID: 3T5E [15]).*

How relevant are G-quadruplexes to living organisms? Although G-quadruplexes were first shown in the 1960s,[14] they did not receive a lot attention until the 1990s after some G-quadruplexes were discovered in living prokaryotic organisms.[13] In the context of the human genome, over 360,000 estimated G-quadruplex forming sequences have been discovered *in silico*,[16] and some of them are located at critical regions of the genome. G-quadruplexes have been found not only to exist in human cells,[17,18] but also participate in key biological processes, including telomere maintenance, transcription and replication.[19-23] In 2016, with the help of high-throughput DNA sequencing methodology and a G-quadruplex targeting antibody, about 10,560 G-quadruplex structures were found in human chromatin, mostly in regulatory, nucleosome-depleted regions.[24] When considering gene promoter regions, G-quadruplex interacting ligands have been shown to interfere with

transcription of the respective gene.[25,26] These reports all indicate undiscovered transcriptional functions of G-quadruplexes in living organisms. Thus G-quadruplexes can be utilised as a potential druggable target for cancer diagnosis and intervention. Apart from anti-cancer therapeutics, G-quadruplexes have also been found in other gene sequences associated with viral infections such as human papilloma virus, HIV and even diabetes.[27-29]

1.3   i-Motifs

G-quadruplexes arise from guanine-rich sequences, however, in the context of the human genome, what will happen to the complementary cytosine-rich sequences? These sequences may form a secondary structure called the intercalated motif, or i-motif. i-Motifs are quite different from G-quadruplexes, especially in terms of their structural arrangement as they are the only known nucleic acid secondary structure held together by systematic base intercalation. Two parallel-stranded hemiprotonated $C \equiv C^+$ base pairs intercalate and lock themselves in an antiparallel manner with the help of Hoogsteen hydrogen bonds, forming a compact structure stable under slightly acidic conditions (Figure 1.4 a).[30] It was not until 1993 that i-motif structures were first resolved by Gehring *et al.* in studying an four stranded d(TCCCC) oligomer structure which possess the intercalated base-pairing. [31,32] As shown in Figure 1.4 b, i-motifs possess two major and two minor grooves, the distance between two paired cytosines is usually only a few Å.[33]

a)

b)
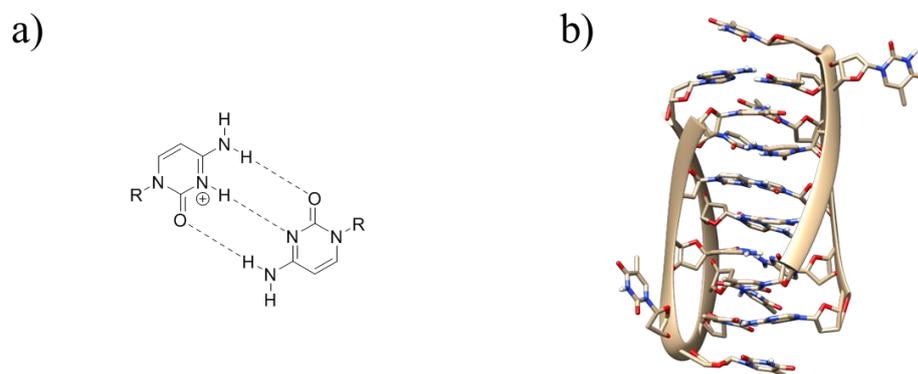
Figure 1.4   *a)* C≡C$^+$ *base pairing b) A side view of dimeric i-motif structure from centromeric alpha-satellite DNA (solution NMR, PDB ID: 2MRZ [34])*

For i-motifs, the slightly acidic conditions which favour folding have paved their way to be employed extensively as pH-sensitive switches in various DNA nanotechnologies[35-38] but this has also provoked scepticism in their potential to play physiological roles in cells. However, this is changing, i-motif forming sequences that can fold at near neutral pH conditions have been found.[39] Accumulating evidence in recent years has indicated that the i-motif may exist in the cell as well.[40] Thus, the potential biological roles of i-motifs are beginning to attract people's attention.

1.3.1   Polymorphism and stability of i-motifs

Like the G-quadruplex, the i-motif can also form intermolecular species such as dimers and tetramers where multiple strands are involved in the structure. The first known i-motif structure was a tetramer (Figure 1.5a). In 1994, Gehring, Leroy and Guron resolved the first tetramer i-motif structure arising from four d(TC$_5$)-strands *via* NMR and gel filtration chromatography.[30] As shown in Figure 1.5 b, when two cytosine-rich hairpins are intercalated, a dimer i-motif is formed. This case was seen with a cytosine-rich dodeca satellite strand in an endogenous *Drosophila*

centromere (Figure 1.4 a).[34] Other putative biologically relevant cytosine-rich sequences exist in other biologically relevant regions, such as in the human telomeric repeat region and sequences within the promoter regions of HIF-1α,[39] c-MYC,[41] and Bcl-2.[42] These are able to form intramolecular i-motif structures *via* folding a single strand with four cytosine stretches back upon itself (Figure 1.5 c).[40]
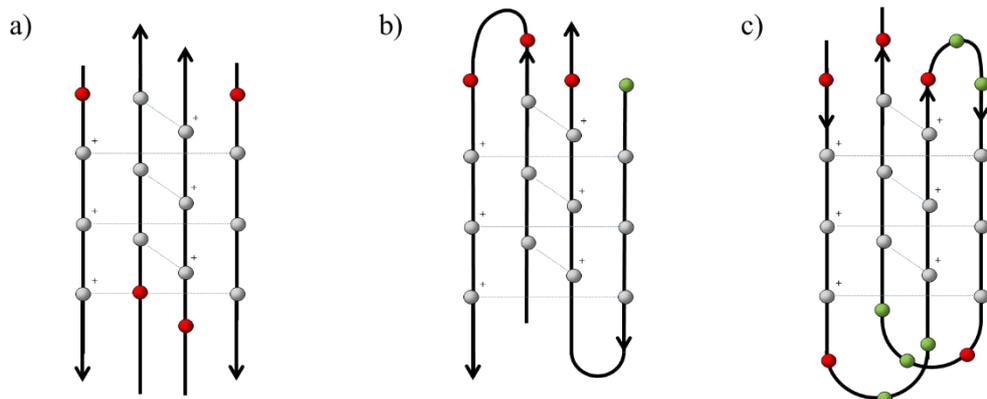
Figure 1.5   *Schematic of different types of i-motif a) Tetramer b) Dimer c)*

Intramolecular

Originally it was suggested that intramolecular i-motifs can be categorized into two classes based on their loop size. Smaller loop-sized (2 to 4 nucleobases on the loops) were denoted 'Class I' and larger loop-sized (over 4 nucleobases on the loops) i-classified as 'Class II'.[43] Initially, it was considered that 'Class II' i-motifs were more stable than their 'Class I' counterparts, since they had a higher chance to gain extra stability *via* interactions between bases in the loops.[43] However, recent years have given rise to a debate on whether stability is in fact related to loop length.[44,45] For example, in the case of the human telomeric repeat region, bases in the loop (TAA) not only form extra A-T base paring, but they stack back on the i-motif's main cytosine core as well however these are classified as 'Class I' i-motifs. This, coupled with the debate in the literature, suggests the classification of i-motif structures based on the loops was premature.

In addition to loop length, other factors, such as the number of cytosine bases, the presence of cations, buffer conditions, and functionalites conjugated to

oligonucleotides, can influence the stability of i-motifs.[40] More $C\equiv C^+$ base pairs present in i-motifs provide more hydrogen bonding, thus enhancing the stability of i-motif structures.[40] In terms of cations, it is clearly shown that presence of Li^+ destabilises human telomeric i-motif.[46] The size of Li^+ is believed to be the reason. Lithium cations possess a radius around 0.7 Å, which facilitate themselves fitting into the pockets in $C\equiv C^+$ base paring,[46] but Li^+ does not have the ability to mimic the quantum mechanical effects of a proton on the hydrogen bonding. So the original packed i-motif structure is disrupted. Previous studies in the Waller group have shown that Ag^+ is able to fold the human telomeric i-motif forming sequence into an i-motif even at pH 7.4, and that this Ag^+ induced i-motif formation is reversible with the addition of cysteine, a chelating agent for silver cations.[47] The mechanism of this interaction is believed to be an insertion of Ag^+ into the $C\equiv C^+$ base pair where it mediates the mismatch in hydrogen-bonding.[48] Another study centred around cations showed recently that Cu^{2+} is able to convert human telomeric i-motif into a hairpin structure, even in acidic conditions.[49]

Unlike G-quadruplexes, where RNA analogues are always more stable than the same sequence in DNA, intramolecular i-motifs formed using RNA sequences are not as stable.[50] This is due to the 2′-OH group on the ribose which sits in the narrow groove of the i-motif, resulting in structural expansion and destabilisation. However, some examples of DNA/RNA hybridised oligonucleotides are able to form i-motifs, but their structural rigidity is not comparable with intramolecular i-motifs formed purely by DNA.[32,51]

1.3.2   Biological relevance of i-motifs

The existence of G-quadruplexes and i-motifs *in vivo* has been constantly questioned ever since their discovery. As estimated from *in silico* experiments,[52] 43% of gene promoters and 69% of oncogene promoters contain guanine/cytosine-rich sequences, capable of forming DNA secondary structures including promoters

of genes involved in cell proliferation, including c-MYC,[41,53,54] PDGF-A,[39,43] pRb,[55] Bcl-2,[42,56-58] K-ras[43,59] and VEGF.[40,52] In 2013, Balasubramanian and co-workers published the first report of imaging G-quadruplex structures in cells *via* immunofluorescence technology.[17]

In contrast, so far there is no direct evidence showing the existence of i-motif in cells. Nevertheless, there is data that supports the existence of i-motifs in biology. Negative supercoiling is a phenomenon introduced by unwinding the DNA double helix during transcription. This phenomenon creates negative superhelical stress on double-stranded DNA which is believed to be relieved by formation of i-motif and G-quadruplex structures.[60] Hurley's group investigated the effect of negative superhelicity *via* incorporating the G/C-rich sequences from c-MYC promoter, which can form i-motif and G-quadruplex, into a supercoiled plasmid. When the plasmid is transcribed, artificial negative superhelicity is created. The result indicated that c-MYC i-motif forms under neutral pH, alongside G-quadruplex formation.[60] Apart from negative supercoiling, molecular crowding conditions are another supporting factor for the formation of i-motif under physiological conditions.[61] By simulating intracellular conditions utilising high molecular weight polyethylene glycol polymers, such as 20% PEG as a cosolute, it was shown that the formation of G-quadruplexes and i-motif was more favoured over duplex or single stranded DNA.[62,63] Under these conditions, modestly stable i-motif structures ($T_m \approx 20\,^{\circ}\mathrm{C}$) were observed even at neutral pH as a result of the increased pK_a of the C≡C$^+$ base pairs.[64]

Telomeres

The telomere regions of chromosomes are one of the most intensively studied regions in the human genome, featuring several hundreds to thousands of tandem repeat sequences at each end of a eukaryotic chromatid.[65] Telomeres are utilised by eukaryotes in maintaining genome integrity and avoiding loss of genetic

information. A telomere consists of 3 to 15 kilobases of tandem $5'$-d(TTAGGG)$_n$-$3'$-repeats (or $5'$-d(AATCCC)$_n$-$3'$ on the cytosine-rich sequences), with a 20- to 150-base-long single-stranded $3'$-overhang at the terminus, and associated capping proteins.[66] Telomeres are truncated as cells divide to a senescent stage when apoptosis is triggered following shortening of telomeric repeats, serving as an indicator of cell age.

Telomeric repeats can be replenished *via* telomerase, an enzyme found only in stem cells, activated lymphocytes and cancer cells.[67] Telomerase immortalises cells and it has been found that about 85% of cancer cells over-express this critical enzyme.[68,69] Thus, reducing the length or otherwise interfering with the synthesis of the telomeric repeats can promote cell apoptosis, with potential applications as an anti-cancer intervention. It has already been shown that G-quadruplex interacting compounds inhibit telomerase activity with a high efficacy.[70,71] Stabilising the i-motif forming sequence in the human telomere using carboxyl-modified single-walled carbon nanotubes (SWNTs) has been shown to inhibit telomerase, followed by telomere uncapping and cell apoptosis.[72] Proteins that bind the telomeric i-motif forming sequences have also been found,[73,74] but whether they bind i-motif structure specifically or just single-stranded DNA is a question that needs further exploration.

Oncogenic Promoters

In addition to the telomeres, non-canonical DNA structures are prevalent in oncogene promoter regions. Several oncogenes whci contain i-motif forming sequences have been studied, including c-MYC,[75,76] Bcl-2,[42,56] VEGF[77] and c-kit.[78,79]

About 90% of c-MYC transcription is controlled by a highly conserved sequence in the P1 promoter called nuclease hypersensitive element (NHE) III$_1$. This consists of a 27 base-pair sequence which is able to form both G-quadruplex and i-motif.[75,76]

In *in vitro* experiments, c-MYC i-motif possesses a transitional pH of 6.6 [43] (pH_T; a pH value where 50% of the DNA population is folded into an i-motif structure), which is relatively close to neutral conditions, It adopts a 5:5:5 loop topology with 8 hemi-protonated base pairs.[76] However, as mentioned in Section 1.2, with the help of artificially induced negative supercoiling to mimic conditions within a cellular context, Laurence Hurley's group showed that c-MYC promoter is able to form a different i-motif structure at neutral pH with a 6:2:6 loop topology.[53,60] Later studies identified two proteins that bind to the C-rich single strand of c-MYC NHE III_1, named heterogeneous ribonucleoprotein K(hnRNP K) and human nonmetastatic 23 isoform 2(NM23-H2).[53]

GC-rich sequences can also be found close to the Bcl-2 P1 promoter, the C-rich sequence is able to form an i-motif structure with a high transitional pH (pH_T 6.6).[42] The Bcl-2 i-motif adopts a 'Class II' loop topology (8:5:7) with a highly dynamic equilibrium towards a hairpin structure.[42,56] Hurley's group found two specific ligands, **NSC 1398948** and **NSC 59276** (Figure 1.6), that are able to interact with this i-motif forming sequence.[56] **NSC 1398948** binds the i-motif structure while **NSC 59276** binds a hairpin structure which exists in equilibrium with the i-motif. Further experiments showed that the Bcl-2 mRNA expression level was up-regulated by the stabilization of i-motif structure and down-regulated by converting the i-motif structure into a hairpin.[56] This is direct evidence that, like G-quadruplexes, ligands that target the i-motif may also have the ability to regulate gene expression.

NSC 1398948   NSC 59276
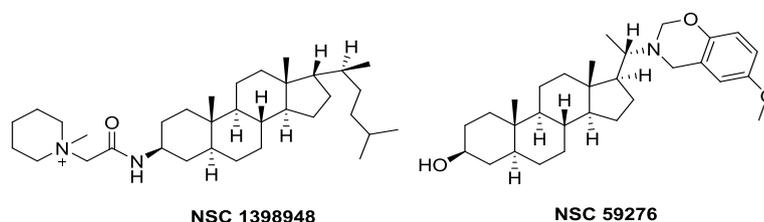
Figure 1.6   Ligands interacting with the i-motif forming sequence from Bcl-2 promoter

A follow-up study carried out by Hurley's group found the protein responsible for converting Bcl-2 i-motif into its hairpin form, hnRNP LL, an analogue of hnRNP K that binds to c-MYC G-quadruplexes (Section 1.2).[57] The dissociation constant of hnRNP LL towards Bcl-2 i-motif was determined at pH 6.5 *via* surface plasmon resonance.The K_d for i-motif was found to be 19.4 pM whilst the K_d for the single stranded DNA was 69.8 pM.[57] Futher studies showed that the RRM1 and RRM2 domains of hnRNP LL contribute to the binding and unfolding of i-motif.[58] As a result, the current working hypothesis is that transcription will increase as the consequence of i-motif unfolding.[57]

1.3.3 Methods for characterising i-motif structures

A wide range of biophysical methods are employed to characterise and investigate non-canonical secondary DNA structure. The examples presented below are commonly used for i-motif research.

Förster resonance energy transfer (FRET) DNA-melting

A FRET DNA-melting assay is a routine primary screen to study the structure and the effects of ligands on the stability of G-quadruplexes and i-motifs.[80] The experiment is carried out with artificially-synthesised nucleic acids tagged with two fluorescent substituents on either end of the oligonucleotide. For example, 6-carboxyfluorescein (**6-FAM** or FAM) can be conjugated to the 5′- end and 6-carboxytetramethyl-rhodamine (**6-TAMRA** or TAMRA) to the 3′- end (Figure 1.7).[81]

FAM
(6-carboxyfluorescein, donor)

TAMRA
(6-carboxytetramethylrhodamine, acceptor)
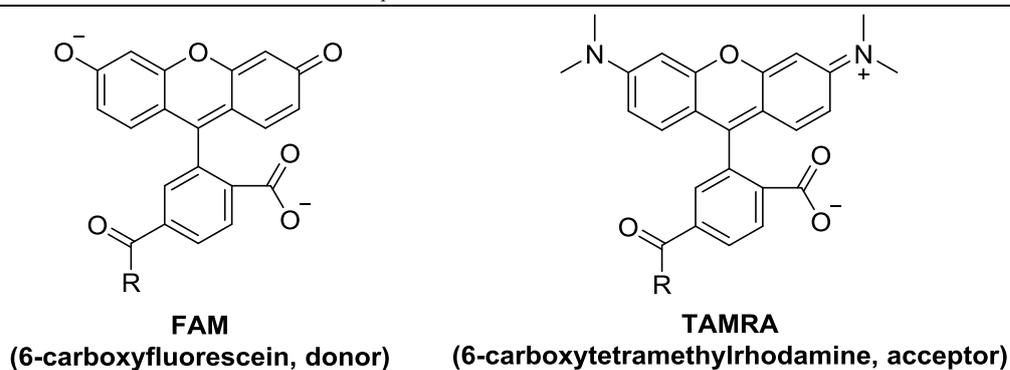
Figure 1.7   Chemical Structure of FAM & TAMRA

In this case the FAM fluorophore behaves as a donor while TAMRA serves as an acceptor (quencher). When the DNA is unfolded, there is no interaction between FAM and TAMRA due to a large distance between them. Exciting the FAM group at 495 nm gives rise to an emission from FAM at 517 nm. Once the DNA structure is folded, FAM can be designed to be proximal to the TAMRA group. In this case, exciting FAM at 495 nm results in the emission of FAM at 517 nm being quenched and replaced by the emission of TAMRA at 580 nm, known as the FRET phenomenon.[80] During the process, the energy received by the donor is transferred to TAMRA *via* a non-radiative dipole-dipole interaction if the distance between the donor and acceptor is within the range of 15 to 60 Å. The FRET efficiency (E_{FRET}) can be described mathematically as the distance between the two fluorophores (Equation 1.1) or the change in fluorescence of the donor (Equation 1.2).[82]

Equation 1.1:
$$E_{\text{FRET}} = \frac{R_0^6}{R^6 + R_0^6}$$

Equation 1.2:
$$E_{\text{FRET}} = 1 - \frac{FI_d}{FI_d^0}$$

Where E_{FRET} is the FRET efficiency, R is the distance between FAM and TAMRA, R_0 is the Förster distance, where the energy transfer between fluorophores is 50%, FI_d is the fluorescence of FAM in the presence of TAMRA, FI_d^0 is the fluorescence of FAM in the absence of TAMRA.

By selectively filtering and measuring the emission intensity of FAM at 517 nm, an indication of how much of the DNA population is folded or unfolded can be determined by the changes in fluorescence (Figure 1.8).
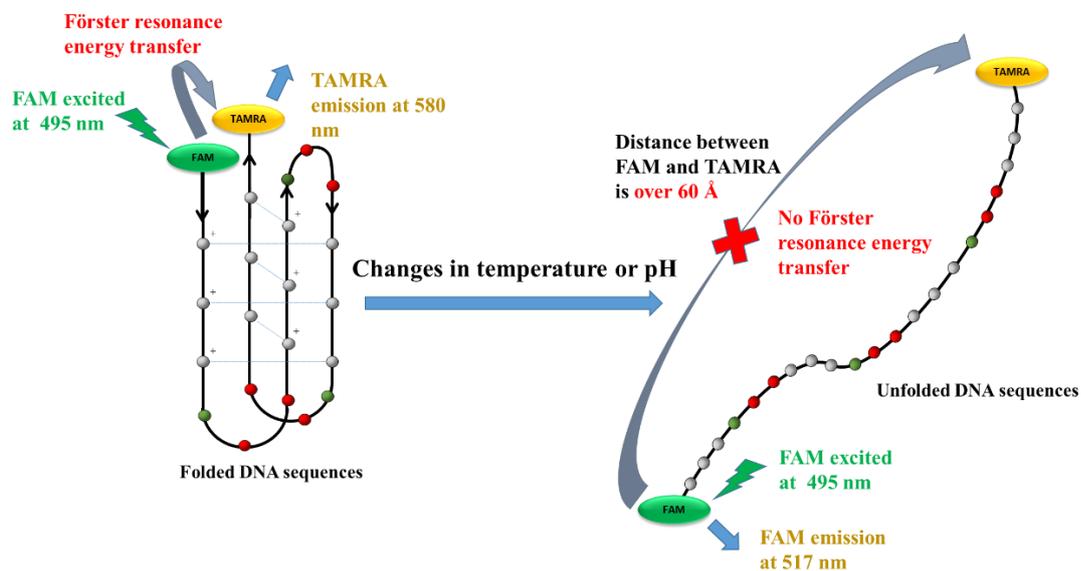
Figure 1.8   *A diagram illustrating how Förster resonance energy transfer can monitor folding of DNA*

Disassociation of oligonucleotides can be achieved by changing their environment or temperature. Thus, thermal stability of a particular oligonucleotide structure can be assessed by measuring and comparing the melting point (T_m) of DNA, the temperature at which 50% of the secondary structure population has unfolded into single strands.[80] Additionally, FRET can also be applied at a static temperature to investigate the folding or unfolding patterns of i-motif on changing the conditions or adding a ligand.

UV-Vis Spectroscopy

Oligonucleotides contain purine and pyrimidine bases. Both of which absorb light around 260 nm as a characteristic property of their aromaticity. However, this property is altered when oligonucleotides are folded due to π-π stacking interactions between nucleobases. This phenomenon can also be employed to identify and

investigate DNA conformational changes.[39,83] In the study of i-motifs, UV-absorption is normally recorded from 220 to 330 nm.[76] When i-motif forming sequences become folded into an i-motif, this results in an increase in absorbance at 240 nm and a decrease in absorbance at 295 nm.[83,84] Just like FRET, UV-Vis spectroscopy can also be used to investigate thermal stability or folding patterns of oligonucleotides. However, results from UV-Vis spectroscopy may be complicated by ligand absorptions which commonly occur in the same region as DNA. [39]

Circular Dichroism (CD)

Circular dichroism is a low-resolution but highly-sensitive experimental technique using circularly polarised electromagnetic rays to detect macro-biological molecules such as peptides and nucleotides. Oligonucleotides are chiral macromolecules that possess different glycosidic bond angles. These glycosidic torsion angles absorb right- and left-handed polarized light differently, which can be measured and quantified empirically as ellipticity.[85,86]

For an intramolecular i-motif, the characteristic CD spectrum gives a large positive peak at 288 nm and a strong negative peak at 255 nm. In contrast, a CD spectrum for a single-stranded DNA molecule has a weaker positive peak at 275 nm and a negative peak at 250 nm.[87] CD can be employed as a diagnostic and dynamic method to monitor the conformational changes induced by altered environment or ligand interactions.

Nuclear Magnetic Resonance (NMR) spectroscopy

NMR spectroscopy is a high-resolution experimental technique for the study of oligonucleotide secondary structure, providing detailed dynamic information, stability, atomic resolution structures and intermolecular interactions *in vitro*.[88] In a $C \equiv C^+$ base pair, the imino proton gives a resonance between 15 to 16 ppm and the two amino protons give resonances from 8 to 10.5 ppm, which can be

considered the characteristic signals for i-motif structures.[89] In a Hoogsteen base pairing, the characteristic peak assigned as imino proton is at 15 to 16 ppm[90,91] while in a Waston-Crick base pairing, this characteristic peak is normally found from 12 to 13 ppm.[56]

Surface Plasmon Resonance (SPR) spectroscopy

SPR spectroscopy is an optical-based, real-time and highly sensitive detecting method for binding affinity and kinetic studies.[92] Its mechanism is based on the changes in refractive index near a sensor surface (metal) induced by ligand binding or dissociation. In an SPR experiment, biotin-tagged oligonucleotides capable of forming i-motif structures are non-covalently attached to streptavidin coated sensor.[93] The side immobilised with oligonucleotides is washed continuously with running buffer and laser light is cast onto the other side of gold plate with the correct angle to create resonant plasmons. The reflected light contains a resulting detectable interference band as an SPR signal. By adding ligands into the running buffer, the interaction between ligands and i-motif will be recorded as changes in the interference band.[92,93] By measuring and analysing variations of refractive index on the interference band detailed ligand binding, information such as binding constants and stoichiometry, can be determined.

1.3.4   i-Motif binding ligands

Compared to the large library of G-quadruplex interacting compounds,[94] the number of i-motif binding ligands is quite limited. Only a handful of compounds have been reported to interact with i-motifs (Figure 1.9). Apart from the Bcl-2 i-motif interacting compound mentioned in Section 1.3.2 (**NSC 1398948**), one of the most prominent ligands is the carboxyl modified single-walled carbon nanotube **(SWNT)**. This reagent is able to specifically stablise human telomeric i-motif even in slightly alkaline conditions (pH = 8) without any interaction with G-quadruplex and duplex DNA.[95] A proposed mechanism for this stablisation is that the nanotube

TMPyP4

NSC 138948

BisA

**Carboxyl-functionalised
single-walled carbon nanotubes
(SWNT)**

Crystal violet

tilorone

[Tb$_2$($_{DL}$-HVal)$_4$(H$_2$O)$_8$]Cl$_6$·2H$_2$O

tamoxifen

mitoxatrone
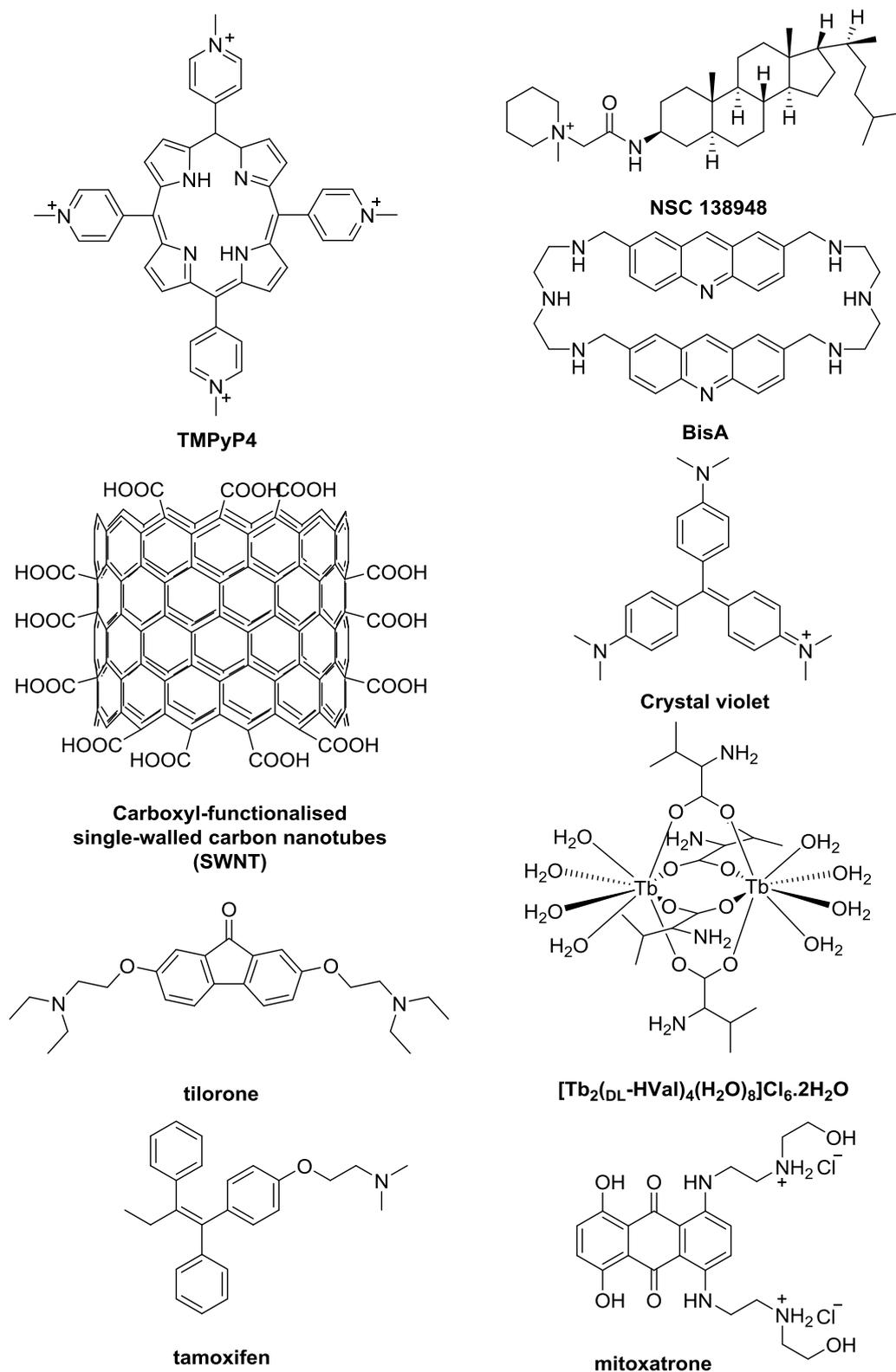
Figure 1.9   Reported and group in-house ligands which have been found to interact with i-motifs.

binds to the 5′-end major groove of i-motif while the carboxyl acids attached to the nanotube interact with the C≡C$^+$ base pairs *via* electrostatics. Later experiments showed that this nanotube can also inhibit telomerase.[72] The cationic porphyrin **TMPyP4** is another ligand that binds to i-motifs with a dissociation constant of 45 μM. Unfortunately, it also binds to duplex DNA (1.2 μM) and G-quadruplexes (0.5 μM), so **TMPyPy** lacks selectivity and binds i-motif nuch weaker than other types of DNA structure.[96] **Crystal violet** has been reported to bind i-motif formed from a poly-cytidine sequence. But its binding against other more complicated i-motif structures are not documented.[97] **BisA** is another i-motif binding ligand based on an acridine structure that has been studied before. It was found that this compound stabilises human telomeric i-motif at pH 6.8.[98] Other ligands, such as phenanthroline derivatives[99] and terbium complexes[100], showed very weak i-motif binding affinities. In our group, several i-motif binding ligands were identified *via* FRET melting and SPR binding studies.[101] These compounds include **mitoxantrone**, **tilorone, tyrothyrcin** and **tamoxifen**. Among them, **mitoxatrone** was the most effective i-motif binder, possessing the lowest micromolar K_d towards human telomeric i-motif ($K_d = 11$ μM).[101,102] meanwhile, **tilorone** was found to bind with a far weaker affinity ($K_d = 158$ μM).[101]

Given the scant literature reporting i-motif interacting ligands, and bearing in mind the potential biological implications of being able to target them, more work needs to be done to discover new compounds which can bind and stabilize i-motif.

1.4 Thermodynamics of oligonucleotides

Oligonucleotides tends to self-assemble into higher order structures, driven by the local thermodynamic and mechanical properties of the molecules. The study of thermodynamics includes binding energies (thermodynamics), transition-state enerties, mechanisms and transformational rates (kinetics), yielding macroscopic data of DNA self-assembly or DNA-ligand interactions.

1.4.1 DNA self-assembly

Single-stranded DNA molecules are not stable in aqueous solution due to the presence of hydrophobic nucleic base. These molecules can self-assemble into higher order structures either between two molecules or on their own. The whole process can be explamed via the classical Gibbs free energy equation and a diagram shown below.
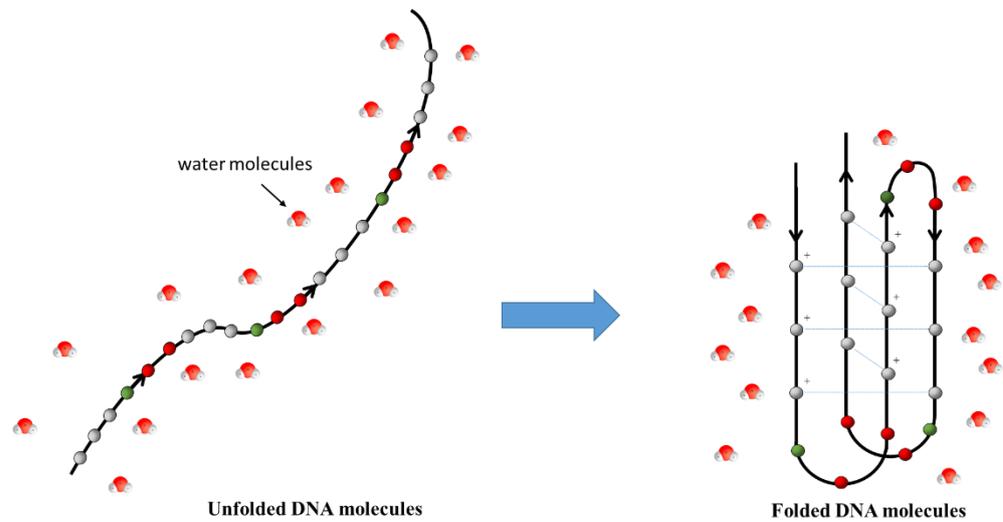
Figure 1.10 DNA folding process pushing the water molecules aside, human telomeric i-motif structure as an example.

Equation 1.3 $$DNA_{unfold} \overset{K}{\Leftrightarrow} DNA_{fold}$$
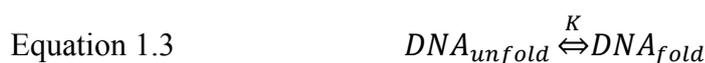
Equation 1.4 $$\Delta G = \Delta H - T\Delta S = -RT\ln K$$

In Equation 1.4, the sign and value of ΔG determines the spontaneity of the reaction. The ΔH is generally attributed to changes in covalent and hydrogen bonds and salt bridges. In the process of DNA self assembly, enthalpy changes are generally negative in duplex DNA, G-quadruplex and i-motif formation due to the energy output in hydrogen bond formation. [103-105] The ΔS attributes to the order of system, which is also generally negative due to the changes from a less ordered structure to an ordered system. However, to the whole environment (including the DNA

molecules and buffer), the entropic changes is increased.

Whether the higher-order DNA structures are present or not is determined by the competitation between entropic ($T\Delta S$) and enthalpic (ΔH) components. By changing the temperature of the environment, the ΔG can reach 0, where oligonucleotides exist in a transit state between folding and unfolding. [105]

1.4.2   DNA-ligand binding.

Intercalation [106] and the minor groove binder [107] are the two principle mode in DNA-ligand binding. [107] In the case of intercalation, flat and aromatic molecules interacts with nucleobase by π-π stacking, which results in a negative enthalpic change while the hydrophobicity of the molecule contributes to a positive entropic change. [107] However, for minor groove binders, such as Hoechst 33258, positive changes in entropy alone contributes to the majority of the Gibbs free energy. [107]

Chapter 2   **Development of a High-throughput Fluorescent Intercalator Displacement Assay for Human Telomeric i-motif**

2.1   Introduction

Despite the fact that research in the field of i-motif is growing rapidly, there are still only a limited number of i-motif interacting compounds recorded in the literature.[40] Additionally, compounds identified to interact with i-motifs often possess higher affinity towards duplex or G-quadruplex structures.[40] Thus, to search for compounds which specifically interact with i-motifs has became increasingly important in the research of i-motifs.

As described in Section 1.3.3, FRET melting based DNA experiments are widely used as a screening method for DNA binding ligands. However, this method still has some limitations. First, oligonucleotide samples used in FRET assays are all modified on both ends of the sequences with fluorophores, such as FAM and TAMRA. These two bulky, planar and aromatic substituents may affect DNA conformations, especially if short sequences are used.[80] Futhermore, the addition of fluorophores to the DNA is more expensive than using non-labelled oligonucleotides. Apart from their price, FRET-labelled oligonucleotides may also possess altered binding affinities towards various DNA binding ligands.[80] Additionally, melting experiments require heating the sample to break down hydrogen bonding interactions between the $C \equiv C^+$ base pairing, which is very different from physiological process observed in the cell and also prevents the use of traditional plate readers for screening different types of ligands.

In this chapter, a screening method based on fluorescent intercalator displacement was developed to identify DNA interacting ligands, and tested in human telomeric i-motif.

2.1.1   Fluorescent Intercalator displacement (FID)

Fluorescent intercalator displacement, also known as fluorescent intercalator replacement, is a simple, cost-effective, nondestructive and high-throughput

method to study DNA sequence selectivity and binding affinity. This assay was first described by Boger, Fink and Hedrick,[108] where DNA was treated with a fluorescent intercalator (Figure 2.1a), such as ethidium bromide,[108] thiazole orange (TO),[109] and Hoechst 33258 (Figure 2.1b),[110] resulting in significant fluorescence enhancement due to DNA intercalation. When a non-fluorescent DNA binding agent was added and replaced the bound fluorescent intercalator, the fluorescence of the DNA-intercalator complex decreased as the result of competitive displacement (Figure 2.1a).[108,109] By calculating the decreasing percentage in fluorescence, relative DNA binding affinities can be calculated. In the same way, when subsequent quantitative titration is applied, accurate absolute binding constants may be obtained by calculation of the changes in fluorescence. By altering the DNA sequence, even more detailed insights such as the size of the binding site on nucleic acids can be investigated.[108] DNA intercalators, such as ethidium bromide and thiazole orange, insert into the gaps between adjacent base, exhibiting drastic changes in their respective fluorescent properties. However, non-fluorescent Hoechst 33258 does not bind DNA in the same way, instead intercalating into the minor-grove and binding specifically to adenine and thymidine base pairs, but still exhibits enhanced fluorescence in a similar manner to thiazole orange and ethidium bromide.

(a)

Gap for intercalators to insert

Space for groove binders

probes (non-fluorescent)

Ligand candidate

Displaced probes (non-fluorescent)

Duplex DNA (non-fluorescent)

DNA-probe complex (fluorescent)

DNA-ligand complex (fluorescent probes displaced)

(b)

Ethidium bromide

Thiazole orange

Hoechst 33258
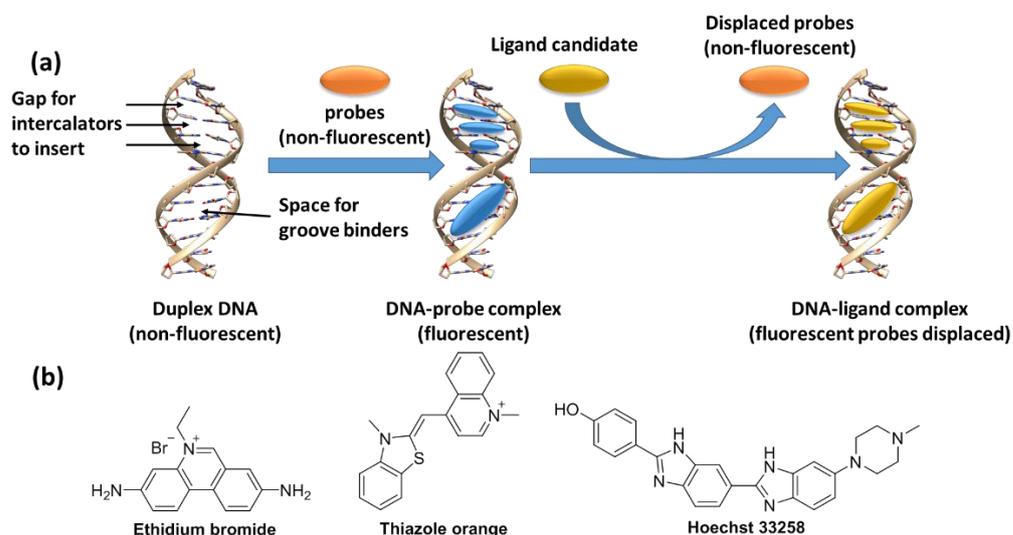
Figure 2.1   a) Example of fluorescent intercalator displacement on B-DNA structure. b) common fluorescent probes for FID assay

FID assays were first established on double-stranded and hairpin structures[111] and later on applied to G-quadruplexes.[112,113] In terms of executing an FID assay, thiazole orange (TO), an asymmetric cyanine, offers many advantages over ethidium bromide.The fluorescence quantum yield (Φ_F) of thiazole orange is very low ($\Phi_F = 2 \times 10^{-4}$) in the absence of DNA, however, when bound to G-quadruplexes and double-helical DNA, the quantum yield increases 100 to 1000 times.[111] For comparison, fluorescence enhancement caused by ethidium bromide is only 20-fold.[109] Apart from that, G-quaduplex-bound thiazole orange possesses a large Stokes shift. It is excited by light at 485 nm but emits fluorescent radiation at 550 nm. This property minimises the possibility of fluorescence self-quenching.

In a series of fluorometric studies carried out by Teulade-Fichou's group, thiazole orange was observed to bind intramolecular human telomeric G-quadruplexes (22AG, 5′-AG$_3$(T$_2$AG$_3$)$_3$-3′) in an anti-parallel conformation with reasonably high affinity ($K_a = 2.1$ to 3.4×10^6 M^{-1}) and an unambiguous 1:1 stoichiometry (Figure 2.2) in both K$^+$ and Na$^+$ buffers.[114] Once thiazole orange was bound to 22AG, it can be excited at 501 nm and its corresponding fluorescent emission peaked at 539 nm.[112] In contrast intermolecular G-quadruplex composed of four oligonucleotides

([TG$_3$T]$_4$) has two available binding sites for thiazole orange and a similar binding affinity (K_a = 1.5 × 10^6 M^{-1}, for each of the binding sites).[114] The only difference between these two G-quadruplex structures is the loop region. Thus thiazole orange locates in the loop-free end of human telomeric G-quadruplex, exerting external π-π stacking on G-tetards.[114] Thiazole orange binds to double-helical DNA with similar affinity (K_a = 2 to 3 × 10^6 M^{-1}).[115] The stoichiometry of thiazole orange binding to G-quadruplexes was further studied with ESI-MS experiments.[115] In these experiments, the result suggested a different binding stoichiometry. The major observation of thiazole orange binding falls in the 1:1 stoichiometry, with a marginal sub-population of 2:1 TO:DNA complex. This might result from a second binding site for thiazole orange which possesses a very weak fluorescence quantum yield.[115]
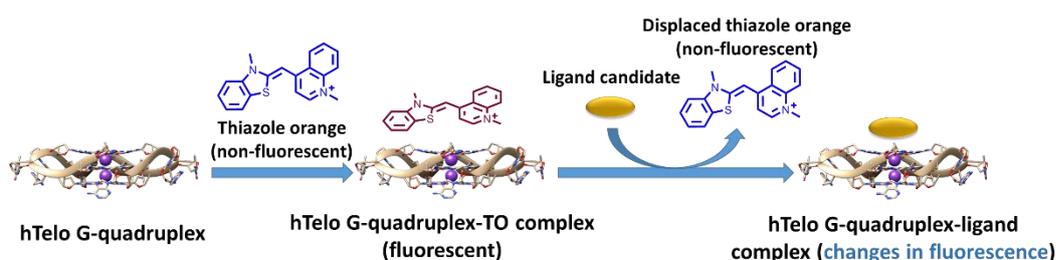
Figure 2.2   *Example of fluorescent intercalator displacement on human telomeric G-quadruplex structure*

After confirming the binding mode between thiazole orange and human telomeric G-quadruplex, Teulade-Fichou *et al.* carried out a series of fluorescent intercalator displacement experiments with different G-quadruplex forming sequences and various G-quadruplex binding ligands.[115-117] These G-quadruplex forming sequences were not restricted to human telomeric repeats but also included other proto-oncogenic sequences, such as VEGF, c-kit, c-MYC and K-ras G-quadruplex forming sequences, indicating that a wide range of G-quaduplexes are suitable for FID assay. Additionally, the ligand to oligonucleotide ratio was set as 2:1 to mimic the two plausible binding sites on G-quadruplexes.[115] Further development led to a high-throughput screening method for G-quadruplex ligands allowing for quick

determination of ligand-G-quadruplex interactions.[117]

2.1.2 Fluorescent probes interacting with i-motifs

Given the success of using FID for DNA binding compounds and the lack of known i-motif binding ligands, we decided to explore the potential of using FID assays to identify ligands for i-motif structures. Thus, the first step for the development of an i-motif-FID assay is to find the ideal fluorescent probes for i-motif structures.

Crystal violet

Crystal violet, a triphenylmethane dye (Figure 2.3), is widely used as a biological stain and probe in biochemical research.
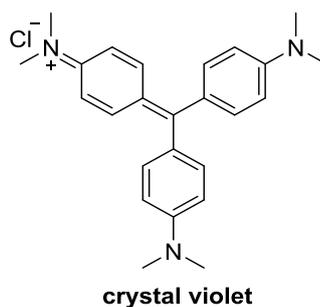
crystal violet

Figure 2.3   Chemical structure of crystal violet

Examples of crystal violet binding to i-motif structures have been reported.[97,118] Oligonucleotide C_{29} ($5'$-$(C_5T_3)_3C_5$-$3'$) was shown to be a stable i-motif with 5 layers of intercalated $C{\equiv}C^+$ base pairs at pH 5.0. Crystal violet binds to this rigid i-motif structure with a similar affinity ($K_a = 1.2{\times}10^6$ M^{-1} 1:1 stoichiometry)[97] as seen for G-quadruplexes.[119] In the UV absorption titration experiment and a drastic enhancement of the fluorescent emission at 580 nm was observed when crystal violet solution was titrated with C_{29} oligomer at pH 5.0.[97] Crystal violet's binding to a tetramolecular i-motif ($5'$-$(AC_3T)_4$-$3'$) was studied *via* molecule docking. The *in silico* calculation results suggested two capping sites on either end of the i-motif (-38.44 and -31.57 kcal·mol^{-1} as binding energy, respectively).[97] The insertion of

crystal violet into this tetramolecular i-motif seems, however, not to be a spontaneous event (23 kcal·mol^{-1} as binding energy). Crystal violet also shows great 'light-on' i-motif selectivity over double-helix and single-stranded structures.[97] In addition, crystal violet was used as a selective electrochemical probe for i-motif structure to achieve pH-driven electrochemical switch.[118]

Thiazole orange

Thiazole orange (TO) has been shown to interact with i-motifs as well (Figure 2.4).[118,120-122]
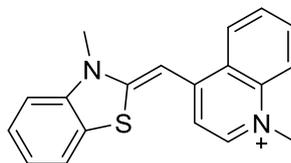
Figure 2.4   Chemical structure of thiazole orange

In 2014, Shao's group reported a pH-gated exciplex using modified i-motif forming sequences. Thiazole orange was covalently conjugated to the human telomeric i-motif, replacing a thymine substituent (T$_{16}$) on the loop. (Figure 2.6a) Later the exciplex was shown to possess a strong fluorescent emission at 580 nm and a Stokes shift of 90 nm at pH 5.5. The fluorescence quantum yield of this exciplex is comparable to that of thiazole orange intercalated into the double-helix DNA (Φ_F = 4 × 10^{-2}).[120] The fluorescence life-time of this exciplex was significantly prolonged (τ = 2.49 ns), which is even higher than thiazole orange in duplex (τ = 1.83 ns). This study suggested that TO can be used as a fluorescent probe if it binds to human telomeric i-motif.[120]

In 2016, a study on i-motif unfolding pathways was carried out using human telomeric i-motifs tethered with thiazole orange as a fluorescent reporter. In this study, single thiazole orange molecules replace multiple nucleic bases on the human telomeric sequences, such as at the 5′-end, loop I, loop II, loop III and the 3′-end

position.[122] All TO-modified nucleic acids adopted i-motif structures at pH 5.0 in sodium cacodylate buffer and were excited at 490 nm in the same condition. When thiazole orange was at the 5′ -end, the 3′-end and loop II, fluorescent emission was observed at 535 nm and when thiazole orange was at the loop II and loop I and loop III, fluorescent emission was observed at 580 nm.[122] By monitoring the change of fluorescence for four TO-modified i-motifs individualy when changing the temperature, it was concluded that the disassociation of intermolecular human telomeric i-motif starts with unzipping $C \equiv C^+$ base pairs initiated from the 3′- to form a triplex-like intermediate first.[122]

Pei *et. al.* questioned whether thiazole orange can be directly used as fluorescent light-up probe for human telomeric i-motif in 2015.[121] In their research, they focused on applying H^+ or Ag^+ to achieve 'OR' logical gate with TO-i-motif complex.[121] Thiazole orange was previously incubated with human telomeric i-motif in PB buffer at pH 6.0. A fluorescent emission at 535 nm was observed when TO bound to pre-formed human telomeric i-motif and excited at 488 nm.[121]

Pyrene

The interaction of pyrene and i-motif structures has also been reported (Figure 2.5).
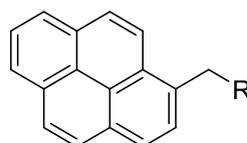
Figure 2.5   Chemical structure of pyrene

In 2011, Kim *et. al.* described an oligonucleotide-based nanostructure close to that observed for thiazole orange with a shorter oligonucleotide (5′-$(TA_2C_4)_2T$-3′).[123] Two adenosines on each of the loop region were modified with pyrene. In acidic buffer, four pyrene motifs on the loop were able to cap both ends of the bimolecular i-motif formed (Figure 2.6b). This capping effect dramatically increased the melting temperature of modified oligonucleotides (ΔT_m = 23.5 ℃) and the

formation of pyrene-capped i-motif in turn enhanced pyrene's fluorescence 4–fold (Figure 2.6b).[123] Pyrene has also been conjugated to i-motif forming sequences closer to human telomeric repeats. In 2013, Majima's group reported an i-motif-based, pH-gated electronic device in which pyrene-appended uridine replaced one thymidine on a TTA loop and successfully achieved photoinduced electron transfer with an anthraquinone substituent conjugated to the other end of human telomeric i-motif.[124]
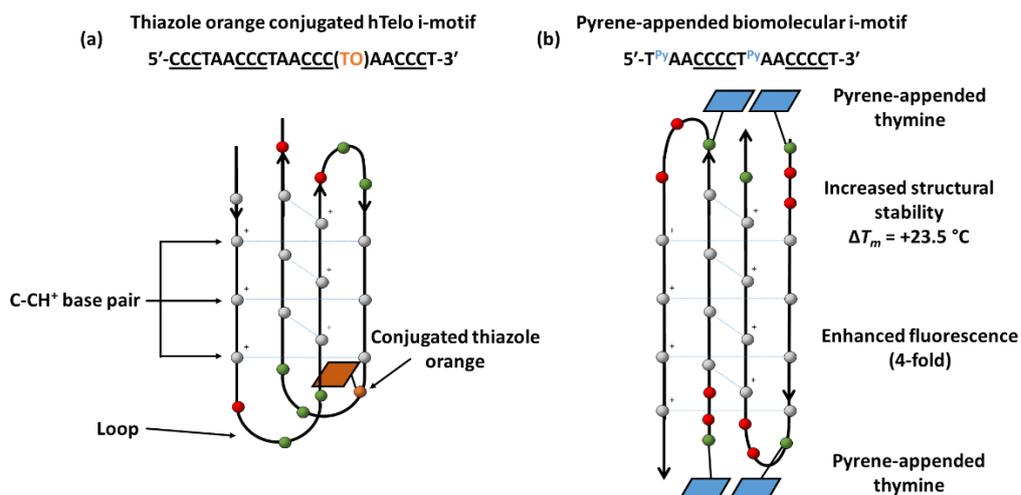
Figure 2.6   Exciplex formed from thiazole orange (a) or pyrene conjugated human telomeric i-motif (b)Pyrene-appended bimoleculear i-motif

2.2   Aims

Although G-quadruplexes-FID assays have been established for over 8 years, there are no reports of i-motif-FID assays. In light of the successful development of FID assays that use thiazole orange as the DNA intercalator and the lack of i-motif binding compounds, it was proposed that a similar assay could be established for screening ligands against i-motif DNA. (Scheme 2.1)
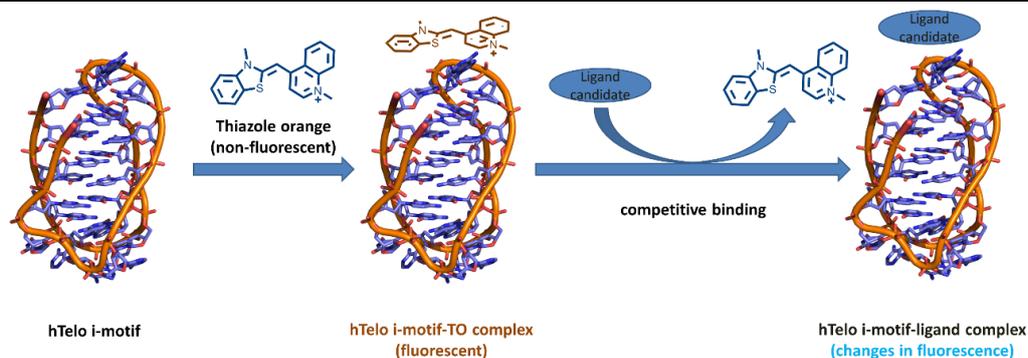
Scheme 2.1   Proposed FID assay on human telomeric i-motif

The aims of this project are:

1) To investigate the potential suitability of different i-motif binding compounds for use as FID fluorescent probes.

2) To characterise the binding properties of the selected fluorescent probes with human telomeric i-motif, to evaluate binding stoichiometry, stabilization of i-motif, binding affinity and the plausible structural alteration of the i-motif.

3) To apply the selected fluorescent probe in an FID assay with some known i-motif binding compounds, to determine methods to screen for possible i-motif binding compounds

4) To estabish a high-throughput ligand screening method for human telomeric i-motif using a compound library.

2.3   Results and discussion

2.3.1   Screening of fluorescent probes suitable for i-motif-FID assay

Ideally a fluorescent probe needs to have a drastic change in fluorescent emission upon binding oligonucleotides. At the beginning of the screening, several G-quadruplex and i-motif fluorescent probes were considered, including acridine

orange, ethidium bromide, crystal violet and 1-pyrenemethanol (Figure 2.7). Acridine orange and ethidium bromide are both used as nuclear stains. Ethidium bromide has been used as a fluorescent probe replacement for thiazole orange in an FID assay previously.[125,126] In the early study of fluorescent intercalator displacement on double helical DNA, acridine orange was reported as an ineffective dye.[127] However, due to the structural differences between double helix and i-motif, acridine orange was still considered as one of the candidate probes.
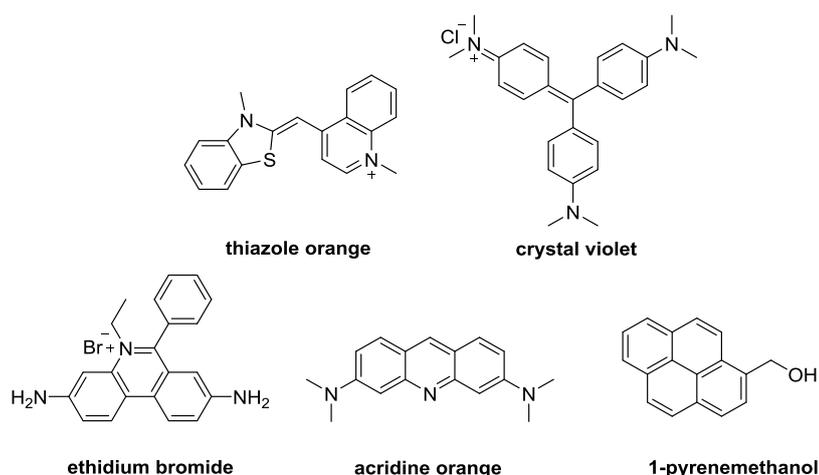
Figure 2.7   Fluorescent probes screened in finding the intercalator for FID assay

The compounds were assessed by monitoring the change in fluorescence after addition of human telomeric i-motif forming DNA (hTeloC, 5′-TAACCCTAACCCTAACCCTAACCC-3′), at pH 5.5, conditions where most i-motif structures are stable. The aim was to observe any changes in fluorescence, such as fluorescence intensity, excitation and emission wavelength, after the probes binds to hTeloC. The result showed that thiazole orange has a dramatic fluorescent enhancement (125-fold) after binding to i-motif with 1:1 concentration ratio (Figure 2.8a). However, the excitation band peaks at 430 nm and gives a narrow Stoke shift of only 21 nm, which is very different from reported i-motif induced fluorescence.[121,122] This might be a result of the fact that thiazole orange in our experiments is not closely stacking on either end of the i-motif.
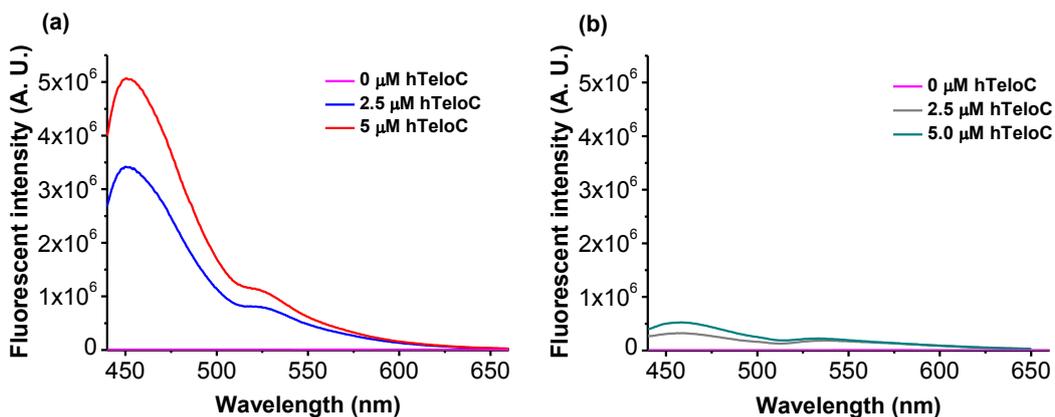
Figure 2.8   Fluorescence emission of TO-i-motif complex in 10 mM sodium cacodylate buffer. The sample was excited at 430 nm. Thiazole orange is 5 μM. a) pH = 5.5, b) pH 7.4

Binding of thiazole orange to hTelo i-motif in pH 7.4 buffer was also studied but no significant fluorescence was observed compared to that in pH 5.5 buffer (Figure 2.8b). This can be explained by the lack of π-π stacking since the human telomeric i-motif forming sequence is mostly in the unfolded conformation at pH 7.4. It is unsurprising that when the pH value was at 5.5 the hTeloC i-motif was rigid enough to provide strong a π-π stacking effect, and thus enhanced the fluorescence of thiazole orange. An additional small shoulder on the emission spectra was observed at 535 nm with an intensity of only 1/3 of the highest emission peak at 451 nm. Thus the emission peak at 451 nm should be selected as the target for fluorescence screening.

In contrast, attempts to use other probes were rather unsuccessful. Although crystal violet had been demonstrated to interact with poly-cytidine DNA, we observed no useful changes in fluorescence on addition of hTeloC (Figure 2.9).
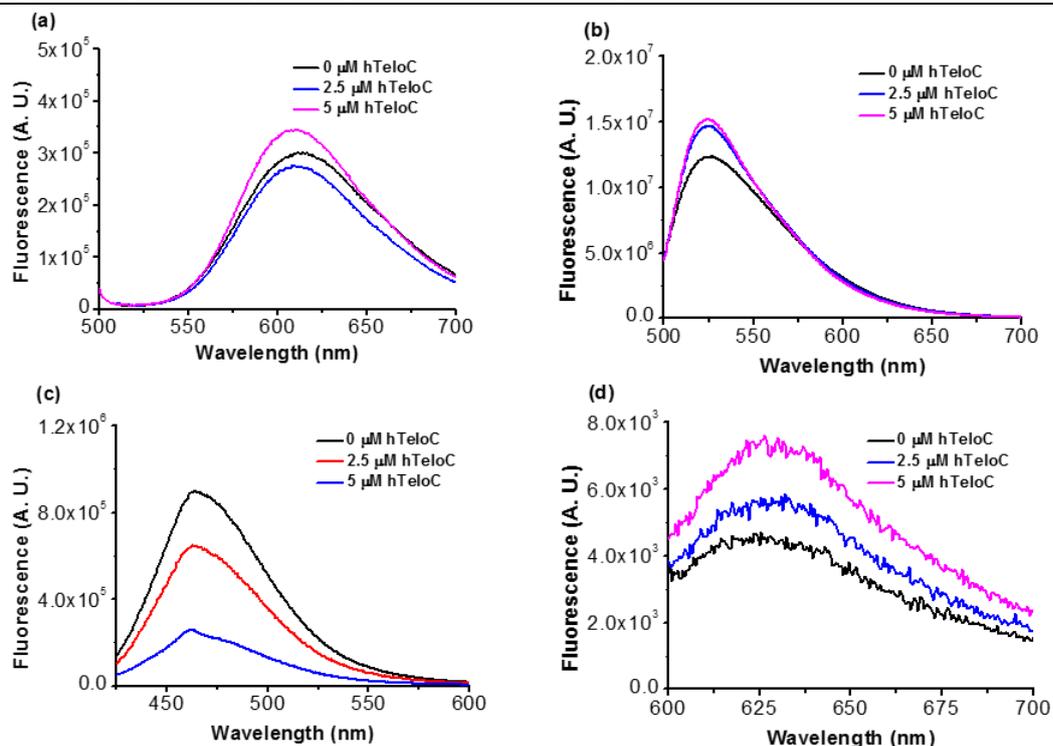
Figure 2.9   Fluoresent emission of probe-i-motif complex in pH 5.5 10 mM sodium cacodylate buffer. a) [ethidium bromide]= 5 μM, excited at 490 nm; b) [acridine orange]=5μM, excited at 490 nm; c) [1-pyrenemethanol] = 5 μM, excited at 440 nm; d) [crystal violet] = 5 μM, excited at 580 nm.

The fluorescence of ethidium bromide and acridine orange only increased by 25 to 70 % upon binding to human telomeric i-motif (1:1 binding), which is not sufficient for a potential FID assay (Figure 2.9a and Figure 2.9b). For 1-pyrenemethanol, its fluorescence decreased by 70% after titration with human telomeric i-motif (Figure 2.9c), indicating a fluorescent quenching on DNA binding. A plausible explaination for this phenomenon is i-motif condensation in the buffer after pyrene binds to human telomeric i-motif. The same phenomenon occurred with ethidium bromide as well. For crystal violet, only a very weak fluorescent emission was detected when the sample was excited at 580 nm (Figure 2.9d). Thus, the other fluorescent probes tested failed to exhibit equal 'light-up' fluorescence properties as thiazole orange. A summary of the fluorescence properties of the tested probes can be found

in Table 2.1. Due to these observed characteristics, thiazole orange was selected as the fluorescent probe for further studies.

Compound	Excitation (nm)	Emission (nm)	Fluorescent Emission enhancement (%)
Acridine orange	490	520	23
Crystal violet	580	625	70
Ethidum bromide	490	620	-12
Thiazole orange	430	450	11860
1-pyrenemethanol	440	470	-74

Table 2.1   Fluorescence probe screening upon binding to hTeloC (1 equivlent). [hTeloC] = 5 μM, [probe] = 5 μM. Buffer: pH 5.5, 10 mM sodium cacodylate.

Fluorescence enhancement of thiazole orange upon i-motif binding

Fluorescent enhancement experiments were performed in collaboration with Tasnim Mahound, a project student previously working in our lab. In order to study the enhancement of fluorescence upon binding to human telomeric i-motif, aliquots of preformed human telomeric i-motif were titrated into a solution of thiazole orange at pH 5.5, resulting in a steady increase in fluorescence emission. This result suggested a concentration-dependant binding event between thiazole orange and hTeloC (Figure 2.10). With 1 equivalent of added DNA, the fluorescence enhancement was found to be 4.52×10^6 (A. U.), 2 equivalence gave 6.14×10^6 (A. U.) and the enhancement started to plateau at 3 equivalence.
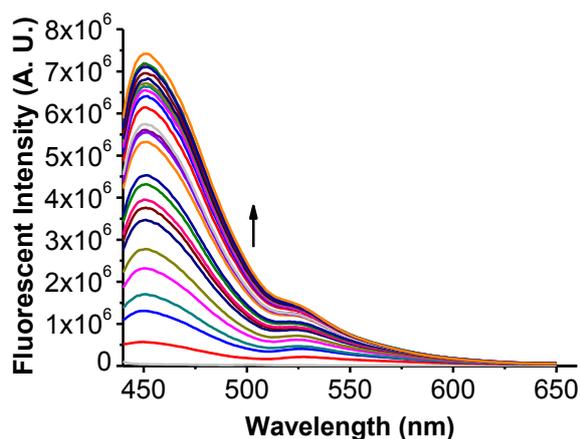
Figure 2.10 Fluorescence emission of thiazole orange titrated with hTeloC at pH 5.5 in 10 mM sodium cacodylate buffer. [TO]= 5μM. [hTeloC] = 0 - 20 μM. Samples were excited at 430 nm.

Additionally, the interaction of thiazole orange and the i-motif forming sequence from the c-MYC promoter (c-MYCC,[41] 5′-TCCCCACCTTCCCCACCCTCCCCA CCCTCCCCA-3′) was also examined by analogous titration experiments. The results also showed a concentration-dependant fluorescence enhancement similar to that observed for hTeloC (Figure 2.11). This suggested that the FID assay developed on thiazole orange with human telomeric i-motif can be applied to other i-motif structures in the future.
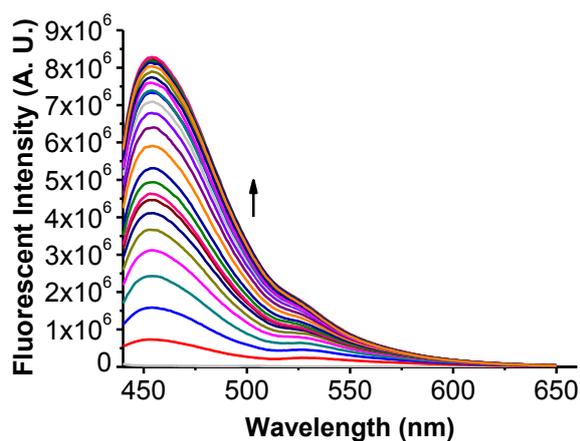
Figure 2.11 Fluorescence emission of thiazole orange titrated with c-MYCC at pH 5.5 in 10 mM sodium cacodylate buffer. [TO]= 5μM. [hTeloC] = 0 ~ 20μM. Samples were excited at 430 nm.

2.3.2 Determining the binding stoichiometry and affinity of thiazole orange

After confirming the potential of thiazole orange as the fluorescent probe in i-motif-FID assays, the binding properties of thiazole orange towards hTelo i-motif needed to be determined. Two UV-based methods were used to characterise the interaction between TO and i-motif. These experiments were performed in collaboration with Joe Neaveason, a project student under my supervision throughout his research in the lab.

Continuous Variation Binding analysis

Using the method of continuous variation binding analysis (also known as a Job plot) with UV spectroscopy we were able to determine the stoichiometry of thiazole orange binding to hTeloC i-motif.[115,128] In these experiments, the total concentration of thiazole orange and hTeloC was kept constant, but the ratio of TO:hTeloC varied. Using maximum concentrations of 20 µM of each component in 50 mM sodium cacodylate buffer at pH 5.5, the samples were measured using UV spectroscopy. In Figure 2.9, the peak at 505 nm increased as more thiazole orange was titrated with the hTeloC, however this absorption peak plateaued when the concentration ratio of thiazole orange to hTeloC was around 6:4 to 7:3. The resulting Job plot indicated a 2:1 binding ratio of TO:hTeloC (Figure 2.12). This result may suggest an end-stacking binding mode between human telomeric i-motif and thiazole orange.
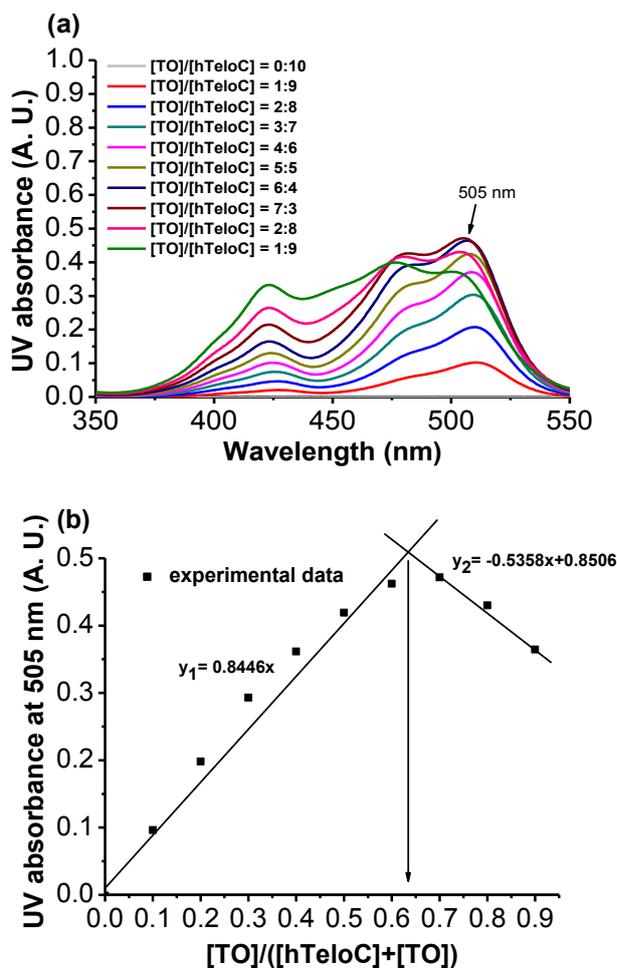
Figure 2.12 a) UV-vis spectra of hTeloC and thiazole orange at pH 5.5 in 50 mM sodium cacodylate buffer; b) UV Job Plot for thiazole orange;hTeloC titration, indicating 2:1 stoichiometry

UV titration and curve fitting

The binding affinity between thiazole orange and hTeloC was determined using UV titrations. Starting with 5 µM of hTeloC in 50 mM sodium cacodylate at pH 5.5, small aliquots of thiazole orange were titrated in and the resulting UV spectrum taken. Using the least squares method,[129,130] a hyperbolic binding curve of fraction bound versus TO concentration was generated. Given the known stoichiometry, the data was fitted with an independent two-site binding model (Equation 2.2) to give two dissociation constants (K_{d1} = 3.61 µM and K_{d2} = 75.6 µM, Figure 2.13). Fitting to a 1:1 binding model (Equation 2.1) was also attempted (Figure 2.13) but did not

give a good fit to the data.

Equation 2.1
$$\theta = \frac{nK_1[Ligand]}{1 + K_1[Ligand]} \qquad n = 1$$

Where θ is the fraction of binding measured, K_1 = the equilibrium association constant for the binding sites. $K_d = 1/K_1$. n = 1 due to the set-up of the model.[131,132]

Equation 2.2
$$\theta = \frac{K_1[Ligand] + 2K_1K_2[Ligand]^2}{1 + K_1[Ligand] + K_1K_2[Ligand]^2}$$

Where θ is the fraction of binding measured; K_1 = the equilibrium association constant for the first binding site; and K_2 = the equilibrium association constants for second binding site. K_{d1} can be determined by the reciprocal of the binding constant (ie. $K_{d1} = 1/K_1$). The results from the binding studies support the model that two thiazole orange molecules bind to one i-motif.[130-132]
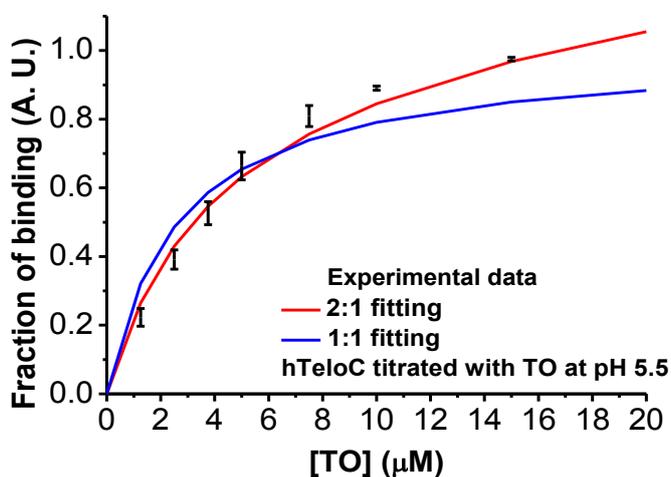
Figure 2.13   UV-vis spectroscopy titration of thiazole orange with hTeloC at pH

5.5 in 50 mM sodium cacodylate buffer. UV absorbance was

measured at 505 nm. [hTeloC] = 10μM

The experiments were also conducted in an opposite manner, by titrating hTeloC into 5 μM TO and monitoring the fluorescence enhancement. The resulting binding constants were found to be the same within error. These results reveal that thiazole orange binds in a weaker affinity to i-motif compared to G-quadruplex and duplex DNA (which demonstrate K_d between 0.3 and 0.67 μM).[114,115] However, this is important, because if thiazole orange were to bind i-motif DNA too tightly, it would be more difficult for a ligand being tested to displace it from the DNA. Since previous assays with other DNA secondary structures have been shown to work with TO binding with similar affinities, these results demonstrate that TO can bind i-motif DNA strongly enough to observe good fluorescence, but weakly enough to be displaced by a ligand.

2.3.3   Effect on i-motif conformation and stability of thiazole orange binding

Experiments in Section 2.3.3 were performed in collaboration with Joe Neaveason, a project student under my supervision throughout his research in the lab.

Given i-motif forming sequence can exist in equilibrium with hairpin structure,[49,57] experiments to determine the nature of the bound complex are required. Circular dichroism (CD)[87] is a technique capable of distinguishing between hairpin and i-motif secondary structures. It was therefore used to investigate the effects of adding thiaozle orange to hTeloC under different pH conditions to determine whether the ligand is able to alter or influence the conformation of the i-motif.

At and below the transitional pH (pH 6 and pH 5.5) intense positive signals are observed around 288 nm accompanied by negative signals around 258 nm for hTeloC, both characteristic of i-motif formation.[87] These signals are less intense at pH 6 compared to pH 5.5, indicating the proportion of i-motif present is higher at pH 5.5. At pH 7.4 the signal at 288 nm is absent, indicating hTeloC is no longer folded into an i-motif. At this pH the equilibrium is shifted towards a mixture of random coil and possible hairpin structures.[87] Thiazole orange was titrated into

hTeloC at different pHs to investigate whether it could alter the conformation of DNA. At pH 5.5, significant changes in the CD spectrum were observed, with a shift in the negative signal from 258 to 270 nm and a decrease in the positive signal at 288 nm (Figure 2.14a). A decrease in the positive signal at 288 nm rather than a shift in the positive signal from 288 nm to 275 nm indicates the changes observed are not consistent with i-motif unfolding or converting to a hairpin structure. The smooth CD spectra obtained when 10 equivalents of thiazole orange were titrated with human telomeric i-motif indicated the signal-to-noise ratio was not affected and unlikely to be the result of precipitation. The signals could be due to an induced CD signal (ICD), which occurs on TO-i-motif binding. A similar ICD effect was also observed at pH 6 (Figure 2.14b). However, at pH 7.4 (Figure 2.14c), a reduction in signal intensity appeared at 10 μM and 50 μM of thiazole orange, coupled with a change to poor signal-to-noise in the spectrum. Additionally, a positive absorbance was also observed at 300-320 nm, indicating a scattering effect Together these features indicate precipitation at pH 7.4. This shows that at higher pH and higher concentrations of thiazole orange the DNA condenses and scatters light. It also reveals that the i-motif is less susceptible to condensation in acidic conditions than the same sequence at physiological pH.
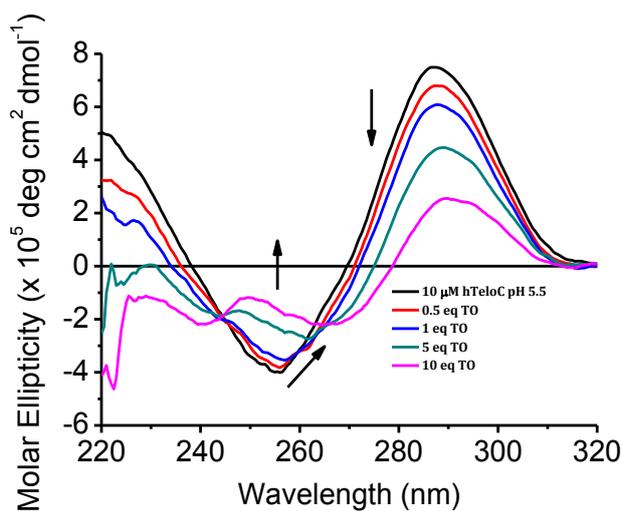
Figure 2.14a: *Circular dichroism titration of hTeloC (10 μM)with thiazole orange (0 – 100 μM) at pH 5.5 10mM sodium cacodylate buffer.*
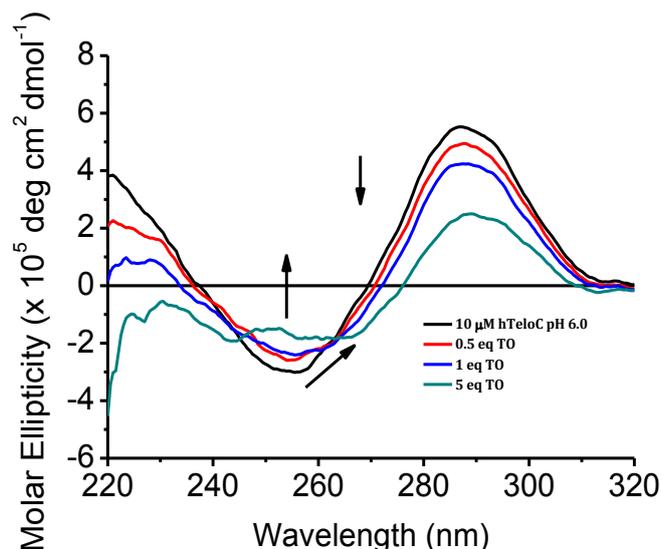
Figure 2.14b: Circular dichroism titration of hTeloC (10 μM)with thiazole
            orange (0 – 50 μM) at pH 6.0 10mM sodium cacodylate buffer.
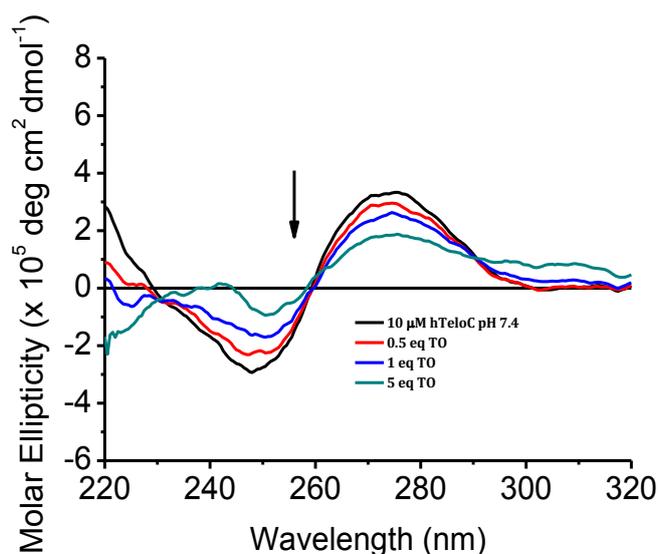
Figure 2.14c   Circular dichroism titration of hTeloC (10 μM) with thiazole
            orange (0 - 50 μM) at pH 7.4 10mM sodium cacodylate buffer.

These experiments highlighted some potential limitations of the conditions possible for use in an FID-type assay; to counter this, acidic pH and lower molar equivalents of thiazole orange were therefore selected. Nevertheless, the results from the CD indicate that the addition of thiazole orange does not alter the conformation of the

hTeloC i-motif within 10 micromolar concentration, so is suitable for use in a FID assay under these conditions.

To determine whether thiazole orange can stabilise i-motif DNA at different pHs, DNA melting experiments were conducted using FRET-melting experiments.
In addition to hTeloC and c-MYCC[41] which are i-motifs that are stable at slightly acidic pH, HIF-1αC[39] (5′-CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGC-3′) was also examined as it is relatively stable at near-physiological pH (Figure 2.12a). For comparison, G-quadruplex (hTeloG, 5′-GGGTTAGGGTTAGGGTTA GGG-3′) and double helical DNA (ds, 5′-TATAGCTATA-HEG(18)a-TATAG CTATA-3′) were also studied (Figure 2.15). For the i-motif forming sequences, melting experiments were performed at the transitional pH (where 50% of the i-motif is folded) and also at pH 5.5. All experiments showed that thiazole orange has a stabilising effect on i-motif DNA, regardless of sequence or pH. For example, adding 12.5 μM of TO to hTeloC at pH 5.5 resulted in an increase in T_m of 13°C.
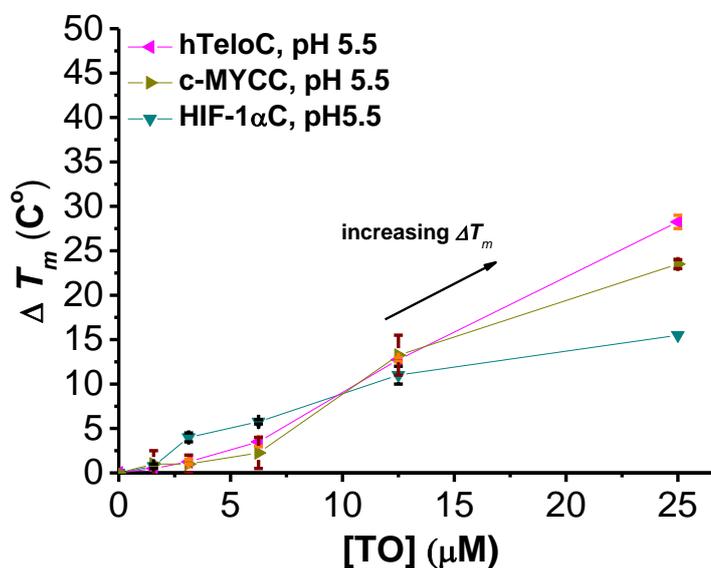
Figure 2.15   *ΔT_m vs concentration on different i-motif forming sequences with TO in 50 mM sodium cacodylate buffer at pH 5.5, TO concentration ranges from 0 to 25 μM. Fluorescent emission was measured at 533 nm.*

The effect of adding thiazole orange is even more noticeable at the transitional pH (Figure 2.16). For example, a large increase in melting temperature of HIF-1αC (ΔT_m = 12.5°C) was observed even with addition of just 3.1 μM TO. In general, with addition of 6.25 μM TO, the thermal stability of three i-motif forming sequences at their transitional pH was increased by 12-18°C (Figure 2.16). The differences in ΔT_m between the different i-motif forming sequences were minimal, suggesting that TO binds a region in the structure which is common to all of these different i-motif sub-types, ie. it does not bind the loop-region, where these i-motif forming sequences vary.
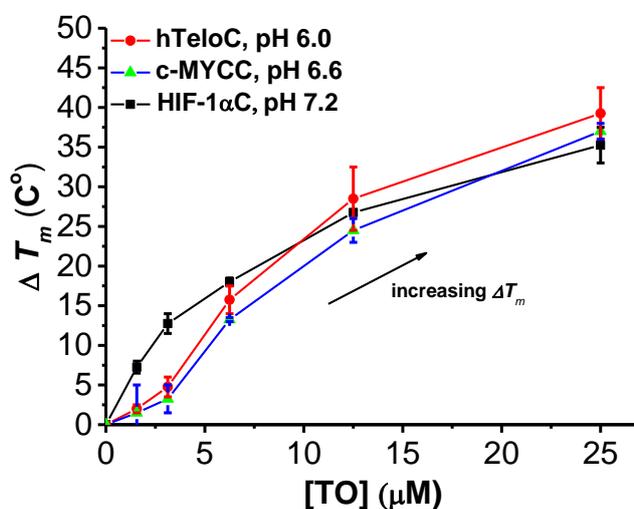
Figure 2.16   *ΔT_m vs concentration on different i-motif forming sequences with TO in 50 mM sodium cacodylate buffer at transitional pH , TO concentration ranges from 0 to 25 μM. Fluorescent emission was measured at 533 nm.*

Results from the experiments carried out on hTeloG and double helical DNA sequences at pH 7.4 also show some degree of stabilistion, as consistent with previous disclosures.[113,133,134] In general, duplex DNA structure was more prone to be stabilised by addition of TO (Figure 2.17). After adding 3.1 μM of TO, the ΔT_m was shown to be above 6°C. In contrast, induced thermal stabilisation of human telomeric G-quadruplex by TO was not that significant, for example a ΔT_m of only 5°C was observed even when 12.5 μM TO was added.
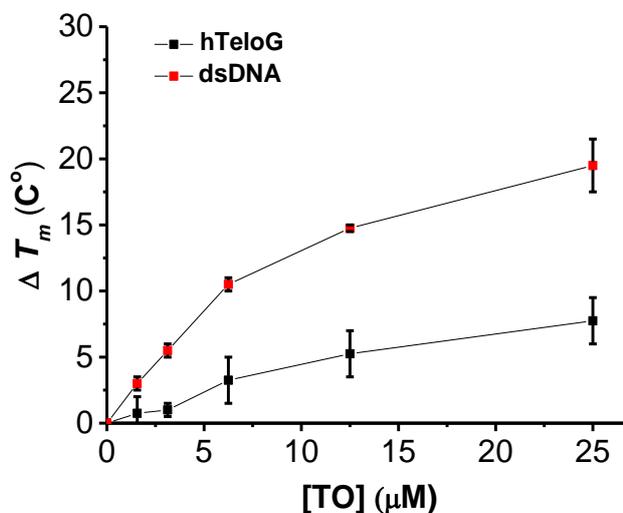
Figure 2.17   *FRET melting experiment on double-stranded DNA and G-quadruplex forming sequences with TO at pH 7.4 in 50 mM sodium cacodylate buffer. TO concentration ranges from 0 to 25 µM. Fluorescent emission was measured at 533 nm.*

2.3.4   Development of an i-motif FID titration assay

After confirming the binding and stabilising effects of TO on human telomeric i-motif, the next step was to develop an assay to explore whether pre-bound TO can be replaced by candidate i-motif binding ligands (Scheme 2.1).

Using the information obtained from the preliminary experiments, 1 µM of pre-annealed hTeloC i-motif in 10 mM sodium cacodylate at pH 5.5 was incubated with thiazole orange. The equivalence between hTeloC and thiazole orange was set as 2:1, consistent with the 2:1 binding mode determined in Section 2.3.2. After an equilibration period of 5 minutes, the samples were excited at 430 nm and the fluorescence spectra of the samples were recorded using a fluorimeter.

Initial studies were performed using the known i-motif binding ligand

mitoxatrone.[101,102] On addition of mitoxatrone, a significant loss in fluorescence emission at 450 nm was observed (Figure 2.18), indicative of displacement of TO from the DNA. Displacement at different ligand concentrations was calculated using Equation 2.3 (Figure 2.18).

Equation 2.3 $$D_x = 1 - \frac{F_x}{F_0} = 1 - \frac{F_{read} - F_{c_0}}{F_{refereence} - F_{c_0}}$$

Where D_x is TO displacement resulting from competitive substitution of ligand; F_0 = the fluorescence intensity of the TO-i-motif complex at 450nm in the absence of any ligands; F_x = the fluorescence intensity at 450 nm after titration with the candidate ligand.

By varying the concentration of ligand added, the ligand concentration where 50% of the TO was displaced by competitive ligand (DC$_{50}$), can be calculated from a plot of D_x against [ligand] (Table 2.2). The results show that $^{i\text{-motif}}$DC$_{50}$ for mitoxantrone was 1.92 ± 0.15 μM. Additionally, other ligands previously found to bind i-motif in the Waller group: tilorone, tamoxifen and tyrothyrcin, were also used in the assessment (Figure 2.19). Tilorone exhibits an $^{i\text{-motif}}$DC$_{50}$ of 2.47 ± 0.44 μM, which is slightly weaker than that of mitoxatrone. The other two compounds have $^{i\text{-motif}}$DC$_{50}$ higher than 5 μM, suggesting far weaker binding properties compared to mitoxatrone and tilorone. This suggests that the TO FID assay might be able to adequately act as a threshold to eliminate weaker i-motif binders.
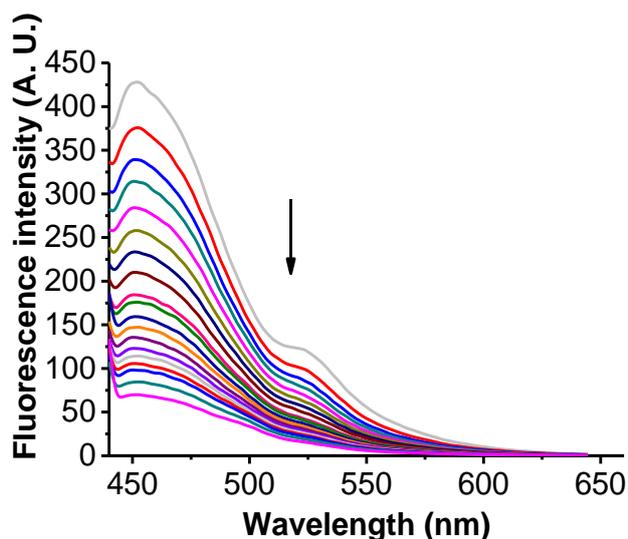
Figure 2.18   *Fluorescent intercalator displacement assay carried out using hTelo i-motif forming sequences and mitoxatrone at pH 5.5 in 10 mM sodium cacodylate buffer. [hTeloC] = 1 μM, [TO] = 2 μM, [mitoxatrone] = 0 – 5 μM. The sample was excited at 430 nm.*
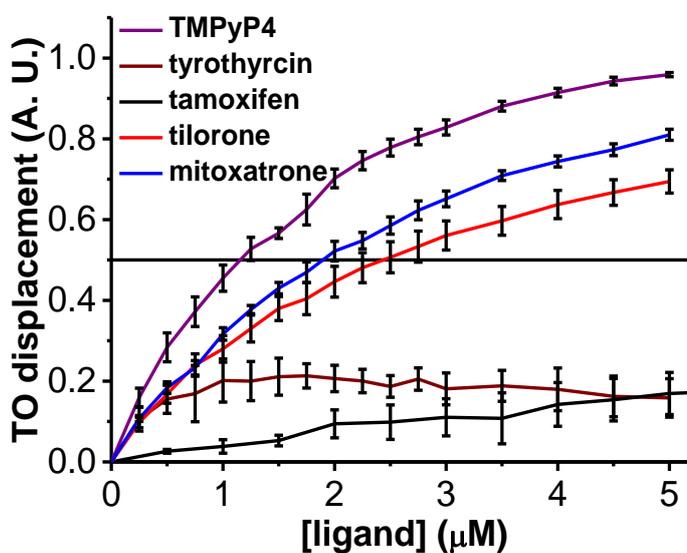
Figure 2.19   *TO displacement calculation based on hTelo i-motif FID assay on mitoxatrone, tilorone, tamoxifen and tyrothrycin at pH 5.5 in 10 mM sodium cacodylate buffer. [hTeloC] = 1 μM, [TO] = 2 μM, [ligand] = 0 – 5 μM. The sample was excited at 430 nm and the fluorescence emission was measured at 450 nm.*

Compound	i-motifDC$_{50}$ (μM)
TMPyP4	1.15±0.15
Mitoxatrone	1.92±0.15
Tilorone	2.47±0.44
Tamoxifen	N/A
Tyrothrycin	N/A

Table 2.2 TO displacement calculation based on hTelo i-motif FID assay on mitoxatrone, tilorone, tamoxifen and tyrothrycin at pH 5.5 in 10 mM sodium cacodylate buffer. [hTeloC] = 1 μM, [TO] = 2 μM, [ligand] = 0 – 5 μM. The sample was excited at 430 nm and the fluorescence emission was measured at 450 nm.

Additionally, TMPyP4, another model compound which binds to the hTelo i-motif,[135] was also studied. Although it showed very good binding affinity in the FID assay (Figure 2.20) with a i-motifDC$_{50}$ of 1.15 ± 0.15 μM. Further investigation indicated that the drop of fluorescence was caused by the UV absorption of TMPyP4 instead of thiazole orange displacement (Figure 2.21). This case indicated that UV absorption of the ligand itself can be a source of a false positive hit due to a 'self-quenching' event. Thus checking the UV absorption of ligand hits would be necessary for the FID assay to be reliable.
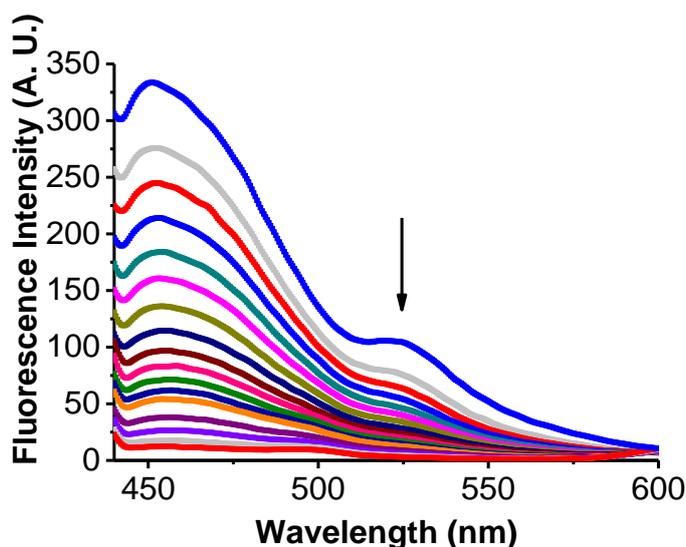
Figute 2.20   *Fluorescent intercalator displacement assay carried out using hTelo i-motif forming sequences and TMPyP4 at pH 5.5 in 10 mM sodium cacodylate buffer. [hTeloC] = 1μM, [TO] = 2μM, [TmPyP4]= 0 - 5μM. The sample was excited at 430nm and fluorescence signal dropped as more TmPyP4 was titrated.*
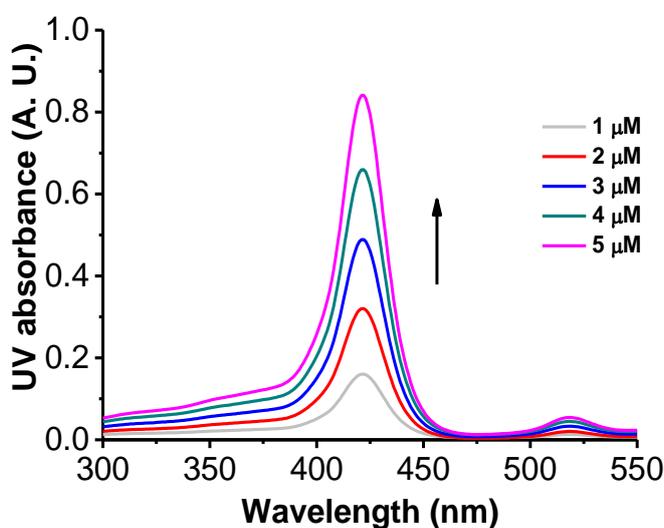
Figure 2.21   *UV-vis spectrum on TmPyP4 at pH 5.5 in 10 mM sodium cacodylate buffer*

In conclusion, non-fluorescent thiazole orange was found to become fluorescent when titrated with pre-formed i-motif structures at acidic pH. The fluorescent

emission was detected at 450 nm when the sample was excited at 430 nm. Although the Stoke's shift in i-motif assay is very narrow when compared to G-quadruplexes (a Stoke's shift of 20 nm compared to 38 nm from the original G-quadruplex-FID assay[112]), it is still possible to conduct FID on i-motif with thiazole orange. Four previously identified i-motif compounds were tested in a trial i-motif FID assay. Their relative binding affinity towards human telomeric i-motif can be quantified by calculating their $^{\text{i-motif}}DC_{50}$ and the results from i-motif FID assay were consistent with studies carried out previously.[101]

2.3.5   Development of a high-throughput screen based on FID i-motif assay (HT-IM-FID)

After confirming the result from the i-motif-FID assay using a fluorimeter, the next stage was to develop TO as a fluorescent indicator in a high-throughput screening FID assay in a microplate format. To establish a low-cost and viable screening method using FID aganist human telomeric i-motif, a microplate and plate-reader were used. A 384-well PCR black microplate was chosen due to its relatively small well volume to reduce the amount of material consumed. The Gen-Plus compound library, was obtained from Microsource Discovery Systems Inc (a gift from Prof. Rob Field, John Innes Center) and contains 960 ligands, including the previously identified mitoxatrone, tilorone, tyrothrycin and tamoxifen. In the screening, 0.5 µM (1 equivalents) pre-folded hTelo i-motif was mixed with 1.0 µM (2 equivalents) TO in the sodium cacodylate buffer in the well and the sample was made to 40 µL. Then 2.5 µM (5 equivalents) ligand from the compound library was added into the well to screen. The microplate was excited at 430 nm and the read of the signal was defined as the intergration of fluorescent intensity from 460 to 480 nm. The value of TO displacement (D_x) was calculated *via* Equation 2.3 shown before. Then the compounds were ranked according to their *Dx* (Table 2.3) to identify the strong i-motif binders. The result from the primary FID assay was then compared with a FRET-melting based ligand screening perform in the group before[101] and futher

additional experiments performed using SPR.

Position	Compound name	Structure	TO displacement (%)
P9H8	Mitroxatrone hydrochloride		85.8
P9C11	Alexidine hydrochloride		82.5
P11H11	Tilorone		70.1
P2B6	Tobramycin		48.6
P1A3	Chlorhexidine		42.3
P3C9	Phenazopyridine hydrochloride		30.5
P7B10	Amodiaquine dihydrochloride		29.7
P1A11	harmalol hydrochloride		26.7
P4A8	Quinalizarin		22.2
P5B11	Minocycline hydrochloride		21.0
P3A4	tyrothrycin		18.6
P2H4	Cadmium acetate		16.7

Table 2.3   FID screening result on NIC library (960 compounds). [hTeloC] = 0.5 μM, [TO] =1.0 μM, [ligand] = 2.5μM. Buffer: pH 5.5, 10 mM sodium cacodylate. Excitation range: 415 to 430 nm, emission recorded: 445 to 495 nm. The data was recorded as the integration of emission curve from 445 to 495 nm.

In the FID screening of 960 ligands (see Table 2.3), **mitoxantrone** and **tilorone**

were confirmed as top hits with FID displacement of over 80%, followed by chlorhexidine and alexidine hydrochloride (83% and 42%, respectively). However, the later are known nucleotide condensing reagents,[101] so they were omitted from further studies. Tyrothricin and tamoxifen were both found to displace TO under the conditions of the experiment but did not meet the 50% disppacement criteria set for hits. This is consistent with their low-affinity towards human telomeric i-motif described in Section 2.2.4.

Other DNA alkylating reagents (such as cisplatin and carboplatin) present in the library previously gave rise to false positive hits in FRET-melting experiments;[101] however, they did not show any activity in the FID assay. This is because the two screens are measuring different DNA interaction properties. In the FRET-melting experiments, pre-annealed i-motif is heated to 95°C in order to melt the DNA and the change in melting temperature on addition of ligands is used as a measure of interaction. In this case, DNA alkylating reagents can hold the DNA conformation by cross-linking the structure. However, in the FID assay the mechanism is based on competitive binding with pre-bound thiazole orange. This will only give a result if the molecule of interest interacts in the same way. This means, alkylating reagents will not give a postitive result in these experiments. However, it does mean that if the tested ligand binds elsewhere on the structure, ie a place not occupied by TO, this may result in a false negative.

Interestingly, tobramycin was found to be a hit using the FID screening method, but not in the previously used FRET assay.[101] This might indicate that if DNA FRET-melting and FID screening were performed side by side, different properties could be revealed on a large scale with ease. Further studies to characterise tobramycin are discussed in Section 2.3.8.

2.3.6 Assessing the high-throughput screening with Z′ factor

The quality of the TO-based FID high-throughput screening can be quantified using

the Z′ factor (Equation 2.4) as demonstrated below.[136,137] Z′ factor gives an indication of the statistical effect size, ie whether a type of response in an assay is large enough to be called a "hit" or worth repeating.

Equation 2.4
$$Z' = 1 - \frac{3 \times (\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

Here, σ_p = the standard deviation of positive control; σ_n = the standard deviation of negative control; μ_p = the mean value of positive control; and μ_n = the mean value negative control

When the Z′ value is between 0.5 and 1.0, this gives a strong indication that the positive responses from the assay are more likely to be actual hits.[137] A Z′ value lower than 0 indicates that the signal from the positive and negative controls overlap and will not provide enough distinction to identify hits; Z′ values between 0 and 0.5 are marginal. In this assay, as we already know tobramycin and mitoxatrone were an effective hit in FID screening, it was possible to use them as a positive control; the wells with only TO and hTeloC were then set as a negative control. The calculated values of Z′ in this assay were found to be 0.66 for mitoxantone and 0.97 for tobramycin, which are both within the range from 0.5 to 1.0, indicative that the result of the screen is a true hit worth following up with further experiments.[137]

2.3.7   Assessing the hit compounds with SPR

In collaboration with Dr. Clare Stevenson at the John Innes Center, the hit compounds were studied using SPR at a single concentration. A moderately high concentration (50 μM) was used to give an indication for a wide range of compounds with different binding capabilities. In this situation, high affinity ligands would be expected to provide a very high response on association with the target oligonucleotides immobilised on the sensor chip; in contrast, low affinity compounds would give a poor response. Also, a concertation higher than 50 μM

would be reaching the solubility limit for some of these compounds.

In this experiment, 50 µM of the 'hit' ligands in buffer were measured using a streptavidin coated chip with immobilised biotinylated duplex DNA, hTeloC and c-MYC i-motifs at pH 5.5. The responses were recorded at equlibrirum and compared to the predicted binding response (R_{max}) calculated with a 1:1 binding stoichiometry (Table 2.4, Equation 2.5).

Equation 2.5 $$R_{max} = \frac{MW_{ligand}}{MW_{oligo}} \times RU_{oligo} \times i.$$

Where: MW_{ligand} = the molecular weight of tested ligands; MW_{oligo} = molecular weight of the immobilised oligonucleotides on the chip. RU_{oligo} = the average response in the SPR experiments in the absence of ligand. i = stoichiometry of binding.

As shown in Table 2.4, the percentage of predicted maximium response (%R_{max}) is an indicator for binding affinity. Of all the compounds tested, mitoxatrone has the highest %R_{max} (2125% for hTeloC, 2534 % for c-MYCC and 4894% for double helical DNA), followed by tobramycin (325% for hTeloC, 419 % for c-MYC and 810% for double helical DNA). Tilorone was found to bind to all three DNA types as well, but not as strongly as tobramycin and mitoxatrone (116% for hTeloC, 121% for c-MYC and 233% for double helical DNA). The fact that tilorone possesss higher affinity towards double helical DNA rather than c-MYCC and hTeloC i-motifs was consistent with previous studies using SPR ($^{hTeloC}K_d$ = 158 µM vs $^{duplex}K_d$ = 14.7µM).[101] Other ligands such as amlodiaquine, harmalol, minocycline and tyrothrycin failed to exert similar affinity towards the i-motifs compared to mitoxatrone and tobramycin. This is consistent with the fluorescent intercalator screening in Section 2.2.5, providing some validation to the procedure and technique used.

Ligands	%R_{max} (%)		
	hTelo i-motif	c-MYC i-motif	Duplex DNA
mitoxantrone	2125	2534	4894
tobramycin	325	420	810
tilorone	116	121	233
phenazopyridine	37	70	136
minocycline	29	49	95
amodiaquine	14	31	59
tyrothrycin	6	14	26
quinalizarin	-0.37	36	69
harmalol	-27	64	123

Table 2.4   SPR screening on the hit compounds from the FID screen against hTelo i-motif, c-MYC i-motif and duplex DNA. [ligands] = 50 μM Running buffer: [sodium cacodylate] = 10mM, [NaCl] = 100 mM, 0.05% tween, pH 5.5

Tobramycin was subjected to further SPR binding experiments across a wider concentration range and the dissociation constants were found to be 23 ± 3.7 μM, 22 ± 6.9 μM and 13 ± 3.4 μM for hTeloC, c-MYCC and duplex DNA respectively (Figure 2.22), calculated using a 1:1 binding model (Equation 2.6).[132]

Equation 2.6

$$\theta = \frac{R_{eq}}{R_{max}} = \frac{nK_1[Ligand]}{1 + K_1[Ligand]}$$

Where: θ = the fraction of binding; R_{eq} = the response at equilibrium from the sensorgram; R_{max} = response at saturation.[132]
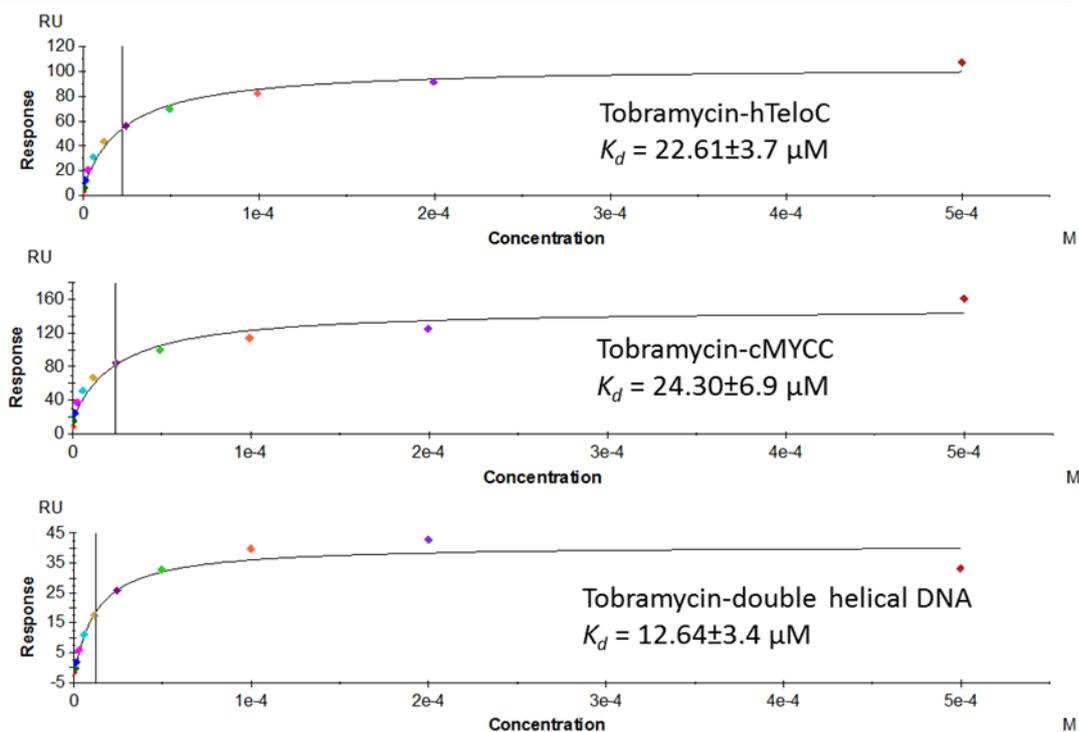
Figure 2.22   *SPR titration of tobramycin against hTeloC, c-MYCC and duplex DNA at pH 5.5 Running buffer; pH 5.5, [sodium cacodylate] = 10mM, [sodium chloride] = 100mM, 0.05% tweet 20, 5% DMSO tweet 20, 5% DMSO*

The K_d for tobramycin binding to the human telomeric i-motif (23 ± 3.7 μM) was the same order of magnitude as mitoxatrone (K_d = 12 ± 3.0 μM).[102] However, although tobramycin binds i-motif structures well, it also binds duplex DNA with almost 2-fold higher affinity. A plausible explaination is that the multiple amino groups on the tobramycin will be protonated in acidic conditions, providing multiple positive charges which can then interact with oligonucleotide sugar phosphate backbones indiscriminately.

2.3.8 Further studies on tobramycin

In the studies above, it was confirmed that tobramycin was the effective hit in the high-throughput screening. Tobramycin is an aminoglycoside used to treat airway infection caused by *Pseudomonas aeruginosa*.[138] To provide another supportive technique, tobramycin binding to the hTeloC i-motif was measured using circular dichroism as well (Figure 2.23). Aliquots of tobramycin were titrated into 10 µM pre-formed human telomric i-motif at pH 5.5. Since tobramycin is a chrial molecule, the molar ellipticity signal of tobramycin was measured and then subtracted from the subsequent spectra. As shown in Figure 2.16, the positive peak at 289 nm and the negative peak at 255 nm indicates intramolecular human telomeric i-motif structure.[86]
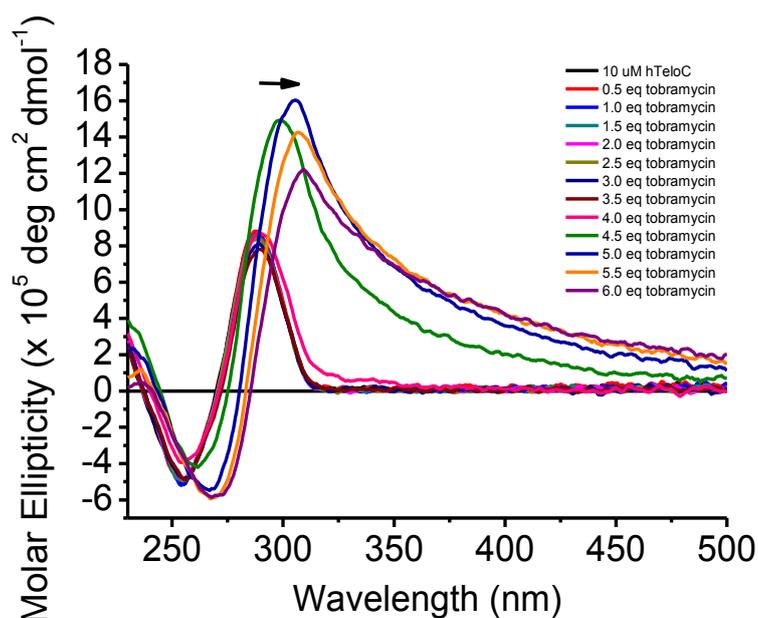
Figure 2.23   *Circular dichroism titration of hTeloC DNA sequences with tobramycin at pH 5.5 10mM sodium cacodylate buffer (tobramycin signal subtracted). The molar ellipticity signal was recorded from 500 nm to 230 nm.*

On addition of tobramycin, the signal is observed to shift towards longer

wavelengths by 20 nm. This is consistent with tobramycin binding i-motif. The changes in the spectrum are not consistent with conformational changes on the i-motif structure, or conversion to a hairpin or unfolded species.[49,56] The increased signal between 300 and 500 nm is unusual, but may arise from the chiral nature of the ligand. Despite substracting control spectra containing equivalent proportions of ligand, these signals remain, indicating perphaps a form of induced signal which arises on binding.

In a previous study of tobramycin-DNA interaction, DNA forms columnar close-packed hexagonal DNA complexes in the presence of tobramycin.[139] Tobramycin is electrostatically bound in these tobramycin-DNA complexes due to its multivalent catonic charges at pH 7.0.[139] Thus as a negatively-charged polyelectrolytes in a packed formation, human telomeric i-motif may also binds to tobramycin and the i-motif structures may be unfolded into other formations due to the presence of tobramycin as well.

2.4   Conclusions

In this chapter, a fluorescent intercalator displacement assay against human telomeric i-motif was developed. In the development of the assay, several DNA-binding fluorescent probes were screened with the C-rich sequence from the human telomere, hTeloC. It was found that thiazole orange exhibited the ideal fluorescent 'light-up' property when bound to hTeloC at pH 5.5. Subsequently, thiazole orange binding to the hTeloC i-motif was studied intensively using UV-vis spectroscopy, Job's analysis, FRET-melting and circular dichoism experiments. These revealed thiazole orange binds to hTeloC at acidic pH with a 2:1 stoichiometry, with a high-affinity binding site and a weaker secondary site (K_{d1} = 3.61 µM and K_{d2} = 75.6 µM). CD experiments indicated the conformation of human telomeric i-motif was not altered when bound to thiazole orange under the conditions used in the FID assay. FRET-melting experiments showed the thermal stability of several i-motif

structures increased upon addition of thiazole orange indicating that TO is able to stabilize multiple i-motif structure types. Fluorescent intercalator displacement experiments were developed using some known i-motif binding ligands previously discovered in the Waller group as positive controls.[101] Finally, a high-throughput i-motif ligand screening method based on fluorescent intercalator displacement was established. A library of 960 known biologically active compounds were tested and the screening result was verified via SPR. The TO displacement assay offers an alternative to using FRET-based DNA melting experiments for discovery of i-motif binding ligands.

Tobramycin was found to be an effective i-motif binding ligand in the high-throughput screening. It binds to human telomeric i-motif with a K_d of a similar magnitude to that of known i-motif binding compound mitoxatrone.

The FID assay as a primary ligand screening method has its own advantages as well as some flaws. In principle, the FRET-bsed DNA melting assay is focused on measuring the the stabilising effect of the tested compound while FID assay was designed to measure ligand binding. Both the FID titration and high-throughput screening experiments are carried out under room temperature, which remove the possibility of ligand degradation caused by heating in DNA-melting based assay. It is much more cost-effective compared to the FRET-melting assay, because there is no requirement for specialised DNA synthesis, furthermore FID experiments use unmodified DNA, and are therefore not affected by the prescence of fluorophores on the DNA. Unfortunately, like all other screening methods, FID has limitations as well. Firstly, FID is not able to distinguish nucleic acid condensing reagents from hit binding compounds but may also be used as a tool to assess nucleic acid condensing reagents. Additionally, its accuracy can be easily affected by compounds which either have UV-absorption or fluorescent emission within the range of 420 to 490 nm. There is also the possibility that tested ligands may have a higher affinity for the thiazole orange itself rather than the i-motif structure, thus

causing false hits. Also if the tested ligand binds i-motif in a region where TO does not, these will not be picked up. Neverthless, every method has its advantages and disadvantages and FID can be a screening and analysis method to support existing i-motif characterization methods, such as circular dichroism and FRET melting experiments.

In the context of tobramycin, it was the first time that an aminoglycoside was shown to be an i-motif binding ligand. However, it was also shown that tobramycin was not able to distinguish duplex DNA with i-motif structures. Thus, tobramycin would not be a suitable ligands for use in i-motif binding studies. However, it is possible to use medicinal chemistry approaches to develop a tobramycin based ligand to target i-motifs with improved selectivity.

Chapter 3   Synthesis of Fluorescent Calix[4[arene Derivatives and Investigation of Binding to DNA

3.1 Introduction

Very few molecules have been documented in the literature to interact with i-motif.[40] However, it is widely known that calixarene derivatives have multiple applications in medicinal chemistry and biology.[140] Herein, an attempt to use calixarene derivatives to interact with DNA secondary structrures is described.

3.1.1 Calix[4]arenes

Calixarenes are a family of macrocyclic oligomers arising from the by-product of phenol-aldehyde resin synthesis.[141,142] The word 'calixarene' (calix: Greek for cup) was coined by Gutsche, indicating a bowl-shaped supramolecular macrocycle with aryl groups in the cyclic array.[143] The number of aromatic rings can be any number above 3 and their structural flexibility increases with increasing ring number.[144] But commonly-used calixarenes usually contain 4 to 8 aromatic rings. The number of rings is given as [n] between '*calix-*' and '*-arene*'.

In the 1940's, an efficient base-induced reaction of p-alkyl phenols with formaldehyde yielding *p-tert*-calix[4]arene was discovered by Zinke and Ziegler, marking the beginning of calixarene chemistry.[142] A calix[4]arene scaffold has four phenol rings linked by four methylene bridges, resembling a cup-shaped molecule bearing a higher rim (wider edge) and a lower rim (narrower edge); both these aspects can also be functionalised. Phenol member rings can be replaced by other substituted phenols[145], polycyclic or even heteroaromatic rings, while methylene bridges can also be replaced by ethers, thiolethers, or tertiary amines according to the different needs. The central cavity inside a calix[4]arene is a hydrophobic space suitable for transporting lipophilic reagents.[146,147]

Calix[4]arenes can adopt four distinct conformers: *cone*, *partial cone*, *1,3-alternate* and *1,2-alternate*; this is due to the free rotation of the carbon bonds on the methylene bridges (Figure 3.1).[148,149] A fixed conformation of calix[4]arene can be

achieved *via* metal templating, followed by alkylating the phenol group to create steric hindrance.[150]
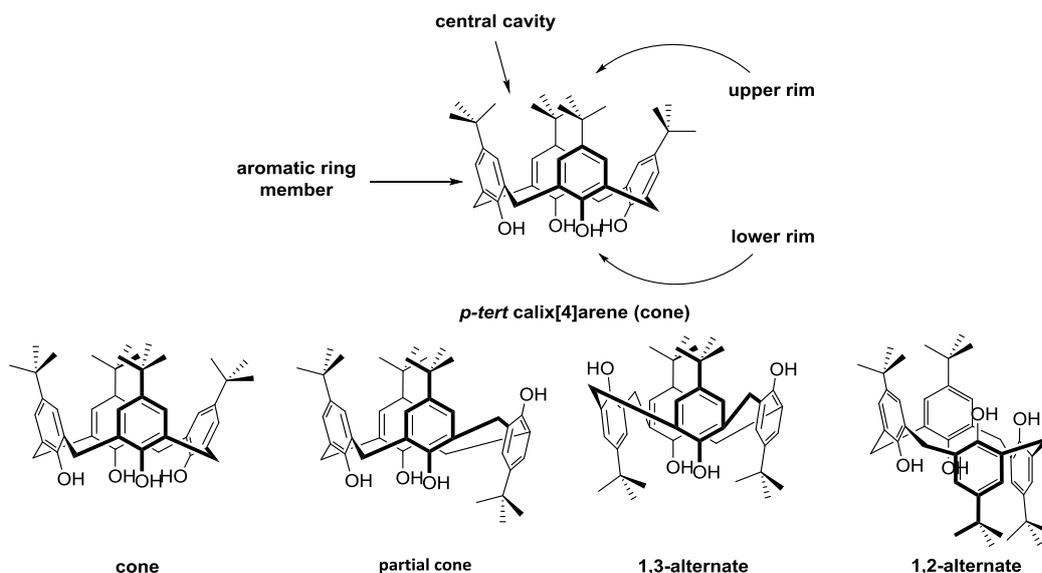
Figure 3.1   p-tert calix[4]arene and its conformer.

Compared to other macrocycles, calix[4]arene platforms, especially the water-soluble versions, have indispensable advantages; including synthetic availability, low toxicity, negligible immunogenicity and efficient regioselective synthesis.[151] Calixarenes also have various biological applications, including being used as antineoplastic, antiviral, antibacterial and antifungal therapeutics,[152] and also in DNA chips[153,154] and biosensing technology.[152,155,156]

Due to the many applications of calix[4]arene derivatives in pharmaceutical studies, these will be described separately in detail.

3.1.2   Calix[4]arenes in drug delivery

Functionalised calix[4]arene derivatives have been investigated both as drugs and as drug delivery vectors. Their delivery mechanisms include host-guest inclusion, micelle and vesicle formation.[157] Host-guest inclusion utilises the complimentary hydrophobicity of drugs and the cavity of calix[4]arenes. The most commonly used amphiphilic calixarene in drug delivery is *p*-sulphonatocalixarene. *p*-

sulphonatocalix[4]arene **1** has been reported by different groups to encapsulate small molecules such as dinuclear platinum complex, topotecan, irinotecan and Vitamin B_6 (Figure 3.2).[146,158,159]
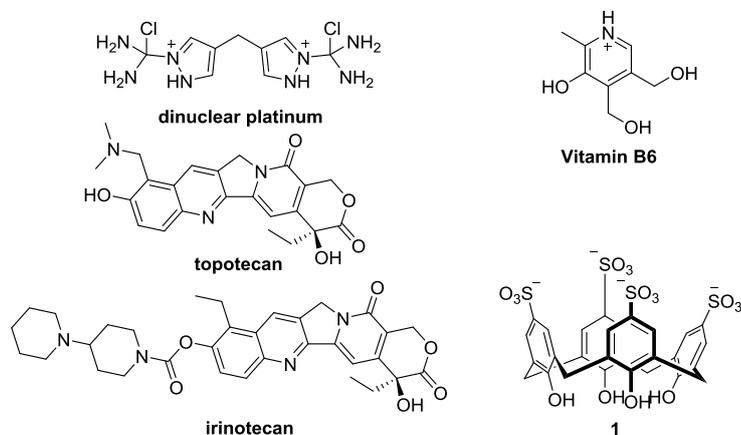
Figure 3.2   *Examples of drugs which have been encapsulated in calixarene **1**.*

Amphiphilic calix[4]arenes have the ability to self-assemble. They can spontaneously aggregate into micelles or vesicles under the right conditions to host lipophilic molecules (Figure 3.3).
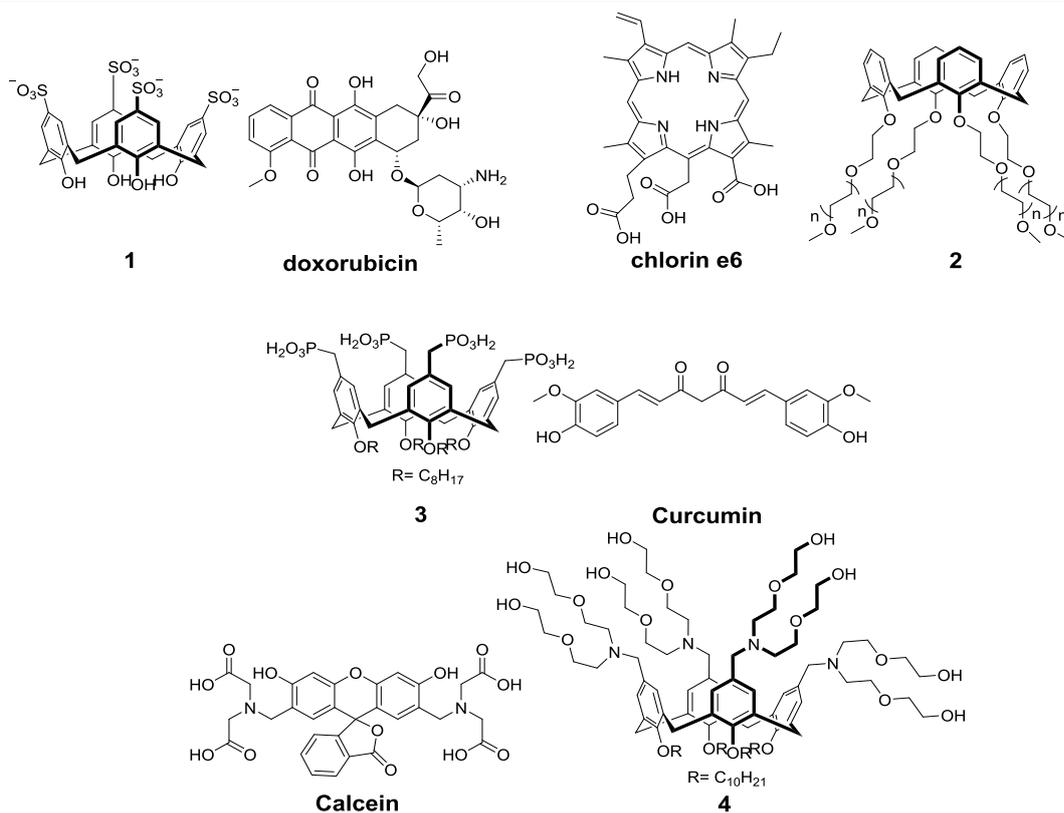
Figure 3.3   Examples of calix[4]arene-based drug delivery strategies dependant on forming micelles

In terms of micelles, Zhu *et al.* described a PEGylated-calix[4]arene **2** micellar drug delivery system to encapsulate the photosensitizer chlorin e6.[160] Another example was reported by Raston *et al.*, using phospholipid calix[4]arene **3** micelles to deliver the antioxidant curcumin under conditions of oxidative stress.[161] Apart from micelles, vesicles formed by amphiphilic calix[4]arenes are larger and often reported to carry hydrophilic ligands. Lee *et al.* reported tunable and stable vesicles based on amphiphilic calix[4]arene **4** to the host hydrophilic dye calcein. These vesicles dissipate to release calcein when the pH value drops from 7 to 5.[162] A more recent and sophisticated vesicle system was reported by Liu *et al.*, where *p*-sulphonatocalix[4]arene **1** vesicles are loaded with doxorubicin, 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS) and cationic protein protamine. The protein guest can be cleaved by trypsin, a natural serine protease. Disaggregation of the resulting vesicles is triggered by addition of trypsin. Controlled release of doxorubicin *via*

this vesicle system has been observed in both cell and mouse models.[163]

3.1.3 Calix[4]arenes in gene delivery

Gene therapy depends on transporting target genes into cells efficiently without causing elevated cytotoxicity. Calix[4]arenes serve as a perfect non-viral, preorganised platform to aid this process. In the field of calixarene-DNA interactions, cationic amphiphilic calixarene conjugates are commonly used as gene delivery agents.[151] In order to interact with DNA/RNA, calix[4]arenes are modified with positively-charged functionalites on one rim. The electrostatic attraction between modified calix[4]arenes and the negatively-charged phosphodiester backbone causes the condensation of DNA, facilitating its uptake by cells. The lower rim (narrow edge) is normally functionalised with long aliphatic chains to minimise the critical micelle concentration (CMC). In order to bind to DNA, calix[4]arenes are required to be in the cone conformation to maximize the unequal distribution of hydrophilic/hydrophobic characteristics.

The most popular hypothesis starts with formation of amphiphilic calix[4]arene micellar nano-particles, followed by the DNA wrapping around the pre-formed particles.[151] In this way, amphiphilic calix[4]arene micelles could be considered acting like artificial histones. More densely packed micelles are favored due to the maximal presentation of positively-charged groups for ion pairing or hydrogen bond formation (Figure 3.4).[151,164] Another proposed mechanism suggests that individual calix[4]arene amphiphiles bind directly to minor grooves or the phosphate backbone of duplex DNA structures. Similarly, this process involves ion pairing and hydrogen bond formation.[126] The DNA-binding calixarenes attract each other and assemble into DNA-calixarene complexes (Figure 3.4).[165]
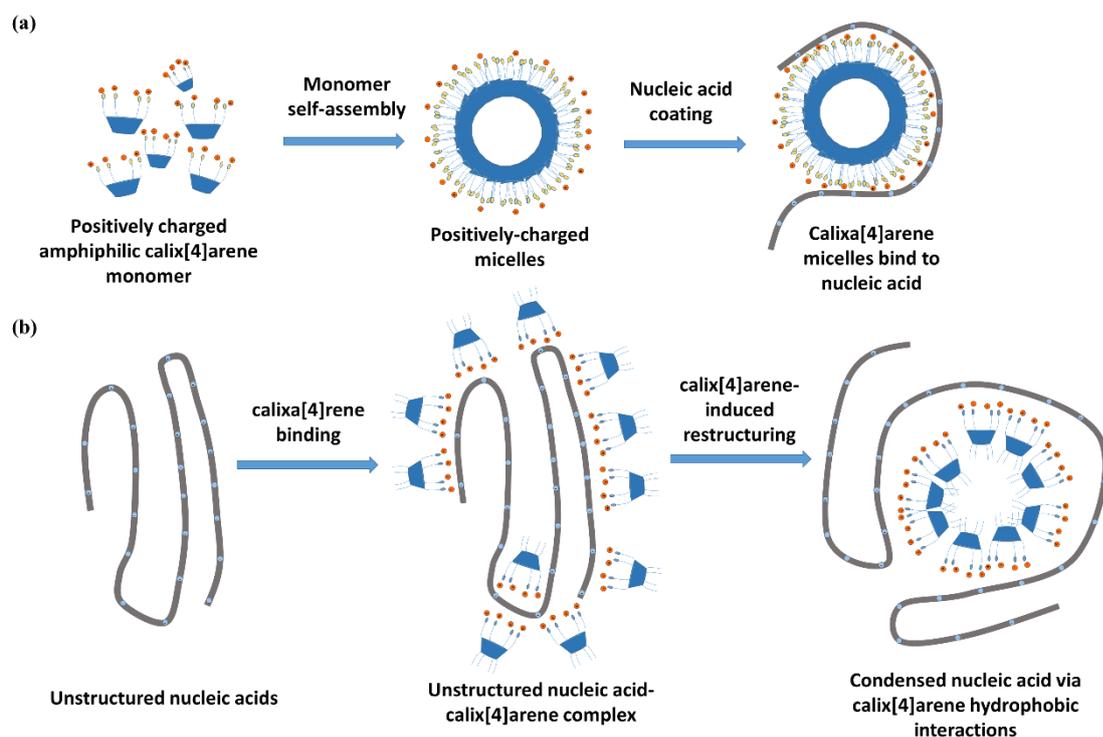
Figure 3.4   *Proposed mechanism of calix[4]arene-induced nucleic acid condensation. (a) Mechanism of nucleic acid binding after micelle formation.[151] (b) Mechanism of calix[4]arene-nucleic acid binding before nucleic acid condenation.[165]*

In order to interact with oligonucleotides, guanidine, amine, imidazolium and polyamine groups can be inserted on the cone-shaped platforms (Figure 3.5).[151,165-169]
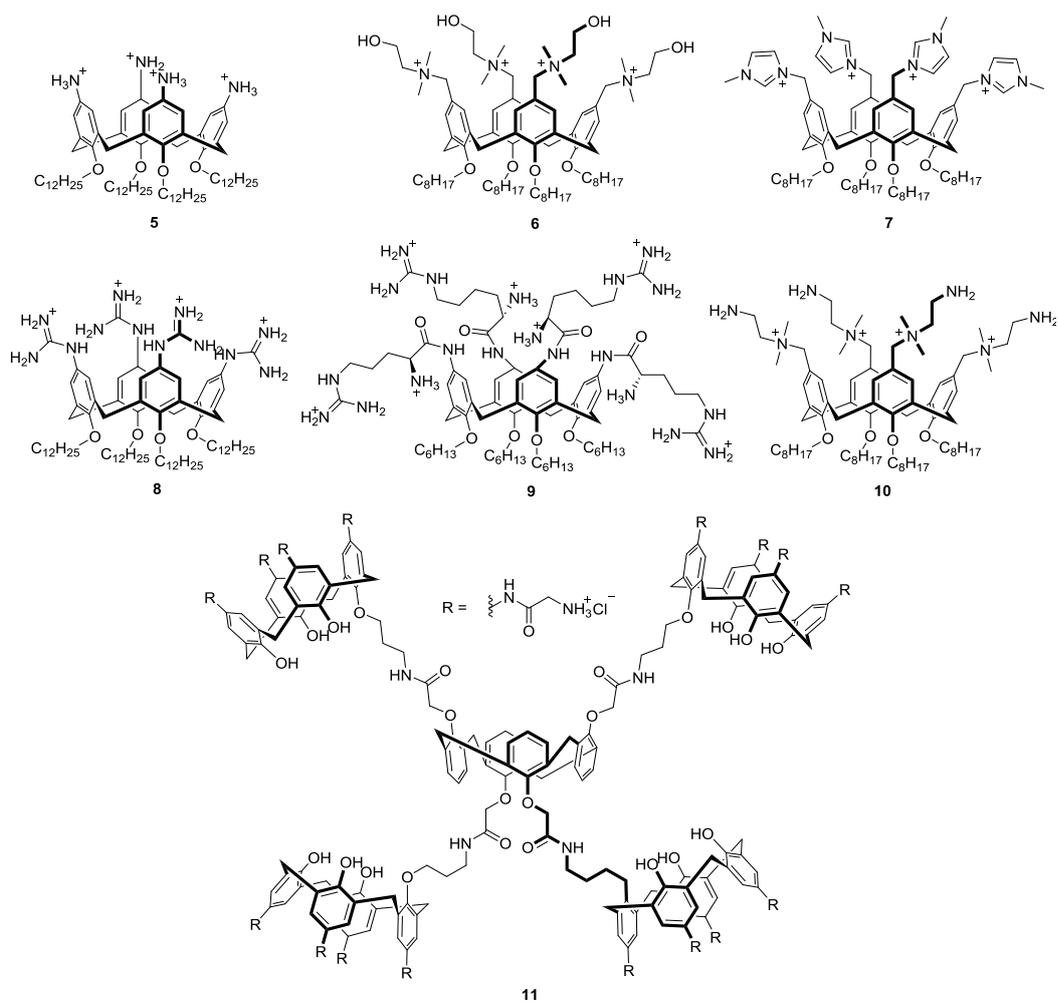
Figure 3.5   *Examples of calix[4]arene-based amphiphiles for gene delivery*

Such modified ligands are capable of transferring DNA (both linear and plasmid) across the cell membrane. As an example, amino-functionalised calixarene **5** was found to form solid lipid nanoparticles via layer-by-layer assembly and was shown to transfect plasmid DNA into the Madin-Darby canine kidney cell line.[166] Other positively-charged, nitrogen containing calixarenes include tetraalkylammonium-functionalised calixarene **6** and methylimidiazolium calix[4]arene **7**, which possess long octyl chains on the lower rim. These are able to self-assemble into micelles of 6 nm in diameter at low CMC (48 μM and 10 μM, respectively). In transfection experiments, calixarenes **6** and **7** were used with plasmid DNA encoded with luciferase. Effective cell transfection was observed in both human cervical cancer cell and monkey kidney fibroblast-like cell lines.

Ungaro's group developed a series of guanidium functionalised calixarenes to study calixarene-induced nucleic acid condensation. Calixarene **8** exhibited efficient plasmid transfection in human rhabdomyosarcoma cell lines.[165] Later Ungaro's group developed arginine grafted tetralysinocalix[4]arene **9** to achieve 55% to 82% DNA transfection in human rhabdomyosarcoma, cervical cancer cell, mouse neuroblastoma and equine adipose-derived stem cell lines.[167]

Other notable examples of gene delivery include the cationic choline-conjugated calixarene **10**, which was able to form virus-sized nanoparticles with plasmids for gene delivery.[168] Apart from small calixarene conjugates, dendritic calixarene-based scafolds were also used in gene delivery. The amino-functionalised multicalix[4]arene **11** was also demonstrated to transfect genes into Chinese hamster ovary cell with negiligable cytoxicity.[169]

3.1.4   Calix[4]arene derivatives as drugs

Apart from delivering drugs and nucleic acids into the cells, calix[4]arenes derivatised with functionalites to achieve receptor or enzyme binding have also been developed. Existing examples have expanded the use of calixarenes in targeting various diseases such as viral infections, bacterial infections and malignancy.

Calix[4]arenes tethered with glycosides

Lectins are a family of specific sugar-binding proteins expressed on the cell surface in various organisms. They are involved in many biological recognition processes in nature. The binding affinity of polysaccharides towards lectin can be enhanced by the number of saccharide units presented *via* non-covalent interactions; this is known as the 'glycoside cluster effect'.[170] The calix[4]arene scaffold itself serves as a perfect dendrimer platform to present a sufficient bulk of diverse carbohydrate motifs for lectin recognition.[171]

In an antiviral study, calix[4]arene **12** conjugated with *N*-acetylneuraminic acid *via* click chemistry (Figure 3.6) was shown to have a 83-fold increase in binding affinity on haemagglutinin expressed on the surface of influenza virus A.[172,173]
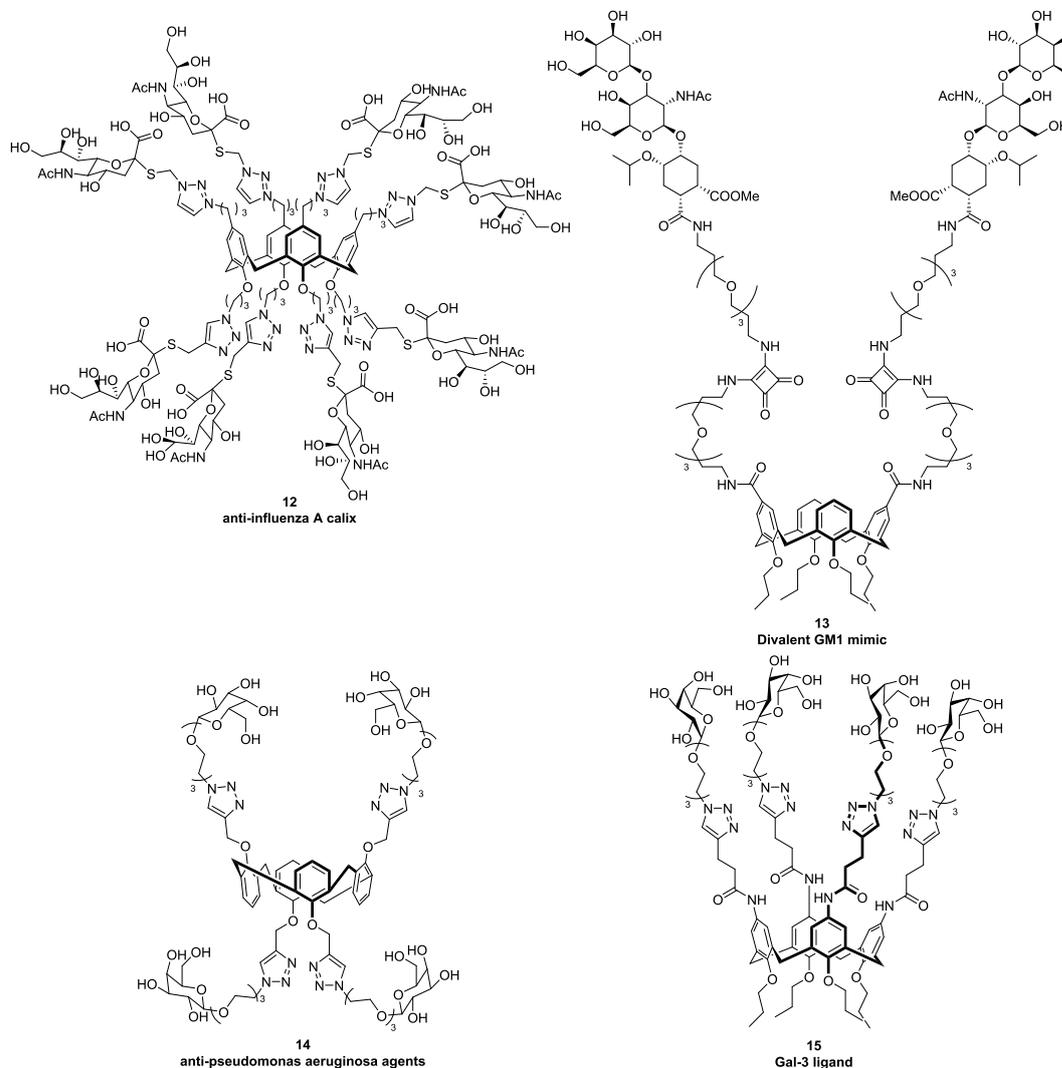
Figure 3.6   Examples of calix[4]arenes tethered with carbonhydrate motifs.

In Ungaro's group, the study of glycocalix[4]arene was focused on treating cholera. Glycocalix[4]arene **13** (Figure 3.6) bearing ganglioside GM1 mimic was synthesed and tested with cholera toxin, a protein complex[174] secreted by bacterium *Vibrio cholerae*. This protein complex is responsible for massive fluid release in a cholera infection.[175] It was found that its affinity towards cholera toxin was increased by 4-fold compared with its mimicking target, GM1 pentasaccharide.[176]

In the study of lung infection, a calix[4]arene tethered with galactose was used as an anti-adhesive agent aganst *Pseudomonas aeruginosa*. Matthews and Vidal reported a calixarene conjugate **14** with a K_d in the nanomolar range against LecA, a soluble lectin expressed by *Pseudomonas aeruginosa*.[177] Futher studies showed that designed glycoclusters on the calixarenes prevented lung infection by *Pseudomonas aeruginosa* in a mouse model.[178]

In 2014, another example using glycocalix[4]arene chemistry was published; lactosyl and galactosyl motifs were attached on either *cone* calix[4]arenes or the *1, 3-alternate* structure. The resulting gylcocalix[4]arenes were then tested with Gal-3 protein, a lectin heavily involved in cancer metastasis and migration. The study showed that cone-shaped lactosylcalixarene **15** (Figure 3.6) possessed the highest affinity for Gal-3 lectin.[179]

Calixarene in protein-binding studies

Calix[4]arenes in the *cone* conformation can be decorated with peptide motifs for peptide-protein interactions; as a result, modified calixarenes have the potential to block protein-protein interactions. In light of the shape of antibodies, Hamilton and co-workers proposed the idea that calix[4]arene analogues can be modified to mimic an artificial binding surface for several proteins.[180] In the late 1990's, calixarene **17** (Figure 3.7) was shown to bind cytochrome C, a critical protein in the electron transport chain in mitochondria.[181] Later studies confirmed that calix[4]arene **16** forms a 1:1 complex with cytochrome C and blocks the heme edge.[182] Another study on calix[4]arene **16** showed that the protein mimic on the calix[4]arene also inhibits α-chymotrypsin at a nanomolar level, suggesting a multifunctional role.[183] Calix[4]arene **17** (Figure 3.7), was shown to achieve potent platelet-derived growth factor (PDGF) inhibition (IC_{50}= 250 nM, Figure 1.17).[184]
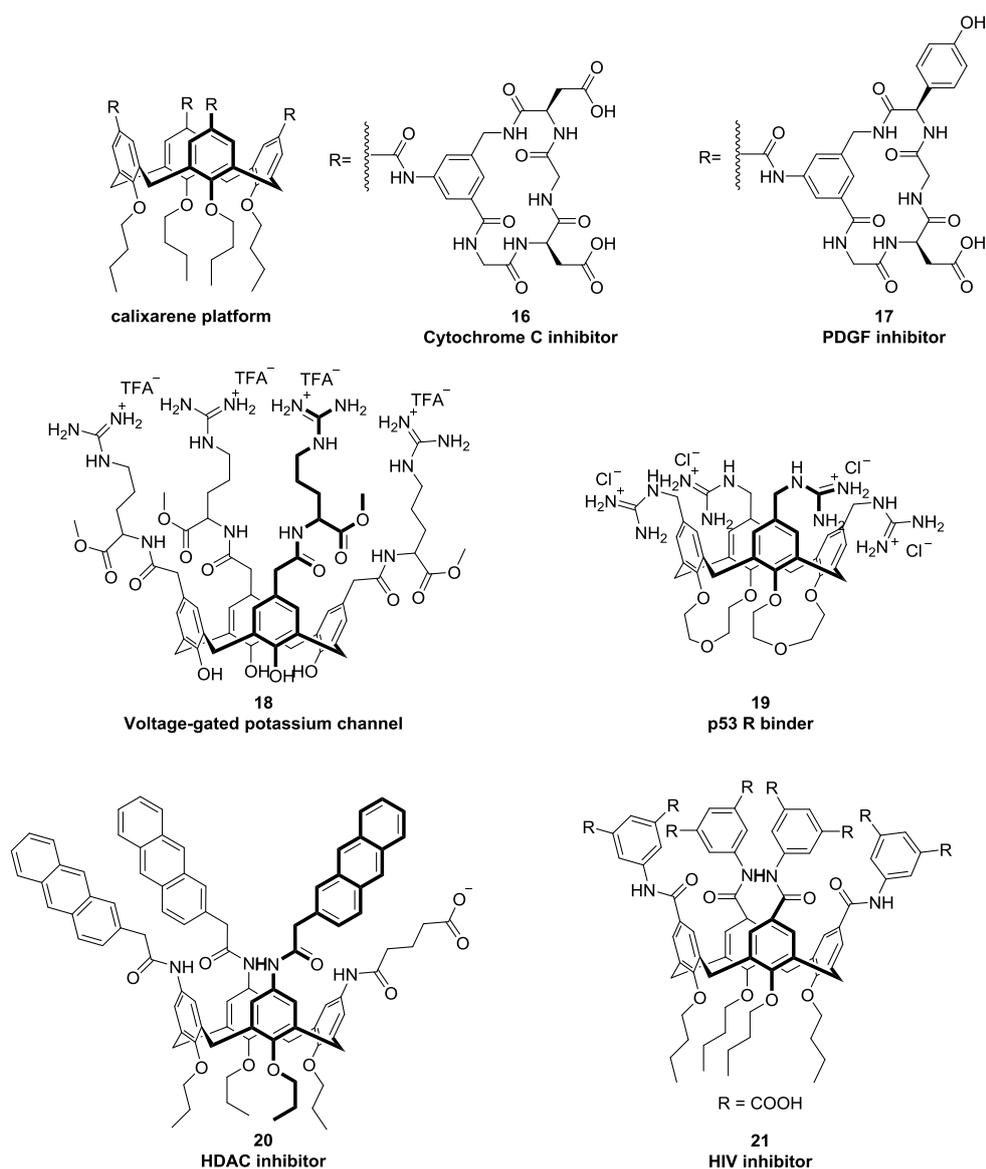
Figure 3.7   *Examples of calix[4]arenes in protein-binding studies*

Other research has focussed on attaching pseudo peptides or amino acid motifs onto calix[4]arenes to achieve biological effect on receptors and proteins. For example, calix[4]arene **18** (Figure 3.7) with an arginine methyl ester on the upper rim was shown to be a reversible inhibitor of the voltage-gated potassium channel.[185] This effect was attributed to the relatively small size of the lower rim, which can enter the channel with ease, while cationic guanidines on the upper rim interact with the Asp-379 residue in the vestibule by ionic and hydrogen bonding.[185] Another similar study used tetraguanidiniomethylcalix[4]arene **19** (Figure 3.7), which had a fixed

lower rim and was shown to bind p53-R337H, a mutant form of the tumour suppressor protein p53.[186] In the mutant p53-R337H, an important salt-bridge between Arg337 and Asp352 is lost, which destabilized the mutant compared to wild-type p53. Furthermore, calixarene **19** can strongly enhance the thermal stability of p53-R337H by wedging into two of the four monomers *via* hydrophobic interaction ion-pairing and hydrogen bonding.[186]

Calixarene derivatives have also been investigated as tools drugs in the studies of epigenetics. As an example, arylamidocalixarene **20** (Figure 3.7) was developed as a potent histone deacetylase (HDAC) inhibitor with three positions on the upper rim functionalised with anthracene while the remaining position is functionalised with an aryl carboxylate. This calixarene was found to possess an *in vitro* IC_{50} of 0.14 ± 0.02 μM against an HDAC enzyme derived from human cervical cancer cell nuclear extracts. *In silico* modeling suggested that its anthracene substituents interacts with two tyrosine residues in pocket A (Tyr 91) and pocket C (Tyr 264) on the surface of HDAC *via* π-π stacking as well as participating in a cation-**π** interaction with a lysine residue in pocket A (Lys19). Meanwhile, the carboxylate binds to the crucial Zn(II) cations in the pocket.[187]

In antiviral studies a proteomimetic calixarene **21** (Figure 3.7) was shown to block HIV-1 entry into cells *via* binding to the exterior surface of gp120. It projects an aromatic isophthalate spacer on the upper rim which is crucial for targeting gp120, possessing a submicromolar-range anti-HIV IC_{50} for 0.65 ± 0.07 μM. Additionally, **21** is active against BMS-378806 resistant HIV strains from human patients.[188,189]

Apart from modifying the calixarene into protein-binding ligands, Consoli and co-workers used a calix[4]arene functionalised with folate acid as a tumour targeting motif and nitrobenzoxadiazole (NBD) as a fluorescent cargo (**22,** Figure 3.8).[190] The folate-acid conjugate was shown to enter neoplastic cells *via* folate receptor-mediated endocytosis.[191] In cell imaging experiments, **22** showed cellular uptake in

human cervical cancer cells, mouse embryonic fibroblasts and human malignant melanoma cell lines. Among them, the cervical cancer cells had the highest level of folate receptor expression whereas the melanoma cell line had only a modest folate receptor expression level. The results showed that after being treated with 10 μM calix[4]arene **22** for 3 hours, mouse embryonic fibroblast cell line had only internalised 10% of the fluorescent ligand while human malignant melanoma cell line had internalised 37% calix[4]arene **22** compared to that internalised by the human cervical cancer cells. Internalisation of compound **22** by human cervical cancer cells was also shown to be blocked by administration of free folic acid.[190] This showed the presence of folic acids on the upper rim of the calix[4]arene was effective at achieving receptor-specific internalisation.

22
Fluorescent folate receptor binder
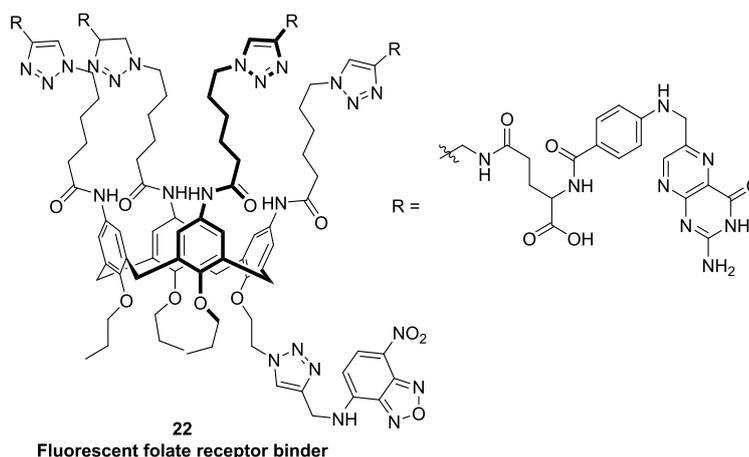
Figure 3.8   Folate-acid appended calix[4]arene-based imaging ligand: FA-C4-NBD

3.1.5   Calixarenes interacting with nucleic acid secondary structures

Calixarene conjugates have been extensively used as nucleic acid condensing reagents, as discussed in Section 1.4.2. But their sequence-specific DNA-binding properties are less well-known. By carefully tuning the functionalites on the calix[4]arene, they can be used to target nucleic acids, for example as DNA intercalators[192] and as major groove binders[126] for duplex DNA. In addition, there

are also some examples of calixarenes targeting G-quadruplexes.[192-196]

Methylazacalix[6]pyridine **23** (Figure 3.9) was the first recorded calixarene compound interacting with G-quadruplex forming sequences.[193] It has been shown *via* circular dichroism that calixarene **23** is able to induce the formation of c-Kit and c-MYC G-quadruplexes.[193] Molecular docking suggested that methylazacalix[6]pyridine **23** exists in a flexible conformation in solution. It binds to the nucleic bases on the loop region of these two G-quadruplexes,[193] but does not binds to double-helical DNA. Additionally, the binding mode towards different intermolecular G-quadruplexes was investigated.[194] It was shown that **23** can increase the thermal stability of intermolecular G-quadruplexes formed from T12 (5′-TAGGGTTAGGGT-3′, dimeric) and H12 (5′-TTAGGGTTAGGG-3′, tetrameric) by 4 and 7 °C, repectively.[194] Docking experiments demonstrated that methyl-azacalix[6]pyridine **23** possesses a twisted cone conformation when bound with one pyridine subunit intercalating into the cavity at the 3'-terminal of the intermolecular G-quadruplex. Computational docking also indicated that compound **23** should additionally be able to bind other intramolecular G-quadruplexes such as thrombin-binding DNA aptamer,[195] c-kit and Bcl-2. Methylazacalix[6]pyridine **23** also possessed the lowest free energy of binding towards intermolecular hTelo G-quadruplexes compared to other methylazacalixpyridines with different pyridine numbers (MACP-n, n= 4, 7, 8, 9).[194]

23
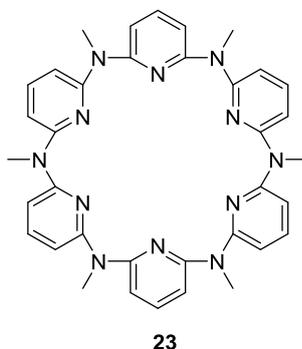
*Figure 3.9   Methylazacalix[6]pyridine **23**, the first studied calixarene-based G-quadruplex interacting compound.*

In another example, calixarenes functionalised at the lower rim with amine groups were shown to interact with the G-quadruplex forming sequences from the human telomere and promoter region of c-MYC.[196] hTeloG and c-MYC G-quadruplexes were treated with calix[4]arene **24** (Figure 3.10) and circular dichroism melting was performed to monitor the conformational changes. Calix[4]arene **24** was found to bind and stabilise both G-quadruplex types. However, the thermal stability of these G-quadruplexes were only increased by 2.7 °C.[196] Further circular dichroism experiments suggested that addition of calix[4]arene **24** altered the conformation of hTeloG from anti-parallel to parallel. In supporting cell-based studies, calix[4]arene **24** also showed significant anti-proliferative effect towards human breast cancer, androgen-sensitive prostate adenocarcinoma, primary glioblastoma and dermal fibroblast cell lines even at 1 μM concentration. However, the anti-proliferative effects of calix[4]arene **24** were not clearly linked to its G-quadruplex stablising effect.[196]

24
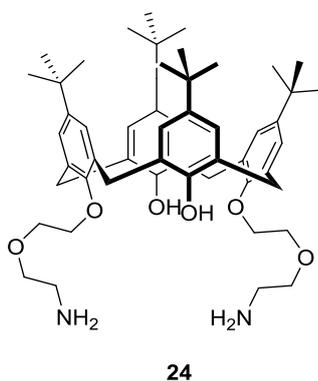
Figure 3.10 Chemical structure of calix[4]arene 24

Calixarenes functionalised with DNA-binding polycyclic-aromatics were shown to bind to duplex DNA (Figure 3.11).[192] *In silico* modeling shows that the pyrene groups on calix[4]arenes **25** and **26** are able to intercalate into DNA base pairs and the whole calix[4]arene structure can insert into the minor groove of B-DNA. Cytotoxicity experiments showed that **26** possessed high potency against the human thyroid cancer cell line, with an IC_{50} of 95 nM.[192]

Figure 3.11   Chemical structure of polycyclic aromatic hydrocarbons-

presenting calix[4]arenes 25 and 26

In Schrader's group, a flexible calixarene dimer **27** functionalised with guanidinium groups was developed (Figure 3.12).[126] **27** was shown to intercalate into the major groove of B-DNA with a 0.6:1 to 0.8:1 (dimer *vs* DNA) stoichiometry. These were determined to have a K_d between 0.2 and 1.3 μM, depending on the DNA sequence used.[126] The binding was also supported by data obtained from ethidum bromide displacement assays, UV-Vis melting and circular dichroism experiments. Further NMR experiments and computer modeling together showed that dimer **27** bridges up to nine base-pairs in B-DNA.[126] The authors indicated that further modification of **27** may lead to sequence-selective DNA recoginition.[126]

27

*Figure 3.12 Calix[4]arene dimer **27** that binds to the major groove of doube-helical DNA*

3.2   Aim

There is an opportunity for investigating the use of the calix[4]arene scaffold in binding DNA structures further. Given the beneficial properties of calix[4]arene derivatives, such as defined conformation,[140] cell-permeablity,[140,169] and potential for multi-valency,[140] a novel family of calixarene-based DNA-interacting ligands can be designed and made. In short, this proof-of-concept project can be divided into 4 small goals to achieve.

1)   To prepare a series of amino functionalised calixarenes featuring suitable substituents for conjugation at the lower rim.

2)   To prepare compementary derivatives of DNA binding motifs which can be attached to the calixarene.

3)   To prepare calix[4]arene conjugates containing DNA binding moieties.

4)   To investigate the interaction of the calix[4]arene-conjugates with DNA using FID, DNA FRET melting and circular dichroism.

3.3   Results and discussions

3.3.1   Calixarene conjugate design

In a previous study in the Matthew's group, a mono-pyrene-substituted amino-calix[4]arene (**28**, Figure 3.13)[197] was prepared and preliminary studies using FRET melting experiments indicated that it interacts with DNA. However, a detailed study had not been undertaken.

28
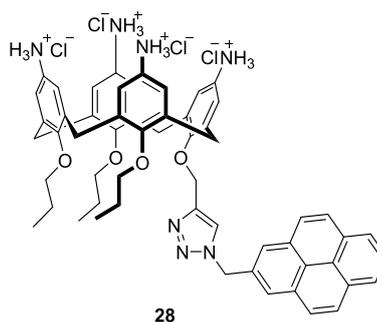
*Figure 3.13 Fluorescent amino-calix[4]arene **28***

Given the previous work in synthesising this compound, it was decided to use a similar approach initially to conjugate DNA binding moieties to a calix[4]arene scaffold. Thus the individual components would need to be compatible with the Huisgen cycloaddition to form the triazole ring.

Copper catalysed azide-alkyne cycloaddition

Copper catalysed azide-alkyne cycloaddition (CuAAC), also commonly reffered to as *the* 'click' reaction, is a widely used chemical reaction to create a covalent linkage between two molecules. The cycloaddition without any catalysts was first discovered by Rolf Huisgen[198] in the 1960s; thereafter in the late 1990s, both Morten Meldal[199] and Karl B. Sharpless[200] in separate studies showed it was possible to achieve regioselectiviy using copper (I) catalysts. This reaction has become a prominent synthetic procedure in various medicinal chemistry research

areas,[201] including small molecule ligands,[202] glycobiology,[203] bioconjugates,[204,205] nanocarriers,[206] and macrocyclics.[207] This reaction can be described as a 1,3-dipolar cycloaddition, where the chemically inert azide group and propargyl group are non-regioselectively bonded together (Scheme 3.1a). However, the presence of Cu (I) cations, the regiochemistry is restricted to only a 1, 4-substituted pattern (Scheme 3.1b),[199,200] as well increasing the rate of reaction by a factor of 10^7. Alternatively, with the help of a Ruthenium catalyst, 1,5-substituted regioselectivity is also available.[206] (Scheme 3.1c). However, to date, CuAAC has been applied far more frequently than its ruthenium complex counterpart.
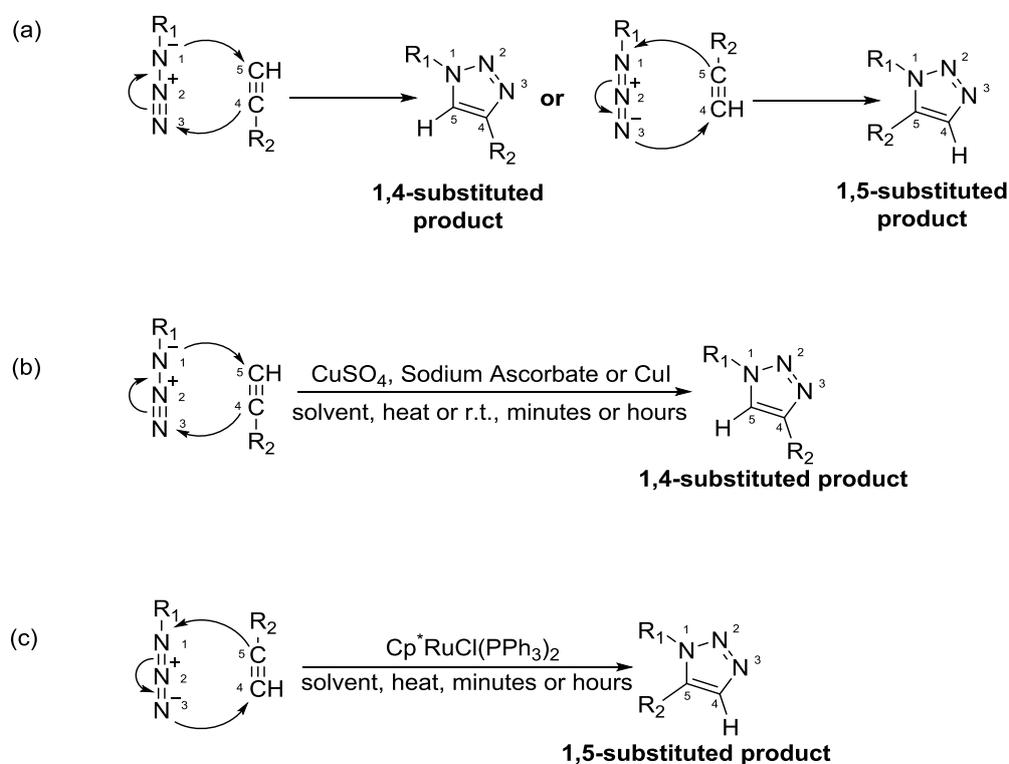
Scheme 3.1   *a) Huigsen alkyne-azide cycloaddition without catalyst, both 1,4-substituted and 1,5-substituted products are formed. b) Copper(I) cation catalysed alkyne-azide cycloaddition. c) Ruthenium complex catalysed alkyne-azide cycloaddition*

In the proposed mechanism, the reaction is initiated by coordination between an alkyne and one Cu (I) cation (Scheme 3.2). This results in a decreased pK_a for the

alkyne-terminal hydrogen. This proton is later replaced by another copper (I) cation to form Cu(I)-acetylide **I**.[208] This structure was stablised by an other copper (I) cation. The source of copper (I) cations can be from either copper (I) iodide[209] or prepared *in situ* from CuSO$_4$ reduction in the presence of sodium ascorbate.[200] Also, normally some organic bases, such as TEA, DIPEA or tris-(benzyl-triazolymethyl)amine (TBTA), are added to ligate copper (I) cations and increase their solubility in organic solvent.[207] The entry of the azide group into the catalytic cycle (Scheme 3.2) immediately gives a metallocycle **2**. The resulting complex undergoes a ring contraction to get a stable copper triazolide **3**. In the last step, the copper (I) cation on the 5′- position of 1,2,3-triazole is replaced by protons from the environment due to pK_a preference again (Scheme 3.2).[208] With the help of copper (I), CuAAC occurs about seven orders of magnitude faster than the uncatalysed cycloaddition.[208]
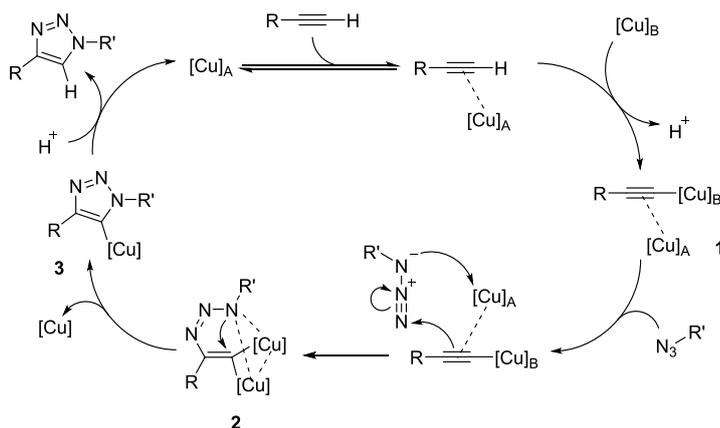
Scheme 3.2   Proposed mechanism of copper catalysed azide-alkyne cycloaddition[208]

In CuAAC, the resulting planar 1,2,3-triazole group is a chemically-inert heteroaromatic group with a proton signal at around 7 ppm on 1D ^{1}H NMR. This group can be also considered as a Z-amide bond mimic (Figure 3.14).[201]
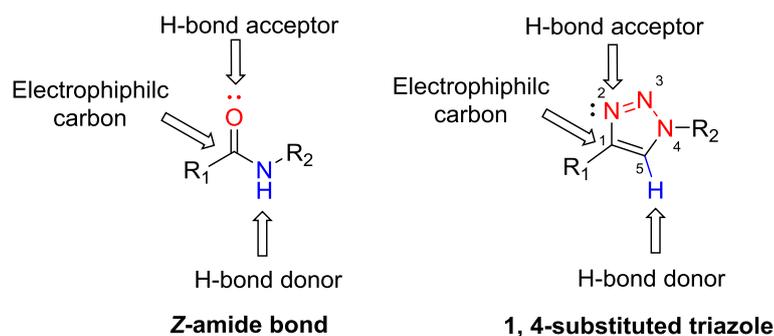
Figure 3.14   Structural comparison between the Z-amide and 1,2,3-triazole

A lone pair on the N_3 atom in the triazole resembles that on the carbonyl oxygen while the C_4 atom possess similar electrophilic and polarised properties to the carbonyl carbon. Unlike the amide bond, 1,2,3-trazole is completely resistant to hydrolysis under physiological conditions. Additionally, the two nitrogen atoms (labelled N_2 and N_3 in Figure 3.14) resemble hydrogen bond acceptors whereas the proton on the C_5 atom can act as a hydrogen bond donor. A higher dipole moment on the 1,2,3-triazole group suggested an increase in peptide bond mimicry compared to the original carbonyl amide.[201] Thus 1,2,3-triazole is widely accepted as an enhanced pharmacophore rather than a conventional linkage.[210] Another difference between Z-amide and 1,2,3-triazole is the distance between two substituents, where the distance in the triazole is about 1.1 times larger than that of Z-amide.

There are further advantages for using CuAAC to assemble molecules. Firstly, large molecules can be easily adorned with the necessary alkyne and azide functionality. Secondly, CuAAC can be achieved in a large variety of solvent systems, such as water, THF, acetone, DMSO, DMF or water/acetonitrile. Thirdly, the reaction usually can be carried out within a large range of temperatures (eg between 0°C to 160 °C) and pH values (4 to 9), indicating the robustness of the reaction.[201] In addition, most products are purified with simple work-up procedures without recrystallisation and chromatography.[201,208] However the toxicity of copper (I) in living systems was recorded as a major drawback for CuAAC, as reactive oxygen

species (ROS) can be generated by copper (I), ascorbate and atmospheric oxygen.[211] Current research has extended the idea of CuAAC towards copper-free 'bioorthogonal chemistry',[212,213] where large biomolecules can be linked with fluorescent molecules without interfering with their native biochemical processes. In bioorthogonal chemistry, the conventional aryl alkyne can be replaced with an alkyne activated by ring strain (Figure 3.15).[214] Another way to improve cell survivability is to add sacrificial reductants to counter ROS.[215-217]
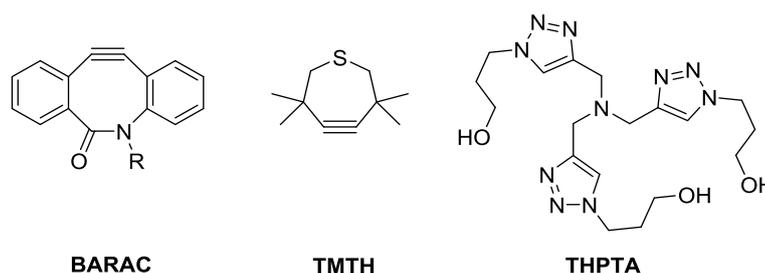
BARAC   **TMTH**   **THPTA**

Figure 3.15   *Examples of activated alkynes (BARAC & TMTH) and reductant (THPTA) in bioorthogonal chemistry*[215,218,219]

Calixarene conjugate design

To investigate the potential of lower-rim functionalised calixarenes to bind to DNA structures further, a series of first-generation ligands were designed incorprating DNA probes, such as pyrene and a derivative of the dye Nile Red (Figure 3.16a). Pyrene was selected as the major DNA interacting motif due to its known interactions with i-motifs and G-quadruplexes (Figure 3.16a).[172,173,187] NRD-2 has previously been shown to bind to c-kit G-quadruplex and down-regulated the corresponding gene expression in human gastric carcinoma cells.[220]

All compounds were fixed in the cone conformation and designed to possess hydrophilic amine groups on the upper rim (Figure 3.16b) to enhance solubility in aqueous media. Three different platforms, mono-substituted, 1,2-disubstituted and 1,3-disubstituted calixarene, were investigated to explore the structure-activity

relationships.
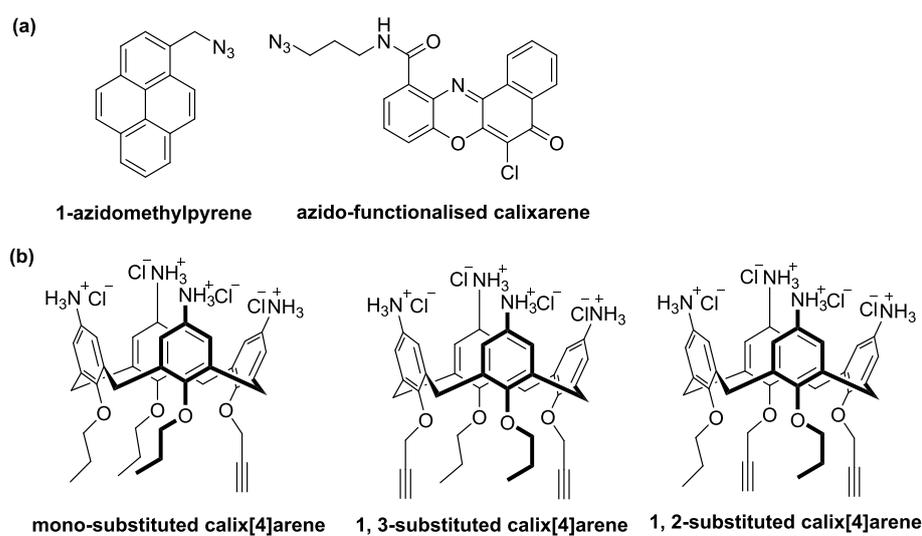
Figure 3.16   a) two DNA binding motifs selected in the proof-of-concept studies; b) three amino-functionalised calixarene platforms bearing alkyne in the proof-of-concept studies

The proposed first generation of calixarene conjugates are shown in Figure 3.17.
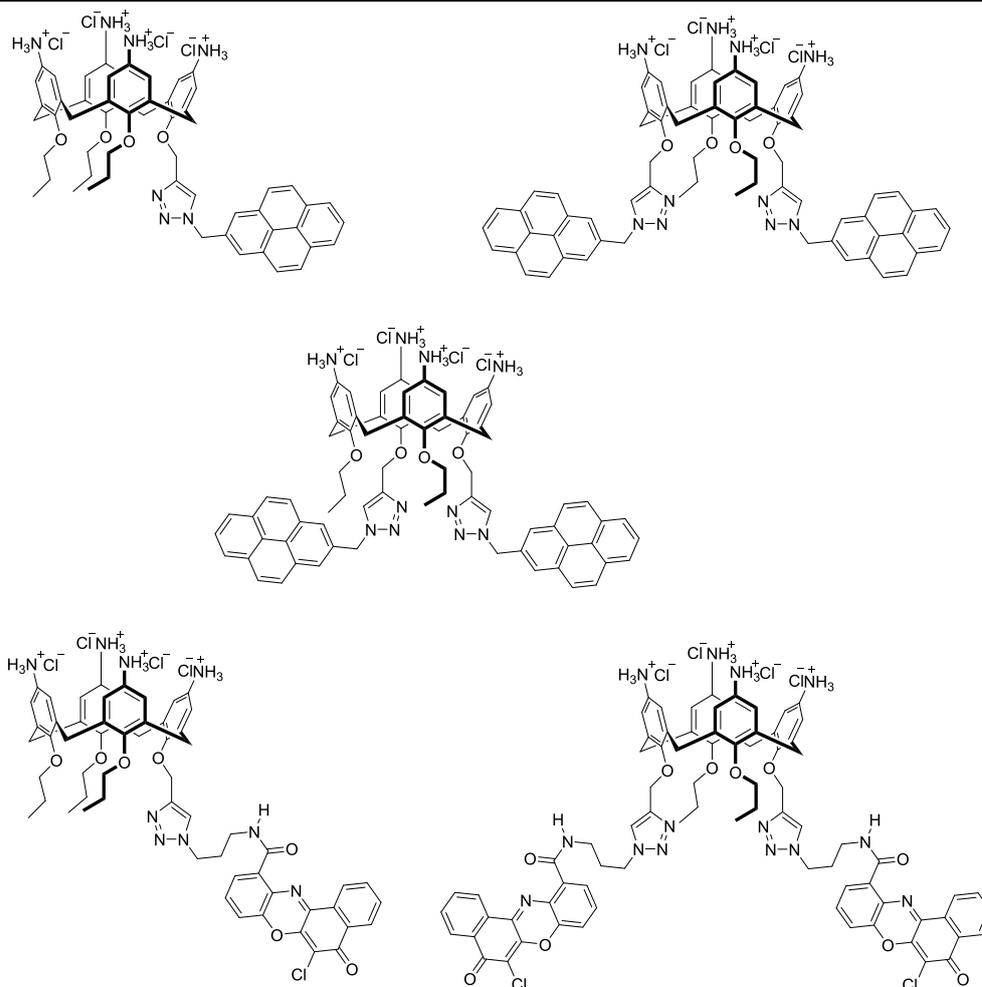
Figure 3.17   *First-generation amino-calix[4]arene conjugates with DNA-*

interacting moieties

3.3.2   Preparation of clickable calix[4]arene platforms

Preparing *p-tert-butyl*calixarene

To synthesis the clickable mono, 1,2- and 1,3-alkyne functionalized calix[4]arenes, a general procedure previously described by Karthauser was followed (Scheme 3.3).[197]
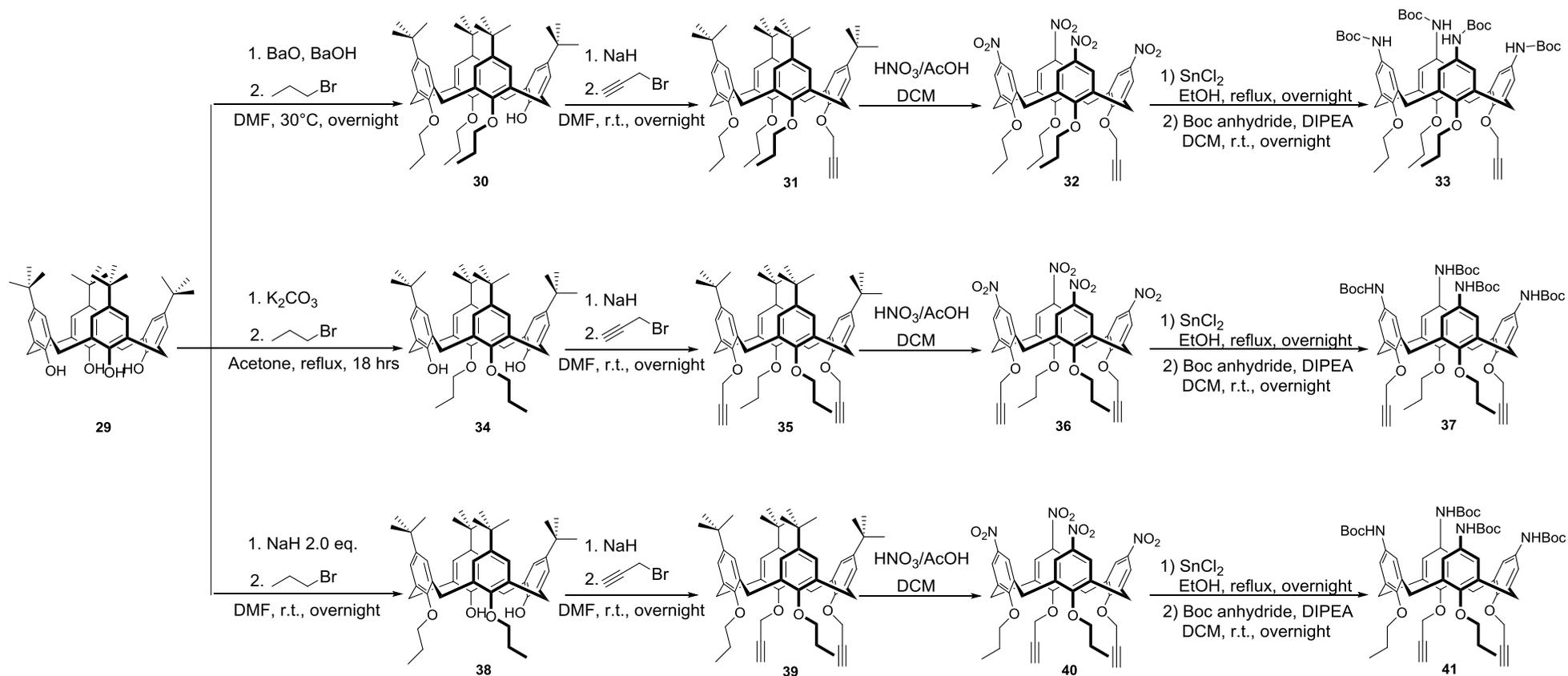
*Scheme 3.3   Synthetic route to prepare mono-propargylated calixarene **33**, 1,3-propargylated calixarene **37** and 1,2-propargylated calixarene **41***

Synthesis of the calixarene conjugates started first by preparation of *p-tert*-butylcalixarene **29**. In this reaction, the first step is base-catalysed phenol-formaldehyde polymerization at 120°C, followed by the breaking of the polymer-chain, and the reorganization of the fragments into four-membered macrocycles which was then precipitated from the mixture with ethyl acetate (Scheme 3.4).[221] The amount of base added is crucial since it determines the number of phenols in the final calixarene system. It was found that addition of 0.03-0.04 equivalents of sodium hydroxide facilitates the formation of calix[4]arene while 0.3 equivalents of sodium hydroxide facilitates the synthesis of calix[6]arene.[222,223]
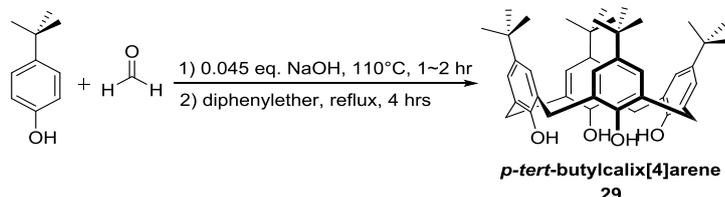
Scheme 3.4   *Synthesis of p-tert-butylcalix[4]arene 29*

The resulting compound **29** was an off-white powder isolated in 50% yield. The structure was confirmed using ^{1}H NMR spectroscopy, which was consistent with the literature (Figure 3.18).[221]
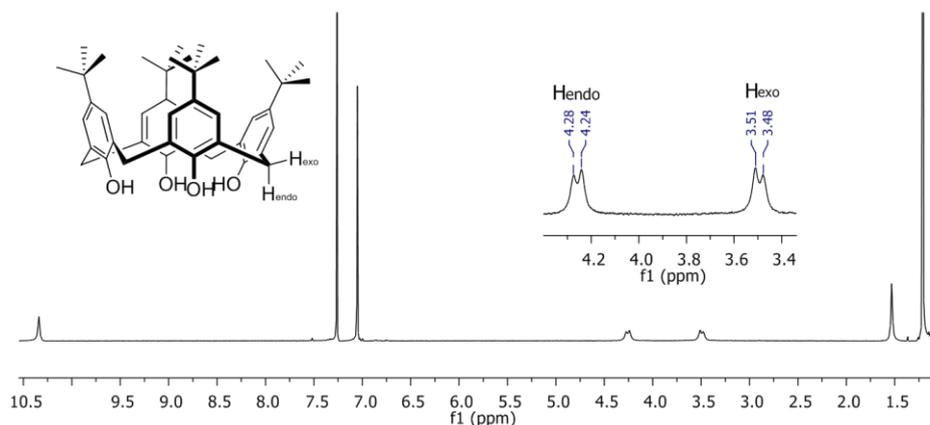
Figure 3.18   *^{1}H-NMR spectrum of Compound 29 (p-tert-butylcalix[4]arene)*

Compound **29** is a highly symmetrical molecule. However, the two protons on the methylene bridge exhibit a very interesting property in the ^{1}H NMR spectrum (Figure 3.18). These two protons experience very different chemical environments.

The proton pointing upwards (H_{exo}) has a lower chemical shift (3.5 ppm) than that of the proton pointing downwards (H_{endo}) (4.2 ppm); since the latter is closer to the O atom (Figure 3.18). Thus the diastereotopic effect results in a pair of doublets with a coupling constant of 13.5 Hz.

The cone conformation of compound **29** was retained at room temperature due to the hydrogen bonding at the lower rim. However, it changes its conformation freely at higher temperatures from cone to partial cone, 1,2-alternate or 1,3-alternate.[149] Alkylating the lower rim helps to block this transannular rotation, essentially locking its conformation[150,224] and consequentially improving the calix[4]arene's solubility in non-polar solvent.[225] In order to lock the conformation of compound **30**, any aliphatic chain longer then ethyl can be applied.[150,224]

Synthesis of mono-propargyl calix[4]arene platform

In order to prepare mono-substituted amino-functionalised calix[4]arene **33**, the synthetic route in Scheme 3.5 was used.[197]
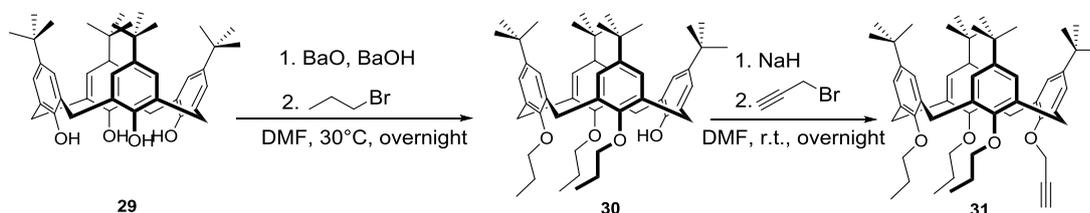
Scheme 3.5   Synthetic route towards clickable calixarene 31

In the first step, **29** was stirred for 30 minutes with a mixture of barium oxide and barium hydroxide at 30 °C prior to the addition of propyl bromide.

In this step, different bases can be used to achieve regioselective alkylation through metal templating effects in the S_N2 Williamson ether formation. Barium was chosen for tri-propylayion due to its ability to strongly co-ordinate only one phenolic residues preventing alkylation due to its size and electropositivity.[226]

After stirring the reaction mixture overnight, water was poured into the reaction mixture to precipitate the crude product. The crude product was collected and recrystallised in DCM/methanol to yield **30** as a white crystalline solid (69% yield).

Propargylation of the remaining phenol was carried out using sodium hydride as a strong base and propargyl bromide (Scheme 3.5).The small Na^+ holds the cone conformation of calixarene by coordinating all four oxygens on a single face. Tthe final phenol anion formed by deprotonation is able to react with bromide nucleophile in an S_N2 reaction mechanism.[177]

After the hydroxyl group of **30** was activated by treatment with sodium hydride for at least one hour, propargyl bromide was added into the reaction mixture and stirred overnight. The crude product was precipitated with the addition of water, then collected and recrystalised in DCM and methanol to yield mono-propargylated compound **31** in 89% yield. Preparation of **31** was confirmed using ^{1}H NMR spectroscopy (Figure 3.19).
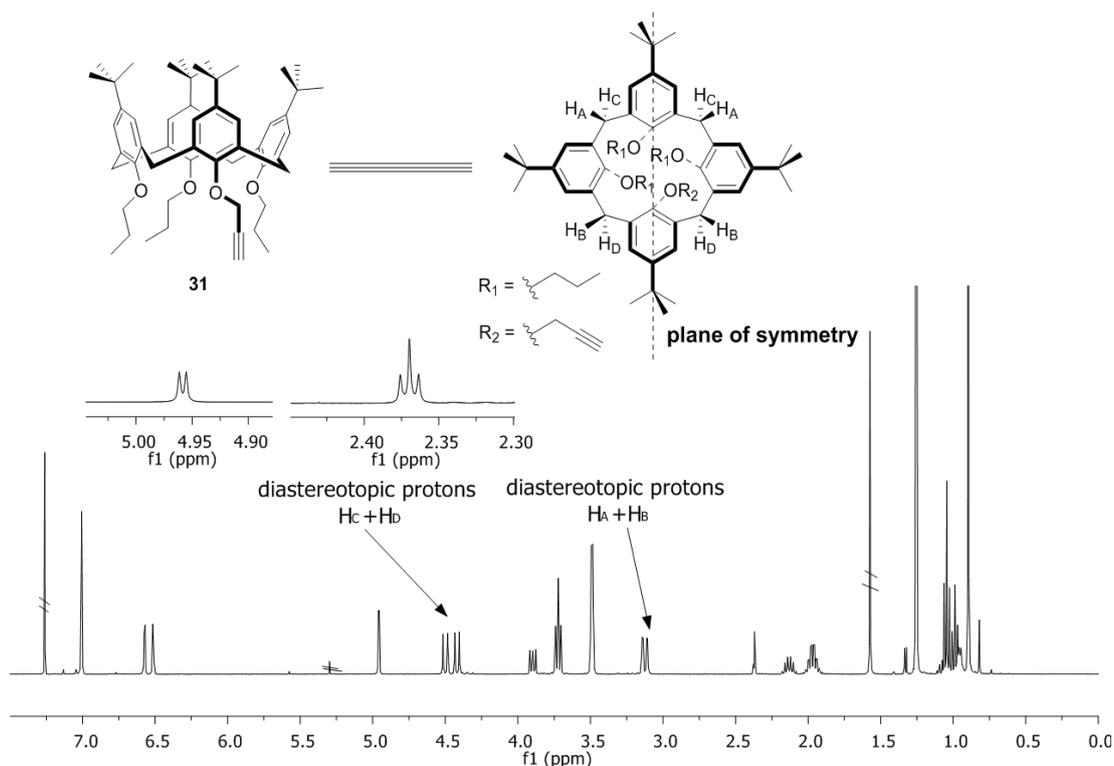
Figure 3.19   Proton splitting pattern in the ^{1}H NMR spectrum of 31

For the monopropargylated derivative **31**, there is a single plane of symmetry resulting in three separate environments within the calixarene (Figure 3.19). Thus additional inequivalence of the methylene bridge protons was observed (H_C and H_D at 4.5 ppm，H_A and H_B at 3.1 ppm). The introduction of propyl groups can be clearly seen from the two signals at 3.7 ppm and 3.9 ppm with an integration of 2 and 4 respectively representing the OC$\underline{H}_2$ group, whilst the propargyl group can be identified at 4.95 ppm and 2.35 ppm for the OC$\underline{H}_2$ and C≡C$\underline{H}$ protons which show characteristic long range coupling of 2.5 Hz.

In order to improve the water solubility of the calixarene, modifications were then applied to the upper rim to introduce an amine group. The initial step is *ipso* nitration to replace the *p-tert*-butyl group (Scheme 3.6). This step can be achieved *via* several methods, such as such as KNO$_3$/AlCl$_3$, HNO$_3$/CH$_3$COOH and HNO$_3$/Ac$_2$O. [227] Among them, the combination of HNO$_3$/CH$_3$COOH is advantagous due to its high yield and reaction kinetics.[227]
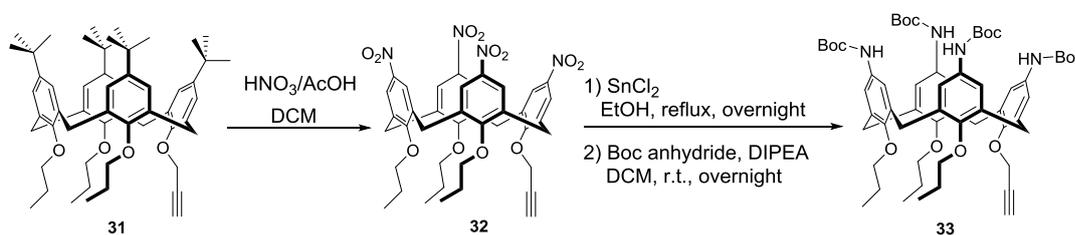
Scheme 3.6   Synthetic route from 31 to 33

Nitration was achieved by adding a mixture of 100% nitric acid and glacial acetic acid into the proparglated compound **31** dissolved in DCM (Scheme 3.6). The reaction mixture turned from colourless to deep-blue immediatedly after the addition of nitric acid, and then slowly changed to orange at which point the reaction was quenched with addition of cold water. After aqueous workup, the product **32** was recrystalised from DCM/methanol to give a off-white crystalline solid (62% yield).

The nitro groups were then reduced using Sn(II)Cl$_2$·2H$_2$O. Other reduction

methods which have been used with calixarenes include hydrazine with Raney nickel or palladium with either hydrazine or hydrogen.[228,229] The later two methods are not suitable for calixarene **31**, as they could potentially reduce the propargyl group on the lower rim as well as the desired nitro groups. Thus the first method was selected as a mild reduction procedure. The insoluble starting material **31** was initially suspended in refluxing ethanol with $Sn(II)Cl_2 \cdot 2H_2O$ overnight. When the reduction of **31** was complete, product **32** was soluble in ethanol. The crude amino-intermediate **32** was obtained as a brown-colored solid with a simple aqueous workup with DCM and used in the next step without further purification.

The amine groups in **32** were then protected as the Boc derivatives by stirring with Boc-anhydride and DIPEA for 18 hours.[230] It is to avoid gradual degraration of the unprotected amine group on **32** over time.[231] The product **33** was purified *via* column chromatography, giving a 34% yield over two steps. The preparation of calixarenes **31**, **32** and **33** were confirmed with ^{1}H NMR spectroscopy (Figure 3.20)
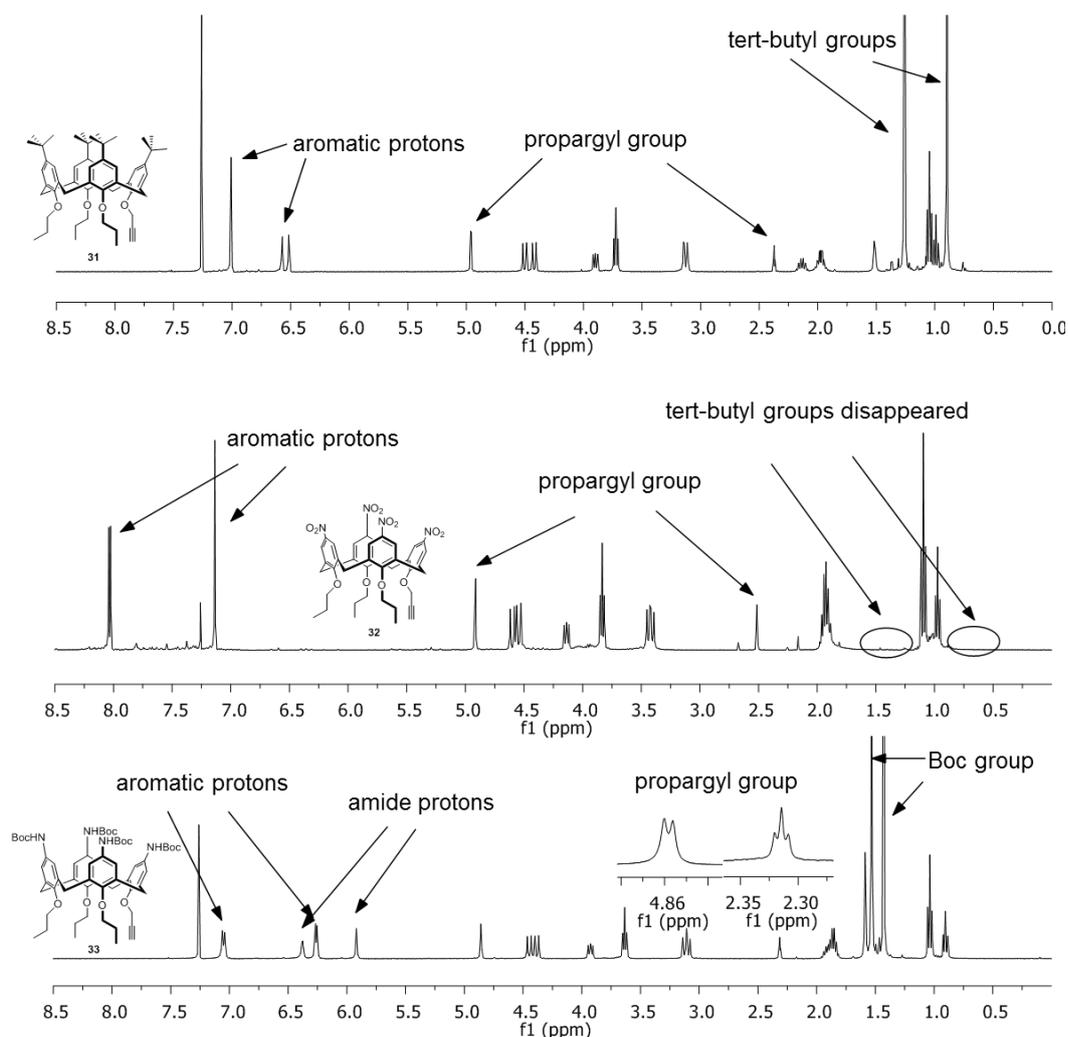
*Figure 3.20   Comparison of chemical shifts in 1H NMR spectrum between compound **31, 32** and **33***

Comparing compound **31** and **32** (Figure 3.20), the three singlets around 1.25 to 0.90 ppm in **31** for *p-tert-butyl* groups are not present in the ^{1}H NMR spectrum for compound **32**, confirming the replacement of the *p-tert*-butyl groups with nitro groups. In addition, the chemical shifts of the aromatic protons were shifted to the higher peak positions due to the electron-withdrawing effects of the nitro group on the aromatic rings.

Comparing **32** with **33** (Figure 3.20), the chemical shifts for the aromatic protons in **33** were shifted back due to the electron-donating effect of the amine. The installation of the Boc-protecting group was observed through three singlets around

1.5 ppm (*tert*-butyl groups) and three small singlets at 6.38 and 5.92 ppm (amide). The critical propargyl group was not affected by the reduction with $Sn(II)Cl_2 \cdot 2H_2O$. This can be revealed by the two proton signals at 4.86 ppm and 2.31 ppm (Figure 3.20).

Preparation 1, 3-dipropargyl calixarene platform

In order to introduce two propargyl groups on the distal phenols, a different alkylation procedure was required for the first step (Scheme 3.7). *p-tert*-Butyl calixarene **29** was heated at reflux in acetone with potassium carbonate for 30 minutes, followed by addition of propyl bromide.[232] K_2CO_3 was chosen as a mild base to deprotonate only two diagonal phenol and hold the whole calixarene structure in a cone conformation.[150,232] After stirring the reaction mixture for 18 hours, water was added to precipitate the crude product. The crude product were collected and recrystalised in DCM/methanol to yield the pure product as a white crystalline solid (46% yield).
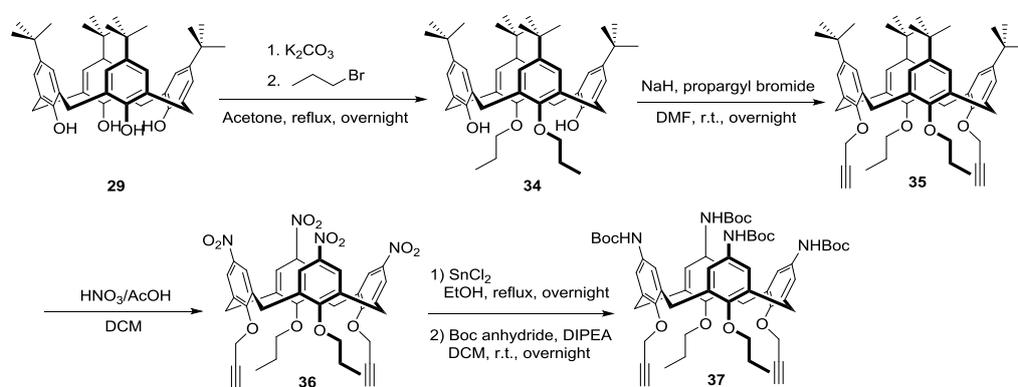
Scheme 3.7   Synthetic route to 1,3-dipropargyl calixarene platform 37

In the next step, as previously sodium hydride was used used as a strong base to deprotonate the remaining phenol groups on **34** to allow the installation of propargyl, yielding the desired product (77%). This is because sodium hydride used in the reaction was a strong base capable of deprotonate the remaining protons. Meanwhile its templating effect is weak as well.[150] Formation of **35** was confirmed with NMR spectroscopy (Figure 3.21).
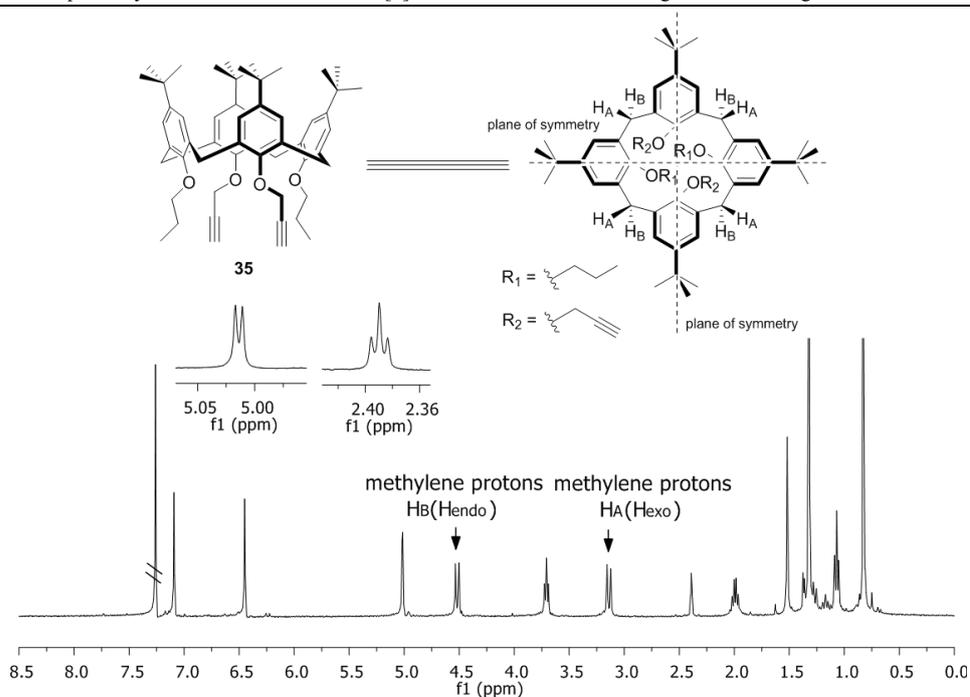
Figure 3.21   *Proton splitting pattern in the ^1H NMR spectrum of compound 35*

For 1,3-dipropargylated compound **35** (Figure 3.21), it has two planes of symmetry (dashed line in Figure 3.21) passing through two diagonally located aromatic rings. The methylene groups connecting two adjacent aromatic rings were spotted as two doublets at 4.5 ppm (H$_A$) and 3.2 ppm (H$_B$). Lower-rim propyl proton signals on compound **35** are demonstrated as three proton peaks at 3.7 ppm, 2.0 ppm and 1.1 ppm. The propargyl group was observed as two proton signals at 5.01 ppm and 2.35 ppm for the OC$\underline{H}_2$ and C≡C$\underline{H}$ protons which show characteristic long range coupling of 2.5 Hz.

Upper rim modification on **35** was achieved by following the same procedures for **33** (Scheme 3.8). The yield for **36** was 48% after recrystalisation from DCM/methanol and the yield for **37** was 25% over two steps after column chromatography.
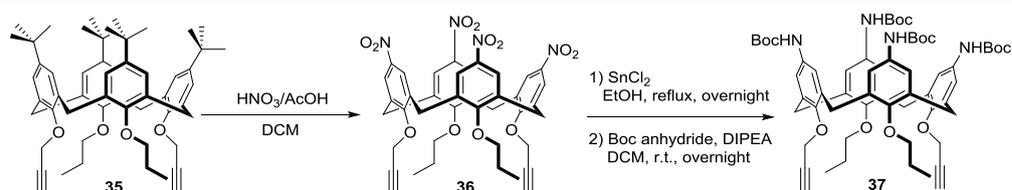
Scheme 3.8   *Synthetic route from 35 to 37*

The structures of products **36** and **37** were confirmed with ^{1}H NMR spectroscopy (Figure 3.22). Conversion of *p-tert*-butyl groups to nitro groups was observed thtough the deletion of two large singlets at 1.30 ppm and 0.8 ppm in the spectrum of product **36**. The chemical shifts of aromatic protons on the compound **36** was moved from lower field to higher field (7.1 ppm to 8.4 ppm and 6.4 ppm to 7.1). The presence of the Boc-protected amino group on the calixarene **37** was verified with the appearance of two singlets around 1.5 ppm. There were also two singlets at 6.5 ppm and 6.0 ppm in the spectrum of **37** representing the amide proton on the Boc-protected amine group. The protons on the propargyl group were not affected by the reduction as verified by the doublets at 4.92 ppm and the triplets at 2.34 ppm.
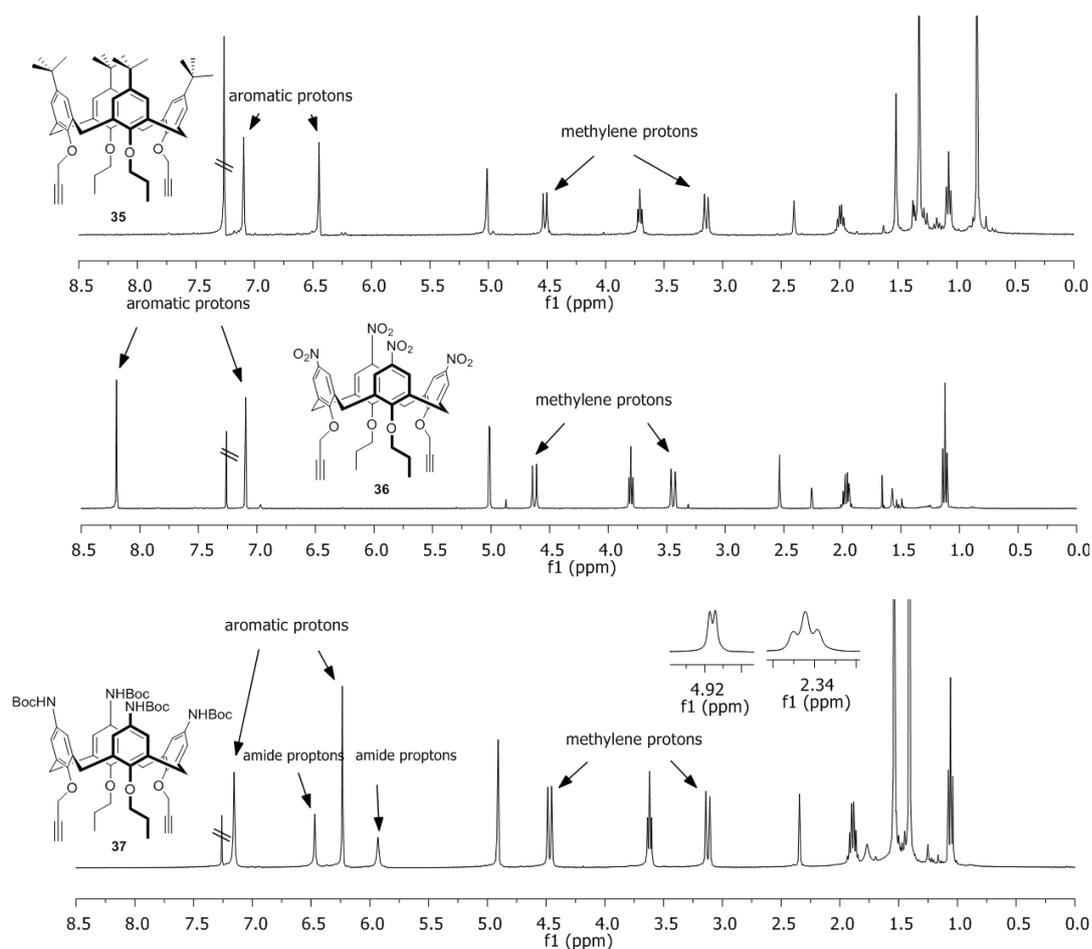
Figure 3.22   *Comparison of chemical shifts in ^{1}H NMR spectrum between compound 35, 36 and 37*

Preparation 1, 2-dipropargyl calixarene platform

As with the distally functionalised molecule **37**, the synthesis of proximally functionalised **41** requires a different first alkylation step (Scheme 3.9). Compound **29** was stirred with a large excess of sodium hydride in DMF, followed by the addition of 2.2 equivalent of propyl bromide.[233] Sodium hydride as a strong base is able to deprotonate all four phenol protons on the lower rim. 2 equivalent of sodium hydride was able to achieve the 1,2-deprotonation by statistics preference. The reaction mixture was precipitated with addition of water after stirring overnight, followed by purification *via* column chromatography to yield a white crystalline solid **38** (31% yield). The next step was to install the propargyl group to the

remaining two adjancet hydroxyl group to prepare **39** in a an analogous manner to **33** and **37** (79% yield).
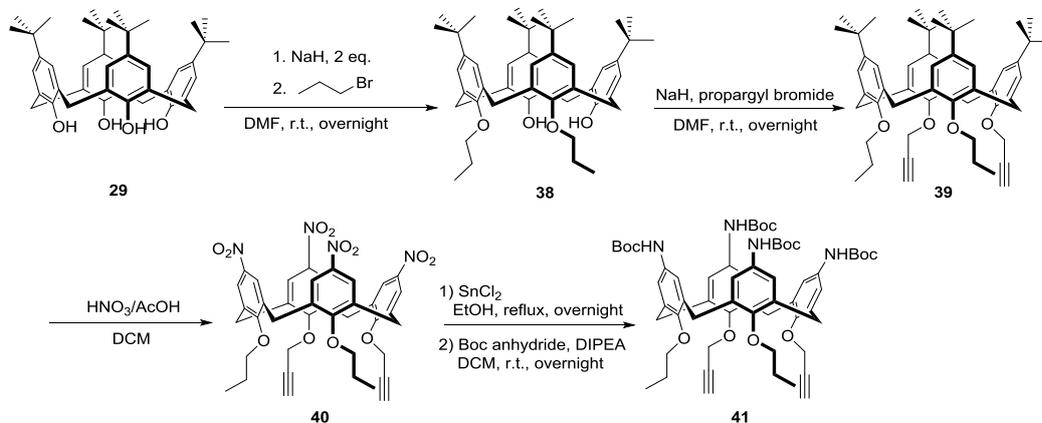
*Scheme 3.9   Synthetic route to 1,3-dipropargyl calixarene platform **41***

The structure of calixarene **39** was confirmed using ^{1}H NMR spectroscopy (Figure 3.23). Compared to **31** and **35**, calixarene **39** gives a more complex spectrum as the plane of symmetry passes through the two methylene bridges linking two propargylated or propylated aromatics (Figure 3.23). Diastereotopic proton signals for the methylene bridge show a clear (1+2+1) pattern in integration at 4.50 ppm and 3.25 ppm, which is consistent with the methylene bridges having 6 different chemical environments. The presence of the propyl groups in compound **39** were demonstrated by the three proton signals at 4.70 ppm, 2.10 ppm and 1.10 ppm, with an integration of 4, 4 and 6, respectively. The presence of propargyl group was confirmed as two proton signals at 4.78 ppm and 2.42 ppm for the OC$\underline{H}_2$ and C≡C$\underline{H}$ protons which show the characteristic long range coupling of 2.5 Hz.
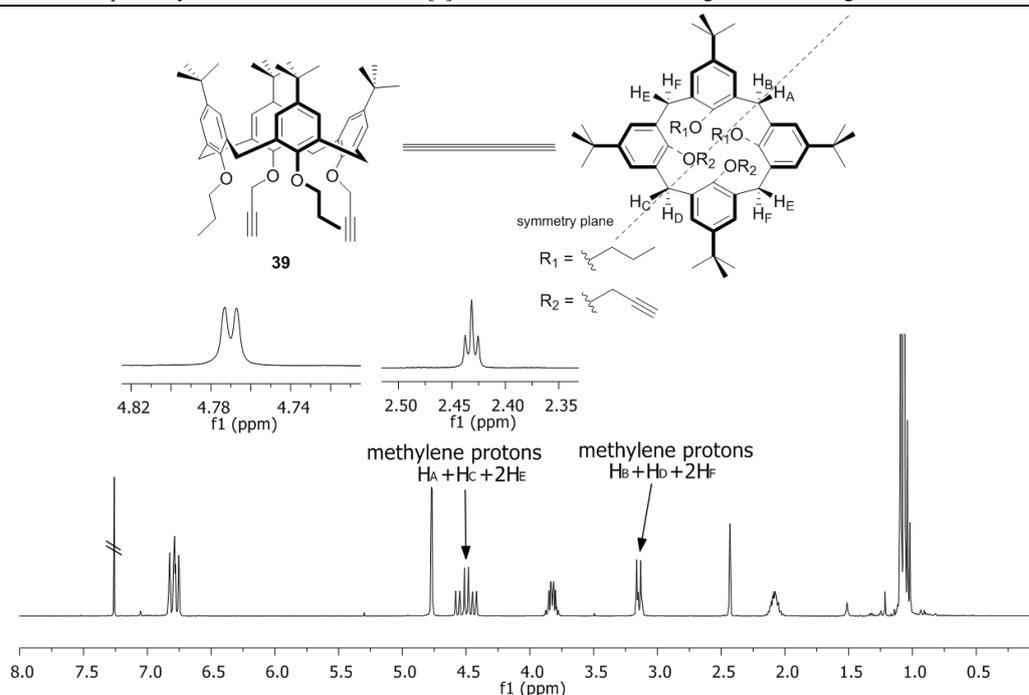
*Figure 3.23   Proton splitting pattern in the ^1H NMR spectrum of compound **39***

Later conversions from **39** to **41** (Scheme 3.10) follow the same procedures as that for Boc-amino calixarenes **33** and **37**. The yield from **39** to **40** was 39% and the yield from **40** to **41** was 45%.
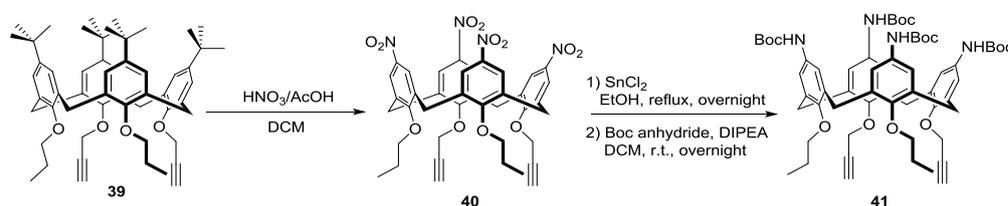
*Scheme 3.10   Synthetic route from **39** to **41***

Preparation of **40** and **41** were also confirmed using ^1H NMR spectoscopy (Figure 3.24). As described previously for **32** and **36**, removal of the *p-tert-* butyl group was shown as the missing two singlets at 1.1 ppm in compound **40** and a shift of the aromatic proton signals towards the higher field comparing to compound **39** (6.80 ppm to 7.60 ppm). The presence of the Boc-amino groups on the upper rim of the calixarene **41** was confirmed with the emergence of two singlets at 1.5 ppm, a shift of aromatic proton signals back towards the lower field (7.60 ppm to 6.60

ppm) and the presence of amide protons at 6.2 ppm. After reducing the nitro group to the amine in **40**, the propargyl group on the lower rim of **41** was not affected. This can be concluded from that the propargyl protons at 4.60 ppm and 2.32 ppm.
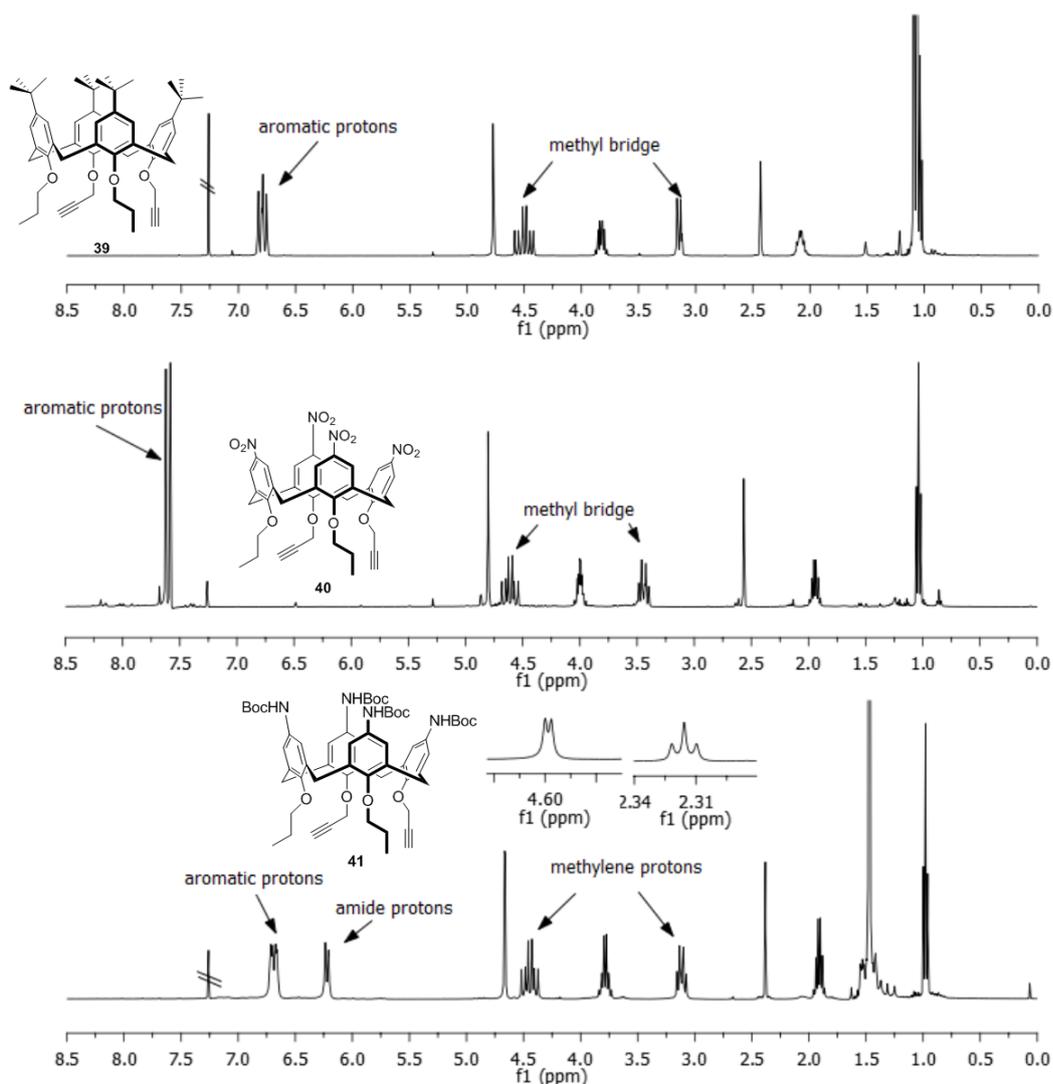
*Figure 3.24   Comparison of chemical shifts in ${}^{1}H$ NMR spectrum between compound **39**, **40** and **41***

3.3.3   Preparation of DNA-interacting moieties

After the preparation of alkyene-bearing calix[4]arene platforms, making DNA-secondary structure-targeting motifs were required. Two DNA-intercalating moieties were proposed (Figure 3.25). Pyrene, when appended to a calix[4]arene has previously been shown to intercalate into double helical DNA[192] and when

covalently attached to oligonucleotides has been demonstrated to cap i-motif structures.[124] Thus pyrene is an interesting group to be investigated further. Additionally, a Nile Red derivatives (NRD) was also chosen as it is a known small, fluroscent G-quadruplex binding compound which has been shown to decrease c-kit gene expression.[25]
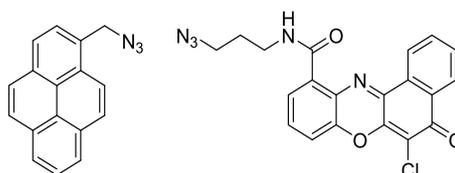
Figure 3.25   Azido-functionalisesd DNA-interacting moieties prepared to append on amino-calix[4]arene

Azido-functionalised Pyrene

In order to incorporate the dyes to the calixarene, functionalisation as the azides was required. Clickable pyrene derivatives were made according to established methods in the Matthews group.[209] 1-Pyrenemethanol **42** was brominated with PBr$_3$ to obtain compound **43**, which was then treated with sodium azide in DMF to obtain the final azido-functionalised pyrene **44** with only one carbon chain (Scheme 3.11). The preparation of **44** was confirmed using ^{1}H NMR spectroscopy. The experimental data was consistent with the existing literature.[209]
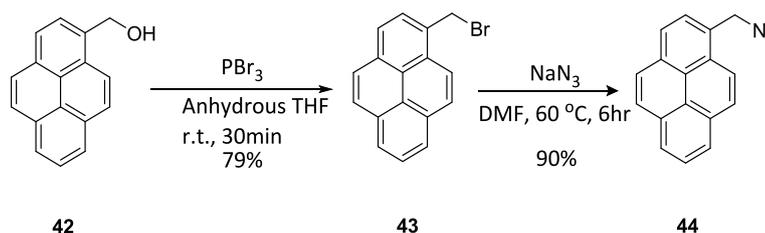
Scheme 3.11   Synthesis of 1-azidomethylpyrene 44

Azido-functionalised Nile red derivatives

The Nile red derivative was synthesized as previously described (Scheme 3.12).[220] The reaction starts with modifying the aliphatic linker **45** bearing bromine to the

azide-containing aliphatic chain **46** by treatment with sodium azide in water. Compound **47** was made from 2,3-dichloro-1,4-naphtoquinone and 3-amino-4-hydroxy anthranilic acid *via* condensation. (I performed this step with Rouven Becker, a project student worked in our lab before.) Then the side chain was introduced under anhydrous conditions in the presence of Ghosez reagent to make compound **48**, which was finally obtained *via* column chromatography to give a red-colored solid (75% yield).
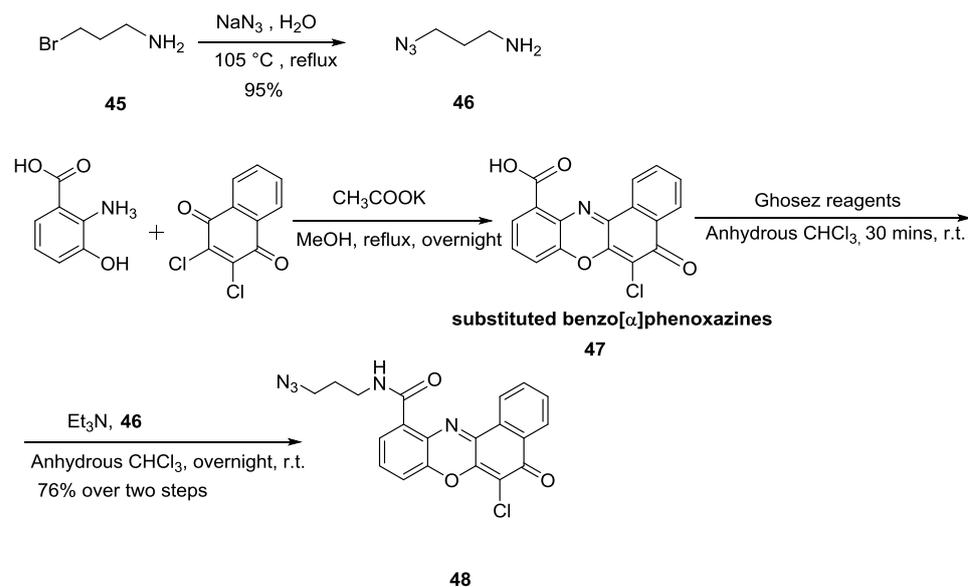
*Scheme 3.12   Synthesis of clickable Nile Red **48***

The structure of **48** was confirmed using ^{1}H NMR and IR spectroscopy (Figure 3.26). In the ^{1}H NMR spectrum of **48**, proton signals on the aliphatic chain can be found at 3.75, 3.50 and 2.05 ppm, while signals for the aromatic protons on the Nile red were observed from 8.5 to 7.5 ppm. The amide proton of **48** was assigned at 9.6 ppm. In the IR spectrum, there was an absorbance at 2107 cm^{-1}, indicating the presence of azide group.
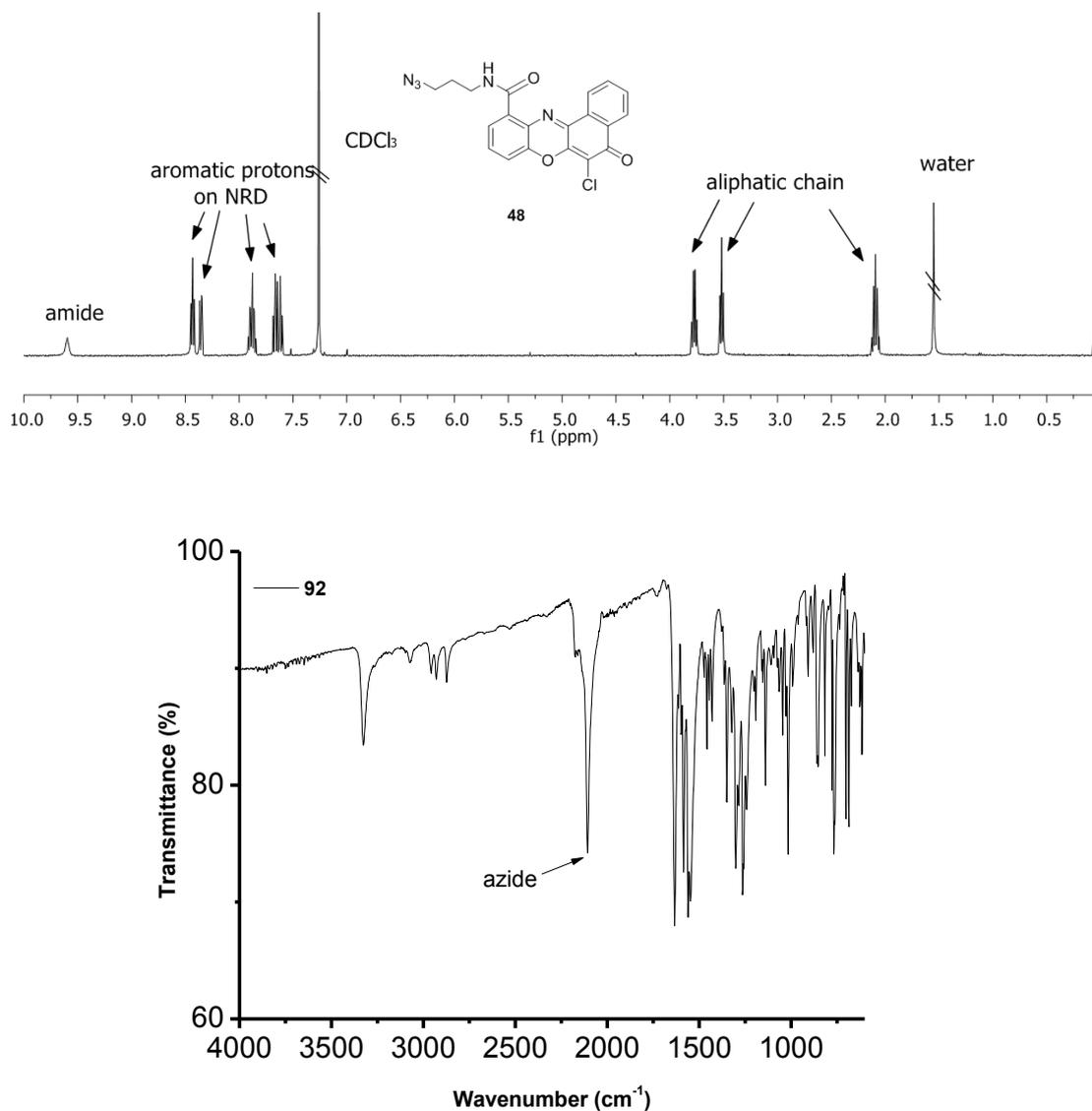
Figure 3.26 ^1H *NMR and IR spectra of compound* **50**

3.3.4   Conjugating DNA-interacting moieties with amino-functionalised calix[4]arenes *via* CuAAC

As discussed in Section 3.1.2, two main methods have been developed for the CuAAC reaction between azides and alkynes.[200,209] Both of them were investigated in a preliminary screen to determine the best conditions for the click reaction on the calixarenes (Scheme 3.13).

The solubility of amino-calix[4]arene **33** in DMF at room temperature was poor, so the reaction required heating to dissolve all the components. Thus **33** and pyrene

were dissolved in DMF at 90°C; sodium ascrobate and $CuSO_4 \cdot 5H_2O$ or CuI was added to catalyse the reaction and after 2 or 3 hours the reaction was quenched by removal from the heat and the addition of cold water. After aqueous workup, the crude brown-colored product was purified with column chromatography in DCM and acetone.[197]

When comparing the results with the catalysts, the combination of sodium ascorbate and $CuSO_5 \cdot 5H_2O$ was chosen for subsequent reactions due to its higher yield (60% *vs* 40%).
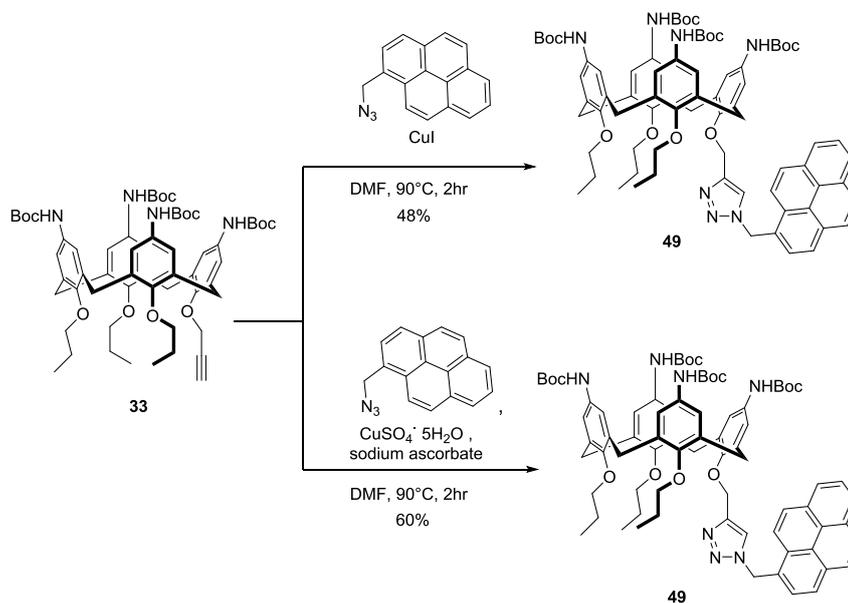
Scheme 3.13   *Two methods for copper-catalysed azide-alkyene cycloaddition (compound **49** an example)*

All other calixarene conjugates were prepared in the analogous way for **49** with $CuSO_4 \cdot 5H_2O$ and sodium ascorbate (Scheme 3.14), reaction yields for each conjugate are provided in Table 3.1.
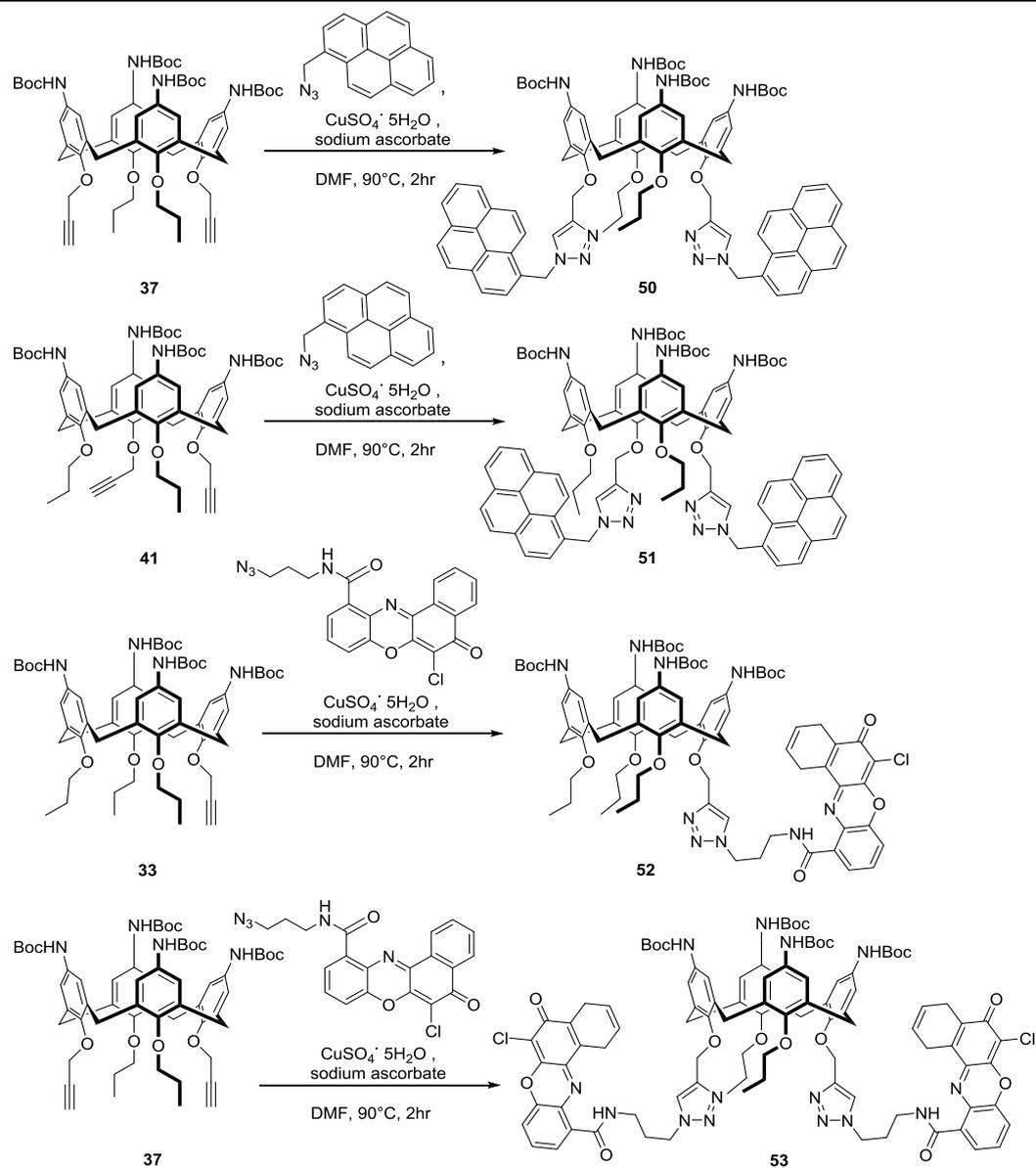
Scheme 3.14   Synthetic routes to different calixarene conjugates with CuAA

DNA binding moiety	Calixarene conjugtes, yield(%)		
	Mono-propargyl **33**	1, 3-dipropargyl **37**	1, 2-dipropargyl **41**
Pyrene **44**	**Calixarene 49**, 60%	**Calixarene 50**, 61 %	**Calixarene 51**, 52%
NRD **48**	**Calixarene 52**, 37%	**Calixarene 53**, 31%	N/A

Table 3.1   Reaction yields of CuAAC for calixarene conjugates

3.3.5   Boc-group deprotection

Mono-pyrene substituted calixarene 28

To remove the Boc protecting groups, **49** was dissolved in DCM and treated with gaseous HCl to obtain **28** as a brown colored solid (Scheme 3.15).[197,234] Other methods to remove the acid-labile Boc protecting groups include addition of TFA or HCl in dioxane.[165,190] However, gaseous HCl was chosen because it reacts with Boc group quickly and the work up involves simply removing excess DCM solvent *in vacuo*.

*Scheme 3.15   Boc group deprotection for calixarene conjugates (compound **28** as an example)*

Compound **28** was purified *via* rp-HPLC for the biophysical assays to give the final compound as a dark yellow crystalline solid (33% over two steps). The chemical structure of compound **28** after purification was confirmed using ¹H NMR (Figure 3.27), with the triazole proton (proton on the 1, 2, 3-triazole group) been seen at 7.79 ppm.

Figure 3.27   *^{1}H NMR spectrum of compound **28***

Protons on the carbon between the 1,2,3-triazole and pyrene were identified as the singlet at 6.38 ppm, while protons on the carbon between the triazole and oxygen were at 5.07 ppm, confirmed by HSQC (Figure 3.28) and HMBC (Figure 3.29). In the HSQC spectrum (Figure 3.28), both the peaks at (6.38, 52.36) and (5.07, 66.27) ppm represents two methylene groups. These are the two methylene groups invoved in the discussion below. In the HMBC spectrum (Figure 3.29), when exciting the carbon at 52.36 ppm, a long-range coupling proton signal was detected in the aromatic region at 8.08 ppm in the pyrene region, suggesting that the methylene group having the carbon signal at 51.82 ppm (H$_B$) is closer to the pyrene. When exciting the proton at 5.06 ppm in the ^{1}H NMR spectrum, another long-range coupling carbon was detected at 154.40 ppm. The carbon signals in turn have an interaction with a proton peak at 6.14, which corresponds to an aromatic proton on the calixarene. This signal at 5.06 ppm in the ^{1}H NMR spectrum was assigned to the protons on the methylene group connecting the triazole group and aromatic ring, assigned as H$_A$. The protons on the calixarene (H$_{calix}$) were also found to be interating with proton H$_A$ in the HMBC spectrum (Figure 3.29) instead of H$_B$, further confirming the idea that H$_A$ was at (5.07, 66.27) ppm and H$_B$ was at (6.38, 52.36) ppm in the HSQC spectrum (Figure 3.28).
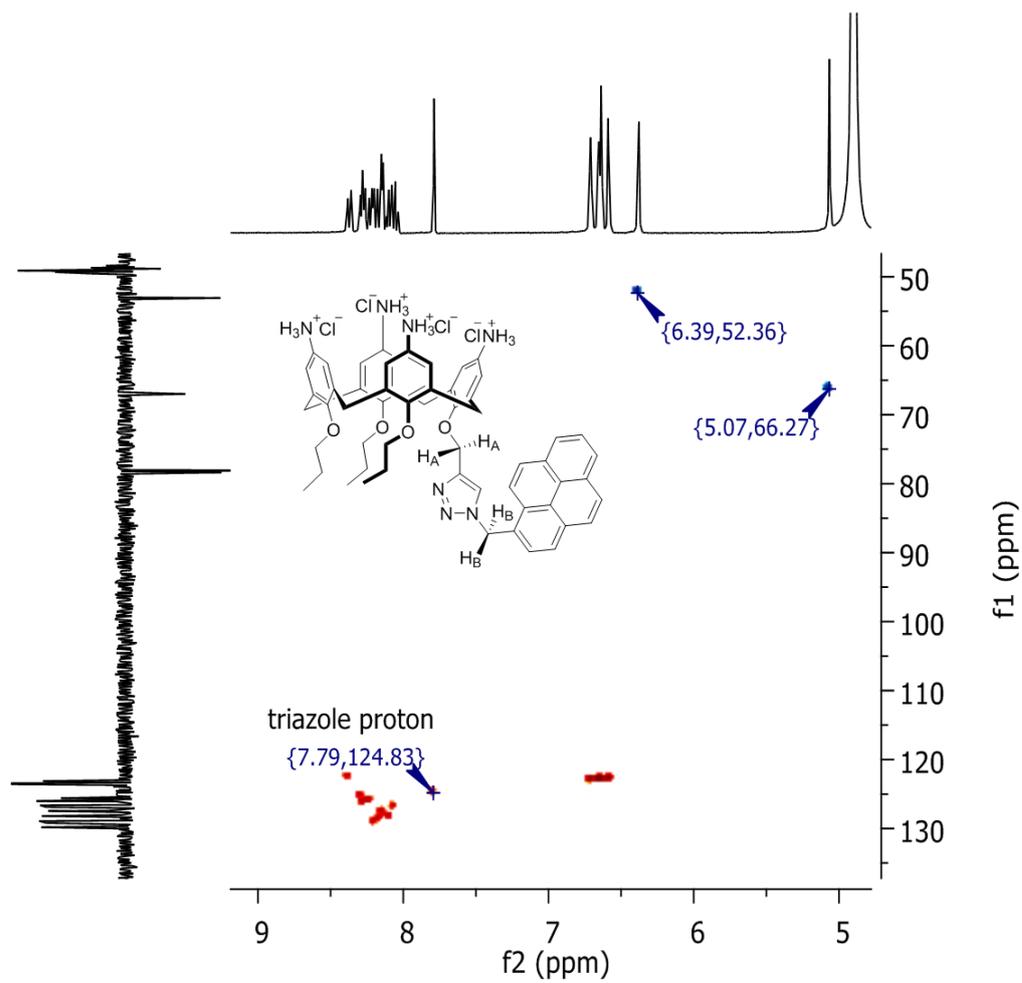
Figure 3.28   HSQC spectrum of compound 28
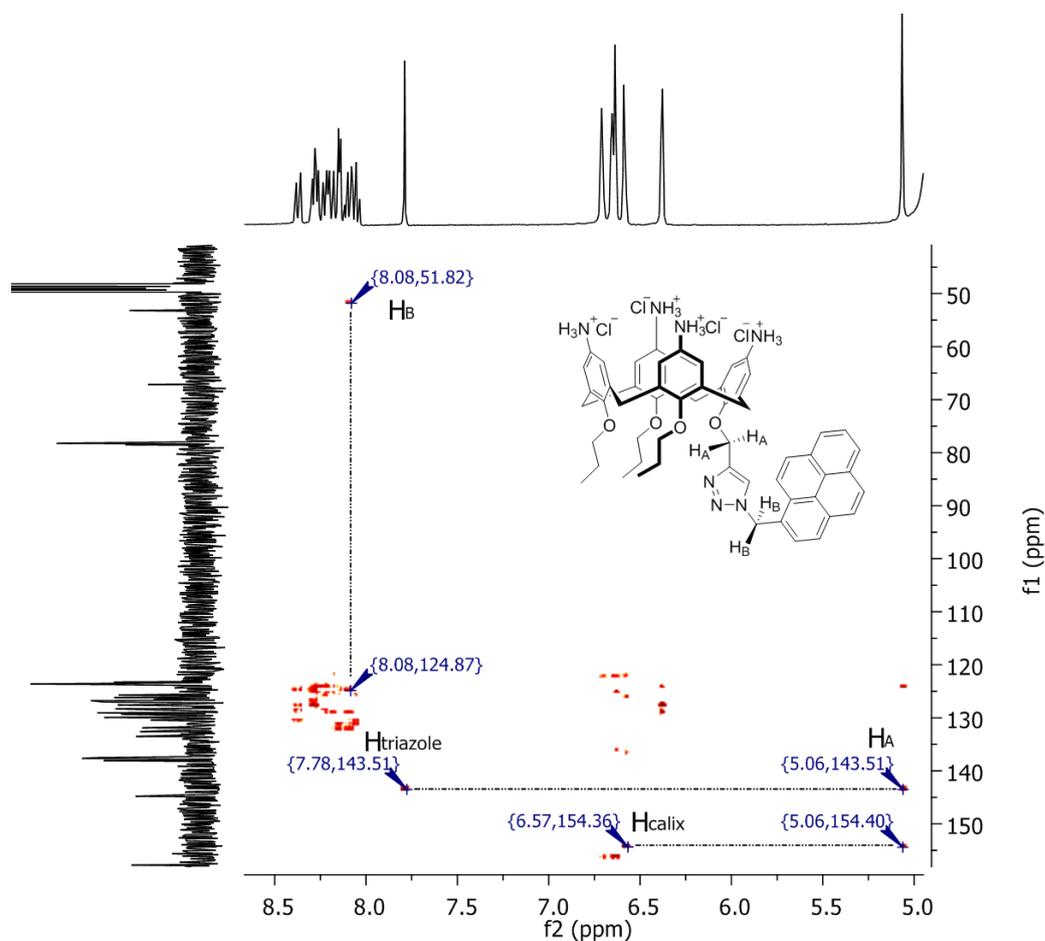
Figure 3.29   HMBC spectrum of compound 28

1,3-dipyrene substituted calixarene 54

Distally pyrene-substituted calixarene **50** prepared from the 1,3-diproparglated calixarene platform **37** and pyrene **44**, was also deprotected with gaseous HCl in DCM/methanol (Scheme 3.16), the crude product after solvent evaporation was purified with precipitating from diethyl ether at 0°C. The structure and purity of calixarene **54** was confirmed with ^{1}H NMR spectroscopy (Figure 3.30).
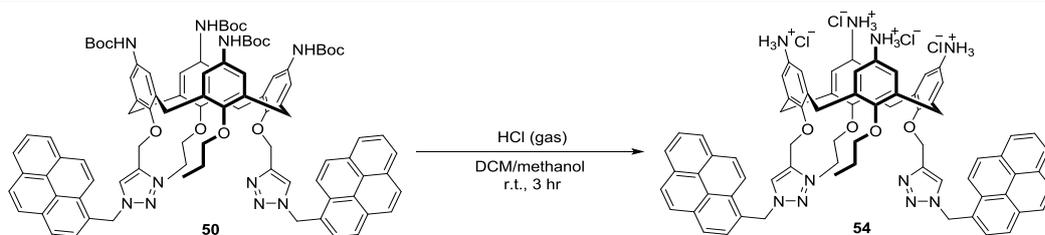
Scheme 3.16   *Synthetic route towards compound **54***
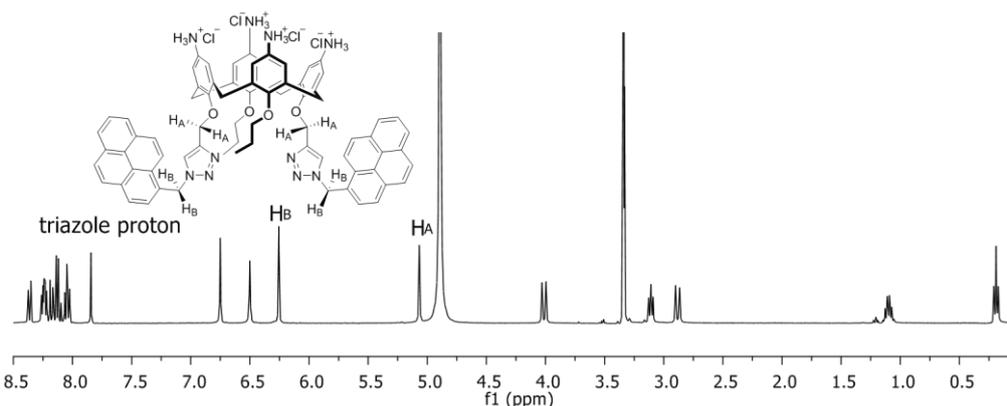
Figure 3.30   *^{1}H NMR spectrum of calixarene **54***

As described in Section 3.3.4, the singlet at 7.85 pm can be assigned as the click protons in the structure, while the singlet at 6.24 ppm belongs to 4 protons on the methylene group connecting pyrene and triazole. Due to the fact that compound **54** possesses two planes of symmetry, the proton signals are reltively simple from 4.0 ppm to 0.1 ppm. Two doublets at 4.0 and 2.8 ppm represent the protons on the methylene bridge of calixarene and the triplet at 3.2 ppm is for the protons on the propyl groups next to the oxygen (Figure 3.30). Other lowfield protons can be assigned to respectiveprotons as described in section 3.3.2 for **37**.

1,2-dipyrene substituted calixarene 55

Compound **51** after column chromatography was deprotected with gaseous HCl in DCM/methanol (Scheme 3.17). The solvent was removed *in vacuo* to obtain a brown colored crystalline solid. The crystals were then purified *via* precipitation

from diethyl ether at 0°C.
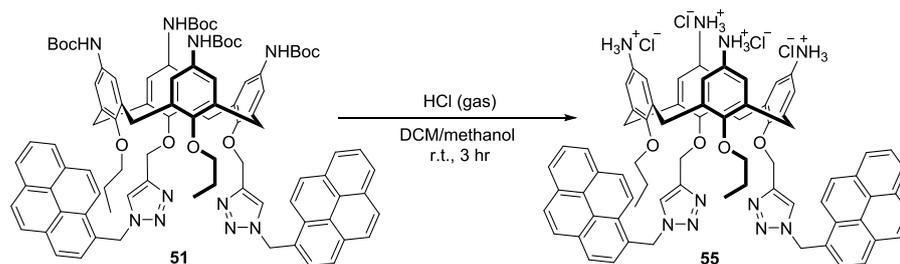
Scheme 3.17 *Synthetic route towards compound* **55**

For compound **55**, attaching the pyrene motif to two adjacent aromatic rings results in changes in the splitting patterns of the proton signals in the ^{1}H NMR spectrum (Figure 3.31). Further investigation of the proton signals from 6.25 to 2.25 ppm was achieved COSY spectscopy (Figure 3.32).
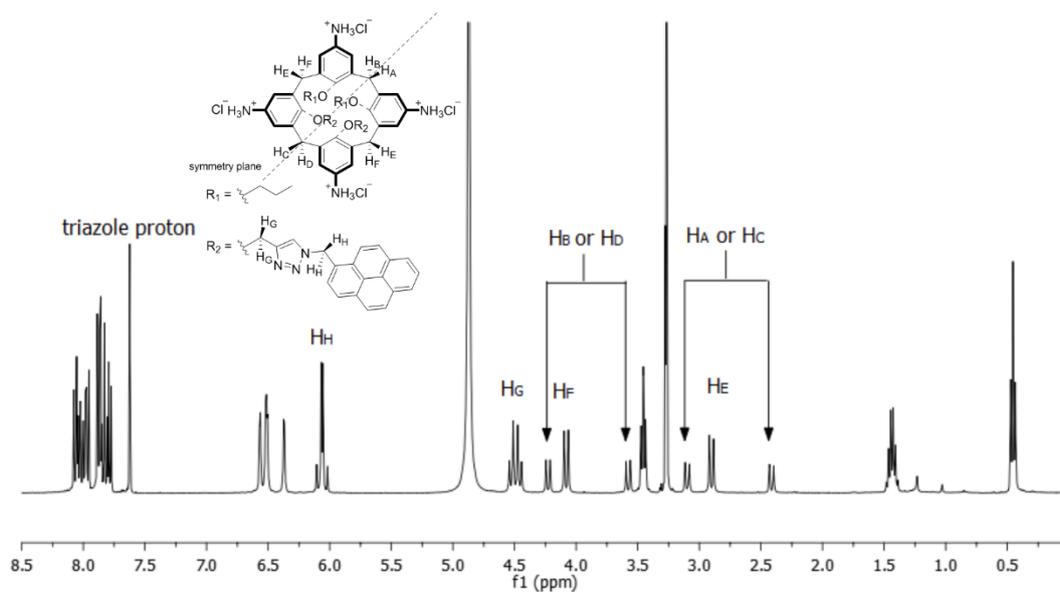
Figure 3.31 *^{1}H NMR spectrum of compound* **55**
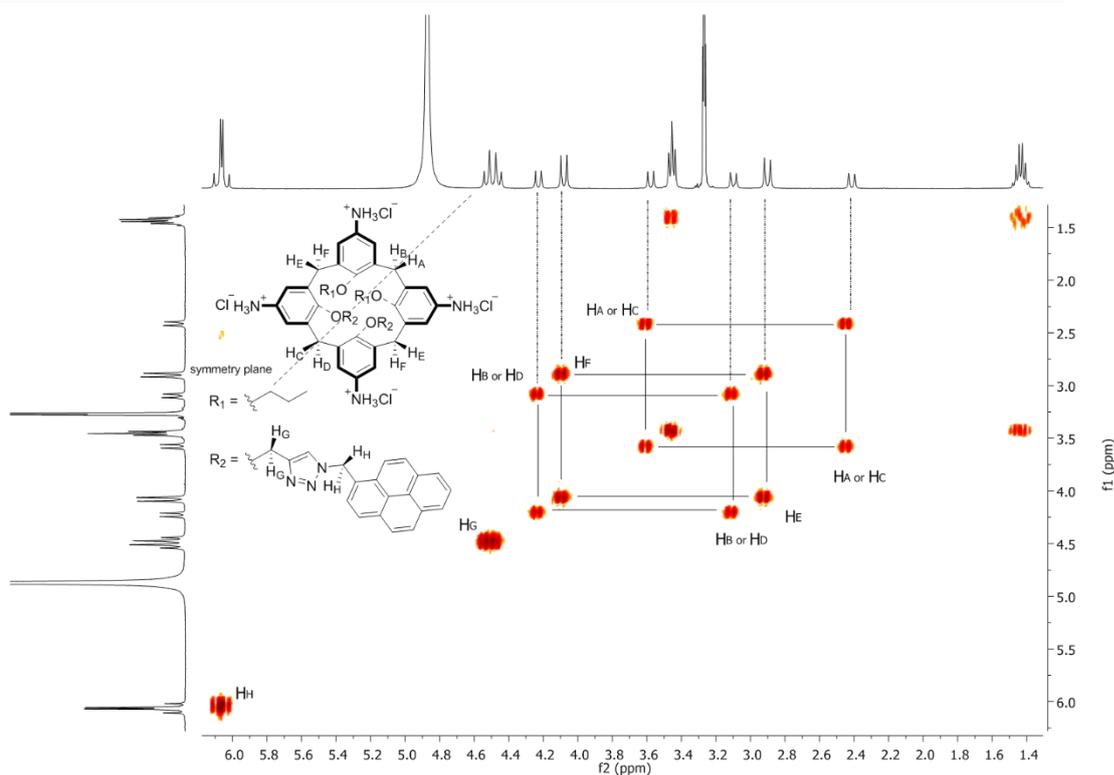
*Figure 3.32 COSY spectrum of compound **55** in deutrated methanol*

As shown in the COSY spectrum (Figure 3.32), protons giving rise to the signals at 4.25, 4.10 and 3.60 ppm are coupled with protons at 3.10, 2.80 and 2.40 ppm. These protons can be assigned as the diastereotopic protons on the methylene bridge. The proton signals at 6.10 and 4.50 ppm do not couple with any other protons signals on the spectrum, thus these protons should be close to the triazole bridge. As described in Section 3.3.4, protons on the methylene bridge connecting triazole and pyrene usually have higher chemical shifts. Thus the signal at 6.10 ppm can be assigned as the methylene between the triazole and the pyrene moiety. The signal at 4.50 ppm can be assigned to the methylene connecting triazole and oxygen atom. Compared to compound **28** and **54**, proton H_H and proton H_G (Figure 3.31 and Figure 3.32) are split.

Mono-NRD tethered calix[4]arene conjugate 56

Mono-NRD appended calixarene **52** was made in an analogous way to **29** and **54** (Scheme 3.18). The red colored crude product after deprotection with gaseous HCl

was redissolved in methanol and precipitated using diethyl ether for 5 to 6 times before structural characterization (Scheme 3.18).
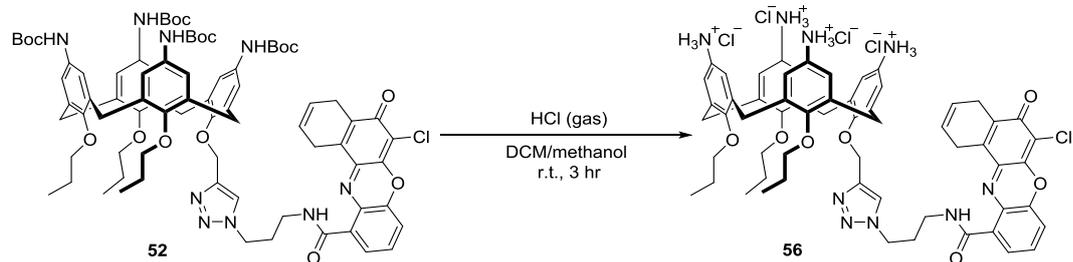
*Scheme 3.18   Preparation of mono-NRD substituted calix[4]arene **56** from **52***

Product **56** was confirmed with ^{1}H NMR spectroscopy with the click proton found at around 8.26 ppm to 8.5 ppm. (Figure 3.33).
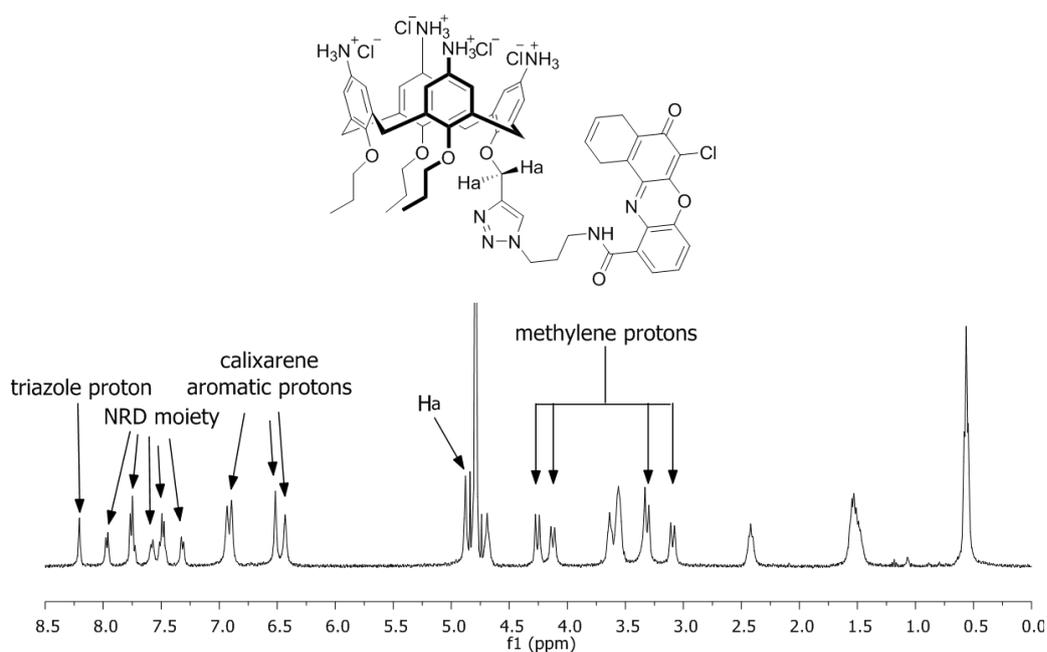
*Figure 3.33   ^{1}H NMR spectrum of calixarene **56***

As shown in Figure 3.33, the NRD was identified as several multiplets from 8.0 ppm to 7.3 ppm, which were accompanied by two triplets (4.6 ppm and 3.7 ppm) and one multiplet at 2.4 ppm representing the aliphatic linker on the NRD binding moiety. The methylene bridges on the mono-substituted calixarene **56** were identified as 4 doublets from 4.25 ppm to 4.0 ppm and 3.4 ppm to 3.0 ppm. Protons

on the carbon atom (Ha) connecting the oxygen on the lower rim and triazole group were found at 4.9 ppm.

1, 3-diNRD substituted calix[4]arene conjugate 57

Distally NRD-substituted calixarene **53** was separated from the reaction mixture *via* column chromatography, which was further deprotected and purified in an anlogues way to prepare calixarene **57** (Scheme 3.19).
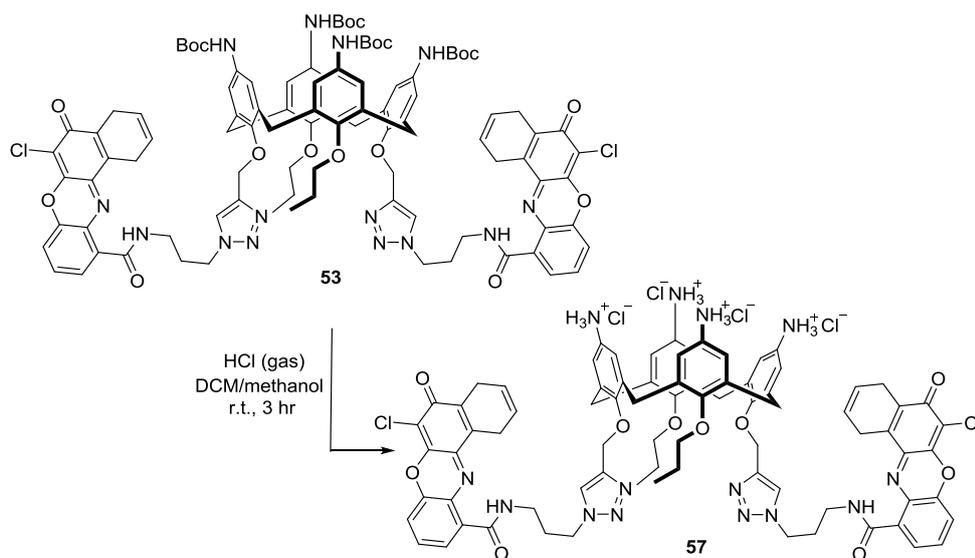
Scheme 3.19   1, 3-diNRD substituted calix[4]arene 57 from 53

The structure of calixarene **57** as its hydrogen chloride salt was confirmed with ^{1}H NMR spectroscopy (Figure 3.34).
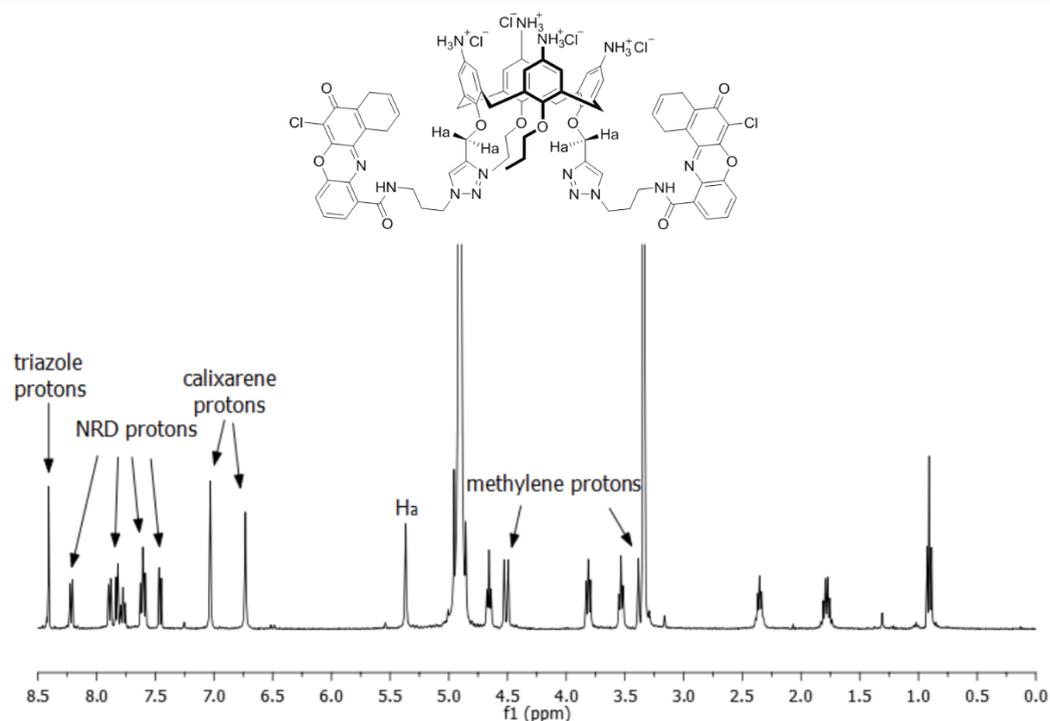
Figure 3.34 ^{1}*H NMR spectrum of calixarene* **57**

As described for **56** (Figure 3.34) the proton on the triazole group was found at 8.40 ppm and the NRD moiety can be recognised as several multiples from 8.2 to 7.4 ppm. Protons on the carbon atom connecting the oxygen on the lower rim and triazole group were found at 5.4 ppm. Compared to **56**, protons on the methylene bridges split into two doulets at 4.5 ppm and 3.3 ppm (overlapped with residual methanol signal).

3.3.6   Preparation for biophysical experiments.

Before any biophysical assays, the prepared ligands (pyrene-appended and NRD-appended calix[4]arene) were tested for their solubility in water and buffer. Solubility results for these five ligands are provided in Table 3.2. As it can be seen on the table, though all ligands prepared are soluble in pure water, their solubility in buffer is generally reduced. With increasing complexity in the structure and sodium concentration in buffer, ligand solubility in buffers containing sodium chloride decreases. Thus compound **28** and compound **54** were the two main

ligands used in the biophysical experiments. Since compound **54** is still not soluble in high NaCl buffer, it was tested in buffers containing 5 mM NaCl instead of 100 mM NaCl. For compound **28**, both 100 mM NaCl and 5mM NaCl buffers were used in the biophysical experiments. In this case, the use of 5mM NaCl buffer for compound **28** was mainly to act as a comparison towards compound **54**.

Compound	MilliQ Water	5mM NaCl, 10 mM NaCaco, pH 5.5	100 mM NaCl, 10mM NaCaco, pH 5.5	5 mM NaCl, 10 mM NaCaco, pH 7.4	100 mM NaCl, 10 mM NaCaco, pH 7.4
28	✓	✓	✓	✓	✓
54	✓	✓	✗	✓	✗
55	✓	✗	✗	✗	✗
56	✓	✗	✗	✗	✗
57	✓	✗	✗	✗	✗

Table 3.2   Solubility table for ligands (ligand concentration: 200 µM, ✓ as soluble and ✗ as insoluble)

3.3.7   FRET melting assay

Calixarene conjugates made were screened *via* FRET based DNA melting analysis to investigate whether they interact with DNA as described in Section 2.2.3.

At the beginning, 100 µM **28** was tested with hTelo G-quadruplex and duplex DNA at pH 7.4 in 100 mM sodium chloride buffer (10 mM sodium cacodylate). This is close to the physiologically relevant ionic and pH conditions in the cell. The result (Figure 3.35 and Figure 3.36) showed that compound **28** does not have any effect on duplex DNA (DS), but has a minimal thermal stabilising effect on human telomeric G-quadruplexes ($\Delta T_m \approx 5°C$). This value is higher than that for the previously disclosed compound **25** ($\Delta T_m \approx 2.7 °C$)**,** using similar concentrations and buffer conditions (K^+ instead of Na^+) measured using CD melting.[196]
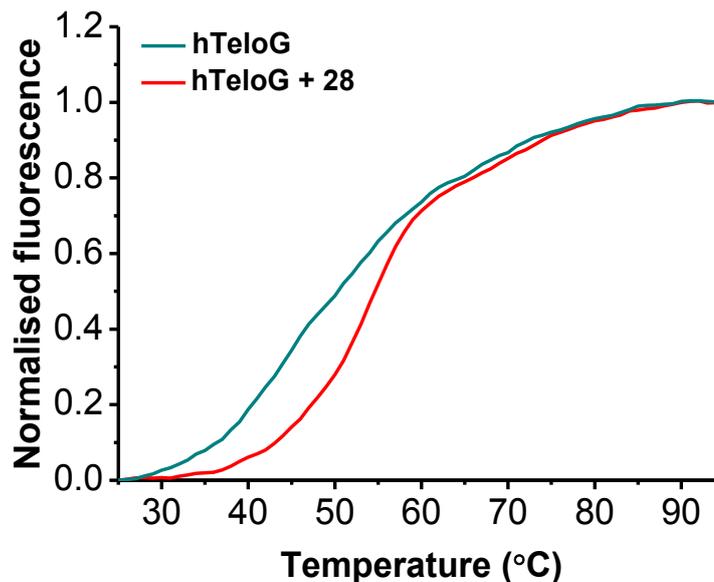
Figure 3.35 *Normalised FRET DNA melting curves for human telomeric G-quadruplexe compound **28**, [$_{FRET}$hTeloG] = 0.2 μM, [Compound **28**] = 100μ M, [sodium cacodylate] = 10 mM, [NaCl] = 100 mM, pH = 7.4, Fluorescent emission was measured at 533 nm.*
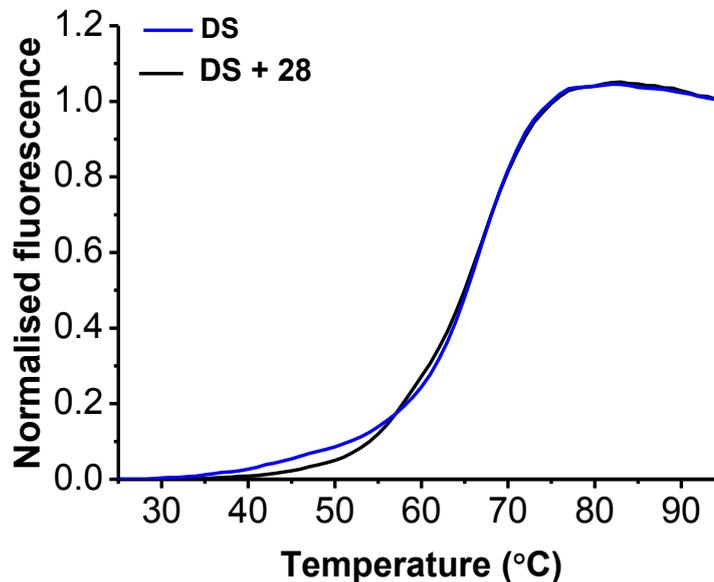
Figure 3.36 *Normalised FRET DNA melting curves for duplex DNA with compound **28**, [$_{FRET}$DS] = 0.2 μM, [Compound **28**] = 100 μM, [sodium cacodylate] = 10 mM, [NaCl] = 100 mM, pH = 7.4, Fluorescent emission was measured at 533 nm.*

Further experiments focused on investigating whether compound **28** has any binding affinity towards i-motif structures. As in Figure 3.37, after addition of **28** to the FRET-labelled DNA, the fluorescence signals for DNA remain quenched regardless of the increase of temperature. Given the previous experiments with the FRET labelled DNA, this indicates that there is not a non-specfic interaction with the fluorophores. Thus, this phenomenon indicates two plausible situations: strong DNA stabilisation at tested concentration or oligonucleotide condensation; both indicate interaction with DNA. Similar phenomena were observed for compound **54** (Figure 3.38). Since compound **54** was not soluble in high salt buffer (100 mM NaCl), it was tested at low salt buffer (5 mM NaCl) instead; compound **28** was tested at the same conditions for comparison. Regardless, in both high and low salt conditions, both **28** and **54** show quenched fluorescence up to 95°C.
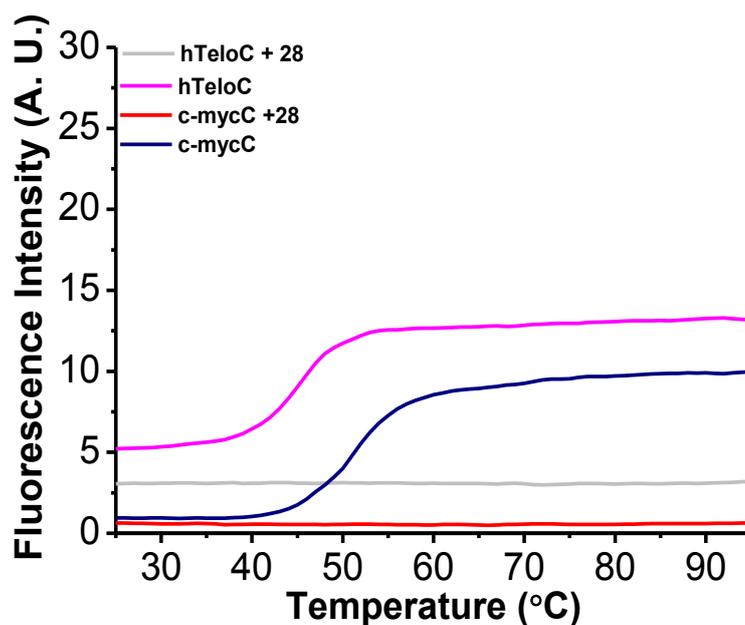
Figure 3.37   *Original FRET DNA melting curves for hTelo and c-myc i-motif with compound **28**, [$_{FRET}$DNA] = 0.2 μM, [Compound **28**] = 100 μM, [sodium cacodylate] = 10 mM, [NaCl]$_{high salt}$ = 100 mM, [NaCl]$_{low salt}$ = 5 mM, pH = 5.5. Fluorescent emission was measured at 533 nm.*
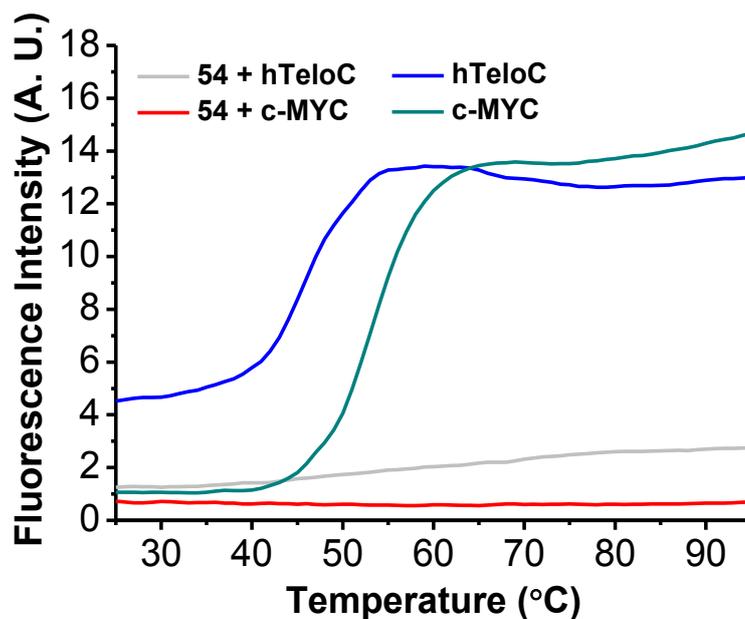
Figure 3.38   *Original FRET DNA melting curves for hTelo and c-myc i-motif with compound **54**, [FRETDNA] = 0.2 μM, [compound **54**] = 100 μM, [sodium cacodylate] = 10 mM, [NaCl] = 5 mM, pH 5.5. Fluorescent emission was measured at 533 nm.*

In order to explore this phenomenon, FRET-melting experiments were implemented with a gradient of ligand concentration for compound **28** (Figure 3.39). The first derivative charts provides the information about rate of change in fluorescence at each temperature on a original FRET-melting curves. The minimum value of the 1st derivative of the melting curves reveals the melting transition. If a ligand stabilises DNA secondary structure, then increasing concentrations of ligand should give a shift of sigmoidal curves in the original FRET-melting data and a corresponding shifts in minimum temperature in the 1st derivatives of the curves. Using this approach, the 1st derivative chart in Figure 3.39 indicates that melting temperature always remains at 45 °C, despite the changes in ligand concentration. This suggests addition of ligand does not stabilize the structure. This phenomenon was consistent with sequences from hTeloC and c-MYCC under both conditions. This indicates that DNA-ligand interactions could result in condensation or precipitation of the DNA.
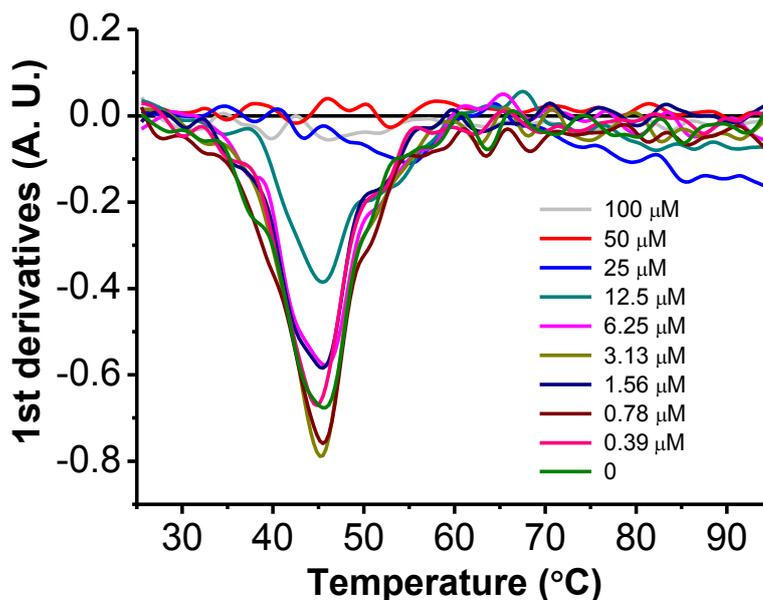
Figure 3.39   *FRET DNA melting experiments on Compound **28** with hTelo i-motif at pH 5.5 (1st derivatives in fluorescence) [FREThTeloC] = 0.2 μM, [sodium cacodylate] = 10 mM, [NaCl] = 100 mM, pH 5.5. Fluorescent emission was measured at 533 nm.*

As the experiments with G-quadruplex and double helical DNA showed very little indication of precipitation, and these were performed at pH 7.4, as opposed to the slightly acidic conditions for the i-motif forming sequences, there are a few variables which could account for the differences observed in the properties on ligand binding. In order to investigate whether calix[4]arene-induced DNA condensation is pH-dependent, the G-quadruplex and duplex DNA sequences were also tested under analogous conditions but instead at pH 5.5, to mimic the experiments performed with i-motif DNA. The data is shown in Figures 3.40 & 3.41. It is clear that there is also a quenching of the fluorescence until the end of the experiment. This may suggest that compound **28** and **54** are able to condense G-quadruplexes and duplex DNA at pH 5.5. This effect was observed in both 100 mM NaCl and 5 mM NaCl buffer conditions. This suggests that the acidic buffer conditions amplify the condensation effect and is not an effect which is necessarily dependent on sequence or structure.
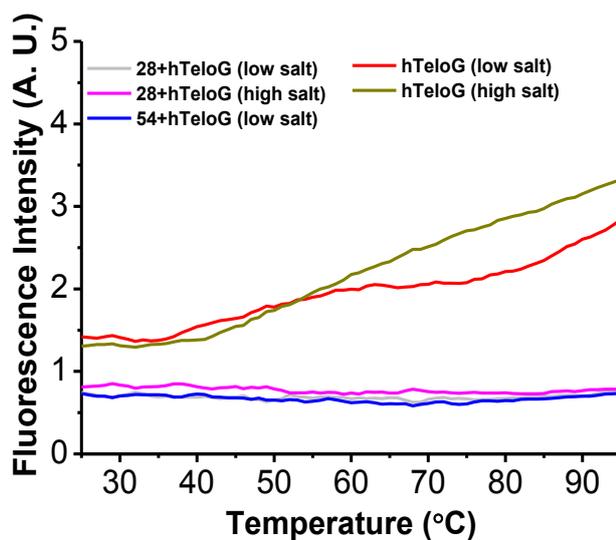
Figure 3.40   *FRET DNA melting experiments on compound **28** and **54** with hTeloG at acidic conditions, [$_{FRET}$hTeloG] = 200 nM, [ligand] = 100mM, [NaCl]$_{high\ salt}$ = 100 mM, [NaCl]$_{low\ salt}$ = 5 mM,, [sodium cacodylate] = 10 mM, pH 5.5. Fluorescent emission was measured at 533 nm.*
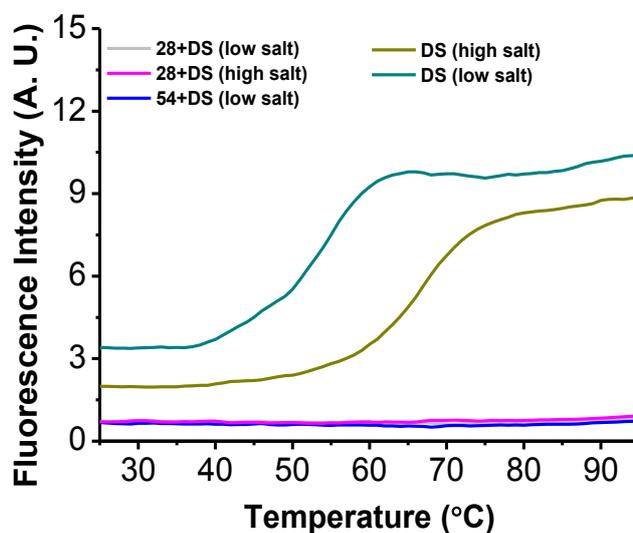
Figure 3.41   *FRET DNA melting experiments on compound **28** and **54** with double helical DNA at acidic conditions, [$_{FRET}$DS] = 200 nM, [ligand] = 100 mM, [NaCl]$_{high\ salt}$ = 100 mM, [NaCl]$_{low\ salt}$ = 5 mM, [sodium cacodylate] = 10 mM, pH 5.5. Fluorescent emission was measured at 533 nm.*

In contrast, the experiments with the i-motif forming sequences were repeated in buffers at pH 7.4 to mimic the previous experiments on G-quadruplex and duplex DNA. At pH 7.4, previously with duplex and quadruplex DNA, no fluorescence quenching effects were observed. Whereas addition of compound **28** and **54** to the i-motif forming sequences, a quenching effect is observed. Typically at physiological pH, the fluorescence is initially high at the start of the experiment for i-motif forming sequences which are more stable under acidic conditions (Figure 3.42a). This is because, at this pH the structure is not completely folded. Addition of either **28** or **54** gives rise to a reduction in fluorescence signal at room temperature (at the start of the melting experiment), implying induced DNA folding. A DNA-melting transition can then be observed in the FRET-melting assay for these samples (Figure 3.42a). This phenomenon was consistent with both hTeloC and the sequence from c-MYC (Figure 3.42b).
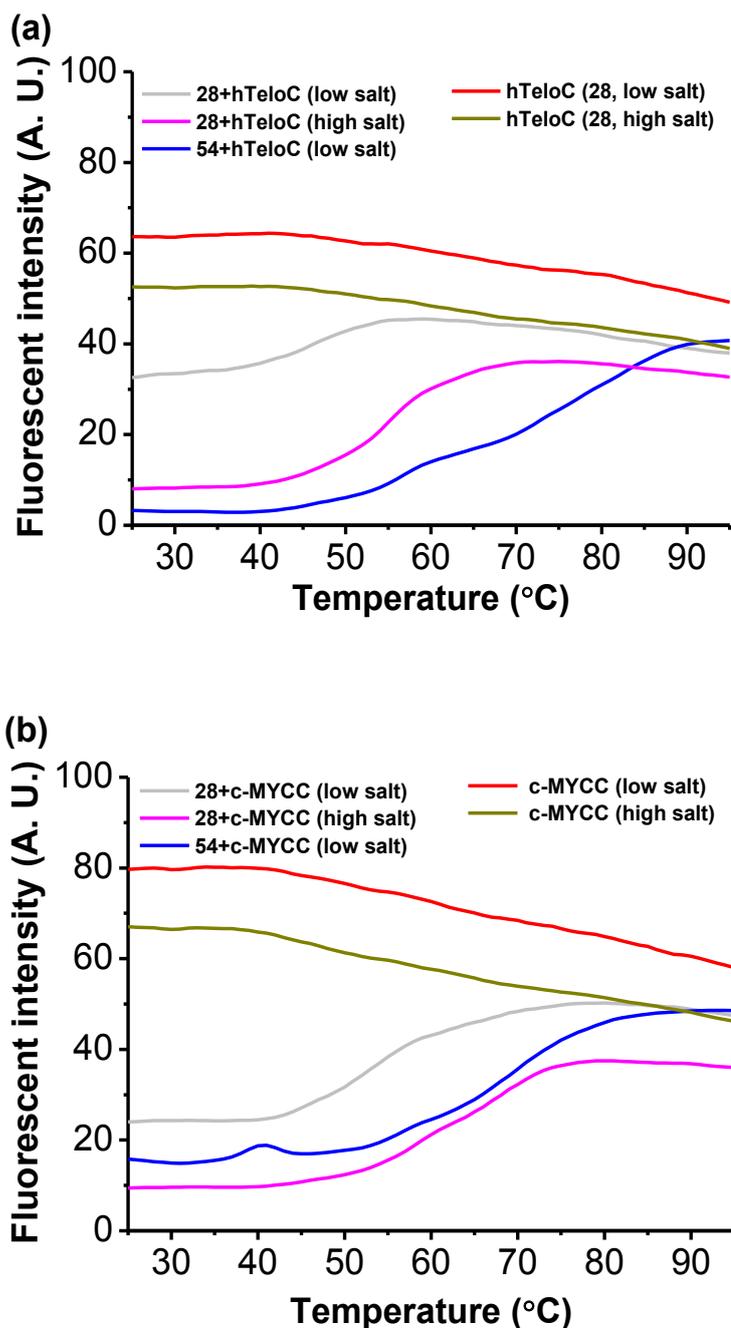
*Figure 3.42   FRET DNA melting experiments on compound **28** and **54** with hTeloC or c-MYCC at pH 7.4, a) hTeloC, b) c-MYCC; [FRETDNA]= 0.2 μM, [NaCl] high salt = 5 mM, [NaCl] low salt = 10 mM , [sodium cacodylate] = 10 mM, [ligand]= 100 μM. Fluorescent emission was measured at 533 nm.*

To explore the nature of this, both of the compounds **28** and **54** were examined at different concentrations (Figure 3.43). For both compounds a dose-dependent

increase in melting temperature was observed. Each time the effect can be reversed by increasing temperature. This is particularly interesting as both hTeloC and c-MYCC typically requires acidic pH to fold completely. These experiments therefore indicate that these maybe examples of ligands which can induce folding
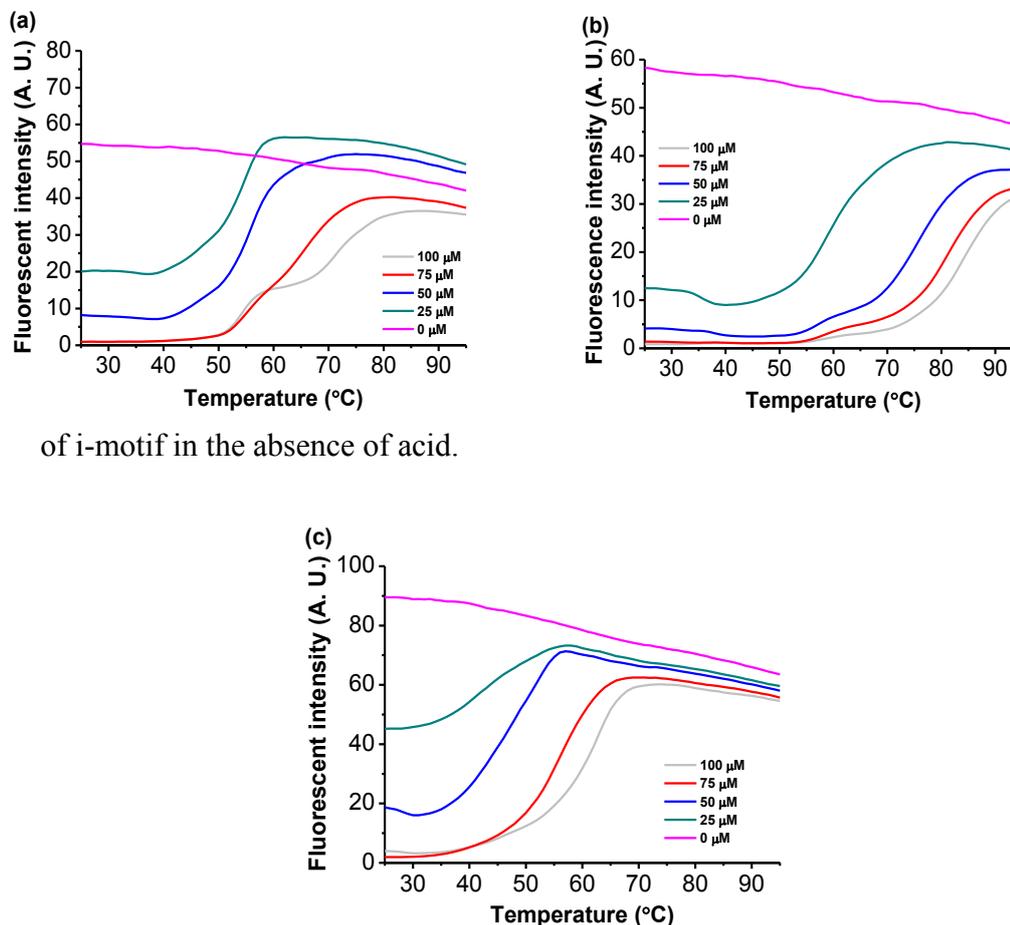
of i-motif in the absence of acid.

*Figure 3.43   FRET DNA melting experiments on compound **28** and **54** huamn telomeric motif at pH 7.4 ; a) compound **28**, [NaCl] = 5 mM, b) compound **54**, [NaCl] = 5 mM, c) compound **28**, [NaCl]= 100mM ; [$_{FRET}$hTeloC]= 0.2 µM, [sodium cacodylate] = 10 mM, [ligand]= 0 - 100 µM. pH = 7.4. Fluorescent emission was measured at 533 nm.*

In order to assess whether this effect was related to the pyrene motif attached to compound **28** and **54**, a reference compound, compound **58** (Figure 3.44), was used as a control. It was found that compound **58** has similar effects compared to

compound **28** and **54** with hTeloC at pH 7.4 both in 100 mM NaCl and 5 mM NaCl buffer. (Figure 3.44a and Figure 3.44b), but not to such a large extent.
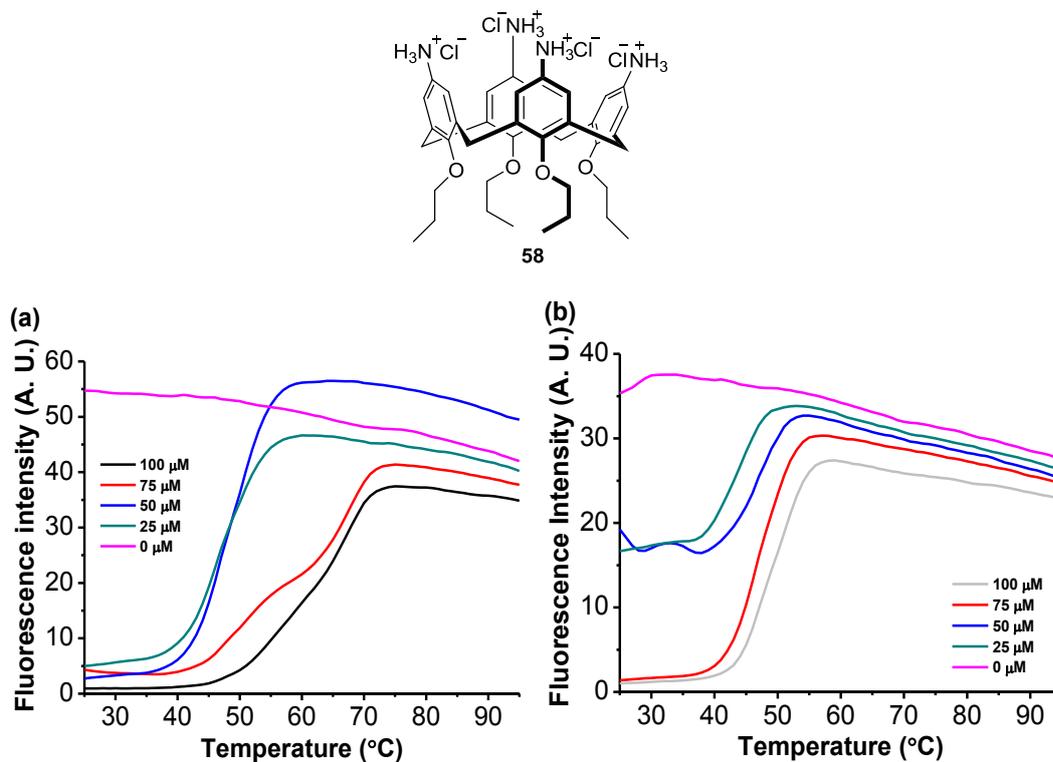
Figure 3.44   *FRET DNA melting experiments on compound **58** on hTeloC at pH 7.4; [FRET hTeloC] = 0.2 µM, [NaCl] = 5 mM, [sodium cacodylate] = 10 mM, [compound **58**] = 5 µM. a) [NaCl] = 100 mM, b) [NaCl] = 5 mM. Fluorescent emission was measured at 533 nm.*

By comparing compounds **28**, **54** and **58** in the FRET-melting experiments in 5 mM NaCl buffer (Figure 3.45), it is possible to see that the compound **28**, bearing one pyrene motif on the lower rim of the calixarene, possesses the highest stabilisation potential to hTeloC at pH 7.4 (T_m = 77°C), followed by compounds **58** (T_m = 61°C) and **54** (T_m = 58°C). These results suggest that the pyrene-motif on the lower rim may facilitate the interaction of hTeloC at pH 7.4. Surprisingly, it seems that the di-pyrene substituted compound **54** possesses reduced stabilization potential compared to the mono-substituted compound **28**. This could be because of steric effects due to the size of the two appended pyrene moieties. In the literature, it has previously been suggested that there was the possibility that the two pyrene moieties on the di-substituted calixarene **54** could have intermolecular π-π stacking interactions.[209] Thus, if the stabilization mechanism is based on π-π stacking on the end of i-motif, the mono-substituted calix[4]arene will behave better than the di-substituted calix[4]arenes, due to the absence of the intermolecular π-π stacking effects. For compounds **28** and **58**, some of the the melting curves appear to have two transitions before the sample is completely unfolded, suggesting a plausible second melting event in the process.
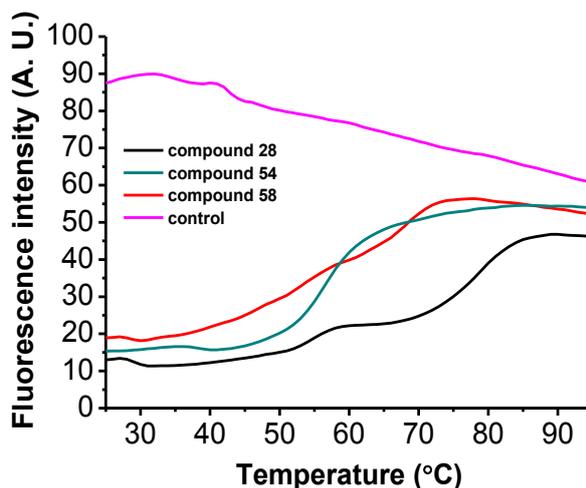
Figure 3.45   *FRET DNA melting experiments on compound **28**, **54** and **58** on hTeloC at pH 7.4; [FRET hTeloC] = 0.2 μM, [NaCl] = 5 mM, [sodium cacodylate] = 10 mM, [ligand] = 5 μM. Fluorescent emission was measured at 533 nm.*

3.3.8 FRET titration assay

FRET melting experiments monitor only one fluorophore attached on the oligonucleotides. As calixarene **58**, **28** and **54** appear to be folding i-motif forming sequences into i-motif even at pH 7.4, further investigations of ligands with DNA were carried out *via* a FRET titration, monitoring both fluorophores. The FRET titration was conducted *via* titrating the ligands into FRET-labelled oligonucleotides and monitoring the fluoroscence changes over a wavelength range on the spectra. By excitating the FAM fluorophore at 490 nm, as decribed in Section 1.3.3, a folded DNA structure will gives a significant fluorescent emission at 580 nm as a TAMRA emission, however, unstructured DNA will gives a corresponding FAM emission at 533 nm, as no FRET is taking place when the fluorophores are far apart. By continuously titrating ligand solution into the FRET-labelled DNA, a change in fluorescence signals will be observed whether there is folding or unfolding upon interaction.

With this technique, compound **28**, **54** and **58** were first studied with FRET-labelled hTeloC (100 nM) at pH 7.4 (Figure 3.46). At pH 7.4, when FAM is excited, the major fluorescence emission is observed at 533 nm. If a ligand helps to fold the hTeloC DNA into i-motif structure (or even a hairpin), the fluorescent emission for TAMRA should increase and the fluorescent emission for FAM should decrease. As compound **28** was titrated into this sample, the fluorescent signal for FAM (emission at 533 nm) was found to decrease; however, no concomitant fluorescent signal enhancement for TAMRA (580 nm) was observed (Figure 3.46a). Similar results were observed for compounds **54** and **58** as well (Figure 3.46b & c). This may suggest that the folding and stabilization effects observed in Figures 3.42, 3.43 and 3.44 result from calixarene-induced condensation of i-motif forming sequence.
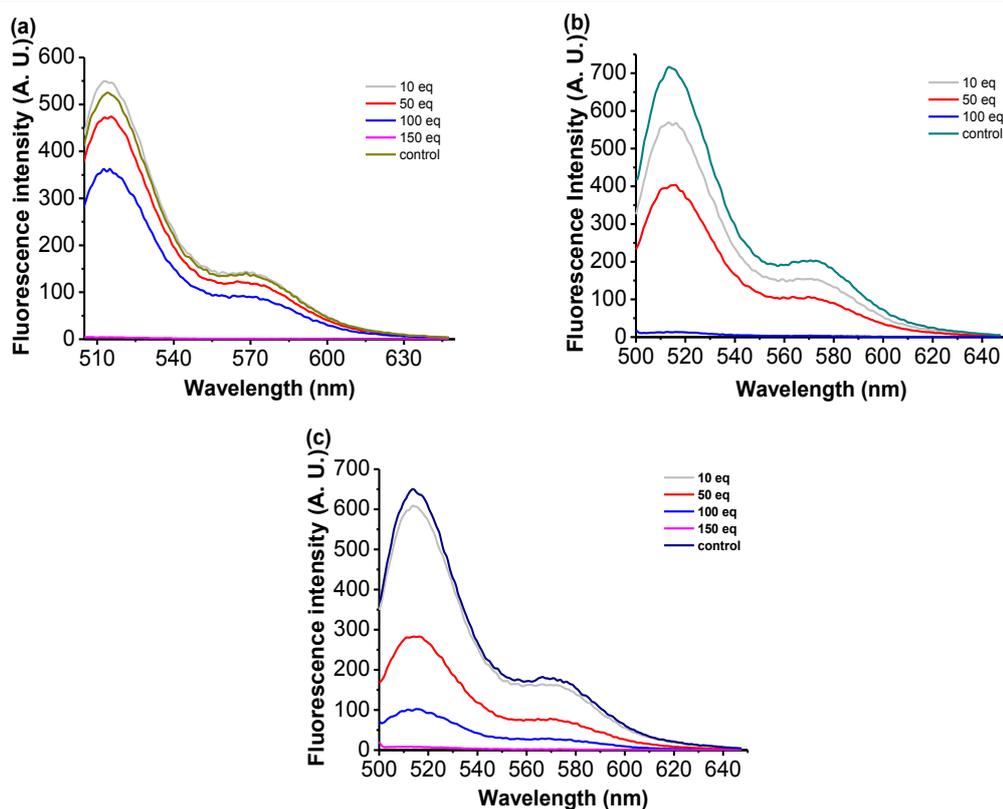
Figure 3.46   *FRET titration of compound 28, 54 and 58 with hTeloC at pH 7.4; (a) compound 28, (b) compound 54, (c) compound 58; [FRET hTeloC] = 0.1 μM, [NaCl] = 5 mM, [sodium cacodylate] = 10 mM. Fluorescent emission was measured from 500 nm to 650 nm.*

In order to explore whether compound **28** has any effects with other DNA secondary structures, analogous titration experiments were performed with the G-quadruplex forming sequence from the human telomere, hTeloG (Figure 3.47).
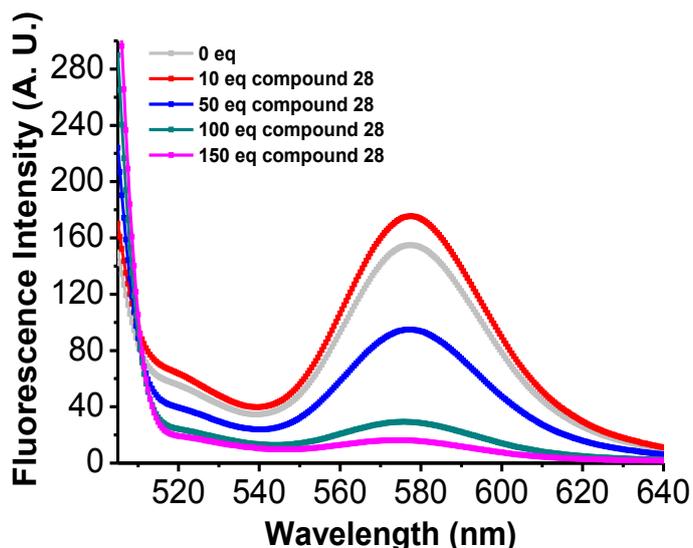
*Figure 3.47   FRET titration of compound **28** with hTeloG at pH 7.4;
[hTeloC] = 0.1 μM, [NaCl] = 100 mM, [sodium cacodylate] = 10
mM.. Fluorescent emission was measured from 500 nm to 650 nm.*

The fluorescence emission at 580 nm indicated that hTeloG sequences forms folded
G-quadruplex at pH 7.4 in the buffer conditions used. As compound **28** was titrated
into the system, the fluorescence at 580 nm drops. This phenomenon is consistent
with the data from Figure 3.35 in Section 3.3.7, implying calixarene-induced DNA
condensation or precipitation is occuring.

3.3.9   FID assay on calix[4]arene conjugates

As described in Section 2.2.4, i-motif fluorescent intercalator displacement is a
potential tool to study the relative affinity of i-motif binding ligands. FID has also
previously been used to give an indication of calixarene-DNA interactions (Section
2.3.5).[126] In order to study the plausible binding between calixarenes and human
telomeric i-motif, the FID assay was selected as a supporting method to study the
relative binding affinity of the calixarene conjugates.

In order to elucidate the possibility of calixarene conjugates interacting with
thiazole orange, the UV absorbance of compounds **28**, **54** and **58** was measured at

pH 5.5 as well (Figure 3.48). From the spectra, compounds **28** and **54** possess a UV absorbance of from 250 to 370 nm while calixarene **58** does not have any noticeable UV absorbance above 350 nm. Since in the FID system, the excitation wavelength is at 440 nm, the fluorescent excitation and emission should not be affected.
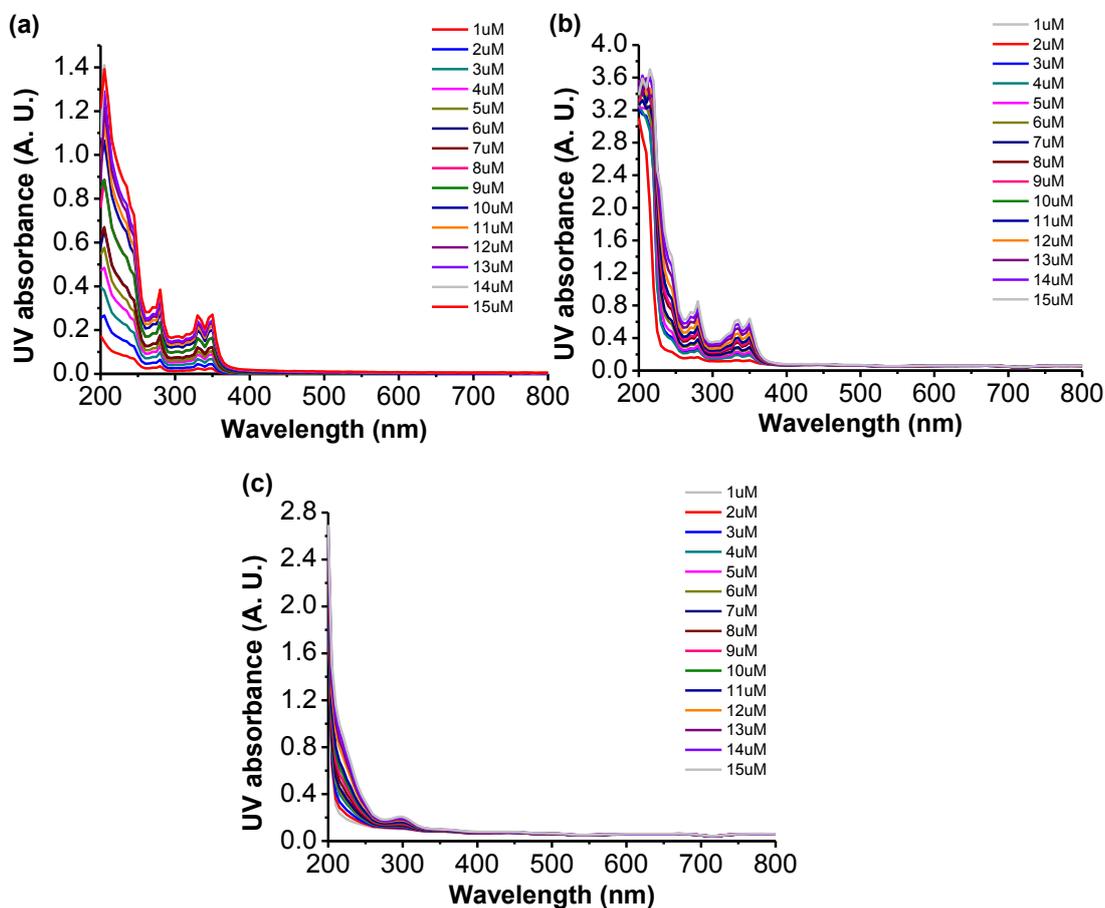
Figure 3.48   UV absorbance of calixarene conjugates at pH 5.5, a) compound 28, b) compound 54, c) compound 58,   [sodium cacodylate] = 10 mM. UV absorbance measured from 200 nm to 800 nm.

Compounds **28**, **54** and **58** were then tested in the FID assay using fluorescence spectroscopy. As shown in Figure 3.49, $_{\text{i-motif}}$DCs for the three compounds were all determined to be around 7.8 μM (**28**, $^{\text{i-motif}}$DC$_{50}$ = 7.8 ± 1.3 μM; **54**, $^{\text{i-motif}}$DC$_{50}$ = 7.6 ± 2.4 μM; **58**, $^{\text{i-motif}}$DC$_{50}$ = 7.9 ± 0.8 μM)and any differences between three calixarenes in binding human telomeric i-motif were within the error of experiment. This indicates the binding affinities of three compounds are almost the same.

However comparing to mitoxatrone and tobramycin discussed in Section 2.2.4, the i-motif DCs for three calixarenes (compound **28**, **54** and **58**) were relatively low, indicating that their binding affinity towards human telomeric i-motif is not as good as mitoxatrone and tobramycin (mitoxatrone, $_{\text{i-motif}}$ DC = 1.8 μM and tobramycin, $_{\text{i-motif}}$ DC = 2.4 μM, Section 2.2.4).
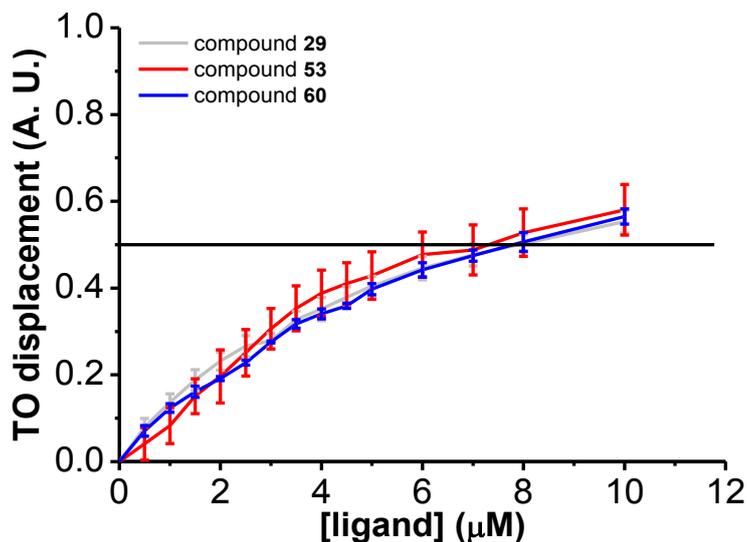
Figure 3.49　　*TO displacement calculation based on hTelo i-motif FID assay on calixarenes 28, 54 and 58 at pH 5.5 in 10 mM sodium cacodylate buffer. [hTeloC] = 1μM, [TO] = 2μM, [ligand] = 0 ~ 5μM. The sample was excited at 430nm and the fluorescence emission was measured at 450 nm.*

3.3.10　Circular dichroism

As the isothermal FRET titration experiments appeared to indicate the sample was precipitating or condensing, but this could have been complicated by the fluorophores, circular dichroism titration assays were also used to provide an insight using a different technique with unmodified DNA (Figure 3.50).
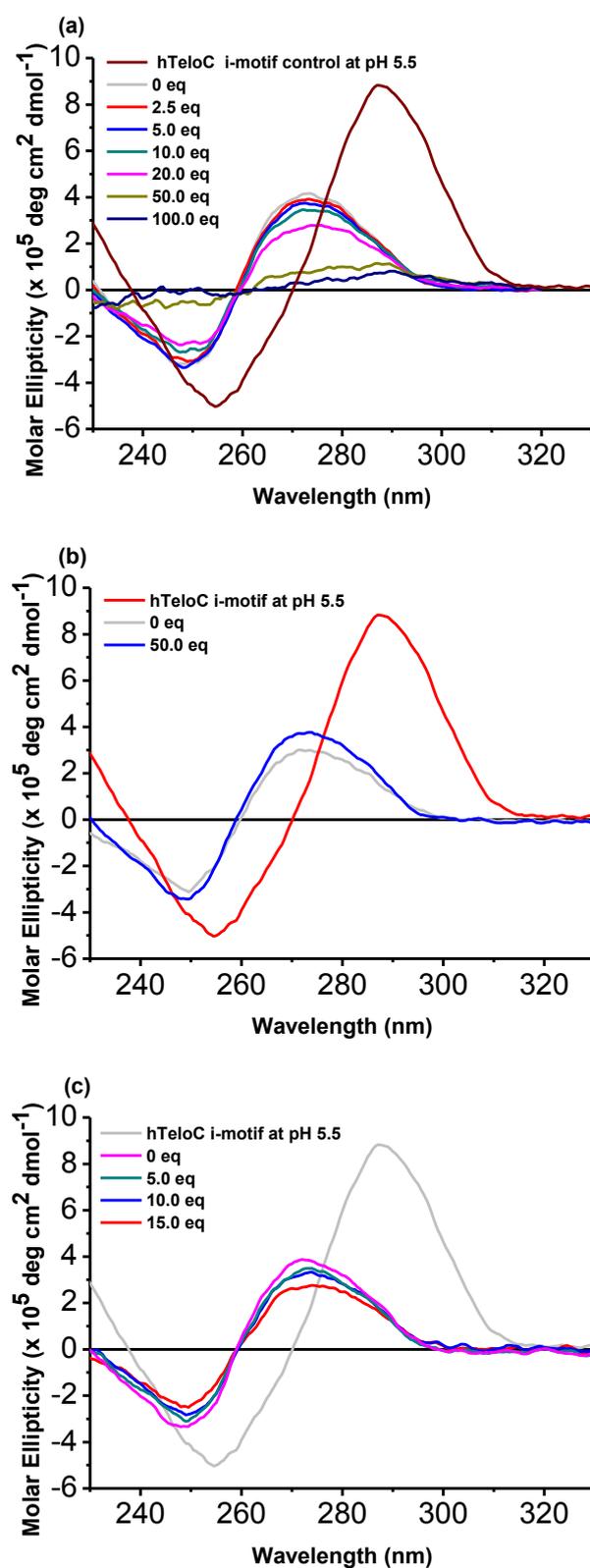
Figure 3.50   Circular dichroism spectra for calixarene with hTeloC at pH 7.4; a) compound **28**, b) compound **54**, c) compound **58**, [hTeloC] = 10 μM, [NaCl] = 100 mM, [sodium cacodylate] = 10 mM.

In the spectrum of hTeloC at physiological pH, the positive band at 274 nm and the negative band at 248 nm are consistent with unstructured, single-stranded DNA. For comparison, folded i-motif, at acidic pH gives rise to a positive band at 288 nm and a negative band at 255 nm (Figure 3.50).

As can be seen in Figure 3.50a, with the increasing concentration of compound **28** in hTeloC containing solution at pH 7.4, the intensity of CD signals for hTeloC decreased, indicating precipitation or DNA condensing events. When performing the circular dichroism on compound **54** and **58**, it is possible to see that the conformation of hTeloC did not change even 50 equivalent of ligands were titrated with the nucleic acids before clear compound precipitation happened in the circular dichroism cuvette (Figure 3.50). In this experiment, clear precipitation was observed, however, the concentration at which these experiments were performed was 10 times higher than those within the FRET-experiments, so this is not particularly surprising.

3.4   Conclusion

In order to prepare the calixarene conjugates, novel amino-functionalised calixarene platforms bearing alkynes, such as calixarene **33**, **37** and **41**, were prepared *via* existing methods from *p-tert*-butylcalixarene. These platforms all adopt the cone conformation and feature Boc-protected amine group on the upper rim and alkynes on the lower rim. Two fluorescent DNA binding moieties bearing azides: pyrene **44** and Nile Red **48**, were also prepared. Conjugating the calixarene platforms with the DNA binding moieties was performed by copper-catalysed alkyene-azide cycloaddition. In total, five calixarene conjugates were made and purified. Among them, compounds **28** and **54** featuring pyrene moieties were tested for binding to several oligonucleotides due to their comparatively good solubility in buffer. Thought the other three calixarene conjugates are all soluble in water, their solubility were very poor in even 5 mM NaCl and 10 mM sodium cacodylate

buffer (both at pH 7.4 and pH 5.5) and unsuited for biophysical studies.

A series of FRET DNA-melting assays and FRET titrations, indicate that compounds **28** and **54** are able to condense several types of oligonucleotides, including sequences that forms G-quadruplexes and i-motifs. Calixarene-conjugate induced DNA condensation is stable at acidic pH, even up to 95°C but at pH 7.4 can be reversed *via* introducing heat, which was observed in melting experiments. This indicates that these experiments, rather than showing a melting event, they are possibly indicating a cloud point for the ligand-DNA complex.[235,236] It has previously been reported that amphiphilic calix[4]arenes with long aliphatic chains can condense oligonucleotides with ease,[140] but previously it was thought that the long chains were required for this to occur. This work indicates that, even with shorter aliphatic chains (where n = 3), this is also possible.

As well as the FRET-based and circular dichroism experiments, the i-motif FID assay developed in the Section 2 was also used to assess the binding affinity of calixarenes **28**, **54**, and **58** at pH 5.5. It was confirmed that the binding affinity between these three calixarenes and hTeloC was the same within error (DC_{50} = 7.5 µM) and were weaker than the model compound involved in the FID assay, such as mitoxatrone ($^{\text{i-motif}}DC_{50}$ = 1.8 µM) and tilorone ($^{\text{i-motif}}DC_{50}$ = 2.1 µM).

Although the original goal in this chaper was to find calixanene conjugates targeting DNA secondary structures, there is still a lot of useful information provided from the study. In this project, it was also first time that FRET DNA-melting experiments, FRET-titrations and the TO-FID assay were used to study oligonucleotide condensation by calix[4]arenes, providing examples of new tools to study ligand-induced DNA condensation. These fluoroscopy-based experiment techniques could provide a quick and good indication of DNA condensation. Among them, FID assay can even be used to quantify the relative binding affinity. From this project, the lesson learned from a medicinal chemistry prospective is that

in designing biologically active calixarene conjugates, solubility is a key issue which requires considerable consideration.

Chapter 4　　Synthesis of cRGD-Conjugated Calix[4]arenes

4.1 Introduction

In this chapter, an attempt to use calixarene as a platform presenting cyclic RGD peptide was disclosed. The aim of this study was to target receptors expressed on tumour tissue, such as integrin.

4.1.1 Metastasis and Integrins

In the early stages of malignancy, cancer cells cluster into a single mass forming a benign tumor, which can be completely removed by surgery if caught at this stage. However, in the later stages, neoplastic cells obtain the ability of metastasis, invading the surrounding tissue and migrating to distant organ sites *via* blood and lymphatic vessels.[237,238] Metastasis is one of the key features for malignances and it contributes to over 90% of the mortality from solid tumors.[239] To eradicate metastasis successfully is a challenging task. Tumour cell metastasis starts with invasion of the surrounding tissue around the primary tumour. [239] Then tumor cells enter the bloodstream directly or *via* lymphatic nodes, to begin their journey towards distant organs. They stop at the first capillary bed encountered, extravasate the bloodstream and land on the 'foreign' tissue.[239]

In order to achieve proliferation, metastasis and migration, neoplastic cells have to interact with the extracellular matrix (ECM), an interlocking network of fibrous proteins, proteoglycans hyaluronic acid and chondroitin sulfate.[240] The ECM is also decorated with a wide range of signalling components that are crucial to cell adhesion, such as fibronectin, laminin, vitronectin and collagen-IV. These signalling components all binds to a family of extracellular receptor, integrins, to achieve their biological impacts.[240]

Integrins are divalent cation-dependent heterodimeric membrane glycoproteins mediating adhesion to the ECM (Figure 4.1).[241] Apart from ECM, integrin also binds to immunoglobulin, growth factors, cytokines and matrix-degrading

proteases. The structure of integrin consists of two non-covalently associated α-subunit and β-subunits. Up to now, 18 α-subunits and 6 β-subunits have been found, assembling into 24 distinct subtypes in the integrin family. Among these, $\alpha_v\beta_5$, $\alpha_v\beta_3$ and $\alpha_5\beta_1$ subtypes have beens show to be involved in metastasis of solid tumor and angiogenesis.[241-243]
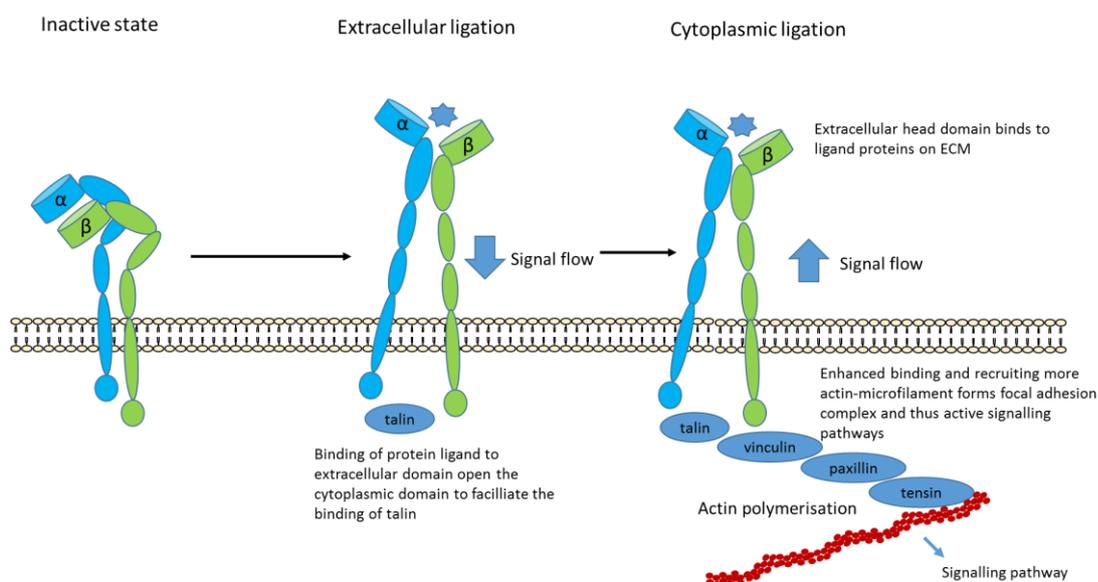
Figure 4.1: Diagram of integrin activation and bidirectional signal tranduction.

The structure of the α- and β-subunits are quite similar. The rod-shaped subunits possess a very large extracellular recognition region for protein binding. A single transmembrane region connects the large extracellular head-domain and the much smaller cytoplasmic tail (Figure 4.1).[241] Integrins are involved in the regulation of multiple consitutively activated pathways in neoplastic cells, such as Src, Rho GTPases, FAK and Erk.[241] Cellular signaling mechanism of integrins are very different from other receptors since they possess unique bi-directional signaling pathways.[244] Extracellular ligation unbends and elongates the dimer and separates the folded cytoplasmic domain on α- and β- subunits, allowing association of the cytoplasmic tail with signaling complexes and the cytoskeleton. Conversely, intracellular ligation between integrins and actin-binding proteins also promotes the separation of the extracellular domain and enhances integrin binding to substrate proteins on ECM.[244] Integrins themselves possess limited kinase activity.

However, by clustering themselves and recruiting multiple kinases and adaptor proteins to form focal adhesion complexes, integrin can play a vital role in cancer metastasis, embryonic development and immunological function.[245] Pathologically relevant subtypes crosstalk with growth factor receptors and exert their influence on oncogene expression as well.[238] Integrins serve as transmembrane links between extracellular contacts and actin-microfilament. Loss of integrin ligation also contributes to initiation of cell apoptosis.

Integrin $\alpha_V\beta_3$

Integrin $\alpha_V\beta_3$ was the first α_V integrin to be correlated with pathological angiogenesis.[241] It is normally expressed on endothelium cells on blood vessels in human biopsy samples is stimulated *via* angiogenic growth factors such as TNFα, IL8 and Basic fibroblast growth factor (bFGF) and cannot be detected in normal human tissue.[245] Inhibition of integrin $\alpha_V\beta_3$ by antagonism has been shown to blocked angiogenesis and inhibits tumour growth in multiple animal models including human tissues. [246-248] In the field of cancer research, Integrin $\alpha_V\beta_3$ is the most intensively studied subtype in the family. Overexpression of integrin $\alpha_V\beta_3$ was found associated with bone metastasis[249] and resistance of chemotherapeutics in breast cancer.[250] In the study of glioblastoma, overexpression of integrin $\alpha_V\beta_3$ was detected in the invasive margins of the tumour.[251] The mRNA of β_3 integrin is detected with a high abundance in tumour associated gastric mucosa, serving as an interesting prognostic marker for gastric carcinoma. [252] In melanoma, $\alpha_V\beta_3$-expressing facilitate tumour survival, invasiveness and growth. [253,254]

Integrin $\alpha_V\beta_5$

Integrin $\alpha_V\beta_3$ crosstalks with the FGF receptor and facilitates angiogenesis downstream. [242] A similar event was observed between integrin $\alpha_V\beta_5$ and VEGF-2 receptor.[242] In pancreatic cancer, hyperactivity of the EGF pathway has been correlated with integrin $\alpha_V\beta_5$. Interestingly, integrin $\alpha_V\beta_5$ is only active after being

activated by EGF triggered SRC phosphorylation of the p130CAS substrate domain and subsequent activation of the GTPase RAP1A.[255] Although pathological angiogenesis does not necessarily require β_5 subunits,[256] integrin $\alpha_V\beta_5$ is more prevalent on cerebral cavernous malformations than integrin $\alpha_V\beta_3$.[257]

4.1.2   cRGD peptide

In studies of fibronectin on the ECM in the early 1980s, Rouslahti and Pierschbacher found a tetrapeptide sequence, Arg-Gly-Asp-Ser (RGDS), which plays a vital role in cell attachment.[258-260] Later, other integrin binding substrates on the ECM, such as laminin, vitronectin and collagen-IV were also found that contained the tripeptide RGD sequence in their binding site.[261-264] This marked the beginning of utilising RGD peptides to target integrins (Figure 4.2). Up to now, there are at least seven integrin subtypes found to bind the RGD peptide.[265,266]
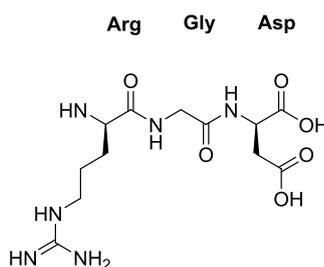
Figure 4.2:   Native RGD-peptide structure

The conformation of the RGD peptide affects its binding affinity drastically since in the natural ligands for integrin, the RGD sequence was presented in a rigid form. Thus in the development of synthetic integrin-binding ligands, cyclic peptide strategies have been applied. In cyclised systems, the RGD sequence is flanked by other amino acids to form a rigid ring, enhancing the binding affinity.[267] Compared to linear RGD sequences, cyclic RGD (cRGD) peptides are around 10-fold more potent.[268] Apart from enhancing affinity, cyclic peptides also offer the advantage of lower proteolytic degradation of the peptide.[267] Cyclic peptide strategies have also been applied and shown to increase the selectivity of the RGD sequence. Linear RGD peptide sequences lack selectivity to specific integrin subtypes while a kinked

cyclic pentapeptide **59** (cRGDfV) possess high selectivity towards integrin $\alpha_V\beta_3$ and $\alpha_V\beta_5$ and stretched cyclic hexapeptide **60** also has integrin $\alpha_{II}\beta_3$ selectivity (Figure 4.3).[269]
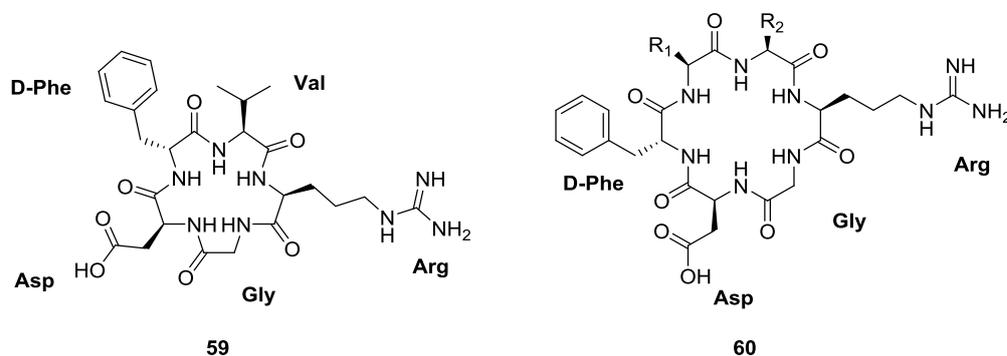
*Figure 4.3: Structures of penta-peptide **59** specific for integrin $\alpha_V\beta_3$ and $\alpha_V\beta_5$ & cyclic hexapeptide **60** specific for integrin $\alpha_{II}\beta_3$*

Cilengitide

In the past 10 years, several RGD-based integrin ligands have been developed. The most important one among them is c(RGDf[N-Me]V), known as Cilengitide (EMD 121974, developed by Merck KGaA, **61** Figure 4.4).[270] It binds integrin $\alpha_V\beta_3$ with sub-nanomolar affinity ($^{\alpha V\beta 3}IC_{50} = 0.58$ nM) and low-nanomolar affinity to integrin $\alpha_V\beta_5$ and $\alpha_5\beta_1$. In terms of selectivity, cilengitide exhibits a 1000-fold preference towards integrin $\alpha_V\beta_3$ over platelate integrin $\alpha_{II}b\beta_3$ ($^{\alpha IIb\beta 3}IC_{50} = 860$ nM).[271] Cilengitide fulfills its antineoplastic effect by inhibiting angiogenesis and inducing growing endothelium cell apoptosis *via* blocking the interaction between integrins with their ECM ligands, inhibiting proliferation and increasing apoptosis in cell lines and causing tumor regression.[247,272-274] Cilengitide is the only anti-integrin cRGD peptide being tested in Phase III clinical trials against glioblastoma[275] and in Phase II clinical studies for metastatic melanoma,[276] non-small cell lung cancer[277] and prostate cancer.[278,279] However, it has failed to show significant improvement in the overall survival in its Phase III clinical studies for glioblastoma.[275] Cilengitide incorporates a *N*-methylated valine. Here, the *N*-methylated amino acid provides further metabolic stability and improved

bioavailability to the peptide drug.[280,281] The guanidinium group on the arginine was shown to bind in a narrow groove in the α_V subunit by forming salt bridges with two aspartate residues on α_V subunit. An aspartate residue on the cilengitide coordinates with divalent cations such as Ca^{2+} and Mg^{2+}, while the glycine in the middle provides steric restrictions to the cyclopentapeptide. These three crucial amino acid residues cannot be changed. The D-phenylalanine on the cilengitide imposes optimal side chain orientation and hydrophobic interactions with integrin.[270]
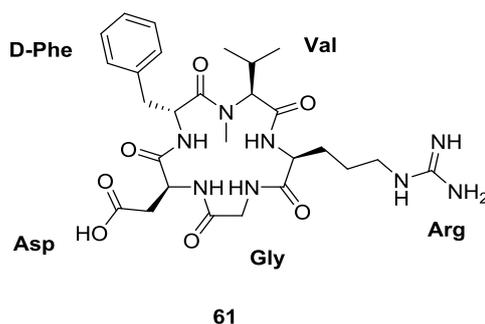
Figure 4.4   Cilengitide 61

c(RGDfK)

Cyclic RGDfK **62** (Figure 4.5) was developed by Kessler and his co-workers.[282] Compared to cligentide, c(RGDfK) is a relatively simple molecule and has a low nanomolar-range binding affinity towards integrin $\alpha_V\beta_3$ ($^{\alpha V\beta 3}IC_{50}$ = 4.2 nM) and high selectivity over integrin $\alpha_{II}\beta_3$ ($^{\alpha II\beta 3}IC_{50}$ = 1800 nM).[282] c(RGDfK) has been extensively used as cell-recognition motif or cell adhesion enhancer after attachment to other biomaterials, such as dendrimers,[283] nanoparticles[284] and radioactive isotope-containing compounds.[285]
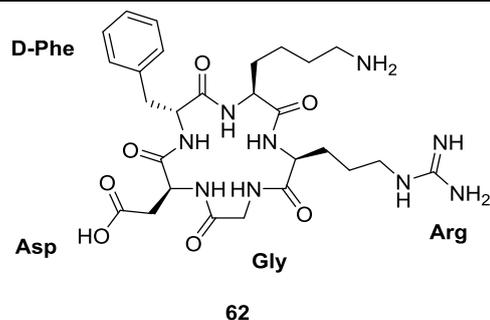
Figure 4.5   *c(RGDfK)* **62**

Peptidomimetics containing RGD peptide

The RGD peptide can also be presented by non-peptide scaffolds. RGD peptide presented by a 4-aminoproline molecule, **63** (Figure 4.6), shows nanomolar inhibition affinity to integrin $\alpha_V\beta_3$ and sub-micromolar inhibition affinity to integrin $\alpha_V\beta_5$.[286] A similar inhibition affinity to integrin $\alpha_V\beta_3$ and $\alpha_V\beta_5$ can also be observed in a peptidomimetic **64** (Figure 4.6) featuring a 7,5-fused bicyclic lactam.[287] Compounds **63** and **64** have even longer half-lives comparing to cilengitide. Peptidomimietics can also be functionalised with other substituents on the non-peptide scaffold, enabling incorporation on other biomaterials.[288]
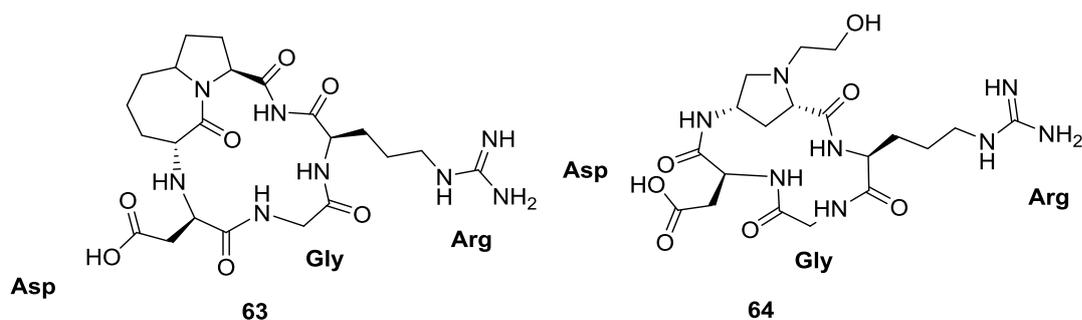
Figure 4.6   *Peptidomimetics* **63** *and* **64**

Cyclic RGD peptide made from cellular recombination

Oxidation of remote thiols on cysteine to form a disulfide bond is a common strategy for peptide macrocyclisation.[289] This method was also applied to cyclic

RGD peptide.[290] Peptide **65** (Figure 4.7) was originally found *via* phage display and is over 200 times more potent than the commonly used linear peptide.[291] Peptide **65** has the potential to conjugate onto virus or proteins by recombinant means. In addition, conformational studies suggested that peptide **65** can adopt two additional conformations possessing lower binding affinity other than its preferential bicyclic structure.[292]
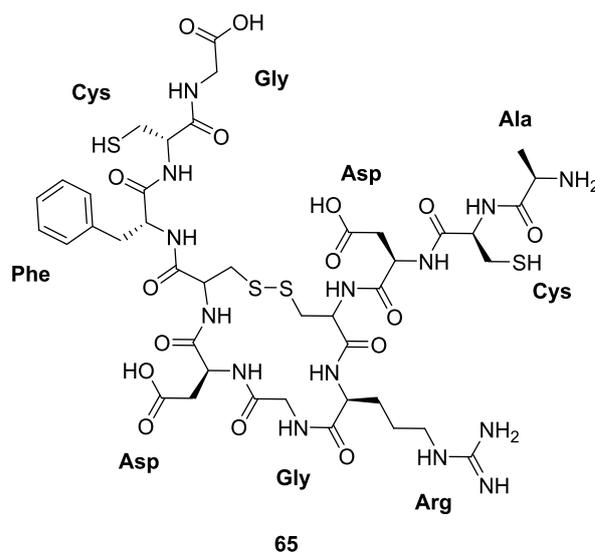
Figure 4.7   *Peptide* **65**

4.1.3   Tumour targeting strategy *via* cRGD peptide

Although cilengitide failed to progress beyond Phase III clinical trials, targeting integrin $\alpha_V\beta_3$ and $\alpha_V\beta_5$ *via* cyclic RGD peptide is still a good anti-metastasis strategy. In recent years, considerable effects have been focused on using cyclic RGD peptides as tumour targeting motifs for drug delivery and imaging purposes.

Diagnostic Imaging

In oncology, early discovery and treatment of malignancy can drastically elongate patients' life expectancy through treatment of the cancer *via* surgery or chemotherapeutics or radiotherapeutics at a very early stage. Oncologists rely on medical imaging methods to make the correct diagnosis. Cancer-related integrins,

such as $\alpha_V\beta_3$ & $\alpha_V\beta_5$, are highly correlated with angiogenesis, thus serving as interesting targets for early tumour diagnostics. In the field of medical imaging, RGD peptide conjugated compounds can be used in positron emission tomography (PET), molecular magnetic resonance imaging (mMRI) and optical imaging.[293]

PET generally requires a radioactive tracer, usually ^{18}F or ^{68}Ga compouds. RGD-based PET tracers have already been used for patients (Figure 4.8). Among them, [^{18}F]Galacto-functionalised peptide **66** was the first FET tracer to be studied in human.[294] The lysine of c(RGDfK) was derivatised with a ^{18}F-labelled sugar moiety *via* an amide bond. It takes around 200 minutes to prepare before being used and the radiochemical yield is about 30%.[295] Compound **66** had been studied for patients with breast cancer, melanoma and sarcoma in clinical trials.[294,296] Another similar example was compound **67**, a PET tracer developed by Siemens Molecular Imaging Inc. In their preparation, radioactive ^{18}F was conjugated to the galactose using 'click' chemistry. Conjugate **67** takes about 90 minutes to prepare and is more suitable for automated synthesis.[297] Conjugate **68** (NOTA-RGD) was made by coupling SCN-Bz-NOTA to c(RGDyK). By incubating with ^{68}Ga cation for 5 minutes, the radioactive form of **68** was made by cation caging.[298]
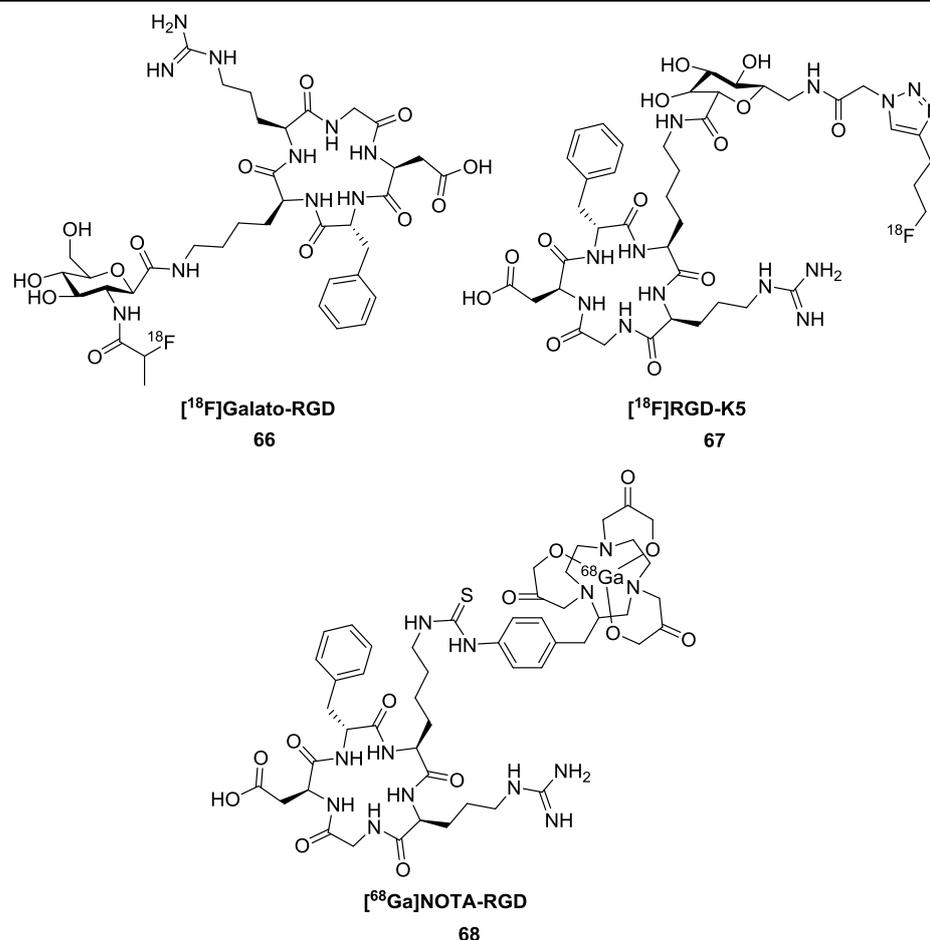
Figure 4.8   Examples of RGD peptide incorporated PET tracer **66**, **67** *and* **68**

When RGD-peptide are conjugated with fluorophores, conjugates can be used for optical screening. Compared to radioactive tracers in tumour imaging, RGD-conjugated fluorophores are easier to prepare, cost-effective and safer.[299] Near infrared (NIR) fluorophores are of great interest to scientists due to their easy synthesis and chemical and optical modifications. Their well-established bio-safety prolife will also benefit translational and clinical studies. However, the majority of the NIR fluorophores are not tumour-specific, requiring conjugation of tumour-recognising motifs for cancer diagnosis.[299] The first cRGD peptide conjugated fluorophore was a Cyanine 5.5 with c(RGDfK) or c(RGDyK) (Figure 4.9). Conjugate **69** was tested targeting on Kaposi's sarcoma in animal model *via* dynamic fluorescence imaging, showing increased uptake in tumour. Another conjugate **70** has been studied in a human orthotopic brain tumour in a mouse

xenograft model *via* 3-D optical imaging,[300,301] showing promising tumour imaging result (IC$_{50}$ = 42.9 nmol/L). The signal-to-noise ratio almost reached 2:1 24 hours after the administration of **70**.

69 (X= H), IC$_{50}$ = 2.5 μmol/L

70 (X=OH), IC$_{50}$ = 42.9 nmol/L
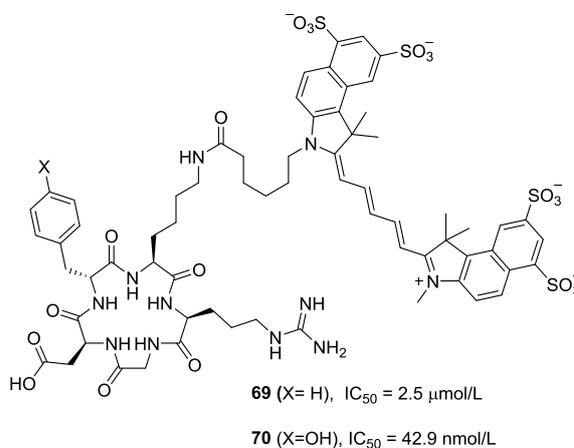
*Figure 4.9   Peptide conjugate **69** and **70**, the first generation Near infrared*

fluorophore incorporated with cyclic RGD peptide

Mono-RGD appended fluorescent ligands normally showed impaired integrin binding compared to cyclic RGD peptides without fluorescent tags.[302] A method to overcome this drawback is to use multivalent platform to increase the number of peptide motifs (Figure 4.10). Conjugate **71** is the peptide-dimer tagged with Cyanine 5.5 while conjugate **72** was the peptide-tetramer tethered with Cyanine 5.5. Both **71** and **72** are derivatives of **70**. In testing with human glioblastoma astrocytoma cells, **71** showed doubled binding affinity (IC$_{50}$ = 27.5 ± 1.2 nmol/L) compared to monomer **70** (IC$_{50}$ = 42.9 ± 1.2   nmol/L) while **72** demonstrated almost 4-fold binding affinity (IC$_{50}$ = 12.1 ± 1.3 nmol/L) compared to **70**.[302] However, both **71** and **72** failed to show significant enhancement in tumour imaging compared to **70**. This might due to the fact that multimerisation of RGD changes the size of the conjugate, overall charges and hydrophilicity of the peptide, thus altering the *in vivo* pharmacokinetics and clearance from non-targeted tissue.[302]

71, IC$_{50}$ = 27.5 nM

72, IC$_{50}$ = 12.1 nM
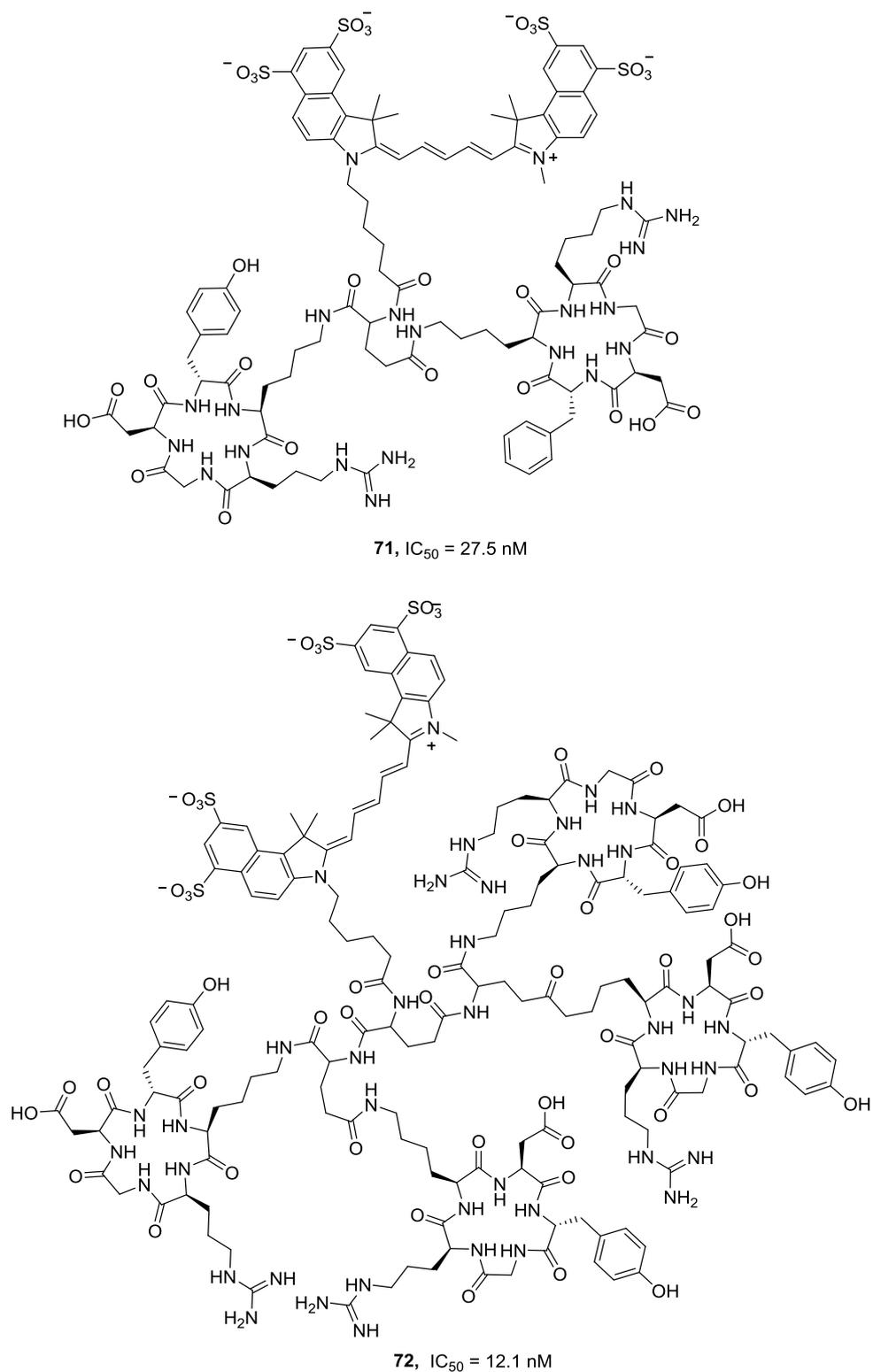
Figure 4.10   cyclic RGDyK tethered dimer 71 & tetramer 72

Apart from fluorescent and radioactive tags, nanoparticles, such as quantum dots (QDs), magnetic iron oxide nanoparticles, polymeric nanoparticles, carbon

nanotubes and gold nanoparticles, can also be used as imaging agent in tumour diagnostics.[299] Quantum dot **73** made from CdTe coated with ZnS was the first RGD peptide-labeled QDs (Figure 4.11). After the attachment of c(RGDyK), **73** showed integrin-specific binding to integrin–positive human breast cancer and primary glioblastoma cell lines in *ex vivo* experiments.[303] In the later experiment in mice, **73** showed increased signal intensity and improved photostability compared to conjugate **70** as a reference agent.[303]

73
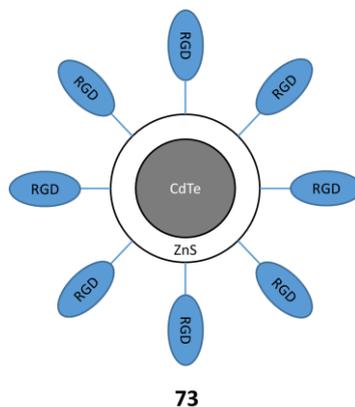
*Figure 4.11   Quantum dot **73** tethered with cyclic RGDyK peptide*

Drug delivery

Nanoparticles within certain sizes, normally 20 to 400 nm in diameter, are able to target tumours in living organisms passively *via* enhanced permeability and retention (EPR) effect. The EPR effect is attributed to abnormal fluid transport dynamics in tumour vessels as a result of uncontrolled angiogenesis, providing second order targeting.[304] Thus, in order to further enhance nanoparticles' tumour targeting, incorporating tumour targeting motifs, such as cyclic RGD, on the surface of nanoparticles is a sensible choice. As it can enhance internalisation by tumour cell presenting high levels of integrin.

Cyclic RGD functionalised nanoparticles **74** encapsulating doxorubicin were found to possess a 15-fold increase in anti-metastatic activity compared to nanoparticles without RGD coating in a doxorubicin-insensitive murine C26 colon carcinoma

model (Figure 4.12).[290,305,306]

74
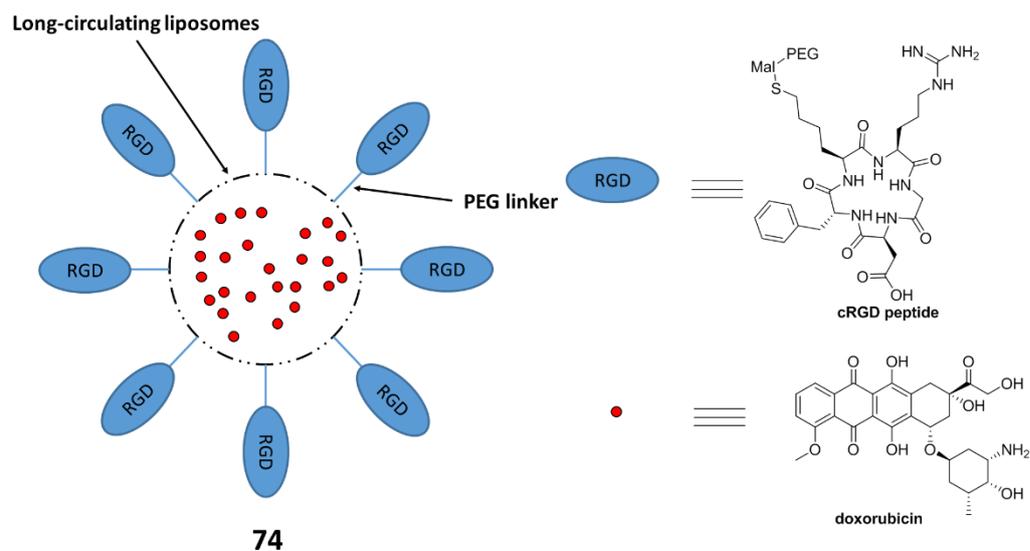
Figure 4.12   cRGD-Mal-PEG-tethered Long circulating liposomes (LCL) system carrying doxorubicin [258,259]

In a later study, biodegradable PLAG-based nanoparticles functionalised with linear RGD peptide (**75)** or RGD peptidomimetics (**76**) have been loaded with paclitaxel and showed targeted cytotoxicity towards neoplastic endothelium both in *in vitro* and *in vivo* models (Figure 4.13).[307]
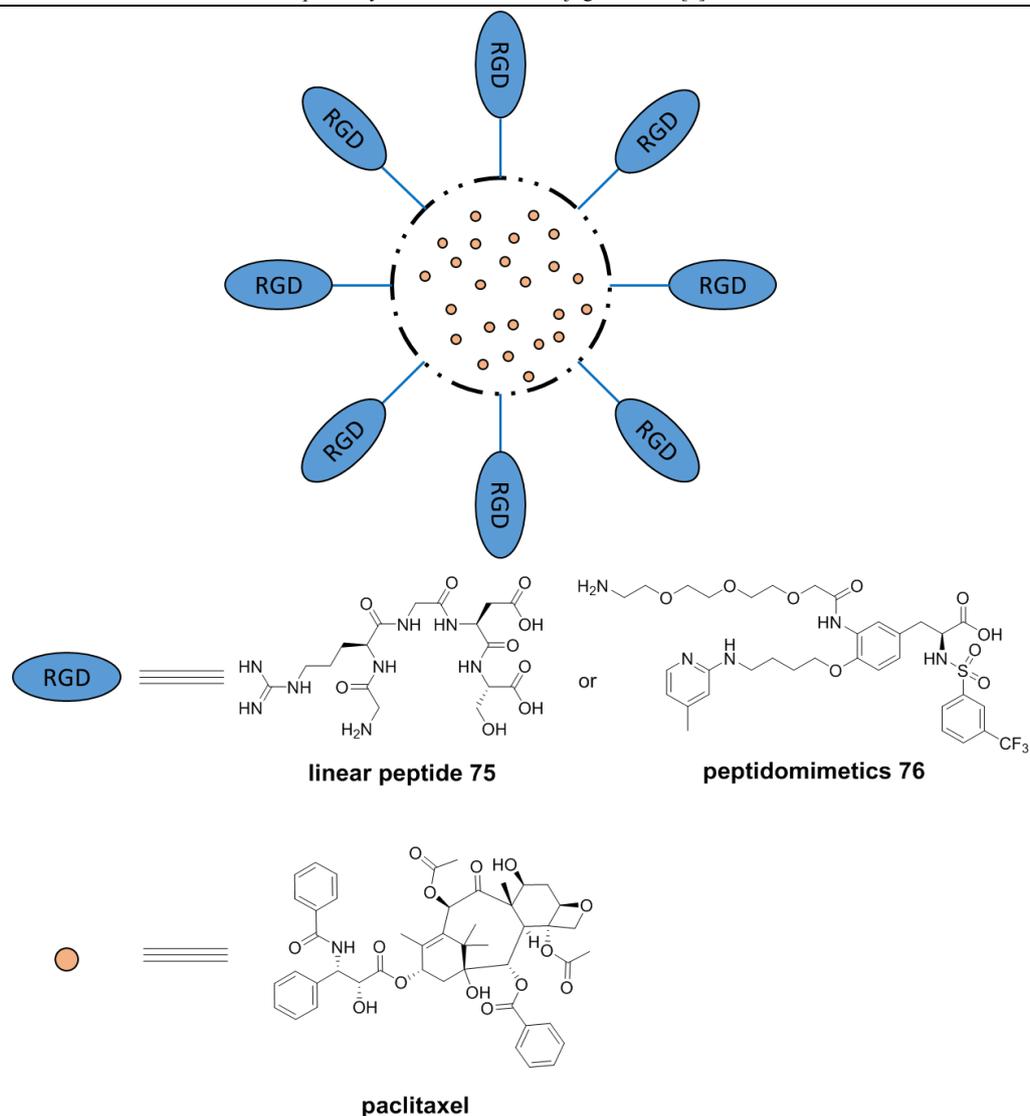
Figure 4.13   *PLAG-based nanoparticle bearing linear RGD peptide **75** and RGD peptidomimetics **76** carrying paclitaxel for drug delivery*

Biomolecules such as proteins and anticancer nucleic acids, can be grafted with RGD peptide to target angiogenesis and metastasis.[241] The clinical use of tumour neocrosis factor-α against malignancy was restricted by its general toxicity. However, by incorporating RGD4C peptide to the protein *via* DNA recombination, modified TNF molecule shows antitumour effects in mice at subnanogram doses when co-administrated with melphalan.[308] The same strategy was applied on TRAIL, endostatin and Interleukin 2 to show targeted antineoplastic effects.[147,309,310]

Another study in the RGD delivery strategy is to deliver nucleic acid, such as DNA or RNA oligomers into the cell, either acting as targeted gene expressing or gene silencing approaches in gene therapeutics. Unmodified DNA and RNA cannot get across the cell membrane easily due to their negative charges. Apart from that, selectivity of gene therapeutics is generally very poor as well. Thus incorporating RGD peptide to nanoparticles loaded with nucleic acids is expected to reduce the side effects as well as increase the efficacy in gene therapeutics.[241] Developed RGD-tagged nucleic acid delivery system includes siVEGFR2, siPLXDC1, anti-HIF-1α siRNA, anti-miR-296 AMO and so on.[311-314]

4.1.4 Cyclic RGD peptide and macrocycles

Although the cyclic RGD strategy is widely used in targeting tumour tissue, examples of cRGD peptide tethered calix[4]arene are quite limited. One example is to use the calix[4]arene as a rigid platform to lock the conformation of the RGD peptide (Figure 4.14).[315] Compound **78** was made *via* solid phase peptide synthesis from a calixarene platform bearing amine and carboxylic acid on the upper rim (**77**). However there is no published biological data on the activity of this compound (**78**).
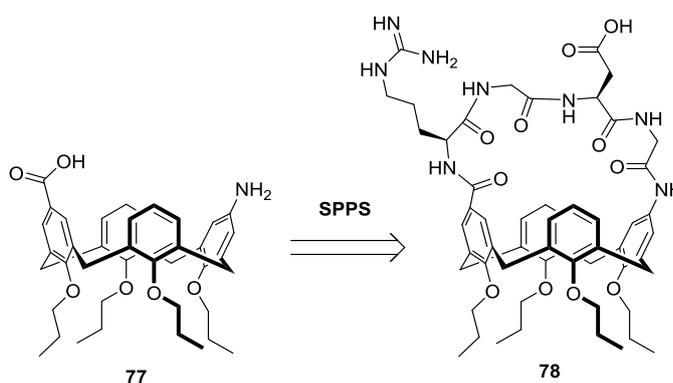
*Figure 4.14 Example of cyclised GRD peptide on calixarene **78***

In another approach, a cryptophane host is conjugated with cyclic RGD peptide (conjugate **79**, Figure 4.15).[316,317] Nano-structure **79** was prepared to perform

fluorescence imaging in cells *via* an Alexa Fluor® 488 fluorophore. **79** binds $\alpha_v\beta_3$ abd $\alpha_{IIb}\beta_3$ integrin with nanomolar affinity (IC$_{50}$ = 20-30 nM) but possessing moderate cytotoxicity to human lung fibroblasts at relevant biosensing concentrations. Cellular internalisation experiments were monitored by confocal laser scanning microscopy and hyperpolarized ^{129}Xe NMR with human pancreas cancer and lung fibroblasts cell lines. It was found that **79** had better internalisation in pancreas cancer cells than in lung fibroblast cells, suggesting a viable targeted delivery of cryptophane to tumour tissue.[316]

As introduced in Section 1.4.3, calixarenes act as a good platform for targeting receptors and enzyme binding. Given the platforms already synthesied within this thesis, this provided a good foundation to develop a new platform for presenting a cyclic RGD peptide.
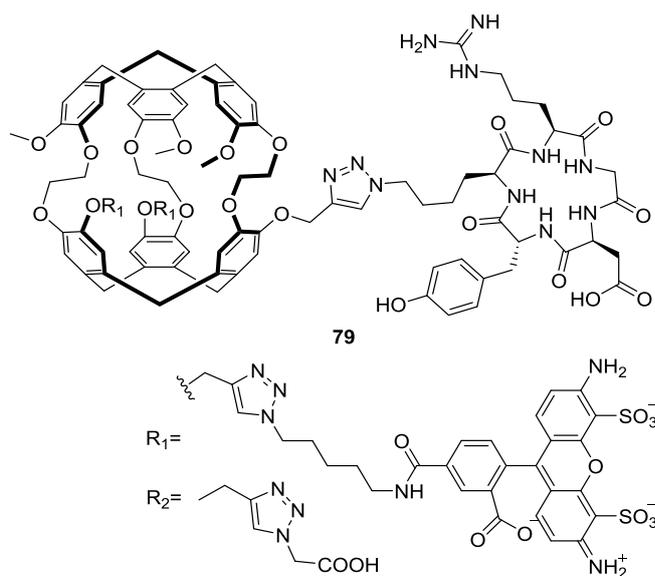
Figure 4.15 cRGD tethered cryptophane 79

4.2   Aims

The aim of this project is to explore the possibility of applying calixarenes as a platform to present cyclic RGD peptide and fluorescent labels to achieve a bioimaging function in a single molecule. The strategy to achieve this and separate aims are:

1)   Synthesise a cyclic RGD peptide functionalised with an alkyne moiety

2)   Synthesis of a derivatised calix[4]arene featuring orthoganol functional groups to enable two consecutive click reactions to introduce two different types of functionalization

3)   Synthesis of an appropriate fluorescent tag which can be added to the calixarene platform

4)   Construct the calixarene target calixarene conjugates bearing fluorophores and cyclic RGD peptides

4.3   Results and discussion

4.3.1   Calixarene conjugate design

In order to develop a bioimagining system, three major components needs to be considered: the targeting ligands, fluorophore and the linker. In this project, calix[4]arene was selected as the linker mainly due to its well-known multivalency (Figure 4.16) and also the prior experience developed in the duration of this PhD. The presence of four binding positions on the upper-rim of the calixarene which can be functionalised with cyclic RGD peptide combined with the potential to functionalise the lower rim of the calixarene with fluorescent substituents or NIR imaging agents makes it a suitable linking molecule.

As described in Section 4.1.3, the direct attachment of fluorophores or toxins to a cyclic RGD peptide has been shown to reduce its relative binding affinity to integrin. Given the calixarene platform can offer multiple sites for recognition, it is hypothesized that the accociated multivalency may overcome this problem.

R_1= cyclic RGD peptide motif
R_2= fluorophore

Calix[4]arene peptide conjugate
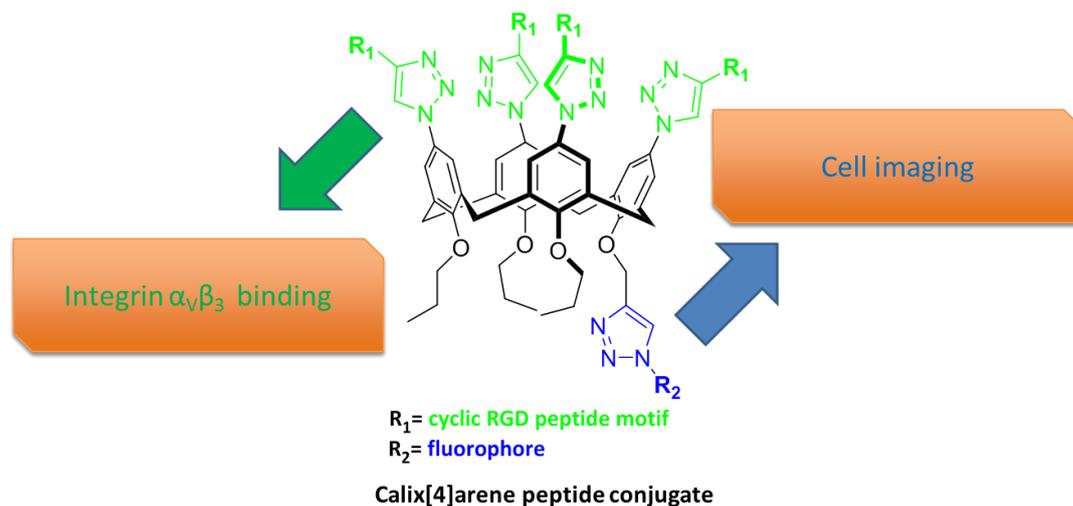
Figure 4.16   Target molecule for cRGD-conjugated calix[4]arene-based imaging ligand.

In order to make the target molecule, a retrosynthetic analysis of calixarene bearing cyclic RGD peptide and fluorophore was implemented (Scheme 4.1). As described in Section 3.3.2, boc-protected aminocalix[4]arene platforms had already been synthesised, it was decided to use the "clickable" calixarene **33** (see Section 3.3.2) as the platform from which to start. The lower rim of **33** can be functionalised with an azide bearing fluorophore. The amino group on the upper rim of the resultant fluorescent calixarene can be deprotected with acid and then converted into diazonium under acidic conditions. The diazonium can be easily replaced by azide *via* electrophilic aromatic substitution. Then four cyclic RGD peptide molecules modified with a propargyl group can then be installed on the upper rim of the calixarene.
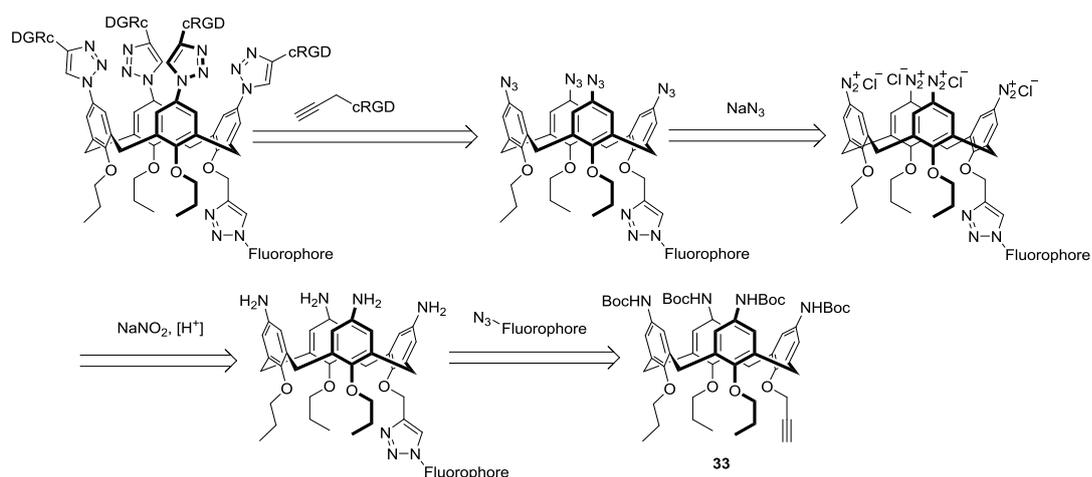
*Scheme 4.1   Retro-synthetic analysis of target molecule to previously synthesised compound **33** (Section 3.3.2).*

Thus the whole project was divided into several small projects as demonstrated below. (Figure 4.17)
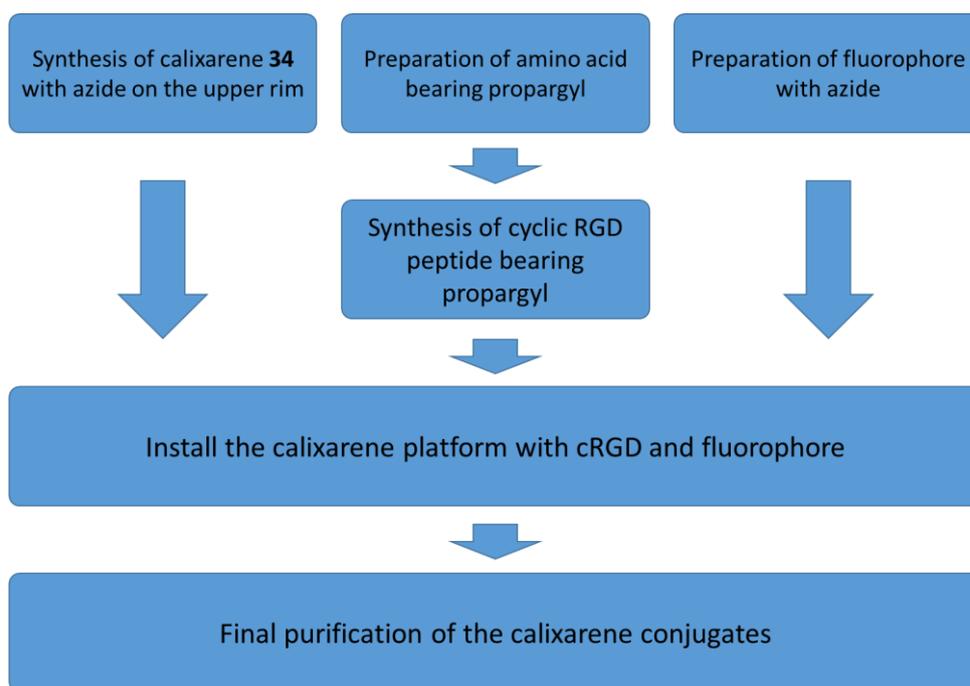
Figure 4.17 Project flow chart for planning.

4.3.2   RGD peptide design

RGD peptides have a very unique structure-activity relationship. Using compound

61 (cilengitide, Section 4.1.2) as an example, the argnine, glycine and aspartate were crucial to the activity of the peptide (Figure 4.18).[269] D-phenylalanine on the ligands can be replaced with D-tyrosine.[318,319]

Cilengitide(c(RGDf[N-Me]V))
J. Med. Chem., **1999**, 42, 3033-3040
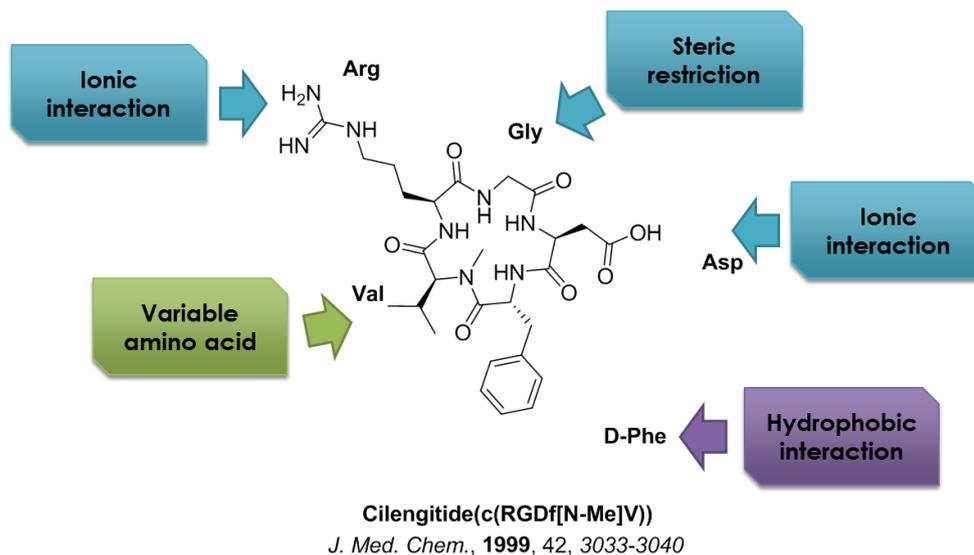
*Figure 4.18 Structure-activity relationship of compound **61***

The only variable amino acid residue is the *N*-methylated valine and the majority of chemical alternation have been made at this position. Lysine is usually chosen as the fifth variable amino acid due to its easy functionalisation on the ε-amino group (compound **80**, Figure 4.19). It can be conjugated with other substituents using peptide coupling itself or through 'click' reactions after being modified to the azide (compound **81**, Figure 4.20).[283,302,320]

80
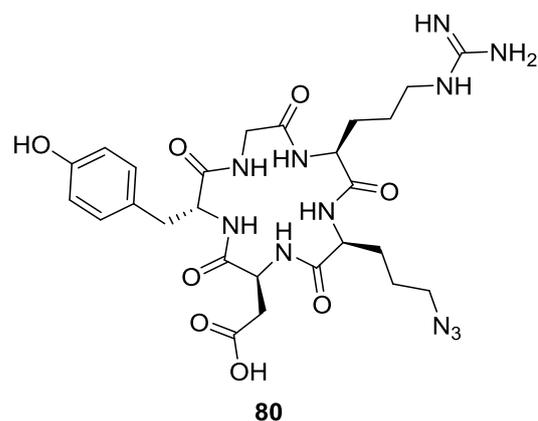
*Figure 4.19   cliclkable azide-c(RGDyK)**80** used in the study.*

81
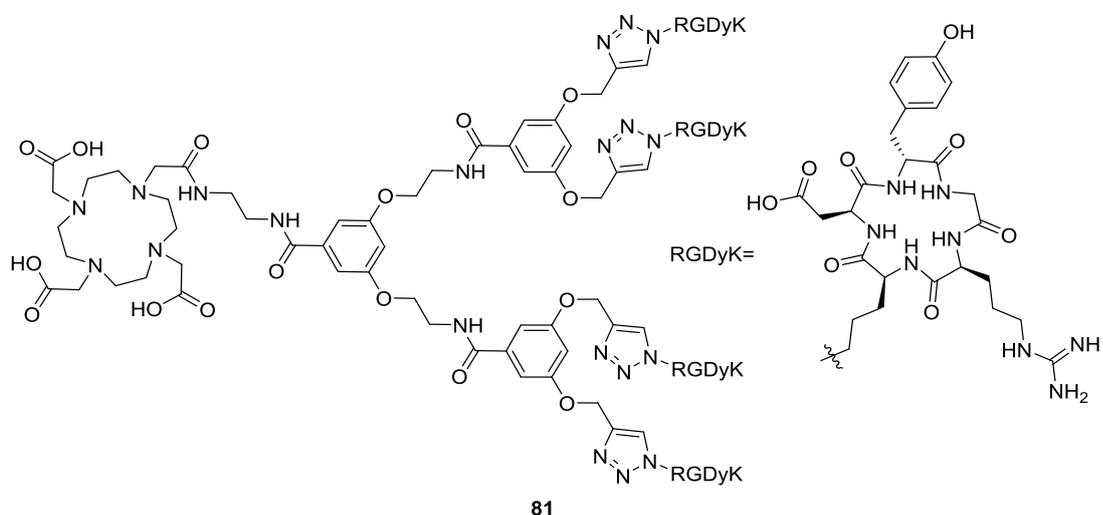
Figure 4.20   Chemical structure of DOTA-conjugated multivalent cyclic-RGD

*dendrimer **81** for tumour imagining purpose*

However, for this project an amino acid capable of being functionalised as an alkyne was required. Thus, serine was chosen as the fifth variable amino residue to replace lysine (compound **83**, Figure 4.21). The binding affinity of **cRGDfS** towards integrin $\alpha_V\beta_3$ was still in low nanomolar range ($^{\alpha V\beta 3}IC_{50}$ = 12 nM).[282]

In addition, the hydroxyl group on the serine need to be propargylated to perform the click reactions required when making the conjugates. This could also be the first example for a propargyl group attached cyclic RGD peptide, possessing many different applications aside the purpose intended for this project.
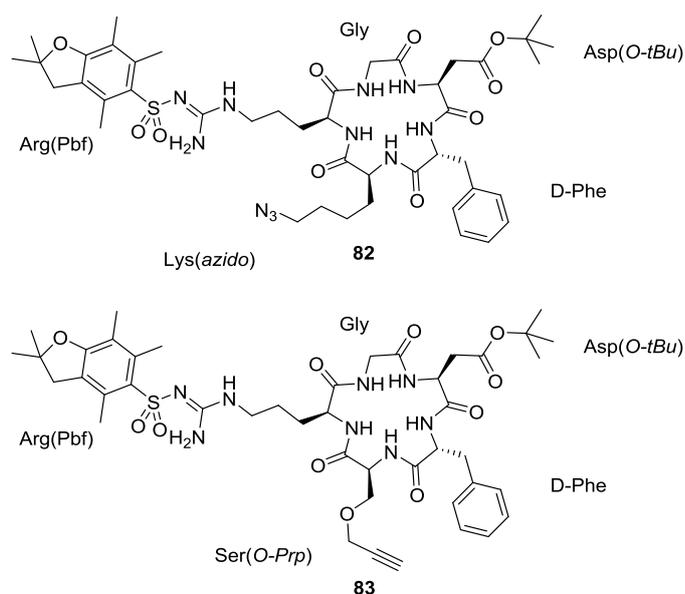
Figure 4.21   *A comparison of protected form of cyclic RGFfS(O-Prp)* **83** *&*
              reported clickable cyclic RGD peptide **82**. [255,273]

The target cyclic pentapeptide was designed to be stable under basic conditions and
for protecting group cleavage on the side chain to be performed with acid. In order
to make this peptide, retrosynthetic analysis was performed. As shown in Scheme
4.2, the Fmoc protected forms of four amino acids are commercially available. Only
the propargylated serine need to be specially prepared before the synthesis of the
peptide.

83
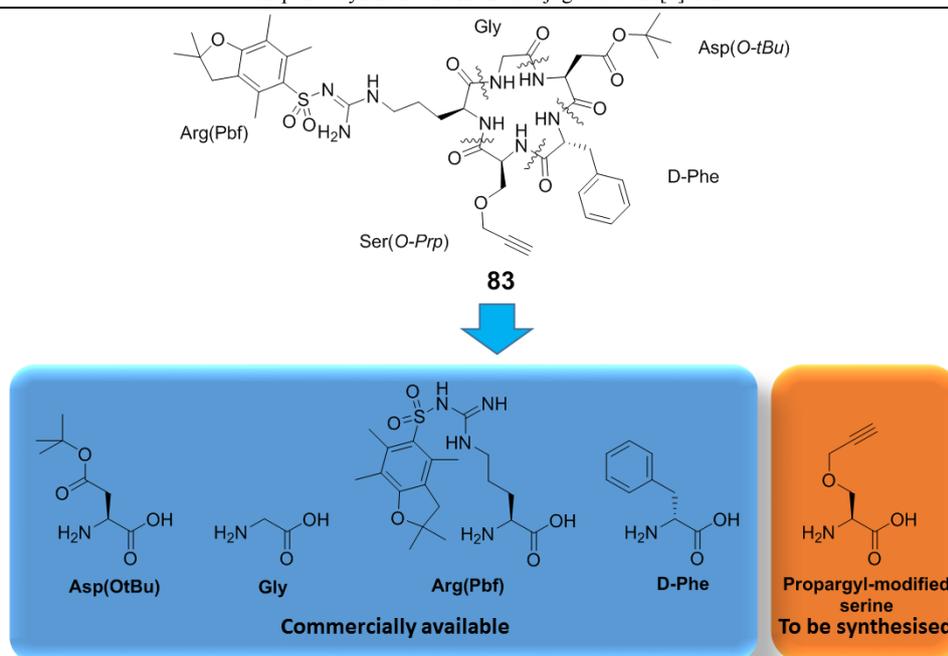
*Scheme 4.2   A retrosynthesis view on proposed cyclicpeptide **83** which would be able to be 'click' onto any suitable azide-functionalised system.*

The preparation of cyclic pentapeptide **83** can be performed *via* liquid phase peptide synthesis strategy or solid phase peptide synthesis (**SPPS**). In this project, **SPPS** was chosen due to its easy handling, relatively fast peptide propagation speed and easier purification.

4.3.3   Preparation of Fmoc-L-Serine (*O*-Propargyl)-OH

O-Propargyl modification of the serine was conducted by alkylating the hydroxyl group of Boc-protected serine **84** (Scheme 4.3).[321] The pK_a value of the hydrogen on the hydroxyl group in **84** is around 17 to 18. Thus in these examples, excess sodium hydride was used as a strong base. Deprotonation of the hydroxyl was carried out at 0°C
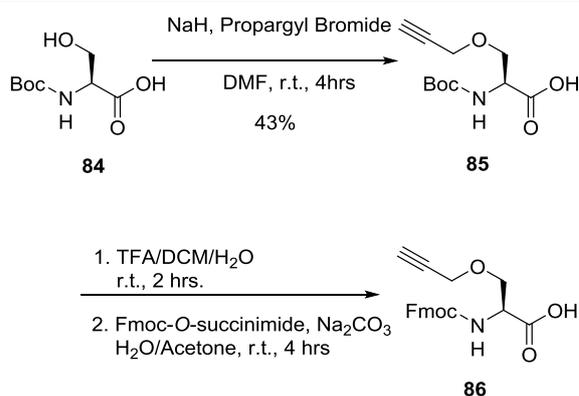
Scheme 4.3 *Preparation of amino acid **86** for SPPS*

When the carboxylic acid is deprotonated ($pK_a \approx 4$), it forms a conjugated system with the carbonyl group (Scheme 4.4) which helps to protect the α proton adjacent to the carbonyl in a strong alkaline conditions.
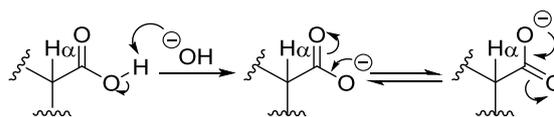
Scheme 4.4 *The conjugation on the carboxylic acid helps to protect α proton on   the amino acid.*

As described in Section 3.3.1, propargyl bromide was used as bromine is a good leaving group. The reaction mixture was monitored *via* thin layer chromatography (TLC), using a ninhydrin stain. However, product **85** after column chromatography still contained a substaintial amount of acetic acid, which was removed by forming an azeotrope with toluene and evaporating several times. The product **85** was confirmed by ^{1}H-NMR spectroscopy (Figure 4.22), where the characteristic terminal proton peak on the triple bond, which exists as a triplet, was found at 2.46 ppm. The obtained spectroscopic data for **85** was in agreement with literature.[321]
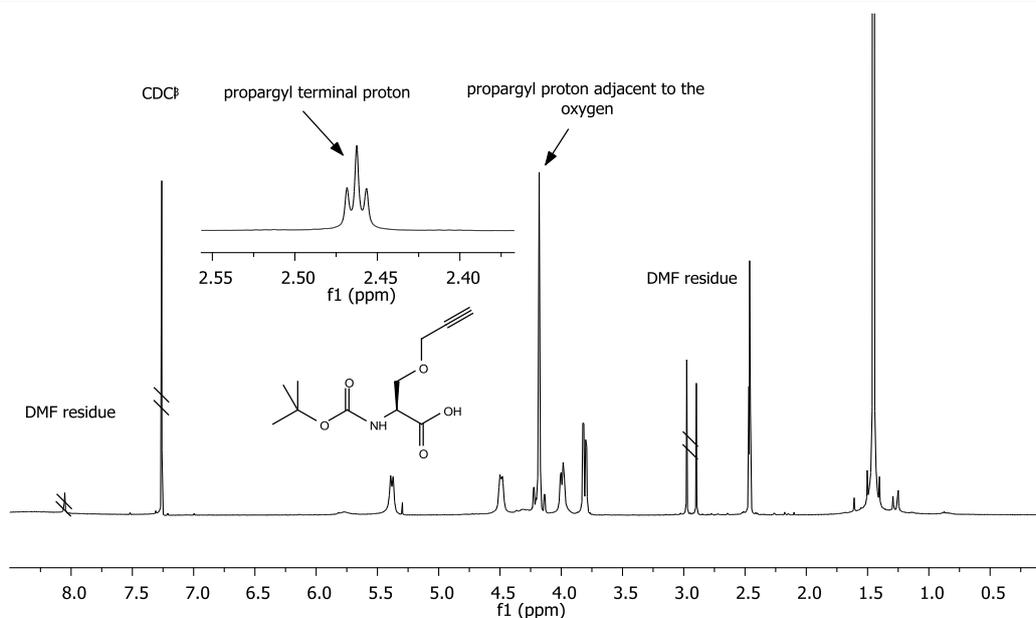
*Figure 4.22   Confirmation of propargyl installation on amino acid **85** via*
*   ¹H NMR spectroscopy*

The next step was to remove the boc protecting group using trifluoroacetic acid (TFA) and then to install an Fmoc protecting group to prepare amino acid **86** which is suitable for automated synthesis.[322] Fmoc-*N*-hydroxysuccinimide ester was used as the Fmoc-donating reagent in this reaction. By deprotonation of the amine on the serine with a mild base, the Fmoc protecting group can be easily attached to the amine end of the amino acid. The product **86** was purified *via* column chromatography and yielded a white solid as the pure product (61% yield). The product **86** was confirmed with using NMR spectroscopy, with a distinctive doublet at 5.65 ppm in deuterated chloroform representing the amide proton. The ¹H-NMR spectrum data for **86** was in consistent with literature data (Figure 4.23).[322] In the COSY spectrum of **86**, the terminal proton on the propargyl group (H_b) was confirmed as the triplex at 2.44 ppm, which was correlated with (dash-line) the singlet at 4.17 ppm. The α proton on the amino acid was annotated as the the doublet at 4.54 ppm (H_c). Which is correlated to (solid line) another three proton signals at 5.68 ppm (H_e), 4.02 ppm (H_d) and 3.81 ppm (H_d) respectively. In the HSQC spectrum, the proton on the amide (H_e) was expected to lose its signal. The α proton

on the amino acid was expected to be the only proton attached to a carbon while the two protons attached the propargyl group (H$_a$) was shown to attached to the same carbon. The HSQC spectrum further confirmed the assignment of protons on the COSY spectrum.
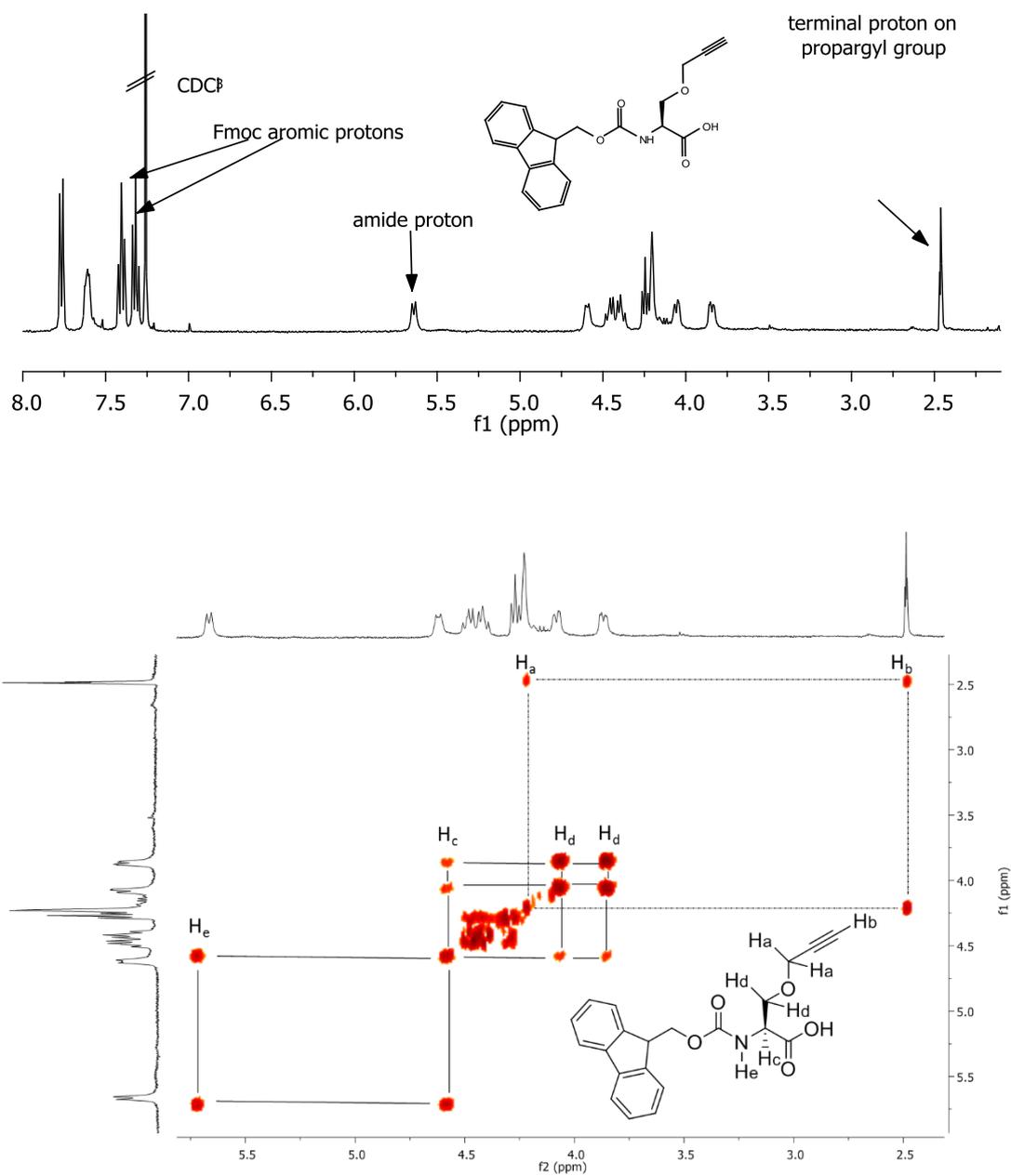
Figure 4.23 Confirmation of 86 via ^{1}H NMR and COSY spectrum
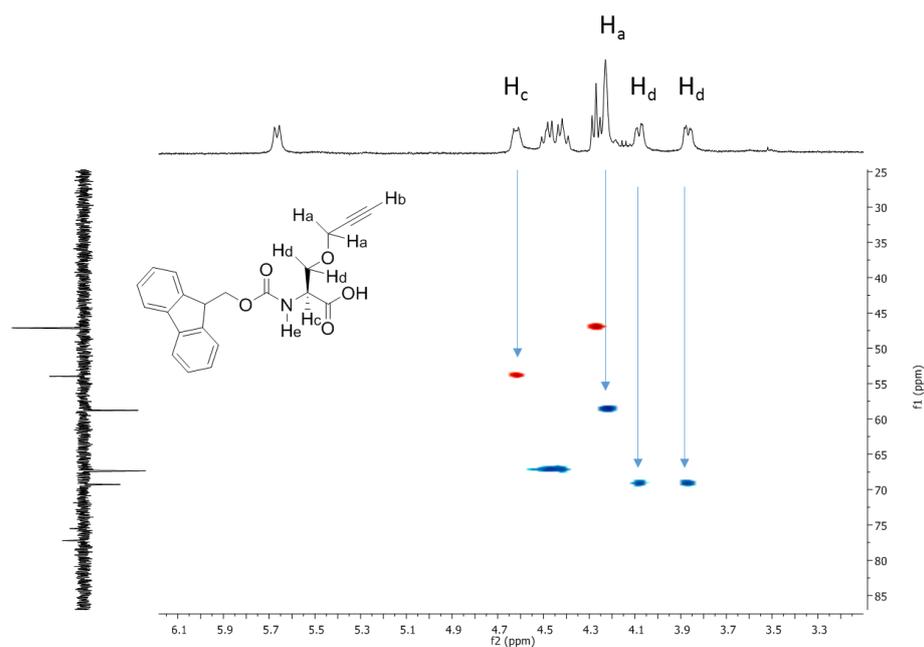
Figure 4.24   *HSQC spectrum of* **86**

4.3.4   SPPS on cyclic RGDfS(*O-Prp*)

There are two main synthetic strategies used in SPPS, the Fmoc (fluoren-9-ylmethloxycarbonyl) strategy and the Boc (*tert*-butoxycarbonyl) strategy, named after the protecting groups on amino acid N-terminal amino group.[323] In this case the synthesis of linear peptide **87** was carried out *via* an Fmoc SPPS strategy (Scheme 4.5) analogous to a previously published synthesis of a cyclic RGD peptide.[323]
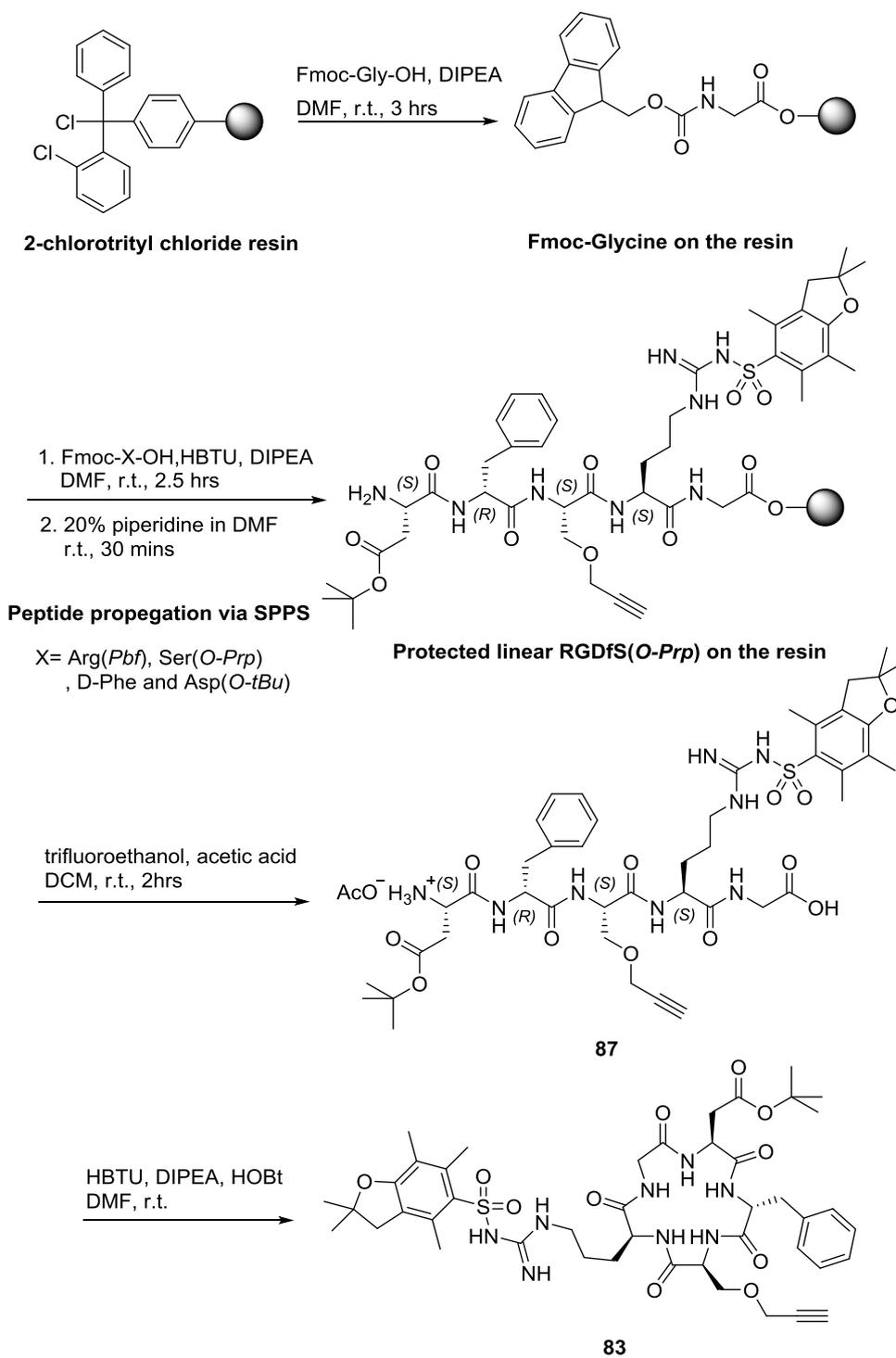
Scheme 4.5   Proposed SPPS-based synthesis routine on cyclic peptide 83

In order to cyclise the peptide, 2-chlorotrityl choride resin was chosen as the immobilized phase in the synthesis due to its ability to provide a free carboxylic acid on the C terminal on the peptide after cleavage (Scheme 4.5).[324-326]

Fmoc-glycine was the first amino acid to be chosen to attach to the resin, which means the further macrocyclisation was performed between glycine and aspartate. The reason to do this is because the free rotation of amino group on the glycine would facilitate the final cyclisation in the future whilst other more restricted side-chain bearing amino acids do not have this advantage.[325]

2-Chlorotritylchloride resin was highly sensitive towards moisture and water. Thus loading of the first amino acid, Fmoc-glycine, on the resin needs to be done under Argon and using dry solvent.[325] After properly drying the resin, UV spectroscopy was used to assess the loading efficiency of Fmoc-glycine (see Section 6.4.2). The quantity of other amino acids used in the later synthesis was calculated based on UV-absorption.[324]

In the next step, 20% piperidine in DMF was used as a nucleophilic base to deprotect the Fmoc-group on the glycine due to its alkalinity (pK_a = 11). The byproduct of this step can be easily washed away with DMF, leaving immobilised glycine with a free amine group on the resin. The Kaiser test[327,328] was used as the indicator to indicate the success of Fmoc removal (see Section 6.4.2).

The Kaiser test was a colorimetric test for the presence of free amine group.[327,328] The Kaiser test requires heating a mixture of resin with the peptide-loaded resin. Incompleted peptide coupling procedure or deprotection on Fmoc-protecting group is normally demonstrated as the development of a blue color (a negative Kaiser test), while a finished peptide coupling procedure was usually illustrated by the naturally yellow color of ninhydrin (positive Kaiser test). The mechanisms of the colour change of ninhydrin stain and Kaiser Test are demonstrated in Scheme 4.6.[323]
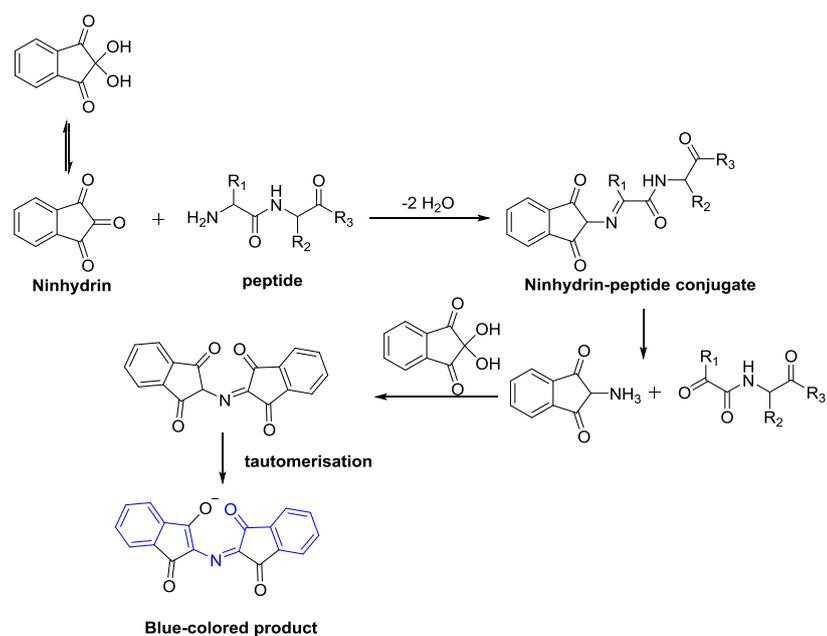
Scheme 4.6   Mechanism of the Kaiser test.[323]

In the next coupling step, HBTU,[329-331] in the presence of DIPEA, was used as the coupling reagents to form the peptide bond between Fmoc-Arg(Pbf)-OH and glycine (Scheme 4.7). Because of the existing conjugated system in a carboxylic acid, synthesis of the peptide bond requires the acid to first be activated.[332] Coupling reagents, such as DCC,[333] EDCI,[334] PyBOP[335] and HBTU,[331] can be to activate the C-terminal carboxylic acid or activated esters or acid halides.[332] HBTU was chosen because of its safety, its good solubility in classical solvent and fast reaction rate. It was also cheaper than HATU, the reagent previously described in the synthesis of the cyclic-RGD peptide.[329]
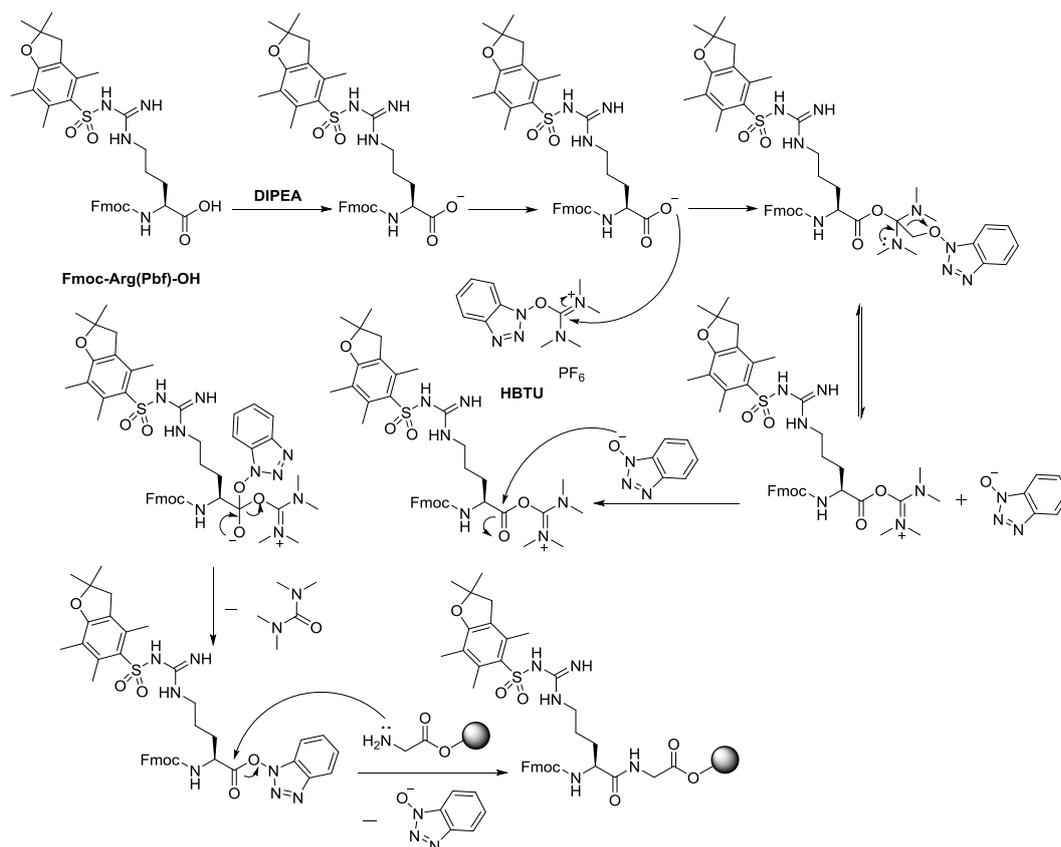
Scheme 4.7   Reaction mechanism in HBTU assisted peptide coupling

The success of the amino acid coupling can also be monitored by Kaiser test (Section 6.4.2) as successful coupling will result in no color change in the Kaiser solution containing the resin.

The next step was to deprotect the Fmoc protecting group on the N-terminal amine group. This was achieved by washing the resin with 20% piperidine in DMF several times, as previously described.[325] Subsequently, Fmoc-Ser(*O*-Prp)-OH, Fmoc-D-Phe-OH and Fmoc-Asp(*O*-tBu)-OH was added on to the peptide chain after the removal of Fmoc group on the previous amino acid was confirmed with the Kaiser test (see Section 6.4.2).

After the final removal of Fmoc protecting group on the aspartate, the linear penta-peptide **87** was cleaved from the resin with trifluoroethanol and acetic acid.[324] This mixture results in a cleavage of the peptide from the resin without affecting the acid

labile protecting group on the aspartate and arginine. After using toluene to azeotrope off the excess acetic acid, the crude product **87** was purified with preparative-HPLC to give a yellow-colored powder (38% yield). The structure of **87** was assigned with ^{1}H-NMR spectroscopy (Figure 4.25) with the propargyl side chain being clearly identified by the terminal proton peak seen at end proton 2.90 ppm. The methyl groups of Pbf-protecting group on the arginine side chain can be found as several singlets between 3.0 to 1.5 ppm. The aromatic protons on D-phenylalanine can be identified as multiplets at 7.3 ppm. Additional high-resolution mass spectroscopy (HRMS) was obtained as well (Figure 4.26). The ionised form of **87** was found at 949.4098 in m/z (Calculated as $[C_{44}H_{63}N_8O_{12}S]^+$).
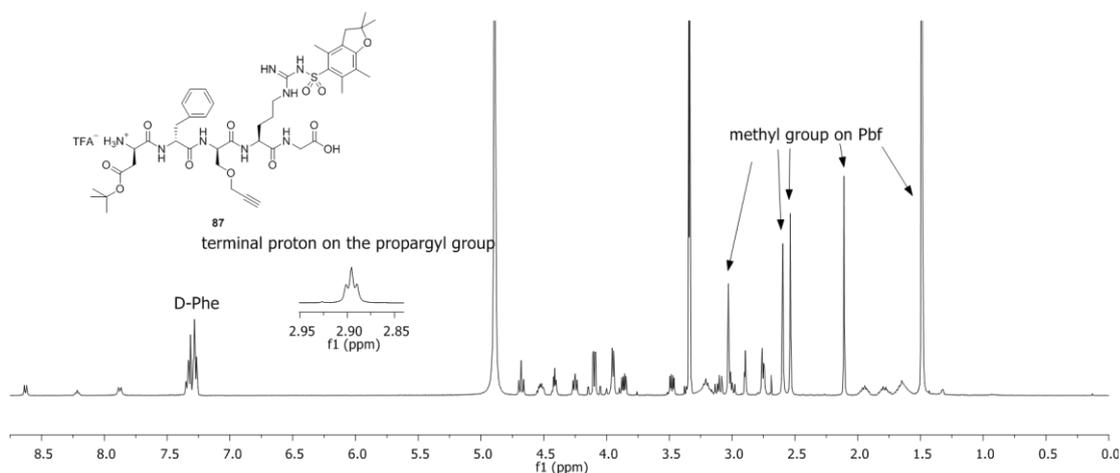
Figure 4.25 ^{1}H NMR spectrum of linear peptide **87**
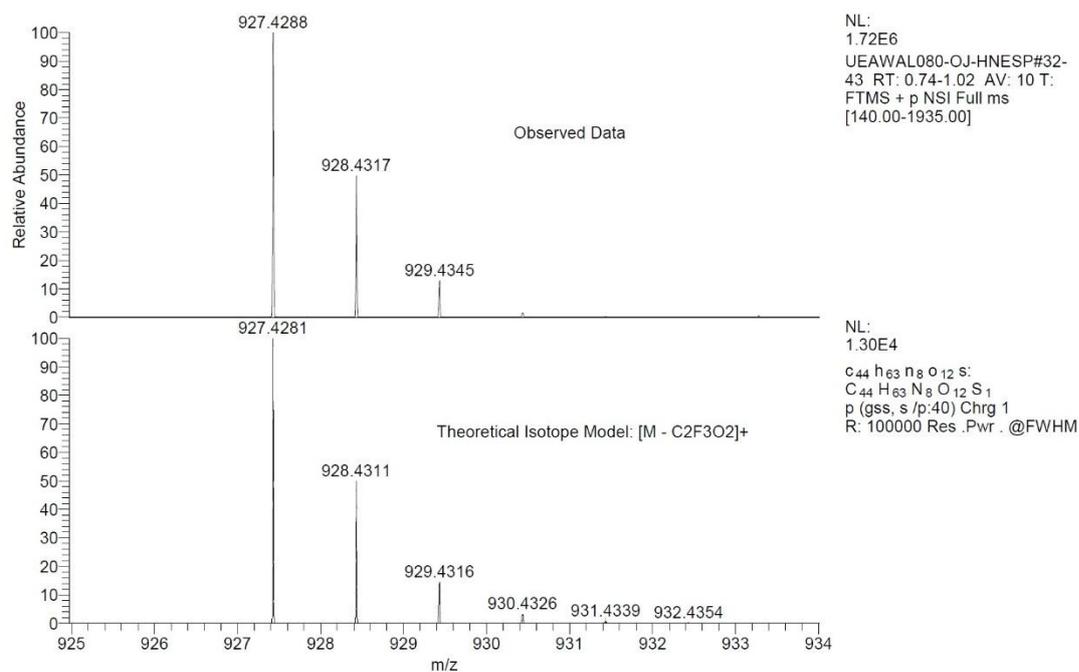
*Figure 4.26   HRMS spectrum of **87***

4.3.5   Cyclisation of the linear peptide

Head-to-tail cyclisation of linear peptide **87** was achieved with HBTU in acetonitrile under very dilute conditions ([linear peptide] $\leq$ 1mmol/L).[325] The reaction was carried out under room temperature for 120 hours with several additions of extra HBTU reagent into the reaction system. The reaction was monitored with Kaiser test to ensure the starting material was depleted as the cyclised peptide should give a negative result in Kaiser test. After the work-up, cyclised, side-chain-protected pentapeptide **83** was purified with preparative HPLC to obtain a yellow-white-coloured solid. The side-chain protected peptide **83** was confirmed with analytical HPLC (98%) and HRMS ([M+Na]$^+$: 931.3982 in m/z) (Figure 4.27).

83
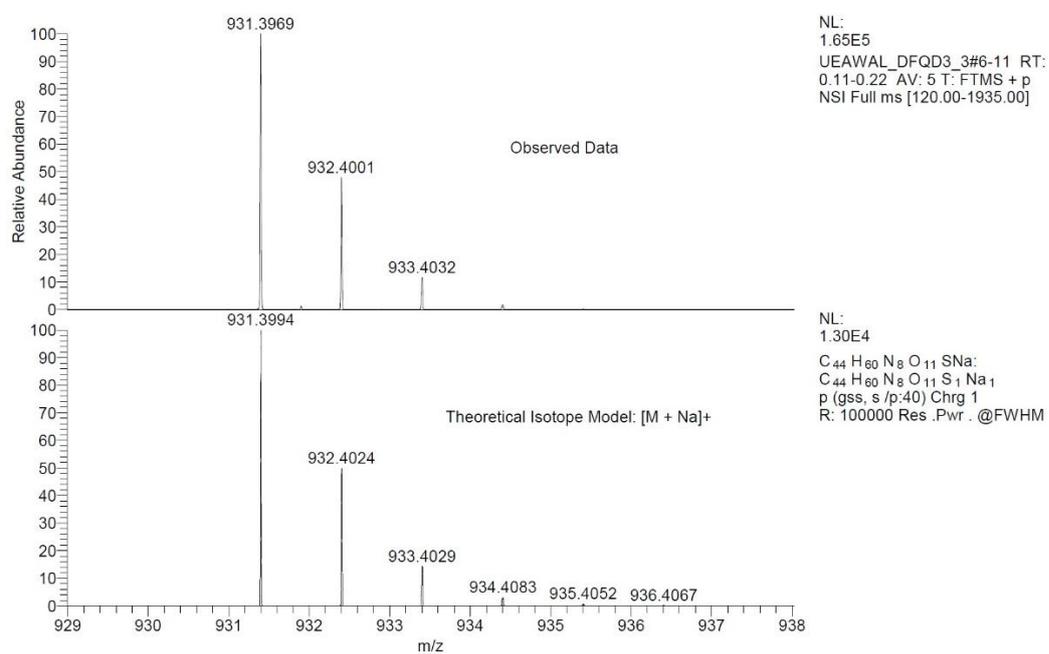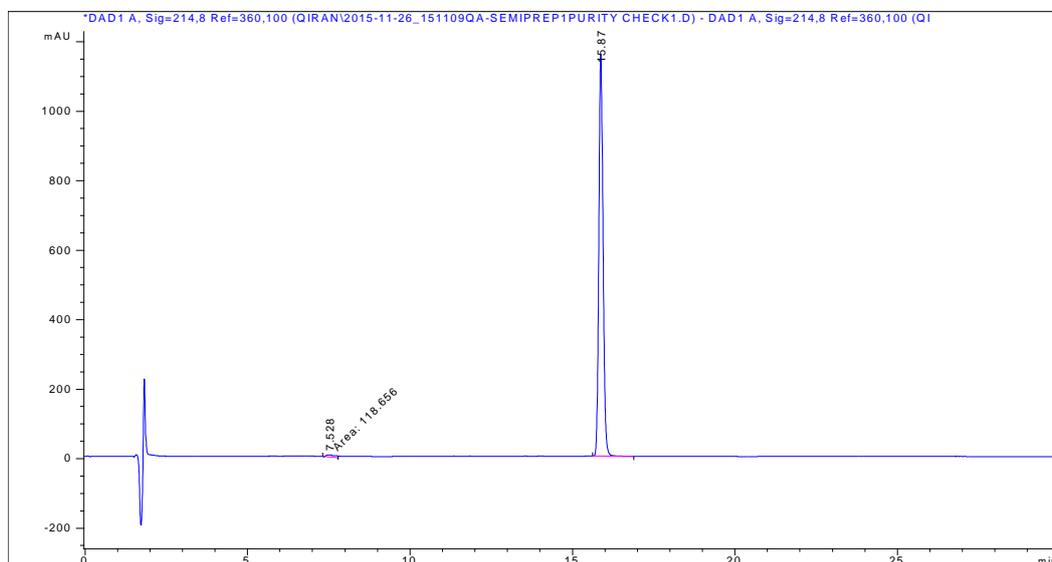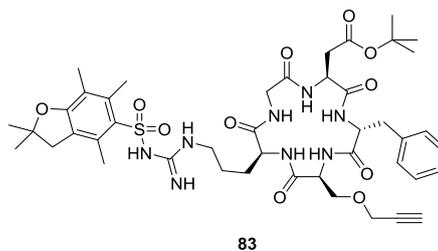
Figure 4.27   Analytical HPLC trace of purified and protetcted cyclic

RGDfS(O-Prp) **83** *and its HRMS spectrum*

4.3.5   Preparation of fluorescent reporter

In terms of the fluorophore, a coumarin motif was chosen due to its strong fluorosence emission. In this project, 3-azido-7-hydroxycoumarin **91** (Scheme 4.) was selected because it is not fluorescent when the azide group on the C$_7$ position is present. After the 'click' reaction, non-fluorescent 3-azido-7-hydroxycoumarin becomes fluorescent, with an emission at 500 nm.[336] This property can also be utilised as an indicator for the attachment of coumarin in the 'click reaction'.

Azide-bearing clickable coumarin was synthesised according to existing methods (Scheme 4.8).[336] By stirring 2,4-dihydroxy benzaldehyde (**88**), *N*-acetyl glycine (**89**) and sodium acetate in acetic anhydride at reflux overnight, 3-acetamido-7-acetoxycoumarin **90** as an intermediate was formed as yellow-colored solid after reaction workup. Then the intermediate was heated at reflux in 2:1 37% HCl/ethanol for 1 hour to deprotect the hydroxyl group and amine. Then the reaction mixture was cooled down to 0°C and sodium nitrite added to form diazonium salt form of the coumarin. The diazonium was then replaced with azide in the solution to form 3-azido-7-hydroxy coumarin **91**. The structure of coumarin **91** was consistent with the existing ^{1}H NMR spectrum in the literature.[336]
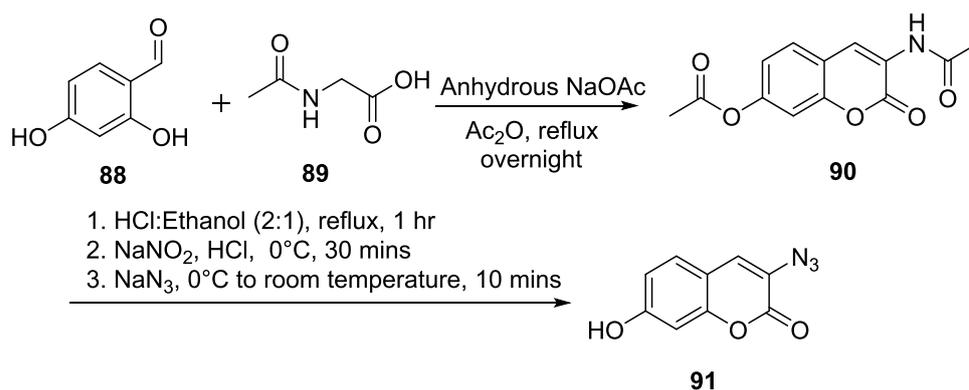
*Scheme 4.8   Preparation of coumarin **91** for click reaction*

4.3.6 Preparing the calixarene-dye conjugate

Two different calixarenes were developed featuring the coumarin dye on the lower-rim. In the first case, azido functionality was introduced directly to the upper-rim of the calixarene **92**, in the second, a spacer was introduced between the upper-rim and calixarene core to form **93**. These two strategies (Scheme 4.9) enabled an investigation of whether four cyclic peptides could be directly appended to the calixarene or whether steric effects would require a spacer to be present. Additionally, it would allow a comparison of the biological effects of more or less flexible calixarene derivatives.
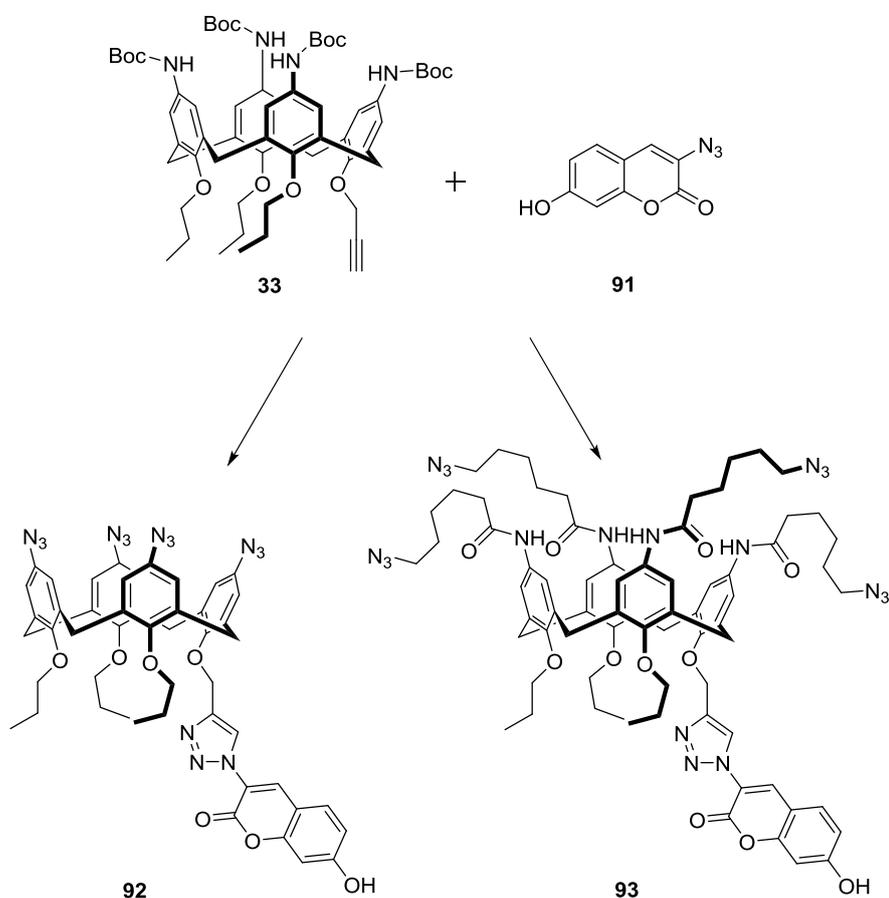
Scheme 4.9   Two proposed calix[4]arene-coumarin conjugates 92 and 93 for click chemistry

Preparation of 92

The azido-coumarin **91** was attached to calix[4]arene **33** using the 'click' reaction conditions described in Section 3.3.4 (Scheme 4.10).[197]

Scheme 4.10   Preparation of calixarene 94

After treating with copper sulphate and sodium ascorbate in DMF at 80°C overnight, the reaction mixture turned fluorescent under UV light, indicating that the 'click' reaction had occured. The reaction was terminated by evaporating the solvent *in vacuo*. And the brown-colored crude product was purified with column chromatography to afford a light-yellow-coloured solid (81.3% yield). The chemical structure of this calixarene-coumarin conjugate **94** was assigned with ^{1}H-NMR spectroscopy, which is in agreement with the literature (Figure 4.28).[197] Calixarene **94** was used as the starting material for **92** and **93** (Scheme 4.9).

Figure 4.28 ^{1}*H-NMR spectrum of Boc-amino calixarene-coumarin conjugate* ***94***

After deprotecting the amino group on the upper rim of calixarene **94** with gaseous HCl, the azide was introduced on the upper rim. In this step, sodium nitrite was used to form diazonium salt under 0°C, then replacing the diazonium with azide to form calixarene **92** as described in Section 4.3.3 (Scheme 4.11).[337]
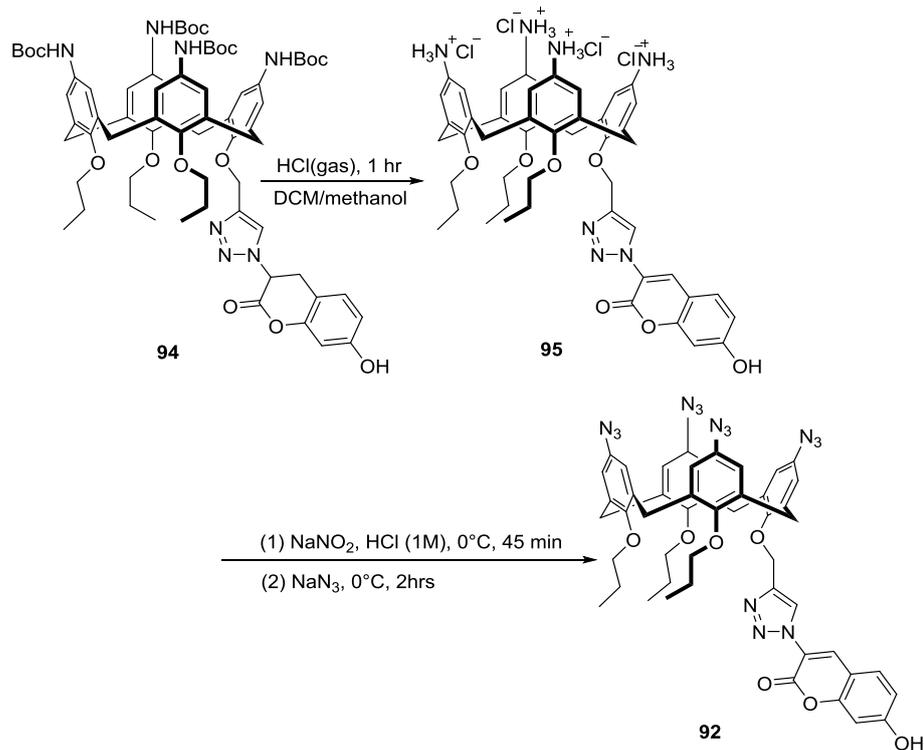
Scheme 4.11   Preparation of calixarene 92

The installation of azide group was confirmed using IR spectroscopy, which gives a distinct and strong transmittance peak at 2106 cm^{-1} (Figure 4.29). As can be seen from Figure 4.29, Boc-protected calixarene **94** does not possess the strong IR absorbance at 2106 cm^{-1} while azido-functionalised calixarene **92** does.
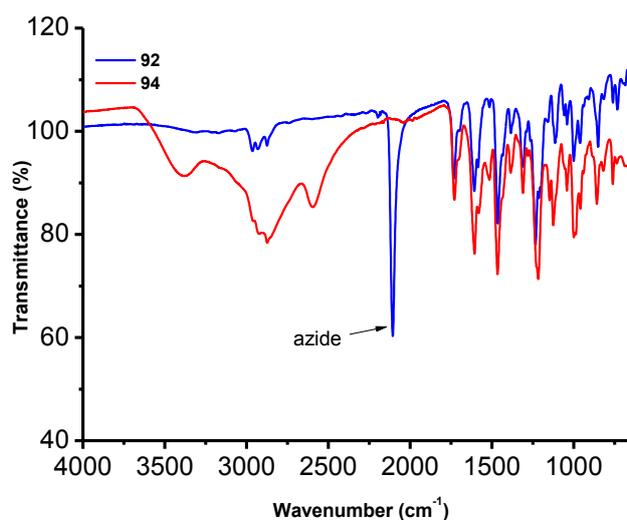
Figure 4.29   IR spectrum of calixarene 94 & calixarene 92

After purification by column chromatography, the product of this reaction was confirmed with ^{1}H-NMR spectroscopy and HRMS (Figure 4.30). This confirmed that the 4 amine groups on the upper rim of the calixarene **95** were all replaced with azide. As seen in Figure 4.30, the peaks from 4.25 to 4.50 ppm and 3.0 to 3.20 ppm are from the protons on the methylene bridges connecting adjacent aromatic rings,confirmeing a symmetrical calixarene structure (as described in Section 3.3.2 and 3.3.5) and installation of 4 azide groups on the upper rim of calixarene **92**. The completion of the reaction was also supported with HRMS, with a mass of 978.3515 for $[C_{49}H_{45}N_{15}O_7Na]^+$ in m/z.
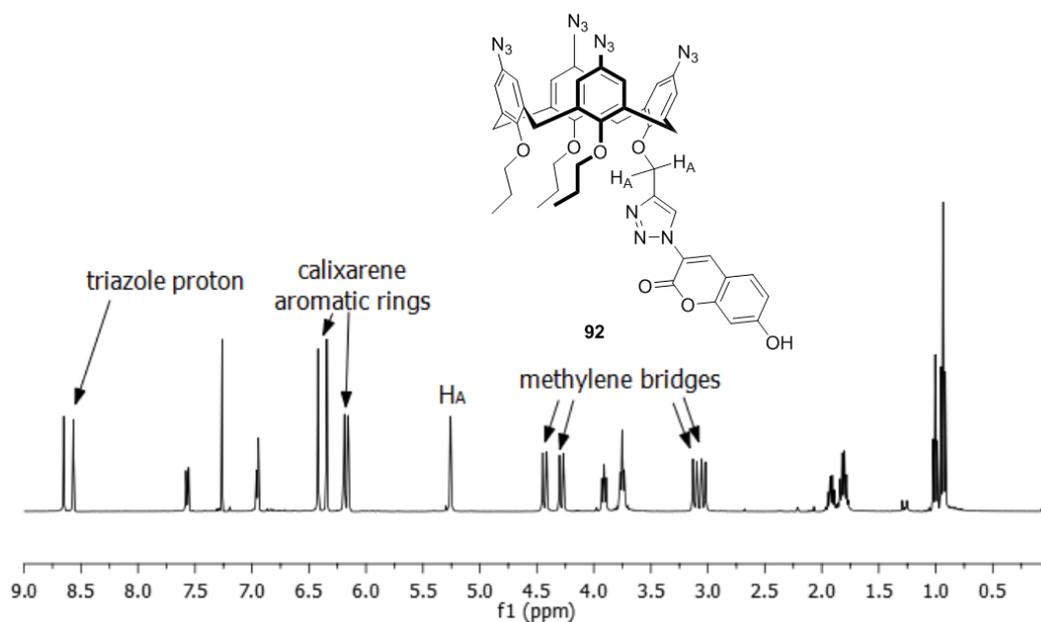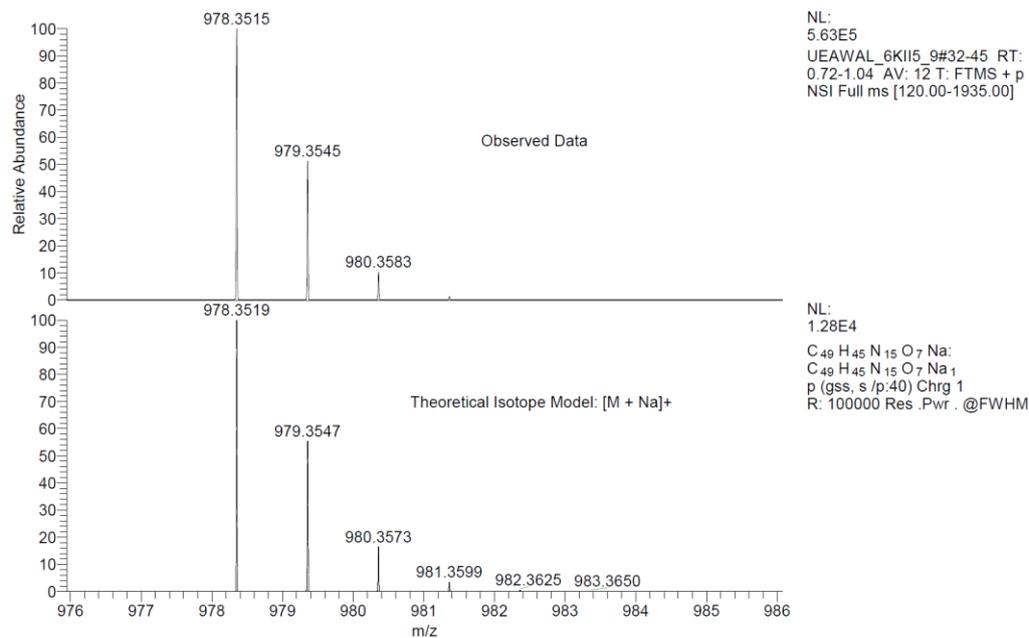
Figure 4.30 ¹*H-NMR spectrum and HRMS spectrum of calixarene* **92**

Preparation of 93

Deprotected calixarene **94** can also be functionalised with 6-azido hexanoic acid (**97**)[190] to create a flexible structure for clickable RGD peptide to bind (Scheme 4.12).
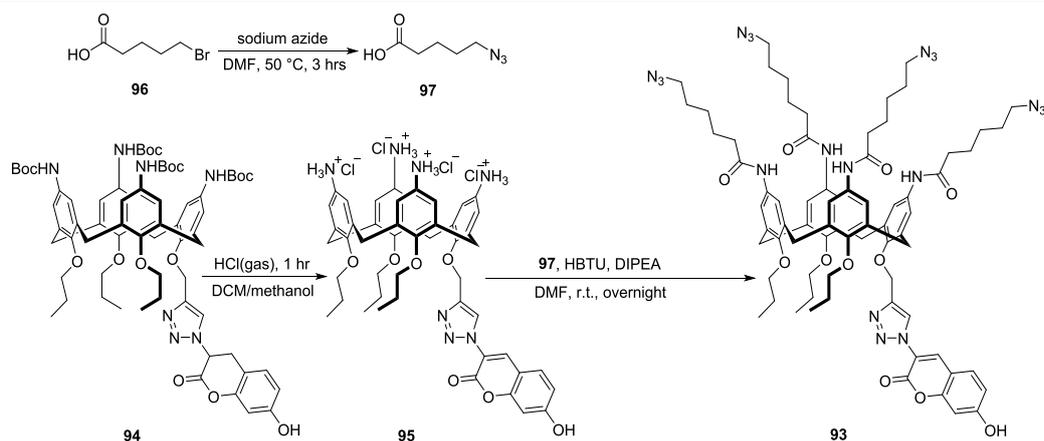
Scheme 4.12   *Synthesis routine for calixarene 93*

Linker **97** was obtained *via* stirring compound **96** with sodium azide in DMF at 50 °C for 3 hours (Scheme 4.12). The bromine was substituted with azide *via* an S_N-2 mechanism. After working up the reaction, the product was obtained as a orange-coloured oil. The introduction of the azide group was confirmed with IR spectroscopy (Figure 4.31) and its NMR spectrum was in accordance with literature data.[190]
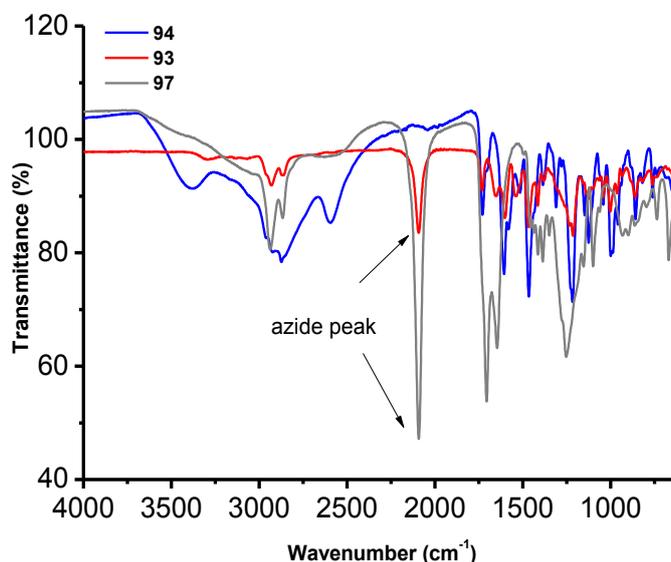
Figure 4.31   *IR spectrum of 6-azido hexanoic acid 93, 94 and linker 97*

Deprotected **94** was then coupled to 6-azido hexanoic acid in the presence of HBTU

and DIPEA as described in Section 4.3.4. After purification by column chromatography, the product was confirmed with IR (Figure 4.31) and ^{1}H NMR spectroscopy in deutrated MeOH (Figure 4.32). The protons pointing inwards on the methylene bridges are clearly spotted from 4.36 to 4.47 ppm as two doublets. This indicated four upper-rim transformations have been made. In the IR spectrum (Figure 4.31), the azide for **93** can be seen at 2100 cm^{-1}, indicating the success of installing azide functionalities on the calixarene platform. The outward-pointing protons on the methylene bridges in ^{1}H-NMR spectrum may overlapped with the solvent residue peaks.
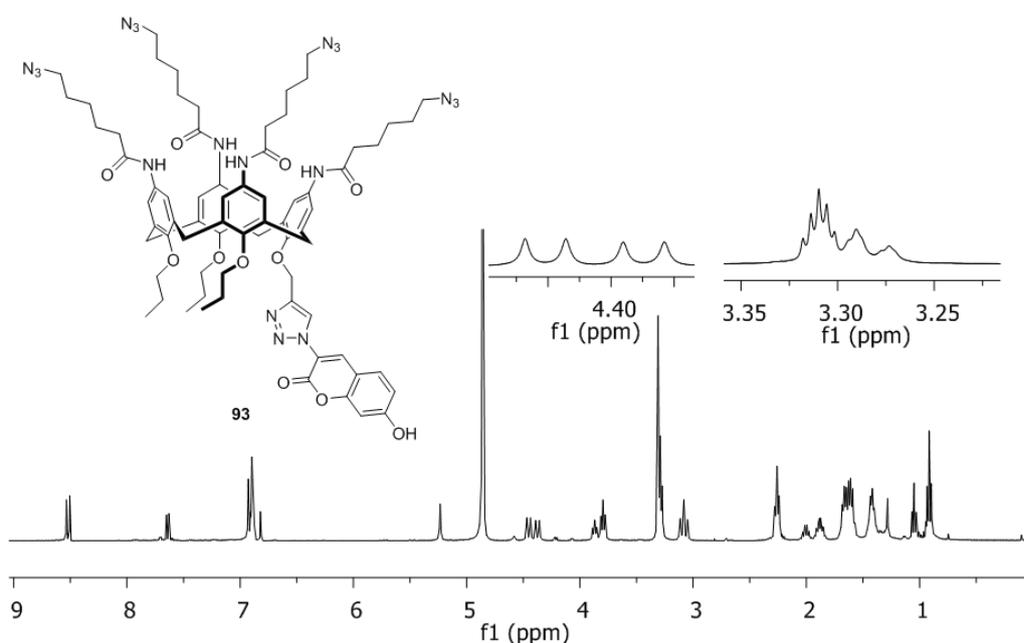
Figure 4.32   *^{1}H-NMR spectrum on **93** in deuterated methanol*

4.3.7   Model reaction: Coupling of cyclic RGDfS(*O-Prp*) to coumarin dye

In order to study whether propargyl-modification of the cyclic RGD peptide was compatible with undertaking a 'click' reaction, a model reaction linking cyclic peptide **83** directly to coumarin **91** first was undertaken (Scheme 4.13). The fluorescence of the coumarin will increase drastically after the 'click' reaction, which serves as a perfect indicator for the reaction progress. After deprotection, the

peptide-coumarin conjugate **99** itself can be used as a reference compound in later cell based experiments.

Unlike the previous click reactions, a chelated copper (I) source, tetrakis(acetonitrile) copper (I) hexafluorophosphate, was used in the click reaction with the peptide.[200,338-340] Using chelated copper (I) source was expected to reduce the possibility that the amide bond on the peptide might chelate the copper (I) cation and thus complicate the purification steps followed.
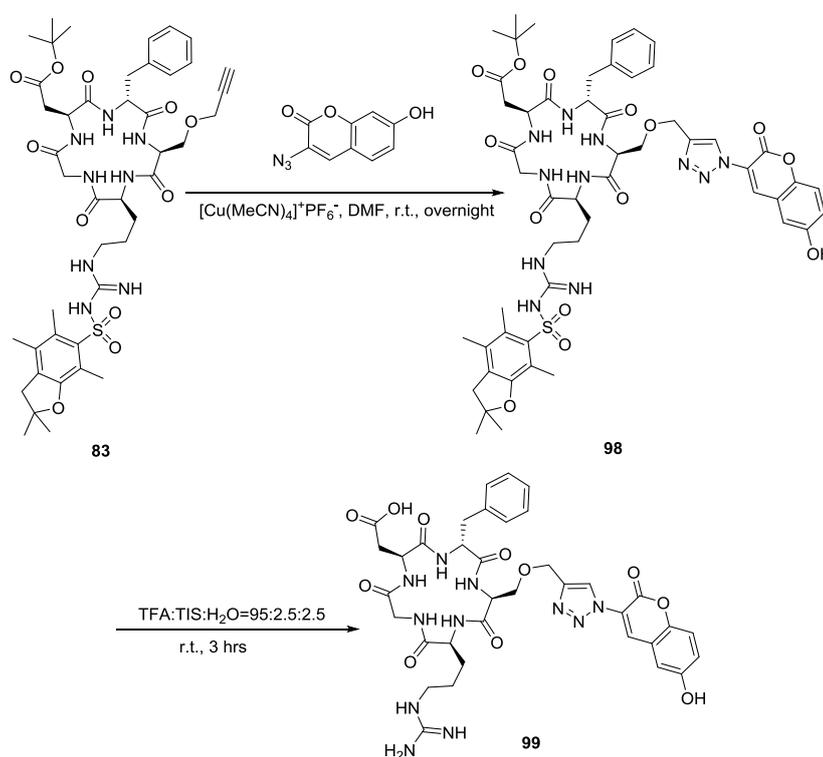
*Scheme 4.13   Reaction scheme of peptide-coumarin conjugates **98** and **99***

Peptide **83**, coumarin **91** and the chelated copper (I) source were stirred under argon in a sealed reaction system at room temperature. When DIPEA was added into the reaction system, the reaction mixture became fluorescent instantly under UV light, indicating the formation of the triazole group on the coumarin.

This confirmed that the protected cRGDfS(*O*-Prp) was suitable for the click

reaction, and can be applied to the two coumarin-appended calixarene platforms.

The resulting crude reaction mixture was deprotected with TFA, triisopropylsilane (TIS) and water. The presence of TFA helps to cleave the acid labile protecting groups on the side chain, such as the *tert*-buytl group aspartate on the and the Pbf group on the arginine.[323] Triisopropylsilane was widely used as a peptide cleavage scavengers in peptide chemistry, protecting the newly deprotected side chain functionalites against the strong acid.[341]

The reaction mixture was then dried on vaccum and ready for further purified by semi-preparative or preparative HPLC.

The product was confirmed using ^{1}H NMR spectroscopy, HSQC and HRMS (Figure 4.33 & Figure 4.34). When dissolving **99** in deuterated DMSO, the amide protons from the peptide structure can also be detected, including the amide proton on arginine side chain (Figure 4.33). The click proton can be assigned as a sharp singlet at 8.5 ppm. Four α protons (apart from the glycine) on the cyclic peptides can be assigned as four red dots shown in the HSQC spectrum (within the circule, Figure 4.33b). The intergration of protons on ^{1}H NMR spectrum matched the number of protons on **99**. Additional HRMS (Figure 4.34) further confirmed the structure of **99**, with a peak found at 804.3058 in m/z as the ionised form of **99** (Calculated as $[C_{36}H_{42}N_{11}O_{11}]^{+}$). Thus it was shown that the alkyene group on cyclic peptide **83** was able to perform CuAAC and the use of chelated copper is able to catalyse the CuAAC in the presence of cyclic RGD peptide.
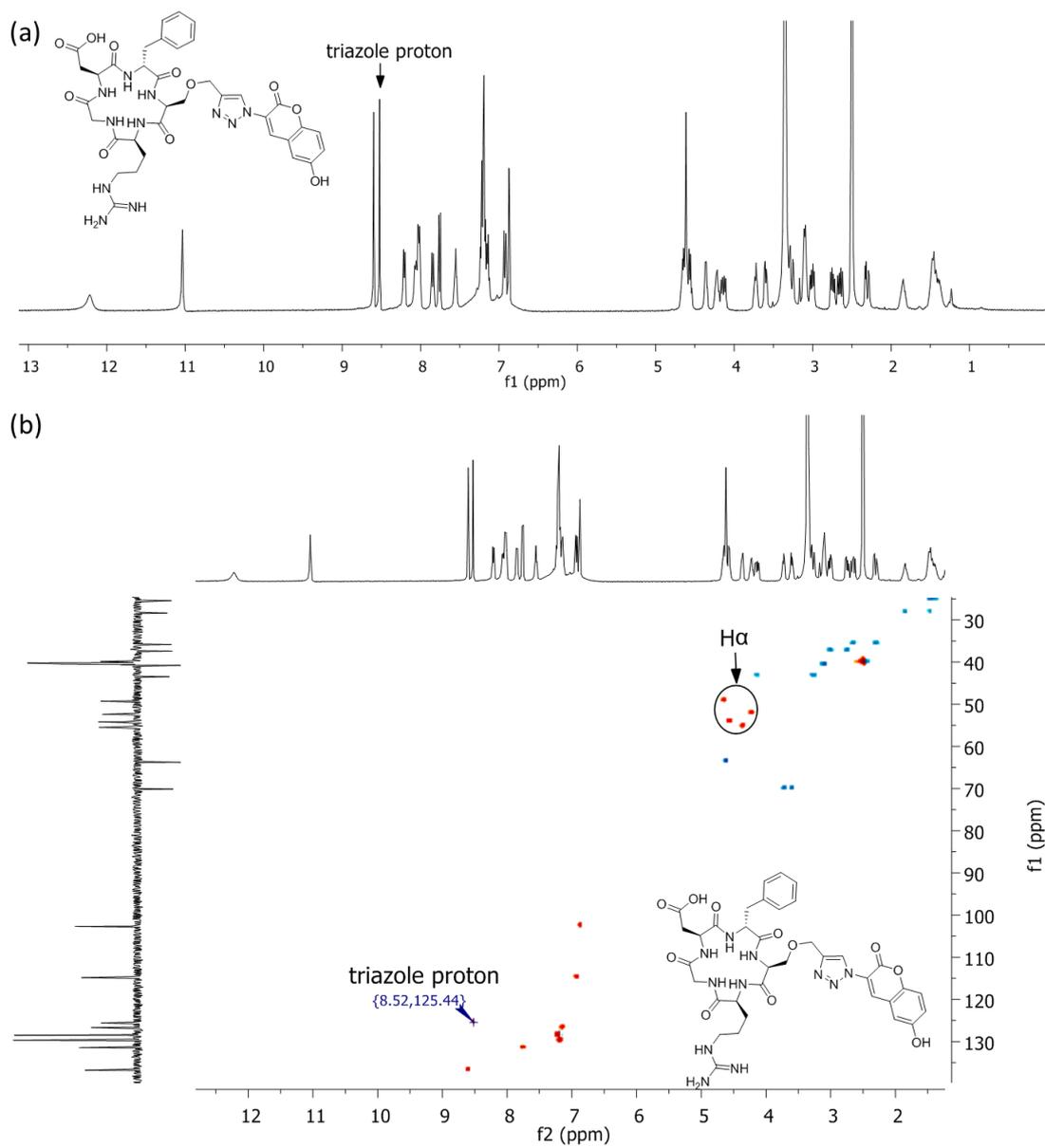
Figure 4.33   *^{1}H NMR and HSQC spectrum on **99** in deutrated-DMSO*
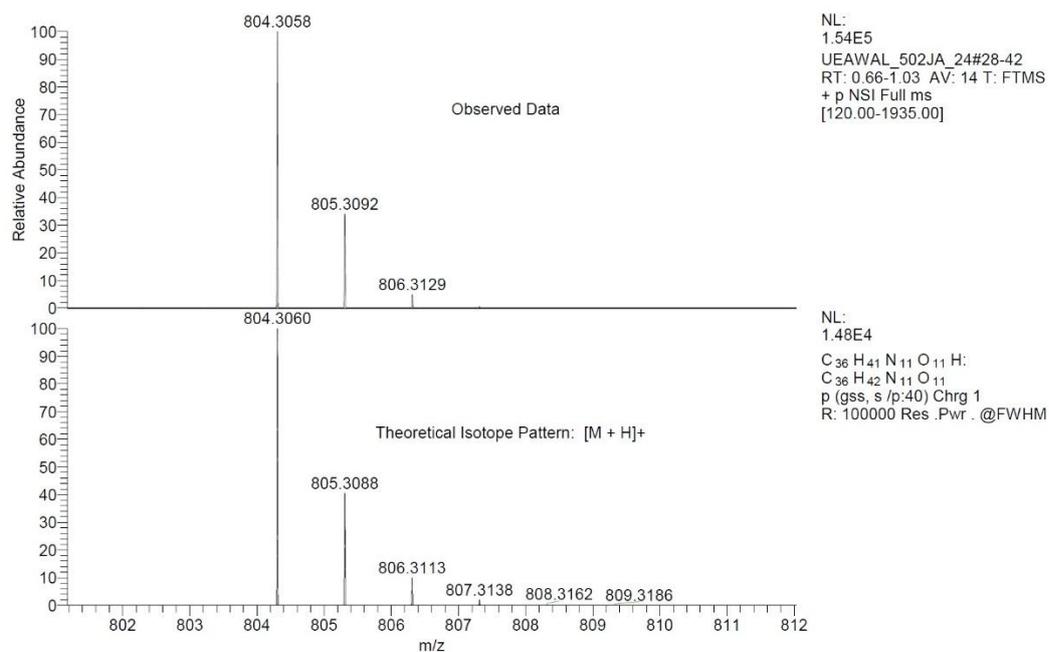
*Figure 4.34   HRMS spectrum for the purified **99***

4.3.8   Incorporating cyclic RGDfS(*O-Prp*) to the calix[4]arene-coumarin conjugate

Click reaction between calixarene 92 and peptide 83

The last and the most critical step in this project was to install cyclic RGD peptide to the fluorescent calixarene-coumarin conjugates, followed by final deprotection in TFA to prepare **100** and **101** (Figure 4.35).

100   **101**

R=
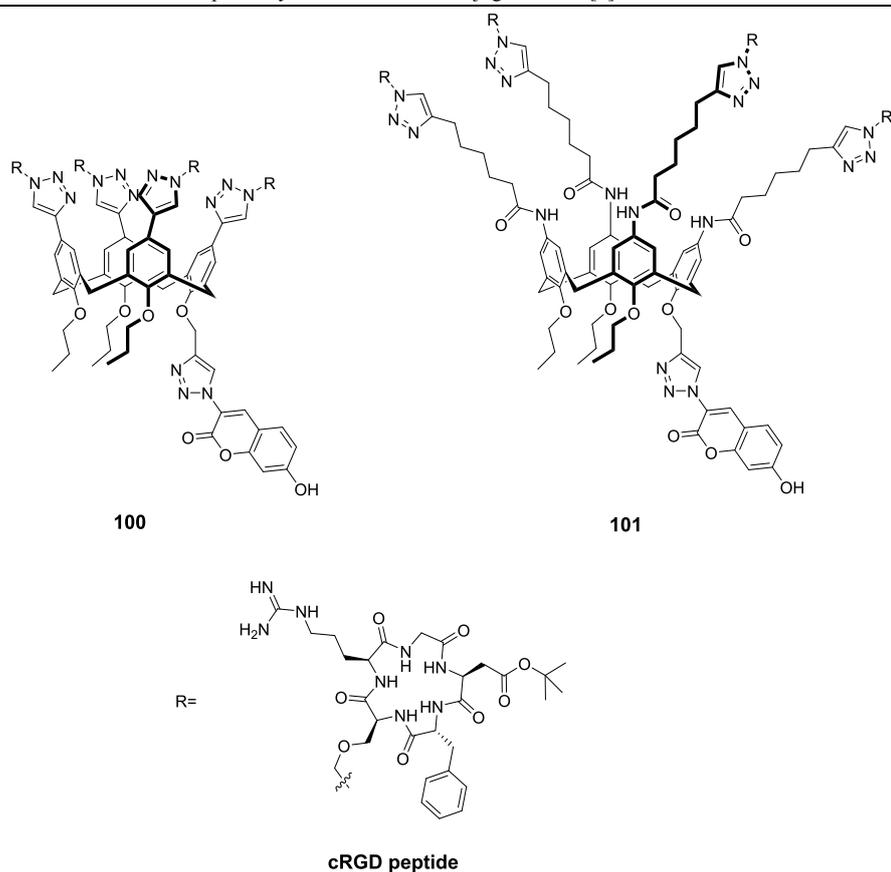
cRGD peptide

*Figure 4.35   Target molecules, **100** and **101**, for synthesis in this project*

In an attempt to make **100**, the fluorescent calixarene platform **92** and protectd peptide **83** was treated with CuSO$_4$·5H$_2$O and sodium ascorbate in DMF under argon at room temperature for four days to make protected conjugates **102** (Scheme 4.14). The reaction mixture was tested by TLC, which indicated that there were still a substantial amount of unreacted **92** left in the reaction mixture (Scheme 4.14). Heating the reaction mixture to 50°C for another 3 days or subsequently 80 °C for another 5 days did not result in any observable changes to the mixture by TLC. This might be due to the relatively bulky **83** not being able to installed on the crowded upper rim of calix[4]arene due to steric hindrance. Alternatively this may be due to the choice of copper (I) source.
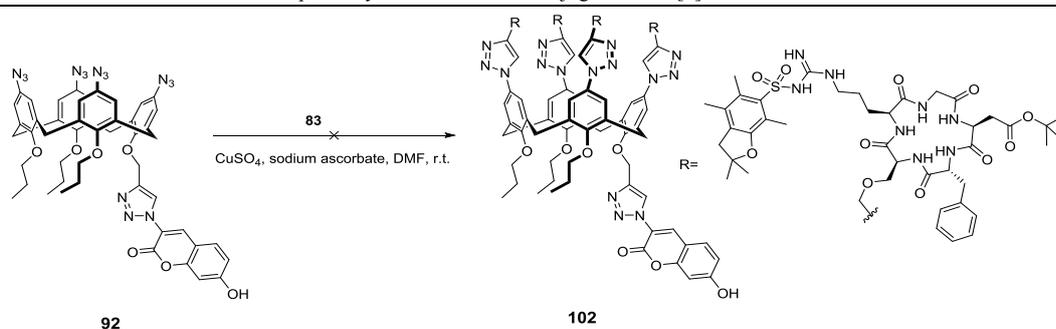
*Scheme 4.14   Click reaction between **92** and **85** did not work.*

Click reaction between calixarene 93 and peptide 83

In contrast treatment of **93** with **81** in DMF in the presence of tetrakis (acetonitrile) copper (I) hexafluorophosphate was more succesful (Scheme 4.15), yielding a green colored crystalline crude product after evapourating the solvent.
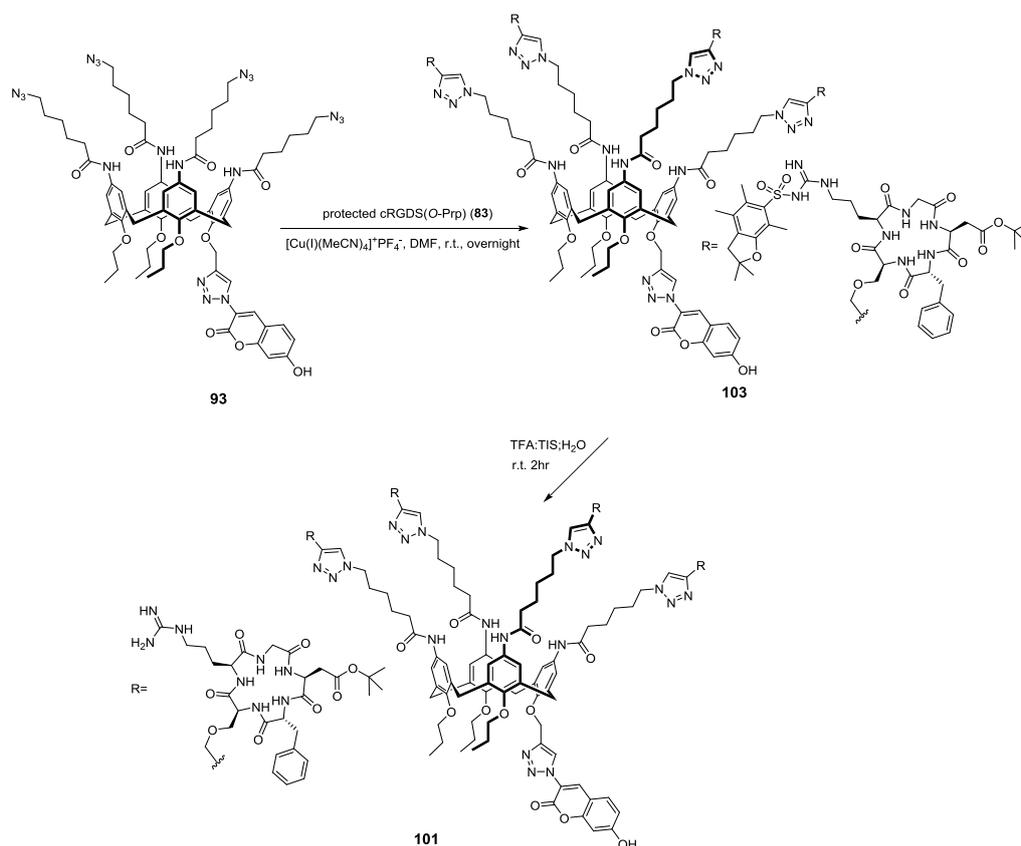
*Scheme 4.15   Click reaction between **93** and peptide **83***

The crude product was deprotected using TFA and TIS and analysed by mass

spectrometry (Figure 4.36).

The HRMS Spectrum of the crude material showed that the starting material for **93** was not detected. Deprotected clickable cRGD peptide was detected as the largest peak ($[M+H]^+$) at 601.2721 in m/z, while the final product were conformed at 953.6990 ($[M+4H]^{4+}$) and 1271.2623 ($[M+3H]^{3+}$) in m/z. Surprisingly, no other significant peaks were observed from 1407 to 3900. This means that all four azides on **93** were reacted with the propargyl group on the peptide, suggesting the completion of the reaction.
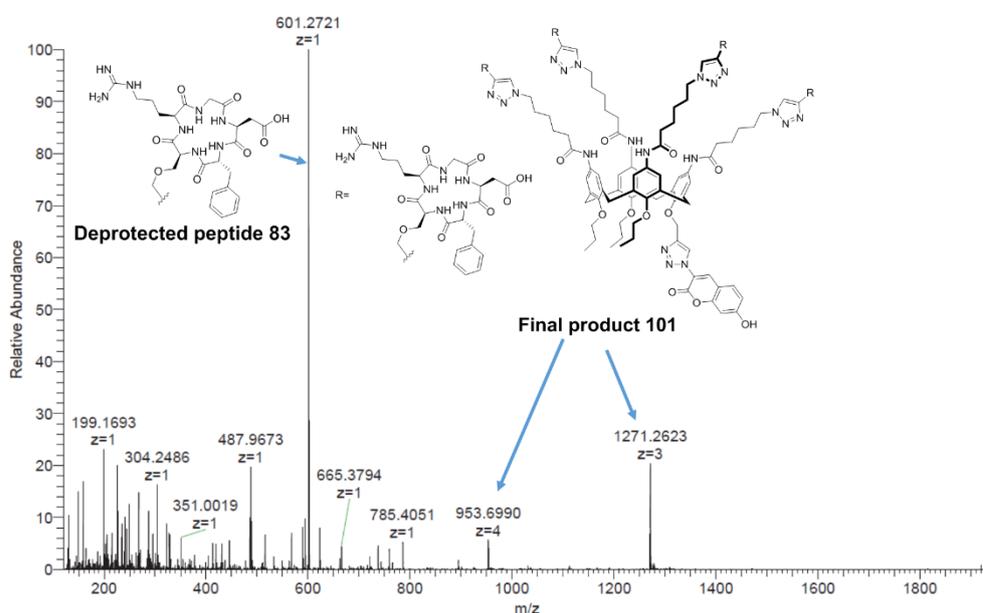
*Figure 4.36   High-Res Mass Spectrometry of **101**(crude sample)*

The crude product was then purified *via* preparative HPLC. Formation of product **101** was then confirmed using MALDI and analytic HPLC (Figure 4.37). The expected ion for calixarene **101** ($[M+H^+]$) was found at 3811.8 in the MALDI spectrum with high purity (96.4%).
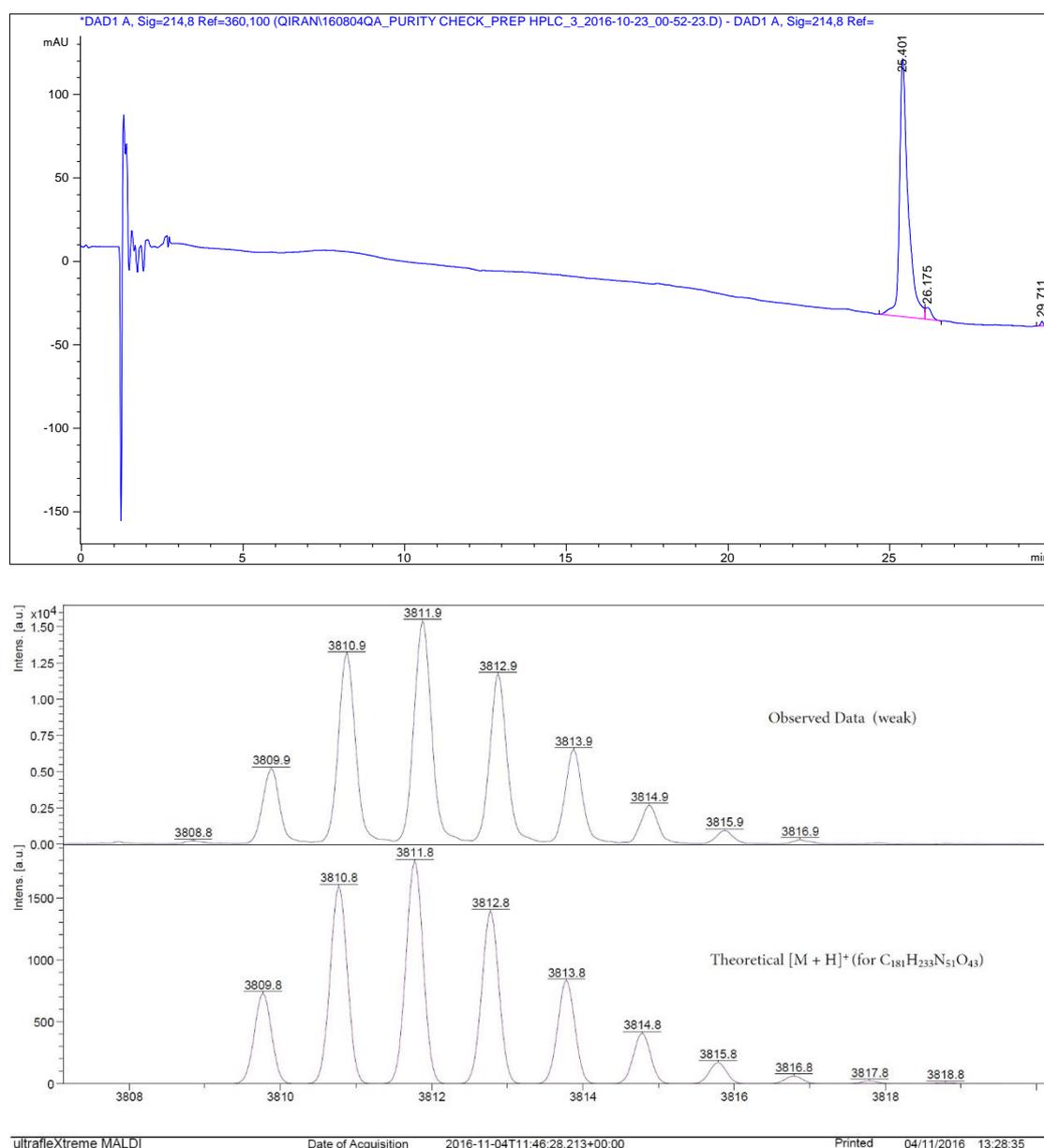
*Figure 4.37   Analytical HPLC trace (solvent trace substracted) of purified calixarene conjugate **101** and MALDI mass spectrum of **101***

4.4   Conclusions

In this study, a novel cyclic peptide-dye conjugate **99** and cyclic RGD-appended calixarene-based fluorescent probe **101** were synthesised and characterised. This was acheived through the synthesis of a novel cyclic RGD peptide, **83**, bearing a

propargyl group for 'click' reactions. It was shown that this peptide can be utilised in 'click' reactions using a chelated copper source. It was also shown that when azido functionalite was directly attached to the upper rim of the calixarene 'click' reactions with cRGDfS(*OPrp*), using $CuSO_4$ and sodium ascorbate were unsucesful. This is likely to be due to steric hinderance on the upper rim although the role of the copper source may also be important.

Chapter 5   Overview and Future Work

5.1 Overview

This thesis has described one i-motif ligand screen method and the application of calix[4]arene to target DNA secondary structures and integrin. *Chapter 1* provides a comprehensive overview of DNA, focusing on DNA secondary structures. Particular emphasis has been placed on i-motifs, including their biological relevance, methods for studying them and current ligands.

In *Chapter 2*, an i-motif ligand screening method was established based on fluorescent intercalator displacement. Five fluorescent probes were tested and thiazole orange was chosen as the fluorescent intercalator due to its 'light-up' fluorescent properties upon binding to human telomeric i-motif. The binding stoichiometry and stabilisation effect of thiazole orange towards human telomeric i-motif was also studied. The feasibility and accuracy of this screening method was evaluated with previously known i-motif binding ligands, such as mitoxatrone, TMPyP4, tilorone and tamoxifen. A plate reader based i-motif ligands screen was then developed using thiazole orange and human telomeric i-motif. The high-throughput screen method was assessed with a compound library containing 960 biologically active ligands. It was found that tobramycin was an effective 'hit' compound in the screen, which interestingly did not show noticeable activity in a comparative screen based on FRET DNA melting. The binding of tobramycin towards human telomeric i-motif was further studied *via* SPR and circular dichroism. The binding affinity of tobramycin was found to be of the same magnitude (23 ± 3.7 µM) as that for mitoxatrone, a previously known i-motif binding ligand. However, it was also found that tobramycin binds to duplex DNA structure with even higher affinity, indicating poor selectivity between different DNA secondary structures.

In *Chapter 3*, the synthesis of five calixarene conjugates tethered with DNA binding moieties was described. The interaction of two compounds, **28** and **54**, were studied

with G-quadruplexes and i-motifs DNA sequences due to their relatively good solubility in water. FRET DNA melting experiments indicated that both **28** and **54** were able to condense G-quadruplex and i-motif forming sequences from human telomere and c-MYC promoter. This calixarene-conjugatre induced DNA condensation is stable at acidic pH even at 95°C. However, it was shown to be reversible at pH 7.4 when heat was applied. The DNA condensation induced by **28** and **54** was verified using other experiments, including FRET titration, circular dichroism and FID assay.

In *Chapter 4*, a proof-of-concept study exploring the binding of peptide functionalised calix[4]arenes to integrin cell membrane receptors was described. A cyclic RGD peptide featuring an alkyne group, **83**, was designed and synthesized using a solid phase Fmoc strategy. This was combined, through 'click' chemistry, with a calixarene functionalized with a fluorescent coumarin dye motif. Direct introduction of the cyclic peptide at the upper-rim was not successful and may be due to either steric hinderance at the upper-rim or the choice of catalytic copper (I) source. In contrast, the 'click'reaction between **83** and **93** was successful and the product **101** was purified using preparative HPLC and confirmed by MALDI.

5.2   Future work

Due to the fact that this thesis is divided into three components, future works will be described in the chapter order.

In *Chapter 2*, a high-throughput i-motif screen had been performed on a compound library containing 960 ligands. The result was the first time to show tobramycin binds to human telomeric i-motif structure. Thus the future work for this chapter will be further exploring other compound libraries with high-throughput screen method as well as the binding studies of tobramycin.

In our group, there is another compound library available for ligand screen, which

contains another 1500 bioligically active ligands. This library can be used with FID high-throughput screen method in the future, aiming to find more i-motif binding compounds as well as further testing the accuracy of this high-thoughput screen method.

Further exploring tobramycin will give us more information about its binding towards human telomeric i-motif. Mass spectrometry,[342] X-Ray diffraction[343,344] and NMR[345,346] are all within the scope in the following studies. These tools will provide detailed information about the binding between human telomeric i-motif and tobramycin, which paves the way to achieve structural selectivity *via* modifying tobramycin in the future.

In *Chapter 3*, **28**, **54** and **58** were proved to reversibly condense various oligonucleotides with at pH 7.4. However, their abilities has not been compared to other commonly used amphiphilic calixarenes in the same conditions yet. Thus in the next step, other calixarenes featuring longer aliphatic chain (> 3) and amine functionalised upper rim can serve as comparisons to **28**, **54** and **58**.

In order to study **28**, **54** and **58**, other experimental techniques such as polyacrylamide gel electrophoresis and dynamic light scattering, can be applied as well. Gel electrophoresis provides more accurate information about the condensing of oligonucleotides while dynamic light scattering provides the size of condensation particles directly.[347,348] Both of these techniques have been used to study calixarene induced DNA condensation before. Apart from these two techniques, SPR can also be used to directly study the binding affinity of calixarene conjugates with oligonucleotides.

For *Chapter 4*, future work will involve the testing of **99** and **101** with integrin positive/negative cell lines and integrin binding ligands or antibodies. Synthetically additional conjugates featuring NIR active dye molecules, to achieve *in situ* tumour imaging in animal models could be prepared. Additionally the cyclic peptide **83** can

be conjugated to other systems, such as nanoparticles or nanocarriers to achieve integrin targeting.

Finally, there are currently no examples of calixarenes functionalised with oligonucleotides or PNA sequences and their application in targeted drug delivery. Building on the integrin targeted calix[4]arene, the lower rim could be functionalised with DNA and drug molecules incorporated through annealing with complimentary DNA/RNA sequences. Dissociation of the structure and release of the drug could then be achieved following internalization through i-motif formation at endosomal pH (Scheme 5.1). A similar approach has previously been described for gold nanoparticles based delivery.[12,349]
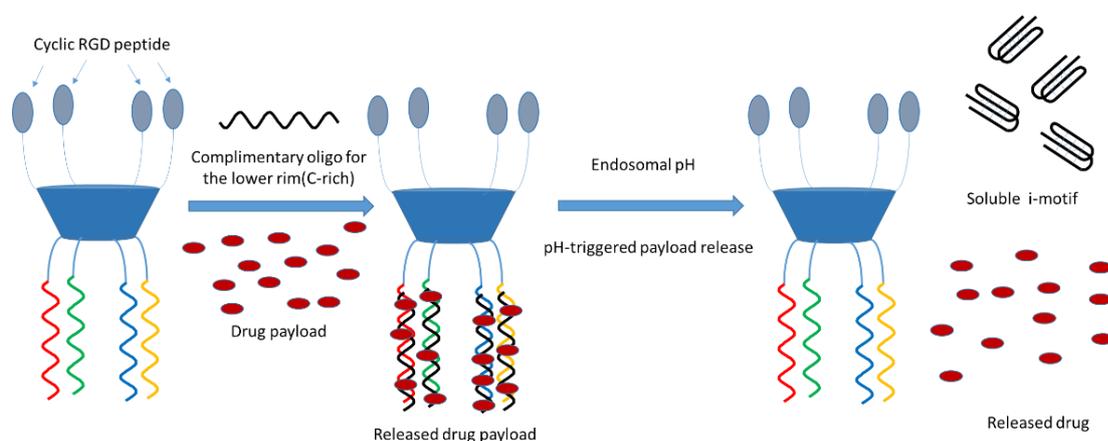
Scheme 5.1 *Proposed oligonucleotide-grafted calixarene conjugates for targeted cytotoxicity*

Chapter 6   Experimental

6.1   General procedures

All chemicals were purchased from Sigmal-Aldrich, Acrôs Organics, Alfa Aesar, MerckMillipore or Thermo Fisher Scientific and used without any further purification.

NMR spectroscopy was performed at 293K, unless otherwise stated, using a Bruker Ultrashield Plus 400 MHz spectrometer (operating at 400 MHz for proton and 100 MHz for carbon). Deutrated solvents for NMR spectroscopy were obtained from VWR international, Sigmal-Aldrich and Goss scientific. The NMR spectra were processed with MestReNova (version 6.0.2-5475). The spectra were calibrated to the residual deuterated solvent peak. Chemical shifts in ^{1}H NMR spectrum are presented in δ (ppm) followed by brackets containing multiplicity(s: singlet, d: doublet, t, triplet, q:quartet, m: multiplet, br:broad). ^{13}C NMR spectras are only presented with chemical shifts.

Thin-layer chromatography was performed with Merck Kieselgel 60 F_{254} plates and spots were visualized with UV-light or stained using ninhydrin or *p*-anisaldehyde.

Column chromatography was performed with Aldrich 60 silica gel at room temperature with a positive pressure of air.

Melting points were measured with an electrothermal Mel-temp melting point apparatus.

Infrared spectroscopy were performed with an FT-IR spectrometer with ATR attachment. The data are presented in wavenumbers (cm^{-1}).

Opitical rotation of chiral compounds were measured *via* an ADP220$^{®}$ polarimeter at 25°C.

Oligonucleotides were purchased from Eurogentec, purified *via* reverse-phase HPLC. Solid DNA samples were dissolved in MilliQ water to prepare stock solution of 100 μM for labelled and 1 mM for unlabelled oligonucleotides. Final concentration of oligonucleotide stock were confirmed *via* UV absorbance at 260 nm with a Nanodrop® ND-1000 spectrophotometer according to the calculated extincition coefficients provided by Eurogenetec.

Name	Sequence 5′ to 3′
hTeloG	d(GGGTTAGGGTTAGGGTTAGGG)
hTeloC	d(TAACCCTAACCCTAACCCTAACCC)
c-MYCC	d(TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA)
FRET hTeloG	FAM-d(GGGTTAGGGTTAGGGTTAGGG)-TAMRA
FRET DS	FAM-d(TATAGCTATA-HEG(18)ª-TATAGCTATA)-TAMRA
FRET hTeloC	FAM-d(TAACCCTAACCCTAACCCTAACCC)-TAMRA
FRET c-MYCC	FAM-d(TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA)-TAMRA
FRET HIF-1αC	FAM-d(CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGC)-TAMRA
Biotin hTeloC	Biotin-d(TAACCCTAACCCTAACCCTAACCC)
Biotin c-MYCC	Biotin-d(CCTTCCCCACCCTCCCCACCCTCCCCA)
Biotin DS	Biotin-d(GGCATAGTGCGTGGGCGTTAGC)
compliment DS	d(GCTAACGCCCACGCACTATGCC)

Table 5.1   Oligonucleotide sequences used in Chapter 2 and Chapter 3

ª HEG(18), Hexaethylene glycol 18 ,is a spacer between two complementary sqeuences of duplex DNA sequences.

DNA annealing was performed by heating the sample at 95°C in a heat block for 5 minutes and then cooling down to room temperature slowly overnight.

High-resolution Mass Spectrometry were performed with LTQ orbitrap XL, ultrafleXtreme (MALDI) or Xevo G2-S and recorded by the EPSRC UK National Mass Spectrometry Facility in Swansea University.

MALDI-TOF mass spectroscopy were carried out with a Kratos Axima CFR MALDI Mass Spectrometer. α-cyno-4-hydroxycinnamic acid was used as matrix.

Analytical RP-HPLC were carried out with an Agilent Technologies 1200 series chromatograph with an Agilent Technologies® ZORBAX Eclipse XDB-C18 (5 μM, 4.6 x 150mm) at 25°C.

Solvent A: 100% HPLC-grade methanol + 0.1% TFA.

Solvent B: 100% MilliQ-water + 0.1% TFA.

Gradient: Solvent A to Solvent B in 25 minutes. Flow rate = 1 mL/min. Detection wavelength = 214 nm/254 nm.

Semi-preparative RP-HPLC were carried out with Agilent Technologies 1200 series chromatograph with an Agilent Technologies ZORBAX Eclipse XDB-C18 (5 μm, 9.4 × 250 mm) column with a flow rate of 4 mL/min (same gradient as used in analytical RP-HPLC).

Preparative RP-HPLC were carried out with an Agilent 1260 Infinity using an Agilent Eclipse XDB-C18 column (5 μM, 21.2 × 150 mm).

Solvent A: 5% methanol in water + 0.05% TFA.

Solvent B: 5% water in methanol + 0.05% TFA.

Gradient: Solvent A to Solvent B in 20 minutes. Flow rate = 20 mL/min. Detection wavelength = 254 nm.

Data obtained in the studies was processed by Microsoft Office Excel and presented in Origin 8.0.

6.2   Experimental for Chapter 2

6.2.1   Fluorescence Spectroscopy

Fluorescence spectra were measured on a Horiba Jobin Yvon Fluorolog®
fluorimeter with a Starna Scientific type 28/9-F quartz cuvette with 1 cm path
length. All experiments were performed in triplicate.

DMSO stock solutions of the probes were diluted to 2.5 µM in 10 mM sodium
cacodylate buffer at pH 5.5. DNA samples were prepared by diluting to 500 µM in
10 mM sodium cacodylate buffer at pH 5.5. In the experiments, 200 µL of probe
(2.5 µM) was added into the cuvette then 1 µL of DNA (at 500 µM) was added to
mimic the 1:1 binding event. After mixing and allowing 5 minutes to equilibrate,
spectra were measured using the following settings: 1.0 nm increment, 0.5
integration time and 5 nm excitation and emission slit widths. The fluorescence was
measured at the respective excitation and emission maxima.

Flouorescence enhancement measurements on TO were performed in collaboration
with Tasnim Mahmound, using TO diluted to 5.0 µM in 10 mM sodium cacodylate
buffer at pH 5.5. DNA samples were diluted to 100 µM in 10 mM sodium
cacodylate buffer at pH 5.5. In the experiments, 200 µL of TO (2.5 µM) was added
to the cuvette and a fluorescence spectrum taken to observe the fluorescence in the
buffer conditions in the absence of DNA. DNA was then titrated into the sample
using the prepared 100 µM stock, mixed, allowed to equilibrate for 5 minutes then
excited at 430 nm; fluorescence emission spectra were measured from 440 to 650
nm.

6.2.2   UV Spectroscopy

UV spectroscopy was performed in collaboration with Joe Neaverson. Samples
were measured using an Agilent Technologies Cary 60 UV-vis spectrophotometer

with a Starna Scientific type 28/9-F quartz cuvette of 1 cm path length. The samples were scanned over a range of 350-550 nm at a rate of 300 nm/s with a data interval of 0.5 nm at room temperature. Samples containing only buffer were measured and subtracted from the data. All experiments were performed in triplicate.

Continuous variation (Job's) binding analysis was performed on measuring the UV absorbance of TO. 200 μL samples containing hTeloC DNA with TO were made up in 10 mM, pH 5.5 sodium cacodylate buffer. The individual concentrations of DNA and TO varied between 0 and 20 μM but the total concentration of DNA + TO remained constant at 20 μM. Job plots were constructed by plotting UV absorbance at 497.5 nm against the ratio of TO to hTeloC i.e. [TO]/([TO] + [hTeloC]).

UV titrations to determine binding affinity were performed using hTeloC (5 μM) in 50 mM sodium cacodylate buffer at pH 5.5. 10 μM stock of TO was prepared in sodium cacodylate 50 mM buffer at pH 5.5 for titration into the 200 μL cuvette filled with 5 μM hTeloC to give concentrations of 0, 1.25, 2.5, 3.75, 5, 7.5, 10, 15 and 20 μM. The UV-vis spectra of 50 mM sodium cacodylate buffer at pH 5.5 with cooresponding amount of DMSO was recorded and subtracted from data obtained with TO. Experiments were performed in triplicate.

Binding curve fitting by least squares method was carried out using different binding models corresponding to 1:1 (Equation 2.1 in Section 2.3.2), 2:1 (Equation 2.2 in Section 2.3.2) and n:1 (Equation 5.1) stoichiometry.[130-132]

1:1 binding model

$$[TO] + [DNA] \rightleftharpoons [TO \cdot DNA] \qquad\qquad K_1$$

Equation 2.1 $\qquad \theta = \dfrac{nK_1[Ligand]}{1 + K_1[Ligand]} \qquad n = 1$

Where: θ = fraction of binding measured, K_1 = equilibrium association constant for the binding sites.[131,132]

2:1 independent binding model
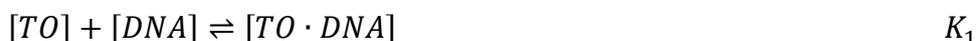
$$[TO] + [DNA] \rightleftharpoons [TO \cdot DNA] \qquad\qquad\qquad K_1$$

$$[TO \cdot DNA] + [TO] \leftrightharpoons [TO \cdot DNA \cdot TO] \qquad\qquad K_2$$

Equation 2.2
$$\theta = \frac{K_1[Ligand] + 2K_1K_2[Ligand]^2}{1 + K_1[Ligand] + K_1K_2[Ligand]^2}$$

θ = fraction of binding measured, K_1 and K_2 = equilibrium association constants for the first and second binding sites.[130-132]

n:1 binding model (multiple independent binding sites) [131,132]

$$n[TO] + [DNA] \rightleftharpoons [nTO \cdot DNA] \qquad\qquad\qquad K_1$$

Equation 5.1
$$\theta = \frac{nK_1[Ligand]}{1 + K_1[Ligand]}$$

θ = fraction of binding measured, K_1 = equilibrium association constant for the binding sites, n = number of independent binding sites [131,132]

6.2.3   Circular dichroism spectroscopy.

Circular dichroism spectroscopy was performed in collaboration with Joe Neaverson. Oligonucleotides (10 µM) were diluted in buffer at the respective pH and concentration as detailed, to give a total volume of 200 µL for each sample and were annealed. Circular dichroism spectra were recorded using a Jasco J-810 spectropolorimeter with a Starna Scientific type-21 quartz cuvette of 1 mm path length. Scans were performed at room temperature from 220 nm to 320 nm with a

scanning speed of 200 nm/min, a response time of 1 s, pitch of 0.5 nm and a bandwidth of 2 nm. Blank samples containing only the respective buffer were recorded and subtracted from the data obtained later. For each sample, TO or tobramycin was titrated directly into the cuvette containing the DNA using pipette and solutions were thoroughly mixed and given an equilibrium time for 5 minutes before recording spectra. All spectra at each aliquot shows an average of three scans.

6.2.4 FRET melting experiments.

FRET melting experiments were performed in collaboration with Joe Neaverson. Solutions of fluorescently labelled DNA (0.4 µM) were made up in sodium cacodylate buffer at the respective concentration and pH as detailed and annealed. Thiazole orange 50 µM was prepared in sodium cacodylate. Serial dilutions of 25, 12.5, 6.25, 3.125, 1.56 and 0 µM thiazole orange were used in combination with 0.4 µM of each FRET-labelled DNA sequence for the FRET melting experiments by pipetting 10 µL of each appropriate DNA and thiazole orange solution (total 20 µL) into PCR tubes from QIAgen. The samlples in the PCR strips were heated from 25 to 95°C at a rate of 1°C per minute. Fluorescence melting curves were obtained using a QIAgen Rotor-Gene Q-series PCR machine using an excitation wavelength of 483 nm and detection wavelength of 533 nm. Each experiment was performed in duplicate.

6.2.5 i-motif FID assay using TO.

TO was diluted to 2 µM sample in 10 mM sodium cacodylate buffer at pH 5.5. hTelo i-motif in water was used without being annealed and made into 50 µM stock in buffer. Tested ligands, including TmPYP$_4$, tilorone, mitoxatrone, tamoxifen and tyrothyricin, was prepared into 50µM in 10 mM sodium cacodylate buffer at pH 5.5. 196 µL of TO sample was added into a Starna Scientific type 28/9-F quartz cuvette of path length of 1 cm. Fluorescence spectra were measured on a Horiba Jobin Yvon Fluorolog fluorimeter, excited at 430 nm. The excitation and emission

slit width were both set as 3 nm. The integration time was set as 0.1 s and increment was set as 10.0 nm. The resulting background emission fluorescence intensity at 450 nm was normalised as 0%. To this solution, 4 µL of hTelo-i-motif was added and mixed with the TO sample. 5 minutes later, a second background fluorescence spectrum was taken when the sample was excited at 430 nm. The fluorescence emission intensity here at 450 nm was normalized as 100% fluorescence. Then the respective ligand was titrated in 1 µL aliquots and the emission spectrum measured from 450 nm to 650 nm.

6.2.6   HT-IM-FID assay

Each IM-FID assay was conducted in a 384-well microplate (Corning® Low Volume 384 well Black Flat Bottom Polystryrene NBS TM Microplate). The temperature was kept constant at 25°C during the whole experiment. Wells on the microplate were filled with 40 µL of a pre-prepared testing cocktail ([hTeloC]= 0.5 µM, [TO]= 1.0 µM, [sodium cacodylate]= 10 mM, pH= 5.5). Then 0.5 µL ligand solution (1 mM in DMSO) was added into each well. At least three control wells on each plate were prepared of just DNA in buffer and another three reference wells on each plate were prepared with DNA and TO in buffer. Plates were read ising a BMG CLARIOstar plate reader. The excitation filter in the experiments were set from 400 to 430 nm, the emission filter was set from 460 to 480 nm. The read of the signal was defined as the intergration of fluorescent intensity from 460 to 480 nm. Each scan was carried out three times. The basal fluorescence signal (F_{c0}) were assigned as the average fluorescence intensity read from the control well. The 100% fluorescence intensity read ($F_{reference}$) was assigned as the average fluorescence intensity read from reference wells. The average of three read (F_{read}) was calculated into D_x for each compound *via* Equation 2.3 in Section 2.3.4.

Equation 5.2
$$D_x = 1 - \frac{F_x}{F_0} = 1 - \frac{F_{read} - F_{c_0}}{F_{refereence} - F_{c_0}}$$

Hit compounds were ranked according to D_x.

6.2.7 Surface Plasmon Resonance

Dr Clare Stevenson in John Innes Center kindly performed SPR experiments for the studies. These experiments were performed on a Biacore T200 instrumnet manufactured by GE Healthcare with a series S streptavidin (SA) coated gold chip (Biacore SA-chip). Experiments were carried out in a filtered and degassed running buffer (10 mM sodium cacodylate, 100 mM sodium chloride and 0.05% Tween, pH 5.5).

Oligonucleotide immobilisation

Biotin-tagged hTeloC (biotinhTeloC) and c-MYCC (biotinc-MYCC) oligonucleotides were diluted to 1 μM in the running buffer while Biotin-tagged DS oligonucleotide (biotinDS) were annealed with its complimentary oligonucleotide (complimentDS) at 1 μM in the running buffer. Biotinlated oligonucleotides were then injected at a flow rate of 10 μL/min in flow cells 2, 3 and 4, leaving the flow cell 1 empty s a blank. The success of oligonucleotide immobilization was checked via the sensor, reaching appromximate 500 response unit (RU).

Tobramycin titration

Tobramycin were diluted in the running buffer from 10 mM stock solution in DMSO to concentrations of 500, 200, 100, 50, 25, 12.5, 6.25, 3.125, 1.56, 0.78, 0.39 and 0 μM *via* serial dilutions.

Tobramycin was tested for its binding affinity towards hTeloC, c-MYCC and duplex DNA using the affinity run wizards in the BIAcore T200 software at 25 °C. In a typical injection, samples were injected for 2 min as a flow rate of 30 μL/min,

followed by two 1 min injection of 1 M NaCl an one 1 min injection of running buffer for chip regeneration.

The SPR experiments started with three startup (blank) cycles. Each of them were performed with blank injections of buffer only. Tobramycin in different concentrations was tested from 0 µM to 500 µM.

The RU obtained was double referenced by substrating the three startup cycles and injections of buffer. Average equilibrium response for each injection was fitted using the affinity fit from the BIAcore T200 evaluation software v2.0 with a 1:1 binding model (Equation 2.6 in Section 2.3.8). The data obtained were the average of two repeat runs.

Equation 2.6
$$\theta = \frac{R_{eq}}{R_{max}} = \frac{nK_1[Ligand]}{1 + K_1[Ligand]} \qquad n = 1$$

Herein, θ is the fraction of binding. R_{eq} is the response read at equilibrium from the sensorgram at a given concentration whilst R_{max} is the largerst response read at saturation. The binding constant K_d was defined as $1/K_1$.[132]

Ligand screening

Hit compounds in Section 2.3.7 were tested at 50 µM in pH 5.5 running buffer (100 mM sodium chloride, 10 mM sodium cacodylate, 0.05% Tween). Oligonucleotide immobilization and ligand testing have been decribed previously. The response (RU) were measured only at 50 µM (ligand concentration) upon binding to DNA, which is propotional to the molecular mass on the chip surface. The theoretical R_{max} was describes in an 1:1 binding stoichiometry (Equation 2.5 in Section 2.3.7)

Equation 2.5
$$R_{max} = \frac{MW_{ligand}}{MW_{oligo}} \times RU_{oligo} \times i.$$

Herein the MW_{ligand} stands for the molecular weight of tested ligands whilst MW_{oligo} stands for the molecular weight of the immobilised oligonucleotides on the chip. RU_{oligo} is the average respond read of startup runs measured in SPR experiments (injections without ligands). i is the pre-set stoichiometry of binding, which equals 1 if it is a 1:1 binding stoichiometry.

6.3   Experimental for Chapter 3

6.3.1   Synthesis of *p-tert*-butylcalix[4]arene

The synthesis of *p-tert*-butylcalix[4]arene was performed according to literature.[221]
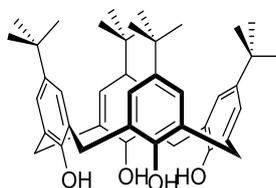
5, 11, 17, 23-*p-tert*-butylcalix[4]arene (29) [221]

p-tert-butylphenol (200.0 g, 1.33 mol), NaOH (2.40 g, 60 mmol) and formaldehyde (130 mL, 37% wt, 1.74 mol) were mixed in a 3 L flange flask. The reaction mixture was heated to 120°C for 2 hours with mechanical stirring. The resulting orange-coloured viscous solid were re-dissolved in diphenyl ether (1 L). A steam of air was introduced to the surface of diphenyl ether solution for at least 1 hour to remove the excess water in the reaction mixture. Then the reaction mixture was heated to reflux for another 3 hours. 1 L of ethyl acetate was added to the reaction mixture after it had been cooled to RT. The resulting crude product precipitated and was filtered. Then the crude product was washed with ethyl acetate (2 × 100 mL), acetic acid (2 × 50 mL), water (2 × 50 mL) and acetone (100 mL). After being dried in the oven (120 °C) overnight, the product was an off-white crystal powder **29** (90.53 g, 42% yield). **¹H NMR** (400MHz, CDCl₃) δ 10.34 (br s, 4H, ArOH), 7.91 (br s, 8H, ArH), 4.26 (d, J = 13 Hz, 4H, ArCH₂Ar),3.49 (d, J = 13 Hz, 4H, ArCH₂Ar), 1.21 (br s, 36H, C(CH₃)₃), **IR** ν 3158, 2952, 2901, 2870, 2351, 1738, 1704, 1605, 1556, 1538, 1480, 1472, 1467, 1454, 1392, 1361, 1305, 1231, 1199, 1123 cm⁻¹,

Melting point: 344-346°C (decomposed). The spectroscopic data are consistent with that reported in the literature.[221]

6.3.2 Preparation of calix[4]arene platforms

Synthesis of calixarene platforms are, including mono-propargylated, 1,2- and 1,3-dipropargylated calixarenes following established methods in Matthews' group. [197,226]
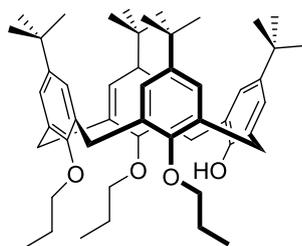
5, 11, 17, 23-Tetra-*tert*-butyl-25, 26, 27-tripropoxy-28-hydroxy-calix[4]arene (30) [226]

A mixture of **29** (20.11 g, 31.03 mmol, 1 eq), barium oxide (7.15 g, 46.78 mmol, 1.5 eq) and barium hydroxide octa-hydrate (33.82 g, 107.35 mmol, 3.45 eq) in 300 mL of DMF was heated to 30°C for 45 minutes. *n*-Propyl bromide (60 mL, 660 mmol, 20.0 eq) was then added into the reaction mixture and stirred for 18 hours at 30°C. The cream colored reaction mixture was quenched with the addition of dilute HCl (10%, 100 mL). The reaction mixture was then extracted with DCM (3 × 200 mL). The resultant organic phase was further washed with dilute HCl (10%, 3 × 100 mL) and brine (2 × 200 mL) and then dried over anhydrous MgSO$_4$. The solvent was removed *in vacuo* and a cream colored solid was obtained. The final white-crystalline product **30** (16.5 g, 69% yield) was recovered by triturating from DCM/methanol. **^{1}H NMR** (CDCl$_3$, 400MHz) δ$_H$ ppm= 7.13 (s,2H), 7.04 (s, 2H), 6.50 (s, 4H), 5.57 (s, 1H), 4.36 (d, *J* = 15 Hz, 2H), 4.32 (d, *J* = 15 Hz, 2H), 3.84 (t, *J* = 8 Hz, 2H), 3.74 (t, *J* = 7.0 Hz, 2H), 3.22 (d, *J* =13.0 Hz, 2H), 3.15 (d, *J* =13.0 Hz, 2H), 2.38-2.27 (m, 2H), 2.07-1.78 (m, 4H), 1.54 (s, 3H), 1.33 (s, 9H), 1.31 (s,

9H), 1.09 (t, J = 7.0 Hz, 6H), 0.94 (t, J =7.0 Hz, 3H), 0.82 (s, 18H), **IR** ν 3520, 2957, 2872, 1694, 1603, 1555, 1537, 1481, 1477, 1386, 1360, 1298, 1240, 1196, 1121, 1108cm^{-1}, **Melting point**: 200-202°C. The spectroscopic data are consistent with those reported in the literature.[226]
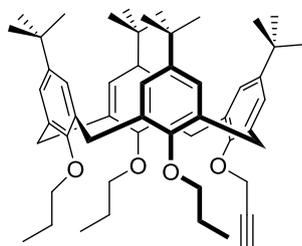
5, 11, 17, 23-Tetra-tert-butyl-25, 26, 27-tripropoxy-28-propargyloxy-calix[4]arene (31) [177]

Calixarene **30** (16.5 g, 21.65 mmol, 1.0 eq) was stirred in 430 mL DMF stirring at room temperature. Sodium hydride (2.03 g, 84.58 mmol) was carefully added into the reaction mixture, followed after one hour by addition of propargyl bromide toluene solution (80% w/w in toluene, 10.63 g, 71.5 mmol). The reaction mixture was stirred for 18 hours, then quenched by addition of 400 mL water. The resultant white precipitate was collected by filtration. The precipitate was redissolved in methanol/DCM and the resultant brown organic phase was washed with dilute HCl (10%, 100 mL) and then brine (3 × 250 mL). The organic phase was dried over MgSO$_4$, filtered and concentrated, to obtain sticky brown-colored oil. After trituration in methanol/DCM, the final product **31** was recovered as white-colored crystals (13.8 g, 80% yield). **^{1}H NMR** (CDCl$_3$, 400 MHz) δ$_H$ ppm= 7.02 (s, 4H), 6.60 (d, J = 2.0 Hz, 2H), 6.54 (d, J = 2.0 Hz, 2H), 4.98 (d, J =2.0 Hz, 2H), 4.53 (d, J =13 Hz, 2H), 4.45(d, J= 13 Hz, 2H), 3.93 (t, J = 7.5Hz, 2H), 3.75 (t, J = 7.5 Hz, 4H), 3.17 (d, J = 2.0 Hz, 2H), 3.14 (d, J =2.0 Hz, 2H), 2.40 (t, J = 2.0 Hz, 1H), 2.22-2.08 (m, 2H), 2.06-1.93 (m, 2H), 1.56 (s, 1H), 1.28(d, J = 2.0 Hz, 18H), 1.07 (t, J =7.5 Hz, 6H), 1.02 (t, J =7.5 Hz, 3H), 0.92 (s, 18H); **IR** ν 3561, 3478, 2956, 2876, 2856, 1973, 1573, 1537, 1479, 1488, 1467, 1385, 1360, 1299, 1238, 1195, 1122 cm^{-1}; **Melting point**: 213-215°C. The spectroscopic data are consistent with

those reported in the literature.[177]
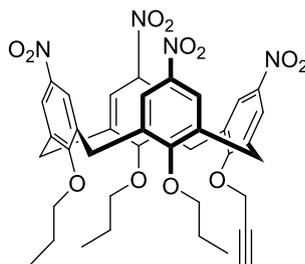
5, 11, 17, 23-Tetra-nitro-25, 26, 27-tripropoxy-28-propargyloxy-calix[4]arene

(32)

Calix[4]arene **31** (6.5 g, 8.45 mmol, 1.0 eq) was dissolved in DCM (100mL) and stirred vigorously in a 250 mL round bottom flask. Then glacial acetic acid (35 mL, 612.0 mmol) and nitric acid (98%, 35 mL, 840.4 mmol) were added into reaction mixture slowly. The color of the reaction mixture turned to dark-blue when concentrated nitric acid was added. The color of the reaction mixture turned slowly from dark-blue to orange over 5 minutes. After stirring for 20 minutes, the reaction mixture was quenched with addition of water (100 mL). After stirring for a further 10 minutes, the reaction mixture was separated and the aquenous layer was extracted with additional DCM (50 mL). The organic layer was combined and then washed with brine (2×75 mL). The organic layer was dried over anhydrous $MgSO_4$ and DCM was evaporated *in vacuo*. The crude solid product was then redissolved in a small amount of DCM (≤ 5 mL) and then precipitated in methanol overnight to afford a pale-yellow colored solid as the pure product **32** (3.9g, 63% yield). **^{1}H NMR** (400MHz, $CDCl_3$) δ 8.07 (s, 2H, Ar$\underline{H}$), 8.06(s, 2H, Ar$\underline{H}$), 7.16 (s, 4H, Ar$\underline{H}$), 4.92(d, J = 2Hz, 2H, OC$\underline{H}_2$CCH), 4.60(d, J = 14Hz, 2H, ArC$\underline{H}_2$Ar), 4.55 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 4.14 (t, J = 8 Hz, 2H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.82 (t, J = 7 Hz, 4H, OC$\underline{H}_2$CH$_2$CH3), 3.44 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 3.42 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 2.51 (t, J = 2 Hz, 1H, OCH$_2$CC$\underline{H}$), 1.98-1.87 (m, 6H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.10 (t, J = 8 Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$), 0.97 (t, J = 7Hz, 3H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C-NMR** (101 MHz, $CDCl_3$) δ 162.8, 160.9, 160.2, 143.8, 142.9, 137.9, 136.6, 134.5, 134.2, 124.8, 124.7, 123.4, 123.3, 78.2, 78.2, 78.0, 77.1, 60.7, 31.5, 31.1, 23.5, 23.4,

10.6, 10.0. **IR**: ν 3417, 3170, 2972, 2923, 1700, 1596, 1532, 1482, 1466, 1370, 1366, 1296, 1240, 1214, 1152, 1060 cm^{-1}; **HRMS (ESI)**: calculated for $C_{40}H_{36}N_4O_{12}H$ ([M+H]$^+$): 769.2721. Found: 769.2708; **Melting Point:** 175-177 °C.
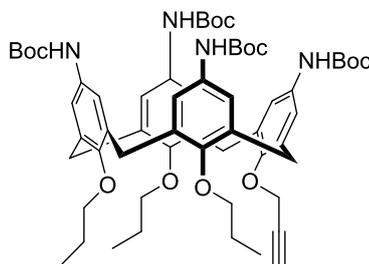
5,11,17,23-Tetra-Boc-amino-25,26,27-tripropoxy-28-propargyloxycalix[4]arene (33)

Tin(II) chloride dehydrate (19.38 g, 85.89 mmol) was added to a mixture of **32** in ethanol (200 mL). The reaction mixture was heated to 90°C and stirred for 24 hours. The solvent was evaporated and the resulting white-pink solid was dissolved in NaOH (200 mL, 10%) and DCM (100 mL). Then the organic phase was separated and the aqueous phase was extracted with DCM (50 mL) and EtOAC (50 mL). The organic phases were combined and dried with Na$_2$SO$_4$ and the solvent was evaporated to obtain a brown-coloured crystalline solid (2.8 g). The solid was dissolved in DCM (100 mL) and Boc anhydride (10 g, 45.6 mmol) and DIPEA (3 mL, 18.2 mmol) were added. The mixture was stirred for 18 hours before removal of the solvent *in vacuo*. Column chromatography (eluent: DCM:EtOAc = 15:1) afforded the product **33** as a white solid (1.23 g, 25 % yield). **^{1}H-NMR** (400MHz, CDCl3) δ 7.05 (s, 2 H, ArH), 7.04 (s, 2 H, ArH), 6.37 (s, 2 H, NHBoc), 6.26 (s, 2 H, ArH), 6.24 (s, 2 H, ArH), 5.90 (s, 2 H, NHBoc), 4.85 (d, *J* = 3 Hz, 2 H, OCH$_2$CCH), 4.43 (d, *J* = 13 Hz, 2 H, ArCH$_2$Ar), 4.36 (d, *J* = 13 Hz, 2 H, ArCH$_2$Ar), 3.92 (t, *J* = 8 Hz, 2H, OCH$_2$CH$_2$CH$_3$), 3.63 (t, *J* = 6 Hz, 4 H, OCH$_2$CH$_2$CH$_3$), 3.10 (d, *J* = 13 Hz, 2 H, ArCH$_2$Ar), 3.08 (d, *J* = 13 Hz, 2 H, ArCH$_2$Ar), 2.31 (t, *J* = 3, 1 H, OCH$_2$CCH), 1.96-1.81 (m, 6 H, OCH$_2$CH$_2$CH$_3$), 1.53 (s, 9 H, C(CH$_3$)$_3$), 1.53 (s, 9 H, C(CH$_3$)$_3$), 1.43 (s,18 H, C(CH$_3$)$_3$), 1.02 (t, *J* = 7 Hz, 6 H, OCH$_2$CH$_2$CH$_3$), 0.90

(t, J = 7 Hz, 3 H, OCH$_2$CH$_2$CH$_3$). ^{13}C NMR (100Hz, CDCl$_3$) δ 153.7, 153.4, 153.2, 153.1, 152.4, 150.9, 137.8, 136.7, 134.0, 133.7, 133.3, 132.5, 132.0, 120.43, 120.4, 119.4, 119.0, 81.0, 80.3, 80.3, 80.0, 74.7, 59.6, 31.7, 31.1, 28.6, 23.5, 23.2, 10.9, 10.2; **IR** ν 3424, 2978, 2930, 2874, 1730, 1700, 1596, 1532, 1526, 1466, 1418, 1366, 1296, 1242, 1214, 1152, 1060 cm^{-1}; **HRMS (ESI):** calculated for C$_{60}$H$_{80}$N$_4$O$_{12}$NH$_4$ [M+NH$_4$]: 1066.6111, Found: 1066.6107 **;Melting Point:** 182-184 °C
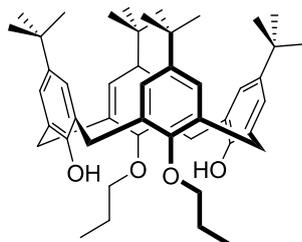
5, 11, 17, 23-Tetra-tert-butyl-25, 26, 27-tripropoxy-28-propargyloxy-calix[4]arene (34) [226]

K$_2$CO$_3$ (5.66 g, 40.95 mmol) was added to a suspension of **29** (12.04 g, 18.74 mmol) in acetone (240 mL). The reaction mixture was heated to reflux for 1 hour. After the addition of 1-bromopropane (8.36 mL, 91.96 mmol), the reaction heated at reflux for 18 hours. The reaction mixture was filtered after cooling and the filtrate was evaporated to obtain the crude product. Then the crude product was re-dissolved in DCM and the organic phase was washed with water (2 × 200 mL) and brine (200 mL). After drying over anhydrous MgSO$_4$, the solvent was removed under reduced pressure to obtain a white-coloured solid. **34** was recrystallize from DCM/methanol as a white crystalline solid (6.05 g , 46% yield). ^{1}**H NMR** (400MHz, CDCl$_3$) δ 7.90 (s, 2H, ArO$\underline{H}$), 7.04 (s, 4H, Ar$\underline{H}$), 6.87 (s, 4H, Ar$\underline{H}$), 4.30 (d, J =13 Hz, 4H, ArC$\underline{H}_2$Ar), 3.95 (t, J = 6.5 Hz, 2H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.31 (d, J =13 Hz, 4H, ArC$\underline{H}_2$Ar), 2.03 (sext, J = 6.5 Hz, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.25 (m, 24H, C(CH$_3$)$_3$ +OCH$_2$CH$_2$C$\underline{H}_3$), 0.99 (s, 18H, C(CH$_3$)$_3$). **IR** ν 3298, 2957, 2826, 1697, 1681, 1592, 1556, 1538, 1485, 1360, 1297, 1200, 1196, 1125 cm^{-}1, **Melting point**: 251-253°C. The spectroscopic data are consistent with that reported in the literature.[226]
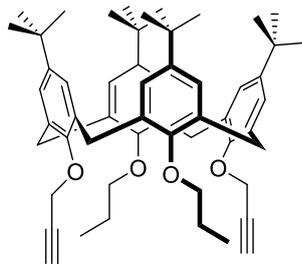
5, 11, 17, 23-*p-tert*-butyl 25, 27-dipropoxy-26, 28-propargylcalix[4]arene (35)

[177]

Sodium hydride (1.3 g, 54.5 mmol) was added to a stirred solution of **34** (5.0 g, 7.06 mmol) in DMF (250mL). After one hour, propargyl bromide (5 mL, 80 wt% in toluene, 56.20 mmol) was added and the reaction mixture was left stirring for another 18 hours. After water was added into the mixture, the resulting white precipitate was collected by filteration and re-dissolved in DCM (200 mL). The organic solution was then washed with water (2 × 100mL) and brine (100 mL). After being dried over MgSO$_4$, the organic phase was evaporated under reduced pressure to obtain the crude product. The crude product was then titurated in DCM/methanol to yield a white-coloured product **35** (4.01 g, 77% yield). **^{1}H NMR** (400MHz, CDCl$_3$) δ 7.09 (s, 4H, Ar$\underline{H}$), 6.45 (s, 4H, Ar$\underline{H}$), 5.01 (d, *J* = 2.5 Hz, 4H, OC$\underline{H}_2$CCH), 4.52 (d, *J* = 2.5 Hz, 4H, ArC$\underline{H}_2$Ar), 3.71 (t, *J* = 7.5 Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.14 (d, *J* = 7.5 Hz, 4H, ArC$\underline{H}_2$Ar), 2.39 (t, *J* = 2.5 Hz, 2H, OCH$_2$CC$\underline{H}$), 1.99 (sext, *J* = 7.5 Hz, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.32 (s, 18H, C(C$\underline{H}_3$)$_3$), 1.07 (t, 6H, *J* = 7.5 Hz, OCH$_2$CH$_2$C$\underline{H}_3$), 0.83 (s, 18H, C(C$\underline{H}_3$)$_3$); **IR** ν 3280, 2957, 2862, 1556, 1538, 1479, 1386, 1360, 1300, 1279, 1261, 1240, 1198, 1121 cm^{-1}; **Melting point**: 190-192°C. The spectroscopic data are consistent with those reported in the literature.[177]
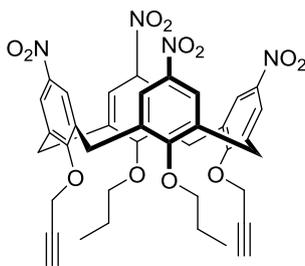
5, 11, 17, 23-tetra-nitro 25, 27-dipropoxy-26, 28-propargylcalixarene (36)

Glacial acetic acid (26.9 mL, 470.0 mmol) and 98% nitric acid (26.9 mL, 649 mmol) were added to a stirred solution of **35** (4.01 g, 4.96 mmol) in 162.5 mL DCM. The resulting blue-dark solution turned into orange after 2 hours. The reaction was quenched with addition of ice cold water (50 mL). The organic phase was washed with saturated $NaHCO_3$ solution (100 mL), additional water (2 × 100 mL) and brine (100mL). After drying over $MgSO_4$, the organic phase was evaporated under reduced pressure. The resulting yellow product **36** was collected and purified by column chromatography (eluent: DCM: Hexane = 50: 50 to 70: 30, 1.8 g, 48% yield). **^{1}H NMR** (400MHz, $CDCl_3$) δ 8.20 (s, 4H, Ar$\underline{H}$), 7.1 (s, 4H, Ar$\underline{H}$), 5.01 (d, $J = 2.5$ Hz, 4H, OC$\underline{H}_2$CCH), 4.63 (d, $J = 14.0$ Hz, 4H, ArC$\underline{H}_2$Ar), 3.81 (t, $J = 7.0$ Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.45 (d, $J = 14.0$ Hz, 4H, ArC$\underline{H}_2$Ar), 2.54 (t, $J = 2.5$ Hz, 2H, OCH$_2$CC$\underline{H}$), 1.96 (sext, $J = 7.0$ Hz, OCH$_2$C$\underline{H}_2$CH$_3$, 4H), 1.12 (t, $J = 7.0$ Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C NMR** (100Hz, $CDCl_3$) 160.6, 160.3, 144.0, 143.10, 138.0, 133.9, 124.8, 123.27, 78.2, 78.2, 77.2, 61.0, 31.4, 23.5, 10.7; **IR** ν 3292, 2970, 2934, 2872, 1586, 1518, 1454, 1342, 1266, 1264, 1210, 1160, 1094 cm^{-1} ; **HRMS (ESI)**: calculated for $C_{40}H_{36}N_4O_{12}H$ ([M+H]$^+$): 765.2408. Found: 765.2386; **Melting Point:** 185°C (decomposed).
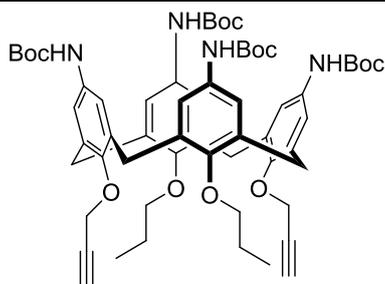
5, 11, 17, 23-Tetra-Boc-amino-25, 27-dipropoxy-26, 28-propargyloxycalix[4]arene (37)

$SnCl_2 \cdot H_2O$ (18.5 g, 82.00 mmol) was added to a stirred suspension of **36** (4.2 g, 5.49 mmol) in ethanol (200 mL) and heated at reflux (90°C) for 18 hours. After cooling, the solvent was evaporated under reduced pressure and the resulting solid was washed with NaOH solution (2 × 100 mL, 10%) and DCM (100 mL). The aqueous phase was extracted with DCM twice (2 × 75mL). The organic layer was combined, dried over anhydrous $MgSO_4$ and evaporated under vacuum. The resulting crude product (2.43 g) was dissolved and stirred directly in DCM (25 mL). Boc anhydride (1.99 g, 9.12 mmol) and 0.33 mL DIPEA (0.33 mL, 2.00 mmol) were added into the reaction mixture. The reaction was left stirring overnight. Then the reaction mixture was evaporated with solvent and purified over column chromatography (eluant: DCM:Ethyl acetate = 15:1) to yield an off-white solid **37** (1.1 g, 19% yield). **^{1}H NMR** (400MHz, $CDCl_3$) δ 7.18 (s, 4H, Ar$\underline{H}$), 6.23 (s, 4H, Ar$\underline{H}$), 4.91 (d, J = 2.5 Hz, 4H, OC$\underline{H}_2$CCH), 4.47 (d, J = 13.5 Hz, 4H, ArC$\underline{H}_2$Ar), 3.62 (t, J = 7.5 Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.13 (d, J = 13.5 Hz, 4H, ArC$\underline{H}_2$Ar), 2.35 (t, J = 2.5 Hz, 2H, OCH$_2$CC$\underline{H}$), 1.89 (sext, J = 7.5 Hz, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.51 (s, 18H, Boc C(C$\underline{H}_3$)$_3$), 1.41 (s, 18H, Boc C(C$\underline{H}_3$)$_3$), 1.06 (t, J = 7.5 Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 153.5, 152.3, 152.2, 150.8, 137.80, 133.5, 133.3, 131.9, 120.3, 118.9, 80.7, 80.2, 79.8, 74.9, 59.9, 31.5, 28.4, 28.4, 23.5, 10.9. **IR**: ν 3298, 2964, 2930, 2876, 1690, 1554, 1548, 1522, 1476, 1416, 1368, 1246, 1210, 1152, 1070 cm^{-1}; **HRMS (ESI):** calculated for $C_{60}H_{76}N_4O_{12}NH_4$ [M+NH$_4^+$]:1062.5798. Found: 1062.5794; **Melting point:** 181°C (decomposed).
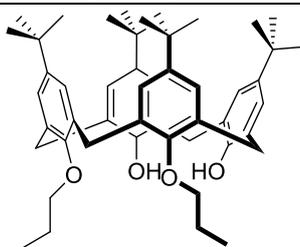
5, 11, 17, 23-*p-tert*-butyl-25, 26-dipropoxy-27, 28-bis-hydroxycalix[4]arene (38)[233]

Sodium hydride (1.71 g, 71.0 mmol) was added to a suspension of **29** (10 g, 15.56 mmol) in DMF (300 mL). After 1 hour, 3.08 mL of 1-bromopropane (33.88 mmol) was added and the reaction mixture was left stirred for 18 hours. Water (500 mL) was added and the mixture was then filtered to obtain a white coloured crude solid product. The crude product was then dried and purified by column chromatography (eluant: DCM: Hexane = 1:1) to yield white crystalline product **38** (3.46 g, 31 %) **¹H NMR** (400MHz, CDCl₃) δ 8.90 (s, 2H, O**H**), 7.02-6.95 (m, 6H, Ar**H**), 6.90 (m, 2H, Ar**H**), 4.50 (d, J = 13.0 Hz, 2H, ArC**H₂**Ar), 4.31 (m, 3H, ArC**H₂**Ar), 4.10-4.00 (m, 2H, OC**H₂**CH₂CH₃), 3.40-3.25 (m, 4H, ArC**H₂**Ar), 2.09 (sext, J = 7.0 Hz, 4H, OCH₂C**H₂**CH₃), 1.22 (s, 9H, C(CH₃)₃), 1.20 (s, 18 H, C(CH₃)₃), 1.10 (m, 27H, C(CH₃)₃+OCH₂CH₂C**H₃**) **IR** ν 3377, 2959,2904, 2872, 1599, 1488, 1437, 1383, 1360, 1316, 1292, 1258, 1210, 1195, 1124, 1103 cm⁻¹; **Melting point**: 205-207°C. The spectroscopic data are consistent with that reported in the literature.[233]
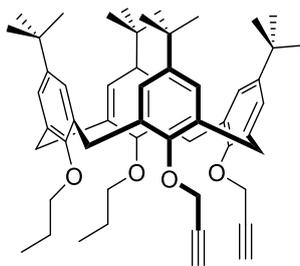
5, 11, 17, 23-*p-tert*-butyl 25,26-dipropoxy-27,28-propargylcalix[4]arene (39)

177

Sodium hydride (0.34 g, 14.4 mmol) was added to a stirred solution of **38** (1.7 g, 2.40 mmol) in DMF (160 mL). After one hour, propargyl bromide solution (80 wt% in toluene, 1.7 g, 19.5 mmol) was added and stirring was continued for 18 hours. Water (100mL) was added to quench the reaction. The resulting precipitate was collected by filteration and re-dissolved in DCM (250 mL). The organic solution was then washed with water (100 mL) and brine (100 mL). After being dried over MgSO$_4$, the solvent was evaporated under reduced pressure to obtain the crude product. The crude product was titurated in DCM/methanol to obtain 1.45 g white solid **39** (1.45g, 78.9 %). **^{1}H NMR** (400MHz, CDCl$_3$) δ 6.83 (d, *J* = 2.5 Hz, 2H, ArH), 6.79 (m, 4H, ArH), 6.75 (d, *J* = 2.5 Hz, 2H, ArH), 4.77 (d, *J* = 2.5 Hz, 4H, OCH$_2$CCH), 4.64-4.38 (m, 4H, ArCH$_2$Ar), 3.93-3.73 (m, 4H, OCH$_2$CH$_2$CH$_3$), 3.2-3.08 (m, 4H, ArCH$_2$Ar), 2.43 (t, *J* = 2.5 Hz, 2H, OCH$_2$CCH), 2.08 (m, 4H, OCH$_2$CH$_2$CH$_3$), 1.09 (s, 18H, C(CH$_3$)$_3$), 1.07 (s, 18H, C(CH$_3$)$_3$), 1.04 (t, *J* = 7.5 Hz, 6H, OCH$_2$CH$_2$CH$_3$); **IR** ν 3311, 2952, 2872, 1600, 1471, 1425, 1410, 1388, 1360, 1296, 1238, 1194, 1122, 1108 cm^{-1}; **Melting point** 173-175 °C. The spectroscopic data are consistent with those reported in the literature.[209]
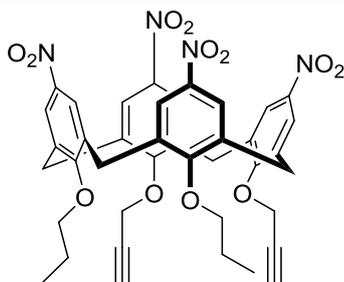
5, 11, 17, 23-tetranitro 25, 26-dipropoxy-27, 28-propargylcalixarene (40)

Glacial acetic acid (25 mL, 436 mmol) and nitric acid (98%, 25 mL, 600 mmol) were added to a stirred solution of **39** (3.77g, 4.66 mmol) in 100 mL DCM. The color of the reaction mixture turned from yellow to dark blue, and then orange. 30 minutes after the color of the reaction mixture turned orange, the reaction was quenched with addition of ice-cold water (100 mL) and saturated NaHCO$_3$ solution (100 mL). The reaction mixture was extracted with DCM and then the organic phase was washed with water (2 × 100 mL) and brine (100 mL). After drying over MgSO$_4$, the organic phase was evaporated under reduced pressure and then purified by column chromatography (eluent : 100% DCM) to yield yellow crystalline solid **40** (1.4 g, 39 %) **^{1}H NMR** (400MHz, CDCl$_3$) δ 7.63 (s, 4H, Ar<u>H</u>), 7.59(s, 4H, Ar<u>H</u>), 4.80(d, J = 2.5 Hz, 4H, OC<u>H</u>$_2$CCH), 4.67(d, J = 14.0 Hz, 1H, ArC<u>H</u>$_2$Ar), 4.61(d, J = 14.0 Hz, 2H, ArC<u>H</u>$_2$Ar), 4.55(d, J = 14.0 Hz, 1H, ArC<u>H</u>$_2$Ar), 3.99 (m, 4H, OC<u>H</u>$_2$CH$_2$CH$_3$), 3.44(m, 4H, ArC<u>H</u>$_2$Ar), 2.57 (t, J = 2.5 Hz, 2 H, OCH$_2$CC<u>H</u>), 1.95 (m, 4H, OCH$_2$C<u>H</u>$_2$CH$_3$), 1.04(t, 6H, OCH$_2$CH$_2$C<u>H</u>$_3$); ^{13}C NMR (100 MHz, CDCl$_3$) δ 161.9, 159.7, 143.8, 143.0, 136.5, 136.4, 135.5, 135.4, 124.2, 124.1, 124.0, 78.2, 78.1, 77.2, 77.0, 61.8, 32.0, 31.5, 31.0, 23.4, 10.3. **IR**: ν 3291, 2966, 2934, 2878, 1586, 1518, 1458, 1341, 1304, 1264, 1206, 1092 cm^{-1} **HRMS (ESI):** calculated for C$_{40}$H$_{36}$N$_4$O$_{12}$H [M+H$^+$]:765.2408; Found: 765.2414; Melting point: 190-192°C (decomposed).
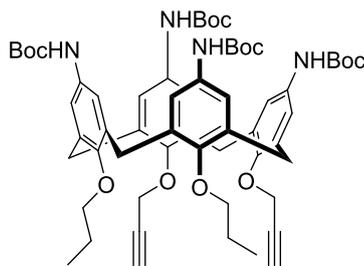
5, 11, 17, 23-Tetra-Boc-amino-25, 26-dipropoxy-27, 28-propargyloxycalix[4]arene (41)

SnCl$_2$·2H$_2$O (9.22 g, 40.86 mmol) was added to a stirred suspension of **40** (1.40 g, 1.8 mmol) in ethanol (100 mL) and heated to reflux (90 °C) for 48 hours. After cooling, the solvent was evaporated under reduced pressure. The resulting solid was washed with 10 % NaOH solution (150 mL) and DCM (50 mL). Then the aqueous phase was extracted with additional DCM (3 × 100 mL) until the precipitates vanished. Then all DCM layers were combined, dried over anhydrous MgSO$_4$ and evaporated under vacuum. The resulting green-colored product (1.01 g) was dissolved and stirred directly in DCM (50 mL). Then 4.0 g Boc anhydride (18.24 mmol) and 0.6 mL DIPEA were added into the reaction mixture. After stirring at room temperature overnight, the solvent was removed by evaporation and purified over column chromatography (eluant: DCM: ethyl acetate = 15:1) to yield an off-white solid **41** (0.862 g, 45% over two steps). **^{1}H NMR** (400MHz, CDCl$_3$) δ 6.69 (m, 8H, Ar$\underline{H}$), 6.17(s, 2H, N$\underline{H}$Boc), 6.15 (s, 2H, N$\underline{H}$Boc), 4.67 (d, J = 2.5 Hz, 4H, OC$\underline{H}_2$CCH), 4.53 (d, J = 13 Hz, 2H, ArC$\underline{H}_2$Ar), 4.44 (d, J = 13 Hz, ArC$\underline{H}_2$Ar), 4.39 (d, J = 13 Hz, ArC$\underline{H}_2$Ar), 3.80 (m, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.13 (m, 4H, ArC$\underline{H}_2$Ar), 2.39 (t, J = 2.5 Hz, 2H, OCH$_2$CC$\underline{H}$), 1.90 (sext, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.52 (br s, 9H, Boc C(CH$_3$)$_3$), 1.48 (br s, 27H, Boc C(CH$_3$)$_3$), 0.98 (t, J = 7.5 Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C NMR** (101 MHz, CDCl$_3$) δ 153.3, 153.2, 152.8, 150.8, 135.83, 135.7, 135.0, 135.0, 133.1, 132.2, 119.8, 119.7, 119.5, 119.5, 80.4, 80.0, 80.0, 77.1, 74.6, 60.7, 32.2, 31.5, 30.9, 28.4, 28.4, 23.3, 10.4; **IR**: ν 3296, 2978, 2932, 2874, 1810, 1756, 1710, 1692, 1600, 1546, 1474, 1418, 1368,

1292, 1246, 1210, 1154, 1066 cm^{-1}; **HRMS** (NSI) m/z: Calculated for C$_{60}$H$_{76}$N$_4$O$_{12}$NH$_4$ [M+NH$_4$]$^+$ 1062.5798; Found : 1062.5792. **Melting Point:** 179-181 °C

6.3.3   Synthesis of clickable DNA-interacting moieties
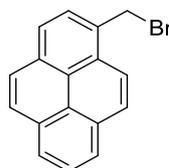
1-bromomethylpyrene (43) [209]

Phosphorus tribromide (1.75 g, 6.45 mmol) was added to a stirred solution of 1-pyrenemethanol (1.00 g, 4.30 mmol) in 5 mL of dry THF. After 30 minutes after, the reaction mixture was filtered and the solid residue was washed with diethyl ether to yield a yellow solid product **43** (932 mg, 79%). **^{1}H NMR** (400MHz, CDCl$_3$) δ 8.39 (d, J = 9 Hz, 1H, Ar<u>H</u>), 8.24 (m, 3H, Ar<u>H</u>), 8.07(m, 4H, Ar<u>H</u>), 5.27 (m, 2H, ArC<u>H</u>$_2$Br); **IR** ν 3332, 3035, 2892.5, 2836, 1601, 1588.5, 1500, 1437, 1416, 1392, 1373, 1314, 1271, 1239, 1220, 1201, 1179, 1141, 1109 cm^{-1}; **Melting point** 155-157 °C. The spectroscopic data are consistent with those reported in the literature. [209]
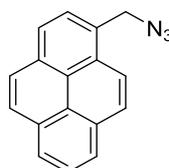
1-azidomethylpyrene (44) [209]

Sodium azide (164 mg, 2.54 mmol) was added to a stirred **43** (0.50 g, 1.69 mmol) in anhydrous DMF (3 mL) at room temperature and heated to 60 °C for 6 hours. The reaction was cooled and quenched with water. The resulting aqueous phase

was extracted with diethyl ether (300 mL). The organic phase was dried over anhydrous $MgSO_4$ and evaporated to obtain a wax like yellow-green solid **44** (387 mg, 89%). **^{1}H NMR** (400MHz, $CDCl_3$) δ 8.32-7.96 (m, 9H, Ar<u>H</u>), 5.06 (br s, 2H, ArC<u>H$_2$</u>N$_3$); **IR** 3042, 2889, 2099, 2079, 1927, 1875, 1603, 1597, 1589, 1509, 1489, 1458, 1435, 1418, 1392, 1340, 1314, 1252k, 1245, 1226, 1194, 1181, 1100 cm^{-1} **Melting point**: 65-67 °C, The spectroscopic data are consistent with those reported in the literature. [209]

5.3.4 Synthesis of clickable NRD for G-quadruplexes binding

3-azidopropan-1-amine (46) [217,350]

A solution of sodium azide (4.0 g, 61.5 mmol, 3.4eq) in H_2O (20 mL) was slowly added into a solution of 3-bromopropylamine hydrobromide **45** (4.0 g, 18.25 mmol, 1.0eq) in H_2O (15 mL). The colorless solution mixture was heated to 105°C for 18 hours. The mixture was cooled to romm temperature and water was removed *in vacuo* until white crystals crashed out. Another 4mL of water was added into the mixture to re-dissolve the crystals and the resulting solution was mixed with diethyl ether (50 mL). Potassium hydroxide (4.0 g, 71.3 mmol) was added into the biphasic mixture at 0°C. After separating the ether phase, the aqueous phase was further extracted with diethyl ether twice (2 × 30 mL). The combined organic phase was dried over $MgSO_4$, filtered and concentrated *in vacuo*, giving a colorless crude oil **46** (1.2 mL, 95%). The product was stored under Argon at -20°C. **^{1}H NMR (**400Hz, $CDCl_3$) δ 3.34(t, J = 7.0 Hz, 2H), 2.80(t, J = 7.0 Hz, 2H), 1.73 (quintet, J = 7.0 Hz, 2H) **^{13}C NMR (**100Hz, $CDCl_3$) δ 49.0, 39.0, 32.2; **IR** ν 3303, 2108 cm^{-1}, The spectroscopic data are consistent with those reported in the literature.[217,350]
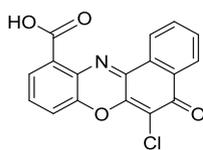
6-chloro-5-oxo-5H-benzo[a]phenoxazine-11-carboxylic acid (47) [220] This chemical was made in collaboration with Rouven Becker, a project student worked in our lab before.

2,3-dichloro-1,4-naphthoquinone (0.42 g, 1.60 mmol), 3-amino-4-hydroxy anthranilic acid (250 mg, 1.60 mmol) and potassium acetate (310 mg, 3.20 mmol) were dissolved in MeOH (20 mL) and then heated to reflux for 12 hours. The reaction mixture was cooled to 0 °C and filtered. The solid precipitate was collected and washed with water and ice cold methanol. After drying over P_2O_5, the product **47** was recovered by recrystalisation from MeOH and chloroform as a red crystalline solid (446 mg, 77%). **^{1}H NMR** (400Hz, DMSO-d_6) δ 8.58 (d, 1H, $J =$ 8.0 Hz), 8.23 (d, 1H, $J = 7.5$ Hz), 7.97 (t, 1H, $J = 7.5$ Hz), 7.90 (t, 1H, $J = 7.5$ Hz), 7.73 − 7.64 (m, 3H). **IR**: ν 2916, 1729, 1643, 1601, 1586, 1561, 1479, 1458, 1441, 1364, 1331, 1300, 1282, 1264, 1183, 1137 cm^{-1} ; **Melting Point**: 274-276°C. The spectroscopic data are consistent with that reported in the literature.[220]
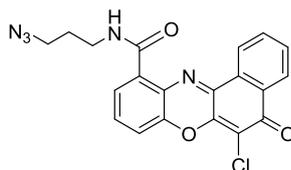
N-(3-azidopropyl)-6-chloro-5-oxo-5H-benzo[a]phenoxazine-11-carboxamide (48)[220]

Pre-dried starting material **47** (0.50 g, 1.535 mmol) was suspended in anhydrous chloroform (30 mL) under argon and cooled to 0°C. 1-chloro-N, N-2-trimethyl-1-propenyl amine (Ghosez's reagent, 800 µL, 4.605 mmol) was added dropwise into the system with syringe. After the suspended material fully dissolved in the

chloroform, triethylamine (2.1 mL, 15.4 mmol) was added, followed by azide linker **46** (420 μL, 4.2 mmol)**.** After stirring overnight, a red precipitate crashed out and additional 30mL of DCM was added to the reaction mixture. The resulting red organic phase was washed with water three times (2 × 40 mL) and the aqueous phase was extracted with DCM three times (3 × 15 mL). The organic phase was dried over $MgSO_4$ and the solvent evaporated. Column chromatography(eluent: DCM: methanol = 9:1) gave **48** (0.613 g, 97.2% yield) as an orange solid. **¹H NMR** (CDCl₃, 400MHz) δ_H ppm= 9.53 (bs, 1H), 8.42(td, 2H, J_1= 7.0 Hz, J_2= 1.5 Hz), 8.34 (dd, 1H, J_1= 7.5 Hz, J_2= 2.0 Hz), 7.90(m, 2H, J_1= 7.5 Hz, J_2= 2.0 Hz), 7.66(t, 1H, J = 6.5 Hz), 7.60(dd, 1H, J_1=8.0 Hz, J_2=1.5 Hz), 3.76(dd, 2H, J_1= 6.5 Hz, J_2= 6.0 Hz), 3.52(t, 2H, J = 6.5 Hz), 2.10(m, 2H); **¹³C NMR & ¹³C-DEPT (CDCl₃, 100MHz)** δ_C ppm 177.0(C), 164.5 (C), 146.5(C), 146.2 (C), 143.6(C), 133.2 (C), 133.0 (CH), 131.7 (CH), 131.6 (CH), 130.5 (C), 129.6(C), 129.3 (C), 128.6(CH), 127.3(CH), 124.3 (CH), 119.4 (CH), 115.4 (C), 49.1(C), 37.3 (C), 28.8 (C); **IR**: ν 3326, 2929, 2873, 2107, 1633, 1597, 1585, 1560, 1547, 1458, 1349, 1301, 1242, 1137, 1016, 1028, 1016 cm⁻¹; **HRMS** (ESI): calculated for $C_{20}H_{14}ClN_5O_3H$ ([M+H]⁺): 408.0858. Found: 408.0864; **Melting point:** 239-243°C

6.3.4   Copper-catalysed alkyene-azide cycloaddition (CuAAC) & Boc group deprotection

General procedures for CUAAC (GP 1)

CuAAC reactions were carried out mainly *via* $CuSO_4 \cdot H_2O$ with sodium ascorbate. 100 mg around Boc-protected amino calix[4]arene platform was dissolved in 5 mL DMF (5 mL). After added 1-azidomethylpyrene, $CuSO_4 \cdot 5H_2O$ (20 mg, 0.8 mmol) and sodium ascorbate (120 mg, 0.61 mmol), or CuI (16.7 mg, 0.10 mmol) was also added into the reaction mixture. Then the reaction heated in a pre-heated oil bath (90°C), for 3 hours. The reaction was monitored with thin layer chromatography and later quenched by withdrawing the heat. After cooling to 50°C, icy cold water

was added to precipitate the crude product. The mixture was filtered and washed with water (50 mL). The precipitate was dissolved in DCM (30 mL) and after being washed with water (2 × 15 mL) and brine (15 mL), the organic phase was dried over MgSO₄, evaporated under reduced pressure. The crude product was purified by column chromatography (eluent: DCM/acetone) to yield the pure product.

General procedures for Boc deprotection (GP 2)

Then the Boc-protected product was dissolved and stirring in 10 mL DCM. The DCM solution was bubbled with HCl gas. 15 minutes later, methanol was added into the reaction flask to dissolve the resulting precipitate. After 3 hours, the solvent was evaporated under reduced pressure to obtain the crude product which was purified *via* semi-prep HPLC purification or recrystallization in diethyl ether and methanol.
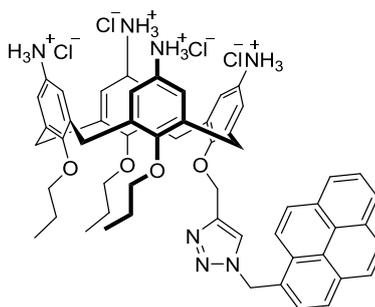
Mono-pyrene substituted calixarene (28) [197]

Using **GP 1**, **33** (102.3 mg, 0.097mmol), azido-pyrene **44** (34.4 mg, 0.133mmol), CuSO₄ 5H₂O (20 mg, 0.8 mmol). Brown-colored Boc-protected calixarene conjugate **49** (61.6 mg) was recovered from column chromatography (eluent: DCM: acetone = 80:20). After deprotecting the Boc group (**GP 2**), 35.2 mg product **28** was recovered as a brown colored salt from preparative RP-HPLC (gradient: starting from Solvent A : Solvent B = 30 : 70 to 90% Solvent B over 15 min; R_t = 3.1 min; 35.2 mg, 33% over two steps) **¹H NMR** (400MHz, CD₃OD) δ 8.40-8.08 (m, 9H, Ar<u>H</u>pyrene), 7.79 (s, 1H, Ar<u>H</u>triazole), 6.75-6.55 (m, 8H, Ar<u>H</u>), 6.38 (s, 2H,

Ar$_{triazole}$CH$_2$Ar$_{pyrene}$), 5.07 (s, 2H, OCH$_2$Ar$_{triazole}$), 4.37 (d, J = 13.5 Hz, 2H, ArCH$_2$Ar), 4.12 (d, J = 13.5 Hz, 2H, ArCH$_2$Ar), 3.81 (t, J = 7.5 Hz, 2H, OCH$_2$CH$_2$CH$_3$), 3.49 (m, 4H, OCH$_2$CH$_2$CH$_3$), 3.20 (d, J = 13.5 Hz, 2H, ArCH$_2$Ar), 2.93 (d, J = 13.5 Hz, 2H, ArCH$_2$Ar), 1.83 (sext, J = 7.5 Hz, 2H, OCH$_2$CH$_2$CH$_3$), 1.48 (sext, J = 7.5 Hz, 2H, OCH$_2$CH$_2$CH$_3$), 0.94 (t, J = 7.5 Hz, 3H, OCH$_2$CH$_2$CH$_3$), 0.52 (t, J = 7.5 Hz, 6H, OCH$_2$CH$_2$CH$_3$) **^{13}C NMR** (400Hz, CD$_3$OD), 157.77, 157.59, 155.66, 144.79, 138.0749, 137.67, 137.63, 137.5146, 133.5073, 132.59, 131.89, 130.44, 129.95, 129.32, 129.08, 129.04, 128.31, 127.57, 127.46, 127.02, 127.00, 126.75, 126.55, 126.53, 126.20, 126.14,125.73, 125.59, 123.71, 123.62, 123.56, 123.25, 78.37, 78.08, 67.00, 53.08, 31.84, 31.44, 24.20, 23.76, 10.51, 10.13 **HRMS** (ESI) m/z: [M+H]$^{+}$ Calcd for C$_{57}$H$_{60}$N$_7$O$_4$ 906.4701 ; Found 906.4700. **Melting point:** 242-244 °C (decomposed) **IR** ν 2931, 2631, 2101, 1702, 1667, 1556, 1471, 1418, 1386, 1311, 1261, 1200, 1131, 1062, 1038 cm^{-1}. The spectroscopic data are consistent with that reported in the literature.[197]
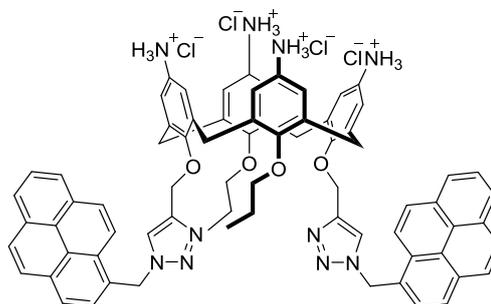
1,3-dipyrene substituted calixarene (**54**)

Using **GP 1**, Boc-protected calix[4]arene **37** (104.1 mg, 0.10 mmol), azido-pyrene **44** (59.6mg, 0.23 mmol) CuSO$_4$•5H$_2$O (27 mg, 1.08 mmol) and sodium ascorbate (128 mg, 0.65 mmol) were used. After the CuAAC and purification *via* column chromatography (eluent: DCM:acetone = 9:1), the product **50** was recovered as a brown-colored solid (92 mg). After being deprotected with HCl gas (**GP 2**), the final compound **54** was purified *via* precipitation from ice-cold methanol and diethyl ether (56 mg, 43% over two steps). **^{1}H NMR** (400MHz, CD$_3$OD) δ 8.40-8.00 (m, 18H, ArH$_{pyrene}$), 7.83 (s, 2H, ArH$_{traizole}$), 6.75 (s, 4H, ArH), 6.51 (s, 4H,

ArH), 6.24 (s, 4H, Ar$_{triazole}$CH$_2$Ar$_{pyrene}$), 5.05 (s, 4H, OCH$_2$Ar$_{triazole}$), 4.00 (d, J = 13.5 Hz, 4H, ArCH$_2$Ar), 3.10 (t, d, J = 7.5 Hz, 4H, OCH$_2$CH$_2$CH$_3$) , 2.87 (d, J = 13.5 Hz, 4H, ArCH$_2$Ar), 1.08 (sext, J = 7.5 Hz, 4H, OCH$_2$CH$_2$CH$_3$), 0.17 (t, J = 7.5 Hz, 6H, OCH$_2$CH$_2$CH$_3$); **¹³C NMR** (400MHz, CD$_3$OD) 157.43, 156.25, 144.84, 138.41, 136.87, 134.68, 133.40, 132.56, 131.88, 130.38, 129.87, 129.28, 129.11, 129.06, 128.34, 127.56, 127.00, 126.77, 126.72, 126.14, 126.11, 125.96, 125.58, 124.30, 123.75, 123.34, 77.82 ,67.01, 52.74, 31.55, 23.31, 9.87 **HRMS** (ESI) m/z: [M+H]$^+$ Calcd for C$_{74}$H$_{67}$N$_{10}$O$_4$ 1159.5347; Found: 1159.5342 **Melting point**: 221-223 °C (decomposed)   **IR** ν 3376, 2869, 2592, 1982, 1696, 1588, 1525, 1470, 1388, 1309, 1281, 1147, 1130, 1049 cm^{-1}
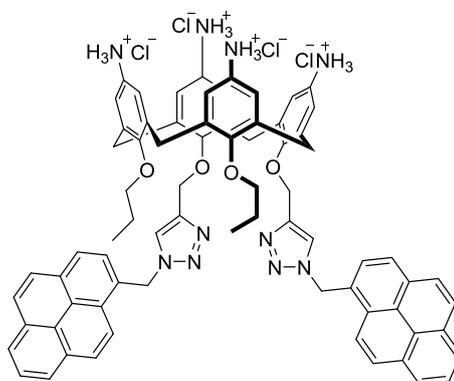
1,2-dipyrene substituted calixarene (55)

Using **GP 1**, Boc-protected calix[4]arene **41** (103 mg, 0.098 mmol), azide-functionalised pyrene **44** (59 mg, 0.23 mmol), CuSO$_4$·5H$_2$O (15 mg, 0.6 mmol) and sodium ascorbate (123 mg, 0.63 mmol) were used. After the reaction, the boc-protected product **51** was recovered as a brown-colored solid (95 mg) *via* column chromatography (eluent: DCM: acetone = 85:15). After Boc deprotection (**GP 2**), the final product **55** as a brown-colored solid was obtained with preparative RP-HPLC. (gradient: starting from Solvent A: Solvent B = 15: 85 to 90% MeOH over 15 min, R$_t$ = 4.2 min, 14 mg, 11 % over two steps) **¹H NMR** (400MHz, CD$_3$OD) δ 8.10-7.92 (m, 9H, ArH$_{pyrene}$), 7.91 (m, 9H, ArH$_{pyrene}$), 7.62 (s, 2H, ArH$_{triazole}$), 6.56 (d, J = 2.5 Hz, 2H, ArH), 6.51 (m, 4H, ArH), 6.37 (d, J = 2.5 Hz, 2H, ArH), 6.06

(m, J = 15 Hz, 4H, Ar$_{pyrene}$CH$_2$Ar$_{triazole}$), 4.50 (q, J = 15 Hz, 4H, OC$\underline{H}_2$Ar$_{triazole}$), 4.23 (d, J = 13.5 Hz, 1H, ArC$\underline{H}_2$Ar), 4.08 (d, J = 13.5 Hz, 2H, ArC$\underline{H}_2$Ar), 3.57 (d, J = 13.5 Hz, 1H, ArC$\underline{H}_2$Ar), 3.45 (t, J = 7.5 Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.09 (d, J = 13.5 Hz, 1H, ArC$\underline{H}_2$Ar), 2.90 (d, J = 13.5 Hz, 2H, ArC$\underline{H}_2$Ar), 2.41 (d, J = 13.5 Hz, 1H, ArC$\underline{H}_2$Ar), 1.44 (sext, d, J = 7.5 Hz, 4H, ArC$\underline{H}_2$Ar), 0.45 (t, J = 7.5 Hz, 6H, ArC$\underline{H}_2$Ar) **^{13}C NMR** (100MHz, CD$_3$OD) 157.60, 155.40, 144.50, 138.07, 137.51, 133.20, 132.38, 131.68, 130.13, 129.64, 129.08, 129.06, 128.75, 128.10, 127.62, 127.60, 127.40, 126.86, 126.69, 126.67, 126.61, 125.95, 125.92, 125.85, 125.35, 123.65, 123.57, 123.52, 123.23, 123.11, 78.36, 66.61, 52.99, 32.01, 31.96, 31.43, 24.10, 10.36   **HRMS** (ESI) m/z: [M+H]$^+$ Calcd for C$_{74}$H$_{67}$N$_{10}$O$_4$ 1159. 5347; Found: 1159.53 **Melting point**: 134-136°C (decomposed). **IR** ν 2923, 1701, 1673, 1652, 1650, 1506, 1470, 1363, 1201, 1130 cm^{-1}.
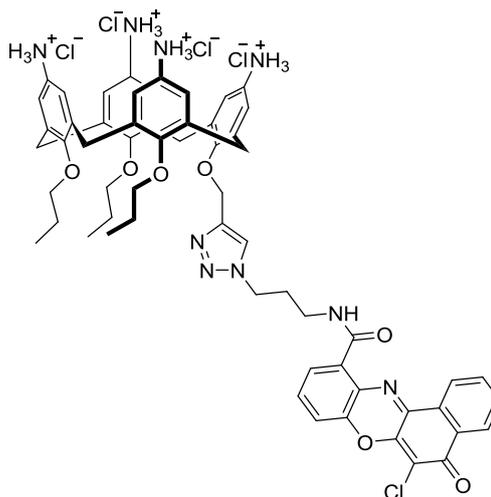
Mono-NRD substituted calixarene (56)

Using **GP 1**, Boc-protected calix[4]arene **33** (101.7 mg, 0.096 mmol), azide-functionlised NRD **48** (46.2 mg, 0.113 mmol), CuSO$_4$·5H$_2$O (27 mg, 1.08 mmol), sodium ascorbate (131.2 mg, 0.672 mmol) were used in the reaction. The boc protected product **52** (56 mg) was purified *via* column chromatography (eluent: DCM: Acetone = 95:5 to 93:7). After de-protection *via* HCl gas (**GP 2**), the final product **56** was recovered as a red-colored solid by precipitation from ice-cold

methanol and diethyl ether (13.1 mg, 11% over two steps). **^{1}H NMR** (400MHz, D$_2$O) δ 8.20 (s, Ar$\underline{H}$$_{NRD}$, 1H), 8.00- 7.25 (m, Ar$\underline{H}$$_{NRD}$+ Ar$\underline{H}$$_{triazole}$, 7H), 6.93 (s, Ar$\underline{H}$, 2H), 6.90 (s, Ar$\underline{H}$, 2H), 6.52 (s, Ar$\underline{H}$, 2H), 6.43 (s, Ar$\underline{H}$, 2H), 4.88(s, OC$\underline{H}_2ArH_{triazole}$, 2H), 4.69 (br s, Ar$_{triazole}C\underline{H}_2CH_2CH_3$NHCO, 2H), 4.26 (d, J = 13.5 Hz, ArC$\underline{H}_2$Ar, 2H), 4.13 (d, J = 13.5 Hz, ArC$\underline{H}_2$Ar, 2H), 3.70-3.50 (m, OCH$_2$CH$_2$CH$_3$+ Ar$_{triazole}$C$\underline{H}_2$CH$_2$CH$_3$NH CO, 8H), 3.31 (d, J = 13.5 Hz, ArC$\underline{H}_2$Ar, 2H), 3.10 (d, J = 13.5 Hz, ArC$\underline{H}_2$Ar, 2H), 2.41 (br s, C$\underline{H}_2$CH$_2$NHCO, 2H), 1.52 (m, OCH$_2$C$\underline{H}_2$CH$_3$, 6H), 0.56 (m, OCH$_2$CH$_2$C$\underline{H}_3$, 9H) **^{13}C NMR** (100MHz, DMSO-d$_6$) 177.10, 165.71, 156.31, 155.76, 154.38, 147.48, 146.45, 143.57, 143.03, 136.27, 136.17, 135.68, 135.42, 133.48, 133.16, 132.33, 131.34, 130.21, 129.78, 126.50, 126.28, 126.16, 126.14, 126.01, 125.99, 125.96, 125.50, 125.17, 123.26, 123.00 122.93, 118.32, 113.23, 76.44, 76.32, 66.17, 47.02, 36.27, 30.16, 29.78, 22.56, 22.21, 10.06, 9.64 **HRMS** (ESI) m/z: Calcd for C$_{60}$H$_{63}$ClN$_7$O$_9$ [M+H]$^+$ 1056.4539; Found : 1056.4536 **Melting point:** 227-229 °C (decomposed) **IR** ν 3381, 3872, 2600, 1644, 1587, 1561, 1464, 1386, 1301, 1262, 1218, 1140, 1044, 1028, 1000 cm^{-1}.

1,3-diNRD substituted calixarene (**57**)

Using **GP 1**, Boc-protected calix[4]arene **37** (100.0 mg, 0.096 mmol), azide-functionalised NRD **48** (93.34 mg, 0.23 mmol), CuSO$_4$·5H$_2$O (12.5 mg, 0.5 mmol)

and sodium ascorbate (120 mg, 0.61 mmol) were used. The boc-protected product **53** (51 mg) was obtained *via* column chromatography (eluent: DCM:Acetone = 9:1 to 5:1). After deprotection and purification *via* precipitation from ice-cold methanol and diethyl ether, 14.5 mg product **57** was recovered as a red crystalline solid (9.4% yield over two steps). **^{1}H NMR** (400MHz, CD$_3$OD) δ 8.37 (s, Ar$\underline{H}$$_{triazole}$, 2H), 8.20 (d, J = 8Hz, Ar$\underline{H}$$_{NRD}$, 2H), 7.88 (d, J = 8Hz, Ar$\underline{H}$$_{NRD}$, 2H), 7.80 (dd, J_1 = 8Hz, J_2 = 1.5 Hz, Ar$\underline{H}$$_{NRD}$, 2H), 7.75 (td, J_1 = 8Hz, J_2 = 1.5 Hz, Ar$\underline{H}$$_{NRD}$, 2H), 7.58 (m, Ar$\underline{H}$$_{NRD}$, 4H), 7.44 (dd, J_1 = 8Hz, J_2 = 1.5 Hz, Ar$\underline{H}$$_{NRD}$, 2H), 7.01 (s, Ar$\underline{H}$, 4H), 6.70 (s, Ar$\underline{H}$, 4H), 5.34 (s, OC$\underline{H}$$_2Ar_{triazole}$, 4H), 4.63 (t, J = 7 Hz, Ar$_{triazole}$C$\underline{H}$$_2CH_2CH_2$NHCO, 4H), 4.48 (d, J = 13.5 Hz, ArC$\underline{H}$$_2$Ar, 4H), 3.78 (t, J = 7.5 Hz, OC$\underline{H}$$_2CH_2CH_3$, 4H), 3.51 (t, J = 7 Hz, Ar$_{triazole}$CH$_2$CH$_2$C$\underline{H}$$_2$NHCO, 4H), 3.32 (d, J = 13.5 Hz, ArC$\underline{H}$$_2$Ar, 4H), 2.32 (quin, J = 7 Hz, Ar$_{triazole}$CH$_2$C$\underline{H}$$_2CH_2$NHCO, 4H), 1.75 (sext, OCH$_2$C$\underline{H}$$_2CH_3$, 4H), 0.88 (t, J = 7.5 Hz, OCH$_2$CH$_2$C$\underline{H}$$_3$, 6H) **^{13}C NMR** (100MHz, DMSO-d$_6$) 176.39, 165.00, 155.77, 154.20, 146.67, 145.72, 142.94, 142.80, 136.07, 135.49, 134.39, 132.92, 132.65, 131.81, 130.70, 129.46, 129.15, 125.93, 124.90, 124.84, 122.56, 122.52, 122.48, 122.44, 117.94, 122.87, 76.40, 66.22, 47.02, 36.27, 30.11, 30.21, 22.04, 9.63. **HRMS** (ESI) m/z: Calcd for C$_{80}$H$_{73}$Cl$_2$N$_{14}$O$_{10}$ [M+H]$^+$ 1459.5011; Found : 1459.5006 **Melting point**: 178 °C (decomposed) **IR** ν 3390, 2875, 1652, 1645, 1634, 1586, 1557, 1471, 1435, 1301.51, 1262, 1214, 1140, 1018 cm^{-1}.

6.3.5   FRET titration

In a typical FRET titration, the specified FRET labelled oligonucleotide (200 μL, 100 nM) in the corresponding buffer were made and used without annealing. The FAM fluorophore was excited at 490 nm and emission was measured from 500 nm to 650 nm. The scanning speed was set to 500 nm per minute and slit width was adjusted according to the relative fluorescence for the control sample. Small aliquots of ligands were prepared in buffer and titrated into the cuvette *via* pipette. The sample was mixed, and after an incubation time of 5 minutes, the fluorescent

spectra of the sample was measured. For each aliquot, the sample was scanned 3 times and the spectra were the average of the scans.

6.3.6   Other biophysical tests

Other biophysical tests, including FRET melting, circular dichroism and fluorescent intercalator displacement on i-motifs were performed as described in Section 6.2.

6.4   Experimental for Chapter 4

6.4.1   Synthesis of Fmoc-Ser(*OPrp*)-OH

The synthesis of propargylated Fmoc-Ser-OH was following the procedure published previously.[321,322]

Boc-Ser(O*Prp*)-OH (85) [293,294]

Boc-protected serine **84** (600 mg, 2.90 mmol) was stirred with sodium hydride (246 mg, 10.0 mmol) in DMF (5 mL) before the addition propargyl bromide (350 µL, 80% w/w toluene solution, 3.1 mmol) at 0°C. The resulting mixture was stirred at room temperature for another 3 hours before being quenched with addition of 5 mL water. The resulting solution was dried *in vacuo* to give a brown-colored solid. The solid was dissolved in water and EtOAc (1:1); the organic layer was separated and the aqueous layer acidified with 2 M HCl solution at 0°C. When the pH value of the aqueous layer was 2, the aqueous layer was extracted with EtOAc (3 × 20 mL). The combined organic phases was then dried over Na_2SO_4 and the solvent was removed *in vacuo* to afford a viscous dark-brown oil as a crude product. The crude

product was then purified *via* column chromatography (eluent Hexane: EtOAc: Acetic acid = 27:23:0.5) to afford Boc-Ser(OPrp)-OH as a yellow-colored viscous oil (279 mg, 43% yield). The product (**85**) contained a small amount of acetic acid (≤ 5%) and was used in the next step without further purification. **¹H NMR** (400 MHz, CDCl₃) δ 5.38 (d, 1H, *J* = 8.0 Hz), 4.53 – 4.44 (m, 1H), 4.18 (t, 1H, *J* = 2.5 Hz), 3.99 (dd, 1H, *J*₁ = 9.0, *J*₂ = 2.5 Hz), 3.81 (dd, 1H, *J*₁ = 9.0, *J*₂ = 2.5 Hz), 2.47 (t, 1H, *J* = 2.5 Hz), 1.46 (s, 5H). The ¹H NMR spectroscopic data are consistent with that reported in the literature.[321,322]

Fmoc-Ser(O*Prp*)-OH (86) [293,294]

Boc-Ser(O*Prp*)-OH **85** (279 mg, 1.15 mmol) was dissolved in DCM (5 mL), stirring at room temperature. Then TFA (5 mL) was added into the reaction, followed by water (100 mL). After being stirred for two hours, the solvent and TFA in the reaction mixture was evaporated in *vacuo* to afford the de-protected amino acid as a crystalline TFA salt. The salt was then re-dissolved in a mixture of water and acetone (2.5 mL of acetone and 2.5 mL of water). Then sodium carbonate (300 mg, 2.83 mmol) and Fmoc-*N*-hydroxysuccinimide ester (470 mg, 1.40 mmol) was added into the mixture and stirred for 4 hours. The solvent was removed *in vacuo* and the resulting solid residue dissolved in 1 M HCl solution with the pH adjusted (pH = 3). The mixture was then washed and extracted with EtOAc (3 × 20 mL) and the organic layer was dried over MgSO₄. After the solvent was removed *in vacuo* the resulintg crude product was purified *via* column chromatography (eluent: hexane:EtOAc = 40:60) to afford Fmoc-Ser(*OPrp*)-OH (**86**) as a white colored solid (258 mg, 61%). **¹H NMR** (400 MHz, CDCl₃) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.65-7.58 (m, 2H), 7.40 (t, , 2H, *J* = 7.5 Hz), 7.32 (t, 2H, *J* = 7.5 Hz), 5.64 (d, 1H, *J* =

8.5 Hz), 4.62-4.56 (m, 1H), 4.49 – 4.36 (m, 2H), 4.28-4.17 (m, 2H), 4.06 (dd, 1H, J = 9.0, 3.0 Hz), 3.84 (dd, 1H, J = 9.0, 3.0 Hz), 2.46 (t, J = 2.5 Hz, 1H); **IR** v 3403, 3259, 3201, 2981, 2387, 2376, 1759, 1720, 1704, 1678, 1651, 1522, 1449, 1438, 1398, 1378, 1355, 1318, 1296, 1261, 1226, 1192, 1170, 1107 cm^{-1}; **Melting point**: 126-128 °C. The spectroscopic data are consistent with that reported in the literature. [321,322]

6.4.2 Solid Phase peptide Synthesis (SPPS)

The synthesis of the linear RGD peptide analogue was performed following the procedure by Liu *et. al.* and Kessler *et. al.*[324-326] The preparation of the linear peptide started with incorporating Fmoc-Gly-OH onto 2-chlorotitryl chloride resin. The loading of Fmoc-Gly-OH on the resin was monitored *via* UV spectroscopy. Then other amino acid residues were added on the peptide chain in standard SPPS coupling methods. Each deprotection and coupling was monitored using the Kaiser test. The final cleavage was conducted *via* treating the peptide-bearing resin with a mixture of trifluoroethanol and acetic acid.

Coupling Fmoc-Gly-OH onto 2-chlorotitryl chloride resin

2-chlorotitryl chloride resin (400 mg) in a peptide synthesis column was sealed under anhydrous conditions. Anhydrous DCM (5 mL) was added into the column *via* syringe and the resin was shaken in anhydrous DCM for 1.5 hours. Fmoc-Gly-OH (400 mg) and DIPEA (350 μL) were added into the column to react with the resin. The column was shaken for 2.5 hours and then the solution was filtered. Then methanol (5 mL) and DIPEA (350 μL) were added into the column to cap unreacted binding sites on the resin. After being shaken for another 2.5 hours, the solution was filtered and the resin on the column was washed with DMF three times, DCM two times and diethyl ether two times. The resin was dried under vaccum for evaluation of Fmoc-Gly-OH loading.

Evaluation of level of the first attachment

Loaded resin (3 × 1 mg) were each weighted into 3 × 10 mm matched silica UV cuvettes. Freshly prepared 20% piperidine in DMF (3 mL) was dispensed into each of the cells. After 1.5 hours, the absorbance at 290 nm was recorded on a UV spectrometer, using 20% piperidine in DMF as a control which was subtracted off the samples. The average loading of amino acid was calculated (Equation 5.3) as the mean value of three samples.[326]

Equation 5.3 $$loading\ (mmol \cdot g^{-1}) = \frac{(Abs_{sample})}{(mg\ of\ sample \times 1.75)}$$

The average loading was 0.687 ± 0.1 mmol/g resin, this equates to 0.275 ± 0.04 mmol Fmoc-Glycin loaded on the resin.

Kaiser test

'Kaiser Solution A' was prepared by dissolving phenol (8 g) in absolute ethanol (2 mL) and then mixing it with KCN solution in pyridine (100µL, 0.0004 mM). 'Kaiser solution B' was prepared using ninhydrin dye (1 g) in absolute ethanol (20 mL). In a typical Kaiser test, small amount of resin (≤1 mg) was mixed with a few drops of Kaiser 'A' and 'B' solutions in a small vial. Then the vial was heated on a heating block for 2 min. The presence of free amino group on the peptide was indicated as a change in the solution color from yellow to dark blue. If the color of the solution did not change, this indicated the absence of free amino group on the peptide.

Peptide synthesis

The Fmoc-protecting group was removed by shaking with 20% piperidine in DMF three times. Then the resin was washed with DMF (5 × 20 mL) and DCM (2 × 20 mL). Then Kaiser test was applied to qualitatively assess the presence of free amine group on the peptide chain. The resin was then shaken in a coupling cocktail

containing Fmoc-Arg(Pbf)-OH acid (454 mg, 0.7 mmol, 2.5 eq), HOBt (94.6 mg, 0.7 mmol, 2.5 eq), HBTU (256.47 mg, 0.7 mmol, 2.5 eq) and DIPEA (244 μL, 180.94 mg, 1.40 mmol, 5 eq) in 5 mL DMF for 2 hours. At the end of the coupling, the resin was filtered and washed with DMF 5 times before performing the Kaiser test to verify removal of the Fmoc. Other amino acids were coupled to the peptide chain using analogous procedures.

Linear precursor cleavage

The resin with the deprotected amine was shaken in a cleavage cocktail composed of DCM, 2,2,2-trifluoroethanol and acetic acid (3:1:1) over 1 hour at room temperature. The resin was filtered and washed with the same cocktail twice (2 × 10 mL). The combined cocktail washes were removed *in vacuo* to give a dark-brown solid residue. The crude solid was re-dissolved in small amount of DCM and precipitated using cold diethyl ether. The crude product **87** was collected by centrifugation as pink-orange colored precipitate (210 mg).

Prep-HPLC purification on linear RGD peptide.

linear RGDfS(*O-Prp*) (87)

87 was purified by prep-HPLC using the setting described in Table 5.3. The linear peptide **87** was obtained as a off-white TFA salt (110 mg, 38%, calculated from the

amount of Fmoc-Glycine loaded on the resin). **^{1}H-NMR** (400 MHz, MeOD) δ 7.52 – 7.39 (m, 5H), 4.83 (t, J = 8.0 Hz, 1H), 4.67 (dd, J_1 = 9.5 Hz, J_2 = 5.0 Hz, 1H), 4.56 (t, J = 4.0 Hz, 1H), 4.43 – 4.38 (m, 1H), 4.30 – 4.19 (m, 2H), 4.15 – 4.04 (m, 2H), 4.01 (dd, J_1 = 9.5 Hz, J_2 = 4.5 Hz, 1H), 3.63 (dd, J_1 = 9.5 Hz, J_2 = 4.0 Hz, 1H), 3.43 – 3.32 (m, 2H), 3.26 (dd, J_1 = 13.5, J_2 = 8.0 Hz, 1H), 3.20-3.13 (m, 3H), 3.04 (t, J = 2.5 Hz, 1H), 2.93 – 2.89 (m, 2H), 2.75 (s, 3H), 2.69 (s, 3H), 2.26 (s, 3H), 2.15-2.06 (m, 1H), 2.01 – 1.88 (m, 1H), 1.88 – 1.70 (m, 2H), 1.64 (s, 9H), 1.63 (s, 6H). **^{13}C NMR** (100 MHz, MeOD) δ 172.8, 172.7, 170.2, 169.0, 168.5, 158.5, 156.8, 138.0, 136.1, 132.8, 132.1, 128.9, 128.97, 128.2, 126.7, 124.6, 117.0, 86.3, 82.5, 78.6, 75.3, 68.3, 57.9, 55.5, 53.9, 52.5, 52.4, 49.4, 42.6, 36.7, 36.7, 35.6, 28.7, 28.7, 27.3, 26.9, 26.9, 18.2, 17.1, 11.1. **IR** ν 3309, 1660, 1650, 1539, 1369, 1201, 1138, 1094 cm^{-1}; **HRMS** (ESI): calculated for $C_{44}H_{63}N_8O_{12}S_1$ ([M-TFA]$^+$): 927.4281. Found: 927.4288; **Melting point:** 133 °C. (decomposed); **Optical rotation**: $[\alpha]_D^{25}$ = -7 (c = 0.5, MeOH)

Time	Percentage of water (%)	Percentage of methanol (%)	Flow (mL/min)
0	50	50	1
15	40	60	1
25	30	70	1
30	5	95	1
35	95	5	1
40	50	50	1

Table 5.2   *Analytical HPLC methods for linear RGDfS(O-Prp)* ***87***

R$_t$ = 23.802 min

Time	Percentage of water (%)	Percentage of methanol (%)	Flow (mL/min)
0	50	50	20
15	40	60	20
25	30	70	20
30	5	95	20
35	95	5	20
40	50	50	20

Table 5.3   Preparative HPLC methods for linear RGDfS(O-Prp) 87

R_t = 13.4 min

Linear peptide cyclisation

Cyclic **RGDfS(*O-Prp*)** (**83**)

Linear peptide **87** (70 mg, 0.062 mmol) was stirred in acetonitrile (80 mL) with HBTU (150 mg, 0.4 mmol) and DIPEA (40 µL, 0.24 mmol) at room temperature for 5 days. ([**87**] ≤ 1 mmol/L.) Additional HBTU (150 mg, 0.4 mmol), was added into the reaction mixture two days after the start of the reaction. The solvent was removed *in vacuo* and the solid residue redissolved in EtOAc (10 mL). The organic phase was washed with 0.1 M HCl solution (2 × 15 mL) and then saturated NaHCO$_3$ solution (20 mL) and brine (20 mL). The organic phase was then dried over MgSO$_4$ and then the solvent was removed *in vacuo* to afford a yellow-colored solid as the crude product amd was purified with preparative RP-HPLC to afford the pure product **83** (39 mg, 64%). The preparative and analytical RP-HPLC methods are

detailed in Table 5.4 & Table 5.5.

Time	Percentage of water (%)	Percentage of acetonitrile (%)	Flow (mL/min)
0	100	0	20
5	60	40	20
20	30	70	20
25	0	100	20
30	100	0	20

Table 5.4   Preparative HPLC method for cyclic RGDfS(O-Prp) 83

R_t = 13.4 min

Time	Percentage of water (%)	Percentage of acetonitrile (%)	Flow (mL/min)
0	95	5	1
25	5	95	1
30	95	5	1

Table 5.5   Analytical HPLC method for cyclic RGDfS(O-Prp) 83

R_t = 15.9 min, Blank acetonitrile solvent trace substracted, **Purity** 98.9%

Peak Number	Rt [min]	Peak width [min]	Area [mAU×s]	Area [%]
1	7.528	0.3324	118.6	1.1207
2	15.873	0.1416	10468.6	98.8793

IR ν 3375, 2868, 2591, 1588, 1525, 1470, 1387, 1309, 1281, 1220, 1147, 1129, 1048 cm^{-1}; **HRMS** (NSI) m/z: Calculated for $C_{44}H_{60}N_8O_{11}SNa$ [M+Na]$^+$ 931.3994; Found : 931.3983. **Melting Point:** 142 °C (decomposed); **Optical rotation**: $[\alpha]_D^{25}$ = -16.4 (c=0.5, MeCN).

6.4.3   Synthesis of coumarin dye for click reaction

The synthesis of clickable coumarin was performed using previously described procedures.[336]

3-acetamido-7-acetoxy-chromen-2-one (**90**) [336]

N-acetyl glycine (2.34 g, 20.00 mmol), 4-Hydroxysalicylaldehyde (2.76 g, 20.0 mmol), anhydrous sodium acetate (16.0 g 19.99 mmol) and acetic anhydride (50 mL) were heated to 120°C for 5 hours. Then the reaction mixture were poured into icy cold water (250 mL). The mixture was then filtered and the precipitate air-dried to afford a yellow powder **90** (1.23, 23%) **^{1}H-NMR** (400MHz, CDCl$_3$): δ 8.67 (s, 1 H), 8.04 (s, 1 H), 8.51 (d, *J* = 8 Hz, 1 H), 7.12 (d, *J* = 2.3 Hz, 1 H), 7.07 (dd, *J* = 8, 2.3 Hz, 1 H), 2.34 (s, 3 H), 2.25 (s, 3 H); **IR** 3340, 3088, 2754, 1757, 1718, 1679, 1624, 1611, 1534, 1431, 1370, 1353, 1287, 1259., 1250, 1203, 1156, 1118 cm^{-1}; **Melting point** 237 °C (decomposed). The spectroscopic data are consistent with that reported in the literature. [336]

3-Azido-7-hydroxy-chromen-2-one (**91**) [336]

3-acetamido-7-acetoxy-chromen-2-one **90** (770 mg, 2.95 mmol) was added in a solution of 37% HCl and ethanol (V$_{HCl}$: V$_{ethanol}$= 2:1, 8mL) and heated to reflux for

1 hour. Then icy cold (15 ml) water were added into the mixture to quench the reaction and then the mixture was stirred at -30 ℃ for 30 minutes. The mixture was warmed to 0 ℃ and sodium nitrite (1 g, 14.5 mmol) was added followed by the slow addition of sodium azide (1.4 g, 21.5 mmol). Then the crude, brown-colored product was filtered and purified by column chromatography (eluent: Hexane:EtOAC = 3:2) to afford **91** as red-coloured needles (486 mg, 82%). **^{1}H-NMR** (400 MHz, DMSO-D) δ 10.52 (s, 1H), 7.59 (s, 1H), 7.48 (d, J = 8.5 Hz, 1H), 6.81 (dd, J = 8.5, 2.0 Hz, 1H), 6.76 (d, J = 2.0 Hz, 1H) ; **IR** ν 3287, 2917, 2849, 2101, 1681, 1623, 1592, 1556, 1514, 1504, 1454, 1371, 1340, 1316, 1257, 1219, 1156, 1119 cm^{-1} **Melting point**: 98-100 °C. The spectroscopic data are consistent with that reported in the literature.[336]

6.4.4   Synthesis of fluorophore-tagged calix[4]arene platform for click reaction

Mono-coumarin substituted calixarene (**94**)[197]

Mono-propargylated calix[4]arene **33** (241 mg, 0.30 mmol) was stirred in DMF (15 mL) at room temperature. Azide-functionalised coumarin **91** (80 mg, 0.40 mmol), CuSO$_4$.5H$_2$O (10 mg, 0.04 mmol) and sodium ascorbate (100 mg, 0.51 mmol) were added to the mixture and heated at 85°C for 20 hours. The DMF solvent was then removed *in vacuo* and the solid was redissolved in DCM (20 mL) and washed with water (2 × 20 mL) and brine (20 mL). The organic phase was dried *in vacuo* to afford a brown-colored crude product which was further purified using column

chromatography (eluent: DCM: methanol = 99:1) to afford a yellow-colored crystalline solid **94** (235mg, 81.3%). **^{1}H-NMR** (400 MHz, CDCl$_3$): δ 8.51 (s, 1H, Ar$\underline{H}$$_{Triazole}$), 8.49 (s, 1H, Ar$\underline{H}$$_{Coumarin}$), 7.50 (d, J = 8 Hz, 1H, Ar$\underline{H}$$_{Coumarin}$), 6.89-6.84 (m, 2H, Ar$\underline{H}$$_{Coumarin}$), 6.71 (s, 4H, Ar$\underline{H}$), 6.52 (s, 4H, Ar$\underline{H}$), 6.28 (s, 2H, N$\underline{H}$Boc), 6.21 (s, 2H, N$\underline{H}$Boc), 6.13 (s, 4H, N$\underline{H}$), 5.21 (s, 2H, OC$\underline{H}_2$C), 4.37 (d, J = 13 Hz, 2H, ArC$\underline{H}_2$Ar), 4.25 (d, J = 13 Hz, 2H, ArC$\underline{H}_2$Ar), 3.85 (t, J = 8 Hz, 2H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.67 (t, J = 7 Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.07 (d, J = 13 Hz, 2H, ArC$\underline{H}_2$Ar), 3.02 (d, J = 13 Hz, 2H, ArC$\underline{H}_2$Ar), 1.90 (sextet, J = 8 Hz, 2H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.81-1.70 (m, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.50 (s, 9H, C(C$\underline{H}_3$)$_3$), 1.49 (s, 9H, C(C$\underline{H}_3$)$_3$), 1.47 (s, 18H, C(C$\underline{H}_3$)$_3$), 0.96 (t, J = 7 Hz, 3H, OCH$_2$CH$_2$C$\underline{H}_3$), 0.86 (t, J = 8 Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C-NMR** (100 MHz, CDCl$_3$): δ 162.3, 156.4, 154.8, 153.9, 153.7, 153.6, 153.1, 153.1, 151.3, 144.5, 136.1, 135.7, 135.1, 135.0, 132.0, 131.8, 120.4, 120.4, 120.2, 120.0, 119.6, 110.9, 130.2, 130.0, 80.5, 80.4, 80.4, 65.9, 36.6, 31.6, 31.6, 31.2, 28.6, 23.3, 23.2, 10.5, 10.4.; **IR** ν 2962, 2919, 2865, 1689, 1602, 1514, 1466, 1412, 1365, 1290, 1219, 1213, 1149 cm^{-1}; **HRMS (NSI)**: m/z calculated for C$_{69}$H$_{85}$N$_7$O$_{15}$ [M+H]$^+$ 1269.6421; Found : 1269.6442; **Melting point**: 168-170 °C. The spectroscopic data is consistent with the exisiting data.[197]
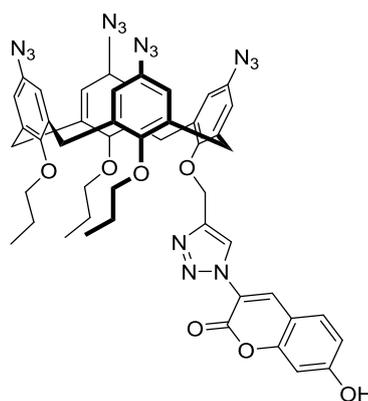
Calixarene-coumarin conjugate 92 [337]

Calixarene **94** (200 mg, 0.11 mmol) was stirred in DCM (10 mL) at room temperature. HCl gas was bubbled through this solution for 10 minutes and 20 mL

MeOH was added to dissolve the precipitated salt. After a further 20 minutes, the solvent was removed *in vacuo* to give a yellow-colored crystalline solid, **95**. After dissolving **95** in a solution of 10% HCl and methanol (10% HCl:methanol = 2:1, 30mL), the mixture was stirred at 0°C for 20 minutes, followed by the addition of sodium nitrite (150 mg, 2.2 mmol) and, after waiting for a further 30 minutes, sodium azide (390 mg, 6.0 mmol) was slowly added into the reaction mixture. After 1.5 hours, the reaction mixture was extracted with DCM (3 × 20mL) and EtOAc (3 × 20mL). The organic phases were combined, washed with brine and dried over anhydrous $MgSO_4$. The solvent was then removed *in vacuo* to affored the brown-colored solid as the crude product which was purified *via* column chromatography (eluent DCM:methanol = 100:1), **92** was recovered as a yellow-colored crystalline solid (111 mg, 73%). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 8.65 (s, 1H, Ar$\underline{H}_{Triazole}$), 8.57 (s, 1H, Ar$\underline{H}_{Coumarin}$), 7.57 (d, J = 7.5 Hz, 1H, Ar$\underline{H}_{Coumarin}$), 6.96-6.94 (m, 2H, Ar$\underline{H}_{Coumarin}$), 6.42 (s, 2H, Ar$\underline{H}_{calixarene}$), 6.34 (s, 2H, Ar$\underline{H}_{calixarene}$), 6.19 (d, J = 3.0 Hz, 2H, Ar$\underline{H}_{calixarene}$), 6.15 (d, J = 3.0 Hz, 2H, Ar$\underline{H}_{calixarene}$), 5.21 (s, 2H, OC$\underline{H}_2$C), 4.42 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 4.28 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 3.91 (t, J = 7.5 Hz, 2H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.75 (t, J = 7.5 Hz, 4H, OC$\underline{H}_2$CH$_2$CH$_3$), 3.11 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 3.04 (d, J = 14 Hz, 2H, ArC$\underline{H}_2$Ar), 1.93 (sextet, J = 7.5 Hz, 2H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.84-1.77 (m, 4H, OCH$_2$C$\underline{H}_2$CH$_3$), 1.01 (t, J = 7.5 Hz, 3H, OCH$_2$CH$_2$C$\underline{H}_3$), 0.94 (t, J = 7.5 Hz, 6H, OCH$_2$CH$_2$C$\underline{H}_3$); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 161.6, 156.3, 154.8, 154.2, 153.73, 152.0, 144.0, 137.0, 136.7, 135.8, 135.7, 134.5, 134.4, 133.7, 130.5, 124.8, 119.6, 118.9, 118.8, 118.5, 118.3, 115.0, 111.1, 103.3, 77.1, 65.6, 31.4, 31.2, 23.2, 23.2, 10.3; **IR** ν 2963, 2913, 2874, 2106, 1732, 1607, 1464, 1384, 1311, 1234, 1114 cm^{-1}; **HRMS (NSI)**: m/z calculated for $C_{49}H_{45}N_{15}O_7Na$ [M+Na]$^+$ 978.3515; Found: 978.3519; **Melting point**: 127-129 °C.
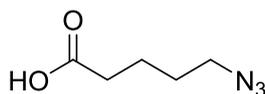
6-azido-hexanoic acid (97) [351]

6-bromohexanoic acid **96** (914 mg, 4.68 mmol) and sodium azide (1.2 g, 18.5 mmol) were mixed in DMF (4 mL) and stirred at 50 °C. Then the mixture was cooled to room temperature and DCM (30 mL) was added. The organic phase was then washed with saturated lithium chloride solution (5 × 30 mL) and brine (2 × 30 mL). The organic phase was then dried over anhydrous $MgSO_4$ and the solvent was removed *in vacuo* to afford the product **97** as a colorless oil (210 mg, 29%). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 3.28(t, J = 7.0 Hz, 2H), 2.37 (t, J = 7.0 Hz, 2H), 1.68-1.60 (m, 4H), 1.45-1.43 (m, 2H); **^{13}C-NMR**(100 MHz, $CDCl_3$) 178.14, , 50.79, 33.44, 28.12, 25.77, 23.86 ppm; **IR** ν 2936, 2867, 2092, 1708, 1646, 1415, 1387, 1253, 1153 cm^{-1}. The spectroscoptic data was in consistent with the literature.[351]
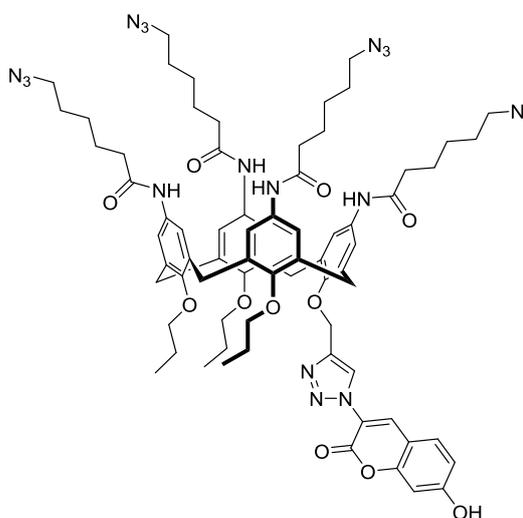
Calixarene-coumarin conjugate 93

Calixarene **94** (180 mg, 0.1 mmol) was stirred in DCM (10 mL) at room temperature. HCl gas was bubbled through this solution for 10 minutes and MeOH (20 mL) was added to dissolve the precipitated salt. After a further 20 minutes, the

solvent was evaporated to give a yellow-colored crystalline solid, **95**. This was dissolved in DMF (5 mL) with 6-azido-hexanoic acid (230 mg, 1.46 mmol), HBTU (500 mg, 1.31 mmol) and DIPEA (600 μL, 3.44 mmol) and stirred overnight at room temperature. The solvent was removed *in vacuo* to afford a brown-colored oil. After dissolving the residue oil in EtOAc (20 mL), the organic phase was washed with 1M HCl (1×10 mL) and brine (2×30 mL), then dried over anhydrous Na_2SO_4 and the organic solvent was removed *in vacuo* to afford a crude oil. Purification with column chromatography (eluent: EtOAc:Hexane = 5:6 to 2:8) gave the final product, **93**, was recovered as a yellow-colored crystalline solid (52 mg, 25%). **^{1}H-NMR** (400 MHz, MeOH): δ 8.53 (s, 1H), 8.50(s, 1H), 7.64(d, 1H, $J = 9$Hz), 6.93-6.88 (m, 9H), 6.82 (d, 1H, $J = 2$ Hz), 5.23 (s, 2H), 4.45 (d, 2H, $J = 13$ Hz), 4.37 (d, 2H, $J = 13$ Hz), 3.87 (t, 2H, $J = 7.5$ Hz), 3.80 (t, 4H, $J = 7.5$ Hz), 3.09 (d, 2H, $J = 13$ Hz), 3.07 (d, 2H, $J = 13$ Hz), 2.26 (t, 9H, $J = 7.5$ Hz), 2.03-1.98 (m, 2H), 1.91-1.85 (m, 4H), 1.66-1.61 (m, 16H), 1.45-1.38 (m, 8H), 1.05 (t, 3H, $J = 7.5$ Hz), 0.92 (t, 6H, $J = 7.5$ Hz); **^{13}C NMR** (100 MHz, MeOD) δ 172.4, 172.4, 167.94 163.0, 156.6, 155.1, 153.1, 152.9, 151.3, 144.0, 135.4, 135.4, 134.8, 134.7, 134.7, 132.9, 132.2, 132.2, 132.1, 131.0, 130.5, 128.4, 125.1, 120.6, 120.5, 120.5, 120.4, 119.2, 118.8, 114.2, 110.5, 102.0, 76.8, 76.7, 67.7, 65.8, 50.9, 38.7, 36.2, 31.5, 31.1, 30.7, 30.2, 29.3, 29.3, 29.3, 29.3, 28.7, 28.2, 26.0, 24.9, 23.5, 23.1, 22.9, 22.6, 13.0, 10.0, 9.5, 9.3; **IR** ν 2931, 2870, 2092, 1731, 1650, 1602, 1535, 1466, 1414, 1382, 1241, 1126, 1041 cm^{-1}; **HRMS (NSI)**: m/z calculated for $C_{146}H_{176}N_{38}O_{22}$ $[M-2H]^{2-}$ 1407.6942; Found : 1407.6936; **Melting point**: 128-130°C.

6.4.5   Synthesis of coumarin-tagged cRGD peptide
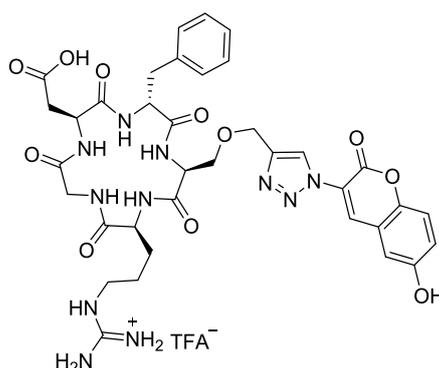
cRGD-coumarin conjugate (99)

Protected **cRGDfS(*O*-Prp) 83** (31.2 mg, 0.034 mmol), clickable coumarin dye **91** (10 mg, 0.049 mmol) and tetrakis(acetonitrile) copper (I) hexafluorophosphate (9.2 mg, 0.024 mmol) were dissolved in 2 mL DMF which had previously been degased with Argon. To this, DIPEA (50 μL, 0.34 mmol) was added. After stirring at room temperature overnight, the solvent was removed *in vacuo*. The brown-colored crude residue **98** was deprotected with a cocktail of TFA, H_2O and TIS (95:2.5:2.5, 3.2 mL). After stirring for 3 hours, the solvent in the reaction mixture were removed *in vacuo* to give a dark-brown-colored residue, which was purified *via* preparative HPLC using the method described in Table 5.4 to give the product, **99**, as the TFA salt which had the appearance of a cream colored foam (10.9 mg, 40%). **¹H-NMR** (400 MHz,DMSO-D): δ 12.23 (s, 1H), 11.04 (s, 1H), 8.60 (s, 1H), 8.52 (s, 1H), 8.21 (d, 1H, *J* = 7.0 Hz), 8.11 − 7.97 (m, 3H), 7.85 (d, *J* = 8.5 Hz, 1H), 7.76 (d, *J* = 8.5 Hz, , 1H), 7.55 (t, *J* = 5.0 Hz, 1H), 7.25-7.14 (m, 5H), 6.93 (d, *J* = 10.5 Hz, 1H), 6.87 (d, *J* = 2.0 Hz, 1H), 4.74 − 4.49 (m, 4H), 4.43 − 4.30 (m, 1H), 4.22 (m,1H), 4.14 (dd, J_1 = 16 Hz, J_2= 8.0 Hz, , 1H), 3.73 (dd, J_1 = 9.5 Hz, J_2= 5.5 Hz, 1H), 3.60 (dd, J_1 = 9.5 Hz, J_2= 3.5 Hz, 1H), 3.27 (d, *J* = 13.0 Hz, 1H), 3.10 (d, *J* = 5.5 Hz, 2H), 3.01 (dd, J_1 = 13.0 Hz, J_2= 7.0 Hz, 1H), 2.74 (dd, J_1 = 13.5, J_2= 7.0 Hz, 1H), 2.65 (dd, J_1 = 16.0, J_2= 9.0 Hz, 1H), 2.30 (dd, J_1 = 16.0, J_2= 5.0 Hz, 1H), 1.85 (m, 1H), 1.53 − 1.35 (m, 3H). **¹³C NMR** (100 MHz, DMSO-d₆) δ 171.7, 171.0, 170.7,

170.0, 169.7, 169.2, 162.6, 156.70 156.4, 154.7, 143.5, 137.6, 136.3, 131.0, 129.2, 128.0, 126.2, 125.1, 119.2, 114.4, 110.3, 102.2, 69.7, 63.3, 55.0, 53.7, 51.9, 48.8, 43.0, 40.3, 37.0, 35.4, 27.9, 25.0. **IR** ν 3265.84, 1637, 1550, 1421, 1318, 1232, 1186, 1128, 1038 cm^{-1}; **HRMS (NSI)**: m/z calculated for $C_{36}H_{41}N_{11}O_{11}H$ $[M+H]^+$ 804.3060, Found : 804.3058; **Melting point**: 144-146 °C decomposed; **Optical rotation**: $[\alpha]_D^{25}$ = -14.2 (c =0.5, DMSO).

Time	Percentage of water (%)	Percentage of methanol (%)	Flow (mL/min)
0	100	0	20
5	60	40	20
20	30	70	20
25	0	100	20
30	100	0	20

*Table 5.6   Preparaive HPLC methods for cRGD-coumarin conjugate **99***

Rt = 10.7 min

6.4.6   Synthesis of cRGD-tagged fluorescent calix[4]arene conjugate
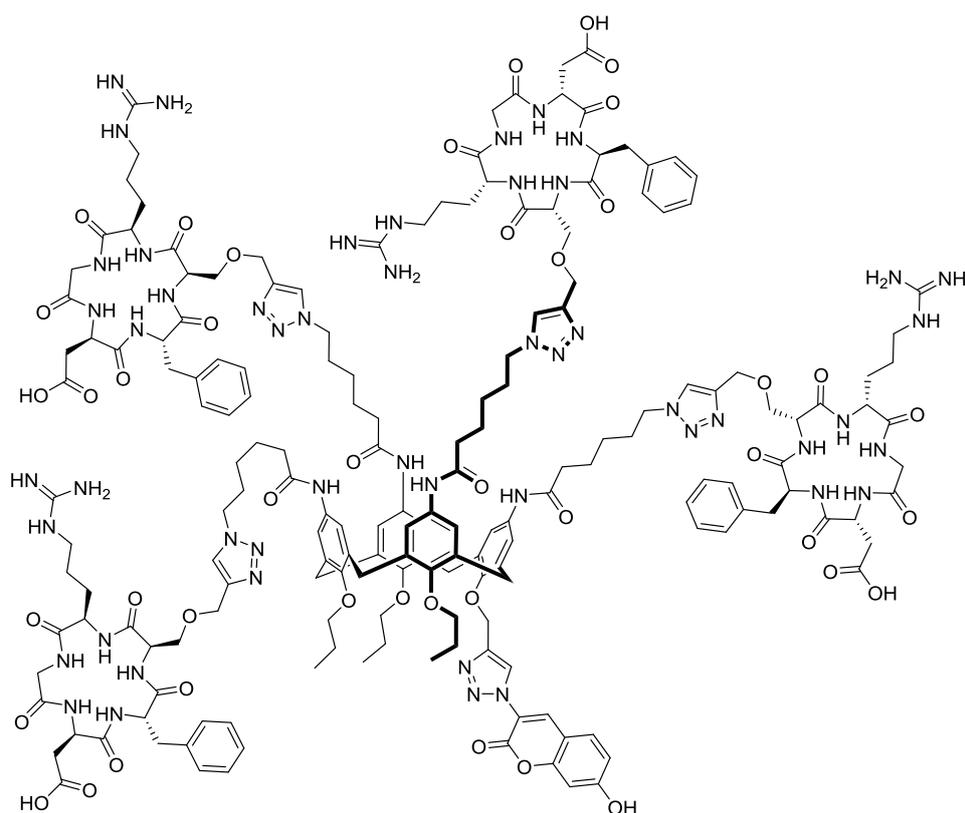
Cyclic RGD peptide-calixarene conjugate (**101**)

Calixarene **93** (13 mg, 0.0092 mmol), peptide **83** (37 mg, 0.041 mmol) and tetrakis(acetonitrile) copper (I) hexafluorophosphate (3.5 mg, 0.0092 mmol) were dissolved DMF under Argon protection at room temperature. A small amout of DIPEA (50 μL, 0.34 mmol) was added into the reaction mixture and the reaction mixture was stirred overnight at room temperature. Then the solvent was removed *in vacuo* and a cleavage cocktail consisting of TFA, H_2O and TIS (95:2.5:2.5, 2.0 mL) was added into the reaction mixture and stirred for another hour at room temperature. After evapourating all the solvent in the reaction mixture, the crude product was purified *via* preparative HPLC as described in Table 5.7 to afford the pure product **101** as a TFA salt and a pink foam (4.9 mg, 15% yield). **IR**: ν 3280, 3203, 3070, 2963, 2936, 2872, 2323, 2161,1982, 1739, 1654, 1649, 1643, 1530, 1467, 1456, 1418, 1379, 1315, 1198, 1174, 1131, 1057, 1039, 1000 cm⁻¹; **Melting**

point: 76-78 ℃; **HRMS (MALDI)**: m/z calculated for $C_{181}H_{232}N_{51}O_{43}H$ $[M+H]^+$ 3811.8, Found: 3811.9

Time	Percentage of water (%)	Percentage of methanol (%)	Flow (mL/min)
0	100	0	20
5	60	40	20
15	30	70	20
20	0	100	20

*Table 5.7   Preparative RP-HPLC methods for conjugate **101***

Rt = 8.2 min

Time	Percentage of water (%)	Percentage of methanol (%)	Flow (mL/min)
0	100	0	1
25	0	100	1
30	100	0	1

*Table 5.8   Analytical HPLC methods for conjugate **101***

Rt = 25.401 min   Blank methanol solvent trace substracted, **Purity** 96.4%
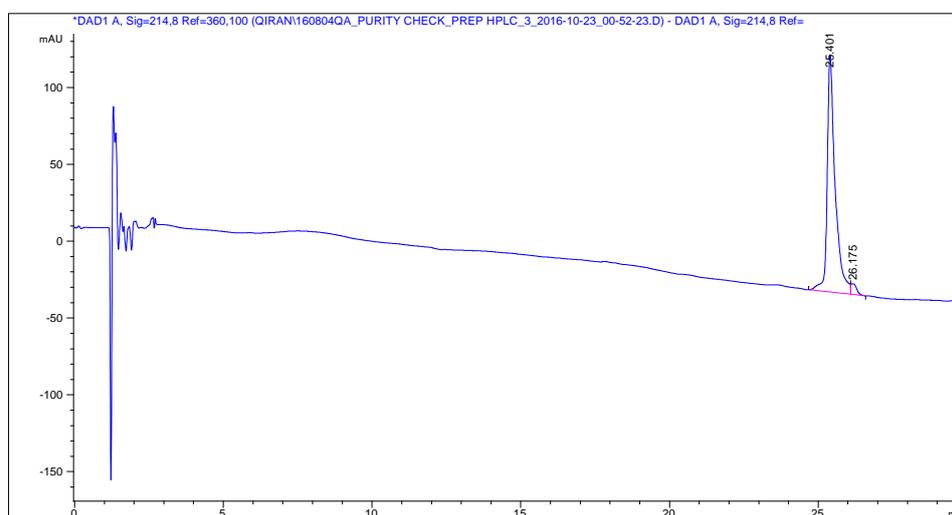
Peak Number	Rt [min]	Peak width [min]	Area [mAU × s]	Area [%]
1	25.401	0.2649	2870.57104	96.4474
2	26.175	0.1943	91.90798	3.0880
3	29.711	0.0814	13.82742	0.4646
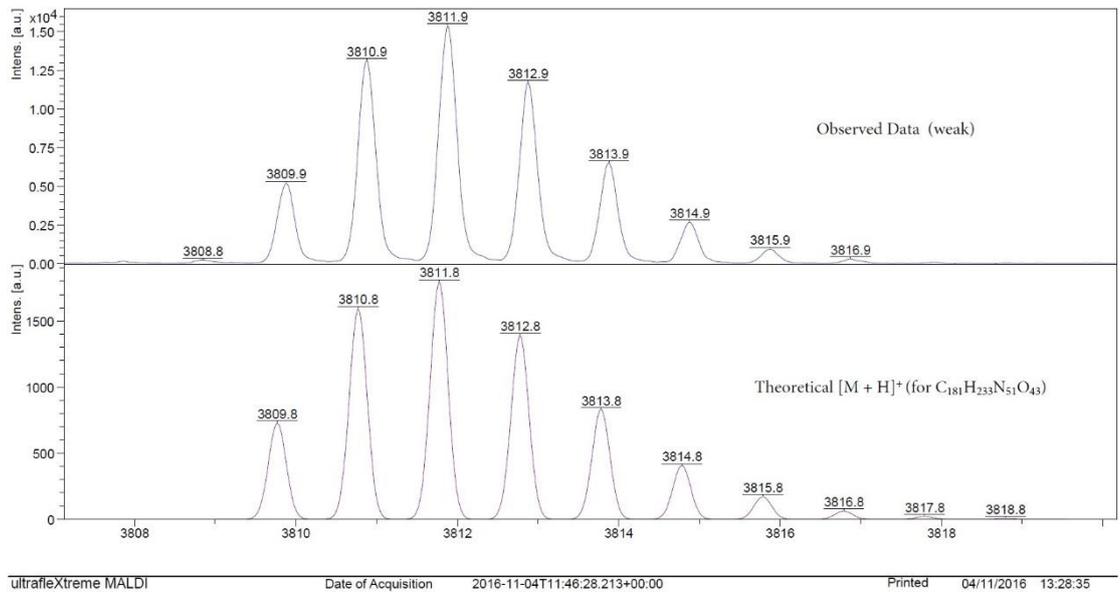
References

1   Alberts, B. *et al. Molecular Biology of the Cell, 4th edition* (Garland Science, 2002).

2   Miescher, F. Ueber die chemische Zusammensetzung der Eiterzellen. *Medicinisch-chemische Untersuchungen* **4**, 441-460 (1871).

3   Dahm, R. From discovering to understanding. Friedrich Miescher's attempts to uncover the function of DNA. *EMBO Rep.* **11**, 153-160, doi:10.1038/embor.2010.14 (2010).

4   Franklin, R. E. & Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **171**, 740-741, doi:10.1038/171740a0 (1953).

5   Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).

6   Drew, H. R. *et al.* Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2179-2183 (1981).

7   Choi, J. & Majima, T. Conformational changes of non-B DNA. *Chem. Soc. Rev.* **40**, 5893-5909, doi:10.1039/c1cs15153c (2011).

8   Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability. *Mutat. Res.* **598**, 103-119, doi:10.1016/j.mrfmmm.2006.01.019 (2006).

9   Bacolla, A. & Wells, R. D. Non-B DNA conformations as determinants of mutagenesis and human disease. *Mol. Carcinog.* **48**, 273-285, doi:10.1002/mc.20507 (2009).

10   Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell Mol. Life Sci.* **67**, 43-62, doi:10.1007/s00018-009-0131-2 (2010).

11   Dong, Y., Yang, Z. & Liu, D. DNA nanotechnology based on i-motif structures. *Acc. Chem. Res.* **47**, 1853-1860, doi:10.1021/ar500073a (2014).

12   Kim, J., Lee, Y. M., Kang, Y. & Kim, W. J. Tumor-homing, size-tunable clustered nanoparticles for anticancer therapeutics. *ACS nano* **8**, 9358-9367, doi:10.1021/nn503349g (2014).

13   Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* **59**, 871-880, doi:10.1016/0092-8674(89)90610-7 (1989).

14   Gellert, M., Lipsett, M. N. & Davies, D. R. HELIX FORMATION BY GUANYLIC ACID. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 2013-2018 (1962).

15   Collie, G. W. *et al.* Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *J. Am. Chem. Soc.* **134**, 2723-2731, doi:10.1021/ja2102423 (2012).

16   Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908-2916, doi:10.1093/nar/gki609 (2005).

17   Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182-186, doi:10.1038/nchem.1548 (2013).

18   Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **6**, 75-80, doi:10.1038/nchem.1805 (2014).

19   Gray, L. T., Vallur, A. C., Eddy, J. & Maizels, N. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat. Chem. Biol.* **10**, 313-318, doi:10.1038/nchembio.1475 (2014).

20   Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex

in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11593-11598, doi:10.1073/pnas.182256799 (2002).

21   Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13**, 770-780, doi:10.1038/nrg3296 (2012).

22   Cheung, I., Schertzer, M., Rose, A. & Lansdorp, P. M. Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. *Nat. Genet.* **31**, 405-409, doi:10.1038/ng928 (2002).

23   Maizels, N. G4-associated human diseases. *EMBO Rep.* **16**, 910-922, doi:10.15252/embr.201540607 (2015).

24   Hansel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, doi:10.1038/ng.3662 (2016).

25   Waller, Z. A. E., Sewitz, S. A., Hsu, S. T. & Balasubramanian, S. A small molecule that disrupts G-quadruplex DNA structure and enhances gene expression. *J. Am. Chem. Soc.* **131**, 12628-12633, doi:10.1021/ja901892u (2009).

26   Rodriguez, R. *et al.* A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **130**, 15758-15759, doi:10.1021/ja805615w (2008).

27   Tlučková, K. *et al.* Human papillomavirus G-quadruplexes. *Biochemistry* **52**, 7207-7216, doi:10.1021/bi400897g (2013).

28   Perrone, R. *et al.* A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J Med Chem* **56**, 6521-6530, doi:10.1021/jm400914r (2013).

29   Paritala, H. & Firestine, S. M. Characterization of insulin ILPR sequences for their ability to adopt a G-quadruplex structure. *Nucleosides, nucleotides & nucleic acids* **29**, 81-90, doi:10.1080/15257771003597691 (2010).

30   Leroy, J. L., Guéron, M., Mergny, J. L. & Hélène, C. Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res.* **22**, 1600-1606 (1994).

31   Gehring, K., Leroy, J. L. & Guéron, M. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature* **363**, 561-565, doi:10.1038/363561a0 (1993).

32   Collin, D. & Gehring, K. Stability of Chimeric DNA/RNA Cytosine Tetrads: Implications for i-Motif Formation by RNA. *J. Am. Chem. Soc.* **120**, 4069-4072, doi:10.1021/ja973346r (1998).

33   Han, X., Leroy, J. L. & Guéron, M. An intramolecular i-motif: the solution structure and base-pair opening kinetics of d(5mCCT3CCT3ACCT3CC). *J. Mol. Biol.* **278**, 949-965, doi:10.1006/jmbi.1998.1740 (1998).

34   Garavis, M., Escaja, N., Gabelica, V., Villasante, A. & Gonzalez, C. Centromeric Alpha-Satellite DNA Adopts Dimeric i-Motif Structures Capped by AT Hoogsteen Base Pairs. *Chem. A Eur. J.* **21**, 9816-9824, doi:10.1002/chem.201500448 (2015).

35   Ghodke, H. B. *et al.* The I-tetraplex building block: rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson-Crick interactions. *Angew. Chem. Int. Ed.* **46**, 2646-2649, doi:10.1002/anie.200604461 (2007).

36   Shu, W. *et al.* DNA molecular motor driven micromechanical cantilever arrays. *J. Am. Chem. Soc.* **127**, 17054-17060, doi:10.1021/ja0554514 (2005).

37   Li, T. & Famulok, M. I-motif-programmed functionalization of DNA nanocircles. *J. Am. Chem. Soc.* **135**, 1593-1599, doi:10.1021/ja3118224 (2013).

38   Idili, A., Vallée-Bélisle, A. & Ricci, F. Programmable pH-triggered DNA nanoswitches. *J. Am.*

Chem. Soc. **136**, 5836-5839, doi:10.1021/ja500619w (2014).

39   Brazier, J. A., Shah, A. & Brown, G. D. I-motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* **48**, 10739-10741, doi:10.1039/c2cc30863k (2012).

40   Day, H. A., Pavlou, P. & Waller, Z. A. E. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **22**, 4407-4418, doi:10.1016/j.bmc.2014.05.047 (2014).

41   Reilly, S. M. *et al.* Folding and hydrodynamics of a DNA i-motif from the c-MYC promoter determined by fluorescent cytidine analogs. *Biophys. J.* **107**, 1703-1711, doi:10.1016/j.bpj.2014.08.014 (2014).

42   Kendrick, S., Akiyama, Y., Hecht, S. M. & Hurley, L. H. The i-motif in the bcl-2 P1 promoter forms an unexpectedly stable structure with a unique 8:5:7 loop folding pattern. *J. Am. Chem. Soc.* **131**, 17667-17676, doi:10.1021/ja9076292 (2009).

43   Brooks, T. A., Kendrick, S. & Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* **277**, 3459-3469, doi:10.1111/j.1742-4658.2010.07759.x (2010).

44   Gurung, S. P., Schwarz, C., Hall, J. P., Cardin, C. J. & Brazier, J. A. The importance of loop length on the stability of i-motif structures. *Chem. Commun.* **51**, 5630-5632, doi:10.1039/c4cc07279k (2015).

45   Reilly, S. M., Morgan, R. K., Brooks, T. A. & Wadkins, R. M. Effect of interior loop length on the thermal stability and pK(a) of i-motif DNA. *Biochemistry* **54**, 1364-1370, doi:10.1021/bi5014722 (2015).

46   Kim, S. E., Lee, I. B., Hyeon, C. & Hong, S. C. Destabilization of i-motif by submolar concentrations of a monovalent cation. *J. Phys. Chem. B* **118**, 4753-4760, doi:10.1021/jp500120d (2014).

47   Day, H. A., Huguin, C. & Waller, Z. A. E. Silver cations fold i-motif at neutral pH. *Chem. Commun.* **49**, 7696-7698, doi:10.1039/c3cc43495h (2013).

48   Ono, A., Torigoe, H., Tanaka, Y. & Okamoto, I. Binding of metal ions by pyrimidine base pairs in DNA duplexes. *Chem. Soc. Rev.* **40**, 5855-5866, doi:10.1039/c1cs15149e (2011).

49   Day, H. A., Wright, E. P., MacDonald, C. J., Gates, A. J. & Waller, Z. A. E. Reversible DNA i-motif to hairpin switching induced by copper(II) cations. *Chem. Commun.* **51**, 14099-14102, doi:10.1039/c5cc05111h (2015).

50   Lacroix, L., Mergny, J. L., Leroy, J. L. & Hélène, C. Inability of RNA to form the i-motif: implications for triplex formation. *Biochemistry* **35**, 8715-8722, doi:10.1021/bi960107s (1996).

51   Chakraborty, S. & Krishnan, Y. Kinetic hybrid i-motifs: intercepting DNA with RNA to form a DNA(2)-RNA(2) i-motif. *Biochimie* **90**, 1088-1095, doi:10.1016/j.biochi.2008.02.022 (2008).

52   Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35**, 406-413, doi:10.1093/nar/gkl1057 (2007).

53   Dexheimer, T. S. *et al.* NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III(1). *Mol. Cancer Ther.* **8**, 1363-1377, doi:10.1158/1535-7163.MCT-08-1093 (2009).

54   Sutherland, C., Cui, Y., Mao, H. & Hurley, L. H. A Mechanosensor Mechanism Controls the G-Quadruplex/i-Motif Molecular Switch in the MYC Promoter NHE III1. *J. Am. Chem. Soc.*, doi:10.1021/jacs.6b09196 (2016).

55   Xu, Y. & Sugiyama, H. Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res.* **34**, 949-954, doi:10.1093/nar/gkj485 (2006).

56   Kendrick, S. *et al.* The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* **136**, 4161-4171, doi:10.1021/ja410934b (2014).

57   Kang, H. J., Kendrick, S., Hecht, S. M. & Hurley, L. H. The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* **136**, 4172-4185, doi:10.1021/ja4109352 (2014).

58   Roy, B. *et al.* Interaction of Individual Structural Domains of hnRNP LL with the BCL2 Promoter i-Motif DNA. *J. Am. Chem. Soc.* **138**, 10950-10962, doi:10.1021/jacs.6b05036 (2016).

59   Manzini, G., Yathindra, N. & Xodo, L. E. Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res.* **22**, 4634-4640, doi:10.1093/nar/22.22.4634 (1994).

60   Sun, D. & Hurley, L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.* **52**, 2863-2874, doi:10.1021/jm900055s (2009).

61   Miyoshi, D. & Sugimoto, N. Molecular crowding effects on structure and stability of DNA. *Biochimie* **90**, 1040-1051, doi:10.1016/j.biochi.2008.02.009 (2008).

62   Miyoshi, D., Matsumura, S., Nakano, S. & Sugimoto, N. Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc.* **126**, 165-169, doi:10.1021/ja036721q (2004).

63   Cui, J., Waltman, P., Le, V. H. & Lewis, E. A. The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules* **18**, 12751-12767, doi:10.3390/molecules181012751 (2013).

64   Rajendran, A., Nakano, S. & Sugimoto, N. Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* **46**, 1299-1301, doi:10.1039/b922050j (2010).

65   Oganesian, L. & Bryan, T. M. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *BioEssays : news and reviews in molecular, cellular and developmental biology* **29**, 155-165, doi:10.1002/bies.20523 (2007).

66   Conomos, D., Pickett, H. A. & Reddel, R. R. Alternative lengthening of telomeres: remodeling the telomere architecture. *Front. Oncol.* **3**, 27, doi:10.3389/fonc.2013.00027 (2013).

67   Xu, Y. Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chem. Soc. Rev.* **40**, 2719-2740, doi:10.1039/c0cs00134a (2011).

68   Mergny, J. L., Riou, J. F., Mailliet, P., Teulade-Fichou, M.-P. & Gilson, E. Natural and pharmacological regulation of telomerase. *Nucleic Acids Res.* **30**, 839-865, doi:10.1093/nar/30.4.839 (2002).

69   Kim, N. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015, doi:10.1126/science.7605428 (1994).

70   Müller, S. *et al.* Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells. *Org. Biomol. Chem.* **10**, 6537-6546, doi:10.1039/c2ob25830g (2012).

71   Burger, A. M. *et al.* The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res.* **65**, 1489-1496, doi:10.1158/0008-5472.CAN-04-2910 (2005).

72   Chen, Y. *et al.* Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.* **3**, 1074, doi:10.1038/ncomms2091 (2012).

73   Marsich, E., Piccini, A., Xodo, L. E. & Manzini, G. Evidence for a HeLa nuclear protein that binds specifically to the single-stranded d(CCCTAA)n telomeric motif. *Nucleic Acids Res.* **24**, 4029-4033, doi:10.1093/nar/24.20.4029 (1996).

74   Lacroix, L. *et al.* Identification of two human nuclear proteins that recognise the cytosine-rich strand of human telomeres in vitro. *Nucleic Acids Res.* **28**, 1564-1575, doi:10.1093/nar/28.7.1564 (2000).

75   Simonsson, T. & Kubista, M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* **26**, 1167-1172, doi:10.1093/nar/26.5.1167 (1998).

76   Simonsson, T., Pribylova, M. & Vorlickova, M. A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Biophys. Res. Commun.* **278**, 158-166, doi:10.1006/bbrc.2000.3783 (2000).

77   Guo, K., Gokhale, V., Hurley, L. H. & Sun, D. Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* **36**, 4598-4608, doi:10.1093/nar/gkn380 (2008).

78   Bucek, P., Jaumot, J., Avino, A., Eritja, R. & Gargallo, R. pH-Modulated Watson-Crick duplex-quadruplex equilibria of guanine-rich and cytosine-rich DNA sequences 140 base pairs upstream of the c-kit transcription initiation site. *Chem. A Eur. J.* **15**, 12663-12671, doi:10.1002/chem.200901631 (2009).

79   Bucek, P., Gargallo, R. & Kudrev, A. Spectrometric study of the folding process of i-motif-forming DNA sequences upstream of the c-kit transcription initiation site. *Anal. Chim. Acta.* **683**, 69-77, doi:10.1016/j.aca.2010.10.008 (2010).

80   Mergny, J. L. Fluorescence energy transfer as a probe for tetraplex formation: the i-motif. *Biochemistry* **38**, 1573-1581, doi:10.1021/bi982208r (1999).

81   De Cian, A. *et al.* Fluorescence-based melting assays for studying quadruplex ligands. *Methods* **42**, 183-195, doi:10.1016/j.ymeth.2006.10.004 (2007).

82   Chirio-Lebrun, M.-C. & Prats, M. Fluorescence resonance energy transfer (FRET): theory and experiments. *Biochem. Educ.* **26**, 320-323, doi:10.1016/s0307-4412(98)80010-1 (1998).

83   Mergny, J. L., Li, J., Lacroix, L., Amrane, S. & Chaires, J. B. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* **33**, e138, doi:10.1093/nar/gni134 (2005).

84   Mergny, J. L. & Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **13**, 515-537, doi:10.1089/154545703322860825 (2003).

85   Nakanishi, K., Berova, N. & Woody, R. W. *Circular Dichroism Principles and Applications*. (VCH Publishers Inc., 1991).

86   Vorlíčková, M. *et al.* Circular dichroism and guanine quadruplexes. *Methods* **57**, 64-75, doi:10.1016/j.ymeth.2012.03.011 (2012).

87   Kypr, J., Kejnovská, I., Renčiuk, D. & Vorlícková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713-1725, doi:10.1093/nar/gkp026 (2009).

88   Webba da Silva, M. NMR methods for studying quadruplex nucleic acids. *Methods* **43**, 264-277, doi:10.1016/j.ymeth.2007.05.007 (2007).

89   Snoussi, K., Nonin-Lecomte, S. & Leroy, J. L. The RNA i-motif. *J. Mol. Biol.* **309**, 139-153, doi:10.1006/jmbi.2001.4618 (2001).

90   Leroy, J.-L. & Guéron, M. Solution structures of the i-motif tetramers of d(TCC), d(5methylCCT) and d(T5methylCC): novel NOE connections between amino protons and sugar protons. *Structure* **3**, 101-120, doi:10.1016/s0969-2126(01)00138-1 (1995).

91   Phan, A. T., Guéron, M. & Leroy, J. L. The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* **299**, 123-144, doi:10.1006/jmbi.2000.3613 (2000).

92   Zeng, S., Baillargeat, D., Ho, H. P. & Yong, K. T. Nanomaterials enhanced surface plasmon resonance for biological and chemical sensing applications. *Chem. Soc. Rev.* **43**, 3426-3452, doi:10.1039/c3cs60479a (2014).

93   Zeng, Z. X., Zhao, Y., Hao, Y. H. & Tan, Z. Tetraplex formation of surface-immobilized human telomere sequence probed by surface plasmon resonance using single-stranded DNA binding protein. *J. Mol. Recognit.* **18**, 267-271, doi:10.1002/jmr.731 (2005).

94   Li, Q. *et al.* G4LDB: a database for discovering and studying G-quadruplex ligands. *Nucleic Acids Res.* **41**, D1115-1123, doi:10.1093/nar/gks1101 (2013).

95   Li, X., Peng, Y., Ren, J. & Qu, X. Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 19658-19663, doi:10.1073/pnas.0607245103 (2006).

96   Martino, L., Pagano, B., Fotticchia, I., Neidle, S. & Giancola, C. Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. *J. Phys. Chem. B* **113**, 14779-14786, doi:10.1021/jp9066394 (2009).

97   Ma, D. L. *et al.* Crystal violet as a fluorescent switch-on probe for i-motif: label-free DNA-based logic gate. *The Analyst* **136**, 2692-2696, doi:10.1039/c1an15091j (2011).

98   Alberti, P. *et al.* Interaction of an acridine dimer with DNA quadruplex structures. *J. Biomol. Struct. Dyn.* **19**, 505-513, doi:10.1080/07391102.2001.10506758 (2001).

99   Wang, L., Wu, Y., Chen, T. & Wei, C. The interactions of phenanthroline compounds with DNAs: preferential binding to telomeric quadruplex over duplex. *Int. J. Biol. Macromol.* **52**, 1-8, doi:10.1016/j.ijbiomac.2012.08.015 (2013).

100   Xu, H., Zhang, H. & Qu, X. Interactions of the human telomeric DNA with terbium-amino acid complexes. *J. Inorg. Biochem.* **100**, 1646-1652, doi:10.1016/j.jinorgbio.2006.05.015 (2006).

101   Day, H. A. *Investigating the Effect of Small Molecule Ligands and Cations on i-mmotif DNA* Doctor of Philosophy thesis, University of East Anglia, (2015).

102   Wright, E. P. *et al. Sci. Reports.* **6**, 39456, doi:10.1038/srep39456 (2016).

103   Mathur, V., Verma, A., Maiti, S. & Chowdhury, S. Thermodynamics of i-tetraplex formation in the nuclease hypersensitive element of human c-myc promoter. *Biochemical and biophysical research communications* **320**, 1220-1227, doi:10.1016/j.bbrc.2004.06.074 (2004).

104   SantaLucia, J., Jr. & Hicks, D. The thermodynamics of DNA structural motifs. *Annual review of biophysics and biomolecular structure* **33**, 415-440, doi:10.1146/annurev.biophys.32.110601.141800 (2004).

105   Hardin, C. C., Perry, A. G. & White, K. Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers* **56**, 147-194, doi:10.1002/1097-0282(2000/2001)56:3<147::aid-bip10011>3.0.co;2-n (2000).

106   Lerman, L. S. Structural considerations in the interaction of DNA and acridines. *Journal of molecular biology* **3**, 18-IN14, doi:10.1016/s0022-2836(61)80004-1 (1961).

107   Haq, I. Thermodynamics of drug–DNA interactions. *Archives of Biochemistry and Biophysics*

403, 1-15, doi:10.1016/s0003-9861(02)00202-3 (2002).

108   Boger, D. L., Fink, B. E. & Hedrick, M. P. Total Synthesis of Distamycin A and 2640 Analogues: A Solution-Phase Combinatorial Approach to the Discovery of New, Bioactive DNA Binding Agents and Development of a Rapid, High-Throughput Screen for Determining Relative DNA Binding Affinity or DNA Binding Sequence Selectivity. *J. Am. Chem. Soc.* **122**, 6382-6394, doi:10.1021/ja994192d (2000).

109   Boger, D. L. & Tse, W. C. Thiazole orange as the fluorescent intercalator in a high resolution fid assay for determining DNA binding affinity and sequence selectivity of small molecules. *Bioorg. Med. Chem.* **9**, 2511-2518, doi:10.1016/s0968-0896(01)00243-7 (2001).

110   Blasko, A., Browne, K. A., He, G. X. & Bruice, T. C. Chemistry of phosphodiesters, DNA and models. 5. Microgonotropens and their interactions with DNA. 3. Structural analysis of the 1:1 complex of d(CGCAAATTTGCG)2 and dien-microgonotropen-c by 2D NMR spectroscopy and restrained molecular modeling. *J. Am. Chem. Soc.* **115**, 7080-7092, doi:10.1021/ja00069a004 (1993).

111   Boger, D. L., Fink, B. E., Brunette, S. R., Tse, W. C. & Hedrick, M. P. A Simple, High-Resolution Method for Establishing DNA Binding Affinity and Sequence Selectivity. *J. Am. Chem. Soc.* **123**, 5878-5891, doi:10.1021/ja010041a (2001).

112   Monchaud, D., Allain, C. & Teulade-Fichou, M. P. Development of a fluorescent intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands. *Bioorg. Med. Chem. Lett.* **16**, 4842-4845, doi:10.1016/j.bmcl.2006.06.067 (2006).

113   Allain, C., Monchaud, D. & Teulade-Fichou, M. P. FRET templated by G-quadruplex DNA: a specific ternary interaction using an original pair of donor/acceptor partners. *J. Am. Chem. Soc.* **128**, 11890-11893, doi:10.1021/ja062193h (2006).

114   Monchaud, D., Allain, C. & Teulade-Fichou, M. P. Thiazole orange: a useful probe for fluorescence sensing of G-quadruplex-ligand interactions. *Nucleosides, nucleotides & nucleic acids* **26**, 1585-1588, doi:10.1080/15257770701548212 (2007).

115   Monchaud, D. *et al.* Ligands playing musical chairs with G-quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex binders. *Biochimie* **90**, 1207-1223, doi:10.1016/j.biochi.2008.02.019 (2008).

116   Largy, E., Hamon, F. & Teulade-Fichou, M. P. Development of a high-throughput G4-FID assay for screening and evaluation of small molecules binding quadruplex nucleic acid structures. *Anal. Bioanal. Chem.* **400**, 3419-3427, doi:10.1007/s00216-011-5018-z (2011).

117   Tran, P. L., Largy, E., Hamon, F., Teulade-Fichou, M. P. & Mergny, J. L. Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-quadruplex structures. *Biochimie* **93**, 1288-1296, doi:10.1016/j.biochi.2011.05.011 (2011).

118   Zhang, X. Y., Luo, H. Q. & Li, N. B. Crystal violet as an i-motif structure probe for reversible and label-free pH-driven electrochemical switch. *Anal. Biochem.* **455**, 55-59, doi:10.1016/j.ab.2014.03.015 (2014).

119   Kong, D. M., Ma, Y. E., Wu, J. & Shen, H. X. Discrimination of G-quadruplexes from duplex and single-stranded DNAs with fluorescence and energy-transfer fluorescence spectra of crystal violet. *Chem. A Eur. J.* **15**, 901-909, doi:10.1002/chem.200801441 (2009).

120   Xu, B., Wu, X., Yeow, E. K. & Shao, F. A single thiazole orange molecule forms an exciplex in a DNA i-motif. *Chem. Commun.* **50**, 6402-6405, doi:10.1039/c4cc01147c (2014).

121   Xu, L. *et al.* Thiazole Orange as a Fluorescent Light-Up Probe for the i-Motif and its Application to the Development of a Molecular Logic System. *Asian J. Org. Chem.* **4**, 1375-1378, doi:10.1002/ajoc.201500347 (2015).

122   Devi, G., He, L., Xu, B., Li, T. & Shao, F. In-stem thiazole orange reveals the same triplex intermediate for pH and thermal unfolding of i-motifs. *Chem. Commun.* **52**, 7261-7264, doi:10.1039/c6cc01643j (2016).

123   Lee, I. J. & Kim, B. H. Monitoring i-motif transitions through the exciplex emission of a fluorescent probe incorporating two (Py)A units. *Chem. Commun.* **48**, 2074-2076, doi:10.1039/c1cc16497j (2012).

124   Choi, J., Tanaka, A., Cho, D. W., Fujitsuka, M. & Majima, T. Efficient electron transfer in i-motif DNA with a tetraplex structure. *Angew. Chem. Int. Ed.* **52**, 12937-12941, doi:10.1002/anie.201306017 (2013).

125   Tse, W. C. & Boger, D. L. A fluorescent intercalator displacement assay for establishing DNA binding selectivity and affinity. *Curr. Protoc. Nucleic Acid Chem.* **Chapter 8**, Unit 8 5, doi:10.1002/0471142700.nc0805s20 (2005).

126   Hu, W. *et al.* Dimeric calixarenes: a new family of major-groove binders. *Chem. A Eur. J.* **18**, 3589-3597, doi:10.1002/chem.201100634 (2012).

127   Eggleston, A. K., Rahim, N. A. & Kowalczykowski, S. C. A Helicase Assay Based on the Displacement of Fluorescent, Nucleic Acid-Binding Ligands. *Nucleic Acids Res.* **24**, 1179-1186, doi:10.1093/nar/24.7.1179 (1996).

128   Job, P. Formation and stability of inorganic complexes in solution. *Ann. Chim.* **9**, 113-203 (1928).

129   Hulme, E. C. & Trevethick, M. A. Ligand binding assays at equilibrium: validation and interpretation. *Br. J. Pharmacol.* **161**, 1219-1237, doi:10.1111/j.1476-5381.2009.00604.x (2010).

130   Kemmer, G. & Keller, S. Nonlinear least-squares data fitting in Excel spreadsheets. *Nat. Protoc.* **5**, 267-281, doi:10.1038/nprot.2009.182 (2010).

131   Carrasco, C. *et al.* Tight Binding of the Antitumor Drug Ditercalinium to Quadruplex DNA. *Chembiochem : a European journal of chemical biology* **3**, 1235-1241, doi:10.1002/1439-7633(20021202)3:12<1235::aid-cbic1235>3.0.co;2-i (2002).

132   Redman, J. E. Surface plasmon resonance for probing quadruplex folding and interactions with proteins and small molecules. *Methods* **43**, 302-312, doi:10.1016/j.ymeth.2007.05.008 (2007).

133   Nygren, J., Svanvik, N. & Kubista, M. The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **46**, 39-51, doi:10.1002/(sici)1097-0282(199807)46:1<39::aid-bip4>3.0.co;2-z (1998).

134   Lubitz, I., Zikich, D. & Kotlyar, A. Specific high-affinity binding of thiazole orange to triplex and G-quadruplex DNA. *Biochemistry* **49**, 3567-3574, doi:10.1021/bi1000849 (2010).

135   Fernandez, S., Eritja, R., Avino, A., Jaumot, J. & Gargallo, R. Influence of pH, temperature and the cationic porphyrin TMPyP4 on the stability of the i-motif formed by the 5'-(C3TA2)4-3' sequence of the human telomere. *Int. J. Biol. Macromol.* **49**, 729-736, doi:10.1016/j.ijbiomac.2011.07.004 (2011).

136   Watkins, D., Ranjan, N., Kumar, S., Gong, C. & Arya, D. P. An assay for human telomeric G-quadruplex DNA binding drugs. *Bioorg. Med. Chem. Lett.* **23**, 6695-6699, doi:10.1016/j.bmcl.2013.10.030 (2013).

137   Zhang, J. H., Chung, T. D. Y. & Oldenbrug, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular*

Screening **4**, 67-73, doi:10.1177/108705719900400206 (1999).

138   Gibson, R. L., Burns, J. L. & Ramsey, B. W. Pathophysiology and management of pulmonary infections in cystic fibrosis. *American journal of respiratory and critical care medicine* **168**, 918-951, doi:10.1164/rccm.200304-505SO (2003).

139   Purdy Drew, K. R., Sanders, L. K., Culumber, Z. W., Zribi, O. & Wong, G. C. Cationic amphiphiles increase activity of aminoglycoside antibiotic tobramycin in the presence of airway polyelectrolytes. *Journal of the American Chemical Society* **131**, 486-493, doi:10.1021/ja803925n (2009).

140   Rodik, R. V., Klymchenko, A. S., Mely, Y. & Kalchenko, V. I. Calixarenes and related macrocycles as gene delivery vehicles. *J. Incl. Phenom. Macrocyclic Chem.* **80**, 189-200, doi:10.1007/s10847-014-0412-8 (2014).

141   Baekeland, L. H. Method of making insoluble products of phenol and formaldehyde. (1909).

142   Zinke, A. *et al.* Zur Kenntnis des Härtungsprozesses von Phenol-Formaldehyd-Harzen. *Monatshefte für Chemie und verwandte Teile anderer Wissenschaften* **83**, 1213-1227, doi:10.1007/bf00899467 (1952).

143   Gutsche, C. D. & Muthukrishnan, R. Calixarenes. 1. Analysis of the product mixtures produced by the base-catalyzed condensation of formaldehyde with para-substituted phenols. *J. Org. Chem.* **43**, 4905-4906, doi:10.1021/jo00419a052 (1978).

144   Yousaf, A., Hamid, S. A., Bunnori, N. M. & Ishola, A. A. Applications of calixarenes in cancer chemotherapy: facts and perspectives. *Drug Des. Devel. Ther.* **9**, 2831-2838, doi:10.2147/DDDT.S83213 (2015).

145   Kobayashi, K. & Yamanaka, M. Self-assembled capsules based on tetrafunctionalized calix[4]resorcinarene cavitands. *Chem. Soc. Rev.* **44**, 449-466, doi:10.1039/c4cs00153b (2015).

146   Wang, G., Zhang, H., Ding, F. & Liu, Y. Preparation and characterization of inclusion complexes of topotecan with sulfonatocalixarene. *J. Incl. Phenom. Macrocyclic Chem.* **69**, 85-89, doi:10.1007/s10847-010-9817-1 (2010).

147   Cao, L. *et al.* Enhancement of antitumor properties of TRAIL by targeted delivery to the tumor neovasculature. *Mol. Cancer Ther.* **7**, 851-861, doi:10.1158/1535-7163.MCT-07-0533 (2008).

148   Böhmer, V., Chhim, P. & Kämmerer, H. *Die Makromolekulare Chemie* **180**, 2503-2506, doi:10.1002/macp.1979.021801025 (1979).

149   Gutsche, C. D. *Calixarene: An Introduction, 2nd Edition*. (The Royal Society of Chemistry, 2008).

150   Iwamoto, K., Araki, K. & Shinkai, S. Conformations and structures of tetra-O-alkyl-p-tert-butylcalix[4]arenes. How is the conformation of calix[4]arenes immobilized? *J. Org. Chem.* **56**, 4955-4962, doi:10.1021/jo00016a027 (1991).

151   Rodik, R. V. *et al.* Virus-sized DNA nanoparticles for gene delivery based on micelles of cationic calixarenes. *Chem. A Eur. J.* **17**, 5526-5538, doi:10.1002/chem.201100154 (2011).

152   Nimse, S. B. & Kim, T. Biological applications of functionalized calixarenes. *Chem. Soc. Rev.* **42**, 366-386, doi:10.1039/c2cs35233h (2013).

153   Nimse, S. B., Song, K. S., Kim, J., Sayyed, D. R. & Kim, T. 9G DNAChip Technology: Self-Assembled Monolayer (SAM) of ssDNA for Ultra-Sensitive Detection of Biomarkers. *Int. J. Mol. Sci.* **14**, 5723-5733, doi:10.3390/ijms14035723 (2013).

154   Song, K. S. *et al.* 9G DNAChip: microarray based on the multiple interactions of 9 consecutive guanines. *Chem. Commun.* **47**, 7101-7103, doi:10.1039/c1cc12489g (2011).

155 Lee, Y. *et al.* ProteoChip: a highly sensitive protein microarray prepared by a novel method of protein immobilization for application of protein-protein interaction studies. *Proteomics* **3**, 2289-2304, doi:10.1002/pmic.200300541 (2003).

156 Amiri, A., Choi, E. Y. & Kim, H. J. Development and molecular recognition of Calixcrownchip as an electrochemical ALT immunosensor. *J. Incl. Phenom. Macrocyclic Chem.* **66**, 185-194, doi:10.1007/s10847-009-9702-y (2009).

157 Zhou, Y., Li, H. & Yang, Y. W. Controlled drug delivery systems based on calixarenes. *Chin. Chem. Lett.* **26**, 825-828, doi:10.1016/j.cclet.2015.01.038 (2015).

158 Wheate, N. J. *et al.* Side-on binding of p-sulphonatocalix[4]arene to the dinuclear platinum complex trans-[{PtCl(NH3)2}2mu-dpzm]2+ and its implications for anticancer drug delivery. *J. Inorg. Biochem.* **103**, 448-454, doi:10.1016/j.jinorgbio.2008.12.011 (2009).

159 Song, J., Li, H., Chao, J., Dong, C. & Shuang, S. Spectroscopic studies on the inclusion interaction of p-sulfonatocalix[6]arene with vitamin B6. *J. Incl. Phenom. Macrocyclic Chem.* **72**, 389-395, doi:10.1007/s10847-011-9994-6 (2011).

160 Tu, C. *et al.* Supramolecular polymeric micelles by the host-guest interaction of star-like calix[4]arene and chlorin e6 for photodynamic therapy. *Chem. Commun.* **47**, 6063-6065, doi:10.1039/c0cc05662f (2011).

161 James, E. *et al.* Antioxidant phospholipid calix[4]arene mimics as micellular delivery systems. *Org. Biomol. Chem.* **11**, 6108-6112, doi:10.1039/c3ob41178h (2013).

162 Lee, M., Lee, S. J. & Jiang, L. H. Stimuli-responsive supramolecular nanocapsules from amphiphilic calixarene assembly. *J. Am. Chem. Soc.* **126**, 12724-12725, doi:10.1021/ja045918v (2004).

163 Wang, K., Guo, D. S., Zhao, M. Y. & Liu, Y. A Supramolecular Vesicle Based on the Complexation of p-Sulfonatocalixarene with Protamine and its Trypsin-Triggered Controllable-Release Properties. *Chem. A Eur. J.* **22**, 1475-1483, doi:10.1002/chem.201303963 (2016).

164 Peters, M. S., Li, M. & Schrader, T. Interactions of calix [n] arenes with nucleic acids. *Natural product communications* **7**, 409-417 (2012).

165 Sansone, F. *et al.* DNA condensation and cell transfection properties of guanidinium calixarenes: dependence on macrocycle lipophilicity, size, and conformation. *J. Am. Chem. Soc.* **128**, 14528-14536, doi:10.1021/ja0634425 (2006).

166 Nault, L., Cumbo, A., Pretôt, R. F., Sciotti, M. A. & Shahgaldian, P. Cell transfection using layer-by-layer (LbL) coated calixarene-based solid lipid nanoparticles (SLNs). *Chem. Commun.* **46**, 5581-5583, doi:10.1039/b926025k (2010).

167 Bagnacani, V. *et al.* Arginine clustering on calix[4]arene macrocycles for improved cell penetration and DNA delivery. *Nat. Commun.* **4**, 1721, doi:10.1038/ncomms2721 (2013).

168 Rodik, R. V., Anthony, A.-S., Kalchenko, V. I., Mély, Y. & Klymchenko, A. S. Cationic amphiphilic calixarenes to compact DNA into small nanoparticles for gene delivery. *New J. Chem.* **39**, 1654-1664, doi:10.1039/c4nj01395f (2015).

169 Lalor, R., DiGesso, J. L., Mueller, A. & Matthews, S. E. Efficient gene transfection with functionalised multicalixarenes. *Chem. Commun.*, 4907, doi:10.1039/b712100h (2007).

170 Lee, Y. C. & Lee, R. T. Carbohydrate-Protein Interactions: Basis of Glycobiology. *Acc. Chem. Res.* **28**, 321-327, doi:10.1021/ar00056a001 (1995).

171 Bojarová, P. & Křen, V. Sugared biomaterial binding lectins: achievements and perspectives. *Biomater. Sci.*, doi:10.1039/c6bm00088f (2016).

172   Air, G. M. Influenza virus-glycan interactions. *Curr. Opin. Virol.* **7**, 128-133, doi:10.1016/j.coviro.2014.06.004 (2014).

173   Marra, A. *et al.* Synthesis of sialoclusters appended to calix[4]arene platforms via multiple azide-alkyne cycloaddition. New inhibitors of hemagglutination and cytopathic effect mediated by BK and influenza A viruses. *Org. Biomol. Chem.* **6**, 1396-1409, doi:10.1039/b800598b (2008).

174   Zhang, R. G. *et al.* The three-dimensional crystal structure of cholera toxin. *J. Mol. Biol.* **251**, 563-573, doi:10.1006/jmbi.1995.0456 (1995).

175   De, S. N. Enterotoxicity of Bacteria-free Culture-filtrate of Vibrio cholerae. *Nature* **183**, 1533-1534, doi:10.1038/1831533a0 (1959).

176   Arosio, D. *et al.* A synthetic divalent cholera toxin glycocalix[4]arene ligand having higher affinity than natural GM1 oligosaccharide. *J. Am. Chem. Soc.* **127**, 3660-3661, doi:10.1021/ja0444029 (2005).

177   Cecioni, S. *et al.* Achieving high affinity towards a bacterial lectin through multivalent topological isomers of calix[4]arene glycoconjugates. *Chem. A Eur. J.* **15**, 13232-13240, doi:10.1002/chem.200901799 (2009).

178   Boukerb, A. M. *et al.* Antiadhesive properties of glycoclusters against Pseudomonas aeruginosa lung infection. *J. Med. Chem.* **57**, 10275-10289, doi:10.1021/jm500038p (2014).

179   Bernardi, S. *et al.* Clicked and long spaced galactosyl- and lactosylcalix[4]arenes: new multivalent galectin-3 ligands. *Beilstein J. Org. Chem.* **10**, 1672-1680, doi:10.3762/bjoc.10.175 (2014).

180   Giuliani, M., Morbioli, I., Sansone, F. & Casnati, A. Moulding calixarenes for biomacromolecule targeting. *Chem. Commun.* **51**, 14140-14159, doi:10.1039/c5cc05204a (2015).

181   Lin, Q., Park, H. S., Hamuro, Y., Lee, C. S. & Hamilton, A. D. Protein surface recognition by synthetic agents: Design and structural requirements of a family of artificial receptors that bind to cytochrome C. *Biopolymers* **47**, 285-297, doi:10.1002/(sici)1097-0282(1998)47:4<285::aid-bip4>3.0.co;2-c (1998).

182   Hamuro, Y., Calama, M. C., Park, H. S. & Hamilton, A. D. A Calixarene with Four Peptide Loops: An Antibody Mimic for Recognition of Protein Surfaces. *Angew. Chem. Int. Ed.* **36**, 2680-2683, doi:10.1002/anie.199726801 (1997).

183   Park, H. S., Lin, Q. & Hamilton, A. D. Protein Surface Recognition by Synthetic Receptors: A Route to Novel Submicromolar Inhibitors for α-Chymotrypsin. *J. Am. Chem. Soc.* **121**, 8-13, doi:10.1021/ja981504o (1999).

184   Blaskovich, M. A. *et al.* Design of GFB-111, a platelet-derived growth factor binding molecule with antiangiogenic and anticancer activity against human tumors in mice. *Nat. Biotechnol.* **18**, 1065-1070, doi:10.1038/80257 (2000).

185   Martos, V. *et al.* Molecular recognition and self-assembly special feature: Calix[4]arene-based conical-shaped ligands for voltage-dependent potassium channels. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10482-10486, doi:10.1073/pnas.0813396106 (2009).

186   Gordo, S. *et al.* Stability and structural recovery of the tetramerization domain of p53-R337H mutant induced by a designed templating ligand. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16426-16431, doi:10.1073/pnas.0805658105 (2008).

187   Chini, M. G. *et al.* Conformationally locked calixarene-based histone deacetylase inhibitors.

Org. Lett. **12**, 5382-5385, doi:10.1021/ol102420b (2010).

188   Tsou, L. K., Chen, C. H., Dutschman, G. E., Cheng, Y. C. & Hamilton, A. D. Blocking HIV-1 entry by a gp120 surface binding inhibitor. *Bioorg. Med. Chem. Lett.* **22**, 3358-3361, doi:10.1016/j.bmcl.2012.02.079 (2012).

189   Tsou, L. K. *et al.* Discovery of a synthetic dual inhibitor of HIV and HCV infection based on a tetrabutoxy-calix[4]arene scaffold. *Bioorg. Med. Chem. Lett.* **20**, 2137-2139, doi:10.1016/j.bmcl.2010.02.043 (2010).

190   Consoli, G. M. *et al.* Design and synthesis of a multivalent fluorescent folate-calix[4]arene conjugate: cancer cell penetration and intracellular localization. *Org. Biomol. Chem.* **13**, 3298-3307, doi:10.1039/c4ob02333a (2015).

191   Leamon, C. P. & Low, P. S. Delivery of macromolecules into living cells: a method that exploits folate receptor endocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5572-5576, doi:10.1073/pnas.88.13.5572 (1991).

192   Rescifina, A. *et al.* DNA Recognition with Polycyclic-Aromatic-Hydrocarbon-Presenting Calixarene Conjugates. *European Journal of Organic Chemistry* **2014**, 7605-7613, doi:10.1002/ejoc.201403050 (2014).

193   Guan, A. J. *et al.* Effects of loops and nucleotides in G-quadruplexes on their interaction with an azacalixarene, methylazacalix[6]pyridine. *J. Phys. Chem. B* **115**, 12584-12590, doi:10.1021/jp204154m (2011).

194   Guan, A. J. *et al.* Stabilizing G-quadruplex DNA by methylazacalix[n]pyridine through shape-complementary interaction. *Bioorg. Med. Chem. Lett.* **26**, 609-612, doi:10.1016/j.bmcl.2015.11.062 (2016).

195   Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. & Feigon, J. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3745-3749, doi:10.1073/pnas.90.8.3745 (1993).

196   Santos, D. *et al.* Cell proliferation effects of calix[4]arene derivatives. *Tetrahedron* **71**, 7593-7599, doi:10.1016/j.tet.2015.07.077 (2015).

197   Karthauser, Z. *A New Approach to Drug Delivery: Non-Peptidic, High Load Macrocyclic Alternatives to Cell Penetrating Peptides* Doctor of Philosophy thesis, University of East Anglia, (2013).

198   Huisgen, R., Szeimies, G. & Möbius, L. 1.3-Dipolare Cycloadditionen, XXXII. Kinetik der Additionen organischer Azide an CC-Mehrfachbindungen. *Chemische Berichte* **100**, 2494-2507, doi:10.1002/cber.19671000806 (1967).

199   Tornoe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **67**, 3057-3064, doi:10.1021/jo011148j (2002).

200   Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **41**, 2596-2599, doi:10.1002/1521-3773(20020715)41:14<2596::aid-anie2596>3.0.co;2-4 (2002).

201   Kacprzak, K., Skiera, I., Piasecka, M. & Paryzek, Z. Alkaloids and Isoprenoids Modification by Copper(I)-Catalyzed Huisgen 1,3-Dipolar Cycloaddition (Click Chemistry): Toward New Functions and Molecular Architectures. *Chem. Rev.* **116**, 5689-5743, doi:10.1021/acs.chemrev.5b00302 (2016).

202   Thirumurugan, P., Matosiuk, D. & Jozwiak, K. Click chemistry for drug development and diverse chemical-biology applications. *Chem. Rev.* **113**, 4905-4979, doi:10.1021/cr200409f (2013).

203   Cheng, C. Y. *et al.* Identification of a new bis-amino acid glycoside selectively toxic to multiple myeloma cells. *Carbohydr. Res.* **394**, 39-42, doi:10.1016/j.carres.2014.05.014 (2014).

204   Bryden, F. *et al.* Regioselective and stoichiometrically controlled conjugation of photodynamic sensitizers to a HER2 targeting antibody fragment. *Bioconjug. Chem.* **25**, 611-617, doi:10.1021/bc5000324 (2014).

205   McKay, C. S. & Finn, M. G. Click chemistry in complex mixtures: bioorthogonal bioconjugation. *Chem. Biol.* **21**, 1075-1101, doi:10.1016/j.chembiol.2014.09.002 (2014).

206   Avti, P. K., Maysinger, D. & Kakkar, A. Alkyne-azide "click" chemistry in designing nanocarriers for applications in biology. *Molecules* **18**, 9531-9549, doi:10.3390/molecules18089531 (2013).

207   Pasini, D. The click reaction as an efficient tool for the construction of macrocyclic structures. *Molecules* **18**, 9512-9530, doi:10.3390/molecules18089512 (2013).

208   Himo, F. *et al.* Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. *J. Am. Chem. Soc.* **127**, 210-216, doi:10.1021/ja0471525 (2005).

209   Park, S. Y. *et al.* A pyrenyl-appended triazole-based calix[4]arene as a fluorescent sensor for Cd2+ and Zn2+. *J. Org. Chem.* **73**, 8212-8218, doi:10.1021/jo8012918 (2008).

210   Agalave, S. G., Maujan, S. R. & Pore, V. S. Click chemistry: 1,2,3-triazoles as pharmacophores. *Chem. Asian. J.* **6**, 2696-2718, doi:10.1002/asia.201100432 (2011).

211   Hong, V., Steinmetz, N. F., Manchester, M. & Finn, M. G. Labeling live cells by copper-catalyzed alkyne--azide click chemistry. *Bioconjug. Chem.* **21**, 1912-1916, doi:10.1021/bc100272z (2010).

212   Prescher, J. A. & Bertozzi, C. R. Chemistry in living systems. *Nat. Chem. Biol.* **1**, 13-21, doi:10.1038/nchembio0605-13 (2005).

213   van Swieten, P. F., Leeuwenburgh, M. A., Kessler, B. M. & Overkleeft, H. S. Bioorthogonal organic chemistry in living cells: novel strategies for labeling biomolecules. *Org. Biomol. Chem.* **3**, 20-27, doi:10.1039/b412558d (2005).

214   Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **39**, 1272, doi:10.1039/b901970g (2010).

215   Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed.* **48**, 9879-9883, doi:10.1002/anie.200905087 (2009).

216   Besanceney-Webler, C. *et al.* Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study. *Angew. Chem. Int. Ed.* **50**, 8051-8056, doi:10.1002/anie.201101817 (2011).

217   Kennedy, D. C. *et al.* Cellular consequences of copper complexes used to catalyze bioorthogonal click reactions. *J. Am. Chem. Soc.* **133**, 17993-18001, doi:10.1021/ja2083027 (2011).

218   Jewett, J. C., Sletten, E. M. & Bertozzi, C. R. Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *J. Am. Chem. Soc.* **132**, 3688-3690, doi:10.1021/ja100014q (2010).

219   de Almeida, G., Sletten, E. M., Nakamura, H., Palaniappan, K. K. & Bertozzi, C. R. Thiacycloalkynes for copper-free click chemistry. *Angew. Chem. Int. Ed.* **51**, 2443-2447,

doi:10.1002/anie.201106325 (2012).

220   McLuckie, K. I. *et al.* G-quadruplex-binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. *J. Am. Chem. Soc.* **133**, 2658-2663, doi:10.1021/ja109474c (2011).

221   Gutsche, C. D. & Iqbal, M. p-tert-BUTYLCALIX[4]ARENE. *Org. Synth.* **68**, 234, doi:10.15227/orgsyn.068.0234 (1990).

222   Gutsche, C. D., Iqbal, M. & Stewart, D. Calixarenes. 19. Syntheses procedures for p-tert-butylcalix[4]arene. *J. Org. Chem.* **51**, 742-745, doi:10.1021/jo00355a033 (1986).

223   Gutsche, C. D., Dhawan, B., No, K. H. & Muthukrishnan, R. Calixarenes. 4. The synthesis, characterization, and properties of the calixarenes from p-tert-butylphenol. *J. Am. Chem. Soc.* **103**, 3782-3792, doi:10.1021/ja00403a028 (1981).

224   Groenen, L. C. *et al.* The 1,2-alternate conformation of calix[4]arenes: a rare conformation? Dynamic 1H NMR studies of flexible tetraalkylated calix[4]arenes. *J. Am. Chem. Soc.* **113**, 2385-2392, doi:10.1021/ja00007a006 (1991).

225   Danil de Namor, A. F., Cleverley, R. M. & Zapata-Ormachea, M. L. Thermodynamics of Calixarene Chemistry. *Chem. Rev.* **98**, 2495-2526, doi:10.1021/cr970095w (1998).

226   Iwamoto, K., Araki, K. & Shinkai, S. Syntheses of all possible conformational isomers of O-alkyl-p-t-butylcalix[4]arenes. *Tetrahedron* **47**, 4325-4342, doi:10.1016/s0040-4020(01)87102-7 (1991).

227   Kumar, S., Varadarajan, R., Chawla, H. M., Hundal, G. & Hundal, M. S. Preparation of p-nitrocalix[n]arene methyl ethers via ipso-nitration and crystal structure of tetramethoxytetra-p-nitrocalix[4]arene. *Tetrahedron* **60**, 1001-1005, doi:10.1016/j.tet.2003.11.057 (2004).

228   Klimentová, J. & Vojtíšek, P. New receptors for anions in water: Synthesis, characterization, X-ray structures of new derivatives of 5,11,17,23-tetraamino-25,26,27,28-tetrapropyloxycalix[4]arene. *Journal of Molecular Structure* **826**, 48-63, doi:10.1016/j.molstruc.2006.04.016 (2007).

229   Sansone, F., Chierici, E., Casnati, A. & Ungaro, R. Thiourea-linked upper rim calix[4]arene neoglycoconjugates: synthesis, conformations and binding properties. *Org. Biomol. Chem.* **1**, 1802-1809, doi:10.1039/b301595e (2003).

230   Saadioui, M., Shivanyuk, A., Böhmer, V. & Vogt, W. Selective N-Protection of a Tetraamino Calix[4]arene Tetraether. *J. Org. Chem.* **64**, 3774-3777, doi:10.1021/jo982524b (1999).

231   Baldini, L. *et al.* CO2 capture by multivalent amino-functionalized calix[4]arenes: self-assembly, absorption, and QCM detection studies. *J. Org. Chem.* **76**, 3720-3732, doi:10.1021/jo200650f (2011).

232   Chetcuti, M. J. *et al.* Synthesis of mono-, di- and tetra-alkyne functionalized calix[4]arenes: reactions of these multipodal ligands with dicobalt octacarbonyl to give complexes which contain up to eight cobalt atoms. *Dalton Trans.*, 2999-3008, doi:10.1039/b821144b (2009).

233   Matthews, S. E. *et al.* Conformationally Mobile Wide Rim Carbamoylmethylphosphine Oxide (CMPO)-Calixarenes. *Journal für praktische Chemie* **341**, 264-273, doi:10.1002/(sici)1521-3897(199904)341:3<264::aid-prac264>3.0.co;2-f (1999).

234   Lalor, R., Baillie-Johnson, H., Redshaw, C., Matthews, S. E. & Mueller, A. Cellular uptake of a fluorescent calix[4]arene derivative. *J. Am. Chem. Soc.* **130**, 2892-2893, doi:10.1021/ja0782596 (2008).

235   Mees, M. *et al.* Partially Hydrolyzed Poly(n-propyl-2-oxazoline): Synthesis, Aqueous Solution

Properties, and Preparation of Gene Delivery Systems. *Biomacromolecules* **17**, 3580-3590, doi:10.1021/acs.biomac.6b01088 (2016).

236   Machado, A. H. *et al.* Interactions between DNA and nonionic ethylene oxide surfactants are predominantly repulsive. *Langmuir : the ACS journal of surfaces and colloids* **26**, 13102-13109, doi:10.1021/la101403j (2010).

237   Ganguly, K. K., Pal, S., Moulik, S. & Chatterjee, A. Integrins and metastasis. *Cell Adh. Migr.* **7**, 251-261, doi:10.4161/cam.23840 (2013).

238   Desgrosellier, J. S. & Cheresh, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nat. Rev. Cancer* **10**, 9-22, doi:10.1038/nrc2748 (2010).

239   Gupta, G. P. & Massague, J. Cancer metastasis: building a framework. *Cell* **127**, 679-695, doi:10.1016/j.cell.2006.11.001 (2006).

240   Gao, Y. *et al.* The ECM-cell interaction of cartilage extracellular matrix on chondrocytes. *Biomed. Res. Int.* **2014**, 648459, doi:10.1155/2014/648459 (2014).

241   Danhier, F., Le Breton, A. & Preat, V. RGD-based strategies to target alpha(v) beta(3) integrin in cancer therapy and diagnosis. *Mol. Pharm.* **9**, 2961-2973, doi:10.1021/mp3002733 (2012).

242   Friedlander, M. *et al.* Definition of Two Angiogenic Pathways by Distinct alpha(v) Integrins. *Science* **270**, 1500-1502, doi:10.1126/science.270.5241.1500 (1995).

243   Boudreau, N. J. & Varner, J. A. The homeobox transcription factor Hox D3 promotes integrin alpha5beta1 expression and function during angiogenesis. *J. Biol. Chem.* **279**, 4862-4868, doi:10.1074/jbc.M305190200 (2004).

244   Takada, Y., Ye, X. & Simon, S. The integrins. *Genome Biol.* **8**, 215, doi:10.1186/gb-2007-8-5-215 (2007).

245   Avraamides, C. J., Garmy-Susini, B. & Varner, J. A. Integrins in angiogenesis and lymphangiogenesis. *Nat. Rev. Cancer* **8**, 604-617, doi:10.1038/nrc2353 (2008).

246   Brooks, P., Clark, R. & Cheresh, D. Requirement of vascular integrin alpha v beta 3 for angiogenesis. *Science* **264**, 569-571, doi:10.1126/science.7512751 (1994).

247   Brooks, P. C. *et al.* Integrin αvβ3 antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels. *Cell* **79**, 1157-1164, doi:10.1016/0092-8674(94)90007-8 (1994).

248   Brooks, P. C. *et al.* Antiintegrin alpha v beta 3 blocks human breast cancer growth and angiogenesis in human skin. *J. Clin. Invest.* **96**, 1815-1822, doi:10.1172/JCI118227 (1995).

249   Takayama, S. *et al.* The relationship between bone metastasis from human breast cancer and integrin alpha(v)beta(3) expression. *Anticancer Res.* **25**, 79-83 (2005).

250   Furger, K. A. *et al.* Beta(3) integrin expression increases breast carcinoma cell responsiveness to the malignancy-enhancing effects of osteopontin. *Mol. Cancer Ther.* **1**, 810-819 (2003).

251   Sheldarke, H. M. & Patterson, L. H. Function and antagonism of beta3 integrins in the development of cancer therapy. *Curr. Cancer Drug Targets* **9**, 519-540, doi:10.2174/156800909788486713 (2009).

252   Li, S. G. *et al.* Correlation of integrin beta3 mRNA and vascular endothelial growth factor protein expression profiles with the clinicopathological features and prognosis of gastric carcinoma. *World J Gastroenterol* **14**, 421-427, doi:10.3748/wjg.14.421 (2008).

253   Hsu, M. *et al.* Adenoviral Gene Transfer of β3 Integrin Subunit Induces Conversion from Radial to Vertical Growth Phase in Primary Human Melanoma. *Am. J. Pathol.* **153**, 1435-1442, doi:10.1016/s0002-9440(10)65730-6 (1998).

254   Petitclerc, E. *et al.* Integrin alpha-v-beta-3 Promotes M21 Melanoma Growth in Human Skin by

Regulating Tumor Cell Survival. *Cancer Res.* **59**, 2724-2730 (1999).

255   Ricono, J. M. *et al.* Specific cross-talk between epidermal growth factor receptor and integrin alphavbeta5 promotes carcinoma cell invasion and metastasis. *Cancer Res.* **69**, 1383-1391, doi:10.1158/0008-5472.CAN-08-3612 (2009).

256   Reynolds, A. R. *et al.* Stimulation of tumor growth and angiogenesis by low concentrations of RGD-mimetic integrin inhibitors. *Nat. Med.* **15**, 392-400, doi:10.1038/nm.1941 (2009).

257   Seker, A. *et al.* Expression of integrins in cerebral arteriovenous and cavernous malformations. *Neurosurgery* **58**, 159-168 (2006).

258   Pierschbacher, M. D., Hayman, E. G. & Ruoslahti, E. Location of the cell-attachment site in fibronectin with monoclonal antibodies and proteolytic fragments of the molecule. *Cell* **26**, 259-267, doi:10.1016/0092-8674(81)90308-1 (1981).

259   Pierschbacher, M. D. & Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **309**, 30-33, doi:10.1038/309030a0 (1984).

260   Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. Identification and isolation of a 140 kd cell surface glycoprotein with properties expected of a fibronectin receptor. *Cell* **40**, 191-198, doi:10.1016/0092-8674(85)90322-8 (1985).

261   Ruoslahti, E. & Pierschbacher, M. New perspectives in cell adhesion: RGD and integrins. *Science* **238**, 491-497, doi:10.1126/science.2821619 (1987).

262   Albelda, S. M. & Buck, C. A. Integrins and other cell adhesion molecules. *FASEB J.* **4**, 2868-2880 (1990).

263   Humphries, M. J. The molecular basis and specificity of integrin-ligand interactions. *J. Cell Sci.* **97**, 585-592 (1990).

264   Ruoslahti, E. Integrins. *J. Clin. Invest.* **87**, 1-5, doi:10.1172/JCI114957 (1991).

265   Hynes, R. O. Integrins:Bidirectional, allosteric signaling machines. *Cell* **110**, 673-687, doi:10.1016/s0092-8674(02)00971-6 (2002).

266   Colombo, M. & Bianchi, A. Click chemistry for the synthesis of RGD-containing integrin ligands. *Molecules* **15**, 178-197, doi:10.3390/molecules15010178 (2010).

267   Joo, S. H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **20**, 19-26, doi:10.4062/biomolther.2012.20.1.019 (2012).

268   Koivunen, E., Gay, D. A. & Ruoslahti, E. Selection of peptides binding to the alpha 5 beta 1 integrin from phage display library. *J. Biol. Chem.* **268**, 20205-20210 (1993).

269   Heckmann, D. & Kessler, H. Design and Chemical Synthesis of Integrin Ligands. *Method Enzymol.* **426**, 463-503, doi:10.1016/s0076-6879(07)26020-3 (2007).

270   Mas-Moruno, C., Rechenmacher, F. & Kessler, H. Cilengitide: The First Anti-Angiogenic Small Molecule Drug Candidate. Design, Synthesis and Clinical Evaluation. *Anticancer Agents Med. Chem.* **10**, 753-768, doi:10.2174/187152010794728639 (2010).

271   Dechantsreiter, M. A. *et al.* N-Methylated cyclic RGD peptides as highly active and selective alpha(V)beta(3) integrin antagonists. *J. Med. Chem.* **42**, 3033-3040, doi:10.1021/jm970832g (1999).

272   Taga, T. *et al.* alpha v-Integrin antagonist EMD 121974 induces apoptosis in brain tumor cells growing on vitronectin and tenascin. *International journal of cancer* **98**, 690-697, doi:10.1002/ijc.10265 (2002).

273   Paolillo, M., Russo, M. A., Serra, M., Colombo, L. & Schinelli, S. Small molecule integrin antagonists   in   cancer   therapy.   *Mini   Rev.   Med.   Chem.*   **9**,   1439-1446,

doi:10.2174/138955709789957404 (2009).

274   Oliveira-Ferrer, L. *et al.* Cilengitide induces cellular detachment and apoptosis in endothelial and glioma cells mediated by inhibition of FAK/src/AKT pathway. *J. Exp. Clin. Cancer Res.* **27**, 86, doi:10.1186/1756-9966-27-86 (2008).

275   Stupp, R. *et al.* Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *The Lancet Oncology* **15**, 1100-1108, doi:10.1016/s1470-2045(14)70379-1 (2014).

276   Kim, K. B. *et al.* A randomized phase II study of cilengitide (EMD 121974) in patients with metastatic melanoma. *Melanoma research* **22**, 294-301, doi:10.1097/CMR.0b013e32835312e4 (2012).

277   Vansteenkiste, J. *et al.* Cilengitide combined with cetuximab and platinum-based chemotherapy as first-line treatment in advanced non-small-cell lung cancer (NSCLC) patients: results of an open-label, randomized, controlled phase II study (CERTO). *Ann. Oncol.* **26**, 1734-1740, doi:10.1093/annonc/mdv219 (2015).

278   Alva, A. *et al.* Phase II study of cilengitide (EMD 121974, NSC 707544) in patients with non-metastatic castration resistant prostate cancer, NCI-6735. A study by the DOD/PCF prostate cancer clinical trials consortium. *Invest. New Drugs* **30**, 749-757, doi:10.1007/s10637-010-9573-5 (2012).

279   Bradley, D. A. *et al.* Cilengitide (EMD 121974, NSC 707544) in asymptomatic metastatic castration resistant prostate cancer patients: a randomized phase II trial by the prostate cancer clinical trials consortium. *Invest. New Drugs* **29**, 1432-1440, doi:10.1007/s10637-010-9420-8 (2011).

280   Ovadia, O. *et al.* Improvement of drug-like properties of peptides: the somatostatin paradigm. *Expert. Opin. Drug Discov.* **5**, 655-671, doi:10.1517/17460441.2010.493935 (2010).

281   Biron, E. *et al.* Improving oral bioavailability of peptides by multiple N-methylation: somatostatin analogues. *Angew. Chem. Int. Ed.* **47**, 2595-2599, doi:10.1002/anie.200705797 (2008).

282   Haubner, R. *et al.* Structural and Functional Aspects of RGD-Containing Cyclic Pentapeptides as Highly Potent and Selective Integrin αVβ3Antagonists. *J. Am. Chem. Soc.* **118**, 7461-7472, doi:10.1021/ja9603721 (1996).

283   Dijkgraaf, I. *et al.* Synthesis of DOTA-conjugated multivalent cyclic-RGD peptide dendrimers via 1,3-dipolar cycloaddition and their biological evaluation: implications for tumor targeting and tumor imaging purposes. *Org. Biomol. Chem.* **5**, 935-944, doi:10.1039/b615940k (2007).

284   Guan, X. *et al.* Cyclic RGD targeting nanoparticles with pH sensitive polymer–drug conjugates for effective treatment of melanoma. *RSC Adv.* **4**, 55187-55194, doi:10.1039/c4ra08537j (2014).

285   Gaertner, F. C., Kessler, H., Wester, H. J., Schwaiger, M. & Beer, A. J. Radiolabelled RGD peptides for imaging and therapy. *Eur. J. Nucl. Med. Mol. Imaging* **39** S126-138, doi:10.1007/s00259-011-2028-1 (2012).

286   Battistini, L. *et al.* 4-Aminoproline-based arginine-glycine-aspartate integrin binders with exposed ligation points: practical in-solution synthesis, conjugation and binding affinity evaluation. *Org. Biomol. Chem.* **7**, 4924-4935, doi:10.1039/b914836a (2009).

287   Belvisi, L. *et al.* Potent Integrin Antagonists from a Small Library of RGD-Including Cyclic Pseudopeptides. *Org. Lett.* **3**, 1001-1004, doi:10.1021/ol007049u (2001).

288   Arosio, D., Manzoni, L., Araldi, E. M. & Scolastico, C. Cyclic RGD functionalized gold nanoparticles for tumor targeting. *Bioconjug. Chem.* **22**, 664-672, doi:10.1021/bc100448r (2011).

289   White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509-524, doi:10.1038/nchem.1062 (2011).

290   Temming, K., Schiffelers, R. M., Molema, G. & Kok, R. J. RGD-based strategies for selective delivery of therapeutics and imaging agents to the tumour vasculature. *Drug Resist. Updat.* **8**, 381-402, doi:10.1016/j.drup.2005.10.002 (2005).

291   Koivunen, E., Wang, B. & Ruoslahti, E. Phage Libraries Displaying Cyclic Peptides with Different Ring Sizes: Ligand Specificities of the RGD-Directed Integrins. *Nature Biotechol.* **13**, 265-270, doi:10.1038/nbt0395-265 (1995).

292   Assa-Munt, N., Jia, X., Laakkonen, P. & Ruoslahti, E. Solution Structures and Integrin Binding Activities of an RGD Peptide with Two Isomers. *Biochemistry* **40**, 2373-2378, doi:10.1021/bi002101f (2001).

293   Chen, H., Niu, G., Wu, H. & Chen, X. Clinical Application of Radiolabeled RGD Peptides for PET Imaging of Integrin alphavbeta3. *Theranostics* **6**, 78-92, doi:10.7150/thno.13242 (2016).

294   Haubner, R. *et al.* Noninvasive visualization of the activated alphavbeta3 integrin in cancer patients by positron emission tomography and [18F]Galacto-RGD. *PLoS medicine* **2**, e70, doi:10.1371/journal.pmed.0020070 (2005).

295   Haubner, R. *et al.* [18F]Galacto-RGD: synthesis, radiolabeling, metabolic stability, and radiation dose estimates. *Bioconjug. Chem.* **15**, 61-69, doi:10.1021/bc034170n (2004).

296   Beer, A. J., Kessler, H., Wester, H. J. & Schwaiger, M. PET imaging of integrin αVβ3 expression. *Cancer Metastasis Rev.* **27**, 631-644, doi:10.7150/thno/v01p0048 (2008).

297   Doss, M. *et al.* Biodistribution and radiation dosimetry of the integrin marker 18F-RGD-K5 determined from whole-body PET/CT in monkeys and humans. *J. Nucl. Med.* **53**, 787-795, doi:10.2967/jnumed.111.088955 (2012).

298   Kim, J. H. *et al.* Whole-body distribution and radiation dosimetry of (68)Ga-NOTA-RGD, a positron emission tomography agent for angiogenesis imaging. *Cancer Biother. Radiopharm.* **27**, 65-71, doi:10.1089/cbr.2011.1061 (2012).

299   Ye, Y. Integrin Targeting for Tumor Optical Imaging. *Theranostics* **1**, 102-126, doi:10.7150/thno/v01p0102 (2011).

300   Gurfinkel, M., Ke, S., Wang, W., Li, C. & Sevick-Muraca, E. M. Quantifying molecular specificity of alphavbeta3 integrin-targeted optical contrast agents with dynamic optical imaging. *J. Biomed. Opt.* **10**, 034019, doi:10.1117/1.1924696 (2005).

301   Chen, X., Conti, P. S. & Moats, R. A. In vivo near-infrared fluorescence imaging of integrin alphavbeta3 in brain tumor xenografts. *Cancer Res.* **64**, 8009-8014, doi:10.1158/0008-5472.CAN-04-1956 (2004).

302   Cheng, Z., Wu, Y., Xiong, Z., Gambhir, S. S. & Chen, X. Near-infrared fluorescent RGD peptides for optical imaging of integrin alphavbeta3 expression in living mice. *Bioconjug. Chem.* **16**, 1433-1441, doi:10.1021/bc0501698 (2005).

303   Cai, W. *et al.* Peptide-labeled near-infrared quantum dots for imaging tumor vasculature in living subjects. *Nano Lett.* **6**, 669-676, doi:10.1021/nl052405t (2006).

304   Chen, G., Roy, I., Yang, C. & Prasad, P. N. Nanochemistry and Nanomedicine for Nanoparticle-based Diagnostics and Therapy. *Chem. Rev.* **116**, 2826-2885, doi:10.1021/acs.chemrev.5b00148

(2016).

305   Janssen, A. *et al.* Peptide-targeted PEG-liposomes in anti-angiogenic therapy. *Int. J. Pharm.* **254**, 55-58, doi:10.1016/s0378-5173(02)00682-8 (2003).

306   Schiffelers, R. *et al.* Anti-tumor efficacy of tumor vasculature-targeted liposomal doxorubicin. *Journal of Controlled Release* **91**, 115-122, doi:10.1016/s0168-3659(03)00240-2 (2003).

307   Danhier, F. *et al.* Targeting of tumor endothelium by RGD-grafted PLGA-nanoparticles loaded with paclitaxel. *J. Control. Release* **140**, 166-173, doi:10.1016/j.jconrel.2009.08.011 (2009).

308   Curnis, F., Gasparri, A., Sacchi, A., Longhi, R. & Corti, A. Coupling tumor necrosis factor-alpha with alphaV integrin ligands improves its antineoplastic activity. *Cancer Res.* **64**, 565-571, doi:10.1158/0008-5472 (2004).

309   Xu, H. M. *et al.* An RGD-modified endostatin-derived synthetic peptide shows antitumor activity in vivo. *Bioconjug. Chem.* **19**, 1980-1986, doi:10.1021/bc800132p (2008).

310   Dickerson, E. B. *et al.* Enhancement of the antiangiogenic activity of interleukin-12 by peptide targeted delivery of the cytokine to alphavbeta3 integrin. *Mol. Cancer Res.* **2**, 663-673 (2004).

311   Schiffelers, R. M. *et al.* Cancer siRNA therapy by tumor selective delivery with ligand-targeted sterically stabilized nanoparticle. *Nucleic Acids Res.* **32**, e149, doi:10.1093/nar/gnh140 (2004).

312   Han, H. D. *et al.* Targeted gene silencing using RGD-labeled chitosan nanoparticles. *Clin. Cancer Res.* **16**, 3910-3922, doi:10.1158/1078-0432.CCR-10-0005 (2010).

313   Wang, X. L. *et al.* Targeted systemic delivery of a therapeutic siRNA with a multifunctional carrier controls tumor proliferation in mice. *Mol. Pharm.* **6**, 738-746, doi:10.1021/mp800192d (2009).

314   Liu, X. Q., Song, W. J., Sun, T. M., Zhang, P. Z. & Wang, J. Targeted delivery of antisense inhibitor of miRNA for antiangiogenesis therapy using cRGD-functionalized nanoparticles. *Mol. Pharm.* **8**, 250-259, doi:10.1021/mp100315q (2011).

315   Baldini, L. *et al.* Solid-phase synthesis of linear and cyclic peptides containing a calix[4]arene amino acid. *Tetrahedron Lett.* **50**, 3450-3453, doi:10.1016/j.tetlet.2009.02.198 (2009).

316   Seward, G. K., Bai, Y., Khan, N. S. & Dmochowski, I. J. Cell-compatible, integrin-targeted cryptophane-129Xe NMR biosensors. *Chem. Sci.* **2**, 1103-1110, doi:10.1039/C1SC00041A (2011).

317   Seward, G. K., Wei, Q. & Dmochowski, I. J. Peptide-mediated cellular uptake of cryptophane. *Bioconjug. Chem.* **19**, 2129-2135, doi:10.1021/bc8002265 (2008).

318   Chen, X., Park, R., Shahinian, A. H., Bading, J. R. & Conti, P. S. Pharmacokinetics and tumor retention of 125I-labeled RGD peptide are improved by PEGylation. *Nucl. Med. Bio.* **31**, 11-19, doi:10.1016/j.nucmedbio.2003.07.003 (2004).

319   Chen, X. *et al.* MicroPET and autoradiographic imaging of breast cancer alpha v-integrin expression using 18F- and 64Cu-labeled RGD peptide. *Bioconjug. Chem.* **15**, 41-49, doi:10.1021/bc0300403 (2004).

320   van Berkel, S. S. *et al.* Application of metal-free triazole formation in the synthesis of cyclic RGD-DTPA conjugates. *Chembiochem : a European journal of chemical biology* **9**, 1805-1815, doi:10.1002/cbic.200800074 (2008).

321   ten Brink, H. T., Rijkers, D. T. & Liskamp, R. M. Synthesis of alkyne-bridged cyclic tripeptides toward constrained mimics of vancomycin. *J. Org. Chem.* **71**, 1817-1824, doi:10.1021/jo051933m (2006).

322   Verlinden, S. *et al.* Oxidative alpha,omega-diyne coupling as an approach towards novel peptidic

macrocycles. *Org. Biomol. Chem.* **13**, 9398-9404, doi:10.1039/c5ob01153a (2015).

323 Palomo, J. M. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC Adv.* **4**, 32658, doi:10.1039/c4ra02458c (2014).

324 Dai, X., Su, Z. & Liu, J. An improved synthesis of a selective αvβ3-integrin antagonist cyclo(-RGDfK-). *Tetrahedron Lett.* **41**, 6295-6298, doi:10.1016/s0040-4039(00)01060-1 (2000).

325 Chatterjee, J., Laufer, B. & Kessler, H. Synthesis of N-methylated cyclic peptides. *Nat. Protoc.* **7**, 432-444, doi:10.1038/nprot.2011.450 (2012).

326 Chan, W. C. & White, P. D. *Fmoc Solid Phase Synthesis* (Oxford University Press, 2000).

327 Sarin, V. K., Kent, S. B. H., Tam, J. P. & Merrifield, R. B. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. *Anal. Biochem.* **117**, 147-157, doi:10.1016/0003-2697(81)90704-1 (1981).

328 Kaiser, E., Colescott, R. L., Bossinger, C. D. & Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34**, 595-598, doi:10.1016/0003-2697(70)90146-6 (1970).

329 Knorr, R., Trzeciak, A., Bannwarth, W. & Gillessen, D. New coupling reagents in peptide chemistry. *Tetrahedron Lett.* **30**, 1927-1930, doi:10.1016/s0040-4039(00)99616-3 (1989).

330 Carpino, L. A. *et al.* The Uronium/Guanidinium Peptide Coupling Reagents: Finally the True Uronium Salts. *Angew. Chem. Int. Ed.* **41**, 441-445, doi:10.1002/1521-3773(20020201)41:3<441::aid-anie441>3.0.co;2-n (2002).

331 Vojkovsky, T. & Drake, B. Improved One-Pot Preparation of Tetramethylfluoroformamidinium Hexafluorophosphate. *Org. Prep. Proced. Int.* **29**, 497-498, doi:10.1080/00304949709355227 (1997).

332 El-Faham, A. & Albericio, F. Peptide coupling reagents, more than a letter soup. *Chem. Rev.* **111**, 6557-6602, doi:10.1021/cr100048w (2011).

333 Sheehan, J. C. & Hess, G. P. A New Method of Forming Peptide Bonds. *J. Am. Chem. Soc.* **77**, 1067-1068, doi:10.1021/j.1399-3011.1994.tb00176.x (1955).

334 Izdebski, J. A. N., Pachulska, M. & Orlowska, A. N-Cyclohexyl-N′-isopropylcarbodiimide: a hybrid that combines the structural features of DCC and DIC. *Int. J. Pept. Protein Res.* **44**, 414-419, doi:10.1111/j.1399-3011.1994.tb00176.x (1994).

335 Frérot, E., Coste, J., Pantaloni, A., Dufour, M.-N. & Jouin, P. PyBOP® and PyBroP: Two reagents for the difficult coupling of the α,α-dialkyl amino acid, Aib. *Tetrahedron* **47**, 259-270, doi:10.1016/s0040-4020(01)80922-4 (1991).

336 Sivakumar, K. *et al.* A fluorogenic 1,3-dipolar cycloaddition reaction of 3-azidocoumarins and acetylenes. *Org. Lett.* **6**, 4603-4606, doi:10.1021/ol047955x (2004).

337 Buttress, J. P. *et al.* "Janus" Calixarenes: Double-Sided Molecular Linkers for Facile, Multianchor Point, Multifunctional, Surface Modification. *Langmuir : the ACS journal of surfaces and colloids* **32**, 7806-7813, doi:10.1021/acs.langmuir.6b02222 (2016).

338 Saha, S., Santra, S., Akhuli, B. & Ghosh, P. [2]Rotaxane with multiple functional groups. *J. Org. Chem.* **79**, 11170-11178, doi:10.1021/jo502235z (2014).

339 Neumann, S., Döhler, D., Ströhl, D. & Binder, W. H. Chelation-assisted CuAAC in star-shaped polymers enables fast self-healing at low temperatures. *Polym. Chem.* **7**, 2342-2351, doi:10.1039/c5py01818h (2016).

340 Wellner, C. & Wagenknecht, H. A. Synthesis of DNA conjugates with metalated tetracationic

porphyrins by postsynthetic cycloadditions. *Org. Lett.* **16**, 1692-1695, doi:10.1021/ol500364j (2014).

341　Pearson, D. A., Blanchette, M., Baker, M. L. & Guindon, C. A. Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis. *Tetrahedron Lett.* **30**, 2739-2742, doi:10.1016/s0040-4039(00)99113-5 (1989).

342　Rosu, F., De Pauw, E. & Gabelica, V. Electrospray mass spectrometry to study drug-nucleic acids interactions. *Biochimie* **90**, 1074-1087, doi:10.1016/j.biochi.2008.01.005 (2008).

343　Howerton, S. B., Nagpal, A. & Williams, L. D. Surprising roles of electrostatic interactions in DNA-ligand complexes. *Biopolymers* **69**, 87-99, doi:10.1002/bip.10319 (2003).

344　Squire, C. J., Backer, L. J., Clark, G. R., Martin, R. F. & White, J. Structures of m-iodo Hoechst-DNA complexes in crystals with reduced solvent content: implications for minor groove binder drug design. *Nucleic Acids Res.* **28**, 1252-1258, doi:10.1093/nar/28.5.1252 (2000).

345　Jenkins, T. C. & Lane, A. N. AT selectivity and DNA minor groove binding: modelling, NMR and structural studies of the interactions of propamidine and pentamidine with d(CGCGAATTCGCG)2. *BBA-Gene Struct. Expr.* **1350**, 189-204, doi:10.1016/s0167-4781(96)00160-1 (1997).

346　Salgado, G. F., Cazenave, C., Kerkour, A. & Mergny, J.-L. G-quadruplex DNA and ligand interaction in living cells using NMR spectroscopy. *Chem. Sci.* **6**, 3314-3320, doi:10.1039/c4sc03853c (2015).

347　Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* **2**, 1849-1861, doi:10.1038/nprot.2007.249 (2007).

348　Nobbmann, U. *et al.* Dynamic light scattering as a relative tool for assessing the molecular integrity and stability of monoclonal antibodies. *Biotech. Genet. Eng. Rev.* **24**, 117-128, doi:10.1080/02648725.2007.10648095 (2007).

349　Son, S., Nam, J., Kim, J., Kim, S. & Kim, W. J. i-motif-driven Au nanomachines in programmed siRNA delivery for gene-silencing and photothermal ablation. *ACS nano* **8**, 5574-5584, doi:10.1021/nn5022567 (2014).

350　Khoukhi, M., Vaultier, M., Benalil, A. & Carboni, B. Curtius Rearrangement of ω-Azido Acid Chlorides: Access to the Corresponding ω-Azido Substituted Amines and Carbamates, Useful Building Blocks for Polyamine Syntheses. *Synthesis* **1996**, 483-487, doi:10.1055/s-1996-4438 (1996).

351　Fletcher, S. A. *et al.* A dual optical and nuclear imaging reagent for peptide labelling via disulfide bridging. *Org. Biomol. Chem.* **13**, 9559-9563, doi:10.1039/c5ob01468a (2015).