UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering, Science and Mathematics

School of Chemistry

Novel Nucleotide Analogues for Forming Stable DNA Triple Helices

by

Simon Richard Gerrard

A thesis submitted for the degree of Doctor of Philosophy.

June 2009

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS
SCHOOL OF CHEMISTRY

Doctor of Philosophy

NOVEL NUCLEOTIDE ANALOGUES FOR FORMING STABLE DNA TRIPLE HELICES

by Simon Richard Gerrard

DNA triple helices are an important tool in a variety of medicinal and biotechnological applications, such as gene therapy and chemotherapeutics. DNA triple helices are formed by binding of a triplex-forming oligonucleotide (TFO) to a DNA duplex, *via* specific recognition of the individual base pairs in the target sequence.

Mixed-sequence recognition of duplex DNA by TFOs is therefore an essential requirement for successful targeting. However, achieving strong, yet specific binding to the pyrimidine.purine (Py.Pu) base pairs CG and TA, by TFOs is a greater challenge than to the purine.pyrimidine (Pu.Py) base pairs (GC, AT), as fewer hydrogen bonds are presented for binding in the major groove of the double helix.

Selective recognition of CG, could be achieved by utilising additional interactions across the CG base pair, *via* amino-modified nucleosides, to form more stable, selective triplets than those which can be formed by the natural base T. Four modified phosphoramidite monomers, *meta*-aminophenyl-modified analogues of the bicyclic nucleosides, (2,3*H*)-furano[2,3-*d*]pyrimidin-2(7*H*)-one and *N*-methyl-(2,3*H*)-pyrrolo-[2,3-*d*]pyrimidin-2(7*H*)-one, were synthesised to address this potential hydrogen-bonding motif.

Biophysical studies demonstrate selective recognition of the CG base pair. Results indicate selectivity for CG and binding affinity are much improved on previous modifications. Their fluorescence properties and general oligonucleotide deprotection conditions were also studied.

In addition, the synthesis of a *bis*-amine modified 6-oxocytidine phosphoramidite monomer for GC recognition was re-investigated.

This research shows significant advances in the field of triplexes for therapeutic use.

Contents

Declaration

I, Simon Richard Gerrard declare that this thesis entitled "Novel Nucleotide Analogues for Forming Stable DNA Triple Helices" and the work presented within are both my own and have been generated by myself as the result of my own original research.

I confirm that:

- this work was done wholly or mainly in candidature for a research degree at this University;

- where I have consulted/quoted the published work of others, this is always clearly attributed/referenced;

- I have acknowledged all main sources of help;

- I have made clear, where the thesis is based on work done in conjunction with another party, what my contribution to that work is;

- work described in Chapters 3 and 4 has been published in the paper below:

  – Gerrard, Simon R.; Srinivasan, Natarajan; Fox, Keith R.; Brown, Tom; "CG Base Pair Recognition Within DNA Triple Helices Using *N*-Methyl-3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one Nucleoside Analogues", *Nucleosides, Nucleotides and Nucleic Acids*, 26(10), 1363-1367.
  DOI: 10.1080/15257770701533958

Signed:……..

Date: 21st June 2009

Acknowledgements

I have many people to thank for making my life at Southampton University so exciting and rewarding. Firstly, thank you to my supervisor Professor Tom Brown for imparting so much knowledge, guidance and enthusiasm, and for his endless project ideas, patience and understanding - this project would have been impossible without him. Many thanks also to Dr. Dorcas Brown and all at ATDBio Ltd. (including Dr. Imenne Bouamaied), for their invaluable work in synthesising and purifying all my many oligonucleotides. I would also like to thank my co-supervisor Professor Keith Fox for his most useful feedback on my reports, posters and publication, and ideas throughout.

Working in the Brown Group, with such a large, welcoming, varied group of people from around the world was fantastic. My thanks go out to all of them for putting up with so many questions from me, and for the useful phrases in many different languages! In particular, I would like to say a big thank you to past members, Dr. Li Hong, Dr. Rohan Ranasinghe, Dr. Vicki Powers, Dr. Wang Yang, Prof. Dr. Xiao Qiang, Chaturong Suparpprom, Dr. Sunil Vadhia, Martina Banchelli, Jan Michels, Jenny Hale, Dr. Lavinia Brennan and Dr. Naomi Hammond for so much useful advice and support throughout, and being genuinely great fun to work with. I must thank both present and recently left members - especially Adeline Durand, Dr. Afaf El-Sagheer, Dr. Peng Guo-mei, Dr. John Zhao, Dr. Imenne Bouamaied, James Richardson, Dr. Nittaya Gale, Dr. Noha Ben-Gaïed and Dr. Petr Kočalka. The PhD wouldn't have been the same without their great sense of humour, and support, and putting up with inane chatter and a little insanity from yours truly from time to time. A massive thank you for invaluable support also goes to Dr. Mastoura Edrees-Abdou, with whom it was a pleasure to work whilst finishing my final melting studies. Lucy, Radha, Chenguang (Harvey) and Paul Douglas (Roach Group) – thanks also!

Many thanks to the Mass Spectrometry and NMR people: John Langley and Julie Herniman, and Neil Wells and Joan Street, for providing such a quality service and answering my many queries. I am also immensely grateful to Dr. Eugen Stulz for his advice and allowing me extra time to finish my thesis.

Finally, I can't thank my girlfriend Sarah enough, for supporting me throughout my PhD, and for just being there!

List of Abbreviations

$^{1\text{-deaza}}$C	1-deazacytosine
$^{2\text{-AE}}$S	2'-aminoethoxy-S
^{AE}T	2'-*O*-aminoethyl-T
$^{2\text{-A}}$P	2-aminopyridine
3-EA	3-ethynylaniline, 3-aminophenyl acetylene
3-HPN	3-hydroxypropionitrile, 2-cyanoethanol
^{4H}T	5-methylpyrimidin-2(1*H*)-one
5'-dTMP	2'-deoxythymidine-5'-monophosphate
$^{5\text{-Me}}$C	5-methylcytosine
$^{5\text{-Me}}$dC, M	5-methyl-2'-deoxycytidine
$^{6\text{-oxo}}$C	6-oxocytosine / 6-oxocytidine
$^{8\text{-oxo}}$A	8-oxoadenine
A	adenine / adenosine
^{Ac}PP	* 6-(3-acetamidophenyl)-*N*-methylpyrrolo-dC
*aeg*PNA	aminoethylglycyl-PNA
^{AE}P	6-(aminoethyl)-*N*-methylpyrrolo-dC
*aep*PNA	aminoethylprolyl-PNA
AIBN	azo-*bis*isobutyronitrile
AMA	conc. aq ammonia:conc. aq methylamine (1:1 v/v)
Anal.	elemental (CHN) thermal combustion analysis
^{AP}P	6-(aminopropyl)-*N*-methylpyrrolo-dC
app	apparent
^{A}PP	* 6-(3-aminophenyl)-*N*-methylpyrrolo-dC
^{apy}C	4*N*-6-aminopyrid-2-yl-dC
aq	aqueous
^{BAox}C	2'-*O*-(2-aminoethyl)-5-(3-aminoprop-1-ynyl)-6-oxocytidine
^{BA}U	2'-*O*-aminoethyl-5-(3-aminoprop-1-ynyl)-uridine (*bis*-amino-U)
BER	borohydride exchange resin (*resin*–NMe_3^+ BH_4^-)
BNA	bridged nucleic acid
BOC	*tert*-butyloxycarbonyl
bp	base pair
br	broad (IR/NMR)
BTEA	benzyltriethylammonium

BTMA	benzyltrimethylammonium
C	cytosine / cytidine
cAMP	adenosine 3',5'-cyclic phosphate
Cbz	benzyloxycarbonyl
CE	Capillary Electrophoresis
CEOC	2-cyanoethoxycarbonyl
COSY	correlation spectroscopy
CPG	controlled pore glass (solid support)
d	2'-deoxy, doublet (NMR)
δ	chemical shift
D_3	4-benzamido-1,3-imidazole
dA	2'-deoxyadenosine
dC	2'-deoxycytidine
DCA	dichloroacetic acid
DCE	1,2-dichloroethane
DCM	dichloromethane
DEA	diethylamine
DEAD	diethyl azodicarboxylate
DEPT	distortionless enhancement through polarization transfer
dF*	2'-deoxy-*β*-D-ribofuranosyl-(2,3*H*)-furano[2,3-*d*]pyrimidin-2(7*H*)-one, furano-dT
dG	2'-deoxyguanosine
DIHT	diisopropylamine hydrotetrazolide
DIPEA	diisopropylethylamine, Hünig's base
DMA	*N,N*-dimethylacetamide
DMAP	4-dimethylamino pyridine
DMF	dimethylformamide
dmf	dimethylformamidine
DMS	dimethyl sulfate
DMSO	dimethyl sulfoxide
DMT	4,4'-dimethoxytrityl
dN	2'-deoxynebularine
DNA	deoxyribonucleic acid
dP	2'-deoxy-*β*-D-ribofuranosyl- (2,3*H*)-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one, pyrrolo-dC

dR	2'-deoxyribofuranose
DSC	disuccinimidyl carbonate
dT	2'-deoxythymidine
dU	2'-deoxyuridine
EDC	ethyl (3-dimethylaminopropyl) carbodiimide
EDTA	ethylenediamine tetra acetate/acetic acid
eq	equivalents
ES	ethylene sulfide
FAM	carboxyfluorescein
FRET	Förster Resonance Energy Transfer
FT	Fourier Transform
G	guanine / guanosine
GC	gas chromatography
^{G}P, ^{G}pdU	5-(3-guanidinylprop-1-ynyl)-dU
^{G}PP	* 6-(3-guanidinylphenyl)-*N*-methylpyrrolo-dC
GR	guanidinylating reagent, *N*,*N'-bis*-[(2-cyanoethoxy)carbonyl]-*S*-methyl-isothiourea
H, HEG	hexaethylene glycol
HMBC	heteromultinuclear bond correlation, long-range ^{1}H–^{13}C COSY
HMDS	hexamethyldisilazane
HMPA	hexamethylphosphoramide
HMQC	heteromultinuclear quantum correlation, ^{1}H–^{13}C COSY
HPLC	high-performance liquid chromatography
HRMS	high resolution MS
IONEX	ion-exchange
IR	infra-red
J	coupling constant (Hz)
K_{assoc}	association constant
LC	LightCycler®
LRMS	low resolution MS
m	medium (IR), multiplet (NMR)
m/*z*	mass to charge ratio
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation Time-of-Flight
^{MA}P	3-methyl-2-aminopyridine
^{MF}P	(2,3*H*)-furano[2,3-*d*]pyrimidin-2(7*H*)-one

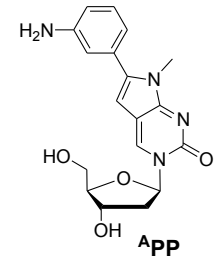
MMT	4-monomethoxytrityl
^{MNH}P	6-methyl-(2,3*H*)-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one, 6-methyl-pyrrolo-dC
^{MOE}pU	5-(3-aminopropynyl)-2'-*O*-methoxyethyl-U
^{mox}C	5-methyl-$^{6\text{-oxo}}$C
Mp	melting point
^{m}P^{B}	5-methyl-2-pyridone
^{M}P	* 6-methyl-*N*-methylpyrrolo-dC
mRNA	messenger RNA
MS	mass spectrometry
n.d.	not determined
N^7-G	N^7-guanine
N^7-I	N^7-inosine
NBS	*N*-bromosuccinimide
NCS	*N*-chlorosuccinimide
NHS	*N*-hydroxysuccinimide
NIS	*N*-iodosuccinimide
NMI	*N*-methylimidazole
NMR	nuclear magnetic resonance
NPhth	phthalimide
OD	optical density (units)
OL	oligonucleotide
P, pdU	5-(3-aminopropynyl)-2'-deoxyuridine
p-TsOH	*para*-toluenesulfonic acid, *p*-tosic acid
P_1	8-aza-9-deaza-9-methyl-G
P^{B}	2-pyridone
PNA	peptide nucleic acid
ppm	parts per million
Pu	purine base
Py	pyrimidine base
pyDDA	6-aminopyrazin-2(1*H*)-one
pyr	pyridine
q	quartet (NMR)
Q^{B}	1-isoquinolone
R_f	retention factor

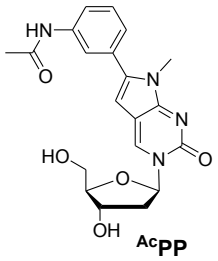
RNA	ribonucleic acid
rt	room temperature
S	*N*-[3-(4-acetamidothiazol-2-yl)phenyl]-acetamide, S base
s	strong (IR), singlet (NMR)
SNP	single nucleotide polymorphism
SPE	solid phase extraction
T	thymine / thymidine
t	triplet (NMR)
TAEA	*tris*(2-aminoethyl)amine
TBAF	tetrabutylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl
tC	1,3-diaza-2-oxophenothiazine
tC^{O}	1,3-diaza-2-oxophenoxazine
TCA	trichloroacetic acid
TFA	trifluoroacetic acid
TFO	triplex-forming oligonucleotide
THF	tetrahydrofuran
TIPDS	tetra-isopropylsiloxane-1,3-diyl
TLC	thin layer chromatography
T_m	DNA melting temperature
TMS	trimethylsilyl
^{TU}PP	* 6-(3-thioureidophenyl)-*N*-methylpyrrolo-dC
U	uracil
^{UN}I	butylureido-naphthimidazole
^{U}PP	* 6-(3-ureidophenyl)-*N*-methylpyrrolo-dC
UV	ultra-violet
vs	very strong (IR)
w	weak (IR)
WNA	W-shape bicyclic nucleic acid
$^{X}PP_f$	(2,3*H*)-furano[2,3-*d*]pyrimidin-2(7*H*)-one (furano-dT) phosphoramidite monomer
$^{X}PP_p$	(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one (*N*-methylpyrrolo-dC) phosphoramidite monomer
$\psi^{iso}C$	pseudoisocytosine

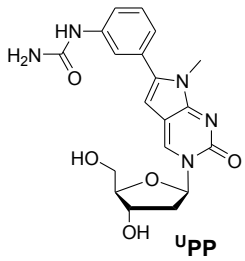
* Monomers for CG recognition.

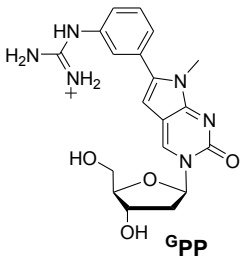
List of Monomers Used

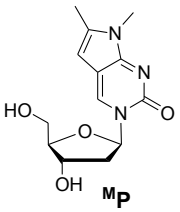
Modified/non-natural monomers, labels (FAM, DABCYL) and spacer (HEG):

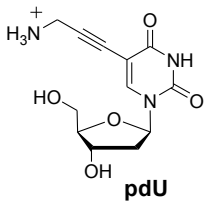


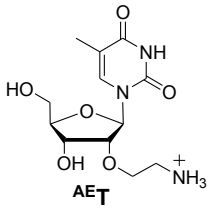


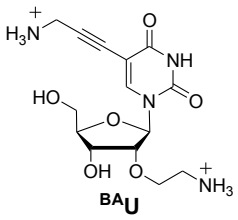


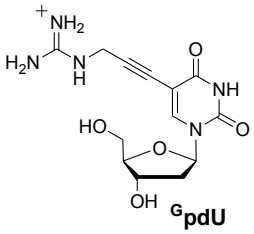


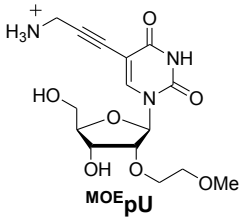


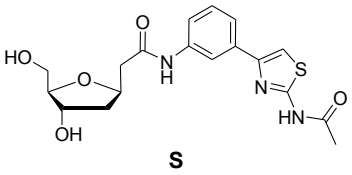


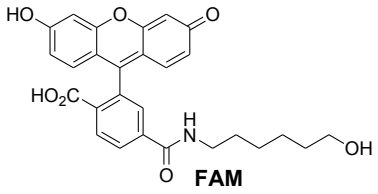


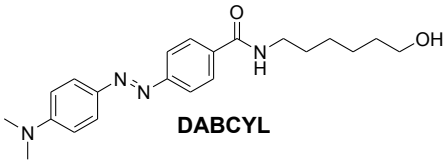


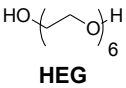












Chapter 1

Base Pair Recognition within DNA Triple Helices – Introduction

1. Base Pair Recognition within DNA Triple Helices: Introduction

1.1 Structure of DNA[1-3]

1.1.1 Primary Structure

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are macromolecular structures comprised of nucleotides (monomeric subunits). Each nucleotide is composed of a heterocyclic nucleobase, a furanose sugar and a phosphate moiety. Human genomic DNA is approximately 3.9×10^9 nucleotides, or 1.33 m in length, and each cell carries two copies.

In DNA, there are two purine bases, adenine **1** (A) and guanine **2** (G); and two pyrimidine bases, cytosine **3** (C) and thymine **4** (T). Nucleobases combine with a 2'-deoxy-ribofuranose sugar (dR) to give the nucleosides, 2'-deoxyadenosine **5** (dA), 2'-deoxyguanosine **6** (dG), 2'-deoxycytidine **7** (dC) and 2'-deoxythymidine **8** (dT) respectively. The bases are covalently bonded to $C^{1'}$ of the sugar moiety by a glycosidic bond; for pyrimidine bases *via* the N^1-atom, and for purine bases, via the N^9-atom (Figure 1.1).

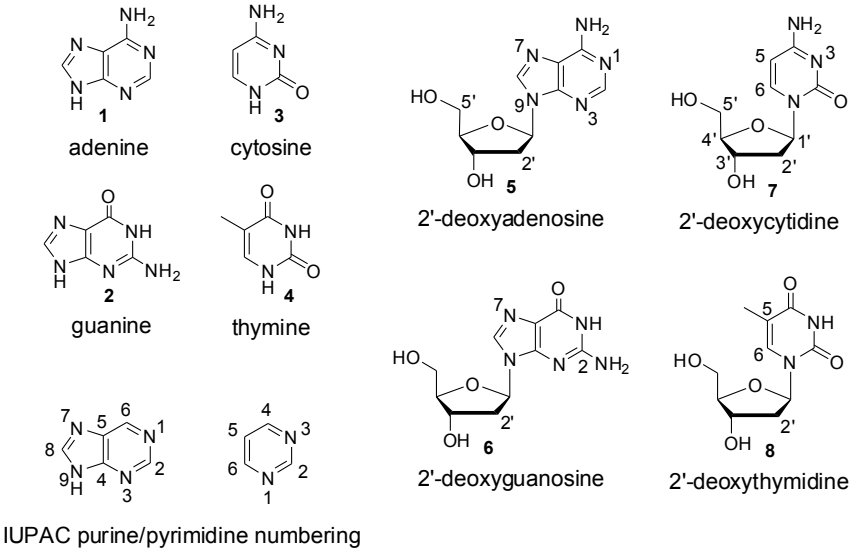


Figure 1.1. DNA nucleobases and corresponding 2'-deoxyribonucleosides.

RNA differs from DNA in two ways. The sugar moiety is ribofuranose, namely it has a hydroxyl group in the 2'-position, and thymine is replaced by uracil **9**, which lacks a methyl group at the 5-position (Figure 1.2).

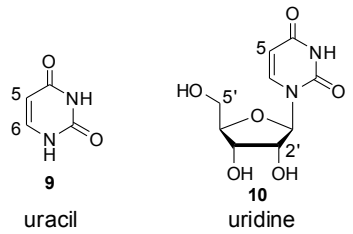


Figure 1.2. Uracil **9** and corresponding ribonucleoside, uridine **10** (U or rU).

Phosphodiesters link the nucleosides together to form the nucleic acid polymer. The sugar-phosphate backbone of this polymer is directional and chiral, by nature of the furanose ring structure. Nucleic acid sequences are therefore by definition, quoted as 5' to 3'; the tetramer ACGT therefore is not the same as TGCA (Figure 1.3).

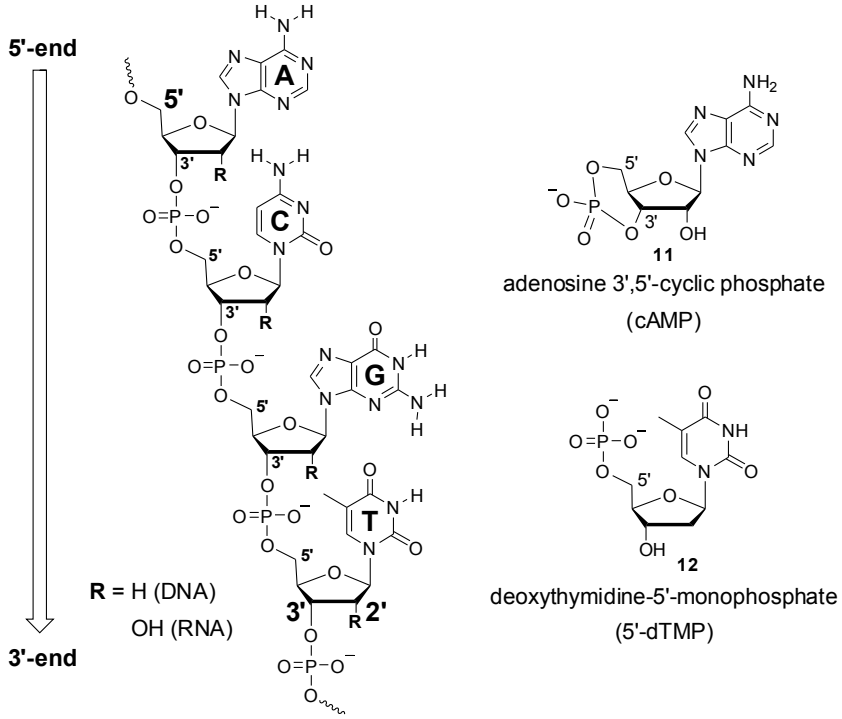


Figure 1.3. Nucleic acid primary structure and examples of simple nucleotides **11,12**.

Geometry of nucleotides within the macromolecule or polymer, and of the nucleosides, is governed by sugar and sugar-base conformations. The furanose sugar ring can adopt two energy-minimised conformations or puckers, the $C^{2'}$-*endo* (*S-type*) and $C^{3'}$-*endo* (*N-type*) (Figure 1.4).

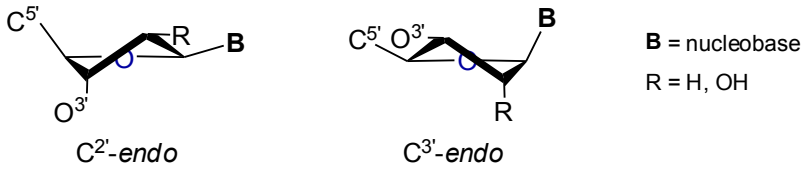


Figure 1.4. C$^{2'}$-*endo* and C$^{3'}$-*endo* sugar puckers.

The nucleobase lies almost perpendicular to the plane of the sugar allowing for different conformations. The nucleoside may adopt one of two principal orientations, either *syn*- or *anti*-conformations, by rotation around the C$^{1'}$-N glycosidic bond (Figure 1.5).

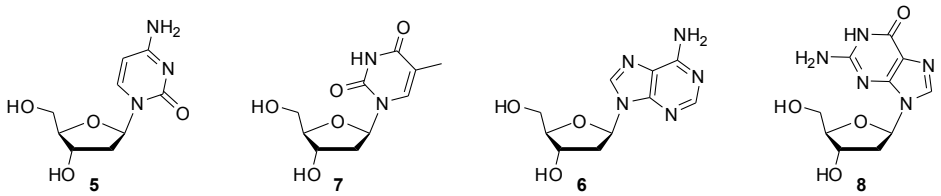


Figure 1.5. *Syn*- and *anti*- conformations of pyrimidine/purine nucleosides.

Pyrimidine nucleobases always occupy the *anti*-conformation due to repulsion between the base carbonyl oxygen and furanose ether oxygen. Purine nucleobases, however, are also able to occupy the *syn*-conformation. Guanine occupies this orientation preferentially in mono-nucleotides, Z-DNA and some other oligomers, where favourable interactions are possible between the NH_2 and 5'-phosphate.

1.1.2 Secondary Structure

The secondary structure of B-DNA was published in 1953 by Watson and Crick,[4] following extensive X-ray diffraction studies by their associates, Franklin and Gosling,[5,6] and Wilkins *et al*.[7]

Two DNA strands are bound in an anti-parallel, double-helical structure by intramolecular hydrogen-bonding between bases. The four bases form Watson-Crick base pairs;[8] G pairs with C *via* three intramolecular hydrogen bonds, and A pairs with T *via* two hydrogen-bonds.

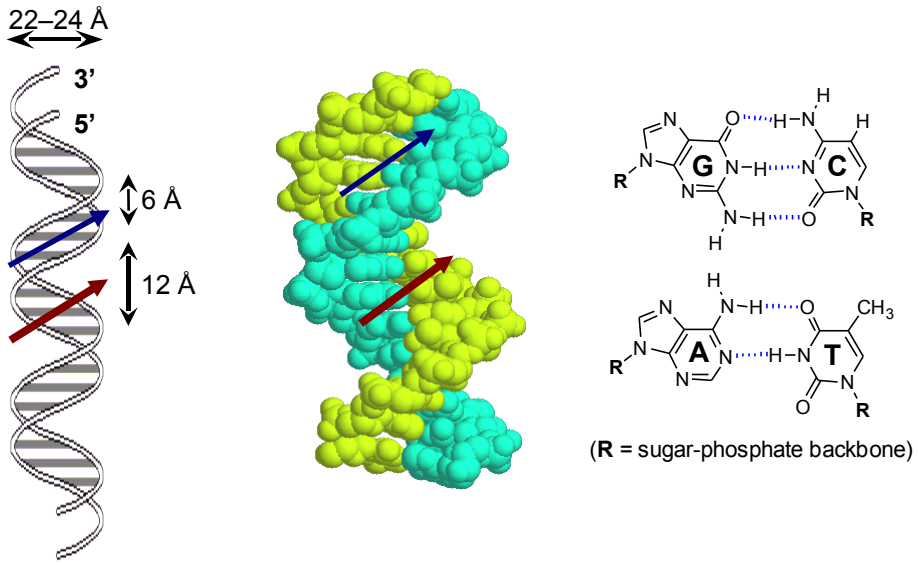


Figure 1.6. B-DNA double helix (schematic/model) and Watson-Crick base pairing (red arrow – major groove, blue arrow – minor groove).

Due to the antiparallel orientation of the two strands within the DNA duplex and non-symmetrical structure of deoxyribofuranose, GC is non-equivalent to CG, and AT is non-equivalent to TA. Therefore four orientation-specific Watson-Crick base pairs GC, CG, AT and TA can be described.

In the DNA double-helix or duplex, the sugar-phosphate backbones form the highly charged (poly-anionic), hydrophilic exterior of the structure, while the base pairs stack on top of each other within the hydrophobic core of the duplex, *via* π-π interactions. There are also two grooves running the entire length of the duplex, a major and a minor groove (Figure 1.6).

RNA exists in both single-stranded (within cell nucleii) and double-stranded forms (viral genome). It can form DNA.RNA hybrids, for example during transcription of genetic code from an unwound, single-strand of DNA to mRNA.

There are two principle conformations for the DNA double helix, A and B, both of which have been characterised by X-ray crystallography. B-DNA, which occurs in conditions of high humidity/low salt (e.g. in solution, *intracellular*), is a right-handed duplex. It has a periodicity of 10 base pairs; a pitch of 34 Å; a major groove 12 Å wide and 9 Å deep; and a minor groove 6 Å wide and 8 Å deep. The sugar pucker is $C^{2'}$-*endo* and nucleosides adopt the *anti*-conformation.

A-DNA occurs under conditions of low humidity/high salt. This form is also adopted by RNA duplexes and DNA.RNA hybrid duplexes. It is a right-handed duplex like B-DNA but is wider; it has a periodicity of 11 base pairs; and the base pairs are tilted 20° from the helical axis. The major groove is now much narrower (3 Å) but very deep (14 Å), and the reverse is true for the minor groove. The sugars adopt the $C^{3'}$-*endo* pucker and the glycosidic bond has the *anti* conformation.

There are also a number of variants on A- and B-forms, some with synthetically-modified bases (methylated, brominated); all are right-handed, with different periodicities, dimensions and sugar conformations.

Z-DNA, a left-handed duplex, was first discovered by Rich *et al*. in 1979 during studies on the DNA hexamer d(CGCGCG).[9] Other sequences of alternating pyrimidine-purine such as $d(CG)_n$ have also been found to adopt this form. Z-DNA has a periodicity of 12 base pairs, a very narrow, deep minor groove, lacks a formal major groove, and is stable in conditions of high salt. Interestingly, for the left-handed duplex $d(CG)_n$, although all nucleosides should now adopt the *syn*-conformation, the pyrimidine nucleoside dC is unable to, hence adopts the *anti*. The sugar pucker is $C^{2'}$-*endo* for dC and $C^{3'}$-*endo* for dG. Due to these conformations, the backbone zigzags around the helical axis.

1.2 DNA Triplexes[1,2,10-13]

The major and minor grooves of the DNA duplex are lined with potential hydrogen-bonding acceptors and donors. The major groove of B-DNA is accessible to DNA binding regions of various proteins and large drug molecules. The minor groove, although smaller, may also host a number of small molecules such as the pyrrole-amidine antibiotics distamycin and netropsin, which have a preference for AT-rich regions.[1,14] Importantly, it is possible for a third strand of DNA (or RNA) to bind inside the major groove to form a local triple-helix or triplex (Figure 1.7).

DNA triple helices were discovered by Felsenfeld and Rich in 1957.[15] In experiments using polyuridylic acid (polyU) and polyadenylic acid (polyA), they found a relatively weak complex (compared with the duplex $A_n.U_n$) was formed in a 2:1 ratio when these were mixed in aqueous solution. Stabilised by Na^+/Mg^{2+} cations, binding was specific

to polyU. No complexation was observed with polyC or other oligonucleotides. A decade later, the same observation was made for $d(CT)_n$ binding to $d(GA)_n.d(CT)_n$, and dG_n binding to $dG_n.dC_n$.[1]

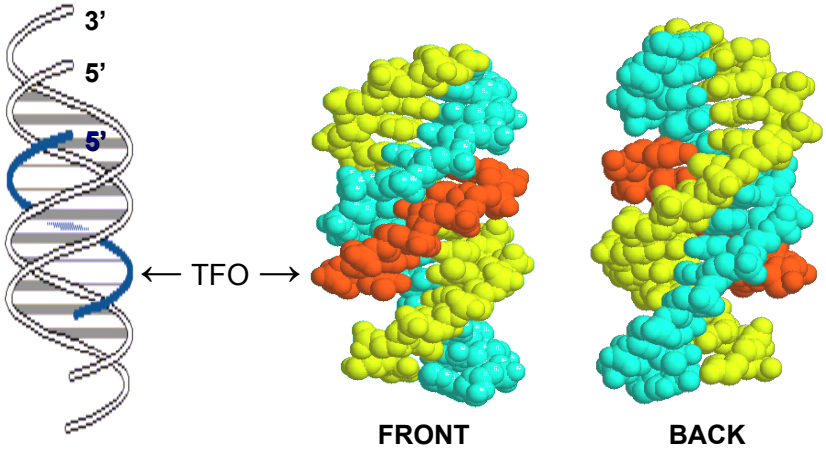


Figure 1.7. Schematic model (TFO in blue) and space-filling molecular model (TFO in red) of a parallel triple helix. Molecular model derived from NMR data on a triplex-duplex junction.[16]

It is only in recent years, however, that the potential applications have been recognised. Triplexes could have important uses in gene therapy (antigene strategy),[17-22] site-directed cleavage[17,23-25] and repair,[17,26-28] or as a tool in molecular biology and biotechnology applications.[17,29,30] Binding of the third strand, or triplex-forming oligonucleotide (TFO), to a specific region of the genome blocks unwinding of the duplex by helicases. Importantly, this prevents transcription thereby inhibiting gene expression (gene knockout), and prevents DNA replication hence interrupts cell division (chemotherapeutics). TFOs for human therapeutic use should *statistically* be *ca.* 16-17 nucleotides long for the recognised duplex sequence to be unique, thus preventing undesired inhibition. In genomic DNA however, the base pairs are not statistically-distributed, hence TFOs may vary in length from 18 to 30 nucleotides or more, to account for local variation and repeating motifs.[31]

In a DNA triplex, the TFO binds in a helical manor around the duplex *via* "Hoogsteen" hydrogen-bonds[32] to exposed hydrogen-bond donors and acceptors in the major groove (Figures 1.7, 1.8).

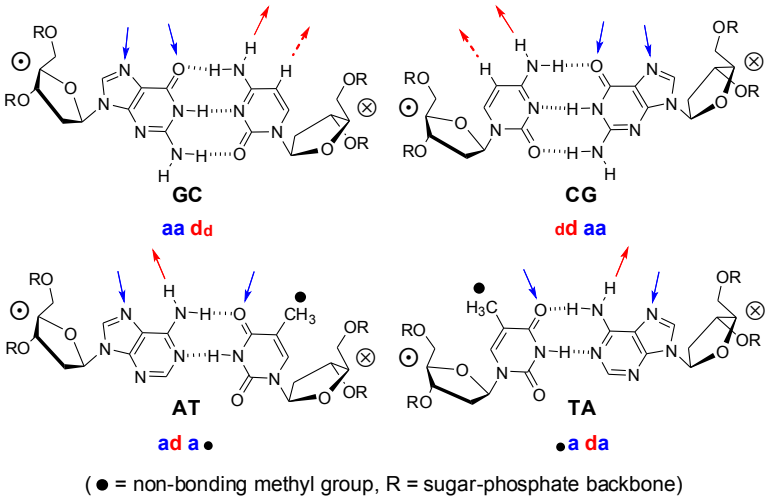


Figure 1.8. Four Watson-Crick base pairs and hydrogen-bond acceptors (a) and donors (d) exposed in major groove.

If the H-bond donors and acceptors on a nucleotide are correctly placed, it can bind to a base pair forming a triplet. Such triplets, written in the form X.YZ, can be formed from the natural bases.

There are two possible triplex configurations, which differ in the orientation of the third strand (Figure 1.9). In the parallel triplex, the third strand (TFO) is orientated in parallel with the strand to which it binds. In antiparallel triplexes, the reverse is true. Antiparallel triplexes are inherently less stable than parallel triplexes, hence most work in this area has involved parallel triple helices. DNA triplexes are structurally-characterised predominantly by NMR studies and circular dichroism (CD). Although the structure of DNA duplexes can be studied by X-ray crystallography (Section 1.1.2), DNA triplexes remain very difficult to crystallise.

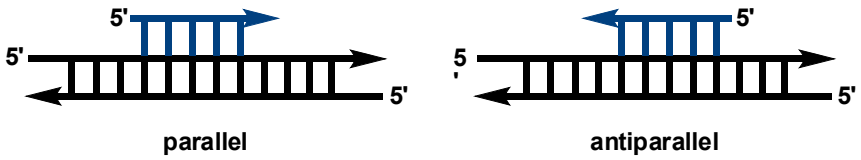


Figure 1.9. TFO orientation in a Parallel and Antiparallel Triplexes.

The most stable, natural parallel triplets are C^+.GC, T.AT, T.CG and G.TA.[33] With all natural triplets, binding can only occur to the nearest base in the pair, hence GC and AT, the purine-pyrimidine base pairs with two available purine H-bond acceptors/donors form the strongest triplets (Figure 1.10).

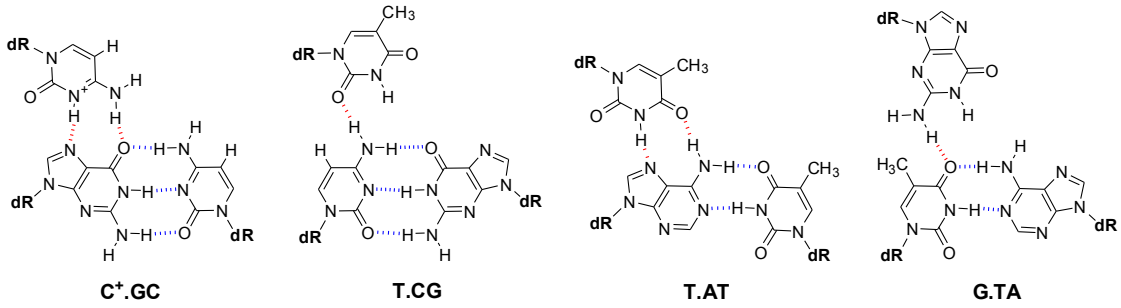


Figure 1.10. Structures of natural parallel triplets C^+.GC, T.CG, T.AT and G.TA (**dR** = deoxyribofuranose).

Antiparallel triplets include G.GC, T.CG, A.AT and T.AT (Figure 1.11). No natural base is capable of specifically recognising a TA base pair in an antiparallel triplex, although C.TA is probably the least unfavourable.[34]

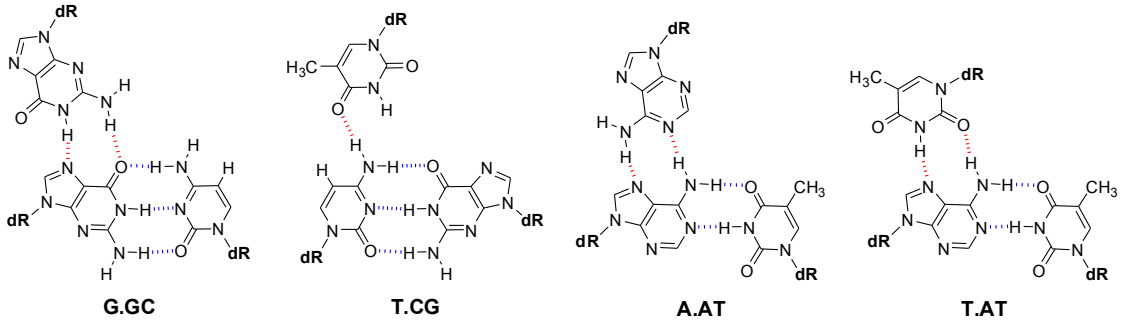


Figure 1.11. Structures of natural "reverse Hoogsteen" antiparallel triplets (**dR** = deoxyribofuranose).

The strongest and most selective natural base triplets, C^+.GC and T.AT (parallel), have been the subject of much research. However, there are intrinsic problems with both of these when used in triplex base pair recognition.

Triplexes are inherently less stable than DNA duplexes due to electrostatic repulsion associated with the close proximity of the three polyanionic strands; Hoogsteen hydrogen-bonds are therefore longer and weaker. Protonation of cytosine at the N^3-position is required for the C^+.GC triplet to form. However, due to the low pK_a of this protonated N^3 atom (pK_{aH} 4.5, as free nucleobase, cytosine), a low pH (<6.0) is required for partial protonation. At low pH, the triplet is very stable as the positive charge aids stabilisation of the negative charges on the phosphate groups, and there may be extra stabilisation from favourable π-stacking of the positively-charged cytosine ring in the complex. At physiological pH, however, few cytosines are protonated. Triplex stability

is then greatly reduced, such that TFOs containing many C bases or contiguous C bases will not form triplexes above pH 6.0. C.GC triplets are much less stable, lacking a second hydrogen-bond and favourable positive charge. The charge is the dominating factor governing stability. The triplet T.AT is less stable than C^+.GC as it lacks a stabilising positive charge, but the stability is not dependent on pH as protonation is not required for its formation.

1.2.1 Purine Recognition

Recognition of purine-pyrimidine base pairs, AT and GC, is potentially easiest because two hydrogen-bonds are possible between the third base and the purine, giving greater selectivity over other base pairs. Pyrimidine bases are more difficult to recognise selectively as only one formal H-bond is possible. Indeed, both purine and pyrimidine recognition have been subject to much research. DNA sequence recognition is still, however, mainly restricted to tracts of contiguous purine-pyrimidine (homopurine.homopyrimidine, Pu.Py) base pairs (AT, GC), as these form the most stable and selective triplexes. Achieving four-base mixed-sequence recognition, however, is essential to be able to target any gene or genetic code of choice. The number of target Pu.Py genomic sequences is limited, and many contain one or several pyrimidine interruptions, or Py.Pu inversions, which must be accommodated.

1.2.1.1 AT Base Pair Recognition

Recognition of the AT base pair was achieved by Sollogoub and Osborne *et al*.,[35-37] with a *bis*-amino-modified analogue of T, 2'-*O*-aminoethyl-5-(3-aminoprop-1-ynyl)-uridine or *bis*-amino-U (^{BA}U) **15** (Figure 1.12).

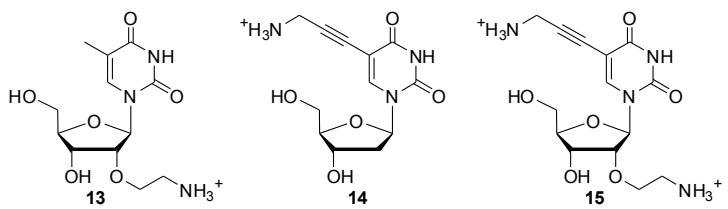


Figure 1.12. Analogues of T and nucleosides for AT recognition; 2'-*O*-aminoethyl-dT (^{AE}T) **13**,[38] 5-(3-aminoprop-1-ynyl)-dU (pdU) **14**[39] and *bis*-amino-U (^{BA}U) **15**.[37]

Cuenoud *et al*.[38] demonstrated that addition of a 2'-aminoethoxy group to dT enhances triplex stability. The amine group is protonated at physiological pH (pH 7.0) hence imparts partial charge-stabilisation of the anionic sugar-phosphate backbone. The replacement of the methyl group of dT, with an aminopropynyl group at the 5-position, significantly enhances stability of the triplex *via* partial charge-stabilisation[39] as described above, and through increased π-π stacking interactions.[40] Combining these modifications afforded ^{BA}U, which whilst retaining the hydrogen-bonding pattern of T, binds to AT more strongly and with much better selectivity.

Guanidinium-based groups have also been assessed as moieties with potential for charge-stabilisation. The guanidinium moiety remains protonated over a large pH range (pK_a ~12.5) and can form multiple hydrogen-bonds. Prakash *et al.*[41] reported that a 2'-*O*-guanidinylethyl group had significant benefits over a 2'-*O*-aminoethyl group. TFOs containing this modification showed high affinity for duplex DNA and RNA, and when in isolated positions, increased the triplex/third strand melting temperature (T_m) by ~3.2 °C per modification. A novel protecting group, *N*-(2-cyanoethoxycarbonyl)-(CEOC) was also used, which is compatible with oligonucleotide synthesis conditions.[42] Sensitivity to base, however, even aqueous sodium bicarbonate or triethylamine, is an issue during synthesis. The guanidinyl group is often added post-synthetically, for example by heating the oligonucleotide with *O*-methylisourea chloride in aqueous ammonia.[43] This is not selective, however, as all exposed amines will be guanidinylated.

Roig *et al.*[44] also reported triplex formation using 2'-deoxyuridine with modification at the 5-position by a range of alkynyl linkers terminating in one or two guanidinium groups. These linkers gave an increase in triplex stability through charge-stabilisation. Incorporation of two non-contiguous monomers gave a small increase in T_m over just one (1.5-2.0 °C). The *bis*-guanidinyl version of ^{BA}U (synthesised *in-house*) was discovered to be as effective as the *mono*-guanidinylated version. This was thought to be due to steric hindrance factors. Both guanidinyl monomers were an improvement on ^{BA}U and these studies are continuing.

1.2.1.2 GC Base Pair Recognition using Pyrimidine Bases

Many modifications could be made to cytosine to give potential nucleotides for GC recognition, and there have also been a variety of other strategies. The first logical analogue was 5-methylcytosine[45-47] ($^{5\text{-Me}}$C) (Figure 1.13). Combining properties of T and C, N^3 has a slightly increased pK_{aH} and triplexes containing a $^{5\text{-Me}}C^+$.GC triplet were stable at a higher pH than C, although still not at physiological pH. Thermodynamic studies indicated the enhanced stability was entropic in origin (methyl groups displacing water molecules from hydrophobic core), however, the methyl groups may also improve base-stacking.[46,48]

Stable triplex formation has also been achieved using $^{5\text{-Me}}$dC modified at the 4*N*-position with spermine[49-51] or tetraethyleneoxyamine.[51] 4-Guanidino analogues of C[52] were also synthesised and assessed against GC.[53,54] However, although the predicted alignment was good, triplexes could not be formed (Figure 1.13).

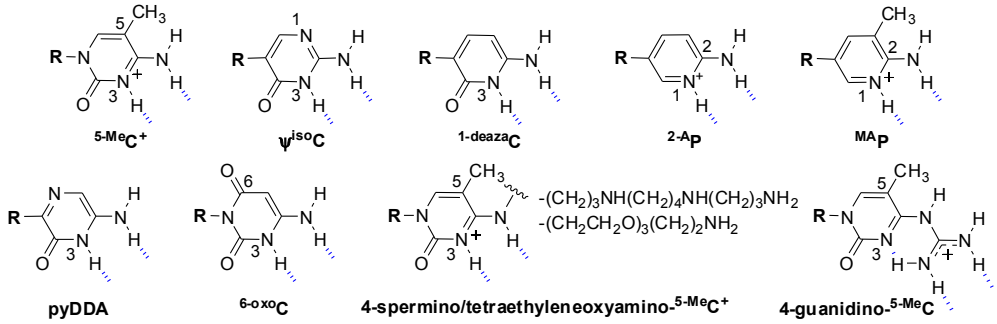


Figure 1.13. Cytosine analogues for GC recognition. **R** = furanose sugar

Pseudoisocytosine (ψ^{iso}C)[55,56] was successfully incorporated into TFOs for recognition of GC. The 2'-*O*-methyl-derivative of ψ^{iso}C formed pH-independent triplets, *via* two permanent N–H bonds. It could also be protonated at the N^1-position thereby increasing stability of ψ^{iso}C.GC triplets by charge-stabilisation. These triplets were also more stable than C^+.GC, at sites containing several contiguous GC pairs. This was presumably due to reduced charge repulsion, as is present between contiguous C^+.GC triplets.

Recently, ψ^{iso}C was incorporated into a peptidic analogue of DNA, aminoethylglycyl peptide nucleic acid (*aeg*PNA) for use in PNA/DNA triplex systems.[57] In PNA[58,59], the

DNA sugar-phosphate backbone is replaced by a flexible, *pseudo*-polypeptide, to which the nucleobases (**B**) are attached (Figure 1.14).

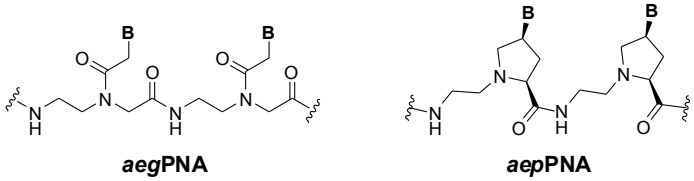


Figure 1.14. Example PNA structures, aegPNA and aepPNA.

1-Deazacytosine[60,61] ($^{1\text{-deaza}}$C), similar in structure to ψ^{iso}C, was investigated by Sollogoub, Powers *et al*. as a non-protonated, *C*-nucleoside analogue of C. Unfortunately, due to the challenging synthesis of the monomer although finally successful, and problems in deprotection of subsequent $^{1\text{-deaza}}$C containing oligonucleotides, biophysical studies were abandoned. There were also potential problems with GC selectivity due to tautomeric ambiguity of the endocyclic amide (Figure 1.15).[61]

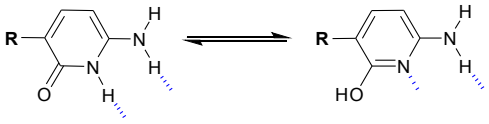


Figure 1.15. Tautomerism in 1-deazacytosine.[61]

2-Aminopyridine[62-66] ($^{2\text{-A}}$P) was successfully employed for recognition of GC. Triplexes containing this C analogue, as the 2'-deoxy-derivative, were found to be far more stable than C or $^{5\text{-Me}}$C at high pH. This is due to the much higher basicity of N^1 (pK_{aH} 5.93), although it was still pH-dependent. This stability was demonstrated when a triplex containing six contiguous $^{2\text{-A}}$P.GC triplets formed at pH 7.0. In addition, it had greater stability at lower pH as well. Surprisingly, Bates *et al.*[62] found the α-anomer more effective in stabilisation of the parallel triplex than the *β*-anomer, although Cassidy *et al.*[63] report otherwise. Cassidy does report, however, that the *α*-anomer can bind in an otherwise *β*-anomeric triplex with minimal structural perturbation.

3-Methyl-2-aminopyridine (^{MA}P)[61,65] also exhibits excellent triplex stability to higher pH than $^{5\text{-Me}}$C. The methyl group, however, does not appear to exact a noticeable difference in properties, hence the hydrophobic/entropic effects associated with $^{5\text{-Me}}$C cannot be a factor in this circumstance. ^{MA}P also demonstrated excellent

selectivity for GC in a mixed four-base sequence,[67] and forms very stable triplexes, together with ^{BA}U, for alternating GC/AT base pairs.[61]

A lesser known pyrazine-based C analogue (pyDDA)[68], demonstrated pH-independent binding between pH 6.3 and 8.0, to GC base pairs. PyDDA was also analysed recently in epimerisation studies as a new base for incorporation into duplex DNA[69].

1.2.1.3 6-Oxocytosine in GC recognition

6-Oxocytosine ($^{6\text{-oxo}}$C) and derivatives were used successfully as non-protonated C-analogues, for pH-independent GC recognition (Figure 1.16).

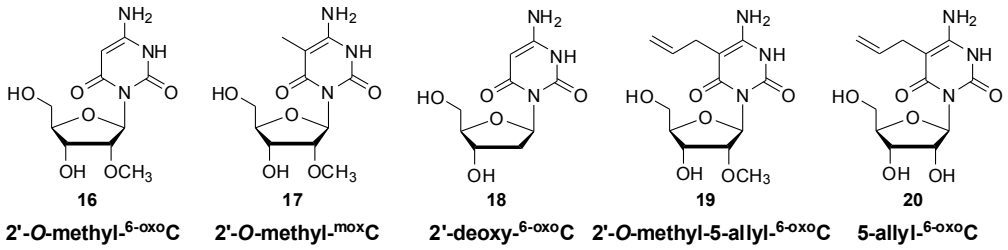


Figure 1.16. Example 6-oxocytidine ($^{6\text{-oxo}}$C) derivatives for GC recognition

Berressem and Engels[70] examined $^{6\text{-oxo}}$C and 5-methyl-6-oxo-C (^{mox}C), as the 2'-*O*-methyl derivatives (**16,17**), for efficacy in triplex formation. Both bases formed stable triplets with a GC inversion in a tract of contiguous AT base pairs. This binding was nearly pH independent with a small loss of stability in basic conditions around pH 8.0; the 5'-methyl group proving slightly destabilising. They also reported that at conditions of low pH, C^+.GC formed the more stable triplet. This was proposed to be due to enhanced base-stacking by the positively-charged cytosine ring, although $^{6\text{-oxo}}$C forms stronger hydrogen bonds to guanine, due to the electron-withdrawing C^6-carbonyl group. Cytosine therefore has a greater binding affinity at low pH due to the positive charge, but forms a much less stable triplet than $^{6\text{-oxo}}$C and ^{mox}C from pH 6.0-8.0. Parsch and Engels[71] later reported results of a number of $^{6\text{-oxo}}$C and 5-allyl-$^{6\text{-oxo}}$C derivatives designed for guanine recognition. UV melting studies showed that 5-allyl-$^{6\text{-oxo}}$C derivatives (inc. **19**) destabilised the triplex. They gave a lower triplex melting temperature (T_m), than $^{6\text{-oxo}}$C; hence, favourable hydrophobic/entropic properties that the 5-methyl group gave to C, also could not play a role in this case (*c.f.* ^{MA}P). An

RNA TFO containing RNA monomers of $^{6\text{-oxo}}$C (inc. **20**), however, did not form triplexes with duplex DNA.

Xiang *et al.*[72,73] reported that 2'-deoxy-^{mox}C (^{mox}dC) had the highest T_m of all analogues tried, and demonstrated pH-independent binding from pH 6.4 to 8.5. However, triplex-formation was weaker over five contiguous GC base pairs. If this base was used alternatively with $^{5\text{-Me}}$C or in presence of the DNA binding agent, spermine, several contiguous GC pairs could be recognised.

Following this work, Xiang and McLaughlin[74] reported use of a simple acyclic carbohydrate linker in place of 2'-deoxyribofuranose (dR) on ^{mox}C. The T_m was increased for this compound at an isolated GC base pair, the linker flexibility aiding orientation for hydrogen-bonding, which may be compromised due to steric repulsions. Over several contiguous GC pairs, however, the T_m was reduced. The triplex may be entropically destabilised by several contiguous flexible linkers.

6-Oxocytosine was even reported to be effective as a sequence-specific HIV-1 integrase competitive inhibitor *in vitro*, in duplex DNA.[75] Integrase is responsible for integrating viral DNA into DNA of the host, thus enabling virus replication.

Studies conducted by Wang, Powers and Osborne[11,37,76] on 2'-*O*-aminoethyl derivatives of $^{6\text{-oxo}}$C and ^{mox}C indicated this base proved ineffective at GC recognition, because triplexes were weak or did not form; the 2'-*O*-aminoethyl group proving non-effective in triplex stabilisation. It was rationalised that a steric clash between the furanose oxygen and C^6-carbonyl oxygen caused the base to twist around the glycosidic bond (Figure 1.17). This makes the triplet non-planar, thereby misaligning the Hoogsteen hydrogen-bonding motif. This repulsion may also have induced a change in sugar pucker.

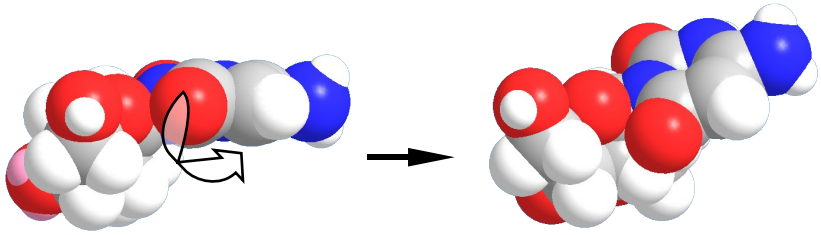


Figure 1.17. Space-filling representation of steric clash and repulsion in 2'-*O*-aminoethyl-6-oxocytidine.

1.2.1.4 GC Base Pair Recognition using Purine Bases

The natural purine base G forms triplets with GC, in both parallel (*anti*-conformer) and antiparallel (*syn*-conformer) triplexes.[33] Another logical approach therefore to analogue development, involved the use of purines as bases for guanine recognition (Figure 1.18).

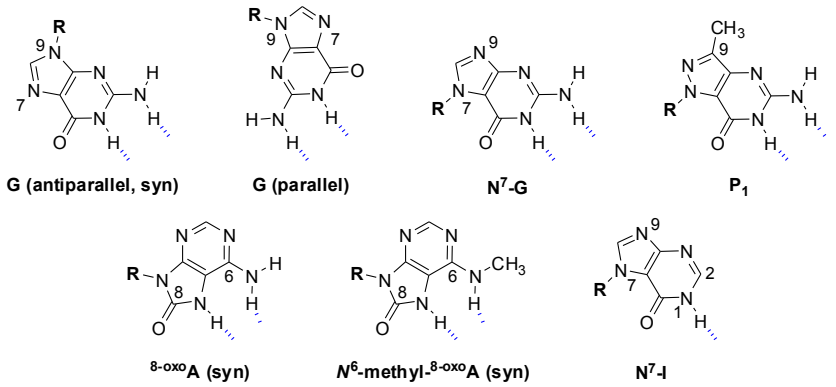


Figure 1.18. Purines as C analogues for GC recognition.

The G-related compounds, N^7-G[77-80] and P_1,[81-83] form triplets with GC, independent of pH. Their stability as isolated X.GC triplets is similar to $^{5\text{-Me}}C^+$.GC, but have much greater stability over $^{5\text{-Me}}C$, in triplexes with contiguous GC base pairs. They also demonstrate selectivity for GC over other base pairs, as was detected by DNase I footprinting studies. D'Costa *et al.*[84] have also incorporated N^7-G into two types of PNA (*aeg*PNA, *aep*PNA) as a C^+ analogue for PNA/DNA triplex formation.

Marfurt *et al.*[85] researched N^7-inosine (N^7-I) as a base for guanine recognition. This structure would form only one hydrogen bond, most likely between N^1H of N^7-inosine and N^7 of guanine, but obtain selectivity over other base pairs by disfavoured steric interactions. The α-anomer was shown to significantly reduce binding affinity relative to a triplex containing a $^{5\text{-Me}}C$ base at this position. The 2'-*O*-methyl- *β*-anomer slightly destabilised the triplex but the 2'-deoxy- *β*-anomer gave a ~10-fold increase in binding affinity on $^{5\text{-Me}}C$ as determined by quantitative DNAse I footprinting analysis. Very high selectivity was also observed over the other base pairs. Further molecular modelling studies also suggested that H-bonding between N^1H and C^2H, with O^6 of guanine cannot be ruled out, as these favourable electrostatic interactions may help stabilise the triplet.

8-Oxoadenine ($^{8\text{-oxo}}$A) and N^6-methyl-8-oxoadenine[86-90] were synthesised for GC recognition and again formed pH-independent triplets (*syn*-conformer). These triplets were almost as stable at a range of pHs, as C^+.GC is at low pH, and could form triplets in DNA tracts containing several contiguous GC pairs. Selectivity for GC was achieved by a combination of steric and electronic factors.

1.2.1.5 Other Novel Base designs

Other heterocyclic bases have been designed, which aim to mimic cytosine in binding to the GC base pair. Christensen *et al.*[57] used a novel 1,8-naphthyridine-2,7(1,8*H*)-dione moiety (as a replacement for C^+) in Hoogsteen-like triplex formation, in a PNA.DNA/PNA system for recognition of a single-strand of DNA (Figure 1.19).

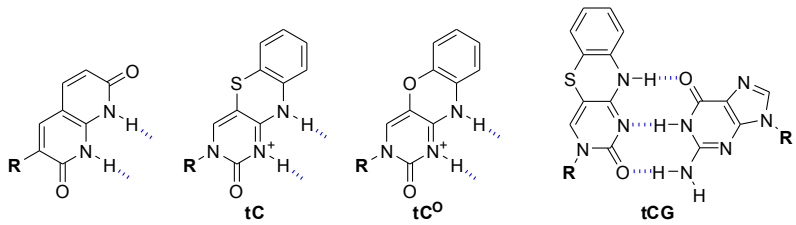


Figure 1.19. Bi- and tricyclic analogues of C; 1,8-naphthyridine-2,7(1,8*H*)-dione, 1,3-diaza-2-oxophenothiazine (tC) and 1,3-diaza-2-oxophenoxazine (tC^O), and the tCG base pair.

Capable of binding with two permanent NH bonds (major tautomer), the larger aromatic system was proposed to improve base-stacking in the triplex, hence impart extra stability. In isolated positions on the PNA it performed comparably with C^+ at pH 5.0, and better by 6.7 °C/modification with C^+ and by 1.5 °C with ψ^{iso}C, at pH 7.0. In adjacent positions, although both C and ψ^{iso}C displayed slightly reduced stability over isolated positions, the naphthyridindione performed significantly better than in isolated positions at pH 7 and 9. This was attributed to the increase in base-stacking over C and ψ^{iso}C. Adjacent C^+ and protonated ψ^{iso}C suffer some electrostatic repulsion due to positive charges. This modified base has been used in triplex recognition studies.[91,92]

Related structures, 1,3-diaza-2-oxo-phenothiazine (tC) and –phenoxazine (tC^O), have also been used in various DNA systems (DNA and DNA.PNA duplexes), as their non-protonated form, as C-substitutes in base pairing,[93-95] Tricyclic base tC has also been studied for its fluorescence properties.[96,97]

1.2.2 Pyrimidine Recognition

Pyrimidine.purine base pairs TA and CG are more difficult to target than purine.pyrimidine base pairs GC and AT, as only one formal hydrogen-bonding residue is offered for binding in the major groove. It may be possible to achieve either selectivity or strong binding affinity, but achieving both in combination is a challenge.

1.2.2.1 TA Base Pair Recognition

The natural base G is capable of recognising TA with only limited affinity and selectivity, hence alternative monomers have been designed to utilise the full hydrogen-bonding pattern of the TA base pair. The most selective nucleoside for recognition of TA base pairs is the S nucleoside **21** (Figure 1.20).[98,99] This non-natural, novel monomer was partly derived from the TA/CG recognition monomer D_3 **22**,[100] which was actually demonstrated to intercalate selectively between a TA-AT step.[101,102]

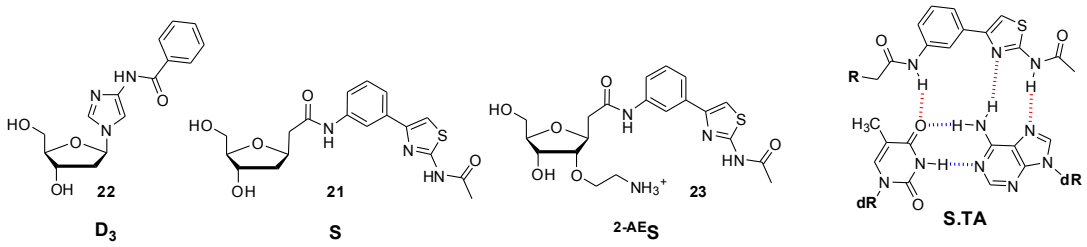


Figure 1.20. TA recognition monomers; D_3 **22**,[100] S **21**[98,99] and 2'-aminoethoxy-S **23** ($^{2\text{-AE}}$S),[76] and the S.TA triplet.

2'-*O*-Aminoethyl-S ($^{2\text{-AE}}$S) **23**, evaluated by Wang *et al.*,[76] showed increased binding affinity to a TA interruption over S. The increase in triplex stability by introduction of a 2'-*O*-aminoethyl group is well-studied,[37,38,103] and clustering of several 2'-*O*-aminoethyl-modified nucleosides affords significant triplex stabilisation and faster rates of triplex formation *in vivo*.[104] The drawback with D_3, S and $^{2\text{-AE}}$S is a significant affinity for CG, although selectivity for TA is increased for S and $^{2\text{-AE}}$S at higher pH.[76] Despite selectivity issues, a TFO containing the recognition monomers, ^{BA}U, ^{MA}P, ^{AP}P (Section 1.2.2.2) and S, was used successfully to target mixed-sequence duplex DNA at pH 7.[67]

TA recognition has also been achieved using 2'-*O*-aminoethyl-derivatives of G and 2-aminopurine,[105] although the 2'-*O*-aminoethyl group in this case proved detrimental to selectivity and binding affinity. This may be due to the $C^{3'}$-*endo* sugar pucker of these RNA-derived nucleosides.

1.2.2.2 CG Base Pair Recognition using modified Pyrimidine bases

The natural base T can recognise CG with a moderate affinity, but also forms triplets with AT. Thymine therefore cannot be used for CG recognition, but was a useful starting point for a series of modified pyrimidine-derived nucleosides, which aimed to utilise the full hydrogen-bonding pattern of the CG base pair.

The C^2-carbonyl group is important in recognition of CG within a parallel triplex. This is demonstrated by the moderate affinity of C, at lower pH, for CG.[106]

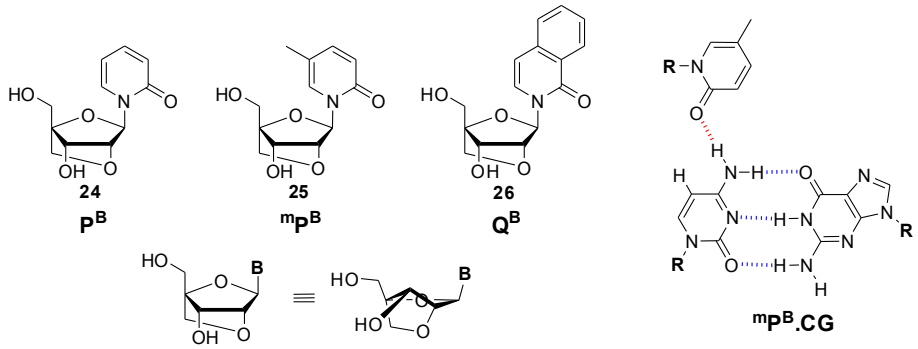


Figure 1.21. 2',4'-BNA-pyridinone nucleosides for CG recognition; 2-pyridone **24** (P^B),[107,108] 5-methyl-2-pyridone **25** ($^mP^B$)[108] and 1-isoquinolone **26** (Q^B).[109]

The simple, non-natural bases 2-pyridone **24** (P^B)[107,108] and 5-methyl-2-pyridone **25** ($^mP^B$)[108] efficiently recognise CG, as the conformationally-locked 2'-*O*,4'-*C*-methylene bridged nucleic acid (2',4'-BNA) nucleosides (Figure 1.21). The sugar is locked in the $C^{3'}$-*endo* conformation, which also affords a generalised increase in triplex stability, presumably through favourable alignment of the third strand base.[110]

Hari *et al.*[109] investigated a benzo-derivative of P^B, the 1-isoquinolone-2',4'-BNA nucleoside, Q^B **26**. P^B was shown to have a slight affinity for AT, thereby reducing selectivity for CG. The analogue Q^B demonstrated a marked reduction in affinity for AT at pH 7.0, appearing no more selective for AT as an abasic site (a sugar with no

base at the 1'-position). They proposed that a steric clash between the isoquinolone C^7H and 5-methyl group of T, when placed against an AT base pair, inhibited binding.

Prévot-Halter and Leumann[111] realised the importance of the N^3-atom in addition to the C^2-carbonyl, in recognition of cytosine. Comparing $^{5\text{-Me}}C$, $^{2\text{-A}}P$, and ^{4H}T **27** (4-oxo-deletion mutant of T), the 4-amino group was shown to play no critical role in CG recognition. In UV melting experiments, ^{4H}T in fact showed much greater selectivity for CG over GC and AT, although binding affinity was reduced.

Later studies, which tested recognition of a homopurine strand with a single C interruption, suggested a weak C–H···O interaction between H^5 of C and $O{=}C^2$ of ^{4H}T was key to the recognition properties.[112] This was partly based on observation that the ^{7}H.GC triplet is as stable as the C^+.GC triplet at pH 7.0, despite forming only one formal H-bond; the two weak C–H···O interactions providing additional stability (Figure 1.22).

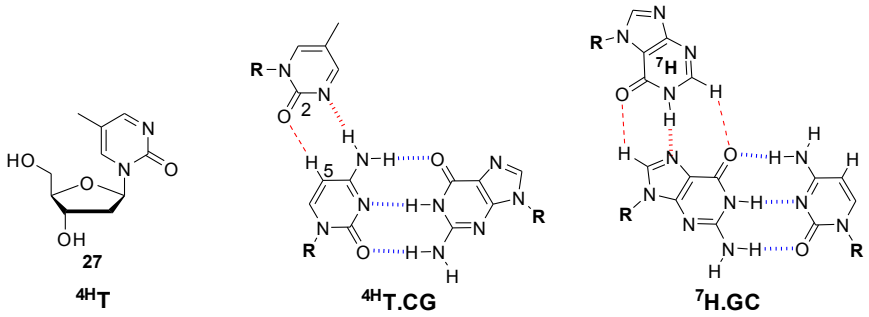


Figure 1.22. Pyrimidin-2-one nucleoside for CG recognition, 5-methylpyrimidin-2(1*H*)-one **27** (^{4H}T),[105,111-114] and putative ^{4H}T.CG and ^{7}H.GC triplets.[112]

Although the proposed interaction (pictured above) has not been confirmed by NMR studies, replacement of C (of the CG base pair) for 5-fluorocytosine destabilised the triplex by 8.3 °C, suggesting this interaction is highly significant.

Buchini and Leumann[105,113,114] subsequently attempted to improve the binding affinity of ^{4H}T for CG and triplex stability, by replacing the deoxyribofuranose (dR) sugar with that of 2'-*O*-aminoethyl-ribofuranose. Targeting a homopurine.homopyrimidine duplex containing a single CG inversion, the 2'-*O*-aminoethyl (2'-AE) group increased the triplex stability by 1.3 °C at pH 7.0, whilst retaining selectivity against the other base pairs. Triplexes were formed using a 15mer TFO, (using ^{AE}T **13** and 2'-*O*-aminoethyl-$^{5\text{-Me}}C$), with up to five CG interruptions in the duplex; selectivity dramatically increasing

with increasing number of CG interruptions, as the T_m of the other triplexes fell sharply. When using dT and $^{5\text{-Me}}$dC in the TFO under the same conditions, however, general triplex stability was poor, and with five interruptions, the T_m could not be determined.

Recognition of CG inversions using a fully 2'-AE-modified TFO produced interesting results. Three and five consecutive CG inversions were tolerated using 2'-AE-^{4H}T (**x**), and position did not affect triplex stability. Neighbouring nucleosides, however, played an important role; 2'-*O*-aminoethyl-$^{5\text{-Me}}$C (c) proving especially destabilising compared to ^{AE}T **13** (t). Changing part of the TFO, from 'tt**x**t**x**t**x**tc' to 'c**x**cc**x**cc**x**c', (and the corresponding duplex sequence accordingly), reduced the triplex T_m from 67 °C to 20 °C. This is likely to be due to electrostatic repulsion between neighbouring protonated cytosine nucleobases. Once the binding motif of ^{4H}T was established, alternative nucleosides were sought, which would increase triplex stability, and binding affinity and selectivity for CG.

The cyclisation reaction, (occurring *via* a 5-*endo*-*dig* mechanism),[115] which leads to the bicyclic nucleoside 3*H*-furano[2,3-*d*]pyrimidin-2(7*H*)-one (furano-dT) was first published by Robins,[116] then Cruickshank *et al*.,[117] as a minor fluorescent side-product of the Sonogashira reaction on 5-iodo-2'-deoxyuridine nucleosides. The 6-unsubstituted furano-dT nucleoside, dF* **28** found use by Woo *et al*.[118] as a fluorescent C analogue, when converted to the 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one (pyrrolo-dC) nucleoside dP **29**, during oligonucleotide deprotection/resin cleavage, using conc. aqueous ammonia. Although dP could not be directly incorporated into oligonucleotides due to instability during the oxidation step, the conversion of dF* to dP was quantitative. Nucleoside dP retains the hydrogen-bonding pattern of C and forms a stable base pair with G, also acting as a fluorescent marker (Figure 1.23).

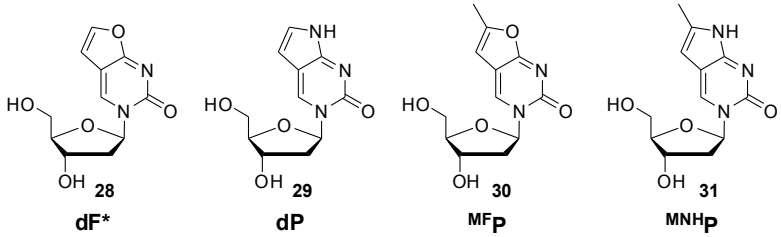


Figure 1.23. Bicyclic furano-dT/pyrrolo-dC nucleosides; dF* **28** and dP **29**,[118] 6-methyl-furano-dT **30** (^{MF}P) and 6-methyl-pyrrolo-dC **31** (^{MNH}P).[119]

The 6-methyl- (^{MNH}P, **31**) and 6-phenyl analogues also found use as fluorescent markers; the fluorescence being sensitive to base-pairing with G.[119,120] These could be directly incorporated into oligonucleotides as the pyrrolo-dC derivatives, or *via* post-synthetic conversion of the furano-dT nucleoside with conc. aqueous ammonia, as previously described. ^{MNH}P and other 6-alkyl derivatives were also tested by McGuigan *et al*. for anti-viral activity.[121]

This core bicyclic structure was modified by Ranasinghe *et al*.,[122] for use as a CG recognition monomer (Figure 1.24). Post-synthetic conversion of the furano-dT nucleotide to the *N*-methyl-pyrrolo-dC derivative using aqueous methylamine, instead of ammonia, alters the hydrogen-bonding pattern. The *N*-methyl group prevents the nucleoside acting as a C-analogue, blocking any recognition of the GC base pair whilst retaining the C^2 carbonyl and N^3, essential for selective binding to CG.[111]

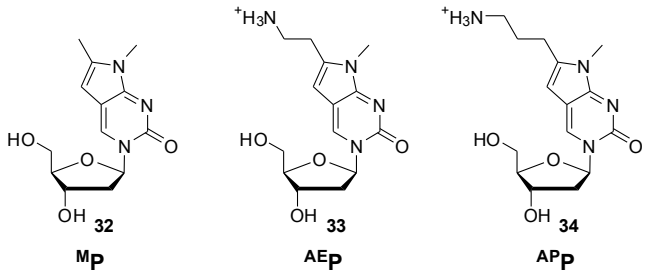


Figure 1.24. *N*-Methylpyrrolo-dC nucleosides for CG recognition; 6-methyl-*N*-methyl-pyrrolo-dC **32** (^{M}P), 6-(aminoethyl)-*N*-methylpyrrolo-dC **33** (^{AE}P) and 6-(aminopropyl)-*N*-methylpyrrolo-dC **34** (^{AP}P).[122]

Utilising the Sonogashira cross-coupling[123] conditions of Hobbs Jr.[124] and the documented cyclisation reaction under reaction conditions published by McGuigan *et al*.,[121] two 6-aminoalkyl-furano-dT nucleosides were synthesised. After incorporation into oligonucleotides, these, and the commercially-available 6-methyl-furano-dT **30** (^{MF}P), were post-synthetically converted to the corresponding *N*-methylpyrrolo-dC analogues **32-34** during deprotection with aque ous methylamine.

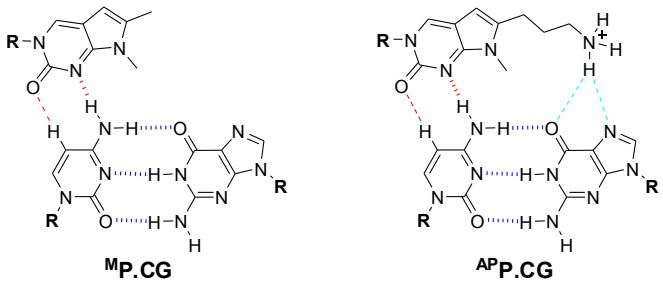


Figure 1.25. Proposed triplet motifs; ^{M}P.CG and ^{AP}P.CG.

The purpose of using this structure was three-fold; firstly, the bicyclic-nucleobase should enhance triplex stability by additional base-stacking; secondly, the amino group would be protonated at pH 7 thus providing stability through charge-stabilisation, and thirdly, it could form extra hydrogen-bonds to guanine. Using the extended hydrogen-bonding motif of the CG base pair could enhance selectivity and binding affinity (Figure 1.25).

Melting studies (fluorescence and UV) compared binding affinity and selectivity of the core structure ^{M}P **32** and nucleosides ^{AE}P **33** and ^{AP}P **34** with the natural base T, against a single CG inversion. At pH 5.5, an increase in T_m for ^{M}P, ^{AE}P and ^{AP}P, of 1.9 °C, 1.1 °C and 2.6 °C respectively, confirmed the aminopropyl linker performed best, although the results are close in value. Quantitative DNase I footprinting studies confirmed selectivity for the duplex target for all three monomers.

In order to compare with ^{4H}T, UV melting was conducted using the same sequence and conditions as used by Buchini and Leumann.[113] At pH 7.0, ^{M}P and ^{AP}P were shown to be comparable with ^{4H}T, and ^{AE}P comparable with T. Therefore, although an improvement on T, and whilst selective for CG, the aminoalkyl moieties did not significantly increase triplex stability. The pendant amino groups most likely did not form hydrogen-bonds with G as originally hoped but contributed some charge-stabilisation. However, this did set precedent for future analogues, which will be discussed in Section 3, and ^{AP}P was used successfully as a CG recognition monomer in a modified TFO against a mixed-sequence DNA target.[67]

1.2.2.3 Novel Pyrimidine-derived Monomers for CG Base Pair Recognition

Modification of cytosine and use of non-natural heterocyclic compounds have been assessed in CG recognition, in parallel and antiparallel triplexes, and in NMR studies in organic solution. This significantly widens the scope for achieving CG recognition, despite very variable binding affinities. Reduction in predicted stability was observed when moving from NMR studies of isolated triplets to melting studies on triplex systems.

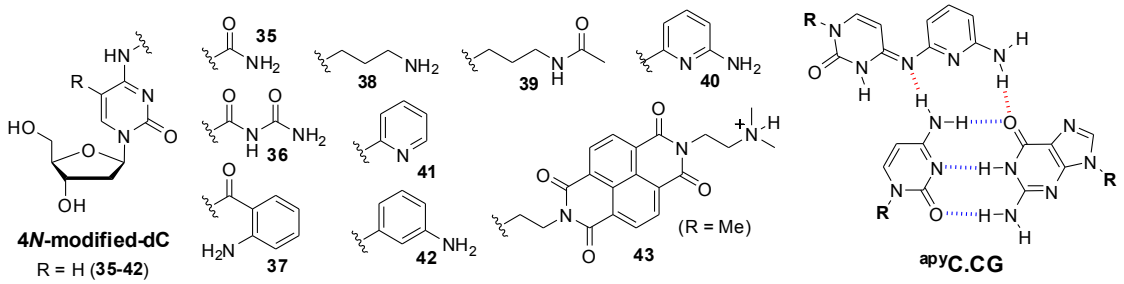


Figure 1.26. Modified C analogues **35-43** for CG recognition in parallel triplex,[125-131] and proposed 4*N*-6-aminopyrid-2-yl-dC.CG (^{apy}C.CG) triplet.[131]

Modification with pendant groups at the 4*N*-position of dC (Figure 1.26) was used to take advantage of the extended hydrogen-bonding motif of CG, and to afford extra stability through increased base-stacking. The 4*N*-carbamoyl- **35** and 4*N*-ureidocarbamoyl-dC **36** nucleosides of Guzzo-Pernell *et al.*[128] were designed to form extra hydrogen-bonds with CG, and could be fitted into reasonable triplets with CG. However, no triplex could be observed, perhaps due to hydrophilicity of the pendant groups preventing binding within the hydrophobic groove. An anthraniloyl-derivative **37** was also evaluated, introducing a hydrophobic benzene ring, but with no triplex formation. A 4*N*-phenyl-carbamoyl derivative did form triplexes at pH 6.0, albeit very weakly, with CG and GC base pairs, but the triplex was too unstable to be of practical use.[132]

4*N*-Aminopropyl- **38** and 4*N*-acetamidopropyl-dC **39** nucleosides of Huang *et al.*,[130] demonstrated selectivity for CG, but relatively weak binding affinity in UV melting studies against a single CG inversion. It was a surprise to the author that the flexible chain could support triplex formation, and more rigid modifications were considered.

Huang *et al*. later investigated a 4*N*-(6-aminopyrid-2-yl)-dC analogue **40** (^{apy}C),[129,131] which forms a triplex with a target containing a single CG inversion. NMR and DMS alkylation studies indicated this base binds as the imino-tautomer, which is also capable of recognising AT. UV melting experiments indicated ^{apy}C had a much stronger binding affinity to CG than their previous modifications. Selectivity against GC and TA was maintained, although it also recognised AT to a lesser extent. A biphasic melting profile against CG was also observed, which indicated other triplex structures, such as intercalation rather than Hoogsteen hydrogen-bonding. Notably, removing the 6-amino group (**41**) essentially eliminated triplex formation, and replacing the pyridyl for a phenyl ring (**42**) had only a minor reduction on triplex stability.

Chin *et al*.[125,126] also evaluated this monomer's CG recognition properties, in an intermolecular ('linear') triplex and an intramolecular ('paper-clip') triplex, against a single CG inversion. The ^{apy}C.CG triplet was shown to be as stable as a C^+.GC triplet, at pH 6.0, and gave pH-independent triplex binding when located internally. The more flexible 'linear' triplex could also tolerate the alternating sequence ^{apy}**C**T^{apy}**C**T^{apy}**C**.

A novel $^{5\text{-Me}}$dC derivative, tethered at the 4*N*-position with an *N*,*N*-dimethylaminoethyl-modified naphthalenediimide **43**, achieves CG recognition through one hydrogen-bond (from N^3 to N^4H of C), and intercalation.[127] UV melting studies demonstrated good selectivity from pH 6.4-7.5, predominantly due to hydrogen-bonding; intercalation was responsible for a general increase in stability compared with C. Melting studies, monitoring at 383 nm (λ_{max} – napthalenediimide), suggested a 'cooperative' un-stacking of the intercalator on third strand dissociation.

1.2.2.4 Novel Non-natural Heterocyclic Monomers for CG Base Pair Recognition

Several non-natural heterocyclic bases for CG recognition have also been studied in organic solution, by NMR of the isolated triplet. Although only an approximation of their behaviour in the aqueous environment, these studies presented some interesting potential candidates for future triplex studies. Some have also been progressed to triplex melting studies.

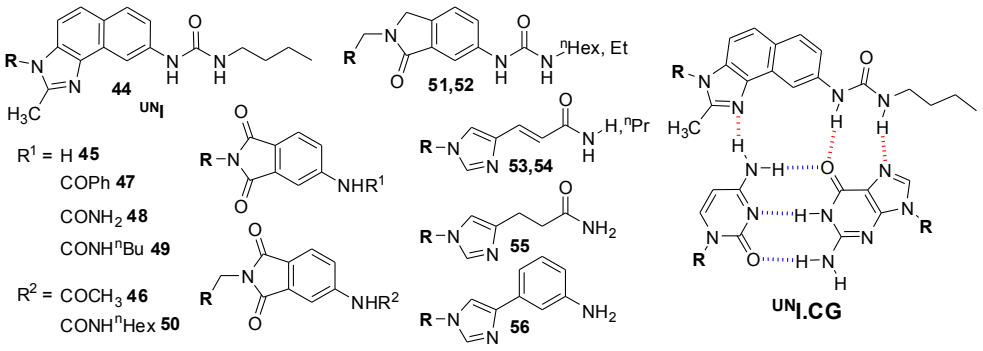


Figure 1.27. Novel heterocyclic CG recognition monomers; ureido-naphthimidazole **44** (^{UN}I),[133] amino- **45**, amido- **46,47**, and ureido-phthalimides **49-50**,[134-136] ureido-isoindolin-1-one **51,52**,[135,137] urocanamide **53,54** and imidazole derivatives **55,56**.[138-140]

A butylureido-naphthimidazole **44** (^{UN}I)[133] was assessed in NMR binding studies in forming a triplet with CG, with hydrogen bonds to 4-NH of C and C=O^6 and N^7 of G. Although incorporated into DNA, biophysical studies were not published.

Several modified phthalimide structures were also evaluated in NMR binding studies with CG.[134-136] Whereas amino- **45**, acetamido- **46** and benzamido- **47** modifications gave relatively weak binding, ureido- **48** and *N*-alkylureido- groups **49,50** demonstrated strong binding *via* three hydrogen-bonds from the phthalimide C=O to 4-NH of C, and from the urea to C=O^6 and N^7 of G.

A hexylureido-isoindolin-1-one structure **51**,[135] analogous to the above monomer, also effectively bound to CG with an association constant (K_{assoc}) ten-times that of Zimmerman's naphthimidazole **44**.[133] This was also progressed into triplex melting studies as the ethylureido-isoindolin-1-one-methylene monomer **52**,[137] however, it proved no more effective at CG recognition than an abasic site, with a stability of approximately 1.8-5.0 kcal/mol lower than the natural T.AT and C$^+$.GC triplets.

Derivatives of urocanamide **53,54** and related imidazole structures **55,56** were evaluated by Purwanto *et al.*[138-140] in NMR studies and UV melting experiments. NOE experiments indicated binding to CG occurred through two hydrogen-bonds, one to C^7 and one to C=O^6 of G. NMR binding studies of the primary amide **53**, *N*-propylamide **54**, alkyl- **55** and phenyl- **56** derivatives indicated strongest binding from the primary amide (urocanamide). Reduced affinity for the other three were attributed to rotation around the amide bond hindering binding for the *N*-ethylamide, too much flexibility for alkyl derivatives, and too little flexibility for phenyl derivatives.

The *N*-propyl urocanamide **54** was progressed to triplex melting studies, targeting a single CG inversion in an AT rich duplex. Despite showing a sharp third strand melting transition, a broad melt transition against TA of a similar T_m was also detected, possibly due to binding to A through protonation of the imidazole ring. An increase in selectivity for CG over TA was observed on increasing pH to 6.5, confirming this pH dependence of TA binding. Triplex stability, however, was approximately 10 °C lower than for the C^+.GC triplet, at pH 6.0, due in part to lack of π-π stacking area and hydrophilicity of the amide moiety. These observations once again highlighted the importance of base-stacking, hydrophobic interactions and conformation within the triplex, which cannot be well approximated in NMR studies of isolated triplets.

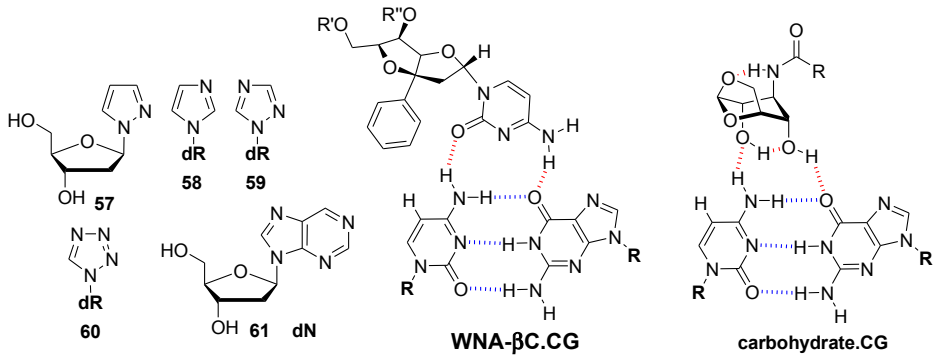


Figure 1.28. Novel carbohydrate and antiparallel CG recognition monomers/triplets.

Recognition of the CG base pair within antiparallel triplexes has also been employed despite issues of stability. The azole monomers **57-60** of Durland *et al.*[141] generally bound to two or more base pairs with moderate affinity, and the purine nucleoside 2'-deoxynebularine **61** (dN) demonstrated equal binding affinity to CG and AT at pH 7.4 *via* one hydrogen bond. The conformationally-locked WNA-βC nucleoside of Sasaki *et al.*[142] effectively recognised CG, whilst the G and T β-anomers, WNA-βG and WNA-βT, bound selectively to TA within an antiparallel triplex, as determined by gel-shift assays. A novel abasic carbohydrate was also assessed in NMR studies as a CG binding agent.[143] Recognition occurs *via* two hydrogen-bonds, although it is doubtful this could be in useful in a triplex context.

Although some success has been afforded within the antiparallel triplex, CG recognition is still most likely to be achieved through the parallel motif. Despite some degree of pH-dependence parallel triplexes are more stable and easier to work with, but antiparallel recognition is still of interest.

1.3 Conclusion

Genomic DNA is a promising target for many small-molecule and oligonucleotide-based therapeutics, aimed at treating a wide variety of genetic diseases and cancers.

The sequence-specificity of DNA allows for oligotherapeutics to be synthesised that target only specific sequences, to block a particular gene function, target cancer cells, and for a range of site-specific DNA strand modifications. It is for these reasons that extensive research has been carried out over a number of years, to achieve selective recognition of each of the four base pairs, whilst understanding the effect of interaction of neighbouring modified nucleosides on the efficacy of triplex formation, and enhancing resistance to *in vivo* degradation.

The therapeutic use of DNA triplexes is an expanding, exciting field, which continues to produce promising results in biophysical studies and against real targets in some biological systems.

Chapter 2

Modified 6-Oxocytidine Nucleosides for GC Recognition

2. Modified 6-Oxocytidine Nucleosides for GC Recognition

2.1. Introduction

Recognition of GC base pairs using the natural base C is both selective and strong below pH 6.0. Due to the low pK_{aH} of N^3 of cytosine, protonation of which is required for Hoogsteen hydrogen-bonding, the C^+.GC or C.GC triplet is not stable, however, at physiological pH (7.0). This severely limits the number of biological applications where this can be used, especially against targets containing contiguous GC base pairs, where repulsion between positive charges destabilises the triplex further.

A variety of pyrimidine- and purine-based monomers for GC recognition have been evaluated as described in Section 1.2.1.2. These include the pyrimidine nucleosides, 5-methylcytosine ($^{5\text{-Me}}$C)[45-47] and the *C*-nucleoside 3-methyl-2-aminopyridine (^{M}AP),[61,65] providing stability through a higher pK_{aH} and hydrophobic methyl group; and N^6-methyl-$^{8\text{-oxo}}$A[87] and pyDDA,[68] which form pH-independent triplets with permanent hydrogen-bonds.

The non-natural base 6-oxocytosine ($^{6\text{-oxo}}$C, **62**)[70-74] also found use as a GC recognition monomer, which forms triplets independent of pH. Unlike some other GC recognition monomers, $^{6\text{-oxo}}$C could potentially be modified at the 5-position on the nucleobase (and 2'-position of the sugar), with pendant amino groups, such as for *bis*-amino-U (^{BA}U).

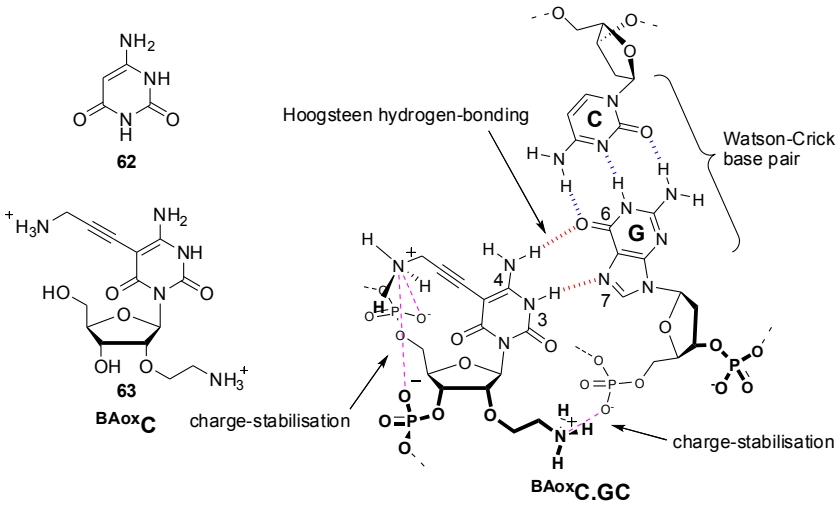


Figure 2.1. Structures of 6-oxocytosine **62**, 2'-*O*-(2-aminoethyl)-5-(3-aminoprop-1-ynyl)-6-oxocytidine **63** (^{BAox}C), and proposed hydrogen-bonding and electrostatic interactions within the putative ^{BAox}C.GC triplet.

In the above diagram (Figure 2.1), charge-stabilising interactions are shown between the protonated 2'-amine moiety and guanidine 5'-phosphate, and also between the protonated 5'-amine moiety and ^{BAox}C phosphates. Hoogsteen hydrogen-bonds are also shown between ^{BAox}C and G.

The initial project aim therefore, was synthesis of a direct analogue of ^{BA}U, a *bis*-amino-modified derivative of 6-oxocytidine, 2'-*O*-(2-aminoethyl)-5-(3-aminoprop-1-ynyl)-6-oxocytidine **63** (^{BAox}C) (Figure 2.1). This would combine the triplex-stabilising effect of the amine groups as exemplified by ^{BA}U and potentially enhanced π-stacking interactions from the alkynyl moiety. GC recognition would be achieved, as previously described, *via* two permanent hydrogen bonds between N^3H and N^4H of the base and N^7 and C^6=O of guanine, giving pH-independent triplex binding. Although poor binding affinity was observed when using 2'-*O*-aminoethyl-modified $^{6\text{-oxo}}$C and ^{mox}C compared to 2'-*O*-methyl- or 2'-deoxy- derivatives (Section 1.2.1.3), combining both amine modifications was proposed to enhance alignment with the GC base pair in addition to the properties described above.

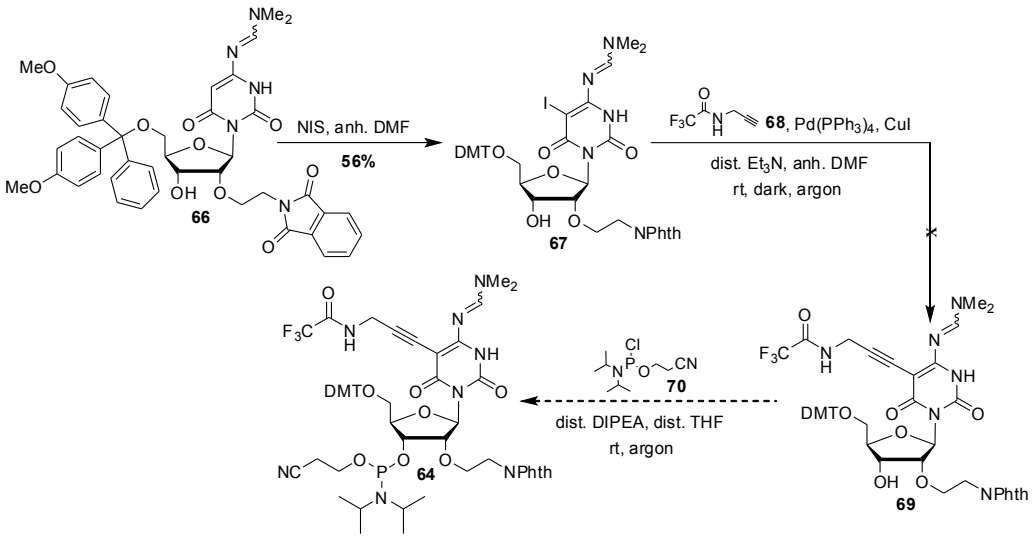
2.2. *Bis*-amino-6-oxocytidine: Synthesis

The key steps in synthesis of a functionalised 6-oxocytidine monomer are coupling of the persilylated 6-oxocytosine (6-aminouracil) base to the modified sugar; sugar deprotection and reprotection of the 5'-position; halogenation at the 5-position; protection of the 4-NH_2 group; and subsequent base modification. The order in which these are to be carried is open to variation. The final compound is then converted into a 3'-*O*-phosphoramidite[144] for incorporation into triplex-forming oligonucleotides (TFOs).

Dr. Osborne conducted much work in house, into synthesis of a ^{BAox}C monomer **64**, for incorporation into TFOs. After the sugar-base coupling using the 2'-*O*-phthalimidoethyl-modified sugar **65**, and three protection/deprotection steps, problems were encountered in two key steps in the synthesis (Scheme 2.1).

Iodination under the above conditions required excess NIS and several days to effect the transformation. The subsequent Sonogashira reaction under Hobbs' conditions[124] was

unsuccessful, thought at the time to be affected by neighbouring functionality rather than electronic factors, and the route was abandoned for re-evaluation.



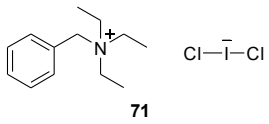
Scheme 2.1: Final three steps in original synthetic route to ^{BAox}C monomer **64**.

2.2.1 Synthesis of 5-Halo-6-oxocytidine nucleosides

It had been observed by Prof. T. Brown that during automated oligonucleotide synthesis, approximately 60% of unmodified $^{6\text{-oxo}}$C nucleosides were iodinated at the 5-position in the emerging oligonucleotide, during the oxidation of P^{III} to P^{V}. Following this observation, Osborne successfully used the iodine/THF/pyridine solution, employed in this oxidation step, to iodinate an acetylated 2'-*O*-phthalimido-$^{6\text{-oxo}}$C derivative. Once reaction feasibility was established, Osbourne reacted a 5'-*O*-(4,4'-dimethoxytrityl)-protected $^{6\text{-oxo}}$C derivative **66** with *N*-iodosuccinimide (NIS) (Scheme 2.1). It was later established that the aqueous workup of 5-iodo-$^{6\text{-oxo}}$C derivatives with sat. aq sodium thiosulfate to remove excess iodine, as used by Dr. Osborne, caused significant deiodination of the product, perhaps a reason for the moderate 56% yield.

Iodine, NIS[145-147] and iodine monochloride,[148,149] are well-known as iodinating agents for nucleosides, although ICl has potential for chlorination[150] and is more difficult to handle. For optimisation of yield/reaction time, an alternative iodinating reagent was sought. Bromination at this point was considered unnecessary as iodinated substrates are more reactive to Pd-catalysed cross-coupling conditions.

Benzyltriethylammonium dichloroiodate (BTEA-ICl_2) **71** (Figure 2.2), was used by Harrowven *et al.*[151] to iodinate a substituted anisole intermediate in synthesis of a target macrocycle, in 99% yield over 5 hours. A co-catalyst $ZnCl_2$ was used, which also helped to solubilise BTEA-ICl_2 in the solvent acetic acid.



71

Figure 2.2. Benzyltriethylammonium dichloroiodate (BTEA-ICl_2) **71**.

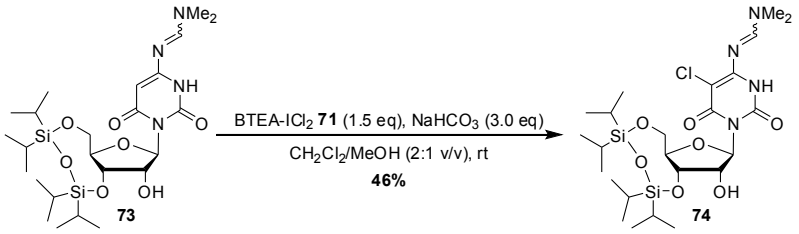
Most ammonium polyhalide-derived reactions such as this, in the literature, have utilised the benzyltrimethylammonium (BTMA) salt, which is commercially available, for halogenation reactions; for example, the dichloroiodate, tetrachloroiodate and tribromide or chlorobromate for iodination, chlorination and bromination respectively. The preferred solvents for halogenation, using BTMA dichloroiodate (BTMA-ICl_2), were methanol or CH_2Cl_2/methanol and several inorganic co-catalysts, $CaCO_3$,[152,153] $NaHCO_3$,[152,154] and $ZnCl_2$[151,155] have been used. Although the BTEA ion has been exploited as a range of salts, it has rarely been used as the dichloroiodate salt. Mitra *et al.*[156-159] however, published several papers discussing halogenation using solid-supported BTEA-polyhalide salts. The BTEA-based reagent was used in preference to BTMA-ICl_2 by the Harrowven group, as it was believed to afford improved yield and selectivity.

2.2.1.1 Synthesis and Reaction of Benzyltriethylammonium dichloroiodate

Literature methods for synthesising BTMA-ICl_2, include reacting BTMA-chloride with iodine monochloride in 86% yield,[152] or with iodine and silver trifluoroacetate.[160] The first of these methods was attempted using BTEA-Cl **72** instead, however, only a gum containing a mixture of salts was isolated. Attention was subsequently turned to a cleaner route, published by Kosynkin and Tour,[154] which used standard-grade reagents (bleach, NaI, HCl) in aqueous medium. This was tested on a 10 g scale, affording BTEA-ICl_2 **71** in 96% yield.

The solution phase iodination conditions of Kosynkin *et al.*[154] and the solvent-free conditions of Hajipour *et al.*[161] (in which the reagent Me_4N^+ ICl_2^- was ground to a fine

powder with the substrate), were both tested in TLC-scale reactions using BTEA-ICl_2 **71**, although the latter method proved inefficient for the BTEA salt. When scaled up on a TIPDS-protected $^{6\text{-oxo}}$C derivative **73**, synthesised by Dr. Osbourne, despite a clean conversion, only the chlorinated nucleoside **74** was afforded in 46% yield.



Scheme 2.2. Halogenation of 6-oxocytidine derivative **73** using BTEA-ICl_2 **71**.

This could be rationalised by dissociation of the dichloroiodate anion forming ICl (brown colouration observed) *in situ*, which then acted as a chlorinating agent. Boden *et al.*[150] published the ability of ICl to chlorinate very electron-rich or sterically hindered substrates, *via* a single electron transfer from substrate to ICl.

2.2.1.2 Other Methods of Halogenation of 6-Oxocytidine derivatives

Pan *et al.*[162] published the halogenation of TIPDS-protected 6-amino-dU derivatives, (structural isomers of 6-oxocytidine) using bromine/iodine in moderate to good yields. Reporting higher yields for bromination, this method was tested on acetylated $^{6\text{-oxo}}$C derivatives. The bromine solution was added dropwise *via* a steel syringe needle, and became darker with a black precipitate of $FeBr_3$, which was thought to catalyse the reaction. Concurrently, Dr. Y. Wang attempted to brominate 6-oxocytosine **62** with no success or very long reaction time by adding the bromine solution *via* glass pipette. Addition of a small quantity of $FeBr_3$ significantly accelerated the reaction, thus supporting this hypothesis. An anomeric mixture of furanose/pyranose $^{6\text{-oxo}}$C nucleosides **75-78** was reacted with bromine, catalysed by 1.0 mol% $FeBr_3$, to afford the brominated products **79-82** in 51% total yield.

After obtaining the anomerically pure 6-oxocytosinyl-β-ribofuranoside **75** (Section 2.2.3), a cleaner and higher yielding halogenation reaction was sought, to produce a large quantity of brominated/iodinated $^{6\text{-oxo}}$C. Despite the moderate yield and long

reaction time necessary for iodination of $^{6\text{-oxo}}$C using NIS, attention was turned back to *N*-halosuccinimides using the alternative conditions of Hirota *et al.*.[163]

Hirota *et al.* brominated several 4*N*-alkyl-1,3-dimethyl-6-oxocytosine derivatives (Figure 2.3) using *N*-bromosuccinimide (NBS), with the radical initiator AIBN, suspended in THF, in 80-95% yield. 6-Oxocytidine derivative **75** was transformed cleanly under these conditions, using only 2.0 mol% AIBN in 30 minutes, to afford the product **79** in 95% yield.

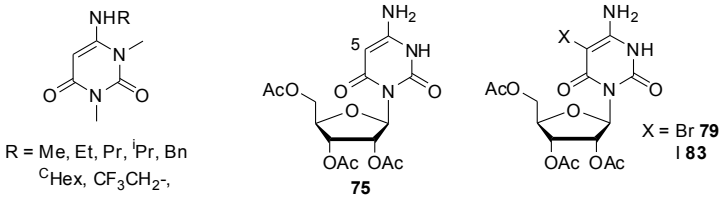
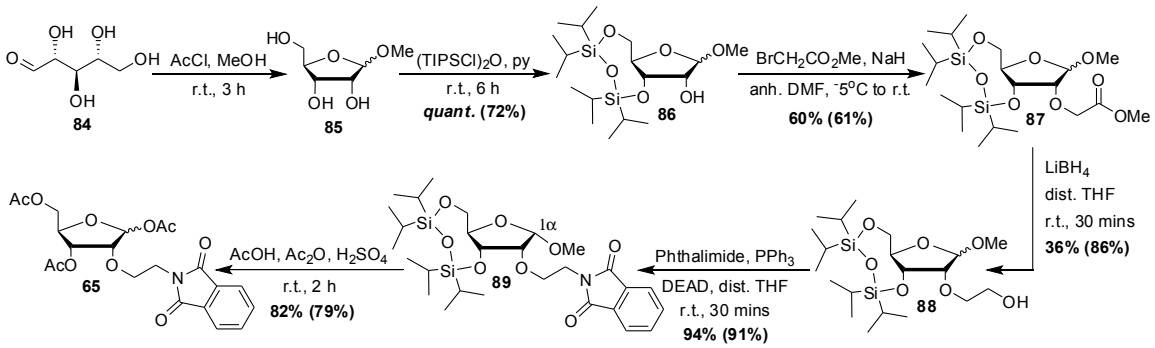


Figure 2.3. 4*N*-Alkyl-1,3-dimethyl-6-oxocytosine derivatives, acetylated $^{6\text{-oxo}}$C **75** and halogenated $^{6\text{-oxo}}$C nucleosides **79,83**.

Iodination using *N*-iodosuccinimide (NIS) under the same conditions resulted in a brown colouration. Presumed to be due to decomposition of NIS in storage, a large quantity was recrystallised (dioxane/chloroform)[164] and the reaction was repeated. Despite the same brown colouration whether conducted at rt or 0 °C, on a 5 g scale the reaction was complete within 15 minutes. Partial or almost complete deiodination of product occurred after aqueous workup, using dilute aq $Na_2S_2O_3$ or $Na_2S_2O_5$ respectively, as previously observed (Section 2.2.1). The sodium thiosulfate ($Na_2S_2O_3$) workup, however, was used successfully on a 0.5 g quantity of the reaction mixture, separating the layers as soon as decolourisation occurred. Suspension of the remainder in ethyl acetate and soaking over 2 days, however, removed all excess iodine and succinimide by-product, affording the product **83** in 82% combined yield.

2.2.2 Synthesis of 2'-*O*-Phthalimidoethyl-modified Ribofuranose Sugar

The 2'-*O*-phthalimidoethyl-modified ribofuranose sugar **65**, synthesised and used by Osborne *et al.*[37] in preparation of the AT-recognition monomer ^{BA}U, was prepared for coupling to 6-oxocytosine **62** (Scheme 2.2), as part of the synthetic route to the ^{BAox}C monomer **64**.

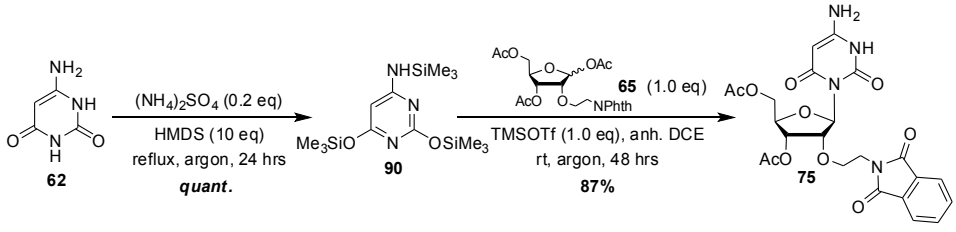


Scheme 2.2. Synthesis of 2'-modified ribofuranose sugar **65** (*lit.*[37] yields in parentheses).

The sugar synthesis starts with formation of the 1'-*O*-methyl acetal of D-ribose **85**, followed by protection of the 3'- and 5'-OH groups with the bidentate tetra-isopropylsiloxane-1,3-diyl (TIPDS) group, using Markiewicz Reagent.[165] The reaction proceeded well over two steps, with minor impurities and solvent traces, which were removed during the following steps (Scheme 2.3).

Alkylation of the 2'-OH with methyl bromoacetate gave the product 2'-*O*-methylethanoyl sugar **87** in a moderate yield, containing a significant methyl bromoacetate impurity which co-eluted with the product. Subsequent reduction of the ester moiety using $LiBH_4$ therefore, afforded a lower yield of **88** than in the literature. Phthalimide substitution *via* Mitsunobu conditions[166] gave the TIPDS-protected sugar **89** (α-anomer only) and subsequent acetolysis gave the final 2'-modified ribofuranose sugar **65** in good yield.

This would then be coupled with 6-oxocytosine **62** (Scheme 2.3), once halogenation and Sonogashira reactions were optimised.

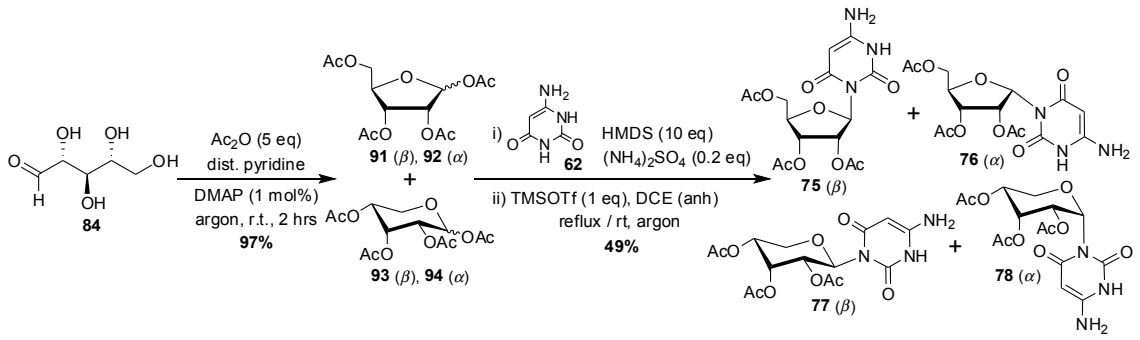


Scheme 2.3. Sugar-base coupling under Vorbrüggen's[167] conditions (in house).

2.2.3 6-Oxocytidine Synthesis

Due to the expensive nature of Markiewicz reagent (£100/g, now ~£25/g), an alternative sugar moiety was needed for synthesis of a 6-oxocytidine nucleoside, on which halogenation and Sonogashira reactions could be evaluated. 1',2',3',5'-Tetra-*O*-acetyl-D-ribose **91** was chosen, as the synthesis is straightforward and the reagents are cheap.

The first stage was acetylation of D-ribose **84**, which was first effected using acetic anhydride in pyridine, with a catalytic amount of DMAP (from in house method). Following purification, this afforded a 97% yield, of a 1:1.6 ratio of ribofuranose (α:β *ca.* 1:7) to ribopyranose (α:β *ca.* 1:6.5). Persilylation of 6-oxocytosine **62** by refluxing in HMDS, followed by coupling to the sugar mixture **91-94** with trimethylsilyl triflate in DCE at rt, afforded an inseparable mixture of β-ribofuranoside **75** and β-ribopyranoside **77** in a 49% yield. Only trace amounts of the α-anomers **76,78** were were isolated due to neighbouring group participation from the 2'-acetato group (Scheme 2.4).



Scheme 2.4. Synthesis of 2',3',5'-tri-*O*-acetyl-6-oxocytosinyl-ribosides **75-78**.

This β-riboside mixture **75,77** was subsequently brominated, as described in Section 2.2.1.2, for Sonogashira cross-coupling.

The desired β-ribofuranoside **75** was synthesised selectively by converting D-ribose **84** into the 1'-*O*-methylribofuranoside **85**, used in synthesis of the 2'-*O*-phthalimidoethyl-β-ribofuranose sugar **65** in Section 2.2.2, followed by acetolysis/acetylation to give the β-anomer of the product sugar **91**, in 39% yield. After crystallisation, the remaining syrup, containing *ca.* 65:35 ratio of α:β-anomers, was not repurified, and further crystallisation of the β-anomer could not be induced by ice-cold ethanol or by seeding.

The sugar-base condensation was conducted under Vorbrüggen's conditions[167] as described in Scheme 2.2. No chromatography was required, as the product nucleoside **75** crystallised out of solution during workup, in 70% yield (16.9 g). The nucleoside was brominated/iodinated using NBS/NIS, as described Section 2.2.1.2.

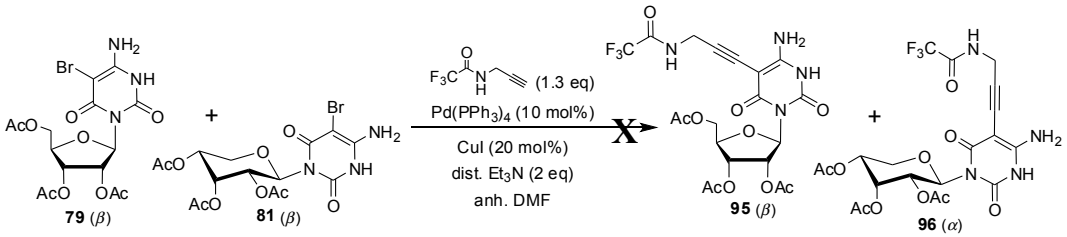
2.2.4 Sonogashira Cross-Coupling[123]

There has been much research on palladium-catalysed alkynylation reactions. Modifying Sonogashira's[123] original conditions (CuI, $(Ph_3P)_2PdCl_2$, Et_2NH) for nucleosides, Robins and Barr's[116,168] use of warm triethylamine and Hobbs' conditions for alkynylation of nucleosides[124] are two popular methods employed in synthesis of modified nucleosides.

Prior to carrying out Sonogashira reactions, the stability of the iodinated **83** and brominated nucleosides **79** to heat, light and base (Et_3N, piperidine) in THF, DMF and methanol was tested. TLC and MS analysis indicated no decomposition of any kind (deacetylation, dehalogenation), hence suggesting they should be stable under Sonogashira conditions.

A range of conditions and reagents for the Sonogashira[123] cross-coupling of a protected aminoalkyne, were tested on the brominated derivatives of the β-riboside mixture **79,81**, brominated nucleoside **79** and iodinated nucleoside **83**.

The brominated β-riboside mixture **79,81** was treated with *N*-trifluoroacetyl propargylamine **68** (2 eq) under Hobbs' conditions;[124] $Pd(PPh_3)_4$ (10 mol%), CuI (20 mol%) and Et_3N (2 eq) in DMF (Scheme 2.5).



Scheme 2.5. Sonogashira cross-coupling reaction between *N*-trifluoroacetyl propargylamine **68** and brominated 6-oxocytosinyl β-ribosides **79,81**.

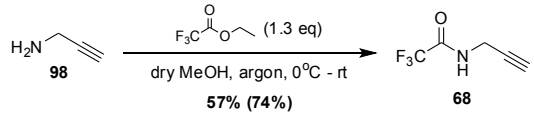
TLC analysis indicated a complex mixture and only starting material and debrominated starting material were isolated after purification (analysed by MS). It is unusual, yet possible that debromination occurred *via* hydride-insertion after Pd-insertion into the C-Br bond.

Triethylamine was the base of choice, because it is non-nucleophilic and there is no possibility of substitution of bromine. It was shown by Wang *et al.*,[169] that triethylamine also inhibits alkyne dimerisation/homo-coupling and promotes the desired cross-coupling reaction. Diethylamine, as used by Kenkichi Sonogashira, was shown to promote homo-coupling under the right conditions.

Using a large excess of Et_3N (20 eq) or as a solvent, on 50 mg and 500 mg scale, resulted only in partial debromination, deacetylation, and triphenylphosphine oxide. A similar observation was made for the iodinated nucleoside **83**.

The last reactions on the brominated *β*-riboside mixture **79,81** were carried out under microwave heating. The cross-coupling reactions were conducted in DMF or THF at 60 °C for 10 mins, and in DMF at 100 °C for 10 mins. Once again, only complex mixtures were formed.

A fresh quantity of *N*-trifluoroacetyl propargylamine[117,170] **68** (Scheme 2.6) was synthesised for subsequent Sonogashira reactions; the quality of which was confirmed by standard characterisation methods and in a test Sonogashira reaction using 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine (DMT-IdU) **97** under Hobbs' conditions.[124] However, subsequent test reactions using the brominated **79** and iodinated **83** $^{6\text{-oxo}}$C derivatives, were unsuccessful. Only partial de-iodination of starting material **83**, and formation of triphenylphosphine oxide (in both cases) were observed.



Scheme 2.6. Synthesis of *N*-trifluoroacetyl propargylamine **68**, (*lit.*[117] yield in parentheses).

It was proposed that the above problems might be exacerbated by the presence of acetyl groups hence a suitable method for deacetylation was sought. The brominated **79** and iodinated **83** $^{6\text{-oxo}}$C nucleosides were deacetylated by heating in 0.4 M K_2CO_3 in

methanol/water (4:1, v/v) (mild reagent occasionally used in oligonucleotide deprotection) until fully dissolved. Treatment with DOWEX 50WX8 ion exchange resin (200/400 mesh, pyridinium form) monitoring the pH until <7, soaking the resulting solid (after filtration and concentration) in cold ethanol or isopropanol and drying, afforded the 'free' bromo- **99** and iodo-nucleosides **100** in good to high yield. If too little DOWEX resin was used, this resulted in a hygroscopic potassium salt of the nucleoside, which required further treatment.

A 0.5 g scale Sonogashira reaction was executed using the deacetylated iodinated nucleoside **100** and alkyne **68**, under the same conditions. However, complete deiodination was observed over 2 hours, to give the deiodinated starting material, 6-oxocytidine **101** in 44% yield (Figure 2.4).

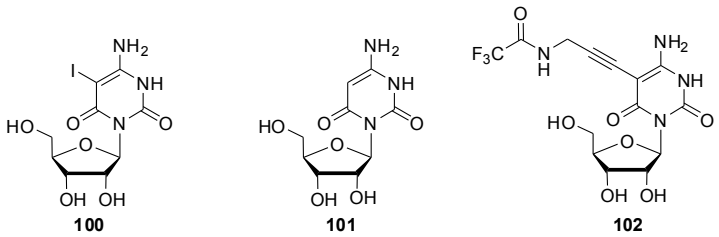


Figure 2.4. Starting material **100**, desired Sonogashira product **102** and 6-oxocytidine **101**.

As little success had been achieved using modifications of Hobbs' conditions, other catalysts and conditions were assessed. Recently, Alami *et al.*[171] published cross-coupling reactions effected by the strongly active (*bis*-benzonitrile)palladium (II) dichloride – copper (I) iodide catalyst system. Using this system they were able to conduct cross-couplings of vinyl chlorides with a variety of alkynes, in dry piperidine. The weakly-coordinated PhCN ligands, allow for much faster oxidative insertion of Pd into the C–X bond (even less reactive C–Cl bonds, at room temperature), and with excellent yields (62-97%) and reaction times (~30 minutes).

The above reaction conditions ($(PhCN)_2PdCl_2$ – 5 mol%, CuI – 10 mol%, alkyne **68** – 2 eq, dry piperidine, rt, argon atmosphere), however proved ineffective on 6-oxocytidine substrates. TLC analysis of the reaction with iodinated nucleoside **100**, indicated only deiodination. TLC and MS analysis of the reaction with 5-bromo-6-oxocytidine **99**, indicated nothing had occurred, with or without DMF as a co-solvent (to aid dissolution). The reactions could not be left longer than 90 minutes or heated due to catalyst sensitivity.

The *bis*-(triphenylphosphine)palladium (II) dichloride (10 mol%) – copper (I) iodide (30 mol%) catalyst system, as used by Robins and Barr[116,168] for cross-coupling reactions with 5-iodo-2'-deoxyuridines in warm triethylamine (50 °C) was also assessed. Even when using 4 equivalents of alkyne, only deiodination and alkyne dimerisation were detected for substrate **100** and nothing was observed for the brominated substrate **99**.

Alternatively, instead of conducting cross-couplings in absence of water, the reaction was conducted in hot water, using 0.5 mol% $(Ph_3P)_4Pd$, as published by Bhattacharya and Sengupta.[172] The authors' principal idea was to utilise the hydrophobic effect, and although the 6-oxocytidine substrates are (partially) water-soluble, it was hoped these conditions might be beneficial. However, again only deiodination of **100** was observed.

These reactions, using various catalysts, solvents, temperatures and ratios of reagents were all unsuccessful. Brominated $^{6\text{-oxo}}$C nucleosides (β-anomers) did not react, and iodinated $^{6\text{-oxo}}$C nucleosides were deiodinated. This illustrates the surprising strength of the C–Br bond and susceptibility of the R–Pd$(PPh_3)_2$–I intermediate to decomposition. There may also be other steric and electronic factors preventing this transformation, such as the neighbouring C^6-carbonyl and C^4-amino groups. Although other alkynes were not tested, *N*-trifluoroacetyl propargylamine **68** couples to protected and unprotected 5-iodo-2'-deoxyuridine nucleosides in excellent yield.[117,170]

Sonogashira-like palladium-catalysed cross-coupling reactions were subsequently abandoned. It is still possible, however, that results could be obtained using a triflato-nucleoside,[59] alkynylzinc reagents[173,174] or even Stille coupling,[175] which have been successfully used to alkynylate other nucleosides.

2.3 Alternative Modifications at 5-position

Attempts to enhance base-stacking by attachment of an alkynyl group failed as described above. It was also found that the 5-allyl-group of $^{6\text{-oxo}}$dC caused destabilisation of the triplex, relative to $^{6\text{-oxo}}$dC alone. An alternative method was sought therefore for attachment of a pendant amine group at the 5-position of $^{6\text{-oxo}}$C, with or without additional base-stacking enhancement, which would increase triplex-stability by charge-stabilisation.

2.3.1 Halide-Amine Exchange

Halide-amine exchange was proposed as an alternative method for introduction of amine functionality to the base. There is much literature precedent for this transformation on aromatic compounds, uridine nucleosides, uracils and aminouracils, using primary and secondary amines and even ammonia. In 1951, Arthur Phillips published a series of substitution reactions on 5-bromouracil, using amines such as methylamine, butylamine, cyclohexylamine, piperidine and benzylmethylamine in typically good to quantitative yields.[176] Benzylmethylamine was used for introduction of a methylamino group (following hydrogenolysis of the benzyl group), as reaction of similar heterocycles with methylamine and ammonia often requires elevated temperatures[177] or high pressure,[178,179] with long reaction times.

5-Bromo-6-oxocytidine **99** was heated at 85 °C in a sealed tube with four different amines (piperidine, benzylamine, propargylamine, TAEA **103**) for 3.5 hours (Figure 2.5). TLC analysis indicated loss of starting material, most likely through degradation, and no positive identification of desired products was made by MS analysis. A tan-coloured solid was isolated from the last reaction with the polyamine *tris*(2-aminoethyl)amine, TAEA **103**, which was confirmed by NMR, however, to be a salt of TAEA.

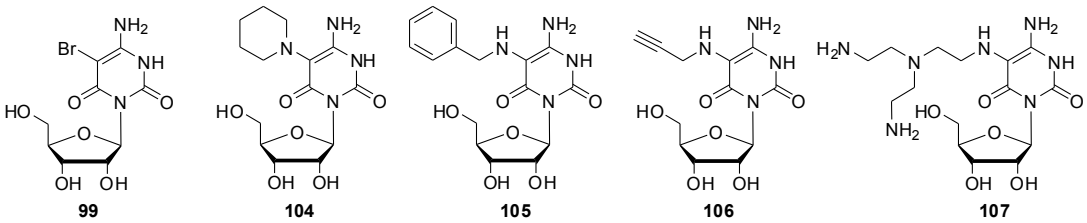


Figure 2.5. 5-Bromo-6-oxocytidine **99** and amine-halide exchange products **104-107**.

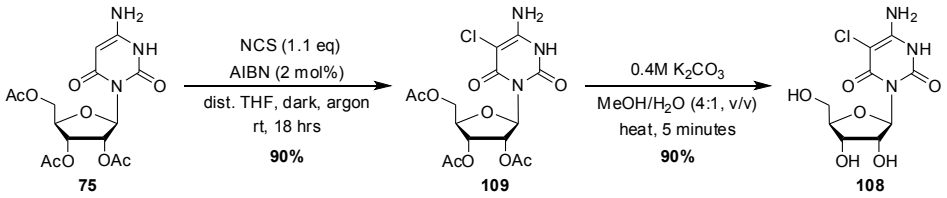
Azam *et al.*[180] attached the polyamine TAEA **103** to an anthraquinone intercalator-modified chimeric TFO, by a simple chloride-amine exchange reaction. The polyamine moiety demonstrated increased triplex stability without altering selectivity, and has since been used, in house, by Dr. Z. Zhao, for synthesising simple polyamine-anthraquinone intercalators for non-specific triplex stabilisation.[181]

Initially, test reactions were performed on 2',3',5'-*O*-triacetyl-5-bromo-6-oxocytidine **79**, heating the nucleoside at 80 °C in a sealed tube, for 2.5 hours in piperidine or benzylamine. Following treatment with DOWEX 50WX8-400 (pyridinium form) and diethyl ether, a white gum containing some deacetylated piperidinyl compound **104**, and deacetylated starting material **101** were obtained respectively. Following this, to avoid complex mixtures due to partial deacetylation, the free nucleoside 5-bromo-6-oxocytidine **99** was used. This was treated with propargylamine, benzylamine and TAEA **103**.

Varying the reaction conditions for bromide-TAEA exchange was unsuccessful. Heating 5-bromo-6-oxocytidine **98** with TAEA **103** at 180 °C for 20 minutes, heating at 75 °C in a 30% aq solution of the amine overnight, and refluxing in distilled pyridine with 5 eq of TAEA **103** produced similar results as obtained previously and isolation of amine salts.

In order to confirm that exchange had occurred successfully originally, when using piperidine, nucleoside **99** was heated in the amine alone, or with pyridine, water or DMF as co-solvent at 70 °C for 17 hours. This proved partially successful, with an estimated 20-30% conversion (TLC analysis) for all reactions, except in water. These results were also confirmed by MS analysis.

The amine-halide exchange was proposed to occur *via* an S_NAr mechanism, which if true would suggest a faster, more efficient reaction with chloro- and fluoro-, rather than bromo- or iodo-groups. To take advantage of this, and to avoid problems due to cleavage of acetyl protecting groups, a large quantity of unprotected 5-chloro-6-oxocytidine **108** was synthesised for comparison with the other 5-halo-$^{6\text{-oxo}}$C derivatives. 2',3',5'-Tri-*O*-acetyl-6-oxocytidine **75** was treated with *N*-chloro-succinimide (NCS) based on the method published by Hirota *et al.*[163] as described in Section 2.2.1.2 (Scheme 2.7). The product was purified by column chromatography, not soaked in ethyl acetate to remove the succinimide by-product, as it is partially soluble in ths solvent.



Scheme 2.7. Synthesis of 5-chloro-6-oxocytidine **108**.

The product **109** was deacetylated as previously described to afford the desired free nucleoside **108** in 81% yield over two steps, containing 1.8 mol % succinimide from the previous step and 35.6 mol % KOAc, not observed for previous deacetylations by this method.

In the next series of test exchange reactions, chlorinated **74,108** and iodinated **100** $^{6\text{-oxo}}$C nucleosides were treated with piperidine, and brominated nucleoside **99** with five 1,n-diaminoalkanes (n = 2, 6, 8, 10, 12), at 70 °C for 18-20 hours (Figure 2.6).

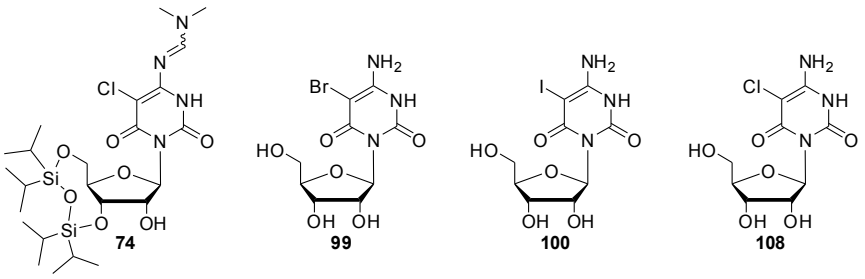


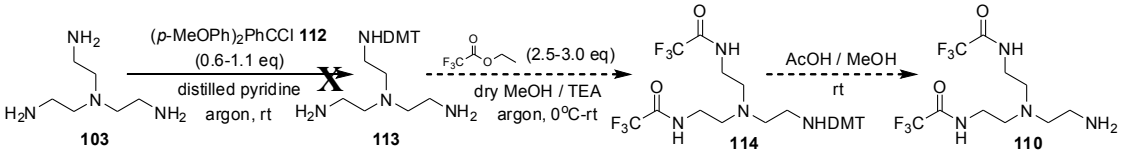
Figure 2.6. Substrates for amine-halide substitution reactions **74,99,100,108**.

5-Iodo-6-oxocytidine **100** reacted partially with piperidine but less cleanly than 5-bromo-6-oxocytidine **99**, suggesting an S_NAr mechanism might apply. The protected chlorinated $^{6\text{-oxo}}$C derivative **74**, however, underwent deprotection of the 4*N*-dimethylformamidine protecting group, with no desired reaction, and none of the 1,n-diaminoalkanes reacted with 5-bromo-6-oxocytidine **99**.

Chloro-$^{6\text{-oxo}}$C derivative **74** was treated with ethylene diamine (42 eq) and TAEA **103** (18 eq) at 75 °C for 90 minutes and also for 18-20 hours. Only formamidine deprotection and amine degradation were observed, and trifluoroacetylation of the latter reaction mixture yielded only trifluoroacetylated amine.

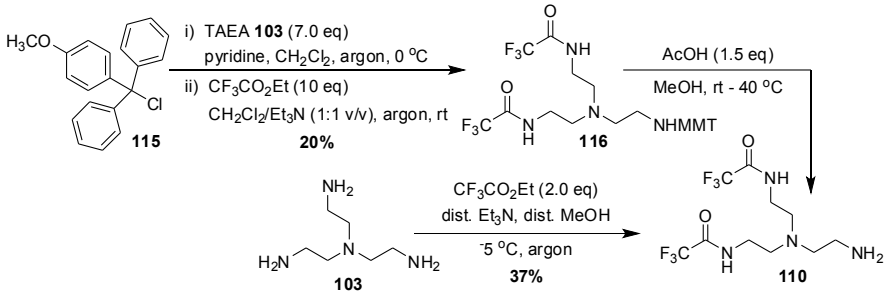
So far, successful amine-halide exchange had been limited to the secondary amine piperidine, which is probably not beneficial to triplex formation. Attempts to couple

with primary amines, and in particular, polyamines, although unsuccessful thus far, were continued using 5-chloro-6-oxocytidine **108**. It was thought, that protection of two of the primary amine moieties of *tris*(2-aminoethyl)amine **103**, could potentially aid the reaction by presenting only one centre for reaction, although the *bis*-protected product was finally not used. Asseline *et al.*[182] published the synthesis of the *bis*-trifluoroacetylated polyamine **110** *via* alkylation of the trifluoroacetic acid salt of *bis*-trifluoroacetylated diethylene triamine **111** with bromoethanol, followed by Mitsunobu reaction with LiN_3 and hydrogenation over Pd/C, in 23% overall yield.



Scheme 2.8. First route to *N*-(2-aminoethyl)-*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine **110**.

The first synthetic procedure attempted (Scheme 2.8) began with mono-protection with a 4,4'-dimethoxytrityl (DMT) group, by addition of a solution of DMTCl **112** in distilled pyridine, dropwise over 20 minutes, to a stirred, ice-cold solution of TAEA **103** (6 eq) in distilled CH_2Cl_2. After aqueous workup, an off-white foam was afforded, which unfortunately contained only *bis*- and *tris*-tritylated products.



Scheme 2.9. Other routes to *N*-(2-aminoethyl)-*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine **110**.

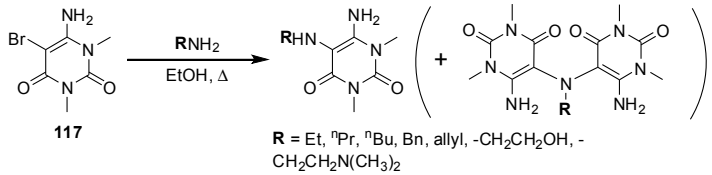
The reaction was repeated using MMTCl **115** due to issues of stability with DMT, using 7 eq of polyamine **103** and dropwise addition of the MMTCl/pyridine solution over a 5½ hour period. This afforded, after workup, a green syrup/resin. Trifluoroacetylation using ethyl trifluoroacetate (10 eq), with excess Et_3N to prevent detritylation, afforded the fully-protected compound **116** in a low 20% overall yield. The detritylation step was carried out on a test scale, using acetic acid (30 eq) in methanol with heating to afford a white oil, which could not be identified by MS analysis (Scheme 2.9).

Free amine **103** was also trifluoroacetylated directly using only 2 equivalents of ethyl trifluoroacetate, which selectively protected only two amines in 37% yield. Indeed, the selectivity obtained using ethyl trifluoroacetate is known, as published by Yang *et al.*,[183] who described the selective protection of two and three chemically-equivalent amines in 1,4,7-triazacyclonane and cyclen respectively, despite using excess reagent.

Finally, after testing the substrate's stability to DMF and pyridine, at 100 °C overnight, a number of test amine-substitution reactions on 5-chloro-6-oxocytidine **108** were carried out. The substrate was heated, under argon as described previously, at 100 °C in DMF, with TAEA **103** (10 eq), diethylene triamine (10 eq) and piperidine (60 eq) for 21 hours; and with piperidine (30 eq), piperidine (30 eq) with KF (2.5 eq), TAEA **103** (30 eq) and 6-aminocaproic acid (30 eq) for up to 2½ days. Once again, only piperidine proved suitably nucleophilic to react. It was difficult to determine if KF had any effect on this reaction.

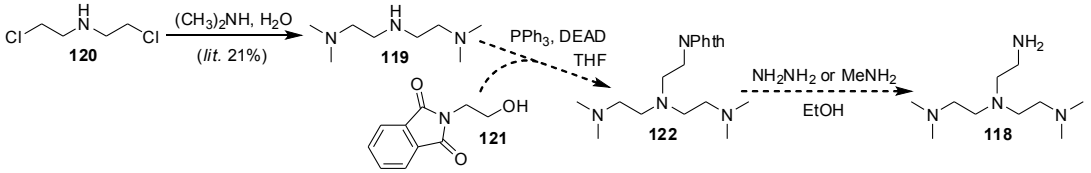
These studies indicated only the strongly nucleophilic amine piperidine was able to displace a halogen at the 5-position of 6-oxocytidine.

There is literature precedent for reaction of primary amines with 5-bromo-6-oxocytosine (5-bromo-6-aminouracil) derivatives (e.g. 5-bromo-1,3-dimethyl-6-aminouracil, **117**). Brederick *et al.*[184] achieved efficient reaction with benzylamine, *n*-butylamine and ethanolamine and Searle & Co.[185] with 2-(*N,N*-dimethylamino)ethylamine, by heating in ethanolic solution. Rybar *et al.*[186] report yields of 50-80%, using simple amines such as propylamine, allylamine and benzylamine, but encountered further reaction with the product to form *bis*-6-oxocytosinylamine compounds, even when using vast excess of the amine. Zhang and Zhang[187] also reported efficient reaction at 120 °C, under microwave irradiation, to afford yields of 70-95% for several arylethylamines (Scheme 2.10).



Scheme 2.10. Bromide-amine exchange reaction on 5-bromo-1,3-dimethyl-6-oxocytosine **117**.[184-187]

Based on the literature evidence, it was thought possible to attach a polyamine or similar amine moiety to the base then couple to the sugar to form the modified nucleoside. However, reaction with the desired amine TAEA **103** may give multiple additions, hence a protecting group strategy or methylated-analogue of TAEA **118** would be needed, synthesis of which could be envisaged to proceed *via* a tetramethyl diethylene triamine **119**.[188]

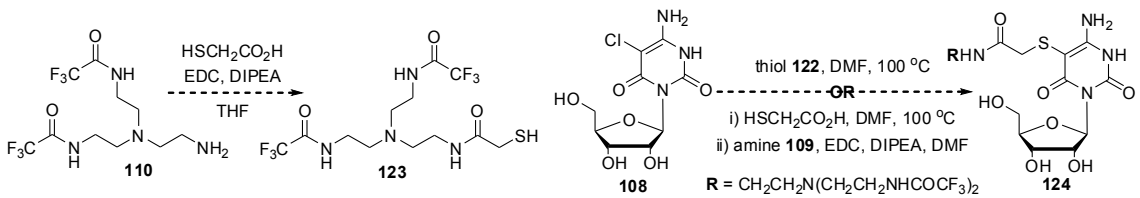


Scheme 2.11. Proposed synthetic route to *N,N,N',N'*-tetramethyl-*tris*-(2-aminoethyl)amine **118**.

It is also unknown whether such a group would affect the sugar-base coupling and this route would require much optimisation. It may also be possible therefore to exchange Br with NH_3, for further base-modification (4*N*,5*N*-cyclisation reactions to increase base-stacking), however, the products (5,6-diaminouracils or 5-amino-6-oxocytosines) are unstable due to oxidative self-condensation.[189]

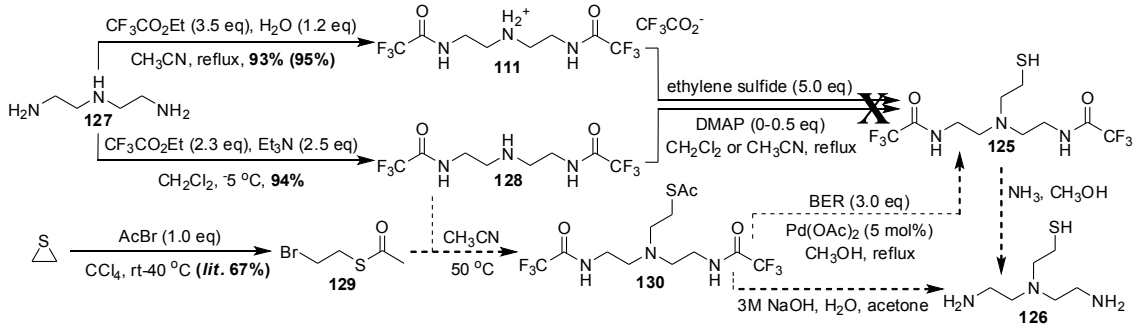
2.3.2 Halide-Thiol Exchange

Amine-modified alkanethiols were chosen as the next candidate for nucleophilic substitution of chloride. Being more nucleophilic than amines and with evidence of chloride-thiol exchange in the literature,[190] this was considered the logical next route. Attachment of a mercaptocarboxylic acid was considered for coupling of various amines, including TAEA **103** or its *bis*-trifluoroacetyl derivative **110**. The amine could be coupled to the acid, before or after halide-thiol exchange (Scheme 2.12).



Scheme 2.12. Proposed synthesis of thiol-modified amine **123** and coupling with 5-chloro-6-oxocytidine **108**.

5-Chloro-6-oxocytidine **108** was treated with mercaptoacetic acid (10 eq) in DMF at 100 °C for 18 hours, resulting in formation of a new product and near complete consumption of starting material. A hygroscopic powder was obtained on precipitation with methanol and diethyl ether, and a second powder was obtained on ether-induced precipitation from the filtrate, the coupled product could not be detected by MS.



Scheme 2.13. Proposed synthetic routes to polyaminethiol derivatives **125,126**, (*lit.*[182,191,192] yields in parentheses).

The second proposal was synthesis of mercapto-analogues of TAEA **125,126**, the attempted synthesis of which is described below (Scheme 2.13). The seven-step synthesis of the non-acetylated analogue **126** was published by Cheng *et al.*[193] in 13% overall yield. Either thiol could then be reacted with the desired 5-chloro-$^{6\text{-oxo}}$C derivative.

Selective trifluoroacetylation of the primary amine groups of diethylene triamine, DETA **127** was first achieved by trifluoroacetylation in presence of water to give the trifluoroacetate salt[182,192] **111**. The large quantity of salt **111** could not be converted to the free amine using Et_3N/heat or polyamine-modified Amberlite IRA-96 resin, hence the free amine **128** was synthesised.[194] This was achieved by reaction of DETA **127** with a slight excess of ethyl trifluoroacetate, under anhydrous conditions. Neither the salt **111** nor free amine **128**, however, reacted with ethylene sulphide (ES) to introduce the 2-mercaptoethyl side-chain, despite strong literature precedent for alkylation of secondary amines with ES,[195,196] and of the salt **111** with alkyl bromides such as bromoethanol.[182] . The alternative route would proceed *via* alkylation using *S*-acetyl-2-bromoethane thiol **129**, synthesised from ES and AcBr (*lit.* 67% yield),[191] followed by selective cleavage of the *S*-acetyl group, using borohydride-functionalised Amberlite IRA-400 resin and $Pd(OAc)_2$.[197] The resin is prepared by stirring the wet chloride-form with an aqueous solution of $NaBH_4$. The deprotection of octyl thioacetate was reported

to proceed cleanly in 98% yield. Alternatively, full deprotection using 3M aq NaOH in acetone[198] should afford the free mercapto-polyamine **126**.

5-Chloro-6-oxocytidine **108** was also treated with cysteamine (2-aminoethane thiol) (10 eq) to assess whether a free amine moiety would affect this transformation. The starting material was consumed over 16 hours and low yields of a white powder and a very hygroscopic tan-coloured powder were obtained by recrystallisation from methanol and precipitation of the filtrate with methanol/diethyl ether respectively. No components could be identified by MS, but this does not disprove possible reaction, which, however, would need substantial optimisation.

Further work was halted at this stage, as attention was transferred to the CG recognition project (Chapters 3-7), which was running alongside this research, and obtaining positive results synthetically and in biophysical studies.

2.4 Alternative Sugar Modification at the 2'-Position

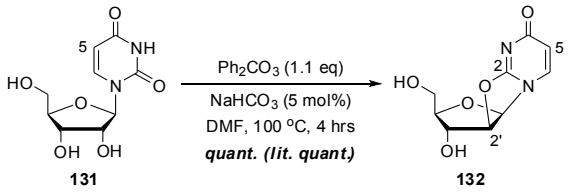
As an alternative sugar moiety, synthesis of a 2'-*O*-methyl-$^{6\text{-oxo}}$C derivative was also investigated, for comparison with published biophysical studies, by a different route to that already published.[70,71] Whilst lacking the charge-stabilising properties of the 2'-*O*-aminoethyl group, the 2'-*O*-methyl moiety still increases stability to enzymic degradation *in vivo*.[199]

2.4.1 2'-Modification *via* Tricyclic Anhydronucleosides

The 2'-*O*-phthalimidoethyl moiety of the monomer ^{BA}U was introduced onto the sugar *via* a six-step synthesis, before coupling to the base.[37] There is scope, however, for introducing 2'-*O*-aminoalkyl modifications on the nucleoside itself, *via* a tricyclic nucleoside intermediate.

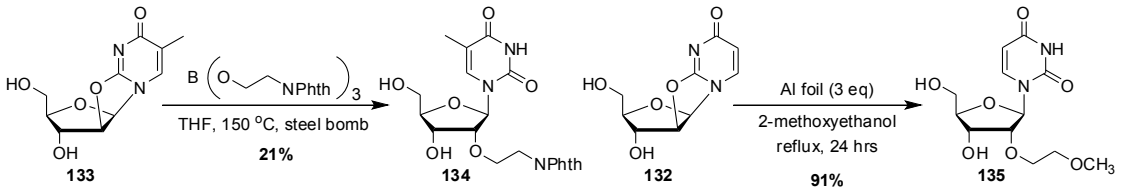
There has been much published into cyclisation reactions of uridine derivatives to form their tricyclic anhydrouridine intermediates. In most cases, HMPA is the chosen solvent, however, due to issues of carcinogenicity, DMA and DMF have been successfully employed, although in varying yields.

In work reported by Reese *et al.*,[200,201] when uridine **131** is heated with diphenyl carbonate (1.1 eq) and a catalytic amount of $NaHCO_3$ (5 mol%) in dry DMF or DMA at 100 °C, for 4-5 hours, the tricyclic compound 2,2'-anhydrouridine **132** can be formed in quantitative yield (Scheme 2.14). The reaction was executed successfully using 5-iodouridine by Dr. L. Brennan, in house, and a methyl group at the 5-position can also be tolerated.[202,203]



Scheme 2.14. Cyclisation of uridine **131** to form 2,2'-anhydrouridine **132**.[200,201]

Manoharan *et al.*[42] and Ross *et al.*.[203] describe introduction of a 2'-*O*-phthalimidoethyl group by the following method. 2-(Phthalimido)ethanol is converted into the trialkoxyborane, by slowly dissolving in a solution of diborane in THF, then reacted with 2,2'-anhydro-5-methyluridine **133** or 2,2'-anhydrouridine **132** in presence of $NaHCO_3$ (cat.) in a steel bomb, at 150 °C for 24 hours or 175 °C for 72 hours. Following purification by column chromatography, only ~20% yield is achieved (Scheme 2.15).



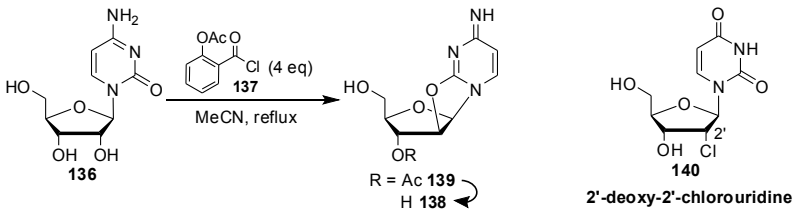
Scheme 2.15. 2'-Modification *via* anhydrouridine derivatives **132,133**.[42,203]

Blommers *et al.*[103] also reported the same coupling using $Ti(O^iPr)_4$ instead of diborane, to form the titanium alkoxide *in situ*, forming an isotopically-labelled version of **134** in 32% yield.

Reese *et al.*[200,201] have also reported the reaction of 2-methoxyethanol with 2,2'-anhydrouridine **132**, *via* the aluminium alkoxide. Refluxing coarsely divided aluminium foil in dry 2-methoxyethanol until dissolution, then reaction with the 2,2'-anhydronucleoside **132**, afforded the 2'-*O*-modified nucleoside after 48 hours in 91% yield. This simpler, yet more effective reaction was executed in house, by Dr. L Brennan, to insert an *N*,*N*-dimethylaminoethoxy group in ~95% crude yield.

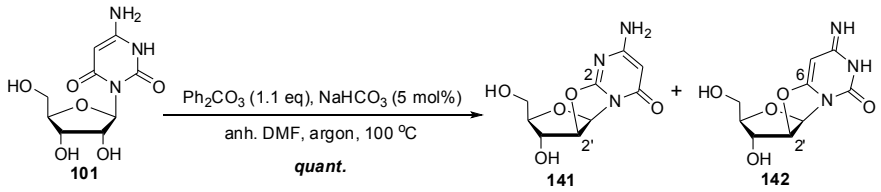
After testing the $Ti(O^iPr)_4$ coupling (mentioned above), 2-(phthalimido)ethanol was successfully coupled to DMT-protected anhydro-5-methyl-uridine in ~45% yield, in house, *via* formation of the aluminium alkoxide, by heating aluminium foil in molten 2-(phthalimido)ethanol. This, however, has yet to be repeated.

Cyclisation of cytidine **136** can be achieved *via* the inexpensive reagent, 2-acetoxybenzoyl chloride **137**, and 2,2'-anhydrocytidine (cyclocytidine) **138** is also commercially available. Reichman *et al.*[204] report a 65% yield of 3'-*O*-acetyl-2,2'-anhydrocytidine **139** after refluxing for 25 minutes in acetonitrile. Deacetylation then afforded the desired product **138**. This method can also be applied to uridine. Heating at 50 °C for 3 minutes affords the cyclised product **132**. Refluxing for 30 minutes affords the 2'-deoxy-2'-chloro- derivative **140** (Scheme 2.16).



Scheme 2.16. Cyclisation of cytidine **136** and 2'-deoxy-2'-chlorouridine **140**.[204]

It was possible that both methods of cyclisation could be employed for 6-oxocytidine **101** and derivatives, as two carbonyl groups are available for reaction. It was also feasible that two products **141,142** are possible through either reaction, although either product could be used for subsequent 2'-modification.



Scheme 2.17. Cyclisation reaction of 6-oxocytidine **101**.

The cyclisation reaction was carried out twice (Scheme 2.17), using diphenyl carbonate, to afford each time a 1:1 mixture of 2,2'- **141** and 6,2'-anhydro-6-oxocytidine **142** in quantitative yield. The ^{1}H-NMR spectrum was complicated, but the sugar protons could be assigned and the spectrum contained two 1'-H signals corresponding to the two isomers.

2'-*O*-Methyl modification could then have been achieved by a similar method to Reese,[200,201] reacting 4 equivalents of $Mg(OMe)_2$ in methanol,[205] with DMT-protected derivatives of 2,2'- **141** and 6,2'-anhydro-6-oxocytidine **142**. The anhydro-nucleosides **141,142** were, however, not modified any further at this point.

2.5 Conclusion

Several synthetic investigations have been carried out, in order to assess reaction feasibility and subsequently to synthesise a *bis*-amino-modified 6-oxocytidine monomer **64**, for GC base pair recognition studies within DNA triple helices. Palladium-catalysed alkynylation was ruled out as a method of introduction of amine functionality at the 5-position due to substrate inactivity or degradation. Instead, modification by amine- or thiol-halide exchange was investigated with some success. Feasibility of 2'-modification was also investigated, although this requires further work.

Synthesis of a *bis*-amino-modified 6-oxocytidine monomer is reliant on several key factors. Attachment of an amine moiety to the 5-position must be achieved, either *via* a

thiol on the nucleoside or an amine on the base. This functionality must then also be tolerated by sugar-base coupling, cyclisation and methoxide insertion reactions. Due to the large number of steps involved, and extensive optimisation needed, unless large quantities of the initial intermediates can be synthesised, poor-moderate yields cannot be tolerated.

It is now unlikely that synthetic work on $^{6\text{-oxo}}$C will continue. However, should this not be the case, the 2'-*O*-methyl group is the most feasible 2'-modification. Whichever route would be chosen, large quantities of some intermediates can be synthesised in good yield.

Chapter 3

Amino-Modified Furano-dT Nucleosides for CG Recognition

3. Amino-Modified Furano-dT Nucleosides for CG Recognition

3.1 Introduction

The pyrimidine.purine base pairs CG and TA remain most difficult to target, as only one formal hydrogen-bond is presented for binding in the major groove. Thymine is the only natural base that can effectively recognise CG, but it cannot be used for CG recognition due to lack of selectivity. This discovery lead to a variety of monomers, including a series of modified pyrimidine-based nucleosides, designed to target the CG base pair, utilising new and extended Hoogsteen hydrogen-bonding patterns.

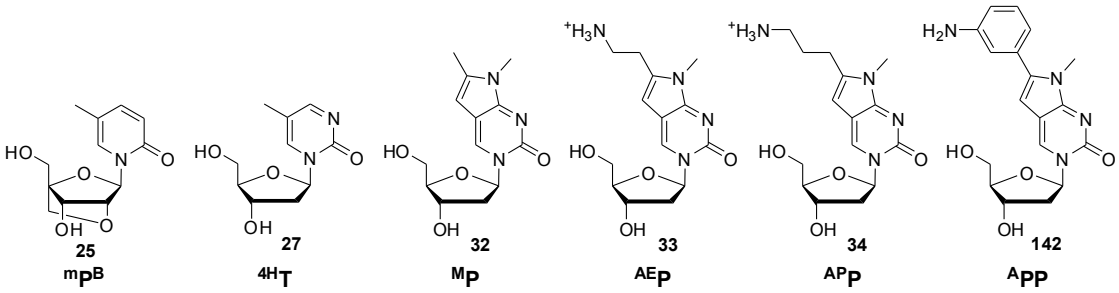


Figure 3.1. Pyrimidine-based CG recognition monomers.

The first generation pyrimidine-based nucleosides P^B **24** (and $^{m}P^B$ **25**)[107,108] and ^{4H}T **27**,[105,111-114] demonstrated the importance of the C^2 carbonyl group and N^3 for effective binding to C of the CG base pair *via* a formal hydrogen-bond and a weaker C–H···O interaction.

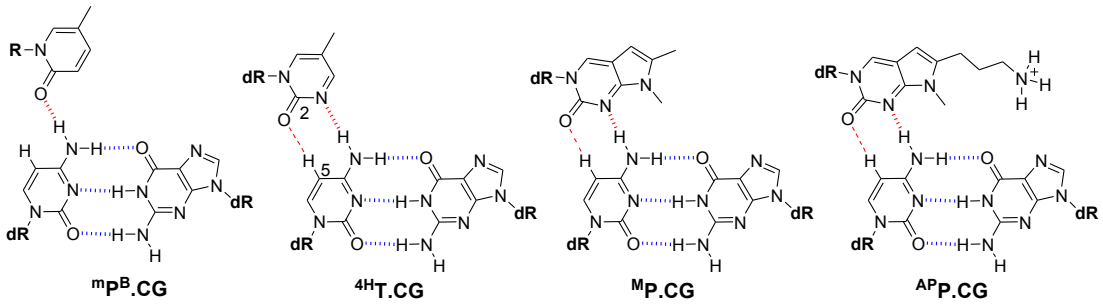


Figure 3.2. Proposed binding motifs for $^{m}P^B$.CG, ^{4H}T.CG, ^{M}P.CG and ^{AP}P.CG triplets.

The second generation nucleosides used the fluorescent bicyclic *N*-methylpyrrolo-dC core of ^{M}P **32**, which retains the C^2 carbonyl and N^3 for binding. An aminoalkyl group was introduced at the 6-position to target C^6=O and N^7 of guanine.[122] Combining the extended aromatic system and amine moiety was proposed to enhance binding affinity

by increased base-stacking and charge-stabilisation due to protonation of the amine group, and also selectivity due to extra hydrogen-bonding to G. A small increase in T_m was obtained for ^{AP}P **34** over T and ^{M}P **32**, against a single CG inversion, indicating the pendant amino group may only contribute stability *via* charge-stabilisation, rather than through additional binding to G.

Modification of ^{AP}P with a 2'-*O*-aminoalkyl-group did not enhance CG binding affinity as expected (through charge-stabilisation), but demonstrated some affinity for AT, presumed to be due to binding of an imino-isomer of the base.[61]

A third generation of monomers was then developed; replacing the flexible alkyl linker with a *meta*-phenylene group, which could increase the potential base-stacking interaction.[206] The new rigid linker may also have entropic benefits, and may place the amino group in the correct position for binding. The new monomer ^{A}PP **143** (aminophenyl-) (Figure 3.1) showed a small increase in T_m over ^{AP}P **34** (aminopropyl-) in UV melting studies at pH 5.8,[206] to give the following order of stability:

$$^{A}PP > ^{AP}P > ^{M}P > T$$

Other modifications were also considered (^{X}PP), and these and alternative methods of synthesis/incorporation of these monomers are discussed below.

3.1.1 Monomer Nomenclature

Furano-dT monomers have been subsequently described by the abbreviation $^{X}PP_{f}$. 'X' describes the aryl functional group (-$C_6H_4NH_2$ – A, -$C_6H_4NHCOCH_3$ – Ac, -$C_6H_4NHCONH_2$ – U) and the subscript 'f', that these monomers were incorporated into TFOs as the furano-dT analogue and post-synthetically modified. The 6-methyl-furano-dT derivative **144** also carries the notation, $^{M}P_{f}$ (Figure 3.4).

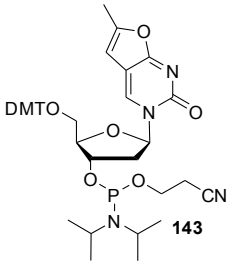


Figure 3.3. Commercially available 6-methyl-furano-dT ($^{M}P_f$) phosphoramidite monomer **144**.

The subscript 'p' denotes *N*-methylpyrrolo-dC monomers ($^{X}PP_p$), which require no further modification after incorporation into TFOs. In addition to the three aryl moieties described by X above, it also describes the guanidinylphenyl moiety (-$C_6H_4NHC(N)NH_2$ – G). The synthesis and biophysical studies of these are discussed in Chapters 5 and 6.

3.2 Synthesis of Furano-dT Monomers for CG Recognition

The first synthetic targets were modified furano-dT nucleosides. In previous work by Ranasinghe *et al.*,[122] the modification was synthesised as the furano-dT phosphoramidite monomer.

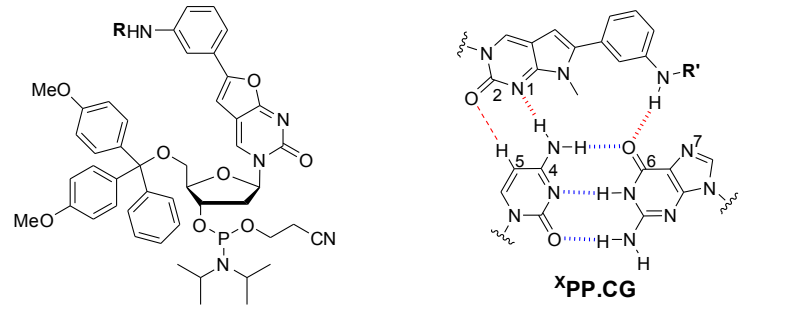
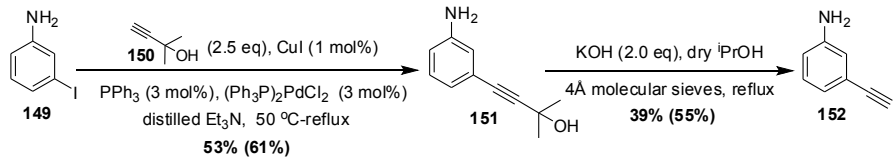


Figure 3.4. Furano-dT monomer structure (**145-147**) and ^{X}PP.CG triplet.

Following incorporation into oligonucleotides, the nucleotide is post-synthetically modified during deprotection and cleavage from the solid support, as described in Section 3.3.

3.2.1 Synthesis of 6-(3-Trifluoroacetamidophenyl)-furano-dT monomer **145**[206,207]

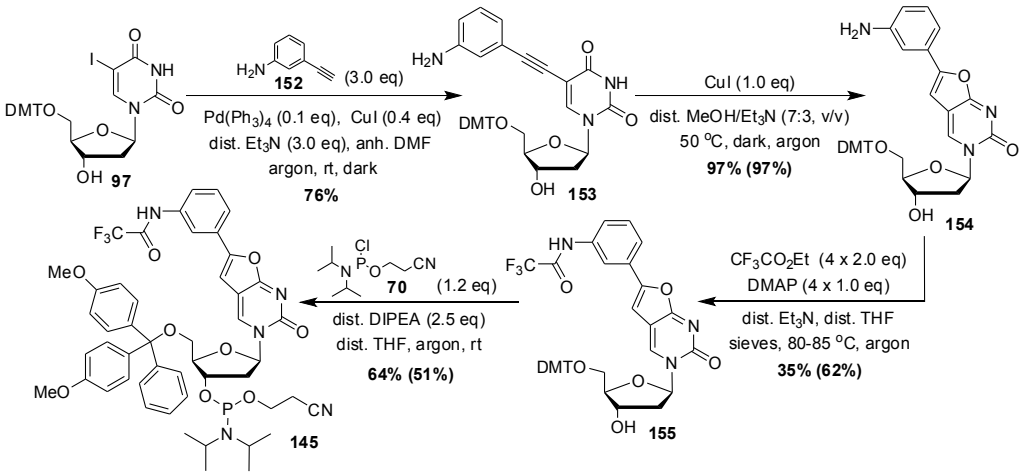
The first aim was to repeat the synthesis of the *meta*-trifluoroacetamido-furano-dT monomer **145**, first synthesised, in house, by Sunil Vadhia,[206] from 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine (DMT-IdU) **97** (DMT-on method) The key steps in the synthetic route to all the furano-dT monomers, are a Sonogashira cross-coupling of an arylalkyne with DMT-IdU **97** or the 'free' nucleoside, 5-iodo-2'-deoxyuridine **148**, followed by a copper (I)-catalysed cyclisation reaction to form the furanopyrimidin-2-one nucleobase core.



Scheme 3.1. Synthesis of 3-ethynylaniline/3-aminophenyl acetylene **152**, (*lit.*[206] yields in parentheses).

The arylalkyne, 3-ethynylaniline (3-aminophenyl acetylene, 3-EA) **152** was synthesised (Scheme 3.1), by the same route as used by Vadhia,[206] which follows a revised version of a method published by Melissaris *et al.*.[208] 3-Iodoaniline **149** was reacted with 2-methyl-but-3-yn-2-ol **150** *via* a Sonogashira cross-coupling reaction, using the catalyst system, $Pd(PPh_3)_2Cl_2$–PPh_3–CuI (3.0:3.0:1.0 mol %), in triethylamine at 50 °C/reflux. Triethylammonium iodide was removed during a lengthy workup with diethyl ether. Stirring in glacial acetic acid, concentration then addition of water, however, did not induce crystallisation or precipitation as described by Melissaris, and aqueous workup with 2M aq HCl did not remove residual starting material. Upon repeating this reaction triethylammonium iodide was removed by trituration/precipitation using diethyl ether, and the protected alkyne **151** was isolated, following purification, in a moderate 53% yield.

Alkyne **151** was deprotected by refluxing in a dry solution of KOH in isopropanol for 3.5 hours, to afford 3-EA **152** in 39% yield. Following this work, a much cheaper commercial supply of 3-ethynylaniline **152** became available, and this compound was subsequently purchased instead.



Scheme 3.2. Synthetic route to 6-(3-trifluoroacetamidophenyl)-furano-dT phosphoramidite monomer **145**, (*lit.*[206] yield in parentheses).

Sonogashira cross-coupling of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **96** (Section 3.2.2) with 3-ethynylaniline **152** was carried out under Hobbs' conditions[124] for alkynylation of nucleosides (Scheme 3.2). Larger quantities of CuI (0.4 eq) and alkyne (3.0 eq) were necessary for successful reaction. Trituration with diethyl ether, to remove unreacted alkyne, and column chromatography afforded the Sonogashira product **153** in 76% yield. Vadhia conducted this cross-coupling using the trifluoroacetyl-protected alkyne followed by deacetylation to give the same product **153**. The NH_2 group was subsequently discovered not to affect this Sonogashira reaction and therefore did not need prior protection, although a slightly higher yield was obtained when protected.[206]

Cyclisation was effected by heating at 50 °C overnight with CuI (1.0 eq) in methanolic triethylamine in absence of light.[121] The reaction could not be monitored by TLC, as the starting material **153** and product **154** had the same R_f, and on isolation, the product contained 3.5 mol % uncyclised material, as determined by ^{1}H-NMR analysis.

The aniline moiety of **154** was protected with a trifluoroacetyl group, by refluxing with ethyl trifluoroacetate (4 × 2.0 eq) and DMAP (4 × 1.0 eq) in THF with Et_3N for greater than 2 days. Despite adding triethylamine to protect against detritylation, the necessary prolonged heating lead to general degradation and a 35% yield. This demonstrates the low reactivity of this aniline and instability to prolonged heating. Vadhia achieved a moderate yield of 62% using a larger excess of ethyl trifluoroacetate. The alkyne amine could have been trifluoroacetylated using trifluoroacetic anhydride, before Sonogashira

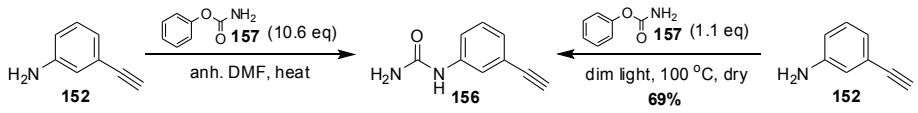
cross-coupling, thus avoiding selectivity/reactivity issues later on. However, it is doubtful that this protecting group would survive reflux in methanol/triethylamine, hence necessitating protection after cyclisation.

Finally, the protected, cyclised nucleoside **155** was phosphitylated at the 3'-position using 2-*O*-cyanoethyl-*N*,*N*-diisopropyl chlorophosphoramidite **70** (1.2 eq) with DIPEA in THF,[144] strictly under an argon atmosphere and anhydrous conditions, and using argon-degassed or distilled solvents at every stage of preparation thereafter (standard procedure for phosphitylation reactions). The air- and acid-sensitive phosphoramidite monomer **155** was obtained in 64% yield after separation of a small amount of uncyclised phosphitylated material carried through from the cyclisation reaction.

3.2.2 Synthesis of 6-(3-Ureidophenyl)-furano-dT Monomer **146**[207]

The anilino-NH_2 of ^{A}PP is only able to form one hydrogen-bond to G (*via* C^6=O). Formation of two hydrogen-bonds with G would not only increase binding affinity, but also selectivity for the CG base pair *via* this extended Hoogsteen binding. A primary urea moiety was proposed as a suitable candidate for this binding motif. Urea moieties have also been used to target G in NMR binding studies of isolated X.CG triplets, using various heterocyclic monomers (Section 1.2.2.4).[133-137] Vadhia had previously attempted the synthesis of the 3-ureidophenyl monomer **146** (Figure 3.3), using the DMT-off synthetic route, where the nucleoside is tritylated after Sonogashira cross-coupling and cyclisation, which proved difficult.

The arylalkyne for synthesis of this monomer, 3-ureidophenyl acetylene **156** was originally synthesised by Vadhia, by heating 3-ethynylaniline **152** with phenyl carbamate **157** (10.6 eq) in DMF.[206] The large quantity of phenyl carbamate used made purification difficult, hence the alkyne/carbamate mixture was used directly in Sonogashira cross-coupling with the free nucleoside 5-iodo-2'-deoxyuridine **148** followed by cyclisation. The phenyl carbamate was removed after cyclisation to give a 43% yield over two steps. Tritylation of the free nucleoside, however, proved problematic affording only a 12% yield.[206] This poor yield could be rationalised by the very high polarity/low solubility of the free, cyclised ureidophenyl nucleoside in the reaction solvent (pyridine).



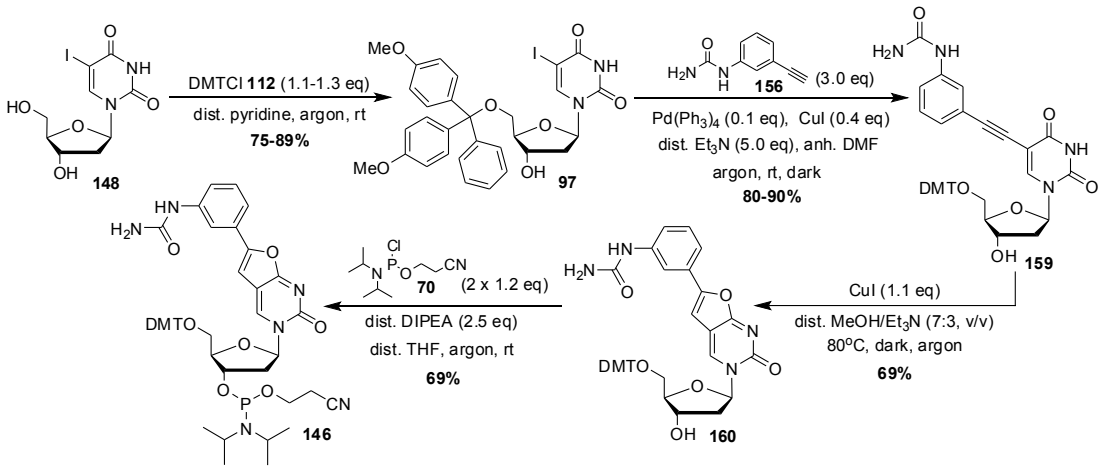
Scheme 3.3. Syntheses of 3-ureidophenyl acetylene **156**.

Conducting the carbamidation in a sealed tube, excluding moisture, without solvent, allowed for use of only a slight excess of phenyl carbamate, and the pure product **156** could be isolated by column chromatography (Scheme 3.3). The reaction was carried out several times at 80-100 °C, using no more than 1.1 eq of phenyl carbamate to ensure adequate mixing of components, however, the yield was limited to a maximum of ~70% due to oxidation of the starting material during the reaction.

Alternatively, reaction with benzoyl isocyanate and deprotection may have afforded the product **156** in higher yield (Section 5.2.5). Using the benzoyl-protected urea in the Sonogashira cross-coupling would be a further improvement. The benzoyl group would not only reduce the polarity of the subsequent products thus aiding purification, but would also block phosphitylation of the urea thereby increasing the potential yield for this reaction. Additionally, the benzoyl group can be removed during oligonucleotide deprotection using ammonia, as was demonstrated, in house, by Dr. Edrees-Abdou in synthesis of the 3-thioureidophenyl-*N*-methylpyrrolo-dC analogue **158** (Section 5.2.5).

Prior to Sonogashira cross-coupling, a large quantity of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine[209] **97** was synthesised by reaction of a slight excess of 4,4'-dimethoxytrityl chloride (DMTCl) **112** with 5-iodo-2'-deoxyuridine **148** (25 g) in 76% yield (Scheme 3.4). Yields varied from 95-75% on a 5-25 g scale, and some *bis*-tritylated compound was usually formed as the reaction is not fully regioselective. The alkyne **156** was coupled to **97** (DMT-on method) in 80% yield, an 89% yield being obtained on a smaller scale, when triethylammonium iodide could be removed by soaking in acetonitrile.

Initially, the product was cyclised by heating at 60 °C overnight. Only 50% conversion was observed, but extra CuI and raising the temperature to 80 °C pushed the reaction to completion, affording only a 39% yield due to significant degradation. A higher yield may also have been attained using column chromatography instead of recrystallisation. Heating at 80 °C under argon pressure for only 2-3½ hours afforded a 69% yield.



Scheme 3.4. Synthetic route to 6-(3-ureidophenyl)-furano-dT phosphoramidite monomer **146**.

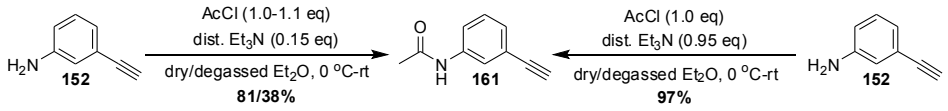
Phosphitylation was pushed to ~90% completion using 2 × 1.2 eq of the chloro-phosphitylating agent **70**. Some of the phosphoramidite co-eluted with the starting material thus reducing the yield slightly further, and some material was lost during precipitation. The phosphoramidite was precipitated after chromatography from CH_2Cl_2 into hexane, to remove remaining hydrolysed reagent (difficult to visualise by TLC), and triethylamine, which would adversely affect subsequent DNA synthesis. A minor nucleoside-based impurity was also detected after reaction, which was removed during purification, presumed later to be due to further phosphitylation of the primary urea in presence of excess reagent. Product **146** was finally afforded in 69% yield, containing negligible impurities.

3.2.3 Synthesis of 6-(3-Acetamidophenyl)-furano-dT Monomer **147**[207]

The third target, a 3-acetamidophenyl-modified furano-dT phosphoramidite **147** was proposed as an alternative to the urea. Although the acetamide can form only one hydrogen-bond with G (*c.f.* ^{A}PP), the amide NH is more acidic than the aniline NH, and may potentially form a stronger H-bond thereby increasing binding affinity.

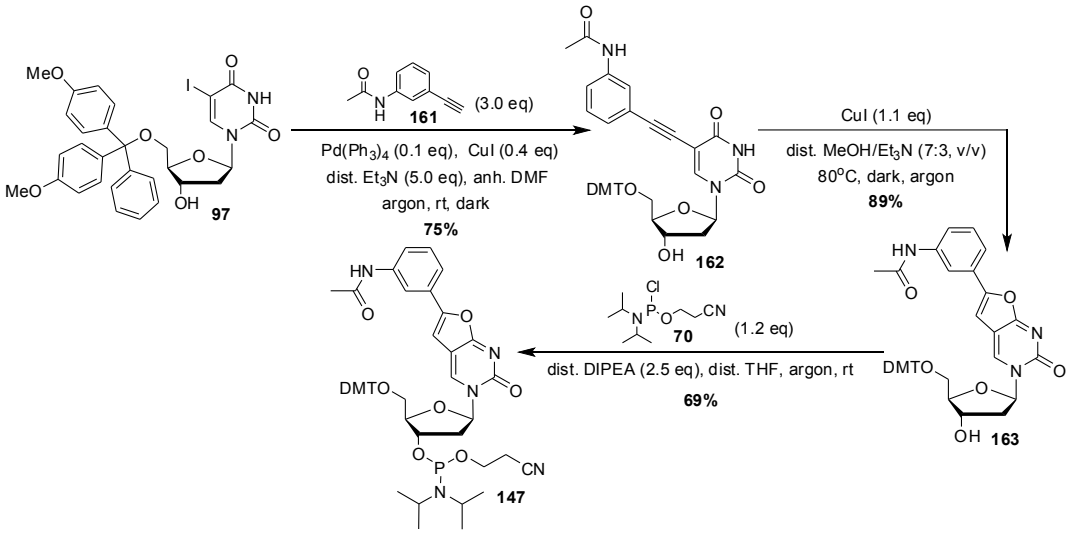
The arylalkyne for this synthesis, 3-acetamidophenyl acetylene **161**, was synthesised in 94% yield from 3-ethynylaniline **152** following a similar method to that published by Kabalka *et al.*[210] for synthesis of *o*-iodoacetanilide. The product was purified by trituration with water as described by Arnold *et al.*,[211] (who prepared **161** using acetic anhydride and acetic acid), followed by aqueous extraction. Using only 0.15 eq Et_3N on

a 1.8 mmol scale, reduced the yield by 11%, and when repeated on an 18 mmol scale, un-neutralised HCl caused degradation to afford a 43% yield, reduced to 38% after recrystallisation (chloroform/hexane) (Scheme 3.5).



Scheme 3.5. Synthesis of 3-acetamidophenyl acetylene **161**.

Sonogashira cross-coupling proceeded to completion using alkyne **161**, but the product **162** proved difficult to separate from a close-running unknown impurity, affording a final yield of 75%. Cyclisation and phosphitylation under the same conditions as previously described afforded the cyclised nucleoside **163** then phosphoramidite monomer **147** in 89% and 69% yields respectively. The phosphitylation was not pushed to completion with extra reagent beyond ~80%, as the hydrolysed reagent can prove difficult to fully remove after purification (Scheme 3.6).

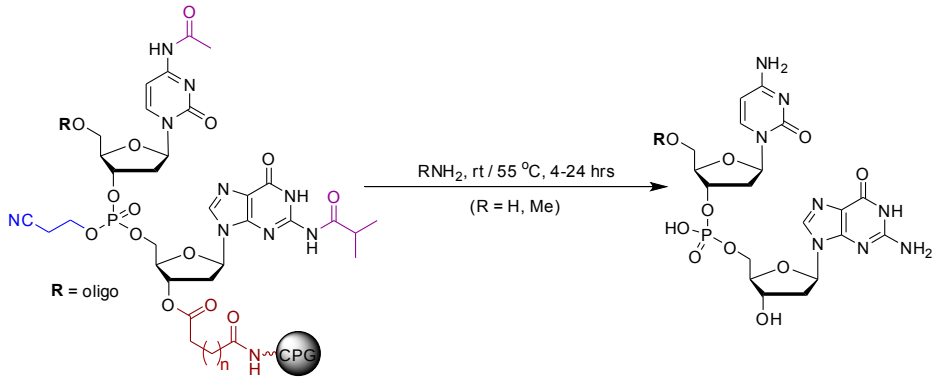


Scheme 3.6. Synthetic route to 6-(3-acetamidophenyl)-furano-dT phosphoramidite monomer **147**.

3.3 Post-Synthetic Conversion of Furano-dT to *N*-Methylpyrrolo-dC

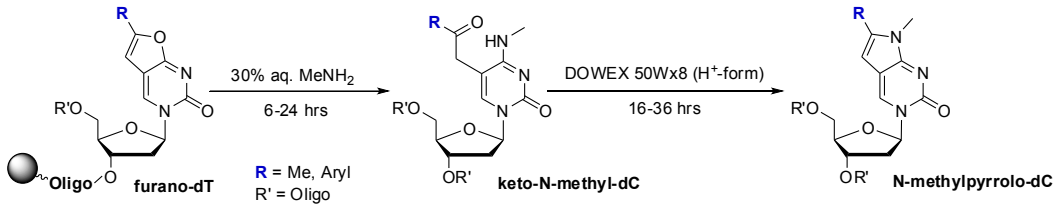
The three furano-dT phosphoramidite monomers **145-147** were incorporated into triplex-forming oligonucleotides (TFOs) for triplex melting studies, by the automated DNA synthesis method on solid-support (CPG), as detailed in Section 9.2.

Following DNA synthesis, the oligonucleotides were deprotected and cleaved from the solid-support using 30 wt % aqueous methylamine at room temperature for 6-24 hours (Scheme 3.7).



Scheme 3.7. General deprotection/cleavage of oligonucleotides highlighting key groups.

Methylamine, in addition to deprotection and resin-cleavage, ring-opens the furano-dT nucleobase to form a 5-aroylmethyl-*N*-methyl-dC derivative. The methylamine insertion reaction occurred quantitatively in 6 hours, although some TFOs were allowed to deprotect overnight or longer to ensure complete reaction (Scheme 3.8).



Scheme 3.8. Post-synthetic conversion of furano-dT to *N*-methylpyrrolo-dC in oligonucleotide.

After removal and washing of the resin, the remaining aqueous solution was concentrated *in vacuo* at 50-60 °C and the residue was re-dissolved in distilled water (1 mL). The oligonucleotide was then treated with DOWEX 50WX8 sulfonic acid resin (200/400 mesh, H^+-form), at room temperature for 18-36 hours, to recyclise the keto-*N*-methyl-dC nucleotide to form the desired *N*-methylpyrrolo-dC derivative (Scheme 3.8).

Initially, oligonucleotides were HPLC-purified before acid-catalysed (DOWEX 50) cyclisation and analysed by MALDI-TOF to check extent of reaction (acidic matrix for MALDI can cause limited additional cyclisation). Once the methylamine insertion reaction was determined to be quantitative, oligonucleotides were treated with DOWEX resin straightaway after deprotection/insertion as described above.

After treatment with DOWEX, the resin was removed by filtration, washed with water and the oligonucleotide solution was concentrated *in vacuo* and desalted (Sephadex NAP™ column). The post-synthetically modified oligonucleotides were analysed by analytical HPLC, capillary gel electrophoresis (CE) (Section 7.3.1), MALDI-TOF and electrospray MS (negative mode). Unfortunately, the acid-catalysed re-cyclisation was not quantitative, only proceeding to 50-80% completion.

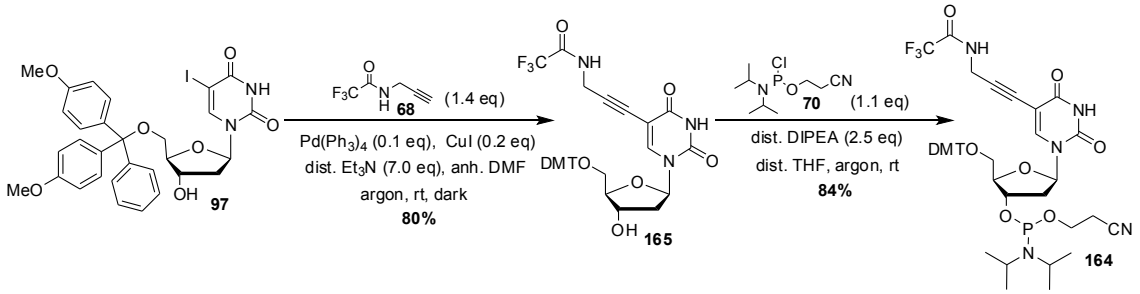
The desired cyclised TFO could be isolated by HPLC for oligonucleotides to be used in UV melting studies (Section 4.1). However, HPLC could not resolve TFOs used for fluorescence melting studies. These TFOs are labelled with a large, lipophilic quencher molecule (DABCYL), which dominates in the HPLC chromatogram thus reducing resolution. The two components were resolved by CE, but cannot be purified by this method hence were used in fluorescence melting studies as a mixture. The uncyclised/ring-opened modification, as a C analogue, could potentially bind to GC not CG as desired.

The standard method for deprotection/cleavage of oligonucleotides is treatment with conc. aqueous ammonia at room temperature or 55 °C for several hours. If these TFOs were treated in this way, ammonia would ring-open the furano-dT modification. The ammonia-insertion product spontaneously recyclises to form the NH-pyrrolo-dC derivative, which as a C analogue, could bind preferentially with GC. The *N*-methyl group, however, prevents the methylamine ring-opened nucleotide from cyclising spontaneously, necessitating acid catalysis. This is presumably due in part to steric bulk of the methyl group.

The post-synthetic conversion of these modifications, although not ideal due to incomplete cyclisation, afforded TFOs that were successfully used in UV and fluorescence melting studies. The results from these are discussed in Chapter 4. Further discussion of oligonucleotide deprotection conditions are given in the study, detailed in Chapter 7.

3.4 Synthesis of 5-(3-aminoprop-1-ynyl)-dU (pdU) monomer 164[39,117]

Prior to melting studies, approximately 5 g of 5-(3-aminoprop-1-ynyl)-dU (pdU) phosphoramidite monomer **164** was synthesised. 5'-*O*-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine **97** and *N*-trifluoroacetyl propargylamine **68** were coupled under Sonogashira Pd-catalysed cross-coupling conditions in two 10 g scale reactions, to afford, after combined purification, the alkynyl-nucleoside product **165** in 80% yield (16.6 g). Following phosphitylation of 5 g of **165** using the chloro-phosphitylating reagent **70**, under standard conditions, the phosphoramidite monomer **164** was afforded in 84% yield (43 bottles for oligonucleotide synthesis).



Scheme 3.9. Synthesis of 5-(3-aminoprop-1-ynyl)-dU (pdU) phosphoramidite monomer **164**.

This phosphoramidite monomer **164** was used for synthesis of TFOs for melting studies (Chapter 4).

3.5 Conclusion

Three furano-dT phosphoramidite monomers **145-147** were synthesised for assessment as monomers for recognition of the CG base pair within a DNA triple helix motif. The common synthetic route (Sonogashira cross-coupling, Cu(I)-catalysed cyclisation and phosphitylation) was successful for each monomer, affording moderate to very good yields for almost every step.

Two main issues were encountered during synthesis. The high polarity of the substrates, once cyclised, introduced problems in purification, even when tritylated. Removal of close-running impurities and starting material also presented minor difficulties. In addition, the cyclised nucleosides demonstrated significantly reduced

reactivity at the 3'-OH and aniline NH_2. This presented problems in trifluoroacetylation and necessitated forcing the phosphitylation to completion with extra reagent, which is not ideal. However, all monomers were synthesised in high purity and incorporated into TFOs with high coupling efficiency (as determined by trityl cation conductivity monitoring – see Section 9.2.1) for triplex melting studies (Chapter 4).

Chapter 4

Biophysical Studies of Furano-dT Modified TFOs for CG Recognition

4. Biophysical Studies of Furano-dT Modified TFOs for CG Recognition

The three modified furano-dT phosphoramidite monomers ($^{A}PP_f$ **145**, $^{Ac}PP_f$ **146**, $^{U}PP_f$ **147**) and commercially available 6-methyl-furano-dT $^{M}P_f$ monomer **144** were incorporated into unlabelled/labelled TFOs for UV and fluorescence melting studies. These studies were designed to test binding affinity for CG and selectivity against the other base pairs (GC, AT, TA) at several pHs. These properties are assessed by determining triplex melting temperatures from graphs of, change in UV absorption or fluorescence emission of fluorophore label respectively, on triplex melting, for different sequences, in several buffer conditions. After incorporation, the modification was post-synthetically converted in the oligonucleotide from a furano-dT to an *N*-methylpyrrolo-dC nucleotide. The procedure for this transformation is described in Section 3.3.

4.1 UV Melting Studies

Methyl- (**144**, $^{M}P_f$), trifluoroacetamidophenyl- (**145**, $^{A}PP_f$) and ureidophenyl- (**147**, $^{U}PP_f$) furano-dT phosphoramidite monomers were incorporated into TFOs (OL3-5) to assess binding affinity for a single CG inversion in a homopurine tract, in the parallel motif. A schematic representation of the UV melting triplex, used originally by Leumann *et al.*[111] is shown in Figure 4.1.

```
     5′-TTT TTM TXT MTM TMT
5′-GCT AAA AAG ACA GAG AGA TCG
   CGA TTT TTC TGT CTC TCT AGC-5′
```

Figure 4.1. UV triplex melting experiment. TFO shown in bold. **M** = $^{5\text{-Me}}$dC, **X** = ^{M}P (OL3), ^{A}PP (OL4), ^{U}PP (OL5), purine duplex strand (**C**, OL1), pyrimidine duplex strand (**G**, OL2).

UV melting experiments were performed, (as described in detail in Section 9.2), on a Varian Cary 400 Scan UV-Vis spectrophotometer monitoring at 260 nm. A 5:1 ratio of TFO:duplex (5:1 μM) was used. Experiments were conducted at pH 6.07 and pH 6.40 (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA). A standard melt programme was used (Programmes 1/2: 10-50-80 °C at 0.5 °C/min) with 3 or 4 heat/anneal cycles in total (Section 9.2). Condensation was not a major problem at low temperature, affecting only the first heat/anneal cycle, hence purging the sample

chamber with argon was not necessary. Disodium EDTA is used to scavenge divalent cations (e.g. Mg^{2+}), which cause general triplex stabilisation.

The results (obtained over six experiments) indicated ^{A}PP demonstrates enhanced binding affinity to CG compared to ^{M}P (ΔT_m +4.0 °C/+4.5 °C at pH 6.07/6.40) (Table 4.1). This represents a significant improvement over ^{M}P, which Ranasinghe *et al.*[122] demonstrated to perform better than T, using the same triplex motif, at pH 7.0 in 10 mM sodium cacodylate buffer (with 200 mM NaCl, 0.25 mM spermine) (Table 4.1).

pH	T_m / °C, melt/anneal		
	^{M}P (6 runs)	^{A}PP (4 runs)	^{U}PP (1 run)
6.07	31.6/23.1, (62.7)	35.6/29.3, (63.5)	32.2/23.0, (62.3)
6.40	24.5/n.d., (63.0)	29.0/23.6, (63.7)	25.6/19.3, (62.7)

Table 4.1. UV triplex melting experiment: $^{X}PP_f$ against CG at pH 6.07 and 6.40 (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA) using UV melting programme No. 1 (Section 9.2). Average T_m values given in °C for TFO melt/anneal and for duplex in parentheses (average of melt and anneal T_m). n.d. = not determined (T_m < 17 °C).

The ^{U}PP TFO (OL5), however, demonstrated reduced affinity compared to ^{A}PP, yet performed better than the control ^{M}P, by +0.6 °C and +1.1 °C, at pH 6.07 and pH 6.40 respectively. The MALDI-TOF analysis of OL5 showed the molecular ion to be 42 Da higher than expected, indicating the oligonucleotide may have acetylated during the capping step of oligonucleotide synthesis (found *m*/*z* 4697.2, expected 4656.2 $[M + H]^+$). It was presumed that acetylation had occurred on the primary urea, thereby reducing the T_m due to increased steric bulk directed into the side of the major groove. The urea was expected to perform better than the aniline, as it has the potential to form two hydrogen-bonds with G instead of one; the NHs would also be more acidic lending to formation of stronger hydrogen-bonds.

In further UV melting studies (Section 6.1), in which TFOs were synthesised without capping, the same order of binding affinity and similar ΔT_ms were observed. It is unlikely therefore that the urea was acetylated, although when synthesised without capping, no peak for $[M + 43]^+$ was observed. Acetylation may have occurred on the 4-*N* position of one of the 5-methyl-dC nucleotides, but this cannot be confirmed.

Significant hysteresis was observed for the melting temperature using this melt programme, with ΔT_m values between 5.4 and 9.2 °C. This could be decreased by reducing the rate of heating/cooling. The duplex hysteresis was within experimental error, at less than 0.5 °C.

A graph displaying melting curves and their derivatives for ^{A}PP (OL4) at pH 6.07, is shown in Figure 4.2 below. The hysteresis between melting and annealing curves is clearly visible.

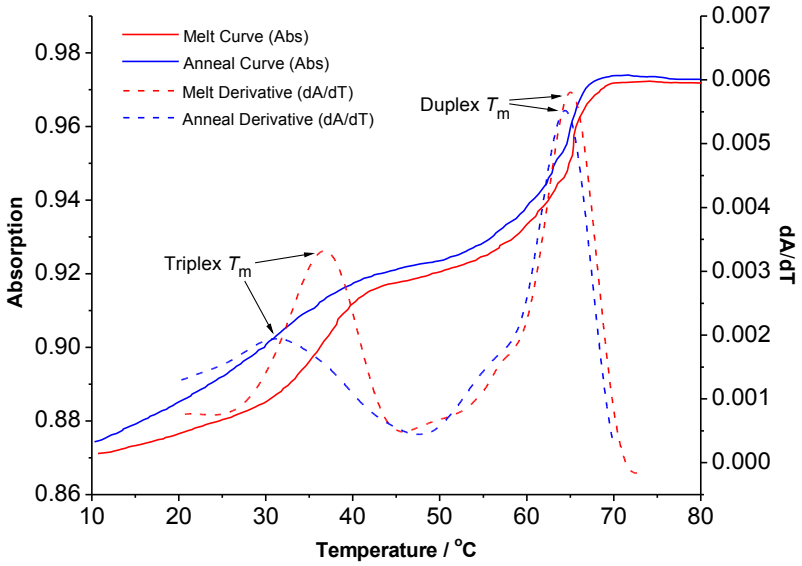


Figure 4.2. UV Melting curves/derivatives for ^{A}PP (OL4) against CG at pH 6.07 (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA). Data is smoothed and un-normalised.

These initial studies indicated that the 3-aminophenyl moiety significantly enhanced binding affinity over a methyl group. The urea moiety, unexpectedly demonstrated poorer binding than the amine. This result necessitated further UV melting and other studies for confirmation and to test other analogues.

4.2. Fluorescence Melting Studies

Fluorescence melting studies were conducted to assess binding affinity and selectivity, and the sequence-dependency of these recognition properties. Fluorescence melting

measures change in fluorescence due separation of fluorophore and fluorescence quencher moieties on duplex and TFO respectively, on triplex melting.

4.2.1 Binding Affinity/Selectivity study

Methyl- (**144**, $^{M}P_{f}$), trifluoroacetamidophenyl- (**145**, $^{A}PP_{f}$), acetamidophenyl (**147**, $^{Ac}PP_{f}$) and ureidophenyl- (**146**, $^{U}PP_{f}$) furano-dT phosphoramidite monomers were incorporated into labelled TFOs for initial fluorescence melting studies. The target was a single CG inversion in a homopurine tract of a hairpin duplex, in the parallel motif. The initial triplex motif is represented below in Figure 4.3.

```
          5′-Q-PMM TPM TXP TPT PTM PT
5′-F-GT GTT AGG AAG AYA AAA AAG AAC TGG T-H )
     CA CAA TCC TTC TZT TTT TTC TTG ACC A-H )
```

Figure 4.3. Fluorescence triplex melting experiment. TFO shown in bold. (**M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **X** = ^{M}P (OL10), ^{A}PP (OL11), ^{Ac}PP (OL12), ^{U}PP (OL13), **Q** = DABCYL (quencher), F = FAM (fluorophore), H = HEG (hexaethylene glycol), hairpin duplex **YZ** = CG (OL6), GC (OL7), AT (OL8), TA (OL9).

Fluorescence melting studies were conducted (as described in detail in Section 9.2), on a Roche LightCycler®. Initial experiments were performed at pH 6.0 (20 mM NaOAc, 200 mM NaCl), and a 10:1 ratio of TFO:duplex (5:0.5 μM) was used. Runs in these experiments were performed up to three times in duplicate, using LC Programme 1 (95-30-95 °C at 0.2 °C/min, see Section 9.2) (Table 4.2). These sequences were used for fluorescence melting only.

There is a stark contrast between some of these results and the previous UV melting study. Whereas ^{A}PP previously demonstrated a significantly higher T_m than ^{M}P against CG (ΔT_m +4.0-4.5 °C), these results place the T_m of ^{A}PP, 1.2 °C lower than ^{M}P. This observation could not be easily explained, although the buffer and sequence had been changed. Despite these observations, the urea (^{U}PP) showed the strongest binding affinity towards CG as originally expected and the order of stability ($^{U}PP > ^{Ac}PP \approx ^{A}PP$), excluding ^{M}P, correlates with the predicted hydrogen-bonding motif.

X / bp	T_m / °C, melt/anneal			
	^{M}P	^{A}PP	^{Ac}PP	^{U}PP
CG	53.1/50.6	51.9/47.9	52.0/47.0	53.9/50.0
GC	40.7/40.3, (67.0)	38.0/36.3, (51.3)	38.8/n.d., (52.5)	39.8/38.1, (53.0)
AT	40.4/38.2	40.3/38.0	40.6/38.0	42.4/40.1
TA	40.1/38.7	42.3/39.7	44.7/40.6	44.6/41.4

Table 4.2. Fluorescence triplex melting experiment: $^{X}PP_f$ against all four base pairs at pH 6.0 (20 mM NaOAc, 200 mM NaCl) using LC Programme 1 (Section 9.2). Average T_m values given in °C for TFO melt/anneal and for secondary melt transition in parentheses. n.d. = not determined (T_m < 36 °C), bp = base pair target.

All monomers demonstrated good selectivity for CG over GC, AT and TA. ^{U}PP was most selective against GC (ΔT_m +14.2 °C) and ^{M}P proved most selective against AT and TA (ΔT_m +12.7 °C, +13.0 °C). Selectivity for CG over TA is most difficult to achieve as both present only one formal hydrogen-bond for binding. Therefore, achieving selectivity against TA is an important criterion in choosing a monomer for CG recognition. However, even the poorest selectivity (^{Ac}PP *vs* TA, ΔT_m +7.3 °C) is still satisfactory.

Hysteresis of 0.42-4.97 °C was observed (average 2.66 °C), despite using a low heating/ cooling rate of 0.2 °C/min (total run time - 12.5 hours). Hysteresis could once more have been reduced by reducing the rate further. However, this would necessitate a run time in excess of 20 hours, which might cause problems due to degradation/evaporation of sample. T_ms differed by no more than 1.2 °C.

The $^{A}PP_f$ modification was evaluated previously by Dr. Powers in a different fluorescence melting triplex motif, at lower pH and a higher buffer concentration. It was shown to be selective for CG, although the order of selectivity over the other base pairs was different from that detected here.[61]

A biphasic melting profile was also observed, against GC, but not the other base pairs or indeed in annealing against any base pair. The presence of the weaker, secondary transition can be rationalised by the proportion of monomer (20-40%), which was not re-cyclised during treatment of TFO with DOWEX-H^+ resin. This C-analogue could, as

previously mentioned (Section 3.3) form a triplet with GC, affording a higher T_m than expected, for a triplex that should contain a disfavourable X.GC triplet (Figure 4.4).

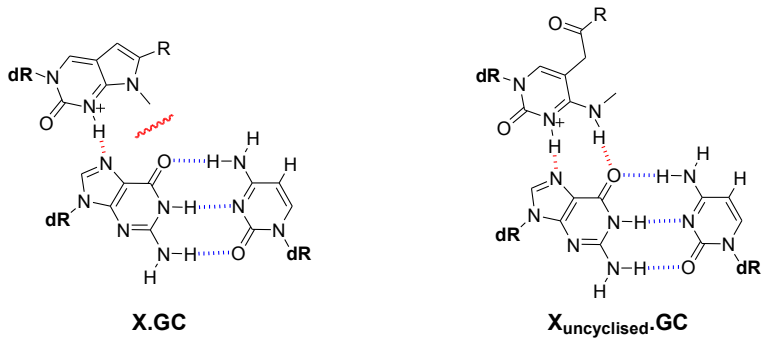


Figure 4.4. Putative X.GC and $X_{uncyclised}$.GC triplets.

The observed secondary T_m value against GC should be similar to that against CG (~52 °C) as was observed for ^{A}PP, ^{Ac}PP and ^{U}PP. However, the value is significantly higher for ^{M}P (67.0 °C), which is difficult to rationalise. Replacing the aryl group for a methyl group should reduce the potential for beneficial interactions within the triple helix, unless the aryl group in fact confers disfavourable steric/hydrophobic interactions. Furthermore, comparing with the many modifications that have been made to the 5-position of C-derivatives for achieving GC recognition (Section 1.2.1.2), it is highly unlikely that the 5-(acetylmethyl)- group of uncyclised/ring-opened $^{M}P_f$ is stabilising enough to account for this T_m. Other secondary structures and interactions may also be contributing factors.

4.2.2 Sequence study

The two possible explanations for reversal in the order of stability of ^{M}P and ^{A}PP between the initial UV and fluorescence melting studies were a change in buffer/pH, and changing the sequence. It was hypothesised, that altering the neighbouring nucleosides to the modification under examination within the TFO may affect binding affinities and selectivity.

New TFOs were therefore synthesised, containing the central trimers, TXT (OL14), PXT (OL15) and PXP (OL16) (X = $^{A}PP_f$, P = pdU) to compare with the TXP-containing TFO (OL11) already investigated. It was thought that presence of pdU next to the ^{X}PP

modification may change binding affinity/selectivity or mask the effects of the modification, and may account for the differences between the two methods.

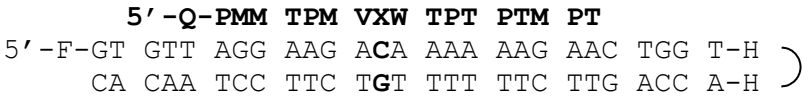


Figure 4.5. Fluorescence triplex melting experiment. TFO shown in bold. (**M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **X** = ^{A}PP & **VXW** = TXT (OL14), PXT (OL15), TXP (OL11), PXP (OL16); **VXW** = TXP & **X** = ^{M}P (OL10), ^{Ac}PP (OL12), ^{U}PP (OL13), **Q** = DABCYL (quencher); F = FAM (fluorophore), H = HEG (hexaethylene glycol), hairpin duplex (**CG**, OL6).

The initial fluorescence melting experiments (Figure 4.5) were repeated using LC Programmes 1 and 2 (95-30-95 °C at 0.2 °C/min) at pH 6.2 (10 mM sodium phosphate, 200 mM NaCl). Experiments were run once or twice in triplicate and a 10:1 ratio of TFO:duplex (5:0.5 μM) was used as before. The aims of these experiments were not only to obtain confirmation of previous results, but to investigate how the change of sequence described above affects binding, and for comparison between analogous furano-dT ($^{X}PP_f$) and *N*-methylpyrrolo-dC ($^{X}PP_p$) modified TFOs (see Section 6.1 for results from $^{X}PP_p$-modified TFOs).

X Seq.	T_m / °C, melt/anneal			
	^{M}P	^{A}PP	^{Ac}PP	^{U}PP
TXT	–	43.2/39.1	*	*
PXT	–	49.3/43.1	*	*
TXP	49.7/43.7	49.9/43.0	48.4/43.0	51.0/45.0
PXP	–	53.6/49.0	*	*

Table 4.3. Fluorescence triplex melting experiment: $^{X}PP_f$ against CG in four TFOs (OL14, OL15, OL11, OL16) at pH 6.2 (10 mM sodium phosphate, 200 mM NaCl) using LC Programme 1 and 2 (Section 9.2). Average T_m values given in °C for TFO melt/anneal. * Data obtained for $^{X}PP_p$-modified TFOs is not shown (Section 6.2).

Hysteresis was much greater than previously observed (4.08-6.91 °C, average 5.61 °C) and annealing data was of poorer quality than in previous experiments. However, the melting data was clear and T_m values differed by no more than 1.0 °C.

These experiments confirm that the urea (^{U}PP) has the highest binding affinity in the TXP-containing TFO. ^{M}P and ^{A}PP appear as effective as each other and slightly better than the acetamide (^{Ac}PP), although experimental error may account for variation as the difference in T_m values is small.

When T is replaced by pdU, the T_m increases because pdU forms additional charge-stabilising interactions with the sugar-phosphate backbone (Section 1.2.1.1). In the ^{A}PP-modified series of TFOs, as for any modification, for each substitution the T_m would be expected to increase by the same amount, and T_m should be equal for the TXP- and PXT-containing TFOs (OL11, OL15) as only the order is changed. However, the T_m for TXP (OL11) is 0.6 °C higher than PXT (OL15). Also the increase in T_m from TXT to TXP is greater than for PXT to PXP.

These observations indicate that placing pdU on the 5'-side of ^{X}PP is slightly destabilising compared to the 3'-side. It is reasonable therefore, to suggest there may be interaction between the 3-aminoprop-1-ynyl group of pdU and the anilino group of ^{A}PP, slightly affecting recognition of CG. This difference was most significant for ^{U}PP ($^{X}PP_p$-modified TFO, see Section 6.2) with a ΔT_m value of 3.0 °C. It is also reasonable to suggest that due to this destabilisation, replacing the second T with pdU has less stabilising effect than expected.

Finally, data for ^{Ac}PP and ^{U}PP confirmed that the T_ms for furano-dT and *N*-methylpyrrolo-dC modified TFOs were approximately the same, only differing by 0.7 and 0.1 °C respectively. These results led into a large study, using *N*-methylpyrrolo-dC monomers, the results of which are discussed in Chapter 6.

4.3 Fluorescence Melting Mismatch Experiment

Experiments were conducted to test if triplex-mediated recognition of a base pair mismatch can occur using the $^{X}PP_f$ modifications tested above. This may also provide information on binding motif of the X.CG triplet. All four modifications were assessed against a hairpin duplex containing a CC mismatch. Fluorescence melting was conducted using LC Programme 4 (95-30-95 °C at 0.25 and 0.2 °C/min) at pH 6.0 (20

mM NaOAc, 200 mM NaCl). Experiments were run once in duplicate and a 10:1 ratio of TFO:duplex (5:0.5 μM) was used as before (Table 4.4).

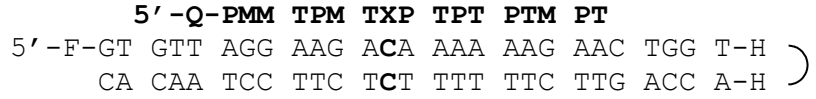


Figure 4.6. Fluorescence triplex melting mismatch experiment. TFO shown in bold. (**M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **X** = ^{M}P (OL10), ^{A}PP (OL11), ^{Ac}PP (OL12), ^{U}PP (OL13), **Q** = DABCYL (quencher), F = FAM (fluorophore), H = HEG (hexaethylene glycol), mismatched hairpin duplex (**CC**, OL17).

These results show a drop in T_m of 4.5 °C for ^{M}P when CG is replaced by a CC mismatch. This is expected as even though ^{M}P is proposed to hydrogen-bond only to one C of the CC base pair as is the case for CG, distortions due to the CC mismatch would cause destabilisation of the triplex as well.

T_m / °C, melt/anneal				
^{M}P.CG	^{M}P.CC	^{A}PP.CC	^{Ac}PP.CC	^{U}PP.CC
53.1/50.8	48.6/47.1	48.9/45.9	49.3/46.2	50.0/48.4

Table 4.4. Fluorescence triplex melting mismatch experiment: $^{X}PP_f$ against CC mismatch hairpin duplex (OL17) at pH 6.0 (20 mM NaOAc, 200 mM NaCl) using LC Programme 4 (Section 9.2). Average T_m values given in °C for TFO melt/anneal. ^{M}P.CG data from this experiment is shown for comparison.

The CC mismatch also reduces the duplex melting temperature by ~8 °C compared to that of the CG hairpin duplex. In addition, the derivative curve is markedly different to the duplexes containing matched base pairs, as it curves towards an apparent melting transition (~74 °C) just below duplex melting. It is possible that other structures may be present, or that this is an unexplained feature of the melting profile of a mismatch-containing hairpin duplex, the latter of which appears more likely.

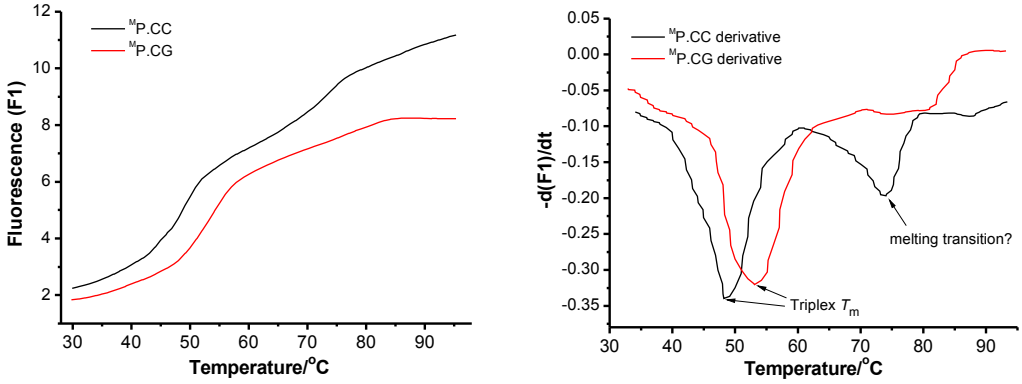


Figure 4.7. Fluorescence melting curves/derivatives for ^{M}P (OL10) against CG (red, OL6) and CC (black, OL17) at pH 6.0 (20 mM NaOAc, 200 mM NaCl). Data is smoothed and un-normalised.

Melting temperatures for all modifications were very similar; the urea performing marginally better than the rest. This is consistent with the proposed binding model, in which differentiation between modifications is achieved through additional hydrogen-bonds to G of the CG base pair. Altering the hydrogen-bond pattern, in conjunction with triplex distortion due to the mismatch, reduces these differences.

4.4 Conclusion

All melting studies demonstrated the effectiveness of the ^{X}PP modifications at recognising CG with similar or better binding affinity/selectivity than ^{M}P, which itself is a large improvement on T.[122] Although it was unknown why order of stability differed between UV and fluorescence melting studies, this was investigated in further studies (Chapter 6). At this stage it was presumed to be caused by change in sequence (particularly of neighbouring nucleotides), buffer and pH, but this was/could not be confirmed.

The significant effect of neighbouring groups on the modification under investigation, was also identified. Results indicated that the modified nucleoside 5-(3-aminoprop-1-ynyl)-dU **14** (pdU, P), interacts differently with ^{X}PP depending on which side it is located, giving rise to a difference in T_m. This neighbouring group interaction introduces important implications for design of TFOs, particularly for use as

oligotherapeutics. Understanding the interactions between modified nucleosides is vital, so that negative interactions can be minimised.

Melting experiments were also conducted against a CC mismatch in the same hairpin duplex motif. It was demonstrated that differences in T_ms between TFOs is reduced as key hydrogen-bonding residues are removed when G is replaced with C. The reduction in overall duplex and triplex T_m is expected, due to the CC mismatch.

It would also be useful to test the CT and CA mismatches to identify if any differentiation could be made between modifications. Both CT and CA contain potential hydrogen-bonding acceptors on T and A, therefore some differentiation or stabilisation may be observed.

Assessing recognition of common base pair mismatches/mutations could also provide valuable information when designing TFOs for therapeutic use. Depending on location of mismatches and on which strand the mutation is located, it may be possible to bind to both desired and undesired sequences. As binding primarily occurs to one side of the base pair, only mutations to the other base should be tolerated.

Chapter 5

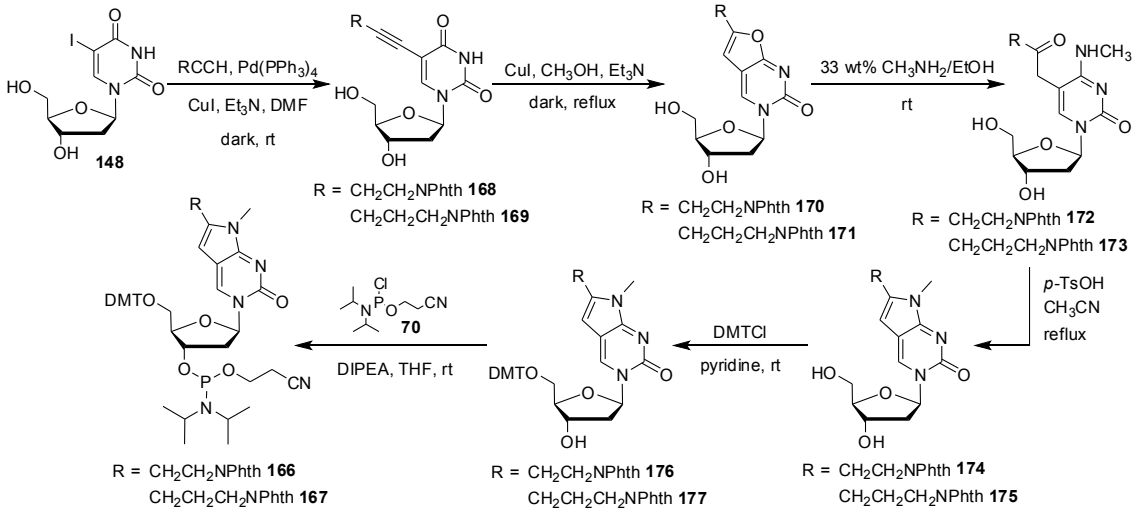
Amine-Modified *N*-Methylpyrrolo-dC Nucleosides for CG Recognition

5. Amine-Modified *N*-Methylpyrrolo-dC Nucleosides for CG Recognition

5.1 Introduction

In previous studies (Sections 3-4), *N*-methylpyrrolo-dC modifications were incorporated into TFOs as the furano-dT analogues then post-synthetically converted. Although the methylamine-insertion reaction was quantitative for each oligonucleotide, the acid-catalysed cyclisation reaction using acidic DOWEX 50WX8 ion-exchange resin, to reform the bicyclic core, did not proceed to completion. The major problem with this method, in addition to reduced yield of TFO, is difficulty in purification when labeled with DABCYL. For this reason, TFOs for fluorescence melting studies were used as the inseparable mixture containing approx. 20-40% uncyclised modification. This did not affect recognition of the CG base pair, although the concentration of correct TFO was lowered as a result. However, in fluorescence melting studies against GC, a second minor melting transition was observed at a similar or substantially higher temperature than against CG. This was proposed to be due to the small proportion of the uncyclised modification, which, as a C-analogue, exhibits GC recognition properties (also Sections 4.2 and 6.2). The very high temperature of one of these transitions could not be rationalised, however, unless the ring-opened modification is exceptionally strong at binding, which is very unlikely.

To overcome this problem, it was proposed, in house, to synthesise monomers as the *N*-methylpyrrolo-dC derivatives. Assuming this bicyclic structure was stable to deprotection conditions (i.e. could not be ring-opened by NH_3 or $MeNH_2$, in aqueous solution), no post-synthetic modification would be required. Ranasinghe synthesised, in house, the protected 6-aminoethyl and 6-aminopropyl- monomers **166,167** as the *N*-methylpyrrolo-dC derivatives ($^{AE}P_p$ and $^{AP}P_p$), *via* methylamine insertion and acid-catalysed cyclisation, on the untritylated furano-dT nucleosides (*c.f.* post-synthetic modification of TFO – Section 3.3) (Scheme 5.1).



Scheme 5.1. Synthesis of $^{AE}P_p$ and $^{AP}P_p$ *N*-methylpyrrolo-dC monomers **166,167** *via* furano-dT intermediates (in house).

Vadhia also applied this method for synthesis of the 6-(3-trifluoroacetamidophenyl)-**178** and 6-(3-ureidophenyl)-*N*-methylpyrrolo-dC monomers **179** ($^{A}PP_p$, $^{U}PP_p$). The trifluoroacetamido monomer $^{A}PP_p$, was synthesised with two low yielding steps, but the ureido monomer $^{U}PP_p$ could not be synthesised. Poor substrate solubility was the most likely cause.

Although this route was successful for one monomer, the route needed to be robust; applicable to synthesis of all derivatives required for biophysical studies. The method therefore, went through optimisation to afford a general synthetic route applicable to all desired analogues (Section 5.2).

5.1.1 Synthesis of 6-(3-Ureidophenyl)-*N*-methylpyrrolo-dC monomer **179**

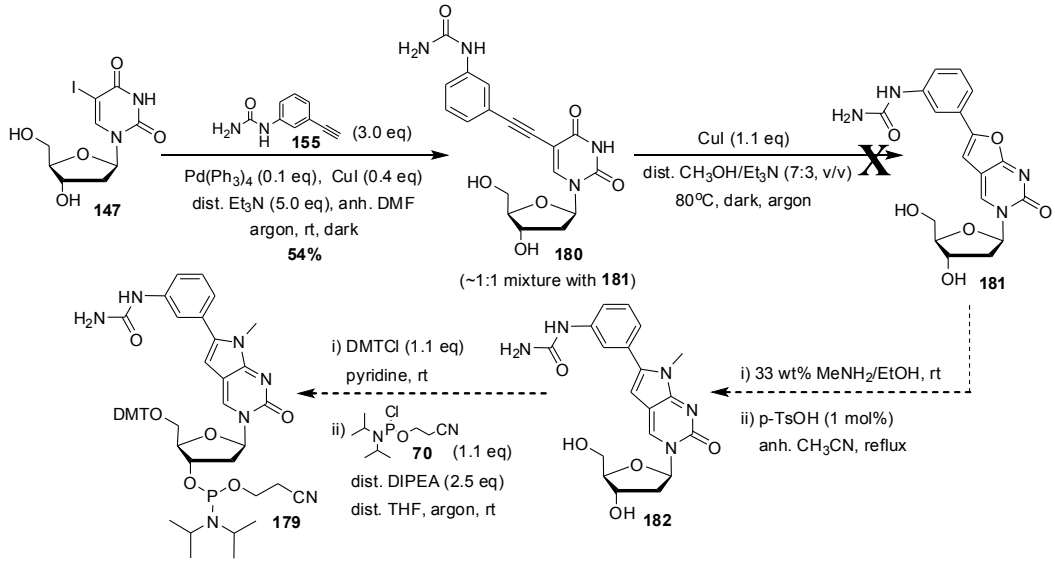
The first target was the ureido-modified monomer **179**, which should exhibit favourable hydrogen-bonding interaction with G of the CG base pair, in addition to the binding motif previously described. The urea moiety increases substrate polarity, which affected purification of the corresponding furano-dT monomer **146** and intermediates, therefore this modification was deemed ideal for testing this synthetic route.

5.1.1.1 Synthesis via untritylated furano-dT derivative

The first method closely followed the synthetic routes used by Ranasinghe[122] and Vadhia.[206] This route proceeds *via* the 'free' (unprotected) furano-dT nucleoside **180**, which was synthesised *via* the Sonogashira and cyclisation reactions discussed in Section 3.2.2.

Following Cu(I)-catalysed cyclisation, Ranasinghe and Vadhia ring-opened the 'free' nucleoside **180** with 33% ethanolic methylamine, before acid-catalysed re-cyclisation. Ranasinghe also, successfully ring-closed a 6-butyl-derivative, by refluxing in acetonitrile with 1 mol% toluene-sulfonic acid, to obtain an 88% yield over these two steps. Vadhia used 80% aqueous acetic acid, at 80 °C to afford the re-cyclised aniline-derivative **143** in 21% yield. Subsequent tritylation proceeded with difficulty in 9% yield. In both methods the acid was difficult to remove, and for the latter case, acetic acid may have also affected column purification and certainly tritylation, although poor to moderate solubility of the product **143** may also have played a role. Vadhia could not re-cyclise the ureido-derivative using acetic acid, most likely due to poor solubility.

In an attempt optimise this route, the above synthesis was first repeated (Scheme 5.2). 5-Iodo-2'-deoxyuridine **148** was coupled three times with 3-ureidophenyl acetylene **156** *via* the previously described Sonogashira reaction, finally yielding a ~1:1 mixture of cyclised **180** and uncyclised nucleoside **181**, containing some alkyne dimer.



Scheme 5.2. First route to $^{U}PP_{p}$ monomer **179**.

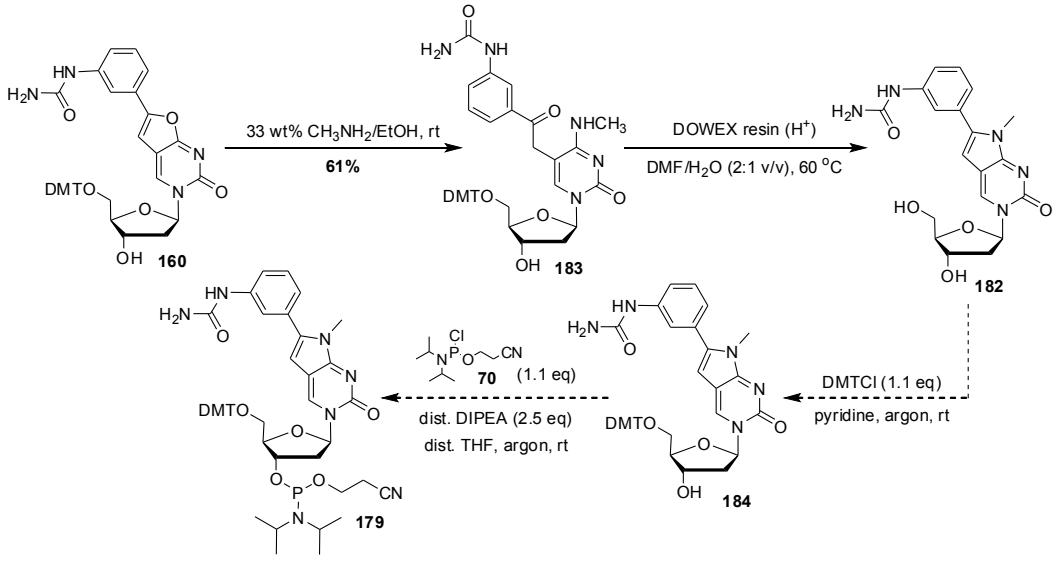
The product mixture **180,181** was refluxed with CuI in dry methanol and triethylamine.[121] However, the mixture was poorly soluble, even at 80 °C under argon pressure, and NMR analysis of the crude material confirmed no change had occurred from the 1:1 mixture.

Addition of DMF solubilised the substrate, but failed to illicit cyclisation, even at 90 °C. Tritylation of the nucleoside mixture was also problematic due to poor solubility in pyridine. Excess DMTCl and some heating were necessary to afford some reaction, but a mixture of *mono*- and *bis*-DMT nucleosides was afforded.

It was noted that leaving the Sonogashira cross-coupling reaction for longer may have resulted in complete cyclisation. However, due to solubility issues, an alternative synthetic route was sought.

5.1.1.2 Synthesis via tritylated furano-dT derivative

In order to overcome problems with solubility, the DMT-protected furano-dT nucleoside **160** was used, synthesised as described in Section 3.2.2.



Scheme 5.3. Second revised route to $^{U}PP_{p}$ monomer **179**.

Alternative sources of acid to *p*-TsOH and acetic acid were sought, to prevent detritylation if possible, but mainly for ease of removal after reaction. Cyclised nucleoside **160** was dissolved in 33 wt% methylamine/ethanol at rt, to afford the ring-

opened product **183** in near quantitative conversion. The crude material was dissolved in anhydrous DMF, divided into two, and agitated with DOWEX-50WX8 resin (pyridinium form) or silica gel/Et_3N, at 30 °C for 1 day, then 50 °C for 1 day and 70 °C for 12 hours. No change was observed using silica gel, and the DOWEX caused some detritylation, despite release of pyridine from the resin. The ring-opened compound **183** was also isolated (61% yield) in a larger-scale reaction.

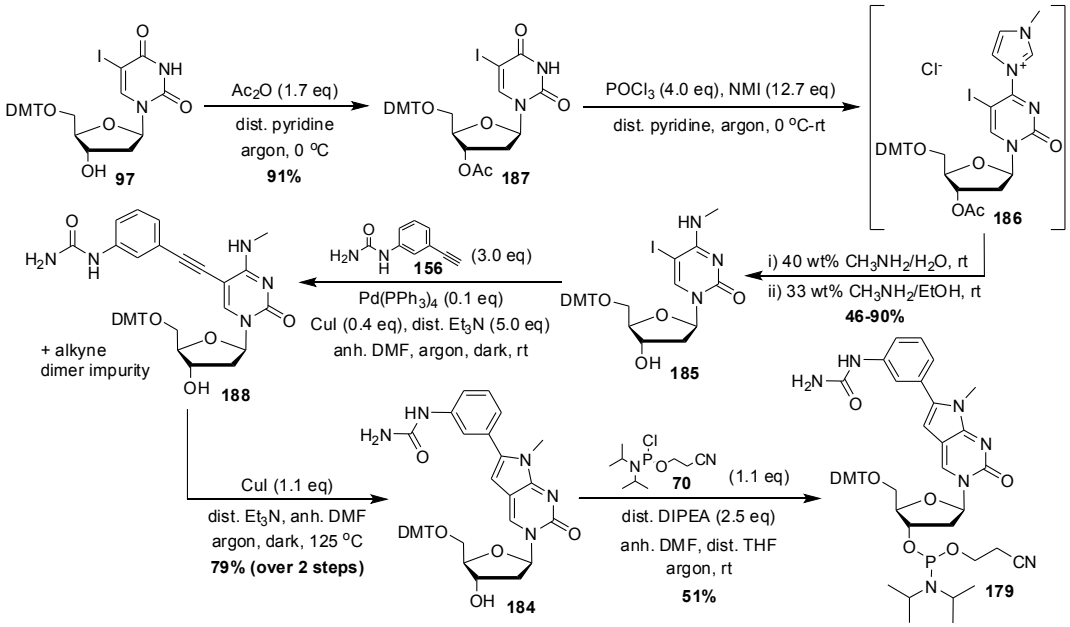
In order to avoid detritylation by exposure to acid, heat alone was assessed as a method of cyclisation. After quantitative ring-opening with methylamine, strong heating of the resulting solution caused significant degradation. Refluxing the crude material in acetonitrile and triethylamine at 90 °C under argon pressure overnight afforded an unidentified product, less polar than the starting material, plus some degradation. This was non-fluorescent on TLC and was presumed to be the 4*N*-methyl-imino-dC derivative. Heating the mixture in ethanolic methylamine at 70 °C overnight also afforded ~50% conversion to the same unidentified, and undesired product, in one step.

Successful cyclisation was achieved using DOWEX 50WX8 resin, in the protonated form. Following ring-opening, the residue was heated at 60 °C overnight in DMF/H_2O (2:1 v/v) with DOWEX 50 (H^+-form) affording the cyclised, detritylated nucleoside **182**. Clearly, protonated DOWEX resin is a suitable, easily removable source of acid, however, it was decided this route was not practical for synthesis of this monomer. The high polarity/poor solubility would cause problems for subsequent purification and tritylation, as observed by Vadhia. Removal then replacement of the DMT protecting group is also not atom-efficient.

*5.1.1.3 Synthesis via derivatisation of 5-iodo-4N-methyl-dC **185***

The third revision to the original route involved initial conversion of tritylated 5-iodo-2'-deoxyuridine substrate **97** to the 4*N*-methyl-2'-deoxycytidine derivative **185**. Although 5-iodo-2'-deoxycytidine is commercially available, the 4*N*-methylated analogue is not, but several methods of uridine to cytidine conversion are presented in the literature.

The 4-carbonyl group is first activated and converted into a good leaving group by a variety of reagents, before reaction with a suitable nucleophile (e.g. NH_3, $MeNH_2$). This conversion was conducted using *N*-methylimidazole (NMI) in presence of $POCl_3$, conditions published by Herdewijn *et al.*.[212] The reaction proceeds *via* the 4-*N*-methylimidazoloyl-dC intermediate **186**, which is not isolated from the reaction (Scheme 5.4).



Scheme 5.4. Third revised route to $^{U}PP_p$ monomer **179**.

After acetylation of the 3'-OH group in 91% yield (Scheme 5.4), the protected nucleoside **187** was treated with phosphorous oxychloride and NMI in pyridine to form the yellow intermediate **186** over 30-60 minutes. Treatment with 40 wt% aqueous methylamine for several hours afforded the partially deacetylated 4*N*-methyl-dC nucleoside. Further treatment with 33 wt% methylamine/ethanol, before or after aqueous workup completed the deacetylation.

On a 1.4 g scale, the product **185** was obtained in a 90% yield, after column chromatography using CH_2Cl_2/CH_3OH as eluent. Reaction on a 6.8 g scale afforded the compound in 80% yield after difficult removal of a significant quantity of NMI. This was removed by 4 columns (different eluents) and recrystallisation from toluene. On an 8.3 g scale, NMI was removed with difficulty by 2 columns and successive extraction with diethyl ether, followed by recrystallisation from ethyl acetate (46% yield).

When later resynthesised in house, by Dr. Mastoura Edrees-Abdou, stirring the reaction at rt overnight after addition of 40 wt% aqueous methylamine was sufficient to completely deacetylated, therefore no further treatment with 33 wt% ethanolic methylamine was required. Also, careful column chromatography using methanol/ CH_2Cl_2 as eluent afforded the product in 84% yield.

Commonly, conversion *via* the 4-(1,2,4-triazol-1-yl)- intermediate is used, using a large excess of 1,2,4-triazole in presence of excess $POCl_3$.[213-215] The intermediate was often isolated, although Sanghvi *et al.*[213] reported an increased yield if this intermediate was not isolated, and have conducted this reaction up to kilo scale, in yields of 65-85%.

Other reagent systems used for this conversion include 3-nitro-1,2,4-triazole/diphenyl phosphorochloridate,[216] l-(mesitylene-2-sulfonyl)-3-nitro-1,2,4-1*H*-triazole (MSNT)/ cat. Ph_2PO_4 (forms the 3-nitro-1,2,4-triazolyl- intermediate),[216] 2,4,6-mesitylenesulfonyl chloride/DMAP,[95] 2,4,6-triisopropylbenzenesulfonyl chloride/DMAP,[217,218] triiso-propylsilyl chloride/DMAP,[219] and 2- and 4-nitrophenol/$POCl_3$.[216,220,221]

Sonogashira cross-coupling with 3-ureidophenyl acetylene **156** gave the desired alkynyl product **188**, with a large quantity of the alkyne dimer impurity, and a minor unknown impurity. The alkyne dimer was partially removed by recrystallisation from methanol. Following cross-coupling, the nucleoside/alkyne dimer mixture was treated with copper iodide in dry, refluxing CH_3OH/Et_3N at 80 °C under McGuigan's conditions,[121] but only partial cyclisation was observed over 2 hours.

One of few literature examples of cyclisation of 5-alkynyl-dC derivatives was published by Inoue *et al.*.[222,223] A protected 4*N*-acetyl-5-ethynyl-dC nucleoside was cyclised by heating at 125 °C in DMF for 10 minutes, in presence of CuI (0.55 eq).

The ureido-modified Sonogashira product **188**, containing the alkyne dimer impurity, was successfully cyclised under similar conditions, to give the desired product in 55% then 79% yield, over two steps, finally using a slight excess of CuI (1.1 eq) and longer reaction time (40 minutes). The alkyne dimer impurity from the previous reaction was then easily removed by column chromatography.

The cyclised nucleoside **184** was subsequently phosphitylated to afford the desired $^{U}PP_{p}$ monomer **179** for incorporation into TFOs. During the first small-scale attempt, using the chloro-phosphitylating reagent **70**, a larger quantity was added in error affording 41% yield of the *bis*-phosphoramidite (reaction on primary urea) plus 19% yield of desired phosphoramidite **179**. Once the reactivity of the urea NH_2 was highlighted, the less reactive *bis*-aminophosphitylating reagent **189** (Figure 5.1) was used, in order to minimise secondary reaction. This necessitated 2.6 eq of reagent, 1.0 eq of activator diisopropylamine hydrotetrazolide (DIHT) **190** and very long reaction time. During this time, significant degradation occurred, with formation of a viscous yellow syrup, from which the product was extracted in 29% yield. No *bis*-phosphitylation was observed.

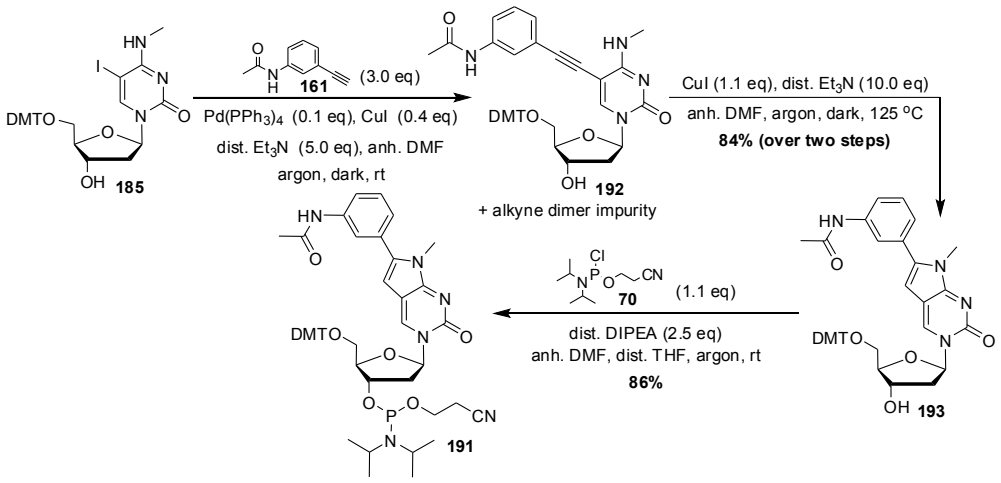
Figure 5.1. *Bis*-amino phosphitylating reagent **189** and activator, DIHT **190**.

DIHT **190** is used in preference to tetrazole due to greater stability and solubility, and lower acidity and hygroscopicity.[224]

Phosphitylation using the chloro-reagent **70** (1.1 eq) was, however, chosen finally, and the product was obtained in 51% yield after column chromatography and precipitation. The reaction was stopped after 2 hrs due to appearance of some *bis*-phosphitylated material. In all phosphitylation reactions, anhydrous DMF was added to solubilise the polar starting material. In total, 6 bottles of monomer (100-120 mg each) were obtained for DNA synthesis.

5.1.2 Synthesis of 6-(3-Acetamidophenyl)-*N*-methylpyrrolo-dC monomer **191**

The acetamide-modified monomer **191** was synthesised by the same route as described above. After Sonogashira cross-coupling of 3-acetamidophenyl acetylene **161** and 5-iodo-4*N*-methyl-dC nucleoside **185**, the resultant product **192**/alkyne-dimer mixture was treated under Inoue's conditions[222,223] and the cyclised acetamido nucleoside **193** was isolated in excellent 84% yield over the two steps (Scheme 5.5).

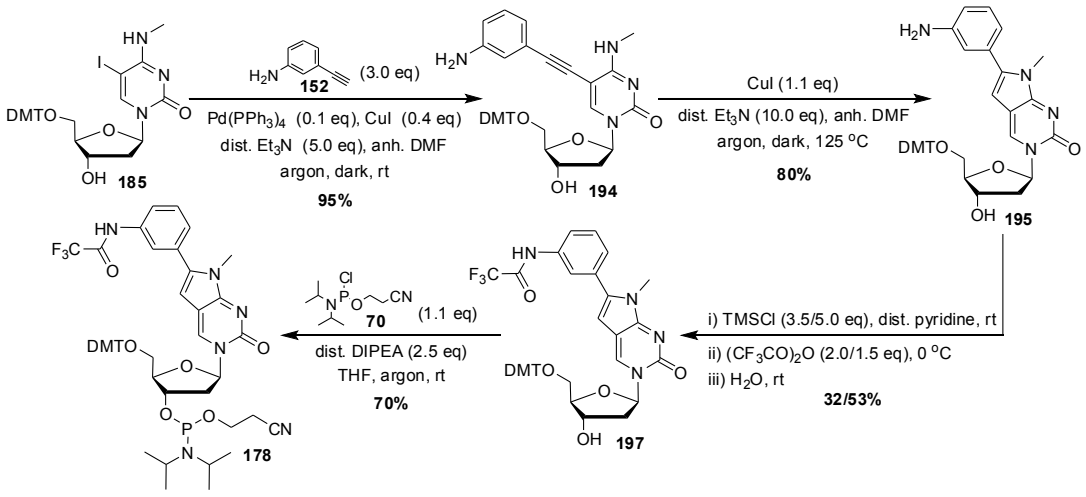


Scheme 5.5. Synthesis of $^{Ac}PP_p$ monomer **191**.

Phosphitylation was carried out using the chloro-reagent **70** in anhydrous THF/DMF to afford the phosphoramidite **191** in excellent 86% yield. This substrate had no competing reaction sites and the reaction proceeded to completion. Despite the high polarity once cyclised, purification proved easier than for the ureido nucleoside **179**. In total, 13 bottles of monomer (70-120 mg each) were obtained for DNA synthesis.

5.1.3 Synthesis of 6-(3-Trifluoroacetamidophenyl)-*N*-methylpyrrolo-dC monomer **178**

Synthesis of the trifluoroacetamido-modified monomer **178** was achieved *via* a slightly altered route. Following Sonogashira coupling of 3-ethynylaniline (3-aminophenyl acetylene) **152** with 5-iodo-4*N*-methyl-dC nucleoside **185** (93-95% yield), cyclisation was carried out under Inoue's conditions.[222,223]



Scheme 5.6. Synthesis of $^{A}PP_p$ monomer **178**.

The first cyclisation attempt afforded only 32% yield of product **195**, plus 12% yield of a *bis*-tritylated product **196** and some untritylated (free) nucleoside **143**. During workup, addition of acetone, methanol or dichloromethane (to precipitate impurities) had caused precipitation of what was thought at that stage to be a mixture of product and copper salts. The mixture was therefore re-dissolved in DMF and filtered, which unfortunately proved unsuccessful in removing significant impurities. After removing the resulting large volume of DMF under high vacuum at elevated temperature (40-50 °C), the dark brown residue was dissolved/suspended in dichloromethane and filtered, successfully removing impurities as originally intended. It was discovered, however, that during DMF removal, some detritylation (through prolonged heating) and retritylation of the product was observed. Retritylation was shown by NMR analysis to have occurred on the less hindered, though less reactive aniline NH_2, not the 3'-OH group. This is possible as the DMT^+ cation is capable of reacting further in absence of water (removed under heat/high vacuum). On repeating this reaction, DCM was used to precipitate impurities, and a yield of 80% was achieved with no DMT migration.

The aniline-NH_2 of **195** was trifluoroacetylated using trifluoroacetic anhydride with transient protection of the 3'-OH with a trimethylsilyl group. This method was used in house, by Dr. Hong Li, to protect the 4-NH_2 group of the *C*-nucleoside and GC recognition monomer, 3-methyl-4-aminopyridine (^{MA}P), during an attempt to optimise the original synthesis.[65] In previously described synthesis (Section 3.2.1), the aniline-NH_2 of the furano-dT analogue **154** was reacted with ethyl trifluoroacetate in only 35% yield with significant degradation due to essential prolonged heating (Section 3.2.1). This alternative method obviated this poor yield but introduced other problems.

During the procedure, the 3'-OH group is protected using trimethylsilyl chloride (TMSCl) in pyridine. Upon completion, trifluoroacetic anhydride is added, followed by water. Addition of water, in presence of pyridine/trifluoroacetic acid/HCl cleaves the TMS group affording the product. On the first attempt, despite adding further equivalents of TMSCl (upto 3.5 eq), some trifluoroacetylation at the 3'-position had occurred; this *bis*-trifluoroacetylated compound **198** proving very difficult to separate by column chromatography. Although the conversion to products was very good, only a 32% yield of the desired product **197** could be obtained pure. Some triethylamine was added after reaction, before removal of solvent, to prevent detritylation.

In the second attempt, in efforts to prevent further reaction, a larger excess of TMSCl (5.0 eq) was added at the start. Although the 3'-protection occurred more quickly than before, after addition of trifluoroacetic anhydride, white fumes were observed and the reaction mixture darkened slightly due to partial detritylation from HCl exposure. Triethylamine was added immediately to prevent further degradation, however, this subsequently prevented complete TMS cleavage by addition of water, as the solution was now too basic. Aqueous workup with 15% aq NH_3 removed excess HCl, affording a 1:1 mixture of TMS-protected product and unprotected nucleoside **197**. Conveniently, the TMS group was cleaved during column chromatography on silica, and the product was afforded in 53% yield. Little *bis*-trifluoroacetylated compound **198** was observed, but the yield was compromised due to detritylation.

Dr. Edrees-Abdou later repeated this reaction, in house, using 20 eq of ethyl trifluoroacetate, 22 eq of Et_3N and 1 eq of DMAP in CH_2Cl_2 at 50 °C overnight. The product **197** was afforded in a maximum yield of 65%, despite some degradation. These conditions again highlight the low reactivity of this aniline due in part, to extended conjugation with the *N*-methylpyrrolopyrimidin-2-one bicycle.

For phosphitylation, the *bis*-amino phosphitylating reagent **189** was tested first. However, poor conversion to products combined with significant degradation on the column led to only a 24% yield of desired monomer **178**. This material was incorporated into an oligoT DNA strand for deprotection tests (Section 7), however, the monomer coupled very poorly (~20%) due to the small quantity obtained (39 mg).

Phosphitylation was most successful using the chloro-phosphitylating reagent **70** in 70% yield. Despite an extended reaction time (4¾ hours), excess reagent (2.6 eq) and prolonged purification time, little degradation was observed and all hydrolysed excess reagent was removed. Two bottles (90 mg each) were obtained for incorporation into TFOs. The monomer contained a significant quantity of $Et_3N.HCl$, which did not affect coupling efficiency during oligonucleotide synthesis. Presence of the free amine would have inhibited coupling (hinders acid-catalysed detritylation step).

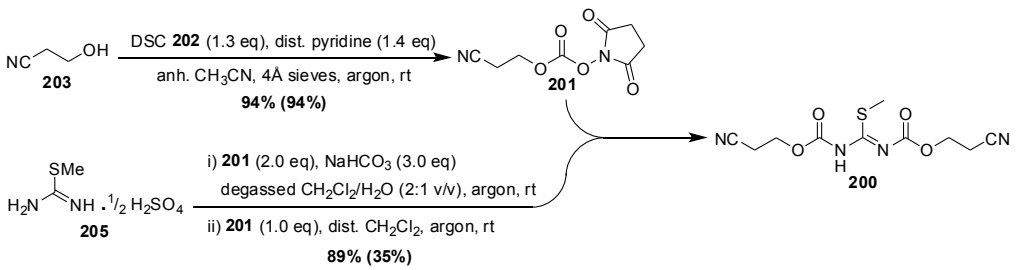
5.1.4 Synthesis of 6-(3-Guanidinylphenyl)-*N*-methylpyrrolo-dC monomer **199**

The guanidinium moiety was expected to exhibit the same binding pattern as the urea, yet contribute additional stability by charge-stabilisation; the guanidine is protonated at physiological pH. The guanidinyl-modified monomer ($^{G}PP_{p}$) proved difficult to synthesise, however, possessing both acid- and base-sensitive groups, necessitating careful manipulation, especially during phosphitylation.

Several methods have been employed for introduction of guanidinyl groups into DNA, both during synthesis[41,225] or post-synthetically on the oligonucleotide.[43,44] Although much easier to introduce post-synthetically, this method is not selective as all exposed primary amines capable of reacting, can react. The main issue with the former approach lies in compatibility with DNA synthesis. Protecting groups must be stable to the acidic (TCA or DCA) and oxidising (iodine) conditions encountered during oligonucleotide synthesis (excludes using BOC and Cbz for example). Using the free-guanidinium moiety significantly reduces the monomer's solubility and therefore the coupling efficiency. The protecting group 2-cyanoethoxycarbonyl- (CEOC) is used as it is stable to oligonucleotide synthesis. However, the groups are base-labile, and removal of just one of these groups from the guanidinyl moiety (most labile on C=N) compromises oligonucleotide synthesis (investigated in house).

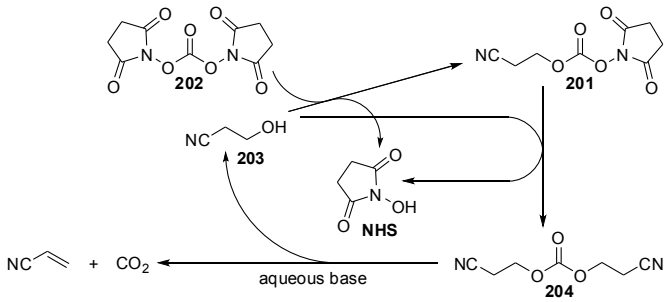
*5.1.4.1 Improved Synthesis of Guanidinylating Reagent **200***

The first step was synthesis of the base-sensitive guanidinylating reagent, *N*,*N'*-*bis*-[(2-cyanoethoxy)carbonyl]-*S*-methyl-isothiourea **200**.



Scheme 5.7. Synthesis of guanidinylating reagent **200**. (*lit.*[42,225,226] yield in parentheses).

The first step in the route to this reagent **200** was synthesis of the NHS-carbonate, (2-cyanoethyl)-*N*-succinimidyl carbonate (CEOC-succinimide) **201**, using di-succinimidyl carbonate (DSC) **202** under anhydrous conditions. The literature protocol[42,226] requires that DSC is added, followed by pyridine, to a solution of 2-cyanoethanol (3-hydroxypropionitrile, 3-HPN) **203** upon which the suspension solubilises and a final yield of 94%, or 87% after further purification (column chromatography) is obtained. In order to reduce reaction time (7 hours), the concentration was increased almost three-fold, however, DSC did not dissolve after 1 hour as expected although the reaction proceeded to completion after 4 hours. After isolation (93%), NMR analysis in d_6-DMSO (much greater solubility than in $CDCl_3$) indicated a large degree of decomposition had occurred. It was hypothesised that the initially high concentration of 3-HPN **203** in the reaction had caused further reaction with the product, due to poor solubility of DSC **202**. The corresponding by-product **204** could decarboxylate in presence of base, producing acrylonitrile and re-forming 3-HPN **203** (Scheme 5.8). This appeared substantiated by significant gas evolution during aqueous workup with aq $NaHCO_3$ (sat. soln:water, 1:1 v/v), and signals for NHS and acrylonitrile in the NMR spectrum. Following later recrystallisation and column chromatography, the compound was isolated in 75% yield, despite fears over loss due to degradation.



Scheme 5.8. Hypothesised mechanism for by-product **204** formation.

The reaction was repeated on a smaller scale at a similar concentration to the literature, and the order of addition was reversed. DSC was stirred in anhydrous CH_3CN for 10 minutes, pyridine was added, followed by 3-HPN **203** dropwise 10 minutes later. The suspension dissolved after 20 minutes and the reaction was complete after 1 hour. Following column chromatography, even after 3 weeks in storage, the pure compound **201** was afforded in 86% yield. MS analysis of both reaction products indicated formation of the correct product with a small quantity of NHS present. NMR analysis in $CDCl_3$ confirmed this also. CEOC-succinimide **201** was discovered to be unstable in

DMSO due in part to presence of water. Also, changing the order of addition of reagents increased the yield by 11% on the previous attempt. On scaling up the reaction to 200 mmol, at twice the concentration in the literature, the reaction was complete within 1.5 hours and a 94% yield was obtained.

The literature protocol[41,225,226] for formation of the guanidinylating reagent is a two-phase reaction (CH_2Cl_2:conc. aq $NaHCO_3$, 1:1 v/v), coupling the organic-soluble NHS-carbonate **201** with water-soluble *S*-methylisothiourea hemisulfate **205**. The literature yield of 35% is poor and not cost-effective on a large-scale due to the price of DSC.

Due to base-sensitivity of the product and in order to increase the yield, the reaction was tested in CH_2Cl_2:H_2O and THF:H_2O (4:1 v/v), with only 6 equivalents of $NaHCO_3$, 2.5 eq of CEOC-succinimide **201**, and using argon-degassed solvents. After workup, TLC indicated a ~3:2 mixture of *mono*- **206** to *bis*-modified **200** isothiourea (Figure 5.1). The sp^2 nitrogen is more difficult to functionalise due to higher protecting group lability under basic conditions. To push the reaction to completion, each reaction was dried, dissolved in CH_2Cl_2 or THF, split into two and treated with 1.5 eq of CEOC-succinimide with or without 1 eq of pyridine. Only reaction in CH_2Cl_2, especially without pyridine, achieved significant further reaction over 26 hours even under non-anhydrous conditions.

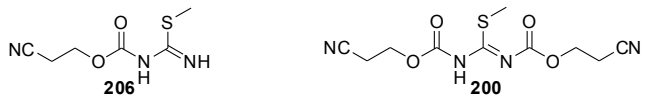
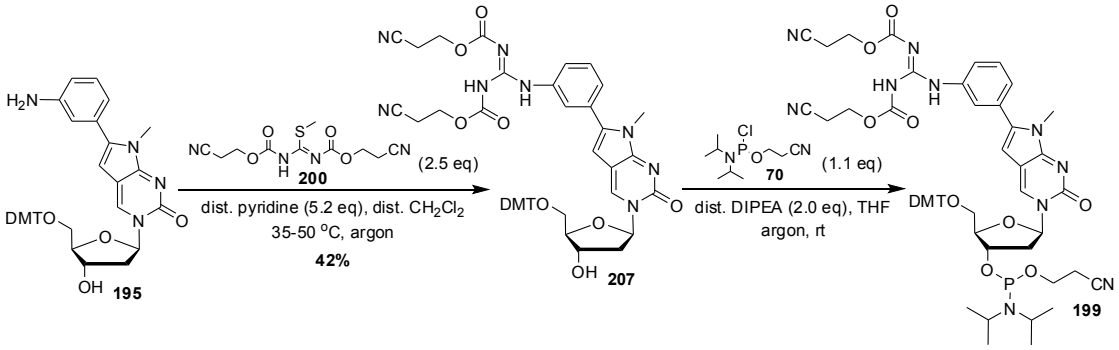


Figure 5.2. *Mono*- **206** and *bis*-modified **200** *S*-methylisothioureas.

Using this knowledge, the reaction was scaled up (87 mmol), at a higher concentration than in the literature under the best conditions from the tests. The two-step procedure worked very efficiently, benefitted by using argon-degassed/distilled solvents, 2 eq then 1 eq of CEOC-succinimide and only 3 eq of $NaHCO_3$ in the first step. A yield of 89% was achieved, a significant improvement on the literature method, and could be optimised further, using less CEOC-succinimide in the second step. A yield greater than 35% appears extremely difficult to achieve without modifying the literature protocol.

*5.1.4.2 Final Steps towards $^{G}PP_p$ Monomer **199***

The anilino-*N*-methylpyrrolo-dC intermediate **195** (synthesised following the synthetic route for the trifluoroacetamidophenyl monomer **178**, see Section 5.1.3) was guanidinylated using guanidinylating reagent **200** under slightly different conditions to the literature, based on experiments in house.



Scheme 5.9. Synthesis of $^{G}PP_p$ monomer **199**.

The reaction was conducted in CH_2Cl_2 with pyridine (instead of DMF with triethylamine) at 35 °C instead of room temperature due to the previously demonstrated low reactivity of the aniline NH_2. Only 1.1 eq of guanidinylating reagent **200** was added initially, using only 0.5 eq of pyridine due to base-sensitivity. Guanidinylation of a primary alkylamine-modified nucleoside, in house, required only 0.5 eq of pyridine. Over more than 1½ days, further reagent and pyridine were added (upto 2.5 eq and 5.2 eq respectively) and the temperature was increased to 50 °C. Although prolonged reaction time and temperature caused some degradation/cleavage of CEOC-groups, these conditions were necessary to ensure adequate reaction occurred. The product was purified by column chromatography, on pyridine-neutralised silica gel, using eluent containing no base, affording a 42% yield. Neutralising the silica gel is important to prevent cleavage of the acid-sensitive DMT group, yet excess base is washed from the column and base is omitted from the eluent to prevent base-mediated cleavage of the CEOC protecting groups from the product **207**.

Phosphitylation was not straight-forward as during storage, some CEOC-cleavage was observed, likely to be caused due to presence of byproduct CH_3SH. The reaction was tested on this material, using 1.1 eq and 1.5 eq of the chloro-phosphitylating reagent **70** and 1.5 eq and 2.0 eq of DIPEA. Literature methods[41,225,226] use the *bis*-amino reagent

189 as DIPEA can cause CEOC-cleavage. However, the low aniline reactivity necessitated these conditions, and less DIPEA was used. The products of both reactions (better conversion for the latter) were combined and purified twice by column chromatography and precipitation. However, due to phosphitylation of the single unprotected guanidinyl nitrogen, the product (**208**) of which eluted similarly to the desired product **199**, the monomer could not be obtained pure. ^{31}P-NMR indicated peaks for both phosphoramidites at ~150 ppm, and signals at ~ 71 ppm corresponding to the *N*-phosphoramidyl group (Figure 5.2).

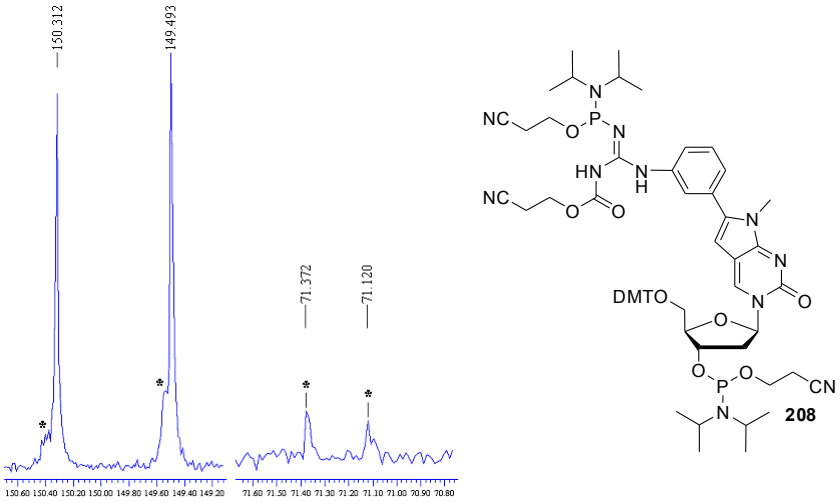


Figure 5.3. Expansions of ^{31}P-NMR of $^{G}PP_p$ phosphoramidite mixture **199,208**, and structure of *bis*-phosphoramidite **208**. *Bis*-phosphoramidyl peaks labelled with an asterisk.

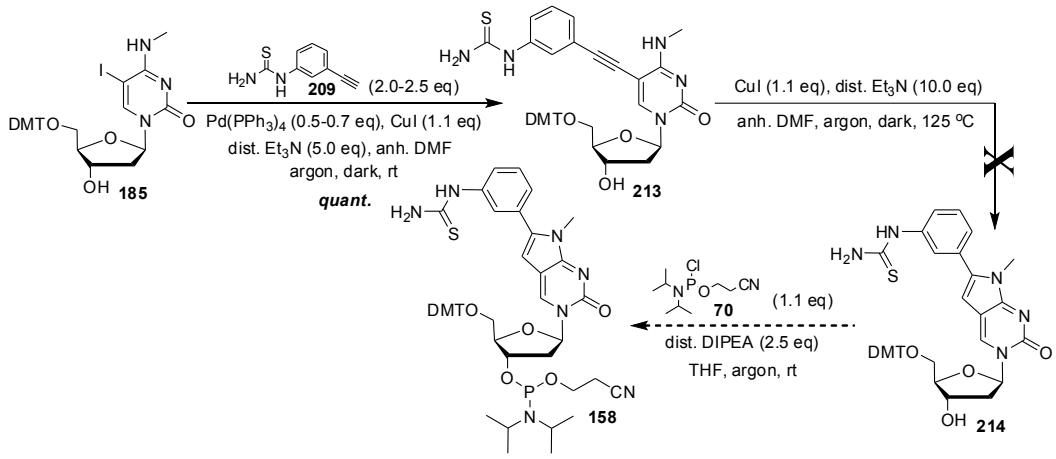
Although the ^{1}H-NMR spectrum was too complicated to assign, both the desired phosphoramidite **199** and *bis*-phosphoramidite **208** were also detected by MS.

When the monomer was re-synthesised in house, by Dr. Edrees-Abdou, guanidinylation was achieved using 2.5 eq of GR **200**, 5 eq of pyridine in refluxing CH_2Cl_2 under argon pressure (50 °C) for 2 days, in 44% yield. Phosphitylation using 1.8 eq of chloro-phosphitylating reagent **70** with 2 eq of DIPEA at rt afforded the desired $^{G}PP_p$ monomer after 2 hours reaction time, in 58% yield. The monomer precursor **207** was precipitated from DCM into hexane to remove residual pyridine and methanethiol and no degradation was encountered during storage.

5.1.5 Synthesis of 6-(3-Thioureidophenyl)-*N*-methylpyrrolo-dC monomer **158**

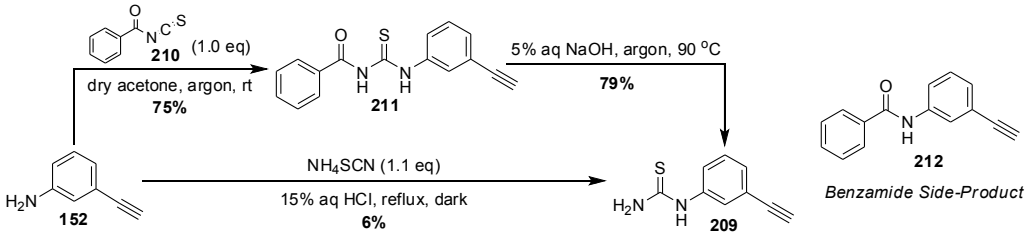
The final modification investigated was a thioureido moiety. The greater acidity of thiourea N-H bonds compared to the analogous urea, should result in strengthening of the proposed hydrogen-bonds with G of the CG base pair.

The initial synthetic route to the thiourea-modified monomer ($^{TU}PP_p$) **158** closely followed the procedure for the $^{U}PP_p$ monomer **179** (Scheme 5.10). After synthesis of the modified alkyne **208**, Sonogashira cross-coupling, cyclisation and phosphitylation were expected to afford the desired monomer **158**.



Scheme 5.10. Initial synthetic route to $^{TU}PP_p$ monomer **158**.

Synthesis of the alkyne for Sonogashira cross-coupling, 3-thioureidophenyl acetylene **209**, was first attempted using one of two methods published by Saxena *et al.*[227] for synthesis of *N*-arylthioureas. Refluxing 3-ethynylaniline **152** in 15% aq HCl with ammonium thiocyanate, however, afforded only 6% yield of product **209** with some recovery of starting material. Poor solubility despite strong heat and vigorous stirring may have hindered the reaction (Scheme 5.11).



Scheme 5.11. Synthetic routes to 3-thioureidophenyl acetylene **209**.

After noting that isothiocyanates could be used for formation of aryl-thioureas,[227,228] 3-EA **152** was treated with trimethylsilyl isothiocyanate (TMS-NCS) in dry Et_2O, with and without base (DIPEA). No desired thiourea could be observed; only thiocyanic acid salts of DIPEA or 3-EA were formed, which were broken down during workup. Although potentially a very convenient route to primary ureas due to facile deprotection of the TMS-thiourea product, TMS-NCS appears to hydrolyse before the poorly reactive aniline can react.

The other method published by Saxena[227] was used successfully to afford the desired alkyne **209** in 59% yield over two steps. 3-Ethynylaniline **152** was reacted with benzoyl isothiocyanate **210** cleanly in 75% yield. Cleavage of the benzoyl group of **211** was achieved by heating at 90 °C in 5% aq NaOH for 15 minutes in 79% yield. Neither step required chromatography and little benzamide byproduct **212** was produced (removed by filtration) (Scheme 5.11).

Sonogashira cross-coupling reactions throughout this work have used the catalyst, $Pd(PPh_3)_4$. This posed a problem for this step, as due to the high affinity of sulphur for palladium, it was likely that the thiourea moiety may bind to palladium, either poisoning or consuming the catalyst.

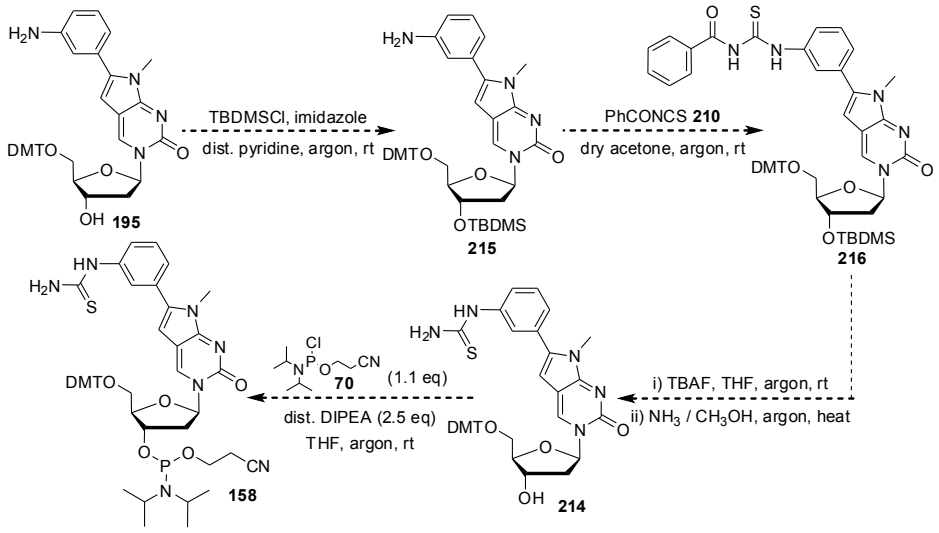
Small-scale test reactions demonstrated that much higher quantities of Pd catalyst and CuI were required to elicit any reaction other than minimal alkyne dimerisation. Two other products were also detected by TLC, which appeared to correspond to Pd-thiourea complexes of the alkynylated product **213** and alkyne **209**.

This reaction was subsequently conducted on a 1.3 g and 0.3 g scale using 0.5-0.7 eq of $Pd(PPh_3)_4$, 1.1 eq of CuI and 2.0-2.5 eq of alkyne, during which, both starting material and alkyne were consumed. After column chromatography, the product from the larger-scale reaction was divided. Aqueous workup was unsuccessful removing inorganic impurities, however, passing through a benzyl-thiol StratoSpheres™ SPE cartridge (PL-BnSH MP Resin) successfully cleaved the residual thiourea complexed-Pd and removed excess Pd from the mixture.

After combining all batches, a portion of the Sonogashira product **213** was treated under Inoue's conditions for cyclisation.[222,223]. Unexpectedly, after 40 minutes, nothing had

occurred except precipitation of a black solid. The mixture was filtered, concentrated *in vacuo*, dried and treated again under the same conditions, after confirming the starting material had remained unchanged. The starting material was mostly consumed, as determined by TLC analysis using several solvent systems, over 1½ hours hence the reaction was worked up and purified twice by column chromatography. TLC analysis of the resulting seven fractions indicated a complex mixture of products, some fluorescent and all UV active. The desired cyclised product **214** could not be detected by MS; only the analogous ureido compound **184** was identified. It was subsequently discovered that in recent work by Shibahara *et al.*,[229] a variety of thiocarbonyl compounds were oxidatively desulfurised in presence of copper (I) chloride (20 mol%), oxygen and heat (80 °C) in excellent yield under the right conditions. Under the conditions of Inoue, it is possible that traces of oxygen, combined with strong heating (125 °C) and CuI (1.1 eq), may have induced oxidative desulfurisation and other degradation products. This explains the presence of the urea "oxidation product" **183** in the mixture.

The remaining uncyclised starting material was treated, without CuI, under the otherwise same conditions affording only complete detritylation of the starting material. This was recovered and treated under Inoue's conditions, at 50 °C this time, but no change was observed after 24 hours. It was decided to abandon this synthetic route and a new route was proposed.

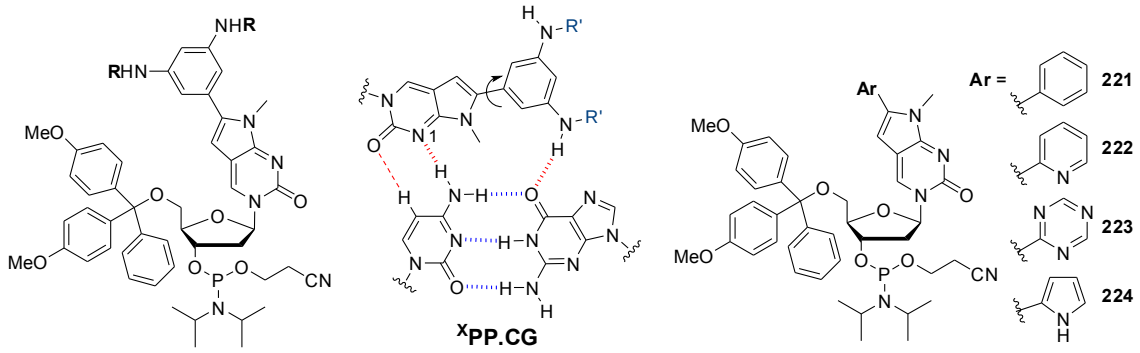


Scheme 5.12. Revised route to $^{TU}PP_p$ monomer **158**.

The new proposed route proceeds *via* modification of the cyclised anilino-nucleoside **195**, with benzoyl isothiocyanate **210**, following suitable protection of the 3'-OH (Scheme 5.12). This synthesis was subsequently investigated by Dr. Edrees-Abdou with success after some further modifications to the route. The phosphoramidite **158**, however, coupled very poorly when incorporated into DNA and MS analysis proved inconclusive for the desired oligonucleotides. The thiourea-derivative ($^{TU}PP_p$) was therefore put on hold. It was proposed that deprotection conditions and/or the iodine oxidation step during oligonucleotide synthesis may have caused degradation/oxidation but this could not be confirmed.

5.2 Future Work

Following this work, other modifications were suggested, which could enhance binding and recognition. The first was substituting the 3-aminophenyl- modification for a 3,5-diaminophenyl- moiety. Due to free rotation around the aryl-bicycle bond, it is possible that the 3-amino group may not always be correctly aligned for binding to G. Employing two groups should present the correct hydrogen-bonding for both rotamers. Any one of the *bis*-acetamide, *bis*-urea, *bis*-trifluoroacetamide or *bis*-guanidinium could be synthesised. However, the *bis*-acetamide and *bis*-urea are likely to be too polar to handle, and the *bis*-guanidinium might be difficult to form, due both to low reactivity and steric hindrance.



R = $COCF_3$ **217**, $COCH_3$ **218**
$CONH_2$ **219**, C(NP)NHP **220**

Figure 5.4. Alternative CG recognition monomer synthetic targets **217-224** and binding motif.

It may also be beneficial to change the aromatic ring to a heterocycle such as pyridine, triazine, pyrrole or imidazole. Altering hydrophilicity and electronic properties may have beneficial effects on binding affinity.

The first target, however, is 6-phenyl-*N*-methylpyrrolo-dC phosphoramidite **220**. It is important to confirm what role the aromatic ring plays in triplex binding and recognition, and whether the modifications do benefit binding or could even be omitted. In addition, it would be necessary therefore to investigate how changing the ring, whether aromatic (phenyl, napthyl), heteroaromatic (pyridine, triazine, pyrrole, thiazole) or non-aromatic (cyclohexyl, cyclopentyl), affects binding, before further derivatisation. The phenyl monomer **221** has been synthesised in house, in excellent overall yield by Dr. Edrees-Abdou, and biophysical studies are in progress.

5.3 Conclusion

Three *N*-methylpyrrolo-dC monomers ($^{A}PP_p$ **178**, $^{Ac}PP_p$ **191**, $^{U}PP_p$ **179**), were synthesised for incorporation into TFOs for biophysical studies. All were synthesised from the key intermediate, 5'-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxy-cytidine **185**, made from the 2'-deoxyuridine analogue **97** in two steps. Sonogashira cross-coupling[124] and a modified copper (I)-catalysed cyclisation[222,223] were used to form the derivatised bicyclic base in typically good to excellent yield.

Problems were again encountered with the low reactivity of the aniline NH_2 (*c.f.* Section 3), necessitating prolonged reaction times, a higher reagent excess, and raised temperature (e.g. trifluoroacetylation, guanidinylation). Care was also taken for the guanidinium monomer ($^{G}PP_p$) and precursor, to avoid cleavage of the base-labile protecting groups, whilst preventing cleavage/activation of the acid-sensitive DMT and phosphoramidite moieties. A revised method for synthesis of the guanidinylating reagent **200** from CEOC-succinimide **201** was also evaluated, affording a final semi-optimised yield of 89%, compared to the literature yield of 35%.[41,225,226] This allows for a much more accessible synthesis, owing to cost of DSC.

The fourth monomer, $^{G}PP_p$ **199**, was synthesised later in house, after synthetic investigation (see above), and a fifth monomer, $^{TU}PP_p$ **158**, although eventually

synthesised later in house, could not be incorporated and studied as desired. Although the thiourea moiety should have imparted extra stability over the urea in the triplex, it is unlikely this modification will be continued. Instability during oligonucleotide synthesis and/or deprotection introduced too many problems. However, there are several new directions for future work, and the syntheses and biophysical studies (Section 6) already undertaken are promising.

The precursors to the four monomers that were incorporated into TFOs, were also deprotected and fluorescence measurements were taken. This is discussed in detail in Section 6.3.

Chapter 6

Biophysical Studies of *N*-Methylpyrrolo-dC Modified TFOs for CG Recognition

6. Biophysical Studies of *N*-Methylpyrrolo-dC Modified TFOs for CG Recognition

Four modified *N*-methylpyrrolo-dC phosphoramidite monomers ($^{A}PP_p$ **178**, $^{Ac}PP_p$ **191**, $^{U}PP_p$ **179**, $^{G}PP_p$ **199**) were incorporated into unlabelled/labelled TFOs for UV and fluorescence melting studies. The initial study (from Section 4.2.2) was designed to compare both post-synthetically modified furano-dT, and *N*-methylpyrrolo-dC nucleosides at CG recognition in the four sequences introduced in Chapter 4. Following this work, the main biophysical studies were carried out. These were again designed to test binding affinity for CG and selectivity against the other base pairs (GC, AT, TA) at several pHs. In addition to determination of the best modification for CG recognition, the other aim was to compare UV and fluorescence melting data (collected using the same buffer conditions), which appeared not to correlate in previous studies.

6.1 Initial Fluorescence Melting Study

This section discusses the *N*-methylpyrrolo-dC results from the 'Sequence Study' introduced in Section 4.2.2.

```
          5′-Q-PMM TPM VXW TPT PTM PT
5′-F-GT GTT AGG AAG ACA AAA AAG AAC TGG T-H )
     CA CAA TCC TTC TGT TTT TTC TTG ACC A-H )
```

Figure 6.1. Fluorescence triplex melting experiment. TFO shown in bold. **M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **X** = ^{Ac}PP & **VXW** = TXT (OL18), PXT (OL19), TXP (OL20), PXP (OL21), or **X** = ^{U}PP & **VXW** = TXT (OL22), PXT (OL23), TXP (OL24), PXP (OL25), **Q** = DABCYL (quencher), F = FAM (fluorophore), H = HEG (hexaethylene glycol), hairpin duplex (**CG**, OL6).

As previously described, these fluorescence melting experiments (Figure 6.1) were conducted using LC Programmes 1 and 2 (95-30-95 °C at 0.2 °C/min) at pH 6.2 (10 mM sodium phosphate, 200 mM NaCl). Experiments were run once in triplicate and a 10:1 ratio of TFO:duplex (5:0.5 μM) was used.

X Seq.	T_m / °C, melt/anneal	
	^{Ac}PP	^{U}PP
TXT	42.1/38.0	44.0/42.0
PXT	49.0/44.0	48.1/45.1
TXP	49.1/44.0 (48.4/43.0)	51.1/47.1 (51.0/45.0)
PXP	54.0/49.0	54.5/50.0

Table 6.1. Fluorescence triplex melting experiment: $^{Ac}PP_p$ and $^{U}PP_p$ against CG, each in four TFOs (^{Ac}PP – OL18-OL21 & ^{U}PP – OL22-OL25) at pH 6.2 (10 mM sodium phosphate, 200 mM NaCl) using LC Programme 1 and 2 (Section 9.2). Average T_m values for TFO melt/anneal given in °C. Data obtained for $^{X}PP_f$-modified TFOs is shown for comparison, in parentheses (Section 4.2.2).

These results confirm the previous findings, that placing pdU (P) on the 5'-side of the modification X (^{X}PP) has a slightly destabilising effect compared to the 3'-side, notably for the urea (^{U}PP). The results for the TXP TFOs for both furano-dT (in parentheses) and *N*-methylpyrrolo-dC modifications were similar (within experimental error), hence previous furano-dT experimental data can be corroborated.

The trends identified when examining the combined $^{X}PP_f$ and $^{X}PP_p$ results, correlate with those in the main fluorescence melting studies, despite differences in buffer composition and pH.

In addition, hysteresis was on average 1.5 °C lower for $^{X}PP_p$ than for $^{X}PP_f$ modifications. This may rise from the higher TFO purity when no post-synthetic modification is required (Sections 3.3, 5.1).

6.2 Primary UV Melting Study

A comprehensive UV melting study was conducted in cooperation with Dr. Edrees-Abdou, to examine the binding affinity/selectivity of the four monomers, in comparison also with the natural base T. Data was collected at pH 6.2, 6.6 and 7.0, and at pH 7.0 with 2 mM spermine added. Spermine is a general/non-specific DNA stabilising agent and component in all eukaryotic cells, which is used in this context, to increase the stability of a duplex/triplex if the T_m is too low to be easily measured. Prior to this

work, large quantities of the pdU phosphoramidite monomer **164** were synthesised by Dr. Edrees-Abdou and Dr. Imenne Bouamaied for synthesis of the TFOs used in this both studies.

```
          5′-TTT TTM VXW MTM TMT
5′-GCT AAA AAG AYA GAG AGA TCG
   CGA TTT TTC TZT CTC TCT AGC-5′
```

Figure 6.2. UV triplex melting experiment. TFO shown in bold. **M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **VXW** = TTT (OL32), TTP (OL33), PTT (OL34), PTP (OL35); **X** = ^{A}PP & **VXW** = TXT (OL36), TXP (OL37), PXT (OL38), PXP (OL39); **X** = ^{Ac}PP & **VXW** = TXT (OL40), TXP (OL41), PXT (OL42), PXP (OL43); **X** = ^{U}PP & **VXW** = TXT (OL44), TXP (OL45), PXT (OL46), PXP (OL47); **X** = ^{G}PP & **VXW** = TXT (OL48), TXP (OL49), PXT (OL50), PXP (OL51); purine strand **Y** = C (OL1), G (OL26), A (OL27), T (OL28), pyrimidine strand **Z** = G (OL2), C (OL23), T (OL30), A (OL31).

UV melting experiments were performed (as described previously, and in detail in Section 9.2), using a 5:1 ratio of TFO:duplex (5:1 μM). Experiments were conducted at pH 6.2, 6.6, 7.0 and pH 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA). All oligonucleotides for this study were synthesised without capping to avoid undesired acetylation of TFOs.

Standard melt programme No. 3 was used (15-50-80 °C at 0.25-0.5 °C/min) with 5 heat/ anneal cycles in total (Section 9.2). Condensation was not a problem at low temperature (15 °C).

There are several trends, which can be drawn from the data given below in Table 6.2. Firstly, at pH 6.2, against CG in sequence TXT, (Table 6.2, Lines 1, 6, 19, 28, 37) the aniline/amine (^{A}PP, OL36) outperforms all other monomers, and the poorest monomer (^{Ac}PP, OL40) gives a T_m that is 1.7 °C greater than T. This result appears to contradict the original order of stability as derived from the proposed hydrogen-bonding triplet motif. The proposed motif places the urea (^{U}PP) and guanidinium (^{G}PP) as the strongest binders, as they can form two hydrogen-bonds and the latter also contributes additional charge-stabilisation. However, these results suggest steric or electronic factors may play a role, where such groups cause destabilisation relative to the sterically non-hindering amino group.

X	Sequence	Base Pair	Buffer	T_m / °C	Hysteresis	
T	TXT	CG	6.2	25.6/22.9	2.7	1
	TXT	CG	6.6	19.4/n.d.	n.d.	2
	TXP			n.d.	n.d.	3
	PXT			23.2/20.1	3.1	4
	PXP			22.4/18.8	3.6	5
$^{A}PP_p$	TXT	CG	6.2	30.0/26.3	3.7	6
		GC		21.1/18.1	3.0	7
		AT		21.4/18.4	3.0	8
		TA		n.d.	n.d.	9
	TXT	CG	6.6	22.7/18.5	4.2	10
	TXP			24.0/19.8	4.2	11
	PXT			25.7/21.2	4.5	12
	PXP			26.8/21.9	4.9	13
	TXT	CG	7.0	n.d.	n.d.	14
	TXT	CG	7.0 SP	30.4/29.4	1.0	15
		GC		20.2/n.d.	n.d.	16
		AT		17.5/n.d.	n.d.	17
		TA		n.d.	n.d.	18
$^{Ac}PP_p$	TXT	CG	6.2	27.3/23.1	4.2	19
		GC		20.4/n.d.	n.d.	20
		AT		20.4/n.d.	n.d.	21
		TA		20.3/n.d.	n.d.	22
	TXT	CG	6.6	21.0/16.8	4.2	23
	TXP			21.1/n.d.	n.d.	24
	PXT			22.6/19.4	3.2	25
	PXP			22.7/18.3	4.4	26
	TXT	CG	7.0	n.d.	n.d.	27
$^{U}PP_p$	TXT	CG	6.2	29.2/24.4	4.8	28
		GC		20.1/n.d.	n.d.	29
		AT		21.5/18.5	3.0	30
		TA		21.6/19.5	2.1	31
	TXT	CG	6.6	22.7/18.5	4.2	32
	TXP			23.4/18.3	5.1	33
	PXT			24.1/18.3	5.8	34
	PXP			24.9/19.0	5.9	35
	TXT	CG	7.0	n.d.	n.d.	36
$^{G}PP_p$	TXT	CG	6.2	27.9/25.0	2.9	37
		GC		19.5/n.d.	n.d.	38
		AT		20.7/n.d.	n.d.	39
		TA		17.8/n.d.	n.d.	40
	TXT	CG	6.6	21.4/18.0	3.4	41
	TXP			21.1/18.0	3.1	42
	PXT			21.6/18.6	3.0	43
	PXP			22.5/20.3	2.2	44
	TXT	CG	7.0	n.d.	n.d.	45

Buffer / Duplex	6.2	6.6	7.0	7.0 SP
CG	61.3	62.8	61.6	63.6
GC	60.5	-	-	63.4
AT	58.4	-	-	61.8
TA	58.6	-	-	61.7

Hysteresis	1.1 – 5.9 °C
Average Hysteresis	3.7 °C
Average duplex ΔT_m due to spermine	+ 2.9 °C

Table 6.2. UV triplex melting experiment: X against YZ at pH 6.2, 6.6, 7.0 and 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA) using UV melting programme No. 3 (Section 9.2). Average T_m values for TFO melt/anneal and duplex (average of melt and anneal T_m), and hysteresis given in °C. n.d. = not determined (T_m < 17 °C).

At pH 6.6 (Lines 2, 10, 23, 32, 41), the urea and aniline (^{A}PP) show similar binding affinity, greater than the guanidinium, which should be the strongest binder.

At pH 6.6, against CG for all the pdU-containing TFOs, the aniline is again the strongest binder, followed by the urea. The order of stability for the other modifications varies, such that for PXT (OL50) and PXP (OL51), the guanidinium is of equal or lower stability than T. This may is attributed to disfavourable steric interaction and also charge-repulsion with the pdU on the 5'-side as previously discussed.

All monomers demonstrated a very high selectivity for CG at pH 6.2, over the other base pairs, by 6.9-8.6 °C. For all monomers except ^{U}PP, the second highest affinity was for AT not TA, the latter of which has often been observed for CG recognition monomers (Section 1.2.2). A putative X.AT triplet can be drawn to support this observation, with hydrogen bonds between N^3 and anilino NH of ^{X}PP and the NH_2 of A and C^4=O of T. Only ^{U}PP and ^{G}PP show significant differences (> 1 °C) between the T_ms against the other base pairs. In addition, the aniline demonstrates very high selectivity for CG, by 10.2 °C at pH 7.0 with 2 mM spermine, where the second highest affinity was for GC (Lines 15-18).

Examination of the data for the pdU-containing TFO series, obtained at pH 6.6, identifies significant differences from the initial fluorescence melting data discussed in Sections 4.2 and 6.1. The main difference is order of stability, where higher T_ms are observed for PXT- rather than TXP-oligonucleotides. It was deduced from previous fluorescence melting data, that disfavourable interaction between the pdU and ^{X}PP amino-modified side-chains occurred when pdU was located on the 5'-side. However, the reverse is observed here, and the difference between TXP and PXT is pronounced for the aniline (1.7 °C) and acetamide (1.4 °C). Otherwise increase in T_m with each T-pdU substitution varies from 0.1 to 1.3 °C. Interestingly, the TXT TFO (X = ^{G}PP, OL48) is 0.2 °C more stable than the corresponding TXP oligonucleotide (OL49), but this may simply be due to experimental error. The results for T at pH 6.6, appear spurious, where the order of stability (TXP < TXT < PXP < PXT) cannot be easily explained.

In summary, this study highlights the high selectivity of all monomers for CG over the other base pairs. The order of binding affinity, however, contradicts the proposed triplet

binding motif, although the urea remains a strong binder. These observations suggest the aniline or urea would be highly suitable candidates for CG recognition, the latter perhaps particularly at higher pH. Selectivity at pH 7.0 with 2 mM spermine should be carried out for the urea to confirm whether the urea does indeed outperform the aniline at higher pH. Without spermine, no T_m could be measured for any monomer at pH 7.0. The acetamide and guanidinium should also be assessed against CG at pH 7.0 + 2 mM spermine to determine whether the effect of the positive-charge of the guanidinium is beneficial to binding at higher pH, where little protonation of other amine functionalities should occur. The acetamide performed poorest throughout and the guanidinium proved less effective than expected.

6.3. Primary Fluorescence Melting Study

The main fluorescence melting study was conducted in cooperation with Dr. Edrees-Abdou, to confirm the previous results and make more accurate comparisons between monomer recognition capabilities. Data was collected using the same buffer systems as for the UV melting study so that direct comparisons between the studies could be made.

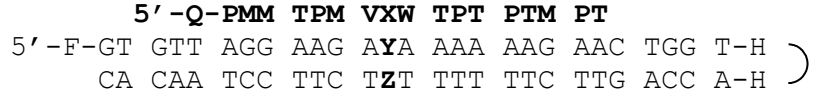


Figure 6.3. Fluorescence triplex melting experiment. TFO shown in bold. **M** = $^{5\text{-Me}}$dC, **P** = 5-(3-aminoprop-1-ynyl)-dU (pdU), **VXW** = TTT (OL60), PTT (OL61), TTP (OL62), PTP (OL63); **X** = ^{A}PP & **VXW** = TXT (OL52), PXT (OL53), TXP (OL54), PXP (OL55); **X** = ^{Ac}PP & **VXW** = TXT (OL18), PXT (OL19), TXP (OL20), PXP (OL21); **X** = ^{U}PP & **VXW** = TXT (OL22), PXT (OL23), TXP (OL24), PXP (OL25); **X** = ^{G}PP & **VXW** = TXT (OL56), PXT (OL57), TXP (OL58), PXP (OL59);. **Q** = DABCYL (quencher), F = FAM (fluorophore), H = HEG (hexaethylene glycol), hairpin duplex **YZ** = CG (OL6), GC (OL7), AT (OL8), TA (OL9).

X	Sequence	Base Pair	Buffer	T_m / °C	Hysteresis	
T	TXT	CG	6.2	44.3/41.0	3.3	1
	TXP			47.5/45.9	1.6	2
	PXT			48.4/47.8	0.7	3
	PXP			51.1/48.7	2.4	4
$^{A}PP_p$	TXT	CG	6.2	44.1/40.9	3.2	5
		GC		36.7/32.9	3.8	6
		AT		33.9/n.d.	n.d.	7
		TA		36.7/33.8	2.9	8
	PXP	CG		56.8/52.9	3.9	9
		GC		45.8/42.9	2.9	10
		AT		43.3/40.9	2.4	11
		TA		47.2/43.9	3.3	12
	TXT	CG	6.6	37.5/32.9	4.6	13
		GC		n.d.	n.d.	14
		AT		n.d.	n.d.	15
		TA		n.d.	n.d.	16
	TXP	CG	6.6	43.2/35.9	7.3	17
	PXT			42.2/35.4	6.8	18
	PXP			48.5/41.4	7.1	19
	TXT	CG	7.0	n.d.	n.d.	20
	PXP			40.6/n.d.	n.d.	21
	TXT	CG	7.0 SP	36.0/33.5	2.5	22
$^{Ac}PP_p$	TXT	CG	6.2	43.9/40.8	3.1	23
		GC		38.8/35.9	2.9	24
		AT		35.0/32.7	2.3	25
		TA		41.9/38.0	3.9	26
	PXP	CG	6.2	55.2/50.2	5.0	27
		GC		46.9/43.8	3.1	28
		AT		45.0/39.9	5.1	29
		TA		50.4/47.0	3.4	30
	TXP	CG	6.2	50.3/46.1	4.2	31
	PXT			51.0/46.0	5.0	32
	TXT	CG	6.6	36.2/32.8	3.4	33
		GC		n.d.	n.d.	34
		AT		n.d.	n.d.	35
		TA		33.9/n.d.	n.d.	36
	TXP	CG	6.6	43.8/35.7	8.1	37
	PXT			43.8/35.7	8.1	38
	PXP			46.5/37.8	8.7	39
	TXT	CG	7.0	n.d.	n.d.	40
	PXP			39.9/n.d.	n.d.	41
	TXT	CG	7.0 SP	36.0/34.2	1.8	42

Table 6.3. Fluorescence triplex melting experiment: X against YZ at pH 6.2, 6.6, 7.0 and 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA) using LC Programme 3 (Section 9.2). Average T_m values for TFO melt/anneal and values for hysteresis given in °C. n.d. = not determined (T_m < 33 °C). Hysteresis: 0.7–9.3 °C, average hysteresis: 4.0 °C. *Continued on following page.*

X	Sequence	Base Pair	Buffer	T_m / °C	Hysteresis	
$^{U}PP_p$	TXT	CG	6.2	44.3/40.9	3.4	43
		GC		39.9/34.4	5.5	44
		AT		35.1/32.8	2.3	45
		TA		40.7/36.9	3.8	46
	PXP	CG	6.2	58.1/53.5	4.6	47
		GC		46.0/45.0	1.0	48
		AT		43.8/40.8	3.0	49
		TA		49.7/45.1	4.6	50
	TXP	CG	6.2	52.0/48.9	3.1	51
	PXT			50.8/45.8	5.0	52
	TXT	CG	6.6	38.0/n.d.	n.d.	53
		GC		n.d.	n.d.	54
		AT		n.d.	n.d.	55
		TA		33.5/n.d.	n.d.	56
	TXP	CG	6.6	44.2/38.8	5.4	57
	PXT			43.0/37.7	5.3	58
	PXP			48.2/43.9	4.3	59
	TXT	CG	7.0	n.d.	n.d.	60
	PXP			39.9/32.8	7.1	61
	TXT	CG	7.0 SP	37.4/34.8	2.6	62
$^{G}PP_p$	TXT	CG	6.2	47.2/45.8	1.4	63
		GC		36.9/35.9	1.0	64
		AT		37.0/35.9	1.1	65
		TA		39.7/38.5	1.2	67
	PXP	CG	6.2	56.0/51.4	4.6	68
		GC		45.0/40.9	4.1	69
		AT		47.0/44.5	2.5	70
		TA		51.0/47.2	3.8	71
	TXT	CG	6.6	40.3/35.0	5.3	72
		GC		n.d.	n.d.	73
		AT		30.9/n.d.	n.d.	74
		TA		33.0/n.d.	n.d.	75
	TXP	CG	6.6	46.2/36.9	9.3	76
	PXT			43.7/36.8	6.9	78
	PXP			47.2/41.9	5.3	79
	TXT	CG	7.0	32.7/n.d.	n.d.	80
	PXP			38.5/n.d.	n.d.	81
	TXT	CG	7.0 SP	39.0/37.9	1.1	82

Table 6.3 *cont.* Fluorescence triplex melting experiment. *Continued from previous page.*

Fluorescence melting studies were performed (as described previously, and in detail in Section 9.2), on a Roche LightCycler®. Experiments were performed at pH 6.2, 6.6 and 7.0, and at pH 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA), and a 10:1 ratio of TFO:duplex (5:0.5 μM) was used. Runs in these experiments were performed up to three times in triplicate, using LC Programme 3 (95-27-95 °C at 0.2 °C/min, see Section 9.2).

At pH 6.2, against CG in sequence TXT, (Table 6.3, Lines 1, 5, 23, 43, 63), ^{G}PP outperforms the other monomers by 2.9 °C, however, the T_m of the poorest monomer (^{Ac}PP) is 0.4 °C below T. The aniline, urea and T have similar T_ms. At pH 6.6 also (Lines 13, 33, 53, 72), the guanidinium monomer has the strongest binding affinity by 2.3 °C, followed by the urea, aniline then acetamide. In the PXP sequence at pH 6.2 (Lines 9, 27, 47, 68), the urea has the highest T_m by 1.3 °C, followed surprisingly by the aniline then guanidinium and acetamide, which is 4.1 °C above T. A similar pattern was observed for TXT and PXP at pH 7.0 and 7.0 with spermine.

The guanidinium monomer performed best, as expected, in the TXT sequence but not for PXP. It is possible that the presence of a protonated amine species either side of the positively-charged guanidinium may have destabilised it significantly, below that of the aniline as observed. Otherwise, apart from the acetamide performing slightly worse than the aniline, the order of stability can be rationalised.

A different situation was encountered for the PXT and TXP sequences. At pH 6.6, for TXP (Lines 17, 37, 57, 76), the guanidinium and urea performed best, whereas for PXT (Lines 18, 38, 58, 77), these appear destabilised such that the acetamide has a similar T_m to ^{G}PP. This indicates that both the guanidinium and urea have a negative interaction with pdU when they are located on the 5'-side (PXT) as originally noted (Section 4.2.2). It is also perhaps surprising, that changing from PXT to PXP has a significant effect on the order of binding affinity, whereas changing from TXT to TXP makes little difference. It was originally thought that the proposed negative interaction was due to steric hindrance between the aminopropynyl group of pdU and modified phenyl group of the monomer. However, charge-repulsion may also play an important role.

All monomers demonstrate good to high selectivity for CG over the other base pairs, at all pHs tested in both TXT and PXP sequences. In all cases, the second most stable binding is to TA as originally expected. The least stable binding occurs to GC then AT, or to AT then GC for ^{G}PP. For TXT at pH 6.2, the guanidinium and aniline demonstrate the highest selectivity, by 7.5 °C, and the acetamide the least, by 2.0 °C. For TXT at pH 6.6, the guanidinium is most selective (ΔT_m 7.4 °C) and the acetamide the least (ΔT_m 2.4 °C), and for PXP at pH 6.2, the aniline is most selective (ΔT_m 9.6 °C), followed by the urea (ΔT_m 8.4 °C), the guanidinium (ΔT_m 5.0 °C) then the acetamide (ΔT_m 4.7 °C).

The high selectivity of ^{G}PP and ^{U}PP compared to ^{Ac}PP can be rationalised by the extra potential hydrogen-bond, which can effect differentiation between the binding motifs of CG and the other base pairs more effectively. It does not explain the good selectivity for the aniline possessing only one potential hydrogen-bond donor, although in the PXP sequence, general destabilisation of ^{G}PP and ^{U}PP, due to neighbouring pdU nucleotides, may reduce their selectivity relative to the aniline.

Data for the pdU-containing TFOs demonstrates greater stability for TXP over PXT for aniline, urea and guanidinium, and the reverse for the acetamide, in all cases (not tested at pH 6.2 for ^{A}PP and ^{G}PP). The greatest difference between TXP and PXT occurs for ^{G}PP (ΔT_m 2.5 °C), is moderate for ^{U}PP (ΔT_m 1.2 °C) and ^{A}PP (ΔT_m 1.1 °C) and small for ^{Ac}PP (av. ΔT_m 0.4 °C).

Charge-repulsion may explain this difference for the guanidinium monomer, and steric hindrance for the urea, but there must be alternative reasons for the observed differences for the aniline and acetamide.

The fluorescence melting study results suggest the aniline, urea or guanidinium would prove suitable candidates. The guanidinium has the greatest binding affinity in most of the conditions tested and the aniline proved the most selective, followed by the guanidinium and urea. Due to sensitivity of intermediates in the synthesis of the guanidinium monomer, either the urea or aniline are probably more suitable.

6.4. Melting Study Comparison

The fluorescence melting study data (Section 6.3) correlates with previous fluorescence melting data (Sections 4.2.2, 6.1), but not with the UV melting study data (Section 6.2). Importantly, the order of binding affinity has stark differences between the two studies.

It is possible that the sequence composition may be partly responsible for the observed differences. Examining the neighbouring triplets to the central three-triplet triplex core of the UV and fluorescence melting motifs, only the triplet on the 5'-side is the same (M^{+}.GC) (Figure 6.4).

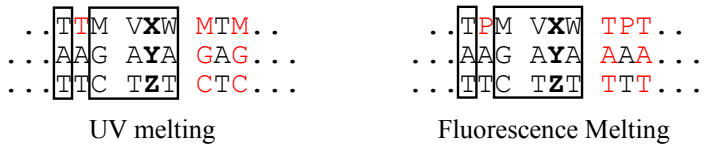


Figure 6.4. Central sections of UV and fluorescence melting triplex motifs highlighting matched triplets (boxed) and differences (red).

It is difficult to explain all the differences between the two studies, however, by analysing the differences in the sequence around the central core, without extensive and time-consuming melting and NMR experiments. It is likely that the general triplex shape, general TFO composition and whether the duplex part is a hairpin or comprises separate strands, contributes to the differences in trends between systems.

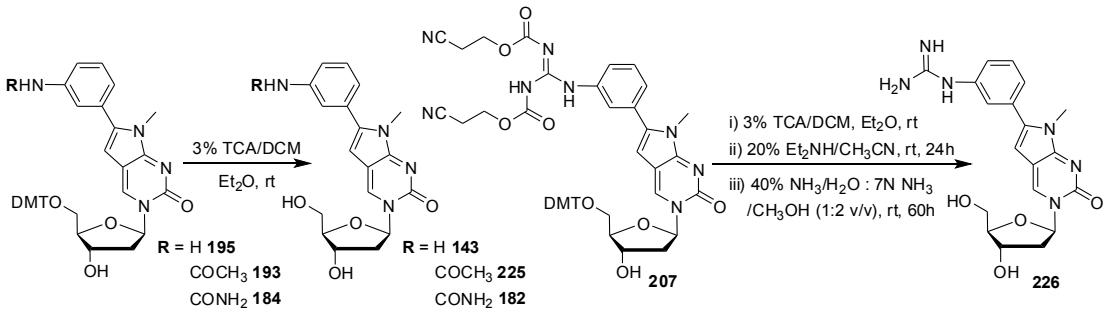
In order to rule out the possibility that the method of analysis is responsible, although unlikely, UV melting experiments should be performed using the fluorescence melting triplex motif. The fluorophore and quencher should not affect the trends, only causing a slight general increase in T_m. The oligonucleotides could also be resynthesised without DABCYL and FAM, although this would create significant work.

If the method of analysis is proven not to be a factor, the differences must be sequence-dependent. This would have considerable implications when designing TFOs for gene-targeting and oligotherapeutics. Depending on the sequence and also interaction with other modified/un-natural nucleotides and recognition monomers, the same monomer may have significant differences in selectivity and binding affinity. Optimising both of these properties is vitally important for achieving mixed-sequence recognition under physiological conditions.

6.5. Fluorescence Properties of CG Recognition Monomers

The synthesis and fluorescent properties of pyrrolo-C nucleobases, nucleosides and PNA monomers have been studied in detail by Robert Hudson *et al.*[120,230-235] and others[119] as they can be used as fluorescent reporter groups in SNP analysis[236] and for fluorimetric detection of guanosine.[120,233] Fluorescence decreases significantly only on duplex hybridisation with G, such that the change is visible to the naked eye on irradiation of the solution with long-wave UV light (365 nm).

These pyrrolo-dC compounds have not, however, been evaluated as probes or fluorophores in triplexes, and only the triplex-forming properties of *N*-methylpyrrolo-dC derivatives have been studied thus far.[67,122,207] For this purpose, the fluorescence properties of the four deprotected nucleosides, ^{A}PP **143**, ^{Ac}PP **225**, ^{U}PP **182** and ^{G}PP **226** were evaluated to afford an indication of their usefulness in this application. Their deprotection procedures are described in Section 9.2.6 (Scheme 6.1)



Scheme 6.1. Deprotection of *N*-methylpyrrolo-dC nucleosides **184,193,195,207**.

The excitation and emission wavelengths were measured in HPLC grade methanol (^{A}PP **143**, ^{Ac}PP **225**) or anhydrous DMF (^{U}PP **182**, ^{G}PP **226**) using a Varian Cary 400 Scan UV-Visible Spectrophotometer and a Varian Cary Eclipse Fluorescence Spectrophotometer respectively (Section 9.2). The extinction coefficient was calculated for each excitation maximum using the Beer-Lambert Law ($A = \varepsilon cL$; A = absorbance, ε = molar extinction coefficient, c = molar concentration, L = pathlength/cm – see Section 9.2) and all data is detailed in Table 6.4 below.

X	λ_{abs1}	ε_1	λ_{abs2}	ε_2	λ_{abs3}	ε_3	λ_{em}	λ_{em}–λ_{abs1}
$^{A}PP_p$	356.3	3730	237.8	16680	203.2	12100	448.5	92.2
$^{Ac}PP_p$	355.5	5530	243.4	33060	202.8	19130	458.0	102.5
$^{U}PP_p$	358.5	15140	287.4	13740	276.4	13410	429.5	71.0
$^{G}PP_p$	366.6	6690	270.2	22380	–	–	458.0	91.4

Table 6.4. UV/Fluorescence data for *N*-methylpyrrolo-dC nucleosides **143,225,182,226**: $\lambda_{abs1,2,3}$ = excitation/absorption maxima (nm), λ_{em} = emission wavelength (nm) from excitation at λ_{abs1}, ε = molar extinction coefficient (L mol^{-1} cm^{-1}), λ_{em}–λ_{abs1} = Stokes Shift (nm).

Absorption and emission spectra are displayed in Figure 6.5. Absorption data has been truncated for samples run in DMF due to noise caused by loss of solvent transparency

below ~270 nm. Emission data has been scaled for easy visual comparisons between curves.

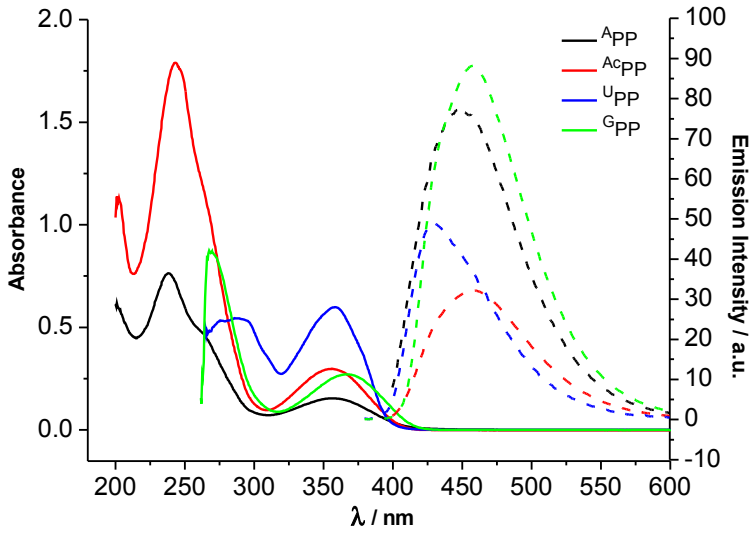


Figure 6.5. UV absorption (solid line) and fluorescence emission (dashed line) spectra for *N*-methylpyrrolo-dC nucleosides. Properties measured at different concentrations. $^{U}PP_p$ and $^{G}PP_p$ absorption data were truncated below ~260 nm. Emission data scaled by scale factor = 1 ($^{A}PP_p$), 0.1 ($^{Ac}PP_p$), 10 ($^{U}PP_p$), 0.6 ($^{G}PP_p$). Fluorescence spectra obtained by excitation at λ_{abs1} maxima (350-370 nm).

The urea, which demonstrated good binding affinity and selectivity, has the highest extinction coefficient (15140 L mol^{-1} cm^{-1}) by greater than a factor of 2. The Stokes shift, although lowest at 71.0 nm, is still very good. The acetamide, despite performing poorest in melting studies, has the highest Stokes Shift (102.5 nm), comparable with that of the 6-arylpyrrolo-C nucleobases **227-229** and PNA monomers **230,231** of Hudson *et al.*.[230,231] The guanidine demonstrated a good Stokes Shift (91.4 nm) and has the second highest extinction coefficient of 6690 L mol^{-1} cm^{-1}, thus may show most potential as a triplex probe. Additionally, all emission spectra maxima have minimal spectral overlap with their corresponding excitation maxima.

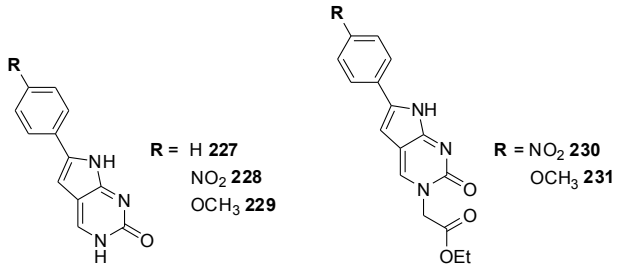


Figure 6.6. 6-Arylpyrrolo-C derivatives **227-231** of Hudson *et al.*.[230,231]

All nucleosides emit close to the excitation wavelength of DABCYL (471 nm), and therefore could be used instead of fluorescein (FAM) in fluorescence melting studies, or in combination with other acceptors for FRET analysis. There may also be a significant change in fluorescence emission on triplex hybridisation, analogous to that observed by Hudson in the duplex, on pairing with G.[120,233] 3'-Pyrene-modified dT has also found use as a probe, showing marked increase in fluorescence on triplex formation, depending on position within the triplex.[237] Fluorescence was strongly dependent on correct hybridisation, which offered the potential for detection of duplex mutations.

The quantum yields should be determined next, for absolute confirmation of the best *N*-methylpyrrolo-dC probe candidate.

6.6 Conclusions & Future Work

The final melting studies confirm the efficacy of these monomers to specifically recognise a CG inversion in a homopurine tract. Comparing the results from both studies, the monomers of choice should be the urea and the aniline. The aniline outperformed the other monomers in the UV melting study for both binding affinity and selectivity, and the urea also demonstrated its efficacy. The guanidinium monomer demonstrated the highest binding affinity in the fluorescence melting study, followed by the urea. However, the aniline proved slightly more selective and the urea less so on average, and these are easier to synthesise than the guanidinium. The acetamide performed relatively poorly throughout both studies, comparibly with T, except where placing pdU on the 5'-side caused destabilisation of the other monomers.

It is difficult to rationalise the differences between the two studies, which should be well correlated. Differences such as the change in order of stability of the pdU-

containing triplexes, and relative binding affinities and selectivity have not yet been explained, but several factors may contribute and there are some studies, which can be carried out to examine this. The three best monomers have been tested against real gene targets, and are currently being re-synthesised for further tests, in conjunction with other recognition monomers. Such mixed-sequence triplex-based recognition has already been evaluated,[67] and gene targeting studies are in progress.

The 6-phenyl-*N*-methylpyrrolo-dC phosphoramidite monomer **221** was synthesised by Dr. Edrees-Abdou, as detailed in Section 5.2, in excellent yield and melting studies are underway. Further modifications for assessment are described also in Section 5.2.

Finally, the fluorescence properties of the four CG recognition monomer nucleosides show potential for their use as probes within triple helices and in fluorescence studies, including melting. The Stokes Shifts and extinction coefficients are comparable with the NH-pyrrolo-C analogues of Hudson *et al.*,[230,231] which have already found use in fluorimetric detection of G.[120,233] Quantum yields must be determined, however, so the most efficient substrate can be chosen.

Chapter 7

Oligonucleotide Deprotection Study

7. Oligonucleotide Deprotection Study

7.1 Introduction[238]

The industry standard method for custom chemical oligonucleotide synthesis is the automated phosphoramidite or 'phosphite triester' cycle on solid-support, first developed by Prof. Marvin Caruthers *et al.* in 1981.[239,240] Modern automated DNA/RNA synthesisers can synthesise oligonucleotides on solid-support, from 2-250 nucleotides long, with stepwise coupling efficiencies of 98.0-99.5% for the standard DNA phosphoramidites, dA, dC, dG and dT. Oligonucleotides may be synthesised on a 40 nmole up to multi-millimole scale for therapeutic applications. Syntheses of up to 0.2 and 1.0 μmole are typical for research applications. The method, to be successful, requires a robust protecting group strategy due to the conditions, which the developing oligonucleotide is exposed to during synthesis. Equally, protecting groups must be labile under easily accessible deprotection conditions after synthesis, and ideally cleavage from the support should be facile and as near to quantitative as possible.

Standard solid-supports include controlled pore glass (CPG) and polystyrene-based resin/beads, to which the oligonucleotide is attached by a base-labile linker. The exocyclic amines in the standard nucleotides are protected with base-labile protecting groups, such as benzoyl and isobutyryl, which are stable during synthesis but easily removed afterwards (Figure 7.1).

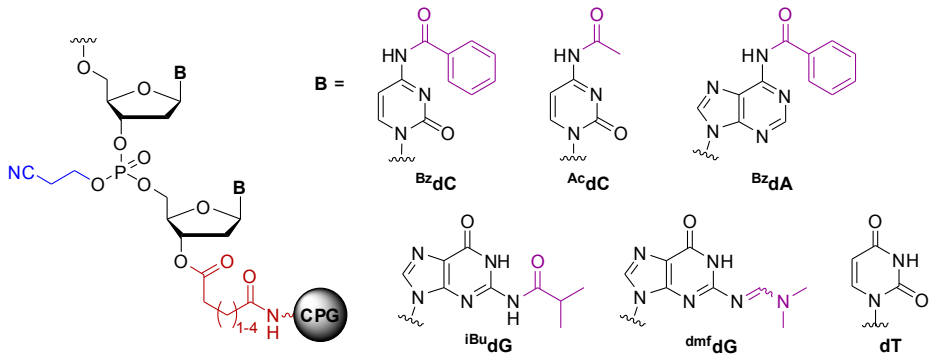


Figure 7.1. Protected oligodeoxynucleotide (ODN) on solid-support and standard DNA nucleobases, illustrating base-labile functionalities (succinyl linker – red; 2-cyanoethyl phosphate protecting group – blue; acyl protecting groups – violet). dT nucleobase requires no protection.

Non-standard nucleosides, either natural or synthetically modified, are incorporated by the same method, appropriately protected with base-labile or other labile groups. Alternative compatible protecting groups for primary/exocyclic amines include phthalimide, trifluoroacetyl and phenoxyacetyl. Hydroxyl groups may be protected as the TBDMS- or allyl-ether for example. The protecting group lability can be altered depending on the oligonucleotide sensitivity and desired deprotection conditions.

7.2 Deprotection Conditions for Oligonucleotides[238]

The standard reagents for deprotection of non-modified oligodeoxynucleotides (ODNs), are concentrated (~33 wt%) aqueous ammonia and concentrated (40 wt%) aqueous methylamine. The oligonucleotide may be cleaved from the solid-support under mild conditions (conc. aq ammonia, rt, 1 hour) before heating or further treatment, or the support may be removed from the synthesis column and all deprotections carried out in one step. Treatment at rt overnight or at 55 °C for 4-6 hours is usually sufficient to ensure full deprotection of shorter ODNs. Using a 1:1 mixture of conc. aq ammonia and conc. aq methylamine (AMA) may also be used to significantly reduce deprotection time for nucleosides such as ^{Ac}dC, ^{Bz}dA and ^{iBu}dG. AMA can also be used for quick cleavage of DNA from the solid-support (5-15 minutes). Methylamine and AMA should not be used for deprotection of ^{Bz}dC, however, as there is precedent for attack of the C^4-position by alkylamines, to form the 4*N*-alkyl-dC derivative. Oligonucleotides may also be deprotected using pressurised ammonia or methylamine gas, following cleavage from the support.

Standard deprotection conditions are also usually sufficient for deprotection of other base-labile groups. Phthalimide may be cleaved using methylamine or hydrazine, and allyl groups can be removed by automated treatment with palladium (0), such as a solution of $Pd(PPh_3)_4$ in chloroform, containing 5% acetic acid and 2.5% *N*-methylmorpholine,[241,242] before removal from the synthesiser for cleavage/deprotection. Oligonucleotides may also be treated with diethylamine before cleavage from the resin,[243,244] or by addition of nitromethane during deprotection.[245] These reagents cleave 2-cyanoethyl groups from the phosphate esters and scavenge acrylonitrile, a by-product of oligonucleotide synthesis. Nitromethane acts predominantly as a scavenger of acrylonitrile. Without such treatments, a small proportion of dT nucleotides are

cyanoethylated at the N^3-position. Such adducts however, do not form readily when deprotecting using methylamine, which also acts as a scavenger. Ammonia is not nucleophilic enough to act as an effective scavenger.

Washing of the solid-support with 20% diethylamine/acetonitrile is incorporated into the automated synthesis procedure on some new DNA/RNA synthesisers. 2-Cyanoethyl groups are cleaved and acrylonitrile scavenged before cleavage from the solid-support.[246] Diethylamine will only cause cleavage from solid-support upon prolonged exposure (>5 hours).

7.3 Oligonucleotide Deprotection Study

7.3.1 Background

Oligonucleotides synthesised for the initial UV and fluorescence melting studies (Chapters 4, 6) were analysed by analytical HPLC and either electrospray or MALDI-TOF MS. Analysis of the non-labelled furano-dT modified TFOs for UV melting, showed that after deprotection with 30 wt% aqueous methylamine (rt, 6-24 hours), and treatment with DOWEX-H^+ resin, cyclisation of the methylamine ring-opened intermediate was not quantitative (Figure 7.2). However, HPLC purification of these oligonucleotides was successful as resolution was obtained.

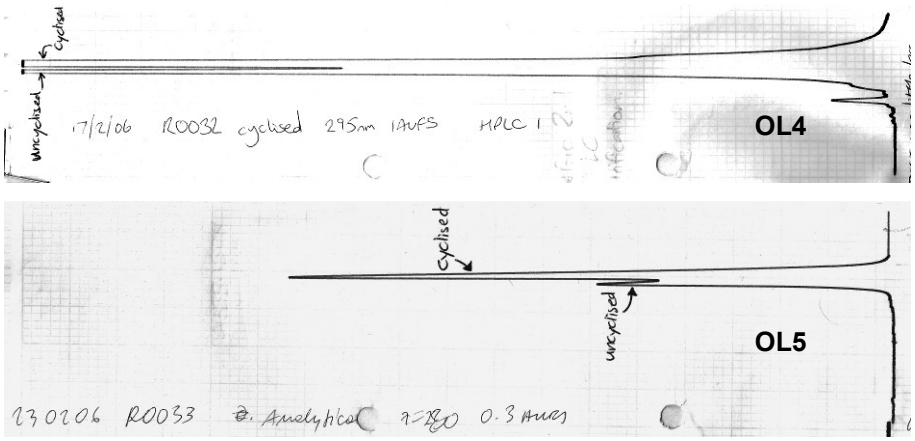


Figure 7.2. HPLC traces of UV melting TFOs (OL4,5) following first HPLC purification and treatment with DOWEX-H^+ resin.

HPLC analysis of the DABCYL-labelled TFOs for fluorescence melting, however, could not resolve the components. The large lipophilic DABCYL moiety causes the labelled DNA to adhere strongly to the stationary phase, thus causing reduction in resolution (”streaking”). MS analysis of several of the first DABCYL-labelled *N*-methylpyrrolo-dC modified TFOs (OL18-21, $^{Ac}PP_p$; OL23, $^{U}PP_p$), indicated they were clean following HPLC purification, and they were subsequently used in melting studies.

HPLC analysis of some of the crude mixtures before purification (OL18-21), identified a normal distribution of minor impurities around the major peak. Analytical ion-exchange (IONEX) chromatography of HPLC-purified TFO OL19, resolved a small possible failure sequence peak, and due to the long retention time was deemed unsuitable as a method of purification (Figure 7.3).

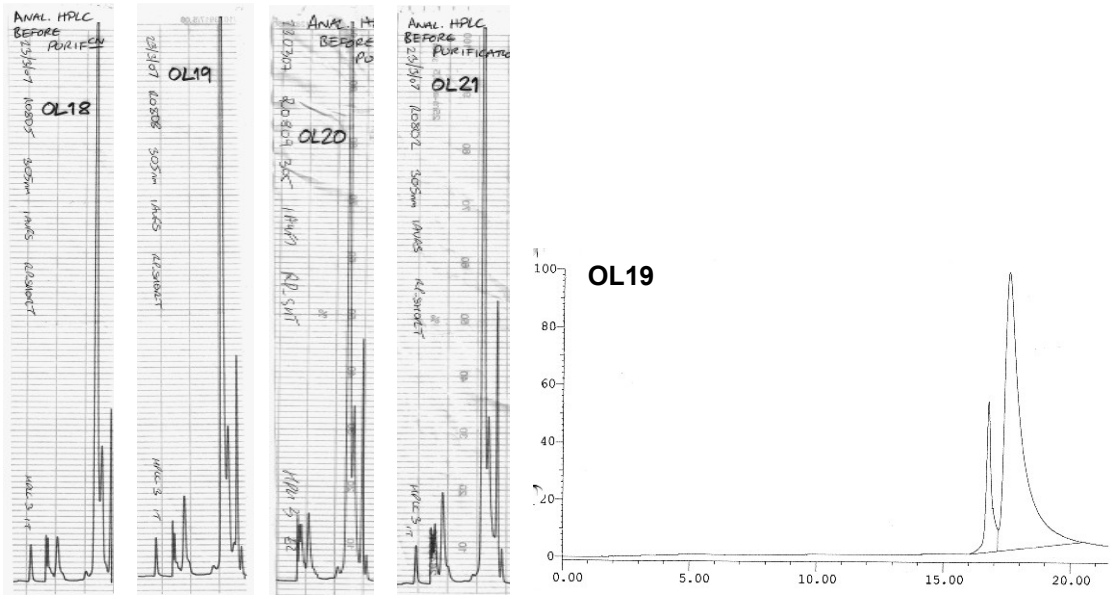


Figure 7.3. HPLC traces of $^{Ac}PP_p$-modified TFOs (OL18-21) before purification and IONEX trace of OL19 after HPLC purification.

Following this work, a Capillary Electrophoresis system was acquired, from which very high resolution chromatograms could be obtained. CE analysis of several of the DABCYL-modified TFOs, resolved each into three peaks for both furano-dT (OL14, $^{A}PP_f$) and *N*-methylpyrrolo-dC modified TFOs (OL20, $^{Ac}PP_p$; OL23, $^{U}PP_p$) (Figure 7.4). This result could be rationalised for the furano-dT modified TFO, as a mixture of cyclised (highest retention time), uncyclised/ring-opened TFOs and a sequence where one phosphoramidite coupling did not occur (failure sequence, lowest retention time). The observation is harder to explain for the *N*-methylpyrrolo-dC modified TFOs

(deprotected using conc. aq NH_3), where only ring-opening by ammonia or failure sequences could be held accountable. Failure sequences are usually capped by treatment with acetic anhydride, before the oxidation step in each cycle of oligonucleotide synthesis. If synthesised without capping, failure sequences, which are usually minor, may become more prominent, especially when using synthetically-modified nucleosides. However, this may account only for the smaller peak of lowest retention time, and ring-opening by ammonia appears unlikely. Secondary structures may also provide an explanation but are even more unlikely.

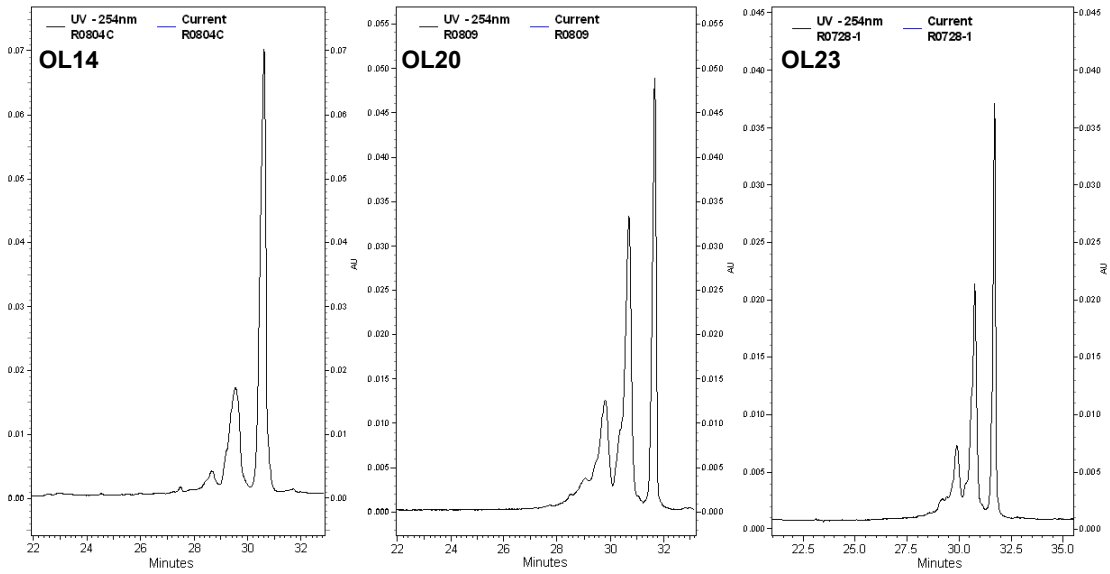
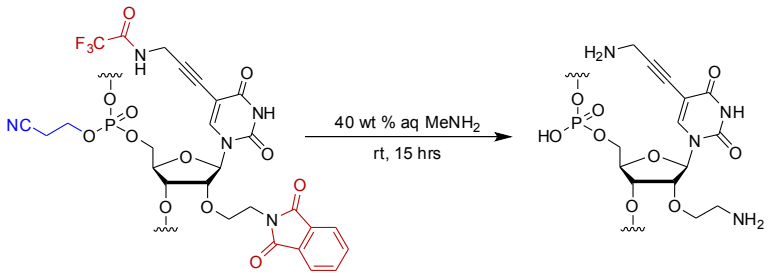


Figure 7.4. CE traces of DABCYL-modified TFOs (OL14, OL20, OL23).

Concurrently there were problems with TFOs containing the GC-recognition monomer, *bis*-amino-U (^{BA}U). Following incorporation into the TFO, the oligonucleotide was deprotected using 40 wt% aqueous methylamine. Methylamine was also required for cleavage of the phthalimide protecting group on the 2'-moiety (Scheme 7.1). Despite previous successful deprotections and melting studies,[35-37] problems were later encountered. Although HPLC analysis of the TFOs indicated they were pure, melting studies afforded poor results, i.e. little or no triplex formation. On receipt of the CE machine, analysis of some of these TFOs revealed a mixture of peaks, which could only be accounted for by excessive degradation of the TFO, not by failure sequences alone. This suggested that the deprotection conditions must have changed since the previous TFOs containing this monomer were synthesised, deprotected and purified.



Scheme 7.1. ^{BA}U nucleotide deprotection in TFO.

After these results, it was decided to investigate oligonucleotide deprotection conditions to determine the optimum conditions for monomers used within the research group.

7.3.2 Deprotection Study

A series of oligonucleotides were synthesised on a 1.0 μmole scale (with capping) based on oligoT (T_{12}), containing the monomers under examination (Figure 7.5).

TTT TTT TT**X** TTT TTT T**Y**T **Y**T**Y** TTT T**ZZ** **Z**T**Z** **Z**T**Z** T**Z**T

Figure 7.5. Sequences for oligonucleotide deprotection study; **X** = T (OL64), pdU (OL65), ^{BA}U (OL68), $^{Ac}PP_p$ (OL69), S (OL70), ^{G}pdU (OL74), ^{MOE}pU (OL71), ^{AE}T (OL72); **Y** = pdU (OL66); **Z** = pdU (OL67).

After each oligonucleotide synthesis, the solid-support was removed from the column. Small portions of the resin (5-8 mg) were deprotected for varying lengths of time, at different temperatures using conc. aq ammonia, 20 wt% aq $MeNH_2$ and 40 wt% aq $MeNH_2$. A list of all deprotections for each oligonucleotide is given in Table 7.1.

Following deprotection, the resin was removed and solvent evaporated, the oligonucleotide residue was dissolved in distilled water (1.0 mL) and desalted using a disposable Sephadex NAP™ 10 column. All samples were then analysed, without HPLC purification, by CE. ES and MALDI-TOF MS analyses were taken as appropriate. Analytical HPLC and IONEX were also used for several samples.

ODN	X	Sequence	Deprotection Conditions				MS	
			Conc[n]	Base	Temp	Time/h	ES	MALDI
OL64	–	TTT TTT TTT TTT	conc	NH_3	55 °C	5		
			20%	$MeNH_2$	rt	15		
			40%	$MeNH_2$	rt	15		
OL65	pdU	TTT TTT TT**X** TTT	**conc**	**NH_3**	**rt**	**5**	✓	✓
			conc	NH_3	rt	15	✓	
			conc	NH_3	55 °C	5		
			conc	NH_3	55 °C	15	✓	
			20%	$MeNH_2$	rt	15		
			40%	$MeNH_2$	rt	15	✓	✓
OL66	pdU	TTT T**X**T **X**T**X** TTT	conc	NH_3	rt	1	✓	✓
			****conc**	**NH_3**	**rt**	**1**		
			conc	NH_3	rt	5		
			conc	NH_3	rt	15		
			conc	NH_3	55 °C	4	✓	✓
			conc	NH_3	55 °C	15	✓	
			20%	$MeNH_2$	rt	5		
			20%	$MeNH_2$	rt	15	✓	
			40%	$MeNH_2$	rt	15		
OL67	pdU	T**XX** **X**T**X** **X**T**X** T**X**T	conc	NH_3	55 °C	1	✓	
			****conc**	**NH_3**	**55 °C**	**1**	✓	
OL68	^{BA}U	TTT TTT TT**X** TTT	conc	NH_3	55 °C	15	✓	✓
			conc	AMA	rt	15		
			20%	**$MeNH_2$**	**rt**	**5**	✓	✓
			20%	$MeNH_2$	rt	8		
			20%	$MeNH_2$	rt	15	✓	
			20%	$MeNH_2$	rt	30	✓	✓
OL69	$^{Ac}PP_p$	TTT TTT TT**X** TTT	**conc**	**NH_3**	**rt**	**5**	✓	✓
			conc	NH_3	55 °C	4	✓	
			20%	$MeNH_2$	rt	5		
			40%	$MeNH_2$	rt	6	✓	
			40%	$MeNH_2$	rt	15	✓	✓
OL70	S base	TTT TTT TT**X** TTT	**conc**	**NH_3**	**rt**	**5**	✓	✓
			conc	NH_3	rt	15	✓	
			conc	NH_3	55 °C	5	✓	
			conc	NH_3	55 °C	15	✓	✓
			20%	$MeNH_2$	rt	5		
			40%	$MeNH_2$	rt	15	✓	
OL71	^{MOE}pU	TTT TTT TT**X** TTT	**conc**	**NH_3**	**rt**	**5**		
			conc	NH_3	55 °C	5		
			20%	$MeNH_2$	rt	5		
OL72	^{AE}T	TTT T**X**T **X**T**X** TTT	**conc**	**NH_3**	**55 °C**	**4**		
OL73	^{AE}T	TTT TTT TT**X** TTT	**conc**	**NH_3**	**55 °C**	**4**		
OL74	^{G}pdU	TTT TTT TT**X** TTT	conc	NH_3	rt	5		
			$^{t}BuNH_2$:MeOH:H_2O (1:1:2)		rt	21		

Table 7.1. Oligonucleotides, deprotection conditions and MS analysis. Optimum conditions evaluated are in bold. **Column washed with 20% Et_2NH/MeCN for 10 minutes before removal of resin for deprotection.

7.3.2.1 Control T_{12} Oligonucleotide

Control oligonucleotide T_{12} (OL64) was deprotected successfully by three methods. No degradation was observed by CE, hence even the strongest deprotection conditions (40 wt% aq $MeNH_2$) caused no degradation of the DNA backbone (Figure 7.6).

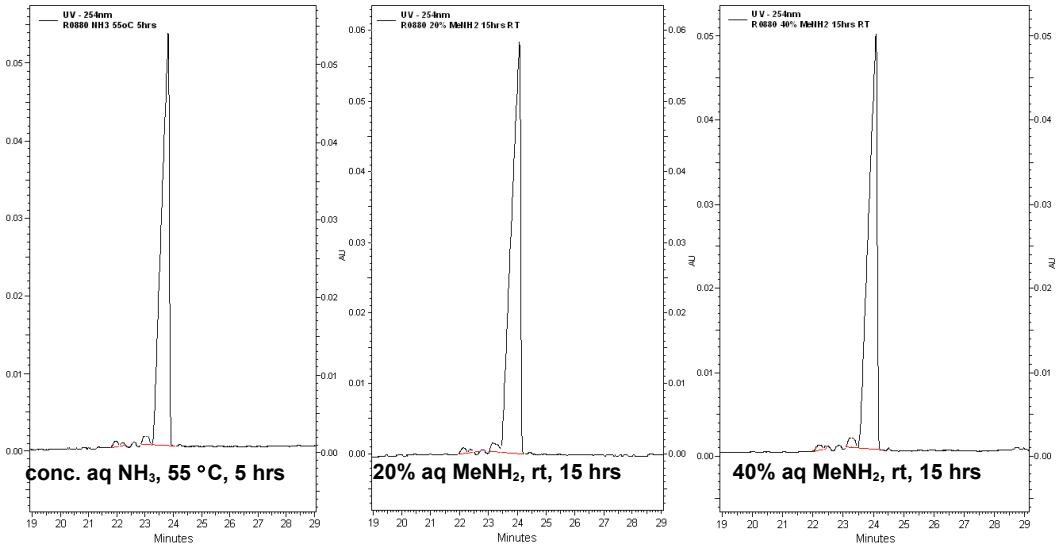


Figure 7.6. CE traces for deprotections of T_{12} (OL64).

It was deduced therefore, that any impurities observed for other oligonucleotides must be due to failure sequences (n-1, n-2 etc.), degradation of the modified nucleotide, or fragmentation of the oligonucleotide due to the modification.

*7.3.2.2 Test Oligonucleotides – 5-(3-Aminopropynyl)-dU (*pdU*)*

The first modified monomer to be tested, pdU, was incorporated into three sequences with one, three and seven insertions.

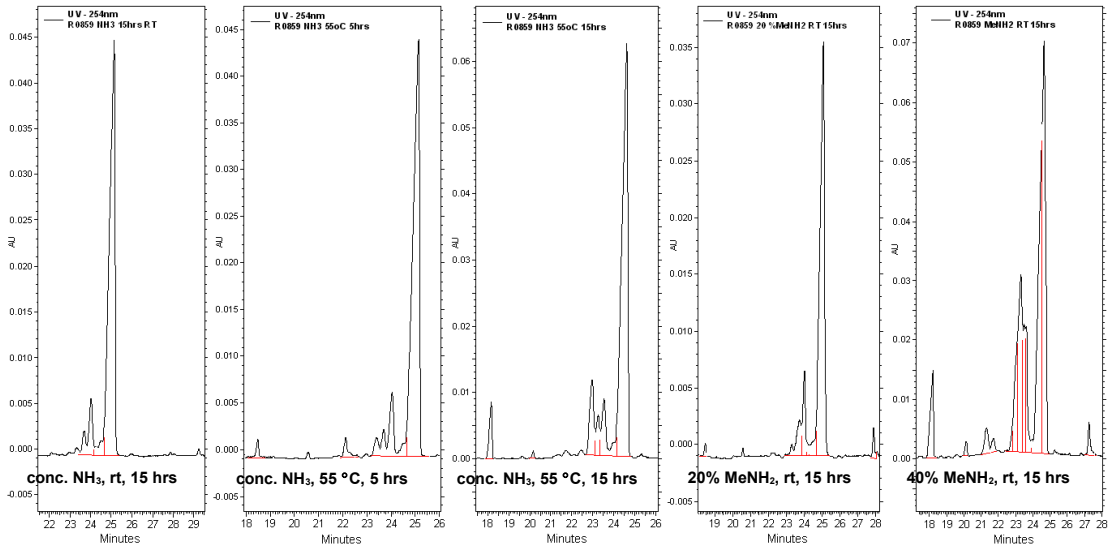
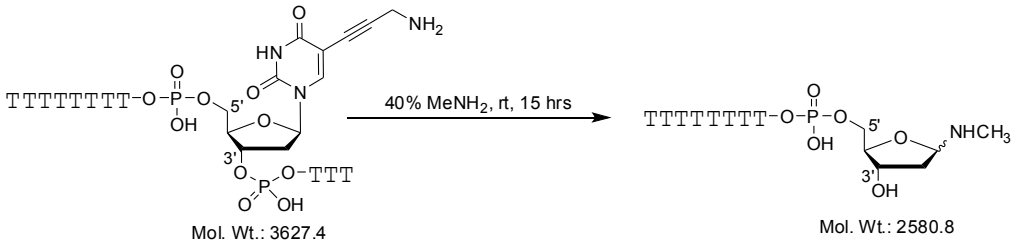


Figure 7.7. CE traces for deprotections of $T_8\mathbf{X}T_3$ (**X** = pdU, OL65).

The best conditions for $T_8\mathbf{X}T_3$ (**X** = pdU, OL65), giving the least degradation, are conc. aq ammonia at rt for 5 hours (Figure 7.7). A longer time does not significantly affect the impurity profile. Heating at 55 °C for 5 hours is also suitable, but heating overnight (15 hrs) caused further appearance of impurities. A 20 wt% aqueous solution of methylamine also deprotected causing little degradation overnight.



Scheme 7.2. Proposed cleavage/degradation of $T_8\mathbf{X}T_3$ by methylamine at high concentration.

However, it is very clear that using conc. aq methylamine (40 wt%) causes formation of a significant impurity (2581 Da) plus other minor peaks. This major impurity can be explained by cleavage of the oligonucleotide at the 3'-side of pdU, and displacement of the alkynyl-U nucleobase by methylamine (Scheme 7.2, Figure 7.8).

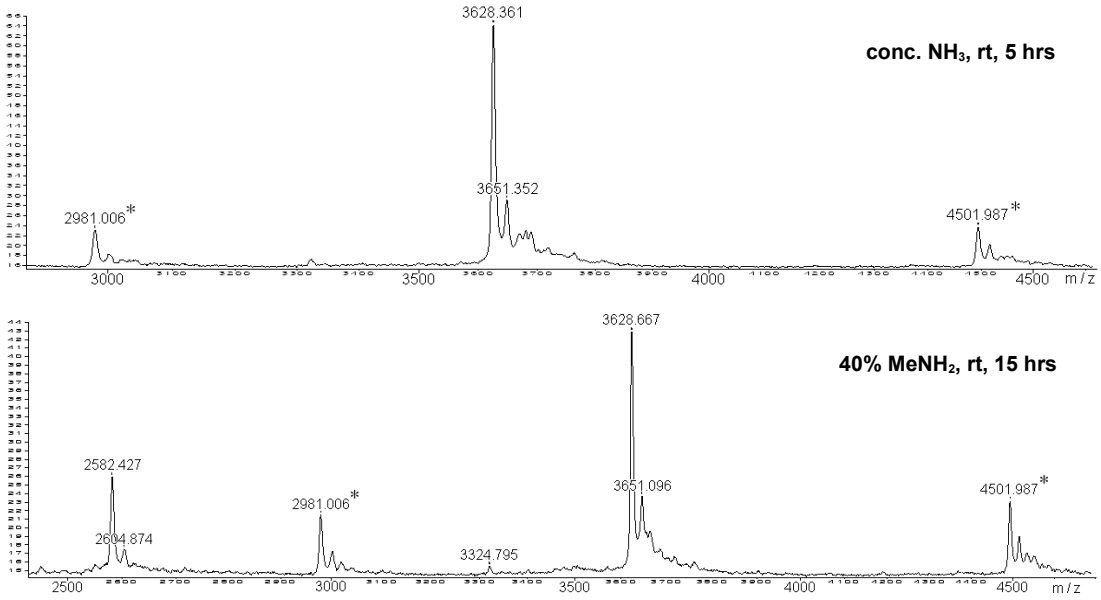


Figure 7.8. MALDI-TOF MS spectra for two deprotections of $T_8\mathbf{X}T_3$ (**X** = pdU, 3627.4 Da, OL65). * Internal T_{10} and T_{15} references.

Similar observations were made for three additions of pdU ($T_4\mathbf{X}T\mathbf{X}T\mathbf{X}T_3$, OL66). The mildest successful deprotection conditions were treatment with conc. aq ammonia at rt for 1 hour only, after washing the resin with 20% Et_2NH/CH_3CN for 10 minutes. Without the diethylamine wash, impurities were present. For this sample, this was due either to poor coupling of pdU, perhaps caused by presence of pyridine in phosphoramidite, or too little time for complete deprotection, not the lack of a diethylamine wash. Under these conditions, 1 hour is the shortest possible time for cleavage from the resin, but may not be long enough for consistent deprotection of all protecting groups (Figures 7.9, 7.10).

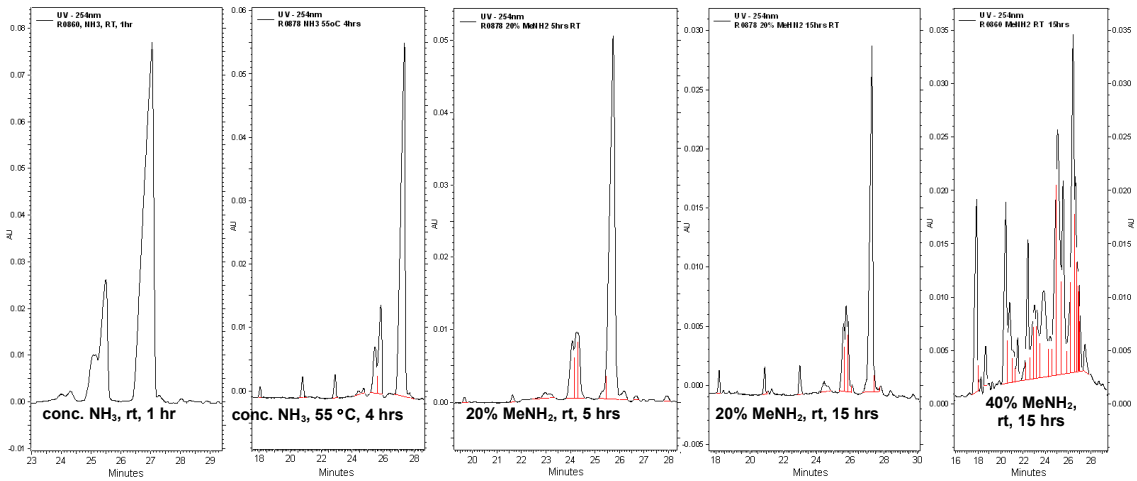


Figure 7.9. CE traces for deprotections of $T_4\mathbf{X}T\mathbf{X}T\mathbf{X}T_3$ (**X** = pdU, OL66).

Heating at 55 °C for only 4 hours with conc. aq ammonia formed several small impurities, and treatment overnight at rt with 40 wt% methylamine caused extensive oligonucleotide degradation. Treatment with 20 wt% methylamine at rt could be used for 5 hours, but impurities formed over 15 hours.

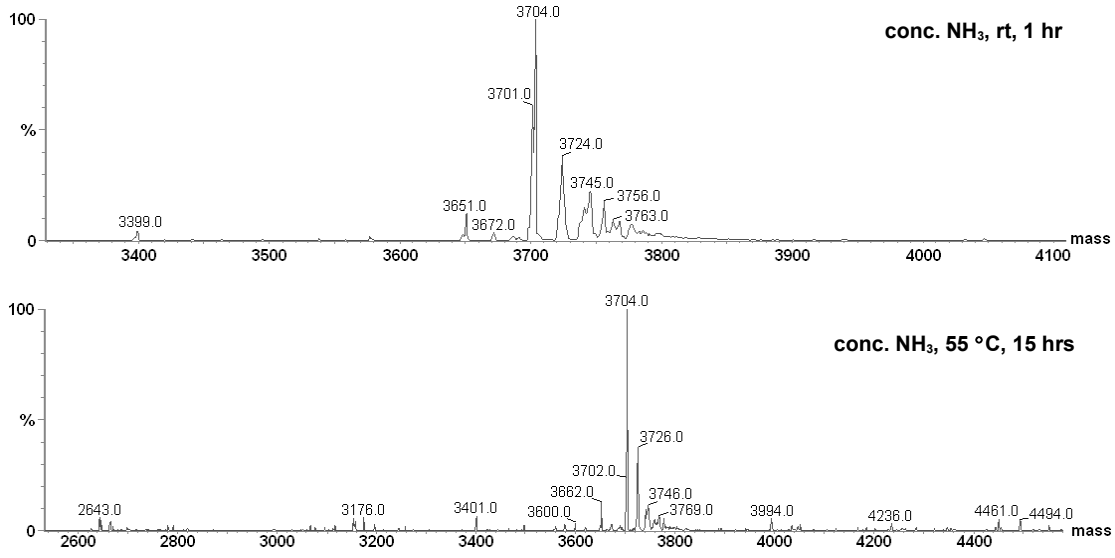


Figure 7.10. ES MS spectra for two deprotections of T_4**X**T**X**T**X**T_3 (**X** = pdU, 3705.5 Da, OL66). Data deconvoluted to resolution of 1 Da.

Oligonucleotide T**X**$_3$T**X**$_2$T**X**T**X**T (X = pdU, OL67) was successfully deprotected under mild conditions (conc. aq ammonia, 55 °C, 1 hr) with and without a prior diethylamine wash. The small impurities were slightly greater without washing and failure sequences became slightly more prominent. A slightly longer deprotection time (~2 hours) may have removed some of these trace impurities. HPLC and IONEX traces were similar for both (Figure 7.11).

A small impurity was present in these oligonucleotides (3651 Da in OL66), due to incorporation of a single dU nucleoside instead of pdU. The dU impurity was formed during the Sonogashira cross-coupling reaction in synthesis of the pdU phosphoramidite monomer **164** and was present at approximately 2-3%. When used in synthesis of the deprotection study oligonucleotides, it only became noticeable on multiple additions. When oligonucleotide OL66 was re-synthesised using a different batch of monomer, with and without capping for comparison, only trace baseline impurities were noticeable. Synthesis was as effective with capping as without capping indicating absence of failure sequences.

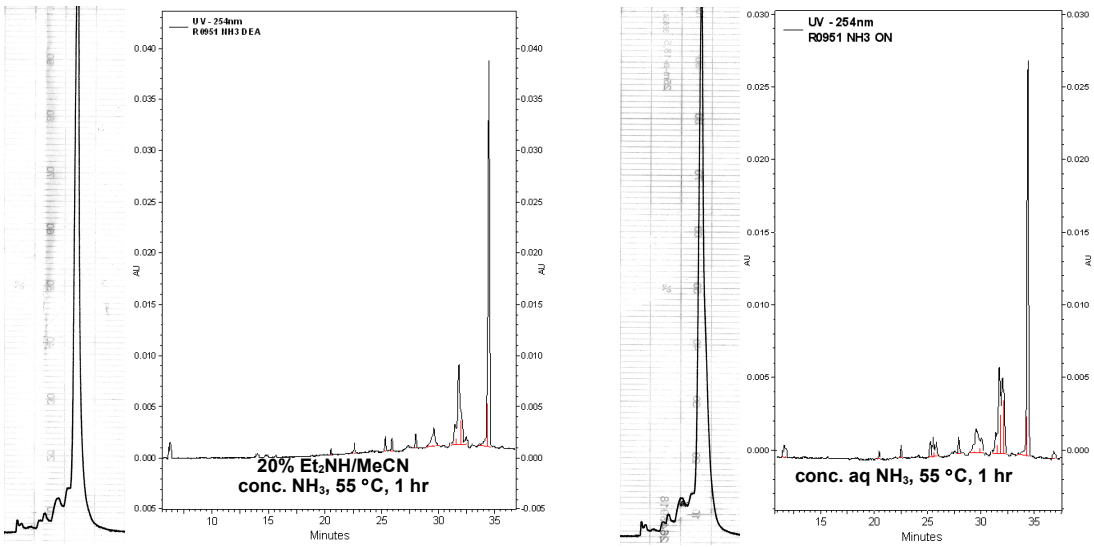


Figure 7.11. HPLC/CE traces for deprotections of T**X**$_3$T**X**$_2$T**X**T**X**T (**X** = pdU, OL67).

7.3.2.3 Test Oligonucleotides – Bis-amino-U (^{BA}U)

The next monomer under test was *bis*-amino-U (^{BA}U). The best conditions for deprotection were discovered to be 20 wt% aq methylamine, at rt for 5 hours.

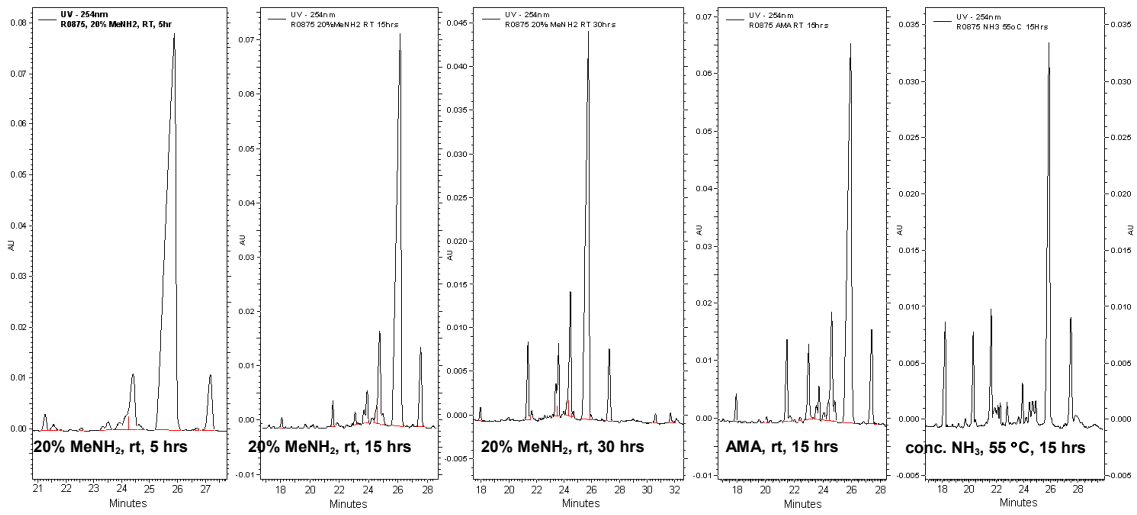


Figure 7.12. CE traces for deprotections of T$_8$**X**T$_3$ (**X** = ^{BA}U, OL68).

Little change was noted after 8 hours, but after 15 and 30 hours notable impurities had formed. Heating at 55 °C in conc. aq ammonia for 15 hours, and treating with AMA (conc. aq NH_3/40 wt% aq $MeNH_2$, 1:1 v/v) at rt for 15 hours produced a similar significant pattern of impurities (Figure 7.12).

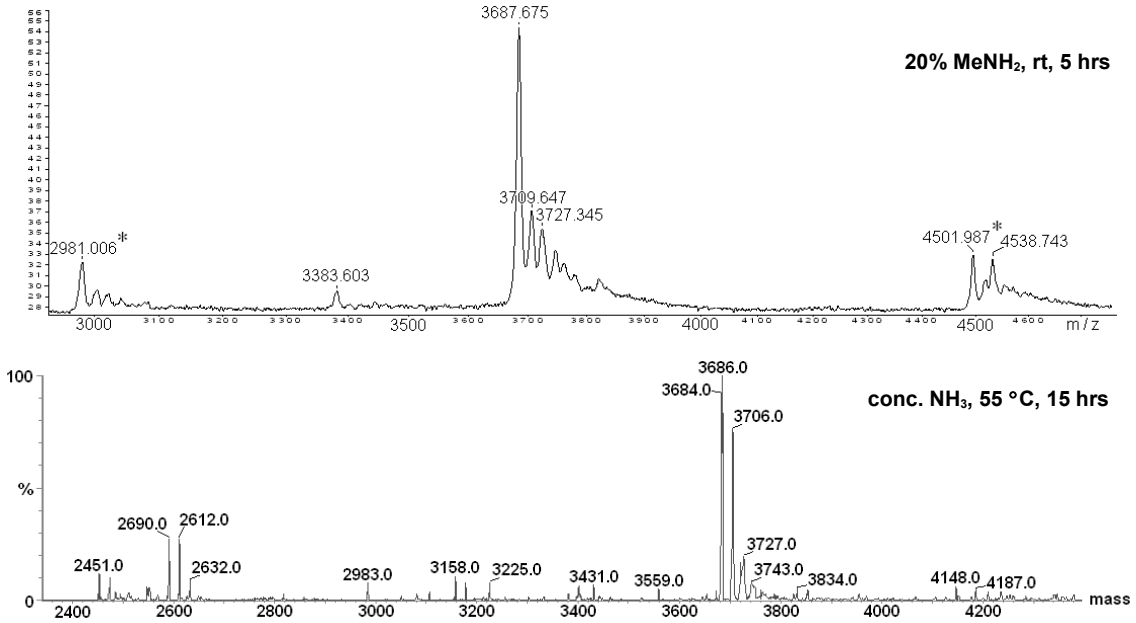


Figure 7.13. MALDI-TOF/ES MS spectra for two deprotections of $T_8\mathbf{X}T_3$ (**X** = ^{BA}U, 3686.5 Da, OL68). ES data deconvoluted to resolution of 1 Da. * Internal T_{10} and T_{15} references.

Only a 'minus T' failure sequence (3382.3 Da) was identified by MS. The other MS peaks could not be assigned (Figure 7.13). Methylamine is required for phthalimide cleavage, but a high concentration causes significant degradation.

7.3.2.4 Test Oligonucleotides – 6-(3-Acetamidophenyl)-N-methylpyrrolo-dC ($^{Ac}PP_p$)

The acetamide monomer for CG recognition was incorporated into a 'T_{12} oligonucleotide' ($T_8\mathbf{X}T_3$, **X** = $^{Ac}PP_p$, OL69) and proved stable to conc. aq ammonia at rt for 5 hours, and 55 °C for 4 hours, and also to 20 wt% aq methylamine at rt for 5 hours (Figure 7.14). Heating at 55 °C in conc. aq ammonia for 15 hours caused more significant degradation than AMA at rt for the same time.

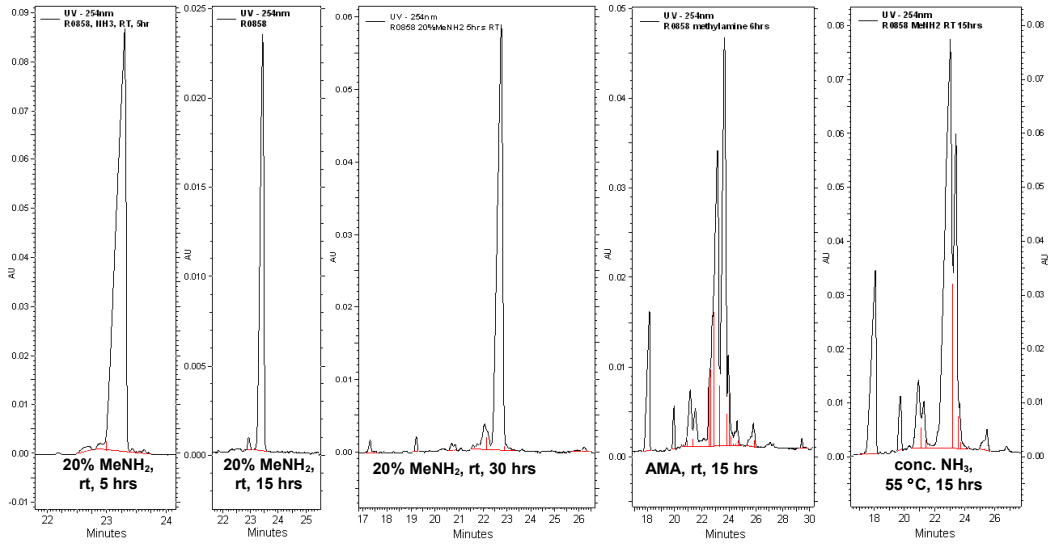


Figure 7.14. CE traces for deprotections of $T_8\mathbf{X}T_3$ (**X** = $^{Ac}PP_p$, OL69).

Concentrated aq methylamine, however, primarily caused major conversion to the cleavage product (2581 Da) shown in Scheme 7.2 (Figure 7.15).

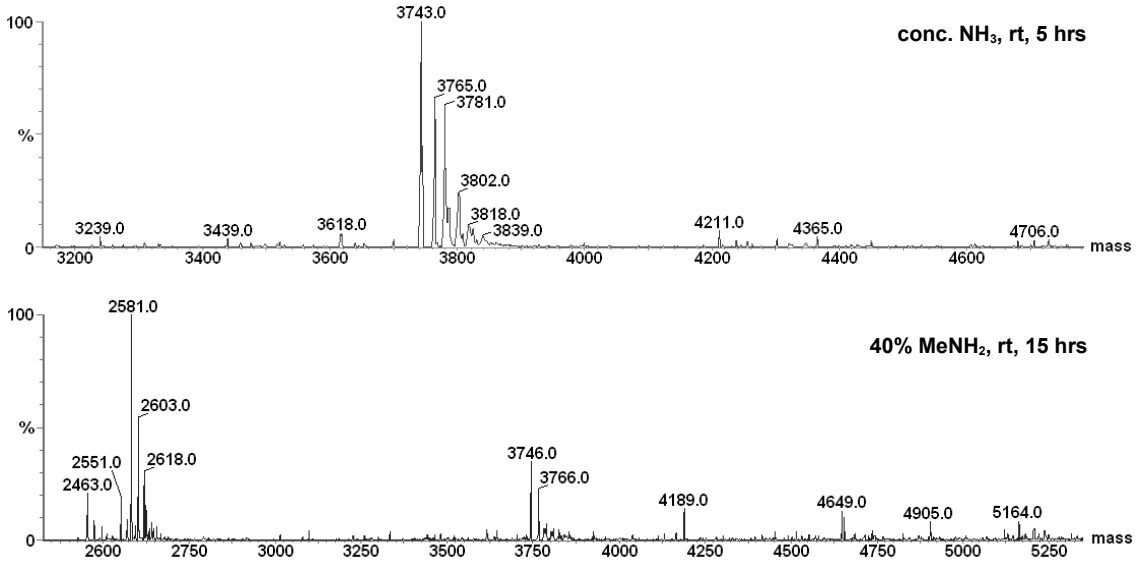


Figure 7.15. ES MS spectra for two deprotections of $T_8\mathbf{X}T_3$ (**X** = $^{Ac}PP_p$, 3744.6 Da, OL69). Data deconvoluted to resolution of 1 Da.

This was originally presumed to be the deacetylated oligonucleotide, but was confirmed otherwise by ES and MALDI-TOF MS. No ring-opening of the *N*-methylpyrrolopyrimidine nucleobase by methylamine was observed.

7.3.2.5 Test Oligonucleotides – S base (TA recognition monomer, S*)*

The 'S base' monomer for TA recognition (Section 1.2.2.1), was incorporated into oligo OL70 (T_8**X**T_3, **X** = S) for deprotection studies. The main aim for this modification was to find conditions, to which the acetamide moiety would be stable. CE analysis produced only a single peak for each sample, but was unable to resolve acetylated and deacetylated products.

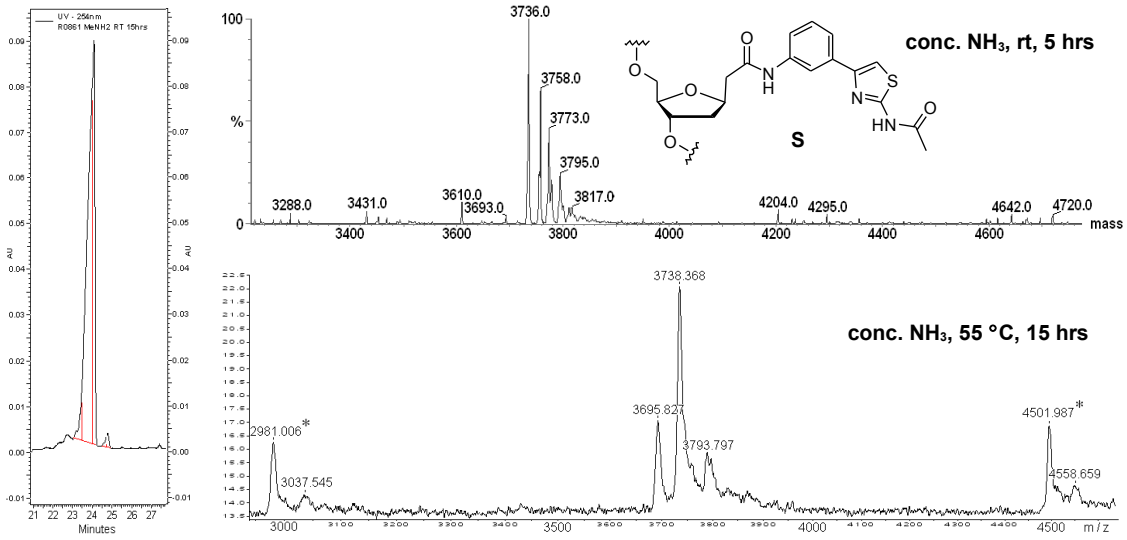


Figure 7.16. CE trace for strongest deprotection conditions (40% aq methylamine, rt, 15 hrs), and ES/MALDI-TOF MS spectra for two deprotections of T_8**X**T_3 (**X** = S, 3737.6 Da, OL70). Deacetylated oligonucleotide mass (3695.6 Da). ES data deconvoluted to resolution of 1 Da. * Internal T_{10} and T_{15} references. Structure of 'S' is shown.

MS analysis highlighted stability in conc. aq ammonia at rt for up to 15 hrs. Heating at 55 °C in conc. ammonia for only 5 hours caused some deacetylation, and in 40 wt% aq methylamine at rt, a significant amount (Figure 7.16). CE analysis of the sample deprotected with 20% aq methylamine (rt, 5 hrs) showed a very sharp peak, but conc. ammonia was concluded as the best deprotection medium.

7.3.2.6 Test Oligonucleotides – 5-(3-aminopropynyl)-2'-O-methoxyethyl-U (^{MOE}pU*)*

AT recognition monomer, ^{MOE}pU, was incorporated into sequence T_8**X**T_3 (OL71) for comparison with its 2'-unmodified analogue, pdU. Poor coupling of the monomer, however, resulted in a major failure sequence (T_{11}) of similar proportion to product by

CE. Of the three conditions tested (Table 7.1), heating at 55 °C in conc. aq ammonia for 5 hrs caused an increase in the apparent failure sequence peak and a small increase in the general impurity profile over the other two.

7.3.2.7 Test Oligonucleotides – Other AT recognition monomers

AT recognition monomer, 2'-aminoethoxy-T (^{AE}T) was incorporated into T$_8$**X**T$_3$ (OL72) and T$_4$**X**T**X**T**X**T$_3$ (OL73) and both were deprotected at 55 °C in conc. aq ammonia for 4 hours. Both samples produced very clean HPLC traces hence this method was deemed suitable, and this should be applicable also at rt.

5-(3-Guanidinylprop-1-ynyl)-dU (^{G}pdU) was incorporated once (T$_8$**X**T$_3$, OL74) and treated under two deprotection conditions. Deprotection with conc. aq ammonia (rt, 5 hrs) caused formation of two major peaks plus impurities by CE (Figure 7.17).

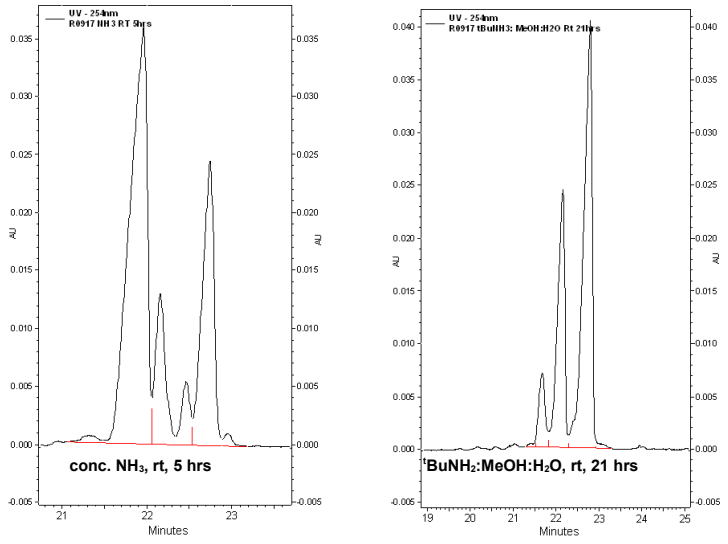


Figure 7.17. CE traces for deprotections of T$_8$**X**T$_3$ (**X** = ^{G}pdU, OL74).

One major peak (~22.8 minutes) appears to correspond to the oligonucleotide with a fully-protected guanidinyl moiety. The other (~21.9 minutes) may correspond to a stable triazinedione derivative, caused by cyclisation due to attack by ammonia. This was observed by Prakash *et al.*[225] when deprotecting oligonucleotides, containing 2-cyanoethyloxycarbonyl- (CEOC) protected guanidinyl moieties, with ammonia (Figure 7.18). Prakash *et al.* cleaved the CEOC groups in 50% aq piperidine at rt for 24 hours, prior to heating at 55 °C in conc. ammonia. Treating with 50% aq piperidine also

cleaved the 2-cyanoethyl phosphate protecting groups and cleaved the oligonucleotide from the resin.

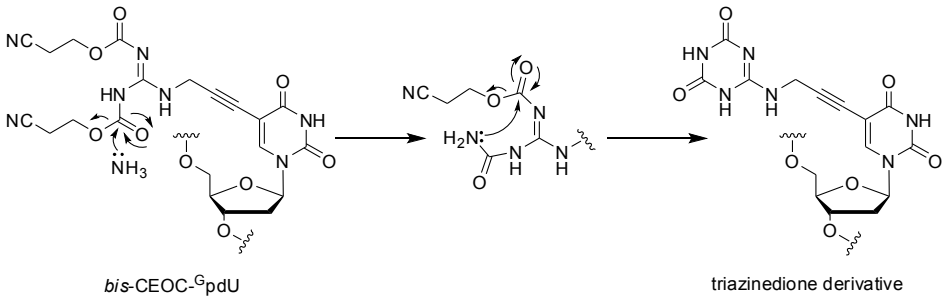


Figure 7.18. Possible mechanism of formation of triazinedione derivative.[225]

It was later discovered that another secondary amine, diethylamine, could be used instead of piperidine. Washing the column with 20% Et_2NH/CH_3CN for at least 30 minutes prior to deprotection, was sufficient to remove the most labile CEOC group thus preventing triazine formation. It is likely that methylamine would attack the CEOC groups in the same manner but the steric bulk of the methyl group might prevent cyclisation after attack.

The second deprotection of OL74 was carried out in $^tBuNH_2:CH_3OH:H_2O$ (1:1:2 v/v/v) at rt for 21 hours. *Tert*-butylamine was used to cleave all protecting groups at the same time. It was presumed that the sterically-hindering *tert*-butyl group would prevent attack of the CEOC groups and cyclisation. These conditions, however, proved too mild, producing a mixture of *bis*-CEOC, *mono*-CEOC and fully deprotected oligonucleotides in an approximate ratio of 9:5:1 (Figure 7.17).

7.4 Conclusion

The results of this study highlight the need for suitable/compatible deprotection conditions for each modified nucleotide within the oligonucleotide. It cannot be assumed that the conditions used for one modification, will be safe to use for another, as proven by the compatibility of most modifications to deprotection with conc. ammonia at rt, except the CEOC-protected guanidinyl-modified nucleotide.

Concentrated (40 wt%) aq methylamine was proven too strong for deprotection of all oligonucleotides containing modified non-natural nucleotides. It is likely that previous deprotections of ^{BA}U-containing oligonucleotides using conc. aq methylamine were successful, due to evaporation of methylamine over time from the bottle in use, thus reducing the effective concentration.

In general, conc. aq ammonia can be used to deprotect most modified oligonucleotides, either at rt or sometimes with heating. Methylamine is needed for cleaving phthalimide groups, and may be required for cleavage of more stable protecting groups.

Washing the column with 20% Et_2NH/CH_3CN solution for 10-20 minutes before removal from the synthesiser, was incorporated into the standard oligonucleotide synthesis protocol, on all synthesisers where it was not already included and could be easily achieved. It cleanly cleaves 2-cyanoethyl groups and scavenges acrylonitrile, and proves highly effective at removing CEOC-protecting groups on guanidinyl-modified nucleotides.

These results were used for choosing deprotection conditions for oligonucleotides used within the research group, including for the TFOs used in the primary melting studies (Sections 6.2, 6.3). For all TFOs used in these melting studies, following synthesis, the solid-support column was washed with 20% Et_2NH/CH_3CN for 20 minutes, or 3 hours for $^{G}PP_p$-containing TFOs (OL48-51, OL56-59). Each oligonucleotide was then deprotected in conc. aq ammonia at rt for 12 hours, with or without 2 hours at 50 °C, or at rt for 24 hours followed by heating at 50 °C for 1 hour to ensure complete deprotection. CE analysis indicated good to high purity for most TFOs after one HPLC purification. Where impurities constituted a significant proportion, the TFO was carefully repurified, collecting only the centre of the peak. Impurities did not, however, hinder performance in melting studies.

Chapter 8

Conclusion

8. Conclusion

8.1. Modified 6-Oxocytidine Nucleosides for GC Recognition

Synthesis of a *bis*-amino-modified 6-oxocytidine monomer **64**, was re-investigated as a potential GC recognition monomer, which exhibits pH-independent binding within a triplex. Previous published studies[70,71] reported pH-independent triplex formation using $^{6\text{-oxo}}$C nucleotides up to pH 8, and enhanced binding to GC at higher pH compared to C. Modifications to this nucleoside,[11,70,76] however, produced varied or negative results, hence further modifications were proposed to introduce protonated amine functionalities, which should enhance triplex stability.

Several synthetic investigations were carried out, in order to assess reaction feasibility. Palladium-catalysed alkynylation was ruled out as a method of introduction of amine functionality at the 5-position due to inactivity of the halogenated substrate or degradation. Modification by amine- or thiol-halide exchange was investigated with limited success, but this requires extensive optimisation if a robust synthetic route is to be found. Feasibility of 2'-modification was also investigated, although this requires further work. Optimised synthetic routes to some key intermediates have been developed.

8.2 Amino-Modified Bicyclic Nucleosides for CG Recognition

Recognition of pyrimidine.purine base pairs is a greater challenge than purine.pyrimidine base pairs as fewer hydrogen-bonds are presented for binding in the major groove. Previous published research describes the development of recognition monomers from T, based around the bicyclic C analogue, *N*-methylpyrrolo-dC.[122] This nucleobase retains the C^2 carbonyl group, vital for CG recognition, whilst blocking GC recognition by replacing an H for a methyl group at the 4*N* position.

The latest design had employed an aminoalkyl group for targeting of hydrogen-bonding residues on G across the base pair. This research examined various *meta*-aminophenyl groups; rigidified linkers better placed for binding.

Furano-dT **145-147** and *N*-methylpyrrolo-dC phosphoramidite monomers **178,179,191** (**199** – synthesis repeated later in house) were synthesised for assessment. All were synthesised by a common synthetic route (Sonogashira cross-coupling, Cu(I)-catalysed cyclisation and phosphitylation) with some modifications, starting from DMT-protected 5-iodo-2'-deoxyuridine **97**, and analogue 5-iodo-*N*-methyl-2'-deoxycytidine **185** respectively. Yields ranged from moderate to excellent and most steps were optimised despite high substrate polarity.

Furano-dT monomers underwent post-synthetic modification to the *N*-methylpyrrolo-dC nucleotide after incorporation into DNA. *N*-Methylpyrrolo-dC monomers could be incorporated with no further modification necessary. Although post-synthetic modification was not complete, and HPLC purification was unable to resolve the mixture for DABCYL-labelled oligonucleotides, both performed equally in melting studies.

Thiourea-modified monomer **158** (synthesis finished later in house) could not be incorporated into TFOs due to issues with stability to oligonucleotide synthesis or deprotection, hence could not be assessed in biophysical studies.

Extensive melting studies demonstrated the effectiveness of these monomers at recognising CG with high selectivity and good binding affinity compared to previous monomers (T, ^{4H}T, ^{M}P). The order of stability was altered, however, between UV and fluorescence melting studies. This was proposed to be due to changes in target sequence and local environment around the CG inversion target, but requires further investigation.

The effect of neighbouring groups on the modification was also highlighted. The CG recognition monomers appear to interact differently with 5-(3-aminoprop-1-ynyl)-dU (pdU, P), depending on which side it is located. This neighbouring group interaction introduces important implications when using TFOs as oligotherapeutics, as all interactions must be taken in to account for successful development.

8.3 Oligonucleotide Deprotection

A study was conducted to determine the optimum deprotection conditions for monomers used within the research group, after experiencing several problems.

Concentrated aq ammonia was determined the best medium for deprotection at rt or sometimes with heating. Most monomers also tolerated 20% aq methylamine for approximately 5 hours. Deprotection times of 15 hours or greater caused some degradation for several monomers, and the standard deprotection media, conc. aq methylamine and AMA, as used for non-modified DNA, caused significant to extensive degradation and strand cleavage for everything except the oligonucleotide T_{12}.

A column wash with 20% Et_2NH/CH_3CN solution was introduced on to DNA synthesisers where not already incorporated, after demonstrating efficient cleavage of 2-cyanoethyl phosphate protecting groups and prevention of cyanoethylation by scavenging of acrylonitrile. This wash also proved very effective at partial deprotection of the guanidine moiety of two monomers, thus avoiding subsequent side-reactions during deprotection.

Chapter 9

Experimental

9. Experimental

9 Synthesis

10 General

HPLC grade acetonitrile, methanol and water were purchased from Fisher Scientific. Pyridine, dichloromethane, triethylamine and diisopropylethylamine were distilled over calcium hydride; methanol, ethanol and isopropanol over magnesium and iodine; phosphorous oxychloride over sodium wire, and tetrahydrofuran over sodium wire and benzophenone before use. Anhydrous *N,N*-dimethylformamide and 1,2-dichloroethane were purchased from Aldrich. Diethyl ether was purchased from Fisher Scientific and dried by passing through a column of dry alumina then degassed at low temperature, or argon-degassed over 4Å molecular sieves. Molecular sieves were oven-dried at greater than 120 °C over 2 days or microwave-activated. Acetone was purchased from Fisher Scientific and dried over Drierite (anhydrous $CaSO_4$). 6-Aminouracil was purchased from Aldrich, recrystallised from boiling water and dried over P_2O_5 before use. Iodine/ pyridine/THF/water solution (0.02 M), phosphitylating reagents and DIHT were purchased from Link Technologies. *N*-Iodosuccinimide was purchased from Aldrich and recrystallised from 1,4-dioxane/carbon tetrachloride before use. Deuterated NMR solvents were purchased from Apollo Scientific Ltd. DOWEX 50WX8 ion-exchange resin was purchased from Aldrich and washed with 0.2 M HCl before treatment. All other reagents were purchased from Aldrich, Lancaster, Avocado, Alfa Aesar, BDH, Fisher Scientific, Acros and Fluka, and used without purification. All reactions requiring absence of oxygen were carried out under an atmosphere of argon. Glassware was oven-dried overnight before use, for reactions requiring exclusion of moisture.

Reactions were monitored by TLC, using Merck Kieselgel 60 F_{254} or Machenary-Nagel Alugram Sil G/UV_{254} silica gel plates (0.22 mm thickness, aluminium backed). Compounds were visualised by irradiation at 254/365 nm, or staining with *p*-anisaldehyde (A'), potassium permanganate (B'), ninhydrin (C', D'), 10% H_2SO_4/ethanol (E') or ceric sulfate (F'), followed by heating.

A' – *p*-anisaldehyde / glacial AcOH / conc. H_2SO_4 / ethanol (9.2:3.8:12.5:338 v/v)

B' – $KMnO_4$ / K_2CO_3 / 5% aq NaOH / water (3:20 g/5:300 mL)

C' – ninhydrin (5 g) / acetone (100 mL)

D' – 0.3% w/v ninhydrin / butan-1-ol / 3% v/v glacial AcOH

E' – 10% v/v conc. sulfuric acid / ethanol

F' – 8% w/v $Ce(SO_4)_2.2H_2SO_4$ hydrate / 15% v/v sulphuric acid

Column chromatography was carried out under air or argon pressure using Fisher Scientific DAVISIL 60Å (35-70 micron) silica gel. Silica gel was pre-equilibrated with triethylamine or pyridine for purification of acid-senstive compounds.

Proton and carbon NMR spectra were recorded at 300/400 MHz and 75/100 MHz, respectively, using either a Bruker AC300 or Bruker DPX400 spectrometer. Fluorine and phosphorous NMR spectra were recorded at 282 MHz and 121 MHz respectively, using a Bruker AC300 spectrometer. NMR spectra were recorded in deuterated chloroform or dimethylsulfoxide. Chemical shifts are given in ppm and spectra are calibrated to the residual solvent peak.[247] *J* values are correct to within 0.5 Hz. Assignment was aided by DEPT-135, ^1H–^1H COSY, HMQC and HMBC experiments and Lorentz-Gauss resolution enhancement data reprocessing (Win-NMR Lite).

Low-resolution mass spectra were recorded using electrospray ionisation (ES) on a Fisons VG platform instrument, or on a Waters ZMD quadrupole mass spectrometer in HPLC grade acetonitrile, methanol or water, or using electron ionisation (EI) on a ThermoQuest TraceMS single quadrupole GC-MS instrument in dichloromethane. High-resolution mass spectra were recorded in HPLC grade acetonitrile, methanol or water using electrospray ionisation on a Bruker APEX III FT-ICR mass spectrometer.

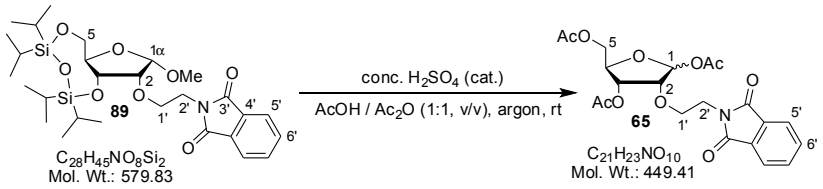
IR spectra were recorded on a Satellite FT-IR instrument using a 'Golden Gate' or 'Smart Orbit' adapter, and visualised using Win First-lite or OMNIC software respectively. All absorptions are measured in cm^{-1} and are described as broad (br), weak (w), medium (m), strong (s) or very strong (vs).

Melting points were recorded on a Gallenkamp Electrothermal melting point apparatus and uncalibrated.

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

9.1.2 Experimental Procedure

1,3,5-Tri-*O*-acetyl-2-*O*-(2-phthalimidoethyl)-D-ribofuranose, 65[37]

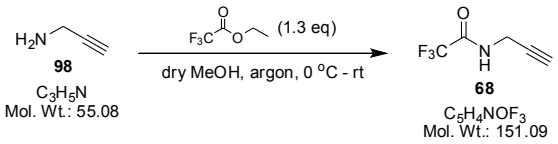


To a stirred solution of **89** (3.31 g, 5.71 mmol) in glacial acetic acid/acetic anhydride (1:1 v/v, 30 mL), under an argon atmosphere, conc. sulfuric acid (0.35 mL, 6.57 mmol, 1.1 eq) was added slowly. The reaction mixture was stirred at rt for 1 hour. Saturated aq sodium bicarbonate (250 mL) was added carefully to the dark green reaction mixture, which was extracted with CH_2Cl_2 (3 × 100 mL). The combined organic fraction was washed with sat. aq sodium bicarbonate and sat. aq KCl, dried (Na_2SO_4) and concentrated *in vacuo* to yield an orange-brown oil/solid. Following purification by column chromatography (1:1, hexane/ethyl acetate), and drying under high vacuum, the product **65** was isolated, as an anomeric mixture (α:β *ca* 2:1), as a very pale yellow oil (2.11 g, 4.70 mmol, 82% yield).

$\mathbf{R}_f$ 0.17 (1:1, hexane/ethyl acetate, A'); **^{1}H NMR** (400 MHz, $CDCl_3$) δ signals due to α anomer 7.86-7.80 (2H, m, **H$^{5'}$**), 7.74-7.68 (2H, m, **H$^{6'}$**), 6.07 (1H, s, **H^1**), 5.00 (1H, dd, *J* = 4.8, 6.8 Hz, **H^3**), 4.35-4.30 (1H, m, **H^4**), 4.29 (1H, t, *J* = 4.8 Hz, **H^5**), 4.14-4.08 (1H, m, **H^5**), 4.12-4.07 (1H, m, **H^2**), 3.94-3.83 (2H, m, **H$^{2'}$**), 3.83-3.72 (2H, m, **H$^{1'}$**), 2.05, 2.04 (3H, s, COC**H$_3$**), 2.00 (3H, s, COC**H$_3$**), signals due to β anomer 7.86-7.80 (2H, m, **H$^{5'}$**), 7.74-7.68 (2H, m, **H$^{6'}$**), 6.29 (1H, d, *J* = 4.5 Hz, **H^1**), 5.13 (1H, dd, *J* = 2.5, 6.5 Hz, **H^3**), 4.38 (1H, br dd, *J* = 3.3, 7.3 Hz, **H^4**), 4.24 (1H, dd, *J* = 3.5, 12.6 Hz, **H^5**), 4.06 (1H, dd, *J* = 5.5, 12.1 Hz, **H^5**), 3.96 (1H, dd, *J* = 4.5, 6.5 Hz, **H^2**), 3.94-3.83 (2H, m, **H$^{2'}$**), 3.83-3.72 (2H, m, **H$^{1'}$**), 2.07, 2.03, 1.96 (3H, s, COC**H$_3$**); **^{13}C NMR** (100 MHz, $CDCl_3$) δ signals due to α anomer 170.6, 170.4, 169.5 (**C**OCH_3), 168.2 (**C$^{3'}$**), 134.2 (**C$^{6'}$**), 132.2 (**C$^{4'}$**), 123.4 (**C$^{5'}$**), 98.9 (**C^1**), 80.8 (**C^2**), 79.5 (**C^4**), 71.9 (**C^3**), 67.8 (**C$^{1'}$**), 63.8 (**C^5**), 37.6 (**C$^{2'}$**), 21.2, 20.8, 20.5 (CO**C**H_3), signals due to β anomer 171.2, 170.5, 170.0 (**C**OCH_3), 168.2 (**C$^{3'}$**), 134.1 (**C$^{6'}$**), 132.3 (**C$^{4'}$**), 123.3 (**C$^{5'}$**), 94.8 (**C^1**), 81.4 (**C^4**), 77.9 (**C^2**), 69.7 (**C^3**), 68.6 (**C$^{1'}$**), 63.4 (**C^5**), 37.6 (**C$^{2'}$**), 21.1, 20.9, 20.7 (CO**C**H_3); **LRMS** (ES^+): *m*/*z* (%): 546 (38), 472 ($[M + Na]^+$, 100).

Analytical results consistent with reported data.[37]

N-Trifluoroacetyl propargylamine, 68[117,170]

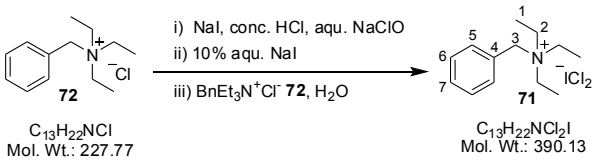


Propargylamine (3.71 g, 67.3 mmol) **98** was added dropwise to an ice-cooled, stirred solution of ethyl trifluoroacetate (10.4 mL, 87.6 mmol, 1.3 eq) in distilled methanol (50 mL) under an argon atmosphere. After addition, the solution was allowed to warm to rt and stirred for 15 hrs. The reaction mixture was concentrated *in vacuo* and redissolved in CH_2Cl_2 (200 mL). The solution was washed with 10% w/v aq citric acid (2 × 100 mL) and sat aq KCl (100 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give an orange/brown oil. Following column chromatography (CH_2Cl_2), and drying *in vacuo*, the product **68** was isolated as a very pale yellow oil (5.81 g, 38.4 mmol, 57%).

R_f 0.27, 0.33 (CH_2Cl_2, B', D'), 0.53, 0.66 (1:9, methanol/CH_2Cl_2, B', D'); **IR** (film): ν 3300 (br m, CC–H), 3092 (br w, N–H), 1703 (vs, C=O), 1551 (m, N–H), 1428, 1363, 1335 (w), 1207 (s), 1152 (br vs, C–F), 1046 (w), 997 (m), 926, 826 (w), 724 (m, C–F), 667 (m), 638 (br s), 602, 556 (m) cm^{-1}; **^{19}F NMR** (282 MHz, $CDCl_3$) δ -76.2 (C**F$_3$**); **^{1}H NMR** (300 MHz, $CDCl_3$) δ 6.73 (1H, br s, N**H**), 4.15 (2H, dd, *J* = 2.6, 5.3 Hz, C**H$_2$**), 2.33 (1H, t, *J* = 2.6 Hz, C≡C**H**); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 157.2 (q, *J* = 37.6 Hz, **C**$COCF_3$), 115.7 (q, *J* = 287.5 Hz, CO**C**F_3), 77.2 (**C**≡CH), 73.3 (C≡**C**H), 29.8 (**C**H_2); **LRMS** (ES^+): *m*/*z* : 206 ($[M + CH_3OH + Na]^+$, 100), 174 ($[M + Na]^+$, 45); (ES^-): *m*/*z* : 150 ($[M – H]^-$, 100), 80 (57).

Analytical results consistent with reported data.[117,170]

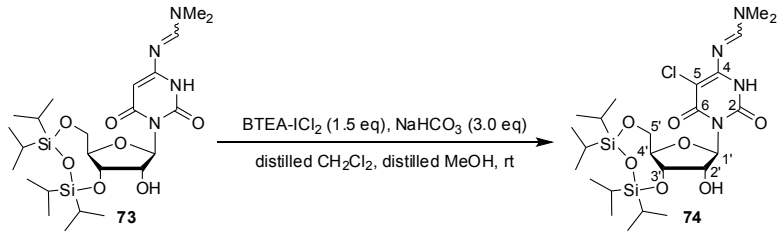
Benzyltriethylammonium dichloroiodate 71[154]



Sodium iodide (6.63 g, 44.2 mmol, 1.0 eq) was dissolved in conc. aq HCl (15.5 mL) and cooled in an ice bath. To this, an 8.8% w/v aqueous solution of NaClO (37.4 mL, 44.1 mmol) was added slowly, with swirling, so that the temperature was kept below 20 °C. Effervescence and a brown residue were observed. Towards the end of addition, the bleach solution was added dropwise until all brown colouration had disappeared. A 10% aq solution of NaI was added dropwise (20 mL) until a trace amount of iodine reappeared. The resulting yellow/orange solution was added portionwise to a solution of benzyltriethylammonium chloride **72** (10.0 g, 44.1 mmol) in water (14 mL) to give a brilliant yellow precipitate. The mixture was filtered, and the precipitate was washed with water and dried under high vacuum over P_2O_5, to give the product salt **71** as a fine, bright yellow powder (16.5 g, 42.2 mmol, 96% yield).

Mp 89-98 °C, *lit.* 84-86 °C;[154] **IR** (solid): 3053 (w, C–H), 2984 (m, C–H), 1479 (s, C–H, Aryl C=C), 1390 (m, C–H), 1360 (m), 1155 (m), 1004 (m), 903 (w), 796 (m), 757 (s, Aryl C–H), 708 (vs Aryl C–H), 605 (w) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO): δ 7.52 (5H, app. t, J = 1.8 Hz, C**H**^{Ar}), 4.48 (2H, s, **H**^{3}), 3.17 (6H, q, J = 7.3 Hz, **H**^{2}), 1.31 (9H, t, J = 7.3 Hz, **H**^{1}); **^{13}C NMR** (75 MHz, d_6-DMSO): δ 132.5 (**C**^{5}), 130.2 (**C**^{6}), 129.0 (**C**^{7}), 127.8 (**C**^{4}), 59.5 (**C**^{3}), 52.0 (**C**^{2}), 7.5 (**C**^{1}); **LRMS** (ES^+): *m*/*z* (%): 192 (M^+, 100); (ES^-): *m*/*z* (%): 380 (I_3^-, 16), 197/199/201 (ICl_2^-, 4, 11:3:1), 126 (I^-, 100).
Analytical results are consistent with the reported data.[154]

N^1-[1-(3',5'-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-β-D-ribofuranosyl)]-(5-chloro-6-oxocytosin-4-yl)-*N*,*N*-dimethylformamidine, 74



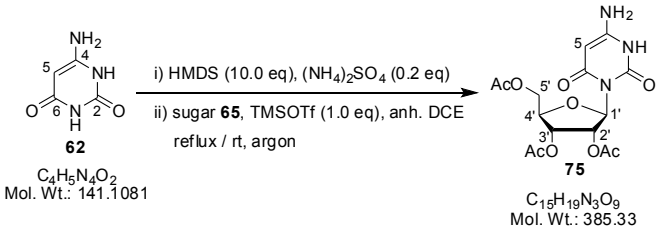
$C_{24}H_{44}N_4O_7Si_2$
Mol. Wt.: 556.80

$C_{24}H_{43}N_4O_7ClSi_2$
Mol. Wt.: 591.24

To a stirred solution of nucleoside **73** (0.66 g, 1.18 mmol) in distilled CH_2Cl_2 (20 mL) and distilled methanol (10 mL), was added anhydrous sodium bicarbonate (0.30 g, 3.56 mmol, 3.0 eq), followed by a solution of benzyltriethylammonium dichloroiodate **71** (0.69 g, 1.77 mmol, 1.5 eq) in distilled CH_2Cl_2 (15 mL), and the reaction mixture was stirred vigorously at rt for 50 minutes. The crude reaction mixture was filtered and concentrated *in vacuo* to give an orange/brown gum. The residue was redissolved in CH_2Cl_2 (50 mL) and washed with water (3 × 50 mL). This was further washed with sat. aq sodium thiosulfate (50 mL), water (2 × 50 mL) and brine (2 × 50 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield a pink foam. The compound was twice purified by column chromatography (0-2% methanol/CH_2Cl_2) to yield after drying *in* vacuo, the chlorinated product **74** as a white crystalline solid (0.37 g, 0.54 mmol, 46%).

R_f 0.50 (14:1, CH_2Cl_2/methanol, A'); **Mp** 94-99 °C (methanol/CH_2Cl_2); **IR** (solid): ν 3485 (w), 2942, 2891, 2864 (m, C–H), 1714 (m, C=O), 1637 (s, C=O, C=C), 1578 (vs, C=N, C=C), 1462 (m), 1418 (s), 1371 (s, O–H), 1296, 1246 (m), 1122 (s), 1074 (s, C–O), 1029 (vs, Si–O), 917 (m), 882, 861, 761 (s), 690 (vs, C–Cl), 587 (s) cm^{-1}; **^{1}H NMR** (300 MHz, $CDCl_3$) δ 8.16 (1H, s, **H^7**), 6.24 (1H, d, *J* = 1.5 Hz, **H$^{1'}$**), 4.98 (1H, t, *J* = 7.0 Hz, **H$^{3'}$**), 4.60 (1H, dd, *J* = 1.5, 6.2 Hz, **H$^{2'}$**), 4.01 (2H, d, *J* = 5.5 Hz, **H$^{5'}$**), 3.84 (1H, td, *J* = 5.3, 7.3 Hz, **H$^{4'}$**), 3.33 (1H, v br s, 2'-O**H**), 3.19, 3.12 (3H, s, NC**H$_3$**), 1.15-1.02 (28H, m, C**H**(C**H$_3$**)$_2$); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 160.0 (**C^6**), 156.8 (N=**C^7**–N), 151.4 (**C^2**, **C^4**), 94.2 (**C^5**–Cl), 89.4 (**C$^{1'}$**), 82.8 (**C$^{4'}$**), 73.7 (**C$^{2'}$**), 73.1 (**C$^{3'}$**), 64.1 (**C$^{5'}$**), 41.6, 35.2 (N–**C**H$_3$), 17.6, 17.5, 17.5, 17.4, 17.3, 17.3, 17.3, 17.2 (Si–CH(**C**H$_3$)$_2$), 13.4, 13.3, 12.9, 12.8 (Si–**C**H(CH$_3$)$_2$); **LRMS** (ES$^+$): *m*/*z* (%): 1203 ([2M + Na]$^+$, 44), 613 ([M + Na]$^+$, 100); **HRMS** (ES$^+$): calcd for $C_{24}H_{43}O_7N_4ClSi_2$ (M), [M + Na]$^+$ = 613.2251, found 613.2267.

2',3',5'-Tri-*O*-acetyl-6-oxocytidine, 75[71,248]



62
$C_4H_5N_4O_2$
Mol. Wt.: 141.1081

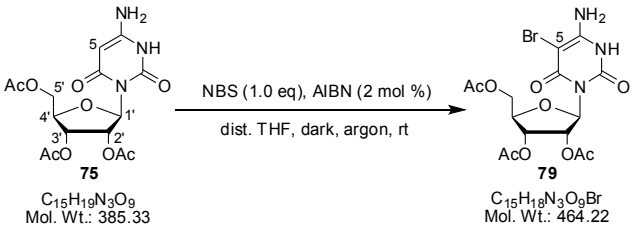
75
$C_{15}H_{19}N_3O_9$
Mol. Wt.: 385.33

A suspension of dry 6-aminouracil **62** (8.00 g, 62.9 mmol, 1.1 eq) and ammonium sulfate (1.67 mg, 12.6 mmol, 0.2 eq) in HMDS (131.5 mL, 631 mmol, 10.5 eq) was heated to reflux, under an argon atmosphere, whilst stirring vigorously, for 24 hours, during which the starting material dissolved. The reaction was allowed to cool to rt and the solvent was removed under high vacuum. The cream-coloured solid was redissolved in anhydrous DCE (150 mL) under an argon atmosphere, to which was added acetylated sugar **65** (19.0 g, 59.8 mmol). Trimethylsilyl triflate (11.4 mL, 63.0 mmol, 1.1 eq) was carefully added dropwise and the reaction was stirred at rt, under an argon atmosphere, for 43 hours. After concentrating *in vacuo* the residue was redissolved in CH_2Cl_2 (250 mL), and washed with water (150 mL) and dilute aq sodium bicarbonate (150 mL), during which a white solid precipitated out of solution. The mixture was filtered, and following drying under high vacuum at 50 °C for 2 days, the product β-anomer **75** was afforded as a fine white solid (16.9 g, 43.8 mmol, 73% yield). The aqueous layer was re-extracted (CH_2Cl_2, 100 mL) and the combined, filtered organic fractions were washed with sat. aq KCl (200 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield a peach-coloured foam/solid (3.52 g) containing the sugar **65** and α- **76** and β-anomers **75** of the product nucleoside.

R_f 0.28 (1:9 methanol/CH_2Cl_2, A'); **Mp** 197-201 °C (CH_2Cl_2), *lit.*[71] 228-229 °C (ethanol/diethyl ether); **IR** (solid): ν 3447, 3324 (w, N–H), 3211, 2992 (w), 2911, 2789 (w, C–H), 1736 (m, C=O), 1718 (s, C=O), 1652 (s, N–H), 1586 (br m), 1537, 1427 (m), 1372 (m, C–H), 1252, 1215 (s, C–O), 1071, 1044 (s), 973, 796, 750 (m), 641 (s), 623, 582 (m), 547 (vs) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.64 (1H, br s, N**H**^{3}), 6.47 (2H, br s, N**H**$_2$), 6.15 (1H, br s, **H**$^{1'}$), 5.64 (1H, dd, *J* = 2.8, 6.8 Hz, **H**$^{2'}$), 5.47 (1H, app. t, *J* = 7.0 Hz, **H**$^{3'}$), 4.55 (1H, s, **H**^{5}), 4.31 (1H, dd, *J* = 3.0, 11.5 Hz, **H**$^{5'}$), 4.08 (1H, ddd, *J* = 3.1, 6.4, 8.0 Hz, **H**$^{4'}$), 4.02 (1H, dd, *J* = 6.5, 11.0 Hz, **H**$^{5'}$), 2.06, 2.03, 2.00 (3H, s, COC**H**$_3$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.1, 169.5, 169.3 (**C**OCH$_3$), 161.9 (**C**^{6}),

154.5 (**C**^{4}–NH_2), 150.4 (**C**^{2}), 84.4 (**C**$^{1'}$), 77.7 (**C**$^{4'}$), 73.6 (**C**^{5}), 72.6 (**C**$^{2'}$), 69.9 (**C**$^{3'}$), 63.1 (**C**$^{5'}$), 20.5, 20.4, 20.2 (CO**C**H_3); **LRMS** (ES^+): *m*/*z* (%): 793 ($[2M + Na]^+$, 12), 430 (32), 408 ($[M + Na]^+$, 100); (ES^-): *m*/*z* (%): 769 ($[2M – H]^-$, 32), 384 ($[M – H]^-$, 100). Analytical results consistent with reported data.[71,248]

5-Bromo-2',3',5'-tri-*O*-acetyl-6-oxocytidine, 79

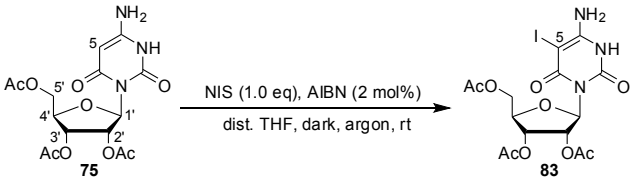


To a stirred suspension of 2',3',5'-tri-*O*-acetyl-6-oxocytidine **75** (3.05 g, 7.91 mmol) and NBS (1.41 g, 7.92 mmol, 1.0 eq) in distilled THF (30.0 mL) was added AIBN (27.0 mg, 0.16 mmol, 2.1 mol %), and the reaction was stirred in absence of light, under an argon atmosphere at rt for 30 minutes. The white slurry was concentrated *in vacuo* to yield a white solid. Following purification by column chromatography (25-80% acetone/toluene), and drying *in vacuo*, the product **79** was afforded as a powdery white solid (3.47 g, 7.48 mmol, 95% yield) containing 2.0 mol % succinimide.

$\mathbf{R}_f$ 0.35 (1:1, acetone/toluene, A'), 0.41 (1:9 methanol/CH_2Cl_2, A'); **IR** (solid): ν 3412 (w, N–H), 3301 (br w, N–H), 3240, 3194, 3137 (w), 3089 (w, N–H), 3020, 2909 (w, CH_2), 1739 (s, C=O), 1718 (vs, C=O), 1628, 1557 (s, N–H), 1426 (m), 1371 (s, CH_3), 1225 (vs, C–O), 1156, 1128, 1091 (s), 1040 (vs), 1013 (s), 941, 930 (m), 897, 765, 727 (s), 678 (m), 617 (s), 538 (s, C–Br) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO) δ 11.00 (1H, br s, N**H**^{3}), 6.70 (2H, s, N$\mathbf{H_2}$), 6.16 (1H, br s, **H**$^{1'}$), 5.62 (1H, dd, *J* = 2.5, 6.5 Hz, **H**$^{2'}$), 5.49 (1H, app. t, *J* = 7.3 Hz, **H**$^{3'}$), 4.33 (1H, dd, *J* = 3.3, 11.8 Hz, **H**$^{5'}$), 4.11 (1H, ddd, *J* = 3.0, 6.5, 8.0 Hz, **H**$^{4'}$), 4.02 (1H, dd, *J* = 6.3, 11.8 Hz, **H**$^{5'}$), 2.07, 2.03, 2.00 (3H, s, CO**C**$\mathbf{H_3}$); **^{13}C NMR** (75 MHz, d_6-DMSO) δ 170.0, 170.0, 169.3 (**C**OCH_3), 158.1 (**C**^{6}), 151.4 (**C**^{4}-NH_2), 149.2 (**C**^{2}), 85.9 (**C**$^{1'}$), 77.9 (**C**$^{4'}$), 72.6 (**C**$^{2'}$), 69.7 (**C**$^{3'}$), 69.1 (**C**^{5}), 62.9 (**C**$^{5'}$), 20.5, 20.3, 20.2 (CO**C**H_3); **LRMS** (ES^+): *m*/*z* : 949/951/953 ($[2M + Na]^+$, 1:2:1, 52), 527/529 ($[M + CH_3CN + Na]^+$, 1:1, 40), 495/497 (1:1, 100), 486/488 ($[M + Na]^+$, 1:1, 99); **HRMS** (ES^+): calcd for $C_{15}H_{18}N_3O_9Br$ (M), $[M + Na]^+$ = 486.0124/488.0104,

found 486.0114/488.0097 (1:1), $[2M + Na]^+$ = 949.0351/951.0330/953.0310, found 951.0254 (1:2:1).

5-Iodo-2',3',5'-tri-*O*-acetyl-6-oxocytidine, 83



$C_{15}H_{19}N_3O_9$
Mol. Wt.: 385.33

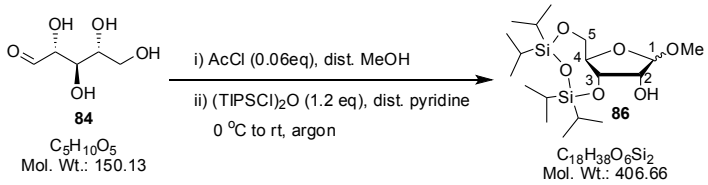
$C_{15}H_{18}N_3O_9I$
Mol. Wt.: 511.22

To a stirred suspension of 2',3',5'-tri-*O*-acetyl-6-oxocytidine **75** (5.00 g, 13.0 mmol) and AIBN (46.3 mg, 0.28 mmol, 2.2 mol %) in distilled THF (50.0 mL) was added NIS (2.95 g, 13.1 mmol, 1.0 eq), and the reaction was stirred in absence of light, under an argon atmosphere, at rt for 15 minutes. The orange/brown slurry was concentrated *in vacuo* to yield an orange/brown solid. A portion of the solid (~0.5 g) was dissolved in ethyl acetate (250 mL) and washed with dilute aq $Na_2S_2O_3$ (100 mL, adjusted to pH 7 using $NaHCO_3$) until decolourised then separated immediately. The solution was further washed with sat. aq KCl (2 × 100 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give an off-white solid (0.34 g). The remainder was suspended in ethyl acetate (550 mL) and allowed to stand in the fridge for 2 hours. The mixture was vacuum filtered, and the residue was washed with ice-cold ethyl acetate (3 × 150 mL) and dried under high vacuum over P_2O_5, to yield the product **83** as a white, powdery solid (5.47 g) containing 0.4 mol % succinimide (recalculated yield - 10.7 mmol, 82%).

R_f 0.35 (1:1, acetone/toluene, A'), 0.54 (tetrahydrofuran, A'); **IR** (solid): ν 3400, 3295 (w, N–H), 3207 (br w, N–H), 1740, 1719 (s, C=O), 1624 (s, C=C), 1551 (s, N–H), 1422 (m), 1371 (s, CH_3), 1228 (vs), 1096, 1040 (s, C–O), 942, 931, 898 (m), 763 (s), 736 (m), 722 (m, CH_2), 677, 647, 630 (m), 602, 571, 537 (s), 512 (m, C–I) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO) δ 10.91 (1H, br s, N**H**^{3}), 6.54 (2H, s, N**H$_2$**), 6.17 (1H, s, **H**$^{1'}$), 5.61 (1H, dd, *J* = 1.8, 6.4 Hz, **H**$^{2'}$), 5.49 (1H, app. t, *J* = 7.2 Hz, **H**$^{3'}$), 4.33 (1H, dd, *J* = 2.7, 11.5 Hz, **H**$^{5'}$), 4.10 (1H, ddd, *J* = 2.9, 6.4, 8.0 Hz, **H**$^{4'}$), 4.02 (1H, dd, *J* = 6.2, 11.5 Hz, **H**$^{5'}$), 2.07, 2.03, 2.00 (3H, s, COC**H$_3$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.0, 169.5, 169.3 (**C**OCH_3), 159.3 (**C**^{6}), 153.3 (**C**^{4}–NH_2), 149.5 (**C**^{2}), 86.1 (**C**$^{1'}$), 77.8 (**C**$^{4'}$), 72.7

(**C**$^{2'}$), 69.6 (**C**$^{3'}$), 62.8 (**C**$^{5'}$), 41.7 (**C**^{5}-I), 20.5 (2'-OCO**C**H_3), 20.3 (3'-OCO**C**H_3), 20.2 (5'-OCO**C**H_3); **LRMS** (ES^+): *m*/*z* : 1061 ($[2M + K]^+$, 3), 1045 ($[2M + Na]^+$, 8), 1023 ($[2M + H]^+$, 2), 550 ($[M + K]^+$, 24), 534 ($[M + Na]^+$, 100), 512 ($[M + H]^+$, 77); (ES^-): *m*/*z* (%): 1021 ($[2M - H]^-$, 2), 608 ($[M + H_3PO_4 - H]^-$, 4), 510 ($[M - H]^-$, 100); **HRMS** (ES^+): calcd for $C_{15}H_{18}N_3O_9I$ (M), $[M + Na]^+$ = 533.9980, found 533.9993, $[M + NH_4]^+$ = 529.0426, found 529.0445, $[M + H]^+$ = 512.0160, found 512.0184.

1-*O*-Methyl-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 86[37]



Acetyl chloride (0.91 mL, 12.8 mmol, 0.1 eq) in distilled methanol (50 mL) was added portionwise to a stirred solution of D-ribose **84** (30.0 g, 200 mmol) in distilled methanol (200 mL). The reaction mixture was stirred at rt, under an argon atmosphere, for 22 hours. Sodium bicarbonate (10.1 g, 120 mmol, 0.6 eq) was added and the mixture was stirred for 10 minutes. The mixture was filtered and concentrated *in vacuo* to yield the crude product, 1'-*O*-methylribofuranoside **85** as an anomeric mixture (α:β *ca* 1:1) as a viscous, orange/red oil (38.3 g), which was used without further purification.

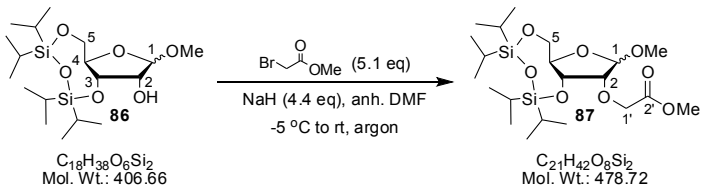
$\mathbf{R}_f$ 0.47, 0.32 (4:1, ethyl acetate/methanol, A'); **LRMS** (ES^+): *m*/*z* (%): 187 ($[M + Na]^+$, 100).

The crude product **85** (5.63 g) was co-evaporated with distilled pyridine and dried under high vacuum before re-dissolving in distilled pyridine (40 mL). The solution was cooled to 0 °C, under an argon atmosphere, and 1,3-dichloro-1,1,3,3-tetraisopropyl-disiloxane (13.2 mL, 41.3 mmol, 1.4 eq) was added dropwise. The reaction mixture was allowed to warm to rt and stirred for 15 hours, then concentrated under high vacuum, co-evaporating with toluene. The resulting cream slurry was partitioned between ethyl acetate (400 mL) and water (200 mL). The organic layer was washed with water (2 × 200 mL), 2 M aq HCl (200 mL) then water (200 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield a yellow oil. Following partial purification by

column chromatography (4:1, hexane/ethyl acetate), and drying under high vacuum the product **86** was afforded, as an anomeric mixture (α:β *ca* 6:1), as a very pale yellow viscous oil (15.1 g, 37.1 mmol, *quant.*).

R_f 0.47, 0.30 (4:1, hexane/ethyl acetate, A'), 0.74 (4:1, ethyl acetate/methanol, A'); **^{1}H NMR** (300 MHz, $CDCl_3$) δ signals due to α anomer 4.85 (1H, s, **H^1**), 4.70 (1H, br t, *J* = 5.3 Hz, **H^3**), 4.13-4.02 (1H, m, **H^2**), 3.99 (1H, dd, *J* = 4.9, 8.9 Hz, **H^4**), 3.85 (1H, dd, *J* = 2.5, 12.3 Hz, **H^5**), 3.69 (1H, dd, *J* = 2.2, 12.3 Hz, **H^5**), 3.40 (3H, s, Me), 2.75 (1H, br s, 2-O**H**), 1.11-0.80 (28H, m, C**H**(C**H$_3$**)$_2$), signals due to β anomer 4.87 (1H, d, *J* = 4.5 Hz, **H^1**), 4.50 (1H, br s, **H^3**), 4.29-4.16 (1H, m, **H^4**), 4.13-4.01 (2H, m, **H^2**, **H^5**), 3.78 (1H, br d, *J* = 10.6 Hz, **H^5**), 3.32 (3H, s, Me), 2.75 (1H, br s, 2-O**H**), 1.10-0.80 (28H, m, C**H**(C**H$_3$**)$_2$); **LRMS** (ES^+): *m*/*z* (%): 447 (100), 429 ([M + Na]$^+$, 50).
Analytical results consistent with reported data.[37]

1-*O*-Methyl-2-*O*-methylethanoyl-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 87[37]



$C_{18}H_{38}O_6Si_2$
Mol. Wt.: 406.66

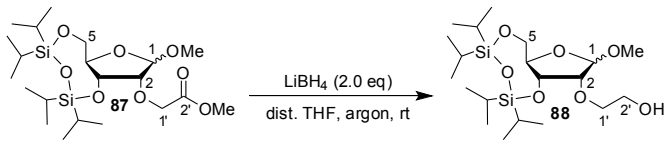
$C_{21}H_{42}O_8Si_2$
Mol. Wt.: 478.72

Compound **86** (15.1 g, 37.1 mmol) was dissolved in anhydrous DMF, and cooled to ⁻5 °C (ice/methanol) under an argon atmosphere. Methyl bromoacetate (18.0 mL, 190.2 mmol, 5.13 eq) was added followed by sodium hydride (60% w/w dispersion in mineral oil) (6.48 g, 162 mmol, 4.4 eq) portionwise. The reaction was stirred under an argon atmosphere at ⁻5 °C for 15 minutes, then allowed to warm and stirred vigorously at rt for 17 hours. Saturated aq KCl (100 mL) was added carefully to the stirred reaction, and the mixture was extracted with diethyl ether (3 × 150 mL). The combined organic fraction was dried (Na_2SO_4) and concentrated under high vacuum, co-evaporating with toluene to yield a two-layered liquid (yellow/colourless). The crude compound was twice-purified by column chromatography (5-20% ethyl acetate/hexane) to yield, after drying under high vacuum, the product **87**, as an anomeric mixture (α:β *ca.* 5:1), as a white solid (10.6 g, 22.1 mmol, 60% yield).

$\mathbf{R}_f$ 0.33, 0.21 (4:1, hexane/ethyl acetate, A'); **^{1}H NMR** (300 MHz, $CDCl_3$) δ signals due to α anomer 4.86 (1H, s, **H^1**), 4.48 (1H, dd, *J* = 4.2, 7.9 Hz, **H^3**), 4.45 (1H, d, *J* = 16.7 Hz, **H$^{1'}$**), 4.32 (1H, d, *J* = 16.7 Hz, **H$^{1'}$**), 4.04 (1H, td, *J* = 2.9, 7.9 Hz, **H^4**), 3.98 (1H, dd, *J* = 2.9, 12.1 Hz, **H^5**), 3.86 (1H, dd, *J* = 5.9, 12.1 Hz, **H^5**), 3.83 (1H, d, *J* = 4.0 Hz, **H^2**), 3.74 (3H, s, $CO_2C\mathbf{H_3}$), 3.32 (3H, s, 1-O**CH$_3$**), 1.10-0.93 (28H, m, C**H**(C**H$_3$**)$_2$), δ signals due to β anomer 4.71 (1H, s, **H^1**), 4.45 (1H, d, *J* = 16.7 Hz, **H$^{1'}$**), 4.32 (1H, d, *J* = 16.7 Hz, **H$^{1'}$**), 4.05-4.02 (1H, m, **H^4**), 4.02-3.77 (4H, m, **H^2**, **H^3**, **H^5**), 3.78 (3H, s, $CO_2C\mathbf{H_3}$), 3.32 (3H, d, *J* = 1.0 Hz, 1-O**CH$_3$**), 1.10-0.93 (28H, m, C**H**(C**H$_3$**)$_2$); **^{13}C NMR** (75 MHz, $CDCl_3$) δ signals due to α anomer 170.5 (**C$^{2'}$**), 106.0 (**C^1**), 83.3 (**C^2**), 81.0 (**C^4**), 74.2 (**C^3**), 68.4 (**C$^{1'}$**), 63.9 (**C^5**), 54.9 (1-O**C**H$_3$), 51.9 ($CO_2\mathbf{C}H_3$), signals due to β anomer 171.0 (**C$^{2'}$**) other signals not visible, indistinguishable signals due to both anomers 17.6, 17.5, 17.5, 17.5, 17.4, 17.4, 17.2, 17.1 (Si–CH(**C**H$_3$)$_2$), 13.5, 13.4, 12.9, 12.8 (Si–**C**H(CH$_3$)$_2$); **LRMS** (ES$^+$): *m*/*z* (%): 501 ([M + Na]$^+$, 100).

Analytical results consistent with reported data.[37]

1-*O*-Methyl-2-*O*-(2-hydroxyethyl)-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 88[37]



$C_{21}H_{42}O_8Si_2$
Mol. Wt.: 478.72

$C_{20}H_{42}O_7Si_2$
Mol. Wt.: 450.71

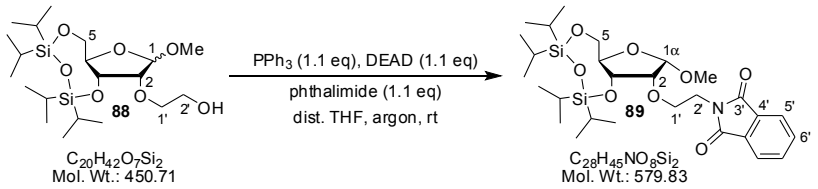
To a solution of compound **87** (9.13 g, 19.1 mmol) in distilled THF (80 mL), under an argon atmosphere, $LiBH_4$ (0.84 g, 38.5 mmol, 2.0 eq) was added portionwise and the reaction was stirred at rt. Further $LiBH_4$ (0.21 g, 9.6 mmol, 0.5 eq) was added after 55 minutes and the reaction was stirred for a further 20 minutes. The reaction mixture was poured into a conical flask, cooled in an ice bath and stirred vigorously. A solution of methanol (7 mL) in THF (50 mL) was added with extreme caution and the reaction mixture was left to stir for 20 minutes. Methanol (30 mL) was added with care, and the mixture was diluted with diethyl ether (150 mL). The solution was washed with water, dried (Na_2SO_4) and concentrated *in vacuo* to yield an orange oil. Following purification by column chromatography (14-20% ethyl acetate/hexane), and drying under high

vacuum, the product **88** was isolated as an anomeric mixture (α:β *ca* 1:3), as a very pale yellow to colourless oil (3.10 g, 6.87 mmol, 36% yield).

R_f 0.09 (4:1, hexane/ethyl acetate, A'); **^{1}H NMR** (300 MHz, $CDCl_3$) δ signals due to α anomer 4.77 (1H, s, **H^1**), 4.50 (1H, dd, *J* = 4.5, 7.8 Hz, **H^3**), 4.05-3.96 (1H, m, **H^4**), 4.00 (1H, dd, *J* = 5.4, 11.3 Hz, **H^5**), 3.91-3.79 (3H, m, **H^5**, **H$^{1'}$**), 3.73-3.67 (2H, m, **H$^{2'}$**), 3.54 (1H, d, *J* = 7.7 Hz, **H^2**), 3.32 (3H, s, OC**H$_3$**), 2.06 (1H, br s, O**H**), 1.10-1.00 (28H, m, C**H**(C**H$_3$**)$_2$), signals due to β anomer 4.63 (1H, d, *J* = 7.5 Hz, **H^1**), 4.12 (1H, app. t, *J* = 3.6 Hz, **H^4**), 3.91-3.79 (1H, m, **H^3**), 3.91-3.73 (2H, m, **H$^{2'}$**), 3.78-3.73 (2H, m, **H$^{1'}$**), 3.75 (1H, t, *J* = 4.8 Hz, **H^2**), 3.73 (1H, dd, *J* = 4.8, 11.2 Hz, **H^5**), 3.66 (1H, dd, *J* = 2.6, 11.2 Hz, **H^5**), 3.47 (3H, s, OC**H$_3$**), 2.06 (1H, br s, O**H**), 1.10-1.00 (28H, m, C**H**(C**H$_3$**)$_2$); **^{13}C NMR** (75 MHz, $CDCl_3$) δ signals due to α anomer 106.1 (**C^1**), 83.8 (**C^2**), 79.4 (**C^4**), 73.2 (**C^5**), 71.8 (**C^3**), 66.2 (**C$^{1'}$**), 62.7 (**C$^{2'}$**), 56.5 (O**C**H$_3$), signals due to β anomer 100.5 (**C^1**), 81.2 (**C^2**), 73.7 (**C^4**), 71.3 (**C^3**), 65.0 (**C$^{1'}$**), 63.8 (**C^5**), 61.8 (**C$^{2'}$**), 54.8 (O**C**H$_3$), indistinguishable signals due to both anomers 17.6, 17.6, 17.5, 17.4, 17.4, 17.4, 17.1, 17.0 (Si–CH(**C**H$_3$)$_2$), 13.6, 13.4, 12.8, 12.7 (Si–**C**H(CH$_3$)$_2$); **LRMS** (ES$^+$): *m*/*z* (%): 473 ([M + Na]$^+$, 100).

Analytical results consistent with reported data.[37]

1-*O*-Methyl-2-*O*-(2-phthalimidoethyl)-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-α-D-ribofuranose, 89[37]

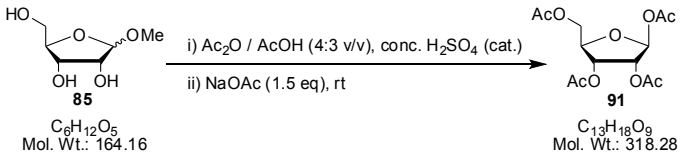


To a solution of compound **88** (2.90 g, 6.43 mmol) in distilled THF (30 mL), under an argon atmosphere, triphenylphosphine (1.86 g, 7.07 mmol, 1.1 eq) and phthalimide (1.04 g, 7.07 mmol, 1.1 eq) were added, followed by DEAD (1.1 mL, 6.99 mmol, 1.1 eq) dropwise, whilst cooling in a water bath. The reaction mixture was then stirred under an argon atmosphere, at rt for 2 hours. The solvent was removed *in vacuo* and the crude material was eluted through a short silica plug (4:1, hexane/ethyl acetate). Following further purification by column chromatography (4:1, hexane/ethyl acetate),

and drying under high vacuum, the product α-anomer **89** was isolated as a cream-coloured solid (3.49 g, 6.0 mmol, 94% yield).

R_f 0.26 (**89**, α), 0.10 (β) (4:1, hexane/ethyl acetate, A'); **^{1}H NMR** (400 MHz, $CDCl_3$) δ signals for α anomer **89** only 7.84 (2H, dd, *J* = 3.0, 5.5 Hz, **H$^{5'}$**), 7.70 (2H, dd, *J* = 3.0, 5.5 Hz, **H$^{6'}$**), 4.67 (1H, s, **H^1**), 4.42 (1H, dd, *J* = 4.3, 7.8 Hz, **H^3**), 4.09 (1H, d, *J* = 4.7, 9.3 Hz, **H$^{1'}$**), 3.98-3.90 (1H, m, **H$^{1'}$**), 3.95-3.90 (1H, m, **H^4**), 3.94 (1H, dd, *J* = 3.0, 12.0 Hz, **H^5**), 3.87-3.80 (2H, m, **H$^{2'}$**), 3.86 (1H, dd, *J* = 5.8, 12.3 Hz, **H^5**), 3.73 (1H, d, *J* = 4.3 Hz, **H^2**), 3.28 (3H, s, O**CH$_3$**), 1.09-0.93 (28H, m, **CH**(**CH$_3$**)$_2$); **^{13}C NMR** (100 MHz, $CDCl_3$) δ signals for α anomer **89** only 168.3 (**C$^{3'}$**), 134.0 (**C$^{6'}$**), 132.3 (**C$^{4'}$**), 123.4 (**C$^{5'}$**), 106.2 (**C^1**), 83.0 (**C^2**), 81.0 (**C^4**), 73.8 (**C^3**), 68.2 (**C$^{1'}$**), 63.7 (**C^5**), 54.8 (O**C**H$_3$), 38.0 (**C$^{2'}$**), 17.6, 17.5, 17.5, 17.5, 17.5, 17.4, 17.2, 17.1 (Si–CH(**C**H$_3$)$_2$), 13.5, 13.4, 12.8, 12.8 (Si-**C**H(CH$_3$)$_2$); **LRMS** (ES$^+$): *m/z* (%): 602 ([M + Na]$^+$, 100).
Analytical results consistent with reported data.[37]

1,2,3,5-Tetra-*O*-acetyl-β-D-ribofuranose, 91[249,250]

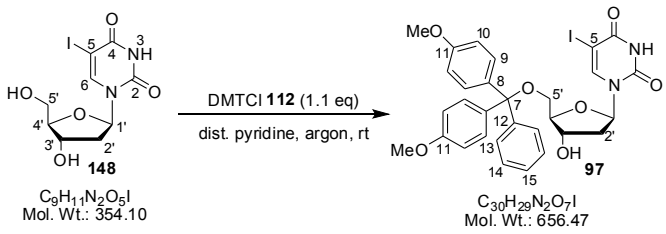


1-*O*-Methylribofuranoside **85** (30.0 g, 183 mmol) was dissolved in acetic anhydride (105 mL) and glacial acetic acid (80 mL), to which was added conc. sulfuric acid (3.00 mL, 56.3 mmol, 0.3 eq) dropwise with ice-cooling. The deep red-brown solution was stirred at rt for 1 hour after which further conc. sulfuric acid (5.00 mL, 93.8 mmol, 0.5 eq) was added dropwise with ice-cooling. After stirring the solution for a further 2 hours at rt, sodium acetate (22.6 g, 275 mmol, 1.5 eq) was added, and the solution was stirred vigorously for 10 minutes. The solvent was removed under high vacuum, co-evaporating with ethanol/isopropanol, to yield a brown, opaque syrup. The syrup was dissolved in chloroform (1.5 L) and washed with water (2 × 400 mL) and sat. aq KCl (2 × 400 mL), dried (Na_2SO_4) and concentrated under high vacuum to afford a clear, brown syrup. To the syrup was added ice-cold ethanol (100 mL) and the mixture was cooled, from which crystallised the product. Following washing with ice-cold ethanol

(2 × 50 mL) and ice-cold diethyl ether (3 × 100 mL), and drying under high vacuum over KOH, the product β-anomer **91** was afforded, as a white, crystalline solid (22.6 g, 71.2 mmol, 39% yield). The remaining fractions and washings were dried under high vacuum to give a brown syrup, as an anomeric mixture (α:β *ca* 65:35) (37.4 g) from which further product could not be crystallised.

R_f 0.47 (**91**, β), 0.40 (α) (2:1, ethyl acetate/hexane, A'), 0.30 (**91**, β), 0.26 (α) (1:1, ethyl acetate/hexane, A'); **Mp** 79-81 °C (ethanol), *lit.*[249,250] 81-83 °C (ethanol); **IR** (solid): ν 3025, 2992, 2961, 2932 (w, C–H), 1751 (s, C=O), 1740 (vs, C=O), 1445 (m), 1374 (m, C–H), 1215 (br vs, C–O), 1095, 1072, 1030 (m, C–O), 956, 913, 893, 883 (s), 815, 729, 679, 643, 621, 603, 589, 552, 536 (m) cm^{-1}; **^{1}H NMR** (300 MHz, $CDCl_3$) δ signals for β anomer **91** only 6.15 (1H, s, **H^1**), 5.34 (1H, dd, *J* = 4.8, 8.1 Hz, **H^3**), 5.33 (1H, app. s, **H^2**), 4.40-4.33 (1H, m, **H^4**), 4.32 (1H, dd, *J* = 3.3, 12.1 Hz, **H^5**), 4.14 (1H, dd, *J* = 5.3, 11.9 Hz, **H^5**), 2.11 (3H, s, COC**H$_3$**), 2.08 (3H, s, COC**H$_3$**), 2.08 (3H, s, COC**H$_3$**), 2.06 (3H, s, COC**H$_3$**); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 170.5, 169.8, 169.5, 169.1 (**C**$COCH_3$), 98.3 (**C^1**), 79.4 (**C^4**), 74.3 (**C^2**), 70.7 (**C^3**), 63.8 (**C^5**), 21.1, 20.8, 20.6, 20.6 (CO**C**H_3); **LRMS** (ES^+): *m/z* (%): 659 ($[2M + Na]^+$, 22), 382 ($[M + CH_3CN + Na]^+$, 87), 341 ($[M + Na]^+$, 100).
Analytical results consistent with reported data.[250]

5'-*O*-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine, 97[209]



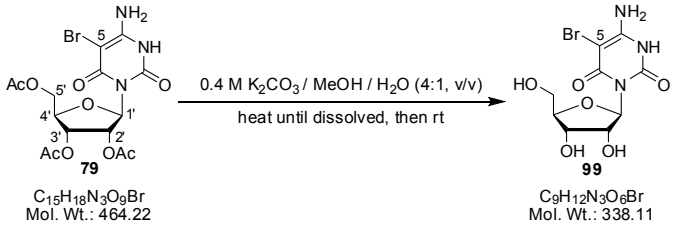
5-Iodo-2'-deoxyuridine **148** (25.0 g, 70.6 mmol) was dried by co-evaporation with distilled pyridine (2 × 50 mL) under high vacuum, then dissolved in distilled pyridine (120 mL). 4,4'-Dimethoxytrityl chloride **112** (25.1 g, 74.1 mol, 1.1 eq) was added in five portions over 4 hours, under an argon atmosphere, whilst stirring vigorously, at rt. The reaction mixture was stirred for a further 3.5 hours cooling in ice, then quenched with methanol/water (1:1 v/v, 20 mL) stirring for 10 minutes at rt. The reaction mixture

was concentrated *in vacuo* and by co-evaporation with methanol/toluene. The foam/syrup was dissolved in CH_2Cl_2 (400 mL), washed with water (2 × 200 mL) during which unreacted material **148** precipitated. The aqueous was filtered and the precipitate was washed with water (50 mL), CH_2Cl_2 (150 mL), ethanol (30 mL) and diethyl ether (2 × 60 mL) and dried over P_2O_5 to give **148** as a white powder (2.41 g. 6.82 mmol, 10%). The combined aqueous fraction was re-extracted (CH_2Cl_2, 150 mL) and combined CH_2Cl_2 layers were washed with sat aq KCl (3 × 150 mL), dried (Na_2SO_4), and concentrated *in vacuo* and dried under high vacuum to give a yellow foam. Following purification by column chromatography (50-95% ethyl acetate/hexane + 0.5% Et_3N → 0.5-1.0% methanol/ethyl acetate + 0.5% Et_3N → 10-15% methanol/ethyl acetate + 0.5-1.0% Et_3N), and drying under high vacuum, the desired mono-tritylated product **97**, the *bis*-tritylated by-product, and a mixture thereof, were afforded as a white foam (37.1 g, 56.5 mmol, 80%), a yellow foam/solid (3.41 g, 3.56 mmol, 5%) and a pale yellow foam (2.19 g) respectively.

R_f 0.19 (**97**), 0.58 (***bis***) (5:95 methanol/CH_2Cl_2 + 0.3% Et_3N, A'); 0.39 (**97**), 0.66 (***bis***) (1:9 methanol/CH_2Cl_2 + 0.3% Et_3N, A'); 0.29 (1:9 acetone/CH_2Cl_2 + 0.7% pyridine, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.67 (1H, N**H**^{3}), 8.01 (1H, s, **H**^{6}), 7.40 (2H, d, J = 7.6 Hz, **H**^{13}), 7.32 (2H, t, J = 7.3 Hz, **H**^{14}), 7.29 (4H, d, J = 9.0 Hz, **H**^{9}), 7.23 (1H, t, J = 7.2 Hz, **H**^{15}), 6.90 (4H, d, J = 8.8 Hz, **H**^{10}), 6.10 (1H, t, J = 6.8 Hz, **H**$^{1'}$), 5.29 (1H, br s, 3'-O**H**), 4.23 (1H, br s, **H**$^{3'}$), 3.90 (1H, dd, J = 3.6, 7.5 Hz, **H**$^{4'}$), 3.74 (6H, s, OC**H**$_3$), 3.20 (1H, dd, J = 4.8, 10.8 Hz, **H**$^{5'}$), 3.16 (1H, dd, J = 3.3, 10.6 Hz, **H**$^{5'}$), 2.25 (1H, td, J = 6.8, 13.6 Hz, **H**$^{2'}$), 2.17 (1H, ddd, J = 3.4, 6.2, 13.3 Hz **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 160.5 (**C**^{4}), 158.1 (**C**^{11}-OCH_3), 150.1 (**C**^{2}), 144.7 (**C**^{12}), 144.2 (**C**^{6}), 135.5, 135.4 (**C**^{8}), 129.7 (**C**^{9}), 127.9 (**C**^{14}), 127.6 (**C**^{13}), 126.7 (**C**^{15}), 113.3 (**C**^{10}), 85.8 (**C**^{7}Ar$_3$), 85.8 (**C**$^{4'}$), 84.9 (**C**$^{1'}$), 70.5 (**C**$^{3'}$), 69.8 (**C**^{5}-I), 63.7 (**C**$^{5'}$), 55.0 (O**C**H_3), 39.8 (**C**$^{2'}$); **LRMS** (ES^+) *m*/*z*: 679 ($[M + Na]^+$, 100).

Analytical results consistent with reported data.[61,170,251]

5-Bromo-6-oxocytidine, 99



$C_{15}H_{18}N_3O_9Br$
Mol. Wt.: 464.22

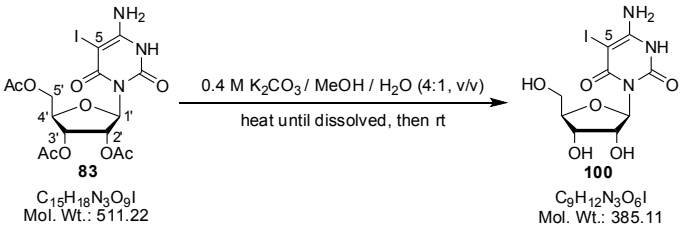
$C_9H_{12}N_3O_6Br$
Mol. Wt.: 338.11

Nucleoside **79** (1.57g, 3.37 mmol) was dissolved with heating in a 0.4 M solution of K_2CO_3 in methanol/water (4:1, v/v) (34.0 mL, 13.6 mmol K_2CO_3, 4.0 eq) over 2 minutes then stirred at rt. After 10 minutes, DOWEX 50Wx8-400 (pyridinium form) was added with stirring until the solution was neutralised (pH 12.0→7.0). The mixture was filtered, the DOWEX washed with methanol (2 × 20 mL), water (2 × 20 mL), and water/methanol (1:1 v/v, 40 mL) and the filtrate was concentrated *in vacuo* to yield a gum. The gum was suspended in ethanol (60 mL) and allowed to stand for 2 days in the fridge. The powder obtained after filtration, washing and drying was hygroscopic and was therefore treated again with DOWEX in methanol/water (1:1 v/v, 25 mL) solution to acidify (pH 8.0→6.0). The DOWEX was washed and filtrate was concentrated *in vacuo* as previously described. The solid was suspended in isopropanol (70 mL) and allowed to stand overnight in the fridge. Following vacuum filtration, the solid was washed with ice-cold isopropanol (10 mL) and ice-cold diethyl ether (2 × 25 mL) and dried under high vacuum over P_2O_5, to afford the desired product **99**, as an off-white chalky solid (1.04 g, 3.06 mmol, 91%).

R_f 0.49 (3:1:1, n-BuOH:AcOH:water, A'); **IR** (solid): ν 3306 (m, N–H), 3207 (m, O–H), 2979 (w, O–H), 2915, 2790 (w, C–H), 1716 (m, C=O), 1604 (vs, C=O), 1538 (s, C=C), 1468, 1444 (m, CH_2), 1372 (m, O–H), 1256, 1172 (w), 1107 (m, C–O), 1037 (s, C–O), 957, 902, 760, 741, 680 (m), 624 (s), 554 (s, C–Br) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.80 (1H, br s, N**H**^{3}), 6.58 (2H, s, N**H$_2$**), 6.02 (1H, d, J = 3.8 Hz, **H**$^{1'}$), 4.96 (1H, br d, J = 4.8 Hz, 2'-O**H**), 4.77 (1H, br d, J = 5.5 Hz, 3'-O**H**), 4.53 (1H, br t, J = 5.5 Hz, 5'-O**H**), 4.46 (1H, br dd, J = 4.8, 9.6 Hz, **H**$^{2'}$), 4.07 (1H, br dd, J = 5.5, 11.1 Hz, **H**$^{3'}$), 3.66 (1H, td, J = 3.3, 6.0 Hz, **H**$^{4'}$), 3.58 (1H, br td, J = 3.6, 11.5 Hz, **H**$^{5'}$), 3.40 (1H, td, J = 6.0, 12.0 Hz, **H**$^{5'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 158.8 (**C**^{6}), 151.1 (**C**^{4}–NH_2), 149.5 (**C**^{2}), 88.1 (**C**$^{1'}$), 84.2 (**C**$^{4'}$), 71.0 (**C**$^{2'}$), 70.1 (**C**$^{3'}$), 69.6 (**C**^{5}), 62.4 (**C**$^{5'}$); **LRMS** (ES^+): m/z: 360/362 ($[M + Na]^+$, 1:1, 100); (ES^-): m/z: 336/338 ($[M – H]^-$, 1:1,

100); **HRMS** (ES^+): calcd for $C_9H_{12}N_3O_6Br$ (M), $[M + Na]^+$ = 359.9802/361.9787, found 359.9802/361.9773 (1:1). Found also 301.1405.

5-Iodo-6-oxocytidine, 100

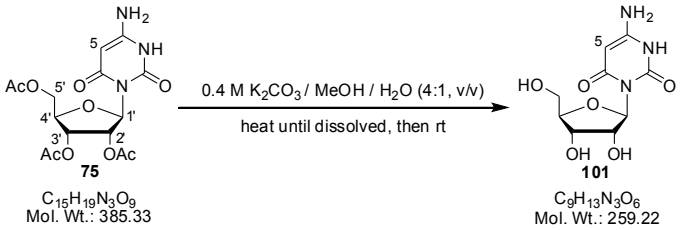


Nucleoside **83** (1.01 g, 1.98 mmol) was dissolved with heating in a 0.4 M solution of K_2CO_3 in methanol/water (4:1, v/v) (17.0 mL, 6.8 mmol K_2CO_3, 3.4 eq) over 1 minute then stirred at rt. After 20 minutes, DOWEX 50WX8-400 (pyridinium form) was added with stirring until the solution was slightly acidified (pH 11.0→5.5). The mixture was filtered, the DOWEX was washed with methanol (8 mL), water (8 mL), then methanol again (20 mL), and the filtrate was concentrated *in vacuo* to yield a white solid. The solid was suspended in ethanol (100 mL) and allowed to stand in the fridge for 2.5 hours. Following vacuum filtration the residue was washed with the filtrate, ethanol (30 mL) and diethyl ether (2 × 20 mL), and dried under high vacuum over P_2O_5, to afford the product **100** as a white, chalky solid (0.59 g, 1.53 mmol, 77%).

$\mathbf{R}_f$ 0.56 (3:1:1, n-BuOH:AcOH:water, A'); **IR** (solid): ν 3368 (br m, N–H, O–H), 3297, 3191 (br m, O–H), 2923 (w, C–H), 1703, 1614 (s, C=O), 1542 (s, C=C), 1422 (s), 1368 (m, O–H), 1102, 1037 (s, C=O), 982, 904, 850 (m), 758, 706, 676, 601 (s), 564 (vs), 524 (s, C–I) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO) δ 10.73 (1H, br s, N**H**^{3}), 6.42 (2H, s, N**H**$_2$), 6.03 (1H, d, *J* = 3.8 Hz, **H**$^{1'}$), 4.96 (1H, br d, *J* = 4.9 Hz, 2'-O**H**), 4.77 (1H, br d, *J* = 5.7 Hz, 3'-O**H**), 4.53 (1H, br t, *J* = 5.4 Hz, 5'-O**H**), 4.44 (1H, br dd, *J* = 4.9, 9.8 Hz, **H**$^{2'}$), 4.06 (1H, br dd, *J* = 5.6, 11.3 Hz, **H**$^{3'}$), 3.65 (1H, td, *J* = 3.5, 6.2 Hz, **H**$^{4'}$), 3.58 (1H, br td, *J* = 3.7, 11.6 Hz, **H**$^{5'}$), 3.40 (1H, td, *J* = 5.9, 11.8 Hz, **H**$^{5'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 160.0 (**C**^{6}), 153.1 (**C**^{4}–NH_2), 150.0 (**C**^{2}), 88.3 (**C**$^{1'}$), 84.2 (**C**$^{4'}$), 71.0 (**C**$^{2'}$), 70.1 (**C**$^{3'}$), 62.4 (**C**$^{5'}$), 42.3 (**C**^{5}); **LRMS** (ES^+): *m*/*z* : 430 ($[M – H + 2Na]^+$, 26), 408 ($[M + Na]^+$, 100); ($ES^–$): *m*/*z* : 384 ($[M – H]^–$, 100); **HRMS** (ES^+): calcd for $C_9H_{12}N_3O_6I$ (M),

$[2M + H]^+$ = 770.9620, found 770.9570, $[M + Na]^+$ = 407.9663, found 407.9654, $[M + H]^+$ = 385.9844, found 385.9834.

6-Oxocytidine, 101



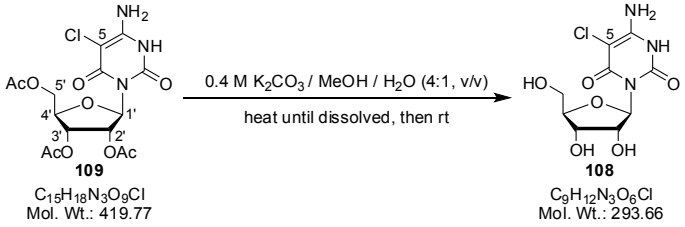
Nucleoside **75** (1.55 g, 4.03 mmol) was dissolved with heating in a 0.4M solution of K_2CO_3 in methanol/water (4:1 v/v) (40.0 mL, 16.0 mmol K_2CO_3, 4.0 eq) over 3 minutes, then stirred at rt for 5 minutes. DOWEX 50WX8-400 (pyridinium form) was added with stirring until the solution was slightly acidified (pH 12.0→5.5). The mixture was filtered, the DOWEX washed with the filtrate, methanol/water (1:1 v/v, 60 mL) then water (20 mL), and the combined filtrate was concentrated *in vacuo*, co-evaporating with methanol/toluene (×5) to yield a sticky white foam. The residue was suspended in isopropanol (50 mL) and allowed to stand in the fridge overnight. The mixture was vacuum filtered, washed with ice-cold isopropanol (25 mL) and diethyl ether (2 × 50 mL), and dried under high vacuum over P_2O_5, to afford the product **101** as a powdery white solid (0.91 g) containing 37.7 mol % potassium acetate (recalculated yield - 2.52 mmol, 71%).

R_f 0.13 (8:2:1, CH_2Cl_2/methanol/acetic acid, A'), 0.11 (1:4, methanol/CH_2Cl_2, A'); **IR** (solid): ν 3334 (br w, O–H, N–H), 3191 (br w, O–H), 1704 (m, C=O), 1603 (s, C=O), 1404 (br m, O–H), 1290 (m), 1200 (w), 1098 (m), 1020 (br s, C–O), 896, 781, 768 (m), 640 (br s), 533 (vs) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 7.18 (2H, br s, N**H_2**), 6.06 (1H, d, *J* = 4.0 Hz, **$H^{1'}$**), 4.47 (1H, dd, *J* = 4.3, 5.8 Hz, **$H^{2'}$**), 4.44 (1H, s, **H^5**), 4.08 (1H, t, *J* = 6.0 Hz, **$H^{3'}$**), 3.65 (1H, td, *J* = 2.8, 8.5 Hz, **$H^{4'}$**), 3.56 (1H, dd, *J* = 2.9, 11.7 Hz, **$H^{5'}$**), 3.40 (1H, dd, *J* = 5.4, 11.7 Hz, **$H^{5'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 163.3 (**C^6**), 156.8 (**C^4**–NH_2), 152.7 (**C^2**), 86.8 (**$C^{1'}$**), 84.1 (**$C^{4'}$**), 73.7 (**C^5**), 71.1 (**$C^{2'}$**), 70.3 (**$C^{3'}$**), 62.5 (**$C^{5'}$**); **LRMS** (ES^+) *m*/*z*: 282 ($[M + Na]^+$, 100); (ES^-) *m*/*z*: 356 ($[M + H_3PO_4 - H]^-$, 53), 258 ($[M - H]^-$, 100); **HRMS** (ES^+): calcd for $C_9H_{13}N_3O_6$ (M), $[3M + Na]^+$ = 800.2311,

found 800.2285, $[2M + Na]^+$ = 541.1506, found 541.1497, $[M + Na]^+$ = 282.0697, found 282.0700.

Analytical results consistent with reported data.[71,248]

5-Chloro-6-oxocytidine, 108



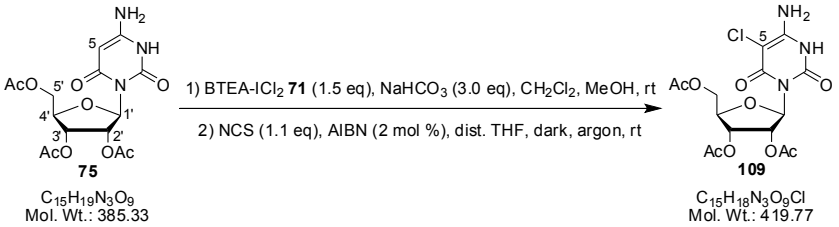
Nucleoside **109** (1.42 g, 3.38 mmol) was dissolved with heating in a 0.4 M solution of K_2CO_3 in methanol/water (4:1 v/v) (33.9 mL, 13.5 mmol K_2CO_3, 4.0 eq) over 3 minutes then stirred at rt. After 5 minutes, DOWEX 50WX8-400 (pyridinium form) was added with stirring, until the solution was slightly acidified (pH 12.0→5.5). The mixture was filtered, the DOWEX washed with methanol/water (1:1 v/v, 2 × 50 mL) then water (50 mL), and the combined filtrate was concentrated *in vacuo* to dryness. The residue was suspended in ethyl acetate (60 mL) and allowed to stand in the fridge overnight. The mixture was vacuum filtered, washed with ethyl acetate (30 mL), ethyl acetate/isopropanol (2:1 v/v, 30 mL) and diethyl ether (2 × 30 mL), and dried under high vacuum over P_2O_5, to afford a powdery white solid (1.24 g).

The solid was retreated with 0.4 M solution of K_2CO_3 in methanol/water (4:1 v/v) (12.7 mL, 5.08 mmol K_2CO_3, 1.5 eq) for 15 minutes, then DOWEX 50WX8-400 (pyridinium form), and concentrated *in vacuo* to dryness. Following suspension in isopropanol (35 mL) overnight, filtration, washing and drying under high vacuum, over P_2O_5, the product **108** was afforded as a white powdery solid (1.07 g) containing. 35.6 mol % potassium acetate and 1.8 mol % succinimide (recalculated yield - 3.05 mmol, 90%).

$\mathbf{R}_f$ 0.52 (3:1:1, n-propanol/conc. aq NH_3/water, A'); **IR** (solid) ν 3306 (br w, O–H, N–H), 3213 (br w, O–H), 2980 (br w, C–H), 1714 (m, C=O), 1613 (br vs, C=O), 1434 (m), 1322, 1267 (w), 1113 (s), 1051, 1023 (s, C–O), 962, 897, 861 (m), 756 (s), 658 (s, C–Cl), 557 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 6.69 (2H, v br s, N**H_2**), 6.12 (1H, d,

J = 4.3 Hz, **H**$^{1'}$), 4.73 (3H, extremely br s, O**H**), 4.53 (1H, dd, J = 4.5, 6.0 Hz, **H**$^{2'}$), 4.13 (1H, t, J = 5.8 Hz, **H**$^{3'}$), 3.67 (1H, dt, J = 2.8, 4.9 Hz, **H**$^{4'}$), 3.56 (1H, dd, J = 2.6, 11.7 Hz, **H**$^{5'}$), 3.41 (1H, dd, J = 4.2, 11.6 Hz, **H**$^{5'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 159.8 (**C**^{6}), 159.2 (**C**^{4}–NH_2), 154.0 (**C**^{2}), 87.9 (**C**$^{1'}$), 84.3 (**C**$^{4'}$), 80.5 (**C**^{5}), 71.1 (**C**$^{3'}$), 70.4 (**C**$^{2'}$), 62.5 (**C**$^{5'}$); **LRMS** (ES^+) m/z: 316/318 ($[M + Na]^+$, 3:1, 100); (ES^-) m/z: 585/587/589 ($[2M - H]^-$, 3:2:1, 9). 292/294 ($[M - H]^-$, 3:1, 100); **HRMS** (ES^+): calcd for $C_{15}H_{18}N_3O_9Cl$ (M), $[2M + Na]^+$ =609.0721/611.0692, found 609.0702 (2:1), $[M + Na]^+$ = 316.0307/318.0277, found 316.0298 (3:1).

5-Chloro-2',3',5'-tri-*O*-acetyl-6-oxocytidine, 109



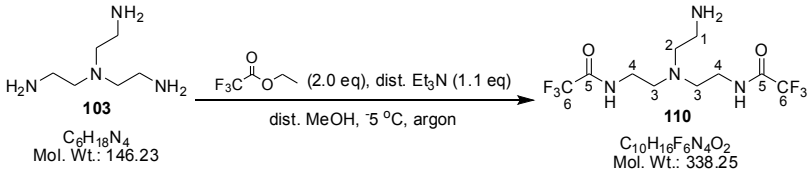
Method 1: To a stirred solution of 2',3',5'-tri-*O*-acetyl-6-oxocytidine **75** (0.39 g, 1.00 mmol) in CH_2Cl_2 (7 mL) and methanol (5 mL), was added sodium bicarbonate (0.25 g, 3.02 mmol, 3.0 eq), followed by a solution of iodinating agent benzyltriethylammonium dichloroiodate **71** (0.59 g, 1.51 mmol, 1.5 eq) in CH_2Cl_2 (8 mL), and the reaction was stirred vigorously at rt for 30 minutes. The reaction mixture was filtered and concentrated *in vacuo* to yield a brown gum. The residue was redissolved in CH_2Cl_2 (40 mL), washed with sat. aq sodium thiosulfate (30 mL) and brine (2 × 40 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield a yellow foam (466 mg). The compound was twice purified by column chromatography (0-5% methanol/CH_2Cl_2) to afford, after drying under high vacuum, the undesired chlorinated product **109** as a white foam (0.11 g, 0.26 mmol, 26% yield).

Method 2: To a stirred suspension of 2',3',5'-tri-*O*-acetyl-6-oxocytidine **75** (1.52 g, 3.95 mmol) and AIBN (13.3 mg, 0.08 mmol, 2.1 mol %) in distilled THF (15.0 mL), was added NCS (0.56 g, 4.15 mmol, 1.1 eq), and the reaction was stirred in absence of light, under an argon atmosphere, at rt for 18 hours. The thick white slurry was dissolved in acetone/methanol (1:1 v/v, 100 mL), and concentrated *in vacuo* to dryness.

Following purification by column chromatography (40-60% acetone/toluene), the desired product **109** was afforded as a white powdery solid (1.64 g) containing 29.6 mol % succinimide (recalculated yield - 3.57 mmol, 90%).

R$_f$ 0.37 (5:95, methanol/CH_2Cl_2, A'), 0.36 (1:1, acetone/toluene, A'); **IR** (solid): ν 3419, 3307 (w, N–H), 3198, 3140, 3087, 3024 (w), 2909 (w, C–H), 1741, 1718 (s, C=O), 1629 (s, N–H), 1559 (br s, N–H), 1430 (m), 1372 (s, C–H), 1228 (br vs, C–O), 1155 (m), 1044 (s, C–O), 939, 897, 766, 747 (m), 629 (C–Cl), 601 (m) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.96 (2H, br s, N**H$_2$**), 6.15 (1H, br s, **H$^{1'}$**), 5.64 (1H, dd, *J* = 2.5, 6.5 Hz, **H$^{2'}$**), 5.49 (1H, t, *J* = 7.3 Hz, **H$^{3'}$**), 4.33 (1H, dd, *J* = 3.3, 11.8 Hz, **H$^{5'}$**), 4.11 (1H, ddd, *J* = 3.3, 6.4, 8.0 Hz, **H$^{4'}$**), 4.03 (1H, dd, *J* = 6.3, 11.8 Hz, **H$^{5'}$**), 2.07 (3H, s, C**H$_3$**), 2.03 (3H, s, C**H$_3$**), 2.00 (3H, s, C**H$_3$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.0, 169.5, 169.3 (**C**OCH$_3$), 158.0 (**C^6**), 150.5 (**C^4**–NH$_2$), 148.7 (**C^2**), 85.6 (**C$^{1'}$**), 80.3 (**C^5**–Cl), 77.9 (**C$^{4'}$**), 72.5 (**C$^{2'}$**), 69.7 (**C$^{3'}$**), 62.9 (**C$^{5'}$**), 20.4, 20.3, 20.2 (CO**C**H$_3$); **LRMS** (ES$^+$) *m*/*z*: 861/863/865 ([2M + Na]$^+$, 6:5:2, 11), 442/444 ([M + Na]$^+$, 3:1, 100), 420/422 ([M + H]$^+$, 3:1, 26); (ES$^-$) *m*/*z*: 418/420 ([M – H]$^-$, 3:1, 100); **HRMS** (ES$^+$): calcd for $C_{15}H_{18}N_3O_9Cl$ (M), [2M + Na]$^+$ = 861.1355/863.1326/865.1296, found 861.1374 (6:4:1), [M + Na]$^+$ = 442.0624/444.0594, found 442.0627 (3:1).

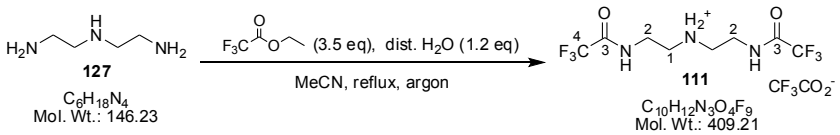
N-(2-aminoethyl)-*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine, 110



To a stirred solution of *tris*(2-aminoethyl)amine **103** (0.50 mL, 3.34 mmol) in distilled methanol (4.0 mL) and distilled Et_3N (0.51 mL, 3.66 mmol, 1.1 eq), under an argon atmosphere, at ⁻5 °C, was added ethyl trifluoroacetate (0.80 mL, 6.72 mmol, 2.0 eq) dropwise over 30 minutes. The mixture was stirred for a further 2.5 hours, then concentrated *in vacuo* and dried under high vacuum co-evaporating with methanol/toluene. Following purification by column chromatography (20-40% acetone/ ethyl acetate + 0.5% Et_3N), *bis*-trifluoroacetamide **110** was afforded, following drying under high vacuum, as a viscous, pale yellow oil (422 mg, 1.25 mmol, 37%).

$\mathbf{R}_f$ 0.21 (2:3, acetone/ethyl acetate + 0.3% Et_3N, D'), 0.15 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, D'), **^{1}H NMR** (300 MHz, d_6-DMSO) δ 9.12 (2H, extremely br s, N**H**$COCF_3$), 3.25 (4H, t, J = 6.4 Hz, **H^3**), 3.18 (2H, app. t, J = 7.0 Hz, **H^2**), 2.65 (4H, q, J = 6.6 Hz, **H^4**), 2.59 (2H, t, J = 6.6 Hz, **H^1**), 1.88 (2H, t, J = 1.3 Hz, N**H$_2$**); **LRMS** (ES^+) *m*/*z*: 379 (100), 339 ([M + H]$^+$, 13); **HRMS** (ES^+): calcd for $C_{10}H_{16}N_4O_2F_6$ (M), [M + H]$^+$ = 339.1250, found 339.1258. Found also 379.1568.

***Bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]ammonium trifluoroacetate, 111**[182,192]



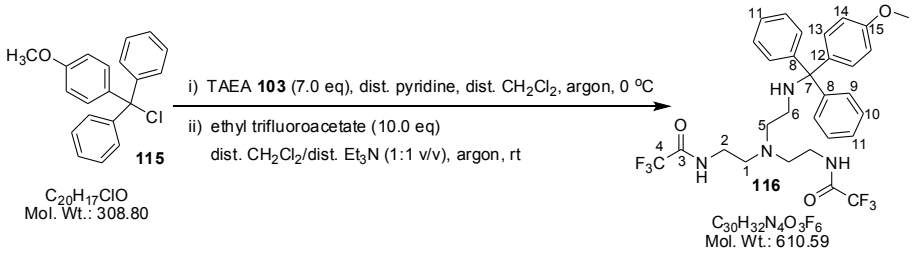
To a stirred solution of diethylene triamine **127** (3.20 mL, 29.6 mmol) in HPLC grade acetonitrile (50.0 mL), was added distilled water (0.64 mL, 35.5 mmol, 1.2 eq), followed by slow addition of ethyl trifluoroacetate (12.3 mL, 104 mmol, 3.5 eq). The solution was heated at reflux, under argon pressure at 105 °C for 16 hours. The reaction mixture was concentrated *in vacuo* and dried under high vacuum co-evaporating with methanol (2 × 50 mL). The syrup was soaked in CH_2Cl_2 (100 mL) overnight, then suspended and the resulting 'gel' was filtered. The solid was washed with CH_2Cl_2 (2 × 100 mL) and dried under high vacuum over P_2O_5, to afford the product salt **111** as an off-white, papery solid (11.3 g, 27.6 mmol, 93%).

$\mathbf{R}_f$ 0.44 (1:1, acetone/ethyl acetate + 0.3% Et_3N, C'), 0.21 (3:7, acetone/CH_2Cl_2 + 0.3% Et_3N, C'), 0.31 (1:1, acetone/CH_2Cl_2 + 0.3% Et_3N, C'); **IR** (solid): ν 3282 (br w, N–H, O-H), 3118 (w, N–H), 2838 (w, N^+–H), 1717, 1673 (s, C=O), 1558 (m, N–H, C=O), 1445 (w, C–H), 1175 (s, C–F), 1142 (vs, C–F), 1015, 837, 797 (m), 724 (s, CH_2), 694 (br m, C–F) cm^{-1}; **^{19}F NMR** (282 MHz, d_6-DMSO) δ -73.5 (C**F$_3$**CO_2^-), -74.2 (NHCOC**F$_3$**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.60 (2H, br s, N**H**$COCF_3$), 8.80 (2H, v br s, R_2N**H$_2^+$**), 3.50 (4H, br dd, J = 5.0, 9.5Hz, **H^2**), 3.13 (4H, t, J = 6.4 Hz, **H^1**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 159.2 (q, J = 31.6 Hz, CF_3**C**O_2^-), 157.3 (q, J = 36.5 Hz, NH**C**OCF_3), 117.5 (q, J = 298.9 Hz, **C**$F_3CO_2^-$), 116.2 (q, J = 287.7 Hz, NHCO**C**F_3), 45.8 (**C^1**), 36.3 (**C^2**); **LRMS** (ES^+) *m*/*z*: 296 ([M + H]$^+$, 100); (ES^-) *m*/*z*: 522 ([M.CF_3CO_2H + CF_3CO_2H – H]$^-$, 7), 408 ([M.CF_3CO_2H – H]$^-$, 100), 392 ([M + H_3PO_4 –

H]$^+$, 56), 294 ([M – H]$^-$, 59); **HRMS** (ES$^+$): calcd $C_8H_{11}N_3O_2F_6$ (M), [M + H]$^+$ = 296.0828, found 296.0823.

Analytical results consistent with reported data.[192]

N-[2-*N'*-(4-Methoxytritylamino)ethyl]-*bis*-*N*,*N*-(2-ethyl trifluoroacetamide), 116



To a stirred solution of *tris*(2-aminoethyl)amine (TAEA) **103** (1.79 mL, 11.9 mmol, 7.0 eq) in distilled CH_2Cl_2 (4.0 mL), under an argon atmosphere at 0 °C, was added very slowly dropwise in portions, a solution of MMTCl **115** (0.53 g, 1.70 mmol) in distilled pyridine (3.5 mL), over 5.5 hours. The reaction was quenched with methanol (2 mL), stirring for 5 minutes, then concentrated *in vacuo*, co-evaporating with CH_2Cl_2 (×5) then under high vacuum, co-evaporating with methanol/toluene (×3). The wet bottle-green powder was washed with hexane (2 × 50 mL) and CH_2Cl_2 (4 × 25 mL). The resulting blue solid was washed a third time with CH_2Cl_2 (4 × 25 mL) and all combined organic fractions were concentrated *in vacuo* and under high vacuum to give a brown/green oil. The oil was dissolved in CH_2Cl_2 (100 mL), washed with water (100 mL), the aqueous was re-extracted (CH_2Cl_2, 3 × 50 mL), and the organic fraction was dried (Na_2SO_4), and concentrated *in vacuo* and under high vacuum, to afford the crude mono-MMT protected polyamine, as a pale bottle green gum. (0.61 g).

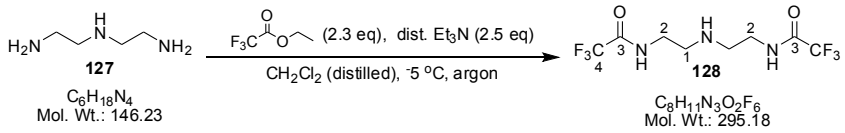
R$_f$ 0.35 (5:1:1, ethyl acetate/methanol/aq NH_3, C');

To a stirred solution of the green resin (0.40 g) in distilled CH_2Cl_2 (5.0 mL) and distilled Et_3N (5.00 mL), under an argon atmosphere, at rt, was added, ethyl trifluoroacetate (1.15 mL, 9.66 mmol, 10.0 eq) dropwise, and the reaction was stirred for 3 hours. The reaction was diluted with CH_2Cl_2 (100 mL), washed with water (2 × 50 mL), the aqueous was re-extracted (CH_2Cl_2, 50 mL). The combined organic fraction was washed with sat. aq KCl (50 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give a

yellow/brown syrup. The compound was twice-purified by column chromatography (1% methanol/CH_2Cl_2 + 0.5% Et_3N; toluene + 0.5% Et_3N → 1.5% methanol/CH_2Cl_2 + 0.5% Et_3N) to afford the product **116** as a pale yellow foam (0.14 g, 0.23 mmol, 20% over two steps).

R_f 0.53 (3% methanol/CH_2Cl_2 + 0.3% Et_3N, C', E'); **IR** (solid) ν 3303 (br w, water), 3057 (br w, N–H), 2952, 2837 (br w, C–H), 1703 (s, C=O), 1607, 1556 (w, Aryl C–H), 1508 (m, Aryl C–H), 1448 (m), 1297 (w), 1248 (m), 1206 (s, C–F), 1148 (br vs, C–F), 1032 (m), 903 (w), 830, 768, 750 (m, Aryl C–H), 703 (s, Aryl C–H), 583 (m) cm^{-1}; **^{19}F NMR** (282 MHz, d_6-DMSO) δ $^{-}$74.1 (COC**F$_3$**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.20 (1H, br s, N**H**COCF_3), 7.38 (4H, d, J = 8.0 Hz, **H^{9}**), 7.29-7.23 (2H, m, **H^{13}**), 7.27 (4H, d, J = 8.5 Hz, **H^{10}**), 7.16 (2H, t, J = 7.3 Hz, **H^{11}**), 6.83 (2H, d, J = 9.0 Hz, **H^{14}**), 3.71 (3H, s, OC**H$_3$**), 3.31 (1H, s, N**H**MMT), 3.19 (4H, br dd, J = 6.3, 10.8 Hz, **H^{2}**), 2.58 (2H, t, J = 6.5 Hz, **H^{6}**), 2.48 (4H, t, J = 6.5 Hz, **H^{1}**), 2.07 (2H, dd, J = 6.0, 12.6 Hz, **H^{5}**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 157.3 (**C^{15}**), 156.2 (q, J = 36.0 Hz, **C**OCF_3), 146.5 (**C^{8}**), 138.1 (**C^{12}**), 129.5 (**C^{13}**), 128.2 (**C^{9}**), 127.6 (**C^{10}**), 125.9 (**C^{11}**), 115.9 (q, J = 288.2 Hz, CO**C**F_3), 113.0 (**C^{14}**), 69.7 (**C^{7}**Ar_3), 54.9 (O**C**H_3), 54.0 (**C^{6}**), 52.1 (**C^{1}**), 41.1 (**C^{5}**), 37.3 (**C^{2}**); **LRMS** (ES^-) *m/z*: 723 ([M + F_3CCO_2H – H]$^-$, 38), 672 (21), 609 ([M – H]$^-$, 13), 145 (100); **HRMS** (ES^+): calcd for $C_{30}H_{32}N_4O_3F_6$ (M), [M + Na]$^+$ = 633.2271, found 633.2274.

***Bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine, 128**[194,252]

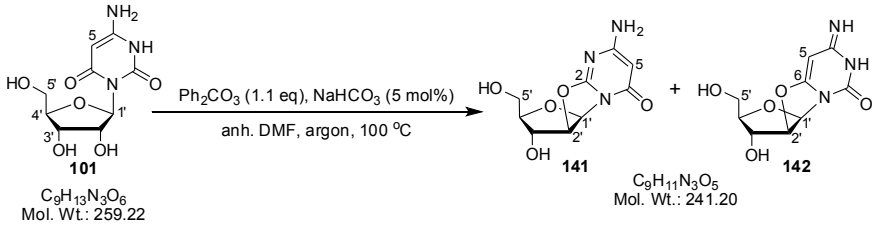


To a stirred solution of diethylene triamine **127** (2.00 mL, 18.5 mmol) in distilled CH_2Cl_2 (40.0 mL), and distilled Et_3N (6.45 mL, 46.3 mmol, 2.5 eq), under an argon atmosphere, at $^{-}$5 °C, ethyl trifluoroacetate (5.07 mL, 42.6 mmol, 2.3 eq), was added dropwise over 1 hour. After stirring for a further 1¼ hours, the reaction mixture was concentrated *in vacuo* and redissolved in argon-degassed CH_2Cl_2 (100 mL). The solution was washed with degassed sat. aq KCl (2 × 50 mL), and the combined aqueous fraction was re-extracted (degassed CH_2Cl_2, 2 × 30 mL). The combined CH_2Cl_2

fractions were dried (Na_2SO_4), concentrated *in vacuo* and dried under high vacuum, to afford the product *bis*-trifluoroacetamide **128**, as an off-white, slightly air-sensitive solid (5.13 g, 17.4 mmol, 94%).

R_f 0.27 (3:7, acetone/CH_2Cl_2 + 0.3% Et_3N, C'), 0.35 (1:1, acetone/CH_2Cl_2 + 0.3% Et_3N, C'); **IR** (solid): ν 3287 (w, N–H), 2945, 2850 (w, C–H), 1704 (s, C=O), 1563 (m, N–H, C=O), 1435 (m, C–H), 1349 (w), 1183, 1140 (br vs, C–F), 950 (w), 866, 843 (m), 724 (s, CH_2), 681 (br m, C–F) cm^{-1}; **^{19}F NMR** (282 MHz, d_6-DMSO) δ -74.1 (COC**F$_3$**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.25 (2H, v br s, N**H**COCF_3), 3.24 (4H, t, J = 6.4 Hz, **H^2**), 2.64 (4H, t, J = 6.5 Hz, **H^1**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 156.3 (q, J = 36.0 Hz, NH**C**OCF_3), 115.9 (q, J = 287.7 Hz, NHCO**C**F_3), 47.1 (**C^1**), 39.3 (**C^2**); **LRMS** (ES^+) *m*/*z*: 296 ($[M + H]^+$, 100); (ES^-) *m*/*z*: 408 ($[M + CF_3CO_2H - H]^-$, 20), 392 ($[M + H_3PO_4 - H]^-$, 100), 294 ($[M - H]^-$, 77); **HRMS** (ES^+): calcd for $C_8H_{11}N_3O_2F_6$ (M), $[M + Na]^+$ = 318.0648, found 318.0672; $[M + H]^+$ = 296.0828, found 296.0830.
Analytical results consistent with reported data.[252]

2,2'-Anhydro-6-oxocytidine and 6,2'-anhydro-6-oxocytidine, 141,142

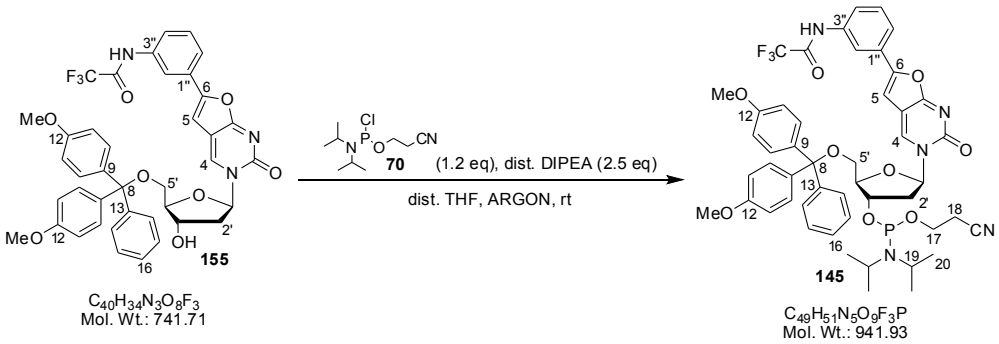


To a stirred solution of 6-oxocytidine **101** (0.20 g, 0.60 mmol) in anhydrous DMF (2.50 mL), under an argon atmosphere, was added diphenyl carbonate (0.14 g, 0.66 mmol, 1.1 eq) and sodium bicarbonate (3.0 mg, 0.04 mmol, 6.0 mol %) and the mixture was stirred at 100 °C for 3 hours. The reaction mixture was allowed to cool to rt, then diethyl ether (35 mL) was slowly added with rapid stirring, and the suspension was stirred for 30 minutes. The ether was decanted, and the precipitate was washed with ether (2 × 15 mL) and dried under high vacuum to afford the crude product as a hygroscopic, powdery solid. The solid was redissolved in water/methanol (2:1, v/v, 30 mL) and concentrated *in vacuo* and under high vacuum. The gum was redissolved in DMF (2 mL) and diethyl ether (25 mL) was added slowly with vigorous stirring. After stirring

for 30 minutes, the ether was decanted, the solid was washed with diethyl ether (2 × 75 mL) and dried *in vacuo* and under high vacuum to give the product as an isomeric mixture **141,142** (*ca.* 1:1), as an off-white chalky solid (0.14 g, 0.59 mmol, 98%).

R_f 0.13, 0.23 (5:1:1, ethyl acetate/methanol/conc. aq NH_3, A'); **LRMS** (ES^+) *m*/*z*: 280 ($[M + Na]^+$, 24), 264 ($[M + Na]^+$, 100).

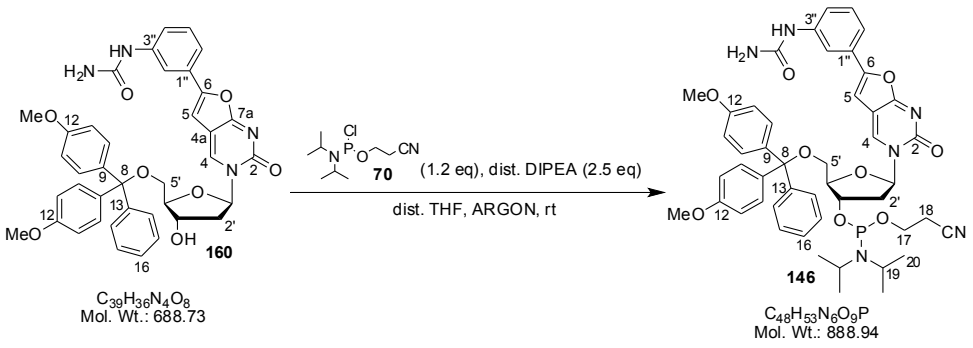
3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-[3-(trifluoroacetamido)phenyl]-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite, 145[206]



To a solution of nucleoside **155** (0.49 g, 0.66 mmol) in distilled THF (4.0 mL) and distilled DIPEA (0.29 mL, 1.67 mmol, 2.5 eq), strictly under an argon atmosphere and excluding moisture, was added chloro-phosphitylating reagent **70** (0.18 mL, 0.81 mmol, 1.2 eq) dropwise, and the reaction was stirred at rt for 80 minutes. The reaction was diluted with argon-degassed ethyl acetate (40 mL), then washed, under an argon atmosphere, with degassed sat aq KCl (3 × 10 mL). The organic fraction was dried (Na_2SO_4), the inorganics were washed with degassed ethyl acetate (3 × 10 mL) and the combined fraction was transferred under argon and concentrated *in vacuo* to give a pale yellow residue. Following purification by column chromatography (degassed 30% ethyl acetate/CH_2Cl_2 + 0.5% Et_3N), on silica gel pre-equilibrated with Et_3N, under argon pressure, and after drying under high vacuum, the product phosphoramidite **145** was afforded, as a diastereomeric mixture (*ca.* 3:2), as an air-sensitive, off-white foam (0.40 g, 0.42 mmol, 64%).

$\mathbf{R}_f$ 0.49, 0.61 (1:1, ethyl acetate/CH_2Cl_2 + 0.3% Et_3N, A'), 0.58 (4:1, ethyl acetate/ CH_2Cl_2 + 0.3% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.2, 149.0 (**P^{III}**); **^{1}H NMR** (300 MHz, d_6-DMSO) δ 11.39 (1H, br s, N**H**$COCF_3$), 8.78 and 8.77 (1H, s, **H^{4}**), 8.05 (1H, t, *J* = 1.5 Hz, **$H^{2"}$**), 7.71 (1H, td, *J* = 1.8, 7.0 Hz, **$H^{4"}$**), 7.55 (1H, br s, **$H^{6"}$**), 7.53 (1H, t, *J* = 7.7 Hz, **$H^{5"}$**), 7.42 and 7.40 (2H, td, *J* = 1.6, 7.1 Hz and *J* = 1.5, 6.6 Hz, **H^{14}**), 7.36-7.19 (3H, m, **H^{15}**, **H^{16}**), 7.28 (4H, d, *J* = 8.8 Hz, **H^{10}**), 6.91 and 6.90 (2H, d, *J* = 9.2 Hz and *J* = 8.8 Hz, **H^{11}**), 6.90 and 6.89 (2H, d, *J* = 9.2 Hz, **H^{11}**), 6.34 and 6.32 (1H, s, **H^{5}**), 6.20 and 6.16 (1H, dd, *J* = 4.2, 6.8 Hz, and *J* = 4.0, 6.6 Hz, **$H^{1'}$**), 4.73-4.57 (1H, m, **$H^{3'}$**), 4.20 and 4.15 (1H, dd, *J* = 3.7, 8.6 Hz and *J* = 3.8, 8.5 Hz, **$H^{4'}$**), 3.81-3.68 and 3.74-3.60 (2H, m, **H^{17}**), 3.71, 3.71 and 3.70 (3H, 3H and 6H, s, O**CH_3**), 3.62-3.45 (2H, m, **H^{19}**), 3.46-3.35 (2H, m, **$H^{5'}$**), 3.36-3.26 (2H, m, **$H^{5'}$**), 2.77 and 2.67 (2H, t, *J* = 5.9 Hz, **H^{18}**), 2.69-2.55 (1H, m, **$H^{2'}$**), 2.55-2.38 (1H, m, **$H^{2'}$**), 1.16 (3H, d, *J* = 7.0 Hz, **H^{20}**), 1.13 (3H, d, *J* = 7.0 Hz, **H^{20}**), 1.12 (3H, d, *J* = 7.0 Hz, **H^{20}**), 1.02 (3H, d, *J* = 6.6 Hz, **H^{20}**); **LRMS** (ES^+) *m*/*z*: 980 ($[M + K]^+$, 41), 964 ($[M + Na]^+$, 100), 942 ($[M + H]^+$, 52). Analytical results consistent with reported data.[206]

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite, 146

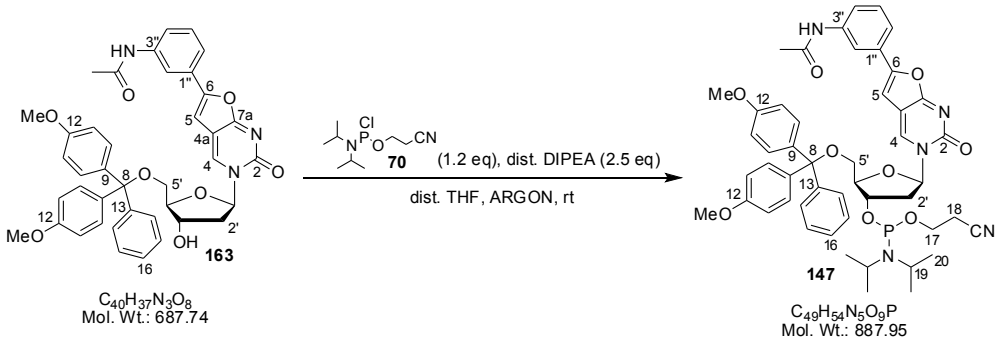


To a solution of nucleoside **160** (1.22 g, 1.77 mmol) in distilled THF (10.0 mL) and distilled DIPEA (0.77 mL, 4.42 mmol, 2.5 eq), strictly under an argon atmosphere and excluding moisture, was added chloro-phosphitylating reagent **70** (0.48 mL, 2.15 mmol, 1.2 eq) dropwise, and the reaction was stirred at rt for 2 hours. Further reagent **70** (0.08 mL, 0.36 mmol, 0.2 eq) was added and the reaction was stirred for 45 minutes. The reaction was quenched with argon-degassed sat aq KCl (1 ml), diluted with degassed

ethyl acetate (60 mL), then washed, under an argon atmosphere, with degassed sat aq KCl (2 × 20 mL). The organic fraction was dried (Na_2SO_4), the inorganics were washed with degassed ethyl acetate (2 × 10 mL) and the combined organic fraction was transferred under argon and concentrated *in vacuo* to give a yellow gum. Following purification by column chromatography (degassed 65% acetone/CH_2Cl_2 + 1.0% Et_3N), on silica gel pre-equilibrated with Et_3N, under argon pressure, then precipitation at rt from distilled CH_2Cl_2 (3 mL) into degassed hexane (150 mL), under an argon atmosphere, the desired product **146** was afforded as a diastereomeric mixture (*ca.* 2:3), as an air-sensitive, powdery white solid (1.08 g, 1.22 mmol, 69%).

R_f 0.33 (65:35, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.44 (9:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.3, 149.1 (**P^{III}**), 15.1 ($HO(H)P^{V}(O)OCH_2CH_2CN$); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 8.75 and 8.74 (1H, s, **H^{4}**), 8.74 (1H, br s, N**H**$CONH_2$), 7.90 (1H, br s, **$H^{2"}$**), 7.42 and 7.40 (2H, d, *J* = 8.5 Hz and *J* = 9.0 Hz, **H^{14}**), 7.38-7.33 (1H, m, **$H^{4"}$**), 7.34 (2H, d, *J* = 7.0 Hz, **H^{15}**), 7.34-7.27 (2H, m, **H^{16}**, **$H^{5"}$**), 7.30 and 7.27 (4H, d, *J* = 9.0 Hz, **H^{10}**), 7.24 (1H, ddd, *J* = 1.8, 3.3, 6.5 Hz, **$H^{6"}$**), 6.91, 6.91 and 6.90 (2H, 2H and 4H, d, *J* = 9.0 Hz, **H^{11}**), 6.21 and 6.17 (1H, dd, *J* = 4.5, 6.5 Hz, **$H^{1'}$**), 6.20 and 6.18 (1H, s, **H^{5}**), 5.93 (2H, br s, CON**H_2**), 4.69 and 4.65 (1H, td, *J* = 6.4, 13.1 Hz and *J* = 6.1, 12.5 Hz, **$H^{3'}$**), 4.19 and 4.15 (1H, td, *J* = 3.5, 5.0 Hz, **$H^{4'}$**), 3.81-3.70 and 3.71-3.62 (2H, m, **H^{17}**), 3.72, 3.72 and 3.71 (3H, 3H and 6H, s, OC**H_3**), 3.63-3.47 (2H, m, **H^{19}**), 3.46 and 3.43 (1H, dd, *J* = 3.4, 10.9 Hz and 3.4, 11.4 Hz, **$H^{5"}$**), 3.43-3.38 and 3.39 (1H, m and dd, *J* = 2.8, 12.1 Hz, **$H^{5'}$**), 2.77 and 2.67 (2H, t, *J* = 6.0 Hz, **H^{18}**), 2.64 and 2.60 (1H, td, *J* = 6.4, 12.8 Hz and *J* = 6.3, 12.4 Hz, **$H^{2"}$**), 2.47-2.39 (1H, m, **$H^{2'}$**), 1.15 (3H, d, *J* = 6.8 Hz, **H^{20}**), 1.13 (3H, d, *J* = 6.8 Hz, **H^{20}**), 1.12 (3H, d, *J* = 6.8 Hz, **H^{20}**), 1.02 (3H, d, *J* = 6.8 Hz, **H^{20}**); **LRMS** (ES^+) *m*/*z*: 927 ($[M + K]^+$, 9), 911 ($[M + Na]^+$, 28), 889 ($[M + H]^+$, 17), 618 (100).

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite, 147

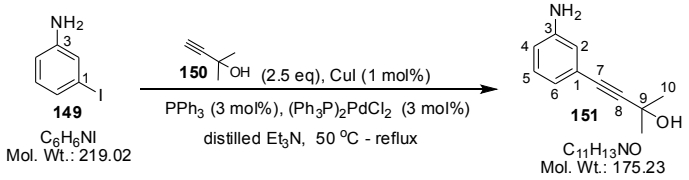


To a solution of nucleoside **163** (1.00 g, 1.46 mmol) in distilled THF (8.0 mL) and distilled DIPEA (0.64 mL, 3.67 mmol, 2.5 eq), strictly under an argon atmosphere and excluding moisture, was added chloro-phosphitylating reagent **70** (0.39 mL, 1.75 mmol, 1.2 eq) dropwise, and the reaction was stirred at rt for 2½ hours. The reaction was quenched with argon-degassed sat aq KCl (2 ml) stirring for 5 minutes, diluted with degassed ethyl acetate (100 mL), then washed, under an argon atmosphere, with degassed sat aq KCl (2 × 20 mL). The organic fraction was dried (Na_2SO_4), the inorganics were washed with degassed ethyl acetate (2 × 15 mL) and the combined fraction was transferred under argon and concentrated *in vacuo* to give a pale yellow foam. Following purification by column chromatography (degassed 10-60% acetone/ CH_2Cl_2 + 0.5% Et_3N), on silica gel pre-equilibrated with Et_3N, under argon pressure, then precipitation twice from distilled CH_2Cl_2 (2 mL) into degassed hexane (200/250 mL) at $^{-}$78 °C/rt, under an argon atmosphere, the desired phosphoramidite **147** was afforded as a diastereomeric mixture (*ca.* 2:3), as an air-sensitive, pale cream foam (0.89 g) containing 18.9 mol % hydrolysed phosphitylating reagent (recalculated yield - 0.97 mmol, 66%).

R$_f$ 0.42, 0.50 (2:3, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.21, 0.30 (1:9, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.3, 149.0 (**P^{III}**), 15.1 (HO(H)P^{V}(O)OCH$_2$CH$_2$CN); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.10 (1H, s, N**H**COCH$_3$), 8.76 and 8.75 (1H, s, **H^{4}**), 8.02 (1H, s, **H$^{2"}$**), 7.59 (1H, br d, *J* = 7.5 Hz, **H$^{4"}$**), 7.44-7.40 (1H, m, **H$^{5"}$**), 7.42-7.37 (2H, m, **H^{14}**), 7.37-7.30 (2H, m, **H^{15}, H$^{6"}$**), 7.30 and 7.27 (4H, d, *J* = 8.5 Hz and *J* = 9.0 Hz, **H^{10}**), 7.30-7.25 (1H, m, **H^{16}**), 6.91, 6.91 and

6.90 (2H, 2H and 4H, d, J = 8.5 Hz, J = 9.0 Hz and J = 9.0 Hz, $\mathbf{H^{11}}$), 6.25 and 6.23 (1H, s, $\mathbf{H^{5}}$), 6.20 and 6.10 (1H, dd, J = 4.3, 6.8 Hz and J = 4.8, 6.3 Hz, $\mathbf{H^{1'}}$), 4.67 and 4.65 (1H, td, J = 6.2, 12.3 Hz and J = 6.0, 12.1 Hz, $\mathbf{H^{3'}}$), 4.19 and 4.15 (1H, td, J = 3.3, 5.0 Hz and J = 3.5, 5.8 Hz, $\mathbf{H^{4'}}$), 3.81-3.69 and 3.73-3.61 (2H, m, $\mathbf{H^{17}}$), 3.72, 3.71 and 3.70 (3H, 3H and 6H, s, OC$\mathbf{H_3}$), 3.62-3.48 (2H, m, $\mathbf{H^{19}}$), 3.49-3.35 (2H, m, $\mathbf{H^{5'}}$), 2.76 and 2.67 (2H, t, J = 6.0 Hz, $\mathbf{H^{18}}$), 2.62 and 2.61 (1H, td, J = 6.4, 12.7 Hz and J = 6.3, 12.6 Hz, $\mathbf{H^{2'}}$), 2.43 and 2.42 (1H, ddd, J = 4.0, 6.8, 13.8 Hz and J = 3.0, 7.0, 13.6 Hz, $\mathbf{H^{2'}}$), 2.08 (3H, s, COC$\mathbf{H_3}$), 1.15 (3H, d, J = 7.0 Hz, $\mathbf{H^{20}}$), 1.13 (3H, d, J = 7.0 Hz, $\mathbf{H^{20}}$), 1.12 (3H, d, J = 6.5 Hz, $\mathbf{H^{20}}$), 1.02 (3H, d, J = 7.0 Hz, $\mathbf{H^{20}}$); **LRMS** (ES^+) m/z: 910 ([M + Na]$^+$, 100), 888 ([M + H]$^+$, 4); (ES^-) m/z: 949 (32), 922/924 ([M + HCl – H]$^-$, 3:1, 33), 886 ([M – H]$^-$, 100).

4-(3-Aminophenyl)-2-methyl-but-3-yn-2-ol, 151[206,253]

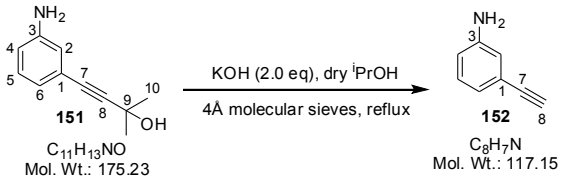


To a solution of 3-iodoaniline **149** (1.00 g, 4.56 mmol) in distilled Et_3N (1.5 mL) under an argon atmosphere, in absence of light, was added CuI (8.7 mg, 0.05 mmol, 1.0 mol %) then 2-methyl-but-3-yn-2-ol **150** (1.10 mL, 11.4 mmol, 2.5 eq), and the mixture was stirred at rt for 5 minutes. Triphenylphosphine (36.0 mg, 0.14 mmol, 3.0 mol %) was added and the reaction was stirred for a further 5 minutes. *Bis*(triphenylphosphine) palladium dichloride (96.2 mg, 0.14 mmol, 3.0 mol %) was added and the reaction was stirred at 50 °C for 20 minutes then refluxed for 50 minutes. The reaction mixture was cooled and concentrated under high vacuum, co-evaporating with methanol/toluene, then the brown syrupy solid was suspended in diethyl ether (70 mL) and allowed to stand in the fridge overnight. The solution was decanted and the solid washed again with diethyl ether (9 × 30 mL), followed by ice-cold ethyl acetate/diethyl ether (1:1 v/v, 100 mL). All fractions were combined, vacuum filtered and the solid was further washed with diethyl ether (30 mL). The combined fractions were concentrated *in vacuo* and redissolved in hot ethyl acetate (10 mL). The hot solution was added dropwise to diethyl ether at –70 °C with vigorous stirring. The suspension was vacuum filtered, the

solid washed with very cold diethyl ether (150 mL) and the filtrate was concentrated *in vacuo*. The resultant syrup was dissolved in CH_2Cl_2 (70 mL) and washed with water (3 × 100 mL). The aqueous fraction was re-extracted (CH_2Cl_2, 200 mL), and combined CH_2Cl_2 fractions were washed with sat aq KCl (100 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give an orange/brown gum. Following purification by column chromatography (0–30% CH_2Cl_2/diethyl ether), and drying under high vacuum, the product alkyne **151** was isolated as a pale yellow solid (0.43 g, 2.43 mmol, 53%).

$\mathbf{R}_f$ 0.29 (1:1, ethyl acetate/hexane, A'), 0.31 (1:1, CH_2Cl_2/diethyl ether, A'), 0.04 (CH_2Cl_2, A'), 0.36 (diethyl ether, A'); **^{1}H NMR** (300 MHz, d_6-DMSO) δ 6.97 (1H, t, J = 7.9 Hz, **H^5**), 6.58 (1H, dd, J = 1.3, 2.0 Hz, **H^2**), 6.53 (1H, ddd, J = 1.1, 2.6, 8.1 Hz, **H^4**), 6.50 (1H, td, J = 1.3, 7.4 Hz, **H^6**), 5.37 (1H, s, O**H**), 5.14 (2H, br s, N**H$_2$**), 1.44 (6H, s, C**H$_3$**); **^{13}C NMR** (75 MHz, d_6-DMSO) δ 148.7 (**C^3**–NH_2), 129.0 (**C^5**), 122.8 (**C^1**), 118.6 (**C^6**), 116.3 (**C^2**), 114.0 (**C^4**), 94.6 (ArC≡**C^8**R), 81.1 (Ar**C^7**≡CR), 63.5 (**C**$(CH_3)_2$OH), 31.7 (**C**H_3); **LRMS** (ES^+): *m*/*z*: 277 ($[M + Et_3N + H]^+$, 10), 217 ($[M + CH_3CN + H]^+$, 35), 176 ($[M + H]^+$, 37), 102 ($[Et_3N + H]^+$, 100).
Analytical results consistent with reported data.[253]

3-Aminophenyl acetylene (or 3-ethynylaniline), 152[206]

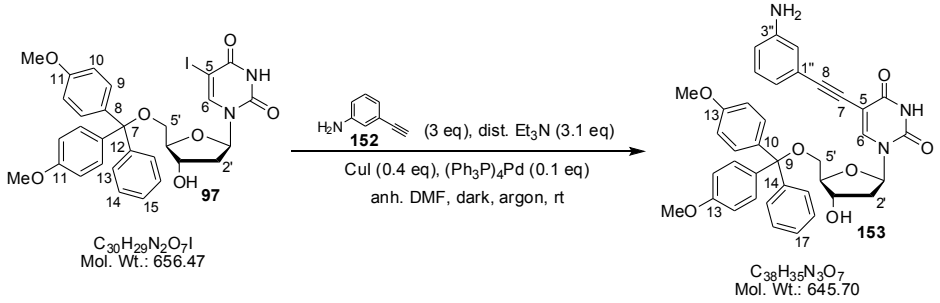


Potassium hydroxide (0.21 g, 3.80 mmol, 2.3 eq) was dissolved in dry, refluxing isopropanol (5 mL) with 4Å molecular sieves. Immediately on dissolution, protected alkyne **151** (0.29 g, 1.63 mmol) was added and the mixture was refluxed for 2½ hours. Further KOH (0.10 g, 1.82 mmol, 1.1 eq) was added and the reaction was refluxed for 1 hour. The mixture was cooled and concentrated to dryness under high vacuum at rt. Dichloromethane (40 mL) was added and the mixture was vacuum filtered. The residue was washed with CH_2Cl_2 (2 × 20 mL), and the filtrate was concentrated to dryness *in vacuo* and under high vacuum to give a tan-coloured solid. Following purification by column chromatography (0–10% diethyl ether/CH_2Cl_2), and drying under high vacuum,

the product alkyne **152** was afforded as a pale orange/brown oil (0.08 g, 0.64 mmol, 39%).

R_f 0.23 (CH_2Cl_2, A'), 0.47 (1:1, diethyl ether/CH_2Cl_2, A'), 0.68 (1:9, methanol/CH_2Cl_2, A'); **^{1}H NMR** (300 MHz, $CDCl_3$) δ 7.10 (1H, t, *J* = 8.1 Hz, **H^5**), 6.90 (1H, td, *J* = 1.2, 7.7 Hz, **H^6**), 6.81 (1H, dd, *J* = 1.5, 2.2 Hz, **H^2**), 6.67 (1H, ddd, *J* = 1.1, 2.6, 8.1 Hz, **H^4**), 3.60 (2H, v br s, 3-N**H$_2$**), 3.02 (1H, s, C≡C**H^8**); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 146.4 (**C^3**–NH_2), 129.4 (**C^5**), 122.9 (**C^1**), 122.6 (**C^6**), 118.4 (**C^2**), 115.9 (**C^4**), 84.0 (**C^7**≡CH), 76.6 (C≡**C^8**H); **LRMS** (EI^+): calcd 177.1 (M^+), found 117.0 (M^+).
Analytical results consistent with reported data.[254]

5-[(3-Aminophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine, 153[206]

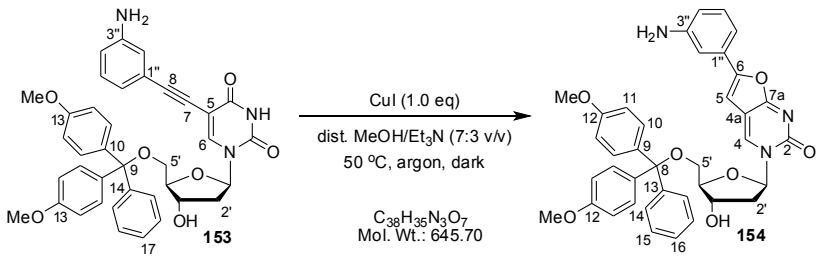


To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **97** (2.32 g, 3.53 mmol) in anhydrous DMF (20.0 mL), under an argon atmosphere, in absence of light, was added CuI (0.27 g, 1.41 mmol, 0.4 eq), distilled Et_3N (1.50 mL, 10.8 mmol, 3.1 eq) then 3-ethynylaniline **152** (1.19 mL, 10.6 mmol, 3.0 eq). The mixture was stirred for 10 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.41 g, 0.35 mmol, 0.1 eq) was added and the reaction was stirred at rt for 75 minutes. The reaction mixture was concentrated under high vacuum, co-evaporating with toluene to give a sticky dark brown gum. The residue was soaked in diethyl ether (2 × 50 mL), the ether was decanted, and residue was co-evaporated with toluene to yield a dark brown foamy resin. The crude material was twice-purified by column chromatography (0–15% methanol/CH_2Cl_2 + 1-5% Et_3N) to afford, and drying under high vacuum, the desired product **153** as an orange/brown glassy solid (2.09 g), containing 57 mol % Et_3N (recalculated yield - 2.67 mmol, 76%).

R_f 0.35 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'); 0.35 (4:1, ethyl acetate/toluene + 0.3% Et_3N, A'); **Mp** 116-120 °C (CH_2Cl_2); **IR** (solid): ν 3342 (w, O–H, N–H), 3232 (w, O–H), 3056, 2973, 2931, 2834 (w, C–H), 2171 (w, C≡C), 1689 (br m, C=O), 1632 (s, N–H), 1596 (s, C=C), 1507 (s, Aryl C=C), 1442 (C–H), 1385 (m, O–H), 1276 (s), 1246 (vs), 1174, 1087, 1030 (s, C–O), 943, 866 (m), 826 (s, C–H), 788 (s), 754, 688 (s, Aryl C–H), 583, 515 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.3 (1H, br s, N**H**^{3}), 7.97 (1H, s, **H**^{6}), 7.41 (2H, d, J = 7.5 Hz, **H**^{15}), 7.32–7.26 (2H, m, **H**^{16}), 7.30 (2H, d, J = 9.0 Hz, **H**^{11}), 7.29 (2H, d, J = 9.0 Hz, **H**^{11}), 7.18 (1H, br t, J = 7.3 Hz, **H**^{17}), 6.95 (1H, t, J = 7.7 Hz, **H**$^{5"}$), 6.86 (2H, d, J = 9.0 Hz, **H**^{12}), 6.84 (2H, d, J = 9.0 Hz, **H**^{12}), 6.55 (1H, ddd, J = 0.6, 2.0, 8.0 Hz, **H**$^{4"}$), 6.52 (1H, app. t, J = 1.6 Hz, **H**$^{2"}$), 6.33 (1H, br d, J = 7.8 Hz, **H**$^{6"}$), 6.14 (1H, t, J = 6.8 Hz, **H**$^{1'}$), 5.31 (1H, br s, 3'-O**H**), 5.13 (2H, s, N**H$_2$**), 4.28 (1H, td, J = 3.3, 6.0 Hz, **H**$^{3'}$), 3.95 (1H, td, J = 3.3, 4.8 Hz, **H**$^{4'}$), 3.68 (3H, s, OC**H$_3$**), 3.67 (3H, s, OC**H$_3$**), 3.24 (1H, dd, J = 5.3, 10.5 Hz, **H**$^{5'}$), 3.15 (1H, dd, J = 2.8, 10.5 Hz, **H**$^{5'}$), 2.30 (1H, td, J = 6.7, 13.6 Hz, **H**$^{2'}$), 2.22 (1H, ddd, J = 3.3, 6.3, 13.6 Hz, **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 161.4 (**C**^{4}), 158.1, 158.0 (**C**^{13}), 149.3 (**C**^{2}), 148.5 (**C**$^{3"}$), 144.6 (**C**^{14}), 142.4 (**C**^{6}), 135.6, 135.3 (**C**^{10}), 129.7, 129.6 (**C**^{11}), 128.8 (**C**$^{5"}$), 127.8 (**C**^{16}), 127.6 (**C**^{15}), 126.6 (**C**^{17}), 122.4 (**C**$^{1"}$), 118.7 (**C**$^{6"}$), 116.0 (**C**$^{2"}$), 114.4 (**C**$^{4"}$), 113.2 (**C**^{12}-OMe), 98.8 (**C**^{5}), 93.0 (**C**^{8}), 85.9 (**C**$^{4'}$), 85.8 (**C**^{9}Ar$_3$), 85.1 (**C**$^{1'}$), 80.3 (**C**^{7}), 70.5 (**C**$^{3'}$), 63.7 (**C**$^{5'}$), 54.9 (O**C**H$_3$), 39.9 (**C**$^{2'}$); **LRMS** (ES$^+$): *m*/*z*: 1314 ([2M + Na]$^+$, 22), 747 ([M + Et_3N + H]$^+$, 100), 719 (41), 668 ([M + Na]$^+$, 44), 646 ([M + H]$^+$, 41); **HRMS** (ES$^+$): calcd for $C_{38}H_{35}N_3O_7$ (M), [M + Na]$^+$ = 668.2367, found 668.2375. Found also DMT$^+$: calcd 303.1380, found 303.1387.

Analytical results consistent with reported data.[206]

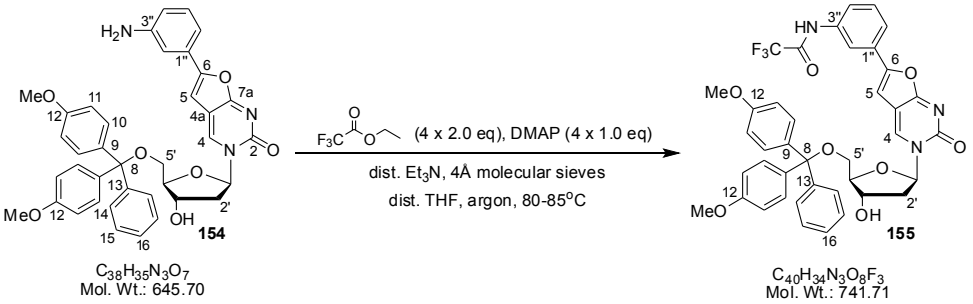
3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-aminophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one, 154[206]



To a solution of nucleoside **153** (1.94 g, 2.48 mmol) in distilled methanol/distilled Et_3N (7:3 v/v, 15.0 mL), under an argon atmosphere, in absence of light, was added CuI (0.57 g, 2.99 mmol, 1.0 eq). The reaction mixture was stirred at 50 °C for 17 hours then cooled to rt and concentrated *in vacuo*, co-evaporating with CH_2Cl_2 (2 × 50 mL). The residue was dissolved/suspended in CH_2Cl_2 (150 mL) and filtered. The filtrate was washed with 5% w/v aq Na_2EDTA (pH 9, 4 × 50 mL), sat aq KCl (2 × 50 mL), dried (Na_2SO_4), concentrated *in vacuo* and dried under high vacuum to afford the product **154** as an orange/brown glassy foam (1.53 g, 2.37 mmol, 96%) containing 3.5 mol % uncyclised material **153**.

$\mathbf{R}_f$ 0.35 (**153**), 0.34 (**154**) (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'); **^{1}H NMR** (300 MHz, d_6-DMSO) δ 8.70 (1H, s, **H^{4}**), 7.40 (2H, d, *J* = 7.5 Hz, **H^{14}**), 7.34 (2H, t, *J* = 7.5 Hz, **H^{15}**), 7.29 (4H, d, *J* = 9.0 Hz, **H^{10}**), 7.31-7.24 (1H, m, **H^{16}**), 7.13 (1H, t, *J* = 7.8 Hz, **H$^{5"}$**), 6.94-6.88 (1H, m, **H$^{2"}$**), 6.92 (2H, d, *J* = 8.8 Hz, **H^{11}**), 6.91 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.84 (1H, br d, *J* = 7.5 Hz, **H$^{4"}$**), 6.62 (1H, br dd, *J* = 1.5, 8.0 Hz, **H$^{6"}$**), 6.13 (1H, dd, *J* = 4.6, 6.4 Hz, **H$^{1'}$**), 5.42 (1H, d, *J* = 4.8 Hz, 3'-O**H**), 5.32 (2H, br s, N**H$_2$**), 4.43 (1H, td, *J* = 5.9, 11.1 Hz, **H$^{3'}$**), 4.04-3.99 (1H, m, **H$^{4'}$**), 3.72 (3H, s, OC**H$_3$**), 3.71 (3H, s, OC**H$_3$**), 3.41 (1H, dd, *J* = 4.0, 10.8 Hz, **H$^{5'}$**), 3.32 (1H, dd, *J* = 2.8, 10.8 Hz, **H$^{5'}$**), 2.46 (1H, td, *J* = 6.6, 13.1 Hz, **H$^{2'}$**), 2.25 (1H, ddd, *J* = 4.8, 6.5, 13.4 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.9 (**C^{7a}**), 158.2 (**C^{12}**–OCH_3), 154.5 (**C^{2}**), 149.2 (**C$^{3"}$**–NH_2), 144.4 (**C^{13}**), 137.3 (**C^{4}**), 135.4 (**C^{9}**), 135.3 (**C^{6}**), 135.1 (**C^{9}**), 129.8, 129.7 (**C^{10}**), 129.5 (**C$^{5"}$**), 128.6 (**C$^{1"}$**), 128.0 (**C^{15}**), 127.8 (**C^{14}**), 126.8 (**C^{16}**), 115.1 (**C$^{6"}$**), 113.3 (**C^{11}**), 112.0 (**C$^{4"}$**), 109.2 (**C$^{2"}$**), 106.7 (**C^{4a}**), 97.8 (**C^{5}**), 87.3 (**C$^{1'}$**), 86.1 (**C^{8}**Ar_3), 85.9 (**C$^{4'}$**), 68.8 (**C$^{3'}$**), 62.4 (**C$^{5'}$**), 55.0 (O**C**H_3), 41.1 (**C$^{2'}$**); **LRMS** (ES^+): *m*/*z*: 917 (20), 663 ($[M + NH_4]^+$, 100), 646 ($[M + H]^+$, 13); **HRMS** (ES^+): calcd for $C_{38}H_{35}N_3O_7$ (M), $[M + Na]^+$ = 668.2367, found 668.2359. Found also DMT^+: calcd 303.1380, found 303.1382.
Analytical results consistent with reported data.[206]

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-[3-(trifluoroacetamido)phenyl]-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one, 155[206]



$C_{38}H_{35}N_3O_7$
Mol. Wt.: 645.70

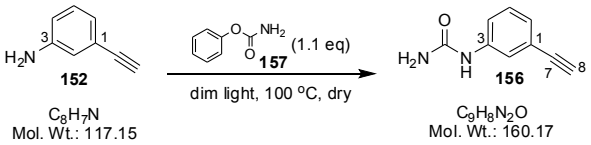
$C_{40}H_{34}N_3O_8F_3$
Mol. Wt.: 741.71

To a stirred solution of nucleoside **154** (1.37 g, 2.12 mmol) and *N*,*N*-dimethylamino pyridine (DMAP) (0.26 g, 2.13 mmol, 1.0 eq) in distilled THF (10.0 mL), under an argon atmosphere, with 4Å molecular sieves, was added distilled Et_3N (0.15 mL) then ethyl trifluoroacetate (0.50 mL, 4.24 mmol, 2.0 eq). The solution was stirred at rt for 2 minutes, then heated to reflux under argon pressure (80 °C). Further DMAP (1.0 eq) and ethyl trifluoroacetate (2.0 eq) were added after 6 hours and 21½ hours. The temperature was then raised to 85 °C. Further DMAP (0.5 eq) and ethyl trifluoroacetate (1.0 eq) were added at 30 hours. At 53 hours, the reaction mixture was cooled to rt, dissolved in CH_2Cl_2/methanol (1:1 v/v, 60 mL) and concentrated *in vacuo*, co-evaporating with CH_2Cl_2 (4 × 40 mL) to give an orange/brown solid. Following purification by column chromatography (70-90% ethyl acetate/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **155** was afforded as an off-white foam (0.55 g, 0.74 mmol, 35%).

R_f 0.23 (9:1, ethyl acetate/CH_2Cl_2 + 0.3% Et_3N, A'); 0.42 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.40 (1H, br s, N**H**COCF$_3$), 8.75 (1H, s, **H^4**), 8.06 (1H, t, *J* = 1.8 Hz, **H$^{2"}$**), 7.72 (1H, td, *J* = 2.0, 7.3 Hz, **H$^{4"}$**), 7.57 (1H, dd, *J* = 1.7, 7.7 Hz, **H$^{6"}$**), 7.53 (1H, t, *J* = 7.7 Hz, **H$^{5"}$**), 7.41 (2H, dd, *J* = 1.5, 8.1 Hz, **H^{14}**), 7.34 (2H, t, *J* = 8.1 Hz, **H^{15}**), 7.31-7.25 (1H, m, **H^{16}**), 7.29 (2H, d, *J* = 8.8 Hz, **H^{10}**), 7.29 (2H, d, *J* = 9.2 Hz, **H^{10}**), 6.92 (2H, d, *J* = 9.1 Hz, **H^{11}**), 6.91 (2H, d, *J* = 9.2 Hz, **H^{11}**), 6.33 (1H, s, **H^5**), 6.16 (1H, dd, *J* = 4.4, 6.6 Hz, **H$^{1'}$**), 5.43 (1H, d, *J* = 4.8 Hz, 3'-O**H**), 4.43 (1H, td, *J* = 5.2, 10.5 Hz, **H$^{3'}$**), 4.04 (1H, dd, *J* = 4.5, 8.1, **H$^{4'}$**), 3.71 (3H, s, OC**H$_3$**), 3.70 (3H, s, OC**H$_3$**), 3.41 (1H, dd, *J* = 4.0, 10.8 Hz, **H$^{5'}$**), 3.32 (1H, dd, *J* = 3.3, 11.0 Hz, **H$^{5'}$**), 2.54-2.44 (1H, m, **H$^{2'}$**), 2.27 (1H, ddd, *J* = 4.6, 6.4, 13.5 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.9 (**C^{7a}**), 158.2 (**C^{12}**–OCH$_3$), 154.7 (q, $^2J_{CF}$ = 36.9 Hz, **C**OCF$_3$), 153.6 (**C^2**), 144.4 (**C^{13}**), 138.2 (**C^4**), 137.4 (**C^6**), 137.3 (**C$^{3"}$**–NHR), 135.4, 135.1 (**C^9**), 129.9

(**C**$^{5''}$), 129.8, 129.7 (**C**^{10}), 128.8 (**C**$^{1''}$), 128.0 (**C**^{15}), 127.8 (**C**^{14}), 126.8 (**C**^{16}), 121.9 (**C**$^{4''}$), 121.5 (**C**$^{6''}$), 116.7 (**C**$^{2''}$), 113.3 (**C**^{11}), 113.2 (q, $^{1}J_{CF}$ = 227.4 Hz, CO**C**F_3), 106.4 (**C**^{4a}), 99.6 (**C**^{5}), 87.6 (**C**$^{1'}$), 86.1 (**C**^{8}Ar$_3$), 86.0 (**C**$^{4'}$), 68.8 (**C**$^{3'}$), 62.4 (**C**$^{5'}$), 54.9 (O**C**H_3), 41.1 (**C**$^{2'}$); **LRMS** (ES$^+$) *m*/*z*: 764 ([M + Na]$^+$, 100); (ES$^-$) *m*/*z*: 740 ([M – H]$^-$, 100).
Analytical results consistent with reported data.[206]

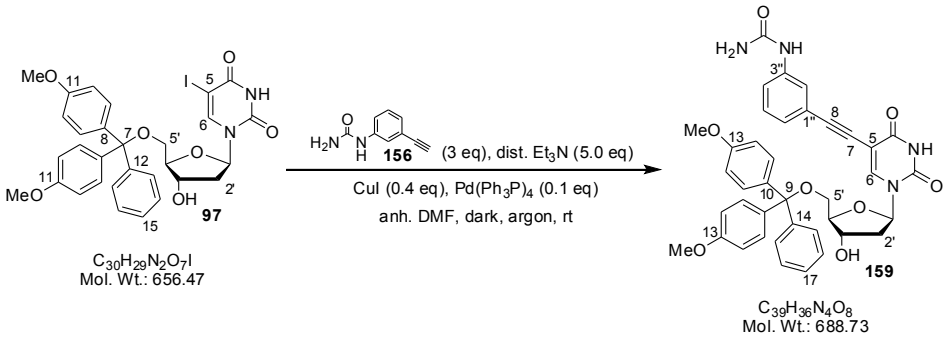
3-Ureidophenyl acetylene, 156[206]



3-Ethynylaniline **152** (3.00 mL, 26.6 mmol) and dry phenyl carbamate **157** (7.31 g, 53.3 mmol, 2.0 eq) were heated in a sealed tube, under an argon atmosphere, excluding moisture, in dim light at 90 °C for 12 hours. The mixture was cooled to rt, transferred using warm methanol/acetone, and dried under high vacuum. Following purification by column chromatography (5-50% acetone/CH_2Cl_2), and drying under high vacuum, the desired urea **156** was afforded, as an off-white, papery solid (2.93 g, 18.3 mmol, 69%).

R$_f$ 0.31 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A', F'), 0.04 (1:9, acetone/CH_2Cl_2 + 0.3% Et_3N, A', F'); **IR** (solid): ν 3467, 3329 (m, N–H), 3302 (w N–H), 3238(m, N–H), 2100 (w, C≡C), 1649 (vs, C=O), 1586, 1575 (s, Aryl C–H), 1540 (br vs, Aryl C–H), 1473 (s), 1413 (m), 1325 (s), 1284, 1256, 1162 (m), 1120 (br w), 1029 (w), 890 (m, Aryl C–H), 801 (m), 773 (m, Aryl C–H), 683 (s), 643 (br s), 587 (s) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO) δ 8.61 (1H, s, N**H**), 7.64 (1H, t, *J* = 1.7 Hz, **H**^{2}), 7.32 (1H, ddd, *J* = 1.1, 2.0, 8.2 Hz, **H**^{4}), 7.22 (1H, t, *J* = 7.9 Hz, **H**^{5}), 7.00 (1H, td *J* = 1.3, 7.3 Hz, **H**^{6}), 5.89 (2H, s, N**H**$_2$), 4.08 (1H, s, C≡C**H**^{8}); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 155.9 (**C**ONH_2), 140.8 (**C**^{3}), 129.0 (**C**^{5}), 124.3 (**C**^{6}), 121.8 (**C**^{1}), 120.5 (**C**^{2}), 118.4 (**C**^{4}), 83.8 (**C**^{7}≡CH), 80.0 (C≡**C**^{8}H); **LRMS** (ES$^+$) *m*/*z*: 343 ([2M + Na]$^+$, 21), 183 ([M + Na]$^+$, 70), 102 ([Et_3N + H]$^+$, 100); (ES$^-$) *m*/*z*: 159 ([M – H]$^-$, 100); **HRMS** (ES$^+$): calcd for $C_9H_8N_2O$ (M), [M + Na]$^+$ = 183.0529, found 183.0530, [2M + Na]$^+$ = 343.1171, found 343.1170.
Analytical results consistent with reported data.[206]

5-[(3-Ureidophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine, 159

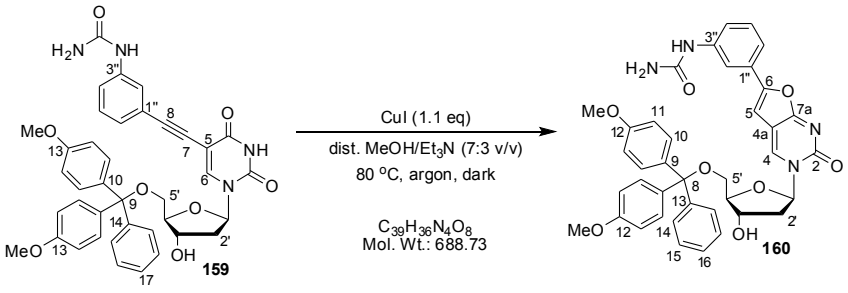


To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **97** (3.01 g, 4.58 mmol) in anhydrous DMF (15.0 mL), under an argon atmosphere, in absence of light, were added distilled Et_3N (3.30 mL, 23.7 mmol, 5.0 eq), CuI (0.35 g, 1.83 mmol, 0.4 eq), then alkyne **156** (2.20 g, 13.7 mmol, 3.0 eq). The mixture was stirred for 15 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.53 g, 0.46 mmol, 0.1 eq) was added and the reaction was stirred at rt for 3 hours. The reaction mixture was concentrated under high vacuum, co-evaporating with CH_3CN to give a dark brown foam/gum. Following purification by column chromatography (5-8% methanol/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **159** was afforded as a pale orange foam (2.71 g) containing 34 mol % Et_3N (recalculated yield - 3.66 mmol, 80%).

R_f 0.23 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.27 (7:3 acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **IR** (solid): ν 3348 (w, O–H), 3208, 3067, 2931 (w, N–H), 2832 (w, C–H), 1669 (br vs, C=O, C=C), 1605, 1580 (s, N–H), 1546, 1506, 1456, 1435 (s), 1342, 1302 (m), 1274 (s), 1247 (vs, C–O), 1175 (s), 1088, 1030 (br s, C–O), 826 (s, Aryl C–H), 789, 753, 698, 585 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.71 (1H, br s, N**H**^{3}), 8.57 (1H, s, N**H**$CONH_2$), 8.02 (1H, s, **H**^{6}), 7.59 (1H, t, *J* = 1.8 Hz, **H**$^{2"}$), 7.41 (2H, d, *J* = 7.5 Hz, **H**^{15}), 7.31-7.26 (1H, m, **H**^{16}), 7.30 (2H, d, *J* = 9.0 Hz, **H**^{11}), 7.29 (2H, d, *J* = 8.8 Hz, **H**^{11}), 7.24 (1H, ddd, *J* = 1.0, 2.1, 8.2 Hz, **H**$^{4"}$), 7.17 (1H, br t, *J* = 7.4 Hz, **H**^{17}), 7.14 (1H, t, *J* = 7.9 Hz, **H**$^{5"}$), 6.85 (2H, d, *J* = 8.8 Hz, **H**^{12}), 6.84 (2H, d, *J* = 9.0 Hz, **H**^{12}), 6.65 (1H, td, *J* = 1.3, 6.5 Hz, **H**$^{6"}$), 6.14 (1H, t, *J* = 6.7 Hz, **H**$^{1'}$), 5.90 (2H, s, CON**H**$_2$), 5.32 (1H, d, *J* = 4.3 Hz, 3'-O**H**), 4.29 (1H, td, *J* = 2.5, 6.5 Hz, **H**$^{3'}$), 3.95 (1H, td, *J* = 3.3, 5.0 Hz, **H**$^{4'}$), 3.67 (6H, s, OC**H**$_3$), 3.25 (1H, dd, *J* = 5.3, 10.5 Hz, **H**$^{5'}$), 3.14 (1H, dd, *J* = 2.8, 10.5 Hz, **H**$^{5'}$), 2.31 (1H, td, *J* = 6.8, 13.6 Hz, **H**$^{2'}$), 2.22 (1H, ddd, *J* = 3.5, 6.3, 13.6 Hz,

H$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 161.4 (**C^{4}**), 158.1, 158.0 (**C^{13}**–OCH_3), 155.8 (**C**ONH_2), 149.3 (**C^{2}**), 144.7 (**C^{14}**), 142.8 (**C^{6}**), 140.6 (**C$^{3"}$**–NHR), 135.6, 135.3 (**C^{10}**), 129.7, 129.6 (**C^{11}**), 128.7 (**C$^{5"}$**), 127.8 (**C^{16}**), 127.6 (**C^{15}**), 126.7 (**C^{17}**), 124.0 (**C$^{6"}$**), 122.4 (**C$^{1"}$**), 120.0 (**C$^{2"}$**), 118.0 (**C$^{4"}$**), 113.2 (**C^{12}**), 98.6 (**C^{5}**), 92.4 (**C^{8}**), 86.0 (**C$^{4'}$**), 85.8 (**C^{9}**Ar_3), 85.1 (**C$^{1'}$**), 81.2 (**C^{7}**), 70.5 (**C$^{3'}$**), 63.7 (**C$^{5'}$**), 54.9 (OCH_3), 39.9 (**C$^{2'}$**); **LRMS** (ES^+) *m*/*z*: 711 ([M + Na]$^+$, 100), 303 ([DMT$^+$], 9); **HRMS** (ES^+): calcd for $C_{39}H_{36}N_4O_8$ (M), [M + Na]$^+$ = 711.2425, found 711.2409, [M + NH_4]$^+$ = 706.2877, found 706.2856. Found also 301.1413.

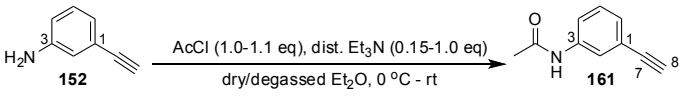
3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one, 160



To a solution of nucleoside **159** (2.13 g, 3.09 mmol) in distilled/argon-degassed methanol/Et_3N (7:3 v/v, 15.0 mL), with 4Å molecular sieves, in absence of light and under an argon atmosphere, was added CuI (0.65 g, 3.40 mmol, 1.1 eq), and the mixture was stirred at reflux under argon pressure (80 °C) for 3.5 hours. After cooling to rt, the reaction cake was dissolved/suspended in CH_2Cl_2 (150 mL), filtered, and the solid was washed with CH_2Cl_2 (2 × 50 mL). The combined filtrate was washed with 5% w/v aq Na_2EDTA (pH 9, 4 × 100 mL) and sat aq KCl (100 mL), dried (Na_2SO_4), and concentrated *in vacuo* to give a yellow solid. The product was crystallised by addition of methanol (10 mL) and the mixture was allowed to stand in the fridge for 2 days. The solid was filtered, washed with the filtrate (×2) then cold diethyl ether (-8 °C, 3 × 10 mL), and dried under high vacuum over KOH, to afford the desired product **160** as a pale cream-coloured fine crystalline solid (1.34 g). The filtrate was concentrated *in vacuo* from which two further crops were obtained (0.11 g, 0.03 g) (combined yield 1.47 g, 2.14 mmol, 69%).

R_f 0.18 (9:1, acetone/CH_2Cl_2 + 0.3% Et_3N), 0.23 (acetone + 0.3% Et_3N); 0.33 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N); **Mp** 160-169 °C (dec.); **IR** (solid): ν 3472, 3305 (m, N–H, O–H), 2932, 2836 (w, C–H), 1663 (s, C=O), 1606 (s, N–H), 1584 (s, C=N), 1547 (m), 1508, 1492 (s, Aryl C=C), 1445, 1408 (m), 1382 (m, C–H), 1342 (m), 1301, 1247 (s), 1174 (vs), 1072 (br s, C–O), 1031 (vs), 972, 912, 878 (m), 828 (s, C–H), 773 (s), 701 (s, Aryl C–H), 583, 532 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 7.73 (1H, br s, N**H**$CONH_2$), 8.72 (1H, s, **H^4**), 7.90 (1H, s, **H$^{2"}$**), 7.40 (2H, d, *J* = 7.5 Hz, **H^{14}**), 7.37-7.31 (4H, m, **H^{15}**, **H$^{4"}$**, **H$^{5"}$**), 7.31-7.24 (1H, m, **H^{16}**), 7.28 (4H, d, *J* = 8.3 Hz, **H^{10}**), 7.28-7.22 (1H, m, **H$^{6"}$**), 6.92 (2H, d, *J* = 8.8 Hz, **H^{11}**), 6.91 (2H, d, *J* = 8.8 Hz, **H^{11}**), 6.19 (1H, s, **H^5**), 6.16 (1H, dd, *J* = 4.5, 6.3 Hz, **H$^{1'}$**), 5.92 (2H, s, CON**H$_2$**), 5.42 (1H, d, *J* = 4.9 Hz, 3'-O**H**), 4.44 (1H, td, *J* = 5.6, 10.8 Hz, **H$^{3'}$**), 4.02 (1H, dd, *J* = 3.5, 7.5 Hz, **H$^{4'}$**), 3.72 (3H, s, OC**H$_3$**), 3.71 (3H, s, OC**H$_3$**), 3.41 (1H, dd, *J* = 3.7, 10.8 Hz, **H$^{5'}$**), 3.34-3.27 (1H, m, **H$^{5'}$**), 2.47 (1H, td, *J* = 6.5, 13.1 Hz, **H$^{2'}$**), 2.26 (1H, ddd, *J* = 4.6, 6.5, 13.6 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.9 (**C^{7a}**), 158.2 (**C^{12}**–OCH_3), 155.9 (**C**ONH_2), 153.7 (**C^2**), 144.5 (**C^6**), 141.3 (**C^{13}**), 137.8 (**C^4**), 137.7 (**C$^{3"}$**–NHR), 135.3, 135.0 (**C^9**), 129.8, 129.8 (**C^{10}**), 129.4 (**C$^{5"}$**), 128.4 (**C$^{1"}$**), 128.0 (**C^{15}**), 127.7 (**C^{14}**), 126.9 (**C^{16}**), 118.7 (**C$^{6"}$**), 117.2 (**C$^{4"}$**), 113.3 (**C^{11}**), 113.2 (**C$^{2"}$**), 106.6 (**C^{4a}**), 98.6 (**C^5**), 87.4 (**C$^{1'}$**), 86.1 (**C^8**Ar_3), 85.9 (**C$^{4'}$**), 68.8 (**C$^{3'}$**), 62.4 (**C$^{5'}$**), 55.0 (O**C**H_3), 41.1 (**C$^{2'}$**); **LRMS** (ES^+) *m*/*z*: 743 ([M + CH_3OH + Na]$^+$, 100), 711 ([M + Na]$^+$, 74); (ES^-) *m*/*z*: 687 ([M – H]$^-$, 100); **HRMS** (ES^+): calcd for $C_{39}H_{36}N_4O_6$ (M), [M + Na]$^+$ = 711.2425, found 711.2389, [M + H]$^+$ = 689.2606, found 689.2572. Found also DMT$^+$ (calcd 303.1380, found 303.1385).

3-Acetamidophenyl acetylene, 161[211]



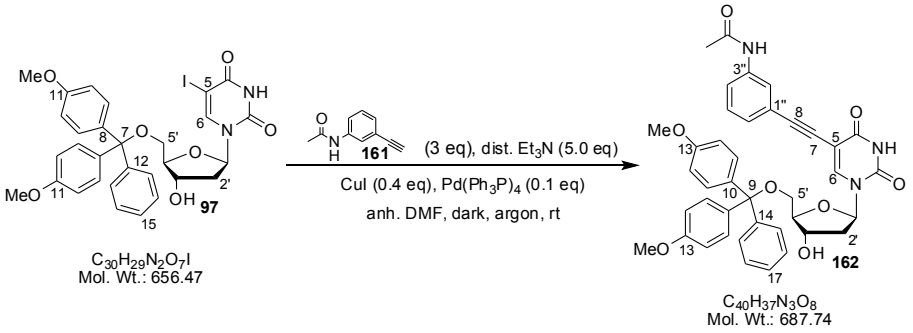
Method 1: To a stirred solution of 3-ethynylaniline **152** (0.20 mL, 1.78 mmol) in dry, degassed diethyl ether (5.0 mL) at 0 °C, under an argon atmosphere, was added distilled Et_3N (0.04 mL, 0.29 mmol, 0.16 eq), followed by acetyl chloride (0.14 mL, 1.97 mmol, 1.1 eq) in dry, degassed diethyl ether (4.0 mL) dropwise. The reaction mixture was stirred for 2 hours, allowed to warm to rt, then stirred for a further 4.5 hours. The reaction mixture was filtered, the solid was washed with diethyl ether (3 ×

10 mL), and combined ether fractions were concentrated *in vacuo* and dried under high vacuum to give a pale yellow oil. Following trituration with water (3.0 mL), the product **161** was afforded, following drying under high vacuum over KOH, as a pale cream-coloured waxy solid (228 mg, 1.43 mmol, 81%).

Method 2: To a stirred solution of 3-ethynylaniline **152** (3.50 mL, 31.1 mmol) and distilled Et_3N (4.37 mL, 31.4 mmol, 1.0 eq) in dry, degassed diethyl ether (15.0 mL), at 0 °C, under an argon atmosphere, was added acetyl chloride (2.23 mL, 31.4 mmol, 1.0 eq) in dry, degassed diethyl ether (25.0 mL), dropwise over 45 minutes. The reaction was stirred for a further 15 minutes then allowed to warm to rt over 30 minutes. The mixture was filtered, the white precipitate was washed with diethyl ether (3 × 35 mL) and combined ether fractions were concentrated *in vacuo* to give a cream-coloured solid. The solid was triturated with water (60 mL) to which was added diethyl ether (100 mL) and the layers were separated. The ether layer was washed with water (50 mL), combined aqueous layers were re-extracted (diethyl ether, 20 mL), and combined ether fractions were concentrated *in vacuo* and under high vacuum to afford the desired product **161** a pale brown oil, which crystallised on standing to an off-white chalky/waxy solid (4.79 g, 30.1 mmol, 97%).

R_f 0.40 (5:95, methanol/CH_2Cl_2 + 0.3% Et_3N, A'); **Mp** 92-94 °C (chloroform/hexane), *lit.*[211] 94-96 °C (CCl_4); **IR** (solid): ν 3300, 3282 (m, CC–H), 1666 (m, C=O), 1605, 1583 (m, Aryl C–H), 1556 (m, C=O), 1481 (m, Aryl C–H), 1425, 1401, 1370, 1308, 1289, 1255 (m), 1015 (w), 880, 786 (m, Aryl C–H), 721, 678, 648 (m), 601 (s), 541 (m) cm^{-1}; **^{1}H NMR** (300 MHz, d_6-DMSO) δ 10.00 (1H, s, N**H**$COCH_3$), 7.77 (1H, app. t, J = 1.7 Hz, **H^2**), 7.54 (1H, ddd, J = 1.1, 2.0, 8.1 Hz, **H^4**), 7.30 (1H, t, J = 7.9 Hz, **H^5**), 7.13 (1H, td, J = 1.3, 7.7 Hz, **H^6**), 4.13 (1H, s, C≡C**H^8**), 2.05 (3H, s, COC**H$_3$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 169.1 (**C**OCH_3), 140.1 (**C^3**), 129.7 (**C^5**), 126.8 (**C^6**), 122.5 (**C^1**), 122.4 (**C^2**), 120.1 (**C^4**), 84.0 (**C^7**≡CH), 81.0 (C≡**C^8**H), 24.6 (CO**C**H_3); **LRMS** (ES^+) *m*/*z*: 160 ($[M + H]^+$, 100), 182 ($[M + Na]^+$, 11), 201 ($[M + CH_3CN + H]^+$, 7), 223 ($[M + CH_3CN + Na]^+$, 17), 242 ($[M + AcOH + Na]^+$, 12); (ES^-) *m*/*z*: 158 ($[M – H]^-$, 100); **HRMS** (ES^+): calcd for $C_{10}H_9NO$ (M), $[2M + Na]^+$ = 341.1260, found 341.1239, $[M + Na]^+$ = 182.0576, found 182.0565, $[M + H]^+$ = 160.0757, found 160.0755.

5-[(3-Acetamidophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine, 162

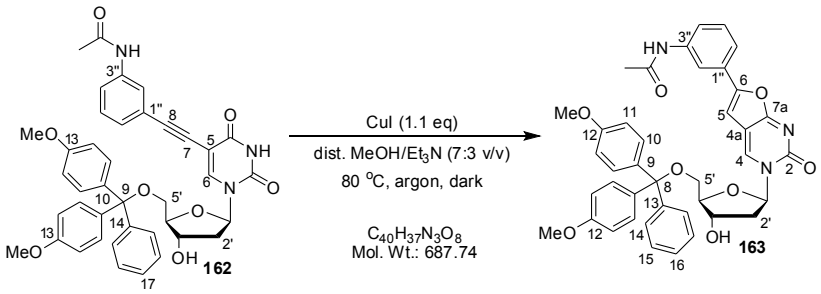


To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **97** (3.02 g, 4.61 mmol) in anhydrous DMF (15.0 mL), under an argon atmosphere, in absence of light, were added distilled Et_3N (3.20 mL, 23.0 mmol, 5.0 eq), CuI (0.35 g, 1.84 mmol, 0.4 eq), then alkyne **156** (2.20 g, 13.8 mmol, 3.0 eq). The mixture was stirred for 15 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.53 g, 0.46 mmol, 0.1 eq) was added and the reaction was stirred at rt for 2 hours. After concentration under high vacuum and co-evaporation with methanol/toluene, the dark brown foam was purified four times by column chromatography (3-10% methanol/CH_2Cl_2 + 0.5% Et_3N; 20-50% acetone/CH_2Cl_2 + 0.5% Et_3N; 40-100% acetone/CH_2Cl_2 + 0.5% Et_3N; 92:6:2-82:16:2, ethyl acetate/methanol/conc. aq NH_3). The desired product **162** was afforded, after drying under high vacuum, as a pale orange foam (2.37 g, 3.44 mmol, 75%).

R_f 0.39 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.33 (82:16:2, ethyl acetate/methanol/conc. aq NH_3, A'), 0.64 (1:4, methanol/ethyl acetate + 0.3% Et_3N, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.63 (1H, extremely br s, N**H**^{3}), 9.98 (1H, s, N**H**COCH$_3$), 8.04 (1H, s, **H**^{6}), 7.74 (1H, t, *J* = 1.5 Hz, **H**$^{2"}$), 7.45 (1H, br dd, *J* = 1.2, 8.0 Hz, **H**$^{4"}$), 7.41 (2H, d, *J* = 7.5 Hz, **H**^{15}), 7.30 (2H, d, *J* = 9.0 Hz, **H**^{11}), 7.30-7.25 (1H, m, **H**$^{5"}$), 7.29 (2H, d, *J* = 9.0 Hz, **H**^{11}), 7.22 (1H, t, *J* = 8.0 Hz, **H**^{16}), 7.16 (1H, t, *J* = 7.3 Hz, **H**^{17}), 6.85 (2H, d, *J* = 8.5 Hz, **H**^{12}), 6.84 (2H, d, *J* = 9.0 Hz, **H**^{12}), 6.74 (1H, br d, *J* = 7.5 Hz, **H**$^{6"}$), 6.14 (1H, t, *J* = 6.8 Hz, **H**$^{1'}$), 5.32 (1H, br s, 3'-O**H**), 4.29 (1H, br s, **H**$^{3'}$), 3.95 (1H, td, *J* = 3.2, 5.0 Hz, **H**$^{4'}$), 3.66 (6H, s, OC**H**$_3$), 3.25 (1H, dd, *J* = 5.3, 10.8 Hz, **H**$^{5'}$), 3.15 (1H, dd, *J* = 2.5, 10.5 Hz, **H**$^{5'}$), 2.32 (1H, td, *J* = 6.8, 13.6 Hz, **H**$^{2'}$), 2.22 (1H, ddd, *J* = 3.5, 6.5, 13.6 Hz, **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 168.4 (**C**OCH$_3$), 161.4 (**C**^{4}), 158.0, 158.0 (**C**^{13}–OCH$_3$), 149.3 (**C**^{2}), 144.7 (**C**^{14}), 142.9 (**C**^{6}), 139.3 (**C**$^{3"}$–NHR), 135.6, 135.3 (**C**^{10}), 129.7, 129.6 (**C**^{11}), 128.8 (**C**$^{5"}$), 127.8 (**C**^{16}), 127.6 (**C**^{15}), 126.6 (**C**^{17}), 125.8 (**C**$^{6"}$), 122.5 (**C**$^{2"}$), 121.3 (**C**$^{4"}$), 119.1 (**C**$^{1"}$), 113.2 (**C**^{12}), 98.4 (**C**^{8}), 92.0 (**C**^{7}), 86.0 (**C**$^{4'}$), 85.8

($\mathbf{C}^{9}Ar_3$), 85.2 ($\mathbf{C}^{1'}$), 81.6 ($\mathbf{C}^{5}$), 70.4 ($\mathbf{C}^{3'}$), 63.7 ($\mathbf{C}^{5'}$), 54.9 ($O\mathbf{C}H_3$), 40.0 ($\mathbf{C}^{2'}$), 24.0 ($CO\mathbf{C}H_3$); **LRMS** (ES^+) *m*/*z*: 710 ($[M + Na]^+$, 100), 705 ($[M + NH_4]^+$, 8); (ES^-) *m*/*z*: 784 ($[M + H_3PO_4 - H]^-$, 100), 732 ($[M + HCO_2H - H]^-$, 45), 686 ($[M - H]^-$, 12); **HRMS** (ES^+): calcd for $C_{40}H_{37}N_3O_8$ (M), $[M + Na]^+$ = 710.2473, found 710.2480.

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-acetamido-phenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one, 163

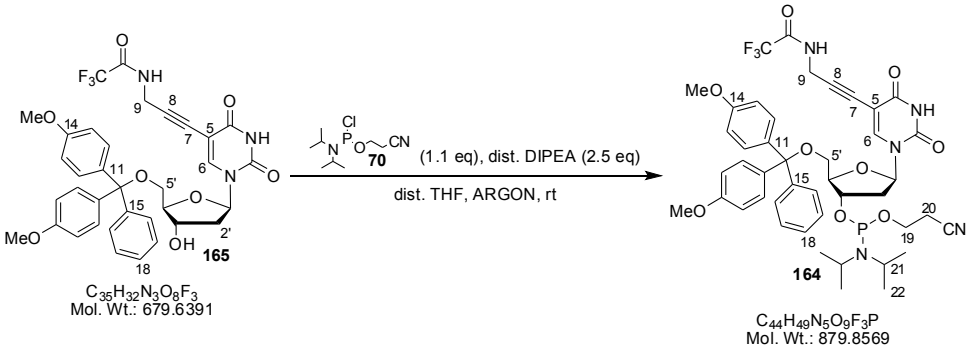


To a solution of nucleoside **159** (1.94 g, 2.82 mmol) in distilled/argon-degassed methanol/Et_3N (7:3 v/v, 20.0 mL), with 4Å molecular sieves, in absence of light and under an argon atmosphere, was added CuI (0.59 g, 3.10 mmol, 1.1 eq), and the mixture was stirred at reflux under argon pressure (80 °C) for 2 hours. The reaction was diluted with methanol (20 mL), filtered, and the filtrate was concentrated *in vacuo*. The green foam was dissolved in CH_2Cl_2 (100 mL) and washed with 5% w/v aq Na_2EDTA (pH 10, 5 × 100 mL). The combined aqueous fraction was re-extracted (CH_2Cl_2, 2 × 70 mL) and the combined CH_2Cl_2 layer was washed with sat aq KCl (2 × 100 mL), dried (Na_2SO_4), and concentrated *in vacuo* to give a green foam. Following purification by column chromatography (20-60% acetone/ethyl acetate + 0.5-1.0% Et_3N), and drying under high vacuum, the desired cyclised product **163** was afforded as a pale yellow foam (1.73 g, 2.52 mmol, 89%).

$\mathbf{R}_f$ 0.31 (1:2, acetone/ethyl acetate + 0.3% Et_3N, A'), 0.62 (1:4, methanol/ethyl acetate + 0.3% Et_3N, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.10 (1H, s, N**H**$COCH_3$), 8.73 (1H, s, $\mathbf{H}^{4}$), 8.03 (1H, t, *J* = 1.5 Hz, $\mathbf{H}^{2''}$), 7.59 (1H, td, *J* = 1.6, 7.8 Hz, $\mathbf{H}^{4''}$), 7.44-7.40 (1H, m, $\mathbf{H}^{5''}$), 7.41 (2H, d, *J* = 7.5 Hz, $\mathbf{H}^{14}$), 7.37 (1H, td, *J* = 1.5, 8.0 Hz, $\mathbf{H}^{6''}$), 7.34 (2H, t, *J* = 8.0 Hz, $\mathbf{H}^{15}$), 7.29 (2H, d, *J* = 9.0 Hz, $\mathbf{H}^{10}$), 7.28 (2H, d, *J* = 9.0 Hz, $\mathbf{H}^{10}$), 7.27 (1H, td, *J* = 1.3, 7.0 Hz, $\mathbf{H}^{16}$), 6.92 (2H, d, *J* = 9.0 Hz, $\mathbf{H}^{11}$), 6.91 (2H, d, *J* = 9.0 Hz, $\mathbf{H}^{11}$), 6.23

(1H, s, **H**^{5}), 6.17 (1H, dd, *J* = 4.5, 6.5 Hz, **H**$^{1'}$), 5.44 (1H, d, *J* = 5.0 Hz, 3'-O**H**), 4.44 (1H, td, *J* = 5.5, 11.0 Hz, **H**$^{3'}$), 4.07 (1H, td, *J* = 3.5, 5.0 Hz, **H**$^{4'}$), 3.72 (3H, s, OC**H**$_3$), 3.70 (3H, s, OC**H**$_3$), 3.41 (1H, dd, *J* = 4.0, 10.5 Hz, **H**$^{5'}$), 3.32 (1H, dd, *J* = 3.0, 10.5 Hz, **H**$^{5'}$), 2.48 (1H, td, *J* = 6.5, 13.3 Hz, **H**$^{2'}$), 2.26 (1H, ddd, *J* = 4.5, 6.5, 13.5 Hz, **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 170.9 (**C**^{7a}), 168.6 (**C**OCH$_3$), 158.2 (**C**^{12}–OCH$_3$), 153.6 (**C**^{2}), 153.4 (**C**^{6}), 144.5 (**C**^{13}), 140.0 (**C**$^{3''}$–NHR), 137.9 (**C**^{4}), 135.4, 135.0 (**C**^{9}), 129.8, 129.7 (**C**^{10}), 129.5 (**C**$^{5''}$), 128.5 (**C**$^{1''}$), 128.0 (**C**^{15}), 127.7 (**C**^{14}), 126.8 (**C**^{16}), 119.8 (**C**$^{4''}$), 119.0 (**C**$^{6''}$), 114.5 (**C**$^{2''}$), 113.3 (**C**^{11}), 106.5 (**C**^{4a}), 98.9 (**C**^{5}), 87.5 (**C**$^{1'}$), 86.1 (**C**^{8}Ar$_3$), 85.9 (**C**$^{4'}$), 68.8 (**C**$^{3'}$), 62.4 (**C**$^{5'}$), 55.0, 55.0 (O**C**H$_3$), 41.1 (**C**$^{2'}$), 24.0 (CO**C**H$_3$); **LRMS** (ES$^+$) *m*/*z*: 710 ([M + Na]$^+$, 100); (ES$^-$) *m*/*z*: 686 ([M – H]$^-$, 100); **HRMS** (ES$^+$): calcd for $C_{40}H_{37}N_3O_8$ (M), [M + Na]$^+$ = 710.2473, found 710.2462.

5'-*O*-(4,4'-Dimethoxytrityl)-5-(3-trifluoroacetamidopropynyl)-2'-deoxyuridine-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite, 164[117,170]

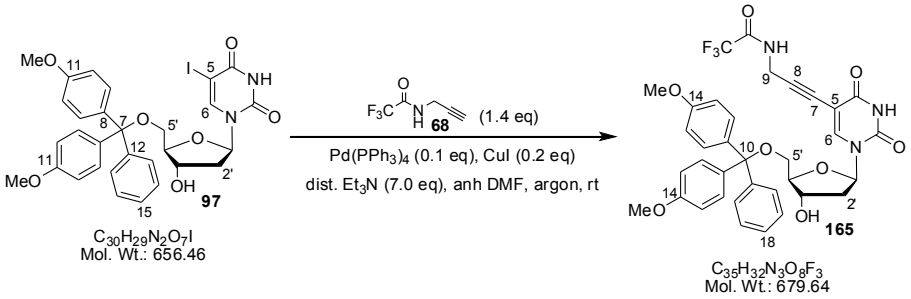


To a stirred solution of nucleoside **165** (5.02 g, 7.39 mmol) in distilled THF (25.0 mL), strictly under an argon atmosphere and excluding moisture, was added distilled DIPEA (3.22 mL, 18.5 mmol, 2.5 eq) followed by chloro-phosphitylating reagent **70** (1.82 mL, 8.14 mmol, 1.1 eq) dropwise, and the reaction was stirred at rt for 1¼ hours. The THF was removed *in vacuo* and the residue was dissolved in argon-degassed ethyl acetate (100 mL). The solution was washed with degassed sat. aq KCl (3 × 50 mL), dried (Na_2SO_4), and concentrated *in vacuo* to give an orange foam. Following purification by column chromatography (10% acetone/CH_2Cl_2 + 1.0% pyridine) on silica gel pre-equilibrated with Et_3N, under argon pressure, the desired phosphoramidite **164** was afforded, after drying under high vacuum, as a diastereomeric mixture (*ca.* 1:2), as an air-sensitive pale peach-coloured foam (5.45 g, 6.19 mmol, 84%).

R_f 0.41 (1:9, acetone/CH_2Cl_2 + 0.5% pyridine, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 148.9, 148.5 (**P^{III}**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.66 (N**H**^{3}), 9.98 (1H, s, CON**H**), 7.97 and 7.96 (1H, s, **H**^{6}), 7.42-7.38 (2H, m, **H**^{16}), 7.33-7.28 (2H, m, **H**^{17}), 7.29 and 7.27 (4H, d, *J* = 7.8 Hz and *J* = 8.0 Hz, **H**^{12}), 7.25-7.19 (1H, m, **H**^{18}), 6.89 and 6.87 (4H, d, *J* = 8.8 Hz and *J* = 9.0 Hz, **H**^{13}), 6.10 and 6.08 (1H, t, *J* = 6.7 Hz and *J* = 6.8 Hz, **H**$^{1'}$), 4.50 and 4.47 (1H, ddd, *J* = 3.6, 6.7, 14.2 Hz and *J* = 3.9, 6.6, 14.2 Hz, **H**$^{3'}$), 4.10 and 4.06 (2H, s, **H**^{9}), 4.08-4.00 (1H, m, **H**$^{4'}$), 3.77-3.67 (1H, m, **H**^{19}), 3.74 (3H, s, OC**H**$_3$), 3.73 (3H, s, OC**H**$_3$), 3.67-3.59 (1H, m, **H**^{19}), 3.60-3.44 (1H, m, **H**^{21}), 3.33-3.24 (1H, m, **H**$^{5'}$), 3.19 and 3.14 (1H, dd, *J* = 3.3, 10.5 Hz and *J* = 3.0, 10.3 Hz, **H**$^{5'}$), 2.75 and 2.64 (2H, t, *J* = 5.9 Hz, **H**^{20}), 2.44 and 2.43 (1H, dd, *J* = 6.8, 13.6 Hz and *J* = 6.9, 13.9 Hz, **H**$^{2'}$), 2.35 and 2.30 (1H, ddd, *J* = 3.9, 6.4, 13.7 Hz and *J* = 3.6, 6.4, 13.9 Hz, **H**$^{2'}$), 1.13 (3H, d, *J* = 6.8 Hz, **H**^{22}), 1.11 (3H, d, *J* = 7.0 Hz, **H**^{22}), 1.10 (3H, d, *J* = 6.8 Hz, **H**^{22}), 0.99 (3H, d, *J* = 6.8 Hz, **H**^{22}); **LRMS** (ES^+) *m/z*: 902 ([M + Na]$^+$, 100), 897 ([M + NH_4]$^+$, 38), 880 ([M + H]$^+$, 32), 303 (DMT^+, 24); (ES^-) *m/z*: 914/916 ([M + HCl – H]$^-$, 2:1, 21), 878 ([M – H]$^-$, 97), 825 ([M – CH_2CHCN – H]$^-$, 100).

Analytical results consistent with reported data.[117,170]

5'-*O*-(4,4'-Dimethoxytrityl)-5-(3-trifluoroacetamidopropynyl)-2'-deoxyuridine, 165[117,170,255]



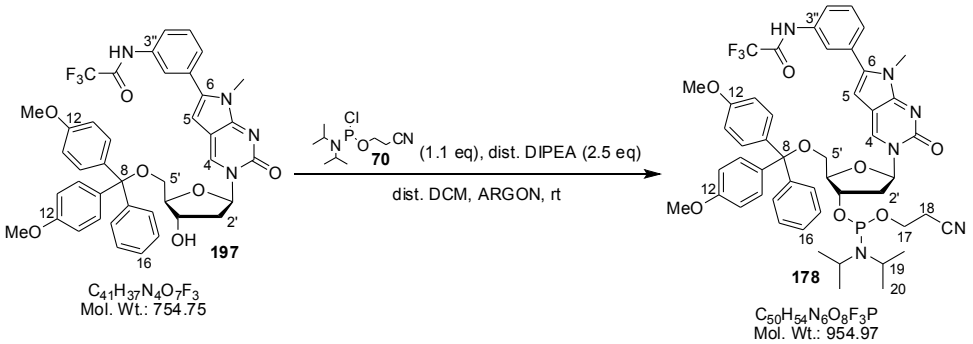
To a stirred solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **97** (10.1 g, 15.4 mmol) in anhydrous DMF (70.0 mL), under an argon atmosphere, was first added distilled Et_3N (15.0 mL, 108 mmol, 7.0 eq) followed by, in absence of light, CuI (0.59 g, 3.09 mmol, 0.2 eq) and 3-(trifluoroacetamido)prop-1-yne **68** (2.41 mL, 21.5 mmol, 1.4 eq). After stirring for 20 minutes, *tetrakis*(triphenylphosphine) palladium (0) (1.78 g, 1.54 mmol, 0.1 eq) was added and the reaction was stirred at rt for 1¾ hours, when solid Na_2EDTA (1.20 g) was added. After stirring for 5 minutes, the reaction

mixture was concentrated under high vacuum to give a dark brown syrup. The syrup was dissolved in ethyl acetate (300 mL) and washed with 5% w/v aq Na_2EDTA (pH 9, 3 × 150 mL) and sat. aq KCl (2 × 150 mL), dried (Na_2SO_4), and concentrated *in vacuo* and under high vacuum to give a golden brown foam. This reaction was repeated on the same scale as described above.

Following joint purification by column chromatography (10-30% acetone/CH_2Cl_2 + 1.0% pyridine), the desired product **165** was afforded, after drying under high vacuum, as a very pale orange foam (16.6 g, 24.4 mmol, 80%). A mixture of starting material **97** and product **165** (1.84 g) was also isolated.

$\mathbf{R}_f$ 0.19 (1:9, acetone/CH_2Cl_2 + 0.5% pyridine, A', B'); **^{19}F NMR** (282 MHz, d_6-DMSO) δ -73.9 (NHCOC**F$_3$**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.66 (N**H**^{3}), 9.98 (1H, s, N**H**COCF$_3$), 7.93 (1H, s, **H**^{6}), 7.40 (2H, d, *J* = 7.5 Hz, **H**^{16}), 7.31 (2H, t, *J* = 7.5 Hz, **H**^{17}), 7.29 (2H, d, *J* = 9.0 Hz, **H**^{12}), 7.28 (2H, d, *J* = 9.0 Hz, **H**^{12}), 7.22 (1H, tt, *J* = 1.5, 7.3 Hz, **H**^{18}), 6.89 (2H, d, *J* = 9.0 Hz, **H**^{13}), 6.89 (2H, d, *J* = 9.0 Hz, **H**^{13}), 6.10 (1H, t, *J* = 6.7 Hz, **H**$^{1'}$), 5.32 (1H, d, *J* = 4.5 Hz, 3'-O**H**), 4.27 (1H, td, *J* = 4.0, 10.3 Hz, **H**$^{3'}$), 4.06 (2H, br d, *J* = 3.8 Hz, **H**^{9}), 3.92 (1H, td, *J* = 3.3, 5.3 Hz, **H**$^{4'}$), 3.74 (6H, s, OC**H$_3$**), 3.26 (1H, dd, *J* = 5.4, 10.4 Hz, **H**$^{5'}$), 3.09 (1H, dd, *J* = 2.9, 10.4 Hz, **H**$^{5'}$), 2.29 (1H, td, *J* = 6.8, 13.6 Hz, **H**$^{2'}$), 2.19 (1H, ddd, *J* = 3.8, 6.3, 13.3 Hz, **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 161.5 (**C**^{4}), 158.1, 158.1 (**C**^{14}–OMe), 156.0 (q, $^{2}J_{CF}$ = 36.9 Hz, **C**OCF$_3$), 149.3 (**C**^{2}), 144.8 (**C**^{15}), 143.7 (**C**^{6}), 135.6, 135.2 (**C**^{11}), 129.7, 129.6 (**C**^{12}), 127.8 (**C**^{17}), 127.5 (**C**^{16}), 126.6 (**C**^{18}), 115.7 (q, $^{1}J_{CF}$ = 288.7 Hz, CO**C**F$_3$), 113.2, 113.2 (**C**^{13}), 97.8 (**C**^{5}), 87.4 (**C**^{10}Ar$_3$), 85.8 (**C**$^{4'}$,**C**^{7}), 85.1 (**C**$^{1'}$), 75.1 (**C**^{8}), 70.4 (**C**$^{3'}$), 63.8 (**C**$^{5'}$), 55.0 (O**C**H$_3$), 39.7 (**C**$^{2'}$), 29.4 (**C**^{9}); **LRMS** (ES$^+$) *m*/*z*: 803 ([M + Et$_3$N + Na]$^+$, 14), 718 ([M + K]$^+$, 14), 702 ([M + Na]$^+$, 100), 697 ([M + NH$_4$]$^+$, 5), 303 (DMT$^+$, 5); (ES$^-$) *m*/*z*: 678 ([M – H]$^-$, 100). Analytical results consistent with reported data.[170,255]

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-trifluoroacetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite, 178



$C_{41}H_{37}N_4O_7F_3$
Mol. Wt.: 754.75

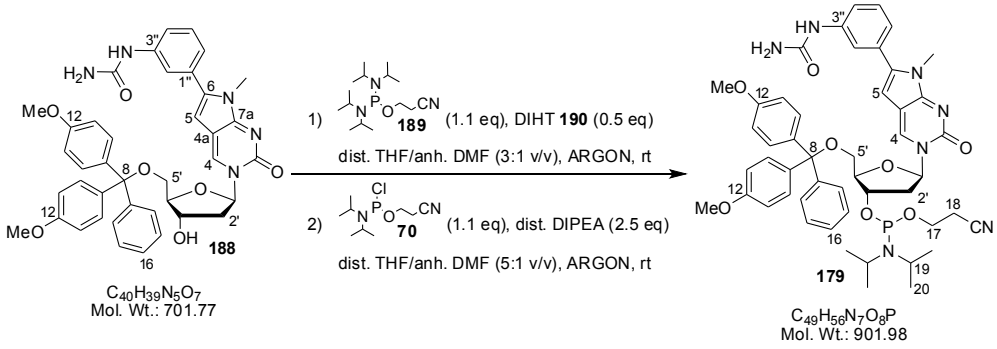
$C_{50}H_{54}N_6O_8F_3P$
Mol. Wt.: 954.97

To a stirred solution of nucleoside **197** (5.02 g, 7.39 mmol) in distilled CH_2Cl_2 (25.0 mL), strictly under an argon atmosphere and excluding moisture, was added distilled DIPEA (3.22 mL, 18.5 mmol, 2.5 eq) followed by chloro-phosphitylating reagent **70** (1.82 mL, 8.14 mmol, 1.1 eq) dropwise, and the reaction was stirred at rt for 4¾ hours. The reaction mixture was diluted with argon-degassed CH_2Cl_2 (15 mL) and washed, under an argon atmosphere, with degassed sat. aq KCl (15 mL). The aqueous was re-extracted (degassed CH_2Cl_2, 3 × 5 mL) and combined CH_2Cl_2 fractions were dried (Na_2SO_4), and concentrated *in vacuo* and dried under high vacuum to give a yellow foam. Following purification by column chromatography (degassed 1% acetone/ CH_2Cl_2 + 0.5% Et_3N → 5% acetone/CH_2Cl_2 + 0.5% Et_3N), on silica gel pre-equilibrated with Et_3N, under argon pressure, then precipitation at rt from degassed CH_2Cl_2 (3 mL) into degassed hexane (180 mL), the desired phosphoramidite **178** was afforded, after drying under high vacuum, as a diastereomeric mixture (*ca.* 2:1), as a pale yellow foam (0.18 g) containing 38.0 mol % Et_3N.HCl (recalculated yield - 0.18 mmol, 70%).

R_f 0.50, 0.57 (1:1, acetone/CH_2Cl_2 + 0.5% Et_3N, A'), 0.29, 0.34 (1:4, acetone/CH_2Cl_2 + 0.5% Et_3N, A'), 0.18, 0.25 (CH_2Cl_2 + 0.5% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.1, 148.9 (**P^{III}**); **^{19}F NMR** (282 MHz, d_6-DMSO) δ -73.5 (NHCOC**F_3**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.47 (N**H**COCF_3), 8.71 and 8.68 (1H, s, **H^4**), 7.85 (1H, br s, **$H^{2"}$**), 7.75 (1H, br dd, *J* = 1.1, 8.2 Hz, **$H^{4"}$**), 7.54 (1H, t, *J* = 8.0 Hz, **$H^{5"}$**), 7.42 and 7.40 (2H, d, *J* = 8.3 Hz and *J* = 8.8 Hz, **H^{14}**), 7.16 (1H, br dd, *J* = 0.9, 7.9 Hz, **$H^{6"}$**), 7.34-7.28 (2H, m, **H^{15}**), 7.30 and 7.27 (4H, d, *J* = 8.5 Hz and *J* = 8.8 Hz, **H^{10}**), 7.28-7.23 (1H, m, **H^{16}**), 6.89, 6.88 and 6.87 (2H, 2H and 4H, d, *J* = 8.8 Hz, *J* = 9.0 Hz and *J* = 8.5 Hz,

$\mathbf{H^{11}}$), 6.30 and 6.26 (1H, dd, J = 5.1, 6.4 Hz and J = 5.0, 6.5 Hz, $\mathbf{H^{1'}}$), 5.69 and 5.65 (1H, s, $\mathbf{H^{5}}$), 4.68 and 4.64 (1H, td, J = 5.3, 10.7 Hz and J = 5.6, 11.2 Hz, $\mathbf{H^{3'}}$), 4.17 and 4.13 (1H, dd, J = 3.5, 7.8 Hz and J = 3.8, 8.0 Hz, $\mathbf{H^{4'}}$), 3.81-3.68 (1H, m, $\mathbf{H^{17}}$), 3.70, 3.69 and 3.69 (3H, 3H and 6H, s, OC$\mathbf{H_3}$), 3.69-3.59 (1H, m, $\mathbf{H^{17}}$), 3.61-3.47 (2H, m, $\mathbf{H^{19}}$), 3.49 (3H, s, NC$\mathbf{H_3}$), 3.46-3.28 (2H, m, $\mathbf{H^{5'}}$), 2.77 and 2.67 (1H, t, J = 5.9 Hz, $\mathbf{H^{18}}$), 2.60 and 2.58 (1H, td, J = 6.7, 13.3 Hz and J = 6.5, 13.2 Hz, $\mathbf{H^{2'}}$), 2.35 (1H, td, J = 6.0, 12.2 Hz, $\mathbf{H^{2'}}$), 1.15 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$), 1.13 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$), 1.12 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$), 1.01 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$); **LRMS** (ES^+) m/z: 977 ([M + Na]$^+$, 100), 955 ([M + H]$^+$, 13); (ES^-) m/z: 953 ([M – H]$^-$, 100); **HRMS** (ES^+): calcd for $C_{50}H_{54}N_6O_8F_3P$ (M), [M + Na]$^+$ = 977.3585, found 977.3591; [M + H]$^+$ = 955.3766, found 955.3783. Found also DMT$^+$ (calcd 303.1380, found 303.1070).

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite, 179



Method 1: To a stirred solution of cyclised nucleoside **188** (0.22 g, 0.32 mmol) in distilled THF (1.5 mL) and anhydrous DMF (0.5 mL), under an argon atmosphere and excluding moisture, was added DIHT **190** (0.03 g, 0.16 mmol, 0.5 eq) followed by phosphitylating reagent **189** (0.11 mL, 0.36 mmol, 1.1 eq), and the reaction was stirred at rt. After 18½ hours, further reagent **189** was added (0.05 mL, 0.16 mmol, 0.5 eq). At 22 hours, further DIHT **190** (27.8 mg, 0.16 mmol, 0.5 eq) was added, then at 24 hours, further reagent **189** was added (0.10 mL, 0.32 mmol, 1.0 eq). After 41 hours, THF was removed *in vacuo*, and the syrup was dissolved in argon-degassed ethyl acetate (5 mL). The sticky yellow precipitate was washed with degassed ethyl acetate (3 × 5 mL), and the combined solution was washed with degassed sat. aq $NaHCO_3$ (3 × 8 mL) and sat.

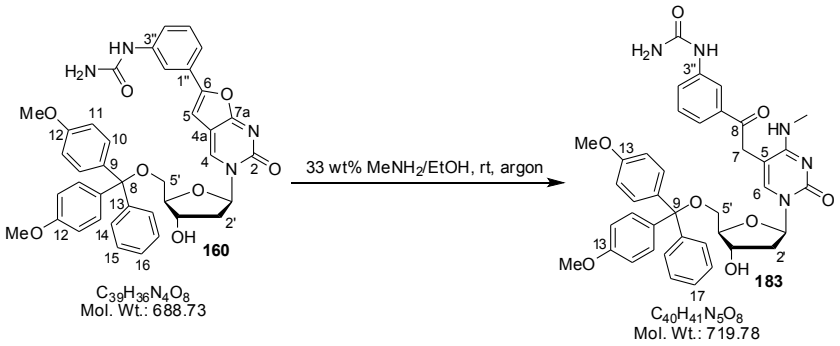
aq KCl (2 × 8 mL), dried (Na_2SO_4), concentrated *in vacuo* and dried under high vacuum to give a viscous yellow syrup. Following purification by column chromatography (degassed 2% methanol/acetone + 0.5% Et_3N) on silica gel pre-equilibrated with Et_3N, under argon pressure, and drying under high vacuum, the desired phosphoramidite **179** was afforded as a diastereomeric mixture (*ca.* 1:1), as an air-sensitive, pale yellow foam (0.08 g, 0.09 mmol, 29%).

Method 2: To a stirred solution of cyclised nucleoside **188** (0.92 g, 1.31 mmol) in distilled THF (5.0 mL) and anhydrous DMF (1.0 mL), under an argon atmosphere and excluding moisture, was added distilled DIPEA (0.58 mL, 3.30 mmol, 2.5 eq) followed by chloro-phosphitylating reagent **70** (0.33 mL, 1.46 mmol, 1.1 eq), and the reaction was stirred at rt for 2 hours. The reaction mixture was concentrated *in vacuo* and under high vacuum to give a viscous orange syrup, which was dissolved in argon-degassed CH_2Cl_2 (80 mL). The solution was washed with degassed sat. aq KCl (60 mL) and the aqueous was re-extracted (degassed CH_2Cl_2, 3 × 10 mL). The combined CH_2Cl_2 fractions were dried (Na_2SO_4) and concentrated *in vacuo* and under high vacuum to give an orange foam. Following purification by column chromatography (degassed acetone + 0.5% Et_3N ⟶ 2% methanol/acetone + 0.5% Et_3N) on silica gel pre-equilibrated with Et_3N, under argon pressure, and precipitation at rt from degassed CH_2Cl_2 (6 mL) into degassed hexane (350 mL), the desired phosphoramidite **179** was afforded, after drying under high vacuum, as a diastereomeric mixture (*ca.* 1:1), as an air-sensitive pale yellow foam (0.60 g, 0.67 mmol, 51%).

$\mathbf{R}_f$ 0.30 (9:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.24, 0.31 (5:95, methanol/acetone + 0.3% Et_3N, A'), 0.17 (acetone + 0.3% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.2, 149.0 (**P^{III}**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 8.69 (1H, s, N**H**CONH$_2$), 8.66 and 8.63 (1H, s, **H^4**), 7.63 (1H, br s, **$H^{2"}$**), 7.44-7.38 (1H, m, **$H^{4"}$**), 7.42 and 7.40 (2H, d, *J* = 8.0 Hz and *J* = 8.5 Hz, **H^{14}**), 7.34 (1H, t, *J* = 7.9 Hz, **$H^{5"}$**), 7.34-7.24 (3H, m, **H^{15}**,**H^{16}**), 7.30 and 7.27 (4H, d, *J* = 8.0 Hz and *J* = 8.8 Hz, **H^{10}**), 7.02 (1H, br d, *J* = 7.5 Hz, **$H^{6"}$**), 6.89 and 6.87 (2H, d, *J* = 8.3 Hz and *J* = 8.6 Hz, **H^{11}**), 6.89 and 6.87 (2H, d, *J* = 8.5 Hz, **H^{11}**), 6.31 and 6.27 (1H, app. t, *J* = 5.8 Hz, **$H^{1'}$**), 5.92 (2H, br s, N**H$_2$**), 5.63 and 5.59 (1H, s, **H^5**), 4.68 and 4.64 (1H, td, *J* = 5.4, 16.1 Hz and *J* = 5.9, 17.9 Hz, **$H^{3'}$**), 4.17 and 4.13 (1H, dd, *J* = 3.4, 7.7 Hz and *J* = 3.8, 8.0 Hz, **$H^{4'}$**), 3.81-3.68 (1H, m, **H^{17}**), 3.71, 3.70 and 3.70 (3H, 3H and 6H, s, OC**H$_3$**), 3.69-3.61 (1H, m, **H^{17}**), 3.63-3.49 (2H, m, **H^{19}**), 3.46 (3H, s, NC**H$_3$**), 3.49-3.44 and 3.41 (1H, m and dd, *J* = 3.5, 11.5 Hz, **$H^{5'}$**),

3.46-3.41 and 3.37 (1H, m and dd, J = 2.9, 10.8 Hz, $\mathbf{H^{5'}}$), 2.77 and 2.67 (1H, t, J = 5.9 Hz, $\mathbf{H^{18}}$), 2.59 (1H, td, J = 6.7, 13.6 Hz, $\mathbf{H^{2'}}$), 2.34 (1H, td, J = 6.0, 11.9 Hz, $\mathbf{H^{2'}}$), 1.15 (3H, d, J = 7.3 Hz, $\mathbf{H^{20}}$), 1.13 (3H, d, J = 7.3 Hz, $\mathbf{H^{20}}$), 1.12 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$), 1.02 (3H, d, J = 6.8 Hz, $\mathbf{H^{20}}$); **LRMS** (ES^+) m/z: 940 ($[M + K]^+$, 8), 924 ($[M + Na]^+$, 45), 902 ($[M + H]^+$, 78), 303 (DMT^+, 100); **HRMS** (ES^+): calcd for $C_{49}H_{56}N_7O_8P$ (M), $[M + Na]^+$ = 924.3820, found 924.3825; $[M + H]^+$ = 902.4001, found 902.4000. Found also DMT^+ (calcd 303.1380, found 303.1381).

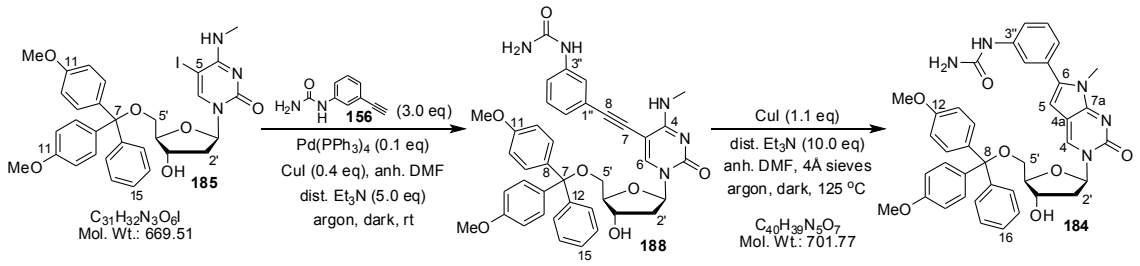
5'-*O*-(4,4'-Dimethoxytrityl)-5-[(3-ureidobenzoyl)methyl]-4*N*-methyl-2'-deoxycytidine, 183



A suspension of nucleoside **160** (0.20 g, 0.29 mmol) in 33 wt% $MeNH_2$/ethanol (2.50 mL) was stirred under an argon atmosphere, at rt for 1 hour. The suspension dissolved over 25 minutes. The solvent was removed *in vacuo* to give a green crystalline solid. Following purification by column chromatography (5-20% methanol/ethyl acetate + 0.5% Et_3N), and drying under high vacuum, the desired ring-opened nucleoside **183** was afforded as a yellow glassy solid (0.13 g, 0.18 mmol, 61%).

$\mathbf{R_f}$ 0.11 (9:1, acetone/ethyl acetate + 0.3% Et_3N, A'), 0.25 (1:4, methanol/ethyl acetate + 0.3% Et_3N, A'); **LRMS** (ES^+) m/z: 758 ($[M + K]^+$, 10), 742 ($[M + Na]^+$, 30), 720 ($[M + H]^+$, 5), 303 (DMT^+, 100); (ES^-) m/z: 816 ($[M + H_3PO_4 - H]^-$, 100), 718 ($[M - H]^-$, 41), 675 ($[M - O{=}C{=}NH - H]^-$, 52); **HRMS** (ES^+): calcd for $C_{40}H_{41}N_5O_8$ (M), $[M + H]^+$ = 720.3028, found 720.3024.

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 184



To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine **185** (1.25 g, 1.87 mmol) in anhydrous DMF (6.0 mL), under an argon atmosphere and in absence of light, were added distilled Et_3N (1.30 mL, 9.33 mmol, 5.0 eq), CuI (0.14 g, 0.75 mmol, 0.4 eq), then alkyne **156** (0.90 g, 5.61 mmol, 3.0 eq). The reaction mixture was stirred for 15 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.53 g, 0.46 mmol, 0.1 eq) was added and the reaction was stirred at rt for 2¼ hours. The reaction mixture was diluted with methanol/acetone (1:1 v/v, 10 mL), filtered and the residue was washed with methanol/acetone (1:1 v/v, 3 × 10 mL). The combined filtrate was concentrated *in vacuo* and under high vacuum, co-evaporating with methanol/toluene to give a dark brown foam. The crude material was partially-purified by column chromatography (80-100% acetone/CH_2Cl_2 + 0.5% Et_3N → 5-20% methanol/acetone + 0.5% Et_3N), to afford, after drying under high vacuum, the product **188** plus alkyne-dimer impurity, as a pale brown foam (1.49 g).

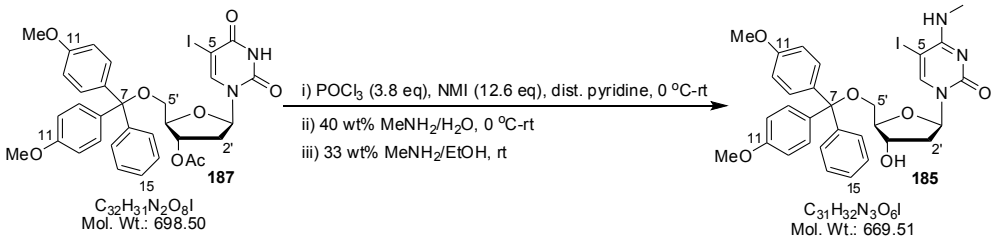
R$_f$ 0.18 (**188**), 0.12 (alkyne dimer) (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.11 (**188**), 0.21 (alkyne dimer) (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.44 (**188** + alkyne dimer), (1:9, methanol/acetone + 0.3% Et_3N, A').

To a stirred solution of partially-purified product **188** (1.48 g) in anhydrous DMF (10.0 mL) and distilled Et_3N (2.60 mL, 18.7 mmol, 10.0 eq), under an argon atmosphere, in absence of light, were added 4Å molecular sieves then CuI (0.39 g, 2.05 mmol, 1.1 eq), and the reaction was stirred at 125 °C for 40 minutes. The mixture was allowed to cool and diluted with methanol (20 mL). The mixture was filtered, the residue was washed with methanol (3 × 10 mL) and the combined filtrate was concentrated *in vacuo* and under high vacuum to give a green foam/resin. The compound was twice-purified by column chromatography (0-10% methanol/acetone + 0.5% Et_3N), then by precipitation

from methanol/acetone (1:1 v/v, 8 mL) into diethyl ether (230 ml), to afford, after drying under high vacuum, the desired *N*-methylpyrrolo-dC product **184**, as a very pale yellow powdery solid (1.03 g, 1.47 mmol, 79% over two steps).

R_f 0.24 (1:9, methanol/acetone + 0.3% Et_3N, A'), 0.49 (1:4, methanol/CH_2Cl_2 + 0.3% Et_3N, A'); **IR** (solid): ν 3324 (br m, O–H, N–H), 2932, 2836 (w, C–H), 1651 (s, C=O), 1606 (s, C=C), 1556 (vs, C=N), 1507 (s, Aryl C=C), 1479 (s), 1441 (m), 1402 (s), 1338, 1303 (m), 1248 (s, O–H), 1176 (s), 1095, 1032 (s, C–O), 904 (m), 828 (s, C–H), 773, 726 (s), 701 (s, Aryl C–H), 659, 584, 534 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 8.71 (1H, s, N**H**$CONH_2$), 8.63 (1H, s, **H^{4}**), 7.63 (1H, t, *J* = 1.8 Hz, **H$^{2"}$**), 7.41 (1H, ddd, *J* = 1.0, 2.0, 7.0 Hz, **H$^{4"}$**), 7.41 (2H, d, *J* = 7.0 Hz, **H^{14}**), 7.34 (1H, t, *J* = 8.0 Hz, **H$^{5"}$**), 7.32 (2H, tt, *J* = 1.5, 7.0 Hz, **H^{15}**), 7.29 (2H, d, *J* = 9.0 Hz, **H^{10}**), 7.28 (2H, d, *J* = 8.5 Hz, **H^{10}**), 7.28 (1H, t, *J* = 8.0 Hz, **H^{16}**), 7.02 (1H, td, *J* = 1.3, 7.5 Hz, **H$^{6"}$**), 6.89 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.89 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.26 (1H, dd, *J* = 5.0, 6.5 Hz, **H$^{1'}$**), 5.92 (2H, s, CON**H$_2$**), 5.59 (1H, s, **H^{5}**), 5.40 (1H, d, *J* = 4.8 Hz, 3'-O**H**), 4.43 (1H, qn, *J* = 5.5 Hz, **H$^{3'}$**), 4.00 (1H, dd, *J* = 4.0, 8.0 Hz, **H$^{4'}$**), 3.70 (3H, s, OC**H$_3$**), 3.69 (3H, s, OC**H$_3$**), 3.46 (3H, s, NC**H$_3$**), 3.40 (1H, dd, *J* = 4.0, 10.5 Hz, **H$^{5'}$**), 3.29 (1H, dd, *J* = 2.5, 10.5 Hz, **H$^{5'}$**), 2.44 (1H, td, *J* = 6.3, 13.1 Hz, **H$^{2'}$**), 2.18 (1H, ddd, *J* = 5.0, 6.5, 13.6 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 158.5 (**C^{7a}**), 158.1 (**C^{12}**–OCH_3), 155.9 (**C**$CONH_2$), 153.5 (**C^{2}**), 144.5 (**C^{13}**), 142.6 (**C^{6}**), 140.9 (**C$^{3"}$**–NHR), 136.2 (**C^{4}**), 135.4, 135.1 (**C^{9}**), 130.8 (**C$^{1"}$**), 129.8, 129.7 (**C^{10}**), 129.1 (**C$^{5"}$**), 127.9 (**C^{15}**), 127.7 (**C^{14}**), 126.8 (**C^{16}**), 120.9 (**C$^{6"}$**), 118.1 (**C$^{4"}$**), 117.4 (**C$^{2"}$**), 113.3 (**C^{11}**), 107.6 (**C^{4a}**), 98.6 (**C^{5}**), 86.7 (**C$^{1'}$**), 86.1 (**C^{8}**Ar_3), 85.7 (**C$^{4'}$**), 69.0 (**C$^{3'}$**), 62.6 (**C$^{5'}$**), 55.0 (O**C**H_3), 41.5 (**C$^{2'}$**), 29.3 (N**C**H_3); **LRMS** (ES^+) *m*/*z*: 1425 ($[2M + Na]^+$, 8), 724 ($[M + Na]^+$, 100), 702 ($[M + H]^+$, 50); **HRMS** (ES^+): calcd for $C_{40}H_{39}N_5O_7$ (M), $[M + Na]^+$ = 724.2742, found 724.2733. Found also DMT^+ (calcd 303.1380, found 303.1382); **Anal.** Calcd $C_{40}H_{39}N_5O_7$ (701.77), C 68.46%, H 5.60%, N 9.97%, found C 65.70%, H 5.82%, N 9.35%.

5'-*O*-(4,4'-Dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine, 185



$C_{32}H_{31}N_2O_8I$
Mol. Wt.: 698.50

$C_{31}H_{32}N_3O_6I$
Mol. Wt.: 669.51

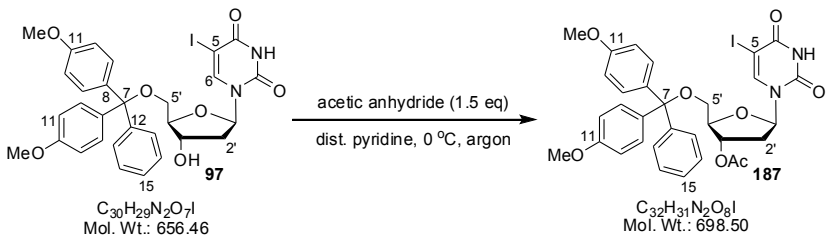
Method 1: To a stirred solution of acetylated nucleoside **187** (1.44 g, 2.07 mmol) and *N*-methylimidazole (2.10 mL, 26.3 mmol, 12.8 eq) in distilled pyridine (50.0 mL), under an argon atmosphere, at 0 °C, distilled phosphorous oxychloride (0.73 mL, 7.83 mmol, 3.8 eq) was added dropwise over 10 minutes. The yellow suspension of intermediate **186** was stirred at rt for 1 hour. A 40 wt% $MeNH_2$/water solution (10.0 mL, 116 mmol $MeNH_2$, 55.9 eq $MeNH_2$) was added whilst cooling in ice, and the cloudy solution was stirred at rt for 90 minutes. The pyridine was removed under high vacuum, and the mixture was diluted with water (200 mL). The aqueous solution was extracted with CH_2Cl_2 (4 × 100 mL), then the combined organic layer was washed with sat. aq KCl (200 mL) and dried (Na_2SO_4). Triethylamine (6.0 mL) was added to prevent detritylation and the solution was concentrated *in vacuo*, and dried under high vacuum. The viscous brown oil was dissolved in 33 wt% $MeNH_2$/ethanol (10.0 mL, 80.3 mmol $MeNH_2$, 38.8 eq $MeNH_2$) and stirred at rt for 40 minutes, then concentrated *in vacuo* and dried under high vacuum. Following purification by column chromatography (0-2% methanol/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the 4*N*-methyl-dC nucleoside **185** was afforded, as a pale yellow foam (1.24 g, 1.86 mmol, 90%).

Method 2: Acetylated nucleoside **187** (6.85 g, 9.80 mmol) was treated with NMI and $POCl_3$ in pyridine, then 40 wt% $MeNH_2$/water and worked up, as described in Method 1. After treatment with 33 wt% $MeNH_2$/ethanol as described above, the resulting syrup was twice-purified by column chromatography (0-3% methanol/CH_2Cl_2 + 0.5% Et_3N, 0-15% acetone/CH_2Cl_2 + 0.5% Et_3N → 60% acetone/CH_2Cl_2 + 0.5% Et_3N) to yield a syrup containing significant NMI. Two further purifications (30-100% acetone/toluene + 0.5% Et_3N; 10-25% acetone/toluene + 1.0% pyridine) failed to remove NMI, however the product was successfully recrystallised from toluene (100 mL), and dried under high vacuum over KOH to afford the product **185** as a fluffy white crystalline solid (5.26 g, 7.86 mmol, 80%).

Method 3: Acetylated nucleoside **187** (8.27 g, 11.8 mmol) was treated with NMI and $POCl_3$ in pyridine, then 40 wt% $MeNH_2$/water and the reaction concentrated as described in Method 1. The resulting syrup was treated with 33 wt% $MeNH_2$/ethanol as described above, without an aqueous workup, then twice-purified by column chromatography (0-100% acetone/CH_2Cl_2 + 0.5% Et_3N → 5-15% methanol/acetone + 1.0% Et_3N). The orange/brown syrup was further purified by a combination of precipitation from toluene into diethyl ether, removal of NMI by soaking in diethyl ether, and recrystallisation from ethyl acetate to afford the product **185** as a white to off-white powdery/crystalline solid (3.62 g, 5.40 mmol, 46%).

R_f 0.40 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.36 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.57 (1:9, methanol/acetone + 0.3% Et_3N, A'); **Mp** 141-144 °C (ethyl acetate/diethyl ether); **IR** (solid): ν 3291 (br w, N–H), 3059, 2930, 2835 (w, C–H), 1634 (s, C=O), 1606 (vs, C=O, C=C), 1547 (s, C=O), 1505 (vs, C=C), 1397 (m, O–H), 1284 (s, C–O), 1246 (vs, C–N), 1173 (s), 1092, 1030 (m, C–O), 825, 779 (s, Aryl C=C), 700, 583 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 7.93 (1H, s, **H^6**), 7.40 (2H, d, *J* = 7.5 Hz, **H^{13}**), 7.32 (2H, t, *J* = 7.3 Hz, **H^{14}**), 7.29 (4H, d, *J* = 9.0 Hz, **H^9**), 7.22 (1H, tt, *J* = 2.0, 7.3 Hz, **H^{15}**), 7.06 (1H, q, *J* = 4.5 Hz, N**H**CH_3), 6.89 (4H, d, *J* = 9.0 Hz, **H^{10}**), 6.12 (1H, dd, *J* = 6.2, 7.4 Hz, **H$^{1'}$**), 5.25 (1H, d, *J* = 4.3 Hz, 3'-O**H**), 4.20 (1H, dt, *J* = 3.3, 6.4 Hz, **H$^{3'}$**), 3.91 (1H, dd, *J* = 3.3, 6.5 Hz, **H$^{4'}$**), 3.74 (6H, s, OC**H$_3$**), 3.20 (1H, dd, *J* = 3.3, 10.8 Hz, **H$^{5'}$**), 3.17 (1H, dd, *J* = 4.5, 10.8 Hz, **H$^{5'}$**), 2.80 (3H, d, *J* = 4.5 Hz, NHC**H$_3$**), 2.21 (1H, ddd, *J* = 3.0, 6.0, 13.3 Hz, **H$^{2'}$**), 2.09 (1H, td, *J* = 6.8, 13.6 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 163.0 (**C^4**), 159.8 (**C^{11}**–OMe), 155.3 (**C^2**), 146.8 (**C^6**), 146.4 (**C^{12}**), 137.2, 137.1 (**C^8**), 131.4 (**C^9**), 129.6 (**C^{14}**), 129.4 (**C^{13}**), 128.4 (**C^{15}**), 115.0 (**C^{10}**), 87.5 (**C^7**Ar_3), 87.4 (**C$^{4'}$**), 87.0 (**C$^{1'}$**), 72.4 (**C$^{3'}$**), 65.4 (**C$^{5'}$**), 59.9 (**C^5**–I), 56.8 (O**C**H_3), 42.4 (**C$^{2'}$**), 30.4 (NH**C**H_3); **LRMS** (ES^+) *m/z*: 692 ($[M + Na]^+$, 12), 670 ($[M + H]^+$, 3), 303 (DMT^+, 100); (ES^-) *m/z*: 668 ($[M – H]^-$, 100); **HRMS** (ES^+): calcd for $C_{31}H_{32}N_3O_6I$ (M), $[M + K]^+$ = 708.0967, found 708.0962, $[M + Na]^+$ = 692.1228, found 692.1219, $[M + H]^+$ = 670.1409, found 670.1414. Found also DMT^+ (calcd 303.1380, found 303.1386); **Anal.** Calcd $C_{31}H_{32}N_3O_6I$ (669.51), C 55.61%, H 4.82%, N 6.27%, found C 54.80%, H 4.75%, N 6.05%.

3'-*O*-Acetyl-5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine, 187[209]



$C_{30}H_{29}N_2O_7I$
Mol. Wt.: 656.46

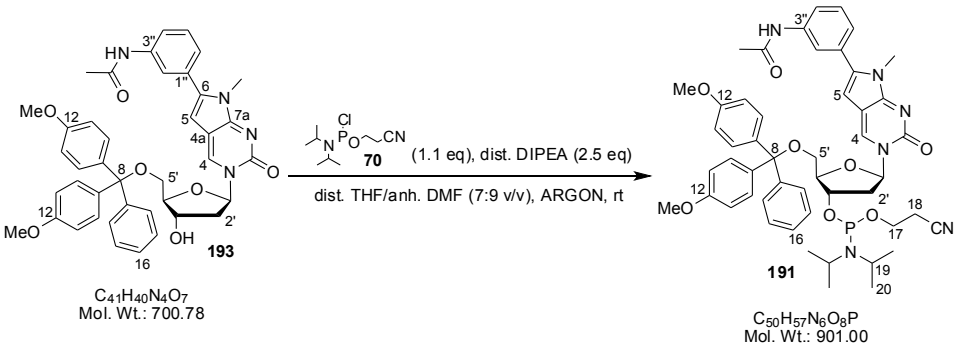
$C_{32}H_{31}N_2O_8I$
Mol. Wt.: 698.50

To a stirred solution of nucleoside **97** (10.0 g, 15.2 mmol) in distilled pyridine (40.0 mL), under an argon atmosphere, at 0 °C was added acetic anhydride (7.20 mL, 76.3 mmol, 5.0 eq) dropwise over 20 minutes. The reaction was stirred at 0 °C for 15 minutes then allowed to warm to rt. After 2¾ hours, the reaction mixture was concentrated *in vacuo* and under high vacuum, co-evaporating with acetone (2 × 50 mL) and diethyl ether (2 × 50 mL), to give white foam/syrup. The syrup was dissolved in ethyl acetate (300 mL) and washed with water (3 × 75 mL). The aqueous was re-extracted (ethyl acetate, 40 mL) and combined organic fractions were washed with sat. aq KCl (120 mL), dried (Na_2SO_4) and concentrated *in vacuo* and under high vacuum to give a white foam. Following purification by column chromatography (0-2% methanol/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **187** was afforded as a white foam (10.3 g, 14.7 mmol, 97%).

R_f 0.31 (1:9, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.48 (5:95, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.33 (7:3, ethyl acetate/hexane, A'); **IR** (solid): ν 3056 (w, C–H, N–H), 2933, 2835 (w, C–H), 1683 (br s, C=O), 1606, 1506 (m, Aryl C=C), 1442 (m, C–H), 1379 (w, CH_3), 1244 (vs, C–O), 1174 (s), 11.05 (m), 1073 (s), 1027 (s, C–O), 826, 754, 700 (s, Aryl C–H), 583 (s), 545 (m) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.75 (1H, s, N**H**^{3}), 8.05 (1H, s, **H**^{6}), 7.39 (2H, d, *J* = 8.0 Hz, **H**^{13}), 7.32 (2H, t, *J* = 7.3 Hz, **H**^{14}), 7.28 (2H, d, *J* = 9.0 Hz, **H**^{9}), 7.28 (2H, d, *J* = 9.0 Hz, **H**^{9}), 7.23 (1H, tt, *J* = 1.1, 7.2 Hz, **H**^{15}), 6.89 (4H, d, *J* = 8.8 Hz, **H**^{10}), 6.10 (1H, dd, *J* = 6.1, 8.2 Hz, **H**$^{1'}$), 5.23 (1H, td, *J* = 2.4, 6.4 Hz, **H**$^{3'}$), 4.07 (1H, td, *J* = 2.9, 4.5 Hz, **H**$^{4'}$), 3.74 (6H, s, OC**H**$_3$), 3.32 (1H, dd, *J* = 4.9, 10.4 Hz, **H**$^{5'}$), 3.22 (1H, dd, *J* = 3.1, 10.4 Hz, **H**$^{5'}$), 2.44 (1H, td, *J* = 7.0, 11.5 Hz, **H**$^{2'}$), 2.33 (1H, ddd, *J* = 2.0, 6.0, 14.3 Hz, **H**$^{2'}$), 2.03 (3H, s, COC**H**$_3$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 169.9 (**C**OCH$_3$), 160.5 (**C**^{4}), 158.1 (**C**^{11}–OCH$_3$), 150.1 (**C**^{2}), 144.6 (**C**^{12}), 144.1 (**C**^{6}), 135.3, 135.2 (**C**^{8}), 129.7 (**C**^{9}), 127.9 (**C**^{14}), 127.6 (**C**^{13}), 126.7 (**C**^{15}), 113.3 (**C**^{10}), 86.0 (**C**^{7}Ar$_3$), 84.7 (**C**$^{1'}$), 83.1 (**C**$^{4'}$), 74.1 (**C**$^{3'}$), 70.2 (**C**^{5}–I), 63.4 (**C**$^{5'}$), 55.0 (O**C**H$_3$), 36.8 (**C**$^{2'}$), 20.7 (CO**C**H$_3$); **LRMS** (ES^+) *m*/*z*: 721 ($[M + Na]^+$, 12), 303 (DMT^+, 100); (ES^-) *m*/*z*: 697 ($[M - H]^-$, 100); **HRMS** (ES^+): calcd for $C_{32}H_{31}N_2O_8I$ (M), [M +

$Et_3N + H]^+$ = 800.2402, found 800.2369; $[M + Na]^+$ = 721.1017, found 721.1007. Found also DMT^+ (calcd 303.1380, found 303.1383).

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite, 191



$C_{41}H_{40}N_4O_7$
Mol. Wt.: 700.78

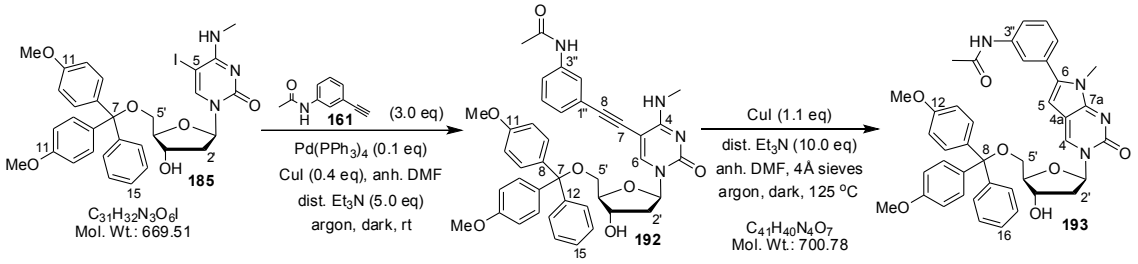
$C_{50}H_{57}N_6O_8P$
Mol. Wt.: 901.00

To a stirred solution of nucleoside **193** (1.37 g, 1.95 mmol) in distilled THF (3.5 mL) and anhydrous DMF (4.5 mL), strictly under an argon atmosphere and excluding moisture, was added distilled DIPEA (0.86 mL, 4.91 mmol, 2.5 eq) followed by chloro-phosphitylating reagent **70** (0.48 mL, 2.15 mmol, 1.1 eq) dropwise, and the reaction was stirred at rt for 1¼ hours. The THF was removed *in vacuo* and the residue was dissolved in argon-degassed ethyl acetate (60 mL). The solution was washed with degassed sat. aq KCl (3 × 40 mL), dried (Na_2SO_4), and concentrated *in vacuo* and dried under high vacuum to give a pale orange foam. Following purification by column chromatography (7:3, acetone/CH_2Cl_2 + 0.5% Et_3N), on silica gel pre-equilibrated with Et_3N, under argon pressure, the desired phosphoramidite **191** was afforded, after drying under high vacuum, as a diastereomeric mixture (*ca.* 1:1), as an air-sensitive, pale milky yellow foam (1.51 g, 1.68 mmol, 86%).

R_f 0.33, 0.46 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **^{31}P NMR** (121 MHz, d_6-DMSO) δ 149.2, 149.0 (**P^{III}**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.08 (1H, s, N**H**$COCH_3$), 8.68 and 8.65 (1H, s, **H^4**), 7.79 (1H, br s, **$H^{2"}$**), 7.62 (1H, br d, J = 8.0 Hz, **$H^{4"}$**), 7.42 (1H, t, J = 7.8 Hz, **$H^{5"}$**), 7.42 and 7.40 (2H, d, J = 8.0 Hz and J = 8.5 Hz, **H^{14}**), 7.34-7.29 (2H, m, **H^{15}**), 7.30 and 7.28 (4H, d, J = 9.0 Hz, **H^{10}**), 7.30-7.23 (1H, m, **H^{16}**), 7.16 (1H, br d, J =

7.5 Hz, $\mathbf{H^{6"}}$), 6.89, 6.88 and 6.87 (2H, 2H and 4H, d, *J* = 9.0 Hz, *J* = 8.5 Hz and *J* = 9.0 Hz, $\mathbf{H^{11}}$), 6.31 and 6.27 (1H, app. t, *J* = 5.8 Hz, $\mathbf{H^{1'}}$), 5.66 and 5.63 (1H, s, $\mathbf{H^{5}}$), 4.68 and 4.65 (1H, ddd, *J* = 5.5, 16.3 Hz and *J* = 6.0, 16.7 Hz, $\mathbf{H^{3'}}$), 4.18 and 4.14 (1H, dd, *J* = 3.3, 7.8 Hz and *J* = 3.8, 8.3 Hz, $\mathbf{H^{4'}}$), 3.82-3.69 (1H, m, $\mathbf{H^{17}}$), 3.70, 3.70 and 3.69 (3H, 3H and 6H, s, OC$\mathbf{H_3}$), 3.70-3.60 (1H, m, $\mathbf{H^{17}}$), 3.62-3.48 (2H, m, $\mathbf{H^{19}}$), 3.47 (3H, s, NC$\mathbf{H_3}$), 3.47 and 3.43-3.39 (1H, dd and m, *J* = 4.5, 11.0 Hz, $\mathbf{H^{5'}}$), 3.44 and 3.38 (1H, dd, *J* = 3.5, 10.5 Hz and *J* = 3.0, 10.5 Hz, $\mathbf{H^{5'}}$), 2.77 and 2.67 (1H, t, *J* = 6.0 Hz, $\mathbf{H^{18}}$), 2.61 and 2.59 (1H, td, *J* = 6.9, 13.4 Hz and *J* = 6.5, 13.1 Hz, $\mathbf{H^{2'}}$), 2.35 (1H, td, *J* = 6.3, 13.1 Hz, $\mathbf{H^{2'}}$), 2.08 (3H, s, COC$\mathbf{H_3}$), 1.15 (3H, d, *J* = 7.0 Hz, $\mathbf{H^{20}}$), 1.13 (3H, d, *J* = 7.0 Hz, $\mathbf{H^{20}}$), 1.12 (3H, d, *J* = 7.0 Hz, $\mathbf{H^{20}}$), 1.02 (3H, d, *J* = 7.0 Hz, $\mathbf{H^{20}}$); **LRMS** (ES^+) *m*/*z*: 939 ($[M + K]^+$, 5), 923 ($[M + Na]^+$, 100), 901 ($[M + H]^+$, 9); **HRMS** (ES^+): calcd for $C_{50}H_{57}N_6O_8P$ (M), $[M + Et_3N + H]^+$ = 1002.5253, found 1002.5261; $[M + Na]^+$ = 923.3868, found 923.3886; $[M + H]^+$ = 901.4048, found 901.4033.

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 193



To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine **185** (1.69 g, 2.53 mmol) in anhydrous DMF (6.0 mL), under an argon atmosphere and in absence of light, were added distilled Et_3N (1.77 mL, 12.7 mmol, 5.0 eq), CuI (0.19 g, 1.01 mmol, 0.4 eq), then alkyne **161** (1.21 g, 7.60 mmol, 3.0 eq). The reaction mixture was stirred for 20 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.29 g, 0.25 mmol, 0.1 eq) was added and the reaction was stirred at rt for 1 hour. The reaction mixture was concentrated under high vacuum, co-evaporating with acetone to give a black/brown foam. The residue was dissolved in CH_2Cl_2 (100 mL) and washed with 5% w/v aq Na_2EDTA (pH 9, 3 × 70 mL). The aqueous was re-extracted (CH_2Cl_2, 2 × 50 mL) and combined organic fractions were washed with sat. aq KCl (100 mL), dried

(Na_2SO_4), and concentrated *in vacuo* and under high vacuum to give a golden brown foam. The crude material was partially-purified by column chromatography (10-30% acetone/CH_2Cl_2 + 1.0% pyridine → 50-70% acetone/CH_2Cl_2 + 1.0% pyridine), to afford, after drying under high vacuum, the product **192** plus alkyne-dimer impurity, as a yellow glassy foam (1.99 g).

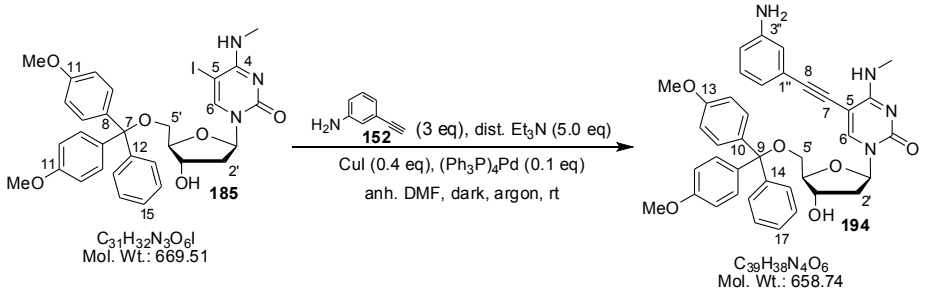
R_f 0.16 (3:2, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.04 (1:4, acetone/CH_2Cl_2 + 0.5% pyridine, A'), 0.28 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.32 (4:1, acetone/CH_2Cl_2 + 1.0% pyridine, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.97 (1H, s, N**H**$COCH_3$), 8.00 (1H, s, **H^6**), 7.79 (1H, s, **H$^{2"}$**), 7.46 (1H, ddd, *J* = 1.0, 2.0, 8.5 Hz, **H$^{4"}$**), 7.43 (1H, q, *J* = 5.0 Hz, N**H**CH_3), 7.40 (1H, dd, *J* = 1.3, 8.3 Hz, **H^{15}**), 7.29 (2H, d, *J* = 9.0 Hz, **H^{11}**), 7.29-7.25 (1H, m, **H$^{5"}$**), 7.28 (2H, d, *J* = 8.5 Hz, **H^{11}**), 7.24 (2H, t, *J* = 7.8 Hz, **H^{16}**), 7.15 (1H, tt, *J* = 1.3, 7.3 Hz, **H^{17}**), 6.90 (1H, td, *J* = 1.3, 7.8 Hz, **H$^{6"}$**), 6.84 (2H, d, *J* = 9.0 Hz, **H^{12}**), 6.83 (2H, d, *J* = 9.0 Hz, **H^{12}**), 6.16 (1H, t, *J* = 6.8 Hz, **H$^{1'}$**), 5.28 (1H, d, *J* = 4.5 Hz, 3'-O**H**), 4.26 (1H, dt, *J* = 3.5, 6.7 Hz, **H$^{3'}$**), 3.97 (1H, td, *J* = 3.0, 5.0 Hz, **H$^{4'}$**), 3.66 (6H, s, OC**H$_3$**), 3.23 (1H, dd, *J* = 5.0, 10.5 Hz, **H$^{5'}$**), 3.15 (1H, dd, *J* = 2.5, 10.5 Hz, **H$^{5'}$**), 2.86 (3H, d, *J* = 4.5 Hz, NHC**H$_3$**), 2.27 (1H, ddd, *J* = 3.5, 6.3, 13.6 Hz, **H$^{2'}$**), 2.16 (1H, td, *J* = 6.7, 13.6 Hz, **H$^{2'}$**), 2.06 (3H, s, COC**H$_3$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 168.4 (**C**OCH_3), 161.7 (**C^4**–$NHCH_3$), 158.0, 158.0 (**C^{13}**–OCH_3), 153.2 (**C^2**), 144.7 (**C^{14}**), 143.0 (**C^6**), 139.1 (**C$^{3"}$**), 135.6, 135.3 (**C^{10}**), 129.7, 129.6 (**C^{11}**), 128.5 (**C$^{5"}$**), 127.8 (**C^{16}**), 127.6 (**C^{15}**), 126.6 (**C^{17}**), 126.4 (**C$^{6"}$**), 122.5 (**C$^{1"}$**), 121.9 (**C$^{2"}$**), 119.4 (**C$^{4"}$**), 113.2, 113.1 (**C^{12}**), 94.3 (**C^8**), 90.3 (**C^7**), 85.9 (**C$^{4'}$**), 85.8 (**C^9**Ar_3), 85.6 (**C$^{1'}$**), 80.4 (**C^5**), 70.7 (**C$^{3'}$**), 63.8 (**C$^{5'}$**), 54.9 (O**C**H_3), 40.8 (**C$^{2'}$**), 27.8 (NH**C**H_3), 23.9 (CO**C**H_3); **LRMS** (ES^+) *m*/*z*: 723 ([M + Na]$^+$, 100), 303 (DMT^+, 29); (ES^-) *m*/*z*: 699 ([M – H]$^-$, 100); **HRMS** (ES^+): calcd for $C_{41}H_{40}N_4O_7$ (M), [M + H]$^+$ = 701.2970, found 701.2974. Found also DMT^+ (calcd 303.1380, found 303.1381).

To a stirred solution of partially-purified alkynylated product **192** (1.92 g) in anhydrous DMF (8.5 mL) and distilled Et_3N (3.40 mL, 24.4 mmol, 10.0 eq), under an argon atmosphere and in absence of light, were added 4Å molecular sieves then CuI (0.51 g, 2.68 mmol, 1.1 eq), and the reaction was stirred at 125 °C for 35 minutes. The mixture was allowed to cool then filtered and the filtrate was concentrated under high vacuum to give a sticky green/brown foam. The residue was dissolved in CH_2Cl_2 (150 mL) and filtered, and the filtrate was washed with 5% w/v aq Na_2EDTA (pH 9, 3 × 100 mL), sat. aq KCl (100 mL), dried (Na_2SO_4), and concentrated *in vacuo* and under high vacuum to

give a green/orange foam. Following purification by column chromatography (0-10% methanol/acetone + 0.5% Et_3N), and drying under high vacuum, the desired cyclised product **193** was afforded, as a yellow powdery solid (1.43 g, 2.04 mmol, 84% over two steps).

R_f 0.09 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.14 (4:1, acetone/CH_2Cl_2 + 1.0% pyridine, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 10.08 (N**H**$COCH_3$), 8.65 (1H, s, **H^{4}**), 7.79 (1H, br s, **H$^{2"}$**), 7.61 (1H, br d, *J* = 8.0 Hz, **H$^{4"}$**), 7.42 (1H, t, *J* = 8.0 Hz, **H$^{5"}$**), 7.40 (2H, d, *J* = 7.0 Hz, **H^{14}**), 7.32 (2H, t, *J* = 7.0 Hz, **H^{15}**), 7.29 (2H, d, *J* = 8.5 Hz, **H^{10}**), 7.28 (2H, d, *J* = 9.0 Hz, **H^{10}**), 7.25 (1H, t, *J* = 7.0 Hz, **H^{16}**), 7.16 (1H, br d, *J* = 8.0 Hz, **H$^{6"}$**), 6.90 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.89 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.26 (1H, dd, *J* = 5.3, 6.3 Hz, **H$^{1'}$**), 5.62 (1H, s, **H^{5}**), 5.39 (1H, d, *J* = 5.0 Hz, 3'-O**H**), 4.43 (1H, td, *J* = 5.5, 11.0 Hz, **H$^{3'}$**), 4.01 (1H, dd, *J* = 4.0, 8.0 Hz, **H$^{4'}$**), 3.70 (3H, s, OC**H$_3$**), 3.69 (3H, s, OC**H$_3$**), 3.47 (3H, s, NC**H$_3$**), 3.40 (1H, dd, *J* = 4.0, 11.0 Hz, **H$^{5'}$**), 3.30 (1H, dd, *J* = 3.0, 11.0 Hz, **H$^{5'}$**), 2.45 (1H, td, *J* = 6.5, 13.6 Hz, **H$^{2'}$**), 2.19 (1H, ddd, *J* = 5.0, 6.5, 13.6 Hz, **H$^{2'}$**), 2.08 (3H, s, COC**H$_3$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 168.5 (**C**OCH_3), 158.5 (**C^{7a}**), 158.1 (**C^{12}**), 153.4 (**C^{2}**), 144.5 (**C^{13}**), 142.2 (**C^{6}**), 139.6 (**C$^{3"}$**), 136.4 (**C^{4}**), 135.4, 135.1 (**C^{9}**), 130.9 (**C$^{1"}$**), 129.8, 129.7 (**C^{10}**), 129.2 (**C$^{5"}$**), 127.9 (**C^{15}**), 127.7 (**C^{14}**), 126.8 (**C^{16}**), 122.8 (**C$^{6"}$**), 119.2 (**C$^{4"}$**), 118.5 (**C$^{2"}$**), 113.3 (**C^{11}**), 107.5 (**C^{4a}**), 98.9 (**C^{5}**), 86.7 (**C$^{1'}$**), 86.1 (**C^{8}**Ar_3), 85.7 (**C$^{4'}$**), 69.0 (**C$^{3'}$**), 62.6 (**C$^{5'}$**), 55.0 (O**C**H_3), 41.5 (**C$^{2'}$**), 29.3 (N**C**H_3), 24.0 (CO**C**H_3); **LRMS** (ES^+) *m*/*z*: 803 ($[M + Et_3N + H]^+$, 14), 723 ($[M + Na]^+$, 100); (ES^-) *m*/*z*: 699 ($[M - H]^-$, 100); **HRMS** (ES^+): calcd for $C_{41}H_{40}N_4O_7$ (M),$[M + H]^+$ = 701.2970, found 701.2965. Found also DMT^+ (calcd 303.1380, found 303.1379).

5'-*O*-(4,4'-Dimethoxytrityl)-5-[(3-aminophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine, 194

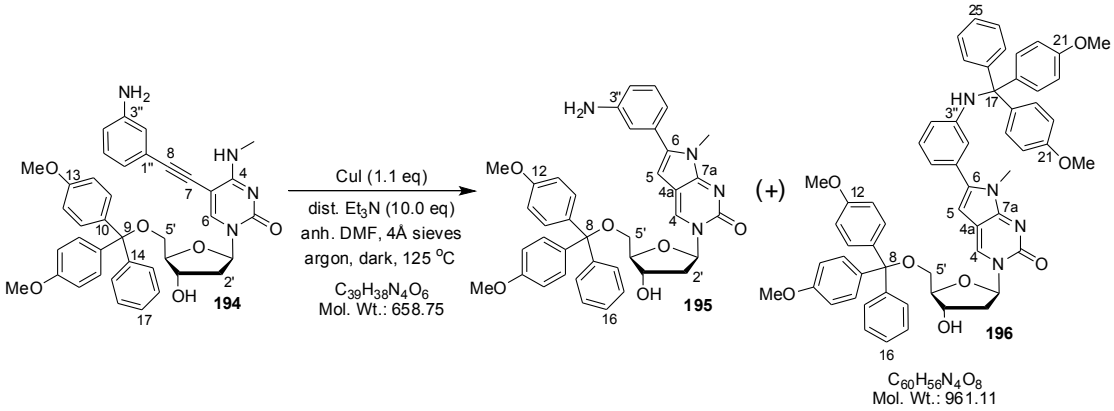


To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine **185** (0.97 g, 1.45 mmol) in anhydrous DMF (4.0 mL), under an argon atmosphere and in absence of light, were added distilled Et_3N (1.0 mL, 7.25 mmol, 5.0 eq) then CuI (0.11 g, 0.58 mmol, 0.4 eq) and alkyne **152** (0.49 mL, 4.35 mmol, 3.0 eq). The reaction mixture was stirred for 20 minutes then *tetrakis*(triphenylphosphine) palladium (0) (0.17 g, 0.15 mmol, 0.1 eq) was added and the reaction was stirred at rt for 1½ hours. The reaction mixture was concentrated under high vacuum, co-evaporating with methanol/toluene/acetone to give an orange/brown foam. The residue was dissolved in methanol/acetone (1:1 v/v, 40 mL) and filtered. The solid was washed with methanol/acetone (1:1 v/v, 2 × 20 mL), and the combined filtrate was concentrated *in vacuo* and dried under high vacuum to give an orange/brown foam. Following purification by column chromatography (0-3% methanol/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the product **194** was afforded, as a pale orange foam (0.91 g, 1.39 mmol, 95%).

R_f 0.31 (1:9, methanol/CH_2Cl_2 + 0.3% Et_3N); 0.31 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'), 0.03 (1:4, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **IR** (solid): ν 3337 (br m, O–H, N–H), 2934, 2835 (w, C–H), 1640 (s, C=O, C=C), 1596 (s, Aryl C–C), 1556 (s), 1505 (vs, C=N), 1444 (m, C–H), 1403 (m), 1352 (m, C–H), 1285 (m, O–H), 1246, 1174 (s), 1086 (br s, C–O), 1029 (br s), 942 (m), 866 (m, C–H), 826 (s, Aryl C–H), 780 (s), 726 (m), 686 (Aryl-C–H), 583, 537, 518 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 7.94 (1H, s, **H^6**), 7.40 (1H, dd, *J* = 1.1, 8.4 Hz, **H^{15}**), 7.35 (1H, q, *J* = 4.8 Hz, N**H**CH_3), 7.30-7.25 (2H, m, **H^{16}**), 7.29 (2H, d, *J* = 8.0 Hz, **H^{11}**), 7.29 (2H, d, *J* = 9.0 Hz, **H^{11}**), 7.17 (1H, tt, *J* = 1.0, 7.3 Hz, **H^{17}**), 6.96 (1H, t, *J* = 7.8 Hz, **H$^{5"}$**), 6.86 (2H, d, *J* = 8.5 Hz, **H^{12}**), 6.83 (2H, d, *J* = 8.8 Hz, **H^{12}**), 6.61 (1H, t, *J* = 1.6 Hz, **H$^{2"}$**), 6.57 (1H, ddd, *J* = 1.1, 2.4, 8.0 Hz, **H$^{4"}$**), 6.46 (1H, ddd, *J* = 1.1, 2.5, 7.5 Hz, **H$^{6"}$**), 6.15 (1H, app. t, *J* = 6.7 Hz, **H$^{1'}$**), 5.30 (1H, d, *J* = 4.3 Hz, 3'-O**H**), 5.11 (1H, br s, N**H$_2$**), 4.25 (1H, td, *J* = 3.1, 9.0 Hz, **H$^{3'}$**), 3.97 (1H, td, *J* = 2.4, 4.8 Hz, **H$^{4'}$**), 3.68 (3H, s, OC**H$_3$**), 3.67 (3H, s, OC**H$_3$**), 3.23 (1H, dd, *J* = 5.0, 10.5 Hz, **H$^{5'}$**), 3.14 (1H, dd, *J* = 2.5, 10.3 Hz, **H$^{5'}$**), 2.86 (3H, d, *J* = 4.8 Hz, NHC**H$_3$**), 2.27 (1H, ddd, *J* = 3.0, 5.9, 13.4 Hz, **H$^{2'}$**), 2.13 (1H, td, *J* = 6.8, 13.6 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 161.7 (**C^4**–$NHCH_3$), 158.0, 158.0 (**C^{13}**–OCH_3), 153.2 (**C^2**), 148.3 (**C^{14}**), 144.6 (**C$^{3"}$**), 142.5 (**C^6**), 135.6, 135.2 (**C^{10}**), 129.7, 129.6 (**C^{11}**), 128.7 (**C$^{5"}$**), 127.8 (**C^{16}**), 127.6 (**C^{15}**), 126.6 (**C^{17}**), 122.4 (**C$^{1"}$**), 119.2 (**C$^{6"}$**), 116.4 (**C$^{2"}$**), 114.5 (**C$^{4"}$**), 113.2, 113.2 (**C^{12}**), 95.3 (**C^8**), 90.6 (**C^7**), 85.9 (**C$^{4'}$**), 85.8 (**C^9**Ar_3), 85.6 (**C$^{1'}$**), 79.0 (**C^5**), 70.8 (**C$^{3'}$**), 63.7 (**C$^{5'}$**), 54.9 (O**C**H_3), 40.8 (**C$^{2'}$**), 27.9 (NH**C**H_3); **LRMS** (ES^+) *m*/*z*:

681 ([M + Na]$^+$, 100); **HRMS** (ES$^+$): calcd for $C_{41}H_{40}N_4O_7$ (M), [M + Na]$^+$ = 681.2684, found 681.2671; [M + H]$^+$ = 659.2864, found 659.2847. Found also DMT$^+$ (calcd 303.1380, found 303.1381).

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-ribofuranosyl]-6-(3-aminophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 195



Method 1: To a stirred solution of alkynylated product **194** (2.21 g, 3.36 mmol) in anhydrous DMF (9.0 mL) and distilled Et_3N (4.70 mL, 33.7 mmol, 10.0 eq), under an argon atmosphere, in absence of light, were added 4Å molecular sieves then CuI (0.70 g, 3.70 mmol, 1.1 eq), and the reaction was stirred at 125 °C for 50 minutes. The mixture was allowed to cool, filtered and the filtrate was concentrated under high vacuum, co-evaporating with methanol/toluene/acetone to give a dark brown/gold foam. The foam was dissolved in DMF (80 mL), filtered, the solid was washed with DMF (3 × 60 mL) and the combined filtrate was concentrated under high vacuum at 35 °C to give a sticky, dark brown foam/gum. The mixture was finally suspended/dissolved in CH_2Cl_2 (100 mL), soaked overnight and filtered. The solid residue was washed (CH_2Cl_2, 3 × 50 mL), and the combined filtrate was concentrated *in vacuo* to give a sticky orange/brown foam. Following purification by column chromatography (50-100% acetone/CH_2Cl_2 + 0.5% Et_3N → 5% methanol/acetone + 0.5% Et_3N), and drying under high vacuum, the desired product **195** and *bis*-tritylated nucleoside **196** were isolated, as pale orange foams (**195**, 0.71 g, 1.07 mmol, 32%; **196**, 0.38 g, 0.39 mmol, 12%).

Method 2: Cyclised nucleoside **195** was prepared using nucleoside **194** (0.88 g, 1.34 mmol), by the same procedure as in Method 1, with a reaction time of 30 minutes. The mixture was allowed to cool, filtered, and the filtrate was concentrated under high vacuum to give a viscous brown syrup. The residue was suspended/dissolved in CH_2Cl_2 (50 mL) and filtered. The solid residue was washed (CH_2Cl_2, 2 × 25 mL) and the combined filtrate was concentrated *in vacuo* and under high vacuum to give a brown foam. Following purification by column chromatography (50-100% acetone/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired cyclised product **195** was afforded, as a pale brown foam (0.71 g, 1.07 mmol, 80%).

Product 195:

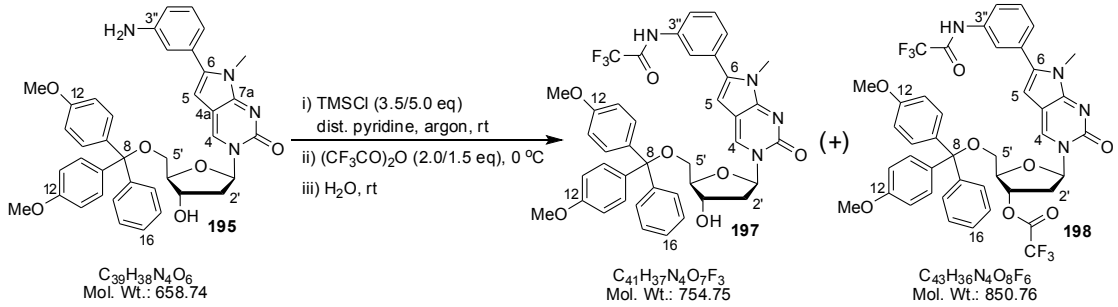
$\mathbf{R}_f$ 0.08 (5:95, methanol/CH_2Cl_2 + 0.3% Et_3N, A'), 0.14 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **IR** (solid): ν 3336 (w, O–H, N–H), 3064, 2930, 2835 (w, C–H), 1699 (w), 1651 (s, C=O), 1605 (s, C=C), 1553 (s, C=N), 1507 (s, Aryl C–C), 1476 (s, N–H), 1445, 1398 (s), 1338 (m), 1301 (s), 1246 (vs), 1174 (s), 1095 (s, C–O), 1030 (vs), 969, 914 (m), 827 (s, C–H), 772, 726 (s), 701 (s, Aryl C–H), 659, 584, 530 (s) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 8.63 (1H, s, **H^{4}**), 7.41 (2H, d, *J* = 7.3 Hz, **H^{14}**), 7.32 (2H, t, *J* = 7.0 Hz, **H^{15}**), 7.29 (4H, d, *J* = 8.8 Hz, **H^{10}**), 7.25 (1H, tt, *J* = 1.8, 7.3 Hz, **H^{16}**), 7.13 (1H, t, *J* = 7.8 Hz, **H$^{5"}$**), 6.90 (2H, d, *J* = 8.8 Hz, **H^{11}**), 6.89 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.67 (1H, t, *J* = 1.8 Hz, **H$^{2"}$**), 6.64 (1H, ddd, *J* = 0.8, 2.2, 8.0 Hz, **H$^{4"}$**), 6.60 (1H, td, *J* = 1.3, 8.0 Hz, **H$^{6"}$**), 6.26 (1H, dd, *J* = 5.0, 6.3 Hz, **H$^{1'}$**), 5.51 (1H, s, **H^{5}**), 5.42 (1H, d, *J* = 4.8 Hz, 3'-O**H**), 5.26 (2H, br s, N**H$_2$**), 4.44 (1H, td, *J* = 5.3, 10.5 Hz, **H$^{3'}$**), 4.00 (1H, dd, *J* = 3.9, 7.9 Hz, **H$^{4'}$**), 3.71 (3H, s, OC**H$_3$**), 3.70 (3H, s, OC**H$_3$**), 3.44 (3H, s, NC**H$_3$**), 3.39 (1H, dd, *J* = 3.9, 10.7 Hz, **H$^{5'}$**), 3.30 (1H, d, *J* = 2.6, 10.9 Hz, **H$^{5'}$**), 2.44 (1H, td, *J* = 6.5, 13.1 Hz, **H$^{2'}$**), 2.19 (1H, ddd, *J* = 5.1, 6.4, 13.4 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 158.5 (**C^{7a}**), 158.2 (**C^{12}**–OCH_3), 153.5 (**C^{2}**), 149.0 (**C$^{3"}$**–NHR), 144.5 (**C^{13}**), 143.4 (**C^{6}**), 135.9 (**C^{4}**), 135.5, 135.2 (**C^{9}**), 131.1 (**C$^{1"}$**), 129.8, 129.7 (**C^{10}**), 129.3 (**C$^{5"}$**), 128.0 (**C^{15}**), 127.8 (**C^{14}**), 126.9 (**C^{16}**), 115.6 (**C$^{6"}$**), 114.3 (**C$^{4"}$**), 113.3 (**C^{11}**,**C$^{2"}$**), 107.7 (**C^{4a}**), 98.0 (**C^{5}**), 86.6 (**C$^{1'}$**), 86.1 (**C^{8}**Ar_3), 85.6 (**C$^{4'}$**), 69.0 (**C$^{3'}$**), 62.6 (**C$^{5'}$**), 55.0 (O**C**H_3), 41.5 (**C$^{2'}$**), 29.4 (N**C**H_3); **LRMS** (ES^+) *m*/*z*: 681 ($[M + Na]^+$, 100); **HRMS** (ES^+): calcd for $C_{39}H_{38}N_4O_6$ (M), $[M + Na]^+$ = 681.2684, found 681.2677. Found also DMT^+ (calcd 303.1380, found 303.1381); **Anal.** Calcd $C_{39}H_{38}N_4O_6$ (658.75), C 71.11%, H 5.81%, N 8.50%, found C 69.90%, H 6.11%, N 8.41%.

Analytical results are consistent with the reported data.[206]

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-[3-(4,4'-dimethoxytrityl)aminophenyl]-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 196:

R$_f$ 0.33 (4:1, acetone/CH_2Cl_2 + 0.3% Et_3N, A'); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 8.57 (1H, s, **H^{4}**), 7.40 (2H, d, *J* = 7.3 Hz, **H^{14}**), 7.33-7.26 (4H, m, **H^{15}**,**H^{23}**), 7.28 (4H, d, *J* = 8.8 Hz, **H^{10}**), 7.26-7.19 (3H, m, **H^{16}**,**H^{24}**), 7.22 (4H, d, *J* = 8.8 Hz, **H^{19}**), 7.16 (1H, tt, *J* = 1.6, 7.0 Hz, **H^{25}**), 6.97 (1H, t, *J* = 7.9 Hz, **H$^{5"}$**), 6.89 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.88 (2H, d, *J* = 9.5 Hz, **H^{11}**), 6.85 (4H, d, *J* = 9.0 Hz, **H^{20}**), 6.64 (1H, br dd, *J* = 1.4, 8.2 Hz, **H$^{6"}$**), 6.54 (1H, d, *J* = 7.8 Hz, **H$^{4"}$**), 6.51 (1H, t, *J* = 1.8 Hz, **H$^{2"}$**), 6.22 (1H, dd, *J* = 5.3, 6.3 Hz, **H$^{1'}$**), 5.39 (1H, d, *J* = 4.5 Hz, 3'-O**H**), 5.38 (1H, s, N**H**DMT), 4.40 (1H, qn, *J* = 5.5 Hz, **H$^{3'}$**), 3.98 (1H, dd, *J* = 3.8, 8.0 Hz, **H$^{4'}$**), 3.69 (6H, s, C^{21}–O**CH_3**), 3.69 (3H, s, C^{12}–O**CH_3**), 3.68 (3H, s, C^{12}–O**CH_3**), 3.36 (1H, dd, *J* = 3.8, 10.4 Hz, **H$^{5'}$**), 3.28 (1H, d, *J* = 2.8, 10.5 Hz, **H$^{5'}$**), 3.12 (3H, s, N**CH_3**), 2.41 (1H, td, *J* = 6.4, 12.8 Hz, **H$^{2'}$**), 2.15 (1H, ddd, *J* = 5.3, 6.3, 13.1 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 158.3 (**C^{7a}**), 158.2, 158.1 (**C^{12}**–OCH_3), 157.6 (**C^{21}**–OCH_3), 153.4 (**C^{2}**), 147.2 (**C$^{3"}$**–NHR), 145.7 (**C^{22}**), 144.4 (**C^{13}**), 143.2 (**C^{6}**), 137.2 (**C^{18}**), 135.9 (**C^{4}**), 135.6, 135.2 (**C^{9}**), 129.9 (**C^{19}**,**C$^{1"}$**), 129.8, 129.6 (**C^{10}**), 128.7 (**C^{23}**), 128.2 (**C$^{5"}$**), 127.9 (**C^{15}**), 127.8 (**C^{14}**,**C^{24}**), 126.8 (**C^{16}**), 126.4 (**C^{25}**), 116.8 (**C$^{6"}$**), 116.4 (**C$^{4"}$**), 115.3 (**C$^{2"}$**), 113.3, 113.1 (**C^{11}**,**C^{20}**), 107.5 (**C^{4a}**), 98.0 (**C^{5}**), 86.6 (**C$^{1'}$**), 86.0 (**C^{8}**Ar_3), 85.6 (**C$^{4'}$**), 69.5 (**C^{17}**Ar_3), 69.0 (**C$^{3'}$**), 62.6 (**C$^{5'}$**), 55.0 (OCH_3), 41.5 (**C$^{2'}$**), 29.1 (NCH_3); **LRMS** (ES^+) *m*/*z*: 1063 ([M + Et_3N + H]$^+$, 16), 983 ([M + Na]$^+$, 59), 961 ([M + H]$^+$, 27), 303 (DMT^+, 100).

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-trifluoroacetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 197



$C_{39}H_{38}N_4O_6$
Mol. Wt.: 658.74

$C_{41}H_{37}N_4O_7F_3$
Mol. Wt.: 754.75

$C_{43}H_{36}N_4O_8F_6$
Mol. Wt.: 850.76

Method 1: To a stirred solution of nucleoside **195** (0.52 g, 0.79 mmol) in distilled pyridine (4.0 mL), under an argon atmosphere and excluding moisture, at rt, was added TMSCl (0.15 mL, 1.5 eq) slowly. After stirring for 2 hours, further TMSCl (0.20 mL, 2.0 eq) was added, and the reaction was stirred for a further 30 minutes. The reaction mixture was cooled to 0 °C, then trifluoroacetic anhydride (0.22 mL, 2.0 eq) was added dropwise. The reaction was stirred at 0 °C for 10 minutes, then allowed to warm to rt for 25 minutes. Water (1.0 mL) was added and the pale yellow solution was stirred for 25 minutes. The reaction mixture was reduced to approximately one-third volume, then was diluted with CH_2Cl_2 (150 mL), adding Et_3N (1.5 mL) to prevent detritylation. The solution was washed with water (60 mL), the aqueous was re-extracted (CH_2Cl_2, 2 × 50 mL), and combined CH_2Cl_2 fractions were washed with sat. aq KCl (140 mL), dried (Na_2SO_4) and concentrated *in vacuo* and dried under high vacuum to give a yellow gum. The compound was twice-purified by column chromatography (50-80% acetone/ CH_2Cl_2 + 0.5% Et_3N; 10-60% acetone/CH_2Cl_2 + 0.5% Et_3N → 5% methanol/acetone + 0.5% Et_3N), from which the product **197** was isolated, after drying under high vacuum, as a pale orange foam (0.19 g, 0.25 mmol, 32%). A mixture containing product **197** and *bis*-trifluoroacetylated compound **198** (*ca.* 11:2) was also isolated (0.31 g).

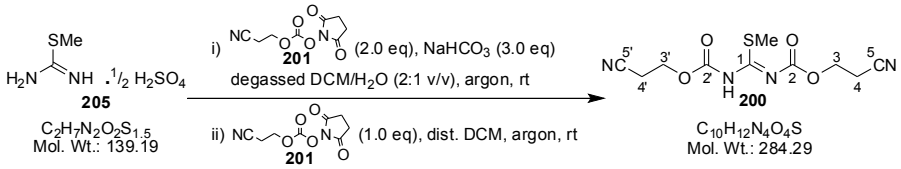
Method 2: Product **197** was prepared using nucleoside **195** (0.51 g, 0.78 mmol) as described in Method 1, using 5 eq of TMSCl, with a total reaction time of 3½ hours. After addition of trifluoroacetic anhydride, white fumes were observed, hence Et_3N (2.0 mL) was added, before water, to prevent detritylation. Water (2.0 mL) was then added and the orange/brown solution was stirred for 2 hours at rt. The reaction mixture was diluted with CH_2Cl_2 (300 mL) then washed with water (300 mL). The aqueous was re-extracted (CH_2Cl_2, 2 × 50 mL), and the combined CH_2Cl_2 fraction was washed with 15% aq NH_3 (100 mL) and sat. aq KCl (140 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give a brown gum. Following purification by column chromatography (5-80% acetone/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **197** was afforded, as a pale orange foam (0.31 g, 0.41 mmol, 53%).

$\mathbf{R}_f$ 0.31 (**197**), 0.41 (**198**) (7:3, acetone/CH_2Cl_2 + 0.5% Et_3N, A'), 0.08 (**197**) (1:4, acetone/CH_2Cl_2 + 0.5% Et_3N, A'); **^{19}F NMR** (282 MHz, d_6-DMSO) δ -73.6 (NHCOC**F$_3$**); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 11.36 (N**H**COCF$_3$), 8.68 (1H, s, **H^4**), 7.83 (1H, br s, **H$^{2"}$**), 7.75 (1H, td, J = 0.9, 8.3 Hz, **H$^{4"}$**), 7.55 (1H, t, J = 8.0 Hz, **H$^{5"}$**), 7.40 (2H, d, J = 8.0 Hz, **H^{14}**), 7.37 (1H, d, J = 8.2 Hz, **H$^{6"}$**), 7.32 (2H, t, J = 7.3 Hz, **H^{15}**),

7.29 (4H, d, J = 8.3 Hz, **H**^{10}), 7.28 (4H, d, J = 9.0 Hz, **H**^{10}), 7.24 (1H, t, J = 7.5 Hz, **H**^{16}), 6.89 (2H, d, J = 8.5 Hz, **H**^{11}), 6.88 (2H, d, J = 9.0 Hz, **H**^{11}), 6.25 (1H, app. t, J = 5.7 Hz, **H**$^{1'}$), 5.65 (1H, s, **H**^{5}), 5.42 (1H, d, J = 4.3 Hz, 3'-O**H**), 4.42 (1H, qn, J = 5.0 Hz, **H**$^{3'}$), 4.00 (1H, dd, J = 3.9, 7.4 Hz, **H**$^{4'}$), 3.69 (3H, s, OC**H**$_3$), 3.69 (3H, s, OC**H**$_3$), 3.49 (3H, s, NC**H**$_3$), 3.39 (1H, d, J = 7.3 Hz, **H**$^{5'}$), 3.35 (1H, d, J = 7.3 Hz, **H**$^{5'}$), 2.45 (1H, td, J = 6.5, 13.1 Hz, **H**$^{2'}$), 2.19 (1H, ddd, J = 5.5, 6.1, 13.2 Hz, **H**$^{2'}$); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 158.6 (**C**^{7a}), 158.2 (**C**^{12}–OCH_3), 154.7 (q, $^2J_{CF}$ = 36.9 Hz, **C**OCF_3), 153.4 (**C**^{2}), 144.5 (**C**^{13}), 141.6 (**C**^{6}), 136.7 (**C**^{4},**C**$^{3''}$), 135.5, 135.2 (**C**^{9}), 131.3 (**C**$^{1''}$), 129.8, 129.7 (**C**^{10}), 129.6 (**C**$^{5''}$), 128.0 (**C**^{15}), 127.8 (**C**^{14}), 126.8 (**C**^{16}), 125.4 (**C**$^{6''}$), 121.3 (**C**$^{4''}$), 120.6 (**C**$^{2''}$), 115.7 (q, $^1J_{CF}$ = 286.7 Hz, CO**C**F_3), 113.3 (**C**^{11}), 107.5 (**C**^{4a}), 99.3 (**C**^{5}), 86.7 (**C**$^{1'}$), 86.1 (**C**^{8}Ar_3), 85.7 (**C**$^{4'}$), 68.9 (**C**$^{3'}$), 62.6 (**C**$^{5'}$), 55.0 (O**C**H_3), 41.5 (**C**$^{2'}$), 29.4 (N**C**H_3); **LRMS** (ES^+) *m/z*: 804 (16), 793 ($[M + K]^+$, 15), 777 ($[M + Na]^+$, 99), 303 (DMT^+, 100); **HRMS** (ES^+): calcd for $C_{41}H_{37}N_4O_7F_3$ (M), $[M + Na]^+$ = 777.2507, found 777.2509. Found also DMT^+ (calcd 303.1380, found 303.1384).

Analytical results are consistent with the reported data.[206]

N,N'-Bis-[(2-cyanoethoxy)carbonyl]-*S*-methyl-isothiourea, 200[41,225,226]

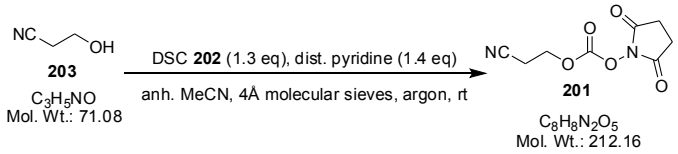


A mixture of *S*-methylisothiourea hemisulfate **205** (6.03 g, 43.3 mmol) and $NaHCO_3$ (10.9 g, 130 mmol, 3.0 eq) in argon-degassed, distilled water (50.0 mL) was stirred, under an argon atmosphere, at rt, for 30 minutes. To this was added degassed CH_2Cl_2 (150.0 mL) followed by (2-cyanoethyl)-*N*-succinimidyl carbonate (CEOC-succinimide) **201** (18.9 g, 86.6 mmol, 2.0 eq) and the mixture was stirred vigorously, under an argon atmosphere, at rt for 8 hours. The mixture was diluted with CH_2Cl_2 (100 mL) and water (100 mL), agitated and separated. The aqueous layer was re-extracted (CH_2Cl_2, 3 × 100 mL) and combined CH_2Cl_2 fractions were dried (Na_2SO_4) and concentrated *in vacuo* and under high vacuum to give a viscous, cloudy yellow syrup.

The syrup was redissolved in distilled CH_2Cl_2 (65.0 mL), to which CEOC-succinimide **201** (9.47 g, 43.3 mmol, 1.0 eq) was added and the mixture was stirred under an argon atmosphere, at rt for 14½ hours. The yellow solution was diluted with CH_2Cl_2 (135 mL) then washed with sat. aq KCl (80 mL). The aqueous fraction was re-extracted (CH_2Cl_2, 2 × 30 mL) and the combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo* to give a cloudy, yellow syrup. Following purification by column chromatography (5:95, ethyl acetate/CH_2Cl_2), and drying under high vacuum, the desired product **200** was afforded as a hard, powdery white solid (11.0 g, 38.7 mmol, 89%). A mixture of reagent **201** and product **200** (*ca.* 7:3) (0.90 g) was also isolated.

R_f 0.46 (1:3, ethyl acetate/CH_2Cl_2, B', F'), 0.28 (1:9, ethyl acetate/CH_2Cl_2, B'); **^{1}H NMR** (400 MHz, $CDCl_3$) δ 11.77 (1H, s, N**H**), 4.40 (2H, t, *J* = 6.3 Hz, **H^3**), 4.36 (2H, t, *J* = 6.0 Hz, **H$^{3'}$**), 2.78 (4H, t, *J* = 6.5 Hz, C**H$_2$**CN), 2.44 (3H, s, SC**H$_3$**); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 173.7 (NH**C^1**=N), 160.4 (N**C^2**=O), 150.8 (NH**C$^{2'}$**=O), 116.8 (**C^5**N), 116.3 (**C$^{5'}$**N), 61.0 (**C^3**), 60.6 (**C$^{3'}$**), 18.1 (**C**H$_2$CN), 14.8 (**S**CH$_3$); **LRMS** (ES$^+$) *m*/*z*: 307 ([M + Na]$^+$, 100), 285 ([M + H]$^+$, 12), 277 (11), 236 (17).
Analytical results are consistent with the reported data.[225]

(2-Cyanoethyl)-*N*-succinimidyl carbonate, 201[42,226]



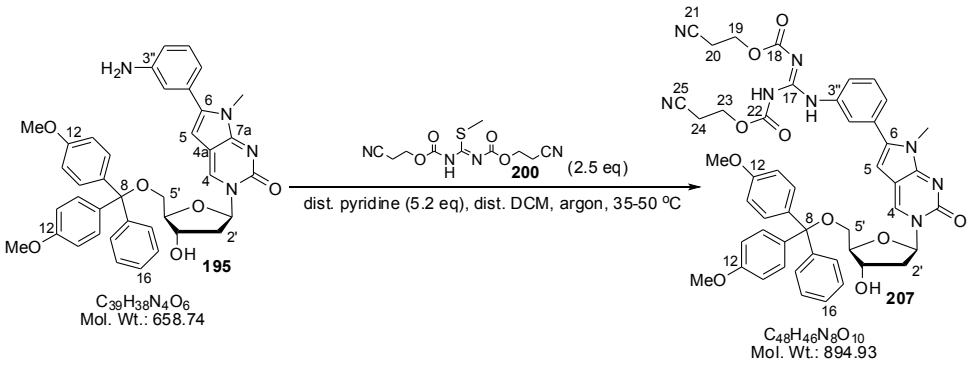
A suspension of di-succinimidyl carbonate (DSC) **202** (69.3 g, 271 mmol, 1.3 eq) in anhydrous acetonitrile (300 mL), with 4Å molecular sieves, was stirred vigorously at rt, under an argon atmosphere for 35 minutes. Distilled pyridine (23.6 mL, 291 mmol, 1.4 eq) was added and the mixture was stirred for 15 minutes. 3-Hydroxypropionitrile (3-HPN) **203** (14.2 mL, 208 mmol) was added dropwise over 35 minutes then the mixture was stirred, at rt, for 1½ hours. The pale yellow mixture was filtered, the residue was washed with CH_3CN (2 × 30 mL) and the combined filtrate was concentrated *in vacuo* and under high vacuum to give a viscous yellow syrup. The syrup was dissolved in argon-degassed CH_2Cl_2 (500 mL) and washed with degassed sat. aq $NaHCO_3$ (2 × 250 mL). The combined aqueous fraction was re-extracted (degassed CH_2Cl_2, 3 × 50 mL)

and combined CH_2Cl_2 fractions were washed with degassed sat. aq KCl (2 × 200 mL), dried (Na_2SO_4) and concentrated *in vacuo* and under high vacuum to give a white waxy solid. The solid was co-evaporated with anhydrous CH_3CN (2 × 50 mL) and dried under high vacuum, then triturated with diethyl ether (2 × 150 mL) and dried under high vacuum, to afford the desired product **201** as a slightly waxy, white solid (42.5 g) containing 5.1 mol % *N*-hydroxysuccinimide (recalculated yield - 195 mmol, 94%).

R_f 0.33 (1:3, ethyl acetate/CH_2Cl_2, B', F'), 0.17 (1:9, ethyl acetate/CH_2Cl_2, B', F'); **IR** (solid): ν 3001, 2968, 2941 (w, C–H), 2262 (w, C≡N), 1810, 1781 (s, C=O), 1727 (br vs, carbonate C=O), 1474, 1430 (w), 1364 (m), 1263, 1231 (s, C–O), 1196 (vs, C–O), 1076 (s, C–O), 992, 968 (s), 861, 813 (m), 763, 718 (s), 642 (s, N–O), 575, 460 (m) cm^{-1}; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.52 (2H, t, *J* = 6.5 Hz, C**H_2**O), 2.85 (4H, s, succinimide-C**H_2**), 2.84 (2H, t, *J* = 6.5 Hz, C**H_2**CN); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 168.4 (succinimide-N**C**O), 151.3 (O**C**O_2), 115.6 (**C**N), 64.8 (**C**H_2O), 25.6 (succinimide-**C**H_2), 18.0 (**C**H_2CN); **LRMS** (ES^+) *m*/*z*: 276 ([M + CH_3CN + Na]$^+$, 30), 235 ([M + Na]$^+$, 100).

Analytical results are consistent with the reported data.[42]

3-*N*-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-{*N*,*N'*-*bis*-[(2-cyanoethoxy)carbonyl]guanidinyl}phenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one, 207



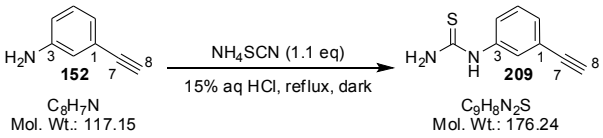
To a stirred solution of nucleoside **195** (0.39 g, 0.60 mmol) and guanidinylating reagent **200** (0.19 g, 0.66 mmol, 1.1 eq) in distilled CH_2Cl_2 (4 mL), under an argon atmosphere, was added distilled pyridine (24.1 μL, 0.30 mmol, 0.5 eq) and the solution was stirred at 35 °C. Further distilled pyridine (0.05 mL, 0.62 mmol, 1.0 eq) was added at 16½ hours

and 19 hours, when distilled CH_2Cl_2 (1.0 mL) was added and the temperature was raised to 50 °C. Molecular sieves (4Å) and distilled pyridine (0.05 mL, 0.62 mmol, 1.0 eq) were added at 20½ hours, distilled CH_2Cl_2 (2.0 mL) at 22½ hours, and further distilled pyridine (0.08 mL, 0.93 mmol, 1.6 eq) and guanidinylating reagent **200** (0.24 g, 0.84 mmol, 1.4 eq) at 23½ hours. After a further 17 hours, the reaction mixture was diluted with CH_2Cl_2 (100 mL), filtered and washed with water (50 mL). The aqueous layer was re-extracted (CH_2Cl_2, 2 × 30 mL) and combined organic fractions were washed with sat. aq KCl (75 mL), dried (Na_2SO_4), and concentrated *in vacuo* and under high vacuum to give a golden brown foam. Following purification by column chromatography (50-100% acetone/CH_2Cl_2), on silica gel pre-equilibrated with pyridine, the desired guanidinylated product **207** was afforded, as a yellow foam (0.23 g, 0.25 mmol, 42%). A mixture of product **207** and starting material **195** (*ca.* 1:2) (0.03 g) was also isolated.

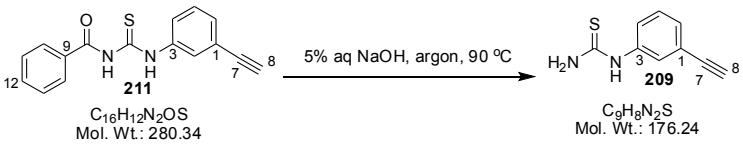
R_f 0.37 (4:1, acetone/CH_2Cl_2, A', E'); **IR** (solid): ν 3271 (br w, O–H, N–H), 2933, 2837 (w, C–H), 2255 (w, C≡N), 1732 (m, C=O), 1650 (s, C=O, C=N), 1622 (s, C=C), 1607, 1555 (s, C=O), 1507 (s, Aryl C–C), 1424, 1398, 1282, 1229, 1175 (s), 1114, 1083 (s, C–O), 1057, 1029 (s), 907 (m), 827 (s, Aryl C–H), 791 (s), 770 (s, Aryl C–H), 727 (s, CH_2), 700 (s, Aryl C–H), 658, 583, 531 (s) cm^{-1}; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 11.57 (1H, s, CON**H**CN), 9.91 (1H, s, N**H**Ar), 8.52 (1H, s, **H^{4}**), 7.48 (1H, t, *J* = 1.8 Hz, **H$^{2"}$**), 7.23 (1H, ddd, *J* = 1.0, 2.0, 8.0 Hz, **H$^{4"}$**), 7.18-7.15 (1H, m, **H$^{5"}$**), 7.15 (2H, d, *J* = 7.0 Hz, **H^{14}**), 7.04 (4H, d, *J* = 9.0 Hz, **H^{10}**), 7.04-6.99 (1H, m, **H^{16}**), 6.98 (2H, t, *J* = 9.0 Hz, **H^{15}**), 6.95 (1H, br d, *J* = 8.5 Hz, **H$^{6"}$**), 6.54 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.53 (2H, d, *J* = 9.0 Hz, **H^{11}**), 6.19 (1H, t, *J* = 5.5 Hz, **H$^{1'}$**), 5.29 (1H, s, **H^{5}**), 4.39 (1H, q, *J* = 5.5 Hz, **H$^{3'}$**), 4.19 (2H, br t, *J* = 5.8 Hz, **H^{24}**), 4.00 (2H, br t, *J* = 5.8 Hz, **H^{20}**), 3.90 (1H, td, *J* = 3.0, 5.0 Hz, **H$^{4'}$**), 3.46 (3H, s, OC**H$_3$**), 3.45 (3H, s, OC**H$_3$**), 3.34 (3H, s, NC**H$_3$**), 3.26 (1H, dd, *J* = 3.0, 11.0 Hz, **H$^{5'}$**), 3.21 (1H, dd, *J* = 3.3, 10.8 Hz, **H$^{5'}$**), 2.59-2.52 (1H, m, **H$^{2'}$**), 2.53 (2H, br t, *J* = 6.0 Hz, **H^{23}**), 2.41 (2H, br t, *J* = 5.8 Hz, **H^{19}**), 2.11 (1H, ddd, *J* = 5.5, 6.5, 13.6 Hz, **H$^{2'}$**); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 163.5 (**C^{17}**=N), 159.3 (**C^{7a}**), 159.1, 158.0 (**C^{12}**), 155.3 (N**C^{18}**=O), 154.0 (**C^{2}**), 153.6 (NH**C^{22}**=O), 144.8 (**C^{13}**), 142.7 (**C^{6}**), 136.8 (**C^{4}**), 136.5 (**C$^{3"}$**–NHR), 136.1, 135.9 (**C^{9}**), 132.3 (**C$^{1"}$**), 130.6, 130.5 (**C^{10}**), 129.9 (**C$^{5"}$**), 128.6 (**C^{15}**), 128.5 (**C^{14}**), 127.5 (**C^{16}**), 126.1 (**C$^{6"}$**), 123.3 (**C$^{2"}$**), 123.2 (**C$^{4"}$**), 117.2 (**C^{21}**≡N), 116.4 (**C^{25}**≡N), 113.8 (**C^{11}**), 109.2 (**C^{4a}**), 100.4 (**C^{5}**), 88.2 (**C$^{1'}$**), 87.4 (**C^{8}**Ar$_3$), 86.6 (**C$^{4'}$**), 70.9 (**C$^{3'}$**), 63.0 (**C$^{5'}$**), 61.5 (**C^{19}**), 60.3 (**C^{23}**), 55.6 (O**C**H$_3$), 43.0 (**C$^{2'}$**), 30.4 (NH**C**H$_3$), 18.4 (**C^{20}**,**C^{24}**); **LRMS** (ES$^+$) *m/z*: 918 ([M + Na]$^+$, 100), 895 ([M + H]$^+$, 10); **HRMS**

(ES$^+$): calcd for $C_{48}H_{46}N_8O_{10}$ (M), $[M + Na]^+$ = 917.3229, found 917.3234. Found also DMT^+: calcd 303.1380, found 303.1386.

3-Thioureidophenyl acetylene, 209



Method 1: Ammonium thiocyanate (1.12 g, 14.7 mmol, 1.1 eq) was dissolved, with care, in 15% aq HCl (1.5 mL). This was added to a stirred emulsion of 3-ethynylaniline **152** (1.50 mL, 13.3 mmol) in 15% aq HCl (1.3 mL), at rt. The mixture was refluxed, in absence of light, for 9¼ hours. The brown reaction mixture was allowed to cool, then diluted with ethyl acetate (180 mL) and washed with water (2 × 40 mL) and sat. aq KCl (40 mL). The organic layer (suspension) was removed and concentrated *in vacuo* to give a golden-brown foam/gum. The gum was dissolved/suspended in acetone (100 mL), filtered, the residue was washed with acetone (2 × 20 mL) and the combined filtrate was concentrated *in vacuo*. The resulting gum was suspended in diethyl ether (50 mL) and allowed to soak overnight. The mixture was filtered, washed with diethyl ether (4 × 25 mL) and combined ether fractions were concentrated *in vacuo* and under high vacuum to give a yellow syrup. Following purification by column chromatography (0-8% methanol/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **209** was afforded as an orange solid (0.14 g, 0.79 mmol, 6%). Starting material **152** was also recovered (0.41 g, 3.53 mmol, 27%).

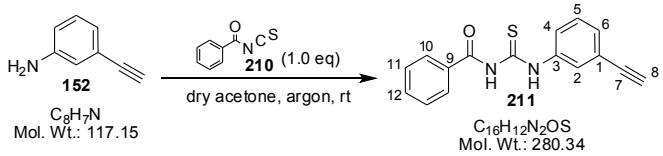


Method 2: Benzoylthiourea **211** (2.54 g, 9.06 mmol) was added, in one portion, to stirred 5% aq NaOH (45.0 mL) under argon, at 90 °C. After stirring for 15 minutes, the suspension was filtered hot to remove the benzamide byproduct **212**, then cooled and acidified to pH 5.0 using 15% aq HCl. The mixture was then basified to pH 8.0 using

conc. aq NH_3 and filtered. The solid was washed with water (3 × 10 mL) then dried under high vacuum over P_2O_5 for 2 days, to afford the desired deprotected thiourea **209** as a pale, cream-coloured powdery solid (1.26 g, 7.15 mmol, 79%).

R_f 0.17 (5:95, acetone/CH_2Cl_2, A'); **Mp** 145-148 °C (CH_2Cl_2); **IR** (solid): ν 3418, 3394 (m, N–H), 3208 (m), 3160 (br s, water), 3015, 2981 (m, Aryl C–H), 2680 (w), 1618 (s, N–H), 1596 (m), 1580 (m, Aryl C–C), 1526, 1514, 1480, 1468 (s), 1423, 1411 (m), 1297, 1248 (s, N–CS), 1155 (m), 1065 (s, C=S), 999, 888, 831 (m), 786, 711, 688, 667, 632, 622, 579 (s), 518 (br vs) cm^{-1}; **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.75 (1H, s, N**H**), 7.62 (1H, t, *J* = 1.8 Hz, **H^2**), 7.54 (2H, extremely br s, **NH_2**), 7.42 (1H, ddd, *J* = 1.0, 2.0, 8.0 Hz, **H^4**), 7.32 (1H, t, *J* = 8.0 Hz, **H^5**), 7.07 (1H, td, *J* = 1.3, 7.7 Hz, **H^6**), 4.17 (1H, s, C≡C**H^8**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 181.2 (**C**=S), 139.5 (**C^3**), 129.0 (**C^5**), 127.4 (**C^6**), 125.6 (**C^2**), 123.5 (**C^4**), 121.8 (**C^1**), 83.2 (**C^7**≡CH), 80.8 (C≡**C^8**H); **LRMS** (ES^+) *m*/*z*: 353 ($[2M + H]^+$, 100), 199 ($[M + Na]^+$, 78), 177 ($[M + H]^+$, 73); **HRMS** (ES^+): calcd for $C_9H_8N_2S$ (M), $[3M + Na]^+$ = 551.1117, found 551.1121; $[2M + Na]^+$ = 375.0709, found 375.0710; $[M + Na]^+$ = 199.0300, found 199.0301.

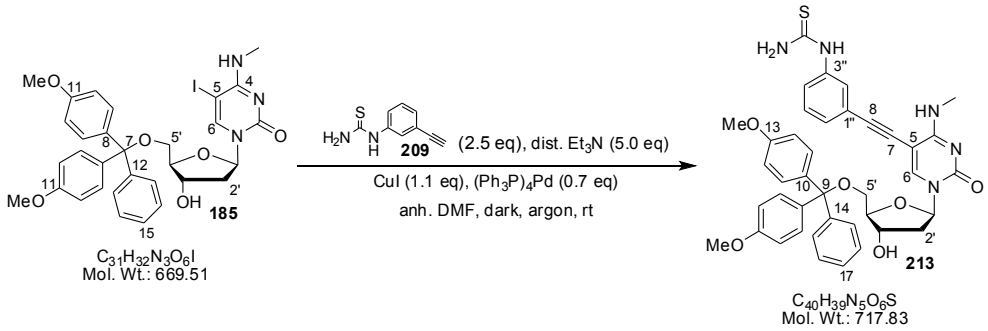
N-Benzoyl-*N'*-(3-ethynylphenyl) thiourea, 211



To a stirred solution of 3-ethynylaniline **152** (1.50 mL, 13.3 mmol) in dry acetone (30.0 mL), under an argon atmosphere, at rt, was added a solution of benzoyl isothiocyanate **210** (1.82 mL, 13.5 mmol, 1.0 eq) in dry acetone (20.0 mL) dropwise over 12 minutes. The solution was stirred at rt for 1 hour, then concentrated *in vacuo* to give a pale yellow papery solid. Cold dry acetone (15 mL) was added and the mixture was cooled at $^-$10 °C for 2 hours. The mixture was filtered cold, then washed with cold, dry acetone (2 × 5 mL) and diethyl ether (2 × 15 mL), then dried under high vacuum over P_2O_5 to afford the desired product **211** as a pale yellow papery solid. Four further crops were obtained from the filtrate. (Combined yield - 2.81 g, 10.0 mmol, 75%).

R_f 0.69 (5:95, acetone/CH_2Cl_2, A'); **Mp** 138-140 °C (acetone); **IR** (solid): ν 3346 (br w, N–H), 3259 (s, alkyne C–H), 2991 (br m, Aryl C–H), 1663 (m, C=O), 1600 (s, Aryl C–C), 1557, 1517 (s, C=O), 1486, 1449 (s), 1338 (s, N–CS), 1299 (s), 1255 (N–CS), 1188, 1165 (m), 1140 (s, C=S), 1098, 1078 (s), 1024, 1000, 975, 929 (m), 878 (s, Aryl C–H), 783 (vs, Aryl C–H), 740 (m), 702 (vs, Aryl C–H), 677, 655 (vs), 591, 533, 515 (s) cm^{-1}; **^{1}H NMR** (300 MHz, $CDCl_3$) δ 12.61 (1H, s, CON**H**CS), 9.09 (1H, s, CSN**H**Ar), 7.90 (2H, dd, *J* = 1.3, 6.0 Hz, **H^{10}**), 7.86 (1H, t, *J* = 1.5 Hz, **H^2**), 7.76 (1H, td, *J* = 2.2, 7.1 Hz, **H^4**), 7.67 (1H, tt, *J* = 1.5, 7.3 Hz, **H^{12}**), 7.55 (2H, tt, *J* = 1.7, 7.5 Hz, **H^{11}**), 7.41 (1H, td, *J* = 1.8, 7.7 Hz, **H^6**), 7.37 (1H, t, *J* = 7.7 Hz, **H^5**), 3.12 (1H, s, C≡C**H^8**); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 178.6 (**C**=S), 167.1 (**C**=O), 137.8 (**C^3**), 134.0 (**C^{12}**), 131.7 (**C^9**), 130.6 (**C^5**), 129.4 (**C^{11}**), 129.0 (**C^6**), 127.6 (**C^{10}**), 127.6 (**C^2**), 124.7 (**C^4**), 123.1 (**C^1**), 82.9 (**C^7**≡CH), 78.2 (C≡**C^8**H); **LRMS** (ES^+) *m*/*z*: 344 ([M + CH_3CN + Na]$^+$, 50), 330 (49), 303 ([M + Na]$^+$, 100), 281 ([M + H]$^+$, 33); **HRMS** (ES^+): calcd for $C_{16}H_{12}N_2OS$ (M), [3M + Cu]$^+$ = 903.1302, found 903.1038, [2M + Cu]$^+$ = 623.0631, found 623.0492; [2M + H]$^+$ = 561.1413, found 561.1358; [M + H]$^+$ = 281.0743, found 281.0742.

5'-*O*-(4,4'-Dimethoxytrityl)-5-[(3-thioureidophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine, 213



To a solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine **185** (0.31 g, 0.46 mmol) in anhydrous DMF (2.0 mL), under an argon atmosphere and in absence of light, were added distilled Et_3N (0.33 mL, 2.33 mmol, 5.0 eq), CuI (0.10 g, 0.51 mmol, 1.1 eq), then alkyne **209** (0.20 g, 1.16 mmol, 2.5 eq). The mixture was stirred for 1 hour then *tetrakis*(triphenylphosphine) palladium (0) (0.38 g, 0.33 mmol, 0.7 eq) was added and the reaction was stirred at rt for 1¼ hours. The reaction mixture was concentrated under high vacuum, co-evaporating with methanol/toluene, to give an

orange foam/glass. The residue was dissolved/suspended in methanol/acetone (1:1 v/v, 40 mL), and filtered. The solid was washed with methanol/acetone (1:1 v/v, 2 × 30 mL) and the combined filtrate was concentrated *in vacuo* and under high vacuum to give a golden orange foam/resin. Following purification by column chromatography (80-100% acetone/CH_2Cl_2 + 0.5% Et_3N), and drying under high vacuum, the desired product **213** was afforded as a yellow foam/resin (0.42 g) containing 21.1 mol % Et_3N.HI, 22.3 mol % triethylamine, and 4.0 mol% starting material **185**, with some Pd-based impurities (recalculated yield - 0.47 mmol, *quant.*).

R_f 0.23 (5:95, methanol/ethyl acetate + 0.5% Et_3N, A'), 0.17 (4:1, acetone/CH_2Cl_2 + 0.5% Et_3N); **^{1}H NMR** (400 MHz, d_6-DMSO) δ 9.68 (1H, s, N**H**), 8.01 (1H, s, **H^6**), 7.49 (1H, br s, **H$^{2"}$**), 7.44-7.39 (1H, m, N**H**CH_3), 7.42 (2H, d, *J* = 7.5 Hz, **H^{15}**), 7.37 (1H, br d, *J* = 8.3 Hz, **H$^{4"}$**), 7.32-7.27 (1H, m, **H$^{5"}$**), 7.30 (2H, d, *J* = 9.0 Hz, **H^{11}**), 7.30 (2H, d, *J* = 9.0 Hz, **H^{11}**), 7.26 (2H, t, *J* = 8.0 Hz, **H^{16}**), 7.17 (1H, t, *J* = 7.3 Hz, **H^{17}**), 6.96 (1H, br d, *J* = 7.5 Hz, **H$^{6"}$**), 6.85 (2H, d, *J* = 9.0 Hz, **H^{12}**), 6.84 (2H, d, *J* = 9.0 Hz, **H^{12}**), 6.16 (1H, t, *J* = 6.5 Hz, **H$^{1'}$**), 5.30 (1H, d, *J* = 4.5 Hz, 3'-O**H**), 4.28 (1H, td, *J* = 3.5, 9.5 Hz, **H$^{3'}$**), 3.98 (1H, td, *J* = 3.3, 4.5 Hz, **H$^{4'}$**), 3.67 (3H, s, OC**H$_3$**), 3.66 (3H, s, OC**H$_3$**), 3.22 (1H, dd, *J* = 5.0, 10.5 Hz, **H$^{5'}$**), 3.16 (1H, dd, *J* = 2.8, 10.3 Hz, **H$^{5'}$**), 2.87 (3H, d, *J* = 5.0 Hz, N**H**CH_3), 2.28 (1H, ddd, *J* = 3.3, 6.0, 13.3 Hz, **H$^{2'}$**), 2.15 (1H, td, *J* = 6.7, 13.6 Hz, **H$^{2'}$**), 2.08 (2H, s, N**H$_2$**); **^{13}C NMR** (100 MHz, d_6-DMSO) δ 181.2 (**C**SNH_2), 161.6 (**C^4**), 158.0, 158.0 (**C^{13}**), 153.2 (**C^2**), 144.7 (**C$^{3"}$**), 143.0 (**C^6**), 139.0 (**C^{14}**), 135.6, 135.3 (**C^{10}**), 129.7, 129.6 (**C^{11}**), 128.5 (**C$^{5"}$**), 127.9 (**C^{15}**), 127.6 (**C^{16}**,**C$^{6"}$**), 126.7 (**C^{17}**), 126.1 (**C$^{2"}$**), 123.9 (**C$^{4"}$**), 122.5 (**C$^{1"}$**), 113.2, 113.2 (**C^{12}**), 94.1 (**C^8**), 90.3 (**C^7**), 85.9 (**C$^{4'}$**), 85.8 (**C^9**Ar_3), 85.7 (**C$^{1'}$**), 80.8 (**C^5**), 70.6 (**C$^{3'}$**), 63.7 (**C$^{5'}$**), 55.0, 54.9 (O**C**H_3), 40.9 (**C$^{2'}$**), 27.9 (NH**C**H_3); **LRMS** (ES^+) *m*/*z*: 740 ($[M + Na]^+$, 100), 177 ($[M + H]^+$, 73); **HRMS** (ES^+): calcd for $C_{40}H_{39}N_5O_6S$ (M), $[M + K]^+$ = 756.2253, found 756.2260; $[M + Na]^+$ = 740.2513, found 740.2550; $[M + H]^+$ = 718.2694, found 718.2696.

9.2 Biophysical Studies

9.2.1 Synthesis of Oligonucleotides

Oligonucleotides were synthesised by ATDBio Ltd. and Dr. Bouamaied on an Applied Biosystems 394 automated DNA/RNA synthesiser or an ÄKTA™ oligopilot plus DNA/ RNA synthesiser using a standard solid-phase 0.2 or 1.0 μmole phosphoramidite synthesis cycle. The cycle comprises acid-catalysed (TCA) detritylation, coupling, capping (Ac_2O) and oxidation (iodine/pyridine/THF). Stepwise coupling efficiencies and overall yields were determined by an automated trityl cation conductivity monitoring function and in all cases were >96%.

All synthesised phosphoramidite monomers were dissolved in anhydrous DNA grade acetonitrile or freshly distilled CH_2Cl_2 and filtered through a Millipore Millex®-FH syringe filter (0.45μm, 13/25 mm). After removing the solvent *in vacuo*, the monomer was dissolved in freshly distilled CH_2Cl_2 and aliquots corresponding to 100-120 mg each were transferred in to dry, sealed, argon-flushed ABI reagent bottles. After drying to constant weight, under high vacuum over KOH for 2-3 days, the monomers were stored under positive argon pressure at $^{-}$20 °C until required.

All phosphoramidite monomers were dissolved in anhydrous DNA grade acetonitrile (or freshly distilled CH_2Cl_2 if poorly soluble) to a concentration of 0.1 M immediately before use.

9.2.2. Deprotection of Oligonucleotides

Oligonucleotides were deprotected/cleaved from solid-support by various conditions. Capping was carried out using acetic anhydride. Washing with 20% diethylamine/ acetonitrile was introduced in the latter half of the project.

OL1, 2, 6-9, 17: conc. aq ammonia, 55 °C, 5 hours (Capping only).
OL3-5: 30% aq methylamine, rt, 6-18 hours (Capping only), followed by treatment with DOWEX-H^+ resin for 15-17 hours.

OL10-13: 30% aq methylamine, rt, 18 hours (Capping only), followed by treatment with DOWEX-H^+ resin for 15-17 hours.

OL14-16: 30% aq methylamine, rt, 6 hours (Capping only), followed by treatment with DOWEX-H^+ resin for 15-17 hours.

OL18-25, 32, 33, 35-37, 40, 41, 43-59: conc. aq ammonia, rt, 24 hours, plus 1 hour at 50 °C (DEA Wash only).

OL26-31: conc. aq ammonia, 55 °C, 5 hours (DEA Wash plus Capping).

OL34, 38, 39, 42: conc. aq ammonia, rt, 12 hours, plus 2 hours at 50 °C (DEA Wash only).

OL60-63: conc. aq ammonia, rt, 12 hours (DEA Wash only).

The deprotection conditions used for oligonucleotides (OL64-70) in the Deprotection Study are discussed in Chapter 7.

9.2.3 Purification of Oligonucleotides

Oligonucleotides were purified by reversed-phase HPLC on a Gilson HPLC system using an Brownlee Aquapore column (C8, 8 × 250 mm, 300Å pore size). The system was controlled by Gilson 7.12 software, and oligonucleotide elution was monitored by UV absorption at 297 nm for non-modified and 310 nm for modified oligonucleotides.

The following purification protocol was used: Gradient of acetonitrile in ammonium acetate buffer from 0-50% buffer B over 30 minutes (flow rate: 4 mL/min; buffer A: 0.1 M ammonium acetate, pH 7.0; buffer B: 0.1 M ammonium acetate with 50% acetonitrile, pH 7.0).

9.2.4 Analysis of Oligonucleotides

Mass spectra of oligonucleotides were recorded in cooperation with Dr. Edrees-Abdou, using electrospray ionisation (ES) on a Fisons VG platform instrument in HPLC grade water, with tripropylamine to aid ionisation; or by Dr. Hong Li, using matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) analysis on a ThermoBioAnalysis Dynamo MALDI-TOF mass spectrometer in positive ion mode

using internal T_n standards.[256] ES MS data was deconvoluted/reprocessed using the Maximum Entropy facility in MassLynx software version 2.22.

CE analysis was conducted by Louise Hagon (ATDBio Ltd.), using a Beckman Coulter P/ACE™ MDQ Capillary Electrophoresis System, using the 32 Karat Software MDQ UV application, at a concentration of ~4 OD/mL. An ssDNA 100-R gel with Tris-Borate-7 M Urea system was used.

Analytical HPLC and ion-exchange analysis were conducted by ATDBio Ltd.

9.2.5 UV Melting Studies

UV Melting experiments were performed, partly in cooperation with Dr. Edrees-Abdou, on a Varian Cary 400 Scan UV-Visible Spectrophotometer in Hellma® SUPRASIL synthetic quartz cuvettes (10 mm pathlength, 1.5 mL sample volume), monitoring at 260 nm, using Cary WinUV Thermal application software. Buffers were prepared according to the literature procedures,[257] allowed to warm to rt and the pH checked before use.

9.2.5.1 Set-Up Procedure

Aqueous solutions of the TFO and duplex strands were mixed in a 5:1 ratio in an eppendorf tube and lyophilised. The residue was redissolved in 1.5 mL of the correct buffer, pH 6.1, 6.2, 6.4, 6.6, 7.0 or 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA) to afford a 5 μM:1 μM concentration of TFO:duplex. The samples were degassed for 1 minute using a sonic-bath, then filtered into the cuvettes though Kinesis regenerated cellulose syringe filters (0.45 μm, 15 mm). The samples were subjected to the desired melt programme alongside a matched cell reference blank.

9.2.5.2 UV Melt Programmes

Programme 1

Stage	Temp (°C)	Rate (°C/min)	Hold time (min)	Data Interval (°C)
Melt 1	10→80	10.0	5	1.0
Anneal 1	80→10	1.0	2	1.0
Melt 2	10→80	0.5	2	0.25
Anneal 2	80→10	0.5	2	0.25
Melt 3	10→50	0.5	2	0.25
Anneal 3	50→20	0.5	2	0.25

Programme 2

Stage	Temp (°C)	Rate (°C/min)	Hold time (min)	Data Interval (°C)
Melt 1	10→80	10.0	2	1.0
Anneal 1	80→10	1.0	2	1.0
Melt 2/3	10→80	0.5	2	0.25
Anneal 2/3	80→10	0.5	2	0.25
Melt 4	10→50	0.5	2	0.25
Anneal 4	50→20	0.5	0	0.25

Programme 3

Stage	Temp (°C)	Rate (°C/min)	Hold time (min)	Data Interval (°C)
Melt 1	20→80	10.0	2	1.0
Anneal 1	80→15	0.5	20	0.1
Melt 2/3	15→80	0.5	2	0.1
Anneal 2/3	80→15	0.5	20	0.1
Melt 4/5	15→50	0.25	2	0.1
Anneal 4	50→15	0.25	20	0.1
Anneal 5	50→20	10.0	0	1.0

Following a fast melt/anneal for equilibration of denaturation of the triplex, heat/cool cycles up to 80 °C were recorded, followed by cycles up to 50 °C, at a slower rate and smaller data interval, for more accurate determination of the triplex T_m and reduction of hysteresis.

9.2.5.3 Data Analysis

T_m values were derived from the derivatives of melting curves for each experiment using Cary WinUV Thermal application software or OriginPro 7.5.

9.2.6 Fluorescence Melting Studies

Fluorescence melting experiments were conducted, partly in cooperation with Dr. Edrees-Abdou, on a Roche LightCycler® 1.5 instrument in LightCycler glass capilliaries (20 μL volume) using Roche LightCycler Software Version 3.5. The LightCycler has one excitation source (488 nm) and change in fluorescence was monitored at 520 nm.

9.2.6.1 Set-Up Procedure

Master solutions (5 μM) of each oligonucleotide involved in each LightCycler run were made. For each experiment within that run, 4.5 μL and 45 μL of the appropriate duplex and TFO respectively were lyophilised in 100 μL eppendorf tubes then redissolved in 90 μL of the correct buffer, pH 6.0 (20 mM NaOAc, 200 mM NaCl), pH 6.2 (10 mM sodium phosphate, 200 mM NaCl) or, pH 6.2, 6.6, 7.0 or 7.0 + 2 mM spermine (10 mM sodium phosphate, 200 mM NaCl, 1 mM Na_2EDTA), to afford a 10:1 ratio of TFO:hairpin duplex (5:0.5 μM). Samples containing only the duplex were used as references within each run. The samples were mixed by vortex then centrifugation, and were degassed in a sonic bath for 1 minute. Aliquots of 20 μL were pipetted into the top of two or three LightCycler tubes per sample, which left one or two spare aliquots. The tubes were capped and centrifugated gently to avoid breakage, then loaded into the carousel and subjected to the desired LC melt programme.

9.2.6.2 LC Melt Programmes

Programme 1

Stage	Temp (°C)	Step Rate (°C/s)	Hold time (s)	Rate (°C/min)
Denaturation	30→95	20.0	300	4.0
Hold	95	–	600	–
Annealing	95→30	1.0	300	0.2
Incubation	30	–	3600	–
Melting	30→95	1.0	300	0.2
Hold	95	–	600	–
Cool	95→30	20.0	1, 5, 300	≥4.0

Programme 2

Stage	Temp (°C)	Step Rate (°C/s)	Hold time (s)	Rate (°C/min)
Denaturation	30→95	20.0	300	4.0
Hold	95	–	600	–
Annealing	95→28	1.0	300	0.2
Incubation	28	–	3600	–
Melting	30→28	1.0	300	0.2
Hold	95	–	600	–
Cool	95→30	20.0	5	240

Programme 3

Stage	Temp (°C)	Step Rate (°C/s)	Hold time (s)	Rate (°C/min)
Denaturation	30→95	20.0	300	4.0
Hold	95	–	600	–
Annealing	95→27	1.0	300	0.2
Incubation	27	–	3600	–
Slow Melting	27→95	1.0	300	0.2
Hold 2	95	–	600	–
Cool	95→30	20.0	300	4.0

Programme 4

Stage	Temp (°C)	Step Rate (°C/s)	Hold time (s)	Rate (°C/min)
Denaturation	30→95	20.0	300	4.0
Hold	95	–	600	–
Annealing	95→30	1.0	240	0.25
Incubation	30	–	3600	–
Slow Melting	30→95	1.0	300	0.2
Hold 2	95	–	600	–
Cool	95→30	20.0	300	0.25

Following fast heating from 30 to 95 °C, and holding at 95 °C for 10 minutes to fully denaturate the DNA, a slow stepwise cool/annealing process was recorded (1.0 °C step, 0.2-0.25 °C/min). After incubation at low temperature (27-30 °C) for 1 hour to equilibrate triplex formation, a slow stepwise heat/melting process was recorded (1.0 °C step, 0.2 °C/min). The samples were held at 95 °C for 10 minutes to equilibrate before rapid cooling back to 30 °C. Approximately ¼ of sample volume (~5 μL) is lost by evaporation during the course of the LC programme.

9.2.6.3 Data Analysis

T_m values were derived from the negative derivatives of melting curves for each experiment using the "Manual T_m" feature in Roche LightCycler Software Version 3.5.

9.2.7 Determining UV/Fluorescence Properties of Fluorescent Nucleosides

9.2.7.1 Deprotection of Nucleosides

DMT-protected nucleosides (**143,182,225,226**) were treated up to three times with 3% TCA/DCM followed by removal of DCM *in vacuo* and trituration with diethyl ether (×3). The aniline **143**, acetamide **225** and urea **182** deprotected nucleosides were afforded in this way as pale yellow to orange powders. The guanidine nucleoside (non-protonated) was further treated with 20% Et_2NH/CH_3CN for 24 hours, concentrated *in vacuo*, then treated with a mixture of 40% NH_3/H_2O and 7N NH_3/CH_3OH (1:2 v/v) for 60 hours, before concentration *in vacuo* and recrystallisation (CH_3OH/Et_2O). The fully deprotected nucleoside **226** was afforded as a yellow-orange powder. Treatment with 20% diethylamine solution cleaved only one 2-cyanoethoxycarbonyl (CEOC) protecting group from the guanidine moiety. Subsequent treatment with ammonia was necessary to cleave the second. Treatment with ammonia only would have resulted in cyclisation to form a stable triazine analogue of the guanidine moiety as reported in work by Prakash *et al*.[225] (Section 7.3.2.7).

9.2.7.2 UV Absorption Measurements

Solutions of nucleosides were made in either HPLC grade methanol ($^{A}PP_p$ **143**, $^{Ac}PP_p$ **225**) or anhydrous DMF ($^{U}PP_p$ **182**, $^{G}PP_p$ **226**) up to 0.174 μM ($^{A}PP_p$), 0.109 μM ($^{Ac}PP_p$), 0.151 μM ($^{U}PP_p$) and 0.154 μM ($^{G}PP_p$) concentrations. The solution being tested was filtered through a Millipore Millex®-FH syringe filter (0.45μm, 13 mm) into a Hellma® SUPRASIL synthetic quartz cuvette (10 mm pathlength, 1.5 mL sample volume). UV absorption/excitation spectra were recorded at this and 2-3 lower concentrations by successive dilution by factor of 2, on a Varian Cary 400 Scan UV-Visible Spectrophotometer using Cary WinUV Scan application software. Spectra for

compounds run in DMF were dominated by noise below 260 nm, due to loss of solvent transparency to UV light below ~270 nm.

Extinction coefficients (ε / L mol^{-1} cm^{-1}) were calculated using the Beer-Lambert Law (Equation 9.1)

$$A = \varepsilon.c.l$$

Equation 9.1. Beer-Lambert Law; A = absorbance (dimensionless), ε = molar extinction coefficient (L mol^{-1} cm^{-1}), c = concentration (M), l = pathlength (cm).

For each absorption/excitation maximum, the extinction coefficient was calculated by a three-step process:

1. At each concentration measured, the value was calculated using Equation 9.1, then the average of those values was taken (ignoring outliers).
2. The value was given as the gradient of a plot of Absorbance *vs* Molar Concentration, ignoring significant outliers.
3. The average of the two values was taken to give an extinction coefficient for each absorption/excitation maximum.

9.2.7.3 Fluorescence Measurements

Solutions of nucleosides were made in either HPLC grade methanol ($^{A}PP_p$ **143**, $^{Ac}PP_p$ **225**) or anhydrous DMF ($^{U}PP_p$ **182**, $^{G}PP_p$ **226**) up to concentrations of 1-3 mg/mL. The solution being tested was filtered through a Millipore Millex®-FH syringe filter (0.45μm, 13 mm) into a Hellma® SUPRASIL synthetic quartz cuvette for fluorimetry (10 mm pathlength, 1.5 mL sample volume). An emission spectrum was run for each sample, by irradiation at the principal absorption maximum (λ ~360 nm) for that sample, on a Varian Cary Eclipse Fluorescence Spectrophotometer using Cary Eclipse Scan application software. Each spectrum was an average of 10 consecutive scans from 380-600 nm.

References

1. Blackburn, G.M., Gait, M.J., Loakes, D., Williams, D.M., *Nucleic Acids in Chemistry and Biology*, 3rd ed., RSC Publishing, **2006**.
2. Neidle, S., *DNA Structure and Recognition*, 1st ed., IRL Press, Oxford University Press Inc., New York, **1994**.
3. Watson, J.D., *The Double Helix: A New Critical Edition*, Weidenfeld and Nicolson Ltd., London, UK, **1981**.
4. Watson, J.D., Crick, F.H.C., *Nature (London)*, **1953**, 171, 737-738.
5. Franklin, R.E., Gosling, R.G., *Nature (London)*, **1953**, 171, 740-741.
6. Franklin, R.E., Gosling, R.G., *Nature (London)*, **1953**, 172, 156-157.
7. Wilkins, M.H.F., Stokes, A.R., Wilson, H.R., *Nature (London)*, **1953**, 171, 738-740.
8. Watson, J.D., Crick, F.H.C., *Nature (London)*, **1953**, 171, 964-967.
9. Wang, A.H.-J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Mare, G., Rich, A., *Nature (London)*, **1979**, 282(5740), 680-686.
10. Fox, K.R., Darby, R.A.J., Triple helix-specific ligands, in *Small Molecule DNA and RNA Binders*, ed. M. Demeunynck, C. Bailly, W.D. Wilson, WILEY-VCH Verlag GmbH & Co., Weinheim, Germany, **2003**, Vol. 2, Ch. 14, pp 360-383.
11. Rusling, D.A., Brown, T., Fox, K.R., DNA Recognition by Triple Helix Formation, in *Sequence-specific DNA Binding Agents*, ed. M.J. Waring, RSC Publishing, Cambridge, UK, **2006**, Ch. 1, pp 1-27.
12. Thuong, N.T., Hélène, C., *Angew. Chem. Int. Ed. Engl.*, **1993**, 32(5), 666-690.
13. Vasquez, K.M., Glazer, P.M., *Q. Rev. Biophys.*, **2002**, 35(1), 89-107.
14. Abu-Daya, A., Brown, P.M., Fox, K.R., *Nucleic Acids Res.*, **1995**, 23(17), 3385-3392.
15. Felsenfeld, G., Davies, D.R., Rich, A., *J. Am. Chem. Soc.*, **1957**, 79, 2023-2024.
16. Asensio, J.L., Brown, T., Lane, A.N., *Structure*, **1999**, 7(1), 1-11.
17. Winters, T.A., *Curr. Opin. Mol. Ther.*, **2000**, 2(6), 670-681.
18. Praseuth, D., Guieysse, A.L., Helene, C., *BBA-Gene Struct. Expr.*, **1999**, 1489(1), 181-206.
19. Karympalis, V., Kalopita, K., Zarros, A., Carageorgiou, H., *Biochemistry (Moscow)*, **2004**, 69(8), 855-860.
20. Carbone, G.M., Napoli, S., Valentini, A., Cavalli, F., Watson, D.K., Catapano, C.V., *Nucleic Acids Res.*, **2004**, 32(14), 4358-4367.

21. Majumdar, A., Puri, N., Cuenoud, B., Natt, F., Martin, P., Khorlin, A., Dyatkina, N., George, A.J., Miller, P.S., Seidman, M.M., *J. Biol. Chem.*, **2003**, 278(13), 11072-11077.
22. Buchini, S., Leumann, C.J., *Curr. Opin. Chem. Biol.*, **2003**, 7(6), 717-726.
23. Moser, H.E., Dervan, P.B., *Science*, **1987**, 238(4827), 645-650.
24. Antony, S., Arimondo, P.B., Sun, J.S., Pommier, Y., *Nucleic Acids Res.*, **2004**, 32(17), 5163–5173.
25. Woynarowski, J.M., *Adv. DNA Seq. Spec. Agents*, **2002**, 4, 1-27.
26. Seidman, M.M., Glazer, P.M., *J. Clin. Invest.*, **2003**, 112(4), 487-494.
27. Guillonneau, F., Guieysse, A.L., Nocentini, S., Giovannangeli, C., Praseuth, D., *Nucleic Acids Res.*, **2004**, 32(3), 1143-1153.
28. Kalish, J.M., Seidman, M.M., Weeks, D.L., Glaze, P.M., *Nucleic Acids Res.*, **2005**, 33(11), 3492–3502.
29. Ma, Z., Taylor, J.-S., *Proc. Natl. Acad. Sci. U. S. A.*, **2000**, 97(21), 11159-11163.
30. Bello-Roufai, M., Roulon, T., Escude, C., *Chem. Biol.*, **2004**, 11(4), 509-516.
31. Whiteford, N., Haslam, N., Weber, G., Prügel-Bennett, A., Essex, J.W., Roach, P.L., Bradley, M., Neylon, C., *Nucleic Acids Res.*, **2005**, 33(19), e171.
32. Hoogsteen, K., *Acta Crystallogr.*, **1959**, 12, 822-823.
33. Fox, K.R., *Curr. Med. Chem.*, **2000**, 7(1), 17-37.
34. Gowers, D.M., Fox, K.R., *Nucleic Acids Res.*, **1999**, 27(7), 1569-1577.
35. Sollogoub, M., Dominguez, B., Fox, K.R., Brown, T., *Chem. Commun.*, **2000**, 23, 2315-2316.
36. Sollogoub, M., Darby, R.A.J., Cuenoud, B., Brown, T., Fox, K.R., *Biochemistry*, **2002**, 41(23), 7224-7231.
37. Osborne, S.D., Powers, V.E.C., Rusling, D.A., Lack, O., Fox, K.R., Brown, T., *Nucleic Acids Res.*, **2004**, 32(15), 4439-4447.
38. Cuenoud, B., Casset, F., Husken, D., Natt, F., Wolf, R.M., Altmann, K.-H., Martin, P., Moser, H.E., *Angew. Chem. Int. Ed.*, **1998**, 37(9), 1288-1291.
39. Bijapur, J., Keppler, M.D., Bergqvist, S., Brown, T., Fox, K.R., *Nucleic Acids Res.*, **1999**, 27(8), 1802-1809.
40. Froehler, B.C., Wadwani, S., Terhorst, T.J., Gerrard, S.R., *Tetrahedron Lett.*, **1992**, 33(37), 5307-5310.
41. Prakash, T.P., Puschl, A., Lesnik, E., Mohan, V., Tereshko, V., Egli, M., Manoharan, M., *Org. Lett.*, **2004**, 6(12), 1971-1974.

42. Manoharan, M., Prakash, T.P., Barber-Peoc'h, I., Bhat, B., Vasquez, G., Ross, B.S., Cook, P.D., *J. Org. Chem.*, **1999**, 64(17), 6468-6472.
43. Deglane, G., Abes, S., Michel, T., Prévot, P., Vives, E., Debart, F., Barvik, I., Lebleu, B., Vasseur, J.-J., *ChemBioChem*, **2006**, 7(4), 684 – 692.
44. Roig, V., Asseline, U., *J. Am. Chem. Soc.*, **2003**, 125(15), 4416-4417.
45. Lee, J.S., Woodsworth, M.L., Latimer, L.J.P., Morgan, A.R., *Nucleic Acids Res.*, **1984**, 12(16), 6603-6614.
46. Xodo, L.E., Manzini, G., Quadrifoglio, F., Van der Marel, G.A., Van Boom, J.H., *Nucleic Acids Res.*, **1991**, 19(20), 5625-31.
47. Povsic, T.J., Dervan, P.B., *J. Am. Chem. Soc.*, **1989**, 111(8), 3059-3061.
48. Singleton, S.F., Dervan, P.B., *Biochemistry*, **1992**, 31(45), 10995-11003.
49. Barawkar, D.A., Kumar, V.A., Ganesh, K.N., *Biochem. Biophys. Res. Commun.*, **1994**, 205(3), 1665-1670.
50. Barawkar, D.A., Rajeev, K.G., Kumar, V.A., Ganesh, K.N., *Nucleic Acids Res.*, **1996**, 24(7), 1229–1237.
51. Rajeev, K.G., Jadhav, V.R., Ganesh, K.N., *Nucleic Acids Res.*, **1997**, 25(21), 4187–4193.
52. Doronina, S.O., Behr, J.-P., *Tetrahedron Lett.*, **1998**, 39(7), 547-550.
53. Robles, J., Grandas, A., Pedroso, E., *Tetrahedron*, **2001**, 57(1), 179-194.
54. Robles, J., Grandas, A., Pedroso, E., *Nucleos. Nucleot. Nucl.*, **2003**, 22(5-8), 1085-1087.
55. Ono, A., Ts'o, P.O.P., Kan, L.-S., *J. Am. Chem. Soc.*, **1991**, 113(10), 4032-4033.
56. Ono, A., Ts'o, P.O.P., Kan, L.-S., *J. Org. Chem.*, **1992**, 57(11), 3225-3230.
57. Christensen, C., Eldrup, A.B., Haaima, G., Nielsen, P.E., *Bioorg. Med. Chem. Lett.*, **2002**, 12(21), 3121-3124.
58. Nielsen, P.E., Egholm, M., Bèrg, R.H., Buchardt, O., *Science*, **1991**, 254(5037), 1497-1500.
59. Egholm, M., Buchardt, O., Nielsen, P.E., Bèrg, R.H., *J. Am. Chem. Soc.*, **1992**, 114(5), 1895-1897.
60. Sollogoub, M., Fox, K.R., Powers, V.E.C., Brown, T., *Tetrahedron Lett.*, **2002**, 43(17), 3121-3123.
61. Powers, V.E.C., Synthesis and studies of nucleoside analogues in triplex forming oligonucleotides, PhD Thesis, University of Southampton, Southampton, UK, **2004**.

62. Bates, P.J., Laughton, C.A., Jenkins, T.C., Capaldi, D.C., Roselt, P.D., Reese, C.B., Neidle, S., *Nucleic Acids Res.*, **1996**, 24(21), 4176-4184.
63. Cassidy, S.A., Slickers, P., Trent, J.O., Capaldi, D.C., Roselt, P.D., Reese, C.B., Neidle, S., Fox, K.R., *Nucleic Acids Res.*, **1997**, 25(24), 4891-4898.
64. Hildbrand, S., Leumann, C.J., *Angew. Chem. Int. Ed. Engl.*, **1996**, 35(17), 1968-1970.
65. Hildbrand, S., Blaser, A., Parel, S.P., Leumann, C.J., *J. Am. Chem. Soc.*, **1997**, 119(24), 5499-5511.
66. Reese, C.B., Wu, Q., *Org. Biomol. Chem.*, **2003**, 1(18), 3160-3172.
67. Rusling, D.A., Powers, V.E.C., Ranasinghe, R.T., Wang, Y., Osborne, S.D., Brown, T., Fox, K.R., *Nucleic Acids Res.*, **2005**, 33(9), 3025-3032.
68. von Krosigk, U., Benner, S.A., *J. Am. Chem. Soc.*, **1995**, 117(19), 5361-5362.
69. Hutter, D., Benner, S.A., *J. Org. Chem.*, **2003**, 68(25), 9839-9842.
70. Berressem, R., Engels, J.W., *Nucleic Acids Res.*, **1995**, 23(17), 3465-3472.
71. Parsch, U., Engels, J.W., *Chem. Eur. J.*, **2000**, 6(13), 2409-2424.
72. Xiang, G., Soussou, W., McLaughlin, L.W., *J. Am. Chem. Soc.*, **1994**, 116(24), 11155-11156.
73. Xiang, G., Bogacki, R., McLaughlin, L.W., *Nucleic Acids Res.*, **1996**, 24(10), 1963-1970.
74. Xiang, G., McLaughlin, L.W., *Tetrahedron*, **1998**, 54(3-4), 375-392.
75. Brodin, P., Pinskaya, M., Parsch, U., Bischerour, J., Leh, H., Romanova, E., Engels, J.W., Gottikh, M., Mouscadet, J.F., *Nucleos. Nucleot. Nucl.*, **2001**, 20(4-7), 481-486.
76. Wang, Y., Rusling, D.A., Powers, V.E.C., Lack, O., Osborne, S.D., Fox, K.R., Brown, T., *Biochemistry*, **2005**, 44(15), 5884-5892.
77. Hunziker, J., Priestley, E.S., Brunar, H., Dervan, P.B., *J. Am. Chem. Soc.*, **1995**, 117(9), 2661-2662.
78. Brunar, H., Dervan, P.B., *Nucleic Acids Res.*, **1996**, 24(11), 1987-1991.
79. Koshlap, K.M., Schultze, P., Brunar, H., Dervan, P.B., Feigon, J., *Biochemistry*, **1997**, 36(9), 2659-2668.
80. St. Clair, A., Xiang, G., McLaughlin, L.W., *Nucleos. Nucleot. Nucl.*, **1998**, 17(5), 925-938.
81. Koh, J.S., Dervan, P.B., *J. Am. Chem. Soc.*, **1992**, 114(4), 1470-1478.
82. Radhakrishnan, I., Patel, D.J., Priestley, E.S., Nash, H.M., Dervan, P.B., *Biochemistry*, **1993**, 32(41), 11228-11234.

83. Priestley, E.S., Dervan, P.B., *J. Am. Chem. Soc.*, **1995**, 117, 4761-4765.
84. D'Costa, M., Kumar, V.A., Ganesh, K.N., *J. Org. Chem.*, **2003**, 68(25), 9839-9842.
85. Marfurt, J., Parel, S.P., Leumann, C.J., *Nucleic Acids Res.*, **1997**, 25(10), 1875-1882.
86. Miller, P.S., Bhan, P., Cushman, C.D., Trapane, T.L., *Biochemistry*, **1992**, 31(29), 6788-93.
87. Krawczyk, S.H., Milligan, J.F., Wadwani, S., Moulds, C., Froehler, B.C., Matteucci, M.D., *Proc. Natl. Acad. Sci. U. S. A.*, **1992**, 89(9), 3761-3764.
88. Jetter, M.C., Hobbs, F.C., *Biochemistry*, **1993**, 32(13), 3249-3254.
89. Lin, S.-B., Wu, C.-M., Lin, W.-C., Kan, L.-S., *J. Chin. Chem. Soc. (Taipei)*, **1999**, 46(5), 693-697.
90. Young, S.L., Krawczyk, S.H., Matteucci, M.D., Toole, J.J., *Proc. Natl. Acad. Sci. U. S. A.*, **1991**, 88, 10023-10026.
91. Durland, R.H., Rao, T.S., Jayaraman, K., Revankar, G.R., *Bioconjugate Chem.*, **1995**, 6(3), 278-282.
92. Staubli, A.B., Dervan, P.B., *Nucleic Acids Res.*, **1994**, 22(13), 2637-2642.
93. Flanagan, W.M., Wagner, R.W., Grant, D., Lin, K.-Y., Matteucci, M.D., *Nature Biotech.*, **1999**, 17, 48-52.
94. Matteucci, M.D., von Krosigk, U., *Tetrahedron Lett.*, **1996**, 37(29), 5057-5060.
95. Lin, K.-Y., Jones, R.J., Matteucci, M., *J. Am. Chem. Soc.*, **1995**, 117(13), 3873-3874.
96. Sandin, P., Wilhelmsson, L.M., Lincoln, P., Powers, V.E.C., Brown, T., Albinsson, B., *Nucleic Acids Res.*, **2005**, 33(16), 5019–5025.
97. Wilhelmsson, L.M., Holmén, A., Lincoln, P., Nielsen, P.E., Nordén, B., *J. Am. Chem. Soc.*, **2001**, 123(10), 2434-2435.
98. Guianvarc'h, D., Benhida, R., Fourrey, J.-L., Maurisse, R., Sun, J.-S., *Chem. Commun.*, **2001**, (18), 1814–1815.
99. Guianvarc'h, D., Fourrey, J.-L., Maurisse, R., Sun, J.S., Benhida, R., *Bioorg. Med. Chem.*, **2003**, 11(13), 2751–2759.
100. Griffin, L.C., Kiessling, L.L., Beal, P.A., Gillespie, P., Dervan, P.B., *J. Am. Chem. Soc.*, **1992**, 114(21), 7976-7982.
101. Koshlap, K.M., Gillespie, P., Dervan, P.B., Feigon, J., *J. Am. Chem. Soc.*, **1993**, 115(17), 7908-7909.

102. Wang, E., Koshlap, K.M., Gillespie, P., Dervan, P.B., Feigon, J., *J. Mol. Biol.*, **1996**, 257(5), 1052–1069.
103. Blommers, M.J.J., Natt, F., Jahnke, W., Cuenoud, B., *Biochemistry*, **1998**, 37(51), 17714-17725.
104. Puri, N., Majumdar, A., Cuenoud, B., Miller, P.S., Seidman, M.M., *Biochemistry*, **2004**, 43(5), 1343-1351.
105. Buchini, S., Leumann, C.J., *Eur. J. Org. Chem.*, **2006**, (14), 3152–3168.
106. Yoon, K., Hobbs, C.A., Koch, J., Sardaro, M., Kutny, R., Weis, A.L., *Proc. Natl. Acad. Sci. U. S. A.*, **1992**, 89(9), 3840-3844.
107. Obika, S., Hari, Y., Sekiguchi, M., Imanishi, T., *Angew. Chem. Int. Ed.*, **2001**, 40(11), 2079-2081.
108. Obika, S., Hari, Y., Sekiguchi, M., Imanishi, T., *Chem. Eur. J.*, **2002**, 8(20), 4796-4802.
109. Hari, Y., Obika, S., Sekiguchi, M., Imanishi, T., *Tetrahedron*, **2003**, 59(27), 5123-5128.
110. Obika, S., Hari, Y., Sugimoto, T., Sekiguchi, M., Imanishi, T., *Tetrahedron Lett.*, **2000**, 41(46), 8923-8927.
111. Prévot-Halter, I., Leumann, C.J., *Bioorg. Med. Chem. Lett.*, **1999**, 9(18), 2657-2660.
112. Prévot, I., Leumann, C.J., *Helv. Chim. Acta*, **2002**, 85(2), 502-515.
113. Buchini, S., Leumann, C.J., *Tetrahedron Lett.*, **2003**, 44(27), 5065-5068.
114. Buchini, S., Leumann, C.J., *Angew. Chem. Int. Ed.*, **2004**, 43(30), 3925 –3928.
115. Rao, M.S., Esho, N., Sergeant, C., Dembinski, R., *J. Org. Chem.*, **2003**, 68(17), 6788-6790.
116. Robins, M.J., Barr, P.J., *J. Org. Chem.*, **1983**, 48(11), 1854-1862.
117. Cruickshank, K.A., Stockwell, D.L., *Tetrahedron Lett.*, **1988**, 29(41), 5221-5224.
118. Woo, J., Meyer Jr., R.B., Gamper, H.B., *Nucleic Acids Res.*, **1996**, 24(13), 2470–2475.
119. Berry, D.A., Jung, K.-Y., Wise, D.S., Sercel, A.D., Pearson, W.H., Mackie, H., Randolph, J.B., Somers, R.L., *Tetrahedron Lett.*, **2004**, 45(11), 2457–2461.
120. Hudson, R.H.E., Ghorbani-Choghamarani, A., *Synlett*, **2007**, (6), 870-873.
121. McGuigan, C., Barucki, H., Blewett, S., Carangio, A., Erichsen, J.T., Andrei, G., Snœck, R., Clercq, E.D., Balzarini, J., *J. Med. Chem.*, **2000**, 43(26), 4993-4997.

122. Ranasinghe, R.T., Rusling, D.A., Powers, V.E.C., Fox, K.R., Brown, T., *Chem. Commun.*, **2005**, (20), 2555-2557.
123. Sonogashira, K., Tohda, Y., Hagihara, N., *Tetrahedron Lett.*, **1975**, 50, 4467-4470.
124. Hobbs Jr., F.W., *J. Org. Chem.*, **1989**, 54(14), 3420-3422.
125. Chin, T.-M., Chung, K.-Y., Chen, J.-J., Lin, W.-C., Kan, L.S., *J. Chin. Chem. Soc. (Taipei)*, **1999**, 46(5), 751-758.
126. Chin, T.-M., Tseng, M.-H., Chung, K.-Y., Hung, F.-S., Lin, S.-B., Kan, L.-S., *J. Biomol. Struct. Dyn.*, **2001**, 19(3), 543-553.
127. Gianolio, D.A., McLaughlin, L.W., *J. Am. Chem. Soc.*, **1999**, 121(26), 6334-6335.
128. Guzzo-Pernell, N., Tregear, G.W., Haralambidis, J., Lawlor, J.M., *Nucleosides Nucleotides*, **1998**, 17(7), 1191-1207.
129. Huang, C.-Y., Bi, G., Miller, P.S., *Nucleic Acids Res.*, **1996**, 24(13), 2606-2613.
130. Huang, C.-Y., Cushman, C.D., Miller, P.S., *J. Org. Chem.*, **1993**, 58(19), 5048-5049.
131. Huang, C.-Y., Miller, P.S., *J. Am. Chem. Soc.*, **1993**, 115(22), 10456-10457.
132. Guzzo-Pernell, N., Lawlor, J.M., Tregear, G.W., Haralambidis, J., *Aust. J. Chem.*, **1998**, 51(11), 965-972.
133. Zimmerman, S.C., Schmitt, P., *J. Am. Chem. Soc.*, **1995**, 117(43), 10769-10770.
134. Lengeler, D., Weisz, K., *Tetrahedron Lett.*, **2001**, 42(8), 1479–1481.
135. Mertz, E., Mattei, S., Zimmerman, S.C., *Org. Lett.*, **2000**, 2(19), 2931-2934.
136. Xiao, Z., Weisz, K., *J. Phys. Org. Chem.*, **2007**, 20(10), 771-777.
137. Mertz, E., Mattei, S., Zimmerman, S.C., *Bioorg. Med. Chem.*, **2004**, 12(6), 1517-1526.
138. Purwanto, M.G.M., Lengeler, D., Weisz, K., *Tetrahedron Lett.*, **2002**, 43(1), 61-64.
139. Purwanto, M.G.M., Weisz, K., *Curr. Org. Chem.*, **2003**, 7(5), 427-446.
140. Purwanto, M.G.M., Weisz, K., *Tetrahedron Lett.*, **2006**, 47(23), 3849–3852.
141. Durland, R.H., Rao, S., Bodepudi, V., Seth, D.M., Jayaraman, K., Revankar, G.R., *Nucleic Acids Res.*, **1995**, 23(4), 647-653.
142. Sasaki, S., Taniguchi, Y., Takahashi, R., Senko, Y., Kodama, K., Nagatsugi, F., Maeda, M., *J. Am. Chem. Soc.*, **2003**, 126(2), 516-528.
143. de la Paz, M.L., González, C., Vicent, C., *Chem. Commun.*, **2000**, (5), 411–412.
144. McBride, L.J., Caruthers, M.H., *Tetrahedron Lett.*, **1983**, 24(3), 245-248.

145. Carreño, M.C., Ruano, J.L.G., Sanz, G., Toledo, M.A., Urbano, A., *Tetrahedron Lett.*, **1996**, 37(23), 4081-4084.
146. Das, B., Kundu, N.G., *Synth. Commun.*, **1988**, 18(8), 855-868.
147. Ge, P., Kirk, K.L., *J. Org. Chem.*, **1997**, 62(10), 3340-3343.
148. Hudson, R.H.E., Li, G., Tse, J., *Tetrahedron Lett.*, **2002**, 43(8), 1381-1386.
149. Robins, M.J., Barr, P.J., Giziewicz, J., *Can. J. Chem.*, **1982**, 60, 554-557.
150. Boden, N., Bushby, R.J., Cammidge, A.N., Headdock, G., *Tetrahedron Lett.*, **1995**, 36(47), 8685-8686.
151. Harrowven, D.C., Woodcock, T., Howes, P.D., *Tetrahedron Lett.*, **2002**, 43(51), 9327-9329.
152. Kajigaeshi, S., Kakinami, T., Yamasaki, H., Fujisaki, S., Manabu, K., Okamoto, T., *Chem. Lett.*, **1987**, (11), 2109-2112.
153. Quintin, J., Lewin, G., *Tetrahedron Lett.*, **2004**, 45, 3635–3638.
154. Kosynkin, D.V., Tour, J.M., *Org. Lett.*, **2001**, 3(7), 991-992.
155. Kajigaeshi, S., Kakinami, T., Moriwaki, M., Watanabe, M., Fujisaki, S., Okamoto, T., *Chem. Lett.*, **1988**, (5), 795-798.
156. Mitra, S.S., Sreekumar, K., *Indian J. Chem., Sect. B*, **1997**, 36B(2), 133-136.
157. Mitra, S.S., Sreekumar, K., *React. Funct. Polym.*, **1997**, 32(3), 281-291.
158. Mitra, S.S., Sreekumar, K., *Polym. Int.*, **1997**, 42(2), 173-178.
159. Mitra, S.S., Sreekumar, K., *Polymer*, **1997**, 38(6), 1363-1366.
160. Haszeldine, R.N., Sharpe, A.G., *J. Chem. Soc.*, **1952**, 993-1001.
161. Hajipour, A.R., Arbabian, M., Ruoho, A.E., *J. Org. Chem.*, **2002**, 67(24), 8622-8624.
162. Pan, B.-C., Chen, Z.-H., Chu, E., Chu, M.-Y.W., Chu, S.-H., *Nucleosides Nucleotides*, **1998**, 17(12), 2367-2382.
163. Hirota, K., Sako, M., Sajiki, H., *Heterocycles*, **1997**, 46, 547-554.
164. Djerassi, C., Lenk, C.T., *J. Am. Chem. Soc.*, **1953**, 75(14), 3493-3495.
165. Markiewicz, W.T., Padyukova, N.S., Samek, Z., Smrt, J., *Collect. Czech. Chem. Commun.*, **1980**, 45(6), 1860-1865.
166. Mitsunobu, O., *Synthesis*, **1981**, 1-28.
167. Vorbrüggen, H., Krolikiewicz, K., Bennua, B., *Chem. Ber.*, **1981**, 114(4), 1235-1245.
168. Robins, M.J., Barr, P.J., *Tetrahedron Lett.*, **1981**, 22(5), 421-424.
169. Wang, L., Yan, J., Li, P.H., Wang, M., Su, C.N., *J. Chem. Res., Synop.*, **2005**, (2), 112-115.

170. Bijapur, J., Factors Affecting the Stability of Nucleic Acids, PhD Thesis, University of Southampton, Southampton, UK, **1999**.
171. Alami, M., Crousse, B., Ferri, F., *J. Organomet. Chem.*, **2001**, 624, 114-123.
172. Bhattacharya, S., Sengupta, S., *Tetrahedron Lett.*, **2004**, 45, 8733-8736.
173. Negishi, E.-i., Qian, M., Zeng, F., Anastasia, L., Babinski, D., *Org. Lett.*, **2003**, 5(10), 1597-1600.
174. Vincent, P., Beaucourt, J.-P., Pichat, L., *Tetrahedron Lett.*, **1981**, 22, 94947.
175. Stille, J.K., Simpson, J.H., *J. Am. Chem. Soc.*, **1987**, 109, 2138-2152.
176. Phillips, A.P., *J. Am. Chem. Soc.*, **1951**, 73(3), 1061-1062.
177. Gütschow, M., Hecker, T., Thiele, A., Hauschildt, S., Eger, K., *Bioorg. Med. Chem.*, **2001**, 9(4), 1059-1065.
178. Lin, T.-S., Guo, J.-Y., Schinazi, R.F., Chu, C.K., Xiang, J.-N., Prusoff, W.H., *J. Med. Chem.*, **1988**, 31(2), 336-340.
179. Storek, M.J., Suciu, A., Verdine, G.L., *Org. Lett.*, **2002**, 4(22), 3867-3869.
180. Azam, A.T.M.Z., Moriguchi, T., Shinozuka, K., *Chem. Commun.*, **2006**, (3), 335-337.
181. Ben-Gaied, N., Zhao, Z., Gerrard, S.R., Fox, K.R., Brown, T., *ChemBioChem*, **2009**, *in press.*
182. Asseline, U., Chassignol, M., Draus, J., Durand, M., Maurizot, J.-C., *Bioorg. Med. Chem.*, **2003**, 11(16), 3499-3511.
183. Yang, W., Giandomenico, C.M., Sartori, M., Moore, D.A., *Tetrahedron Lett.*, **2003**, 44(12), 2481-2483.
184. Brederick, H., Gottfried, K., Wieland, H., *Chem. Ber.*, **1959**, 92(3), 583-594.
185. Schroeder, E.F., G. D. Searle & Co. *5-Halo-6-Amino-Uracils and Derivatives Thereof.* US 2731465, 17th January 1956.
186. Rybár, A., Hesek, D., Szemes, F., Alföldi, J., *Collect. Czech. Chem. Commun.*, **1990**, 55(9), 2257-2269.
187. Zhang, L., Zhang, Y.J., *Tetrahedron Lett.*, **2006**, 47(5), 775-778.
188. Kupchan, S.M., Bondesson, G., Davies, A.P., *J. Med. Chem.*, **1972**, 15(1), 65-68.
189. Taylor, E.C., Loux, H.M., Falco, E.A., Hitchings, G.H., *J. Am. Chem. Soc.*, **1955**, 77(8), 2243 - 2248.
190. Fieser, L.F., Brown, R.H., *J. Am. Chem. Soc.*, **1949**, 71(11), 3609-3614.
191. Hansen, B., *Acta Chem. Scand.*, **1957**, 11, 537-540.

192. O'Sullivan, M.C., Dalrymple, D.M., *Tetrahedron Lett.*, **1995**, 36(20), 3451-3452.
193. Cheng, C.-C., Lu, Y.-L., *Chem. Commun.*, **1998**, 253-254.
194. Koščová, S., Budějšínský, M., Hodačová, J., *Collect. Czech. Chem. Commun.*, **2003**, 68(4), 744-750.
195. Lee, Y., Sarjeant, A.A.N., Karlin, K.D., *Chem. Commun.*, **2006**, 621-623.
196. Zoller, U., Three-Membered Rings Containing Sulfur, in *Small Ring Heterocycles*, Chemistry of Heterocyclic Compounds: A Series of Monographs, ed. A. Hassner, John Wiley & Sons **1983**, Vol. 42, Part 1, Ch. III, pp 333-629.
197. Choi, J., Yoon, N.M., *Synth. Commun.*, **1995**, 25(17), 2655-2663.
198. Zheng, T.-C., Burkart, M., Richardson, D.E., *Tetrahedron Lett.*, **1999**, 40(4), 603-606.
199. Cummins, L.L., Owens, S.R., Risen, L.M., Lesnik, E.A., Freier, S.M., McGee, D., Guinosso, C.J., Cook, P.D., *Nucleic Acids Res.*, **1995**, 23(11), 2019-2024.
200. Legorburu, U., Reese, C.B., Song, Q., *Tetrahedron*, **1999**, 55(17), 5635-5640.
201. Reese, C.B., Song, Q. *2'-Substituted RNA Preparation*. WO 00/56747, 28th September 2000.
202. Manfredini, S., Baraldi, P.G., Durini, E., Porcu, L., Angusti, A., Vertuani, S., Solaroli, N., De Clercq, E., Karlsson, A., Balzarini, J., *Bioorg. Med. Chem. Lett.*, **2001**, 11(10), 1329-1332.
203. Ross, B.S., Springer, R.H., Tortorici, Z., Dimock, S., *Nucleos. Nucleot. Nucl.*, **1997**, 16(7-9), 1641-1644.
204. Reichman, U., Chu, c.K., Hollenberg, K.A., Watanabe, K.A., Fox, J.J., *Synthesis*, **1976**, (8), 533-534.
205. McGee, D.P.C., Zhai, Y., *Nucleos. Nucleot. Nucl.*, **1996**, 15(11-12), 1797-1803.
206. Vadhia, S.J., The Effects of Modified Nucleosides on DNA Duplexes and Triplexes, PhD Thesis, University of Southampton, Southampton, UK, **2007**.
207. Gerrard, S.R., Srinivasan, N., Fox, K.R., Brown, T., *Nucleos. Nucleot. Nucl.*, **2007**, 26(10-12), 1363-1367.
208. Melissaris, A.P., Litt, M.H., *J. Org. Chem.*, **1994**, 59(19), 5818-5821.
209. Classon, B., Samuelsson, B., *Acta Chem. Scand. Ser. B*, **1985**, 39(6), 501-504.
210. Kabalka, G.W., Wang, L., M., P.R., *Tetrahedron*, **2001**, 57(38), 8017-8028.
211. Arnold, F.E., Reinhardt, B.A., Hedberg, F.L. *Aromatic enyne compounds and their synthesis*. US 4162265, 24th July 1979.

212. Herdewijn, P., Balzarini, J., Baba, M., Pauwels, R., Aerschot, A.V., Janssen, G., Clercq, E.D., *J. Med. Chem.*, **1988**, 31(10), 2040-2048.

213. Divakar, K.J., Sawant, C.M., Mulla, Y.A., Zemse, D.V., Sitabkhan, S.M., Ross, B.S., Sanghvi, Y.S., *Nucleos. Nucleot. Nucl.*, **2003**, 22(5-8), 1321-1325.

214. Lyttle, M.H., Walton, T.A., Dick, D.J., Carter, T.G., Beckman, J.H., Cook, R.M., *Bioconjugate Chem.*, **2002**, 13(5), 1146-1154.

215. San-Félix, A., Velázquez, S., Pérez-Pérez, M.J., Balzarini, J., Clercq, E.D., Camarasa, M.J., *J. Med. Chem.*, **1994**, 37(4), 453-460.

216. Miah, A., Reese, C.B., Song, Q., *Nucleosides Nucleotides*, **1997**, 16(1&2), 53-65.

217. Jin, Y.H., Liu, P., Wang, J., Baker, R., Huggins, J., Chu, C.K., *J. Org. Chem.*, **2003**, 68(23), 9012-9018.

218. Song, G.Y., Paul, V., Choo, H., Morrey, J., Sidwell, R.W., Schinazi, R.F., Chu, C.K., *J. Med. Chem.*, **2001**, 44(23), 3985-3993.

219. Shi, J., Du, J., Ma, T., Pankiewicz, K.W., Patterson, S.E., Tharnish, P.M., McBrayer, T.R., Stuyver, L.J., Otto, M.J., Chu, C.K., Schinazic, R.F., Watanabe, K.A., *Bioorg. Med. Chem.*, **2005**, 13, 1641-1652.

220. Holmes, S.C., Gait, M.J., *Nucleos. Nucleot. Nucl.*, **2003**, 22(5-8), 1259-1262.

221. Nyilas, Á., Chattopadhyaya, J., *Acta Chem. Scand. Ser. B*, **1986**, 40(10), 826-830.

222. Inoue, H., Imura, A., Ohtsuka, E., *Nippon Kagaku Kaishi*, **1987**, (7), 1214-1220.

223. Inoue, H., Otsuka, E., Imura, A. *Preparation of fluorescent nucleosides or nucleotides and their use for DNA hybridization probes*. JP99-255499, 11th July 1987.

224. Nurminen, E.J., Mattinen, J.K., Lönnberg, H., *J. Chem. Soc., Perkin Trans. 2*, **2001**, (11), 2159–2165.

225. Prakash, T.P., Püschl, A., Manoharan, M., *Nucleos. Nucleot. Nucl.*, **2007**, 26(2), 149-159.

226. Segev, D. *System for delivering therapeutic agents into living cells and cell nuclei*. US 20060160763, 20th July 2006.

227. Saxena, A.K., Pandey, S.K., Seth, P., Singh, M.P., Dikshit, M., Carpy, A., *Bioorg. Med. Chem.*, **2001**, 9, 2025-2034.

228. Bernat, J., Kristian, P., Guspanova, J., Imrich, J., Busova, T., *Collect. Czech. Chem. Commun.*, **1997**, 62, 1491-1496.

229. Shibahara, F., Suenami, A., Yoshida, A., Murai, T., *Chem. Commun.*, **2007**, (23), 2354-2356.
230. Hudson, R.H.E., Dambenieks, A.K., *Heterocycles*, **2006**, 68(7), 1325-1328.
231. Hudson, R.H.E., Dambenieks, A.K., Viirre, R.D., *Synlett*, **2004**, (13), 2400-2402.
232. Hudson, R.H.E., Dambenieks, A.K., Viirre, R.D., *Nucleos. Nucleot. Nucl.*, **2005**, 24(5-7), 581-584.
233. Hudson, R.H.E., Ghorbani-Choghamarani, A., *Nucleos. Nucleot. Nucl.*, **2007**, 26(6&7), 533-537.
234. Hudson, R.H.E., Moszynski, J.M., *Synlett*, **2006**, (18), 2997-3000.
235. Wojciechowski, F., Hudson, R.H.E., *J. Am. Chem. Soc.*, **2008**, 130(38), 12574-12575.
236. Yoshimoto, K., Xu, C.-Y., Nishizawa, S., Haga, T., Satake, H., Teramae, N., *Chem. Commun.*, **2003**, (24), 2960-2961.
237. Van Daele, I., Bomholt, N., Filichev, V.V., Van Calenbergh, S., Pedersen, E.B., *ChemBioChem*, **2008**, 9, 791-801.
238. Caruthers, M.H., Beaucage, S.L., Unit 3.3, in *Current Protocols in Nucleic Acid Chemistry*, ed. S.L. Beaucage, D.E. Bergstrom, G.D. Glick, R.A. Jones, John Wiley & Sons Inc. **2002**, Vol. 1.
239. Beaucage, S.L., Caruthers, M.H., *Tetrahedron Lett.*, **1981**, 22(20), 1859-1862.
240. Matteucci, M., Caruthers, M.H., *J. Am. Chem. Soc.*, **1981**, 103(11), 3185-3191.
241. Albericio, F., Kates, S.A. *Automated allyl deprotection in solid-phase synthesis*. US 5777077, 7th July 1998.
242. Albericio, F., Kates, S.A. *Automated allyl deprotection in solid-phase synthesis*. US 5861351, 19th January 1999.
243. Hsiung, H.M. *Process for de-cyanoethylating blocked nucleotides*. US 4419509, 30th March 1983.
244. Hsiung, H.M. *Process for de-cyanoethylating blocked nucleotides*. US 4426517, 24th August 1984.
245. Umemoto, T., Wada, T., *Tetrahedron Lett.*, **2005**, 46(24), 4251-4253
246. Bhan, A., Holmberg, L. *Method for deprotecting oligonucleotides*. US 6887990, 3rd May 2005.
247. Gottlieb, H.E., Kotlyar, V., Nudelman, A., *J. Org. Chem.*, **1997**, 62(21), 7512-7515.
248. Winkley, M.W., Robins, R.K., *J. Chem. Soc. C*, **1969**, (5), 791 - 796.

249. Guthrie, R.D., Smith, S.C., *Biochem. Prep.*, **1971**, 13, 1-3.

250. Kam, B.L., Barascut, J.-L., Imbach, J.-L., *Carbohydr. Res.*, **1979**, 69(1), 135-142.

251. Pike, A.R., Ryder, L.C., Horrocks, B.R., Clegg, W., Elsegood, M.R.J., Connolly, B.A., Houlton, A., *Chem. Eur. J.*, **2002**, 8(13), 2891-2899.

252. Hay, M.P., Pruijn, F.B., Gamage, S.A., Liyanage, H.D.S., Kovacs, M.S., Patterson, A.V., Wilson, W.R., Brown, J.M., Denny, W.A., *J. Med. Chem.*, **2004**, 47(2), 475-488.

253. Bleicher, L.S., Cosford, N.D.P., Herbaut, A., McCallum, J.S., McDonald, I.A., *J. Org. Chem.*, **1998**, 63(4), 1109-1118.

254. Bosch, E., Jeffries, L., *Tetrahedron Lett.*, **2001**, 42(46), 8141–8142.

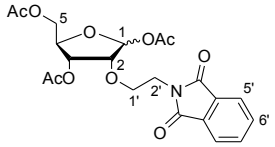
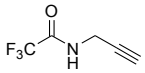
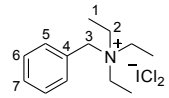
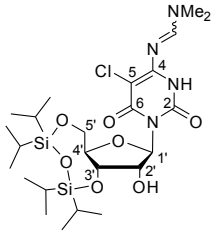
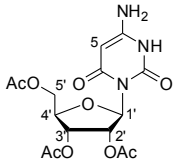
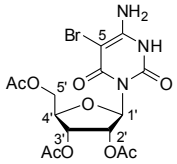
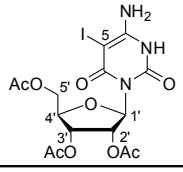
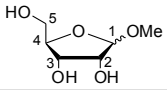
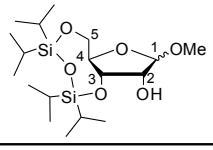
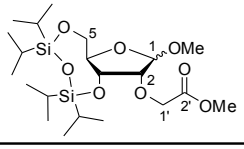
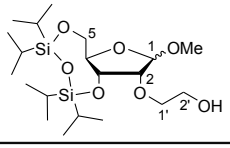
255. Battersby, T.R., Ang, D.N., Burgstaller, P., Jurczyk, S.C., Bowser, M.T., Buchanan, D.D., Kennedy, R.T., Benner, S.A., *J. Am. Chem. Soc.*, **1999**, 121(42), 9781-9789.

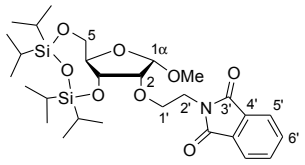
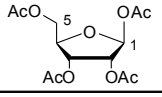
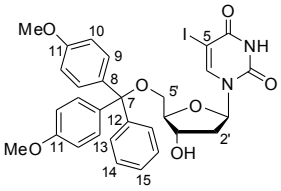
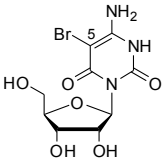
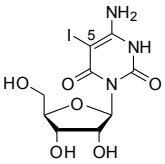
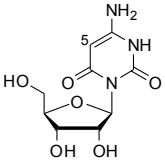
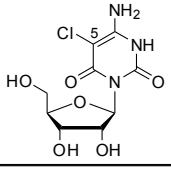
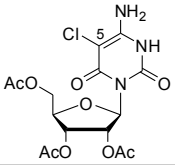
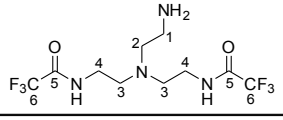
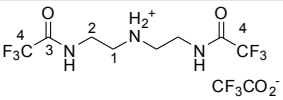
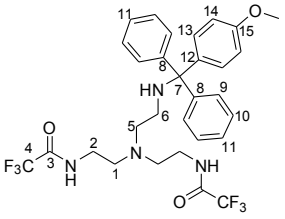
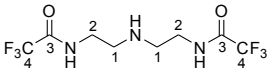
256. Langley, G.J., Herniman, J.M., Davies, N.L., Brown, T., *Rapid Commun. Mass Spectrom.*, **1999**, 13(17), 1717-1723.

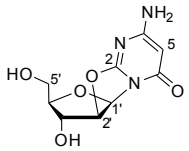
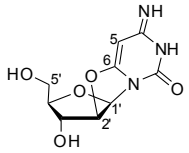
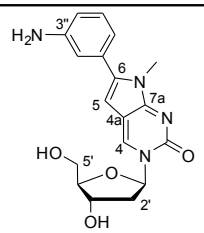
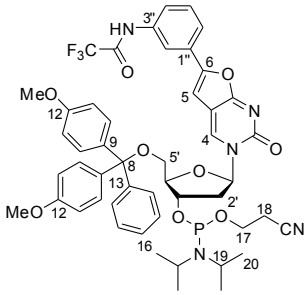
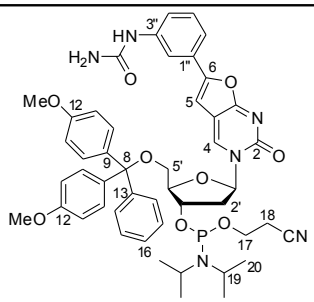
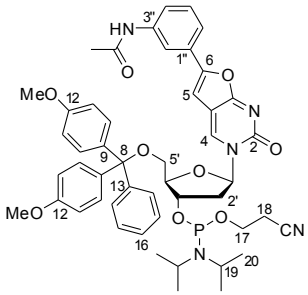
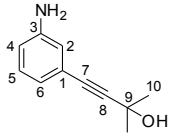
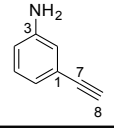
257. Sambrook, J., Fritsch, E.F., Maniatis, T., *Molecular Cloning: a laboratory manual, Vol. 3*, 2nd ed., Cold Spring Harbor Laboratory Press, New York, U.S.A., **1989**.

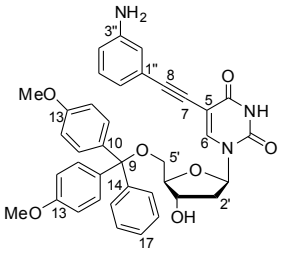
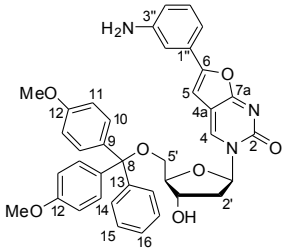
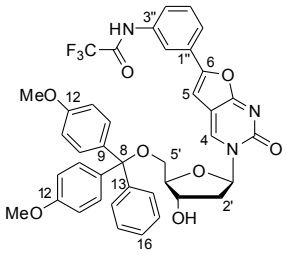
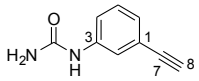
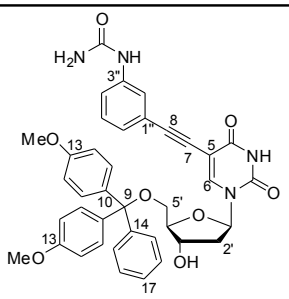
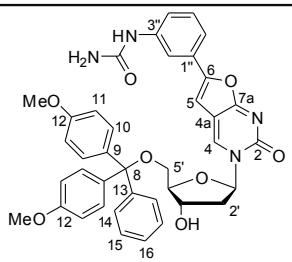
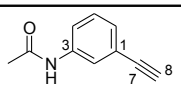
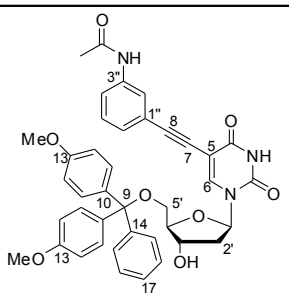
Appendices

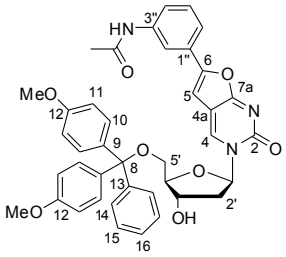
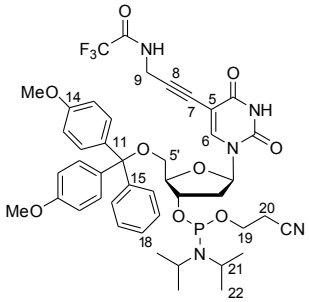
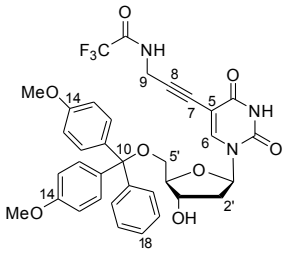
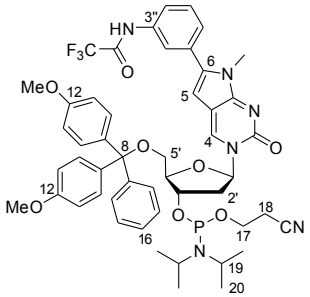
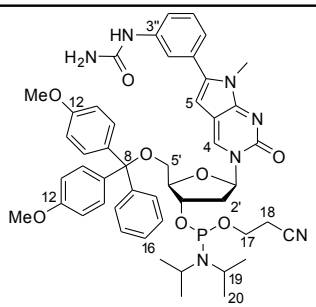
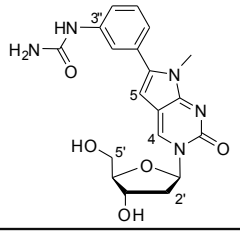
A: List of Synthesised Compounds

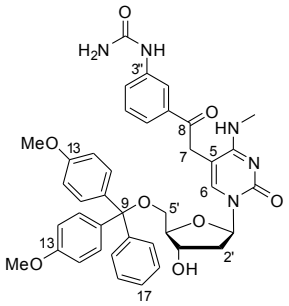
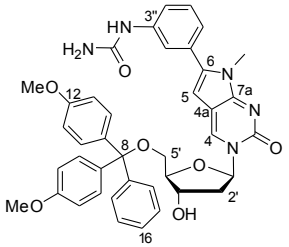
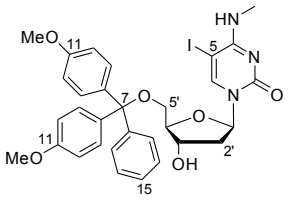
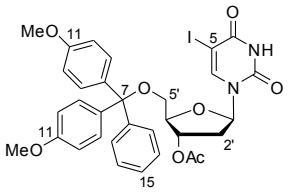
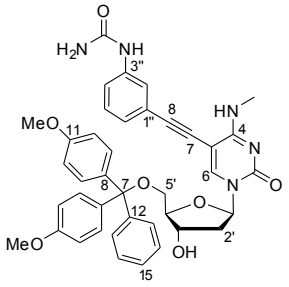
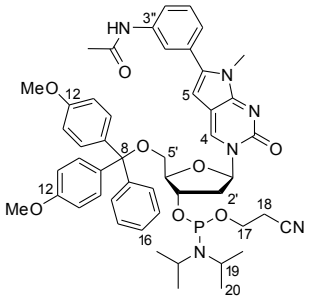
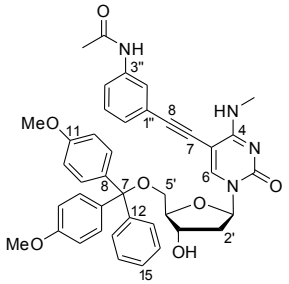
Cmpd No.	Structure	Name
65		1,3,5-tri-*O*-acetyl-2-*O*-(2-phthalimidoethyl)-D-ribofuranose
68		*N*-trifluoroacetyl propargylamine
71		benzyltriethylammonium dichloroiodate
74		N^1-[1-(3',5'-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-β-D-ribofuranosyl)]-(5-chloro-6-oxocytosin-4-yl)-*N,N*-dimethylformamidine
75		2',3',5'-tri-*O*-acetyl-6-oxocytidine
79		5-bromo-2',3',5'-tri-*O*-acetyl-6-oxocytidine
83		5-iodo-2',3',5'-tri-*O*-acetyl-6-oxocytidine
85		1-*O*-methyl-D-ribofuranose
86		1-*O*-methyl-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose
87		1-*O*-methyl-2-*O*-methylethanoyl-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose
88		1-*O*-methyl-2-*O*-(2-hydroxyethyl)-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose

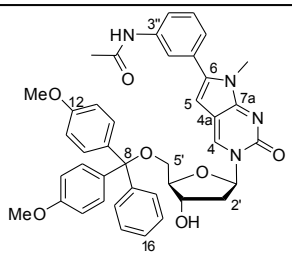
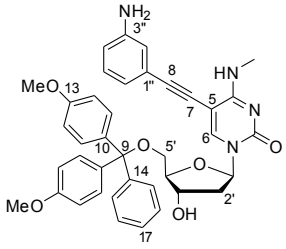
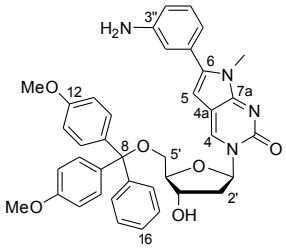
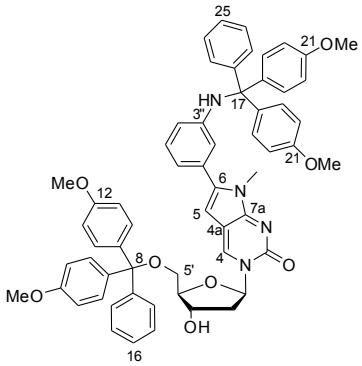
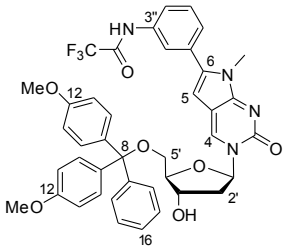
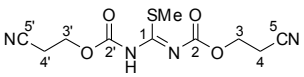
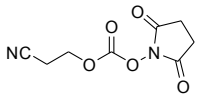
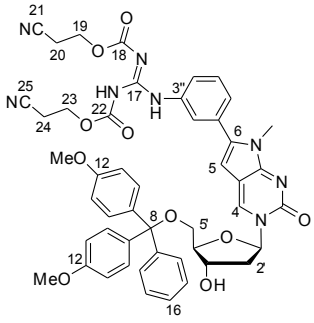
89		1-*O*-methyl-2-*O*-(2-phthalimidoethyl)-3,5-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-α-D-ribofuranose
91		1,2,3,5-tetra-*O*-acetyl-β-D-ribofuranose
97		5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine
99		5-bromo-6-oxocytidine
100		5-iodo-6-oxocytidine
101		6-oxocytidine
108		5-chloro-6-oxocytidine
109		5-chloro-2',3',5'-tri-*O*-acetyl-6-oxocytidine
110		*N*-(2-aminoethyl)-*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine
111		*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]ammonium trifluoroacetate
116		*N*-[2-*N'*-(4-methoxytritylamino)ethyl]-*bis*-*N*,*N*-(2-ethyl trifluoroacetamide)
128		*bis*-*N*,*N*-[2-(trifluoroacetamido)ethyl]amine

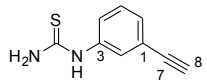
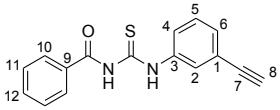
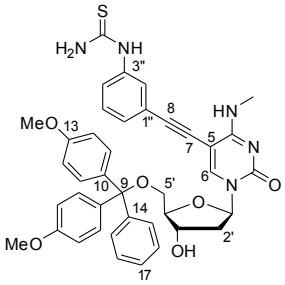
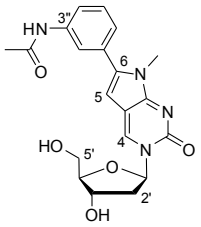
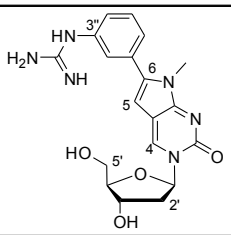
141		2,2'-anhydro-6-oxocytidine
142		6,2'-anhydro-6-oxocytidine
143		3-*N*-[2'-deoxy-β-D-ribofuranosyl]-6-(3-aminophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
145		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-[3-(trifluoroacetamido)phenyl]}-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite
146		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite
147		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-*O*-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite
151		4-(3-aminophenyl)-2-methyl-but-3-yn-2-ol
152		3-aminophenyl acetylene, 3-ethynylaniline

153		5-[(3-aminophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine
154		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-aminophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one
155		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-[3-(trifluoroacetamido)phenyl]}-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one
156		3-ureidophenyl acetylene
159		5-[(3-ureidophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine
160		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one
161		3-acetamidophenyl acetylene
162		5-[(3-acetamidophenyl)ethynyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyuridine

163		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-furano[2,3-*d*]pyrimidine-2-one
164		5'-*O*-(4,4'-dimethoxytrityl)-5-(3-trifluoroacetamidopropynyl)-2'-deoxyuridine-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite
165		5'-*O*-(4,4'-dimethoxytrityl)-5-(3-trifluoroacetamidopropynyl)-2'-deoxyuridine
178		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-trifluoroacetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite
179		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite
182		3-*N*-[2'-deoxy-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one

183		5'-*O*-(4,4'-dimethoxytrityl)-5-[(3-ureidobenzoyl)methyl]-4*N*-methyl-2'-deoxycytidine
184		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-ureidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
185		5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-4*N*-methyl-2'-deoxycytidine
187		3'-*O*-acetyl-5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine
188		5'-*O*-(4,4'-dimethoxytrityl)-5-[(3-ureidophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine
191		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one-3'-*O*-(2-*O*-cyanoethyl-*N,N*-diisopropyl) phosphoramidite
192		5'-*O*-(4,4'-dimethoxytrityl)-5-[(3-acetamidophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine

193		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
194		5'-*O*-(4,4'-dimethoxytrityl)-5-[(3-aminophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine
195		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-aminophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
196		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-[3-(4,4'-dimethoxytrityl)aminophenyl]-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
197		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-trifluoroacetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
200		*N*,*N'-bis*-[(2-cyanoethoxy)carbonyl]-*S*-methyl-isothiourea
201		(2-cyanoethyl)-*N*-succinimidyl carbonate
207		3-*N*-[2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranosyl]-6-(3-{*N*,*N'-bis*-[(2-cyanoethoxy)carbonyl]guanidinyl}phenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one

209		3-thioureidophenyl acetylene
211		*N*-benzoyl-*N'*-(3-ethynylphenyl)-thiourea
213		5'-*O*-(4,4'-dimethoxytrityl)-5-[(3-thioureidophenyl)ethynyl]-4*N*-methyl-2'-deoxycytidine
225		3-*N*-[2'-deoxy-β-D-ribofuranosyl]-6-(3-acetamidophenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one
226		3-*N*-[2'-deoxy-β-D-ribofuranosyl]-6-(3-guanidinylphenyl)-(2,3*H*)-*N*-methylpyrrolo[2,3-*d*]pyrimidine-2(7*H*)-one

B: Oligonucleotide List

ODN No.	Sequence	Expected Mass	Expected $[M+H]^+$	Observed Mass	Notes
1	GCT AAA AAG A**C**A GAG AGA TCG	6521.34	-	*	
2	CGA TCT CTC T**G**T CTT TTT AGC	6329.17	-	*	
3	TTT TTM T^{M}**P**$_{f}$T MTM TMT	4535.04	4536.05	4535.16	
4	TTT TTM T^{A}**PP**$_{f}$T MTM TMT	4612.13	4613.13	4611.98	
5	TTT TTM T^{U}**PP**$_{f}$T MTM TMT	4655.15	4656.16	4654.83 / 4697.25	partial acetylation
6	**F** GTG TTA GGA AGA **C**AA AAA AGA ACT GGT **HH**A CCA GTT CTT TTT T**G**T CTT CCT AAC AC	17843.96	-	*	CG hd
7	**F** GTG TTA GGA AGA **G**AA AAA AGA ACT GGT **HH**A CCA GTT CTT TTT T**C**T CTT CCT AAC AC	17843.96	-	*	GC hd
8	**F** GTG TTA GGA AGA **A**AA AAA AGA ACT GGT **HH**A CCA GTT CTT TTT T**T**T CTT CCT AAC AC	17842.97	-	*	AT hd
9	**F** GTG TTA GGA AGA **T**AA AAA AGA ACT GGT **HH**A CCA GTT CTT TTT T**A**T CTT CCT AAC AC	17842.97	-	*	TA hd
10	**Q** PMM TPM T^{M}**P**$_{f}$P TPT PTM PT	5807.10	5808.10	5807.41 / 5827.63	partially uncyclised
11	**Q** PMM TPM T^{A}**PP**$_{f}$P TPT PTM PT	5884.18	5885.19	5884.66 / 5902.71	partially uncyclised
12	**Q** PMM TPM T^{Ac}**PP**$_{f}$P TPT PTM PT	5926.22	5927.23	5926.67 / 5943.79	partially uncyclised
13	**Q** PMM TPM T^{U}**PP**$_{f}$P TPT PTM PT	5927.21	5928.21	5927.31 / 5945.78	partially uncyclised
14	**Q** PMM TPM T^{A}**PP**$_{f}$T TPT PTM PT	5845.14	-	5846.7	
15	**Q** PMM TPM P^{A}**PP**$_{f}$T TPT PTM PT	5884.18	-	5882.5	
16	**Q** PMM TPM P^{A}**PP**$_{f}$P TPT PTM PT	5923.17	-	5963.6	[M+2Na-2H]
17	**F** GTG TTA GGA AGA **C**AA AAA AGA ACT GGT **HH**A CCA GTT CTT TTT T**C**T CTT CCT AAC AC	17803.93	-	*	CC hd
18	**Q** PMM TPM T^{Ac}**PP**$_{p}$T TPT PTM PT	5887.04	-	5889.1	
19	**Q** PMM TPM P^{Ac}**PP**$_{p}$T TPT PTM PT	5926.11	-	5929.6	
20	**Q** PMM TPM T^{Ac}**PP**$_{p}$P TPT PTM PT	5926.11	-	5929.9	
21	**Q** PMM TPM P^{Ac}**PP**$_{p}$P TPT PTM PT	5965.14	-	5966.1	
22	**Q** PMM TPM T^{U}**PP**$_{p}$T TPT PTM PT	5888.07	-	5890.5	
23	**Q** PMM TPM P^{U}**PP**$_{p}$T TPT PTM PT	5927.10	-	5929.3	(+ 5875.4)
24	**Q** PMM TPM T^{U}**PP**$_{p}$P TPT PTM PT	5927.10	-	5928.8	
25	**Q** PMM TPM P^{U}**PP**$_{p}$P TPT PTM PT	5966.13	-	5968.3	
26	GCT AAA AAG A**G**A GAG AGA TCG	6561.37	-	*	
27	GCT AAA AAG A**A**A GAG AGA TCG	6545.37	-	*	
28	GCT AAA AAG A**T**A GAG AGA TCG	6536.36	-	*	
29	CGA TCT CTC T**C**T CTT TTT AGC	6289.14	-	*	
30	CGA TCT CTC T**T**T CTT TTT AGC	6304.15	-	*	
31	CGA TCT CTC T**A**T CTT TTT AGC	6313.17	-	*	
32	TTT TTM T**T**T MTM TMT	4496.99	-	4499.2	

33	TTT TTM T**T**P MTM TMT	4536.03	-	4541.0	(+ 4236.8)
34	TTT TTM P**T**T MTM TMT	4536.03	-	4537.4	
35	TTT TTM P**T**P MTM TMT	4575.07	-	4578.0	
36	TTT TTM T^{A}**PP**$_{p}$T MTM TMT	4611.14	-	4611.9	
37	TTT TTM T^{A}**PP**$_{p}$P MTM TMT	4650.18	-	4652.6	
38	TTT TTM P^{A}**PP**$_{p}$T MTM TMT	4650.18	-	4651.8	
39	TTT TTM P^{A}**PP**$_{p}$P MTM TMT	4689.21	-	4690.0	
40	TTT TTM T^{Ac}**PP**$_{p}$T MTM TMT	4653.18	-	4653.8	
41	TTT TTM T^{Ac}**PP**$_{p}$P MTM TMT	4692.21	-	4696.4	
42	TTT TTM P^{Ac}**PP**$_{p}$T MTM TMT	4692.21	-	4693.2	
43	TTT TTM P^{Ac}**PP**$_{p}$P MTM TMT	4731.25	-	4733.1	
44	TTT TTM T^{U}**PP**$_{p}$T MTM TMT	4654.17	-	4655.3	
45	TTT TTM T^{U}**PP**$_{p}$P MTM TMT	4693.20	-	4696.6	
46	TTT TTM P^{U}**PP**$_{p}$T MTM TMT	4693.20	-	4697.0	
47	TTT TTM P^{U}**PP**$_{p}$P MTM TMT	4732.24	-	4734.1	
48	TTT TTM T^{G}**PP**$_{p}$T MTM TMT	4653.18	-	4653.9	
49	TTT TTM T^{G}**PP**$_{p}$P MTM TMT	4692.22	-	4694.4	
50	TTT TTM P^{G}**PP**$_{p}$T MTM TMT	4692.22	-	4694.1	
51	TTT TTM P^{G}**PP**$_{p}$P MTM TMT	4731.25	-	4732.2	
52	**Q** PMM TPM T^{A}**PP**$_{p}$T TPT PTM PT	5845.17	-	5849.2	
53	**Q** PMM TPM P^{A}**PP**$_{p}$T TPT PTM PT	5884.21	-	5884.8	
54	**Q** PMM TPM T^{A}**PP**$_{p}$P TPT PTM PT	5884.21	-	5886.0	
55	**Q** PMM TPM P^{A}**PP**$_{p}$P TPT PTM PT	5923.24	-	5924.2	
56	**Q** PMM TPM T^{G}**PP**$_{p}$T TPT PTM PT	5887.21	-	5887.8	
57	**Q** PMM TPM P^{G}**PP**$_{p}$T TPT PTM PT	5926.25	-	5928.6	
58	**Q** PMM TPM T^{G}**PP**$_{p}$P TPT PTM PT	5926.25	-	5926.1	
59	**Q** PMM TPM P^{G}**PP**$_{p}$P TPT PTM PT	5965.28	-	5967.5	
60	**Q** PMM TPM T**T**T TPT PTM PT	5731.00	-	5730.4	
61	**Q** PMM TPM P**T**T TPT PTM PT	5770.03	-	5771.0	
62	**Q** PMM TPM T**T**P TPT PTM PT	5770.03	-	5770.5	
63	**Q** PMM TPM P**T**P TPT PTM PT	5809.07	-	5809.7	
64	TTT TTT TTT TTT	3588.35	-	Oligonucleotides for Deprotection Study (Chapter 7)	
65	TTT TTT TT**P** TTT	3627.39	-		
66	TTT T**P**T **P**T**P** TTT	3705.46	-		
67	T**PP** **P**T**P** **P**T**P** T**P**T	3861.61	-		
68	TTT TTT TT^{BA}**U** TTT	3686.46	-		
69	TTT TTT TT^{Ac}**PP**$_{p}$ TTT	3744.54	-		
70	TTT TTT TT**S** TTT	3737.57	-		
71	TTT TTT TT^{MOE}**pU** TTT	3701.47	-		
72	TTT TTT TT^{AE}**T** TTT	3647.42	-		
73	TTT T^{AE}**T**T ^{AE}**T**T^{AE}**T** TTT	3765.56	-		
74	TTT TTT TT^{G}**P** TTT	3669.43	-		

Table. ES MS (1 dp), MALDI-TOF MS (2 dp), **Q** = DABCYL (quencher), **F** = FAM (fluorophore), **H** = HEG. Key nucleotides marked in bold. * No MS analysis necessary.

C: Conferences Attended, Posters, Presentations and Publication

July 2007	Year 3 Postgraduate Talks, School of Chemistry, University of Southampton, UK. – Presentation
June 2007	3rd Nucleic Acids Summer School, University of Southern Denmark, Odense, Denmark. – Presentation
March 2007	SCI Postgraduate Symposium, UCL, London, UK. – Presentation
November 2006	Year 2 Postgraduate Poster Day, School of Chemistry, University of Southampton, UK. – Poster
September 2006	XVII International Roundtable Conference on Nucleosides, Nucleotides and Nucleic Acids, Bern, Switzerland. – Poster

– Gerrard, S.R., Srinivasan N., Fox, K.R., Brown, T. 'CG Base Pair Recognition within DNA Triple Helices using *N*-Methyl-3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one Nucleoside Analogues.' *Nucleos. Nucleot. Nucl.* **2007**, 26(10-12), 1363-1367.
DOI: 10.1080/15257770701533958

– Ben Gaied, N., Zhao, Z., Gerrard, S.R., Fox, K.R., Brown, T. 'Potent Triple Helix Stabilisation by 3' and 5', 3'-Modified Triplex Forming Oligonucleotides.' *ChemBioChem* **2009**, *in press.*
DOI: 10.1002/cbic.200900232