

**SYNTHETIC APPROACHES TO NOVEL,  
TRIAZOLE-CONTAINING OLIGONUCLEOTIDE  
ANALOGUES**

Lia Ashley Maria Tazioli

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University,  
School of Engineering and Physical Sciences,  
September 2010

The copyright of this thesis is owned by the author. Any quotation from this thesis or any use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.

## ABSTRACT

The work reported in this thesis focuses on the development of synthetic approaches to prepare novel triazole-containing nucleic acid (TCNA) monomers for subsequent incorporation into oligomers. The triazole moiety was designed to be prepared using “click chemistry”. Initial studies involved development of viable synthetic pathways for preparation of both the required azide component, derived from L-serine methyl ester and the nucleobase-containing alkyne component. The azide has been successfully synthesised from either protected or unprotected L-serine methyl ester by direct diazotransfer employing the novel ‘diazo donor’, imidazole-1-sulfonyl azide **152**. Synthesis of the four protected nucleobase-containing alkyne components has been achieved in overall yields ranging from 55-89%. The key step involved alkylation of the appropriately protected nucleobase with propargyl bromide. A series of model ‘click’ reactions were performed in which it was found that the best yields of triazole products were obtained using CuSO<sub>4</sub>·5H<sub>2</sub>O and sodium ascorbate in a 1:2 ratio. These conditions have been applied to the ‘click’ reaction employing the thymine alkyne component **161** and L-serine derived azide **158** to afford the desired thymine-derivatised triazole product **246** in a 44% yield. Preliminary studies into converting the resulting triazole compound into the required phosphoramidite thymynyl TCNA monomer **252** have been undertaken.

## ACKNOWLEDGEMENTS

I wish here to express my sincere appreciation to Dr Nicola Howarth for inspiring my enthusiasm for this research area and for lending valuable support and encouragement in the course of countless informal discussions. I would like to thank my second supervisor Professor Dave Adams for his advice and support and Dr Richard Wightman, Dr Arno Kraft and Dr Kevin McCullough for their help.

Thanks also to Dr Alan Boyd for NMR spectroscopy, Mrs Christina Graham for elemental analysis, Dr Georgina Rosair for x-ray diffractometry, the National Mass Spectrometry service centre, Swansea for mass spectrometry and the Engineering and Physical Science Research Council for generous financial support of my studies.

I would also like to thank the friends that I have made over the course of my postgraduate research for making my time more enjoyable, namely Barbara, Ben, Koen and Mary. In particular, I would like to thank Jennyfer for her help and patience on a daily basis in the lab. Without her calming and positive words, I would have had a more stressful experience.

Finally, a special thanks to my parents and grandparents for their encouragement and patience throughout all of my studies. Without them I wouldn't be where I am today.

In this thesis, I hope to make a worthwhile contribution to the research field which I am both hopeful and indeed confident will, in the not too distant future, find a cure for all cancers.

Lia Tazioli

## ACADEMIC REGISTRY Research Thesis Submission



Name:	LIA ASHLEY MARIA TAZIOLI		
School/PGI:	EPS/CHEMISTRY		
Version: <i>(i.e. First, Resubmission, Final)</i>	FINAL	Degree Sought (Award <b>and</b> Subject area)	PhD/CHEMISTRY

### Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

- 1) the thesis embodies the results of my own work and has been composed by myself
- 2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
- 3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted\*.
- 4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
- 5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

\* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	
-------------------------	--	-------	--

### Submission

Submitted By <i>(name in capitals)</i> :	
Signature of Individual Submitting:	
Date Submitted:	

### For Completion in Academic Registry

Received in the Academic  
Registry by *(name in capitals)*:

Method of Submission

*(Handed in to Academic Registry; posted  
through internal/external mail):*

E-thesis Submitted (**mandatory for  
final theses from January 2009**)

Signature:

Date:

## CONTENTS

<b>ABSTRACT</b> .....	<b>i</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>ii</b>
<b>CONTENTS</b> .....	<b>iv</b>
<b>ABBREVIATIONS</b> .....	<b>vi</b>
<b>1.1 INTRODUCTION</b> .....	<b>2</b>
<i>1.1.1 Latest cancer statistics</i> .....	2
<i>1.1.2 Cancer</i> .....	3
<i>1.1.3 The cell cycle</i> .....	3
<i>1.1.4 Chemotherapy</i> .....	5
<b>1.2 CURRENT CHEMOTHERAPEUTICS</b> .....	<b>6</b>
<i>1.2.1 Antimetabolites</i> .....	6
<i>1.2.2 Spindle inhibitors</i> .....	7
<i>1.2.3 Genotoxic drugs</i> .....	8
<b>1.3 THE COMPONENTS OF DNA</b> .....	<b>12</b>
<i>1.3.1 Shapes of nucleotides</i> .....	15
<i>1.3.2 Secondary structure of DNA</i> .....	16
<i>1.3.3 Alternative secondary structures of DNA</i> .....	19
<i>1.3.4 Higher order DNA structures</i> .....	22
<b>1.4 DNA REPLICATION</b> .....	<b>23</b>
<b>1.5 GENE EXPRESSION</b> .....	<b>25</b>
<i>1.5.1 Transcription</i> .....	26
<i>1.5.2 Translation</i> .....	29
<b>1.6 METHODS FOR CONTROLLING GENE EXPRESSION</b> .....	<b>32</b>
<i>1.6.1 Antisense therapy</i> .....	32
<i>1.6.2 Antigene therapy</i> .....	35
<b>1.7 GROOVE BINDERS</b> .....	<b>35</b>
<i>1.7.1 Minor groove binders</i> .....	36
<i>1.7.2 Major groove binders</i> .....	41
<b>1.8 TECHNIQUES FOR DETERMINING THE BINDING AFFINITIES AND SELECTIVITIES OF OLIGONUCLEOTIDE ANALOGUES</b> .....	<b>45</b>
<b>1.9 OLIGONUCLEOTIDE ANALOGUES</b> .....	<b>46</b>
<i>1.9.1 Natural oligonucleotides</i> .....	46
<i>1.9.2 Base modifications</i> .....	47
<i>1.9.3 Sugar modifications</i> .....	50
<i>1.9.4 Phosphate modifications</i> .....	60
<i>1.9.5 Sugar-phosphate backbone modifications</i> .....	64
<b>1.10 RESEARCH OUTLINE</b> .....	<b>67</b>
<b>2. RESULTS AND DISCUSSION</b> .....	<b>72</b>

<b>2.1 INTRODUCTION.....</b>	<b>72</b>
<b>2.2 DEVELOPMENT OF A SYNTHETIC ROUTE TO AZIDE COMPONENT</b>	
<b>(91).....</b>	<b>73</b>
2.2.1 <i>Glycerol route (route A).....</i>	73
2.2.2 <i>Serine route (route B) .....</i>	82
2.2.3 <i>Conclusions.....</i>	96
<b>2.3 DEVELOPMENT OF THE ALKYNE COMPONENT SYNTHESIS .....</b>	<b>99</b>
2.3.1 <i>Thymine alkyne component synthesis.....</i>	100
2.3.2 <i>Cytosine alkyne component synthesis .....</i>	103
2.3.3 <i>Adenine alkyne component synthesis .....</i>	107
2.3.4 <i>Guanine alkyne component synthesis.....</i>	112
2.3.5 <i>Conclusions.....</i>	122
<b>2.4 DEVELOPMENT OF THE ‘CLICK’ CHEMISTRY .....</b>	<b>125</b>
2.4.1 <i>Introduction.....</i>	125
2.4.2 <i>Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition of azides and terminal</i> <i>alkynes.....</i>	127
2.4.3 <i>Model ‘click’ reactions .....</i>	132
2.4.4 <i>‘Click’ studies using serine-derived azide 158 and thymine alkyne 161. ....</i>	138
2.4.5 <i>Cytosine and adenine alkyne component ‘click’ reactions.....</i>	141
2.4.6 <i>Synthesis of the thymine TCNA phosphoramidite monomer.....</i>	142
2.4.7 <i>Conclusions.....</i>	144
<b>3.1 CONCLUSIONS .....</b>	<b>149</b>
<b>3.2 FUTURE WORK.....</b>	<b>153</b>
<b>4. EXPERIMENTAL .....</b>	<b>156</b>
<b>4.1 EXPERIMENTAL INTRODUCTION .....</b>	<b>156</b>
<b>Appendix A. The Genetic code.....</b>	<b>193</b>
<b>Appendix B. Table of the Amino acid structures.....</b>	<b>195</b>
<b>Appendix C. 160 Crystal structure.....</b>	<b>197</b>
<b>Appendix D. 161 Crystal structure.....</b>	<b>204</b>
<b>Appendix E. 174 Crystal structure.....</b>	<b>209</b>
<b>Appendix F. 166 Crystal structure .....</b>	<b>212</b>
<b>Appendix G. 175 Crystal structure. ....</b>	<b>217</b>
<b>Appendix H. 176 Crystal structure .....</b>	<b>222</b>
<b>REFERENCES.....</b>	<b>229</b>

## ABBREVIATIONS

A.....	adenine	DMSO.....	dimethylsulphoxide
Ar.....	aromatic	DMTrCl.....	
Abs.....	absorbance	..4,4'-dimethoxytriphenylmethylchloride	
Å.....	angstroms	DNA.....	deoxyribonucleic acid
AcCl.....	acetyl chloride	D <sub>2</sub> O.....	deuterated water
Ac <sub>2</sub> O.....	acetic anhydride	ds.....	double stranded
BuLi.....	butyllithium	EI.....	electron impact
Bn.....	benzyl	EIF.....	eukaryotic initiation factor
Bu <sup>t</sup> .....	butyl	EF.....	elongation factor
Bz.....	benzoyl	ESI.....	electron spray ionisation
C.....	cytosine	Et.....	ethyl
CDCl <sub>3</sub> .....	deuterated chloroform	EtOAc.....	ethyl acetate
CD.....	circular dichroism	EtOH.....	ethanol
CHCl <sub>3</sub> .....	chloroform	Et <sub>2</sub> O.....	diethyl ether
CH <sub>3</sub> CN.....	acetonitrile	ESI.....	electron spray ionisation
cm.....	centimetres	ESMA.....	
Cu.....	copper	.....	electrophoretic mobility shift assays
CuI.....	copper iodide	eq.....	equivalents
CuBr.....	copper bromide	FH <sub>2</sub> .....	dihydrofolate
CuSO <sub>4</sub> ·5H <sub>2</sub> O.....		FH <sub>4</sub> .....	tetrahydrofolate
.....	copper sulphate pentahydrate	g.....	grams
COSY.....	correlation spectroscopy	G.....	guanine
Cs <sub>2</sub> CO <sub>3</sub> .....	caesium carbonate	GNA.....	glycol nucleic acid
d.....	doublet	h.....	hours
dd.....	doublet of doublets	HIV.....	human immunodeficiency virus
DABCO... 1,4-diazabicyclo[2.2.2]octane		HNES.....	hard nano-electron spray
DBU.....		HWU.....	Heriot-Watt University
..... 1,8-diazabicyclo[5.4.0]undec-7-ene		HSV.....	herpes simplex virus
DCM.....	dichloromethane	Hz.....	hertz
DIPEA.....	<i>N,N</i> -diisopropylethylamine	H <sub>2</sub> O.....	water
DHFR.....	dihydrofolate reductase	IR.....	infra red
DMAP.....	4-dimethylaminopyridine	<sup>i</sup> PrCOCl.....	isobutyryl chloride
DMF.....	dimethylformamide	LiAlH <sub>4</sub> .....	lithium aluminium hydride

LiBH <sub>4</sub> .....	lithium borohydride	RNA.....	ribonucleic acid
<i>m</i> -.....	meta	RNP.....	ribonucleoprotein
<i>m</i> .....	multiplet	rRNA.....	ribosomal RNA
<i>M</i> .....	molar	<i>s</i> .....	singlet
Me.....	methyl	<i>ss</i> .....	single stranded
MeOD.....	deuterated methanol	<i>st</i> .....	stretch
MeOH.....	methanol	snRNA.....	small nuclear RNA
MEMCl.....	.....	snoRNA.....	small nucleolar RNA
.....	methoxyethoxymethyl chloride	T.....	thymine
Min.....	min	<i>t</i> .....	triplet
miRNA.....	micro RNA	<i>T<sub>m</sub></i> .....	melting temperature
mols.....	moles	tRNA.....	transfer RNA
mmols.....	millimoles	TBDPS.....	<i>tert</i> -butyldiphenylsilyl
mg.....	milligrams	TBDMS.....	<i>tert</i> -butyldimethylsilyl
mRNA.....	messenger RNA	TEA.....	triethylamine
MHz.....	mega hertz	<i>tert</i> .....	tertiary
MsCl.....	methane sulfonyl chloride	TFA.....	trifluoroacetic acid
N <sub>2</sub> .....	nitrogen	TFO.....	triplex forming oligonucleotide
NMR.....	Nuclear Magnetic Resonance	Tf <sub>2</sub> O.....	triflic anhydride
nm.....	nanometre	TfN <sub>3</sub> .....	triflic azide
NaHCO <sub>3</sub> .....	sodium bicarbonate	THF.....	tetrahydrofuran
NaBH <sub>4</sub> .....	sodium borohydride	THPTA.....	.....
NaH.....	sodium hydride	tris(hydroxypropyltriazolylmethyl)amine	
NaN <sub>3</sub> .....	sodium azide	TCNA... triazole-containing nucleic acid	
NaOH.....	sodium hydroxide	TLC.....	thin layer chromatography
<i>o</i> -.....	ortho	TMS.....	trimethylsilyl
<i>p</i> -.....	para	TMP.....	thymidine monophosphate
PE.....	petrol ether	TMG.....	1,1,3,3-tetramethylguanidinium
Ph.....	phenyl	Tol.....	toluene
PNA.....	peptide nucleic acid	UV.....	ultra-violet
ppm.....	parts per million	VSV.....	vesicular stomatitis virus
<i>q</i> .....	quartet	w/w.....	weight/weight
<i>quint</i> .....	quintet	xDNA.....	expanded DNA
R <sub>f</sub> .....	retention factor	yDNA.....	wide DNA
rt.....	room temperature	yyDNA.....	double wide



# **Chapter 1**

## **Introduction**

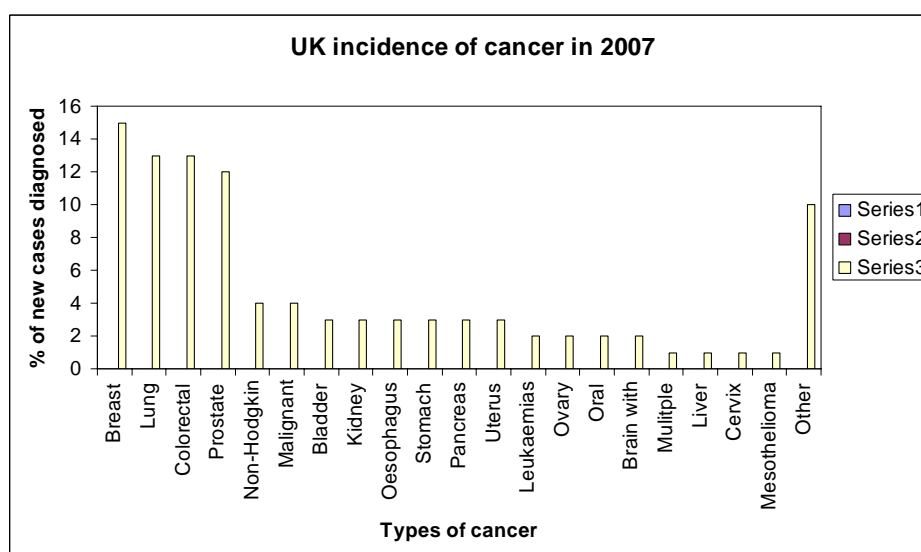
## 1.1 INTRODUCTION

Cancer is recognised as being one of the biggest killers in the developed world today. It is expected to affect all of us in some way, at some stage in our lives. Understanding why cancer arises in the body, how it can be prevented and how it can be ‘cured’ are the key aspects of current cancer research. Increased public awareness through education and publicity, advances in diagnostic testing and screening, and the development of new and improved treatments, has undoubtedly resulted in greater survival rates being observed. However, no widespread, single preventative measure or cure has been found. This failing is the driving force behind the extensive array of cancer research carried out in the world today and forms the basis from which our research group has stemmed.

This research will focus on the development of next generation anti-cancer therapeutics; by first establishing a greater understanding of current cancer treatments and then applying these findings to our own scientific endeavours.

### 1.1.1 Latest cancer statistics

With 298,000 people being diagnosed with cancer in the UK each year, it is expected that greater than 1 in 3 people will develop some form of the disease during their lifetime.<sup>1</sup> The twenty most common cancers identified in 2007 are shown in Graph 1. Although there are over 200 different types of cancer known, breast, lung, colorectal and prostate together account for over half (54%) of all new cases.<sup>1</sup>



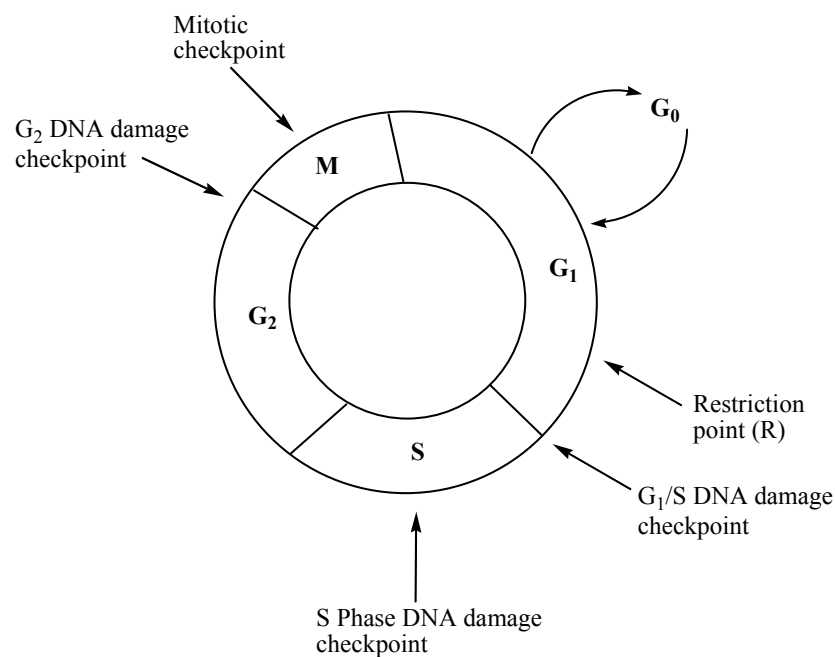
**Graph 1.** UK incidence of cancer in 2007.

### 1.1.2 Cancer

Cancer involves loss of control of a number of processes and, ultimately, results in an abnormal growth of cells which proliferate in an uncontrollable fashion and, in some cases, metastasize. Eventually this unrestrained growth and division of cancer cells interferes with the normal functioning of the body and results in the death of the sufferer.

### 1.1.3 The cell cycle

At the heart of cellular proliferation is the cell division cycle. This is the process by which a cell grows, replicates its DNA and then undergoes mitosis into two daughter cells.<sup>2</sup> A simplified diagram of the four main stages of the cell cycle is depicted in Figure 1.



**Figure 1.** The stages of the cell cycle (redrawn from ref 2).

The two most important stages of this cycle are considered to be the S phase, when DNA replication occurs, and the M phase (mitosis) when the cell divides into two.<sup>2</sup> The S phase is the longest part of the cell cycle and typically takes 10-12 hours of the 24 hours normally required for one complete cycle.<sup>2</sup> In the M phase, which is normally

---

complete within 1-2 hours, there is an ordered sequence of events that occur leading to the alignment and separation of duplicated chromosomes.<sup>2</sup> In order to ensure that the integrity of the genome is maintained, DNA replication cannot begin until mitosis has ended and mitosis cannot commence until DNA replication is complete.<sup>2</sup> Between these two phases are  $G_1$  and  $G_2$ , which are two gap phases.  $G_1$  supersedes mitosis and is the point at which the cell is responsive to both negative and positive growth signals.<sup>2</sup> It is here that the necessary proteins for DNA replication are synthesised.<sup>2</sup>  $G_2$  follows on from the S phase and is where the cell prepares to enter mitosis.<sup>2</sup> At this point, the cell undertakes the synthesis of key proteins required to assemble all of the machinery which is essential for mitosis.<sup>2</sup> The fifth stage,  $G_0$  (or quiescence), is the point at which the cell may reversibly leave  $G_1$  if it is deprived of the appropriate growth-promoting signals.<sup>2</sup>

Transition from one phase to another through the cell cycle is monitored and regulated at various sites known as checkpoints.<sup>3</sup> These checkpoints are sensor mechanisms, contained within the cell, which screen the cellular environment and establish if appropriate conditions have been satisfied before allowing further progression through the cycle.<sup>2</sup> Every checkpoint consists of three components: (i) a sensor mechanism to identify aberrant or incomplete cell cycle events such as DNA damage; (ii) a signal transduction pathway to carry the signal from the sensor to the third component, and; (iii) an effector, which has the ability to induce a cell cycle arrest should a malfunction be found and need to be resolved.<sup>2</sup>

The first of these checkpoints occurs at the  $G_1/S$  transition phase and is responsible for sensing DNA damage. The checkpoint at the  $G_2/M$  transition phase monitors the fidelity of DNA replication and, like the  $G_1/S$  transition, is also involved in detection of DNA damage.<sup>2</sup> A third checkpoint, known as the spindle checkpoint, is initiated during mitosis if a functional mitotic spindle is formed incorrectly.<sup>2</sup> In Figure 1, the restriction point, R, can be seen between mid and late  $G_1$ . It is here that the cell ascertains whether it has received sufficient growth signals to exit  $G_1$  and enter the S phase, replicate its DNA and fulfil one successful round of cell division.<sup>2</sup> If the cell fails this inspection, it will enter  $G_0$ .

Cancer occurs when these checkpoints malfunction, allowing inappropriate proliferation to take place.

---

**1.1.4 Chemotherapy**

Many of the compounds which have been investigated as anticancer agents act at multiple sites in the cell cycle.<sup>4</sup> Their action can be either cytostatic or cytotoxic depending on the status of the target's cell cycle.<sup>4</sup> There are three main classes of chemotherapeutic agents:

**(1) Antimetabolites:** These are drugs that interfere with the formation of key biomolecules within the cell, e.g. nucleotides. These drugs interrupt DNA replication and, consequently, cell division. Examples of antimetabolites include; folate antagonists (methotrexate),<sup>5</sup> purine antagonists (6-mercaptopurine)<sup>5</sup> and pyrimidine antagonists (5-fluorouracil).<sup>5</sup>

**(2) Genotoxic drugs:** These are drugs that interfere with DNA. These agents hinder DNA replication, and therefore, cell division. Examples of genotoxic drugs include the family of alkylating agents (mitomycin C), intercalating agents (doxorubicin), enzyme inhibitors (bleomycin) and groove binders (distamycin).

**(3) Spindle inhibitors:** Such agents prevent correct cell division by interfering with the cytoskeletal components enabling one cell to divide into two. An example of a spindle inhibitor is Taxol® (paclitaxel).

The common side effects associated with chemotherapy include; nausea and vomiting, constipation or diarrhoea, dry skin or rashes, lung problems and difficulty breathing, bleeding or bruising, muscle and nerve problems, fertility and sexuality problems, anaemia, a low number of white cells in blood counts and hair loss. Although, the damaged normal tissue usually recovers and repairs, these unpleasant side effects are often found to be more debilitating to the patient than the cancer itself. The lack of specificity and selectivity of chemotherapeutic agents towards cancer cells only, is the main source of these current clinical therapeutic problems. Chemotherapeutics used in cancer treatment target all cells undergoing cell division. Cell types that rapidly divide, such as those found in the bone marrow and in the lining of the small intestine, are the most susceptible to attack. Sites in the body where there is a steady demand for new cells, such as the skin, hair and nails are also unwanted targets.

## 1.2 CURRENT CHEMOTHERAPEUTICS

As outlined in section 1.1.4, page 5, there are three main classes of chemotherapeutic agents: antimetabolites, spindle inhibitors and genotoxic drugs.

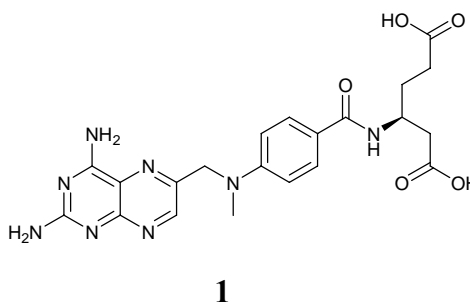
### 1.2.1 Antimetabolites

Antimetabolites are drugs capable of being processed, within a cell, under the disguise of being a normal metabolite, and therefore, interfere with regular cell division and function.<sup>5</sup> The majority of antimetabolite drugs currently used to treat cancer interfere with the production of either DNA or RNA.<sup>5</sup> The main classes of antimetabolites are: (i) folate antagonists, and; (ii) purine/pyrimidine antagonists.

#### (i) Folate antagonists

Folate antagonists (antifolates) inhibit the enzymes dihydrofolate reductase (DHFR) and thymidylate synthetase (used to generate thymidine monophosphate (TMP)), which participate in the formation of nucleotides.<sup>5</sup> DHFR converts dihydrofolate (FH<sub>2</sub>) into tetrahydrofolate (FH<sub>4</sub>) and subsequently, 10-formyl tetrahydrofolate (CH<sub>2</sub>FH<sub>4</sub>). By blocking these enzymes, the folate antagonists prevent nucleotide production and hence DNA replication and cell division are disrupted.<sup>5</sup>

An example of a folate antagonist currently used clinically is methotrexate [Figure 2].



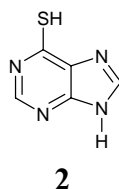
**Figure 2.** Structure of methotrexate.

Methotrexate is used to treat many different forms of cancer such as choriocarcinoma, breast, bladder, head and neck cancers, large cell and high grade lymphoma, acute lymphocytic leukaemia and osteogenic cancers.<sup>5</sup> Methotrexate's exact mechanism of

action is, as yet, uncertain.<sup>5</sup> However, it is known to enter the cell through an active transport system, which is available to the natural substrate folate, and then bind to and inhibit the enzymes.<sup>5</sup>

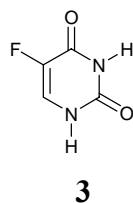
*(ii) Purine/pyrimidine antagonists*

Purine antagonists function by replacing purine nucleotides (adenine and guanine) in one or more normal cell functions.<sup>5</sup> They are ‘decoys’ for the natural nucleotides. Purine antagonists inhibit DNA synthesis in two ways: (1) they inhibit the synthesis of purine nucleotides, and; (2) they may become incorporated into DNA during its synthesis and further interfere with cell division.<sup>5</sup> An example of a purine antagonist drug which is currently used to treat many forms of leukaemia is 6-mercaptopurine [Figure 3].<sup>5</sup>



**Figure 3.** Structure of 6-mercaptopurine.

Pyrimidine antagonists function by blocking the synthesis of pyrimidine-containing nucleotides (thymine and cytosine).<sup>5</sup> As was the case for purine antagonists, they act as ‘decoys’ of the natural nucleotides and their mode of action is identical too. An example of a pyrimidine antagonist which is used clinically is 5-fluorouracil [Figure 4].



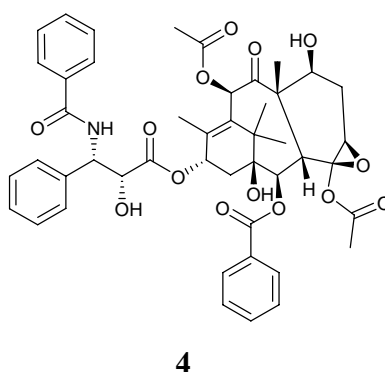
**Figure 4.** Structure of 5-fluorouracil.

### **1.2.2 Spindle inhibitors**

Spindle inhibitors are different to the previous examples in that they do not alter DNA structure or function. Instead they act on the mechanics of cell division.

As stated previously, mitosis is the process by which a cell is replicated and then divided into two separate daughter cells.<sup>6</sup> This process involves spindle fibres which attach themselves to the replicated chromosome and facilitate the movement of one copy of the chromosome to each side of the dividing cell.<sup>6</sup> Without such spindle fibres, the cell cannot divide and cell apoptosis results.<sup>6</sup>

Spindle inhibitor drugs act in a cell-cycle specific manner, halting cell division in the early stages of mitosis.<sup>6</sup> Spindle fibres are composed of microtubules which are composed of smaller subunits of proteins known as tubulins.<sup>6</sup> Certain types of spindle inhibitors associate with the tubulin monomers and prevent their synthesis.<sup>6</sup> This complex stops the correct formation of such spindle microtubules and disables the transportation of chromosomes.<sup>6</sup> The movement of replicated chromosomes is dependent on two processes: (1) the polymerisation of tubulins to form microtubules, and; (2) the breakdown of such microtubules.<sup>6</sup> An example of a current spindle inhibitor drug is taxol® [Figure 5].



**Figure 5.** *Structure of Taxol®.*

The mode of action for Taxol® is to bind to microtubules and consequently prevent their breakdown.<sup>6</sup>

### **1.2.3 Genotoxic drugs**

Genotoxic anticancer drugs can be classified into three main groups: (i) alkylating agents; (ii) enzyme inhibitors, and; (iii) intercalators.



---

*(i) Alkylating agents*

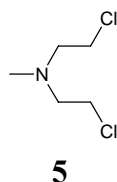
Alkylating groups act by adding alkyl groups to negatively charged groups.<sup>7</sup> They are extremely electrophilic compounds capable of reacting with nucleophiles to form strong covalent bonds.<sup>7</sup> There are two types of alkylating agents: mono and dialkylating species. Monoalkylating agents react with one guanine nucleobase, and are less common clinically, as they have been superseded by more efficient dialkylating compounds.

Alkylating agents act by three different mechanisms to cause disruption of DNA function and results in cell apoptosis. The three mechanisms are outlined as follows:

In the first mechanism, alkylating drugs, attach alkyl groups to guanine nucleobases.<sup>7</sup> Repair enzymes attempt to restore the alkylated bases, and in so doing, cause the DNA to become fragmented.<sup>7</sup>

The second mechanism involves cross-linking between two guanine nucleobases on DNA and is only seen for dialkylating agents.<sup>7</sup> These cross-links may occur within a single molecule of DNA (intrastrand cross-linking), or by bridging between two different molecules of DNA (interstrand cross-linking).<sup>7</sup> Ultimately, this course of action prevents DNA from separating and therefore prevents transcription from occurring.<sup>7</sup>

The first alkylating agent used clinically, was the nitrogen mustard compound, mechlorethamine (Mustargen) [Figure 6].<sup>7</sup>

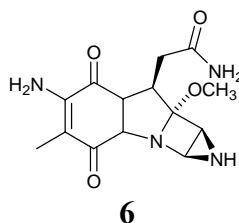


**Figure 6.** *Structure of mechlorethamine.*

Mechlorethamine is an example of this second mode of action and is known as a bifunctional agent, as it is capable of binding and reacting at two different sites.<sup>7</sup> The chloride ion is displaced intramolecularly by the nitrogen atom, to form a substantially electrophilic aziridine ion intermediate.<sup>7</sup> This reactive intermediate then alkylates the

*N7* of guanine in one or both strands of the DNA molecule, which results in cross-linking between the guanine residues in DNA chains, and/or depurination which facilitates DNA strand breakages.<sup>7</sup> It is currently used to treat Hodgkin's lymphoma, non-Hodgkins lymphoma and some types of chronic leukaemia.

The third mechanism of action of alkylating agents involves inducing the mispairing of nucleotides resulting in mutations.<sup>7</sup> Guanine nucleobases erroneously pair with thymine instead of its Watson-Crick base pair cytosine, and the resulting mismatch, if not corrected, has the potential to become a permanent mutation.<sup>7</sup> Mitomycin C [Figure 7] is an example of a current drug which alkylates using primarily, this third mechanism.



**Figure 7.** Structure of mitomycin C

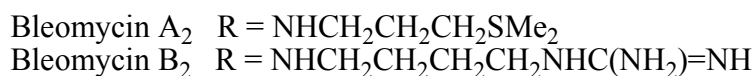
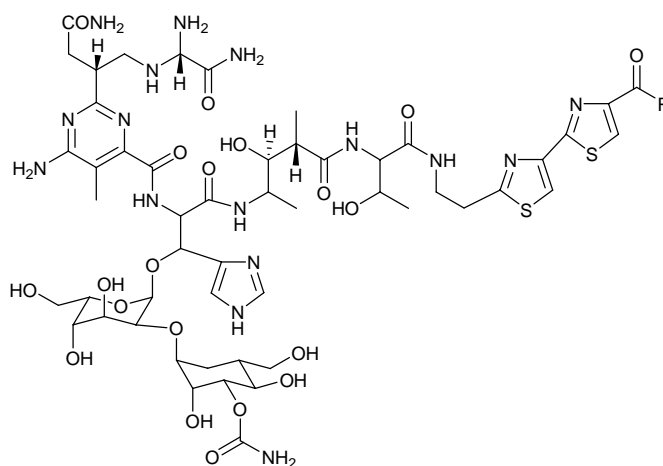
It is derived from the bacterium *Streptomyces caespitosus* and is commonly used as an antibiotic, however, it can also function as an alkylating agent.<sup>7</sup>

Alkylating agents are not selective in their action, and therefore, have the ability to react with all nucleophiles in the body, such as proteins or other macromolecules.<sup>7</sup> Alkylation can occur in both cycling and resting cells and are cell-cycle non-specific.<sup>7</sup> However, proliferating cells are more susceptible to attack from these agents.<sup>7</sup> For these reasons, it is vital that the alkylating agents are applied locally, targeting the tumour.<sup>7</sup>

#### (ii) Enzyme inhibitors

Genotoxic enzyme inhibitors are drugs which interfere with key enzymes, e.g. topoisomerases, which are involved in DNA replication, thereby inducing DNA damage. Bleomycin (Blenoxane®) [Figure 8] is one such example of a compound. It has the ability to cleave the phosphodiester backbones of duplex DNA and, subsequently, avert the enzyme DNA ligase from repairing the damage.<sup>8</sup> Bleomycin is a glycopeptide antibiotic, which is produced by the bacterium *Streptomyces verticillus*. Its

mode of action is to form complexes with iron, which go on to reduce molecular oxygen to superoxide and hydroxyl radicals. These radicals subsequently cause single- and double-stranded breaks in DNA.<sup>8</sup> Further, these reactive oxygen species induce lipid peroxidation, carbohydrate oxidation and alterations in prostaglandin synthesis and degradation.<sup>9</sup> Bleomycin is currently used to treat many forms of cancer e.g. Hodgkin's lymphoma, non-Hodgkin's lymphoma, testicular cancer, cancers of the head and neck and also, fluid on the lung.

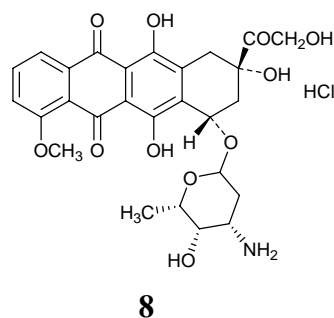


7

**Figure 8.** Structure of Bleomycin.

### (iii) Intercalators

Intercalators bind in the spaces between nucleotide base pairs in the double helix of DNA.<sup>10</sup> Their flat,  $\pi$ -electron chromophore binds tightly but reversibly to DNA through a combination of dipolar forces such as electrostatic, hydrophobic and hydrogen-bonding interactions.<sup>10</sup> Through binding, intercalating drug molecules alter the structure of DNA, which consequently prevents polymerases and other DNA binding proteins from functioning correctly.<sup>10</sup> They interfere with transcription and replication and induce mutations in DNA.<sup>10</sup> An example of an intercalating drug currently used clinically is doxorubicin (Adriamycin®) [Figure 9].



**Figure 9.** Structure of doxorubicin.

Doxorubicin was first isolated from the fungus *Streptomyces peucetius* and is an anthracycline antibiotic. It acts as both an intercalator and an enzyme inhibitor and is used to treat many forms of cancer including Hodgkins and non-Hodgkins lymphoma, breast, lung, bladder, ovarian, testicular, gastric, thyroid cancer, soft tissue sarcoma, hepatoma, Wilm's tumour, acute leukaemia and neuroblastoma.<sup>10</sup>

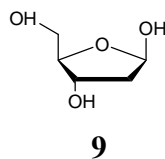
The main disadvantage for all the chemotherapeutic agents described in this section, is that they fail to demonstrate high specificity to cancer cells, mainly, as a result of them being small molecules. As stated previously, all forms of rapidly dividing cells are targeted, including normal healthy cells, and this results in the common side-effects of chemotherapy described in section 1.1.4, on page 5. They also fail to discriminate effectively between cancer-causing and normal genes. In order to address the later issue, selectivity must be improved. To develop drugs that inhibit specifically one particular gene, much larger compounds have to be employed (see the later section 1.7.1. on pages 37-42).

The purpose of the research reported in this thesis was to develop the next generation of cancer chemotherapeutics which would be more selective for cancer cells. We were interested in regulating gene expression at the level of DNA. In order for us to discuss the specific drugs that target DNA, it is first necessary to understand the structure and function of DNA itself.

### 1.3 THE COMPONENTS OF DNA

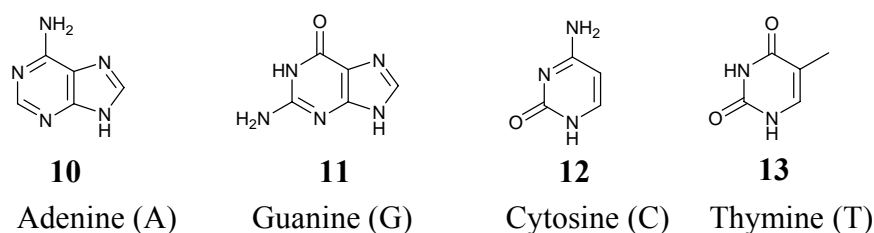
DNA is essentially a highly negatively charged biopolymer, the monomers of which are called nucleotides. A nucleotide consists of three components; a five-carbon sugar, a

heterocyclic base and a phosphate group. The sugar in DNA is 2-deoxy-D-ribose, a furanose, as shown in Figure 10.



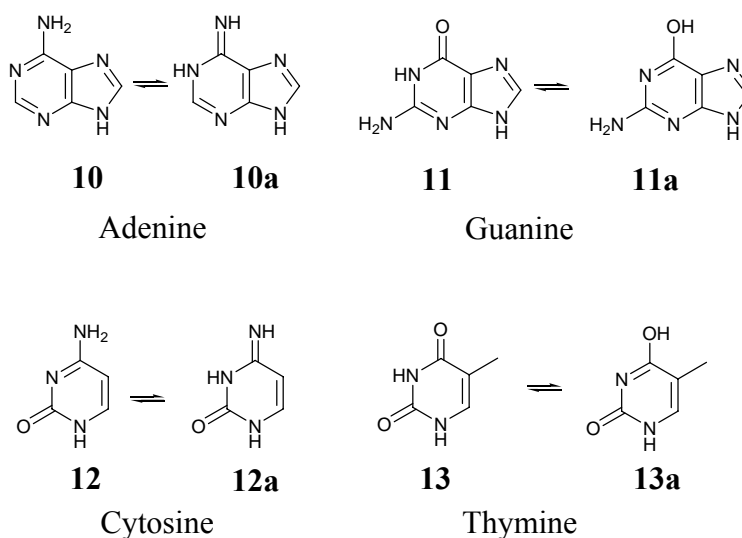
**Figure 10.** Structure of 2-deoxy-D-ribose.

DNA consists of four unique, organic heterocyclic bases; adenine (A), guanine (G), thymine (T) and cytosine (C). As stated previously, adenine and guanine are bicyclic purines and cytosine and thymine are monocyclic pyrimidines, the structures of which are depicted in Figure 11.



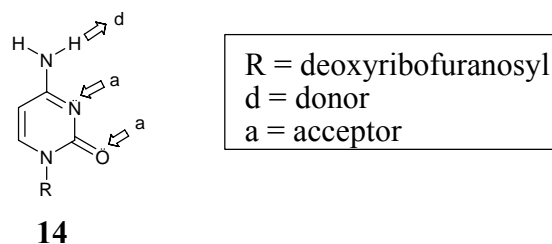
**Figure 11.** DNA bases.

At physiological pH ( $5 < \text{pH} < 9$ ), the bases exist in the forms shown in Figure 11 above; however, at pH greater than 3, keto-enol and amino-imino, tautomeric forms may also exist.<sup>6</sup> These tautomeric forms are shown in Figure 12.



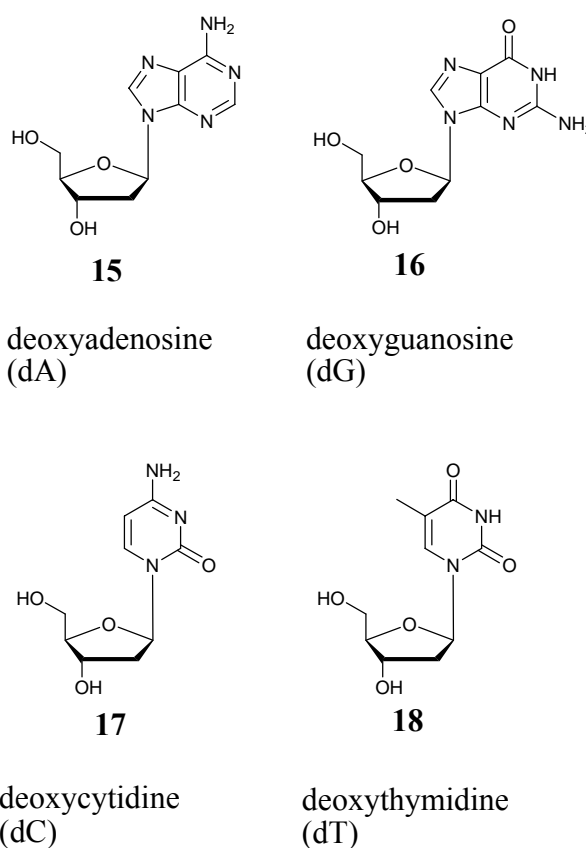
**Figure 12.** Tautomeric forms.

The planar faces of each of the bases make them hydrophobic.<sup>6</sup> The N-H groups are good hydrogen-bond donors while the  $sp^2$ -hybridized electron pairs, on both the oxygens of the carbonyl groups and ring nitrogens, are much better hydrogen-bond acceptors than the oxygens of either the phosphate or the pentose sugar.<sup>6</sup> The hydrogen donor/acceptor abilities of the nucleotide bases is illustrated in Figure 13 taking deoxycytidine as an example.<sup>6</sup>



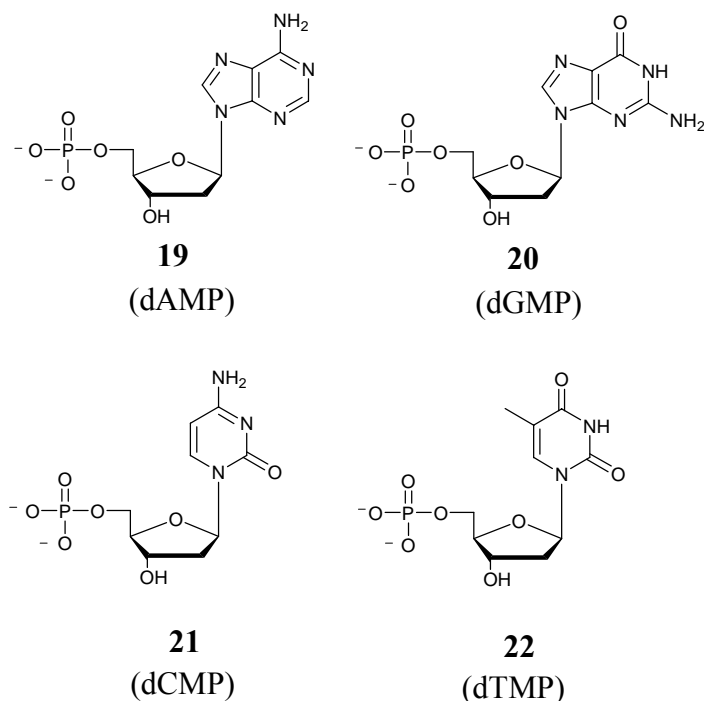
**Figure 13.** Donor/acceptor ability of deoxycytidine.<sup>6</sup>

The nucleotide bases are attached to the sugar moiety through  $\beta$ -glycosidic links between  $C1$  of the sugar and either  $N1$  or  $N9$  of a pyrimidine or purine base, respectively.<sup>6</sup> The resulting structures are known generally as deoxyribonucleosides, and the four deoxyribonucleosides present in DNA are shown in Figure 14.



**Figure 14.** Structures of the deoxyribonucleosides.

The name nucleotide is given to phosphate esters of nucleosides. In the case of DNA, it refers to a nucleoside bearing a 5'-phosphate group on C5 of the sugar ring, as shown in Figure 15.



(MP = 5'-monophosphate)

**Figure 15.** Structures of the deoxyribonucleotides.

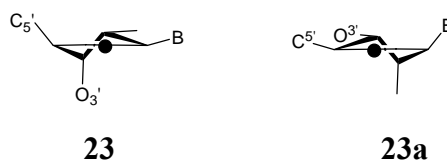
### 1.3.1 Shapes of nucleotides

Nucleotides incur many interactions between non-bonded atoms as a result of their bulky substituents.<sup>6</sup> In order to minimise these non-bonding interactions, nucleotides adopt compressed structures through use of sugar puckers and *syn*- and *anti*-conformers of the glycosidic bond.

#### *The sugar pucker*

In order to minimize the interactions between non-bonded substituents, the furanose rings of nucleotides are twisted out of the plane.<sup>6</sup> This 'puckering' is defined by identifying the main displacement of C'<sub>2</sub> and C'<sub>3</sub> from the median plane denoted by C'<sub>1</sub>-O'<sub>4</sub>-C'<sub>4</sub>.<sup>6</sup> The displacement of these two atoms is not equal.<sup>6</sup> If for example, the *endo*-displacement of C'<sub>2</sub> is found to be greater than the *exo*-displacement of C'<sub>3</sub>, then the

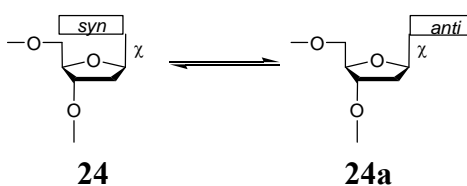
conformation is called  $C^{2'}$ -*endo*.<sup>6</sup> The *endo* face of the sugar ring contains the nucleotide base together with  $C'_{5'}$  and the *exo* face is the opposite side.<sup>6</sup> The sugar puckers are positioned in the north and south domains of the pseudorotation cycle of the sugar ring.<sup>6</sup> Fortunately, the relative shape of the C-C-C-C bonds in the  $C^{2'}$ -*endo* and -*exo*- forms reflect the (S) and (N) designations.<sup>6</sup> In regular dsDNA, the sugar moiety adopts the  $C^{2'}$ -*endo* (S) pucker.<sup>6</sup> The sugar puckers are shown in Figure 16.



**Figure 16.** Structures of  $C^{2'}$ -*endo* (S) **23** and  $C^{3'}$ -*endo* (N) **23a** sugar puckers.

#### *The syn-anti conformation of the glycosidic bond*

*Syn*- and *anti*-conformers result from rotation about the glycosidic bond through the torsion angle  $\chi$ .<sup>6</sup> The heterocyclic bases are contained within a plane perpendicular to the furanose ring and approximately intersect the  $O'_{4'}-C'_{1'}-C'_{2'}$  angle.<sup>6</sup> As a result, the bases can reside in two orientations, the *syn*- conformer and the *anti*-conformer.<sup>6</sup> For the *anti*-conformer, the *H6* (pyrimidine) or *H8* (purine) atom is positioned above the sugar ring [Figure 17] and in the case of the *syn*-conformer, it is the *O2* (pyrimidine) or *N3* (purine) atom which is positioned above the sugar ring [Figure 17].<sup>6</sup> In regular dsDNA, nucleotides adopt the *anti*-conformation.<sup>6</sup>

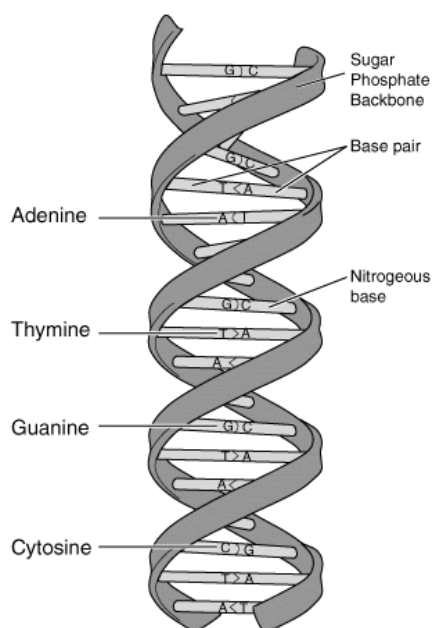


**Figure 17.** Structures of the *syn*- **24** and *anti*- **24a** conformations.

### 1.3.2 Secondary structure of DNA

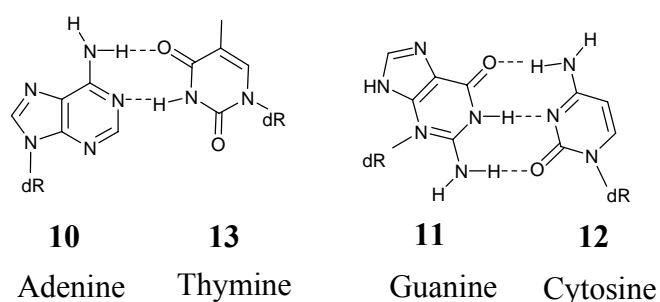
DNA generally exists in the form of a double-stranded helix, with the exception of a few viruses where it is found as a single strand. Two polynucleotide chains, come together in an anti-parallel fashion to form a double helix and are stabilised by hydrogen bonding that ensues from complementary base pairing.<sup>11</sup> The double helix of DNA is depicted on the next page in Figure 18.





**Figure 18.** Structure of the double helix (taken from [www.tutorvista.com](http://www.tutorvista.com)).

From Figure 18, it can be seen that the sugar-phosphate backbone is on the outside of the molecule encapsulating the nucleobase pairs in the core of the structure. Two hydrogen bonds form between an adenine of one strand and a thymine on the other, when they are located opposite each other,<sup>6</sup> whereas three hydrogen bonds form between a cytosine and a guanine. Adenine only base pairs with thymine and cytosine only base pairs with guanine. This base association shown in Figure 19, is known as Watson-Crick base pairing after the scientists who first reported the structure of DNA.<sup>6</sup>



**Figure 19.** Watson-Crick base pairing, with hydrogen bond formation.

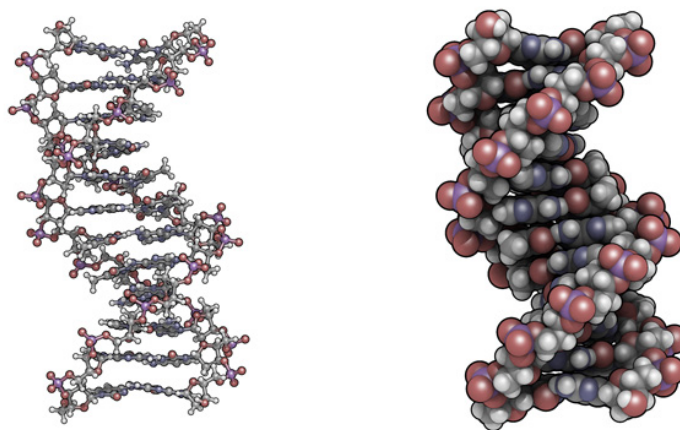
The flat base pairs arrange themselves in the structure of dsDNA in such a way, so as to eradicate water molecules. They achieve this by stacking and rotating slightly, relative to one another.<sup>6</sup>

## B-DNA

B-DNA is the commonest form of dsDNA. Its key feature is that it has a major and minor groove of approximately equal intensity.<sup>6</sup> The depth of the major and minor grooves are 8.8Å and 7.5Å respectively, and the width of the major and minor grooves are 11.7Å and 5.7Å respectively.<sup>6</sup> The heterocyclic bases stack directly above their neighbours in the same strand and are perpendicular to the axis of the helix.<sup>6</sup> As mentioned previously, the  $C^{2'}$ -*endo* sugar pucker is the predominant conformation adopted by the sugar moiety and all the glycosides adopt the *anti*-conformation.<sup>6</sup> Phosphate groups in the same chain are 6.7Å apart and there are 10 base pairs per turn in B-DNA.<sup>6</sup>

The wide major groove of B-DNA is surrounded by a unimolecular surface of water molecules which interacts with all the C=O, N and NH groups that are present in this region, and is responsible for solvating the phosphate backbone.<sup>6</sup> In the smaller minor groove, two water molecules per base-pair can be found, in a well-ordered zig-zag pattern.<sup>6</sup>

The average parameters for B-DNA are shown in Table 1<sup>6</sup> (page 21) and diagrams of its structure are shown in Figure 20 below.



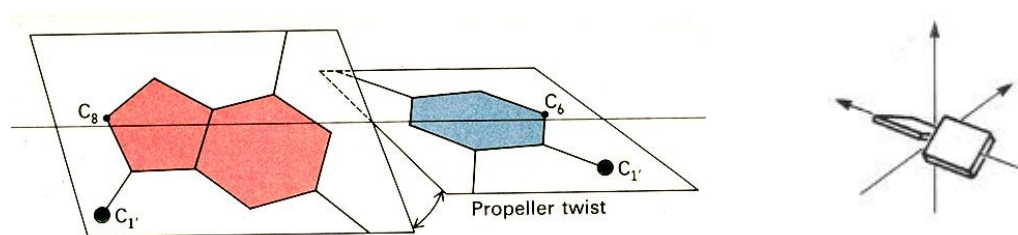
**Figure 20.** Structure of B-DNA (taken from [www.somewhereville.com](http://www.somewhereville.com)).

In addition to B-DNA, alternative secondary structures of DNA are known. These include A-DNA and Z-DNA.

### 1.3.3 Alternative secondary structures of DNA

There are several conformations that dsDNA can adopt apart from B-DNA as summarised in Table 1 on page 21.<sup>6</sup> Of these, the two most important alternative double helical structures are; the A- and Z- form (Figure 22, page 20). It is believed, that both the A- and Z- form of the double helix may exist in localized sections of the chromosome, but not exclusively.<sup>1</sup>

The A-form arises when dsDNA becomes dehydrated.<sup>6</sup> This causes the heterocyclic bases to tilt by  $20^\circ$  (propeller twist [Figure 21]) relative to the helical axis and the overall perception of the double helix to become stocky in shape.<sup>6</sup> The A-form follows the Watson-Crick model with anti-parallel, right-handed double-helices.<sup>11</sup> All the parameters for the A-form are shown in Table 1 on page 21.<sup>6</sup>

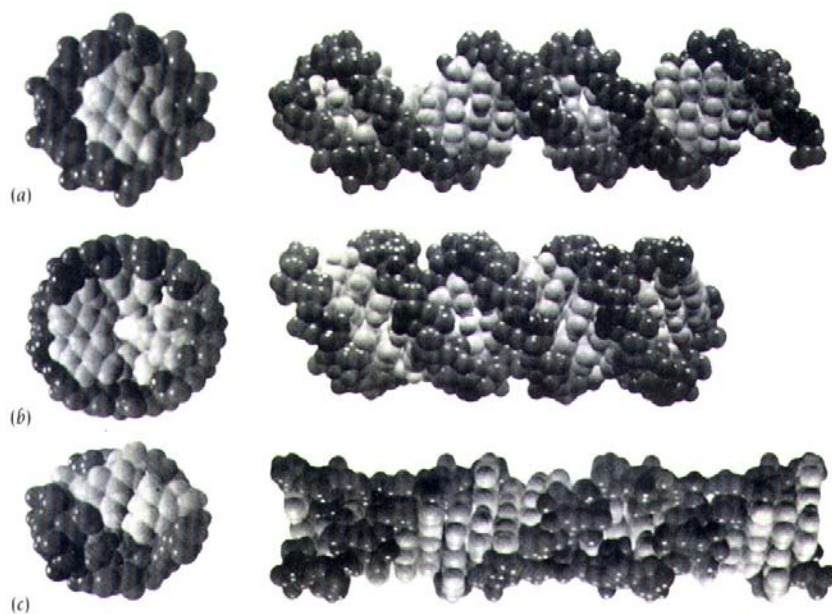


**Figure 21.** Schematic diagrams of the propeller twist that arises in A-DNA (taken from [www.ed.stanford.edu](http://www.ed.stanford.edu)).

In the A-form, the major groove is deeper by  $4.7\text{\AA}$  and narrower by  $9.0\text{\AA}$  compared to B-DNA, there are 11 base pairs per turn and the  $C^{3'}$ -endo sugar pucker is predominant.<sup>6</sup> The structure of A-DNA is shown in Figure 22 on page 20.

The Z-form of dsDNA is a left-handed double helical structure. This arises due to a switch in the glycosidic torsion angle, from the regular *anti*- to the atypical *syn*-conformation, for one nucleoside (the purine) in each base pair.<sup>6</sup> This *anti-syn* relationship alternates along the DNA backbone which, as a consequence, causes the phosphorous atoms to follow a zig-zag path.<sup>6</sup> Further, the sugar pucker alternates from  $C^{2'}$ -endo for the *anti*-residues to  $C^{3'}$ -endo for the *syn*-nucleosides.<sup>6</sup> It is the combination of these two changes that causes a significant difference to the base-stacking abilities of Z-DNA compared to the other forms of dsDNA.<sup>6</sup> As a consequence of this, the minor

groove of Z-DNA is so cavernous that it contains both the helical axis and the major groove. The surface of the minor groove is also convex exposing the *C5* of the cytosine residues and the *N7* of the guanine *C8* moieties.<sup>6</sup> The major groove is shallower by 5.1Å and narrower by 2.9Å compared to B-DNA, there are 12 base pairs per turn and the *C*<sup>3'</sup>-*endo(syn)* sugar pucker is predominant.<sup>6</sup> The parameters for Z-DNA are shown in Table 1 on page 21 and its structure is shown in Figure 22 below.



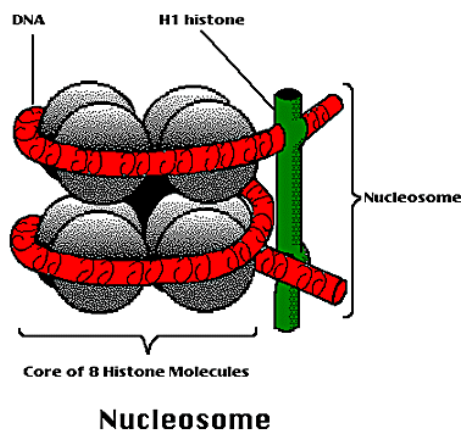
**Figure 22.** Structures of (a) B-DNA, (b) A-DNA and (c) Z-DNA (taken from [www.nthu.edu](http://www.nthu.edu)).

Structure Type	Helix sense	Residues Per turn	Twist per bp $t^\circ$	Displacement bp / Å	Rise per bp / Å	Base tilt $\tau^\circ$	Sugar pucker	Groove Width / Å		Groove Depth / Å	
								minor	major	minor	major
A-DNA	R	11	32.7	4.5	2.56	20	<i>C-3'-endo</i>	11.0	2.7	2.8	13.5
B-DNA	R	10	36	-0.2 to -1.8	3.3-3.4	-6	<i>C-2'-endo</i>	5.7	11.7	7.5	8.8
Z-DNA	L	12	-9, -51	-2 to -3	3.7	-7	<i>C-3'-endo(syn)</i>	2.0	8.8	13.8	3.7

**Table 1.** Average helix parameters for the major DNA conformations.

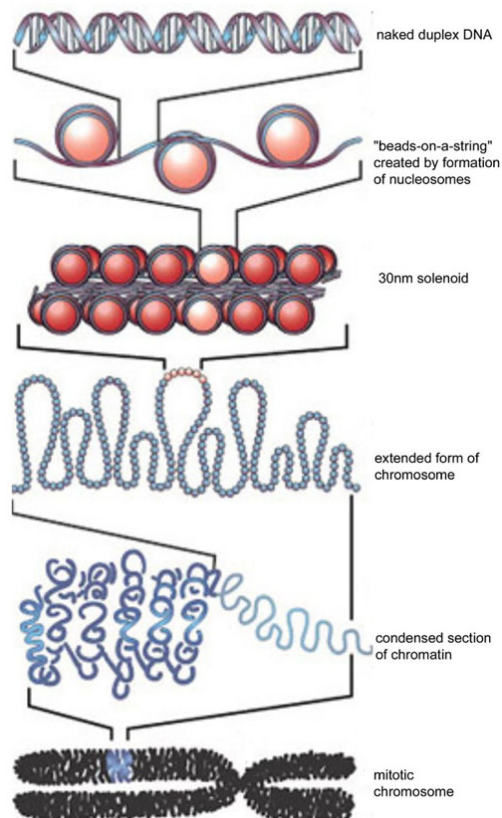
### 1.3.4 Higher order DNA structures

DNA in eukaryotic cells exists as a DNA-protein complex known as chromatin. Small basic proteins called histones, abundant in the amino acids arginine and lysine, bind to the phosphate groups on the DNA backbone through formation of ionic bonds.<sup>6</sup> Four of these proteins, H2A, H2B, H3 and H4, (two of each) come together to form an octamer protein complex called a nucleosome core [Figure 23].<sup>6</sup>



**Figure 23.** Structure of a nucleosome (taken from [www.accessexcellence.org](http://www.accessexcellence.org)).

The DNA duplex (red) wraps around the block of eight histones (grey) to give 1.75 turns of a left-handed superhelix.<sup>6</sup> Each successive nucleosome core is linked *via* an intervening stretch of linker DNA of length 30-40 base-pairs. The structure that results is termed ‘beads on a string’ and is shown schematically in Figure 24 on page 23. Although this process condenses the DNA molecule, further compaction is still required.<sup>6</sup> The nucleosomes form a 10 nm fibre which is further condensed by a fifth histone, H1 (green), to form what is known as the 30 nm fibre or a solenoid.<sup>6</sup> This fibre is arranged as long loops attached to a central chromosomal protein scaffold, which forms the highly dense chromosome structure. Chromosomes are located in the nucleus of a cell.



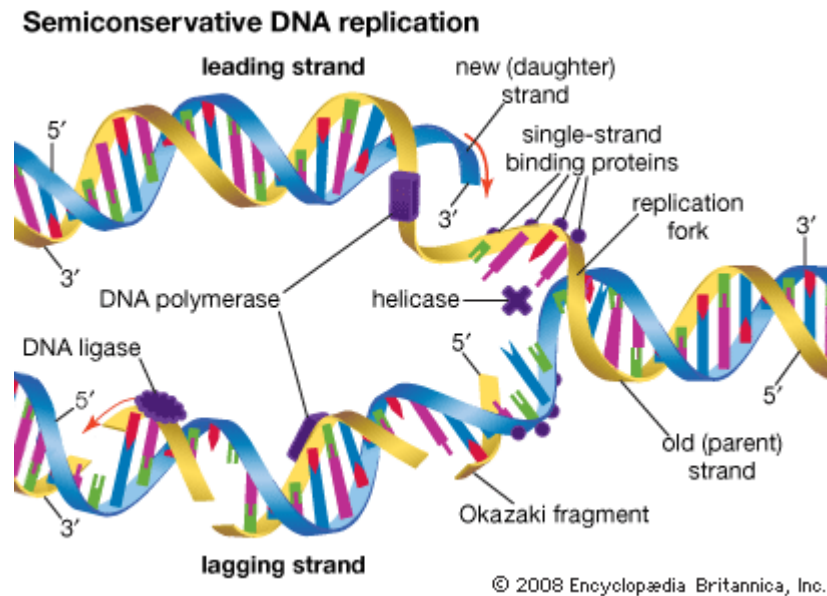
**Figure 24.** *Hierarchy of chromatin (taken from [www.themedicalbiochemistrypage.org](http://www.themedicalbiochemistrypage.org)).*

When cells divide, each chromosome must be carefully replicated to ensure that each daughter cell obtains a complete and accurate set of information.

## 1.4 DNA REPLICATION

Double-stranded DNA (dsDNA) replication is known to be semi-conservative, in that the two original parental strands are separated and act as a template for synthesising the new strand.<sup>6</sup> As a result of this, there is precise duplication of DNA in terms of both quantity and sequence.<sup>6</sup>

Firstly, the two original polynucleotide chains of dsDNA, parental strands, are separated at the origin of replication by the enzyme helicase (Figure 25, page 24).



**Figure 25.** *Semi-conservative DNA replication (taken from [www.britannica.com](http://www.britannica.com)).*

Due to the vast length of DNA and the fact that it is attached to protein structures (previously described in section 1.3.4, page 22), separation presents a topological problem known as overwinding.<sup>12</sup> This results in the formation of positive supercoils that if unrelieved, would halt strand separation and therefore DNA replication.<sup>12</sup> Such supercoils are removed by topoisomerases enzymes of which there are two types; type I and type II, their action is summarised in Table 2 below.

<u>Topoisomerase</u>	<u>Action on DNA</u>	<u>Effect on supercoiling</u>
Type I	Cuts one DNA strand	Relaxes positive and negative supercoils.
Type II (gyrase)	Cuts two DNA strands	Relaxes positive supercoils but cannot insert negative supercoils.

**Table 2.** *Summary of the types of topoisomerases and their action.*

At the point at which the two strands are separated, the replicative fork [Figure 25], DNA is synthesised at a rate of 50 base pairs per second.<sup>12</sup> There are hundreds of origins of replication situated along the chromosome, allowing for replicative forks to



work in both directions.<sup>12</sup> Enzymes called DNA polymerases catalyse the polymerisation of nucleotides into DNA.<sup>12</sup> DNA polymerase can only synthesis DNA in a 5'- to 3'- direction, which creates a problem due to the two parental strands running in anti-parallel directions, hence, only one strand has the required directionality.<sup>12</sup> From Figure 25, the leading strand is synthesised continuously in the 5'- to 3'- direction and the lagging strand is synthesised discontinuously in a 5'- to 3'- direction.<sup>6</sup>

DNA polymerase cannot initiate chain synthesis as it is unable to join together two free nucleotides. This problem is overcome by the preliminary links, phosphodiester bonds, being prepared by an RNA polymerase known as primase.<sup>12</sup>

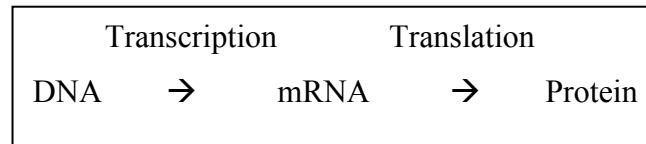
As elongation of the leading strand continues, a greater portion of the lagging strand becomes single-stranded.<sup>6</sup> DNA polymerase then acts on this lagging strand to synthesise short, complementary strands called Okazaki fragments<sup>6</sup> [Figure 25]. As DNA unwinds, primase repeats the initiation section on the lagging strand, to which these Okazaki fragments become attached [Figure 25].<sup>12</sup> The lagging strand is synthesised as a series of these Okazaki fragments in which the gaps between have been filled by DNA polymerase. Another enzyme called DNA ligase, then joins the fragments together forming the complete DNA strand.<sup>12</sup>

Due to the necessity for a short segment of single-stranded DNA, on both the leading and lagging strand close to the replicative fork, this area is protected by single-strand binding proteins.<sup>6</sup> Single-stranded binding proteins, which themselves have no base sequence specificity, bind to the separated strands, the energy release of which drives strand separation to completeness.<sup>12</sup> A product of the *ssb* gene, these proteins also protect this highly susceptible region from hydrolysis by nucleases.<sup>6</sup> DNA polymerase displaces these proteins in the process of synthesising a new complementary strand and the overall result is the production of two daughter double helices containing identical genetic content to their parent.<sup>6</sup>

## 1.5 GENE EXPRESSION

The genetic code, derived from the sequence of the four bases, A, T, C and G, is used to direct the assembly of each of the twenty amino acids required to produce the protein for which a specific gene is responsible.<sup>12</sup>

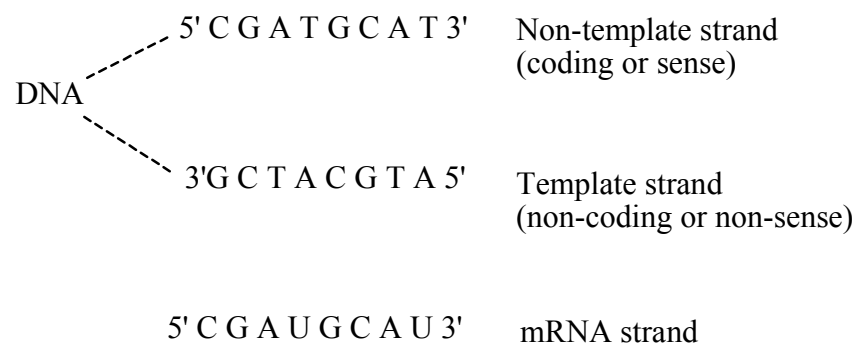
A gene does not directly participate in protein synthesis. In eukaryotes, the DNA is enclosed inside the nuclear membrane whilst the protein-synthesising machinery is situated in the cell cytoplasm.<sup>12</sup> Instead, a copy of the code is transcribed into mRNA transcripts through a process known as transcription.<sup>12</sup> The manner in which mRNA transcripts guide the correct sequencing of amino acids to synthesise proteins is called translation.<sup>6</sup> The overall flow of genetic information involved in gene expression is shown schematically in Figure 26 below.



**Figure 26.** Schematic diagram showing the flow of genetic information involved in gene expression.

### **1.5.1 Transcription**

Transcription is the first step in the biosynthetic process of protein synthesis within a cell and, as stated above, is the process of transcribing the information stored in DNA to mRNA transcripts. The strand of DNA double helix that is copied by base pair complementarity is known as a the template strand [Figure 27].<sup>12</sup>

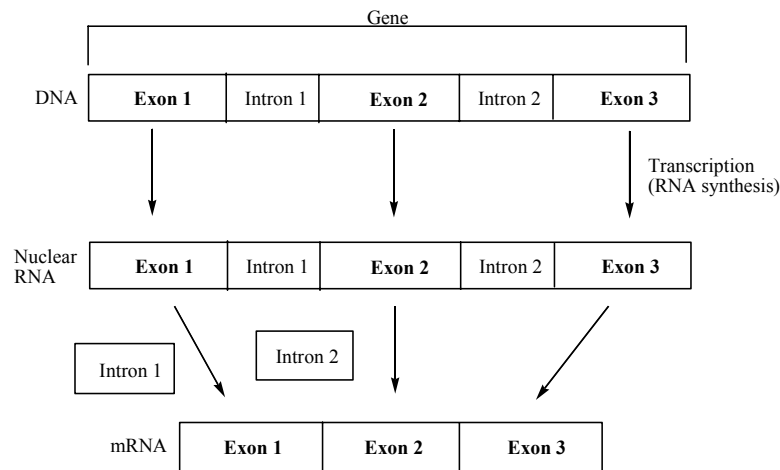


**Figure 27.** The relationship between mRNA and the template/non-template strands of DNA.

As previously mentioned in section 1.3.4, page 22, DNA undergoes a great degree of packing in order to fit its vast length into the nucleus of a cell. Before (or during) transcription, the higher order structure of DNA must selectively loosen at a particular

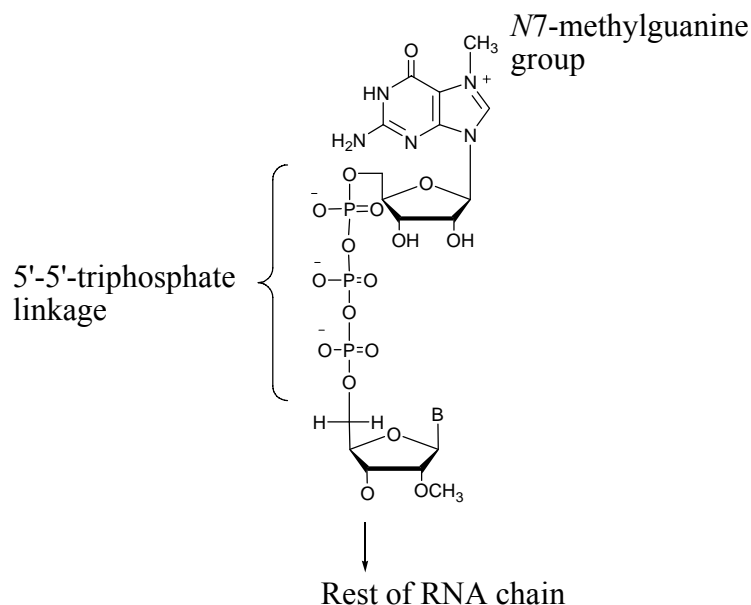
gene to be transcribed.<sup>12</sup> This is achieved by a large number of protein transcription factors assembling on the 5'- side of a gene at the promoter site and recruiting RNA polymerase from the surrounding medium, to bind to this complex.<sup>12</sup> Together, the protein transcription factors and RNA polymerase facilitate the opening of the DNA double helix at a specific gene.<sup>12</sup> The progression from this initiation to the actual synthesis of mRNA is complicated and to date, incompletely understood.<sup>12</sup> However, RNA polymerase is known to move along the template strand in a 3'- to 5'- direction.<sup>12</sup> In doing so, the enzyme reads the exposed DNA nucleotide bases and assembles the corresponding Watson-Crick complementary ribonucleotide (supplied as triphosphates) to the RNA strand being formed.<sup>12</sup> When RNA polymerase reaches the end of the gene, it synthesises the sequence AAUAAA (which is coded for by the template strand).<sup>12</sup> The enzyme continues to transcribe beyond this point and then terminates by an as yet not understood mechanism.<sup>12</sup> A separate enzyme cuts the RNA transcript a short distance beyond this termination signal and then another enzyme adds adenine nucleotides to form a polyA tail.<sup>12</sup> As many as 200 adenines are added to the 3'-end of the mRNA transcript and the main reason for this is believed to be to protect the transcripts from rapid degradation from 3'- to 5'- *exonucleases*.<sup>12</sup> Further, it may also be necessary for the transport of the mRNA transcripts and in their subsequent translation.<sup>12</sup>

The initial mRNA (pre-RNA) produced by transcription is proportionate in length to the size of the protein to which it codes.<sup>12</sup> It is divided into several segments linked by intervening stretches of RNA which do not code for the amino acid sequence.<sup>12</sup> These intervening stretches of RNA are called introns and the coding segments are called exons.<sup>12</sup> There are between 2 and 500 introns in every gene, which vary in length from 50 to 20,000 base pairs, in contrast to exons which are normally less than 1000 base-pairs in length.<sup>12</sup> Due to pre-RNA being longer than that required for subsequent translation, following transcription, a maturation procedure takes place to remove the introns and thus shorten the molecule.<sup>12</sup> The process by which the introns are eliminated and the exons are linked together is known as RNA splicing<sup>12</sup> and is depicted in Figure 28 on page 28.



**Figure 28.** mRNA splicing (redrawn from [www.unm.edu](http://www.unm.edu)).

mRNA, once prepared, immediately undergoes a modification at the 5'-end known as capping.<sup>12</sup> The first nucleotide triphosphate incorporated into the RNA simply accepts the nucleotide on its 3'-OH group, which results in the initial triphosphate being unchanged.<sup>12</sup> The capping process removes the terminal triphosphate and replaces it with a GMP residue from GTP, the 5'-5'triphosphate which results, is extremely rare in nature [Figure 29].<sup>12</sup> The guanine nucleobase is then methylated at the *N7* position and a methyl group is added to the sugar 2'-OH of the second, and occasionally third, nucleotide.<sup>12</sup> The purpose of this 'cap' is to both protect the mRNA from exonucleases and to facilitate the initiation of translation.<sup>12</sup>



25

**Figure 29.** Structure of the 5'-cap in eukaryotic mRNA.

Unlike DNA, which to a large extent is immortal in cellular terms, mRNA has a half-life of twenty minutes to several hours.<sup>6</sup> This means that mRNA has to be continuously produced.<sup>6</sup> Destruction of mRNA molecules in the cytoplasm marks the end point of the synthesis of the specific proteins encoded by the gene.<sup>6</sup>

mRNA is not the only form of RNA that is produced from transcription of a gene: Ribosomal RNA (rRNA) for use in the synthesis of ribosomes (see the later section 1.5.2, page 30); transfer RNA (tRNA) used to transport amino acids to the growing polypeptide chain (see the later section 1.5.2, page 30); small nuclear RNA (snRNA) used in the processing of the primary transcripts produced from transcription of mRNA, rRNA and tRNA to functional molecules exported to the cytosol; small nucleolar RNA (snoRNA) that have a variety of functions and are located in the nucleolus of the cell; and microRNA (miRNA) that regulate the expression of mRNA molecules are also synthesised through transcription.<sup>12</sup> However, it is mRNA that is translated into a polypeptide.

### **1.5.2 Translation**

The second step in the sequence of genes directing the synthesis of proteins is translation and is summarised in Figure 30 below.

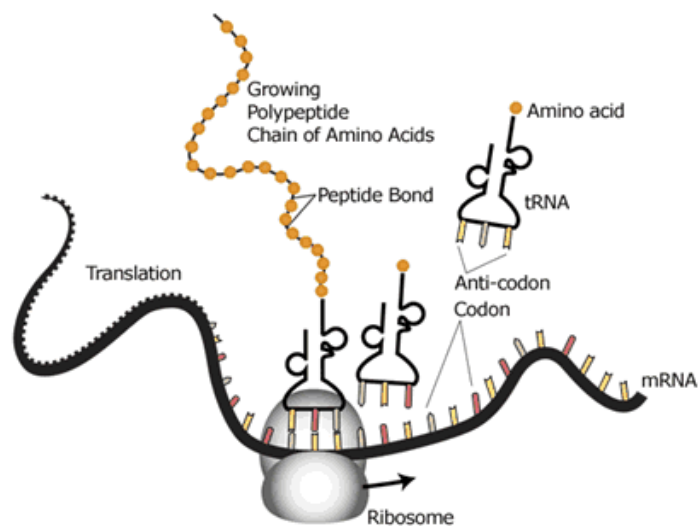


Image adapted from: National Human Genome Research Institute.

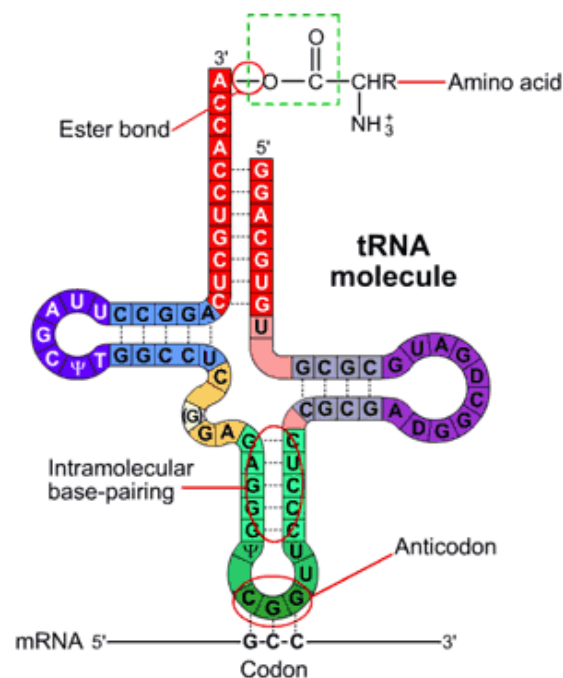
**Figure 30.** Schematic diagram of translation (taken from [www.le.ac.uk](http://www.le.ac.uk)).

Translation is a three step process involving the use of mature mRNA, ribosomes and tRNA. Proteins are synthesised from twenty different varieties of amino acids and the sequence of bases in the mRNA, designates the sequence of amino acids in the protein.<sup>12</sup> A codon is a triplet of bases which represents each amino acid included on the mRNA.<sup>12</sup> Identification of the codons responsible for each amino acid can be found in the genetic code, a list of all these constituents<sup>12</sup> [Appendix A].

Proteins are synthesised on particles called ribosomes.<sup>12</sup> In every cell, there is a large quantity of these multi-subunit particles, they are ~20 nm in diameter and are composed of both ribosomal rRNA and a ribosomal protein, called ribonucleoprotein (RNP).<sup>6</sup> rRNAs conform to a highly folded, compact structure and are associated with the formation of numerous proteins.<sup>6</sup>

Due to no direct link, either physically or chemically, being established between the association of a codon and its corresponding amino acid, it was postulated that an adaptor molecule known as tRNA was involved.<sup>12</sup>

tRNAs are small RNA molecules of less than 100 nucleotides in length.<sup>12</sup> They constitute a cloverleaf structure, with the unpaired nucleic bases arranged as a direct consequence of the internal base-pairing that exists. The complete structure is called an anticodon and is depicted in Figure 31 below.



**Figure 31.** *The cloverleaf structure of tRNA (taken from [www.interscience.wiley.co.uk](http://www.interscience.wiley.co.uk)).*

The tRNA molecules have two main sections; an anticodon and an amino acid accepting site.<sup>12</sup> The anticodon is situated at a hairpin bend and is a triplet of unpaired bases complementary to a codon.<sup>12</sup> These bases are therefore available for hydrogen bonding to a complementary codon on mRNA.<sup>12</sup> At the 3'-accepting site, the sequence CCA marks the point at which the ribose sugar becomes covalently bound to the amino acid through the formation of an ester linkage.<sup>12</sup>

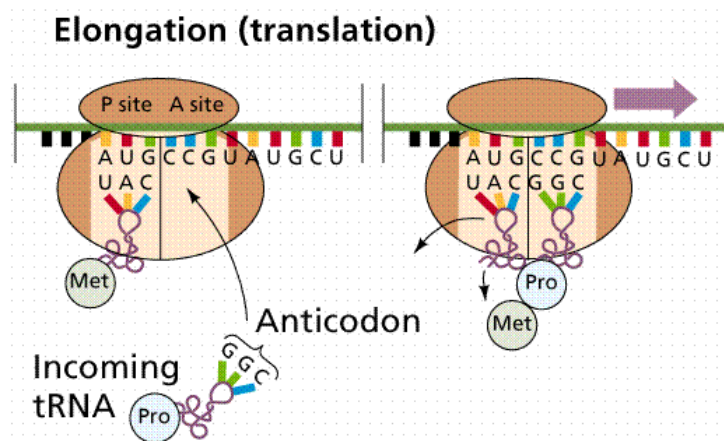
The synthesis of a protein molecule can be divided into three phases; initiation, elongation and termination.<sup>12</sup>

*Initiation*

As discussed in section 1.5.1 on page 28, mRNAs are 'capped' with a methylated guanine nucleotide at the 5' end.<sup>12</sup> A collection of eukaryotic initiation factors (EIF) and an amino acid containing-tRNA (in this case Met) [Figure 32], attach to the cap, and are joined by a 40 S ribosomal subunit.<sup>12</sup>

The ribosome in the initiation complex is situated at the cap site, some distance from the AUG start codon.<sup>12</sup> The 40 S subunit complex advances along the mRNA to the initiating AUG codon and at this point, the 60 S subunit joins the complex which completes the initiation process.<sup>12</sup> After initiation is achieved, polypeptide synthesis proceeds involving two cytosolic elongation factors.<sup>12</sup>

*Elongation*



**Figure 32.** *Elongation in eukaryotes (taken from www.emcmaricopa.edu).*

---

The P site only accepts tRNAs charged with an amino acid from the initiation process [Figure 32].<sup>12</sup> All subsequent aminoacyl-tRNAs enter the A site after being complexed in the cytosol with the elongation factor EF1 $\alpha$ .<sup>12</sup>

The aminoacyl groups on the two tRNA molecules in the P and A sites are in the vicinity of the catalytic site of the enzyme peptidyl transferase. This RNA molecule is responsible for transferring the amino acid in the P site (Met) to the free amino group of the aminoacyl-tRNA in the A site (Pro), the result of which is a dipeptide (Met-Pro) attached to the tRNA [Figure 32].<sup>12</sup>

Once this first peptide bond is synthesised, the A site is filled with a peptidyl-tRNA and the P site contains the uncharged tRNA.<sup>12</sup> The ribosome is repositioned one codon further along the mRNA in a process known as translocation.<sup>12</sup> The discharged tRNA is transported to a third exit site and the peptidyl-tRNA in the A site is moved to the P site.<sup>12</sup> The elongation process continues in an infinite fashion until a stop codon is encountered on the mRNA.<sup>12</sup>

### *Termination*

At the end of the mRNA, one of three stop codons; UAG, UAA or UGA is encountered.<sup>12</sup> Once this occurs, a specific cytoplasmic protein release factor attaches and evokes the release of the final protein from the tRNA.<sup>12</sup> Hydrolysis of the ester bond occurs and the ribosome detaches from the mRNA and dissociates into subunits ready for the next initiation process.<sup>12</sup>

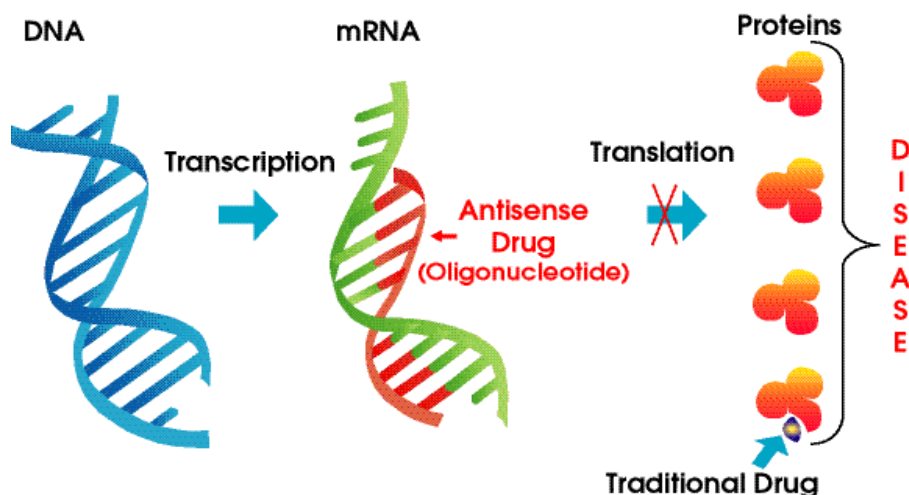
## **1.6 METHODS FOR CONTROLLING GENE EXPRESSION**

Gene expression can be controlled in two ways using either antisense therapy or antigene therapy.

### ***1.6.1 Antisense therapy***

Antisense oligonucleotides are short single-stranded DNA molecules, of complementary sequence to a target mRNA strand, which are capable of specifically binding to and thereby inhibiting translation of this mRNA [Figure 33].

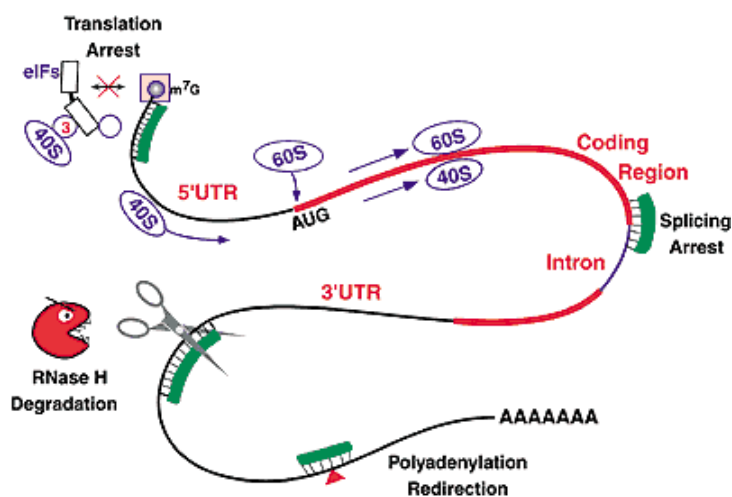




**Figure 33.** Schematic diagram of Antisense therapy (taken from [www.csbsju.edu](http://www.csbsju.edu)).

In the 1970s, it was discovered that an oligodeoxynucleotide, bearing a complementary sequence to part of the Rous Sarcoma virus's mRNA, was capable of preventing its translation and, therefore, its viral production.<sup>13</sup> These findings afforded researchers with a new drug-design strategy to treat diseases that arose from over-expression of unwanted genetic information.<sup>14</sup> Thus, providing the sequence of the gene is known, then a synthetic oligonucleotide can be prepared to form a stable Watson-Crick base paired duplex with the target mRNA and, subsequently, avert its translation.

Once bound to the mRNA target, inhibition can occur through three main modes of action; cleavage by RNase H, translational arrest and directed alternative splicing.<sup>15</sup> Figure 34 depicts these modes of action.



**Figure 34.** Translation modes of action (taken from [www.journalsprou.com](http://www.journalsprou.com)).

---

Of these modes of action, gene inhibition is substantially due to endogenous RNase H. This ubiquitous endonuclease recognizes mRNA-oligonucleotide duplexes and cleaves the RNA strand, leaving intact the antisense strand, which is then free to bind other molecules of the mRNA target.<sup>15</sup> Due the antisense strand remaining intact during this process, an antisense approach to exploit this mode of action would be catalytic.<sup>15</sup>

Translational arrest [Figure 34] results in hybrid arrest or the inhibition of the initiation complex (see the previous section 1.5.2, page 31) and therefore, translation is halted.<sup>15</sup>

Directed alternative splicing [Figure 34] involves the re-establishment of correct splicing sites by oligonucleotides that target incorrect mutant splicing sites.<sup>15</sup> The result of this process is to alter the protein produced and not, as in the case of the other two modes of action, prevent its production.<sup>15</sup>

Although inhibition of gene expression is predominantly the biological mechanism, non-antisense effects may also contribute to the overall response.<sup>15</sup>

The main targets for antisense therapy include numerous cancers, viral infections and some maladies that start from parasites and bacteria.

Any potential modified oligonucleotides designed to act as antisense drugs, have to fulfil the following criteria:

- (1) bind to target mRNA with high affinity and sequence selectivity;
- (2) be resistant to enzyme degradation;
- (3) be soluble in aqueous media;
- (4) be able to penetrate cells, and;
- (5) be able to halt the *in vivo* translation of mRNA.

Requirement four has proved problematic with respect to transporting the antisense oligonucleotides to their specific target. The preferred methods for a synthetic oligonucleotide to pass through the cell membrane in cell cultures (*in vitro*) are: (i) by active or passive transport;<sup>13, 16, 17</sup> (ii) synthesis of antisense RNA within the target cell,<sup>18</sup> or (iii) by external microinjection into a single cell. Although such techniques are

---

worthwhile under laboratory conditions, none have valuable or practical applications *in vivo*.

Another limitation of antisense oligonucleotides is that they tend to be easily degraded by nucleolytic enzymes.

The flexibility and selectivity of antisense therapy makes it a valuable research tool and an exciting new class of therapeutic compounds for serious diseases with known genetic etiologies but poor therapeutic alternatives.<sup>19</sup>

### **1.6.2 Antigene therapy**

In the antigene strategy, gene expression is specifically inhibited at the transcriptional level through a DNA-binding agent. The main advantage of antigene therapy over the antisense strategy (see previous section 1.6.1, pages 32-34) is that there are only two copies of the target sequences of DNA per diploid cell, compared to thousands of copies of the target mRNA sequence.<sup>20</sup> This potentially reduces the quantity of DNA-binding agent required for activity.<sup>20</sup> As a result, this should increase efficacy and reduce side effects. Further, by blocking transcription the re-population of mRNA is prevented and therefore the DNA-binding agent would not have to be re-administered.<sup>20</sup>

Unfortunately, the potential benefits of antigene therapy are currently counteracted by several common restrictions that apply to all oligonucleotide-based treatments.<sup>20</sup> These include bioinstability, poor cellular permeability, and non-specific binding of the oligonucleotides used.<sup>20</sup>

Despite these disadvantages, the capability for many cellular functions to be controlled by antigene therapy is still very attractive. Antigene therapy forms the basis from which the work reported in this thesis has developed.

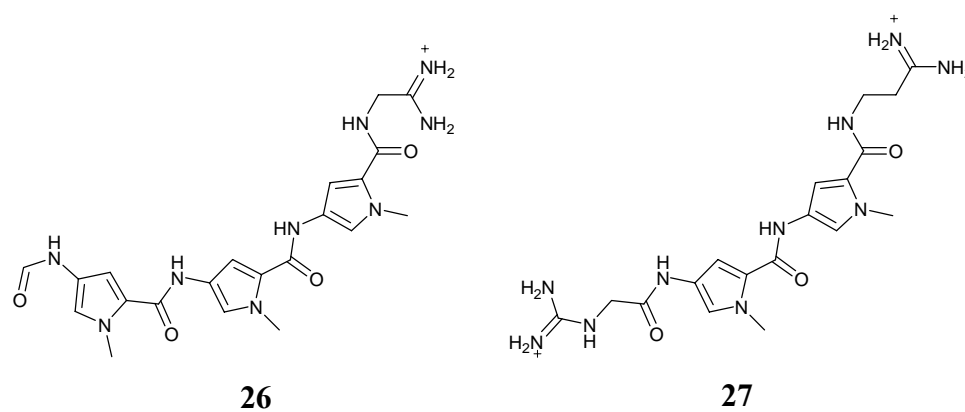
## **1.7 GROOVE BINDERS**

The design and synthesis of molecules capable of gene discrimination and down-regulating the expression of any required gene(s) associated in a disease process, would be a major accomplishment. Historically, there have been two broad approaches

described in the literature to achieve this: (1) to devise molecules that interact non-covalently in the minor groove of DNA, and; (2) to use oligonucleotides targeted to the major groove of DNA, which form triple helical structures.

### 1.7.1 Minor groove binders

Several natural products have been found to bind in the minor groove of DNA with some degree of sequence specificity. Although the structures of such groove binders are diverse they all share a common feature, a crescent shape. Examples of minor groove binders with antitumour capabilities are distamycin and netropsin [Figure 35].<sup>21</sup>



**Figure 35.** Structures of distamycin **26** and netropsin **27**.

Although distamycin and netropsin are too toxic for clinical use, they have been extensively studied.<sup>22</sup> They are characterised by the presence of an oligopeptidic pyrrolocarbamoyl frame ending with an amidino moiety, which reversibly binds to the minor groove of DNA by hydrogen-bonding, Van der Waals contacts and electrostatic interactions, with a strong preference for adenine-thymine (AT)-rich regions containing at least four AT base pairs.<sup>21</sup> They block the binding of transcription factors required for the initiation of transcription by RNA polymerase (see the previous section 1.5.2, page 31).<sup>21</sup>

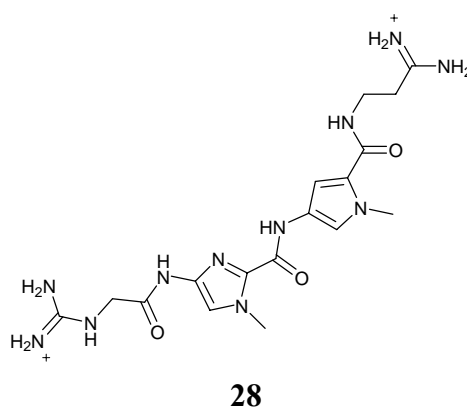
Their mode of binding was first revealed from an x-ray crystal structure of netropsin binding to a B-DNA dodecamer of sequence d(CGCGAATTCGCG).<sup>22</sup> Binding in the minor groove caused the displacement of water molecules from the spine of hydration, which aid the stabilisation of B-DNA, and were replaced with the covalently bonded backbone chain of the drug molecule.<sup>22</sup> The amide NH participates in hydrogen-

bonding, to bridge the DNA, with the *N3* of adenine and *O2* of thymine present on adjacent base pairs and opposed helical strands.<sup>22</sup> As a consequence of the narrowness of the minor groove, netropsin was found to locate itself symmetrically in the midpoint of the groove with the two pyrrole rings, to some extent, non-coplanar.<sup>22</sup> Drug binding forces the minor groove to open by 0.5-2.0Å and causes the helical axis to bend by 8° at the binding region.<sup>22</sup> Although netropsin has an intrinsic twist that favours insertion into the minor groove, this is further enhanced upon binding, to produce the final structure in which the two pyrrole groups are twisted 33° relative to one another.<sup>22</sup>

The structure of distamycin is different to netropsin in that the guanidinium tail has been replaced by another pyrrole ring and a terminal NH-CHO.<sup>22</sup> Therefore, the optimal distamycin binding site encompasses five base pairs, as opposed to the four observed with netropsin, as a direct result of the additional amide NH.<sup>22</sup>

In an attempt to tailor base specificity, Kopka *et al.* decided to substitute the methylpyrrole in netropsin for methylimidazole.<sup>22</sup> It was hoped that by replacing the ring CH by N, the drug molecule would then favour guanine-cytosine (GC) regions by being able to accommodate the NH<sub>2</sub> on the purine and a hydrogen bond acceptor.<sup>22</sup>

Kopka *et al.*<sup>22</sup> therefore prepared the lexitropsin molecule shown in Figure 36 below.



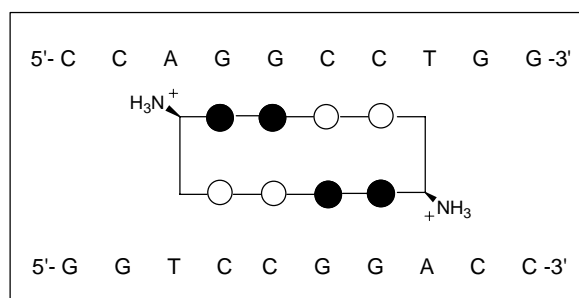
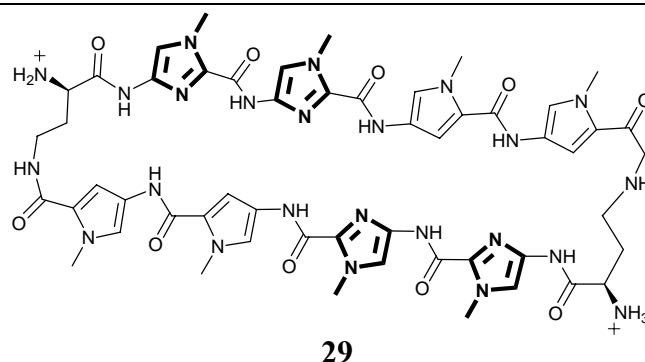
**Figure 36.** Structure of Lexitropsin 28.

This imidazole analogue of netropsin was shown to exhibit a mild decrease in specificity towards AT rich regions, and a concomitant increase in the acceptance of GC sites.<sup>22</sup> This finding implied that new hydrogen bonds formed between the 2-amino group of guanine and the *N3* imidazole nitrogen.<sup>22</sup>

---

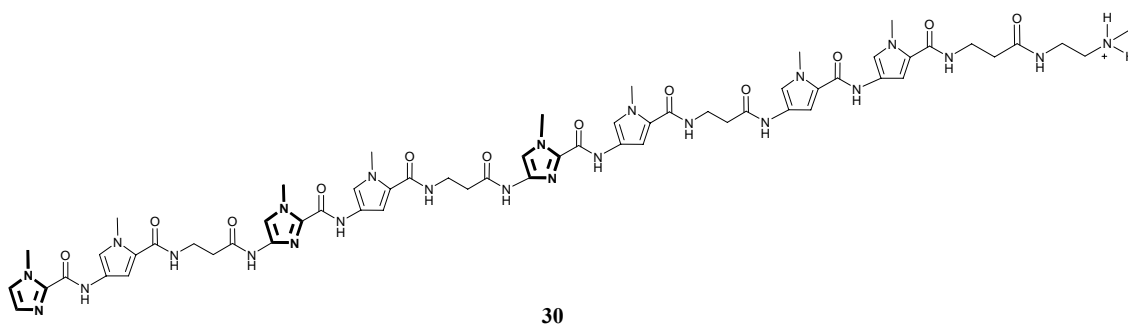
The number of lexitropsins that have since been prepared and evaluated is vast.<sup>23</sup> In general, they all exhibit some ability to recognise GC base pairs, although they also retain AT recognition.<sup>23</sup> The lack of GC specificity was initially ascribed to the differences in molecular electrostatic potential between AT and GC minor groove regions, with the strongly negative potential well of the later preferring the formally dicationic charged netropsin-based lexitropsins.<sup>23</sup> However, recent experimental and theoretical data on the thermodynamics of netropsin binding has indicated that Van der Waals interactions may be responsible instead of electrostatics.<sup>23</sup> In addition, lexitropsins commonly demonstrate reduced overall DNA affinity.<sup>23</sup>

As stated previously, it has been estimated that 15-17 base pairs have to be specifically targeted to inhibit gene expression of a single gene. As the natural minor groove binders' netropsin and distamycin are sequence selective for much shorter segments of DNA, Dervan *et al.*<sup>24</sup> decided to investigate if longer sequences could be targeted using extended distamycin structures. Thus, they prepared a series of six pyrroleimidazole polyamides, containing three to eight rings, designed to bind to 5-10 base pair sites respectively.<sup>24</sup> The DNA binding sites were based on the 5'-TGACA-3' core sequence and contained sequential AT base pair inserts in the central region of the binding site to enable recognition by the additional pyrrolocarboxamides.<sup>24</sup> The results of their study showed that DNA sequences of up to nine base pairs can be specifically recognised by pyrroleimidazole polyamides, containing 3-7 rings, by a 2:1 polyamide-DNA complex formation in the minor groove.<sup>24</sup> The maximum binding affinity was seen for the polyamide containing five rings and the subsequent addition of up to two more rings, had no further effect.<sup>24</sup> The polyamide containing eight rings failed to recognise a ten base pair site.<sup>24</sup> On closer inspection of their x-ray crystal structure, Dervan *et al.* discovered that upon binding of the eight ring polyamide, the minor groove widened by 4Å and consequentially, the major groove was compressed.<sup>24</sup> Further, an 18° bend in the helix towards the major groove was also observed.<sup>24</sup> Figure 37 on page 39 shows the structure of their eight ring pyrrolepolyamide and also a schematic ball and stick model of its binding with the target sequence.



**Figure 37.** Structure of the eight ring pyrrolepolyamide **29** and the schematic ball and stick model of its binding with the binding target.

Since these discoveries in the 1980s, Dervan *et al.*<sup>25</sup> have continued their search for a polyamide capable of recognising and binding to a longer sequence of bases. In 1998, they reported on ImPy- $\beta$ -ImPy- $\beta$ -ImPy- $\beta$ -PyPy- $\beta$ -Dp [Figure 38], a pyrrole-imidazole polyamide dimer that was able to distinguish the 16 base pair sequence 5'-ATAAGCAGCTGCTTTT-3' which is present in the regulatory region of the HIV-1 genome.<sup>25</sup>

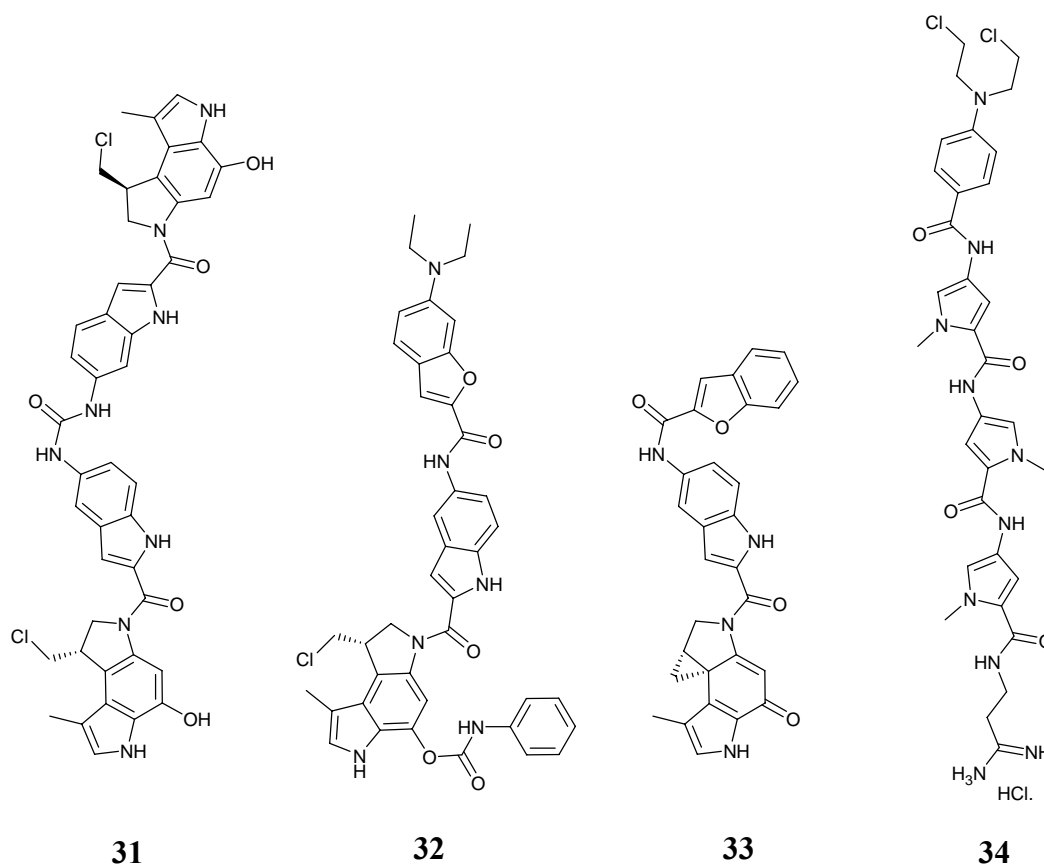


**Figure 38.** Structure of ImPy- $\beta$ -ImPy- $\beta$ -ImPy- $\beta$ -PyPy- $\beta$ -Dp **30**.

Their results showed that this polyamide bound to the target sequence as a cooperative antiparallel dimer.<sup>25</sup> Further, the high binding affinity together with the affinity cleavage pattern observed for the polyamide/DNA complex, implicated that the eight pairs of

amide residues formed a completely overlapped core, which correctly positioned the six Im-Py pairs for recognition of the six GC base pairs and two  $\beta/\beta$  pairs for recognition of the two AT base pairs.<sup>25</sup>

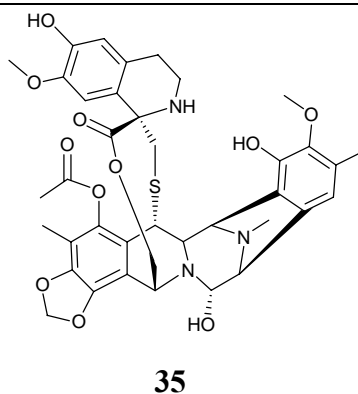
The minor groove binding strategy, thanks to the work by the Dervan group, now has the potential to produce viable antigene therapeutics. However, despite several promising analogues exhibiting antitumour properties in preclinical trials and having entered clinical phase trials, for example, adozelesin,<sup>26</sup> carzelesin,<sup>27</sup> bizelesin<sup>28</sup> and tallimustine<sup>29</sup> [Figure 39], their clinical efficacies have yet to be established.<sup>30</sup>



**Figure 39.** Structures of adozelesin **31**, carzelesin **32**, bizelesin **33** and tallimustine **34**.

An example of a minor groove binder used currently in the treatment of advanced soft tissue sarcoma is trabectedin [Figure 40, page 41].<sup>31</sup>





**Figure 40.** Structure of trabectedin.

Trabectedin was originally derived from the Caribbean marine tunicate *Ecteinascidia turbinata* but can now be prepared synthetically.<sup>31</sup> Although its mode of action is as yet, not clear, it appears to bend DNA away from the site of action, unlike other minor groove binders that bend DNA towards the site of action.<sup>31</sup> Trabectedin is thought to interfere with the transcription-coupled nucleotide excision repair pathway and in addition to this, blocks the G<sub>2</sub>/M phase of the cell cycle.<sup>31</sup> *In vitro* and *in vivo* studies have demonstrated activity against a range of solid tumour cell lines, human tumour explants (soft tissue sarcoma, prostate, ovarian, renal cancers, melanoma, breast and non-small cell lung cancer) and human xenografts.<sup>31</sup>

### 1.7.2 Major groove binders

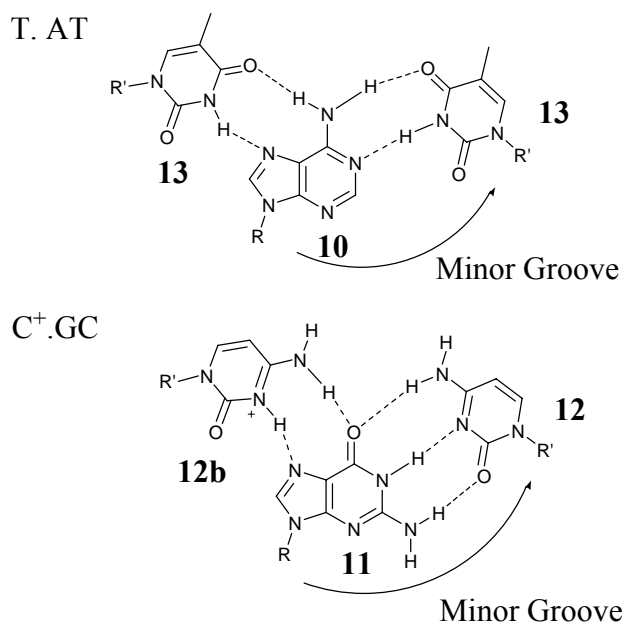
The commonest example of major groove binders are ss-oligonucleotides that are capable of binding to dsDNA through Hoogsteen hydrogen bonding to form triple helices. This ability was first discovered by Felsenfeld *et al.*<sup>32</sup> and much research has since been carried out. The information contained within this sub-section should be regarded as a précis.

Although third strand binding was found to be weaker than double-stranded Watson-Crick binding, it has been shown to be sequence-specific and easily stabilised by divalent cations.<sup>33</sup> It is this potential for sequence-specific binding that has led to an extensive study of molecules capable of forming triple helices. Such triplex targets include genomic duplex DNA and ss-mRNA.<sup>33</sup> Third strand binding to these targets can be achieved by natural and modified ssDNA oligonucleotides and ssDNA

oligonucleotide mimics such as peptide nucleic acids (PNAs) (see the later section 1.9.5, pages 63-65).<sup>33</sup>

Triple-helix forming oligonucleotides (TFOs) are typically 10 to 30 nucleotides in length.<sup>33</sup> Even though the sequence of TFOs can be composed of either polypurine or polypyrimidine molecules, TFOs are limited by their inability to recognise anything other than purine bases.<sup>33</sup> This limitation arises as a direct consequence of purines containing two vacant Hoogsteen hydrogen bonding sites compared to one in the case of pyrimidines and therefore, means that purines are bound more efficiently. The formation of triple-stranded structures is for that reason, restricted to areas of dsDNA which are abundant in homopurine sequences.<sup>34</sup>

A typical triple helix motif occurs as a direct result of a polypurine target duplex being recognised by a pyrimidine-rich third strand.<sup>35</sup> Such an oligonucleotide orientates itself parallel to the polypurine strand in the dsDNA target in the major groove and forms complementary T-A-T and C<sup>+</sup>-G-C Hoogsteen base triplets (where C<sup>+</sup> indicates the N3-protonated cytosine) [Figure 41].<sup>35</sup>

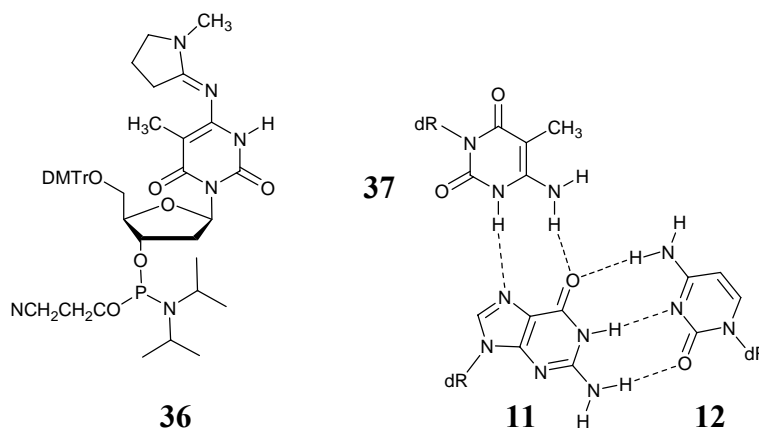


**Figure 41.** Hoogsteen hydrogen bonding.

The T-A-T base triplet involves formation of two Hoogsteen hydrogen bonds between the thymidine in the third strand and the adenine residue in the Watson-Crick A-T base pair. Compared to the C<sup>+</sup>-G-C base triplet involving formation of two Hoogsteen

hydrogen bonds between a protonated cytosine in the third strand and the guanine residue in the Watson-Crick G-C base pair. The prerequisite for prior protonation of the cytosine residue is a pH-dependent process and therefore, restricts the formation of these triplexes in living cells. Further, as a result of the electrostatic repulsion which occurs when two protonated cytosines are located next to each other in the third strand, the ensuing triplex would be destabilised.<sup>36</sup> It is for this reason that several pH-independent cytosine analogues have been designed and synthesised in order to allow for the formation of such parallel-stranded pyrimidine-purine-pyrimidine recognition motifs at physiological pH.

One example is that reported by Xiang *et al.*<sup>37</sup> involving the pyrimidine nucleoside  $m^{5ox}C$  [Figure 42]. This was found to permit G-C base pair recognition under slightly acidic, neutral and slightly basic conditions.<sup>37</sup>

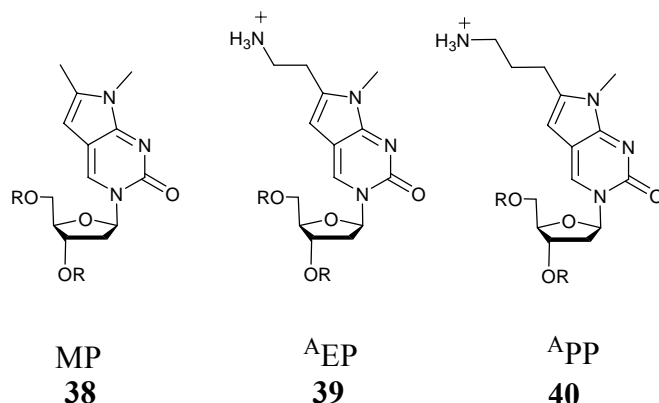


**Figure 42.** Structure of  $m^{5ox}C$  **36** and the fully protected  $m^{5ox}C$  **37** G-C base triplet.

$m^{5ox}C$  includes: (i) A pyrimidine-like ring system to minimize anomalous conformational changes in the backbone of the triplex strand; (ii) a bidentate hydrogen bond donor to interact effectively with the O6-oxygen and N7-nitrogen of the target G-C base pair, and; (iii) a 5-methyl group to contribute to the formation of a potentially important, stabilising, hydrophobic spine.<sup>37</sup> The nucleoside was prepared and then incorporated into a series of 15-mer oligonucleotides, of which the triplex formation was studied using a 25-mer dsDNA target containing four G-C base pairs.<sup>37</sup> The triplex stabilities were assayed at five different pHs i.e. 6.4, 7.0, 7.5, 8.0 and 8.5, and their results showed that for the cytosine-containing 15-mer control reaction, the  $T_m$  values (described in more detail in the later section 1.8 on page 44) for the resulting triplex decreased from 39.0 °C at pH 6.4 to 14.0 °C at pH 8.5. However, the  $m^{5ox}C$ -containing

15-mer oligomer, showed pH-independent triplex transitions that varied by only 2 °C over the entire pH range investigated (28 °C at both pH 6.4 and 8.5).<sup>37</sup> The triplex containing m<sup>5ox</sup>C was less stable at all pHs compared to the triplex containing the control third strand.

The ability to develop the potential uses of TFOs as therapeutics is currently hindered by a limited recognition alphabet. In attempt to overcome this problem, Brown *et al.*<sup>38</sup> have recently reported on the three nucleobases shown below in Figure 43.



**Figure 43.** Structures of MP 38, <sup>A</sup>EP 39 and <sup>A</sup>PP 40.

MP was designed primarily to assess the suitability of the 3*H*-pyrrole[2,3-*d*]pyrimidin-2(7*H*)-one core for GC recognition. By incorporating an acyclic amino group, as in the cases of <sup>A</sup>EP and <sup>A</sup>PP, it was hoped that this moiety would increase the hydrogen bonding potential with *N7* or *O6* of guanine, and/or participate in electrostatic interactions with the anionic phosphate groups and, thereby, increase stability.<sup>38</sup>

From their results shown in Table 3 on page 45, the 3*H*-pyrrole[2,3-*d*]pyrimidin-2(7*H*)-one core was found to recognise GC base pairs and the 18-mer TFO 5'TCTCTCTTXXCCTCCTCC-3' formed stable triplexes with the target duplexes 5'-AGAGAGAGGYAGGAGGAGG-3' and 5'-CCTCCTCCTZTTCTCTCT-3'.<sup>39</sup> The control thymine oligomer  $T_m$  was 44.5 °C at pH 5.5 when Y=C and Z=G, compared to MP, <sup>A</sup>EP and <sup>A</sup>PP which had melting temperatures of 46.4, 45.6 and 47.1 °C respectively.<sup>39</sup> Attempts to increase the thermodynamic stability by including pendant protonated amino groups, as in the case with <sup>A</sup>EP and <sup>A</sup>PP, failed to increase the thermodynamic stability of the resulting triplexes.<sup>39</sup> From Table 3, it can be seen that the differences in melting temperatures of the triplexes studied are negligible apart from when Y=C and Z=G.<sup>39</sup>

X	YZ =	$T_m/^\circ\text{C}$			
		AT	TA	GC	CG
T		54.5	39.2	43.7	44.5
MP		42.7	38.7	41.4	46.4
<sup>A</sup> EP		40.4	37.8	40.2	45.6
<sup>A</sup> PP		42.1	37.6	41.4	47.1

**Table 3.** Melting temperatures of T, MP, <sup>A</sup>EP and <sup>A</sup>PP.

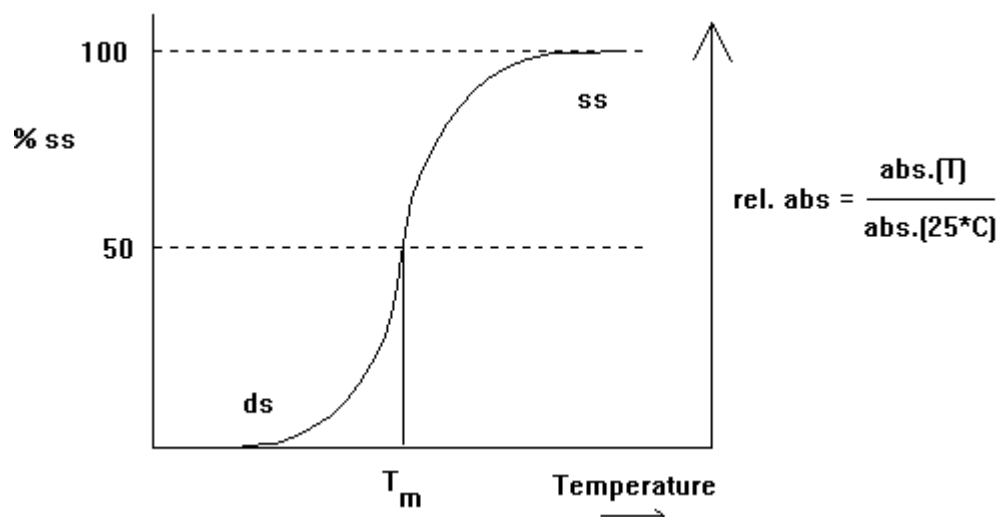
## 1.8 TECHNIQUES FOR DETERMINING THE BINDING AFFINITIES AND SELECTIVITIES OF OLIGONUCLEOTIDE ANALOGUES

The binding affinity and selectivity of any newly synthesised oligonucleotide analogues are normally investigated by melting temperature experiments ( $T_m$ ).

*The melting temperature experiments,  $T_m$*

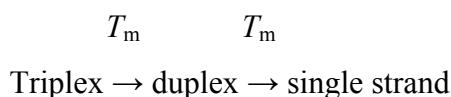
The melting temperature ( $T_m$ ) of a double stranded nucleic acid complex is the temperature at which 50% of the duplex has dissociated into two separate strands.<sup>40</sup> The  $T_m$  depends on the length of the molecule, its specific sequence composition, the concentration of the oligonucleotide and the properties of the solvent that it is contained within.<sup>40</sup> This parameter gives an excellent indication of the degree of affinity that a species has for binding to target oligonucleotides.<sup>40</sup> The  $T_m$  is directly proportional to the guanine-cytosine content within the molecule i.e. the greater the guanine-cytosine content, the higher the  $T_m$  value.<sup>40</sup> This is due to the difference in hydrogen bonding capacities of guanine-cytosine base pairs, compared to adenine-thymine pairs.<sup>40</sup>

A common way to deduce the  $T_m$  of a nucleic acid complex is to use a thermostatted cell in a UV spectrophotometer. A plot of temperature *versus* absorbance affords an S-shaped curve containing two plateaus [Graph 2, page 46]. The temperature is gradually increased over the range 20 °C to 90 °C whilst the absorbance is monitored at a wavelength of 260 nm.<sup>40</sup> The absorbance measurement at the half-way point between the two plateaus yields the  $T_m$  value.<sup>40</sup>



**Graph 2.** Denaturation behaviour of DNA (taken from the what is life website).

Melting temperature experiments can also be used to prove triplex formation. In this case, two  $T_m$  values will be observed as the third strand first dissociates to yield the Watson-Crick duplex, and then further desolves to afford the two single strands [Figure 44].



**Figure 44.** Schematic transition of DNA composition during denaturation.

## 1.9 OLIGONUCLEOTIDE ANALOGUES

### 1.9.1 Natural oligonucleotides

Natural strands of oligodeoxyribonucleotides cannot be used as therapeutic agents due to the fact that their sugar-phosphodiester backbone, is readily hydrolysed by intracellular enzymes, namely nucleases.<sup>6</sup> They are easily degraded and cleared *in vivo* and generally, have poor bioavailability.<sup>18</sup>

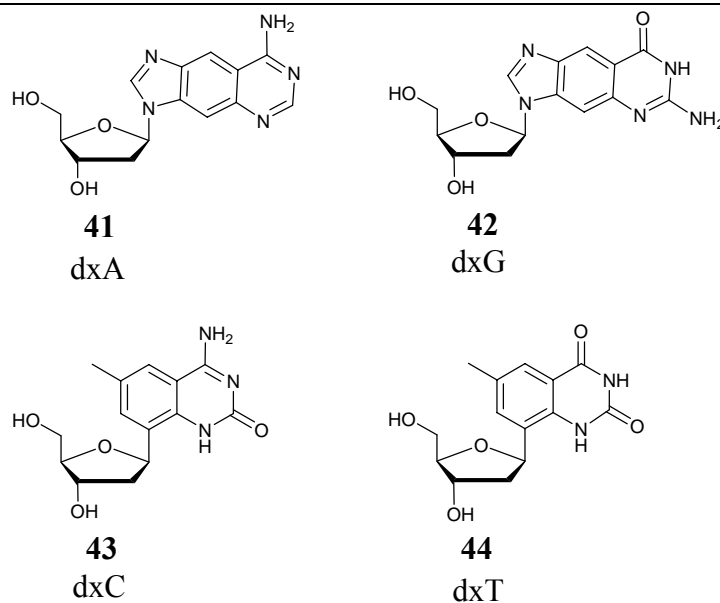
In order to overcome the problems associated with natural oligonucleotides, their analogues have been the focus of much attention in recent years. To date, these studies have typically involved modification of one of the four components of an

oligonucleotide i.e. the base, the sugar moiety, the phosphate group, or the sugar-phosphate moiety. A selection of the most promising modifications reported will be discussed over the next few sections.

### 1.9.2 Base modifications

The inclusion of unnatural nucleobases into oligonucleotides has been extensively researched. For example, Kool *et al.*<sup>41, 42</sup> have created a range of base-expanded nucleobases in order to evaluate their effect on Watson-Crick base pairing.<sup>34</sup> These analogues have been categorised into three main classes: xDNA (“expanded DNA”);<sup>43-48</sup> yDNA (“wide DNA”)<sup>49-51</sup> and yyDNA (“double wide DNA”)<sup>52</sup> and these will be discussed below.

In the case of xDNA [Figure 45 on page 48], the width of the nucleobase was increased by 2.4Å by adding a phenyl ring to each of the four natural bases.<sup>41</sup> Subsequently, the expanded adenine nucleotide dxA, together with the natural thymidine nucleotide dT, was incorporated into the 10-mer oligomer of sequence 5'-d(xATxAxATxATTxAT). The hybridisation ability of the oligomer was then investigated and it was found that it formed a more thermodynamically stable duplex with itself compared to the corresponding natural oligonucleotide i.e. the  $T_m$  for xDNA was found to be 54.9 °C compared to 20.1 °C for the natural dsDNA.<sup>53</sup> Even though the helical diameter of this modified 10-mer homoduplex had increased by approximately 3.0Å (as shown by CD studies), NMR solution studies uncovered that it bore several structural similarities relative to natural DNA.<sup>41</sup> In particular, the homoduplex containing the expanded xA nucleobase was found to be of B-DNA type, the helix was right-handed, the deoxysugar conformations were C<sup>2'</sup>-*endo* and the glycosidic bonds were in the *anti* conformation.<sup>34, 41</sup>



**Figure 45.** Expanded nucleobases (*xDNA set*).

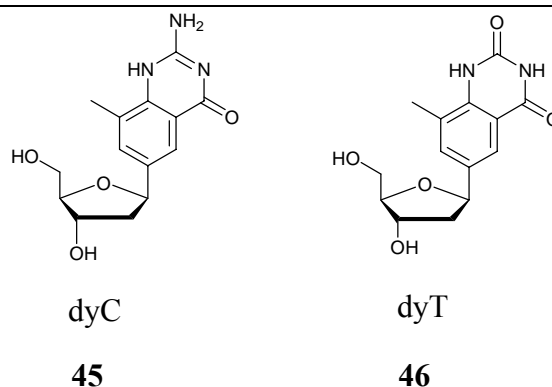
Regarding studies of yDNA, the benzohomologated nucleotide dyC [Figure 46, page 49] has been incorporated into an 18-mer self-complementary oligonucleotide sequence as shown in Table 4 below.<sup>54</sup> Comparisons with the corresponding natural dsDNA showed that the yDNA complex was more stable ( $\Delta T_m = 17.7$  °C).<sup>34, 41</sup> and further, studies have shown that yDNA selectivity was also greater than that of the natural DNA control.<sup>54</sup>

When the oligomer containing dyT was investigated, it too was found to form a complex with increased stability [Table 4].<sup>54</sup> This complex exhibited a  $T_m$  which was 43 °C higher than natural DNA and had an estimated Gibbs' free energy which was 11 kcal mol<sup>-1</sup> more favourable.<sup>34, 41</sup>

	<u>Sequence</u>	<u><math>T_m</math>/°C</u>
yDNA	5'-yC <sup>y</sup> C <sup>y</sup> CG <sup>y</sup> CGGG 3'-GGG <sup>y</sup> CG <sup>y</sup> C <sup>y</sup> C <sup>y</sup> C	71.1
DNA	5'-CCCGCGGG 3'-GGGCGCCC	53.4
yDNA	5'-yT <sup>y</sup> T <sup>y</sup> TAA <sup>y</sup> TA <sup>y</sup> TAA 3'-AAA <sup>y</sup> T <sup>y</sup> TA <sup>y</sup> TA <sup>y</sup> T <sup>y</sup> T	58.2
DNA	5'-TTTAATATAA 3'-AAATTATATT	15.2

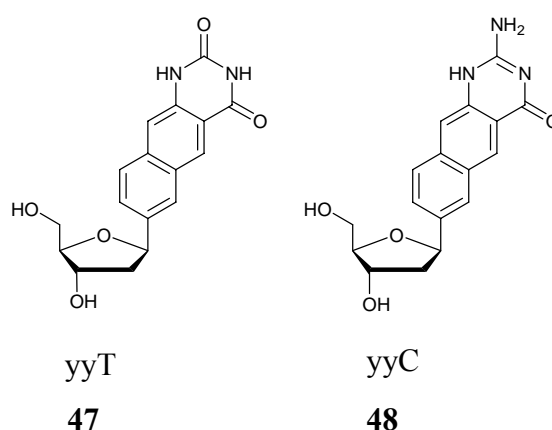
**Table 4.** yDNA containing dyC, dyT and control DNA sequences with their  $T_m$ .





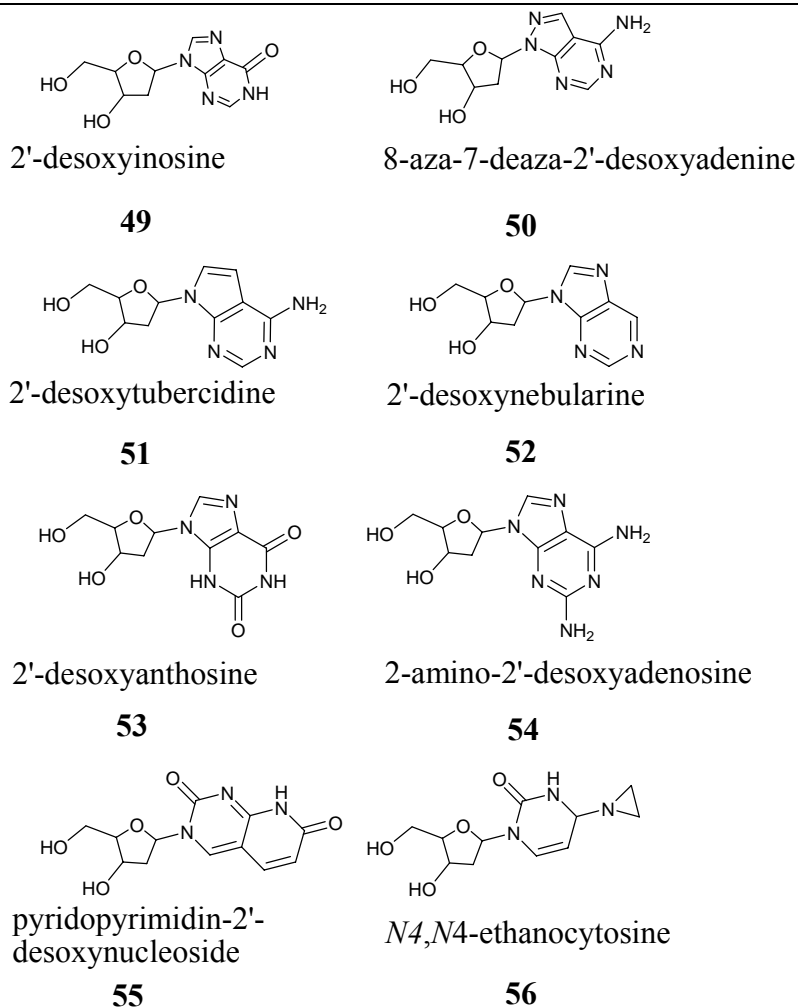
**Figure 46.** Structures of dyC **45** and dyT **46**.

yyDNA is derived from the class of naphtho-homologated nucleobases.<sup>34</sup> This set of compounds represents the limit, in terms of size, that a DNA-like molecule can have.<sup>41</sup> dydT and dydC [Figure 47] have been incorporated into an oligonucleotide of sequence 5'-dXCGCGCG, where X is either dydT or dydC respectively. The corresponding  $T_m$  of the oligomers when bound to 3'-AGCGCGC and 3'-GGCGCGC, were 66.3 °C and 60.8 °C respectively.<sup>55</sup> The doubly widened nucleobases were found to stabilise the resulting duplexes by 1.4 - 1.6 kcal/mol (37 °C) compared to their smaller yDNA homologues.<sup>55</sup> Surprisingly, yyT and yyC were also found to show negligible sequence selectivity towards the four natural bases.<sup>55</sup> The reason for this is thought to be due to their larger size which may prevent Watson-Crick-like base pairing from occurring.<sup>55</sup>



**Figure 47.** Structures of dydT **47** and dydC **48**.

In addition to these expanded nucleobases reported by Kool *et al.*<sup>41</sup> a vast number of other base-modified nucleosides have been reported in the literature for inclusion into oligonucleotide analogues. A selection of these modified bases are shown in Figure 48 on page 50.



**Figure 48.** A selection of base-modified nucleosides.<sup>40</sup>

In general, the oligomers bearing the above base-modified nucleosides have been found to form less stable duplexes with complementary oligonucleotides than those formed from natural oligonucleotides of the same sequence.<sup>40</sup> As a result of this, their incorporation into antisense or antigene therapeutics is hindered.<sup>40</sup>

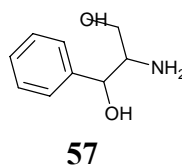
The base modifications described in this section should be regarded as a précis and by no means a comprehensive outline of current research in this area. For a more detailed discussion, the reader is referred to the review article by Uhlmann *et al.*<sup>40</sup>

### 1.9.3 Sugar modifications

There are many reasons as to why the deoxyribose component of a nucleic acid is a popular site for modification. It is hoped that manipulation at this site will lead to increased enzymatic stability, improved nucleic acid-protein binding and, generally,

additional functionality in DNA. Thus, there is an enormous quantity of literature available on sugar modifications which have been investigated to date.<sup>34,56</sup> The information contained within this sub-section should be regarded as a brief outline, with specific focus on acyclic sugar modifications as these have direct bearing on the research work reported in this thesis.

In an attempt to overcome the common problem reported in the literature, of acyclic oligonucleotide analogues not forming duplexes with natural oligonucleotides of acceptable stability, Rana *et al.*<sup>57</sup> decided to impart backbone rigidity into serinol derived acyclic analogues by introducing substituents in the acyclic backbone.<sup>57</sup> They therefore, synthesised and studied oligomers bearing either the (1*R*,2*R*)- or (1*S*,2*S*)-stereoisomers of the serinol derived acyclic analogue **57** [Figure 49].<sup>57</sup>



**Figure 49.** Structure of 2-amino-1-phenyl-1,3-propanediol.

The oligonucleotides examined are shown in Table 5 below. Oligomer (1) is the fully unmodified 18-mer control whereas oligomers (2)-(5) contain the serinol derived monomer **x**.

#### Oligonucleotide sequences

(1) 5'-TTCTTCTTCTTTTCTTTT-3'

(2) 5'-TTCTTCTTCTTTTCTT**x**T-3'

(3) 5'-TTCTTCTT**Cx**TTTCTTTT-3'

(4) 5'-T**x**CTTCTTCTTTTCTTTT-3'

(5) 5'-TTCTTCTTCTTTT**Cxxx**T-3'

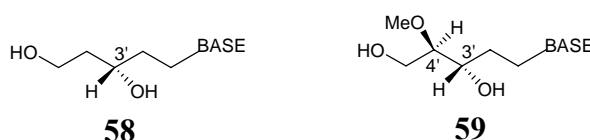
(**x** = modified monomer)

**Table 5.** Oligonucleotide sequences.

The abilities of these oligomers to form duplexes and triplexes were explored using UV melting experiments. From their results, it was found that incorporating the chiral,

acyclic phenyl derivative into the oligonucleotides did not hinder the formation of either duplexes or triplexes.<sup>57</sup> These findings also demonstrated that the *SS*-diastereoisomer was slightly better than the corresponding *RR*-stereoisomer with respect to forming stable triplexes with dsDNA at pH 5.8 ( $\Delta T_m = 4$  °C).<sup>57</sup> The reason for their modified oligonucleotide not interrupting the formation of stable duplexes and triplexes is reportedly due to the phenyl substituent enforcing conformational rigidity of the acyclic chain, and also by it causing local hydrophobic dissolution in the major groove.<sup>57</sup> Further, the location in the major groove of the phenyl substituent in the *SS*-diastereoisomer may be more favourable than that corresponding to the *RR*-stereoisomer.

Vandendriessche *et al.*<sup>14</sup> have incorporated two different types of acyclic nucleosides into oligonucleotides; those with the 3*S*,5-dihydroxypentyl side chain **58** and those with the 4*R*-methoxy-3*S*,5-dihydroxypentyl side chain **59** [Figures 50].



**Figure 50.** Structure of the 3*S*,5-dihydroxypentyl side chain **58** and 4*R*-methoxy-3*S*,5-dihydroxypentyl side chain **59**.

Both acyclic nucleosides lack the C<sub>4</sub>'-O-C<sub>1</sub>' moiety of natural DNA sugar. Their reason for investigating these nucleosides was that it had been postulated that the ribose sugar oxygen may be involved in the recognition processes of both *exo*- and *endonucleases*. Thus, by removing it, it was hoped that the nuclease stability of the resulting acyclic derivative would be enhanced. In the second acyclic nucleoside **59** studied, the ribose sugar oxygen is disguised within a methoxy group.<sup>14</sup> A common feature of both acyclic nucleosides **58** and **59** is the lack of a glycosidic linkage.<sup>56</sup> However, both the number of bonds between the nucleobase and the primary and secondary hydroxyl groups, and the configuration of the stereocentres are consistent with those found in natural 2'-deoxy-D-ribose.<sup>56</sup>

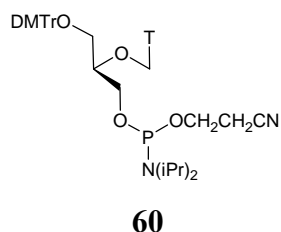
Both acyclic nucleosides were incorporated into oligomers and their nuclease stabilities were studied. Their results showed that both a 13-mer and 12-mer homoadenine oligomer, composed entirely of the 3*S*,5-dihydroxypentyl or 4*R*-methoxy-

3*S*,5-dihydroxypentyl-nucleosides respectively, were stable towards snake venom phosphodiesterase for more than two hours.<sup>56</sup> The corresponding unmodified natural homoadenine oligonucleotides, when treated under the same conditions, showed complete degradation after 3.1 minutes.<sup>56</sup> In addition to this result, the 3*S*,5-dihydroxypentyl-nucleoside when incorporated at the centre of their oligonucleotide, demonstrated hyperchromicity values of approximately 50% that demonstrated by the unmodified oligonucleotide.<sup>56</sup> In attempt to understand their finding, they postulated that the 13-mer oligomer was being degraded to a 7-mer oligomer, which was subsequently stabilised by the presence of the 3*S*,5-dihydroxypentyl-nucleoside at its 3'-end.<sup>56</sup>

The ability of heteroligonucleotides bearing acyclic nucleosides **58** and **59** to hybridise to complementary ss-oligonucleotides has been studied using melting temperature experiments. The data showed that the 13-mer homo-adenine oligomer containing a single 4*R*-methoxy-3*S*,5-dihydroxypentyl-nucleoside at the centre had a  $T_m$  of 26.7 °C when hybridised to the complementary unmodified oligomer.<sup>56</sup> The corresponding control experiment had a  $T_m$  of 33.3 °C, indicating that incorporation of this analogue was destabilising ( $\Delta T_m = -6.6$  °C). A lesser degree of destabilisation was observed when two 4*R*-methoxy-3*S*,5-dihydroxypentyl-nucleosides were incorporated at each end of the homo-adenine oligomers as the melting temperature observed was 30.9 °C which afforded a net  $\Delta T_m$  of -2.4 °C.<sup>56</sup> Further, the 4*R*-methoxy-3*S*,5-dihydroxypentyl analogues were shown to destabilise the resulting *S*-heteroduplex formed with the corresponding ss-oligonucleotide, slightly more than the 3*S*,5-dihydroxypentyl analogues.<sup>56</sup> Incorporation of a single 3*S*,5-dihydroxypentyl-nucleoside at the centre of the 13-mer homo-adenine oligomers had a  $T_m$  of 28.6 °C ( $\Delta T_m = 1.9$  °C compared to the 4*R*-methoxy-3*S*,5-dihydroxypentyl analogue) and a  $T_m$  of 28.5 °C ( $\Delta T_m = 1.7$  °C compared to the 4*R*-methoxy-3*S*,5-dihydroxypentyl analogue) when incorporated at either end of the homo-adenine oligomers.<sup>56</sup>

Vandendriessche *et al.*<sup>14</sup> hypothesised that due to the conformational mobility of their acyclic nucleosides, the base pairing properties may be different from Watson-Crick base pairing and this could contribute to the destabilisation observed.<sup>56</sup> Increased flexibility may prevent the acyclic nucleosides from successfully discriminating between the natural bases.<sup>56</sup>

Schneider *et al.* have prepared and introduced flexible acyclic nucleosides of the type **60** [Figure 51] into the centre of a series of oligonucleotides (Table 6, page 55).<sup>58</sup>



**Figure 51.** Flexible nucleoside *t*.

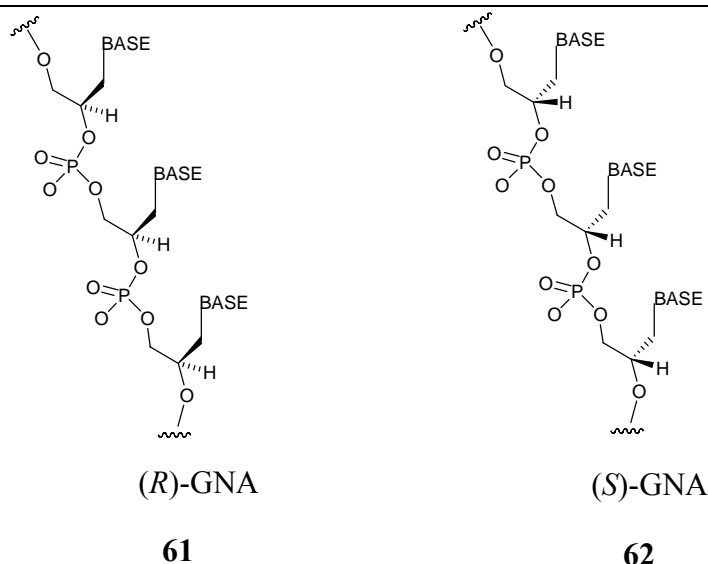
The ability of these oligomers to hybridise to ss-oligonucleotides was studied (Table 6, page 55). Their observations showed that the stabilities of the resulting complexes formed from oligomers bearing the modified acyclic nucleoside **60** decreased relative to the corresponding unmodified dsDNA. The  $T_m$  values of these heteroduplexes were found to decrease by 9-15 °C for each flexible nucleoside incorporated.<sup>58</sup> Although the effect was not additive, duplexes formed from oligomers containing two adjacent, acyclic monomers **60** were slightly more destabilised than for those generated from oligomers containing the two acyclic monomers **60**, separated by a natural nucleoside.<sup>58</sup>

Sequence	$T_m / ^\circ\text{C}$
(1) 5'-CTTTTTTTG-3' 3'-GAAAAAAC-5'	40.0
(2) 5'-CTTTtTTTG-3' 3'-GAAAAAAC-5'	25.0
(3) 5'-CTTtTtTTG-3' 3'-GAAAAAAC-5'	13.0
(4) 5'-CTTttTTTG-3' 3'-GAAAAAAC-5'	11.0
(5) 5'-CTTTTTTTG-3' 3'-GAAAGAAC-5'	21.0
(6) 5'-CTTTTTTTTTTTG-3' 3'-GAAAAAAAAAAC-5'	55.0
(7) 5'-CAAATAAAG-3' 3'-GTTTATTTC-5'	37.0
(8) 5'-CAAAtAAAG-3' 3'-GTTTATTTC-5'	25.0
(9) 5'-CAAAtAtAAG-3' 3'-GTTATATTC-5'	12.0
(10) 5'-CAAAtAAAG-3' 3'-GAAAAAAC-5'	11.0
(11) 5'-CTTTtTTTG-3' 3'-GAAAGAAC-5'	12.0

**Table 6.** Table of sequences and  $T_m$  values.

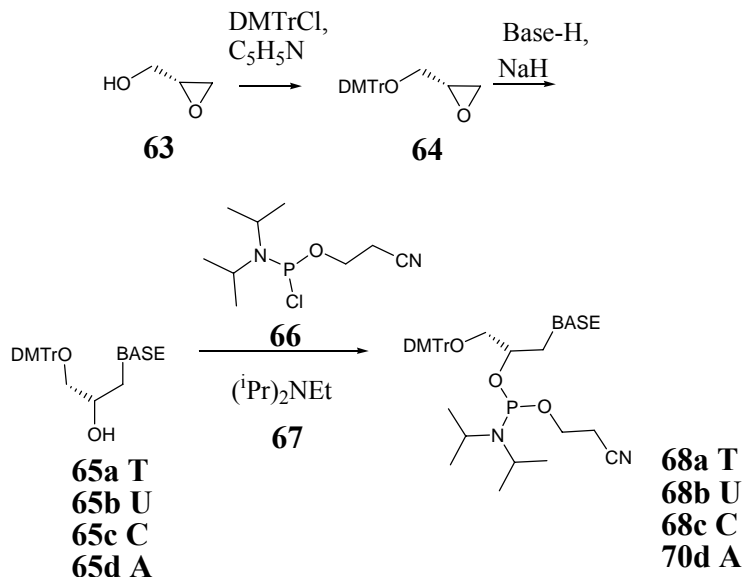
In attempt to explain their findings, they hypothesised that, by incorporating a flexible nucleoside analogue into an oligonucleotide, the ability to form duplex structures was decreased. This was due to the entropy lost upon going from two oligonucleotides to a duplex being larger in a system containing flexible nucleosides, opposed to one containing natural oligonucleosides.<sup>58</sup>

Recently, Meggers *et al*<sup>59, 60</sup> have reported the design and synthesis of glycol nucleic acid (GNA) (Figure 52, page 56).



**Figure 52.** Structures of (R)- **61** and (S)- **62** glycol nucleic acid (GNA).

This propylene glycol phosphodiester backbone is conceivably the most atom economical example of modifications to natural DNA. The acyclic glycol nucleosides were prepared by stereospecific and regioselective nucleophilic ring opening of epoxide **64** by the appropriately protected nucleobase as shown in Scheme 1 below.<sup>60</sup>

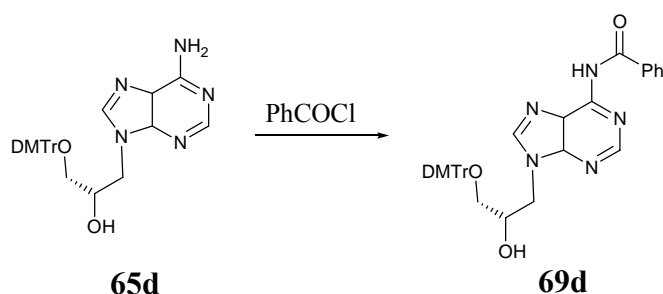


**Scheme 1.** Formation of (S)-T **68a**, U **68b**, C **68c** and A **70d**.

**63** was tritylated to give **64**, which was then ring opened in the presence the appropriate nucleobase, thymine, uracil or *N*4-benzoylcytosine.<sup>59</sup> The pyrimidine nucleosides **65a**, **65b** and **65c** were then directly converted into the corresponding phosphoramidites **68a**, **68b** and **68c**, in the presence of 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite **66**

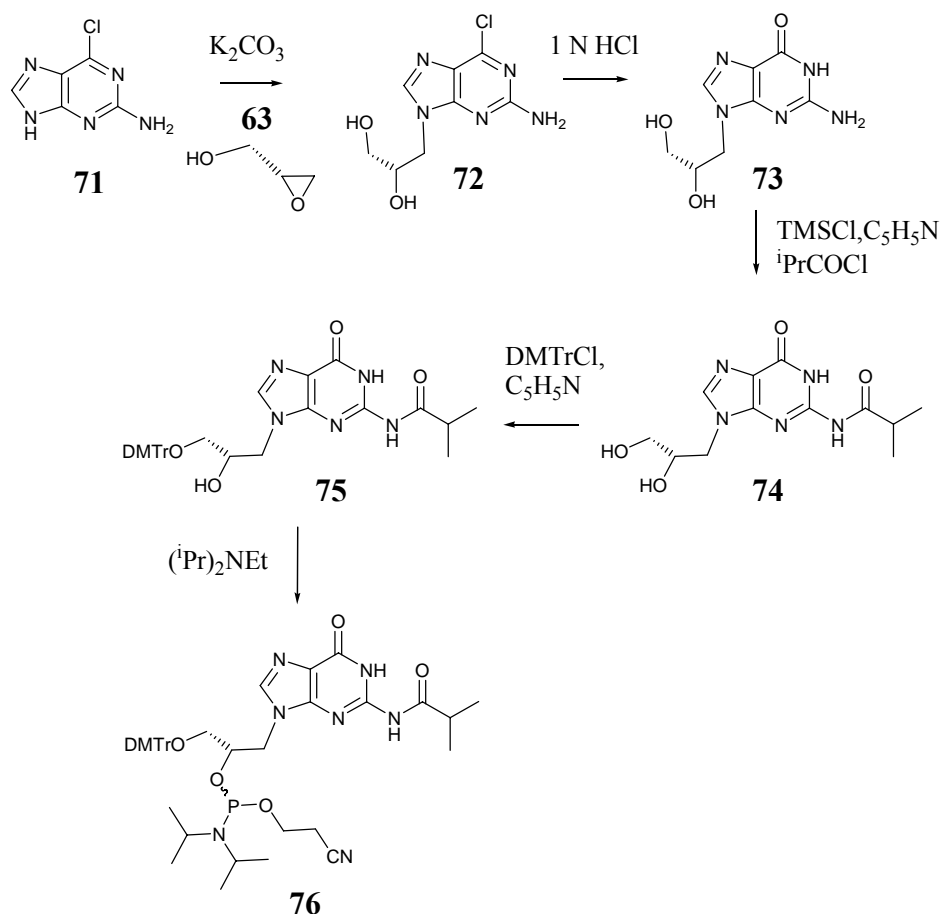


and Hünig's base **67**.<sup>59</sup> The adenine nucleoside **65d** was benzoylated to give **69d** as shown in Scheme 2, and then directly converted to its corresponding phosphoramidite **70d** (Scheme 1, page 56).<sup>59</sup>



**Scheme 2.** Benzoylation of **65d**.

To synthesise the guanine nucleoside **76**, natural guanine was substituted with commercially available 2-amino-6-chloropurine **71** for reasons discussed in more detail in chapter 2, section 2.3.4, page 111.<sup>59</sup> Treatment with **63** gave intermediate **72** to which the exocyclic amino group was subsequently protected with an isobutyryl group to afford **74**.<sup>59</sup> This was then tritylated and treated with Hünig's base as in the previous cases, to afford the corresponding phosphoramidite [Scheme 3].<sup>59</sup>



**Scheme 3.** Synthesis of (*S*)-G **76**.

Subsequently, seven 18-mer oligonucleotides consisting entirely of each enantiomer were prepared in high yields [Table 7].

3'-TTTTAAATTTTAATATAT-2'	<b>a</b>	Antiparallel duplex
2'-AAAATTTAAAATTATATA-3'	<b>b</b>	( $T_m = 63$ °C)
3'-TAAAATTTATATTATTAA-2'	<b>c</b>	Antiparallel duplex
2'-ATTTTAAATATAATAATT-3'	<b>d</b>	( $T_m = 63$ °C)
3'-TAAAATTTATATTATTAA-2'	<b>c</b>	Parallel duplex
3'-ATTTTAAATATAATAATT-2'	<b>e</b>	(No $T_m$ )
3'-TAAAATTTATATTATTAA-2'	<b>c</b>	One mismatch
2'-ATTTTAAATTTAATAATT-3'	<b>f</b>	( $T_m = 55$ °C)
3'-TAAAATTTATATTATTAA-2'	<b>c</b>	Two mismatches
2'-ATTTAAAATTTAATAATT-3'	<b>g</b>	( $T_m = 44$ °C)

**Table 7.** Table of 18-mer GNA oligomers prepared.

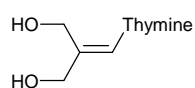
Temperature-dependent UV spectroscopy at 260 nm was carried out to investigate the ability of these GNA oligomers to hybridise to themselves.<sup>59</sup> It was found that, for 1:1 mixtures of two complementary (*S*)-GNA strands i.e. **a** and **b** and **c** and **d** [Table 7], sigmoidal curves with melting temperatures of 63 °C were observed.<sup>60</sup> Furthermore, the thermal stability of the resulting GNA duplex formed from oligomers **c** and **d**, noticeably surpassed the thermal stabilities of both the analogous DNA duplex ( $T_m$  40.5 °C) and RNA ( $T_m$  42.5 °C, where U was used in place of T).<sup>60</sup> This was a surprising result as nucleic acids normally require a cyclic structure (such as a ribose) to impart structural rigidity into the nucleic acid backbone and enable hybridisation to occur.<sup>59</sup> The requirement for the two GNA strands to bind in an antiparallel orientation was demonstrated since no sigmoidal melting or only weak hyperchromicities were observed for solutions containing the single strands of the 18-mer (*S*)- or (*R*)- GNA oligonucleotides alone.<sup>60</sup> In order to confirm that the (*S*)- or (*R*)- GNA homoduplexes relied entirely on the formation of Watson-Crick base pairs, mismatch experiments were performed on an 18-mer GNA duplex derived from oligomers **c** and **f**, which included a

single T:T mismatch [Table 7]. It was found that the  $T_m$  value for this duplex was reduced by 8 °C compared to the corresponding fully complementary homoduplex. When two mismatches were incorporated, one T:T and one A:A, as in the case with oligomers **c** and **g** (Table 7, page 58) the  $T_m$  decreased by 19 °C.<sup>60</sup> Similar results were obtained for the analogous enantiomeric (*R*)-GNA oligomers.<sup>60</sup>

Both the (*S*)-GNA and (*R*)-GNA homochiral duplexes formed from oligomers **c** and **d** [Table 7], have been evaluated using circular dichroism (CD). The CD spectra recorded displayed strongly amplified CD bands at 205, 220 and 275 nm upon mixing of oligomers **c** and **d** in a 1:1 fashion. These findings clearly showed that helical duplexes had formed.<sup>60</sup> It was further found that the CD signal was strongest for a 1:1 ratio of GNA oligomers **c** and **d** and that the Cotton Effect decreased as the temperature increased, additional evidence of homoduplex formation.<sup>60</sup> The analogous (*S*)-GNA and (*R*)-GNA homochiral homoduplexes exhibited mirror image CD spectra and both were distinguishable from the CD spectra recorded for a DNA duplex of the same sequence.<sup>60</sup> Meggers *et al.*<sup>60</sup> have postulated that GNA strands **c** and **d** exhibited a large Cotton Effect due to pre-organisation of the GNA backbone taking place in the single strand prior to hybridisation.<sup>60</sup>

The antiparallel cross pairing between (*S*)-GNA, (*R*)-GNA, DNA and RNA has also been investigated.<sup>60</sup> It was found that complementary (*S*)-GNA and (*R*)-GNA oligomers do not undergo antiparallel cross-pairing with each other, and nor do they form stable antiparallel cross-pairs with complementary ssDNA.<sup>60</sup> Interestingly, however, (*S*)-GNA oligomers were shown to hybridise to complementary RNA. This was apparent from their UV melting temperature experiment curves, for duplexes with the sequences **c** and **d** of RNA, (*S*)-GNA and the RNA/(*S*)-GNA hybrid.<sup>60</sup>

Other acyclic sugar modified nucleosides which have been reported to date for inclusion into oligonucleotide analogues include the acyclic nucleoside  $T^\beta$ , **77** which is conformationally restricted about the C=C double bond [Figure 53].



77

**Figure 53.** Structure of 1-[3-hydroxy-2-(hydroxymethyl)prop-1-enyl]thymine ( $T^\beta$ ).

It was thought that this compound would be an effective mimic of a natural nucleoside according to molecular modelling and geometry calculations.<sup>15, 61</sup> These theoretical studies implied that compound **77** would be able to replace natural thymidines in both A- and B-type DNA helices, with no disruption to their resulting helical structure.<sup>15, 61</sup> However, preliminary hybridisation studies on oligonucleotides in which one or two thymidines had been substituted for **77**, with complementary ssDNA or RNA targets in fact showed reduced binding affinities compared to the corresponding unmodified oligonucleotides.<sup>15</sup> When one T<sup>β</sup> nucleoside was incorporated into the centre of the oligonucleotide of sequence dT<sub>7</sub>T<sup>β</sup>T<sub>6</sub>, the  $T_m$  of the resulting duplex decreased from 36.0 °C to 31.0 °C. This change in  $T_m$  was also observed when T<sup>β</sup> was located near the end of the oligonucleotide (dT<sub>11</sub>TT<sup>β</sup>T).<sup>15, 61</sup> Thus, it can be concluded from this example that merely restricting the conformational freedom of the acyclic nucleoside using a C=C double bond is not sufficient to produce stable hybridisation.

Despite an already vast sum of research having been carried out, there is still scope to improve both resistance to enzymatic degradation and to enhance binding capabilities in this area. Having learnt from the work reported by Vandendriessche *et al.*<sup>14</sup> and Schneider *et al.*<sup>58</sup> in particular, that increased flexibility may lead to both acyclic nucleosides unsuccessfully discriminating between their natural bases, and that decreased ability to form duplex structures may arise due to the entropy lost being larger in a system containing flexible nucleosides, opposed to one containing natural oligonucleosides. Future endeavours within this research area are of benefit.

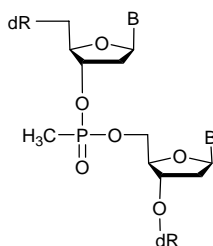
#### **1.9.4 Phosphate modifications**

As stated previously, a major problem that has been encountered in the use of natural phosphodiester oligonucleotides as potential regulators of gene expression is that they are readily degraded ( $t_{1/2} = 30$  minutes) by nucleolytic activities in cells and culture medium.<sup>62</sup> Therefore, in recent years, substantial effort has gone into trying to develop oligonucleotide analogues which show increased resistance towards nucleases. Due to the phosphorus centre being the main site of attack for nucleases, it is not surprising that many modifications have been made to the natural internucleotide linkage.<sup>62</sup> Of the modifications which have been reported to date, methylphosphonates<sup>63</sup> and phosphorothioates<sup>62</sup> have been extensively evaluated. These modifications have been shown to increase stability of the resulting oligomers to both *exo*- and *endo*-enzymatic

degradation.<sup>62</sup> The reviews by Goodchild,<sup>64</sup> Varma<sup>65</sup> and Uhlmann and Peyman.<sup>40</sup> provide excellent accounts of the progress achieved to date in this research field.

(i) *Methylphosphonates*

Methylphosphonates [Figure 54] are oligonucleotide analogues in which the non-bridging oxygen of the natural phosphodiester linkage has been substituted for a methyl group.



78

**Figure 54.** Structure of a methylphosphonate linkage.

They were the first example of an antisense, backbone modified oligonucleotide which displayed increased enzymatic stability compared to the natural counterpart.<sup>66</sup> Methylphosphonates are much more lipophilic than natural oligonucleotides due to their lack of a formal negative charge and this results in increased binding affinities as a direct consequence of having removed electrostatic repulsion.<sup>66</sup>

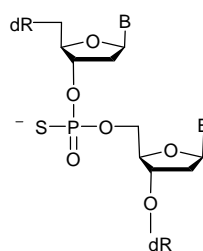
Miller *et al.*<sup>66</sup> have demonstrated that oligomethylphosphonates are accomplished inhibitors of Vesicular Stomatitis Virus (VSV) and type 1 Herpes Simplex Virus (HSV-1) in tissue culture, using an antisense approach. An eight residue oligomer d(TpCCTCCTG) (all nucleotides linked through methylphosphonate linkages apart from the 5'-thymidine residue) was designed to be complementary to the acceptor splice intersection of HSV-1 IE.<sup>66</sup> It was shown that, not only did this 8-mer oligomer inhibit HSV-1 viral replication in a sequence specific and dose-dependent fashion ( $IC_{50} = 25$  Mm;  $IC_{90} = 75$   $\mu$ M),<sup>66</sup> but that it was also stable to nucleases and exhibited increased cell permeability compared to the complementary control.<sup>66</sup> Oligomethylphosphonates enter cells in a passive manner, unlike active transport in the case of charged oligonucleotides.<sup>66</sup>

The use of methylphosphonate oligonucleotides is limited by weak hybridisation.<sup>66</sup> This problem has been overcome by employing oligonucleotides bearing alternating methylphosphonate and phosphodiester linkages.<sup>67</sup> Thus, a chimera of sequence pCpUpCpCpCpGpGpApCpUpCpApGpApU where **p** = methylphosphonate and p = phosphodiester linkage was prepared which was complementary to the upper hairpin region of the HIV virus TAR RNA.<sup>67</sup> The 15-mer was found to form a stable duplex with its complementary RNA target and had a  $T_m$  of 71 °C.<sup>67</sup>

Methylphosphonates have been superseded by phosphorothioates.

(ii) *Phosphorothioates*

Phosphorothioates [Figure 55] are oligonucleotide analogues in which the non-bridging oxygen of the natural phosphodiester linkage has been substituted for a sulphur.



79

**Figure 55.** Structure of a phosphorothioate linkage.

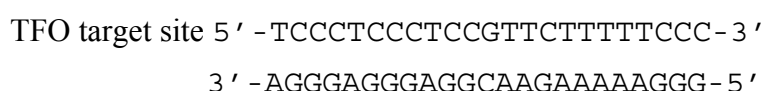
Phosphorothioates are relatively easy to prepare. The phosphorothioate linkage is typically synthesised as a mixture of a pair of diastereoisomers (i.e.  $R_p$  and  $S_p$ ), and these are not subsequently separated.<sup>68</sup> Different nucleases show differing preferences for  $R_p$  or  $S_p$  linkages depending on the nature of the active site of the enzyme.<sup>69</sup>

Phosphorothioates afford the best hybridisation properties of the phosphate modification studied to date.<sup>68</sup> In an antisense approach, Hoke *et al.*<sup>62</sup> prepared a 21-mer oligonucleotide phosphorothioate compound, ISIS 1080, of sequence A\*C\*C\*G\*A\*G\*G\*T\*C\*C\*A\*T\*G\*T\*C\*G\*T\*A\*C\*G\*C which was complementary to the internal AUG codon of UL13 mRNA in HSV-1. They found that by incorporating this oligonucleotide, the ability to form stable hybrids with mRNA was reduced.<sup>62</sup> The  $T_m$  of the control phosphodiester oligonucleotide was 68.9 °C and the  $T_m$  of the modified

oligomer was 61.1 °C.<sup>62</sup> The corresponding  $\Delta T_m$  of 7.8 °C was thought to be as a direct result of incorporating the phosphorothioate linkage.<sup>62</sup> Thus, there was an overall loss in hybridisation stability when ISIS 1080 was incorporated. This loss in stability led to a reduction in the infectious yield of HSV-1 in HeLa cells, to 9.0%  $\pm$  11% by supporting RNase H cleavage of the HSV mRNA to which it was targeted.<sup>62</sup>

Phosphorothioate modified oligonucleotides are known to affect the stability of triplexes, formed from an antigene perspective, with complementary single stranded DNA, in different ways depending on the environment of the modification.<sup>66</sup> Phosphorothioates linkages 5'- to a purine increases triplex stability, whereas phosphorothioate linkages 5'- to a pyrimidine decreases triplex stability.<sup>66</sup>

An example of a TFO with phosphorothioate linkages is described by McGuffie *et al.*<sup>69</sup> They report on the phosphorothioate oligonucleotide of sequence 5'-TGGGTGGGTGGTTTGT TTTTGGG-3', which is a mixed purine-pyrimidine oligonucleotide designed to bind in antiparallel orientation to the purine strand of the target duplex. This oligonucleotide formed a triplex *in vitro* with the target site shown below [Figure 56], and was sufficiently taken up by leukaemia cells where it remained intact for up to 72 hours.<sup>69</sup>

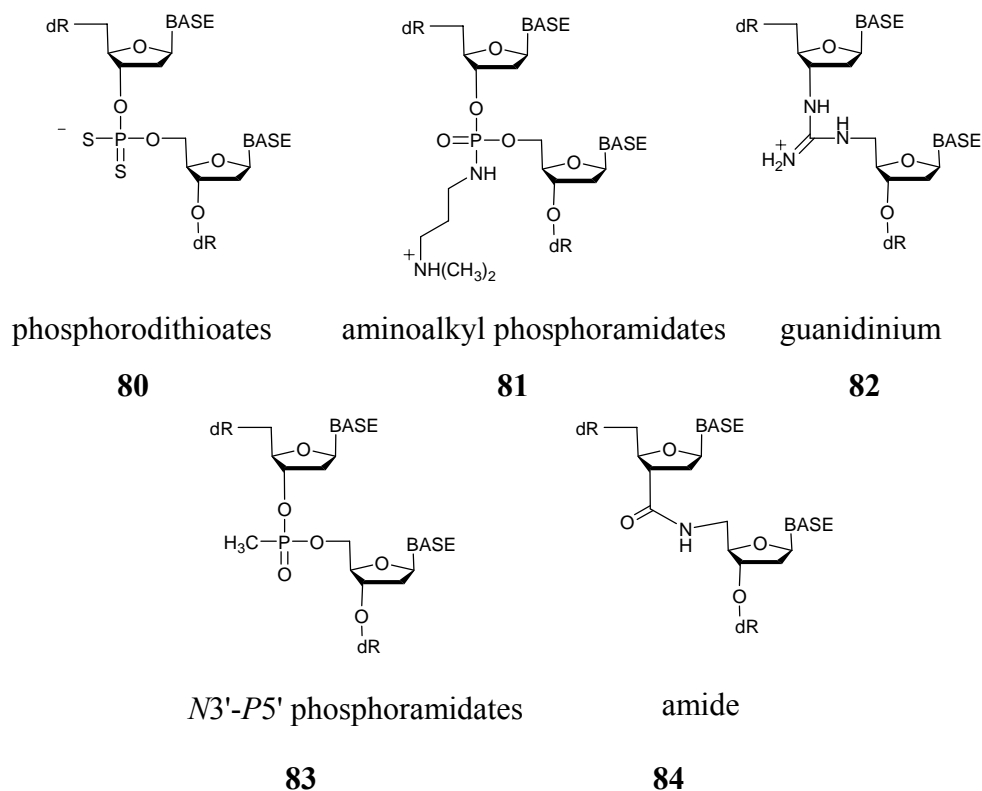


**Figure 56.** Target TFO binding site.

Electrophoretic mobility shift assays (ESMA) were used to determine that a triplex had formed between the target site and the phosphorothioate oligonucleotide.<sup>69</sup> This involved labelling the 5'- end of the oligonucleotide corresponding to the pyrimidine strand of the duplex DNA with <sup>32</sup>P and annealing it to the complementary purine strand.<sup>69</sup> Their oligonucleotide was shown to inhibit endogenous *c-myc* gene expression, reduce cell proliferation, disrupt cell cycle progression and induce cell apoptosis.<sup>69</sup>

The limitations associated with phosphorothioate modifications are that they are toxic as a direct result of non-specific binding interactions with numerous enzymes and they are degraded by nucleases over a period of time.<sup>70</sup>

As stated previously, the number of examples available in the literature for phosphate modifications is vast. Figure 57 depicts a selection of other such modifications.



**Figure 57.** A selection of phosphate modifications.

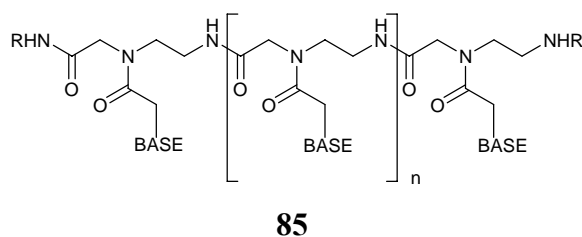
### 1.9.5 Sugar-phosphate backbone modifications

In attempt to impart synthetic oligonucleotides with resistance towards intracellular enzymes, namely nucleases, numerous analogues have been designed in which both the sugar and phosphate groups in the backbone have been altered together at the same time. Due to the extent of research already carried out in this field, only the most promising example reported to date will be discussed here. The reader is referred to a review by Varma for a more comprehensive insight into this research field.<sup>65</sup>

#### *Peptide nucleic acids (PNAs)*

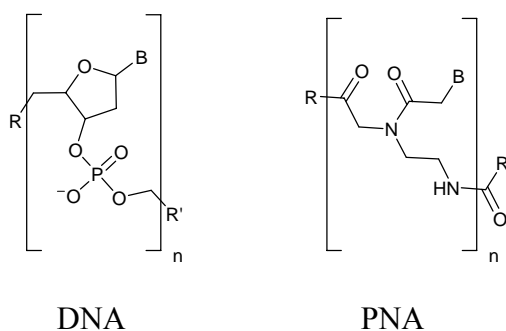
Peptide nucleic acids are analogues of DNA in which the conventional sugar-phosphate backbone has been replaced with an achiral polyamide backbone composed of 2-aminoethyl glycine subunits [Figure 58].<sup>71</sup>





**Figure 58.** Schematic diagram of the structure of PNA.

The structure of PNA, although radical at the time, was designed to maintain key features of DNA such as the base-to-base and base-to-backbone distances. Egholm *et al.*<sup>72</sup> discovered from computational models, that the optimal number of bonds between two nucleosides was six (consistent with the number of bonds between two nucleosides in natural DNA) [Figure 59]. Further, the optimum number of bonds between the base and a backbone is two or three.<sup>72</sup>



**Figure 59.** Schematic comparison of the structures of DNA and PNA ( $B = \text{base}$ ).

PNAs stiff, achiral, polyamide structure allows for excellent sequence specific binding with high affinity to ssDNA, RNA and dsDNA.<sup>71</sup> The tertiary amide linker of PNA is conformationally labile and exists as both the *E*- and *Z*-rotameric forms in uncomplexed PNA.<sup>71</sup> Due to the lack of charge repulsion<sup>71</sup> and the gain of entropy, following on from the pre-organisation of PNA, DNA/PNA complexes exhibit closely packed binding structures.<sup>71</sup>

Since their development, PNAs have demonstrated that they possess numerous significant functional properties and these have resulted in them being useful tools in molecular biology, prebiotic chemistry and antisense/antigene research.<sup>71</sup>

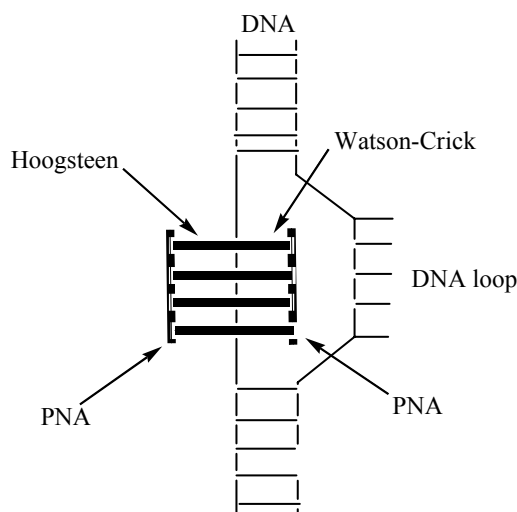
PNA is limited by its reduced aqueous solubility and consequent minimal uptake by cells.<sup>73</sup> To overcome these problems, PNA is generally synthesised as *C*-terminal lysine

amides.<sup>74</sup> The binding affinity and aqueous solubility are therefore aided as a result of this positive charge having been incorporated.<sup>74</sup>

PNAs form triplexes with homopurine ssDNA (or RNA) targets and two complementary homopyrimidine PNA strands, (PNA•DNA•PNA or PNA•RNA•PNA). These hybrids exhibit unprecedented high thermal stability, resulting in a  $T_m$  increase of 1 °C per base for DNA/PNA hybrids and 1.5 °C per base for DNA/RNA complexes compared to DNA binding to itself.<sup>75</sup> Such hybrids can be formed in either parallel or antiparallel conformations, but the latter are more stable by 1-2 °C.<sup>76</sup>

Nielsen *et al*<sup>75</sup> have previously shown<sup>77, 78</sup> that PNA binds to homopurine targets in dsDNA by a novel mechanism, triplex-invasion, and not by the more conventional formation of DNA<sub>2</sub>/PNA triplexes in the major groove through Hoogsteen hydrogen bonding.

Triplex-invasion involves homopyrimidine PNAs binding through Watson-Crick base pairing to its complementary oligonucleotide strand in an antiparallel orientation, resulting in the displacement of the non-complementary strand, to afford the PNA/DNA heteroduplex and the displaced strand.<sup>79</sup> A second strand of PNA then binds in the major groove to give a PNA<sub>2</sub>/DNA triplex. Triplex-invasion is depicted schematically in Figure 60 on page 67.



**Figure 60.** PNA triplex invasion (redrawn from the Kobenhavens University website).

As mentioned previously, PNA oligomers can potentially be used as antisense/antigene therapeutics. However, they are substantially hydrophobic in nature,

compared to the hydrophilicity observed for peptides or oligonucleotides, and it is for this reason that they are not readily taken up by cells.<sup>76</sup> This therefore, has restricted their utility in such applications to date.

## 1.10 RESEARCH OUTLINE

This purpose of this chapter was to gain a greater understanding of the structure, replication and molecular interactions of DNA. As well as the processes by which gene expression leads to the synthesis of proteins in cells. Having achieved this, it was decided to investigate further the role that antigene therapy can play in the future of drug research, through inhibition of gene expression. The focus of the work reported in this thesis relates specifically to improving the regulation of gene expression in a sequence-selective manner.

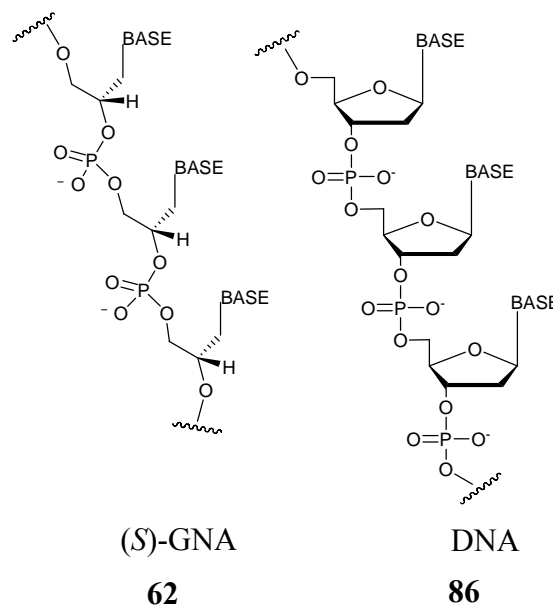
As already discussed in section 1.7.2 on pages 41-45, one potential approach for achieving this involves the use of triplex-forming oligonucleotides. These oligonucleotides bind in the major groove of the double-stranded target *via* Hoogsteen hydrogen bonding to form a triple helix. This complex subsequently interferes with regulatory proteins binding to the same and effectively inhibits transcription. Such compounds could find clinical value as antiviral and anticancer agents.

With a view to developing novel TFOs, we were inspired by the findings reported by Meggers *et al.*<sup>59, 60</sup> (detailed on pages 54-58) on the structurally simplified glycol nucleic acid (GNA) mimics. As stated earlier, these researchers have demonstrated that, in addition to forming stable homochiral GNA:GNA duplexes, (*S*)-GNA oligomers hybridised to complementary, anti-parallel ssRNA; however, no binding to complementary, anti-parallel ssDNA was observed. The reason for this is perhaps due to the fact that, in GNA, both the base-to-backbone and base-to-base intrastrand distances are shorter than those found in natural DNA [Table 8].

<u>Nucleic acid</u>	<u>base-to-base/N<sup>o</sup> of bonds</u>	<u>base-to-backbone/N<sup>o</sup>. of bonds</u>
( <i>S</i> )-GNA	5	2
DNA	6	3

**Table 8.** *Base-to-base and base-to-backbone distances of (*S*)-GNA 62 and DNA 86.*

Figure 61 shows the structure of (*S*)-GNA **62** and DNA **86**.



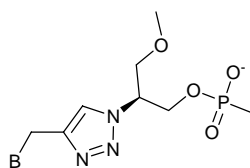
**Figure 61.** Structures of (*S*)-GNA **62** and DNA **86**.

These bond distances, and in particular the base-to-base intrastrand distance, are important and should be carefully considered in the design of oligonucleotide analogues. As stated previously, Egholm *et al.*<sup>72</sup> discovered from computational models that the optimal number of bonds between two nucleosides was six, and that the optimal number of bonds between the base and the backbone was two or three (see the previous section 1.9.5, page 64).

In seeking to improve upon the findings reported by Meggers *et al.*<sup>80</sup> by simply increasing the number of bonds in both the base-to-base and base-to-backbone of GNA, to mirror those found in DNA, the number of degrees of flexibility of the molecule would also increase as a consequence. This may have a detrimental effect on its binding due to entropic reasons. Since it is generally assumed that nucleic acid analogues containing a phosphodiester backbone must be cyclic in order to produce the required conformational preorganisation for duplex formation.<sup>80</sup> It was already astonishing that the thermal stability of the GNA duplex noticeably surpassed that of the analogous DNA duplex.<sup>80</sup>

By adhering to the optimal number of base-to-base and base-to-backbone bonds outlined by Egholm *et al.*<sup>72</sup> and in attempt to prevent any consequential entropic issues, we designed the triazole-containing nucleic acid (TCNA) phosphoramidite monomer

shown in Figure 62 below, to explore side chain and backbone expanded analogues of GNA.



87

**Figure 62.** Structure of our TCNA phosphoramidite monomer.

As the name suggests, a triazole group is included. This was incorporated in an attempt to counteract any entropic problems that may accrue from increasing the side chain length and therefore, flexibility, by restricting the side chains movement. Further, the added advantage of using a triazole moiety is that it can be incorporated using the increasingly popular “click chemistry” approach, which was first discovered by Sharpless *et al.*<sup>81</sup> ‘Click’ chemistry, for our purpose, is a Cu<sup>I</sup>-catalysed variant of the Huisgen 1,3-dipolar cycloaddition of terminal alkynes and azides. This reaction will be discussed in greater detail in chapter 2, (section 2.4.2, page 128).

The six bonds between the bases of our TCNA analogue are consistent with those in DNA. It is hoped that this will prevent the bases from becoming out-of-phase with each other and, therefore, aid its hybridisation with DNA. The number of bonds between the base and the backbone, however, has been increased to five in our TCNA compared to three in regular DNA. Although the number of bonds between these two moieties is larger, the actual distance is reflective of the incorporation of the triazole group and not the analogous carbon-carbon distance. The combined crude bond distance for the shortest part of the triazole group (N1-C5=C4) is 2.72Å, and the analogous crude C-C=C distance is 2.88Å. Thus, the actual distance between the base and the backbone is shorter than first anticipated.

The work reported in this thesis will, therefore, focus on the development of viable synthetic approaches for the preparation of the required novel triazole-containing phosphoramidite monomers, for incorporation into TCNA oligomers. The main aims were: (1) to develop a suitable synthetic pathway for the synthesis of the azide component; (2) to develop an appropriate synthetic route for the preparation of the nucleobase-containing alkyne component; (3) to investigate the use of the ‘click

reaction' for coupling our azide and alkyne components, and; (4) to develop a viable route for the construction of the TCNA monomers. Finally, if time permitted, it was planned to examine the introduction of the TCNA phosphoramidite monomers into oligomers using solid phase synthesis. Our progress towards achieving these objectives will be detailed in the next chapter.

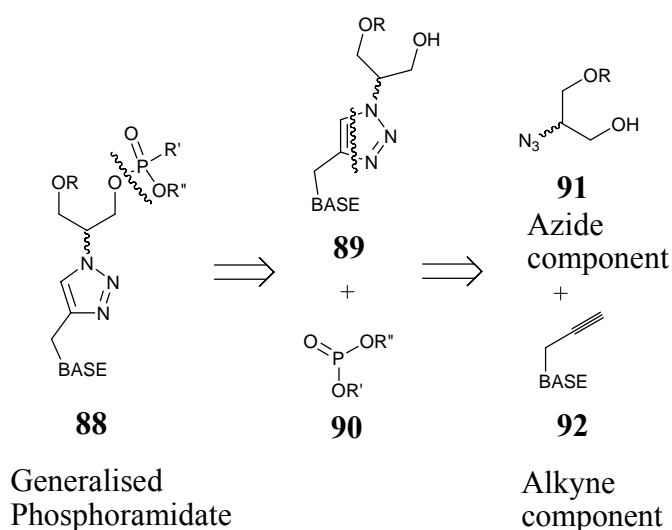
**Chapter 2**  
**Results and Discussion**

## 2. RESULTS AND DISCUSSION

### 2.1 INTRODUCTION

As discussed in our research outline in chapter 1 (section 1.10, pages 67-70), the aim of this research was to design a synthetic approach to all four novel, triazole-containing phosphoramidite monomers [Scheme 4] required for integration into oligomers.

It was envisaged that the triazole unit required for these monomers could be easily prepared using “click chemistry” (see later section 2.4, pages 125-147) which was first introduced by Sharpless *et al.*<sup>81</sup> in 2001. There are many processes which have been identified that fall into the category of ‘click’ chemistry, but for our purposes, we focused on the Cu<sup>I</sup>-catalysed variant of the Huisgen 1,3-dipolar cycloaddition of terminal alkynes and azides. Using this approach, we proposed that the phosphoramidite monomers **88** could be simply prepared from the retrosynthetic analysis, shown in Scheme 4, using the appropriate nucleobase-containing azide **91** and alkyne component **92**. [Scheme 4]



BASE = A, C, G, T

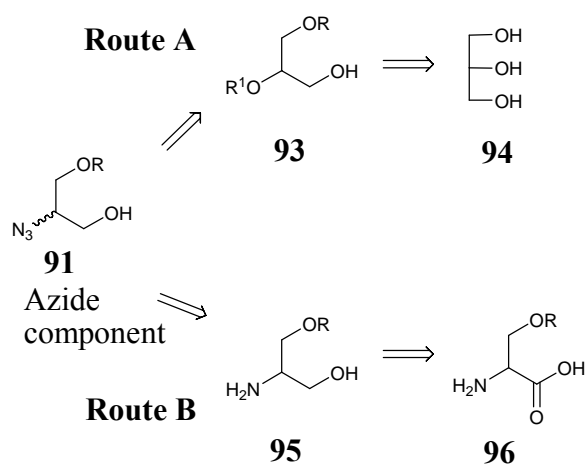
**Scheme 4.** Retrosynthetic analysis pathway of phosphoramidite monomer **88**.

This chapter will focus on: (i) the development of a viable synthetic route to the azide component **91**; (ii) development of synthetic routes to each of the four nucleobase-containing terminal alkyne components, **92**, and; finally, (iii) the development of the ‘click’ reaction to afford the target TCNA phosphoramidite monomers.



## 2.2 DEVELOPMENT OF A SYNTHETIC ROUTE TO AZIDE COMPONENT (91)

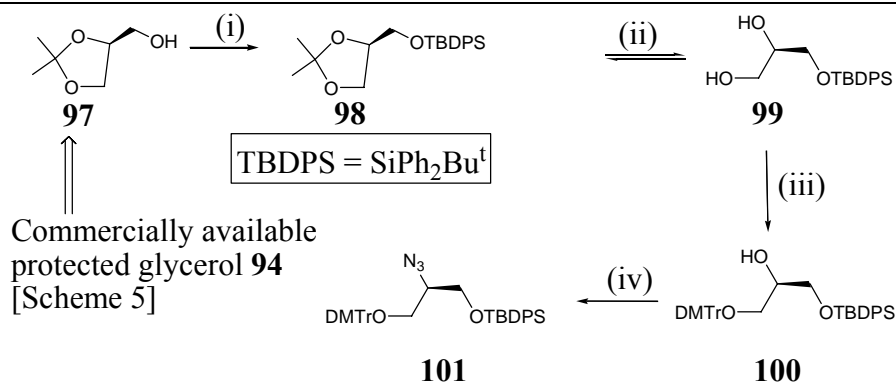
It was envisaged that the azide component **91** could be prepared according to the retrosynthetic analysis pathway shown in Scheme 5 below. In route A, the azide **91** undergoes a functional group interconversion to the differentially protected triol **93** which can be obtained from naturally occurring glycerol **94**. In route B, the azide **91** undergoes a functional group interconversion to afford the amine alcohol **95** which can be obtained from the carboxylic acid **96**, which is commercially available serine.



**Scheme 5.** Retrosynthetic analysis pathway of the azide component **91**.

### 2.2.1 Glycerol route (route A)

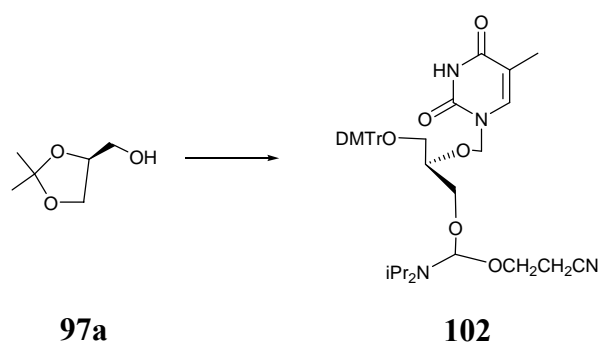
In order to obtain the required azide component **91**, we decided to investigate the four step reaction pathway shown in Scheme 6 on page 74, starting from commercially available (4*S*)-2,2-dimethyl-1,3-dioxolane-4-methanol **97**. [Scheme 6]



**Scheme 6.** Reagents and conditions: (i) TBDPSCl, DMAP, TEA, DCM, rt, 24 h; (ii) TFA, THF/water, rt, 45 h; (iii) DMTrCl, DBU, DCM, rt, 27 h; (iv) (1) MsCl, DCM, rt, 24 h, (2)  $\text{NaN}_3$ , DMF, rt, 24 h.

The first step involved protecting the primary alcohol with a TBDPS group to afford **98**. The acetal protecting group was then removed to form diol **99**, which subsequently had the primary alcohol function selectively protected over the secondary alcohol with a DMTr group to give **100**. The final step in this reaction scheme was to convert the remaining alcohol to azide **101** via an *in situ* mesylation nucleophilic substitution.

Scheme 6 was based on work reported by Xiang *et al.*<sup>82</sup> These researchers had successfully obtained the cytosine analogue **102**, containing a conformationally flexible acyclic linker, starting from the alternative isomer (4*R*)-2,2-dimethyl-1,3-dioxolane-4-methanol **97a** [Scheme 7].

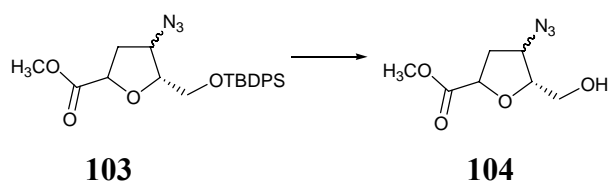


**Scheme 7.** Synthesis of **102**.

Inspired by their finding, we decided to apply the same conditions for the preparation of **98**. Thus, DMAP and TEA were added to a stirred solution of **97** in dry DCM at rt followed by *tert*-butyldiphenylsilyl chloride. After aqueous work-up and purification by flash chromatography, **98** was afforded in a 97% yield. Formation of **98** was confirmed

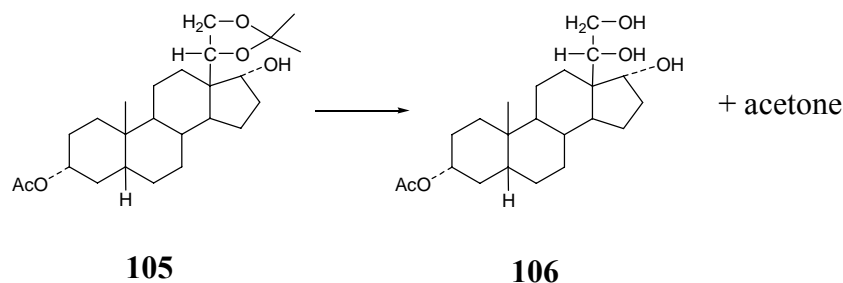
from the  $^1\text{H}$  NMR spectrum recorded which showed a singlet at 0.92 ppm corresponding to the nine protons of the *tert*-butyl group attached to the silicon functionality, and two multiplets at 7.16-7.29 ppm and 7.50-7.58 ppm, corresponding to the ten aromatic protons of the TBDPS protecting group.

Acetals are typically cleaved under acidic conditions. However, we were aware of the fact that TBDPS groups can be labile under certain acidic conditions because Edwards *et al.*<sup>83</sup> had reported the successful removal of the equivalent silicon group in a 76% yield using a 3% methanolic solution of HCl at rt [Scheme 8]. Therefore, less harsh acidic conditions were sought.



**Scheme 8.** Reagents and conditions: 3% HCl in  $\text{CH}_3\text{OH}$ , rt.

Lewbart *et al.*<sup>84</sup> described the successful removal of an acetal protecting group from 5 $\beta$ ,21-pregnane-3 $\alpha$ ,17,20 $\beta$ ,21-tetrol 3-acetate, **105** using the milder organic acid, acetic acid. [Scheme 9]

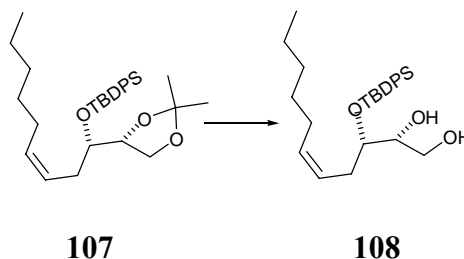


**Scheme 9.** Synthesis of 5 $\beta$ ,21-pregnane-3 $\alpha$ ,17,20 $\beta$ ,21-tetrol 3-acetate, **106**. Reagents and conditions: 60% acetic acid, 4 h, rt.

This group of researchers treated compound **105** with a 60% acetic acid solution for 5.5 hours at rt to afford **106** in a 97% yield. Inspired by their work, we decided to apply the same conditions for the removal of the acetal protecting group on **98**. Unfortunately, in this case the acetal protecting group was not cleaved and only unreacted starting material **98** was recovered. Further attempts at this reaction including increasing the reaction time to a maximum of 24 hours and heating the reaction to 60  $^\circ\text{C}$  for a

maximum of 6 hours, also failed to afford the desired diol, **99**. We therefore sought another method.

Leblanc *et al.*<sup>85</sup> had described the successful cleavage of an acetal protecting group on **107** using TFA to afford the diol, **108** in an 83% yield. Interestingly, the TBDPS protecting group was unaffected by these conditions. [Scheme 10]



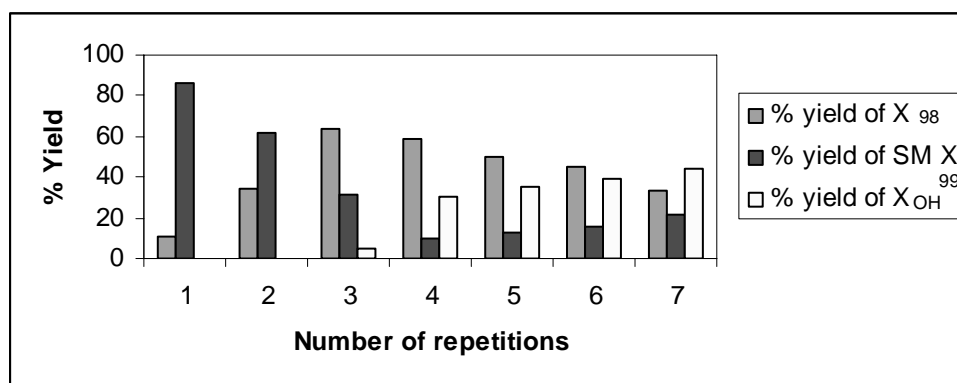
**Scheme 10.** Diol preparation. Reagents and conditions: TFA, THF/H<sub>2</sub>O, 0 °C, 24 h.

Thus these conditions were applied to our compound **98**. A solution of **98** in THF and H<sub>2</sub>O was stirred at 0 °C, and TFA was added before leaving the resulting solution to stir overnight. After work-up and purification by column chromatography, diol **99** was afforded in an 11% yield. This finding was after one addition of reagents.

We were aware of the fact that this step was an equilibrium process and therefore, we decided to explore whether the equilibrium could be driven in favour of diol formation by periodically evaporating the reaction mixture *in vacuo* and adding fresh reagents. The results of this investigation are shown in Table 9. It was found that three repetitions afforded the best balance between a clean reaction and a good percentage yield of 64%.

N <sup>o</sup> . of repetitions	% Yield of <b>99</b>	% Yield of SM <b>98</b>	% Yield of TBDPSOH
1	11	86	0
2	34	62	0
<b>3</b>	<b>64</b>	<b>31</b>	<b>5</b>
4	59	10	30
5	50	13	35
6	45	16	39
7	33	22	44

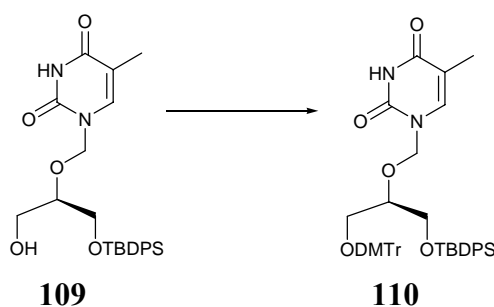
**Table 9.** N<sup>o</sup>. of repetitions, % yield and % of recovered materials.



**Graph 2.** *N<sup>o</sup>. of repetitions versus % yield of materials recovered.*

The formation of the desired product **99** was confirmed by <sup>1</sup>H NMR spectroscopy. The spectrum recorded showed a broad singlet at 2.26 ppm integrating to two protons which was assigned to the diol protons. Further, the nine protons corresponding to the *tert*-butyl group were observed as a singlet at 0.91 ppm, and the ten aromatic protons from the diphenyl constituent were observed as two multiplets at 7.18-7.33 ppm and 7.45-7.53 ppm. The presence of both of the later signals confirmed that the TBDPS group had remained intact.

The next step in the synthetic pathway involved protecting the primary alcohol of **99** with a DMTr protecting group. This ensured that the triazole-containing monomers were protected with the same group used in DNA/RNA synthesis and so these monomers could be incorporated into nucleic acid oligomers.

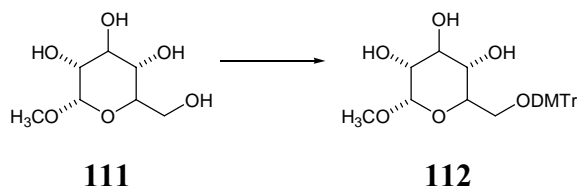


**Scheme 11.** *DMTr protection. Reagents and conditions: DMTrCl, pyridine, rt, 4 h.*

Xiang *et al.*<sup>82</sup> showed that a DMTr protecting group could be introduced into alcohol **109** using DMTrCl in the presence of pyridine [Scheme 11]. Their DMTr protected compound **110** was obtained in a 92% yield. Therefore, a solution of our diol **99** in anhydrous pyridine was treated with DMTrCl at rt and the progress of the reaction was

monitored periodically by TLC. Unfortunately, after 24 hours, only a spot corresponding to starting diol **99** could be observed and there was no indication of any product formation. Therefore, we decided to apply heat to the reaction and so the mixture was gradually warmed to a maximum of 70 °C over 12 hours. Following work-up and purification by flash chromatography, still no product could be identified from the  $^1\text{H}$  NMR spectra collected for the fractions obtained. Only the starting diol **99** was recovered in a 96% yield.

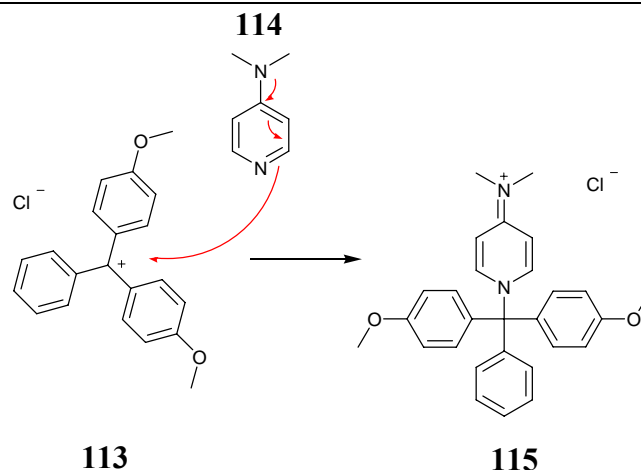
Chaudhary *et al.*<sup>86</sup> reported that the primary alcohol of  $\alpha$ -methylglucoside **111** could be selectively tritylated to give **112** in an 88% yield [Scheme 12]. This was accomplished by treating a solution of **111** in DMF with DMTrCl in the presence of DMAP and TEA.



**Scheme 12.** Selective tritylation of  $\alpha$ -methylglucoside. Reagents and conditions: DMTrCl, DMAP, TEA, DMF, rt, 24 h.

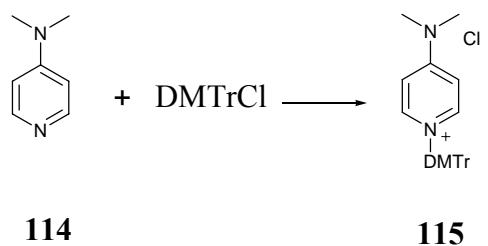
Thus, these conditions were applied to diol **99**. Following work-up and purification by flash chromatography, none of the  $^1\text{H}$  NMR spectra recorded for each of the fractions isolated corresponded to the desired DMTr protected product **100**. Again, only unreacted diol **99** could be identified.

It was hypothesised that this tritylation reaction should occur by preliminary formation of the *N*-tritylpyridinium salt **115** [Scheme 13].



**Scheme 13.** Proposed mechanism for formation of **115**. Reagents and conditions: DMAP, DMTrCl, DCM, rt, 30 mins.

It was thought that if this salt had not formed *in situ*, or if its formation was rate determining then this could be the source of our problems. In order to avoid this issue, we decided to explore the procedure reported by Hernandez *et al.*<sup>87</sup> They had described the preparation and isolation of **115** prior to its subsequent use in the immediate tritylation of primary alcohols [Scheme 14].



**Scheme 14.** Preparation of **115**. Reagents and conditions: (i) DMAP, DMTrCl, DCM, rt, 30 mins.

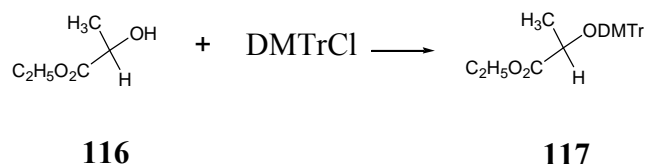
Thus, salt **115** was prepared by reaction of DMAP with DMTrCl in DCM at rt for 30 minutes. The <sup>1</sup>H NMR spectra recorded on the white solid isolated after addition of Et<sub>2</sub>O to the reaction mixture, confirmed that alkylation of the heterocyclic nitrogen had occurred.

Subsequently, salt **115** was dissolved in DCM and treated with diol **99** and the reaction was left for 24 hours. Unfortunately, following work-up and purification by flash chromatography, still no alkylated product **100** was observed. Again, unreacted starting diol **99** was recovered quantitatively. The outcome of this reaction suggested

that our alkylation problems did not lie with the formation of the required DMAP salt. There had to be underlying issues. Thus, our attempt to make the ‘electrophile more electrophilic’ had failed.

In order to explore other possible reasons which may be preventing DMTr alkylation of diol **99**, we decided to investigate the substitution of DMAP for NaH in the hope that use of this stronger base would deprotonate the alcohol and increase the speed of the rate determining step. Thereby, making the ‘nucleophile more nucleophilic’. Therefore, one equivalent of NaH was added to a stirred solution of diol **99** in DCM followed by DMTrCl and the reaction was left at rt for 24 hours. Regrettably, only unreacted starting diol **99** was again recovered. The same result was achieved when two equivalents of NaH were added.

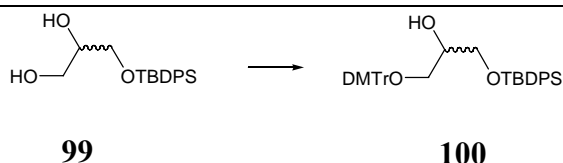
An alternative procedure for performing this reaction has been reported by Messenger *et al.*<sup>88</sup> They describe a convenient method for the preparation of trityl ether **117** from the secondary alcohol **116** in an 81% yield using DBU as a sterically hindered amidine base [Scheme 15].



**Scheme 15.** *Tritylation from secondary alcohols. Reagents and conditions: DBU, DMTrCl, DCM, rt, 27 h.*

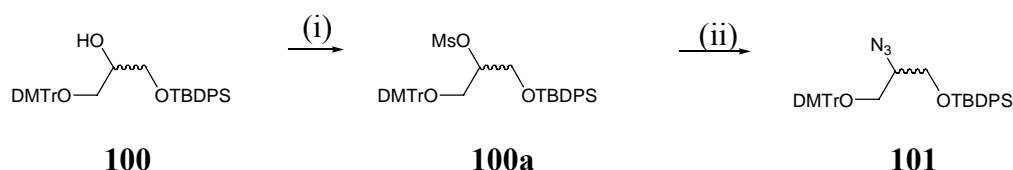
Thus, we decided to apply the same reaction conditions for DMTr protection of diol **99**. Therefore, **99** was added to a solution of DMTrCl and DBU in DCM. Following work-up and purification by flash chromatography, **100** was afforded in a 49% yield [Scheme 16]. Although this yield was poor, it was sufficient to progress with the synthetic scheme and could be optimised at a later date. The <sup>1</sup>H NMR spectrum recorded of the product **100** showed two multiplet signals corresponding to the aromatic protons of the DMTr substituent at 6.56–6.83 ppm and 7.06–7.67 ppm, a singlet at 3.71 ppm corresponding to the methoxy protons and a third multiplet at 3.53–3.69 ppm corresponding to the CH and CH<sub>2</sub> groups near the OH and silicon protecting group respectively.





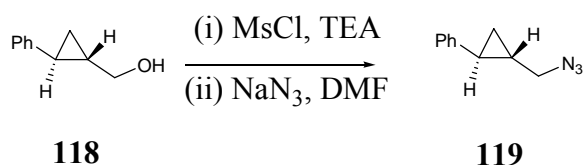
**Scheme 16.** DMTr alkylation. Reagents and conditions: DBU, DMTrCl, DCM, rt, 27 h.

Having successfully obtained **100**, the final step in our synthetic scheme was to convert the remaining secondary alcohol in **100** to azide **101** via a mesylation and nucleophilic substitution with azide [Scheme 17].



**Scheme 17.** Azide formation via mesylation. Reagents and conditions: (i) MsCl, DCM, rt, 24 h; (ii) NaN<sub>3</sub>, DMF, rt, 24 h.

From our initial inspection of the literature, a method involving two reactions being carried out in one pot by Mohapatra *et al.*<sup>89</sup> was appealing. They described the *in situ* conversion of (1*S*,2*S*)-cyclopropyl alcohol **118** into the corresponding mesylate followed by its immediate transformation into azide **119** in an overall 92% yield [Scheme 18].



**Scheme 18.** Alcohol to azide conversion via an *in situ* mesylate intermediate. Reagents and conditions: (i) MsCl, TEA, DCM, 0 °C, 6 h; (ii) NaN<sub>3</sub>, DMF, 60 °C, 24 h.

It was decided to follow the strategy of Mohapatra *et al.*<sup>89</sup> however, not initially as a one-pot reaction. Instead we sought to isolate and characterise the intermediate mesylate product **100a** and once certain of its formation, continue with its conversion to azide **101**. Having established that both steps were successful individually, the one-pot reaction would then be carried out.

In order to isolate the mesylate product **100a**, alcohol **100** was dissolved in DCM and cooled to 0 °C. TEA and MsCl were added and the reaction mixture was stirred at rt for 24 hours. Regrettably, after work-up and purification of the crude product by flash chromatography, no mesylate product **100a** was obtained. Only the starting alcohol **100** was recovered quantitatively. Despite several repeat attempts of this reaction [Table 10], we were never able to isolate **100a**. The reason for this failure is currently uncertain. The reaction conditions attempted are shown in Table 10.

Experiment N°	Reaction Time/hours	Reagent equivalents			% Yield of 100a
		Alcohol	MsCl	TEA	
1	24	1	3	6	0
2	48	1	3	6	0
3	62	1	3	6	0
4	24	1	4	6	0
5	24	1	5	6	0
6	24	1	6	6	0
7	48	1	4	6	0
8	48	1	5	6	0
9	48	1	6	6	0

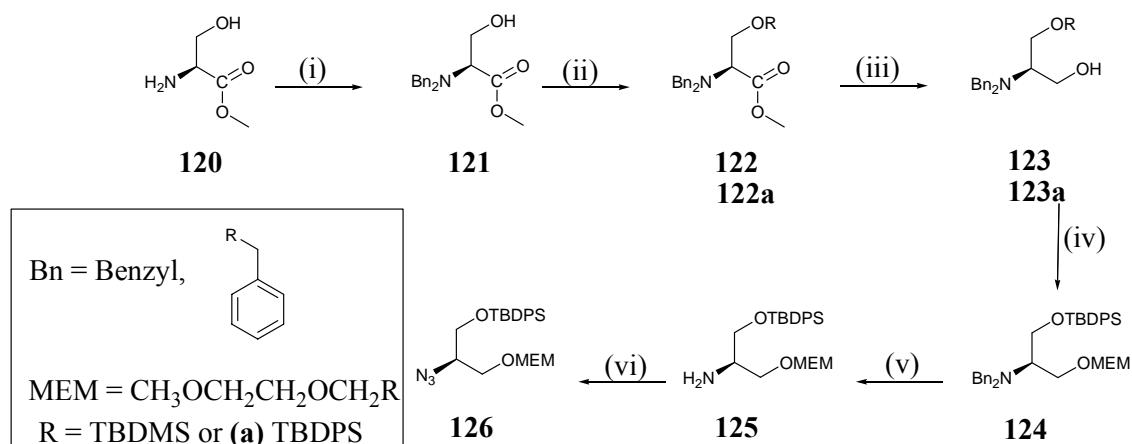
**Table 10.** Mesylate reaction conditions attempted.

The synthetic scheme faced problems throughout its entirety; the difficulty in removing acetal **98** to afford diol **99**, albeit that this was overcome, the problems involved in protecting the primary alcohol with DMTr and the low percentage yield obtained when the DMTr product **100** was finally achieved. These difficulties coupled with the fact that the mesylate could not be produced in our hands, led us to abandon the synthetic route to afford azide **101** (Scheme 6, page 74) and focus our efforts on designing and developing a new synthetic scheme.

### 2.2.2 Serine route (route B)

The alternative route B to target azide **91**, as described in our retrosynthesis in Scheme 6, (see the previous section 2.2.1, page 74) starts from commercially available L-serine

methyl ester **120** and involves reduction of the ester to the primary alcohol **95**, followed by a diazotransfer reaction to azide **91**. Scheme 19 shows our initial synthetic pathway to target azide **91**.

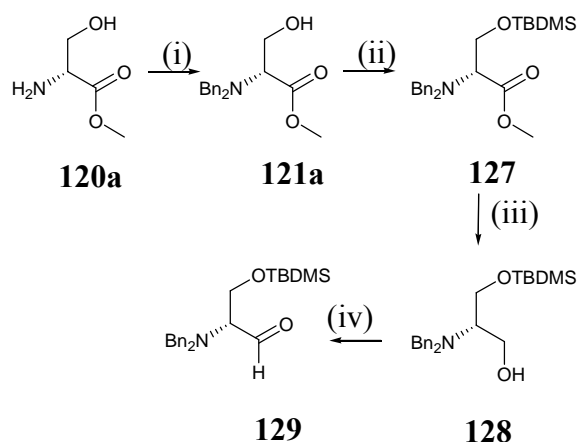


**Scheme 19.** Reagents and conditions: (i) *BnBr*, *NaHCO*<sub>3</sub>, *THF/DMF*, reflux, 14 h; (ii) *TBDMSCl* or (a) *TBDPSCl*, imidazole, *DCM*, rt, 6 h; (iii) *LiBH*<sub>4</sub>, *THF*, rt, 24 h; (iv) *MEM-Cl*, *NaH*, *THF*, rt, 24 h; (v) *H-cube* (see later discussion for details); (vi) ‘diazotransfer’, *K*<sub>2</sub>*CO*<sub>3</sub>, *CuSO*<sub>4</sub>·5*H*<sub>2</sub>*O*, *MeOH*, rt, 24 h.

This scheme preliminary involved protecting the free amine of L-serine methyl ester **120** with a benzyl group to afford **121**, the alcohol of which was then protected with a silyl group to afford **122**. The choice of silyl protecting group was uncertain at this stage as we were unsure about the ability of TBDMS to withstand the following reduction conditions. Therefore, TBDMS and TBDPS were tested. Subsequently, the ester of **122/a** was reduced using *LiBH*<sub>4</sub> to give alcohol **123/a**. The third step involved protection of the newly formed alcohol using *MEM-Cl* to give **124**. This was followed by hydrogenation to remove the benzyl protecting groups and give **125**. The final step in this synthetic scheme used a diazotransfer reaction to produce azide **126**. The protecting groups were carefully selected in order to afford a means of removing each substituent independently of each other; benzyl group removal by hydrogenation should not interfere with the silicon or MEM protecting groups and likewise the silicon removal with fluorides should not interfere with MEM or benzyl groups.

The first three steps in our synthetic scheme use similar conditions to those reported by Laïb *et al.*<sup>90</sup> for the preparation of (2*R*)-2-(*N,N*-dibenzylamino)-3-*O*-[(*tert*-

butyldimethylsilyl)oxy]propionaldehyde **129** from commercially available D-serine methyl ester **120a** [Scheme 20].

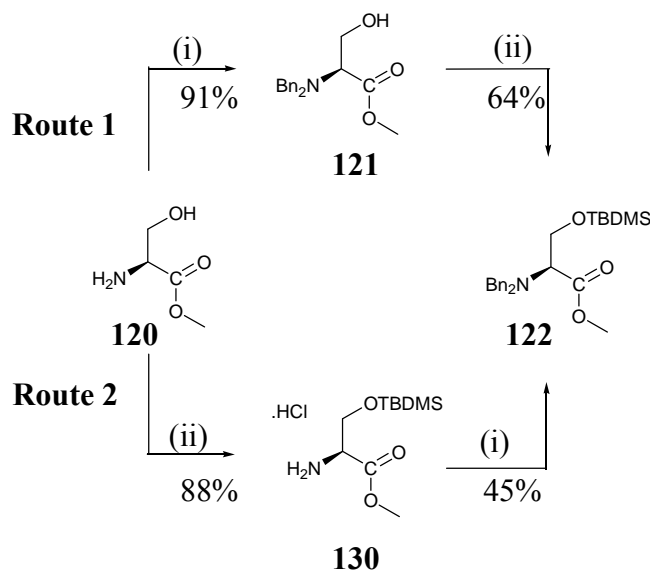


**Scheme 20.** Synthetic route to **129**. Reagents and conditions: (i)  $\text{NaHCO}_3$ ,  $\text{BnBr}$ ,  $\text{DMSO}/\text{THF}$ , reflux, 15 h; (ii)  $\text{TBDMSCl}$ ,  $\text{DMF}$ , imidazole, rt, 15 h; (iii)  $\text{LiBH}_4$ ,  $\text{Et}_2\text{O}-\text{MeOH}$ ,  $0^\circ\text{C}$ -reflux, 3 h; (iv) Swern.

The first step in our synthetic scheme followed the procedure reported by Laïb *et al.*<sup>90</sup> for the quantitative synthesis of **121a** with the exception of using  $\text{DMF}$  instead of  $\text{DMSO}$ . Therefore, L-serine methyl ester **120**,  $\text{NaHCO}_3$  and  $\text{BnBr}$  were dissolved in a 1:1 mixture of  $\text{THF}$  and  $\text{DMF}$  and heated to reflux for 14 hours. Following work-up and purification by flash chromatography, **121** was afforded in a 91% yield. Characterisation data for this compound was found to be consistent with that recorded by Laïb *et al.*<sup>90</sup>

The next step in our reaction scheme was to protect the primary alcohol **121** using a silicon group. Laïb *et al.*<sup>90</sup> had chosen to protect the equivalent alcohol using a  $\text{TBDMS}$  group and it was decided to follow their approach. Therefore, **121** was suspended in  $\text{DCM}$  and treated with  $\text{TBDMSCl}$  and imidazole for 15 hours at rt. After work-up and purification by flash chromatography, **122** was given in a 64% yield as opposed to the 90% yield reported by Laïb *et al.*<sup>90</sup> The reason for the difference in yields may be due to  $\text{DCM}$  being used as solvent instead of  $\text{DMF}$ . The characterisation data for our compound **122** was comparable to their data reported. When the exact reaction conditions by Laïb *et al.*<sup>90</sup> were followed, i.e.  $\text{DMF}$  as solvent, an unidentifiable impurity was present alongside product **122**. The source of this impurity was thought to be from the  $\text{DMF}$  used and therefore, it was decided to alter the solvent choice.

An investigation into the order of these two reactions was then carried out to ascertain whether by reversing these two steps, the percentage yields could be optimised.



**Scheme 21.** Route 1 and 2 with percentage yields. Reagents and conditions: (i) *BnBr*, *NaHCO*<sub>3</sub>, *THF/DMF*, reflux, 14 h; (ii) *TBDMSCl*, *imidazole*, *DCM*, rt, 6 h.

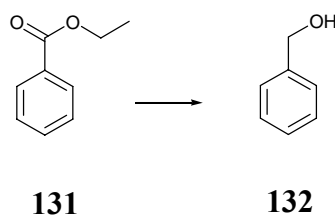
In route 1, the first two reaction steps were carried out in the order described by Laïb *et al.*<sup>90</sup> In route 2, the first two steps were reversed.

From scheme 21 above, it can be seen that route 1 gave product **122** in an overall yield of 60% whereas route 2 gave **122** in an overall yield of 44%. Thus route 1 represented the optimised ordering of these two steps.

The next step in our synthetic scheme involved reduction of the methyl ester of **122** to give the primary alcohol **123**. The conditions for this step reported by Laïb *et al.*<sup>90</sup> seemed the obvious choice. Therefore, *LiBH*<sub>4</sub> and *MeOH* were added to a solution of ester **122** in dry *Et*<sub>2</sub>*O* at 0 °C and the resulting mixture was heated to reflux for 3 hours. After this time, as no new product spot was observed by TLC, the mixture was heated at reflux for a further 3 hours. Following work-up and purification by flash chromatography, the only material that was recovered was the unreacted starting ester **122**. This was surprising given that Laïb *et al.*<sup>90</sup> had reported a 99% yield for their corresponding reaction. This reaction was attempted several times under different conditions (including increased reaction time at reflux and increased equivalents of reagents added) but these reactions similarly failed to afford the alcohol **123**. Drying ether is problematic in terms of being able to remove all traces of water. It was reasoned

that the presence of the remaining water may have influenced the outcome of the  $\text{LiBH}_4$  reaction. Therefore, we sought different reduction conditions in which the water content could be reduced more efficiently.

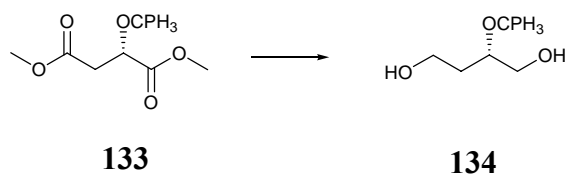
We were inspired by a procedure by da Costa *et al.*<sup>91</sup> which reported the successful reduction of the ethyl ester **131** to primary alcohol **132** in a 96% yield using  $\text{NaBH}_4$  as the reducing agent [Scheme 22].



**Scheme 22.** Ester reduction to primary alcohol. Reagents and conditions:  $\text{NaBH}_4$ ,  $\text{MeOH}$ ,  $\text{THF}$ ,  $70^\circ\text{C}$ , 20 min.

Thus, we decided to apply this method to our ester **122**. Therefore,  $\text{NaBH}_4$  was suspended in  $\text{THF}$ , added to **122** and the mixture was heated to reflux. After one hour,  $\text{MeOH}$  was carefully added and the reaction mixture was maintained at reflux for a further four hours and the progress was monitored periodically by TLC. However, although no new product was observed after this time, it was decided to work-up the reaction. Following purification of the residue by flash chromatography, only starting ester **122** was recovered. This finding coupled with the outcome of the  $\text{LiBH}_4$  reaction, implied that borohydrides alone were not suitable reducing agents for ester **122** in our hands.

A procedure reported by Walker *et al.*<sup>92</sup> using  $\text{LiAlH}_4$  was then attempted in the hope that use of this stronger reducing agent would overcome our difficulties. Walker *et al.*<sup>92</sup> had demonstrated the successful use of  $\text{LiAlH}_4$  for the reduction of methyl ester **133** to the corresponding diol **134** in an 83% yield [Scheme 23].

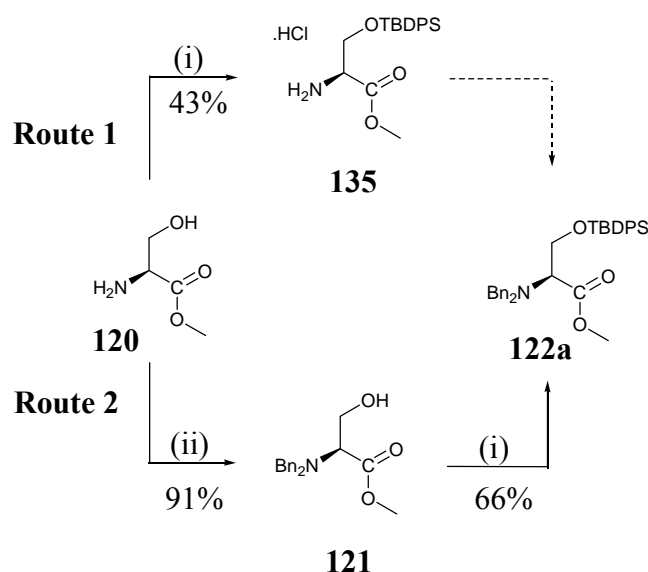


**Scheme 23.**  $\text{LiAlH}_4$  ester reduction. Reagents and conditions:  $\text{LiAlH}_4$ ,  $\text{THF}$ ,  $0^\circ\text{C}$ -rt, 24 h.

Thus, a solution of ester **122** in THF was added dropwise to a stirred suspension of  $\text{LiAlH}_4$  in THF at  $0\text{ }^\circ\text{C}$ . After 24 hours at rt, the reaction mixture was quenched by careful addition of saturated aqueous  $\text{Et}_2\text{O}$  solution. Despite using a stronger reducing agent, following work-up and purification by flash chromatography, unreacted starting ester **122** was still recovered in a 46% yield. Further, TBDMSOH and the desilylated alcohol **123** were recovered in a 24% and 28% yield respectively. It was unclear at this stage whether desilylation had occurred before or after the reduction had taken place.

Our initial doubt with regard to the ability of the TBDMS protecting group to withstand our reduction conditions had been confirmed. It was therefore decided to incorporate a TBDPS protecting group in order try and minimise desilylation from occurring.

Thus, the reaction conditions reported by Laib *et al.*<sup>90</sup> were followed with the exception of TBDPSCl having been used instead of TBDMSCl (Scheme 20, page 84). Again, an investigation into the optimal reaction order was carried out, the results of which are shown below in scheme 24.



**Scheme 24.** Route 1 and 2 with percentage yields. Reagents and conditions: (i) TBDPSCl, imidazole, DCM, rt, 6 h; (ii) BnBr,  $\text{NaHCO}_3$ , THF/DMF, reflux, 14 h.

In route 1, the primary alcohol of L-serine methyl ester **120** was protected using TBDPSCl to afford **135**, with the intention of then protecting the  $\text{NH}_2$  group with BnBr, to determine the optimal reaction order in terms of percentage yields. In route 2, the

previously synthesised benzyl protected **121** (see Scheme 21, page 86) was protected with a TBDPS group to afford **122a**.

The first step of route 1, involved L-serine methyl ester **120**, being taken up in DCM and treated with TBDPSCl and NaHCO<sub>3</sub> for 6 hours at rt. This resulted in a 43% yield of **135** being achieved. TBDPS protection was identified by the nine proton singlet at 0.89 ppm corresponding to the *tert*-butyl group and the two multiplets at 7.18-7.30 ppm and 7.50-7.52 ppm corresponding to six and four protons respectively, for the aromatic portion of the TBDPS protecting group.

The next step involved protecting the NH<sub>2</sub> group of **135** with BnBr to afford **122a**. This was not carried out due to the percentage yield attained for step one in this route, being lower than the corresponding step in route 2, to which it was being compared. Our efforts were therefore concentrated on route 2.

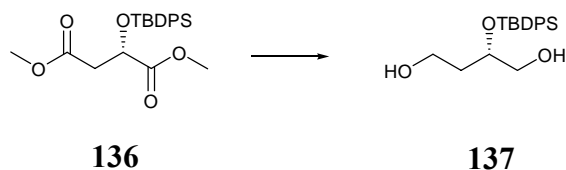
In route 2, the previously prepared benzyl protected serine derivative **121** was subjected to the silylation conditions described previously on page 86. Alcohol **121** was dissolved in DCM and treated with imidazole and TBDPSCl for 6 hours at rt. Following work-up and purification by flash chromatography, product **122a** was afforded in a 66% yield. Evidence of the silicon protection was observed from the nine proton singlet situated at 0.89 ppm corresponding to the *tert*-butyl group and the two multiplets at 7.06-7.34 ppm and 7.45-7.51 ppm corresponding to the aromatic portion of this species.

Having obtained TBDPS protected **122a** the subsequent reduction with LiAlH<sub>4</sub> could be compared. Thus, LiAlH<sub>4</sub> in THF was added slowly to a stirred solution of ester **122a** dissolved in THF at 0 °C. The reaction mixture was then stirred at rt overnight and following work-up and purification by flash chromatography, no desired product was observed. Starting ester **122a** was recovered quantitatively.

It was apparent from this result that reduction of **122** must have taken place after desilylation had occurred, in the case of the TBDMS scheme (page 87). Increasing the stability of the silicon protecting group by employing TBDPS had prevented desilylation taking place, and consequently stopped the ester reduction. In other words, steric crowding of the silicon protecting group was inhibiting alcohol formation. Alternative ester reduction conditions were therefore sought.

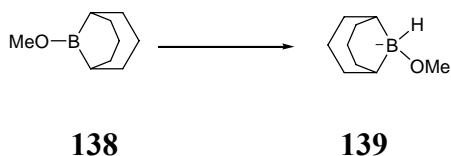


A second procedure reported by Walker *et al.*<sup>92</sup> in which they successfully converted **136** to **137** in a 99% yield [Scheme 25] was then followed.



**Scheme 25.**  $\text{LiBH}_4$ /B-methoxy 9-BBN ester reduction. Reagents and conditions:  $\text{LiBH}_4$ , B-methoxy 9-BBN, THF, 0 °C-rt, 24 h.

These reduction conditions were further supported by Brown *et al.*<sup>39</sup> in which they had reported that ester reduction with  $\text{LiBH}_4$  could be enhanced using B-methoxy 9-BBN. The ester reduction is understood to occur by first activating the B-methoxy 9-BBN **138**, to form its hydride **139** and then this species acts as the reducing agent [Scheme 26].



**Scheme 26.** Activation of B-methoxy 9-BBN. Reagents and conditions:  $\text{LiBH}_4$ , THF.

Thus, we were inspired to try this strategy for our reduction. Ester **122a** in dry THF was added to a stirred solution of  $\text{LiBH}_4$  in dry THF at 0 °C. Subsequently, a solution of B-methoxy 9-BBN in hexane was added and the reaction mixture was left to stir at rt overnight. Following work-up and purification by flash chromatography, the desired product **123a** was afforded in a 48% yield (unoptimised). From the  $^1\text{H}$  NMR spectrum recorded of **123a**, a nine proton singlet at 1.08 ppm, a sixteen and a four proton multiplet at 7.20-7.49 ppm and 7.57-7.75 ppm respectively, corresponding to the aromatic protons were observed, which concluded that the *tert*-butyl group was untouched. The broad singlet at 1.65 ppm was indicative of the OH group generated.

The next step in the synthetic pathway was to protect the newly formed primary alcohol of **123a** with a MEM group. This reaction was attempted by dissolving alcohol **123a** in dry THF with NaH. After an hour, MEM-Cl was added and the reaction mixture was stirred at rt for 24 hours. Following work-up and purification by flash chromatography, the desired product **124** was afforded. However, from the  $^1\text{H}$  NMR

spectrum recorded and subsequent TLC analysis of the product isolated, it was discovered that the desired material had co-eluted with another unidentifiable side product. Product **124** was identified by a three proton singlet at 1.26 ppm for the MEM CH<sub>3</sub> group and a ten proton multiplet at 3.45-3.95 ppm, encompassing overlap of the four benzyl CH<sub>2</sub> protons and six protons from the remaining three MEM CH<sub>2</sub> groups. Several attempts using a number of different solvent systems were used to separate the product from the impurity, but unfortunately these two compounds always co-eluted from the column together.

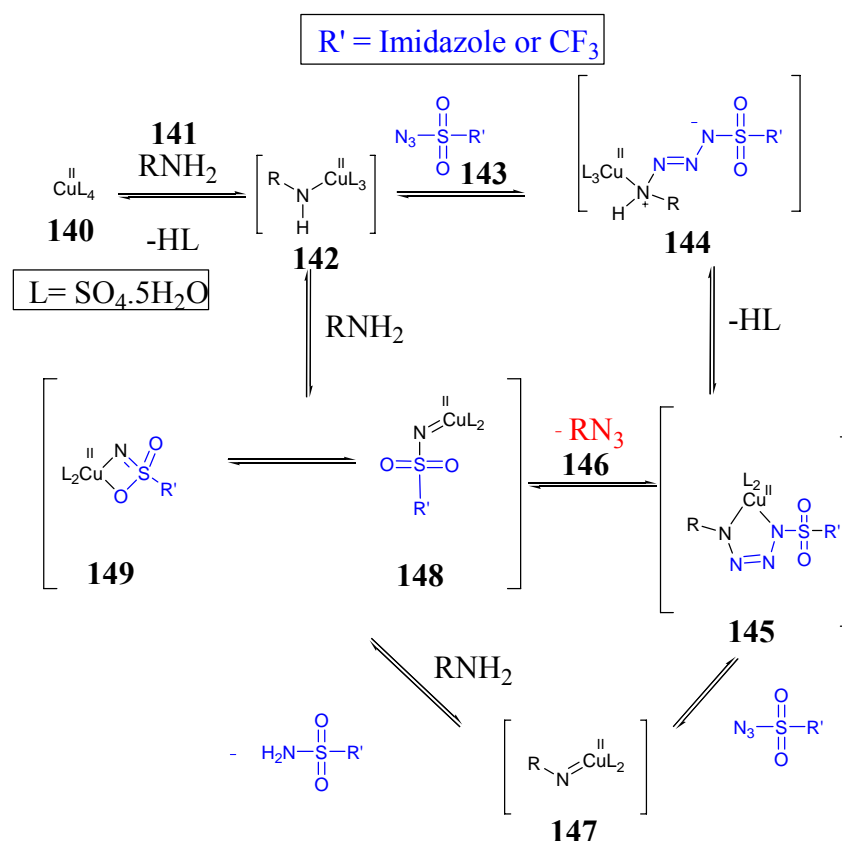
Alternative MEM protection reactions were investigated in order to see if formation of the side product could be prevented. Firstly, the reaction was repeated without the addition of NaH. After the reaction mixture was quenched with aqueous NaHCO<sub>3</sub> solution, the pH was increased to 7.0 using aqueous NaOH solution. Following work-up and purification by flash chromatography, the impurity was still observed in the <sup>1</sup>H NMR spectrum recorded.

We were then inspired by conditions reported by Corey *et al.*<sup>93</sup> in which they successfully prepared several MEM ethers from primary, secondary, tertiary and hindered alcohols. Their MEM ethers were prepared by treating the alcohol with DIPEA and MEM chloride in DCM. The percentage yields for all their examples reported were above 95% and the reactions were complete within 30 to 60 minutes. Thus, a solution of alcohol **124** in dry DCM was cooled to 0 °C and DIPEA along with MEM-Cl were added. The reaction mixture was stirred at rt for 24 hours and the progress was monitored periodically by TLC. Unfortunately, following work-up and purification by flash chromatography, the impurity could still be observed alongside the desired product **124** in the <sup>1</sup>H NMR spectrum recorded. At this point, since it did not seem to be possible to prepare or separate the desired MEM compound **124** from this side product, it was decided to continue with the subsequent step in the synthetic pathway in the hope that this impurity could be separated at a later stage.

The next step in our strategy for preparing the azide component involved removing the benzyl protecting groups to afford amine **125** (see the previous Scheme 19, page 83). This was attempted using the in-house H-cube® hydrogenator, which combines continuous flow microchemistry with on demand hydrogen generation from electrolysis of H<sub>2</sub>O. Thus, a solution of impure **124** in absolute EtOH was cycled once through the

system at 0.5 mL/min. The temperature of the hydrogenator was set to 50 °C and the catcart® used was 10% Pd/C. The  $^1\text{H}$  NMR spectrum recorded of the crude product no longer showed the presence of benzyl protons which had appeared at 3.49-4.05 ppm, 7.20-7.49 ppm and 7.57-7.75 ppm as multiplet signals in the  $^1\text{H}$  NMR spectrum recorded for impure starting material **124**. The newly generated  $\text{NH}_2$  group of amine **125** was observed as a broad singlet at 4.40 ppm. Having envisaged that the subsequent step to form azide **126** would need further purification, it was decided to continue with the next step and then attempt to separate the unknown impurity from the product, in order to keep the quantity of product lost to a minimum.

The next step in our synthetic pathway was to transform the amine into an azide through a diazotransfer reaction [Scheme 19]. The proposed mechanism of such a reaction is shown below in Scheme 27.<sup>94</sup>

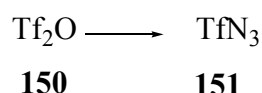


**Scheme 27.** Proposed mechanism for the diazotransfer reaction (redrawn and adapted from ref 95)

Firstly, the copper (II) catalyst **140** binds reversibly to the  $\text{NH}_2$  group of the organic amine starting material **141** to give **142**. The ‘diazo donor’ **143** then associated itself reversibly with this complex, to form **144**. Loss of one of the copper ligands L,

produces complex **145** which can then either break up to regenerate the ‘diazo donor’ **143** and species **147** or form the desired azide **146**. In order to carry out this transformation, a ‘diazo donor’ species is commonly used. Two such donors are: (1) trifluoromethanesulfonyl azide (TfN<sub>3</sub>) **151**, and; (2) imidazole-1-sulfonyl azide hydrochloride, **152**.

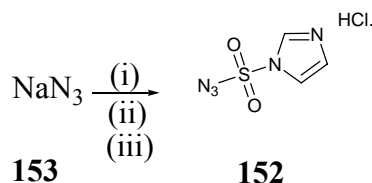
TfN<sub>3</sub> **151** can be prepared from commercially available triflic anhydride **150**<sup>95</sup> [Scheme 28].



**Scheme 28.** Preparation of TfN<sub>3</sub> **151**. Reagents and conditions: NaN<sub>3</sub>, toluene/H<sub>2</sub>O, 0 – 10 °C, 2.5 h.

**151** is the commonest ‘diazo donor’ for the Cu<sup>II</sup> catalysed conversion of an amine to an azide, and it generally produces high yields and preserves the stereochemistry of the starting material in the product.<sup>96</sup> However, neat **151** is explosive in nature and has a poor shelf life, which necessitates it being prepared in solution and used immediately.<sup>96</sup> Further, it is known that preparation of **151** from the expensive starting reagent, trifluoromethanesulfonic anhydride (Tf<sub>2</sub>O) **150**, affords inconsistent yields and therefore, the reagent must be used in a liberal excess.<sup>96</sup>

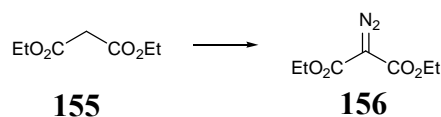
At the time of this research, imidazole-1-sulfonyl azide hydrochloride **152** was emerging as an alternative, efficient, inexpensive and shelf-stable diazotransfer reagent.<sup>96</sup> Goddard-Borger *et al.*<sup>96</sup> designed and prepared this new diazotransfer reagent which had the potential to surpass TfN<sub>3</sub> as a ‘diazo donor’ in terms of its stability, ease of use and reproducible yields.<sup>96</sup> Further, this crystalline reagent can be easily prepared [Scheme 30] in a one-pot reaction on a large scale (several grams) from relatively inexpensive starting materials NaN<sub>3</sub> **153**, sulfonyl chloride **154** and imidazole.<sup>96</sup>



**Scheme 30.** Synthesis of imidazole-1-sulfonyl azide hydrochloride **152**. Reagents and conditions: (i) SO<sub>2</sub>Cl<sub>2</sub>, MeCN, rt, 24 h; (ii) imidazole, rt, 3 h; (iii) HCl in EtOH, 0 °C, 1 h.

This involves treating **153** with sulfonyl chloride in pyridine and subsequent addition of imidazole. Acidification of the reaction mixture then precipitates **152** as a hydrochloride salt.<sup>96</sup>

An example of **152** being used as a ‘diazo donor’ has been described by Goddard-Borger *et al.*<sup>96</sup> Here **156** from **155** was successfully prepared in a 65% yield using this reagent [Scheme 31].



**Scheme 31.** Synthesis of diazo compound **156** using **152**. Reagents and conditions: **154**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{H}_2\text{O}$ ,  $\text{MeOH}$ , rt, 16h.

We were inspired by this new, safe diazotransfer method, and we decided to investigate the conversion of amine **125** to azide **126** using **152** in an analogous manner. In order to accomplish this, **152** first had to be prepared. This was achieved following the procedure described by Goddard-Borger *et al.*<sup>96</sup> Firstly, sulfonyl chloride **154** was added dropwise to a cooled suspension of sodium azide **153** in dry pyridine and the resulting mixture was left at rt overnight. Imidazole was then added to the cooled mixture and the slurry was stirred for a further 3 hours at rt. Subsequent work-up and acidification afforded **152** as a white solid in a 69% yield. This compared favourably to the 63% yield reported by Goddard-Borger *et al.*<sup>96</sup> The I.R. signal corresponding to the azide was of strong intensity at  $2172\text{ cm}^{-1}$  and the  $^1\text{H}$  NMR spectrum recorded of **152** confirmed the presence of the imidazole moiety. Finally, the purity of **152** was verified by elemental analysis.

Having successfully prepared azide **152**, we were now in a position to use this in our diazotransfer reaction to form azide **126** from amine **125**. Therefore, a solution of azide **152** in  $\text{MeOH}$  was added to amine **125** in the presence of  $\text{K}_2\text{CO}_3$  and  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  and stirred at rt overnight according to the procedure by Goddard-Borger *et al.*<sup>96</sup> [Scheme 31]. Following work-up and purification by flash chromatography, formation of the desired azide product **126** was confirmed by I.R. spectroscopy on the material afforded. The spectrum showed a strong signal at  $2096\text{ cm}^{-1}$  which corresponded to the presence of an azide group. Furthermore, the  $^1\text{H}$  NMR spectrum recorded of **126** no longer showed a broad singlet at 4.40 ppm which had been attributed to the  $\text{NH}_2$  group in the

starting material **125**. The impurity that had been produced during the MEM reaction (page 91), despite our best endeavours to remove it at that purification stage, was disappointingly still present alongside azide **126**, following the purification step in this diazotransfer procedure. We therefore decided to abandon this pathway in search of an alternative approach avoiding MEM protection.

#### Alternative serine approach

It was realised that the methyl ester of **120** could be viewed as a protected hydroxyl function, and if this was the case, then this would avoid having to use an alcohol protecting group such as MEM in our pathway. Therefore, it was reasoned that Scheme 19 (page 83) could be both modified and shortened so that following protection of the side chain hydroxyl group of **120** with a silicon protecting group to give **135** (previously prepared), the diazotransfer reaction could then be achieved to generate the desired azide **157** [Scheme 32]. This could then be employed in the subsequent ‘click’ reactions with the appropriate alkyne components before liberating the hydroxyl function to produce our TCNA monomer.



**Scheme 32.** Direct diazotransfer. Reagents and conditions: **152**,  $\text{K}_2\text{CO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{MeOH}$ , rt, 16h.

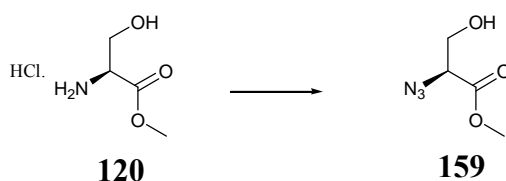
Thus, direct diazotransfer was carried out using previously prepared **135** (Scheme 24, page 88). The amine group of **135** was converted into azide **157** following the same reaction procedure described previously. Thus, amine **135**, azide **152**,  $\text{K}_2\text{CO}_3$  and  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  were taken up in  $\text{MeOH}$  and stirred at rt overnight. Following work-up, the I.R. spectrum recorded of **157** showed a strong azide signal at  $2110 \text{ cm}^{-1}$ . From the crude  $^1\text{H}$  NMR spectrum recorded of **157**, it appeared that some of the silicon protecting group had been removed. In light of this finding, it was then decided to employ previously prepared TBDMS protected **130**, as this silicon protecting group was deemed to be more stable to these reaction conditions [Scheme 33].<sup>97</sup>



**Scheme 33.** Direct diazotransfer. Reagents and conditions: **152**,  $\text{K}_2\text{CO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{MeOH}$ , *rt*, 16h.

Therefore, amine **130** was subjected to the same diazotransfer conditions as **135**, and following the work-up, the crude I.R. spectrum recorded for **158** showed a strong signal at  $2117 \text{ cm}^{-1}$  corresponding to the azide. From the  $^1\text{H}$  NMR spectrum recorded for **158** it was evident that the silicon protecting group was untouched.

Further investigation into desilylation in the presence of  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  and  $\text{K}_2\text{CO}_3$  was carried out by Whittingham *et al.*<sup>98</sup> in our research group. It had initially been thought that the copper catalysts may be responsible for desilylation. However, the results of this investigation, confirmed that  $\text{K}_2\text{CO}_3$  was responsible. TBDPS groups are removed quickly in the presence of this base, and furthermore the results suggested that, if left for long enough, TBDMS groups may also be removed, contrary to that described by Kocienski.<sup>97</sup> Although we had seen no evidence for TBDMS desilylation having occurred, it was decided to investigate if the alcohol of **120** needed to be protected in order for a successful diazotransfer reaction to occur. Therefore, the direct diazotransfer of **120** was investigated [Scheme 34].



**Scheme 34.** Direct diazotransfer. Reagents and conditions: **152**,  $\text{K}_2\text{CO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{MeOH}$ , *rt*, 16h.

Thus, having subjected **120** to the same reaction conditions and following work-up, the crude I.R. spectrum recorded of **159** showed a strong azide signal at  $2110 \text{ cm}^{-1}$ . The crude  $^1\text{H}$  NMR spectrum recorded of **159** showed singlets at 3.64 ppm, 3.70 ppm and 3.82 ppm which confirmed the methoxy  $\text{CH}_3$ , the  $\text{CH}_2$  and  $\text{CH}$  substituents respectively, and the  $\text{OH}$  group was identified as a broad singlet at 5.74 ppm.

Azides **158** and **159** were then immediately used in the subsequent ‘click’ reactions described in section 2.4.4 on pages 138-140.

### 2.2.3 Conclusions

In order to develop a suitable synthetic pathway to afford our azide component, we initially investigated route A (Scheme 6, page 74) derived from commercially available glycerol **94**. This scheme unfortunately afforded limited results.

The first problem faced in this synthetic pathway was to remove the acetal protecting group of **98** to generate the diol **99** without desilylation occurring simultaneously. This difficulty was overcome using a procedure reported by Leblanc *et al.*<sup>85</sup> and having carried out an optimisation investigation into the number of repetitions of fresh reagents added, diol **99** was afforded in a 64% yield.

Difficulties in selectively alkylating the newly generated primary alcohol **99** using DMTrCl were then encountered. After two main investigations into the appropriate conditions for such a task: (i) making the ‘electrophile more electrophilic’ by both isolating and creating *in situ*, the 4-dimethylamino-N-dimethoxytritylpyridinium chloride salt **115** required for such an alkylation, and; (ii) making the ‘nucleophile more nucleophilic’ by employing NaH to deprotonate alcohol **99** and thereby increasing the speed of the rate determining step. Unfortunately, the alkylated species **100** could not be obtained. A procedure reported by Messenger *et al.*<sup>88</sup> was then investigated and this allowed us to successfully afford **100**, albeit in an unoptimised yield of 49% using the sterically hindered base DBU.

The final step in this reaction pathway was to obtain azide **101** and this again proved problematic. It was decided to first investigate the one-pot conversion of alcohol **100** to azide **101** via an *in situ* mesylate from which we had been inspired by Mohapatra *et al.*<sup>89</sup> Having decided to first isolate the mesylate product **100a** to verify its production, and then subject it to the subsequent azide conversion, in our hands, the mesylate product could not be obtained. Several attempts to obtain this product were carried out, for example, increasing the reaction time and altering the equivalents of reagents added, but despite this, all attempts failed to afford us with the desired product **101**. In light of all these problems faced with this synthetic pathway, it was decided to focus our efforts on



route B, which had also been outlined as a potential way to achieve our azide component, at the time of our retrosynthetic analysis of azide **91**.

The second synthetic pathway to achieve azide **126** was route B (Scheme 19, page 84), starting from commercially available L-serine methyl ester **120**. The first three steps of this scheme had been inspired by Laïb *et al.*<sup>90</sup> in which they had carried out the same processes on D-serine methyl ester **120a**, and we were confident that these conditions could be applied to our starting material. However, this was not the case.

The first step to protect the amine moiety with a benzyl group was successful in a 92% yield and the subsequent protection of the alcohol function also was successful in a 64% yield. However, the third step to reduce the methyl ester to the corresponding primary alcohol **123** was not as straightforward. In our hands, using the conditions reported by Laïb *et al.*<sup>90</sup> the reduction was unsuccessful, despite considerable effort. Investigation into use of another borohydride was then undertaken, but to no avail. It was then decided to explore the use of a stronger reducing agent, LiAlH<sub>4</sub>, which had been described by Walker *et al.*<sup>92</sup> However, despite the reduction having taken place, this had unfortunately the detrimental affect of removing the silicon protecting group. What wasn't clear at this stage was whether desilylation had occurred prior to, or after, the reduction of the methyl ester.

In an attempt to prevent the loss of the silicon group, we substituted the TBDMS moiety for a TBDPS protecting group, as this silicon group was deemed to be more stable under these reaction conditions. The outcome of this investigation was that no reduction had taken place. This allowed us to conclude that in the example with the TBDMS protecting group, reduction of the ester had taken place after desilylation had occurred. It was postulated that perhaps the silicon protecting group was sterically hindering the reduction reaction taking place and that by employing the more labile TBDMS group, this reduction was achieved on the corresponding desilylated alcohol. This problem was later overcome by employing conditions reported by Walker *et al.*<sup>92</sup> and further confirmed by Brown *et al.*<sup>39</sup> for the enhancement of LiBH<sub>4</sub> in reduction reactions by use of *B*-methoxy 9-BBN. These conditions allowed us to produce alcohol **123a** in an unoptimised yield of 49%. Although this yield was low, it was decided to continue with the next steps in the reaction pathway and optimise the individual steps at a later date.

Therefore, an investigation into protecting the newly generated alcohol **123a** with MEM-Cl was then carried out. Despite this reaction having been successful in terms of protecting the alcohol function of **123a**, the desired product could not be separated from an unidentifiable product which had co-eluted. Despite trying to find a solvent system that allowed for the two species to be separated, we had no success. The impurity was not U.V. active and was thought to be a MEM type derivative, although this could not be verified. In an attempt to eradicate the side reaction from occurring, the reaction was repeated without NaH being present. Unfortunately this did not solve our problem. We were then inspired by a different procedure reported by Corey *et al.*<sup>93</sup> which employed DIPEA along with MEM-Cl. Again however, co-elution of the impurity with the desired product occurred.

At this point, it was decided to pursue the subsequent step in the reaction pathway in the hope that the impurity could be removed in a later purification step. Thus, the next step to remove the benzyl protecting groups was investigated using an in-house hydrogenator and afforded amine **125** quantitatively. No purification was carried out at this stage in attempt to keep product loss to a minimum as purification of the subsequent step was envisaged.

The final step in this reaction pathway was to convert the amine into an azide using a diazotransfer reaction. We were inspired by Goddard-Borger *et al.*<sup>96</sup> on the development of the ‘diazo donor’ **152** for use in such a conversion. Thus, this reagent was prepared and used together with our amine **125**. From <sup>1</sup>H NMR spectroscopy, it was clear that the desired product had been formed, but unfortunately, despite considerable effort to separate the product, the previously prepared impurity had been carried through. Given that we could not isolate the desired azide, due to the unidentified impurity that always co-eluted with the product. In the interest of time, it was decided to abandon this route, and focus our efforts on a new synthetic scheme which would avoid the need to use MEM-Cl.

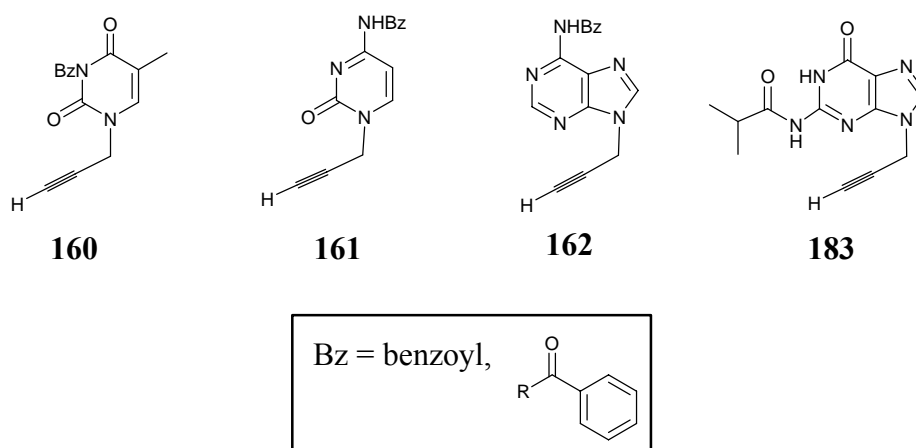
The alternative route designed investigated the viability of direct diazotransfer to take place on the previously prepared TBDPS protected compound **135**. Under these conditions, the silicon protecting group was labile (initially it was thought that the copper catalyst was responsible for such cleavage) and in light of this finding, the reaction was repeated using the previously prepared TBDMS protected species, as this

was reported by Kocienski<sup>97</sup> to be more stable to these conditions. Analysis of the <sup>1</sup>H NMR spectrum recorded for **135** showed that the silicon group had been untouched, however, an investigation carried out by Whittingham *et al.*<sup>98</sup> in our research group, concluded that removal of the TBDMS silicon protecting group may also result if given long enough time. As a result of this separate investigation, it was decided to ascertain whether direct diazotransfer could occur in the presence of the alcohol moiety contained within L-serine methyl ester **120**, and hence prevent the silicon removal being an issue. Direct diazotransfer was once again successful and azides **158** and **159** were now ready to be coupled with the nucleobase-containing alkyne component in a subsequent ‘click’ reaction (see the later section 2.4.4, pages 139-141).

Having successfully obtained our azide components, the focus of this chapter now moves to the development of the alkyne component synthesis.

### 2.3 DEVELOPMENT OF THE ALKYNE COMPONENT SYNTHESIS

The four natural nucleobases were synthetically modified to include an alkyne moiety, in order to make them suitable components for the ‘click’ reaction. The target alkyne components we envisaged preparing for our ‘click’ reaction are shown below in Figure 63.



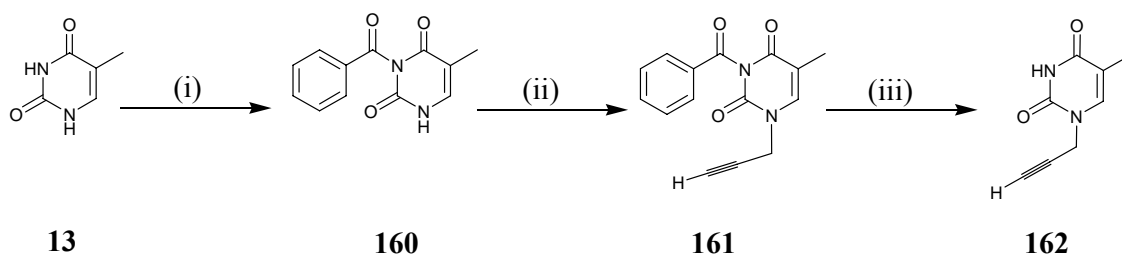
**Figure 63.** The target alkyne components.

From Figure 63, **160** is benzoylated in the N3 position for reasons discussed later in section 2.3.1, (pages 100-103). It was essential to protect the exocyclic NH<sub>2</sub> groups on **161** and **162** with a benzoyl group, not only to increase the lipophilicity of the monomer, but also to adhere to the traditional protecting group strategy when

incorporating the monomers into oligomers. In the case of **183**, the exocyclic NH<sub>2</sub> is alternatively protected with an isobutyryl group for reasons discussed later in section 2.3.4.

This section will focus on the individual synthetic approach for each nucleic base and the problems we encountered and overcame.

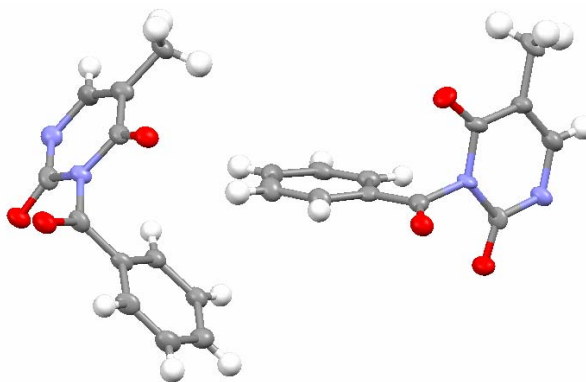
### 2.3.1 Thymine alkyne component synthesis



**Scheme 35.** Reagents and conditions: (i) BzCl, CH<sub>3</sub>CN/C<sub>5</sub>H<sub>5</sub>N (1:1), rt, 24 h; (ii) HC≡CCH<sub>2</sub>Br, NaH, DMF, rt, 48 h; (iii) BnOH, 90°C.

Scheme 35 shows the planned synthetic scheme designed to prepare the alkyne component. The N3 position is first protected using BzCl to increase the lipophilicity of the compound and also to eradicate any possibility of N3 alkylation in the subsequent step. Alkylation is then carried out in the N1 position using commercially available propargyl bromide to give **161** and finally, debenzoylation is carried out (optional) to form **162**.

Therefore, the first step involved treating a suspension of commercially available thymine **13**, in a mixture of CH<sub>3</sub>CN and pyridine, with BzCl. Following work-up, the crude yellow solid afforded was purified by recrystallisation to give **160** in a 73% yield. N3 benzoylation was confirmed by x-ray crystallography. As shown in Figure 64 on page 101, two molecules were found to reside within the asymmetric unit which differed slightly in the orientation of both the phenyl and methyl substituents.



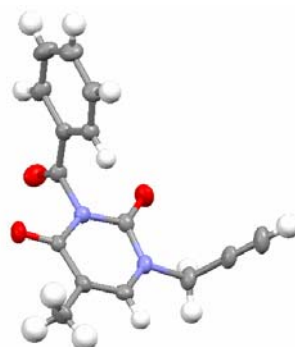
**Figure 64.** Ball and stick representation of the crystal structure of **160** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

From this crystal structure, the angle between the two rings was found to be  $84.39^\circ$  and one of the carbonyl oxygens forms a close contact with the nitrogen in the ring, which results in a dimer (information obtained from Dr Georgina Rosair, HWU).

The next step in the synthetic pathway involved alkylation of the *N3* benzoylated derivative **161**. We decided to adapt a procedure found in a paper by Choi *et al.*<sup>99</sup> Thus, a solution of **160** in dry DMF was treated with NaH. Subsequently, propargyl bromide was added and after 24 hours in the dark, the TLC analysis showed that, unfortunately, no reaction had occurred. Therefore, it was decided to heat the reaction. The temperature was gradually raised in  $10^\circ\text{C}$  increments up to a maximum of  $80^\circ\text{C}$  over a four hour period. Regrettably, following work-up, the  $^1\text{H}$  NMR spectrum recorded of the crude reaction mixture confirmed that the reaction had failed and only unreacted starting material could be identified.

An extensive search of the literature provided evidence for use of both  $\text{Cs}_2\text{CO}_3$ <sup>100</sup> and  $\text{K}_2\text{CO}_3$ <sup>101</sup> as bases for alkylation purposes. Due to  $\text{Cs}_2\text{CO}_3$  generally being more soluble in organic solvents than  $\text{K}_2\text{CO}_3$ , we substituted, NaH for  $\text{Cs}_2\text{CO}_3$ . After work-up, the crude residue was purified by flash chromatography, to successfully afford **161** in a 55% yield. From the I.R. spectrum recorded of **161**, we could see a weak alkyne stretch at  $2129\text{ cm}^{-1}$  and three strong carbonyl signals at  $1747\text{ cm}^{-1}$ ,  $1696\text{ cm}^{-1}$  and  $1644\text{ cm}^{-1}$ . The  $^1\text{H}$  NMR data obtained, displayed a clear triplet at 2.67 ppm which was assigned to the terminal propargyl proton and a doublet at 4.56 ppm which was attributed to the propargyl 'CH<sub>2</sub>' attachment. The absence of the *N*-1H, which had previously been observed at 11.58 ppm in the  $^1\text{H}$  NMR spectrum recorded for **161**, was further proof

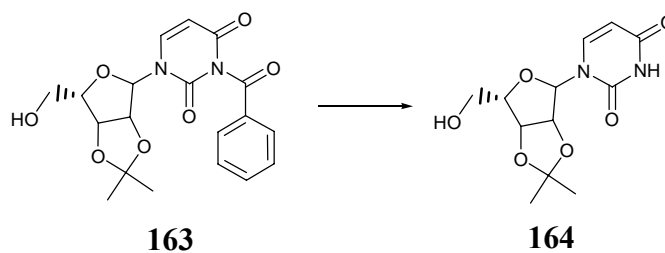
that *N*1 alkylation had taken place. From the crystal structure obtained [Figure 65], alkylation had occurred exclusively at the desired position and there was only one molecule within the asymmetric unit.



**Figure 65.** Ball and stick representation of the crystal structure of **161** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

From this crystal structure, it was observed that a twist angle of  $84.06^\circ$  exists between the planes defined by the two rings (information obtained from Dr Georgina Rosair, HWU).

The final step in our synthetic strategy to **162** involved removal of the *N*3 benzoyl protecting group of **161**. There are a number of ways to do this reported in the literature. We were attracted to the simple approach described by Maguire *et al.*<sup>102</sup> These researchers had reported that the benzoyl group can be easily removed by heating the compound **163** to reflux in benzyl alcohol [Scheme 36].

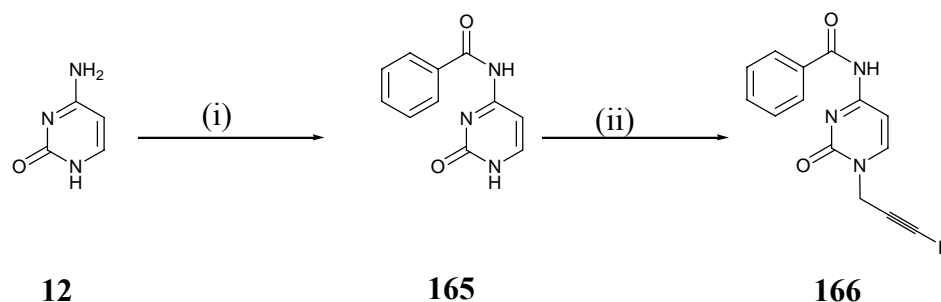


**Scheme 36.** Debenzoylation. Reagents and conditions: *BnOH*, reflux, 24 h.

Thus, **161** was dissolved in benzyl alcohol and heated to reflux for 30 hours. Removal of the solvent and purification using flash chromatography afforded **162** in an 82% yield. From the  $^1\text{H}$  NMR spectrum recorded, the absence of the aromatic protons and the new 'NH' proton signal at 8.12 ppm proved that debenzoylation had taken place.

Further, from the I.R. spectrum recorded, there were only two strong carbonyl signals observed at  $1701\text{ cm}^{-1}$  and  $1698\text{ cm}^{-1}$ , in contrast to the three observed at  $1747\text{ cm}^{-1}$ ,  $1696\text{ cm}^{-1}$  and  $1644\text{ cm}^{-1}$  for **162**. This compound was then directly subjected to our ‘click’ conditions as described in the later section 2.4.4 (pages 139-141).

### 2.3.2 Cytosine alkyne component synthesis

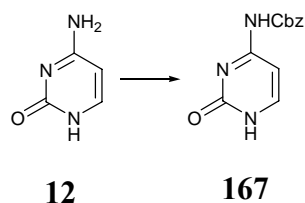


**Scheme 37.** Reagents and conditions: (i)  $\text{BzCl}$ ,  $\text{C}_5\text{H}_5\text{N}$ , rt, 24 h; (ii)  $\text{HC}\equiv\text{CCH}_2\text{Br}$ , DMF, DBU, rt, 24 h.

Scheme 37 shows the planned synthetic scheme designed to prepare the cytosine alkyne component **165**. The initial step in Scheme 37 involves the benzoylation of commercially available cytosine **12** to afford **165**. The exocyclic  $\text{NH}_2$  group must be protected in order to prevent interference in later oligomerisation steps. Usually the protecting group for cytosine (based on oligodeoxynucleotide chemistry) is a benzoyl group. We decided to keep the protecting group the same for our purposes. The final step involves alkylating **165**, using propargyl bromide, to obtain the final cytosine alkyne component **166**.

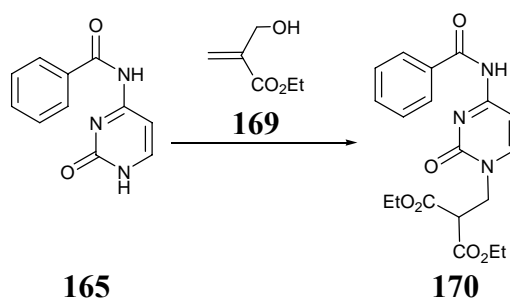
Thus, the first step investigated towards our target cytosine derivative **165** involved suspending commercially available cytosine **12** in dry pyridine in the presence of  $\text{BzCl}$ ; these conditions had been previously used in order to benzoylate the  $N3$ -position of thymine (Section 2.3.1, pages 98-100). Following work-up and purification by recrystallisation, **165** was obtained in a 92% yield. This product proved to be insoluble in all deuterated solvents available in the laboratory and therefore, structural verification by NMR analysis could not be carried out. However, the I.R. spectrum recorded indicated the presence of a medium NH stretching band at  $2964\text{ cm}^{-1}$  and two strong carbonyls bands at  $1738\text{ cm}^{-1}$  and  $1699\text{ cm}^{-1}$ . Elemental analysis also afforded results within experimental parameters and mass spectrometry showed that the desired molecular ion peak was present in 100% relative abundance. Although none of these

techniques proved the site of benzylation, elemental analysis and mass spectrometry results proved that mono-alkylation had occurred. Our elemental analysis was consistent with that reported for the same compound by Brown *et al.*<sup>103</sup> Their elemental analysis found C 61.2, H 4.0, N 19.7 and our data was found to be C 61.4, H 4.2, N 19.4. The required data for C<sub>11</sub>H<sub>9</sub>O<sub>2</sub>N<sub>3</sub> was C 61.4, H 4.2, N 19.5. Based on evidence reported by Nielsen *et al.*<sup>75</sup> on their Cbz protection of *N*4-cytosine [Scheme 38], we surmised that the benzoyl group was most likely attached to the exocyclic *N*4-amino group.



**Scheme 38.** Cbz protection of *N*4-cytosine. Reagents and conditions: PhCH<sub>2</sub>OCOCl, C<sub>5</sub>H<sub>5</sub>N, rt, 24 h.

The next step in our synthetic pathway to **166** was alkylation using propargyl bromide. Initially, we were inspired by conditions reported by Guillaume *et al.*<sup>104</sup> in which **165** was alkylated with the Michael acceptor **169** to give **170** in a 72% yield [Scheme 39].

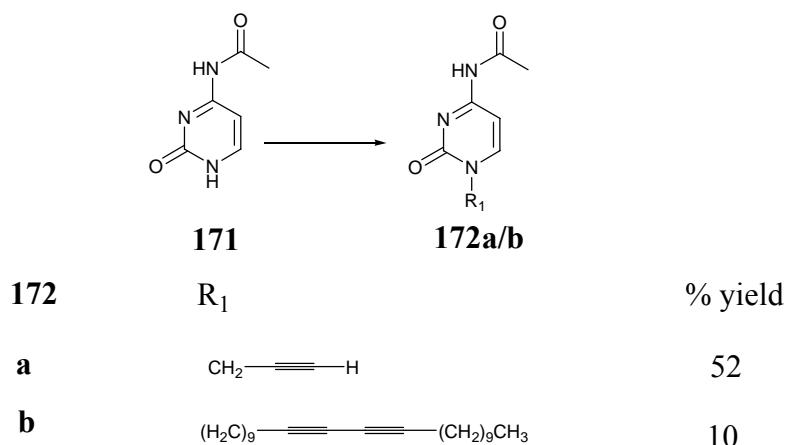


**Scheme 39.** Preparation of **170**. Reagents and conditions: **169**, DBU, DMF, rt, 24 h.

We decided to investigate employing the same conditions for the preparation of **165**. Thus, **12** was suspended in dry DMF in the presence of DBU and propargyl bromide overnight at rt. Following work-up and purification by flash chromatography, none of the desired product **165** was present in any fractions obtained, according to <sup>1</sup>H NMR spectroscopy. Unfortunately, only the starting material was recovered.

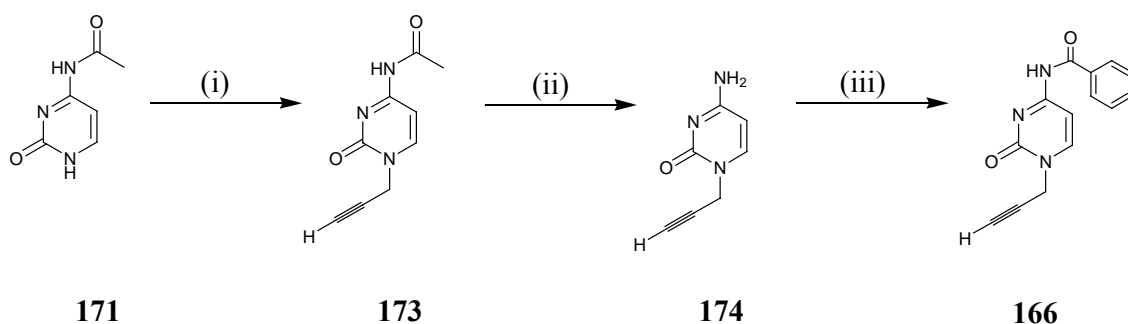


During a search of the literature to find viable alternative routes, we discovered a report by Lindsell *et al.*<sup>105</sup> in which it was described how competing *O2* alkylation had hampered *N1* alkylation during their attempts to prepare a series of cytosine derivatives. These problems were alleviated by starting from the commercially available *N4*-acetylcytosine **171** [Scheme 40].



**Scheme 40.** Cytosine derivatives. Reagents and conditions: **171**, **a** or **b**,  $\text{K}_2\text{CO}_3$ , DMF, *rt*, 24 *h*.

We decided that use of this compound for the preparation of our cytosine monomer may be worthy of exploration, in order to investigate whether the nature of the *N4*-acetylcytosine protecting group of cytosine could influence *N1*-alkylation. Thus, we devised a new synthetic scheme to our cytosine alkyne component **166** starting from *N4*-acetyl cytosine as shown below in Scheme 41.



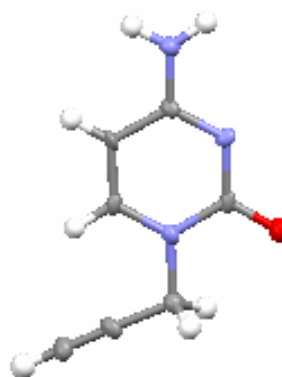
**Scheme 41.** Reagents and conditions: (i)  $\text{HC}\equiv\text{CCH}_2\text{Br}$ , DMF,  $\text{K}_2\text{CO}_3$ , *rt*, 24 *h*; (ii) 5% (*v/v*) methanolic ammonia solution, 24 *h*, *rt*; (iii)  $\text{BzCl}$ ,  $\text{C}_5\text{H}_5\text{N}$ , *rt*, 24 *h*.

This new synthetic route involved first alkylation of the commercially available starting material **171** to afford **173** followed by removal of the acetyl protecting group to

unmask the exocyclic amino function **174**, and finally, re-protection using benzoyl chloride to generate the target compound **166**.

Thus, **171** was suspended in dry DMF and to this anhydrous  $K_2CO_3$  and propargyl bromide were added. Following work-up and purification by recrystallisation from water, **173** was afforded in the slightly reduced yield, compared to the reported yield, of 40%. Full characterisation of this compound was consistent with that published by Lindsell *et al.*<sup>105</sup> From the  $^1H$  NMR spectrum recorded, the alkyne CH proton of **173** was observed as a triplet at 3.45-3.48 ppm and the alkyne  $CH_2$  protons were present as a doublet at 4.63 ppm.

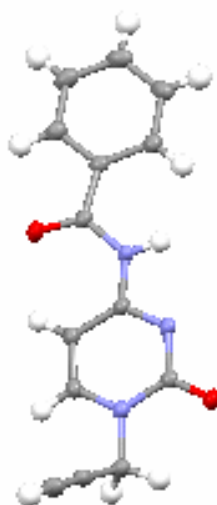
The second step in this new synthetic scheme was also based on the work reported by Lindsell *et al.*<sup>105</sup> They had successfully removed the acetyl protecting group of **173** with a methanolic ammonia solution to afford **174** quantitatively [Scheme 41]. Therefore, it was decided to employ these conditions for our purposes. Thus, **173** was dissolved in a 5% (v/v) methanolic ammonia solution and the resulting mixture was left to stir at rt for 24 h [Scheme 41]. Following solvent removal under reduced pressure, **174** was afforded in a 94% yield. Again, full characterisation of this compound was consistent with that published by Lindsell *et al.*<sup>105</sup> The  $NH_2$  group was observed as a broad singlet at 7.16 ppm in the  $^1H$  NMR spectrum recorded of **174**. Further, a crystal structure of **174** [Figure 66] was obtained which clearly confirmed that alkylation had indeed occurred at the *N1* position exclusively.



**Figure 66.** Ball and stick representation of the crystal structure of **174** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

The final step in this strategy involved re-protection of the exocyclic amine of **174** with a benzoyl group, to afford our final cytosine alkyne component **166** [Scheme 41].

As before, the reaction conditions previously used to protect the *N*3 position of thymine were applied to **174**. Thus, BzCl was added to a cooled solution of **174** in dry pyridine and the solution was stirred at rt overnight. Following work-up and purification by flash chromatography, **166** was afforded in an 84% yield. The  $^1\text{H}$  NMR spectrum recorded of **166** showed five aromatic protons as two multiplet signals at 7.46-7.67 ppm and 7.98-8.02 ppm. In addition to the spectral data collected for **166**, we obtained a crystal structure [Figure 67] in which it was observed that there was only one molecule within the asymmetric unit and that the single benzoyl group was attached to the desired *N*4-amino group.



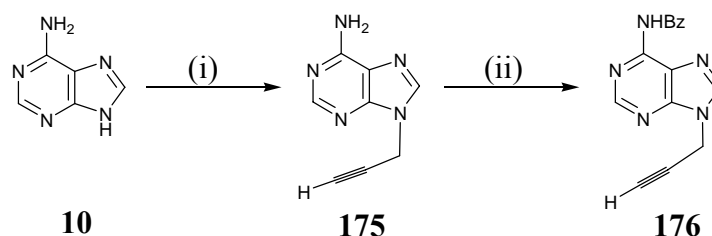
**Figure 67.** Ball and stick representation of the crystal structure of **166** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

From this crystal structure, it was observed that ribbons formed between the  $\text{NH}\cdots\text{N}$  hydrogen bonds forming a ring R2,2(8) and alternating with a R4,4(12)  $\text{NH}\cdots\text{O}$  and  $\text{NH}\cdots\text{N}$ . Furthermore, the pyrimidine ring is coplanar with the alkyne constituent and the layers stack in a herringbone array (information obtained from Dr Georgina Rosair, HWU).

### 2.3.3 Adenine alkyne component synthesis

In the design of the adenine alkyne component synthetic scheme, our approach in terms of protection and alkylation was slightly different to the previous two cases of thymine and cytosine. This was mainly due to the regioselective problems associated with adenine. The order that the exocyclic amine protection and alkylation steps are

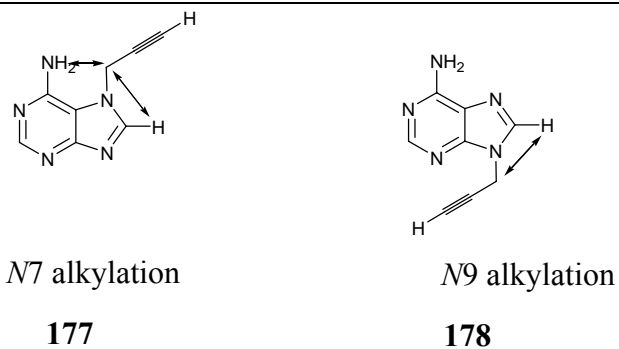
carried out is crucial in influencing the regioselective outcome. It is generally known from the literature, that alkylation must be carried out first on commercially available adenine **10** in order to afford mainly, the *N*9 alkylated product. Protection of the exocyclic amine may then be achieved. Bearing this in mind, we designed the synthetic scheme 42 below to afford the adenine alkyne component **176**.



**Scheme 42.** Reagents and conditions: (i)  $\text{HC}\equiv\text{CCH}_2\text{Br}$ , DMF, NaH, rt, 24 h; (ii)  $\text{BzCl}$ ,  $\text{C}_5\text{H}_5\text{N}$ , rt, 4 h.

This strategy initially involves the selective alkylation of the *N*9 position of commercially available **10** with propargyl bromide to afford **175**. The subsequent step involved protecting the exocyclic amine group with a benzoyl group to obtain our target adenine alkyne component **176**. Again, the benzoyl group was chosen to be consistent with oligodeoxynucleotide synthesis.

For the first step of this strategy, we were again initially inspired by the procedure reported by Choi *et al.*<sup>99</sup> as had previously been used in the case of the thymine alkylation (page 99). This involved dissolving commercially available adenine in dry DMF in the presence of NaH and propargyl bromide. Following work-up and purification by flash chromatography, we obtained **175** in a 36% yield. From the  $^1\text{H}$  NMR spectrum recorded of **175**, the alkyne CH proton was observed as a triplet at 3.54 ppm and the alkyne  $\text{CH}_2$  protons were seen as a doublet at 5.17 ppm. At this point, we needed to verify that the *N*9 alkylated product had been isolated and this was achieved from 2D spectroscopy. It was reasoned that if *N*7 alkylation had taken place, then cross-correlation peaks would be seen for the alkyne  $\text{CH}_2$  group interacting with both *H*8 and the  $\text{NH}_2$  group [Figure 68]. If on the other hand, *N*9 alkylation had taken place, then cross-correlation peaks would only be seen for the interaction of the  $\text{CH}_2$  group with *H*8 [Figure 68].

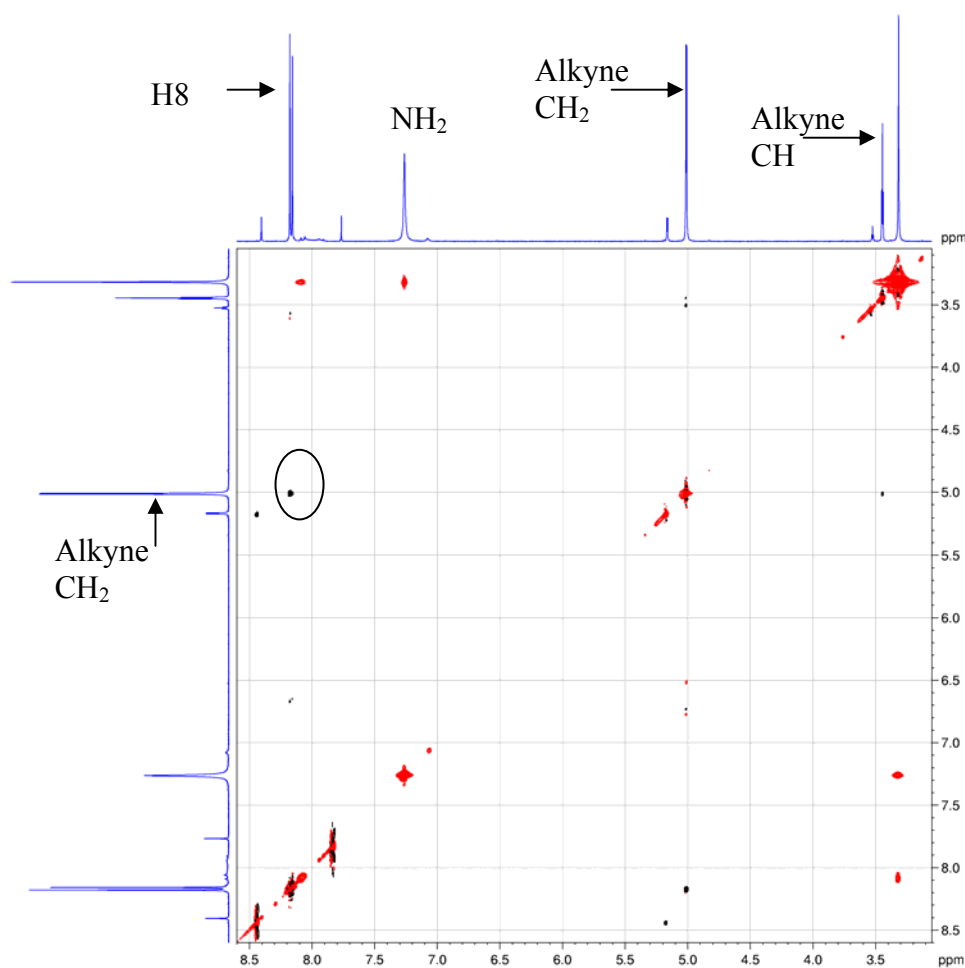


**Figure 68.** *N7 versus N9 alkylation.*

From the NOESY\* spectrum recorded [Figure 69], we could deduce that the NH<sub>2</sub> group was on the opposite side of the molecule to the propargyl group. Cross-correlation peaks were only observed between the alkyne CH<sub>2</sub> group and *H8* and therefore, *N9* alkylation was confirmed.

---

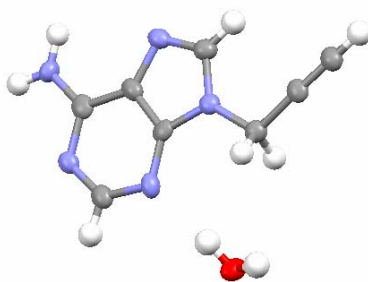
\*NOESY spectrum obtained from Dr Alan Boyd, Heriot-Watt University



**Figure 69.** NOESY\* spectrum of alkylated adenine product **175**.

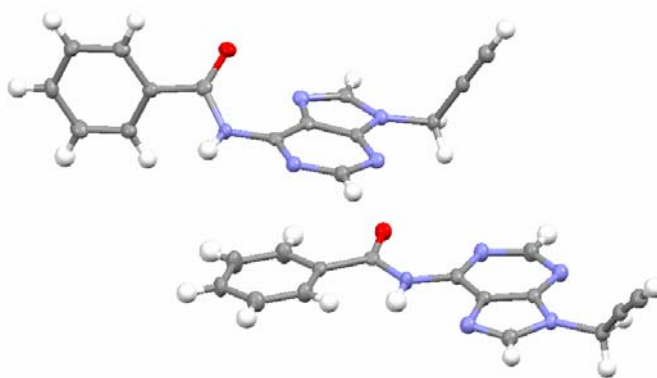
Further, we obtained a crystal structure of **175** which showed that there was only one molecule within the asymmetric unit [Figure 70]. The alkyne group was seen to be *anti* rather than *syn* with respect to the fused ring system. Also, there was one water molecule which engaged in hydrogen bonding with *N3*, and another water molecule, generated by symmetry, for every organic molecule.

\*NOESY spectrum obtained from Dr Alan Boyd, Heriot-Watt University.



**Figure 70.** Ball and stick representation of the crystal structure of **175** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

The final step to produce our adenine alkyne analogue, **176**, involved benzoylating the free amine substituent of **175**. This was achieved according to a procedure reported by Lindsell *et al.*<sup>105</sup> previously discussed in section 2.3.1 (page 99). **175** was suspended in dry pyridine, cooled to 0 °C and BzCl was added. The reaction was stirred at rt for 4 hours and following work-up and purification by flash chromatography, **176** was obtained in a 67% yield. In the crystal structure [Figure 71], we can see that there are two molecules within the asymmetric unit and that they differ in orientation of the adenine ring system with respect to the carbonyl group.

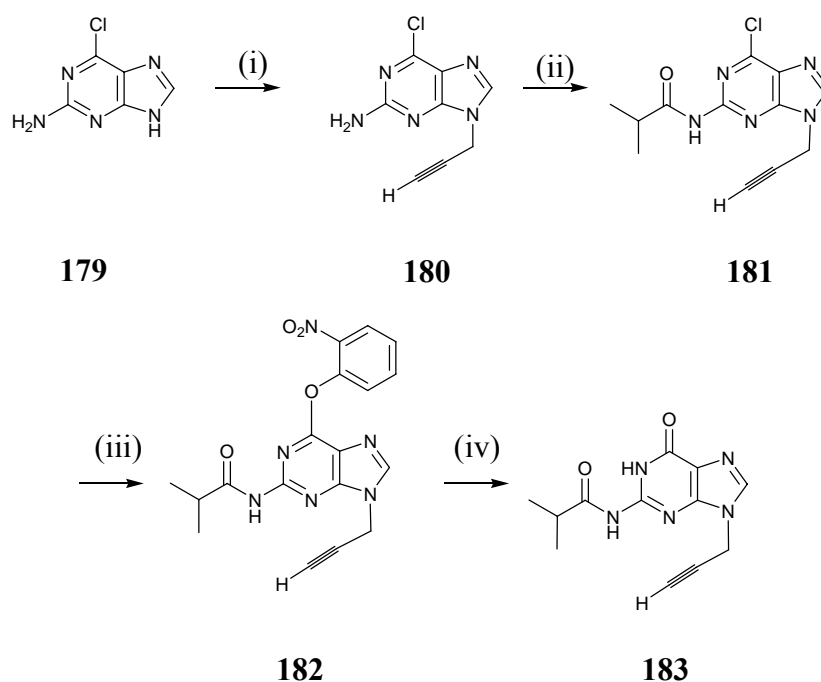


**Figure 71.** Ball and stick representation of the crystal structure of **176** (red = oxygen, white = hydrogen, grey = carbon and blue = nitrogen).

From this crystal structure, the distance between the alkyne CH to the *N*7 position was 2.283Å. Further, the distance between the CH and the centroid of the phenyl ring was 2.82Å. The phenyl and alkyne components lie out of the plane formed by the purine unit (information obtained from Dr Georgina Rosair, HWU).

## 2.3.4 Guanine alkyne component synthesis

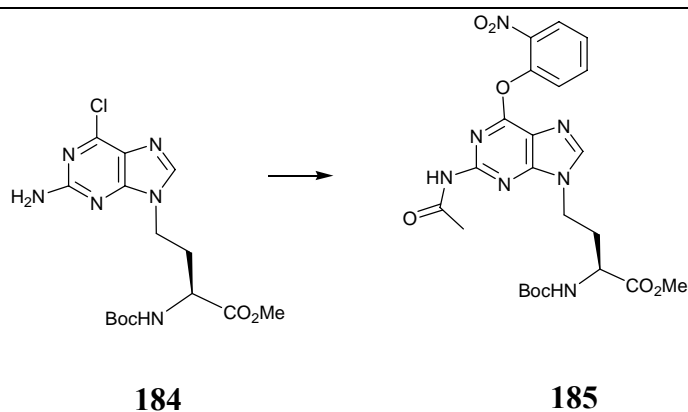
It is known from the literature that guanine is very difficult to work with, and it has been generally accepted that in order to manipulate its structure, it is best to start from the commercially available, although expensive, 2-amino-6-chloropurine, **179** [Scheme 43].<sup>106</sup> This intermediate has been shown to improve the regioselectivity of *N*9 versus *N*7 alkylation. Thus, a synthetic route to guanine alkyne component **183** was devised based on this starting material as shown below in Scheme 43.



**Scheme 43.** Reagents and conditions: (i) HC≡CCH<sub>2</sub>Br, K<sub>2</sub>CO<sub>3</sub>, DMF, rt, 24 h; (ii) *i*PrCOCl, DMAP, C<sub>5</sub>H<sub>5</sub>N, rt, 4 h; (iii) 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h; (iv) 1,1,3,3-tetramethylguanidine, 2-nitrobenzaloxime, CH<sub>3</sub>CN, rt, 48 h.

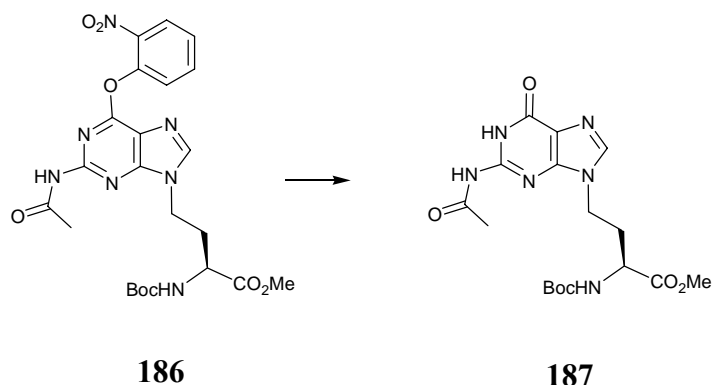
The first step in this pathway involved alkylation of the *N*9 position of **179** with propargyl bromide to afford **180** and then protecting the exocyclic NH<sub>2</sub> group with an isobutyryl group to give **181**. Unlike the benzoyl protecting group, in the case of the other nucleobases, the isobutyryl group is the desired NH<sub>2</sub> protecting group of guanine for oligonucleotide chemistry. The next step in the synthetic pathway involved replacing the 6-chloro group for an oxygen functionality that would subsequently yield the corresponding guanine carbonyl. There are many methods to achieve this, but we were inspired by a procedure reported by Howarth *et al.*<sup>106</sup> They successfully achieved this substitution quantitatively to afford **185** [Scheme 44], based on work by Reese *et al.*<sup>107, 108</sup> and Sproat *et al.*<sup>109</sup>





**Scheme 44.** Substitution of the chloro substituent. Reagents and conditions: 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h.

The final step in this synthetic pathway involves replacement of the 6-chloro group to give the 6-oxygen function required for the guanine alkyne component **182**. We were again inspired by a procedure reported by Howarth *et al.*<sup>106</sup> in which they had successfully achieved this to produce **187** [Scheme 45].

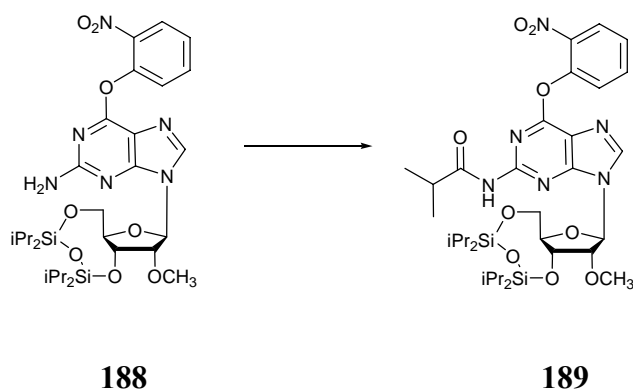


**Scheme 45.** Regeneration of the carbonyl moiety. Reagents and conditions: 1,1,3,3-tetramethylguanidine (TMG), 2-nitrobenzaloxime, CH<sub>3</sub>CN, rt, 24 h.

For the first alkylation step in this scheme, we were inspired by a procedure reported by Lindsell *et al.*<sup>105</sup> in which they had successfully prepared the same alkylated species **180** in a 57% yield from 2-amino-6-chloropurine **179** in the presence of propargyl bromide and Na<sub>2</sub>CO<sub>3</sub>. Thus, a suspension of **179** in dry DMF was treated with propargyl bromide in the presence of anhydrous K<sub>2</sub>CO<sub>3</sub>. It was decided to employ K<sub>2</sub>CO<sub>3</sub> in place of Na<sub>2</sub>CO<sub>3</sub>, despite this having been reported by Lindsell *et al.*<sup>105</sup> due to having successfully used K<sub>2</sub>CO<sub>3</sub> in the case of our cytosine alkylation (see the previous section 2.3.2, page 103). After 24 hours at rt, the reaction was worked up and purified by flash chromatography, to give **180** in a 77% yield. From the <sup>1</sup>H NMR spectrum recorded, the

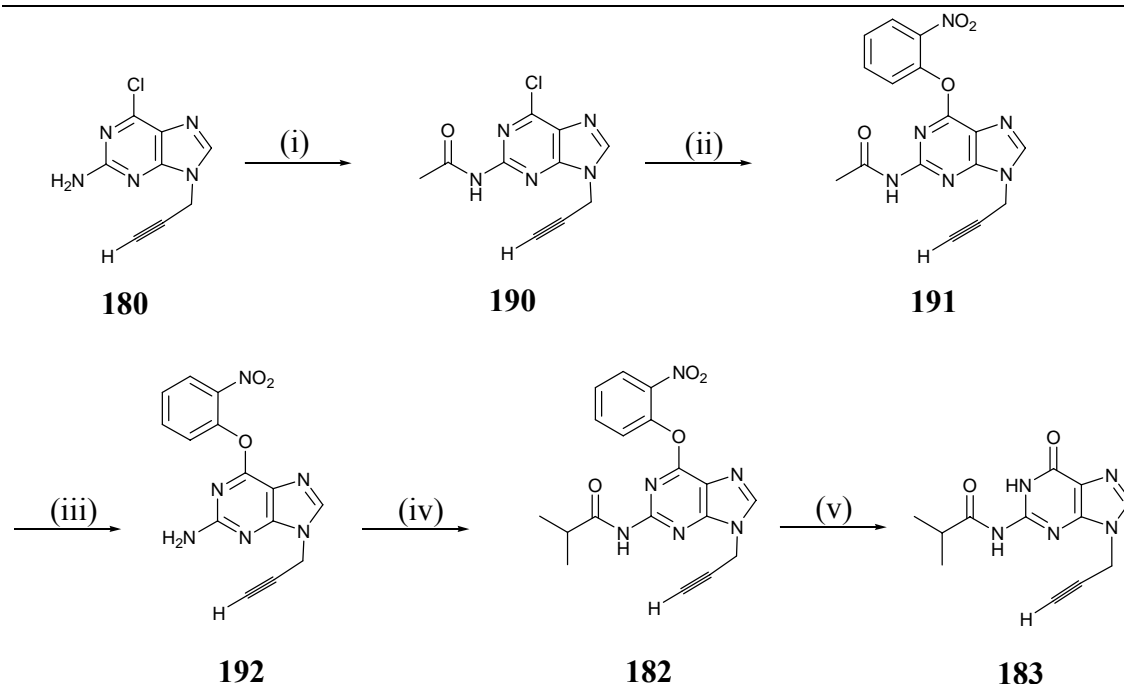
alkyne CH proton was observed as a triplet at 3.48 ppm and the alkyne CH<sub>2</sub> protons appeared as a doublet at 4.93 ppm. Whilst mass spectrometry and elemental analysis data confirmed that mono-alkylation had been achieved, we were unsure whether *N*9 or *N*7 alkylation had taken place exclusively. Due to there being no protons in the corresponding 6 position, it was difficult to confirm the alkylation site from 2D spectroscopy. It was hoped that characterisation of subsequent products in the reaction pathway would afford us with evidence of *N*9 alkylation.

Having successfully produced our alkylated compound **180**, the next step was to protect the exocyclic amine to give **181**. We were inspired by the procedure reported by Grötli *et al.*<sup>110</sup> in which they had successfully converted **188** into **189** in an 81% yield [Scheme 46].



**Scheme 46.** Isobutyryl protection. Reagents and conditions: *i*PrCOCl, DMAP, C<sub>5</sub>H<sub>5</sub>N, 40 °C, 8h.

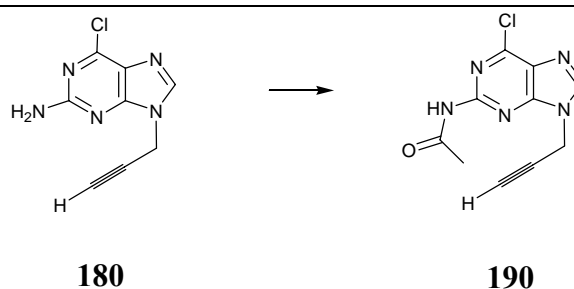
At this point we encountered problems in obtaining isobutyryl chloride quickly from our suppliers. In the interest of time, and in order to investigate the final two steps in our guanine alkyne component synthetic pathway (Scheme 43, page 109), it was decided to employ a different protecting group that could be replaced with an isobutyryl group at a later date. Thus, the modified Scheme 47 on the next page was devised and explored.



**Scheme 47.** Modified guanine route. Reagents and conditions: (i)  $Ac_2O$ , reflux, 3 h; (ii) 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h; (iii) 5% (v/v) methanolic ammonia, rt, 48 h; (iv)  $iPrCOCl$ , DMAP,  $C_5H_5N$ , 40 °C, 8h; (v) 1,1,3,3-tetramethylguanidine (TMG), 2-nitrobenzaloxime,  $CH_3CN$ , rt, 48 h.

Starting from the alkylated guanine **180**, we proposed to protect the exocyclic  $NH_2$  group with an acetyl group to give **190**. We then planned to substitute the 6-chloro group of **191** with a 2-nitrophenyl group as described before (page 109). The next step would then involve removal of the acetyl protecting group to give **192**. It was rationalised that this could be achieved using a solution of methanolic ammonia, as achieved in the case of the cytosine analogue (see the previous section 2.3.2, page 103). Subsequent re-protection with the desired isobutyryl protecting group, would then afford **182**. In the final step, regeneration of the carbonyl moiety could be achieved to give the final guanine alkyne component **183** as described before (page 113).

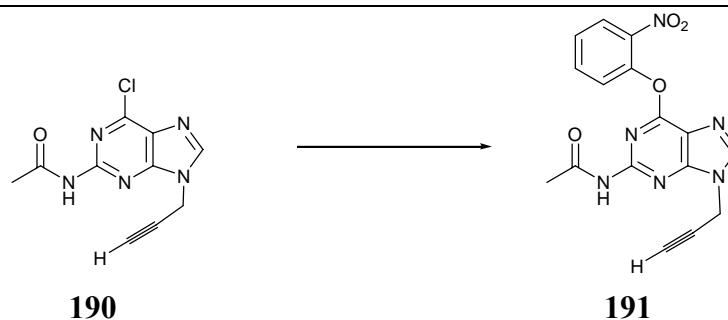
Thus, having already prepared the alkylated product **180**, from Scheme 43 (page 109), we needed to protect this with an acetyl group. To achieve this, we were inspired by a procedure reported by Lindsell *et al.*<sup>105</sup> in which they had successfully protected the exocyclic  $NH_2$  group of **180** using acetic anhydride to give **190** in a 32% yield [Scheme 48].



**Scheme 48.** Exocyclic  $\text{NH}_2$  protection with acetic anhydride. Reagents and conditions:  $\text{Ac}_2\text{O}$ , reflux, 3 h.

Thus, a solution of **180** in acetic anhydride was heated to reflux for 3 hours. After solvent evaporation *in vacuo* the crude material obtained was purified by flash chromatography to give **190** in an 82% yield. This yield was surprising since it was a lot higher than that reported by Lindsell *et al.*<sup>105</sup> They had reported having obtained a mixture of both the mono- and diacetylated species (15%), which made us question the purity of our product **190**. The I.R. spectrum obtained for **190** showed a strong carbonyl signal at  $1677\text{ cm}^{-1}$  and a medium NH stretching signal at  $3302\text{ cm}^{-1}$ . The  $^1\text{H}$  NMR spectrum recorded of **190** showed a singlet at 2.21 ppm consistent with the  $\text{CH}_3$  and a broad singlet at 10.90 ppm corresponding to the NH of the amide group. Further, the mass spectrometry data obtained was consistent with the mono-acetylated compound. None of the diacetylated species was detected.

The next step in our reaction pathway was to introduce a 2-nitrophenyl group into the C6 position by displacing the chloro substituent. As described earlier (page 113) we were inspired by the procedure reported by Howarth *et al.*<sup>106</sup> Thus, the protected 6-chloro derivative **190** was dissolved in dry DCM and to this TEA, 2-nitrophenol and DABCO were added. The reaction mixture was stirred for 24 hours at rt and following work-up and purification by flash chromatography, **195** was isolated in an 89% yield [Scheme 49].



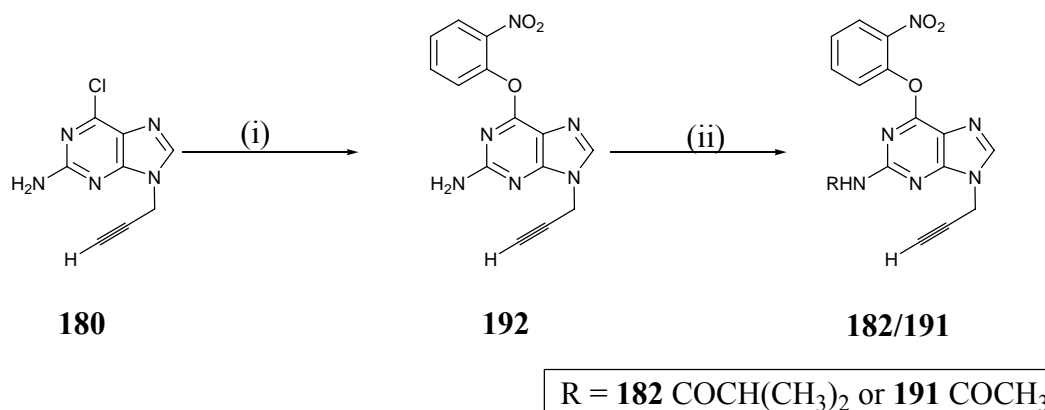
**Scheme 49.** Substitution of the chloro substituent. Reagents and conditions: 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h.

From the I.R. spectrum recorded for **191**, a weak C=O stretching signal at  $1623\text{ cm}^{-1}$  and a strong signal at  $1408\text{ cm}^{-1}$  were observed. In the  $^1\text{H}$  NMR spectrum recorded of **191**, the aromatic protons of the 2-nitrophenyl substituent were observed as two multiplet signals at 7.42-7.51 ppm and 7.71-7.75 ppm. In the  $^{13}\text{C}$  NMR spectrum recorded, the two quaternary carbons for the 2-nitrophenyl substituent were observed at 145 ppm and 151 ppm. Although there was a negligible shift from 159.91 ppm to 159.10 ppm for the C6 position, there was no indication of the chloro substituent being present in the molecular weight obtained from mass spectroscopy data collected.

The next step in Scheme 47 (page 115), involved removing the acetyl protecting group to regenerate the  $\text{NH}_2$  group of **192**. Having successfully achieved the same outcome with our cytosine analogue **174**, we were inspired to employ the same reaction conditions as before (see section 2.3.2, page 107). Thus, **191** was dissolved in a 5% (v/v) methanolic ammonia solution and the resulting mixture was left to stir at rt for 24 h [Scheme 47]. Disappointingly, from TLC analysis and the  $^1\text{H}$  NMR spectrum recorded on the crude product isolated after the work-up, we concluded that no reaction had occurred. Thus, another aliquot of methanolic ammonia solution was added and the reaction mixture was left for a further 24 hours at rt. Unfortunately, after removing the solvent under reduced pressure, still no product **192** was present according to the  $^1\text{H}$  NMR spectrum subsequently recorded.

Having not been able to remove the acetyl protecting group of **191**, we then decided to investigate if direct substitution of the chloro substituent from **180** could be achieved to afford **192** [Scheme 50]. Not only would this shorten the overall reaction scheme, but would also alleviate our problems encountered in removing the acetyl protecting group

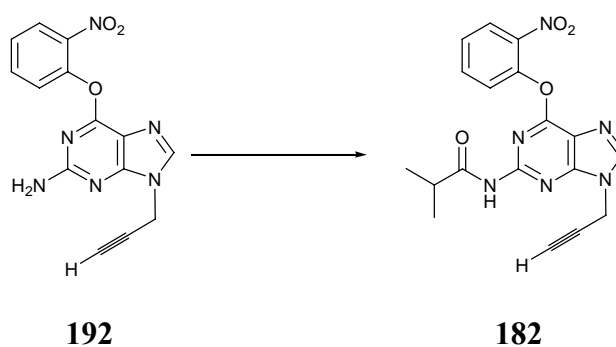
of **195**. Therefore, we exposed our alkylated 2-amino-6-chloropurine derivative, **180**, to the same reaction conditions as shown in step 2 (Scheme 47, page 117) [Scheme 50].



**Scheme 50.** Substitution of the chloro substituent. Reagents and conditions: (i) 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h; (ii) AcCl or <sup>i</sup>PrCOCl, DMAP, C<sub>5</sub>H<sub>5</sub>N, 40 °C, 8h.

Thus, **180** was dissolved in dry DCM and to this TEA, 2-nitrophenol and DABCO were added. The reaction mixture was stirred for 24 hours at rt and following work-up and purification by flash chromatography, **192** was afforded in a 43% yield. The NH<sub>2</sub> peak was strong at 3293 cm<sup>-1</sup>. The <sup>1</sup>H NMR spectrum again showed, the aromatic protons of the 2-nitrophenyl substituent were two multiplet signals at 7.49-7.62 ppm and 7.80-7.89 ppm.

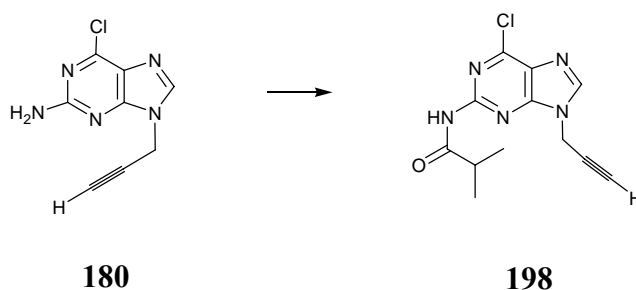
Now that we had prepared **192**, and isobutyryl chloride had been obtained from our suppliers, protection of the NH<sub>2</sub> group was attempted using the procedure reported by Grötli *et al.*<sup>110</sup> (Scheme 46, page 115) [Scheme 51].



**Scheme 51.** NH<sub>2</sub> protection. Reagents and conditions: <sup>i</sup>PrCOCl, DMAP, C<sub>5</sub>H<sub>5</sub>N, 40 °C, 8h.

Thus, **192** was taken up in dry pyridine and to this, DMAP and isobutyryl chloride were added. Following work-up and purification by flash chromatography, **182** was afforded in a 16% yield. From the  $^1\text{H}$  NMR spectrum recorded, the two  $\text{CH}_3$  groups from the isobutyryl protecting group were observed as a doublet at 1.06 ppm and the CH was seen as a septet at 2.95 ppm.

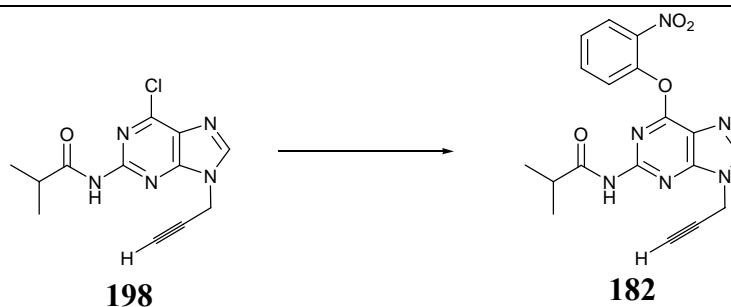
In order to understand this low yield, it was postulated that perhaps the nature of the substituent at the  $C6$  position influenced the nucleophilicity of the  $N2$  amino function, and/or steric hindrance was accountable. We therefore decided to revert back to our initial guanine synthetic pathway (Scheme 43, page 114) and repeat this step using our previously prepared alkylated compound **180** [Scheme 52].



**Scheme 52.** Exocyclic  $\text{NH}_2$  protection with isobutyryl chloride. Reagents and conditions:  $i\text{PrCOCl}$ , DMAP,  $\text{C}_5\text{H}_5\text{N}$ , rt, 24 h.

Thus, **180**, DMAP and isobutyryl chloride were taken up in dry pyridine at rt and left stirring for 24 h [Scheme 52]. Following work-up and purification by flash chromatography, **198** was afforded in an 88% yield. The  $^1\text{H}$  NMR spectrum recorded of **198**, showed a methyl doublet at 1.29 ppm and a septet at 2.93 ppm for the CH of the isobutyryl group. The increase in percentage yield obtained when  $\text{NH}_2$  protection was carried out in the presence of a 6-chloro substituent supported our theory that either activation of the  $\text{NH}_2$  group was achieved or that the smaller chloro group provided less steric hindrance.

The next step involved substitution of the 6-chloro substituent for a 2-nitrophenol group. This was achieved using the same reaction conditions as shown in Scheme 43 (page 114) [Scheme 53].

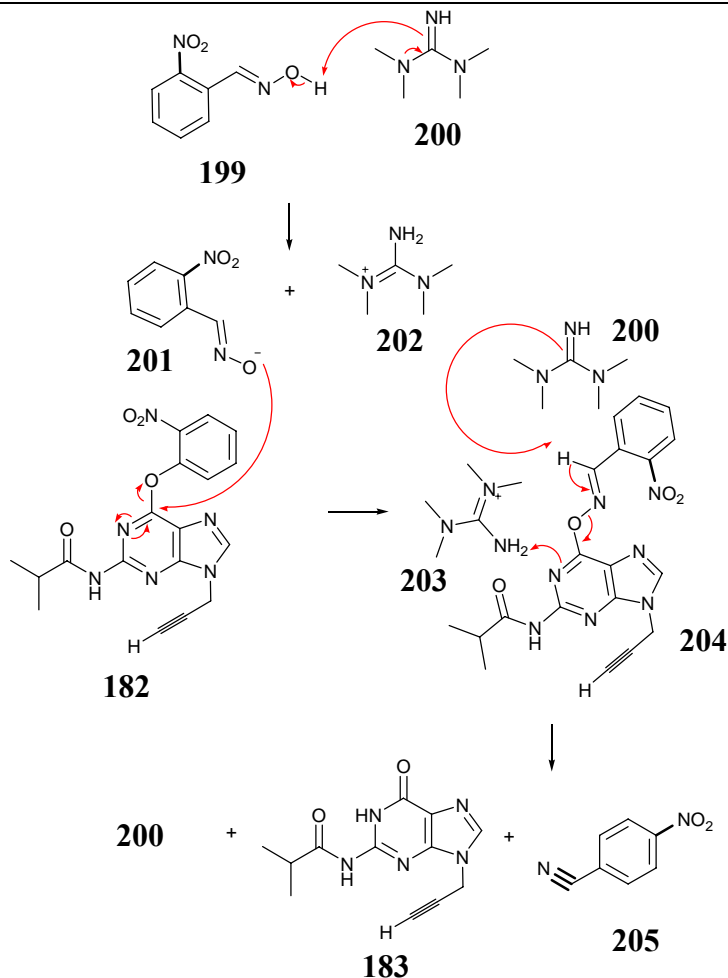


**Scheme 53.** Substitution of the chloro substituent. Reagents and conditions: 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h.

Following work-up and purification by flash chromatography, **182** was afforded in a 89% yield. Both the I.R. and the  $^1\text{H}$  NMR spectrum recorded for **182** were consistent with the formation of **182**.

The final step to afford our protected guanine alkyne component **183**, required us to regenerate the carbonyl substituent at the C6 position to give **183**. As described earlier (page 113) we were inspired by the procedure by Howarth *et al.*<sup>106</sup> Thus, a solution of 1,1,3,3-tetramethylguanidine (TMG) **200** in dry acetonitrile was added to a stirred solution of **182** and 2-nitrobenzaloxime **199** in dry acetonitrile. The reaction mixture was stirred at rt and monitored periodically by TLC over a 48 hour period. TLC and the  $^1\text{H}$  NMR spectrum recorded of crude reaction mixture confirmed that the reaction had not taken place and, unfortunately, the starting material **182** was recovered quantitatively. In attempt to understand this failure, we decided to evaluate a possible mechanism behind this specific reaction.





**Scheme 54.** Proposed mechanism for the regeneration of the carbonyl moiety.

Scheme 54 above shows our proposed mechanism for the regeneration of the carbonyl function of **183** from **182**. 2-Nitrobenzaldehyde **199** and TMG **200** react together to form **201** and **202**. **201** goes on to substitute the 2-nitrophenol group in the C6-position to give **204**. Subsequently, excess **200** removes the acidic proton of **204** to afford the target compound **183**. This mechanism is not simple and may provide the basis from which future endeavours into regenerating the carbonyl moiety of guanine can begin.

Although our efforts at producing our final guanine alkyne component **183** were unsuccessful, we were satisfied that having obtained several protected-guanine derivatives, namely **182**, with the alkyne group present, that this would be sufficient to at least attempt our guanine analogue ‘click’ reaction (see the later section 2.4.5, page 142). It was hypothesised that once the ‘click’ product was obtained, that further modification, regeneration of the carbonyl moiety, could then be attempted.

---

### 2.3.5 Conclusions

The thymine alkyne component was achieved by first protecting the *N3* position with a benzoyl group to increase the lipophilicity of the compound. Having successfully achieved this in a 73% yield, we then alkylated **150** using propargyl bromide to regenerate our alkylated thymine monomer **161** in a 55% yield. Our initial failure using NaH had led us to substitute this base with K<sub>2</sub>CO<sub>3</sub> which had resulted in this successful alkylation. The final (optional) step involved removal of the benzoyl protecting group to generate the NH in the *N3* position. We had been inspired by conditions reported by Maguire *et al.*<sup>102</sup> in which debenzoylation had been achieved using benzyl alcohol. Thus, these conditions were applied to **161** and the thymine monomer **162** was afforded in an 82% yield.

Development of the cytosine alkyne component **166** was less straightforward. Our initial route towards this target involved benzoylating the exocyclic NH<sub>2</sub> group of commercially available cytosine. Despite having been unable to verify the structure of **165** by <sup>1</sup>H NMR spectroscopy, due to its poor solubility in any of the deuterated solvents available, it was decided to pursue the next step in the pathway and achieve the alkylated product **166**. A procedure by Guillaume *et al.*<sup>104</sup> was then investigated using DBU, however, this reaction failed. In the process of investigating alternative alkylation conditions in the literature, we came across the research of Lindsell *et al.*<sup>105</sup> in which they advised using commercially available *N4*-acetylcytosine **171** to prevent competing *O2* alkylation from occurring when trying to alkylate in the *N1* position. Since this group of researchers had successfully prepared the same alkylated species **173** using this starting material, we decided to replicate their synthetic pathway towards **173**. Re-protection of the ensuing exocyclic NH<sub>2</sub> group could then be achieved using the benzoylation conditions previously used in the case of our thymine monomer. Thus, the alkylated species **166** and then the amine species **174** were afforded in a 40% and 94% yield respectively, following their work. Protection of the NH<sub>2</sub> group was then achieved with a benzoyl group to afford **166** in an 84% yield, in an analogous manner to that achieved for the thymine component.

In the case of developing a synthetic route towards the adenine alkyne component **176**, our approach was slightly different. For purine manipulation, alkylation has to be carried out prior to the protection of the exocyclic amine function, contrary to the

sequence of steps discussed previously in the case of the two pyrimidine nucleobases. Thus, *N9* alkylation with propargyl bromide was first investigated. Alkylation conditions used in the case of the thymine monomer were applied and the adenine alkylated product **175** was afforded in a 36% yield. This yield was comparable to others in terms of typical adenine percentage yields. Once we had confirmed that *N9* alkylation had occurred exclusively, from NOESY spectroscopy, we were then in the position to protect the exocyclic NH<sub>2</sub> group with a benzoyl group to afford the target adenine alkyne component **176**. This was achieved in a 67% yield following a procedure by Lindsell *et al.*<sup>105</sup> using BzCl in dry pyridine.

The final nucleobase investigated towards the development of its alkyne component was guanine. It is widely accepted in the literature that when manipulating guanine, it is best to start from commercially available 2-amino-6-chloropurine **179**. Therefore, the *N9* position of **179** was first alkylated using conditions reported by Lindsell *et al.*<sup>105</sup> in which they had successfully prepared the same alkylated species. Thus, **180** was obtained in a 77% yield.

The next step involved protecting the exocyclic amine function with an isobutyryl group. However, we experienced problems getting this reagent from our suppliers. In light of this, we decided to explore this protection using an alternative group which could be replaced with an isobutyryl group once it was available. Lindsell *et al.*<sup>105</sup> had opted to protect this function with acetic anhydride and so we decided to reproduce their conditions. **190** was obtained in an 82% yield.

The next investigation involved substituting the 6-chloro group for a 2-nitrophenol group. This was achieved using conditions reported by Howarth *et al.*<sup>106</sup> and **191** was afforded in an 89% yield.

It was then attempted to remove the acetyl protecting group, which had been successful in the cytosine example described previously. Thus, the same deprotection conditions were employed however, unfortunately, no reaction had taken place.

It was then decided to investigate if direct substitution of the chloro substituent with the 2-nitrophenol group could occur from the alkylated compound **180**. Therefore, **180** was subjected to the same substitution conditions previously reported and the desired

product **193** was obtained in a 43% yield. At this point, we had managed to obtain isobutyryl chloride from our suppliers and therefore, the NH<sub>2</sub> protection using this reagent was carried out on **182**. Although the reaction had been successful, the desired product was only obtained in a 16% yield which was unsatisfactory.

It was then decided to investigate if the nature of the substituent at the C6 position influenced the nucleophilicity of the N2 amino function. Thus, the isobutyryl chloride conditions were repeated using the 6-chloro alkylated compound **180**. In this case, the desired product was produced in an 88% yield which confirmed our suspicions that the C6 substituent does influence the nucleophilicity of the N2 amino function. In the later case, the smaller chloro substituent allowed for easier protection of the NH<sub>2</sub> group, and in the case of the 2-nitrophenyl substituent in the C6 position, protection was more difficult due to steric hindrance. Further, electronic reasons could also influence the nucleophilicity of the N2 amino function.

Substitution of the chloro group for a 2-nitrophenol group was then achieved using the same conditions reported previously in the case of the alkylated isobutyryl protected compound **198**. The desired product **182** was afforded in an 89% yield.

The next step in the synthetic pathway towards the guanine alkyne-containing component was to remove the 2-nitrophenol group and regenerate the carbonyl substituent of guanine. To achieve this, we were inspired by the procedure reported by Howarth *et al.*<sup>106</sup> By treating **182** with 2-nitrobenzaldoxime **199** and TMG **200** as their procedure suggested, unfortunately, resulted in no product **183** being observed. Due to time constraints, further investigation into the reaction was not undertaken. However, the proposed mechanism (Scheme 54, page 123) is not simple and, upon further inspection, may account for the problems we encountered. Instead, the starting materials were recovered quantitatively. Despite having not been able to generate the target guanine alkyne component **183**, we had however, prepared several protected guanine intermediates (namely **182**) containing an alkyne moiety which could, in theory, be used in the ‘click’ reaction. Subsequent modification of the nucleobase component could then be achieved following the ‘click’ reaction.

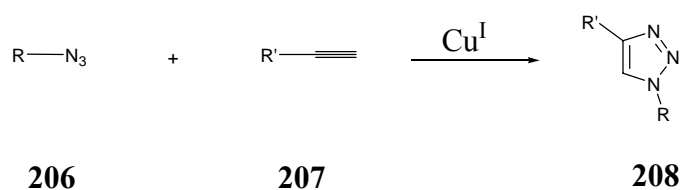
## 2.4 DEVELOPMENT OF THE 'CLICK' CHEMISTRY

### 2.4.1 Introduction

In 2001, a landmark review was published by Sharpless *et al.*<sup>81</sup> describing a novel strategy for organic chemistry, coined 'click' chemistry. Dr Sharpless defined 'click' chemistry as a group of reactions that "...must be modular, wide in scope, give very high yields, generate only inoffensive byproducts that can be removed by nonchromatographic methods, and be stereospecific (but not necessarily enantioselective). The required process characteristics include simple reaction conditions (ideally, the process should be insensitive to oxygen and water), readily available starting materials and reagents, the use of no solvent or a solvent that is benign (such as water) or easily removed, and simple product isolation. Purification, if required, must be by non-chromatographic methods, such as crystallisation or distillation, and the product must be stable under physiological conditions."<sup>81</sup>

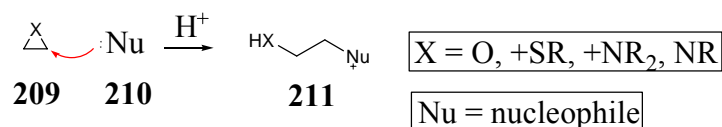
As previously stated, 'click' chemistry encompasses a group of reactions and to date, these reactions can be classified into four main categories:

**(1) Cycloadditions:** These are reactions which primarily refer to 1,3-dipolar cycloadditions [Scheme 55]. However, they also include hetero-Diels Alder cycloadditions.<sup>111</sup>



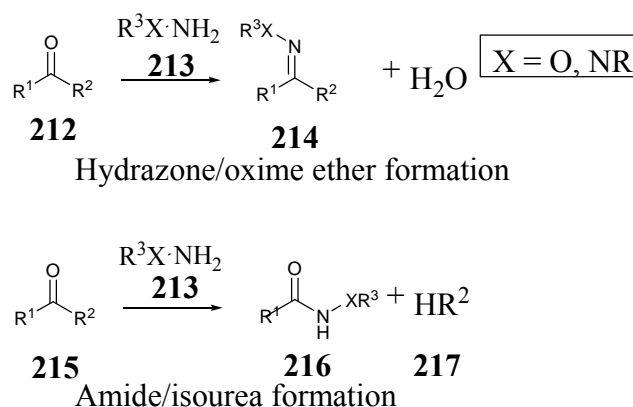
**Scheme 55.** Cycloadditions ( $\text{Cu}^{\text{I}}$  catalysed Huisgen 1,3-dipolar cycloaddition of azides and terminal alkynes).

**(2) Nucleophilic ring-openings:** These are reactions which refer to the opening of strained heterocyclic electrophiles, e.g. epoxides, cyclic sulfates, aziridines, aziridinium ions and episulfonium ions [Scheme 56].<sup>111</sup>



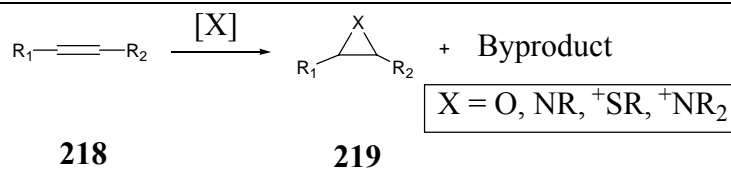
Scheme 56. Nucleophilic ring openings

**(3) Carbonyl chemistry of the non-aldol type:** These reactions include for example, the formation of hydrazones, ureas, oximes, thioureas, amides, ethers and aromatic heterocycles [Scheme 57].<sup>111</sup> Aldol type reactions cannot be considered ‘click’ reactions due to their thermodynamic driving forces being typically slow, resulting in long reaction times, and due to the formation of side products.<sup>111</sup>

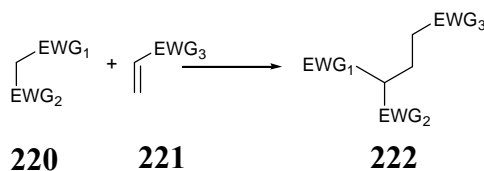


Scheme 57. Non-Aldol carbonyl chemistry.

**(4) Additions to carbon-carbon multiple bonds:** These reactions include for example, aziridinations, epoxidations, nitrosyl halide additions, dihydroxylations, sulfonyl halide additions and certain Michael additions [Scheme 58].<sup>111</sup>



Formation of three membered rings



Certain Michael additions

EWG = electron withdrawing group

**Scheme 58.** Carbon multiple bond additions.

Amongst these four main classifications, cycloadditions and, in particular, the Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition of azides and terminal alkynes is the most studied and widely used exemplar.<sup>111</sup>

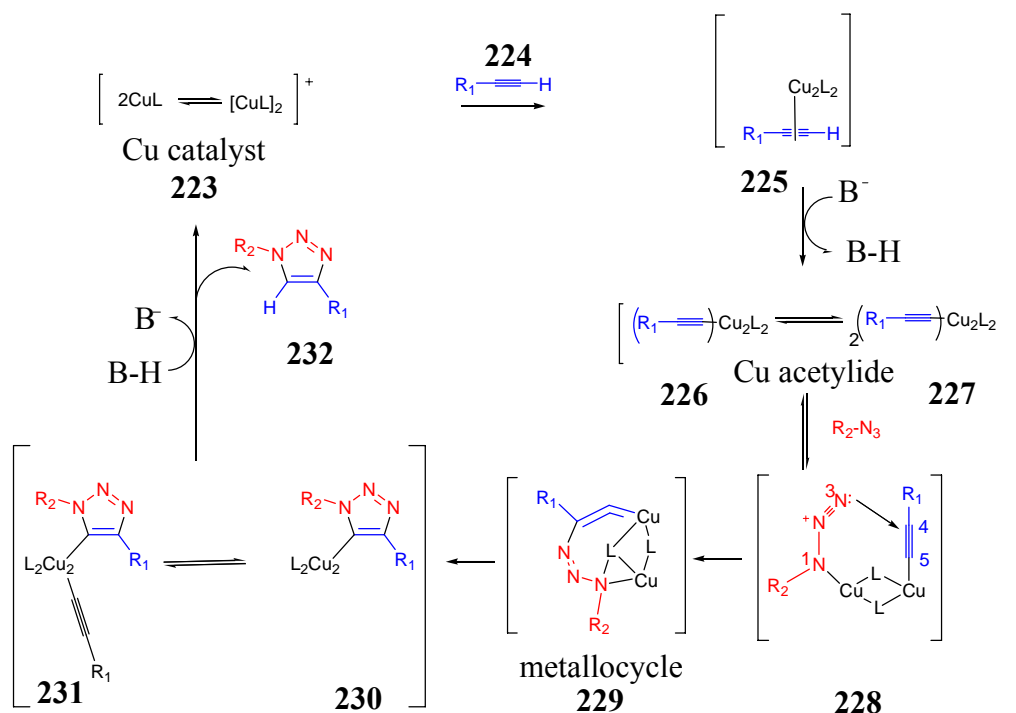
#### 2.4.2 Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition of azides and terminal alkynes

This particular cycloaddition fulfils the ‘click’ chemistry criteria completely and results in the formation of 1,2,3-triazoles. Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloadditions are reliable, easy to use and are regiospecific, forming the 1,4-substituted product exclusively.<sup>111</sup> The reaction can be performed at temperatures between 0 and 160°C (although the reactions typically do not require heat elevation), over a pH range of 5 to 12 and in several different solvents (including water).<sup>111</sup> This catalysed version proceeds 10<sup>7</sup> times faster than the uncatalysed version and is unaffected by steric factors.<sup>111</sup> Primary, secondary, tertiary and aromatic azides all participate in this cycloaddition.<sup>111</sup> Further, azides and terminal alkynes can both tolerate oxygen, water, biological molecules, common organic synthetic conditions, a variety of pHs and solvents and the reaction conditions of living systems e.g. hydrolysis, reducing environment.<sup>111</sup> Although the decomposition of aliphatic azides is thermodynamically favoured, they are stable under the aforementioned conditions due to a kinetic barrier that exists.<sup>111</sup> Thus, they are invisible in solution until they come into contact with a dipolarophile, the alkyne.<sup>111</sup>

*The mechanism of the Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition*

Cycloadditions generally proceed in a concerted mechanism, however, both molecular modelling and experimental data on this particular cycloaddition has shown it to be a stepwise process.<sup>111</sup> The catalysed concerted reaction has been shown, using density functional theory calculations, to have a greater activation barrier by 1.8 kcal/mol, than for the uncatalysed concerted reaction.<sup>111</sup> Furthermore, the stepwise-catalysed Huisgen 1,3-dipolar cycloaddition reaction has an activation barrier which is 11 kcal/mol lower than that for the catalysed concerted reaction.<sup>111</sup>

Due to these findings, and given that Cu<sup>I</sup> is known to readily insert itself into terminal alkynes in a Sonogashira type coupling, it has been proposed that the mechanism of the Huisgen 1,3-dipolar cycloaddition is that shown in Scheme 59.<sup>111</sup>



**Scheme 59.** Proposed mechanism for the Cu<sup>I</sup> catalysed Huisgen 1,3-dipolar cycloaddition (redrawn from ref 107).

From Scheme 59 above, the first step involves  $\pi$  complexation of the Cu<sup>I</sup> dimer **223** to the alkyne **224** to afford **225**.<sup>111</sup> Deprotonation of the terminal hydrogen then occurs, which forms the Cu-acetylide **226/227** equilibrium.<sup>111</sup> Depending on the reaction conditions used, there are a variety of different kinds of Cu-acetylide complexes which can result.<sup>111</sup>  $\pi$  complexation of the Cu<sup>I</sup> serves to lower the pK<sub>a</sub> of the terminal alkyne



by up to 9.8 pH units, which allows the deprotonation to occur in the absence of base and in aqueous solvent.<sup>111</sup> The next step involves displacing one of the ligands, L, from the second Cu in the Cu-acetylide **226/227** complex with nitrogen, to form the dicopper species **228** which activates the azide for nucleophilic attack.<sup>111</sup> N3 is then in close enough proximity to attack C4 of the alkyne, resulting in the formation of a metallocycle **229**.<sup>111</sup> This then contracts once the lone pair of electrons on N1 has attacked C5, forming the respective triazole in **230**.<sup>111</sup> The attached Cu dimer then complexes to a second terminal alkyne, which cannot undergo a further cycloaddition due to the unfavourable structure of the ensuing complex **231**, and dissociates upon protonations to reform **223**.<sup>111</sup> A final protonation releases the Cu<sup>I</sup> catalyst to generate the 1,2,3-triazole product **232**, and allows the catalyst to undergo a second cycle with other substrates.<sup>111</sup> Although further studies are required to conclusively confirm that both these protonations are as a direct result of interactions with protonated external base and/or solvent, it is expected.<sup>111</sup>

#### *The copper catalyst*

The catalyst required for the 1,3-dipolar cycloaddition of azides and terminal alkynes can be generated in several different ways. Perhaps the most common of these techniques is to reduce Cu<sup>II</sup> salts *in situ* to produce Cu<sup>I</sup> salts.<sup>111</sup> CuSO<sub>4</sub>·5H<sub>2</sub>O is normally reduced by sodium ascorbate, but other reducing agents including tris(2-carboxyethyl)phosphine and hydrazine can also be used.<sup>111</sup> The reagents in this strategy are relatively cheap and the reaction can be carried out in water, which consequently prevents the atmosphere having to be deoxygenated.<sup>111</sup> The reaction is also therefore, environmentally safe.<sup>111</sup> By being able to remove the need for using base in the reaction, by using water as a solvent, the need for protecting groups is also eliminated.<sup>111</sup> The major limitation of this method is that the reducing agent may additionally reduce the Cu<sup>II</sup> salt to Cu<sup>0</sup>, although this can normally be prevented by using a correct ratio of reducing agent to catalyst or by the addition of a copper stabilising agent such as tris-(hydroxypropyltriazolylmethyl)amine (THPTA) to the reaction.<sup>111</sup>

The catalyst can also be generated by the direct addition of Cu<sup>I</sup> salts to the reaction. The copper salts that can be added include, CuBr, CuI, CuOTf·C<sub>6</sub>H<sub>6</sub> and [Cu(NCCH<sub>3</sub>)<sub>4</sub>][PF<sub>6</sub>] for example.<sup>111</sup> A reducing agent is not required for this technique, but an organic solvent (or a mixed solvent) and hence a deoxygenated atmosphere is

necessary.<sup>111</sup> Protecting groups are also required along with base.<sup>111</sup> Although direct addition of Cu<sup>I</sup> salts can be successful, this method is not as reliable as the Cu<sup>II</sup> technique described previously.

Oxidation of copper metal with an amine salt can also be used to form the active catalyst, albeit with several disadvantages. Longer reaction times as well as large quantities of expensive copper are the main problems with this technique.<sup>111</sup> Slightly acidic environments are also necessary to dissolve the metal, which may be detrimental if acid sensitive protecting groups are contained within the reactants.<sup>111</sup>

Another way of producing the active catalyst is to use Cu<sup>0</sup> nanosize clusters.<sup>112</sup> These clusters effectively catalyse the cycloaddition reaction without the need for an amine hydrochloride salt.<sup>112</sup> The reaction is thought to occur on the surface of the nanoclusters rather than in solution and the oxidation state of the copper during this process is likely to be Cu<sup>I</sup>.<sup>112</sup> Cycloadditions using nanosize Cu<sup>0</sup> demonstrate the same broad range of high yielding reactions as expected with ‘click’ chemistry however, the major disadvantage is that solvation of these nanosize particles necessitates a slightly acidic environment.<sup>112</sup> Hence, the need for protection of acid sensitive functional groups.<sup>112</sup> Further, Cu<sup>0</sup> nanosize clusters are currently not commercially available, and the Cu<sup>0</sup> nanoscale powder required to produce them, is greater than seven times more expensive than the other copper sources discussed in this section.<sup>112</sup>

A catalyst may not always be necessary for this reaction, when the reactant alkyne is sufficiently electron deficient. However, electron deficient alkynes, are themselves, extremely reactive towards nucleophiles and may result in unwanted side reactions which would defeat the previously outlined set of criteria for a ‘click’ reaction (page 125).<sup>111</sup>

#### *‘Click’ chemistry applications*

Since its debut in 2001, there has been an explosion of growth in publications relating to a wealth of applications of ‘click’ chemistry. The fields of bioconjugation, material science and drug discovery have perhaps benefited the most from this practical and sensible chemical approach. The reader is referred to a review by Moses *et al.*<sup>113</sup> for a comprehensive guide to the recent applications of ‘click’ chemistry.

---

*The problems associated with 'click' chemistry*

There are several problems associated with 'click' chemistry and in particular, the Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition variant. Firstly, if the diene, the azide in this case, is too electron deficient then no reaction will take place.<sup>111</sup> This is due to the ground state configuration being too low for it to successfully interact with the dienophile, the terminal alkyne.<sup>111</sup> Similarly, the dienophile cannot be too electron rich.<sup>111</sup> This circumstance is unlikely though as it would require functional groups not commonly found in biological systems or administered drugs.<sup>111</sup>

Alkyne homocoupling is a more common problem. This arises when an alkyne reacts with a second alkyne in place of the azide.<sup>111</sup>

Another problem is Cu<sup>I</sup> saturation. The Cu<sup>I</sup>-acetylide complex must have physical contact with the azide in order for a successful 'click' reaction to result.<sup>111</sup> If however, the complex becomes surrounded by terminal alkynes, a chelation effect with the alkynes results and the complex becomes saturated.<sup>111</sup> Thus, azide functional groups are effectively prevented from having contact with the complex and performing the displacement.<sup>111</sup> Although Cu<sup>I</sup> saturation is rare, it occurs when the dipolarophile contains multiple terminal alkynes that are capable of coordinating to a single site.<sup>111</sup>

Azide stability may also be a limitation in these cycloaddition reactions. When the ratio of nitrogen atoms to carbon atoms exceeds, or is equal to one, in an organic molecule, then the molecule is deemed explosive and therefore, very dangerous.<sup>111</sup> A slight input of external energy would have disastrous consequences.<sup>111</sup> Fortunately, large molecules with high carbon content are usually used in pharmaceutical research.

Despite copper being required in the body, excessive intake can have fatal consequences. The side effects of increased copper levels include kidney diseases, neurological disorders, hepatitis and Alzheimer's disease.<sup>111</sup> Copper is toxic at high levels due to it easily accepting and donating single electrons, thereby, changing electronic states which allows it to catalyse toxic reactions, such as the *in vivo* reduction of hydrogen peroxide to form hydroxyl free radicals.<sup>111</sup> The copper catalyst must therefore be completely removed if *in vivo* applications of 'click' chemistry are to be found.<sup>111</sup>

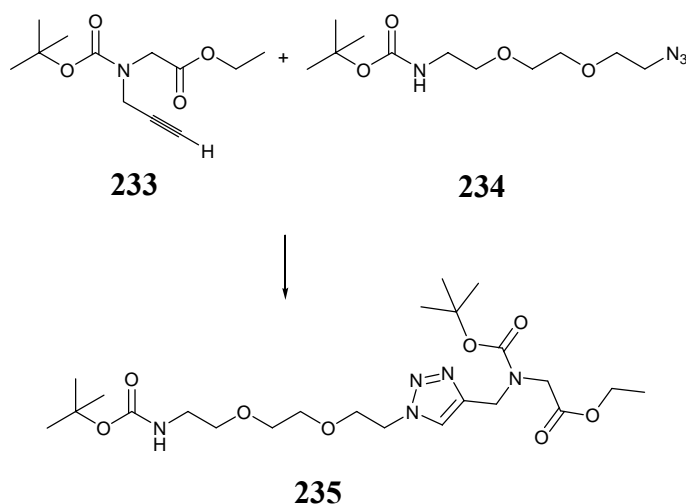
---

Another challenge to overcome is the biocompatibility of 1,2,3-triazoles. Although they were first discovered over a decade ago, little information is known about their biological pathway.<sup>111</sup> Individual toxicities of the 1,2,3-triazoles have been extensively researched, however, there are currently no reports reviewing their *in vivo* metabolism.<sup>111</sup> This is surprising since 1,2,3-triazoles are known bioisosteric replacements for amides.<sup>111</sup>

The final limitation, or mere inconvenience, of ‘click’ chemistry at present is that, currently, many of the ‘click’ building blocks are not commercially available. Although in the future this will no doubt not be an issue, at the present time it can be seen as a daunting task.

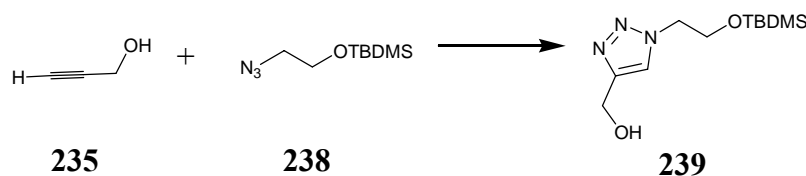
#### **2.4.3 Model ‘click’ reactions**

In order to minimise the potential loss of both our precious serine derived azide components and nucleobase containing alkyne components, it was decided to first investigate the ‘click’ reaction conditions using structurally similar, readily available models. Once the conditions have been optimised, these would be applied to our desired ‘click’ reactions. From the findings discussed in section 2.4.1 of this chapter (pages 126-133), it was decided to initially adopt the approach of reducing Cu<sup>II</sup> salts *in situ* to Cu<sup>I</sup> salts using sodium ascorbate. This is the commonest strategy used for accomplishing this cycloaddition reaction. We were inspired by the procedure reported by Goujon *et al.*<sup>114</sup> in our research group, in which **235** had been prepared from azide **233** and alkyne **234** in a 76% yield [Scheme 60].



**Scheme 60.** 'Click' conditions. Reagents and conditions: **233**, **234**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 24 h.

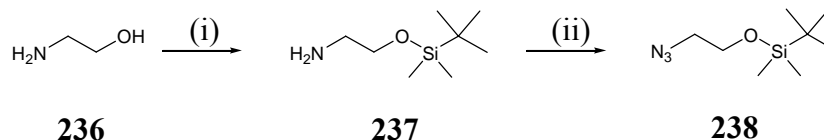
The first of our model reactions involved employing commercially available propargyl alcohol **235**, as the substitute thymine-containing alkyne component **161**, and the model azide **238** shown below [Scheme 61].



**Scheme 61.** Model 'click' reactions. Reagents and conditions:  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ .

Propargyl alcohol **235** was selected as a suitable alkyne component substitute as it was both commercially available and would prevent complication of the resulting triazole proton in the respective  $^1\text{H}$  NMR spectrum recorded of product **239**. It was decided to employ the model azide **238** as this would also prevent complication of the  $^1\text{H}$  NMR spectrum obtained for **239** and further, it would allow the silicon lability to be assessed.

Azide **238** first had to be prepared and this was achieved starting from commercially available ethanolamine **236** [Scheme 62].



**Scheme 62.** Preparation of the model azide **238**. Reagents and conditions: (i) TBDMSCl, imidazole, DCM, rt, 24 h; (ii) **152**, NaHCO<sub>3</sub>, CuSO<sub>4</sub>·5H<sub>2</sub>O, H<sub>2</sub>O, MeOH, rt, 24h.

The first step in this synthetic pathway to model azide **238** required us to protect the primary alcohol of **236** with a silicon group to produce **237** and then secondly, to convert the amine into azide **238** using the previously prepared ‘diazo donor’ **152** (see the previous section 2.2.2, page 94).

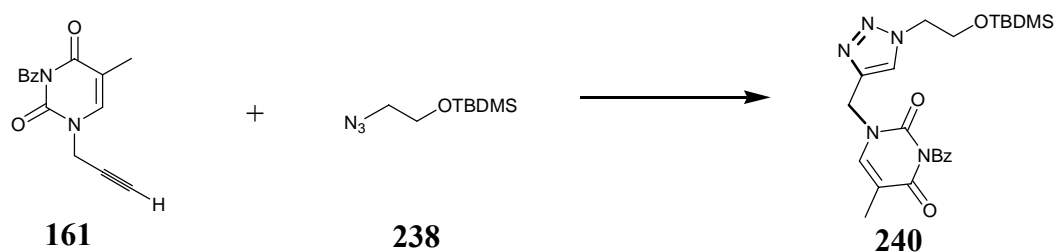
Thus, **237** was afforded in a 53% yield by treating **236** with TBDMSCl in DCM for 24 h at rt. From the <sup>1</sup>H NMR spectrum recorded of **237**, a singlet corresponding to nine *tert*-butyl protons at 0.31 ppm and a second singlet at 1.14 ppm corresponding to six protons of the two CH<sub>3</sub> groups attached to the silicon, indicated that successful alcohol protection had occurred.

**237** was then subjected to the diazotransfer conditions described previously (see section 2.2.2, page 94). The azide signal was observed at 2114 cm<sup>-1</sup> in the crude I.R. spectrum recorded of **238**. The crude <sup>1</sup>H NMR spectrum recorded showed two singlets at 0.09 ppm and 0.91 ppm corresponding to the two and three methyl groups respectively on the silicon moiety, and two triplets at 3.27 ppm and 3.80 ppm consistent with the protons from the two CH<sub>2</sub> groups. There was also a broad singlet at 5.30 ppm which indicated that the diazotransfer reaction had not been quantitative, and that some starting amine **237** was still present.

Thus, azide **238** and alkyne **235** were then subjected to the ‘click’ conditions reported by Goujon *et al.*<sup>114</sup> [Scheme 61]. **238** was taken up in an equal mixture of DCM, *tert*-butyl alcohol and H<sub>2</sub>O and to this, **235**, sodium ascorbate, and CuSO<sub>4</sub>·5H<sub>2</sub>O (1:2) were added. After 24 h stirring at rt, and following work-up and purification by flash chromatography, **239** was afforded in a 70% yield. From the <sup>1</sup>H NMR spectrum recorded of **239**, the triazole proton was clearly seen as a singlet at 7.61 ppm. The OH

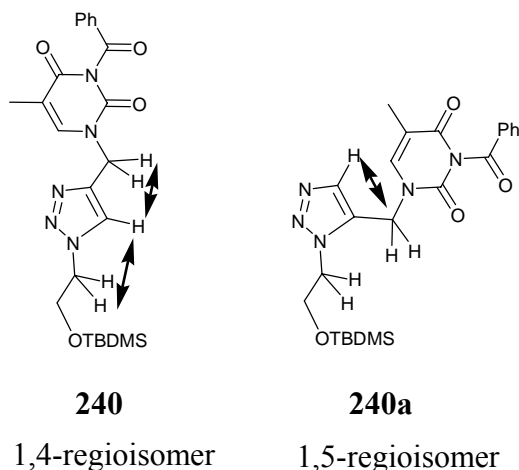
was observed as a broad singlet at 3.12 ppm and the silicon moiety was shown to be untouched.

The next model reaction involved employing the actual thymine alkyne component **161** together with the previously prepared model azide **238** [Scheme 63].



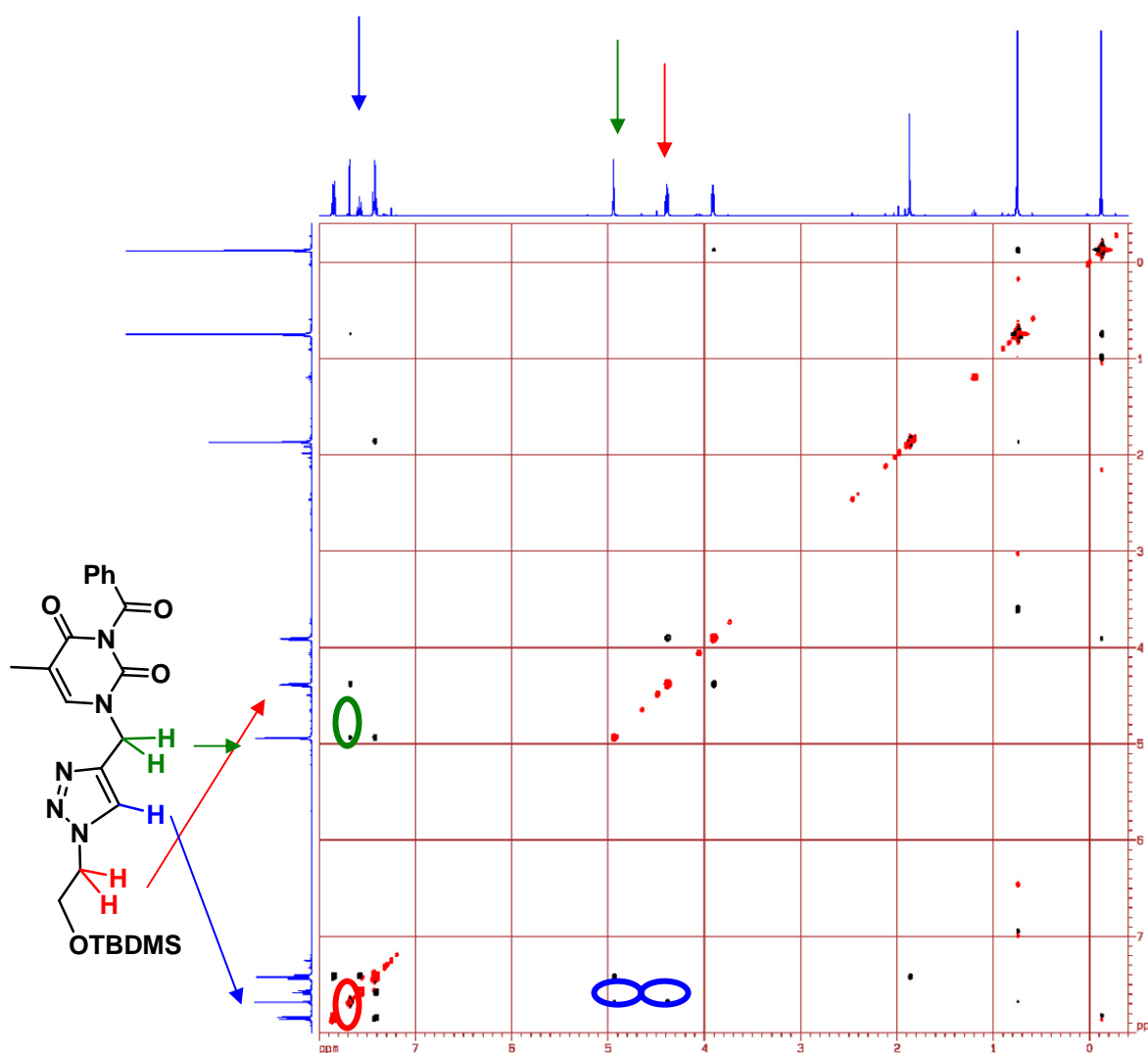
**Scheme 63.** Model reaction using the actual thymine alkyne component **161**. Reagents and conditions: **152**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 24 h.

Thus, the actual thymine alkyne component **161** was subjected to exactly the same conditions as described previously. The reaction mixture was stirred at rt for 24 h, worked-up and purified by flash chromatography to afford **240** in a substantially reduced yield of 24%. From the  $^1\text{H}$  NMR spectra recorded of **240**, the product was identified by a triazole signal at 7.52 ppm and a double doublet at 4.78 ppm corresponding to the  $\text{CH}_2$  group between the triazole and thymine moieties. All other protons were consistent with product **240**. Despite being confident that the ‘click’ conditions employed would favour the formation of the 1,4-regioisomer **240** over the 1,5-regioisomer **240a** [Figure 72], it was decided to carry out NOESY\* spectroscopy to confirm this.



**Figure 72.** 1,4- and 1,5-regioisomers of **240**.

For 1,4-regioisomer formation, two cross-correlation peaks would be expected resulting from the interaction between the CH<sub>2</sub> groups either side of the triazole moiety, with the triazole proton. If the 1,5-regioisomer had formed then only one cross-correlation peak should be observed for the interaction of the triazole proton with the CH<sub>2</sub> group of the nucleobase portion of the molecule. The NOESY\* spectrum recorded for **240** is shown below in Figure 73.

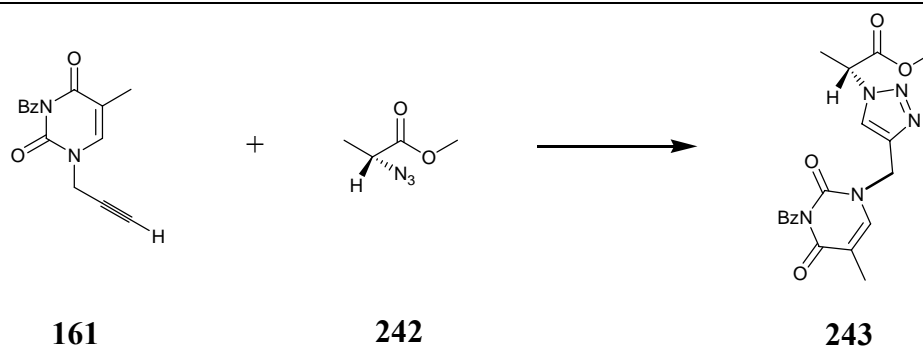


**Figure 73.** NOESY\* spectra of product **240**.

Having been able to form the triazole product **240**, it was then decided to employ a secondary azide, and investigate the viability of a more sterically hindered azide **242** in the ‘click’ reaction with our thymine alkyne component **161** [Scheme 64].

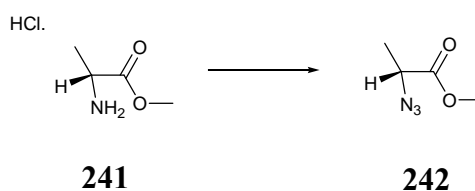
\*NOESY spectrum obtained from Dr Alan Boyd, Heriot-Watt University.





**Scheme 64.** Model reaction using the actual thymine alkyne component **161**. Reagents and conditions: **152**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 24 h.

This model azide first had to be prepared from commercially available L-alanine methyl ester **241** [Scheme 65].



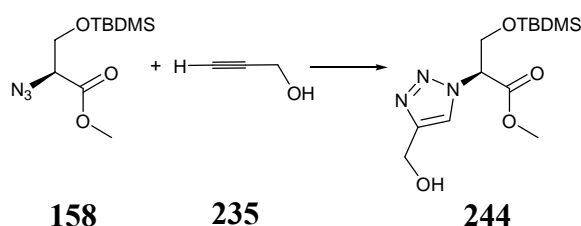
**Scheme 65.** Preparation of the model azide **242**. Reagents and conditions: **152**,  $\text{NaHCO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{H}_2\text{O}$ ,  $\text{MeOH}$ , rt, 24h.

Following the procedure by Goddard-Borger *et al.*<sup>96</sup> azide **242** was prepared. The strong azide signal at  $2106\text{ cm}^{-1}$  in the crude I.R. spectrum recorded of **241** proved that the diazotransfer reaction had been successful. Further, the crude  $^1\text{H}$  NMR also confirmed the correct structure of **242**. The methyl group was observed as a singlet at 1.69 ppm, the methoxy  $\text{CH}_3$  protons were present as a singlet at 3.98 ppm and the remaining CH proton was observed as a singlet at 4.07 ppm.

Azide **242** was then subjected to the same ‘click’ conditions reported previously and following work-up, product **243** was afforded in a 13% yield. The formation of the triazole product was confirmed by  $^1\text{H}$  NMR and NOESY spectroscopy. The triazole proton was observed as a singlet at 7.61 ppm, and a double doublet at 4.79 ppm confirmed the presence of the  $\text{CH}_2$  group between the triazole and thymine moieties. It was postulated that the poor yield observed for this ‘click’ reaction, and indeed the prior example discussed on page 136, may be due to solubility problems of our nucleobase-

containing alkyne component **161**. Despite the yields of these reactions being disappointingly low, it was decided in the interest of time to continue to investigate the viability of our actual azides **158** and **159** with the substitute alkyne component **235**.

Thus, the final model ‘click’ reaction undertaken employed use of our actual azide **158** (see the previous section 2.2.2, page 97) and the alkyne substitute, propargyl alcohol **235** [Scheme 66].

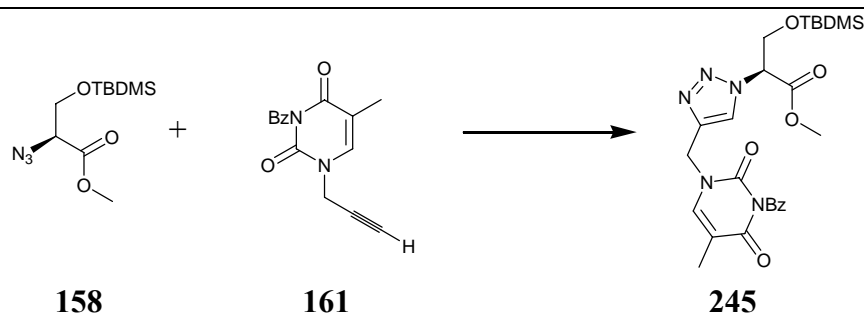


**Scheme 66.** Model ‘click’ reaction with the actual azide **158**. Reagents and conditions:  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 24 h.

Thus, the thymine alkyne substitute **235** was added to a mixture of **158**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , and sodium ascorbate (1:2) in an equal mixture of DCM, *tert*-butyl alcohol and  $\text{H}_2\text{O}$ . Following work-up and purification by flash chromatography, the triazole product **244** was afforded in a 62% yield. From the  $^1\text{H}$  NMR spectrum recorded of **244**, the OH group was seen as a broad singlet at 3.18 ppm, the  $\text{CH}_2$  group next to the OH group was present as a singlet at 4.80 ppm and the triazole signal was observed at 7.92 ppm. The silicon portion of the molecule was seen to be intact and the rest of the signals were consistent with product **244**.

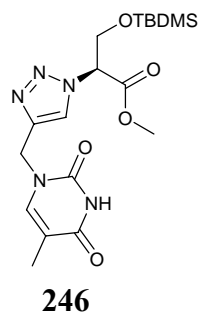
#### 2.4.4 ‘Click’ studies using serine-derived azide **158** and thymine alkyne **161**.

We decided that we were now in a position to begin to conduct preliminary investigations into the ‘click’ reaction using our actual monomers, i.e. the serine-derived azide **158** and thymine alkyne component **161** [Scheme 67].



**Scheme 67.** Actual ‘click’ reaction. Reagents and conditions:  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 24 h.

We chose to first examine the use of the same reaction conditions as we had previously used for our model ‘click’ reactions. Thus, the thymine alkyne component **161** was added to a mixture of azide **152**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  and sodium ascorbate (1:2) in an equal mixture of DCM, *tert*-butyl alcohol and  $\text{H}_2\text{O}$ . The reaction mixture was stirred at rt for 24 hours, worked-up and purified by flash chromatography. From the  $^1\text{H}$  NMR spectra recorded of the product containing fractions, a ‘click’ product was identified. However, this lacked the benzoyl nucleobase protecting group. The debenzoylated ‘click’ product **246** was formed in the extremely low yield of 15% [Figure 74].



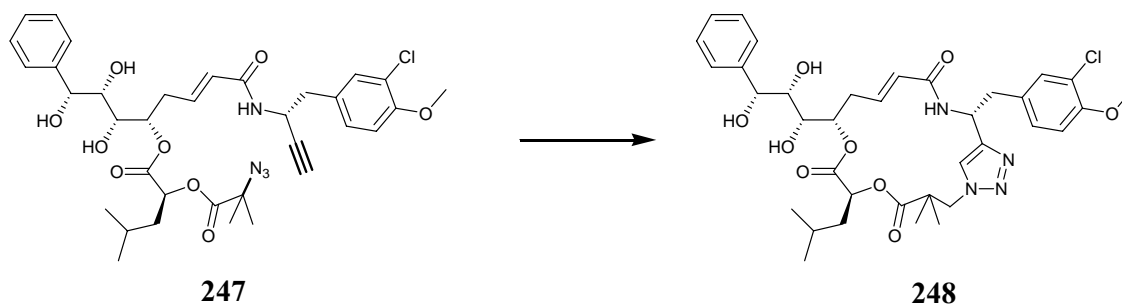
**Figure 74.** Debzoylated ‘click’ product **246**.

None of the product containing fractions contained a click product with the benzoyl group attached (**245**). At this point, we were not concerned that a debenzoylation side reaction had taken place, since removal of this protecting group had been planned for the next stage in our synthetic pathway for the synthesis of the thymine monomer. However, we felt that the yield of this ‘click’ product was currently unacceptable, and so we embarked upon an optimisation strategy in an attempt to improve the yield of the desired ‘click’ reaction. We envisaged that this could be achieved in two different ways; either through optimisation of the ‘click’ reaction itself or through the elimination of the competing debenzoylation side reaction.

Our initial attempt to optimise the ‘click’ reaction, centred around using the same conditions reported previously with the exception of employing a longer reaction time. Thus, after 30 h, following work-up and purification by flash chromatography, the debenzoylated ‘click’ product **246** was afforded in the improved yield of 44%.

Whilst looking for further ways to optimise this reaction, we came across a review by Hein *et al.*<sup>111</sup> in which it was suggested that the sodium ascorbate:Cu<sup>II</sup> ratio, could be increased to a 10-fold excess of the reducing agent, thereby increasing the levels of Cu<sup>I</sup> in the reaction mixture. It was decided that this was worthy of investigation and, therefore, this idea was explored by increasing our ratio from 2:1 sodium ascorbate:Cu<sup>II</sup> to 10:1. Unfortunately, the yield of the debenzoylated ‘click’ product **246** achieved for this reaction was 37%.

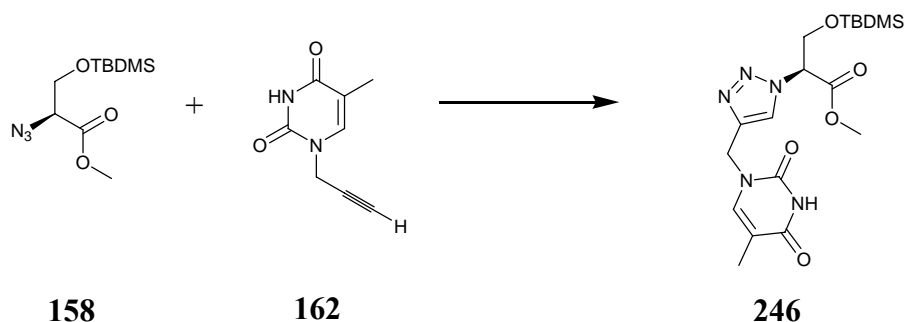
Our final attempts to further improve the yield of the ‘click’ reaction involved replacing the CuSO<sub>4</sub>·5H<sub>2</sub>O catalyst for other copper (I) sources. We were inspired by the procedure reported by Nahrwold *et al.*<sup>115</sup> in which they had successfully conducted the intramolecular 1,3-dipolar cycloaddition using **247** and a copper (I) catalyst, CuI. The desired triazole **248** had been afforded in a 32% yield [Scheme 68]. Therefore, we decided to apply these conditions to our ‘click’ reaction.



**Scheme 68.** Cu<sup>I</sup> ‘click’ reaction. Reagents and conditions: CuI, DIPEA, toluene, rt, 20 h.

Thus, the alkyne component **161** was added to a stirred suspension of azide **158**, CuI and DIPEA in dry DMF for 24 hours at rt. It was decided to substitute toluene, which had been utilised in the literature procedure, for DMF in order to enhance the solubility of alkyne **161**. Following work-up and purification by flash chromatography, the triazole product **245** was afforded in a 15% yield. Here, none of the debenzoylated ‘click’ product **246** was observed in any of the product containing fractions.

In attempt to eliminate the competing debenzoylation of the thymine moiety side reaction from occurring, it was then decided to explore this ‘click’ reaction using the deprotected thymine alkyne component **162** (see the previous section 2.3.1, page 103). Thus, **162** was subjected to the same reaction conditions as described previously [Scheme 69].



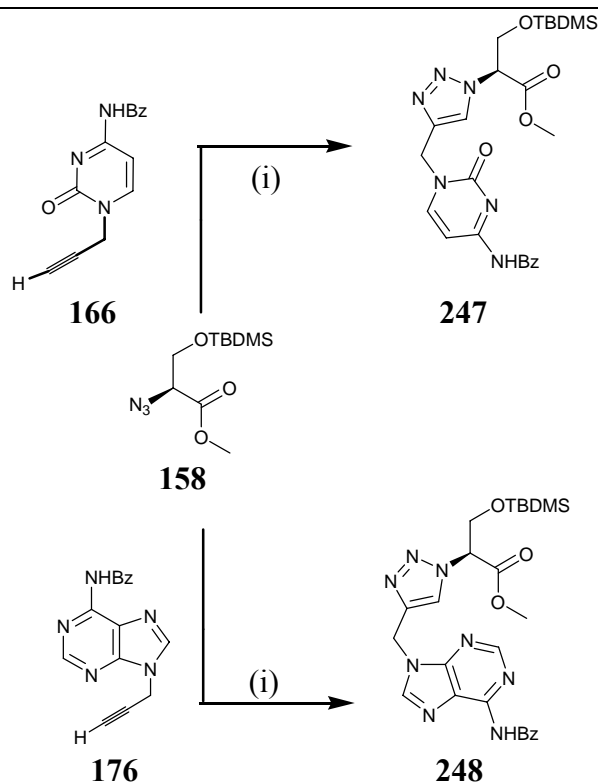
**Scheme 69.** Debenzoylated thymine ‘click’ reaction. Reagents and conditions:  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 30 h.

Following work-up, the crude  $^1\text{H}$  NMR spectrum recorded on the residue afforded indicated that the ‘click’ product **246** was present as a singlet appeared at 7.86 ppm corresponding to the triazole proton. However, this spectrum showed that unreacted starting materials were the main components with **246** only being produced in an estimated 8% yield.

Having investigated several methods for optimising the ‘click’ reaction and elimination of the debenzoylation side reaction, the best ‘click’ reaction conditions encountered for our particular azide and alkyne components, involved the *in situ* reduction of  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  using sodium ascorbate (in a 1:2 ratio) in a 1:1:1 mixture of DCM, *tert*-butyl alcohol and water at rt for 30 h. In the interest of time, it was decided to pursue the ‘click’ reactions between our azide **158** and both the cytosine and adenine alkyne components **166** and **176**.

#### 2.4.5 Cytosine and adenine alkyne component ‘click’ reactions

Despite being aware that the ‘click’ reactions with the cytosine and adenine alkyne components **166** and **176**, respectively, may require subtle differences in terms of the ‘click’ reaction conditions used with the thymine alkyne component **161** (see the previous section 2.4.4, page 140), it was decided to carry out a preliminary investigation into these two reactions [Scheme 70].



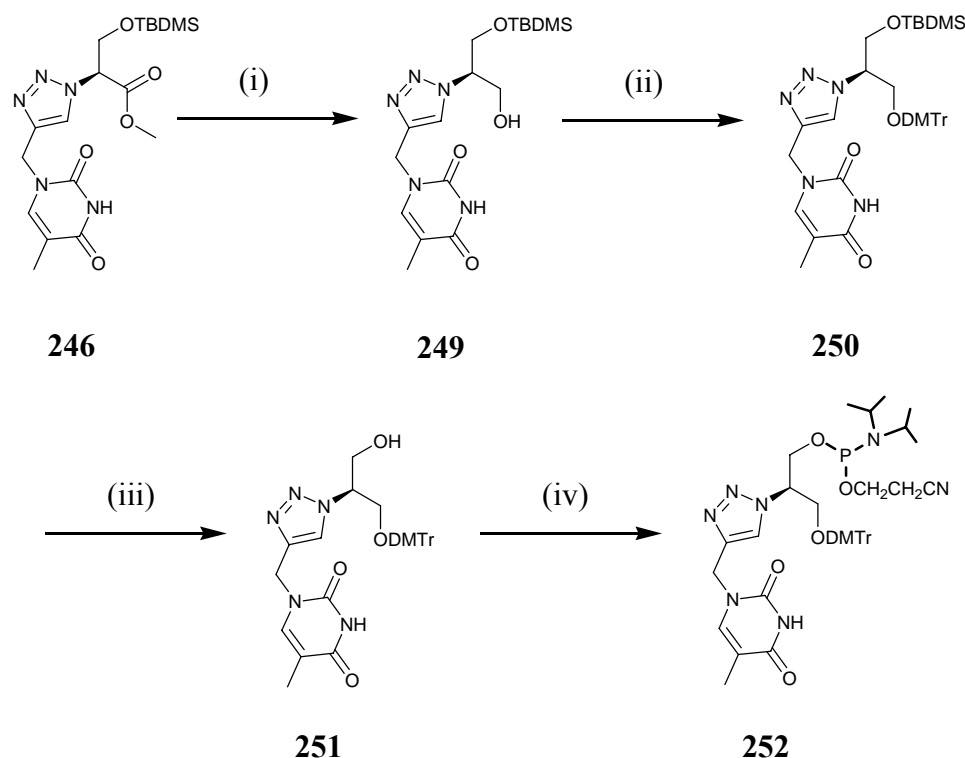
**Scheme 70.** Adenine and cytosine ‘click’ reactions. Reagents and conditions: (i)  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 30 h.

Thus, alkynes **166** and **176** were added to separate mixtures of azide **158**,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate (1:2) in an equal mixture of DCM, *tert*-butyl alcohol and  $\text{H}_2\text{O}$ . The reaction mixtures were stirred at rt and subsequently worked-up. Unfortunately, from the crude  $^1\text{H}$  NMR spectra recorded of **247** and **248**, no triazole signals were observed in the characteristic region. The alkyne starting materials **166** and **176** were recovered quantitatively following flash chromatography.

Despite being confident that the cytosine and adenine ‘click’ reactions could be achieved, due to time constraints further investigation into the appropriate conditions could not be carried out at this stage. Instead, it was decided to pursue the post ‘click’ reactions to afford the thymine TCNA phosphoramidite monomer **252**.

#### 2.4.6 Synthesis of the thymine TCNA phosphoramidite monomer

Having successfully obtained the thymine ‘click’ products **245** and **246** (see the previous section 2.4.4, page 140), the final steps to produce the thymine-containing TCNA phosphoramidite monomer **252** were designed and are shown in Scheme 71.



**Scheme 71.** Post 'click' reaction pathway. Reagents and conditions: (i)  $\text{LiBH}_4$ , *B*-methoxy 9-*BBN*, THF, 0 °C-rt, 24 h; (ii) DMTrCl, DBU, DCM, rt, 24 h.

The first step in this post 'click' pathway involves reducing the ester of **246** to the primary alcohol **249** followed by re-protection with a DMTr substituent to afford **250**. Removal of the silicon group is then achieved to give **251** which is subsequently phosphorylated to afford **252**, ready for the incorporation into oligomers.

In order to perform the first step, we selected to employ the same conditions as we had used previously for the reduction of **122a** to **123a** (page 86). Thus, a solution of **246** in dry THF was added to a solution of  $\text{LiBH}_4$  in dry THF at 0 °C. *B*-Methoxy 9-*BBN* was then added to the mixture and the reaction temperature was allowed to warm to rt. After 24 h, the reaction mixture was worked-up and purified by flash chromatography. The desired alcohol was obtained in a 25% yield. The  $^1\text{H}$  NMR spectrum recorded for **249** showed the newly generated primary alcohol as a broad singlet at 1.60 ppm. The remainder of the peaks were consistent with that expected for **249**.

The next step in this pathway, involved protection of the alcohol of **249** with a DMTr group. Thus, alcohol **249** was added to a suspension of DMTrCl in dry DCM in the presence of DBU. After 24 h stirring at rt, the reaction mixture was worked-up. Disappointingly, the  $^1\text{H}$  NMR spectrum recorded for the crude residue recovered

indicated that this protection reaction had not occurred. This was further confirmed by the fact that the alcohol starting material **249** was recovered quantitatively following flash chromatography. Due to time constraints, further investigation into this step was not carried out, however, it was postulated that perhaps steric hindrance was preventing the reaction from occurring and that a different, smaller protecting group could be employed instead, in the future.

#### **2.4.7 Conclusions**

In an attempt to prevent loss of our precious azide and alkyne components during the preliminary investigation of our ‘click’ reaction, it was decided to carry out some model reactions.

The first of these model reactions involved investigating the ‘click’ reaction using commercially available propargyl alcohol **235** and a silicon protected ethanolamine-derived azide **238**. **235** was chosen as not only was it easy to obtain, but it would also allow us to determine the outcome of the ‘click’ reaction easily, by simplifying the <sup>1</sup>H NMR region expected for the triazole proton. This azide was chosen due to it being a primary azide and also because it allowed us to investigate the stability of the silicon group to the ‘click’ conditions. We had been inspired by conditions reported by Goujon *et al.*<sup>114</sup> in our research group, and successfully went on to prepare the model ‘click’ product **239** in a 70% yield.

Using this same azide **238**, we then attempted the ‘click’ reaction using the same conditions as before, except our actual thymine alkyne component **161** was employed. In this case, the desired ‘click’ product was only afforded in a 24% yield. This yield was unacceptably low and was attributed to poor solubility of the nucleobase-containing alkyne component **161**. Due to this reaction only being a model, it was decided to continue and investigate the ‘click’ reaction between the actual alkyne component **161** and another model azide. This time, a secondary azide **242** was employed to replicate our actual azide, and further, to allow us to evaluate the stability of the methyl ester portion of our actual azide **158** during the ‘click’ reaction. Again, although the desired ‘click’ product **243** was obtained, the yield of 13% was extremely low.



---

The last model reaction investigated involved using our actual azide **158** with propargyl alcohol. This successfully produced the triazole product **244** in a 62% yield.

Despite our preliminary ‘click’ reactions affording the corresponding triazole product in poor yields when the thymine-containing alkyne component **161** was employed. We were confident that now we had established reproducible ‘click’ reaction conditions, and knew how to verify if the desired reaction had taken place, that our precious starting materials would not be wasted. ‘Click’ studies using the serine-derived azide **158** and the thymine-containing alkyne component **161** could now be investigated.

Thus, using the same conditions described previously in our model ‘click’ reactions, no desired triazole-containing product **245** was formed between our actual azide **158** and our actual thymine-containing alkyne component **161**. Instead, another ‘click’ product, in which the benzoyl group had been cleaved (**246**) was afforded in a 15% yield. Although this result was surprising, the fact that the benzoyl group had been removed wasn’t deemed a problem, as its removal had been intended at a later stage, prior to the monomers incorporation into oligomers. However, the poor yield was unacceptably low and, therefore, an optimisation strategy was embarked upon.

This investigation centred around two different ways of achieving a greater percentage yield of the ‘click’ reaction. Firstly, through optimisation of the ‘click’ reaction itself, and secondly, through elimination of the competing debenzoylation side reaction. To optimise the ‘click’ reaction, our first endeavour involved increasing the reaction time. The best yield obtained from this investigation was 44% after 30 h.

Next, we explored the role that the ratio of sodium ascorbate to copper catalyst could play in the ‘click’ reaction. Thus, this ratio was increased from 2:1 to 10:1 and the outcome was compared. The observed yield from this attempt was 37% however, no further investigation into use of different ratios was carried out at this time.

Our final attempt to optimise the ‘click’ reaction involved using a different source of Cu<sup>I</sup>. Instead of reducing Cu<sup>II</sup> to Cu<sup>I</sup> *in situ* using sodium ascorbate, as in the previous cases, it was decided to explore the direct addition of Cu<sup>I</sup> to the reaction mixture. We were inspired by a procedure reported by Nahrwold *et al.*<sup>115</sup> to use CuI. The desired

'click' product **245** was obtained in a 15% yield, but surprisingly, the benzoyl group was intact.

From our 'click' optimisation investigation, the highest yield possible of 44%, albeit of the debenzoylated species **246**, was obtained after 30 h using a 2:1 ratio of sodium ascorbate to copper (II) catalyst.

In an attempt to eradicate the competing debenzoylation side reaction from occurring, the previously prepared thymine alkyne component **161**, in which the benzoyl group had already been removed, was subjected to these optimised 'click' reaction conditions. The main constituent of the reaction mixture observed from the crude <sup>1</sup>H NMR spectrum recorded was the starting materials and the desired triazole-containing 'click' product **245** was only present in an estimated 8% yield.

Preliminary investigations into the viability of the other nucleobase-containing alkyne components in the 'click' reaction was then carried out. Azide **158** was subjected to the same optimised 'click' reactions discussed previously, with both the cytosine and adenine-containing alkyne components **166** and **176**. Unfortunately, these initial reactions failed and the starting materials were recovered quantitatively in each case. Due to these results, the 'click' reaction with the guanine alkyne component was not attempted. Despite these findings, we are confident that the cytosine, adenine and guanine-containing alkyne components will be able to be incorporated into our triazole-containing monomers in the future. However, due to time constraints, it was not possible for us to investigate this here. It was deemed a more appropriate use of the remaining time, to investigate the post 'click' reactions and develop a route towards our thymine-containing TCNA phosphoramidite monomer **252**.

Having obtained our thymine triazole-containing monomer **161**, the next step in the synthetic pathway to afford our TCNA phosphoramidite monomer **252** was to reduce the methyl ester to the primary alcohol **249**. This was achieved in a 25% yield using conditions previously encountered in the development of a synthetic route towards the azide component in route B (page 82). Although this yield was poor, no optimisation attempts were carried out at this time and instead, it was decided to pursue the next step.

---

This involved protection of the newly generated primary alcohol with a DMTr group, which is a prerequisite for the monomers subsequent incorporation into oligomers. Conditions previously used in the development of a synthetic route towards the azide component in route A (page 73) were employed. Regrettably, this reaction failed. Again, no optimisation attempts were made at this time, however, it was postulated that steric hindrance between the 'click' alcohol product and the DMTr constituent itself, was responsible for this reactions failure. In the future, a smaller protecting group could be employed to clarify if this hypothesis is correct, the results of which can then be applied to find a suitable alternative or to have the confidence to pursue alternative DMTr conditions.

## **Chapter 3**

### **Conclusions and future work**

### 3.1 CONCLUSIONS

The work described in this research project has provided a fundamental basis from which our TCNA monomers could, with further development in the future, be prepared for incorporation into oligomers for investigations of their potential utility in the field of antigene therapy. As outlined in chapter 1 (see the previous section 1.10, pages 67-70), the main aims were: (1) to develop a suitable synthetic pathway for the synthesis of the azide component; (2) to develop an appropriate synthetic route for the preparation of the nucleobase-containing alkyne component; (3) to investigate the use of the ‘click reaction’ for coupling our azide and alkyne components, and; (4) to develop a viable route for the construction of the TCNA phosphoramidite monomers. The target compounds are shown in Figure 75 below.

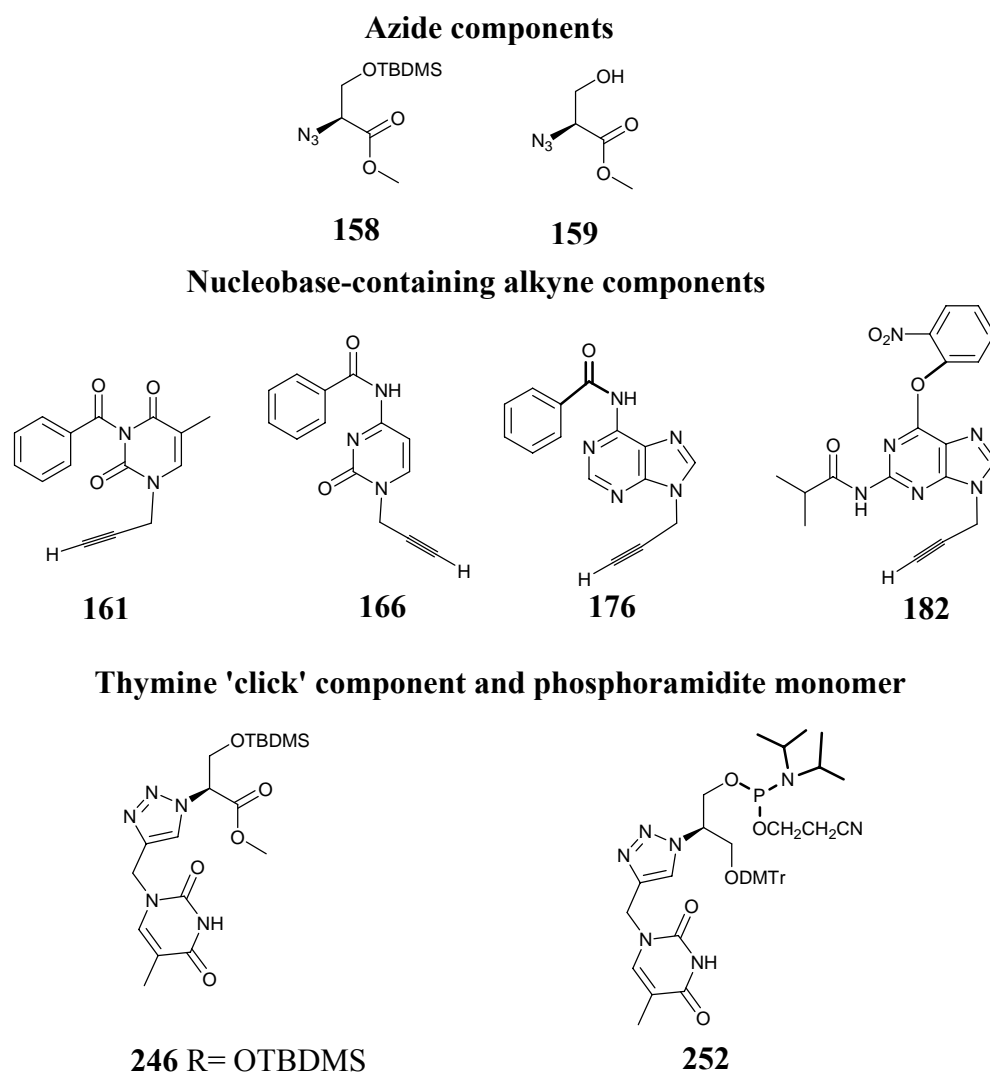
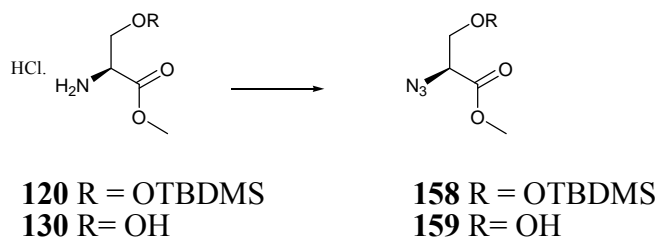


Figure 75. Target compounds.

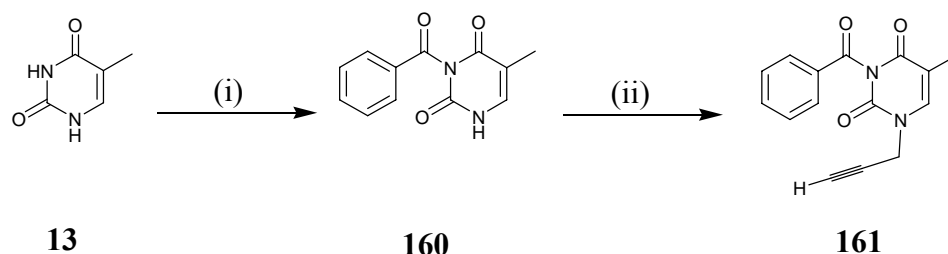
After investigating several strategies to develop a viable, synthetic scheme towards the azide component, this was finally achieved by direct diazotransfer of both the silicon protected **158** and natural L-serine methyl ester **159**. The novel ‘diazo donor’ imidazole-1-sulfonyl azide **152**, as inspired by Goddard-Borger *et al.*<sup>96</sup>, was used in this diazotransfer reaction [Scheme 72]. The presence of azides **158** and **159** were confirmed by strong signals at 2117  $\text{cm}^{-1}$  and 2110  $\text{cm}^{-1}$  respectively, in the I.R spectra recorded (see the previous section 2.2.2, page 97).



**Scheme 72.** Final synthetic route to the azide components **158** and **159**. Reagents and conditions: **130** or **120**,  $\text{K}_2\text{CO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , MeOH, rt, 16h.

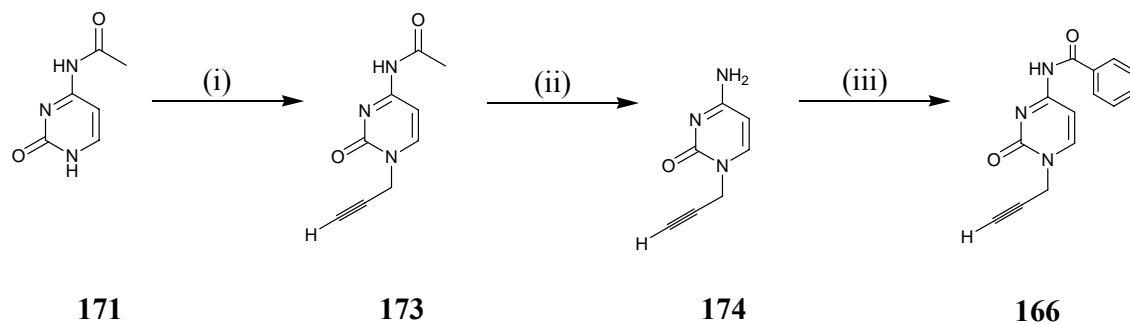
In parallel to the development of a viable synthetic route to afford our azide components, we were also developing the synthesis to afford us with each of the nucleobase-containing alkyne components for use in the ‘click’ reaction. Despite the structural similarity of the alkyne components, each nucleobase required individual investigation into the appropriate conditions to afford such a component.

The thymine-containing alkyne component **161** was prepared by protecting the *N3* amine function with BzCl to give **160** in a 73% yield (see the previous section 2.3.1, page 101). **160** was subsequently alkylated with propargyl bromide in the *N1* position to give **161** in a 55% yield [Scheme 73].



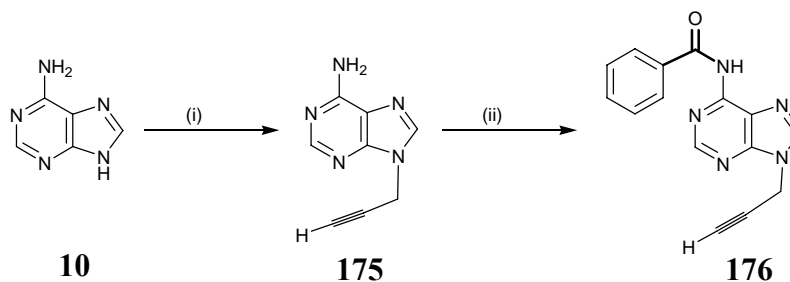
**Scheme 73.** Final synthetic route to the thymine-containing alkyne component **161**. Reagents and conditions: (i) BzCl,  $\text{CH}_3\text{CN}/\text{C}_5\text{H}_5\text{N}$  (1:1), rt, 24 h; (ii)  $\text{HC}\equiv\text{CCH}_2\text{Br}$ ,  $\text{Cs}_2\text{CO}_3$ , DMF, rt, 24 h.

The cytosine-containing alkyne component **166** was synthesised from commercially available *N*4-acetylcytosine **171** as inspired by Lindsell *et al.*<sup>105</sup> (see the previous section 2.3.2, page 104). This reagent was alkylated with propargyl bromide in the presence of  $K_2CO_3$  to afford **173** in a 40% yield. Removal of the acetyl protecting group was then carried out by treating **173** with a methanolic ammonia solution to afford **174** in a 94% yield. Re-protection of the exocyclic  $NH_2$  group was then achieved with a benzoyl group to afford the target cytosine monomer **166** in a 67% yield [Scheme 74].



**Scheme 74.** Final synthetic route to the cytosine-containing alkyne component **166**. Reagents and conditions: (i)  $HC\equiv CCH_2Br$ , DMF,  $K_2CO_3$ , rt, 24 h; (ii) 5% (v/v) methanolic ammonia solution, 24 h, rt; (iii)  $BzCl$ ,  $C_5H_5N$ , rt, 24 h.

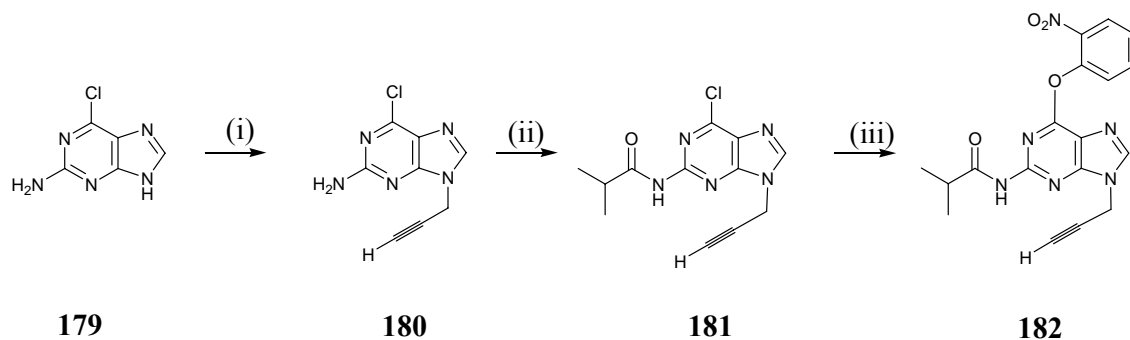
The adenine-containing alkyne component **176** was achieved by first alkylating with propargyl bromide in the *N*9 position to give **175** in a 36% yield (see section 2.3.3, page 109). Protection of the exocyclic amine function was then achieved by treating **175** with  $BzCl$  in pyridine to afford the target adenine alkyne product **176** in a 67% yield [Scheme 75].



**Scheme 75.** Final synthetic route to the adenine-containing alkyne component **176**. Reagents and conditions: (i)  $HC\equiv CCH_2Br$ , DMF,  $NaH$ , rt, 24 h; (ii)  $BzCl$ ,  $C_5H_5N$ , rt, 4 h.

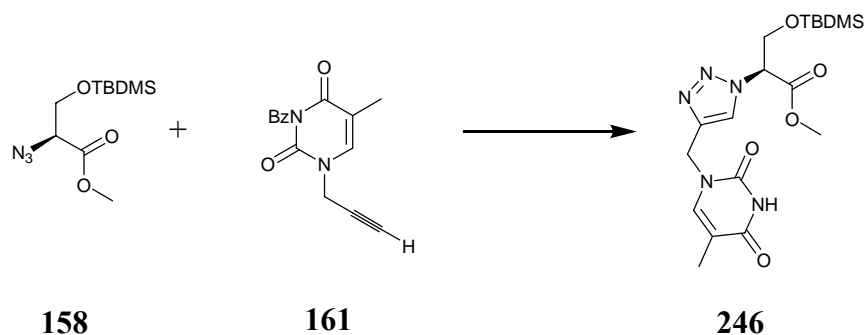
The protected guanine-containing alkyne component **182** was successfully prepared starting from commercially available 2-amino-6-chloropurine **179** (see the previous

section 2.3.4, page 114). Having been inspired by Lindsell *et al.*<sup>105</sup> alkylation in the N9 position was achieved using propargyl bromide to afford **180** in a 77% yield. The exocyclic amine function was then protected using an isobutyryl group to afford **181** in an 88% yield. Substitution of the 6-chloro group for a 2-nitrophenol group was then successfully carried out in an 89% yield to give **182** [Scheme 76]. This product was now suitable to be used in the ‘click’ reaction with our azide component.



**Scheme 76.** Final synthetic route to the protected guanine-containing alkyne component **182**. Reagents and conditions: (i)  $\text{HC}\equiv\text{CCH}_2\text{Br}$ ,  $\text{K}_2\text{CO}_3$ , DMF, rt, 24 h; (ii)  $^i\text{PrCOCl}$ , DMAP,  $\text{C}_5\text{H}_5\text{N}$ , rt, 4 h; (iii) 2-nitrophenol, DABCO, TEA, DCM, rt, 24 h.

Having prepared the azide components and the nucleobase-containing alkyne components, we were now in a position to evaluate the ‘click’ reaction. After investigating the ‘click’ conditions using model reactions (see the previous section 2.4.3, pages 134-140), the azide components **158** and **159**, together with the thymine-containing alkyne component **161**, were subjected to the same ‘click’ conditions. It was found, during our brief attempt to optimise this reaction, that azide **158** and thymine-containing alkyne component **161** in a 1:1:1 mixture of DCM, *tert*-butylalcohol and  $\text{H}_2\text{O}$  in the presence of a 1:2 ratio of  $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$  to sodium ascorbate, at room temperature for 30 h afforded **246** in a 44% yield [Scheme 77].



**Scheme 77.** Final conditions for the actual ‘click’ reaction to prepare **246**. Reagents and conditions:  $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$ , rt, 30 h.



A preliminary investigation into the development of the thymine-containing TCNA phosphoramidite monomer **252** was carried out (see the previous section 2.4.6, pages 144-146). The thymine ‘click’ product was reduced to the corresponding primary alcohol **249** in a 25% yield. Due to time constraints, further optimisation of this step and indeed investigation into the remaining three steps in this synthetic scheme, towards the final target phosphoramidite monomer was not achieved.

### **3.2 FUTURE WORK**

If time had permitted, a more substantial investigation into the optimisation of the ‘click’ reaction would have been undertaken, primarily for the improved preparation of the thymine monomer, but also for the remaining nucleobase-containing monomers.

To date, there has been no comprehensive study into the ideal conditions for the Cu<sup>I</sup>-catalysed azide-alkyne coupling in the literature or indeed in this research, however, the results outlined in both suggest that Cu<sup>I</sup>-catalysed azide-alkyne couplings afford triazoles under a variety of conditions, underscoring the robustness of the reaction.<sup>112</sup> Typically, ‘click’ reactions take 6-36 hours to occur at ambient temperature, in a variety of different solvent mixtures, the catalyst is efficient in a pH range of 4-12 and the percentage yields are generally greater than 70%.<sup>112</sup>

The problems encountered by other researchers, with the Cu<sup>I</sup>-catalysed Huisgen 1,3-dipolar cycloaddition variant, have been outlined previously (section 2.4.1 on page 129) and are mainly due to; (i) the azide being too electron deficient; (ii) alkyne homocoupling occurring; (iii) Cu<sup>I</sup> saturation or fluctuation of appropriate levels of Cu<sup>I</sup> in the reaction mixture and; (iiii) poor azide stability.<sup>112</sup> It was postulated, in our case, that the poor yields and need for the ‘click’ products to be purified was due to the combination of poor nucleobase-containing alkyne component solubility in the solvent mixtures used and fluctuations in the levels of Cu<sup>I</sup> catalyst throughout the reaction.

With this in mind, we would have carried out a more thorough investigation into other ‘click’ conditions that could potentially, have optimised our particular reaction. This investigation would have focused on; (i) continued alteration of the copper (II) catalyst to sodium ascorbate ratio. Development of a methodical examination into different Cu<sup>II</sup> catalysts:sodium ascorbate ratios in the range (1:2) to (1:10), in order to eliminate the

---

fluctuating or insufficient starting levels of Cu<sup>I</sup> in the reaction mixture; (ii) the usefulness of oxidating copper metal to Cu<sup>I</sup> or use of Cu<sup>0</sup> nanoclusters; (iii) direct addition of a variety of different Cu<sup>I</sup> salts (e.g. CuBr, CuCl), using a wide selection of nitrogen-containing bases; (iiii) microwave assisted ‘click’ reactions; (v) the use of a Cu<sup>I</sup>-stabilising agent (e.g. THPTA) to eliminate Cu<sup>I</sup> saturation and; (vi) the use of different solvent mixtures (perhaps employing DMF) to increase the solubility of our nucleobase-containing alkyne components.<sup>112</sup>

Only when the above analysis has been carried out, with our particular azide and nucleobase-containing alkyne components, will comparisons with the literature be able to be made. Further, such an investigation would be vital for progressing the work described in this thesis, as the poor yields and need for purification encountered here would hinder advances towards achieving a triazole-containing nucleic acid oligomer for future use in antigene therapy.

**Chapter 4**  
**Experimental**

---

## 4. EXPERIMENTAL

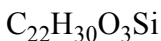
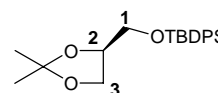
### 4.1 EXPERIMENTAL INTRODUCTION

Chemicals were purchased from Aldrich, Acros and Alfa Aesar chemical companies and were used as supplied without further purification. Standard procedures were used to dry organic solvents where stated. Reactions were routinely carried out under an inert atmosphere of argon or N<sub>2</sub>. ‘Light petroleum’ refers to the fraction boiling between 40 °C and 60 °C. Analytical thin layer chromatography was performed on aluminium backed plates pre-coated with Merck Kieselgel 60 GF<sub>254</sub> (Art. 05554). Visualization was achieved by ultra-violet light (254 nm) and/or alkaline potassium permanganate stain followed by heating. Flash chromatography was performed using DAVISIL® silica (60 Å; 35-70 µm) from Fisher (cat. S/0693/60). Fully characterised compounds were chromatographically homogeneous (conventional columns).

<sup>1</sup>H NMR spectra were recorded at 200 and 400 MHz on Bruker AC200 and DPX400 spectrometers; <sup>13</sup>C NMR spectra were recorded at 50 and 100 MHz on the same instruments. Chemical shifts are reported in parts per million ( $\delta$  in ppm) and are referenced against solvent signals ( $\delta_C$  77.16 for chloroform and  $\delta_C$  39.52 for dimethyl sulfoxide for <sup>13</sup>C spectra;  $\delta_H$  7.26 for chloroform and  $\delta_H$  2.50 methyl sulfoxide for <sup>1</sup>H spectra). Chemical shift values are accurate to  $\pm 0.01$  ppm and  $\pm 0.1$  ppm for <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra, respectively. *J* values are given in Hz. Multiplicity designations used are: s, d, t, q, quint and m for singlet, doublet, triplet, quartet, quintet and multiplet, respectively. In <sup>13</sup>C NMR spectra, signals corresponding to C, CH, CH<sub>2</sub>, or CH<sub>3</sub> groups are assigned from DEPT. High-resolution mass spectra were obtained from the National Mass Spectrometry Service Centre, Swansea. Melting points were recorded on a Stuart Scientific SMP10 and are uncorrected. Infrared spectra were recorded on a Perkin Elmer 1600 FT IR spectrometer. These spectra were recorded as either potassium bromide discs, as solutions in CHCl<sub>3</sub>, or as films between sodium chloride plates. Bands are defined as: w, br, m, s, broad, weak, medium, strong, respectively. Microanalytical data was recorded on an Exeter CE-440 Elemental Analyser, and  $[\alpha_D]$  analysis was recorded on a Thorn Bendix NPL Automatic Polarimeter type 143D with a yellow sodium filter, through the analytical services housed in Chemistry, at Heriot-Watt University school of Engineering and Physical Sciences.

### Preparation of (S)-2,3-O-isopropylidene-1-o-TBDPS-glycerol (**98**)<sup>116</sup>

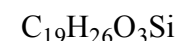
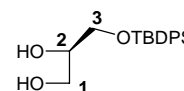
(4S)-2,2-Dimethyl-1,3-dioxolane-4-methanol (2.50 g, 18.92 mmol, 1.00 eq), DMAP (0.13 g, 1.02 mmol, 0.05 eq) and TEA (3 mL, 21.52 mmol, 1.14 eq) were dissolved in dry DCM (100 mL) and to this, TBDPSCl (4.83 mL, 18.85 mmol, 0.99 eq) was added dropwise with continual stirring. The mixture was stirred at rt for 24 h and washed with saturated NaHCO<sub>3</sub> solution (50 mL). The resulting solution was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and the extract was evaporated to dryness *in vacuo*. The orange residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc (80: 20) as eluant. Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, orange/brown oil. (**98**, 6.56 g, 16.90 mmol, 97%):  $R_f = 0.43$  (light petroleum:EtOAc, 80:20);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3071 m, 3050 m, 2931 s, 2858 s, 1589 w, 1472 m, 1428 m, 1379 m, 1255 m, 1215 m, 1143 m, 1112 m;  $\delta_H$  (200 MHz; CDCl<sub>3</sub>) 0.92 (9 H, s, (CH<sub>3</sub>)<sub>3</sub>C), 1.23 (3 H, s, CH<sub>3</sub>), 1.24 (3 H, s, CH<sub>3</sub>), 3.44-3.65 (2 H, m, C-1H<sub>2</sub>), 3.74-3.81 (1 H, m, C-2H), 3.89-4.11 (2 H, m, C-3H<sub>2</sub>), 7.16-7.29 (6 H, m, Ar-H), 7.50-7.58 (4 H, m, Ar-H);  $\delta_c$  (50 MHz; CDCl<sub>3</sub>) 19.1 (C), 25.3 (CH<sub>3</sub>), 26.4 (3 × CH<sub>3</sub>), 26.6 (CH<sub>3</sub>), 64.4 (CH<sub>2</sub>), 66.6 (CH<sub>2</sub>), 75.9 (CH), 109.0 (C), 127.6 (4 × CH), 129.6 (2 × CH), 133.1 (2 × C), 135.4 (4 × CH);  $m/z$  (ESI) 388 (M+NH<sub>4</sub><sup>+</sup>, 38%), 293 (M-Ph)<sup>+</sup>, 100%), 43 (M-(C(CH<sub>3</sub>)<sub>2</sub>)<sup>+</sup>, 100%); [Found (ESI) 388.2302, C<sub>22</sub>H<sub>34</sub>O<sub>3</sub>SiN requires 388.2304]; (Found: C, 71.35; H, 8.24. C<sub>22</sub>H<sub>30</sub>O<sub>3</sub>Si requires C, 71.31; H, 8.16 (c. 19%);  $[\alpha]_D^{24} + 4.02$  (19.90 mg / 2 mL in DCM).



### Attempted preparation of (S)-3-O-(TBDPS)-glycerol (**99**)

#### Method A

**98** (0.10 g, 0.28 mmol, 1.00 eq) was dissolved in 60% aqueous acetic acid solution (60 mL) and allowed to stir at rt for 4 h. The solvent was removed under reduced pressure to afford a clear, colourless oil. <sup>1</sup>H NMR conducted on this crude residue confirmed that the reaction had failed and that the unreacted starting acetal was the sole compound present.

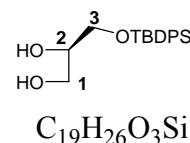


## Method B

The reaction was set up as described in method A except that the reaction was left to stir at rt for 24 h instead of 4 h. Subsequently, the solvent was removed under reduced pressure to afford a clear, colourless oil.  $^1\text{H}$  NMR confirmed that the reaction had failed and the unreacted starting material was the sole compound present.

Preparation of (S)-3-O-(TBDPS)-glycerol (**99**)

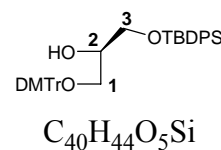
To a stirred solution of **98** (1.80 g, 5.05 mmol, 1.00 eq) in a mixture of THF/H<sub>2</sub>O ((4:1) v/v, 180 mL) at 0 °C was added TFA (0.897 mL, 11.71 mmol, 2.30 eq). The resulting solution was allowed to warm to rt before being allowed to stir for 45 h. The solvent was removed *in vacuo* and a further solution of THF/H<sub>2</sub>O ((4:1) v/v, 180 mL) was added. The resulting mixture was cooled to 0 °C and TFA (0.89 mL, 11.71 mmol, 2.30 eq) was added. The mixture was allowed to warm to rt before being allowed to stir for 24 h. The solvent was removed *in vacuo* and a further quantity of THF/H<sub>2</sub>O ((4:1) v/v, 180 mL) was added. The resulting mixture was cooled to 0 °C and TFA (0.89 mL, 11.71 mmol, 2.30 eq) was added. The mixture was allowed to warm to rt before being left to stir for 24 h. The solvent was removed *in vacuo* and the residue was purified by flash chromatography on a silica column, eluting with (light petroleum:EtOAc:TEA (50:50:1) to light petroleum:EtOAc:TEA (25:75:1) to light petroleum:EtOAc:TEA (0:100:1) gradient elution). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white solid. (**99**, 1.02 g, 2.92 mmol, 64%):  $R_f = 0.23$  (light petroleum:EtOAc, 50:50); m.p. 51-52 °C;  $\nu_{\text{max}}$  (KBr)/cm<sup>-1</sup> 3391 br s (OH), 3071 w, 2929 m, 2857 m, 1657 w, 1463 w, 1428 w, 1391 w;  $\delta_{\text{H}}$  (200 MHz; CDCl<sub>3</sub>) 0.91 (9 H, s, (CH<sub>3</sub>)<sub>3</sub>C), 2.26 (2 H, br s, 2 × OH), 3.36-3.67 (5 H, m, C-1,2 and 3H), 7.18-7.33 (6 H, m, Ar-H), 7.45-7.53 (4 H, m, Ar-H);  $\delta_{\text{C}}$  (50 MHz; CDCl<sub>3</sub>) 19.0 (C), 26.5 (3 × CH<sub>3</sub>), 63.7 (CH<sub>2</sub>), 65.0 (CH<sub>2</sub>), 71.6 (CH), 127.7 (4 × CH), 129.8 (2 × CH), 132.6 (2 × C), 135.3 (4 × CH);  $m/z$  (ESI) 348 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (ESI) 348.1990, C<sub>19</sub>H<sub>30</sub>O<sub>3</sub>SiN requires 348.1989]; (Found: C, 68.74; H, 7.93. C<sub>19</sub>H<sub>30</sub>O<sub>3</sub>SiN requires C, 69.05; H, 7.93%);  $[\alpha]_{\text{D}}^{24} + 6.00$  (20.00 mg / 2 mL in DCM).



---

**Attempted preparation of (S)-1-O-DMTr-3-O-TBDPS-glycerol (100)**
*Method A*

The diol, **99** (0.20 g, 0.63 mmol, 1.00 eq) was dissolved in dry  $C_5H_5N$  (20 mL) and, to this, DMTrCl (0.21 g, 0.63 mmol, 1.00 eq) was added. The reaction mixture was stirred at rt under  $N_2$  for 24 h. TLC confirmed no reaction had occurred. The reaction mixture was heated gradually to a maximum of 70 °C for 12 h, and the solvent was removed *in vacuo* to afford a yellow/orange oil. The crude product was purified by flash chromatography on a silica column eluting with dichloromethane:diethyl ether:TEA (50:50:1). None of the  $^1H$  NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **99**; unreacted starting diol was the sole compound present.

*Method B*

A solution of **99** (0.20 g, 0.63 mmol, 1.00 eq), DMTrCl (0.24 g, 0.69 mmol, 1.10 eq), TEA (0.13 mL, 0.95 mmol, 1.50 eq) and DMAP (0.004 g, 0.03 mmol, 0.04 eq) in dry DMF (20 mL) was stirred at rt under  $N_2$  for 25 h. The yellow solution was poured into ice- $H_2O$  (20 mL) and extracted with DCM ( $4 \times 30$  mL). The combined organic layers were washed with saturated  $NH_4Cl$  solution ( $2 \times 20$  mL) and then  $H_2O$  (20 mL). The separated organic phase was dried over anhydrous  $MgSO_4$ , filtered and concentrated *in vacuo* to afford a yellow oil. The crude product was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (50:50:1). None of the  $^1H$  NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **99**; unreacted starting diol was the sole compound present.

*Method C*

**99** (0.10 g, 0.32 mmol, 1.00 eq) was dissolved in dry DMF (10 mL) and, to this, NaH (60% in mineral oil) (0.02 g, 0.35 mmol, 1.10 eq) was added. The reaction mixture was stirred at rt under  $N_2$  for 20 min. DMTrCl (0.12 g, 0.35 mmol, 1.10 eq) and TEA (0.07 mL, 0.47 mmol, 1.50 eq) were then added and the reaction mixture was maintained at rt for 24 h. A further aliquot of NaH (60% in mineral oil, 0.01 g, 0.14 mmol, 0.42 eq) was added and the reaction mixture was maintained at rt for 24 h. The solvent was removed under reduced pressure and the crude product was purified by flash chromatography on

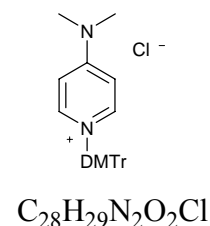
a silica column eluting with light petroleum:EtOAc:TEA (50:50:1). None of the  $^1\text{H}$  NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **99**; unreacted starting diol was the sole compound present.

#### Method D

4-Dimethylamino-*N*-triphenyldimethoxypyridinium chloride, **115** (0.35 g, 0.76 mmol, 1.20 eq) and diol **99** (0.20 g, 0.63 mmol, 1.00 eq) were dissolved in dry DCM (30 mL) before being left to stir at rt under  $\text{N}_2$  for 24 h. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (50:50:1). None of the  $^1\text{H}$  NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **99**; unreacted starting diol was the sole compound present.

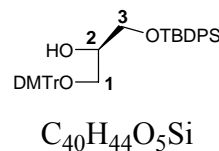
#### Preparation of 4-dimethylamino-*N*-triphenyldimethoxypyridinium chloride (**115**)<sup>87</sup>

DMTrCl (0.40 g, 1.19 mmol, 1.10 eq) and DMAP (0.13 g, 1.09 mmol, 1.00 eq) were mixed in dry DCM (30 mL) before being left to stir at rt under  $\text{N}_2$  for 30 min.  $\text{Et}_2\text{O}$  (100 mL) was added and the solvent was removed *in vacuo* to afford an off-white solid. **115** was used immediately in the subsequent step to attempt the preparation of **100** without further purification.



#### Preparation of (*S*)-1-*O*-DMTr-3-*O*-(TBDPS)propan-2-ol (**100**)

The diol, **99** (0.25 g, 0.73 mmol, 1.00 eq) was added with stirring to a solution of DMTrCl (0.29 g, 0.87 mmol, 1.20 eq) and DBU (0.16 g, 1.02 mmol, 1.40 eq) in dry DCM (15 mL). After 27 h at rt the reaction was quenched by addition of  $\text{H}_2\text{O}$  (20 mL). The organic layer was separated, washed with  $\text{H}_2\text{O}$  ( $2 \times 20$  mL) and dried over anhydrous  $\text{Na}_2\text{SO}_4$ . Filtration, followed by evaporation of the solvent *in vacuo* afforded a yellow oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white solid. (**100**, 0.24 g, 0.37

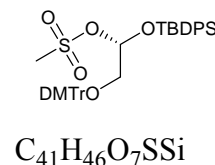




mmol, 49%):  $R_f = 0.45$  (light petroleum:EtOAc, 50:50); m.p. 40-42 °C;  $\nu_{\max}$  (KBr)/ $\text{cm}^{-1}$  3564 br s (OH), 3069 m, 2999 m, 2931 s, 2857 m, 2836 m, 1608 s, 1509 s, 1463 m, 1445 m, 1427 m, 1250 s, 1113 s, 1078 m;  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) 0.97 (9 H, s,  $(\text{CH}_3)_3\text{C}$ ), 2.44 (1 H, br s, OH), 3.12-3.30 (2 H, m, C-3 $H_2$ ), 3.53-3.69 (6 H, m, C-1 $H_2$ , C-2 $H$ ), 3.71 (3 H, s,  $\text{OCH}_3$ ), 6.56-6.83 (5 H, m, Ar- $H$ ), 7.06-7.67 (18 H, m, Ar- $H$ );  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) 19.4 (C), 27.0 (3  $\times$   $\text{CH}_3$ ), 55.4 (2  $\times$   $\text{CH}_3$ ), 64.2 ( $\text{CH}_2$ ), 65.1 ( $\text{CH}_2$ ), 71.4 (CH), 113.3 (4  $\times$  CH), 126.9 (CH), 127.2 (2  $\times$  CH), 127.9 (8  $\times$  CH), 128.3 (2  $\times$  CH), 129.9 (2  $\times$  CH), 130.2 (4  $\times$  CH), 133.3 (C), 135.7 (4  $\times$  CH), 136.2 (C), 145.1 (C), 158.6 (C);  $m/z$  (ESI) 655 ( $\text{M}+\text{Na}^+$ , 100%), 303 ( $\text{M}-\text{DMTr}-\text{O}^+$ , 100%); [Found (ESI) 655.2848,  $\text{C}_{40}\text{H}_{44}\text{O}_5\text{SiNa}$  requires 655.2850]; (Found: C, 76.16; H, 6.71.  $\text{C}_{40}\text{H}_{44}\text{O}_5\text{Si}$  requires C, 75.91; H, 7.01%);  $[\alpha]_{\text{D}}^{24} + 3.00$  (20.00 mg / 2 mL in DCM).

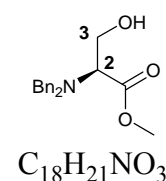
#### Attempted preparation of (S)-1-O-DMTr-3-O-(TBDPS)propan-2-yl methanesulfonate (100a)

Alcohol **100** (0.62 g, 1.00 mmol, 1.00 eq) was dissolved in dry DCM (20 mL) and cooled to 0 °C. To this, TEA (0.84 mL, 6.01 mmol, 6.00 eq) and  $\text{CH}_3\text{SO}_2\text{Cl}$  (0.23 mL, 3.01 mmol, 3.00 eq) was added. The reaction mixture was stirred at rt under  $\text{N}_2$  for 24 h. The reaction mixture was treated with saturated  $\text{NH}_4\text{Cl}$  solution (30 mL) and extracted with diethyl ether (3  $\times$  40 mL). The combined organic layers were washed with saturated  $\text{NaHCO}_3$  solution (30 mL) followed by saturated  $\text{NaCl}$  solution (30 mL). The organic phase was dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford a yellow oil. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). None of the  $^1\text{H}$  NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **100a**, unreacted starting material **100** was the sole compound present.



#### Preparation of (S)-methyl 2-(dibenzylamino)-3-hydroxypropanoate (121)

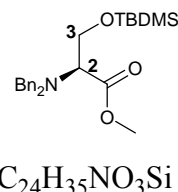
L-Serine methyl ester HCl salt (0.25 g, 1.61 mmol, 1.00 eq) was dissolved in a mixture of dry THF/DMF ((1:1) v/v, 20 mL) and to this, BnBr (0.57 mL, 4.82 mmol, 3.00 eq) and anhydrous  $\text{NaHCO}_3$  (0.54 g, 6.423 mmol, 4.00 eq) was added. The mixture was heated to reflux for 14 h. The reaction mixture was concentrated *in vacuo* and then  $\text{H}_2\text{O}$  (20 mL) was added.



The aqueous layer was extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL) and dried over anhydrous MgSO<sub>4</sub>. Filtration followed by concentration *in vacuo* afforded a clear, colourless liquid. The crude product was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (80:20:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**121**, 0.44 g, 1.47 mmol, 91%):  $R_f = 0.71$  (light petroleum:EtOAc, 80:20);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3439 br s (OH), 3085 m, 3062 m, 3028 s, 2950 s, 1731 s (C=O), 1667 s, 1602 m, 1585 w, 1197 s, 1139 s, 1075 s, 699 s;  $\delta_H$  (200 MHz; CDCl<sub>3</sub>) 3.76 (2 H, d  $J$  5.8, CH<sub>2</sub>), 2.95 (1 H, br s, OH), 3.39-3.89 (8 H, m, OCH<sub>3</sub>, C-2H, and 2 × Benzyl-CH<sub>2</sub>), 7.14-7.31 (10 H, m, Ar-H);  $\delta_C$  (50 MHz; CDCl<sub>3</sub>) 51.2 (CH<sub>3</sub>), 54.5 (2 × CH<sub>2</sub>), 59.0 (CH<sub>2</sub>), 61.4 (CH), 126.6 (2 × CH), 127.2 (4 × CH), 128.3 (4 × CH), 138.4 (2 × C), 171.5 (C);  $m/z$  (ESI) 300 (M+H<sup>+</sup>, 100%), 234 (M-(CO<sub>2</sub>CH<sub>3</sub>)<sup>+</sup>, 59%); [Found (ESI) 300.1597, C<sub>18</sub>H<sub>21</sub>NO<sub>3</sub> requires 300.1594]; (Found: C, 72.02; H, 7.34, N, 4.61. C<sub>18</sub>H<sub>21</sub>NO<sub>3</sub> requires C, 72.22; H, 7.07; N, 4.68%).

#### Preparation of (S)-methyl 3-O-(TBDMS)-2-(dibenzylamino)propanoate (**122**)

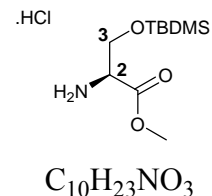
**121** (0.76 g, 2.54 mmol, 1.00 eq) was dissolved in dry DCM (20 mL) and to this imidazole (0.23 g, 3.81 mmol, 1.50 eq) and TBDMSCl (0.49 g, 3.30 mmol, 1.30 eq) was added. The reaction mixture was stirred at rt for 18 h and then H<sub>2</sub>O (20 mL) was added to quench the reaction. The aqueous phase was extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless liquid. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (80:20:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**122**, 0.68 g, 1.64 mmol, 64%):  $R_f = 0.75$  (light petroleum:EtOAc, 80:20);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3086 m, 3063 s, 3029 s, 1737 s (C=O), 1602 w, 1586 w, 1198 s, 1147 s, 1077 s, 1006 m, 837 s, 778 s, 698 s;  $\delta_H$  (200 MHz; CDCl<sub>3</sub>) 0.01 (6 H, s, 2 × CH<sub>3</sub>), 0.87 (9 H, s, (CH<sub>3</sub>)<sub>3</sub>C), 3.53-4.05 (10 H, m, CH<sub>3</sub>, 2 × Benzyl-CH<sub>2</sub>, C-2H and C-3H<sub>2</sub>) 7.12-7.43 (10 H, m, Ar-H);  $\delta_C$  (50 MHz; CDCl<sub>3</sub>) 18.2 (CH<sub>3</sub>), 25.9 (3 × CH<sub>3</sub>), 33.6 (CH<sub>3</sub>), 51.1 (CH<sub>3</sub>), 55.6 (2 × CH<sub>2</sub>), 62.8 (CH<sub>2</sub>), 63.1 (CH), 127.0 (2 × CH), 128.3 (2 × CH), 128.4 (2 × CH), 128.8 (2 ×



CH), 129.1 (2 × CH), 139.9 (2 × C), 172.0 (C);  $m/z$  414 (ESI) ( $M+H^+$ , 100%), 91 ( $M-PhCH_2^+$ , 100%); [Found 414.2459, (ESI)  $C_{24}H_{35}NO_3Si$  requires 414.2458].

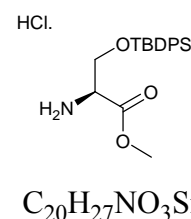
### Preparation of (S)-methyl 2-amino-3-O-(TBDMS)propanoate (**130**)

L-Serine methyl ester HCl salt (3.00 g, 19.28 mmol, 1.00 eq) was suspended in dry DCM (40 mL) and to this, imidazole (1.97 g, 28.92 mmol, 1.50 eq) and TBDMSCl (3.78 g, 25.06 mmol, 1.30 eq) was added. The reaction mixture was stirred at rt for 18 h and then  $H_2O$  (40 mL) was added to quench the reaction. The aqueous phase was extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous  $MgSO_4$ , filtered and concentrated *in vacuo* to afford the title compound, as a white solid. (**130**, 3.93 g, 16.78 mmol, 88%):  $R_f$  = 0.28 (light petroleum:EtOAc, 80:20);  $\nu_{max}$  (KBr)/ $cm^{-1}$  3360 m ( $NH_2$ ), 2929 s, 1749 s (C=O), 1006 m;  $\delta_H$  (200 MHz;  $CDCl_3$ ) 0.02 (6 H, m, 2 ×  $CH_3$ ), 0.86 (9 H, s,  $(CH_3)_3C$ ), 3.76 (3 H, s,  $CH_3$ ), 3.90-4.05 (3 H, m, C-2  $H$  and C-3 $H_2$ ), 5.85 (2 H, br s, 2 ×  $NH$ );  $\delta_c$  (50 MHz;  $CDCl_3$ ) 17.6 (2 ×  $CH_3$ ), 25.1 (3 ×  $CH_3$ ), 25.7 (C), 52.0 ( $CH_3$ ), 55.0 (CH), 62.5 ( $CH_2$ ), 170.2 (C);  $m/z$  (ESI) 234 ( $M+H^+$ , 100%); [Found (ESI) 234.1523,  $C_{10}H_{24}NO_3Si$  requires 234.1520].



### Preparation of (S)-methyl 2-amino-3-O-(TBDPS)propanoate (**135**)

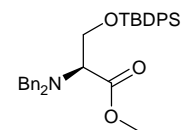
L-Serine methyl ester HCl salt (3.00 g, 19.28 mmol, 1.00 eq) was taken up in dry DCM (30 mL) and to this, imidazole (1.99 g, 28.92 mmol, 1.50 eq) and TBDPSCl (6.89 g, 25.06 mmol, 1.30 eq) was added. The reaction mixture was allowed to stir under  $N_2$  at rt for 6 h and then  $H_2O$  (30 mL) was added to quench the reaction. The aqueous phase was extracted with EtOAc (5 × 50 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous  $MgSO_4$ , filtered and concentrated *in vacuo* to afford a clear, yellow oil. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**135**, 2.93 g, 8.18 mmol, 43%):  $R_f$  = 0.12 (light petroleum:EtOAc, 50:50);  $\nu_{max}$  (neat)/ $cm^{-1}$  3386 m ( $NH_2$ ), 1742 s (C=O);  $\delta_H$  (200 MHz;  $CDCl_3$ ) 0.89 (9 H, s,  $(CH_3)_3C$ ), 1.67 (2 H, br s,  $NH_2$ ), 3.56 (1 H, t,  $J$  4.15,



7.89, *CH*), 3.56 (3 H, s, *CH*<sub>3</sub>), 3.68–3.87 (2 H, dd, *J* 13.9, 4.15, 7.89, *CH*<sub>2</sub>), 7.18–7.30 (6 H, m, *Ar-H*), 7.50–7.52 (4 H, m, *Ar-H*);  $\delta_c$  (50 MHz; CDCl<sub>3</sub>) 19.2 (*CH*<sub>3</sub>), 26.7 (3 × *CH*<sub>3</sub>), 51.9 (*C*), 56.4 (*CH*), 66.1 (*CH*<sub>2</sub>), 127.7 (4 × *CH*), 129.8 (2 × *CH*), 133.0 (2 × *C*), 135.5 (4 × *CH*), 175.1 (*C*); *m/z* 358 (ESI) (*M-HCl*<sup>+</sup>, 26%), 715 (*M+M*<sup>+</sup>, 100%); [Found 358.1827, (ESI) C<sub>20</sub>H<sub>28</sub>NO<sub>3</sub>Si requires 358.1833].

### Preparation of (*S*)-methyl 3-O-(TBDPS)-2-(dibenzylamino)propanoate (**122a**)

**121** (0.45 g, 1.08 mmol, 1.00 eq) was taken up in dry DCM (20 mL) and to this, imidazole (0.110 g, 1.62 mmol, 1.50 eq) and TBDPSCI (0.21 g, 1.40 mmol, 1.30 eq) were added. The reaction mixture was stirred at rt under N<sub>2</sub> for 6 h and H<sub>2</sub>O (20 mL) was added to quench the reaction. The aqueous layer was extracted with EtOAc (4 × 50 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a white oil. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (80:20:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**122a**, 0.38 g, 0.71 mmol, 66%); *R*<sub>f</sub> = 0.71 (light petroleum:EtOAc, 80:20);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3069 s, 3028 s, 1798 s (C=O);  $\delta_H$  (200 MHz; CDCl<sub>3</sub>) 0.89 (9 H, s, (*CH*<sub>3</sub>)<sub>3</sub>C), 3.51–3.65 (6 H, m, *CH*, *CH*<sub>2</sub> and *CH*<sub>3</sub>), 3.78–3.96 (4 H, m, 2 × Benzyl-*CH*<sub>2</sub>), 7.06–7.34 (16 H, m, *Ar-H*), 7.45–7.51 (4 H, m, *Ar-H*);  $\delta_c$  (50 MHz; CDCl<sub>3</sub>) 19.3 (*CH*<sub>3</sub>), 26.8 (*CH*<sub>3</sub>), 51.3 (*CH*), 55.6 (*CH*<sub>2</sub>), 63.2 (*CH*<sub>2</sub>), 63.5 (*CH*<sub>2</sub>), 127.1 (2 × *CH*), 127.9 (4 × *CH*), 128.4 (4 × *CH*), 128.9 (4 × *CH*), 129.9 (2 × *CH*), 133.3 (2 × *C*), 135.8 (4 × *CH*), 139.9 (2 × *C*) 172.1 (*C*); *m/z* 538 (HNES) (*M+NH*<sub>4</sub><sup>+</sup>, 100%); [Found 538.2763, (HNES) C<sub>34</sub>H<sub>39</sub>NO<sub>3</sub>Si+NH<sub>4</sub><sup>+</sup> requires 538.2772].



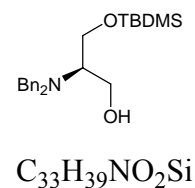
---

**Attempted preparation of (R)-3-O-(TBDMS)-2-(dibenzylamino)propan-1-ol (123a)**


---

*Method A*

**122** (0.11 g, 0.26 mmol, 1.00 eq) was dissolved in Et<sub>2</sub>O (20 mL). The solution was cooled to 0 °C and LiBH<sub>4</sub> (0.02 g, 1.06 mmol, 4.00 eq) and MeOH (0.33 mL) were added. The reaction mixture was heated to reflux for 6 h. Saturated NH<sub>4</sub>Cl solution (20 mL) was added to quench the reaction. The aqueous phase was separated and extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless oil. The residue was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (80:20:1). None of the <sup>1</sup>H NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **123**; unreacted starting material was the sole compound present.

*Method B*

**122** (0.20 g, 0.48 mmol, 1.00 eq) was dissolved in diethyl ether (20 mL). The solution was cooled to 0 °C and then LiBH<sub>4</sub> (0.04 g, 1.92 mmol, 4.00 eq) and MeOH (0.60 mL) were added. The reaction mixture was heated to reflux for 24 h. Saturated NH<sub>4</sub>Cl solution (20 mL) was added to quench the reaction. The aqueous phase was separated and extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (80:20:1). None of the <sup>1</sup>H NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **123**, unreacted starting material was the sole compound present.

*Method C*

NaBH<sub>4</sub> (0.99 g, 26.42 mmol, 6.00 eq) was suspended in dry THF (20 mL) in the presence of **122** (1.83 g, 4.40 mmol, 1.00 eq) and was heated to reflux for 1 h. MeOH (10 mL) was added dropwise down the condenser over 30 min and reflux was maintained for a further 4 h. The reaction was quenched with saturated NH<sub>4</sub>Cl solution

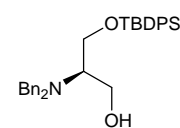
(15 mL) before being left to stir at rt for a further 1 h. The organic phase was aqueous phase was extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, yellow liquid. The residue was purified by flash chromatography on a silica column eluting with light petroleum ether:EtOAc:TEA (80:20:1). None of the <sup>1</sup>H NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **123**, unreacted starting material was the sole compound present.

*Method D*

**122** (0.68 g, 1.64 mmol, 1.00 eq) in dry THF (10 mL) was added dropwise to a stirred solution of LiAlH<sub>4</sub> (0.06 g, 1.64 mmol, 1.00 eq) in dry THF (10 mL) at 0 °C under N<sub>2</sub>. The reaction mixture was stirred at rt for 24 h. Subsequently, the reaction was quenched by careful addition of an aqueous saturated solution of diethyl ether (20 mL). The inorganic precipitates were removed by filtration and washed with EtOAc (50 mL). The filtrate was concentrated *in vacuo* to afford a yellow oil. The crude product was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (50:50:1). None of the <sup>1</sup>H NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **123**. Starting ester and the desilylated alcohol and TBDMSOH were recovered.

**Attempted preparation of (R)-3-O-(TBDPS)-2-(dibenzylamino)propan-1-ol (123a)**

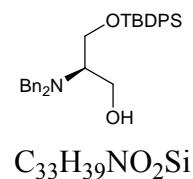
A solution of **122a** (0.42 g, 0.78 mmol, 1.00 eq) was added slowly to a stirred solution of LiAlH<sub>4</sub> (0.01 g, 0.39 mmol, 0.50 eq) in dry THF (10 mL) at 0 °C under N<sub>2</sub>. The reaction mixture was stirred at rt for 24 h.



The reaction was quenched by the careful addition of an aqueous saturated solution of diethyl ether (20 mL). The resulting solution was dried over anhydrous MgSO<sub>4</sub>, filtered and washed with EtOAc (20 mL). The filtrate was concentrated *in vacuo* to afford a clear, colourless oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). None of the <sup>1</sup>H NMR spectra recorded from the fractions obtained were consistent with formation of the desired compound **123a**, the ester **122a** was the sole compound present.

### Preparation of (*R*)-3-O-(TBDPS)-2-(dibenzylamino)propan-1-ol (**123a**)

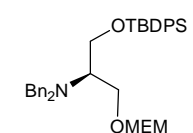
A solution of **122a** (0.27 g, 0.50 mmol, 1.00 eq) in dry THF (5 mL) was added to a stirred solution of LiBH<sub>4</sub> (0.02 g, 21.78 mmol, 2.00 eq) in dry THF (15 mL) at 0 °C under N<sub>2</sub>. A solution of β-methoxy-9-borabicyclo[3.3.1]nonane (9-BBN) (1M in hexane, 0.38 mL, 1.10 mmol, 2.20 eq) was added dropwise to the mixture. The reaction mixture was allowed to warm to rt and maintained for 24 h. Solvent was removed *in vacuo* and H<sub>2</sub>O (50 mL) was added. The aqueous layer was extracted with EtOAc (5 × 50 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow oil. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**123a**, 0.12 g, 0.24 mmol, 48%): *R*<sub>f</sub> = 0.67 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3400 w (OH), 3069 m, 3028 m;  $\delta_{\text{H}}$  (200 MHz; CDCl<sub>3</sub>) 1.08 (9 H, s, (CH<sub>3</sub>)<sub>3</sub>C), 1.65 (1 H, br s, OH), 3.18 (1 H, t, *J* 6, CH), 3.49–4.05 (8 H, m, CH<sub>2</sub>-OH, CH<sub>2</sub> and 2 × Benzyl-CH<sub>2</sub>), 7.20–7.49 (16 H, m, Ar-*H*), 7.57–7.75 (4 H, m, Ar-*H*);  $\delta_{\text{C}}$  (50 MHz; CDCl<sub>3</sub>) 19.1 (C), 26.6 (3 × CH<sub>3</sub>), 53.9 (2 × CH<sub>2</sub>), 59.4 (CH<sub>2</sub>), 59.0 (CH<sub>2</sub>), 61.3 (CH), 127.2 (2 × CH), 127.8 (4 × CH), 128.6 (4 × CH), 128.9 (4 × CH), 129.6 (2 × CH), 134.8 (2 × C), 135.5 (4 × CH), 139.4 (2 × C (benzyl)); *m/z* 510 (HNES) (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found 510.2823, (HNES) C<sub>33</sub>H<sub>39</sub>NO<sub>2</sub>Si requires 510.2823].



### Preparation of (*R*)-3-O-(TBDPS)-2-(dibenzylamino)-1-((2-methoxyethoxy)methoxy)propane (**124**)

#### Method A

**123a** (1.16 g, 2.28 mmol, 1.00 eq) was taken up in dry THF (20 mL) and to this, NaH (60 % in mineral oil) (0.88 g, 2.28 mmol, 1.00 eq) was added. The reaction mixture was allowed to stir at rt under N<sub>2</sub> for 3 h. 2-Methoxyethoxymethyl chloride (MEM-Cl) (0.26 mL, 2.28 mmol, 1.00 eq) was added and the reaction mixture was maintained for 24 h. Solvent was removed *in vacuo* to afford the crude product as a clear, colourless oil. The residue was



purified by flash chromatography on a silica column eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**124**, 1.00 g, 1.67 mmol, 98%):  $R_f = 0.63$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$  3069 w, 3028 w, 2928 s, 2856 s;  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) 0.83–0.88 (2 H, m,  $\text{CH}_2$ ), 1.05–1.08 (9 H, d,  $J$  6,  $(\text{CH}_3)_3\text{C}$ ), 1.26 (3 H, s,  $\text{CH}_3$ ), 3.11 (1 H, t,  $J$  6, CH), 3.36 (2 H, s,  $\text{CH}_2$ ), 3.45–3.95 (10 H, m, Benzyl- $\text{CH}_2$ ,  $3 \times \text{CH}_2$ ), 7.18–7.47 (16 H, m, Ar- $H$ ), 7.62–7.74 (4 H, m, Ar- $H$ );  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) 19.1 (C), 26.6 ( $3 \times \text{CH}_3$ ), 53.9 ( $\text{CH}_2$ ), 55.1 ( $\text{CH}_2$ ), 58.9 (CH), 60.0 ( $\text{CH}_2$ ), 61.3 ( $\text{CH}_2$ ), 62.6 ( $\text{CH}_2$ ), 67.1 ( $\text{CH}_2$ ), 71.7 ( $\text{CH}_2$ ), 95.6 ( $\text{CH}_2$ ), 126.6 ( $2 \times \text{CH}$ ), 128.1 ( $4 \times \text{CH}$ ), 128.5 ( $4 \times \text{CH}$ ), 129.8 ( $4 \times \text{CH}$ ), 129.8 ( $2 \times \text{CH}$ ), 134.8 ( $2 \times \text{C}$ ), 135.6 ( $4 \times \text{CH}$ ), 140.5 ( $2 \times \text{C}$ ); (impurities present in both spectra)  $m/z$  598 (HNES) ( $\text{M}+\text{NH}_4^+$ , 100%), 510 ( $\text{M}-(\text{MEM})^+$ , 54%); [Found 598.3333, (HNES)  $\text{C}_{37}\text{H}_{51}\text{N}_2\text{O}_4\text{Si}$  requires 598.3347].

#### Method B

**123a** (0.20 g, 0.39 mmol, 1.00 eq) was dissolved in dry DCM (10 mL) and cooled to 0 °C. To this, MEM-Cl (0.07 mL, 0.59 mmol, 1.50 eq) was added and the reaction mixture was stirred at rt under  $\text{N}_2$  for 24 h. The reaction mixture was cooled to 0 °C and quenched with saturated aqueous  $\text{NaHCO}_3$  solution (6 mL). The pH was increased to 7.0 using aqueous NaOH solution (10 mL) and the organic phase was separated. The aqueous phase was extracted with DCM ( $4 \times 40$  mL) and the combined organic layers were washed with saturated NaCl solution (40 mL), dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford a clear, yellow oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo*.

#### Method C

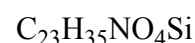
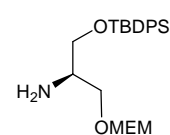
**123a** (0.05 g, 0.19 mmol, 1.00 eq) was dissolved in dry DCM (5 mL) and the reaction mixture was cooled to 0 °C. To this, *N*-ethyl-diisopropylamine (DIPEA) (0.05 mL, 0.29 mmol, 1.50 eq) and MEM-Cl (0.03 mL, 0.29 mmol, 1.50 eq) were added and the reaction mixture was allowed to warm to rt before being left to stir under  $\text{N}_2$  for 24 h. The reaction mixture was cooled to 0 °C and quenched with an aqueous  $\text{NaHCO}_3$  solution (20%, 1 mL). The aqueous phase was extracted with DCM ( $4 \times 10$  mL) and the



combined organic layers were washed with saturated NaCl solution (20 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless oil. The residue was purified by flash chromatography on a silica column eluting with dichloromethane:TEA (99:1). Fractions containing product were combined and concentrated *in vacuo*. <sup>1</sup>H NMR confirmed that the reaction had failed, with respect to the impurities present in the spectra of method A, still being present.

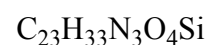
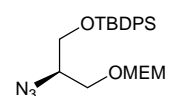
**Preparation of (R)-3-O-(TBDPS)-1-((2-methoxyethoxy)methoxy)propan-2-amine (125)**

**124** (0.48 g, 1.08 mmol, 1.00 eq) was dissolved in EtOH (9 mL). The solution was flushed through the H-cube at 0.5 mL/min, 50 °C, 70 bar, controlled H<sub>2</sub> 79 bar, HPLC pump 78 bar. The product was collected and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**125**, 0.48 g, 1.15 mmol, 100%): *R<sub>f</sub>* = 0.0 (light petroleum:EtOAc, 50:50); *v*<sub>max</sub> (neat)/cm<sup>-1</sup> 3351 s, 3071 s, 2926 s, 1651 m, 1471 m, 1427m, 823 m, 700 m. <sup>1</sup>H NMR spectroscopy confirmed that the title compound, **125**, was present (100%) along with the impurities observed in previous steps.



**Preparation of (R)-2-azido-1-O-((2-methoxyethoxy)methoxy)-3-O-(TBDPS)propoxy (126)**

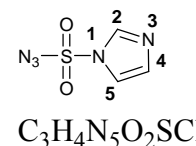
**125** (0.29 g, 1.08 mmol, 1.00 eq), imidazole-1-sulfonyl azide HCl salt **152** (0.21 g, 1.20 mmol, 1.10 eq), K<sub>2</sub>CO<sub>3</sub> (0.24 g, 1.28 mmol, 1.00 eq) and copper sulfate pentahydrate (CuSO<sub>4</sub>·5H<sub>2</sub>O) (0.003 g, 0.01 mmol, 0.01 eq) were taken up in dry MeOH (20 mL) and the reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. The solvent was removed *in vacuo* and subsequently, H<sub>2</sub>O (20 mL) was added and the resulting solution was extracted with EtOAc (4 × 30 mL). The combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:EtOH (25:72:3). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound. *v*<sub>max</sub> (neat)/cm<sup>-1</sup> 3368 s, 3071 s, 2930 s, 2096 s, 1718 m, 1472 m, 1428m, 823 m, 702 m. <sup>1</sup>H NMR spectroscopy



confirmed that the title compound, **126** was present along with the impurities observed in previous steps.

### Preparation of imidazole-1-sulfonyl azide HCl salt (**152**)<sup>96</sup>

Sulfonyl chloride **154** (SO<sub>2</sub>Cl<sub>2</sub>) (3.71 mL, 46.15 mmol, 1.00 eq) <sup>HCl.</sup> was added dropwise to an ice-cooled suspension of NaN<sub>3</sub> (3.00 g, 46.15 mmol, 1.00 eq) in dry acetonitrile (30 mL) and the reaction mixture was stirred at rt for 24 h under N<sub>2</sub>. Imidazole (5.97 g, 87.69 mmol, 1.90 eq) was added portion wise to the ice-cooled mixture and the resulting slurry was stirred for 3 h at rt. The mixture was diluted with EtOAc (100 mL), washed with H<sub>2</sub>O (2 × 100 mL) and the aqueous phase was back-extracted with EtOAc (100 mL). The combined organic layers were washed with saturated NaHCO<sub>3</sub> solution (2 × 100 mL), dried over anhydrous MgSO<sub>4</sub> and filtered. A solution of HCl in EtOH [obtained by the drop-wise addition of acetyl chloride (CH<sub>3</sub>COCl) (4.92 mL, 69.23 mmol, 1.50 eq) to ice-cooled dry EtOH (30mL)] was added drop-wise to the filtrate with stirring, the mixture chilled in an ice-bath, filtered and the filter cake was washed with EtOAc (3 × 100 mL) to afford the title compound as a white solid. (**152**, 5.52 g, 26.33 mmol, 69%): *R*<sub>f</sub> = 0.35 (light petroleum:EtOAc, 50:50); *ν*<sub>max</sub> (neat)/cm<sup>-1</sup> 2172 s, 1581 s, 1508 s, 1429 s, 1322 s, 1162 s; *δ*<sub>H</sub> (200 MHz; D<sub>2</sub>O) 7.59-7.65 (1 H, m, C-4/5H), 7.98-8.05 (1 H, m, C-4/5H), 9.49-9.55 (1 H, m, C-2H); *δ*<sub>c</sub> (50 MHz; D<sub>2</sub>O) 119.9 (CH), 121.9 (CH), 137.3 (C-2); (Found: C, 17.30; H, 1.92; N, 33.57. C<sub>3</sub>H<sub>4</sub>N<sub>5</sub>O<sub>2</sub>SCl requires C, 17.19; H, 1.92; N, 33.41%).



### Attempted preparation of (S)-methyl 2-azido-3-O-(TBDPS) propanoate (**157**)

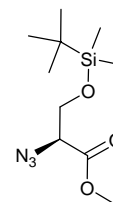
#### Method A

**135** (0.38 g, 2.17 mmol, 1.00 eq), **152** (0.41 g, 2.39 mmol, 1.10 eq), K<sub>2</sub>CO<sub>3</sub> (0.30 g, 2.17 mmol, 1.00 eq) and CuSO<sub>4</sub>·5H<sub>2</sub>O (0.005 g, 0.02 mmol, 0.01 eq) were taken up in dry MeOH (20 mL) and the reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. The solvent was removed *in vacuo* and subsequently, H<sub>2</sub>O (5 mL) was added and the resulting solution was extracted with EtOAc (4 × 30 mL). The combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous MgSO<sub>4</sub>,

filtered and concentrated *in vacuo* to afford a blue/green residue.  $^1\text{H}$  NMR confirmed that the reaction had failed. The desilylated compound was recovered (~92%).

### Preparation of (S)-methyl 2-azido-3-O-(TBDMS) propanoate (158)

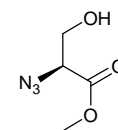
**130** (0.50 g, 2.85 mmol, 1.00 eq),  $\text{K}_2\text{CO}_3$  (0.39 g, 2.85 mmol, 1.00 eq),  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (0.007 g, 0.03 mmol, 0.01 eq) and **152** (0.54 g, 3.14 mmol, 1.10 eq) were suspended in dry MeOH (10 mL) and the reaction mixture was stirred at rt under  $\text{N}_2$  for 24h. (The colour



changed from green to blue) The reaction mixture was concentrated *in vacuo* and  $\text{H}_2\text{O}$  (10 mL) was added. The aqueous phase was extracted with EtOAc ( $4 \times 40$  mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was used immediately in the subsequent step;  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$  2117 s ( $\text{N}_3$ ) (solvent peaks present in spectrum);  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) 0.03 (6H, s,  $2 \times \text{CH}_3$ ), 0.83 (9H, s,  $3 \times \text{CH}_3$ ), 3.74 (3H, s,  $\text{OCH}_3$ ), 3.81 (1H, s, CH), 3.98-4.02 (2H, m (confused by solvent),  $\text{CH}_2$ ) (solvent/imidazole peaks present on spectrum);  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) 13.9 ( $2 \times \text{CH}_3$ ), 16.4 (C), 25.3 ( $3 \times \text{CH}_3$ ), 60.1 (CH), 62.8 ( $\text{CH}_3$ ), 64.2 ( $\text{CH}_2$ ), 171.1 (C) (solvent/imidazole peaks present on spectrum);  $m/z$  (HNESP) 260 ( $\text{M}+\text{H}^+$ , 100%); [Found (HNESP) 260.1429,  $\text{C}_{10}\text{H}_{21}\text{N}_3\text{O}_3\text{Si}$  requires 260.1425].

### Preparation of (S)-methyl 2-azido-3-hydroxypropanoate (159)

L-Serine methyl ester HCl salt (0.50 g, 3.21 mmol, 1.00 eq),  $\text{K}_2\text{CO}_3$  (0.44 g, 3.21 mmol, 1.00 eq) and  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (0.01 g, 0.03 mmol, 0.01 eq) were suspended in dry MeOH (10 mL) before being left to stir at rt under  $\text{N}_2$  for 24 h. The solvent was removed *in vacuo* and  $\text{H}_2\text{O}$  (10 mL)

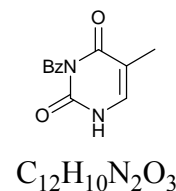


was added. The aqueous phase was extracted with EtOAc ( $4 \times 20$  mL) and the combined organic layers were washed with saturated NaCl solution (30 mL). The organic phase was dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was used immediately in the subsequent;  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$  2110 s ( $\text{N}_3$ ) (solvent peaks present in spectrum);  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) 3.64 (3H, s,  $\text{OCH}_3$ ), 3.70 (2H, s,  $\text{CH}_2$ ), 3.82 (1H, s, CH), 5.74 (1H, br s, OH) (solvent/imidazole peaks present on spectrum);  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) 52.5 ( $\text{CH}_3$ ), 62.3 ( $\text{CH}_2$ ), 62.9 (CH), 168.9 (C) (solvent/imidazole peaks present on spectrum);  $m/z$

(ASAP) 163 ( $M+NH_4^+$ , 100%); [Found (ASAP) 163.0825,  $C_4H_{11}N_4O_3$  requires 163.0826].

### Preparation of *N*3-benzoyl thymine (**160**)

Benzoyl chloride (BzCl) (1.01 mL, 8.72 mmol, 2.20 eq) was added to a mixture of thymine (0.55 g, 4.36 mmol, 1.00 eq) in dry  $CH_3CN/C_5H_5N$  ((5:2) v/v, 20 mL). The mixture was stirred at rt for 24 h, filtered and concentrated *in vacuo* to afford a yellow solid. A solution of anhydrous  $K_2CO_3$  (0.60 g, 4.36 mmol, 1.00 eq) dissolved in dioxane/ $H_2O$  ((1:1) v/v, 10 mL) was used to dissolve the solid and the mixture was allowed to stir for 1 h at rt. The mixture was acidified to pH 5.0 using acetic acid and then concentrated *in vacuo*. The yellow solid was taken up in a saturated  $NaHCO_3$  solution (30 mL) before being left to stir for 2 h at rt. Filtration followed by recrystallisation with  $MeOH/Et_2O$  afforded the title compound as an off white solid. (**160**, 1.34 g, 5.80 mmol, 73%):  $R_f = 0.39$  (DCM:MeOH, 94:6); m.p. 166-167 °C;  $\nu_{max}$  (KBr)/ $cm^{-1}$  3180 m, 3054 m, 2825 m, 1734 s (C=O), 1677 s (C=O), 1460 m, 1452 m, 1421 m;  $\delta_H$  (200 MHz;  $DMSO-d^6$ ) 1.86 (3 H, s,  $CH_3$ ), 7.45-7.56 (3 H, m, Ar-*H*), 7.61-7.71 (1 H, m, C-6*H*), 7.77-7.86 (2 H, m, Ar-*H*), 11.58 (1 H, b s, NH);  $\delta_c$  (50 MHz;  $DMSO-d^6$ ) 12.4 ( $CH_3$ ), 111.8 (C), 128.8 (2 × C), 129.8 (2 × C), 133.7 (C), 133.9 (C), 136.1 (C), 149.8 (C), 164.5 (C), 170.5 (C);  $m/z$  (ESI) 231 ( $M+H^+$ , 11%), 248 ( $M-(Methyl)^+$ , 50%); [Found (ESI) 231.0762,  $C_{12}H_{11}N_2O_3$  requires 231.0764]; (Found: C, 62.41; H, 4.26; N, 12.09.  $C_{12}H_{12}N_2O_3$  requires C, 62.61; H, 4.38; N 12.17%); The X-Ray crystal structure (Appendix C) was acquired from the crystals obtained from  $Et_2O$  and MeOH.

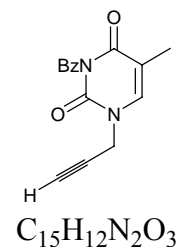


### Attempted preparation of 1-(prop-2-ynyl)-*N*3-benzoyl thymine (**161**)

**160** (0.170 g, 0.738 mmol, 1.00 eq) was added to dry DMF (20 mL) and to this NaH (0.024 g, 1.45 mmol, 1.00 eq) and propargyl bromide (0.067 mL, 7.38 mmol, 0.95 eq) was added. The resulting mixture was stirred at rt for 24 h. TLC analysis showed that no reaction had occurred. The reaction mixture was heated in increments of 10 °C up to a maximum of 80 °C over a 4 h period. The  $^1H$  NMR spectrum recorded of the crude mixture showed that the reaction had failed.

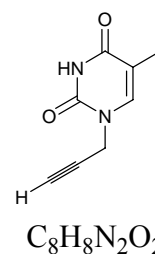
### Preparation of 1-(prop-2-ynyl)-*N*-benzoyl thymine (**161**)

**160** (1.00 g, 4.34 mmol, 3.00 eq) was added to dry DMF (20 mL) and to this propargyl bromide (0.129 mL, 1.45 mmol, 1.00 eq) and Cs<sub>2</sub>CO<sub>3</sub> (0.94 g, 2.89 mmol, 2.00 eq) was added. The resulting mixture was stirred at rt for 48 h. The solvent was removed *in vacuo* and H<sub>2</sub>O was added (30 mL). The aqueous phase was extracted with EtOAc (4 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL). The organic phase was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a white solid. The crude product was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc:TEA (50:50:1). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white solid. (**161**, 0.64 g, 2.38 mmol, 55%): *R*<sub>f</sub> = 0.19 (light petroleum:EtOAc, 50:50); m.p. 178-180 °C;  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3261 m, 3073 w, 2928 w, 2129 w (alkyne), 1747 s (C=O), 1696 s (C=O), 1644 s (C=O), 1492 w, 1442 s;  $\delta_{\text{H}}$  (200 MHz; CDCl<sub>3</sub>) 1.01 (3 H, d, *J* 0.8, CH<sub>3</sub>), 2.67 (1 H, t, *J* 2.5, Alkyne-*H*), 4.56 (2 H, d, *J* 2.5, N-CH<sub>2</sub>), 7.52-7.70 (2 H, m, Ar-*H*), 7.73-7.96 (4 H, m, Ar-*H*);  $\delta_{\text{C}}$  (50 MHz; CDCl<sub>3</sub>) 11.8 (CH<sub>3</sub>), 37.0 (CH<sub>2</sub>), 76.3 (C), 77.9 (C), 109.3 (CH), 129.4 (2 × CH), 130.2 (2 × CH), 130.9 (C), 135.5 (CH), 141.1 (CH), 148.8 (C), 162.7 (C), 169.5 (C); *m/z* (ESI) 269 (M+H<sup>+</sup>, 71%), 286 (M+(Methyl)<sup>+</sup>, 100%), 554 (M+(M+Methyl)<sup>+</sup>, 86%); [Found (ESI) 269.0922, C<sub>15</sub>H<sub>13</sub>N<sub>2</sub>O<sub>3</sub> requires 269.0921]; (Found: C, 66.99; H, 4.48, N, 10.24. C<sub>15</sub>H<sub>12</sub>N<sub>2</sub>O<sub>3</sub> requires C, 67.16; H, 4.51; 10.44%); The X-Ray crystal structure (Appendix D) was acquired from the crystals obtained from Et<sub>2</sub>O and MeOH.



### Preparation of 1-(prop-2-ynyl)-thymine(**162**)

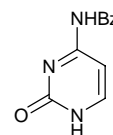
**161** (0.10 g, 0.37 mmol, 1.00 eq) was dissolved in benzyl alcohol (BnOH) (1 mL) and heated to 90 °C for 30 h. The solvent was removed *in vacuo* and the residue was purified by flash chromatography on a silica column eluting with (light petroleum:EtOAc (100) to light petroleum:EtOAc (80:20) to light petroleum:EtOAc (50:50) gradient elution). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white solid. (**162**, 0.05 g, 82%): *R*<sub>f</sub> = 0.44 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3251 m, 3014 w, 2934 w, 2125 w (alkyne), 1701 s (C=O), 1698 s (C=O), 1473 s, 1423 s;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) 1.95 (3H, d, *J* 0.8,



$CH_3$ ), 2.47 (1H, t,  $J$  2.5, Alkyne- $H$ ), 4.51 (2H, d,  $J$  2.49, N- $CH_2$ ), 7.21 (1H, d, C-6H), 8.12 (1H, NH);  $\delta_c$  (50 MHz;  $CDCl_3$ ) 12.4 ( $CH_3$ ), 36.6 ( $CH_2$ ), 75.2 (C), 77.6 (C), 111.3 (C), 138.8 (CH), 150.1 (C), 164.3 (C);  $m/z$  (ASAP) 165 ( $M+H^+$ , 100%); [Found (ASAP) 165.0654,  $C_8H_9N_2O_2$  requires 165.0659].

### Preparation of *N*-benzoyl cytosine (**165**)

Cytosine (1.40 g, 12.60 mmol, 1.00 eq) was suspended in dry  $C_5H_5N$  (30 mL) and was stirred at rt under  $N_2$  in the presence of BzCl (3.45 mL, 27.72 mmol, 2.20 eq) for 24 h. The reaction mixture was acidified with HCl solution to pH 5.0 and stirred for a further 3 h. The precipitate was collected and washed with hot EtOH (100 mL). The solid was recrystallised from an aqueous acetic acid solution and the title compound was collected as an off-white solid. (**165**, 2.49 g, 92%):  $\nu_{max}$  (neat)/ $cm^{-1}$  2964 m, 2323 w, 2015 w (alkyne), 1738 s (C=O), 1699 s (C=O), 1504 s, 1422 s; (Found: C, 61.47; H, 4.15, N, 19.36.  $C_{11}H_9N_3O_2$  requires C, 61.39; H, 4.22; N 19.53%).



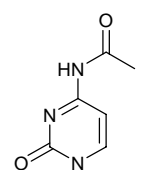
$C_{11}H_9N_3O_2$

### Attempted preparation of 1-(prop-2-ynyl)-*N*4-benzoyl cytosine (**174**)

A suspension of **165** (0.50 g, 2.32 mmol, 1.00 eq) and DBU (0.316 mL, 2.11 mmol, 0.91 eq) in dry DMF (20 mL) was stirred at rt for 1 h. To this, propargyl bromide (0.248 mL, 2.32 mmol, 1.00 eq) was added and the reaction mixture was left overnight in the dark. The solvent was removed in vacuo and the crude residue was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc (60:40). From the  $^1H$  NMR spectra obtained, the title compound **174** was not observed in any of the product containing fractions.

### Preparation of 1-(prop-2-ynyl)-4-acetylamino pyrimidin-2(1H)-one<sup>105</sup>

A suspension of *N*<sup>4</sup>-acetylcytosine **171** (1.00 g, 6.53 mmol, 1.00 eq), anhydrous  $K_2CO_3$  (1.04 g, 7.50 mmol, 1.33 eq) and propargyl bromide (80% in toluene, 0.50 mL, 5.64 mmol, 1.00 eq) in dry DMF (20 mL) was stirred at rt under  $N_2$  for 24 h. The mixture was filtered and the solvent was removed *in vacuo* to afford a brown solid. The residue was dissolved

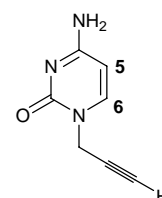


$C_9H_9N_3O_2$

in hot H<sub>2</sub>O (20 mL) and treated with decolourising charcoal (0.20 g). The mixture was filtered through celite© and the clear filtrate was reduced in volume by half, by gentle heating. The cooled filtrate precipitated the title compound as an off white solid. (**173**, 0.51 g, 2.66 mmol, 40%):  $R_f = 0.24$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3197 s (NH<sub>2</sub>), 3028 w, 2114 w (Alkyne), 1695 s (C=O), 1659 s (C=O);  $\delta_H$  (200 MHz; DMSO-*d*<sup>6</sup>) 2.09 (3 H, s, CH<sub>3</sub>), 3.45 (1 H, t, *J* 3.2, Alkyne-CH), 4.63–4.64 (2 H, d, *J* 3.4, Alkyne-CH<sub>2</sub>), 7.19–7.22 (1 H, d, *J* 7.2, CH), 8.13–8.17 (1 H, d, *J* 7.3, CH), 10.89 (1 H, bs, NH);  $\delta_C$  (50 MHz; DMSO-*d*<sup>6</sup>) 24.3 (CH<sub>3</sub>), 38.4 (CH<sub>2</sub>), 76.4 (C), 78.3 (C), 95.7 (CH), 149.1 (CH), 154.6 (C), 162.6 (C), 170.9 (C); *m/z* (HNES) 192 (M+NH<sub>4</sub><sup>+</sup>, 100%); (Found (HNES) 192.0766, C<sub>9</sub>H<sub>13</sub>N<sub>4</sub>O<sub>2</sub> requires 192.0768).

#### Preparation of 4-amino-1-(prop-2-ynyl)pyrimidin-2(1H)-one (**174**)<sup>105</sup>

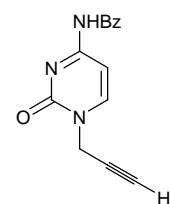
**173** (0.19 g, 0.99 mmol, 1.00 eq) was stirred in a methanolic ammonia solution (5% v/v, 20 mL) at rt under N<sub>2</sub> for 24 h. The solvent was removed *in vacuo* to afford the title compound as an off white solid. (**174**, 0.14 g, 0.94 mmol, 94%);  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3381 s (NH<sub>2</sub>), 2122 w (Alkyne), 1666 s (C=O), 1614 s;  $\delta_H$  (200 MHz; DMSO-*d*<sup>6</sup>) 3.45 (1 H, t, *J* 2.6, Alkyne-CH), 4.47 (2 H, d, *J* 2.6, CH<sub>2</sub>), 5.73 (1 H, d, *J* 7.3, C-5H), 7.16 (2 H, d, NH<sub>2</sub>), 7.64 (1 H, d, *J* 7.5, C-6H);  $\delta_C$  (50 MHz; DMSO-*d*<sup>6</sup>) 37.3 (CH<sub>2</sub>), 75.3 (C), 79.4 (C), 94.1 (CH), 144.8 (CH), 155.1 (C), 165.9 (C); *m/z* (EI) 149 (M<sup>+</sup>, 52%), 112 (M-(Propargyl group plus N and NH<sub>2</sub> being protonated)<sup>+</sup>, 94%), 77 (M-(NH<sub>2</sub> and Propargyl group)<sup>+</sup>, 100%); [Found (EI) 149.0584, C<sub>7</sub>H<sub>7</sub>N<sub>3</sub>O requires 149.0584]; The X-Ray crystal structure (Appendix E) was acquired from the crystals obtained from Et<sub>2</sub>O and MeOH.



C<sub>7</sub>H<sub>7</sub>N<sub>3</sub>O

#### Preparation of 4-(benzoylamino)-1-(prop-2-ynyl)pyrimidin-2(1H)-one (**166**)

**174** (0.26 g, 1.73 mmol, 1.00 eq) was dissolved in dry C<sub>5</sub>H<sub>5</sub>N (20 mL) and cooled to 0 °C. To this, BzCl (0.30 mL, 2.59 mmol, 1.50 eq) was added and the reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. The solvent was removed *in vacuo* and the residue was washed with hot EtOH (150 mL). The solvent was removed *in vacuo* and diluted with H<sub>2</sub>O (30 mL). The aqueous phase was extracted with EtOAc (3 × 40 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over

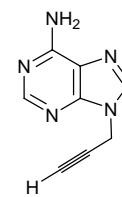


C<sub>14</sub>H<sub>11</sub>N<sub>3</sub>O<sub>2</sub>

anhydrous MgSO<sub>4</sub> and filtered. Concentration *in vacuo* afforded the crude product as a brown oil. The crude product was purified by flash chromatography on a silica column eluting with light petroleum ether:EtOAc (50:50). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a pale yellow solid. (**166**, 0.13 g, 0.51 mmol, 84%):  $R_f = 0.08$  (light petroleum ether:EtOAc, 50:50);  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3268 m (NH), 2121 w (Alkyne), 1719 s (C=O), 1697 s (C=O), 1657 s;  $\delta_H$  (400 MHz; DMSO-*d*<sup>6</sup>) 3.49 (1 H, t,  $J$  2.4, Alkyne-CH), 4.69 (2 H, d,  $J$  2.5, Alkyne-CH<sub>2</sub>), 7.35 (1 H, d,  $J$  7.1, C-5H), 7.46-7.67 (3 H, m, Ar-H), 7.98-8.02 (2 H, m, Ar-H), 8.23 (1 H, d,  $J$  7.5, CH-6), 11.21 (1H, br s, NH);  $\delta_c$  (100 MHz; DMSO-*d*<sup>6</sup>) 38.5 (CH<sub>2</sub>), 76.4 (C), 78.2 (C), 96.5 (CH), 128.4 (C), 128.7 (2 × CH), 128.8 (CH), 129.1 (C), 132.7 (CH), 133.1 (C), 133.4 (CH), 149.1 (C), 165.6 (C);  $m/z$  (HNESP) 254 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 254.0924, C<sub>14</sub>H<sub>15</sub>N<sub>4</sub>O<sub>2</sub> requires 254.0926]; The X-Ray crystal structure (Appendix F) was acquired from the crystals obtained from Et<sub>2</sub>O and MeOH.

#### Preparation of *N*2-amino-*N*9-prop-2-ynyl adenine (**175**)

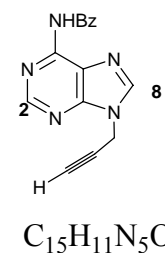
Adenine (2.00 g, 14.80 mmol, 1.00 eq) was taken up in dry DMF (30 mL) and stirred at rt under N<sub>2</sub> in the presence of NaH (60% in mineral oil, 0.49 g, 14.80 mmol, 1.00 eq) for 4 h. To this, propargyl bromide (80% in toluene, 1.58 mL, 14.80 mmol, 1.00 eq) was added dropwise and the reaction mixture was maintained for 24 h. Solvent was removed *in vacuo* and H<sub>2</sub>O (20 mL) was added. The aqueous phase was extracted with EtOAc (5 × 50 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL), dried over anhydrous MgSO<sub>4</sub> and filtered. Concentration *in vacuo* afforded the title compound as an off white solid. (**175**, 0.93 g, 5.34 mmol, 36%):  $R_f = 0.14$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3584 m, 3148 s, 2111 w (Alkyne), 1420 w, 772 w;  $\delta_H$  (400 MHz; DMSO-*d*<sup>6</sup>) 3.53 (1 H, t,  $J$  2.4, Alkyne-CH), 5.16–5.18 (2 H, d,  $J$  2.4, Alkyne-CH<sub>2</sub>), 7.29 (2 H, br s, NH<sub>2</sub>), 7.77 (1 H, s, CH), 8.42 (1 H, s, CH);  $\delta_c$  (100 MHz; DMSO-*d*<sup>6</sup>) 32.1 (CH<sub>2</sub>), 75.7 (C), 78.2 (C), 118.4 (C), 139.9 (CH), 148.9 (C), 152.6 (CH), 155.9 (C);  $m/z$  (HNES) 174 (M+NH<sub>4</sub><sup>+</sup>, 100%); (Found (HNES) 174.0771, C<sub>8</sub>H<sub>7</sub>N<sub>5</sub> requires 174.0774); The X-Ray crystal structure (Appendix G) was acquired from the crystals obtained from diethyl ether and MeOH. The identity of the product was confirmed by NOESY and COSY spectroscopy.





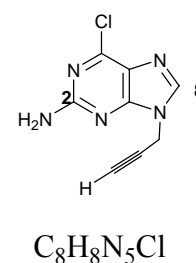
### Preparation of *N*-benzoylamino-*N*-prop-2-ynyl-adenine (**176**)

**175** (0.05 g, 0.22 mmol, 1.00 eq) was taken up in dry C<sub>5</sub>H<sub>5</sub>N (10 mL) and cooled to 0 °C. To this, BzCl (0.03 mL, 0.33 mmol, 1.50 eq) was added and the mixture was warmed to rt and stirred under N<sub>2</sub> for 4 h. Solvent was removed *in vacuo* and the residue was dissolved in DCM (20 mL). The residue was washed with an aqueous NaHCO<sub>3</sub> solution (5%, 20 mL) and the organic phase was dried over anhydrous MgSO<sub>4</sub>. Filtration followed by concentration *in vacuo* afforded the crude product as a brown/orange solid. The residue was purified by flash chromatography on a silica column, eluting with MeOH:DCM (3:97). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white solid. (**176**, 0.04 g, 0.14 mmol, 67%): *R*<sub>f</sub> = 0.24 (MeOH:DCM, 3:97); m.p. 138-140 °C (turned brown);  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3407 m, 3207 s, 3054 m, 2982 m, 2125 w (Alkyne), 1691 s (C=O), 1615 s, 1583 s, 1408 s, 1248 s, 1217 s, 946 w, 752 m;  $\delta_{\text{H}}$  (400 MHz; DMSO-*d*<sup>6</sup>) 2.67 (1 H, t, *J* 2.6, Alkyne-CH), 5.33 (2 H, d, *J* 2.6, Alkyne-CH<sub>2</sub>), 7.44-7.56 (3 H, m, Ar-*H*), 7.96-8.12 (2 H, m, Ar-*H*), 8.16 (1 H, s, C-8*H*), 8.71 (1 H, s, C-2*H*), 9.51 (1 H, br s, NH);  $\delta_{\text{C}}$  (100 MHz; DMSO-*d*<sup>6</sup>) 33.3 (CH<sub>2</sub>), 75.4 (C), 78.4 (C), 122.9 (C), 127.9 (2 × CH), 128.7 (2 × CH), 132.7 (C), 132.9 (CH), 142.1 (CH), 149.71 (C), 151.6 (CH), 152.6 (C), 170.4 (C); *m/z* (HNES) 278 (M+NH<sub>4</sub><sup>+</sup>, 100%); (Found (HNES) 278.1033, C<sub>15</sub>H<sub>15</sub>N<sub>6</sub>O requires 278.1036); The X-Ray crystal structure (Appendix H) was acquired from the crystals obtained from Et<sub>2</sub>O and MeOH.



### Preparation of 9-(prop-2-ynyl)-6-chloro-2-amino-9*H*-purine (**180**)

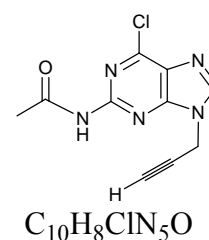
To a suspension of 2-amino-6-chloropurine **179** (0.50 g, 5.89 mmol, 1.00 eq) in dry DMF (30 mL) was added anhydrous K<sub>2</sub>CO<sub>3</sub> (0.41 g, 5.86 mmol, 0.99 eq). The suspension was stirred at rt under N<sub>2</sub> for 2 h. To this, propargyl bromide (80% in toluene, 0.63 mL, 5.89 mmol, 1.00 eq) was added and the reaction mixture was left for 24 h. Solvent was removed *in vacuo* to afford the crude product as an orange/yellow solid. The residue was purified by flash chromatography on a silica column eluting with MeOH:DCM (2:98). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a pale yellow solid. (**180**, 0.47 g, 2.26 mmol, 77%): *R*<sub>f</sub> = 0.17 (MeOH:DCM, 2:98); m.p. 237-238 °C;  $\nu_{\max}$  (KBr)/cm<sup>-1</sup> 3457 s (NH<sub>2</sub>), 3392 w, 3309 s



(NH<sub>2</sub>), 3203 s (NH<sub>2</sub>), 2127 w (Alkyne), 1404 m, 1139 m, 1000 m; δ<sub>H</sub> (400 MHz; DMSO-*d*<sup>6</sup>) 3.48 (1 H, t, *J* 2.5, Alkyne-*H*), 4.93 (2 H, d, *J* 2.5, Alkyne-CH<sub>2</sub>), 7.02 (2 H, br s, NH<sub>2</sub>), 8.18 (1 H, s, C-8*H*); δ<sub>c</sub> (100 MHz; DMSO-*d*<sup>6</sup>) 32.4 (CH<sub>2</sub>), 76.1 (C), 77.8 (C), 123.1 (C), 142.3 (CH), 149.5 (CH), 153.6 (C), 159.9 (C); *m/z* (HNES) 208 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNES) 278.1033, C<sub>8</sub>H<sub>8</sub>ClN<sub>5</sub> requires 208.0384]; (Found: C, 46.03; H, 2.72, N, 33.46. C<sub>8</sub>H<sub>8</sub>ClN<sub>5</sub> requires C, 46.28; H, 2.91; 33.73%).

### Preparation of 9-(prop-2-ynyl)-6-chloro-2-acetylamino-9*H*-purine (190)

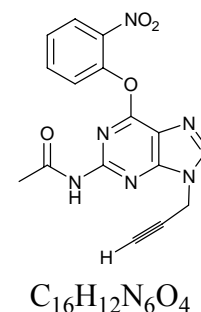
**180** (0.13 g, 0.62 mmol, 1.00 eq) was taken up in acetic anhydride (Ac<sub>2</sub>O) (7 mL) and was heated to reflux for 3 h. Solvent was removed *in vacuo* to afford the crude product as a pale yellow solid. The residue was purified by flash chromatography on a silica column, eluting with MeOH:DCM (3:97). Fractions containing product were



combined and concentrated *in vacuo* to afford the title compound as a pale yellow solid. (**180**, 0.12 g, 0.48 mmol, 82%): *R*<sub>f</sub> = 0.16 (MeOH:DCM, 3:97); m.p. 156-157 °C; ν<sub>max</sub> (KBr)/cm<sup>-1</sup> 3302 m (NH), 2123 w (Alkyne), 1677 s (C=O); δ<sub>H</sub> (200 MHz; DMSO-*d*<sup>6</sup>) 2.21 (3 H, s, CH<sub>3</sub>), 3.56 (1 H, t, *J* 3.2, Alkyne-CH), 5.08 (2 H, d, *J* 3.2, Alkyne-CH<sub>2</sub>), 8.58 (1 H, s, CH), 10.90 (1 H, br s, NH); δ<sub>c</sub> (50 MHz; DMSO-*d*<sup>6</sup>) 21.9 (CH<sub>3</sub>), 32.8 (CH<sub>2</sub>), 76.5 (C), 77.4 (C), 122.8 (C), 142.8 (CH), 148.3 (CH), 154.2 (C), 159.7 (C), 170.5 (C); *m/z* (EI) 249 (M<sup>+</sup>, 16%), 209 (M-(Propargyl group)<sup>+</sup>, 30%), 207 (M-(Acetyl)<sup>+</sup>, 100%); [Found (EI) 249.0411, C<sub>10</sub>H<sub>10</sub>ClN<sub>5</sub>O requires 249.0412].

### Preparation of 9-(prop-2-ynyl)-6-(2-nitrophenyl)-2-acetylamino-9*H*-purine (191)

To a stirred solution of **190** (0.23 g, 0.99 mmol, 1.00 eq) in dry DCM (20 mL) was added TEA (0.4 mL, 2.90 mmol, 3.00 eq), 2-nitrophenol (0.40 g, 2.90 mmol, 3.00 eq) and 1,4-diazabicyclo-(2,2,2)-octane (DABCO) (0.11 g, 0.97 mmol, 1.00 eq). The reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. DCM (20 mL) was added to dilute the reaction mixture and the organic phase was washed with saturated

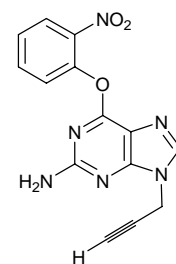


NaHCO<sub>3</sub> solution (60 mL). The aqueous phase was re-extracted with DCM (4 × 60 mL) and the combined organic layers were dried over anhydrous MgSO<sub>4</sub>. Filtration followed by evaporation afforded a clear, yellow oil. The residue was purified by flash chromatography on a silica column, eluting with EtOAc. Fractions containing product

were combined and concentrated *in vacuo* to afford the title compound as a white foam. (**191**, 0.25 g, 0.71 mmol, 89%):  $R_f = 0.58$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (KBr)/ $\text{cm}^{-1}$  3385 m (NH), 3237 m (NH), 2122 w (Alkyne), 1677 m, 1624 s (C=O), 1408 m, 1233 s;  $\delta_{\text{H}}$  (400 MHz;  $\text{CDCl}_3$ ) 1.29 (3 H, s,  $\text{CH}_3$ ), 1.67 (1 H, t,  $J$  2.5, 5.4, Alkyne- $H$ ), 4.07 (2 H, d,  $J$  2.4, Alkyne- $\text{CH}_2$ ), 6.52 (2 H, m, Ar- $H$ ), 6.82 (1 H, s,  $\text{CH}$ ), 7.23 (2 H, d,  $J$  10.8, Ar- $H$ ), 7.36 (1 H, br s,  $\text{NH}$ );  $\delta_{\text{C}}$  (100 MHz;  $\text{CDCl}_3$ ) 24.6 ( $\text{CH}_3$ ), 33.3 ( $\text{CH}_2$ ), 60.3 ( $\text{CH}$ ), 75.3 (C), 75.4 (C), 117.3 (C), 125.4 (CH), 125.7 (CH), 126.7 (CH), 134.8 (CH), 142.0 (C-8), 145.3 (CH), 151.8 (C), 153.8 (C), 159.1 (C), 171.1 (C);  $m/z$  (HNESP) 353 ( $\text{M}+\text{NH}_4^+$ , 100%); [Found (HNESP) 353.0995,  $\text{C}_{16}\text{H}_{16}\text{N}_7\text{O}_4$  requires 353.0993].

### Attempted preparation of 9-(prop-2-ynyl)-6-(2-nitrophenyl)-2-amino-9H-purine (**192**)

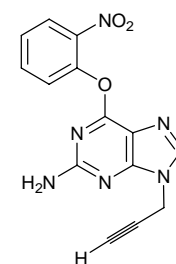
**191** (0.16 g, 0.45 mmol, 1.00 eq) was stirred in a methanolic ammonia solution (5% v/v, 20 mL) at rt under  $\text{N}_2$  for 48 h. Solvent was removed *in vacuo* to afford a yellow/brown solid. A further aliquot of methanolic ammonia solution (5% v/v, 20 mL) was added and the reaction mixture was stirred under the same conditions for 24 h. Solvent was removed *in vacuo*.  $^1\text{H}$  NMR confirmed that the reaction had failed and the starting material was the sole compound present.



$\text{C}_{14}\text{H}_{10}\text{N}_6\text{O}_3$

### Preparation of 9-(prop-2-ynyl)-6-(2-nitrophenyl)-2-amino-9H-purine (**192**)

To a stirred solution of **180** (0.28 g, 1.35 mmol, 1.00 eq) in dry DCM (10 mL) was added DABCO (0.15 g, 1.35 mmol, 1.00 eq), TEA (0.56 mL, 4.05 mmol, 3.00 eq) and 2-nitrophenol (0.56 g, 4.05 mmol, 3.00 eq). The reaction mixture was stirred at rt under  $\text{N}_2$  for 24 h. DCM (20 mL) was added to dilute the mixture and the organic phase was washed with a saturated  $\text{NaHCO}_3$  solution (20 mL). The aqueous phase was re-extracted with DCM ( $3 \times 20$  mL) and the combined organic layers were



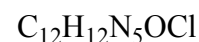
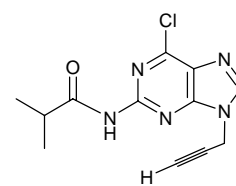
$\text{C}_{14}\text{H}_{10}\text{N}_6\text{O}_3$

dried over anhydrous  $\text{MgSO}_4$ . Filtration followed by evaporation *in vacuo*, afforded an orange oil. The crude product was purified by flash chromatography on a silica column, eluting with EtOAc. Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a pale yellow solid. (**192**, 0.18 g, 0.58 mmol, 43%):  $R_f = 0.39$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$  3293 m ( $\text{NH}_2$ ), 2124 w

(Alkyne), 1626 m, 1576 m, 1527 m, 1407 m, 1238 m;  $\delta_{\text{H}}$  (400 MHz; DMSO- $d^6$ ) 3.47 (1 H, t,  $J$  2.4, Alkyne-CH), 4.93 (2 H, d,  $J$  2.5, Alkyne-CH<sub>2</sub>), 6.56 (2 H, br s, NH<sub>2</sub>), 7.49-7.62 (2 H, m, Ar-H), 7.80-7.89 (2 H, m, Ar-H), 8.06 (1 H, s, CH);  $\delta_{\text{C}}$  (100 MHz; DMSO- $d^6$ ) 32.9 (CH<sub>2</sub>), 76.5 (C), 78.9 (C), 126.1 (CH), 126.3 (CH), 127.3 (C), 135.4 (CH), 136.1 (C), 142.3 (CH), 142.7 (C-8), 145.6 (C), 155.7 (C), 159 (C), 160.2 (C);  $m/z$  (HNESP) 311 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 311.0887, C<sub>14</sub>H<sub>14</sub>N<sub>7</sub>O<sub>3</sub> requires 311.0890].

### Preparation of 9-(prop-2-ynyl)-6-chloro-2-isobutyrylamino-9H-purine (198)

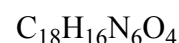
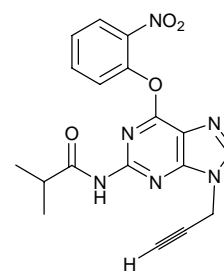
**180** (0.27 g, 1.30 mmol, 1.00 eq) was dissolved in dry C<sub>5</sub>H<sub>5</sub>N (20 mL) and to this, DMAP (0.02 g, 0.13 mmol, 0.10 eq) and isobutyryl chloride (0.24 g, 2.30 mmol, 1.77 eq) was added. The reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. MeOH (20 mL) was added to quench the reaction and the solvent was removed *in*



*vacuo*. The residue was dissolved in DCM (20 mL) and the solution was washed with aqueous NaHCO<sub>3</sub> (1 mol dm<sup>-3</sup>, 20 mL). The organic phase was separated, dried over anhydrous MgSO<sub>4</sub> and concentrated *in vacuo* to afford a red oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc (50:50). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a white oil. (**198**, 0.31 g, 1.13 mmol, 88%):  $R_{\text{f}}$  = 0.45 (light petroleum:EtOAc, 50:50);  $\nu_{\text{max}}$  (neat)/cm<sup>-1</sup> 3302 m, 3054 m, 2981 m, 2306 w (Alkyne), 1722 m (C=O), 1402 m, 1134 m, 1001 w;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) 1.29 (6 H, d,  $J$  6.7, 2 × CH<sub>3</sub>), 2.56 (1 H, t,  $J$  2.6, Alkyne-H), 2.91 (1 H, sep,  $J$  6.4, 13.2, CH), 5.02 (2 H, d,  $J$  2.4, Alkyne-CH<sub>2</sub>), 8.13 (1 H, br s, NH), 8.22 (1 H, s, CH);  $\delta_{\text{C}}$  (100 MHz; CDCl<sub>3</sub>) 19.2 (2 × CH<sub>3</sub>), 33.6 (CH<sub>2</sub>), 36.2 (CH), 75.1 (C), 75.9 (C), 128.0 (C), 143.6 (C-8), 151.3 (C), 152.1 (C), 152.3 (C), 175.5 (C);  $m/z$  (HNESP) 278 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 278.0803, C<sub>12</sub>H<sub>12</sub>ClN<sub>5</sub>O requires 278.0808].

### Preparation of 9-(prop-2-ynyl)-6-(2-nitrophenyl)-2-isobutyrylamino-9H-purine (182)

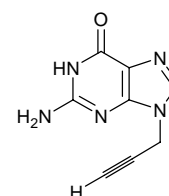
To a stirred solution of **198** (0.09 g, 0.33 mmol, 1.00 eq) in dry DCM (20 mL) was added, TEA (0.14 mL, 0.975 mmol, 3 eq), 2-nitrophenol (0.14 g, 0.98 mmol, 3.00 eq) and DABCO (0.04 g, 0.33 mmol, 1.00 eq). The reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. DCM (20 mL) was added to dilute the reaction mixture and the organic phase was washed with saturated NaHCO<sub>3</sub> solution (30 mL). The aqueous phase was re-extracted with DCM (4 × 40 mL)



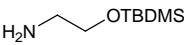
and the combined organic layers were dried over anhydrous MgSO<sub>4</sub>. Filtration followed by evaporation afforded the crude product as a clear, yellow oil. The residue was purified by flash chromatography on a silica column, eluting with EtOAc. Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a yellow oil. (**182**, 0.11 g, 0.29 mmol, 89%):  $R_f = 0.21$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3465 w (NH), 2985 m, 2122 w (Alkyne), 1740 s (C=O), 1623 w, 1374 m, 1240 s;  $\delta_H$  (400 MHz; CDCl<sub>3</sub>) 1.06 (6 H, d,  $J$  6.7, 2 × CH<sub>3</sub>), 2.55 (1 H, t,  $J$  2.6, Alkyne-CH), 3.12 (1H, sep, CH), 4.98 (2 H, d,  $J$  2.6, Alkyne-CH<sub>2</sub>), 7.42-7.51 (2 H, m, Ar-H), 7.71-7.75 (2 H, m, Ar-H), 8.10 (1 H, br s, NH), 8.15 (1 H, s, CH);  $\delta_c$  (100 MHz; CDCl<sub>3</sub>) 18.9 (2 × CH<sub>3</sub>), 33.5 (CH<sub>2</sub>), 34.8 (CH), 75.4 (C), 75.6 (C), 125.9 (CH), 126.7 (CH), 126.7 (C), 128.4 (CH), 130.1 (CH), 134.9 (C), 142.1 (C-8), 145.5 (C), 151.8 (C), 159.2 (C), 169.9 (C), 175.9 (C);  $m/z$  (HNESP) 381 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 381.1306, C<sub>18</sub>H<sub>20</sub>N<sub>7</sub>O<sub>4</sub> requires 381.1308].

### Attempted preparation of 9-(prop-2-ynyl)- 2-amino-9H-guanine (183)

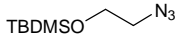
A solution of 1,1,3,3-tetramethylguanidine (TMG) **200** (0.18 mL, 1.41 mmol, 9.00 eq) in dry CH<sub>3</sub>CN (5 mL) was added to a stirred solution of **182** (0.06 g, 0.16 mmol, 1.00 eq) and 2-nitrobenzaloxime **199** (0.26 g, 1.58 mmol, 10.00 eq) in dry CH<sub>3</sub>CN (5 mL). The reaction mixture was stirred at rt under N<sub>2</sub> for 48 h. Solvent was removed *in vacuo*. <sup>1</sup>H NMR confirmed that the reaction had failed and the starting material, **182**, was the sole compound present.



### Preparation of 2-(TBDMS)ethanamine (237)

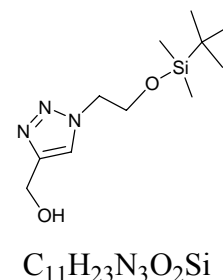
Ethanolamine (1.00 g, 16.37 mmol, 1.00 eq) was taken up in dry  DCM (20 mL) and to this, imidazole (1.45 g, 24.56 mmol, 1.30 eq)  $C_8H_{21}NOSi$  and TBDMSCl (3.70 g, 24.56 mmol, 1.50 eq) was added. The reaction mixture was stirred at rt under  $N_2$  for 24 h.  $H_2O$  (30 mL) was added to quench the reaction and the aqueous phase was separated and extracted with EtOAc ( $4 \times 40$  mL). The combined organic layers were washed with saturated NaCl solution (30 mL), dried over  $MgSO_4$ , filtered and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (237, 1.53 g, 8.69 mmol, 53%):  $R_f = 0.1$  (light petroleum:EtOAc, 50:50);  $\nu_{max}$  (neat)/ $cm^{-1}$  2942 s ( $NH_2$ ), 1736 m, 1590 m;  $\delta_H$  (200 MHz;  $CDCl_3$ ) 0.31 (9 H, s,  $(CH_3)_3C$ ), 1.14 (6 H, s,  $2 \times CH_3$ ), 3.08 (2 H, t,  $J$  5.4,  $CH_2$ ), 3.95 (2 H, t,  $J$  5.4, 10.8,  $CH_2$ ), 5.92 (2 H, br s,  $NH_2$ );  $\delta_c$  (50 MHz;  $CDCl_3$ ) 18.4 ( $2 \times CH_3$ ), 26.0 ( $3 \times CH_3$ ), 43.5 ( $CH_2$ ), 63.6 ( $CH_2$ ), 121.7 (C);  $m/z$  (HNESP) 176 ( $M+NH_4^+$ , 100%), 351 ( $M+(M)^+$ , 73%); (Found (HNESP) 176.1462,  $C_8H_{25}N_2OSi$  requires 176.1465).

### Preparation of (2-azido)-1-*o*-(TBDMS)ethane (238)

237 (1.53 g, 8.75 mmol, 1.00 eq) was added to imidazole-1-sulfonyl  azide HCl salt 152 (1.67 g, 9.62 mmol, 1.10 eq),  $K_2CO_3$  (1.21 g, 8.75 mmol, 1.00 eq) and  $CuSO_4 \cdot 5H_2O$  (0.02 g, 0.09 mmol, 0.01 eq) in  $MeOH$  (20 mL) and the reaction mixture was stirred at rt under  $N_2$  for 24 h. The mixture was reduced in volume by concentration *in vacuo* and then diluted with  $H_2O$  (20 mL). The solution was extracted with EtOAc ( $4 \times 20$  mL) and the combined organic layers were washed with saturated NaCl solution (20 mL), dried over anhydrous  $MgSO_4$ , filtered and concentrated *in vacuo*. The pale blue solid was used immediately in the subsequent step.  $\nu_{max}$  (neat)/ $cm^{-1}$  2114 m ( $N_3$ ) (solvent peaks on spectrum);  $\delta_H$  (200 MHz;  $CDCl_3$ ) 0.09 (6H, s,  $2 \times CH_3$ ), 0.91 (9H, s,  $3 \times CH_3$ ), 3.27 (2H, t,  $J$  5.0,  $CH_2$ ), 3.80 (2H, t,  $J$  5.4,  $CH_2$ ), (solvent peaks present on spectrum);  $\delta_c$  (50 MHz;  $CDCl_3$ ) -5.50 ( $2 \times CH_3$ ), 25.7 ( $3 \times CH_3$ ), 53.2 ( $CH_2$ ), 60.4 (C), 62.6 ( $CH_2$ ) (solvent peaks present on spectrum);  $m/z$  (ASAP) 202 ( $M+H^+$ , 100%); [Found (ASAP) 202.1368,  $C_8H_{20}N_3OSi$  requires 202.1370].

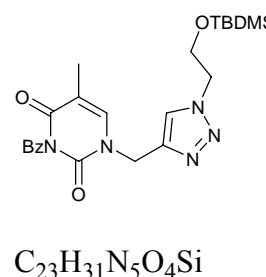
### Preparation of (1-O-(TBDMS ethane)-1H-1,2,3-triazol-4-yl)methanol (**239**)

**238** (0.57 g, 2.85 mmol, 1.00 eq) was taken up in a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and to this, sodium ascorbate (0.57 mL, 0.57 mmol, 0.20 eq), CuSO<sub>4</sub>·5H<sub>2</sub>O (0.01 g, 0.26 mmol, 0.09 eq) and propargyl alcohol (0.18 mL, 3.11 mmol, 1.09 eq) were added. The reaction mixture was stirred under N<sub>2</sub> at rt for 24 h. H<sub>2</sub>O (10 mL) was added to quench the reaction and the aqueous phase was extracted with EtOAc (4 × 20 mL). The combined organic layers were washed with saturated NaCl solution (20 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford the title compound as a yellow/brown oil. (**239**, 0.48 g, 1.99 mmol, 70%) *R*<sub>f</sub> = 0.42 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 3269 w (OH), 1471 w, 1361 m, 1182 m, 1111 m;  $\delta_{\text{H}}$  (200 MHz; CDCl<sub>3</sub>) -0.139 (6H, s, 2 × CH<sub>3</sub>), 0.737 (9H, s, 3 × CH<sub>3</sub>), 3.12 (1H, br s, OH), 3.87 (2H, t, *J* 5.0, CH<sub>2</sub>), 4.36 (2H, t, *J* 5.0, CH<sub>2</sub>), 4.65 (2H, s, CH<sub>2</sub>), 7.61 (1H, s, triazole-*H*);  $\delta_{\text{C}}$  (50 MHz; CDCl<sub>3</sub>) -5.9 (2 × CH<sub>3</sub>), 25.4 (3 × CH<sub>3</sub>), 52.5 (CH<sub>3</sub>), 56.1 (CH<sub>2</sub>), 61.7 (CH<sub>2</sub>), 63.3 (CH<sub>2</sub>), 64.0 (CH), 123.1 (CH); *m/z* (HNESP) 258 (M+H<sup>+</sup>, 100%); [Found (HNESP) 258.1630, C<sub>11</sub>H<sub>24</sub>N<sub>3</sub>O<sub>2</sub>Si requires 258.1632].



### Preparation of 3-benzoyl-1-((1-(2-(*tert*-butyldimethylsilyloxy)ethyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrimidine-2,4(1H,3H)-dione (**240**)

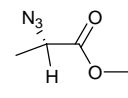
**238** (1.76 g, 8.74 mmol, 1.00 eq) in EtOAc (6 mL) was added to a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and to this, **161** (1.85 g, 8.02 mmol, 1.09 eq), CuSO<sub>4</sub>·5H<sub>2</sub>O (0.19 g, 0.79 mmol, 0.09 eq) and sodium ascorbate (2.19 mL, 1.75 mmol, 0.20 eq) were added. The deep yellow solution was stirred for 24 h at rt. The solvent was removed *in vacuo* and H<sub>2</sub>O (20 mL) was added. The solution was extracted with EtOAc (4 × 30 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc (50:50). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, yellow oil. (**240**, 0.94 g, 2.06 mmol, 24%); *R*<sub>f</sub> = 0.15 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 2929m,



1748 s (C=O), 1699 s (C=O), 1658 s (C=O), 1461 m, 1440 s;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) - 0.28 (6 H, s, 2 × CH<sub>3</sub>), 0.59 (9 H, s, 3 × CH<sub>3</sub>), 1.71 (3 H, d, *J* 1.2, CH<sub>3</sub>), 3.75 (2 H, t, *J* 4.7, CH<sub>2</sub>), 4.23 (2 H, t, *J* 5.2, CH<sub>2</sub>), 4.78 (2 H, s, CH<sub>2</sub>), 7.24-7.29 (3 H, m, Ar-*H*), 7.39-7.44 (1 H, m, Ar-*H*), 7.52 (1 H, s, Triazole-CH), 7.67-7.71 (2 H, m, Ar-*H*);  $\delta_{\text{C}}$  (100 MHz; CDCl<sub>3</sub>) -5.9 (2 × CH<sub>3</sub>), 12.2 (CH<sub>3</sub>), 17.9 (CH), 25.5 (3 × CH<sub>3</sub>), 42.8 (CH<sub>2</sub>), 52.7 (CH<sub>2</sub>), 61.6 (CH<sub>2</sub>), 110.8 (C), 124.7 (Triazole-CH), 130.2 (2 × CH), 130.3 (2 × CH), 131.4 (C), 134.8 (CH), 139.9 (CH), 141.2 (C), 149.6 (C=O), 162.9 (C=O), 168.8 (C=O); *m/z* (HNESP) 470 (M+NH<sub>4</sub><sup>+</sup>, 100%), 939 (M+(M<sup>+</sup>), 44%); [Found (HNESP) 470.2208, C<sub>23</sub>H<sub>35</sub>N<sub>6</sub>O<sub>4</sub>Si requires 470.2218]; The identity of the product was confirmed by NOESY spectroscopy.

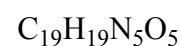
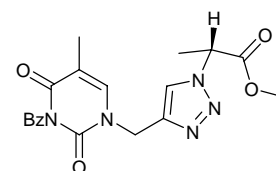
### Preparation of (*R*)-methyl 2-azidopropanoate (**242**)

L-Alanine methyl ester HCl salt (1.00 g, 7.16 mmol, 1.00 eq) was added to a mixture of imidazole-1-sulfonyl azide HCl salt **152** (1.36 g, 7.88 mmol, 1.10 eq), K<sub>2</sub>CO<sub>3</sub> (0.99 g, 7.16 mmol, 1.00 eq) and CuSO<sub>4</sub>·5H<sub>2</sub>O (0.02 g, 0.07 mmol, 0.01 eq) and the reaction mixture was stirred for 24 h at rt under N<sub>2</sub>. The mixture was reduced in volume by concentration *in vacuo* and then diluted with H<sub>2</sub>O (20 mL). The solution was extracted with EtOAc (4 × 20 mL) and the combined organic layers were washed with saturated NaCl solution (20 mL). The organic phase was dried over anhydrous MgSO<sub>4</sub>, filtered and reduced in volume by concentration *in vacuo*. The residue was used immediately in the subsequent step;  $\nu_{\text{max}}$  (neat)/cm<sup>-1</sup> 2106 s (N<sub>3</sub>), 1738 s (C=O) (solvent peaks on spectrum);  $\delta_{\text{H}}$  (200 MHz; CDCl<sub>3</sub>) 1.69 (3H, s, CH<sub>3</sub>), 3.98 (3H, s, OCH<sub>3</sub>), 4.07 (1H, s, CH) (solvent/imidazole peaks present on spectrum);  $\delta_{\text{C}}$  (50 MHz; CDCl<sub>3</sub>) 16.4 (CH<sub>3</sub>), 52.3 (CH), 56.7 (CH<sub>3</sub>), 151.1 (C) (solvent peaks present on spectrum); *m/z* (ASAP) 147 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (ASAP) 147.0877, C<sub>4</sub>H<sub>15</sub>N<sub>3</sub>O<sub>2</sub> requires 147.0877].



### Preparation of (*R*)-methyl 2-(4-((3-benzoyl-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)methyl)-1*H*-1,2,3-triazol-1-yl)propanoate (**243**)

**242** (0.59 g, 3.05 mmol, 1.00 eq) in EtOAc (6.5 mL) was added to a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and to this, **161** (0.89 g, 3.32 mmol, 1.09 eq),

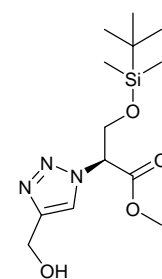




CuSO<sub>4</sub>·5H<sub>2</sub>O (0.02 g, 0.65 mmol, 0.09 eq) and sodium ascorbate (0.74 mL, 0.61 mmol, 0.20 eq) were added. The deep yellow solution was stirred for 24 h at rt. The solvent was removed *in vacuo* and H<sub>2</sub>O (10 mL) was added. The solution was extracted with EtOAc (6 × 15 mL) and the combined organic layers were washed with saturated NaCl solution (20 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was purified by flash chromatography on a silica column, eluting with light petroleum:EtOAc (80:20). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, colourless oil. (**243**, 0.18 g, 0.45 mmol, 13%): *R*<sub>f</sub> = 0.34 (light petroleum:EtOAc, 80:20);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 2922 s, 1746 s (C=O), 1697 s (C=O), 1656 s (C=O), 1461 s, 1377 s;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) 1.61 (3 H, d, *J* 7.5, CH<sub>3</sub>-CH), 1.71 (3 H, d, *J* 1.25, Thymine-CH<sub>3</sub>), 3.55 (3 H, s, OCH<sub>3</sub>), 4.79 (2 H, ab, *J* 14.9, CH<sub>2</sub>), 5.22 (1 H, q, *J* 7.5, CH<sub>3</sub>-CH), 7.22-7.30 (3 H, m, 3 × Ar-CH), 7.38-7.47 (1 H, m, Ar-CH), 7.61 (1 H, s, Triazole-CH), 7.66-7.72 (2 H, m, 1 × Ar-CH and Thymine-CH);  $\delta_{\text{C}}$  (100 MHz; CDCl<sub>3</sub>) 12.7 (CH<sub>3</sub>), 18.2 (CH<sub>3</sub>), 43.4 (CH<sub>2</sub>), 53.5 (OCH<sub>3</sub>), 58.6 (CH), 111.5 (C), 123.4 (CH), 129.4 (2 × CH), 130.7 (2 × CH), 131.8 (C), 135.3 (CH), 140.3 (CH), 142.1 (C), 150.1 (C=O), 163.3 (C=O), 169.3 (C=O), 169.7 (C=O); *m/z* (HNESP) 398 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 398.1459, C<sub>19</sub>H<sub>23</sub>N<sub>6</sub>O<sub>5</sub> requires 398.1459]; The identity of the product was confirmed by NOESY spectroscopy.

#### Preparation of (S)-methyl 3-(tert-butyldimethylsilyloxy)-2-(4-(hydroxymethyl)-1H-1,2,3-triazol-1-yl)propanoate (**244**)

**158** (0.74 g, 2.85 mmol, 1.00 eq), CuSO<sub>4</sub>·5H<sub>2</sub>O (0.06 g, 0.26 mmol, 0.09 eq), propargyl alcohol (0.18 mL, 3.11 mmol, 1.09 eq) and sodium ascorbate (0.92 mL, 0.57 mmol, 0.20 eq) were suspended in a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and stirred at rt under N<sub>2</sub> for 24 h. Solvent was removed *in vacuo* and H<sub>2</sub>O (10 mL) was added. The aqueous phase was extracted with EtOAc (6 × 20 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL). The organic phase was dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was purified by flash chromatography on a silica column eluting with (light petroleum:EtOAc (80:20) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (30:70) gradient elution). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, yellow oil. (**244**, 0.56 g, 1.76 mmol, 62%): *R*<sub>f</sub> = 0.15 (light

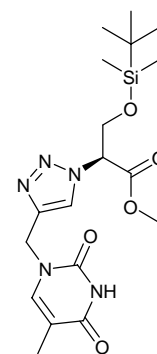


C<sub>13</sub>H<sub>25</sub>N<sub>3</sub>O<sub>4</sub>Si

petroleum:EtOAc, 30:70);  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$  3286 w (OH), 1749 s, (C=O), 1362 m, 1253 s;  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) 0.02 (6H, s,  $2 \times \text{CH}_3$ ), 0.85 (9H, s,  $3 \times \text{CH}_3$ ), 3.18 (1H, br s, OH), 3.78 (3H, s,  $\text{OCH}_3$ ), 4.02 (1H, d,  $J$  10.4,  $\text{CH}_2\text{-OTBDPS}$ ), 4.37 (1H, d,  $J$  10.4,  $\text{CH}_2\text{-OTBDPS}$ ), 4.80 (2H, s,  $\text{CH}_2\text{-OH}$ ), 5.53-5.57 (1H, m, CH), 7.92 (1H, s, triazole-H);  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) 17.7 ( $2 \times \text{CH}_3$ ), 25.2 ( $3 \times \text{CH}_3$ ), 27.8 (C), 52.7 ( $\text{CH}_3$ ), 56.0 ( $\text{CH}_2$ ), 63.3 ( $\text{CH}_2$ ), 64.0 (CH), 122.6 (CH), 147.1 (C), 167.1 (C);  $m/z$  (HNESP) 316 ( $\text{M}+\text{H}^+$ , 100%); [Found (HNESP) 316.1688,  $\text{C}_{13}\text{H}_{26}\text{N}_3\text{O}_4\text{Si}$  requires 316.1687]; (Found C, 49.46; H, 8.05; N, 13.56.  $\text{C}_{13}\text{H}_{25}\text{N}_3\text{O}_4\text{Si}$  requires C, 49.50; H, 7.99; N, 13.32%).

### Preparation of (S)-methyl 3-(tert-butyldimethylsilyloxy)-2-(4-((5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)methyl)-1H-1,2,3-triazol-1-yl)propanoate (246)

**158** (1.62 g, 6.26 mmol, 1.00 eq),  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (0.14 g, 0.56 mmol, 0.09 eq), sodium ascorbate (2.02 mL, 1.25 mmol, 0.2 eq) and **161** (1.83 g, 6.82 mmol, 1.09 eq) were suspended in a mixture of DCM/*tert*-butyl alcohol/ $\text{H}_2\text{O}$  ((1:1:1) v/v, 12 mL) and stirred at rt under  $\text{N}_2$  for 30 h. Solvent was removed *in vacuo* and  $\text{H}_2\text{O}$  (30 mL) was added. The aqueous phase was extracted with EtOAc ( $6 \times 40$  mL) and the combined organic layers were washed with saturated NaCl solution (50 mL). The organic phase was dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford a yellow/brown



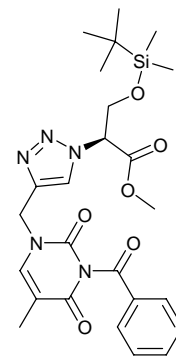
$\text{C}_{18}\text{H}_{29}\text{N}_5\text{O}_5\text{Si}$

oil. The residue was purified by flash chromatography on a silica column eluting with (light petroleum:EtOAc (75:25) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (0:100) gradient elution). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, yellow oil. (**246**, 1.17 g, 0.93 mmol, 44%):  $R_f = 0.25$  (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$  2956 w, 2858 w, 1754 w, 1674 s, 1470 w, 1126 w;  $\delta_{\text{H}}$  (200 MHz;  $\text{CDCl}_3$ ) -0.15 (6H, s,  $2 \times \text{CH}_3$ ), 0.66 (9H, s,  $3 \times \text{CH}_3$ ), 1.74 (3H, s, thymine- $\text{CH}_3$ ), 3.65 (3H, s,  $\text{OCH}_3$ ), 3.92-3.95 (1H, d,  $J$  13.3,  $\text{CH}_2\text{-OTBDMS}$ ), 4.18-4.25 (1H, d,  $J$  13.3,  $\text{CH}_2\text{-OTBDMS}$ ), 4.82-4.98 (2H, dd,  $J$  13.3,  $\text{CH}_2$ ), 5.42-5.46 (1H, m, CH), 7.26 (1H, s, thymine-CH), 7.91 (1H, s, triazole-CH), 10.08 (1H, br s, NH);  $\delta_{\text{C}}$  (50 MHz;  $\text{CDCl}_3$ ) -6.1 ( $\text{CH}_3$ ), -5.9 ( $\text{CH}_3$ ), 12.1 ( $\text{CH}_3$ ), 17.8 (C), 25.3 ( $3 \times \text{CH}_3$ ), 42.4 ( $\text{CH}_2$ ), 52.9 ( $\text{OCH}_3$ ), 63.3 ( $\text{CH}_2$ ), 64.4 (CH), 110.9 (C), 124.6 (triazole-CH), 140.0 (CH), 141.8 (C), 150.9 (C=O), 164.4 (C=O), 166.9 (C=O);  $m/z$  (HNESP) 424 ( $\text{M}+\text{H}^+$ , 100%); [Found (HNESP)

424.2007, C<sub>18</sub>H<sub>30</sub>N<sub>5</sub>O<sub>5</sub>Si requires 424.2011]. The identity of the product was confirmed by NOESY and COSY spectroscopy.

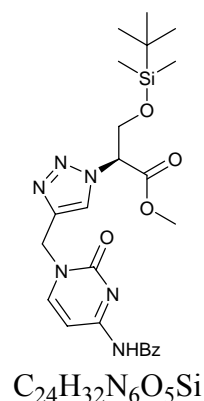
**Preparation of (2S)-methyl 2-(4-((3-benzoyl-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)methyl)-4,5-dihydro-1,2,3-triazol-1-yl)-3-(tert-butyl-dimethylsilyloxy)propanoate (245)**

**158** (0.74 g, 2.85 mmol, 1.00 eq), CuI (0.82 g, 4.28 mmol, 1.50 eq), **161** (0.77 g, 2.85 mmol, 1.00 eq) and DIPEA (4.96 mL, 28.50 mmol, 10.00 eq) were suspended in dry DMF (7 mL). The reaction mixture was stirred at rt under N<sub>2</sub> for 30 h. The dark green mixture was taken up in EtOAc (100 mL) and H<sub>2</sub>O (100 mL) and the layers were separated. The aqueous phase was extracted with EtOAc (3 × 50 mL) and the combined organic layers were washed with saturated NaCl solution (50 mL). The organic phase was dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a brown/yellow liquid. The residue was purified by flash chromatography on a silica column eluting with light petroleum:EtOAc (75:25) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (25:75) gradient elution. Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, yellow oil. (**245**, 0.14 g, 0.27 mmol, 15%): *R*<sub>f</sub> = 0.29 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 1746 s (C=O), 1698 s (C=O), 1599 s (C=O), 1435 m, 1250 s, 1224 s;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) -0.04 (6H, s, 2 × CH<sub>3</sub>), 0.77 (9H, s, 3 × CH<sub>3</sub>), 1.92 (3H, s, CH<sub>3</sub>-thymine), 3.77 (3H, s, OCH<sub>3</sub>), 3.99-4.05 (1H, dd, *J* 13.7, CH<sub>2</sub>), 4.33-4.38 (1H, d, *J* 13.7, CH<sub>2</sub>), 4.89-5.13 (2H, d, *J* 16.6, CH<sub>2</sub>), 5.55 (1H, m, CH), 7.41-7.51 (3H, m, Ar-CH), 7.57-7.66 (1H, m, thymine-CH), 7.86-7.92 (2H, m, Ar-CH), 8.00 (1H, s, triazole-CH);  $\delta_{\text{C}}$  (100 MHz; CDCl<sub>3</sub>) -5.9 (CH<sub>3</sub>), -5.7 (CH<sub>3</sub>), 12.3 (CH<sub>3</sub>), 17.9 (C), 25.4 (3 × CH<sub>3</sub>), 42.8 (CH<sub>2</sub>), 53.0 (OCH<sub>3</sub>), 63.4 (CH<sub>2</sub>), 64.6 (CH), 111.0 (C), 127.3 (C), 128.5 (C), 129.0 (2 × CH), 130.3 (2 × CH), 131.6 (CH), 134.8 (CH), 139.8 (CH), 149.7 (C=O), 163.0 (C=O), 167.0 (C=O), 168.8 (C=O); *m/z* (HNESP) 528 (M+NH<sub>4</sub><sup>+</sup>, 100%); [Found (HNESP) 528.2259, C<sub>25</sub>H<sub>33</sub>N<sub>5</sub>O<sub>6</sub>Si requires 528.2263]. The identity of the product was confirmed by NOESY and COSY spectroscopy.



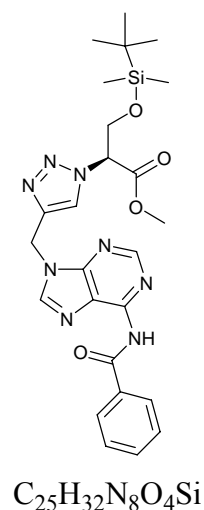
**Attempted preparation of (S)-methyl 2-(4-((4-benzamido-2-oxopyrimidin-1(2H)-yl)methyl)-1H-1,2,3-triazol-1-yl)-3-(tert-butyldimethylsilyloxy)propanoate (247)**

**158** (0.166 g, 0.641 mmol, 1.00 eq), in EtOAc (3 mL) was added to a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and to this, **166** (0.177 g, 0.699 mmol, 1.09 eq), CuSO<sub>4</sub>·5H<sub>2</sub>O (0.015 g, 0.06 mmol, 0.09 eq) and sodium ascorbate (0.21 mL, 0.013 mmol, 0.20 eq) were added. The deep yellow solution was stirred for 30 h at rt. The solvent was removed *in vacuo* and H<sub>2</sub>O (10 mL) was added. The solution was extracted with EtOAc (4 × 30 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was purified by flash chromatography on a silica column, eluting (light petroleum:EtOAc (80:20) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (20:80) gradient elution). <sup>1</sup>H NMR confirmed that none of the fractions contained the title compound **247** and the starting materials **158** and **166** were recovered.



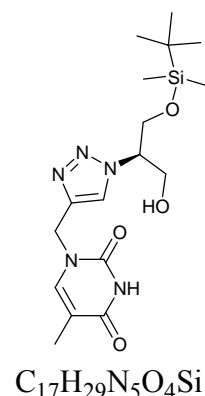
**Attempted preparation of (S)-methyl 2-(4-((6-benzamido-9H-purin-9-yl)methyl)-1H-1,2,3-triazol-1-yl)-3-(tert-butyldimethylsilyloxy)propanoate (248)**

**158** (0.59 g, 1.90 mmol, 1.00 eq), in EtOAc (3 mL) was added to a mixture of DCM/*tert*-butyl alcohol/H<sub>2</sub>O ((1:1:1) v/v, 6 mL) and to this, **176** (0.58 g, 2.07 mmol, 1.09 eq), CuSO<sub>4</sub>·5H<sub>2</sub>O (0.04 g, 0.17 mmol, 0.09 eq) and sodium ascorbate (0.74 mL, 0.38 mmol, 0.20 eq) were added. The deep yellow solution was stirred for 30 h at rt. The solvent was removed *in vacuo* and H<sub>2</sub>O (10 mL) was added. The solution was extracted with EtOAc (4 × 30 mL) and the combined organic layers were washed with saturated NaCl solution (30 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a yellow/brown oil. The residue was purified by flash chromatography on a silica column, eluting (light petroleum:EtOAc (80:20) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (20:80) gradient elution). <sup>1</sup>H NMR confirmed that none of the fractions contained the title compound **248** and the starting materials **158** and **176** were recovered.



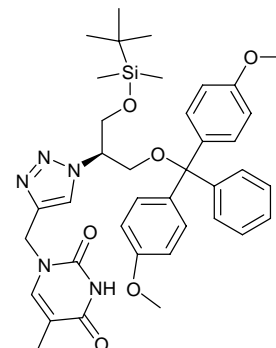
**Preparation of (*R*)-1-((1-(3-(tert-butyldimethylsilyloxy)-1-hydroxypropan-2-yl)-1*H*-1,2,3-triazol-4-yl)methyl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (249)**

**246** (0.10 g, 0.19 mmol, 1.00 eq) was added to a stirred solution of LiBH<sub>4</sub> (0.01 g, 0.38 mmol, 2.00 eq) in dry THF (10 mL) at 0 °C under N<sub>2</sub>. To this, 9-BBN (1M in hexane, 0.14 mL, 0.42 mmol, 2.20 eq) was added dropwise to the mixture. The reaction mixture was allowed to warm to rt and maintained for 24 h. The solvent was removed *in vacuo* and H<sub>2</sub>O (10 mL) was added. The aqueous phase was extracted with EtOAc (4 × 20 mL) and the combined organic layers were washed with saturated NaCl solution (20 mL). The organic phase was dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless oil. The residue was purified by flash chromatography on a silica column eluting with (light petroleum:EtOAc (70:30) to light petroleum:EtOAc (50:50) to light petroleum:EtOAc (30:70) gradient elution). Fractions containing product were combined and concentrated *in vacuo* to afford the title compound as a clear, yellow oil. (**249**, 0.02 g, 0.10 mmol, 25%): *R*<sub>f</sub> = 0.52 (light petroleum:EtOAc, 50:50);  $\nu_{\max}$  (neat)/cm<sup>-1</sup> 2985 s, 1740 s, 1373 m, 1240 s;  $\delta_{\text{H}}$  (400 MHz; CDCl<sub>3</sub>) -0.03 (6H, s, 2 × CH<sub>3</sub>), 0.81 (9H, s, 3 × CH<sub>3</sub>), 1.60 (1H, br s, OH), 1.90 (3H, s, thymine-CH<sub>3</sub>), 4.01-4.05 (2H, m, CH<sub>2</sub>-OTBDMS), 4.11-4.18 (2H, m, CH<sub>2</sub>OH), 4.61-4.68 (1H, m, CH), 4.97 (2H, s, CH<sub>2</sub>), 7.33 (1H, s, thymine-CH), 7.86 (1H, s, triazole-CH), 8.46 (1H, br s, NH);  $\delta_{\text{C}}$  (100 MHz; CDCl<sub>3</sub>) -5.8 (CH<sub>3</sub>), -5.7 (CH<sub>3</sub>), 12.3 (CH<sub>3</sub>), 18.0 (C), 25.6 (3 × CH<sub>3</sub>), 42.8 (CH<sub>2</sub>), 62.0 (CH<sub>2</sub>), 63.1 (CH<sub>2</sub>), 64.2 (CH), 111.2 (C), 124.3 (CH), 140.1 (CH), 141.6 (C), 150.6 (C), 163.8 (C); *m/z* (HNESP) 396 (M+H<sup>+</sup>, 100%); [Found (HNESP) 396.2059, C<sub>17</sub>H<sub>30</sub>N<sub>5</sub>O<sub>4</sub>Si requires 396.2062].



**Attempted preparation (R)-1-((1-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)-1-(tert-butyldimethylsilyloxy)propan-2-yl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrimidine-2,4(1H,3H)-dione (250)**

**249** (0.01 g,  $2.53 \times 10^{-3}$  mmol, 1.00 eq) and DMTrCl (0.01 g,  $3.03 \times 10^{-3}$  mmol, 1.20 eq) was suspended in dry DCM (3 mL) and to this, DBU (0.01 mL,  $3.54 \times 10^{-3}$  mmol, 1.40 eq) was added. The reaction mixture was stirred at rt under N<sub>2</sub> for 24 h. TLC confirmed no reaction had occurred and the reaction mixture was maintained for a further 24 h. H<sub>2</sub>O (3 mL) was added and the organic phase was separated and washed with H<sub>2</sub>O (2 × 6 mL), dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to afford a clear, colourless oil. <sup>1</sup>H NMR showed no product **250** was present and the starting material was the sole compound present.



**Chapter 5**  
**Appendices**

## **Appendix A**



Appendix A. The Genetic code

	<b>T</b>	<b>C</b>	<b>A</b>	<b>G</b>
<b>T</b>	TTT Phe (F) TTC Phe (F) TTA Leu (L) TTG Leu (L)	TCT Ser (S) TCC Ser (S) TCA Ser (S) TCG Ser (S)	TAT Tyr (Y) TAC <u>TAA STOP</u> <u>TAG STOP</u>	TGT Cys (C) TGC <u>TGA STOP</u> TGG Trp (W)
<b>C</b>	CTT Leu (L) CTC Leu (L) CTA Leu (L) CTG Leu (L)	CCT Pro (P) CCC Pro (P) CCA Pro (P) CCG Pro (P)	CAT His (H) CAC His (H) CAA Gln (Q) CAG Gln (Q)	CGT Arg (R) CGC Arg (R) CGA Arg (R) CGG Arg (R)
<b>A</b>	ATT Ile (I) ATC Ile (I) ATA Ile (I) ATG Met (M) START	ACT Thr (T) ACC Thr (T) ACA Thr (T) ACG Thr (T)	AAT Asn (N) AAC Asn (N) AAA Lys (K) AAG Lys (K)	AGT Ser (S) AGC Ser (S) AGA Arg (R) AGG Arg (R)
<b>G</b>	GTT Val (V) GTC Val (V) GTA Val (V) GTG Val (V)	GCT Ala (A) GCC Ala (A) GCA Ala (A) GCG Ala (A)	GAT Asp (D) GAC Asp (D) GAA Glu (E) GAG Glu (E)	GGT Gly (G) GGC Gly (G) GGA Gly (G) GGG Gly (G)

Table 11. The genetic code.

## **Appendix B**

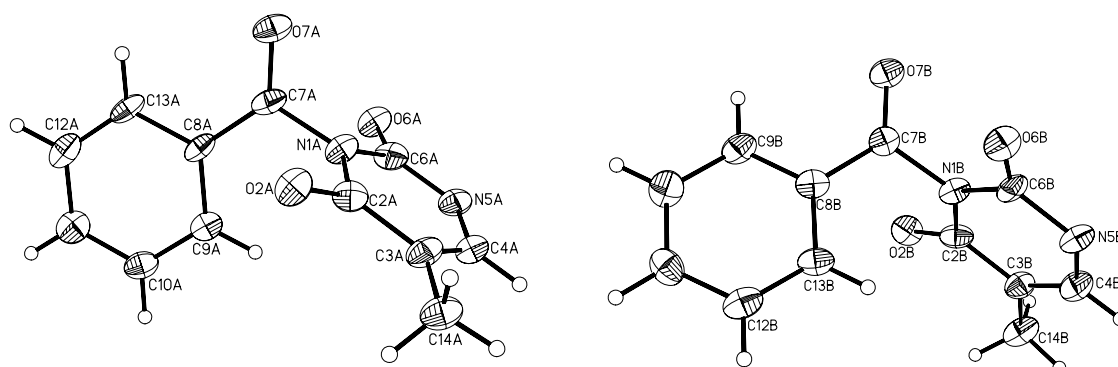
Appendix B. Table of the Amino acid structures

Amino Acids	Ala Glu Leu Ser	Arg Gln Lys Thr	Asn Gly Met Trp	Asp His Phe Tyr	Cys Ile Pro Val
Ala Arg Asn Asp Cys					
Glu Gln Gly His Ile					
Leu Lys Met Phe Pro					
Ser Thr Trp Tyr Val					

Table 12. Amino acid structure

## **Appendix C**

### Appendix C. 160 crystal structure



There are two molecules in the asymmetric unit with small differences in the orientation of the phenyl and methyl groups.

Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 13. Crystal data and structure refinement for 160.**

Identification code	x82556o	
Empirical formula	C <sub>12</sub> H <sub>14</sub> N <sub>2</sub> O <sub>3</sub>	
Formula weight	234.25	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	Pca2(1)	
Unit cell dimensions	a = 9.485(8) Å	α = 90°.
	b = 8.678(8) Å	β = 90°.
	c = 25.67(2) Å	γ = 90°.
Volume	2113(3) Å <sup>3</sup>	
Z	8	
Density (calculated)	1.473 Mg/m <sup>3</sup>	
Absorption coefficient	0.107 mm <sup>-1</sup>	
F(000)	992	
Crystal size	0.64 x 0.30 x 0.04 mm <sup>3</sup>	
Theta range for data collection	2.35 to 26.53°.	

Index ranges	-10<=h<=11, -10<=k<=10, -32<=l<=32
Reflections collected	25634
Independent reflections	2191 [R(int) = 0.1411]
Completeness to theta = 25.00°	98.3 %
Absorption correction	None
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	2191 / 1 / 309
Goodness-of-fit on F <sup>2</sup>	1.125
Final R indices [I>2sigma(I)]	R1 = 0.0777, wR2 = 0.1946
R indices (all data)	R1 = 0.1628, wR2 = 0.2665
Largest diff. peak and hole	0.410 and -0.498 e.Å <sup>-3</sup>

**Table 14. Bond lengths [Å] and angles [°] for 160.**

N(1A)-C(2A)	1.395(12)	C(12A)-C(13A)	1.386(18)
N(1A)-C(6A)	1.399(13)	C(12A)-H(12A)	0.9500
N(1A)-C(7A)	1.460(13)	C(13A)-H(13A)	0.9500
C(2A)-O(2A)	1.243(13)	C(14A)-H(14A)	0.9800
C(2A)-C(3A)	1.451(14)	C(14A)-H(14B)	0.9800
C(3A)-C(4A)	1.354(15)	C(14A)-H(14C)	0.9800
C(3A)-C(14A)	1.476(13)	N(1B)-C(6B)	1.388(13)
C(4A)-N(5A)	1.375(14)	N(1B)-C(2B)	1.420(11)
C(4A)-H(4A)	0.9500	N(1B)-C(7B)	1.477(12)
N(5A)-C(6A)	1.349(13)	C(2B)-O(2B)	1.228(13)
C(6A)-O(6A)	1.243(12)	C(2B)-C(3B)	1.443(14)
C(7A)-O(7A)	1.222(13)	C(3B)-C(4B)	1.342(16)
C(7A)-C(8A)	1.467(16)	C(3B)-C(14B)	1.514(13)
C(8A)-C(9A)	1.398(14)	C(4B)-N(5B)	1.403(13)
C(8A)-C(13A)	1.405(14)	C(4B)-H(4B)	0.9500
C(9A)-C(10A)	1.404(15)	N(5B)-C(6B)	1.368(12)
C(9A)-H(9A)	0.9500	C(6B)-O(6B)	1.229(12)
C(10A)-C(11A)	1.372(15)	C(7B)-O(7B)	1.208(12)
C(10A)-H(10A)	0.9500	C(7B)-C(8B)	1.471(15)
C(11A)-C(12A)	1.381(16)	C(8B)-C(9B)	1.392(15)
C(11A)-H(11A)	0.9500	C(8B)-C(13B)	1.408(15)

C(9B)-C(10B)	1.400(16)	C(8A)-C(9A)-H(9A)	120.3
C(9B)-H(9B)	0.9500	C(10A)-C(9A)-H(9A)	120.3
C(10B)-C(11B)	1.397(16)	C(11A)-C(10A)-C(9A)	119.7(10)
C(10B)-H(10B)	0.9500	C(11A)-C(10A)-H(10A)	120.2
C(11B)-C(12B)	1.392(16)	C(9A)-C(10A)-H(10A)	120.2
C(11B)-H(11B)	0.9500	C(10A)-C(11A)-C(12A)	121.2(11)
C(12B)-C(13B)	1.375(16)	C(10A)-C(11A)-H(11A)	119.4
C(12B)-H(12B)	0.9500	C(12A)-C(11A)-H(11A)	119.4
C(13B)-H(13B)	0.9500	C(11A)-C(12A)-C(13A)	120.5(10)
C(14B)-H(14D)	0.9800	C(11A)-C(12A)-H(12A)	119.7
C(14B)-H(14E)	0.9800	C(13A)-C(12A)-H(12A)	119.7
C(14B)-H(14F)	0.9800	C(12A)-C(13A)-C(8A)	119.0(10)
C(2A)-N(1A)-C(6A)	124.4(8)	C(12A)-C(13A)-H(13A)	120.5
C(2A)-N(1A)-C(7A)	119.2(8)	C(8A)-C(13A)-H(13A)	120.5
C(6A)-N(1A)-C(7A)	115.7(8)	C(3A)-C(14A)-H(14A)	109.5
O(2A)-C(2A)-N(1A)	117.7(9)	C(3A)-C(14A)-H(14B)	109.5
O(2A)-C(2A)-C(3A)	125.4(9)	H(14A)-C(14A)-H(14B)	109.5
N(1A)-C(2A)-C(3A)	116.9(9)	C(3A)-C(14A)-H(14C)	109.5
C(4A)-C(3A)-C(2A)	116.2(9)	H(14A)-C(14A)-H(14C)	109.5
C(4A)-C(3A)-C(14A)	124.6(9)	H(14B)-C(14A)-H(14C)	109.5
C(2A)-C(3A)-C(14A)	119.1(10)	C(6B)-N(1B)-C(2B)	127.1(8)
C(3A)-C(4A)-N(5A)	124.9(10)	C(6B)-N(1B)-C(7B)	115.7(7)
C(3A)-C(4A)-H(4A)	117.5	C(2B)-N(1B)-C(7B)	117.0(8)
N(5A)-C(4A)-H(4A)	117.5	O(2B)-C(2B)-N(1B)	118.9(9)
C(6A)-N(5A)-C(4A)	121.0(9)	O(2B)-C(2B)-C(3B)	126.3(9)
O(6A)-C(6A)-N(5A)	123.7(9)	N(1B)-C(2B)-C(3B)	114.8(9)
O(6A)-C(6A)-N(1A)	120.1(9)	C(4B)-C(3B)-C(2B)	118.4(9)
N(5A)-C(6A)-N(1A)	116.3(9)	C(4B)-C(3B)-C(14B)	123.3(9)
O(7A)-C(7A)-N(1A)	117.9(10)	C(2B)-C(3B)-C(14B)	118.2(9)
O(7A)-C(7A)-C(8A)	126.5(10)	C(3B)-C(4B)-N(5B)	123.5(10)
N(1A)-C(7A)-C(8A)	115.5(9)	C(3B)-C(4B)-H(4B)	118.2
C(9A)-C(8A)-C(13A)	120.2(10)	N(5B)-C(4B)-H(4B)	118.2
C(9A)-C(8A)-C(7A)	121.2(9)	C(6B)-N(5B)-C(4B)	122.0(9)
C(13A)-C(8A)-C(7A)	118.5(10)	O(6B)-C(6B)-N(5B)	125.5(9)
C(8A)-C(9A)-C(10A)	119.4(10)	O(6B)-C(6B)-N(1B)	120.4(9)

N(5B)-C(6B)-N(1B)	114.1(9)	C(12B)-C(11B)-H(11B)	120.0
O(7B)-C(7B)-C(8B)	125.6(9)	C(10B)-C(11B)-H(11B)	120.0
O(7B)-C(7B)-N(1B)	120.7(9)	C(13B)-C(12B)-C(11B)	121.0(11)
C(8B)-C(7B)-N(1B)	113.8(9)	C(13B)-C(12B)-H(12B)	119.5
C(9B)-C(8B)-C(13B)	119.7(10)	C(11B)-C(12B)-H(12B)	119.5
C(9B)-C(8B)-C(7B)	118.5(10)	C(12B)-C(13B)-C(8B)	119.6(10)
C(13B)-C(8B)-C(7B)	121.8(10)	C(12B)-C(13B)-H(13B)	120.2
C(8B)-C(9B)-C(10B)	120.4(10)	C(8B)-C(13B)-H(13B)	120.2
C(8B)-C(9B)-H(9B)	119.8	C(3B)-C(14B)-H(14D)	109.5
C(10B)-C(9B)-H(9B)	119.8	C(3B)-C(14B)-H(14E)	109.5
C(11B)-C(10B)-C(9B)	119.3(11)	H(14D)-C(14B)-H(14E)	109.5
C(11B)-C(10B)-H(10B)	120.4	C(3B)-C(14B)-H(14F)	109.5
C(9B)-C(10B)-H(10B)	120.4	H(14D)-C(14B)-H(14F)	109.5
C(12B)-C(11B)-C(10B)	120.0(11)	H(14E)-C(14B)-H(14F)	109.5

**Table 15. Torsion angles [°] for 160.**

C(6A)-N(1A)-C(2A)-O(2A)	-177.0(9)
C(7A)-N(1A)-C(2A)-O(2A)	-7.8(14)
C(6A)-N(1A)-C(2A)-C(3A)	4.7(14)
C(7A)-N(1A)-C(2A)-C(3A)	174.0(9)
O(2A)-C(2A)-C(3A)-C(4A)	-179.1(10)
N(1A)-C(2A)-C(3A)-C(4A)	-1.0(15)
O(2A)-C(2A)-C(3A)-C(14A)	2.6(17)
N(1A)-C(2A)-C(3A)-C(14A)	-179.4(9)
C(2A)-C(3A)-C(4A)-N(5A)	-0.3(17)
C(14A)-C(3A)-C(4A)-N(5A)	177.9(10)
C(3A)-C(4A)-N(5A)-C(6A)	-1.9(17)
C(4A)-N(5A)-C(6A)-O(6A)	-175.2(10)
C(4A)-N(5A)-C(6A)-N(1A)	5.1(14)
C(2A)-N(1A)-C(6A)-O(6A)	173.5(9)
C(7A)-N(1A)-C(6A)-O(6A)	3.9(14)
C(2A)-N(1A)-C(6A)-N(5A)	-6.7(14)
C(7A)-N(1A)-C(6A)-N(5A)	-176.4(9)
C(2A)-N(1A)-C(7A)-O(7A)	102.3(11)



---

C(6A)-N(1A)-C(7A)-O(7A)	-87.5(12)
C(2A)-N(1A)-C(7A)-C(8A)	-76.3(11)
C(6A)-N(1A)-C(7A)-C(8A)	93.9(11)
O(7A)-C(7A)-C(8A)-C(9A)	172.4(10)
N(1A)-C(7A)-C(8A)-C(9A)	-9.2(14)
O(7A)-C(7A)-C(8A)-C(13A)	-7.4(16)
N(1A)-C(7A)-C(8A)-C(13A)	171.0(9)
C(13A)-C(8A)-C(9A)-C(10A)	-2.2(14)
C(7A)-C(8A)-C(9A)-C(10A)	178.0(9)
C(8A)-C(9A)-C(10A)-C(11A)	2.1(15)
C(9A)-C(10A)-C(11A)-C(12A)	-0.8(17)
C(10A)-C(11A)-C(12A)-C(13A)	-0.4(17)
C(11A)-C(12A)-C(13A)-C(8A)	0.2(16)
C(9A)-C(8A)-C(13A)-C(12A)	1.1(15)
C(7A)-C(8A)-C(13A)-C(12A)	-179.2(9)
C(6B)-N(1B)-C(2B)-O(2B)	179.6(10)
C(7B)-N(1B)-C(2B)-O(2B)	3.1(14)
C(6B)-N(1B)-C(2B)-C(3B)	1.2(14)
C(7B)-N(1B)-C(2B)-C(3B)	-175.3(8)
O(2B)-C(2B)-C(3B)-C(4B)	179.1(10)
N(1B)-C(2B)-C(3B)-C(4B)	-2.6(14)
O(2B)-C(2B)-C(3B)-C(14B)	-1.3(16)
N(1B)-C(2B)-C(3B)-C(14B)	176.9(9)
C(2B)-C(3B)-C(4B)-N(5B)	1.5(17)
C(14B)-C(3B)-C(4B)-N(5B)	-178.1(10)
C(3B)-C(4B)-N(5B)-C(6B)	1.5(16)
C(4B)-N(5B)-C(6B)-O(6B)	174.1(10)
C(4B)-N(5B)-C(6B)-N(1B)	-2.9(14)
C(2B)-N(1B)-C(6B)-O(6B)	-175.6(9)
C(7B)-N(1B)-C(6B)-O(6B)	1.0(14)
C(2B)-N(1B)-C(6B)-N(5B)	1.5(14)
C(7B)-N(1B)-C(6B)-N(5B)	178.1(8)
C(6B)-N(1B)-C(7B)-O(7B)	82.4(11)
C(2B)-N(1B)-C(7B)-O(7B)	-100.6(11)
C(6B)-N(1B)-C(7B)-C(8B)	-98.8(10)

---

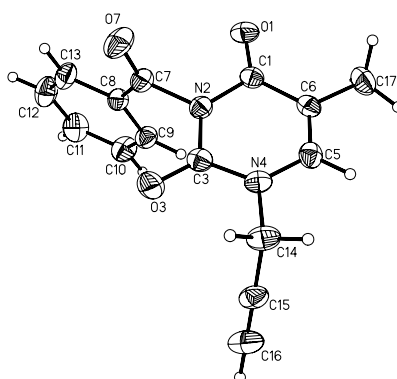
C(2B)-N(1B)-C(7B)-C(8B)	78.1(10)
O(7B)-C(7B)-C(8B)-C(9B)	15.0(15)
N(1B)-C(7B)-C(8B)-C(9B)	-163.6(9)
O(7B)-C(7B)-C(8B)-C(13B)	-164.4(10)
N(1B)-C(7B)-C(8B)-C(13B)	17.0(13)
C(13B)-C(8B)-C(9B)-C(10B)	-2.7(15)
C(7B)-C(8B)-C(9B)-C(10B)	177.9(9)
C(8B)-C(9B)-C(10B)-C(11B)	1.6(16)
C(9B)-C(10B)-C(11B)-C(12B)	1.1(16)
C(10B)-C(11B)-C(12B)-C(13B)	-2.7(16)
C(11B)-C(12B)-C(13B)-C(8B)	1.6(15)
C(9B)-C(8B)-C(13B)-C(12B)	1.1(14)
C(7B)-C(8B)-C(13B)-C(12B)	-179

**Symmetry transformations used to generate equivalent atoms:**

Donor --- H....Acceptor [ ARU ]	D - H	H...A	D...A	D - H...A
C(13B)--H(13B) ..O(2B) [ -1/2+x,-y,z]	0.95	2.54	3.432(12)	156'
C(14B)--H(14F) ..O(7B) [ -1/2+x,-y,z]	0.98	2.58	3.326(13)	133

## **Appendix D**

---

**Appendix D. 161 crystal structure**


Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 16. Crystal data and structure refinement for 161.**

Identification code	x82469o	
Empirical formula	C <sub>15</sub> H <sub>12</sub> N <sub>2</sub> O <sub>3</sub>	
Formula weight	268.27	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	Pna2(1)	
Unit cell dimensions	a = 11.656(2) Å	α = 90°.
	b = 12.237(2) Å	β = 90°.
	c = 9.2778(16) Å	γ = 90°.
Volume	1323.4(4) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.346 Mg/m <sup>3</sup>	
Absorption coefficient	0.096 mm <sup>-1</sup>	
F(000)	560	
Crystal size	0.58 x 0.44 x 0.40 mm <sup>3</sup>	
Theta range for data collection	2.41 to 30.76°.	
Index ranges	-16 ≤ h ≤ 16, -17 ≤ k ≤ 17, -13 ≤ l ≤ 13	
Reflections collected	29233	
Independent reflections	2173 [R(int) = 0.0360]	

Completeness to theta = 25.00°	100.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.962 and 0.863
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	2173 / 1 / 182
Goodness-of-fit on F <sup>2</sup>	1.106
Final R indices [I>2sigma(I)]	R1 = 0.0444, wR2 = 0.1176
R indices (all data)	R1 = 0.0543, wR2 = 0.1234
Largest diff. peak and hole	0.333 and -0.185 e.Å <sup>-3</sup>

**Table 17. Bond lengths [Å] and angles [°] for 161.**

C(1)-O(1)	1.232(3)	C(12)-H(12)	0.9500
C(1)-N(2)	1.412(3)	C(13)-H(13)	0.9500
C(1)-C(6)	1.446(3)	C(14)-C(15)	1.471(4)
N(2)-C(3)	1.385(3)	C(14)-H(14A)	0.9900
N(2)-C(7)	1.476(3)	C(14)-H(14B)	0.9900
C(3)-O(3)	1.223(3)	C(15)-C(16)	1.179(4)
C(3)-N(4)	1.375(3)	C(16)-H(16)	0.9500
N(4)-C(5)	1.376(3)	C(17)-H(17A)	0.9800
N(4)-C(14)	1.470(3)	C(17)-H(17B)	0.9800
C(5)-C(6)	1.343(3)	C(17)-H(17C)	0.9800
C(5)-H(5)	0.9500	O(1)-C(1)-N(2)	118.9(2)
C(6)-C(17)	1.503(3)	O(1)-C(1)-C(6)	126.4(2)
C(7)-O(7)	1.205(3)	N(2)-C(1)-C(6)	114.74(19)
C(7)-C(8)	1.465(3)	C(3)-N(2)-C(1)	126.54(18)
C(8)-C(9)	1.395(3)	C(3)-N(2)-C(7)	116.07(18)
C(8)-C(13)	1.412(3)	C(1)-N(2)-C(7)	117.29(19)
C(9)-C(10)	1.386(3)	O(3)-C(3)-N(4)	123.8(2)
C(9)-H(9)	0.9500	O(3)-C(3)-N(2)	122.0(2)
C(10)-C(11)	1.385(4)	N(4)-C(3)-N(2)	114.24(18)
C(10)-H(10)	0.9500	C(3)-N(4)-C(5)	122.02(19)
C(11)-C(12)	1.391(4)	C(3)-N(4)-C(14)	117.97(19)
C(11)-H(11)	0.9500	C(5)-N(4)-C(14)	120.0(2)
C(12)-C(13)	1.379(4)	C(6)-C(5)-N(4)	123.6(2)

C(6)-C(5)-H(5)	118.2	C(13)-C(12)-C(11)	119.6(2)
N(4)-C(5)-H(5)	118.2	C(13)-C(12)-H(12)	120.2
C(5)-C(6)-C(1)	118.32(19)	C(11)-C(12)-H(12)	120.2
C(5)-C(6)-C(17)	122.9(2)	C(12)-C(13)-C(8)	120.0(2)
C(1)-C(6)-C(17)	118.8(2)	C(12)-C(13)-H(13)	120.0
O(7)-C(7)-C(8)	125.6(2)	C(8)-C(13)-H(13)	120.0
O(7)-C(7)-N(2)	118.4(2)	N(4)-C(14)-C(15)	112.68(18)
C(8)-C(7)-N(2)	115.92(18)	N(4)-C(14)-H(14A)	109.1
C(9)-C(8)-C(13)	119.8(2)	C(15)-C(14)-H(14A)	109.1
C(9)-C(8)-C(7)	122.1(2)	N(4)-C(14)-H(14B)	109.1
C(13)-C(8)-C(7)	118.1(2)	C(15)-C(14)-H(14B)	109.1
C(10)-C(9)-C(8)	119.5(2)	H(14A)-C(14)-H(14B)	107.8
C(10)-C(9)-H(9)	120.2	C(16)-C(15)-C(14)	179.3(3)
C(8)-C(9)-H(9)	120.2	C(15)-C(16)-H(16)	180.0
C(11)-C(10)-C(9)	120.3(2)	C(6)-C(17)-H(17A)	109.5
C(11)-C(10)-H(10)	119.9	C(6)-C(17)-H(17B)	109.5
C(9)-C(10)-H(10)	119.9	H(17A)-C(17)-H(17B)	109.5
C(10)-C(11)-C(12)	120.7(2)	C(6)-C(17)-H(17C)	109.5
C(10)-C(11)-H(11)	119.6	H(17A)-C(17)-H(17C)	109.5
C(12)-C(11)-H(11)	119.6	H(17B)-C(17)-H(17C)	109.5

Symmetry transformations used to generate equivalent atoms:

**Table 18. Torsion angles [°] for 161.**

O(1)-C(1)-N(2)-C(3)	-176.24(19)
C(6)-C(1)-N(2)-C(3)	4.0(3)
O(1)-C(1)-N(2)-C(7)	0.0(3)
C(6)-C(1)-N(2)-C(7)	-179.74(18)
C(1)-N(2)-C(3)-O(3)	-178.6(2)
C(7)-N(2)-C(3)-O(3)	5.1(3)
C(1)-N(2)-C(3)-N(4)	2.5(3)
C(7)-N(2)-C(3)-N(4)	-173.76(18)
O(3)-C(3)-N(4)-C(5)	174.3(2)
N(2)-C(3)-N(4)-C(5)	-6.9(3)

---

O(3)-C(3)-N(4)-C(14)	-5.2(3)
N(2)-C(3)-N(4)-C(14)	173.64(18)
C(3)-N(4)-C(5)-C(6)	4.6(3)
C(14)-N(4)-C(5)-C(6)	-175.9(2)
N(4)-C(5)-C(6)-C(1)	2.6(3)
N(4)-C(5)-C(6)-C(17)	-178.7(2)
O(1)-C(1)-C(6)-C(5)	173.9(2)
N(2)-C(1)-C(6)-C(5)	-6.4(3)
O(1)-C(1)-C(6)-C(17)	-4.9(3)
N(2)-C(1)-C(6)-C(17)	174.8(2)
C(3)-N(2)-C(7)-O(7)	86.3(3)
C(1)-N(2)-C(7)-O(7)	-90.3(3)
C(3)-N(2)-C(7)-C(8)	-94.6(2)
C(1)-N(2)-C(7)-C(8)	88.7(2)
O(7)-C(7)-C(8)-C(9)	178.4(2)
N(2)-C(7)-C(8)-C(9)	-0.6(3)
O(7)-C(7)-C(8)-C(13)	-0.5(4)
N(2)-C(7)-C(8)-C(13)	-179.5(2)
C(13)-C(8)-C(9)-C(10)	1.3(3)
C(7)-C(8)-C(9)-C(10)	-177.5(2)
C(8)-C(9)-C(10)-C(11)	-2.1(4)
C(9)-C(10)-C(11)-C(12)	1.3(5)
C(10)-C(11)-C(12)-C(13)	0.3(5)
C(11)-C(12)-C(13)-C(8)	-1.0(4)
C(9)-C(8)-C(13)-C(12)	0.2(4)
C(7)-C(8)-C(13)-C(12)	179.1(2)
C(3)-N(4)-C(14)-C(15)	82.2(3)
C(5)-N(4)-C(14)-C(15)	-97.3(3)
N(4)-C(14)-C(15)-C(16)	-132(27)

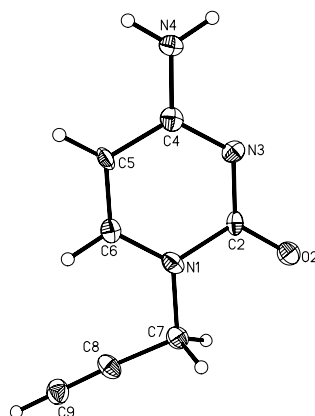
---

Symmetry transformations used to generate equivalent atoms:

## **Appendix E**



---

**Appendix E. 174 crystal structure**


Displacement ellipsoids drawn at the 50% probability level

Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 19. Crystal data and structure refinement for 174.**

Identification code	x82673	
Empirical formula	C7 H7 N3 O	
Formula weight	149.16	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)/c	
Unit cell dimensions	a = 5.3400(11) Å	$\alpha = 90^\circ$ .
	b = 17.674(4) Å	$\beta = 97.46(3)^\circ$ .
	c = 6.9390(14) Å	$\gamma = 90^\circ$ .
Volume	649.3(2) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.526 Mg/m <sup>3</sup>	
Absorption coefficient	0.109 mm <sup>-1</sup>	
F(000)	312	
Crystal size	0.38 x 0.22 x 0.03 mm <sup>3</sup>	
Theta range for data collection	2.30 to 25.18°.	
Index ranges	-6 ≤ h ≤ 6, 0 ≤ k ≤ 21, 0 ≤ l ≤ 8	

Reflections collected	1169
Independent reflections	1169 [R(int) = 0.0000]
Completeness to theta = 25.00°	97.3 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.997 and 0.819
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	1169 / 2 / 108
Goodness-of-fit on F <sup>2</sup>	0.942
Final R indices [I > 2sigma(I)]	R1 = 0.0824, wR2 = 0.2029
R indices (all data)	R1 = 0.1903, wR2 = 0.2799
Largest diff. peak and hole	0.460 and -0.439 e.Å <sup>-3</sup>

**Table 20. Bond lengths [Å] and angles [°] for 174.**

N(1)-C(6)	1.341(8)	C(2)-N(1)-C(7)	117.7(5)
N(1)-C(2)	1.381(8)	O(2)-C(2)-N(3)	122.7(6)
N(1)-C(7)	1.467(8)	O(2)-C(2)-N(1)	118.2(6)
C(2)-O(2)	1.235(7)	N(3)-C(2)-N(1)	119.0(5)
C(2)-N(3)	1.363(8)	C(4)-N(3)-C(2)	121.0(5)
N(3)-C(4)	1.317(8)	N(3)-C(4)-N(4)	119.7(5)
C(4)-N(4)	1.330(8)	N(3)-C(4)-C(5)	121.2(6)
C(4)-C(5)	1.404(8)	N(4)-C(4)-C(5)	119.1(6)
N(4)-H(41)	0.89(2)	C(4)-N(4)-H(41)	116(4)
N(4)-H(42)	0.90(2)	C(4)-N(4)-H(42)	124(4)
C(5)-C(6)	1.354(9)	H(41)-N(4)-H(42)	118(6)
C(5)-H(5)	0.9500	C(6)-C(5)-C(4)	116.7(6)
C(6)-H(6)	0.9500	C(6)-C(5)-H(5)	121.6
C(7)-C(8)	1.459(9)	C(4)-C(5)-H(5)	121.6
C(7)-H(7A)	0.9900	N(1)-C(6)-C(5)	122.6(5)
C(7)-H(7B)	0.9900	N(1)-C(6)-H(6)	118.7
C(8)-C(9)	1.151(9)	C(5)-C(6)-H(6)	118.7
C(9)-H(9)	0.9500	C(8)-C(7)-N(1)	112.1(5)
		C(8)-C(7)-H(7A)	109.2
C(6)-N(1)-C(2)	119.3(5)	N(1)-C(7)-H(7A)	109.2
C(6)-N(1)-C(7)	123.0(5)	C(8)-C(7)-H(7B)	109.2

N(1)-C(7)-H(7B)	109.2	C(9)-C(8)-C(7)	177.1(7)
H(7A)-C(7)-H(7B)	107.9	C(8)-C(9)-H(9)	180.0

Symmetry transformations used to generate equivalent atoms:

**Table 21. Torsion angles [°] for 174.**

C(6)-N(1)-C(2)-O(2)	179.2(5)	N(3)-C(4)-C(5)-C(6)	2.4(9)
C(7)-N(1)-C(2)-O(2)	1.2(8)	N(4)-C(4)-C(5)-C(6)	-178.0(6)
C(6)-N(1)-C(2)-N(3)	0.0(8)	C(2)-N(1)-C(6)-C(5)	2.8(9)
C(7)-N(1)-C(2)-N(3)	-177.9(5)	C(7)-N(1)-C(6)-C(5)	-179.4(6)
O(2)-C(2)-N(3)-C(4)	179.4(6)	C(4)-C(5)-C(6)-N(1)	-4.0(9)
N(1)-C(2)-N(3)-C(4)	-1.4(8)	C(6)-N(1)-C(7)-C(8)	16.4(8)
C(2)-N(3)-C(4)-N(4)	-179.4(5)	C(2)-N(1)-C(7)-C(8)	-165.8(5)
C(2)-N(3)-C(4)-C(5)	0.2(8)		

Symmetry transformations used to generate equivalent atoms:

**Table 22. Hydrogen bonds for 174 [Å and °].**

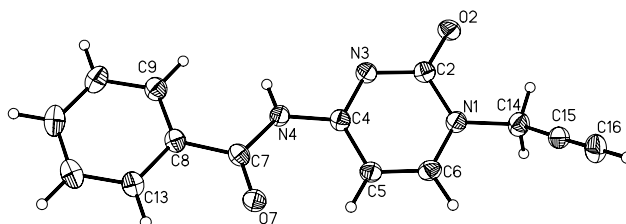
D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(4)-H(41)...O(2)#1	0.89(2)	2.05(3)	2.924(7)	166(6)
N(4)-H(42)...N(3)#2	0.90(2)	2.08(3)	2.946(7)	162(6)

Symmetry transformations used to generate equivalent atoms:

#1 x,y,z-1 #2 -x+3,-y+1,-z+1

## **Appendix F**

---

**Appendix F. 166 crystal structure**


Displacement ellipsoids drawn at the 50% probability level

Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 23. Crystal data and structure refinement for 166**

Identification code	2009src0266	
Empirical formula	C <sub>14</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	
Formula weight	253.26	
Temperature	120(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2 <sub>1</sub> /n	
Unit cell dimensions	a = 12.8138(3) Å	α = 90°.
	b = 6.0048(2) Å	β = 90.599(2)°.
	c = 15.2212(4) Å	γ = 90°.
Volume	1171.12(6) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.436 Mg/m <sup>3</sup>	
Absorption coefficient	0.100 mm <sup>-1</sup>	
F(000)	528	
Crystal size	0.21 x 0.1 x 0.03 mm <sup>3</sup>	
Theta range for data collection	3.18 to 27.48°.	
Index ranges	-16 ≤ h ≤ 16, -7 ≤ k ≤ 6, -19 ≤ l ≤ 17	
Reflections collected	14858	

Independent reflections	2669 [R(int) = 0.0504]
Completeness to theta = 25.00°	99.8 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.997 and 0.721
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	2669 / 0 / 175
en Goodness-of-fit on F <sup>2</sup>	1.118
Final R indices [I>2sigma(I)]	R1 = 0.0593, wR2 = 0.1063
R indices (all data)	R1 = 0.0830, wR2 = 0.1187
Largest diff. peak and hole	0.264 and -0.270 e.Å <sup>-3</sup>

**Table 24. Bond lengths [Å] and angles [°] for 166.**

N(1)-C(6)	1.352(3)	C(11)-H(11)	0.9500
N(1)-C(2)	1.400(3)	C(12)-C(13)	1.384(3)
N(1)-C(14)	1.483(3)	C(12)-H(12)	0.9500
C(2)-O(2)	1.239(2)	C(13)-H(13)	0.9500
C(2)-N(3)	1.366(3)	C(14)-C(15)	1.463(3)
N(3)-C(4)	1.326(2)	C(14)-H(14A)	0.9900
C(4)-N(4)	1.392(3)	C(14)-H(14B)	0.9900
C(4)-C(5)	1.410(3)	C(15)-C(16)	1.186(3)
N(4)-C(7)	1.388(2)	C(16)-H(16)	0.9500
N(4)-H(4N)	0.87(3)	C(6)-N(1)-C(2)	120.99(18)
C(5)-C(6)	1.353(3)	C(6)-N(1)-C(14)	118.82(18)
C(5)-H(5)	0.9500	C(2)-N(1)-C(14)	120.14(17)
C(6)-H(6)	0.9500	O(2)-C(2)-N(3)	122.1(2)
C(7)-O(7)	1.220(2)	O(2)-C(2)-N(1)	120.04(19)
C(7)-C(8)	1.491(3)	N(3)-C(2)-N(1)	117.87(17)
C(8)-C(9)	1.391(3)	C(4)-N(3)-C(2)	120.08(18)
C(8)-C(13)	1.399(3)	N(3)-C(4)-N(4)	113.74(18)
C(9)-C(10)	1.388(3)	N(3)-C(4)-C(5)	123.18(19)
C(9)-H(9)	0.9500	N(4)-C(4)-C(5)	123.06(18)
C(10)-C(11)	1.379(3)	C(7)-N(4)-C(4)	126.80(18)
C(10)-H(10)	0.9500	C(7)-N(4)-H(4N)	118.7(16)
C(11)-C(12)	1.379(4)	C(4)-N(4)-H(4N)	114.3(16)

---

C(6)-C(5)-C(4)	116.22(19)
C(6)-C(5)-H(5)	121.9
C(4)-C(5)-H(5)	121.9
N(1)-C(6)-C(5)	121.6(2)
N(1)-C(6)-H(6)	119.2
C(5)-C(6)-H(6)	119.2
O(7)-C(7)-N(4)	122.0(2)
O(7)-C(7)-C(8)	121.10(18)
N(4)-C(7)-C(8)	116.90(18)
C(9)-C(8)-C(13)	119.0(2)
C(9)-C(8)-C(7)	124.60(19)
C(13)-C(8)-C(7)	116.38(19)
C(10)-C(9)-C(8)	119.9(2)
C(10)-C(9)-H(9)	120.0
C(8)-C(9)-H(9)	120.0
C(11)-C(10)-C(9)	120.5(2)
C(11)-C(10)-H(10)	119.7
C(9)-C(10)-H(10)	119.7
C(10)-C(11)-C(12)	120.1(2)
C(10)-C(11)-H(11)	119.9
C(12)-C(11)-H(11)	119.9
C(11)-C(12)-C(13)	119.9(2)
C(11)-C(12)-H(12)	120.0
C(13)-C(12)-H(12)	120.0
C(12)-C(13)-C(8)	120.5(2)
C(12)-C(13)-H(13)	119.8
C(8)-C(13)-H(13)	119.8
C(15)-C(14)-N(1)	112.24(17)
C(15)-C(14)-H(14A)	109.2
N(1)-C(14)-H(14A)	109.2
C(15)-C(14)-H(14B)	109.2
N(1)-C(14)-H(14B)	109.2
H(14A)-C(14)-H(14B)	107.9
C(16)-C(15)-C(14)	177.5(2)
C(15)-C(16)-H(16)	180.0

---

Symmetry transformations used to generate equivalent atoms:

**Table 25. Torsion angles [°] for 166**

C(6)-N(1)-C(2)-O(2) 179.44(19)	C(4)-N(4)-C(7)-C(8) -179.44(19)
C(14)-N(1)-C(2)-O(2) -3.2(3)	O(7)-C(7)-C(8)-C(9) 171.7(2)
C(6)-N(1)-C(2)-N(3) -0.5(3)	N(4)-C(7)-C(8)-C(9) -8.1(3)
C(14)-N(1)-C(2)-N(3) 176.89(18)	O(7)-C(7)-C(8)-C(13) -6.5(3)
O(2)-C(2)-N(3)-C(4) 179.47(19)	N(4)-C(7)-C(8)-C(13) 173.73(19)
N(1)-C(2)-N(3)-C(4) -0.6(3)	C(13)-C(8)-C(9)-C(10) 1.0(3)
C(2)-N(3)-C(4)-N(4) -176.44(18)	C(7)-C(8)-C(9)-C(10) -177.1(2)
C(2)-N(3)-C(4)-C(5) 1.7(3)	C(8)-C(9)-C(10)-C(11) -0.3(3)
N(3)-C(4)-N(4)-C(7) -167.77(19)	C(9)-C(10)-C(11)-C(12) -0.6(4)
C(5)-C(4)-N(4)-C(7) 14.1(3)	C(10)-C(11)-C(12)-C(13) 0.7(4)
N(3)-C(4)-C(5)-C(6) -1.7(3)	C(11)-C(12)-C(13)-C(8) 0.1(3)
N(4)-C(4)-C(5)-C(6) 176.3(2)	C(9)-C(8)-C(13)-C(12) -0.9(3)
C(2)-N(1)-C(6)-C(5) 0.5(3)	C(7)-C(8)-C(13)-C(12) 177.4(2)
C(14)-N(1)-C(6)-C(5) -177.0(2)	C(6)-N(1)-C(14)-C(15) -61.5(3)
C(4)-C(5)-C(6)-N(1) 0.6(3)	C(2)-N(1)-C(14)-C(15) 121.1(2)
C(4)-N(4)-C(7)-O(7) 0.8(3)	

---

Symmetry transformations used to generate equivalent atoms:

**Table 26. Hydrogen bonds for 166 [Å and °].**

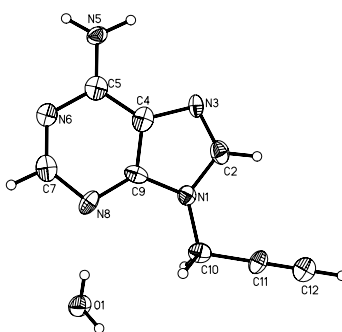
D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(4)-H(4N)...O(2)#1	0.87(3)	2.21(3)	3.056(2)	164(2)

---

Symmetry transformations used to generate equivalent atoms:



## **Appendix G**

**Appendix G. 175 crystal structure.**

The Alkyne group is anti rather than syn to the fused ring system. There is a water solvent molecule which engages in hydrogen bonding with N8 and with another water molecule generated by symmetry. There is one water molecule per organic molecule.

Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 27. Crystal data and structure refinement for 175.**

Identification code	x82526 LT304	
Empirical formula	C <sub>8</sub> H <sub>9</sub> N <sub>5</sub> O	
Formula weight	191.20	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)/c	
Unit cell dimensions	a = 6.826(2) Å	α = 90°.
	b = 26.976(9) Å	β = 96.104(14)°.
	c = 4.8221(15) Å	γ = 90°.
Volume	882.9(5) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.438 Mg/m <sup>3</sup>	
Absorption coefficient	0.103 mm <sup>-1</sup>	
F(000)	400	
Crystal size	0.38 x 0.08 x 0.06 mm <sup>3</sup>	
Theta range for data collection	3.00 to 24.82°.	

Index ranges	-8<=h<=7, -31<=k<=31, -5<=l<=5
Reflections collected	13197
Independent reflections	1520 [R(int) = 0.0572]
Completeness to theta = 24.82°	99.3 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.9938 and 0.9618
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	1520 / 6 / 139
Goodness-of-fit on F <sup>2</sup>	1.040
Final R indices [I>2sigma(I)]	R1 = 0.0541, wR2 = 0.1217
R indices (all data)	R1 = 0.0873, wR2 = 0.1371
Largest diff. peak and hole	0.438 and -0.203 e.Å <sup>-3</sup>

**Table 28. Bond lengths [Å] and angles [°] for 175.**

N(1)-C(9)	1.362(4)	C(12)-H(12)	0.9500
N(1)-C(2)	1.369(4)	O(1)-H(11W)	0.903(17)
N(1)-C(10)	1.455(4)	O(1)-H(12W)	0.881(17)
C(2)-N(3)	1.315(4)	C(9)-N(1)-C(2)	106.4(2)
C(2)-H(2)	0.9500	C(9)-N(1)-C(10)	124.8(2)
N(3)-C(4)	1.379(4)	C(2)-N(1)-C(10)	128.5(2)
C(4)-C(9)	1.395(4)	N(3)-C(2)-N(1)	114.0(3)
C(4)-C(5)	1.399(4)	N(3)-C(2)-H(2)	123.0
C(5)-N(5)	1.335(4)	N(1)-C(2)-H(2)	123.0
C(5)-N(6)	1.380(4)	C(2)-N(3)-C(4)	103.2(2)
N(5)-H(51)	0.903(17)	N(3)-C(4)-C(9)	111.4(3)
N(5)-H(52)	0.887(17)	N(3)-C(4)-C(5)	132.1(3)
N(6)-C(7)	1.322(4)	C(9)-C(4)-C(5)	116.5(3)
C(7)-N(8)	1.345(4)	N(5)-C(5)-N(6)	120.3(3)
C(7)-H(7)	0.9500	N(5)-C(5)-C(4)	122.5(3)
N(8)-C(9)	1.346(4)	N(6)-C(5)-C(4)	117.2(3)
C(10)-C(11)	1.472(4)	C(5)-N(5)-H(51)	119.4(18)
C(10)-H(10A)	0.9900	C(5)-N(5)-H(52)	120.2(19)
C(10)-H(10B)	0.9900	H(51)-N(5)-H(52)	120(3)
C(11)-C(12)	1.169(4)	C(7)-N(6)-C(5)	119.3(2)

N(6)-C(7)-N(8)	129.0(3)	N(1)-C(10)-H(10A)	109.3
N(6)-C(7)-H(7)	115.5	C(11)-C(10)-H(10A)	109.3
N(8)-C(7)-H(7)	115.5	N(1)-C(10)-H(10B)	109.3
C(7)-N(8)-C(9)	110.2(3)	C(11)-C(10)-H(10B)	109.3
N(8)-C(9)-N(1)	127.2(3)	H(10A)-C(10)-H(10B)	107.9
N(8)-C(9)-C(4)	127.8(3)	C(12)-C(11)-C(10)	178.0(3)
N(1)-C(9)-C(4)	105.0(2)	C(11)-C(12)-H(12)	180.0
N(1)-C(10)-C(11)	111.7(2)	H(11W)-O(1)-H(12W)	110(3)

Symmetry transformations used to generate equivalent atoms:

**Table 29. Torsion angles [°] for 175.**

C(9)-N(1)-C(2)-N(3)	1.4(3)	C(7)-N(8)-C(9)-N(1)	-177.7(3)
C(10)-N(1)-C(2)-N(3)	175.5(2)	C(7)-N(8)-C(9)-C(4)	0.7(4)
N(1)-C(2)-N(3)-C(4)	-1.0(3)	C(2)-N(1)-C(9)-N(8)	177.5(3)
C(2)-N(3)-C(4)-C(9)	0.2(3)	C(10)-N(1)-C(9)-N(8)	3.1(4)
C(2)-N(3)-C(4)-C(5)	-177.2(3)	C(2)-N(1)-C(9)-C(4)	-1.1(3)
N(3)-C(4)-C(5)-N(5)	-2.6(5)	C(10)-N(1)-C(9)-C(4)	-175.5(2)
C(9)-C(4)-C(5)-N(5)	-179.9(3)	N(3)-C(4)-C(9)-N(8)	-178.0(2)
N(3)-C(4)-C(5)-N(6)	176.4(3)	C(5)-C(4)-C(9)-N(8)	-0.2(4)
C(9)-C(4)-C(5)-N(6)	-0.9(4)	N(3)-C(4)-C(9)-N(1)	0.6(3)
N(5)-C(5)-N(6)-C(7)	-179.6(2)	C(5)-C(4)-C(9)-N(1)	178.4(2)
C(4)-C(5)-N(6)-C(7)	1.3(4)	C(9)-N(1)-C(10)-C(11)	-175.1(2)
C(5)-N(6)-C(7)-N(8)	-0.9(4)	C(2)-N(1)-C(10)-C(11)	11.8(4)
N(6)-C(7)-N(8)-C(9)	-0.1(4)	N(1)-C(10)-C(11)-C(12)	-107(10)

Symmetry transformations used to generate equivalent atoms:

**Table 30. Hydrogen bonds for 175 [Å and °].**

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(5)-H(51)...N(3)#1	0.903(17)	2.20(2)	3.065(3)	161(3)
N(5)-H(52)...N(6)#2	0.887(17)	2.131(18)	3.014(3)	173(3)

---

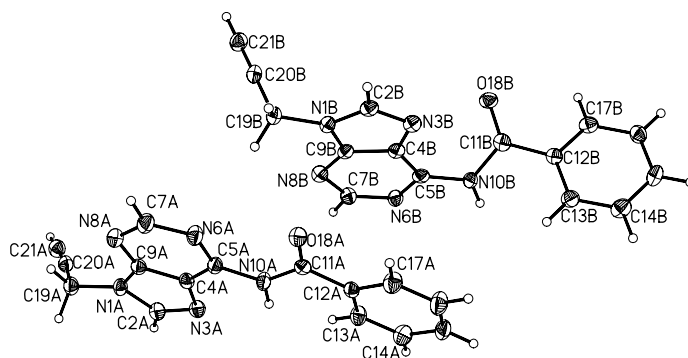
O(1)-H(11W)...N(8)	0.903(17)	1.941(18)	2.837(3)	171(3)
O(1)-H(12W)...O(1)#3	0.881(17)	1.861(18)	2.741(2)	176(3)

Symmetry transformations used to generate equivalent atoms:

#1  $-x, -y+1, -z+1$  #2  $-x+1, -y+1, -z$  #3  $x, -y+3/2, z+1/2$

## Appendix H

## Appendix H. 176 crystal structure



There are two molecules in the asymmetric unit. They differ by the orientation of the adenine ring system with respect to the carbonyl.

Data was collected by Dr Georgina Rosair on a Bruker X8 Apex 2 CCD single crystal diffractometer at 100 K.

**Table 31. Crystal data and structure refinement for 176.**

Identification code	x82562_0m	
Empirical formula	C <sub>15</sub> H <sub>11</sub> N <sub>5</sub> O	
Formula weight	277.29	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	P-1	
Unit cell dimensions	a = 8.2899(6) Å	α = 78.615(3)°.
	b = 12.2253(10) Å	β = 77.172(4)°.
	c = 13.6802(11) Å	γ = 75.627(4)°.
Volume	1294.58(18) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.423 Mg/m <sup>3</sup>	
Absorption coefficient	0.096 mm <sup>-1</sup>	
F(000)	576	
Crystal size	0.62 x 0.35 x 0.04 mm <sup>3</sup>	
Theta range for data collection	2.58 to 29.10°.	

Index ranges	-10<=h<=8, -16<=k<=15, -18<=l<=16
Reflections collected	17848
Independent reflections	5488 [R(int) = 0.0528]
Completeness to theta = 25.00°	94.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.9962 and 0.7464
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	5488 / 0 / 387
Goodness-of-fit on F <sup>2</sup>	0.980
Final R indices [I>2sigma(I)]	R1 = 0.0516, wR2 = 0.0984
R indices (all data)	R1 = 0.1473, wR2 = 0.1285
Largest diff. peak and hole	0.273 and -0.307 e.Å <sup>-3</sup>

**Table 32. Bond lengths [Å] and angles [°] for 176.**

N(1A)-C(9A)	1.377(3)	C(13A)-C(14A)	1.384(3)
N(1A)-C(2A)	1.380(3)	C(13A)-H(13A)	0.9500
N(1A)-C(19A)	1.465(3)	C(14A)-C(15A)	1.382(4)
C(2A)-N(3A)	1.316(3)	C(14A)-H(14A)	0.9500
C(2A)-H(2A)	0.9500	C(15A)-C(16A)	1.379(3)
N(3A)-C(4A)	1.392(3)	C(15A)-H(15A)	0.9500
C(4A)-C(9A)	1.383(3)	C(16A)-C(17A)	1.384(4)
C(4A)-C(5A)	1.393(3)	C(16A)-H(16A)	0.9500
C(5A)-N(6A)	1.338(3)	C(17A)-H(17A)	0.9500
C(5A)-N(10A)	1.388(3)	O(18B)-C(11B)	1.226(3)
N(6A)-C(7A)	1.351(3)	C(19A)-C(20A)	1.462(4)
C(7A)-N(8A)	1.338(3)	C(19A)-H(19A)	0.9900
C(7A)-H(7A)	0.9500	C(19A)-H(19B)	0.9900
N(8A)-C(9A)	1.337(3)	C(20A)-C(21A)	1.184(4)
N(10A)-C(11A)	1.377(3)	C(21A)-H(21A)	0.9500
N(10A)-H(10A)	0.85(3)	N(1B)-C(9B)	1.369(3)
C(11A)-O(18A)	1.218(3)	N(1B)-C(2B)	1.371(3)
C(11A)-C(12A)	1.492(3)	N(1B)-C(19B)	1.468(3)
C(12A)-C(13A)	1.387(3)	C(2B)-N(3B)	1.315(3)
C(12A)-C(17A)	1.392(3)	C(2B)-H(2B)	0.9500



N(3B)-C(4B)	1.396(3)	C(9A)-C(4A)-N(3A)	111.1(2)
C(4B)-C(9B)	1.391(3)	C(9A)-C(4A)-C(5A)	117.2(2)
C(4B)-C(5B)	1.394(3)	N(3A)-C(4A)-C(5A)	131.7(2)
C(5B)-N(6B)	1.349(3)	N(6A)-C(5A)-N(10A)	122.0(2)
C(5B)-N(10B)	1.385(3)	N(6A)-C(5A)-C(4A)	119.2(2)
N(6B)-C(7B)	1.346(3)	N(10A)-C(5A)-C(4A)	118.7(2)
C(7B)-N(8B)	1.334(3)	C(5A)-N(6A)-C(7A)	117.2(2)
C(7B)-H(7B)	0.9500	N(8A)-C(7A)-N(6A)	129.2(2)
N(8B)-C(9B)	1.348(3)	N(8A)-C(7A)-H(7A)	115.4
N(10B)-C(11B)	1.386(3)	N(6A)-C(7A)-H(7A)	115.4
N(10B)-H(10B)	0.89(3)	C(9A)-N(8A)-C(7A)	111.0(2)
C(11B)-C(12B)	1.484(3)	N(8A)-C(9A)-N(1A)	127.8(2)
C(12B)-C(17B)	1.396(3)	N(8A)-C(9A)-C(4A)	126.2(2)
C(12B)-C(13B)	1.398(3)	N(1A)-C(9A)-C(4A)	106.0(2)
C(13B)-C(14B)	1.381(4)	C(11A)-N(10A)-C(5A)	130.0(2)
C(13B)-H(13B)	0.9500	C(11A)-N(10A)-H(10A)	117.6(18)
C(14B)-C(15B)	1.381(3)	C(5A)-N(10A)-H(10A)	112.3(18)
C(14B)-H(14B)	0.9500	O(18A)-C(11A)-N(10A)	122.5(2)
C(15B)-C(16B)	1.388(4)	O(18A)-C(11A)-C(12A)	122.2(2)
C(15B)-H(15B)	0.9500	N(10A)-C(11A)-C(12A)	115.3(2)
C(16B)-C(17B)	1.387(3)	C(13A)-C(12A)-C(17A)	118.3(2)
C(16B)-H(16B)	0.9500	C(13A)-C(12A)-C(11A)	123.9(2)
C(17B)-H(17B)	0.9500	C(17A)-C(12A)-C(11A)	117.8(2)
C(19B)-C(20B)	1.469(4)	C(14A)-C(13A)-C(12A)	120.8(2)
C(19B)-H(19C)	0.9900	C(14A)-C(13A)-H(13A)	119.6
C(19B)-H(19D)	0.9900	C(12A)-C(13A)-H(13A)	119.6
C(20B)-C(21B)	1.179(3)	C(15A)-C(14A)-C(13A)	120.4(2)
C(21B)-H(21B)	0.9500	C(15A)-C(14A)-H(14A)	119.8
C(9A)-N(1A)-C(2A)	105.3(2)	C(13A)-C(14A)-H(14A)	119.8
C(9A)-N(1A)-C(19A)	126.2(2)	C(16A)-C(15A)-C(14A)	119.4(2)
C(2A)-N(1A)-C(19A)	128.1(2)	C(16A)-C(15A)-H(15A)	120.3
N(3A)-C(2A)-N(1A)	114.2(2)	C(14A)-C(15A)-H(15A)	120.3
N(3A)-C(2A)-H(2A)	122.9	C(15A)-C(16A)-C(17A)	120.3(3)
N(1A)-C(2A)-H(2A)	122.9	C(15A)-C(16A)-H(16A)	119.8
C(2A)-N(3A)-C(4A)	103.29(19)	C(17A)-C(16A)-H(16A)	119.8

C(16A)-C(17A)-C(12A)	120.8(2)	C(5B)-N(10B)-C(11B)	123.4(2)
C(16A)-C(17A)-H(17A)	119.6	C(5B)-N(10B)-H(10B)	114.8(17)
C(12A)-C(17A)-H(17A)	119.6	C(11B)-N(10B)-H(10B)	117.1(16)
C(20A)-C(19A)-N(1A)	114.6(2)	O(18B)-C(11B)-N(10B)	122.5(2)
C(20A)-C(19A)-H(19A)	108.6	O(18B)-C(11B)-C(12B)	122.4(2)
N(1A)-C(19A)-H(19A)	108.6	N(10B)-C(11B)-C(12B)	115.1(2)
C(20A)-C(19A)-H(19B)	108.6	C(17B)-C(12B)-C(13B)	119.1(2)
N(1A)-C(19A)-H(19B)	108.6	C(17B)-C(12B)-C(11B)	118.8(2)
H(19A)-C(19A)-H(19B)	107.6	C(13B)-C(12B)-C(11B)	121.9(2)
C(21A)-C(20A)-C(19A)	173.2(3)	C(14B)-C(13B)-C(12B)	119.9(2)
C(20A)-C(21A)-H(21A)	180.0	C(14B)-C(13B)-H(13B)	120.1
C(9B)-N(1B)-C(2B)	105.91(19)	C(12B)-C(13B)-H(13B)	120.1
C(9B)-N(1B)-C(19B)	126.3(2)	C(13B)-C(14B)-C(15B)	120.6(2)
C(2B)-N(1B)-C(19B)	127.76(19)	C(13B)-C(14B)-H(14B)	119.7
N(3B)-C(2B)-N(1B)	114.2(2)	C(15B)-C(14B)-H(14B)	119.7
N(3B)-C(2B)-H(2B)	122.9	C(14B)-C(15B)-C(16B)	120.4(2)
N(1B)-C(2B)-H(2B)	122.9	C(14B)-C(15B)-H(15B)	119.8
C(2B)-N(3B)-C(4B)	103.4(2)	C(16B)-C(15B)-H(15B)	119.8
C(9B)-C(4B)-C(5B)	116.2(2)	C(17B)-C(16B)-C(15B)	119.3(2)
C(9B)-C(4B)-N(3B)	110.5(2)	C(17B)-C(16B)-H(16B)	120.4
C(5B)-C(4B)-N(3B)	133.2(2)	C(15B)-C(16B)-H(16B)	120.4
N(6B)-C(5B)-N(10B)	117.3(2)	C(16B)-C(17B)-C(12B)	120.8(2)
N(6B)-C(5B)-C(4B)	118.9(2)	C(16B)-C(17B)-H(17B)	119.6
N(10B)-C(5B)-C(4B)	123.8(2)	C(12B)-C(17B)-H(17B)	119.6
C(7B)-N(6B)-C(5B)	118.2(2)	N(1B)-C(19B)-C(20B)	112.1(2)
N(8B)-C(7B)-N(6B)	129.0(2)	N(1B)-C(19B)-H(19C)	109.2
N(8B)-C(7B)-H(7B)	115.5	C(20B)-C(19B)-H(19C)	109.2
N(6B)-C(7B)-H(7B)	115.5	N(1B)-C(19B)-H(19D)	109.2
C(7B)-N(8B)-C(9B)	110.5(2)	C(20B)-C(19B)-H(19D)	109.2
N(8B)-C(9B)-N(1B)	126.9(2)	H(19C)-C(19B)-H(19D)	107.9
N(8B)-C(9B)-C(4B)	127.1(2)	C(21B)-C(20B)-C(19B)	179.5(3)
N(1B)-C(9B)-C(4B)	106.0(2)	C(20B)-C(21B)-H(21B)	180.0

---

Symmetry transformations used to generate equivalent atoms:

**Table 33. Torsion angles [°] for 176.**

C(9A)-N(1A)-C(2A)-N(3A) -0.5(3)	N(10A)-C(11A)-C(12A)-C(17A)
C(19A)-N(1A)-C(2A)-N(3A) -174.3(2)	-170.7(2)
N(1A)-C(2A)-N(3A)-C(4A) 0.6(3)	C(17A)-C(12A)-C(13A)-C(14A) 1.7(4)
C(2A)-N(3A)-C(4A)-C(9A) -0.5(3)	C(11A)-C(12A)-C(13A)-C(14A)
C(2A)-N(3A)-C(4A)-C(5A) 177.3(3)	-179.5(2)
C(9A)-C(4A)-C(5A)-N(6A) -2.4(3)	C(12A)-C(13A)-C(14A)-C(15A) -0.6(4)
N(3A)-C(4A)-C(5A)-N(6A) 179.8(2)	C(13A)-C(14A)-C(15A)-C(16A) -0.5(4)
C(9A)-C(4A)-C(5A)-N(10A) 176.8(2)	C(14A)-C(15A)-C(16A)-C(17A) 0.6(4)
N(3A)-C(4A)-C(5A)-N(10A) -0.9(4)	C(15A)-C(16A)-C(17A)-C(12A) 0.5(4)
N(10A)-C(5A)-N(6A)-C(7A) -177.8(2)	C(13A)-C(12A)-C(17A)-C(16A) -1.6(4)
C(4A)-C(5A)-N(6A)-C(7A) 1.5(3)	C(11A)-C(12A)-C(17A)-C(16A)
C(5A)-N(6A)-C(7A)-N(8A) 0.4(4)	179.5(2)
N(6A)-C(7A)-N(8A)-C(9A) -1.1(4)	C(9A)-N(1A)-C(19A)-C(20A) 120.0(3)
C(7A)-N(8A)-C(9A)-N(1A) 179.6(2)	C(2A)-N(1A)-C(19A)-C(20A) -67.4(3)
C(7A)-N(8A)-C(9A)-C(4A) -0.1(3)	C(9B)-N(1B)-C(2B)-N(3B) -0.2(3)
C(2A)-N(1A)-C(9A)-N(8A) -179.6(2)	C(19B)-N(1B)-C(2B)-N(3B) 177.5(2)
C(19A)-N(1A)-C(9A)-N(8A) -5.7(4)	N(1B)-C(2B)-N(3B)-C(4B) 0.7(3)
C(2A)-N(1A)-C(9A)-C(4A) 0.1(3)	C(2B)-N(3B)-C(4B)-C(9B) -1.0(3)
C(19A)-N(1A)-C(9A)-C(4A) 174.1(2)	C(2B)-N(3B)-C(4B)-C(5B) -178.2(3)
N(3A)-C(4A)-C(9A)-N(8A) 180.0(2)	C(9B)-C(4B)-C(5B)-N(6B) -0.1(3)
C(5A)-C(4A)-C(9A)-N(8A) 1.8(4)	N(3B)-C(4B)-C(5B)-N(6B) 177.1(2)
N(3A)-C(4A)-C(9A)-N(1A) 0.2(3)	C(9B)-C(4B)-C(5B)-N(10B) -177.7(2)
C(5A)-C(4A)-C(9A)-N(1A) -178.0(2)	N(3B)-C(4B)-C(5B)-N(10B) -0.6(4)
N(6A)-C(5A)-N(10A)-C(11A) -9.0(4)	N(10B)-C(5B)-N(6B)-C(7B) 177.5(2)
C(4A)-C(5A)-N(10A)-C(11A) 171.7(2)	C(4B)-C(5B)-N(6B)-C(7B) -0.3(3)
C(5A)-N(10A)-C(11A)-O(18A) 4.4(4)	C(5B)-N(6B)-C(7B)-N(8B) 1.1(4)
C(5A)-N(10A)-C(11A)-C(12A)	N(6B)-C(7B)-N(8B)-C(9B) -1.4(4)
-175.9(2)	C(7B)-N(8B)-C(9B)-N(1B) -177.9(2)
O(18A)-C(11A)-C(12A)-C(13A)	C(7B)-N(8B)-C(9B)-C(4B) 0.9(3)
-169.8(2)	C(2B)-N(1B)-C(9B)-N(8B) 178.5(2)
N(10A)-C(11A)-C(12A)-C(13A)	C(19B)-N(1B)-C(9B)-N(8B) 0.8(4)
10.4(3)	C(2B)-N(1B)-C(9B)-C(4B) -0.5(2)
O(18A)-C(11A)-C(12A)-C(17A) 9.1(3)	C(19B)-N(1B)-C(9B)-C(4B) -178.2(2)

---

C(5B)-C(4B)-C(9B)-N(8B) -0.3(4)  
 N(3B)-C(4B)-C(9B)-N(8B) -178.1(2)  
 C(5B)-C(4B)-C(9B)-N(1B) 178.69(19)  
 N(3B)-C(4B)-C(9B)-N(1B) 0.9(3)  
 N(6B)-C(5B)-N(10B)-C(11B) 139.6(2)  
 C(4B)-C(5B)-N(10B)-C(11B) -42.7(3)  
 C(5B)-N(10B)-C(11B)-O(18B) -12.7(4)  
 C(5B)-N(10B)-C(11B)-C(12B) 167.2(2)  
 O(18B)-C(11B)-C(12B)-C(17B)  
 -33.7(3)  
 N(10B)-C(11B)-C(12B)-C(17B)  
 146.4(2)  
 O(18B)-C(11B)-C(12B)-C(13B)  
 141.3(2)  
 N(10B)-C(11B)-C(12B)-C(13B)  
 -38.7(3)  
 C(17B)-C(12B)-C(13B)-C(14B) 1.3(3)  
 C(11B)-C(12B)-C(13B)-C(14B)  
 -173.6(2)  
 C(12B)-C(13B)-C(14B)-C(15B) -0.7(4)  
 C(13B)-C(14B)-C(15B)-C(16B) -0.8(4)  
 C(14B)-C(15B)-C(16B)-C(17B) 1.7(4)  
 C(15B)-C(16B)-C(17B)-C(12B) -1.1(4)  
 C(13B)-C(12B)-C(17B)-C(16B) -0.4(3)  
 C(11B)-C(12B)-C(17B)-C(16B)  
 174.7(2)  
 C(9B)-N(1B)-C(19B)-C(20B) -84.7(3)  
 C(2B)-N(1B)-C(19B)-C(20B) 98.0(3)

---

Symmetry transformations used to generate equivalent atoms:

**Table 34. Hydrogen bonds for 176 [Å and °].**

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(10A)-H(10A)...N(3A)	0.85(3)	2.61(3)	2.991(3)	109(2)
N(10B)-H(10B)...N(6B)#10	0.89(3)	2.15(3)	3.020(3)	166(2)

Symmetry transformations used to generate equivalent atoms:

#1-x+1,-y+1,-z+1

---

**REFERENCES**

1. C. R. UK, *Cancer in the UK:2010*.
2. M. D. Garrett, *Current Science*, 2001, **81**, 515-522.
3. L. H. Hartwell and T. A. Weinert, *Science*, 1989, **246**, 629-634.
4. G. I. Shapiro and J. W. Harper, *The journal of clinical investigation*, 1999, **104**, 1645-1653.
5. S. B. Kaye, *British Journal of Cancer*, 1998, **78**.
6. G. M. Blackburn and M. J. Gait, eds., *Nucleic Acids in Chemistry and Biology*, Oxford University Press, Oxford, 1990.
7. P. Brookes and P. D. Lawley, *J.Biochem*, 1961, **80**, 496-503.
8. G. L. Patrick, ed., *An Introduction to Medicinal Chemistry*, Oxford University Press, Oxford, 2001.
9. H. Umezawa, *J. Antibiot. Ser A*, 1966, **19**, 200-203.
10. J. Feigon, W. A. Denny, W. Leupin and D. J. Kearns, *J. Med. Chem*, 1984, **27**, 450-456.
11. C. R. Calladine, H. R. Drew, B. F. Luisi and A. A. Travers, *Understanding DNA* Elsevier academic press, Editon edn., 2004.
12. W. H. Elliot and D. C. Elliot, eds., *Biochemistry and Molecular Biology*, second edn., Oxford University Press, 2001.
13. P. C. Zamecnik, J. Goodchild and P. S. Y. Taguchi, *Proc. Natl. Acad. Sci. U.S.A.*, 1986, **83**, 4143-4146.
14. F. Vandendriessche, A. V. Aerschot, M. Voortmans, G. Janssen, R. Busson, A. V. Overbeke, W. V. D. Bossche, J. Hoogmartens and P. Herdewijn, *J. Chem. Soc. Perkin Trans* 1993, 1567-1575.
15. T. Boesen, D. S. Pedersen, B. M. Nielsen, A. B. Petersen, M. A. Petersen, M. Munck, U. Henriksen, C. Nielsen and O. Dahl, *Bioorg. & Med. Chem. Lett*, 2003, **13**, 847-850.
16. P. C. Zamecnik and M. L. Stephenson, *Proc. Natl. Acad. Sci. U.S.A.*, 1978, **75**, 280-284
17. D. Jaskulski, J. K. DeReil, W. E. Mercer, B. Calabretta and R. Baserga, *Science*, 1988, **240**, 1544-1546.
18. Z. Huang, K. C. Schnieder and S. A. Benner, *J.Org. Chem*, 1991, **56**, 3869-3882
19. J. P. Shaw, K. Kent, J. Bird, J. Fishback and B. Froehler, *Nucleic Acids Res*, 1991, **19**, 747-750.

20. S. Buchini and C. J. Leumann, *Current Opinion in Chemical Biology*, 2003, **7**.
21. M. J. Cocco, L. A. Hanakahi, M. D. Huber and N. Maizels, *Nucleic Acids Res*, 2003, **31**, 2944-2951.
22. M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura and R. E. Dickenson, *Proc. Natl. Acad. Sci. U.S.A*, 1985, **82**, 1376-1380.
23. S. Neidle, *Nat. Prod. Rep*, 2001, **18**, 291-309.
24. J. J. Kelly, E. E. Baird and P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A*, 1996, **93**, 6981-6985.
25. J. W. Trauger, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc*, 1998, **120**, 3534-3535.
26. B. K. Bhugan, K. S. Smith, E. G. Adams, T. L. Wallace, D. D. V. Hoff and L. H. Li, *Cancer Chemotherapy pharmacology*, 1992, **30**, 348-354.
27. L. H. Li, T. F. DeKoning, R. C. Kelly, W. C. Krueger, J. P. McGovren, G. E. Padbury, G. L. Petzold, T. L. Wallace, R. J. Ouding, M. D. Prairies and I. Gebhard, *Cancer Research*, 1992, 4904-4913.
28. C. A. Carter, W. R. Ward, L. H. Li, T. F. DeKoning, J. P. McGovren and J. Plowman, *Clinical Cancer Research*, 1996, **2**, 1143-1149.
29. G. Pezzoni, M. Grandi, G. Biasoli, L. Capolongo, D. Ballinari, F. C. Giuliani, B. Barbieri, E. Pesenti, A. Paston, N. Mongelli and F. Spreafico, *J. Cancer. Research*, 1991, **64**, 1047-1050.
30. T. Yamori, A. Matsunaga, S. Sato, K. Yamazaki, A. Komi, K. Ishizu, I. Mita, H. Edatsugi, Y. Matsuba, K. Takezawa, O. Nakanishi, H. Kohno, Y. Nakajima, H. Komatsu, T. Andoh and T. Tsuruo, *Cancer Research*, 1999, **59**, 4042-4049.
31. N. J. Carter and S. J. Keam, *Drugs*, 2007, **67**, 2257-2276.
32. G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc*, 1957, 2023-2024.
33. M. P. Knauert and P. M. Glazer, 2001, **10**, 2243-2251.
34. A. J. A. Cobb, *Org. Biomol. Chem*, 2007, **5**, 3260-3275.
35. G. Xiang, R. Bogacki and L. W. McLaughlin, *Nucleic Acids Res*, 1996, **24**, 1963-1970.
36. A. Ono, P. O. P. Ts'O and L. S. Kan, *J. Am. Chem. Soc*, 1991, **113**, 4032-4033.
37. G. Xiang, W. Soussou and L. W. McLaughlin, *J. Am. Chem. Soc*, 1994, **116**, 11155-11156.
38. R. T. Ranasinghe, D. A. Rusling, V. E. C. Pavers, K. R. Fox and T. Brown, *Chem. Commun*, 2005, 2555-2557.
39. H. C. Brown and S. J. Narasimhan, *J. Org. Chem*, 1984, **49**, 3891-3898.

40. E. Uhlmann and A. Peyman, *Chemical Reviews*, 1990, **90**, 544-579.
41. E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 1-22.
42. A. T. Krueger, H. Lu, A. H. F. Lee and E. T. Kool, *Acc. Chem. Res.*, 2007, **40**, 141-145.
43. S. R. Lynch, H. B. Liu, J. M. Gao and E. T. Kool, *J. Am. Chem. Soc.*, 2006, **128**, 14704-14707.
44. J. Gao, H. Lui and E. T. Kool, *Angew. Chem. Int. Ed.*, 2005, **44**, 3118-3120.
45. H. B. Liu, J. M. Gao and E. T. Kool, *J. Am. Chem. Soc.*, 2005, **127**, 1396-1399.
46. H. Liu, J. Gao and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 639-642.
47. H. Liu, S. R. Lynch and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 6900-6903.
48. H. Liu, J. Gao, L. Maynard, Y. D. Saito and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 1102-1105.
49. A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2005, **127**, 3332-3334.
50. A. H. F. Lee and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 132-135.
51. H. Lu, K. He and E. T. Kool, *Angew. Chem. Int. Ed.*, 2004, **43**, 5834-5836.
52. A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **128**, 9219-9223.
53. H. Liu, S. R. Lynch and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 6900-6905.
54. A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2005, **127**, 3332-3338.
55. A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2006, **128**, 9219-9230.
56. F. Vandendriessche, K. Augustyns, A. V. Aerschot, R. Busson, J. Hoogmartens and P. Herdewijn, *Tetrahedron*, 1993, **49**, 7223-7238.
57. V. S. Rana, V. A. Kumar and K. N. Ganesh, *Bioorg. & Med. Chem. Lett.*, , 1997, **7** 2837-2842.
58. K. C. Schneider and S. A. Benner, *J. Am. Chem. Soc.*, 1990, **112**, 453-455.
59. L. Zhang, A. E. Peritz, P. J. Carroll and E. Meggers, *Synthesis*, 2006, **4**, 645-653.
60. L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4175-4176.
61. T. Boesen, C. Madsen, D. S. Pedersen, B. M. Nielsen, A. B. Petersen, M. A. Petersen, M. Munck, U. Henriksen, C. Nielsen and O. Dahl, *Org. Biomol. Chem.*, 2004, **2**, 1245-1254
62. G. D. Hoke, K. Draper, S. M. Freier, C. Gonzalez, V. B. Driver, M. C. Zounes and D. J. Ecker, *Nucleic Acids Res.*, 1991, **19** 5743-5748
63. P. S. Miller, C. H. Agris, L. Aurelian, K. R. Blake, A. Murakami, M. P. Reddy, S. A. Spitz and P. O. P. Ts'o, *Biochemie*, , 1985, **67**, , 769-776, .
64. J. Goodchild, *Bioconjugate Chemistry*, 1990, **1**, 165-186.



- 
65. R. S. Varma, *Synlett*, 1992, 621-636.
  66. P. Miller, C. Agris, L. Aurelian, K. Blake, T. Kelly, A. Murakami, M. P. Reddy, S. Spitz, P. O. P. Ts'o and R. Wides, *Fed. Proc.*, 1984, **43**, 1727-1730.
  67. P. S. Miller, R. A. Cassidy, T. Hamma and N. S. Kondo, *Pharma. Thera.*, 2000, **85**, 159-163.
  68. E. Larsen, K. Danel and E. B. Pedersen, *Nucleosides & Nucleotides*, 1995, **14**, 1905-1912.
  69. E. M. McGuffie, D. Pacheco, G. M. R. Carbone and C. V. Catapano, *Cancer Research*, 2000, **60**, 3790-3799.
  70. C. A. Stein and A. M. Kreig, *Antisense Res. Dev.*, 1994, 67-71.
  71. R. Schutz, M. Cantin, C. Roberts, B. Greinear, E. Uhlmann and C. Leumann, *Angew. Chem. Int. Ed.*, 2000, **39**, 1250-1253.
  72. M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895-1897.
  73. V. S. Rana, V. A. Kumar and K. N. Ganesh, *Tetrahedron*, 2001, **57**, 1311-1321.
  74. V. S. Rana, V. A. Kumar and K. N. Ganesh, *Tetrahedron*, 2001, **57** 1311-1321
  75. K. L. Dueholm, M. Egholm, C. Behrens, L. Christensen, H. F. Hansen, T. Vulpius, K. H. Petersen, R. H. Berg, P. E. Nielsen and O. Buchardt, *J. Org. Chem.*, 1994, **59**, 5767-5773.
  76. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566-568.
  77. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1501.
  78. D. Y. Cherny, B. P. Belotserkovskii, M. D. F. Kamenetskii, M. Egholm, O. Buchardt and P. E. Nielsen, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 1667-1671.
  79. P. Nielsen, F. Kirpekar and J. Wengel, *Nucleic Acids Res.*, 1994, **22**, 703-710.
  80. Y. Zhang, R. Hsung and M. Tracey, *Org. Letters*, 2004, **6**, 1151-1154.
  81. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.
  82. G. Xiang and L. W. McLaughlin, *Tetrahedron*, 1998, **54**, 375-392.
  83. A. A. Edwards, O. Ichihara, S. Murfin, R. Wilkes, M. Whittaker, D. J. Watkin and G. W. J. Fleet, *J. Comb. Chem.*, 2004, **6**, 230-238.
  84. M. L. Lewbart and J. J. Schneider, *J. Org. Chem.*, 1969, **11**, 3505-3512.

85. Y. Leblanc, B. J. Fitzsimmons, J. Adams, F. Perez and J. Rokach, *J. Org. Chem*, 1986, **51**, 789-793
86. S.K.Chaudhary and O. Hernandez, *Tett. Lett*, 1979, 95-98.
87. O. Hernandez, S.K.Chaudhary, R. H. Cox and J. Porter, *Tetrahedron Letters*, 1981, **22**, 1491-1494.
88. S. C. Messenger, J. P. Girard and J. C. Rossi, *Tett. Lett.*, , 1992, **33**, , 2689-2691, .
89. D. K. Mohapatra, *Synthetic Communications* 1999, **29**, 4261-4268.
90. T. Laib, J. Chastanet and J. Zhu, *J.Org. Chem*, 1998, **63**, 1709-1713.
91. J. C. S. d. Costa, *Arkivoc*, 2006, **(i)**, 128-133.
92. D. Walker, Heriot-Watt, 2004.
93. E. J. Corey, J. L. Gras and P. Ulrich, *Tetrahedron Letters*, 1976, **11**, 809-812.
94. P. T. Nyffeler, C. Liang, K. M. Koeller and C. Wong, *J. Am. Chem. Soc*, 2002, **124**, 10773-10778.
95. A. Titz, Z. Radic, O. Schwardt and B. Ernst, *Tetrahedron Lett.*, 2006, **47**, 2383-2385.
96. E. D. Goddard-Borger and R. V. Stick, *Org. Letters*, 2007, **9**, 3797-3800.
97. P. J. Kocienski, ed., *Protecting groups*, First edn., Thieme Medical publishers inc, New York, 2000.
98. N. Whittingham and N. M. Howarth, Editon edn., 2010.
99. M. Choi and H. Kim, *Arch Pharm Res*, 2003, **26**, 990-996.
100. H. B. Lazrek, M. Taourirte, T. Oulih, J. L. Barascut, J. L. Imbach, C. Pannecouque, M. Witrouw and E. D. Clerqc, *Nucleosides, Nucleotides & Nucleic acids*, 2001, **20**, 1949-1960.
101. A. A. Makinsky, A. M. Kritzyn, E. A. Uljanova, O. D. Zakharova and G. A. Nevinsky, *Russian journal of Bioorganic chemistry*, 2000, **26**, 662-668.
102. A. R. Maguire, I. Hladezuk and A. Ford, *Carbohydrate Res*, 2002, **337**, 369-372.
103. D. M. Brown, A. Todd and S. Varadarajan, *J. Org. Chem*, 1956, 2384-2387.
104. S. Guillarme, S. Legoupy, A.-M. Aubertin, C. Olicard, N. Bourgougnon and F. Huet, *Tetrahedron*, 2003, **59**, 2177-2184.
105. W. E. Lindsell, C. Murray, P. N. Preston and T. A. J. Woodman, *Tetrahedron*, 2000, **56**, 1233-1245.
106. N. M. Howarth and L. P. G. Wakelin, *J. Org. Chem*, 1997, **62**, 5441-5450.
107. C. B. Reese and P. A. Stone, *J. Am. Chem. Soc. Perkin Trans. I*, 1984, 1263-1271.

- 
108. S. S. Jones, C. B. Reese, S. Sibanda and A. Ubasawa, *Tett. Lett.*, 1981, **22**, 4755-4758.
109. B. S. Sproat, B. Beijer and A. Iribarren, *Nucleic Acids Res.*, 1990, **18**, 41-49.
110. M. Grötli, M. Douglas, B. Beijer, R. G. Garcia, R. Eritja and B. Sproat, *J.Chem.Soc., Perkin Trans. 1*, 1997, 2779-2788.
111. C. D. Hein, X. Liu and D. Wang, *Pharmaceutical Research*, 2008, **25**, 2216-2230.
112. V. D. Bock, H. Hiemstra and J. H. V. Maarseveen, *Eur. J. Org. Chem.*, 2006, 51-68.
113. J. E. Moses and A. D. Moorhouse, *Chem.Soc. Rev.*, 2007, **36**, 1249-1262.
114. J. Goujon, 2009.
115. M. Nahrwold, T. Bogner, S. Eissler, S. Verma and N. Sewald, *Org. Lett.*, 2010, **x**, x.
116. G. Xiang and L. W. McLaughlin, *Tetrahedron*, 1998, **54**, 375-392.