

University of Southampton Research Repository ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering, Science and Mathematics

School of Chemistry

**Synthesis of oligonucleotide analogues for use in
DNA nanostructures**

By

Adeline Durand

A thesis submitted for the degree of Doctor of Philosophy.

March 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS
SCHOOL OF CHEMISTRY

Doctor of Philosophy

**SYNTHESIS OF OLIGONUCLEOTIDE ANALOGUES FOR USE IN DNA
NANOSTRUCTURES**

by Adeline Durand

Thanks to its ability to form duplexes through selective base-pair recognition, DNA is a unique material for orderly self-assembled construction at the nanoscale. To develop a nanotechnology platform on a grid of addressable molecular building blocks using DNA node structures, DNA complexes need to be fixed onto surfaces. To fulfil this requirement on lipid membranes, phosphoramidites monomers modified with a cholesterol moiety and a spacer unit were synthesised. The hydrophobic spacer provides separation between the hydrophobic cholesterol moiety and the phosphate backbone of the DNA strand. For better anchorage of the network to a lipid surface, a hydrophilic spacer was also used to link the cholesterol to the nucleoside.

Solution studies demonstrated that the melting temperature (T_m) of the duplex with adjacent cholesterols on each strand is much higher than that of the unmodified duplex.

The reliable and highly selective copper(I)-catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC), the best known example of click chemistry, has proven to be of remarkably broad utility in synthetic chemistry and in nucleic acid chemistry in particular. CuAAC was exploited for the synthesis of a very stable double stranded catenane duplex. The catenane was formed from a single stranded cyclic template and its linear complement, using a third short oligonucleotide (ODN) as a helper for the circularisation of the second ODN. The copper(I)-catalysed cyclisation of oligonucleotides occurred by reaction between their terminal azide and their opposite terminal alkyne, to produce a 1,2,3-triazole linkage between the two reactants. The catenane was characterised by denaturing polyacrylamide gel electrophoresis, UV melting studies and enzyme digestion.

Contents

Abstracts	i
Contents	ii
Declaration	vi
Acknowledgements	vii
List of abbreviations	viii

Chapter 1 Introduction

1.1 DNA structure¹	2
1.1.1 Primary structure	2
1.1.2 Secondary structure	4
1.2 Nanotechnology	6
1.2.1 Introduction on nanotechnology	6
1.2.2 DNA and nanotechnology	6
1.2.2.1 Consideration of DNA-sequence design	7
1.2.2.2 Design and assembly	9
1.2.2.2.1 DNA toolbox: junctions and crossovers	9
1.2.2.2.2 DNA tiles	11
1.2.2.2.3 DNA objects	14
1.2.2.3 Applications	15
1.2.2.3.1 Mechanical devices	15
1.2.2.3.2 DNA assembly of biomolecules	17
1.2.2.3.3 DNA Scaffolds for nanoelectronics	18
1.2.2.3.4 DNA-based computation and algorithmic assembly	19
1.3 AMNA project	21
1.3.1 General	21
1.3.2 Lipophilic oligonucleotides	23
1.4 Click chemistry	24
1.4.1 Introduction	24

1.4.2	Huisgen reaction	27
1.4.3	Mecanism	29
1.4.4	Selected applications	30
1.4.4.1	Use in organic chemistry	30
1.4.4.2	Drug discovery	31
1.4.4.3	Bioconjugation	32
1.4.4.4	Click chemistry and DNA	35
1.5	Gel Electrophoresis	37
1.5.1	PAGE	37
1.5.2	Ferguson plot	38
1.6	Methods for studying the properties of nucleic acids	39
1.6.1	UV melting	39
Chapter 2	Synthesis and properties of hydrophobic monomers and oligonucleotides for Addressable Molecular Node Assembly	
2.1	Introduction	42
2.2	Cholesterol-linked benzyl alcohol monomer, 8	43
2.2.1	Synthesis	43
2.2.2	Biophysical studies	45
2.3	Synthesis of the thymidine monomer linked to cholesterol by a hydrophobic spacer, 13	47
2.3.1	First route	47
2.3.2	Second route	49
2.3.3	Biophysical studies	53
2.3.3.1	10-mers	53
2.3.3.2	18-mers	54
2.4	Synthesis of 6-(cholesteryloxycarbonyl)-hexanol phosphoramidite, 23	57
2.4.1	Synthesis	57
2.4.2	Biophysical studies	57
2.5	Synthesis of the thymidine monomer, linked to cholesterol by a hydrophilic spacer	58

2.5.1	Synthesis	59
2.5.1.1	Tosylation of hexaethylene glycol	59
2.5.1.2	Displacement of the tosyl group with phthalimide	59
2.5.1.3	Oxidation of the primary alcohol	60
2.5.1.4	Amide coupling reaction	61
2.5.1.5	Hydrazinolysis	63
2.5.1.6	Phosphitylation	66
2.6	Conclusions	68
2.7	Future work	69

Chapter 3 Synthesis of DNA nanostructures: use of click chemistry for the closure of catenanes

3.1	Introduction	71
3.2	Synthesis of oligonucleotides for click chemistry	72
3.2.1	Synthesis of alkyne oligonucleotides	72
3.2.2	Synthesis of azide oligonucleotides	73
3.3.	Click chemistry	74
3.3.1	First catenane	74
3.3.1.1	Cyclisation	74
3.3.1.2	Catenation of cyclic oligonucleotides with their linear complements	75
3.3.2	Ligation of shorter catenane sequence	79
3.3.3	Alternative sequence of the catenane	83
3.3.3.1	Catenation	83
3.3.3.2	Enzyme digestion	85
3.3.3.3	Native gel	87
3.3.3.4	Thermodynamic studies of the catenane	88
3.4	Conclusions	90
3.5	Future work	91

Chapter 4 Experimental

4.1	General experimental	93
4.2	Synthesis	95
4.3	Biophysical studies	136

4.3.1	Preparation of synthetic oligonucleotides	136
4.3.2	Oligonucleotides synthesised	137
4.3.3	Oligonucleotides characterisation	138
4.3.4	UV melting studies	139
4.3.5	Fluorescence melting studies	140
4.3.6	Click chemistry experimental	140
4.3.6.1	Synthesis of azidohexanamide-labeled oligonucleotides	140
4.3.6.2	Click cyclisation of ssDNA	141
4.3.6.3	Preparation of dsDNA catenane	141
4.3.6.4	Denaturing PAGE gel electrophoresis analysis and purification	142
4.3.6.5	Ferguson plot analysis	142
4.3.6.6	Native PAGE gel electrophoresis	143
4.3.6.7	EcoR1 digestion	143
References		144
Appendix		151

Declaration

I, Adeline Durand declare that this thesis entitled “Synthesis of oligonucleotides for Addressable Molecular Node Assembly and DNA Nanostructures” and the work presented within are both my own and have been generated by myself as the result of my own original research.

I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted the published work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- work described in Chapter 2 has been published in the paper below:

Durand A.; Brown T.; “Synthesis and properties of oligonucleotides containing a cholesterol thymidine monomer”, *Nucleosides, Nucleotides and Nucleic Acids*, **2007**, 26, 785-794.

DOI: 10.1080/15257770701501534

Signed:.....

Date: 9th March 2010

Acknowledgements

I would like to express my gratitude to my supervisor Professor Tom Brown for giving me the chance to do a PhD in England and for all his ideas and assistance without which the research would not have been possible. I also would like to thank Dr. Dorcas Brown and all at ATDBio LTd for their precious work in the preparation of the oligonucleotides used during my PhD.

Many thanks to Dr. John Langley and Julie Herniman, and Joan Street and Dr. Neils Wells for mass spectrometry and NMR facilities, and for providing an excellent quality service.

There are many people that I would like to acknowledge from the Brown group for their support and help throughout my three years of PhD. I would especially like to thank Simon Gerrard for being such a joyful and reliable PhD partner throughout these three years, Dr. Rohan Ranasinghe, for his intellectual acuity and proof reading, and Dr Naomi Hammond for proof reading this thesis. I am indebted to Dr. Afaf El-Sagheer for her invaluable help in my work in click chemistry. A huge thank you goes to the past and present members of the Brown group, for their great support and sense of humour and for making my life in Southampton so lively: Martina Banchelli, Dr. Noha Ben-Gaïed, Dr. Lavinia Brennan, Dr Imenne Bouamaied, Dr. Nittaya Gale, Dr. Guo-mei Peng, Lucy Hall, Dr. Hong Li, Harvey Lou, Dr Petr Kočalka, Jan Michels, Dr. Mastoura Edrees-Abdou, Dr. Vicky Powers, James Richardson, Korn Suprapprom, Radha Taylor, Dr. Sunil Vadhia and Dr. John Zhao. I must also mention the members of other groups Filipa, Martin, Stephen, Penny and Paul. Thank you for being such good friends.

This work was supported by the European Commission's 6th Framework Programme to which I am very grateful.

I am very thankful to my parents for their tireless emotional and financial support, in particular through the difficulties I had to face with my health.

Finally I would like thank Nick for his love and support during the challenging time of the writing up.

List of abbreviations

ABI	Applied Biosystems International
Ac	Acetyl
Ac₂O	Acetic anhydride
AcOH	Acetic acid
AMNA	addressable molecular node assembly
BAIB	[bis(acetoxy)-iodo]benzene
br	broad
Chol	cholesterol
COSY	correlated spectroscopy
DEPT	distortionless enhancement through polarisation transfer
DFT	density functional theory
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethyl sulfoxide
DMT	dimethoxytrityl

DNA	deoxyribonucleic acid
DX	double-crossover
EDC	<i>N</i> -(3-Dimethylaminopropyl)- <i>N</i> '-ethylcarbodiimide
EDTA	ethylenediamine tetraacetic acid
eq	equivalents
ES⁺	electrospray positive
EtOAc	ethyl acetate
FRET	fluorescence resonance energy transfert
IR	Infrared
HEG	hexaethylene glycol
HMQC	¹ H- ¹³ C COSY
HRMS	high resolution mass spectrometry
HPLC	high performance liquid chromatography
LRMS	low resolution mass spectrometry
MALDI-TOF	matrix assisted laser desorption-ionisation-time of flight
MeOH	methanol
NMR	nuclear magnetic resonance

ODN	oligonucleotide
R_f	retention factor
rt	room temperature
S_N	nucleophile substitution
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy
TFA	trifluoroacetyl
THF	tetrahydrofuran
TLC	thin layer chromatography
T_m	melting temperature
Ts	tosyl
TX	triple-crossover
UV	ultraviolet

CHAPTER 1

Introduction

1. Introduction

1.1 DNA structure¹

1.1.1 Primary structure

The nucleic acids DNA and RNA are macromolecules composed of nucleotide monomer subunits, linked together by a sugar-phosphate backbone. Each nucleotide consists of three building blocks: a furanose sugar, a phosphate moiety and a heterocyclic base. The heterocyclic bases in nucleic acids are derivatives of purine or pyrimidine. In DNA there are two bicyclic purine bases, adenine (A) and guanine (G); and two monocyclic pyrimidine bases, cytosine (C) and thymine (T). Nucleobases are covalently attached between a nitrogen atom and the C^{1'} of a 2'-deoxy-ribofuranose sugar (dR) to form the nucleosides, deoxyadenosine (dA), deoxyguanosine (dG), deoxycytidine (dC) and thymidine (T) respectively. The pyrimidine bases are covalently bonded *via* the N¹-atom, and the purine bases, *via* the N⁹-atom (Figure 1.1).

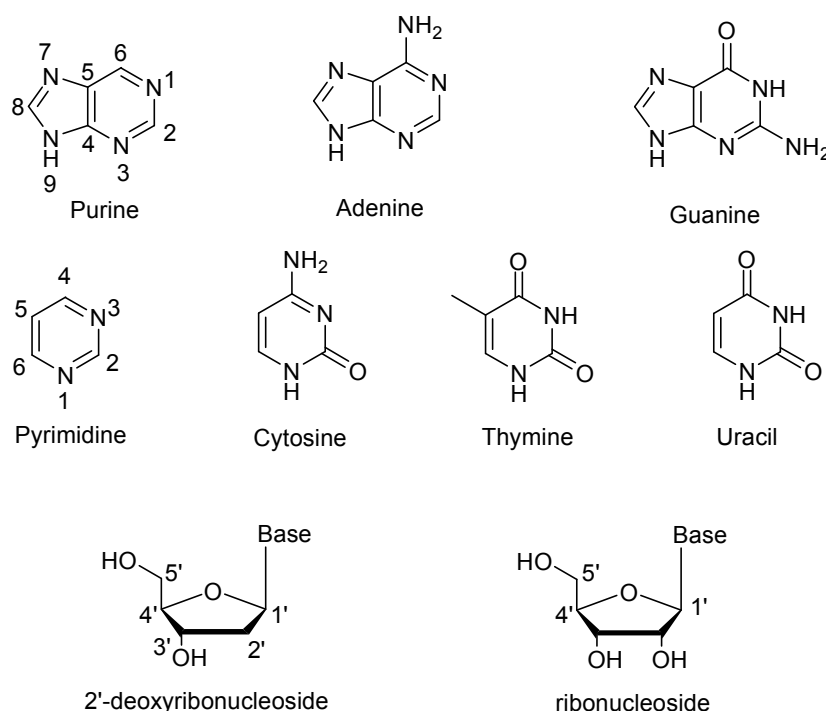


Figure 1.1: Structures of the purine and pyrimidine heterocyclic bases present in DNA and RNA; and deoxyribonucleoside and ribonucleoside.

For RNA, the sugar moiety is ribofuranose, namely it has a hydroxyl group in the 2'-position, and thymine is replaced by uracil, which lacks a methyl group at the 5-position (Figure 1.1).

Geometry of nucleotides within the macromolecule or polymer, and of the nucleosides, is governed by sugar and sugar-base conformations. The sugars can adopt two characteristic non-planar conformations where the furanose ring twists out of plane to minimise non-bonding interactions between its substituents: the C^{2'} endo (S type) and the C^{3'} endo (N-type) (Figure 1.2).

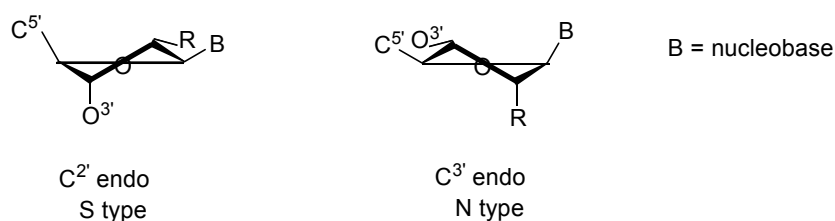


Figure 1.2: Structures of the C^{2'}-endo (S) and C^{3'}-endo (N) sugar puckers.

In standard orientation the base lies above the plane of the sugar to give a β N-glycosidic bond. There are two principal orientations of the base and sugar in a nucleoside, which are structurally important. These are known as the *anti* and *syn* conformations (Figure 1.3).

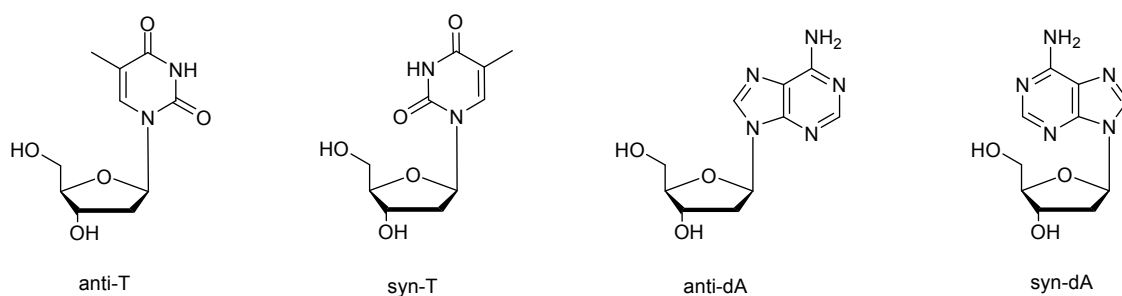


Figure 1.3: *Syn*- and *anti*- conformations of thymidine and deoxyadenosine nucleosides

A nucleotide is a nucleoside with a phosphate covalently bound to the 3' or 5'-end of the sugar, giving the oligonucleotide its directionality. At neutral pH each phosphate group carries a single negative charge; hence nucleic acids are highly charged polymers.

Convention denotes that the DNA sequence is recorded from the 5'-terminus to the 3'-terminus using the single letter abbreviations of the different nucleotides.

1.1.2 Secondary structure

The basic features of the double helix were deduced and published by James Watson and Francis Crick in 1953,² following extensive X-ray diffraction studies by their associates, Franklin and Gosling,³ and Wilkins *et al.*⁴ The two DNA strands were observed to be wound around each other in a helical path with the negatively charged sugar-phosphate backbone on the outside of the structure, which forms the hydrophilic part of the structure. The planar bases were found to lie perpendicular to the axis of the helix on the inside, stacking above one other *via* π - π interactions and forming the hydrophobic core of the structure. The four bases form Watson-Crick base pairs;⁵ G pairs with C *via* three intermolecular hydrogen bonds, and A pairs with T *via* two hydrogen-bonds (Figure 1.4).

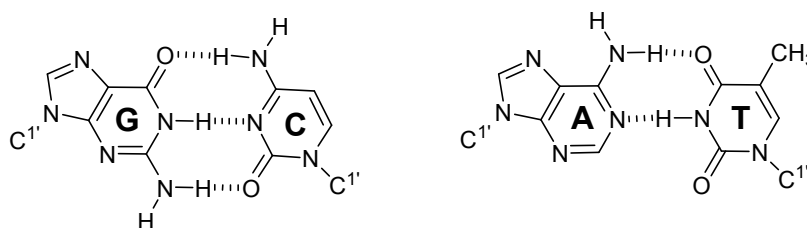


Figure 1.4: Structures of the GC and AT Watson-Crick base pairs. The GC base pair has three hydrogen bonds and the AT base pair has two.

The two strands are antiparallel as their 5'-3' phosphodiester backbones run in opposite directions (Figure 1.5).

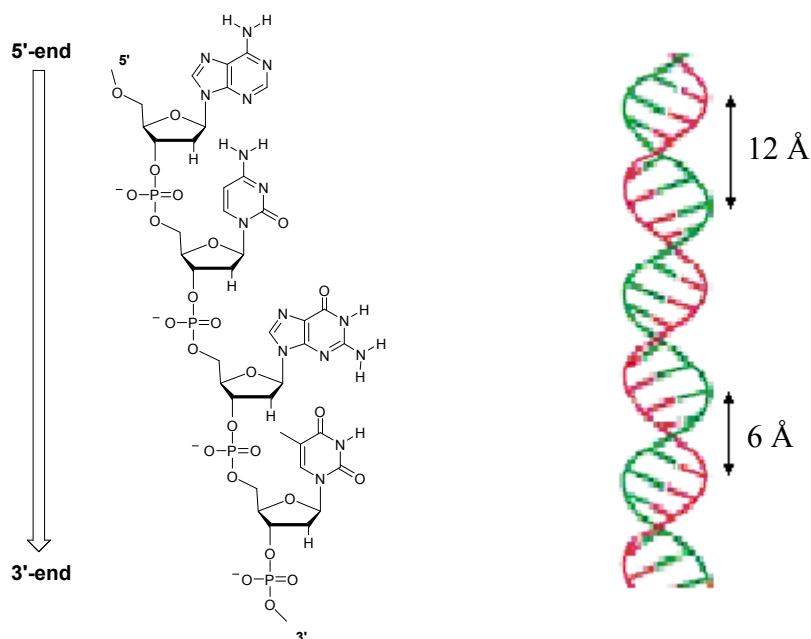


Figure 1.5: DNA single strand structure and B-DNA double helix showing the major and minor groove.

There are two principal conformations for the DNA double helix, A and B, both of which have been characterised by X-ray crystallography.

B-DNA, the most biologically relevant form, occurs in conditions of high humidity/low salt. It is an anti-parallel right-handed duplex with a 10 base pair (34 Å) per turn periodicity, in which the base pairs sit directly on the helical axis and the sugars adopt the $C^{2'}$ -*endo* pucker. The attachment of bases to the sugar-phosphate backbone through glycosidic bonds is asymmetrical, resulting in the formation of two different grooves, the major groove and the minor groove. The major groove is 12 Å wide and 9 Å deep; and the minor groove is 6 Å wide and 8 Å deep. Both are lined with potential hydrogen bond acceptors and donors but these are more exposed in the major groove.

The A-form duplex (A-DNA) is also an antiparallel right-handed helix with an 11 base pair per turn periodicity with sugar adopting the $C^{3'}$ -*endo* pucker, and occurring under conditions of low humidity/high salt. This form of duplex is most commonly adopted by DNA. RNA hybrids and RNA duplexes and the A-form is also adopted by DNA duplexes in G-rich sequences. The major groove is narrow and deep and the minor groove wide and shallow and the phosphate backbone is stiffer than that of B-DNA.

1.2 Nanotechnology

1.2.1 Introduction on nanotechnology

The first mention of nanotechnology (not yet using that name) occurred in a talk given by Richard Feynman in 1959, entitled “*There's Plenty of Room at the Bottom*”.

The term “Nanotechnology” was created by Tokyo Science University professor Norio Taniguchi in 1974 to describe the precision manufacture of materials with nanometre tolerances. In the 1980s the term was reinvented and its definition expanded by K Eric Drexler, particularly in his 1986 book *Engines of Creation: The Coming Era of Nanotechnology*.⁶ Since then the word “nanotechnology” has become very popular and is used to describe many types of research where the characteristic dimensions are less than about 1,000 nanometres.

At the nanometer scale, matter behaves differently to the meter scale. It often becomes more reactive and, for example, a material's electrical conductivity, strength and melting point may all change. Since the invention of Scanning Tunnelling Microscopy (STM) in the 1980s, researchers have developed increasingly powerful techniques for observing and manipulating surfaces at the nanoscale. This has led to the development of new materials and even to the conceptual design of molecular pumps and motors.⁷⁻¹⁰

A number of commercial applications have already reached the market. These include biocompatible medical implants, high-performance computer hard drives, scratch-resistant paints and self-sterilising surfaces.^{11,12} Many of these applications are the results of unexpected discoveries made in the course of fundamental research.

1.2.2 DNA and nanotechnology

An increasing number of researchers within nanoscience are using nucleic acids as building-blocks at the bottom-up approach towards novel materials with predictable 2D or 3D structures.¹³ DNA is indeed a promising candidate for achieving the desired structural formation and arrangement in nanosciences. Despite its simplicity the outstanding specificity of the A-T and G-C base-pairing system allows convenient programming. It is a building block with a very high information content. It has well-known diameter scale structural geometry, a 3.4 nm helical repeat and about 2 nm diameter (as described in section 1.1.2) and it has combined stiffness and flexibility; the

persistence length, that is, the length up to which polymers can be considered essentially rigid and straight, is about 50 nm, which corresponds to 150 base pairs in the double helix. Below this size, double-stranded DNA (dsDNA) behaves as a rather rigid polymer, on the other hand, single-stranded DNA (ssDNA) is very flexible. Furthermore, it is easy to synthesise using automated phosphoramidite chemistry,¹⁴ and to amplify *via* the polymerase chain reaction (PCR). Lastly, it is particularly noteworthy that nature provides us with a variety of highly specific enzymes which allow the processing of DNA material with atomic precision and accuracy at the Angstrom level (e.g. restriction endonucleases).

1.2.2.1 Consideration of DNA-sequence design

When planning to assemble materials by simple DNA-programmed processes the first thing to consider is how to design the sequences to obtain the desired assembly.¹⁵ Several major steps are required for the design and construction of DNA-based nanoarchitecture. Firstly, any requirements to be satisfied by the envisioned construct and its potential application should be identified. For example, it may eventually be wished to use the scaffold for the directed attachment of non-nucleic acid components, such as nanoparticles or proteins. In this case it needs to be determined whether the DNA scaffold should be rigid or flexible. The next question might concern whether the construct is to be used as a static scaffold or as a dynamic mechanical device that undergoes reversible or irreversible transitions between distinct and predictable conformations. It must also be determined whether symmetric or nonsymmetric patterns are to be produced by the assembly of the array, what kind of periodicity is desired, in how many dimensions the scaffold should grow, and whether this growth is to be terminated or potentially infinite.¹⁶

It is also necessary to tackle the conjugation of the material and DNA (or analogs). The power of automated oligonucleotide synthesis makes small oligonucleotides easily accessible and many DNA-modifiers are commercially available and can be incorporated during automated DNA-synthesis. These include phosphoramidites with modified nucleoside bases, fluorescent dyes, bio-labels such as biotin and a variety of functional chemical groups (FCG).¹⁷ Modifications can be incorporated at the 5' or 3'-

ends of the DNA sequence using a modifier with a phosphoramidite group for coupling with the 5' or 3'-OH of the DNA sequence and a protected FCG; but internal modifications within the sequence can also be included using the relevant protected FCG. It is relatively easy to conjugate DNA with organic molecules that carry functional groups which react with the FCGs.¹⁷ Biomolecules and chemically modified biomolecules can be conjugated with DNA in a similar manner (e.g. biotin for subsequent association with streptavidin).^{18,19} The inorganic material most frequently used in conjugation with DNA is nanoparticulate gold.²⁰ DNA sequences with a thiol modifier can react directly with the naked gold nanoparticle or, by ligand exchange, react with monolayer-covered particles to form thiolate-Au bonds.²¹

In addition to these considerations, it is required that DNA-sequences are of sufficient length. Typically 10- to 25-mer DNA sequences are used. A 10-mer dsDNA helix has a melting temperature between 25 and 35 °C. Guanine (G) and cytosine (C) interact *via* three hydrogen bonds leading to higher stabilization (*ca.* 6 kcal mol⁻¹) than the two-hydrogen bond interaction between thymine (T) and adenine (A) (*ca.* 4.5 kcal mol⁻¹) (see section 1.1.2).

Finally, the various different DNA oligomers designed have to be assembled. The way to self-assemble complex patterns can generally be divided into four categories. Traditionally, DNA-based nanostructures self-assemble from a set of unique oligonucleotides with carefully chosen base sequences. The oligonucleotides are mixed in buffer solution and annealed by heating the solution to a high temperature (around 95 °C) and slowly letting it cool to room temperature. During the cooling process, complementary sequences hybridize with each other and lead the nanostructure into a unique structure. The second method is hierarchical self-assembly. Selected groups of tiles are mixed in separate tubes and then combined sequentially. The algorithmic self-assembly programs the nanostructure with specific binding domains that can be specifically and cooperatively bound to form a complex pattern following an algorithm rule. A few years ago, a radically different approach towards DNA assembly was developed by Rothemund²² (and previously introduced by Yan *et al.*²³), which is based on folding a single, long “scaffold” strand into a desired shape with the help of multiple “staple” strands (Figure 1.6). In contrast to the traditional assembly approach, there is no need to observe strict stoichiometric ratios, and assembly occurs quite quickly and with an extraordinary yield. This is because after initial attachment of the staple strands

to the scaffold, folding occurs locally within the scaffold. When a large excess of staple strands are used, mismatches in the assembly are “healed” by strand displacement.²⁴ Many self-assembly systems with complex patterns actually use a combination of methods.

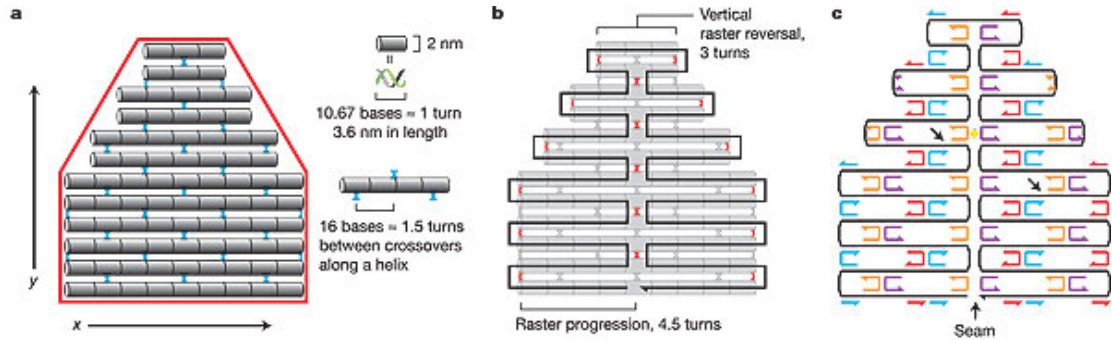


Figure 1.6: Design of DNA origami: **a**, A shape (red) approximated by parallel double helices joined by periodic crossovers (blue). **b**, A scaffold (black) runs through every helix and forms more crossovers (red). **c**, As first designed, more staples bind two helices and are 16-mers. Adapted with permissions from ref. 22.

1.2.2.2 Design and assembly

1.2.2.2.1 DNA toolbox: junctions and crossovers

The study of artificial DNA structures for applications in nanotechnology began in the early 80s when Seeman conducted pioneering work on the construction of artificial nucleic acid architectures combining sticky-end cohesion and branched DNA junction motifs.²⁵ The branched molecule used is an analog of a biological intermediate in genetic recombination, the Holliday junction. However, the naturally occurring Holliday junction is usually formed between two duplex molecules with the same sequence. This sequence symmetry destabilises the position of the branch point so that it is free to migrate. Seeman showed that by specifically designing sequences which were able to exchange strands at a single specified point and by breaking the sequence symmetry, this allowed the branch junction to migrate. This idea is illustrated in Figure 1.7, which shows that a four-arm branched DNA junction with complementary sticky ends could self-assemble into 2D square lattices.

These lattices are potentially useful as scaffolds with nanometer-spaced repetitive structural features.

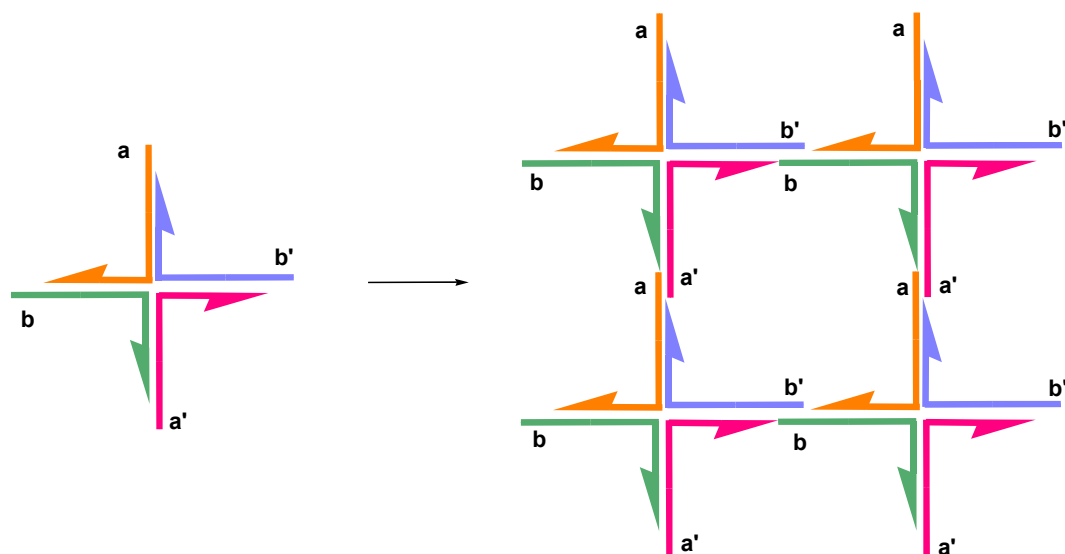


Figure 1.7: Sticky-ended assembly of branched molecules. A branched molecule is shown on the left with four sticky ends; a complementary to a' and b complementary to b' . Four of them are shown to self-assemble in a 2D lattice, with further sticky ends on the ends on the outside, so that an infinite lattice could be formed by the addition of further components (adapted from ref²⁶)

DNA duplexes or Holliday junction structures were found to be too flexible to achieve 2D or even 3D structures. To solve this problem, Seeman and co-workers constructed inherently rigid building blocks for DNA construction, the crossover tiles. Double crossover (DX) consist of two double-stranded helices, which interchange single strands at two crossover points.²⁷ DX tiles are usually designed such that they comprise sticky ends at their four arms which can be used for self-assembly. They became the initial building blocks for the construction of periodic assemblies and the formation of the first two-dimensional crystals of DNA tiles.²⁸ These large lattices provide multiple attachment sites, so the basic structure of a DX motif could possibly be addressed in many different ways. The use of paired crossovers greatly increases the stiffness of the tiles over that of linear dsDNA. Following the success of DX lattices, triple crossover (TX)²⁹ and paranemic crossover (PX)³⁰ were demonstrated and used in the construction of nanostructures. (Figure 1.8)

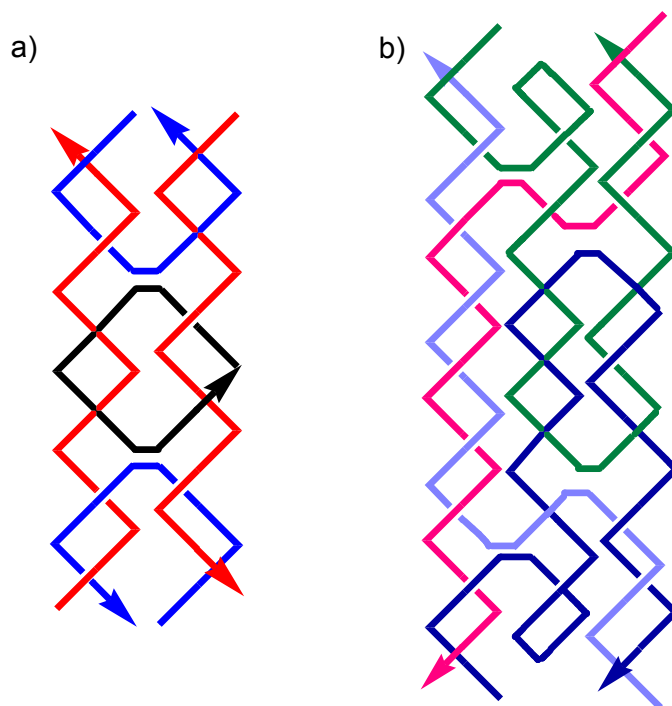


Figure 1.8: Key motifs of DNA Nanotechnology. a) DX is a double crossover molecule, resulting from a double exchange of DNA backbones; b) TX is a triple crossover that results from two successive double reciprocal exchanges.

1.2.2.2.2 DNA tiles

Crossover tiles have also been used as building blocks for the generation of larger motifs, which were then used as building blocks themselves. For example, a cyclic double-stranded supramolecular structure which included three flexible bulges was constructed using DX motifs only at first,³¹ and then DX and PX tiles.³² TX molecules were used to build linear arrays,³³ 2D lattices,²⁹ and DNA tubes.³⁴ Lattices with rhombus-like units have been made in which the helix crossing angles are about 60° . A set of rhombids was subsequently assembled into 1D “railroad tracks” as well as into 2D lattices (Figure 1.9).³⁵

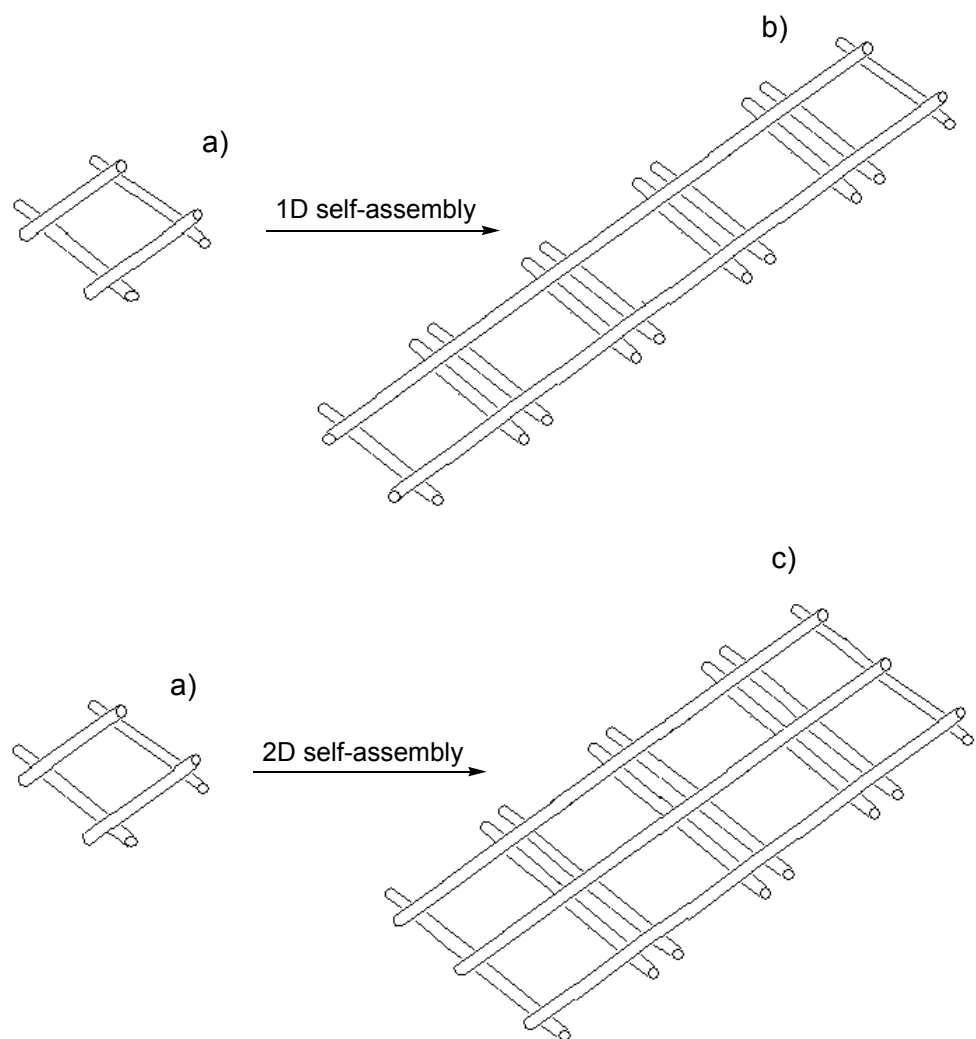


Figure 1.9: The rhombus motif. Combination of four junctions into a rhombus-like motif. Four molecules are combined. There are six turns of DNA in each helix, and four turns between crossover points, leading to one-turn overhangs on the ends (a), which can be used as a basic component for the self-assembly of a 1D linear railroad track-like arrangement (b), or a 2D latticework of DNA (c) (adapted from ref³⁵).

Mao *et al.* constructed an array of triangular tiles made of flexible four-arm DNA junctions,³⁶ and two versions of triangular tiles which form hexagonal patterns were also prototyped.^{37,38}

While all crossover tiles can be considered as flat and rigid building blocks, the structural variety of individual types offers a large range of options for modulating the spacing and connectivity of the building blocks and of the final superstructure desired.

Moreover, crossover motifs can be further modified and expanded to include additional functionalities.

Yan *et al.* constructed a DNA nanostructure, so called 4×4 tiles, consisting of four four-arm junctions oriented with a square aspect ratio.¹⁸ The tiles were assembled to form linear ribbons and 2D nanogrids. Arrays of the 4×4 tiles were used for the fabrication of ohmic conductors through metalisation of a ribbon-type assembly with silver.^{18,39} Moreover, protein periodic arrays were achieved by templated self-assembly of streptavidin onto the DNA nanogrids containing biotinylated oligonucleotides.^{18,19}

Typically, the assembly of complex structures requires a large number of DNA strands with distinct base sequences. The sequences must be chosen carefully to avoid unwanted interactions between the strands and this makes sequence design challenging. Recently, Mao and co-workers introduced the concept of “sequence symmetry”. Utilising the four-fold symmetry of the 4×4 tiles developed by Yan *et al.*, which is composed of nine distinct strands of DNA, they developed a related structure, which can be formed by only three types of DNA strands with symmetric sequences (Figure 1.10).⁴⁰

It is argued that sequence symmetry reduces the number of unique DNA strands needed, thus minimizing both cost and experimental errors. It also helps to minimise the required unique sequence space, and thus greatly simplifies the sequence design for the possible formation of very large structures. Though useful, structures with complex patterns are hard to achieve using the fully symmetric design.

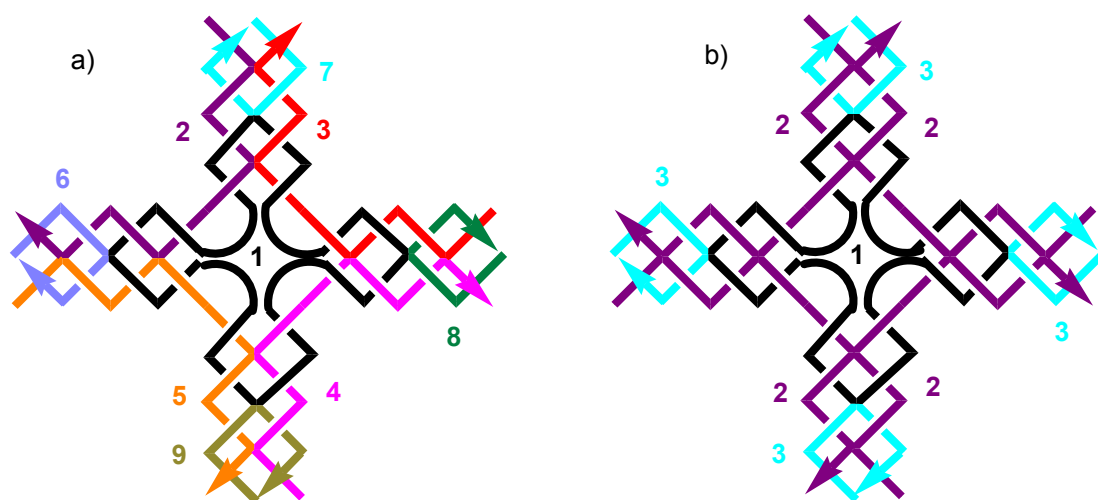


Figure 1.10: Sequence symmetry reduces the complexity of assemblies. a) An asymmetric crossbar structure composed of four Holliday junctions, which are formed by nine strands of DNA with distinct sequences. The arrows indicate the 3' ends of the DNA strands. b) A symmetric crossbar structure that is formed by nine DNA structures with symmetric sequences. The number of strands with distinct base sequences is reduced to three (adapted from ref⁴⁰).

1.2.2.2.3 DNA objects

Seeman and co-workers constructed the first object with a non-trivial connectivity of 3 or greater: a DNA molecule whose helix axes were connected like the edges of a cube;⁴¹ followed by a truncated octahedron using 4-arm junctions.⁴² Turberfield and co-workers reported the single-step synthesis of a tetrahedron, constructed out of four 55-mer ssDNA oligomers.⁴³ This object is a more-rigid wire frame than the cube because its triangular sides fulfil the tensegrity concept: a combination of tense and integrity, instead of rigidly joining rigid building blocks.⁴⁴ Tensegrity is again used by Mao and co-workers who prepared supramolecules using DNA triangles made of rods (DNA duplexes) and connected by junction motifs, which are stable thanks to the rigid rods pushing the 3D structure outward and the flexible junctions pulling it inward.³⁶ The conformational stability of the tetrahedron was recently validated through AFM imaging with an ultrasharp tip.⁴⁵ Tensegrity clearly proves to be a promising tool for the design of 3D DNA nanoarchitecture.⁴⁶

Another example of a putative DNA tetrahedron was reported by von Kiedrowski *et al.*⁴⁷ They synthesised trisoligonucleotides in which three strands of DNA are covalently connected to an organic linker molecule with a three fold symmetry axis.

Recently, using 20 distinct trisoligonucleotides as vertices, they succeeded in assembling mechanically stable DNA-based molecular dodecahedra.⁴⁸

Shih *et al.* designed a 1.7-kilobase single-stranded DNA which after addition of five 40-mer ODNs, readily folded into an octahedron structure by a simple denaturation-renaturation procedure.⁴⁹

A set of Borromean rings was built from the combination of a right-handed B-DNA 3-arm junction connected to a left-handed Z-DNA 3-arm junction. They are all disconnected by cleavage of just a single ring.⁵⁰

In a recent breakthrough, He *et al.* built molecular tetrahedra, dodecahedra, and bucky balls, from a minimal of intrinsically stiff crossover structures, adopting the rules of sequence symmetry.⁵¹

1.2.2.3 Applications

The 2D arrays and 3D structures assembled from DNA represent interesting objects in their own right, but their real usefulness will come from their application as scaffolds and templates upon which chemical reactions are performed, their use as mechanical elements or the utilisation of the information content of DNA constructs for computing.

1.2.2.3.1 Mechanical devices

DNA is not only useful for the construction of static scaffolds but it can also be used for the fabrication of dynamic assemblies for nanomechanical action. Environmental changes can induce a conformation change in a DNA nanostructure. For example, one of the first switches, created by Seeman and co-workers, was relatively small and consists of two DX molecules connected by a shaft (Figure 1.11). Double-stranded DNA with the sequence (CG)_n can be flipped from the usual right-handed helix B-DNA to a left-handed conformation Z-DNA. Conditions which promote the transition are high ionic strength or an effector that emulates high ionic strength, such as Co(NH₃)₆Cl₃. The yellow nucleotides on the shaft represent 20 nucleotide pairs that are capable of undergoing the B-Z transition. In the absence of Z-promoting conditions, these nucleotides will be in the B state, as shown at the top of the panel, but in the presence of Z-promoting conditions they convert to Z-DNA. The transition was demonstrated by

measuring FRET differences between the two states.⁵² Alteration of pH value has also been used to induce the conformational change of a DNA triplex, used as nanomechanical actuator.^{53,54}

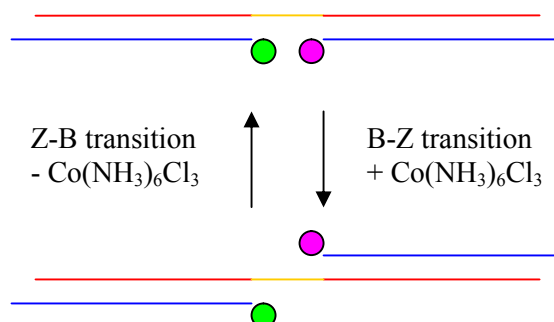


Figure 1.11: A DNA nanomechanical device based on the B-Z transition.

Another method, called “Strand Replacement” can also be made reversible by the simple addition of a well-designed DNA ODN. The strand replacement mechanism is demonstrated with the DNA tweezers from Yurke et al (Figure 1.12).⁵⁵ This principle has been used to effect conformational changes in a wide variety of systems and a number of variations on the tweezers have been reported.^{56,57}

Simmel and co-workers also employed this method for the construction of a protein-bearing molecular machine based on a DNA aptamer. They opened and closed a G-quadruplex-shaped thrombin-binding aptamer, allowing the release or binding of a thrombin molecule in the presence of a specific DNA sequence.⁵⁸ There also have been a variety of elegant designs of walking devices produced.^{8,59,60}

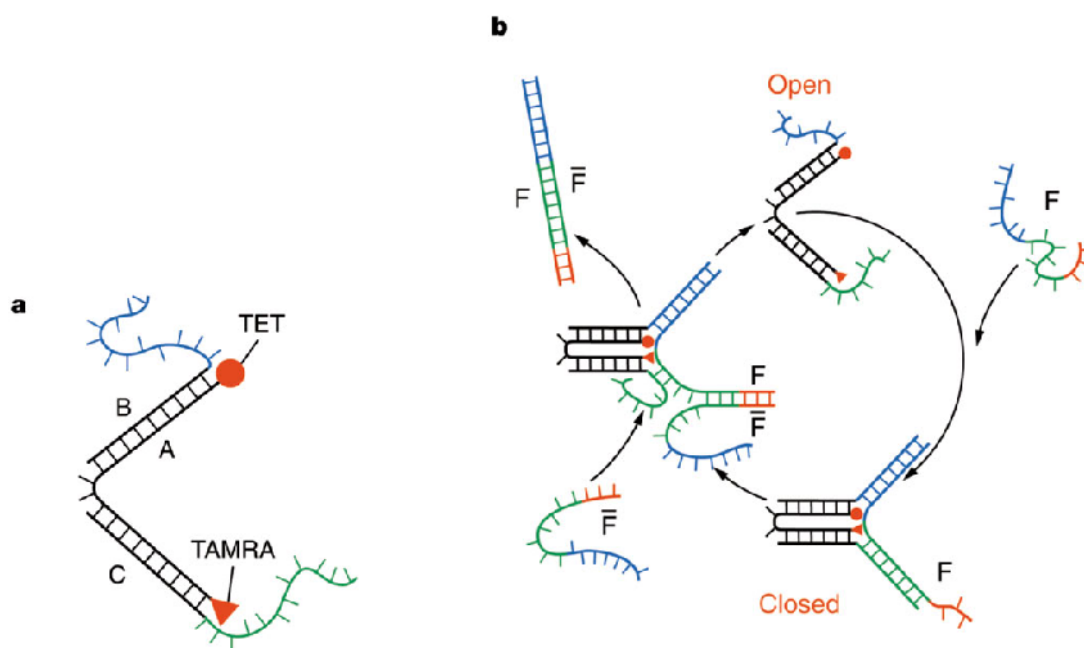


Figure 1.12: Construction and operation of the molecular tweezers. **a**, Molecular tweezer structure formed by hybridisation of ODN A, B and C. **b**, Closing and opening the molecular tweezers. Closing strand F hybridises with the dangling ends of strands B and C (blue and green) to pull the tweezers closed. Hybridisation with the overhang section of F (red) allows F strand to remove F from the tweezers, forming a double-stranded waste product FF and allowing the tweezers to open. Adapted with permission from ref. 48.

1.2.2.3.2 DNA assembly of biomolecules

Assembly of other biomolecules on DNA templates and arrays may prove useful for fabrication of biomimetics and other devices with applications such as biochips, immunoassays, biosensors, and a variety of nanopatterned materials.

The conjugation of DNA and streptavidin *via* a covalent linker was reported by Niemeyer *et al.*, and these conjugates were applied to DNA-programmed assembly on a macroscopic DNA array on a surface and in a nanoscale array made by aligning DNA-tagged proteins to specific positions along a oligonucleotide template.^{61,62}

Biotin labelled oligonucleotides are commercially available, routinely used in biotechnology, and their application for nanostructuring is growing in popularity.^{63,64} As described previously (see section 1.2.2.2.2), Yan *et al.* used the 4×4 tiles to template the self-assembly of streptavidin proteins into periodic 2D arrays through a biotin group attached in the centre of the tile.¹⁸ The streptavidin protein array can be programmed with controlled spatial distance and density of the proteins, as seen on AFM pictures.¹⁹

A number of applications of materials assembled by structural DNA nanotechnology are already starting to emerge in biotechnology. Tuberfield and co-workers showed the binding of the protein RuvA to the Holliday junction of a 2D lattice. This resulted in a 2D crystal array of this protein, which allowed for its structural elucidation by using electron microscopy.⁶⁵ By assembling different DNA tile arrays, each with a specific recognition molecule and "bar-coded" with a specific fluorescent dye, Yan developed a platform that allows the simultaneous detection of multiple biological analytes. This method may be faster than DNA or protein microarrays for small-scale profiling of bioanalytes.⁶⁶

1.2.2.3.3 DNA scaffolds for nanoelectronics

The conjugation of DNA with metal, semiconductor and magnetic nanoparticles is becoming increasingly important since DNA self-assembly provides a framework for the spacing and hierarchical organisation of nanomaterials. The periodicities and interparticle spacings defined by DNA nanostructures are readily adjustable, with sub-nanometer spatial resolutions. This level of precision provides exquisite control in the construction of rationally defined 2D and 3D nanoscale assemblies.

Pioneering reports on the assembly of gold nanoparticles by hybridisation of DNA-nanoparticle conjugates were published in 1996 by two groups. Mirkin *et al.* developed a method for assembling colloidal gold nanoparticles rationally and reversibly into macroscopic aggregates.⁶⁷ When they attached non-complementary ODNs on the surface of two batches of gold nanoparticles, using thiol groups, and added a DNA target complementary to both DNA-nanoparticles sequences, hybridisation forced the particles to aggregate. The resulting colour change in the presence of the target is a very efficient and easy method for DNA-detection. At the same time, Alivisatos and co-workers applied gold nanoparticles labelled with only single ODN molecules for the synthesis of homodimeric or homotrimeric nanoparticle assemblies.⁶⁸

Another strategy of using DNA as a scaffold for gold nanoparticles relies on the biotin-streptavidin binding. Yan *et al.* used TX molecules functionalised with biotinylated DNA strands, which were incubated with streptavidin-coated gold nanoparticles to assemble them in periodic arrays.⁶⁹

Sleiman *et al.* developed DNA building blocks, containing two arms of different sequences, and a rigid organic corner unit, which selectively assembled into a discrete DNA hexagon. This approach was used to organise six gold nanoparticles into a hexagon.⁷⁰

1.2.2.3.4 DNA-based computation and algorithmic assembly

In 1994 Adleman demonstrated that DNA can be used for computationally solving a small example of a Hamiltonian path problem in computer science.¹¹ The idea behind DNA-based computation is that for some classes of computational problems the parallelism of molecular assembly overcomes the slow speed of the required macroscopic manipulations. An example of DNA computing based on regular assemblies of DNA motifs as building blocks is the work done by Winfree *et al.* on the “Wang tiles”. He described the similarity of DNA branched junctions with sticky ends to Wang tiles. Two such tiles only assemble if the joint edges have the same colour specific sticky ends (Figure 1.13). The assembly of Wang tiles can be used to simulate a cellular automaton. This is a theoretical model of a computer which can be used to execute, in principle, any calculations done by a conventional computer. As an example, they designed a set of DX tiles capable of solving a small example of the Hamilton-path problem.⁷¹

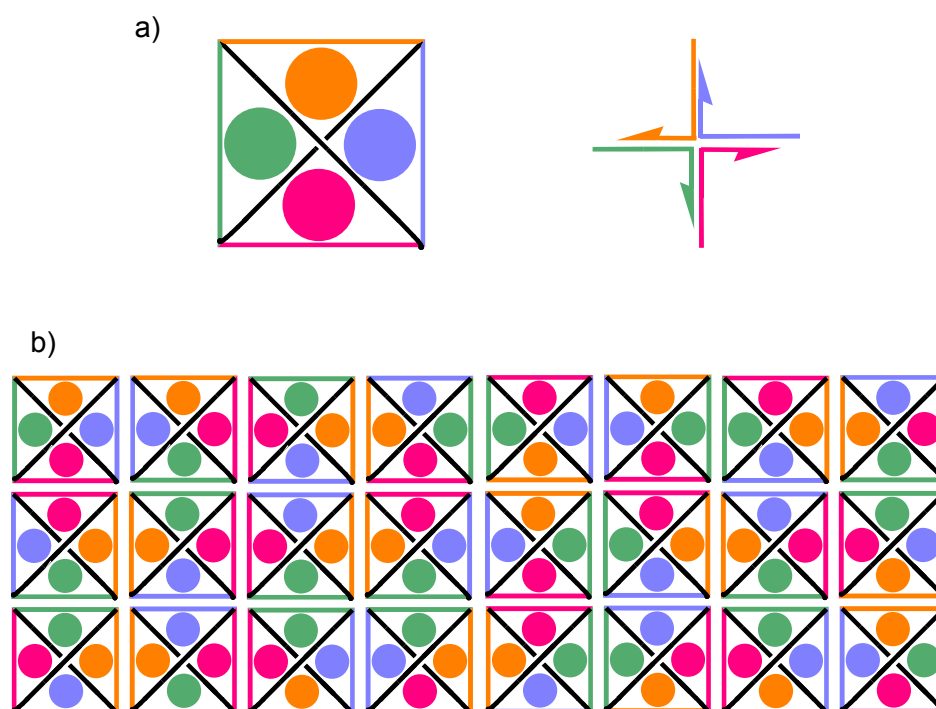


Figure 1.13: Computational assembly of the DNA tile motifs. a) Relationship between Wang tiles and branched junction, indicating that the sticky ends on a branched junction can emulate a Wang tile. b) Wang tiles with colour-coded edges can be assembled according to the edge codes, thereby executing computations.

Algorithmic DNA self assembly has been used to realise simple logic functions such as XOR. The XOR calculation yields a 1 if the two inputs are different and a 0 if they are the same. The concept of algorithmic assembly has been extended in recent years and most importantly by Rothemund *et al.*, who assembled DX tiles into Sierpinski triangles.¹² These are symmetric triangle structures which represent the iteration of XOR calculations. A polymeric nucleating scaffold strand was used at the starting point of the assembly, to suppress spurious nucleation and carry the input information for the computation.

Based on the concept of nucleated DNA self-assembly and previous DNA junction structures, a recent breakthrough in structural DNA nanotechnology toward creating complex shapes and patterns came from Rothemund's scaffolds, DNA origami.²² This method is based on folding a simple, long single-stranded DNA into a desired shape with the help of multiple "staple" strands. In contrast to the traditional assembly approach, there is no need to observe strict stoichiometric ratios, and assembly occurs

quite quickly and in good yields. A variety of 2D shapes are created with the same long strand and different sets of the “staple” strands: squares, rectangles, five-pointed stars, a smiley face, and different triangles, all roughly 100 nm in dimensions.

1.3 AMNA project

1.3.1 General

The objective of the Addressable Molecular Node Assembly (AMNA) project is to develop a nanotechnology platform based on a sub-micron-size 2-D grid of addressable molecular building blocks (“nodes”). DNA base-pair recognition will then be used to construct the network, with each node having typically three oligonucleotide strands. Oligonucleotide synthesis starts with a single oligomer strand, then incorporation of a simple branching monomer followed by an orthogonal deprotection strategy, allows the synthesis of two independent oligonucleotide strands from this point, forming a node as shown in Figure 1.14. The two different oligonucleotide sequences synthesised from the branch point, can be constructed 3’ to 5’ or 5’ to 3’ (using reverse amidites) as desired. This provides a great deal of flexibility in the ways that the nodes can be linked together.

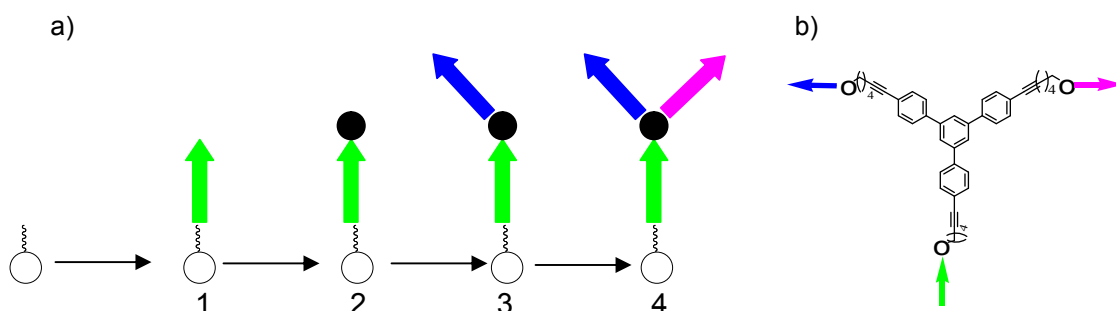


Figure 1.14: a) Oligonucleotide synthesis: The first strand is synthesised as in 1, then the branching monomer is incorporated as in 2 and selective deprotection allows the synthesis of the second arm (blue) as in 3 then the third arm (pink) as in 4. b) Three way symmetrical node used.

A group of six nodes will connect to form a hexagon (Figure 1.15) providing a planar network of hexagons. The aim will be to incorporate interesting functional groups which could be used to attach various moieties on the grid map, including chromophoric “radiation sources”, extending perpendicular to the grid surface.

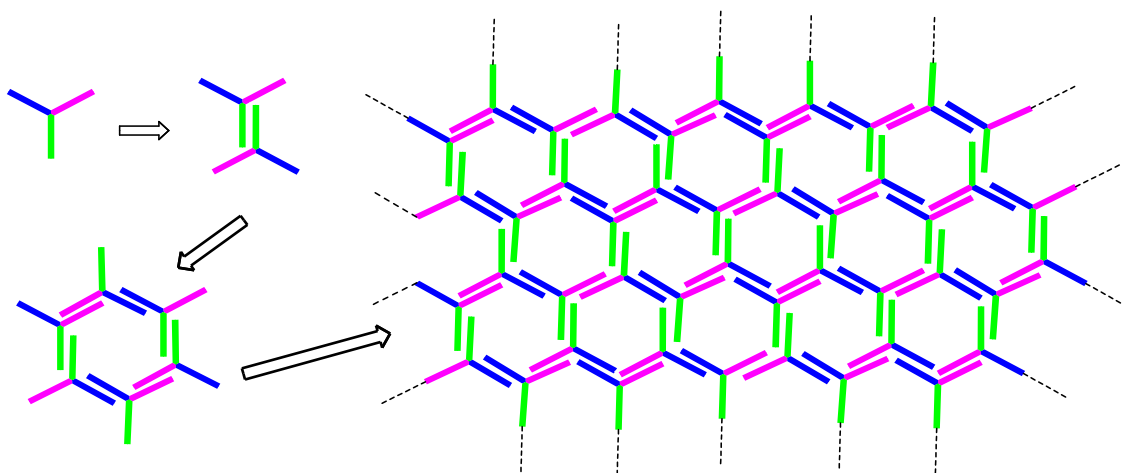


Figure 1.15: Formation of the hexagons by using branching points and Watson-Crick base-pair recognition

Incorporation of fluorescent labels and lipophilic groups (to anchor the network to a lipid surface) is to be achieved using modified phosphoramidite building blocks.

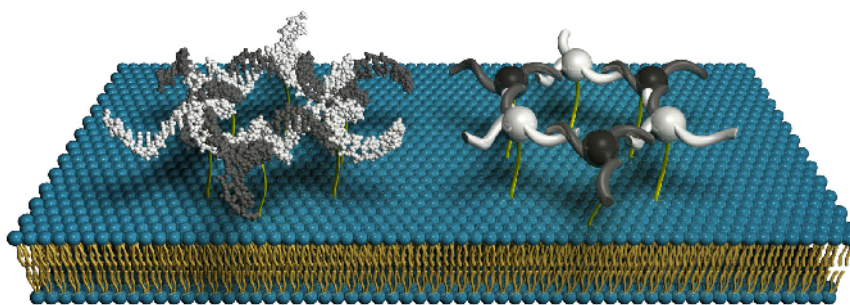


Figure 1.16: To the left - molecular model, and to the right - schematic picture of start hexagon consisting of six nodes, each bearing three deoxyribonucleotide (DNA) arms with mutually complementary sequences of ten bases. Lipid chains, covalently tethered to the DNA strands, serve the role of anchoring the hexagon to a lipid matrix. The grid will be built up, radially from the start hexagon, by a net of hexagons and nodes, each being uniquely addressable by its base-sequence code. ©AMNA project.

1.3.2 Lipophilic oligonucleotides

Chemical modifications are introduced using the appropriate phosphoramidite monomers and incorporating them during oligonucleotide assembly. Attachment of lipid moieties is carried out using cholesterol with a hydrophilic group as a spacer. In this way the hydrophilic group will form a bridge between the lipophilic group and the phosphate backbone of the DNA chain and will enable the oligonucleotide to interact with lipid membranes and at the same time remain in a hydrophilic environment at a fixed distance from the lipid surface.

Lipid-conjugated oligonucleotides (L-ODN) will be studied with respect to assembly in an effectively planar lipid bilayer. Highly monodisperse liposomes made of different natural and synthetic lipids will be prepared through extrusion techniques and L-ODN incorporation will be investigated as a function of the lipid composition, insertion of double bonds, charge, solution ionic strength and in the presence of lyotropic agents.

Hexagonal phases are commonly encountered as lipid mesophases and should be affected by the presence of three-armed oligonucleotides tethered to lipids and by the length of the spacer. In this case the hexagonal network is not only confined to a 2D space, but may also propagate into 3D. At this stage a proper tailoring of the spacing

agent should be crucial in imparting appropriate protrusion length to the three-branched oligonucleotidic groups.

Parameters of the binding affinity (lipid chain length, membrane charge and ionic strength) between the membrane and nodes will have to be fine-tuned. A very strong membrane affinity of hexagons and an optimal, medium binding affinity of the monomer nodes are desirable. In this case the nodes will ideally be present at a high concentration on the membrane surface, but still be able to dissociate and associate quickly, until they find and hybridise to their membrane-bound target sequences of the growing network.

1.4 Click chemistry

1.4.1 Introduction

Following nature's lead, a new concept in organic chemistry has been defined which has the potential to accelerate the synthesis of substances with useful purposes. The way this is being done is by using nature's small molecule building blocks that have C-C bonds to create heteroatom links, which are much easier to make efficiently than carbon-carbon bonds. This approach has been termed by Sharpless and co-workers "click chemistry".

To be called "click", the reactions must be modular, have a high thermodynamic driving force (give very high yields) and the starting material and reagent should be readily available. They should proceed under simple conditions using solvents that are benign (such as water), generating only inoffensive by-products and necessitate simple work-up and purification procedures. This makes click chemistry both clean and green.

These click reactions achieve their required characteristics because they start with "high energy" compounds, proceed to virtually complete conversion of these reagents and tend to be highly selective. Such reactions are often termed "spring-loaded". The ideal click chemistry reactions are based on stereospecific processes with virtually all of the control elements and the enabling "energy packets" present in the reactive components.⁷²

Olefins are among the most attractive starting materials available to the synthetic chemist as their chemistry provides for the creation of diverse scaffolds, and the

attachment and display of various functionalities through the oxidative addition of heteroatoms to specifically placed olefinic sites (Figure 1.17). For example, their importance is enhanced by their role as progenitors of still higher-energy intermediates such as epoxides, aziridines, episulfonium ions and aziridinium ions, which are all suitable for a click chemistry approach.⁷³

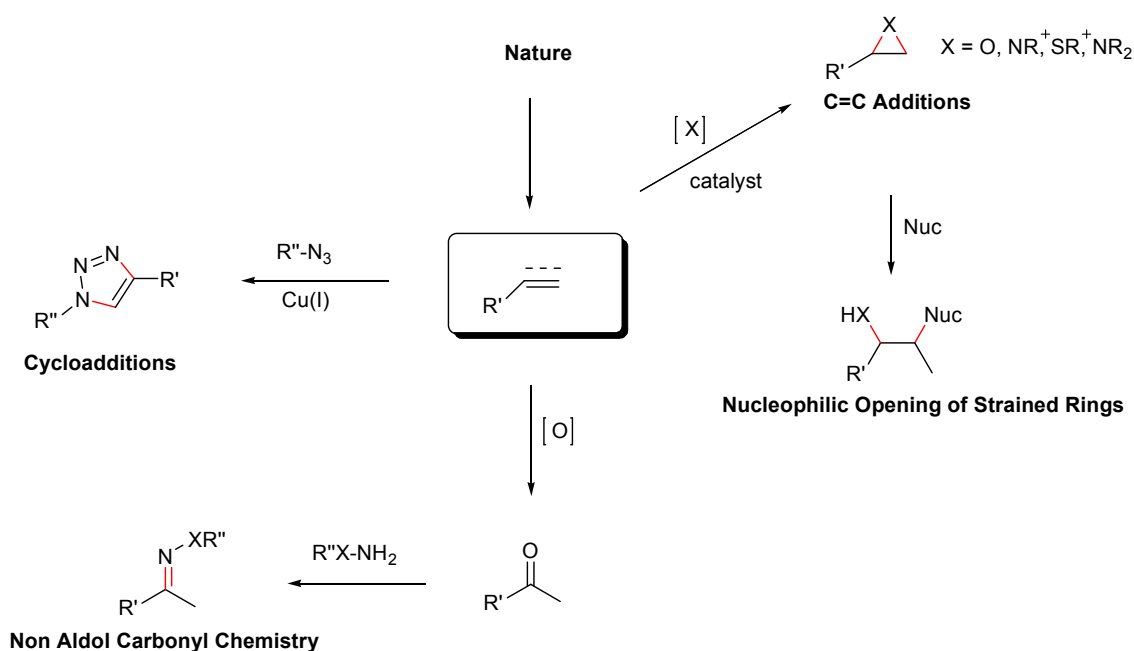


Figure 1.17: A selection of reactions which meet the Click Chemistry criteria. Unsaturated compounds provide the carbon framework. New groups are attached via carbon-heteroatom bonds (shown in red).

Sharpless and co-workers have identified a set of the most useful click reactions where the best are pure fusion processes. This characteristic is based on the understanding of the limitation of raw material, combined with environmental concerns, which necessitate rethinking the strategies towards complex organic synthesis and recognising the advantages of atom-economical reactions.⁷⁴

- Nucleophilic opening of Spring-Loaded Rings:

These are the $\text{S}_{\text{N}}2$ ring-opening of epoxides, aziridines, cyclic sulfates, cyclic sulfamidates, aziridinium ions and episulfoniums ions, which are strained heterocyclic electrophiles from olefins or their oxidation products (Figure 1.18).

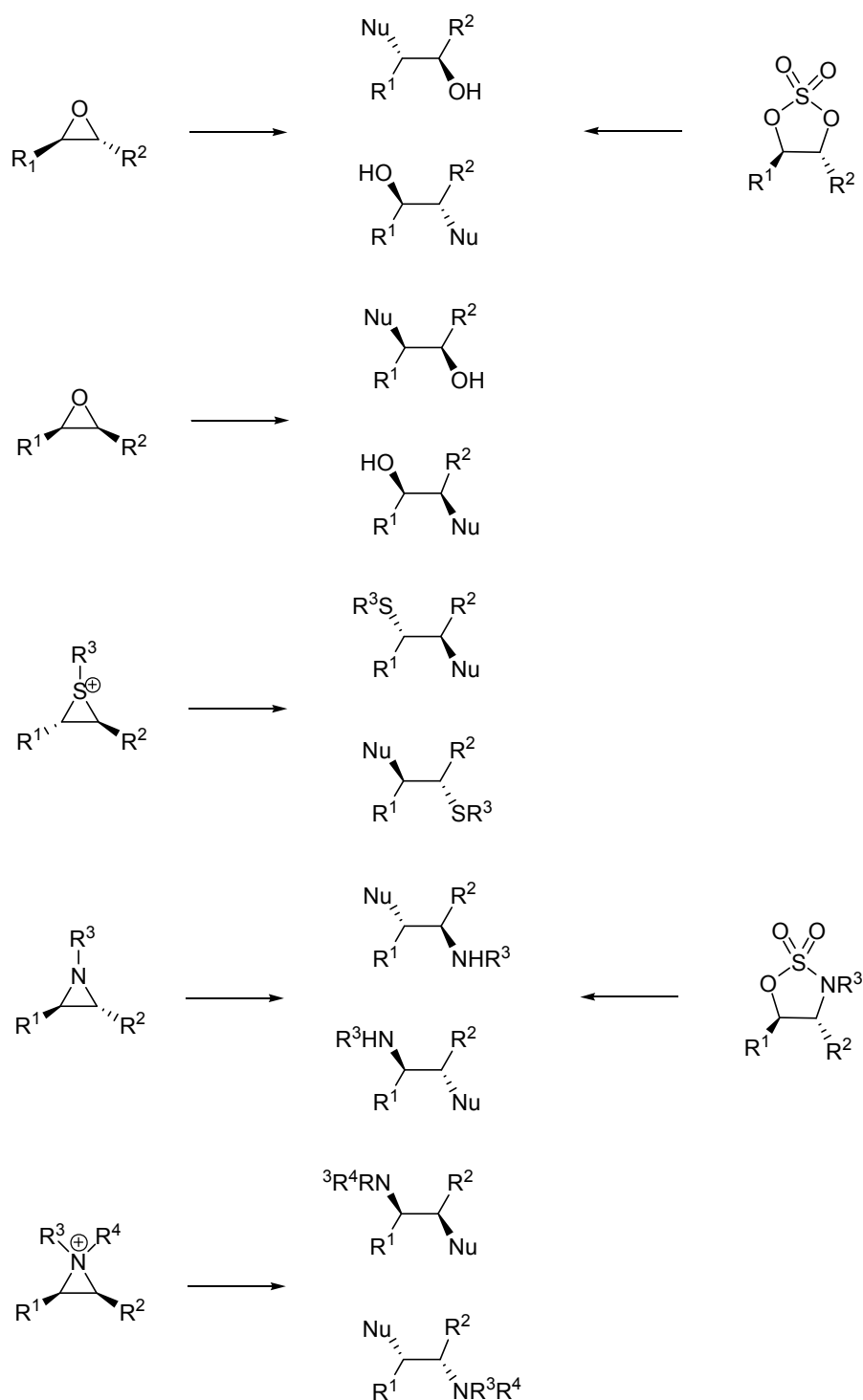


Figure 1.18: Opening of spring-loaded cyclic electrophiles

- **Cycloaddition Reactions:**

The two typical examples are hetero-Diels-Alder and 1,3-dipolar cycloadditions. The latter is the most useful and reliable of the click reactions because the azide group is by

far the most convenient of the 1,3-dipolar components to introduce and to “carry” until needed.

- Non aldol carbonyl chemistry:

Acetals, ketals and some of their aza-analogues, usually used as diol protecting groups, can be viewed as an attractive class of heterocycles for medicinal chemistry applications. They represent one of the rare click chemistry module based on reversible carbonyl chemistry.

- Addition to multiple carbon-carbon bonds:

These reactions are typically oxidation, such as epoxidation, dihydroxylation, aziridination, and nitrosyl and sulfenyl halide additions, but also certain Michael additions.

1.4.2 Huisgen reaction

The Cu^I-catalysed 1,3-dipolar cycloaddition has become, within the past few years, a premier component of the click chemistry paradigm.

The prime example of a click reaction is the copper catalysed variant of the ubiquitous Huisgen 1,3 dipolar cycloaddition of alkynes and azides.^{75,76} The azides and alkynes are easily installed into a molecule and despite being some of the more energetic species known, they are also very stable under a variety of conditions. Indeed, they are essentially inert to most biological and organic conditions and the majority of common conditions in organic synthesis.^{77,78} This stability, being purely kinetic in origin, is responsible for the slow nature of the cycloaddition reaction, which requires elevated temperatures and long reaction times⁷⁹⁻⁸³ Unless the acetylene component is attached to a highly electron-deficient group, the cycloaddition reaction affords mixtures of the 1,4- and 1,5-disubstituted-regioisomers (Figure 1.19).

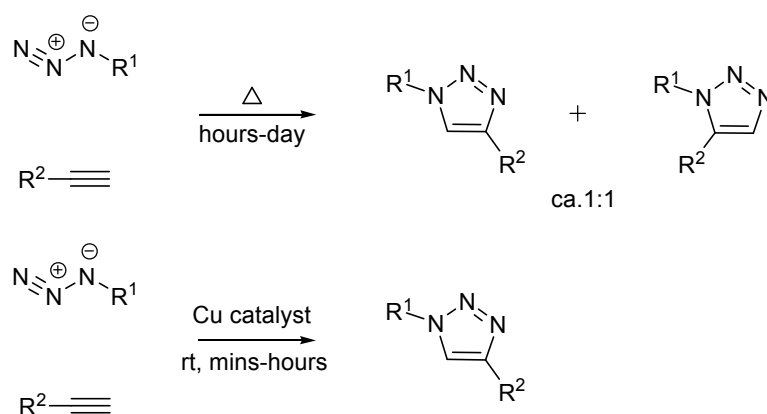


Figure 1.19: The thermally induced Huisgen cycloaddition of azides to alkynes usually results in mixtures of the 1,4 and 1,5-triazole stereoisomers. In contrast, Cu (I) catalysed reaction results in exclusive formation of the 1,4-triazole.

The reaction has benefited from a dramatic rate enhancement (up to 10^7) by copper catalysis (Figure 1.19), reported independently by the Sharpless⁷⁸ and Meldal groups.⁸⁴ This new reaction process is experimentally simple, proceeds with almost complete conversion and selectivity for the 1,4-disubstituted-1,2,3-triazole, and appears to have enormous scope for many reactions.

The reaction performs best in aqueous systems, is tolerant to a broad range of temperatures, is remarkably insensitive to pH, and is not significantly affected by the steric and electronic properties of the groups attached to the azide and alkyne reactive centres.^{85,86}

Different sources of copper (I) can be utilised in the reaction. Sources of Cu^{I} are copper salts (CuI , CuBr), coordination complexes (such as $[\text{Cu}(\text{CH}_3\text{CN})_4]\text{PF}_6$, $\text{CuOTf}\cdot\text{C}_6\text{H}_6$ ⁷⁸, $[\text{Cu}(\text{PPh}_3)_3]\text{Br}$ ⁸⁷) and with a mixture of Cu^0 and Cu^{II} ⁸⁵. As Cu^{I} is thermodynamically unstable, the catalyst is better prepared *in situ* by reduction of Cu^{II} salts by ascorbic acid and/or sodium ascorbate. However, at very low concentration of the reagents, the reaction is not fast enough, and copper- or ascorbate-mediated degradation of the biological scaffolds has been observed. The use of the ligand tris-(benzyltriazolylmethyl)amine (TBTA) significantly accelerates the reaction and increases the redox potential of $\text{Cu}^{\text{I}}/\text{Cu}^{\text{II}}$ (Figure 1.20).⁸⁸

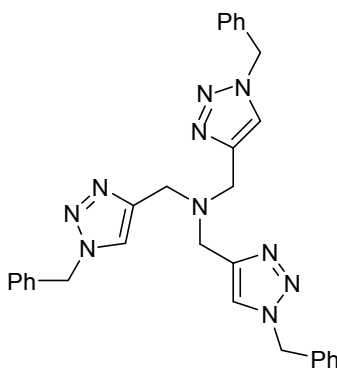


Figure 1.20: *Tris*-(benzyltriazolylmethyl)amine, TBTA.

1.4.3 Mechanism

The process of this catalytic reaction has been postulated to occur by a stepwise mechanism as opposed to the thermal dipolar cycloaddition, which occurs through a concerted mechanism. The lowest calculated barrier of any concerted 1,3-dipolar cycloaddition of a copper-acetylene π complex is 23.7 kcal/mol, too high to be responsible for significant rate effect of Cu^{I} catalysis. The proposed catalytic cycle lowers the activation barrier relative to that of the non-catalysed by as much as 11 kcal/mol (Figure 1.21).

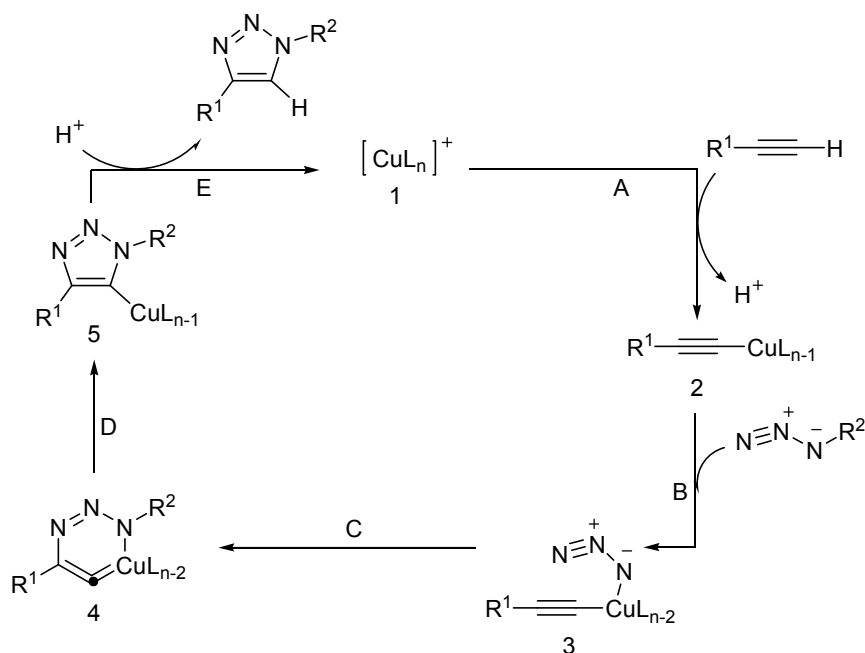


Figure 1.21: Proposed catalytic cycle for 1,3-dipolar cycloaddition

The sequence begins with the complexation of the alkyne (step A) to the Cu^I metal centre to form a copper acetylide **2** *via* the π complex. Alkyne π complexation was calculated to be endothermic by 0.6 kcal/mol in acetonitrile and exothermic by 11.7 kcal/mol in water. This result is in good agreement with experimental observation of a rate acceleration of a ligand dissociation in water. Moreover, the copper coordination of acetylene is calculated to lower the pK_a of the alkyne terminal proton by up to 9.8 pH units, making deprotonation in aqueous system possible without the use of strong base.

Formation of the Cu^I-acetylide species allows for subsequent azide displacement of one ligand and results in the intermediate **3** (step B). The key bond is created in the next step (step C) by nucleophilic attack of the nitrogen distal to the acetylide, generating the unusual six membered metallacycle **4**. This step is endothermic by 12.6 kcal/mol with a calculated barrier of 18.7 kcal/mol, which is considerably lower than the barrier for the non-catalytic reaction (26.0 kcal/mol), and explains the tremendous rate acceleration of the copper catalysed process. The regiospecificity can be explained by the fact that both azide and alkyne are linked to the copper prior to the formation of the C-C bonds.

The resulting metallacycle **4** undergoes facile ring contraction (step D). Indeed, the energy barrier which leads to the triazolyl-copper derivative **5** is especially low (3.2 kcal/mol). Protonation (step E) of the triazole-copper derivative affords the 1,4-disubstituted 1,2,3-triazole, thus regenerating the catalyst and ending the cycle.^{85,86,89}

1.4.4 Selected applications

Within a short time-frame, click chemistry, and essentially the Copper(I)-catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC), has proven to be of remarkable utility and broad scope not only in organic synthesis, but also in bioscience (drug discovery, bioconjugation) and material science.⁹⁰⁻⁹²

1.4.4.1 Use in organic chemistry

CuAAC has been used for the construction of a library of asymmetric organocatalysts for the highly diastereo- and enantioselective Michael addition^{93,94}. Pericàs *et al.*

anchored a hydroxyproline derivative to a Merrifield resin using click chemistry and used the resulting catalyst to direct the aldol reaction in water.⁹⁵

1.4.4.2 Drug discovery

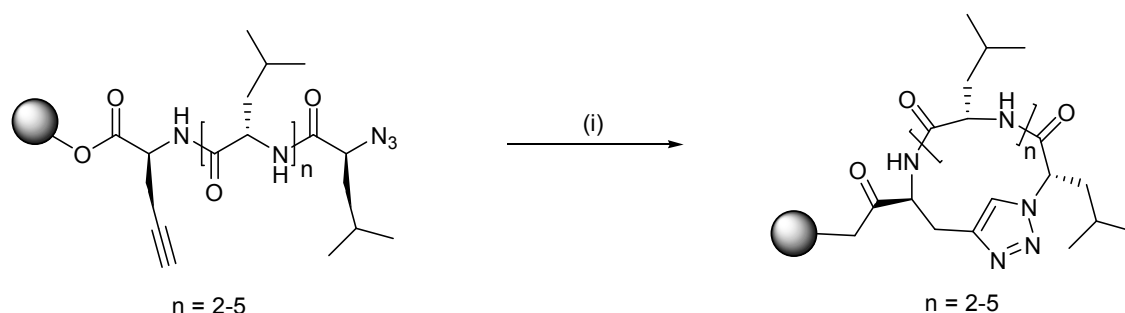
Click chemistry is playing an important role in biomedical research, facilitating the development of lead discovery libraries, assisting optimisation and also tagging biological systems. The lack of by-products and high conversion often allows screening of reaction products without further purification⁹⁰

Using established, extensive internal medicinal chemistry capabilities that utilise click chemistry, laboratories at Lexicon Pharmaceuticals, have generated a variety of sophisticated libraries of small molecule chemical compounds with drug-like characteristics. Each library compound was produced in only one or two synthetic steps, from key building blocks reagents, using solution-phase chemistry.

Examples include nucleophilic opening of strained epoxides and aziridines for the formation of 1,2-difunctionalised compounds, generation of five-membered aromatic heterocycles with imidoesters⁹⁶, synthesis of 1,2,3-triazole-derived libraries with azides via 1,3-dipolar cycloaddition with β -ketoesters⁹⁷, and preparation of non-aromatic heterocycle libraries with 3-aminoazetidines. A targeted library led to the discovery of Peroxisome Proliferator-Activated Receptor γ (PPAR- γ) agonists.⁹⁸

The reliability of click chemistry has had a significant impact on high throughput methodology. For instance, Wong et al identified a nanomolar inhibitor of α -1,3-Fucosyltransferase. The library of 85 donor substrates were synthesised *via* CuAAC and screened *in situ* within the same well in which they were synthesised.⁹⁹

Lokey *et al.* have demonstrated the utility of click chemistry in the on-resin cyclisation of a series of leucine-rich tetra-, penta-, hexa-, and heptapeptides, which makes possible the generation of vast libraries of cyclic peptides. The cyclisation is complete within 6 hours, utilising mild conditions that are orthogonal to the protection groups common to solid phase peptide synthesis (Scheme 1.1).¹⁰⁰



Scheme 1.1: Reagents and conditions (i) CuBr (1eq), Na ascorbate (3eq), DIPEA (10 eq), 2,6-lutidine (10eq), DMF, 6 h, 25 °C.

Two highly versatile cycloadditions (aza-Diels-Alder and CuAAC) have recently been used for the rapid assembly of proline mimetic libraries.¹⁰¹

The 1,2,3-triazole moiety is a potential pharmacophore that can mimic the atom placement and electronic properties of a peptide bond without the same susceptibility to hydrolysis.^{72,102} Therefore it can be readily incorporated into a design strategy, rather than being used merely as a linking functionality. Several examples exist within the literature where the biological activity of 1,2,3-triazoles have been described, including HIV-1 protease inhibitor,^{103,104} antibacterial,^{105,106} anti-inflammatory,¹⁰⁷ tuberculosis inhibition^{108,109} to name a few.

1.4.4.3 Bioconjugation

The highly chemoselectivity of click chemistry has recently emerged as a powerful conjugation strategy for the preparation of diverse bioconjugates including conjugation with peptides, proteins, oligosaccharides, glycoconjugates and DNA.

The original demonstration of versatility of CuAAC towards bioconjugation was hinted by Meldal and co-workers in their breakthrough article as the joint discoverers of the CuAAC.⁸⁴ They used the triazole as a peptide bond mimic and showed that the reaction conditions of CuAAC were fully compatible with solid-phase peptide synthesis. Triazole ligations allow a broad possibility of oligopeptide designs such as cyclic peptides using triazole-based spacers or triazoles as peptide isosteres.^{110,111} Maarseveen *et al.* proved that 1,4-disubstituted 1,2,3-triazoles can serve as transoid amide bond

mimics and demonstrated the retention of biological activity of the cyclic peptidomimetic.¹¹¹

Following the discovery and development of CuAAC in water, the potential for attachment of synthetic labels to a biomolecular environment has been realised. *Finn et al.* applied CuAAC to conjugation of fluorescein dye molecules to the cowpea mosaic virus.¹¹² The virus itself is a structurally rigid assembly of 60 identical copies of a two-protein asymmetric unit around a single-stranded RNA genome. The virus particle presents functionality on its exterior surface, including reactive lysine or cysteine residues, which were decorated with azide and alkynes, which were therefore used for conjugation *via* CuAAC.

The procedure was developed in a further study by *Finn et al.*, where they used a new water-soluble sulfonated bathophenanthroline ligand (Figure 1.22). The ligand displayed higher reaction rates than TBTA with lower concentration of labelling substrate. However, as the ligand increases the sensitivity of the copper catalyst to oxidation, the reaction has to be carried out in an inert atmosphere.¹¹³

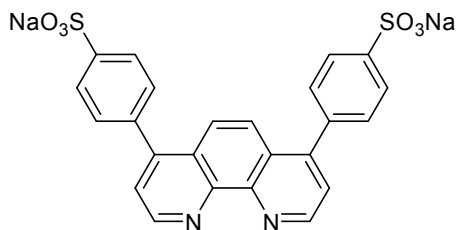


Figure 1.22: Water sulfonated bathophenanthroline ligand

In carbohydrate chemistry, several examples exist where one or more saccharides have been linked to themselves, to a central core (drug, peptide, etc), or functional groups by CuAAC. Indeed, the synthesis of triazole-linked glycosaccharides has solved the problem of sensitivity of glycosidic linkages towards chemical and enzymatic hydrolysis which make them poor lead-compounds.^{114 115} Such ligation is versatile enough to allow for example the formation of hybrid structures such as defined glycopeptides.^{116,117} For instance, Danishefsky and co-workers used multiple click triazole formation to combine three tumor-associated carbohydrate antigens on a single polypeptide scaffold (Figure 1.23).¹¹⁸

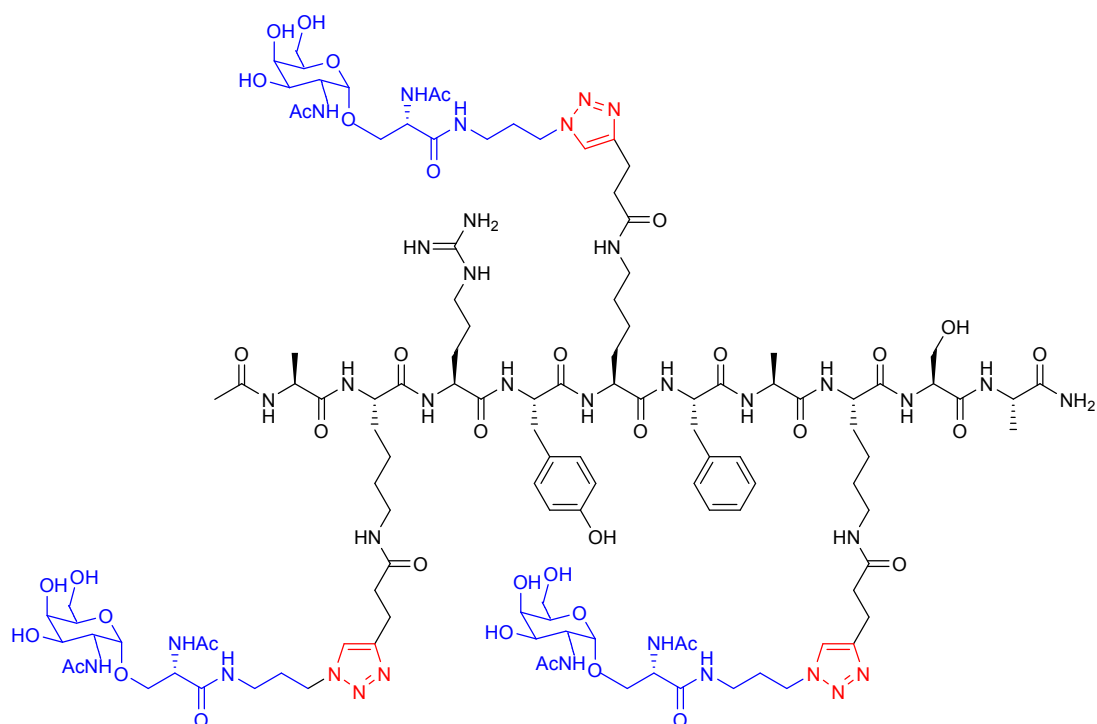


Figure 1.23: Synthetic glycopeptide fragments conjugated to a modified carrier protein *via* click chemistry.

The power of bioconjugation with click chemistry extends to the labelling of biomolecules *in vivo*. The bioorthogonal azide and alkyne functionalities have been exploited by *Cravatt et al.* for use in activity-based profiling of proteins (ABPP). ABPP is a rather new postgenomic method in which affinity labels are used to profile proteins on the basis of their function in biological systems. Commonly used reporter groups in ABPP (fluorophores, biotin, etc) are quite bulky and may inhibit cellular uptake. A rhodamine alkyne was reacted with sterically inconspicuous azido-modified enzymes using CuAAC. Mice were used to prove that the click chemistry-ABPP method could identify changes in enzyme activity *in vivo*. However, the click reaction was still only performed after sacrifice of the animal and cell lysis, due to the toxic nature of the copper catalysis (Figure 1.24).^{119,120}

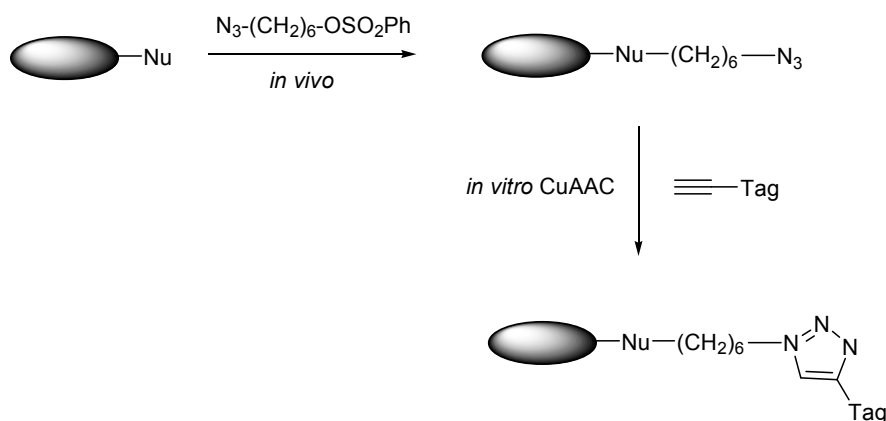


Figure 1.24: Click activity-based protein profiling

Although the reactivity of the participating groups are almost perfectly orthogonal to those of the diverse functional groups in biological systems, this reaction still does not provide a final solution. In fact, the ideal bioconjugation reaction can be carried out without the requirement of Cu^{I} catalysis (which exhibits considerable cell toxicity) and ancillary ligands. Bertozzi *et al.* investigated novel substituted “spring-loaded” cyclooctynes as an alternative means of activating alkynes for catalyst free [3 + 2] cycloaddition with azides. They performed *in vivo* labelling and the validity of this copper-free approach in living cells was demonstrated.^{121,122} Cycloaddition reactions involving cycloalkynes are not regioselective but present some “click” features such as chemoselectivity and ready applicability in physiological conditions.

1.4.4.4 Click chemistry and DNA

Click chemistry has been used in the field of DNA sequencing to improve the efficiency of the Sanger method in mild conditions without the need of additives or the formation of by-products.¹²³ The M13-40 universal forward sequencing primer was tagged at its 5'-end with alkynyl 6-carboxyfluorescein (FAM) without further purification by gel electrophoresis or HPLC required for conventional fluorescent oligonucleotides synthesis. The FAM-triazole-M13-40 primer was successfully used in the Sanger method.¹²⁴ However, in this early work, a standard metal-free Huisgen cycloaddition was performed. Carell *et al.* used CuAAC to post-synthetically decorate alkyne modified DNA with a high density of azido-labels. They demonstrated the utility of the

TBTA ligand which stabilises the Cu^I oxidation state since copper ions in the presence of O₂ induce DNA cleavage.^{125,126}

CuAAC has also proven to be useful for the immobilisation of acetylene-bearing ssDNA to gold-coated azide-functionalised self-assembled monolayers in the presence of TBTA. The ligand here also prevented the production of potentially DNA damaging oxygen radicals *via* the redox chemistry of copper (I).¹²⁷

CuAAC has been used in the nucleic acid field to attach various labels to modified DNA such as polypeptides¹²⁸, fluorescent labels¹²⁹, or after methyltransferase-mediated site-specific alkylation¹³⁰. It has also been used with microwave assistance, for the preparation of DNA-glycoconjugates using alkynes attached to phosphoramidate linkages present in synthetic oligonucleotides.¹³¹ Microwave irradiation has also been used to accelerate CuAAC reactions between pyren-1-yl azides and ODNs incorporating several insertions of azides to form twisted intercalating nucleic acids.¹³²

Kocalka *et al.* used CuAAC to covalently cross-link complementary DNA strands internally utilising several alkyne and azide-modified uracil monomers to form very stable duplexes.¹³³

The synthesis of large covalently closed DNA catenanes was achieved by Kumar *et al.* who developed an efficient method of oligonucleotide ligation using CuAAC (Figure 1.25). First, two different strands, one containing a 3'-azide and the other a 5'-alkyne, were connected by a templated triazole cyclisation. This approach was applied for the intramolecular circularisation of a single stranded ODN to form macrocyclic DNA without the use of template. When the complementary DNA strand, also possessing 3'-azide and 5'-alkyne termini, was allowed to anneal to the cyclic DNA and treated with Cu(I), a double-stranded DNA catenane resulted, connected with two stable triazole units.¹³⁴ This strategy has also been used to produce extremely stable and very short duplexes, containing hexaethylene glycol bridges, by forming triazole linkages between the sugar-phosphate backbone at the termini.¹³⁵

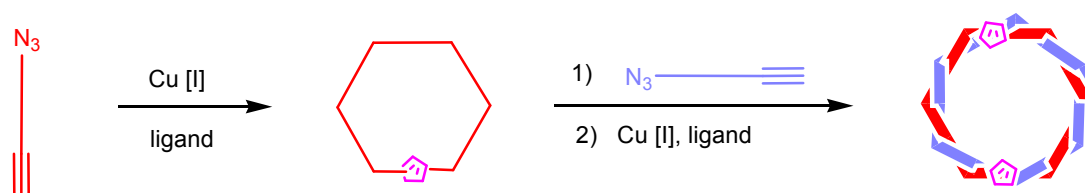


Figure 1.25: Synthesis of a double-stranded DNA catenane from a single-stranded cyclic template (shown in red), followed by annealing of the complementary strand (shown in blue) and ligation (triazoles are represented by pink pentagons).

1.5 Gel electrophoresis¹³⁶

Electrophoresis is generally used for separation and sometimes purification of macromolecules that differ in size, charge or conformation. This widely-used technique can result in very high resolution. The charged macromolecules are placed in an electric field and travel through a matrix of inert material, which acts as a molecular sieve, allowing separation of macromolecules.

1.5.1 PAGE

Nucleic acids are electrophoresed within a gel matrix composed of either agarose or polyacrylamide. Polyacrylamide is a cross-linked polymer of acrylamide. The length of the polymer chains is dictated by the concentration of acrylamide used, typically between 5 and 20%. Acrylamide polymerises in head-to-tail fashion to form long polymers, which form a complex network held together by bis-acrylamide crosslinks.

The cross-linked polymers create pores in the gel; the size of pores is determined by the acrylamide concentration, therefore the separation can be tuned depending on the size range of the molecules being separated.

1.5.2 Ferguson plot

The electrophoretic mobility (μ) observed for a macromolecule in solution is determined by Equation 1.1:

$$\mu = v/E = q/f$$

Equation 1.1: Electrophoretic mobility in solution. v = velocity, E = electric field strength, q = total charge of the macromolecule, f = frictional coefficient.

For large DNA molecules, both q and f are directly proportional to DNA molecular weight. Hence, μ is expected to be independent of molecular weight. In gels, μ of a macromolecule is determined primarily by the volume fraction of pores within the gel that the macromolecule can enter. Small DNA molecules migrate through the matrix faster than larger DNA molecules because they encounter less frictional drag in the gel. The relationship between mobility in free electrophoresis and that through a gel was developed by Ferguson¹³⁷ (Equation 1.2)

$$\log \mu_i = \log \mu_{0i} - K_i C$$

Equation 1.2: Electrophoretic mobility of a macromolecule (i) in gel matrix. μ_0 = free solution mobility, C = gel concentration, K_R = retardation coefficient (constant specific of i).

Semilogarithmic plots of mobility vs gel concentration are called Ferguson plots and illustrate the behaviour of various macromolecules (Figure 126). The relative mobility (μ_{ri}) of a macromolecule can be determined relative to a high mobility dye (μ_d) (Equation 1.3).

$$\mu_{ri} = \mu_i / \mu_d = d_i / d_d$$

Equation 1.3: Electrophoretic relative mobility of a macromolecule (i) in gel matrix. μ = mobility, d = distance that the macromolecule (i) or the dye (d) travels.

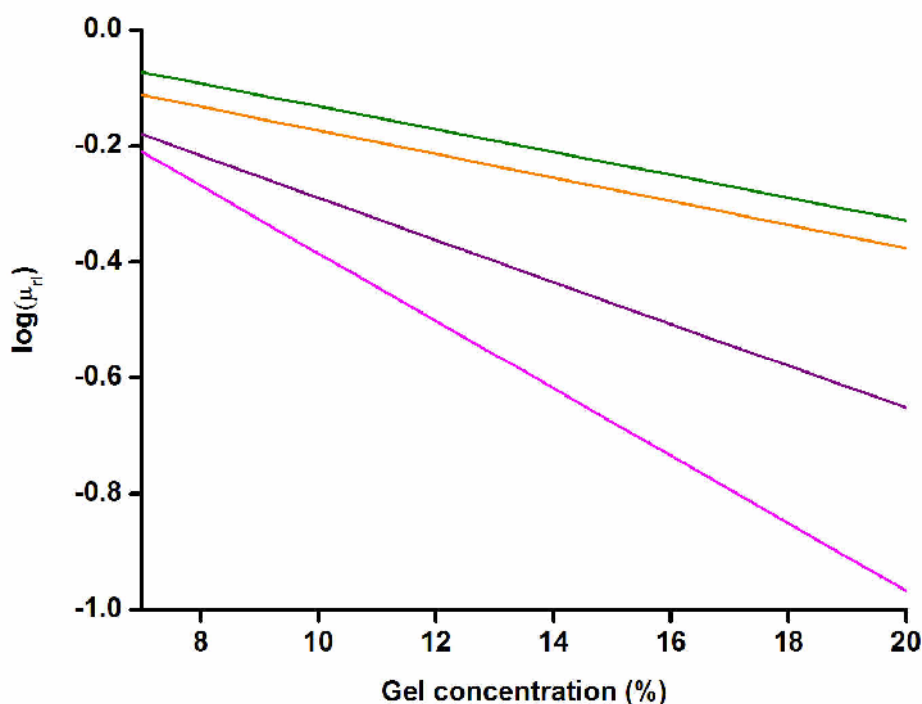


Figure 1.26: Ferguson plot of a mixture of macromolecules of different molecular weights. This situation applies to DNA fragments.

1.6 Methods for studying the properties of nucleic acids

1.6.1 UV melting¹³⁸

Ultraviolet spectroscopy (UV) is a commonly used technique to determine the quantity of absorbing species (chromophores). The chromophores in nucleic acids are the purine and pyrimidine heterocyclic bases and the measure of UV absorption determines the concentration of nucleic acids in solution according to the Beer-Lambert law given below.

$$\text{Absorbance } (A_n) = \log_{10}(I_0/I) = \epsilon c l$$

Equation 1.4: The Beer-Lambert law. A_n = absorbance at n nm, I_0 = intensity of the incident light, I = intensity of the transmitted light, ϵ = molar absorption coefficient (extinction coefficient), c = concentration of the chromophore in mol/L and l = path length in cm.

The UV absorption profiles of nucleosides are due to base compositions which give rise to absorption in the 240-280 nm range at neutral pH. ODNs have strong absorption

maxima (λ_{max}) at 260 nm. The base nitrogen atoms can protonate and therefore such spectra are highly pH dependent. The absorption at 260 nm (A_{260} or OD, optical density) is commonly used as a measure of total nucleic acid concentration in a buffered solution.

The temperature dependence of UV absorption can also be used for determining the stability of nucleic acids structures. The stacked bases in multi-stranded nucleic acid structures absorb less UV light than do unstacked bases. This effect is called hypochromism. The input of thermal energy disrupts the base pairing and causes the strands to separate. The T_m of a duplex is defined as the temperature at which 50% of the ODN strands are hybridised, whilst the other 50% have been separated into single strands. The T_m value is affected by the length of the ODN, its sequence composition and base or sugar modifications.

The T_m of a duplex is determined by calculating the maximum point of the first derivative of the melting curve. A typical melting curve is shown in Figure 1.27.

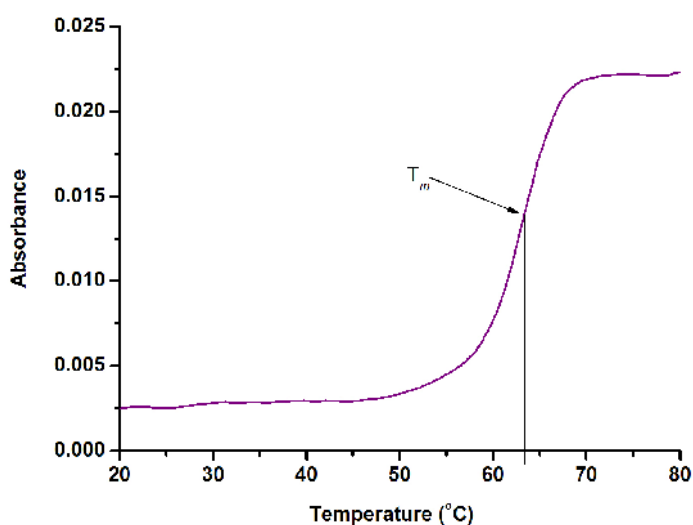


Figure 1.27. Illustration of the melting temperature (T_m) of a DNA duplex.

CHAPTER 2

Synthesis and properties of hydrophobic monomers and oligonucleotides for Addressable Molecular Node Assembly

2. Synthesis and properties of hydrophobic monomers and oligonucleotides for Addressable Molecular Node Assembly

2.1 Introduction

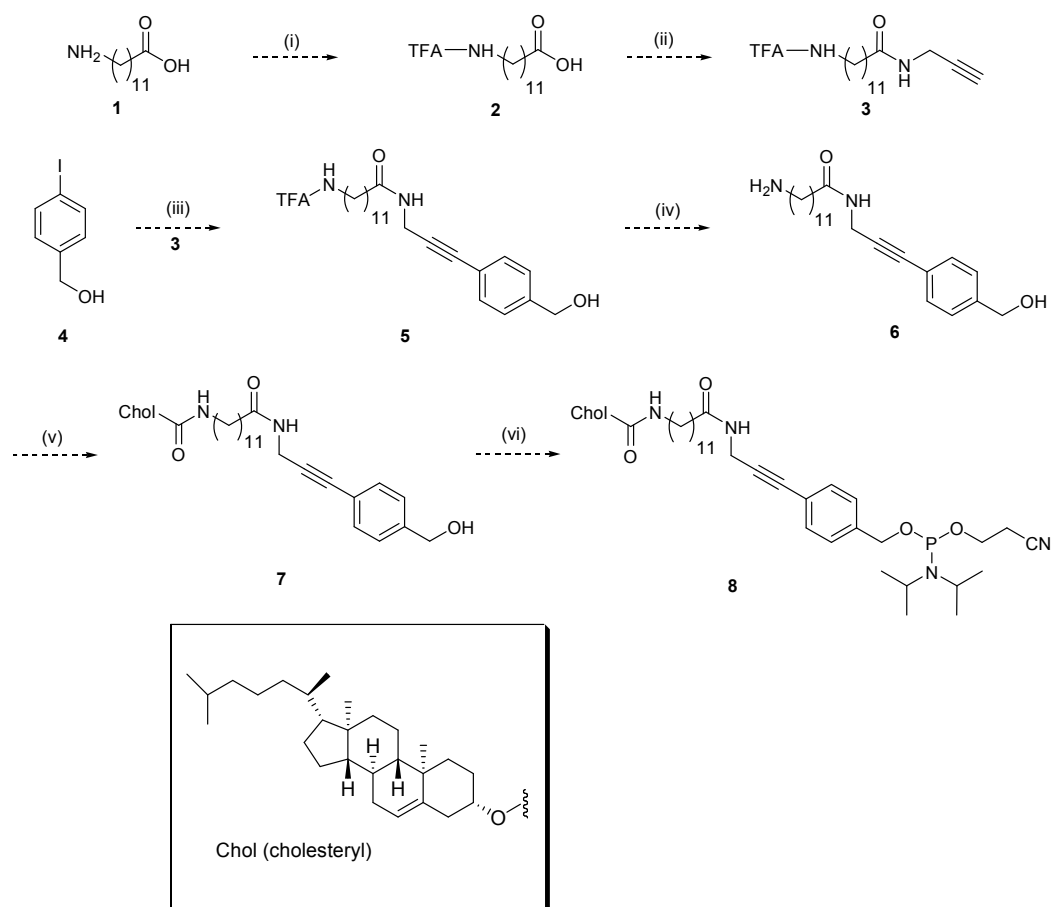
The objective of the AMNA project, of which this PhD research project is a part thereof, is to develop a nanotechnology platform on a grid of addressable molecular building blocks using DNA node structures. Lipophilic groups are incorporated into the oligonucleotides (ODNs) for subsequent anchorage of the network to a lipid surface.

This chapter describes the syntheses of phosphoramidite monomers functionalised with cholesterol and a spacer unit. The spacer provides separation between the hydrophobic cholesterol moiety and the phosphate backbone of the DNA strand. The monomers are subsequently incorporated into oligonucleotides for studying the assembly of Lipid-conjugated oligonucleotides (L-ODNs) in the planar lipid bi-layer.

The main objective was to realise the synthesis of lipophilic phosphoramidite monomers in order to incorporate lipophilic groups into oligonucleotides.

2.2 Cholesterol-linked benzyl alcohol monomer, **8**

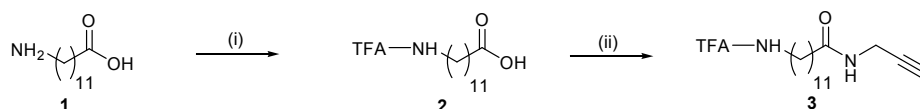
The first monomer to be synthesised was a benzyl alcohol derivative, attached to a cholesterol group, separated by a spacer unit **3** (Scheme 2.1).



Scheme 2.1: *Reagents and conditions* (i) Ethyl trifluoroacetate (1.5 eq), Et₃N (3 eq), MeOH, rt; (ii) prop-2-yn-1-amine (1 eq), EDC (1.3 eq), CH₂Cl₂, rt; (iii) **3** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, rt, dark; (iv) ethanolic methylamine (80 eq), rt; (v) cholesteryl chloroformate (1.2 eq), CH₂Cl₂, rt; (vi) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt.

2.2.1 Synthesis

The amine of commercially available 12-aminododecanoic acid **1**, was protected by reaction with ethyl trifluoroacetate in MeOH with Et₃N at room temperature. The product **2** was coupled with prop-2-yn-1-amine using EDC in CH₂Cl₂ at room temperature. Purification by column chromatography gave **3**, in a 79% overall yield for the two steps.



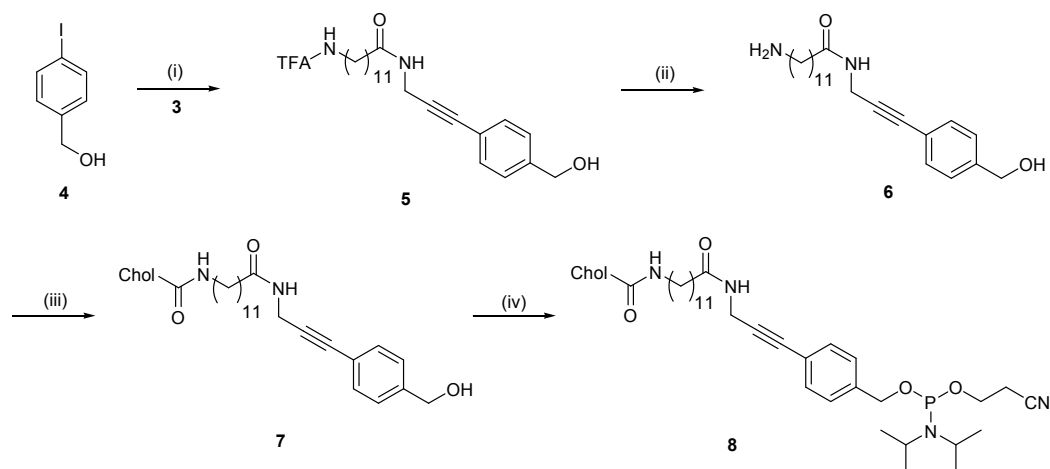
Scheme 2.2: *Reagents and conditions* (i) Ethyl trifluoroacetate (1.5 eq), Et₃N (3 eq), MeOH, 22.5 h, rt; (ii) prop-2-yn-1-amine (1 eq), EDC (1.3 eq), CH₂Cl₂, 24 h, rt, 79% overall.

Spacer synthesis was repeated on a substantial scale (57 mmol) in order to have a sufficient quantity of compound **3** to use in future derivatisation with various lipophilic groups.

Spacer **3** was attached to 4-iodobenzylalcohol **4** *via* a Sonogashira coupling in DMF at room temperature. Purification by column chromatography gave **5** in 75% yield.

The terminal TFA-protected amine of **5** was deprotected using ethanolic methylamine at room temperature to generate 12-amino-*N*-{3'-[4'''-(hydroxymethyl)phenyl]prop-2'-ynyl}dodecanamide **6**. This was coupled with cholesteryl chloroformate in CH₂Cl₂ at room temperature to afford compound **7** in a 19% overall yield for the two steps, after purification by column chromatography. An increased overall yield of 33% was achieved by altering the reaction conditions.¹³⁹ Compound **6** was dried overnight under high vacuum to remove traces of ethanolic methylamine and the quantity of cholesteryl chloroformate was increased to an excess, with 1.2 equivalents.

Compound **7** was then phosphitylated with 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite and DIPEA in THF at room temperature. This reaction successfully afforded the phosphoramidite monomer **8** in a 43% yield after purification by column chromatography (Scheme 2.3).



Scheme 2.3: *Reagents and conditions* (i) **3** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, 2 h, rt, dark, 75%; (ii) ethanolic methylamine (80 eq), 24 h, rt; (iii) cholesteryl chloroformate (1.2 eq), CH₂Cl₂, 2 h at 0 °C then 1.5 h at rt, 33% overall; (iv) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, 1.5 h, rt, 43%.

2.2.2 Biophysical studies

Monomer **8** (Bz^{Chol}) was incorporated into the oligonucleotide ODN-1 (Bz^{Chol} TTTTTTTTTTTT) using standard solid phase synthesis conditions as described in Chapter 4 (see section 4.3.1), to assess coupling efficiency, and to optimise purification protocols. The hydrophobicity of the cholesterol aided HPLC purification by increasing retention time on the column (see Appendix). Analysis by MALDI-TOF MS showed successful incorporation of Bz^{Chol} into the oligonucleotide ([M+H]⁺ calcd, 4421.5; found, 4420.0).

The stability of Bz^{Chol} to aqueous ammonia was studied using ODN-1 and ODN-2 (Bz^{Chol} TTT). ODN-2 was heated at 55 °C in aqueous ammonia for 3 hours; one sample of ODN-1 was also heated at 55 °C in ammonia for 3 hours and another sample for 15 hours. Analysis of all samples by MALDI-TOF MS indicated in all cases, both the intact oligonucleotide (ODN-1, ODN-2) and the corresponding oligonucleotides without the cholesterol modification.

The sample of ODN-1 which was heated for 15 hours, was purified by HPLC and the main fraction was analysed by MALDI-TOF MS. The resultant peak corresponded to the loss of Bz^{Chol}.

ODNs / Conditions		Expected mass [M+H] ⁺	Observed mass [M+H] ⁺
ODN-2 heated 3 h		1683.8	1680.7 and 933.8
ODN-1	heated 3 h	4421.5	4416.4 and 3666.4
	heated 15 h	4421.5	4421.7 and 3669.3
	heated 15 h and purified	4421.5	3669.0

Table 2.1: MALDI-TOF MS analysis on a ThermoBioAnalysis Dynamo MALDI-TOF mass spectrometer in positive ion mode using internal T_n standards.¹⁴⁰ (see Section 4.1).

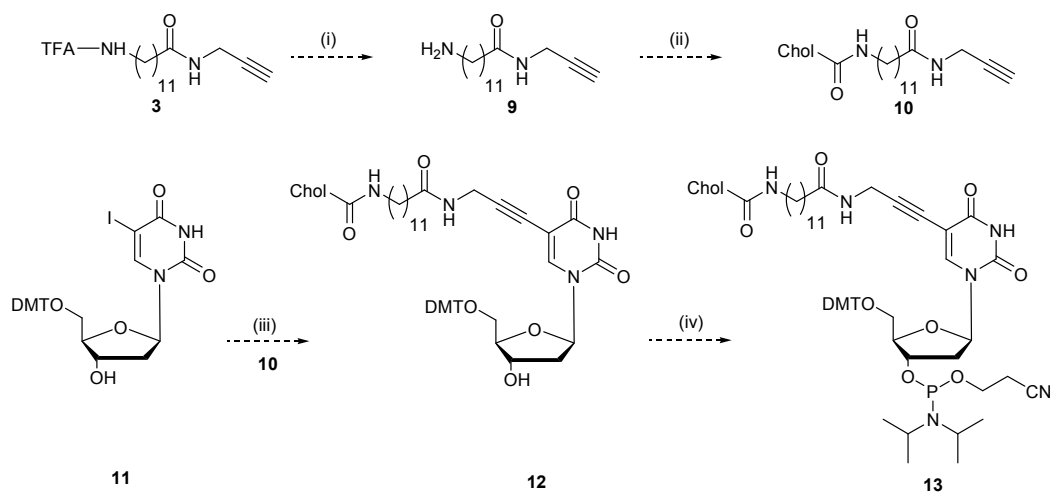
It is possible that the peak at mol wt 3669.0 could have been due to incomplete coupling caused by presence of residual phosphitylating agent, therefore Bz^{Chol} was purified a second time and then incorporated into ODN-2. One sample was treated in aqueous ammonia for 30 min and a second was heated in aqueous ammonia for 65 hours. MALDI-TOF analysis showed two peaks, which were identical to those obtained in the first experiment. Since there is no difference in the ratio of oligonucleotide to decomposition product under 30 min or 65 hours at 55 °C in aqueous ammonia, this is not a problem caused by instability in ammonia. Decomposition could be occurring due to a problem with purification of the monomer. It is also possible that the desired oligonucleotide was retained on the HPLC column, or that coupling of Bz^{Chol} was incomplete.

ODNs / Conditions	Expected mass [M+H] ⁺	Observed mass [M+H] ⁺
ODN-2 treated 30 min	1683.8	1684.8 and 946.6
ODN-2 treated 65 h	1683.8	1685.5 and 931.5

Table 2.2 : MALDI-TOF MS analysis on a ThermoBioAnalysis Dynamo MALDI-TOF mass spectrometer in positive ion mode using internal T_n standards.¹⁴⁰

2.3. Synthesis of the Thymidine monomer linked to Cholesterol by a hydrophobic spacer, 13

The cholesterol moiety was attached this time to a 2'-deoxythymidine nucleoside, separated by the same hydrophobic spacer unit **3** (Scheme 2.4).



Scheme 2.4: *Reagents and conditions* (i) Ethanolic methylamine (90 eq), rt; (ii) cholesteryl chloroformate (1.2 eq), CH₂Cl₂, 2 h at 0 °C then at rt; (iii) **10** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, rt, dark; (iv) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt.

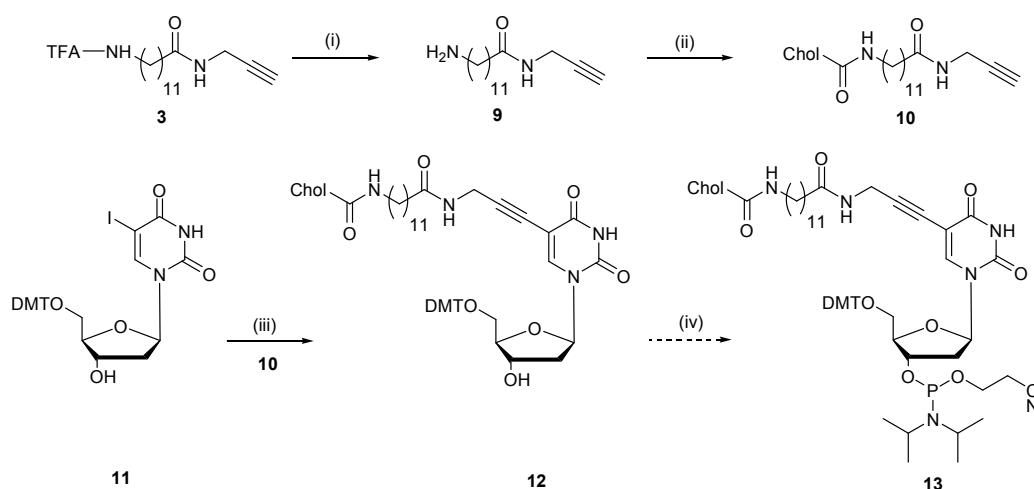
2.3.1 First route

The first step toward synthesis of this cholesterol-linked thymidine monomer was to deprotect the amine then attach the cholesterol prior to the Sonogashira reaction. This prevents reaction between cholesteryl chloroformate and the secondary 3'-alcohol of the nucleoside. Secondly, the acidic nature of the cholesteryl chloroformate could cleave the acid-labile DMT of the 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-deoxyuridine **11**.

The amine of **3** was deprotected with ethanolic methylamine to give 12-amino-*N*-(prop-2-ynyl)dodecanamide **9**. This was subsequently coupled with cholesteryl chloroformate by the same method as in Scheme 2.3. Following purification by column

chromatography, 12-(cholesteryloxycarbonylamino)-*N*-(prop-2-ynyl) dodecanamide **10**, was afforded in a 40% overall yield over the two steps (Scheme 2.5).

The conditions for the Sonogashira reaction between alkyne **10** and 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **11**¹⁴¹ (synthesised by Chaturong Suparpprom) were investigated, because standard conditions^{142,143} afforded **12** in only 17% yield.

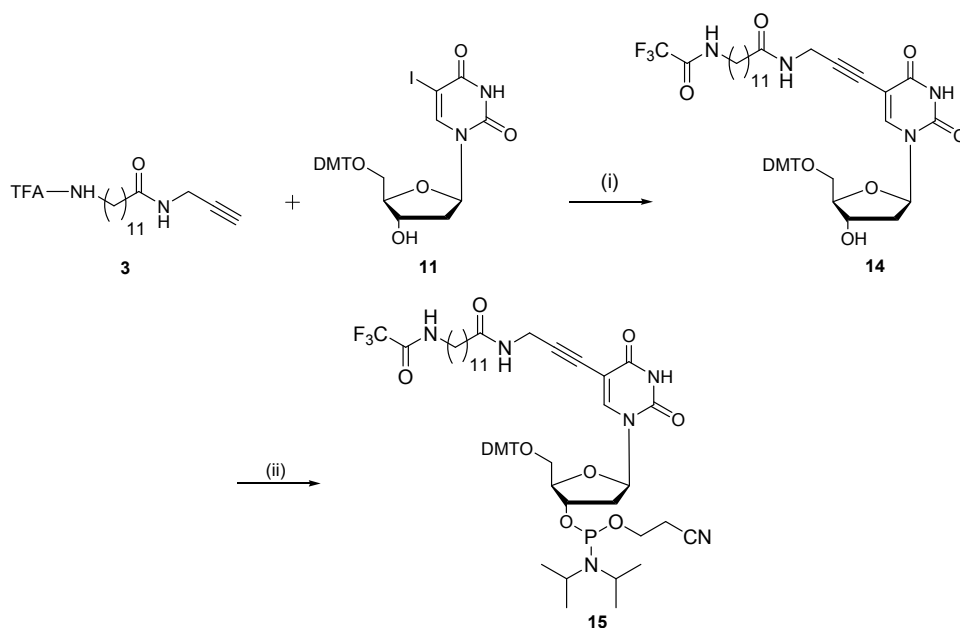


Scheme 2.5: *Reagents and conditions* (i) Ethanolic methylamine (90 eq), 31 h, rt; (ii) cholesteryl chloroformate (1.2 eq), CH₂Cl₂, 2 h at 0 °C then 2 h at rt, 40% over two steps; (iii) **10** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, 2 h, rt, dark.; (iv) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophoramidite (1.1 eq), DIPEA (2.5 eq), THF, 2 h, rt.

The first approach in solving this problem was to examine whether the number of equivalents of Et₃N or the temperature made any difference to the yield obtained. Reactions were carried out using either 5 or 20 equivalents of Et₃N and the temperature at either rt or 60 °C. However, the results of these tests were inconclusive. The second approach employed microwave techniques, but this did not improve the yield either. The reaction was repeated using unprotected thymidine and 50 equivalents of Et₃N, but the compound obtained was not pure and the yield was low. This poor yield could be rationalised by the bulkiness of the cholesterol, which may have prevented the coupling to take place.

2.3.2 Second route

As the first route did not work, an alternative route to monomer **13** was considered. The Sonogashira cross-coupling was achieved by using the protected thymidine **11** and a shorter alkyne **3**, as shown in Scheme 2.6. Purification by column chromatography gave **14** in a 66% yield. This compound was then subjected to conventional phosphitylation conditions: addition of 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite and DIPEA in THF under an argon atmosphere to give **15** in a 47% yield (Scheme 2.6).

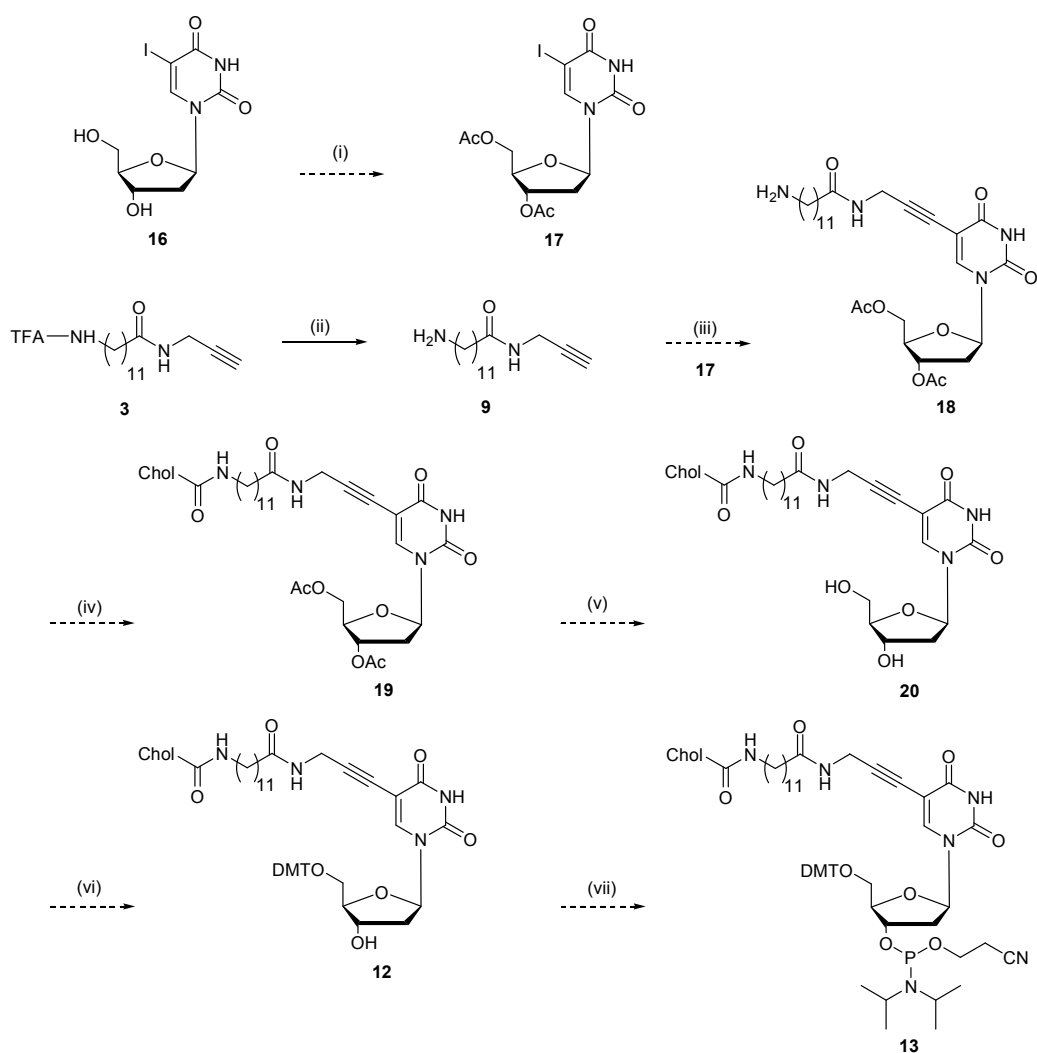


Scheme 2.6: *Reagents and conditions* (i) CuI (0.4 eq), Pd(PPh₃)₄ (0.2 eq), Et₃N (10 eq), DMF, rt, dark, 23 h, 66%; (ii) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, 1.5 h, rt, 47%.

The amine could be easily deprotected to enable the attachment of a dye, such as rhodamine, to compound **14**. After phosphitylation, the monomer could be incorporated into oligonucleotides using standard solid phase oligonucleotide synthesis and then deprotected under standard deprotection conditions.

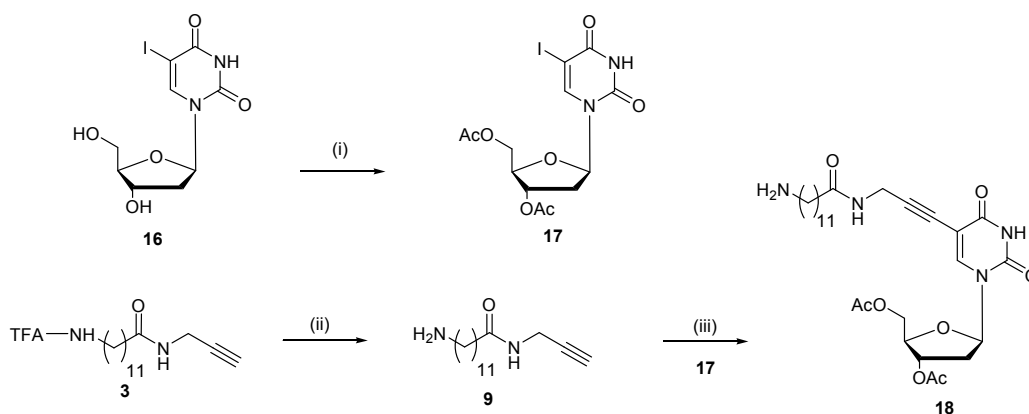
Synthesis of the cholesterol-modified monomer using the shorter alkyne **3**, requires that the Sonogashira reaction between the deprotected alkyne **9** and the 5-iodo-2'-

deoxyuridine derivative, is carried out before the cholesterol is coupled to the alkyne. It is necessary therefore, to protect the nucleoside at the 3'- and 5'-positions (acetyl) to avoid a possible reaction between these hydroxyl groups and cholesteryl chloroformate. Then the 5'-position would have to be tritylated following deprotection of the acetyl groups. The last step is the phosphitylation of the 3'-position to give monomer **13**, as shown in Scheme 2.7.



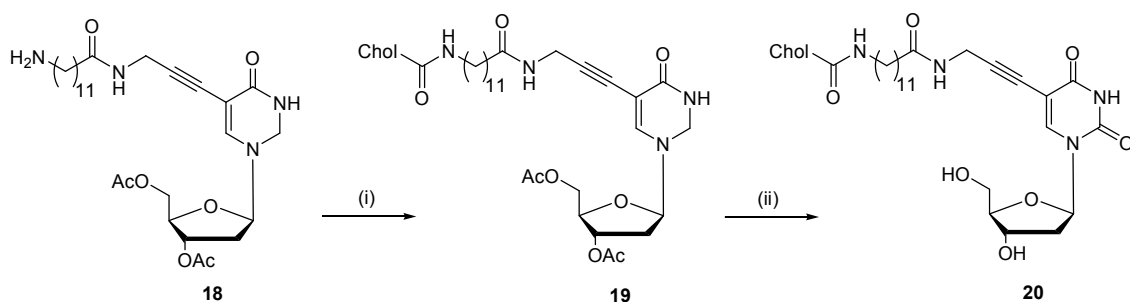
Scheme 2.7: *Reagents and conditions:* (i) Ac_2O (3 eq), pyridine, rt; (ii) ethanolic methylamine (70 eq), rt, 24 h; (iii) **17** (1.1 eq), CuI (0.2 eq), $\text{Pd(PPh}_3)_4$ (0.1 eq), Et_3N (5 eq), DMF, rt, dark; (iv) cholesteryl chloroformate (1.2 eq), CH_2Cl_2 ; (v) ethanolic methylamine; (vi) 4,4'-dimethoxytrityl chloride (1.2 eq), pyridine, rt; (vii) 2-*O*-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt.

5-Iodo-2'-deoxyuridine **16** was protected in the 5'- and 3'-positions using acetic anhydride in pyridine to generate **17** quantitatively after column chromatography (Scheme 2.8). The amino group of the spacer **3** was deprotected using ethanolic methylamine at 55 °C to give 12-amino-*N*-(prop-2-ynyl)dodecanamide **9**. After further investigations into the conditions of the Sonogashira coupling, compound **9** was attached to the protected nucleoside **17** in DMF at room temperature followed by a simple filtration through celite as workup. Purification by column chromatography gave **18** in a 73% overall yield (Scheme 2.8).



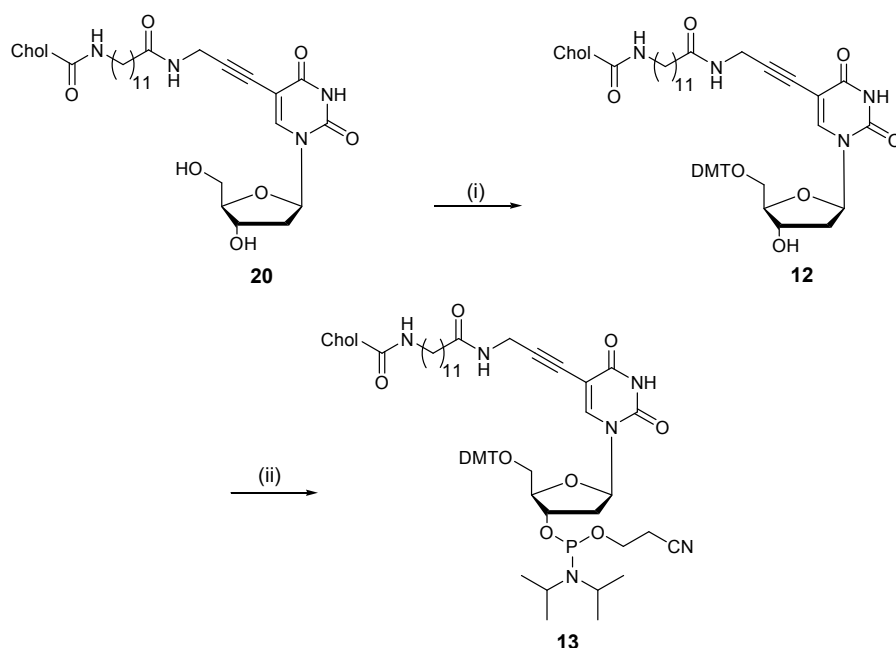
Scheme 2.8: *Reagents and conditions:* (i) Ac₂O (3 eq), pyridine, rt, 16.5 h, quant.; (ii) ethanolic methylamine (70 eq), 55 °C, 20 h; (iii) **17** (2.2 eq), CuI (0.4 eq), Pd(PPh₃)₄ (0.2 eq), Et₃N (40 eq), DMF, rt, dark, 30.5 h, 73% overall.

Compound **18** was coupled with cholesteryl chloroformate in a mixture of CH₂Cl₂ and DMF at room temperature with DIPEA as a base. Purification by column chromatography gave **19** in only 32% yield. The first conditions used for the deacetylation of **19** were ethanolic methylamine in CH₂Cl₂, but analysis of the product did not show formation of the free nucleoside. Reactions were carried out using either 15 or 62 equivalents of ethanolic methylamine and the temperature at either rt or 55 °C. However, the results of these tests were inconclusive. Another attempt at deprotecting **19** using potassium carbonate in a mixture of methanol and water successfully resulted in **20**.



Scheme 2.9: *Reagents and conditions:* (i) Cholesteryl chloroformate (1.2 eq), DIPEA (1.2 eq), $\text{CH}_2\text{Cl}_2/\text{DMF}$, 2 h at 0 °C and 2 h at rt, 32%; (ii) 0.4 M K_2CO_3 in $\text{MeOH}:\text{H}_2\text{O}$ (4:1, v/v) (6 eq), 1 h 45 min, rt.

Without further purification, **20** was tritylated with 4,4'-dimethoxytrityl chloride in pyridine and CH_2Cl_2 . Purification by column chromatography gave **12** in a 45% overall yield. Compound **12** was then phosphitylated with 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite and DIPEA in THF at room temperature. This reaction successfully afforded the phosphoramidite monomer **13** in a 82% yield after purification by column chromatography.



Scheme 2.10: *Reagents and conditions:* (i) 4,4'-dimethoxytrityl chloride (3 eq), pyridine, CH_2Cl_2 , rt, 5 h, 45% overall; (vii) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt, 0.5 h, 82%.

2.3.3 Biophysical studies

2.3.3.1 10-mers

Monomer **13** (T^{chol}) was incorporated into oligonucleotides ODN-3 and ODN-4 using standard automated DNA synthesis conditions as described in section 4.3.1. UV melting experiments were performed on a Varian Cary 400 Scan UV-Vis spectrophotometer, monitoring at 260 nm, in buffer at pH 7.0 (10 mM sodium phosphate with 0.2 M NaCl and 1 mM Na_2EDTA). A 1:1 ratio of each oligonucleotide at 1 μM concentration was used. Full details are given in section 4.3.4. A schematic representation of the UV melting duplex is shown below, in Figure 2.1.

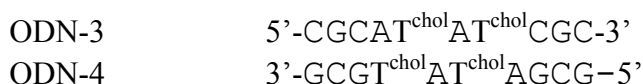


Figure 2.1: Duplex UV melting experiment. T^{chol} = monomer **13**.

Using a standard melt programme (20-80-20 °C at 1 °C/min, see section 4.3.4), UV melting of these two oligonucleotides and each with their non-modified complementary strand showed no duplex formation. The concentration of salt was then increased to 0.5 M and 1 M NaCl but no duplex was formed with ODN-3 and ODN-4 together, or with their non-modified complementary strands. A buffer solution of 0.2 M NaCl with 1% sodium dodecyl sulphate was used, in order to aid solubilisation of the cholesterol moiety. Again no duplex was formed.

These oligonucleotides were studied by fluorescence melting using a Roche LightCycler[®] and the fluorescent intercalator SYBR Green I, which was used to measure the fluorescence change upon heating the oligomer mixtures. The studies were performed using a 10 mM sodium phosphate buffer with 0.2 M NaCl and 1 mM Na_2EDTA at pH 7.0 with intercalator:total volume 1:10 v/v. The fluorescence melting results did not show any duplex formed.

2.3.3.2 18-mers

T^{Chol} was then incorporated into the oligonucleotides ODN-5 and ODN-6, 18-mers with only one modification. Duplex formation between ODN-5 and two probe oligonucleotides was analysed by UV melting experiments. A schematic representation of the UV melting duplex is shown below, in Figure 2.2.

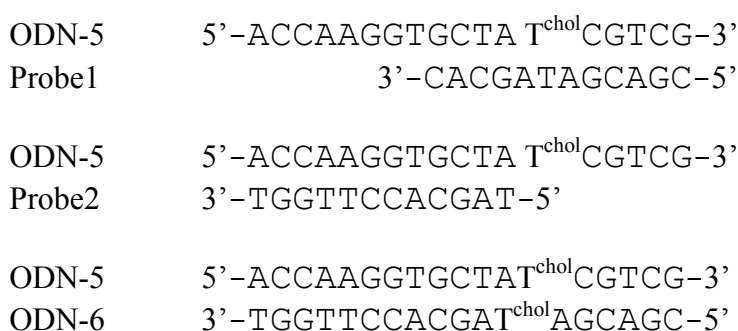


Figure 2.2: Duplex UV melting experiment. T^{chol} = monomer 13.

The UV melting experiment was performed at pH 7.0 (10 mM sodium phosphate buffer with 0.2 M NaCl and 1 mM Na₂EDTA). A standard duplex melt programme was used (20-80-20 °C at 1 °C/min, see section 4.3.4).

ODNs	ODN-5/Probe 1	ODN-5/Probe 2	ODN-5/ODN-6
T_m (°C)	48.7	54.0	nd

Table 2.3: Duplex UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer with 0.2 M NaCl and 1 mM Na₂EDTA) using a standard melt programme (20-80-20°C at 1°C/min). nd = not determined- T_m above 70 °C.

These results show that the T_m of the duplex between the two modified oligonucleotides is too high to be measured by UV melting.

These oligonucleotides were studied by fluorescence melting using the fluorescent intercalator SYBR Green I. Experiments were conducted at pH 7.0 (10 mM sodium

phosphate buffer with 0.2 M, 0.4 M or 0.6 M NaCl, and 1 mM Na₂EDTA) with intercalator:total volume 1:10 v/v .

buffer with NaCl 0.2 M			
ODNs	ODN-5/Probe1	ODN-5/Probe2	ODN-5/ODN-6
T_m (°C)	44.5	46.2	67.6
buffer with NaCl 0.4 M			
ODNs	ODN-5/Probe1	ODN-5/Probe2	ODN-5/ODN-6
T_m (°C)	42.6	46.9	75.7
buffer with NaCl 0.6 M			
ODNs	ODN-5/Probe1	ODN-5/Probe2	ODN-5/ODN-6
T_m (°C)	39.7	46.5	78.4

Table 2.4: Fluorescence melting experiment at pH 7.0 (10 mM sodium phosphate buffer with 0.2 M, 0.4 M or 0.6 M NaCl, and 1 mM Na₂EDTA). Each T_m is an average of two separate runs.

The fluorescence melting data shows the dependence of the T_m of the duplex ODN-5/Probe1 on salt concentration. This is likely due to increased hydrophobic interaction between the cholesterol moiety and the DNA bases at high salt concentration and possible formation of a hairpin structure. However, the duplex formation of ODN-5/ODN-6 is enhanced by increasing the concentration of NaCl. In this case, the two cholesterol groups interact more favourably with each other with increasing salt concentration.

Duplex formation between ODN-5 and ODN-6 was studied by fluorescence melting (see Figure 2.2). A comparison of the data obtained with the T^{Chol} modified oligonucleotides (ODN-5, ODN-6) and the control T oligonucleotides (ODN-5' and ODN-6') is showed in Figure 2.4. Experiments were performed using the same conditions as described above. A schematic representation of the UV melting duplex is shown below, in Figure 2.3

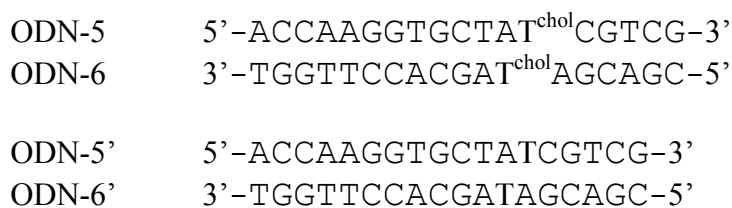


Figure 2.3: Duplex UV melting experiment. T^{chol} = monomer **13**.

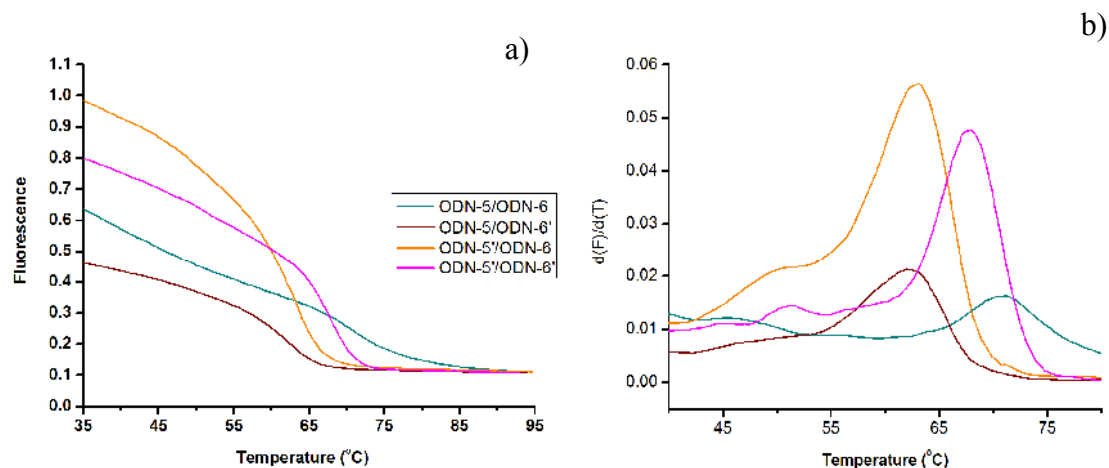


Figure 2.4: (a) Fluorescence melting curves at pH 7.0 (10 mM sodium phosphate buffer with 0.2 M NaCl and 1 mM Na₂EDTA). The concentration of each ODN was 0.5 μ M. (b) d(F)/d(T) Melting data were processed using the OriginPro 7.5 software.

The T_m of the duplex ODN-5/ODN-6, is higher than that of the natural duplex ODN-5'/ODN-6' (+2.4 $^{\circ}$ C) and much higher than that of ODN-5'/ODN-6 and ODN-5/ODN-6' (+8.1 $^{\circ}$ C and +8.8 $^{\circ}$ C respectively) (Table 2.5).

buffer with NaCl 0.2 M				
ODNs	ODN-5'/ODN-6'	ODN-5'/ODN-6	ODN-5/ODN-6'	ODN-5/ODN-6
T_m ($^{\circ}$ C)	67.6	61.9	61.2	70.0

Table 2.5: Fluorescence melting experiment at pH 7.0 (10 mM sodium phosphate buffer with 0.2 M NaCl and 1 mM Na₂EDTA). Each T_m is an average of two separate runs.

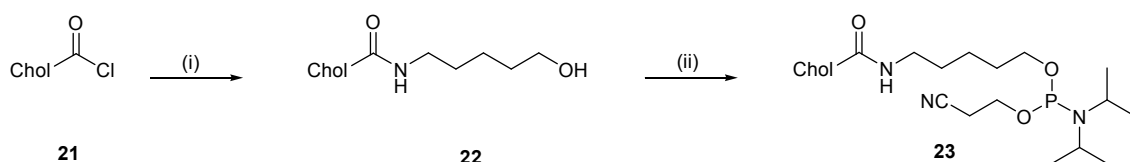
The fluorescence melting data in Table 2.5 clearly illustrates a lower T_m of the duplex with a cholesterol moiety on one DNA strand, but higher melting temperature when cholesterol is present on both strands.

2.4. Synthesis of 6-(cholesteryloxycarbonyl)-hexanol phosphoramidite, **23**

Since the first phosphoramidite **8** was not stable under standard oligonucleotide deprotection conditions, a monomer was proposed with a simple alkyne chain as linker between the cholesterol and the phosphoramidite moiety.

2.4.1 Synthesis

Commercially available 6-amino-hexanol **21** was reacted with cholesteryl chloroformate in CH_2Cl_2 at room temperature. Purification by column chromatography gave **22** in 86% yield. The next and final step was the phosphitylation of **22** with 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite and DIPEA in THF under an argon atmosphere which gave **23** in a 43% yield.



Scheme 2.11: *Reagents and conditions:* (i) 6-amino-hexanol (2.2 eq), CH_2Cl_2 , 2 h at 0°C then 16.5 h at rt, 86%; (ii) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt, 0.5 h, 43%.

2.4.2 Biophysical studies

Monomer **23** (H^{Chol}) was incorporated into oligonucleotide ODN-7, using standard solid phase oligonucleotide synthesis to assess the coupling efficiency, and to optimise the purification protocols. Due to the very high hydrophobic nature of the cholesterol,

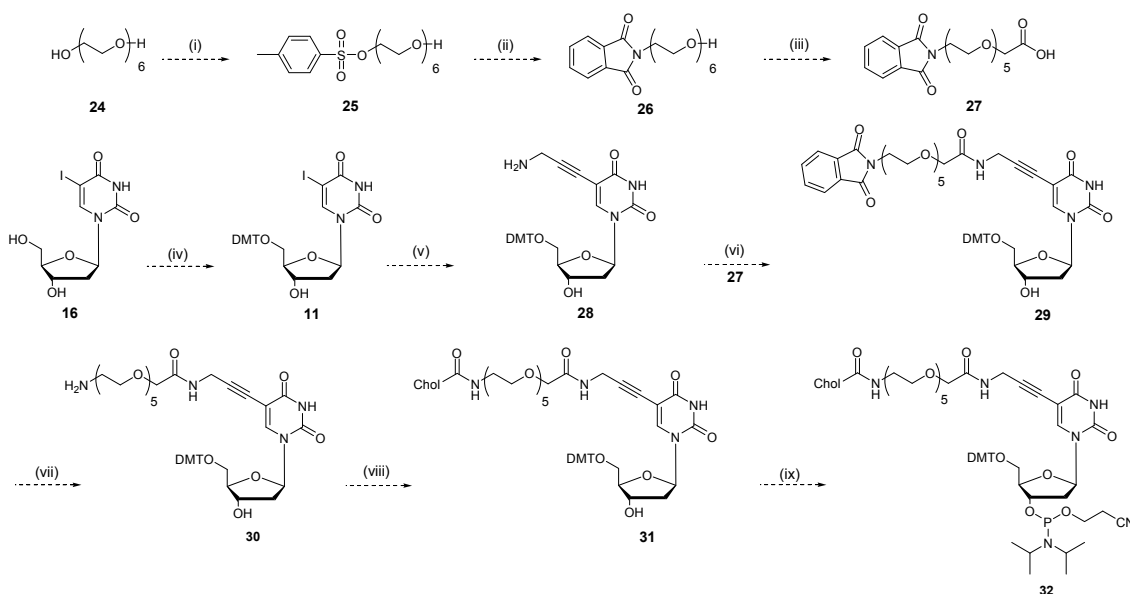
the purification of the oligonucleotide on a reverse phase column has only been successful analytically. Analysis by MALDI-TOF MS showed the successful incorporation of the H^{Chol} into ODN-7 ($[M+H]^+$ calcd 4235.45, found 4234.38).

2.5. Synthesis of the thymidine monomer, linked to cholesterol by a hydrophilic spacer

The use of a hydrophobic spacer for monomer **13** would allow its penetration into the lipid bilayer. A linker such as poly(ethylene glycol), which is hydrophilic, would allow the monomer to anchor to the lipid bilayer *via* the cholesterol moiety only, rather than the cholesterol-linker unit.

A new nucleoside-based cholesteryl 2'-deoxythymidine phosphoramidite monomer **34** was synthesised, with a hydrophilic hexa(ethylene glycol) linker for use in oligonucleotide synthesis.

The route to the target is described below in Scheme 2.12.



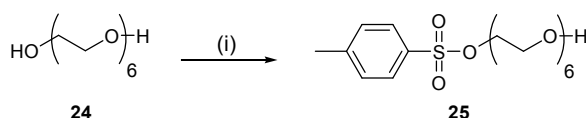
Scheme 2.12: *Reagents and conditions* (i) TsCl (0.25 eq), Et₃N (1 eq), DMAP (0.12 eq), CH₂Cl₂, rt; (ii) potassium phthalimide (1.5 eq), DMF, reflux; (iii) TEMPO (0.2 eq), BAIB (3eq), H₂O:CH₃CN (1:1 v./v), rt; (iv) 4,4'-dimethoxytrityl chloride (1.2 eq), pyridine, rt; (v) prop-2-yn-1-amine (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, rt, dark; (vi) **27** (1eq), EDC (1.3 eq), CH₂Cl₂, rt; (vii) hydrazine monohydrate (98% w/v) (8 eq), EtOH, rt; (viii) cholesteryl chloroformate (1.2 eq), DIPEA (1.2 eq),

CH₂Cl₂, 2 h at 0 °C then rt; (ix) 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt.

2.5.1 Synthesis

2.5.1.1 Tosylation of hexaethylene glycol

Sutherland¹⁴⁴ described the synthesis of hexaethylene glycol mono-*p*-toluenesulfonate **25** using a workup procedure which avoids purification by column chromatography and allows recovery of the starting material. Using this method, one terminal alcohol of **24** was selectively tosylated using toluenesulfonyl chloride with Et₃N and DMAP in CH₂Cl₂. The reaction mixture was washed with distilled water, saturated sodium bicarbonate and citric acid solutions. However, the extraction did not produce a pure product. The reaction was repeated following the same procedure, but an attempt to recrystallise from ethyl acetate and petroleum ether 40-60 °C was not successful. Finally, column chromatography afforded **25** in 54% yield.



Scheme 2.13: *Reagents and conditions* (i) TsCl (0.25 eq), Et₃N (1 eq), DMAP (0.12 eq), CH₂Cl₂, 2 h at 0 °C then 5 h at rt, 54%.

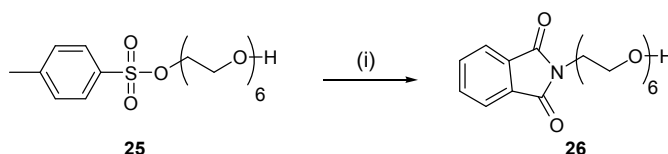
This step has been repeated to improve the yield of the tosylation from 52% to 60%.

2.5.1.2 Displacement of the tosyl group with phthalimide

Following Muller's procedure,¹⁴⁵ **25** was reacted with potassium phthalimide in DMF at reflux, but no precipitate was observed after 4h, contrary to the description in the paper. Diethyl ether was added and the reaction was stirred for a further 30 mins in an ice bath. The precipitate was then filtered, washed with ether, and then dissolved in CH₂Cl₂. The insoluble impurities were filtered off and the filtrate was concentrated. The residue was then dissolved in CH₂Cl₂ but no precipitate was formed with addition of ether, and a

pure product could not be obtained. Another attempt of making HEG-phthalimide **26** was carried out by modifying the workup procedure. After removal of DMF, the residue was redissolved in ethyl acetate and washed with water and saturated sodium bicarbonate. Again, a pure product was not obtained.

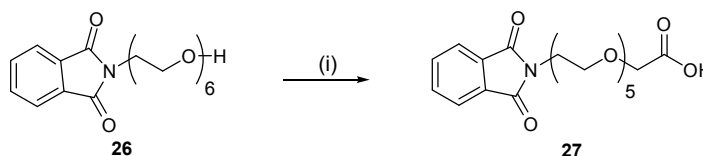
The reaction was repeated a third time and the workup procedure altered again. After removing DMF, the residue was dissolved in CH₂Cl₂ and washed with water and then saturated potassium chloride solution. The aqueous layer was re-extracted three times with CH₂Cl₂. The resultant residue was purified by column chromatography and gave **26** in 78% yield.



Scheme 2.14: *Reagents and conditions* (i) Potassium phthalimide (1.5 eq), DMF, 8 h, reflux, 78%.

2.5.1.3 Oxidation of the primary alcohol

The oxidation of the primary alcohol was carried out according to the procedure described by Bessodes¹⁴⁶ on a 1 g scale. Compound **26** was dissolved in a 1:1 v/v mixture of water and acetonitrile with TEMPO as a catalyst and BAIB as co-oxidant. Purification by column chromatography gave **27** in only 28% yield, contrary to the quantitative yield obtained in the procedure. No other product and starting material was seen on TLC except BAIB which has been added in excess.

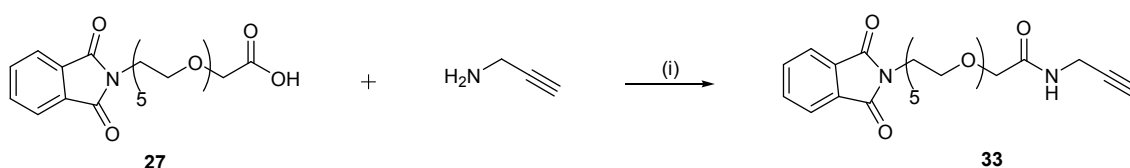


Scheme 2.15: *Reagents and conditions* (i) TEMPO (0.2 eq), BAIB (3eq), H₂O:CH₃CN (1:1 v/v), 3.5 h, rt, 72%.

An alternative oxidation using Jones' reagent was then carried out. The reaction was successful on a small scale, as observed by MS. At the same time, the catalytic oxidation with BAIB was carried out again and the conditions for the silica gel column chromatographic purification were changed. The solvent was changed from a gradient of MeOH in CH₂Cl₂ to isocratic elution using CH₂Cl₂ with 3% MeOH to give **27** in 72% yield.

2.5.1.4 Amide coupling reaction

It was decided that firstly the carboxylic acid **27** would be coupled with prop-2-yn-1-amine.



Scheme 2.16

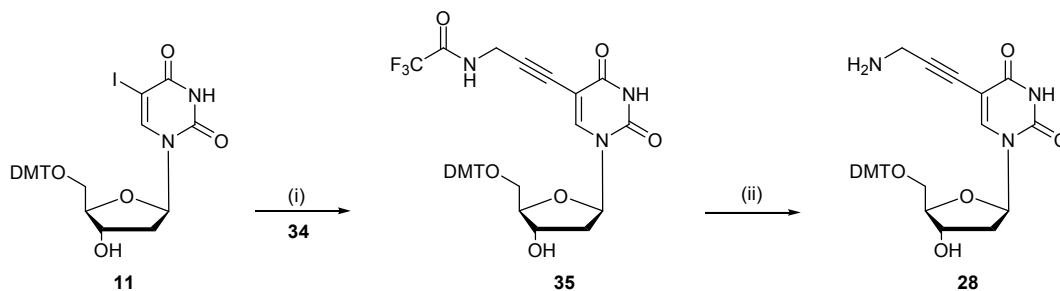
Several conditions using different coupling reagents were tried as detailed below (Scheme 2.16):

- CDI with DIPEA in THF: 21% yield
- HBTU and HOBT with DIPEA in MeCN: 37% yield
- EDC in CH₂Cl₂: 22% yield
- EDC and HOBT with DIPEA in CH₂Cl₂: not purified

All these conditions showed several spots on TLC and gave low yields.

Following this, the order of the steps in the synthetic protocol was changed. The Sonogashira reaction was carried out with unprotected prop-2-yn-1-amine but column chromatographic purification over silica gel was difficult, due to the product's high polarity. Thus, the reaction was carried out with 2',2',2'-trifluoro-*N*-(prop-2-ynyl)acetamide **34** previously synthesised in the laboratory by Simon Gerrard) and

5'-*O*-(4,4-dimethoxytrityl)-5-iodo-2'-deoxyuridine **11** (repurified from a previous synthesis in the laboratory by Simon Gerrard).



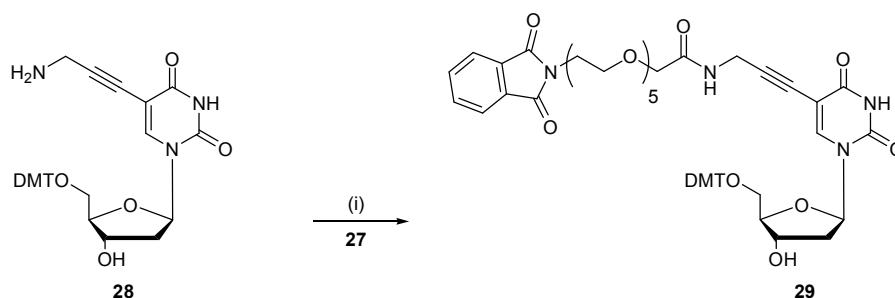
Scheme 2.17: *Reagents and conditions* (i) 2',2',2'-trifluoro-*N*-(prop-2-ynyl)acetamide **34** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, 1.5 h, rt, dark, 96%; (ii) 33 wt % ethanolic methylamine (50 eq), rt, 2 h.

The nucleoside has an acid-labile dimethoxytrityl group on the 5'-position, therefore column chromatography was carried out in the presence of a small amount of Et₃N. However, the TLC showed that cyclisation to the furanopyrimidine occurred during or after purification. This is a known side reaction that occurs if the product is left in the presence of CuI and/or Et₃N for an extended period.^{147,148} In order to avoid this issue, the column chromatography was carried out with pyridine affording **35** in an excellent 96% yield. The amine was then deprotected with ethanolic methylamine to give **28**, which was used without purification (Scheme 2.17).

In order to investigate the reactivity of the amine in the coupling reaction, a similar nucleoside with two more carbons between the free amine and the alkyne, was compared with **28**, using two different conditions:

- EDC and HOBT with DIPEA in CH₂Cl₂
- EDC and DMAP with DIPEA in CH₂Cl₂.

The reactions worked in all four cases, but with several byproducts, as shown by TLC.

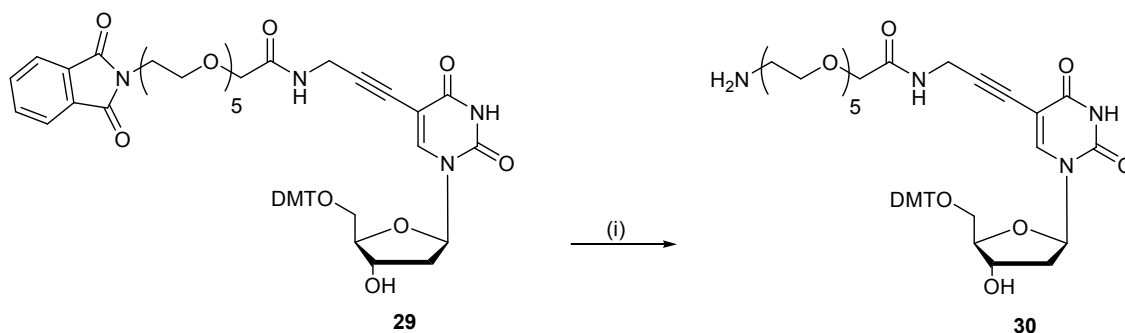


Scheme 2.18: Reagents and conditions (i) **27** (1.1 eq), PyBrOP (1.1 eq), DIPEA (3 eq), CH₂Cl₂, 1.5 h, rt, 49%.

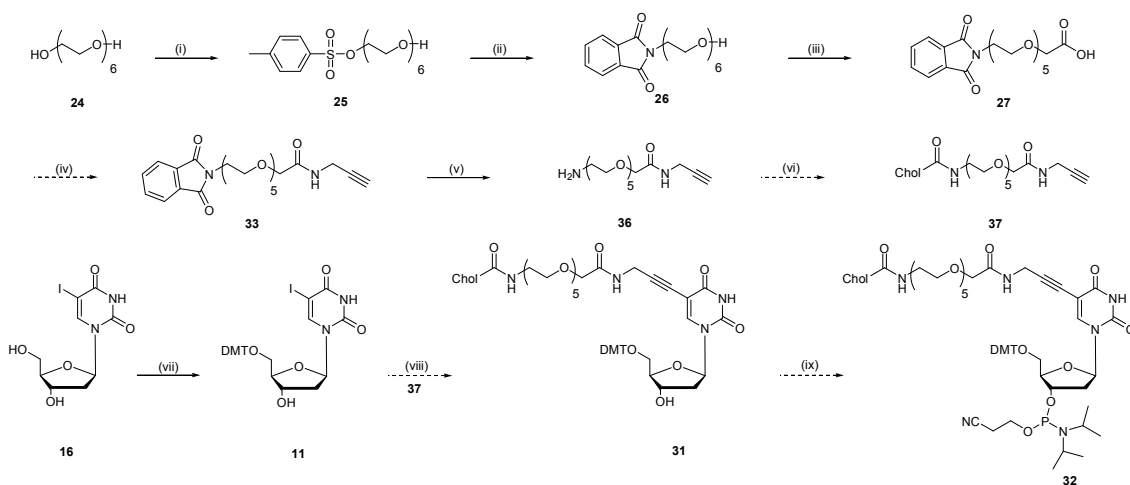
Two other coupling reagents, PyBOP and PyBrOP¹⁴⁹ were then compared with HBTU in CH₂Cl₂ in the above reaction on 5-[3-amino-(propynyl)]-2'-deoxyuridine **28**. Both showed clean formation of the correct product by TLC, but PyBrOP was much quicker to react. The reaction was carried out on 1 g scale and gave **29** in 49% yield (Scheme 2.18).

2.5.1.5 Hydrazinolysis

Cleavage of the phthalimide protecting group to give the primary amine was carried out in 96% yield using hydrazine monohydrate in a mixture of CH₂Cl₂ and ethanol. The product of the reaction is not readily soluble, which gives broad NMR spectra.

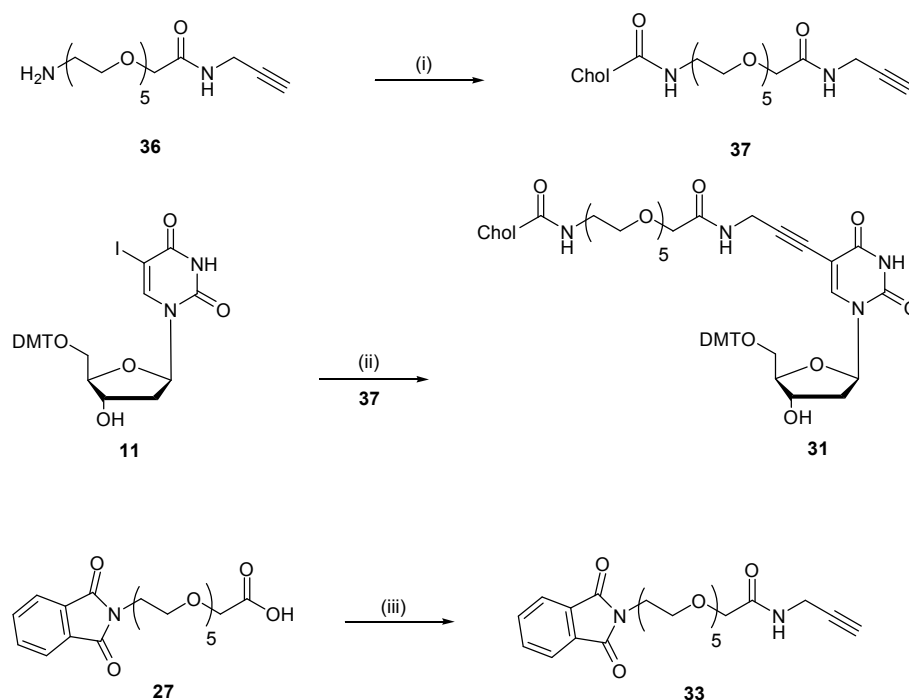


Scheme 2.19: Reagents and conditions (i) Hydrazine monohydrate (98% w/v) (8 eq), CH₂Cl₂/ethanol, 3 h, rt, 96%.



Scheme 2.21: Reagents and conditions (i) TsCl (0.25 eq), Et₃N (1 eq), DMAP (0.12 eq), CH₂Cl₂, 2 h at 0°C then 5 h at rt, 60%; (ii) potassium phthalimide (1.5 eq), DMF, 8 h, reflux, 80%; (iii) TEMPO (0.2 eq), BAIB (3 eq), H₂O:CH₃CN (1:1 v/v), 3.5 h, rt, 72%; (iv) prop-2-yn-1-amine (1.1 eq), PyBrOP (1.1 eq), DIPEA (3 eq), CH₂Cl₂, rt; (v) hydrazine monohydrate (98% w/v) (8 eq), EtOH, 1.25 h, rt, 71%; (vi)olesteryl chloroformate (1.2 eq), DIPEA (1.2 eq), CH₂Cl₂, 2 h at 0°C then rt; (vii) 4,4'-dimethoxytrityl chloride (1.2 eq), pyridine, 3.5 h, rt, 75%; (viii) **37** (1.1 eq), CuI (0.2 eq), Pd(PPh₃)₄ (0.1 eq), Et₃N (5 eq), DMF, rt, dark; (ix) 2-*O*-cyanoethyl-*N,N*-diisopropylamino chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), THF, rt.

Compound **36** was coupled witholesteryl chloroformate in CH₂Cl₂ at rt, with DIPEA as base, to give **37** in 88% yield. Sonogashira coupling reaction of **37** and protected nucleoside **11** afforded **31** in 63% yield. At this stage, the whole synthetic scheme was repeated on an 11 g scale. The coupling reaction between the carboxylic acid **27** and prop-2-yn-1-amine was carried out using PyBrOP as coupling reagent, affording **33** in 53% yield (Scheme 2.22).



Scheme 2.22: *Reagents and conditions* (i) cholesteryl chloroformate (1.7 eq), DIPEA (1.2 eq), CH_2Cl_2 , 2 h at 0°C then 19 h at rt, 88%; (ii) **37** (1.5 eq), CuI (0.4 eq), $\text{Pd}(\text{PPh}_3)_4$ (0.1 eq), Et_3N (7 eq), DMF, rt, dark, 3.5 h, 63%; (iii) prop-2-yn-1-amine (1.5 eq), PyBrOP (1.1 eq), DIPEA (3 eq), CH_2Cl_2 , 1h 15 min, rt, 53%.

2.5.1.6 Phosphitylation

Compound **31** was then subjected to conventional phosphitylation conditions by addition of 2-*O*-cyanoethyl-*N,N*-diisopropylamino chlorophosphoramidite and DIPEA in CH_2Cl_2 under an argon atmosphere. However, ^{31}P NMR analysis in d_6 -DMSO showed a chemical shift at 15.1 ppm in addition to the desired phosphoramidite diastereoisomeric mixture at 148.8 and 148.3 ppm. This inferred the presence of a different phosphorous species and therefore the conditions for this reaction were investigated.

The first approach to resolve this problem was to change the conditions of column chromatography. The new method employed CH_2Cl_2 with an acetone gradient of 25-45%, instead of CH_2Cl_2 with a MeOH gradient of 1-2%. However, after purification, ^{31}P -NMR didn't show the presence of any phosphoramidite. Another attempt was carried out using the same reaction conditions and ^{31}P NMR of the crude product again showed a chemical shift at 15.1 ppm, in addition to the phosphoramidite

diastereoisomeric mixture. TLC showed the degradation of the phosphoramidite into two more polar products.

Thus, to investigate this fully, the phosphitylation was conducted on 5'-*O*-(4,4-dimethoxytrityl)-thymidine **11**. The addition of an excess of phosphitylating reagent again showed the degradation of the phosphoramidite into two more polar products, one with a similar R_f as the starting material.

It was considered that 2-*O*-cyanoethyl-*N,N*,diisopropylamino chlorophosphoramidite was too reactive for the phosphitylation of **31**. Therefore 2-*O*-cyanoethyl-*N,N,N,N*'-tetraisopropylphosphoramidite was used as phosphitylating reagent, in the presence of the activator diisopropylamine hydrotetrazolide (DIHT). Column chromatography in hexane with an EtOAc gradient of 80-100% allowed the separation of the two compounds which were visualised on TLC. NMR and MS analysis confirmed the isolation of the pure phosphoramidite in a 66% yield, but the nature of the other product was not identified. It is likely to be a mixture of starting material, phosphonothymidine and a phosphoramidite at position N³, all of which have the same R_f .

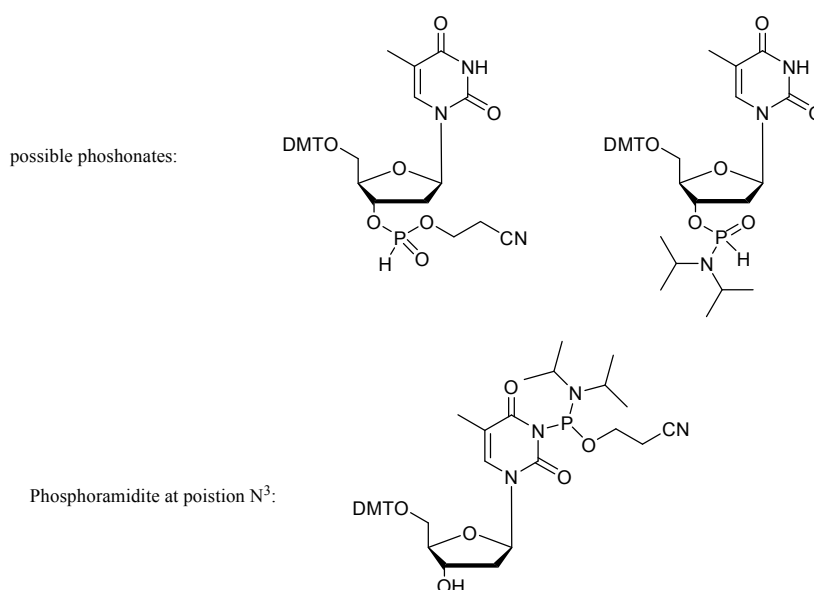
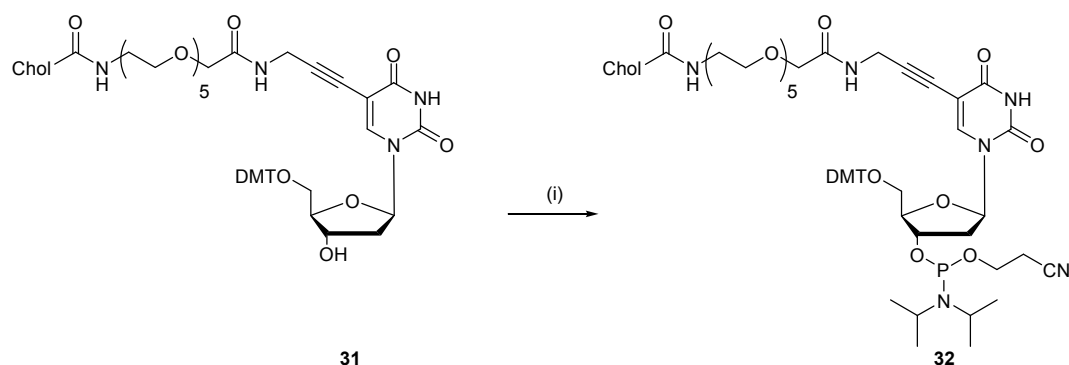


Figure 2.5: Possibles phosphonates and phosphoramidites as by-products

Finally, **31** was successfully converted to its 2-*O*-cyanoethyl-*N,N*-diisopropyl phosphoramidite **32** in 66% yield, using these same conditions (Scheme 2.23).



Scheme 2.23: *Reagents and conditions* (i) 2-*O*-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphoramidite (1.1 eq), DIHT (0.5 eq), CH₂Cl₂, 2h 20 min, rt, 66%.

2.6 Conclusions

Four cholesterol monomers (Bz^{Chol}, T^{Chol}, H^{Chol} and monomer **32**) were synthesised for subsequent incorporation into oligonucleotides, in order to assess the assembly of lipid-conjugated oligonucleotides (L-ODNs) in the planar lipid bi-layer. The first two monomers were synthesised using the hydrophobic spacer **3** to link the cholesterol to a benzyl or 2'-deoxyuridine monomer. In both cases, Sonogashira cross-coupling was used to attach the spacer to the monomers. T^{Chol} was incorporated into ODNs and biophysical studies showed a stabilisation of the duplex when cholesterol was present on both strands. The simple monomer H^{Chol}, synthesised using only cholesteryl chloroformate and 6-amino-hexanol, was incorporated into ODN-7 with success.

For better anchorage of the network to a lipid surface, a hydrophilic spacer was used to link the cholesterol to the 2'-deoxyuridine nucleoside. Problems were encountered with the amide coupling and the hydrazinolysis (see sections 2.5.1.4 and 2.5.1.5) during the synthesis of monomer **32** and several amendments in the order of the synthesis were required. Study and modification of the method of phosphitylation was also necessary, namely a modification of the phosphitylating reagent

2.7 Future work

Problems with the synthesis of phosphoramidate monomer **32** resulted in a number of alterations being made to the original procedure. Now that a reliable synthesis of the target monomer has been found, the next step is to incorporate the monomer into a series of appropriately designed oligonucleotides. Biophysical studies would be carried out in order to analyse the monomer's influence on duplexes and compare it with the hydrophobic linked nucleoside-based cholesterol-dT phosphoramidite. It would then be necessary to study incorporation into the lipid bi-layer and to assess the stability given by this ligation to the hexagonal array.

The AMNA project was halted along with the collaboration with the other European laboratories. In the future, a positive development for this work would be to finalise the attachment of single hexagons and then the array of the hexagons to the lipid surface.

CHAPTER 3

Synthesis of DNA nanostructures: use of click chemistry for the closure of catenanes

3. Synthesis of DNA nanostructures: use of click chemistry for the closure of catenanes

3.1. Introduction

Click chemistry has been applied to DNA systems in a variety of applications as described in Section 1.4.4.3. These include coupling ODNs to self-assembled monolayers,¹²⁷ labelling ODNs,^{125,131,132,151,152} inter-strand cross linking¹³³ and intra-strand cyclisation reaction and catenation of DNA duplexes¹³⁴ and to produce very stable cyclic DNA miniduplexes.¹³⁵

Cyclic duplexes have potential applications in biology. They could be for instance used for drug delivery as they are resistant to DNase enzymic degradation. Because intracellular enzymatic degradation occurs principally from the 3' end, cyclic duplexes, which by definition have no ends, are therefore resistant to exonucleases.^{152,153} Kumar *et al.* have exploited Copper(I)-catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC) for the chemical ligation of two oligonucleotide strands, with or without template, and also for the circularisation of oligonucleotides, for use in the synthesis of pseudohexagonal catenanes.¹³⁴ The catenanes could be used as DNA building blocks for the construction of nanoarrays as their covalent linkers offer more stability than that of simple linear oligonucleotide nanostructures.

CuAAC is very fast and clean, tolerant of other functional groups and functions efficiently in aqueous media. The cyclisation of an oligonucleotide occurs by reaction between its terminal azide and its opposite terminal alkyne. It is catalysed by Cu[I] to produce a 1,2,3-triazole linkage between the two reactants.⁸⁶

The aim of the work described in this chapter was to use the CuAAC, the best example of click chemistry,⁷⁸ to make a double stranded catenane duplex from a single stranded cyclic template and its linear complement.

The triazole is linked to the single stranded cyclic ODN by multiple hexa(ethylene glycol) units (HEG) or multiple thymidines (T) in order to let the complement strand thread through the centre of the circularised ODN and form the required duplex before

cyclising itself (Figure 3.1).

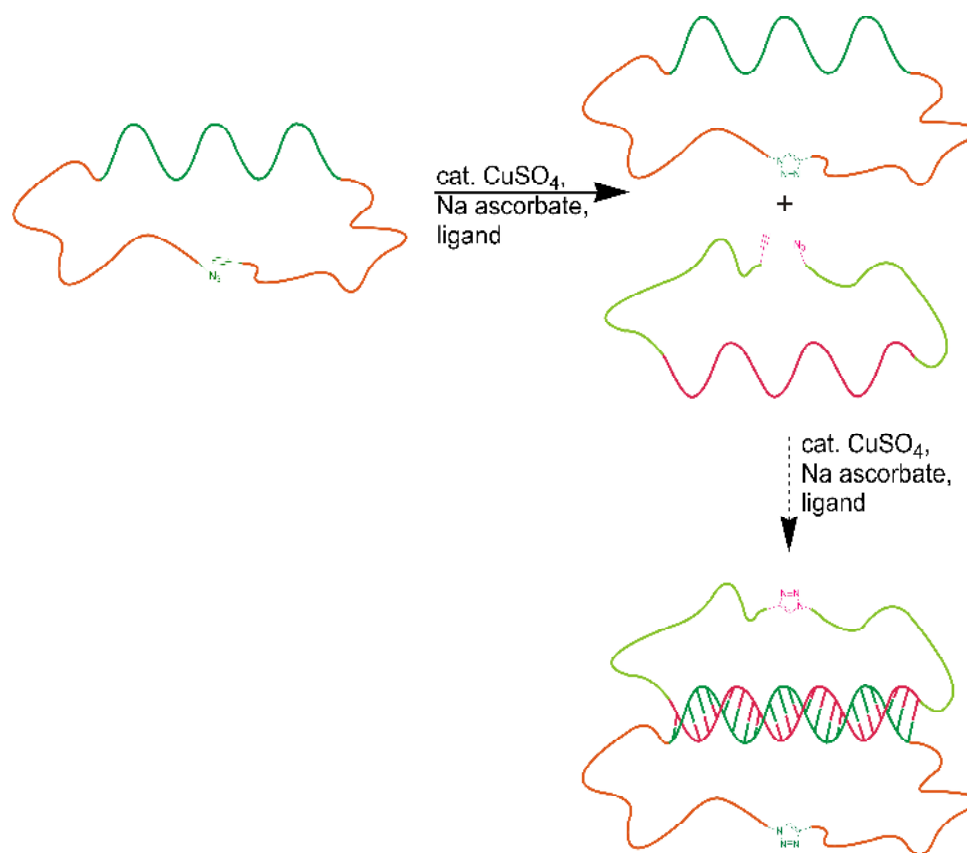


Figure 3.1: Formation of double-stranded catenane from single-stranded cyclic template and its linear complement strand. In this representation the DNA is shown in dark green and pink, whilst multiple T bases and multiple HEG units are in orange and yellow/green respectively.

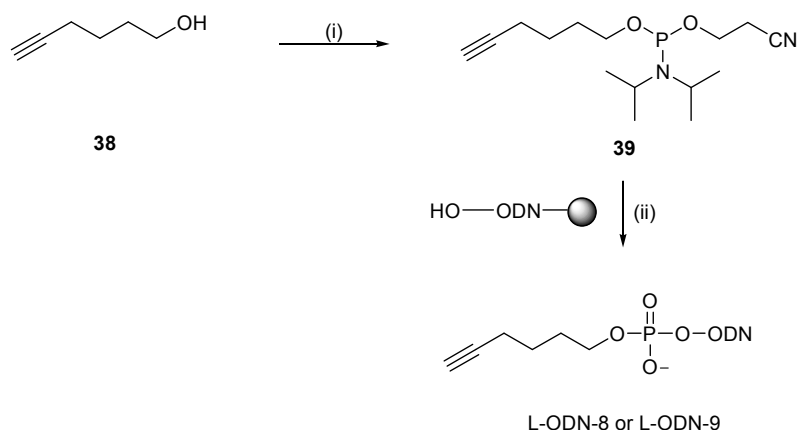
3.2 Synthesis of oligonucleotides for click chemistry

3.2.1 Synthesis of alkyne oligonucleotides

5-Hexyn-1-ol was subjected to conventional phosphitylation conditions by addition of 2-*O*-cyanoethyl-*N,N*,diisopropylamino chlorophosphoramidite and DIPEA in CH₂Cl₂ under an argon atmosphere. The pure phosphoramidite **42** was isolated in 20% yield.

This monomer **42** was then incorporated into ODNs at the 5'-terminus by standard oligodeoxynucleotide synthesis method^{14,154} as described in section 4.3.1, to give ODN-8 and ODN-9 which were purified by HPLC and characterised by MALDI-TOF MS for

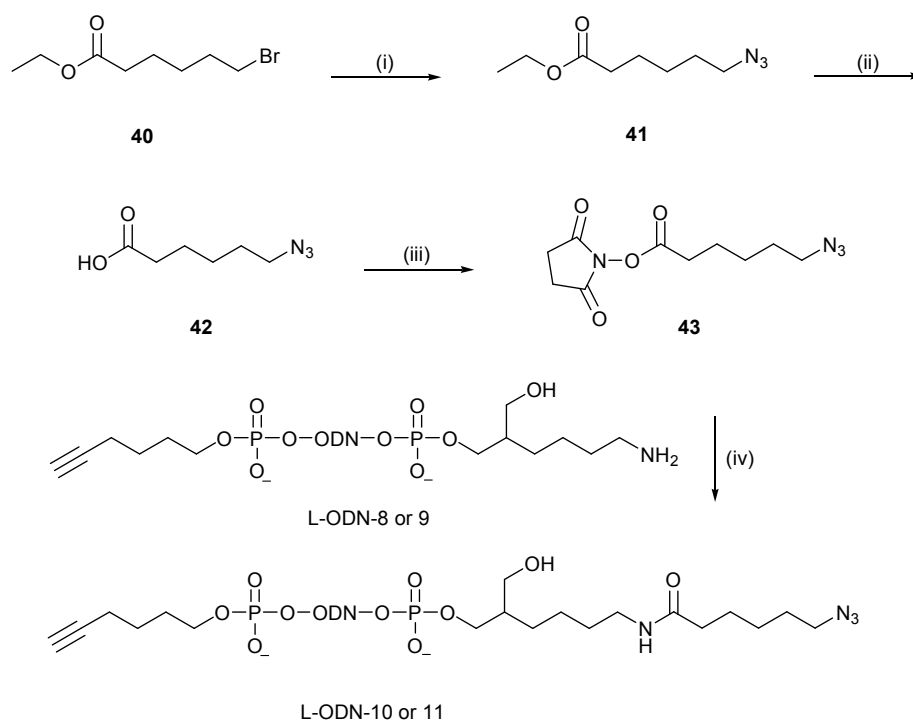
ODN-8 ($[M+H]^+$ calcd, 12876; found, 12884) and ES⁻ MS for ODN-9 ($[M-H]^-$ calcd, 8860; found, 8867) (see section 4.3.3). All ODN sequences are given in Chapter 4 (see section 4.3.2).



Scheme 3.2: *Reagents and conditions* (i) 2-*O*-cyanoethyl-*N,N*-diisopropylamino chlorophosphoramidite (1.1 eq), DIPEA (2.5 eq), CH₂Cl₂, 40 min, rt, 20%; (ii) oligonucleotide synthesis: OH-ODN-resin/tetrazole; Ac₂O/*N*-methylinidazole, I₂/H₂O/THF/Pyridine; NH₃/H₂O.

3.2.2 Synthesis of azide oligonucleotides

Successful azidation of ethyl-6-bromohexanoate **40** was achieved using sodium azide and potassium bicarbonate at 50 °C, in 86% yield. Compound **41** was then hydrolysed using a solution of sodium hydroxide in a mixture of water and dioxane in 91% yield. The following esterification of **42** was carried out with *N*-hydroxysuccinimide and EDC as an activating agent and afforded **43** in a 52% yield (Scheme 3.3).



Scheme 3.3: *Reagents and conditions* (i) NaN₃ (2 eq), K₂CO₃ (1.5 eq), DMF, 5 h 45 min, 50 °C, 86%; (ii) NaOH 2 M, H₂O:dioxane 1:2 v/v, 1 h, rt, 91%; (iii) *N*-hydroxysuccinimide (1.3 eq), EDC (1.4 eq), 6.5 h, rt, 52%; (iv) Na₂CO₃/NaHCO₃ 0.5 M (pH 8.75), DMSO.

Azide L-ODN-10 and 11 were prepared by labelling 3'-amino-modified L-ODN-8 and 9 with **43** in bicarbonate buffer at pH 8.75. The resultant 3'-azide-labelled ODNs were then purified by reversed-phase HPLC (see Appendix) and characterised by MALDI-TOF MS for L-ODN-10 ([M+H]⁺ calcd, 13015; found, 130227) and ES⁻ MS for ODN-11 ([M-H]⁻ calcd, 8999; found, 9006) (see section 4.3.3).

3.3. Click chemistry

3.3.1 First catenane

3.3.1.1 Cyclisation

The Cu[I] complex was generated *in situ* from Cu[II] sulfate and sodium ascorbate. All ligation reactions were carried out in 0.2 M aqueous NaCl in the presence of the water soluble *tris*-(hydroxypropyltriazolylmethyl)amine ligand.⁸⁸ Self-circularisation reactions of L-ODN-10 and its complementary L-ODN-11 gave C-ODN-10 and C-ODN-11 respectively, characterised by MALDI-TOF MS (see section 4.3.3).

ODNs sequences:

L-ODN-10 5'-K-**TTTTTTTTTTGCACCAGAATTCATCACGGAG**TTTTTTTTTTT-Z-3'
 L-ODN-11 3'-Z-**HHHCGTGGTCTTAAGTAGTGCCTC**HHH-K-5'

A quantitative conversion of L-ODN-10 to the cyclic C-ODN-10 was observed by capillary electrophoresis (CE). (Figure 3.2)

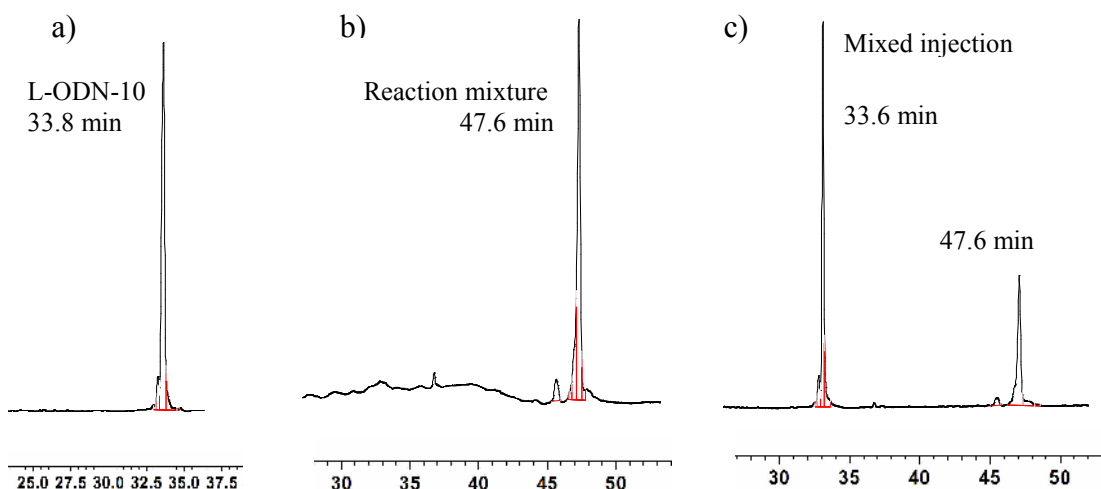


Figure 3.2: Capillary electrophoresis analysis: (a) L-ODN-10; (b) cyclisation reaction mixture of L-ODN-10; (c) mixture of L-ODN-10 and reaction mixture of C-ODN-10.

Similar CE analyses were obtained after cyclisation of L-ODN-11 in C-ODN-11.

3.3.1.2 Catenation of cyclic oligonucleotides with their linear complements

Formation of the double stranded catenane from C-ODN-10 and the L-ODN-11 was carried out in parallel with C-ODN-11 and the L-ODN-10. The purified cyclic ODNs were mixed with their respective linear complementary strands and the click reaction was carried out at 0.2 M aqueous NaCl. The reaction was monitored by denaturing polyacrylamide gel electrophoresis (PAGE). (Figure 3.3)

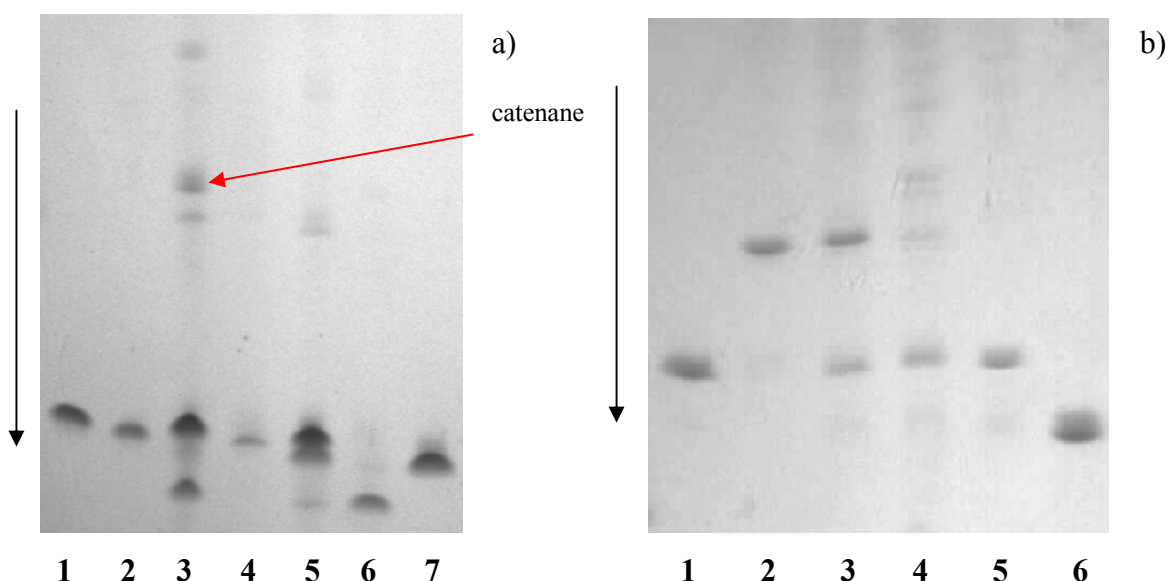


Figure 3.3: Gels visualised by UV shadowing (254 nm) (a) Denaturing 8% PAGE analysis: **1**-L-ODN-10, **2**-C-ODN-10, **3**-crude mixture of the catenation of L-ODN-10 and C-ODN-11, **4**-crude mixture of the circularisation of L-ODN-10 into C-ODN-10, **5**-crude mixture of the catenation of L-ODN-11 and C-ODN-10, **6**-C-ODN-11, **7**-L-ODN-11. (b) Denaturing 20% PAGE analysis: **1**-L-ODN-10, **2**-C-ODN-10, **3**-crude mixture of the catenation of L-ODN-10 and C-ODN-11, **4**-crude mixture of the catenation of L-ODN-11 and C-ODN-10, **5**-C-ODN-11, **6**-L-ODN-11.

On the 8% PAGE analysis, L-ODN-10 and C-ODN-10 run at the same velocity but C-ODN-11 runs faster than L-ODN-11. On the 20% PAGE analysis, C-ODN-11 runs slower than L-ODN-11 and also C-ODN-10 runs slower than L-ODN-10.

On the 8% PAGE analysis, there is a new retarded band, lane **3**, but this band cannot confidently be attributed to a catenane as this band is not present on 20% PAGE analysis. Analysis by capillary electrophoresis (CE) also showed no additional peaks.

The catenation reaction was repeated on a bigger scale with similar results.

UV melting of C-ODN-11 and L-ODN-10 followed by C-ODN-10 and L-ODN-11 using a standard melt programme (20-80-20 °C at 0.5 °C/min), showed only a small decrease in thermal stability ($\Delta T_m = -2$ or -3 °C) compared with the duplex of the two linear oligonucleotides L-ODN-10 and L-ODN-11 (Table 3.1). This suggests that the failure of the catenane formation was not due to thermal instability of the circularised ODN.

The attempted formation of the catenane duplex at 1 M NaCl was also unsuccessful.

buffer with NaCl (0.2 M)				
ODNs	L-ODN-10/L-ODN-11	L-ODN-10/C-ODN-11	C-ODN-10/L-ODN-11	C-ODN-10/C-ODN-11
T_m (°C)	63.4	61.9	61.3	69.0
0.5 °C intervals				
buffer with NaCl (1 M)				
ODNs	L-ODN-10/L-ODN-11	L-ODN-10/C-ODN-11	C-ODN-10/L-ODN-11	C-ODN-10/C-ODN-11
T_m (°C)	71.1	68.3	68.1	nd
0.25 °C intervals				

Table 3.1: Duplex (0.25 μ M) UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer containing 0.2 M or 1 M NaCl and 1 mM Na₂EDTA). Each T_m is an average of three separate runs. nd = not determined.

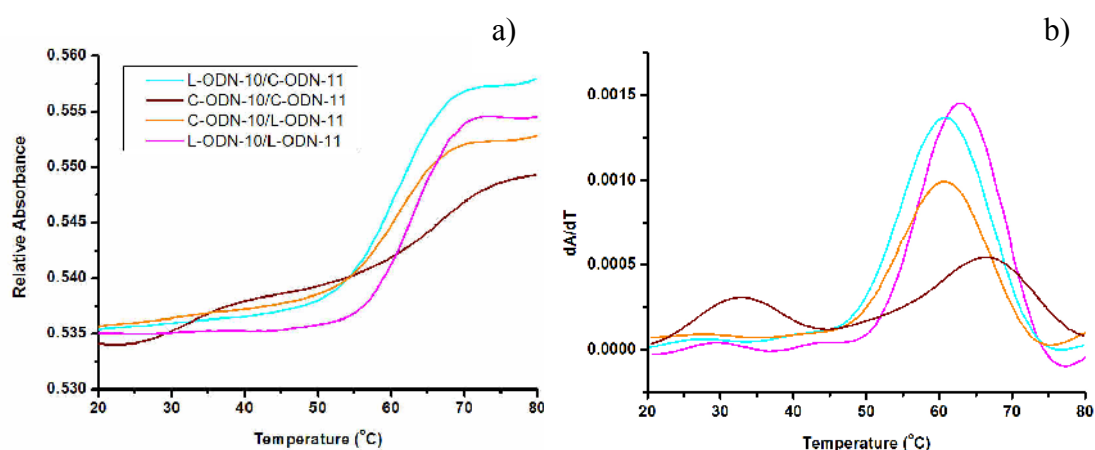


Figure 3.4: (a) Duplex UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer containing 0.2 M NaCl and 1 mM Na₂EDTA). The concentration of each ODN was 0.25 μ M. (b) derivative d(A)/d(T). Melting data were processed using the OriginPro 7.5 software.

The ODNs have 21 matching base pairs, constituting two turns of a double helix. This may have caused the ODN to be too rigid to ligate to its complementary ODN (Figure 3.5).

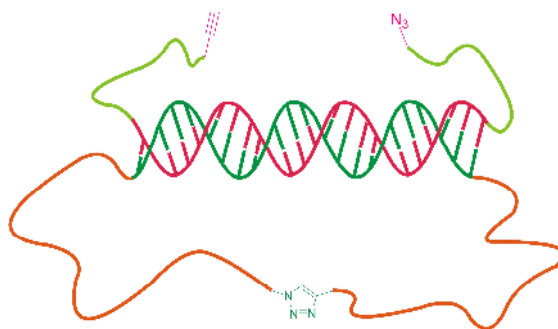


Figure 3.5: Non formation of double-stranded catenane from single-stranded cyclic template and its linear complement strand as the alkyne and azide of the second ODN are too far from each other. In this representation the DNA is shown in dark green and pink, whilst multiple T bases and multiple HEG units are in orange and yellow/green respectively.

For this reason, two new ODNs were synthesised with only 12 matching base pairs (just over one helix turn). A fluorescein label was incorporated into the linear ODN, in order to improve the visualisation of the products by gel electrophoresis, using transillumination in addition to UV-shadowing. A third short ODN, complementary to the extremities of this second ODN to cyclise, was also used to help the second circularisation (Figure 3.6).

ODNs sequences:

L-ODN-14 5'-K-TTTTTTTTTTTTTTTT**CCAGAATTCATC**TTTTTTTTTTTTTTTTT-Z-3'
 L*-ODN-15 3'-Z-GCGTHHH**GGTCTTAAGTAG**HHHT*GCG-K-5'

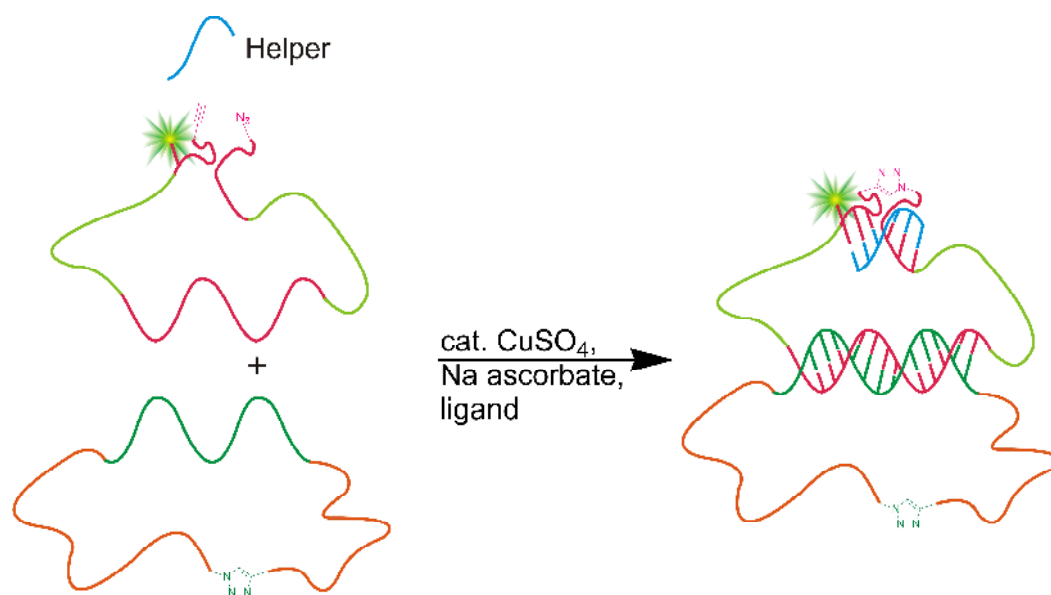


Figure 3.6: Formation of double-stranded catenane from single-stranded cyclic template and its linear complement strand with a short ODN strand as a helper. In this representation the DNA is shown in dark green and pink and blue (for the helper), multiple T bases are in orange, multiple HEG units are in yellow/green and the fluorescein is a yellow/green star.

3.3.2 Ligation of shorter catenane sequence

L-ODN-14 and L*-ODN-15 were prepared using standard automated DNA synthesis as described in section 4.3.1 and labelled with 3'-azide and 5'-alkyne as described in section 3.2.2. Self-circularisation of L-ODN-14 and L*-ODN-15 gave C-ODN-14 and C*-ODN-15 respectively.

ODNs sequences:

L-ODN-14 5'-K-**TTTTTTTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTTTTT**-Z-3'
 L*-ODN-15 3'-Z-GCGTHHH**GGTCTTAAGTAG**HHHT*GCG-K-5'

The catenane was prepared by mixing the C-ODN-14, the L*-ODN-15, and the helper ODN-16. The click reaction was carried out at 200 mM aqueous NaCl. A new retarded band appeared on the denaturing 20% PAGE gel on lane **3** (Figure 3.7), which was attributed to the formation of the covalently linked double stranded catenane. Lane **3** showed also no more L*-ODN-15 and C-ODN-14 but a band corresponding to C*-ODN-15, which means that some of L*-ODN-15 cyclised during the catenation step.

However, the crude mixture of the circularisation of L*-ODN-15, lane 4, showed also a faint retarded band at the same retention factor which could be attributed to a dimer of L*-ODN-15 on 20% PAGE (Figure 3.7 b).

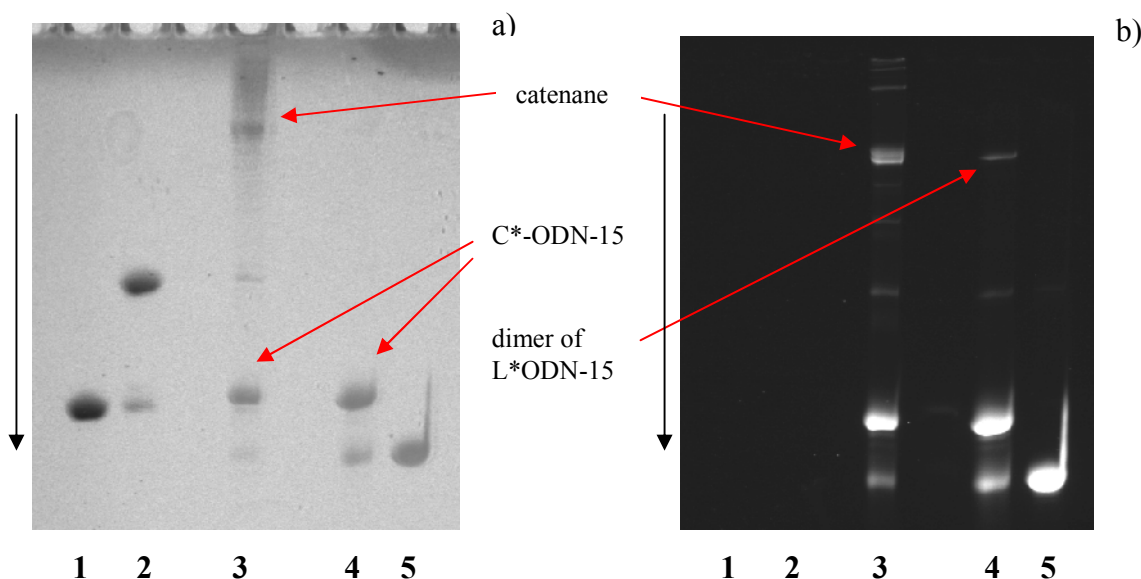


Figure 3.7: Denaturing 20% PAGE analysis: 1-L-ODN-14, 2-C-ODN-14 not purified by HPLC, 3-crude mixture of the catenation of C-ODN-14 and L*-ODN-15, 4-C*-ODN-15 not purified by HPLC, 5-L*-ODN-15. (a): gel visualised by UV shadowing (254 nm) (b) gel trans-illuminated at 302 nm.

The reaction conditions of the self-circularisation were investigated in an attempt to improve the yield. Firstly, a 20-mer template of poly A (A_{20}) was added to the reaction mixture during the ligation of L-ODN-14. It was hoped that it would form a duplex with the multiple T bases at both extremities of L-ODN-14 and therefore help its closure, in the same way as the helper in Figure 3.5. However, this did not improve the yield as shown in Figure 3.8.

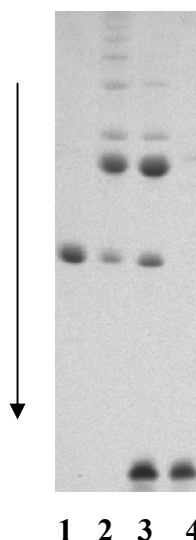


Figure 3.8: Denaturing 20% PAGE analysis visualised by UV shadowing (254 nm): **1**-linear L-ODN-14, **2**-circularisation of L-ODN-14 into C-ODN-14, **3**-circularisation of L-ODN-14 into cyclic C-ODN-14 with A₂₀. (1 eq), **4**- A₂₀.

Following this, it was established experimentally that when the concentration of the ODN was low (0.8 μ M) a concentration of 8 μ M (10 equivalents) of copper was sufficient to ensure cyclisation. However at high ODN concentration (8 μ M), a very high concentration of copper (400 μ M, 50 equivalents) was necessary to obtain a quantitative conversion of linear oligonucleotides into cyclic oligonucleotides (Figure 3.9). A possible explanation of this concentration effect is that L-ODN-14 could form a 6 base pairs dimer at the EcoR1 site, at high ODN concentration, that would prevent cyclisation.

ODNs sequences:

L-ODN-14 5'-K-TTTTTTTTTTTTTTTTCCAGAAATTCATCTTTTTTTTTTTTTTTT-Z-3'
L-ODN-14 3'-Z-TTTTTTTTTTTTTTTTCTACTTAAGACCTTTTTTTTTTTTTTTT-K-5'

Catenation was also carried out in the presence of an increased amount of copper (50 equivalents), but again no improvement of the yield was seen by PAGE analysis.

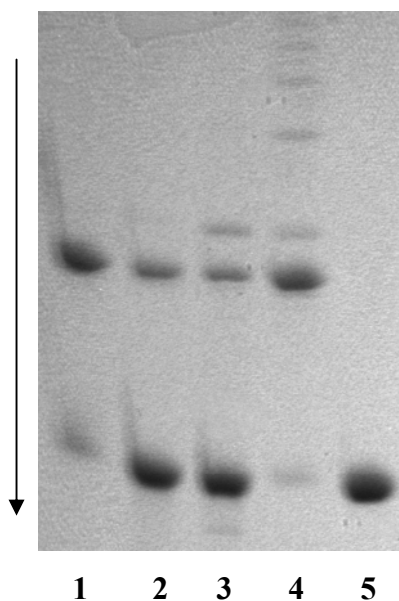


Figure 3.9: Denaturing 20% PAGE analysis visualised by UV shadowing (254 nm): **1**-circularisation of L-ODN-14 (8 μ mol) into C-ODN-14 with copper (400 μ mol), **2**-circularisation of L-ODN-14 (8 μ mol) into C-ODN-14 with copper (80 μ mol), **3**-circularisation of L-ODN-14 (8 μ mol) into C-ODN-14 with copper (80 μ mol), **4**-circularisation of L-ODN-14 (0.8 μ mol) into C-ODN-14 with copper (8 μ mol), **5**- L-ODN-14.

Finally, Ferguson plot analysis^{70,136} on linear and cyclic oligonucleotides (log mobility vs gel concentration) showed an increasing gel retardation coefficient with cyclisation of the ODN, while the catenane had a much higher retardation coefficient in comparison to the single stranded ODNs. This therefore made it possible to distinguish between the catenane and the dimerised ODN-17, confirming the success of the catenation reaction (Figure 3.10).

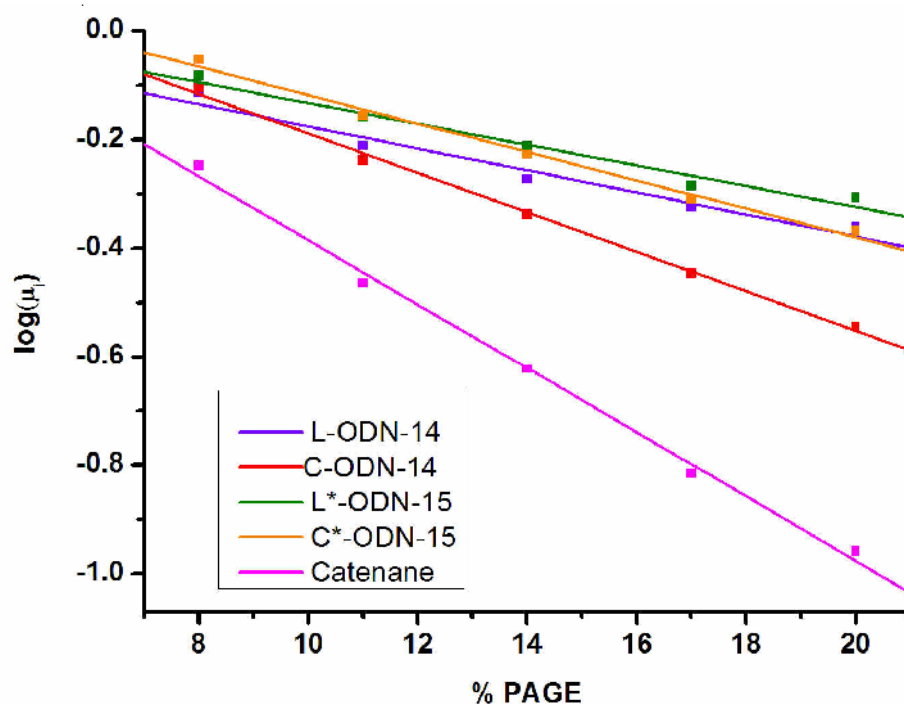


Figure 3.10: Ferguson plot analysis of L-ODN-14, C-ODN-14, L*-ODN-15, C*-ODN-15 and the catenane from L*-ODN-15 and C-ODN-14 in polyacrylamide gels. The logarithm of the absolute mobility is plotted against polyacrylamide gel concentration. Melting data were processed using the OriginPro 7.5 software.

In order to disprove unambiguously that the retarded band was not due to a dimer of the linear ODN, the strategy was modified so that the first ODN that is cyclised contained a fluorescein. Therefore new sequences were designed for the catenation: L*-ODN-19 was used instead of L-ODN-14 and L-ODN-20 instead of L*-ODN-15

ODNs sequences:

L*-ODN-19 5'-K-TTT*TTTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTTTTT-Z-3'
 L-ODN-20 3'-Z-GCGTHHHGGTCTTAAGTAGHHHTGCG-K-5'

3.3.3 Alternative sequence of the catenane

3.3.3.1 Catenation

L*-ODN-19 and L-ODN-20 were prepared using standard automated DNA synthesis as described in section 4.3.1 and labelled with 3'-azide and 5'-alkyne as described in

section 3.2. Self-circularisation reaction of L*-ODN-19 and L-ODN-20 were then carried out to give C*-ODN-19 and C-ODN-20 respectively.

The catenane was prepared by mixing the cyclic C*-ODN-19, the linear L-ODN-20, and the helper ODN-16. The click reaction was carried out at 0.2 M aqueous NaCl. Figure 3.11 shows a clear retarded band which was attributed to the formation of the covalently closed double stranded catenane. According to the intensities on the gel, the yield of the reaction appeared to be around 50%.

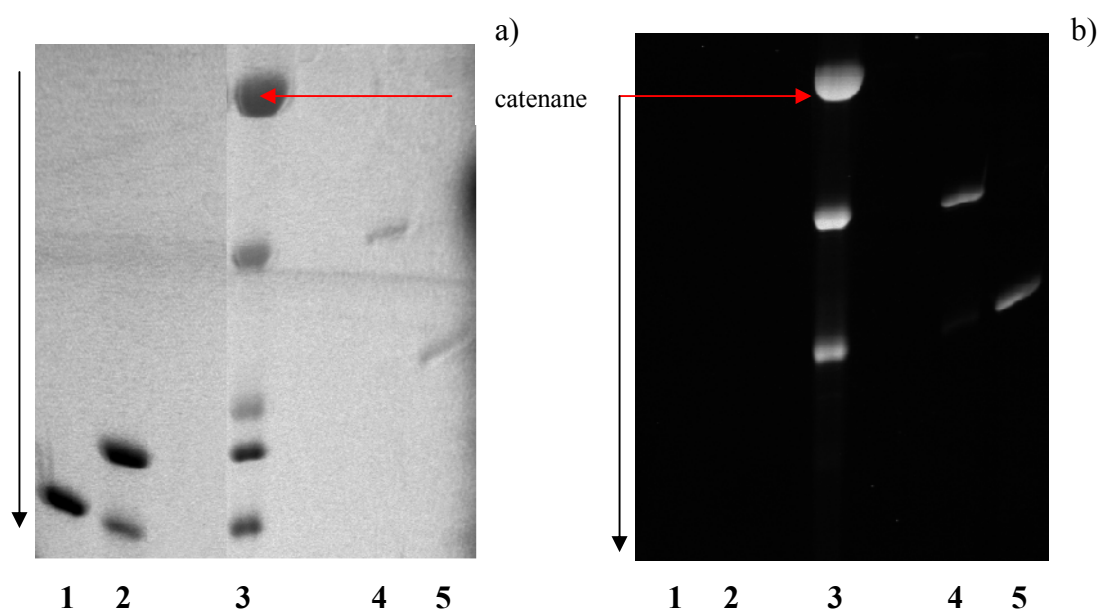


Figure 3.11: Denaturing 20% PAGE analysis: **1**-L-ODN-20, **2**-C-ODN-20 non purified, **3**-crude mixture of the catenation of cyclic C*-ODN-19 and L-ODN-20, **4**-C*-ODN-19 not purified, **5**-L*-ODN-19. (a): gel visualised by UV shadowing (254 nm) (b) gel trans-illuminated at 302 nm.

Purification of the catenane by reversed-phase HPLC and anion-exchange HPLC was attempted but failed because the catenane has a similar retention times than C*-ODN-19 and C-ODN-20. However, the catenane was successfully purified by gel-electrophoresis extraction with an overall yield of 25% (catenation and purification). The catenation reaction and purification were repeated on a bigger scale and gave similar results.

3.3.3.2 Enzyme digestion

L*-ODN-19 and L-ODN-20 were designed so that their duplex would contain an EcoR1 restriction site (Figure 3.12 a). As EcoR1 only recognises and cuts double strands, the formation of the catenane can then be proved.

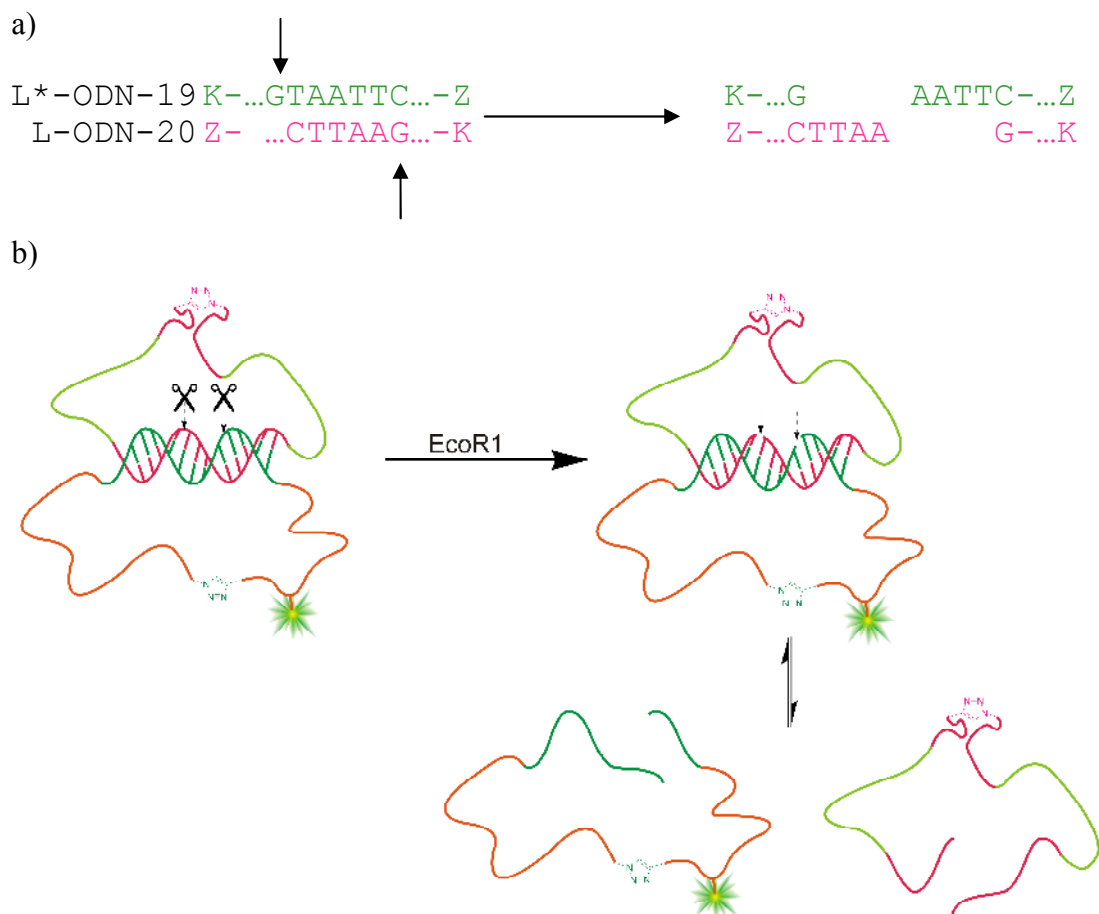


Figure 3.12: Digestion of the catenane by EcoR1. a) EcoR1 restriction site on the ODNs. K = alkyne, Z = azide. b) In this representation the DNA is shown in dark green and pink, multiple T bases are in orange, multiple HEG units are in yellow/green and the fluorescein is a yellow/green star.

Figure 3.13 shows a gel analysis of the digestion of the oligonucleotides with and without the enzyme (controls). The catenane was cut by EcoR1 and shows the formation of the digested cyclic C*-ODN-19 and the digested cyclic C-ODN-20, in lane 7, proving the existence of the catenane. There is a band for the catenane in lane 9 and 10 on the gel b) stained with SYBR Green I, which is due to the horizontal migration of the catenane.

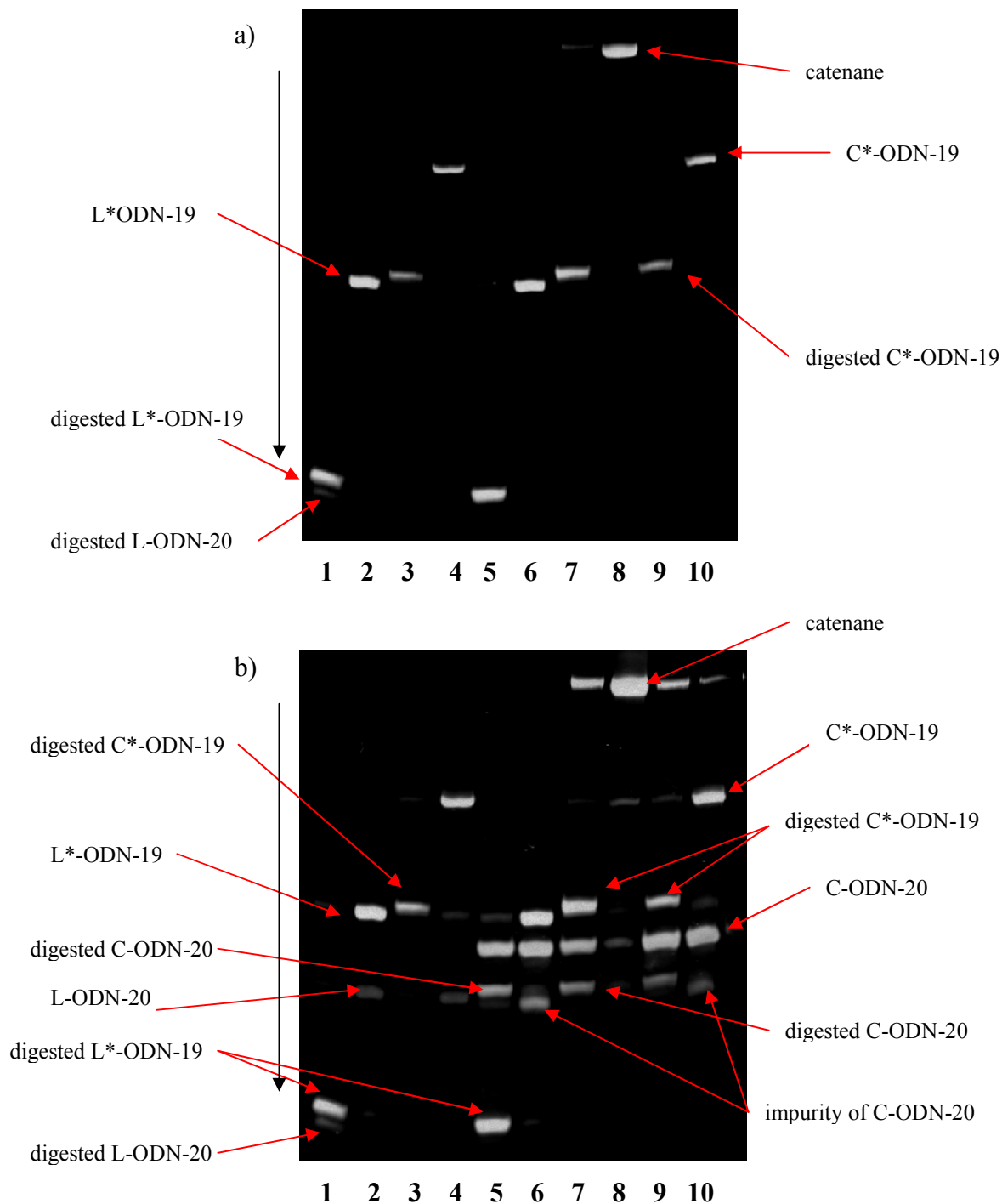


Figure 3.13: Denaturing 20% PAGE analysis of EcoR1 ODN digestion: Even lanes are controls: DNA-only without EcoR1. **1**-L*-ODN-19 and L-ODN-20 + EcoR1; **2**-L*-ODN-19 and L-ODN-20, **3**-C*-ODN-19 and L-ODN-20 + EcoR1; **4**-C*-ODN-19 and L-ODN-20, **5**-L*-ODN-19 and C-ODN-20 + EcoR1; **6**-L*-ODN-19 and C-ODN-20, **7**-catenane (from C*-ODN-19 and L-ODN-20) + EcoR1; **8**-catenane (from C*-ODN-19 and L-ODN-20), **9**-C*-ODN-19 and C-ODN-20 + EcoR1; **10**- C*-ODN-19 and C-ODN-20. (a) gel trans-illuminated at 302 nm (b) gel trans-illuminated at 302 nm after 30 min staining with SYBR Green I

The results of the EcoR1 digestions suggest a few points:

- C-ODN-20 is not digested as well as C*-ODN-19. This may be due to the presence of HEG in the oligonucleotide, which increases the oligonucleotide stability towards the enzyme.
- The impurity of C-ODN-20 is totally digested by EcoR1, implying it is in a double stranded conformation.

3.3.3.3 Native gel

The catenane was analysed on a native 14% PAGE gel, giving same retention factors to those obtained by denaturing PAGE gel analysis. (Figure 3.14) This result invalidates the existence of a secondary structure for the catenane. On Figure 3.14 a), the catenane appeared pure, but some impurities appeared on Figure 3.14 b) because fluorescence is a very sensitive technique.

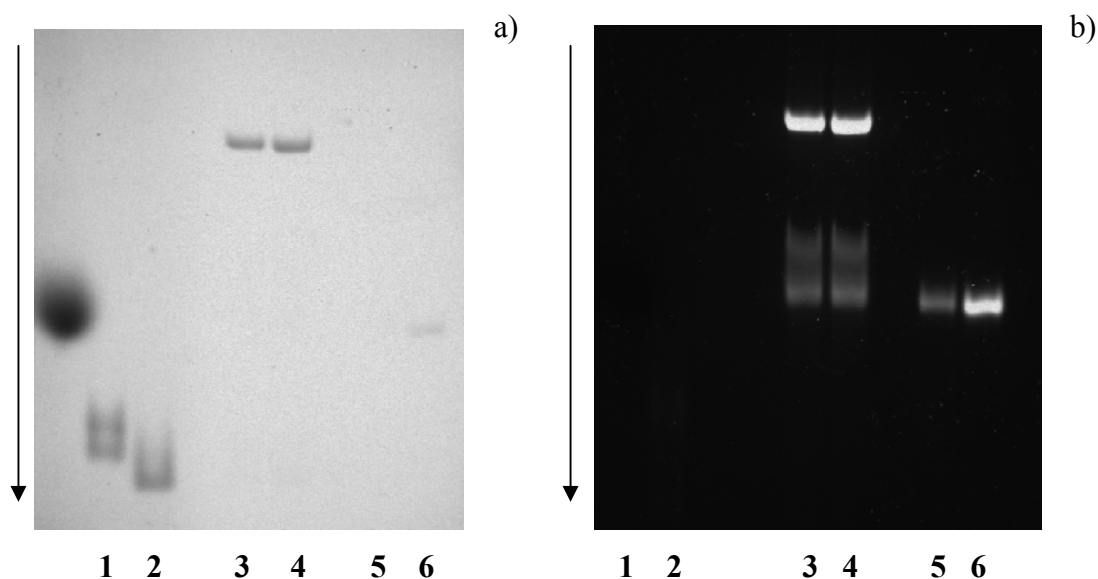


Figure 3.14: Native 14% PAGE analysis: 1-L-ODN-20, 2-C-ODN-20 purified by gel, 3&4-catenane purified by gel, 5-C*-ODN-19 purified by gel, 6-L*-ODN-19. (a): gel visualised by UV shadowing (254 nm) (b) gel trans-illuminated at 302 nm.

3.3.3.4 Thermodynamic studies of the catenane

Using a standard melt programme (20-80-20 °C at 0.5 °C/min with a hold time between melting and annealing), UV melting of the catenane showed no duplex-random coil transition in the range of UV melting temperatures, in contrast with the duplexes of L-ODN-20/L*-ODN-19, L-ODN-20/C*-ODN-19 and C-ODN-20/L*-ODN-19.

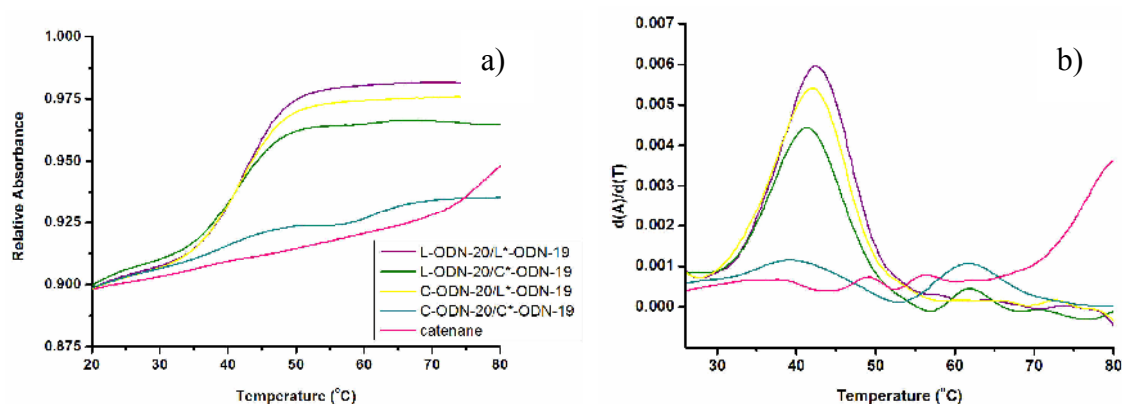


Figure 3.15: (a) Duplex UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer containing 200 mM NaCl and 1 mM Na₂EDTA). The concentration of each ODN was 2 μM. (b) derivative d(A)/d(T). Melting data were processed using the OriginPro 7.5 software.

A control melting experiment with single strands also showed no duplex formation. Therefore, formamide was added to the buffer to reduce the T_m of the catenane, enabling observation of its melting transition.

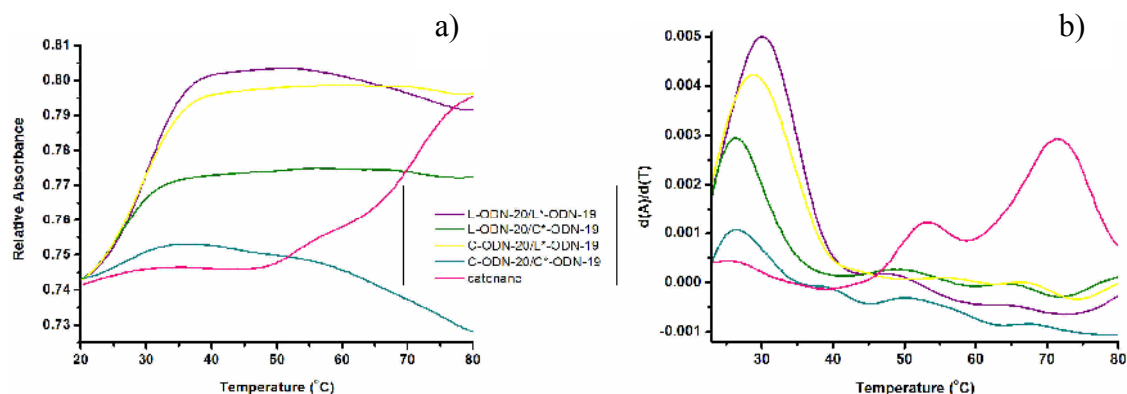


Figure 3.16: (a) Duplex UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer containing 200 mM NaCl and 1 mM Na₂EDTA with 20% formamide). The concentration of each ODN was 2 μ M. (b) derivative d(A)/d(T). Melting data were processed using the OriginPro 7.5 software.

The addition of formamide caused the derivative of the catenane to show another transition with a lower T_m , but corresponding to a very stable duplex. This is unexplained at this stage as no other evidence shows the presence of a secondary structure of the catenane or any relevant impurities. Indeed, both denaturing and native gels showed the presence of only one significant spot for the purified catenane. A hypothetic explanation would be that this transition is due to the formation of a catenane from 2 L-ODN-20 that would form a duplex at the Eco R1 site and both cyclised; although no significant band appeared on PAGE analysis that would confirm this hypothesis.

ODNs sequences:

L-ODN-20 5'-K-GCGTHHHGATGAATTC TGGHHHTGCG-Z-3'
 L-ODN-20 3'-Z-GCGTHHHGGTCTTAAG TAGHHHTGCG-5'

Catenane			
buffer	no formamide	20% formamide	40% formamide
T_m (°C)	nd	71.6	61.4

Table 3.2: Duplex UV melting experiment at pH 7.0 (10 mM sodium phosphate buffer containing 200mM NaCl and 1mM Na₂EDTA without or with 20% or 40% formamide). Each T_m is an average of five separate runs.

The evolution of the melting curves and derivatives by the addition of formamide is shown in Figure 3.17.

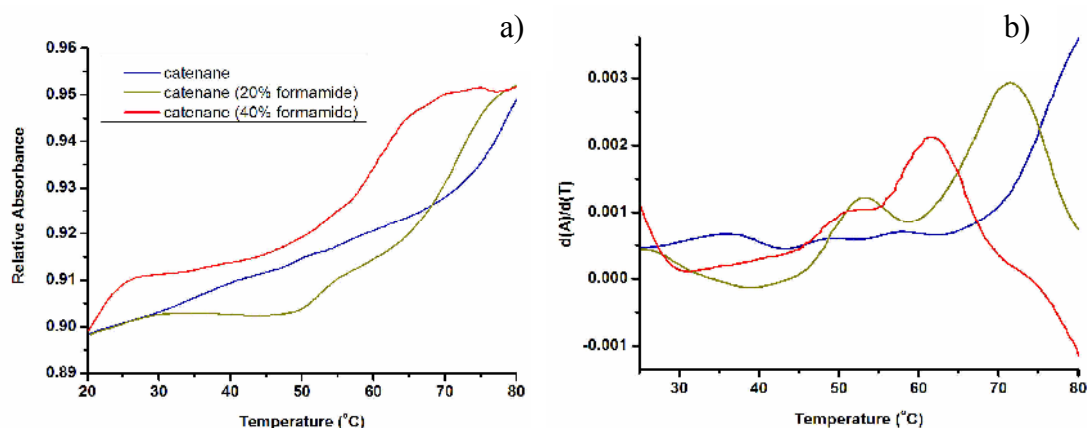


Figure 3.17: Duplex UV melting experiment at pH 7.0 and 2 μ M of duplex in 10 mM sodium phosphate buffer containing 200mM NaCl and 1mM Na_2EDTA without or with 20% or 40% of formamide. Melting data were processed using the OriginPro 7.5 software.

3.4 Conclusions

Azides were incorporated to alkyne-modified oligonucleotides by post-synthetic modification in order to circularise them, using CuAAC. Catenation of two complementary DNA strands was eventually completed after modification of the oligonucleotides sequences. The ODNs were designed with only twelve matching base pairs to make just over one helix turn and a fluorescein was incorporated in one of them to facilitate visualisation. A third short ODN was also used as a helper for the circularisation of the second ODN.

These results clearly show that the design of the ODNs' sequences is crucial for the realisation of the catenation and that a maximum of one helix turn is required for the success of the second closure, but it might be possible to create a catenane of less than one helix turn. The hypothetical formation of catenane made by the two L-ODN-20 could be prevented by omitting the EcoR1 site that is avoiding self-complementarity.

3.5 Future work

The resulting double stranded catenane duplex is very stable since the 1,2,3-triazole unit has a high chemical stability (in general, being inert to severe hydrolytic, oxidising, and reducing conditions even at high temperature).⁸⁶ It is therefore a suitable candidate as a DNA building block for the construction of various nanostructures.

Owing to the interesting biological activity of 1,2,3-triazoles and the nature of the ODNs, the catenanes have vast potential biological uses. It would be interesting to use the catenane for drug delivery by designing the sequences in order to elicit cleavage of one of the ODNs after cell penetration. A cysteine or a thiol group could be easily incorporated into the ODNs during chemical synthesis for subsequent synthesis of a disulfide bond. The reversible nature of the disulfide bond and its stability in plasma make it an attractive tool for drug delivery. Indeed, the bond is stable in the oxidising extracellular space, but can undergo facile cytoplasmic scission due to reducing agents in the intracellular space.¹⁵⁵ This route could be investigated for later application of the catenane as a delivery tool.

CHAPTER 4

Experimental

4. Experimental

4.1 General experimental

All reagents used were purchased from Aldrich, Fluka, Avocado, Alfa Aesar, Fisher Scientific, Acros or Lancaster and used without further purification, except for the reaction solvents, CH₂Cl₂ and DIPEA, pyridine and Et₃N which were purified by distillation (over calcium hydride), MeOH (over iodine and magnesium) and THF (over sodium wire and benzophenone). All reactions were carried out under an argon atmosphere. Glassware was oven-dried overnight before use, for reactions requiring exclusion of moisture. Column chromatography was carried out under air or argon pressure using Fisher Scientific DAVSIL 60Å (35-70 micron) silica. Thin layer chromatography was performed using Merck Kieselgel 60 F₂₅₄ silica plates (0.22 mm thickness, aluminium backed). Compounds were visualised by irradiation at 254/365 nm or by staining with potassium permanganate (A), PMA (B), *p*-anisaldehyde (C), ninhydrin (D, G) 2,4-DNPH (E), Mary's reagent (F), followed by heating.

A: KMnO₄:K₂CO₃ (3:20 g) / aqueous NaOH (5% w/v):water (5:300 mL)

B: phosphomolybdic acid 10% in ethanol

C: *p*-anisaldehyde / glacial AcOH / conc. H₂SO₄ / ethanol 95% (9.2:3.8:12.5:338 v/v)

D: ninhydrin (0.3% w/v) / butanol: glacial AcOH (97:3 v/v)

E: 2,4-dinitrophenyl hydrazine/ethanol/H₂SO₄

F: 4,4'-bis-dimethylamino benzhydrol/acetone

G: ninhydrin (5 g) / acetone (100 mL)

¹H NMR spectra were measured at 300 MHz on a Bruker AC300 spectrometer or 400 MHz on a Bruker DPX400 spectrometer. ¹³C NMR spectra were measured at 100 MHz on a Bruker DPX400 spectrometer. Chemical shifts are given in ppm relative to tetramethylsilane, and *J* values are given in Hz and are correct to within 0.5 Hz. All ¹H NMR spectra are internally referenced to the appropriate residual undeuterated solvent signal¹⁵⁶ and the ¹³C NMR spectra to the appropriate residual deuterated solvent signal. Multiplicities of ¹³C signals were determined using DEPT spectral editing technique and

are described as s (quaternary carbon), d (CH), t (CH₂) and q (CH₃). Assignment was also aided by COSY (¹H-¹H) and HMQC experiments. ³¹P NMR spectra were recorded on a Bruker AC300 spectrometer at 121 MHz and were externally referenced to 85% phosphoric acid in D₂O.

Low-resolution mass spectra were recorded using electrospray ionisation (ESI) on a Fisons VG platform instrument or a Waters ZMD quadrupole mass spectrometer in HPLC grade acetonitrile or methanol. High-resolution mass spectra were recorded using electrospray ionisation on a Bruker APEX III FT-ICR mass spectrometer in HPLC grade acetonitrile or methanol.

Mass spectra of oligonucleotides were recorded using electrospray ionisation (ESI) on a Fisons VG platform instrument in HPLC grade water, with tripropylamine to aid ionisation; or using MALDI-TOF MS analysis on a ThermoBioAnalysis Dynamo MALDI-TOF mass spectrometer in positive ion mode using internal T_n standards.¹⁴⁰

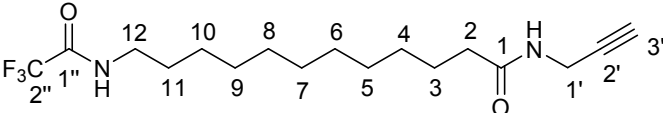
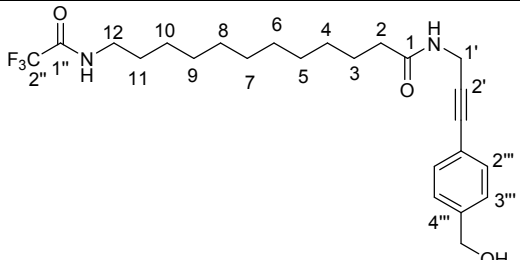
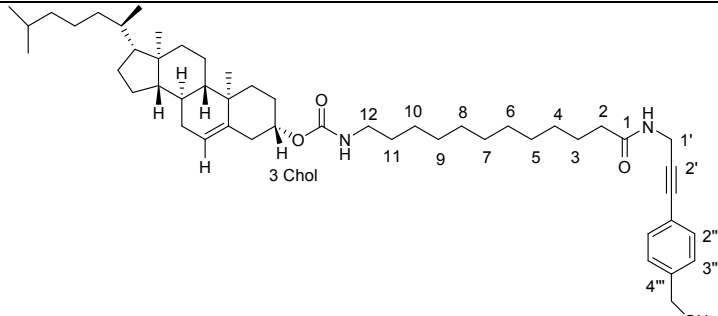
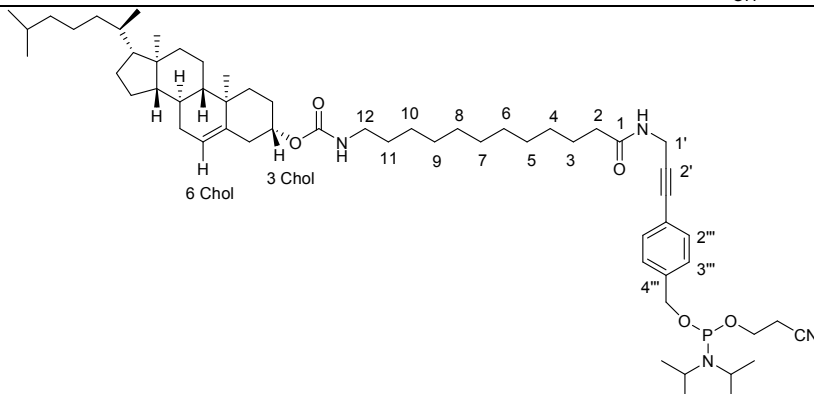
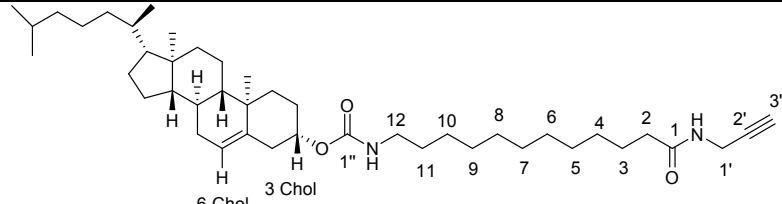
IR spectra were recorded on a Satellite FT-IR using a Golden Gate adapter and WIN FIRST-lite software or using a Smart Orbit adapter and OMNIC software. Absorptions are described as strong (s), medium (m) and weak (w).

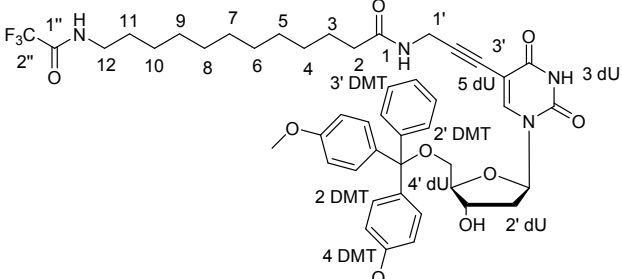
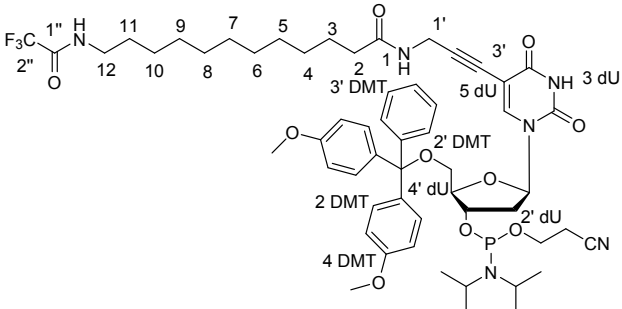
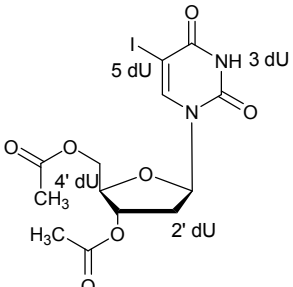
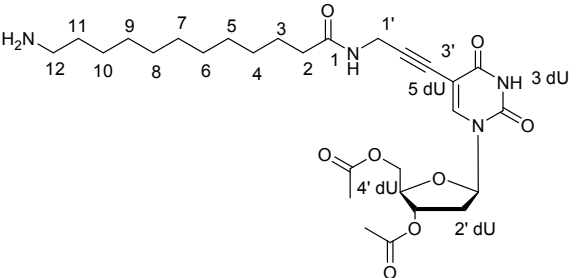
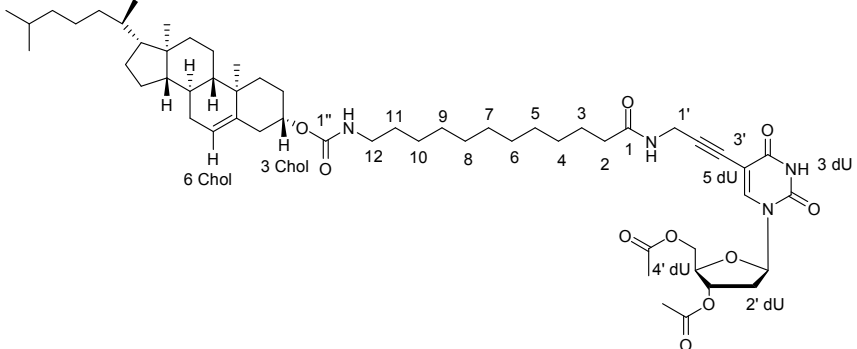
Melting points were measured on a Gallenkamp electrothermal melting point apparatus and are uncorrected.

Elemental (CHN) Thermal Combustion Analysis was carried out by MEDAC Ltd., Egham, Surrey, UK.

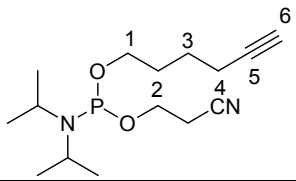
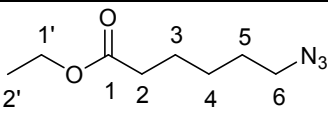
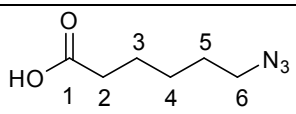
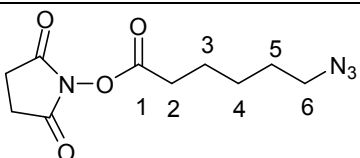
4.2 Synthesis

Compounds synthesised:

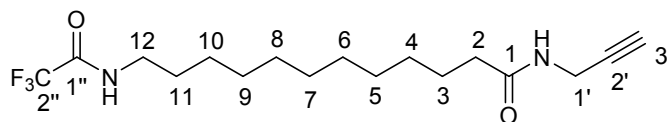
No	Structure	Page
3		100
5		101
7		102
8		104
10		105

14		106
15		108
17		109
18		110
19		112

12		114
13		116
22		117
23		119
25		120
26		121
27		122

39		132
41		133
42		134
46		135

12-(2'',2'',2''-Trifluoroacetamido)-N-(prop-2'-ynyl)dodecanamide, 3



12-Aminododecanoic acid (5.00 g, 23.2 mmol) was dissolved in MeOH (55 mL) and Et₃N (10 mL) was added. Ethyl trifluoroacetate (4.14 mL, 34.8 mmol) was added and the solution was stirred under an argon atmosphere, at rt. The reaction was complete after 22.5 hours. The solvents were removed *in vacuo* to give 12-(2'',2'',2''-trifluoroacetamido)'dodecanoic acid (**2**) as a pale yellow solid. This was then redissolved in distilled CH₂Cl₂ (100 mL), EDC.HCl (7.14 g, 37.3 mmol) and prop-2-yn-1-amine (1.96 mL, 28.7 mmol) were added and the solution was stirred at rt for 24 hours. The reaction was diluted with CH₂Cl₂, washed with saturated aqueous NaHCO₃, then saturated aqueous KCl, dried (Na₂SO₄), filtered and the solvents removed *in vacuo* to give a pale yellow solid. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 6:4) to give **3** as a colourless solid (6.35 g, 79% over two steps).

R_f (EtOAc:hexane 6:4, A) 0.33; (CH₂Cl₂:MeOH 4:1,A) 0.74.

mp 91-92 °C.

v_{max}(neat)/cm⁻¹ 3315 (m, NH), 3290 (m, NH), 2918 (m, CH₂), 2851 (w, CH₂), 1697 (s, CO), 1644 (s, CO), 1570 (w), 1531 (m), 1470 (m, CH₂), 1448 (w, CH₂), 1416 (w), 1373 (w), 1207 (s, CF₃), 1179 (s, CF₃), 719 (s).

¹H NMR (400MHz, CDCl₃): δ_H 6.43 (1H, br s, NH), 5.65 (1H, s, NH), 4.04 (2H, dd, *J* = 5.0, 2.5 Hz, H^{1'}), 3.35 (2H, q, *J* = 6.9 Hz, H¹²), 2.22 (1H, t, *J* = 2.5 Hz, H^{3'}), 2.18 (2H, t, *J* = 7.5 Hz, H²), 1.59 (4H, m, H^{3,11}), 1.28 (14H, m, H⁴⁻¹⁰).

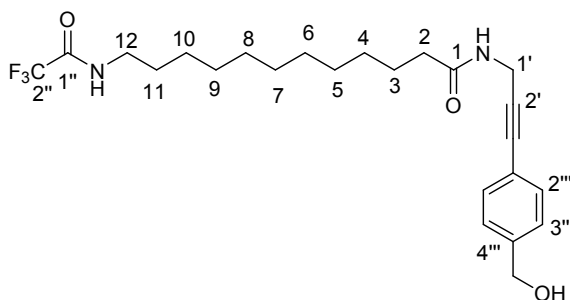
¹³C NMR (100MHz, CDCl₃): δ_C 172.8 (s, C¹O), 157.3 (s, C^{1''}O), 117.5 (s, C^{2''}), 79.8 (d, C^{3'}), 71.6 (s, C^{2'}), 40.1 (t, C¹²), 36.6 (t, C²), 29.0-29.4 (8t, C^{1',11,4-9}), 26.7 (t, C¹⁰), 25.6 (t, C³).

LRMS [ES⁺]: *m/z* 371 ([M+Na]⁺, 100%), 349 ([M+H]⁺, 80%).

HRMS [ES⁺]: *m/z* for C₁₇H₂₈F₃N₂O₂ [M+H]⁺: calcd 349.2098, found 349.2093.

Anal calculated: C, 58.61; H, 7.81; N, 8.04, found: C, 58.78; H, 7.84; N, 7.95.

12-(2'',2'',2''-Trifluoroacetamido)-N-{3'-[4'''-(hydroxymethyl)phenyl]prop-2'-ynyl}dodecanamide, 5



4-Iodobenzyl alcohol (500 mg, 2.14 mmol) was dissolved in anhydrous DMF (5 mL) to which **3** (819 mg, 2.35 mmol), Et₃N (1.49 mL, 10.7 mmol) and CuI (81 mg, 0.43 mmol) were added. The solution was stirred in the dark, at rt, under an argon atmosphere for 10 min before adding tetrakis(triphenylphosphine)palladium (247 mg, 0.21 mmol). The reaction went to completion after 1 hour 50 min. The solvents were removed *in vacuo* and the residue was redissolved in CH₂Cl₂. After washing with saturated aqueous NaHCO₃, then saturated aqueous KCl, the solution was dried (Na₂SO₄), filtered and the solvents removed *in vacuo* to give a yellow solid. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 7:3) to give **5** as a yellow solid (0.73 g, 75%).

R_f(EtOAc:hexane 7:3, B) 0.35.

mp 100 °C.

v_{max}(neat)/cm⁻¹ 3305 (m, NH), 2915 (m, CH₂), 2849 (w, CH₂), 1701 (s, CO), 1639 (s, CO), 1565 (w), 1525 (s, Aryl C=C), 1446 (m, CH₂), 1413 (w), 1372 (w), 1210 (s, CF₃), 1167 (s, CF₃), 1092 (m), 1008 (s, CH^{Ar}), 853 (m), 814 (m), 720 (s).

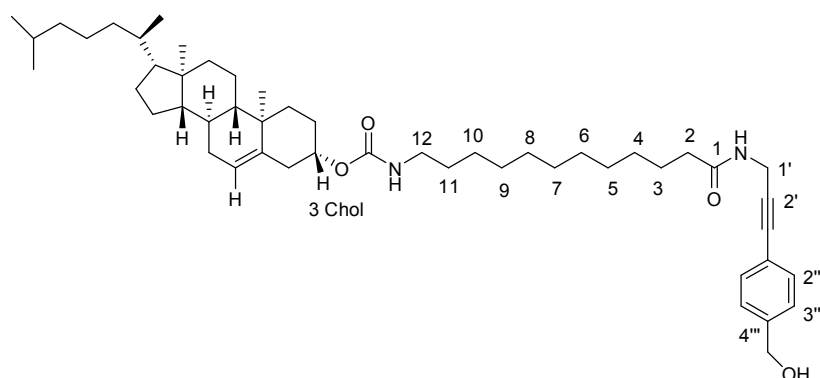
¹H NMR (400MHz, *d*₆-DMSO): δ_H 9.35 (1H, br s, NH), 8.28 (1H, s, NH), 7.36 (2H, d, *J* = 8.0 Hz, H^{2'''}), 7.30 (2H, d, *J* = 8.0 Hz, H^{3'''}), 5.23 (1H, t, *J* = 5.2 Hz, OH), 4.50 (2H, d, *J* = 5.0 Hz, CH₂OH), 4.09 (2H, d, *J* = 5.0 Hz, H^{1'}), 3.16 (2H, q, *J* = 6.5 Hz, H¹²), 2.10 (2H, t, *J* = 7.3 Hz, H²), 1.48 (4H, m, H^{3,11}), 1.24 (14H, s, H⁴⁻¹⁰).

¹³C NMR (100MHz, *d*₆-DMSO): δ_C 181.1 (s, C¹O), 171.9 (s, C^{1''}O), 143.0 (s, C^{4'''}), 131.0 (d, 2C^{2'''}), 126.5 (d, 2C^{3'''}), 128.8 (s, C^{2''}), 120.5 (s, C^{1'''}), 86.7 (s, C^{3'}), 81.5 (s, C^{2'}), 62.4 (t, CH₂OH), 38.9 (t, C¹²), 35.1 (t, C²), 28.1-28.9 (8t, C^{1',11,4-9}), 26.1 (t, C¹⁰), 25.1 (t, C³).

LRMS [ES⁺]: *m/z* 477 ([M+Na]⁺, 100%).

HRMS [ES⁺]: *m/z* for C₂₄H₃₃F₃N₂O₃Na [M+Na]⁺: calcd 477.2335, found 477.2334.

12-(Cholesteryloxycarbonylamino)-*N*-{3'-[4'''-(hydroxymethyl)phenyl]prop-2'-ynyl}dodecanamide, 7



Compound **5** (2.12 g, 4.68 mmol) was dissolved in ethanolic methylamine (33% w/v, 45.2 mL, 368 mmol). The reaction mixture was stirred at rt for 24 hours. The solvent was removed *in vacuo* and the residue was dried overnight under high vacuum to give 12-amino-*N*-{3'-[4'''-(hydroxymethyl)phenyl]prop-2'-ynyl}dodecanamide (**6**) as a brown solid (1.88 g). This compound was dissolved in distilled CH₂Cl₂ (40 mL) to which a solution of cholesteryl chloroformate (2.52 g, 5.62 mmol) in distilled CH₂Cl₂ (20 mL) was added. The solution was stirred at 0 °C, under an argon atmosphere for 2

hours and then at rt for a further 1.5 hours. The reaction was diluted with CH₂Cl₂, washed with water and saturated aqueous KCl, dried (Na₂SO₄), filtered and the solvents removed *in vacuo* to give a yellow solid. The crude mixture was purified by silica gel column chromatography (gradient of EtOAc:hexane 2:3 to 3:2) to give **7** as a colourless solid (1.19 g, 33% overall).

R_f (EtOAc:hexane 3:2, A) 0.30.

mp 110-111 °C.

v_{max} (neat)/cm⁻¹ 3284 (m, NH), 2919 (s, CH₂), 2848 (m, CH₂), 1686 (m, CO), 1636 (s, CO), 1528 (s, C=C Ar), 1464 (m, CH₂), 1253 (s, C-O), 1012 (s, C-OH), 652 (m, C=C).

¹H NMR (400 MHz, CDCl₃): δ_H 7.39 (2H, d, *J* = 8.0 Hz, **H^{2'''}**), 7.30 (2H, d, *J* = 8.0 Hz, **H^{3'''}**), 5.80 (1H, br s, NH), 5.36 (1H, m, **H^{6 Chol}**), 4.68 (2H, s, CH₂OH), 4.61 (1H, br s, NH), 4.49 (1H, m, **H^{3 Chol}**), 4.26 (2H, d, *J* = 5.0 Hz, **H^{1'}**), 3.12 (2H, q, *J* = 6.0 Hz, **H¹²**), 2.21 (2H, t, *J* = 7.5 Hz, **H²**), 1.48 (4H, m, **H^{3,11}**), 1.25 (14H, s, **H⁴⁻¹⁰**), 0.67-2.37 (44H, br m).

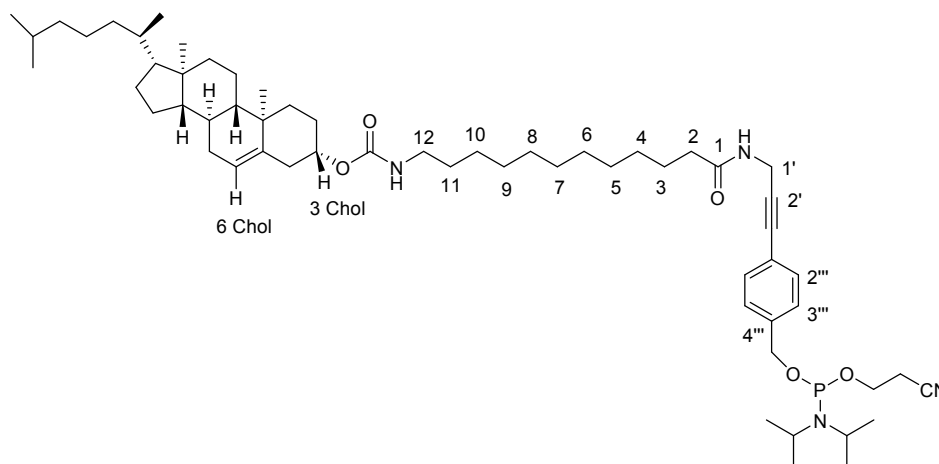
¹³C NMR (100 MHz, CDCl₃): δ_C 173.4 (s, **C¹O**), 157.5 (s, **C^{1'''}O**), 142.1 (s, **C^{4'''}**), 140.5 (s, **C^{5 Chol}**), 132.5 (2d, **C^{2'''}**), 127.4 (2d, **C^{3'''}**), 123.1 (d, **C^{6 Chol}**), 122.3 (s, **C^{1'''}**), 85.6 (s, **C^{2'}**), 83.9 (s, **C^{3'}**), 65.4 (t, CH₂OH), 57.4 (d), 56.8 (d), 50.7 (d), 43.0 (s), 40.4 (s and t), 40.2 (2t), 39.3 (2t), 37.2 (2t), 36.9 (2t), 36.5 (d), 32.6 (2t and d) 29.9 (2t), 28.9 (3t), 28.7(d), 26.2 (2t), 24.9 (2t), 24.5 (2t), 23.5 (d), 23.2 (2q), 21.7 (t), 20.0 (q), 19.4 (q), 12.5(q).

LRMS [ES⁺]: *m/z* 793 ([M+Na]⁺, 60%).

HRMS [ES⁺]: *m/z* for C₅₀H₇₈N₂O₄Na [M+Na]⁺: calcd 793.5854, for found 793.5865.

Anal calculated: C, 77.87; H, 10.19; N, 3.63, found: C, 77.43; H, 10.19; N, 3.52.

12-(Cholesteryloxycarbonylamino)-N-{3'-[4'''-(O-(2-cyanoethyl-N,N-diisopropyl phosphoramidyl)oxymethyl)phenyl]prop-2'-ynyl}dodecanamide, **8**



Under an argon atmosphere, 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (47 μ L, 0.21 mmol) was added dropwise to an argon-degassed solution of **7** (147 mg, 0.19 mmol) in THF (2 mL) with DIPEA (83 μ L, 0.48 mmol). After stirring at rt for 1.5 hours, the solution was transferred *via* cannula under an argon atmosphere to a separating funnel containing degassed ethyl acetate. The mixture was washed with degassed saturated aqueous KCl, dried (Na_2SO_4) and the solvent was removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc: hexane 1:1, 1% pyridine) gave **8** as a colourless air- and acid-sensitive foam (83 mg, 43%).

R_f (EtOAc:hexane 1:1) 0.35.

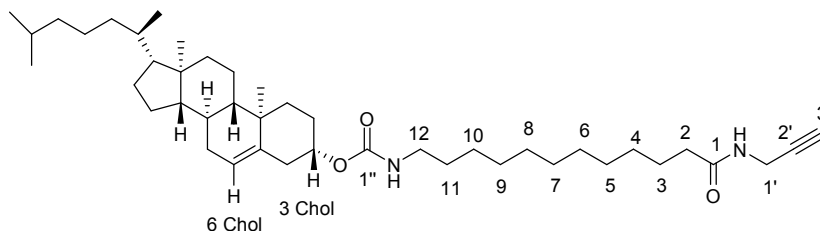
¹H NMR (300 MHz, *d*₆-DMSO): δ_{H} 8.27 (1H, br s, NH), 7.38 (2H, d, J = 8.4 Hz, **H**^{2'''}), 7.33 (2H, d, J = 8.4 Hz, **H**^{3'''}), 6.98 (1H, br s, NH), 5.33 (1H, d, J = 5.1 Hz, **H**^{6 Chol}), 4.68 (2H, m, CH₂-OP), 4.22 (1H, m, **H**^{3 Chol}), 4.09 (2H, d, J = 5.1 Hz, **H**^{1'}), 3.76 (2H, m, OCH₂CH₂CN), 3.60 (2H, m, *i*Pr-CH), 2.92 (2H, q, J = 6.6 Hz, **H**¹²), 2.77 (2H, t, J = 5.9 Hz, CH₂CN), 2.09 (2H, t, J = 7.5 Hz, **H**²), 0.69-2.13 (73 **H**, br m).

³¹P NMR (121 MHz, *d*₆-DMSO): δ_{P} 148.9.

LRMS [ES⁺]: m/z 994 ([M+Na]⁺, 50%).

HRMS [ES⁺]: m/z for C₅₉H₉₅N₄O₅PNa [M+Na]⁺: calcd 993.6932, found 993.6916.

12-Cholesteryloxycarbonylamino-*N*-(prop-2'-ynyl)dodecanamide, 10



Compound **3** (1.00 g, 2.87 mmol) was dissolved in ethanolic methylamine (33% w/v) (31.88 mL, 256 mmol). The mixture reaction was stirred at rt for 31 hours. The solvent was removed *in vacuo* and the residue was dried overnight under high vacuum to give **9** as a beige solid (980 mg). This compound was dissolved in distilled CH₂Cl₂ (20 mL) to which a solution of cholesteryl chloroformate (2.09 g, 4.66 mmol) in distilled CH₂Cl₂ (20 mL) was added. The solution was stirred at 0 °C, under an argon atmosphere for 2 hours and then at rt for 2 hours again. The reaction was diluted with CH₂Cl₂, washed with water and saturated aqueous KCl, dried (Na₂SO₄), filtered and the solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 2:3) to give **10** as a colourless solid (759 mg, 40% overall).

R_f (EtOAc:hexane 2:3, A) 0.29.

mp 108 °C.

v_{max}(neat)/cm⁻¹ 3298 (m, NH), 2919 (s, CH₂), 2846 (m, CH₂), 1696 (m, CO), 1655 (s, C=O), 1463 (m, CH₂), 1270 (m, C-O), 1235 (m, C-O), 649 (m, C=C), 618 (m, C-H alkyne).

¹H NMR (400 MHz, CDCl₃): δ_H 5.62 (1H, br s, NH), 5.42 (1H, m), 5.37 (1H, t, *J* = 2.5, H^{6 Chol}), 4.68 (1H, m), 4.57 (1H, br s, NH), 4.49 (1H, m, H^{3 Chol}), 4.05 (2H, dd, *J* = 5.3, 2.7 Hz, H^{1'}), 3.14 (2H, m, H¹²), 2.22 (1H, t, *J* = 2.7 Hz, H^{3'}), 2.19 (2H, t, *J* = 7.5 Hz, H²), 1.48 (4H, m, H^{3,11}), 1.26 (14H, s, H⁴⁻¹⁰), 0.68-2.47 (41H, br m).

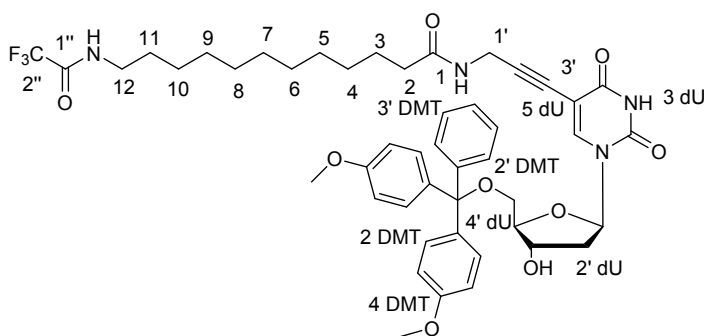
¹³C NMR (100 MHz, CDCl₃): δ_C 171.7 (s, C¹O), 153.7 (s, C^{1''}O), 138.9 (s, C^{5 Chol}), 122.9 (d, C^{6 Chol}), 82.1 (d, C^{3'}), 70.6 (s, C^{2'}), 55.7 (d), 55.2 (d), 49.1 (d), 41.4 (s), 38.8 (s and t), 38.6 (2t), 37.6 (t), 35.8 (t), 35.5 (2t), 35.2 (2t), 34.8 (d), 31.0 (t), 30.9 (t and

d), 28.4 (2t), 28.3 (2t), 28.2 (t), 27.3 (t), 27.2 (t), 27.0 (d), 26.4 (t), 23.3 (2t), 22.9 (t), 21.8 (d), 21.6 (2q), 20.1 (t), 18.4 (q), 17.8 (q), 10.9 (q).

LRMS [ES⁺]: *m/z* 688 ([M+Na]⁺, 60%), 666 ([M+H]⁺, 30%).

HRMS [ES⁺]: *m/z* for C₄₃H₇₂N₂O₃Na [M+Na]⁺: calcd 687.5435, found 687.5441.

5'-O-(4,4-Dimethoxytrityl)-5-[12-(2'',2'',2''-trifluoroacetamido)-N-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine, **14**



Nucleoside **11** (500 mg, 885 μ mol) was dissolved in anhydrous DMF (3.5 mL) to which **3** (678 mg, 2.00 mmol), Et₃N (1.23 mL, 8.84 mmol) and CuI (68 mg, 354 μ mol) were added. The solution was stirred in the dark, at rt, under an argon atmosphere for 10 min before adding tetrakis(triphenylphosphine)palladium (204 mg, 176 μ mol). The reaction had gone to completion after 23 hours. The solvents were removed *in vacuo*, the residue was dissolved in CH₂Cl₂, washed with saturated aqueous KCl and water, dried (Na₂SO₄), filtered and the solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 4:1, 0.5 % pyridine) to give **14** as yellow solid (510 mg, 66%).

R_f (EtOAc:hexane 4:1, 0.5% Et₃N, A) 0.24; (EtOAc:hexane 9:1, 0.5% Et₃N, C) 0.43.

v_{max} (neat)/cm⁻¹ 3297 (m, NH), 3058 (m, C-H), 2923 (m, CH₂), 2851 (m, CH₃), 1704 (s, CO), 1665 (s, CO), 1507 (s, C=C), 1436 (s, CH₂), 1174 (s, C-O), 1152 (s, C-O), 1091 (s, C-O), 721 (s, CH Ar).

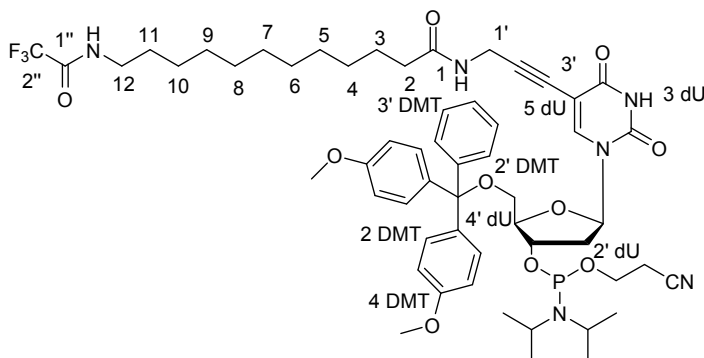
¹H NMR (400 MHz, CDCl₃): δ_H 9.09 (1H, br s, NH^{3 dU}), 8.12 (1H, s, H^{6 dU}), 7.09-7.61 (9H, m, H^{DMT}), 6.77 (4H, dd, *J* = 8.8, 1.3 Hz, H^{3 DMT}), 6.62 (1H, s, NH¹), 6.25 (1H, dd, *J* = 6.8, 6.8 Hz, H^{1' dU}), 5.38 (1H, t, *J* = 5 Hz, NH^{1''}), 4.47 (1H, dt, *J* = 6.0, 3.0 Hz, H^{3' dU}), 4.02 (1H, m, H^{4' dU}), 3.83 (1H, dd, *J* = 17.6, 5.0 Hz, H^{1'}), 3.77 (1H, dd, *J* = 17.6, 5.0 Hz, H^{1'}), 3.71 (6H, s, O-CH₃), 3.39 (1H, dd, *J* = 10.8, 2.8 Hz, H^{5' dU}), 3.30 (2H, m, H^{12, 5' dU}), 2.44 (1H, ddd, *J* = 13.6, 6.0, 3.0 Hz, H^{2' dU}), 2.21 (1H, m, H^{2' dU}), 1.82 (2H, t, *J* = 7.3 Hz, H¹¹), 1.49 (2H, t, *J* = 7.0 Hz, H²), 1.42 (2H, br t, *J* = 7.0 Hz, H³), 1.25 (14H, m, H⁴⁻¹⁰).

¹³C NMR (100MHz, CDCl₃): δ_C 172.9 (s, C¹O), 162.1 (s, C^{4 dU}O), 159.0 (3s, C^{1''}O, C^{4 DMT}), 149.4 (s, C^{2 dU}O), 144.9 (s, C^{1' DMT}), 143.5 (d, C^{6 dU}), 135.9 (2s, C^{1 DMT}), 132.4 (2d, C^{DMT}), 132.3 (d, C^{DMT}), 132.2 (d, C^{DMT}), 130.4 (d, C^{DMT}), 128.9 (d, C^{DMT}), 128.8 (d, C^{DMT}), 128.3 (d, C^{DMT}), 128.1 (d, C^{DMT}), 117.7 (s, C^{2''}), 113.7 (4d, C^{3 DMT}), 99.8 (s, C^{5 dU}), 89.9 (s, C^{3'}), 87.4 (d, C^{4' dU}), 87.1 (s, C^{0 DMT}), 86.2 (d, C^{1' dU}), 74.5 (s, C^{2'}), 72.4 (d, C^{3' dU}), 63.6 (t, C^{5' dU}), 55.6 (2q, OCH₃), 54.2 (t), 42.0 (t), 40.3 (t, C²), 29.4-36.8 (7t), 29.2 (t, C^{1'}), 26.9 (t, C¹⁰), 25.7 (t, C³).

LRMS [ES⁺]: *m/z* 900 ([M+Na]⁺, 30%).

HRMS [ES⁺]: *m/z* for C₄₃H₇₂N₂O₃Na [M+Na]⁺: calcd 899.3813, found 899.3825.

5'-*O*-(4,4-Dimethoxytrityl)-5-[12-(2'',2'',2''-trifluoroacetamido)-*N*-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine 3'-*O*-(2-cyanoethyl)-*N,N*-diisopropyl phosphoramidite, 15



Under an argon atmosphere, 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (122 μ L, 548 μ mol) was added dropwise to a degassed solution of **14** (437 mg, 498 μ mol) in THF (1.5 mL) with DIPEA (217 μ L, 1.25 mmol). After stirring at rt for 1.5 hours, the solution was transferred *via* cannula under an argon atmosphere to a separating funnel containing degassed ethyl acetate. The mixture was washed with degassed saturated aqueous KCl, dried (Na_2SO_4) and the solvent was removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc: hexane 7:3, 0.5 % pyridine) gave **15** as a diastereomeric mixture as a colourless (*ca.* 1:1) air- and acid-sensitive foam (251 mg, 47%).

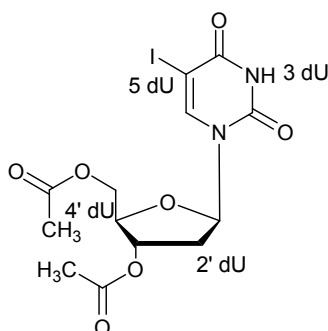
R_f (EtOAc:hexane 7:3, 0.5% Et₃N, C) 0.26 and 0.31.

¹H NMR (400 MHz, *d*₆-DMSO): δ_H 11.6 (1H, br s, NH³ dU), 9.36 (2H, br s, NH^{1''}), 8.14 (1H, m, NH¹), 7.93 (1H, s, H⁶ dU), 7.21-7.41(9H, m, H^{DMT}), 6.88 (4H, dd, *J* = 8.8, 1.3 Hz, H³ DMT), 6.09 (1H, dd, *J* = 15.9, 6.7 Hz, H^{1'} dU), 4.48 (1H, m, H^{3'} dU), 4.04 (1H, m, H^{4'} dU), 3.95 (1H, dd, *J* = 17.9, 5.2 Hz, H^{1'}), 3.89 (1H, dd, *J* = 17.9, 5.2 Hz, H^{1'}), 3.74 (6H, s, O-CH₃), 3.63 (2H, m, OCH₂CH₂CN), 3.53 (2H, m, *i*Pr-CH), 3.32 (2H, m, H^{5'} dU), 3.16 (2H, q, *J* = 6.6 Hz, H¹²), 2.75 (1H, t, *J* = 5.9 Hz, CH₂CN), 2.65 (1H, t, *J* = 5.9 Hz, CH₂CN), 2.30-2.43 (2H, m, H^{2'} dU), 2.05 (2H, t, *J* = 7.6 Hz, H²), 1.46 (4H, m, H^{3,11}), 1.22 (14H, s, H⁴⁻¹⁰), 0.98-1.14 (12H, m, *i*Pr-CH₃).

LRMS [ES⁺]: *m/z* 1100 ([M+Na]⁺, 100%).

 ^{31}P NMR (162 MHz, d_6 -DMSO) δ_{p} 148.3, 147.9.

3',5'-*O*-Acetyl-5-iodo-2'-deoxyuridine, **17**



5-Iodo-2'-deoxyuridine **16** (500 mg, 1.41 mmol) was dissolved in the minimum volume of pyridine (3 mL) and acetic anhydride (40 μ L, 4.24 mmol) was added. The reaction was stirred under an argon atmosphere for 16.5 hours. The pyridine was removed and the residue was co-evaporated with toluene *in vacuo*. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 3:2) to give **17** as a colourless solid (610 mg, 99%).

R_f (EtOAc:hexane 3:2, C) 0.28; (EtOAc:MeOH:NH₃ 5:1:1, C) 0.25.

mp 156-157 °C.

ν_{\max} (neat)/cm⁻¹ 3236 (m, NH), 1736 (s, CO), 1673 (s, CO), 1362 (s, CH), 1235 (s, C-O), 1194 (s, C-O), 1106 (s, C-O).

¹H NMR (400 MHz, CDCl₃): δ_{H} 9.20 (1H, s, NH^{3 dU}), 7.96 (1H, s, H^{6 dU}), 6.28 (1H, dd, $J = 8.3, 5.6$ Hz, H^{1' dU}), 5.23 (1H, dt, $J = 5.0, 2.0$ Hz, H^{3' dU}), 4.40 (1H, dd, $J = 12.6, 3.0$ Hz, H^{5' dU}), 4.33 (1H, dd, $J = 12.6, 3.0$ Hz, H^{5' dU}), 4.29 (1H, dt, $J = 5.0, 3.0$ Hz, H^{4' dU}), 2.54 (1H, ddd, $J = 14.1, 5.5, 2.0$ Hz, H^{2' dU}), 2.17 (1H, m, H^{2'}), 2.20 (3H, s, CH₃), 2.11 (3H, s, CH₃).

¹³C NMR (100 MHz, CDCl₃): δ_{C} 170.7 (s, CO-CH₃), 170.5 (s, CO-CH₃), 160.0 (s, C^{2 dU}O), 150.2 (s, C^{4 dU}O), 144.1 (d, C^{6 dU}), 85.8 (d, C^{1' dU}), 83.0 (d, C^{4' dU}), 74.4 (d, C^{3' dU}), 69.3 (s, C^{5 dU}), 64.2 (t, C^{5' dU}), 38.6 (t, C^{2' dU}), 21.4 (q, CO-CH₃), 21.2 (q, CO-CH₃).

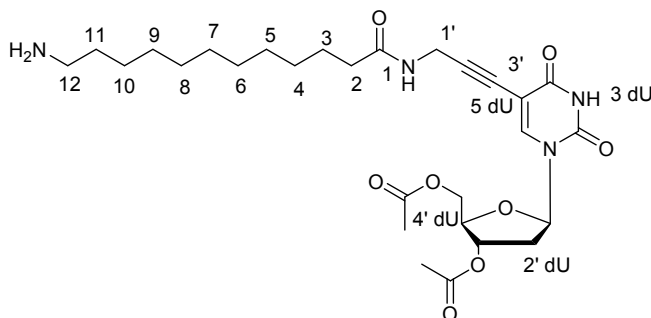
LRMS [ES⁺]: m/z 461 ([M+Na]⁺, 20%).

HRMS [ES⁺]: m/z for C₁₃H₁₅IN₂O₇Na [M+Na]⁺: calcd 460.9816, found 460.9823.

Anal calculated: C, 35.63; H, 3.45; N, 6.39, found: C, 35.60; H, 3.38; N, 6.31.

Analytical results consistent with reported data.¹⁴⁷

3',5'-O-Acetyl-5-[12-amino-N-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine, 18



Compound **3** (1.00 g, 2.87 mmol) was dissolved in ethanolic methylamine (33% w/v, 25 mL, 201 mmol) and the reaction mixture was stirred at 55 °C for 20 hours. The solvent was removed *in vacuo* and the residue was dried overnight under high vacuum to give **9** as a white solid (929 mg). Compound **9** (244 mg, 966 µmol), was then added to a solution of **17** (193 mg, 440 µmol) in anhydrous DMF (2 mL) followed by Et₃N (2.46 mL, 17.6 mmol) and CuI (34 mg, 179 µmol). The solution was stirred in the dark, at rt, under an argon atmosphere for 10 min before adding tetrakis(triphenylphosphine)palladium (102 mg, 88 µmol). The reaction had gone to completion after 30.5 hours. The solvents were removed *in vacuo* and the residue dissolved in a mixture of CH₂Cl₂ and MeOH and filtered through celite and the solvents were again removed *in vacuo*. The crude material was preadsorbed onto silica gel with MeOH and purified by silica gel column chromatography (EtOAc:hexane 8:2) to give **18** as brown syrup (238 mg, 73% overall).

R_f (EtOAc:MeOH:NH₃ 5:1:1, C) 0.20; (CH₂Cl₂:MeOH 6:4,C) 0.14.

ν_{max} (neat)/cm⁻¹ 3254 (w, NH), 2925 (m, CH₂), 2844 (w), 1710 (s, CO), 1651 (s, CO), 1456 (m), 1362 (m, CH), 1228 (s, C-O), 1101 (m), 1023 (m), 977 (C-O), 776 (m), 603 (m), 564 (m).

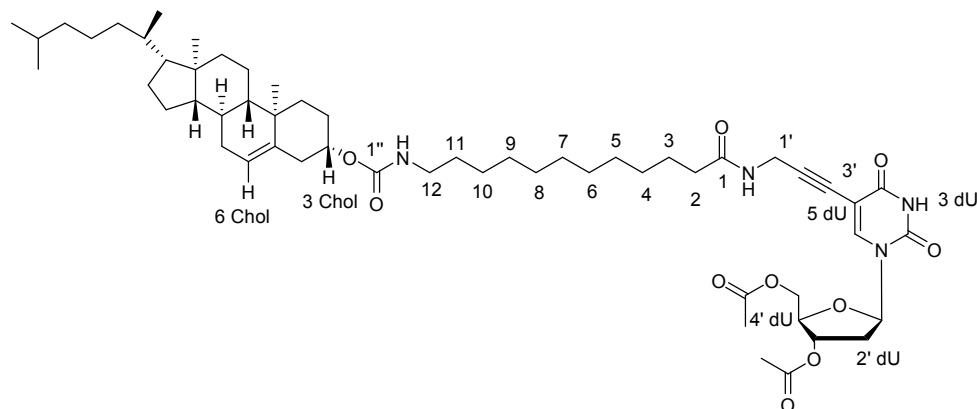
¹H NMR (400 MHz, MeOH): δ_H 7.93 (1H, s, **H**^{6 dU}), 6.18 (1H, dd, *J* = 7.7, 6.2 Hz, **H**^{1' dU}), 5.24 (1H, dt, *J* = 6.6, 2.6 Hz, **H**^{3' dU}), 4.26-4.37 (3H, m, **H**^{5' dU, 4' dU}), 4.10 (2H, s, **H**^{1'}), 2.89 (2H, t, *J* = 7.7 Hz, **H**¹²), 2.47 (1H, ddd, *J* = 14.6, 6.2, 2.7 Hz, **H**^{2' dU}), 2.36 (1H, m, **H**^{2' dU}), 2.15 (2H, m, **H**²), 2.11 (3H, s, CH₃), 2.11 (3H, s, CH₃), 1.60 (4H, m, **H**^{3,11}), 1.28 (14H, m, **H**⁴⁻¹⁰).

¹³C NMR (100MHz, MeOH): δ_c 177.0 (s, C¹O), 173.2 (2s, COCH₃), 145.6 (d, C⁶ dU), 134.3 (s), 134.1 (s), 131.2 (s), 105.8 (s), 101.6 (s, C⁵ dU), 88.4 (d, C^{1'} dU), 85.2 (d, C^{4'} dU), 76.7 (d, C^{3'} dU), 66.1 (t, C^{5'} dU), 42.0 (t, C²), 39.6 (2t, C^{2'} dU,12), 38.0 (t, C¹¹), 31.6 (t, C⁴⁻⁹), 31.5 (2t, C⁴⁻⁹), 31.4 (t, C⁴⁻⁹), 31.3 (2t, C⁴⁻⁹), 29.7 (t, C^{3,10,1'}), 28.5 (t, C^{3,10,1'}), 27.9 (t, C^{3,10,1'}), 21.9 (2q, CH₃).

LRMS [ES⁺]: m/z 563 ([M+H]⁺, 30%).

HRMS [ES⁺]: m/z for C₂₈H₄₃N₄O₈ [M+H]⁺: calcd 563.3075, found 563.3066.

3',5'-*O*-Acetyl-5-[12-cholesteryloxycarbonylamino-*N*-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine, **19**



A solution of cholesteryl chloroformate (133 mg, 297 μmol) in distilled CH_2Cl_2 (0.6 mL) was added to a stirred solution of **18** (139 mg, 247 μmol) in anhydrous DMF (1 mL) with DIPEA (52 μL , 297 μmol). The solution was stirred at 0 $^\circ\text{C}$, under an argon atmosphere for 2 hours and then at rt for 2 hours again. The solvents were removed *in vacuo* and the residue was dissolved in CH_2Cl_2 , washed with aqueous citric acid (10% w/v) and saturated aqueous KCl then dried (Na_2SO_4), filtered and the solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (gradient of EtOAc:hexane 3:2 to 4:1) to give **19** as a yellow solid (77 mg, 32%).

R_f (EtOAc:hexane 9:1, C) 0.30.

mp 132 $^\circ\text{C}$.

ν_{max} (neat)/ cm^{-1} 3273 (w, NH), 2924 (m, CH_2), 2845 (m, CH_3), 2358 (w, $\text{C}\equiv\text{C}$ alkyne), 1695 (s, CO), 1628 (s), 1542 (m), 1459 (m, CH_2), 1361 (m, CH_3), 1224 (s, C-O), 1039 (m), 953 (m, C-O), 776 w), 604 (w).

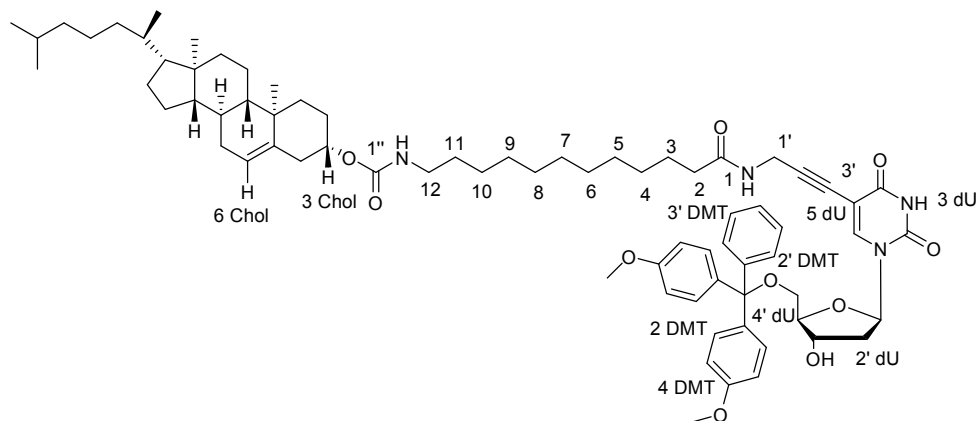
^1H NMR (400 MHz, CDCl_3): δ_{H} 9.06 (1H, s, $\text{NH}^{3\text{dU}}$), 7.80 (1H, s, $\text{H}^{6\text{dU}}$), 6.26 (1H, dd, $J = 7.5, 6.0$ Hz, $\text{H}^{1'\text{dU}}$), 6.06 (1H, br s, NH), 5.36 (1H, m, $\text{H}^{6\text{Chol}}$), 5.23 (1H, dt, $J = 7.0, 3.0$ Hz, $\text{H}^{3'\text{dU}}$), 4.64 (1H, br s, NH), 4.48 (1H, br s, $\text{H}^{3\text{Chol}}$), 4.36 (2H, m, $\text{H}^{5'\text{dU}}$), 4.28 (1H, m, $\text{H}^{4'\text{dU}}$), 4.23 (2H, d, $J = 5.0$ Hz, $\text{H}^{1'}$), 3.13 (2H, m, H^{12}), 2.54 (1H, ddd, $J = 14.1, 6.0, 2.5$ Hz, $\text{H}^{2'\text{dU},2}$), 2.20-2.25 (3H, m, $\text{H}^{2'\text{dU},2}$), 2.16 (3H, s, CH_3), 2.11 (3H, s, CH_3), 1.47 (4H, m, $\text{H}^{3,11}$), 1.25 (14H, s, H^{4-10}), 1.07-2.56 (43H, br m).

¹³C NMR (100MHz, CDCl₃): δ_C 173.0 (2s, CO), 170.4 (2s, COCH₃), 149.1 (2s, CO), 142.3 (d, C⁶ dU), 140.0 (s, C⁵ Chol), 122.6 (d, C⁶ Chol), 100.1 (s, C⁵ dU), 90.3 (s, C^{3'}), 85.8 (d, C^{1'} dU), 82.8 (d, C^{4'} dU), 74.1 (s and d, C^{2', 3'} dU), 63.8 (t, C^{5'} dU), 56.8 (d), 56.3 (d), 50.2 (d), 42.5 (s), 39.9 (t), 39.7 (2t), 38.7 (t), 38.4 (t), 37.1 (t), 36.7 (t), 36.6 (t), 36.3 (t), 35.9 (d), 32.0 (t and d), 30.1 (t), 30.0 (t), 29.8 (t), 29.6 (t), 29.5 (t), 29.4 (2t), 28.4 (2t), 28.1 (d), 25.6 (t), 24.4 (t), 23.9 (t), 22.9 (d), 22.7 (2q), 21.2 (t), 21.0 (t), 20.9 (2q), 19.5 (q), 18.9 (q), 12.0 (q).

LRMS [ES⁺]: *m/z* 997 ([M+Na]⁺, 100%).

HRMS [ES⁺]: *m/z* for C₅₆H₈₆N₄O₁₀Na [M+Na]⁺: calcd 997.6236, found 997.6244.

5'-O-(4,4-Dimethoxytrityl)-5-[12-cholesteryloxycarbonylamino-N-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine, **12**



Compound **19** (800 mg, 0.82 mmol) was dissolved in 0.4 M potassium carbonate in MeOH:H₂O (4:1 v/v, 13 mL, 5.20 mmol). The reaction mixture was stirred at rt for 1 hour 45 min and then diluted with CH₂Cl₂. DOWEX 50WX8-400 resin (pyridinium form) was added and the solution was stirred until pH 7. The mixture was filtered and the solvents were evaporated and co-evaporated with pyridine *in vacuo* to give **20** as a solid (891 mg). Compound **20** (653 mg, 0.73 mmol) was dissolved in CH₂Cl₂ (5 mL) to which 3 mL of pyridine was added followed by 4,4'-dimethoxytrityl chloride (298 mg, 0.88 mmol) in small portions. The reaction was stirred for 5 hours and 5 mL of MeOH was added to the solution. The solvents were evaporated and co-evaporated with toluene *in vacuo* and the crude mixture was purified by silica gel column chromatography (EtOAc:hexane 4:1, 1% pyridine) to give **12** as a yellow solid (358 mg, 45% overall)

R_f (EtOAc:hexane 4:1, 0.5% Et₃N, C) 0.41.

mp 104 °C.

ν_{max} (neat)/cm⁻¹ 3346 (w, NH), 3056 (w, CH Ar), 2925 (m, CH₂), 2844 (m, CH₃), 2357 (w, C-C alkyne), 1687 (s, CO), 1601 (m), 1503 (s, Aryl C=C), 1446 (s, CH₂), 1283 (s), 1242 (s, C-O), 1172 (s), 1172 (s), 1086 (m), 1025 (s, C-O), 825 (m), 723 (m), 694 (m), 576 (s).

¹H NMR (400 MHz, CDCl₃): δ_H 9.40 (1H, s, NH^{3 dU}), 8.19 (1H, s, H^{6 dU}), 7.17-7.44 (9H, m, H^{DMT}), 6.84 (4H, d, *J* = 8.5 Hz, H^{3 DMT}), 6.32 (1H, br t, *J* = 6.5 Hz, H^{1' dU}), 6.18

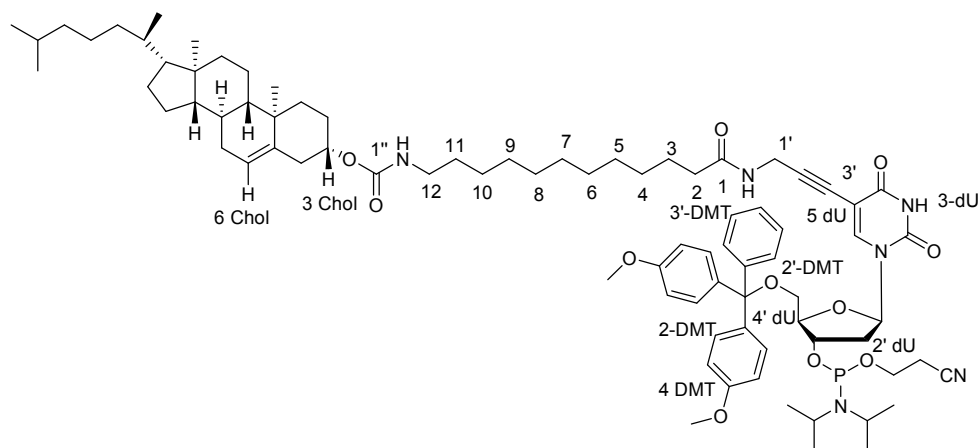
(1H, br s, NH), 5.53 (1H, br s, NH), 5.37 (1H, br s, $\text{H}^{6\text{ Chol}}$), 4.83 (1H, br s, NH), 4.69 (1H, br s, $\text{H}^{3'\text{ dU}}$), 4.11 (1H, br s, $\text{H}^{4'\text{ dU}}$), 3.87 (2H, d, $J = 5.5$ Hz, $\text{H}^{1'}$), 3.81 (3H, s, CH_3), 3.79 (3H, s, CH_3), 3.43 (1H, br d, $J = 9.0$ Hz, $\text{H}^{5'\text{ dU}}$), 3.33 (1H, br d, $J = 8.0$ Hz, $\text{H}^{5'\text{ dU}}$), 3.13 (2H, br s, H^{12}), 2.51 (1H, m, $\text{H}^{2'\text{ dU}}$), 2.22-2.36 (3H, m, $\text{H}^{2'\text{ dU}, 2}$), 1.52 (4H, m, $\text{H}^{3,11}$), 1.26 (14H, m, H^{4-10}), 0.97-6.16 (44H, br m).

^{13}C NMR (100MHz, CDCl_3): δ_{C} 172.8 (s, CO), 172.7 (s, CO), 158.8 (2s, $\text{C}^4\text{ DMT}$), 147.5 (2s, CO), 144.7 (s, $\text{C}^{1'\text{ DMT}}$), 143.3 (d, $\text{C}^{6\text{ dU}}$), 140.0 (s, $\text{C}^{5\text{ -Chol}}$), 135.7 (s, $\text{C}^{1\text{ DMT}}$), 135.6 (s, $\text{C}^{1\text{ DMT}}$), 130.2 (2d, C^{DMT}), 130.1 (d, C^{DMT}), 129.3 (d, C^{DMT}), 129.2 (d, C^{DMT}), 128.2 (d, C^{DMT}), 128.1 (d, C^{DMT}), 127.9 (d, C^{DMT}), 127.2 (d, C^{DMT}), 122.6 (d, $\text{C}^{6\text{-Chol}}$), 113.3 (4d, $\text{C}^{3\text{ DMT}}$), 99.1 (s, $\text{C}^{5\text{ dU}}$), 89.9 (s, $\text{C}^{3'}$), 86.3 (d, $\text{C}^{4'\text{ dU}}$), 86.0 (s, $\text{C}^{0\text{ DMT}}$), 85.9 (d, $\text{C}^{1'\text{ dU}}$), 74.4 (d, $\text{C}^{3'\text{ dU}}$), 74.2 (s, $\text{C}^{2'}$), 72.2 (t, $\text{C}^{5'\text{ dU}}$), 56.3 (d), 55.4 (d), 55.4 (q, O CH_3), 55.3 (q, O CH_3), 50.2 (d), 42.5 (s), 39.9 (t), 39.7 (t), 38.2 (t), 37.1 (t), 36.7 (t), 36.4 (2t), 36.3 (2t), 35.9 (d), 32.0 (t), 31.0 (d), 30.1 (t), 30.0 (t), 29.6 (2t), 29.5 (t), 29.4 (2t), 28.4 (t), 28.3 (t), 28.1 (d), 25.7 (t), 24.4 (t), 24.0 (t), 22.9 (d), 22.7 (2q), 21.2 (2t), 19.5 (2q), 18.9 (2q), 12.0 (q).

LRMS [ES^+]: m/z 1216 ($[\text{M}+\text{Na}]^+$, 40%).

Anal calculated: C, 73.46; H, 8.44; N, 4.69, found: C, 73.16; H, 8.46; N, 4.59.

5'-O-(4,4-Dimethoxytrityl)-5-[12-cholesteryloxycarbonylamino-*N*-(prop-2'-ynyl)dodecanamido]-2'-deoxyuridine-3'-O-(2-cyanoethyl)-*N,N*-diisopropyl phosphoramidite, **13**



Under an argon atmosphere, 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (61 μ L, 0.28 mmol) was added dropwise to a degassed solution of **12** (298 mg, 0.25 mmol) in THF (1.5 mL) with DIPEA (109 μ L, 0.62 mmol). After stirring at rt for 30 min, the solution was transferred *via* cannula under an argon atmosphere to a separating funnel containing degassed ethyl acetate. The mixture was washed with degassed saturated aqueous KCl, dried (Na_2SO_4) and the solvent was removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc: hexane 4:1, 0.5 % pyridine) gave **13** as a diastereomeric mixture (*ca.* 1:1) colourless and air- and acid-sensitive foam (284 mg, 82%).

R_f (EtOAc:hexane 4:1, 0.5% Et_3N , C) 0.20 and 0.33.

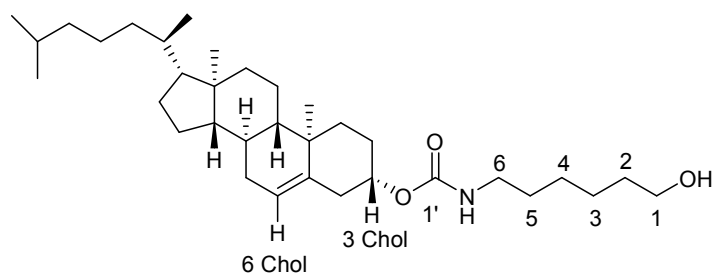
^1H NMR (400 MHz, d_6 -DMSO): δ_{H} 11.61 (1H, br s, $\text{NH}^{3\text{dU}}$), 8.12 (1H, dd, $J = 12.8, 5.3$ Hz, NH^1), 7.92 (1H, s, $\text{H}^{6\text{dU}}$), 7.22-7.40 (9H, m, H^{DMT}), 6.96 (1H, br s, $\text{NH}^{1''}$), 6.88 (4H, dd, $J = 8.3; 5.8$ Hz, $\text{H}^{3\text{DMT}}$), 6.09 (1H, dd, $J = 15.6; 7.0$ Hz, $\text{H}^{1'\text{dU}}$), 5.31 (1H, br s, $\text{H}^{6\text{Chol}}$), 4.48 (1H, br s, $\text{H}^{3\text{Chol}}$), 4.29 (1H, m, $\text{H}^{3'\text{dU}}$), 4.02 (1H, m, $\text{H}^{4'\text{dU}}$), 3.94 (1H, dd, $J = 17.7, 5.0$ Hz, $\text{H}^{1'}$), 3.88 (1H, d, $J = 17.7, 5.0$ Hz, $\text{H}^{1'}$), 3.74 (3H, s, O- CH_3), 3.73 (3H, s, O- CH_3), 3.67 (2H, m, OCH $_2$ CH $_2$ CN), 3.52 (2H, m, *i*Pr-CH), 3.19 (1H, dd, $J = 10.5, 3.0$ Hz, $\text{H}^{5'}$), 3.14 (2H, dd, $J = 10.5, 3.0$ Hz, $\text{H}^{5'}$), 2.92 (2H, q, $J = 6.5$ Hz,

H^{12}), 2.75 (1, t, $J = 5.8$ Hz, CH_2CN), 2.64 (1, t, $J = 5.8$ Hz, CH_2CN), 2.40 (1H, m, $\text{H}^{2' \text{ dU}}$), 2.23 (1H, m, $\text{H}^{2' \text{ dU}}$), 1.21 (14H, m, H^{6-12}), 0.64-2.05 (58H, br m).

^{31}P NMR (162 MHz, d_6 -DMSO) δ_{p} 148.5, 148.9.

LRMS $[\text{ES}^+]$: m/z 1216 ($[\text{M}+\text{Na}]^+$, 40%).

6-(Cholesteryloxycarbonylamino)-hexanol, **22**



A solution of cholesteryl chloroformate **21** (1.10 g, 2.45 mmol) in distilled CH_2Cl_2 (5 mL) was added to a stirred solution of 6-aminohexanol (631 mg, 5.38 mmol) in distilled CH_2Cl_2 (5 mL). The solution was stirred at 0 °C, under an argon atmosphere for 2 hours, then at rt for a further 16.5 hours. The reaction was diluted with CH_2Cl_2 , washed with aqueous citric acid (10% w/v) and saturated aqueous NaHCO_3 , dried (Na_2SO_4), filtered and the solvents were removed *in vacuo* to give a yellow solid. The crude mixture was purified by silica gel column chromatography (EtOAc:hexane 3:2) to give **22** as a colourless solid (1.12 g, 86%).

R_f (EtOAc:hexane 3:2, A) 0.41.

mp 126-127 °C.

ν_{max} (neat)/ cm^{-1} 3350 (w, NH), 2925 (s, CH_2), 2864 (m), 1687 (s, CO), 1538 (m), 1467 (m, CH_2). 1373 (m), 1251(s, CO), 1144 (m), 1005 (m, C-O), 645 (m, C=C).

^1H NMR (400 MHz, CDCl_3): δ_{H} 5.36 (1H, m, $\text{H}^{4 \text{ Chol}}$), 4.60 (1H, br s, NH), 4.49 (1H, m, $\text{H}^{3 \text{ Chol}}$), 3.63 (2H, t, $J = 6.5$ Hz, H^1), 3.15 (2H, t, $J = 6.8$ Hz, H^6), 2.35 (1H, ddd,

$J = 13.0, 5.0, 1.5 \text{ Hz, } \mathbf{H}^{2 \text{ Chol}}$), 2.26 (1H, m, $\mathbf{H}^{2 \text{ Chol}}$), 1.52 (2H, m, \mathbf{H}^2), 1.34 (2H, m, \mathbf{H}^5), 0.67-2.16 (46H, br m).

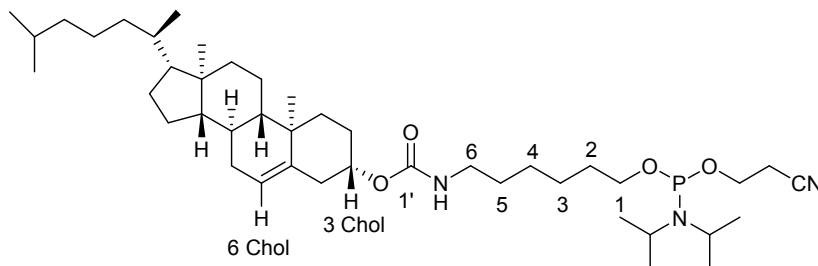
$^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 182.7 (s, $\text{C}^{1' \text{ O}}$), 142.0 (s, $\text{C}^{5 \text{ Chol}}$), 124.6 (d, $\text{C}^{6 \text{ Chol}}$), 64.9 (t, C^2), 58.9 (d), 58.3 (d), 52.2 (d), 44.5 (2s), 42.9 (t, C^4), 41.9 (t), 41.7 (t), 40.8 (t), 39.2 (t), 38.7 (t), 38.4 (t), 38.0 (d), 34.7 (t), 34.1 (t and d), 32.2 (t), 30.4 (t and d), 30.2 (d), 28.6 (t), 27.5 (t), 26.5 (t), 26.0 (t), 25.0 (2q), 23.2 (t), 21.5 (q), 20.9 (q), 14.0 (q).

LRMS [ES^+]: m/z 553 ($[\text{M}+\text{Na}]^+$, 100%), 1082 ($[\text{2M}+\text{Na}]^+$, 70%).

HRMS [ES^+]: m/z for $\text{C}_{34}\text{H}_{60}\text{NO}_3$ $[\text{M}+\text{H}]^+$: calcd 530.4567, found 530.4571.

Anal calculated: C, 77.07; H, 11.22; N, 2.64, found: C, 77.19; H, 11.33; N, 2.66.

6-(Cholesteryloxycarbonylamino)-hexanol 1-*O*-(2-cyanoethyl)-*N,N*-diisopropyl phosphoramidite, **23**



Under an argon atmosphere, 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (491 μ L, 2.20 mmol) was added dropwise to a degassed solution of **22** (1.06 g, 2.00 mmol) in THF (5 mL) with DIPEA (872 μ L, 5.00 mmol). After stirring at rt for 30 min, the solution was transferred *via* cannula under an argon atmosphere to a separating funnel containing degassed ethyl acetate. The mixture was washed with degassed saturated aqueous KCl, dried (Na_2SO_4) and the solvents were removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc:hexane 3:7, 0.5 % pyridine) gave **23** as a colourless air- and acid-sensitive foam (626 mg, 43%).

R_f (EtOAc:hexane 3:7, A) 0.46.

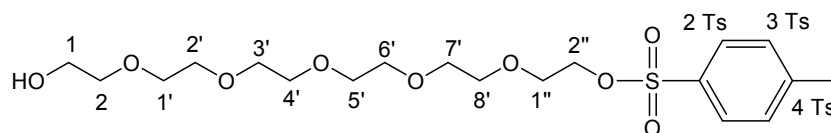
^1H NMR (400 MHz, d_6 -DMSO): δ_{H} 6.98 (1H, t, $J = 5.5$ Hz, NH), 5.33 (1H, m, $\text{H}^{4\text{ Chol}}$), 4.29 (1H, m, $\text{H}^{1\text{ Chol}}$), 3.71 (2H, m, $\text{OCH}_2\text{CH}_2\text{CN}$), 3.51-3.60 (4H, m, *i*Pr-CH, H^1), 2.93 (2H, q, $J = 6.8$ Hz, H^6), 2.79 (2H, t, $J = 6.0$ Hz, CH_2CN), 2.28 (1H, dd, $J = 12.8, 3.9$ Hz, $\text{H}^{2\text{ Chol}}$), 2.19 (1H, m, $\text{H}^{2\text{ Chol}}$), 1.51 (2H, m, H^2), 1.31 (2H, m, H^5), 0.69-2.30 (57H, br m).

^{31}P NMR (162 MHz, d_6 -DMSO) δ_{p} 147.6.

LRMS [ES^+]: m/z 753 ($[\text{M}+\text{Na}]^+$, 100%).

HRMS [ES^+]: m/z for $\text{C}_{43}\text{H}_{76}\text{N}_3\text{O}_4\text{PNa}$ $[\text{M}+\text{Na}]^+$: calcd 752.5466, found 752.5468.

Hexa(ethylene glycol) mono-*p*-toluenesulfonate, **25**



A solution of *p*-toluenesulfonyl chloride (2.53 g, 13.28 mmol) in CH₂Cl₂ (5 mL) was added dropwise over 2 hours to an ice-cooled solution of hexa(ethylene glycol) **24** (15 g, 53.13 mmol), Et₃N (7.41 mL, 53.13 mmol) and DMAP (779 mg, 6.37 mmol) in CH₂Cl₂ (10 mL) and the reaction mixture was stirred for a further 5 h at rt. The reaction mixture was dissolved in CH₂Cl₂, washed with distilled water, saturated aqueous NaHCO₃, and aqueous citric acid (10% w/v), dried (Na₂SO₄), filtered and solvent was removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (gradient of EtOAc:toluene 7:3 to neat EtOAc, then EtOAc:MeOH 95:5) to give **25** as a yellow oil (3.47 g, 60 %).

R_f (EtOAc) 0.18, (EtOAc:MeOH:Et₃N 5:1:1) 0.49, (CH₂Cl₂:MeOH:AcOH 10:2:1) 0.30.
v_{max}(neat)/cm⁻¹ 3450 (w, OH), 2869 (m, CH₂), 1598 (w, Aryl C=C), 1452 (w, CH₂), 1351 (m, SO₂), 1175 (s, SO₂), 1095 (s, C-O).

¹H NMR (400 MHz, CDCl₃): δ_H 7.79 (2H, dd, *J* = 8.3, 1.8 Hz, **H² Ts**), 7.33 (2H, br d, *J* = 8.5 Hz, **H³ Ts**), 4.13-4.17 (4H, m, **H¹, 2''**), 3.57-3.71 (20H, m, **H^{2,1'-8',1''}**), 2.44 (3H, s, **CH₃**).

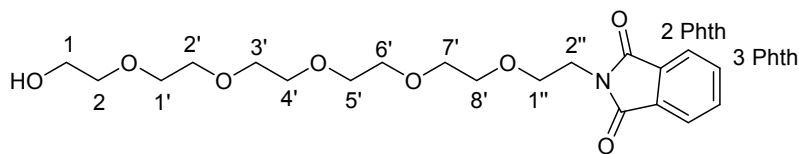
¹³C NMR (100MHz, CDCl₃): δ_C 144.9 (s, **C⁴ Ts**), 133.2 (s, **C¹ Ts**), 129.9 (2d, **C³ Ts**), 128.1 (2d, **C² Ts**), 72.7 (t, **C²**), 70.9 (t, **C^{1'-8'}**), 70.7 (4t, **C^{1'-8'}**), 70.6 (2t, **C^{1'-8'}**), 70.4 (t, **C^{1'-8'}**), 69.4 (t, **C^{1''}**), 68.8 (t, **C^{2''}**), 61.8 (t, **C¹**), 21.7 (q, **CH₃**).

LRMS [ES⁺]: *m/z* 454.5 ([M+NH₄]⁺, 50%), 459.4 ([M+Na]⁺, 100%).

HRMS [ES⁺]: *m/z* for C₁₉H₃₂O₉NaS [M+Na]⁺: calcd 459.1659, found 459.1659.

Analytical results consistent with reported data.¹⁴⁴

2-O-[(2''-Phthalimido-ethoxy)tetra(ethylene oxy)]ethanol, 26



Compound **25** (1.94 g, 4.44 mmol) and potassium phthalimide (1.24g, 6.66 mmol) were heated in anhydrous DMF (8 mL) under reflux in an argon atmosphere for 8 hours. The solvent was removed *in vacuo* and the residue was dissolved in CH₂Cl₂, washed with water and with saturated aqueous KCl, dried (Na₂SO₄), filtered and solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (CH₂Cl₂ with a gradient of MeOH, 1-3%) to give **26** as yellow oil (1.41 g, 78 %).

R_f (CH₂Cl₂:MeOH 19:1, D) 0.23, (CH₂Cl₂:MeOH 8:2, D) 0.65, (EtOAc:MeOH:Et₃N 5:1:1, E) 0.63.

v_{max}(neat)/cm⁻¹ 3468 (w, OH), 2869 (m, CH₂), 1773 (m, C=O), 1707 (s, C=O), 1467 (w, CH₂), 1095 (s, C-O), 720 (s, CH₂).

¹H NMR (400 MHz, CDCl₃): δ_H 7.82 (2H, dd, *J* = 5.5, 3.0 Hz, **H² Phth**), 7.69 (2H, dd, *J* = 5.0, 3.0 Hz, **H³ Phth**), 3.88 (2H, t, *J* = 5.8 Hz, **H^{2''}**), 3.69-3.73 (4H, m, **H¹, 1''**), 3.54-3.66 (18H, m, **H^{2,1'-8'}**).

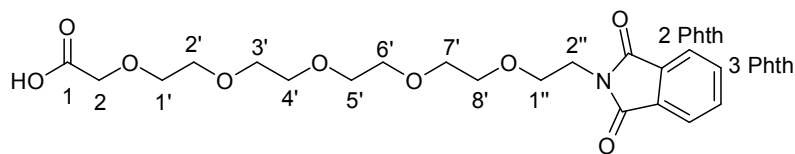
¹³C NMR (100MHz, CDCl₃): δ_C 168.3 (2s, CO), 134.0 (2d, **C³ Phth**), 132.3 (2s, **C¹ Phth**), 123.3 (2d, **C² Phth**), 72.6 (t, **C²**), 70.7 (2t, **C^{1'-8'}**), 70.6 (4t, **C^{1'-8'}**), 70.4 (t, **C^{1'-8'}**), 70.2 (t, **C^{1'-8'}**), 68.0 (t, **C^{1''}**), 61.8 (t, **C¹**), 37.4 (t, **C^{2''}**).

LRMS [ES⁺]: *m/z* 429.5 ([M+NH₄]⁺, 10%), 434.4 ([M+Na]⁺, 100%).

HRMS [ES⁺]: *m/z* for C₁₉H₃₂O₉SNa [M+Na]⁺: calcd 434.1785, found 434.1789.

Analytical results consistent with reported data.¹⁵⁷

2-O-[(2''-Phthalimidoethoxy)tetra(ethylene oxy)]acetic acid, 27



Compound **26** (1.00 g, 2.43 mmol) was dissolved in a 1:1 v/v mixture of water and acetonitrile (7 mL). TEMPO (76 mg, 0.49 mmol) was added followed by BAIB (2.35 g, 7.29 mmol). The reaction mixture was stirred at rt for 3.5 hours, then the solvents were removed. Purification by silica gel column chromatography (CH₂Cl₂, 3% MeOH) gave **27** as yellow oil (748 mg, 72 %).

R_f (CH₂Cl₂, MeOH 9:1, F) 0.30, (EtOAc:MeOH:Et₃N 5:1:1, F) 0.21, (CH₂Cl₂:MeOH:AcOH 10:2:1, F) 0.73.

v_{max}(neat)/cm⁻¹ 2872 (m, CH₂), 1706 (s, C=O), 1467 (w, CH₂), 1393 (m), 1353 (w), 1101 (s, C-O), 1024 (m), 719 (s, CH₂).

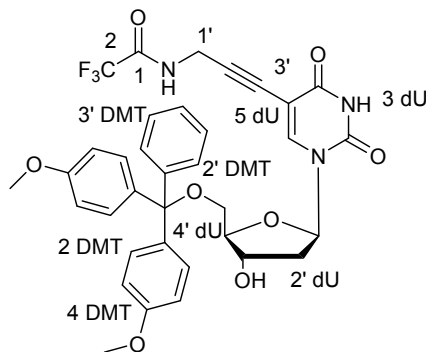
¹H NMR (400 MHz, CDCl₃): δ_H 7.82 (2H, dd, *J* = 5.0, 3.0 Hz, **H² Phth**), 7.70 (2H, dd, *J* = 5.3, 3.3 Hz, **H³ Phth**), 4.12 (2H, s, **H²**), 3.88 (2H, t, *J* = 5.8 Hz, **H^{2''}**), 3.72 (2H, t, *J* = 5.8 Hz, **H^{1''}**), 3.55-3.68 (16H, m, **H^{1'-8'}**).

¹³C NMR (100 MHz, CDCl₃): δ_C 172.3 (s, **C¹**), 168.4 (2s, CO), 134.1 (2d, **C³ Phth**), 132.3 (2s, **C¹ Phth**), 123.4 (2d, **C² Phth**), 71.2 (t, **C^{1'-8'}**), 70.7 (t, **C^{1'-8'}**), 70.6 (2t, **C^{1'-8'}**), 70.4 (2t, **C^{1'-8'}**), 70.3 (t, **C^{1'-8'}**), 70.1 (t, **C^{1'-8'}**), 69.2 (t, **C²**), 68.1 (t, **C^{1''}**), 37.4 (t, **C^{2''}**).

LRMS [ES⁺]: *m/z* 443.4 ([M+NH₄]⁺, 40%), 448.4 ([M+Na]⁺, 100%), 464.3 ([M+K]⁺, 10%).

HRMS [ES⁺]: *m/z* for C₁₉H₃₂O₉NNa [M+Na]⁺: calcd 448.1578, found 448.1576.

5'-O-(4,4-Dimethoxytrityl)-5-[2,2,2-trifluoro-N-(prop-2'-ynyl)-acetamido]-2'-deoxyuridine, **35**



5'-O-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine **11**¹⁴¹ (5.00 g, 7.62 mmol) was dissolved in anhydrous DMF (50 mL) to which 2',2',2'-trifluoro-N-(prop-2'-ynyl)acetamide **34** (1.20 mL, 10.8 mmol), Et₃N (7.43 mL, 53.3 mmol) and CuI (338 mg, 1.77 mmol) were added. The solution was stirred in the dark, at rt, under an argon atmosphere for 10 min before tetrakis(triphenylphosphine)palladium (880 mg, 0.76 mmol) was added. The reaction had gone to completion after 1.5 hours. The solvents were removed *in vacuo*, the residue dissolved in EtOAc, washed with aqueous Na₂EDTA (5% w/v), dried (Na₂SO₄), filtered and the solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (gradient of CH₂Cl₂:acetone 9:1 to 7:3, 1% pyridine) to give **35** as a yellow foam (4.96 g, 96%).

R_f (CH₂Cl₂:acetone 4:1, C) 0.33; (EtOAc:MeOH:NH₃ 5:1:1, A) 0.29.

¹H NMR (400 MHz, CDCl₃): δ_H 9.78 (1H, br s, NH), 8.23 (1H, s, H⁶ dU), 7.16-7.42 (9H, m, H^{DMT}), 6.84 (4H, d, *J* = 8.5 Hz, H³ DMT), 6.35 (1H, dd, *J* = 7.7, 6.0 Hz, H^{1'} dU), 4.59 (1H, dt, *J* = 5.8, 2.8 Hz, H^{3'} dU), 4.14 (1H, m, H^{4'} dU), 3.95 (1H, dd, *J* = 17.9, 4.8 Hz, H^{1'}), 3.88 (1H, dd, *J* = 17.9, 4.7 Hz, H^{1'}), 3.77 (6H, s, O-CH₃), 3.38 (1H, dd, *J* = 11.3, 3.0 Hz, H^{5'} dU), 3.35 (1H, dd, *J* = 11.0, 3.0 Hz, H^{5'} dU), 2.57 (1H, ddd, *J* = 13.8, 5.8, 2.8, H^{2'} dU), 2.32 (1H, m, H^{2'} dU).

¹³C NMR (100 MHz, CDCl₃): δ_C 162.5 (s, CO), 158.8 (2s, C^{4-DMT}), 149.5 (2s, CO), 144.6 (s, C^{1'-DMT}), 143.7 (d, C⁶ dU), 135.6 (2s, C^{1-DMT}), 130.2 (d, C^{DMT}), 130.1 (d, C^{DMT}), 129.2 (2d, C^{DMT}), 128.4 (2d, C^{DMT}), 128.2 (2d, C^{DMT}), 128.0 (d, C^{4'-DMT}), 124.0

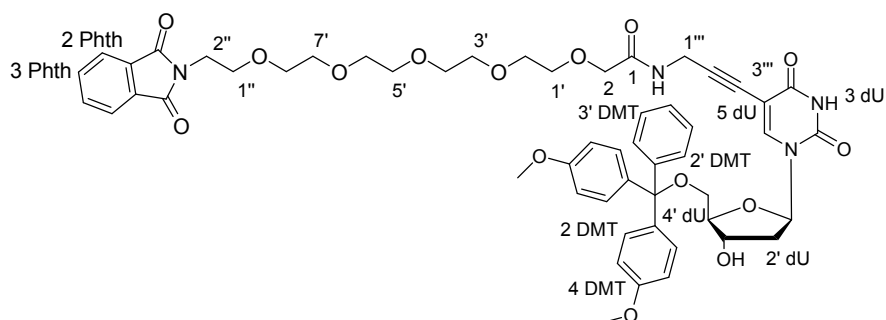
(s, C²), 113.5 (4d, C^{3-DMT}), 99.2 (s, C^{5 dU}), 87.6 (s, C^{0-DMT}), 87.2 (d, C^{4' dU}), 86.2 (d, C^{1' dU}), 77.6 (s, C^{3'}), 75.5 (s, C^{2'}), 72.5 (d, C^{3' dU}), 63.7 (t, C^{5' dU}), 55.4 (2q, CO-CH₃), 41.7 (t, C^{2' dU}), 30.5 (t, C^{1'}).

LRMS [ES⁺]: *m/z* 702 ([M + Na]⁺, 100); 697 ([M + NH₄]⁺, 15).

HRMS [ES⁺]: *m/z* for C₃₅H₃₂F₃N₃O₈ [M + Na]⁺: calcd 702.2034, found 702.2043.

Analytical results consistent with reported data.¹⁴⁸

5'-O-(4,4-Dimethoxytrityl)-5-{2-O-[(2''-phthalimido-ethoxy)tetra(ethylene oxy)]-N-prop-2'''-ynyl-acetamido}-2'-deoxyuridine, **29**



Compound **35** (3.11g, 4.58 mmol) was dissolved in ethanolic methylamine (33% w/v) (28.48 mL, 229 mmol) and the reaction mixture was stirred at rt for 2 hours. The solvent was removed *in vacuo* and the product was dried overnight under high vacuum to give **28** as a yellow foam (2.97 g, 5.09 mmol) which was used without further purification.

2-[(2-Phthalimido-ethoxy)tetra(ethylene oxy)]acetic acid **27** was dissolved in CH₂Cl₂ (10 mL) to which PyBroP (1.21 g, 2.59 mmol) and DIPEA (1.23 mL, 7.05 mmol) were added. The solution was stirred at rt under an argon atmosphere for 20 mins before adding the nucleoside **28** (1.51 g, 2.59 mmol). The reaction had gone to completion after 1.5 hours. The solvents were removed *in vacuo* and the residue was dissolved in CH₂Cl₂, washed with saturated aqueous NaHCO₃, dried (Na₂SO₄) and the solvent was removed *in vacuo*. Purification by silica gel column chromatography under argon

pressure (CH₂Cl₂ with a gradient of MeOH from 0.5% to 2%, 0.5 % Et₃N) gave **29** as a pale yellow foam (1.13 g, 49%).

R_f (CH₂Cl₂:acetone 4:1, A) 0.11; (CH₂Cl₂:MeOH 95:5, C) 0.19.

ν_{max} (solid)/cm⁻¹ 3100 (w, CH^{Phth}), 2868 (w, CH₂), 1772 (w), 1708 (s, CO), 1607 (m), 1508 (m, C=C), 1456 (m, CH₂), 1393 (m), 1348 (w), 1280 (m), 1247 (s, CO), 1175 (m, CO), 1089 (s, CO), 1028 (s, C–O), 838 (s), 721 (s, CH^{Phth}), 583 (m).

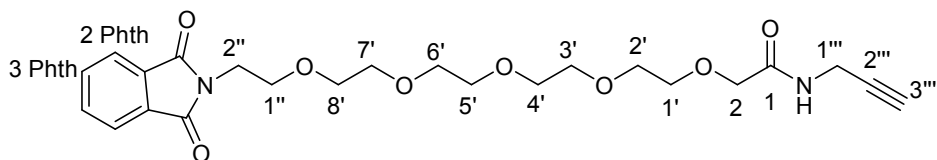
¹H NMR (400 MHz, CDCl₃): δ_H 7.98 (1H, s, **H**^{6 dU}), 7.82 (2H, dd, *J* = 5.5, 3.0 Hz, **H**^{2 Phth}), 7.68 (2H, dd, *J* = 5.5, 3.0 Hz, **H**^{3 Phth}), 7.18-7.42 (10 H, m, **H**^{DMT}, **NH**), 6.83 (4H, d, *J* = 8.5 Hz, **H**^{3 DMT}), 6.26 (1H, dd, *J* = 7.3, 6.0 Hz, **H**^{1' dU}), 4.54 (1H, dt, *J* = 6.0, 3.0 Hz, **H**^{3' dU}), 4.08 (1H, m, **H**^{4' dU}), 4.03 (1H, dd, *J* = 18.0, 5.5 Hz, **H**^{1'''}), 3.93 (1H, dd, *J* = 18.0, 5.5 Hz, **H**^{1'''}), 3.94 (2H, s, **H**²), 3.87 (2H, t, *J* = 5.8 Hz, **H**^{2''}), 3.76 (6H, s, O-CH₃), 3.71 (2H, t, *J* = 5.8 Hz, **H**^{1''}), 3.55-3.62 (16H, m, **H**^{1'-8'}), 3.38 (1H, dd, *J* = 10.6, 4.0 Hz, **H**^{5' dU}), 3.31 (1H, dd, *J* = 10.5, 3.0 Hz, **H**^{5' dU}), 2.49 (1H, ddd, *J* = 13.6, 6.0, 3.0 Hz, **H**^{2' dU}), 2.26 (1H, m, **H**^{2' dU}).

¹³C NMR (100 MHz, CDCl₃): δ_C 169.7 (s, CO), 168.5 (2s, CO), 161.6 (s, CO), 158.8 (2s, **C**^{4-DMT}), 149.2 (s, CO), 144.7 (s, **C**^{1'-DMT}), 143.2 (d, **C**^{6 dU}), 135.7 (s, **C**^{1-DMT}), 135.6 (s, **C**^{1-DMT}), 134.1 (2d, **C**^{3 Phth}), 132.2 (2s, **C**^{1 Phth}), 130.2 (2d), 130.1 (4d), 128.1 (d), 128.0 (d), 127.1 (d, **C**^{4'-DMT}), 123.4 (2d, **C**^{2 Phth}), 113.5 (4d, **C**^{3-DMT}), 99.6 (s, **C**^{5 dU}), 87.1 (s, **C**^{0-DMT}), 86.5 (d, **C**^{4' dU}), 85.9 (d, **C**^{1' dU}), 77.3 (s, **C**^{3'''}), 74.1 (s, **C**^{2'''}), 72.2 (d, **C**^{3' dU}), 71.1 (2t, **C**^{1'-8'}), 70.6 (4t, **C**^{1'-8'}), 70.5 (t, **C**^{1'-8'}), 70.3 (t, **C**^{3'-8'}), 70.2 (t, **C**²), 68.0 (t, **C**^{1''}), 63.8 (t, **C**^{5' dU}), 55.4 (2q, CO-CH₃), 41.5 (t, **C**^{2' dU}), 37.4 (t, **C**^{2''}), 29.4 (t, **C**^{1'''}).

LRMS [ES⁺]: *m/z* 1014 ([M + Na]⁺, 100%); 1009 ([M + NH₄]⁺, 81).

HRMS [ES⁺]: *m/z* for C₅₃H₅₈N₄O₁₅Na [M + Na]⁺: calcd 1013.3791, found 1013.3787.

2-O-[(2''-Phthalimido-ethoxy)tetra(ethylene oxy)]-N-prop-2'''-ynyl-acetamide, 33



2-[(2-Phthalimido-ethoxy)tetra(ethylene oxy)]acetic acid **27** (150 mg) was dissolved in CH_2Cl_2 (1.5 mL) to which PyBroP (181 mg, 0.39 mmol) and DIPEA (184 μL , 1.06 mmol) were added. The solution was stirred at rt under an argon atmosphere for 20 mins before adding prop-2-yn-1-amine (36 μL , 0.53 mmol). The reaction had gone to completion after 1 hour 15 min. The reaction mixture was diluted in CH_2Cl_2 , washed with saturated aqueous NaHCO_3 , distilled water and aqueous citric acid (10% w/v), dried (Na_2SO_4), filtered and the solvent was removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc to EtOAc:acetone 1:4) gave **33** as a pale yellow oil (85 mg, 53%).

R_f (CH_2Cl_2 :MeOH, 9:1, G) 0.55; (EtOAc:MeOH: NH_3 5:1:1, G) 0.54; (EtOAc:acetone, 9:1, A) 0.25.

ν_{max} (neat)/ cm^{-1} 3255 (m, CH), 2886 (w, CH), 2864 (w, CH), 1773 (w, CO), 1707 (s, CO), 1664, (m), 1519 (w), 1468 (w, CH_2), 1425 (w), 1342 (m), 1320 (w), 1286 (w), 1141 (w), 1107 (s, C–O), 1096 (s, C–O), 1058 (m), 1031 (m), 956 (w), 935 (w), 841 (m), 714 (s, CH_2).

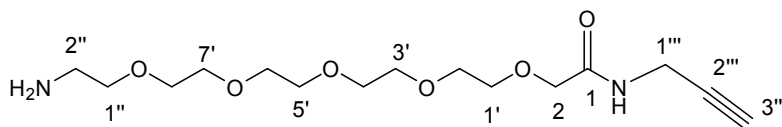
^1H NMR (400 MHz, CDCl_3): δ_{H} 7.82 (2H, dd, $J = 5.5, 3.0$ Hz, H^2 Phth), 7.70 (2H, dd, $J = 5.5, 3.0$ Hz, H^3 Phth), 7.38 (1H, br s, NH), 4.06 (2H, dd, $J = 5.5, 2.5$ Hz, $\text{H}^{1''}$), 4.00 (2H, s, H^2), 3.88 (2H, t, $J = 5.8$ Hz, $\text{H}^{2''}$), 3.72 (2H, t, $J = 5.8$ Hz, $\text{H}^{1''}$), 3.56–3.67 (16H, m, $\text{H}^{1'-8'}$), 2.21 (1H, t, $J = 2.5$ Hz, $\text{H}^{3'''}$).

^{13}C NMR (100 MHz, CDCl_3): δ_{C} 170.0 (s, CO), 168.4 (2s, CO), 134.0 (2d, C^3 Phth), 132.3 (2s, C^1 Phth), 123.3 (2d, C^2 Phth), 79.3 (s, $\text{C}^{2'''}$), 71.3 (d, $\text{C}^{1'''}$), 70.7 (8t, $\text{C}^{1'-8'}$), 70.2 (t, C^2), 68.0 (t, $\text{C}^{1''}$), 37.4 (t, $\text{C}^{2''}$), 28.5 (t, $\text{C}^{3'''}$).

LRMS [ES^+]: m/z 480.5 ($[\text{M} + \text{NH}_4]^+$, 10%), 485.4 ($[\text{M} + \text{Na}]^+$, 100%).

HRMS [ES^+]: m/z for $\text{C}_{23}\text{H}_{30}\text{N}_2\text{NaO}_8$ $[\text{M} + \text{Na}]^+$: calcd 485.1894, found 485.1894.

2-O-[(2''-Amino-ethoxy)tetra(ethylene oxy)]-N-prop-2'''-ynyl-acetamide, 36



Compound **33** (490 mg, 1.06 mmol) was dissolved in EtOH (5 mL) to which hydrazine monohydrate (98% w/v, 0.41 mL, 8.48 mmol) was added. The reaction mixture was stirred at rt for 1 hour 20 min. The solvent was removed *in vacuo*, the residue was dissolved in EtOAc, the precipitate was removed by filtration and solvent was removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (CH₂Cl₂:MeOH, 95:5, 0.5 % Et₃N) to give **36** as brown syrup (249 mg, 71%).

R_f (CH₂Cl₂:MeOH 9:1, G) 0.29; (CH₂Cl₂:MeOH 19:1, G) 0.21; (acetone:MeOH 3:2, G) 0.18.

ν_{max} (neat)/cm⁻¹ 3293 (w), 2871 (m, CH), 1665 (CO), 1531 (w, NH), 1455 (w), 1348 (w), 1284 (w), 1248 (w), 1096 (s, C–O), 944 (w), 843 (m).

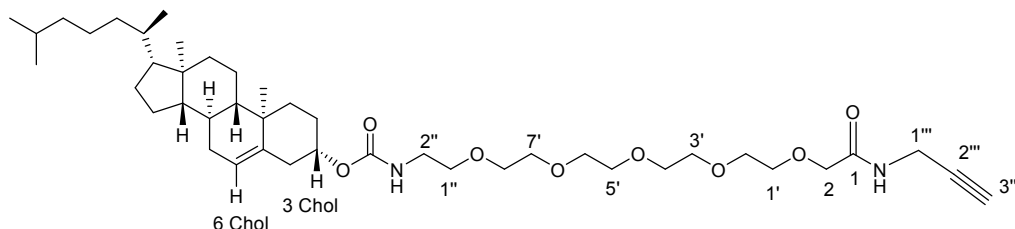
¹H NMR (400 MHz, CDCl₃): δ_H 7.44 (1H, br s, NH), 4.06 (2H, dd, *J* = 5.5, 2.5 Hz, H^{1'''}), 3.99 (2H, s, H²), 3.60-3.66 (16H, m, H^{1'-8'}), 3.48 (2H, t, *J* = 5.3 Hz, H^{1''}), 2.83 (2H, t, *J* = 5.3 Hz, H^{2''}), 2.21 (1H, t, *J* = 2.5 Hz, H^{3'''}), 1.64 (2H, br s, NH₂).

¹³C NMR (100 MHz, CDCl₃): δ_C 170.0 (s, CO), 79.8 (s, C^{2'''}), 73.5 (t, C^{1''}), 71.2 (d, C^{3'''}), 70.7 (8t, C^{1'-8'}), 70.2 (t, C²), 41.9 (t, C^{2''}), 28.4 (t, C^{1'''}).

LRMS [ES⁺]: *m/z* 333.2 ([M + H]⁺, 100%).

HRMS [ES⁺]: *m/z* for C₁₅H₂₉N₂O₆ [M + H]⁺: calcd 333.2020, found 333.2017.

2-*O*-[(2''-Cholesteryloxycarbonylamino-ethoxy)tetra(ethylene oxy)]-*N*-prop-2'''-ynyl-acetamide, 37



A solution of cholesteryl chloroformate (404 mg, 0.90 mmol) in distilled CH_2Cl_2 (2.5 mL) was added to a stirred solution of **36** (176 mg, 0.53 mmol) in distilled CH_2Cl_2 (2 mL). The solution was stirred at 0°C , under an argon atmosphere for 2 hours, then at rt for a further 19 hours. The reaction mixture was diluted with CH_2Cl_2 , washed with aqueous citric acid (10% w/v) and saturated aqueous KCl, dried (Na_2SO_4), filtered and the solvent was removed *in vacuo*. The crude mixture was purified by column chromatography (CH_2Cl_2 :MeOH 97:3) to give **37** as a pale pink oil (347 mg, 88%).

R_f (CH_2Cl_2 :MeOH 19:1, G) 0.21; (CH_2Cl_2 :MeOH 9:1, G) 0.57; (EtOAc:MeOH: NH_3 5:1:1, A) 0.59.

ν_{max} (neat)/ cm^{-1} 3306 (w), 2933 (m, CH), 2866 (m, CH), 1674 (m, CO), 1530 (m, NH), 1442 (m), 1398 (m), 1376 (m), 1344 (m), 1233 (m), 1105 (s, C–O), 949 (w), 843 (w).

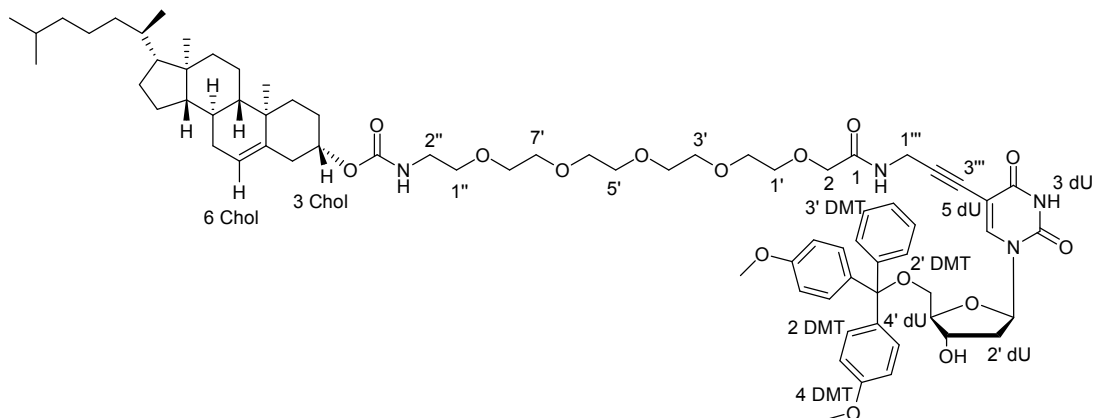
^1H NMR (400 MHz, CDCl_3): δ_{H} 7.38 (1H, br s, NH), 5.35 (1H, m, $\text{H}^{6\text{ Chol}}$), 4.50 (1H, m, $\text{H}^{3\text{ Chol}}$), 4.08 (2H, dd, $J = 5.5, 2.5$ Hz, $\text{H}^{1''}$), 4.00 (2H, s, H^2), 3.60–3.68 (16H, m, $\text{H}^{1'-8'}$), 3.53 (2H, t, $J = 5.0$ Hz, $\text{H}^{2''}$), 3.34 (2H, br d, $\text{H}^{1''}$), 2.22 (1H, t, $J = 2.5$ Hz, $\text{H}^{3''}$), 0.67–2.38 (br m, 44H).

^{13}C NMR (100 MHz, CDCl_3): δ_{C} 170.0 (s, CO), 154.6 (s, CO), 140.0 (s, $\text{C}^{5\text{ Chol}}$), 122.6 (d, $\text{C}^{6\text{ Chol}}$), 79.9 (s, $\text{C}^{2'''}$), 75.2 (t, $\text{C}^{3\text{ Chol}}$), 71.3 (t, $\text{C}^{1''}$), 71.2 (d, $\text{C}^{3'''}$), 70.7 (8t, $\text{C}^{1'-8'}$), 69.1 (d, C^2), 56.8 (t), 56.3 (t), 50.2 (t), 42.5 (s and t, $\text{C}^{2''}$), 39.9 (s), 38.7 (t), 38.6 (t), 36.7 (t), 36.3 (t), 35.9 (2d), 32.0 (d and t), 28.3 (t, $\text{C}^{3'''}$), 28.2 (d), 28.1 (d), 24.4 (t), 24.0 (t), 22.9 (d), 22.7 (q), 21.4 (t), 21.2 (q), 19.5 (q), 18.8 (q), 12.0 (q).

LRMS [ES^+]: m/z 762.6 ($[\text{M} + \text{NH}_4]^+$, 10%), 767.6 ($[\text{M} + \text{Na}]^+$, 100%).

HRMS [ES^+]: m/z for $\text{C}_{43}\text{H}_{72}\text{N}_2\text{NaO}_8$ $[\text{M} + \text{Na}]^+$: calcd 767.5181, found 767.5176.

5'-O-(4,4-Dimethoxytrityl)-5-{2-O-[(2''-cholesteryloxycarbonylamino-ethoxy)tetra(ethylene oxy)]-N-prop-2'''-ynyl-acetamido}-2'-deoxyuridine, **31**



5'-O-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine **11** (106 mg, 0.16 mmol) was dissolved in anhydrous DMF (1.5 mL) to which Et₃N (7.43 mL, 53.3 mmol), CuI (7 mg, 0.04 mmol) and a solution of **37** (181 mg, 0.24 mmol) in anhydrous DMF (2 mL) were added. The solution was stirred in the dark, at rt, under an argon atmosphere for 10 min before tetrakis(triphenylphosphine) palladium (18 mg, 0.02 mmol) was added. After 1.5 hours TLC analysis indicated that the reaction had not gone to completion, hence further CuI (7 mg, 0.04 mmol) was added to the reaction mixture. The reaction was complete after 3.5 hours. The solvents were removed *in vacuo*, the residue was dissolved in EtOAc, washed with aqueous Na₂EDTA (5% w/v), the organic layer was dried (Na₂SO₄), filtered and the solvents were removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (EtOAc:MeOH 95:5, 0.5 % Et₃N) to give **31** as a pale yellow foam (129 mg, 63%).

R_f (CH₂Cl₂:MeOH 19:1, C) 0.21; (CH₂Cl₂:MeOH 97:3, C) 0.14; (EtOAc:MeOH:NH₃ 5:1:1, C) 0.29.

ν_{max} (solid)/cm⁻¹ 3350 (w), 2932 (m, CH), 2866 (m, CH), 1694 (s, CO), 1607 (w), 1508 (m, NH), 1455 (m), 1402 (w), 1344 (m), 1280 (m), 1248 (s, CO), 1175 (m), 1095 (s, CO), 1032 (m), 949 (w), 827 (m).

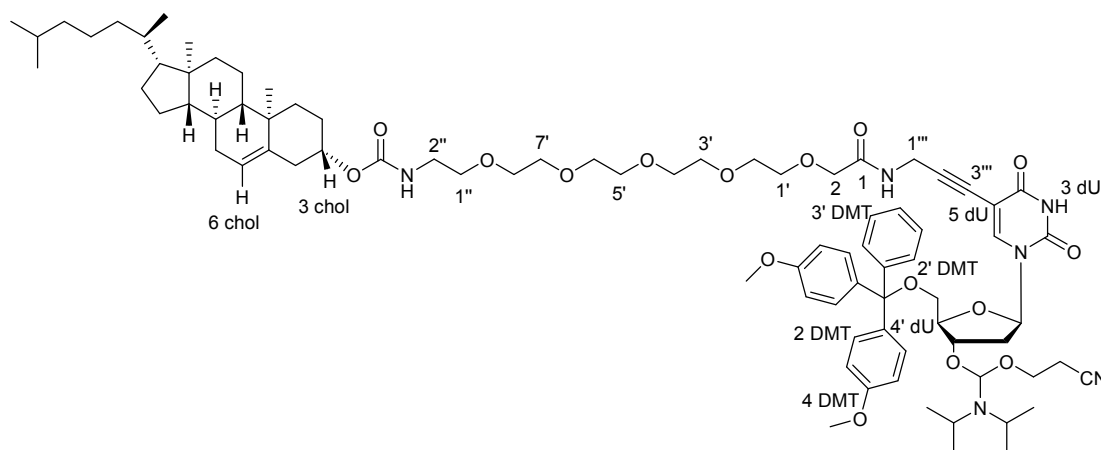
¹H NMR (400 MHz, CDCl₃): δ_H 8.04 (1H, s, **H⁶ dU**), 7.31-7.47 (9H, m, **H^{DMT}**), 7.20 (1H, br s, **NH**), 6.89 (4H, dd, *J* = 8.5, 1.5 Hz, **H³ DMT**), 6.32 (1H, dd, *J* = 6.5, 6.5 Hz,

$\text{H}^{1'} \text{ dU}$), 5.39 (1H, m, $\text{H}^6 \text{ Chol}$), 4.52-4.58 (2H, m, $\text{H}^3 \text{ Chol}$, $3' \text{ dU}$), 4.12 (1H, br m, $\text{H}^{4'} \text{ dU}$), 4.05 (2H, br m, $\text{H}^{1''}$), 3.99 (2H, s, H^2), 3.83 (6H, s, O-CH₃), 3.64-3.68 (18H, m, $\text{H}^{1'-8',2''}$), 3.56 (2H, br s, $\text{H}^{1''}$), 3.44 (1H, dd, $J = 10.5, 3.5 \text{ Hz}$, $\text{H}^{5'} \text{ dU}$), 3.38 (1H, dd, $J = 10.5, 3.5 \text{ Hz}$, $\text{H}^{5'} \text{ dU}$), 2.53 (1H, ddd, $J = 13.5, 5.5, 3.0 \text{ Hz}$, $\text{H}^{2'} \text{ dU}$), 2.30 (1H, m, $\text{H}^{2'} \text{ dU}$), 0.72-2.39 (br m, 43H).

^{13}C NMR (100 MHz, CDCl₃): δ_{C} 169.70 (2s, CO), 161.5 (s, CO), 158.8 (2s, $\text{C}^4 \text{ DMT}$), 149.2 (s, CO), 144.7 (s, $\text{C}^{1'} \text{ DMT}$), 143.1 (d, $\text{C}^6 \text{ dU}$), 140.0 (s, $\text{C}^5 \text{ Chol}$), 135.6 (2s, C^{DMT}), 130.1 (2d, C^{DMT}), 128.2 (4d, C^{DMT}), 128.0 (2d, C^{DMT}), 127.1 (s, C^{DMT}), 122.6 (d, $\text{C}^6 \text{ Chol}$), 113.5 (2d, C^{DMT}), 99.7 (s, $\text{C}^5 \text{ dU}$), 89.5 (s, $\text{C}^0 \text{ DMT}$), 87.2 (s, $\text{C}^{3''}$), 86.5 (d, $\text{C}^{4'} \text{ dU}$), 85.8 (d, $\text{C}^{1'} \text{ dU}$), 75.3 (d), 74.1 (s), 72.3 (d), 70.3-71.1 (9t, $\text{C}^{1'-8',2}$), 69.1 (t), 63.8 (t, $\text{C}^{5'} \text{ dU}$), 56.8 (d), 56.3 (d), 55.4 (2q, CO-CH₃), 51.1 (2d), 48.9 (t), 42.5 (s), 41.5 (t, $\text{C}^{2'} \text{ dU}$), 39.9 (t), 39.7 (t), 38.6 (t), 37.1 (t), 36.7 (t), 36.3 (t), 35.9 (d), 32.0 (d and t), 29.4 (d), 28.3 (t, $\text{C}^{1''}$), 28.3 (t), 28.1 (d), 24.4 (t), 24.0 (t), 23.0 (d), 22.7 (q), 21.4 (t), 21.2 (q), 19.5 (q), 18.9 (q), 12.0 (q).

LRMS [ES⁺]: m/z 1295.4 ([M + Na]⁺, 100%).

5'-O-(4,4-Dimethoxytrityl)-5-{2-O-(2''-cholesteryloxycarbonylamino-ethoxy)tetra(ethyleneoxy)]-N-prop-2'''-ynyl-acetamido}-2'-deoxyuridine 3'-O-(2-cyanoethyl)-N,N-diisopropyl phosphoramidite, **32**



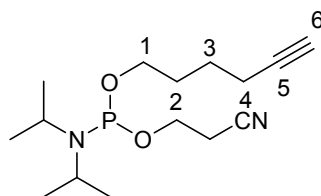
Under an argon atmosphere, 2-O-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphoramidite (137 μ L, 0.43 mmol) was added dropwise to a degassed solution of **31** (500 mg, 0.39 mmol) in CH_2Cl_2 (3 mL) with dry DIHT (33 mg, 0.20 mmol). After stirring at rt for 2 hours 20 min, the reaction mixture was diluted with degassed CH_2Cl_2 and washed with degassed saturated aqueous NaHCO_3 . The aqueous layer was washed with degassed CH_2Cl_2 and the combined organic layers were combined, dried (Na_2SO_4), filtered and the solvents removed *in vacuo*. Purification by silica gel column chromatography under argon pressure (EtOAc:hexane 4:1, 0.5 % Et_3N) gave **32** as diastereomeric mixture (*ca.* 1:1) pale yellow air- and acid-sensitive foam (378 mg, 66%).

R_f (CH_2Cl_2 :Acetone 3:2, C) 0.46; (CH_2Cl_2 :MeOH 19:1,C) 0.61; (EtOAc:MeOH: NH_3 5:1:1, C) 0.35.

¹H NMR (400 MHz, d_6 -DMSO): δ_{H} 11.62 (1H, br s, NH), 8.01 (1H, dd, $J = 12.1$, 6.2 Hz, NH), 7.93 (1H, s, $\text{H}^6 \text{dU}$), 7.23-7.41 (9H, m, H^{DMT}), 6.88 (4H, m, $\text{H}^3 \text{DMT}$), 6.08 (1H, dd, $J = 13.0$, 6.0 Hz, $\text{H}^{1'} \text{dU}$), 5.33 (1H, m, $\text{H}^6 \text{Chol}$), 4.51 (1H, m, $\text{H}^3 \text{Chol}$), 4.35 (1H, m, $\text{H}^{3'} \text{dU}$), 3.95-4.06 (5H, m, $\text{H}^{3''',4'} \text{dU}$, $1'''$), 3.89 (2H, s, H^2), 3.73 (6H, s, O-CH₃), 3.55-3.73 (4H, m, OCH₂CH₂CN and *i*Pr-CH), 3.48-3.55 (18H, m, $\text{H}^{1'-8', 2''}$), 3.56 (2H, br s, $\text{H}^{1''}$), 3.15 (2H, m, $\text{H}^{5'} \text{dU}$), 2.64 (1H, t, $J = 6.0$ Hz, CH₂CN), 2.75 (1H, t, $J = 6.0$ Hz, CH₂CN), 2.53 (1H, m, $\text{H}^{2'} \text{dU}$), 2.36 (1H, m, $\text{H}^{2'} \text{dU}$), 0.64-2.18 (br m, 57H).

³¹P NMR (162 MHz, d_6 -DMSO) δ_{P} 148.4, 148.8.

Hex-5-yn-1-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite, **39**



Under an argon atmosphere, 2-*O*-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (2.50 mL, 11.21 mmol) was added dropwise to a degassed solution of 5-hexyn-1-ol **38** (1.00 g, 10.2 mmol) in CH₂Cl₂ (10 mL) with DIPEA (4.44 mL, 25.5 mmol). After stirring at rt for 40 mins, the reaction mixture was diluted with degassed CH₂Cl₂ and washed with degassed saturated aqueous KCl, dried (Na₂SO₄), filtered and the solvent was removed *in vacuo*. Purification by column chromatography under argon pressure (Et₂O:hexane 1:1, 1 % pyridine) gave **39** as a colourless, air- and acid-sensitive sensitive oil (609 mg, 20%).

R_f (EtOAc :hexane 9:1, 0.5% Et₃N, A) 0.27; (EtOAc :hexane 9:1, 1 % pyridine, A) 0.23; (Et₂O :hexane 1:1, 1 % pyridine, A) 0.48.

¹H NMR (400 MHz, *d*₆-DMSO): δ_H 3.83 (2H, m, OCH₂CH₂CN) 3.76-3.53 (4H, m, *i*Pr-CH, **H**¹), 2.63 (2H, t, *J* = 6.5 Hz, CH₂CN), 2.23 (2H, td, *J* = 7.0, 2.5 Hz, **H**⁴), 1.94 (1H, t, *J* = 2.5 Hz, **H**⁶), 1.71 (2H, tt, *J* = 14.7, 6.5 Hz, **H**²), 1.61 (2H, tt, *J* = 14.7, 6.5 Hz, **H**³), 1.18 (3H, d, *J* = 3.5 Hz, *i*Pr-CH₃), 1.17 (3H, d, *J* = 3.5 Hz, *i*Pr-CH₃).

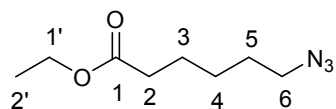
³¹P NMR (162 MHz, *d*₆-DMSO) δ_P 148.1.

LRMS [ES⁻]: *m/z* 299 ([M+H]⁺, 100%).

HRMS [ES⁺]: for C₁₅H₂₇N₂NaO₂P [M + Na]⁺: calcd 321.1702, found 321.1698.

Analytical results consistent with reported data.¹³⁵

Ethyl 6-azido-hexanoate, **41**



To a stirred suspension of sodium azide (1.47 g, 22.6 mmol) and potassium carbonate (2.34 g, 16.9 mmol) in anhydrous DMF (25 mL), ethyl-6-bromohexanoate **40** (2 mL, 11.3 mmol) was added under an argon atmosphere. The reaction mixture was stirred at 50 °C for 5 hours 45 min. The reaction mixture was then filtered through a celite plug and washed with toluene. Purification by silica gel column chromatography (hexane to hexane:EtOAc 95:5) gave **41** as a colourless oil (1.80 mg, 86%).

R_f (EtOAc:hexane, 1:4, A) 0.48.

ν_{max} (neat)/cm⁻¹ 2939 (w, CH), 2867 (w, CH), 2091 (s, N₃), 1731 (m, CO), 1457 (w, CH), 1373 (m), 1249 (m), 1179 (m), 1153 (m), 1096 (s, CO), 1029 (m).

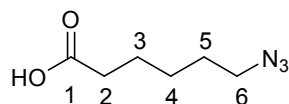
¹H NMR (400 MHz, CDCl₃): δ_H 4.12 (2H, q, *J* = 7.2 Hz, **H**^{1'}), 3.26 (2H, t, *J* = 7.2 Hz, **H**⁶), 2.30 (2H, t, *J* = 7.3 Hz, **H**²), 1.65 (2H, p, *J* = 7.7 Hz, **H**³), 1.61 (2H, p, *J* = 7.5 Hz, **H**⁵), 1.40 (2H, p, *J* = 7.5 Hz, **H**⁴), 1.25 (3H, t, *J* = 7.3 Hz, **H**^{2'}).

¹³C NMR (100 MHz, CDCl₃): δ_C 173.6 (s, **C**¹), 60.4 (t, **C**^{1'}), 51.4 (t, **C**⁶), 34.2 (t, **C**²), 28.7 (t, **C**⁵), 26.4 (t, **C**⁴), 24.6 (t, **C**³), 14.4 (q, **C**^{2'}).

LRMS [ES⁺]: *m/z* 208 ([M + Na]⁺, 100%).

HRMS [ES⁺]: for C₈H₁₅N₃NaO₂ [M + Na]⁺: calcd 208.1056, found 208.1058.

6-Azidohexanoic acid, **42**



Compound **41** (1.68 mg, 9.08 mmol) was dissolved in a mixture of H₂O and dioxane (2:1 v/v) (13.5 mL) to which 2 M aqueous sodium hydroxide (9.08 mL, 18.2 mmol) was added. The reaction mixture was stirred at rt for 1 hour. The crude mixture was purified by filtration through DOWEX resin (pyridinium form) (H₂O:MeOH, 1:1 v/v) to give **42** as a colourless oil (1.23 mg, 91%).

R_f (EtOAc:hexane, 1:4, A) 0.11.

ν_{max} (neat)/cm⁻¹ 2939 (m, CH), 2866 (m, CH), 2090 (s, N₃), 1702 (m, CO), 1412 (w, CH), 1252 (m), 1154 (w), 931 (m).

¹H NMR (400 MHz, CDCl₃): δ_H 10.38 (1H, br s, CO₂H), 3.28 (2H, t, *J* = 6.8 Hz, **H**⁶), 2.37 (2H, t, *J* = 7.5 Hz, **H**²), 1.65 (2H, p, *J* = 7.5 Hz, **H**³), 1.61 (2H, p, *J* = 7.3 Hz, **H**⁵), 1.40 (2H, p, *J* = 7.4 Hz, **H**⁴).

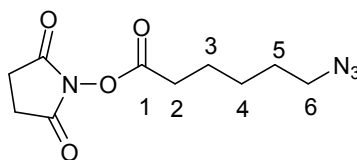
¹³C NMR (100 MHz, CDCl₃): δ_C 179.9 (s, **C**¹), 51.3 (t, **C**⁶), 34.0 (t, **C**²), 28.7 (t, **C**⁵), 26.3 (t, **C**⁴), 24.3 (t, **C**³).

LRMS [ES⁻]: *m/z* 156 ([M - H]⁻, 60%).

HRMS [ES⁺]: for C₆H₁₁N₃NaO₂ [M + Na]⁺: calcd 180.0743, found 180.0745.

Analytical results consistent with reported data.¹⁵⁸

N-Succinimidyl 6-azido-hexanoate, **43**



Compound **42** (1.20g, 7.62 mmol) was dissolved in CH₂Cl₂ (20 mL) to which *N*-hydroxysuccinimide (1.11 mg, 9.65 mmol) and EDC (2.04g, 10.7 mmol) were added. The reaction was complete after 6.5 hours. The reaction mixture was diluted with CH₂Cl₂, washed with water, dried (Na₂SO₄), filtered and the solvent was removed *in vacuo*. The crude mixture was purified by silica gel column chromatography (CH₂Cl₂ to CH₂Cl₂:MeOH 99:1) to give **43** as a yellow oil (1.00 g, 52%).

R_f (EtOAc: CH₂Cl₂, 3:7, A) 0.67, (CH₂Cl₂, A) 0.38, (CH₂Cl₂:MeOH 99:1, A) 0.56.

ν_{max} (neat)/cm⁻¹ 2944 (w, CH), 2869 (w, CH), 2094 (s, N₃), 1812 (w), 1782 (m), 1731 (m, CO), 1363 (m), 1200 (s), 1063 (s), 645 (m).

¹H NMR (400 MHz, CDCl₃): δ_H 3.28 (2H, t, *J* = 6.8 Hz, **H**⁶), 2.82 (4H, s, **H**^{Succ}), 2.62 (2H, t, *J* = 7.3 Hz, **H**²), 1.78 (2H, p, *J* = 7.5 Hz, **H**³), 1.63 (2H, p, *J* = 7.2 Hz, **H**⁵), 1.49 (2H, p, *J* = 7.2 Hz, **H**⁴).

¹³C NMR (100 MHz, CDCl₃): δ_C 169.2 (2s, CO^{Succ}), 168.5 (s, **C**¹), 51.2 (t, **C**⁶), 30.9 (t, **C**²), 28.5 (t, **C**⁵), 26.0 (t, **C**⁴), 25.7 (2t, **C**^{Succ}), 24.3 (t, **C**³).

LRMS [ES⁺]: *m/z* 277 ([M + Na]⁺, 50%).

HRMS [ES⁺]: for C₁₀H₁₄N₄NaO₄ [M + Na]⁺: calcd 277.0907, found 277.0904.

Analytical results consistent with reported data.¹⁵⁸

4.3 Biophysical studies

4.3.1 Preparation of synthetic oligonucleotides

The synthesised monomers were treated as follows: After purification by column chromatography the monomer was dissolved in DNA grade anhydrous acetonitrile except for the cholesterol monomers which were dissolved in anhydrous CH_2Cl_2 :acetonitrile (3:1 v/v) and filtered through a Millipore Millex®-FH syringe filter (0.45 μm , 25 mm). The solvent was then removed and the monomer was redissolved in freshly distilled CH_2Cl_2 . Aliquots corresponding to 100 mg were transferred to ABI monomer reagent bottles and dried in a desiccator overnight under high vacuum before being stored under a positive pressure of argon at $-20\text{ }^\circ\text{C}$ until use.

Standard A, G, C and T DNA phosphoramidites, solid supports and additional reagents including the C-7 aminoalkyl CPG were purchased from Link Technologies or Applied Biosystems Ltd. All oligonucleotides were synthesised on an Applied Biosystems 394 automated DNA/RNA synthesiser using a standard 0.2 μmole or 1.0 μmole phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and were $>98.0\%$. All β -cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use.

Cleavage of the oligonucleotide from the solid support was achieved by exposure to concentrated ammonia solution (30 min at room temp) in a sealed vial and deprotection achieved by heating for various time periods and monitoring by mass spectrometry.

Purification of oligonucleotides was carried out by reversed phase HPLC on a Gilson system using an ABI Aquapore column (C8), 8 mm \times 250 mm, pore size 300 Å. The system was controlled by Gilson 7.12 software and the following protocol was used: run time 30 mins, flow rate 4 mL per min, binary system, gradient: time in mins (% buffer B); 0 (0); 3(0); 5(20); 21 (100); 25(100); 27 (0); 30(0). Elution buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate with 35% acetonitrile pH 7.0. Elution of oligonucleotides was monitored by ultraviolet absorption at 295 nm.

After HPLC purification, oligonucleotides were desalted using a disposable NAP-10 Sephadex column purchased from GE Healthcare, aliquoted into an eppendorf tube and stored at -20°C . Purified oligonucleotides were then analysed by MALDI-TOF MS.

4.3.2 Oligonucleotides synthesised

ODNs	Sequences (5'-3')
	(Bz^{Chol} = monomer 8 , T^{Chol} = monomer 13 , H^{Chol} = monomer 23 , H = HEG, K = alkyne, Z = azide, NH_2 = amino C7, T^* = fluorescein dT, L = linear, L^* = linear and fluorescent, C = cyclic, C^* = cyclic and fluorescent)
ODN-1	Bz^{Chol} -TTTTTTTTTTTT
ODN-2	Bz^{Chol} -TTT
ODN-3	CGCAT $^{\text{Chol}}$ AT $^{\text{Chol}}$ CGC
ODN-4	GCGAT $^{\text{Chol}}$ AT $^{\text{Chol}}$ GCG
ODN-5	ACCAAGGTGCTAT $^{\text{Chol}}$ CGTCG
ODN-6	CGACGAT $^{\text{Chol}}$ AGCACCTTGGT
ODN-7	H^{Chol} -TGACTGTGACAC
L-ODN-8	K-TTTTTTTTTTGCACCAGAATTCATCACGGAGTTTTTTTTTT-NH ₂
L-ODN-9	K-HHHCTCCGTGATGAATTCTGGTGCHHH-NH ₂
L-ODN-10	K-TTTTTTTTTTGCACCAGAATTCATCACGGAGTTTTTTTTTT-Z
L-ODN-11	K-HHHCTCCGTGATGAATTCTGGTGCHHH-Z
C-ODN-10	Circular ssDNA form of ODN-10
C-ODN-11	Circular ssDNA form of ODN-11
L-ODN-12	K-TTTTTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTT-NH ₂
L^* -ODN-13	K-GCGT*HHHGATGAATTCTGGHHHTGCG-NH ₂
L-ODN-14	K-TTTTTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTT-Z
L^* -ODN-15	K-GCGT*HHHGATGAATTCTGGHHHTGCG-Z
C-ODN-14	Circular ssDNA form of ODN-14
C^* -ODN-15	Circular ssDNA form of ODN-15
ODN-16	CGCCGC
L^* -ODN-17	K-TTT*TTTTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTT-NH ₂

L-ODN-18	K-GCGTHHHGATGAATTCTGGHHHTGCG-NH ₂
L*-ODN-19	K-TTT*TTTTTTTTTTTTCCAGAATTCATCTTTTTTTTTTTTTTT-Z
L-ODN-20	K-GCGTHHHGATGAATTCTGGHHHTGCG-Z
C*-ODN-19	Circular ssDNA form of ODN-19
C-ODN-20	Circular ssDNA form of ODN-20
Catenane	Catenane form of C-ODN-19 and L-ODN-20

4.3.3 Oligonucleotide characterisation

ODNs	Expected mass	Observed mass
ODN-1	4421.5	4420.0 ^{MALDI}
ODN-2	1683.8	1680.7 ^{MALDI}
ODN-3	4287.0	4286.9 ^{MALDI}
ODN-4	4287.0	4366.6 ^{MALDI}
ODN-5	6149.6	6149.6 ^{MALDI}
ODN-6	6149.6	6150.7 ^{MALDI}
ODN-7	4235.5	4237.8 ^{MALDI}
L-ODN-8	12876	12883.9 ^{MALDI}
L-ODN-9	8860	8867 ^{ES}
L-ODN-10	13015	13223.6 ^{MALDI}
L-ODN-11	8999	9006 ^{ES}
C-ODN-10	13015	13017.9 ^{MALDI}
C-ODN-11	9001	9002.8 ^{MALDI}
L-ODN-14	13222	13226.6 ^{MALDI}
L*-ODN-15	9289	9293.8 ^{MALDI}
C-ODN-14	13222	nd
C*-ODN-15	9289	nd
L*-ODN-19	13757	13761.7 ^{MALDI}
L-ODN-20	8777	8779.7 ^{MALDI}
C*-ODN-19	13735	nd
C-ODN-20	8777	8778.7 ^{MALDI}
ODN-27	22511	nd

nd = non determined

4.3.4 UV melting studies

UV melting experiments were performed on a Varian Cary 400 Scan UV-Visible spectrophotometer, in Hellma® SUPRASIL synthetic quartz, 10 mm pathlength cuvettes, monitoring at 260 nm, at a 1 μ M concentration in a volume of 1.5 mL, or at 0.25 or 2 μ M concentration in a volume of 1.2 mL. Buffers were prepared according to literature procedure¹⁵⁹ and kept at 5 °C until needed. The buffers were allowed to equilibrate to rt and the pH was checked before use.

Samples were prepared as follows: The complementary strands were mixed in a 1:1 ratio in 2 mL eppendorf tubes then lyophilised before resuspending in 1.5 or 1.2 mL of the appropriate buffer solution pH 7.0 (10 mM sodium phosphate, 1 mM Na₂EDTA and 200 mM, 500 mM or 1 M NaCl) to give a final concentration of 0.25, 1 or 2 μ M of each strand. The samples were then filtered into the cuvettes through Kinesis regenerated cellulose 13 mm, 0.45 μ M syringe filters.

For experiments in chapter 1, the following procedure was followed:

Following an initial rapid heat (20 to 80 °C at 10 °C/min) and cool cycle (80 to 20 °C at 1 °C/min) to ensure uniform annealing of the strands, the UV-melting curves were recorded for 3 consecutive heat and cool cycles (20 to 80 to 20 °C at 1 °C/min).

For experiments in chapter 2, the following procedure was followed;

Following an initial rapid heat (20 to 80 °C at 10 °C/min), the UV-melting curves were recorded for 3 cool and heat cycles: 80 to 20 °C at 0.5 °C/min with a hold time of 20 min to favour the duplex formation and to 80 °C at 0.5 °C/min with a hold time of 2 min.

T_m values were calculated for each experiment using Cary WinUV Thermal Application software.

4.3.5 Fluorescence melting studies

Fluorescence melting experiments were performed using a Roche LightCycler[®] 1.5 instrument, in LightCycler glass capillaries (20 μ L volume). The Lightcycler has one excitation source (488 nm) and the changes in fluorescence emission were measured at 520 nm.

The capillaries contained 0.5 μ M of each strand using the fluorescent intercalator SYBR Green I (intercalator:total volume 1:10 v/v), in 10 mM sodium phosphate buffer (pH 7.0) with 1 mM Na₂EDTA and the appropriate concentration of NaCl (200 mM, 400 mM or 600 mM NaCl). The complexes were denaturated by heating to 95 °C at a rate of 20 °C/sec and were maintained at this temperature for 5 mins before cooling to 30 °C at 0.1 °C/sec. Samples were then held at 30 °C for 5 mins before melting again by heating to 95 °C at 0.1 °C/sec. Data was collected during both annealing and melting steps and compared. Each reaction was performed in duplicate. T_m values were determined from the first derivatives of the melting profile using the Roche LightCycler[®] software version 3.5.

4.3.6 Click chemistry experimental

Water soluble *tris*-(hydroxypropyltriazolylmethyl)amine ligand was synthesised by the reported literature method.⁸⁸

4.3.6.1 Synthesis of azidohexanamide-labeled oligonucleotides

The following procedure was used to incorporate the azide at the 3'-end of C7-aminoalkyl ODN. A 1.00 μ mol synthesis of the ODN was incubated for 4 hours in 0.5 M Na₂CO₃/NaHCO₃ buffer (pH 8.75, 70 μ L) with *N*-succinimidyl 6-azido-hexanoate **43** (3 μ L) and DMSO (100 μ L). The crude ODN was desalted by NAP-10 gel-filtration according to the manufacturer's instructions (GE Healthcare) and purified by reversed-phase HPLC. The purified ODNs were characterised by ES-MS: L-ODN-11 ($[M-H]^-$ calcd, 8999; found, 9006) and MALDI-TOF MS: L-ODN-10 ($[M+H]^+$ calcd, 13015;

found, 13223.6); L-ODN-14: ($[M+H]^+$ calcd, 13022; found, 13226.6); L*-ODN-15 ($[M+H]^+$ calcd, 9289 ; found, 9294); L*-ODN-19 ($[M+Na]^+$ calcd, 13757 ; found, 13761.7) and L-ODN-20 ($[M+H]^+$ calcd, 877; found, 8779.7).

4.3.6.2 Click cyclisation of ssDNA

Excess Cu[II] (10, 20 and 50 eq) was used in order to obtain high yields of cyclic ODNs whilst keeping the ratio of Cu[II]/sodium ascorbate/ligand at 1/10/7. The following method produced the best yields:

To a solution of *tris*-(hydroxypropyltriazolylmethyl)amine ligand⁸⁸ (1.22 mg, 2.80 μ mol in 590 μ L 0.2 M NaCl) under an argon atmosphere was added 0.4 M NaCl (955 μ L), sodium ascorbate (8 μ L, 4 μ mol) followed by CuSO₄.H₂O (4 μ L, 0.40 μ mol). The azido-labelled alkyne ODN (943 μ L, 20 nmol) was added and the solution was incubated at rt for 2 hours. The reaction mixture was desalted by NAP-25 gel-filtration according to the manufacturer's instructions (GE Healthcare) and purified by reversed-phase HPLC. The purified ODNs were characterised by MALDI-TOF MS C-ODN-10 ($[M+H]^+$ calcd, 13015; found, 13017.9); C-ODN-11 ($[M+H]^+$ calcd, 9001; found, 9002.8); C-ODN-20 ($[M+H]^+$ calcd, 8777; found, 8778.7). No MS analysis was possible for very high molecular weight ODNs: C-ODN-14, C*-ODN-15 and C*-ODN-19.

4.3.6.3 Preparation of dsDNA catenane

Excess Cu[II] (10, 20 and 50 eq) was used in order to give high yields of cyclic ODNs whilst keeping the ratio of Cu[II]/sodium ascorbate/ligand at 1/10/7. The following method gave the best yields:

A solution of cyclised ssDNA (7 nmol) and its complementary azido-labelled alkyne ODN (7 nmol) in 5 mL of 0.2 M NaCl was prepared and heated to 80 °C then cooled down slowly to rt for 2 hours. ODN-16 (14 nmol) and 0.4 M NaCl (21.0 μ L) were added and the reaction was incubated at rt for 30 min (solution 1). To a solution of *tris*-(hydroxypropyltriazolylmethyl)amine ligand (1.06 mg, 2.45 μ mol in 0.2 M NaCl,

1937.0 μL) under an argon atmosphere was added 0.4 M NaCl (10.5 μL), sodium ascorbate (7.0 μL , 3.50 μmol) followed by $\text{CuSO}_4 \cdot \text{H}_2\text{O}$ (3.5 μL , 350 nmol) (solution 2). Solution 1 was added to the above prepared catalyst solution (solution 2) and the reaction was incubated at rt for 2 hours. The reaction mixture was desalted by NAP-10 gel-filtration according to the manufacturer's instructions (GE Healthcare) and lyophilized, then analysed and purified by denaturing 20% PAGE.

4.3.6.4 Denaturing PAGE gel electrophoresis analysis and purification

The ODNs were analysed or purified on 20% polyacrylamide/7M urea gels (up to 20 A_{260} of crude DNA per gel) at a constant power of 20 W, for 3-4 hours, using 0.09M tris-borate-EDTA buffer (pH 8.0). Gels were visualised on a fluorescent TLC plate and illuminated with a UV lamp (254 nm) or trans-illuminated with a UV light when a fluorescein dT modification was present. For purification of the catenane, the bands were excised, and the gel pieces were crushed and incubated with 1 mL of sterile water at 37°C for 16 hours. The tubes were then agitated and centrifuged and the supernatants were desalted using NAP-10 gel-filtration according to the manufacturer's instructions (GE Healthcare).

4.3.6.5 Ferguson plot analysis

L-ODN-14, L*-ODN-15, C-ODN-14, C*-ODN-15 and the catenane were subjected to analysis using a 5 point Ferguson-plot. The 5 different acrylamide concentrations used for this study were 8, 11, 14, 17 and 20%. The protocol involved a run of 2, 2.25, 2.33, 2.5 and 3.33 hours respectively, at rt, with a constant power of 20 W. Using the linear plots in Figure 1, the experimentally obtained retardation coefficients (i.e negative slopes) for L-ODN-14, L*-ODN-15, C-ODN-14, C*-ODN-15 and the catenane were found to be 0.036, 0.020, 0.020, 0.026 and 0.058 respectively.

4.3.6.6 Native PAGE gel electrophoresis analysis

The ODNs were analysed on 14% polyacrylamide/7M urea gels (up to 20 A₂₆₀ of crude DNA per gel) at a constant voltage of 130 V, for 11 hours, using 0.09M tris-borate buffer (pH 8.0). The gel was visualised on a fluorescent TLC plate and illuminated with a UV lamp (254 nm) or trans-illuminated with a UV light when a fluorescein dT modification was present.

4.3.6.7 EcoR1 digestion

ODNs (3 µg) containing an EcoR1 restriction site were digested with or without 18 U EcoR1 (1.5 µL) (Promega) at 37 °C for 3.15 hours. The reaction mixture also contained 6 µL EcoR1 buffer (× 10, Buffer H), 6 µL BSA (1 µg/mL) and 46.5 µL H₂O. Products were subsequently separated by 20% denaturing PAGE electrophoresis.

References

1. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. *Nucleic Acids in Chemistry and Biology*; 3rd Ed.; RSC Publishing, **2006**.
2. Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 737-738.
3. Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740-741.
4. Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738-740.
5. Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 964-967.
6. Drexler. *Engines of Creation: The Coming Era of Nanotechnology*; Anchor Books, **1986**.
7. Chen, Y.; Wang, M. S.; Mao, C. D. *Angew. Chem., Int. Ed.* **2004**, *43*, 3554-3557.
8. Shin, J. S.; Pierce, N. A. *J. Am. Chem. Soc.* **2004**, *126*, 10834-10835.
9. Bath, J.; Green, S. J.; Turberfield, A. J. *Angew. Chem., Int. Ed.* **2005**, *44*, 4358-4361.
10. Tian, Y.; He, Y.; Chen, Y.; Yin, P.; Mao, C. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4355-4358.
11. Adleman, L. M. *Science* **1994**, *266*, 1021-1024.
12. Rothmund, P. W. K.; Papadakis, N.; Winfree, E. *PLoS. Biol.* **2004**, *2*, 2041-2053.
13. Seeman, N. C. *Nature* **2003**, *421*, 427-431.
14. Caruthers, M. H. *Science* **1985**, *230*, 281-285.
15. Pancoska, P.; Moravek, Z.; Moll, U. M. *Nucleic Acids Res.* **2004**, *32*, 4630-4645.
16. Feldkamp, U.; Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2006**, *45*, 1856-1876.
17. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 10441-10488.
18. Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. *Science* **2003**, *301*, 1882-1884.
19. Park, S. H.; Yin, P.; Liu, Y.; Reif, J. H.; LaBean, T. H.; Yan, H. *Nano Lett.* **2005**, *5*, 729-733.
20. Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2001**, *40*, 4128-4158.
21. Gothelf, K. V.; LaBean, T. H. *Org. Biomol. Chem.* **2005**, *3*, 4023-4037.
22. Rothmund, P. W. K. *Nature* **2006**, *440*, 297-302.
23. Yan, H.; LaBean, T. H.; Feng, L. P.; Reif, J. H. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 8103-8108.
24. Simmel, F. C. *Angew. Chem., Int. Ed.* **2008**, *47*, 5884-5887.
25. Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237-247.
26. Seeman, N. C.; Wang, H.; Yang, X. P.; Liu, F. R.; Mao, C. D.; Sun, W. Q.; Wenzler, L.; Shen, Z. Y.; Sha, R. J.; Yan, H.; Wong, M. H.; Sa-Ardyen, P.; Liu, B.; Qiu, H. X.; Li, X. J.; Qi, J.; Du, S. M.; Zhang, Y. W.; Mueller, J. E.; Fu, T. J.; Wang, Y. L.; Chen, J. H. *Nanotechnology* **1998**, *9*, 257-273.
27. Fu, T. J.; Seeman, N. C. *Biochemistry* **1993**, *32*, 3211-3220.
28. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539-544.
29. LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F. R.; Winfree, E.; Reif, J. H.; Seeman, N. C. *J. Am. Chem. Soc.* **2000**, *122*, 1848-1860.

30. Shen, Z. Y.; Yan, H.; Wang, T.; Seeman, N. C. *J. Am. Chem. Soc.* **2004**, *126*, 1666-1674.
31. Yang, X. P.; Wenzler, L. A.; Qi, J.; Li, X. J.; Seeman, N. C. *J. Am. Chem. Soc.* **1998**, *120*, 9779-9786.
32. Zhang, X. P.; Yan, H.; Shen, Z. Y.; Seeman, N. C. *J. Am. Chem. Soc.* **2002**, *124*, 12940-12941.
33. Liu, Y.; Lin, C. X.; Li, H. Y.; Yan, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 4333-4338.
34. Liu, D.; Park, S. H.; Reif, J. H.; LaBean, T. H. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 717-722.
35. Mao, C. D.; Sun, W. Q.; Seeman, N. C. *J. Am. Chem. Soc.* **1999**, *121*, 5437-5443.
36. Liu, D.; Wang, M. S.; Deng, Z. X.; Walulu, R.; Mao, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 2324-2325.
37. Ding, B. Q.; Sha, R. J.; Seeman, N. C. *J. Am. Chem. Soc.* **2004**, *126*, 10230-10231.
38. Chelyapov, N.; Brun, Y.; Gopalkrishnan, M.; Reishus, D.; Shaw, B.; Adleman, L. *J. Am. Chem. Soc.* **2004**, *126*, 13924-13925.
39. Park, S. H.; Yan, H.; Reif, J. H.; LaBean, T. H.; Finkelstein, G. *Nanotechnology* **2004**, *15*, S525-S527.
40. He, Y.; Tian, Y.; Chen, Y.; Deng, Z. X.; Ribbe, A. E.; Mao, C. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 6694-6696.
41. Chen, J. H.; Seeman, N. C. *Nature* **1991**, *350*, 631-633.
42. Zhang, Y. W.; Seeman, N. C. *J. Am. Chem. Soc.* **1994**, *116*, 1661-1669.
43. Goodman, R. P.; Berry, R. M.; Turberfield, A. J. *Chem. Commun.* **2004**, 1372-1373.
44. Pugh, A. *An Introduction to Tensegrity*; University of California Press, **1976**; pp. 121.
45. Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. *Science* **2005**, *310*, 1661-1665.
46. Zhang, C.; Su, M.; He, Y.; Zhao, X.; Fang, P. A.; Ribbe, A. E.; Jiang, W.; Mao, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 10665-10669.
47. von Kiedrowski, G.; Eckardt, L. H.; Naumann, K.; Pankau, W. M.; Reimold, M.; Rein, M. *Pure Appl. Chem.* **2003**, *75*, 609-619.
48. Zimmermann, J.; Cebulla, M. R. J.; Monninghoff, S.; von Kiedrowski, G. *Angew. Chem., Int. Ed.* **2008**, *47*, 3626-3630.
49. Shih, W. M.; Quispe, J. D.; Joyce, G. F. *Nature* **2004**, *427*, 618-621.
50. Mao, C. D.; Sun, W. Q.; Seeman, N. C. *Nature* **1997**, *386*, 137-138.
51. He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. D. *Nature* **2008**, *452*, 198-U41.
52. Mao, C. D.; Sun, W. Q.; Shen, Z. Y.; Seeman, N. C. *Nature* **1999**, *397*, 144-146.
53. Chen, Y.; Lee, S. H.; Mao, C. *Angew. Chem., Int. Ed.* **2004**, *43*, 5335-5338.
54. Bruciale, M.; Zuccheri, G.; Samori, B. *Org. Biomol. Chem.* **2005**, *3*, 575-577.
55. Yurke, B.; Turberfield, A. J.; Mills, A. P.; Simmel, F. C.; Neumann, J. L. *Nature* **2000**, *406*, 605-608.
56. Simmel, F. C.; Yurke, B. *Phys. Rev. E* **2001**, *6304*.
57. Simmel, F. C.; Yurke, B. *Appl. Phys. Lett.* **2002**, *80*, 883-885.
58. Dittmer, W. U.; Reuter, A.; Simmel, F. C. *Angew. Chem., Int. Ed.* **2004**, *43*, 3550-3553.
59. Sherman, W. B.; Seeman, N. C. *Nano Lett.* **2004**, *4*, 1203-1207.

60. Bath, J.; Turberfield, A. J. *Nat. Nanotechnol.* **2007**, *2*, 275-284.
61. Niemeyer, C. M.; Sano, T.; Smith, C. L.; Cantor, C. R. *Nucleic Acids Res.* **1994**, *22*, 5530-5539.
62. Niemeyer, C. M.; Burger, W.; Peplies, J. *Angew. Chem., Int. Ed.* **1998**, *37*, 2265-2268.
63. Niemeyer, C. M. *Trends Biotechnol.* **2002**, *20*, 395-401.
64. Park, S. J.; Lazarides, A. A.; Mirkin, C. A.; Letsinger, R. L. *Angew. Chem., Int. Ed.* **2001**, *40*, 2909-2912.
65. Malo, J.; Mitchell, J. C.; Venien-Bryan, C.; Harris, J. R.; Wille, H.; Sherratt, D. J.; Turberfield, A. J. *Angew. Chem., Int. Ed.* **2005**, *44*, 3057-3061.
66. Lin, C. X.; Liu, Y.; Yan, H. *Nano Lett.* **2007**, *7*, 507-512.
67. Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. *Nature* **1996**, *382*, 607-609.
68. Alivisatos, A. P.; Johnsson, K. P.; Peng, X. G.; Wilson, T. E.; Loweth, C. J.; Bruchez, M. P.; Schultz, P. G. *Nature* **1996**, *382*, 609-611.
69. Li, H. Y.; Park, S. H.; Reif, J. H.; LaBean, T. H.; Yan, H. *J. Am. Chem. Soc.* **2004**, *126*, 418-419.
70. Aldaye, F. A.; Sleiman, H. F. *Angew. Chem., Int. Ed.* **2006**, *45*, 2204-2209.
71. Winfree, E. On the computational power of DNA annealing and ligation. In *DNA Based Computing*; Lipton E.J; Baum E.B. Eds., 1996; pp. 199-219.
72. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004-+.
73. Fioravanti, S.; Pellacani, L.; Tardella, P. A.; Morreale, A.; Del Signore, G. *J. Comb. Chem.* **2006**, *8*, 808-811.
74. Trost, B. M. *Science* **1991**, *254*, 1471-1477.
75. Huisgen, R. *Angew. Chem., Int. Ed.* **1963**, *2*, 565-632.
76. Huisgen, R. *Pure Appl. Chem.* **1989**, *61*, 613-628.
77. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010.
78. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596-+.
79. Gothelf, K. V.; Jorgensen, K. A. *Chem. Rev.* **1998**, *98*, 863-909.
80. Mulzer, J.; Altenbach, H. J.; Braun, M.; Krohn, K.; Reissig, H. U. *Org. Synt. High.*; mbH VCH Verlagsgesellschaft Ed., **1991**.
81. Katritzky, A. R.; Singh, S. K. *J. Org. Chem.* **2002**, *67*, 9077-9079.
82. Harju, K.; Vahermo, M.; Mutikainen, I.; Yli-Kauhaluoma, J. *J. Comb. Chem.* **2003**, *5*, 826-833.
83. Akritopoulou-Zanze, I.; Gracias, V.; Djuric, S. W. *Tetrahedron Lett.* **2004**, *45*, 8439-8441.
84. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.
85. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. *J. Am. Chem. Soc.* **2005**, *127*, 210-216.
86. Wu, P.; Fokin, V. V. *Aldrichimica Acta* **2007**, *40*, 7-17.
87. Wu, P.; Feldman, A. K.; Nugent, A. K.; Hawker, C. J.; Scheel, A.; Voit, B.; Pyun, J.; Frechet, J. M. J.; Sharpless, K. B.; Fokin, V. V. *Angew. Chem., Int. Ed.* **2004**, *43*, 3928-3932.
88. Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.* **2004**, *6*, 2853-2855.
89. Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *European Journal of Organic Chemistry* **2005**, 51-68.
90. Kolb, H. C.; Sharpless, K. B. *Drug Discov. Today* **2003**, *8*, 1128-1137.

91. Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36*, 1249-1262.
92. Hawker, C. J.; Fokin, V. V.; Finn, M. G.; Sharpless, K. B. *Aust. J. Chem.* **2007**, *60*, 381-383.
93. Luo, S. Z.; Xu, H.; Mi, X. L.; Li, J. Y.; Zheng, X. X.; Cheng, J. P. *J. Org. Chem.* **2006**, *71*, 9244-9247.
94. Alza, E.; Cambeiro, X. C.; Jimeno, C.; Pericas, M. A. *Org. Lett.* **2007**, *9*, 3717-3720.
95. Font, D.; Jimeno, C.; Pericas, M. A. *Org. Lett.* **2006**, *8*, 4653-4655.
96. Kolb, H. C.; Kanamarlapudi, R. C.; Richardson, P. F.; Khan, G. Modified safe and efficient process for the environmentally friendly synthesis of imidoesters; Lexicon Pharmaceuticals Inc.
97. Kolb, H. C.; Cavallaro, C.; Shi, Z.-C.; Gontcharov, A.; Wang, Z.-M.; Richardson, P. F.; Kanamarlapudi, R. C. One step synthesis of 1,2,3-triazole carboxylic acids; Lexicon Pharmaceuticals Inc. New Brunswick NJ USA.
98. Kolb, H. C.; McGeehan, G.; Shi, Z.-C.; Kolla, L. R.; Cavallaro, C. PPAR-gamma agonists as agents for the treatment of type II diabetes; Lexicon Pharmaceuticals Inc. Princeton NJ USA.
99. Lee, L. V.; Mitchell, M. L.; Huang, S. J.; Fokin, V. V.; Sharpless, K. B.; Wong, C. H. *J. Am. Chem. Soc.* **2003**, *125*, 9588-9589.
100. Turner, R. A.; Oliver, A. G.; Lokey, R. S. *Org. Lett.* **2007**, *9*, 5011-5014.
101. Rodriguez-Borges, J. E.; Goncalves, S.; Do Vale, M. L.; Garcia-Mera, X.; Coelho, A.; Sotelo, E. *J. Comb. Chem.* **2008**, *10*, 372-375.
102. Horne, W. S.; Yadav, M. K.; Stout, C. D.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2004**, *126*, 15366-15367.
103. Whiting, M.; Tripp, J. C.; Lin, Y. C.; Lindstrom, W.; Olson, A. J.; Elder, J. H.; Sharpless, K. B.; Fokin, V. V. *J. Med. Chem.* **2006**, *49*, 7697-7710.
104. Gopi, H. N.; Tirupula, K. C.; Baxter, S.; Ajith, S.; Chaiken, I. M. *ChemMedChem* **2006**, *1*, 54-+.
105. Genin, M. J.; Allwine, D. A.; Anderson, D. J.; Barbachyn, M. R.; Emmert, D. E.; Garmon, S. A.; Graber, D. R.; Grega, K. C.; Hester, J. B.; Hutchinson, D. K.; Morris, J.; Reischer, R. J.; Ford, C. W.; Zurenko, G. E.; Hamel, J. C.; Schaadt, R. D.; Stapert, D.; Yagi, B. H. *J. Med. Chem.* **2000**, *43*, 953-970.
106. Vatmurge, N. S.; Hazra, B. G.; Pore, V. S.; Shirazi, F.; Chavan, P. S.; Deshpande, M. V. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2043-2047.
107. Kim, D. K.; Kim, J.; Park, H. J. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2401-2405.
108. Dabak, K.; Sezer, O.; Akar, A.; Anac, O. *Eur. J. Med. Chem.* **2003**, *38*, 215-218.
109. Costa, M. S.; Boechat, N.; Rangel, E. A.; Da Silva, F. D.; de Souza, A. M. T.; Rodrigues, C. R.; Castro, H. C.; Junior, I. N.; Lourenco, M. C. S.; Wardell, S.; Ferreira, V. F. *Bioorg. Med. Chem.* **2006**, *14*, 8644-8653.
110. Punna, S.; Kuzelka, J.; Wang, Q.; Finn, M. G. *Angew. Chem., Int. Ed.* **2005**, *44*, 2215-2220.
111. Bock, V. D.; Speijer, D.; Hiemstra, H.; van Maarseveen, J. H. *Org. Biomol. Chem.* **2007**, *5*, 971-975.
112. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.
113. Sen Gupta, S.; Kuzelka, J.; Singh, P.; Lewis, W. G.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2005**, *16*, 1572-1579.
114. Gouin, S. G.; Bultel, L.; Falentin, C.; Kovensky, J. *Eur. J. Org. Chem.* **2007**, 1160-1167.

115. Baron, A.; Bleriot, Y.; Sollogoub, M.; Vauzeilles, B. *Org. Biomol. Chem.* **2008**, *6*, 1898-1901.
116. Kuijpers, B. H. M.; Groothuys, S.; Keereweer, A. R.; Quaedflieg, P.; Blaauw, R. H.; van Delft, F. L.; Rutjes, F. *Org. Lett.* **2004**, *6*, 3123-3126.
117. Hotha, S.; Kashyap, S. *J. Org. Chem.* **2006**, *71*, 364-367.
118. Wan, Q.; Chen, J. H.; Chen, G.; Danishefsky, S. J. *J. Org. Chem.* **2006**, *71*, 8244-8249.
119. Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686-4687.
120. Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535-546.
121. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.
122. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644-648.
123. Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 5463-5467.
124. Seo, T. S.; Li, Z. M.; Ruparel, H.; Ju, J. Y. *J. Org. Chem.* **2003**, *68*, 609-612.
125. Burley, G. A.; Gierlich, J.; Mofid, M. R.; Nir, H.; Tal, S.; Eichen, Y.; Carell, T. *J. Am. Chem. Soc.* **2006**, *128*, 1398-1399.
126. Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. *Org. Lett.* **2006**, *8*, 3639-3642.
127. Devaraj, N. K.; Miller, G. P.; Ebina, W.; Kakaradov, B.; Collman, J. P.; Kool, E. T.; Chidsey, C. E. D. *J. Am. Chem. Soc.* **2005**, *127*, 8600-8601.
128. Humenik, M.; Huang, Y. W.; Wang, Y. R.; Sprinzl, M. *Chembiochem* **2007**, *8*, 1103-1106.
129. Seela, F.; Sirivolu, V. R. *Org. Biomol. Chem.* **2008**, *6*, 1674-1687.
130. Weller, R. L.; Rajski, S. R. *Org. Lett.* **2005**, *7*, 2141-2144.
131. Bouillon, C.; Meyer, A.; Vidal, S.; Jochum, A.; Chevolot, Y.; Cloarec, J. P.; Praly, J. P.; Vasseur, J. J.; Morvan, F. *J. Org. Chem.* **2006**, *71*, 4700-4702.
132. Geci, I.; Filichev, V. V.; Pedersen, E. B. *Chem.-Eur. J.* **2007**, *13*, 6379-6386.
133. Kocalka, P.; El-Sagheer, A. H.; Brown, T. *Chembiochem* **2008**, *9*, 1280-1285.
134. Kumar, R.; El-Sagheer, A.; Tumpene, J.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T. *J. Am. Chem. Soc.* **2007**, *129*, 6859-6864.
135. El-Sagheer, A. H.; Kumar, R.; Findlow, S.; Werner, J. M.; Lane, A. N.; Brown, T. *Chembiochem* **2008**, *9*, 50-52.
136. Stellwagen, N. C.; Kraev, A.; Snowdon, R. J.; Langsdorf, A.; Coulthard, M. B.; Johnson, W. M.; Ashton, F. E. *Nucleic Acid Electrophoresis*; Tietz Dietmar Ed.; Springer, **1998**; pp. 337.
137. Ferguson, K. A. *Metabolism* **1964**, *13*.
138. Gore, M. G. *Spectrophotometry and Spectrofluorimetry - A practical approach*; Oxford University Press, **2000**; pp. 368.
139. Perouzel, E.; Jorgensen, M. R.; Keller, M.; Miller, A. D. *Bioconjugate Chem.* **2003**, *14*, 884-898.
140. Langley, G. J.; Herniman, J. M.; Davies, N. L.; Brown, T. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1717-1723.
141. Ahmadian, M.; Bergstrom, D. E.; Zhang, P. *Nucleic Acids Res.* **1998**, *26*, 3127-3135.
142. Osborne, S. D.; Powers, V. E. C.; Rusling, D. A.; Lack, O.; Fox, K. R.; Brown, T. *Nucleic Acids Res.* **2004**, *32*, 4439-4447.

143. Ranasinghe, R. T.; Rusling, D. A.; Powers, V. E. C.; Fox, K. R.; Brown, T. *Chem. Commun.* **2005**, 2555-2557.
144. McCairn, M. C.; Tonge, S. R.; Sutherland, A. J. *J. Org. Chem.* **2002**, 67, 4847-4855.
145. Pillai, V. N. R.; Mutter, M.; Bayer, E.; Gatfield, I. *J. Org. Chem.* **1980**, 45, 5364-5370.
146. Masson, C.; Scherman, D.; Bessodes, M. *J. Polym. Sci., Part A: Polym. Chem.* **2001**, 39, 4022-4024.
147. Robins, M. J.; Barr, P. J. *J. Org. Chem.* **1983**, 48, 1854-1862.
148. Cruickshank, K. A.; Stockwell, D. L. *Tetrahedron Lett.* **1988**, 29, 5221-5224.
149. Montalbetti, C.; Falque, V. *Tetrahedron* **2005**, 61, 10827-10852.
150. Blommers, M. J. J.; Natt, F.; Jahnke, W.; Cuenoud, B. *Biochemistry* **1998**, 37, 17714-17725.
151. Seela, F.; Sirivolu, V. R. *Helv. Chim. Acta* **2007**, 90, 535-552.
152. Lee, I. K.; Ahn, J. D.; Kim, H. S.; Park, J. Y.; Lee, K. U. *Curr. Drug Targets* **2003**, 4, 619-623.
153. Kool, E. T. *Annu. Rev. Biophys. Biomolec. Struct.* **1996**, 25, 1-28.
154. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, 22, 1859-1862.
155. Saito, G.; Swanson, J. A.; Lee, K. D. *Adv. Drug Deliv. Rev.* **2003**, 55, 199-215.
156. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, 62, 7512-7515.
157. Schlatterer, J. C.; Jaschke, A. *Biochem. Biophys. Res. Commun.* **2006**, 344, 887-892.
158. Grandjean, C.; Boutonnier, A.; Guerreiro, C.; Fournier, J. M.; Mulard, L. A. *J. Org. Chem.* **2005**, 70, 7123-7132.
159. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning - A laboratory Manual*; Cold Spring Harbour Laboratory Press, **1989**; Vol. 3, B-21.

Appendix

HPLC trace of ODN-1

