

**The Modification of Nucleosides to Probe
Decarboxylation and Denitration Processes**



UNIVERSITY OF
LIVERPOOL

Thesis submitted in accordance with the requirements of the University of
Liverpool for the degree of Doctor in Philosophy

By

Kathryn Rose Williams

September 2018

Abstract

The work presented in this thesis explored the modification of naturally occurring nucleosides and is split into two main sections investigating the processes of denitration and decarboxylation. A common theme of the work investigated the nitration of nucleosides and their biological role. Both purines and pyrimidine nucleosides have been investigated, with the initial focus being on guanosine and a potential repair mechanism of a common DNA lesion associated with nitration. Secondly, an essential decarboxylation step in the biosynthesis of pyrimidine nucleosides was investigated to probe whether formation of bioisosteric analogues of a key intermediate in this process could aid in the treatment of disease.

8-Nitroguanine is a DNA lesion strongly associated with inflammation-related carcinogenesis. Nitration of the guanine base greatly labilises the glycosidic bond, often resulting in the formation of abasic sites which can lead to mispairing during DNA replication. A potential repair mechanism of the lesion was investigated involving a reductive denitration reaction. Synthesis of the ribonucleoside form of 8-nitroguanosine was achieved and the process of reductive denitration was investigated using sodium borohydride. The main product of the reduction was found to be 8-aminoguanosine, but a small amount of guanosine was found to form showing reductive denitration of the lesion is a chemically feasible reaction and thus a potential repair mechanism. A deuterium labelling study proved the origin of the guanosine formed was from the 8-nitroguanosine starting material.

The final step in the *de novo* biosynthesis of pyrimidine nucleotides involves a decarboxylation reaction to produce uridine 5'-monophosphate, catalysed by the enzyme orotidine monophosphate decarboxylase (ODCase). Certain disease causing organisms, such as the malaria parasite, rely on this pathway to obtain the pyrimidines they require. Humans utilise two pathways to obtain pyrimidines which introduces the possibility of inhibiting the *de novo* pathway as a means of therapeutic intervention. A series of eleven bioisosteric analogues of the natural substrate of ODCase were synthesised as potential inhibitors of the enzyme and antimalarial agents. Computational docking of the eleven compounds into a crystal structure of ODCase was carried out and showed that all are predicted to fit into the active site. All eleven compounds synthesised have been sent for biological testing in a 3D7 assay to assess their antimalarial activity.

Contents

Abstract	i
Contents	ii
Acknowledgements	v
Abbreviations	vi
Chapter 1 Introduction	2
1.1 The primary structure of nucleic acids.....	2
1.1.1 The nucleobases.....	3
1.1.2 The sugar conformation.....	5
1.2 The secondary structure of nucleic acids.....	7
1.3 The double-helix.....	8
1.3.1 A-form helix.....	9
1.3.2 B-form helix.....	10
1.3.3 Z-form helix.....	11
1.4 Multi-stranded structures.....	12
1.4.1 Triple helices.....	12
1.4.2 <i>i</i> -Motifs.....	13
1.4.3 G-Quadruplexes.....	14
Chapter 2 Introduction	17
2.1 How does 8-nitroguanine form?.....	17
2.2 Why is the 8-nitroguanine lesion a problem?.....	21
2.3 DNA repair.....	24
2.3.1 Excision repair.....	24
2.3.2 Direct repair.....	25
2.4 Reductive denitration.....	27
2.5 Project aims.....	29
Chapter 2 Results and discussion 1	30
2.6 Synthesis of 8-nitroguanosine.....	30
2.7 Reductive denitration of 8-nitroguanosine.....	34

2.8 Control reactions.....	40
2.9 Deuterium study.....	40
2.10 Change of reaction conditions.....	41
2.11 Conclusions	43
2.12 Future work.....	43
Chapter 3 Introduction	46
3.1 Pyrimidine nucleotide biosynthesis	46
3.2 OMP decarboxylase (ODCase).....	48
3.2.1 ODCase catalytic mechanism	48
3.2.2 ODCase promiscuity	53
3.2.3 Known inhibitors of ODCase	54
3.2.4 Design of inhibitors of ODCase	57
3.3 Project aims	60
Chapter 3 Results and discussion 2.....	61
3.4 Synthesis of 6-iodouridine.....	61
3.5 Potential routes to 6-nitrouridine	66
3.5.1 Nucleophilic displacement using nitrite.....	66
3.5.2 Oxidation of C-6 azidouridine	68
3.5.3 Miscellaneous nitration attempts	70
3.6 Click chemistry.....	71
3.7 Formation of C-6 substituted tetrazole	81
3.8 Alternative electrophiles for use in the double deprotonation reaction.....	85
3.8.1 Methyl chloroformate - Synthesis of orotidine.....	85
3.8.2 Methyl chloroformate - Formation of a hydroxamic acid	87
3.8.3 Diethyl chlorophosphate – Formation of a C-6 phosphonate.....	89
3.9 Palladium chemistry	90
3.9.1 Route to a C-6 substituted propargylic acid.....	90
3.9.2 Formation of a carbon linked triazole	94
3.9.3 Isolation of 6-ethynyl uridine	97

3.9.4 Attempts to form a C-6 substituted boronic acid	98
3.10 Synthesis of 6-aminouridine	99
3.11 Conclusions and future work	100
Chapter 4 Results and discussion 3	104
4.1 Validation docking of BMP	104
4.2 Docking of final compounds	106
4.3 Biological testing	111
4.4 Conclusions	113
Chapter 5 Experimental	115
5.1.1 General techniques	115
5.1.2 HPLC solvents	116
5.1.3 HPLC methods	116
5.1.4 Solvents	119
5.1.5 General reagents	120
5.2 Experimental procedures for results and discussion 1	121
5.3 Experimental procedures for results and discussion 2	127
5.4 Computational protocols for results and discussion 3	158
Chapter 6 Bibliography	161
Chapter 7 Appendices	174
Chapter 8 Publication	181

Acknowledgements

First and foremost I would like to express my deepest gratitude to my supervisor, Professor Rick Cosstick. Your constant support, guidance and unfailing enthusiasm over the past four years have been an inspiration. I could not have hoped for a better supervisor.

Additional thanks are given to the following: my secondary supervisor, Dr Ian O'Neil for all of the help you have given me, particularly in relation to the work carried out on 8-nitroguanosine; Dr Andrew Carnell for the helpful discussions in group meetings; Dr Neil Berry for assistance with the computational aspects of the work presented; The MicroBioRefinery at the University of Liverpool, particularly Stephen Moss, for allowing and helping me to use their HPLC-MS equipment; The Liverpool School of Tropical Medicine for testing my final compounds; the technical staff and Analytical Services department at the University of Liverpool, particularly Moya McCarron and Dr Konstantin Luzyanin for their help with mass spectrometry and NMR, and the EPSRC for funding my work.

To the past and present members of the Cosstick and Carnell groups, I would like to say a big thank you for making Lab 1.84 such a fun place to work. I am going to miss our lunchtime PuzzGrid sessions and doing Quiz of the Week! I am particularly grateful to Dr Neil Kershaw for his help and patience with me in the lab and Dr Katie Alexander for helping me to get started in the field of nucleic acid chemistry.

I would like to thank my friends Rudi Grosman, Stephanie Yip, Chris Janot, Manuel Barday, Teresa Almeida and Maria Pin No for providing a welcome distraction from work and always being there when I need them. I thoroughly enjoyed our yearly camping trips to climb various mountains around the UK, even if we did always pick the hardest routes. A special thank you to Rudi. There have been many bad chemistry days, but your support and encouragement at the end of them have helped me to keep going.

Finally, my biggest thank you is to my family. To my Mum and Dad, I am eternally grateful for your unconditional love, support and unwavering belief in me. Your work ethic is a source of inspiration to me and I wouldn't be where I am today without the sacrifices you have made. To my little sister Geor, thank you for always being there for me. To my two Nans, thank you for the endless supplies of food and cups of tea! Seeing you both on a Saturday is always a highlight of my week. You never fail to make smile and I will be forever grateful for your advice, guidance and support.

Abbreviations

A	Adenine
ADP	Adenosine-5'-diphosphate
AGT	O ⁶ -Alkylguanine-DNA alkyl-transferase
AIBN	Azobisisobutyronitrile
AP	Apurinic/apyrimidinic
app	Apparent
Arg	Arginine
Asp	Aspartate
ATP	Adenosine-5'-triphosphate
BAIB	Bis(acetoxy)iodobenzene
BER	Base excision repair
BMP	Barbiturate monophosphate
br	Broad
C	Cytosine
CAIR	4-Carboxy-5-aminoimidazole ribonucleotide
CMP	Cytidine-5'-monophosphate
COSY	Correlation spectroscopy
CTP	Cytidine-5'-triphosphate
CuAAC	Copper-catalysed azide-alkyne cycloaddition
Cys	Cysteine
d	Doublet
dd	Doublet of doublets
DCM	Dichloromethane
Ddn	Deazaflavin-dependent nitroreductase
DMAP	4-Dimethylaminopyridine

DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethyl sulfoxide
DMTr	4,4'-Dimethoxytriphenylmethyl
DNA	Deoxyribonucleic acid
eNOS	Endothelial nitric oxide synthase
ES	Electrospray
FAD	Flavin adenine dinucleotide
G	Guanine
8-HDF	8-Hydroxy-7,8-didemethyl-5-deazariboflavin
His	Histidine
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
<i>i</i>	<i>iso</i>
<i>i</i> -Motif	Intercalated motif
iNOS	Inducible nitric oxide synthase
IR	Infrared
IUPAC-IUB	International Union of Pure and Applied Chemistry – International Union of Biochemistry
<i>J</i>	Coupling constant
K_i	Inhibition constant
LDA	Lithium diisopropylamide
Lys	Lysine
m	Multiplet
MMFF	Molecular mechanics force field

MMR	Mismatch repair
MS	Mass spectrometry
MTHF	Methenyltetrahydrofolate
m/z	Mass to charge ratio
N	North
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NAIR	4-Nitro-5-aminoimidazole ribonucleotide
NER	Nucleotide excision repair
NMR	Nuclear magnetic resonance
nNOS	Neuronal nitric oxide synthase
NOS	Nitric oxide synthase
ODCase	Orotidine monophosphate decarboxylase
OMP	Orotidine monophosphate
<i>p</i>	<i>para</i>
PDB	Protein data bank
pK _a	Acid dissociation constant
ppm	Parts per million
Pro	Proline
PRPP	5-Phosphoribosyl-1-pyrophosphate
PRRs	Pattern recognition receptors
q	Quartet
quat	Quaternary
R _f	Retention factor
RNA	Ribonucleic acid
RNS	Reactive nitrogen species

ROS	Reactive oxygen species
RP	Reverse phase
RSM	Recovered starting material
S	South
s	Singlet
S _N 2	Bimolecular nucleophilic substitution
T	Thymine
<i>t</i>	<i>tert</i>
t	Triplet
t _½	Half-time
TBDMS	<i>tert</i> -Butyldimethylsilyl
TBDPS	<i>tert</i> -Butyldiphenylsilyl
td	Triplet of doublets
TEAB	Triethylammonium bicarbonate
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
<i>T_m</i>	Thermal melting temperature
TFO	Triplex forming oligonucleotide
U	Uracil
UMP	Uridine monophosphate
UV	Ultraviolet
XMP	Xanthosine-5'-monophosphate

Chapter 1

Introduction

Chapter 1 Introduction

1.1 The primary structure of nucleic acids

Nucleotides are the building blocks of the long chain polymers deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).¹ The basic structure of a nucleotide is comprised of three key components, namely, a heterocyclic nitrogenous base (nucleobase), a furanose ring formed from a pentose sugar and at least one phosphate group.² In the absence of any phosphate residues the molecules are referred to as nucleosides rather than nucleotides.² Figure 1 shows the general structure of a nucleoside and the numbering system used to refer to each atom of the sugar.³

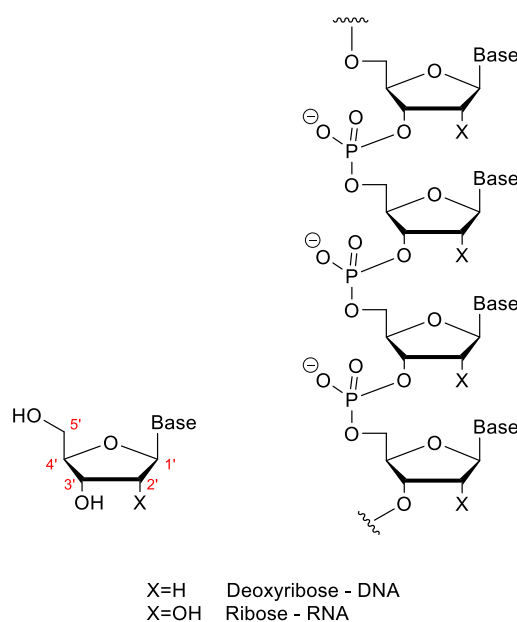


Figure 1 General structure of a nucleoside (left). The numbers in red refer to each atom of the sugar according to the rules set out by IUPAC-IUB.³ General structure of a nucleic acid chain (right).⁴

The atoms of the sugar moiety are differentiated from the atoms of the base by a superscript prime mark after the number.³ The nucleotide monomers that make up DNA are formed using the pentose sugar 2-deoxy-D-ribose whereas the nucleotides that make up RNA are derived from D-ribose.² The two pentoses differ only at C-2' where D-ribose has a hydroxyl group and 2-deoxy-D-ribose has a proton.² In both cases, a β -glycosidic bond attaches the sugar at C-1' to a base residue.² Both DNA and RNA chains are formed when their constituent nucleotide monomers polymerise to form a phosphodiester linkage as shown on the right of Figure 1.⁴ Work carried out by Klein and Thannhauser helped to establish that in DNA, phosphodiester linkages can only form between the 5'-OH of one nucleotide and the 3'-OH of another.² The extra hydroxyl group at C-2' of RNA nucleotides means that,

unlike DNA, they have the potential to form phosphate esters at this position. However, examples of 2'- to 5'- phosphodiester linkages are rare and typically RNA forms 3'- to 5'- phosphate esters just like in DNA.² The 2'-hydroxyl group present in RNA also has an effect on its overall stability. Whilst both DNA and RNA are liable to decompose in solution, the presence of the additional hydroxyl group in RNA greatly enhances the susceptibility of its phosphodiester bonds towards hydrolysis.⁴ Nevertheless, this vulnerability is somewhat offset by the additional stability the 2'-OH provides towards the glycosidic bond in RNA as compared to DNA where in its absence, the lability of the glycosidic bond towards hydrolysis is significantly increased.

1.1.1 The nucleobases

There are five naturally occurring bases which can be split into two classes known as the purines and the pyrimidines. Figure 2 shows the structure of each base and the numbering system used to refer to each atom.³

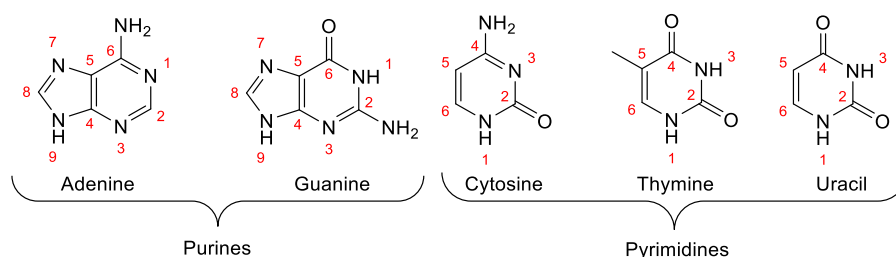


Figure 2 Structures of the five primary nucleobases found in DNA and RNA. The numbers in red refer to the atom label according to the rules set out by IUPAC-IUB.³

The purines consist of adenine (A) and guanine (G), bases which are common to both DNA and RNA. They are differentiated from the pyrimidines by their bicyclic structure and form glycosidic bonds at the *N*-9 position.² The pyrimidines are smaller, monocyclic bases and of the three, only cytosine (C) can be found in both DNA and RNA. Thymine (T), which only differs from uracil (U) by the presence of a methyl group at the *C*-5 position, is found in DNA whereas uracil is found in RNA. All three pyrimidines use *N*-1 to form glycosidic bonds to the anomeric carbon of the pentose sugar.² Given that they are the only part of a nucleotide unit to vary, the exact sequence of the nucleobases in a DNA or RNA chain is what encodes the biological and genetic information, and is what constitutes the primary structure of nucleic acids.

Each of the five nucleobases has the ability to display either keto-enol or amine-imine tautomerisation and in some cases, displays both. Through the use of multiple

spectroscopic techniques, it has been established that under physiological conditions all five nucleobases exist predominantly in their keto and amino forms, as drawn in Figure 2.² The ability of the nucleobases to tautomerize was one of the biggest obstacles that stood in the way of early pioneers in nucleic acid research establishing the exact structure and bonding behaviour of DNA and RNA. Right up until the early 1950's, it was widely believed that the bases existed in their enolic form and it was not until their keto forms were fully considered and appreciated that Watson and Crick were able to publish their seminal work on the structure of DNA in 1953.^{5,6}

As previously mentioned, the nucleobases are attached to C-1' of the sugar *via* a β -glycosidic bond. This means that the bases are on the same face of the sugar as the 5'-carbon. They lie virtually perpendicular to the plane of the sugar and it is approximated that they bisect the angle created by O4'-C1'-C2'.² There are two major conformations that the bases can adopt known as *syn* and *anti*.⁷ Of the two, the *anti* conformation is the most stable.⁸ In this conformer, the smaller part of the base moiety, *i.e.* H-8 in the purines and H-6 in the pyrimidines, lies above the sugar.² The *syn* conformation has the base orientated in such a way that the larger bulk of the bases, *i.e.* N-3 in the purines and O-2 in the pyrimidines, is located above the sugar which introduces unfavourable steric effects.²

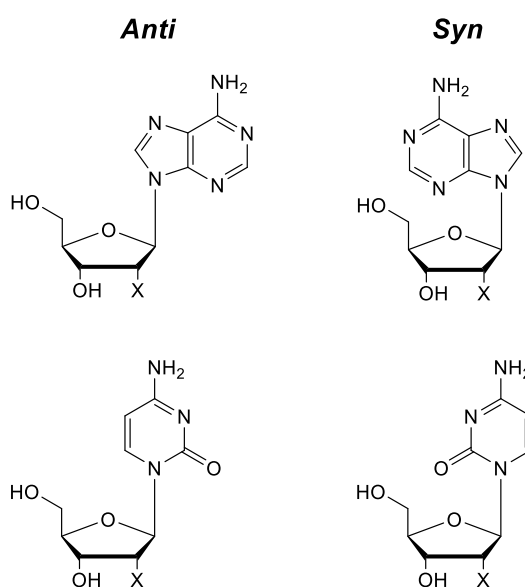


Figure 3 Top: *Anti* and *syn* conformations adopted by the purines, in this example adenine. Bottom: *Anti* and *syn* conformations adopted by the pyrimidines, in this example cytosine. X = H or OH.

It is due to these steric effects that nucleobases predominantly adopt the *anti* conformation. Whilst not unheard of, examples of pyrimidine nucleotides in the *syn* conformation are rarer than examples of purines.⁸ In fact, guanine has been found

to display a preference for the *syn* conformer under certain conditions. It is thought that guanine is able to form a stabilising electrostatic interaction between its amino group at C-2 and its 5'-phosphate which makes the *syn* conformer overall less unfavourable.⁸ Furthermore, the introduction of steric bulk at the C-8 position of guanine has been found to also result in the nucleobase showing a preference for the *syn* conformer. It has been determined that the presence of a halogen atom at C-8 of 2'-deoxyguanosine results in a destabilisation of the base pairing to 2'-deoxycytidine.⁹ This is due to a destabilisation of the *anti* conformer caused by steric interactions and the preference of the halogenated guanine to now adopt the *syn* conformation.⁹ These findings have important biological implications as several well-known DNA lesions contain modifications at the C-8 position of guanine and it is possible some of their mutagenic and carcinogenic properties are due to the increased steric bulk at this position causing destabilising effects. One such lesion is 8-nitroguanosine which is the subject of part of this thesis and will be discussed in greater detail in the next chapter.

1.1.2 The sugar conformation

It is both energetically and sterically unfavourable for the pentose sugar in nucleic acids to adopt a planar structure and so to help alleviate ring strain and unfavourable eclipsing steric interactions, it adopts a puckered conformation.⁷ There are two main conformations that it can adopt known as the envelope (E) and the twist (T).⁷ In the envelope conformation, four of the five ring atoms lie in the same plane with the fifth atom displaced. The twist conformation has two adjacent atoms displaced in opposite directions from the plane formed by the other three atoms. The conformations are described as *exo* when the atom displaced furthest from the plane is on the opposing face to the 5'-carbon and termed *endo* when it is on the same side.²

In work carried out to investigate the concept of pseudorotation in nucleosides and nucleotides by Altona *et al.*, it was established that nucleic acids largely adopt two conformations.¹⁰ These conformations are known as the *C2'-endo* and *C3'-endo*, both of which are derived from a twist pucker arrangement and are shown in Figure 4. Due to their geographic locations on the pseudorotational cycle, in addition to the shape formed by their C4'-C3'-C2'-C1' atoms, the conformations can also be referred to as South (S) and North (N) respectively.¹⁰

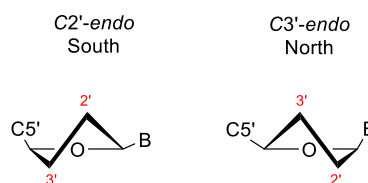


Figure 4 Two most commonly occurring nucleic acid sugar pucker conformations; *C2'*-endo (*S*) on the left and *C3'*-endo (*N*) on the right

The stability of both *S* and *N* conformations comes in the main from the gauche interactions (60° dihedral angles) between the $O-4'$ of the sugar ring and the oxygen atom at $C-3'$ as well as the additional oxygen at $C-2'$ in RNA.¹¹⁻¹³ In solution, interconversion between the *S* and *N* conformations is rapid due to the energy barrier separating them being relatively low (less than 20 kJ mol^{-1}).¹⁴ However, due to the reasons about to be discussed, both DNA and RNA show a preference for which conformation they adopt.

In the case of DNA, the *C2'*-endo conformation predominates. As DNA lacks the $2'$ -OH found in RNA, the gauche interaction between $O-4'$ and $O-3'$ has the largest contribution to its overall stability. As this is not present in the *N* conformation it preferentially adopts the *S* conformation instead.

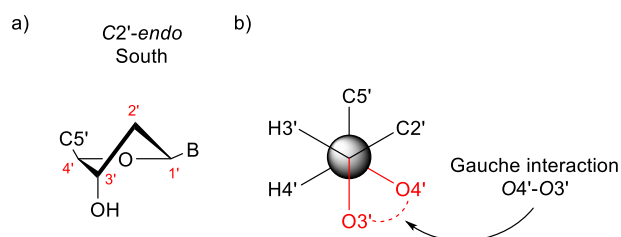


Figure 5 a) DNA sugar pucker b) Newman projection viewing along $C3'-C4'$ bond showing the key stabilising gauche interaction.

RNA on the other hand shows a preference to adopt the *N* conformation. The extra hydroxyl group it possesses at $C-2'$ introduces additional gauche interactions which makes both the *S* and *N* conformations equally stable. In order to understand RNA's preference for *C3'*-endo, an additional stereoelectronic effect has to be considered. The anomeric effect, although somewhat weaker than the gauche effect, provides extra stabilisation through donation of the lone pair of electrons on $O-4'$ into the σ^* orbital of the glycosidic bond.¹² This anomeric interaction is only possible in the *N* conformation where the nucleobase adopts a pseudo-axial position. This allows for a better overlap between its anti-bonding orbital and the lone pair of electrons at $O-4'$ than is possible in the *S* conformation. In addition, through adopting the *C3'*-endo conformation, RNA can form an intra-strand hydrogen bond between the $O-2'$ of one

residue and the O-4' of another.² This is not possible in the C2'-endo conformation and so again adds to its overall stability and preference for N.

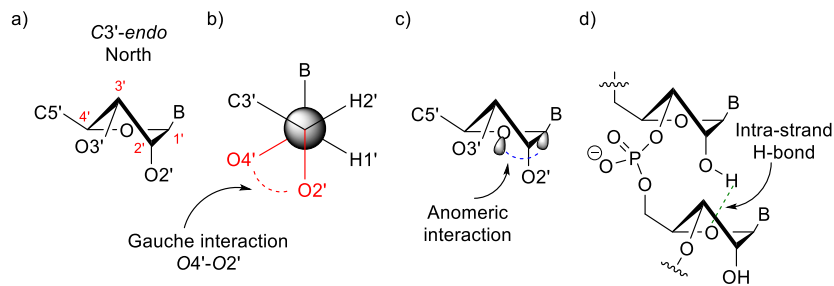


Figure 6 a) RNA sugar pucker b) Newman projection viewing along C1'-C2' bond showing a stabilising gauche interaction c) Anomeric interaction between O-4' lone pair and σ^* orbital of glycosidic bond d) Intra-strand hydrogen bond between O-2' of one residue and O-4' of another.

1.2 The secondary structure of nucleic acids

Nucleic acids are able to form secondary structures such as duplexes through a process known as base-pairing in which specific hydrogen bonds form between nucleobases. The combination of NH groups and lone pairs on their carbonyl oxygens means the nucleobases make both good hydrogen bond donors and good hydrogen bond acceptors.

The most common hydrogen bonding pattern found in nature is known as Watson-Crick pairing.⁵ In this pairing model, guanine specifically pairs to cytosine in both DNA and RNA by forming three hydrogen bonds whilst adenine pairs to thymine in DNA and uracil in RNA.⁵ In both A•T and A•U base-pairs, there are only two hydrogen bonds present meaning they are thermodynamically less stable than G•C pairs.^{2,5} Figure 7 shows the Watson-Crick hydrogen bonding pattern.

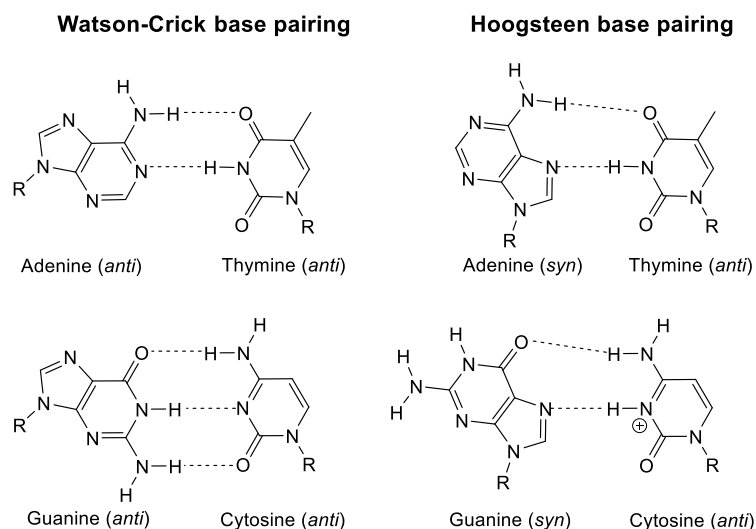


Figure 7 The two most commonly found base-pairing patterns; Watson-Crick (left) and Hoogsteen (right).

Also shown in Figure 7, is the Hoogsteen base-pairing pattern. This is an alternative hydrogen bonding pattern first reported by Karst Hoogsteen.^{15,16} This mode of pairing again has G pairing with C and A with T/U, but differs from the Watson-Crick model in the conformation the purine nucleobase adopts. In Hoogsteen base pairs, adenine and guanine adopt the *syn* conformer which results in these base pairs having quite different properties compared to Watson-Crick pairs. In the *syn* conformation, the purine nucleobases of Hoogsteen base pairs utilise the *N*-7 of their imidazole ring to form hydrogen bonds and for the G•C base pair, only 2 hydrogen bonds can form and it is required that cytosine is protonated. Hoogsteen base-pairs are much less common in nature than Watson-Crick pairs, but they can be found in some higher order multi-stranded structures.

The ability of the nucleobases to pair so specifically is what allows the formation of more complex nucleic acid structural motifs.

1.3 The double-helix

The double helical structure of DNA was first proposed in 1953 by Watson and Crick and is comprised of two anti-parallel, complementary strands of nucleic acids.^{5,6} The strands are bonded to each other by the aforementioned Watson-Crick base-pairing interactions and coil into a right-handed helix known as B-DNA. The helix orientates itself in such a way that the sugar-phosphate backbone is on the outside of the duplex and the nucleobase pairs are stacked inside. In this way, the helix is able to minimise electrostatic repulsions by having the hydrophobic nucleobases shielded by the negatively charged, hydrophilic sugar-phosphate backbone. This arrangement also contributes to the overall stability of the duplex. There are several key interactions that help to stabilise the structure of the double helix with one of the most important being the π - π stacking interactions that occur between adjacent base pairs.¹⁷ Further stability is imparted through solvent interactions between water molecules and the hydrophilic sugar-phosphate backbone. The exclusion of solvent water from the hydrophobic core created by the base stacking interactions also results in the gain of favourable entropy.²

There are two grooves present on the surface of nucleic acid double helices. They are known as the major and minor grooves and are formed by the gaps that open up as the two sugar-phosphate backbones coil around each other.² The size of these grooves depends on which helical conformation is adopted. Nucleic acid duplexes are capable of displaying structural polymorphism meaning that they can adopt different helical conformations depending on the surrounding environment

and the composition of their bases.¹⁸ It has been established *via* X-ray diffraction studies that a wide variety of conformations are able to be adopted by nucleic acid duplexes.¹⁹ Of these, there are three main forms that predominate known as A, B and Z, each of which will now be looked at in closer detail.

1.3.1 A-form helix

The A-form helix is a right-handed double helix that has its strands running anti-parallel to each other, bonded together through Watson-Crick base-pairing. It contains sugars that display the *C3'-endo* pucker making it the helical conformation most favoured and adopted by RNA duplexes.²⁰ Under conditions of low-humidity (high-salt) it is the conformation adopted by DNA duplexes as well.²¹ The A-form helix is characterised by having a major groove that is deep and narrow and a minor groove that is wide and shallow.²² The base pairs in A-form helices are displaced away from the helix axis which creates a hollow core of approximately 3 Å running through the middle of the helix.² Both A-RNA and A-DNA display similar structural parameters with both helices containing 11 base pairs per turn. A-form DNA was first revealed in diffraction patterns of DNA fibres obtained by Rosalind Franklin and Raymond Gosling under conditions of low humidity in the early 1950's.^{21,23} Since then, the development of technology has allowed single crystal x-ray structures of A-form DNA to be obtained.²⁴ From these single crystal structures, a vast amount of structural information about A-helices has been determined (Table 1). Analysis of DNA using single crystal x-ray structures was initially held back by difficulties in isolating pure, homogenous DNA samples, but advancements in oligonucleotide syntheses during the 1970's means obtaining short, uniform oligonucleotides is now relatively straight forward.²⁵

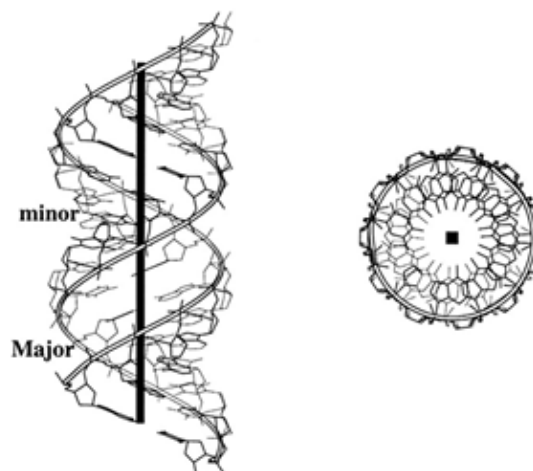


Figure 8 A-form helix viewed from the side (left) and down the helix axis (right) adapted from reference 26.²⁶

1.3.2 B-form helix

The B-form is the most commonly observed helix and is the one adopted by DNA under physiological conditions.² It shares some similarities with the A-type helix in that they are both right-handed helices whose nucleic acid strands run anti-parallel to each other and are bonded by Watson-Crick hydrogen bonds. However, in B-form helices, the sugars adopt the *C2'-endo* pucker and the base-pairs, of which there are 10 per turn, are not displaced but rather sit directly on the helix axis.^{19,27} This results in the B-type helix having a narrower diameter compared to the A-type and major and minor grooves that are more similar in depth.²⁷ Like the A-form, the structure of B-form DNA was first revealed in diffraction patterns of DNA fibres obtained by Franklin and Gosling.^{21,23} The B-form was found to exist under conditions of high humidity and the data from the diffraction patterns helped enable Watson and Crick to devise their model of the DNA double helix.⁵ Similarly again to the A-form, many single crystal x-ray structures of B-form DNA are now in existence.²⁵ Dickerson *et al.* were amongst the first to report single crystal diffraction data for B-form DNA which was obtained using the dodecamer sequence CGCGAATTCGCG.²⁸ These single crystal diffraction analyses have not only confirmed the double helix model put forward by Watson and Crick, but also revealed a wealth of data regarding the structural parameters of the helix, with some of the key parameters summarised in Table 1.

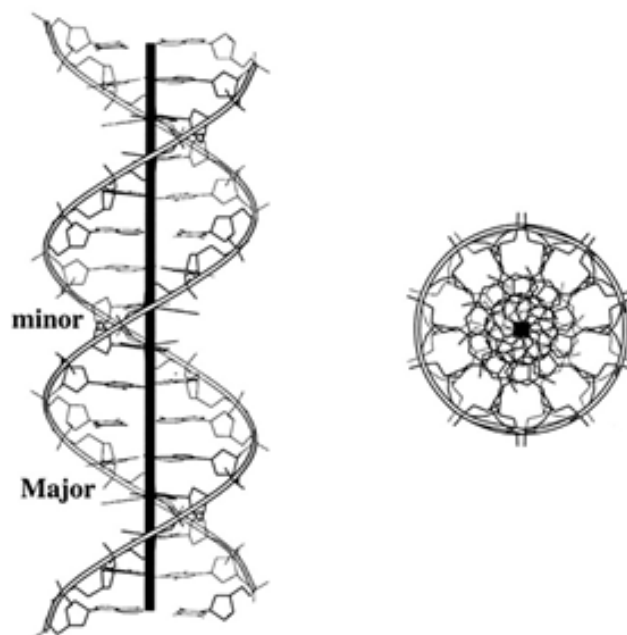


Figure 9 B-form helix viewed from the side (left) and down the helix (right) adapted from reference 26.²⁶

1.3.3 Z-form helix

The Z-type helix differs significantly from both A and B helices. Its structure was first elucidated in 1979 by Rich *et al.* using single crystal x-ray diffraction and found to be a left-handed, anti-parallel duplex.²⁹ Although not a requirement for its formation, Z-form helices are most often observed for alternating purine-pyrimidine sequences, typically poly(dG-dC), and high salt concentrations are known to help stabilise them.^{2,30} Their sugar-phosphate backbone adopts an unusual zig-zag appearance which is the origin of the helix name. The reason for this zig-zagging appearance is due to the unusual combination of conformations its nucleobases and sugars adopt. Its purine nucleotides have been found to have their nucleobases in the *syn* conformation and to display a *C3'-endo* pucker whereas its pyrimidines have their nucleobases in the more usual *anti* conformation and adopt a *C2'-endo* sugar pucker.²⁵ It is much less commonly observed than either A or B form helices and its exact biological significance is still not yet fully understood.³¹ Work on Z-form helices has discovered that negative supercoiling of DNA is known to stabilise Z-DNA helix formations and so a biological role for Z-DNA is most commonly associated with processes that utilise this phenomenon, such as transcription.³² The findings that several proteins can bind specifically to Z-DNA also points towards a biological role for this helix.³³⁻³⁵

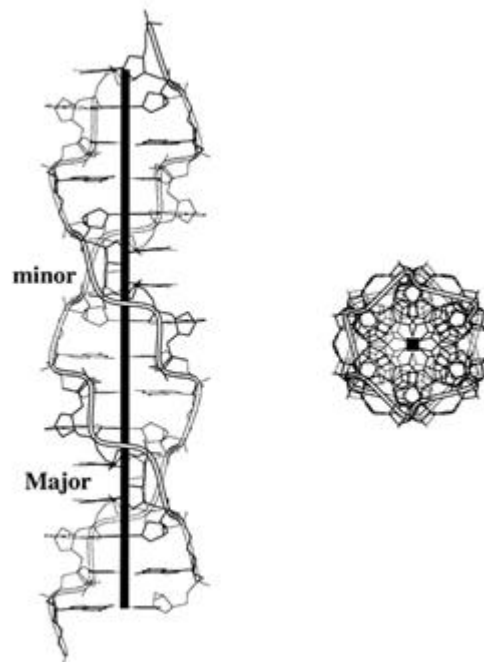


Figure 10 Z-form helix viewed from the side (left) and down the helix axis (right) adapted from reference 26.²⁶

Table 1 compares and summarises some of the key structural parameters of the helices discussed.^{2,27}

Table 1 Key structural parameters of the most common nucleic acid helices.

Parameter	A-DNA	A-RNA	B-DNA	Z-DNA
Helix sense	Right-handed	Right-handed	Right-handed	Left-handed
Sugar pucker	<i>C3'-endo</i>	<i>C3'-endo</i>	<i>C2'-endo</i>	Purines: <i>C3'-endo</i> Pyrimidines: <i>C2'-endo</i>
Nucleobase conformation	<i>Anti</i>	<i>Anti</i>	<i>Anti</i>	Purines: <i>Syn</i> Pyrimidines: <i>Anti</i>
Base-pairs per turn	11	11	10	12
Twist per base pair (°)	32.7	32.7	36	-9, -51
Helix diameter (Å)	23	~23	20	18
Axial rise per base pair (Å)	2.56	2.8	3.3-3.4	3.7

1.4 Multi-stranded structures

Through the many studies carried out into elucidating the structure of DNA, it has been established that several higher order structures exist that involve bonding of multiple DNA strands. These multi-stranded DNA structures are of interest due to their potential to be exploited as therapeutic agents. The three most widely studied multi-stranded structures are the triple helix, the *i*-motif and the G-quadruplex.

1.4.1 Triple helices

In DNA triple helices, a third strand of DNA is bonded *via* Hoogsteen hydrogen bonding to a B-form DNA duplex.³⁶ The bonding of the third strand has been found to be weak compared to the Watson-Crick bonding present in duplexes, but the binding of the third strand is sequence specific which is why triple helices are of therapeutic interest.³⁷ Triple helices can be classified as either intramolecular or intermolecular depending on the origin of the third strand. Intramolecular triple helices form from polypurine-polypyrimidine duplexes that contain mirror repeat sequences.³⁶ Under certain conditions, such as low pH, these duplexes can dissociate.³⁸ Intramolecular triple helices can then form if one of the single strands folds back on one of the mirror repeats and binds to it.³⁸ Intermolecular triple helices form when a different strand of DNA bonds to a duplex. The third strand is typically a triplex forming oligonucleotide (TFO) and it binds at the major groove of the DNA

duplex.³⁷ It is restricted to binding to the duplex at sites which have runs of purines on one strand and pyrimidines on the other.³⁹ The TFO will always bind to the purine rich strand of the duplex regardless of whether it is a polypurine or polypyrimidine molecule.³⁶ Polypurine TFO's bind in an antiparallel sense to the purine strand of the duplex using reverse Hoogsteen bonding whereas polypyrimidine TFO's bond in a parallel fashion using regular Hoogsteen hydrogen bonding.³⁶ The ability of TFO's to bind to DNA duplexes with high affinity and a high degree of specificity means they possess great potential to be used for genetic manipulation.

1.4.2 *i*-Motifs

i-Motifs, short for intercalated motifs, are four stranded DNA structures that form in regions rich in cytosine.⁴⁰ The structure is comprised of two parallel DNA duplexes held together in an antiparallel orientation by intercalated base pairs.⁴⁰ The base pairs that hold the structure together are C•C⁺ pairs in which one of the bases is protonated (Figure 11).⁴⁰

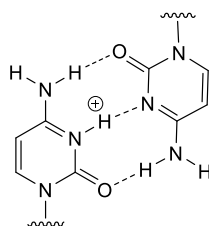


Figure 11 Structure of a C•C⁺ base pair.

The structures were first discovered by Gehring *et al.* for the sequence d(TCCCCC).⁴¹ *In vitro* investigations showed they are particularly stable under acidic conditions, becoming less stable as the pH rises to physiological pH or higher.⁴² For this reason, for many years a biological role for *i*-motifs was uncertain, but in recent years studies have shown that certain conditions such as molecular crowding and negative superhelicity during transcription can induce formation of the structure at physiological pH.⁴³⁻⁴⁵ Very recently, a study by Christ *et al.* showed that a human antibody fragment can recognise *i*-motif structures with high selectivity and affinity.⁴⁶ They then used this fragment in immunofluorescent staining experiments to provide the first direct evidence that *i*-motif structures are present in the nuclei of human cells.⁴⁶ Most work in the early years following their discovery focused on using *i*-motifs for applications in DNA nanotechnology.⁴⁰ But, with the discovery that they can exist for certain *in vivo*, a whole range of other biological roles for this structure can now be explored.

1.4.3 G-Quadruplexes

Of all the higher order nucleic acid structures, probably the most widely studied is the G-quadruplex. They form in regions of nucleic acid sequences that are rich in guanine.⁴⁷ The single stranded guanine rich strands can fold to form four stranded G-quadruplexes either intermolecularly or intramolecularly.⁴⁸ G-Quadruplexes arise through the stacking of G-quartets, which form when four guanine bases assemble into a tetramer held together by Hoogsteen hydrogen bonds (Figure 12).⁴⁹

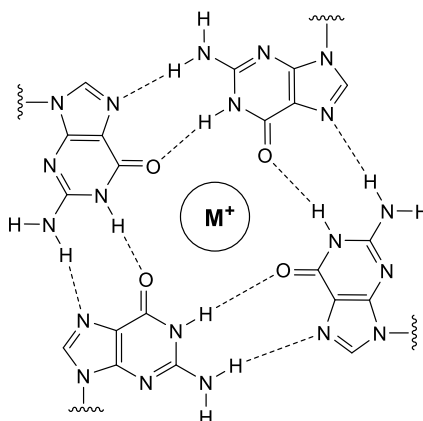


Figure 12 Structure of a G-quartet with a monovalent cation at its centre. The dashed bonds are Hoogsteen hydrogen bonds.

The G-quadruplex formed when two or more G-quartets self-stack is stabilised by the presence of a monovalent cation.⁴⁹ The cation, which is typically located between two quartets, helps to reduce the electrostatic repulsion in the centre of the tetramer that is created by the carbonyl oxygen atoms and so provides a stability to the overall structure.⁵⁰ Their stability and ability to form at physiological pH is what has attracted so much attention to G-quadruplexes.

Simple computational algorithms have predicted that the human genome has the potential to form around 376 000 G-quadruplex structures, although more sophisticated high-resolution sequencing-based methods have predicted almost double this number can form.^{51,52} Their presence in human cells was confirmed through visualisation upon binding of a highly specific DNA G-quadruplex antibody.⁵³ Further computational analysis has revealed that G-quadruplexes or sequences capable of forming G-quadruplexes are highly prevalent in certain key regulatory regions of the human genome that are involved in processes such as transcription regulation and translation.⁴⁹ This has made G-quadruplexes an attractive target for potential new therapeutic treatments, particularly the targeting of G-quadruplexes by small molecules for use in anti-cancer therapies.⁴⁹

Telomeres are repetitive sequences of nucleotides found at the ends of chromosomes that are required to protect them from degradation and fusing with other neighbouring chromosomes.⁵⁴ They are made up of a double stranded region and a single stranded overhang.⁵⁵ The overhang is guanine rich and capable of forming G-quadruplexes.⁵⁵ Telomeric repeats are lost during each cycle of cell division causing shortening of the telomere which ultimately leads to cell death.⁵⁵ However, the reverse transcriptase ribonucleoprotein telomerase helps to lengthen and maintain telomeres through addition of telomeric repeats.⁵⁶ It is inactive in most somatic cells, but has been found to be upregulated in 80-90% of tumour cells which can lead to cell immortality of malignant cells.^{55,56} The formation of G-quadruplexes has been found to block the activity of telomerase and so much work has gone into trying to develop small molecules that can bind and stabilise G-quadruplex structures as a means of developing new anti-cancer therapies.^{48,49}

Chapter 2

Reductive denitration as a potential repair mechanism of the DNA lesion 8-nitroguanine

Chapter 2 Introduction

The aim of this project was to investigate whether it was possible to carry out a reductive denitration reaction on 8-nitroguanine, thereby converting it directly back to guanine. Such a conversion would suggest that a direct repair of this lesion was a chemically feasible process. It is therefore important to first discuss how the lesion occurs, why the lesion is such a problem, the established mechanisms of direct DNA repair and the literature precedent for reductive denitration reactions.

2.1 How does 8-nitroguanine form?

The DNA lesion 8-nitroguanine is strongly associated with inflammation-related carcinogenesis.^{57,58} It has been established that areas of chronic inflammation are at high risk of becoming cancerous, due in the main to the high amounts of reactive oxygen species (ROS) and reactive nitrogen species (RNS) concentrated in these areas.⁵⁹⁻⁶¹ The link between inflammation and cancer was first suggested back in 1863 by Rudolf Virchow who identified white blood cells (leucocytes) in samples of carcinogenic tissues.⁶² In the years since, there has been much research to confirm the existence of such a link and there is now a general acceptance that there is a connection between inflammation and carcinogenesis. One of the most compelling pieces of evidence to support the link comes from the finding that inflammatory cells and mediators are present in the majority of tumours, if not all, irrespective of what triggered the tumour to develop.⁵⁹ It has been estimated that between 15-25% of all cancer cases are as result of chronic inflammation, or, infections that can induce and trigger an inflammatory response.^{59,62,63} Table 2 summarises some of the cancers known to be associated with inflammation.⁶²⁻⁶⁴

Table 2 Causes of chronic inflammation and their associated cancers.

Cancer Site	Inflammatory Cause
Stomach	<i>Helicobacter pylori</i> infection
Skin (Melanoma)	Sunburn caused by exposure to UV light
Lungs	Repeated exposure to tobacco smoke
Cervix	Human papilloma virus
Pancreas	Chronic pancreatitis caused by tobacco, alcohol and genetics
Bladder	Schistosomiasis
Liver	Hepatitis

Inflammation is the natural physiological response to infection or tissue injury that initiates the healing process. It can be classified into one of two categories known

as acute or chronic inflammation. Acute inflammation is the first response of the host which involves activation and migration of plasma and leucocytes to the damaged tissue.⁶⁵ Medzhitov described a generic inflammatory pathway which comprises of four main components, namely inducers, sensors, mediators and effectors.⁶⁶ Inducers can be categorised as either exogenous or endogenous. Stressed, damaged or malfunctioning tissues induce an endogenous inflammatory signal.⁶⁶ Exogenous inducers can be subdivided further into either microbial or non-microbial inducers. As the name suggests, microbial inducers are derived from invading microorganisms whereas examples of non-microbial inducers include allergens, irritants and foreign bodies.⁶⁶ Tissue-resident immune cells, which include macrophages, dendritic cells and mast cells, contain receptors on their surface known as pattern recognition receptors (PRRs).⁶⁶ These receptors act as sensors and are able to recognise the signals given off by the inducers and stimulate cellular mediators.⁶⁷ The amount and combination of inflammatory mediators produced depends on the exact nature of the inducer.⁶⁸ The purpose of the mediators is to act on the effectors, which in most cases are the target tissues and cells, to produce a response that restores the affected tissue to homeostasis and removes the inducer.^{66,68} Most commonly, the inflammatory response involves release of potent oxidising and highly reactive species that attack and destroy the inducers of inflammation.⁶⁹ A successful response to acute inflammation is deemed as removal of the inflammatory inducer followed by repair of the damaged tissue and termination of the immune response.⁶⁶

However, problems can arise when the inflammatory response fails to eliminate the inducer, termination of the immune response fails or continual exposure to an irritant results in recurring bouts of inflammation. These issues lead to chronic inflammation which can ultimately lead to carcinogenesis. As discussed, when an inflammatory response is required, mediators induce the release of highly reactive species in order to eliminate the inducer. Unfortunately, these highly reactive and potent species do not discriminate between the host and the pathogen and so there is invariably collateral damage to otherwise healthy host cells.^{66,70} This collateral damage is minimised in acute inflammation where overall the inflammatory response is short lived and well-regulated, but in chronic inflammation the prolonged exposure to RNS and ROS can have serious implications. ROS and RNS are capable of causing both oxidative and nitrative damage to DNA which can lead to mutagenesis and carcinogenesis.^{69,71} Nitric oxide (NO) is one such RNS and its

chemistry at sites of inflammation makes it the most relevant with regard to the work carried out in this thesis.

NO is a key signalling molecule that participates in several physiological processes such as vasodilation, neurotransmission and the aforementioned inflammatory response.⁷² It is produced endogenously by the enzyme nitric oxide synthase (NOS) which displays three isoforms.⁷² Two of these isoforms are Ca^{2+} dependent and are known as the endothelial form (eNOS) and the neuronal form (nNOS) whereas the third form is Ca^{2+} independent and is known as the inducible form (iNOS).⁷² The two calcium dependent isoforms are constitutive enzymes that act transiently and when activated emit only nanomolar levels of NO.⁷³ The inducible form is what participates in the immune response and is capable of producing highly concentrated bursts of NO.⁷² At sites of chronic inflammation there is an upregulation of iNOS which is problematic as excess NO is produced for extended periods of time.⁷³ In the immune response, NO is produced as a cytotoxic agent to attack and remove the invading pathogen, but when it is present for prolonged periods it is capable of causing significant damage to neighbouring cells.⁷⁴ Oxidative and nitrative damage caused to DNA at sites of chronic inflammation are rarely due to NO itself as individually it is fairly unreactive towards DNA.⁵⁷ However, several RNS are formed through reaction of NO with other highly reactive species present and it is these RNS that can cause DNA lesions.

Under aerobic conditions, NO is known to autoxidise to produce the nitrosating agent N_2O_3 which can cause DNA deamination and alkylation.^{57,75} Deamination of DNA results in the conversion of cytosine into uracil, guanine into xanthine and adenine into hypoxanthine which can lead to DNA mispairing and mutations.⁷³ Deamination caused by N_2O_3 occurs through direct attack of N_2O_3 at nucleobases containing an amine moiety.⁵⁷ Alkylation occurs when N_2O_3 reacts with a secondary amine to form a *N*-nitrosamine which can alkylate DNA and cause mutagenic lesions such as O^6 -alkylguanine.⁵⁷

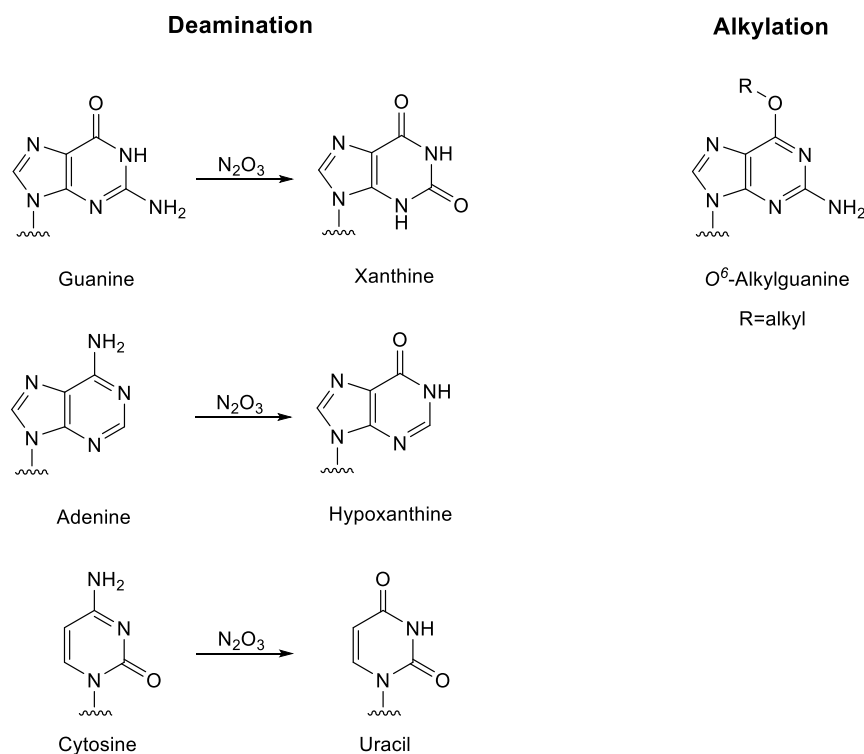
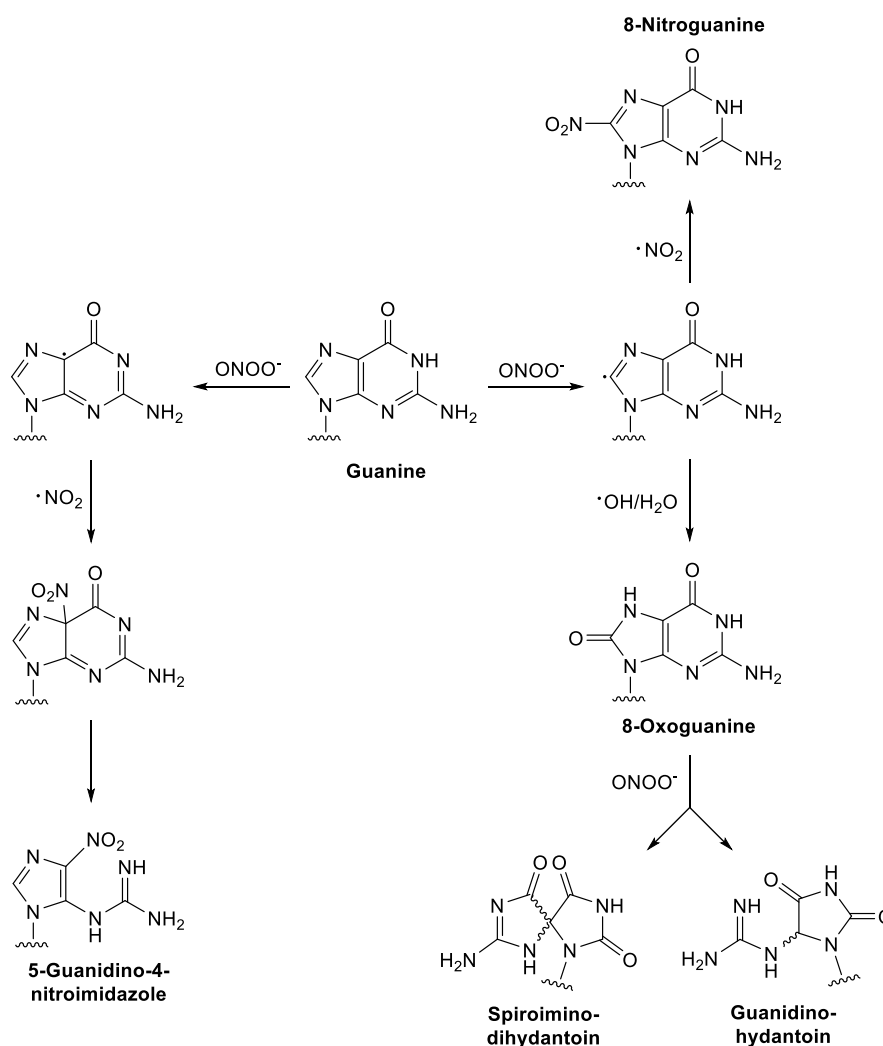


Figure 13 Deamination and alkylation products formed through reaction of the RNS N_2O_3 with DNA.

As mentioned previously, at sites of chronic inflammation, numerous highly reactive species are present. The DNA lesion under investigation in this thesis, 8-nitroguanine, forms as a result of the reaction between NO and the ROS, superoxide ($O_2^{\cdot-}$).⁷⁶ Superoxide is generated by activated macrophages and reacts with NO to form the highly reactive species peroxynitrite ($ONOO^{\cdot}$) in a reaction that is only limited by diffusion.^{60,77} Of all the nucleobases, it has been established that guanine has the lowest oxidation potential and so is the most susceptible to attack by oxidising and nitrative species.^{78,79} Peroxynitrite therefore readily oxidises guanine to first form the widely studied DNA lesion, 8-oxoguanine.⁷⁹ The oxidation process does not stop there however, as it has been estimated that 8-oxoguanine is approximately 1000-fold more reactive than guanine and so it can undergo further oxidation to produce products such as guanidinohydantoin and spiroiminodihydantoin.⁵⁷ 8-Nitroguanine forms as a result of peroxynitrite decomposition. When peroxynitrite decomposes, it can either be proton catalysed or CO_2 catalysed. In the proton-catalysed decomposition, $^{\cdot}OH$ and NO_2^{\cdot} form whereas in the CO_2 decomposition pathway $CO_3^{\cdot-}$ forms in addition to NO_2^{\cdot} .⁷³ Regardless of the decomposition pathway, one of the products formed is NO_2^{\cdot} which nitrates guanine to form 8-nitroguanine as well as 5-guanidino-4-nitroimidazole.⁵⁷ Scheme 1 summarises the reactions that can take place between peroxynitrite and guanine.



Scheme 1 Major products formed through reaction of peroxynitrite with guanine.⁵⁷

2.2 Why is the 8-nitroguanine lesion a problem?

The formation of 8-nitroguanine at sites of chronic inflammation is problematic as it is a mutagenic lesion.⁸⁰ The presence of the electron withdrawing nitro group at C-8 makes the lesion chemically unstable. The lability of the glycosidic bond is greatly increased which can result in the spontaneous release of 8-nitroguanine and formation of an abasic site, as shown in Figure 14.

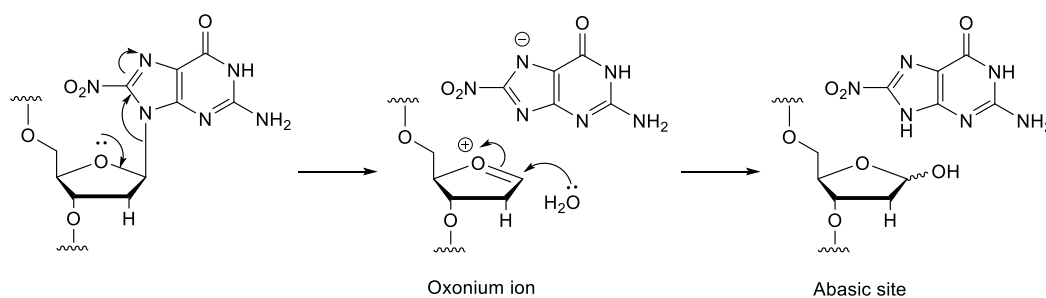


Figure 14 Spontaneous loss of 8-nitroguanine to form an abasic site.

As an individual nucleoside, 8-nitrodeoxyguanosine has been estimated to depurinate with a half-life of 1 hour under physiological conditions although the half-life increases slightly to ~4 hours when the lesion is part of a DNA strand.^{79,81} The stability of 8-nitroguanine is much improved in RNA where after 6 hours, only 5% depurination was found to have occurred.⁸² Its increased stability in RNA is due to the presence of the C-2' OH group which destabilises the oxonium ion making the depurination process much less favourable.

Abasic sites are known to pair preferentially with adenine during DNA synthesis and so loss of 8-nitroguanine results in a G→T transversion mutation.⁸³ Due to its inherent instability, incorporation of the lesion into chemically synthesised oligodeoxynucleotide sequences has proved challenging and so relatively little is known about its base-pairing preferences. Suzuki *et al.* sought to establish the miscoding potential of 8-nitroguanine by preparing an oligodeoxynucleotide strand containing one 8-nitroguanine adduct.⁷⁵ The adduct was prepared photochemically and the oligo strand was then used as a template in primer extension reactions to investigate whether polymerisation would extend past the lesion and if so, determine what deoxyribonucleotide was incorporated opposite.⁷⁵ Four different mammalian polymerases, α , β , η , and $\kappa\Delta C$ were tested. It was found that when polymerases α and β catalysed the reaction, the majority of syntheses were retarded opposite the lesion or in some cases one base prior.⁷⁵ When synthesis was able to extend past the lesion, cytosine was preferentially incorporated, but pairing with adenine was also observed.⁷⁵ For η and $\kappa\Delta C$ polymerases, primer extension reactions were extended past the 8-nitroguanine lesion, but broad miscoding specificity and high miscoding frequency were observed.⁷⁵ The results showed that whilst both polymerases were able to incorporate the correct base cytosine opposite the lesion, a significant amount of misincorporation also took place. Primer extension reactions using polymerase η showed adenine was incorporated opposite the lesion with almost the same probability as the correct base cytosine and polymerase $\kappa\Delta C$ actually displayed a preference for incorporation of adenine across from 8-nitroguanine.⁷⁵ The results obtained from this study show the mutagenic nature of 8-nitroguanine as they suggest the lesion displays a preference for base-pairing with adenine which would again induce a G→T transversion. However, when considering the results of this study it should be noted that the template-primer system used in the experiments readily releases 8-nitroguanine to leave abasic sites. As discussed, abasic sites show a preferences for pairing with adenine so it is

possible that the polymerases were pairing adenine with an abasic site although the authors do state that observed depurination was minimal (less than 0.8%).⁷⁵

A further study by Bhamra *et al.* carried out similar primer extension experiments, but rather than use 8-nitrodeoxyguanosine in the oligodeoxynucleotide, the more stable RNA analogue, 8-nitro-2'-O-methylguanosine, was used instead.⁸⁴ The use of 8-nitro-2'-O-methylguanosine rather than the natural lesion was rationalised as its glycosidic bond is more stable due to the 2'-OMe group which helps to minimise depurination and it is more readily incorporated into oligodeoxynucleotides by conventional chemical methods.⁸⁴ For the primer extension reactions in this study, two polymerases were used. The polymerase AMV-RT was chosen as it is known to extend past 2'-OMe groups and DNA polymerase β was also investigated to allow for a direct comparison to the results of Suzuki.⁸⁴ The results obtained showed that both polymerases stalled significantly when they reached the 8-nitro lesion.⁸⁴ The AMV-RT polymerase showed a preference for incorporating the correct base cytosine across from the lesion, but the results using polymerase β were more in agreement with Suzuki's findings of 8-nitroguanosine displaying a base-pairing preference for adenosine.^{75,84} Polymerase β in Bhamra's study showed a preference for incorporation of adenine opposite the lesion, although in a different quantitative ratio to that reported by Suzuki for polymerase β .^{75,84} Possible reasons for the difference in observed ratios could be due to the use of two different 8-nitro analogues and the use of different pH's by the two different research groups.

Bhamra *et al.* also carried out some thermal melting studies to try and determine what nucleotide formed the most stable pairing opposite 8-nitroguanine.⁸⁴ They again used 8-nitro-2'-O-methylguanosine as a model compound for their study. They found the most stable pairing to be 8-nitro-G•G.⁸⁴ The thermal melting temperature (T_m) was found to be significantly lower for the 8-nitro-G•C pair compared to the control G•C pair.⁸⁴ The introduction of the nitro group was also found to destabilise the G•T pair, but had no effect on the G•A pair.⁸⁴ NMR studies determined that the presence of the nitro group at C-8 induces a change in conformation about the glycosidic bond from *anti* to *syn*.⁸⁴ This means that in the 8-nitroguanine lesion, it is the Hoogsteen face rather than the Watson-Crick face that takes part in hydrogen bonding. The Hoogsteen face forms only two hydrogen bonds to an opposite nucleotide which helps to explain why the 8-nitro-G•C pair is less stable than the G•C pair which forms three Watson-Crick hydrogen bonds. Despite the finding that the 8-nitro-G•G pair is the most stable, the polymerase study indicated that adenosine was preferentially added opposite the lesion.⁸⁴ This

is thought to be because the *syn*-8-nitroG•G pair is a poor genetic match for the natural Watson-Crick pairing and so the polymerases select against it. Further work on the subject is required to determine why exactly adenosine seems to be the nucleobase of choice incorporated opposite the lesion.

2.3 DNA repair

DNA repair and the mechanisms by which it takes place are areas of nucleic acid chemistry that have been extensively researched and continue to be of great interest. A full discussion on the topic goes beyond the scope of the work carried out in this thesis, but the following section gives an overview of some of the main mechanisms of repair.

2.3.1 Excision repair

It has been estimated that $\sim 10^5$ DNA lesions occur in a human cell per day and so it is of the utmost importance that repair pathways exist in order to maintain the genome integrity.⁸⁵ Several different DNA repair mechanisms are known to exist, with some acting on specific lesions whereas others are able to act on and repair a wider, more general range.

One of the most important mechanisms of DNA repair is excision repair. There are two main types of excision repair known as base excision (BER) and nucleotide excision (NER). In BER, a single damaged nucleobase is repaired whereas NER acts on bulkier lesions which are often formed by radiation and chemical exposure.⁸⁶ BER utilises a DNA glycosylase to act on the damaged base by cleaving the glycosidic bond to leave an abasic site.⁸⁷ Various DNA glycosylases are known to exist that can recognise lesions caused by damage such as deamination, methylation and oxidation.⁸⁷ Glycosylases can either be monofunctional or bifunctional. Monofunctional glycosylases act simply by cleaving the glycosidic bond of a lesion whereas bifunctional glycosylases can carry out AP lyase activity as well.⁸⁵ Following removal of the damaged base, an AP endonuclease then creates an incision in the sugar-phosphate backbone at the abasic site, unless, one has already been made by a bifunctional glycosylase.^{86,87} At this stage, BER can either carry out a short patch repair or a long patch repair. In short patch repair, the abasic sugar is then removed by a DNA polymerase, typically polymerase β , which concurrently fills in the gap, followed by ligation of the nick in the DNA strand by a ligase.^{86,87} In long patch repair, following incision by an AP endonuclease, DNA polymerase δ/ϵ acts with a clamp loading factor and

processivity factor to produce a flap of nucleotides, typically 2-10 nucleotides long.^{85,86} This flap is then removed using an endonuclease and polymerase δ/ϵ synthesises a new strand which can then be ligated with DNA ligase.^{86,87}

NER typically repairs lesions that induce helical distortions in the DNA structure, such as UV-induced dimer lesions, although it is known to have a broad substrate acceptability.^{85,88} It is a highly conserved process, known to be used by both prokaryotes and eukaryotes.⁸⁸ NER in eukaryotes requires a combination of more complex enzymes and repair factors than is required for NER in prokaryotes, but overall the general principle is very similar.⁸⁸ It primarily involves (i) recognition of the damaged site, (ii) dual incisions that bracket the lesion and form a 24-32 nucleotide oligomer in eukaryotes or a 12-13 nucleotide oligomer in prokaryotes, (iii) removal of the lesion-containing oligomer, (iv) gap-filling synthesis and finally ligation to restore the DNA strand.^{85,86,88}

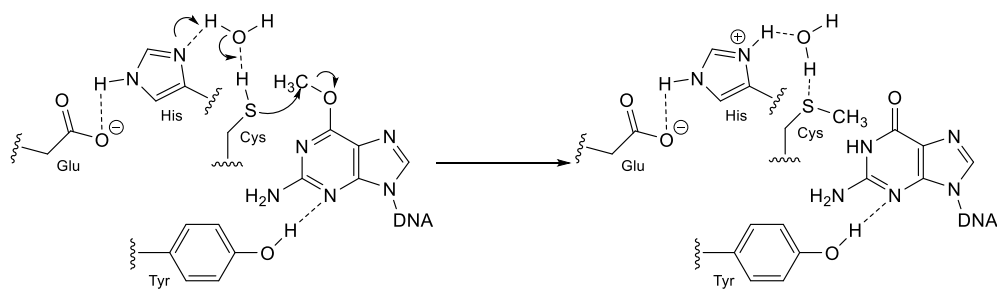
Base-mismatch repair (MMR) is another mechanism of correcting damage to DNA and broadly falls into the category of an excision-type repair. MMR primarily looks for and repairs base-base mismatches and nucleotide insertions/deletions that occur during DNA replication.^{89,90} It is a highly conserved process and much of the current understanding of the way MMR works in man is derived from studies on *Escherichia coli*.⁹¹ For MMR to be carried out successfully, it must first recognise the mismatched pair and then direct the repair proteins to the misincorporated base on the daughter strand.⁹² MMR is thus strand specific as it carries out repairs on the daughter strand rather than the template strand of DNA synthesis.⁹² *E. coli* uses three key proteins known as MutS, MutL and MutH to carry out MMR, each of which has a human homolog.^{89,92} MutS and MutL are required to recognise the mismatch and MutH functions as an endonuclease that creates the incision that acts as a starting point for excision of the mismatched base.^{89,90} A helicase is then used to unwind the DNA from the nick site past the mismatch followed by digestion of the error-containing unwound daughter strand by an exonuclease.⁹¹ Like BER and NER, MMR also requires a DNA polymerase to then fill in the gap and finally a ligase to seal the nick.⁸⁹

2.3.2 Direct repair

Relatively few DNA lesions are known to be repaired by direct chemical reversal, but the repair of *O*⁶-alkylguanine lesions and the repair of UV-induced pyrimidine dimer lesions are two examples. Pyrimidine dimers are repaired directly in bacteria using the process of photoreactivation catalysed by a photolyase.⁹³ Photolyases

contain two non-covalently bound chromophore cofactors that are essential to their catalytic ability.⁹³ The cofactor FAD is present in all photolyases whereas the other cofactor is either methenyltetrahydrofolate (MTHF) or 8-hydroxy-7,8-didemethyl-5-deazariboflavin (8-HDF) depending on the class of enzyme.⁹³ The mechanism of action first requires the enzyme to recognise the dimer lesion and bind to it. Exposure to blue light is then required for activation as absorption of a blue light photon by either MTHF or 8-HDF is essential for providing the energy required for the process.⁹³ The excitation energy absorbed from the blue photon is then passed to the flavin cofactor which transfers an electron to the pyrimidine dimer to form a radical.⁹³ This dimer radical is highly unstable and so decays quickly, splitting to form the two original pyrimidines while concomitantly transferring an electron back to the flavin cofactor so it can regenerate.⁹³ This process cannot take place in humans as photolyase is no longer functioning in man and so UV-induced dimer lesions are repaired as mentioned previously by NER.

O^6 -Alkylguanine lesions are considered mutagenic as the presence of the alkyl group on the nucleobase can result in mispairing of the lesion with thymine and thus a transition of G→A can occur.⁹⁴ Repair of the lesion takes place by means of a direct chemical reversal mediated by the enzyme O^6 -alkylguanine-DNA alkyltransferase (AGT).^{95,96} The enzyme acts sacrificially as once it carries out the dealkylation, it cannot be regenerated.^{94,97} The enzyme contains a highly conserved amino acid sequence consisting of Pro-Cys-His-Arg which is responsible for carrying out the catalysis.⁹⁵ The SH group of the Cys residue acts as an alkyl acceptor by carrying out an S_N2 displacement of the alkyl group as shown in Figure 15.⁹⁵ The S-alkylcysteine that forms then renders the enzyme inactive which is why it is said to act sacrificially. The enzyme most commonly carries out demethylation reactions on O^6 -methylguanine, but it has also been found to dealkylate ethyl-, 2-hydroxyethyl- and 2-chloroethyl guanines.⁹⁵



DNA is repaired (dealkylated) by sacrificial enzyme AGT

AGT= O^6 -alkylguanine-DNA transferase

Figure 15 Cartoon showing repair of alkylated guanine by AGT.

It is assumed that 8-nitroguanine lesions are repaired *via* one of the excision repair mechanisms given that they readily depurinate to form abasic sites which can be detected by these methods. As the lesion has a relatively short half-life and is prone to depurination, it is possible that the 8-nitroguanine remains undetected until an abasic site forms and the abasic site is what is repaired. However, repair of the lesion by direct chemical reversal is an attractive possibility as it would directly regenerate guanine from the lesion.

2.4 Reductive denitration

Reduction of a nitro group typically yields an amine. Indeed, when 8-nitroguanosine has been exposed to classical reducing agents such as zinc-HCl and sodium hydrosulfite, the expected 8-aminoguanosine has been found to form.^{81,98} The formation of 8-aminoguanosine was also observed when Chen *et al.* carried out an enzymatic reduction on 8-nitroguanine using lipoyl dehydrogenase with NADPH as a cofactor.⁹⁹ Nevertheless, the work carried out in this thesis sought to investigate the feasibility of converting 8-nitroguanosine directly back to guanosine as this could suggest a potential repair mechanism for the lesion.

The previous enzymatic work carried out by Chen made mention of only 8-aminoguanosine forming in the reduction by lipoyl dehydrogenase.⁹⁹ However, the discovery that a deazaflavin-dependent nitroreductase (Ddn) enzyme can carry out a reductive denitration on a nitroimidazole compound prompted the thought that a related enzyme capable of binding to DNA could carry out a similar reaction on the 8-nitroguanine lesion.¹⁰⁰ The reductive denitration of the nitroimidazole compound was reported by Singh *et al.* whilst investigating new treatments for tuberculosis.¹⁰⁰ The postulated mechanism of action is believed to proceed through hydride transfer from the enzyme to the nitroimidazole ring followed by protonation and finally elimination of nitrous acid, as shown in Figure 16.¹⁰⁰

The overall reaction is a denitration in which NO_2 is replaced by H and it was hypothesised that a similar reaction on the 8-nitroguanine lesion in DNA could convert it back to guanosine. If a hydride equivalent could be delivered to the C-8 position, then an addition-elimination reaction could take place that would expel the nitro group.

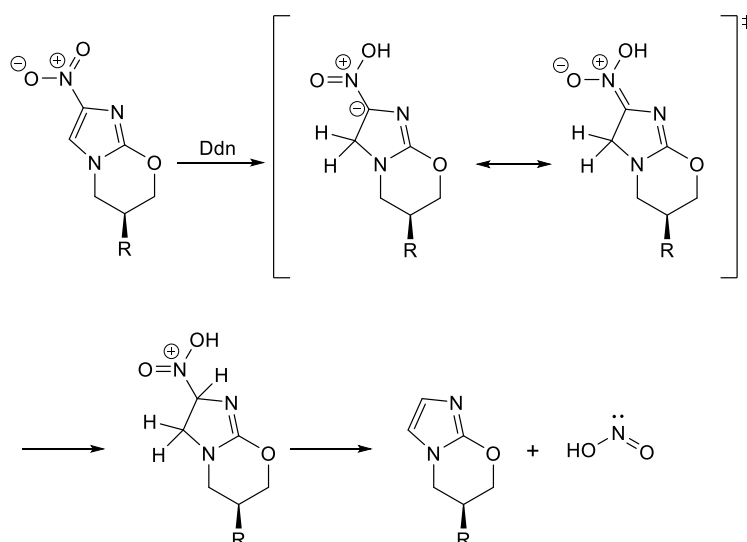
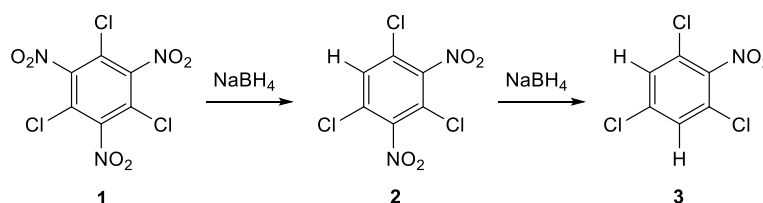


Figure 16 Suggested mechanism for the reductive denitration of a nitroimidazole by a deazaflavin-dependent nitroreductase (Ddn). The R group of the nitroimidazole refers to a number of *p*-(trifluoromethoxy)benzene derivatives.¹⁰⁰

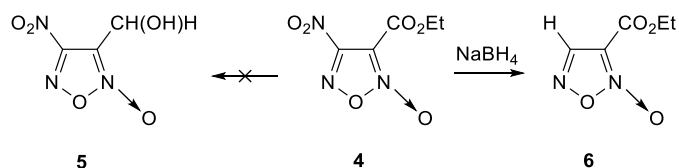
In terms of carrying out a reductive denitration, there is literature precedent to show sodium borohydride carrying out the reaction on both heterocyclic systems and polysubstituted nitro aromatic compounds. Work by Kaplan found that when 1,3,5-trichloro-2,4,6-trinitrobenzene **1** was exposed to sodium borohydride under alkaline conditions, a reduction took place that produced 1,3,5-trichloro-2,4-dinitrobenzene **2** as the only isolable product rather than any reduction products.¹⁰¹ Further reduction of **2** with sodium borohydride resulted in the substitution of another nitro group with hydride as shown in Scheme 2.¹⁰¹ Further work in this area has determined that the products formed from reaction of poly-substituted nitro aromatic compounds with NaBH_4 is strongly influenced by the precise substrate and reaction conditions, but these results indicate the reaction is possible.¹⁰²



Scheme 2 Reductive denitration of a polysubstituted nitro aromatic compound using NaBH_4 .¹⁰¹

A study by Blinnikov and Makhova found that a nitro substituent in a furoxan derivative **4** was replaced by hydride using NaBH_4 .¹⁰³ The initial aim of their work was to produce a nitro alcohol through reduction of a nitro ester, but they

unexpectedly found that a reductive denitration occurred instead, as shown in Scheme 3.¹⁰³



Scheme 3 Reductive denitration of heterocyclic system.¹⁰³

Again, the results of this work show displacement of a nitro group with hydride is possible and so this project aimed to investigate whether a similar reaction was feasible using a nitro nucleoside as a substrate.

2.5 Project aims

The DNA lesion 8-nitroguanine is known to form at sites of chronic inflammation and is strongly associated with mutagenic and carcinogenic events that occur at these sites. It is currently unknown whether the lesion is repaired by its own specific DNA repair pathway, but given its propensity to depurinate, it is likely it is more generally recognised and repaired *via* an excision repair mechanism. The purpose of this work was to investigate the possibility of carrying out a direct repair on the lesion. More specifically, this project sought to investigate the chemical feasibility of using a hydride source to carry out a reductive denitration reaction on 8-nitroguanosine that would convert it directly back to guanosine and therefore suggest that a direct repair of the lesion was possible by an enzyme able to deliver a hydride equivalent.

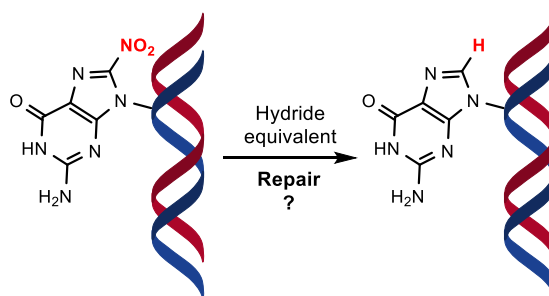
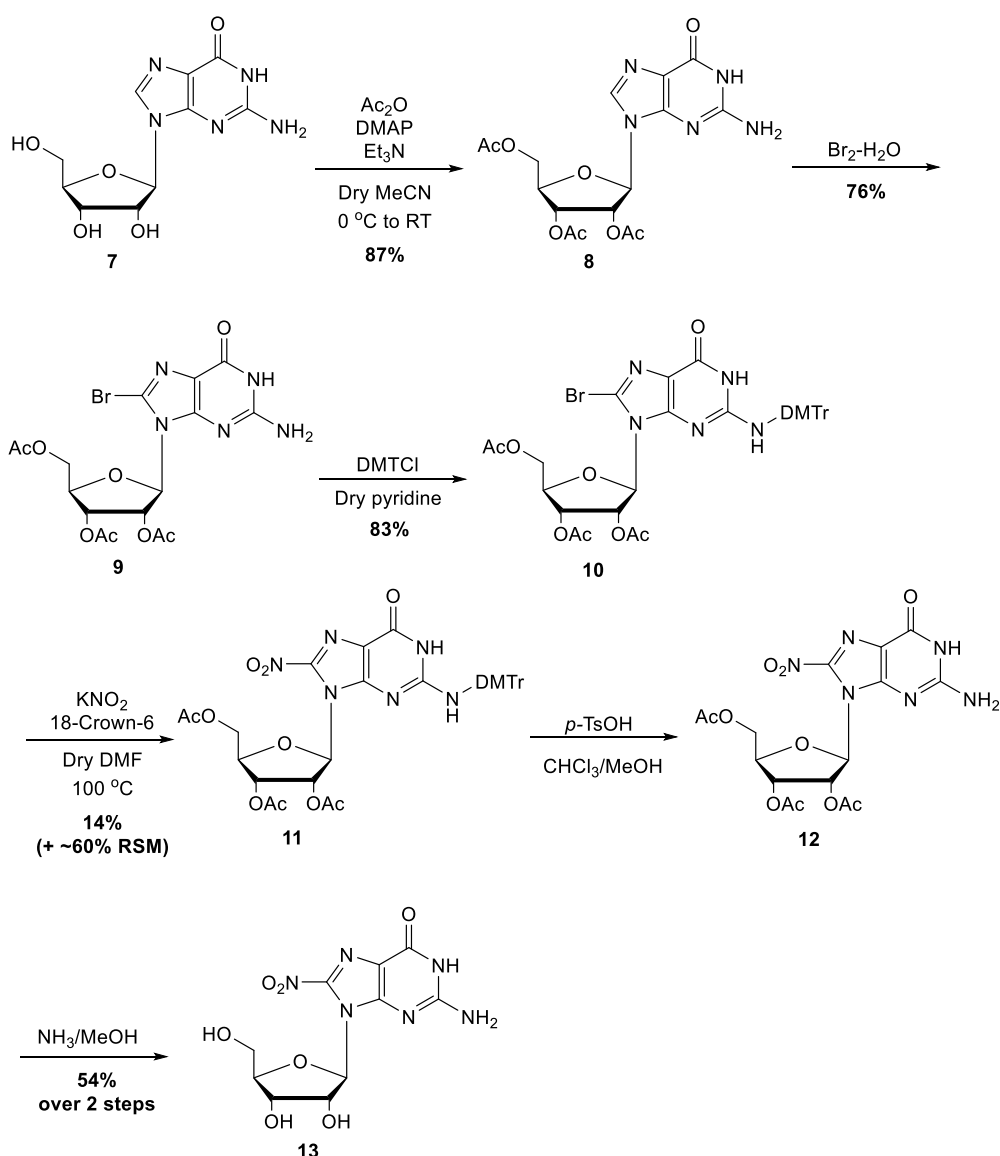


Figure 17 Hypothesised direct repair of 8-nitroguanosine by a hydride source or equivalent.

Chapter 2 Results and discussion 1

2.6 Synthesis of 8-nitroguanosine

Before investigations into the feasibility of carrying out a reductive denitration on 8-nitroguanosine could begin, the compound itself first had to be synthesised. A synthetic route to 8-nitroguanosine was published by Saito *et al.* in 2008 and has been more widely used.^{84,104} This route was therefore adopted for use in this project. Scheme 4 outlines the route and the following section will discuss each step in more detail.



Scheme 4 Synthetic route to 8-nitroguanosine.

As can be seen from Scheme 4, it was decided to use the ribonucleoside form of guanosine as the starting material as the presence of the extra hydroxyl group at

the 2' position helps add stability, making depurination following nitration less likely to occur. This hydroxyl group, along with the hydroxyls at C-3' and C-5' require protection to stop them from participating in any unwanted side reactions. The first step of the synthesis therefore was protection of the three hydroxyl groups of the sugar ring using the acetyl protecting group. The acetyl protecting group was chosen as it is removed under basic conditions and it is electron withdrawing. These properties help to reduce depurination as they minimise exposure to acid later in the synthesis and the electron withdrawing nature of acetyl groups makes formation of an oxonium ion less favourable.

Reaction of guanosine with acetic anhydride in the presence of triethylamine with DMAP as a catalyst produced the desired tri-*O*-acetyl nucleoside in 87% yield. The DMAP increases the rate of reaction by acting as a nucleophilic catalyst and the triethylamine helps to neutralise the acetic acid that forms as a by-product. The formation of the acetylated product **8** was confirmed by ¹H NMR which showed three singlets at 2.03, 2.04 and 2.11 ppm corresponding to the incorporation of the three CH₃ groups from the acetyls.

The next step of the synthesis was an electrophilic bromination reaction at the C-8 position. Insertion of bromine at this position is required so that it can be displaced by nitrite in the key step of the synthesis. The bromination was carried out by treating tri-*O*-acetylguanosine **8** with aliquots of bromine water which produced the desired product **9** in 76% yield. This reaction is somewhat unusual in that both the starting material and product are insoluble in the reaction media. Nevertheless, formation of the product was confirmed by mass spectrometry which showed the characteristic isotope pattern for incorporation of a bromine atom. In addition, the ¹H NMR lacked the peak for *H*-8 that was present in the starting material at 7.93 ppm which indicated successful incorporation of bromine at the correct position.

Protection of the exocyclic amine of the nucleobase was then required as, like the hydroxyl groups of the sugar, it has the potential to take part in unwanted side reactions. The dimethoxytrityl (DMTr) protecting group was chosen for this purpose because, as mentioned previously, compound **9** is fairly insoluble and so introduction of the DMTr group helps to increase the solubility of the nucleoside. Reaction of compound **9** with DMTrCl in anhydrous pyridine afforded the fully protected nucleoside **10** in 83% yield. Confirmation the desired product had been successfully isolated was provided by ¹H NMR which indicated the presence of the DMTr aromatic protons at 6.83-7.30 ppm.

For the critical nitration step, compound **10** was reacted with KNO_2 in DMF at 100°C in the presence of 18-crown-6. The high temperature for the reaction was required in order to solubilise the KNO_2 . The 18-crown-6 makes the NO_2^- anion more nucleophilic as it acts as a chelating agent for the potassium cation, for which it has a particular affinity. However, the yield of nitrated product **11** obtained was disappointingly low (14%) and not in agreement with the 60% yield reported by Saito *et al.* using the same conditions.¹⁰⁴ Similarly low yields have also been found for this reaction previously within our research group and a great amount of effort has gone into trying to improve the reaction conditions. Unfortunately, the exact mechanism by which the nitration takes place under laboratory conditions is unknown which makes it difficult to know how best to optimise the reaction. It has been postulated that the nitration with KNO_2 could proceed *via* an addition-elimination type mechanism, as illustrated in Figure 18, but so far this could not be proven.

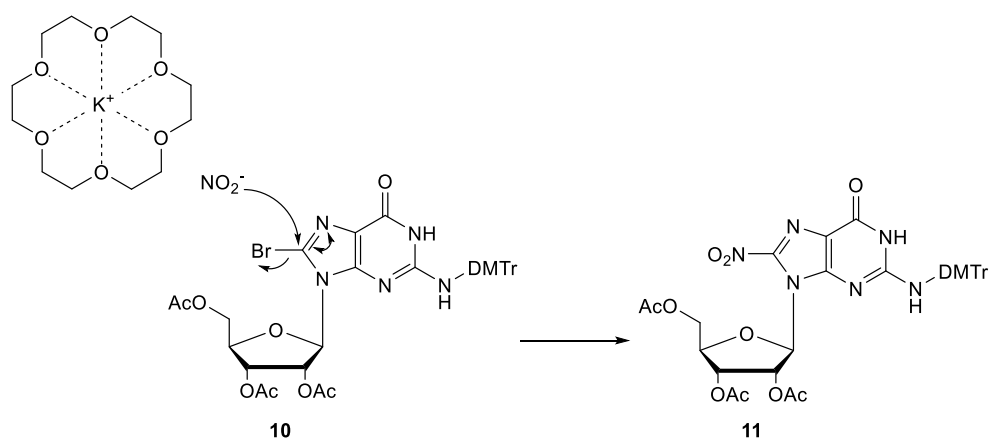


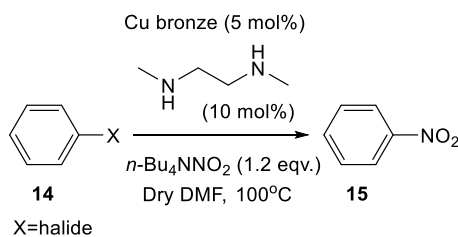
Figure 18 Formation of 8-nitroguanosine via an addition-elimination mechanism.

As discussed in the introduction, it is thought that *in vivo*, the reaction between guanosine and peroxyxynitrite and its decomposition products proceeds *via* a radical pathway. Previous work within the group investigated whether the reaction with KNO_2 could potentially proceed in a similar fashion by adding the radical initiator azobisisobutyronitrile (AIBN). However, it was found that this had no overall effect on the reaction rate or yield implying the reaction was unlikely to involve radical species.

Other previous attempts to increase the yield of nitration include; varying the reaction concentration, varying the reaction solvent, altering the amounts of reagents used, changing the halogen at C-8 and use of silver (I) salts to sequester

the displaced halogen. All had little or no effect on increasing the yield to those stated in the literature.

Due to the consistently low yields of ~15% obtained in this project, several further attempts were made to improve the yield. These included; investigating the need for anhydrous conditions, using a different source of nitrite and the use of a copper catalyst. Addition of a known amount of water to the reaction resulted in no product formation which indicates anhydrous conditions are required. Upon changing the source of nitrite from KNO_2 to $n\text{-Bu}_4\text{NNO}_2$, again, no product formation was observed, although $n\text{-Bu}_4\text{NNO}_2$ was found to be much more soluble in DMF than KNO_2 . A paper by Koizumi *et al.* reported that aromatic halides could be nitrated in good yield using a copper catalyst, amine ligand and $n\text{-Bu}_4\text{NNO}_2$, as shown in Scheme 5.¹⁰⁵ However, when these conditions were tried using compound **10** as a starting material no nitro product was formed.



Scheme 5 Nitration of an aromatic halide.¹⁰⁵

As the main aim of the project was to investigate the reductive denitration of 8-nitroguanosine, it was decided at this point to continue with the original conditions reported by Saito as no progress was being made in optimising the reaction. Saito's reaction conditions allow for recovery of a significant proportion of unreacted starting material (~60%) which can be recycled in further reactions. So, despite the disappointing yields, enough nitrated product **11** was able to be synthesised and taken forward to the next step.

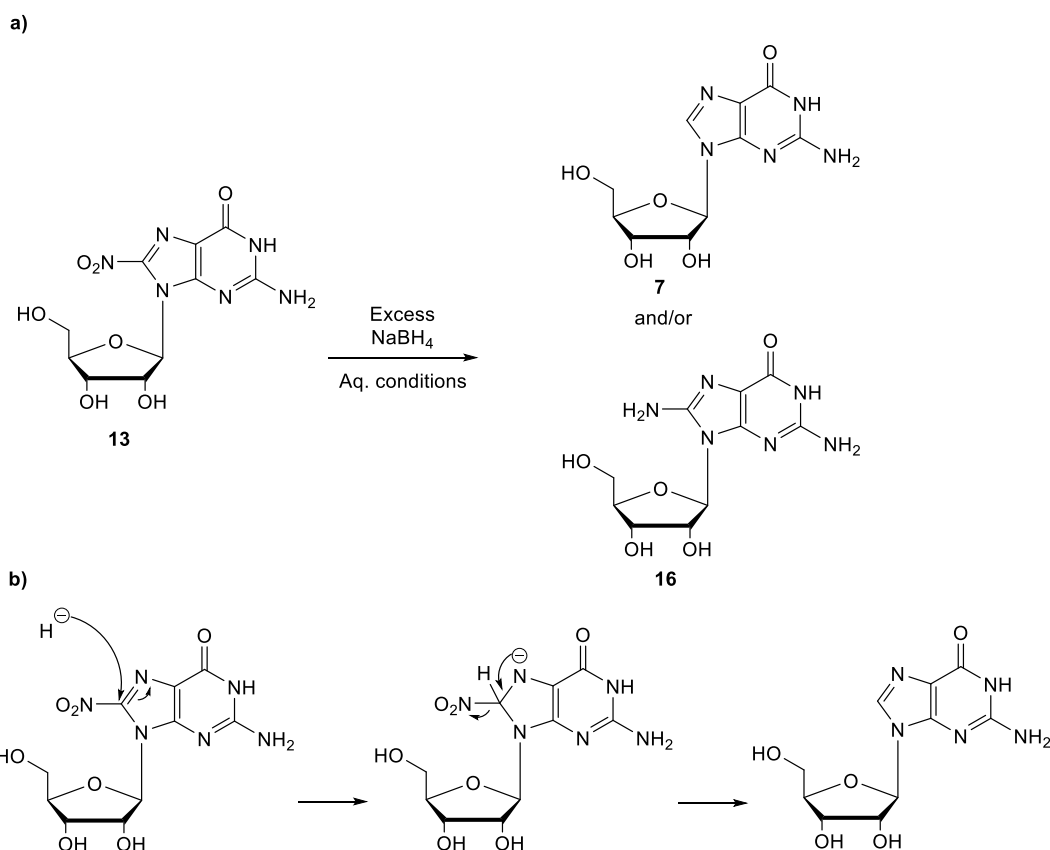
Formation of the nitrated product **11** was observed through monitoring of the reaction by reverse phase high performance liquid chromatography (RP-HPLC). Once nitrated, the nucleoside possesses a distinctive UV absorption at ~400 nm that is not present in the starting material and so provides an effective method for identifying product formation. Following purification by column chromatography, successful isolation of the desired product **11** was confirmed by mass spectrometry. An accurate mass, showing only one isotope, was obtained that was consistent with a mass corresponding to that of the desired nitro compound.

The final stage of the synthesis was removal of the protecting groups. As the two different protecting groups employed are removed under different conditions, an orthogonal deprotection strategy was required. The DMTr group requires acidic conditions to be removed which can promote depurination in the nitrated nucleoside. It was therefore decided to deprotect the exocyclic amine first whilst the electron withdrawing acetyl groups were still in place to offer some stabilisation to the glycosidic bond. Removal of the DMTr group was achieved by adding *p*TsOH to a solution of compound **11** in CHCl₃/MeOH. Following removal of the solvent and trituration, the DMTr deprotected nucleoside was carried through to the next step without further purification. The acetyl groups were then removed using methanolic ammonia. These conditions are typical for removal of acetyl protecting groups and are ideal for use with nucleosides as they do not cause any significant decomposition. Upon completion of the deprotection, the nucleoside was trituated to remove any acetamide that formed as a by-product. Over the course of the two deprotection steps, 8-nitroguanosine **13** was obtained in a 54% yield. Spectroscopic evidence for its formation was provided by ¹H NMR which lacked the aromatic signals of the DMTr group and the three CH₃ signals of the acetyl groups showing they had been successfully removed.

2.7 Reductive denitration of 8-nitroguanosine

Once 8-nitroguanosine had been synthesised, investigations into reductively denitrating it began. It was decided to test out the reaction using NaBH₄ as the hydride source. As discussed in the introduction, NaBH₄ has previously been utilised to successfully carry out reductive denitration reactions.^{101,103} Although it is not a compound that would be found *in vivo*, it was thought it would be a good starting point for assessing whether reduction of 8-nitroguanosine to guanosine by a hydride equivalent is chemically feasible. There are several enzymes known to be capable of delivering a hydride equivalent, an example being the deazaflavin-dependent nitroreductase discussed earlier.¹⁰⁰ It was therefore thought that if it could be shown that 8-nitroguanosine could be reductively denitrated with NaBH₄, then a similar process could be envisioned *in vivo* using an enzyme capable of delivering a hydride equivalent.

Scheme 6 summarises the hypothesised reductive denitration reaction and the potential products it could form.



Scheme 6 a) Hypothesised reductive denitration reaction of 8-nitroguanosine b) Mechanism showing how attack of hydride at C-8 could result in loss of nitro group.

The reduction reactions were carried out on a relatively small scale, due in part to the difficulties associated with bringing through large quantities of 8-nitroguanosine. Initially, it was decided to carry out the reductive denitration reactions under aqueous conditions using 0.1 M triethylammonium bicarbonate (TEAB) solution. The reasoning behind this decision was that the aqueous conditions would make the study more biologically relevant and the slightly basic pH (~pH 8) of the TEAB solution would help to reduce any depurination occurring, although it is not possible to stop this happening altogether. A large excess of NaBH_4 (50 eqv.) was added at regular intervals over the course of the reactions (every 30 minutes for 90 minutes) to compensate for hydrolysis under the aqueous conditions. RP-HPLC was used to monitor the reactions using a gradient of MeCN in TEAB (0.1 M) which had previously been shown to be an effective eluent for separating mixtures of nucleosides. Figure 19 shows two RP-HPLC traces recorded under these conditions.

The results obtained were promising as they showed 8-nitroguanosine (retention time=11.56 min) being consumed in the reaction with the appearance of a new peak in the chromatogram for the reduced product (retention time=8.80 min). 8-

Nitroguanosine was readily identified in the chromatograms due its distinctive absorbance at ~400 nm. As can be seen from the traces, despite best efforts to minimise depurination, some trace impurities were present. The new peak forming over the course of the reaction did not absorb at 400 nm implying it was a denitrated product. However, it was not clear from the RP-HPLC data whether the product forming was guanosine *via* a reductive denitration or whether it was 8-aminoguanosine *via* simple reduction. A series of coinjections of the reaction mixture with standard samples of guanosine and 8-aminoguanosine proved inconclusive, as under these eluent conditions, guanosine and 8-aminoguanosine were found to both elute at very similar retention times. Efforts to isolate the product formed in the reactions so an NMR spectrum could be recorded were unsuccessful as they had not been carried out on a large enough scale to obtain sufficient material for a good NMR sample.

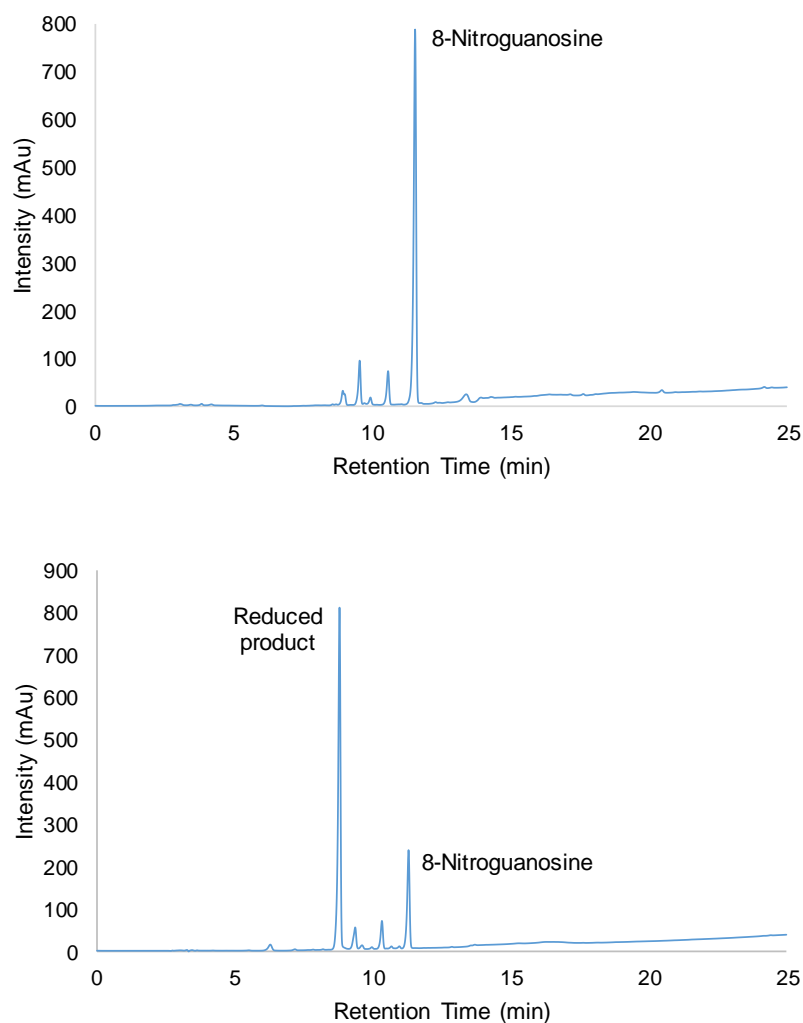


Figure 19 Reduction of 8-nitroguanosine in 0.1 M TEAB with NaBH_4 . Top: HPLC trace at time 0. Bottom: HPLC trace at time 90 minutes. Both recorded on a reverse-phase C18 column with an elution gradient of MeCN in TEAB (0.1 M). Chromatograms recorded at 254 nm using HPLC Method 1 (p 116).

Therefore, a change in the RP-HPLC eluents was required in order to find a suitable system that would give a good separation between the two potential products of the reaction. At this stage of the project, an opportunity to use a HPLC-MS system through a collaboration with The MicroBioRefinery at The University of Liverpool arose. So, in addition to wanting to find a better resolving eluent system for separation of the reductive denitration reaction mixture, a move towards conditions that would be compatible with HPLC-MS was also required.

A gradient of MeCN in 0.1% AcOH in water was found to give excellent separation between the two potential products of the reactions, guanosine and 8-aminoguanosine, and so was used to monitor all subsequent reactions. These eluent conditions were also ideal for use with HPLC-MS systems as the AcOH and MeCN are volatile enough not to cause issues in the spectra obtained from the MS system. It was also decided at this point to carry out the reduction reactions in distilled water rather than 0.1 M TEAB, as the presence of triethylammonium cations from TEAB tend to dominate the mass spectra obtained from the HPLC-MS. Figure 20 shows a series of 3 chromatograms recorded using the optimised conditions.

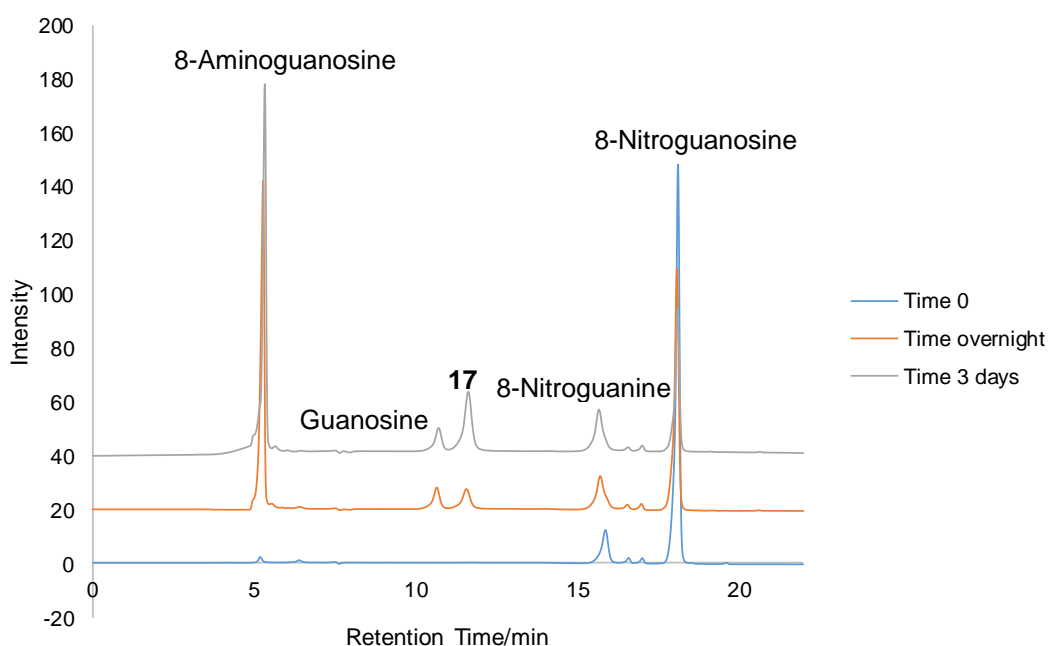


Figure 20 Reduction of 8-nitroguanosine in water with NaBH_4 . All three traces were recorded using a reverse-phase C-18 column and an elution gradient of 0.1% AcOH in water/MeCN. Chromatograms recorded at 254 nm using HPLC Method 2 (p 117).

The chromatograms show that addition of NaBH_4 to 8-nitroguanosine in water causes a slow reduction to occur that produces 8-aminoguanosine as the major product in addition to a small amount (8-10%) of guanosine. The identity of both

guanosine and 8-aminoguanosine was confirmed with the aid of HPLC-MS. Table 3 shows the masses obtained for all the peaks.

Table 3 Reduction products identified using HPLC-MS with both their calculated and found masses.

Compound	Calculated mass for [M-H] ⁻	Mass found for [M-H] ⁻
8-Nitroguanosine 13	327.0695	327.0685
8-Aminoguanosine 16	297.0953	297.0943
8-Nitroguanine	195.0272	195.0265
Guanosine 7	282.0844	282.0827
Compound 17	-	280.0672

Interestingly, the chromatograms show that the small amount of nitrated base, 8-nitroguanine, present in the starting material due to depurination, persists throughout the reaction and shows no sign of reducing to either 8-aminoguanine or guanine.

As can be seen in the chromatograms in Figure 20 and the mass data in Table 3, in addition to the two reduction products, a third product labelled **17** was found to form (~20%). Using the accurate mass data obtained from its peak in the chromatogram it is believed compound **17** is 8,5'-O-cycloguanosine (Figure 21).

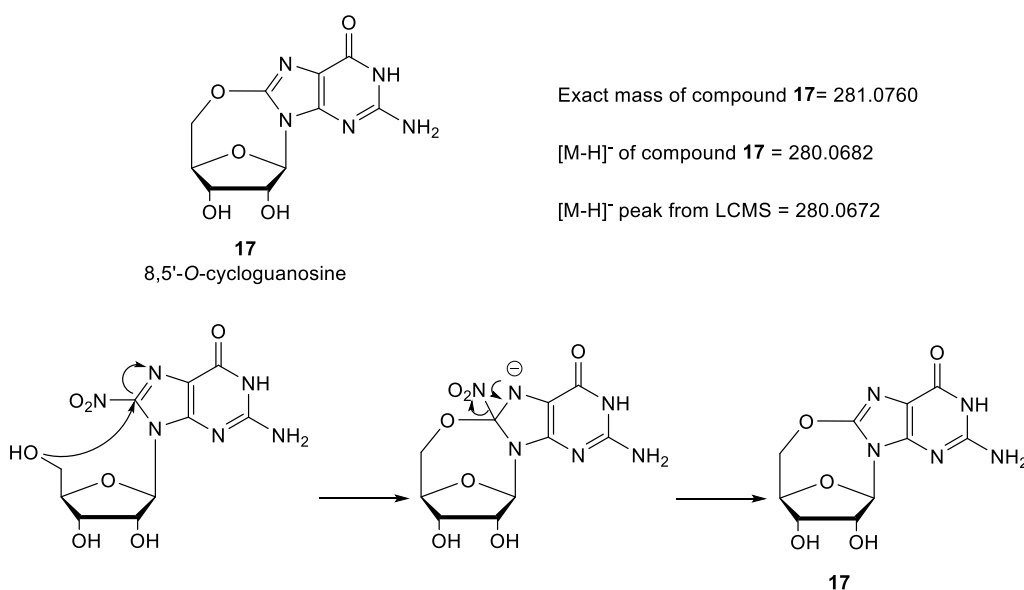
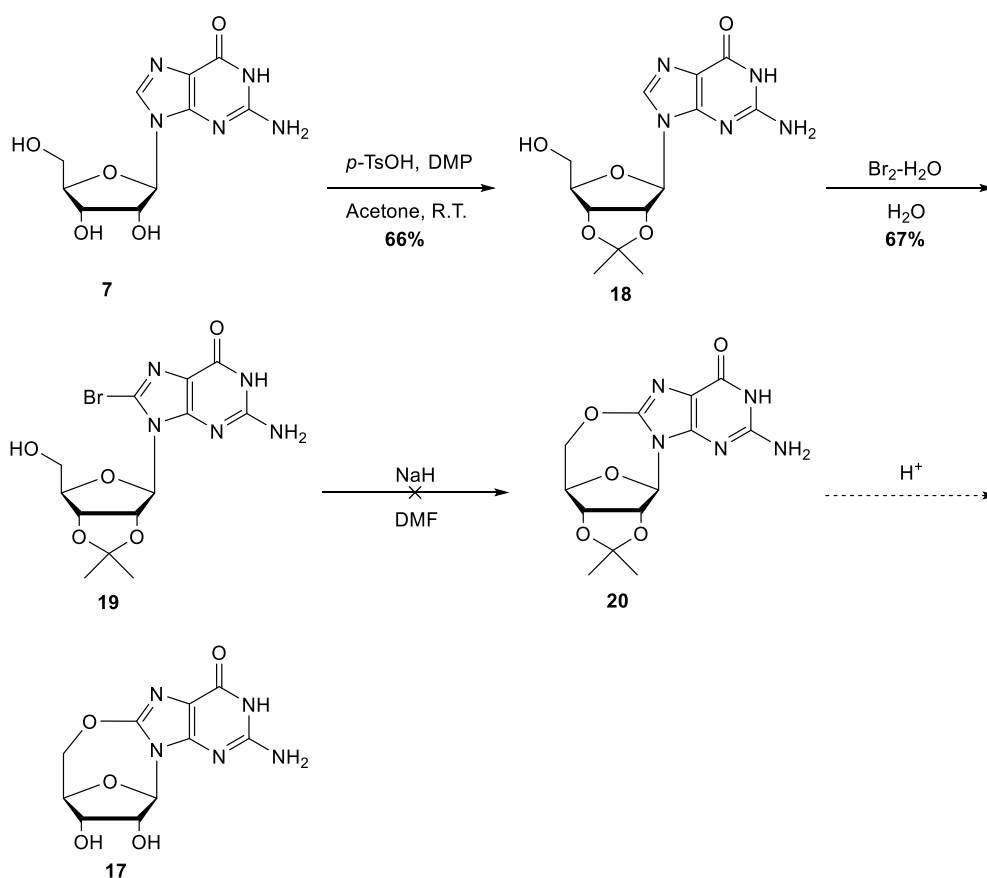


Figure 21 Structure of 8,5'-O-cycloguanosine with its measured and calculated masses for the [M-H]⁻ ion (top). Proposed mechanism of formation of 8,5'-O-cycloguanosine (bottom).

Over the course of the reaction, the pH rises as the NaBH_4 is hydrolysed. Given that the nitro functionality is known to be a very good leaving group, particularly with thiol nucleophiles, it is believed that the 5'-OH carries out a displacement of the nitro group facilitated by the basic pH leading to compound **17**, as shown in Figure 21.^{106,107}

Efforts were made to synthesise a standard sample of 8,5'-O-cycloguanosine so that a coinjection with the reaction mixture could be run to confirm that it is compound **17**. Scheme 7 shows the planned synthetic route which involves using cyclisation conditions originally reported by Ikehara.¹⁰⁸



Scheme 7 Planned synthetic route to 8,5'-O-cycloguanosine.

However, several attempts to perform the cyclisation using NaH in DMF failed. Despite the lack of a standard compound, the literature precedent for this type of cyclisation combined with the accurate mass for compound **17** was considered solid evidence that compound **17** was 8,5'-O-cycloguanosine. In addition, it is not a compound that would readily form *in vivo*, as the 5'-OH that is thought to displace the nitro moiety forming compound **17** would not be free to react as it is bonded to the next nucleotide forming the sugar phosphate backbone.

2.8 Control reactions

In order to confirm the products formed during the reductive denitration reactions were due to the reducing agent being added, a series of control reactions were carried out. Table 4 summarises the control reactions performed and their outcomes.

Table 4 Summary of control reactions carried out, showing the outcome of the reaction and the reason why it was performed.

Control conditions	Reason	Outcome
8-NO ₂ -rG dissolved in distilled water, sodium dithionite added.	To produce a standard sample of 8-NH ₂ -rG using a known reduction method.	HPLC showed only one peak for 8-NH ₂ -rG following reduction.
8-NO ₂ -rG dissolved in 0.1 M TEAB, no NaBH ₄ added.	To rule out reduction products forming due to some sort of hydrolysis.	After 24 hours, HPLC showed no formation of 8-NH ₂ -rG, guanosine or compound 17 .
8-NO ₂ -rG dissolved in distilled water, no NaBH ₄ added.	To rule out reduction products forming due to some sort of hydrolysis.	After 24 hours, HPLC showed no formation of 8-NH ₂ -rG, guanosine or compound 17 . A slight increase in the amount of 8-nitroguanine was observed.
8-NH ₂ -rG dissolved in 0.1 M TEAB, NaBH ₄ (50 eqv.) added every 30 mins for 90 mins.	To determine whether the guanosine forming was coming from 8-NO ₂ -rG or breakdown of 8-NH ₂ -rG.	After 24 hours, HPLC showed presence of only 8-NH ₂ -rG.
8-NH ₂ -rG dissolved in distilled water, NaBH ₄ (50 eqv.) added every 30 mins for 90 mins.	To determine whether the guanosine forming was coming from 8-NO ₂ -rG or breakdown of 8-NH ₂ -rG.	After 24 hours, HPLC showed presence of only 8-NH ₂ -rG.
8-NH ₂ -rG dissolved in distilled water, no NaBH ₄ added.	To determine whether the guanosine forming was coming from 8-NO ₂ -G or breakdown of 8-NH ₂ -rG.	After 24 hours, HPLC showed presence of only 8-NH ₂ -rG.

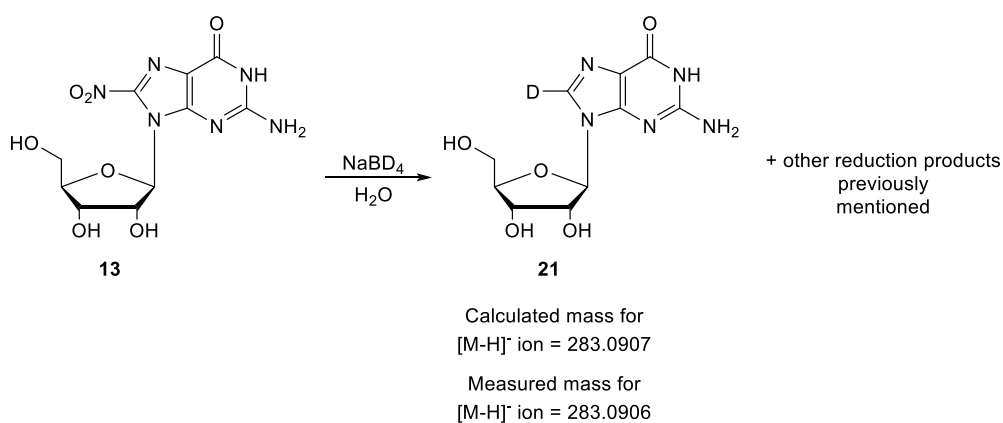
The results of the control reactions indicate that the observed products formed during the reductive denitration reactions because of the NaBH₄ added. To further confirm the results a deuterium study was carried out.

2.9 Deuterium study

A deuterium study was carried out in order to prove that the guanosine forming over the course of the reductive denitration reactions was being produced as a result of the hydride delivering reducing agent being added.

Analogous reactions to those performed previously were carried out in which 8-nitroguanosine was dissolved in distilled water and then 50 equivalents of reducing agent was added every 30 minutes for 90 minutes. The only difference in reaction conditions to those performed previously was the substitution of NaBH₄ for NaBD₄. HPLC was again used to monitor the reactions and pleasingly identical chromatograms were observed to those obtained for reaction with NaBH₄. 8-Aminoguanosine, guanosine and compound **17** were all found to form following reaction of 8-nitroguanosine with NaBD₄.

Analysis by HPLC-MS showed the guanosine peak to have the correct mass for incorporation of one deuterium atom (Scheme 8). This proved that the guanosine forming over the course of the reaction was due to the action of the reducing agent being added.



Scheme 8 Reduction to form 8-deuteroaminoguanosine. Deuterium incorporation identified using HPLC-MS. Both its calculated and measured masses for the [M-H]⁻ ion are shown.

As expected, the mass obtained for the peak corresponding to 8-aminoguanosine showed no inclusion of deuterium as any incorporated during the reduction reaction would be exchanged through protonation by the solvent water. This exchange cannot take place in the case of guanosine as the proton/deuteron on C-8 is not labile.

2.10 Change of reaction conditions

In an attempt to increase the proportion of guanosine formed, the reaction was investigated under a variety of conditions. Aqueous buffered solutions between pH 6-9 were prepared (Table 5) and the reduction reactions were carried out as before to see if any increase in the proportion of guanosine could be observed.

Table 5 Table to show pH of aqueous buffers used in the reductive denitration reactions.

pH required	Aqueous solution used
6	0.1 M Ammonium acetate adjusted to pH 6
7	Distilled water
8	0.1 M TEAB
9	0.1 M Ammonium acetate adjusted to pH 9

All of the reactions at the four different pH's were monitored by HPLC, with the pH 7 reaction acting as a control as the reaction media (distilled water) was the same as used previously. However, very little difference in product distribution at pH's 6, 8 and 9 were observed compared to that obtained in water. The only real difference observed between the reactions was that at pH 6 the reduction appeared to progress slightly faster. The fact that all the aqueous reduction reactions started at different pH's gave essentially the same product distribution can be rationalised by considering their finishing pH. The large excess of borohydride added regularly over the course of the reaction raises the pH as it is hydrolysed meaning that all of the reactions ended up pH ~9.6 by the end of reaction.

The slightly milder reducing agent, NaBH_3CN was tried in the reaction in place of NaBH_4 to see whether this had an effect on the product distribution obtained. Initial reaction of 8-nitroguanosine with NaBH_3CN was carried out in an analogous way to the NaBH_4 reactions, by adding 50 equivalents every half an hour for 90 minutes to a solution of the nucleoside in distilled water. Analysis by HPLC showed that under these conditions no reduction took place, even after prolonged stirring, with the only change in the chromatograms being an increase in the amount of 8-nitroguanine base observed. The reaction with NaBH_3CN was then repeated at the slightly acidic pH 6, obtained by using 0.1 M ammonium acetate solution adjusted to pH 6. NaBH_3CN is known to be more activated at lower pH's and so it was thought that by using pH 6, some reduction may be observed. However, HPLC analysis of the reaction again showed only starting material indicating that NaBH_3CN is too weak a reducing agent to carry out the reductive denitration. Due to the potential for depurination to occur more readily under acidic conditions, no pH lower than 6 was investigated.

Work carried out by Lamson *et al.* on the denitration of polysubstituted nitro aromatic compounds suggests that the reaction occurs quicker in DMSO than in protic solvents.¹⁰² However, when DMSO was used in favour of distilled water for the reaction in this project, denitration was found to occur either extremely slowly or

not at all. An initial reaction found that after addition of NaBH_4 and a further week of stirring no 8-aminoguanosine or guanosine had formed at all. The only product observed was a trace amount of the cyclisation product **17**. A second attempt at the reaction in DMSO was heated to $50\text{ }^\circ\text{C}$ and left to stir for even longer (10 days). HPLC analysis of this reaction showed a small a peak for 8-aminoguanosine and a larger peak for compound **17**. It is believed that as the 8-nitroguanosine starting material is not being reduced to either 8-aminoguanosine or guanosine in DMSO as quickly as it does in aqueous conditions, then the competing reaction of displacement by the 5'-OH group can occur more, producing compound **17**.

2.11 Conclusions

The results of this study show that whilst the major product that forms in the reductive denitration of 8-nitroguanosine is 8-aminoguanosine, a small proportion of guanosine does also form. Although efforts to increase the proportion of guanosine formed have so far been unsuccessful, its formation by displacement of the nitro group by a hydride equivalent, in aqueous solution, does clearly show that a direct repair of 8-nitroguanosine is a chemically feasible reaction. Control reactions and a deuterium study have proven that the origin of the guanosine formed is from the 8-nitroguanosine starting material used in the reactions and not from breakdown of 8-aminoguanosine or an alternative non-reductive pathway.

2.12 Future work

Now that it has been shown that NaBH_4 is capable of forming guanosine from 8-nitroguanosine, albeit it in a small amount, it would be interesting to see what other hydride delivering agents can perform the transformation. A move towards investigation of more biologically relevant hydride sources would be favourable. Of the highest priority, would be investigations into whether an enzyme similar to Ddn could also convert 8-nitroguanosine to guanosine. Ddn is dependent on a flavin derivative to function and so studies into the synthesis of flavin related molecules would also be an interesting line of enquiry. Other sources of physiological hydride, such as nicotinamide adenine dinucleotide hydride (NADH), are also of interest to this project. NADH is a coenzyme and naturally occurring source of hydride found in all living cells. Along with its phosphate ester, NADPH, they take part in a variety of cellular redox reactions and processes. Several synthetic analogues of NADH are known, with most commonly derived from the nicotinamide moiety as this is what acts as the hydride donor, an example of which is shown in Figure 22. Future work

would look to see if any NADH or flavin analogues can carry out the reductive denitration process.

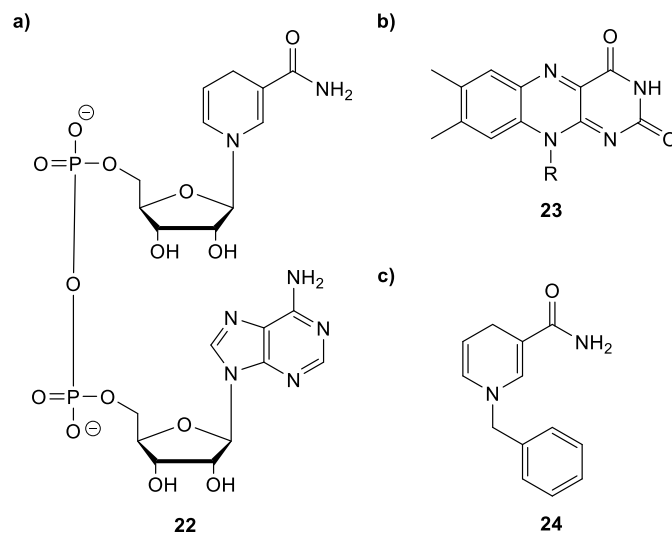


Figure 22 a) Structure of NADH b) General structure of flavin from which analogues can be derived by altering the R group c) Example of an NADH analogue based on the structure of nicotinamide.

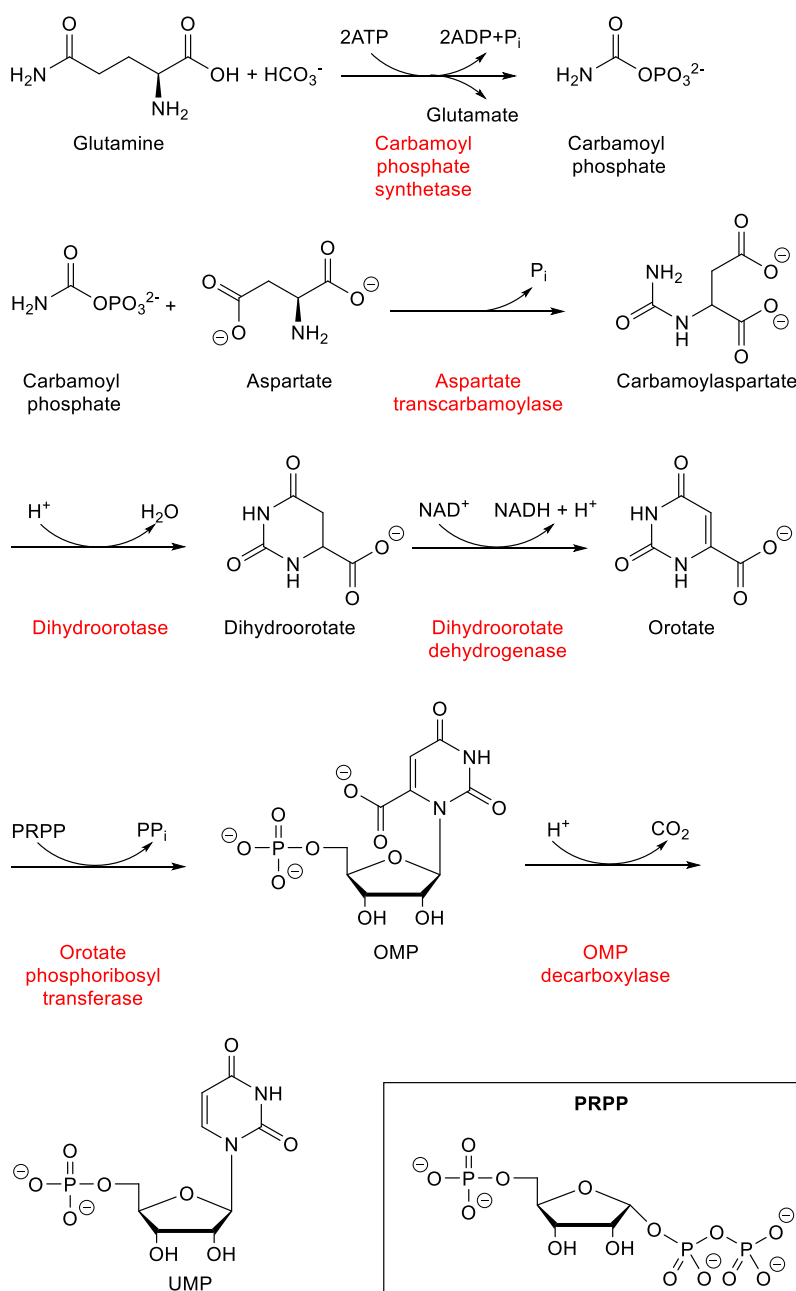
Chapter 3

Modifications of a pyrimidine nucleoside to probe the process of decarboxylation

Chapter 3 Introduction

3.1 Pyrimidine nucleotide biosynthesis

There are two routes through which pyrimidine nucleotides can be synthesised, namely the *de novo* pathway and the salvage pathway. The salvage pathway makes use of intermediates and by-products from the degradation of nucleic acids and recycles them to reform the desired nucleotide.² In the *de novo* pathway, the nucleotide is synthesised from its constituent starting materials in a series of enzyme catalysed reactions, as shown in Scheme 9.^{2,109}



Scheme 9 The *de novo* pyrimidine biosynthetic pathway. Key enzymes are shown in red.

In the first step of the synthesis, glutamine reacts with bicarbonate to produce carbamoyl phosphate in a reaction that requires two molecules of ATP and is catalysed by carbamoyl phosphate synthetase.¹⁰⁹ Aspartate transcarbamoylase then catalyses the irreversible formation of carbamoylaspartate in a step that is subject to feedback inhibition by the pathway's final product, cytidine triphosphate (CTP).² Carbamoylaspartate cyclises with loss of water to give dihydroorotate which is dehydrogenated in the next step by dihydroorotate dehydrogenase to produce the pyrimidine ring orotate.^{2,109} 5-Phosphoribosyl-1-pyrophosphate (PRPP), a key intermediate in the pathway, then reacts with orotate to form orotidine monophosphate (OMP).^{2,109} In this step, there is an inversion of configuration at C-1' which results in the formation of the desired β -glycosidic linkage.² The final step is the decarboxylation of OMP by OMP decarboxylase (ODCase) to produce uridine monophosphate (UMP), the nucleotide from which all other pyrimidine nucleotides are derived.^{2,109}

Not only are pyrimidine nucleotides pivotal to life as the monomers that form DNA and RNA, they also play an essential role in cell metabolism as metabolic regulators and extracellular mediators.¹¹⁰ Most organisms, humans included, use a combination of both pathways in order to form the pyrimidine nucleotides they require. Certain organisms, however, such as the malaria parasite for example, lack the ability to salvage pyrimidine nucleotides and so rely completely on the *de novo* pathway to obtain the pyrimidines that they need.¹¹¹ This presents the potential to target the pyrimidine *de novo* pathway as a means to combat diseases such as malaria. Interestingly, although the malaria parasite is unable to salvage pyrimidine nucleotides, it relies solely on the salvage pathway to obtain purines as it lacks the ability to form purine rings *de novo*.¹¹²

If the pyrimidine *de novo* pathway could be disrupted, then certain disease-causing organisms would be starved of the pyrimidine nucleotides they need to survive, without causing too much damage to human cells which are still able to salvage them. Inhibition of one or more of the enzymes that catalyse the pyrimidine *de novo* pathway is one way this could be achieved. Disruption of one of the reactions in the pathway would cause an accumulation of intermediates prior to that reaction with the knock-on effect of causing a depletion in the subsequent intermediates and ultimately a reduction in the amount of nucleotides formed.¹¹³ Any imbalance capable of causing a shortfall in nucleotide synthesis can lead to mutations, miscoding and cell death.¹¹³

3.2 OMP decarboxylase (ODCase)

As shown in Scheme 9, ODCase catalyses the final step in the pyrimidine *de novo* pathway resulting in the formation of UMP from OMP. Enzymes earlier in the pathway have been extensively investigated, particularly in relation to potential cancer treatments although issues with toxicity have arisen.¹¹³ Only in relatively recent years has there been a renewed interest in ODCase as a therapeutic target, due largely to the discovery of the fact that of the six main enzymes that participate in the *de novo* pathway, ODCase has been discovered to be exceptionally proficient.¹¹⁴ It has been found to display a rate enhancement of over 17 orders of magnitude compared to the corresponding uncatalysed reaction at room temperature and neutral pH.^{114,115} To put this into context, it is estimated that the half-time ($t_{1/2}$) for the decarboxylation of OMP to occur in the absence of ODCase is approximately 78 million years compared to the 18 ms it takes when ODCase is present.¹¹⁶

ODCase is found in the majority of species, from bacteria to parasites to humans, with viruses being the only exception currently known.¹¹⁷ Viruses rely on their host cells to provide them with the nucleotides they need to replicate. In humans, and certain other high-level organisms, ODCase forms part of a bifunctional enzyme known as UMP synthase whereas it is a monofunctional enzyme in parasites and bacteria.¹¹⁸

3.2.1 ODCase catalytic mechanism

The mechanism by which ODCase decarboxylates OMP has been the subject of numerous studies, particularly since the enzyme has been found to carry out the process without the aid of any metal ions or co-factors making it unique amongst decarboxylases.^{119,120} Through the many studies into the mechanism of decarboxylation, much information has been gained about the active site of ODCase. The enzyme is known to exist as a dimer and two conserved lysine residues and two conserved aspartate residues have been identified as being particularly important to its overall catalytic ability.^{121,122} It has been found that upon binding, the C-6 position of OMP and these four residues are in close proximity to each other and although the specific decarboxylation mechanism remains unknown, several proposals have been put forward.^{123,124}

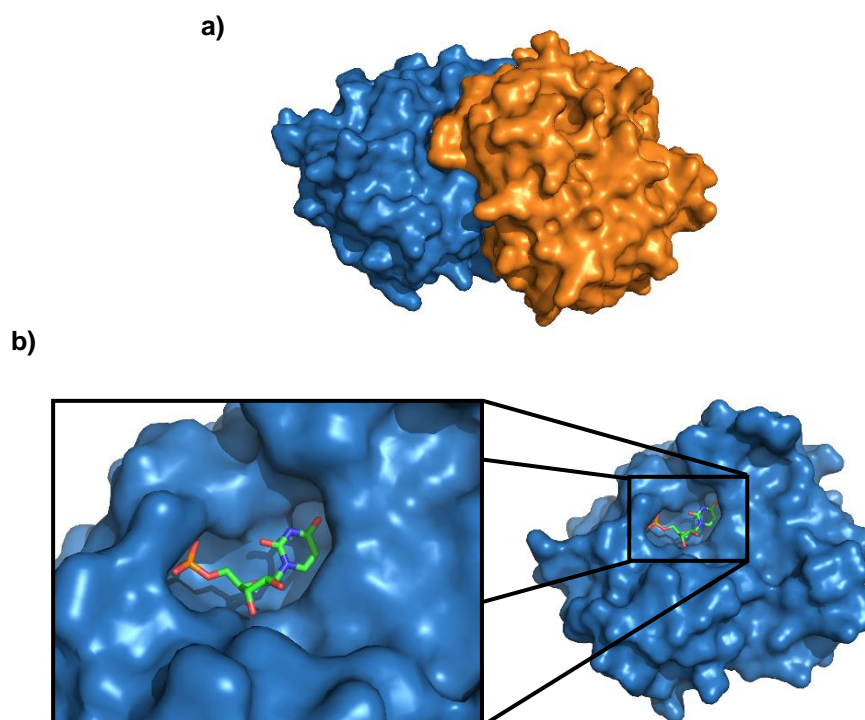


Figure 23 a) Surface structure of ODCase from *Methanothermobacter thermoautotrophicus* (one monomer coloured blue, the other coloured orange) b) Close up image of active site of one of the monomers in complex with UMP (coloured by atom type, carbon-green, nitrogen-blue, oxygen-red, phosphorus-orange). Images from a) and b) were rendered using PyMOL from data deposited in the protein data bank (PDB code: 4NUW).¹²⁵

Early work on the decarboxylation mechanism was carried out by Beak and Siegel.¹²⁶ Through studying the way 1,3-dimethylorotic acid is decarboxylated in the absence of any enzyme, they postulated that the formation of a zwitterionic intermediate leading to a nitrogen ylide is key to the reaction mechanism.¹²⁶ They proposed that protonation occurs at the carbonyl oxygen of C-2 resulting in the formation of an intermediate with a positively charged N-1.¹²⁶ They believed that this intermediate would offer stabilisation to the carbanion formed following decarboxylation therefore making the process more likely to occur.¹²⁶ Figure 24 shows their proposed mechanism.

This mechanism was later ruled out by Rishavy and Cleland in a paper that showed formation of a nitrogen ylide at this position was unlikely to occur.¹²⁷ They carried out a study to investigate the ¹⁵N kinetic isotope effect on N-1 and found that during the decarboxylation reaction there is no change in bond order that would imply formation of a quaternary N-1.¹²⁷

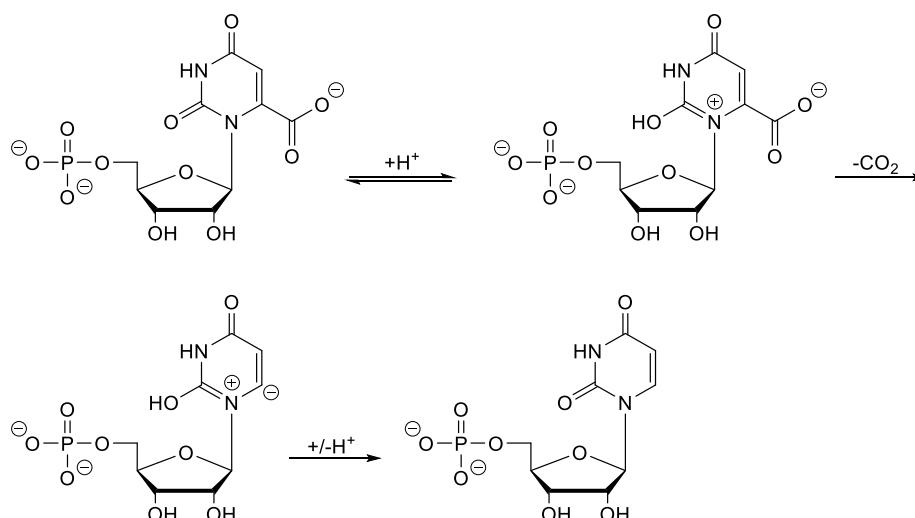


Figure 24 Proposed mechanism of decarboxylation put forward by Beak and Siegel.¹²⁶

An alternative mechanism involving a Michael addition to form a covalently bound intermediate which can then undergo a decarboxylative elimination was put forward by Silverman and Groziak.¹²⁸ An active site residue was suggested to perform a nucleophilic attack at C-5 creating a *trans* intermediate species that is then able to decarboxylate, eliminating carbon dioxide and the enzymic nucleophile as shown in Figure 25.¹²⁸ Work carried out by Acheson *et al.* disproved this theory as they were able to show there are no secondary deuterium isotope effects at C-5.¹²⁹ If the Michael addition mechanism was correct, a change in geometry from sp² to sp³ hybridization would be expected at C-5 upon addition of the enzyme active site residue. In order to monitor whether this change in geometry occurs, Acheson *et al.* replaced H-5 with a deuterium atom and monitored the effect on the rate constant for decarboxylation.¹²⁹ They found no observable secondary deuterium isotope effects suggesting that the geometry at C-5 does not alter.¹²⁹

As mentioned, to date, the exact catalytic mechanism is still not known for certain, but there is an increasing amount of evidence that the mechanism proceeds *via* a direct decarboxylation, as shown in Figure 26. Evidence for the viability of the different stages of this mechanism has been put forward by several different research groups.¹³⁰⁻¹³² There are now over 200 crystal structures of the enzyme deposited in the Protein Data Bank (PDB) which in combination with computational studies and biochemical assays have helped to determine that transition state stabilisation and substrate distortion are the two most significant factors in the rate that OMP decarboxylates.^{118,133-135}

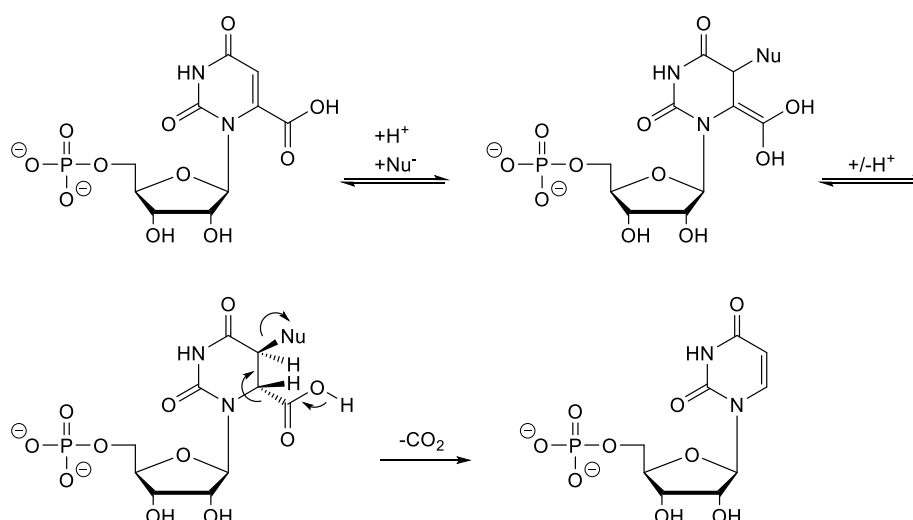


Figure 25 Michael addition mechanism of decarboxylation proposed by Silverman and Groziak.¹²⁸

The direct decarboxylation mechanism is believed to be initiated through distortion of the OMP carboxylate moiety caused by repulsive interactions with one of the key aspartate residues in the enzyme's active site.¹²⁴ This results in a loss of carbon dioxide and the formation of a vinyl carbanion intermediate. Stabilisation of the vinyl carbanion is believed to be the most important factor in the decarboxylation catalysis.¹³⁶⁻¹³⁹ The negative charge is thought to be stabilised by the amino group of one of the key lysine residues in the active site. In the final step, the same lysine is proposed to transfer a solvent derived proton to the C-6 position thereby neutralising the negative charge to form the desired product, UMP.

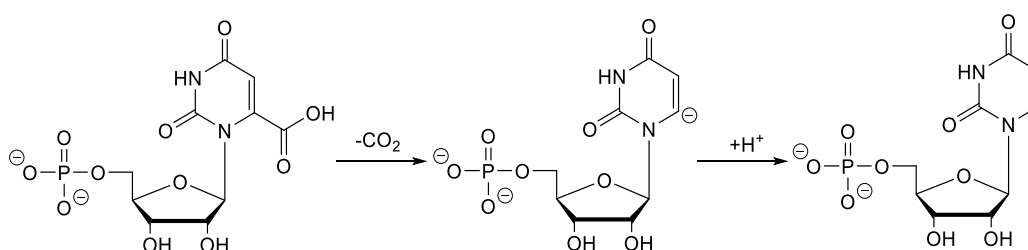


Figure 26 Direct decarboxylation mechanism.

A similar mechanism has been suggested in which the decarboxylation is postulated to occur in a concerted manner as shown in Figure 27.¹⁴⁰ In the concerted mechanism, the loss of carbon dioxide is proposed to occur simultaneously with protonation.¹⁴⁰ However, evidence supporting the existence of a vinyl carbanion disfavors the concerted mechanism and makes the stepwise mechanism set out in the direct decarboxylation theory the more likely route to UMP.

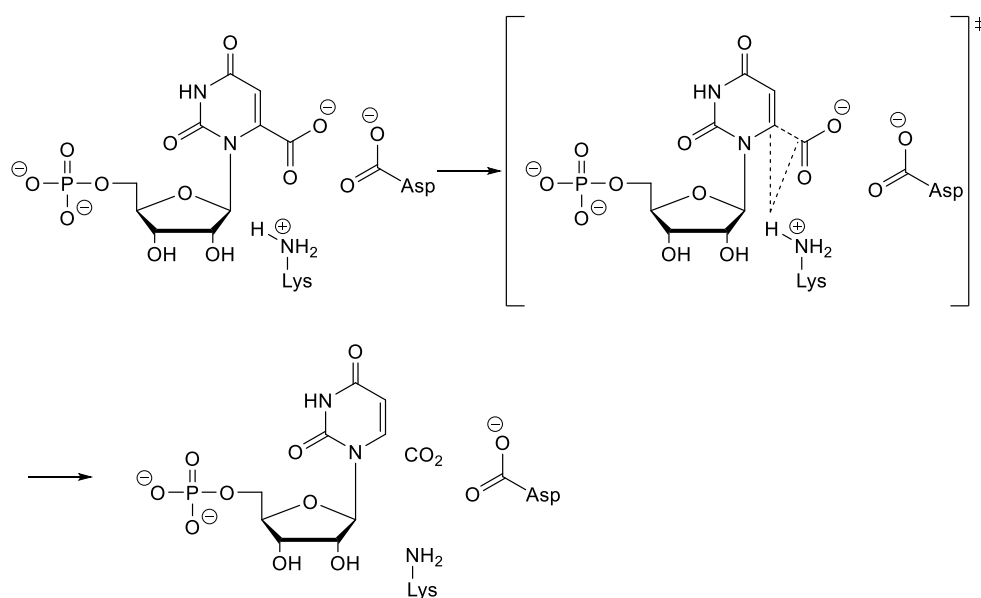


Figure 27 Concerted mechanism proposed by Appleby *et al.*¹⁴⁰

A study by Toth *et al.* found that when the decarboxylation of OMP by ODCase takes place in a mixed H₂O/D₂O buffer, the H/D isotope ratios of the product UMP matches the H/D ratio of the buffer.^{137,138} The result implies that the rate determining step of the reaction does not involve proton transfer making the concerted mechanism seem unlikely and lending support to the stepwise decarboxylation proceeding through formation of a vinyl carbanion. Further evidence for a vinyl carbanion was provided in a study that showed ODCase is capable of catalysing the exchange of the C-6 proton of UMP for a deuterium atom from solvent D₂O.¹⁴¹ The exchange goes *via* the formation of a vinyl carbanion in a reaction that is the reverse of the proton transfer that takes place in the decarboxylation of OMP.¹⁴¹ The pK_a of the C-6 proton of enzyme bound UMP has been calculated to be ≤22 which is around 10 units lower than the C-6 proton of 1,3-dimethyl uracil (pK_a 30-34) in water.^{141,142} This increase in acidity of H-6 in UMP when it is bound to the enzyme shows that ODCase must be able to stabilise a vinyl carbanion intermediate. Furthermore, it was found that ODCase from yeast catalyses the replacement of the C-6 proton in 5-fluoro-UMP approximately 3400 times faster than it does in UMP.¹³⁹ The presence of an atom of fluorine, the most electronegative element, at C-5 is thought to stabilise the negative charge that is built up at C-6 of the vinyl carbanion and increases the rate of exchange.

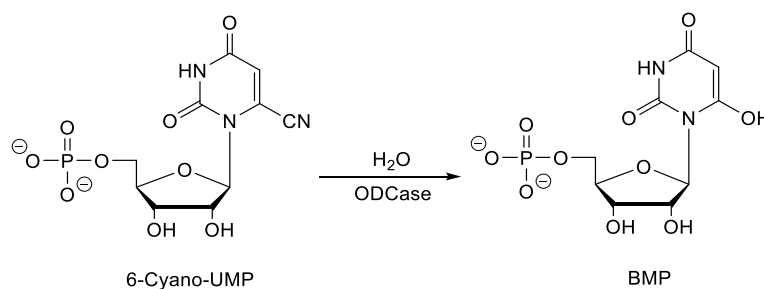
The 5'-phosphate group of OMP is also considered to play an important role in the rate of decarboxylation. It is believed that substrate OMP gains an intrinsic binding

energy from interaction of its 5'-phosphate with surrounding enzymic residues and that it is this energy that helps to force the negatively charged carboxylate group into the active site where decarboxylation can take place.^{115,143-145} Evidence of the importance of the 5'-phosphate group was obtained in a study that showed the rate of decarboxylation is dramatically decreased for orotidine as compared to OMP.^{146,147}

The other main contributing factor to decarboxylation is believed to be substrate distortion. Several crystal structures of ODCase in complex with OMP and other analogues have been obtained that show that C-6 substituents are not co-planar to the pyrimidine ring in the active site.¹³⁴ Molecular and quantum mechanics calculations predict that substrate distortion contributes between 10-15% to the overall rate of ODCase catalysis.¹³⁴

3.2.2 ODCase promiscuity

As well as carrying out the decarboxylation of OMP, ODCase displays enzymatic promiscuity as it has also been shown to carry out a pseudo-hydrolysis reaction. Whilst investigating potential inhibitors of ODCase, Fujihashi *et al.* found that 6-cyano-UMP was converted to 6-hydroxy-UMP (barbiturate-5'-monophosphate, BMP) by ODCase (Scheme 10).¹⁴⁸ Crystals obtained following incubation of 6-cyano-UMP with ODCase derived from *Methanobacterium thermoautotrophicum* were analysed by X-ray crystallography and showed the enzyme's active site bound to BMP.¹⁴⁸



Scheme 10 Pseudo hydrolysis reaction carried out by ODCase that converts 6-cyano-UMP to 6-hydroxy-UMP (BMP).

Incubation of 6-cyano-UMP under the same reaction conditions, but in the absence of any enzyme failed to produce any BMP showing it was ODCase that was catalysing the reaction.¹⁴⁸

As shown in the previous example, there is a degree of plasticity associated with the active site of ODCase meaning it is able to accept and interact with other substrates

apart from OMP. This ability to interact with other ligands is one of the reasons ODCase is such an interesting and attractive target when investigating new therapeutic treatments. Some of the currently known inhibitors of ODCase will now be examined in more detail as well as looking into the factors to be considered when designing new ones.

3.2.3 Known inhibitors of ODCase

Numerous research groups have investigated the design of inhibitors to target ODCase in an effort to probe its reaction mechanism and to develop new drug candidates in the treatment of malaria, cancer and viral infections. The majority of the inhibitors that have presently been developed are based on the structure of the enzyme's natural substrate OMP, but contain modifications at the C-5 and C-6 positions of the pyrimidine ring. The carboxylate group that is lost from OMP during the decarboxylation process is located at C-6 so most currently known inhibitors have looked to mimic the natural substrate by replacing CO₂ at this position with a similarly sized group. As discussed, when CO₂ is lost from OMP there is a build-up of negative charge around the C-6 position so functionalities that are capable of being transition state analogues have been of high interest. However, inhibitors of ODCase are not limited to being derivatives of OMP. The plasticity of the enzymes active site allows for incorporation of other nucleotide structures, with inhibitors derived from cytidine known, as well as inhibitors derived from purine nucleotides. In addition, there are also reports of some non-nucleotide inhibitors of ODCase.

The most potent inhibitor of ODCase currently known is BMP, which has been determined to have an inhibition constant (K_i) of 9×10^{-12} M against ODCase from yeast.¹⁴⁹ BMP is believed to be a transition state inhibitor with the negative charge from the hydroxyl group at the C-6 position mimicking both the negative charge of the carboxylate group of OMP and the negative charge that is present at C-6 during decarboxylation. Similarly, 6-aza-UMP, another of the most potent inhibitors of ODCase known, is also thought to mimic the carbanion transition state.¹⁵⁰

6-Aza-UMP has been investigated in anti-cancer therapies and has been shown to exhibit good activity against several clinical tumour models.^{150,151} Another analogue that has been extensively studied as a potential treatment for cancer is the unusual C-nucleoside, pyrazofurin. Like 6-aza-UMP, it has been found to display good anti-cancer activity, particularly against leukaemia cell lines.^{152,153} In addition, both compounds have also been identified as effective anti-viral agents, principally in the treatment of West Nile virus.¹⁵⁴ The compounds are thought to act as competitive

inhibitors of ODCase and pyrazofurin has even been evaluated in phase I clinical trials relating to its use in cancer treatments.¹⁵⁵ However, issues with patient toxicity have meant that further investigations into this compound as a treatment for cancer have been abandoned.^{155,156}

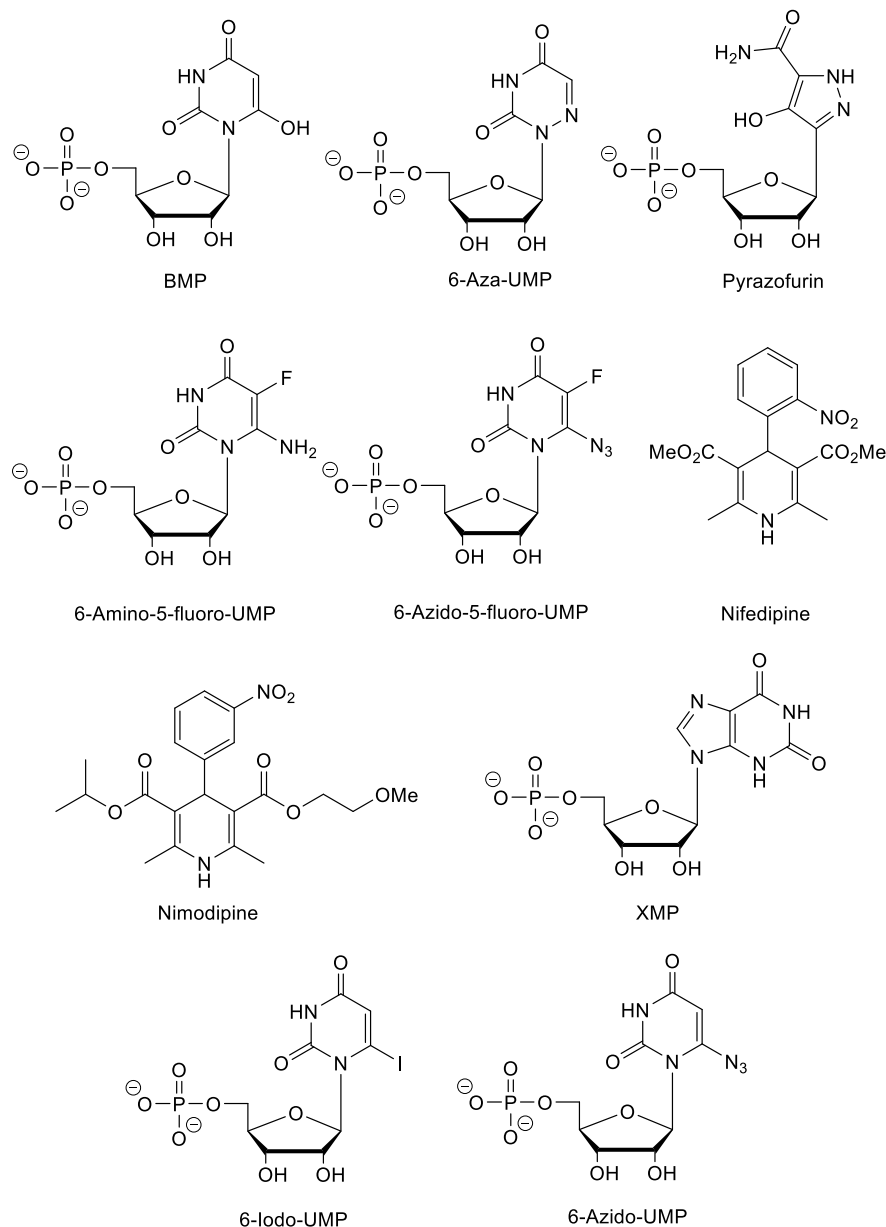


Figure 28 A selection of known inhibitors of ODCase shown in their monophosphate form.

Bello *et al.* found that derivatives of UMP that carry a fluorine atom at C-5 as well as being modified at C-6 show promise as chemotherapeutic agents against cancer.¹⁵⁷ They found that 6-azido-5-fluoro-UMP and 6-amino-5-fluoro-UMP covalently inhibit ODCase and their nucleoside derivatives show potent anti-cancer activity in cell based assays.¹⁵⁷ Their results show promise that ODCase inhibition could one day be used in the treatment of cancers.

Nifedipine and nimodipine are two non-nucleotide inhibitors of ODCase. Both drugs are known to be calcium channel blockers, but both have been found to display moderate inhibitory activity against ODCase with K_i values of 105 and 18 μM , respectively.¹⁵⁸ The binding of these two compounds was assessed in detail by Meza-Avina *et al.* in order to try and elucidate more information on how the active site of ODCase is able to accommodate ligands quite different from its natural substrate.¹⁵⁹ They were unable to crystallise the enzyme in complex with either of the two compounds so carried out computational docking studies instead to determine the key interactions between the ligands and the active site.¹⁵⁹ They found that both compounds fit nicely into the active site and the nitro moiety present in both structures interacts with an arginine residue known to be important for binding of the monophosphate group of the natural substrate OMP.¹⁵⁹ The side chain ester functionalities of both compounds were found to bind at the same site as where C-5 substitutions in the natural substrate interact.¹⁵⁹ These findings show that inhibitors of ODCase do not necessarily need to be derived from a nucleotide structure and so expands the possibilities and structural features to be considered when designing new inhibitor scaffolds.

Despite its primary function being the decarboxylation of a pyrimidine nucleotide, ODCase has been found to also bind purine nucleotides in its active site with xanthosine-5'-monophosphate (XMP) being known to inhibit yeast ODCase.^{117,160} While cytidine-5'-monophosphate (CMP) is known to be a weak inhibitor of ODCase, *N*-3 and *N*-4 oxygen derivatives of CMP show much improved potency which again goes to show the wide scope of ligands ODCase is able to accommodate.¹⁶¹

The development of ODCase inhibitors as a means to target the malaria parasite has received increasing amounts of attention in recent years and is indeed one of the main focuses of this thesis.^{162,163} Research into new antimalarial drugs is of critical importance as resistance to current treatments continues to grow. Kotra *et al.* discovered that 6-iodo-UMP binds irreversibly to *P. falciparum* ODCase by forming a covalent bond through displacement of the iodo moiety at C-6 by one of the key lysine residues of the active site.¹⁶⁴ This was a surprising result given that ODCase catalysis is not known to involve any covalent species. Subsequently, additional compounds capable of covalently inhibiting ODCase were identified, as detailed earlier in the section for the cases of 6-azido-5-fluoro-UMP and 6-amino-5-fluoro-UMP.¹⁵⁷ This finding prompted further studies of the nucleoside analogue, 6-iodouridine, and this compound was found to display potent anti-plasmodial activity

in *in vitro* testing.¹⁶⁵ Additional testing of 6-iodouridine as a potential new drug was carried out in malaria mouse models where it was assessed in combination with the well-known antimalarial artemisinin.¹⁶⁵ The compound was found to show an additive effect *in vitro* when combined with artemisinin as well as good efficacy in *in vivo* testing.¹⁶⁵ The same research group investigated other potential ODCase inhibitors to target malaria based on introducing functionalities at C-6 that are of a similar size to the carboxylate in the natural substrate.¹⁶⁶ They synthesised and tested 6-cyano, 6-azido, 6-amino and 6-methyl derivatives as inhibitors of plasmoidal ODCase. Of these compounds, the most promising appeared to be 6-azido-UMP which was found to also covalently inhibit *P. falciparum* ODCase.¹⁶⁶ However, unlike the case of 6-iodouridine, its nucleoside analogue showed no significant activity against *Plasmodia* cultures.¹⁶⁶

These results show promise that targeting of plasmoidal ODCase could produce a potential new treatment or drug candidate in the fight against malaria. The observation that ODCase can be irreversibly inhibited by C-6 substituted compounds containing good leaving groups opens up a new class of potential therapeutic agents to be considered when designing inhibitor molecules.

3.2.4 Design of inhibitors of ODCase

The purpose of this project was to investigate the design and synthesis of inhibitor molecules to target the enzyme ODCase, with the ultimate aim of generating molecules capable of exerting an anti-malarial effect. The previous section highlighted some of the known inhibitors of ODCase and before detailing the work carried out in this study, it is important to first consider what structural aspects could contribute towards the design of new inhibitor molecules.

Work carried out by Meza-Avina *et al.* sought to identify the characteristics required by an inhibitor in order to achieve potent inhibition of ODCase.¹⁵⁹ The generic pharmacophore they proposed was based on nucleotide ligands although they do note their awareness that non-nucleotide structures can act as ODCase inhibitors.¹⁵⁹ They identified three regions critical to binding in their empirical model as shown in Figure 29.¹⁵⁹

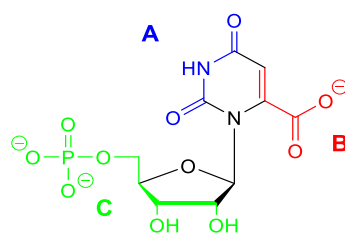


Figure 29 Empirical pharmacophore proposed by Meza-Avina *et al.* for inhibitors of ODCase.¹⁵⁹

Region A is thought to be important in terms of obtaining the correct orientation of the nucleobase in the enzyme active site.¹⁵⁹ Region B has been identified as important as the properties of the substituents at this site will greatly influence the potency and overall inhibitory ability of the ligand.¹⁵⁹ Region C, as mentioned earlier, has been found to confer a significant amount of binding energy and so is vital to obtaining and maintaining tight binding of the inhibitor molecule in the active site.¹⁵⁹ For the purpose of this thesis, all molecules were synthesised in their nucleoside form so that they have the potential to be tested in cell-based assays. The 5'-monophosphate group is highly polar and so would prevent the potential inhibitor molecules from entering the cell. As is usually the case when assessing drug molecules derived from nucleic acids, it is assumed that the nucleoside form will be phosphorylated by a nucleoside kinase upon entering the cell.

As discussed in the previous section, there are many inhibitors of ODCase that have already been reported which display varying degrees of potency. However, the replacement of the carboxylate group of the natural substrate with a nitro group has yet to be detailed which is somewhat surprising given that the nitro functionality is both isosteric and isoelectronic to a carboxylate. One possible reason for the lack of reports of 6-nitouridine could be due to the difficulties associated with nitrating nucleosides. More often than not, nitration reactions require harsh reaction conditions which are not always compatible with nucleoside chemistry. The glycosidic bond linking the sugar and base of nucleosides is particularly sensitive to acidic conditions making the common procedure of nitrating molecules with a mix of concentrated nitric and concentrated sulphuric acids totally unsuitable.

There is literature precedent for a nitro group replacing a carboxylate group and successfully inhibiting an enzyme. Firestine *et al.* investigated the enzyme aminoimidazole ribonucleotide carboxylase which catalyses the formation of 4-carboxy-5-aminoimidazole ribonucleotide (CAIR), shown in Figure 30.¹⁶⁷ They found that by replacing the carboxylate group of the product with a nitro group to produce

4-nitro-5-aminoimidazole ribonucleotide (NAIR), they could achieve potent inhibition of the enzyme.¹⁶⁷

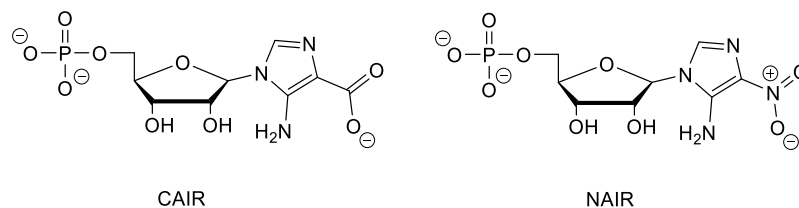


Figure 30 Structure of 4-carboxy-5-aminoimidazole ribonucleotide (CAIR) and its nitro analogue 4-nitro-5-aminoimidazole ribonucleotide (NAIR).

Work by Judice *et al.* provides a further example of a nitro group successfully replacing and mimicking a carboxylate group.¹⁶⁸ They found that replacement of a glutamate residue in the active site of Staphylococcal nuclease with its nitro analogue had little effect on catalytic activity over a certain pH range.¹⁶⁸ Therefore, if milder nitration conditions can be found then the synthesis of 6-nitrouridine could prove useful in the search for new inhibitors of ODCase.

In addition to 6-nitrouridine, another potential inhibitor of ODCase could be a uridine molecule substituted at the C-6 position with a tetrazole moiety as tetrazoles can act as a bioisostere for the carboxylate group.^{169,170} There is one previous report in the literature of the preparation of the C-6 tetrazole adduct, formed from 6-cyanouridine and NaN_3 .¹⁷¹ However, its ability to inhibit ODCase was not tested or, not reported, and so further work on this compound is required in order to assess its potential as an inhibitor.

Use of click chemistry could prove useful in the search for other new inhibitors of ODCase as click reactions have become increasingly popular in medicinal chemistry over recent years, due in part to their mild reaction conditions and high yields.¹⁷² One of the most commonly used click reactions is the copper catalysed azide alkyne cycloaddition (CuAAC) which produces triazole compounds.¹⁷³ Triazoles are stable under both reductive and oxidative conditions and are able to take part in hydrogen bonding making them useful functionalities to include in potential drug molecules.¹⁷² As yet, no C-6 substituted uridine triazoles have been reported. This could be due to them being ruled out as potential inhibitors due to the larger size of triazoles compared to carboxylates, but nevertheless the introduction of a triazole at the C-6 position of uridine would be an interesting possibility for a new inhibitor molecule.

Hydroxamic acids and boronic acids can also be used as carboxylic acid bioisosteres and so introduction of these functionalities at the C-6 position of uridine is another option worth considering when designing potential ODCase inhibitors.¹⁷⁰ Compounds that are able to mimic the transition state of the decarboxylation have already been shown to be important as BMP is currently the best known inhibitor of ODCase. Therefore, design of alternative compounds that can remain stable with a negative charge at or close to C-6 are of high priority.

3.3 Project aims

ODCase is clearly a remarkable enzyme that plays a vital role in the *de novo* synthesis of pyrimidine nucleotides. The reliance of certain pathogens on the pyrimidine *de novo* pathway presents an opportunity to target and inhibit the enzymes of the pathway as a means of therapeutic intervention. The vast amount of work that has gone into elucidating the reaction mechanism of ODCase has produced a variety of crystal structures of the enzyme, derived from several different species. The active site of the enzyme is now well mapped which makes it an attractive target for the design of potential new drug candidates. The aims of this project were to:

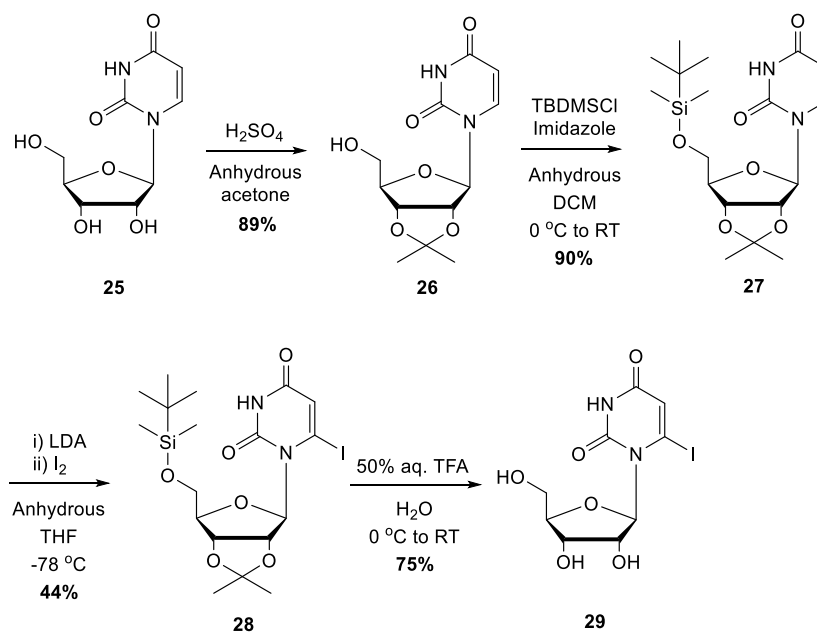
- 1) Design a series of inhibitors to target the enzyme ODCase by introducing moieties at the C-6 position of uridine that have the potential to mimic the carboxylate group that is in the natural substrate.
- 2) Investigate the use of molecular modelling and docking studies to predict the ability of potential inhibitors to interact with the active site of ODCase.
- 3) Assess the anti-plasmodial activity of final nucleoside compounds through biological testing.

Chapter 3 Results and discussion 2

The first stage of this project investigated the synthesis of 6-iodouridine. This compound has been reported previously in the literature by Kotra *et al.* and as mentioned in the introduction, it is known to be a potent inhibitor of ODCase.¹⁶⁴ It was identified as a key intermediate for use in this project as having an iodo group at the C-6 position allows for a range of potential transformations to take place. It was thought that it would therefore act as a good starting point for the synthesis of a variety of different analogues of OMP. In addition, one of the overall aims of the project was to send final compounds for biological testing and so it was reasoned that having a sample of 6-iodouridine tested would act as a good comparison for validating the results of the testing against those of previous studies.

3.4 Synthesis of 6-iodouridine

Scheme 11 outlines the synthetic route to 6-iodouridine. The preparation utilised conditions originally reported by Kotra *et al.* with some modifications made to the critical iodination step, which will now be discussed in more detail.¹⁶⁴



Scheme 11 Synthetic route to 6-iodouridine.

The synthesis began with protection of the three hydroxyl groups of the ribose sugar ring. This was required to prevent them from taking part in unwanted side reactions. One of the main advantages of nucleoside syntheses involving uridine is that protection of the nucleobase is rarely required unlike in the cases of adenosine, guanosine and cytidine. Protection of the three hydroxyl groups was achieved using

an isopropylidene group for the 2' and 3' positions and the silyl protecting group TBDMS for the 5'-OH. These groups were chosen as they are stable under basic conditions which is critical as the all-important iodination step involves the use of the strong base LDA.

The isopropylidene group is one of the most commonly used groups for the protection of 1,2-diols. It was introduced prior to the TBDMS group by suspending uridine in anhydrous acetone with a catalytic amount of concentrated sulphuric acid. Compound **26** was then isolated in 89% yield following column chromatography. Confirmation compound **26** had been successfully synthesised was provided by ¹H NMR which showed the presence of two singlets at 1.28 and 1.48 ppm, indicative of the two CH₃ groups of the isopropylidene group. In addition, the spectrum lacked signals for the 2' and 3'-OH groups, but still showed the 5'-OH as a triplet at 5.08 ppm.

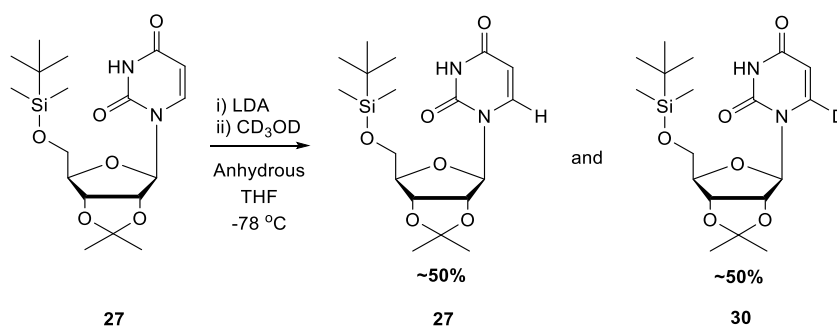
Protection of the 5'-OH of compound **26** was carried out using TBDMSCl and imidazole in anhydrous DCM. It has been found that in the absence of imidazole, the reaction between TBDMSCl and alcohols is very slow and low yielding whereas in its presence, reactions are found to proceed in high yield under mild conditions.¹⁷⁴ The imidazole functions by reacting with TBDMSCl to displace the chloride and produce an intermediate that is more reactive than the initial silyl compound. In this reaction, this intermediate was then attacked by the 5'-OH to produce the desired compound **27** in 90% yield. Analysis of the product by ¹H NMR spectroscopy showed singlet peaks at 0.09, 0.10 and 0.90 ppm, integrating to 3, 3 and 9 protons respectively. These peaks correspond to the two methyls and the *t*Bu of the TBDMS group showing that the reaction had been successful.

With the nucleoside now protected, the next stage was iodination of the 6 position of the pyrimidine ring. The procedure reported by Kotra *et al.* for the formation of 6-iodouridine involves a double deprotonation of compound **27** by LDA, firstly at the acidic ring NH then at C-6, after which I₂ is added as the electrophile.¹⁶⁴ For this reaction, it was decided to generate the LDA *in situ* from diisopropylamine and *n*-BuLi rather than buy it in ready-made and so the reaction was therefore carried out under strictly anhydrous conditions. The reaction however proved quite problematic, with consistently low yields being obtained. A one-off yield of 64% was consistent with that reported in the literature, but the yields of all other initial attempts were considerably lower, averaging at 24%. As the protected 6-iodo compound had been

identified as being a key intermediate for use in further reactions, it was decided to try and improve the reaction yield through optimisation of the reaction conditions.

An initial attempt to improve the yield investigated whether a change in the source of iodine would have any effect. However, when I₂ was substituted for 1,2-diiodoethane, no significant improvement in the average yield for the reaction was observed. Upon addition of I₂ to the reaction mixture containing the anionic nucleoside, a distinctive colour change shows whether the reaction is working. When the LDA has successfully generated the anionic nucleoside, the characteristic deep red colour of the iodine added is lost as it is consumed in the reaction. This colour consumption is not observed when 1,2-diiodoethane is used as the iodo source and since it was found not to offer any vast improvement to the yield obtained, it was decided to continue using I₂ as the electrophile.

It is believed that one of the main barriers to obtaining a good yield for this reaction is the need to carry out a double deprotonation. A deuterium labelling study was carried out in order to try and ascertain the percentage of deprotonation that was occurring at the C-6 position. For this study, compound **27** was deprotonated in an analogous fashion to the iodination reaction, but rather than trap out the anion formed with I₂, an excess of deuterated methanol (CD₃OD) was added. Scheme 12 shows the reaction.



Scheme 12 Deuterium labelling study carried out to determine the amount of deprotonation occurring at C-6.

Following work-up, the reaction was analysed by ¹H NMR which showed the presence of only the starting material and the C-6 deuterated analogue. As expected, almost all of the signals of the two compounds overlapped. The percentage of deuteration was determined by analysing the integration values for the H-5 and H-6 signals. The integration value of the H-1' signal was found to be 1 as it represents H-1' from both compounds and it was used as reference for the other signals. The signal corresponding to H-5 was observed as an apparent triplet that also integrated to 1 as it represents H-5 from both compounds. It appeared as

an apparent triplet due to the overlap between the doublet of compound **27** and the singlet of compound **30**. *H*-5 appears as a doublet in compound **27** due to splitting by the proton at the 6 position whereas when deuterium is incorporated at this position, as for compound **30**, *H*-5 appears as a singlet. The amount of deprotonation that took place was determined by integrating the signal for *H*-6. This was found to have an integration value of 0.5 indicating that approximately 50% of the reaction media analysed was starting material with the other 50% being deuterated. This implies that around 50% of the starting material is deprotonated at C-6.

The results of the deuterium study correlate with work that has been carried out to determine the pK_a of the C-6 proton. Sievers and Wolfenden used 1,3-dimethyluracil as a model compound for UMP to determine the pK_a of *H*-6.¹⁴² They used 1,3-dimethyluracil to avoid any difficulties that might have arisen due to deprotonation at other positions. From their results, they estimated that in water at 25 °C, the pK_a of 1,3-dimethyluracil is 34 ± 2 .¹⁴² As the structure of 1,3-dimethyluracil and UMP are very similar, it is highly probable that the pK_a of *H*-6 in UMP is also approximately 34. The base used in the experiments, LDA, has a pK_a of ~ 36 and so the position of equilibrium for the deprotonation of *H*-6 is likely to lie only slightly to the right-hand side.

The optimisation of the original reaction conditions had so far shown that changing the iodo electrophile had no great effect on the yield and that the base being used was only just strong enough to get deprotonation at the desired position to occur. In an effort to avoid using an even stronger, more hazardous base than LDA unnecessarily, further modifications of the reaction conditions were attempted. Initially, 2.2 equivalents of LDA were being generated at 0 °C before being cooled to -78 °C for the reaction. It was found that increasing the number of equivalents of LDA to 2.6 and generating and maintaining it at -78 °C increased the overall average yield from 24% to a point where yields in the region of 40% could be consistently achieved. Whilst ideally higher yields than 40% would be desirable, it was found that the reaction could be carried out under these conditions safely on a scale that allowed the production of multigram quantities of the desired protected iodo compound. The conditions also allowed for recovery of any unreacted starting material which meant it could be recycled in further reactions. These conditions were therefore adopted for use as standard for this reaction throughout the rest of the project.

Confirmation compound **28** had been made was provided by ^1H NMR and mass spectrometry. The ^1H NMR spectrum lacked the signal for *H*-6 at 7.70 ppm found in the starting material showing that a proton was no longer present at this position. Mass spectrometry then confirmed the presence of the iodo moiety as the correct mass for the desired product was obtained.

Deprotection of compound **28** was then required to obtain a sample of 6-iodouridine for biological testing. As both the isopropylidene and TBDMS groups are acid labile, it was found that both could be removed in tandem by suspending compound **28** in a 50% aqueous trifluoroacetic acid solution. Following purification, 6-iodouridine was isolated in 75% yield. A ^1H NMR spectrum of the isolated product lacked the two methyl group signals of the isopropylidene at 1.34 and 1.55 ppm and the TBDMS signals at 0.05 and 0.88 ppm that were present in the starting material showing compound **29** had been successfully synthesised.

Once the protected iodo compound **28** had been successfully isolated and an optimised route to it obtained, a list of OMP analogues that could potentially be derived from it was drawn up.

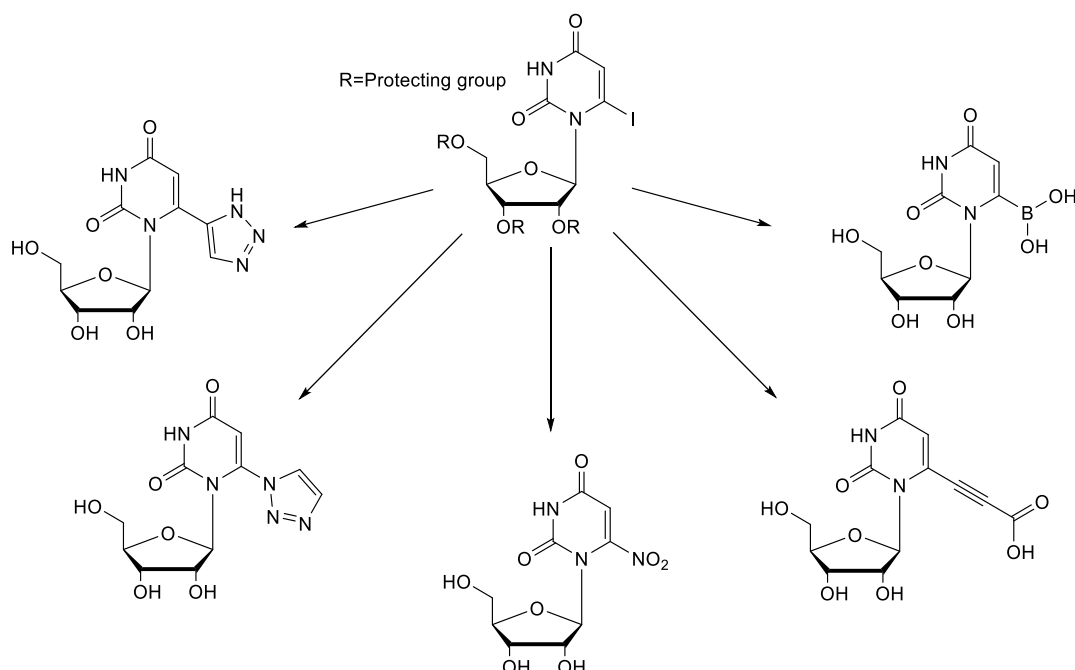


Figure 31 Potential analogues of OMP that could be derived from compound **28**.

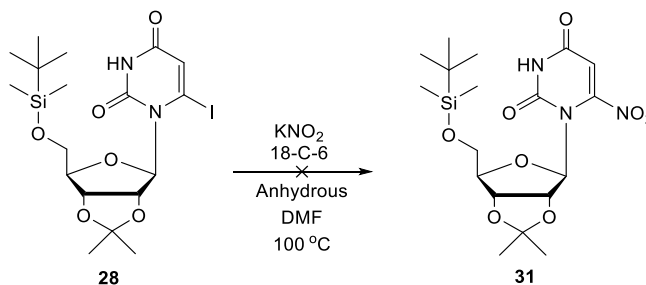
The main aim of the project was to carry out modifications to the C-6 position of uridine that would allow a range of OMP analogues to be synthesised. As discussed in the introduction, there are several moieties that are known to be good bioisosteres of a carboxylate group, but amongst them, the nitro group is known to

be one of the best as it is both isosteric and isoelectronic to a carboxylate. It was therefore decided to investigate the synthesis of 6-nitrouridine first.

3.5 Potential routes to 6-nitrouridine

3.5.1 Nucleophilic displacement using nitrite

There have been no previous reports in the literature of the preparation of 6-nitrouridine so the first step in the synthesis was to investigate potential routes of nitration. Having synthesised compound **28** with its halogen at C-6, a substitution reaction involving displacement of the halogen by nitrite was thought to be the simplest route to 6-nitrouridine. It was decided to first attempt nitration using the reaction conditions reported by Saito *et al.* for the nitration of guanosine described in Chapter 2 as these conditions have been proven to be mild enough for use in nucleoside chemistry.¹⁰⁴ Although in the case of Saito's nitration the reaction was carried out on a bromo substituted nucleoside, it was thought use of an iodo nucleoside instead would not have much of an effect overall. Scheme 13 shows the reaction.



Scheme 13 Attempted formation of 6-nitrouridine using KNO₂ and 18-crown-6.

The conditions for conversion of 8-bromoguanosine to 8-nitroguanosine reported by Saito were 1 equivalent of nucleoside to 10 equivalents of KNO₂ and 18-crown-6 with DMF as solvent, all heated to 100 °C.¹⁰⁴ Unfortunately, no nitrated product could be isolated when they were applied to compound **28**. The reaction was monitored by TLC which showed the formation of several new components in the reaction mixture, but none could be isolated and characterised by spectroscopic means.

The reaction was repeated, but rather than heat to 100 °C again, this time the reaction mixture was kept at room temperature and half the number of equivalents of KNO₂ and 18-crown-6 were used in an attempt to make monitoring of the reaction

and interpretation of the compounds forming easier. Again, however, no nitrated nucleoside was isolated from the reaction mixture. There was one component on the TLC plate that was not visible under UV light, but appeared as a black spot when the plate was stained using sugar stain. Initially, it was thought that this could be the desired nitro compound as nitration was expected to alter the absorption characteristics of the uracil base. However, when isolated, a combination of NMR and mass spectrometry showed that the compound was actually a protected sugar that had lost the nitrogenous base (Figure 32).

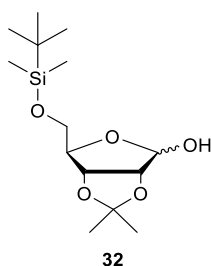
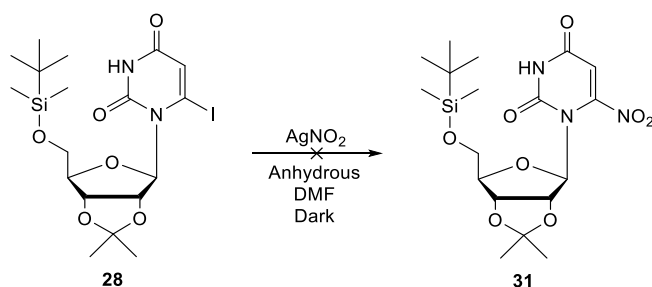


Figure 32 Abasic site formed during attempted nitration.

This was a disappointing result, although the fact that the base has been lost from the sugar ring could indicate that the nitro compound formed, but the glycosidic bond was subsequently cleaved. Nitration of the base will increase the nucleosides susceptibility to be hydrolysed due to the electron withdrawing power of the nitro group making it a better leaving group. This is consistent with what has previously been observed for 8-nitroguanosine which readily depurinates. However, given that none of the desired compound was isolated, alternative reaction conditions were tried.

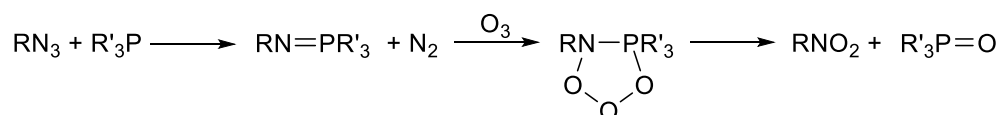
It was decided to attempt the reaction using AgNO_2 as the nitro source. The well-known affinity of silver for halide ions was the reasoning behind the switch from KNO_2 to AgNO_2 as it was hoped that the Ag^+ ions may be able to interact with the iodine atom at C-6 and encourage the desired reaction to take place (Scheme 14). Again, unfortunately no nitrated uridine was obtained. After stirring overnight, only starting material could be detected by TLC.

Scheme 14 Attempted nitration using AgNO_2 .

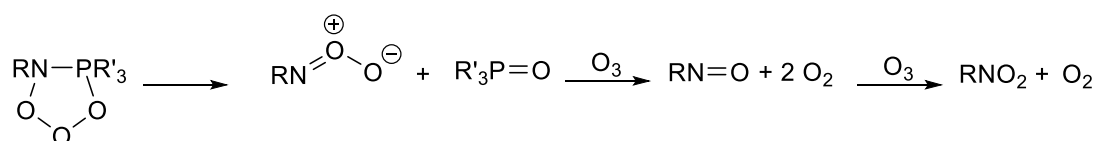
3.5.2 Oxidation of C-6 azidouridine

A literature search showed that Corey *et al.* had reported a procedure that produced organic nitro compounds under relatively mild conditions.¹⁷⁵ They had devised the method for production of a nitro galactose intermediate that they found could not be synthesised using conventional methods such as silver nitrite displacement.¹⁷⁵ The general reaction scheme for the method, shown in Scheme 15, involves oxidation of an azide starting material. This azide is then reacted with triphenylphosphine to give a phosphine imine that when oxidised with ozone was found to give a nitrated product. The exact mechanism by which the phosphine imine is converted to a nitro compound is not known exactly, but Corey speculated that it requires at least 3 equivalents of ozone and could possibly take place as shown in Scheme 15.¹⁷⁵

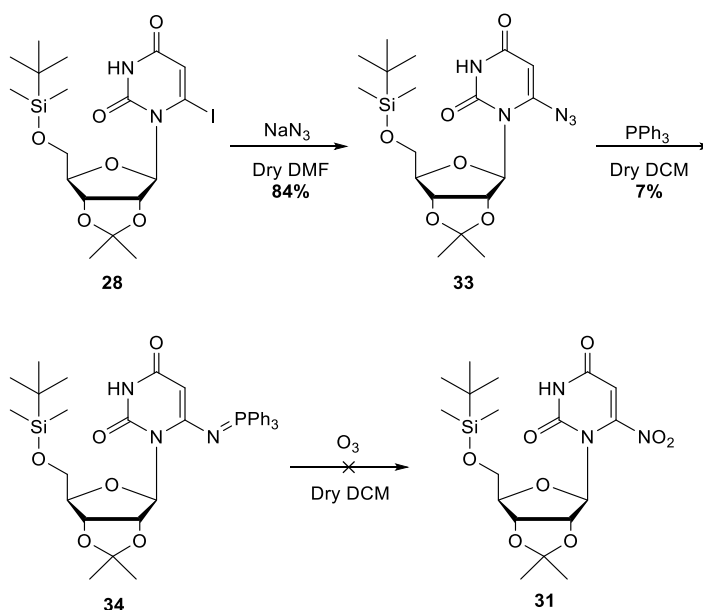
a)



b)

Scheme 15 a) Overall reaction to produce organic nitro compound starting from organic azide b) Possible mechanism by which ozone causes formation of organic nitro compound.¹⁷⁵

Corey *et al.* found that simple $\text{S}_{\text{N}}2$ displacement of either a halide or a sulfonate by azide produced the compounds they needed for their starting material.¹⁷⁵ It was therefore decided to attempt this method of nitration by again starting from the protected iodo compound **28**, reacting with sodium azide then carrying out the Corey nitration using ozonolysis. Scheme 16 shows the route followed.



Scheme 16 Attempted route to 6-nitouridine.

Reaction of compound **28** with sodium azide in DMF produced the desired protected 6-azidouridine **33** in a good yield of 84%. Compound **33** was confirmed to have been made by mass spectrometry which showed a peak at 462.18 which is consistent with the $[M+Na]^+$ ion of the desired compound. A strong absorption at 2136.43 cm^{-1} , indicative of an azide, in the IR spectrum of compound **33** further confirmed its identity.

An initial attempt at the nitration was carried out without attempting to isolate any intermediates. The azido nucleoside was reacted with PPh_3 and then directly exposed to a saturated solution of ozone in DCM. After subjecting the reaction mixture to several short bursts of ozone it was purged and worked up. However, analysis of the crude product showed predominantly starting material and so it was decided to repeat the reaction, but attempt to isolate the phosphine imine to show that it was forming.

Compound **33** and PPh_3 were stirred in anhydrous DCM for 1 hour after which TLC showed no remaining starting material. Following column chromatography, one compound was isolated which a combination of ^1H NMR, ^{13}C NMR, ^{31}P NMR and mass spectrometry showed to be the protected phosphine imine **34**. A peak in the ^{31}P NMR at 14.39 ppm was too high for PPh_3 which comes around -6 ppm and too low for triphenylphosphine oxide as this peak comes around 29 ppm.¹⁷⁶ It was in the right region however for a $\text{N}=\text{P}$ bond which literature values state are ~ 7 ppm depending on what else is bonded to the nitrogen.¹⁷⁶ Significant changes in the ^1H NMR were observed for $H-1'$ and $H-5$ on going from the azide to the phosphine

imine. The *H*-5 proton shifted by approximately 1 ppm from 5.50 ppm in the azide to 4.43 ppm in the phosphine imine. The *H*-1' proton showed a similar shift, but in the other direction going from 6.09 ppm in the azide to approximately 7.26 ppm in the phosphine imine. The peak for *H*-1' in the phosphine imine came at the same shift as the CDCl₃ peak, but its presence was confirmed by HSQC which showed that C-1' coupled to this peak. Figure 33 shows potential resonance forms of the phosphine imine that could explain these changes in chemical shift.

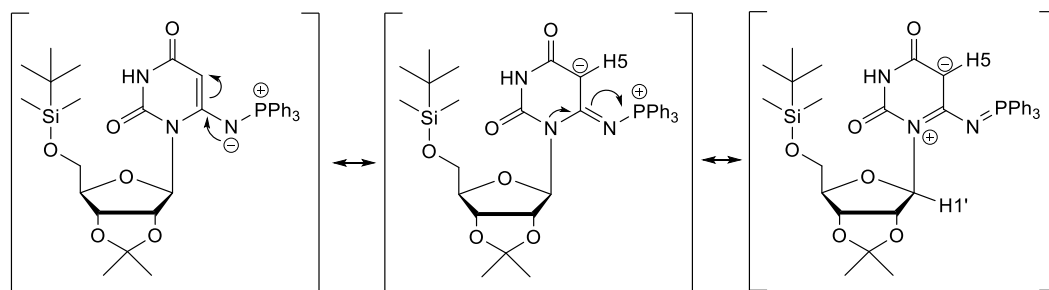


Figure 33 Potential resonance forms of compound **34**.

A resonance form for the phosphine imine can be drawn that puts a negative charge at C-5 and a positive charge at N-1. This would explain the shift to a lower ppm for *H*-5 as it is directly next to a negative charge and the shift to higher ppm for *H*-1' as it is directly next to a positively charged nitrogen. The ¹³C NMR provided further proof that the phosphine imine had formed as the C-6 signal, which is in close proximity to the phosphine, was split into a doublet by the phosphorus atom. Mass spectrometry also showed a peak at 696.26 for the [M+Na]⁺ ion of the desired protected phosphine imine.

Despite obtaining clear spectroscopic evidence that the protected phosphine imine **34** could be obtained, it was found that it could only be isolated in low yield (7%). Nevertheless, a second attempt at the nitration was carried out. Formation of the phosphine imine was monitored by TLC. Upon formation, the reaction mixture was transferred to a saturated solution of ozone in DCM. Unfortunately, following work up, no nitrated product could be detected in the crude material. This could be due to the fact that although the phosphine imine forms, it does so in only a small amount and further reaction does not take place.

3.5.3 Miscellaneous nitration attempts

Several further attempts at nitration were carried out which include; use of TEMPO as a catalyst in a radical based method, a conjugate addition to the 5,6-double bond and use of Saito's nitration conditions with alternative protecting groups.^{104,177-179}

However, no nitrated product has been obtained. Whilst this is a disappointing result, the lack of reports in the literature on the formation of 6-nitouridine shows achieving nitration at the C-6 position of uridine is not easily accomplished and further work in the area is required. It was decided at this stage to move on with the project and investigate other analogues of OMP that could be synthesised.

3.6 Click chemistry

Having successfully synthesised the protected 6-azido compound **33** in good yield, it was decided to utilise this compound further to carry out some click chemistry. Prior to the start of the project, as discussed in the introduction, the mild conditions used in click reactions were identified as showing great potential for synthesising triazole and tetrazole moieties at C-6 that could produce promising OMP analogues. Of particular interest was formation of the carboxy triazole uridine derivative shown in Figure 34 due to its similarity to OMP. But, before a synthetic route to it was devised, a test reaction was carried out between compound **33** and 5-chloro-1-pentyne which was immediately available in the laboratory.

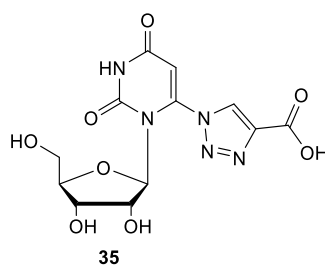
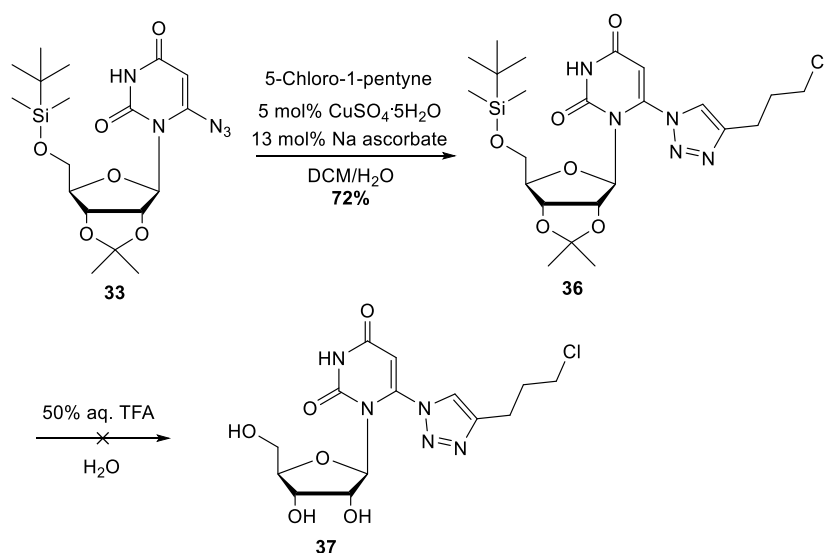


Figure 34 C-6 substituted uridine carboxy triazole identified as potential inhibitor of ODCase.

Scheme 17 shows the reaction between compound **33** and 5-chloro-1-pentyne. Reaction conditions used were based on those reported by Ferreira *et al.* for the formation of 1,2,3-triazole glycoconjugates.¹⁸⁰



Scheme 17 Test reaction between compound **33** and 5-chloro-1-pentyne.

Reactions of this type are known as copper-catalysed azide-alkyne cycloadditions (CuAAC). The presence of the copper catalyst in the reaction mixture allows formation of only the 1,4-isomer of the product, whereas, the reaction done in the absence of copper gives both 1,4 and 1,5-isomers from the 1,3-dipolar cycloaddition. Figure 35 shows the current most widely accepted mechanism for CuAAC reactions.¹⁸¹ The reaction is initiated through coordination of Cu(I) to the alkyne to give a copper acetylide intermediate. The Cu(I) species is most commonly generated *in situ* through reduction of a Cu(II) species by a reducing agent such as sodium ascorbate. A second Cu(I) is then recruited to form the catalytically active complex which bears both a σ and π bound copper atom. This complex can reversibly bind to the organic azide forming a copper-azide-acetylide metalocycle. Cyclisation can then take place followed by protonation and release of the triazole product.

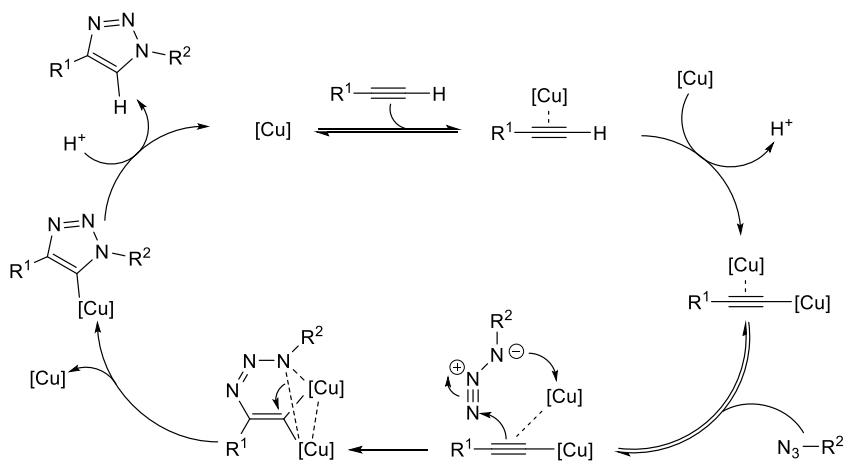
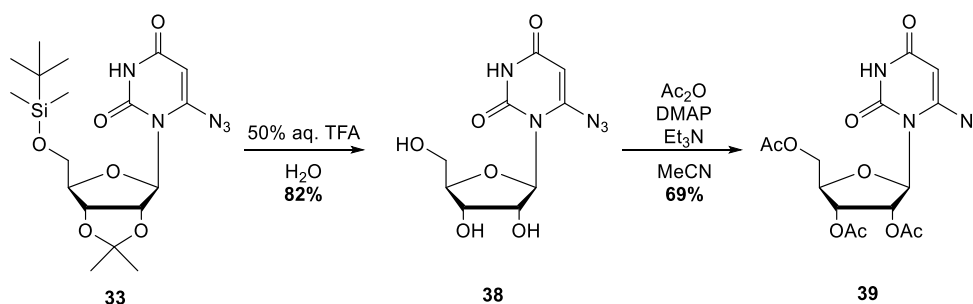


Figure 35 Catalytic cycle for CuAAC reactions.¹⁸¹

Formation of the desired triazole **36** was carried out successfully in 72% yield. However, when deprotection of the isopropylidene and TBDMS groups was attempted using the same conditions as had removed them from the 6-iodo compound **28**, it was found compound **36** broke down. None of the deprotected triazolic nucleoside could be detected or isolated, possibly due to loss of the nucleobase or interference in the deprotection from the chloro propyl side chain. It was decided to switch to different protecting groups that could be removed under milder conditions. Acetyl protecting groups were chosen as they can be removed using methanolic ammonia and their electron withdrawing nature helps to stabilise the glycosidic bond. There is a requirement during the iodination reaction for the protecting groups to be base stable so it was decided to keep using the isopropylidene and TBDMS groups until the azide stage, then perform a protecting group swap. Scheme 18 summarises these reactions.



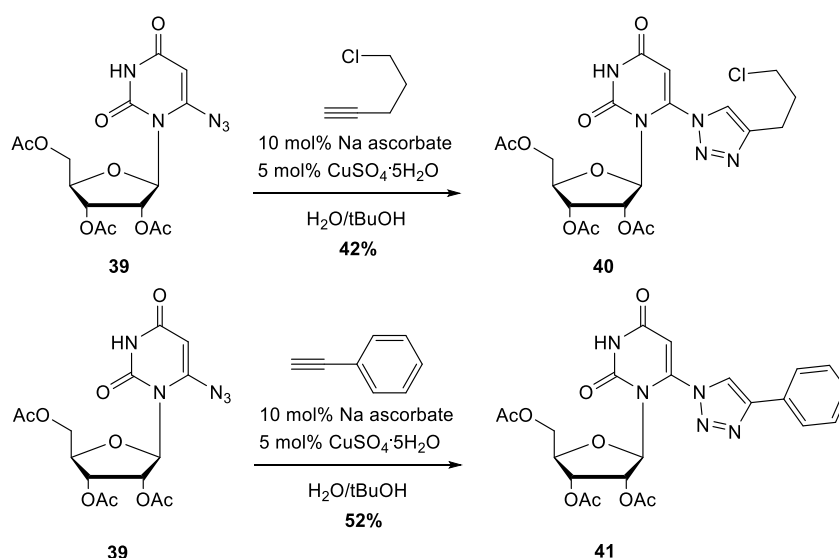
Scheme 18 Protecting group swap performed on 6-azidouridine.

The isopropylidene and TBDMS groups were successfully removed from compound **33** using a 50% aqueous TFA solution in 82% yield. ¹H NMR showed no protecting group signals and a mass spectrum of the product had a peak at 308.06 for the [M+Na]⁺ ion of the desired deprotected nucleoside. A sample of 6-azidouridine **38** was kept for biological testing. In order to reprotect with acetyl groups, compound **38** was suspended in anhydrous MeCN with acetic anhydride, DMAP and triethylamine using the same procedure that had been used for formation of tri-O-acetylguanosine. The reaction proceeded in a good yield of 69% and confirmation the reaction had been successful was provided by ¹H NMR which showed the presence of three singlets at 2.05, 2.08 and 2.10 ppm for the three methyl groups of the acetyls.

Having successfully changed the protecting groups, compound **39** was reacted separately with both 5-chloro-1-pentyne and phenyl acetylene, again in the presence of copper sulphate and sodium ascorbate (Scheme 19). For both of these

reactions, it was decided to swap the reaction solvent from DCM/H₂O to *t*BuOH/H₂O to achieve better solubility of the reactants.

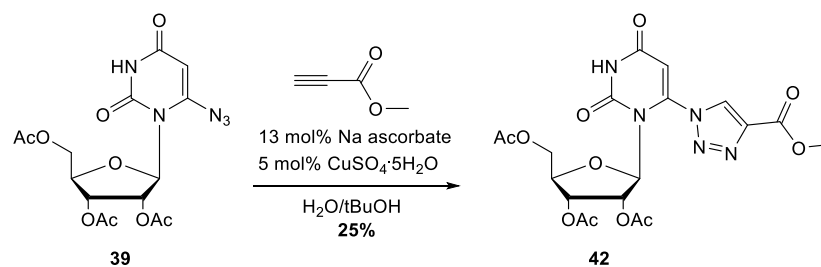
Both reactions were found to produce the desired triazolic product, although only in fairly modest yields. Confirmation compound **40** had been successfully isolated was provided by ¹H NMR which showed peaks due to the propyl side chain of the triazole at 2.20-2.27, 2.99 and 3.62 ppm as well as its mass spectrum which contained a peak at 514.13 for the [M+H]⁺ ion. ¹H NMR was also used to confirm compound **41** had been isolated. It showed characteristic signals between 7.42 and 7.88 ppm in the aromatic region integrating to five protons, indicating successful incorporation of the phenyl ring.



Scheme 19 Click reactions between compound **39** and 5-chloro-1-pentyne (top) and phenyl acetylene (bottom)

The change in protecting groups may have contributed to the moderate yields, but this was necessary as acidic deprotection with a triazole present had been found to be unsuccessful.

Having established click reaction conditions suitable for synthesising substituted triazoles at the 6-position of uridine, attempts were made to synthesise the carboxy triazole **35** by reaction with methyl propiolate.



Scheme 20 Reaction between compound **39** and methyl propiolate.

TLC showed the formation of new components in the reaction mixture, but it was found to be extremely sluggish. Isolation of the new components was carried out using column chromatography, but the desired triazole **42** was found to form as only the minor component of the reaction, in a disappointing yield of 25%. Its formation was confirmed by ^1H NMR which showed a singlet at 3.97 ppm for the methoxy group in addition to a singlet at 8.98 ppm for the single proton of the triazole ring. A peak in its mass spectrum at 518.11 is correct for the $[\text{M}+\text{Na}]^+$ ion of the desired triazole compound **42**.

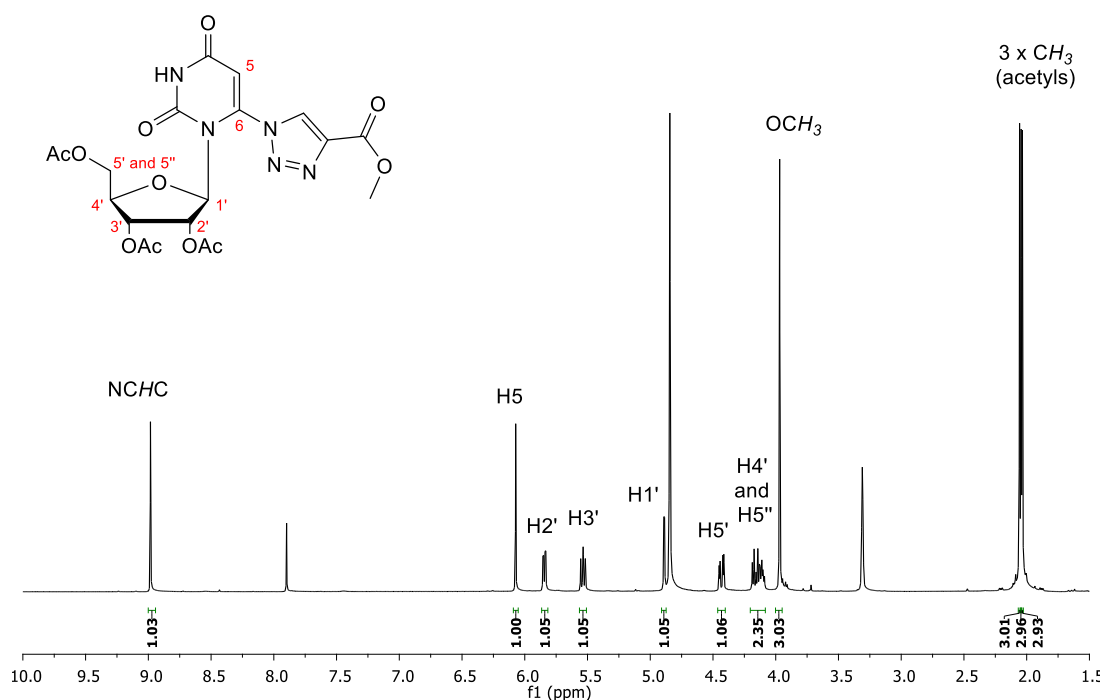


Figure 36 ^1H NMR of the desired methoxy ester triazole **42** recorded in CD_3OD .

The major component isolated following column chromatography and extensive spectroscopic analysis (^1H NMR, ^{13}C NMR, COSY, HSQC, HMBC, mass spectrometry) was found to be the compound shown in Figure 37.

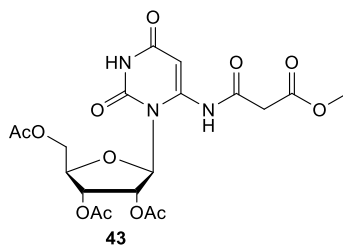


Figure 37 Unexpected product isolated from click reaction between compound **39** and methyl propiolate.

Formation of compound **43**, and as the major product of the reaction, was unexpected as CuAAC reactions generally produce only 1,4-substituted triazoles in high yield. Upon isolation of compound **43**, determination of its structure proved quite tricky. Its ^1H and ^{13}C NMR spectra showed the presence of signals corresponding to the sugar, nucleobase and protecting groups implying that it was not a decomposition product.

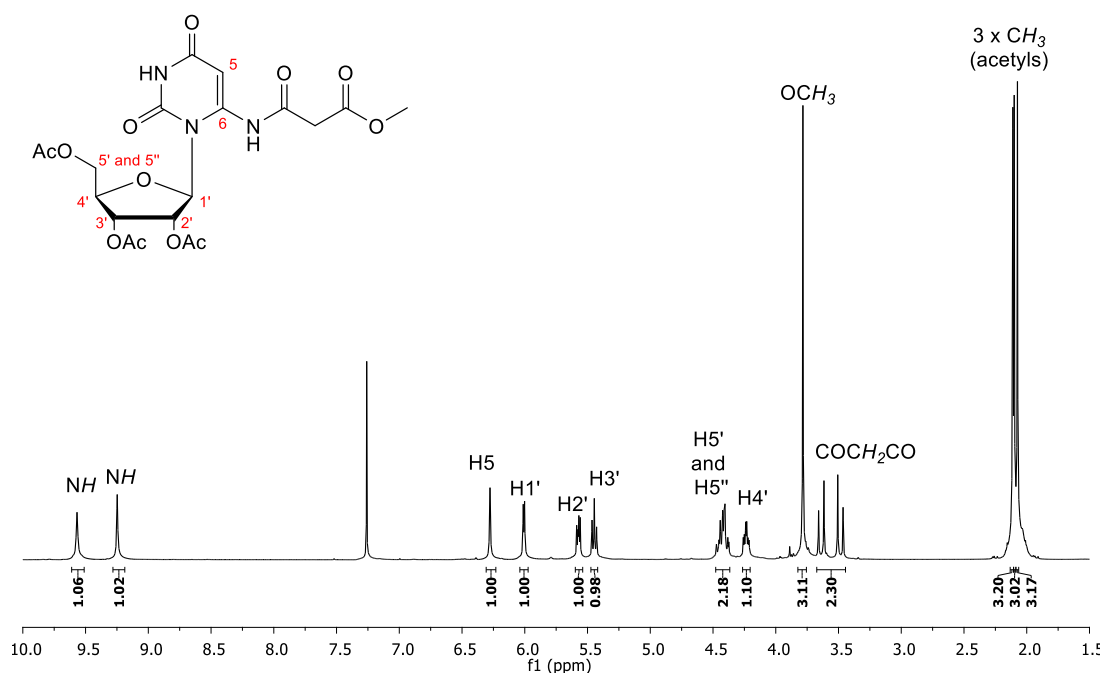


Figure 38 ^1H NMR of the unexpected product **43** recorded in CDCl_3 .

However, it lacked a signal that could be due to the triazole proton, but did show signals that suggested the presence of an ester. Addition of D_2O to the NMR sample revealed that two exchangeable protons were present at 9.25 and 9.57 ppm. In addition, the ^1H NMR spectrum contained an AB quartet system that integrated to two protons at 3.46-3.66 ppm which 2D NMR analysis showed was due to a CH_2 group. When the D_2O exchange was carried out, this AB quartet was found to slowly

diminish, indicating likely proximity to a C=O group. The presence of a singlet corresponding to an OCH₃ group indicated a reaction between the nucleoside and methyl propiolate had occurred, but the absence of a triazole proton and the presence of a CH₂ moiety implied it was not the desired reaction that had taken place.

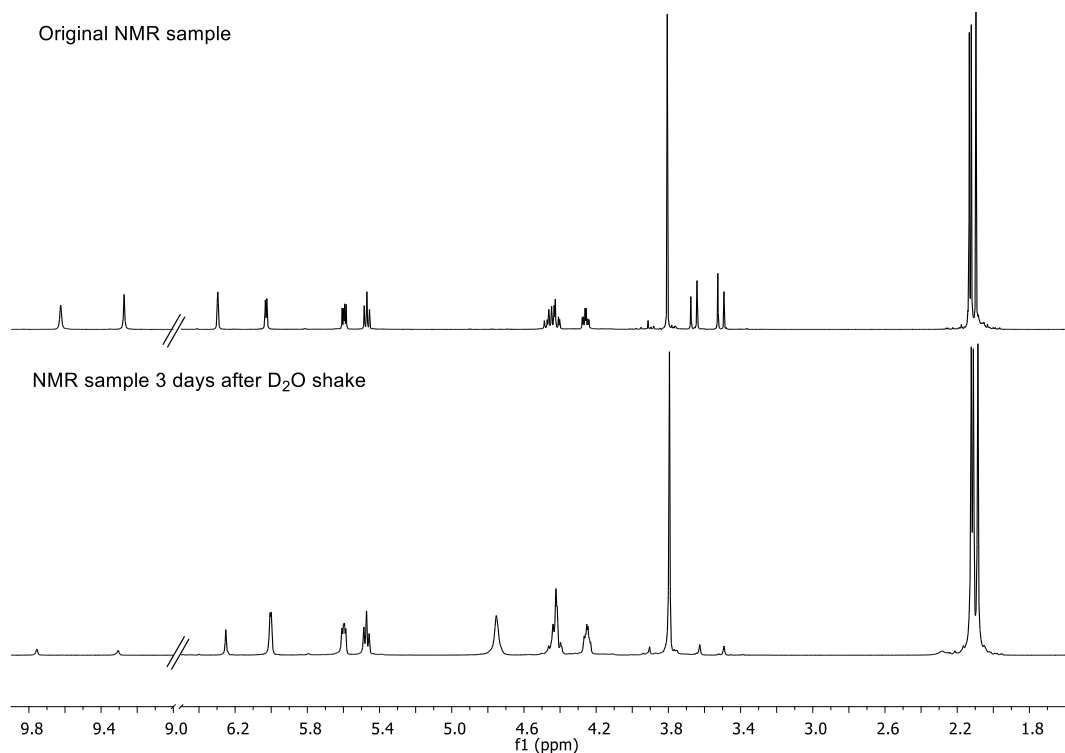


Figure 39 D₂O exchange carried out on ¹H NMR sample of compound **43**. The top spectrum shows the initial sample in CDCl₃ and the bottom spectrum is the same sample 3 days after addition of D₂O.

A mass spectrum of compound **43** found a peak at 508.12 which is consistent with the mass of the [M+Na]⁺ ion of the structure shown in Figure 37. By considering what possible side reactions could occur given the reagents used, in combination with the spectroscopic data, the structure of compound **43** was deduced. It was initially thought to form following the mechanism shown in Figure 40. Following reaction of the nucleosidic azide component with the alkyne to form the triazole, it was thought the lone pair on N-1 of the triazole could delocalize resulting in loss of molecular nitrogen and formation of a carbene. Carbenes are known to be highly reactive and so rapid shift of the adjacent proton was thought to occur in order to stabilize the structure. Attack of solvent water can then take place which following proton rearrangement would produce the unexpected compound **43**.

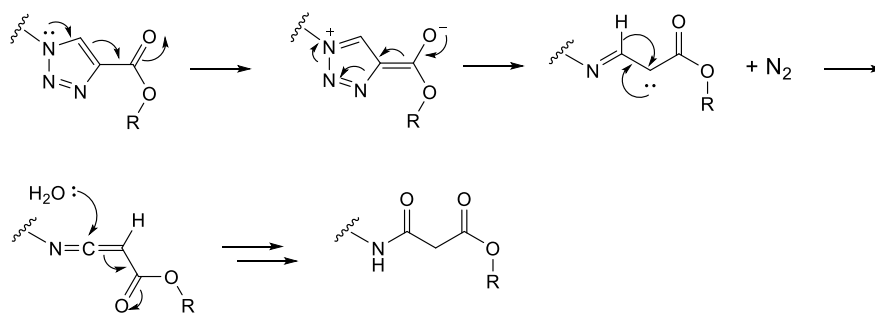


Figure 40 Possible mechanism by which compound **43** forms.

A search of the literature showed compounds similar to compound **43** are known, but to form them from triazoles requires flash vacuum pyrolysis conditions and presumably proceeds *via* the mechanism shown in Figure 40.¹⁸² This called into doubt whether the structure elucidated for compound **43** was correct as the reaction was carried out at room temperature and not the high temperatures required for pyrolysis to occur. A control reaction was therefore carried out on the small amount of pure methyl ester triazole **42** obtained from the reaction. It was first dissolved in *t*BuOH/H₂O to see whether any decomposition occurred and then re-subjected to the reaction conditions and monitored to see if any rearrangement to compound **43** could be observed. However, it was found that compound **42** was quite stable to the reactions conditions with no decomposition at all observed. This was in agreement with the finding that harsh vacuum pyrolysis conditions are required for the triazole to decompose.

A further search of the literature found studies carried out by Chang *et al.* and Baskaran *et al.* had reported findings that when reacted together in the presence of a copper catalyst, azides and alkynes do not always give 1,4-triazoles.^{183,184} The work of both groups investigated the synthesis of amidines and they found that coupling of a sulfonyl azide and an alkyne with a copper catalyst followed by treatment with an amine produced amidines in good yield. A role for copper in the mechanism was advocated by both groups.^{183,184} Figure 41 shows one example from the work of Baskaran.¹⁸³ In the example, the electron withdrawing effect of the tosyl group connected to the azide, in combination with the presence of the ester connected to the alkyne, is thought to help promote this type of reaction. The reactivity of the azide component towards the Cu-acetylide formed in the reaction is thought to play a part in determining what product is formed. For the amidine product to form, following loss of nitrogen, the amine present must attack the ketenimine intermediate to give the desired product.

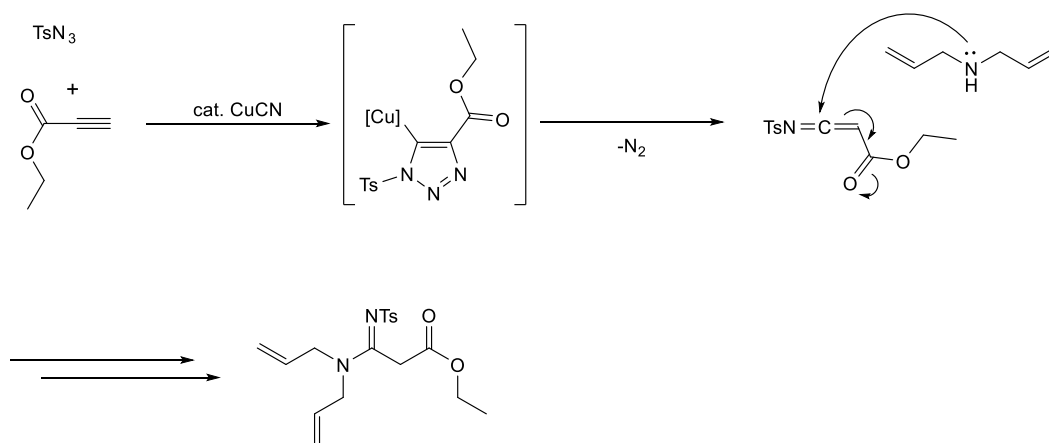
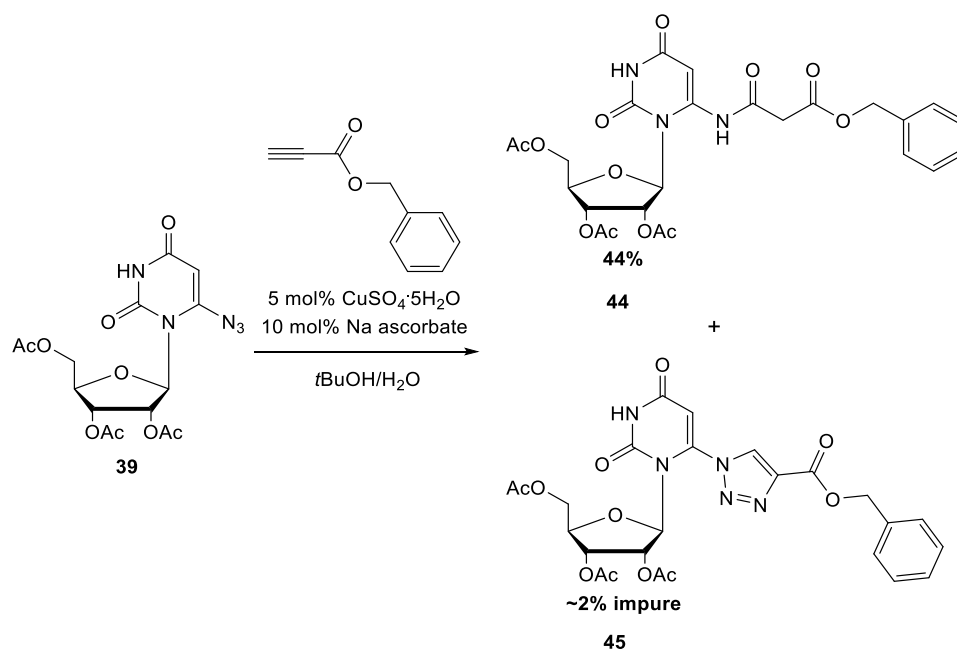


Figure 41 Possible mechanism by which amidines form from reaction of azides, alkynes and amines.¹⁸³

In the reaction reported here, although a sulfonyl azide is not used, the nucleosidic azide is somewhat electron deficient due to the electron withdrawing effect of the carbonyl at C-4 in conjugation with the 5,6-double bond. Although some formation of the desired triazole product is observed, formation of compound **43** through a pathway similar to the one described by both Chang and Baskaran can be envisaged. Formation of an amidine product does not occur as no amine is present, but the water present as a co-solvent could carry out a similar attack on a ketenimine intermediate to form compound **43**. The initial mechanism postulated in Figure 40 is likely to be the route through which compound **43** forms, apart from that copper is thought to remain associated with the triazole at the beginning of the mechanism and helps to initiate loss of the nitrogen.

An analogous reaction was carried out using benzyl propiolate rather than methyl propiolate to see whether it had any effect on the amount of 1,4-substituted triazole produced (Scheme 21). Once again, the major component of the reaction was the analogous benzyl derivative of compound **43** which was isolated in 44% yield and only milligram quantities of impure 1,4-benzyl ester triazole **45** could be isolated.

Although a small amount of the methoxy triazole **42** was isolated, it seems the reaction conditions and reagents used disfavour formation of 1,4-triazoles of this type. The earlier success in click reactions with 5-chloro-1-pentyne and phenyl acetylene show the nucleoside azide **39** is able to take part in click reactions, but it seems when in combination with alkynyl esters in the presence of copper, the favoured reaction pathway leads primarily to the formation of the unexpected products described.

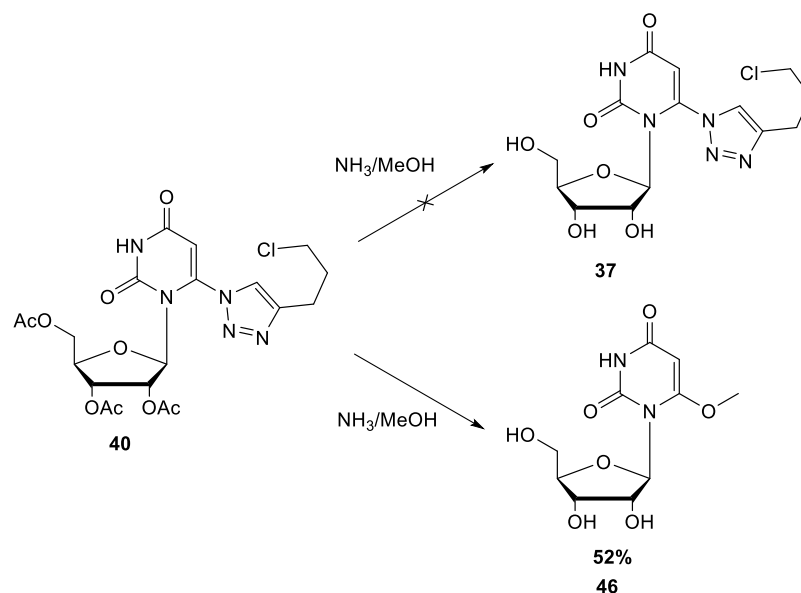


Scheme 21 Click reaction between compound **39** and benzyl propiolate.

It appears unlikely at present that the reaction could be scaled to produce larger quantities of the desired triazoles as there does not seem to be a way to prevent the side reaction which results in loss of nitrogen and formation of a ketenimine intermediate from happening. The apparent low reactivity of the starting azide material is thought to also contribute to the issues encountered. Although successful reactions were achieved with 5-chloro-1-pentyne and phenyl acetylene, the yields obtained for the click reactions are much lower than reported literature values. The reactions were found to be extremely slow which could be due to decreased reactivity of the azide due to conjugation with the uracil base. The slowness of the reaction could contribute to the formation of these unexpected products as the desired triazoles are not forming quickly enough allowing alternative reactions to compete. Further work on optimising the conditions when carrying out click reactions with these reagents is required in order to produce the desired 1,4-ester triazoles, but unfortunately due to time constraints this was not pursued further in this project.

Although triazole **40** with its chloro-propyl side chain was not originally a target as an OMP analogue, it was decided that as it had been successfully synthesised it should be deprotected for testing. The earlier attempt at deprotecting the isopropylidene and TBDMS analogue under acidic conditions had failed, but it was hoped that removal of the acetyl protecting groups by methanolic ammonia would be more successful (Scheme 22). However, it was found that following column

chromatography, the only product that could be isolated was 6-methoxy uridine. This was a quite a surprising result, as although loss of the triazole moiety from the 6 position is feasible, displacement by methanol was unexpected. Given the reaction conditions, and that ammonia is a better nucleophile than methanol, it was thought that 6-aminouridine was the more likely product.



Scheme 22 Deprotection of compound **40** using methanolic ammonia.

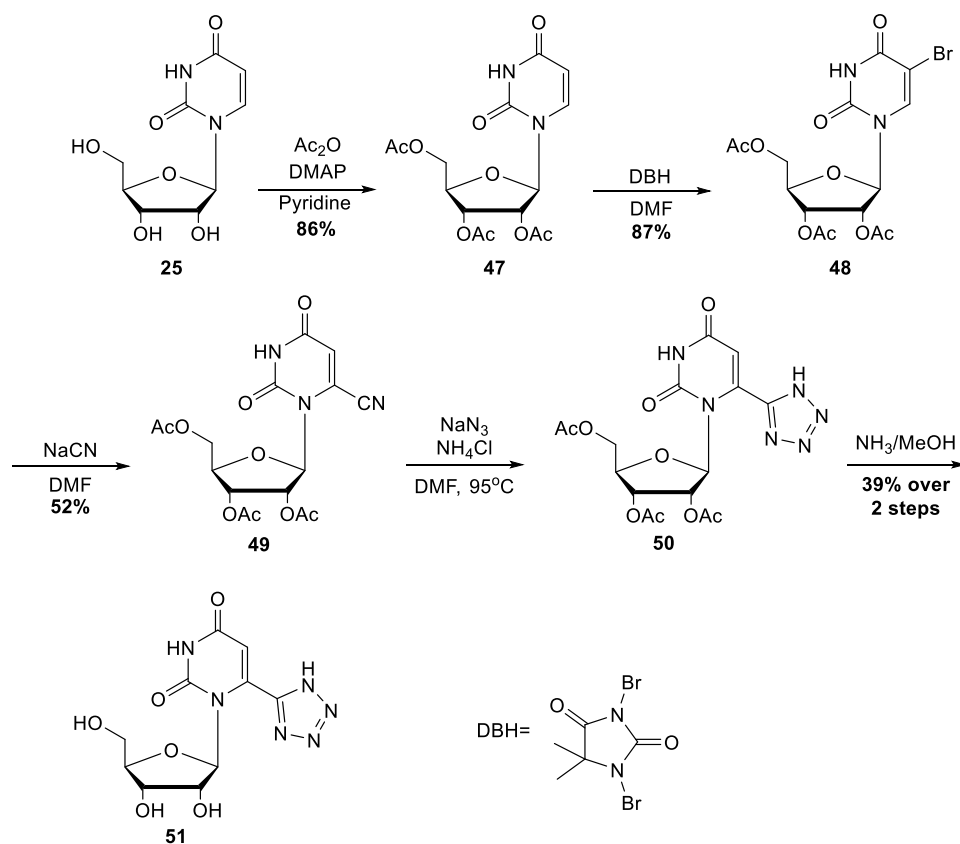
The ¹H NMR of compound **46** lacked any protecting group signals showing that they had been successfully removed, but unfortunately it also lacked any signals for the chloro-propyl triazole moiety. All sugar and base signals were present and a singlet at 3.93 ppm integrating to three protons was indicative of an OMe group. The compound's identity was confirmed by mass spectrometry which showed a peak at 273.07 which corresponds to the [M-H]⁻ ion of 6-methoxy uridine.

The results obtained from the reaction show that the C6-N bond is easily broken through addition of an external nucleophile and loss of the triazole moiety occurs readily. For this reason, in addition to the problems described earlier, it was decided to shift focus away from synthesising triazoles of this type.

3.7 Formation of C-6 substituted tetrazole

Prior to the start of the project, incorporation of a tetrazole moiety at the C-6 position of uridine was identified as being one of the key analogues desired to synthesise. Tetrazoles are excellent bioisosteres of carboxylic acids and so a tetrazole group at the C-6 position was potentially an interesting mimic of OMP. There has been one previous report in the literature of formation of a 6-tetrazole derivative of uridine, but

it was not assessed for its ability to bind to ODCase or in any *in vitro* biological tests and thus it was decided to design a route to 6-(1*H*-tetrazol-5-yl) uridine (Scheme 23).¹⁷¹



Scheme 23 Synthetic route to tetrazole derivative 51.

Formation of tetrazoles quite often proceeds through reaction of azides with cyano compounds. As shown for the triazole click chemistry, the bond between C-6 of uridine and the nitrogen of a heterocycle is susceptible to cleavage *via* an addition-elimination pathway. It was therefore decided to react a nucleosidic nitrile with azide rather than *vice versa* to form the tetrazole moiety as this would lead to a carbon-carbon bond between the nucleoside and the heterocycle which should be more stable.

Formation of protected 6-cyanouridine could be achieved starting from the 6-iodo compound **28**, but a high yielding route *via* protected 5-bromouridine has previously been published and so these conditions were adopted.^{179,185} It was decided to use acetyl protecting groups for the synthesis and so the synthetic route began by protecting the three hydroxyl groups of uridine. The reaction proceeded in 86% yield and confirmation the desired product **47** had been obtained was provided by ¹H

NMR which showed the presence of 3 singlets integrating to three protons each corresponding to the acetyl groups.

Formation of 5-bromo-2',3',5'-tri-*O*-acetyluridine **48** was achieved through reaction of compound **47** with 1,3-dibromo-5,5-dimethylhydantoin.¹⁸⁵ The reaction proceeded in a very good yield of 87%. Successful isolation of the desired product was confirmed by ¹H NMR and mass spectrometry. The mass spectrum of compound **48** showed an isotope pattern that is distinctive for incorporation of bromine, with peaks at 471.00 and 473.00 for the [M+Na]⁺ ion. The ¹H NMR of compound **48** lacked the signal for *H*-5 at 5.80 ppm that was present in the starting material, showing the bromination had occurred at the correct position.

Cyanation was achieved by reacting compound **48** with sodium cyanide in DMF. The procedure followed was originally detailed by Inoue *et al.* and proceeds *via* an interesting addition-elimination type mechanism (Figure 42).¹⁷⁹

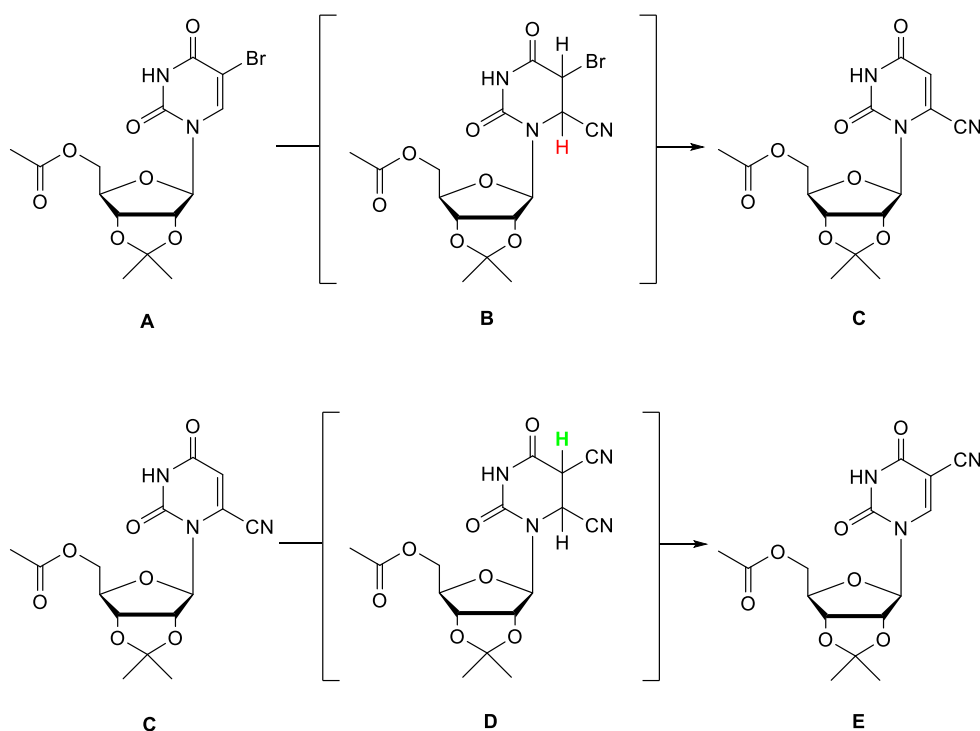


Figure 42 Cyanation reaction originally reported by Inoue *et al.*¹⁷⁹

Inoue *et al.* found that when 1 equivalent of sodium cyanide is added to compound A, intermediate B is formed.¹⁷⁹ The *H*-6 proton shown in red is more acidic than the *H*-5 proton and so is eliminated as HBr to give protected 6-cyano uridine. Additionally, they reported that when a further equivalent of NaCN is added to compound C, intermediate D then forms.¹⁷⁹ Now, the *H*-5 proton shown in green is more acidic so gets eliminated as HCN to give protected 5-cyano uridine. For this

project, only one equivalent of NaCN was used and the desired product was found to form in 52% yield. Confirmation cyanation had been achieved was provided by ^1H NMR and mass spectrometry. The mass spectrum of compound **49** lacked the characteristic isotope peaks of bromine showing that it was no longer present and contained a peak at 418.09 which is consistent with the $[\text{M}+\text{Na}]^+$ ion of the desired product. The ^1H NMR spectrum of compound **49** lacked a signal for *H*-6 at 7.83 ppm that was present in the starting material, but did contain a signal at 6.31 ppm for *H*-5.

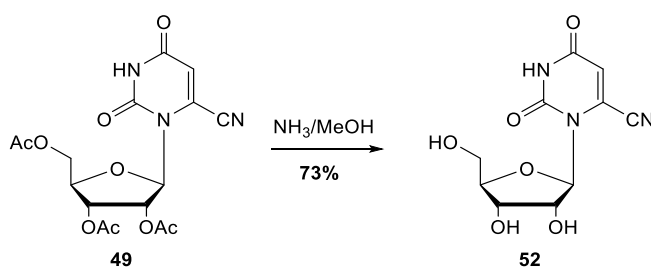
Following the success of this reaction, an additional analogous reaction was carried out, swapping NaCN for KNO_2 and 18-crown-6 to see if nitration of the 6-position could be achieved using this method. Unfortunately however, only starting material **48** was recovered from the reaction.

Formation of the tetrazole moiety was achieved by reacting compound **49** with sodium azide at 95 °C in DMF in the presence of ammonium chloride.¹⁷¹ Mass spectrometry was used to determine the success of the reaction as ^1H NMR was found to be less informative in this case. Compound **50** was found to be quite polar, even with the presence of the three acetyl groups and so NMR analysis was performed using CD_3OD . Under these conditions the tetrazole *NH* proton was not observable and could not be used to determine whether the reaction had worked. The mass spectrum however showed the desired increase in mass corresponding to formation of the tetrazole with a peak at 461.10 for the $[\text{M}+\text{Na}]^+$ ion of the desired product. NMR analysis showed the presence of dimethylamine as an impurity in the product, which forms as a result of breakdown of the reaction solvent DMF at elevated temperatures. Due to the high polarity of the product and its observed preference to go into the aqueous layer during extractions, it was decided to carry the product through without further purification to the deprotection step.

Deprotection of the acetyl protecting groups of compound **50** was achieved using methanolic ammonia. The desired deprotected product was isolated by filtration and purified by trituration with diethyl ether. ^1H NMR analysis of the isolated product lacked the three singlets associated with the acetyl groups indicating they had been successfully removed. The spectrum also indicated successful removal of the dimethylamine impurity and so the yield over the final two steps of the synthesis was calculated as 39%. No displacement of the heterocyclic moiety by methanol was observed, as the mass spectrum showed a mass of 311.07 which is correct for

the $[M-H]^-$ ion of the desired product. A sample of compound **51** was sent for biological testing.

A sample of 6-cyano-2',3',5'-tri-*O*-acetyluridine was also deprotected so that it could be sent for biological testing (Scheme 24). 6-Cyanouridine is a known inhibitor of ODCase, but as its protected form had been successfully synthesised en route to the tetrazole derivative, it was decided to deprotect a sample. It was thought that, like the 6-iodouridine and 6-azidouridine synthesised previously, although it is not a novel compound, sending a sample for biological testing would be good for comparison with other compounds synthesised in the project. Removal of the acetyl protecting groups from compound **49** was again achieved using methanolic ammonia. The deprotection proceeded in 73% yield and confirmation the isolated product was correct was provided by ^1H NMR and mass spectrometry. The ^1H NMR spectrum showed all required sugar and nucleobase signals, but it did not contain the three singlet signals that correspond to the methyl groups of the acetyls. The mass spectrum of compound **52** contained a peak at 268.06 which is correct for the $[M-H]^-$ ion of 6-cyanouridine.



Scheme 24 Deprotection of compound **49** using methanolic ammonia.

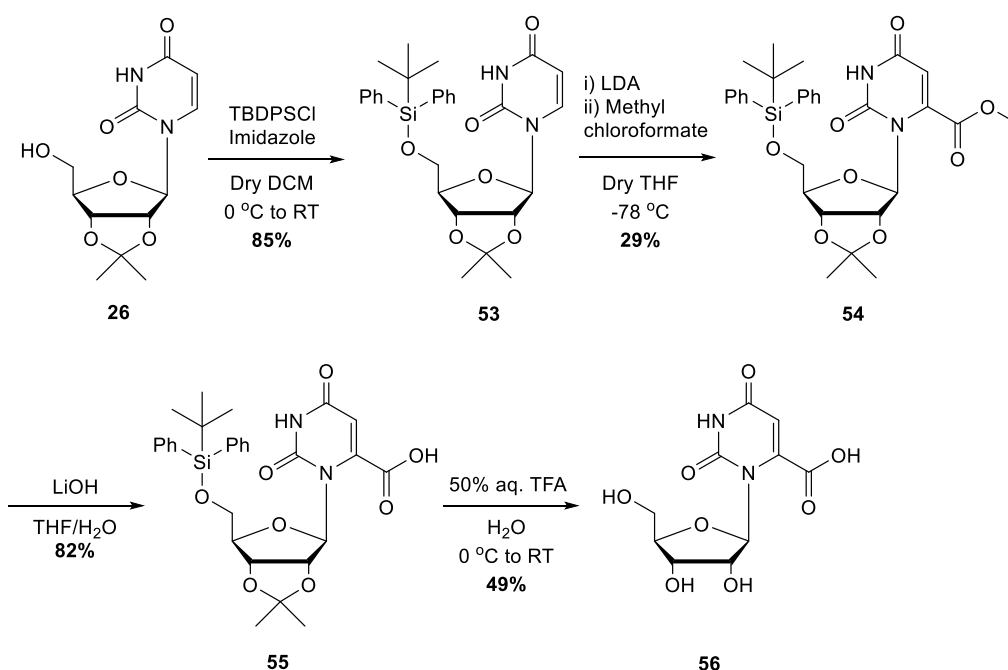
3.8 Alternative electrophiles for use in the double deprotonation reaction

3.8.1 Methyl chloroformate - Synthesis of orotidine

Synthesis of orotidine, the nucleoside derived from the natural substrate of ODCase, was desired so that a sample could be sent for biological testing and the results could be compared to those of other compounds synthesised. A route to the formation of orotidine was devised using a similar method to that used to obtain 6-iodouridine. It was thought that by carrying out an analogous double deprotonation of protected uridine **27** with LDA, then trapping out the anion formed with methyl chloroformate, that following hydrolysis of the ester and deprotection, the desired product would be isolated.

Scheme 25 shows the route followed. The 5'-silyl protecting group was switched from TBDMS to TBDPS for the synthesis as the dimethyl silyl protecting group was found to be unstable to the ester hydrolysis conditions. Although removal of the protecting groups was required in the next step, it was desirable to have all three ring hydroxyl groups protected following hydrolysis of the ester to make purification of the carboxylic acid easier.

The silylation of compound **26** with TBDPSCI was achieved using the same method as had been used to form the TBDMS derivative **27**. The reaction proceeded in a very good yield of 85% and following column chromatography the isolated product was characterised using NMR and mass spectrometry. The ^1H NMR spectrum showed peaks in the aromatic region between 7.36-7.47 and 7.61-7.65 ppm integrating to ten protons in total for the two phenyl rings, indicating incorporation of the TBDPS group had been successful. In addition, mass spectrometry of compound **53** showed a peak at 545.21 which is correct for the $[\text{M}+\text{Na}]^+$ ion of the desired product.



Scheme 25 Synthetic route to orotidine.

Insertion of the methyl ester functionality at the C-6 position was achieved by carrying out a double deprotonation of compound **53** using LDA, then adding methyl chloroformate. The 29% yield obtained for the reaction is quite low, but as previously discussed for the synthesis of the 6-iodo compound **28**, this is presumably due to the need to carry out a double deprotonation. As described for the iodination reaction, a significant portion of starting material can be recovered

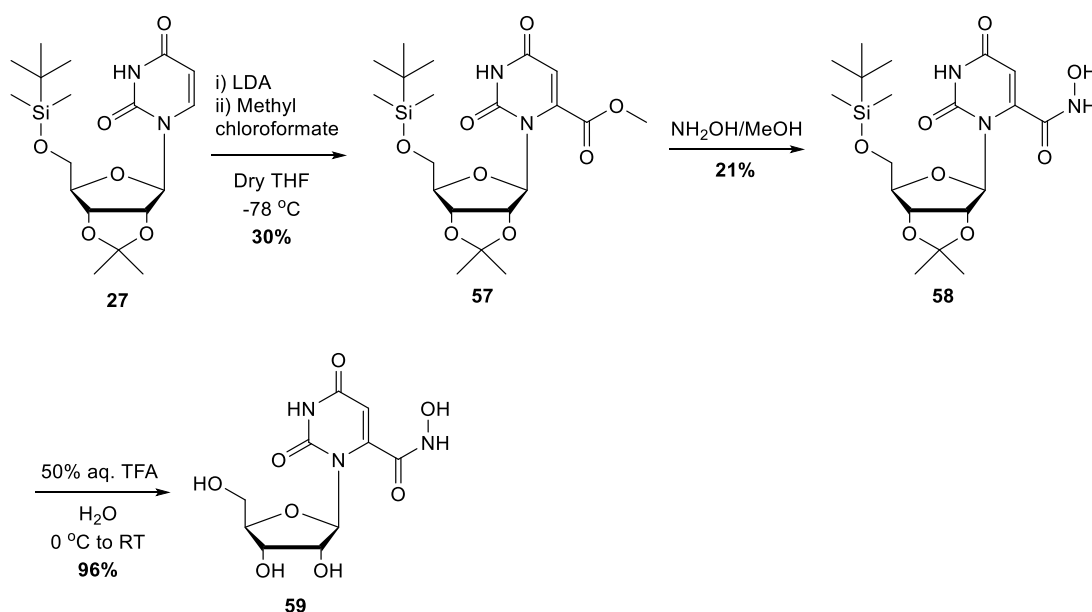
from the reaction for use in subsequent reactions. Confirmation compound **54** had been successfully synthesised was obtained from ^1H NMR which showed the presence of a singlet at 3.81 ppm integrating to three protons corresponding to the OMe group. It also lacked the signal for *H*-6 at 7.57 ppm found in the starting material showing the reaction had taken place at the desired position. Mass spectrometry further confirmed the success of the reaction as a peak at 603.21 was obtained which is correct for the $[\text{M}+\text{Na}]^+$ ion of the desired product.

Hydrolysis of the methyl ester was achieved using lithium hydroxide in a mixture of THF and water as solvent. These conditions were found to result in cleavage of a 5'-TBDMS group but the diphenyl silyl equivalent was quite stable which allowed for successful purification of compound **55**. A ^1H NMR of compound **55** lacked a singlet for the methoxy group showing the hydrolysis had been successful and mass spectrometry showed the desired decrease in mass on going from the methyl ester to the carboxylic acid.

Full deprotection of all protecting groups to obtain the final desired compound **56** was carried out under acidic conditions, using a 50% aqueous TFA solution. The ^1H NMR of compound **56** lacked both the isopropylidene and TBDPS signals present in the starting material and a mass spectrum showed a peak at 287.05 which corresponds to the $[\text{M}-\text{H}]^-$ ion of orotidine. A sample of the final unprotected material was sent for biological testing.

3.8.2 Methyl chloroformate - Formation of a hydroxamic acid

Hydroxamic acids are mildly acidic ($\text{p}K_{\text{a}} \sim 8-9$) and are known to be strong metal chelators.¹⁷⁰ Due to their similarity in structure, they are often considered as isosteres of carboxylic acids and have been used successfully as such in previous studies.^{170,186} They are most commonly formed through reaction of esters or activated esters with hydroxylamine salts. As formation of a C-6 methyl ester derivative of uridine had been obtained whilst synthesising orotidine, it was decided to utilise this compound further by forming a hydroxamic acid that could act as an inhibitor of ODCase. Scheme 26 shows the synthetic route followed to obtain a C-6 substituted hydroxamic acid.



Scheme 26 Synthetic route to C-6 substituted hydroxamic acid derivative.

Starting from compound **27**, again a double deprotonation reaction was carried out using methyl chloroformate as the electrophile. The 5'-silyl group used was the TBDMS group as it was found to be stable to the subsequent reaction conditions. The reaction proceeded in a very similar yield (30%) to that done using the 5'-TBDPS protected compound **53** (29%) during the synthesis of orotidine. Formation of the C-6 methyl ester derivative **57** was confirmed by ¹H NMR which showed a singlet at 3.94 ppm for the OMe group.

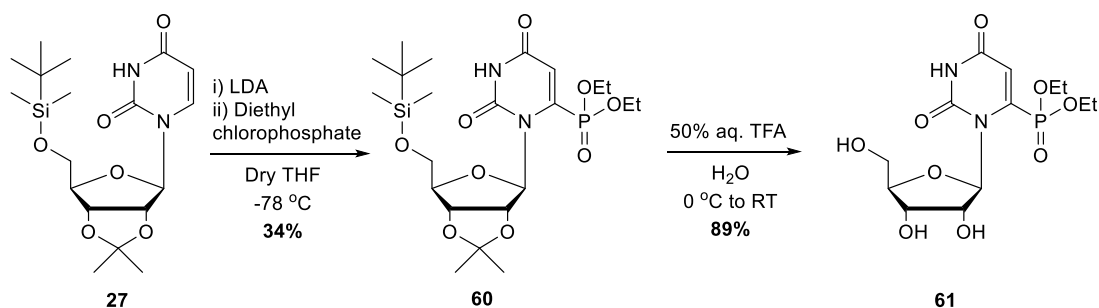
Formation of the hydroxamic acid was achieved by reacting compound **57** with a 1.5 M solution of hydroxylamine in methanol, prepared according to a patent submitted by Mathew and Ulmer.¹⁸⁷ The reaction was found to be quite low yielding (21%), but several previous attempts to form the hydroxamic acid *via* other methods had been unsuccessful. Confirmation the reaction had been successful was provided by mass spectrometry. Whilst the ¹H NMR of the product obtained lacked the methyl singlet due to the ester of the starting material, it was not possible to determine whether the reaction had formed the desired hydroxamic acid or had hydrolysed to the carboxylic acid. However, the mass spectrum of compound **58** showed a peak at 480.18 which is correct for the [M+Na]⁺ ion of the hydroxamic acid.

Deprotection of the isopropylidene and TBDMS groups was achieved using a 50% aqueous TFA solution in almost quantitative yield (96%). Removal of the protecting groups was confirmed by ¹H NMR which lacked the two singlets of the isopropylidene group present in the starting material at 1.33 and 1.50 ppm and the

TBDMS signals at 0.07 and 0.90 ppm. A mass spectrum of compound **59** contained a peak at 302.06 which is correct for the $[M-H]^-$ ion of the desired hydroxamic acid. A sample of compound **59** was submitted for biological testing.

3.8.3 Diethyl chlorophosphate – Formation of a C-6 phosphonate

Although problems have arisen carrying out a double deprotonation on protected uridine in order to insert different functionalities at the 6 position, as discussed, a procedure that can reliably produce ~30-40% yields has now been obtained that also allows for recovery of unreacted starting material. The reaction was therefore utilised further to produce a C-6 substituted phosphonate compound. It was decided to investigate whether the presence of a P=O bond at C-6 was able to mimic the C=O bond of OMP. Scheme 27 shows the synthetic route followed.



Scheme 27 Synthetic route C-6 substituted diethyl phosphonate **61**.

The first step was to carry out the double deprotonation of compound **27** using LDA. Diethyl chlorophosphate was then added as the electrophile and compound **60** was found to form in 34% yield. Diethyl chlorophosphate was used as it has been shown previously to react with C-6 of uridine, but the compound formed has not been tested for its ability to bind to ODCase or as an antimalarial.¹⁸⁸ Confirmation compound **60** had been correctly isolated was provided by NMR. The ¹H NMR lacked the signal for H-6 at 7.70 ppm present in the starting material but showed signals at 1.31 and 4.11-4.23 ppm corresponding to the presence of two ethyl groups. The ¹³C NMR showed splitting of the C-6 carbon into a doublet due to its close proximity to the phosphorus atom. A ³¹P NMR showed only one peak at 5.61 ppm indicating there was only one phosphorus environment and this shift was in the correct region for a phosphonate of this type.

Again, the isopropylidene and TBDMS protecting groups were removed using a 50% aqueous TFA solution to yield the diethyl phosphonate product **61** in 89% yield. A ¹H NMR spectrum of compound **61** contained no protecting group signals

and a peak at 379.09 in its mass spectrum for its $[M-H]^-$ ion confirmed it had been successfully deprotected. A sample of this compound was submitted for biological testing.

Due to time constraints, no further work was carried out on compound **61**, but in theory it is possible to remove the diethyl functionality of the phosphonate using sodium iodide.¹⁸⁸ Future work would investigate synthesis of this compound so it could be tested for its antimalarial properties *in vitro*.

3.9 Palladium chemistry

3.9.1 Route to a C-6 substituted propargylic acid

Formation of potential inhibitors of ODCase that can mimic the transition state of the decarboxylation process were identified as important targets. It was thought that insertion of a carbon-carbon triple bond at the C-6 position would help to mimic the elongation of the bond that occurs as CO_2 is lost from this position in the natural substrate OMP. The initial synthetic target was that shown in Figure 43, due to its similarity in structure to orotidine.

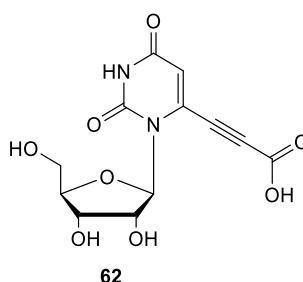


Figure 43 Target in search of new ODCase inhibitor.

It was decided to utilise the Sonogashira reaction to form the desired compound **62**.¹⁸⁹ The Sonogashira reaction is used to form carbon-carbon bonds and due to its mild operating conditions, it was deemed to be suitable for use in this project. The reaction involves coupling of a terminal alkyne with either an aryl or vinyl halide under basic conditions in the presence of both a palladium catalyst and a copper catalyst (Figure 44).

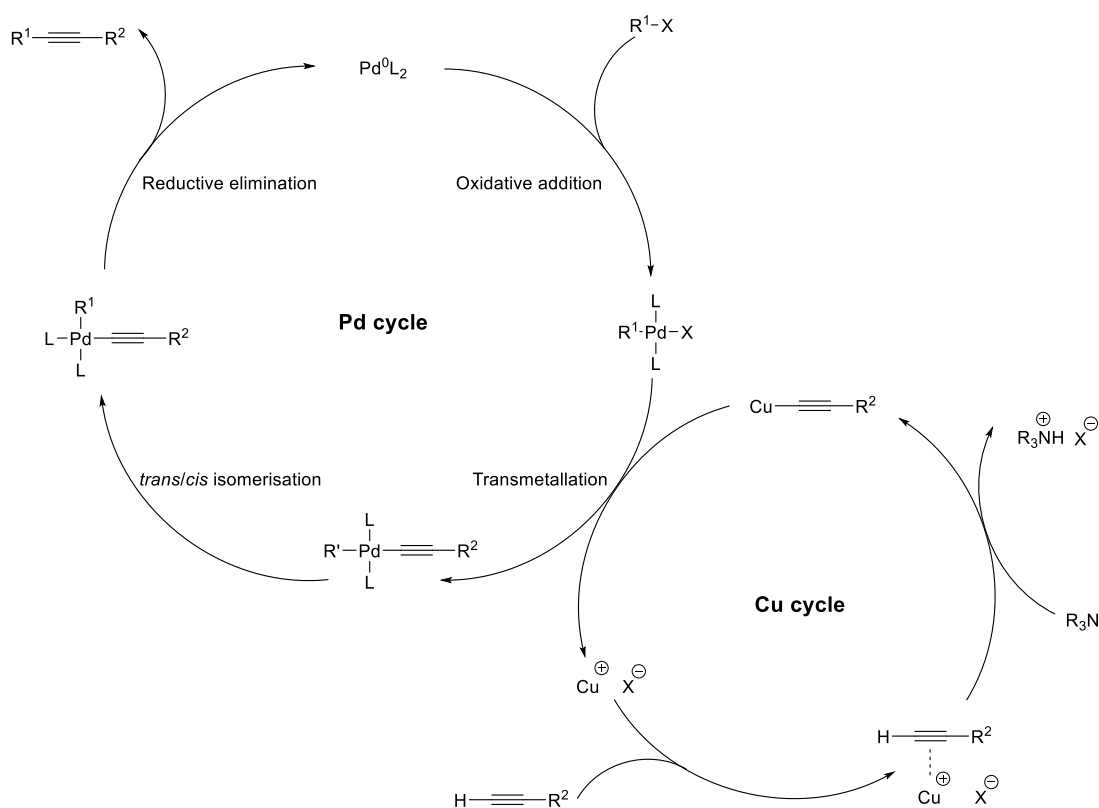
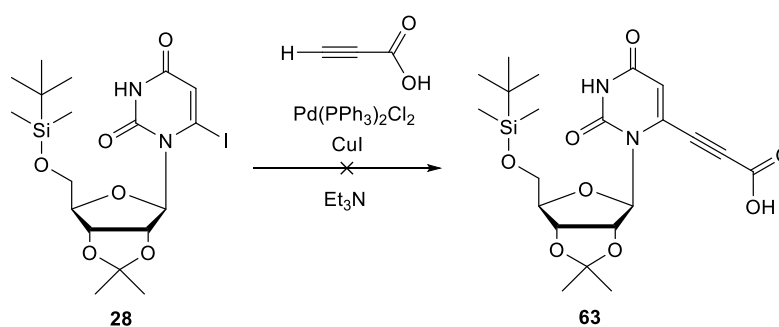


Figure 44 Sonogashira reaction mechanism.¹⁹⁰

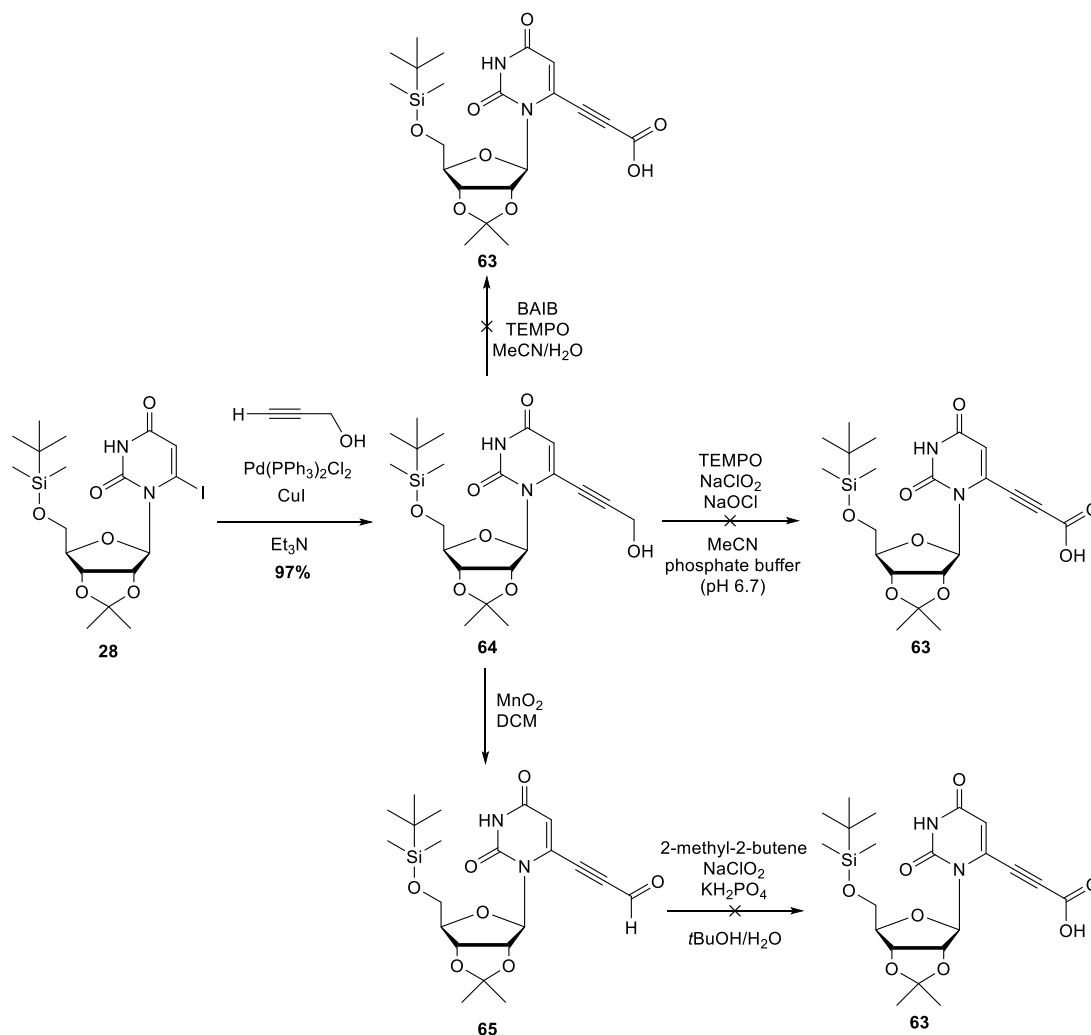
The protected iodo compound **28**, previously identified as a key intermediate in the project, was thought to be an ideal substrate for the reaction as it is a vinyl halide and is stable to basic conditions. A Sonogashira reaction between compound **28** and propiolic acid was attempted as a successful reaction between the two would directly give the desired analogue.



Scheme 28 Unsuccessful Sonogashira reaction between compound **28** and propiolic acid.

Unfortunately, the reaction was unsuccessful. Although TLC showed the formation of new components, none could be isolated. It was speculated that decarboxylation could have occurred.

Nevertheless, alternative routes to compound **62** were attempted. A search of the literature showed a Sonogashira reaction between a 6-iodo uridine derivative and propargyl alcohol was feasible.¹⁹¹ It was therefore decided to attempt to form the C-6 substituted propargyl alcohol and then oxidise it to the desired carboxylic acid (Scheme 29).



Scheme 29 Formation of C-6 substituted propargyl alcohol and subsequent oxidation attempts.

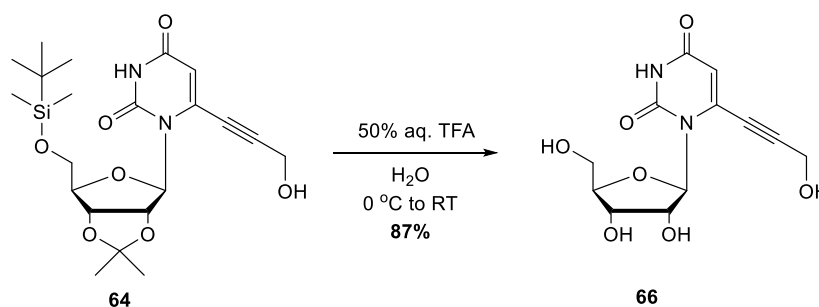
The Sonogashira reaction between compound **28** and propargyl alcohol was carried out in the presence of $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ and CuI using triethylamine as solvent and was found to proceed in almost quantitative yield (97%). Confirmation the reaction had been successful was provided by ^1H NMR and mass spectrometry. A ^1H NMR of the isolated compound showed the presence of a CH_2 group at 4.52 ppm appearing as a doublet as well as a signal for the OH group at 3.20 ppm that had been split into a triplet due to the coupling between the two groups. A mass spectrum of the

compound contained a peak at 475.19 which is consistent with the mass that should be obtained for the $[M+Na]^+$ ion of the desired compound.

Having formed the nucleoside propargylic alcohol **64**, the next step was to investigate ways of oxidising it to the corresponding carboxylic acid. It was decided to first attempt a stepwise oxidation in which the alcohol would be converted to the aldehyde and then oxidised further to the carboxylic acid. For this task, the oxidising agent manganese dioxide was chosen as it is known to be fairly mild and is selective for allylic, benzylic and propargylic alcohols. Reaction of compound **64** with MnO_2 was found to be quite slow, but TLC monitoring of the reaction showed the formation of a new component and eventual consumption of all starting material. Isolation of the new component was achieved using column chromatography and although a preliminary 1H and ^{13}C NMR spectrum seemed to indicate formation of the desired aldehyde, it was found to be quite unstable. Despite showing signs of breaking down, the crude aldehyde material obtained was subjected to Pinnick-Lindgren oxidation conditions in the hope that some carboxylic acid derivative could be isolated. The Pinnick-Lindgren oxidation is a mild reaction used for converting aldehydes to carboxylic acids. Unfortunately however, no desired product could be isolated from the reaction.

Due to the apparent instability of the propargylic aldehyde, it was decided to attempt formation of the carboxylic acid directly from the alcohol in a one-pot method. Two attempts were made to carry out this conversion; one involving oxidation by the hypervalent iodine species bis(acetoxy)iodobenzene (BAIB) and one using a modified Pinnick oxidation with TEMPO and bleach.^{192,193} However, neither method was found to give the desired carboxy compound **63**. The most promising of the one-pot methods was the BAIB reaction which consumed all of the starting material, but attempted isolation of new compounds formed in the reaction was unsuccessful. A 1H NMR of the crude reaction material showed it contained a nucleosidic component, but it appeared loss of the TBDMS protecting group was occurring which made purification difficult and isolation unachievable.

Whilst it was not possible to obtain the C-6 substituted propargylic acid, it was decided that as the propargylic alcohol **64** had been synthesised, a sample should be deprotected for biological testing (Scheme 30). Although C-6 substituted propargyl alcohol uridine has been synthesised previously, it has not been evaluated as an antimalarial or as an inhibitor of ODCase.



Scheme 30 Deprotection of compound 64 under acidic conditions.

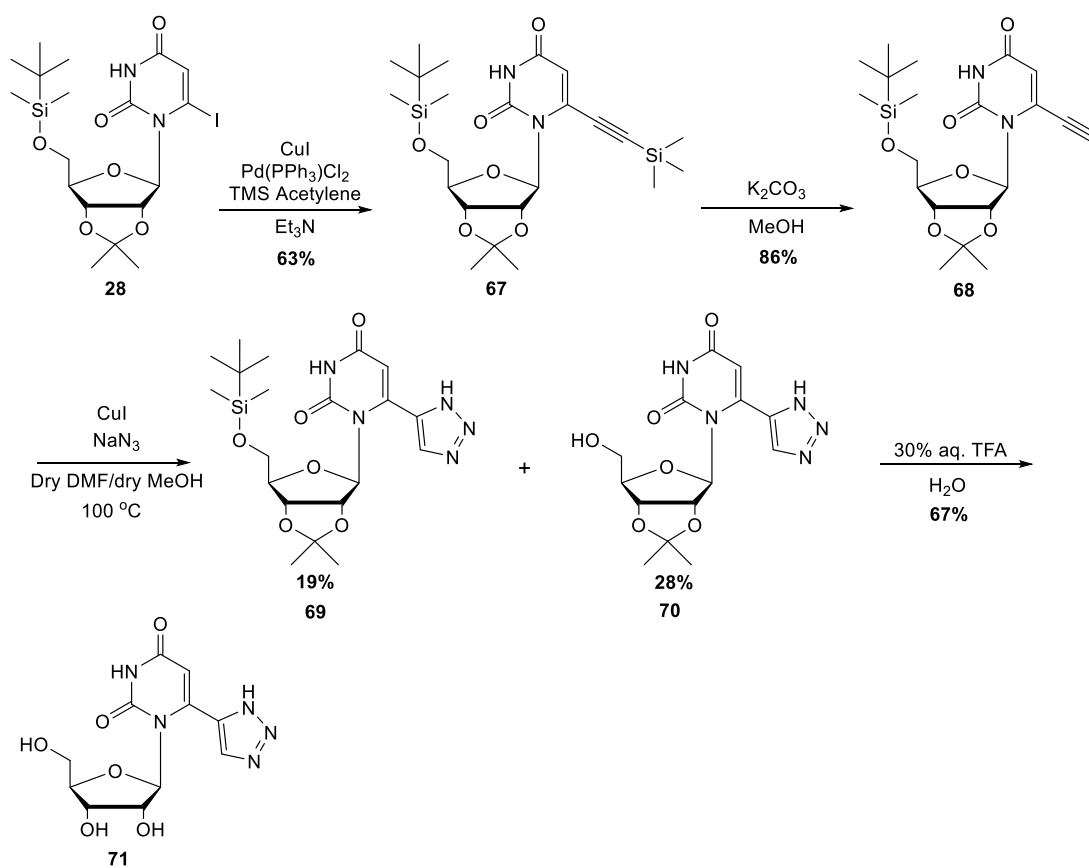
Removal of the protecting groups was carried out using a 50% aqueous TFA solution and proceeded in a good yield of 87%. A ^1H NMR of the isolated compound lacked the protecting group signals present in the starting material and a mass spectrum showed a peak at 297.07 which is correct for the $[\text{M}-\text{H}]^-$ ion of the desired deprotected nucleoside **66**.

3.9.2 Formation of a carbon linked triazole

As discussed earlier in the section on click chemistry, formation of a nitrogen linked triazole to the C-6 position encountered several problems. One of the main issues was deprotection of the compounds. Under acidic conditions, the compounds were found to decompose and under basic conditions, an unusual addition-elimination reaction of the triazole moiety was found to occur. As demonstrated for the C-6 substituted tetrazole described earlier, the formation of a carbon-carbon bond between C-6 and a heterocycle is much more stable than a carbon-nitrogen one. It was decided therefore to investigate whether a carbon linked triazole could be formed at the C-6 position *via* reaction of azide with a nucleosidic alkyne.

Following on from the successful Sonogashira reaction between compound **28** and propargyl alcohol, the same reaction conditions were used to react compound **28** with TMS acetylene. This reaction was required to insert the desired carbon-carbon triple bond at C-6 and as a TMS group can be selectively cleaved, it is an ideal route to forming the C-6 substituted alkyne required for forming a carbon linked triazole.

The Sonogashira reaction between compound **28** and TMS acetylene proceeded in a good yield of 63%. Its success was confirmed by ^1H NMR which contained a singlet at 0.89 ppm integrating to nine protons corresponding to the three methyls of the TMS group. In addition, the ^{13}C NMR contained signals for the alkyne carbons and a mass spectrum showed a peak at 517.22 for the $[\text{M}+\text{Na}]^+$ ion of the desired compound.



Scheme 31 Formation of a carbon linked triazole at C-6.

Selective removal of the TMS group in the presence of the TBDMS group was achieved using potassium carbonate in methanol. The reaction proceeded in 86% yield and the isolated compound was characterised by NMR and mass spectrometry. A ^1H NMR lacked the singlet due to the three methyl groups of TMS present in the starting material, but contained a signal for the terminal alkyne proton that appeared at the same chemical shift as the C-5' protons (3.76-3.84 ppm). An HSQC confirmed that the terminal alkyne proton correlated to a carbon at 90.91 ppm which corresponds to the terminal alkyne carbon. A mass spectrum of the isolated compound showed the desired reduction in mass for loss of the TMS group with a peak appearing at 445.18 for the $[\text{M}+\text{Na}]^+$ ion.

With a terminal alkyne now in place at C-6, a [3+2] cycloaddition between compound **68** and an appropriate azide could take place. Yamamoto *et al.* reported a procedure that utilises a copper catalyst to perform the addition of TMS azide to an alkyne to give a 1,2,3-triazole.¹⁹⁴ The TMS functionality is lost during the reaction so the 1,2,3-triazole produced is unsubstituted. Figure 45 shows the authors

proposed mechanism for the reaction.¹⁹⁴ Reaction of compound **68** with TMS azide in the presence of copper iodide at 100 °C produced the desired triazole compound in 19% yield. The low yield obtained for the desired product is due in part to the fact that some loss of TBDMS occurred under the reaction conditions, so as well as isolating the fully protected nucleoside triazole **69**, a proportion of 6-(1*H*-1,2,3-triazol-5-yl)-2',3'-isopropylidene uridine **70** was also isolated. Confirmation the triazole had formed was provided by ¹H NMR and mass spectrometry. A ¹H NMR spectrum showed the terminal alkyne proton was no longer present, but a signal at 8.21 ppm, integrating to one, corresponding to the triazole proton, was present. A ¹³C spectrum lacked the signals for the two carbons of the triazole which called into question whether the desired product had been made, but a mass spectrum showed a peak at 488.19 ppm which is correct for the [M+Na]⁺ ion of the desired triazole. Due to the small amount of compound **69** isolated, it is possible that the NMR sample was just not strong enough for these signals to appear in the carbon spectrum, especially given that one of them is a quaternary carbon.

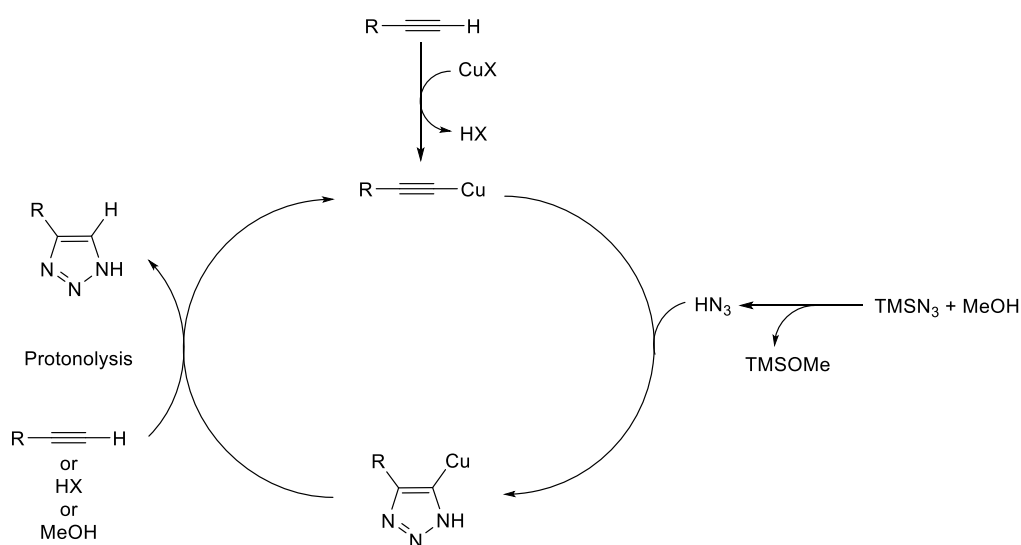


Figure 45 Mechanism of copper catalysed [3+2] cycloaddition proposed by Yamamoto et al. for reaction between a terminal alkyne and TMS azide.¹⁹⁴

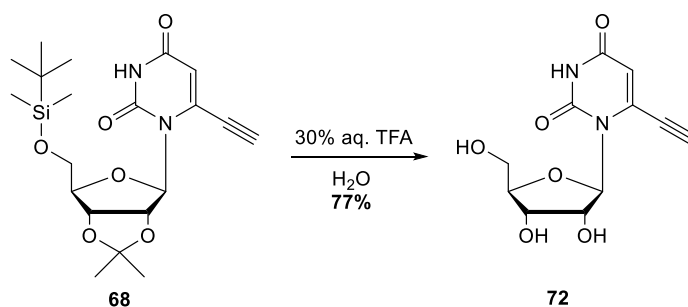
For the deprotection reaction, the fully protected triazole compound **69** and compound **70** lacking the TBDMS group were combined. A 30 % aqueous TFA solution was used to achieve full deprotection to compound **71**. A ¹H NMR lacked any protecting group signals and a peak at 310.08 in the mass spectrum for the [M-H]⁻ ion confirmed the desired carbon linked triazole had been formed. A sample was submitted for biological testing.

The success of this 1,2,3-triazole forming reaction was pleasing, as prior to the start of the project, insertion of a triazole moiety at the C-6 position was identified as a

promising target. Due to the difficulties described in forming a nitrogen linked triazole starting from a nucleoside azide, it was thought isolation of a C-6 substituted triazole may not be achievable. But, the formation of the carbon linked triazole shows promise that further compounds of this type could be synthesised, depending on how well the molecular modelling studies predicts it to bind and how active the *in vitro* biological testing finds it.

3.9.3 Isolation of 6-ethynyl uridine

Having successfully isolated protected 6-ethynyl uridine **68** during the synthesis of the carbon linked triazole, it was decided to fully deprotect a sample so it too could be sent for biological testing. Although 6-ethynyl uridine has been prepared previously, it has not been tested for its antimalarial properties.¹⁹¹ Scheme 32 shows the deprotection of compound **68**.



Scheme 32 Deprotection of compound **68** to yield 6-ethynyl uridine.

A 30% aqueous TFA solution was used to remove the TBDMS and isopropylidene groups. Isolation of the desired compound was achieved using column chromatography. Initial spectroscopic analysis indicated the reaction had been a success as a ¹H NMR recorded in CD₃OD contained all of the desired signals, although the terminal acetylene proton was observed at a higher chemical shift (4.65 ppm) than where it had been observed for the protected analogue **68** in CDCl₃ (3.76-3.84 ppm). Compound **72** was found to be quite insoluble in methanol and so in order to record a ¹³C spectrum the solvent was changed to D₂O. However, upon changing NMR solvents, signals corresponding to the acetylene moiety could not be observed in either ¹H or ¹³C spectra. There was no signal in the ¹H NMR for *H*-6 indicating that a substituent was still present at this position, but the terminal alkyne proton was not observed and the two acetylenic carbons in the ¹³C spectrum could also not be detected. A second ¹H NMR spectrum in CD₃OD was recorded following exposure to D₂O and the signal thought to correspond to the terminal alkyne was no longer present. However, a mass spectrum of the isolated compound showed a

peak at 267.06 which is correct for the $[M-H]^-$ ion of 6-ethynyl uridine, indicating the product was correct. It is thought that NMR signals for the acetylene moiety are not observed possibly due to some form of exchange reaction with solvent D_2O . Due to lack of time and material, it was not possible to record any spectra in $DMSO-d_6$. The acetylenic signals are observable in the protected nucleoside, when $CDCl_3$ is used to record the NMR spectrum, but they appear to disappear following deprotection despite 1H NMR indicating there is a substituent at C-6 and mass spectrometry showing the correct mass for 6-ethynyl uridine. Although the acetylene moiety could not be observed in the NMR data, it is believed 6-ethynyl uridine has been successfully isolated as the accurate mass data obtained for compound **72** is correct for 6-ethynyl uridine and under the deprotection conditions used, there is no plausible mechanism for loss of the acetylene moiety.

3.9.4 Attempts to form a C-6 substituted boronic acid

Whilst carrying out some palladium chemistry, it was decided to investigate whether a C-6 substituted boronic acid could be synthesised (Figure 46).

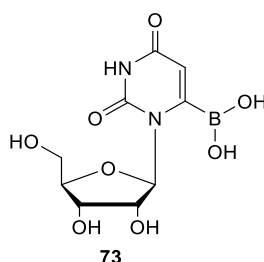
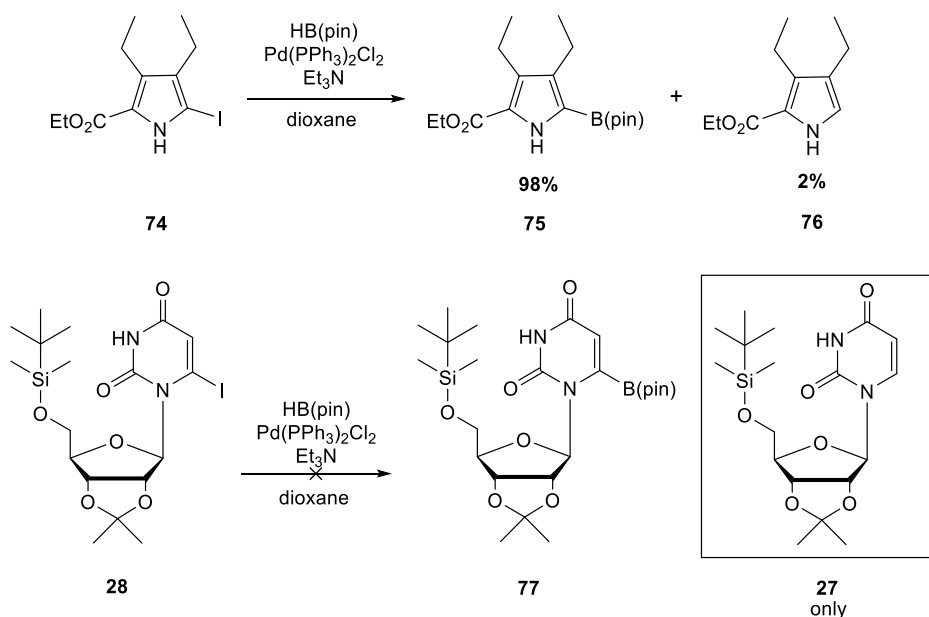


Figure 46 C-6 substituted boronic acid.

Boronic acids are known to be good bioisosteres of carboxylic acids and can be formed through palladium catalysed coupling of an aryl halide or triflate with a diboronyl ester such as bis(pinacolato)diboron.¹⁷⁰ Hydrolysis of the boronate ester formed yields the boronic acid.

Several attempts were made to synthesise compound **73** using palladium cross-coupling reactions starting from the protected iodo intermediate **28**, but unfortunately, none were found to yield the desired boronic acid **73**. The most promising of the reactions was a Pd catalysed borylation reaction using conditions reported by Yoshida *et al.* (Scheme 33).¹⁹⁵

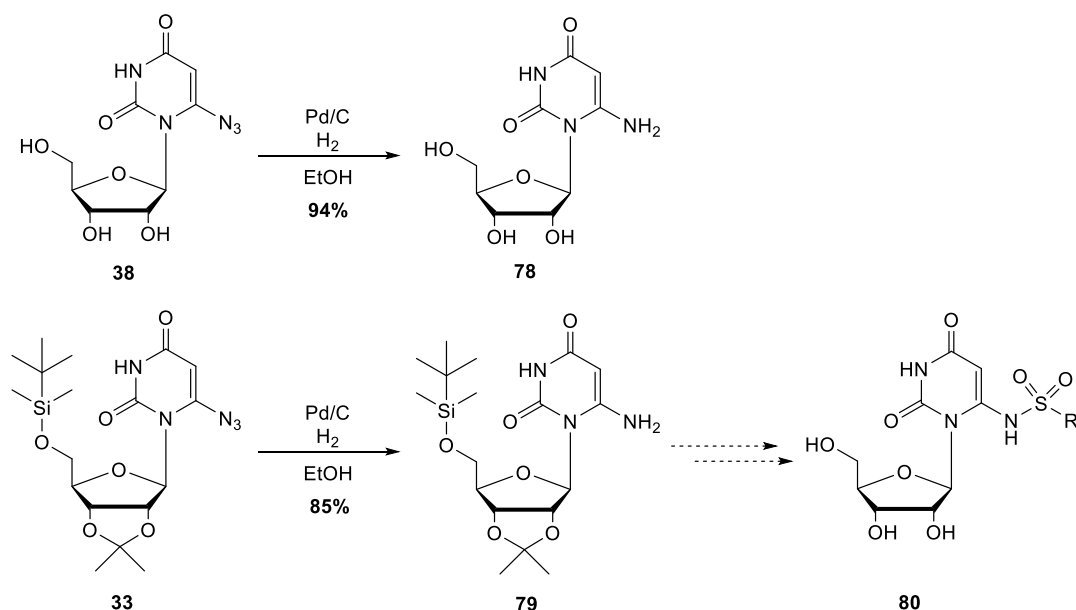


Scheme 33 Borylation conditions reported by Yoshida *et al.*¹⁹⁵

When the iodo compound **28** was subjected to these reaction conditions, the only isolatable product obtained was the protected uridine compound **27**. This implies that the desired borylation may have taken place, but the boron ester formed then protodeboronated. Although Yoshida *et al.* reported only minor formation of the dehalogenated pyrrole **76**, this was not the case using compound **28** as the substrate.¹⁹⁵ An attempt to make the C-6 substituted boronic acid **73** was also attempted in the absence of Pd by carrying out a deprotonation of compound **27** and adding $(i\text{PrO})_3\text{B}$, but again this was found to be unsuccessful. 2-Pyridyl boronic acids and esters are known to be highly susceptible to decomposition *via* protodeboronation.¹⁹⁶ A C-6 substituted uridine boronic acid is very similar in structure to 2-pyridyl boronic acid and so it is highly likely that it would also be prone to undergo protodeboronation readily. For this reason, this class of analogues was not pursued further in this project.

3.10 Synthesis of 6-aminouridine

The final analogue synthesised was 6-aminouridine. This compound has been isolated previously and tested as an inhibitor of ODCase and as an antimalarial.¹⁶⁶ It has been found to be a moderate inhibitor of ODCase from both *Methanobacterium thermoautotrophicum* and *Plasmodium falciparum*, but to show very little antimalarial activity.¹⁶⁶ It was synthesised in this project to act as a comparison for other analogues in the biological testing and as it was identified as an intermediate for the synthesis of further analogues.



Scheme 34 Synthesis of 6-aminouridine (top) Potential route to other analogues via 6-amino derivative **79** (bottom).

To obtain a sample of 6-aminouridine for biological testing, 6-azidouridine was hydrogenated over a palladium on carbon catalyst. The resulting 6-aminouridine was characterised by NMR and confirmation it had been successfully formed was provided by mass spectrometry which showed a peak at 282.07 which corresponds to its $[M+Na]^+$ ion.

Formation of a sulphonamide at the C-6 position was identified as another potential inhibitor of ODCase. Sulphonamides are very good isosteres of carboxylic acids so it would be interesting to investigate how one interacts with ODCase.¹⁷⁰ There was not time in this project to fully investigate the synthesis of a sulphonamide uridine derivative, but it is a compound that is of high priority for future work. Preliminary work has shown that the hydrogenation of fully protected azido compound **33** gives the protected amine **79** in good yield and so future work will look to continue on the synthesis from this point.

3.11 Conclusions and future work

The work carried out in this project sought to generate a series of analogues of the compound OMP that could act as inhibitors of the enzyme ODCase and potentially display antimalarial properties. In total, eleven compounds were generated and they have been sent for biological testing in a *Plasmodium falciparum* 3D7 assay.

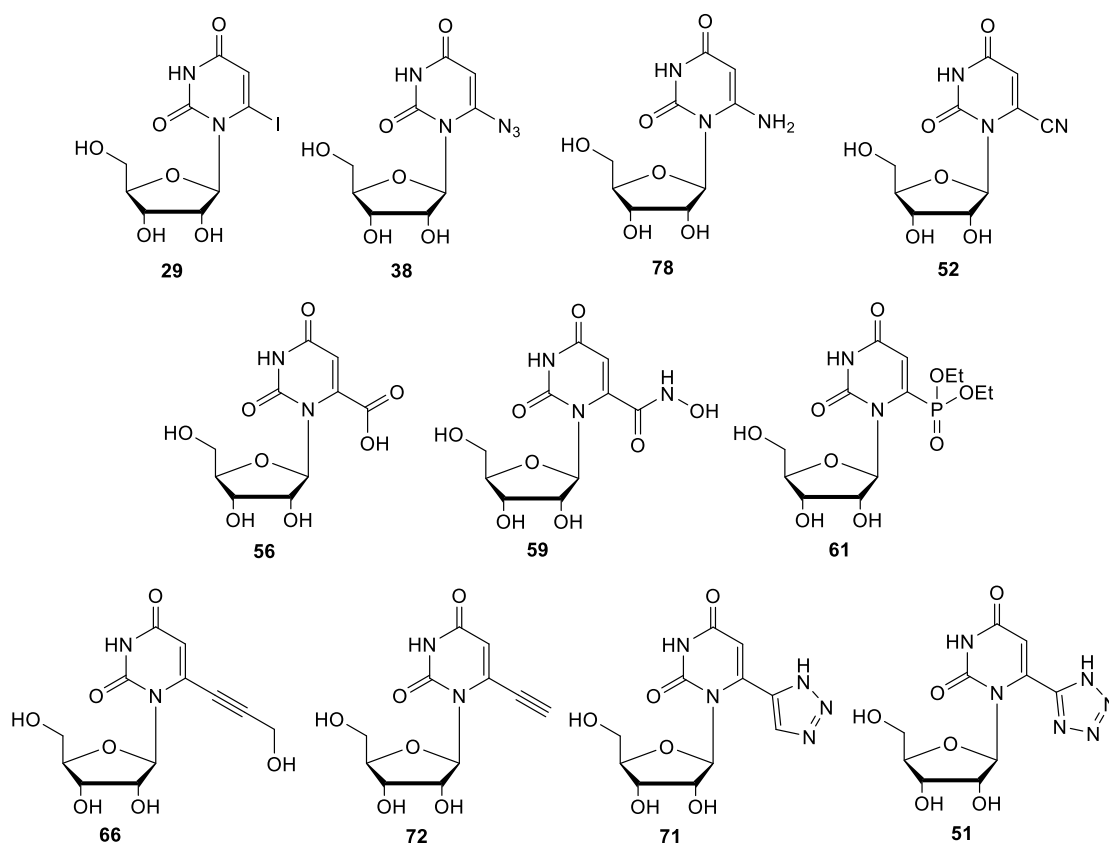


Figure 47 Final compounds synthesised.

Of the compounds synthesised, four (**29**, **38**, **78**, **52**) have been tested previously as ODCase inhibitors and for their antimalarial properties.¹⁶⁶ They were synthesised again presently so that the results obtained from the biological testing could be compared to those obtained previously and as a comparison for the other analogues synthesised. The nucleoside form of ODCase's natural substrate **56** was also generated during the course of the project so its results could be compared to those of the other analogues.

Each of the compounds synthesised was also computationally docked into the crystal structure of ODCase to see how well they might fit into the active site. More details on the computational work can be found in the following chapter.

As discussed, several problems have been encountered during this work. The initial low yields obtained for the synthesis of the key iodo intermediate **28** were overcome through optimisation of the reaction conditions and it can now be synthesised in sufficiently large quantities so that it does not act as a bottleneck to further reactions. Although not all analogues identified prior to the start of the project as being promising inhibitors were able to be synthesised, such as 6-nitrouridine,

eleven analogues were able to be isolated. Of these eleven, the carbon linked triazole **71** and the hydroxamic acid **59** are completely novel.

Future work would look to build on the synthetic work that has already taken place. As mentioned, formation of a C-6 substituted sulphonamide is of high interest as is formation of a C-6 substituted phosphonic acid through further reaction of compound **61**. The results obtained from the *in vitro* testing will also largely influence what compounds are investigated next.

Chapter 4

Computational docking and biological testing

Chapter 4 Results and discussion 3

Molecular modelling was performed in order to try and predict how well the eleven final compounds synthesised might interact with ODCase. The project used three computational programs to model, dock and visualise the ligands and ODCase crystal structure. The modelling program Spartan '16 was used to construct all compounds that were to be docked into ODCase.¹⁹⁷ To carry out the docking, the program GOLD was used which stands for Genetic Optimisation for Ligand Docking.¹⁹⁸ It is designed to take into account the flexibility of ligands when scoring them. GOLD contains four different scoring functions which allows the user to compare the performance of the ligands scored using different criteria. For this work, the scoring function ChemPLP was chosen to assess the binding of the ligands.¹⁹⁸ This scoring function is now the default scoring function used in GOLD as validation tests have found that overall it is more effective than the other scoring functions for predicting the poses of ligands effectively.^{198,199} The program PyMOL was used to visualise the crystal structure of ODCase and the docked ligands.²⁰⁰

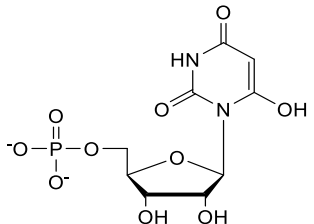
4.1 Validation docking of BMP

Before docking of the eleven final compounds into an ODCase crystal structure could be carried out, a validation of the docking protocol was required. As mentioned in the introduction, there are now more than 200 ODCase crystal structures deposited in the Protein Data Bank, over half of which are derived from *Methanothermobacter thermautotrophicus* (*Mt*) and one of these entries, PDB code: 1X1Z, was selected for use in this study.²⁰¹ This crystal structure is of ODCase in complex with 6-hydroxyuridine-5'-monophosphate (BMP), the most potent inhibitor of ODCase currently known.¹⁴⁸ The reasons for choosing this crystal structure were:

- 1) It was thought if GOLD could accurately match the binding pose of the best known inhibitor of ODCase, then it would act as a benchmark for evaluating how well the compounds synthesised in this study might bind.
- 2) ODCase has four key active site residues, two Lys and two Asp, known to be important for binding. These residues are conserved across species and so any compounds predicted to bind well to *Mt* should bind well to other species.^{131,164}
- 3) Previously it has been found that the K_i values obtained for inhibitors against ODCase from *Mt* are comparable to their K_i values against ODCase from other species.¹⁶¹

The first stage of the validation involved downloading the 1X1Z PDB file and treating it according to Protocol 1 (see Experimental section).²⁰¹ ODCase is a dimer and often in molecular modelling, one monomer of a dimer crystal structure is deleted prior to beginning docking studies, especially if the dimer is symmetrical. However, in this case, it was decided to leave ODCase as a dimer for the docking study as previous investigations have found that both monomers contribute key residues to the active site.¹¹⁸ The structure of BMP was then modelled in Spartan '16 following Protocol 2.¹⁹⁷ GOLD was used to extract the crystallographic BMP molecule and then the BMP molecule modelled in Spartan '16 was re-docked as described in Protocol 3, generating 25 solutions.^{197,198} This was to test whether GOLD was able to match the binding pose of the crystallographic data. In order to evaluate how well GOLD carried out this task, the root mean square deviation (RMSD) values generated from the docking were considered. RMSD is a measure of the distance between the pose of a ligand predicted by the docking program and the pose of the ligand in the crystallographic structure, so, the lower this value is, the better. Ideally, RMSD values should be less than 2 Å.²⁰² Table 6 summarises the data obtained from the validation docking.

Table 6 Summary of data obtained from validation docking of BMP in ODCase crystal structure (PDB code: 1X1Z).²⁰¹

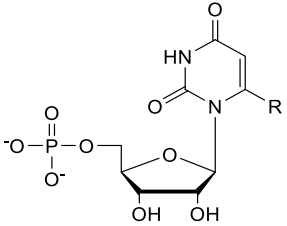

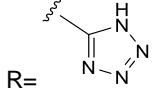
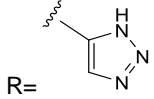
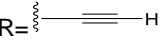
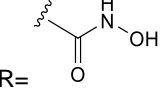
Structure	Average ChemPLP score (\pm standard deviation)	Average RMSD (\pm standard deviation) (Å)
	139.50 (3.31)	0.27 (0.06)

The data obtained from the validation docking shows that GOLD is suitable for carrying out further docking into ODCase.¹⁹⁸ The average RMSD value of 0.27 Å is well below the threshold value of 2 Å showing GOLD is able to match the crystallographic pose of BMP very well. Visual inspection of the 25 poses generated by GOLD in PyMOL showed good matching of the poses to the crystallographic BMP molecule which correlated with the low RMSD values obtained.^{198,200} The average ChemPLP score was also calculated so it could be compared to the scores from the eleven final compounds.

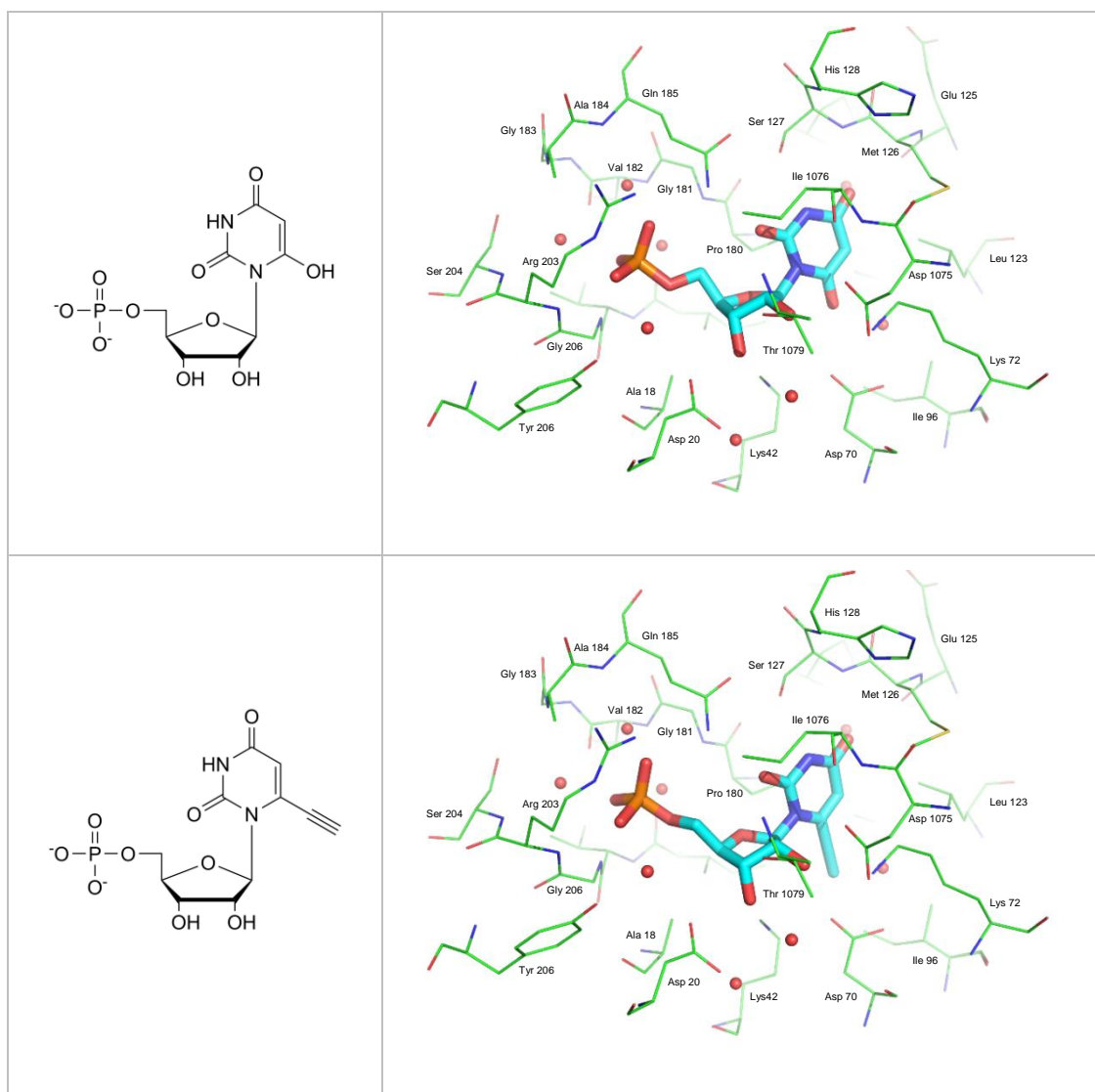
4.2 Docking of final compounds

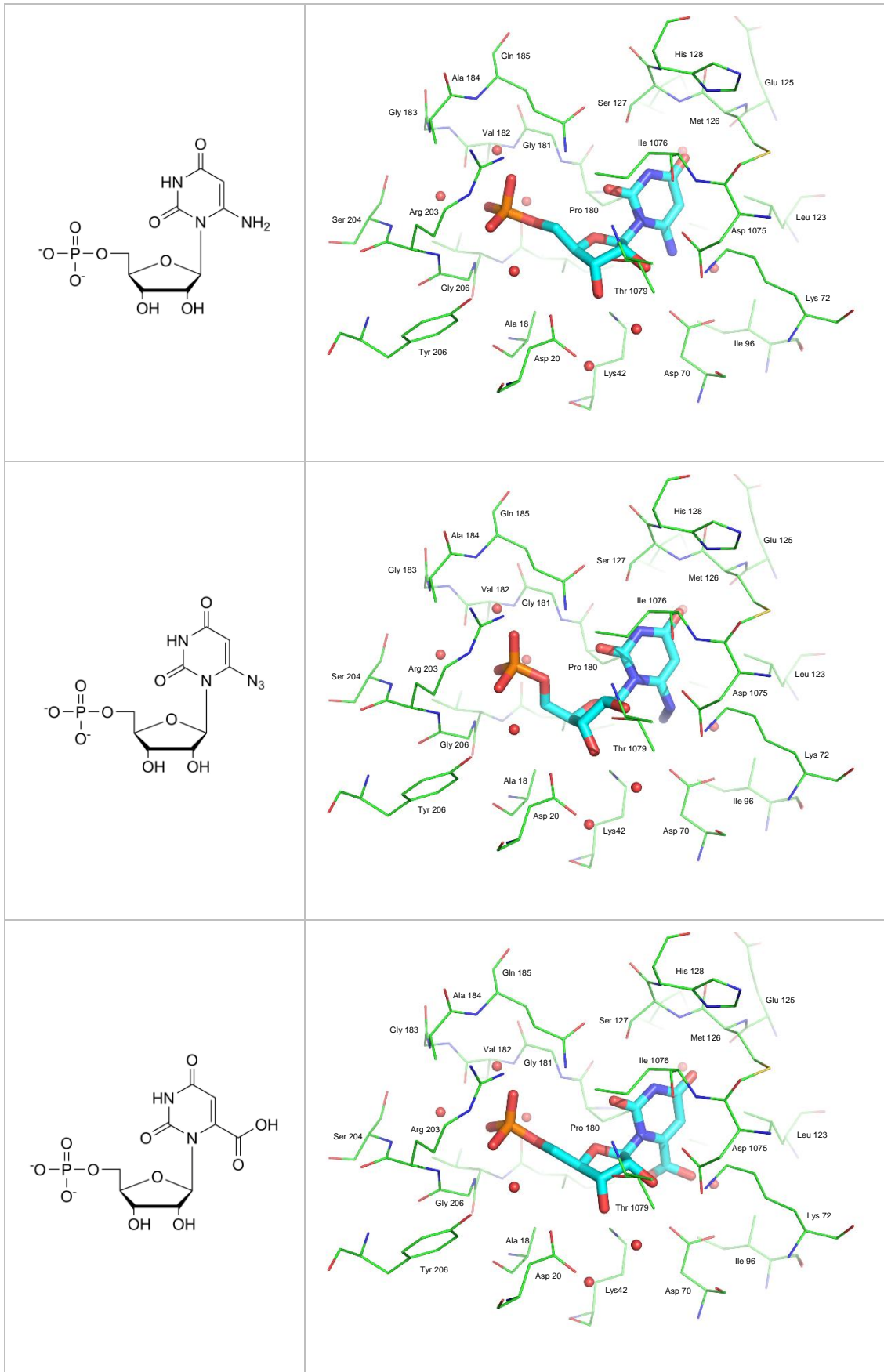
Having established that GOLD can successfully reproduce the binding pose of a known inhibitor of ODCase, the eleven final compounds synthesised in this project were docked and evaluated.¹⁹⁸ The compounds were modelled in Spartan '16 according to Protocol 2 as their 5' monophosphates.¹⁹⁷ It has been established that the 5'-monophosphate moiety of the enzyme's natural substrate, OMP, contributes a significant amount of intrinsic binding energy. The eleven compounds synthesised in this project were sent for biological testing in their nucleoside form as the presence of a 5' phosphate can often prevent them from entering cells. It is assumed that cellular kinases will phosphorylate the nucleosides upon entry to the cell and so for this computational work they were constructed as 5'-monophosphates. Each compound in turn was docked into the ODCase crystal structure as described in Protocol 4.¹⁹⁸ Table 7 summarises the results obtained.

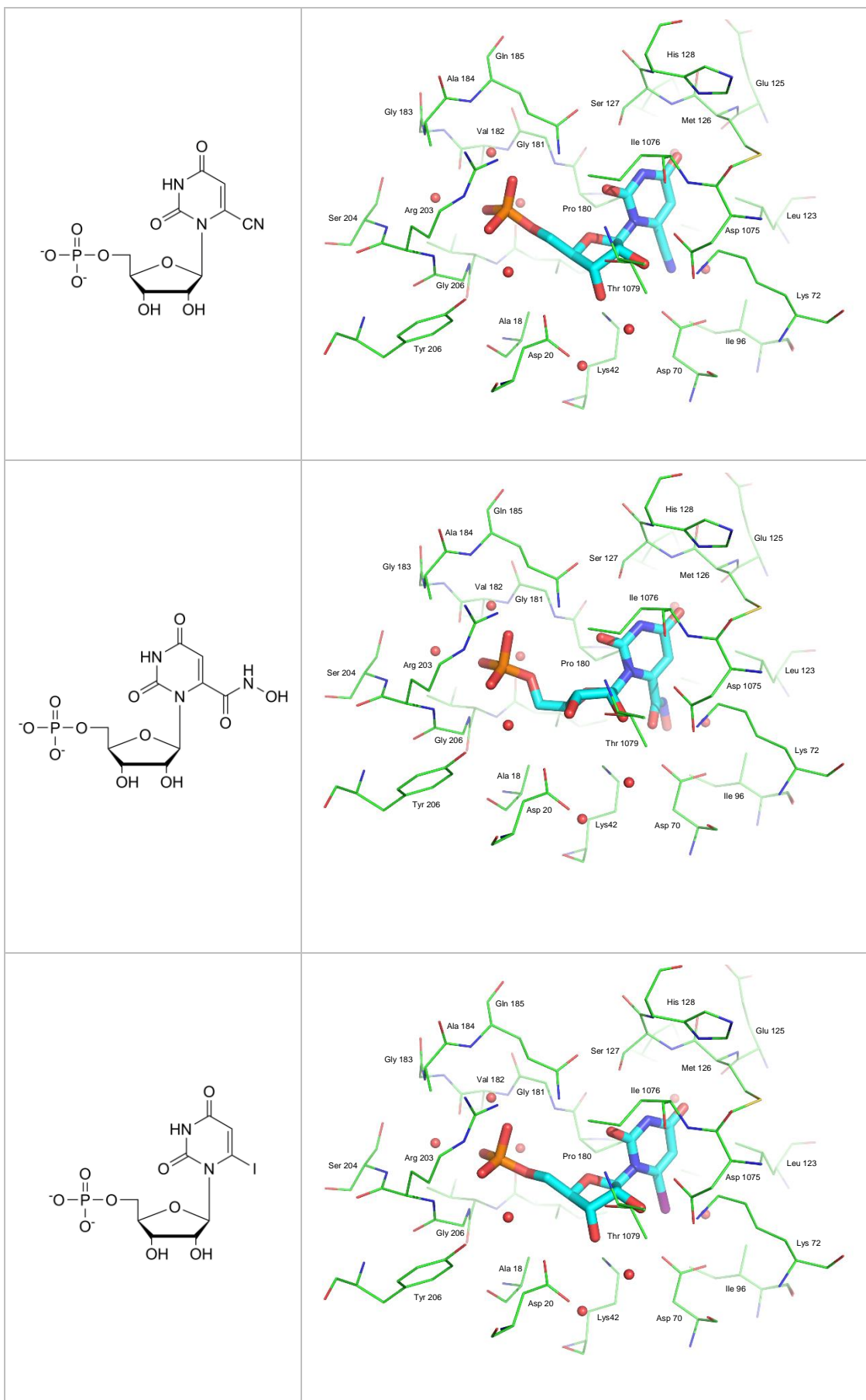
Table 7 Average ChemPLP score for the top 10 poses of each compound.¹⁹⁸

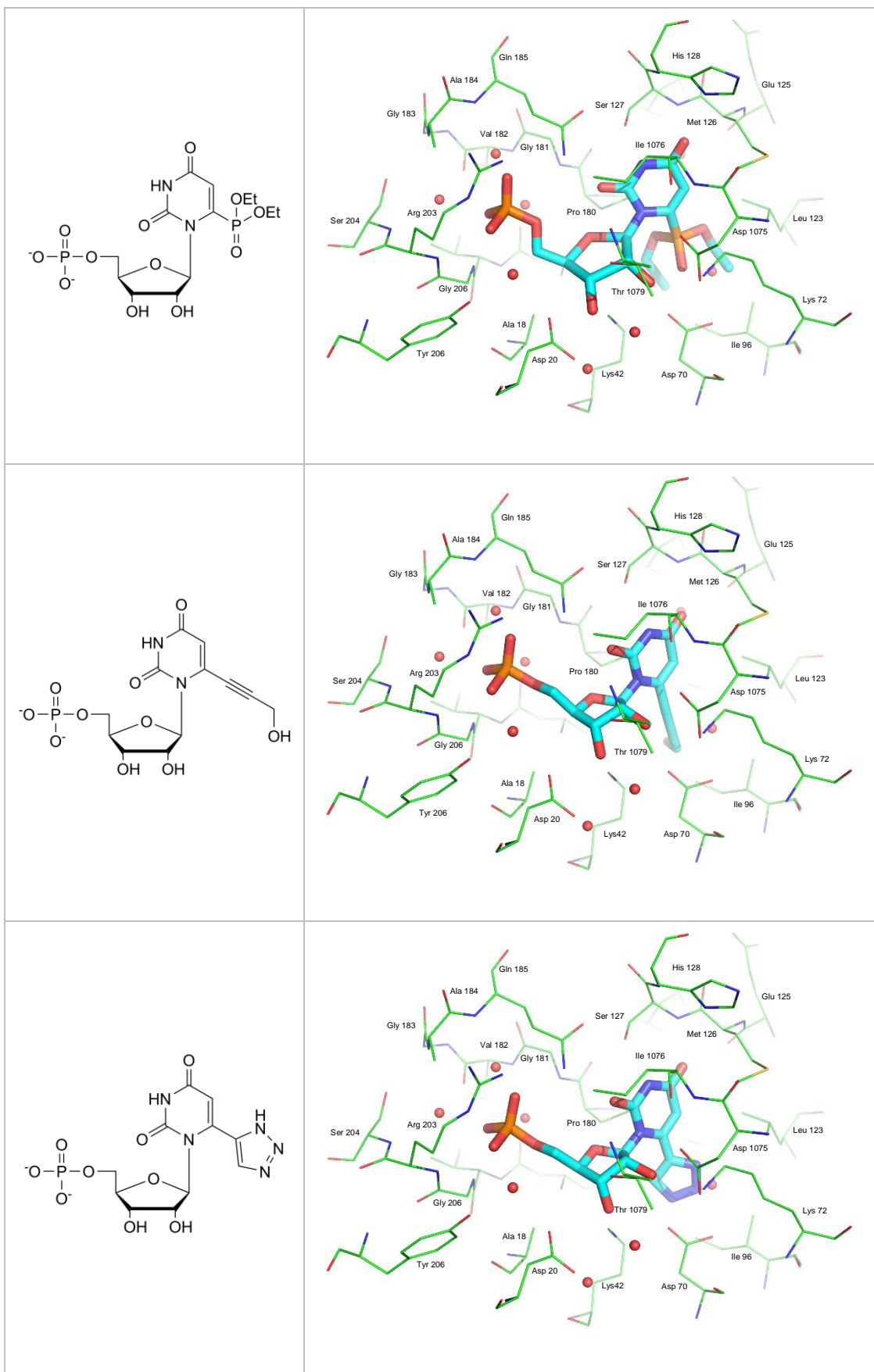
Compound	Average ChemPLP score for top 10 poses (\pm standard deviation)
	
R=OH	139.50 (3.31)
R=I	129.71 (1.50)
R=N ₃	135.85 (1.89)
R=COOH	138.21 (1.80)
R=CN	126.73 (1.49)
R=NH ₂	142.95 (2.45)
	124.23 (2.46)
	107.57 (1.27)
	110.99 (4.04)
R=PO(OEt) ₂	84.84 (3.10)
	135.79 (2.76)
	131.41 (1.96)

For the ChemPLP scoring function, the higher the value obtained, the better the binding of the compound is predicted to be. The results show that GOLD predicts 6-amino-UMP to bind most strongly to the active site of ODCase with the diethyl phosphonate derivative predicted to be the worst. The natural substrate, OMP, has a score just slightly higher than BMP and is second only to 6-amino-UMP. Visualisation of the poses generated by GOLD in PyMOL seemed to indicate that all eleven compounds fitted into the active site in a similar manner to BMP.²⁰⁰ The only compound that differed was the tetrazole derivative. All but two of its top ten poses had the 5' phosphate and the uracil base transposed compared to BMP and the other compounds. The images on the following pages show the top binding pose for each compound in the active site, visualised from the same orientation. The binding pose of BMP is shown for reference.









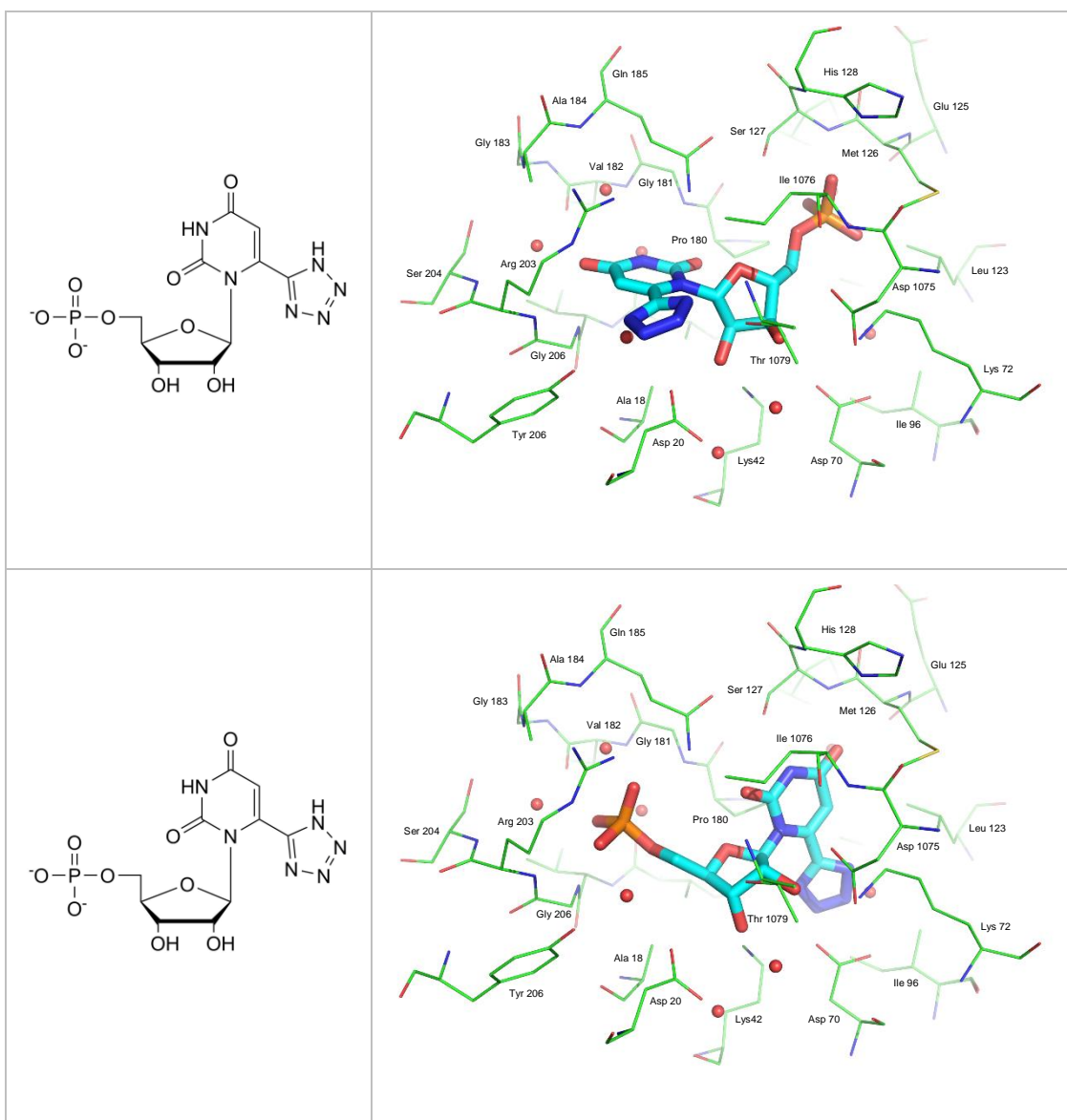


Figure 48 Top binding pose for each compound docked into ODCase crystal structure (PDB code: 1X1Z).²⁰¹ First entry shows binding of BMP into active site as a reference. Last two entries show the binding pose of the tetrazole derivative which is incorrectly orientated in one (top binding pose of this compound) and correctly orientated in the other. The images show residues within 6 Å of the active site as lines and crystallographic waters as red spheres.²⁰⁰ The docked compounds are shown as sticks (coloured by atom type; carbon-light blue, nitrogen-dark blue, oxygen-red, phosphorus-orange, iodine-purple).²⁰⁰

Despite having the lowest average ChemPLP score, as can be seen in the image for its top binding pose, even the diethyl phosphonate derivative appears to fit into the active site. However, as can be seen for the tetrazole derivative, a higher docking score can sometimes be misleading as GOLD predicted the majority of its docking poses in the wrong orientation.

4.3 Biological testing

As mentioned, four of the eleven final compounds (iodo **29**, amino **78**, azido **38**, cyano **52**) sent for biological testing have been synthesised previously and are

known to be inhibitors of ODCase. All were obtained in either one or two steps from key intermediates used in the project. All four have previously been tested in a *P. falciparum* 3D7 assay. They were re-synthesised and sent for biological testing in this project to provide benchmarks against which the novel compounds synthesised, such as the triazole derivative **71**, could be compared.

The previous assay was carried out by Kotra *et al.* and they found that 6-iodouridine exhibits potent *in vitro* anti-plasmodial activity.¹⁶⁴ They additionally found that its mononucleotide derivative 6-iodo-UMP is a covalent inhibitor of ODCase from both *Methanobacterium thermautotrophicum* and *Plasmodium falciparum*.¹⁶⁴ However, 6-cyanouridine, 6-aminouridine and 6-azidouridine showed no significant *in vitro* antimalarial activity despite all three mononucleotide forms showing inhibitory activity against ODCase.¹⁶⁶ These results show that good inhibitory activity against ODCase does not always result in antimalarial activity. Nevertheless, the example of 6-iodouridine gives confidence that molecules capable of inhibiting ODCase can produce an antimalarial effect.

Prior to sending for biological testing, the eleven compounds were assessed for purity by HPLC. Table 8 summarises the results.

Table 8 Purity of compounds as assessed by HPLC. Retention times denoted with ^a were obtained using HPLC Method 3 and those denoted ^b were obtained using HPLC Method 4.

Compound	Retention time (min)	Purity by HPLC (%)
6-Iodouridine 29	23.15 ^a	100
6-Azidouridine 38	18.43 ^a	93
6-Aminouridine 78	13.61 ^a	89
6-Ethynyl uridine 72	21.25 ^a	96
6-(1 <i>H</i> -1,2,3-Triazol-5-yl) uridine 71	15.77 ^a	100
6-(3-Hydroxyprop-1-yn-1-yl) uridine 66	20.80 ^a	98
6-Hydroxamic acid uridine 59	6.67 ^a	100
6-Cyanouridine 52	18.31 ^a	93
6-Diethyl phosphonate uridine 61	31.01 ^a	100
6-Carboxyuridine 56	7.69 ^b	100
6-(1 <i>H</i> -Tetrazol-5-yl) uridine 51	12.54 ^b	100

The compounds all showed a high level of purity and were deemed suitable to be assessed biologically. The compounds were sent to The Liverpool School of Tropical Medicine for biological assessment in an assay that uses 3D7 *Plasmodium falciparum* parasites. The results of the assay are currently being awaited.

4.4 Conclusions

The computational work carried out has shown that the docking program GOLD can accurately re-dock the known inhibitor BMP into the crystal structure of ODCase. Docking of compounds synthesised over the course of this project was carried out and visualisation of the results generated indicate that they all fit into the active site of the enzyme. The amino derivative was predicted to bind most strongly with the diethyl phosphonate derivative predicted to bind least strongly, possibly due to its greater size.

The compounds have also been sent for biological testing to assess their antimalarial properties. Once the results of the testing have been received, further development of any compounds showing antimalarial activity will be conducted.

Chapter 5
Experimental

Chapter 5 Experimental

5.1.1 General techniques

Analytical thin layer chromatography (TLC)

TLC was performed on UV₂₅₄ sensitive, silica gel coated, aluminium TLC plates purchased from Merck. Plates were visualised under UV light and treated with a sugar stain. Following treatment with the sugar stain and heating, compounds containing a ribose moiety stained on the TLC plates as black spots.

Flash column chromatography

The required quantity of silica gel was made into a homogenous slurry using the column eluent. The slurry was applied to the column over a base layer of sand. The crude material to be purified was then introduced to the column in either the minimum volume of column eluent, or, as a powder due to pre-absorption onto silica. Eluted fractions were collected and analysed by TLC.

Nuclear magnetic resonance (NMR) spectroscopy

All NMR spectra were recorded on either a Bruker Avance 400 MHz or 500 MHz spectrometer in the deuterated solvent stated. Chemical shifts are reported in ppm and coupling constants (*J*) are reported in Hz. All ¹³C and ³¹P NMR spectra were proton decoupled. The chemical shifts reported for all samples run in deuterated chloroform (CDCl₃) are relative to an internal standard of tetramethylsilane.

Mass spectrometry

Unless otherwise stated (please see HPLC-MS), all mass spectra were recorded by Ms Moya McCarron in the Mass Spectrometry Department at the University of Liverpool using a Micromass LCT Mass Spectrometer in the ES ionisation mode. Samples were injected using a direct infusion syringe pump.

Infrared spectroscopy

IR spectra were recorded using a Bruker FTIR Alpha spectrometer.

pH Measurements

Measurements of pH were recorded using a pH probe connected to a Mettler Toledo T50 autotitrator system.

High performance liquid chromatography (HPLC)

HPLC was performed on an Agilent 1260 Infinity system equipped with an autoinjector, a photodiode array detector and quaternary pump. Chromatographic data was controlled and handled using Agilent OpenLAB Chemstation software.

Reverse phase (RP) HPLC was performed on a Gemini® 5 µm C18 column (110 Å, 250 mm x 4.6 mm) purchased from Phenomenex®. The eluent system and control method (see HPLC methods) used were changed depending on the polarity of the compounds being analysed. The RP-HPLC column was stored in HPLC grade MeCN when not in use.

High performance liquid chromatography – mass spectrometry (HPLC-MS)

HPLC-MS was performed on an Agilent 6530B accurate mass Q-TOF mass spectrometer connected to an Agilent 1260 Infinity HPLC system (as described in HPLC section). Samples were run on the RP-HPLC column described in the HPLC section. HPLC-MS data was obtained using a multimode ion source and processed using Agilent MassHunter software.

5.1.2 HPLC solvents

Acetonitrile – HPLC gradient grade was purchased from Fisher Scientific.

Acetic acid – HPLC gradient grade was purchased from Sigma Aldrich. The desired concentration of aqueous acetic acid solution was then made up using distilled water.

Triethylammonium bicarbonate solution (TEAB) – prepared by bubbling CO₂ gas through a solution of triethylamine (HPLC gradient grade purchase from Sigma Aldrich) and distilled water until a pH between 7.5 and 7.7 was obtained. This solution was diluted to give a final concentration of 1 M which was then further diluted as necessary.

5.1.3 HPLC methods

HPLC Method 1

The eluent was gradually changed from 100% 0.1 M TEAB solution to 100% of a 40% MeCN in 0.1 M TEAB solution over 20 minutes before being changed back to 100% 0.1 M TEAB solution over 3 minutes. The eluent was held at 100% 0.1 M TEAB solution for the final 2 minutes of the run.

Column temperature – 25 °C

Flow rate – 1 mL/min

Injection volume – 5 µL

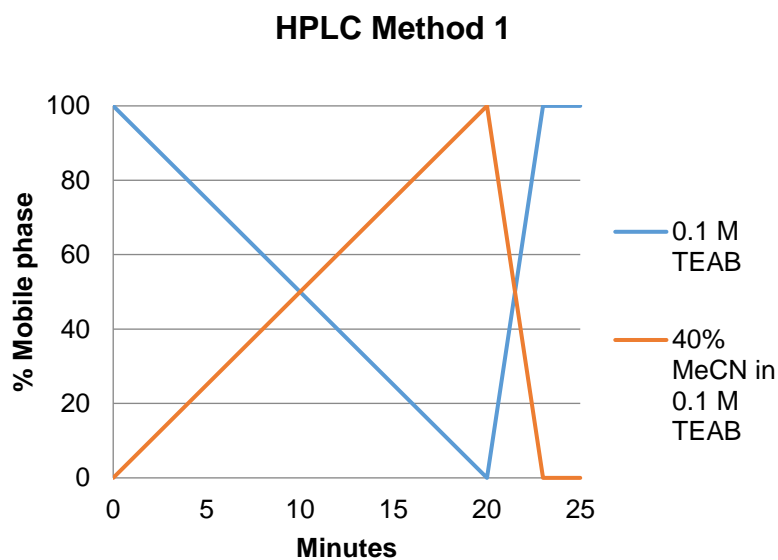


Figure 49 HPLC method 1

HPLC Method 2

The eluent was held at 96% 0.1% AcOH in water and 4% MeCN for the first 2 minutes of the run. It was then gradually altered to 80% 0.1% AcOH in water and 20% MeCN over the next 10 minutes and held at this ratio for 3 minutes. The eluent was then changed back to 96% 0.1% AcOH in water and 4% MeCN over 3 minutes and held at this ratio for the final 4 minutes of the run.

Column temperature – 40 °C

Flow rate – 0.5 mL/min

Injection volume – 5 µL

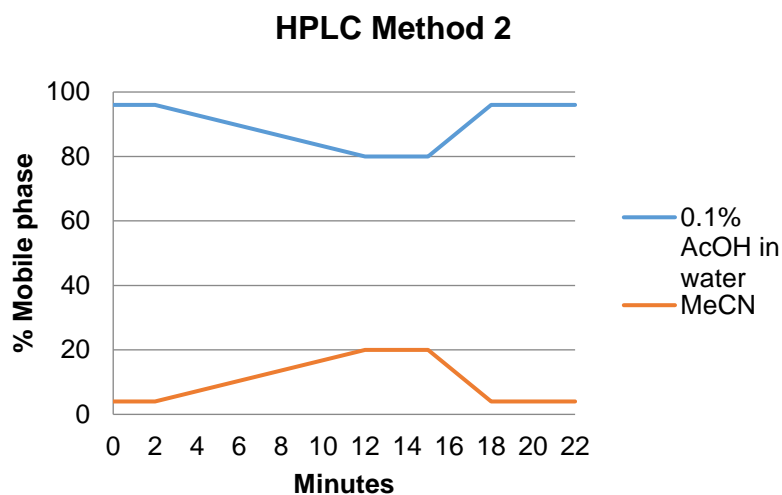


Figure 50 HPLC method 2

HPLC Method 3

The eluent was held at 98% 0.1% AcOH in water and 2% MeCN for the first 5 minutes of the run. It was then gradually altered to 80% 0.1% AcOH in water and 20% MeCN over the next 20 minutes and held at this ratio for 2 minutes. The eluent was then changed back to 98% 0.1% AcOH in water and 2% MeCN over 3 minutes and held at this ratio for the final 5 minutes of the run.

Column temperature – 40 °C

Flow rate – 0.5 mL/min

Injection volume – 5 μ L

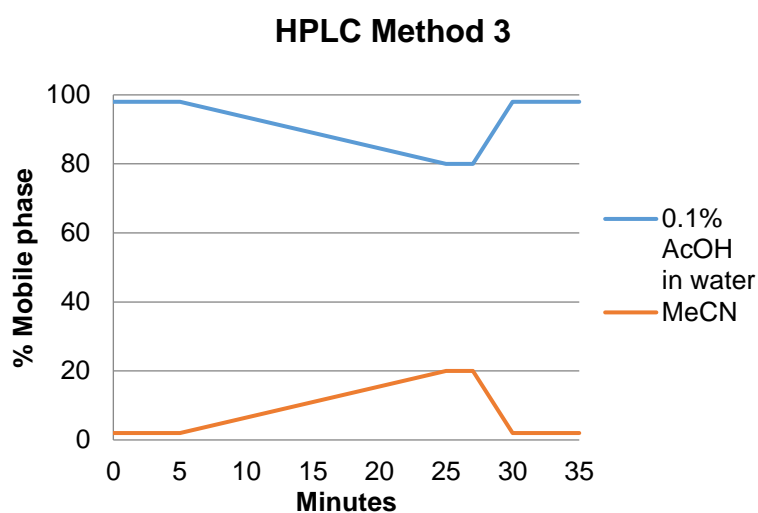


Figure 51 HPLC method 3

HPLC Method 4

The eluent was held at 98% 0.1 M TEAB solution and 2% MeCN for the first 2 minutes of the run. It was then gradually altered to 90% 0.1 M TEAB solution and 10% MeCN over the next 10 minutes and held at this ratio for 2 minutes. The eluent was then changed back to 98% 0.1 M TEAB solution and 2% MeCN over 2 minutes and held at this ratio for the final 2 minutes of the run.

Column temperature – 40 °C

Flow rate – 0.5 mL/min

Injection volume – 5 µL

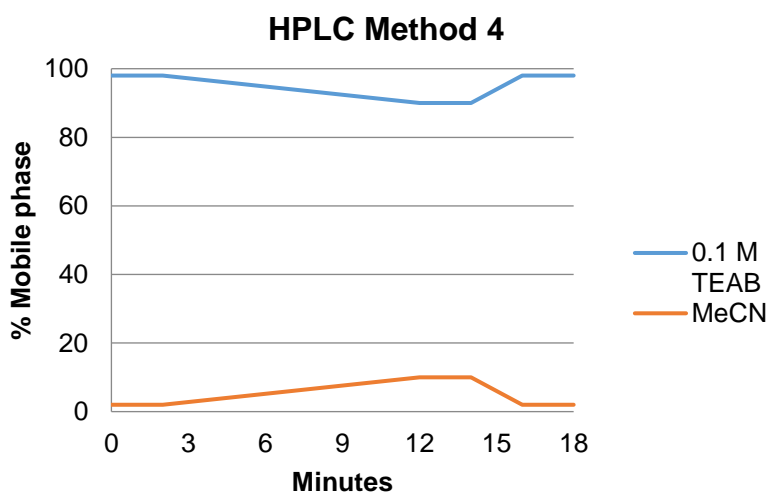


Figure 52 HPLC method 4

5.1.4 Solvents

Unless otherwise stated, all solvents were purchased from Fisher Scientific. Anhydrous solvents were obtained as follows:

Acetone – AcroSeal™ anhydrous solvent purchased from Acros Organics.

Acetonitrile – Sure/Seal™ anhydrous solvent purchased from Sigma Aldrich.

Dichloromethane – anhydrous solvent obtained *via* solvent passage through drying columns supplied by Innovative Technology Ltd.

N,N-Dimethylformamide – Sure/Seal™ anhydrous solvent purchased from Sigma Aldrich.

Pyridine – Sure/Seal™ anhydrous solvent purchased from Sigma Aldrich.

THF – anhydrous solvent obtained *via* solvent passage through drying columns supplied by Innovative Technology Ltd.

Triethylamine – Sure/Seal™ anhydrous solvent purchased from Sigma Aldrich.

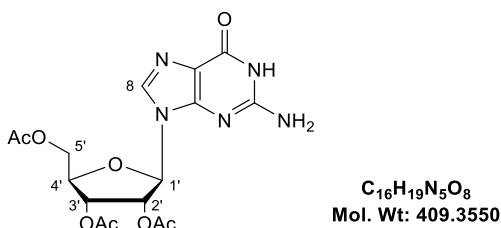
5.1.5 General reagents

Unless otherwise stated, general reagents were all purchased from Acros Organics, Fluorochem or Sigma Aldrich.

Sugar stain – Phenol (3 g) was dissolved in a solution of ethanol (95 mL) and concentrated sulphuric acid (5 mL).

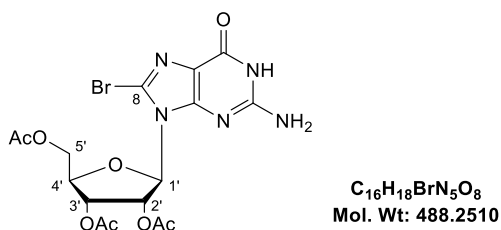
5.2 Experimental procedures for results and discussion 1

2',3',5'-Tri-*O*-acetylguanosine **8**^{84,104}



DMAP (0.13 g, 1.06 mmol, 0.1 eqv.) and Et₃N (8.86 mL, 63.6 mmol, 6 eqv.) were added to a suspension of guanosine (3.09 g, 10.9 mmol, 1 eqv.) in anhydrous MeCN (40 mL) and cooled to 0 °C. Acetic anhydride (3.00 mL, 31.7 mmol, 3 eqv.) was added dropwise and the resulting reaction mixture was allowed to warm to room temperature. The reaction was found to be complete following full solubilisation of the reactants. MeOH (~5 mL) was added to quench the reaction which was then concentrated *in vacuo* to leave a thick, cloudy oil. The desired product was isolated by recrystallisation from the minimum amount of hot *i*PrOH and then washed with Et₂O to remove residual Et₃N. This yielded a white solid as product (3.90 g, 87%). ¹H NMR (400 MHz, DMSO-*d*₆): δ (ppm) 10.78 (1H, br s, NH), 7.93 (1H, s, H8), 6.55 (2H, br s, NH₂), 5.99 (1H, d, *J* 6.0, H1'), 5.79 (1H, app t, *J* 6.0, H2'), 5.49 (1H, dd, *J* 5.6 and 4.4, H3'), 4.38 (1H, dd, *J* 11.2 and 3.6, H5'), 4.33-4.30 (1H, m, H4'), 4.26 (1H, dd, *J* 11.2 and 5.6, H5''), 2.11 (3H, s, COCH₃), 2.04 (3H, s, COCH₃), 2.03 (3H, s, COCH₃). ¹³C NMR (100 MHz, DMSO-*d*₆): δ (ppm) 170.10 (CO), 169.45 (CO), 169.28 (CO), 156.69 (C6), 153.94 (C2), 151.12 (C4), 135.62 (C8), 116.83 (C5), 84.41 (C1'), 79.54 (C4'), 72.05 (C2'), 70.31 (C3'), 63.08 (C5'), 20.53 (COCH₃), 20.38 (COCH₃), 20.19 (COCH₃). HRMS (ES⁺) (*m/z*): 432.1129 ([M+Na]⁺); C₁₆H₁₉N₅O₈Na requires 432.1126 (+0.6943 ppm).

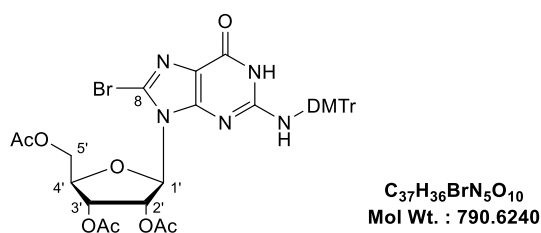
8-Bromo-2',3',5'-tri-*O*-acetylguanosine **9**^{84,104}



Aliquots of bromine water (2 mL, ~30 mL in total) were added to a vigorously stirred suspension of 2',3',5'-tri-*O*-acetylguanosine (3.39 g, 8.28 mmol) in distilled H₂O (30

mL) until the colour of the reaction remained permanently yellow. The reaction mixture was then filtered under vacuum and the solid collected was washed with cold *i*PrOH and Et₂O. This yielded a pale orange solid as product (3.09 g, 76%). ¹H NMR (400 MHz, DMSO-d₆): δ (ppm) 10.91 (1H, br s, NH), 6.60 (2H, br s, NH₂), 6.02 (1H, dd, *J* 6.2 and 4.6, H2'), 5.89 (1H, d, *J* 4.4, H1'), 5.65 (1H, app t, *J* 6.0, H3'), 4.42 (1H, dd, *J* 11.8 and 3.8, H5'), 4.34-4.30 (1H, m, H4'), 4.21 (1H, dd, *J* 11.8 and 6.2, H5''), 2.11 (3H, s, COCH₃), 2.07 (3H, s, COCH₃), 2.00 (3H, s, COCH₃). ¹³C NMR (100 MHz, DMSO-d₆): δ (ppm) 170.09 (CO), 169.47 (CO), 169.40 (CO), 155.40 (C6), 153.81 (C2), 151.92 (C4), 120.10 (C8), 117.18 (C5), 87.63 (C1'), 79.30 (C4'), 71.27 (C2'), 69.92 (C3'), 62.76 (C5'), 20.47 (COCH₃), 20.30 (COCH₃), 20.24 (COCH₃). HRMS (ES⁺) (*m/z*): 510.0231 and 512.0211 ([M+Na]⁺); C₁₆H₁₈BrN₅O₈Na requires 510.0231 and 512.0213 (0.0000 and -0.3906 ppm).

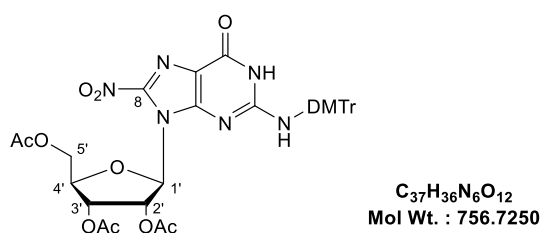
N²-Dimethoxytrityl-8-bromo-2',3',5'-tri-O-acetylguanosine 10^{84,104}



8-Bromo-2',3',5'-tri-O-acetylguanosine (3.23 g, 6.62 mmol, 1 eqv.) was dried by co-evaporation of water with anhydrous pyridine (3 x 10 mL) and then dissolved in anhydrous pyridine (30 mL). DMTrCl (4.49 g, 13.3 mmol, 2 eqv.) was added in portions, under a N₂ atmosphere, over the course of 5 minutes and the resulting reaction mixture was left to stir overnight. Upon completion, the reaction was partitioned between H₂O (30 mL) and DCM (30 mL). The aqueous layer was extracted with DCM (2 x 30 mL) and the combined organic extracts were washed with H₂O (3 x 50 mL) and saturated NaHCO₃ solution (30 mL) before being dried over Na₂SO₄. All solvent was removed *in vacuo* to leave a brown oil which was purified by flash column chromatography (100% DCM – 2% MeOH/98% DCM) to give a yellow solid as product (4.33 g, 83%). R_f = 0.61 (5% MeOH/95% DCM). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 7.30-7.19 (9H, m, DMTr Ar-H), 6.85-6.83 (4H, m, DMTr CHCOCH₃), 5.58 (1H, d, *J* 6.0, H1'), 5.23 (1H, app t, *J* 6.4, H2'), 4.97 (1H, dd, *J* 6.6 and 4.6, H3'), 4.02-3.98 (1H, m, H4'), 3.90-3.82 (2H, m, H5' and H5''), 3.76 (6H, s, DMTr Ar-OCH₃), 2.09 (3H, s, COCH₃), 2.01 (3H, s, COCH₃), 1.94 (3H, s, COCH₃). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 172.01 (CO), 170.98 (CO), 170.79

(CO), 160.06 (Ar-C), 160.00 (Ar-C), 158.07 (C6), 153.03 (C2), 152.83 (C4), 131.16 (Ar-C), 131.03 (Ar-C), 129.73 (Ar-C), 129.09 (Ar-C), 128.12 (Ar-C), 123.23 (C8), 119.18 (C5), 114.32 (Ar-C), 114.26 (Ar-C), 89.32 (C1'), 80.65 (C4'), 71.89 (C3'), 71.78 (NHC), 71.36 (C2'), 64.20 (C5'), 55.68 (Ar-OCH₃), 20.70 (COCH₃), 20.45 (COCH₃), 20.16 (COCH₃). Signals not observed: 3 x quaternary carbons from DMTr group. HRMS (ES⁺) (m/z): 812.1556 and 814.1500 ([M+Na]⁺); C₃₇H₃₆BrN₅O₁₀Na requires 812.1538 and 814.1525 (+2.2163 and -3.0707 ppm).

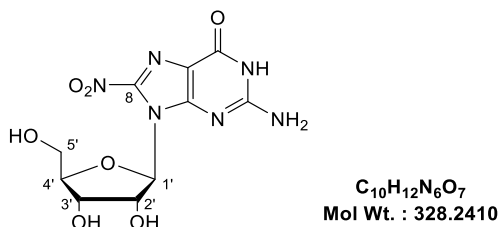
N2-Dimethoxytrityl-8-nitro-2',3',5'-tri-*O*-acetylguanosine **11**^{84,104}



*N*2-Dimethoxytrityl-8-bromo-2',3',5'-tri-*O*-acetylguanosine (1.50 g, 1.90 mmol, 1 eqv.) was dissolved in anhydrous DMF (105 mL). KNO₂ (1.62 g, 19.0 mmol, 10 eqv.) and 18-crown-6 (5.01g, 19.0 mmol, 10 eqv.) were added. The reaction mixture was heated to 100 °C for 6 hours then cooled to room temperature and poured into brine (100 mL). The aqueous solution was extracted with EtOAc (3 x 100 mL) and the combined organic extracts were washed with saturated NaHCO₃ solution (100 mL) and dried over Na₂SO₄. All solvent was removed *in vacuo* to leave an orange oil. The crude product was purified by flash column chromatography (50% Hex/50% EtOAc – 100% EtOAc) to yield a yellow solid as product (0.20 g, 14%). ¹H NMR (400 MHz, DMSO-*d*₆): δ (ppm) 11.19 (1H, br s, NH), 8.06 (1H, br s, NH), 7.35-7.12 (9H, m, DMTr Ar-H), 6.90-6.87 (4H, m, DMTr CHCOCH₃), 5.58 (1H, d, *J* 4.8, H1'), 5.46 (1H, dd, *J* 7.0 and 5.0, H2'), 5.15 (1H, app t, *J* 6.8, H3'), 4.14 (1H, dd, *J* 12.0 and 3.6, H5'), 4.01-3.97 (1H, m, H4'), 3.86 (1H, dd, *J* 12.0 and 6.0, H5''), 3.73 (3H, s, DMTr Ar-OCH₃), 3.72 (3H, s, DMTr Ar-OCH₃), 2.05 (3H, s, COCH₃), 2.00 (3H, s, COCH₃), 1.78 (3H, s, COCH₃). ¹³C NMR (100 MHz, DMSO-*d*₆): δ (ppm) 169.87 (CO), 169.13 (CO), 168.82 (CO), 157.97 (quat C), 157.91 (quat C), 156.02 (quat C), 153.31 (quat C), 151.28 (quat C), 144.66 (quat C), 142.89 (quat C), 136.46 (quat C), 135.98 (quat C), 129.68 (Ar-C), 129.63 (Ar-C), 128.23 (Ar-C), 127.89 (Ar-C), 126.83 (Ar-C), 115.24 (quat C), 113.16 (2 x CHCOCH₃), 113.10 (2 x CHCOCH₃), 87.25 (C1'), 78.24 (C4'), 70.88 (C2'), 70.00 (NHC), 68.62 (C3'), 61.83 (C5'), 55.02 (Ar-

OCH₃), 55.00 (Ar-OCH₃), 20.37 (COCH₃), 20.20 (COCH₃), 19.84 (COCH₃). HRMS (ES⁺) (m/z): 779.2282 ([M+Na]⁺); C₃₇H₃₆N₆O₁₂Na requires 779.2283 (-0.1283 ppm).

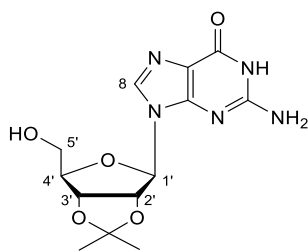
8-Nitroguanosine 13^{84,104}



To a solution of *N*2-dimethoxytrityl-8-nitro-2',3',5'-tri-*O*-acetylguanosine (0.20 g, 0.26 mmol, 1 eqv.) in CHCl₃ (10 mL) was added a solution of *p*TsOH (10 mg, 0.05 mmol, 0.2 eqv.) in MeOH (1 mL). After 20 minutes the reaction was concentrated under reduced pressure to leave an orange residue as crude. This was then triturated with Et₂O and carried forward to the final deacetylation step without further purification. The detritylated nucleoside was dissolved in NH₃ (7 M) in MeOH (2 mL) and allowed to stir at room temperature for 48 hours in a sealed reaction vial. All solvent was then removed *in vacuo* and the resulting orange solid was triturated several times with Et₂O to yield the desired product (46 mg, 54% over two steps). ¹H NMR (400 MHz, DMSO-d₆): δ (ppm) 6.30 (1H, d, *J* 5.6, H1'), 4.93 (1H, app t, *J* 5.8, H2'), 4.19 (1H, app t, *J* 5.2, H3'), 3.88-3.84 (1H, m, H4'), 3.68 (1H, dd, *J* 11.8 and 4.2, H5'), 3.53 (1H, dd, *J* 12.0 and 5.6, H5''). Signals not observed: NH, NH₂, 2'-OH, 3'-OH, 5'-OH. ¹³C NMR (100 MHz, DMSO-d₆): δ (ppm) 159.91 (quat C), 157.64 (quat C), 152.78 (quat C), 143.18 (quat C), 115.70 (quat C), 90.40 (C1'), 85.84 (C4'), 71.10 (C2'), 70.28 (C3'), 61.98 (C5'). HRMS (ES⁻) (m/z): 327.0704 ([M-H]⁻); C₁₀H₁₁N₆O₇ requires 327.0695 (+ 2.7517 ppm)

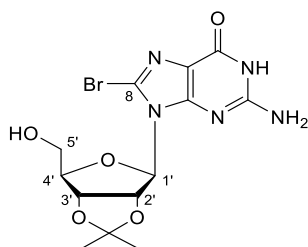
General procedure for 8-nitroguanosine reduction reactions

8-Nitroguanosine (1.0 mg, 3.1 μmoles) was dissolved in distilled H₂O (1 mL) with stirring. Reducing agent (50 eqv.) was added, with additional portions (50 eqv.) subsequently added every 30 minutes for 90 minutes. Reactions were monitored by HPLC with samples taken every 30 minutes from the reaction start.

2',3'-Isopropylidene guanosine 18

$C_{13}H_{17}N_5O_5$
Mol Wt. : 323.3090

To a suspension of guanosine (3.01 g, 10.6 mmol, 1 eqv.) in anhydrous acetone (120 mL) was added *p*TsOH (1.83 g, 9.62 mmol, 0.9 eqv.) and DMP (30 mL, 245 mmol, 23.1 eqv). The resulting reaction mixture was allowed to stir at room temperature overnight. All solvent was then removed under reduced pressure to leave an off-white solid. The solid was dissolved in water (20 mL) and solid $NaHCO_3$ (0.89 g, 10.6 mmol, 1 eqv.) was added in portions which produced a white foam that was stirred for 2 hours. A saturated $NaHCO_3$ solution (10 mL) was then added and the reaction was stirred for a further 2 hours, over which time a suspension formed. The reaction mixture was filtered under vacuum and the white solid isolated was washed several times with cold water then dried under vacuum to yield the desired product (2.27 g, 66%). 1H NMR (400 MHz, $DMSO-d_6$): δ (ppm) 10.87 (1H, br s, NH), 7.89 (1H, s, H8), 6.58 (2H, br s, NH_2), 5.92 (1H, d, J 2.8, H1'), 5.19 (1H, dd, J 6.4 and 2.8, H2'), 5.08 (1H, br s, 5'-OH), 4.96 (1H, dd, J 6.2 and 3.0, H3'), 4.13-4.10 (1H, m, H4'), 3.57-3.49 (2H, m, H5' and H5''), 1.51 (3H, s, CH_3), 1.31 (3H, s, CH_3). ^{13}C NMR (100 MHz, $DMSO-d_6$): δ (ppm) 157.72 (quat C), 154.38 (quat C), 150.80 (quat C), 135.69 (C8), 116.81 (quat C), 113.03 (quat C), 88.50 (C1'), 86.59 (C4'), 83.52 (C2'), 81.21 (C3'), 61.64 (C5'), 27.08 (CH_3), 25.25 (CH_3). HRMS (ES⁺) (m/z): 346.1122 ([M+Na]⁺); $C_{13}H_{17}N_5O_5Na$ requires 346.1122 (0.0000 ppm).

8-Bromo-2',3'-isopropylidene guanosine 19

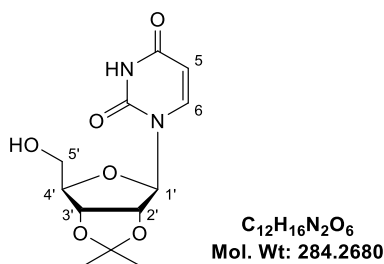
$C_{13}H_{16}BrN_5O_5$
Mol Wt. : 402.2050

2',3'-Isopropylidene guanosine (2.38 g, 7.36 mmol) was suspended in distilled H_2O (25 mL). Aliquots of bromine water (5 mL) were added until the reaction mixture remained permanently yellow. The reaction mixture was stirred for 6 hours at room

temperature and then filtered under vacuum. The solid obtained was washed with the reaction liquor then dried under vacuum to yield the desired product as a pale yellow solid (1.98 g, 67%). ^1H NMR (400 MHz, DMSO-d_6): δ (ppm) 10.86 (1H, s, NH), 6.64 (2H, br s, NH_2), 5.89 (1H, d, J 1.6, H1'), 5.44 (1H, dd, J 6.4 and 1.6, H2'), 5.12 (1H, dd, J 6.4 and 4.0, H3'), 4.08-4.03 (1H, m, H4'), 3.54 (1H, dd, J 11.4 and 5.8, H5'), 3.47 (1H, dd, J 11.6 and 6.8, H5''), 1.52 (3H, s, CH_3), 1.32 (3H, s, CH_3). Signal not observed: 5'-OH. ^{13}C NMR (100 MHz, DMSO-d_6): δ (ppm) 155.47 (quat C), 153.66 (quat C), 151.51 (quat C), 120.29 (quat C), 117.09 (quat C), 113.08 (quat C), 89.80 (C1'), 88.33 (C4'), 82.83 (C2'), 81.51 (C3'), 61.83 (C5'), 27.07 (CH_3), 25.35 (CH_3). HRMS (ES^+) (m/z): 424.0244 and 426.0216 ($[\text{M}+\text{Na}]^+$); $\text{C}_{13}\text{H}_{16}\text{BrN}_5\text{O}_5\text{Na}$ requires 424.0227 and 426.0208 (+4.0092 and +1.8778 ppm).

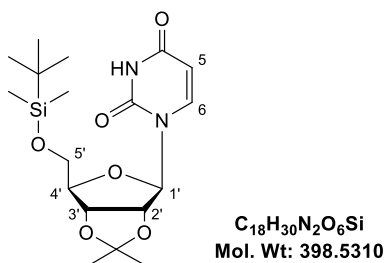
5.3 Experimental procedures for results and discussion 2

2',3'-Isopropylidene uridine 26¹⁶⁴



Concentrated H_2SO_4 (1 mL) was added dropwise to a suspension of uridine (2.00 g, 8.19 mmol) in anhydrous acetone (85 mL) and stirred for 3 hours. The reaction was quenched by the addition of Et_3N (5.2 mL) and all solvent was removed *in vacuo* to leave a thick, yellowish oil as crude. The crude material was purified by flash column chromatography (100% DCM-5% MeOH/95% DCM) to yield the desired product as a white solid (2.08 g, 89%). $R_f = 0.25$ (5% MeOH/95% DCM). 1H NMR (400 MHz, $DMSO-d_6$): δ (ppm) 11.37 (1H, s, NH), 7.79 (1H, d, J 8.0, H6), 5.83 (1H, d, J 2.4, H1'), 5.63 (1H, d, J 8.0, H5), 5.08 (1H, t, J 4.6, 5'-OH), 4.89 (1H, dd, J 6.0 and 2.4, H2'), 4.74 (1H, dd, J 6.4 and 3.6, H3'), 4.08-4.05 (1H, m, H4'), 3.62-3.52 (2H, m, H5' and H5''), 1.48 (3H, s, CH_3), 1.28 (3H, s, CH_3). ^{13}C NMR (100 MHz, $DMSO-d_6$): δ (ppm) 163.18 (CO), 150.34 (CO), 141.92 (C6), 112.97 (quat C), 101.74 (C5), 91.12 (C1'), 86.52 (C4'), 83.68 (C2'), 80.48 (C3'), 61.27 (C5'), 27.05 (CH_3), 25.19 (CH_3). HRMS (ES⁺) (m/z): 307.0904 ([M+Na]⁺); $C_{12}H_{16}N_2O_6Na$ requires 307.0901 (+0.9769 ppm).

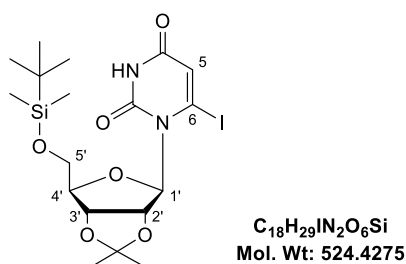
5'-*tert*-Butyldimethylsilyl-2',3'-isopropylidene uridine 27¹⁶⁴



2',3'-Isopropylidene uridine (3.00 g, 10.6 mmol, 1 eqv.) was suspended in anhydrous DCM (30 mL) and cooled to 0 °C. Imidazole (1.45 g, 21.2 mmol, 2 eqv.) and TBDMSCl (2.39 g, 15.9 mmol, 1.5 eqv.) were added under N_2 and the reaction mixture was allowed to warm to room temperature. After 3 hours, TLC showed the

reaction to be complete. All solvent was removed *in vacuo* and the resulting crude material was dissolved in EtOAc (50 mL) and washed with water (30 mL) and brine (30 mL). The organic layer was dried over Na₂SO₄ and then concentrated under reduced pressure. The crude material was purified by flash column chromatography (100% DCM – 3% MeOH/97% DCM) to yield a white solid as product (3.79 g, 90%). R_f = 0.44 (5% MeOH/95% DCM). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 8.59 (1H, br s, NH), 7.70 (1H, d, *J* 8.0, H6), 5.99 (1H, d, *J* 2.8, H1'), 5.68 (1H, dd, *J* 8.0 and 2.0, H5), 4.76 (1H, dd, *J* 6.0 and 2.8, H3'), 4.68 (1H, dd, *J* 6.4 and 2.8, H2'), 4.33-4.31 (1H, m, H4'), 3.93 (1H, dd, *J* 11.4 and 2.2, H5'), 3.80 (1H, dd, *J* 11.4 and 3.0, H5''), 1.59 (3H, s, CH₃), 1.36 (3H, s, CH₃), 0.90 (9H, s, 3 x CH₃), 0.10 (3H, s, SiCH₃), 0.09 (3H, s, SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 163.04 (CO), 150.11 (CO), 140.66 (C6), 114.30 (quat C), 102.32 (C5), 92.07 (C1'), 86.80 (C4'), 85.55 (C2'), 80.43 (C3'), 63.50 (C5'), 27.43 (CH₃), 26.00 (3 x CH₃), 25.51 (CH₃), 18.49 (quat C), -5.30 (SiCH₃), -5.40 (SiCH₃). HRMS (ES⁺) (*m/z*): 421.1767 ([M+Na]⁺); C₁₈H₃₀N₂O₆SiNa requires 421.1765 (+0.4749 ppm).

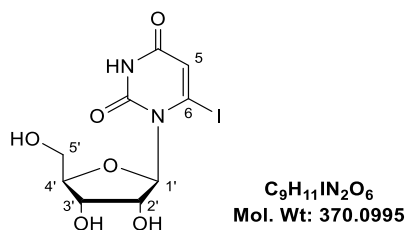
6-Iodo-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine **28**¹⁶⁴



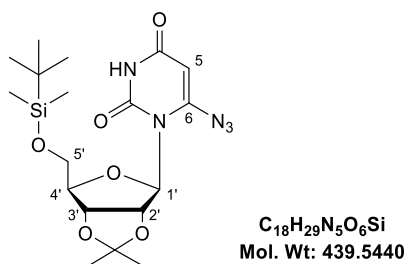
A solution of anhydrous diisopropylamine (1.03 mL, 7.37 mmol, 2.6 eqv.) in anhydrous THF (10 mL) was cooled to -78 °C. To this solution was added 2.5 M *n*-BuLi (2.95 mL, 7.37 mmol, 2.6 eqv.). The resulting LDA solution was allowed to stir at -78 °C for 10 minutes and then a solution of 5'-TBDMS-2',3'-isopropylidene uridine (1.13 g, 2.84 mmol, 1 eqv.) in anhydrous THF (5 mL) was added. The reaction mixture was stirred for 45 minutes and then a solution of iodine (0.77 g, 3.03 mmol, 1.1 eqv.) in anhydrous THF (5 mL) was added. The reaction mixture was maintained at -78 °C with stirring for a further 4 hours and then was quenched by the addition of AcOH (0.4 mL). EtOAc (50 mL) was then added and the reaction mixture was washed with 1 M Na₂S₂O₃ solution (50 mL), saturated NaHCO₃ solution (50 mL) and brine (50 mL). The organic layer was dried over Na₂SO₄ and all solvent was removed under reduced pressure to leave a crude residue that was purified by

flash column chromatography (100% Hex – 70% Hex/30% EtOAc). The desired product was isolated as a yellow solid (0.66 g, 44%). $R_f = 0.26$ (70% Hex/30% EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3): δ (ppm) 9.34 (1H, br s, NH), 6.45 (1H, s, H5), 6.09 (1H, d, J 1.2, H1'), 5.18 (1H, dd, J 6.4 and 1.2, H2'), 4.81 (1H, dd, J 6.0 and 4.4, H3'), 4.19-4.15 (1H, m, H4'), 3.83-3.75 (2H, m, H5' and H5''), 1.55 (3H, s, CH_3), 1.34 (3H, s, CH_3), 0.88 (9H, s, 3 x CH_3), 0.05 (3H, s, SiCH_3), 0.05 (3H, s, SiCH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3): δ (ppm) 161.44 (CO), 147.19 (CO), 116.93 (C5), 114.00 (quat C), 113.78 (quat C), 102.08 (C1'), 89.96 (C4'), 84.54 (C2'), 82.16 (C3'), 64.12 (C5'), 27.37 (CH_3), 26.09 (3 x CH_3), 25.53 (CH_3), 18.61 (quat C), -5.07 (SiCH_3), -5.09 (SiCH_3). HRMS (ES^+) (m/z): 547.0734 ($[\text{M}+\text{Na}]^+$); $\text{C}_{18}\text{H}_{29}\text{IN}_2\text{O}_6\text{SiNa}$ requires 547.0732 (+0.3656 ppm).

6-Iodouridine 29¹⁶⁴

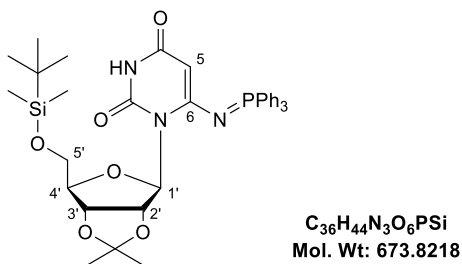


6-Iodo-5'-TBDMS-2',3'-isopropylidene uridine (0.10 g, 0.19 mmol) was suspended in distilled H_2O (2 mL) and cooled to 0 °C. A 50% aqueous TFA solution (3 mL) was added dropwise and the reaction mixture was allowed to warm to room temperature and left to stir in the dark for 45 minutes. The reaction mixture was concentrated *in vacuo* and purified by flash column chromatography (10% EtOH/90% CHCl_3 – 15% EtOH/85% CHCl_3) to yield the desired product as a white solid (53 mg, 75%). $^1\text{H NMR}$ (400 MHz, CD_3OD): δ (ppm) 6.43 (1H, s, H5), 5.96 (1H, d, J 3.6, H1'), 4.71 (1H, dd, J 6.4 and 3.2, H2'), 4.33 (1H, app t, J 6.6, H3'), 3.88 (1H, td, J 6.3 and 3.1, H4'), 3.80 (1H, dd, J 12.0 and 2.8, H5'), 3.67 (1H, dd, J 12.0 and 6.0, H5''). $^{13}\text{C NMR}$ (100 MHz, CD_3OD): δ (ppm) 163.94 (CO), 149.00 (CO), 117.50 (C6), 117.26 (C5), 104.04 (C1'), 86.29 (C4'), 73.39 (C2'), 71.28 (C3'), 63.71 (C5'). HRMS (ES^+) (m/z): 392.9545 ($[\text{M}+\text{Na}]^+$); $\text{C}_9\text{H}_{11}\text{IN}_2\text{O}_6\text{Na}$ requires 392.9554 (-2.2903 ppm).

6-Azido-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 33²⁰³

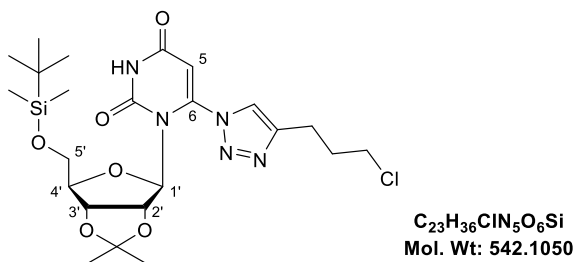
6-Iodo-5'-TBDMS-2',3'-isopropylidene uridine (0.20 g, 0.38 mmol, 1 eqv.) was dissolved in anhydrous DMF (3 mL). Sodium azide (0.03 g, 0.46 mmol, 1.2 eqv.) was then added and the reaction mixture was stirred in the dark for 1 hour at room temperature, after which time TLC showed complete disappearance of starting material. All solvent was then removed *in vacuo* and the resulting crude residue dissolved in EtOAc (30 mL). The crude was washed with cold H₂O (3 x 30 mL) and brine (3 x 30 mL) and the organic layer was dried over Na₂SO₄. All solvent was removed under reduced pressure and the crude material was purified by flash column chromatography (70% Hex/30% EtOAc – 60% Hex/40% EtOAc) to give the desired compound as an off-white solid (0.14 g, 84%). $R_f = 0.27$ (60% Hex/40% EtOAc). IR ν_{max}/cm^{-1} 2136 (N₃). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.41 (1H, br s, NH), 6.09 (1H, d, J 1.2, H1'), 5.50 (1H, s, H5), 5.14 (1H, dd, J 6.6 and 1.4, H2'), 4.80 (1H, dd, J 6.4 and 4.8, H3'), 4.13-4.09 (1H, m, H4'), 3.83-3.75 (2H, m, H5' and H5''), 1.54 (3H, s, CH₃), 1.33 (3H, s, CH₃), 0.88 (9H, s, 3 x CH₃), 0.05 (6H, s, 2 x SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 162.01 (quat C), 152.35 (quat C), 149.30 (quat C), 114.12 (quat C), 89.96 (C1'), 89.40 (C4'), 88.69 (C5), 84.28 (C2'), 81.88 (C3'), 64.09 (C5'), 27.38 (CH₃), 26.08 (3 x CH₃), 25.54 (CH₃), 18.61 (quat C), -5.10 (SiCH₃), -5.13 (SiCH₃). HRMS (ES⁺) (m/z): 462.1784 ([M+Na]⁺); C₁₈H₂₉N₅O₆SiNa requires 462.1779 (+1.0818 ppm).

6-((Triphenyl- λ^5 -phosphanylidene)amino)-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine **34**¹⁷⁵

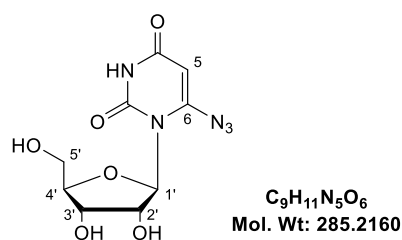


To a solution of 6-azido-5'-TBDMS-2',3'-isopropylidene uridine (0.15 g, 0.34 mmol, 1 eqv.) in anhydrous DCM (2 mL) was added a solution of triphenylphosphine (89 mg, 0.34 mmol, 1 eqv.) in anhydrous DCM (2 mL). The resulting bright yellow reaction mixture was stirred for 1 hour and then concentrated *in vacuo*. The crude material was purified by flash column chromatography (40% Pet. Ether/60% EtOAc – 100% EtOAc) to yield the title compound as a white solid (17 mg, 7%). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 8.07 (1H, br s, NH), 7.70 (6H, app dd, *J* 12.6 and 7.4, Ar-H), 7.64-7.60 (3H, m, Ar-H), 7.51 (6H, td, *J* 7.6 and 3.2, Ar-H), 5.27 (1H, app d, *J* 6.8, H2'), 4.85 (1H, dd, *J* 6.2 and 4.2, H3'), 4.43 (1H, s, H5), 4.21-4.17 (1H, m, H4'), 3.87-3.85 (2H, m, H5' and H5''), 1.58 (3H, s, CH₃), 1.37 (3H, s, CH₃), 0.88 (9H, s, 3 x CH₃), 0.05 (6H, s, 2 x SiCH₃). Signal not observed: H1' (correlations present but signal hidden under CDCl₃ peak). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 163.55 (CO), 158.37 (d, *J* 12.2, 3 x Ar-C), 151.21 (CO), 133.20 (d, *J* 2.8, 3 x Ar-C), 132.84 (d, *J* 10.3, 6 x Ar-C), 129.34 (d, *J* 12.6, 6 x Ar-C), 126.95 (d, *J* 102.6, C6), 113.12 (quat C), 89.33 (d, *J* 1.5, C1'), 89.26 (C4'), 84.97 (C2'), 84.46 (d, *J* 8.1, C5), 82.94 (C3'), 64.85 (C5'), 27.46 (CH₃), 26.17 (3 x CH₃), 25.87 (CH₃), 18.68 (quat C), -4.97 (SiCH₃), -5.04 (SiCH₃). ³¹P NMR (162 MHz, CDCl₃): δ (ppm) 14.39. HRMS (ES⁺) (*m/z*): 696.2604 ([M+Na]⁺); C₃₆H₄₄N₃O₆PSiNa requires 696.2629 (-3.5906 ppm).

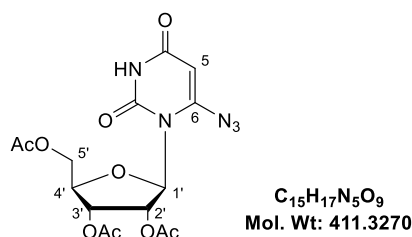
6-(4-(3-Chloropropyl)-1*H*-1,2,3-triazol-1-yl)-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 36¹⁸⁰



6-Azido-5'-TBDMS-2',3'-isopropylidene uridine (0.2 g, 0.46 mmol, 1 eqv.) was dissolved in a mixture of DCM (0.7 mL) and distilled H₂O (0.7 mL). To this solution was added 5-chloro-1-pentyne (43 μ L, 0.41 mmol, 0.9 eqv.) followed by CuSO₄·5H₂O (6 mg, 24.0 μ mol, 5 mol%) and sodium ascorbate (13 mg, 65.6 μ mol, 13 mol%). The resulting reaction mixture was stirred vigorously at room temperature overnight. The reaction mixture was then diluted with more DCM (10 mL) and washed with water (10 mL). The organic extract was dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (80% Hex/20% EtOAc – 60% Hex/40% EtOAc) to yield the desired product as a white solid (0.16g, 72%). *R*_f = 0.24 (60% Hex/40% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.39 (1H, br s, NH), 7.85 (1H, s, NCHC), 5.99 (1H, s, H5), 5.24 (1H, app d, *J* 6.4, H2'), 5.10 (1H, app s, H1'), 4.79 (1H, dd, *J* 6.4 and 4.4, H3'), 4.11-4.07 (1H, m, H4'), 3.83-3.81 (2H, m, H5' and H5''), 3.62 (2H, t, *J* 6.2, CH₂CH₂CH₂Cl), 3.01 (2H, t, *J* 7.4, CH₂CH₂CH₂Cl), 2.25 (2H, m, CH₂CH₂CH₂Cl), 1.43 (3H, s, CH₃), 1.31 (3H, s, CH₃), 0.89 (9H, s, 3 x CH₃), 0.07 (3H, s, SiCH₃), 0.06 (3H, s, SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 161.42 (quat C), 149.11 (quat C), 147.91 (quat C), 146.17 (quat C), 123.34 (NCHC), 114.13 (quat C), 100.85 (C5), 91.96 (C1'), 89.97 (C4'), 84.14 (C2'), 81.85 (C3'), 64.03 (C5'), 43.97 (CH₂CH₂CH₂Cl), 31.42 (CH₂CH₂CH₂Cl), 27.22 (CH₃), 26.05 (3 x CH₃), 25.42 (CH₃), 22.63 (CH₂CH₂CH₂Cl), 18.58 (quat C), -5.07 (SiCH₃), -5.09 (SiCH₃). HRMS (ES⁺) (*m/z*): 564.2015 ([M+Na]⁺); C₂₃H₃₆ClN₅O₆SiNa requires 564.2016 (-0.1772 ppm).

6-Azidouridine 38²⁰³

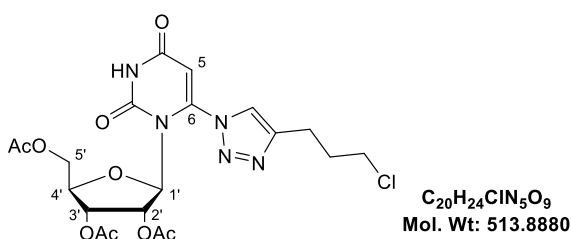
6-Azido-5'-TBDMS-2',3'-isopropylidene uridine (0.88 g, 2.00 mmol) was suspended in distilled H₂O (6 mL) and cooled to 0 °C. A 50% aqueous TFA solution (3 mL) was added dropwise and the reaction mixture was allowed to warm to room temperature and left to stir in the dark for 45 minutes. The reaction mixture was concentrated *in vacuo* and purified by flash column chromatography (5% EtOH/95% CHCl₃ – 20% EtOH/80% CHCl₃) to yield the desired product as an off-white solid (0.47 g, 82%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 5.96 (1H, d, *J* 3.6, H1'), 5.58 (1H, s, H5), 4.64 (1H, dd, *J* 6.4 and 3.6, H2'), 4.33 (1H, app t, *J* 6.4, H3'), 3.85-3.79 (2H, m, H4' and H5'), 3.69-3.65 (1H, m, H5''). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 164.57 (quat C), 154.55 (quat C), 151.32 (quat C), 92.00 (C1'), 89.62 (C5), 85.96 (C4'), 73.22 (C2'), 70.93 (C3'), 63.36 (C5'). HRMS (ES⁺) (*m/z*): 308.0603 ([M+Na]⁺); C₉H₁₁N₅O₆Na requires 308.0602 (+0.3246 ppm).

6-Azido-2',3',5'-tri-*O*-acetyl uridine 39

6-Azidouridine (0.47 g, 1.65 mmol, 1 eqv.) and DMAP (20 mg, 0.16 mmol, 0.1 eqv.) were added to anhydrous MeCN (5 mL) and cooled to 0 °C. Et₃N (1.37 mL, 9.83 mmol, 6 eqv.) was added followed by the dropwise addition of acetic anhydride (0.46 mL, 4.88 mmol, 3 eqv.). The reaction mixture was then warmed to room temperature and stirred for 3 hours. Once TLC indicated there was no remaining starting material, MeOH (5 mL) was added and all solvent was then removed under reduced pressure. The crude residue was purified by flash column chromatography (100% DCM – 2% MeOH/98% DCM) to yield the desired product as a white solid (0.47 g, 69%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 6.01 (1H, d, *J* 2.4, H1'), 5.73

(1H, dd, J 6.8 and 2.8, H2'), 5.59 (1H, s, H5), 5.55 (1H, app t, J 7.2, H3'), 4.45-4.40 (1H, m, H5'), 4.23-4.17 (2H, m, H4' and H5''), 2.10 (3H, s, COCH₃), 2.08 (3H, s, COCH₃), 2.05 (3H, s, COCH₃). ¹³C NMR (100MHz, CD₃OD): δ (ppm) 172.45 (CO), 171.71 (CO), 171.34 (CO), 164.37 (quat C), 153.66 (quat C), 150.93 (quat C), 89.71 (C5), 89.62 (C1'), 80.29 (C4'), 74.46 (C2'), 71.06 (C3'), 63.86 (C5'), 20.60 (COCH₃), 20.40 (COCH₃), 20.31 (COCH₃). HRMS (ES⁺) (m/z): 434.0912 ([M+Na]⁺); C₁₅H₁₇N₅O₉Na requires 434.0918 (-1.3822 ppm).

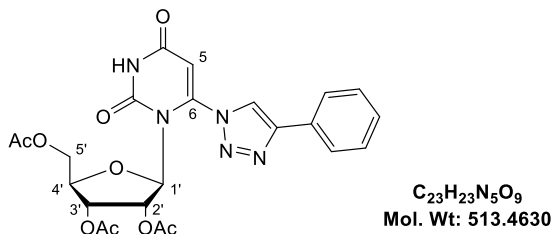
6-(4-(3-Chloropropyl)-1H-1,2,3-triazol-1-yl)-2',3',5'-tri-O-acetyl uridine 40



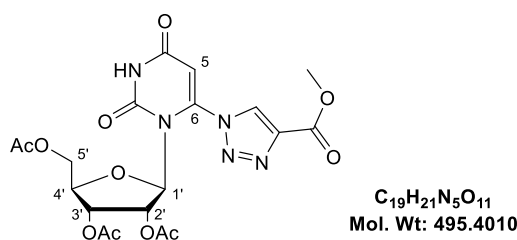
6-Azido-2',3',5'-tri-O-acetyl uridine (0.10 g, 0.24 mmol, 1 eqv.) was dissolved in a mixture of *t*BuOH (1 mL) and distilled H₂O (1 mL). To this solution was added 5-chloro-1-pentyne (28 μ L, 0.28 mmol, 1.2 eqv.), a 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and a 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%). The resulting reaction mixture was stirred vigorously at room temperature overnight. TLC showed starting material to still be present so more 5-chloro-1-pentyne (28 μ L, 0.25 mmol), 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%) were added. The reaction was left to stir for 5 days and then diluted with water (20 mL) and extracted with EtOAc (3 x 20 mL). The organic extract was dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue obtained was purified by flash column chromatography (90% Hex/10% EtOAc – 40% Hex/60% EtOAc) to yield the desired product as a white solid (52 mg, 42%). R_f = 0.55 (20% Hex/80% EtOAc) ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.37 (1H, br s, NH), 7.74 (1H, s, NCHC), 5.83-5.80 (2H, m, H5 and H2'), 5.49 (1H, app t, J 7.4, H3'), 4.90 (1H, d, J 2.4, H1'), 4.45 (1H, dd, J 11.8 and 3.0, H5'), 4.21 (1H, dd, J 11.8 and 6.6, H5''), 4.17-4.12 (1H, m, H4'), 3.62 (2H, t, J 6.2, CH₂CH₂CH₂Cl), 2.99 (2H, t, J 7.4, CH₂CH₂CH₂Cl), 2.27-2.20 (2H, m, CH₂CH₂CH₂Cl), 2.08 (3H, s, COCH₃), 2.06 (3H, s, COCH₃), 2.03 (3H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 170.84 (CO), 170.17 (CO), 169.57 (CO), 161.10 (quat C), 148.78 (quat C), 147.58 (quat C), 145.45 (quat C), 123.80 (NCHC), 101.21 (C5), 90.99 (C1'), 79.95 (C4'), 73.42 (C2'),

70.11 (C3'), 63.22 (C5'), 43.96 (CH₂CH₂CH₂Cl), 31.47 (CH₂CH₂CH₂Cl), 22.57 (CH₂CH₂CH₂Cl), 20.88 (COCH₃), 20.61 (COCH₃), 20.46 (COCH₃). HRMS (ES⁺) (m/z): 514.1332 ([M+H]⁺); C₂₀H₂₅ClN₅O₉ requires 514.1335 (-0.5835 ppm).

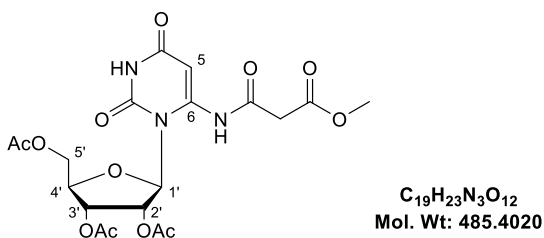
6-(4-Phenyl-1*H*-1,2,3-triazol-1-yl)-2',3',5'-tri-*O*-acetyl uridine 41



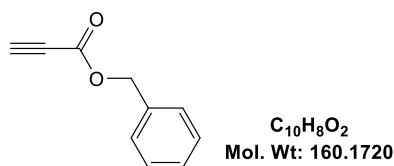
6-Azido-2',3',5'-tri-*O*-acetyl uridine (0.1 g, 0.24 mmol, 1 eqv.) was dissolved in a mixture of *t*BuOH (0.5 mL) and distilled H₂O (0.5 mL). To this solution was added phenylacetylene (29 μ L, 0.26 mmol, 1.1 eqv.), a 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and a 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%). The reaction was stirred at room temperature overnight. Further phenylacetylene (29 μ L, 0.26 mmol, 1.1 eqv.), 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%) were added and the reaction was stirred for another 3 days. All solvent was then removed *in vacuo* and the crude material was purified by flash column chromatography (60% Hex/40% EtOAc – 50% Hex/50% EtOAc). The desired product was isolated as a white solid (64 mg, 52%). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.34 (1H, br s, NH), 8.16 (1H, s, NCHC), 7.90-7.88 (2H, m, Ar-H), 7.51-7.42 (3H, m, Ar-H), 5.92 (1H, d, *J* 1.6, H5), 5.85 (1H, dd, *J* 6.8 and 2.4, H2'), 5.51 (1H, dd, *J* 7.6 and 7.2, H3'), 4.94 (1H, d, *J* 2.4, H1'), 4.46 (1H, dd, *J* 12.0 and 3.2, H5'), 4.21 (1H, dd, *J* 12.0 and 6.4, H5''), 4.16-4.13 (1H, m, H4'), 2.09 (3H, s, COCH₃), 2.05 (3H, s, COCH₃), 1.99 (3H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 170.84 (CO), 170.32 (CO), 169.60 (CO), 161.05 (quat C), 148.78 (quat C), 148.71 (quat C), 145.40 (quat C), 129.50 (Ar-C), 129.34 (2 x Ar-C), 128.77 (quat C), 126.13 (2 x Ar-C), 121.95 (NCHC), 101.49 (C5), 91.17 (C1'), 79.93 (C4'), 73.49 (C2'), 70.06 (C3'), 63.17 (C5'), 20.88 (COCH₃), 20.61 (COCH₃), 20.46 (COCH₃). HRMS (ES⁺) (m/z): 536.1395 ([M+Na]⁺); C₂₃H₂₃N₅O₉Na requires 536.1388 (+1.3056 ppm).

6-Methoxy carbonyl triazole- 2',3',5'-tri-O-acetyl uridine 42

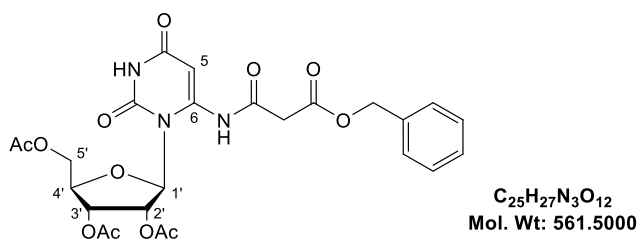
6-Azido-2',3',5'-tri-O-acetyl uridine (160 mg, 0.39 mmol, 1 eqv.) was dissolved in a mixture of *t*BuOH (1 mL) and distilled H₂O (1 mL). Methyl propiolate (35 μ L, 0.39 mmol, 1 eqv.), a 0.1 M aqueous CuSO₄·5H₂O solution (192 μ L, 5 mol%) and a 0.5 M aqueous sodium ascorbate solution (100 μ L, 13 mol%) were added and the reaction was stirred overnight. TLC showed the reaction to be progressing so more methyl propiolate (35 μ L, 0.39 mmol, 1 eqv.), 0.1 M aqueous CuSO₄·5H₂O solution (192 μ L, 5 mol%) and 0.5 M aqueous sodium ascorbate solution (100 μ L, 13 mol%) were added and the reaction was stirred overnight once more. All solvent was then removed *in vacuo*. The crude residue was dissolved in distilled H₂O (30 mL) and extracted with EtOAc (3 x 30 mL). The combined organic extracts were washed with distilled H₂O (2 x 20 mL) and dried over Na₂SO₄. The crude material was purified by flash column chromatography (80% Hex/20% EtOAc – 40% Hex/60% EtOAc) to yield the desired product as an off-white solid (49 mg, 25%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 8.98 (1H, s, NCHC), 6.07 (1H, s, H5), 5.85 (1H, dd, *J* 6.8 and 2.8, H2'), 5.54 (1H, app t, *J* 7.4, H3'), 4.89 (1H, d, *J* 2.4, H1'), 4.43 (1H, dd, *J* 11.8 and 3.0, H5'), 4.19-4.09 (2H, m, H4' and H5''), 3.97 (3H, s, OCH₃), 2.06 (3H, s, COCH₃), 2.04 (3H, s, COCH₃), 2.03 (3H, s, COCH₃). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 172.50 (CO), 171.68 (CO), 171.24 (CO), 163.49 (quat C), 161.57 (quat C), 150.70 (quat C), 145.90 (quat C), 141.23 (quat C), 132.58 (NCHC), 103.22 (C5), 91.82 (C1'), 80.76 (C4'), 74.43 (C2'), 70.89 (C3'), 63.64 (C5'), 52.94 (OCH₃), 20.63 (COCH₃), 20.39 (COCH₃), 20.25 (COCH₃). HRMS (ES⁺) (*m/z*): 518.1138 ([M+Na]⁺); C₁₉H₂₁N₅O₁₁Na requires 518.1130 (+1.5541 ppm).

6-(3-(Methoxy)-3-oxopropanamido)-2',3',5'-tri-O-acetyl uridine 43

6-Azido-2',3',5'-tri-O-acetyl uridine (0.1 g, 0.24 mmol, 1 eqv.) was dissolved in a mixture of *t*BuOH (0.5 mL) and distilled H₂O (0.5 mL). To this solution was added methyl propiolate (24 μ L, 0.27 mmol, 1.1 eqv.), a 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and a 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%). The reaction was stirred at room temperature overnight. TLC showed only starting material so another equivalent of methyl propiolate (24 μ L, 0.27 mmol, 1.1 eqv.), 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%) was added. The reaction was stirred for 2 days and monitored by TLC. More methyl propiolate (0.21 mL, 2.40 mmol, 10 eqv.), 0.1 M aqueous CuSO₄·5H₂O solution (122 μ L, 5 mol%) and 0.5 M aqueous sodium ascorbate solution (49 μ L, 10 mol%) were added and the reaction mixture was stirred for a further 2 days. All solvent was then removed *in vacuo* and the crude material was purified by flash column chromatography (70% Hex/30% EtOAc – 50% Hex/50% EtOAc). The title compound was isolated as an off-white solid (40 mg, 34%). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.57 (1H, br s, NH), 9.25 (1H, br s, NH), 6.28 (1H, s, H5), 6.01 (1H, d, *J* 4.0, H1'), 5.58 (1H, dd, *J* 7.0 and 4.2, H2'), 5.45 (1H, app t, *J* 7.2, H3'), 4.45 (1H, dd, *J* 12.2 and 7.8, H5'), 4.39 (1H, dd, *J* 12.2 and 3.4, H5''), 4.24 (1H, td, *J* 7.5 and 3.3, H4'), 3.78 (3H, s, OCH₃), 3.66-3.46 (2H, ABq, *J* 17.2, COCH₂CO), 2.11 (3H, s, COCH₃), 2.10 (3H, s, COCH₃), 2.07 (3H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 171.44 (quat C), 171.14 (quat C), 169.78 (quat C), 169.45 (quat C), 163.30 (quat C), 163.19 (quat C), 149.98 (quat C), 145.43 (quat C), 95.58 (C5), 88.51 (C1'), 80.01 (C4'), 73.22 (C2'), 69.78 (C3'), 63.27 (C5'), 53.15 (OCH₃), 42.42 (COCH₂CO), 20.88 (COCH₃), 20.72 (COCH₃), 20.50 (COCH₃). HRMS (ES⁺) (*m/z*): 508.1195 ([M+Na]⁺); C₁₉H₂₃N₃O₁₂Na requires 508.1174 (+4.1329 ppm).

Benzyl Propiolate

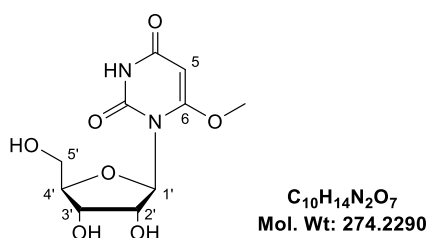
Propiolic acid (0.62 mL, 10.1 mmol, 1 eqv.) was added to acetone (50 mL) followed by potassium carbonate (2.07 g, 15.0 mmol, 1.5 eqv.) and benzyl bromide (1.55 mL, 13.0 mmol, 1.3 eqv.). The resulting reaction mixture was stirred at room temperature overnight. All solvent was removed and the crude material was dissolved in distilled H₂O (30 mL) and extracted with Et₂O (3 x 30 mL). The combined organic extracts were dried over Na₂SO₄ and purified by column chromatography (100% Hex – 90% Hex/10% EtOAc). The title compound was isolated as a clear liquid (0.36 g, 22%). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 7.39-7.35 (5H, m, Ar-H), 5.22 (2H, s, CH₂Ph), 2.89 (1H, s, HCC). ¹³C NMR (400 MHz, CDCl₃): δ (ppm) 152.68 (quat C), 134.66 (quat C), 128.87 (Ar-C), 128.83 (2 x Ar-C), 128.72 (2 x Ar-C), 75.18 (HCC), 74.68 (quat C), 68.06 (CH₂Ph).

6-(3-(Benzyloxy)-3-oxopropanamido)-2',3',5'-tri-O-acetyl uridine 44

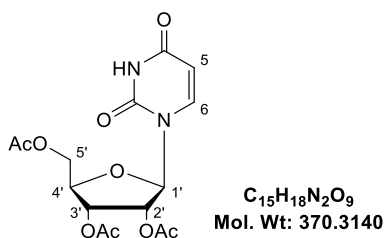
6-Azido-2',3',5'-tri-O-acetyl uridine (0.1 g, 0.24 mmol, 1 eqv.) was dissolved in a mixture of *t*BuOH (0.5 mL) and distilled H₂O (0.5 mL). To this solution was added benzyl propiolate (38 μL, 0.24 mmol, 1 eqv.), a 0.1 M aqueous CuSO₄·5H₂O solution (122 μL, 5 mol%) and a 0.5 M aqueous sodium ascorbate solution (49 μL, 10 mol%). The reaction was stirred at room temperature for 4 days at which point TLC showed no remaining starting material. All solvent was removed *in vacuo* and the crude material was purified by flash column chromatography (90% Hex/10% EtOAc – 50% Hex/50% EtOAc). The title compound was isolated as a white solid (59 mg, 44%). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.41 (1H, br s, NH), 9.22 (1H, br s, NH), 7.40-7.34 (5H, m, Ar-H), 6.32 (1H, d, *J* 1.2, H5), 6.05 (1H, d, *J* 4.4, H1'), 5.58 (1H, dd, *J* 7.2 and 4.4, H2'), 5.43 (1H, app t, *J* 7.2, H3'), 5.24-5.17 (2H, ABq, *J* 12.0,

CH_2Ph), 4.42 (1H, dd, J 12.2 and 8.2, H_5'), 4.31 (1H, dd, J 12.2 and 3.0, H_5''), 4.22 (1H, td, J 7.8 and 2.7, H_4'), 3.72-3.50 (2H, ABq, J 16.8, COCH_2CO), 2.12 (3H, s, COCH_3), 2.08 (3H, s, COCH_3), 2.08 (3H, s, COCH_3). ^{13}C NMR (100 MHz, CDCl_3): δ (ppm) 171.44 (quat C), 171.14 (quat C), 169.78 (quat C), 168.95 (quat C), 163.15 (quat C), 162.97 (quat C), 149.97 (quat C), 145.35 (quat C), 134.61 (quat C), 128.96 (Ar-C), 128.89 (2 x Ar-C), 128.55 (2 x Ar-C), 95.48 (C5), 88.24 (C1'), 80.13 (C4'), 73.11 (C2'), 69.71 (C3'), 68.17 (CH_2Ph), 63.16 (C5'), 42.74 (COCH_2CO), 20.92 (COCH_3), 20.71 (COCH_3), 20.55 (COCH_3). HRMS (ES^+) (m/z): 584.1500 ($[\text{M}+\text{Na}]^+$); $\text{C}_{25}\text{H}_{27}\text{N}_3\text{O}_{12}\text{Na}$ requires 584.1487 (+2.2255 ppm).

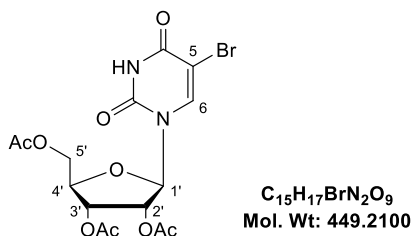
6-Methoxy uridine 46



6-(4-(3-Chloropropyl)-1*H*-1,2,3-triazol-1-yl)-2',3',5'-tri-*O*-acetyl uridine (74 mg, 0.14 mmol) was dissolved in NH_3 (7 M) in MeOH (2 mL) and stirred in a sealed vial at room temperature for 36 hours. All solvent was removed *in vacuo* and the crude material was purified by flash column chromatography (5% MeOH/95% DCM – 10% MeOH/90% DCM). The title product was isolated as a white solid (20 mg, 52%). ^1H NMR (400 MHz, CD_3OD): δ (ppm) 6.11 (1H, d, J 3.2, $\text{H}_{1'}$), 5.15 (1H, s, H_5), 4.57 (1H, dd, J 5.8 and 3.4, $\text{H}_{2'}$), 4.25 (1H, app t, J 6.4, $\text{H}_{3'}$), 3.93 (3H, s, OCH_3), 3.83-3.79 (2H, m, $\text{H}_{4'}$ and $\text{H}_{5'}$), 3.65 (1H, dd, J 12.8 and 6.0, $\text{H}_{5''}$). ^{13}C NMR (100 MHz, CD_3OD): δ (ppm) 166.29 (quat C), 164.55 (quat C), 151.49 (quat C), 90.61 (C1'), 85.51 (C4'), 79.52 (C5), 73.44 (C2'), 70.83 (C3'), 63.21 (C5'), 58.20 (OCH_3). HRMS (ES^-) (m/z): 273.0727 ($[\text{M}-\text{H}]^-$); $\text{C}_{10}\text{H}_{13}\text{N}_2\text{O}_7$ requires 273.0728 (-0.3662 ppm).

2',3',5'-Tri-O-acetyl uridine 47

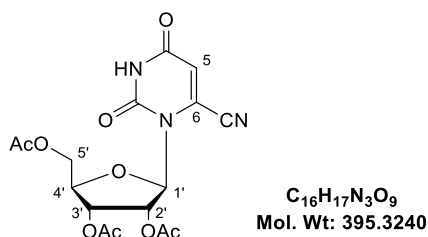
Uridine (3.00 g, 12.3 mmol, 1 eqv.) and DMAP (0.03 g, 0.25 mmol, 0.1 eqv.) were dissolved in anhydrous pyridine (75 mL) and cooled to 0 °C. Acetic anhydride (3.95 mL, 41.9 mmol, 3.4 eqv.) was then added dropwise with stirring, after which the reaction mixture was allowed to warm to room temperature. After 3 hours, TLC showed no remaining starting material. The reaction was quenched with H₂O (30 mL) and then partitioned between CHCl₃ (100 mL) and H₂O (100 mL). The organic layer was washed with 1 M HCl (100 mL) and brine (100 mL) and dried over Na₂SO₄. All solvent was removed *in vacuo* and the resulting residue was co-evaporated twice with toluene. The crude material was purified by flash column chromatography (30% Hex/ 70% EtOAc – 100% EtOAc) to yield the desired product as a white foam (3.91 g, 86%). *R_f* = 0.60 (100% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.15 (1H, br s, NH), 7.40 (1H, d, *J* 8.4, H6), 6.05 (1H, d, *J* 4.4, H1'), 5.80 (1H, dd, *J* 8.2 and 1.8, H5), 5.36-5.30 (2H, m, H2' and H3'), 4.40-4.32 (3H, m, H4', H5' and H5''), 2.15 (3H, s, COCH₃), 2.14 (3H, s, COCH₃), 2.11 (3H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 170.26 (CO), 169.78 (CO), 169.77 (CO), 162.82 (CO), 150.31 (CO), 139.41 (C6), 103.57 (C5), 87.59 (C1'), 80.08 (C4'), 72.85 (C2'), 70.32 (C3'), 63.28 (C5'), 20.91 (COCH₃), 20.64 (COCH₃), 20.55 (COCH₃). HRMS (ES⁺) (*m/z*): 393.0908 ([M+Na]⁺); C₁₅H₁₈N₂O₉Na requires 393.0905 (+0.7632 ppm).

5-Bromo-2',3',5'-tri-O-acetyl uridine 48¹⁸⁵

2',3',5'-Tri-O-acetyl uridine (1.00 g, 2.70 mmol, 1 eqv.) was dissolved in anhydrous DMF (8 mL). 1,3-Dibromo-5,5-dimethylhydantoin (0.42 g, 1.47 mmol, 0.55 eqv.) was added and the resulting clear, yellow reaction mixture was stirred for 1 hour. Cold

H₂O (30 mL) was then added and the reaction mixture was extracted with DCM (3 x 30 mL). The organic extracts were washed with saturated NaHCO₃ solution (30 mL) and brine (30 mL) and dried over Na₂SO₄. All solvent was removed *in vacuo* to leave a thick, yellow oil which was purified by flash column chromatography (50% Hex/50% EtOAc – 45% Hex/55% EtOAc) to yield a white foam (1.05 g, 87%). R_f = 0.36 (40% Hex/60% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.27 (1H, br s, NH), 7.83 (1H, s, H6), 6.10-6.06 (1H, m, H1'), 5.34-5.31 (2H, m, H2' and H3'), 4.41-4.32 (3H, m, H4', H5' and H5''), 2.21 (3H, s, COCH₃), 2.12 (3H, s, COCH₃), 2.10 (3H, s, COCH₃). ¹³C NMR (400 MHz, CDCl₃): δ (ppm) 170.20 (CO), 169.77 (CO), 169.72 (CO), 158.61 (CO), 149.65 (CO), 138.65 (C6), 98.08 (C5), 87.42 (C1'), 80.40 (C4'), 73.23 (C2'), 70.21 (C3'), 63.07 (C5'), 21.04 (COCH₃), 20.63 (COCH₃), 20.52 (COCH₃). HRMS (ES⁺) (m/z): 471.0011 and 473.0004 ([M+Na]⁺); C₁₅H₁₇BrN₂O₉Na requires 471.0010 and 472.9992 (+0.2123 and +2.5370 ppm).

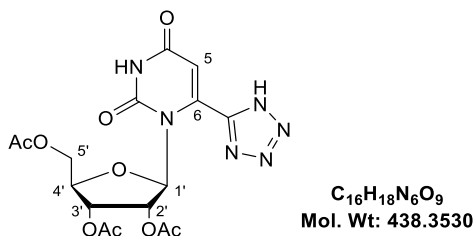
6-Cyano-2',3',5'-tri-O-acetyl uridine 49¹⁷⁹



5-Bromo-2',3',5'-tri-O-acetyl uridine (1.00 g, 2.23 mmol, 1 eqv.) was dissolved in anhydrous DMF (10 mL). NaCN (0.17 g, 3.47 mmol, 1.6 eqv.) was added and the reaction mixture was stirred overnight at room temperature. Cold water (30 mL) was added and then the product was extracted with DCM (3 x 30 mL). The organic extracts were washed with brine (3 x 30 mL) and then dried over Na₂SO₄. All solvent was removed *in vacuo* and the crude material was purified by flash column chromatography (70% Pet. Ether/30% EtOAc – 50% Pet. Ether/50% EtOAc) to yield the desired product as a white solid (0.46 g, 52%). R_f = 0.19 (60% Pet. Ether/40% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.46 (1H, br s, NH), 6.31 (1H, s, H5), 5.82 (1H, d, *J* 3.6, H1'), 5.64 (1H, dd, *J* 7.0 and 3.4, H2'), 5.42 (1H, app t, *J* 7.4, H3'), 4.51 (1H, dd, *J* 11.8 and 2.6, H5'), 4.29 (1H, td, *J* 7.1 and 2.5, H4'), 4.23 (1H, dd, *J* 12.0 and 6.4, H5''), 2.12 (6H, s, 2 x COCH₃), 2.09 (3H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 170.81 (CO), 170.49 (CO), 169.70 (CO), 160.22 (quat C), 148.37 (quat C), 127.45 (quat C), 113.38 (C5), 110.70 (quat C), 92.91 (C1'), 79.92 (C4'), 73.06 (C2'), 69.75 (C3'), 62.88 (C5'), 20.83 (COCH₃), 20.56 (COCH₃),

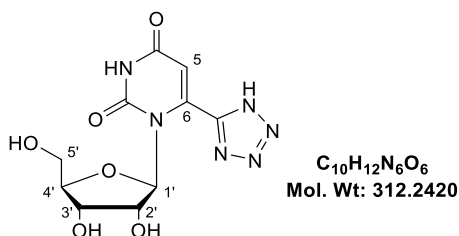
20.45 (COCH₃). HRMS (ES⁺) (m/z): 418.0851 ([M+Na]⁺); C₁₆H₁₇N₃O₉Na requires 418.0857 (-1.4351 ppm).

6-(1*H*-Tetrazol-5-yl)-2',3',5'-tri-*O*-acetyl uridine 50¹⁷¹



6-Cyano-2',3',5'-tri-*O*-acetyl uridine (0.20 g, 0.51 mmol, 1 eqv.), NaN₃ (40 mg, 0.62 mmol, 1.2 eqv.) and NH₄Cl (33 mg, 0.62 mmol, 1.2 eqv.) were dissolved in anhydrous DMF (3 mL). The reaction mixture was heated to 95 °C for 5 hours then stirred at room temperature overnight. All solvent was then removed under reduced pressure and the material obtained carried through to the next step without further purification. ¹H NMR (400 MHz, CD₃OD): δ (ppm) 6.07 (1H, d, *J* 2.4, H1'), 6.01 (1H, s, H5), 5.93 (1H, dd, *J* 6.6 and 2.6, H2'), 5.59 (1H, app t, *J* 7.4, H3'), 4.42 (1H, dd, *J* 11.6 and 3.2, H5'), 4.17 (1H, dd, *J* 11.8 and 5.8, H5''), 4.13-4.09 (1H, m, H4'), 2.06 (3H, s, COCH₃), 2.05 (3H, s, COCH₃), 2.04 (3H, s, COCH₃). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 172.61 (CO), 171.64 (CO), 171.39 (CO), 164.95 (quat C), 156.82 (quat C), 152.01 (quat C), 148.16 (quat C), 104.94 (C5), 93.03 (C1'), 80.25 (C4'), 74.34 (C2'), 71.26 (C3'), 64.02 (C5'), 20.66 (COCH₃), 20.48 (COCH₃), 20.35 (COCH₃). HRMS (ES⁺) (m/z): 461.1034 ([M+Na]⁺); C₁₆H₁₈N₆O₉Na requires 461.1027 (+1.5181 ppm).

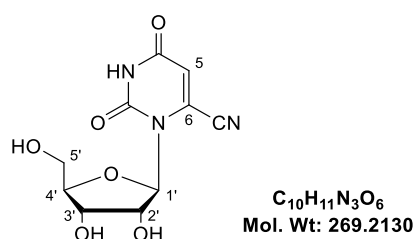
6-(1*H*-Tetrazol-5-yl) uridine 51



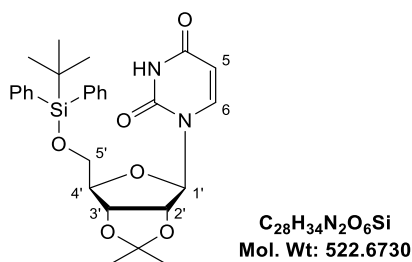
6-(1*H*-Tetrazol-5-yl)-2',3',5'-tri-*O*-acetyl uridine (carried through crude) was dissolved in NH₃ (7 M) in MeOH (2 mL) and stirred in a sealed vial for 5 hours, over which time a white precipitate formed. The reaction mixture was then filtered and

the white solid obtained was triturated several times with Et₂O and dried under vacuum. The desired product was isolated as a white solid (62 mg, 39% over 2 steps). ¹H NMR (400 MHz, D₂O): δ (ppm) 6.03 (1H, s, H5), 5.43 (1H, d, *J* 3.2, H1'), 4.73 (1H, dd, *J* 6.4 and 3.2, H2'), 4.23 (1H, app t, *J* 7.0 H3'), 3.82-3.76 (2H, m, H4' and H5'), 3.63 (1H, dd, *J* 11.8 and 6.4, H5''). ¹³C NMR (100 MHz, D₂O): δ (ppm) 164.93 (quat C), 151.39 (quat C), 147.10 (quat C), 105.15 (C5), 93.68 (C1'), 83.15 (C4'), 71.65 (C2'), 69.15 (C3'), 61.49 (C5'). Signal not observed: 1 x quat C. HRMS (ES⁻) (m/z): 311.0744 ([M-H]⁻); C₁₀H₁₁N₆O₆ requires 311.0746 (-0.6429 ppm).

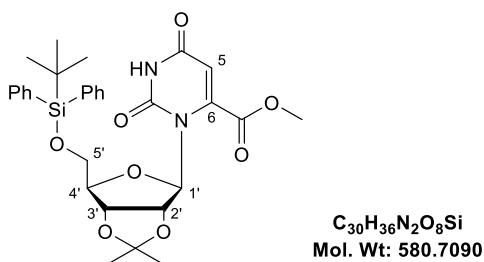
6-Cyanouridine 52



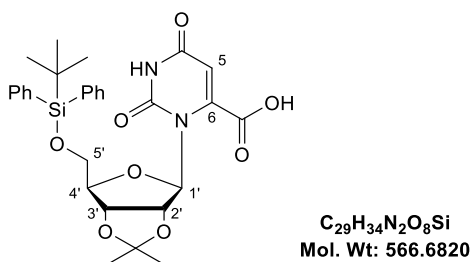
6-Cyano-2',3',5'-tri-O-acetyl uridine (0.20 g, 0.51 mmol) was dissolved in NH₃ (7 M) in MeOH (4 mL) and stirred in a sealed vial for 36 hours. All solvent was then removed *in vacuo* and the crude residue was purified by flash column chromatography (100% CHCl₃ – 10% EtOH/90% CHCl₃) to yield the desired product as a thick, pale yellow oil (0.10 g, 73%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 6.47 (1H, s, H5), 5.88 (1H, d, *J* 4.4, H1'), 4.65 (1H, dd, *J* 6.6 and 4.6, H2'), 4.26 (1H, app t, *J* 6.2, H3'), 3.93 (1H, td, *J* 6.2 and 3.6, H4'), 3.82 (1H, dd, *J* 12.0 and 3.6, H5'), 3.74 (1H, dd, *J* 11.4 and 6.0, H5''). ¹³C NMR (CD₃OD): δ (ppm) 162.92 (quat C), 150.73 (quat C), 128.80 (quat C), 114.59 (C5), 112.52 (quat C), 95.19 (C1'), 86.67 (C4'), 73.12 (C2'), 70.94 (C3'), 63.28 (C5'). HRMS (ES⁻) (m/z): 268.0578 ([M-H]⁻); C₁₀H₁₀N₃O₆ requires 268.0575 (+1.1192 ppm).

5'-*tert*-Butyldiphenylsilyl-2',3'-isopropylidene uridine 53

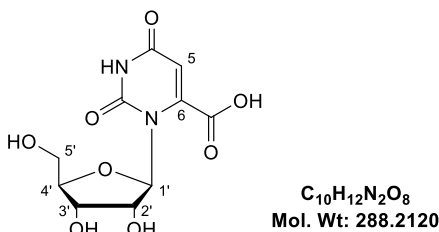
2',3'-Isopropylidene uridine (1.95 g, 6.86 mmol, 1 eqv.) was added to anhydrous DCM (20 mL) and cooled to 0 °C. Imidazole (0.96 g, 14.1 mmol, 2 eqv.) and TBDPSCI (2.7 mL, 10.4 mmol, 1.5 eqv.) were added and the reaction mixture was warmed to room temperature. After 2 hours, TLC indicated complete absence of starting material. All solvent was removed *in vacuo* to leave a crude residue which was re-dissolved in EtOAc (60 mL) and washed with H₂O (40 mL) and brine (40 mL). The organic layer was dried over Na₂SO₄ and all solvent removed under reduced pressure. The crude material was purified by flash column chromatography (80% Pet. Ether/20% EtOAc – 60% Pet. Ether/40% EtOAc) to yield the desired product as a white foam (3.03 g, 85%). *R_f* = 0.44 (5% MeOH/95% CHCl₃). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 8.79 (1H, br s, NH), 7.65-7.61 (4H, m, Ar-H), 7.57 (1H, d, *J* 8.4, H6), 7.47-7.36 (6H, m, Ar-H), 5.98 (1H, d, *J* 2.8, H1'), 5.43 (1H, d, *J* 8.4, H5), 4.82 (1H, dd, *J* 6.4 and 3.2, H3'), 4.73 (1H, dd, *J* 6.4 and 2.8, H2'), 4.28-4.26 (1H, m, H4'), 4.00 (1H, dd, *J* 11.6 and 2.4, H5'), 3.85 (1H, dd, *J* 12.0 and 3.6, H5''), 1.58 (3H, s, CH₃), 1.35 (3H, s, CH₃), 1.09 (9H, s, 3 x CH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 163.08 (CO), 150.10 (CO), 140.68 (C6), 135.75 (2 x Ar-C), 135.52 (2 x Ar-C), 132.97 (Ar-C), 132.45 (Ar-C), 130.30 (Ar-C), 130.23 (Ar-C), 128.11 (2 x Ar-C), 128.07 (2 x Ar-C), 114.55 (quat C), 102.60 (C5), 91.72 (C1'), 86.51 (C4'), 85.12 (C2'), 80.26 (C3'), 64.07 (C5'), 27.41 (CH₃), 27.10 (3 x CH₃), 25.53 (CH₃), 19.45 (quat C). HRMS (ES⁺) (*m/z*): 545.2079 ([M+Na]⁺); C₂₈H₃₄N₂O₆SiNa requires 545.2078 (+0.1834 ppm).

6-Methoxy carbonyl-5'-*tert*-butyldiphenylsilyl-2',3'-isopropylidene uridine 54

Anhydrous diisopropylamine (0.31 mL, 2.21 mmol, 2.3 eqv.) was dissolved in anhydrous THF (6 mL) and cooled to 0 °C. 2.5 M *n*-BuLi (0.88 mL, 2.20 mmol, 2.3 eqv.) was added dropwise and the resulting LDA solution was stirred for 20 minutes at 0 °C. The LDA solution was then cooled to -78 °C and a solution of 5'-TBDPS-2',3'-isopropylidene (0.50 g, 0.96 mmol, 1 eqv.) in anhydrous THF (6 mL) was added to it. The reaction mixture was stirred at -78 °C for 1 hour then a solution of methyl chloroformate (0.09 mL, 1.16 mmol, 1.2 eqv.) in anhydrous THF (3 mL) was added. The reaction mixture was maintained at -78 °C for 4 hours and then quenched by the addition of AcOH (0.3 mL). The reaction was warmed to room temperature and then diluted with EtOAc (40 mL) and washed with saturated NaHCO₃ solution (20 mL) and brine (20 mL). The organic layer was dried over Na₂SO₄ and all solvent was removed *in vacuo*. The crude product was purified by flash column chromatography (5% EtOAc/95% Toluene – 30% EtOAc/70% Toluene) to yield the desired product as a white solid (0.16 g, 29%). *R*_f = 0.37 (30% EtOAc/70% Toluene). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 8.25 (1H, br s, NH), 7.68-7.64 (4H, m, Ar-H), 7.44-7.31 (6H, m, Ar-H), 6.06 (1H, s, H5), 5.91 (1H, app s, H1'), 5.12 (1H, app d, *J* 6.8, H2'), 4.70 (1H, app t, *J* 5.6, H3'), 4.21-4.17 (1H, m, H4'), 3.84 (1H, dd, *J* 10.8 and 5.2, H5'), 3.81 (3H, s, OCH₃), 3.77 (1H, dd, *J* 10.8 and 7.6, H5''), 1.55 (3H, s, CH₃), 1.33 (3H, s, CH₃), 1.04 (9H, s, 3 x CH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 162.08 (CO), 161.44 (CO), 149.59 (CO), 145.11 (C6), 135.83 (2 x Ar-C), 135.79 (2 x Ar-C), 133.56 (Ar-C), 133.49 (Ar-C), 129.85 (Ar-C), 129.80 (Ar-C), 127.82 (2 x Ar-C), 127.70 (2 x Ar-C), 114.60 (quat C), 106.13 (C5), 93.27 (C1'), 88.63 (C4'), 84.82 (C2'), 81.67 (C3'), 64.69 (C5'), 53.92 (OCH₃), 27.39 (CH₃), 26.93 (3 x CH₃), 25.49 (CH₃), 19.36 (quat C). HRMS (ES⁺) (*m/z*): 603.2134 ([M+Na]⁺); C₃₀H₃₆N₂O₈SiNa requires 603.2133 (+0.1658 ppm).

6-Carboxy-5'-tert-butyl-diphenylsilyl-2',3'-isopropylidene uridine 55

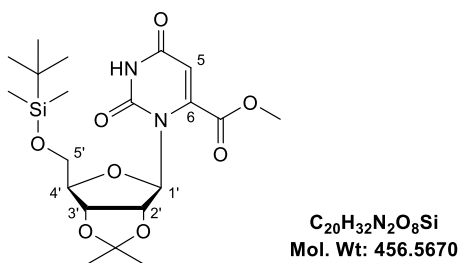
6-Methoxy carbonyl-5'-TBDPS-2',3'-isopropylidene uridine (0.16 g, 0.28 mmol, 1 eqv.) was dissolved in a mixture of THF (4 mL) and distilled H₂O (4 mL). LiOH (16 mg, 0.67 mmol, 2.4 eqv.) was added and the reaction mixture was stirred for 2 hours at which point TLC indicated no starting material remained. AcOH (0.5 mL) was added to quench and neutralise the reaction and all solvent was removed *in vacuo*. The crude residue was purified by flash column chromatography (3% MeOH/97% DCM – 12% MeOH/88% DCM) to give the desired product as a white solid (0.13 g, 82%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 7.69-7.65 (4H, m, Ar-H), 7.44-7.30 (6H, m, Ar-H), 6.01 (1H, app s, H1'), 5.73 (1H, s, H5), 5.17 (1H, d, *J* 6.4, H2'), 4.81 (1H, dd, *J* 6.2 and 4.6, H3'), 4.20-4.15 (1H, m, H4'), 3.88-3.86 (2H, m, H5' and H5''), 1.49 (3H, s, CH₃), 1.31 (3H, s, CH₃), 1.02 (9H, s, 3 x CH₃). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 166.35 (CO), 151.88 (CO), 136.77 (2 x Ar-C), 136.69 (2 x Ar-C), 134.68 (Ar-C), 134.55 (Ar-C), 130.84 (Ar-C), 130.78 (Ar-C), 128.76 (2 x Ar-C), 128.61 (2 x Ar-C), 114.62 (quat C), 100.43 (C5), 95.29 (C1'), 91.14 (C4'), 86.19 (C2'), 83.69 (C3'), 66.24 (C5'), 27.63 (CH₃), 27.31 (3 x CH₃), 25.55 (CH₃), 19.98 (quat C). Signals not observed: 1 x CO and C6. HRMS (ES⁻) (*m/z*): 565.2021 ([M-H]⁻); C₂₉H₃₃N₂O₈Si requires 565.2012 (+1.5924 ppm).

6-Carboxyuridine 56

6-Carboxy-5'-TBDPS-2',3'-isopropylidene uridine (87 mg, 0.15 mmol) was suspended in distilled H₂O (1 mL) and cooled to 0 °C. A 50% aqueous TFA solution (1 mL) was added and the reaction mixture was then warmed to room temperature and stirred for 3 hours. The reaction mixture was concentrated *in vacuo* and purified

by flash column chromatography (10% MeOH/90% DCM – 30% MeOH/70% DCM with 1% AcOH). The desired product was isolated as an off-white solid (21 mg, 49%). ¹H NMR (400 MHz, D₂O): δ (ppm) 5.78 (1H, s, H5), 5.57 (1H, d, *J* 3.2, H1'), 4.75 (1H, dd, *J* 6.6 and 3.4, H2'), 4.36 (1H, app t, *J* 6.8, H3'), 3.97-3.93 (1H, m, H4'), 3.89 (1H, dd, *J* 12.4 and 2.8, H5'), 3.77 (1H, dd, *J* 12.4 and 6.4, H5''). ¹³C NMR (100 MHz, D₂O): δ (ppm) 167.29 (quat C), 166.22 (quat C), 154.91 (quat C), 151.22 (quat C), 98.64 (C5), 94.59 (C1'), 83.74 (C4'), 71.97 (C2'), 69.16 (C3'), 61.48 (C5'). HRMS (ES⁻) (*m/z*): 287.0525 ([M-H]⁻); C₁₀H₁₁N₂O₈ requires 287.0521 (+1.3935 ppm).

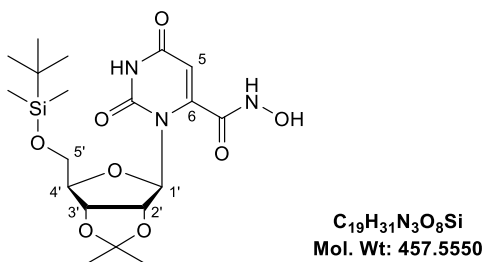
6-Methoxy carbonyl-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 57



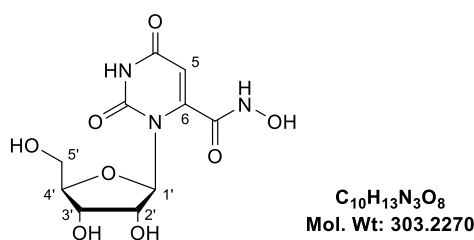
Anhydrous diisopropylamine (0.91 mL, 6.49 mmol, 2.6 eqv.) was dissolved in anhydrous THF (4 mL) and cooled to -78 °C. 1.6 M *n*-BuLi (4.10 mL, 6.56 mmol, 2.6 eqv.) was added and the resulting LDA mixture was stirred at -78 °C for 20 mins. A solution of 5'-TBDMS-2',3'-isopropylidene uridine (1.00 g, 2.51 mmol, 1 eqv.) in anhydrous THF (3 mL) was added and the reaction mixture was stirred for 1 hour at -78 °C. A solution of methyl chloroformate (0.23 mL, 2.98 mmol, 1.2 eqv.) in anhydrous THF (2 mL) was added and the reaction mixture was maintained at -78 °C for 4 hours. The reaction was quenched by the addition of a saturated solution of NH₄Cl (1 mL). The reaction was then diluted with EtOAc (50 mL) and washed with a saturated NaHCO₃ solution (50 mL) and brine (50 mL). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo* to leave a crude residue. The crude was purified by flash column chromatography (100% Hex – 60% Hex/40% EtOAc) to yield the desired product as a white solid (0.34 g, 30%). *R_f* = 0.43 (30% EtOAc/70% Toluene). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.59 (1H, br s, NH), 6.08 (1H, app s, H5), 5.90 (1H, s, H1'), 5.17 (1H, app d, *J* 6.4, H2'), 4.73 (1H, app t, *J* 5.6, H3'), 4.12-4.08 (1H, m, H4'), 3.94 (3H, s, OCH₃), 3.83-3.74 (2H, m, H5' and H5''), 1.54 (3H, s, CH₃), 1.34 (3H, s, CH₃), 0.87 (9H, s, 3 x CH₃), 0.04 (6H, s, 2 x SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 162.35 (CO), 162.05 (CO), 150.06 (CO), 145.26 (C6),

114.45 (quat C), 106.13 (C5), 93.50 (C1'), 88.85 (C4'), 84.83 (C2'), 81.61 (C3'), 64.01 (C5'), 54.01 (OCH₃), 27.36 (CH₃), 26.05 (3 x CH₃), 25.47 (CH₃), 18.59 (quat C), -5.13 (SiCH₃), -5.16 (SiCH₃). HRMS (ES⁺) (m/z): 479.1823 ([M+Na]⁺); C₂₀H₃₂N₂O₈SiNa requires 479.1820 (+0.6261 ppm).

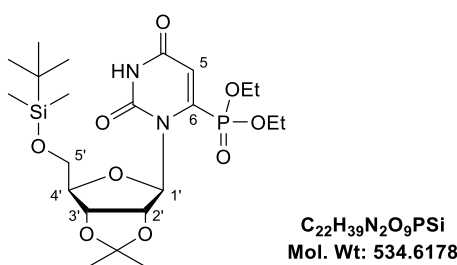
6-Hydroxamic acid-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine **58**¹⁸⁷



6-Methoxy carbonyl-5'-TBDMS-2',3'-isopropylidene uridine (0.34 g, 0.74 mmol, 1 eqv.) was dissolved in a 1.5 M solution of hydroxylamine in MeOH (5 mL, 10 eqv.) and stirred in a sealed reaction vial for 2 days. All solvent was then removed *in vacuo* and the crude residue obtained was purified by flash column chromatography (0.5% MeOH/99.5% DCM – 5% MeOH/95% DCM) to yield the desired product as an off-white solid (70 mg, 21%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 5.80 (1H, d, *J* 1.2, H1'), 5.75 (1H, s, H5), 5.24 (1H, dd, *J* 6.6 and 1.4, H2'), 4.78 (1H, dd, *J* 6.6 and 4.6, H3'), 4.03 (1H, td, *J* 6.3 and 4.7, H4'), 3.82-3.80 (2H, m, H5' and H5''), 1.50 (3H, s, CH₃), 1.33 (3H, s, CH₃), 0.90 (9H, s, 3 x CH₃), 0.07 (3H, s, CH₃), 0.07 (3H, s, CH₃). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 164.84 (quat C), 161.08 (quat C), 151.65 (quat C), 149.14 (quat C), 115.05 (quat C), 103.98 (C5), 94.65 (C1'), 90.30 (C4'), 85.94 (C2'), 83.43 (C3'), 65.05 (C5'), 27.52 (CH₃), 26.38 (3 x CH₃), 25.53 (CH₃), 19.20 (quat C), -5.08 (SiCH₃), -5.20 (SiCH₃). HRMS (ES⁺) (m/z): 480.1771 ([M+Na]⁺); C₁₉H₃₁N₃O₈SiNa requires 480.1773 (-0.4165 ppm).

6-Hydroxamic acid uridine 59

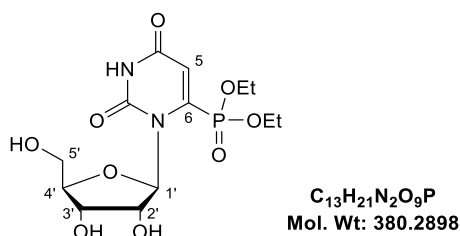
6-Hydroxamic acid-5'-TBDMS-2',3'-isopropylidene uridine (95 mg, 0.21 mmol) was suspended in distilled H₂O (1 mL) and cooled to 0 °C. A 50% aqueous TFA solution (1 mL) was added and the reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction was then concentrated *in vacuo* and purified by flash column chromatography (10% MeOH/90% DCM – 20% MeOH/80% DCM) to yield the desired product as an off-white solid (61 mg, 96%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 5.77 (1H, s, H5), 5.58 (1H, d, *J* 3.2, H1'), 4.69 (1H, dd, *J* 5.8 and 3.8, H2'), 4.28 (1H, app t, *J* 6.0, H3'), 3.88-3.84 (1H, m, H4'), 3.80 (1H, dd, *J* 12.0 and 2.8, H5'), 3.70 (1H, dd, *J* 12.0 and 6.0, H5''). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 164.97 (quat C), 161.29 (quat C), 151.86 (quat C), 150.05 (quat C), 104.10 (C5), 95.56 (C1'), 86.39 (C4'), 73.47 (C2'), 71.00 (C3'), 63.48 (C5'). HRMS (ES⁻) (*m/z*): 302.0635 ([M-H]⁻); C₁₀H₁₂N₃O₈ requires 302.0630 (+1.6553 ppm).

6-Diethyl phosphonate-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 60¹⁸⁸

Anhydrous diisopropylamine (0.91 mL, 6.49 mmol, 2.6 eqv.) was dissolved in anhydrous THF (4 mL) and cooled to -78 °C. 1.6 M *n*-BuLi (4.10 mL, 6.56 mmol, 2.6 eqv.) was added and the resulting LDA mixture was stirred at -78 °C for 20 mins. A solution of 5'-TBDMS-2',3'-isopropylidene uridine (0.99 g, 2.48 mmol, 1 eqv.) in anhydrous THF (3 mL) was added and the reaction mixture was stirred for 1 hour at -78 °C. A solution of diethyl chlorophosphate (0.44 mL, 3.04 mmol, 1.2 eqv.) in anhydrous THF (2 mL) was added and the reaction mixture was maintained at -78 °C for 4 hours. The reaction was quenched by the addition of a saturated solution of

NH_4Cl (1 mL). The reaction was then diluted with EtOAc (50 mL) and washed with a saturated NaHCO_3 solution (50 mL) and brine (50 mL). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* to leave a crude residue. The crude was purified by flash column chromatography (90% Hex/10% EtOAc – 50% Hex/50% EtOAc) to yield the desired product as a white solid (0.45 g, 34%). ^1H NMR (400 MHz, CDCl_3): δ (ppm) 10.50 (1H, s, NH), 6.37 (1H, d, J 14.0, H5), 6.06 (1H, app s, H1'), 5.08 (1H, dd, J 6.6 and 1.0, H2'), 4.68 (1H, dd, J 6.4 and 4.8, H3'), 4.23-4.11 (4H, m, 2 x CH_2CH_3), 4.03-3.99 (1H, m, H4'), 3.75-3.67 (2H, m, H5' and H5''), 1.42 (3H, s, CH_3), 1.31 (6H, t, J 7.2, 2 x CH_2CH_3), 1.22 (3H, s, CH_3), 0.76 (9H, s, 3 x CH_3), -0.08 (3H, s, SiCH_3), -0.08 (3H, s, SiCH_3). ^{13}C NMR (100 MHz, CDCl_3): δ (ppm) 162.24 (d, J 19.7, CO), 150.31 (d, J 12.3, CO), 145.47 (d, J 189.9, C6), 113.53 (quat C), 112.21 (d, J 10.7, C5), 94.39 (C1'), 89.14 (C4'), 84.15 (C2'), 81.92 (C3'), 64.63 (d, J 5.6, CH_2CH_3), 64.37 (d, J 6.2, CH_2CH_3), 64.04 (C5'), 27.04 (CH_3), 25.74 (3 x CH_3), 25.29 (CH_3), 18.19 (quat C), 15.96 (d, J 6.0, CH_2CH_3), 15.90 (d, J 6.1, CH_2CH_3), -5.36 (SiCH_3), -5.43 (SiCH_3). ^{31}P NMR (162 MHz, CDCl_3): δ (ppm) 5.61. HRMS (ES⁺) (m/z): 557.2069 ([M+Na]⁺): $\text{C}_{22}\text{H}_{39}\text{N}_2\text{O}_9\text{PSiNa}$ requires 557.2055 (+2.5125 ppm).

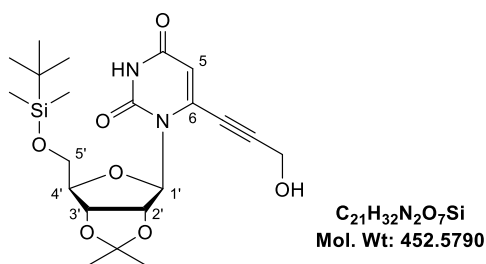
6-Diethyl phosphonate uridine 61



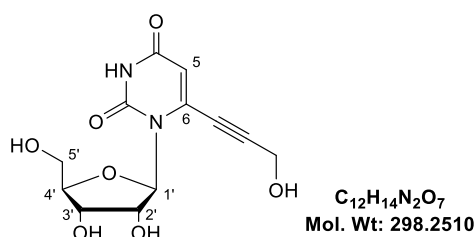
6-Diethyl phosphonate-5'-TBDMS-2',3'-isopropylidene uridine (0.38 g, 0.71 mmol) was suspended in distilled H_2O (3 mL) and cooled to 0 °C. A 50% aqueous TFA solution (2 mL) was added and the reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction was then concentrated under reduced pressure and the crude material purified by flash column chromatography (3% MeOH/97% DCM – 8% MeOH/92% DCM). The desired product was isolated as a white solid (0.24 g, 89%). ^1H NMR (400 MHz, CD_3OD): δ (ppm) 6.37 (1H, d, J 14.4, H5), 5.86 (1H, d, J 2.8, H1'), 4.70 (1H, dd, J 6.4 and 3.2, H2'), 4.35 (1H, app t, J 6.4, H3'), 4.32-4.24 (4H, m, 2 x CH_2CH_3), 3.86 (1H, td, J 6.4 and 2.9, H4'), 3.81 (1H, dd, J 12.0 and 2.8, H5'), 3.68 (1H, dd, J 12.0 and 6.0, H5''), 1.42 (6H, td, J 7.1 and 2.5, 2 x CH_2CH_3). ^{13}C NMR (100 MHz, CD_3OD): δ (ppm) 163.88 (d, J 19.6, CO),

151.80 (d, J 12.2, CO), 146.95 (d, J 191.4, C6), 113.49 (d, J 10.5, C5), 97.08 (d, J 2.7, C1'), 86.06 (C4'), 73.35 (C2'), 71.20 (C3'), 66.21 (d, J 6.4, CH₂CH₃), 66.15 (d, J 6.5, CH₂CH₃), 63.73 (C5'), 16.46 (d, J 3.0, CH₂CH₃), 16.40 (d, J 3.2, CH₂CH₃). ³¹P NMR (162 MHz, CD₃OD): δ (ppm) 6.18. HRMS (ES⁻) (m/z): 379.0911 ([M-H]⁻); C₁₃H₂₀N₂O₉P requires 379.0912 (-0.2638 ppm).

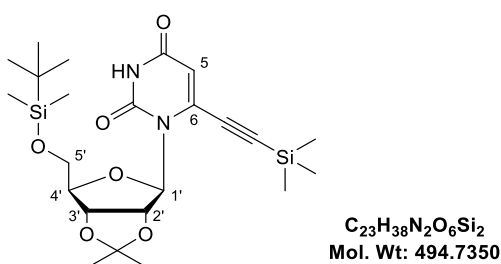
6-(3-Hydroxyprop-1-yn-1-yl)-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 64



6-Iodo-5'-TBDMS-2',3'-isopropylidene uridine (0.30 g, 0.57 mmol, 1 eqv.) was dissolved in degassed, anhydrous Et₃N (6 mL). To this solution was added Pd(PPh₃)₂Cl₂ (20 mg, 5 mol%), CuI (6 mg, 5 mol%) and propargyl alcohol (43 μ L, 0.74 mmol, 1.3 eqv.). The resulting reaction mixture was stirred at room temperature overnight. All solvent was removed *in vacuo* and the crude residue produced was re-dissolved in EtOAc (50 mL) and washed with H₂O (50 mL). The aqueous layer was extracted further with EtOAc (3 x 30 mL) and then the combined organic extracts were dried over Na₂SO₄. The crude material was concentrated under reduced pressure and purified by flash column chromatography (60% Hex/40% EtOAc – 50% Hex/50% EtOAc) to give a pale yellow solid as product (0.25 g, 97%). R_f = 0.11 (60% Hex/40% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.83 (1H, s, NH), 6.21 (1H, s, H1'), 5.95 (1H, s, H5), 5.17 (1H, dd, J 6.4 and 0.8, H2'), 4.81 (1H, dd, J 6.2 and 4.6, H3'), 4.52 (2H, d, J 5.6, CH₂OH), 4.17-4.13 (1H, m, H4'), 3.81-3.80 (2H, m, H5' and H5''), 3.20 (1H, t, J 5.8, CH₂OH), 1.54 (3H, s, CH₃), 1.33 (3H, s, CH₃), 0.87 (9H, s, 3 x CH₃), 0.05 (6H, s, 2 x SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 162.90 (CO), 149.74 (CO), 137.98 (quat C), 113.97 (quat C), 108.02 (C5), 101.80 (quat C), 93.81 (C1'), 89.64 (C4'), 84.29 (C2'), 82.04 (C3'), 75.62 (quat C), 64.23 (C5'), 51.18 (CH₂OH), 27.33 (CH₃), 26.06 (3 x CH₃), 25.47 (CH₃), 18.59 (quat C), -5.12 (SiCH₃), -5.13 (SiCH₃). HRMS (ES⁺) (m/z): 475.1884 ([M+Na]⁺); C₂₁H₃₂N₂O₇SiNa requires 475.1871 (+2.7358 ppm).

6-(3-Hydroxyprop-1-yn-1-yl) uridine 66

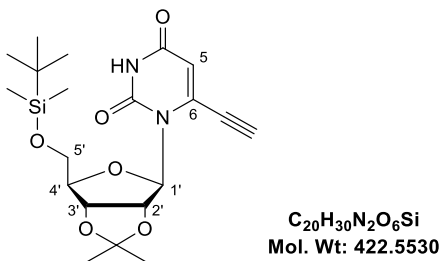
6-(3-Hydroxyprop-1-yn-1-yl)-5'-TBDMS-2',3'-isopropylidene uridine (0.11 g, 0.24 mmol) was suspended in distilled H₂O (1 mL) and cooled to 0 °C. A 50% aqueous TFA solution (1 mL) was added and the reaction mixture was allowed to warm to room temperature and stir for 2 hours. The reaction mixture was then concentrated under reduced pressure and the crude material was purified by flash column chromatography (5% MeOH/95% DCM – 12% MeOH/88% DCM) to yield the desired product as a white solid (62 mg, 87%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 6.09 (1H, d, *J* 4.0, H1'), 5.94 (1H, s, H5), 4.74 (1H, dd, *J* 6.0 and 4.0, H2'), 4.46 (2H, s, CH₂OH), 4.32 (1H, app t, *J* 6.2, H3'), 3.89 (1H, td, *J* 6.0 and 3.2, H4'), 3.80 (1H, dd, *J* 12.0 and 3.2, H5'), 3.68 (1H, dd, *J* 12.0 and 6.0, H5''). ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 164.47 (CO), 151.53 (CO), 139.75 (quat C), 108.64 (C5), 102.95 (quat C), 95.91 (C1'), 86.35 (C4'), 75.92 (quat C), 72.92 (C2'), 71.29 (C3'), 63.64 (C5'), 51.03 (CH₂OH). HRMS (ES⁻) (*m/z*): 297.0728 ([M-H]⁻); C₁₂H₁₃N₂O₇ requires 297.0728 (0.0000 ppm).

6-((Trimethylsilyl)ethynyl)-5'-tert-butylidimethylsilyl-2',3'-isopropylidene uridine 67

6-Iodo-5'-TBDMS-2',3'-isopropylidene uridine (0.15 g, 0.29 mmol, 1 eqv.) was dissolved in degassed, anhydrous Et₃N (3.5 mL). To this solution was added Pd(PPh₃)₂Cl₂ (20 mg, 10 mol%), CuI (5 mg, 10 mol%) and TMS acetylene (95 μL, 0.69 mmol, 2.4 eqv.). The reaction mixture was stirred at room temperature overnight. All solvent was then removed *in vacuo* and the crude residue was dissolved in EtOAc (30 mL). The reaction mixture was washed with water (30 mL)

and the aqueous layer was then extracted with EtOAc (2 x 30 mL) and the combined organic extracts were washed further with brine (30 mL). H₂O₂ (~10 drops) was added to the organic extracts which were then washed once more with brine (20 mL) and dried over Na₂SO₄. The crude material was concentrated under reduced pressure and purified by flash column chromatography (100% Hex – 80% Hex/20% EtOAc) to yield the desired product (90 mg, 63%). R_f = 0.35 (70% Hex/30% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 8.73 (1H, br s, NH), 6.32 (1H, d, *J* 1.2, H1'), 5.93 (1H, d, *J* 2.4, H5), 5.19 (1H, dd, *J* 6.4 and 1.6, H2'), 4.79 (1H, dd, *J* 6.6 and 4.2, H3'), 4.19-4.15 (1H, m, H4'), 3.84-3.77 (2H, m, H5' and H5''), 1.53 (3H, s, CH₃), 1.35 (3H, s, CH₃), 0.89 (9H, s, 3 x CH₃), 0.29 (9H, s, 3 x SiCH₃), 0.06 (6H, s, 2 x SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 162.12 (CO), 149.57 (CO), 137.86 (C6), 113.88 (quat C), 111.17 (quat C), 107.90 (C5), 94.48 (C1'), 93.89 (quat C), 89.42 (C4'), 84.12 (C2'), 82.36 (C3'), 64.26 (C5'), 27.38 (CH₃), 26.12 (3 x CH₃), 25.73 (CH₃), 18.65 (quat C), -0.71 (3 x SiCH₃), -5.06 (SiCH₃), -5.09 (SiCH₃). HRMS (ES⁺) (m/z): 517.2162 ([M+Na]⁺); C₂₃H₃₈N₂O₆Si₂Na requires 517.2161 (+0.1933 ppm).

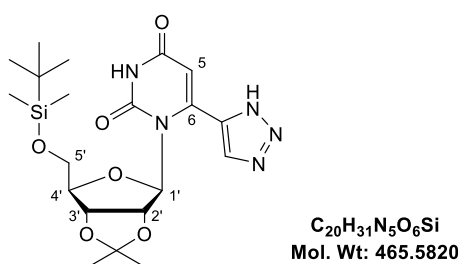
6-Ethynyl-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 68



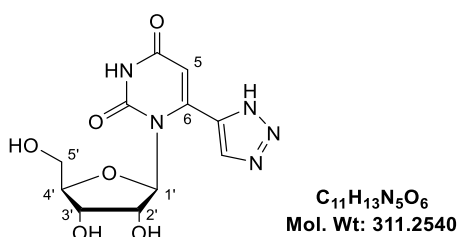
6-((Trimethylsilyl)ethynyl)-5'-TBDMS-2',3'-isopropylidene uridine (0.10 g, 0.20 mmol, 1 eqv.) was dissolved in MeOH (1.5 mL). To this solution was added potassium carbonate (36 mg, 0.26 mmol, 1.3 eqv.) and the resulting reaction mixture was stirred for 1 hour. The reaction mixture was then diluted with water (10 mL) and extracted with EtOAc (3 x 10 mL). The organic extracts were dried over Na₂SO₄ and concentrated *in vacuo*. The crude material was purified by flash column chromatography (100% Hex – 60% Hex/40% EtOAc) to yield the desired product as a pale yellow solid (73 mg, 86%). R_f = 0.18 (70% Hex/30% EtOAc). ¹H NMR (400 MHz, CDCl₃): δ (ppm) 9.24 (1H, br s, NH), 6.26 (1H, d, *J* 1.2, H1'), 6.02 (1H, d, *J* 2.4, H5), 5.19 (1H, dd, *J* 6.4 and 1.2, H2'), 4.82 (1H, dd, *J* 6.4 and 4.4, H3'), 4.18-4.14 (1H, m, H4'), 3.84-3.76 (3H, m, H5', H5'' and CCH), 1.55 (3H, s, CH₃), 1.34

(3H, s, CH_3), 0.88 (9H, s, 3 x CH_3), 0.05 (6H, s, 2 x $SiCH_3$). ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 162.14 (CO), 149.58 (CO), 137.28 (C6), 113.97 (quat C), 109.16 (C5), 94.04 (C1'), 90.91 (CCH), 89.70 (C4'), 84.31 (C2'), 82.12 (C3'), 73.81 (CCH), 64.21 (C5'), 27.38 (CH_3), 26.09 (3 x CH_3), 25.52 (CH_3), 18.62 (quat C), -5.09 ($SiCH_3$), -5.10 ($SiCH_3$). HRMS (ES^+) (m/z): 445.1768 ($[M+Na]^+$); $C_{20}H_{30}N_2O_6SiNa$ requires 445.1765 (+0.6739 ppm).

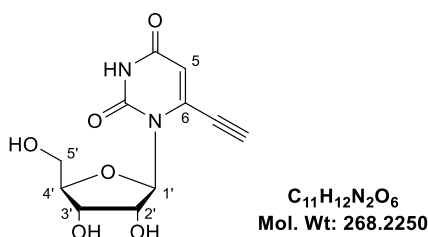
6-(1*H*-1,2,3-Triazol-5-yl)-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine
69¹⁹⁴



6-Ethynyl-5'-TBDMS-2',3'-isopropylidene uridine (41 mg, 0.10 mmol, 1 eqv.) was dissolved in a mixture of anhydrous DMF (0.9 mL) and anhydrous MeOH (0.1 mL). CuI (1 mg, 5.25 μ mol, 5 mol%) was then added followed by TMS azide (0.02 mL, 0.15 mmol, 1.5 eqv.). The reaction mixture was heated to 100 °C and stirred overnight. Once TLC showed complete absence of starting material, the reaction was cooled to room temperature and all solvent was removed *in vacuo*. The crude material was purified by flash column chromatography (0.5% MeOH/99.5% DCM – 5% MeOH/95% DCM) to yield the desired product (9 mg, 19%). 1H NMR (400 MHz, CD_3OD): δ (ppm) 8.21 (1H, s, CCHN), 5.90 (1H, s, H5), 5.83 (1H, app s, H1'), 5.32 (1H, dd, J 6.4 and 0.8, H2'), 4.08 (1H, td, J 6.6 and 4.0, H4'), 3.87-3.80 (2H, m, H5' and H5''), 1.43 (3H, s, CH_3), 1.32 (3H, s, CH_3), 0.90 (9H, s, 3 x CH_3), 0.07 (3H, s, $SiCH_3$), 0.06 (3H, s, $SiCH_3$). Signal not observed: H3' (correlation present but peak hidden under water peak). ^{13}C NMR (100 MHz, CD_3OD): δ (ppm) 164.98 (quat C), 152.26 (quat C), 148.13 (quat C), 114.53 (quat C), 105.03 (C5), 94.48 (C1'), 91.12 (C4'), 85.81 (C2'), 83.84 (C3'), 65.30 (C5'), 27.40 (CH_3), 26.38 (3 x CH_3), 25.47 (CH_3), 19.25 (quat C), -5.07 ($SiCH_3$), -5.14 ($SiCH_3$). Signals not observed: 2 x C from triazole. HRMS (ES^+) (m/z): 488.1942 ($[M+Na]^+$); $C_{20}H_{31}N_5O_6SiNa$ requires 488.1936 (+1.2290 ppm).

6-(1*H*-1,2,3-Triazol-5-yl) uridine 71

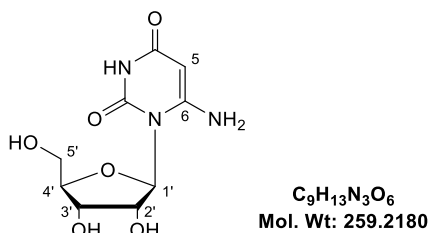
6-(1*H*-1,2,3-Triazol-5-yl)-5'-TBDMS-2',3'-isopropylidene uridine (9 mg, 19.3 μ mol) and 6-(1*H*-1,2,3-triazol-5-yl)-2',3'-isopropylidene uridine (10 mg, 28.5 μ mol) were suspended in a 30 % aqueous TFA solution (1 mL) and stirred for 2 hours. The reaction mixture was then concentrated *in vacuo* and the crude residue was purified by flash column chromatography (5% MeOH/95% DCM – 12% MeOH/88% DCM). The desired product was isolated as a pale yellow solid (10 mg, 67%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 8.26 (1H, s, CCHN), 5.91 (1H, s, H5), 5.52 (1H, d, *J* 3.2, H1'), 4.77 (1H, dd, *J* 6.0 and 3.6, H2'), 4.31 (1H, app t, *J* 6.2, H3'), 3.85-3.79 (2H, m, H4' and H5'), 3.71-3.66 (1H, m, H5''). ¹³C NMR (400 MHz, CD₃OD): δ (ppm) 165.00 (quat C), 152.31 (quat C), 149.06 (quat C), 140.17 (quat C), 130.66 (CCHN), 105.03 (C5), 95.41 (C1'), 86.16 (C4'), 73.05 (C2'), 71.48 (C3'), 63.85 (C5'). HRMS (ES⁻) (*m/z*): 310.0792 ([M-H]⁻); C₁₁H₁₂N₅O₆ requires 310.0793 (-0.3225 ppm).

6-Ethynyl uridine 72

6-Ethynyl-5'-TBDMS-2',3'-isopropylidene uridine (68 mg, 0.16 mmol) was suspended in a 30% aqueous TFA solution (2 mL) and stirred at room temperature for 2 hours. The reaction mixture was then concentrated under reduced pressure and the crude residue purified by flash column chromatography (5% MeOH/95% DCM – 10% MeOH/90% DCM). The desired product was isolated as an off-white solid (33 mg, 77%). ¹H NMR (400 MHz, D₂O): δ (ppm) 6.23 (1H, s, H5), 6.21 (1H, d, *J* 3.6, H1'), 4.85 (1H, dd, *J* 6.4 and 4.0, H2'), 4.40 (1H, app t, *J* 6.6, H3'), 3.99 (1H, td, *J* 6.3 and 3.1, H4'), 3.90 (1H, dd, *J* 12.4 and 2.8, H5'), 3.77 (1H, dd, *J* 12.4 and 6.0, H5''). Signal not observed: CCH.¹³C NMR (100 MHz, D₂O): δ (ppm) 164.78

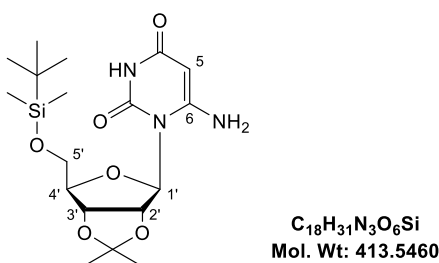
(CO), 150.72 (CO), 137.91 (C6), 109.07 (C5), 93.66 (C1'), 83.75 (C4'), 71.39 (C2'), 69.21 (C3'), 61.35 (C5'). Signals not observed: CCH and CCH. HRMS (ES-) (m/z): 267.0621 ([M-H]⁻); C₁₁H₁₁N₂O₆ requires 267.0623 (-0.7489 ppm).

6-Aminouridine 78



6-Azidouridine (0.2 g, 0.70 mmol, 1 eqv.) was dissolved in EtOH (5 mL) and the mixture was purged with N₂. 10% Pd/C (75 mg, 10 mol%) was added and the reaction mixture was placed under a H₂ atmosphere and left to stir overnight. The reaction mixture was filtered through Celite® and all solvent removed *in vacuo* to give the desired product as a white solid (0.17 g, 94%). ¹H NMR (400 MHz, CD₃OD): δ (ppm) 6.41 (1H, d, *J* 7.2, H1'), 4.62 (1H, app t, *J* 7.0, H2'), 4.25 (1H, dd, *J* 6.4 and 3.2, H3'), 4.01-3.99 (1H, m, H4'), 3.78-3.77 (2H, m, H5' and H5''). Signal not observed: H5. ¹³C NMR (100 MHz, CD₃OD): δ (ppm) 158.63 (quat C), 152.90 (quat C), 89.90 (C1'), 86.92 (C4'), 71.39 (C2'), 71.18 (C3'), 61.80 (C5'). Signals not observed: 2 x quaternary carbons. HRMS (ES⁺) (m/z): 282.0694 ([M+Na]⁺); C₉H₁₃N₃O₆ requires 282.0697 (-1.0636 ppm).

6-Amino-5'-*tert*-butyldimethylsilyl-2',3'-isopropylidene uridine 79



6-Azido-5'-TBDMS-2',3'-isopropylidene uridine (0.12 g, 0.27 mmol, 1 eqv.) was dissolved in EtOH (5 mL) and the mixture was purged with N₂. 10% Pd/C (0.03 g, 10 mol%) was added and the reaction mixture was placed under a H₂ atmosphere and left to stir overnight. The reaction mixture was filtered through Celite® and concentrated to yield the desired product (95 mg, 85%). ¹H NMR (400 MHz, CDCl₃):

δ (ppm) 9.30 (1H, br s, NH), 6.54 (1H, d, J 4.0, H1'), 5.68 (2H, br s, NH₂), 5.07 (1H, dd, J 6.6 and 4.2, H2'), 4.95 (1H, s, H5), 4.84-4.81 (1H, m, H3'), 4.04-4.03 (1H, m, H4'), 3.96-3.82 (2H, m, H5' and H5''), 1.55 (3H, s, CH₃), 1.33 (3H, s, CH₃), 0.89 (9H, s, 3 x CH₃), 0.09 (3H, s, SiCH₃), 0.08 (3H, s, SiCH₃). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) 163.54 (quat C), 155.81 (quat C), 150.60 (quat C), 115.59 (quat C), 88.98 (C1'), 84.38 (C4'), 82.16 (C2'), 80.35 (C5), 78.69 (C3'), 62.08 (C5'), 27.34 (CH₃), 26.08 (3 x CH₃), 25.55 (CH₃), 18.71 (quat C), -5.32 (SiCH₃), -5.37 (SiCH₃). HRMS (ES⁺) (m/z): 436.1868 ([M+Na]⁺); C₁₈H₃₁N₃O₆SiNa requires 436.1874 (-1.3756 ppm).

5.4 Computational protocols for results and discussion 3

The molecular docking program GOLD was used to validate the binding of an existing inhibitor, BMP, and to investigate the binding ability of the eleven final compounds synthesised.¹⁹⁸ Default settings provided by GOLD were used unless otherwise stated. The visualisation program PyMOL was used to view crystal structures and binding poses generated by GOLD.²⁰⁰ The molecular modelling program Spartan '16 was used to construct the eleven final compounds as their 5' monophosphate ready for docking into the ODCase crystal structure.¹⁹⁷

Protocol 1 - Visualisation of ODCase crystal structure in complex with the known inhibitor BMP

The crystal structure of ODCase in complex with the known inhibitor BMP was downloaded from the protein data bank (PDB code: 1X1Z).²⁰¹ The 1X1Z PDB file was loaded into PyMOL which showed both chains of the dimer with a BMP ligand in each active site.^{200,201} It was decided to delete the ligand from monomer B and concentrate on the active site of monomer A. Water molecules that had polar contacts with the BMP ligand were identified (eight in total) and all others were deleted along with three glycerols. The edited 1X1Z dimer was saved as .pdb file.

Protocol 2 – Modelling of compounds for docking

Spartan '16 was used to model the eleven final compounds as their 5' phosphates.¹⁹⁷ In addition, BMP was also modelled so it could be redocked into the edited crystal structure obtained in Protocol 1 to validate binding. An energy minimisation was performed on all structures using the molecular mechanics force field (MMFF) and they were then saved as .sdf files.

Protocol 3 - Validation of docking

The edited 1X1Z file, obtained from Protocol 1, was loaded into GOLD and all hydrogen atoms were added.¹⁹⁸ Eight water molecules (HOH402, HOH403, HOH406, HOH407, HOH408, HOH409, HOH411 and HOH461) were extracted. Ligand A, the crystallographic BMP molecule, was then extracted. The binding site was defined as being all atoms within 6 Å of where the BMP molecule was bound. The BMP molecule constructed according to Protocol 2 was then loaded using the 'Select Ligands' option and 25 GA runs were specified. The crystallographic BMP molecule was set as a reference ligand. In 'Configure Waters', the eight water molecules were set to toggle and spin states. Under 'Fitness & Search Options',

docking was selected with ChemPLP chosen as the scoring function. The parameter file was set to default and early termination was not allowed. The search efficiency was increased to 200% in 'GA settings' and the output files were set to save as one file, in the SD format, in the relevant directory with no lone pairs set to save.

Protocol 4 - Docking of final compounds

The edited 1X1Z file, obtained from Protocol 1, was loaded into GOLD and all hydrogen atoms were added.¹⁹⁸ Eight water molecules (HOH402, HOH403, HOH406, HOH407, HOH408, HOH409, HOH411 and HOH461) were extracted. Ligand A, the crystallographic BMP molecule, was then extracted. The binding site was defined as being all atoms within 6 Å of where the BMP molecule was bound. In turn, the file for each of the final compounds obtained from Protocol 2 was then loaded using the 'Select Ligands' option and 25 GA runs were specified. The crystallographic BMP molecule was set as a reference ligand. In 'Configure Waters', the eight water molecules were set to toggle and spin states. Under 'Fitness & Search Options', docking was selected with ChemPLP chosen as the scoring function. The parameter file was set to default and early termination was not allowed. The search efficiency was increased to 200% in 'GA settings' and the output files were set to save as one file, in the SD format, in the relevant directory with no lone pairs set to save.

Chapter 6
Bibliography

Chapter 6 Bibliography

1. W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, 1984.
2. G. M. Blackburn, M. J. Gait, D. Loakes and D. M. Williams, *Nucleic acids in chemistry and biology*, Royal Society of Chemistry, Cambridge, 3rd edn., 2006.
3. IUPAC-IUB Joint Commission on Biochemical Nomenclature, *Eur. J. Biochem.*, 1983, **131**, 9-15.
4. V. A. Bloomfield, D. M. Crothers and I. Tinoco, *Nucleic Acids. Structures, Properties, and Functions*, University Science Books, 2000.
5. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737-738.
6. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964-967.
7. W. Saenger, *Angew. Chem., Int. Ed. Engl.*, 1973, **12**, 591-601.
8. N. Foloppe, B. Hartmann, L. Nilsson and A. D. MacKerell, *Biophys. J.*, 2002, **82**, 1554-1569.
9. M. L. Hamm, S. Rajguru, A. M. Downs and R. Cholera, *J. Am. Chem. Soc.*, 2005, **127**, 12220-12221.
10. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205-8212.
11. J. Clayden, N. Greeves and S. G. Warren, *Organic Chemistry*, Oxford University Press, 2nd edn., 2012.
12. J. Plavec, W. Tong and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 1993, **115**, 9734-9746.
13. J. Plavec, C. Thibaudeau and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 1994, **116**, 6558-6560.
14. K. Bondensgaard, M. Petersen, S. K. Singh, V. K. Rajwanshi, R. Kumar, J. Wengel and J. P. Jacobsen, *Chem. Eur. J.*, 2000, **6**, 2687-2695.
15. K. Hoogsteen, *Acta Cryst.*, 1959, **12**, 822-823.
16. K. Hoogsteen, *Acta Cryst.*, 1963, **16**, 907-916.
17. E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 1-22.
18. A. Udvardy and P. Schedl, *J. Mol. Biol.*, 1983, **166**, 159-181.

19. A. Ghosh and M. Bansal, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 620-626.
20. S. Arnott and D. W. L. Hukins, *Biochem. Biophys. Res. Commun.*, 1972, **47**, 1504-1509.
21. R. E. Franklin and R. G. Gosling, *Nature*, 1953, **171**, 740.
22. R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini and M. L. Kopka, *Science*, 1982, **216**, 475.
23. R. E. Franklin and R. G. Gosling, *Acta Cryst.*, 1953, **6**, 673-677.
24. C. Courseille, A. Dautant, M. Hospital, B. Langlois d'Estaintot, G. Precigoux, D. Molko and R. Teoule, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1990, **46**, FC9-FC12.
25. R. E. Dickerson, in *Methods Enzymol.*, Academic Press, 1992, vol. 211, pp. 67-111.
26. X. J. Lu and W. K. Olson, *Nucleic Acids Res.*, 2003, **31**, 5108-5121.
27. P. Belmont, J.-F. Constant and M. Demeunynck, *Chem. Soc. Rev.*, 2001, **30**, 70-81.
28. R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Nature*, 1980, **287**, 755.
29. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, *Nature*, 1979, **282**, 680.
30. T. J. Thamann, R. C. Lord, A. H. J. Wang and A. Rich, *Nucleic Acids Res.*, 1981, **9**, 5443-5458.
31. A. Herbert and A. Rich, *J. Biol. Chem.*, 1996, **271**, 11595-11598.
32. A. Rich and S. Zhang, *Nat. Rev. Genet.*, 2003, **4**, 566.
33. A. Herbert, J. Alfken, Y.-G. Kim, I. S. Mian, K. Nishikura and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 8421-8426.
34. T. Schwartz, J. Behlke, K. Lowenhaupt, U. Heinemann and A. Rich, *Nature Struct. Biol.*, 2001, **8**, 761.
35. Y.-G. Kim, M. Muralinath, T. Brandt, M. Percy, K. Hauns, K. Lowenhaupt, B. L. Jacobs and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6974-6979.

36. A. Jain, G. Wang and K. M. Vasquez, *Biochimie*, 2008, **90**, 1117-1130.
37. I. Radhakrishnan and D. J. Patel, *Biochemistry*, 1994, **33**, 11405-11416.
38. J.-s. Sun, T. Carestier and C. Hélène, *Curr. Opin. Struct. Biol.*, 1996, **6**, 327-333.
39. M. P. Knauert and P. M. Glazer, *Hum. Mol. Genet.*, 2001, **10**, 2243-2251.
40. H. A. Day, P. Pavlou and Z. A. E. Waller, *Bioorg. Med. Chem.*, 2014, **22**, 4407-4418.
41. K. Gehring, J.-L. Leroy and M. Guéron, *Nature*, 1993, **363**, 561.
42. S. Benabou, A. Aviñó, R. Eritja, C. González and R. Gargallo, *RSC Advances*, 2014, **4**, 26956-26980.
43. A. Rajendran, S.-i. Nakano and N. Sugimoto, *Chem. Commun.*, 2010, **46**, 1299-1301.
44. J. Cui, P. Waltman, V. Le and E. Lewis, *Molecules*, 2013, **18**, 12751.
45. D. Sun and L. H. Hurley, *J. Med. Chem.*, 2009, **52**, 2863-2874.
46. M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger and D. Christ, *Nature Chem.*, 2018, **10**, 631-637.
47. D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364.
48. M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770.
49. R. Hänsel-Hertsch, M. Di Antonio and S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.*, 2017, **18**, 279.
50. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402-5415.
51. J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2005, **33**, 2908-2916.
52. V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith and S. Balasubramanian, *Nat. Biotechnol.*, 2015, **33**, 877.
53. G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nature Chem.*, 2013, **5**, 182.

54. Y. Xu, *Chem. Soc. Rev.*, 2011, **40**, 2719-2740.
55. A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2006, **34**, 2723-2735.
56. N. Kim, M. Piatyszek, K. Prowse, C. Harley, M. West, P. Ho, G. Coviello, W. Wright, S. Weinrich and J. Shay, *Science*, 1994, **266**, 2011-2015.
57. T. Sawa and H. Ohshima, *Nitric Oxide*, 2006, **14**, 91-100.
58. Y. Hiraku, *Environ. Health Prev. Med.*, 2010, **15**, 63-72.
59. A. Mantovani, P. Allavena, A. Sica and F. Balkwill, *Nature*, 2008, **454**, 436.
60. P. Lonkar and P. C. Dedon, *Int. J. Cancer*, 2011, **128**, 1999-2009.
61. H. Ohshima and H. Bartsch, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, 1994, **305**, 253-264.
62. F. Balkwill and A. Mantovani, *The Lancet*, 2001, **357**, 539-545.
63. G. Multhoff, M. Molls and J. Radons, *Front. Immunol.*, 2012, **2**, 1-17.
64. L. M. Coussens, Werb, Z., *Nature*, 2002, **420**, 860.
65. R. Kumar, G. Clermont, Y. Vodovotz and C. C. Chow, *J. Theor. Biol.*, 2004, **230**, 145-155.
66. R. Medzhitov, *Nature*, 2008, **454**, 428.
67. M. Karin, T. Lawrence and V. Nizet, *Cell*, 2006, **124**, 823-835.
68. D. Okin and R. Medzhitov, *Curr. Biol.*, 2012, **22**, R733-R740.
69. H. Ohshima, *Toxicol. Lett.*, 2003, **140-141**, 99-104.
70. C. Nathan, *Nature*, 2002, **420**, 846.
71. N. R. Jena, *J. Biosci.*, 2012, **37**, 503-517.
72. L. J. Hofseth, S. P. Hussain, G. N. Wogan and C. C. Harris, *Free Radical Biol. Med.*, 2003, **34**, 955-968.
73. P. C. Dedon and S. R. Tannenbaum, *Arch. Biochem. Biophys.*, 2004, **423**, 12-22.

74. U. Förstermann and W. C. Sessa, *Eur. Heart J.*, 2012, **33**, 829-837.
75. N. Suzuki, M. Yasui, N. E. Geacintov, V. Shafirovich and S. Shibutani, *Biochemistry*, 2005, **44**, 9238-9245.
76. S. Kawanishi, S. Ohnishi, N. Ma, Y. Hiraku and M. Murata, *Int. J. Mol. Sci.*, 2017, **18**, 1808.
77. D. A. Wink and J. B. Mitchell, *Free Radical Biol. Med.*, 1998, **25**, 434-456.
78. S. Steenken and S. V. Jovanovic, *J. Am. Chem. Soc.*, 1997, **119**, 617-618.
79. W. L. Neeley and J. M. Essigmann, *Chem. Res. Toxicol.*, 2006, **19**, 491-505.
80. K. Kaneko, T. Akuta, T. Sawa, H. W. Kim, S. Fujii, T. Okamoto, H. Nakayama, H. Ohigashi, A. Murakami and T. Akaike, *Cancer Letters*, 2008, **262**, 239-247.
81. V. Yermilov, J. Rubio and H. Ohshima, *FEBS Lett.*, 1995, **376**, 207-210.
82. M. Masuda, H. Nishino and H. Ohshima, *Chem.-Biol. Interact.*, 2002, **139**, 187-197.
83. S. Obeid, N. Blatter, R. Kranaster, A. Schnur, K. Diederichs, W. Welte and A. Marx, *EMBO J.*, 2010, **29**, 1738-1747.
84. I. Bhamra, P. Compagnone-Post, I. A. O'Neil, L. A. Iwanejko, A. D. Bates and R. Cosstick, *Nucleic Acids Res.*, 2012, **40**, 11126-11138.
85. T. Iyama and D. M. Wilson, *DNA Repair*, 2013, **12**, 620-636.
86. Aziz Sancar, Laura A. Lindsey-Boltz, Keziban Ünsal-Kaçmaz and S. Linn, *Annu. Rev. Biochem.*, 2004, **73**, 39-85.
87. K. Yun-Jeong and M. W. David, III, *Curr. Mol. Pharmacol.*, 2012, **5**, 3-13.
88. J. T. Reardon and A. Sancar, in *Methods Enzymol.*, Academic Press, 2006, vol. 408, pp. 189-213.
89. G.-M. Li, *Cell Res.*, 2007, **18**, 85-98.
90. R. R. Iyer, A. Pluciennik, V. Burdett and P. L. Modrich, *Chem. Rev.*, 2006, **106**, 302-323.
91. J. Jiricny, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 335-346.
92. W. Yang, *Mutat. Res.*, 2000, **460**, 245-256.

93. A. Sancar, *Chem. Rev.*, 2003, **103**, 2203-2238.
94. S. Mitra, *DNA Repair*, 2007, **6**, 1064-1070.
95. A. E. Pegg, *Mutat. Res., Rev. Mutat. Res.*, 2000, **462**, 83-100.
96. A. E. Pegg, M. E. Dolan and R. C. Moschel, in *Prog. Nucleic Acid Res. Mol. Biol.*, eds. W. E. Cohn and K. Moldave, Academic Press, 1995, vol. 51, pp. 167-223.
97. B. Demple, B. Sedgwick, P. Robins, N. Totty, M. D. Waterfield and T. Lindahl, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 2688-2692.
98. V. Yermilov, J. Rubio, M. Becchi, M. D. Friesen, B. Pignatelli and H. Ohshima, *Carcinogenesis*, 1995, **16**, 2045-2050.
99. H.-J. C. Chen, Y.-M. Chen and C.-M. Chang, *Chem.-Biol. Interact.*, 2002, **140**, 199-213.
100. R. Singh, U. Manjunatha, H. I. M. Boshoff, Y. H. Ha, P. Niyomrattanakit, R. Ledwidge, C. S. Dowd, I. Y. Lee, P. Kim, L. Zhang, S. Kang, T. H. Keller, J. Jiricek and C. E. Barry, *Science*, 2008, **322**, 1392-1395.
101. L. A. Kaplan, *J. Am. Chem. Soc.*, 1964, **86**, 740-741.
102. D. W. Lamson, P. Ulrich and R. O. Hutchins, *J. Org. Chem.*, 1973, **38**, 2928-2930.
103. A. N. Blinnikov and N. N. Makhova, *Mendeleev Commun.*, 1999, **9**, 13-14.
104. Y. Saito, H. Taguchi, S. Fujii, T. Sawa, E. Kida, C. Kabuto, T. Akaike and H. Arimoto, *Chem. Commun.*, 2008, 5984-5986.
105. S. Saito and Y. Koizumi, *Tetrahedron Lett.*, 2005, **46**, 4715-4717.
106. V. Terzic, D. Padovani, V. Balland, I. Artaud and E. Galardon, *Org. Biomol. Chem.*, 2014, **12**, 5360-5364.
107. Y. Fuchi and S. Sasaki, *Org. Lett.*, 2014, **16**, 1760-1763.
108. M. Ikehara and K. E. I. Muneyama, *Chem. Pharm. Bull.*, 1970, **18**, 1196-1200.
109. I. M. Lagoja, *Chem. Biodiversity*, 2005, **2**, 1-50.

110. M. Lucas-Hourani, D. Dauzonne, H. Munier-Lehmann, S. Khiar, S. Nisole, J. Dairou, O. Helynck, P. V. Afonso, F. Tangy and P.-O. Vidalain, *Antimicrob. Agents Chemother.*, 2017, **61**.
111. P. Reyes, P. K. Rathod, D. J. Sanchez, J. E. K. Mrema, K. H. Rieckmann and H.-G. Heidrich, *Mol. Biochem. Parasitol.*, 1982, **5**, 275-290.
112. M. J. Downie, K. Kirk and C. B. Mamoun, *Eukaryot. Cell*, 2008, **7**, 1231-1237.
113. R. I. Christopherson, S. D. Lyons and P. K. Wilson, *Acc. Chem. Res.*, 2002, **35**, 961-971.
114. A. Radzicka and R. Wolfenden, *Science*, 1995, **267**, 90.
115. N. Wu, W. Gillon and E. F. Pai, *Biochemistry*, 2002, **41**, 4002-4011.
116. B. P. Callahan and B. G. Miller, *Bioorg. Chem.*, 2007, **35**, 465-469.
117. E. Poduch, L. Wei, E. F. Pai and L. P. Kotra, *J. Med. Chem.*, 2008, **51**, 432-438.
118. M. Fujihashi, J. S. Mnpotra, R. K. Mishra, E. F. Pai and L. P. Kotra, *J. Genet. Genomics*, 2015, **42**, 221-234.
119. B. G. Miller and R. Wolfenden, *Annu. Rev. Biochem.*, 2002, **71**, 847-885.
120. W. Cui, J. G. DeWitt, S. M. Miller and W. Wu, *Biochem. Biophys. Res. Commun.*, 1999, **259**, 133-135.
121. M. Fujihashi, L. Wei, L. P. Kotra and E. F. Pai, *J. Mol. Biol.*, 2009, **387**, 1199-1210.
122. B. G. Miller, M. J. Snider, R. Wolfenden and S. A. Short, *J. Biol. Chem.*, 2001, **276**, 15174-15176.
123. P. Harris, J.-C. Navarro Poulsen, K. F. Jensen and S. Larsen, *Biochemistry*, 2000, **39**, 4217-4224.
124. S. Mundra and L. P. Kotra, *Future Med. Chem.*, 2014, **6**, 165-177.
125. 4NUW, <https://www.rcsb.org/structure/4NUW>, (accessed 27/07/18).
126. P. Beak and B. Siegel, *J. Am. Chem. Soc.*, 1976, **98**, 3601-3606.
127. M. A. Rishavy and W. W. Cleland, *Biochemistry*, 2000, **39**, 4569-4574.

128. R. B. Silverman and M. P. Groziak, *J. Am. Chem. Soc.*, 1982, **104**, 6434-6439.
129. S. A. Acheson, J. B. Bell, M. E. Jones and R. Wolfenden, *Biochemistry*, 1990, **29**, 3198-3202.
130. A. Vardi-Kilshtain, D. Doron and D. T. Major, *Biochemistry*, 2013, **52**, 4382-4390.
131. N. Wu, Y. Mo, J. Gao and E. F. Pai, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 2017-2022.
132. H. Hu, A. Boone and W. Yang, *J. Am. Chem. Soc.*, 2008, **130**, 14493-14503.
133. M. Fujihashi, K. Mito, E. F. Pai and K. Miki, *J. Biol. Chem.*, 2013, **288**, 9011-9016.
134. M. Fujihashi, T. Ishida, S. Kuroda, L. P. Kotra, E. F. Pai and K. Miki, *J. Am. Chem. Soc.*, 2013, **135**, 17432-17443.
135. A. Warshel, M. Štrajbl, J. Villà and J. Florián, *Biochemistry*, 2000, **39**, 14728-14738.
136. J. L. Van Vleet, L. A. Reinhardt, B. G. Miller, A. Sievers and W. W. Cleland, *Biochemistry*, 2008, **47**, 798-803.
137. K. Toth, T. L. Amyes, B. M. Wood, K. Chan, J. A. Gerlt and J. P. Richard, *J. Am. Chem. Soc.*, 2007, **129**, 12946-12947.
138. K. Toth, T. L. Amyes, B. M. Wood, K. Chan, J. A. Gerlt and J. P. Richard, *J. Am. Chem. Soc.*, 2010, **132**, 7018-7024.
139. W.-Y. Tsang, B. M. Wood, F. M. Wong, W. Wu, J. A. Gerlt, T. L. Amyes and J. P. Richard, *J. Am. Chem. Soc.*, 2012, **134**, 14580-14594.
140. T. C. Appleby, C. Kinsland, T. P. Begley and S. E. Ealick, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 2005-2010.
141. T. L. Amyes, B. M. Wood, K. Chan, J. A. Gerlt and J. P. Richard, *J. Am. Chem. Soc.*, 2008, **130**, 1574-1575.
142. A. Sievers and R. Wolfenden, *J. Am. Chem. Soc.*, 2002, **124**, 13986-13987.
143. B. Goryanova, T. L. Amyes, J. A. Gerlt and J. P. Richard, *J. Am. Chem. Soc.*, 2011, **133**, 6545-6548.

144. B. J. Desai, B. M. Wood, A. A. Fedorov, E. V. Fedorov, B. Goryanova, T. L. Amyes, J. P. Richard, S. C. Almo and J. A. Gerlt, *Biochemistry*, 2012, **51**, 8665-8678.
145. B. Goryanova, L. M. Goldman, T. L. Amyes, J. A. Gerlt and J. P. Richard, *Biochemistry*, 2013, **52**, 7500-7511.
146. B. M. Wood, T. L. Amyes, A. A. Fedorov, E. V. Fedorov, A. Shabila, S. C. Almo, J. P. Richard and J. A. Gerlt, *Biochemistry*, 2010, **49**, 3514-3516.
147. B. G. Miller, M. J. Snider, S. A. Short and R. Wolfenden, *Biochemistry*, 2000, **39**, 8113-8118.
148. M. Fujihashi, A. M. Bello, E. Poduch, L. Wei, S. C. Annedi, E. F. Pai and L. P. Kotra, *J. Am. Chem. Soc.*, 2005, **127**, 15048-15050.
149. H. L. Levine, R. S. Brody and F. H. Westheimer, *Biochemistry*, 1980, **19**, 4993-4999.
150. R. E. Handschumacher, *J. Biol. Chem.*, 1960, **235**, 2917-2919.
151. A. Čihák, J. Veselý and J. Škoda, *Advances in Enzyme Regulation*, 1985, **24**, 335-354.
152. S. S. Matsumoto, J. M. Fujihaki, L. D. Nord, R. C. Willis, V. M. Lee, B. S. Sharma, Y. S. Sanghvi, G. D. Kini, G. R. Revankar, R. K. Robins, W. B. Jolley and R. A. Smith, *Biochem. Pharmacol.*, 1990, **39**, 455-462.
153. E. Cadman, F. Eiferman, R. Heimer and L. Davis, *Cancer Res.*, 1978, **38**, 4610-4617.
154. J. D. Morrey, D. F. Smee, R. W. Sidwell and C. Tseng, *Antiviral Res.*, 2002, **55**, 107-116.
155. P. A. Salem, G. P. Bodey, M. A. Burgess, W. K. Murphy and E. J. Freireich, *Cancer*, 1977, **40**, 2806-2809.
156. E. C. Cadman, D. E. Dix and R. E. Handschumacher, *Cancer Res.*, 1978, **38**, 682-688.
157. A. M. Bello, D. Konforte, E. Poduch, C. Furlonger, L. Wei, Y. Liu, M. Lewis, E. F. Pai, C. J. Paige and L. P. Kotra, *J. Med. Chem.*, 2009, **52**, 1648-1658.
158. T. Najarian and T. W. Traut, *Neurorehab. Neural Re.*, 2000, **14**, 237-241.
159. M. E. Meza-Avina, L. Wei, Y. Liu, E. Poduch, A. M. Bello, R. K. Mishra, E. F. Pai and L. P. Kotra, *Bioorg. Med. Chem.*, 2010, **18**, 4032-4041.

160. N. Wu and E. F. Pai, *J. Biol. Chem.*, 2002, **277**, 28080-28087.
161. M. K. Purohit, E. Poduch, L. W. Wei, I. E. Crandall, T. To, K. C. Kain, E. F. Pai and L. P. Kotra, *J. Med. Chem.*, 2012, **55**, 9988-9997.
162. G. J. Crowther, A. J. Napuli, J. H. Gilligan, K. Gagaring, R. Borboa, C. Francek, Z. Chen, E. F. Dagostino, J. B. Stockmyer, Y. Wang, P. P. Rodenbough, L. J. Castaneda, D. J. Leibly, J. Bhandari, M. H. Gelb, A. Brinker, I. H. Engels, J. Taylor, A. K. Chatterjee, P. Fantauzzi, R. J. Glynne, W. C. Van Voorhis and K. L. Kuhen, *Mol. Biochem. Parasitol.*, 2011, **175**, 21-29.
163. M. K. Yadav, S. K. Pandey and D. Swati, *J. Bioinform. Comput. Biol.*, 2013, **11**, 1350003.
164. A. M. Bello, E. Poduch, M. Fujihashi, M. Amani, Y. Li, I. Crandall, R. Hui, P. I. Lee, K. C. Kain, E. F. Pai and L. P. Kotra, *J. Med. Chem.*, 2007, **50**, 915-921.
165. I. E. Crandall, E. Wasilewski, A. M. Bello, A. Mohmmed, P. Malhotra, E. F. Pai, K. C. Kain and L. P. Kotra, *J. Med. Chem.*, 2013, **56**, 2348-2358.
166. A. M. Bello, E. Poduch, Y. Liu, L. Wei, I. Crandall, X. Wang, C. Dyanand, K. C. Kain, E. F. Pai and L. P. Kotra, *J. Med. Chem.*, 2008, **51**, 439-448.
167. S. M. Firestine, W. Wu, H. Youn and V. Jo Davisson, *Bioorg. Med. Chem.*, 2009, **17**, 794-803.
168. J. K. Judice, T. R. Gamble, E. C. Murphy, A. M. de Vos and P. G. Schultz, *Science*, 1993, **261**, 1578-1581.
169. I. Muegge, *Med. Res. Rev.*, 2003, **23**, 302-321.
170. C. Ballatore, D. M. Huryn and A. B. Smith, *ChemMedChem*, 2013, **8**, 385-395.
171. J.-B. Lin, J.-L. He, Y.-C. Shih and T.-C. Chien, *Tetrahedron Lett.*, 2011, **52**, 3969-3972.
172. H. C. Kolb and K. B. Sharpless, *Drug Discov. Today*, 2003, **8**, 1128-1137.
173. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596-2599.
174. E. J. Corey and A. Venkateswarlu, *J. Am. Chem. Soc.*, 1972, **94**, 6190-6191.
175. E. J. Corey, B. Samuelsson and F. A. Luzzio, *J. Am. Chem. Soc.*, 1984, **106**, 3682-2683.

176. E. Pretsch, P. Bühlmann and M. Badertscher, *Structure Determination of Organic Compounds - Tables of Spectral Data*, Springer, 4th edn., 2009.
177. S. Maity, S. Manna, S. Rana, T. Naveen, A. Mallick and D. Maiti, *J. Am. Chem. Soc.*, 2013, **135**, 3355-3358.
178. T. Naveen, S. Maity, U. Sharma and D. Maiti, *J. Org. Chem.*, 2013, **78**, 5949-5954.
179. H. Inoue and T. Ueda, *Chem. Pharm. Bull.*, 1978, **26**, 2657-2663.
180. D. R. da Rocha, W. C. Santos, E. S. Lima and V. F. Ferreira, *Carbohydr. Res.*, 2012, **350**, 14-19.
181. B. T. Worrell, J. A. Malik and V. V. Fokin, *Science*, 2013, **340**, 457-460.
182. V. V. R. Rao, B. E. Fulloon, P. V. Bernhardt, R. Koch and C. Wentrup, *J. Org. Chem.*, 1998, **63**, 5779-5786.
183. K. D. Veeranna, K. K. Das and S. Baskaran, *Angew. Chem., Int. Ed.*, 2017, **56**, 16197-16201.
184. I. Bae, H. Han and S. Chang, *J. Am. Chem. Soc.*, 2005, **127**, 2038-2039.
185. R. Rayala and S. F. Wnuk, *Tetrahedron Lett.*, 2012, **53**, 3333.
186. S. D. Barrett, A. J. Bridges, D. T. Dudley, A. R. Saltiel, J. H. Fergus, C. M. Flamme, A. M. Delaney, M. Kaufman, S. LePage, W. R. Leopold, S. A. Przybranowski, J. Sebolt-Leopold, K. Van Becelaere, A. M. Doherty, R. M. Kennedy, D. Marston, W. A. Howard, Y. Smith, J. S. Warmus and H. Tecle, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 6501-6504.
187. C. T. Mathew and H. E. Ulmer, US4707294, 1987.
188. M. Honjo, T. Maruyama, M. Horikawa, J. A. N. Balzarini and E. D. Clercq, *Chem. Pharm. Bull.*, 1987, **35**, 3227-3234.
189. K. Sonogashira, Y. Tohda and N. Hagihara, *Tetrahedron Lett.*, 1975, **16**, 4467-4470.
190. R. Chinchilla and C. Nájera, *Chem. Soc. Rev.*, 2011, **40**, 5084-5121.
191. H. Tanaka, K. Haraguchi, Y. Koizumi, M. Fukui and T. Miyasaka, *Can. J. Chem.*, 1986, **64**, 1560-1563.
192. T. J. Donohoe, A. Ironmonger and N. M. Kershaw, *Angew. Chem., Int. Ed.*, 2008, **47**, 7314-7316.

193. J. B. Epp and T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293-295.
194. T. Jin, S. Kamijo and Y. Yamamoto, *Eur. J. Org. Chem.*, 2004, **2004**, 3789-3791.
195. J.-i. Setsune, M. Toda, K. Watanabe, P. K. Panda and T. Yoshida, *Tetrahedron Lett.*, 2006, **47**, 7541-7544.
196. P. A. Cox, A. G. Leach, A. D. Campbell and G. C. Lloyd-Jones, *J. Am. Chem. Soc.*, 2016, **138**, 9145-9157.
197. Spartan, <https://www.wavefun.com/>, (accessed 08/04/18).
198. GOLD, <https://www.ccdc.cam.ac.uk/>, (accessed 08/04/18).
199. J. W. Liebeschuetz, J. C. Cole and O. Korb, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 737-748.
200. PyMOL, <https://pymol.org/>, (accessed 08/04/18).
201. 1X1Z, <https://www.rcsb.org/structure/1x1z>, (accessed 12/04/18).
202. J. C. Cole, C. W. Murray, J. W. M. Nissink, R. D. Taylor and R. Taylor, *Proteins: Struct., Funct., Bioinf.*, 2005, **60**, 325-332.
203. E. Poduch, A. M. Bello, S. Tang, M. Fujihashi, E. F. Pai and L. P. Kotra, *J. Med. Chem.*, 2006, **49**, 4937-4945.

Chapter 7
Appendices

Chapter 7 Appendices

Appendix A: Masses obtained from HPLC-MS

8-Nitroguanosine 13

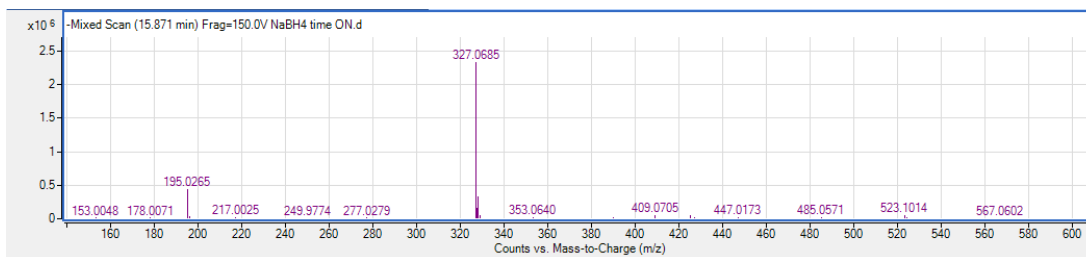


Figure 53 Mass peak from HPLC-MS for 8-nitroguanosine 13

8-Aminoguanosine 16

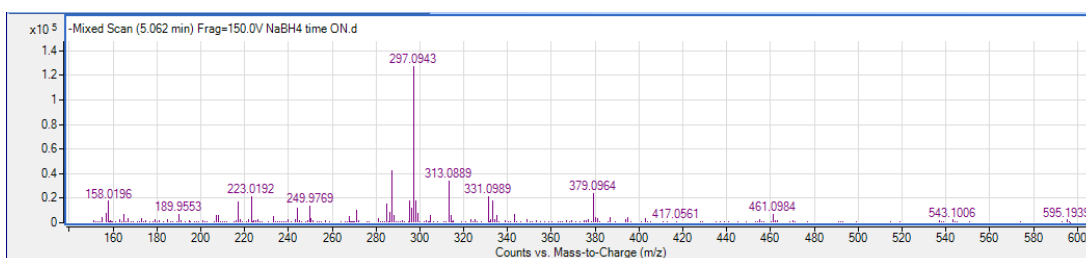


Figure 54 Mass peak from HPLC-MS for 8-aminoguanosine 16

8-Nitroguanine

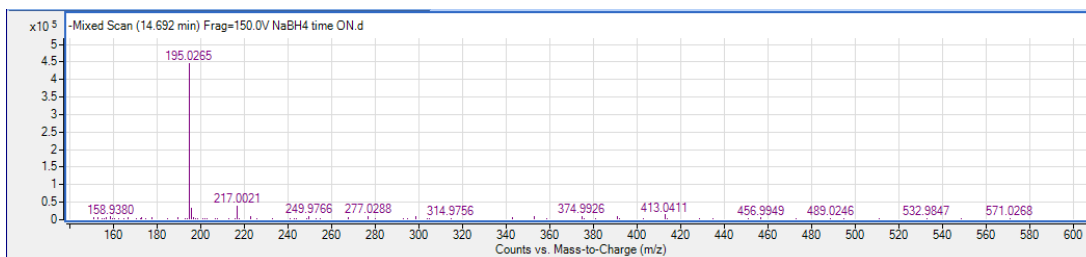


Figure 55 Mass peak from HPLC-MS for 8-nitroguanine

Guanosine 7

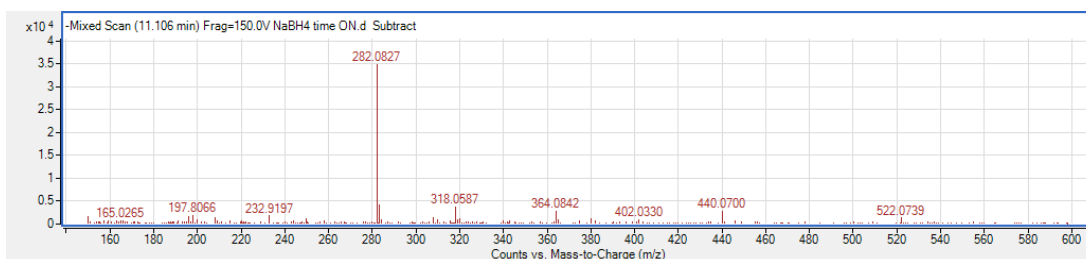


Figure 56 Mass peak from HPLC-MS for guanosine 7

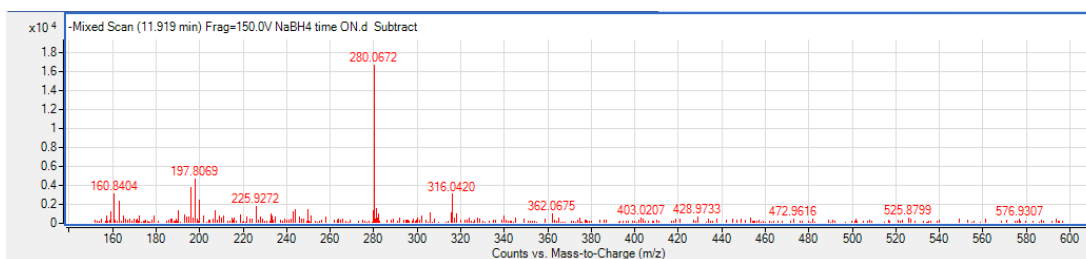
8,5'-O-Cycloguanosine 17

Figure 57 Mass peak from HPLC-MS for 8,5'-O-cycloguanosine 17

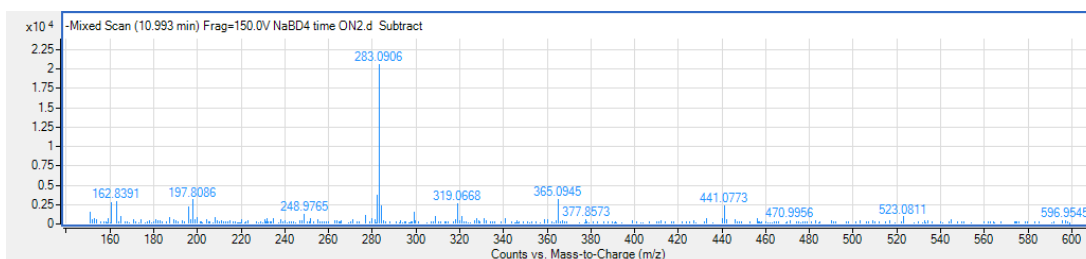
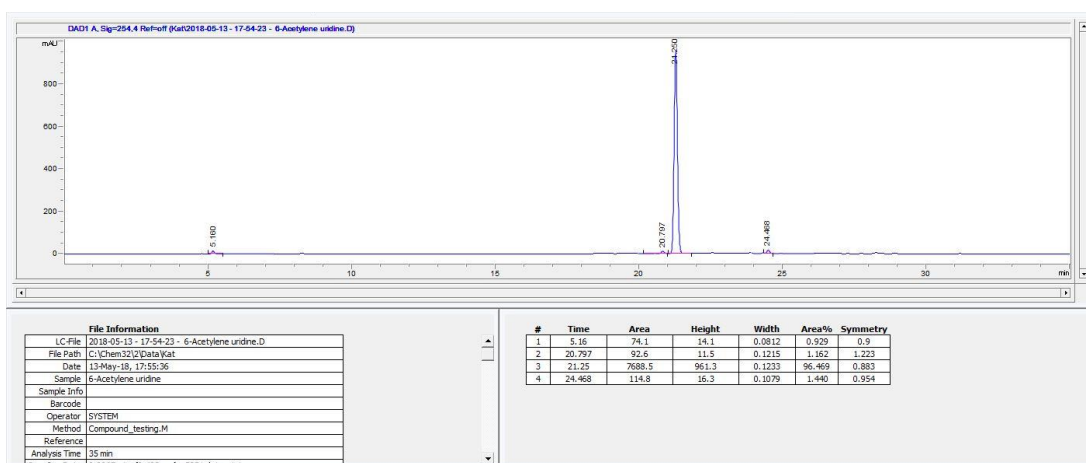
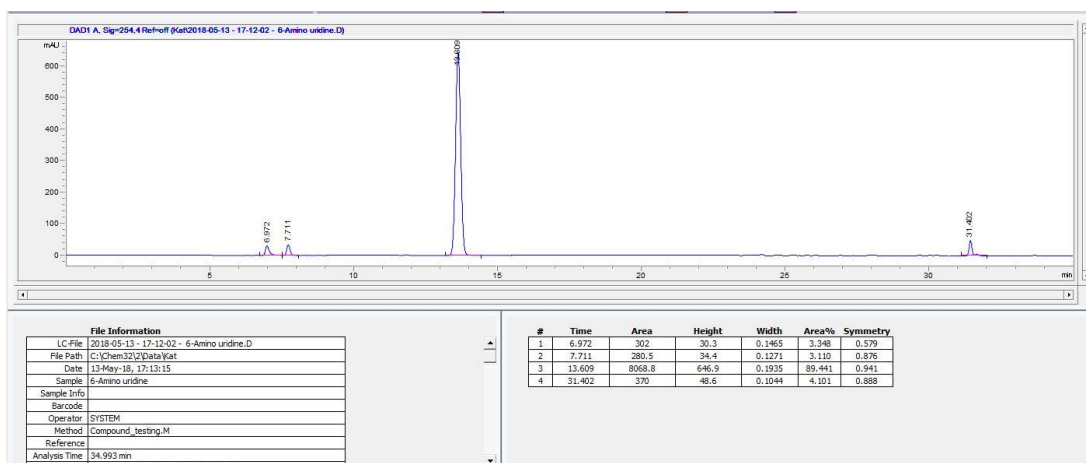
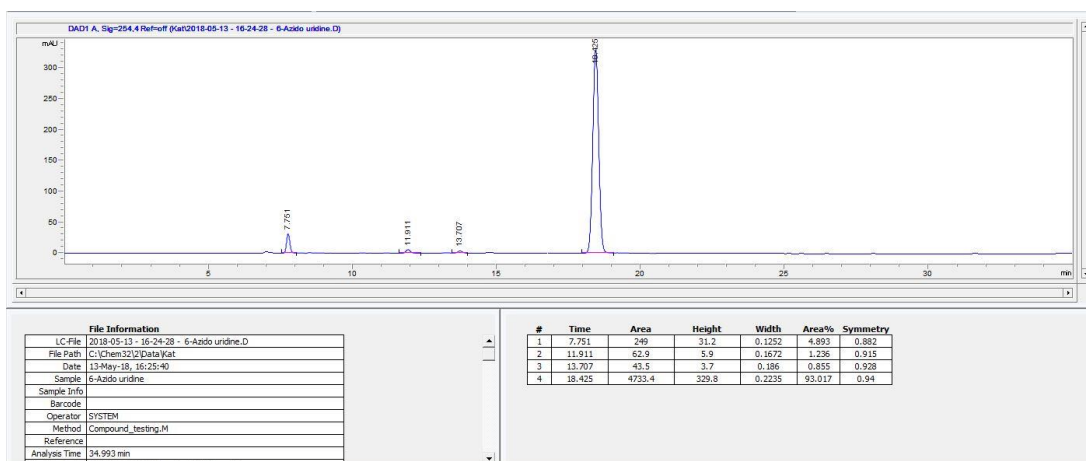
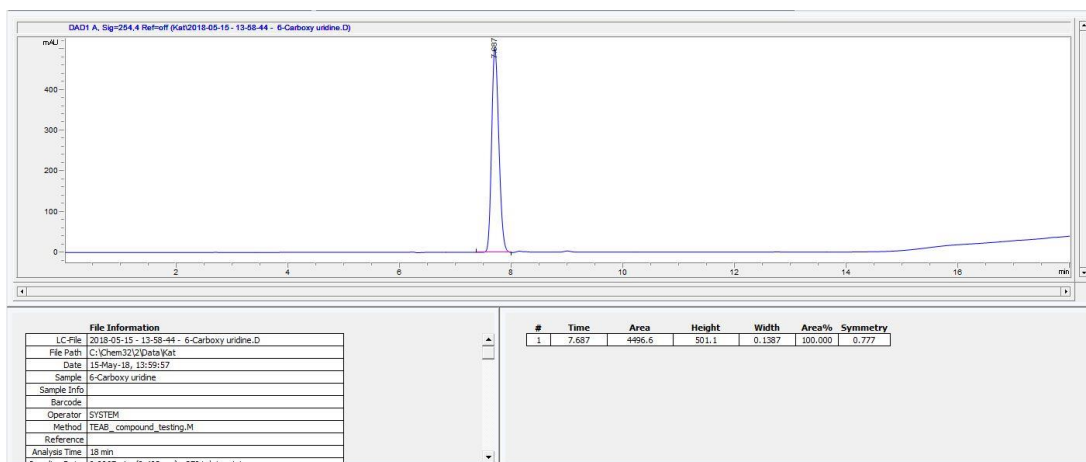
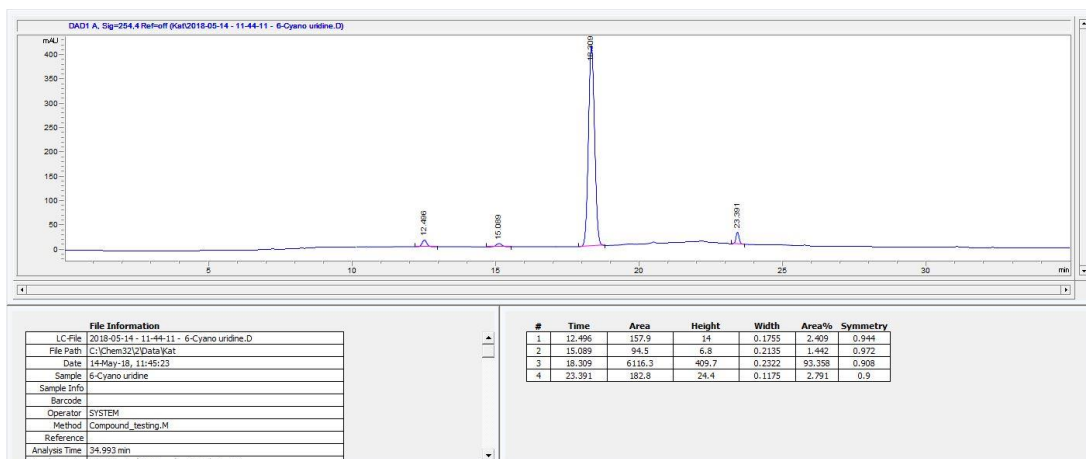
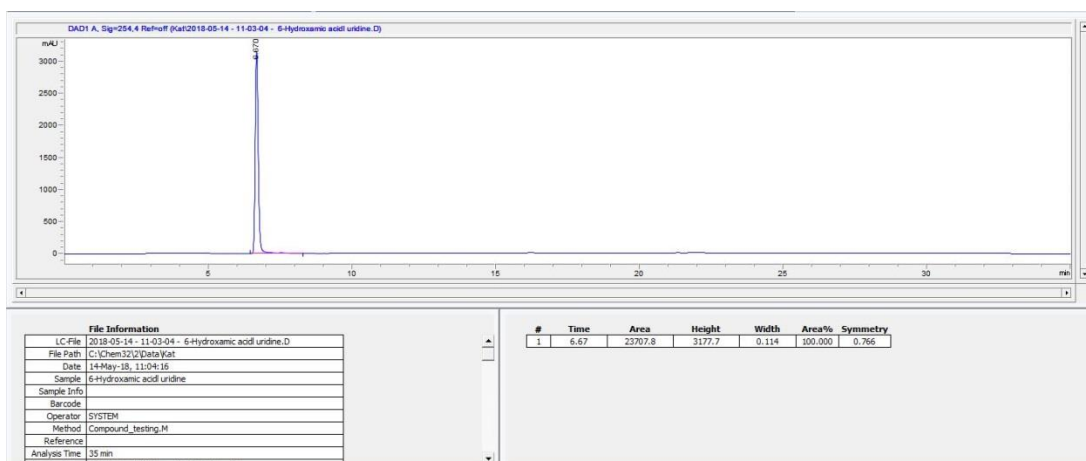
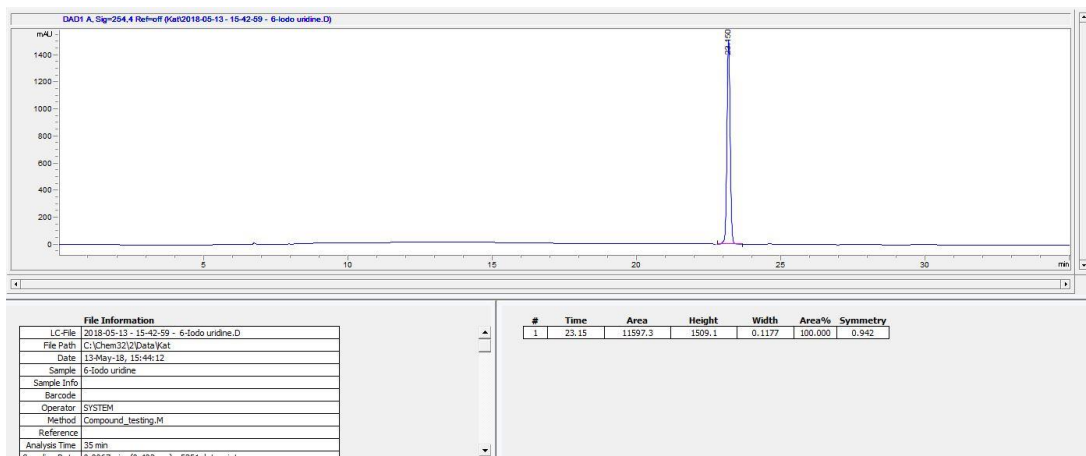
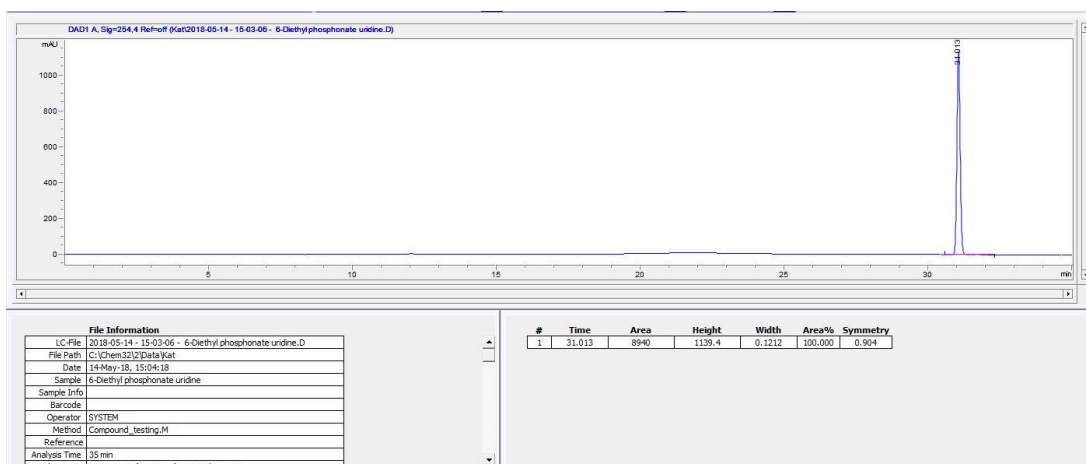
8-Deuteroguanosine 21

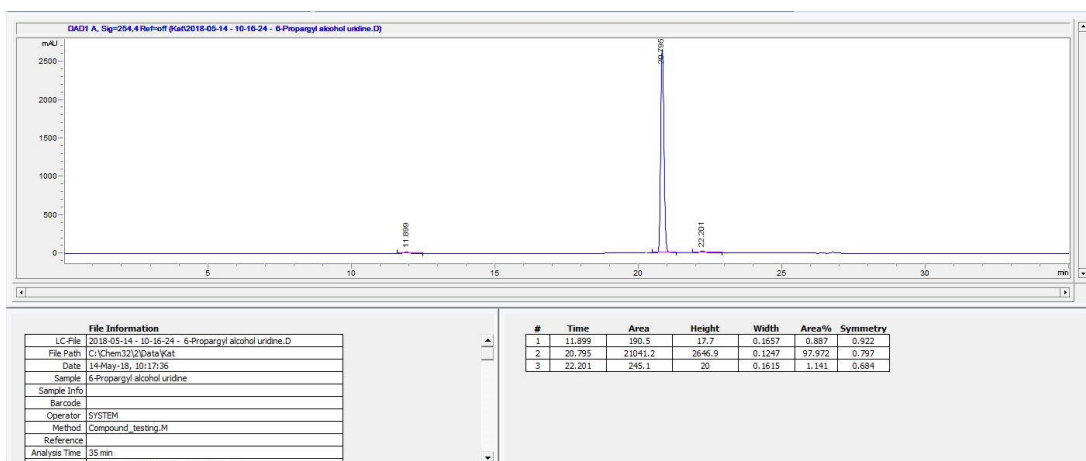
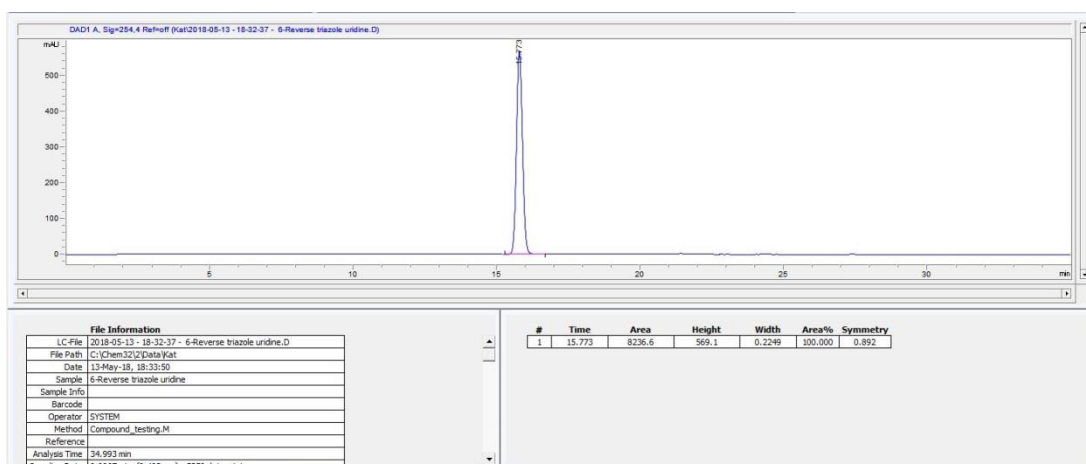
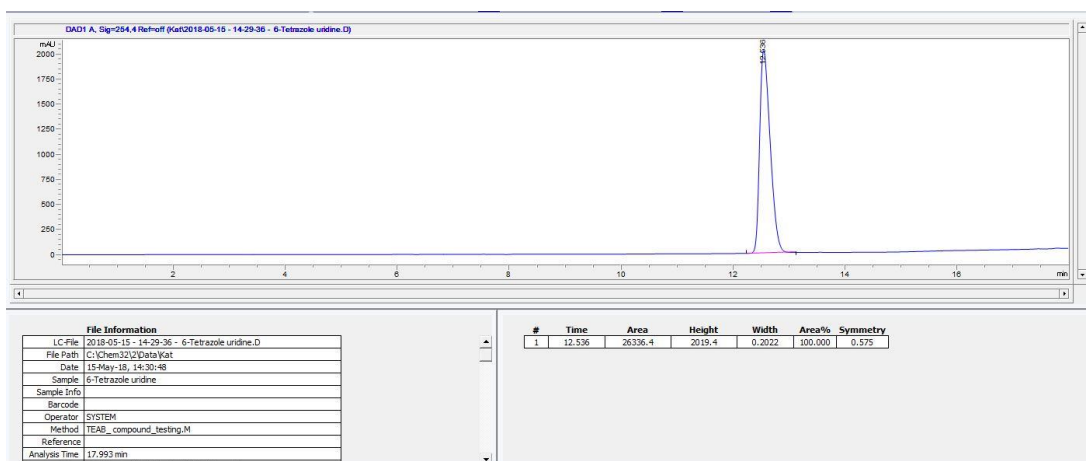
Figure 58 Mass peak from HPLC-MS for 8-deuteroguanosine 21

Appendix B: HPLC traces of eleven final compounds

6-Ethynyl uridine **72**Figure 59 HPLC trace of 6-ethynyl uridine **72**6-Aminouridine **78**Figure 60 HPLC trace of 6-aminouridine **78**

6-Azidouridine **38**Figure 61 HPLC trace of 6-azidouridine **38**6-Carboxyuridine **56**Figure 62 HPLC trace of 6-carboxyuridine **56**6-Cyanouridine **52**Figure 63 HPLC trace of 6-cyanouridine **52**

6-Hydroxamic acid uridine **59**Figure 64 HPLC trace of 6-hydroxamic acid uridine **59**6-Iodouridine **29**Figure 65 HPLC trace of 6-iodouridine **29**6-Diethyl phosphonate uridine **61**Figure 66 HPLC trace of 6-Diethyl phosphonate uridine **61**

6-(3-Hydroxyprop-1-yn-1-yl) uridine **66**Figure 67 HPLC trace of 6-(3-hydroxyprop-1-yn-1-yl) uridine **66**6-(1H-1,2,3-Triazol-5-yl) uridine **71**Figure 68 HPLC trace of 6-(1H-1,2,3-triazol-5-yl) uridine **71**6-(1H-Tetrazol-5-yl) uridine **51**Figure 69 HPLC trace of 6-(1H-Tetrazol-5-yl) uridine **51**

Chapter 8
Publication

Chapter 8 Publication



DOI: 10.1002/chem.201705541

CHEMISTRY
 A European Journal
 Full Paper

Biochemistry

Chemistry of the 8-Nitroguanine DNA Lesion: Reactivity, Labelling and Repair

 Katie J. Alexander, Matthew McConville, Kathryn R. Williams, Konstantin V. Luzyanin, Ian A. O'Neil, and Richard Cosstick^{✉[a]}

Abstract: The 8-nitroguanine lesion in DNA is increasingly associated with inflammation-related carcinogenesis, whereas the same modification on guanosine 3',5'-cyclic monophosphate generates a second messenger in NO-mediated signal transduction. Very little is known about the chemistry of 8-nitroguanine nucleotides, despite the fact that their biological effects are closely linked to their chemical properties. To this end, a selection of chemical reactions have been performed on 8-nitroguanine nucleosides and oligodeoxynucleotides. Reactions with alkylating reagents reveal how the

8-nitro substituent affects the reactivity of the purine ring, by significantly decreasing the reactivity of the N2 position, whilst the relative reactivity at N1 appears to be enhanced. Interestingly, the displacement of the nitro group with thiols results in an efficient and specific method of labelling this lesion and is demonstrated in oligodeoxynucleotides. Additionally, the repair of this lesion is also shown to be a chemically feasible reaction through a reductive denitration with a hydride source.

Introduction

The guanine base in DNA is particularly sensitive to chemical modification^[1] by reactive oxygen species and reactive nitrogen species (RNS) and a number of mutagenic lesions, such as 8-nitro-2'-deoxyguanosine^[2] (8-nitro-dG, **1a**, Figure 1) and the more widely studied 8-oxo-2'-deoxyguanosine^[3] (8-oxo-dG, **2a**)

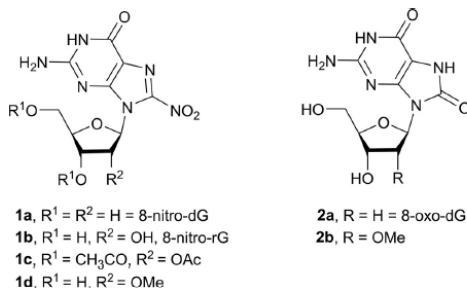


Figure 1. Structures of nucleosides involved in this study.

[a] Dr. K. J. Alexander, Dr. M. McConville, K. R. Williams, Dr. K. V. Luzyanin, Dr. I. A. O'Neil, Prof. R. Cosstick
 Department of Chemistry, University of Liverpool
 Crown Street, Liverpool, L69 7ZD (UK)
 E-mail: rcosstick@liv.ac.uk

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/chem.201705541>.

are associated with these reactive species. A characteristic of 8-nitrodG is that the nitro group labilises the glycosidic bond to release 8-nitroguanine (8-nitroG), leaving an apurinic site in the DNA.^[4] The mutagenic nature of 8-nitrodG results from both the generation of apurinic sites^[5] in the DNA and the mispairing of 8-nitrodG^[6–8] during DNA synthesis; both these events lead to G to T transversion mutations.

The RNS, which are responsible for the nitrate damage are generated particularly at sites of inflammation, due to over-expression of nitric oxide synthase and given the mutagenic nature of 8-nitro-dG, it is not surprising that this lesion is increasingly linked with inflammation-related carcinogenesis.^[9] Specific antibodies raised against 8-nitroG have been used to show that 8-nitroG is present in surgically-removed human tissues from the sites of cancers that are caused by chronic inflammation^[9] and the accumulation of 8-nitroG is also associated with a poor prognosis of inflammation-induced cancers.^[9,10] Thus, 8-nitroG is progressively seen as a biochemical marker for DNA damage and as potential indicator of inflammation-induced cancers. Nitration of the guanine base also occurs in the ribonucleotide pool and 8-nitroguanosine-3',5'-cyclic phosphate^[11] (8-nitro-cGMP) has attracted attention as an electrophilic second messenger that is involved in an unusual post translational modification of proteins. This process of protein S-ganylation,^[11] results from the nucleophilic sulfhydryl group of a cysteine residue displacing the nitro group of 8-nitro cGMP. More recently H₂S, a proposed gaseous transmitter in mammals, has been shown to react with and sequester 8-nitro-cGMP, and provides a potential mechanism to shut down this electrophile-mediated signalling pathway.^[12]

Collectively, these diverse studies have shown that the chemical properties of 8-nitroguanine nucleotides are inherent-

ly related to their biological function and have encouraged us to look more closely at their chemistry. We now report our studies on some chemical reactions performed on 8-nitroguanine nucleosides and oligonucleotides. The results shed some light on how the 8-nitro substituent affects the reactivity of the purine ring, they demonstrate that the 8-nitroguanine lesion in DNA can be fluorescently labelled and raise the possibility of a repair mechanism for this lesion.

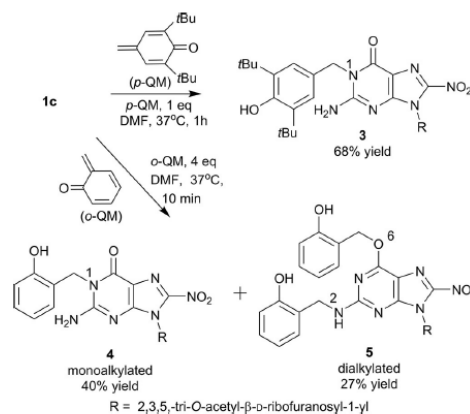
Results and Discussion

Alkylation reactions

Previous studies have demonstrated that some chemical properties of guanine nucleosides are significantly altered by introduction of the 8-nitro group. The nitro group greatly enhances the hydrolytic lability of the glycosidic bond^[4] and also significantly lowers the pK_a of the proton on $N1$ ($pK_a = 8.4$ for 8-nitroguanosine compared to 9.4 for guanosine).^[8] In addition, we have made attempts to introduce the usual acyl or formamidine protecting groups at the $N2$ position of 8-nitroG nucleosides, but were unable to prepare these derivatives using routine procedures.^[8] This failure to functionalise the $N2$ position implied that the reactivity of the guanine base is significantly reduced because of the strongly electron-withdrawing nitro group and suggested that the chemistry of 8-nitroguanine nucleosides required wider exploration.

We chose to investigate alkylation reactions on 2',3',5'-tri-*O*-acetyl-8-nitroguanosine (**1c**, Figure 1), as in this substrate the electron withdrawing acetyl groups, particularly on the 2'-OH, help to further stabilise the glycosidic bond. Given the seemingly inert nature of the 8-nitroguanine base, *ortho*- and *para*-quinone methides were selected as alkylating agents as they possess a benzylic carbon which is highly electron deficient and extremely susceptible to nucleophilic attack.^[13,14] In addition, the products of the reactions of these quinone methides with 2'-deoxyguanosine have been well characterised.^[13,14]

In initial experiments, acetylated 8-nitroguanosine (**1c**) was reacted with a molar equivalent of the *para*-quinone methide (*p*-QM) prepared from 2,6-bis(*tert*-butyl)-4-methylphenol by oxidation with silver(I) oxide (Scheme 1).^[13] Under these conditions TLC and HPLC showed that a single nucleoside product (**3**) was formed and was subsequently isolated in 68% yield following column chromatography. Mass spectrometry ascertained that (**3**) was a monoalkylated product and $N1$ was established as the position of alkylation through NMR techniques. DEPT 135 established the presence of two CH_2 groups and through HSQC and COSY experiments, benzylic protons from the quinone methide adduct were identified as a singlet at 5.11 ppm. In the ¹³C NMR spectrum the benzylic signal resonated at 45.7 ppm, characteristic of a carbon-nitrogen bond; a chemical shift of approximately 70 ppm would be expected should the benzylic carbon have been attached to an oxygen atom. From the ¹H NMR it was possible to determine which nitrogen atom was alkylated. Acquisition of the spectra in a non-protic solvent indicated the absence of the signal corresponding to the $N1$ proton which was present in the starting materi-



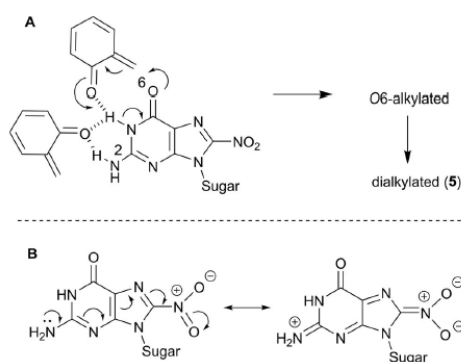
Scheme 1. Reactions of 2',3',5'-tri-*O*-acetyl-8-nitroguanosine with quinone methides. Experimental details in the Supporting Information.

al at 11.34 ppm. Additionally, the broad peak at 6.06 ppm corresponding to the $N2$ signal integrated to 2 protons. These observations confirm that alkylation occurred at the $N1$ position. This assignment was further strengthened when compared to the $N2$ -alkylated product previously characterised from the reaction between deoxyguanosine and a *para*-quinone methide,^[13] in which these observations were reversed; the peak at 10.35 ppm associated with the $N1$ proton was present and the signal associated with the $N2$ proton integrated as a single proton.

We also examined the alkylation of acetylated 8-nitroguanosine with an *ortho*-quinone methide (*o*-QM) produced from 2-bromomethyl-*O*-*tert*-butyldimethylsilylphenol using potassium fluoride (Scheme 1).^[15,16] In an initial experiment with one equivalent of the *o*-QM, two products were observed by HPLC; a more polar product ($\approx 40\%$) and a less polar product ($\approx 5\%$). When four molar equivalents of *o*-QM were used a more even distribution of the products was obtained and the more polar and less polar products were isolated in yields of 40 and 27%, respectively. Mass spectrometry revealed the more polar (**4**) and less polar (**5**) products to be mono-alkylated and di-alkylated species, respectively (Scheme 1). NMR techniques revealed the mono-alkylated product (**4**) to be entirely consistent with the $N1$ -*ortho* quinone methide, based on similarities to the $N1$ -*para* quinone methide product (**3**).

For the dialkylated product (**5**), the presence of two non-nucleosidic CH_2 groups with chemical shifts of 67.4 ppm and 40.0 ppm in the ¹³C NMR spectrum, was established. These resonances are characteristic of a benzylic carbon-oxygen bond and a benzylic carbon-nitrogen bond, respectively and therefore indicated that $O6$ is one of the alkylation sites.^[14,17] With regard to the second alkylation site on nitrogen, alkylation at the $N1$ position seemed unlikely. This was supported by ¹H NMR spectra recorded in a non-protic solvent revealed the presence of a single *NH* signal which integrated to one proton.

The results of alkylation reactions performed here with 2',3',5'-tri-*O*-acetyl-8-nitroguanosine (1c) show significant differences to those previously reported for 2'-deoxyguanosine (dG), both in terms of the products formed and their distribution. For example, the sole nucleoside product isolated from alkylation of (1c) with *p*-QM is the *N*1 adduct (3), whereas previous results obtained with dG show *N*2 to be the favoured site of alkylation.¹⁴ That the 8-nitro group leads to preferential reaction at *N*1 rather than *N*2 is not so surprising given that the lone pair on *N*2 is directly conjugated through to the nitro group (Scheme 2) and that reaction at *N*1 of the nitrated nucleoside might be enhanced by deprotonation, due to its significantly lower *p*K_a.



Scheme 2. Panel A shows *O*6 versus *N*2 hydrogen bond-directed alkylation, resulting in initial alkylation at the presumably more reactive *O*6 position and subsequent alkylation at *N*2, to give dialkylated 5. Panel B shows reduced nucleophilicity of *N*2 because of conjugation through to the nitro group.

The reaction of excess *o*-QM with nitrated nucleoside (1c) once again gave the *N*1 adduct (4) as the major product (40%), together with the *O*6,*N*2-dialkylated product (5, 27%). Alkylation at *N*1 was expected based on the result obtained with *p*-QM, particularly since the use of excess potassium fluoride to generate the *o*-QM in unbuffered conditions, would be expected to lead to mildly basic conditions that favour reaction at *N*1 due to deprotonation. The *O*6,*N*2-dialkylated product was unexpected, since no *O*6 or *N*2 mono-alkylated products were observed. The *ortho* geometry of the *o*-QM means that hydrogen-bonding of *NH*1 with the carbonyl group of the quinone methide could assist nucleophilic attack by either *O*6 or *N*2 (Scheme 2). Given that *N*2 appears to be highly deactivated, attack by *O*6 is likely to be favoured from the hydrogen-bonded complex. To explain why no *O*6 mono-alkylated product is observed, the subsequent alkylation on *N*2, would have to be relatively fast and might be assisted by the alkylated and therefore electron donating *O*6. In contrast, studies on the reaction between dG and *o*-QM have been shown to give *N*2 alkylation as either the sole product¹⁸ or to favour *N*2 alkylation over *N*1 by a factor of about 5.¹⁴

The alkylation reactions reported here have been performed on 2',3',5'-tri-*O*-acetyl-8-nitroguanosine in order to reduce depurination and were also optimised to generate good yields of the products to aid their characterisation, therefore the reaction outcomes cannot be directly compared with quantitative studies previously obtained with dG. However, given that the sites of alkylation for DNA bases with quinone methides remain constant in their deoxynucleoside, single-stranded oligodeoxynucleotide and duplex polydeoxynucleotide forms,¹⁸ the presence of the acetyl protecting groups on the sugar is not expected to alter the site of alkylation. Thus, the results presented here indicate that the reactivity of the guanine base is significantly altered by the introduction of the 8-nitro group; the reactivity of the *N*2 position significantly decreases whilst *N*1 appears to be enhanced, presumably due to its considerably lower *p*K_a.

Reductive denitration of 8-nitroguanosine

Aromatic nitro compounds such as nitroimidazoles have emerged as potent *anti*-tuberculosis pro-drugs that exert their activity through the release of nitric oxide (NO), which is the effective agent.¹⁹ NO released from these drugs is initiated by nucleophilic attack of a hydride equivalent from a deazaflavin cofactor that is present in a deazaflavin-dependent nitroreductase and the overall reaction is a denitration in which NO₂ is replaced by H.¹⁹ A related reaction performed on the 8-nitroG lesion in DNA, in which a hydride equivalent was delivered to the C8 position, could remove the nitro group through an addition-elimination reaction, thus suggesting a potential DNA repair mechanism by directly reversing the lesion. Relatively few DNA lesions are repaired by direct chemical reversal, but this route is well established for the repair of UV-induced pyrimidine dimer lesions²⁰ and the repair of guanine *O*6 alkylation products by *O*6-alkylguanine DNA alkyltransferase.²¹ We were therefore interested in investigating the chemical feasibility of reducing 8-nitroguanosine back to guanosine as this could suggest a potential repair mechanism for this lesion by direct chemical reversal.

There are a few examples where borohydride has been used to effect aromatic denitration and these have been reported for both polysubstituted nitrobenzenes^{22,23} and heterocyclic systems.²⁴ These reactions take place through an addition elimination mechanism in which nitrite is displaced by hydride.²³ Earlier studies^{14,25} on the characterisation of the 8-nitroguanine base have shown that reducing agents such as: sodium hydrosulfite, zinc–HCl, hydrazine hydrate–Raney nickel; all yielded the expected 8-aminoguanine, but to our knowledge the reaction has not been studied with hydride reducing agents.

Prior to studying the reduction, 8-nitroguanosine (1b, Figure 1) was purified by reverse-phase (RP)-HPLC, although as can be seen in Figure 2, a small amount of the depurination product 8-nitroguanine was unavoidably introduced into the starting samples of 8-nitroguanosine during evaporation of the fractions. It was also necessary to develop RP-HPLC conditions that would separate the two most likely products; namely 8-

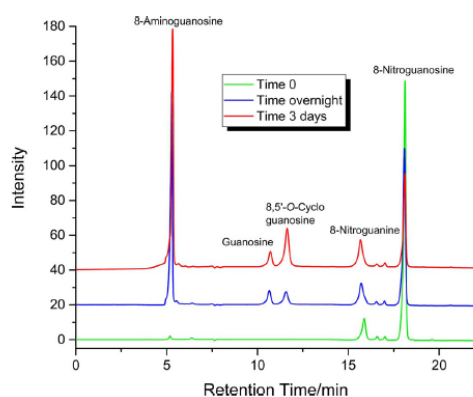
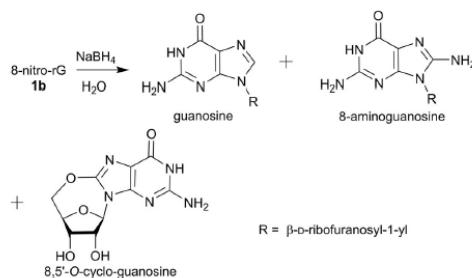


Figure 2. RP-HPLC analysis of the reduction of 8-nitroguanosine with sodium borohydride. Details of the reaction conditions are described in the Experimental section. HPLC chromatograms were obtained using an elution gradient of 0–20% acetonitrile in 0.1% aqueous acetic acid over 20 min, with a flow rate of 0.5 mL·min⁻¹. Chromatograms were recorded at 254 nm.

aminoguanosine and guanosine. An extensive investigation of eluents revealed that 8-aminoguanosine and guanosine were readily separated using a gradient of acetonitrile in aqueous acetic acid (0.1% AcOH) and these elution conditions were also compatible with HPLC-MS.

Reduction reactions (Scheme 3) were generally performed in unbuffered distilled water using a very large excess of sodium borohydride that was repeatedly added at regular intervals to compensate for hydrolysis. Under these conditions, a slow reduction occurred (Figure 2) that gave 8-aminoguanosine as the major product together with a small amount (8–10%) of guanosine (Scheme 3). The identity of both products was confirmed by co-elution with authentic samples of guanosine and 8-aminoguanosine and also by HPLC-MS analysis (see Figure 2, Table 1 and Supporting Information). In addition to the reduction products, an additional nucleoside product (≈20%) was observed that was identified by HPLC-MS as 8,5'-O-cycloguanosine and presumably results from displacement of the nitro group by the 5'-hydroxyl group. Formation of the cyclo-



Scheme 3. Reduction of 8-nitroG with sodium borohydride.

Table 1. High-resolution mass data for the borohydride (either NaBH₄ or NaBD₄) reduction of 8-nitroguanosine as shown in Scheme 3. Accurate mass measurements were obtained by LC-MS as described in the experimental section. Mass spectra for individual compounds are shown in the Supporting Information.

Nucleoside	Formula	Calculated exact mass for [M-H] ⁻	Measured accurate mass for [M-H] ⁻
guanosine	C ₁₀ H ₁₃ N ₅ O ₅	282.0844	282.0827
8-deuteroguanosine	C ₁₀ H ₁₂ DN ₅ O ₅	283.0907	283.0906
8,5'-O-cycloguanosine	C ₁₀ H ₁₁ N ₅ O ₅	280.0687	280.0693
8-aminoguanosine	C ₁₀ H ₁₄ N ₆ O ₅	297.0953	297.0963

cleoside is probably facilitated by the rise in pH caused by hydrolysis of the borohydride.

Further proof for the direct displacement of the nitro group by hydride was obtained by performing the reaction using NaBD₄ in H₂O, and in this case HPLC-MS revealed the guanosine fraction to have the correct mass for the incorporation of one deuterium atom (see Table 1). As expected, the peak corresponding to 8-aminoguanosine showed no inclusion of deuterium as any incorporated during reduction would exchange through protonation.

In an attempt to increase the proportion of guanosine, the reduction was investigated under a variety of conditions. Using buffered aqueous conditions between pHs 6.0–9.0 (0.1 M ammonium acetate) there was little difference in terms of the product distribution from those performed in water, although the reaction at lower pH did progress slightly faster. It is not surprising that all the aqueous reduction reactions gave similar results as the very large excess of borohydride, which raises the pH as it is hydrolysed, resulted in all these reactions becoming alkaline as the reaction proceeded (pH ≈ 9.6 at the end of the reaction). No reduction products were observed when the weaker reducing agent sodium cyanoborohydride was used at pH 6.0. The work of Lamson^[23a] has established that the denitration of polysubstituted nitrobenzenes proceeded more rapidly in DMSO than in protic solvents. However, when the reaction was performed in DMSO we were unable to detect any reduction products, even during a prolonged reaction (> 1 week) and only a trace amount of 8,5'-O-cycloguanosine was formed.

Although we have been unable to increase the proportion of guanosine above about 12%, its formation by displacement of the nitro group by a hydride equivalent, in aqueous solution, does clearly show that a direct repair of the 8-nitroguanine lesion is a chemically feasible reaction.

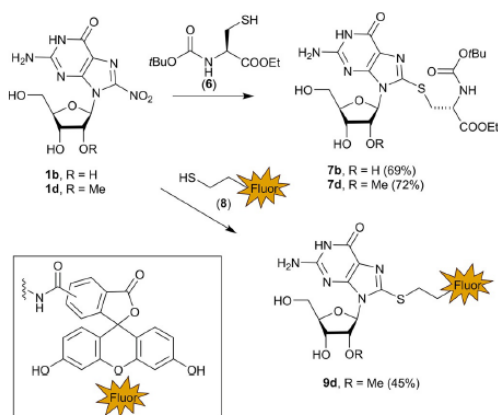
Detection of 8-nitroguanine through the displacement of the nitro group by thiols

Methods have been developed to detect the 8-nitroguanine base in tissues^[9,10] and urine,^[20] but these methods are only applicable once the 8-nitroguanine base has been hydrolytically released from DNA and they give no information as to the

steady-state levels of this lesion in DNA samples. We have recently shown that surface-enhanced Raman spectroscopy is a very sensitive technique for detecting this lesion due to the nitro group and can distinguish between the 8-nitroguanine that is covalently bound to DNA or has been released by hydrolysis.^[27]

The displacement of the nitro group in 8-nitroguanine nucleosides by sulfur nucleophiles is well established and specific examples include the capture of 8-nitroguanosine by the nucleoside derivative nitroG-Grasp,^[28] sulfhydrylation^[12] and protein S-guanylation^[29] of 8-nitro-cGMP. The mild conditions for these displacement reactions suggested that it might be possible to develop a detection procedure for this lesion based on displacement of the nitro group with a fluorescent thiol. In theory, each lesion site would be converted to a fluorescently labelled guanine base that would be stable to depurination (due to the loss of the strongly electron-withdrawing nitro group) and thus create a permanent and quantifiable record of the lesion sites, which in theory could be identifiable through sequencing methods.

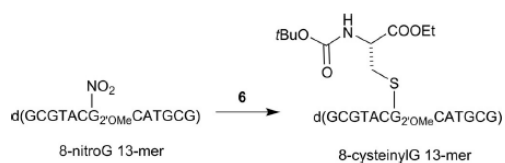
As 8-nitro-dG (1a), which is the naturally occurring lesion in DNA readily undergoes depurination, our initial experiments to explore this labelling technique were performed on 8-nitroguanosine (1b, Figure 1) and 8-nitro-2'-O-methylguanosine (1d, Figure 1). Both of these ribonucleosides have a more stable glycosidic bond than their 2'-deoxy counter part, due to the electronegative 2'-oxygen. The 2'-O-methylriboside is also ideally suited as a means to incorporate this lesion into oligonucleotides, as no protection of the 2'-position is necessary during DNA synthesis and it has previously been used as a model for 8-nitro-dG in oligonucleotides.^[8] We chose to use *N*-(*tert*-butoxycarbonyl)cysteine ethyl ester (6, Scheme 4) as a model thiol as its hydrophobicity was expected to significantly increase the retention time of the resulting nucleoside/oligonucleotide adducts on reverse-phase HPLC and its low extinction coefficient at 260 nm, was not expected to significantly in-



Scheme 4. Reaction of thiols with 8-nitroguanosine nucleosides. Compound 8 is shown as a mixture of the 5(6)-regioisomers.

terfere with the HPLC analysis. Thus, reaction of 8-nitroguanosine with a slight excess of cysteine thiol (6) in aqueous buffer (pH 7.5), gave the S-guanylated cysteine derivative (7b) in good yield (69%).^[30] Analysis of ¹H-¹³C HMBC contour map revealed a long range ¹H-¹³C correlation between the C8 resonance and the protons adjacent to the sulfur atom, thus confirming that the sulfur atom had displaced the nitro group. A similar yield (72%) was obtained for the conversion 8-nitro-2'-O-methylguanosine (1d) to its corresponding 8-S-cysteinylnucleoside derivative (7d, Scheme 4).

To apply this labelling technique to model DNA sequences, 8-nitro-2'-O-methylguanosine (1d) was incorporated into 13-mer, 27-mer and 37-mer oligodeoxynucleotides. These model sequences each contained a single 8-nitroguanine base and were prepared using previously reported procedures.^[8] Reaction of the nitro 13-mer sequence (Scheme 5) with a 100-fold



Scheme 5. Reaction of nitro-13-mer with cysteine nucleophile 6.

excess of cysteine thiol (6), in Tris-HCl buffer (pH 8.5) was monitored by RP-HPLC. A relatively smooth reaction was observed with greater than 90% conversion to the less polar cysteinylnucleoside 13-mer after 6 hours, together with formation of a small amount of the depurinated 13-mer, which runs slightly faster than the nitro-13-mer (Figure 3, Table 2 for retention times.).

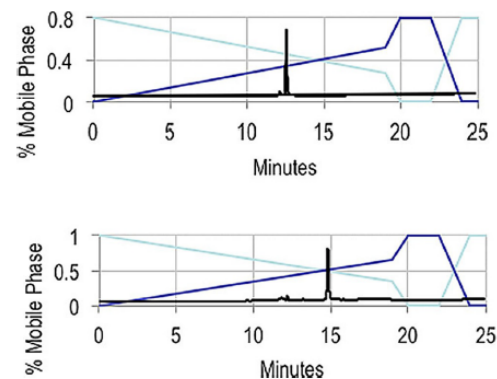


Figure 3. RP-HPLC chromatogram showing reaction of 8-nitro-13-mer with cysteine derivative 6. Top panel shows the nitro-13-mer starting material and the bottom panel shows the reaction mixture after 6 hours. RP-HPLC was performed as described in the experimental section using a linear gradient of 0% to 65% eluent B (blue line) in eluent A, over 19 min, with a flow rate of 1.0 mL min⁻¹. Eluent A = aq TEAB (0.1 M), eluent B = 40% CH₃CN in aq TEAB (0.1 M). Chromatograms were recorded at 254 nm.

Table 2. Reverse-phase HPLC retention times (min) for nitro oligomers and their reaction products with thiol nucleophiles. HPLC analysis conditions are as described in the legend to Figure 3. Nitro-13-mer = d(GCGTACGATGCG); nitro-27-mer = d(GTGACTGGGAAAACCCCTGGCGTTACCC); nitro-37-mer = d(GCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG). G = position of 8-nitro-2'-OMe-guanosine.

Oligomer length	nitro-oligomers	cysteinyloligomers	8-fluoro oligomers
13-mer	12.55	14.78	18.63
27-mer	12.82	14.13	–
37-mer	12.52	13.74	–

Table 3. Mass data for all the 13-mer oligonucleotides, obtained by ES ionisation as described above. 13-mer sequence = d(GCGTACGATGCG), G = position of modification.

Oligonucleotide	Calculated exact mass	Measured accurate mass
8-nitroG-13mer	4050.7025	4050.6133
cysteiny 13-mer	4250.7996	4250.6850
fluoroscein 13-mer	4438.7810	4438.7090

The cysteinyl 13-mer was isolated by RP-HPLC and was shown by negative mode electrospray mass spectrometry to have the expected mass (Table 3 in the Experimental section).

The cysteine nucleophile (6) was also used in displacement reactions on the longer 27-mer and 37-mer sequences, which are more representative of the naturally occurring 8-nitroG lesion in DNA. These reactions were carried out under the same conditions as those described for the nitro 13-mer and although they proceeded more slowly, >90% conversion was observed after 11 hours. Both the nitro 27-mer and nitro 37-mer were predominantly converted to a single product with the expected longer retention times when compared to the nitro-containing oligonucleotides (see Table 2 for retention times). As we were unable to obtain mass spectrometry data for the products from these longer oligomers, the presence of the 8-S-cysteinyl group in the 27-mer and 37-mer was confirmed by enzymatic digestion. Treatment of the product oligomers with phosphodiesterase and alkaline phosphatase (pH 8.5), followed by HPLC analysis showed the presence of 7d in the digestion mix, together with the four naturally occurring 2'-deoxynucleosides. When the chromatogram from the digestion mix was analysed to obtain the composition ratios of the component nucleosides, the peak area for 7d was lower than expected and an unidentified peak, with the UV absorption characteristics of a nucleoside (retention time 9.10 min) was also observed. It was initially thought that the unidentified peak might be a partially digested dinucleotide, but its proportion in the digestion mixture increased as the digestion time increased. Given that the original cysteine nucleophile (6, retention time = 19.10 min) could also be detected in the digestion mix, the most likely cause of this unidentified nucleoside was hydrolysis of the 8-S-cysteinyl derivative (7d) to give 8-oxo-2'-O-methylguanosine (2b, Figure 1) and releasing (6). A

standard of 8-oxo-2'-O-methylguanosine was prepared^[31] and it was shown to have the same retention time on HPLC as the unidentified nucleoside in the digestion mixture, thus implying that 8-S-cysteinyl derivative (7d) was breaking down in the digestion mixture.

Finally, we wanted to use a fluorescent thiol derivative in order to demonstrate the labelling of the 8-nitroG lesion with a readily quantifiable reporter group. We chose to use a cystamine derivative of carboxyfluorescein (8), which had previously been prepared as a mixture of regioisomers and used to label cell-penetrating oligoguanidinium compounds.^[30] This fluorescent thiol (8) was subsequently used to displace the nitro group in 8-nitro-2'-O-methylguanosine (1d), using similar conditions to those described for the preparation of (7d, Scheme 4). As the fluorescent thiol (8) is a mixture of regioisomers, the resulting fluorescein adduct formed with 8-nitro-2'-O-methylguanosine (9d) was more difficult to isolate and characterise than the 8-S-cysteinyl derivatives, but the identity of the product was confirmed by high resolution mass spectrometry. The fluorescent thiol (8) was also used to label the nitro 13-mer and the resulting fluorescein-tagged 13-mer was isolated by HPLC and identified by electrospray mass spectrometry.

The mild conditions and clean reaction of the cysteine thiol (6) with the nitro 13-mer (Figure 3) and longer sequences, clearly indicates the potential for using this reaction to label the site of 8-nitroG lesions in DNA. However, whilst the hydrolysis of the 8-S-cysteinyl guanosine derivative (7d) during the enzymatic digestion was not evident when analysing and manipulating the 8-S-cysteinyl oligodeoxynucleotides, this reaction does suggest that the labelling conditions might have to be optimised to eliminate hydrolysis to the 8-oxo guanine nucleoside. The observed hydrolytic breakdown of 5-guanylated nucleosides does suggest a potential additional source of 8-oxoguanosine nucleosides/nucleotides in biological systems.

Conclusion

These studies reveal that the 8-nitro group exerts a significant effect on the reactivity of the guanine base. In particular, the reactivity of the guanine NH₂ group towards electrophiles is significantly reduced, whilst the reactivity of the NH group appears to be enhanced, due to its lower pK_a. The ease with which the nitro group is replaced by thiol nucleophiles under mild conditions and the high efficiency of this reaction show that sites of guanine nitration in DNA can be fluorescently labelled using suitably derivatised thiols. The derivatisation process leaves a permanent marker of the original lesion that can be quantified. By using the appropriate thiol (e.g. a thiol-derivatised crown ether) it should be relatively easy to adapt this labelling technique so that the 8-nitroguanine base is converted into a current-modulating marker that can be read through nanopore sequencing.^[32]

Finally, we have shown that reduction of 8-nitroguanosine with sodium borohydride in aqueous solution gives 8-amino-guanosine as the major product, but is also accompanied by a denitration reaction to give guanosine. The formation of gua-

nosine, in the presence of a hydride equivalent, demonstrates that the direct repair of the 8-nitroguanine lesion is a chemically feasible reaction.

These results significantly enhance our knowledge of the chemical properties of this base modification and provide new methods to detect and study this lesion.

Experimental Section

8-Nitroguanosine (1 b) and 8-nitro-2'-*O*-methylguanosine (1 d) were prepared as previously described.^[8] 2'3'5'-Tri-*O*-acetyl-8-nitroguanosine (1 c) was prepared as described by Saito et al.^[29b] *N*-(*tert*-butoxycarbonyl)cysteine ethyl ester (6) was prepared by esterification of *N*-(*tert*-butoxycarbonyl)cystine.^[33] 8-Oxo-2'-*O*-methylguanosine was prepared from 8-bromo-2'-*O*-methylguanosine^[35] following the procedure of Nampalli and Kumar.^[31] 5(6)-Carboxyfluorescein was purchased from Sigma-Aldrich as a mixture of 5- and 6-carboxy regioisomers.

All RP-HPLC was performed using Gemini™ C18 columns (110 Å, 250 mm x 4.6 mm) purchased from Phenomenex. RP-HPLC analysis of oligonucleotides was performed on an automated Gilson HPLC system equipped with an autoinjector, a photodiode array detector and dual hydraulic pump, operating with a flow rate 1 mL min⁻¹. Chromatograms were recorded at 254 nm. Chromatographic data were controlled and processed using UniPoint® Version 3.0. All oligonucleotides were purified by RP-HPLC using a triethylammonium bicarbonate (TEAB) buffer system. A 1 M stock solution of TEAB was prepared by bubbling CO₂ through a 1 M aqueous solution of Et₃N for 6–8 hours until the pH measured 7.5–7.6. Stock solutions of TEAB were refrigerated until needed. All UV measurements were taken on a PerkinElmer Lambda 25 spectrophotometer.

RP-HPLC analysis of 8-nitroguanosine reduction reactions were performed on an Agilent 1260 Infinity system equipped with an autoinjector, a photodiode array detector and quaternary pump. Chromatographic data were controlled and processed using Agilent Open LAB Chemstation software. An elution gradient of 0–20% acetonitrile in 0.1% aqueous acetic acid over 20 min, with a flow rate of 0.5 mL min⁻¹. Chromatograms were recorded at 254 nm.

LC-MS was performed on an Agilent 6530B accurate mass Q-TOF G6530B mass spectrometer connected to an Agilent 1260 Infinity HPLC system as described above, with a flow rate of 0.5 mL min⁻¹. LC-MS data was obtained using a multimode ion source and processed using Agilent Masshunter software. Spectra were recorded in negative ionisation mode.

All NMR spectra were recorded on a Bruker 400 MHz spectrometer; ¹H (400 MHz) and ¹³C (100 MHz). All chemical shifts are reported in p.p.m. and coupling constants (J) in Hz. ¹³C spectra were ¹H decoupled. ¹H and ¹³C chemical shifts are relative to an internal standard of tetramethylsilane. Mass spectra of nucleosides and chemically synthesised ODNs were recorded on a MicroMass LCT mass spectrometer using electrospray ionisation (ES) and direct infusion syringe pump sampling. All small molecules were dissolved in methanol. Oligonucleotides were dissolved in water and analysed using negative mode ES ionisation.

Oligodeoxynucleotide synthesis and purification

Synthesis and purification of 13- and 27- and 37-mer oligodeoxynucleotides containing 8-nitro-2'-*O*-methyl-guanosine was performed as previously described.^[8] All oligodeoxynucleotides containing the 8-nitroguanine modification showed the characteristic absorption at 395 nm.

Borohydride reduction of 8-nitroguanosine

8-Nitroguanosine (1.0 mg, 3.05 μmoles) was dissolved in distilled water (1 mL) with stirring. Sodium borohydride (5.8 mg, 0.153 mmoles, 50 equiv) was added, with additional portions (50 equiv) subsequently added every 30 mins for 90 mins. Reaction monitored by HPLC, with samples taken every 30 mins.

8-Cysteinylyl labelled oligomers

8-cysteinylyl 13-mer = d(GCGTACG^(tBu-Cys)CATGCG) 8-cysteinylyl 27-mer = d(GTGACTGG^(tBu-Cys)AAAACCCCTGGCG-TTACCC) 8-cysteinylyl 37-mer = d(GCGCAACGCAATTAATGTG^(tBu-Cys)AGT-TAGCTCACTCAT-TAG)

To a solution of the appropriate 8-nitro oligomer (3.0 OD (260 nm) units) dissolved in distilled H₂O (100 μl) was added 1 M TRIS.HCl (pH 8.5, 10 μl) and ethyl *N*-(*tert*-butoxycarbonyl)-cysteinate (6, 0.80 mg, >100 equiv). The resulting mixture was stirred at 37 °C for 6 hours during which time the progress of the reaction was monitored by RP-HPLC using a gradient of 0–26% acetonitrile in TEAB (0.1 M). Retention times for the oligonucleotide substrates and products are presented in Table 2. Oligonucleotide products were purified using the same gradient and the fractions containing the products were evaporated 3 times with distilled water in a vacuum centrifuge.

8-Fluorescein labelled oligomer

8-Fluorescein 13-mer = d(GCGTACG^(Fluoro)CATGCG). To a solution of the 8-nitro 13-mer (3.1 OD units) dissolved in distilled H₂O (100 μl) was added 1 M TRIS.HCl (pH 8.5, 10 μl) and 3',6'-dihydroxy-3-oxo-*N*-(2-sulfanylethyl)-3*H*-spiro[2-benzofuran-1,9'-xanthene]-5(6)-carboxamide (8, 1.7 mg). The resulting mixture was stirred at 37 °C for 24 hours during which time the progress of the reaction was monitored by RP-HPLC using a gradient of 0–26% acetonitrile in TEAB (0.1 M). Retention times for the oligonucleotide substrates and products are presented in Table 2. Oligonucleotide products were purified using the same gradient and the fractions containing the products were evaporated 3 times with distilled water in a vacuum centrifuge.

Acknowledgements

MM was supported by the Engineering and Physical Sciences Research Council (EPSRC) (research grant EP/H002464/1) and KJA and KRW were funded by EPSRC studentships. We would like to thank the University of Liverpool MicroBioRefinery facility for access to LC-MS and to Stephen Moss for LC-MS analysis. We would also like to thank Alex Bretherton for performing preliminary studies on the reduction of 8-nitroguanosine.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biological chemistry · DNA lesion · DNA modification · guanine · oligonucleotides

- [1] S. Steenken, S. V. Jovanovic, *J. Am. Chem. Soc.* 1997, 119, 617.
[2] S. Kawanishi, Y. Hiraku, *Antioxid. Redox Signaling* 2006, 8, 1047.
[3] J. Cadet, T. Douki, J.-L. Ravanat, *Acc. Chem. Res.* 2008, 41, 1075.

- [4] V. Yermilov, J. Rubio, H. Ohshima, *Febs. Lett.* **1995**, *376*, 207.
- [5] S. Obeid, N. Blatter, R. Kranaster, A. Schnur, K. Diederichs, W. Welte, A. Marx, *EMBO J.* **2010**, *29*, 1738.
- [6] N. Suzuki, M. Yasui, N. E. Geacintov, V. Shafirovich, S. Shibutani, *Biochemistry* **2005**, *44*, 9238–9245.
- [7] K. Kaneko, T. Akuta, T. Sawa, H. W. Kim, S. Fujii, T. Okamoto, H. Nakayama, H. Ohgashi, A. Murakami, T. Akaike, *Cancer Lett.* **2008**, *262*, 239.
- [8] I. Bhamra, P. Compagnone-Post, I. A. O'Neil, L. Iwanejko, A. D. Bates, R. Cosstick, *Nucl. Acids Res.* **2012**, *40*, 11126.
- [9] a) Y. Hiraku, *Environ. Health Prev. Med.* **2010**, *15*, 63; b) Y. Hiraku, K. Sakai, E. Shibata, M. Kamijima, N. Hisanaga, N. Ma, S. Kawanishi, M. Murata, *J. Occup. Health* **2014**, *56*, 186–196.
- [10] Y. Hoki, M. Murata, Y. Hiraku, N. Ma, A. Matsumine, A. Uchida, S. Kawanishi, *Oncol. Rep.* **2007**, *18*, 1165.
- [11] T. Sawa, M. H. Zaki, T. Okamoto, T. Akuta, Y. Tokutomi, S. K. Mitsuyama, H. Ihara, A. Kobayashi, M. Yamamoto, S. Fujii, H. Arimoto, T. Akaike, *Nat. Chem. Biol.* **2007**, *3*, 727.
- [12] V. Terzić, D. Padovani, V. Balland, I. Artaud, E. Galardon, *Org. Biomol. Chem.* **2014**, *12*, 5360.
- [13] M. A. Lewis, D. Graff, J. L. Bolton, J. A. Thompson, *Chem. Res. Toxicol.* **1996**, *9*, 1368.
- [14] W. F. Veldhuyzen, Y.-F. Lam, S. E. Rokita, *Chem. Res. Toxicol.* **2001**, *14*, 1345.
- [15] T. Li, S. E. Rokita, *J. Am. Chem. Soc.* **1991**, *113*, 7771.
- [16] S. E. Rokita, J. Yang, P. Pande, W. A. Greenberg, *J. Org. Chem.* **1997**, *62*, 3010.
- [17] J. Reinhard, W. E. Hull, C. W. Von Der Lieth, U. Eichhorn, H. C. Klier, B. Kaina, M. Wiessler, *J. Med. Chem.* **2001**, *44*, 4050.
- [18] P. Pande, J. Shearer, J. Yang, W. A. Greenberg, S. E. Rokita, *J. Am. Chem. Soc.* **1997**, *119*, 6773.
- [19] a) R. Singh, U. Manjunatha, H. I. Boshoff, Y. H. Ha, P. Niyomrattanakit, R. Ledwidge, C. S. Dowd, I. Y. Lee, P. Kim, L. Zhang, *Science* **2008**, *322*, 1392; b) R. Tiwari, G. C. Moraski, V. Krchňák, P. A. Miller, M. Colon-Martinez, E. Herrero, A. G. Oliver, M. J. Miller, *J. Am. Chem. Soc.* **2013**, *135*, 3539.
- [20] A. Sancar, *Chem. Rev.* **2003**, *103*, 2203.
- [21] a) A. E. Pegg, M. E. Dolan, R. C. Moschel, *Prog. Nucleic Acid Res. Mol. Biol.* **1995**, *51*, 167; b) S. Mitra, *DNA Repair* **2007**, *6*, 1064.
- [22] L. A. Kaplan, *J. Am. Chem. Soc.* **1964**, *86*, 740.
- [23] a) D. W. Lamson, P. Ulrich, R. O. Hutchins, *J. Org. Chem.* **1973**, *38*, 2928; b) V. Gold, A. Y. Miri, S. R. Robinson, *J. Chem. Soc. Perkin Trans. 2* **1980**, 243.
- [24] A. N. Blinnikov, N. N. Makhova, *Mendeleev Commun.* **1999**, *9*, 13.
- [25] V. Yermilov, J. Rubio, M. Becchi, M. D. Friesen, B. Pignatelli, H. Ohshima, *Carcinogenesis* **1995**, *16*, 2045.
- [26] T. Sawa, M. Tatemichi, T. Akaike, A. Barbin, H. Ohshima, *Free Radical Biol. Med.* **2006**, *40*, 711.
- [27] S. Dick, S. E. J. Bell, K. J. Alexander, I. A. O'Neil, R. Cosstick, *Chem. Eur. J.* in press, <https://doi.org/10.1002/chem.20170179>.
- [28] Y. Fuchi, S. Sasaki, *Org. Lett.* **2014**, *16*, 1760.
- [29] a) Y. Ishima, H. Hoshino, T. Shinagawa, K. Watanabe, T. Akaike, T. Sawa, U. Kragh-Hansen, T. Kai, H. Watanabe, T. Maruyama, M. Otagiri, *J. Pharm. Sci.* **2012**, *101*, 3222; b) Y. Saito, H. Taguchi, S. Fujii, T. Sawa, E. Kida, C. Kabuto, T. Akaike, H. Arimoto, *Chem. Commun.* **2008**, 5984.
- [30] Similar yields were obtained when glutathione was used as the nucleophile.
- [31] S. Nampalli, S. Kumar, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1677.
- [32] N. An, A. M. Fleming, H. S. White, C. J. Burrows, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11504.
- [33] J. Fernández-Carneado, M. Van Gool, V. Martos, S. Castel, P. Prados, J. De Mendoza, E. Giral, *J. Am. Chem. Soc.* **2005**, *127*, 869.

Manuscript received: November 21, 2017

Accepted manuscript online: January 3, 2018

Version of record online: February 5, 2018