Synthesis and Study of Neutral Backbones for Therapeutic Antisense Oligonucleotides

Lillian Lie

Corpus Christi College

University of Oxford

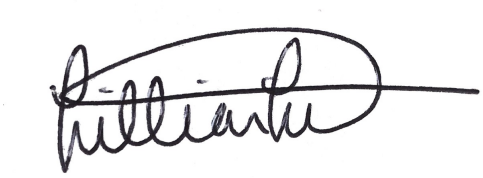
A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

Declaration

I, Lillian Lie, declare that this thesis and all work herein, carried out from October 2020 until July 2024, is entirely my own work, except where indicated in the text and has not been submitted previously for any other degree at this or any other university.



Lillian Lie

11$^{\text{th}}$ July, 2024

Acknowledgements

First and foremost I would like to acknowledge Prof. Tom Brown for providing me with the opportunity to conduct my DPhil in his lab, a wonderful group of co-workers, and his supervision. Of course, I would also like to thank Prof. Afaf El-Sagheer for her guidance as well as her aid in synthesising the oligonucleotides on the trusty (yet occasionally temperamental) ABI. Many of the faculty of the Department of Chemistry, as well as other departments within Oxford, have provided me with the necessary advice and resources to complete my research. Among them, Dr. Catheryn Lim was instrumental in training myself and others on cell culture and the biological assays, and I thank her heartily for always providing me with her time and attention. Dr. Marjorie Fournier was vital in conducting all proteomics work; I thank her for offering her time and expertise on advanced mass spectrometry. The cell work was conducted in facilities run by Prof. Matthew Wood at the Institute of Developmental and Regenerative Medicine and Prof. Angela Russell in the Department of Pharmacology and I thank them both for providing these resources.

I want to thank Dr. Alice Kennett from the bottom of my heart − this thesis would not nearly be as coherent, my scientific skills nearly as competent, nor my experience nearly as joyful without her continual input, emotional support, and true friendship. I thank Martin Flerin for conducting fantastic MD work and being a friend in addition to a key contributor in our collaboration. Within the Tom Brown group, I thank Ben Lagan for his companionship as we embarked on similar project paths, as well as Sebastian, Diallo, Chunsen, Jerry, and all other Brownies for making the office and lab a place full of friendly faces.

I thank Dr. Tristan Johnston-Wood; meeting you changed my whole experience of the PTCL, my DPhil, and Oxford. I owe much of what gave my every-day joy − running, T-room coffee breaks, and various high table dinners − to our

friendship. I thank the incredibly long list of friends who supported me from afar during this difficult life experience − Bronte, Lauren, Nick, and Kate. Max, that first year was not easy and I thank you for your love and the friendship that we've sustained beyond that year. Sid, of the things my time at 1910 gave me, our friendship was by far the most important. Ryan, thank you for being a friend through all those wonderful (but painful) SLRs. Dani, thank you for supporting me through the last few years and surely, the future − whether it was at horse box, in Somerset, in Münich, or wherever we meet next, it feels like sisterhood.

The staff and students at Corpus Christi College provided me with a wonderful community and made what could have been a lonely and difficult time easier − incredible food, beautiful housing, financial support, and most of all, the feeling that I always had a greater family to rely upon. Peter and Matt, thank you for being like brothers − I am sure our little Corpus family will extend well beyond the quad. Rowley, thank you for our sudden yet powerful friendship, sitting with me in the LLL pew while I frantically wrote, and making those last days in Oxford hard to say goodbye to.

I thank the Bickford family for providing me with a second family and a home-away-from home. And I thank Mr. Galinski − when people ask "how did you come to love chemistry?", a great teacher was always the answer.

I thank Tenzin Charles Rosson for coming into my life with such reckless enthusiasm. Single-handedly, your unwavering belief healed the parts of me that were broken in time for me to see this through.

Und bei meiner Familie − meine Eltern und meine drei Schwestern − bedanke ich mich für die Unterstützung, das Verständnis und die Liebe, die ihr mir bedingungslos geschenkt haben. Trotz der Ferne war ich nie alleine.

⁂

List of Publications

This thesis is partially based on the following publications or manuscripts prepared:

Chapter 2

- Baker, Y.R., Thorpe, C., Chen, J., Poller L., Cox, L., Kumar, P., Lie, L., McClorey, G., Epple, S., Singleton, D., McDonough, M., Hardwick, J., Christensen, K., Wood, M. J. A., Hall, J. P., El-Sagheer, A., Brown, T. An LNA-amide modification that enhances the cell uptake and activity of phosphorothioate exon-skipping oligonucleotides. *Nat Communications* **13**, 4036 (2022).

  Contribution: Synthesis of small molecule monomers.

Chapter 3

- Lie, L., Kennett, A., Flerin, M., Zengin-Kurt, B., Baker, Y., El-Sagheer, A., Duarte, F., Brown, T. Synthesis, biological activity, and molecular dynamics simulations of LNA-neutral linkages for enhanced splice-switching antisense oligonucleotides. **Manuscript to be submitted**.

  Contribution: Conceptualisation, small molecule synthesis, oligonucleotide synthesis, biophysical experiments, biological experiments, writing.

⁂

Declaration of Collaborations

This thesis contains work conducted in collaboration with other researchers. Specific contributions are outlined below for each relevant chapter:

- **Chapter 2**: Amide coupling optimisation was conducted together with Dr. Alice Kennett.
- **Chapter 3**: Molecular dynamics simulations were conducted by Martin Flerin. Two dinucleotide phosphoramidites were synthesised by Dr. Belma Zengin-Kurt and Dr. Alice Kennett, respectively. Biophysical and biological experiments of the amide-inspired backbones were carried out together with Dr. Alice Kennett.
- **Chapter 5**: Proteomics work was conducted by Dr. Marjorie Fournier and her team at the Advanced Proteomics Facility in the Department of Biochemistry.

While this statement provides an initial overview, additional details regarding individual contributions are also clearly indicated within the relevant sections of each chapter.

Abstract

Antisense oligonucleotides (ASOs) are a therapeutic modality which represent a transformative approach to drug development. By altering RNA processing through Watson-Crick base-pairing and therefore modulating down-stream protein expression, ASOs have provided heretofore undruggable diseases such as Duchenne muscular dystrophy with first-of-its-kind treatment. However, many ASOs rely solely upon chemical modifications to attain drug-like properties. This doctoral thesis focuses on expanding the chemical space of ASOs by combining neutral backbones and the locked nucleic acid (LNA) ribose modification to enhance their drug-like properties and therapeutic potential.

The work detailed in this thesis includes the synthesis of various dinucleotide and monomer phosphoramidites featuring neutral linkages, such as the amide (Chapter 2), carbamate (Chapter 3), and alkyl triester (Chapter 4) in combination with LNA. Biophysical studies such as thermal melting stability and circular dichroism reveal that linkages which are both flexible and good bioisosteres of the native phosphodiester adopt stable ASO:RNA heteroduplexes. Splice-switching was chosen as the therapeutic mechanism of action to evaluate these chemical modifications, as these ASOs have the capability to restore or up-regulate protein expression which is historically difficult to achieve using traditional small molecule discovery. The LNA-amide and LNA-isopropyl triester linkages were found to be more active than the clinically-validated 2'OMe ribose and phosphorothioate (PS) backbone modifications, indicating their potential in enhancing the drug-like properties of ASOs. Finally, the protein-binding profiles of the most promising ASOs were investigated (Chapter 5), offering preliminary insights into their enhanced activity. This thesis hopes that by developing and investigating LNA-neutral backbones, it may contribute to the medicinal chemistry of ASOs, an area in which continual innovation may improve their clinical success.

"DNA neither cares nor knows. DNA just is. And we dance to its music."

– Richard Dawkins

To my family, with whom I share DNA, and to Duke, who shared less but was still family. To all the women in science who created a world in which mine hasn't held me back. And to my younger self, who asked questions like "what is matter?" and who set out for answers.

⁂

Contents

List of Abbreviations

Abbreviation	Definition
A	adenine
AAV	adeno-associated virus
Ac	acetyl
ACN	acetonitrile
ACTB	β-actin
ADME	adsorption, distribution, metabolism, and excretion
ADR	adverse drug reactions
ASGPR	asialoglycoprotein receptor
ASO	antisense oligonucleotide
BCA	bicinchoninic acid
BMD	Becker muscular dystrophy
Bn	benzyl
bp	base pair
BSA	bis(trimethylsilyl)acetamide
BTT	5-benzylthio-1H-tetrazole
Bz	benzoyl
°C	degrees Celsius
^{Me}C	5-methylcytosine
C	cytosine
calcd.	calculated
CD	circular dichroism

Abbreviation	Definition
CDI	1,1'-carbonyldiimidazole
cEt	constrained ethyl
CISY	citrate synthase
CPG	controlled-pore glass
CPP	cell-penetrating peptide
C_q	quantification cycle
CRISPR	clustered regularly interspaced short palindromic repeats
CSF	cerebrospinal fluid
Da	Dalton
DAP	dimethylaminopropyl
DCC	*N, N*-dicyclohexylcarbodiimide
DDQ	2,3-dichloro-5,6-dicyano-benzoquinone
DDX-5	DEAD box protein 5
DGC	dystrophin glycoprotein complex
DMD	Duchenne muscular dystrophy
DMAP	4-dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle Medium
dmf	dimethylformamidine
DMF	dimethylformamide
DMFDMA	*N, N*-dimethylformamide dimethyacetal
DMP	Dess-Martin periodinane
DMTr	dimethoxytrityl
DNA	deoxyribonucleic acid

Abbreviation	Definition
dsDNA	double-stranded DNA
EDA	ethylenediamine
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDITH	3-ethoxy-1,2,4-dithiazoline-5-one
EE	early endosome
eq.	equivalents
ESI	electrospray ionisation
FBS	foetal bovine serum
FDA	Food and Drug Administration (USA)
FUBP1	far upstream element binding protein 1
g	gram
G	guanine
G3BP1	Ras GTPase-activating protein-binding protein 1
kb	kilobase
kDa	kilodalton
kg	kilogram
h	hour(s)
HATU	*O*-(7-Azabenzotriazol-l-yl)-*N,N,N,N*-tetramethyluronium hexafluorophosphate
HFIP	hexafluoroisopropanol
HIV	human immunodeficiency virus
HMG (A1, B1, B2)	high-mobility group proteins
hnRNP	heterogeneous ribonucleoproteins

Abbreviation	Definition
HRMS	high-resolution mass spectrometry
HPLC	high-performance liquid chromatography
Hz	Hertz
iBu	isobutyryl
iPr	isopropyl
LC-MS	liquid chromatography-mass spectrometry
LE	late endosome
LFQ	label-free quantification
LNA	locked nucleic acid
LNP	lipid nano-particle
LRMS	low-resolution mass spectrometry
M	molar
M6PR	mannose-6-phosphate receptor
MD	molecular dynamics
µM	micromolar
mg	milligram
min	minute(s)
mL	millilitre
MMI	methylene methylimino
MMTr	methoxytrityl
MOE	methoxyethyl
mol	mole
MOP	methoxypropyl

Abbreviation	**Definition**
MPTE	methyl phosphate triester
mRNA	mature RNA
MS	mass spectrometry
MS/MS	tandem mass spectrometry
m/z	mass-to-charge ratio
Nap	naphthalene
NB	*nota bene*
NHS	*N*-hydroxysuccinimide
nm	nanometer
nM	nanomolar
nmol	nanomole
NMR	nuclear magnetic resonance
NSAF	normalised spectral abundance factor
nt	nucleotide
NUCL	nucleolin
PAM	protospace-adjacent motif
PBS	phosphate-buffered saline
(q)PCR	(quantitative) polymerase chain reaction
PMO	phosphorodiamidate morpholino
PNA	peptide nucleic acid
ppm	parts per million
PS	phosphorothioate
PSF	PTB-associated splicing factor

Abbreviation	Definition
PSM	peptide spectral match
PyBOP	benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
R_f	retention factor
RAN	ras-related nuclear protein
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RP-HPLC	reverse-phase high-performance liquid chromatography
SAF	spectral abundance factor
SAR	structure-activity relationship
SARNP	SAP domain-containing ribonucleoprotein
sat.	saturated
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SHMT2	serine hydroxymethyltransferase
siRNA	short interfering RNA
SMA	spinal muscular atrophy
SNA	spherical nucleic acid
SPOS	solid-phase oligonucleotide synthesis
SPS	solvent purification system
SRSF1	serine and arginine rich splicing factor 1
SSBP	single-strand DNA-binding protein
ssDNA	single-stranded DNA

Abbreviation	Definition
T	thymine
T_{m}	melting temperature
TBAF	tert-butylammonium fluoride
TBDPS	tert-butyldiphenylsilyl
TCA	trichloroacetic acid
TEAA	triethylammonium acetate
THF	tetrahydrofuran
THP	tetrahydropyran
TLC	thin-layer chromatography
TMS(OTf)	trimethylsilyl (trifluoromethanesulfonate)
TOTU	*O*-[(Ethoxycarbonyl)cyanomethylenamino]-*N,N,N',N'*-tetramethyluronium tetrafluoroborate
U	uracil
UV-vis	ultra-violet visible
v/v	volume/volume
w/v	weight/volume
RMSD	root mean squared deviation
rt	room temperature

1

Introduction

Contents

1.1 Nucleic acids: biological function, structure, and synthesis

1.1.1 The central dogma

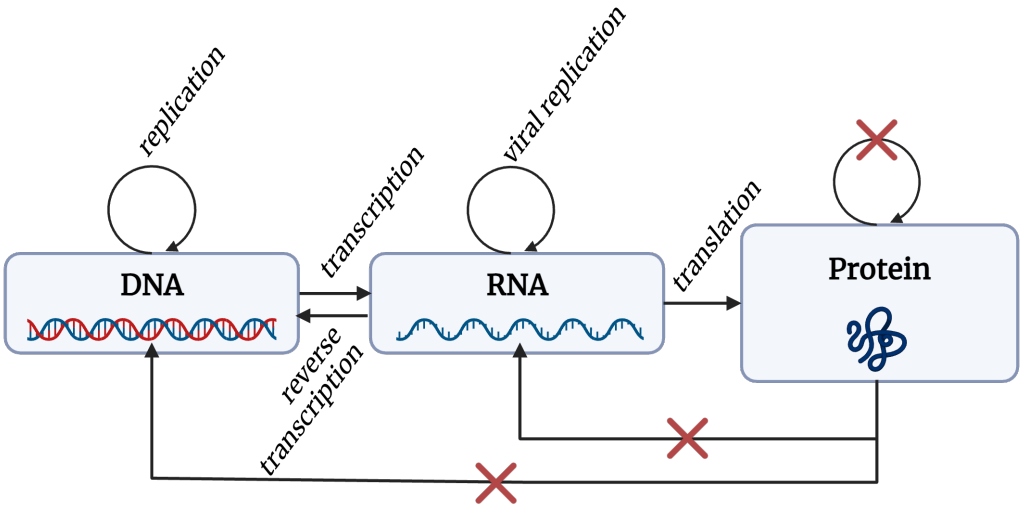


Figure 1.1: The flow of genetic information proposed by the central dogma. Figure adapted from Crick, 1958.[1]

"[James] Watson said to me, a few years ago, 'the most significant thing about nucleic acids is that we don't know what they do'" Francis Crick writes in his seminal publication "On Protein Synthesis" in 1958, published after his lecture by the same name.[1] In this important work, Francis Crick proposes a simple relationship between genetic information and protein synthesis that was heretofore a matter of great intrigue. Termed by him as the "central dogma" and depicted in Figure 1.1, he proposed that "the main function of genetic material is to control (not necessarily directly) the synthesis of proteins." Information flows in a singular direction, from deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) to proteins – this is now taught in schools all around the world and forms the bedrock of molecular and chemical biology, genetics, biochemistry, (certainly this thesis) and much more. Almost 70 years later, we know a considerable amount more about this "flow of information" – the mechanisms of transcription of DNA to RNA and translation of RNA to proteins are

well-elucidated.[a] A detailed account of all cellular processes involved is outside the scope of this thesis, but the most relevant and important aspects are briefly highlighted here.

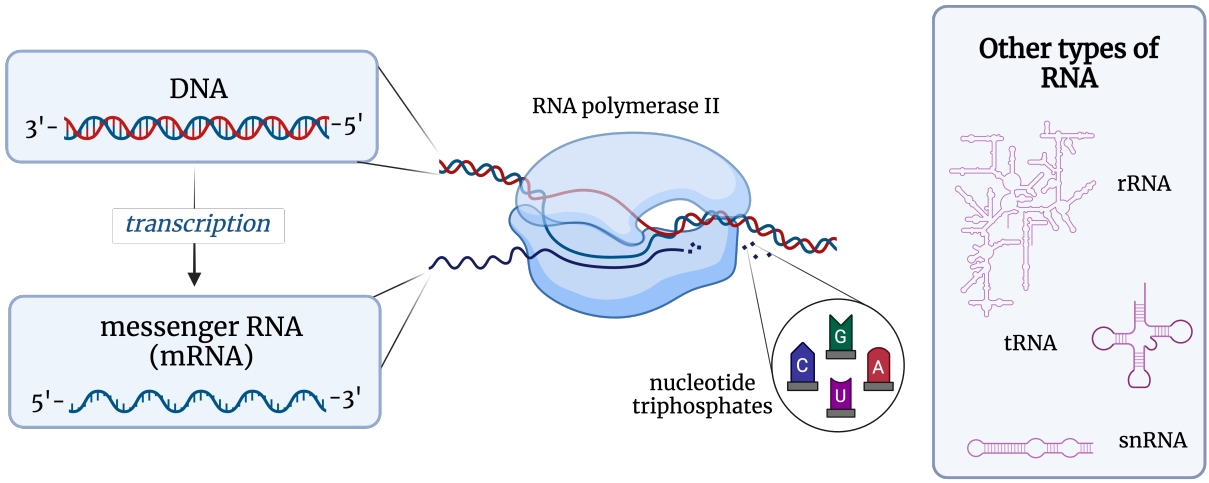


Figure 1.2: A closer look at transcription: genomic DNA is read by RNA polymerase II (in eukaryotes) in the 5'→3' direction, recruiting nucleotide triphosphates and producing messenger RNA in the 5'→3' direction. Other types of RNA carry out cellular functions which are indirectly involved in the flow of information.

Protein synthesis begins with the transfer of information in genomic DNA to RNA; interestingly, messenger RNA (mRNA) was not discovered until 1961 as being the "unstable intermediate" between genes and ribosomes.[3] Because both biopolymers use the same language of nucleotide monomers, this transfer from double-stranded DNA to intermediary single-stranded RNA is termed transcription. Summarised in Figure 1.2, genomic DNA is read by an RNA polymerase enzyme in a 5'→3' direction. At the site of transcription, the DNA double helix is unwound by the polymerase, which recruits the nucleoside triphosphates complementary to the template strand and catalyses the formation of the phosphodiester bonds, producing the growing single-stranded RNA strand.[4] This forms the precursor mRNA (pre-mRNA). However, other types of RNA have also been discovered, classes and types of which have a role alternative to encoding genetic material. For example, rRNA (ribosomal RNA)[b], tRNA (transfer RNA)[c], snRNA (small nuclear RNA)[d], and more were discovered in

[a] The more we discover, the more the debate about the continued validity of the central dogma exactly as it was written continues. For example, prions are proteins which, some argue, enable some information flow from proteins to the genome.[2]

[b] forms part of the ribosomal machinery

[c] adaptor molecule used in translation to recruit amino acids to the mRNA and ribosome

[d] involved in regulation of intron splicing and RNA processing

due time (representative structures are seen in Figure 1.2) – these RNA classes are crucial to the production of proteins but not directly involved in the flow of information.[4]

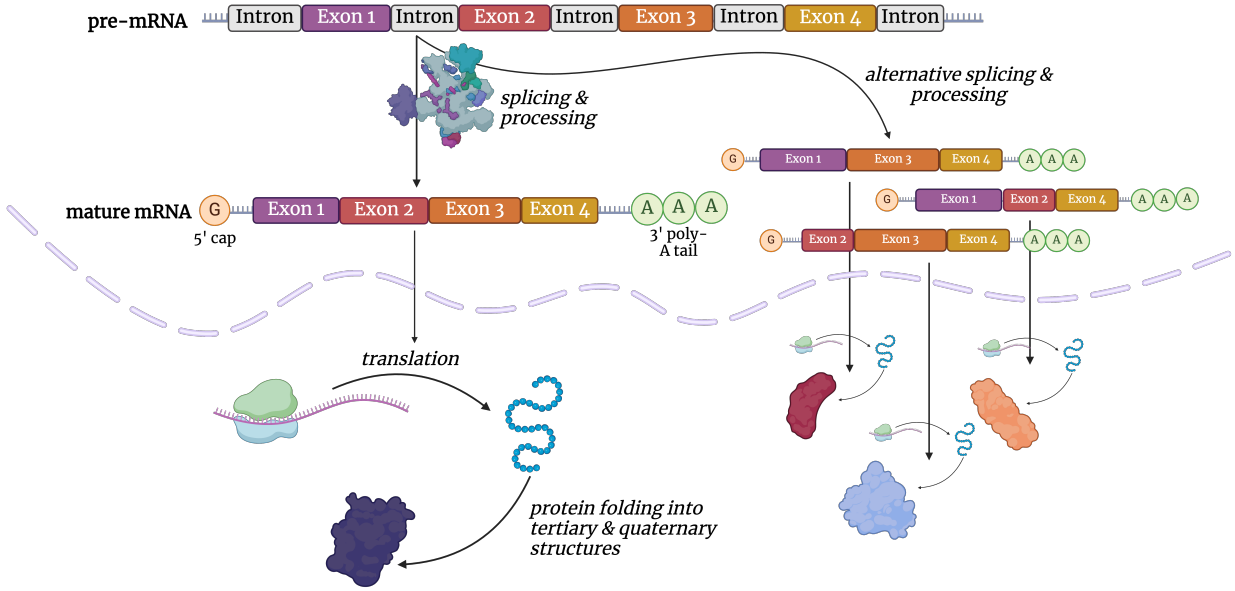


Figure 1.3: Schematic representation of splicing and alternative splicing: introns of the pre-mRNA are spliced out by the spliceosome and processed via 5'-capping and 3'-polyadenylation. The mature mRNA is exported across the nuclear membrane into the cytoplasm where it undergoes translation by a ribosome into a peptide which folds into a protein. Alternative splicing occurs when alternative exon combinations are spliced together, producing a variety of proteins.

Eukaryotic genes include additional processing steps to form mature messenger RNA from the pre-mRNA produced by the RNA polymerase. This processing includes "splicing", in which the non-coding regions (introns) of the pre-mRNA are removed and the coding regions (exons) are ligated together (depicted in Figure 1.3). The existence of exons and introns, discovered in 1977,[5] was very surprising because the vast majority of work had until then been focused on simple bacterial genes, whose mRNA does not have this feature. The origin of these intervening, non-coding sequences remains a long-standing debate due to their evolutionary significance, since introns and exons are seen across all eukaryotic life, but not in prokaryotic genomes.[6] Regardless of the intron-early or intron-late debate (referring to the origin of introns in the evolutionary timeline), it is widely accepted that introns, and exon splicing in general, is extremely important to the complexification of life in adaptive evolution due to alternative splicing.[7] Alternative splicing, in which exons from the same gene

are spliced together in different combinations, represented in Figure 1.3, is a crucial function in the eukaryotic genome which allows it to be more efficient − a single gene creates many different mRNA sequences, thus diversifying protein expression. The proposal of alternative splicing by Gilbert in 1978[8] addressed the discrepancy between the number of known protein-coding genes (∼25,000) vs. known proteins (∼>90,000).[6] In addition to splicing, processing of the RNA at the 5'-end (capping) and 3'-end (poly-adenylation and cleavage) produces mRNA which is exported from the nucleus to cytoplasm for translation.[4]

Splicing is a process which must be highly accurate − mis-splicing events as a result of genetic mutations may lead to altered protein expression and contribute to a wide variety of diseases such as muscular dystrophies, neurodegenerative conditions, and cancer.[6] Harnessing alternative splicing and inducing an altered splicing pattern to produce functional (or partially functional) protein is a treatment strategy for diseases such as spinal muscular atrophy (SMA) and Duchenne muscular dystrophy (DMD) which suffer from loss-of-function genes.[9]

Translation is the final and highly conserved step in the synthesis of proteins; the information contained in the nucleic acid sequence is translated to an amino acid "language" sequence through the (nearly universal) genetic code (see Figure A.1 for the code, a schematic representation, and a more detailed description of translation). Upon termination of translation, a polypeptide is released which folds to form a final functional protein. This primary, one-dimensional sequence of amino acids dictates secondary, tertiary, and in fact even quaternary structure of complex 3-dimensional proteins. This "sequence hypothesis", first hypothesised by Crick in 1958 in the same seminal paper, still holds mostly true to this day.[10] It is no longer surprising that a simple 20-amino acid language, encoded by an even simpler 4-unit genetic one, should produce such complexity, but at the time prior to the discovery of nucleic and amino acids, this produced much controversy: how is a biopolymer consisting of only four repeated subunits capable of storing and self-replicating genetic material?

1.1.2 The structure of DNA and RNA

Arguably one of the most important discoveries of the 20th century, the elucidation of the structure of DNA, began to answer this question. Photograph 51, an x-ray based diffraction image taken by graduate student Raymond Gosling working under the supervision of Dr. Rosalind Franklin at King's College London in 1952 first demonstrated that DNA was a double helix (not a protein as some thought or a triple helix as Pauling first suggested).[11] The double helix model, put together by Crick and Watson and presented in a Nature paper in 1953, explained how nucleic acids could both store and self-replicate genetic information.[12]

1.1.2.1 Primary structure

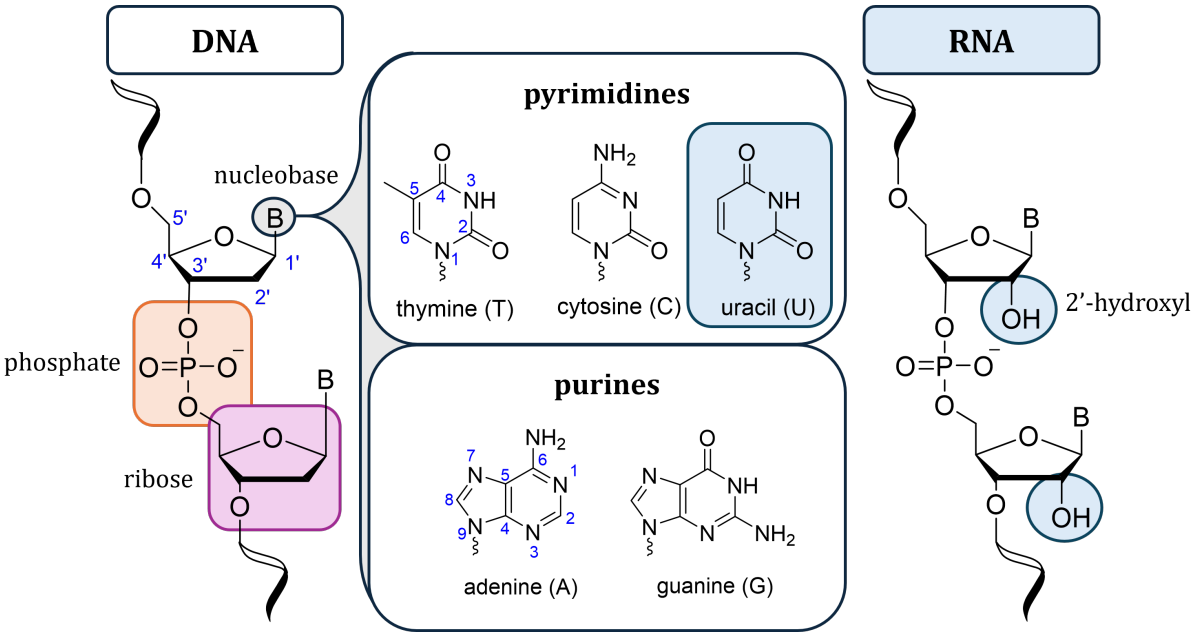


Figure 1.4: Primary structure of DNA is made up of repeating nucleotide units, each containing a ribose (pink), phosphate linkage (orange), and nucleobase (grey). Structure of the nucleobases are categorised into pyrimidines and purines. Structural differences of RNA are highlighted (blue) – the addition of a 2'-OH group and the use of uracil (U) instead of thymine.

DNA and RNA are biological polymers; each monomer subunit includes a ribose, phosphate backbone, and a nitrogen-containing nucleobase (thymine (T), cytosine (C), adenine (A), or guanine (G)) (Figure 1.4). A 5-membered ribose ring with a nucleobase attached is termed

a nucleoside[e], while the inclusion of the phosphate backbone at the 3'-hydroxyl position is termed a nucleotide.[13]

The structural difference between DNA and RNA is two-fold and highlighted in Figure 1.4: RNA ribose units include a 2'-hydroxyl group and thymine nucleobases are replaced by uracil nucleobases, which differ by a methyl group. RNA occurs naturally as a single strand due to its role in genetic expression and regulation, while DNA occurs naturally in cells as double stranded due to its primary role as information storage. RNA is much more susceptible to degradation due to the hydroxyl group. Interestingly, the "RNA world" hypothesis posits that RNA existed prior to DNA; DNA evolved as a less-reactive long-term storage solution while RNA was a reactive molecule capable of self-replicating as well as storing genetic information in the short term.[14]

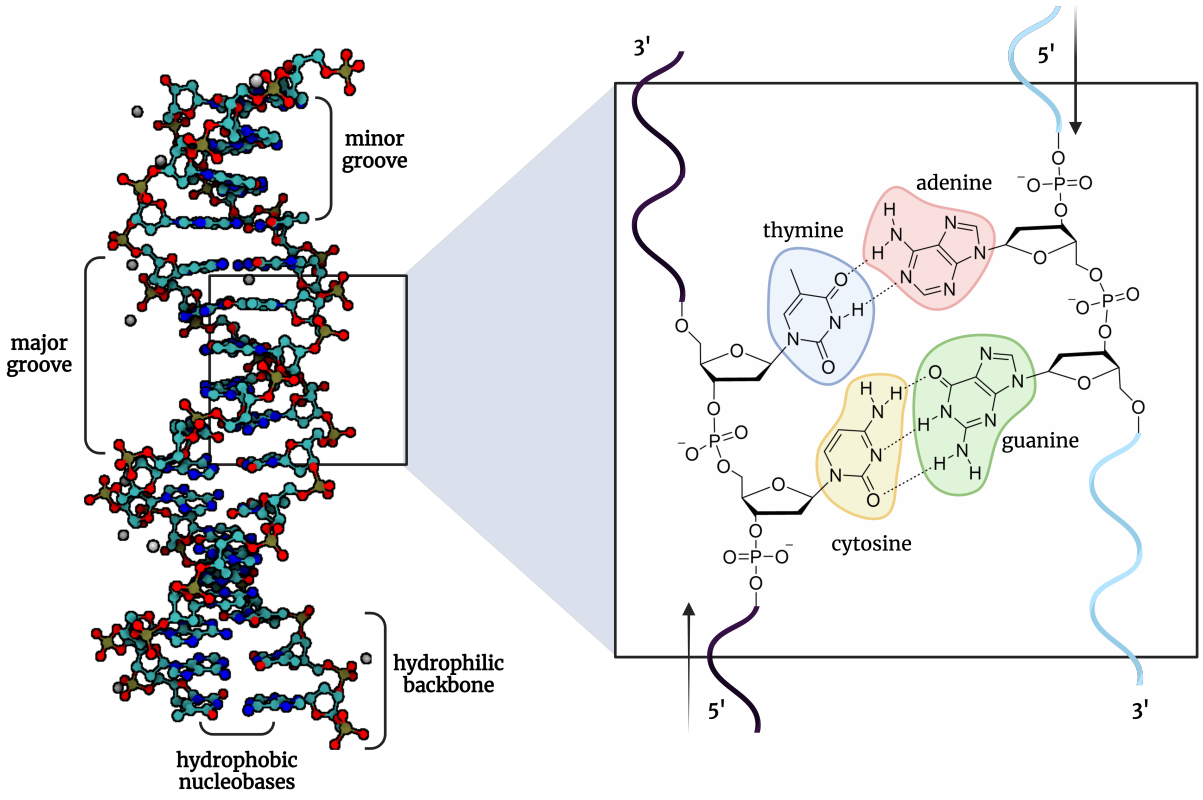


Figure 1.5: Crystal structure of a 16-mer B-form DNA duplex (PDB, 3BSE), with schematic representation of Watson-Crick hydrogen bonding between A-T and C-G nucleobases.

[e]Numbering convention uses prime (') to differentiate the furanose carbons from the atoms in the 6- or 9-membered nucleobases, which are numbered starting with a nitrogen in a counter-clockwise direction.

Chargaff first observed that the ratio of A:T and C:G was always 1:1 in DNA, which ultimately aided Watson and Crick in hypothesising that purines (A and G) and pyrimidines (T and C) hydrogen bonded to each other.[15] Now called Watson-Crick base pairing, the hydrogen bonding between thymine and adenine and between cytosine and guanine drives the formation of the double helix between two single strands of DNA (ssDNA) to form double-stranded DNA (dsDNA), visualised in Figure 1.5. The hydrogen bonds between the two strands sequester the hydrophobic nucleobases to the inside of the helix, while the hydrophilic sugar-phosphate linkages form the solvent-exposed backbone.[16] Sequences rich in G-C base pairs require more energy to dissociate from their antiparallel strand, as there are three hydrogen bonds between G-C rather than the two found in A-T base pairing.[13] Finally, the phosphate backbone is negatively charged; in biological systems, this charge is neutralised in solution by salts such as Na^+, Mg^{2+}, polyamines such as spermine, or the lysine and arginine side-chains in many DNA-binding proteins.[16]

The dsDNA is extremely stable and can store genetic information for millenia – DNA is inherently a functional coding system only *because* of its helical structure that can be unwound and due to the conserved nature of Watson-Crick base-pairing, be used to replicate itself or encode for the more reactive RNA biopolymer.[14]

1.1.2.2 Secondary structure

The primary structure of subunits (nucleotides) was understood since the 1940s, but the secondary helical structure was only discovered in 1953.[13] Part of the difficulty in elucidating these structures via diffraction images lies in the fact that nucleic acid helices can adopt different global helical structures, namely A-form and B-form.[f]

The hydrophobic nucleobases hydrogen-bond to each other according to Chargaff's rule in order to minimise exposure to aqueous solution with the more hydrophilic sugar-phosphate groups organised to the outside, giving rise to a helix. On the inside of the helix, the base pairs stack on top of each other (termed a "step") through the π-π-interactions of their

[f] Z-form is also discussed for completion, but has less relevance in this thesis.

aromatic rings. The nucleobases are planar, but when they engage in hydrogen bonding, they have geometric parameters relative to each other (Figure 1.6). For example, they can hydrogen-bond at an angle twisted away from each other, called "propeller twist". Additional parameters include stretch (displacement along the x-axis), shear (z-axis), and stagger (y-axis). Each nucleobase step also has geometric parameters with regards to the step it stacks on top of or below (see base step parameters in Figure 1.6). For example, these steps are stacked at an angle relative to each other, the "twist angle", which can range from 27.9° to 40°, to maximise hydrophobic interactions.[16] Additional base step geometry parameters include base pair tilt (angle at which they are vertically stacked), rise (distance in the y-axis between the base pairs), or slide (displacement along the x-axis).[g17]

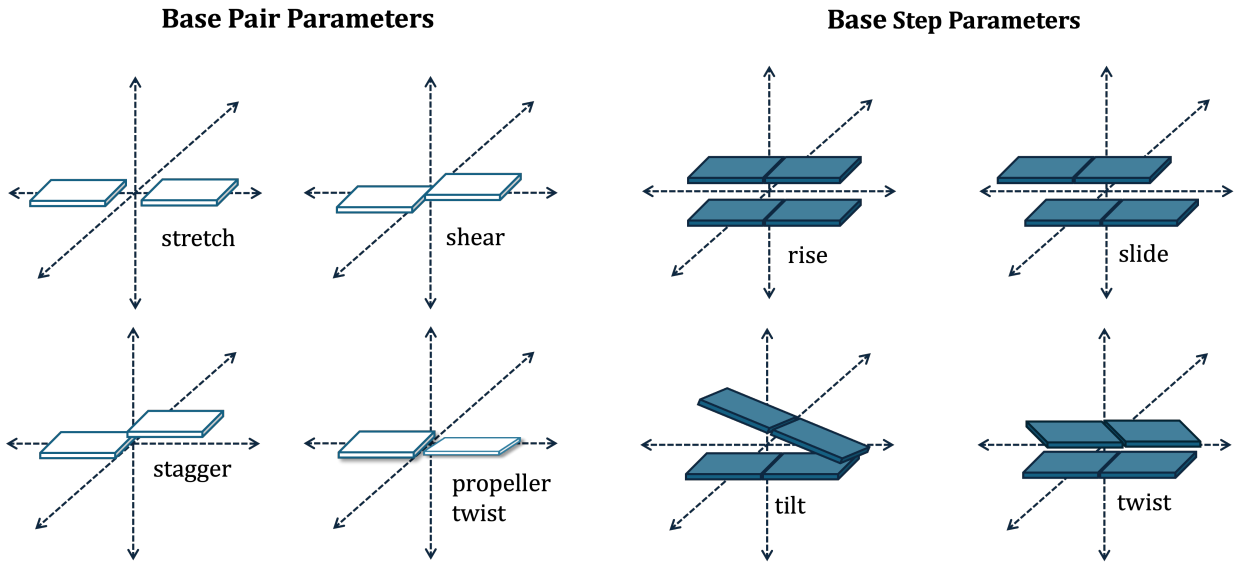


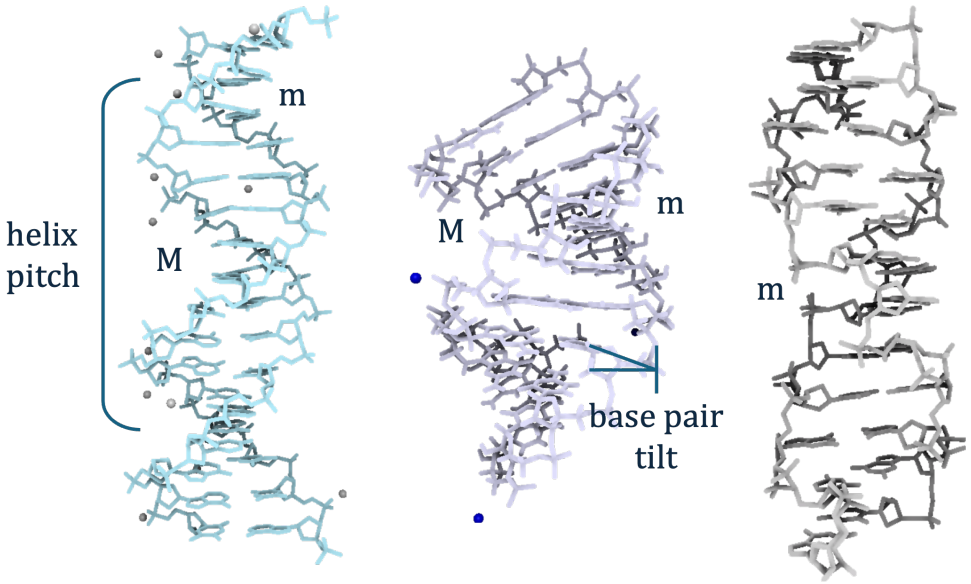
Figure 1.6: Examples of important base pair and base step parameters. Together, the geometries at which the base pairs hydrogen-bond to each other and in which the base steps stack inform the global helical conformation.

Together, the geometries between the paired planar nucleobases (base pair) and the geometries in which each nucleobase pair is stacked (base step) greatly affect the global conformation of the double helix. For example, twist angles of 40° corresponds to a helix of only 9 base pairs (bp) per helical turn, while a reduced twist angle corresponds to an elongated helix.[16] Each nucleobase pairing (16 in total) adopts slightly different angles, as well as stacking energies (the energy required to dissociate the stacked pairs). Therefore, the sequence of

[g]There are additional base pair and base step geometries, but some are ommitted for the sake of brevity.[17]

the DNA greatly effects the global conformation. Indeed (considering only DNA), A-form helices are favoured by purine-rich sequences, B-form is favoured by mixed sequences, and Z-form is observed for stretches of pyrimidine-purine-alternating sequences.[16]

A.



B.

Helical Structure	B-form	A-form	Z-form
Handedness	Right-handed	Right-handed	Left-handed
Base pairs per turn	10	11	12
Diameter	20 Å	23-26 Å	18 Å
Helical rise per bp	3.4 Å	2.6 Å	3.7 Å
Pitch/turn of helix	33.2 Å	24.6 Å	44.6 Å
Base pair tilt to axis	0-1°	19°	9°
Sugar pucker	C2'-endo	C3'-endo	C2'-endo (C, T) C3'-endo (A, G)
Major groove (M)	Wide, deep	Narrow, very deep	One groove: deep, narrow
Minor groove (m)	Narrow, deep	Very broad, shallow	

Figure 1.7: **A.** Crystal structures of a 16-mer B-form DNA duplex (blue, PDB, 3BSE), 18-mer RNA A-form duplex (purple, PDB, 1QCU), and 12-mer Z-form DNA duplex (grey, PDB, 4OCB). **B.** Table of helical structure parameters for B-, A-, and Z-form duplexes, compiled from literature sources.[16,18,19]

In B-form, the canonical form adopted by DNA-DNA homoduplexes, the nucleobases are almost perpendicular to the helical axis (base pair tilt to axis = $0-1°$), making it a fairly symmetrical long cylinder 20 Å in diameter, with a rise between base pairs of 3.4 Å (Figure 1.7B). Most importantly, the distribution of sugar puckers in B-form DNA tend toward the

C2'-endo type, with a distance of 7.0 Å between the two phosphates per nucleotide (nt). C2'-endo refers to the carbon which points above the ribose plane; sugar conformations are depicted in Figure 1.8.[18] One of the structural changes which induces a macroscopic B-form to A-form shift is the conformational change from C2'-endo to C3'-endo sugars, in which the C3'-carbon points above the ribose plane. This causes the distance between two phosphates to be reduced to 5.9 Å and the overall helix becomes more compact (11 bp per turn compared to 10, and a rise ber bp of 2.6 Å).[18,19] Additionally, the base pairs are stacked further off centre, increasing the width of A-form helices (diameter of 23-26 Å) and the base pair tilt is higher (19°). Due to the 2'-hydroxyl group, the furanose rings in RNA are most stable in the C3'-endo ribose conformations. Therefore, RNA:DNA heteroduplexes and RNA:RNA homoduplexes are known to adopt A-form helical structures. A-form helices also have a

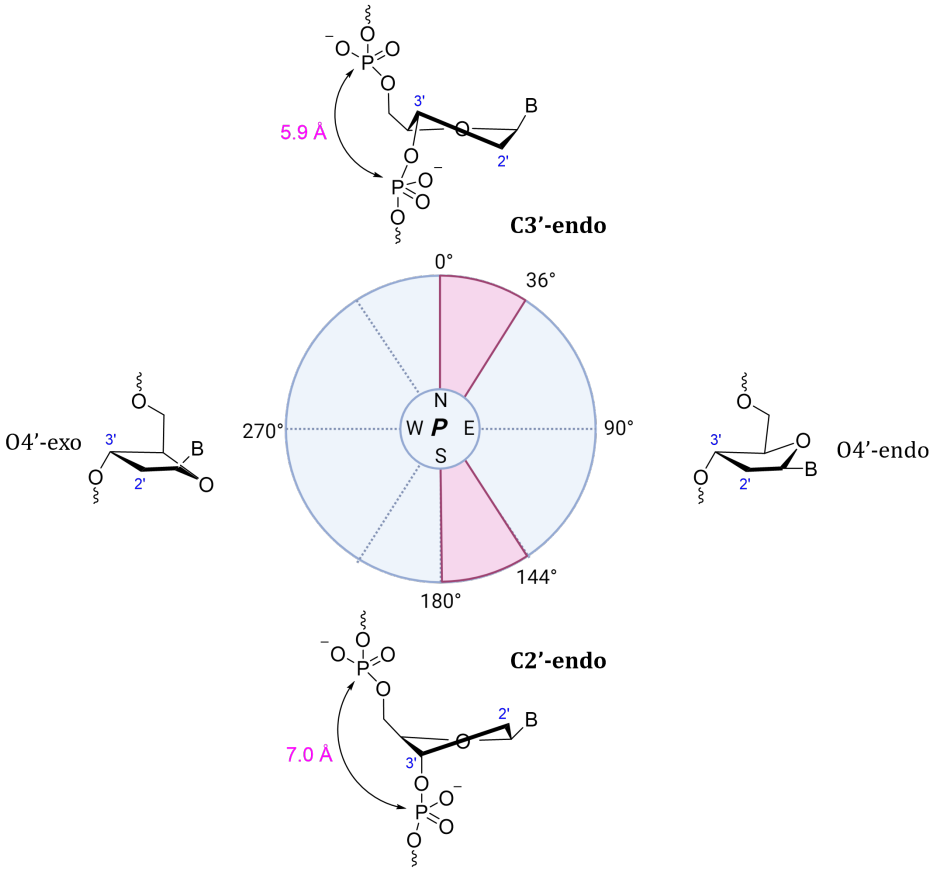


Figure 1.8: Conformations of the ribose ring in a nucleotide, oriented along the pseudorotation phase angle (*P*, or pucker). C3'-endo conformations are of the *North*-type between 0-36° out-of-plane distortion while C2'-endo conformations are of the *South*-type between 144-180°. Figure adapted from the Damha et al.[20]

larger open center which is able to accommodate these 2'-OH groups.

Finally, Z-form DNA is a strikingly different form of DNA helix which is usually seen in stretches of alternating pyrimidine/purine sequences (e.g. CGCGCG). This helix is left-handed, with the phosphate backbones adopting a zig-zag pattern. The zig-zag is formed by alternating G in a *syn* C3'-endo conformation with the C in an *anti* C2'-endo conformation (see *syn* and *anti* nucleosides in Figure A.2).[21] While Z-form DNA was long thought to be a structure only induced by high salt concentrations, it is now being recognised as having cellular functions such as regulating eukaryotic gene transcription.[17]

As a duplex forms, a major and minor groove become apparent, dictated by the space between the two sugar-phosphate backbones (distance is largest for major, smallest for minor). The hydrogen-bonded nucleobases sequestered to the core of the helix are exposed to the aqueous environment more in the major groove than the minor, which makes the grooves' conformation important to protein-binding which may or may not be sequence-specific.[22] In B-form DNA, the major and minor groove widths exist in a ratio of 7:4;[19] the width of the major groove can approximately accommodate a third nucleobase (for triplex formation) or expose the nucleobase sequence to proteins which recognise the hydrogen bond donor and acceptor pattern that is complementary to their amino acid side-chains. For example, insertions of α-helixes of a DNA-binding protein which recognises only certain DNA sequences occurs in the major groove of B-form DNA.[16,17] In A-form helices, the base pairs are tilted with respect to the helical axis by $\sim 20^{\circ}$, causing a very narrow and deep major groove and a shallow minor groove (see Figure 1.7 for crystal structures with annotated major (M) and minor (m) grooves).[22] This highly accessible minor groove has been shown to be involved in DNA-replication – the DNA polymerase, which is not sequence-specific, will induce A-form like duplexes whose minor grooves are very perturbed by mismatched base-pairing to increase the fidelity of replication.[23]

1.1.3 Solid-phase oligonucleotide synthesis

As soon as the structure of DNA was understood, the first attempts were made to synthesise DNA chemically. The very first dinucleotide was synthesised by Michelson and Todd in the 1950s by coupling a 3'-phosphorochloridate thymine monomer with a 5'-OH thymine monomer (Figure A.1A).[24] The phosphodiester method developed by Khorana in the 1960s utilised a 5'-trityl protected phosphate monomer (3'-OH), which was activated by *N,N'*-dicyclohexylcarbodiimide (DCC) and coupled to a 5'-phosphate monomer (Figure A.1B). This method introduced very important concepts such as the use of an acid-labile trityl protecting group, as well as the standard nucleobase protecting groups. However, the phosphate linkage synthesised was itself not protected and branching of the oligonucleotide was a problem. The continual need to purify or precipitate between couplings in solution also meant the method was incredibly time-consuming.[25]

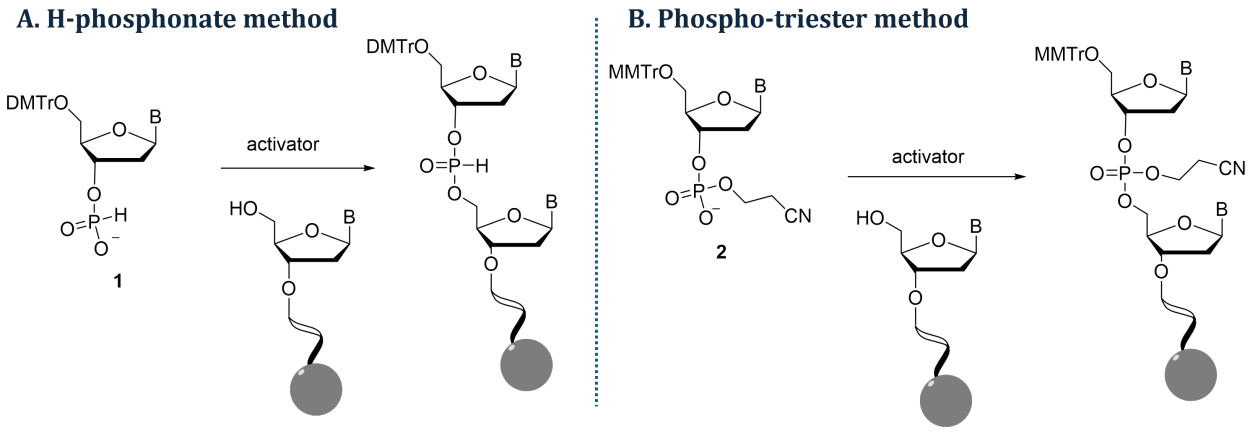


Figure 1.9: Early solid-phase oligonucleotide synthesis methods: **A.** H-phosphonate method to SPOS using a P(V) H-phoshonate monomer 1. **B.** Phosphotriester method to SPOS using a β-cyanoethyl-protected P(V) monomer 2.

The ability to chemically synthesise longer oligonucleotides in a high-yielding and time-efficient manner was revolutionised by the synthesis of a tetrapeptide on a solid support by Merrifield in 1963.[26h] The application of solid-phase chemical synthesis to oligonucleotides (now termed solid-phase oligonucleotide synthesis (SPOS)) had varying success depending on the method used. An H-phosphonate method, depicted in Figure 1.9A, used a 5'-H-phosphonate monomer **1**, which was activated by an acetylating agent and coupled to a

[h]The establishment of solid-phase chemistry won Merrifield the Nobel Prize in 1984.

growing 5'-OH oligonucleotide on solid support. This produced an H-phosphonate backbone which was oxidised through the use of iodine at the very end of the synthesis.[27] Letsinger introduced the phosphotriester approach (Figure 1.9B), in which the phosphorus center of a phosphotriester monomer **2** was protected with a β-cyanoethyl group, solving the problem of internucleoside branching produced by coupling a phosphodiester. Combined with the solid-phase approach, this approach yielded the very first semi-automatic oligonucleotide synthesisers.[28] However, in both approaches, the reduced reactivity of the P(V) centre meant that coupling times were long and coupling yields could not reach near-quantitative levels, limiting the length of oligonucleotides possible.[29]

Finally, a breakthrough was made by Caruthers with the switch to highly reactive P(III) chemistry using a 3'-cyanoethyl-protected phosphoramidite monomer (Figure 1.10, monomer **3**).[30] This "phosphoramidite" method, developed by Caruthers in the 1980s, remains the gold standard for SPOS today, enabling highly efficient automated synthesis of oligonucleotides up to 200 base pairs (bp) in length.[29,31,32]

The synthesis cycle builds synthetic DNA in the 3' to 5' direction (opposite to DNA/RNA synthesis in the cell), adding each nucleoside in a cycle of detritylation, coupling, capping, and oxidation.[29] It begins with a controlled pore glass (CPG) resin bead, which has the first 5'-DMTr-protected nucleoside covalently attached ("start" in Figure 1.10, oligonucleotide **I**). Detritylation in acidic conditions (commonly 3% trichloroacetic acid in dichloromethane) reveals the 5'-OH group (oligonucleotide **II**). The next nucleoside phosphoramidite **3** is introduced in large excess; the P(III) monomer must be kept under inert conditions. Tetrazole, or a derivative thereof, is used as an activator which displaces the protonated diisopropylamino group and is rapidly displaced by the incoming 5'-OH group on the growing oligonucleotide chain to produce the longer oligucleotide **III**. The cyanoethyl group remains as a protecting group of the P(III) moiety until the end of SPOS. Unreacted 5'-OH groups are capped with the use of acetic anhydride (mixed with *N*-methylimidazole[i]) to produce a halted oligonucleotide **IV**. Capping prevents failed sequences from reacting in following steps, which would

[i]Used to maintain a basic pH to prevent detritylation.

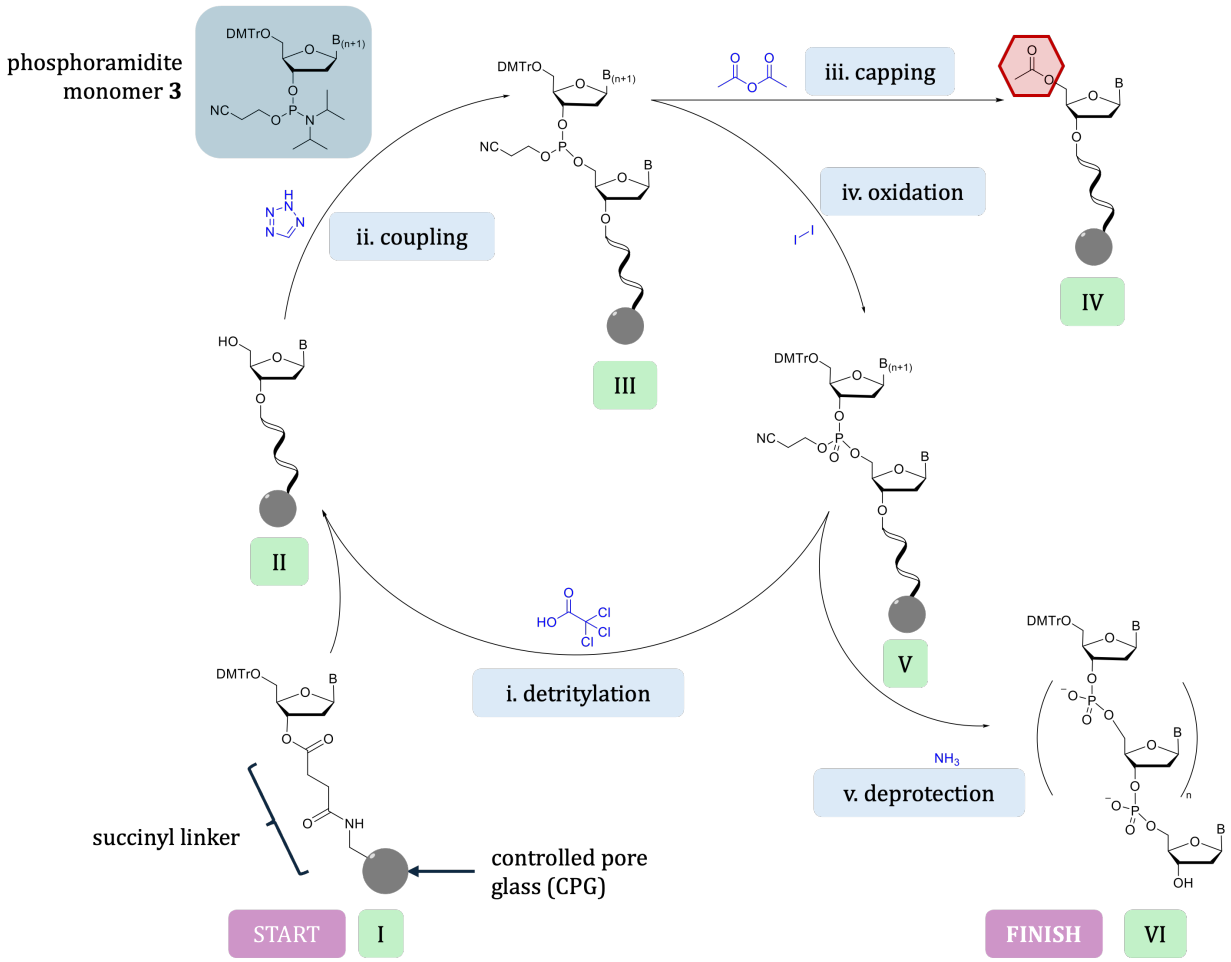


Figure 1.10: Solid-phase oligonucleotide synthesis (SPOS) cycle using P(III) phosphoramidite chemistry.

produce a mixture of by-products ever-growing in complexity as the potential combinations of n-1 failures grow. Finally, the P(III) linkage is oxidised to P(V) by iodine in the presence of water and pyridine to produce oligonucleotide **V**. The P(III) linkage is unstable to acidic conditions (required by the detritylation of the next step) and so, must be converted to the stable P(V) prior to the next coupling. Finally, the cycle begins again with 5'-detritylation in acidic conditions (Figure 1.10).[32,33]

After the final coupling, the oligonucleotide must be cleaved from the solid-support. Linkers to the solid support have evolved over the years, but one of the most common remains the succinyl linker, which can be rapidly cleaved by ester hydrolysis using concentrated ammonium hydroxide. This same solution is used to hydrolyse the cyanoethyl protecting groups from each P(V) linkage, as well as the nucleobase protecting groups, to produce final

oligonucleotide (now in-solution) **VI** ("finish" of Figure 1.10).[33]

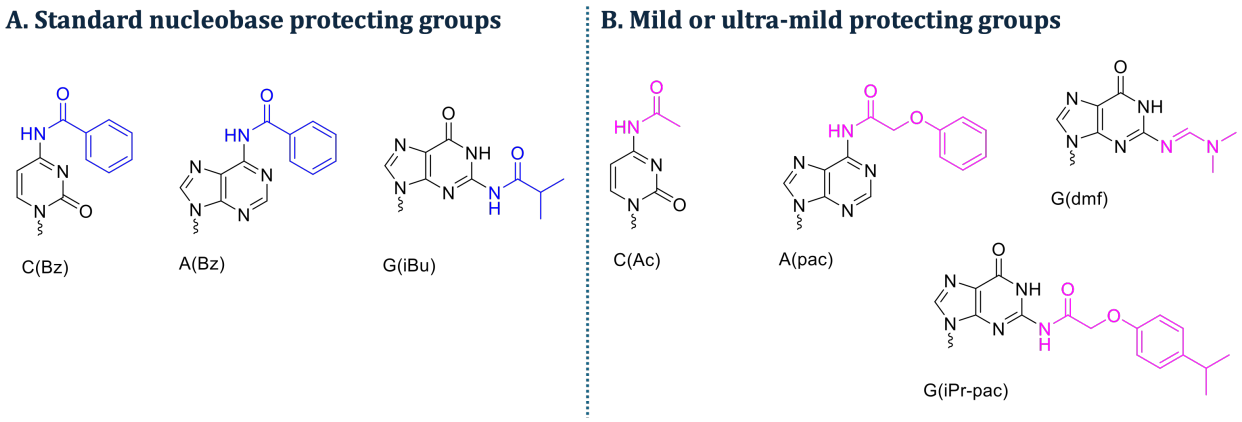


Figure 1.11: A. Standard protecting groups cleaved during oligonucleotide deprotection with concentrated aqueous ammonia at 55 ° C. **B.** Milder protecting groups which facilitate the use of 33% aqueous ammonia, 40% methylamine, or K_2CO_3/MeOH in oligonucleotide deprotection.

Nucleobases other than thymine require protecting groups due to primary amines which could react in the SPOS cycle.[33] Standard protecting groups are usually benzoyl (Bz) for cytosine and adenine and isobutyryl (iBu) for guanine, which can be removed by heating the oligonucleotide in concentrated aqueous ammonia at 55 °C for 5 h (Figure 1.11A). The isobutyryl group is by far the most resistant to cleavage and thus, a more labile protecting group such as dimethylformamidine (dmf) can also be used to protect guanine when the oligonucleotide requires milder deprotection conditions (reducing the time at 55 °C to 1 h).[34] Some oligonucleotide chemistries require much less aggressive conditions to prevent degradation by aqueous ammonia. These "ultramild" protecting groups include acetyl (Ac) for cytosine, phenoxydeacetyl (pac) for adenine, and isopropyl phenoxyacetyl (iPr pac) for guanine (Figure 1.11B). These can be cleaved in a mixture of 33% aqueous ammonia and 40% aqueous methylamine or potassium carbonate in methanol.[35] Acetyl is required as a protecting group for cytosine in these cases, as transamination by methylamine can occur if the less-labile benzoyl group is used.[33]

Oligonucleotides can now be produced on a nanomole (nmol) to kilogram (kg) scale, depending on their required use and length. Synthetic DNA has arguably revolutionised molecular

biology and biochemistry as an "indispensable component of basic research"[j] – polymerase chain reaction (PCR) assays are now routine and techniques such as cloning plasmids and CRISPR-genome editing are impossible without synthetic oligonucleotides.[31]

Oligonucleotides are most commonly purified by high-performance liquid chromatography (HPLC), using different methods such as reverse-phase or ion-exchange chromatography. The reverse-phase technique (RP-HPLC), used in this thesis, elutes the oligonucleotide with increasing percentage of non-polar eluent (e.g. 20 → 80% ACN in 0.1 M TEAA) separating the products based primarily on the oligonucleotides' interactions with a hydrophobic column (e.g. C18), while maintaining constant ion concentration (see Figure 1.12). Ion-exchange (IE-HPLC) is also a popular method of separating desired oligonucleotides from their n-1 or n+1 (and all other failed sequences) impurities. A gradient of salt concentration in the eluent is used (e.g. 0 → 1M NaBr), which competes for adsorption onto the column, and the oligonucleotide is eluted primarily based on charge distribution.[36] IE-HPLC is often used when the oligonucleotide forms secondary structure such as hairpins and loops, as the mobile phase is highly basic and disrupts hydrogen bonding.[36]

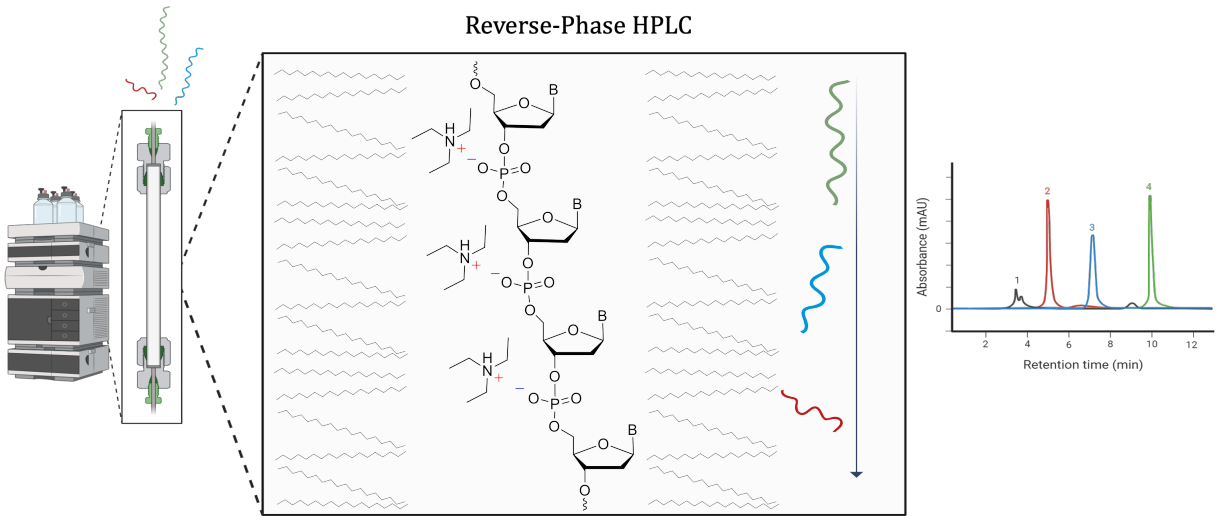


Figure 1.12: Schematic representation of reverse-phase HPLC. Oligonucleotides are separated on a column containing a hydrophobic stationary phase (e.g. C18). They are ion-paired to a buffer such as TEAA and separated on the basis of hydrophobicity (and therefore also size).

[j] Caruthers, in a reflection essay in 2013.[31]

Oftentimes, the oligonucleotide is purified "DMTr-on", in which the 5'-monomer remains tritylated to provide a useful purification handle; full sequences will be much more hydrophobic with a 5'-DMTr group than any capped failed sequences without. Post-purification, it is easily removed in acidic conditions (3% trichloroacetic acid (TCA) in dichloromethane). Due to the negatively charged phosphate or phosphorothioate backbone, the purified oligonucleotide is always "paired" with a counter-ion. Depending on the buffer used during purification, this is most commonly Et_3N^+. If intended for use in living systems, such as cells or animals, this salt must be "exchanged" for a biologically compatible salt such as Na^+. This is done often by passing the oligonucleotide through a Dowex column or a RP-HPLC column in which a sodium acetate buffer is used.[37]

1.1.4 Characterising oligonucleotides

1.1.4.1 Determining molecular weight and purity

Oligonucleotide purity and characterisation is most commonly determined by liquid chromatography-mass spectrometry (LC-MS), coupled to negative electrospray ionisation (ESI^-). Ion-pair reverse-phase liquid chromatography (IP-RPLC) retains the oligonucleotide based on hydrophobicity in a gradient of polar → non-polar solvent, while including the "ion-pairs" of an alkylamine (e.g. Et_3N) and an acidic modifier (e.g. hexafluoroisopropanol (HFIP)).[38,39] The introduction of HFIP by Apffel et al. was a breakthrough, as it is volatile compared to buffers used previously such as TEAA, can adjust the pH of the solution, and enhances the MS signal intensity by minimising cation adduction.[40] As an LC-MS chromatogram is produced, the molecular weight of each eluted species is determined by ESI^-. The solvent is evaporated and electrospray ionises the oligonucleotide into multiple charged precursor ions (e.g. 4^-, 5^-, 6^- and so on). These mass-to-charge (m/z) ratios can be "deconvoluted" to calculate the oligonucleotide's high molecular weight which is above the detectable range of the mass spectrometer.[41,42] Combining LC-MS and ESI^- is a powerful tool which can simultaneously quantify oligonucleotide purity (and impurities), while characterising these species by mass − Sigma-Aldrich report ESI-MS systems which have a mass resolution of ∼0.03%,

corresponding to ±3 Da for an oligonucleotide weighing 10 kDa.[43]

1.1.4.2 Thermal duplex stability

An important biophysical property of a synthetic oligonucleotide is the stability of the duplex it forms with its complementary DNA or RNA strand. Duplex stability is assessed by the temperature required to separate the strands (melting temperature, or T_m). The oligonucleotide is annealed to its complementary DNA or RNA sequence by combining equimolar concentrations in an aqueous solvent. As the temperature of the sample is increased, the hydrogen bonds between the two complementary strands break and the duplex melts (represented in Figure 1.13A). This is measured by a ultraviolet visible spectrophotometer (UV-Vis) at 260 nm. Single-stranded DNA absorbs more intensely than double-stranded DNA; observing this difference over a temperature gradient of 25 → 75 °C will indicate melting, as shown in Figure 1.13B. A reported melting temperature (T_m) is the temperature at which 50% of the duplex is melted and half remains annealed.[44] Often, the melting of an oligonucleotide duplex is graphically presented as the first derivative of the absorbance at

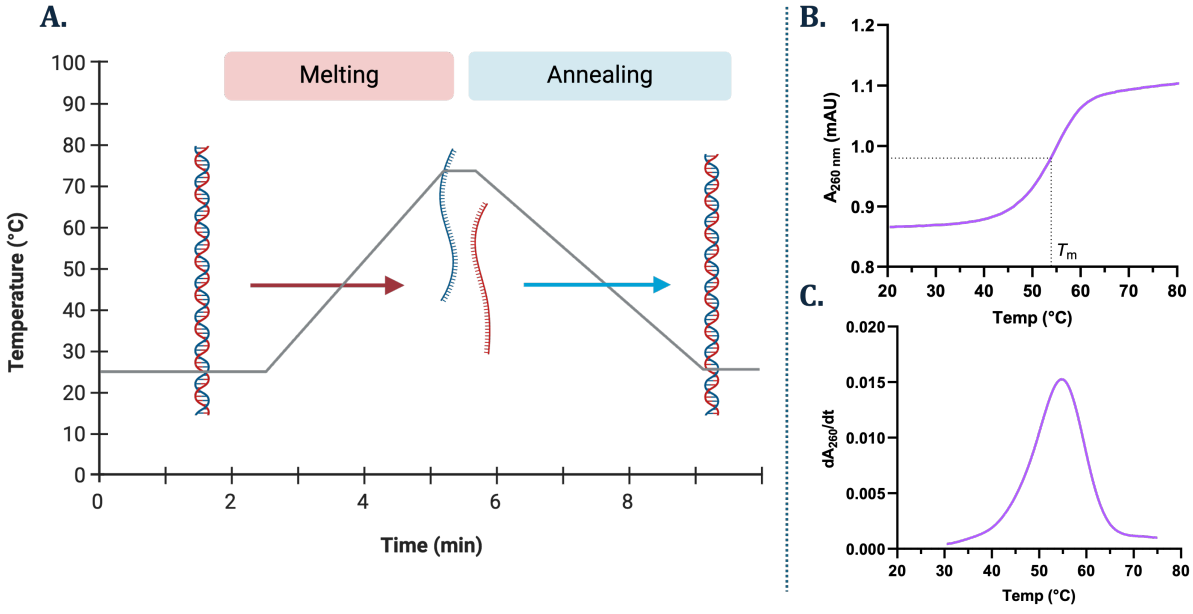


Figure 1.13: A. Representative process of the melting and annealing cycles conducted to measure T_m. **B** Experimental determination of T_m by measuring absorbance. Given T_m is temperature at which 50% of the duplex is single-stranded (inflection point) **C.** T_m presented as the first derivative of absorbance in **B**.

260 nm over the temperature gradient − in this case, the inflection point (the given T_m) is the maxima (see Figure 1.13C).

Melting temperature is correlated to an oligonucleotide's GC content, as oligonucleotides with higher GC content melt at higher temperatures due to the increased hydrogen bonding. T_m can also be predicted[k] through combined consideration of GC-content, base-stacking energies, and oligonucleotide and salt concentrations.[45] T_m is used to inform many molecular biology techniques, such as determining annealing temperatures for PCR, Southern blotting, and in situ hybridisation. In the therapeutic application of oligonucleotides and in this thesis, melting temperature is used to assess whether a chemical modification renders a duplex more or less stable.[44]

1.1.4.3 Circular dichroism

Thermal melting temperature is largely dependent on an oligonucleotide's primary structure − both the nucleobase sequence and structural modifications to the repeating units. However, the secondary structure of oligonucleotides can also be characterised. Circular dichroism (CD) spectroscopy is used to evaluate global duplex conformation and differentiate among A-, B-, Z-form helices (or indicate if the helix adopts a mixed topology).[46]

The founding principle behind CD spectroscopy is that chiral molecules absorb right- and left-handed polarised light differently; the difference in absorption is called ellipticity and is measured in degrees (θ).[46] DNA and RNA duplexes are chiral and have a "handedness"; as seen previously, A-form and B-form helices are right-handed while Z-form is left-handed. Wavelengths of 200-340 nm of circularly polarised light are used (the range at which nucleobases absorb light), and the resulting CD spectra provide a trace or signature for different nucleic acid secondary structures (represented in Figure 1.14A). B-form DNA is the least chiral, as the nucleobases are most perpendicular to the helical axis, and as such, the ellipticity intensities are lower. Most B-form spectra are characterised by maxima at 280 nm and minima at 245 nm (an ASO:DNA heteroduplex exhibiting B-form characteristics is seen

[k] Online calculators such as the OligoAnalyzer™ Tool by Integrated DNA Technologies are publicly available.

in Figure 1.14B). A-form helices are more chiral, due to increased base tilt with respect to the helical axis, and are characterised by maxima at 260 nm and a very intense minima at 210 nm (an ASO:RNA heteroduplex exhibiting A-form characteristics is seen in Figure 1.14B).[47] Finally, Z-form DNA was observed for the first time through CD years before it was observed through crystallography; it is left-handed and is characterised by a signature almost exactly opposite to that of B-form (negative band at 290 nm, positive band at 260 nm, and an intense negative band at 205 nm).[48]

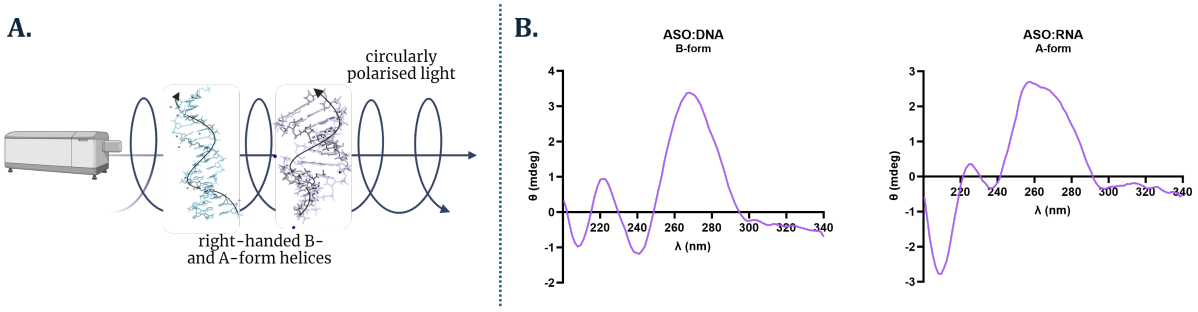


Figure 1.14: **A.** Schematic representation of circularly polarised light shone upon right-handed A-form and B-form duplexes. **B.** Example circular dichroism spectra of an ASO:DNA heteroduplex (exhibiting B-form characteristics) and an ASO:RNA heteroduplex (exhibiting A-form characteristics).

The resolution of these large macromolecules using CD is much lower than for chiral small molecules (which can provide information at the bond-level), and so CD spectra are primarily used to monitor global structural changes.[47] For example, triplex and quadruplex formation can be monitored, as well as condensation.[46] CD is a powerful tool to observe conformational changes through manipulation of variables such as ionic strength, temperature, and concentration.[49,50] Due to the Beer-Lambert law (given in A.3), CD spectra are also extremely sequence-specific, as the extinction coefficient (ε) is dependent on nucleobase content. This specificity and sensitivity means single spectra are not sufficient to make structural interpretations – they should only be interpreted relative to each other through the variation of a single variable. In this thesis, CD spectra are used to assess potential conformational changes to helices of the same sequence and concentration, but carrying various chemical modifications.

1.2 Antisense oligonucleotides: nucleic acids as therapies

The principles of the central dogma and the development of chemical DNA synthesis combined pose a powerful therapeutic premise − a DNA or RNA molecule which can interfere with protein synthesis processes prior to translation of mRNA. This idea, first put into practice in the late 1970s by Zamnecik and Stephenson[1] to inhibit Rous sarcomal viral replication,[51] garnered great interest, but remained surprisingly unsuccessful for decades. The primary reason behind this was the lack of chemical modifications. Native DNA fragments are rapidly degraded by endonucleases ubiquitously present in human serum and cells. Therefore, the potential of oligonucleotides serving as therapeutic agents which alter RNA processing through Watson-Crick base pairing was abandoned for a long time. However, structural modification of DNA has enabled development of antisense oligonucleotides (ASOs) which remain stable in plasma long enough to reach their target RNA and have a therapeutic effect.[52]

1.2.1 ASO mechanisms of action

As a large molecule therapeutic modality, ASOs act through one of two mechanisms − by steric hindrance or protein-mediated degradation (both represented in Figure 1.15).

1.2.1.1 Steric hindrance

ASOs which have entered the cell and been successfully trafficked to the cytoplasm (and/or nucleus) can hybridise to their complementary mRNA and form ASO:RNA heteroduplexes (Figure 1.15, left). This sterically hinders the ribosomal units from binding, stopping translation or blocking reading frames. As the ribosome cannot continue translation, this results in down-regulated or halted protein production.[53] Alternatively, the ASO can enter the nucleus and bind to its complementary pre-mRNA target prior to splicing. These "splice-switching"

[1]Zamecnik and Stephenson report the first use of chemically synthesised DNA as an "antisense" agent − they transfected chick embryonic fibroblasts with a 13-mer which was complementary to the 5' and 3'-ends of the 35S RNA in the retrovirus and demonstrated inhibition of viral replication.

ASOs commonly bind to intron-exon junctions and block the RNA-spliceosome interactions, causing the protein to skip this region, inducing either intron or exon skipping (the exon skipping mechanism is depicted in Figure 1.15, left). This "switches" the splicing pattern in a way that produces alternative mRNA – e.g. inducing a frame-shift which can down-regulate protein production or restoring the reading frame and inducing functional protein production.[54]

Splice-switching ASOs pose a very powerful method to restore or up-regulate protein production. Historically, this therapeutic approach is much more difficult than developing drugs which are antagonists or inhibitors and down-regulate protein function. Diseases such as spinal muscular atrophy (SMA), in which patients do not produce sufficient functional SMN protein, had no FDA-approved treatment until 2016. Nusinersen, the first ASO to reach

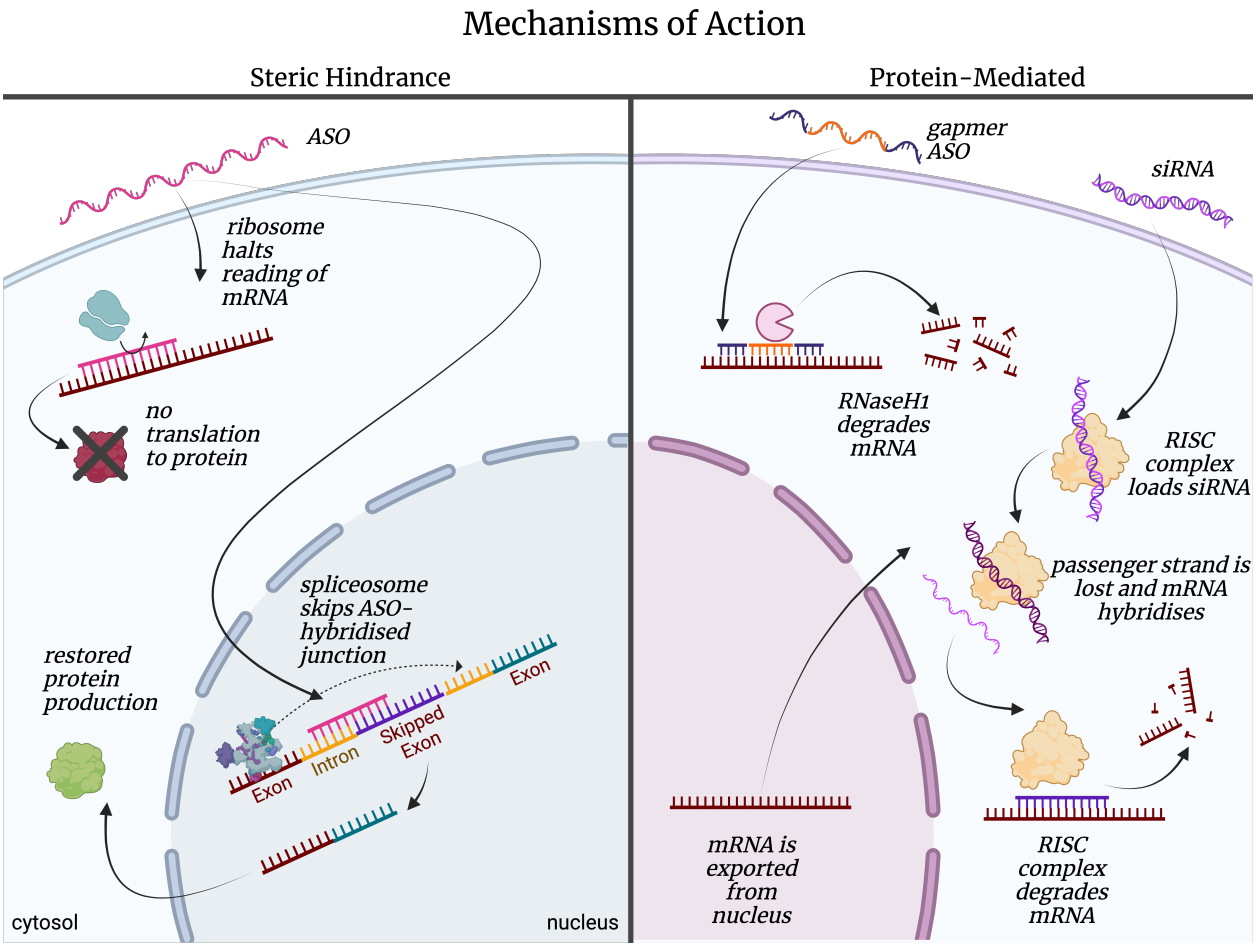


Figure 1.15: Mechanisms of action for therapeutic oligonucleotides, split into steric blocking (left) and protein-mediated (right) types of RNA interactions.

"blockbuster" status, is a first-of-its-kind treatment which induces exon 7 inclusion and restores the production of full SMN2 mRNA and therefore increased SMN protein through splice-switching.[55]

1.2.1.2 Protein-mediated degradation

The main alternative ASO mechanism of action is RNaseH-mediated mRNA degradation, depicted in Figure 1.15 (right). These ASOs also bind to mRNA in the cytoplasm (or nucleus), but rather than acting through steric blocking, rely on the recruitment of the protein RNaseH1.[56] This protein recognises DNA:RNA heteroduplexes and degrades exclusively the RNA strand, thus destroying the mRNA target and preventing its translation to protein.[57] RNaseH1 does not recognise duplexes in which the ASO is modified at the C2'-position[m], as it requires a DNA:RNA heteroduplex. Therefore, ASOs are designed with a central region of DNA sugars (termed the "gap"). For additional stability and sequence specificity, the 2'-modifications are added at the 5' and 3'-ends of the oligonucleotide (termed the "wings"). The majority of "gapmer" motifs are 3-10-3 or 5-10-5, indicating the number of oligonucleotides in the wings and the gap, respectively. Gapmers are a powerful tool to down-regulate protein expression.[58]

Finally, small interfering RNA (siRNA), also sometimes known as silencing RNA, are therapeutic entities which also rely on DNA-protein interactions to degrade mRNA. siRNA molecules are not considered ASOs, because they are double-stranded rather than single-stranded.[59] However, they impart their therapeutic action by annealing to target mRNA in an "antisense" way, and so, they deserve mention. siRNA enters the cell as a double-stranded duplex consisting of the "passenger" strand and the "guide" strand; this duplex associates with the RNA-induced silencing complex (RISC), specifically the Argonaute (Ago) protein. The passenger strand dissociates and the guide strand is free to hybridise to the target mRNA. The Ago2 protein then degrades the target mRNA, preventing its translation to protein. One key difference between Ago2-mediated and RNaseH-mediated mRNA degra-

[m]Chemical modifications to ASOs (and which are/are not compatible with RNaseH protein-mediated degradation) are further discussed in Section 1.2.2.

dation is that the therapeutic siRNA is complexed with the Ago2 protein prior to antisense binding, while RNaseH complexes with the ASO:mRNA duplex.[56,60]

Small molecule therapeutics are traditionally targeted at proteins to inhibit or antagonise their function – functional protein production or up-regulation is a unique therapeutic mechanism of action being explored by additional strategies such as protein replacement, gene replacement, gene editing, and small molecules targeting RNA.[55] Among these however, nucleic-acid based therapies – specifically splice-switching ASOs – are among the most clinically advanced, with six FDA-approved drugs as of 2024.[56] For genetic diseases in which patients suffer from a loss-of-function mutation, splice-switching ASOs pose a powerful and promising solution. The clinical success and importance of protein up-regulation encouraged us to use splice-switching as the mechanism of action for ASOs synthesised in this thesis.

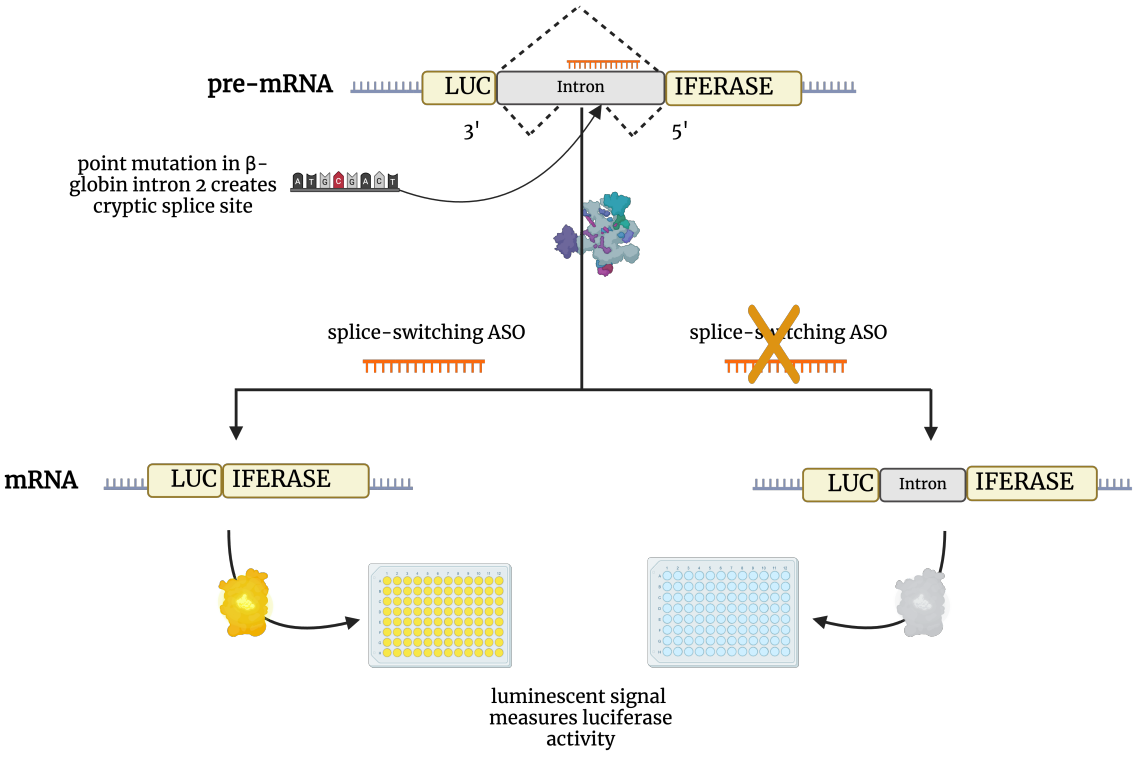


Figure 1.16: Schematic representation of the luciferase assay: pre-mRNA of HeLa pLuc/705 cells contains a mutated intron site in the luciferase gene; an active splice-switching ASO is able to restore the reading frame and produce functional luciferase protein, which is measured by luminescence. Figure adapted from literature.[61]

Chemical modifications reported here are evaluated for their therapeutic potential using a reporter assay which measures the restoration of luciferase protein production. The luciferase assay, depicted in Figure 1.16, uses HeLa pLuc/705 cells − HeLa cells were transfected with a recombinant plasmid that encodes the luciferase gene (pLuc/705). However, a point-mutation in the β-globin intron causes aberrant splicing and lack of luciferase production (Figure 1.16, right).[61] Treatment with ASOs complementary to this splice site induces skipping of the aberrant intron, restores the reading frame, and causes the cell to produce luciferase which is quantified by luminescence (Figure 1.16, left). This luminescence is normalised to protein quantity to account for cell number variation. The assay is a quick and validated method to evaluate splice-switching ASOs;[62] however, other exon-skipping therapeutic models were also investigated, such as a dystrophic murine cell line (H2k *mdx*) which is discussed further in Chapter 5.

1.2.2 ASO chemical modifications

As mentioned previously, synthetic DNA was only transformed into a clinically-viable therapeutic modality through the development of chemical modifications. These structural changes impart drug-like properties to the oligonucleotide such as increased target binding affinity and nuclease resistance.

1.2.2.1 Backbone modifications

The first ASO modification which arguably revolutionised the industry was the phosphorothioate (PS) backbone, first reported by Eckstein et al. in 1970.[63] The simple replacement of a non-bridging oxygen with a sulfur in the phosphodiester backbone was found to render the oligonucleotide much more stable to endonucleases (see backbones in Figure 1.17). PS linkages remain the backbone of choice in ASO design and are clinically validated − they make up the backbones of multiple FDA-approved ASOs on the market.[64] This modification is also one of the few which are compatible with the RNAseH-mediated degradation mechanism.[57] Furthermore, they have been found to extensively bind to proteins in the plasma,

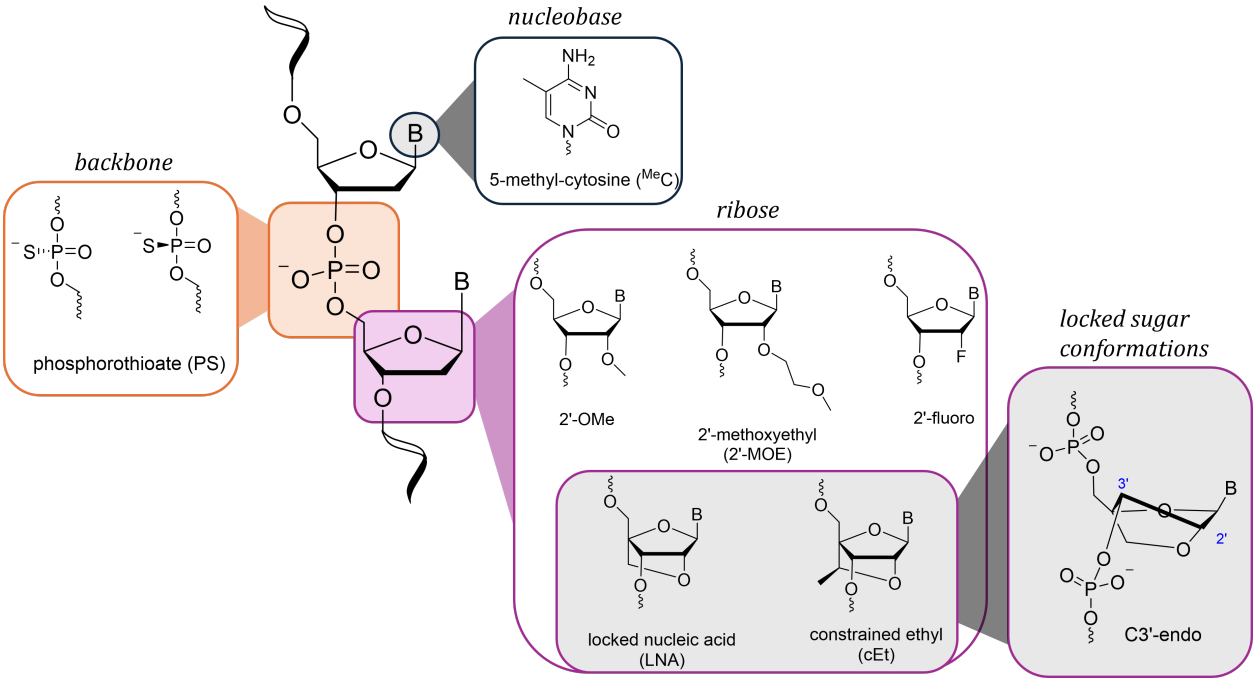


Figure 1.17: The most common chemical modifications to enhance ASO drug-like properties, broken down into modifications to the phosphate backbone (orange), ribose (purple), or nucleobase (grey).

e.g. albumin, which increases their circulation time, reduces their renal clearance, and increases their absorption into tissues.[65] They additionally bind intracellular proteins[n] which may facilitate nuclear accumulation.[66] One of the few drawbacks of the PS backbone is that its incorporation can slightly reduce the thermal stability of an ASO:RNA heteroduplex.[67]

Furthermore, every phosphorothioate incorporation creates a chiral (R- or S-) centre at the phosphorus – a fully-PS-modified 18-mer therapeutic ASO such as nusinersen is made up of more than 100,000 diastereomers. There is evidence to suggest that the stereochemistry of an ASO can have effects on the pharmacological properties – Iwamoto et al. demonstrate that S_p-PS linkages decrease thermal stability relative to R_p-PS linkages and that the RNaseH1 protein has some stereochemical preferences for contact with a backbone with the motif 3'-$S_pS_pR_p$-5'.[68] Of course, purification of an FDA-approved 20-mer ASO such as mipomersen with 524,288 stereoisomers is out of the question. The ability to install stereopure linkages during SPOS has been pioneered by research groups such as Baran's, who develop a chiral P(V) reagent (termed as "(+)-Ψ" or "(-)-Ψ") to program the stereochemistry of the linkage

[n] Protein-binding of ASOs containing the PS backbone is discussed more extensively in Chapter 5.

at the point of nucleoside coupling (see reagent structures in Figure A.3).[69] The ease of synthesis and reduction in synthetic steps (no need for oxidation or sulfurisation of a P(V) linkage) makes it an attractive option for producing stereopure ASOs. However, the sheer number of diastereomers which could potentially be synthesised and tested for pharmacological effect remains vast. As rational ASO design advances and certain universal "rules" about stereopurity may be elucidated, this technology will certainly become extremely useful in bringing stereopure (or partially stereopure) ASOs to the clinic.

1.2.2.2 Ribose modifications

Modifications to the ribose moiety are incorporated into therapeutic ASOs in order to increase their thermal duplex stability, which is slightly reduced by the phosphorothioate. The most common sugar modifications are at the 2'-carbon on the ribose and are depicted in Figure 1.17. They include the 2'-O-methyl (2'-OMe), 2'-methoxy-ethyl (2'-MOE), or 2'-fluoro (2'-F) modifications. By adding steric bulk or electronegativity to the C2'-position, the modifications induce a C3'-endo pucker and the ASO is able to form an A-form heteroduplex (characteristic of RNA) with its RNA target which is more thermodynamically stable than a native DNA:RNA heteroduplex.[56] However, the A-form characteristic of the ASO:RNA duplex also means that they are incompatible with the RNaseH-mediated degradation mechanism which only recognise mixed DNA:RNA (therefore mixed A/B-form) helices.[70] Therefore, all of these modifications are placed in the wings of the gapmers, discussed previously in Section 1.2.1.

The locked nucleic acid (LNA) modification is note-worthy of further discussion – as a bicyclic ribose sugar it "locks" the ribose into the C3'-endo conformation (see Figure 1.17). LNA moieties greatly increase the thermal duplex stability of an ASO:RNA heteroduplex, at ∼7 °C per incorporation.[71] The stabilisation effect can vary from 3-9 °C depending on position and oligonucleotide design, but importantly, it is a "portable" modification which increases the thermal melting temperature of an ASO, regardless of sequence or position.[72] While 2'-OMe modifications also induce the C3'-endo sugar pucker (and thus, A-form helices), they

are not as structurally rigid and the base pairs maintain stacking identical to native RNA.[73]

For both the LNA and related bridged nucleic acid modifications such as constrained ethyl (cEt), the unique bicycle conformation is the driving factor behind their improved binding properties. The bicycle ring system restricts the rotational degrees of freedom around the internal bonds, pre-organising the furanose ring, and reduces the entropy change for hybridisation.[74] Additionally, a crystal structure of an all-LNA helix observes that the LNA nucleotides change base pair geometries so as to increase hydrophobic stacking interactions between bases.[75] This induces a perturbed A-form-like conformation which is "underwound" and the major groove is widened.[72] Finally, molecular dynamics (MD) simulations suggest that the free energy of solvation is 1-2 kcal/mol lower for an LNA moiety than a DNA moiety, making the formation of a duplex thermodynamically favourable.[76]

The increased binding affinity of LNA-containing ASOs is also linked to enhanced potency.[77] Interestingly, the length of gapmers can be reduced as a result of their high affinity − even 12-13-mer gapmers with LNA in the wings demonstrated in vivo potency in murine models.[78] Oligonucleotides containing LNA form such stable duplexes with good mismatch discrimination that they are routinely used as probes[o] for detection of short or low-abundance targets.[79] Unfortunately, the LNA modification has also been associated with increased hepatotoxicity due to this increased binding affinity.[80] However, some gapmer sequences containing LNA have shown good tolerance and no dose-limiting hepatotoxicity, demonstrating that hepatotoxic potential is difficult to predict based on modification only.[72] Indeed, a gapmer with LNA wings is in clinical trials targeting transforming growth factor beta 2 (TGF-β2).[81]

1.2.2.3 Nucleobase modifications

The final moiety of a nucleotide which can be altered is the nucleobase. Interestingly, few chemical modifications to the nucleobase have become cannon due to the importance of Watson-Crick base pairing with target RNA; extremely modified nucleobases may reduce affinity and specificity, as well as alter ASO metabolism.[58] The only chemical modification

[o]In applications such as PCR, microarrays, and in situ hybridisation

clinically validated is methylation at the C5 position of pyrimidines cytosine and uracil (resulting in nucleobases ^{Me}C and ^{Me}U). This modification is found naturally in both bacteria and eukaryotes; methylation of the cytosine is an epigenetic mechanism which has been shown to play a role in gene regulation and cell differentiation.[82] The additional methyl group enhances the thermal stability of the ASO by ∼0.5 °C per modification, as the hydrophobic methyl group can enhance the stacking interactions between planar nucleobase pairs.[83] It is also used to minimise potential immunostimulatory responses; for example, the Toll-like receptor TLR-9 mediates inflammatory responses and recognises unmethylated (but not methylated) CpG motifs.[84]

1.2.2.4 Next-generation modifications and neutral backbones

Increasing thermal duplex stability and nuclease resistance of ASOs via chemical modification has been very successful. However, ASOs continue to face challenges in important drug-like properties such as cellular uptake. Therefore, novel "next-generation" modifications are continually investigated to improve the medicinal chemistry of ASOs. Specifically, many charge-neutral backbones have been heavily explored in the literature as potentially reducing anionic character and increasing naked cellular uptake. Examples of charge-neutral linkages (depicted in Figure 1.18A) include:

1. Phosphorodiamidate morpholino (PMO): this modification has found clinical success, specifically in exon-skipping mechanisms of actions. FDA-approved ASOs such as eteplirsen, golodirsen, and casimersen are fully PMO-modified − all are approved for the treatment of Duchenne muscular dystrophy (DMD). PMO ASOs display good binding affinity, enhanced nuclease resistance, and a low protein-binding profile.[54] However, solid-phase synthesis of PMOs is not conducted through the well-established P(III) phosphoramidite chemistry and is not as straight-forward or high yielding − this is rapidly changing, however, through increased interest in improving automated PMO synthesis.[85]

2. Peptide nucleic acid (PNA): this modification entirely replaces the ribose with a peptide

scaffold, yet the nucleobases are still able to engage in Watson-Crick base pairing. They display enhanced binding affinity and very good nuclease resistance. PNA has been extensively explored in therapeutic ASOs and shown to be effective and safe; a PNA for the treatment of pain in osteoarthritis is currently in an Australian Phase II clinical trial.[86,87]

3. Methoxyphosphonate (MOP): this is a good example of a site-specific fine-tuned chemical modification which can modulate ASO-protein interactions. Ionis Pharmaceuticals discover a single incorporation of this neutral alkylphosphonate linkage at the 2-position in the DNA gap can significantly reduce a gapmer's toxicity.[88]

4. Formacetal: synthesised first by Rozners et al. in 1997,[89] the formacetal linkage has been investigated in gapmers and was found to be stabilising relative to a PS-backbone and have good in vivo potency (ED_{50}=0.5 mg/kg).[90]

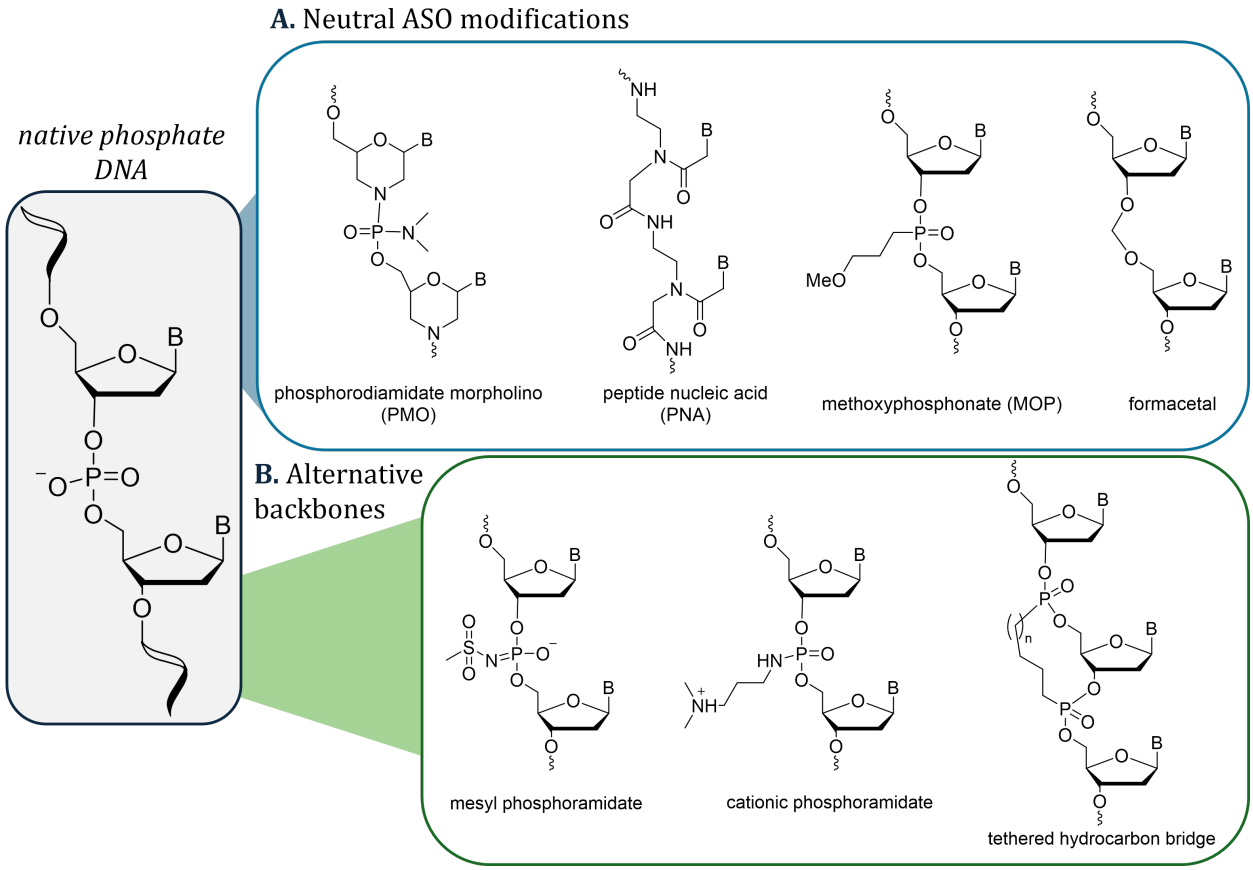


Figure 1.18: A. Examples of next-generation ASO modifications with charge-neutral linkages. **B.** Examples of alternative backbone chemistries explored in the literature.[91–93]

Many other unnatural backbone structures have been explored in the literature. For brevity, a few interesting examples which expand the ASO chemical space are included in Figure 1.18B:

1. Mesyl phosphoramidate: while a tosyl-phosphoramidate linkage was shown to decrease binding affinity, presumably due to steric effects, the reduction in size to a mesyl phosphoramidate linkage has been shown to give an ASO increased thermal stability and nuclease resistance. Importantly, this linkage is compatible with the recruitment of RNAseH protein and enhanced in vivo potency make it a promising alternative to the phosphorothioate backbone.[91]

2. Rather than an anionic or neutral-charge backbone, a cationic linkage such as the cationic phosphoramidate have been incorporated into ASOs that have good affinity and selectivity.[92]

3. "Tethered" trinucleotides were synthesised by Rajasekaran et al. to test the effects of a structurally pre-organised backbone on the on and off-rates of duplex hybridisation. Unfortunately, 15-membered macrocyclic constrained analogues increased off-rates and reduced duplex stability compared to the structurally pre-organised LNA.[93]

Given the clinical and commercial success of fully neutral PMO ASOs and the therapeutic promise of others such as PNA, this thesis aims to investigate neutral backbones − among those explored are the amide (Chapter 2), carbamate and other amide-like backbones (Chapter 3), and alkyl triester (Chapter 4) linkages. However, all linkages are also combined with the pre-organised and rigid LNA modification. It is hypothesised that the combination of the stabilising LNA with an unnatural and neutral backbone would produce an ASO with improved biophysical and biological properties.

1.2.3 Clinical challenges facing ASOs

Given their pre-determined binding principles, well-established chemical modifications, and the advancement of genomic sequencing, it is perhaps surprising that ASOs have not yet

been rapidly approved for a larger variety of diseases. Only 11 ASOs[p] and 6 siRNA are FDA-approved to date. However, many more are in clinical trials. A table of FDA-approved oligonucleotide therapies[q] is given in Table 1.1, including the disease of interest and the

[p]12 ASOs are approved if including the personalised medicine milasen.
[q]Including siRNA for completion.

Oligonucleotide	Market Name and Company	Approval Year	Chemical Modifications	MOA	Delivery	Target	Disease Indication
fomiversen	Vitravene™ by Ionis	1998	21-mer PS	Gapmer	/	CMV IE-2 mRNA (eye)	CMV retinitis
mipomersen*	Kynamro™ by Ionis	2013	20-mer, 2'-MOE, PS	Gapmer	/	Apo-B mRNA (liver)	HoFH
eteplirsen	Exondys 51® by Sarpeta	2016	30-mer PMO	Exon Skipping	/	DMD exon-51 pre-mRNA (muscle)	DMD
nusinersen	Spinraza® by Ionis	2016	18-mer 2'-MOE, PS	Exon Inclusion	/	SMN2 pre-mRNA (CNS)	SMA
patisiran	Onpattro® by Alnylam	2018	2'OMe, 2'F, PS	siRNA	LNP	TTR mRNA (liver)	hATTR
inotersen	Tegsedi® by Ionis	2018	20-mer 2'-MOE, PS	Gapmer	/	TTR mRNA (liver)	hATTR
milasen**	Boston Children's Hospital	2018	22-mer 2'-MOE, PS	Splice switching	/	pre-mRNA cryptic splice site (CNS)	CLN7
volanesorsen	Wayvlivra® by Ionis	2019	20-mer, 2'MOE, PS	Gapmer	/	Apo-C mRNA (liver)	FCS
givosiran	Givlaari® by Alnylam	2019	PS	siRNA	GalNac	ALAS1 mRNA (liver)	AHP
golodirsen	Vyondys 53™ by Sarepta	2019	25-mer PMO	Exon skipping	/	DMD exon-53 pre-mRNA (muscle)	DMD
viltolarsen	Viltepos™ by NS Pharma	2020	21-mer PMO	Exon Skipping	/	DMD exon-53 pre-mRNA (muscle)	DMD
lumasiran	Oxlumo™ by Alnylam	2020	2'-F, 2'OMe, PS	siRNA	GalNac	HA01 mRNA (liver)	PH1
inclisiran	Leqvio® by The Medicines Company	2020	2'-F, 2'-OMe, PS	siRNA	GalNac	PCSK9 mRNA (liver)	FH
casimersen	Amondys 45™ by Sarepta	2021	22-mer PMO	Exon Skipping	/	DMD exon-45 pre-mRNA (muscle)	DMD
vutrisiran	Amvuttra™ by Alnylam	2022	2'-F, 2'-OMe, PS	siRNA	GalNac	TTR mRNA (liver)	hATTR
tofersen	Qalsody® by Ionis	2023	20-mer 2'-MOE, PS	Gapmer	/	SOD1 mRNA (brain)	ALS
nedosiran	Rivfloza® by Dicerna	2023	2'-F, 2-OMe, PS	siRNA	GalNac	LDH mRNA (liver)	PH1
eplontersen***	Wainua™ by Ionis	2023	20-mer 2'-MOE, PS	Gapmer	GalNac	TTR mRNA (liver)	hATTR

Table 1.1: ASOs approved by the FDA. *withdrawn for hepatoxicity in 2018. **milasen is an individualised medicine designed for a single patient. ***approved Dec, 2023, eplontersen is the first ASO which is approved for self-administration via an auto-injector. AHP: acute hepatic porphyria, ALS: amyotrophic lateral sclerosis, FH: familial hypercholesterolemia, CLN7: neuronal ceroid lipofuscinosis 7 (a type of Batten's disease), CMV: cytomegalovirus, DMD: Duchenne muscular dystrophy, hATTr: hereditary transthyretin amyloidosis, HoFH: homozygous familial hypercholesterolaemia, PH1: primary hyperoxaluria, SMA: spinal muscular atrophy. Table compiled from literature sources.[94–96]

chemical modifications used.

While they are an incredibly powerful drug modality, with the potential to target many heretofore "undruggable" diseases, ASOs face a few large challenges before reaching widespread clinical success. Among these clinical challenges are:

- Cellular uptake: a typical 20-mer ASO will have a molecular weight of ∼7 kilodaltons (kDa) and 19 negative charges. Therefore, ASOs are unable to passively diffuse across the cell membrane and must be actively transported through endocytotic trafficking pathways.[97] There are certain receptor-mediated uptake pathways which lead to "productive" uptake vs. "nonproductive" uptake, highlighting the importance of intracellular trafficking[r] in addition to the initial cellular uptake.[98]

- Clinical efficacy and long-term safety: huge improvements have been made via chemical modifications and conjugation or delivery strategies to achieve ASOs that produce clinically efficacious knockdown (via RNA-degradation) or up-regulation (splice-switching). However, restored protein production is often still quite low relative to the dose administered − the three ASOs approved for treatment of DMD only increase dystrophin production by ∼0.9%.[99] It is impressive that such a small increase can see clinical benefit, but many ASOs fail clinical trials due to a lack of efficacy. Additionally, ASOs with lower potencies must be administered in high doses, leading to concerns about their long-term safety.[55]

- ADME: every ASO has a different biodistribution profile depending on route of administration (intravenous vs. intrathecal vs. subcutaneous, etc.). Chemistries can significantly impact plasma protein-binding of the ASO, with some such as nusinersen and mipomersen being >90% protein-bound in plasma while others such as eteplirsen and golodirsen are <40% protein-bound. While the mechanisms of metabolism (primarily via endonucleases) and excretion (primarily renal clearance) are known, there exists a knowledge gap in how (as well as how much) the current drugs on the market are taken up into tissue and cells after local or systemic distribution (as well as the

[r] ASO-protein interactions involved in cellular uptake and trafficking are further discussed in Chapter 5.

proportion that enters via productive vs. non-productive pathways). Most of what we know about uptake mechanisms and subcellular distribution of ASOs are studied via in vitro models.[98]

- Toxicity: in order to reach high plasma concentrations for sufficient uptake and clinical efficacy, the usual doses of "naked" ASO are very large (e.g. the ASOs treating DMD are dosed at 30 mg/kg once weekly).[98] This causes the primary side-effects of ASO to be hepatotoxicity (especially for those ASOs targeting hepatic gene expression) and kidney toxicity. Indeed, the FDA withdrew mipomersen in 2019 due to severe hepatotoxicity.[100] The adverse drug reactions (ADR) of ASOs were thought to be primarily due to off-target hybridisation and therefore, off-target protein effects. However, there is increasing emphasis on the protein-binding profiles of ASOs causing toxic effects.[101] For example, the tight binding of toxic ASOs to paraspeckle proteins is indicated in inducing apoptosis and nepthrotoxicity.[102]

Many different approaches have been explored to increase an ASO's cellular uptake; a select number are depicted in Figure 1.19. The most clinically promising ASO delivery strategies are:

- Conjugation (Figure 1.19, left)
  - Sugar conjugation: trivalent conjugation of *N*-acetylgalactosamine (GalNac) sugars to a therapeutic oligonucleotide is, for now, the only clinically validated method of receptor-mediated targeted delivery; GalNac binds avidly to the Asialoglycoprotein receptor (ASGPR), which is predominantly expressed on hepatocyte cell surfaces (Figure 1.19). Upon binding, the receptor internalises the therapeutic cargo via clathrin-mediated endocytosis. The siRNA-GalNac conjugate givosiran (one of six FDA-approved siRNA-GalNac conjugates) is approved for acute hepatic porphyria (AHP).[56] More than 10 ASO-GalNac and 15 siRNA-GalNac conjugates are in clinical trials.[103]
  - Lipid conjugation: cholesterol conjugation on the 3'-strand of siRNA has been

shown to increase lipoprotein binding and receptor-mediated targeted delivery to hepatocytes (a cholesterol-conjugated siRNA is shown binding to low-density lipoprotein (LDL) in Figure 1.19). A cholesterol-conjugated siRNA (RXi-109) is in Phase II clinical trials, and other lipid conjugates such as vitamin E are being investigated.[97,104]

– Antibodies and peptides: Conjugation strategies to target tissues other than the liver are in high demand. For example, Avidity Biosciences is developing an antibody-siRNA conjugate which is entering a Phase III clinical trial for myotonic dystrophy type 1. Conjugation to cell-penetrating peptides (CPPs) is also in clinical trials (the use of a Tat-CPP is depicted in Figure 1.19); Sarepta Therapeutics is investigating a peptide-conjugated-PMO ASO (PPMO), currently in Phase II clinical trials for the treatment of DMD.[97]

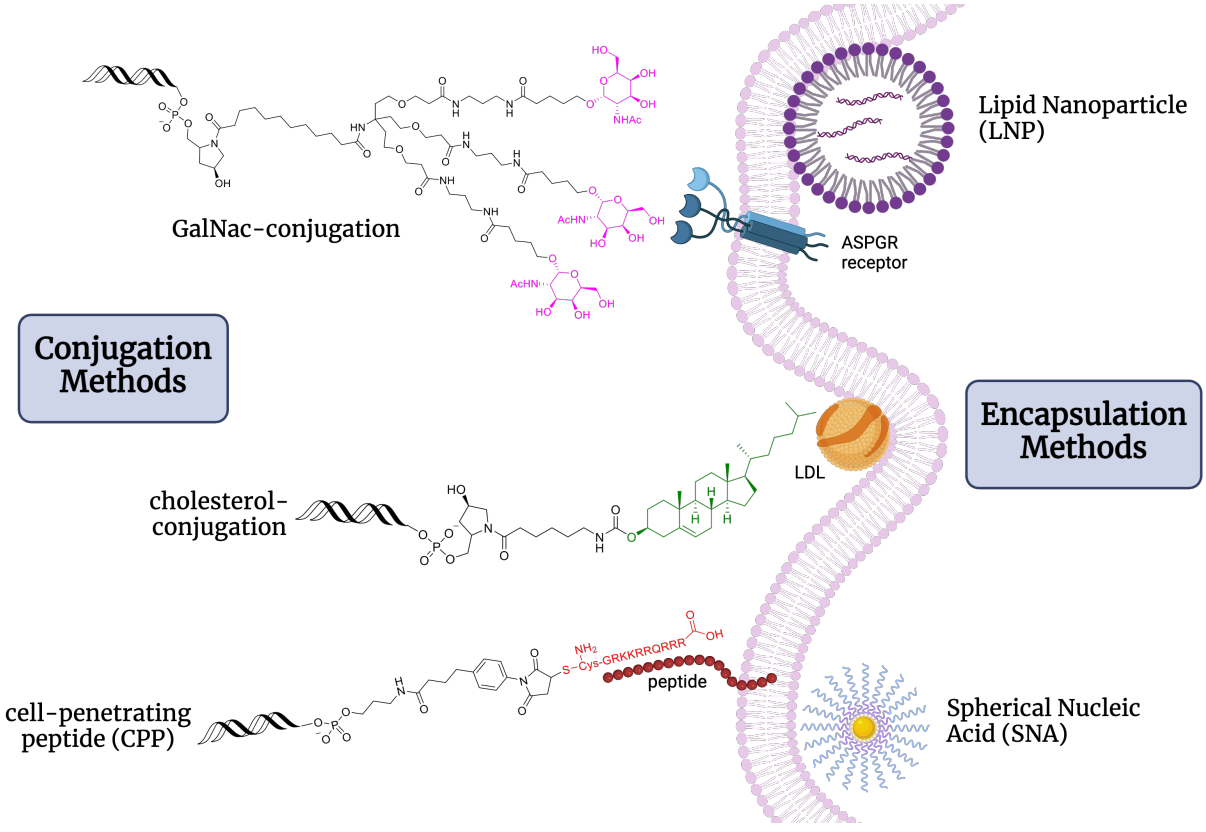


Figure 1.19: Conjugation methods (left) and encapsulation methods (right) explored in the literature and in the clinic for the delivery of therapeutic oligonucleotides to cells.

- Encapsulation (Figure 1.19, right)

- Lipid nanoparticles (LNP): spherical vesicles comprised of ionisable lipids, LNPs are a highly efficient and novel drug formulation which have been applied to the delivery of small molecules and mRNA; famously, two COVID vaccines delivered mRNA to cells via LNP encapsulation.[105] Their application to oligonucleotide therapies is validated in the siRNA drug patisiran, which is delivered to cells in an LNP for the treatment of hereditary ATTR amyloidosis (depicted in Figure 1.19).[56]

- Exosomes, SNAs, nanocages: nano-structures such as exosomes (a type of extracellular vesicle), spherical nucleic acids (SNA) (a hydrophobic nanoparticle covered with hydrophilic ASOs, depicted in Figure 1.19), or DNA-based cages (structures which self-assemble and incorporate the ASO) are among the most recent and innovative ASO-delivery systems under investigation.[97]

While there is an unmet clinical need for delivery systems which are efficient, targeted, and high-loading, many ASOs on the market remain administered unconjugated and unencapsulated – essentially "naked".[54] Chemical modifications to the sugar, backbone, and nucleobase are solely relied upon to give the ASO drug-like properties. In this vein, medicinal chemistry of ASOs remains an area in which innovation can improve their clinical success.

1.3 Medicinal chemistry of ASOs: aims of this thesis

The power of ASOs vs. traditional small molecules is that the binding to the target is dictated by known Watson-Crick base-pairing. Amongst the most difficult processes in small molecule pharmaceutical development is designing a drug to fit a protein pocket – many iterations are required to optimise a molecule which selectively and potently (but not too potently) binds the protein of interest.[106] When a novel disease target is selected, this iterative Design-Make-Test-Analyse (DMTA) process to test chemical modifications must be conducted again, because proteins and their structures and conformations are incredibly diverse and complex.[106,107]

Successful ASO medicinal chemistry can be applied time and time again – the same PS modification is used in ASOs targeting a variety of genes. Although there can be sequence- and position-dependence, the most successful structural modifications impart an ASO with improved properties regardless of sequence (and therefore disease). Rather than begin the small molecule DMTA cycle again for every novel pharmacophore, ASO medicinal chemistry begins with the same DNA structure, and thus, is much more target- and disease-agnostic.

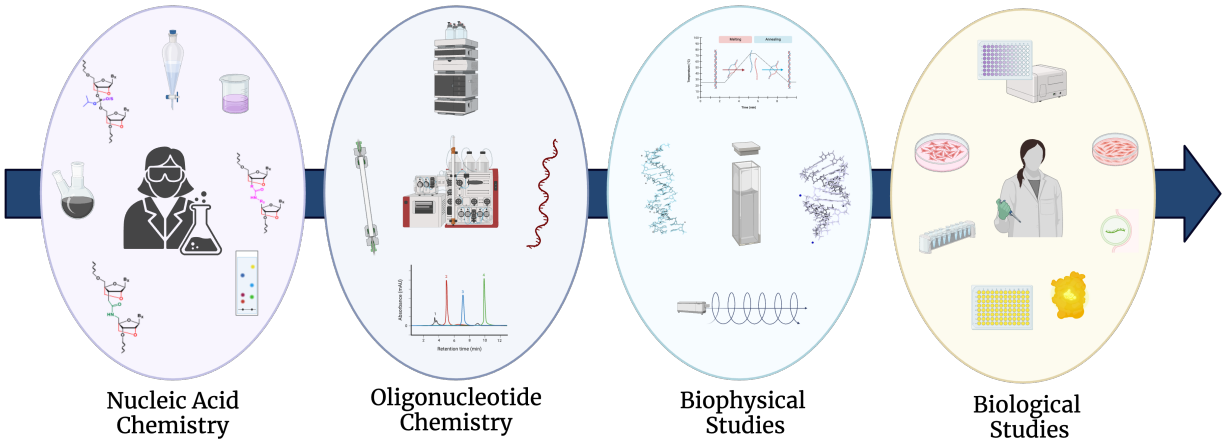


Figure 1.20: Schematic representation of the thesis workflow. Chapters 2, 3, and 4 follow this workflow for different LNA-neutral modifications. Chapter 5 will compare the most promising modifications in further biological work.

This thesis aims to contribute to ASO medicinal chemistry through the investigation of various neutral backbone structures in combination with the LNA modification. Each chapter follows a similar workflow, beginning with the synthesis of monomer nucleosides or dinucleotides, the synthesis of chimeric ASOs, and the study of their biophysical and biological properties (Figure 1.20). Overall, the primary aims of this thesis are:

1. Expand the ASO chemical space by combining neutral linkages and locked nucleic acid modifications in splice-switching ASOs.

2. Evaluate these structural modifications for their effects on the biophysical properties of ASOs, such as thermal duplex stability and global conformation.

3. Evaluate these structural modifications for their effects on the biological properties of ASOs, such as splice-switching activity and protein-binding profiles.

2

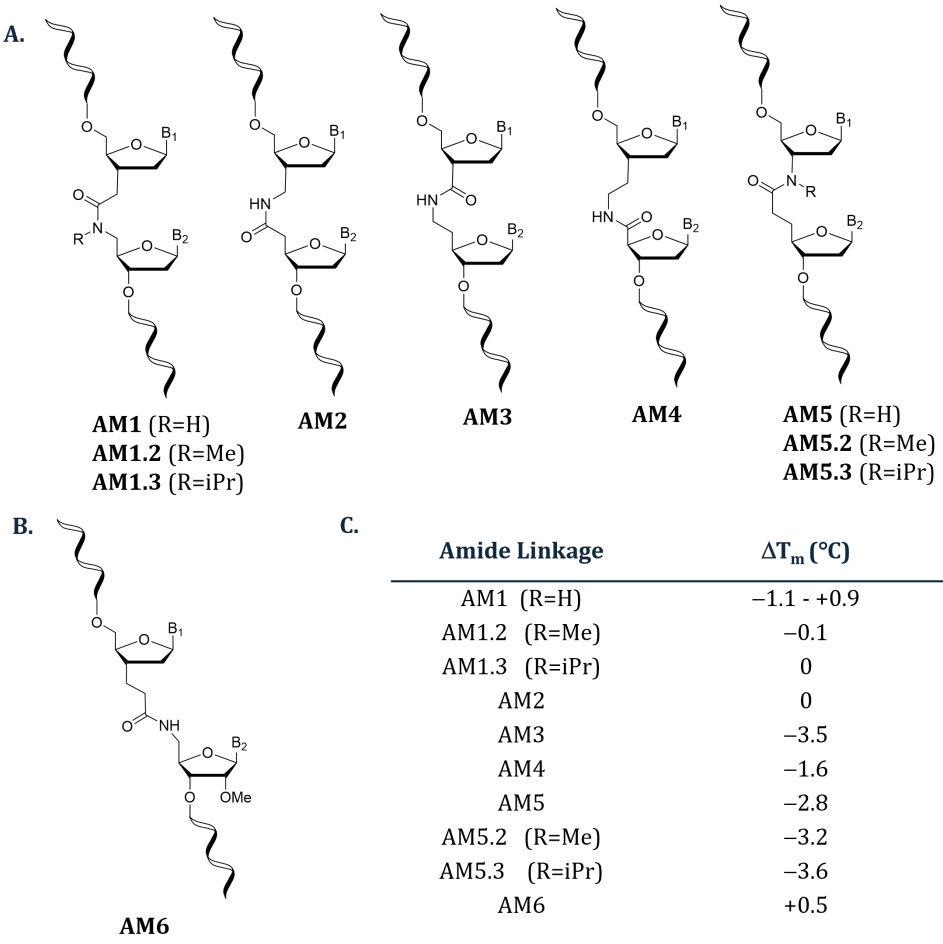
The synthesis and study of ASOs containing LNA-amide backbones

Contents

2.1 Introduction

This first experimental chapter will discuss the LNA-amide backbone modification − the LNA-amide linkage is characterised by linking two locked nucleic acid nucleosides by 5'-$CH_2CONHCH_2$-3' (Structure AM1, Figure 2.1). Various constitutional isomers of this four-atom, five-bond backbone have been explored by the literature (Figure 2.1A); introducing an amide linker as a neutral and achiral backbone between DNA nucleosides was first reported in 1993 by Idziak et a1.[108] Oligonucleotides containing this 5'-$CH_2CONHCH_2$-3' linkage demonstrated the ability to form stable duplexes with both DNA and RNA reverse complements. Since then, many research groups, including our own, have investigated the thermal stability of the AM1 linkage − melting temperatures reported for a single incorporation of the AM1 linkage within a DNA:RNA heteroduplex vary from –1.1 °C[108] to +0.9 °C,[67,109] demonstrating that oligonucleotides with an AM1 backbone can form equally stable heteroduplexes as DNA controls containing only phosphodiester backbones. Additionally, it was shown to have good sequence specificity and enzymatic resistance, by multiple accounts.[62,109,110]

Various constitutional isomers of the amide linkage, as depicted in Figure 2.1A, have previously been explored, especially by groups such as Lebreton's and Rozners'.[109] For example, the same four-atom linkage but with the orientation of the amide moiety inverted (AM2) was synthesised and is shown to have a similar duplex stabilisation as the original amide.[113] Both of these linkages adopt *trans* geometries and due to resonance around the CO-N amide bond, the linkage forms a planar structure (see Figure B.1 for amide geometries).[116] The restriction of free rotation around the CO-N bond makes it much more rigid than the natural phosphodiester bond, but ASOs containing amide linkages formed stable duplexes with RNA and the linkage was shown by crystal structure to be a close mimic of the phosphodiester.[52,62,117] Constitutional isomers of the four-atom linkage (e.g. AM3, AM4, and AM5), in which the amide bonds are immediately adjacent to the furanose rings, were destabilising − potentially indicating the importance of restricting the bonds with reduced rotational freedom to the middle of the linkage.[52,114] Interestingly, increasing the size of the substituent on the nitrogen atom (e.g. R=Me or R=iPr) was increasingly destabilising when the linkage



Amide Linkage	ΔT_m (°C)
AM1 (R=H)	−1.1 - +0.9
AM1.2 (R=Me)	−0.1
AM1.3 (R=iPr)	0
AM2	0
AM3	−3.5
AM4	−1.6
AM5	−2.8
AM5.2 (R=Me)	−3.2
AM5.3 (R=iPr)	−3.6
AM6	+0.5

Figure 2.1: A. Structures of various four-atom amide linkages incorporated into DNA/PO oligonucleotides in previous work. **B.** An example structure of a longer five-atom amide linkage previously explored in the literature.[111] **C.** Reported melting temperatures of oligonucleotides containing backbones AM1-6. ΔT_m are given as the reported ΔT_m for a DNA/RNA heteroduplex to the reference's own DNA phosphate control. References: AM1,[109,112] AM2,[113] AM3,[52] AM4,[114] AM5,[115] AM6.[111]

was directly adjacent to either ribose, as in the case of structures AM5.2 and AM5.3,[67,115] but less so when the amide was placed in the middle of the linkage, presumably due to steric effects.[52] Amide-containing backbones of longer linker length, for example AM6, have also been explored. Pallan et al. report this amide to be relatively stabilising, but it should be noted it is already combined with the stabilising 2'OMe modification on the 3'-end of the linkage, and when five additions were incorporated, it had a destabilising effect by –0.7 °C per modification.[a111]

[a]Different sequences and chemistries also affect melting temperature variation and thus, small changes should be interpreted with caution.

Given the promising thermal duplex stability of the AM1 linkage, its biological compatibility has been explored more extensively than other constitutional isomers. Its compatibility with ASO mechanisms of action such as siRNA and RNaseH-mediated RNA degradation has been investigated and is outlined below:

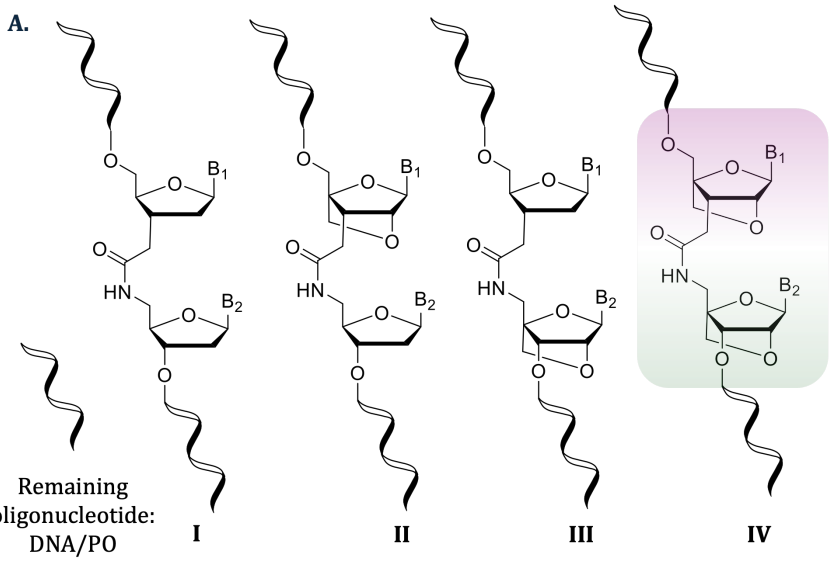
1. In RNAseH-mediated degradation: work by our own group has introduced consecutive amide linkages in the wings of gapmers and demonstrated that the RNaseH cleavage efficiency remains intact;[118] work by Ionis Pharmaceuticals has advanced gapmers containing a single AM1 linkage in the gap region to in vivo work − these therapeutic ASOs had good potency, which is very promising for its adoption in the clinic.[90]

2. In siRNA: the amide linkage was well tolerated at internal positions of both guide and passengers strands,[117,119] increased silencing activity at the 5'-end of the passenger strand, and even consecutive linkages had little effect on silencing activity.[120,121] Additionally, it was well tolerated in some positions of the guide strand and there was even an increase in silencing activity when placed between nucleosides in positions 10 and 12, the catalytic site of the Argonaute protein.[122]

3. In CRISPR: most recently, their incorporation into the protospace-adjacent motif (PAM) of the distal region revealed the amide was well-tolerated in the DNA-recognition region of CRISPR RNA.[123]

The amide is also a close bioisostere of the phosphate diester. A crystal structure of an amide-containing RNA:DNA duplex with RNaseH revealed that the carbonyl effectively mimics the PO orientation in a phosphodiester bond and that the nitrogen is able to act as a H-bond donor to a main-chain oxygen, stabilising the duplex-protein interaction.[122] The literature reports that amides are cumulatively destabilising − as additional amides were added to an oligonucleotide, the duplex stability was decreased (five additions resulted in –0.3 °C per modification, −1.5 °C total).[109] It was hypothesised that combining a biocompatible linkage such as the amide with a stabilising ribose modification such as the LNA, introduced in Chapter 1, would result in a linkage that confers an oligonucleotide with higher thermal

A.



B.

Amide Linkage	DNA Complement		RNA Complement	
	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)
unmodified control	63.5	–	61.4	–
I	61.2	−2.3	61.0	−0.4
II	63.1	−0.3	64.4	+3.0
III	60.9	−2.5	63.9	+2.5
IV	63.4	−0.1	66.5	+5.1

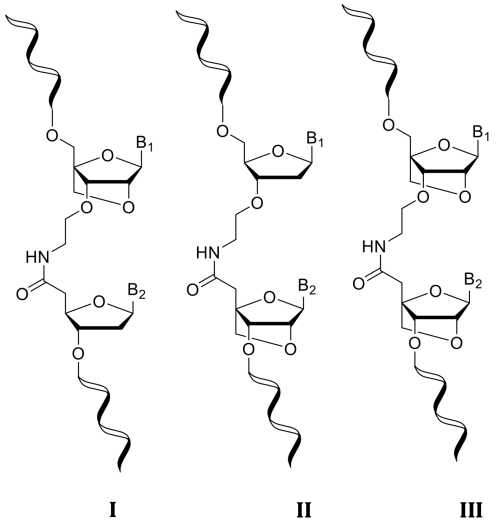
Figure 2.2: The LNA-amide combinations previously studied by the Tom Brown group.[62] **A.** Amide linkages in DNA/PO oligonucleotides containing zero, one, or two LNA moieties. **B.** Melting temperatures of DNA/PO oligonucleotides containing LNA-amide structures.

stability, global A-form conformation, and maintains its biocompatibility.

Therefore, the amide linkage (AM1) was further investigated by our own group, in combination with the LNA modification. The LNA was placed at the 5', 3', or both ends of the amide linkage, as shown in Figure 2.2A. Unsurprisingly, the modification with two LNA moities flanking the amide (Figure 2.2, structure IV) resulted in the highest thermal stability, increased the T_m of the duplex with the RNA complement by 5.1 °C and effectively induced no change (−0.1 °C) when in a duplex with DNA.[62] A single LNA on either side resulted in no change or moderate destabilisation against the DNA complement (−0.3-2.5 °C) but resulted in moderate stability against the RNA complement (+2.5-3.0 °C), which is promising as the thermal stability of an ASO:RNA duplex is more therapeutically relevant. It is interesting to note the amide was stabilising in duplex II, where LNA is placed on the 5' nucleoside such that the modification is "inside" of the linkage; the conformational rigidities

of both the LNA and the amide did not conflict to destabilise the duplex.

A.



B.

Amide Linkage	DNA Complement ΔT_m (°C)	RNA Complement ΔT_m (°C)
I	−8.0	−4.0
II	+1.5	+3.5
III	−6.0	±0

Figure 2.3: **A.**The longer 6-atom amide linkages explored in combination with LNA by Pallan et al.[111] **B.** The ΔT_m of linkages I-III against the DNA and RNA complementary sequences (remaining oligonucleotide is DNA/PO modified).

Combining LNA with an amide backbone has also been investigated in the literature by others; Wengel investigated a 6-atom amide-containing linkage (5'-$OCH_2CH_2NHCOCH_2$-3') in combination with LNA (Figure 2.3A).[124] Placing the LNA solely on the 5'-monomer was destabilising, against both DNA (−8 °C) and RNA (−4 °C), but when incorporated solely on the 3'-monomer, the LNA was able to "rescue" the thermal stability of the oligonucleotide which became more stable against RNA (+3.5 °C). Flanking the amide on both ends with LNA gave intermediate stability; while it was destabilising against DNA (−6 °C), it induced no change against RNA. Potentially the amide linkage alone was so conformationally unfavourable that incorporating LNA on the 5'-monomer is not as beneficial as on the 3'-monomer due to the steric clash within the linkage.[124] In this vein, LNA may be added at the 3'-end of conformationally destabilising linkages to "rescue" their stability − alternatively, in cases where the linker already mimics the natural phosphodiester such as AM1, flanking the linkage on either side can dramatically increase stability. The impact of LNA flanking

a destabilising linkage on both ends vs. solely the 3'-end is further discussed in Chapter 3 when combined with the carbamate linkage.

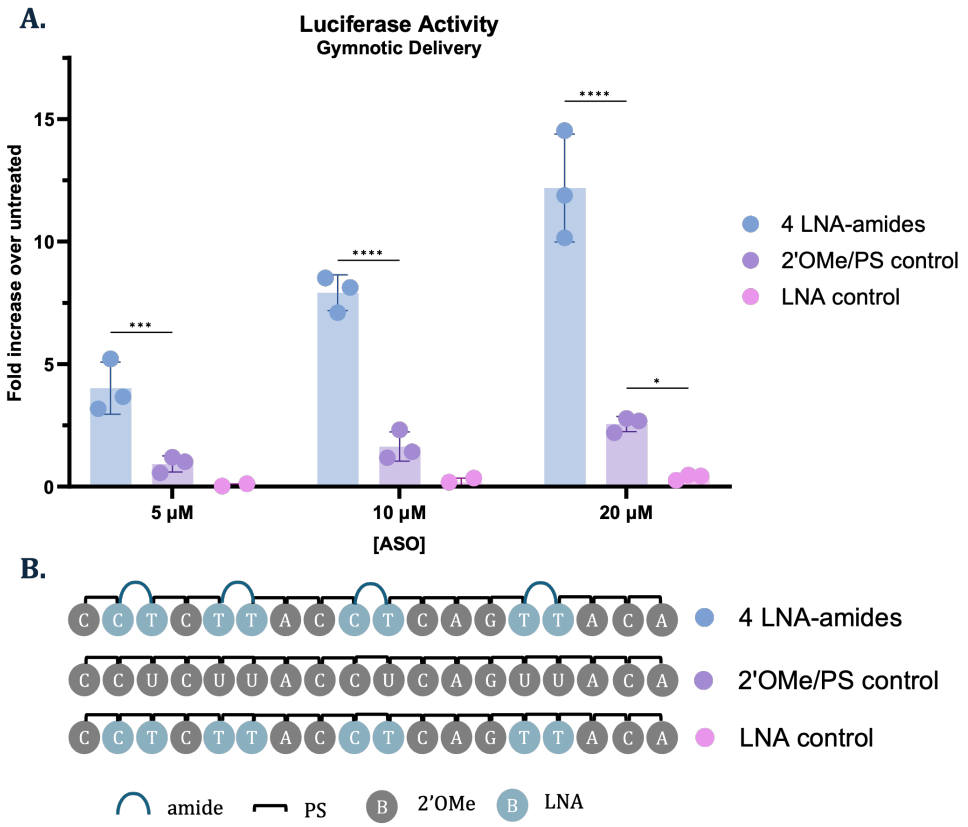


Figure 2.4: **A.** Splice-switching activity of ASOs luciferase reporter assay, delivered by gymnosis. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using the 2'OMe/PS control as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. Each dot represents one distinct technical replicate ($n = 3$). Data reproduced with permission from Baker, et al.[62] **B.** Schematic representation of the chimeric oligonucleotides containing 2'OMe, LNA, PS and amide modifications.

Given the promising thermal stability of duplex IV (Figure 2.2, from here known as the LNA-amide modification), it was incorporated into chimeric oligonucleotides in which the remaining ASO contained the 2'OMe ribose modification and phosphorothioate backbones (2'OMe/PS) (Figure 2.4B). ASOs containing 2'OMe/PS modifications are distinguished as "therapeutic" due to their ability to be applied in biological systems, differentiating them from DNA/PO oligonucleotides which are not stable or sufficiently enzyme-resistant for in vitro work. In a preliminary splice-switching reporter assay conducted by our group,

therapeutic ASOs containing four incorporations of the LNA-amide were found to have increased activity, both when delivered with transfection agents and without (naked delivery/gymnosis) (Figure 2.4A). The LNA-amide-containing ASOs are significantly more active at lower transfected concentrations than the 2'OMe/PS control, signifying their increased potency. Most importantly, they are more active across all doses when delivered by gymnosis (see gymnotic activity in Figure 2.4).[62] These preliminary findings were exciting and merited further investigation of the LNA-amide modification.

2.2 Aims of the Chapter

The aims of this chapter are to synthesise therapeutic ASOs containing LNA-amide backbones to further investigate the promising biological properties demonstrated thus far. The chapter aims to:

1. Synthesise locked 3'-carboxylic acid monomers and locked 5'-amino phosphoramidites in order to install the LNA-amide backbone using on-resin amide coupling (termed the monomer approach).

2. Optimise the on-resin amide coupling. A dimer approach is also considered in which the LNA-amide linkage is installed via the synthesis of LNA-amide dinucleotide phosphoramidites.

3. Synthesise ASOs in which the number of LNA-amide incorporations is varied (from one to four incorporations) and the position of the LNA-amide modification is varied within the oligonucleotide to preliminarily investigate positional effects.

4. Evaluate the biophysical properties of ASOs containing LNA-amide linkages through duplex melting temperature and circular dichroism.

5. Evaluate the biological activity of ASOs containing LNA-amide linkages in a luciferase reporter assay.

2.3 Synthesis of LNA-amide nucleosides

Previous reports of ASOs containing amide backbones have used one of two methods – coupling of a dimer phosphoramidite during standard solid-phase oligonucleotide synthesis[108,124] or via solid-phase peptide coupling of two monomers.[118,125,126] The report of automated solid-phase amide coupling on an oligonucleotide synthesiser by the Rozners group in 2022 is especially exciting, as it paves the way for this linkage to be produced on an industrial and clinically-viable scale.[125] However, they report 91-95% coupling yields between RNA monomers, which is not sufficient for a high-yielding synthesis of longer oligonucleotides; furthermore, it is well-known that LNA monomers are less reactive and more difficult to handle.[77]

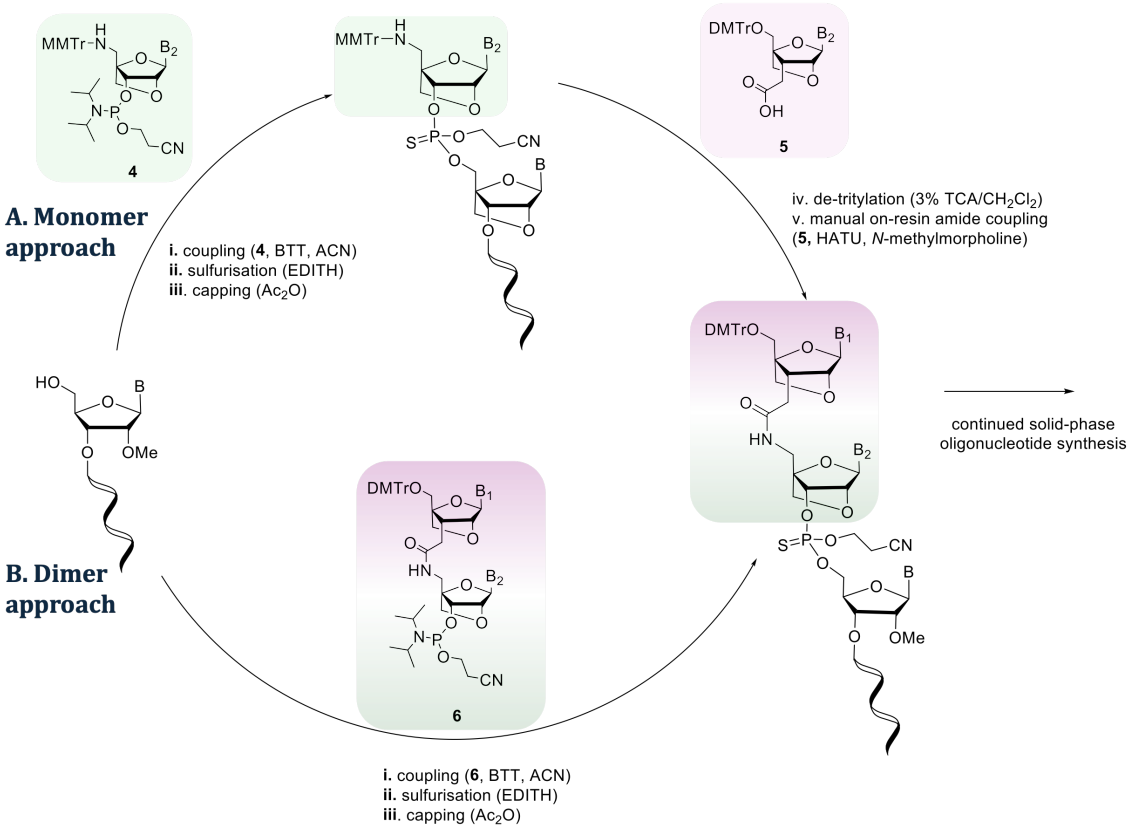


Figure 2.5: Methods to install the LNA-amide modification in chimeric 2'OMe/PS ASOs via the monomer approach (**A**) or dimer approach (**B**).

To construct the therapeutic ASOs, these two different approaches to install the LNA-amide linkage were considered. The monomer approach (Figure 2.5A), installs the amide linkage via a manual peptide coupling step. First, a 5'-MMTr-protected locked amino phosphoramidite

4 is coupled to the growing oligonucleotide (step i., Figure 2.5A). Standard solid-phase synthesis steps proceed: sulfurisation (ii.), then capping (iii.), followed by detritylation (iv.). The resin is removed from the synthesiser and the locked 3'-carboxylic acid monomer **5** is manually coupled using a peptide coupling step (v.). The resin is then returned to the synthesiser for continued solid-phase synthesis. The dimer approach (Figure 2.5B) installs the LNA-amide linkage directly via the coupling of a 5'-DMTr-protected LNA-amide dimer phosphoramidite **6** (i.), followed by the normal solid-phase synthesis steps, sulfurisation (ii.) and capping (iii.).

The monomer approach was first pursued with the aim of optimising the amide coupling efficiency to reach that of solid-phase synthesis (>95%).[127] While the dimer approach requires 16 distinct phosphoramidites to synthesise all nucleoside combinations, the monomer approach has fewer synthetic requirements, requiring only the four locked 5'-MMTr-protected-amino phosphoramidites and four locked 3'-COOH monomers (Figure 2.5A). If a near-quantitative and efficient solid-phase amide coupling protocol can be achieved, this method could, in future, be automated on the same synthesiser as SPOS. This would allow the LNA-amide linkage to be as readily installed as other modified nucleobases, using only 8 different nucleosides.

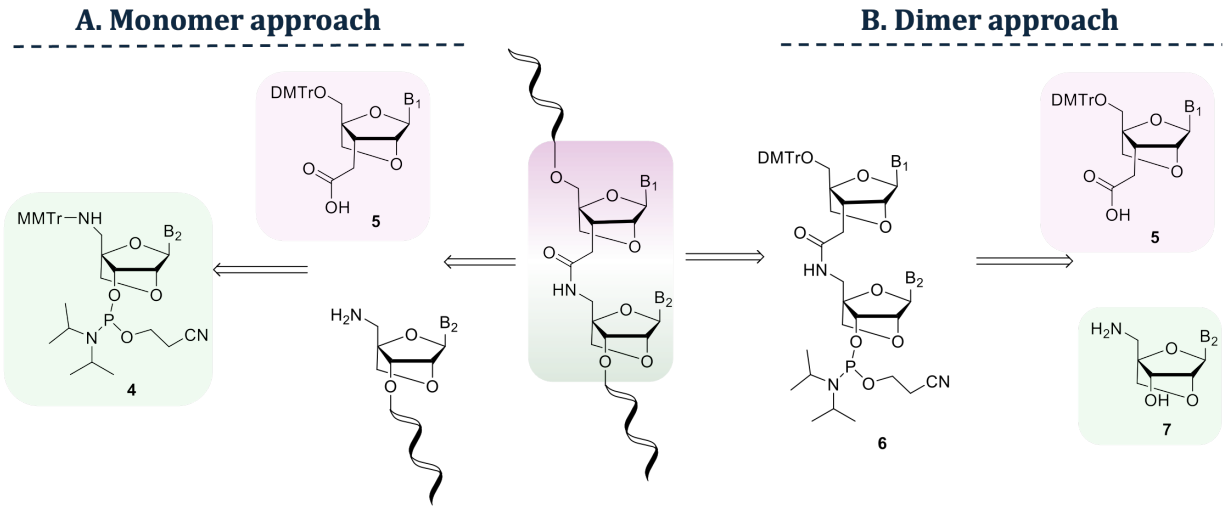


Figure 2.6: Retrosynthetic schemes to the monomer (**A.**) and dimer (**B.**) approaches to installing the LNA-amide linkage.

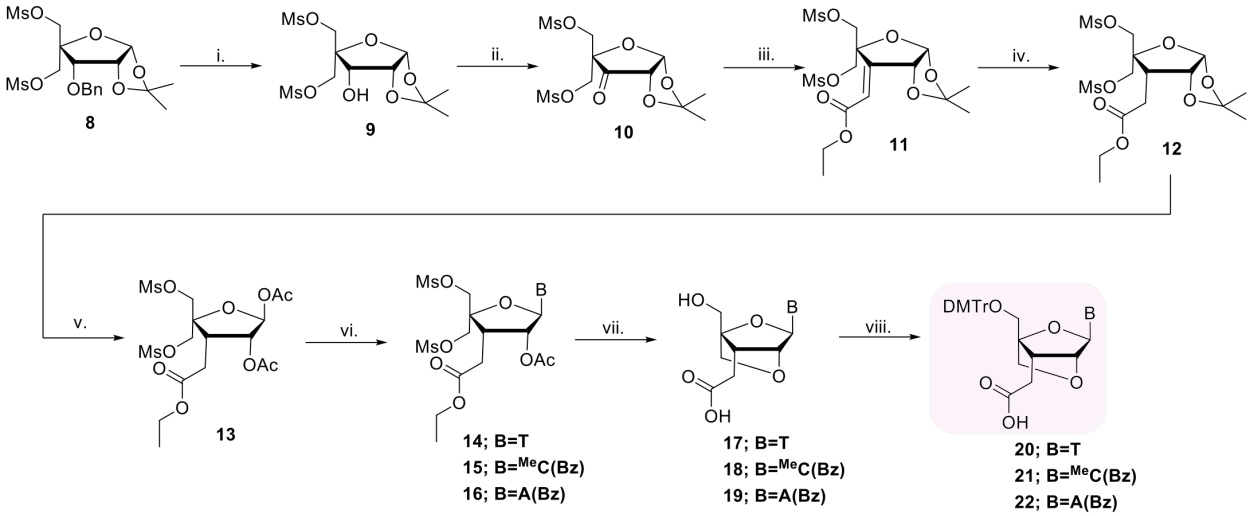
The synthesis of 5'-monomer building block **5** is required by both approaches, while the 5-MMTr-protected-amino phosphoramidite **4** can be synthesised from the 5'-NH_2 monomer **7**. If a switch to the dimer approach is required, the 3'-carboxylic acid **5** and intermediate 5'-NH_2 monomer **7** can be coupled in solution and a phosphitylation of the dinucleotide would give SPOS-compatible dimer **6** (Figure 2.6).

2.3.1 Synthesis of 3'-carboxylic acid LNA monomers

The synthesis of a DNA version of the 3'-COOH monomer **5** was reported in 1994, but suffered from a few drawbacks, among them: each route to synthesise all four nucleobases started with its own distinct nucleoside, the routes used toxic reagents such as OsO_4, were not compatible with benzoyl protecting groups required for adenine, and experienced difficulty reacting G(dmf).[112] An improved route to synthesise the 2'-OAc equivalent of the 3'-carboxylic acid monomer **5** was reported in the literature for all four nucleobases.[128] This route shared a common intermediate prior to glycosylation, which reduced the synthetic requirements to achieve all four nucleosides. Here, a similar route is presented which includes the formation of a bicyclic ring, achieving all four LNA 3'-COOH nucleosides with SPOS-compatible protecting groups (e.g. 5'-DMTr) (Scheme 2.1).[62b]

Beginning with commercially available compound **8**, hydrogenation of the 3'-OBn gave alcohol **9** in 89% yield which was oxidised using Dess-Martin periodinane (DMP) to give ketone **10** in 92% yield. The 3'-C=C bond was installed using the Wittig reaction with (carbethoxymethylene)triphenylphosphorane, which selectively gave the *E*-stereoisomer (confirmed by X-ray crystal structure[62]) of alkene **11** in moderately good yield (64%). A second hydrogenation using Pd/C and H_2 gave the ester **12** in near-quantitative yield; until this step, all reactions proceeded in sufficient purity to avoid column chromatography and enable multi-gram synthesis. Ester **12** was produced as a single enantiomer with the 3'-C-C bond oriented on the α face of the ring (confirmed by X-ray crystal structure[62]); this is hypothesised to be due to sterics from the 3'-axial substituents on the bottom face of the

[b] I acknowledge Dr. Ysobel Baker for the design of this synthesis route.



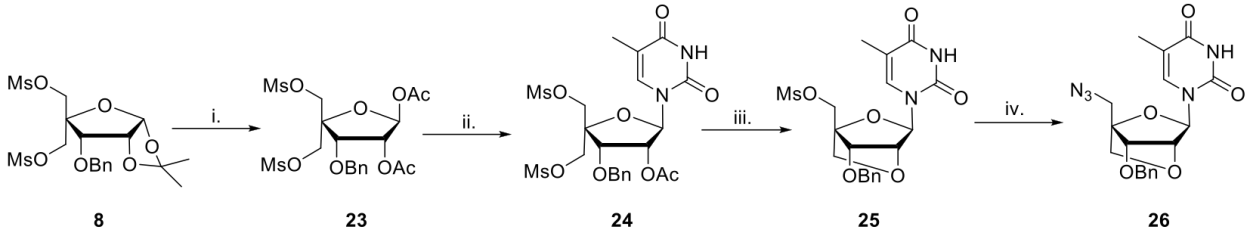
Scheme 2.1: Synthesis of locked 3'-COOH nucleosides: i. 20% $Pd(OH)_2/C$, NH_4CO_2, MeOH, 60 °C, 24 h, 89%; ii. DMP, CH_2Cl_2, rt, 24 h, 92%; iii. (carbethoxymethylene)triphenylphosphorane, CH_2Cl_2, rt, 24 h, 64%; iv. 10% wt. Pd/C, H_2, rt, 24 h, 97%; v. camphorsulfonic acid, Ac_2O, AcOH, 80 °C, 24 h, 96%, vi. nucleobase, BSA, TMSOTf, MeCN or CH_2Cl_2, reflux, 24 h, 83% (T), 55% (^{Me}C(Bz)), 63% (A(Bz)); vii. aqueous base (2 M NaOH or 1 M LiOH), 1,4-dioxane, rt then 55 °C, 3-5 h, 96% (T), 77% (^{Me}C(Bz)), 63% (A(Bz)); viii. DMTrCl, DMAP, Et_3N, pyridine, rt, 24 h, 70% (T), 17% (^{Me}C(Bz)), 58% (A(Bz)).

furanose ring directing the incoming H_2 to the β face. The acetal hydrolysis was conducted using a procedure reported by Arzel et al.,[129] which used a one-pot acid-catalysed ring opening and bis-acetylation of the resulting diol to give the 1,2-diacetyl **13** in 96% yield. It was important to avoid the use of aqueous acetic acid; they demonstrate an intramolecular ring closure can occur to form a lactone by-product, by the elimination of water.[129] Glycosylation with nucleobases thymine, N^4-benzoylmethylcytosine, and N^6-benzoyladenine using well-known Vorbrüggen conditions gave the nucleosides **14-16** in moderately good yields.[130] Attack of the nucleobase at the anomeric carbon results only in the β-anomer due to neighbouring group participation. Treatment with aqueous base hydrolysed both the 3'-ester and the 2'-ester to produce a 3'-carboxylic acid and a 2'-alkoxide which is able to attack the 5'-carbon and eliminate a mesylate in an intramolecular S_N2 reaction to form the 2'-4'-oxymethylene bridge (LNA). Subsequent heating then additionally hydrolysed the remaining 5'-mesyl group to give locked carboyxlic acid monomers **17-19**, in yields ranging from 63-96%. A final protection of the 5'-hydroxyl group with 4,4-dimethoxytrityl chloride

gave monomers **20**-**22**.

2.3.2 Synthesis of 5'-amino LNA monomers

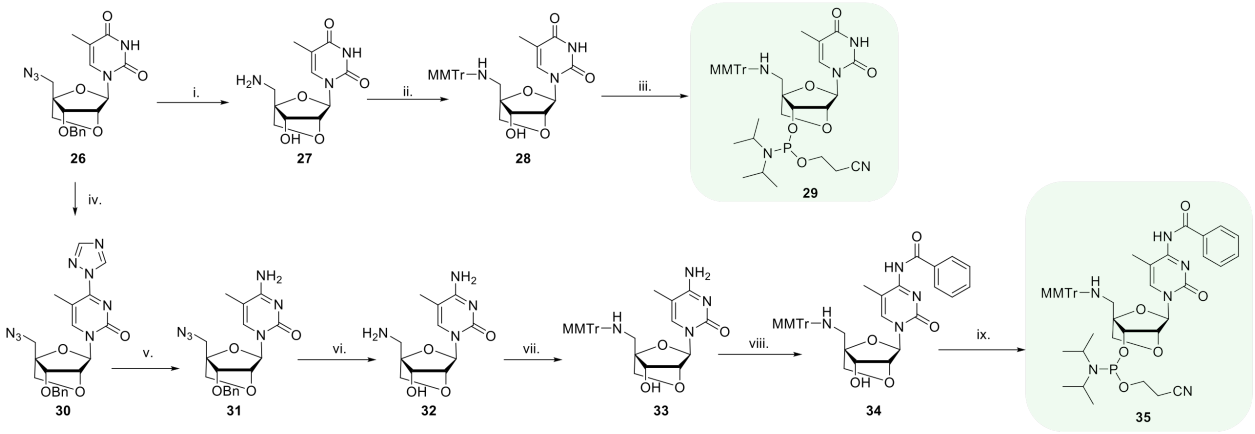
With the 5'-nucleosides in hand, the corresponding 3'-nucleosides of the amide linkage were synthesised. These 3'-monomers require a 5'-MMTr-protected amino group and a 3'-phosphoramidite group (see compound **4**, Figure 2.6) in order to be installed on the solid-phase oligonucleotide synthesiser.



Scheme 2.2: Synthesis of the locked 5'-N_3 thymine nucleoside: i. trifluoroacetic acid, Ac_2O, pyridine, rt, 24 h, 91%; ii. thymine, BSA, TMSOTf, MeCN, reflux, 24 h, 74%; iii. 2 M NaOH, 1,4-dioxane/H_2O (1:1, v:v), rt, 24 h, quant.; iv. NaN_3, DMF, 50 °C, 24 h, 79%.

The routes to synthesise the 5'-amino LNA pyrimidine monomers thymine and ^{Me}cytosine share a 5'-azide intermediate **26** (Scheme 2.2). Beginning with the same commercially available building block **8** as in the acid route, the acetal ring in **8** was hydrolysed under acidic conditions and the diol acetylated to give **23** in 91% yield. Vorbrüggen conditions with thymine gave nucleobase **24** as a single β-enantiomer in 74% yield and a similar base-mediated "locking" strategy was employed to form the 2'-4'-oxymethylene ring (LNA) in **25** in quantitative yield. To introduce the 5'-amino functionality, the mesylate was treated with NaN_3 to furnish the 5'-azide **26** in 79% yield.

Here, the routes diverged for the synthesis of the thymine and ^{Me}cytosine phosphoramidites. The 5'-N_3 thymine monomer **26** can be converted to the 5'-N_3 ^{Me}cytosine monomer **31** through treatment with phosphorus oxychloride and 1,2,4-triazole to give the triazole derivative **30**, which was used crude and converted to the free amine with aqueous ammonium hydroxide in 77% yield over two steps.[131] A "global hydrogenation" on the 5'-N_3 nucleosides **26** and **31** using palladium hydroxide on carbon reduced the azide and deprotects the



Scheme 2.3: Synthesis of locked 5'-NH_2 phosphoramidites: i. 20% $Pd(OH)_2/C$, NH_4CO_2, MeOH, reflux, 4 h, quant.; ii. MMTrCl, pyridine, 2 h, rt, 45%; iii. 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, CH_2Cl_2, rt, 1 h, 85%; iv. $POCl_3$, 1,2,4-triazole, Et_3N, MeCN, 0 °C to rt, 4 h, used crude; v. aq. NH_3, dioxane, rt, 24 h, 77% over two steps; vi. 20% $Pd(OH)_2/C$, NH_4CO_2, MeOH, reflux, 5 h, quant.; vii. MMTrCl, DMF/Et_3N (1:1, v:v), rt, 4 h, 73%; viii. Bz_2O, pyridine, rt, 5 h, then 2 M NaOH, EtOH, rt, 2 h, 77%; ix. 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, CH_2Cl_2, rt, 89%.

3'-benzyl group to give the locked 5'-NH_2 monomers **27** and **32** in quantitative yields.[132] The 5'-amino groups were protected with 4-methoxytrityl chloride (MMTrCl) − the MMTr group is routinely employed to protect 5'-amines because it less labile than the commonly used dimethoxytrityl (DMTr) group (used routinely to protect 5'-OH groups).[c] The thymine and ^{Me}cytosine monomers **27** and **32** were protected at the 5'-NH_2 to give the 5'-MMTr-protected monomers **28** and **33**. The ^{Me}cytosine monomer required benzoyl protection at the N4-position of the cytosine to furnish the nucleobase with the appropriate protecting group for SPOS; treatment with benzoic anhydride and in situ base-mediated ester hydrolysis of a 3'-OBz intermediate gave the 5'-MMTr-protected ^{Me}C(Bz) monomer **34** in 77% yield. Finally, phosphitylation of the two 3'-hydroxyl monomers **28** and **34** using 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite afforded the phosphoramidites **29** and **35** in yields ranging from 85-89%.

It should be noted that a route to give the 5'-NH_2 adenine(Bz) monomer was attempted. It proceeded in a similar fashion from the glycoyslation of diacetyl **10** with N^6-benzoyladenine; however, removal of the 3'-benzyl group during the global hydrogenation step was unsuc-

[c] As amines are more basic than alcohols, they are more likely to become protonated (leading to deprotection) so the more stable protecting group is used.[133]

cessful (see Appendix Scheme B.1). This is supported by reports from Rozners et al., who are unable to remove a benzyl group from an N^6-benzoyladenine monomer.[132] They hypothesise that the adenine nucleobase complexed with the Pd catalyst as an electron-rich ligand and thus reduced its activity. A second benzoylation of the adenine nucleoase decreases the electron density, but even in this case they required 100% w/w Pd/C to achieve benzyl-deprotection. As locked nucleic acids are less reactive and the benzyl group in this case was very hindered by the conformationally rigid bicycle, hydrogenation using these harsh conditions remained unsuccessful. Going forward, the LNA-amide positions within a sequence were chosen so as to avoid requiring the 5'-NH_2-A(Bz) monomer.

2.4 On-resin amide coupling optimisation

It is important that a manual amide coupling, if widely implemented, should give coupling yields as close to that of automated peptide or oligonucleotide coupling as possible, which are in the range of 98-99% for automated synthesisers.[127] The amide coupling conditions reported previously by our group were investigated here.[62]

Coupling yields were determined by integrating the crude oligonucleotide HPLC chromatogram and calculating

$$yield = \left(\frac{m}{n+m}\right) * 100$$

where m = area under the successfully coupled oligonucleotide peak (extended by +1 nucleoside) and n = area under 5'-NH_2 oligonucleotide peak (unreacted) and failed sequences (represented in Figure 2.7B).

First, the published amide coupling conditions (described in Figure 2.7A) were tested on the simplest oligonucleotide "substrate" − an unmodified poly-thymine oligonucleotide containing DNA monomers and a phosphate backbone with an unprotected 5'-NH_2 locked monomer (T) at its growing 5'-end (Figure 2.7C, entries 1-4). Coupling with all four 3'-COOH locked nucleobases (T (**20**), ^{Me}C(Bz) (**21**), A(Bz) (**22**), G(iBu)[d], revealed good coupling yields for

[d]I thank Dr. Alice Kennett for the synthesis of this monomer.

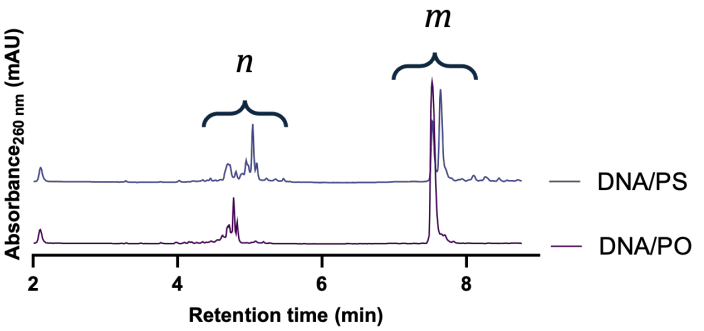
pyrimidine 3'-COOH monomers (99% (T), 95% (^{Me}C(Bz)), but these decreased when coupling the purine acids (91% (A(Bz), 84% (G(iBu)). When the oligonucleotide substrate was made more complex by incorporating other nucleobases in the sequence, the average amide coupling yield for the four LNA acid monomers dropped from 92% to 78% (Figure 2.7C, entries 5-8).

A.

Published Conditions:

- 10 eq. PyBOP, 10 eq. 3'-COOH monomer, 30 eq. N-methylmorpholine)
- 1 h, rt

B.



C.

Entry	Sequence	Modifications	Coupling Acid	Yield (%)	Average Yield (%)
1	NH_2-TTTTTTTTTTTT	DNA/PO	T	99	92
2			^{Me}C(Bz)	95	
3			A(Bz)	91	
4			G(iBu)	84	
5	NH_2-TACCTCAGTTACA	DNA/PO	T	80	78
6			^{Me}C(Bz)	75	
7			A(Bz)	78	
8			G(iBu)	74	
9	NH_2-TACCTCAGTTACA	DNA/PS	T	84	70
10			^{Me}C(Bz)	64	
11			A(Bz)	66	
12			G(iBu)	65	
13	NH_2-TACCUCAGUUACA	2'OMe/PS	T	73	71
14			^{Me}C(Bz)	70	
15			A(Bz)	70	
16			G(iBu)	70	

Figure 2.7: A. Published on-resin amide coupling conditions.[62] **B.** Crude HPLC traces of 3'-COOH G(iBu) monomer coupling to oligonucleotides of increasing backbone and ribose complexity. The area under failed coupling sequences are given as n and the area under the successfully coupled oligonucleotide is given as $n + 1$. **C.** Table of oligonucleotide sequences, 3'-COOH nucleosides undergoing amide coupling, and the corresponding yields.

In order to make the substrate oligonucleotide more relevant to the aims of synthesising chimeric ASOs, the backbone modification was swapped from the unmodified phosphate for phosphorothioate linkages, but the sugars remained as DNA. An additional decrease in yield was observed, with all LNA acid monomers other than thymine having a coupling efficiency

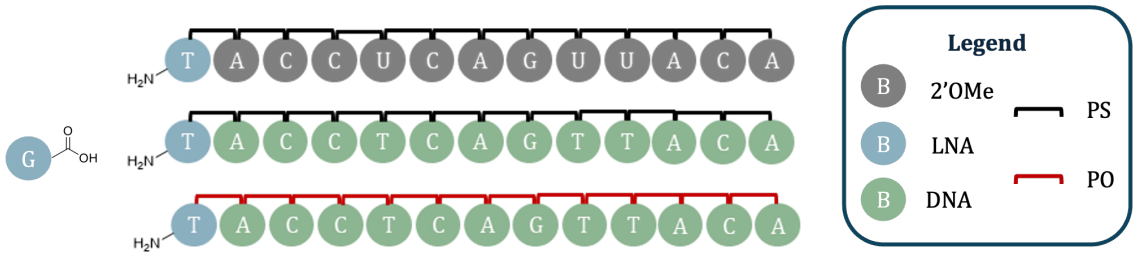
of ∼65% (Figure 2.7C, entries 9-12). An example of this lowered coupling efficiency as the backbone becomes more complex can be seen in the crude HPLC chromatograms of Figure 2.7B, which show the coupling between locked G(iBu) acid monomer and an oligonucleotide with a DNA/PS or DNA/PO backbone (entry 12 and entry 8, respectively). While the coupling of locked G(iBu) monomer to an unmodified DNA/PO oligonucleotide proceeded in 74% yield, the switch to the DNA/PS oligonucleotide reduced this yield to 65%.

Finally, the DNA sugars were changed to the 2'OMe monomers to give an oligonucleotide substrate with the most relevant modifications (2'OMe/PS). Interestingly, this produced a similar coupling yield to the DNA/PS sequence, with all LNA acid monomers coupling in approximately 70% conversion (Figure 2.7C, entries 13-16). Fluctuation in crude oligonucleotide purity is reported from synthesis to synthesis;[127] the slight increase in yield from the DNA/PS substrate to the 2'OMe/PS substrate is attributed to each "round" of synthesis being slightly different, rather than 2'OMe modifications improving the amide coupling efficiency.[127]

Although the conditions published in previous work are highly efficient for coupling a locked thymine acid to a DNA/PO oligonucleotide (entry 1),[62] complexifying the growing oligonucleotide to be more clinically relevant in its modifications and including the coupling of other carboxylic acid nucleobases can see a drop to 64%. Assuming 99% coupling for all other phosphoramidites, this would reduce the yield of a 2'OMe/PS 20-mer from 83% to 53% with a single LNA-amide incorporation. The most important factors which reduced the amide coupling yield were whether the oligonucleotide was *a)* composed of a mixed sequence, *b)* containing a phosphorothioate backbone, and *c)* the nature of the nucleobase acid being coupled. In order to achieve a robust method which is high-yielding and can be used on any sequence, of any modification, and with any of the nucleobases, further optimisation was needed.

Across all conditions, the coupling of the locked 3'-COOH G(iBu) monomer always observed the lowest coupling conversion. Thus, the concentration and time of the reaction were varied using this monomer as a benchmark (Figure 2.8). Across all variations of the growing

A.



B.

Entry	Sequence	Modifications	Special Conditions	Time	Acid Conc.	Yield (%)
1	NH_2-TACCTCAGTTACA coupling with G(iBu) acid	DNA/PO	None	1 h	25 mM	74
2				2 h	25 mM	78
3				1 h	50 mM	78
4		DNA/PS	None	1 h	25 mM	65
5				2 h	25 mM	71
6				1 h	50 mM	71
7		2'OMe/PS	None	1 h	25 mM	70
8				2 h	25 mM	75
9				1 h	50 mM	74
10			Pre-activation for 30 min	1 h	25 mM	77
11			15 eq PyBOP		25 mM	74
12			50 °C		25 mM	74

Figure 2.8: **A.** Schematic representation of the amide coupling between the 3'-COOH G(iBu) monomer and oligonucleotides of increasing backbone and sugar complexity. **B.** Table of amide coupling conditions and corresponding yields. Yellow = conditions previously listed in Figure 2.7. Blue = best conversion achieved.

oligonucleotide sequence (DNA/PO, DNA/PS, 2'OMe/PS), doubling the reaction time or doubling the concentration of acid only moderately increased coupling yields by ~5% (Figure 2.8B, entries 2, 3, 5, 6, 8, 9). For the relevant 2'OMe/PS modified sequence, the most promising coupling conditions included a pre-activation period of 30 minutes wherein the 3'-COOH monomer, benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), and *N*-methylmorpholine are reacted prior to loading onto the resin (Figure 2.8B, entry 10). Heating and increasing coupling reagent equivalents (Figure 2.8B, Entries 11 and 12) did not make a large improvement in coupling efficiency. Despite these changes, the on-resin amide coupling yield only marginally improved (from 70% to 77% for a 2'OMe/PS backbone).

It was hypothesised that the difficulty in achieving efficient amide couplings was due to capping of the free amine through migrating acetyl or benzoyl protecting groups. This

would explain the dramatic decrease in yield when the oligonucleotide was changed from poly-thymine to a mixed sequence. To investigate this further, the 5'-NH_2 oligonucleotide was subjected to the amide coupling reaction conditions without the acid coupling partner in the following conditions: *1)* DMF only, *2)* DMF, *N*-methylmorpholine, *3)* DMF, *N*-methylmorpholine, PyBOP. The ratio of oligonucleotide ESI^- mass peaks corresponding to the acetyl amide vs. unprotected amine grew once the oligo was solvated in DMF. Addition of *N*-methylmorpholine led to a single ESI^- mass peak corresponding solely to the dead-end acetylated product, while addition of PyBOP also caused a second LCMS peak to appear whose mass corresponded to the benzoylated amide (23% benzoylated and 67% acetylated). Unfortunately, there was no way to avoid protecting group migration; the only viable protocol change was to limit the time between automatic detritylation and carrying out the manual on-resin amide coupling. Therefore, detritylation (3 min in 3% TCA/CH_2Cl_2) and an acetonitrile wash (5 min) were conducted on the synthesiser, followed by immediate treatment of the column with a solution containing the pre-activated locked acid monomer, coupling agent HATU, and *N*-methylmorpholine in DMF. In future, this time could be even further reduced if the amide coupling reagents were pre-activated and reacted with the column by an automatic hybrid oligonucleotide/peptide solid-phase synthesiser.

The splice-switching ASO sequence which is compatible with the luciferase assay, described in Chapter 1 and given in Figure 2.4, was fortunately rich in pyrimidines − there were enough positions in which the LNA-amide could be placed between thymine and ^{Me}cytosine(Bz) nucleotides. Therefore, the focus was placed on optimising the coupling of the locked 3'-COOH ^{Me}C(Bz) monomer with locked 5'-NH_2 T-functionalised oligonucleotide (Figure 2.9, entries 1-3) and the coupling of the locked 3'-COOH T monomer with a 5'-NH_2 ^{Me}C(Bz)-functionalised oligonucleotide (Figure 2.9, entries 4-6). Unsurprisingly, when the 5'-terminal nucleoside contains a benzoyl-protected nucleobase (as is the case with ^{Me}cytosine), the coupling yield using the coupling reagent PyBOP is reduced (80% vs. 85% in Figure 2.9A, entries 4 vs. 1). Additional amide coupling reagents were investigated which are known in the literature to be more reactive − this was the most effective change to the previously-reported

A.

Entry	Sequence (2'OMe/PS)	Coupling Acid	Coupling Reagent	Time	Acid Conc.	Yield (%)
1	NH_2-TCGGCUUACCU	^{Me}C(Bz)	PyBOP	1 h	25 mM	85
2			TOTU			90
3			HATU			95
4	NH_2-CGGCUUACCU	T	PyBOP	1 h	25 mM	80
5			TOTU			80
6			HATU			85

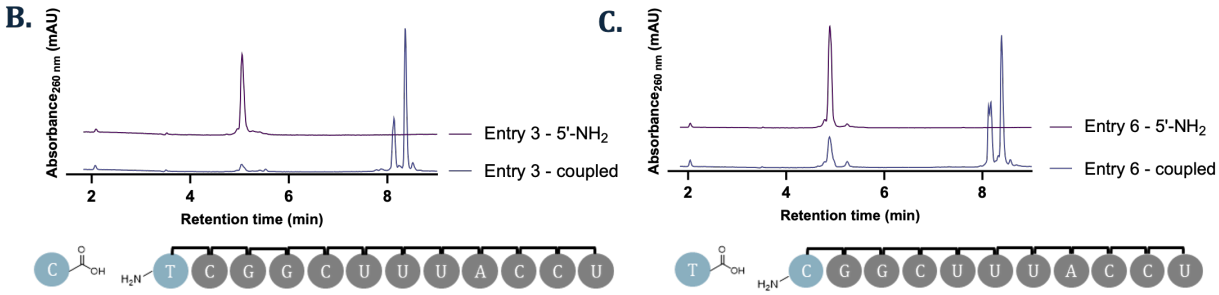


Figure 2.9: **A.** Table of oligonucleotide sequences, 3'-COOH nucleoside undergoing coupling, various amide coupling reagents, and the corresponding yields. **B.** Crude HPLC chromatogram of Entry 3 before and after coupling. **C.** Crude HPLC chromatogram of Entry 6 before and after coupling.

coupling conditions (Figure 2.10).[e] O-((Ethoxycarbonyl)cyanomethylenamino)-N,N,N',N'-tetramethyluronium tetrafluoroborate (TOTU) (Figure 2.10) produces an acyl oxime ester and has precedence in the construction of PNA – it was more efficient than PyBOP for coupling the locked ^{Me}C(Bz) acid (90%) but was still second-best.[134] The use of hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) (Figure 2.10) produced the best results, increasing yields for the coupling of locked ^{Me}C(Bz) and T acids to 95% and 85%, respectively. HATU produces a more reactive azobenzotriazole (-OAt) ester (rather than the benzotriazole (-OBt) ester produced by PyBOP) and the pyrimidine nitrogen has been shown to be able to stabilise the incoming amine which can increase amide coupling yields.[135]

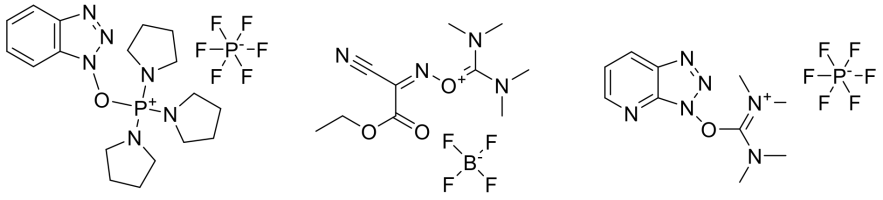


Figure 2.10: Structures of amide coupling reagents used in the on-resin amide optimisation of Figure 2.9.

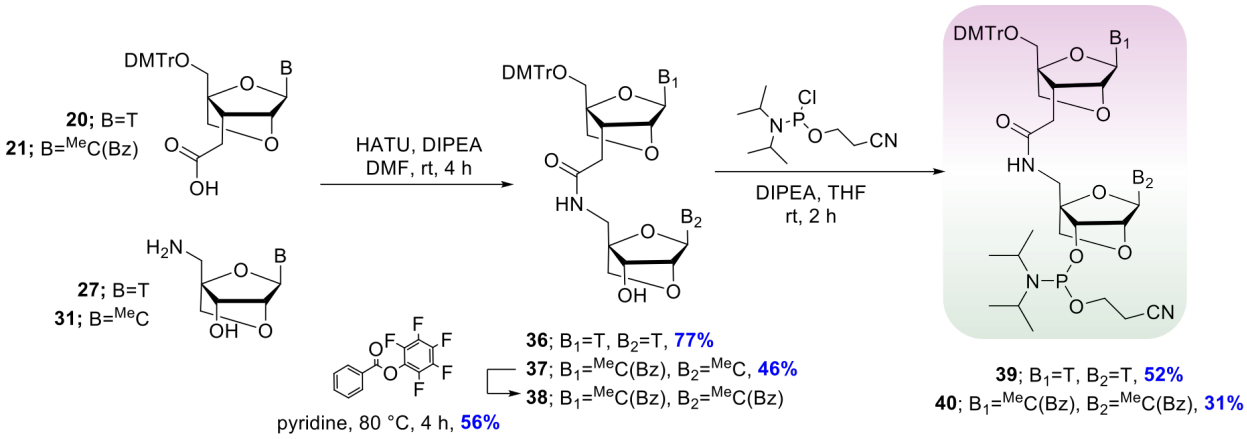
[e]I thank Dr. Alice Kennett for conducting these six on-resin amide coupling reactions.

By changing the coupling reagent to HATU, pre-activating the acid prior to treatment of the resin, and being mindful of protecting group migration, coupling yields between relevant nucleobases ^{Me}C(Bz) and T were improved to 85%+ (seen in the crude HPLC chromatograms of Figure 2.9B and C). In future work, the coupling between 5'-NH_2 and 3'-COOH purines would be carried out with these conditions to investigate the potential improvement in their amide coupling efficiency. However, an amide coupling efficiency of 85%, or indeed even 95%, is still below that of the standard phosphoramidite coupling of automated solid-phase oligonucleotide synthesisers (98-99%). The monomer approach also required more time-consuming involvement in the synthesis of the ASO, especially if the number of incorporations is increased; with coupling yields below near-quantitative, the crude mixture of failed sequences would become increasingly complex as more LNA-amide linkages were included. Therefore, the dimer approach was considered as an alternative method of installing the LNA-amide backbone in chimeric therapeutic oligonucleotides.

2.5 Synthesis of LNA-amide dinucleotides

The benefits of the dimer approach are that the solid-phase synthesis can go forward without interruption or manual coupling; it was hypothesised that the resulting crude ASO would have fewer failed sequences (a result of dead-end acetylation or benzoylation observed in the monomer approach) and therefore the purification would be more facile. However, to be considered for clinical adoption and to outweigh the reduced synthetic demands of the monomer approach, all 16 dimers must be synthetically accessible and their synthesis relatively facile.

To synthesise a T-T LNA-amide dimer (Scheme 2.4), the 3'-COOH thymine monomer **20** and the 5'-NH_2 thymine monomer **27** were coupled using HATU to give the dinucleotide **36**, isolated in 77% yield. The following phosphitylation to phosphoramidite **39** was performed in THF due to the limited solubility of alcohol **36** in CH_2Cl_2. The second dimer synthesised was the ^{Me}C(Bz)-^{Me}C(Bz) dimer **38**. Coupling of the 3'-COOH ^{Me}C(Bz) monomer **21** with the 5'-NH_2 ^{Me}C monomer **32** gave dimer **37** in 46% yield. Benzoylation of the unprotected N4-position of the 3'-nucleoside cytosine (B_2) using pentafluorophenylbenzoate



Scheme 2.4: Synthesis of LNA-amide dimer phosphoramidites **39** and **40**.

was conducted after dimer coupling to give dinucleotide **38** in which both nucleobases were benzoyl-protected. The second protection was conducted post-coupling to avoid undesired benzoyl protection of the 5'-NH_2 ^{Me}C monomer at the reactive N4-position. A phosphitylation of the dimer **38** to give phosphoramidite **40** proceeded in even further reduced yields (31%), as it was more prone to oxidation during handling (inert aqueous workup and inert column chromatography).

2.6 Oligonucleotide synthesis

With both the on-resin amide coupling conditions and dimer phosphoramidites in hand, four LNA-amide-containing ASOs were synthesised (Figure 2.11A). These ASOs are complementary to the abberant splice site of the luciferase assay. Positive controls include a fully 2'OMe/PS modified oligonucleotide (ASO-1) and an LNA control oligonucleotide (ASO-2), which contains LNA in the same eight positions as ASO-6 but no amide linkages. A scramble oligonucleotide (ASO-7) was also synthesised to ensure the sequence-specificity of the assay. As previously reported, the LNA-amide linkage was compatible with standard ASO deprotection methods (heating at 55 °C in concentrated ammonium hydroxide for 5 hours)[62] and the ASOs were purified, de-tritylated, and the Et_3N^+ salt was exchanged for Na^+, as described in Chapter 7.

A.

ASO#	Sequence (5' → 3')	Expected Mass (Da)	Found Mass (Da)	Modifications x = LNA-amide
ASO-1	CCUCUUACCUCAGUUACA	6096.3	6096.5	2'OMe/PS
ASO-2	**C**CT**C**TT**A**C**C**T**C**AG**T**T**A**CA	6192.4	6195.5	2'OMe/PS/LNA
ASO-3	CCUC**TxT**ACCUCAGUUACA	6066.4	6068.5	1 LNA-amide
ASO-4	CCUC**TxT**ACCUCAG**TxT**ACA	6036.5	6037.0	2 LNA-amides
ASO-5*	**CxC**UCUUA**CxC**UCAGUUACA	6036.5	6037.0	2 LNA-amides (CC)
ASO-6	C**CxT**C**TxT**AC**CxT**CAG**TxT**ACA	5976.6	5974.0	4 LNA-amides
ASO-7	UCACUCAGAUAGUUGAAGCC	6935.4	6935.3	scramble

B.

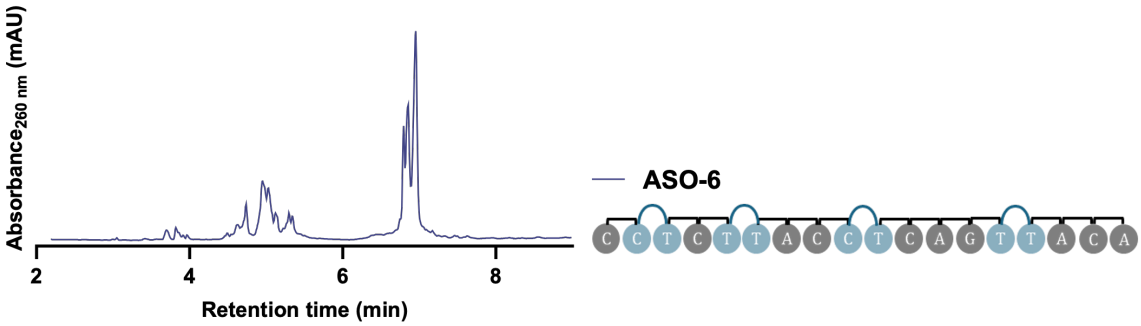


Figure 2.11: **A.** Table of chimeric 2'OMe/PS ASOs synthesised containing LNA-amide linkages. **B** = LNA, **x** = amide linkage. *synthesised via dimer approach. **B.** Crude HPLC chromatograms of ASO-6, containing four incorporations of the LNA-amide backbone, installed via the monomer approach.

2.6.0.1 Comparing the dimer vs. monomer approach

The crude HPLC chromatogram of ASO-6, synthesised via the monomer approach, revealed only 59.6% conversion to desired full-length oligonucleotide after four manual LNA-amide couplings (Figure 2.11B). The separation between desired oligonucleotide product and failed sequences was large enough to facilitate purification and so, it was isolated in good purity (>95%). However, ASO-6 was also very time-consuming to synthesise, as manual intervention in the SPOS cycle occurred every 2-3 nucleoside couplings. Together with the low crude conversion, these were reasonable grounds for a comparison to the dimer approach for less time-intensive and higher-yielding oligonucleotide synthesis.

ASO-4 and ASO-5 both contain two LNA-amide backbone incorporations, but were synthesised via these two different methods (Figure 2.12). The crude HPLC chromatogram of

ASO-4, synthesised by monomer coupling, gave a crude conversion of 78.3% to the desired full-length oligonucleotide, while ASO-5, synthesised by dimer coupling, gave a crude conversion of 83.4%. It should be noted that ASO-4 and ASO-5 are not directly comparable, as they include the LNA-amide linkage in different positions (TT vs. $^{Me}C^{Me}C$); however, it is likely that if the monomer approach was used to couple between ^{Me}C nucleosides, this crude conversion would be even lower, as the on-resin amide couplings between these two nucleosides was less efficient. The hypothesis that the dimer approach would result in a less complex and cleaner crude HPLC chromatogram was confirmed, as the monomer approach produced more failed sequences. Investigation of ASOs containing various amide-inspired linkages, discussed in Chapter 3, are synthesised via dimer phosphoramidites – the dimer approach is less time-intensive during SPOS and yields higher crude conversions to the desired oligonucleotide product.

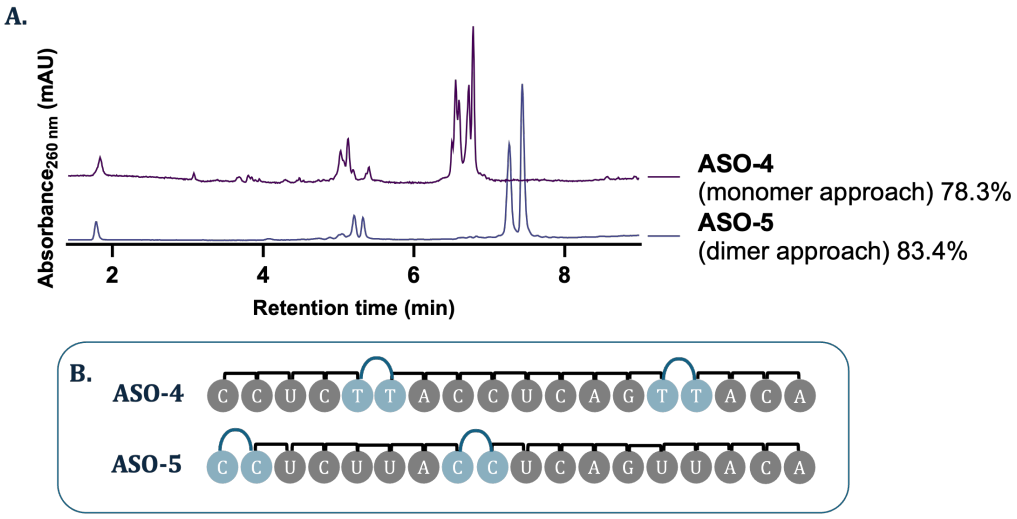


Figure 2.12: A. Crude HPLC chromatograms of the two ASOs containing 2 LNA-amide incorporations, synthesised via monomer approach (ASO-4) or via dimer approach (ASO-5). **B.** Schematic representations of the two ASOs containing LNA-amide backbones.

2.7 Biophysical properties of ASOs containing LNA-amide linkages

In order to investigate the duplex stability of chimeric ASOs containing LNA-amide backbones, the thermal melting temperatures of ASOs 1-6 were measured (Figure 2.13).[f] The LNA-amide modification, as previously reported, had a clear stabilising effect in all heteroduplexes it forms (Figure 2.13A). Against its DNA reverse complement, the melting temperature of the duplex increased by +3.4 °C for a single incorporation, and by ∼3.3 °C per additional incorporation, demonstrating an additive effect.

A.

Amide Linkage	DNA Complement		RNA Complement						
	200 mM		25 mM (full)		200 mM (fragment)		25 mM (fragment)		
	T_m (°C)	$\Delta T_{m\,(2'OMe)}$	T_m (°C)	$\Delta T_{m\,(2'OMe)}$	T_m (°C)	$\Delta T_{m\,(2'OMe)}$	T_m (°C)	$\Delta T_{m\,(2'OMe)}$	$\Delta T_{m\,(LNA)}$
ASO-1: 2'OMe/PS	54.8 ±0.1	-	62.0 ±0.2	-	54.3 ±0.2	-	41.0 ±0.1	-	-27.5
ASO-3: 1 LNA-amide	58.2 ±0.1	+3.4	65.0 ±0.1	+3.0	-	-	-	-	-
ASO-4: 2 LNA-amides	61.5 ±0.1	+6.7	67.5 ±0.1	+5.5	58.0 ±0.1	+3.7	44.8 ±0.1	+3.7	-23.7
ASO-6: 4 LNA-amides	68.0 ±0.1	+13.2	>75	-	>75	-	49.5 ±0.2	+8.5	-19.0
ASO-2: LNA control	>75	-	>75	-	>75	-	68.5 ±0.2	+27.5	-

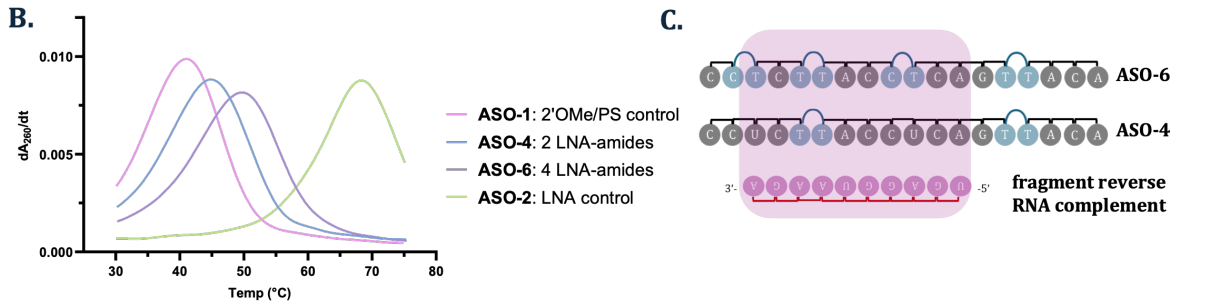


Figure 2.13: **A.** Melting temperatures of ASOs containing LNA-amide linkages against DNA and RNA targets. **B.** First derivative curves of melting temperatures for ASOs containing LNA-amide linkages against complementary RNA fragment (25 mM NaCl, 10 mM phosphate, pH = 7.0). **C.** Schematic representations of the ASO-6 and ASO-4 against the complementary RNA fragment.

When forming a heteroduplex with the RNA complement, the ASOs containing LNA-amide linkages were so stable that at standard salt concentrations (200 mM NaCl), all melting temperatures were above the measurable 75 °C. It is well-known that a reduced salt concentration

[f] For complementary sequences and oligonucleotide masses see Table B.1.

will reduce the melting temperature of a duplex, due to an increased electrostatic repulsion between strands.[136] Therefore, a series of UV melting studies were performed with reduced concentrations of NaCl until the ASO-1:RNA duplex melting temperature was sufficiently low enough that the more stable LNA-amide-containing ASOs would be melt below 75 °C (see Figure B.2 for the first derivative melting curves). At 25 mM NaCl, the 2'OMe/PS control (ASO-1) had a melting temperature of 62 °C − the LNA-amide containing ASOs which were more stable against the full RNA complement observed a similar (if slightly reduced) trend, at ∼+3 °C per linkage (ASO-6 with four LNA-amide incorporations and ASO-2 with equal LNA content remained too stable, T_m >75 °C. In order to ensure this trend was upheld at more physiological salt concentrations such as 200 mM NaCl, a fragment approach was also used,[62] in which the ASO was hybridised to a complementary 10-mer. The fragment spans an internal region of the ASO sequence, represented in Figure 2.13C. A single LNA-amide spanning the hybridised region maintained a stabilisation of +3.7 °C.

It is clear that the most stable ASO is the LNA control (ASO-2), an ASO with LNA in the same eight positions as ASO 6, but lacking amide linkages (see Figure 2.11A for LNA placement). To deepen the understanding of how the amide linker effects thermal stability in isolation, the reduced salt and fragment approaches were combined to achieve a melting temperature for the LNA control (ASO-2) (T_m = 68.5 °C). The difference in thermal duplex stability between the 2'OMe/PS control and the LNA control containing eight LNA modifications was ΔT_m = 27.5 °C, indicating that each additional LNA stabilises the ASO by ∼7 °C (in line with literature results).[137] The ΔT_m between the ASO with four LNA-amide linkages (ASO-6) compared to the ASO with identical LNA quantities (ASO-2) was −19 °C, or −9.5 °C per amide linkage (given the fragment only spans two amide linkages). Potentially, the conformational rigidity of the LNA and the amide structure are in fact clashing, but they are sufficiently compatible to maintain an overall stabilisation effect. In future work, a chimeric 2'OMe/PS ASO containing a single amide linkage between 2'OMe nucleotides could further evaluate the amide in isolation in therapeutic ASOs without the complication of considering the LNA's stabilising effect.

However, the relationship between thermal stability and biological activity is not a straightforward one,[90,101] and barring an extremely unstable linkage, there is very little way to predict based on T_m alone whether the modification will be significantly more biologically active. Important other factors will certainly play a role in whether a modification can improve therapeutic properties, including global helical structure. The circular dichroism (CD) spectra of ASOs 1, 2, and 6 against their DNA and RNA reverse complements are shown in Figure 2.14. The CD spectra of ASO-3 and ASO-4 are given in Figure B.3.

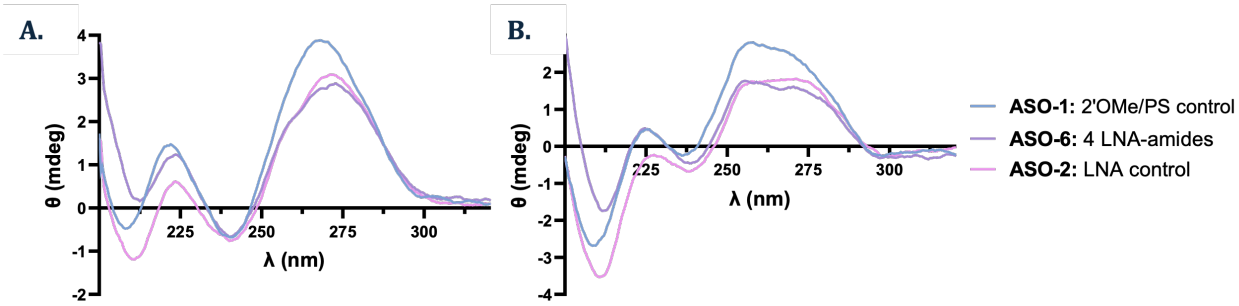


Figure 2.14: **A.** Circular dichroism spectra of ASOs against the DNA complement (200 mM Nacl). **B.** Circular dichroism spectra of ASOs against the RNA complement (25 mM NaCl).

Considering first the spectra of the ASO:DNA duplexes (Figure 2.14A), all three oligonucleotides demonstrated global B-type helices, with a maximum around 270 nm and a minimum at 245 nm. However, an additional small maxima (shoulder-like shift from 270 nm to 260 nm) was observed for the oligonucleotides containing LNA moieties which is hypsochromic, indicating some additional A-form character. This is consistent with the literature which observes LNA content inducing a B $\rightarrow$ A-form shift in ASO:DNA duplexes due to an increase in the number of sugars which are restrained to the C3'-endo conformation.[138]

The spectra of the duplexes against the RNA target (Figure 2.14B) showed a mixed A/B-form conformation, which is expected for an ASO:RNA heteroduplex.[139] While A-form helices are characterised by a maximum of 260 nm and an intense minimum of 210 nm,[50] A/B mixtures will have maxima that are blue-shifted toward 270 nm (B-like), while having additional negative bands at 245 nm (B-like) and 210 nm (A-like). In Figure 2.14B, the oligonucleotides have similar maxima at 260 nm, indicating the A-like conformation. Interestingly, there

is a second maxima or "plateau"-like peak that can be observed for the oligonucleotides containing LNA moieties whose shift is bathochromic (rather than hypsochromic), indicating some potential A → B-form character or structural perturbation. Both 2'OMe and LNA modifications encourage C3'-endo conformation (*N*-type sugars), so this difference may be due to the LNA (rather than the amide) causing some structural displacement. This is supported by a crystal structure solved by the group previously[g], in which an LNA-amide was incorporated into the DNA strand of a DNA:RNA heteroduplex; the 5'-LNA moiety was seen slightly shifted outward to stick out of the duplex, but the amide's backbone torsion angles remained unchanged (Figure 2.15A).[62]

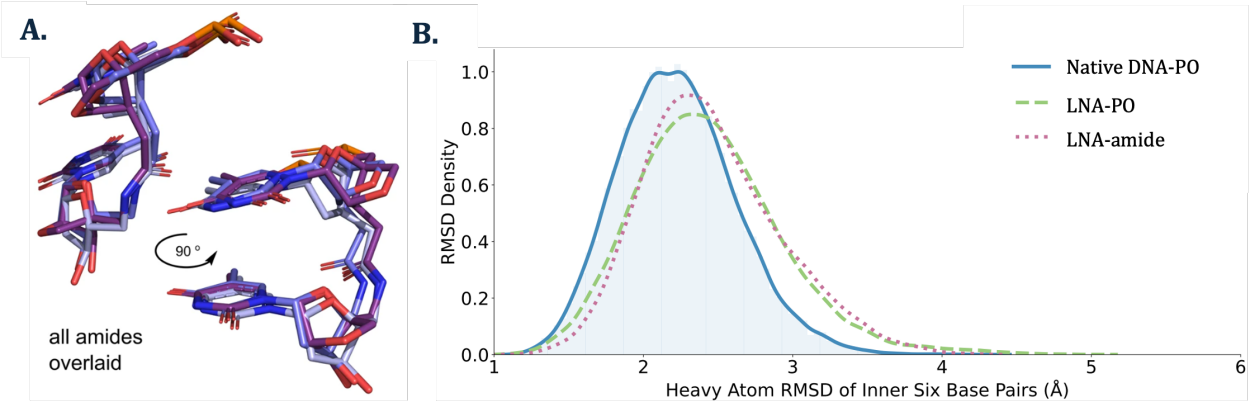


Figure 2.15: A. Crystal structures overlaid of DNA:RNA duplexes containing DNA-Am-DNA (light blue), LNA-Am-DNA (dark blue), and LNA-Am-DNA (purple) linkages. Reproduced with permission from Baker et al.[62] **B.** RMSD density histograms for the inner six base pairs of a DNA:RNA heteroduplex containing native DNA-phosphate, LNA-phosphate, and LNA-amide linkages.[a]

[a]I thank Martin Flerin, DPhil candidate supervised by Prof. Fernanda Duarte, for producing the RMSD distribution figure.

CD spectra are known to change based on slight variations in temperature and salt concentrations[50] – other techniques such as x-ray crystallography should additionally be used to confirm a modification's impact on global duplex conformation. The aforementioned crystal structure of the LNA-amide in a DNA:RNA heteroduplex were found to support Rozner's previous crystal structure of an amide demonstrating that the carbonyl bond is an excellent mimic of a PO bond in the phosphodiester linkage (Figure 2.15A).[62,117] Additionally, molecular dynamics (MD) simulations of a DNA:RNA heteroduplex containing a single LNA-amide

[g]Conducted by Dr. Ysobel Baker, in collaboration with Dr. Hall at the University of Reading.

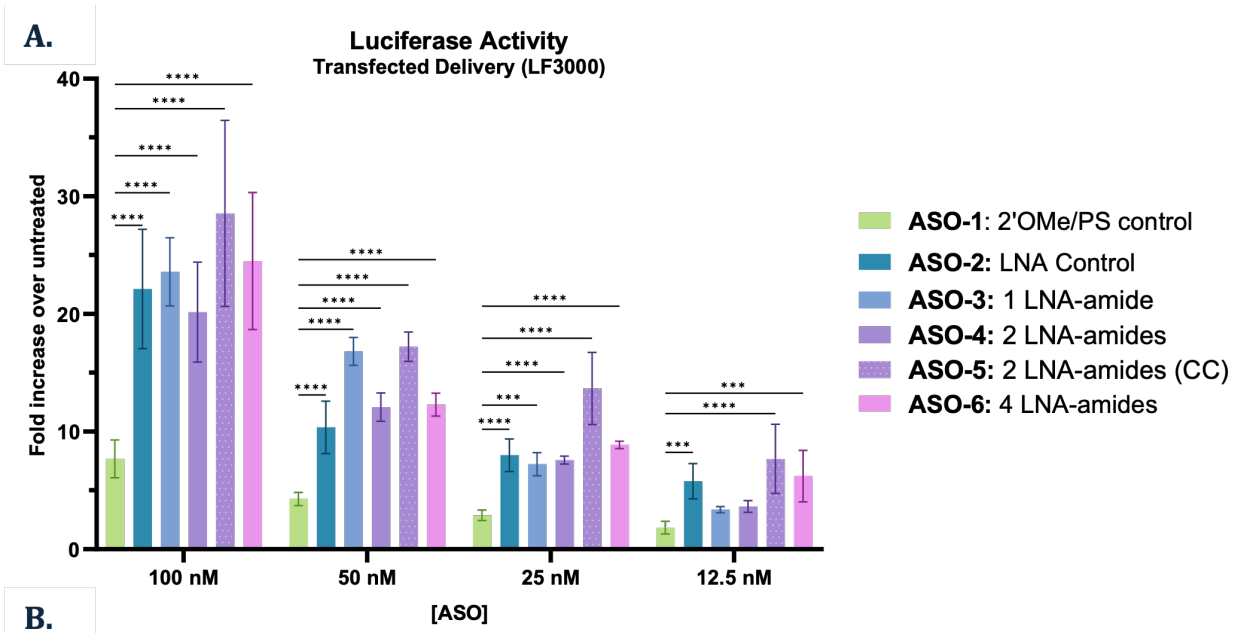
backbone was conducted in collaboration.[h] The root-mean square deviation (RMSD) distribution of the duplex containing an LNA-amide (Figure 2.15B) was almost identical to the RMSD distribution of a duplex containing two locked sugars but no amide linkage and very similar to the RMSD distribution of the unmodified native DNA/PO duplex . The similarity in unimodal RMSD distributions over a 1 µs simulation demonstrated that the inclusion of both the locked sugars and the amide backbone was not inducing a large conformational change in bond distances, nor were they inducing a populated second structure. It was promising to observe that computational results agree with experimental crystal structures − the use of MD simulations was again implemented through collaboration in Chapter 3.

2.8 Biological activity of ASOs containing LNA-amide linkages

In order to test the splice-switching activity of the LNA-amide ASOs, the luciferase reporter assay was used, described in Chapter 1. HeLa pLuc/705 cells were treated with ASOs 1-6 using the transfection agent Lipofectamine 3000 (Figure 2.16) and naked delivery (Figure 2.17), at four different doses. Activity of the ASO via transfection is a good indicator of the ASO's "inherent" activity, as transfection forms a liposome carrying the ASO which bypasses the cell membrane and delivers the ASO into the cell via endocytosis.[140] However, gymnotic activity is arguably a more important benchmark for advancement of novel chemistries in therapeutic ASOs, because this delivery method more closely represent in vivo or clinical methods where transfection reagents are not used. The scramble ASO-7 was also tested in a dose-response assay using both delivery methods and it had no activity (see Figure B.4) , demonstrating the sequence-specificity of the assay.

ASOs containing any number of LNA-amides were inherently more active than the 2'OMe/PS control (ASO-1) (Figure 2.16A). When transfected, they demonstrated significantly higher activity than ASO-1 at almost all doses (25 nM − 100 nM). At the highest dose (100 nM),

[h] I thank Martin Flerin, DPhil candidate in the Prof. Fernanda Duarte group, for conducting the MDS.



B.

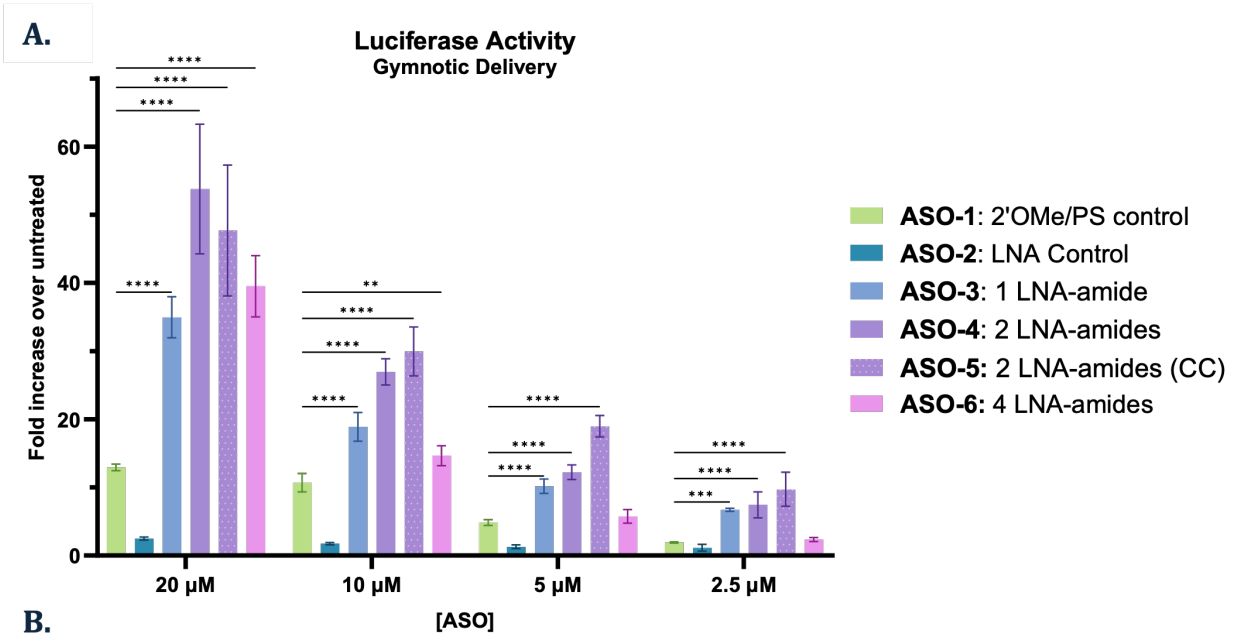
Fold increase over **ASO-1:** 2'OMe/PS control	**100 nM**	**50 nM**	**25 nM**	**12.5 nM**
ASO-2: LNA control	2.9 ±0.9	2.4 ±0.6	2.8 ±0.6	3.1 ±1.2
ASO-3: 1 LNA-amide	3.1 ±0.7	3.9 ±0.6	2.5 ±0.5	1.8 ±0.6
ASO-4: 2 LNA-amides	2.6 ±0.8	2.8 ±0.5	2.6 ±0.4	2.0 ±0.6
ASO-5: 2 LNA-amides (CC)	3.7 ±1.3	4.0 ±0.6	4.7 ±1.3	4.2 ±2.0
ASO-6: 4 LNA-amides	3.2 ±1.0	2.9 ±0.4	3.1 ±0.5	3.4 ±1.6

Figure 2.16: **A.** Dose response splice-switching activity of ASOs containing LNA-amide linkages using the luciferase reporter assay, delivered by transfection with Lipofectamine 3000. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using ASO-1 (2'OMe/PS control) as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing LNA and LNA-amide backbones.

the activity of the ASOs containing LNA-amides ranged from 2.6 to 3.7-fold higher than ASO-1 (Figure 2.16B). Comparison between the transfected activities of ASO-4 and ASO-5 revealed a slight positional effect – the activity of the ASO with incorporations at the TT positions (ASO-4) ranged from 2 to 2.8-fold higher than ASO-1 vs. 3.7 to 4.7-fold higher for the ASO with the incorporations at the CC positions (ASO-5). Interestingly, the activity of the ASO did not increase linearly with multiple incorporations of the LNA-amide. Both the fold-increase over ASO-1 for the single incorporation of the LNA-amide (ASO-3)

vs. four incorporations (ASO-6) were ∼3-fold at the highest dose. Finally, the LNA-control (ASO-2) (lacking any amide linkages) had comparable transfected activity to the LNA-amide containing ASOs (ASOs 3-6), ranging from 2.4 to 3.1-fold more active than ASO-1.

A second dose response assay was conducted using gymnotic delivery (Figure 2.17) (higher concentrations of ASO are required due to the limited cellular uptake of naked ASOs). By introducing the cell membrane as an obstacle to the ASO's splice-switching activity, it is pos-



B.

Fold increase over **ASO-1:** 2'OMe/PS control	**20 μM**	**10 μM**	**5 μM**	**2.5 μM**
ASO-2: LNA control	0.2 ±0.0	0.2 ±0.0	0.3 ±0.1	0.6 ±0.3
ASO-3: 1 LNA-amide	2.7 ±0.3	1.8 ±0.3	2.1 ±0.3	3.5 ±0.2
ASO-4: 2 LNA-amides	4.2 ±0.8	2.5 ±0.4	2.5 ±0.3	3.9 ±1.0
ASO-5: 2 LNA-amides (CC)	3.7 ±0.8	2.8 ±0.5	3.9 ±0.5	5.0 ±1.3
ASO-6: 4 LNA-amides	3.1 ±0.4	1.4 ±0.2	1.2 ±0.2	1.2 ±0.2

Figure 2.17: **A.** Dose response splice-switching activity of ASOs containing LNA-amide linkages using the luciferase reporter assay, delivered by gymnosis. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using ASO-1 (2'OMe/PS control) as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$. ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing LNA and LNA-amide backbones.

sible to better evaluate if an ASO modification can improve activity via a clinically-relevant route. It was promising to see that the ASOs containing LNA-amide linkages retained their significantly increased activity compared to the 2'OMe/PS control in line with previous reports (Figure 2.17A).[62] However, the LNA-control (ASO-2) was surprisingly inactive when treated gymnotically − at all concentrations, it was significantly less active than ASO-1 with fold-change in activity below 1. It is hypothesised that this is due to either cytotoxicity[i] when delivered via gymnosis or reduced cell permeability, or both. A few trends remained consistent between delivery methods − the ASOs containing two incorporations of the LNA-amide (ASO-4 and ASO-5) were again observed to be the most active, producing a 4.2 and 3.7-fold increase over ASO-1 at the highest dose, respectively. The ASO containing four incorporations of the LNA-amide (ASO-6), while still significantly more active at the highest doses (3.1-fold increase at 20 µM) was less potent when delivered via gymnosis, as the activity was reduced to only ∼1.3-fold increase over ASO-1 at the lower doses (2.5 to 10 µM). However, it was promising to observe that while the LNA-control lost all of its gymnotic activity, this LNA-amide ASO had equal LNA content and still observed equal, if not slightly greater, splice-switching activity than the 2'OMe/PS control (ASO-1). Potentially, the amide linkage may be altering the protein-binding profile of the LNA-containing ASO to encourage productive uptake or reduce cytotoxicity.[80] For therapeutic applications, it was promising to see that potentially only 1 or 2 incorporations of an LNA-amide backbone would be required to observe beneficial activity in a splice-switching ASO − in future work, it would be important to "walk" the modification along the sequence to optimise the ideal position within a pre-clinical candidate sequence.

An increase observed from inherent transfected activity to gymnotic activity (which requires the passage through the cell wall) may signal an ASO has increased cellular uptake. However, ASOs containing one or two incorporations of the LNA-amide (ASOs 3-5) have similar improved activity when transfected or treated via gymnosis (the differences between fold-increase over ASO-1 was often within error). ASO-6, containing four LNA-amide incor-

[i] A preliminary cytotoxicity assay for the 2'OMe/PS control and LNA control ASOs is given in Figure B.5, but more biological replicates are needed.

porations, observed a decrease when delivered via gymnosis, (similar to the LNA control), potentially due to slight cytotoxicity or the altered protein-binding profiles related to the higher LNA-content. Therefore, whether the LNA-amide linkage unilaterally increases cell uptake (regardless of position and number) is not as straightforward as once imagined. However, it was encouraging to report that inherently active ASOs often have similar naked activity, especially at higher doses. It is proposed that transfected assays continue to be used as a first indication of a modification's potential success, as these require lower ASO quantities and results could inform a larger scale synthesis for gymnotic testing. Across both delivery methods, the inclusion of an LNA-amide linkage significantly improved an ASO's splice-switching activity compared to the 2'OMe/PS control, which merits further investigation.

2.9 Conclusions

In conclusion, the therapeutic potential of ASOs containing LNA-amide backbones were explored in this chapter. A monomer approach to installing this linkage was used − locked 3'-carboxylic acid nucleosides and locked 5'-amino phosphoramidites were synthesised and the on-resin amide coupling between the pyrimidine thymine and ^{Me}cytosine monomers was optimised. However, amide coupling yields remained below that of automated solid-phase oligonucleotide synthesisers and so, LNA-amide dimer phosphoramidites were synthesised to use a dimer approach. LNA-amide dinucleotide phosphoramidites TT and ^{Me}C(Bz)-^{Me}C(Bz) were synthesised and coupled on the oligonucleotide synthesiser. The incorporation of dimers rather than manual on-resin coupling of monomers required less time-consuming solid-phase synthesis and the crude conversion to desired final oligonucleotide product (by HPLC integration) was higher than in the monomer approach.

Various ASOs containing the LNA-amide modification were synthesised, ranging in incorporation number from one to four. The biophysical properties of these ASOs were studied and are consistent with what is reported in the literature; the LNA-amide backbone increases the thermal duplex stability of both ASO:DNA and ASO:RNA duplexes and does not in-

duce large global conformational changes in a duplex compared to a native phosphodiester (by crystal structure and MD simulations) or compared to a 2'OMe/PS-modified ASO (by CD).[117] Next, the ASOs were tested in the luciferase reporter assay to further explore the LNA-amide's promising biological activity demonstrated by our group previously.[62] All ASOs containing LNA-amide linkages were significantly more active than the 2'OMe/PS control at almost all transfected doses, ranging from 2.6 to 3.7-fold increase at the highest dose. This promising improvement was mirrored in the gymnotic assay, where the ASOs containing LNA-amide linkages produced a fold-increase over ASO-1 ranging from 2.7 to 4.2-fold at the highest dose (20 µM). A positional effect was observed for the ASOs containing two incorporations of the LNA-amide; ASO-5 was more potent than ASO-4 when delivered via transfection. Finally, the incorporation number additionally has an effect, as both ASOs with two incorporations (ASO-4 and ASO-5) were the most active after gymnotic treatment (producing 3.7 to 4.2-fold increase at 20 µM). The LNA-control (ASO-2) and ASO containing four LNA-amide backbones (ASO-6) contain identical LNA quantities but only the ASO containing amide backbones was able to retain gymnotic activity, while ASO-2 was inactive. Overall, the LNA-amide linkage can easily be installed into therapeutic chimeric ASOs via eight monomers, and its biological activity holds great therapetic promise.

3

The synthesis and study of ASOs containing LNA-amide-inspired linkages

Contents

3.1 Introduction

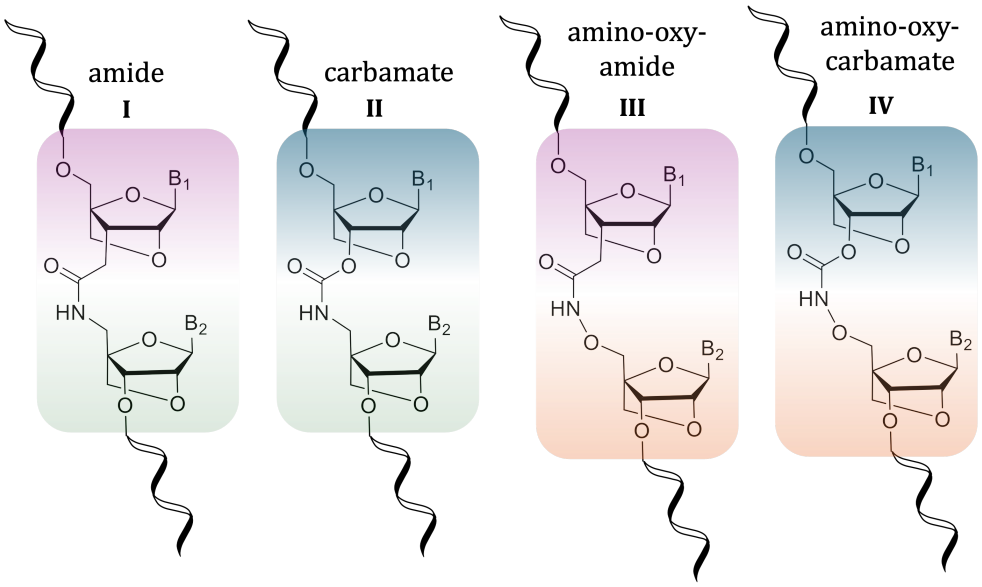


Figure 3.1: Summary of LNA-amide-inspired linkages discussed in this chapter: Linkage I (amide) was discussed in Chapter 2 and the synthesis of amide-inspired backbones II (carbamate), III (amino-oxy-amide) and IV (amino-oxy-carbamate) are discussed here in Chapter 3.

The LNA-amide linkage is an excellent mimic of the phosphate diester with promising biological activity, and so, this chapter will investigate other amide analogues with varied structures in an effort to identify additional mimics that may be more synthetically accessible. This chapter will investigate "amide-inspired" backbones which vary slightly from the amide in hetero-atom quantity, composition, and linker length. The LNA moieties will be flanking the linkage to enable comparison to the LNA-amide and increase the thermal duplex stability of potentially destabilising linkages. The three backbones investigated in this chapter are a

carbamate linkage (5'-$OCONHCH_2$-3'), an amino-oxy-amide linkage (5'-$CH_2CON(H)OCH_2$-3'), and an amino-oxy-carbamate linkage (5'-$OCON(H)OCH_2$-3') (Structures II, III, and IV in Figure 3.1).

3.1.1 Introduction to the carbamate linkage

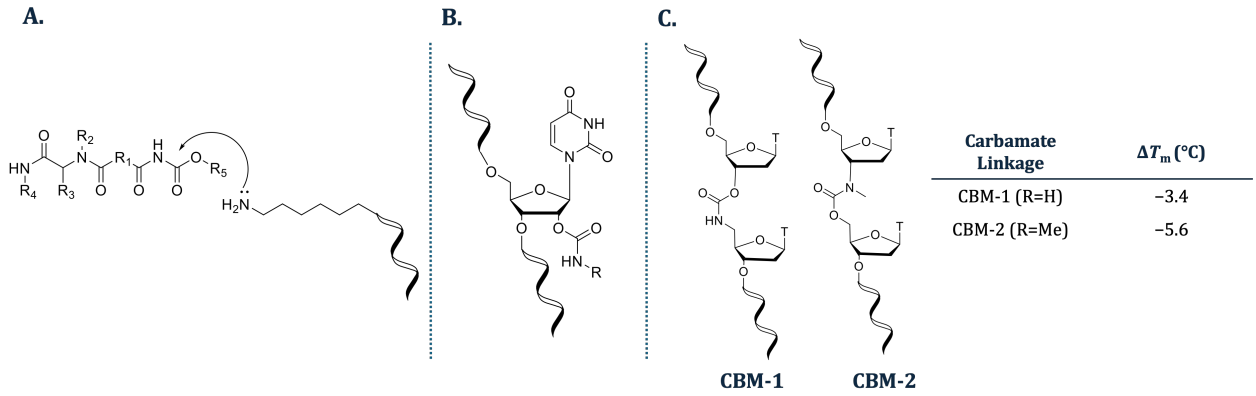
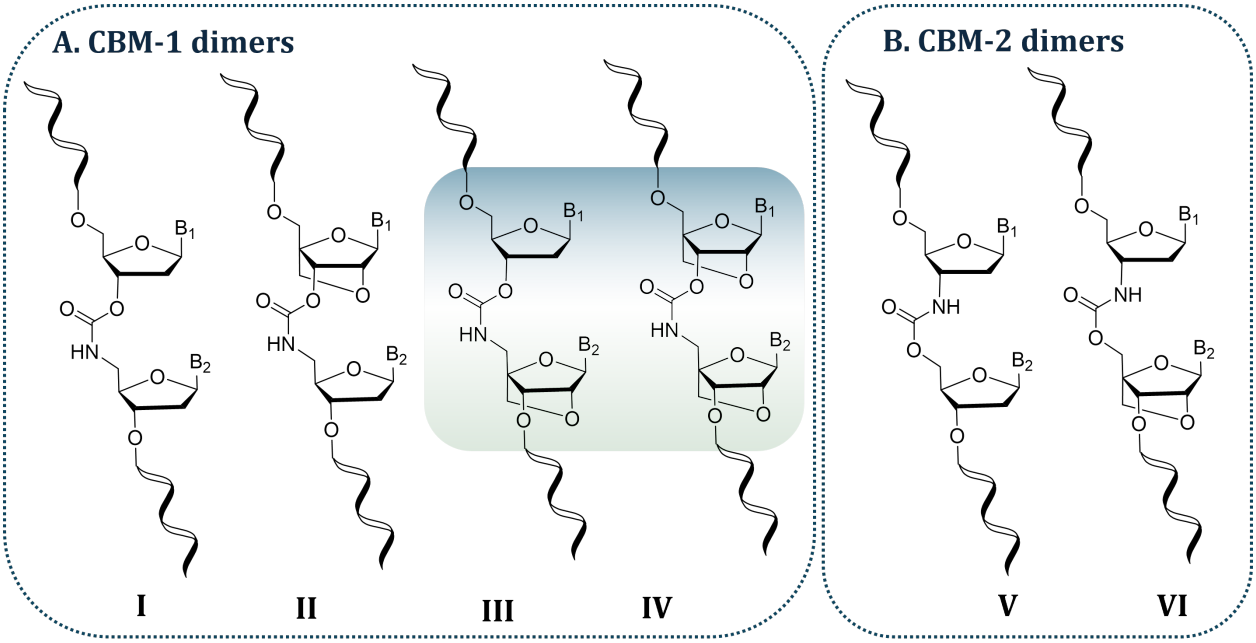


Figure 3.2: Examples of the carbamate linkage as a functional group in ASOs, e.g. bioconjugation strategy (**A.**),[141] as a linkage in a 2'-modification (**B.**),[142] or the first incorporation as a backbone by Waldner et al. (**C.**),[143] featuring both constitutional isomers CBM-1 and CBM-2.

The carbamate linkage is a common functional group in medicinal chemistry for small molecule drug discovery;[144] it has also been used in ASOs in various applications. For example, Kita et al. conjugate oligonucleotides to peptoids bearing various R groups by reaction of a 5'-NH_2 group with an activated carbamate-functionalised tag (Figure 3.2A).[141] The carbamate group has also been used as a linkage to modify the 2'-position in an oligonucleotide − however, this was unfortunately destabilising according to thermal melting studies (ΔT_m = −3 to −5 °C)(Figure 3.2B).[142] It was first investigated for incorporation as a backbone into oligonucleotides by Waldner et al. in 1994 (Figure 3.2C). Compared to the DNA phosphodiester control, the CBM-1 linkage was destabilising in a DNA:RNA heteroduplex (ΔT_m = −3.4 °C), but it displayed good enzymatic resistance.[143] The constitutional isomer CBM-2 (additionally containing a methylated amino-moiety) was more destabilising in a DNA-RNA heteroduplex (ΔT_m = −5.6 °C)

The Tom Brown group has investigated the carbamate backbone further by "rescuing" the

destabilisation effect through combination with LNA, in a similar strategy to the amide discussed in Chapter 2. The backbones synthesised previously include the 5'-OCON(H)CH_2-3' linkage (CBM-1) in combination with LNA flanking the 5', 3', or both sides of the dinucleotide; the constitutional isomer 5'-$NH_2C(O)OCH_2$-3' (CBM-2) was also synthesised in combination with LNA on the 3'-end (Figure 3.3A).[145] Stabilisation by the LNA was found



C.

Carbamate Linkage	DNA Complement		RNA Complement	
	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)
unmodified control	60.2	–	63.6	–
I	57.5	−2.7	57.2	−6.4
II	50.5	−9.7	56.1	−7.5
III	60.3	+0.1	61.3	−2.3
IV	53.4	−6.8	62.1	−1.5
V	53.6	−6.6	55.0	−8.6
VI	51.7	−8.5	57.1	−6.5

Figure 3.3: Summary of carbamate linkages in combination with LNA explored previously by the Tom Brown group. **A.** Oligonucleotides previously synthesised containing a 5'-OCON(H)CH_2-3' linkage (CBM-1). **B.** Oligonucleotides previously synthesised containing a 5'-NHC(O)OCH_2-3' linkage (CBM-2). **C.** Duplex stability of oligonucleotides containing one of these 6 linkages against their DNA and RNA reverse complements. The remaining chemistry of the oligonucleotide was unmodified DNA/PO. Linkages III and IV are highlighted (blue) as they were the least destabilising relative to the control oligonucleotide and are further investigated in therapeutic ASOs in this chapter. Data reproduced with permission from Thorpe et al.[145]

to be best when flanking both sides of the CBM-1 carbamate linkage (ΔT_m of the DNA-carbamate-DNA linkage (linkage I, Figure 3.3A) = −6.4 °C vs. ΔT_m = −1.5 °C for the LNA-carbamate-LNA linkage (linkage IV, Figure 3.3A). The LNA was not tolerated well at the 5'-end on the "inside" of the linkage − the ΔT_m of the LNA-carbamate-DNA (linkage II, Figure 3.3A) = −7.5 °C, potentially due to the clash between two structurally rigid groups (Figure 3.3C). As is a similar case to the destabilising five-atom amide linkage discussed in Figure 2.3 (Chapter 2) and reported by Pallan et al.,[111] placing the LNA solely on the 3'-side of the linkage can also mitigate its destabilisation (ΔT_m of DNA-carbamate-LNA (linkage III, Figure 3.3A) = −2.3 °C).[111] A constitutional isomer of the CBM-1 linkage, the CBM-2 linkage (5'-NHC(O)OCH_2-3') was also investigated (Figure 3.3B), but it was very destabilising. The addition of one LNA moiety flanking its 3'-side did increase the thermal stability (ΔT_m = +2.1 °C from DNA-CBM-2-DNA (linkage V, Figure 3.3B) to DNA-CBM-2-LNA (linkage VI, Figure 3.3B), but remained overall destabilising (ΔT_m = −6.5 °C compared to the native phosphate diester) and so, it was not further investigated here.

It was interesting that the CBM-1 linkage has very different biophysical properties from the amide, despite being only a one-atom change. Thorpe et al. hypothesised that the carbamate would have very similar atom positioning as the amide (with the carbonyl in the same position) − it might be expected that the carbamate should continue to be a good bioisostere of the natural phosphate linkage, yet it is much less stable than the amide. The rotational barrier of the carbamate C-N bond is 3-4 kcal mol^{-1} (15-20%) lower than in the C-N bond in structurally analogous amides due to the additional sterics and electronic effects of the oxygen.[67] It is hypothesised that the carbamate's electronic perturbation of the oxygen (compared to the carbon) make it less stable than the amide, but the LNA-carbamate (structure IV, Figure 3.3) is not so destabilising (ΔT_m = −1.5 °C against the RNA complement) that it does not merit further investigation. It has not yet been investigated in a biological system, and thus, linkages III and IV (DNA-carbamate-LNA and LNA-carbamate-LNA) were synthesised in this chapter for inclusion into 2'OMe/PS therapeutic ASOs for the first time.

3.1.2 Introduction to the amino-oxy linkage

The two additional LNA-amide-inspired linkages included amino-oxy-moieties − these backbones were inspired by previous works which use the amino-oxy functional group as a ligation method − Datta et al. report amino-oxy click chemistry to functionalise an oligonucleotide with various bis-homo and bis-hetero ligands, for applications such as delivery conjugates and fluorescent labelling (Figure 3.4A).[146] In fact, the majority of the literature reports the synthesis of the extremely reactive amino-oxy moiety (NH_2-O-R) in order to conduct oxime ligation with aldehydes, which is established as a click-like bio-orthogonal ligation method.[147] For example, the Bertozzi research group discovered a novel epimerase inhibitor by constructing a 1338-member library through the use of a 5'-amino-oxy uridine monomer (Figure 3.4B).[148]

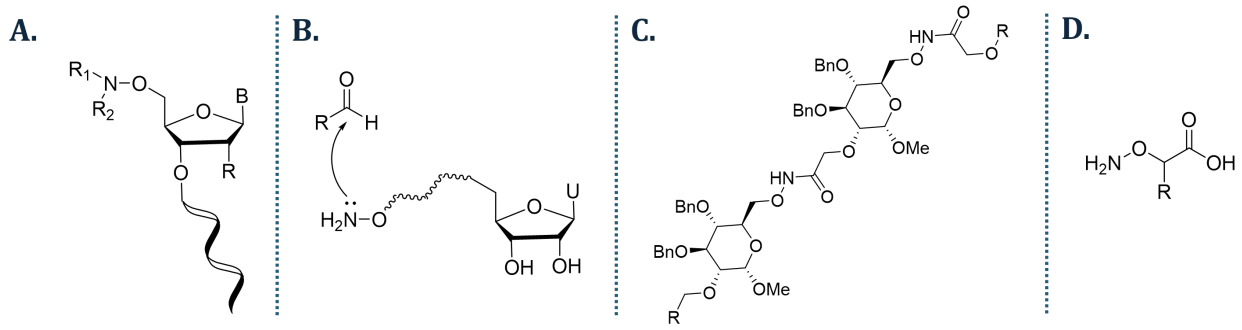


Figure 3.4: Examples of amino-oxy linkages used outside of dinucleotide backbones: **A.** Datta et al. functionalise an oligonucleotide with bis-homo or bis-hetero ligands.[146] **B.** Winans et al. construct a library of small molecules through click-like ligation of various 5'-amino-oxy uridine monomers with various aldehydes.[148] **C.** Amino-oxy-amide linkage between glycolipids for the construction of oligosaccharide mimetics.[149] **D.** α-amino-oxy acid for the construction of novel peptides.[150]

The inclusion of the amino-oxy moiety as a linkage rather than a conjugation strategy has limited but attested exploration in various biopolymers. For example, the linkage (5'-$OCH_2CONOCH_2$-3') has been reported between 6-membered glycolipids in order to construct novel oligosaccharide mimetics (Figure 3.4C).[149] The N-O bond has been used to alter the properties of amino acids in the synthesis of novel peptides, e.g. the *N*-oxyamide (NH_2-O-CH_2(R)-COOH) amino acid building block has been used in the construction of foldamers, anion receptors, and ion channels; using α-aminooxy acids rather than amino acids increases

the peptide's ability to form intramolecular hydrogen bonds and thus, interesting secondary structures (Figure 3.4D).[150]

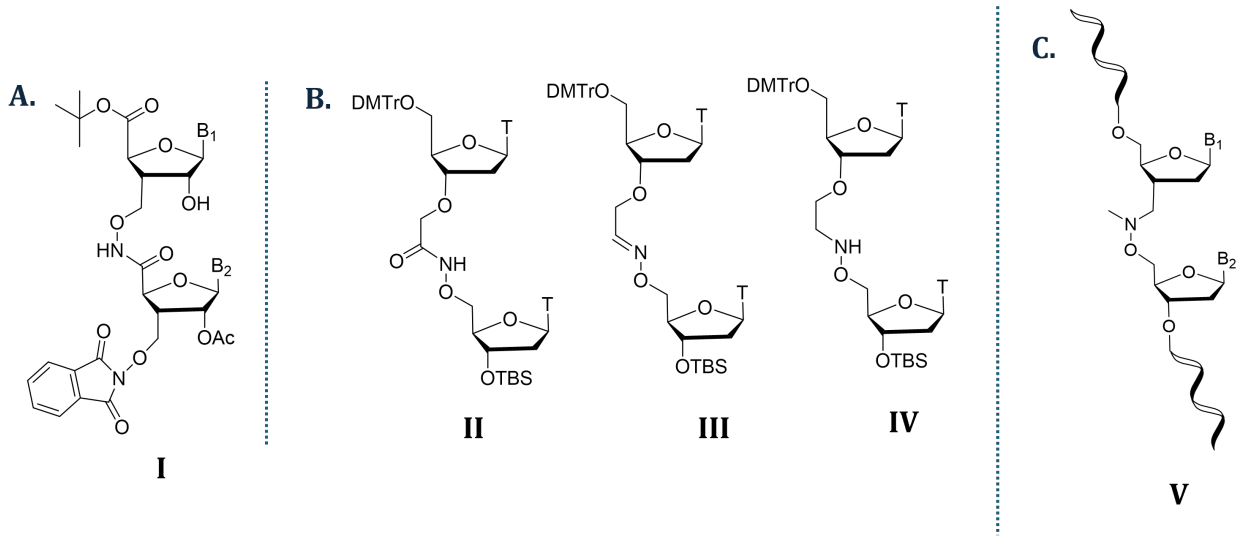


Figure 3.5: A. *N*-oxyamide dinucleotide synthesised previously in the literature.[151] **B.** 6-atom linkages containing an amino-oxy moiety synthesised previously in the literature.[152] **C.** The methylene methylimino (MMI) linkage containing an amino-oxy moiety. This linkage has been incorporated successfully into oligonucleotides via both monomer and dimer approaches.

Within the oligonucleotide field specifically, the use of the amino-oxy functional group as a backbone is less prevalent; the ones found in the literature are described in Figure 3.5. The synthesis of the *N*-oxyamide dinucleotide I (Figure 3.5A) was described by Gong et al. in 2011, but this was not incorporated into oligonucleotides.[151] It was hypothesised that linkage I would be quite destabilising, as the carbonyl is placed directly adjacent to the 3'-monomer; amide linkages with this orientation were poorly tolerated in the literature.[52] Longer, 6-atom linkages derived from a shared 5'-NH_2-O-thymidine monomer were described by Peyrat et al (linkages II-IV, Figure 3.5).[152] However, these dinucleotides were never incorporated into oligonucleotides, and their synthesis was also very lengthy (>10 steps per dinucleotide). Furthermore, these 6-atom linkages are likely to be poorly tolerated within oligonucleotides due to the extended internucleotide distance. Therefore, the investigation of linkages I-IV was not pursued.

The most promising amino-oxy-containing linkage that is well-explored is the methylene methylimino (MMI) linkage (5'-$CH_2N(Me)OCH_2$-3') (linkage V, Figure 3.5C). This is also

a 4-atom, 5-bond neutral achiral linkage (as the carbamates and amides) and was first reported as a dinucleotide and incorporated into oligonucleotides in 1992 by Vasseur et al.[153] A ΔT_m of -0.3 °C per modification (against the RNA complement) was reported, as well as good enzymatic resistance and base mismatch discrimination. The promising biophysical properties led others to explore this linkage; it has been installed via a dimer approach[90] or via solid-phase synthesis (hydrazinolysis, coupling, reduction, and methylation occurring on-resin),[154] or included as a linker in 2'-modifications.[155] Despite the initial promising biophysical properties, when a single incorporation of this linkage was incorporated in a gapmer, it had the highest ED_{50} value in vivo compared to other neutral achiral linkages (3.7 mg/kg vs 0.5 mg/kg achieved with an amide in the same position).[90]

The reactivity of the amino-oxy moiety, its widespread usage as a conjugation ligand, and the promising biophysical properties of the MMI linkage inspired the amino-oxy-amide and amino-oxy-carbamate linkages (see Figure 3.1). These combine the amino-oxy moiety with the carbonyl moiety of the amide and carbamates, which has been shown to be a good phosphate bioisostere. In order to facilitate good comparison to the other hetero-atom-containing linkages explored in this thesis, they were also combined with two flanking LNA sugars. The expansion of the chemical space available to oligonucleotide chemists is increasingly important for the rational design of ASOs.

3.2 Aims of the Chapter

The promising properties of the LNA-amide warranted further investigation of the structure-activity relationship (SAR) of neutral, achiral, amide-like backbones. This chapter aims to synthesise splice-switching therapeutic ASOs containing three different amide-inspired linkages. The chapter aims to:

1. Synthesise DNA-LNA and LNA-LNA carbamate dimer phosphoramidites and incorporate these into 2'OMe/PS chimeric ASOs for the first time.

2. Synthesise novel locked amino-oxy-amide and amino-oxy-carbamate dimer phospho-

ramidites, beginning with a route to synthesise a novel locked 5'-amino-oxy monomer.

3. Evaluate the biophysical properties of ASOs containing these amide-inspired linkages.

4. Evaluate the biological activity of ASOs containing amide-inspired linkages.

The chapter also aims to discuss two additional LNA-amide-inspired linkages, the LNA-alkoxyamide and LNA-sulfamate, which were added to the investigation of LNA-amide-inspired linkages through collaboration.[a] The chapter aims to further understand the SAR behind LNA-neutral linkages discussed thus far through MD simulations, also conducted through collaboration.[b] Finally, the chapter aims to discuss the biological activity of all ASOs containing LNA-amide-inspired linkages in the luciferase splice-switching reporter assay.

3.3 Synthesis of LNA-amide-inspired dinucleotides

3.3.1 Synthesis of carbamate dinucleotide phosphoramidites

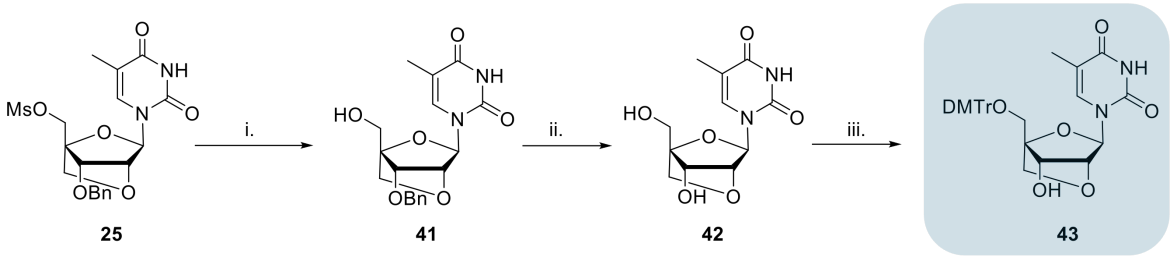
Synthesis of the envisioned carbamate and amino-oxy-carbamate dimers require a 5'-DMTr-protected 3'-hydroxyl locked thymine monomer **43**[c] (Scheme 3.1) for the 5'-end of the linkages. This was synthesised following a protocol by Koshkin et al.[156] The 5'-mesylate **25**, synthesised in Chapter 2, was reacted with sodium benzoate to give an intermediate 5'-benzoate ester and base-mediated hydrolysis gave the 5'-hydroxyl monomer **41** in 87% yield over two steps. The 3'-benzyl group was deprotected using Pd-catalysed hydrogenation to give the diol **42** in near-quantitative yield. Finally, protection of the 5'-hydroxyl group with 4-dimethoxytrityl chloride gave the 5'-DMTr-LNA-T-3'-OH monomer **43** in 73% yield.

The synthesis of the DNA-carbamate-LNA dimer **50** was achieved following a protocol inspired by Thorpe et al.[145] The route began with the deprotonation of commercially available 5'-DMTr-protected thymidine **44** by sodium hydride, followed by reaction with carbonyl di-

[a] I thank Dr. Alice Kennett and Dr. Belma Zengin-Kurt who were involved in the collaboration on LNA-neutral linkages.

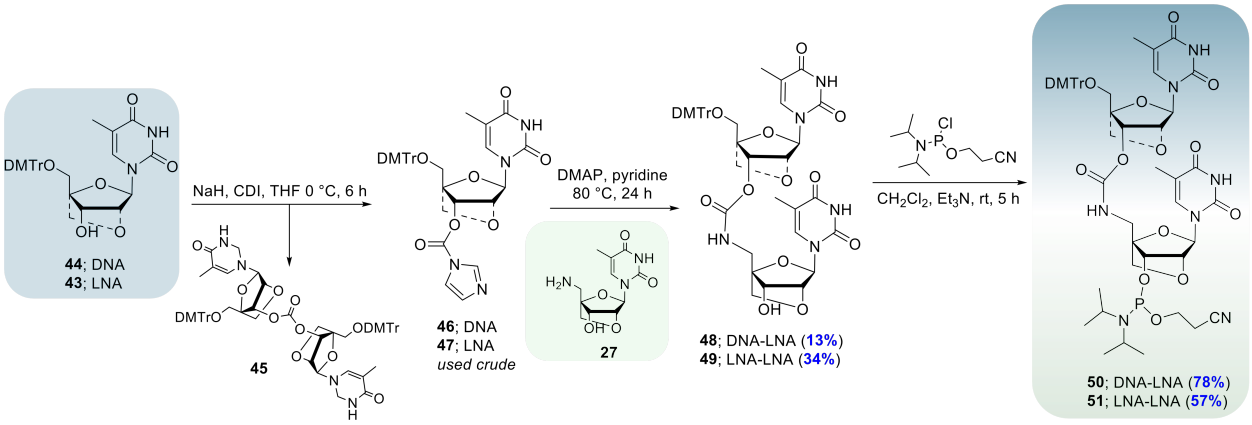
[b] I thank Martin Flerin, DPhil candidate supervised by Prof. Fernanda Duarte, for collaborating on this project. He conducted the MDS of these backbones.

[c] It is commercially available upon request, but was also synthesised on a large scale in seven steps beginning from the commercially-available bis-acetyl monomer **8**.



Scheme 3.1: Synthesis of 5'-DMTr-LNA-3'-OH monomer **43**. i. NaOBz, DMF, 100 °C, 24 h, then 2 M NaOH, THF:H_2O (1:1, v:v), rt, 16 h, 87% over two steps; ii. 20% $Pd(OH)_2/C$, H_2, MeOH, rt, 24 h, 96%; iii. DMTrCl, pyridine, rt, 24 h, 73%.

imidazole (CDI) to give the activated imidazole carbamate **46** which was used crude following aqueous work-up. The activated electrophile was reacted with the locked 5'-NH_2-monomer **27** (synthesised in Chapter 2) to give a carbamate-linked DNA-LNA dinucleotide alcohol **48** in 13% yield.[d] This was converted to the phosphoramidite **50** by standard phosphitylation conditions in 78% yield.



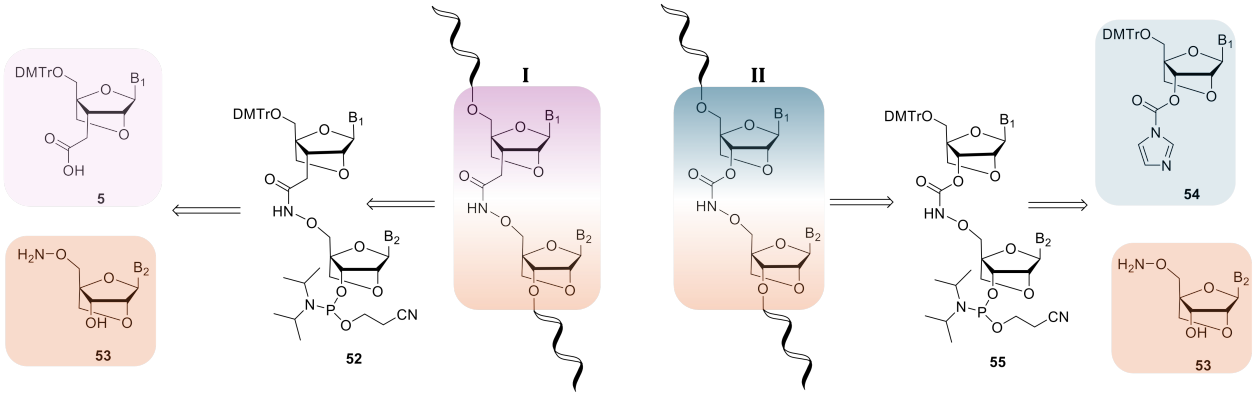
Scheme 3.2: Synthesis of DNA-carbamate-LNA dimer phosphoramidite **50** and LNA-carbamate-LNA dimer phosphoramidite **51.**

The LNA-carbamate-LNA dimer **51** was synthesised by an identical route, beginning with the deprotonation of the locked alcohol **43** (synthesised in Scheme 3.1) with sodium hydride. Reaction of the alkoxide with CDI produced the activated imidazole carbamate **47** which was reacted in a DMAP-catalysed electrophilic substitution reaction with the same locked 5'-NH_2 monomer **27** to give the carbamate-linked dimer **49**, isolated in a yield of 34%. The carbamate coupling proceeded in low yield (34%), due to the formation of a carbonate dimer byproduct **45** upon deprotonation of the alcohol. The carbamate-linked LNA-LNA

[d] I thank Autumn Usher, Part II, for the synthesis of this compound.

dinucleotide alcohol **49** was similarly converted to the phosphoramidite **51** by standard phosphitylation conditions in 57% yield, a lower crude conversion (monitored by TLC) due to the reduced reactivity of compounds with increased number of LNA moieties.

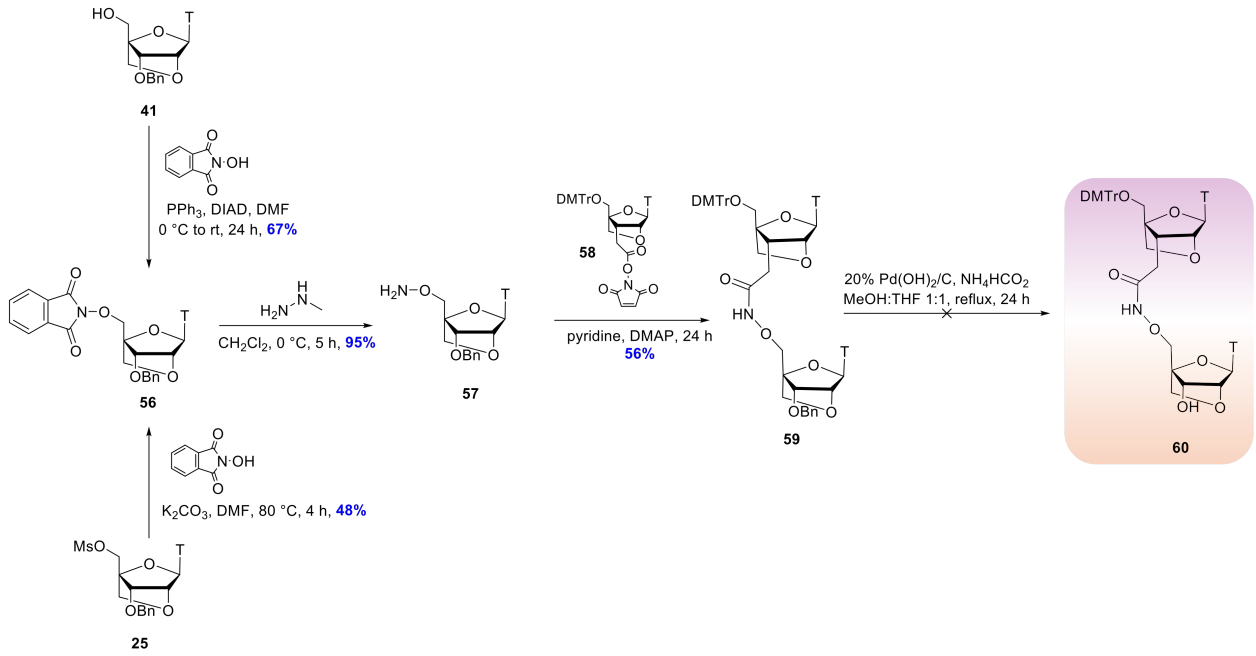
3.3.2 Synthesis of amino-oxy dinucleotide phosphoramidites



Scheme 3.3: Retrosynthetic route to achieve the LNA-amino-oxy-amide linkage (**I**) and LNA-amino-oxy-carbamate linkage (**II**). Both linkages share the common 5'-amino-oxy monomer **53**, while linkage **I** requires the locked 3'-COOH monomer **5** and linkage **II** requires the locked 3'-activated electrophile **54**, accessible via the locked 3'-OH monomer **43**.

In order to install the amino-oxy linkages, two phosphoramidite dimers were envisioned whose retrosyntheses share the same monomer at the 3'-end of the linkage, a locked 5'-NH_2-O-R monomer building block **53** (Scheme 3.3). To synthesise the amino-oxy-amide linkage (I), the 5'-monomer is comprised of the locked 3'-COOH monomer **5**, synthesised in Chapter 2, Scheme 2.1. To synthesise the amino-oxy-carbamate linkage (II), the 5'-monomer **54** is comprised of the activated imidazole carbamate achieved by activation of the locked 3'-OH monomer **43** with CDI (synthesised in Scheme 3.2).

In order to achieve the 5'-amino-oxy monomer **53**, a Gabriel-like synthesis was envisioned in which *N*-hydroxyphthalimide is used to install the N-O bond, as reported in the literature (Scheme 3.4).[146] Both an S_N2 reaction from the 5'-mesylate **25** and a Mitsonobu reaction from the 5'-alcohol **41** using *N*-hydroxyphthalimide as the nucleophile were attempted. The S_N2 to produce the 5'-oxyphthalimide monomer **56** proceeded in lower yields (48%) than the Mitsonobu reaction (67%) (Scheme 3.4). Despite proceeding in higher yields, the Mitsonobu reaction suffered from poor separation between the 5'-oxyphthalimide product **56** and the



Scheme 3.4: Attempted synthesis of the amino-oxy-amide dimer **60** via a 3'-OBn protected dimer **59**.

triphenylphosphine oxide byproduct; multiple purifications by column chromatography were required, even when the triphenylphosphine source was resin-bound. Due to the relative ease of the purification of the S_N2 reaction compared to the Mitsonobu, it was the preferred method for scaling up the reaction. Next, nucleophilic acyl substitution using methylhydrazine liberated the 5'-amino-oxy compound **57** in 95% yield, producing phthalhydrazide as a precipitate which was removed by filtration.

In order to form the amino-oxy-amide dimer **59**, the locked thymine 3'-COOH monomer **20** was first activated by reaction with *N*-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), and after an aqueous work-up, this activated ester **58** was reacted with the 5'-amino-oxy monomer **57** in a DMAP-catalysed substitution reaction to produce the LNA-LNA amino-oxy-amide dimer **59** in 64% yield. Unfortunately, the deprotection of the 3'-benzyl protecting group via hydrogenation conditions using palladium hydroxide on carbon was unsuccessful.

There are several reports of subjecting an amino-oxy-functionalised compound to palladium-catalysed hydrogenation conditions; however, these reports have mixed success. Benzyl de-

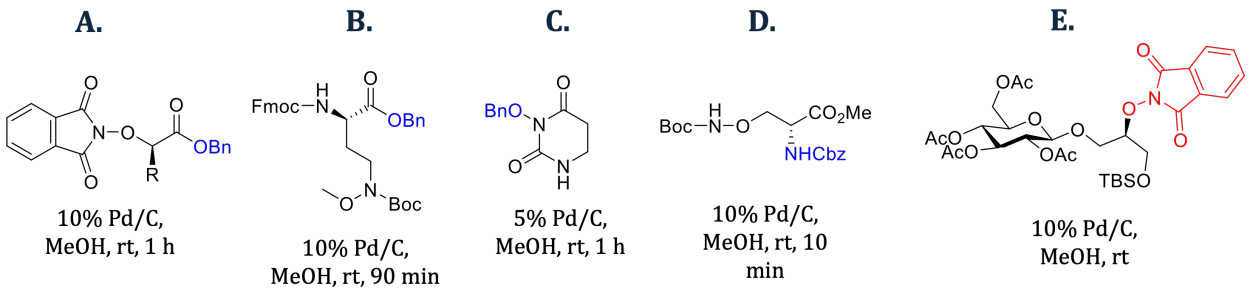
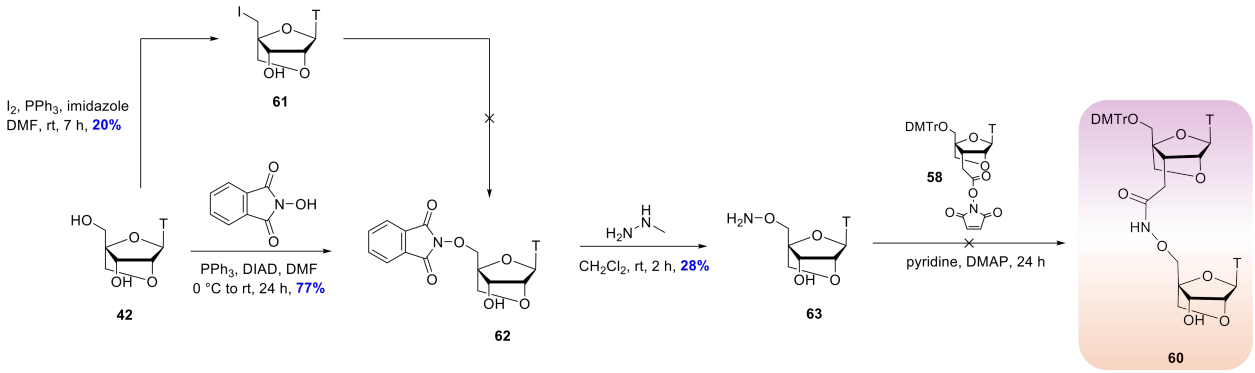


Figure 3.6: Compounds containing N-O bonds subjected to palladium-catalysed hydrogenation conditions in the literature.[157–161] Compounds A-D were successfully deprotected (blue protecting groups), while the N-O bond in compound E was entirely cleaved (red).

protection using hydrogenation conditions was successful in reports of a hydroxyphthalimide compound[157] (compound A, Figure 3.6), a methoxyamine-Boc-protected amino acid[158] (compound B, Figure 3.6) and benzyloxyhydrouracil (compound C, Figure 3.6).[159] Therefore, it was hypothesised that perhaps the amino-oxy-amide functional group in dimer **59** should be stable to these conditions. However, it is most likely that cleavage of the amino-oxy-amide backbone occurred; there are reports of this N-O bond cleavage in Pd/carbon conditions occurring, for example when Noel et al. attempt to remove benzyl chloroformate (CBz) protecting groups from Boc-protected-NH-O-amino-oxy acid (compound D, Figure 3.6) – they require a reduced reaction time of 10 min to mitigate yield loss.[161] Furthermore, Chen et al. use hydrogenolysis conditions to cleave an oxy-phthalimide group from a functionalised glycolipid (compound E, Figure 3.6) in order to liberate the hydroxyl group.[160]

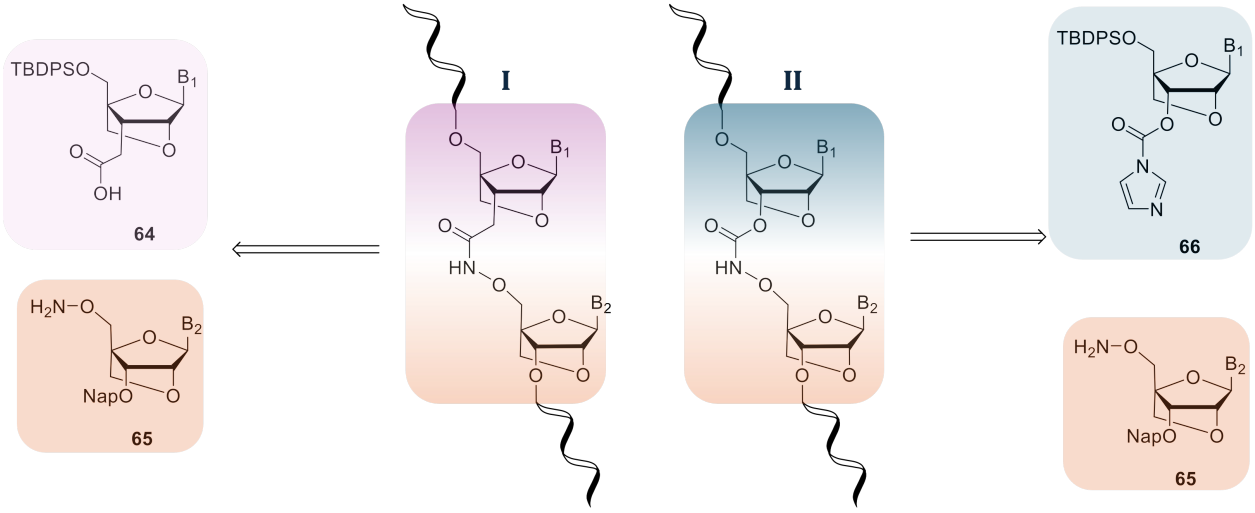


Scheme 3.5: Attempted synthesis of amino-oxy amide dimer via a 5'-amino-oxy-3'-hydroxyl monomer **63**.

A second strategy starting with the unprotected 3'-OH diol **42** was attempted to avoid the use of the 3'-OBn protecting group (Scheme 3.5). An iodide intermediate **61** was synthesised in the first instance, but this was isolated in low yield (20%) and the following S_N2 reaction with *N*-hydroxyphthalimide was unsuccessful. Pleasingly, a Mitsonobu reaction of the 5'-alcohol in **42** with *N*-hydroxyphthalimide did proceed in 77% yield. However, the nucleophilic acyl substitution using methylhydrazine to produce the 5'-amino-oxy monomer **63** was very low yielding (28%) due to the difficulty in handling the extremely polar compound. It was hypothesised that a lipophilic protecting group on the 3'-hydroxyl group would improve the purification of the amino-oxy monomer **63**. The coupling of this polar and reactive compound with the activated ester **58** was unsuccessful.

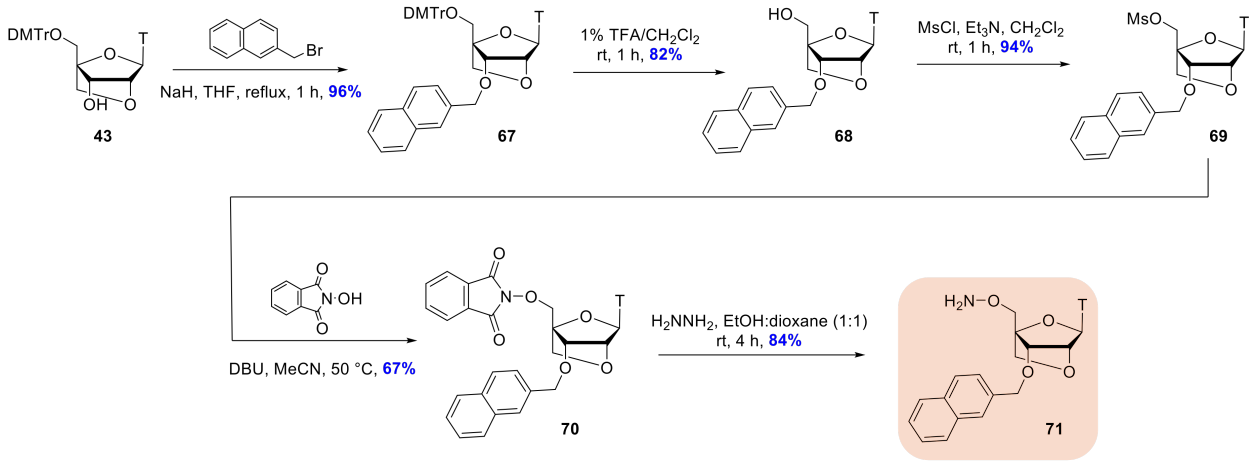
The synthesis of the amino-oxy-carbamate dimer was also attempted; a coupling between the same reactive 5'-amino-oxy monomer **63** and the activated 3'-imidazole carbamate electrophile **54** (activated in the same way as the carbamate synthesis in Scheme 3.2) was attempted (Scheme C.1). However, the same carbonate by-product dimer **45** was formed upon activation by CDI, and no amino-oxy-carbamate product was isolated. For both the amino-oxy-amide and amino-oxy-carbamate, coupling between the two locked monomers was much easier to monitor (by TLC) and isolate when the 5'-amino-oxy compound was 3'-protected with a lipophilic group; as the 3'-alcohol, compound **63** was too polar and difficult to handle.

An alternative protecting group to the benzyl group was investigated; the naphthalene (Nap) protecting group has precedence in 6-membered sugars.[162] In nucleotides, the naphthalene group is used less routinely, but has been used in the patent literature for the protection of the 3'-OH groups when installing the MMI linkage − therefore, it was hypothesised to be a viable route to achieving the similar amino-oxy linkage.[163] Therefore, the synthetic aims were revised to include a 5'-amino-oxy-3'-naphthalene-protected monomer **65** (Scheme 3.6). However, the 5'-DMTr protecting group used in the previous 5'-monomers was incompatible with naphthalene deprotection conditions which required 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ). Therefore, the synthetic aims were also revised to include the 5'-TBDPS-protected electrophiles **64** and **66** (Scheme 3.6).



Scheme 3.6: Revised retrosyntheses to achieve the amino-oxy-amide linkage (I) and amino-oxy-carbamate linkage (II). The protecting group strategy was changed to include the syntheses of 5'-TBDPS-protected electrophiles **64** and **66** and a shared 5'-amino-oxy-3'-naphthalene-protected nucleoside **65**.

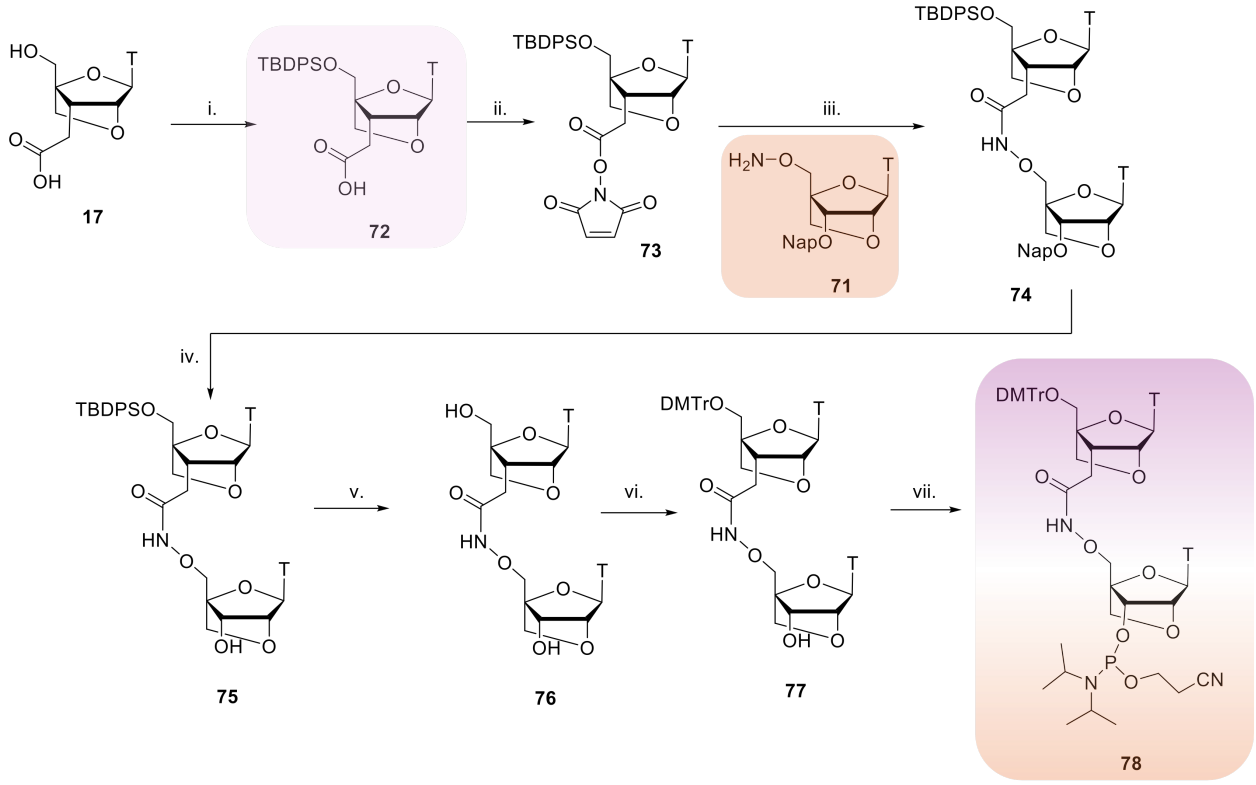
The synthetic route to achieve the 5'-amino-oxy-3'-naphthalene thymine monomer **71** began with the previously-synthesised 5'-DMTr-protected locked monomer **43** (Scheme 3.7). Alkylation of the 3'-alcohol with 2-bromonaphthalene gave the 3'-naphthalene protected monomer **67** in very good yield (96%). The trityl group was deprotected in acidic conditions to give the 5'-hydroxyl in 82% yield, which was subsequently protected by reaction with methanesulfonyl chloride to give the 5'-mesylate **69** in 94% yield. An S_N2 reaction with *N*-hydroxyphthalimide gave the 5'-oxyphthalimide monomer **70** in 67% yield. A Mitsunobu



Scheme 3.7: Synthetic route to achieve 5'-amino-oxy-3'-naphthalene locked thymidine monomer **71**. The route yields 42% over 5 steps.

reaction was attempted on the intermediate 5'-alcohol **68**, but similar problems in purification arose as in the case of the 3'-OBn protected monomer **56**; the retention factors of 3'-protected, 5'-oxyphthalimide monomers are too similar to the triphenylphosphine oxide by-product. The route via an intermediate mesylate was higher yielding despite being one step longer. Finally, hydrazine was used to liberate the 5'-amino-oxy compound **71** in 84% yield.

The synthesis of the amino-oxy-amide dimer using this novel locked 5'-NH_2-O-3'-ONap nucleoside **71** was attempted first (rather than the amino-oxy-carbamate dimer), as the activation of the 3'-carboxylic acid **72** would not produce the carbonate dimer by-product as it might with the carbamate. In order to enable later 3'-naphthalene deprotection in DDQ, the 5'-DMTr protecting group of the 3'-COOH monomer was replaced with the compatible bulky tert-butyldiphenylsilyl protecting group (TBDPS). The 5'-OH-T-3'-COOH monomer



Scheme 3.8: Synthetic route to achieve amino-oxy-amide dimer phosphoramidite **78**. i. TBDPSCl, imidazole, DMF, rt, 24 h, 46%; ii. NHS, EDC, CH_2Cl_2, rt, 4 h, used crude; iii. compound **71**, pyridine, DMAP, rt, 16 h, 89% over two steps; iv. DDQ, CH_2Cl_2, H_2O, rt, 24 h, 76%; v. TREAT-HF, Et_3N, THF, rt, 1 h, 88%; vi. DMTrCl, pyridine, rt, 24 h, 53%; 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, CH_2Cl_2, Et_3N, rt, 4 h, 22%.

17, synthesised in Chapter 2, was silylated to give 5'-TBDPS-3'-COOH monomer **72** in 43% yield. This was activated in a similar fashion as before, using NHS and EDC, and amide coupling with the amino-oxy compound **71** was much more successful than previous amino-oxy-coupling reactions, to give the amino-oxy-amide-linked dinucleotide **74** in 89% yield. The 3'-naphthalene group was first deprotected using DDQ (76% yield) and the TBDPS group was deprotected secondly using triethylamine trihydrofluoride (TREAT-HF) to give the diol dinucleotide **76** in 88% yield. The trityl protection of the 5'-hydroxyl group unfortunately only isolated dimer **77** in a yield of 53%, as the LNA-LNA diol was very difficult to handle; this is substantiated by literature which report difficulty in protecting the MMI-linked diol dinucleotide which is similarly polar.[163] Finally, the phosphitylaton of the 3'-hydroxyl group was attempted in standard phosphitylating conditions using 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite to give phosphoramidite **78**, but was unfortunately very low yielding (22%). Upon additional equivalents of the chlorophosphine, a second phosphoramidite species began to form, which was hypothesised to be a bis-phosphoramidite product, as the reactive nitrogen in the amino-oxy-amide linkage is also available to conduct nucleophilic attack of the electron-deficient phosphorus (see Scheme C.2).

Due to this low yield (22 mg), the phosphoramidite dimer **78** was unable to be successfully incorporated into an oligonucleotide.[e] The synthesis of the amino-oxy-carbamate dimer using the novel 3'-naphthalene monomer **71** was attempted as well, but the 5'-TBDPS-protected alcohol was not sufficiently activated to give sufficient amino-oxy-carbamate dimer quantities (see Scheme C.3). It is likely this dimer would have suffered from similar issues with bis-phosphitylation due to the reactive nitrogen. In future work, it is proposed to alkylate the nitrogen in the amino-oxy-amide linkage prior to coupling; the reactivity of the nitrogen in the amino-oxy bond could therefore be reduced and potentially make the compound easier to handle, as well as enable higher-yielding phosphitylations that do not suffer from by-product formation.

[e]The quantititive demands of the synthesiser are at least 500 µL of a 0.1 M solution.

3.4 Additional amide-inspired backbones

In a collaboration with other group members, additional amide-inspired backbones were investigated in order to attain a deeper understanding of neutral achiral hetero-atom-containing linkages and how they impact an ASO's stability, duplex conformation, and biological activity (Figure 3.7). The LNA-amide (linkage I) was used as a benchmark – the LNA-carbamate (linkage II) was included as a linkage which differs by one atom (Figure 3.7A). The amino-oxy linkages were envisioned as part of this investigation, but the difficult synthesis precluded them from incorporation into ASOs. Instead, phosphoramidite dimers with LNA-alkoxyamide (linkage III) and LNA-sulfamate (linkage IV) linkages were synthesised (Figure 3.7B).[f] The stability of the alkoxy-amide linkage (between DNA monomers) has been investigated in a DNA:RNA heteroduplex, and despite being a one-atom extension, the 6-bond linkage is only minimally destabilising ($\Delta T_m = -2$ °C).[111] Structurally, this linkage is a hybrid between the amide and the carbamate as it contains both the 3'-O and methylene moieties. Finally, the sulfamate linkage was chosen for investigation. The sulfamate was hypothesised to be a very good bioisostere of the natural phosphodiester, and the inclusion

[f]The LNA-alkoxyamide phosphoramidite dimer was synthesised by Dr. Belma Zengin-Kurt and the LNA-sulfamate phosphoramidite dimer was synthesised by Dr. Alice Kennett.

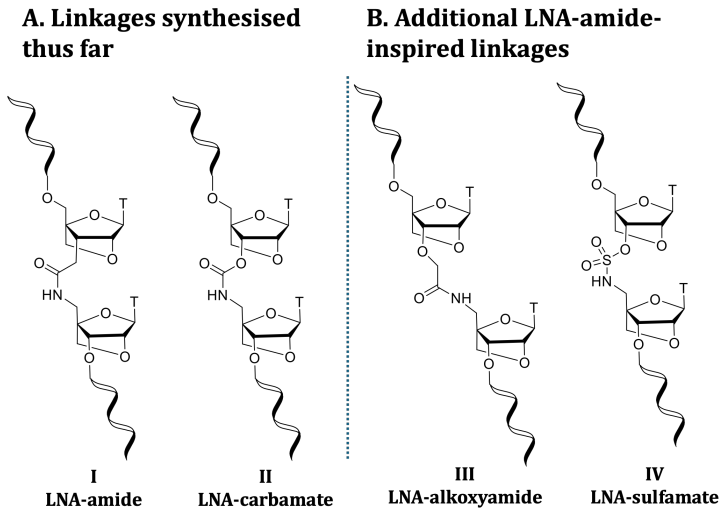


Figure 3.7: Summary of additional LNA-neutral linkages investigated: **A.** LNA-amide (I) and LNA-carbamate (II) linkages incorporated thus far in therapeutic ASOs. **B.** LNA-alkoyxamide (III) and LNA-sulfamate (IV) linkages were also installed via additional dimer phosphoramidites into therapeutic ASOs.

of the 3'-O and 5'-NH moieties enables comparison to amide-inspired backbones II and III (Figure 3.7).

3.5 Oligonucleotide synthesis

The carbamate-linked DNA-LNA dimer phosphoramidite **50** and the carbamate-linked LNA-LNA dimer phosphoramidite **51** were incorporated into chimeric 2'OMe/PS ASOs (Figure 3.8A, ASOs 8-11). This was the first time the carbamate linkage is included in therapeutic ASOs and ASOs 8-11 will serve to investigate the effect of a DNA vs. LNA sugar on the 5'-monomer of the carbamate linkage, as well as any differences between one or two incorporations. A schematic representation of the ASO modifications and their placement within the sequence is given in Figure 3.8B.

A.

ASO#	Sequence (5 → 3')	Expected Mass (Da)	Found Mass (Da)	Modifications x = linkage
ASO-1	CCUCUUACCUCAGUUACA	6096.3	6096.5	2'OMe/PS
ASO-8*	CCUCT**x**TACCUCAGUUACA	6042.4	6042.5	x = DNA-LNA carbamate
ASO-9*	CCUCT**x**TACCUCAGT**x**TACA	5988.5	5985.5	x = DNA-LNA carbamate
ASO-10	CCUC**TxT**ACCUCAGUUACA	6070.5	6071.5	x = LNA-LNA carbamate
ASO-11	CCUC**TxT**ACCUCAG**TxT**ACA	6044.5	6042.0	x = LNA-LNA carbamate
ASO-12**	CCUC**TxT**ACCUCAG**TxT**ACA	6070.1	6070.0	x = alkoxyamide
ASO-13	CCUC**TxT**ACCUCAG**TxT**ACA	6112.4	6112.5	x = sulfamate

B.

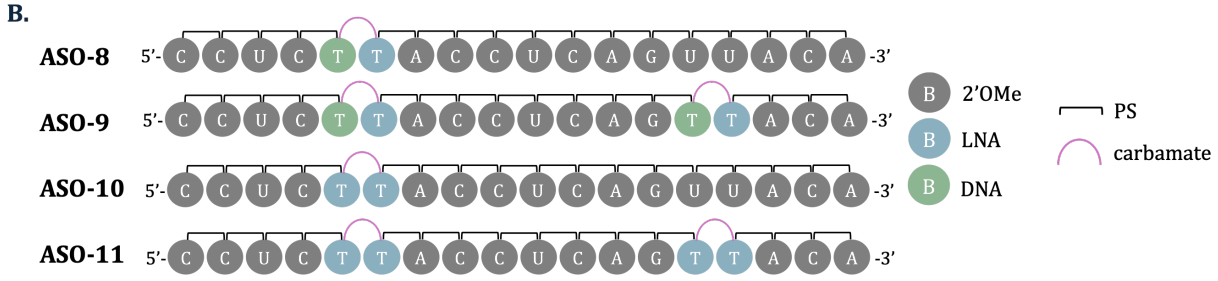


Figure 3.8: A. Table of ASOs synthesised containing carbamate linkages, LNA-sulfamate, and LNA-alkoxyamide linkages. **B** = LNA. *synthesised and purified by Autumn Usher. **synthesised and purified by Dr. Alice Kennett. **B.** Schematic representation of the ASOs synthesised containing carbamate linkages.

ASO-12 and ASO-13 were synthesised using LNA-alkoxyamide and LNA-sulfamate phosphoramidite dimers, synthesised by Dr. Alice Kennett and Dr. Belma Zengin-Kurt, respectively. Given the promising biological activity of the ASO containing two incorporations of the LNA-amide (ASO-4) seen in Chapter 2, the LNA-alkoxyamide and LNA-sulfamate linkages were incorporated in the same two positions. It was hypothesised that at least two incorporations would be sufficient to see noticeable differences among neutral linkages (as it was with the LNA-amide), while more than two incorporations would double the required phosphoramidites to include dimers beyond TT. Thus, for a preliminary study, two incorporations was used as a benchmark. ASO-12 and ASO-13 will be compared in a greater discussion to ASO-11 (2 LNA-carbamates) and ASO-4 (2 LNA-amides). All LNA-neutral linkages were compatible with standard ASO deprotection methods (heating at 55 °C in concentrated ammonium hydroxide for 5 h) and the ASOs were purified, de-tritylated, and the Et_3N^+ salt was exchanged for the Na^+, as described in Chapter 7.

3.6 Biophysical properties of ASOs containing amide-inspired linkages

3.6.1 Biophysical properties of ASOs containing carbamate linkages

In the literature, one incorporation of the DNA-LNA carbamate in an unmodified oligonucleotide was approximately equally stable to the DNA/PO control when hybridised to its DNA reverse complement ($\Delta T_m = +0.1$ °C); here, in a 2'OMe/PS chimera, the DNA-LNA carbamate linkage rendered the ASO:DNA heteroduplex slightly less stable ($\Delta T_m = -0.9$ °C) (Figure 3.9A). However, the linkage is much less tolerated by an ASO:RNA heteroduplex; when hybridised to the complementary reverse RNA, the $\Delta T_m = -2.8$ °C.[g] Incorporating additional DNA-LNA carbamate linkages increased the destabilisation effect, lowering the

[g]The melting temperatures against the full RNA complement were once again measured at reduced salt concentrations (25 mM NaCl) in order to attain a melting temperature for the stable 2'OMe/PS control.

thermal stability by ∼1-2 °C per incorporation against DNA and ∼3 °C per incorporation against RNA (Figure 3.9B).

A.

Carbamate Linkage	DNA Complement		RNA Complement			
	200 mM		25 mM (full)		200 mM (fragment)	
	T_m (°C)	ΔT_m	T_m (°C)	ΔT_m	T_m (°C)	ΔT_m
ASO-1: 2'OMe/PS control	54.8 ±0.1	-	62.0 ±0.1	-	54.3 ±0.2	-
ASO-8: 1 DNA-LNA carb	53.9 ±0.1*	−0.9	59.2 ±0.1	−2.8	-	-
ASO-9: 2 DNA-LNA carbs	51.5 ±0.1*	−3.3	54.5 ±0.2	−7.5	48.2 ±0.4	−6.1
ASO-10: 1 LNA-LNA carb	-**	-	61.9 ±0.1	−0.1	-	-
ASO-11: 2 LNA-LNA carbs	53.5 ±0.2	−1.3	61.3 ±0.1	−0.7	51.6 ±0.2	−2.7

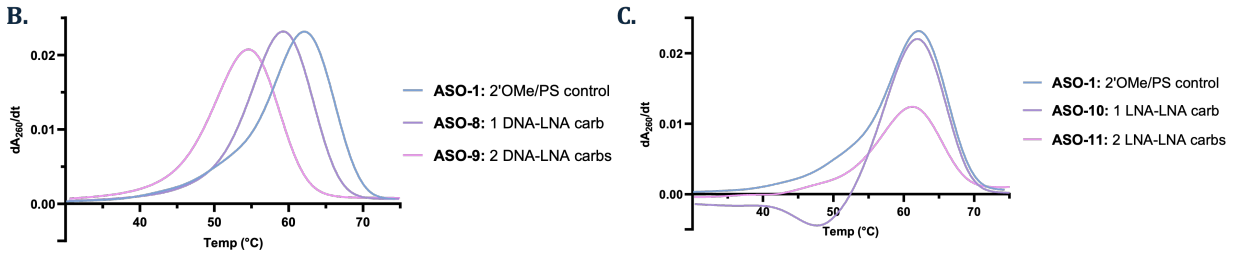


Figure 3.9: A. Table of melting temperatures for oligonucleotides containing carbamate backbones. *melting temperatures measured by Autumn Usher, Part II. **melting temperature not measured due to material insufficiency. Representative first derivative curves of duplex melting for oligonucleotides containing DNA-LNA carbamate backbones (**B.**) or LNA-LNA carbamate backbones (**C.**) against complementary RNA (full sequence) at reduced salt concentration (25 mM NaCl, 10 mM phosphate, pH = 7.0).

When an additional stabilising LNA moiety was added to the 5'-side of the linkage (LNA-LNA carbamate linkage) the melting temperature of the ASO:RNA heteroduplex was increased, as expected (from $\Delta T_m = -2.8$ °C to an improved −0.1 °C). Interestingly, additional LNA-LNA carbamate linkages (increased from one incorporation to two) only had a slight additive effect in destabilisation (from −0.1 °C to −0.7 °C, see Figure 3.9C). The inclusion of an additional LNA moiety also rescued the destabilisation of the carbamate at more physiological salt concentrations; at 200 mM NaCl using the fragment approach (described in Chapter 2), the ΔT_m of ASO-9 containing a single DNA-carbamate-LNA linkage = −6.1 °C, but which was increased to −2.7 °C in ASO-11 (single LNA-carbamte-LNA linkage). However, the stability of this LNA-LNA carbamate linkage in a heteroduplex with RNA does not confer to a heteroduplex with the DNA complement, as the ASO with two LNA-LNA

carbamate linkages was destabilising by −1.3 °C against DNA.[h]

It is interesting at this point to compare the thermal duplex stabilisation of the carbamate to the amide − while the amide linkage is also neutral, achiral and 5-bonds in length, the duplex melting temperature of a single "LNA-carbamate" incorporation (containing LNA on both sides) was 3.1 °C less stabilising against the RNA target than the LNA-amide ($\Delta T_m = +3.0$ vs. −0.1 °C for single incorporations of the LNA-amide and LNA-carbamate, respectively). Furthermore, the LNA-amide's stabilising properties were additive, but the LNA-carbamate was not stabilising nor largely additive; once two incorporations of these LNA-neutral linkages were added, the difference between the thermal stabilitiy of ASO-4 (2 LNA-amides) and ASO-11 (2 LNA-carbamates) grew to 6.2 °C − simply due to two atom changes from carbons to oxygens. This could be attributed to the increased electron density (potentially causing electrostatic repulsion) and increased rigidity of the linkage.

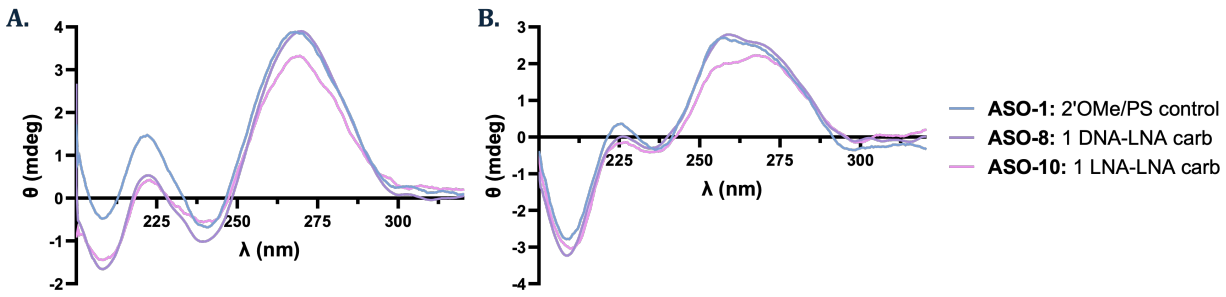


Figure 3.10: **A.** Circular dichroism spectra of ASOs against the DNA complement (200 mM NaCl, 10 mM phosphate, pH = 7.0). **B.** Circular dichroism spectra of ASOs against the RNA complement (25 mM NaCl, 10 mM phosphate, pH = 7.0).

The global duplex conformations of carbamate-containing ASOs were also measured and the circular dichroism spectra of ASO-1, ASO-8, and ASO-10 against complementary DNA and RNA are shown in Figure 3.10. The CD spectra of ASO-9 and ASO-11 are given in Figure C.1C and D. As was the case with the LNA-amide, all three oligonucleotides demonstrated global B-type helices against the DNA targent, with a maximum at 270 nm and a minimum at 245 nm. It is interesting to see the intensity of the maxima for ASO-10, containing the LNA-LNA carbamate linkage was reduced compared to ASO-8, containing the DNA-LNA

[h]Figures C.1A and B display the first derivative melting curves of all ASOs containing carbamate linkages against DNA and against the RNA fragment.

carbamate linkage. This indicates a shift toward more A-like behaviour, which is expected for an oligonucleotide containing an additional LNA moiety.[139] For the ASO:RNA heteroduplex, the 2'OMe/PS control and the DNA-LNA carbamate ASO continued to have similar CD signatures, while ASO-10, containing the additional LNA moiety, observed a second maxima or "plateau", seen with LNA-amides as well (Figure 3.10B). The bathochromic shift from the maxima at 260 nm to 275 nm indicates some potential A→B character or structural perturbation caused by the 5'-LNA (which is shown by crystal structure in the LNA-amide).[62]

3.6.2 Biophysical properties of additional LNA-neutral linkages

A.

LNA-neutral Linkage	DNA Complement (200 mM)		RNA Complement (25 mM)	
	T_m (°C)	ΔT_m	T_m (°C)	ΔT_m
ASO-1: 2'OMe/PS control	54.7 ±0.2	-	63.3 ±0.1	-
ASO-4: LNA-amide	61.7 ±0.2	+7.0	67.5 ±0.1	+4.2
ASO-11: LNA-carbamate	53.5 ±0.2	−2.0	61.3 ±0.1	−2.0
ASO-12: LNA-alkoxyamide	56.2 ±0.1	+1.5	65.0 ±0.1	+1.7
ASO-13: LNA-sulfamate	65.2 ±0.1	+10.5	73.3 ±0.1	+10.0

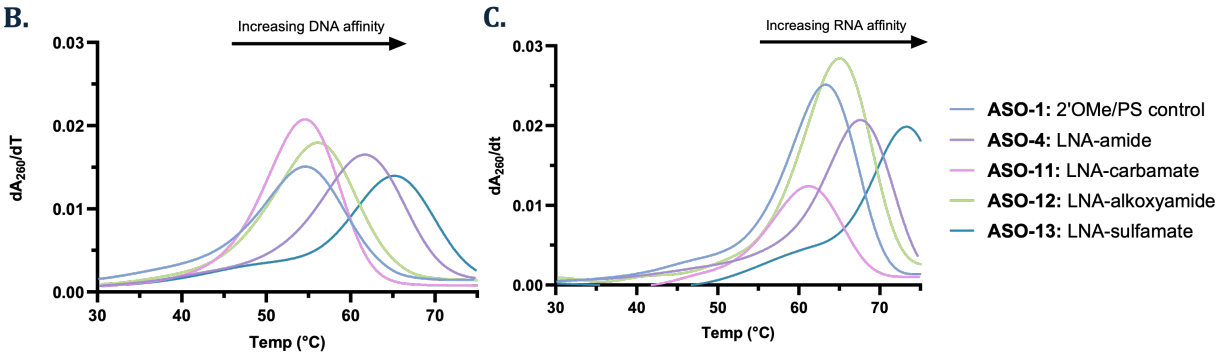


Figure 3.11: Thermal duplex stability of ASOs containing LNA-neutral linkages: **A.** Melting temperatures against the complementary DNA (10 mM phosphate, 200 mM NaCl, pH = 7.0) and complementary RNA (10 mM phopshate, 25 mM NaCl, pH = 7.0). First derivative curves of melting temperatures against complementary DNA (**B.**) and complementary RNA (**C.**).

The biophysical properties of ASOs containing the two additional LNA-neutral linkages (LNA-alkoxyamide and LNA-sulfamate) were explored in direct comparison to ASO-4 (2 LNA-amides) and ASO-11 (2 LNA-carbamates) (Figure 3.11A). The LNA-amide and LNA-carbamate linkage behaved as previously reported, with the LNA-amide linkage stabilising the ASO against both the DNA and RNA complements, and the LNA-carbamate linkage

mildly destabilising the ASO. Incorporating the LNA-alkoxyamide linkage in an ASO (ASO-12) was minimally stabilising against both the DNA and RNA targets (+1.5 and +1.7 °C, respectively), which was interesting as it was tolerated well despite being a longer 5-atom linkage. Potentially there is sufficient flexibility in the bonds of this linkage to adopt a conformation in which the internucleotide distance is not too large compared to the internucleotide distance of phosphodiesters. Finally, the ASO containing the LNA-sulfamate linkages (ASO-13) was surprisingly greatly stabilising against both the DNA and RNA targets (+10.5 and +10.0 °C, respectively), demonstrating that most likely, the sulfamate is the best mimic of a natural phosphodiester, followed by the amide linkage. The sulfur atom in the sulfamate contains a similar diester tetrahedral geometry to the phosphorus, and the amide contains a carbonyl group which is shown to mimic one of the P=O bonds.

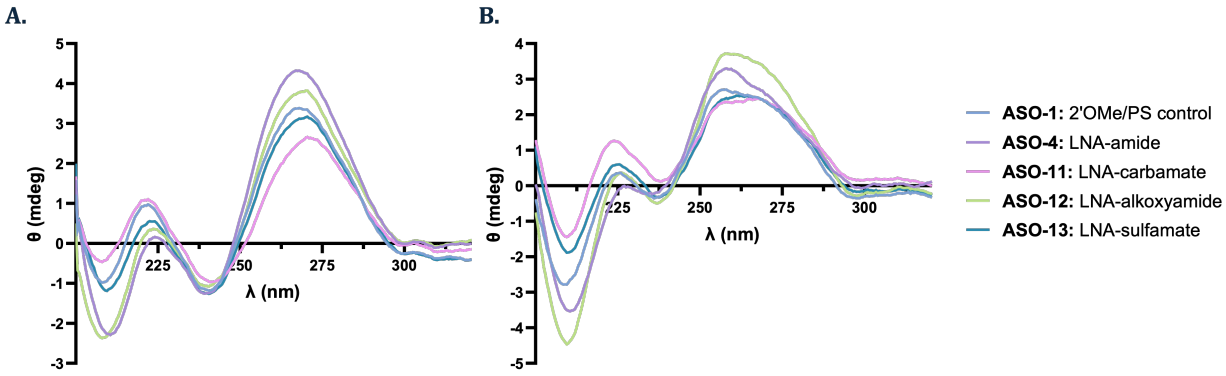


Figure 3.12: Circular dichroism spectra of ASOs containing LNA-neutral linkages in an ASO:DNA heteroduplex (**A.**) (200 mM NaCl, 10 mM phosphate, pH = 7.0) and ASO:RNA heteroduplex (**B.**) (25 mM NaCl, 10 mM phosphate, pH = 7.0).

The global conformation of ASOs containing these LNA-neutral linkages was also measured via circular dichroism (Figure 3.12). In line with previous reports of the LNA-amide and LNA-carbamate, the LNA-neutral linkages did not significantly perturb the overall duplex structure of the ASO against a DNA target, which is consistent with a B-type helix (Figure 3.12A), observing a maxima at 260 nm and a minima at 245 nm. However, differences among linkages can be observed when forming the ASO:RNA heteroduplex (Figure 3.12B). Interestingly, the LNA-amide appeared to have a greater maxima toward 260 nm, indicating a larger A-like character than the control oligonucleotide (ASO-1), while incorporation of the LNA-sulfamate linkage shows the greatest mixed A/B topology, characterised by the same

plateau-like maxima with a bathochromic shift from 260 nm towards 280 nm.

Finally, the nuclease resistance of these LNA-neutral linkages was tested in a serum stability study. All of the LNA-neutral linkages were very stable to endonucleases, as was the 2'OMe/PS control, over 96 h at 37 °C (Figure 3.13). A DNA control was used to verify the activity of the human serum endonucleases, which was digested within 96 h.[i]

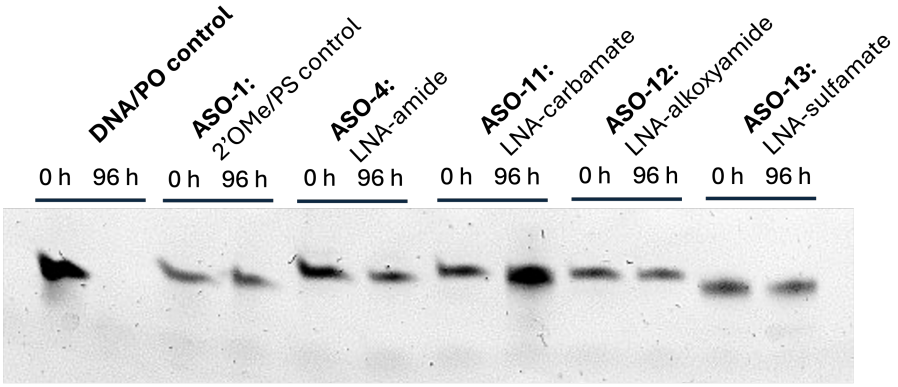


Figure 3.13: Nuclease resistance of ASOs containing LNA-neutral linkages and a DNA phosphate control oligonucleotide against human serum (96 h, 37 °C).

3.6.3 Molecular dynamics simulations of LNA-neutral linkages

Molecular dynamics (MD) simulations of these LNA-neutral linkages were conducted in a project collaboration with Martin Flerin[j]. A summary of the simulations' results are given in Figure 3.14 – briefly, the simulations started with the published crystal structure of a DNA:RNA heteroduplex (d-CTTTCTTTG/r-CAAAGAAAG)[62] which served as the DNA phosphodiester control. This was compared to a heteroduplex in which the DNA strand was altered to include two LNA sugars (LNA-phosphate) and heteroduplexes in which the same two LNA sugars and a neutral linker were included (LNA-neutral). The position of modification in the DNA:RNA heteroduplex was d-CT**T*T**CTTTG/r-CAAAGAAAG. MD simulations were run for three repeats of 1 µs simulations for each duplex, beginning with a randomly seeded set of initial velocities; the average structures, root mean square deviations

[i]The nuclease resistance of the other ASOs containing LNA-amide and LNA-carbamate linkages (e.g. ASO-3 or ASO-6 containing one or four LNA-amides, respectively) were assumed to be equally stable to these ASOs containing two incorporations.

[j]DPhil candidate supervised by Prof. Fernanda Duarte

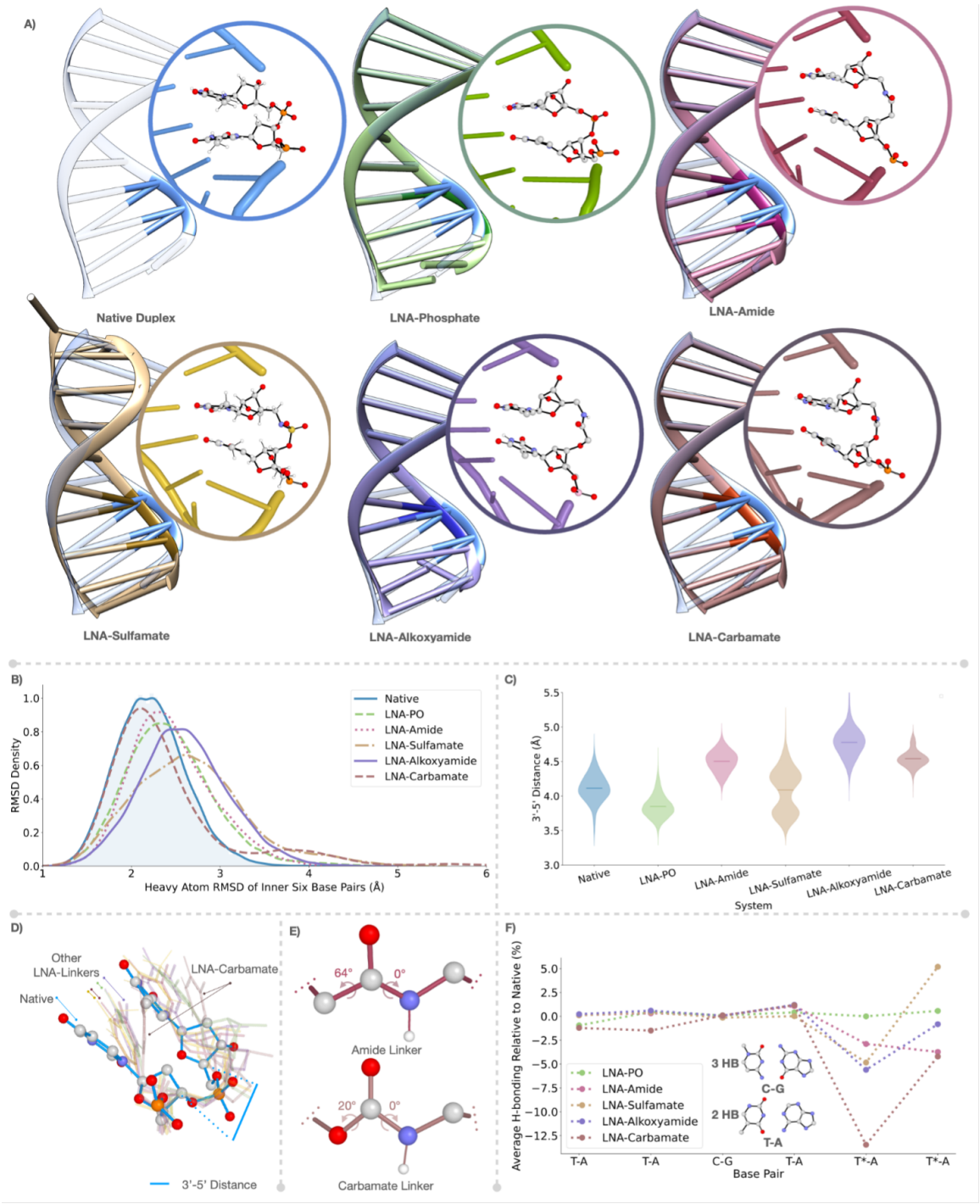


Figure 3.14: Summary of molecular dynamics (MD) simulations. **A.** Average structure of the most populated cluster for duplexes containing the LNA-neutral linkages, overlayed with the average structure of the native phosphate duplex. **B.** Histogram of RMSD distributions of the inner 6 base pairs of the duplexes containing LNA-neutral linkers. **C.** Violin plot of 3'-5' distance distributions between the two thymidine nucleotides; the horizontal line indicates the average distance. **D.** Representation of the 3'-5' internucleotide distance; overlayed structure of native TT dinucleotide (blue, stark) and all modified TT dinucleotides (transparent), aligned at the 5' monomer. **E.** Comparison of key median dihedral angles of amide and carbamate linkers over the simulation. **F.** Duration of H-bonding, as a percentage of simulation time, between each of the inner six base pairs in all modified duplexes, relative to the native duplex. Figure produced by Martin Flerin.

(RMSD) of the inner 6 base pairs, linker distance and angles, and hydrogen bonding between strands were compared (Figure 3.14A-F).

It was promising to note the overall good agreement among the MD simulation data and observed biophysical properties. The RMSDs of the inner six base pairs over an extended simulation time are a computational prediction of the duplex stability − as seen in the histogram of RMSD distributions (Figure 3.14B), all linkages formed stable duplexes over the course of the simulation. A slight bimodal distribution was shown for the LNA-carbamate, demonstrating a second less populated conformational cluster − this linkage was also the least stable in a duplex. The most noticeable difference among the LNA-neutral linkages, however, was observed at the nucleotide level.

Measuring the 3'-5' distances between the two thymines of the LNA-neutral dinucleotide also investigated the linker length and flexibility over the course of the simulation (represented in Figure 3.14D). The LNA-sulfamate ASO had the most similar average 3'-5' distance to the native phosphate, with a seemingly higher degree of flexibility as it also had the widest distribution of distances (Figure 3.14C). Experimentally, this linkage also demonstrated the largest A→B mixed topology in its CD spectrum against RNA (Figure 3.11C). B-form duplexes are globally more flexible than A-form duplexes;[Liu] the A→B-form shift induced by the LNA-sulfamate linkage demonstrates that MD simulations are a useful tool to further understand the global conformation of the duplex. The LNA-carbamate linkage had a very narrow distribution of distances between the 5'-3' thymines in the dinucleotide; simulation also revealed a lack of flexibility in the median dihedral angle (reduced from 64° around the α-carbon in the amide to 20° around the around the α-oxygen in the carbamate) (Figure 3.14E). This increased rigidity may be an explanation for its significantly lower thermal duplex stability compared to the amide.

Finally, the hydrogen bonding between the inner base pairs was considered (Figure 3.14E). Including two locked nucleotides (LNA-PO) with a phosphate backbone did not change the H-bonding of the system compared to the native phosphate diester, but the neutral linkages did affect the hydrogen bonding duration. At the modified position, the LNA-carbamate

caused the greatest disruption to H-bonding duration (up to 13%) at both the 3' and 5' position, potentially due to its increased rigidity. Surprisingly, the LNA-sulfamate slightly improved H-bonding duration at the 5'-position; this linkage was also the most stabilising when forming an ASO:RNA heteroduplex (ΔT_m of ASO-13 = +10.0 °C).

Experimental results indicated that the structural differences among amide-inspired neutral linkages can cause great ranges of duplex stabilisation and conformations − it was promising to see a computational model also generated similar trends. These preliminary results indicate a good correlation between molecular dynamics simulations and experimental biophysical properties; this joint collaboration proposes that computational modelling of artificial linkages and chemical modifications in oligonucleotide duplexes is a useful tool and should be utilised to predict favourable properties and guide rational ASO design.[k]

3.7 Biological activity of ASOs containing LNA-amide-inspired linkages

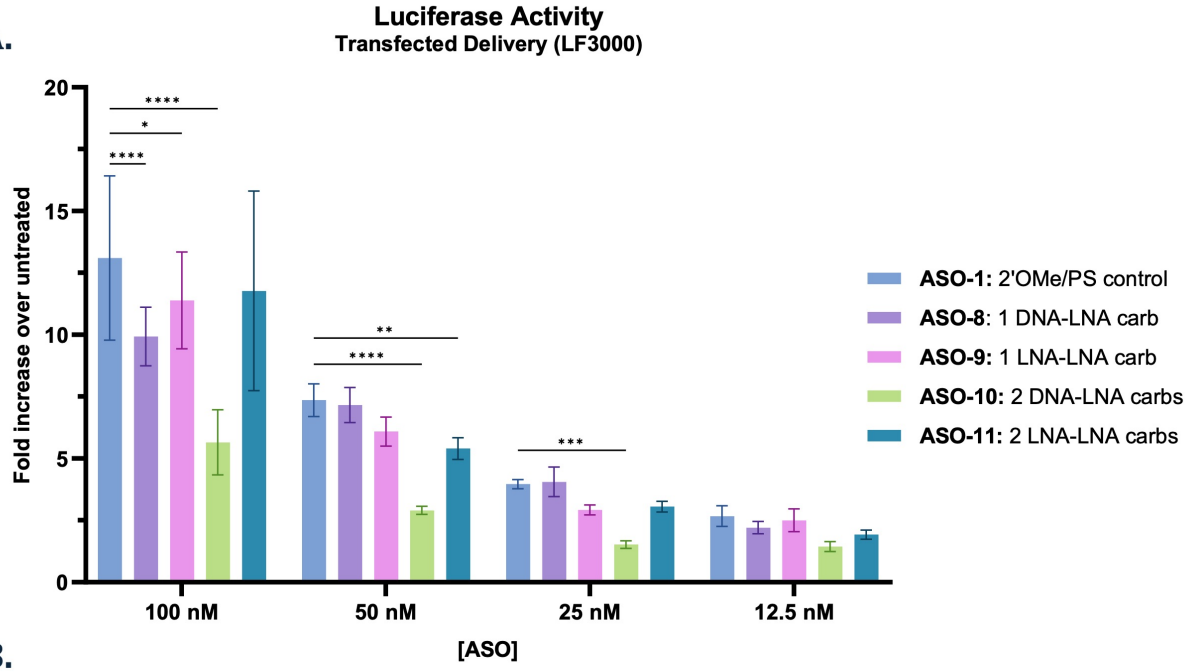
3.7.1 The splice-switching activity of ASOs containing carbamate linkages

The luciferase reporter assay used in Chapter 2 was also used to evaluate the splice-switching activity of ASOs 8-11 containing carbamate linkages. For the first time, DNA-LNA carbamate and LNA-LNA carbamate linkages were tested in a biological system; it was hypothesised that the LNA-LNA carbamate would be more active due to its increased thermal duplex stability. However, if the DNA-LNA carbamate was found to be equally active, this would be a useful finding to reduce the complexity of syntheses going forward − DNA-LNA neutral linkages are synthetically more accessible than LNA-LNA dinucleotides.

Unfortunately, both the ASOs containing DNA-LNA carbamate and LNA-LNA carbamate linkages were equally or less active than the 2'OMe/PS control (ASO-1) using transfection

[k]I thank Martin Flerin for optimising, modeling, running, and interpreting all of the MD simulations data in this project collaboration.

A.



B.

Fold increase over **ASO-1**: 2'OMe/PS control	**100 nM**	**50 nM**	**25 nM**	**12.5 nM**
ASO-8: 1 DNA-LNA carb	0.8 ±0.2	1.0 ±0.1	1.0 ±0.2	0.8 ±0.2
ASO-9: 1 LNA-LNA carb	0.9 ±0.3	0.8 ±0.1	0.7 ±0.1	0.9 ±0.2
ASO-10: 2 DNA-LNA carbs	0.4 ±0.1	0.4 ±0.0	0.4 ±0.0	0.5 ±0.1
ASO-11: 2 LNA-LNA carbs	0.9 ±0.4	0.7 ±0.1	0.8 ±0.1	0.7 ±0.1

Figure 3.15: **A.** Dose-response splice-switching activity of ASOs containing carbamate linkages, delivered by transfection with Lipofectamine 3000. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS control (ASO-1) as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing carbamate backbones.

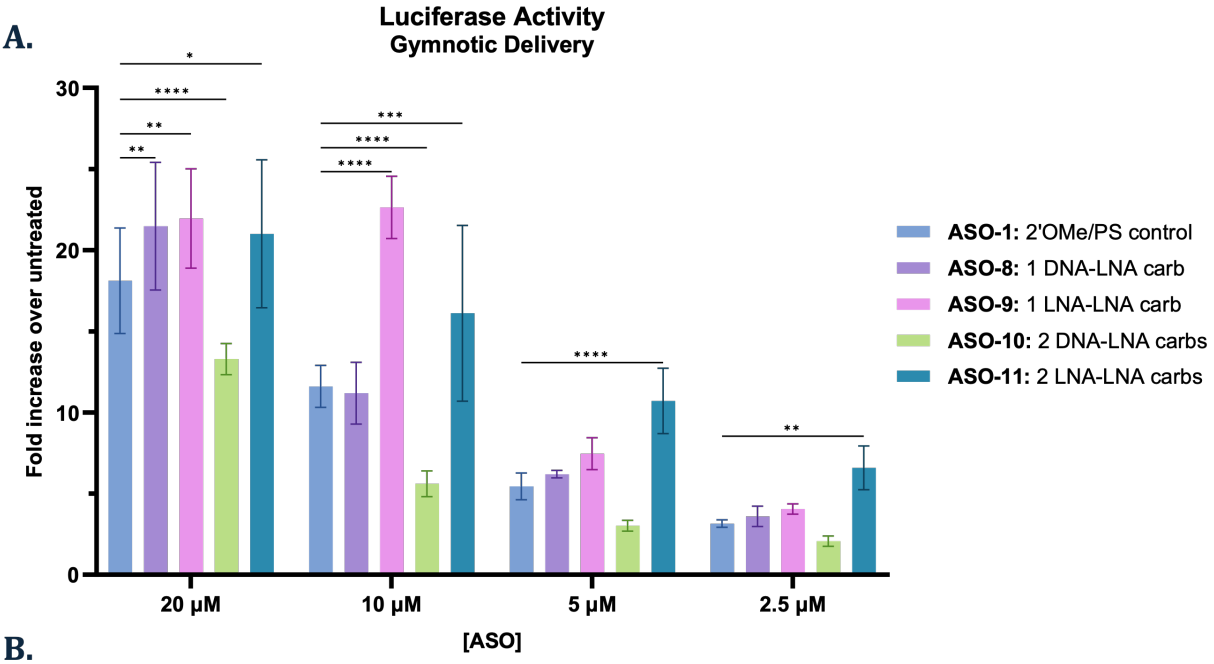
(Figure 3.15); no carbamate-containing ASO was more active than the positive control. A single incorporation of the DNA-LNA carbamate linkage (ASO-8) resulted in a range of 0.8 to 1.0-fold change compared to ASO-1, demonstrating equal if not reduced activity after intracellular delivery. An additional DNA-LNA carbamate linkage (ASO-10) significantly reduced the inherent activity (0.4 to 0.5-fold change compared to ASO-1), demonstrating that ASOs with multiple incorporations are increasingly less active with additional DNA-

carbamate-LNA linkages. The LNA-LNA carbamate counterparts (ASO-10 and ASO-11) are similarly active; a single incorporation (ASO-10) resulted in a range of 0.7 to 0.9-fold change over ASO-1. The additional LNA-moiety in ASO-10 increased its thermal duplex stability compared to ASO-8, but did not render the ASO more active. However, interestingly, the activity of the ASOs containing LNA-LNA carbamates was not decreased with an additional incorporation. The activity of ASO-11 (containing two LNA-carbamate-LNA linkages) also ranged from 0.7 to 0.9-fold change over ASO-1, a different trend observed to the DNA-carbamate-LNA linkage. Additional work should investigate ASOs with three, four, or more incorporations of these two linkages to ascertain whether this effect continues to be observed as the ASO becomes increasingly modified. While not having nearly as much inherent potency as the LNA-amides, the LNA-carbamates remain worth investigating in a gymnotic assay to evaluate their activity in a clinically-relevant delivery method.

The ASOs containing carbamate linkages displayed some increased splice-switching activity relative to the 2'OMe/PS control (ASO-1) once delivered via gymnosis (Figure 3.16A) − this slight difference in activity between the two delivery methods may be attributed to some increased cellular uptake by the ASOs containing carbamate backbones. However, their activity remained lackluster compared to the previously-explored linkages such as the LNA-amide. ASO-8 (0.8 to 1.0-fold change over ASO-1 via transfection) improved its activity to be equal or slightly above the control ASO (1.0 to 1.2-fold change via gymnosis). This linkage remains poorly tolerated however, in additional quantities − the gymnotic activity of ASO-9 remained well below that of ASO-1 (0.5 to 0.7-fold change). ASOs containing LNA moieties on both sides of the carbamate observed an increase in activity − ASO-11 (containing two incorporations), was the most potent ASO delivered by this method, and a 2.1-fold change over ASO-1 was observed at the lowest gymnotic dose (2.5 µM). ASOs containing LNA-LNA carbamate linkages may be more active in a gymnotic assay due to increased cellular uptake − potentially through increased productive protein-binding profiles given their increased number of LNA-moieties, but this requires further exploration.

Unfortunately, the DNA-LNA carbamate linkage was not a very promising candidate for

A.



B.

Fold increase over **ASO-1**: 2'OMe/PS control	**20 μM**	**10 μM**	**5 μM**	**2.5 μM**
ASO-8: 1 DNA-LNA carb	1.2 ±0.3	1.0 ±0.2	1.1 ±0.2	1.1 ±0.2
ASO-9: 1 LNA-LNA carb	1.2 ±0.3	2.0 ±0.3	1.4 ±0.3	1.3 ±0.1
ASO-10: 2 DNA-LNA carbs	0.7 ±0.1	0.5 ±0.0	0.6 ±0.1	0.7 ±0.1
ASO-11: 2 LNA-LNA carbs	1.2 ±0.3	1.4 ±0.5	2.0 ±0.5	2.1 ±0.5

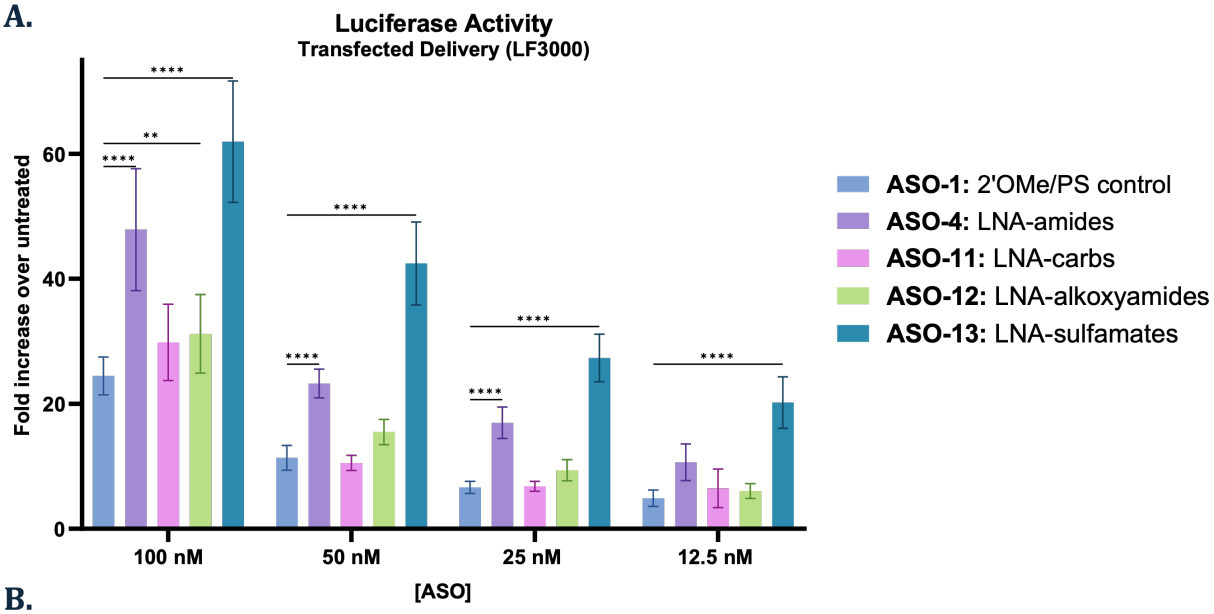
Figure 3.16: **A.** Dose-response splice-switching activity of ASOs containing carbamate linkages, delivered by gymnosis. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS ASO as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n \geq 2$); each biological replicate is the mean ±SD of technical replicates ($n \geq 2$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing carbamate backbones.

improving the splice-switching activity of an ASO; while DNA-LNA dinucleotides are synthetically more accessible, their equal, or reduced, activity compared to a 2'OMe/PS control underscores the importance of flanking both sides of a destabilising linkage, such as the carbamate, with stabilising LNA sugars. Combination of LNA with linkages such as these are under investigation – for example, various triazole linkers have been combined with LNA on the 5', 3', and 5' and 3'-flanks of the linkage – LNA on the 3'-side (or both) improved the binding affinity of an ASO:RNA heteroduplex,[83] and oligonucleotides with LNA on either

side of a triazole maintained gene silencing activity in siRNA.[164] Additionally, the magnitude of the effect of the linkage became more observable once two linkages were incorporated – this supports the hypothesis that two incorporations of a novel linkage are sufficient to evaluate them. Building on this, the additional LNA-neutral linkages (LNA-alkoxyamide and LNA-sulfamate) were evaluated in ASOs with two incorporations each.

3.7.2 The splice-switching activity of ASOs containing additional LNA-neutral linkages

A.



B.

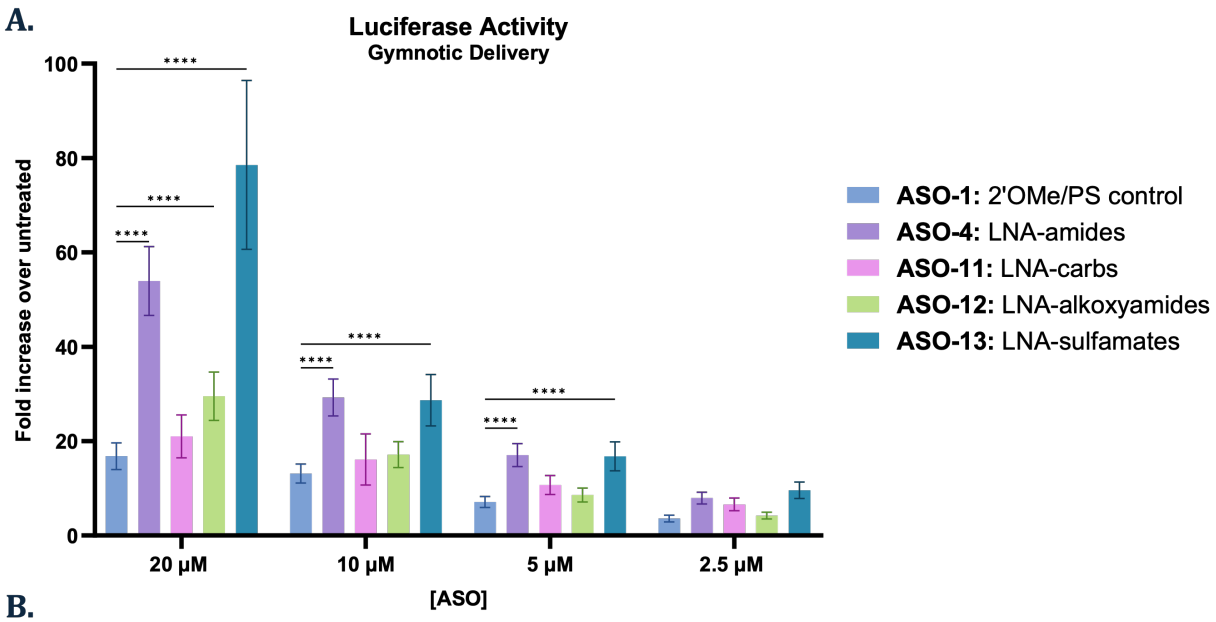
Fold increase over **ASO-1**: 2'OMe/PS control	**100 nM**	**50 nM**	**25 nM**	**12.5 nM**
ASO-4: LNA-amide	2.0 ±0.5	2.0 ±0.4	2.6 ±0.5	2.2 ±0.8
ASO-11: LNA-carbamate	1.2 ±0.3	0.9 ±0.2	1.0 ±0.2	1.3 ±0.7
ASO-12: LNA-alkoxyamide	1.3 ±0.3	1.4 ±0.3	1.4 ±0.3	1.2 ±0.4
ASO-13: LNA-sulfamate	2.5 ±0.5	3.7 ±0.9	4.1 ±0.8	4.1 ±1.4

Figure 3.17: A. Dose-response splice-switching activity of ASOs containing LNA-neutral linkages, transfected with Lipofectamine 3000. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS ASO as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing LNA-neutral backbones.

The luciferase reporter assay was repeated for ASO-4 (2 incorporations of the LNA-amide) and ASO-11 (2 incorporations of the LNA-carbamate) in order to compare the activities to ASO-12[1] (2 incorporations of the LNA-alkoxyamide) and ASO-13 (2 incorporations of the LNA-sulfamate). The inherent activities of each ASO were compared first via delivery by transfection (Figure 3.17A). The LNA-carbamate did not demonstrate significantly more splice switching than the 2'OMe/PS control, ranging from 0.9 to 1.3-fold increase over ASO-1 – these values are slightly higher than reported in the previous assay (0.7 to 0.9-fold increase). Despite conducting three biological replicates (consisting of three technical replicates) for each assay, there remained some variability among plates for the extent of fold increase in normalised luminescence over untreated cells. However, all ASOs demonstrated higher fold increase in normalised luminescence over the untreated cells, including the positive control ASO-1, and the fold increase over the ASO-1 remain within error. The ASO containing LNA-alkoxyamides (ASO-12) did not demonstrate significantly more splice-switching than the 2'OMe/PS control (ranging from 1.2 to 1.4-fold change over ASO-1). The LNA-amide ASO (ASO-4) had very high inherent activity, in line with the results of Chapter 2 (2.0 to 2.6-fold increase over ASO-1) and the LNA-sulfamate ASO (ASO-13) was even more active at all doses (ranging from 2.5 to 4.1-fold increase). ASO-13, containing two LNA-sulfamate linkages, was also the most potent, demonstrating ∼4-fold increase over the 2'OMe/PS-modified ASO at the lowest transfected doses (25 and 12.5 nM).

All ASOs were also evaluated by gymnotic delivery (Figure 3.18) – as expected, inherently more active ASOs (LNA-amide and LNA-sulfamate) also had improved gymnotic activity compared to the 2'OMe/PS control (ASO-1). The LNA-carbamate ASO, as discussed previously, observed some increase in activity when delivered via gymnosis (1.2 to 1.8-fold increase over ASO-1), but its activity remained low compared to the linkages such as the LNA-amide and LNA-sulfamate. The LNA-amide and LNA-sulfamate ASOs were equally potent at gymnotic doses below 20 µM (∼2.5-fold more active than ASO-1). According to MD simulations and experimental data, these two linkages are likely the best isosteres of the natural phosphodiester linkage, while increasing the thermal duplex stability of the

[1]Synthesised by Dr. Alice Kennett



B.

Fold increase over **ASO-1**: 2'OMe/PS control	**20 µM**	**10 µM**	**5 µM**	**2.5 µM**
ASO-4: LNA-amide	3.2 ±0.7	2.2 ±0.5	2.4 ±0.5	2.2 ±0.6
ASO-11: LNA-carbamate	1.2 ±0.3	1.2 ±0.5	1.5 ±0.4	1.8 ±0.5
ASO-12: LNA-alkoxyamide	1.8 ±0.4	1.3 ±0.3	1.2 ±0.3	1.2 ±0.3
ASO-13: LNA-sulfamate	4.7 ±1.3	2.2 ±0.5	2.4 ±0.6	2.7 ±0.7

Figure 3.18: A. Dose-response splice-switching activity of ASOs containing LNA-neutral linkages, delivered via gymnosis. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS control as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n \geq 2$) with the exception of ASO-13 at 20 µM, in which n=1; each biological replicate is the mean ±SD of technical replicates ($n \geq 2$). **B.** Table of fold-increase over ASO-1 (2'OMe/PS control) for ASOs containing LNA-neutral backbones.

ASO:RNA heteroduplex. The LNA-sulfamate, by MD, was also the most flexible linker and even increased the H-bonding duration of the 5'-thymidine residue to its adenine pair.

The aims of the LNA-amide-inspired project were to further investigate the structure-activity-relationship of LNA-neutral hetero-atom-containing linkages. An alternative linkage to the LNA-amide was sought with equal, if not improved, therapeutic properties and a more facile synthesis. The LNA-sulfamate demonstrated such a promising candidate, with high thermal stability, significantly improved splice-switching activity over the 2'OMe/PS control,

and a shorter synthesis than the LNA-amide dimer phosphoramidite.[m]

3.8 Conclusions

In conclusion, phosphoramidite dimers containing carbamate linkages (flanked by DNA-LNA and LNA-LNA sugars) were synthesised and incorporated into therapeutic 2'OMe/PS chimeric ASOs for the first time. The biophysical properties of these ASOs were studied − while the DNA-LNA carbamate was destabilising in an ASO:RNA heteroduplex, the LNA-moiety at the 5'-monomer in the LNA-LNA carbamate ASO could "rescue" some of this instability. The splice-switching activity of the ASOs containing DNA-LNA and LNA-LNA carbamates were compared in a luciferase reporter assay − ASOs containing DNA-LNA carbamates were similarly or less active the 2'OMe/PS control ASO, while LNA-LNA carbamate ASOs were slightly more active than the control ASO during gymnotic delivery. The ASO with two incorporations of the LNA-carbamate was the most potent when delivered via gymnosis − it was therefore included in the comparative study of LNA-neutral linkages.

A novel dinucleotide containing an amino-oxy-amide linkage was also synthesised via a novel locked 5'-amino-oxy thymidine monomer. Unfortunately, handling the reactive amino-oxy-linked dimers was difficult and phosphitylation produced by-products. Low yields precluded the incorporation of the amino-oxy linkage in ASOs. Instead, other neutral amide-inspired linkages such as the LNA-alkoxyamide and LNA-sulfamate were investigated[n] and ASOs containing two incorporations of each amide-inspired linkage were synthesised. Among these, the LNA-sulfamate was the most promising linkage. It had the highest duplex stabilisation (+10.0 °C in an ASO:RNA heteroduplex), highest flexibility (according to simulated 3'-5'-dinucleotide distances), and the highest splice-switching activity in both the transfected and gymnotic assays, outperforming even the LNA-amide at the highest gymnotic dose (4.7-fold increase over ASO-1 at 20 µM). It was additionally encouraging to see a correlation

[m] While both the LNA-amide and LNA-sulfamate share a locked 5'-NH_2 monomer, the LNA-amide additionally requires the synthesis of the locked 3'-COOH monomer over eight steps, while the LNA-sulfamate requires the locked 3'-hydroxyl monomer which is commercially available.

[n] In collaboration with Dr. Alice Kennett, Dr. Belma Zengin-Kurt, and Martin Flerin.

among trends across computational and experimental approaches − this collaborative project promotes an interdisciplinary approach when investigating new ASO chemistries for rational ASO design.

4

The synthesis and study of ASOs containing LNA-triester linkages

Contents

4.1 Introduction

This chapter aims to discuss a final neutral linkage – an alkylated phosphodiester (5'-OP(O)(OR)O-3') or alkylated phosphorothioate (5'-OP(S)(OR)O-3') linkage. Functionalisation of one of the non-bridging oxygens with an alkyl-R group in the phosphate backbone has precedence in the literature, with many different functional groups for various purposes. These include, but are not limited to, DNA probes, siRNA pro-drugs, and gapmers.

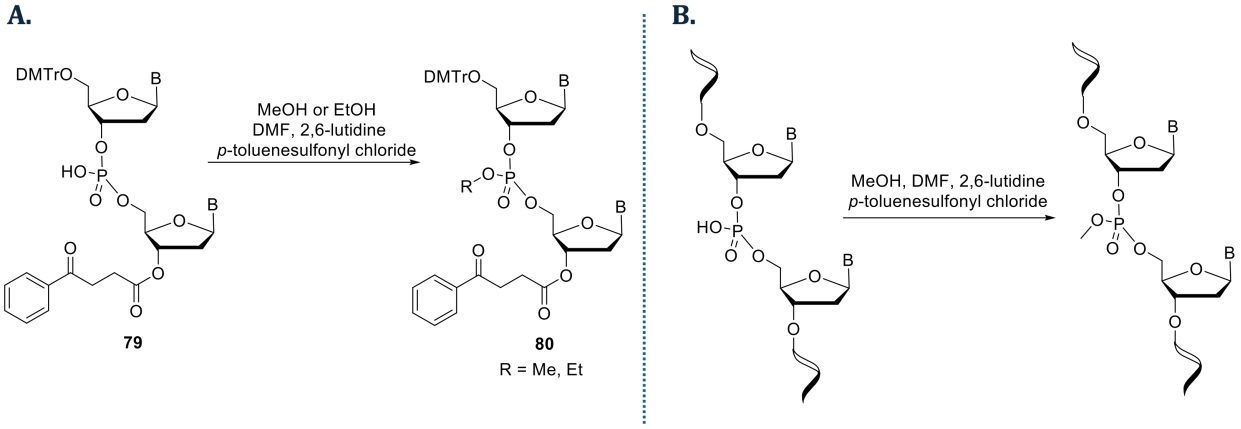
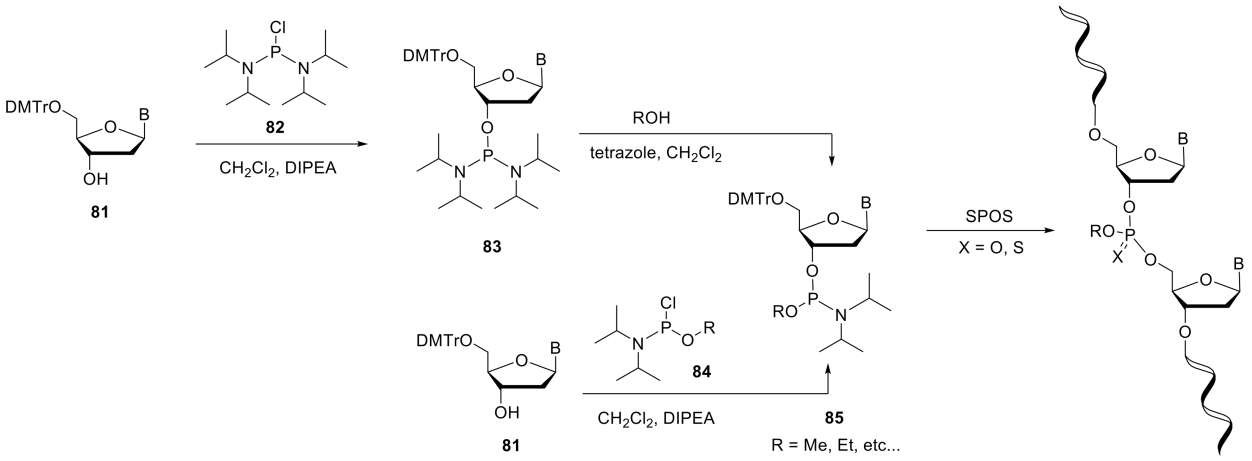


Figure 4.1: Installation of the triester backbone by alkylation of a phosphodiester linkage in dinucleotide **79** (**A.**), reported by T'so and Miller[165] or in an oligonucleotide (**B.**), reported by Moody et al.[166]

Synthesis of dinucleotides containing an alkylated triester backbone, the "neutral analogues of dinucleotide monophosphates", was first reported in 1971 by T'so and Miller.[165] The alkyl R-groups (in this preliminary case, methyl and ethyl) were installed by removal of the β-cyano-ethyl protecting group and alkylation of the free alcohol in the phosphodiester linkage of dimer **79** by treatment with methanol or ethanol and *p*-toluenesulfonyl chloride (giving dimer **80**) (Figure 4.1A).[165] These dinucleotides displayed significant resistance to snake venom and spleen phosphodiesterases, and preliminary biophysical studies demonstrated extremely reduced electrostatic repulsion during duplex formation with the complementary dinucleotides.[167] Moody et al. describe the synthesis of a fully methylated neutral oligonucleotide by a similar synthesis, beginning with a native phosphodiester oligonucleotide and



Scheme 4.1: Installation of the triester backbone by incorporation of a monomer phosphoramidite **85**, synthesised via bis(diisopropylamino)phosphoramidite monomer **83** or via an alkylated di-diisopropylamino-phosphine reagent **84**.

treatment of all phosphodiester linkages with methanol and *p*-toluenesulfonyl chloride (Figure 4.1B). In a 12-mer, they report a ΔT_m of +14 °C compared to an unmodified oligonucleotide. At very low salt concentrations, this fully charge-neutral oligonucleotide was still able to form duplexes due to the lack of electrostatic repulsion.[166] This "MPTE" (methyl phosphate triester) linkage has since been explored in antisense mechanisms, for example the site-selective blocking of the viral HIV-1 RNA in order to inhibit its infectivity.[168,169] Due to its extremely high binding affinity, it has also been utilised in DNA probes (termed "nDNA", or neutral DNA). Probes partially modified by MPTE linkages have also been shown to inhibit the reverse transcription of oncogenic miRNA (miR-21), with an efficiency comparable to probes modified by LNA.[140]

An alternative approach to synthesising alkyl triesters uses an alkylated phosphoramidite monomer (**85**) which is coupled during solid-phase synthesis, as shown in Scheme 4.1. Literature reports have synthesised this monomer via phosphitylation of a DNA monomer **81** by reaction with bis(diisopropylamino)chlorophosphine **82**; one of the diisopropylamino-moieties in the phosphordiamidite **83** is then replaced by the desired alcohol in the presence of tetrazole to give the diester phosphoramidite **85** (Scheme 4.1).[170] Alternatively, treatment of DNA monomer **81** with a phosphitylating reagent already functionalised with the alkyl R-group is also reported; the alkylated chloro-isopropylamino-phosphine reagent **84** is synthesised by re-

action of bis(diisopropylamino)chlorophosphine **83** with the desired alcohol.[90] In 1996, Iyer et al. coupled a methoxy *N*-pent-4-enoyl protected ("MEPNT") phosphoramidite monomer.[171] They conducted both oxidation of the linkage with iodine to give a P(O)OMe backbone and sulfurisation of the linkage with EDITH reagent to give a P(S)OMe backbone. A similar strategy will be adopted in this chapter to compare the effects of the second non-bridging atom (O vs. S) in an alkylated triester linkage. In the literature, multiple phosphorothioate oligonucleotides were synthesised (ranging from one to three incorporations of the methoxy-triester); they reported very marginal changes to the thermal stability of the ASO:DNA or ASO:RNA duplexes.[171]

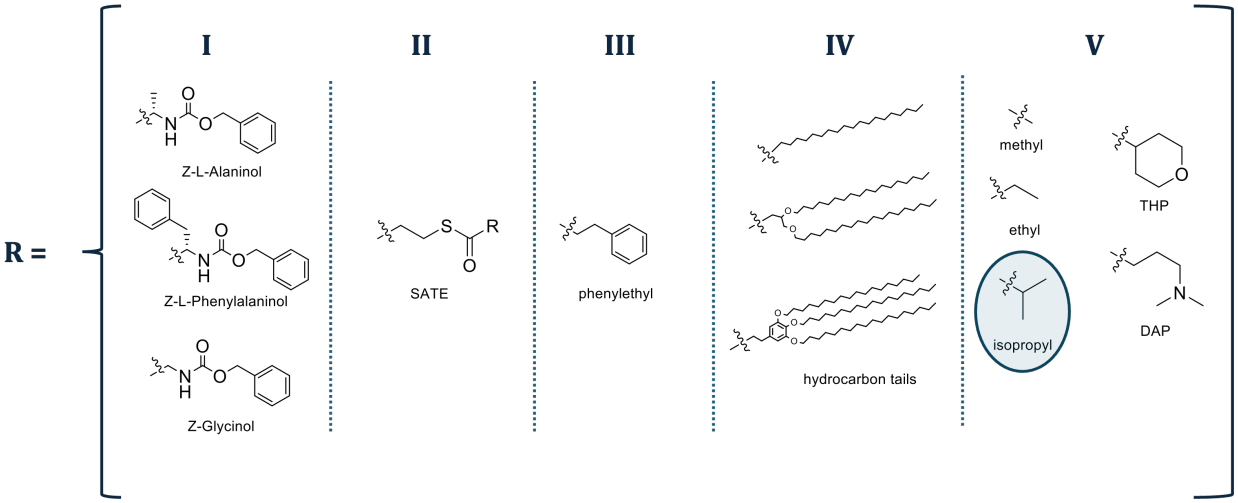


Figure 4.2: Alkyl triester R groups previously explored in the literature. I: Cbz-protected amino acids in ASOs,[170] II: S-acyl-2-thioethyl groups for a pro-drug approach,[172] III: phenylethyl in siRNA,[173] IV: hydrocarbon tails on siRNA ends,[174] V: methyl, ethyl, isopropyl, THP, and DAP in gapmers.[90] Based on reports by the literature, the isopropyl group (highlighted in blue) was selected for further investigation.

Of course, R-groups beyond methyl and ethyl have been explored in the literature. Examples of promising R-group functionalisation are given in Figure 4.2:

- In 2015, Monfregola et al. included a variety of Cbz-protected amino acids in order to synthesise oligonucleotides containing neutral or cationic hydrophobic linkages – they reported little change in thermal stability, but increased HeLa cell uptake by fluorescence-labelled oligonucleotides containing four incorporations of Cbz-protected amino acid alcohols (demonstrated by flow cytometry) (Figure 4.2(I)).[170]

- A pro-drug approach − Meade et al. synthesised oligonucleotides containing neutral S-acyl-2-thioethyl (SATE) phosphotriester groups which were cleaved by thioesterases once the charge-neutral ASO entered the cytoplasm. These "siRNNs" (short interfering ribonucleic neutrals) demonstrated extended albumin serum binding and had promising in vivo RNAi activation in mice (Figure 4.2(II)).[172]

- Large aromatic R-groups, such as the phenylethyl group (Figure 4.2(III)) were incorporated into siRNA oligonucleotides by Tsubaki et al. and demonstrated good gene-silencing activity in a dual firefly luciferase assay. They reported minimal changes to the thermal stability within double stranded siRNA as well as enhanced nuclease stability at the 3'-end of the passenger strand.[173]

- Large hydrocarbon chains have also been reported in the tail ends of siRNA (Figure 4.2(IV)). Three novel hydrophobic groups, stearyl phosphate ester, 2,3-distearyloxyphosate ester, and 3,4,5-tris(stearyloxy)phenethyl phosphate ester, were shown to increase gene silencing activity.[174]

Alkylation of the phosphate linkage has been widely established, but there still remain infinite R-groups that are possible − not all can be evaluated for their in vivo therapeutic properties. Ionis Pharmaceuticals reported a thorough investigation of select triester phosphate linkages in a gapmer mechanism of action. They synthesised gapmers containing single incorporations of ethyl, isopropyl (iPr), tetrahydrofuranosyl (THP), and dimethylaminopropyl (DAP) phosphate triesters in multiple positions in the gap (Figure 4.2(V)). Incorporated into a DNA/PO oligonucleotide, all were moderately destabilising (ΔT_m = 1-2 °C) compared to the control DNA, with the OiPr and OTHP groups being most stable. These modifications were evaluated in position four of toxic 3-10-3 2'-cEt gapmers; they demonstrated equally good potency for reducing mRNA in the liver; the ASOs containing OiPr and OTHP linkages were slightly more potent than a control methoxyphosphate (R = Me) control. Hepatotoxicity was also reduced, demonstrating an improved therapeutic index.[90]

Here, the triester was utilised as a neutral backbone in a therapeutic 2'OMe/PS single-stranded ASO (rather than a pro-drug, siRNA, or probe) − encouraged by the results of Ionis Pharmaceutical's gapmer study, the isopropyl R group was selected for further investigation in a splice-switching mechanism of action which has previously not yet been investigated. Its combination with LNA was hypothesised to increase thermal duplex stability (which was previously reported for the OiPr group as minimally de-stabilising), improve A-form characteristics, and enable good comparison to other charge-neutral linkages synthesised thus far. Finally, it would be interesting to evaluate the impact of oxidising or sulfurising the linkage, as previous in vivo gapmer work uses phosphate triesters (P(O)OR) rather than phosphorothioates (P(S)OR).

4.2 Aims of the Chapter

The aims of this chapter were to investigate a final neutral linkage, the isopropyl phosphate or isopropyl phosphorothioate, flanked by LNA in splice-switching therapeutic ASOs. The LNA-neutral linkages studied thus far, such as the amide and sulfamate, had very promising therapeutic properties, but require either dinucleotide phosphoramidites (16 distinct compounds) or two different structurally-modified monomers (8 distinct compounds). It was envisioned that a neutral linkage which was synthetically more accessible (requiring only the four monomers), but retained these promising properties, could more immediately bring the benefits of LNA-neutral ASOs to in vivo work and beyond.

The chapter aims to:

1. Synthesise novel locked isopropyl triester phosphoramidite monomers. Synthesise additional phosphoramidites required by alkyl-amine-based oligonucleotide deprotection conditions.

2. Incorporate the triester phosphoramidites in 2'OMe/PS chimeric ASOs in which the remaining non-bridging atom is oxygen or sulfur (see Figure 4.3).

3. Compare ASOs containing "LNA-OiPr (PO)" vs. "LNA-OiPr (PS)" linkages for their

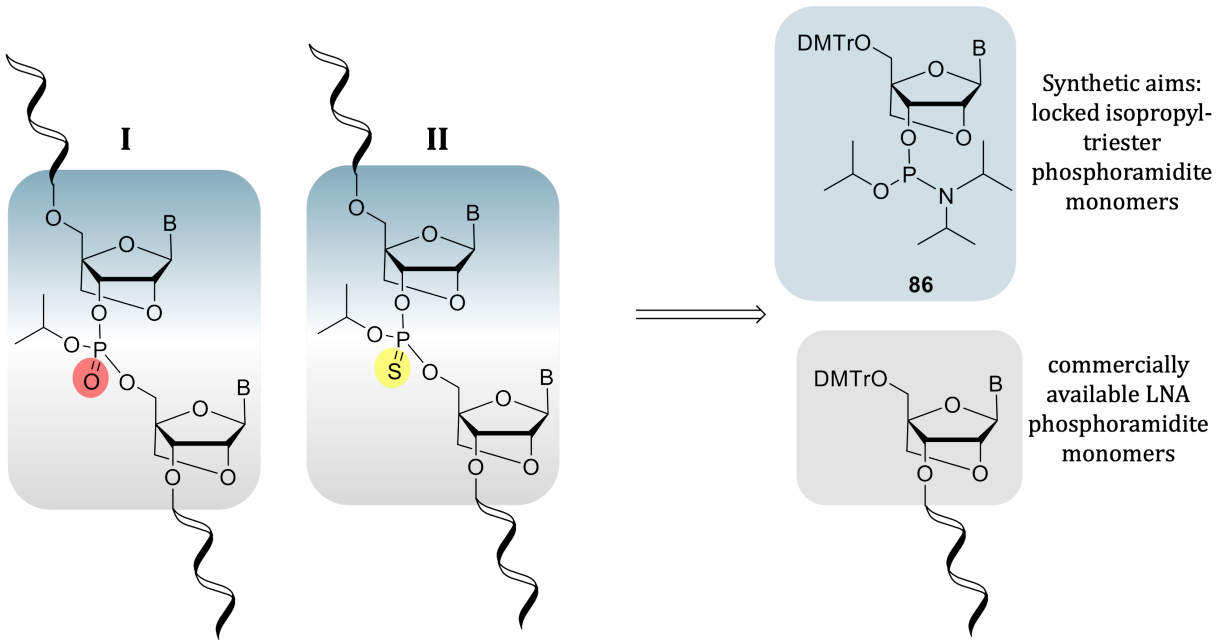


Figure 4.3: Synthetic aims of the chapter: A neutral isopropyl linkage between two locked nucleic acids achieved by coupling a novel 5'-DMTr-protected locked triester phosphoramidite **86** onto a 3'-locked monomer (commercially available) in the growing oligonucleotide chain. The P(III) backbone can be oxidised in the SPOS cycle to give a phosphodiester linkage (I) or sulfurised to give a phosphorothioate linkage (II).

biophysical and biological properties.

4. Take forward the most promising isopropyl triester linkage and synthesise additional ASOs in which the number of LNA-OiPr incorporations is varied (one to four incorporations).

    - Evaluate the biophysical properties of ASOs containing LNA-OiPr linkages.
    - Evaluate the biological properties of ASOs containing LNA-OiPr linkages.

4.3 Synthesis of LNA-OiPr phosphoramidite monomers

Installation of the alkyl linkage in the ASO was envisioned by a "monomer approach" – coupling an isopropyl phosphoramidite monomer **86**, rather than alkylation of a phosphate or phosphorothioate bond as described in Figure 4.1. A one-pot in situ reaction of 5'-DMTr-LNA-T-3'-OH monomer **43** with bis(diisopropylamino)chlorophosphine (**82**), followed by additional of isopropyl alcohol was found to be low-yielding and produced many phosphity-

lated by-products which could go on to react during oligonucleotide synthesis.

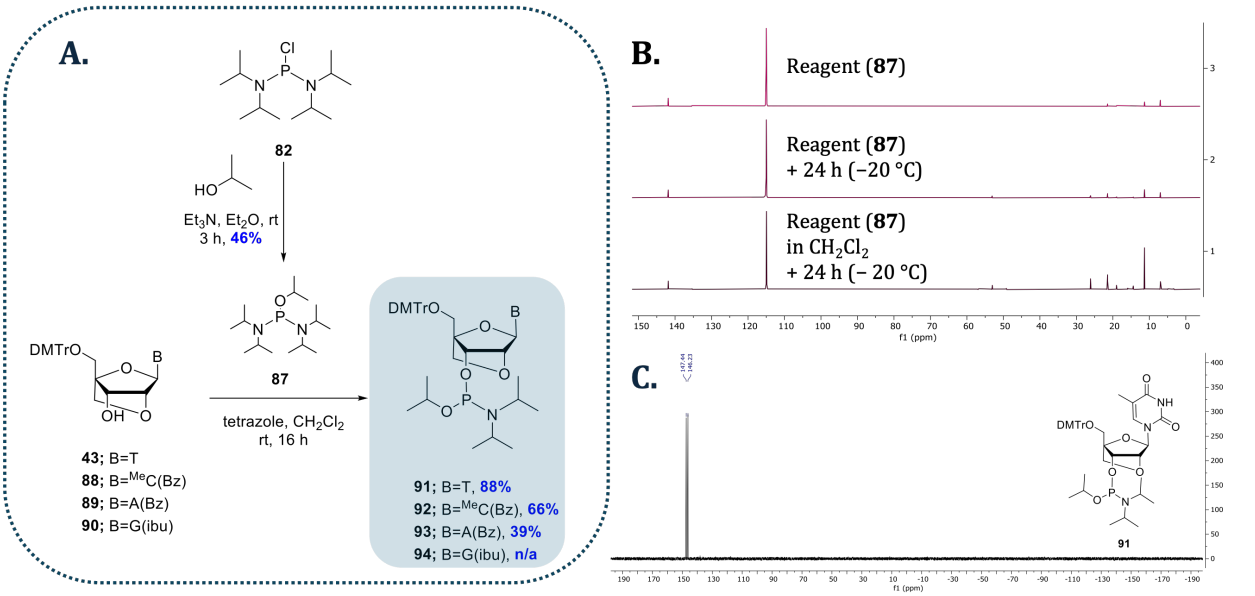


Figure 4.4: **A.** Synthetic scheme of isopropyl phosphoramidite reagent **87** and synthesis of desired locked isopropyl triester phosphoramidite monomers **90-94**. **B.** ^{31}P NMR of reagent **87** immediately after inert distillation, after 24 h at −20 °C, and after 24 h at −20 °C in anhydrous, degassed CH_2Cl_2. **C.** ^{31}P NMR of the thymidine isopropyl triester phosphoramidite monomer **91** prior to use in solid-phase oligonucleotide synthesis.

Therefore, the synthesis of the bis(diisopropylamino)isopropoxy-phosphine **87** was envisioned as a phosphitylating reagent which already carried the isopropyl triester moiety (Figure 4.4A). **87** was synthesised by nucleophilic substitution of bis(diisopropylamino)chlorophosphine **82** with isopropyl alcohol and subsequently purified by distillation. This purification step is crucial, as using the reagent crude in a one-step reaction as previously reported for DNA monomers[90] gave reduced phosphitylation yields of the locked monomers. Unfortunately, the reagent was highly unstable to oxygen, as oxidation began to occur just 24 h post-distillation, worsened by storage in anhydrous, degassed CH_2Cl_2. Degradation of the reagent (shown by ^{31}P NMR in Figure 4.4B) meant the phosphine reagent **87** was synthesised and freshly distilled prior to every monomer synthesis. Phosphitylation of commercially available or synthesised locked monomers **43**[a], **88**, **89**, **90**[b] with the phosphine reagent **87** gave the triester phosphoramidites **91-93**. These too required purification by inert column chromatography in order to ensure only one P(III) species was reacted with the oligonucleotide

[a]synthesised in Scheme 3.1
[b]Purchased from Hongene Biotechnologies

synthesiser (see Figure 4.4C for the ^{31}P NMR of purified locked thymine phosphoramidite **91**); although conducting a second inert purification was time-consuming, it enabled a much cleaner oligonucleotide synthesis.

Unfortunately, the oxidation of the locked isopropyl-triester monomers (**91-93**) increased with the difficulty of reactivity of the nucleobases − the locked thymidine phosphoramidite **91** was isolated in the highest yield (88%) while the recovery of the locked adenine(Bz) monomer **93** was low (39%) and no locked guanosine(iBu) monomer **94** was isolated due to degradation during the handling of the oxygen-sensitive product (inert aqueous work-up and column chromatography). In future work, it is proposed to optimise this reaction or re-visit the one-pot step on a larger scale in order to achieve the LNA G(iBu)-OiPr monomer **94**.

4.4 Oligonucleotide synthesis

4.4.1 Synthesis of LNA-OiPr (PO) vs. LNA-OiPr (PS) linkages

ASOs with neutral LNA-isopropyl triester linkages were synthesised as therapeutic 2'OMe/PS chimeras, with LNA flanking both ends of the OiPr neutral linkage. However, the P(III) centre of the envisioned triester linkage can be oxidised or sulfurised during SPOS to give an alkyl-phosphate or alkyl-phosphorothioate linkage.

Oligonucleotide construction of these two variants is described in Figure 4.5. First, a commercially-available locked monomer was coupled to the growing oliognucleotide using extended coupling time (10 min). Standard oligonucleotide steps (capping, sulfurisation, and detritylation) were used to achieve the locked 3'-monomer depicted on the growing oligonucleotide (I in Figure 4.5). Next, the custom 5'-DMTr-protected LNA-OiPr phosphoramidite (**86**) was coupled using extended coupling time (10 min) (step i, Figure 4.5). This produced the P(III) linkage (oligonucleotide II, Figure 4.5). Sequences failing to couple are capped as per standard SPOS to produce capped oligonucleotide (III, Figure 4.5). The successfully-coupled linkage II was oxidised using iodine in THF, pyridine, and water to produce the isopropyl-phosphate linkage IV. Alternatively, the sulfurising reagent 3-ethoxy-

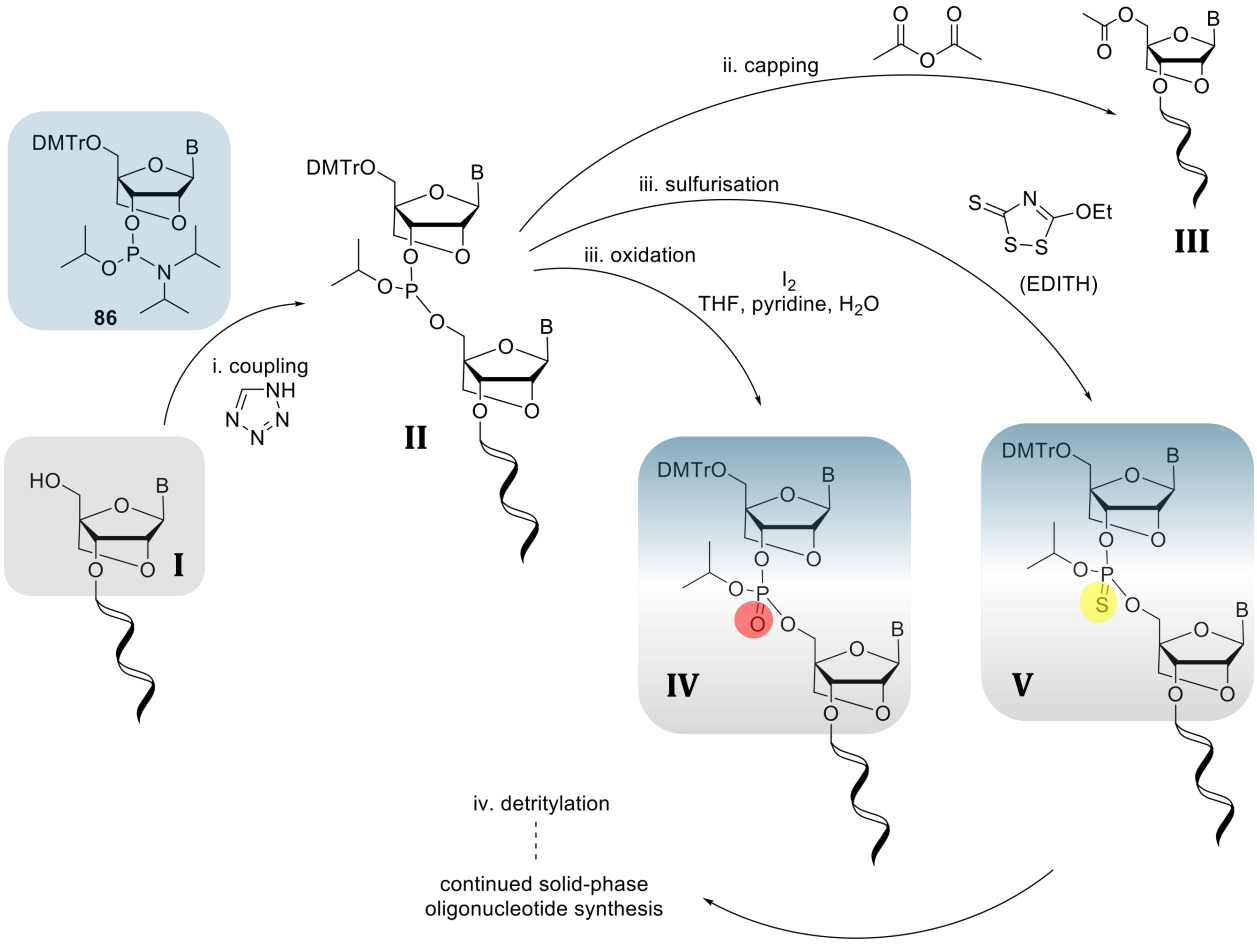


Figure 4.5: The oligonucleotide synthesis cycle required to install an isopropyl-phosphate triester linkage (I) or an isopropyl-phosphorothioate linkage (II) through the coupling of custom 5'-DMTr-LNA-isopropyl triester phosphoramidite **86**.

1,2,4-dithiazoline-5-one (EDITH) was used to produce the isopropyl-phosphorothiate linkage V. The remaining steps of oligonucleotide synthesis cycle (detritylation) was conducted as normal, and the solid-phase synthesis was continued (Figure 4.5).

4.4.2 Oligonucleotide deprotection

ASOs were synthesised containing incorporations of the LNA-OiPr linkage, using either oxidation or sulfurisation (Table 4.1). In the first instance, the standard deprotection conditions (aqueous ammonium hydroxide at 55 °C for 5 h) were used. While ASO-14 and ASO-15 containing single LNA-OiPr linkages were stable, degradation of the triester group was observed for ASOs with multiple triester incorporations (ASOs 16-18), rendering the oligonucleotides

ASO#	Sequence (5 → 3')	Expected Mass (Da)	Found Mass (Da)	Modifications x = linkage	Deprotection Conditions
ASO-1	CCUCUUACCUCAGUUACA	6096.3	6096.5	2'OMe/PS	
ASO-2	C**CTCTT**AC**CT**CAG**TT**ACA	6192.4	6195.5	2'OMe/PS/LNA	
ASO-14	CCUC**TxT**ACCUCAGUUACA	6147.4	6149.6	x = OiPr (PO)	≤1 LNA-OiPr incorporation Stable to NH_3, 55 °C, 5 h
ASO-15	CCUC**TxT**ACCUCAGUUACA	6163.5	6168.0	x = OiPr (PS)	
ASO-16	CCUC**TxT**ACCUCAG**TxT**ACA	6230.6	6231.0	x = OiPr (PS)	>1 LNA-OiPr incorporations **Not** stable to NH_3, 55 °C, 5 h Requires EDA/THF, rt, 2 h
ASO-17	C**CxTCTxTACCxTCAGTxTACA**	6300.6	6298.5	x = OiPr (PO)	
ASO-18	C**CxTCTxTACCxTCAGTxTACA**	6364.9	6365.0	x = OiPr (PS)	

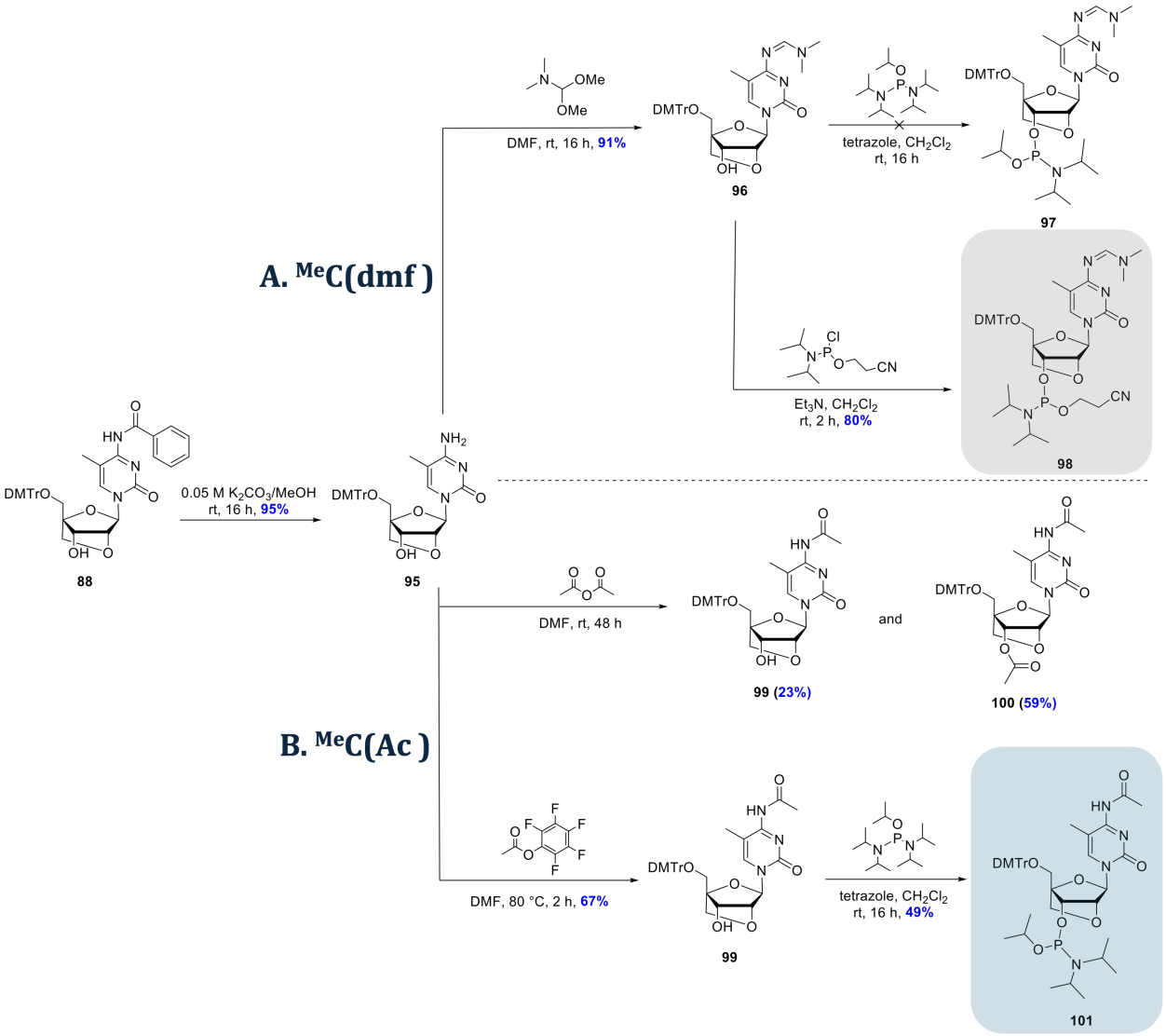
Table 4.1: Table of ASOs synthesised containing LNA-OiPr triester linkages. ASOs in which the linkage was oxidised (PO) are highlighted in blue.

with the standard phosphate or phosphorothioate linkages. Very little oligonucleotide material was recovered for ASO-17 and ASO-18, as these ASOs suffered from the greatest quantity of degradation. This phenomenom is substantiated by the literature, which reports triester linkages, as well as alkyl linkages such as the methoxypropyl phosphonate (MOP) as being susceptible to degradation during standard conditions when incorporated in one or more positions.[90]

Alternative deprotection conditions for alkylphosphonates and alkyl phosphotriesters have been reported which do not degrade the neutral linkage. First, the β-cyanoethyl groups are deprotected by treatment with 20% diethylamine in THF, followed by nucleobase deprotection by treatment with ethylenediamine (EDA) (50% in THF, 15 min at 55 °C).[90] However, a different protecting group than the standard benzoyl is required for the ^{Me}cytosine monomers – ^{Me}C(Bz) suffers from transamination if alkylamines are used during oligonucleotide deprotection[175] (up to 15% transamination at the N4-position of cytosine by EDA was observed by Hogrefe et al.)[176] Therefore, the protecting group strategy for locked ^{Me}C monomers had to be revised to replace the benzoyl protecting group with groups such as acetyl (Ac) or dimethylformamidine (dmf), which were reported as labile in EDA without transamination.[177]

4.5 Synthesis of LNA ^{Me}cytosine nucleosides

Both the standard locked ^{Me}cytosine phosphoramidite and the locked ^{Me}cytosine-OiPr phosphoramidites require synthesis with alkyl-amine-compatible protecting groups. Where cytosine is placed at other points of the oligonucleotide (bearing 2'OMe modification), commercially available monomers can be used as they are acetyl-protected. A comprehensive overview of which modified ^{Me}cytosine monomers are commercially available with which N4-protecting groups can be found in Table D.1.



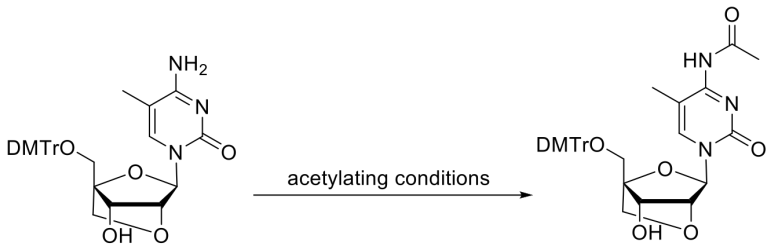
Scheme 4.2: Synthesis of locked LNA-^{Me}C nucleosides and phosphoramidites with EDA-compatible protecting groups: LNA-^{Me}C(dmf) monomers (**A.**) and LNA-^{Me}C(Ac) monomers (**B.**).

First, the commercially available locked ^{Me}C(Bz) monomer **43** was deprotected using potassium carbonate in methanol to give the monomer **95** in near-quantitative yields (95%) (Scheme 4.2). Dimethylformamidine (dmf) was the first protecting group considered, as it is selective for the N4-amino position of cytosine and would not protect the 3'-alcohol moiety.[178,179] Protection with *N,N*-dimethylformamide dimethylacetal (DMFDMA) proceeded in very good yields (91%) to give the locked ^{Me}C(dmf) monomer **96** (Scheme 4.2A). Unfortunately, phosphitylation of this monomer using the custom chloro-diisopropylamino-isopropylphosphine reagent **87** was unsuccessful; this was not due to a problem with reactivity, but rather in separating the excess reagent from the phosphoramidite product which had the same retention factor. Instead, the locked ^{Me}C(dmf) monomer **96** was reacted with standard phosphitylating reagent chloro(diisopropylamino)-β-cyanoethoxyphosphine in 80% yield to give phosphoramidite **96**. This monomer was utilised when installing an LNA monomer at the 3'-end of the LNA-LNA triester linkage (e.g. when the linkage is envisioned as 5'-**Bx^{Me}C**-3').

The commonly-used acetyl (Ac) protecting group was considered as an alternative (Scheme 4.2B).[180] Unfortunately, the use of acetic anhydride resulted in bis-acetylation of the monomer **95**, acetylating both the free N4-amine and 3'-OH positions; only 23% of desired product **99** was recovered while 59% was the bis-side-product **100** (Scheme 4.2B). Various acetylation conditions and reagents were tested in order to reduce bis-side-product formation and increase selectivity, summarised in Table 4.2.

Drop-wise addition of acetic anhydride was only able to achieve 35% yield of the desired monomer (entry 2). Temporary protection of the 3'-hydroxyl group using the trimethylsilyl (TMS) protecting group was also attempted. 3'-TMS protection was conducted using TMSCl, followed by acetylation with acetic anhydride and aqueous workup. Silyl deprotection using tetrabutylammonium fluoride (TBAF) improved this yield to 59% over two steps (entry 3). However, this was a lengthier route which still suffered from some bis-product formation due to acetic anhydride causing some TMS-deprotection. A one-pot reaction using BSA as the silyl protecting group source and acetyl chloride as the acetylating reagent yielded little desired product (24%, entry 4). Phenol acetate has shown to be selective for primary

A.



95 99

B.

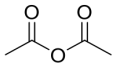
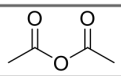
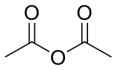
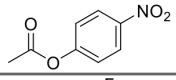
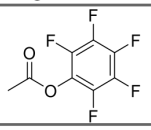
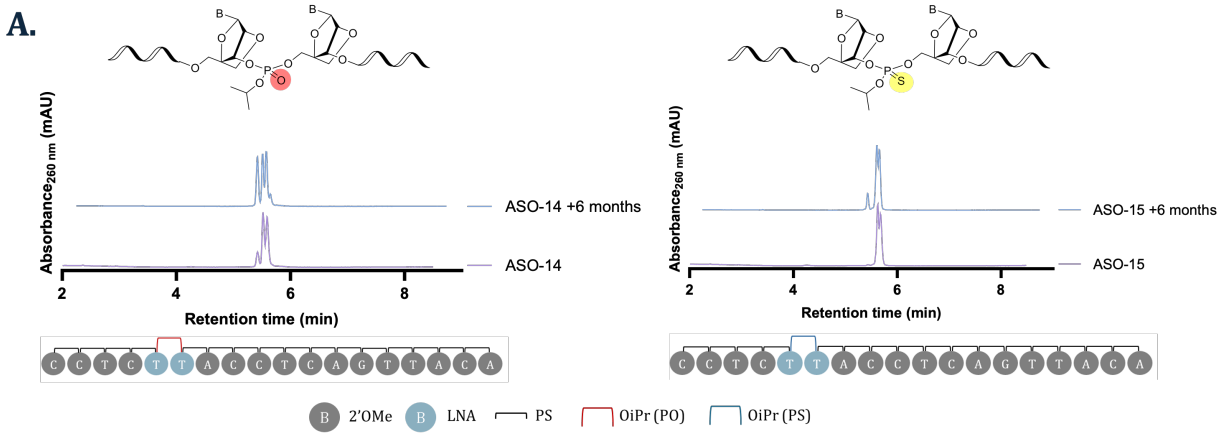
#	Acetylating Reagent	Reaction Conditions	3'-OH Protecting Group	Yield (%)
1		DMF, 3 eq. over 48 h, rt	n/a	23
2		DMF, 1.4 eq. drop-wise, rt, 24 h	n/a	35
3		i. TMSCl, pyridine, rt, 3 h, then Ac_2O, rt, 2 h ii. TBAF, THF, rt, 5 h	TMS	59
4		BSA, ACN, 80 °C, 5 h, then AcCl, Et_3N, rt, 16 h, then TBAF, rt, 1 h	TMS	24
5		Et_3N, DMF, rt to 50 °C, 24 h	n/a	53
6		DMF, rt to 80 °C, 2 h	n/a	**67**

Table 4.2: A. Scheme of acetyl-protection at the N4-position of locked ^{Me}cytosine monomer **95** to yield **99**. **B.** Table of conditions attempted to achieve locked ^{Me}cytosine(Ac) monomer **99**.

amines over alcohols;[181] the addition of electron-withdrawing groups such as a para-nitro group increases its reactivity. The use of 4-nitrophenyl acetate as the acetylating reagent increased the yield to 53% (Entry 5). The use of pentafluorophenyl acetate was most successful, with 67% yield for the desired mono-acetylated product **99** (Entry 6, Table 4.2). This was subsequently reacted with the custom isopropyl phosphine **87** in 49% yield to give locked ^{Me}C(Ac) isopropyl phosphoramidite **101** (Scheme 4.2B). Going forward, additional ASOs incorporating the LNA-OiPr linkages were synthesised using the custom ^{Me}C(dmf) LNA phosphoramidite **98** and custom ^{Me}C(Ac) LNA-OiPr phosphoramidite **101**.

4.6 Properties of ASOs containing LNA-OiPr (PO) vs. LNA-OiPr (PS) linkages



B.

OiPr Linkage	RNA Complement		
	T_m (°C)	$\Delta T_{m\ (2'OMe)}$	$\Delta T_{m\ (LNA)}$
ASO-1: 2'OMe/PS control	41.0 ±0.1	-	+27.5
ASO-2: LNA control	68.5 ±0.2	+27.5	-
ASO-17: 4 LNA-OiPr (PO)	65.1 ±0.5	+24.1	−3.4
ASO-18: 4 LNA-OiPr (PS)	65.2 ±0.6	+24.2	−3.3

Figure 4.6: Comparison of ASOs containing LNA-OiPr (PO) vs. LNA-OiPr (PS) linkages: **A.** HPLC chromatograms of purified ASO-14 (one LNA-OiPr (PO) linkage) vs. ASO-15 (one LNA-OiPr (PS) linkage), before and after 6 months storage in water at 4 °C. **B.** Melting temperatures of ASOs containing 4 LNA-OiPr linkages against the complementary RNA fragment (25 mM NaCl, 10 mM phosphate, pH = 7.0).

ASOs containing LNA-OiPr (PO) and LNA-OiPr (PS) linkages were compared in order to evaluate the effect of oxidising or sulfurising the isopropyl triester linkage. Preliminary biophysical and biological studies were conducted. The LNA-OiPr (PS) linkage was deemed more therapeutically promising, and thus taken forward, due to three factors:

1. **Degradation:** The stability of ASO-14 and ASO-15 were compared in Figure 4.6A, and the LNA-OiPr (PO) linkage was more susceptible to degradation. Both purified oligonucleotides were stored in water at 4 °C over 6 months. Both ASOs observe a growing quantity of degraded oligonucleotide via HPLC; the left-most peak with the lowest retention time corresponded to the mass of the degraded product (P-OiPr was cleaved to the P-O^-). The increase in this peak area was significant for ASO-14 (and

exists even at the point after purification, from 13% to 31% degraded product), while only 11% of degraded oligonucleotide was apparent after 6 months for ASO-15.

2. **Thermal stability:** The ASO containing four LNA-OiPr (PO) linkages (ASO-17) displayed a similar duplex thermal stability as the ASO containing four LNA-OiPr (PS) linkages (ASO-18) (Figure 4.6B). The reduction in phosphorothioate content for ASO-18 had no effect on duplex stability.[67]

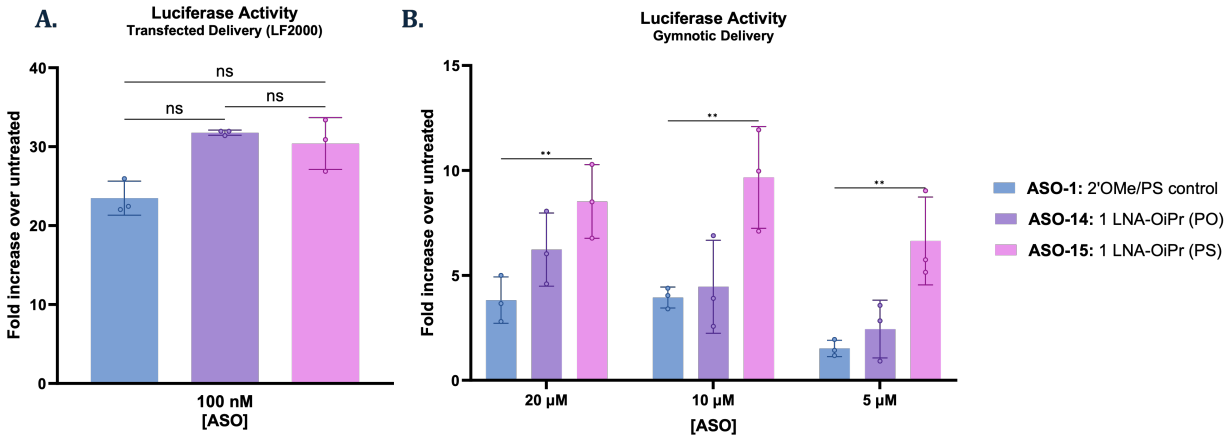


Figure 4.7: Splice-switching activity of ASOs containing a single LNA-OiPr (PS) and LNA-OiPr (PO) linkage, delivered by transfection with Lipofectamine 2000 (100 nM) (**A.**) or gymnotic delivery (**B.**). Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 1-way ANOVA test (**A.**) or 2-way ANOVA test (**B.**) using the 2'OMe/PS control ASO as the control within each concentration. **represents $P < 0.01$. Data are represented as mean ±SD; each dot represents a technical replicate ($n = 3$).

3. **Splice-switching activity:** A preliminary biological study of ASO-14 and ASO-15 was conducted using the luciferase assay (Figure 4.7). When using Lipofectamine 2000 at the highest dose (100 nM) (Figure 4.7A), the ASOs containing LNA-OiPr linkages (regardless of oxidation or sulfurisation of the linkage) were similarly active to the 2'OMe/PS control (ASO-1). However, when delivering the ASOs via gymnosis, the oligonucleotide with a full PS backbone had significantly more activity over the 2'OMe/PS control, while the ASO with a singular PO linkage did not (Figure 4.7B). Phosphorothioates are well-known in the literature to increase productive cellular uptake through protein interactions, but it was interesting to observe a single atom change could impart a gymnotic activity difference.[182] However, only a single biological repli-

cate was conducted – in future, additional replicates should be conducted to further explore this effect.

Due to the differences in stability and biological activity between LNA-OiPr (PO) and LNA-OiPr (PS) linkages, ASOs with the LNA-OiPr (PS) linkages were prioritised. ASOs incorporating one, two, and four linkages were re-synthesised (using the custom ^{Me}C(Ac) LNA-OiPr phosphoramidite **101** and custom ^{Me}C(dmf) LNA phosphoramidite **98**) (Table 4.1). The mild deprotection conditions (50% EDA/THF, 2 h, rt) were used, and the ASOs were purified, de-tritylated, and the salt was exchanged for the Na^+, as described in Chapter 7.

4.7 Biophysical properties of ASOs containing LNA-OiPr linkages

A.

OiPr Linkage	DNA Complement		RNA Complement						
	200 mM		25 mM (full)		200 mM (fragment)		25 mM (fragment)		
	T_m (°C)	$\Delta T_{m\ (2'OMe)}$	T_m (°C)	$\Delta T_{m\ (2'OMe)}$	T_m (°C)	$\Delta T_{m\ (2'OMe)}$	T_m (°C)	$\Delta T_{m\ (2'OMe)}$	$\Delta T_{m\ (LNA)}$
ASO-1: 2'OMe/PS control	54.8 ±0.1	-	62.1 ±0.1	-	54.3 ±0.2	-	41.0 ±0.1	−27.5	-
ASO-15: 1 LNA-OiPr	61.2 ±0.3	+6.4	68.6 ±0.1	+6.5	-	-	-	-	-
ASO-16: 2 LNA-OiPrs	68.3 ±0.2	+13.5	74.8 ±0.1	+12.7	64.9 ±0.1	+10.6	51.5 ±0.2	+10.5	−17
ASO-18: 4 LNA-OiPrs	>75	-	>75	-	>75	-	64.3 ±0.5	+23.3	−4.2
ASO-2: LNA control	>75	-	>75	-	>75	-	68.5 ±0.2	+27.5	-

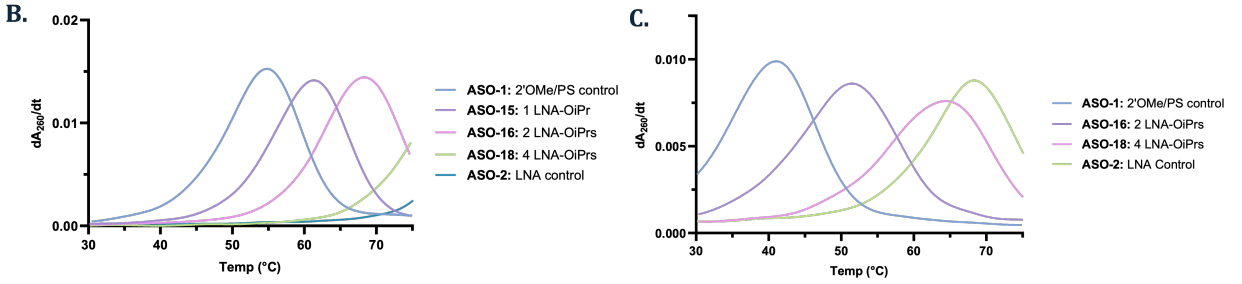


Figure 4.8: A. Table of melting temperatures for ASOs containing LNA-OiPr (PS) backbones. Representative first derivative curves of duplex melting for ASOs containing LNA-OiPr (PS) backbones against complementary DNA (200 mM NaCl, 10 mM phosphate, pH = 7.0) (**B.**) and against a complementary RNA fragment (25 mM NaCl, 10 mM phosphate, pH = 7.0) (**C.**).

Previously, the isopropyl triester linkage was reported as having a small destabilising effect on the thermal duplex stability of an ASO – Ionis Pharmaceuticals report a range of −2.2 to −0.6 °C for a single OiPr linkage in a gapmer (depending on position in the gap). However,

combining the isopropyl linkage with the locked nucleic acid modification rendered the ASO stabilising both against the DNA and the RNA complement. A single LNA-OiPr linkage (ASO-15) increased the ASO:DNA heteroduplex melting temperature by 6.4 °C; this effect was additive, as the ΔT_{m} of the ASO containing two LNA-OiPr linkages (ASO-16) = 13.4 °C (Figure 4.8A). The linkage was stabilising to an even greater extent by an ASO:RNA heteroduplex; the ΔT_{m} of ASO-15, containing one incorporation of LNA-OiPr linkage, against a fragment RNA complement = +10.6 °C. When using the full RNA complement (at reduced salt concentrations (25 mM NaCl)), this effect was reduced and only a ΔT_{m} of ∼6.5 per incorporation °C was observed. This is approximately the stabilisation range one would expect for a single LNA, which indicates that the isopropyl triester linkage had little to no conflicting or destabilising effects.

Finally, the fragment and reduced salt approaches were combined in order to reduce the melting temperature of ASO-18, containing four LNA-OiPr linkages, which was too stable to melt at measurable temperatures using either of these approaches alone. By comparing ASO-18 to ASO-2, the LNA-control which contains eight LNA moieties in the same positions, the effects of the OiPr linkage can be tentatively separated from the stabilising effects of the LNA. The ΔT_{m} of ASO-18 was only −4.2 °C compared to ASO-2 − the OiPr linkage had a destabilisation effect compared to LNA of around −2.1 °C per modification (in the fragment approach, the complementary RNA fragment only spans two out of the four LNA-OiPr linkages). The LNA-OiPr is much more stable than the LNA-amide, LNA-carbamate, and even more stabilising than the LNA-sulfamate, investigated in Chapters 2 and 3.

Therefore, the "neutrality" of the linkage of a given linkage does not automatically improve duplex stability by reduction of electrostatic repulsion. Potentially, the LNA-triester was the most stable LNA-neutral linkage, because it perturbed the effects of the stabilising LNA the least. Compared to the most stable LNA-PS-LNA linkage, the LNA-triester linkages maintain the bridging bonds' lengths and flexibilities. Inclusion of the alkyl isopropyl group may not be bulky enough to interfere with the conformational rigidity of the LNA moiety in the 5'-side of the linkage. It would be interesting to synthesise ASOs with larger R-groups,

both as chains or branched hydrophobic groups, to investigate the size at which the sterics clash with the LNA. Investigations of other R-groups is in progress in our group, as SAR with alkyl triester R groups is thus far limited to DNA monomers by the literature.

The CD spectra of ASOs containing LNA-OiPr linkages were measured to further investigate the differences in global duplex conformation compared to ASOs with 2'OMe/PS and LNA/PS linkages. The CD spectra of ASO:DNA duplexes (Figure 4.9A) revealed that the increased quantity of LNA induced a B→A-form shift, as was the case with the LNA-amide and is consistent with the literature.[138] The appearance of a second additional maxima (shoulder-like shift from 270 nm to 260 nm) in the CD of ASO-2 and ASO-18 indicated some additional A-form characteristics.

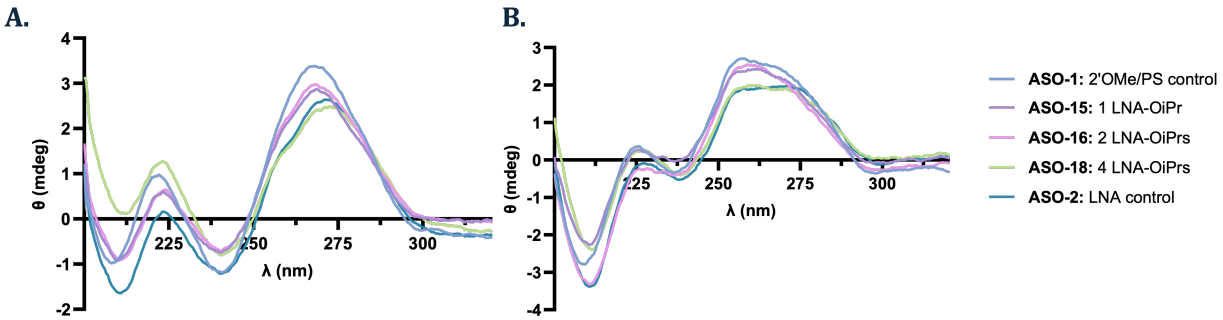


Figure 4.9: Circular dichroism spectra of ASOs containing LNA-OiPr linkages against DNA complement (**A.**) (200 mM NaCl, 10 mM phosphate, pH = 7.0) and RNA complement (**B.**) (25 mM NaCl, 10 mM phosphate, pH = 7.0).

As expected, the CD of the ASO:RNA heteroduplexes revealed mixed A/B form conformation (Figure 4.9B). Shifts from A/B mixed conformations toward A-form are characterised by blue-shifted maxima toward 260 nm and intense minima at 210 nm.[139] The 2'OMe/PS control (ASO-1), as well as the ASOs with one or two incorporations of LNA-OiPr linkage (ASO-15 and ASO-16) revealed A-like characteristics, displaying maxima at 260 nm and intense minima at 210 nm. However, as LNA content increased, a plateau-like maxima with a shift in the bathochromic direction (260 nm toward 275 nm), also observed with the LNA-amide and LNA-sulfamate, was observed for the ASOs with four LNA-OiPr linkages (ASO-18) and the LNA-control (ASO-2) (containing 8 LNA moieties each). As hypothesised previously, some A→B-form character or structural perturbation may be caused by the LNA

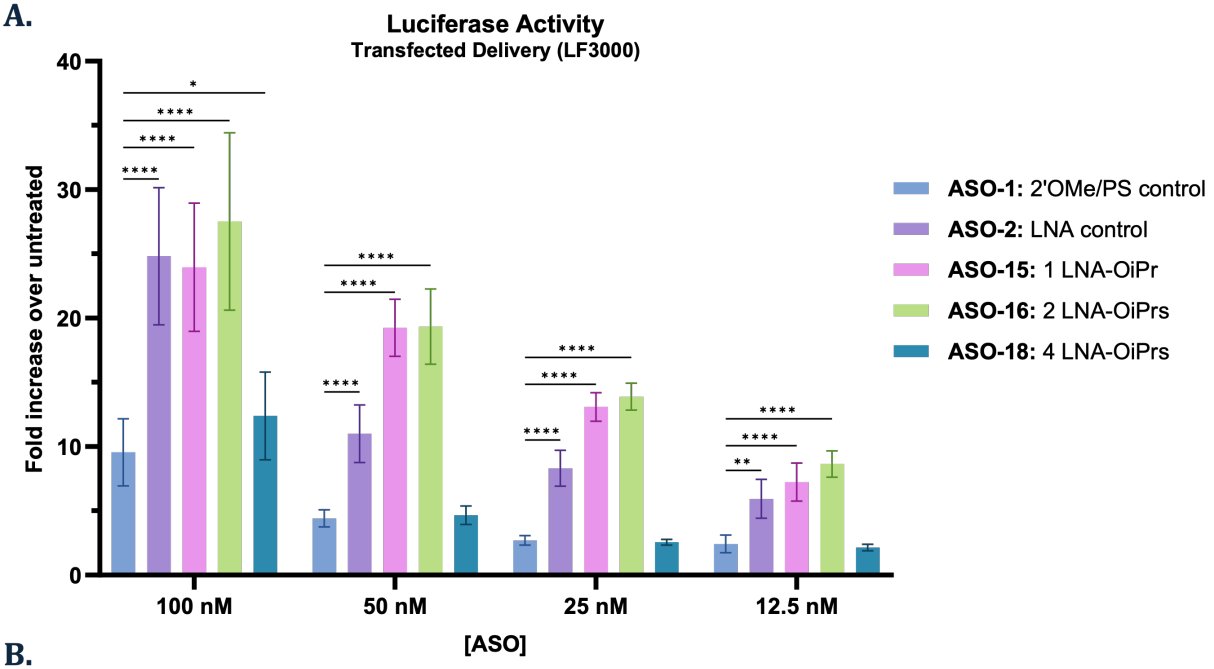
in the 5'-position, as shown by crystal structure.[62]

It would be interesting to apply computational techniques to the LNA-OiPr linkage in future work − the MD simulations (applied to the LNA-neutral linkages in Chapters 2 and 3) could provide additional information on the structural perturbation, if any, an alkyl triester imparts on a native phosphodiester linkage. Potentially, MD simulations could additionally predict the effects of various R groups on the conformation and flexibility of the duplex using RMSD, 3'-5' dinucleotide distances, and H-bonding duration.

4.8 Biological activity of ASOs containing LNA-OiPr linkages

Finally, the ASOs containing LNA-OiPr (PS) linkages were evaluated in the luciferase splice-switching assay, also used in Chapters 2 and 3. Treatment of the HeLa pLuc/705 cells by transfection with Lipofectamine 3000 demonstrated that all ASOs containing LNA-OiPr linkages had equal, if not increased, activity compared to ASO-1, containing the clinically-validated 2'OMe/PS modifications (Figure 4.10A). It was interesting to note that only ASOs containing one and two incorporations of the LNA-OiPr linkage (ASO-15 and ASO-16, respectively), had increased transfected activity compared to ASO-1; they displayed a range of fold increase over ASO-1 from 2.5-fold to 5.1-fold (Figure 4.10B). However, ASO-18, containing four LNA-OiPr linkages, was approximately equal in splice-switching activity to ASO-1 (0.9 to 1.3-fold increase). Similar to Chapter 2, the LNA control (ASO-2) was inherently very active when delivered via transfection agent (2.4 to 3.1-fold increase). Given that the only difference in modification between ASO-2 and ASO-18 is four alkyl triesters (both ASOs contain eight LNA moieties), potentially numerous alkyl triester groups (beyond two incorporations) was the cause for this reduced transfected activity. A plausible hypothesis is that these ASOs with increased quantity of hydrophobic groups are binding to proteins which sequester them in intracellular compartments in non-productive uptake pathways.[183] A comprehensive characterisation of the protein-binding profiles of ASOs containing LNA-OiPr vs.

A.



B.

Fold increase over **ASO-1**: 2'OMe/PS control	**100 nM**	**50 nM**	**25 nM**	**12.5 nM**
ASO-2: LNA control	2.6 ±0.9	2.5 ±0.6	3.1 ±0.7	2.4 ±0.9
ASO-15: 1 LNA-OiPrs	2.5 ±0.9	4.4 ±0.8	4.9 ±0.8	3.0 ±1.0
ASO-16: 2 LNA-OiPrs	2.9 ±1.1	4.4 ±0.9	5.1 ±0.8	3.6 ±1.1
ASO-18: 4 LNA-OiPrs	1.3 ±0.5	1.1 ±0.2	0.9 ±0.2	0.9 ±0.3

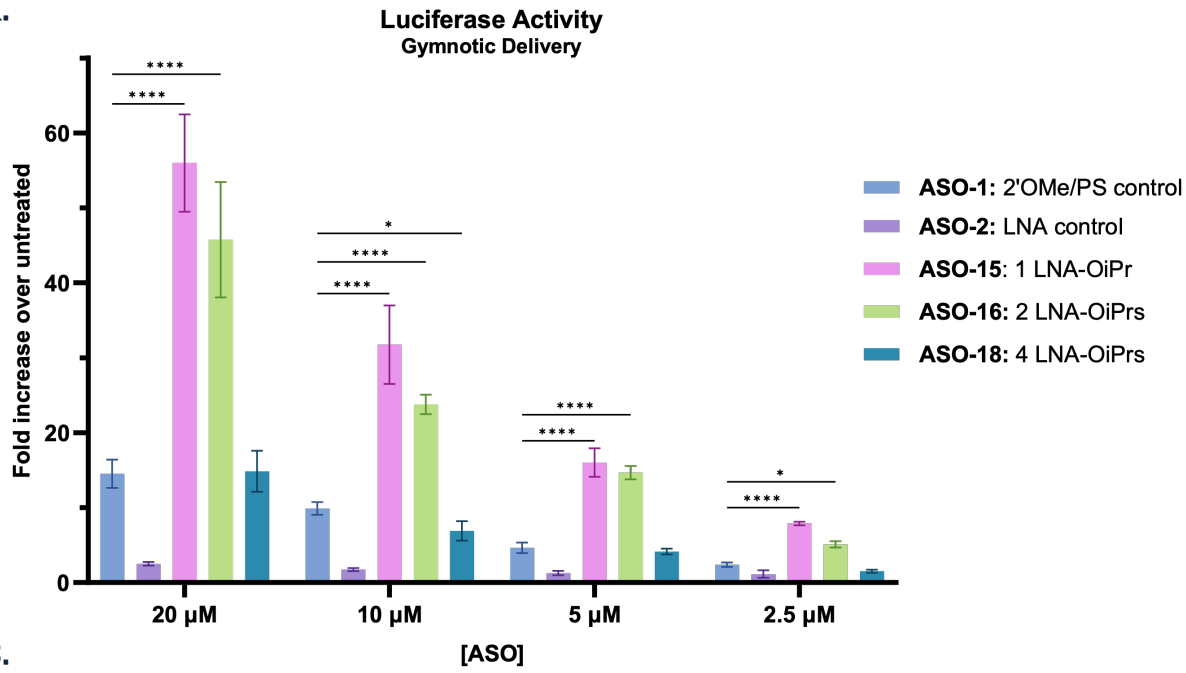
Figure 4.10: Splice-switching activity of ASOs containing LNA-OiPr linkages using the luciferase reporter assay, delivered via transfection with Lipofectamine 3000. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS control as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

LNA-PS linkages could begin to investigate this effect, or, alternatively, cellular trafficking experiments using fluorescently-labelled ASOs.

Consistent with the results in Chapter 2 when investigating the LNA-amide, the LNA control (ASO-2) was much less active when delivered to the cells via gymnosis (Figure 4.11A); its activity was reduced compared to the 2'OMe/PS control ASO (0.2 to 0.5-fold increase over ASO-1). This was hypothesised to be due to increased cytotoxicity or a difference in protein binding properties which are involved in productive cellular uptake. For ASOs containing

LNA-OiPr linkages, the activity trends seen in the transfected assay were continued when the assay was conducted with gymnotic delivery. ASOs containing one or two LNA-OiPr linkages (ASO-15 and ASO-16) were significantly more active at all treatment doses. ASO-15 was 3.2 to 3.9-fold more active than ASO-1, while ASO-16 was 2.1 to 3.2-fold more active. It was promising to observe a single (or indeed, merely two) alkyl triester linkage(s) could so greatly impact the splice-switching activity of an ASO. However, large effects in activity by small or limited chemical modifications are not unusual – Ionis Pharmaceuticals report an IC_{50} of 0.09 µM for a gapmer containing a single incorporation of a DNA OiPr linkage

A.



B.

Fold increase over **ASO-1**: 2'OMe/PS control	20 µM	10 µM	5 µM	2.5 µM
ASO-2: LNA control	0.2 ±0.0	0.2 ±0.0	0.3 ±0.1	0.5 ±0.2
ASO-15: 1 LNA-OiPrs	3.9 ±0.7	3.2 ±0.6	3.5 ±0.7	3.3 ±0.4
ASO-16: 2 LNA-OiPrs	3.1 ±0.7	2.4 ±0.2	3.2 ±0.5	2.1 ±0.3
ASO-18: 4 LNA-OiPrs	1.0 ±0.2	0.7 ±0.1	0.9 ±0.2	0.6 ±0.1

Figure 4.11: Splice-switching activity of ASOs containing LNA-OiPr linkages using the luciferase reporter assay, delivered via gymnosis. Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS control as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

compared to 0.17 µM for a fully PS-modified gapmer.[90]

Interestingly, ASO-18, containing four LNA-OiPr linkages, maintained a similar gymnotic activity to the 2'OMe/PS control (1.0-fold increase at the highest dose (20 µM)); its activity was only slightly reduced (0.6 to 1.0-fold increase) when delivered via gymnosis vs. transfection (0.9 to 1.3-fold increase). The activity of the LNA-control (ASO-2) was significantly reduced when delivered gymnotically. Potentially, the presence of the isopropyl alkyl groups was "rescuing" the gymnotic activity of ASO-18, which contains the same quantity of LNA moieties as ASO-2 through altering the protein-binding profile of the LNA-modified ASO − for example, the triester group may reduce binding to non-productive or cytotoxicity-related proteins, such as paraspeckles.[101]

As with the LNA-amide, trends which are seen in the transfected assay (e.g. 1 incorporation is more active compared to the positive control ASO than 4 incorporations) can be seen again in the gymnotic assay. This was promising, as the gymnotic assay required large quantities of material − ASOs may be reasonably evaluated for further investigation using transfection. However, each novel LNA-neutral linkage must be evaluated independently, as the ASO containing four LNA-amides was much more active than the corresponding LNA-OiPr ASO with four incorporations (in the same positions). Therefore, no universal "rule" has yet to be established for the optimal incorporation number (nor position) of LNA-neutral linkages. Nevertheless, the incorporation of one or two LNA-OiPr linkages demonstrated ∼3-fold increase in gymnotic activity over a 2'OMe/PS control, encouraging further investigation of this linkage.

4.9 Conclusions

A final LNA-neutral linkage was investigated in this chapter, the "LNA-OiPr", in which the remaining non-bridging oxygen of a phosphorothioate (or phosphodiester) linkage was functionalised with an alkyl isopropyl group. Novel locked isopropyl triester phosphoramidites were synthesised and incorporated into ASOs. Alternative deprotection conditions for ASOs

containing >1 LNA-OiPr linkage were required, which led to necessary protecting group manipulation for the locked ^{Me}cytosine monomers. ^{Me}C(Bz) monomers suffer from transamination during deprotection with alkylamines such as EDA; therefore, novel locked ^{Me}C(Ac) and ^{Me}C(dmf) monomers were synthesised.

Oxidation of the isopropyl triester linkage (LNA-OiPr (PO)) and sulfurisation of the isopropyl triester linkage (LNA-OiPr (PS)) were conducted on ASOs containing one and four of these linkages, respectively. The LNA-OiPr (PS) linkage had some advantages − among them reduced degradation, equal thermal stability, and increased gymnotic activity. Therefore, ASOs containing the LNA-OiPr (PS) (one, two, and four incorporations) were further investigated. The LNA-OiPr linkage rendered an ASO the most stable of all LNA-neutral linkages investigated thus far; the ΔT_m of ASO-16 (containing two incorporations of LNA-OiPr (PS)) = +12.7 °C against the full RNA complement. Consistent with other LNA-neutral linkages investigated thus far, the LNA-OiPr linkage induced increased A-form characteristics in the CD spectra of ASO:DNA heteroduplexes and, as LNA-content on the 5'-end increased, mixed A/B-form topology was observed for ASO:RNA heteroduplexes. Finally, the biological activity of the ASOs were evaluated in the luciferase splice-switching reporter assay. The ASOs containing one or two (but not four) LNA-OiPr linkages were significantly more active than the 2'OMe/PS control during both transfected and gymnotic delivery.

It is clear that the LNA-OiPr linkage has very promising therapeutic properties, as does the LNA-amide. Thus far, the structure of an LNA-neutral linkage has been shown to have large effects on the biophysical and biological properties of an ASO. Each LNA-neutral linkage observed slightly different trends for biological activity for a given position or incorporation-number. However, through thorough SAR evaluation of many more LNA-neutral linkages, in future it may be possible to determine the ideal position, number, conformation, flexibility, etc. of modified neutral backbones for rational design of successful therapeutic ASOs.

5

Further biological properties of ASOs containing LNA-neutral linkages

Contents

5.1 Introduction

In Chapters 2, 3, and 4, ASOs containing LNA-neutral linkages were synthesised and individually evaluated for their splice-switching activity in the luciferase reporter assay. In this final chapter, their splice-switching activities are directly compared and possible structure-activity relationships are briefly discussed. The chapter aims to progress the most promising LNA-neutral linkages into two biological studies:

1. First, evaluate splice-switching in a therapeutically relevant assay, the in vitro Duchenne muscular dystrophy model (H2k *mdx*).

2. Second, preliminarily investigate protein-binding profiles of ASOs with LNA-neutral linkages through a protein-binding assay and proteomics.

5.1.1 Introduction to Duchenne muscular dystrophy (DMD)

Duchenne muscular dystrophy (DMD) is an X-linked recessive genetic disorder caused by mutations in the dystrophin gene and is the most common muscular dystrophy, affecting 1 in 5,000 male births worldwide.[184,185] The dystrophin gene is the largest gene in the human genome − approximately 2.6 million base pairs long and containing 79 exons.[185] Various mutations in the gene, the majority of which are deletions and the minority of which are point mutations or smaller insertions/deletions, cause premature termination of the protein translation and a loss of dystrophin. Becker muscular dystrophy (BMD) is a milder and rarer form of DMD caused by mutations which shorten the reading frame (not terminating it); therefore, patients produce truncated forms of the dystrophin protein and/or reduced expression.[186]

The dystrophin protein plays a crucial role in the dystrophin-glycoprotein complex (DGC) in striated muscle cells − the DGC is present on the cytoplasmic side of the sarcolemma and

links the intracellular cytoskeleton network of proteins to the transmembrane components. As a whole, the complex connects the cytoskeleton to the extracellular matrix, acting as a stabiliser during muscle contraction.[187] Children diagnosed with DMD will have progressive muscular degeneration and atrophy as well as cardiomyopathy (see Figure 5.1 for a schematic representation of healthy muscle vs. dystrophic muscle). The average life-span is 20-40 years of age; patients die prematurely most oftenly from cardiac and/or respiratory failure.[185]

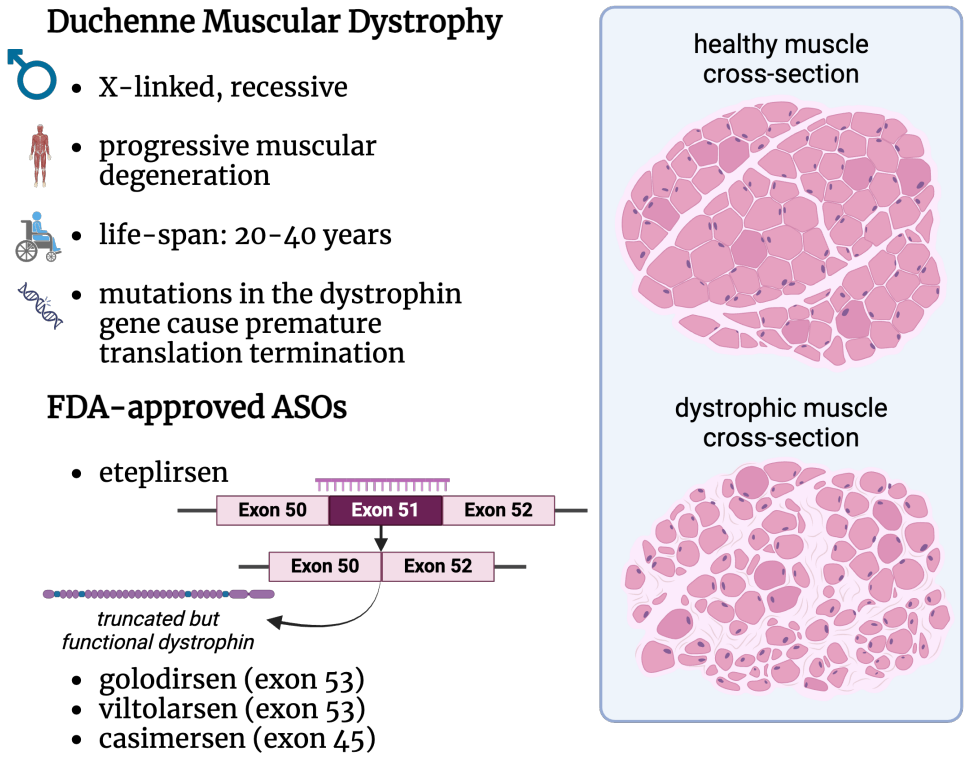


Figure 5.1: An introduction to Duchenne Muscular dystrophy (DMD): key aspects of the genetic disease and FDA-approved ASO-based treatments are listed.

Sporadic cases of DMD (1 in 10,000 gametes) means genetic screening is not a possible way to eliminate it.[188] The size of the dystrophin protein (427 kDa) also make a protein replacement strategy difficult. Viral vector technology to replace the cDNA is in development, but suffers from size constraints. The carrying capacity of AAV-viral gene replacement therapy is ∼4.7 kb, while the cDNA encoding dystrophin is ∼14 kb.[189] In 2023, the FDA approved an AAV for the treatment of DMD which encodes for a shortened version of DMD ("micro-dystrophin", which is 138 kDa in size rather than 427 kDa).[190] However, exon skipping ASOs remain a promising strategy to correct aberrant splicing and provide close to full-length dystrophin,

which is potentially more beneficial. Indeed, an exon skipping ASO (eteplirsen, aka Exondys 51, approved in 2016) provided the first ever treatment for DMD.[191] Eteplirsen is a 30-mer phosphorodiamidate morpholino (PMO) oligonucleotide which hybridises to exon 51, induces exon skipping, and restores the translational reading frame – this causes the production of shortened but functional dystrophin (see Figure 5.1).[191] Since its approval by the FDA in 2017, other ASOs such as golodirsen, viltolarsen, and casimersen have been approved for DMD patients who suffer from mutations in other exons (only 14% of patients have amenable mutations for eteplirsen).[192]

5.1.1.1 The H2k *mdx* murine model

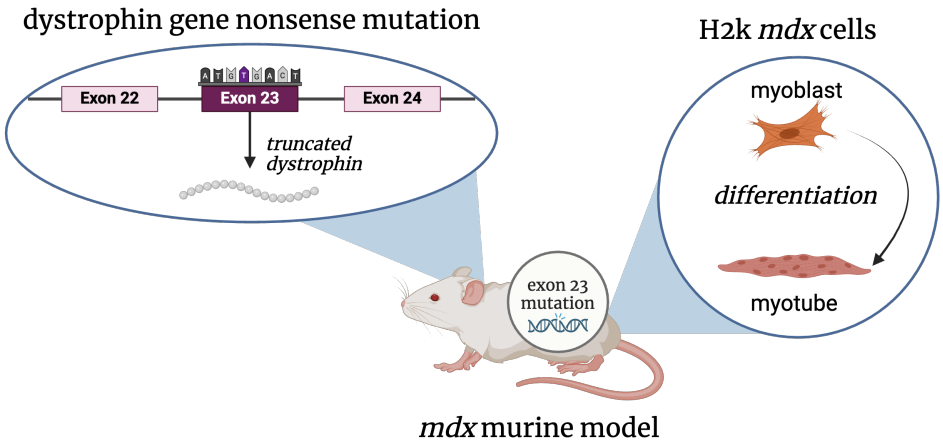


Figure 5.2: The *mdx* murine model carries a C-to-T nonsense mutation in exon 23 of its dystrophin gene, leading to a premature stop codon, nonsense mediated decay, and no production of dystrophin protein. H2k *mdx* cells (dystrophin-deficient myoblasts) are derived from this murine model.

There exist a large number of animal models for DMD (almost 60, including a canine model), but most commonly, the *mdx* murine model is used to evaluate novel chemistries for exon skipping activity.[193] The *mdx* mouse, first developed in 1984,[194] has a nonsense point mutation (C-to-T) in exon 23 which stops full-length dystrophin production and produces truncated, non-functional dystrophin (Figure 5.2). The *mdx* mice have mild clinical symptoms, remain fertile, and their life-span is only reduced by ∼25%.[193] Skeletal myoblasts from a

H-2Kb-tsA58 transgenic mouse[a] were crossed with nude *mdx* mice to result in dystrophin-deficient myoblasts which were immortalised in 1994.[196] These H2k-*mdx* myoblasts are often used in the literature to evaluate exon skipping ASO chemistries in vitro prior to animal work (Figure 5.2).[197] They are used in this chapter to evaluate the effect of the LNA-amide and LNA-triester linkage on exon 23 skipping in 2'OMe/PS chimeric therapeutic ASOs.

5.1.2 Introduction to ASO-protein binding interactions

Development of ASO medicinal chemistry is an important objective, but makes up only one of the areas of focus for ASO research and development.[64,198] As the chemical space has been explored, so too has the structure-activity relationship of various chemistries on an ASO's binding with its target RNA. However, recent work has broadened the focus from solely ASO-RNA target binding to include ASO-protein interactions, despite proteins not being the intended binding target.[b] The phosphorothioate modification, for example, is successful not just due to its enzymatic stability, but its ability to bind many plasma and membrane proteins, which can increase bio-availability and cellular uptake.[199] Stanley Crooke, founder of Ionis Pharmaceuticals, writes "proteins determine the fate of PS-ASOs in all biological systems", truly highlighting the importance of these interactions.[198]

There are various ASO-protein interactions during an ASO's journey from administration to engagement with its target RNA – beginning with plasma proteins (if delivered systemically), membrane-bound proteins, intracellular proteins, and (where relevant) nuclear proteins (Figure 5.3). The main focus of this brief review is on proteins interacting with phosphorothioate-containing ASOs, as all ASOs synthesised herein contain PS linkages. If administered into the systemic circulation (steps 1-2, Figure 5.3), PS-ASOs bind in the blood-stream to plasma proteins, most extensively to albumin (>95% PS-ASOs in circulation are protein-bound). These ASO-protein interactions are vital for the biodistribution

[a]These mice express the thermolabile large tumour antigen (TAg) (encoded by the SV40 tsA58 mutant strain), capable of conditional immortalisation. Production is directed by the mouse antigen promoter H-2K, which is inducible by exposure to interferon-γ. Therefore, myoblasts derived from these mice are able to be conditionally immortalised based on temperature and exposure to interferon-γ.[195]

[b]Protein interactions for RNaseH1 and Argonaute are well-studied, but the scope is being expanded to include proteins not directly involved in ASO mechanisms of action.

to tissues and delaying renal clearance.[94,200] Meaningful interactions with plasma proteins require 10-12 phosphorothioate linkages, and it has been observed that sugar modifications and ionic interactions can play a role in total plasma protein binding.[201]

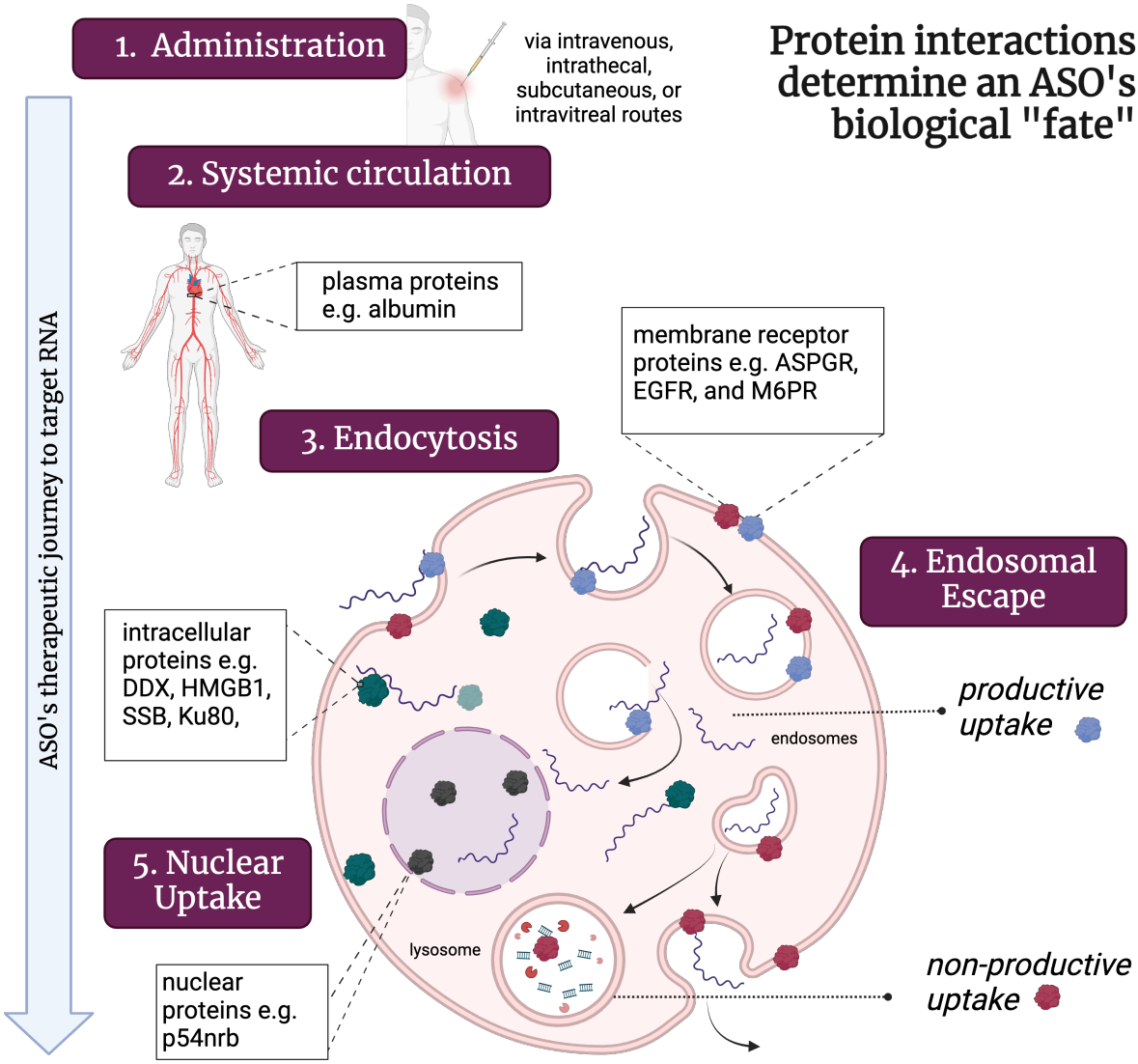


Figure 5.3: Schematic pathway of an ASO's biological journey, from administration to its target RNA. ASO-protein interactions are vital to bio-distribution and cellular trafficking. Some proteins are membrane-bound or intracellular chaperones. Some may facilitate productive or non-productive uptake.

The vast majority of ASOs enter the cell via endocytosis, first beginning with adsorption to the cell membrane by association with specific membrane proteins and then internalisation via an endocytic pathway (these pathways can be clathrin- or calveolin- dependent, or other pathways, depending on the surface protein) (step 3, Figure 5.3). They are trafficked through

an endocytic pathway from early endosome (EE) to late endosomes (LE), but in order to reach target RNA or mRNA, they must be trafficked out of the endosome (termed endosomal escape, step 4, Figure 5.3) to reach the cytosol or nucleus.[202] Finally, if the target is pre-mRNA in the nucleus, nuclear uptake is the final step before the ASO can hybridise and exert pharmacological effect (step 5, Figure 5.3).

Proteins can be deemed "productive" or "non-productive" depending on whether they enhance an ASO's pharmacological effect and help or hinder its trafficking from initial injection to target RNA.[65] For example, the mannose-6-phosphate receptor (M6PR) is a ubiquitously expressed transmembrane glycoprotein which transports enzymes to the lysosome, but has also been shown to facilitate ASO release from the late endosome.[203] Once in the cytosol or nucleus, intracellular proteins are very important to an ASO's activity. It is not surprising that proteins with a nucleic acid binding domain also bind ASOs; they may reduce an ASO's activity (e.g. Ku70 and Ku80 inhibit gapmer activity due to competition for RNAseH binding) or enhance activity (e.g. SSB and NPM1 affect subcellular distribution and nuclear accumulation).[66]

Some PS-ASOs are inherently more toxic than others − the mechanism behind this toxicity is partially elucidated by the study of ASO-protein interactions. It is known that PS-ASOs form ASO-protein aggregates once released from membraned organelles. One type of these aggregates is a "PS body"; first discovered in 1998 in the nucleus,[204] PS bodies are comprised of TCP1β protein subunits and PS-ASOs.[205] Some nuclear aggregates have been found to be involved in cytotoxicity, most notably paraspeckles. Paraspeckles are composed of ribonucleoproteins such as P54nrb and PTB-associated splicing factor (PSF), which form around a scaffold of specific long, non-coding RNA (NEAT1). However, it has been observed that PS-ASOs can co-localise with these proteins, mainly P54nrb and PSF, and form their own paraspeckle structures (without NEAT1).[206] Toxic PS-ASOs bind paraspeckle proteins more extensively than non-toxic PS-ASOs − these ASO-protein interactions and co-localisation to the nucleolus is critical to inducing nucleolar toxicity.[101] However, chemical modification to the PS-ASO has been shown to be an effective strategy to alter ASO toxicity as well as

its protein binding profile.

5.1.2.1 Using chemical modifications to alter ASO-protein binding

2'-modifications can alter ASO-protein binding profiles − it has even been shown that some proteins can bind *specifically* to some 2'-modifications but not others. For example, the Hsp90 protein binds to the 5'-LNA or cEt wings of PS-ASO gapmers and this binding enhances their activity, but does not do so to a 5'-MOE wing.[207] Furthermore, Ionis Pharmaceuticals demonstrate that the toxicity of some PS-ASO gapmers can be significantly reduced with the simple inclusion of a 2'-modification. A single 2'OMe nucleotide on the 5'-side of the gap (in position 2) reduced overall protein binding and improved the ASO hepatotoxicity profile.[101] Additionally, an alkyl methyl phosphonate in the same position also changed toxic PS-ASO gapmers' protein binding profiles; nucleolar mislocalisation of P54nrb was absent and in vivo hepatotoxicity was reduced.[88]

Studying ASO-protein binding interactions can be used to further explore chemical modifications − elucidating which structures increase binding to productive proteins and which decrease binding to non-productive proteins or proteins involved in inducing toxicity will aid in rational ASO design. The exciting preliminary work by Ionis Pharmaceuticals demonstrates that it is possible to utilise structural modifications to alter these interactions and thus, produce safer, more effective ASO therapies.

5.2 Aims of the Chapter

This chapter aims to provide a deeper discussion of the biological properties of ASOs containing LNA-neutral linkages synthesised thus far. Further investigation of the most active chemical modifications in a therapeutic exon-skipping model would provide proof-of-concept for their therapeutic relevance. Furthermore, ASO-binding proteins clearly play an extremely important role in an ASO's therapeutic profile; exploring the interactions of proteins with ASOs containing LNA-neutral linkages may give additional insight into the SAR behind these structures. The chapter aims to:

1. Compare the splice-switching activity of the LNA-amide, LNA-carbamate, and LNA-OiPr linkages in the luciferase assay, taking the most active linkages forward into a more therapeutic splice-switching assay.

2. Evaluate ASOs containing the LNA-amide and LNA-OiPr linkage in a therapeutically relevant model (the dystrophic H2k *mdx* cell line) and quantify skipping of exon 23.

3. Qualitatively determine any differences in protein-binding profiles among ASOs through a biotin pull-down assay.

4. Apply mass spectrometry techniques to identify and quantify proteins bound by ASOs containing LNA-neutral linkages.

5.3 Splice-switching activity of ASOs containing LNA-neutral linkages

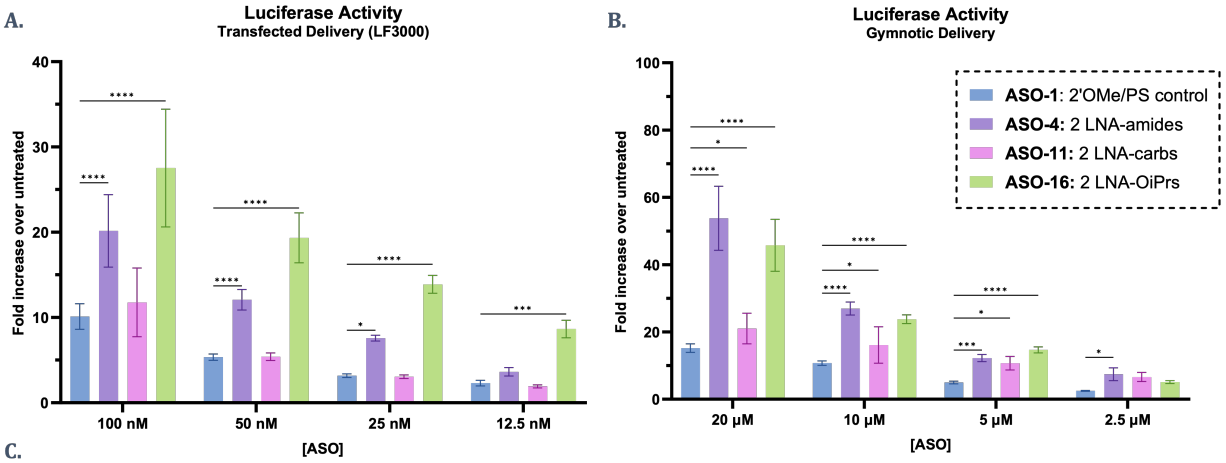
5.3.1 A reporter model – splice switching in the luciferase assay

In Chapters 2, 3, and 4, LNA-neutral linkages were individually investigated for their biophysical properties and biological activities in the luciferase splice-switching assay. ASOs with two incorporations of each LNA-neutral linkage (LNA-amide, LNA-carbamate, and LNA-OiPr) are compared here, one from each chapter. Across all chapters, the ASOs with two incorporations of each LNA-neutral linkage were consistently the most active in comparison to the 2'OMe/PS control (ASO-1), both during transfected and gymnotic delivery.[c] For comparisons of ASOs containing one LNA-neutral linkage each, see Figure E.1, and for comparisons of ASOs containing four LNA-neutral linkages each, see Figure E.2.

Using transfection, ASO-16, containing two LNA-OiPr linkages, had the highest increase in activity compared to the 2'OMe/PS control (2.7 to 4.4-fold increase over ASO-1), a significant difference at all doses. ASO-4, containing two LNA-amide linkages, also displayed

[c]For the LNA-OiPr linkage, both ASOs containing one and two incorporations were similarly active. For sake of comparison, ASO-16 (two incorporations) is represented here.

significantly improved activity at the higher doses, and a range of 1.6 to 2.4-fold increase over ASO-1 across all doses (Figure 5.4A). The LNA-carbamate, the most rigid and destabilising structure, was similarly active compared to the the 2'OMe/PS control at all concentrations (0.8 to 1.2-fold increase). It was interesting to note that ASOs with the highest potency were also those with the highest thermal duplex stability; the T_ms and ΔT_ms of the ASOs containing 2 LNA-neutral linkages are given in Table 5.1 and their thermal stability follows a similar trend to their transfected activity (LNA-OiPr > LNA-amide > LNA-carbamate).



C.

100 nM	50 nM	25 nM	12.5 nM	Fold increase over **ASO-1**: 2'OMe/PS control	20 μM	10 μM	5 μM	2.5 μM
2.0 ±0.5	2.3 ±0.3	2.4 ±0.4	1.6 ±0.3	**ASO-4:** 2 LNA-amides	3.5 ±0.7	2.5 ±0.2	2.5 ±0.3	3.0 ±0.8
1.2 ±0.4	1.0 ±0.1	1.0 ±0.1	0.8 ±0.1	**ASO-11:** 2 LNA-carbamates	1.4 ±0.3	1.5 ±0.5	2.2 ±0.4	2.6 ±0.6
2.7 ±0.8	3.6 ±0.6	4.4 ±0.4	3.7 ±0.7	**ASO-16:** 2 LNA-OiPrs	3.0 ±0.6	2.2 ±0.2	2.9 ±0.3	2.1 ±0.2

Figure 5.4: Dose-response splice-switching activity of ASOs containing two LNA-neutral linkages transfected with Lipofectamine 3000 (**A.**) and delivered via gymnosis (**B.**). Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using ASO-1 as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

When delivered via gymnosis, both ASOs with LNA-amide and LNA-OiPr continued to be significantly more active than the 2'OMe/PS control (Figure 5.4B). The activity of ASO-4 (2 LNA-amides) and ASO-16 (2 LNA-OiPrs) compared to ASO-1 ranged from 2.1 to 3.5-fold more active (across all doses). While the ASO containing LNA-OiPr linkages was the most potent when delivered via transfection, the activities of ASO-4 and ASO-16 are similar when delivered via gymnosis – the fold-increase over ASO-1 are within error at each dose.

ASO#	RNA Complement	
	T_m (°C)	ΔT_m (°C)
ASO-1: 2'OMe/PS control	62.0 ±0.2	–
ASO-4: 2 LNA-amides	57.5 ±0.2	+5.5
ASO-11: 2 LNA-carbamates	61.3 ±0.1	−0.07
ASO-16: 2 LNA-OiPrs	74.8 ±0.1	+12.8

Table 5.1: Melting temperatures of ASOs containing 2 LNA-neutral linkages against the full complementary RNA (25 mM NaCl, 10 mM phosphate, pH = 7.0)

Finally, ASO-11, containing 2 LNA-carbamate linkages, observed some increased activity when delivered via gymnosis; the range in fold increase in activity over ASO-1 increased to 1.4 to 2.6-fold, but remains the least active out of the three LNA-neutral linkages. Additional factors beyond thermal melting temperature must be playing a role in gymnotic activity, as ASO-16 was significantly more stabilising than ASO-4 (ΔT_m = +12.8 °C vs. +5.5 °C) (Table 5.1), yet their gymnotic activity was similar. The trend seen in transfected activity (LNA-OiPr > LNA-amide > LNA-carbamate) no longer applies − splice-switching activity which includes the cell membrane as an obstacle follows the trend LNA-OiPr = LNA-amide > LNA-carbamate. One hypothesis is that protein-binding profiles of an LNA-amide and LNA-OiPr ASO might both have increased binding to productive uptake proteins compared to that of an ASO containing only 2'OMe/PS linkages.

Given the promising activity of the LNA-amide and LNA-OiPr linkages in the luciferase reporter assay, a second assay was selected in order to demonstrate that these backbone chemistries are effective at splice-switching in a therapeutically relevant model. The luciferase reporter assay is an efficient in vitro screening method which has many useful advantages − first, the HeLa pLuc/705 cell line is simple to culture and the cells are simple to treat, requiring no changes in morphology (e.g. differentiation). Executing the assay requires only a single luciferase reagent and the commonly-used BCA protein assay kit (as opposed to more involved procedures such as PCR which require RNA extraction kits, various enzymes, and primers, probes, etc.). From cell lysis to data collection, the assay only requires ∼120 minutes. The luciferase reporter assay is, therefore, a facile, quick, and cheap method of

evaluating many different ASOs with chemical modifications. However, it also has a number of limitations:

1. The in vitro model (HeLa pLuc/705 cells) does not inherently have a disease pathology caused by exon or intron skipping mutations and the mRNA and protein produced by successful splice-switching are not therapeutically relevant.[208]

2. The assay uses luciferase protein quantity as a measure of pre-mRNA splice-switching and does not take into consideration the rate of translation of this spliced mRNA into protein. Therefore, it assumes all mRNA with a restored reading frame has been translated at the same efficiency as mRNA without a restored reading frame and that the protein quantity directly reflects mRNA quantity.

3. Luminescence is normalised to protein quantity which is determined by BCA assay (using separate lysate samples from the identical well). However, protein quantity can vary, not just based on cell number, but also on health and other nonspecific cellular responses − an assay using a co-reporter internal control would minimise these effects.

4. The luciferase measurement is very time-sensitive; while luminescence should decrease in all wells at the same rate, plate-to-plate variation can occur due to differences in time between reagent addition and plate reading. Therefore, the plate measurement was always conducted immediately after the luciferase reagent was added.

Due to these limitations and in the interest of evaluating LNA-neutral linkages in a therapeutic model, the most promising modifications (LNA-amide and LNA-OiPr) were investigated in a dystrophic murine model, the H2k *mdx* cell line.

5.3.2 A therapeutic model – splice-switching in the H2k *mdx* in vitro model

20-mer ASOs were synthesised with the sequence complementary to the exon-23 splice site in the H2k *mdx* myoblast dystrophin gene (Table 5.2). ASOs containing one, two, and four incorporations of the LNA-amide and LNA-OiPr linkages were synthesised (ASOs 21-26), as well as 2'OMe/PS and LNA positive controls (ASO-19 and ASO-20).

ASO#	Sequence (5' → 3')	Expected Mass (Da)	Found Mass (Da)	Modifications x = linkage
ASO-19	GGCCAAACCUCGGCUUACCU	6885.0	6885.5	2'OMe/PS
ASO-20	GG**C**CAAACC**T**CGGC**TT**AC**C**T	6981.2	6986.0	2'OMe/PS/LNA
ASO-21	GGCCAAACC**TxC**GGCUUACCU	6855.0	6856.5	
ASO-22	GGCCAAACC**TxC**GGC**TxT**ACCU	6825.0	6824.0	x = LNA-amide
ASO-23	GG**CxC**AAACC**TxC**GGC**TxTACxC**U	6765.0	6766.0	
ASO-24	GGCCAAACC**TxC**GGCUUACCU	6952.1	6952.5	
ASO-25	GGCCAAACC**TxC**GGC**TxT**ACCU	7019.2	7020.5	x = LNA-OiPr
ASO-26	GG**CxC**AAACC**TxC**GGC**TxTACxC**U	7153.5	7150.0	

Table 5.2: ASOs synthesised containing 2'OMe/PS, LNA, LNA-amide, and LNA-OiPr modifications for evaluation in an H2k *mdx* cell line.

The workflow for the culture and differentiation of H2k *mdx* cells is depicted in Figure 5.5. Briefly, myoblasts were seeded, and at 60-70% confluency, differentiation was initiated using a media change which lacks interferon-γ.[196] Myoblasts require 3-4 days to fully differentiate into striated myotubes. After differentiation, the striated myotubes were transfected with ASO (timeline 1, Figure 5.5)[d], the media was changed after 24 h (to minimise transfection toxicity), and the mRNA was harvested 48 h later.

However, timeline 1 did not yield sufficient dystrophin mRNA transcripts – it is known in the literature that it is more difficult to transfect fully differentiated myotubes with ASO than myoblasts.[209] Therefore, a second timeline (timeline 2, Figure 5.5) was conducted in which the myoblasts were transfected during the differentiation process (therefore overlapping the

[d]Using RNAiMax transfection agent. Gymnotic treatment in differentiation media was also conducted.

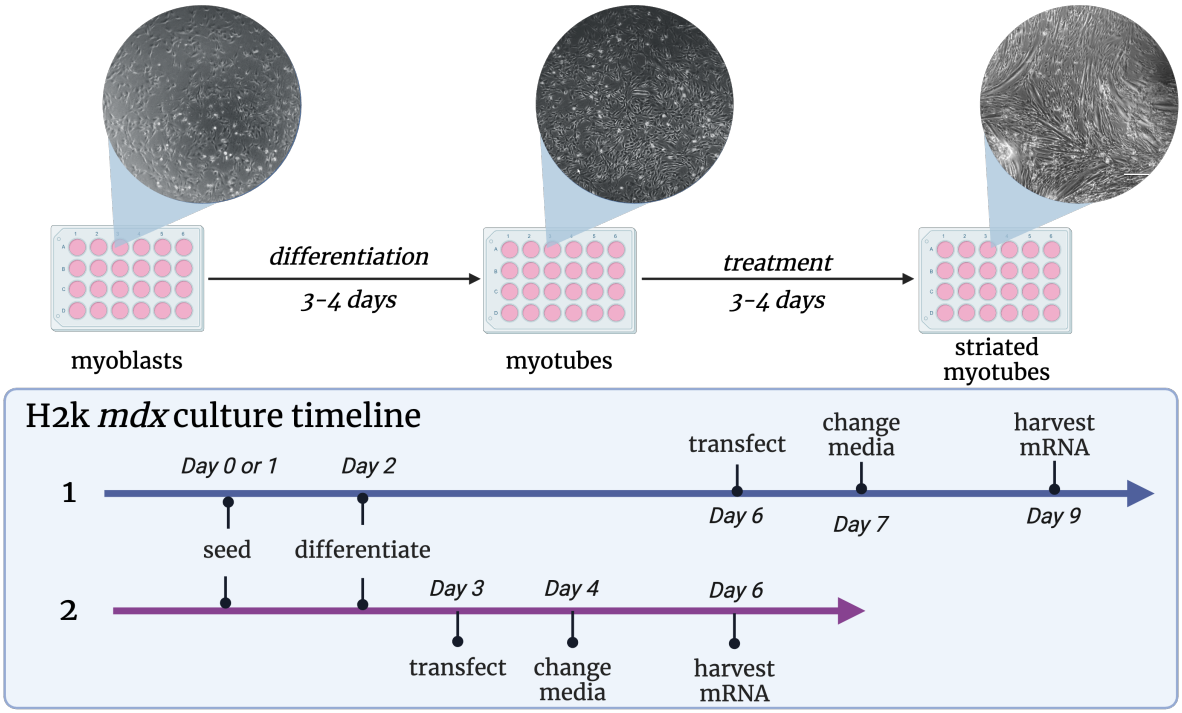


Figure 5.5: Schematic representation of the protocol for culturing H2k *mdx* cells and their treatment with splice-switching ASOs.

differentiation and treatment timelines). Additional difficulties arose in culturing these cells, briefly discussed here:

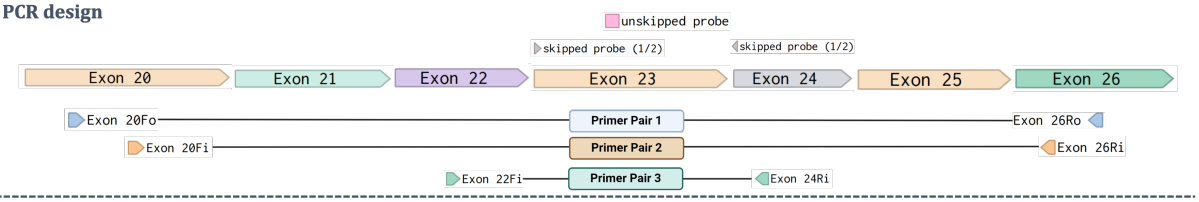
- A 24-well plate was necessary for cell culture, as the mRNA harvested was not sufficient when culturing in 48- or 96-well plates. This significantly increased the quantity of ASO required to treat the cells, especially when delivered via gymnosis in a 20 µM dose – the minimum media volume for a 24-well sample is 500 µL (vs. 100 µL for a 96-well plate in the luciferase assay).
- The cells were very sensitive to confluency at the point of their differentiation; if the myoblasts were differentiated at a low confluency (~60%) the cells did not have sufficient confluency to fuse successfully, but if they were differentiated at a higher confluency (~70-75%), they were often overconfluent by the end of the assay timeline.

Unfortunately, the mRNA which was harvested from these experiments was not successfully reverse-transcribed and amplified. In the first instance, qPCR was attempted using primers derived from the literature, which cover the region exon 20-26.[197,209,210] Probes were designed

to cover exon 23 and the exon 22-24 junction, shown in Figure 5.6A.[e] Using multiplex PCR (entry 1, Table 5.6), amplifying the exon 20-26 region and including both probes, the number of quantification cycles (C_q) was unfortunately >30, which was not reliable. The amplification of the housekeeping gene GAPDPH was used to ensure the quality of the mRNA; the average C_q value was 17 for the 2'OMe/PS control sample used to investigate PCR conditions, precluding the degradation of the mRNA as a cause.

Nested PCR is routinely used in the literature due to the very low abundance of full dystrophic mRNA.[197,211,212] Therefore, reverse transcription and the first amplification using primer pair 1 covering exon 20 (forward, outer) and exon 26 (reverse, outer) were first carried out, followed by a second amplification using primer pair 2 covering exon 20 (forward, inner) and exon 26 (reverse, inner) or primer pair 3 covering exon 22 (forward, inner) and exon 24 (reverse, inner) (see Figure 5.6). It was hypothesised that a shorter amplicon may be more successful. Unfortunately, neither nested qPCR conditions (entries 2, 3, Table 5.6) were able to give C_q values within a reasonable range of the standard curve.

A. PCR design



B. PCR conditions

Entry	Polymerase Enzyme	mRNA or cDNA quantity	Primers Outer: 20Fo/26Ro	Probes?	Conditions	Notes
1	Luna One-Step	25, 50, 100 ng mRNA	Multiplex	Y	Manufacturer (20 µL)	Cq > 30
2	Luna One-Step	50 ng mRNA	20Fi/20Ri	Y	Manufacturer (20 µL)	Cq > 30
3	Luna One-Step	50 ng mRNA	22Fi/24Ri	Y	Manufacturer (20 µL)	Cq > 30
4	Luna One-Step	100 ng cDNA	20Fi/20Ri	N	Manufacturer (20 µL)	Amplification, not replicable
5	Luna One-Step	100 ng cDNA	22Fi/24Ri	N	Manufacturer (20 µL)	No amplification
6	Q5	200 ng cDNA	20Fi/20Ri	N	Manufacturer (25 µL)	No amplification
7	Q5	200 ng cDNA	22Fi/24Ri	N	Manufacturer (25 µL)	No amplification
8	Amplitaq Gold	400 ng cDNA	20Fi/20Ri	N	Extended (50 µL reaction)	Amplification, not replicable
9	Amplitaq Gold	400 ng cDNA	22Fi/24Ri	N	Extended (50 µL reaction)	No amplification
10	Amplitaq Gold	200 ng cDNA	20Fi/20Ri	N	Extended (25 µL reaction)	No amplification

Figure 5.6: Table of (q)PCR conditions tested for the mRNA harvested from H2k *mdx* cells.

[e]I thank Ben Lagan for the design of the probes.

Finally, nested PCR was attempted without probes (entries 4-10, Table 5.6)[f] and the PCR products were visualised by 2% agarose gel, which has precedence in the literature.[197,211,212] Unfortunately, amplification of a product was only seen in two conditions (entries 5 and 8, Table 5.6, agarose gels can be seen in Figure E.3 and Figure E.4); when technical replicates on the same cDNA sample were conducted, this amplification was not replicable. Furthermore, the amplification of two distinctive bands (skipped vs. unskipped) was also not replicable.

It is hypothesised that the abundance of total dystrophin mRNA transcripts from the H2k *mdx* was simply too low to achieve replicable and reliable (q)PCR results.[g] In skeletal muscle, it is estimated that the fraction of dystrophin transcripts relative to total mRNA is 0.25-1.2%, while in cultured myoblasts/myotubes, this number is <0.1%.[213] Therefore, it is suggested in future work to evaluate ASOs containing LNA-amide and LNA-OiPr linkages in the H2k *mdx* murine model. Harvest of skeletal mouse tissue yield higher dystrophin mRNA transcripts and additionally serve as a better indicator of true therapeutic potential − in vivo models pose the ASO with additional barriers such as systemic distribution and tissue perfusion, as well as a better understanding of the absorption, distribution, metabolism, and excretion (ADME) for ASOs containing LNA-neutral linkages. The LNA-OiPr and LNA-amide linkages demonstrate very promising splice-switching activity in the luciferase assay with just two incorporations; while in vivo work requires large amounts of ASO, the synthetic requirements of the LNA-OiPr (just four custom monomer phosphoramidites) make it especially well-suited to further exploration. Furthermore, the optimised on-resin amide coupling conditions (Chapter 2) could be combined with automated solid-phase synthesis in future work to enable synthesis of larger quantities of LNA-amide-containing ASOs.

[f] cDNA was used during these reactions, transcribed from the mRNA using a high-capacity reverse transcription kit immediately after harvest, precluding mRNA degradation as a factor.

[g] Additional experience with this cell line may also increase the quality of the mRNA harvested. Unfortunately, time constraints precluded continued culture of these cells.

5.4 Protein-binding profiles of ASOs containing LNA-neutral linkages

While efforts to assess the exon-skipping efficacy of ASOs containing LNA-neutral linkages are ongoing, the differences in activity observed in the luciferase assay highlight the importance of studying ASO-protein binding profiles. A biotin pull-down method, established in the literature, can be used to "capture" intracellular proteins which bind to ASOs, depicted in Figure 5.7.[66,88,101]

5.4.1 Biotin pull-down assay

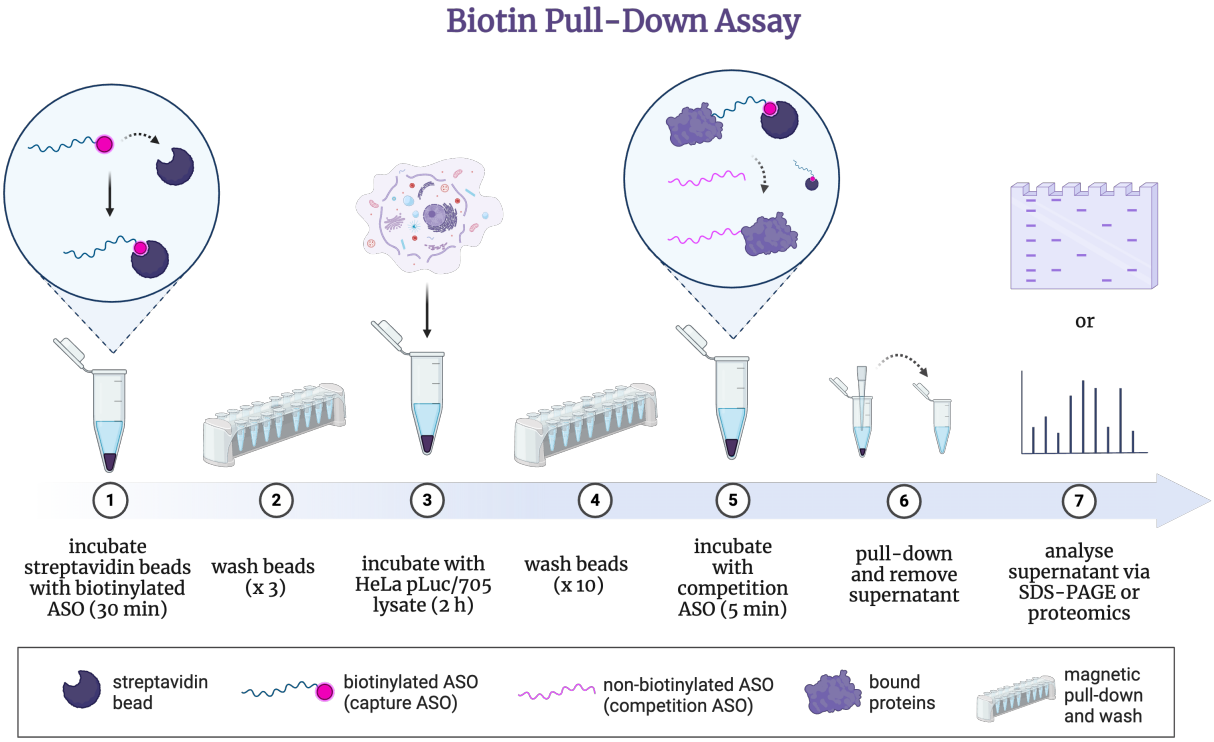
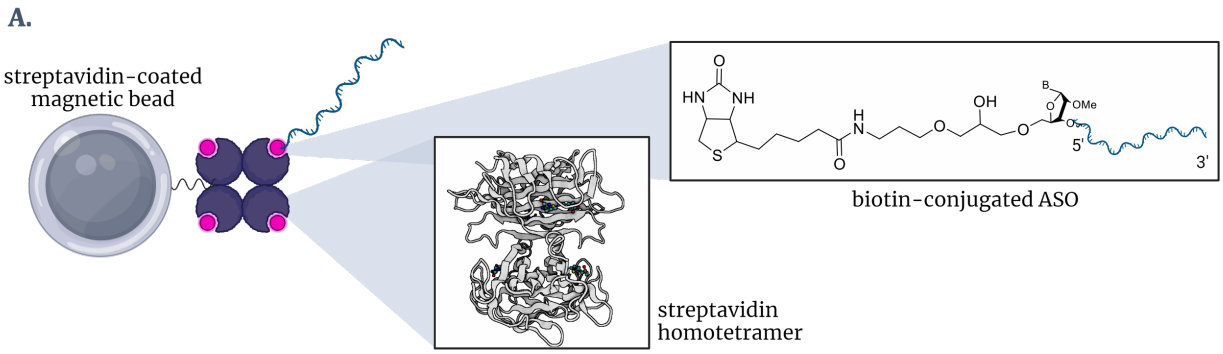


Figure 5.7: Schematic representation of the biotin pull-down assay protocol.

Briefly, the ASO containing the LNA-neutral linkage(s) is synthesised with a 5'-biotin handle and these are incubated with streptavidin-coated magnetic beads. The binding between the small molecule biotin and the streptavidin protein is amongst the strongest non-covalent interactions found in nature ($K_d = 10^{-14/-15}$ M), and is commonly used in many biological

applications; in this case, it renders the beads ASO-conjugated (step 1, Figure 5.7).[214] The beads are washed and incubated with cell lysate; proteins which bind ASOs remain associated with the streptavidin bead and intracellular proteins which do not have ASO-binding interactions are washed off (step 2-4, Figure 5.7). The proteins which remain bound to the ASO-bead complex are competitively eluted using ASOs which do not have biotin conjugation (termed "competition ASOs") (step 5, Figure 5.7). This method reduces non-specific protein binding, as the ASO-protein binding interaction must happen twice. The beads are pulled-down and the supernatant containing ASOs and their bound proteins can be removed



B.

	ASO#	Sequence (5'-3')	Expected Mass (Da)	Found Mass (Da)	Modifications x = LNA-neutral linkage
Capture ASOs	ASO-27	Biotin - CCUCUUACCUCAGUUACA	6518.8	6520.0	2'OMe/PS
	ASO-28	Biotin - CCTCTTACCTCAGTTACA	6614.8	6617.5	2'OMe/PS/LNA
	ASO-29	Biotin - CCUCTxTACCUCAGUUACA	6488.9	6490.5	x = LNA-amide
	ASO-30	Biotin - CCUCTxTACCUCAGTxTACA	6458.9	6460.0	x = LNA-amide
	ASO-31	Biotin - CCUCTxTACCUCAGUUACA	6585.9	6586.4	x = LNA-OiPr
	ASO-32	Biotin - CCUCTxTACCUCAGTxTACA	6653.1	6654.4	x = LNA-OiPr
Competition ASOs	ASO-1	CCUCUUACCUCAGUUACA	6096.3	6096.5	2'OMe/PS
	ASO-2	CCTCTTACCTCAGTTACA	6192.4	6195.5	2'OMe/PS/LNA
	ASO-3	CCUCTxTACCUCAGUUACA	6066.4	6068.5	x = LNA-amide
	ASO-4	CCUCTxTACCUCAGTxTACA	6036.5	6037.0	x = LNA-amide
	ASO-15	CCUCTxTACCUCAGUUACA	6163.5	6168.0	x = LNA-OiPr
	ASO-16	CCUCTxTACCUCAGTxTACA	6230.6	6231.0	x = LNA-OiPr

Figure 5.8: A. Schematic representation of streptavidin-coated magnetic beads, expanded to show the crystal structure of biotin-bound streptavidin (PDB, 6J6J) and structure of the 5'-biotin handle conjugated to ASOs. **B.** Table of ASOs synthesised having the 5'-biotin conjugated moiety (capture ASOs) and identical ASOs synthesised lacking the 5'-biotin moiety (competition ASOs).

and visualised by SDS-PAGE electrophoresis or alternatively analysed in other techniques such as immunostaining or mass spectrometry (steps 6-7, Figure 5.7).

Control ASOs (2'OMe/PS and LNA controls) and ASOs containing LNA-neutral linkages (LNA-amide and LNA-OiPr) were synthesised with a biotin handle attached to the 5'-end (termed "capture" ASOs; Figure 5.8A). Only ASOs containing one and two incorporations were synthesised (Figure 5.8B) for ease of synthetic requirements, as these incorporation positions required only TT dimer phosphoramidites or T monomer phosphoramidites − these ASOs were also the most biologically promising according to the luciferase splice-switching data. The identical ASOs synthesised previously without the 5'-biotin handle are termed "competition" ASOs (Figure 5.8B).

In the first instance, the proteins captured using this method were visualised by SDS-PAGE.[101] In the literature, 56 intracellular proteins which bind PS-ASOs[66] were identified using this method and the protocol has also been used to qualitatively assess toxic gapmers, because the extent of overall protein-binding is higher in toxic gapmers than non-toxic gapmers.[88] In this use-case, the biotin-pull down assay was applied to assess the qualititative differences, if any, that could be found among intracellular proteins binding to the ASOs containing LNA-neutral linkages. The advantages of this protocol are that it is relatively simple and quick − the protocol could be carried out in one day, including SDS-PAGE and silver staining, to qualitatively assess protein-binding of various ASOs. Furthermore, the biotin-handle phosphoramidite is cheap and the ASO purification is facile.

SDS-PAGE gels showing the protein-binding profiles of control ASOs and ASOs containing one or two LNA-neutral linkages are given in Figure 5.9. Important differences in the protein binding profile are visible between the LNA control and the 2'OMe/PS control − this was very promising, as it corresponds to literature findings which reveal that LNA-containing oligonucleotides bind to different proteins than 2'OMe-modified oligonucleotides.[215] Distinct bands are visible for the LNA control ASO but not as visible for the 2'OMe/PS control (band B, Figure 5.9) and vice versa (band A, Figure 5.9).[215] Interestingly, the LNA-neutral ASOs demonstrate an intermediate protein binding profile, which is reasonable as they in-

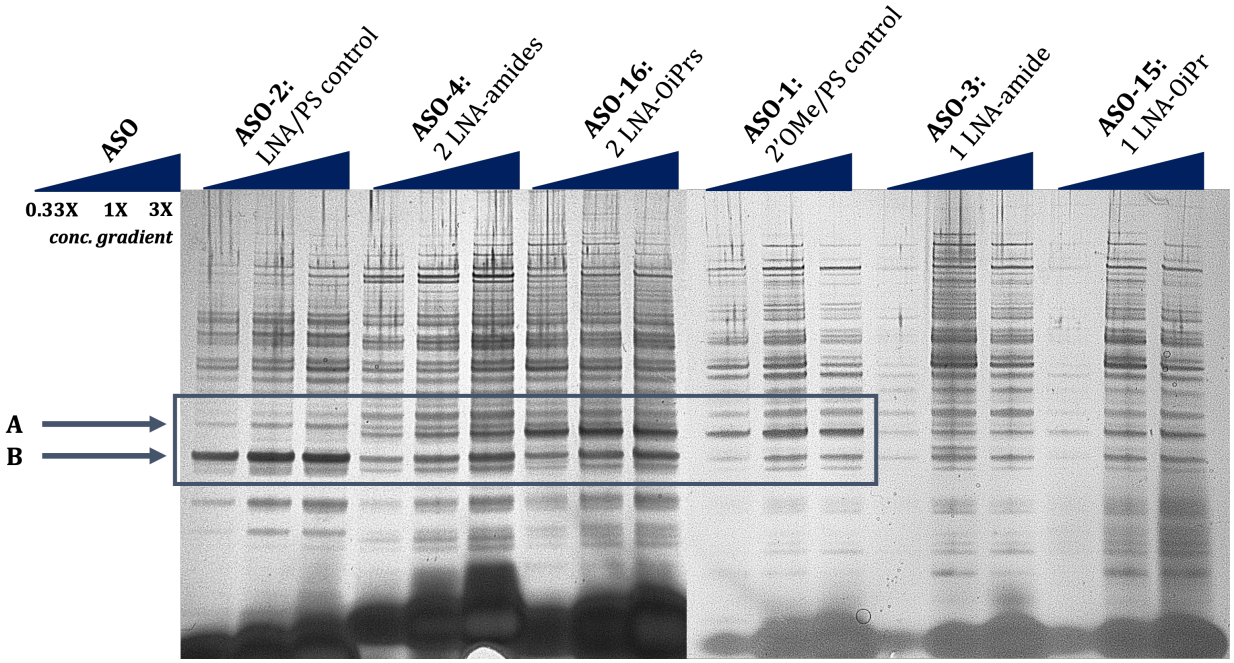


Figure 5.9: SDS-PAGE gel (4-12%) separating the proteins bound by ASOs containing LNA-neutral linkages and control ASOs, visualised by silver staining. A concentration gradient of competition ASO is -dependent response is observed, as the concentration of competition ASO is increased (0.33X, 1X, 3X).

clude an intermediate quantity of LNA moieties. However, differences between amide and OiPr-linkages can be observed; the ASO containing 2 LNA-amide modifications has greater quantities of the band B (abundant in the LNA control), while the ASO containing 2 LNA-OiPr modifications has slightly greater quantities of band A (abundant in the 2'OMe control). The ASOs containing one incorporation of the LNA-amide and LNA-OiPr each have a fainter protein-binding profile (likely due to differences in washing from one experimental assay to the next). However, when discounting band strength, they too reveal intermediate presences of bands A and B.

However, the drawbacks of this protocol are that it is qualitative. Despite attempts to maintain consistency in the washing steps, some beads may be washed more vigorously than others, leading to lower overall protein quantities in one eluent over another. Varying the concentration of the competition ASO is one method to increase confidence that a certain protein size group binds more extensively to an ASO than another, as this trend should be

seen across all concentrations. However, the lack of consistency during the washing step remains, as it can be seen in Figure 5.9 − the samples for one LNA-amide (ASO-4) contains more overall proteins in the sample eluted with 20 µM ASO (1X concentration) than the 60 µM (3X) sample. Additionally, the protocol cannot identify the proteins and only give an indication as to their size.[h]

The hypothesis that the ASOs studied thus far would have different protein-binding profiles was substantiated, and intrigued by these qualitative differences, a quantitative approach was pursued. The ability to identify and quantify these proteins by proteomics offered a reasonable further investigation.[i]

5.4.2 Label-free quantification

Proteomics, the study of proteomes and their functions, has been revolutionised by the advent of advanced mass spectrometry. Advanced mass spectrometry techniques can identify and quantify proteins, even without the use of peptide or protein labels ("label-free").[216] Label-free quantification (LFQ) of the eluted proteins was carried out to identify and quantify the relative amounts of each protein in the samples. Briefly, the supernatant produced by competitive elution and final pull-down step (step 6, Figure 5.7) was digested using a benzonase endonuclease to reduce DNA and RNA background noise. The proteins were digested to smaller peptide fragments using endoproteinases lys-c and trypsin, commonly-used serine proteases which cleave proteins into fragments ranging from 700-1500 Da.[217,218] Peptide fragments were separated by mass and compared to a database of known protein/peptide fragments to identify proteins present in the sample. A peptide-spectrum match (PSM) is identified when the recorded MS/MS spectrum matches the theoretical MS/MS spectrum of the protein.[219] Each PSM is a probability, evaluated for "false positives" using a 1% false discovery rate (FDR) threshold.[220] A spectral abundance factor (SAF) is calculated for each

[h]The first gels conducted included a protein size ladder; however, the silver staining was extremely sensitive − even the smallest quantities of ladder loaded onto the gel (0.5 µL, diluted) caused bleeding into the first few lanes. Therefore, it was omitted for the final gels.

[i]I thank Dr. Marjorie Fournier in the Department of Biochemistry Advanced Proteomics Facility for her guidance and the Advanced Proteomics Facility for conducting these experiments.

protein by normalising the total number of identified peptide sequence matches (PSMs) to the size of the protein (Equation 5.1), indicating the abundance of a given protein's presence in the sample.

$$\mathrm{SAF}^{\text{specific protein}} = \frac{\#PSM}{\#AA} \tag{5.1}$$

A normalised spectral abundance factor (NSAF) considers this abundance in relation to the total protein quantity of that sample (Equation 5.2).[216]

$$\mathrm{NSAF} = \frac{\mathrm{SAF}^{\text{specific protein}}}{\sum \mathrm{SAF}^{\text{all proteins}}} \tag{5.2}$$

The biotin-pull down protocol was repeated for four capture and competition ASO pairs – the 2'OMe/PS control (ASO-1), LNA control (ASO-2), ASO containing 2 LNA-amide linkages (ASO-4), and ASO containing 2 LNA-OiPr linkages (ASO-16). LFQ was performed on the supernatant containing the proteins bound by these ASOs.[j]

In order to perform LFQ, the final three washes of the beads and the final competitive elution was conducted in a PBS-only buffer lacking Tween-20 due to the detergent being incompatible with mass spectrometers.[221] This increase in surface tension led to a slower magnetic pull-down and different quantities in beads in the final competitive elution step. Unfortunately, samples pulled down using the 2'OMe control and LNA control contained greater final total protein quantities than samples pulled down using the ASOs with LNA-neutral linkages (see SAF sums in Table 5.3). Although the NSAF metric normalises the protein abundance to the total protein quantity within the sample, comparing samples with very different total protein quantities is not advised. A protein with a small SAF in a sample with very few proteins may lead to an inflated sense of its abundance, compared to its SAF in a different sample with a larger variety and number of total proteins.

Therefore, the proteins isolated using the 2'OMe control and LNA control are compared

[j]I thank the Advanced Proteomics Facility for conducting these LFQ experiments.

SAF (protein)	ASO-1 2'OMe/PS control	ASO-2 LNA/PS control	ASO-4 2 LNA-amides	ASO-16 2 LNA-OiPrs
SAF (HMGB1)	0.153	0.181	0.019	0.028
SAF (SSBP)	0.128	0.047	0.047	0.047
SAF (hnRNP A1)	0.110	0.051	0.024	0.013
SAF (hnRNP A/B)	0.099	0.074	0.045	0.019
....	...	...	...	...
∑ SAF	4.11	3.71	0.85	0.45

Table 5.3: Table of spectral abundance factors (SAF) for the most abundant proteins and SAF sums calculated for the samples produced by the biotin pull-down assay.

first, with limited discussion of the proteins isolated by LNA-neutral ASOs. In future work, the necessary lack of Tween-20 would be taken into consideration; potentially, solely the beads could be submitted for digestion, but this would likely increase non-specific binding quantities.[k]

5.4.2.1 Protein-binding profiles of 2'OMe vs. LNA-containing ASOs

20 proteins of interest were identified and selected for discussion: eight proteins of interest were identified in which the NSAF for the 2'OMe control was greater than the LNA control, four proteins in which the NSAF was similar, and eight proteins in which the NSAF for the LNA control was greater than the 2'OMe control. PSM count was also taken into consideration; the higher the PSM count, the more probable that this protein was identified. Therefore, proteins with higher PSM counts were prioritised as proteins of interest. However, it should be noted that larger proteins tend to have higher number of PSMs; repeated experiments should be conducted to increase confidence in the proteins which are identified; this may also enable discussion of smaller proteins with lower PSM counts.

Most of the proteins which were identified were DNA- or RNA-binding proteins (Figure 5.10, left), which was not surprising. However, some proteins of interests also include enzymes such as citrate synthase and proteins involved in transport or structure, such as Ran or Actin, respectively.[198] The vast majority of the proteins are localised to the nucleus or cytoplasm (or

[k] Unfortunately, the high cost of LFQ prevented a repeat experiment.

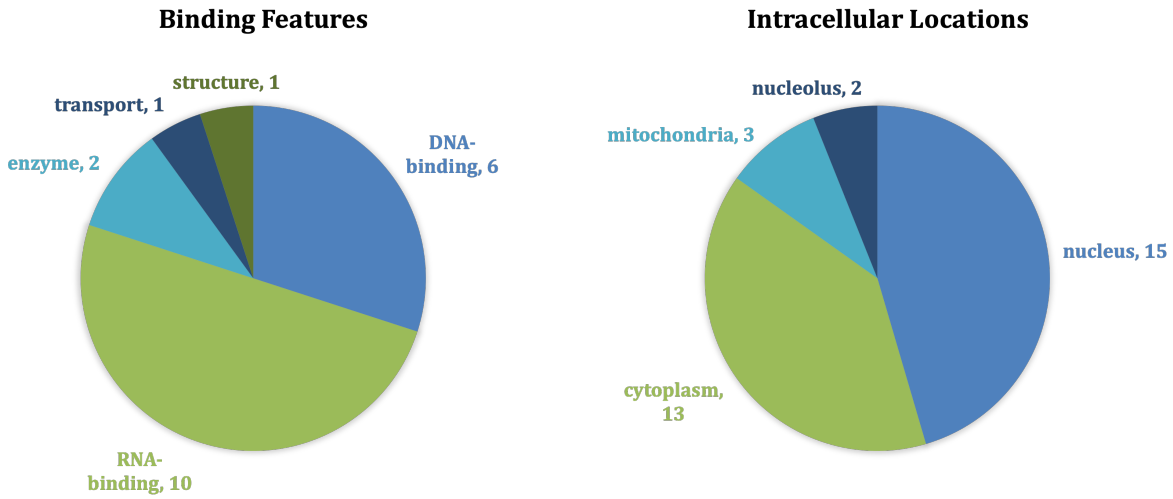


Figure 5.10: 20 proteins of interest were selected when comparing the binding profiles of the 2'OMe and LNA control ASOs. The total count of proteins which share the same characteristics are categorised by binding features (left) and intracellular location (right).

both), with some proteins, specifically the enzymes, present in the mitochondria. Nucleolin and the helicase DDX5 are both primarily localised in the nucleolus (Figure 5.10, right).[222,223]

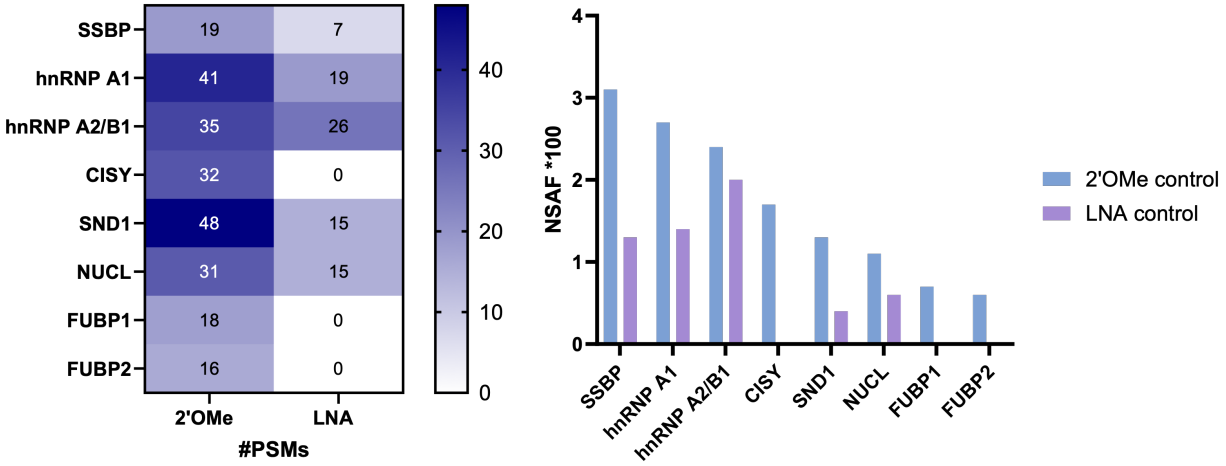


Figure 5.11: Eight proteins of interest were identified whose abundance was greater in the 2'OMe sample than in the LNA sample, by #PSMs (left) and NSAF (right).

Using this biotin pull-down protocol coupled with mass spectrometry LFQ, it was promising to identify many proteins which are reported in the literature as ASO-binding.[66,198] 8 proteins of interest were selected for discussion that were more abundant in the sample pulled down by the 2'OMe control ASO than the LNA control (Figure 5.11):

- Single-stranded DNA binding protein (SSBP) – this protein plays a key role in DNA

repair and replication, is known to affect ASO localisation in the nucleus, and enhances an ASO's activity.[224] A greater abundance found bound to the 2'OMe control could be one of the reasons behind the increased gymnotic activity of ASO-1.

- Heterogenous nuclear ribonucleoproteins (hnRNP A1, A2): well-known to bind to PS-ASOs, their effect on ASO activity has yet to be characterised. These proteins are among the most abundant in the nucleus, involved in pre-mRNA processing, and shuttle between the nucleus and cytoplasm.[225]

- Citrate synthase (CISY) is reported as an ASO-binder and no change in ASO activity is reported upon knockdown nor is it involved in ASO localisation.[198]

- Staphylococcal nuclease domain-containing protein 1 (SND1): this protein composes part of the RISC complex and is indicated in the RNAi intereference pathway. It would be interesting to investigate this protein-binding interaction further for siRNA development.[226]

- Nucleolin (NUCL): this protein has been shown to bind PS ASOs with various 2'-modifications (2'-F, 2'-MOE, 2'-cEt) with similar affinity[227] but in this case, it was more abundant in the 2'OMe sample than in the sample pulled down by the LNA control.

- Far upstream element-binding protein 1 and 2 (FUBP1 and FUBP2): these proteins were identified exclusively in the 2'OMe ASO sample. Upstream elements have been selected as ASO targets in the literature;[228] the potential higher extent with which they bind to 2'OMe ASOs over LNA-modified ASOs could be of future investigation for ASOs designed to target these regions.

Proteins in which the abundance for 2'OMe and LNA ASOs were relatively equal may indicate that they have no preference for 2'-chemical modifications, compared to other less promiscuous proteins. Four proteins of interest were identified with the highest PSM and NSAF scores (Figure 5.12):

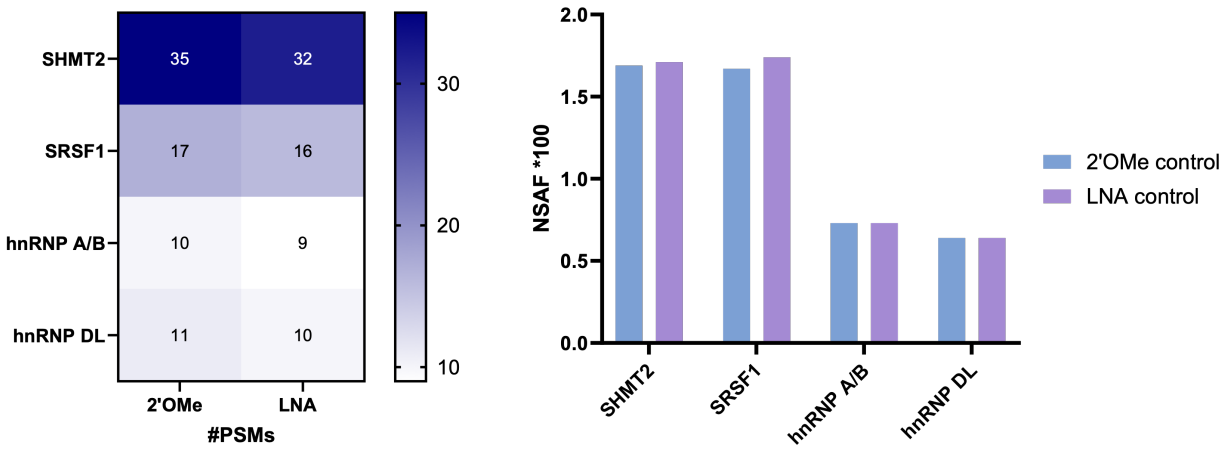


Figure 5.12: Four proteins of interest were identified whose abundance was relatively similar in the 2'OMe and LNA samples, by #PSMs (left) and NSAF (right).

- Serine hydroxymethyltransferase (SHMT2): this enzyme was among the most abundant of all proteins in both the 2'OMe and LNA samples, with PSM counts of 35 and 32, respectively. It is known in the literature as an ASO-binder but its effects on activity are not characterised.
- Serine/arginine-rich splicing factor 1 (SRSF1): also known as alternative splicing factor 1 (ASF), this protein shuttles between the cytoplasm and nucleus and is a regulator of both mRNA splicing and translation – binding to this protein should be further investigated for splice-switching ASOs.[229]
- hnRNP A/B and hnRNP DL: it is interesting to note that while hnRNP A1 and A2 were more abundant in the 2'OMe sample, hnRNP A/B and D-like were equally abundant in the 2'OMe and LNA samples. This preliminary result could indicate potential 2'-modification-specific binding in this family of proteins.

Finally, proteins were observed which were more abundant in the LNA control sample than in the 2'OMe control (Figure 5.13). Among these proteins are some which are associated with ASO co-localisation, aggregation, and subsequent cell toxicity – these ASO-protein interactions could provide preliminary reasoning behind the reduced gymnotic activity for the LNA control (ASO-2).

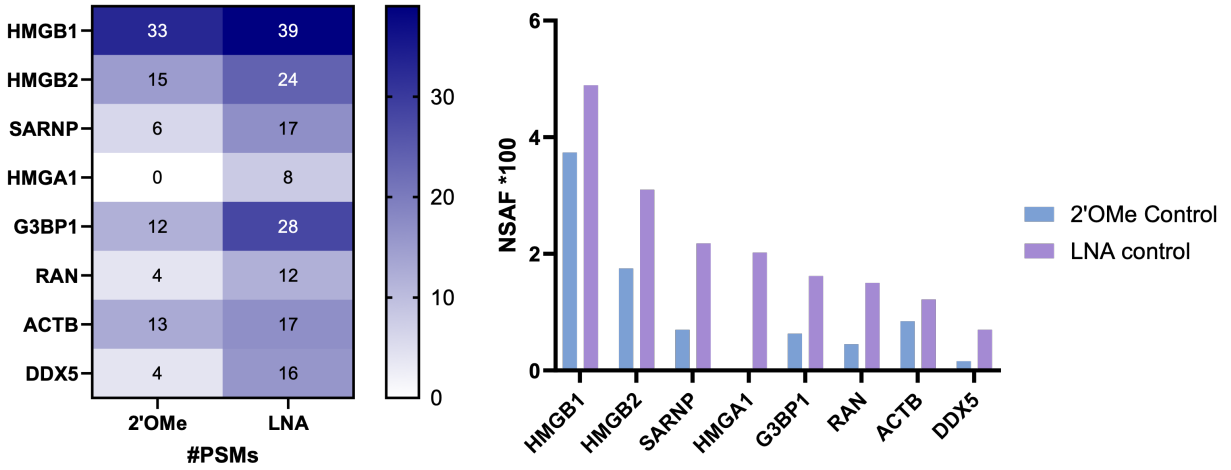


Figure 5.13: Eight proteins of interest were identified whose abundance was greater in the LNA sample than in the 2'OMe sample, by #PSMs (left) and NSAF (right

- High mobility group proteins (HMG B1, B2, A1): this family of DNA-binding proteins are involved in the regulation of transcription.[230] HMGB1 is known to induce no change in ASO activity, nor does it co-localise with PS-ASOs, but it is interesting to note that this entire family of proteins was more abundant for the LNA-control sample, pointing to some potential 2'-modification preference.[198]

- SAP domain-containing ribonucleoprotein (SARNP): this ribonucleoprotein is upregulated in response to various cytokines and is involved in mRNA transport but not yet reported as a protein which binds to ASOs in a therapeutically significant way.[231]

- Ras GTPase-activating protein-binding protein 1 (G3BP1): this protein is involved in the formation of stress granules and is well-known to bind ASOs with more hydrophobic 2'-modifications (e.g. cEt), even redistributing cEt-ASOs from the nucleus to cytoplasm aggregates. This cellular trafficking out of the nucleus may be one reason the LNA control ASO observed reduced splice-switching gymnotic activity.[183]

- Probable ATP-dependent RNA helicase (DDX5): other members of the DDX family (e.g. DDX6 and DDX21) are known to bind to PS-ASOs, but DDX5 is not yet reported. DDX21 is known to co-localise with toxic ASOs, and its de-localisation is linked with nucleolar stress and p53 activation, leading to cytotoxicity.[101]

- GTP-binding nuclear protein Ran (RAN): this protein is involved in nucleo-cytoplasmic shuttling and is known to bind proteins in aggregates.[232]
- Beta actin (ACTB): this cytosolic structural protein is known to reduce ASO activity upon knockdown, but its effects on ASO localisation are not yet characterised.[198]

It was promising to observe that the biotin pull-down protocol coupled to LFQ was able to identify proteins which are known to bind ASOs, verifying this method. Additional proteins which are DNA- or RNA-binding were identified, which require further investigation to confirm their ASO:protein interactions. Amongst those known in the literature, it was interesting to see a marked difference in abundance of proteins between 2'OMe and LNA-containing PS-ASOs, especially those such as SSBP, nucleolin, hnRNPs, and HMGPs.

5.4.2.2 Protein-binding profiles of ASOs containing LNA-neutral linkages

Due to the overall lower quantity of proteins in the samples produced by LNA-neutral linkage biotin pull-down (see Table 5.3), only a few proteins of interest could confidently be selected for discussion. However, it was encouraging to see proteins pulled down using ASOs containing two LNA-neutral linkages were all known in the literature to bind ASOs and were also found in the 2'OMe and LNA-containing ASOs in high abundance (Figure 5.14). The NSAF for these proteins was artificially high, as the total protein quantity is lower in these two samples (Figure 5.14, right). However, certain proteins, despite being part of a smaller total number, did have large PSM counts (Figure 5.14, left) − this increases the confidence that they would be identified again as abundantly bound if repeated, and so, they are briefly discussed.

The LNA-amide sample is unique amongst the four samples in having reduced PS-content, as the amide replaces the phosphorothioate moiety entirely. It is interesting to note that nucleolin (NUCL) is very abundant in the LNA-amide sample by PSM count, but much less so in the sample containing LNA-OiPr linkages (21 PSMs vs. 4 PSMs), which may indicate additional affinity for other chemical backbones besides PS. Similar affinity for the LNA-amide ASO is seen for proteins such as SHMT2 and hnRNP A2/B1 which are known

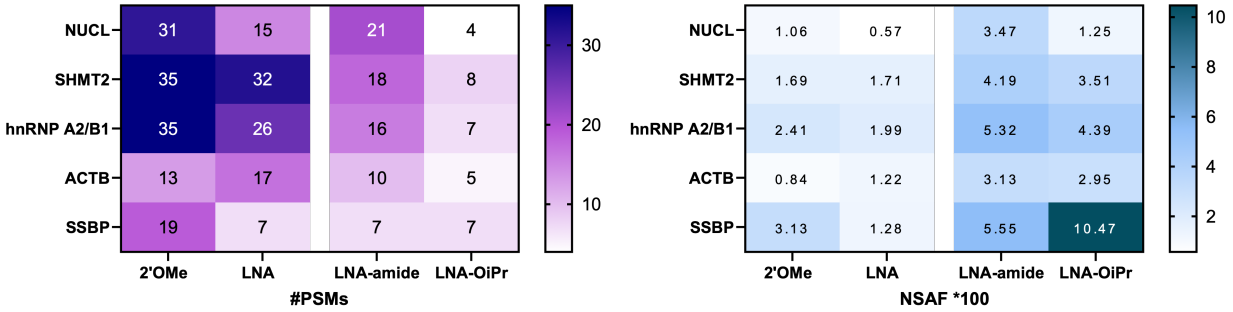


Figure 5.14: Five proteins of interest were identified whose abundance was high in samples pulled down using LNA-neutral linkages, by #PSMs (left) and NSAF (right).

to bind PS-ASOs but are also greatly abundant in the LNA-amide sample.

The knockdown of cytoplasmic actin is known to reduce ASO activity, indicating it enhances ASO activity − both the LNA-amide and LNA-OiPr samples had high PSM matches for this protein (10 and 5 PSM, respectively), potentially contributing to their enhanced gymnotic activities. Finally, the SSBP protein was identified in the LNA-neutral ASO samples − indeed, the PSM count of SSBP proteins in LNA-neutral ASOs was equal to the LNA control (7 PSMs), despite these samples being significantly lower in protein quantity. Therefore, the NSAF for SSBP in the LNA-neutral samples was significantly higher than both the 2'OMe and LNA control ASOs. If ASOs containing LNA-OiPr and LNA-amide modifications do indeed bind SSBP with more affinity than the 2'OMe or LNA modifications alone, this may be one indication as to their promising splice-switching activity − SSBP is a protein known to localise ASOs, increase nuclear accumulation, and enhance ASO activity.

5.5 Conclusions

This final experimental chapter aimed to delve deeper into the most promising LNA-neutral linkages, first comparing the LNA-neutral backbones synthesised in Chapters 2, 3, and 4 in the luciferase splice-switching assay. ASOs containing two incorporations of the LNA-amide and LNA-OiPr linkages were most active and it was envisioned to test them in a therapeutic in vitro model. Unfortunately, this aim was not met due to the difficulty in culturing H2k *mdx* cells and difficulty in extracting exon skipping data from the harvested mRNA. This

remains an important objective for future work, as the results may strengthen the promising preliminary results of the LNA-amide and LNA-OiPr and shed further insight on design rules such as incorporation number and position in a therapeutic exon skipping ASO.

Finally, the chapter aimed to preliminarily investigate the protein binding profiles of the ASOs containing LNA-neutral linkages. The aim was to determine first qualitatively, and second quantitatively, the differences (if any) in intracellular proteins bound by these ASOs. The biotin pull-down assay was successfully implemented from the literature[101] and qualitative differences among ASOs containing 2'OMe, LNA, and LNA-neutral linkages were indeed observed via SDS-PAGE gel. The protocol was applied again to identify and quantify proteins of interest by proteomic mass spectrometry methods (LFQ). Unfortunately, not all samples were comparable due to variation in total protein quantity. However, many proteins identified and quantified were known in the literature to bind PS-ASOs, some indeed having preferences for certain 2'-modification chemistries. A preliminary analysis of the proteins found bound to ASOs containing LNA-neutral linkages, such as actin and SSBP, did indicate that these promising ASOs do bind certain proteins known to enhance activity. These protocols should be repeated in order to account for the differences in total protein quantity and of course, it would be interesting to observe quantifiable differences among ASOs containing one, two, and four incorporations or ASOs containing less active linkages such as the LNA-carbamate. However, preliminary results indicate that this is a viable method to further elucidate ASO-protein interactions for novel chemical modifications.

6

Conclusions

Contents

6.1 Main research outcomes

In this thesis, the successful synthesis and study of chemically-modified antisense oligonucleotides (ASOs) is described. This work contributes to the growing field of therapeutic ASOs, large molecules with the potential to treat previously undruggable diseases. This thesis sought to expand the ASO chemical space by combining neutral backbones with the stabilising LNA ribose modification to obtain ASOs with drug-like properties. The main research outcomes of this thesis, as well as specific future work which may expand its scope, are summarised:

- Synthesis of ASOs containing LNA-neutral backbones
  - Dinucleotide phosphoramidites containing LNA-amide linkages (Chapter 2) and LNA-carbamate linkages (Chapter 3) as well as alkyl isopropyl triester monomer phosphoramidites (Chapter 4) were synthesised.
  - A dinucleotide containing an LNA-amino-oxy-amide linkage was synthesised (Chapter 3); in future work, alkylation of the reactive nitrogen in this linkage may provide higher phosphitylation yields and enable ASO synthesis and further investigation.
  - Clinical-scale oligonucleotide synthesis requires larger quantities; on-resin amide coupling was optimised (Chapter 2), which could enable solid-phase synthesis of the LNA-amide linkage in future work.
- Biophysical studies of ASOs containing LNA-neutral backbones
  - While all LNA-neutral linkages reduce the electrostatic repulsion between oligonucleotide strands, not all LNA-neutral linkages therefore stabilise an ASO:RNA heteroduplex. Conformationally unfavourable neutral linkages which are not sufficiently flexible may reduce duplex stability despite lowering the electronic repulsion. The most thermally stabilising linkage was the LNA-OiPr linkage, followed by the LNA-sulfamate[a] and LNA-amide − these linkages were also most active in

[a]I thank Dr. Alice Kennett for synthesising this dimer phosphoramidite.

the luciferase splice-switching assay. The LNA-carbamate and LNA-alkoxyamide[b] were similarly stable to ASOs containing only 2'OMe/PS modifications. Circular dichroism studies revealed that LNA-neutral linkages encouraged a B→A-form shift in ASO:DNA heteroduplexes and a mixed A/B topology in ASO:RNA heteroduplexes; perturbations in the ASO:RNA heteroduplex were likely due to the rigid C3'-endo conformation forced by the LNA modification in the 5'-nucleotide.[62]

- Through collaboration, a computational approach was applied to many of the LNA-neutral linkages in this thesis.[c] Molecular dynamics simulations results, which lent insight into the global conformational parameters of the helix, corresponded with observed experimental data − in future work, extending the modelled duplex to include 2'OMe/PS modifications in the ASO strand could provide an even more therapeutically-relevant model. Modelling additional novel chemical backbones in ASO:RNA duplexes could be applied to further establish this correlation and potentially, MD could begin to be used to predict a modification's success.

- Biological studies of ASOs containing LNA-neutral backbones

  - ASOs containing LNA-neutral linkages were evaluated for splice-switching activity in a luciferase reporter assay. Transfection was used as a delivery method to evaluate the activity of the ASOs if delivered intracellularly, while gymnotic delivery was used to determine if ASOs containing LNA-neutral linkages remained active once cellular uptake was required. It became clear that backbone chemistry greatly affects both inherent splice-switching activity, but also may affect cellular uptake. The most active linkages (LNA-OiPr, LNA-sulfamate, and LNA-amide) are good bioisosteres of the natural phosphodiester as well as flexible in their conformation. Future work on these LNA-neutral linkages in splice-switching models should prioritise therapeutic relevance by conducting in vivo work with

[b] I thank Dr. Belma Kurt-Zengin for synthesising this dimer phosphoramidite.

[c] I thank Martin Flerin, DPhil candidate supervised by Prof. Fernanda Duarte, for conducting the computational experiments.

the dystrophic murine model. Furthermore, study of additional ASOs containing LNA-neutral linkages in varying positions and incorporation number could aid in future rational design.

– ASOs have different protein-binding profiles depending on chemical modification. A biotin pull-down assay, coupled to mass spectrometry LFQ, was applied to ASOs containing 2'OMe, LNA, and LNA-neutral modifications. Multiple proteins were identified which bound these ASOs to different extents, such as SSBP and nucleolin.

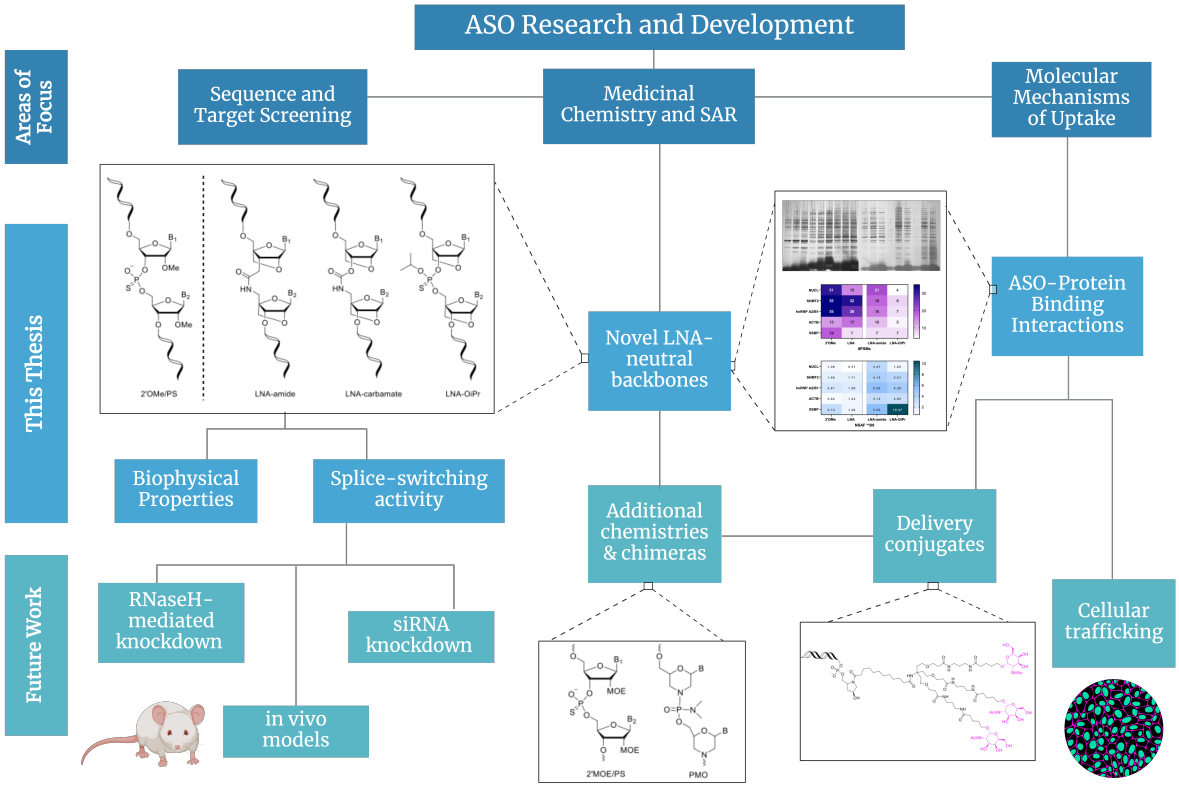


Figure 6.1: Flowchart depicting the three main areas of ASO research and development, mapping the main research outcomes of this thesis and future work.

6.2 Future work

As is the case with all scientific research, one door opened leads to many more. A variety of additional projects could provide a deeper investigation of LNA-neutral linkages in therapeutic ASOs, could begin to progress novel modifications to the clinic, or tackle some of the

clinical challenges facing ASOs as a therapeutic modality (Figure 6.1). Of many possible projects, some of the most relevant future work is suggested here:

- Additional chimera chemistries − LNA-neutral linkages were combined with 2'OMe/PS modifications to synthesise chimeric therapeutic ASOs. However, their combination with other clinically-validated modifications such as PMO or 2'MOE should be explored. If ASOs containing LNA-neutral linkages combined with PMO or 2'MOE/PS display higher splice-switching activity than their respective positive control (e.g. an ASO which is solely PMO or 2'MOE/PS modified), it may help progress the modification to pre-clinical work. Novel chemical modifications are still sought, but they should render the ASO even more drug-like than all "standard" modifications which have demonstrated clinical efficacy in the past.

- Additional mechanisms of action − LNA-neutral linkages should be incorporated into gapmers or siRNA and these ASOs evaluated for knockdown efficacy. If ASOs containing LNA-neutral linkages demonstrate significantly increased activity in gapmers, as well as siRNA, this would extend their therapeutic potential to all mechanisms of action, making them more "universal" modifications which can be applied to many more preclinical candidates.

- Additional ASO-protein interactions − the biotin pull-down assay described and conducted in Chapter 5 captures only intracellular proteins. Plasma protein-binding by LNA-neutral ASOs (e.g. extent to which the ASO is albumin-bound) may be important to understand biodistribution of these ASOs if administered systemically. Additionally, interactions with membrane-bound receptors which are known to associate with ASOs (e.g. stabilin or integrin) should be studied to further elucidate the endocytotic pathway in which LNA-neutral ASOs are intracellularised (e.g. clathrin- or caveolin-dependent endocytosis (or other)).[202]

- Cellular trafficking − in order to better understand the promising gymnotic activity of ASOs containing LNA-neutral linkages, their intracellular distribution could be ob-

served by techniques such as confocal fluorescence microscopy. Fluorescence-labelled-ASOs could be tracked and observed in various subcellular compartments such as the cytoplasm, nucleus, endosomes, lysosomes and vesicles produced by various endocytotic pathways.[233]

- Delivery conjugates − transfected data suggest that the ASOs containing LNA-amide and LNA-OiPr linkages are inherently more active than the 2'OMe/PS control once intracellularised. Conjugation to a delivery moiety such as GalNac, discussed in Chapter 1, might continue to progress these modifications to the clinic. A more potent drug cargo would reduce the therapeutic dose required and potentially reduce the hepatotoxicity associated with many liver-targeting ASOs.

In conclusion, this thesis has expanded the chemical space of antisense oligonucleotides through the synthesis and study of novel LNA-neutral backbones; it is my hope that this small contribution to their medicinal chemistry will add to the broader advancements of ASOs as a promising and effective class of therapeutic drugs.

7

Materials and Methods

Contents

7.1 Small molecule synthesis

7.1.1 General procedures

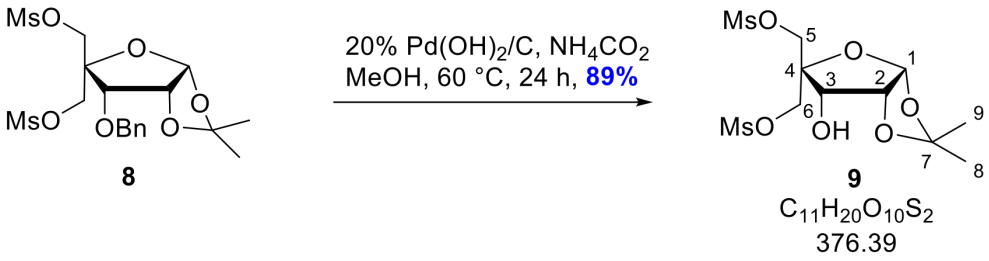
All reagents for organic synthesis were purchased from Sigma-Aldrich, Fluorochem, Carbosynth, Thermo Fisher Scientific, Alfa Aesar, Hongene, and used without further purification (unless otherwise stated). Anhydrous solvents (CH_2Cl_2, pyridine, EtOAc, THF, Et_3N) were obtained from an MBraun benchtop solvent purification system (SPS). Air and moisture sensitive reactions were carried out under inert atmosphere using argon in oven-dried glassware. Solvents used in phosphitylations (reaction, work-up, and purification) were degassed with argon for at least 20 min. Thin-layer chromatography (TLC) was used to monitor reactions using Kieselgel 60 F24 silica gel plates (*Merck)*; plates were visualised by UV light at 254/265 nm and stained with *p*-anisaldehyde followed by gentle heating. Column chromatography was carried out using a Biotage SP4 system using Biotage Sfär Duo or Biotage Sfär C18 Duo columns, unless otherwise specified.

^{1}H NMR (400 MHz, 500 MHz, or 600 MHz), ^{13}C NMR (101 MHz, 126 MHz, or 151 MHz) and ^{31}P NMR (162 MHz) were recorded on a Bruker AVIIHD Nanobay, Bruker AVIIIHD 500, or Bruker NEO 600, and ^{1}H and ^{13}C spectra were referenced to the appropriate deuterated solvent signal. Chemical shifts are given in ppm and coupling constants (J) are given in Hertz (Hz). COSY, HSQC, and HMBC spectra were acquired for assignment of compounds when possible. MestreNova software was used for data processing.

Low-resolution mass spectrometry (LRMS) using electrospray ionisation positive (ESI^+) and negative (ESI^-) modes was recorded on a Waters LCT Premier XE bench-top acceleration time-of-flight LC-MS system. High resolution mass spectrometry (HRMS) using electrospray ionisation positive (ESI^+) mode was recorded on a BioAccord LC-MS system. High-resolution values are given to 4 decimal places from the calculated (calcd.) molecular formula and values are within 5 ppm.

7.1.2 Compounds in Chapter 2

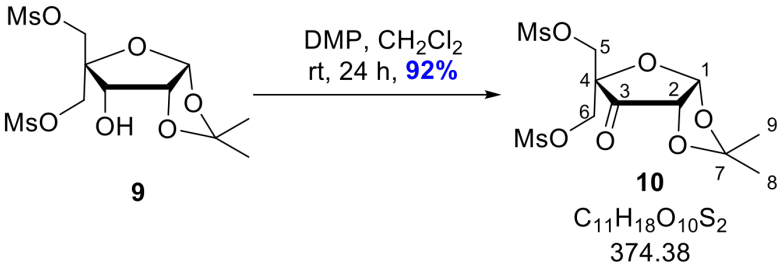
((3αR,6S,6αR)-6-hydroxy-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxole-5,5-diyl)bis(methylene) dimethanesulfonate (**9**)



3-O-benzyl-4-C-(methanesulfonyloxymethyl)-5-O-methanesulfonyl-1,2-O-isopropylidene-α-D-ribofuranose **8**, purchased from Carbosynth, (10.0 g, 21.4 mmol) and ammonium formate (10.0 g, 159 mmol) were dissolved in MeOH (250 mL). 20% $Pd(OH)_2$ on carbon (1.50 g, 2.41 mmol, 10 mol%) was added. The reaction was kept under argon, stirred, and refluxed to 60 °C overnight. The reaction was filtered through Celite and the solvent was removed under vacuum. The resulting solid was dissolved in EtOAc (100 mL), washed with a half-sat. solution of brine (2 × 60 mL), dried over Na_2SO_4, and the solvent was removed under vacuum to give **9** (7.15 g, 19.0 mmol, 89%) as a white powder.

R_f: 0.40 (EtOAc:Pet. Ether, 7:3); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.88 (d, J = 4.0 Hz, 1H, C(1)H), 4.75 (dd, J = 6.1, 4.0 Hz, 1H, C(3)H), 4.65 (d, J = 11.6 Hz, 1H, C(5 or 6)H_A), 4.49 (d, J = 11.6 Hz, 1H, C(5 or 6)H_B), 4.39 (m, 1H, C(2)H), 4.33 (d, J = 10.9 Hz, 1H, C(5 or 6)H_A), 4.29 (d, J = 10.9 Hz, 1H, C(5 or 6)H_B), 3.12 (s, 3H, CH_3-OMs), 3.08 (s, 3H, CH_3-OMs), 1.66 (s, 3H, C(8 or 9)H_3), 1.38 (s, 3H, C(8 or 9)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 114.2 (C-7), 104.9 (C-1), 84.5 (C-4), 79.3 (C-3), 72.8 (C-2), 69.5 (C-5 or 6), 68.8 (C-5 or 6), 38.2 (CH$_3$-OMs), 37.8 (CH$_3$-OMs), 26.4 (C-8 or 9), 26.2 (C-8 or 9); LRMS-ESI (m/z): $[M+Na]^+$ 399 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{11}H_{20}O_{10}S_2Na$, 399.0390; found, 399.0391.

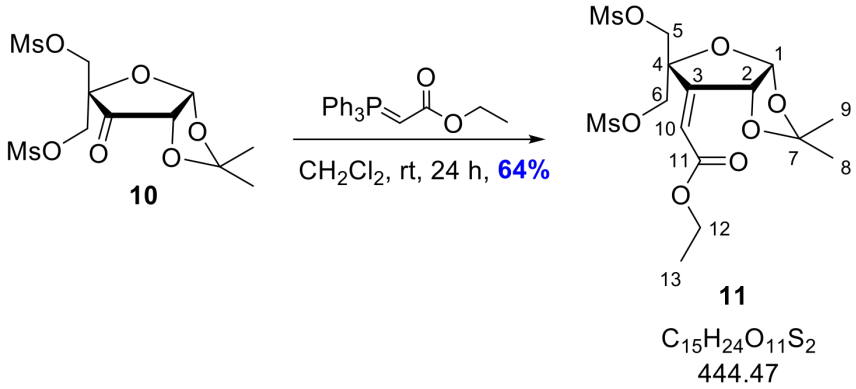
((3αR,6αS)-2,2-dimethyl-6-oxotetrahydrofuro[2,3-*d*][1,3]dioxole-5,5-diyl)bis(methylene) dimethanesulfonate (**10**)



Compound **9** (15.1 g, 40.2 mmol) and Dess-Martin periodinane (25.6 g, 60.4 mmol) were dissolved in CH_2Cl_2 (250 mL) and stirred at rt overnight. A solution of 10% $Na_2S_2O_3$ (w/v) in sat. $NaHCO_3$ (100 mL) was added slowly to the reaction and stirring was continued until gas evolution ceased. The organic layer was washed with sat. $NaHCO_3$ (2 × 100 mL), dried over Na_2SO_4, and the solvent was removed under reduced pressure to give **10** (13.8 g, 36.9 mmol, 92%) as a white foam.

R_f: 0.60 (EtOAc:Pet. Ether, 6:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.17 (d, J = 4.1 Hz, 1H, C(1)H), 4.52 (d, J = 4.2 Hz, 1H, C(2)H), 4.48 (d, J = 7.4 Hz, 1H, C(5 or 6)H_A), 4.46 (d, J = 8.1 Hz, 1H, C(5 or 6)H_B), 4.34 (d, J = 7.2 Hz, 1H, C(5 or 6)H_A), 4.31 (d, J = 6.6 Hz, 1H, C(5 or 6)H_B), 3.12 (s, 3H, CH_3-OMs), 3.03 (s, 3H, CH_3-OMs), 1.57 (d, J = 0.7 Hz, 3H, C(8 or 9)H_3), 1.40 (d, J = 0.7 Hz, 3H, C(8 or 9)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 205.0 (C-3), 115.6 (C-7), 103.0 (C-1), 84.3 (C-4), 76.7 (C-2), 69.5 (C-5 or 6), 68.9 (C-5 or 6), 38.3 (CH$_3$-OMs), 38.0 (CH$_3$-OMs), 27.2 (C-8 or 9), 26.8 (C-8 or 9); LRMS-ESI (m/z): $[M+Na]^+$ 397 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{11}H_{18}O_{10}S_2Na$, 397.0234; found, 397.0224.

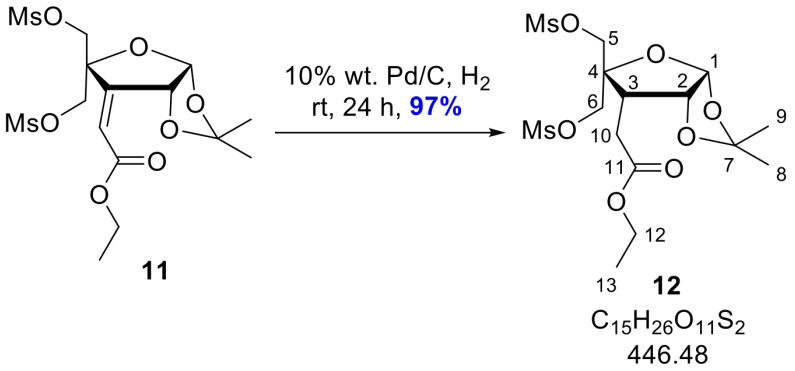
Ethyl 2-((3αR,6αR)-2,2-dimethyl-5,5-bis(((methylsulfonyl)oxy)methyl)dihydrofuro[2,3-d]-dioxol-6(5H)-ylidene)acetate (**11**)



Compound **10** (13.8 g, 36.9 mmol) and (carbethoxymethylene)triphenylphosphorane (18.0 g, 15.6 mmol) were dissolved in CH_2Cl_2 (125 mL) and stirred at rt overnight. The solvent was removed under vacuum and EtOH (200 mL) was added to the residue. The resulting precipitate was filtered, washed with EtOH, and dried to yield **11** (10.4 g, 23.4 mmol, 64%) as a white powder.

R_f: 0.50 (EtOAc:Pet. Ether, 6:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.07 (d, J = 1.4 Hz, 1H, C(10)H), 5.93 (d, J = 3.7 Hz, 1H, C(1)H), 5.80 (dd, J = 3.7, 1.4 Hz, 1H, C(2)H), 4.53 (d, J = 11.0 Hz, 1H, C(5 or 6)H_A), 4.37 (d, J = 10.9 Hz, 1H, C(5 or 6)H_B), 4.22-4.29 (m, 4H, C(5 or 6)H_2 and C(12)H_2), 3.11 (s, 3H, CH_3-OMs), 3.07 (s, 3H, CH_3-OMs), 1.60 (s, 3H, C(8 or 9)H_3), 1.41 (s, 3H, C(8 or 9)H_3), 1.33 (t, J = 7.1 Hz, 3H, C(13)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.0 (C-11), 151.2 (C-3), 121.4 (C-8), 114.5 (C-7), 105.6 (C-1), 85.0 (C-4), 78.6 (C-2), 70.0 (C-5 or 6), 69.5 (C-5 or 6), 61.5 (C-12), 38.2 (CH$_3$-OMs), 37.9 (CH$_3$-OMs), 27.2 (C-8 or 9), 26.4 (C-8 or 9), 14.2 (C-13); LRMS-ESI (m/z): $[M+Na]^+$ 467 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{15}H_{24}O_{11}S_2Na$, 467.0652; found, 467.0659.

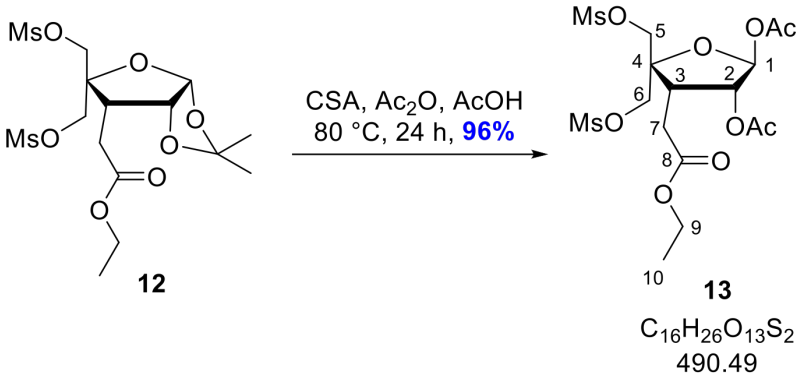
Ethyl 2-((3α*R*,6*S*,6α*R*)-2,2-dimethyl-5,5-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuro[2,3-*d*][1,3]dioxol-6-yl)acetate (**12**)



To a solution of compound **11** (2.60 g, 5.85 mmol) in EtOAc (100 mL) under argon was added 10% wt. palladium on activated carbon (311 mg, 5 mol%). The argon atmosphere was replaced by a hydrogen atmosphere through multiple rounds of vacuum and filling. The solution was stirred vigorously overnight under a hydrogen atmosphere. The solution was filtered through a pad of Celite and the solvent removed under reduced pressure to give **12** (2.54 g, 5.69 mmol, 97%) as a white powder. NB: ^{1}H NMR was used to ensure reaction was finished; no R_f change was observed.

R_f: 0.60 (EtOAc:Pet. Ether, 6:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.85 (d, J = 3.8 Hz, 1H, C(1)H), 4.88 (dd, J = 5.2, 3.8 Hz, 1H, C(2)H), 4.59 (d, J = 10.8 Hz, 1H, C(5 or 6)H_A), 4.31 (m, 2H, C(5 or 6)H_2), 4.23 (d, J = 10.8 Hz, 1H, C(5 or 6)H_B), 4.15-4.19 (m, 2H, C(12)H_2), 3.12 (s, 3H, CH_3-OMs), 3.07 (s, 3H, CH_3-OMs), 2.79-2.84 (m, 1H, C(3)H), 2.71 (dd, J = 16.9, 9.6 Hz, 1H, C(10)H_A), 2.57 (dd, J = 16.9, 5.5 Hz, 1H, C(10)H_B), 1.61 (s, 3H, C(8 or 9)H_3), 1.30 (s, 3H, C(8 or 9)H_3), 1.28 (t, J = 7.2 Hz, 3H, C(13)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.3 (*C*-11), 112.9 (*C*-7), 105.3 (*C*-1), 84.1 (*C*-4), 81.7 (*C*-2), 70.1 (*C*-5 or 6), 68.2 (*C*-5 or 6), 61.3 (*C*-12), 42.6 (*C*-3), 38.1 (*C*H$_3$-OMs), 37.7 (*C*H$_3$-OMs), 28.6 (*C*-10), 26.3 (*C*-8 or 9), 25.5 (*C*-8 or 9), 14.2 (*C*-13); LRMS-ESI (*m/z*): $[M+Na]^+$ 469 (100%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{15}H_{26}O_{11}S_2Na$, 469.0809; found, 469.0815.

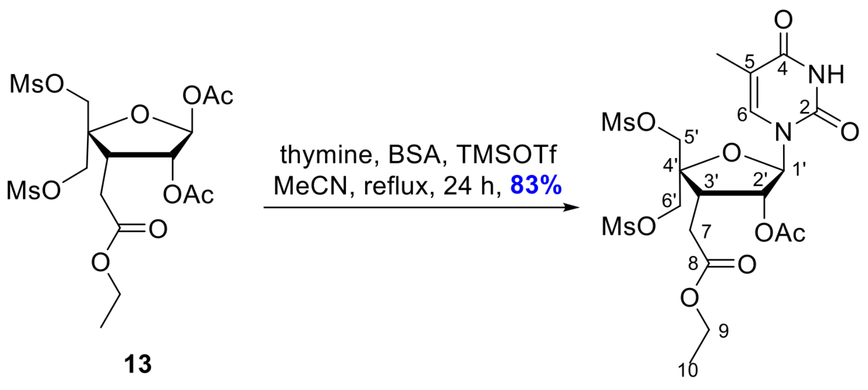
(3*R*,4*S*)-4-(2-ethoxy-2-oxoethyl)-5,5-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate (**13**)



Compound **12** (1.59 g, 3.56 mmol) was dissolved in a solution of AcOH (20 mL) and acetic anhydride (15 mL). Camphorsulfonic acid (CSA) was added (41 mg, 0.178 mmol) and the reaction was stirred at 80 °C for 90 min. Another measure of CSA (41 mg, 0.178 mmol) was added and the reaction continued stirring at 80 °C for another 90 min. Another measure of CSA (41 mg, 0.178 mmol) was added and after 90 min at 80 °C the reaction was complete. The solvent was removed under reduced pressure and the residue diluted in EtOAc (100 mL). The organic layer was washed with sat. $NaHCO_3$ (2 × 50 mL), brine (50 mL) and dried over Na_2SO_4. The solvent was removed under reduced pressure to give **13** (1.68 g, 3.43 mmol, 96%) as an amber sticky oil and used without further purification.

R_f: 0.60 (EtOAc:Pet. Ether, 6:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.12 (s, 1H, C(1)*H*), 5.33 (d, J = 5.0 Hz, 1H, C(2)*H*), 5.02 (d, J = 3.7 Hz, 2H, C(5 or 6)H_2), 4.41 (s, 2H, C(5 or 6)H_2), 4.14-4.18 (m, 2H, C(9)H_2), 3.12 (s, 1H, C(3)*H*), 3.09 (s, 3H, CH_3-OMs), 3.07 (s, 3H, CH_3-OMs), 2.63 (t, J = 7.5 Hz, 2H, C(7)H_2), 2.15 (s, 3H, CH_3-OAc), 2.13 (s, 3H, CH_3-OAc), 1.28 (t, J = 7.2 Hz, 3H, C(10)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) 170.9 (*C*-8), 169.5 (*C*OCH$_3$-OAc),169.3 (*C*OCH$_3$-OAc), 97.8 (*C*-1), 84.6 (*C*-4), 77.8 (*C*-2), 71.4 (*C*-5 or 6), 66.9 (*C*-5 or 6), 61.6 (*C*-9), 40.5 (*C*-3), 38.1 (*C*H$_3$-OMs), 37.7 (*C*H$_3$-OMs), 28.6 (*C*-7), 21.2 (*C*H$_3$-OAc), 20.8 (*C*H$_3$-OAc), 14.2 (*C*-10); LRMS-ESI (*m/z*): $[M+Na]^+$ 513 (100%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{16}H_{26}O_{13}S_2Na$, 513.0707; found, 513.0728.

Ethyl 2-((3*S*,4*R*,5*R*)-4-acetoxy-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,2-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-3-yl)acetate (**14**)

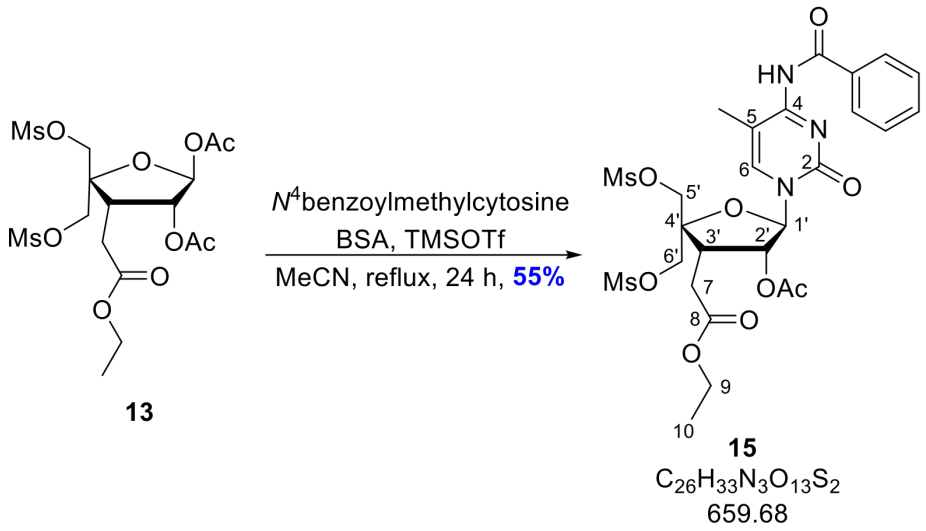


Compound **13** (2.37 g, 4.83 mmol) and thymine (762 mg, 6.04 mmol) were co-evaporated with anhydrous MeCN (2 × 15 mL). It was re-dissolved in anhydrous MeCN (15 mL) and bis(trimethylsilyl)acetamide (BSA) (3.25 mL, 13.28 mmol) was added. The solution was heated to reflux for 1 hour. The reaction was cooled to rt and trimethylsilyl trifluoromethane-sulfonate (TMSOTf) (1.14 mL, 6.28 mmol) was added. The reaction was heated to reflux overnight resulting in a dark red solution. CH_2Cl_2 (50 mL) was added to dilute the solution and a solution of sat. $NaHCO_3/H_2O$ (1:1, v/v) (60 mL) was added. The solution was stirred for 10 min or until gas evolution ceased. The organic layer was washed with sat. $NaHCO_3$ (2 × 50 mL), brine (50 mL), dried over Na_2SO_4, and the solvent was evaporated under reduced pressure to give an amber oil. This was purified via column chromatography (0-5% MeOH/CH_2Cl_2) to give **14** (2.24 g, 4.02 mmol, 83%) as a light yellow foam.

R_f: 0.67 (MeOH:CH_2Cl_2, 95:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.53 (s, 1H, NH), 7.04 (q, J = 1.2 Hz, 1H, C(6)H), 5.58 (dd, J = 7.6, 1.7 Hz, 1H, C(2')H), 5.51 (d, J = 1.7 Hz, 1H, C(1')H), 4.46 (s, 2H, C(5' or 6')H_2), 4.39 (d, J = 11.0 Hz, 1H, C(5' or 6')H_A), 4.31 (d, J = 11.0 Hz, 1H, C(5' or 6')H_B), 4.16 (q, J = 7.1 Hz, 2H, C(9)H_2), 3.49-3.57 (m, 1H, C(3')H), 3.11 (s, 3H, CH_3-OMs), 3.08 (s, 3H, CH_3-OMs), 2.58-2.65 (m, 2H, C(7)H_2), 2.15 (s, 3H, CH_3-OAc), 1.93 (d, J = 1.2 Hz, 3H, CH_3-thymine), 1.26 (t, J = 7.2 Hz, 3H,

C(10)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.8 (C-8), 170.3 (COCH$_3$-OAc), 163.6 (C-4), 150.0 (C-2), 138.3 (C-6), 111.7 (C-5), 95.6 (C-4'), 85.5 (C-1'), 78.5 (C-2'), 70.0 (C-5' or 6'), 67.3 (C-5' or 6'), 61.5 (C-9), 41.4 (C-3'), 38.0 (CH$_3$-OMs), 37.7 (CH$_3$-OMs), 29.0 (C-7), 20.8 (CH$_3$-OAc), 14.3 (C-10), 12.4 (CH$_3$-thymine); LRMS-ESI (m/z): $[M+H]^+$ 557 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{19}H_{29}N_2O_{13}S_2$, 557.1106; found, 557.1112; $[M+Na]^+$ calcd. for $C_{19}H_{28}N_2O_{13}S_2Na$, 579.0925; found, 579.0928.

Ethyl-2-((3*S*,4*R*,5*R*)-4-acetoxy-5-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-2,2-bis (((methylsulfonyl)oxy)methyl)tetrahydrofuran-3-yl)acetate (**15**)

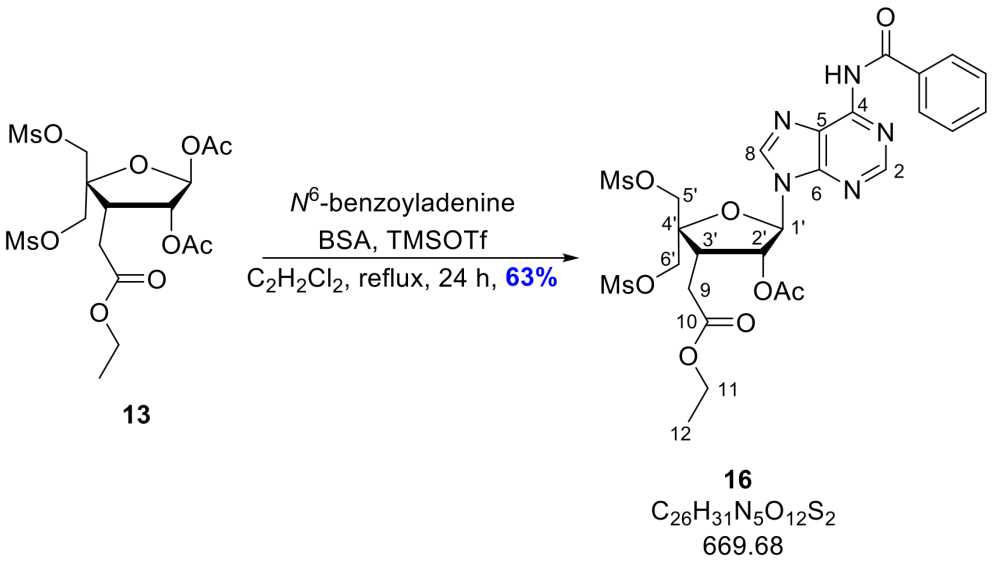


Compound **13** (3.25 g, 6.63 mmol), N^4-benzoylmethylcytosine (2.50 g, 9.94 mmol), and BSA (4.81 mL, 17.9 mmol) were suspended in anhydrous MeCN (70 mL) and refluxed for 1 hour. The reaction was left to cool to rt and TMSOTf (1.58 mL, 8.62 mmol) was added. The reaction mixture was further refluxed overnight. After cooling to rt, the reaction mixture was diluted with EtOAc (50 mL) and sat. $NaHCO_3$ (50 mL) was added slowly (generates gas). The organic layer was washed with sat. $NaHCO_3$ (2 × 70 mL), brine (70 mL), dried over Na_2SO_4, and reduced to give an amber oil. This was purified by column chromatography (0-10% MeOH/CH_2Cl_2) to yield **15** (2.39 g, 3.62 mmol, 55%) as a pale yellow foam.

R_f: 0.67 (CH_2Cl_2:MeOH, 19:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.32 – 8.29 (m, 1H, H-Ar), 7.57 – 7.49 (m, 2H, H-Ar), 7.49 – 7.39 (m, 2H, H-Ar), 7.21 (s, 1H, C(6)H), 5.61 (dd, J

= 7.5, 1.7 Hz, 1H, C(2')H), 5.55 (d, J = 1.7 Hz, 1H, C(1')H), 4.51 – 4.43 (m, 2H, C(5' or 6')H_2), 4.41 (d, J = 11.0 Hz, 1H, C(5' or 6')H_A), 4.33 (d, J = 11.0 Hz, 1H, C(5' or 6')H_B), 4.16 (q, J = 7.2 Hz, 2H, C(9)H_2), 3.56 (ddd, J = 9.3, 7.5, 6.6 Hz, 1H, C(3')H), 3.11 (s, 3H, CH_3-OMs), 3.09 (s, 3H, CH_3-OMs),), 2.67 (dd, J = 16.8, 9.3 Hz, 1H, C(7)H_A), 2.57 (dd, J = 16.7, 6.5 Hz, 1H, C(7)H_B), 2.16 (s, 3H, CH_3-OAc), 2.11 (d, J = 1.2 Hz, 3H, CH_3-^{Me}cytosine), 1.26 (t, J = 7.1 Hz, 3H, C(10)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.7 (C-8), 170.2 (COCH$_3$-OAc), 163.6 (C-4), 150.0 (C-2), 139.1 (C-6), 132.8 (C-Ar), 130.1 (C-Ar), 128.3 (C-Ar), 112.6 (C-5), 95.9 (C-4'), 85.7 (C-1'), 78.4 (C-2'), 69.8 (C-5' or 6'), 67.4 (C-5' or 6'), 61.5 (C-9), 41.2 (C-3'), 38.0 (CH$_3$-OMs), 37.7 (CH$_3$-OMs), 29.1 (C-7), 20.8 (CH$_3$-OAc), 14.3 (C-10), 13.5 (CH$_3$-^{Me}cytosine); LRMS-ESI (m/z): $[M]^+$ 659 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{26}H_{33}N_3O_{13}S_2Na$, 682.1347; found, 682.1371.

Ethyl-2-((3*S*,4*R*,5*R*)-4-acetoxy-5-(6-benzamido-9*H*-purin-9-yl)-2,2-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-3-yl)acetate (**16**)

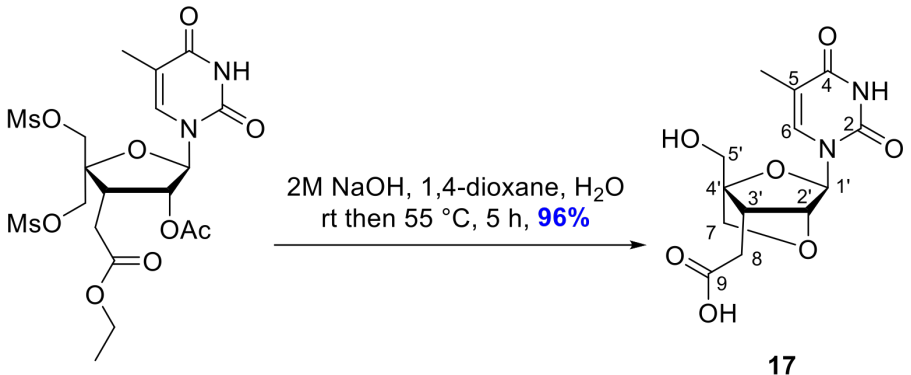


Compound **13** (1.30 g, 2.65 mmol) and N^6-benzoyladenine (790 mg, 3.30 mmol) were dissolved in anhydrous CH_2Cl_2 (25 mL) and BSA (1.75 mL, 7.16 mmol) was added. The reaction mixture was stirred at reflux for one hour and then allowed to cool to rt. TMSOTf (1.12 mL, 6.10 mmol) was added and the mixture was refluxed overnight. The solution was cooled to rt, diluted with CH_2Cl_2 (20 mL) and sat. $NaHCO_3$:H_2O (1:1, v/v) (20 mL)

was added and stirred vigorously until gas evolution ceased. The organic layer was washed with $NaHCO_3$ (2 × 50 mL) and brine (50 mL). The organic layers were combined, dried over Na_2SO_4, and reduced under vacuum to yield a crude amber oil. This was purified via column chromatography (0-60% EtOAc/Pet. Ether) to yield **16** (1.12 g, 1.67 mmol, 63%) as a yellow foam.

R_f: 0.30 (CH_2Cl_2:MeOH, 8:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.83 (s, 1H, C(2)H), 8.09 (s, 1H, C(8)H), 8.01-8.04 (m, 2H, H-Ar), 7.59-7.63 (m, 1H, H-Ar), 7.47-7.55 (m, 2H, H-Ar), 6.12 (d, J = 1.1 Hz, 1H, C(1')H), 5.89 (dd, J = 6.7, 1.1 Hz, 1H, C(2')H), 4.56 (m, 2H, C(5' or 6')H_2), 4.49 (d, J = 10.8 Hz, 1H, C(5' or 6')H_A), 4.40 (d, J = 10.8 Hz, C(5' or 6')H_B), 4.19 (d, J = 7.1 Hz, 1H, C(11)H_A), 4.15 (d, J = 7.1 Hz, 1H, C(11)H_B), 4.11 (m, 1H, C(3')H), 3.12 (s, 3H, CH_3-OMs), 2.96 (s, 3H, CH_3-OMs), 2.73 (dd, J = 7.8 Hz, 5.1 Hz, 2H, C(9)H_2) 2.20 (s, 3H, CH_3-OAc), 1.27 (t, J = 5.4 Hz, 3H, C(12)H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.7 (C-9), 170.2 (COCH$_3$-OAc), 152.9 (C-2), 150.1 (C-6), 150.0 (C-4), 142.6 (C-8), 133.0 (C-Ar), 129.1 (C-Ar), 128.0 (C-Ar), 124.9 (C-5), 90.8 (C-4'), 86.0 (C-1'), 79.0 (C-2'), 69.6 (C-5' or 6'), 66.9 (C-5' or 6'), 61.6 (C-11), 41.7 (C-3'), 38.0 (CH$_3$-OMs), 37.6 (CH$_3$-OMs), 28.9 (C-9), 20.8 (CH$_3$-OAc), 14.3 (C-12); LRMS-ESI (m/z): $[M]^+$ 669 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{26}H_{31}N_5O_{12}S_2Na$, 692.1303; found, 692.1324.

2-((1*S*,3*R*,4*R*,7*S*)-1-(hydroxymethyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**17**)

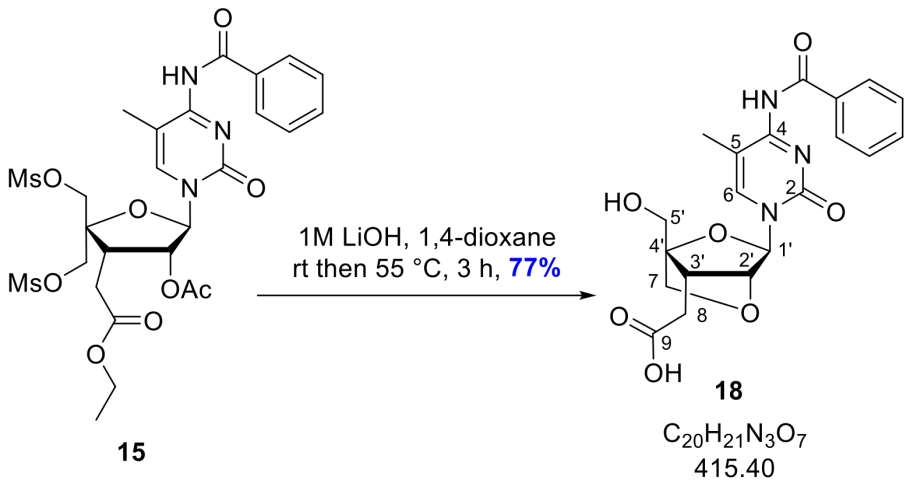


Compound **14** (3.00 g, 5.39 mmol) was dissolved in 1,4-dioxane/H_2O (1:1, v/v) (20 mL) and

2 M NaOH (20 mL) was added. The solution was stirred at rt for 4 h and then heated to 55 °C for 1 hour. The reaction was cooled to rt and acidified with 1 M HCl (40 mL) and CH_2Cl_2 (20 mL) was added and stirred vigorously. The aqueous layer was washed with 20% isopropanol in CH_2Cl_2 (10 × 50 mL) until the aqueous layer was clear of product via TLC. The solvent was removed under reduced pressure to yield **17** (1.62 g, 5.19 mmol, 96%) as a pale yellow solid.

R_f: 0.35 (CH_2Cl_2:MeOH:Et_3N, 60:38:2); ^{1}H NMR (400 MHz, D2O) δ 7.69 (t, J = 1.3 Hz, 1H, C(6)H), 5.69 (s, 1H, C(1')H), 4.67 (s, 1H, C(2')H), 4.03 (d, J = 2.1 Hz, 2H, C(7)H_2), 3.90 (q, J = 8.8 Hz, 2H, C(5')H_2), 2.50-2.53 (m, 2H, C(8)H_2), 2.43 (q, J = 7.5 Hz, 1H, C(3')H), 1.91 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, D2O) δ 179.5 (C-9), 166.7 (C-4), 151.2 (C-2), 136.6 (C-6), 110.4 (C-5), 91.1 (C-4'), 86.7 (C-1'), 80.7 (C-2'), 71.4 (C-5'), 57.3 (C-7), 40.7 (C-3'), 31.9 (C-8), 11.7 (CH$_3$-thymine); LRMS-ESI (m/z): $[M]^+$ 312 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{13}H_{17}N_2O_7$, 313.1030; found, 313.1029; $[M+Na]^+$ calcd. for $C_{13}H_{16}N_2O_7Na$, 335.0850; found, 335.0851.

2-((1*S*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**18**)

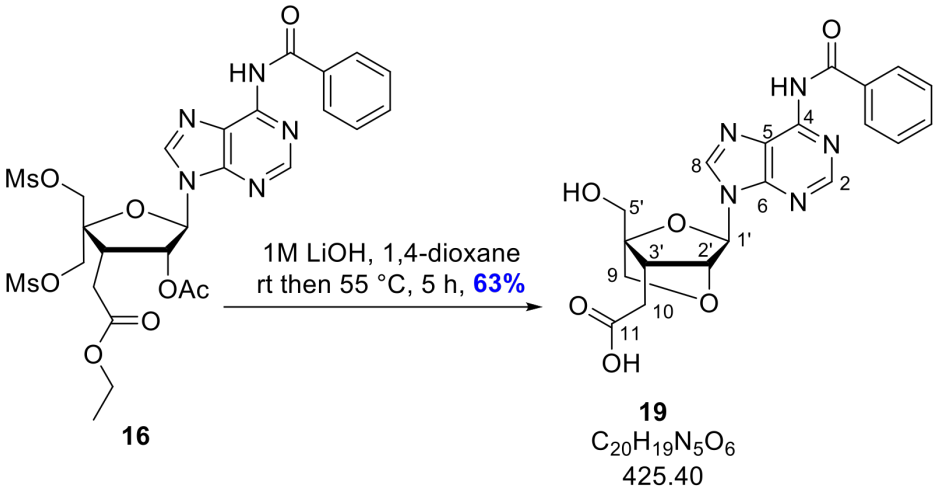


Compound **15** (2.34 g, 3.55 mmol) was dissolved in 1,4-dioxane (30 mL) and 1 M LiOH (12 mL, 11.7 mmol) was added. The reaction was left to stir at rt for 2 h and then heated to 55 °C for a further hour until the reaction was completed as determined by LCMS. The reaction

mixture was reduced, re-dissolved in CH_2Cl_2 and water (50 mL each), and the aqueous layer was washed with CH_2Cl_2 (2 × 50 mL). The aqueous layer was acidified with 1 M HCl and the product was then extracted from the aqueous layer using 15% isopropanol in CH_2Cl_2 until the aqueous layer was clear by TLC. The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to give **18** (1.13 g, 2.72 mmol, 77%) as a light yellow foam.

R_f: 0.50 (CH_2Cl_2:MeOH,Et_3N, 60:38:2); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.21 – 8.11 (m, 2H, *H*-Ar), 7.96 (d, J = 1.2 Hz, 1H, C(6)*H*), 7.64 – 7.55 (m, 1H, *H*-Ar), 7.50 (td, J = 7.7, 1.5 Hz, 2H, *H*-Ar), 5.51 (s, 1H, C(1')*H*), 4.43 (s, 1H, C(2')*H*), 3.82 (d, J = 13.1 Hz, 1H, C(5')H_A), 3.76 (d, J = 13.0 Hz, 1H, C(5')H_B), 3.75 (d, J = 8.4 Hz, 1H, C(7)H_A), 3.62 (d, J = 8.5 Hz, 1H, C(7)H_B), 3.25 (s, 1H, C(3')*H*), 2.34 – 2.15 (m, 2H, C(8)H_2), 2.03 (d, J = 1.0 Hz, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 177.9 (*C*-9), 172.9 (*C*O-Bz), 159.1 (*C*-4), 147.2 (*C*-2), 139.0 (*C*-6), 132.6 (*C*-Ar) 129.3 (*C*-Ar), 128.3 (*C*-Ar), 108.9 (*C*-5), 91.4 (*C*-4'), 87.0 (*C*-1'), 79.8 (*C*-2'), 71.1 (*C*-7), 56.7 (*C*-5'), 36.9 (*C*-3'), 29.9 (*C*-8), 13.7 (*C*H_3-^{Me}cytosine); LRMS-ESI (*m/z*): $[M]^+$ 415 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{20}H_{22}N_3O_7$, 416.1452; found, 416.1428; $[M+Na]^+$ calcd. for $C_{20}H_{21}N_3O_7Na$, 438.1272; found, 438.1295.

2-((1*S*,3*R*,4*R*,7*S*)-3-(6-benzamido-9*H*-purin-9-yl)-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**19**)

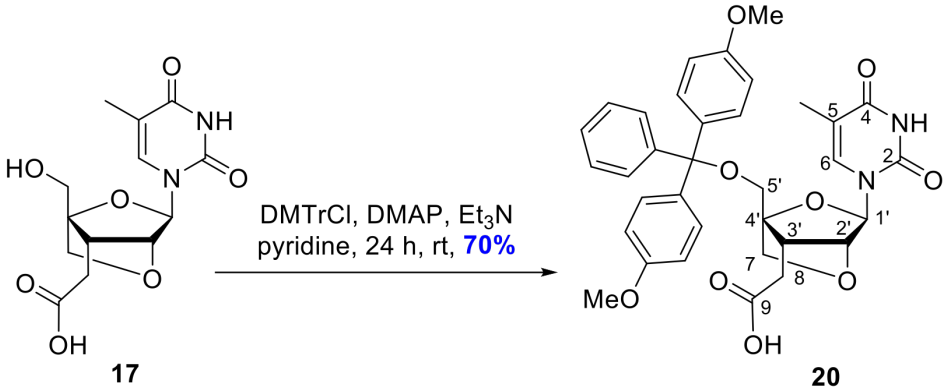


Compound **16** (1.12 g, 1.67 mmol) was dissolved in 1,4-dioxane (20 mL) and 1 M LiOH (5.50

mL, 5.50 mmol) and stirred at rt for 3 h. The reaction was heated to 55 °C for a further two h until the reaction was complete as determined by LCMS. The solvent was removed under reduced pressure, re-dissolved in CH_2Cl_2 and water (20 mL each), and the aqueous layer was washed with CH_2Cl_2 (2 × 20 mL). The aqueous layer was acidified with 1 M HCl and the product was then extracted from the aqueous layer using 15% isopropanol in CH_2Cl_2 until the aqueous layer was clear by TLC. The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to yield **19** (450 mg, 1.06 mmol, 63%) as a light yellow foam.

R_f: 0.40 (CH_2Cl_2:MeOH, 8:2); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.76 (s, 1H, C(2)*H*), 8.53 (s, 1H, C(8)*H*), 8.15 – 8.02 (m, 2H, *H*-Ar), 7.68 – 7.57 (m, 2H, *H*-Ar), 7.61 – 7.46 (m, 2H, *H*-Ar), 6.08 (s, 1H, C(1')*H*), 4.73 (s, 1H, C(2')*H*), 3.86 (d, J = 8.6 Hz, 1H, C(5')H_A), 3.84 (s, 2H, C(9)H_2), 3.77 (d, J = 8.5 Hz, 1H, C(5')H_B), 2.57 (dd, J = 9.7, 4.2 Hz, 1H, C(3')*H*), 2.50 (dd, J = 17.1, 4.0 Hz, 1H, C(10)H_A), 2.33 (dd, J = 17.1, 9.7 Hz, 1H, C(10)H_B); ^{13}C NMR (101 MHz, DMSO-d_6) δ 172.9 (*C*-11), 149.6 (*C*-2), 141.0 (*C*-8), 132.4 (*C*-Ar), 128.5 (*C*-Ar), 128.4 (*C*-Ar), 123.9 (*C*-5), 90.3 (*C*-4'), 85.5 (*C*-1'), 80.0 (*C*-2'), 71.5 (*C*-5'), 57.1 (*C*-9), 40.8 (*C*-3'), 28.7 (*C*-10); LRMS-ESI (*m/z*): $[M]^+$ 425 (100%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{20}H_{19}N_5O_6Na$, 448.1228; found, 448.1217.

2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**20**)

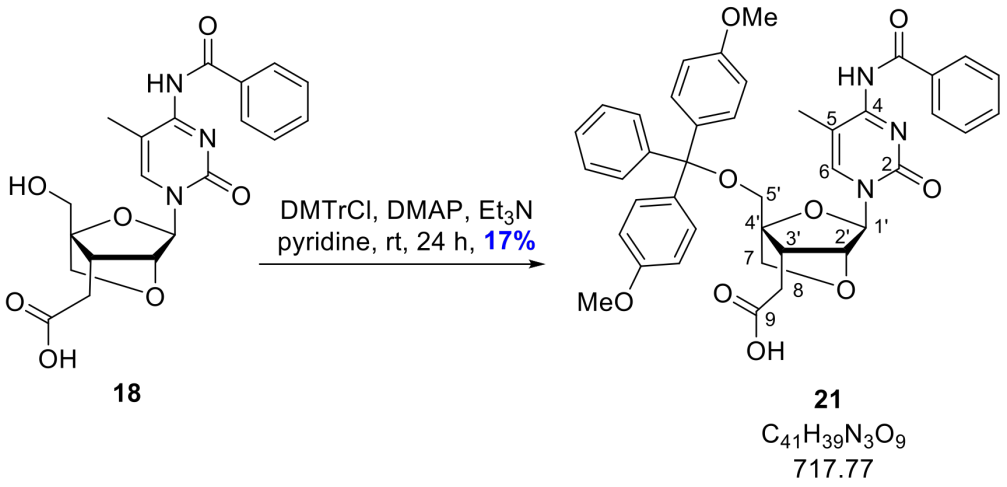


Compound **17** (200 mg, 0.641 mmol) was co-evaporated with anhydrous pyridine (3 × 10

mL), dissolved in anhydrous pyridine (9 mL) and anhydrous Et_3N (125 µL, 0.897 mmol) and 4Å activated molecular sieves were added. The solution was stirred at rt for 15 min then 4-dimethylaminopyridine (DMAP) (188 mg, 1.54 mmol) and 4,4-dimethoxytritylchloride (DMTrCl) (869 mg, 2.56 mmol) were added and stirred at rt overnight. The solvent was removed under reduced pressure and the residue was purified via column chromatography (0-30% MeOH/EtOAc with 2% pyridine) to give **20** (274 mg, 0.446 mmol, 70%) as a yellow sticky powder.

R_f: 0.70 (EtOAc:MeOH:pyridine, 70:28:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (s, 1H, C(6)H), 7.30-7.34 (m, 10H, H-Ar), 6.84 (dd, J = 9.0, 3.9 Hz, 4H, H-Ar), 5.69 (s, 1H, C(1')H), 4.67 (s, 1H, C(2')H), 3.79 (d, J = 1.6 Hz, 6H, 2 × CH_3-OMe), 3.76 (d, J = 8.4 Hz, 1H, C(5')H_A), 3.65 (d, J = 11.0 Hz, 1H, C(7)H_A), 3.60 (d, J = 8.4 Hz, 1H, C(5')H_B), 3.28 (d, J = 11.1 Hz, 1H, C(7)H_B), 2.58 (dd, J = 9.7, 4.1 Hz, 1H, C(3')H), 2.31 (dd, J = 16.6, 9.8 Hz, 1H, C(8)H_A), 2.16 (dd, J = 16.4, 4.2 Hz, 1H, C(8)H_B), 1.73 (s, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.3 (C-9), 164.4 (C-4), 158.8 (C-2), 137.1 (C-6), 130.1 (C-Ar), 129.3 (C-Ar), 128.2 (C-Ar), 127.9 (C-Ar), 127.2 (C-Ar), 113.5 (C-Ar), 110.3 (C-5), 89.8 (C-4'), 87.5 (C-1'), 86.8 (C-DMTr), 80.5 (C-2'), 71.9 (C-5'), 59.0 (C-7), 55.3 (CH$_3$-OMe), 40.2 (C-3'), 29.1 (C-8), 12.7 (CH$_3$-thymine); LRMS-ESI (m/z): $[M]^+$ 614 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{34}H_{34}N_2O_9Na$, 637.2157; found, 637.2158.

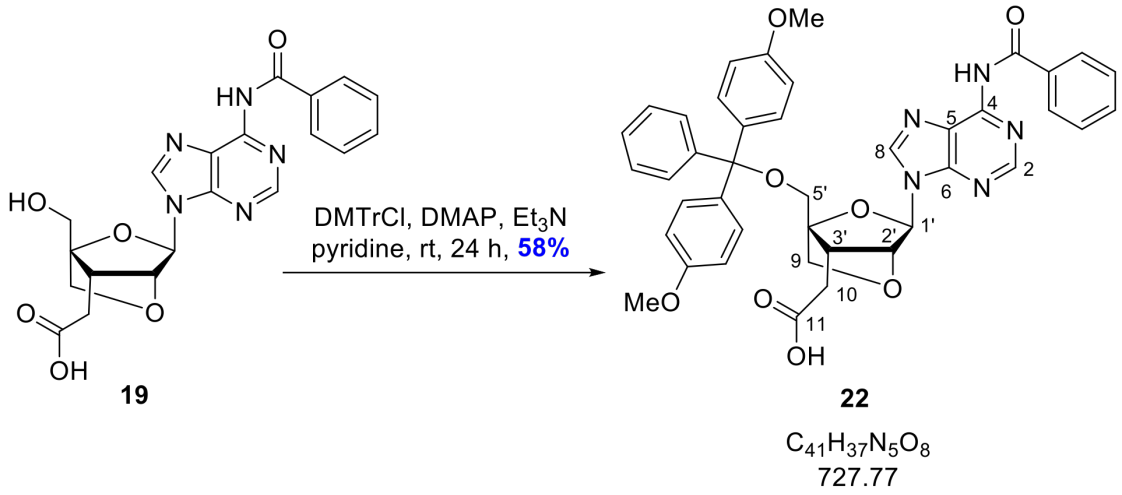
2-((1R,3R,4R,7S)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2H)-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**21**)



Compound **18** (56 mg, 0.135 mmol) was dissolved in anhydrous pyridine (5 mL) and anhydrous Et_3N (28 µL, 0.201 mmol) and 4Å activated molecular sieves were added. The solution was stirred at rt for 15 min then DMAP (25 mg, 0.203 mmol) and DMTrCl (114 mg, 0.337 mmol) were added and stirred at rt overnight. The solvent was removed under reduced pressure and the residue was purified via column chromatography (0-30% MeOH/EtOAc with 2% pyridine) to yield **21** (16 mg, 0.022 mmol, 17%) as a yellow sticky powder.

R_f: 0.75 (EtOAc:MeOH:pyridine, 70:28:2); ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.15 (s, 1H, N*H*), 7.89 (s, 1H, C(6)*H*), 7.60 (t, J = 7.3 Hz, 1H, *H*-Ar(Bz)), 7.51 (t, J = 7.6 Hz, 2H, *H*-Ar(Bz)), 7.47 – 7.43 (m, 2H, *H*-Ar(DMTr)), 7.37 – 7.29 (m, 7H, *H*-Ar(DMTr)), 7.28 – 7.25 (m, 1H, *H*-Ar(Bz)), 6.96 – 6.89 (m, 4H, *H*-DMTr), 5.57 (s, 1H, C(1')*H*), 4.53 (s, 1H, C(2')*H*), 3.75 (s, 6H, 2 x C*H*$_3$-OMe), 3.67 (q, J = 8.6 Hz, 2H, C(7)H_2), 3.54 (d, J = 11.3 Hz, 1H, C(5')H_A), 3.36 (d, J = 11.3 Hz, 1H, C(5')H_B), 2.47 (dd, J = 9.1, 4.4 Hz, 1H, C(3')*H*), 2.22 (dd, J = 17.0, 9.0 Hz, 1H, C(8)H_A), 2.03 (dd, J = 16.9, 4.6 Hz, 1H, C(8)H_B), 1.83 (s, 3H, C*H*$_3$-^{Me}cytosine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 170.3 (*C*-9), 158.2 (*C*-4), 144.6 (*C*-2), 135.1 (*C*-6), 132.6 (*C*-Ar), 129.7 (*C*-Ar), 129.2 (*C*-5), 128.5 (*C*-Ar), 128.4 (*C*-Ar), 128.0 (*C*-Ar), 127.6 (*C*-Ar), 126.9 (*C*-Ar), 113.4 (*C*-Ar), 89.6 (*C*-4'), 87.3 (*C*-1'), 85.9 (*C*-DMTr), 79.5 (*C*-2'), 71.3 (*C*-7), 58.5 (*C*-5'), 55.1 (*C*H$_3$-OMe), 40.1 (*C*-3'), 28.5 (*C*-8), 14.1 (*C*H$_3$-^{Me}cytosine); LRMS-ESI (*m/z*): $[M]^+$ 717 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{41}H_{40}N_3O_9$, 718.2759; found, 718.2755; $[M+Na]^+$ calcd. for $C_{41}H_{39}N_3O_9Na$, 740.2579; found, 740.2558.

2-((1*R*,3*R*,4*R*,7*S*)-3-(6-benzamido-9*H*-purin-9-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**22**)

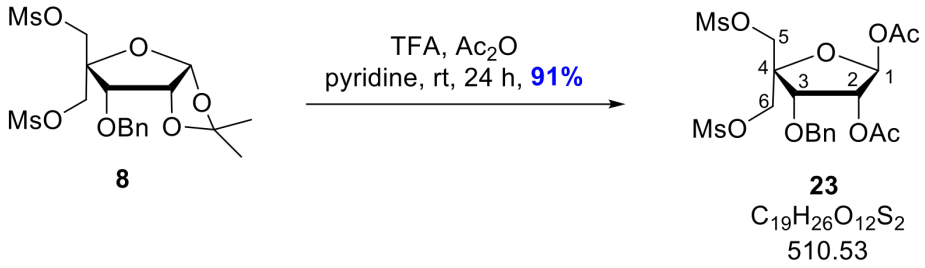


Compound **19** (264 mg, 0.621 mmol) was dissolved in anhydrous pyridine (5 mL) and anhydrous Et_3N (123 µL, 0.882 mmol) and 4Å activated molecular sieves were added. The solution was stirred at rt for 15 min then DMAP (99 mg, 0.807 mmol) and DMTrCl (842 mg, 2.45 mmol) were added and stirred at rt overnight. The solvent was removed under reduced pressure and the residue was purified via column chromatography (0-30% MeOH/EtOAc with 2% pyridine) to give **22** (263 mg, 0.361 mmol, 58%) as a yellow sticky powder.

R_f: 0.52 (EtOAc:MeOH:pyridine, 70:28:2); ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 8.78 (s, 1H, C(2)*H*), 8.50 (s, 1H, C(8)*H*), 8.09 – 8.02 (m, 2H, *H*-Ar(Bz)), 7.69 – 7.61 (m, 1H, *H*-Ar(Bz)), 7.61 – 7.47 (m, 2H, *H*-Ar(Bz)), 7.34 – 7.19 (m, 10H, *H*-Ar(DMTr)), 6.89 (dt, J = 9.3, 2.1 Hz, 4H, *H*-Ar(DMTr)), 6.14 (s, 1H, C(1')*H*), 4.81 (s, 1H, C(2')*H*), 3.84 (d, J = 8.6 Hz, 1H, C(5')H_A), 3.75 (d, J = 8.3 Hz, 1H, C(5')H_B), 3.73 (s, 6H, 2 × CH_3-OMe), 3.44 (m, 2H, C(9)H_2), 2.74 (dd, J = 9.8, 4.0 Hz, 1H, C(3')*H*), 2.24-2.16 (m, 1H, C(10)H_A), 2.08 (dd, J = 16.8, 4.2, 1H, C(10)H_B); ^{13}C NMR (101 MHz, DMSO-d$_6$) δ 173.2 (*C*-11), 165.7 (*C*OBz), 158.2 (*C*-5), 151.4 (*C*-2), 150.4 (*C*-6), 144.6 (*C*-4), 140.8 (*C*-8), 135.2 (*C*-Ar), 132.5 (*C*-Ar), 129.7 (*C*-Ar), 128.5 (*C*-Ar), 128.5 (*C*-Ar), 127.9 (*C*-Ar), 127.6 (*C*-Ar), 126.8 (*C*-Ar), 125.7 (*C*-Ar), 113.3 (*C*-Ar), 88.8 (*C*-4'), 85.7 (*C*-1'), 80.0 (*C*-2'), 71.8 (*C*-9), 59.5 (*C*-5'), 55.0 (*C*H$_3$-OMe), 40.2 (*C*-3'), 31.1 (*C*-10); LRMS-ESI (m/z): [M]$^+$ 727 (100%); HRMS-ESI

(*m/z*): $[M+Na]^+$ calcd. for $C_{41}H_{37}N_5O_8Na$, 750.2534; found, 750.2532.

(2*S*,3*R*,4*S*)-4-(benzyloxy)-5,5-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate (**23**)

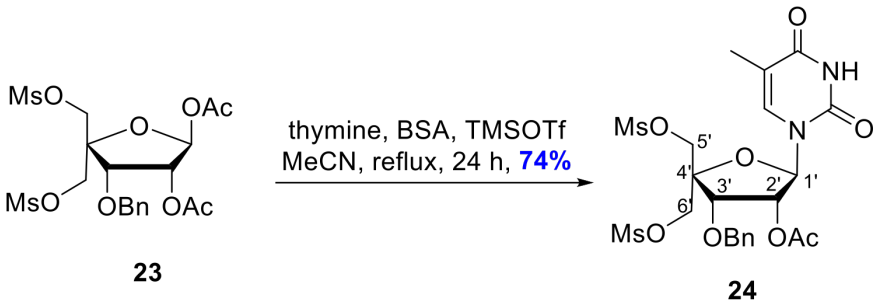


A solution of **8** (5.00 g, 10.7 mmol) in 80% aqueous TFA (25 mL) was stirred at rt for 1 h. The solvent was removed under reduced pressure and the residue dissolved in CH_2Cl_2 (50 mL) and washed with sat. $NaHCO_3$ (2 × 50 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was co-evaporated with anhydrous pyridine (2 × 25 mL), dissolved in anhydrous pyridine (25 mL), and treated with acetic anhydride (3.75 mL, 39.7 mmol) overnight. The reaction was quenched by the addition of sat. $NaHCO_3$ (50 mL) and washed with EtOAc (2 × 50 mL). The organic layers were combined, washed with brine (50 mL), dried over Na_2SO_4, and concentrated under reduced pressure to give **23** as a sticky yellow residue (4.97 g, 9.73 mmol, 91%).

R_f: 0.40 (EtOAc:Pet. Ether, 1:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.31- 7.39 (m, 5H, *H*-Ar), 6.17 (s, 1H, C(1)*H*), 5.38 (d, J = 4.8 Hz, 1H, C(2)*H*), 4.62 (d, J = 11.1 Hz, 1H, CH_A-OBn), 4.52 (d, J = 9.0 Hz, 1H, CH_B-OBn), 4.50 (d, J = 9.5 Hz, 1H, C(5 or 6)H_A), 4.42 (d, J = 4.8 Hz, 1H, C(3)*H*), 4.38 (d, J = 11.5 Hz, 1H, C(5 or 6)H_B), 4.30 (d, J = 10.6 Hz, 1H, C(5 or 6)H_A), 4.19 (d, J = 10.6 Hz, 1H, C(5 or 6)H_B), 3.02 (s, 3H, CH_3-OMs), 3.01 (s, 3H, CH_3-OMs), 2.15 (s, 3H, CH_3-OAc), 2.09 (s, 3H, CH_3-OAc); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.3 (*C*OCH$_3$-OAc), 168.8 (*C*OCH$_3$-OAc), 128.7 (*C*-Ar), 128.5 (*C*-Ar), 97.4 (*C*-1), 82.8 (*C*-4), 78.8 (*C*-3), 74.1 (*C*H$_2$-OBn), 73.5 (*C*-2), 68.9 (*C*-5 or 6), 68.5 (*C*-5 or 6), 37.8 (*C*H$_3$-OMs), 37.6 (*C*H$_3$-OMs), 21.0 (*C*H$_3$-OAc), 20.7 (*C*H$_3$-OAc); LRMS-ESI (*m/z*): $[M+Na]^+$ 533 (100%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{19}H_{26}O_{12}S_2Na$, 533.0758;

found, 533.0755.

(2*R*,3*R*,4*S*)-4-(benzyloxy)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-5,5-bis (((methylsulfonyl)oxy)methyl)tetrahydrofuran-3-yl acetate (**24**)



23

24

$C_{22}H_{28}N_2O_{12}S_2$

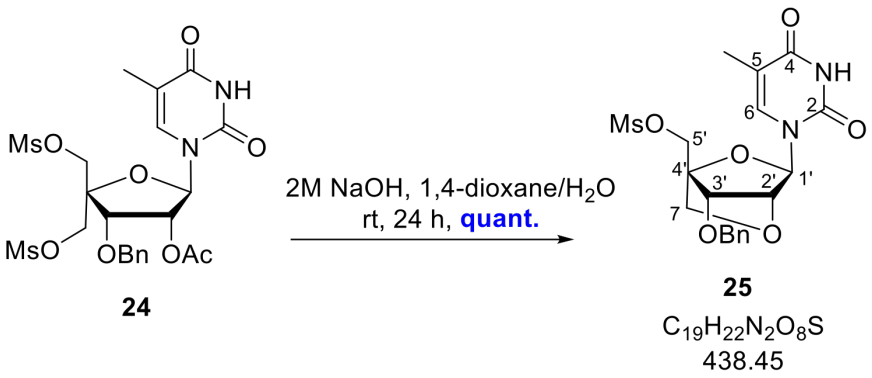
576.59

Compound **23** (2.85 g, 5.58 mmol), BSA (3.80 mL, 15.42 mmol), and thymine (881 mg, 6.99 mmol) were suspended in anhydrous MeCN (50 mL). The reaction mixture was refluxed under argon for 1 h. The solution was cooled to rt and TMSOTf (1.31 mL, 7.26 mmol) was added and the reaction was refluxed overnight. The solution was cooled to rt, diluted with CH_2Cl_2 (50 mL) and washed with sat. $NaHCO_3$ (2 × 50 mL). The aqueous layer was washed with CH_2Cl_2 (2 × 50 mL). The organic layers were combined, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (0-5% MeOH/CH_2Cl_2) to give **24** as a yellow foam (2.37 g, 4.11 mmol, 74% yield).

R_f: 0.40 (CH_2Cl_2:MeOH, 95:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.93 (s, 1H, N*H*), 7.30-7.39 (m, 5H, *H*-Ar), 7.07 (s, 1H, C(6)*H*), 5.71 (d, J = 3.4 Hz, 1H, C(1')*H*), 5.56 (dd, J = 6.5, 3.4 Hz, 1H, C(2')*H*), 4.69 (d, J = 6.6 Hz, 1H, C(3')*H*), 4.57 (q, J = 10.9 Hz, 2H, C(5' or 6')H_2), 4.54 (d, J = 11.7 Hz, 1H, CH_A-OBn), 4.35 (q, J = 10.8 Hz, 2H, C(5' or 6')H_2), 4.31 (d, J = 11.7 Hz, 1H, CH_B-OBn), 3.02 (s, 3H, CH_3-OMs), 3.00 (s, 3H, CH_3-OMs), 2.12 (s, 3H, CH_3-OAc), 1.92 (s, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.1 (*C*OCH$_3$-OAc), 163.6 (*C*-4), 150.2 (*C*-2), 138.0 (*C*-6), 136.8 (*C*-Ar), 128.7 (*C*-Ar), 112.0 (*C*-5), 92.5 (*C*-1'), 84.1 (*C*-4'), 78.0 (*C*-3'), 75.0 (*C*H$_2$-OBn), 73.8 (*C*-2'), 68.5 (*C*-5' or 6'),

67.6 (*C*-5' or 6'), 37.9 (*C*H_3-OMs), 37.8 (*C*H_3-OMs), 20.8 (*C*H_3-OAc), 12.4 (*C*H_3-thymine); LRMS-ESI (*m/z*): $[M]^+$ 576 (100%); HRMS-ESI (*m/z*): $[M]^+$ calcd. for $C_{22}H_{28}N_2O_{12}S_2$, 576.1078; found, 576.1054; $[M+H]^+$ calcd. for $C_{22}H_{29}N_2O_{12}S_2$, 577.1157; found, 577.1166; $[M+Na]^+$ calcd. for $C_{22}H_{28}N_2O_{12}S_2Na$, 599.0976; found, 599.0979.

((1*R*,3*R*,4*R*,7*S*)-7-(benzyloxy)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl methanesulfonate (**25**)

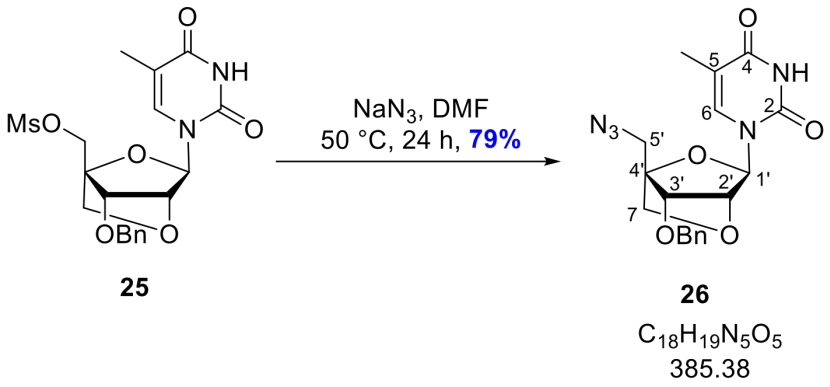


Compound **24** (2.37 g, 4.11 mmol) was dissolved in 1,4-dioxane/water (1/1, v/v) (50 mL). 2 M NaOH (50 mL) was added. The solution was stirred at rt for 24 h. The reaction was diluted with sat. $NaHCO_3$ (100 mL) and extracted with CH_2Cl_2 (2 × 100 mL), then extracted with 10% isopropanol in CH_2Cl_2 (2 × 50 mL) with the addition of 1 M HCl (2 × 20 mL) in the wash. The organic layers were combined, dried over Na_2SO_4, and concentrated under reduced pressure to give **25** (1.89 g, 4.31 mmol, quant.) as a white foam.

R_f: 0.60 (CH_2Cl_2:MeOH, 95:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.50 (s, 1H, N*H*), 7.41 (d, J = 1.2 Hz, 1H, C(6)*H*), 7.29-7.37 (m, 5H, *H*-Ar), 5.66 (d, J = 0.7 Hz, 1H, C(1')*H*), 4.62-4.70 (m, 2H, CH_2-OBn), 4.58 (s, 1H, C(2')*H*), 4.51-4.55 (m, 2H, C(5')H_2), 4.09 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.92 (s, 1H, C(3')*H*), 3.88 (d, J = 7.8 Hz, 1H, C(7)H_B), 3.07 (s, 3H, CH_3-OMs), 1.93 (s, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.8 (*C*-4), 150.0 (*C*-2), 136.9 (*C*-Ar), 134.6 (*C*-6), 128.5-129.0 (*C*-Ar), 111.3 (*C*-5), 88.2 (*C*-1'), 86.1 (*C*-4'), 77.7 (*C*-2'), 76.5 (*C*-3'), 73.0 (*C*H_2-OBn), 72.1 (*C*-7), 64.5 (*C*-5'), 38.4 (*C*H_3-OMs), 12.9 (*C*H_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 439 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{19}H_{23}N_2O_8S$, 439.1170; found, 439.1168; $[M+Na]^+$ calcd. for $C_{19}H_{22}N_2O_8SNa$,

461.0989; found, 461.0990.

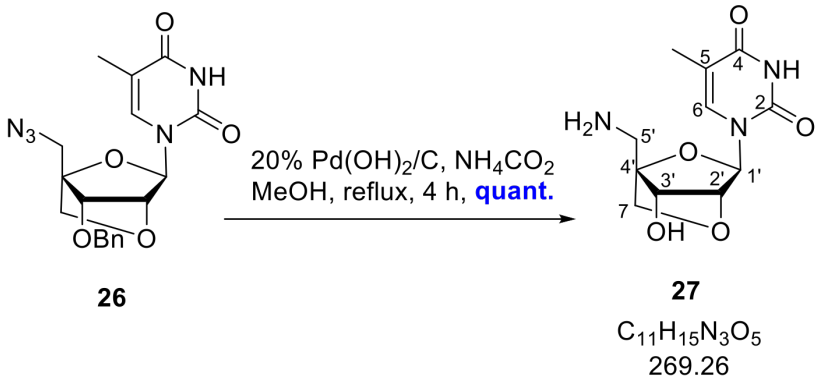
1-((1*S*,3*R*,4*R*,7*S*)-1-(azidomethyl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**26**)



Compound **25** (2.34 g, 5.34 mmol) was dissolved in DMF (50 mL). Sodium azide (694 mg, 10.7 mmol) was added and the mixture was stirred at 50 °C for 24 h. The solvent was removed under reduced pressure, re-suspended in EtOAc (100 mL), washed with sat. $NaHCO_3$ (2 × 50 mL) and dried over Na_2SO_4. Evaporation of the solvent under reduced pressure gave 26 as a white sticky compound (1.63 g, 4.23 mmol, 79%).

R_f: 0.60 (EtOAc:Pet. Ether, 7:3); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.52 (s, 1H, N*H*), 7.40 (d, J = 1.3 Hz, 1H, C(6)*H*), 7.28-7.38 (m, 5H, *H*-Ar), 5.64 (s, 1H, C(1')*H*), 4.66 (d, J = 11.4 Hz, 1H, CH_A-OBn), 4.57 (s, 1H, C(2')*H*), 4.55 (d, J = 11.4 Hz, 1H, CH_B-OBn), 4.03 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.87 (s, 1H, C(3')*H*), 3.84 (d, J = 7.8 Hz, C(7)H_B), 3.75 (d, J = 13.7 Hz, 1H, C(5')H_A), 3.66 (d, J = 13.7 Hz, 1H, C(5')H_B), 1.94 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.5 (*C*-4), 149.7 (*C*-2), 136.7 (*C*-Ar), 134.4 (*C*-6), 127.9-128.6 (*C*-Ar), 110.8 (*C*-5), 87.8 (*C*-4'), 86.9 (*C*-1'), 76.8 (*C*-2'), 76.5 (*C*-3'), 72.6 (*C*H$_2$-OBn), 72.5 (*C*-7), 47.5 (*C*-5'), 12.9 (*C*H$_3$-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 386 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{18}H_{20}N_5O_5$, 386.1459; found, 386.1458; $[M+Na]^+$ calcd. for $C_{18}H_{19}N_5O_5Na$, 408.1278; found, 408.1275.

1-((1*S*,3*R*,4*R*,7*S*)-1-(aminomethyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-pyrimidine-2,4(1*H*,3*H*)-dione (**27**)



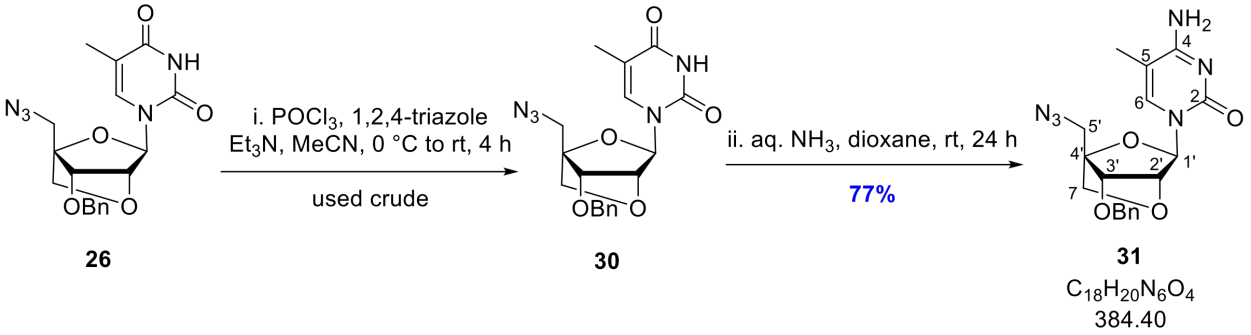
26 **27**

$C_{11}H_{15}N_3O_5$
269.26

To a solution of compound **26** (1.32 g, 3.43 mmol) dissolved in MeOH (60 mL) was added 20% $Pd(OH)_2/C$ (296 mg, 2.12 mmol). Ammonium formate (1.09 g, 17.15 mmol) was added and the reaction was refluxed under argon for 4 h. The reaction mixture was diluted with MeOH (10 mL), filtered over Celite, and concentrated under reduced pressure to give **27** (957 mg, 3.55 mmol, quant.) as a white fluffy powder.

R_f: 0.20 (CH_2Cl_2:MeOH, 85:15); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.68 (d, J = 1.3 Hz, 1H, C(6)H), 5.41 (s, 1H, C(1')H), 4.10 (s, 1H, C(3')H), 3.92 (s, 1H, C(2')H), 3.81 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.62 (d, J = 7.8 Hz, 1H, C(7)H_B), 2.89-2.97 (m, 2H, C(5')H_2), 1.78 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.9 (C-4), 150.0 (C-2), 135.1 (C-6), 108.5 (C-5), 87.9 (C-4'), 86.4 (C-1'), 79.1 (C-3'), 71.6 (C-7), 69.2 (C-2'), 36.8 (C-5'), 12.3 (CH$_3$-thymine); LRMS-ESI (m/z): $[M+H]^+$ 270 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{11}H_{16}N_3O_5$, 270.1085; found, 270.1117; $[M+Na]^+$ calcd. for $C_{11}H_{15}N_3O_5Na$, 292.0904; found, 292.0893.

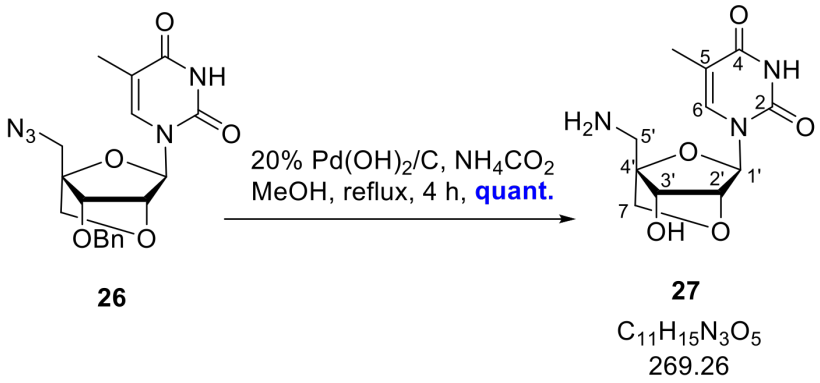
4-amino-1-((1*S*,3*R*,4*R*,7*S*)-1-(azidomethyl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidin-2(1*H*)-one (**31**)



$POCl_3$ was freshly distilled at 80 °C under 200 mbar pressure. $POCl_3$(1.51 mL, 16.2 mmol) was slowly added to an ice-cold suspension of 1,2,4-triazole (4.89 g, 70.8 mmol) in anhydrous MeCN (88 mL). The white foaming solution was left to stir at 0 °C for 10 min. Et_3N (11.3 mL, 80.9 mmol) was added dropwise and the solution was left to stir at rt for 30 min. A solution of compound **26** (779 mg, 2.02 mmol) in anhydrous MeCN (12 mL) was added and the reaction mixture was left to stir for 4 h. The reaction was poured into stirring EtOAc (50 mL) and washed with water (2 × 50 mL), brine (2 × 50 mL), dried over Na_2SO_4, and the solvent was removed to produce a yellow oil. This was dissolved in 1,4-dioxane (30 mL) and aqueous ammonia (6.40 mL) was added to the solution which was left to stir at rt overnight. The solvent was reduced and the crude oil was purified by column chromatography (0-15% MeOH/EtOAc) to yield compound **31** (597 mg, 1.55 mmol, 77%) as a white foam.

R_f: 0.50 (EtOAc:MeOH, 9:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.19 (s, 1H, N*H*), 7.46 (d, J = 1.1 Hz, 1H, C(6)*H*), 7.27-7.35 (m, 5H, *H*-Ar), 5.70 (s, 1H, C(1')*H*), 4.70 (s, 1H, C(2')*H*), 4.61 (d, J = 11.4 Hz, CH_A-OBn), 4.46 (d, J = 11.4 Hz, CH_B-OBn), 4.01 (m, 1H, C(7)H_A), 3.85 (d, J = 7.7 Hz, 1H, C(7)H_B), 3.83 (s, 1H, C(3')*H*), 3.75 (d, J = 13.7 Hz, 1H, C(5')H_A), 3.65 (d, J = 13.6 Hz, 1H, C(5')H_B), 1.96 (s, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.0 (*C*-4), 155.8 (*C*-2), 137.1 (*C*-Ar), 136.9 (*C*-6), 128.6 (*C*-Ar), 128.3 (*C*-Ar), 127.9 (*C*-Ar), 102.1 (*C*-5), 88.3 (*C*-1'), 86.7 (*C*-4'), 76.7 (*C*-2'), 76.4 (*C*-3'), 72.5 (*C*-7), 72.3 (*C*H_2-OBn), 47.7 (*C*-5'), 13.6 (*C*H_3-^{Me}cytosine); LRMS-ESI (m/z): $[M]^+$ 384 (100%); HRMS-ESI (m/z): $[M\text{-}H]^+$ calcd. for $C_{18}H_{19}N_6O_4$, 383.1473; found, 383.1474.

1-((1*S*,3*R*,4*R*,7*S*)-1-(aminomethyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-pyrimidine-2,4(1*H*,3*H*)-dione (**27**)



26

27

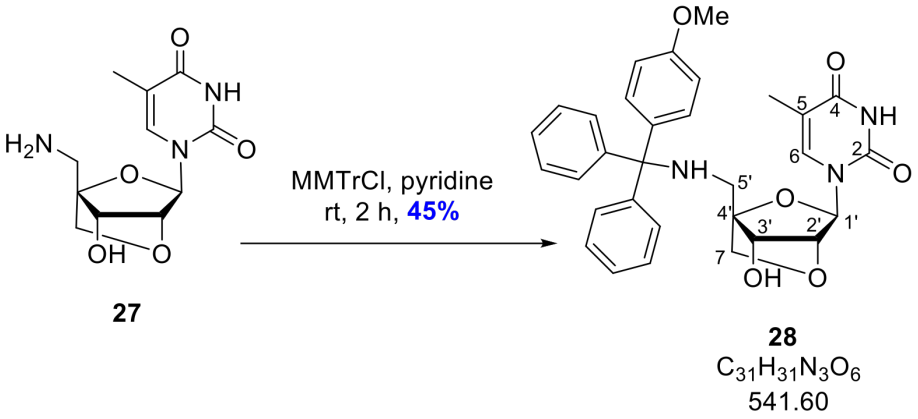
$C_{11}H_{15}N_3O_5$

269.26

To a solution of compound **26** (1.32 g, 3.43 mmol) dissolved in MeOH (60 mL) was added 20% $Pd(OH)_2/C$ (296 mg, 2.12 mmol). Ammonium formate (1.09 g, 17.15 mmol) was added and the reaction was refluxed under argon for 4 h. The reaction mixture was diluted with MeOH (10 mL), filtered over Celite, and concentrated under reduced pressure to give **27** (957 mg, 3.55 mmol, quant.) as a white fluffy powder.

R_f: 0.20 (CH_2Cl_2:MeOH, 85:15); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.68 (d, J = 1.3 Hz, 1H, C(6)H), 5.41 (s, 1H, C(1')H), 4.10 (s, 1H, C(3')H), 3.92 (s, 1H, C(2')H), 3.81 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.62 (d, J = 7.8 Hz, 1H, C(7)H_B), 2.89-2.97 (m, 2H, C(5')H_2), 1.78 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.9 (C-4), 150.0 (C-2), 135.1 (C-6), 108.5 (C-5), 87.9 (C-4'), 86.4 (C-1'), 79.1 (C-3'), 71.6 (C-7), 69.2 (C-2'), 36.8 (C-5'), 12.3 (CH$_3$-thymine); LRMS-ESI (m/z): $[M+H]^+$ 270 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{11}H_{16}N_3O_5$, 270.1085; found, 270.1117; $[M+Na]^+$ calcd. for $C_{11}H_{15}N_3O_5Na$, 292.0904; found, 292.0893.

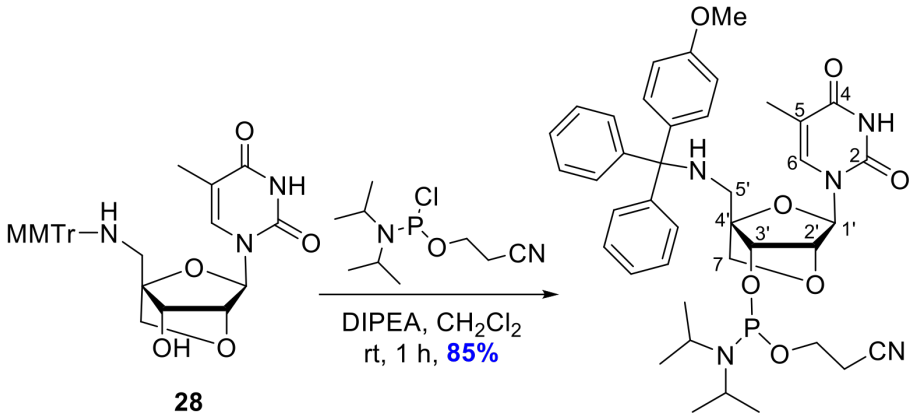
1-((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-1-((((4-methoxyphenyl)diphenylmethyl)amino)methyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**28**)



Compound **27** (585 mg, 2.17 mmol) was dissolved in anhydrous pyridine (20 mL) and 4-methoxytriphenylmethyl chloride (MMTrCl) (805 mg, 2.61 mmol) was added portion-wise over 1 h. The reaction was stirred for 2 h at rt and pyridine was removed under pressure. The reaction was purified by column chromatography (0-70% EtOAc/Pet. Ether, 0.1% pyr) to give **28** (534 mg, 0.98 mmol, 45%) as a white foam.

R_f: 0.70 (EtOAc:pyr, 99:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.65 (d, J = 1.3 Hz, 1H, C(6)H), 7.51 – 7.44 (m, 4H, H-Ar), 7.42 – 7.32 (m, 2H, H-Ar), 7.23 – 7.14 (m, 2H, H-Ar), 6.86 – 6.78 (m, 2H, H-Ar), 5.63 (s, 1H, C(1')H), 4.46 (s, 1H, C(3')H), 4.27 (s, 1H, C(2')H), 3.93 (d, J = 8.3 Hz, 1H, C(7)H_A), 3.80 (d, J = 8.3 Hz, 1H, C(7)H_B), 3.76 (s, 3H, CH_3-OMe), 2.58 (d, J = 4.0 Hz, 2H, C(5')H_2), 1.94 (s, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.1 (C-4), 158.2 (C-Ar), 150.1 (C-2), 145.9 (C-Ar), 145.8 (C-Ar), 137.4 (C-6), 136.3 (C-Ar), 134.7 (C-Ar), 129.9 (C-Ar), 128.5 (C-Ar), 128.2 (C-Ar), 126.7 (C-Ar), 113.5 (C-Ar), 110.4 (C-5), 88.8 (C-4'), 87.2 (C-1'), 79.8 (C-3'), 72.7 (C-7), 70.7 (C-2'), 70.4 (C-MMTr), 55.3 (CH$_3$-OMe), 40.2 (C-5'), 12.8 (CH$_3$-thymine); LRMS-ESI (m/z): $[M]^+$ 541 (100%), $[M+H]^+$ 542 (33%), $[2M]^+$ 1082 (96%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{31}H_{31}N_3O_6Na$, 564.2105; found, 564.2107.

2-cyanoethyl-((1*S*,3*R*,4*R*,7*S*)-1-((((4-methoxyphenyl)diphenylmethyl)amino)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl) diisopropylphosphoramidite (**29**)



28

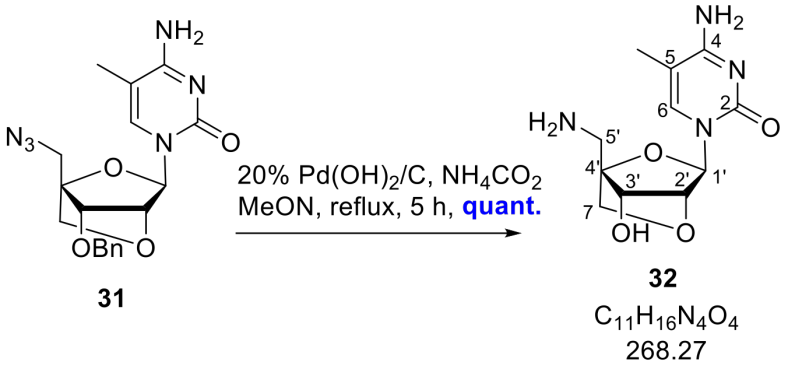
29

$C_{40}H_{48}N_5O_7P$
741.83

Compound **28** (534 mg, 0.986 mmol) was dissolved in anhydrous, degassed CH_2Cl_2 (5 mL) and anhydrous, degassed DIPEA (429 µL, 2.46 mmol) and 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite (330 µL, 1.48 mmol) were added. The reaction was stirred at rt under an inert atmosphere for 1 h. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was submitted to inert column chromatography (70% EtOAc/Pet. Ether with 1% pyridine) to give phosphoramidite **29** (622 mg, 0.838 mmol, 85%) as a white foam.

R_f: 0.55 (EtOAc:Pet. Ether:pyr, 70:29:1); ^{31}P NMR (162 MHz, $CDCl_3$) δ 148.67, 148.30, 138.85 (reagent, small impurity); HRMS-ESI (m/z): $[M+Na]^+$ calcd for $C_{40}H_{48}N_5O_7PNa$, 764.3184; found, 764.3209.

4-amino-1-((1*S*,3*R*,4*R*,7*S*)-1-(aminomethyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidin-2(1*H*)-one (**32**)



31

32

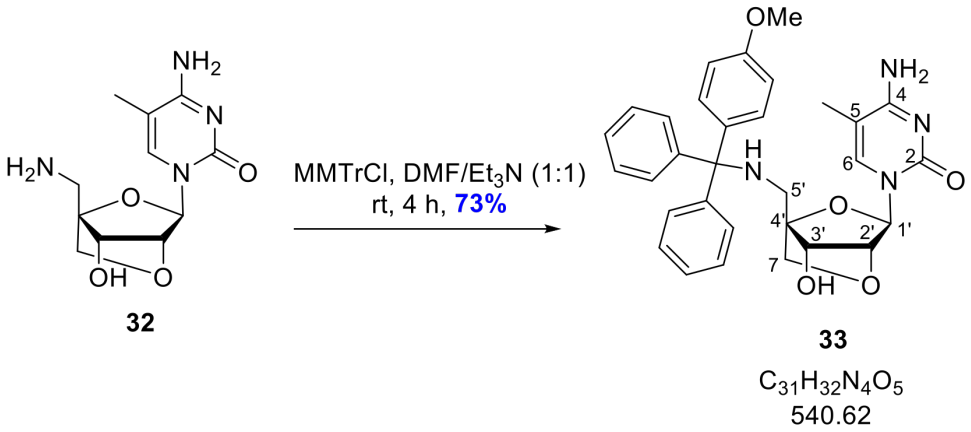
$C_{11}H_{16}N_4O_4$

268.27

Compound **31** (930 mg, 2.42 mmol) was dissolved in MeOH (25 mL) and 20% $Pd(OH)_2/C$ (168 mg, 10 mol%) and NH_4HCO_2 (1.06 g, 16.8 mmol) was added and refluxed for 5 h. The reaction was allowed to cool to rt and filtered through a layer of Celite. The solvent was removed under pressure and the crude was submitted to flash column chromatography (50-70% MeOH/EtOAc) to give **32** (676 mg, 2.51 mmol, quant.) as a white powder.

R_f: 0.18 (MeOH:EtOAc, 3:7); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.33 (s, 1H, N*H*), 7.57 (d, J = 1.2 Hz, 1H, C(6)*H*), 5.42 (s, 1H, C(1')*H*), 4.09 (s, 1H, C(3')*H*), 3.92 (s, 1H, C(2')*H*), 3.83 (d, J = 7.9 Hz, 1H, C(7)H_A), 3.67 (d, J = 7.9 Hz, 1H, C(7)H_B), 3.07 (q, J = 14.2 Hz, 2H, C(5')H_2), 1.86 (d, J = 1.1 Hz, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 165.6 (*C*-4), 154.7 (*C*-2), 137.3 (*C*-6), 100.9 (*C*-5), 87.3 (*C*-1'), 87.1 (*C*-4'), 79.3 (*C*-3'), 71.6 (*C*-7), 69.0 (*C*-2'), 36.8 (*C*-5'), 13.3 (*C*H$_3$-^{Me}cytosine); LRMS-ESI (*m/z*): $[M]^+$ 268 (56%), [2M]+ 536 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{11}H_{17}N_4O_4$, 269.1244; found 269.1241.

4-amino-1-((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-1-((((4-methoxyphenyl)diphenylmethyl)amino)methyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidin-2(1*H*)-one (**33**)



32

33

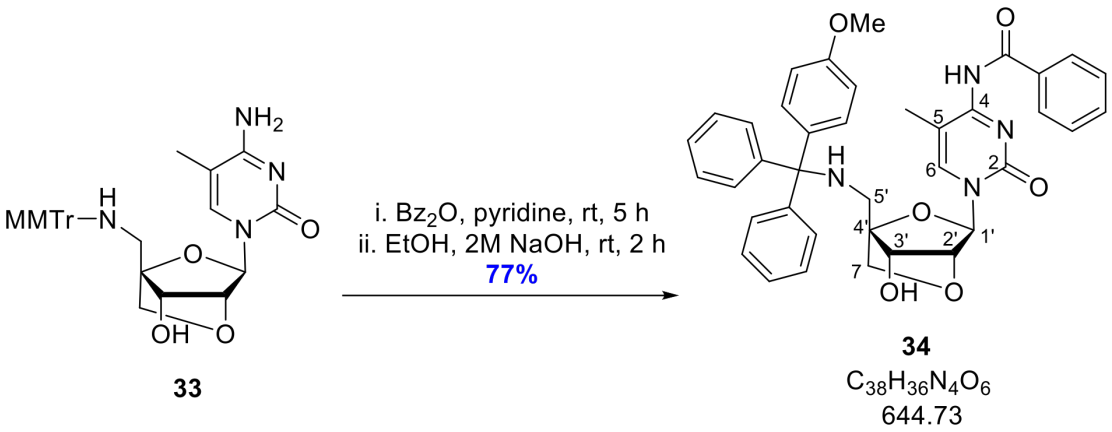
$C_{31}H_{32}N_4O_5$

540.62

Compound **32** (677 mg, 2.52 mmol) was dissolved in a solution of anhydrous DMF:anhydrous Et_3N (1:1, v:v) (3 mL) and MMTrCl (857 mg, 2.78 mmol) was added, and the reaction was stirred at rt for 4 h. The solvent was removed under reduced pressure and the residue redissolved in CH_2Cl_2 (10 mL), which was washed with H_2O (10 mL), brine (10 mL), dried over Na_2SO_4, and the solvent removed under reduced pressure. The residue was purified by column chromatography (0-20% MeOH/EtOAc with 1% pyridine) to give **33** (997 mg, 1.84 mmol, 73%) as a white foam.

R_f: 0.45 (EtOAc:MeOH:pyr, 80:19:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (d, J = 1.2 Hz, 1H, C(6)*H*), 7.52 – 7.44 (m, 4H, *H*-Ar), 7.42 – 7.34 (m, 2H, *H*-Ar), 7.34 – 7.23 (m, 4H, *H*-Ar), 7.17 (t, J = 7.2 Hz, 2H, *H*-Ar), 6.85 – 6.76 (m, 2H, *H*-Ar), 5.57 (s, 1H, C(1')*H*), 4.45 (s, 1H, C(3')*H*), 4.13 (s, 1H, C(2')*H*), 3.91 (d, J = 8.0 Hz, 1H, C(7)H_A), 3.77 (d, J = 8.0 Hz, 1H, C(7)H_B), 3.75 (s, 3H, CH_3-OMe), 2.67 – 2.46 (m, 2H, C(5')H_2), 1.96 – 1.91 (m, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.8 (*C*-4), 157.9 (*C*-Ar), 149.1 (*C*-2), 145.9 (*C*-Ar), 137.4 (*C*-6), 136.7 (*C*-Ar), 129.8 (*C*-Ar), 128.4 (*C*-Ar), 127.9 (*C*-Ar), 126.4 (*C*-Ar), 124.1 (*C*-Ar), 113.2 (*C*-Ar), 102.2 (*C*-5), 88.3 (*C*-4'), 87.8 (*C*-1'), 79.6 (*C*-3'), 72.7 (*C*-7), 70.4 (*C*-MMTr), 70.2 (*C*-2'), 55.2 (*C*H$_3$-OMe), 40.2 (*C*-5'), 13.2 (*C*H$_3$-^{Me}cytosine); LRMS-ESI (*m/z*): $[M]^+$ 540 (88%), $[2M]^+$ 1080 (100%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{31}H_{32}N_4O_5Na$, 563.2265; found, 563.2280

N-(1-((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-1-((((4-methoxyphenyl)diphenylmethyl)amino)methyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (**34**)



33

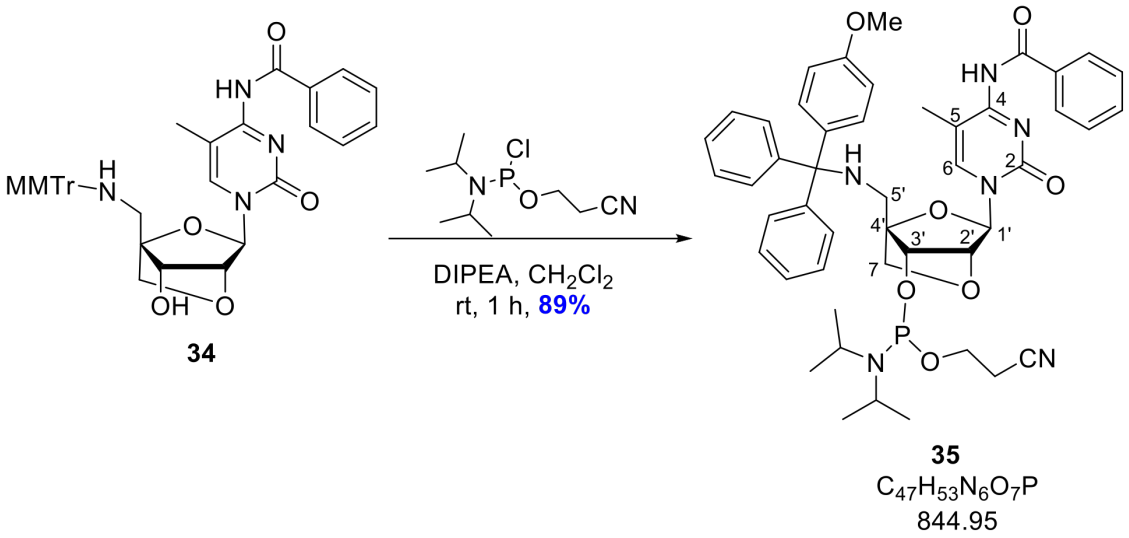
34
$C_{38}H_{36}N_4O_6$
644.73

Compound **33** (450 mg, 0.832 mmol) was dissolved in anhydrous pyridine (10 mL) and benzoic anhydride (Bz_2O) was added (377 mg, 1.66 mmol) and the solution was stirred at rt for 5 h. The solution was diluted in EtOH (10 mL) and treated with 2M NaOH (10 mL) and left to stir at rt for 2 h. The solution was diluted with $NaHCO_3$ (50 mL) and the product was extracted with CH_2Cl_2 (3 × 50 mL), dried over Na_2SO_4, and reduced under pressure. The residue was purified by column chromatography (0-50% EtOAc/Pet. Ether, 1% pyridine) to give **34** (411 mg, 0.637 mmol, 77%) as a white foam.

R_f: 0.75 (EtOAc:Pet. Ether:pyr, 50:50:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33-8.30 (m, 2H, *H*-Ar(Bz)), 7.81 (d, J = 1.3 Hz, 1H, C(6)*H*), 7.54 – 7.49 (m, 4H, *H*-Ar(MMTr)), 7.46 – 7.42 (m, 2H, *H*-Ar(Bz)), 7.41– 7.39 (m, 2H, *H*-Ar(MMTr)), 7.33 – 7.28 (m, 5H, 4 x *H*-Ar(MMTr) and Bz), 6.83– 6.87 (m, 2H, *H*-Ar(MMTr)), 5.70 (s, 1H, C(1')*H*), 4.45 (s, 1H, C(3')*H*), 4.27 (s, 1H, C(2')H, 3.90 (d, J = 8.5 Hz, 1H, C(7)H_A), 3.83 (d, J = 8.5 Hz, 1H, C(7)H_B), 2.58 (d, J = 15.8 Hz, 2H, C(5')H_2), 2.16 (d, J = 1.1 Hz, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.9 (*C*-4), 158.3 (*C*-Ar), 149.9 (*C*-2), 145.8 (*C*-Ar), 137.4 (*C*-6), 132.7 (*C*-Ar), 130.1 (*C*-Ar), 129.9 (*C*-Ar), 128.6 (*C*-Ar), 128.3 (*C*-Ar), 128.3 (*C*-Ar), 126.8 (*C*-Ar), 123.9 (*C*-Ar), 113.6 (*C*-Ar), 111.7 (*C*-5), 88.9 (*C*-4'), 87.4 (*C*-1'), 79.6 (*C*-3'), 72.6 (*C*-7), 70.9 (*C*-2'), 70.4 (*C*-MMTr), 55.4 (*C*H$_3$-OMe), 40.1 (*C*-5'), 14.0 (*C*H$_3$-^{Me}cytosine); LRMS-ESI (m/z): $[2M]^+$ 1288 (66%), $[2M+Na]^+$ 1289 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for

$C_{38}H_{37}N_4O_6$, 645.2708; found, 645.2735.

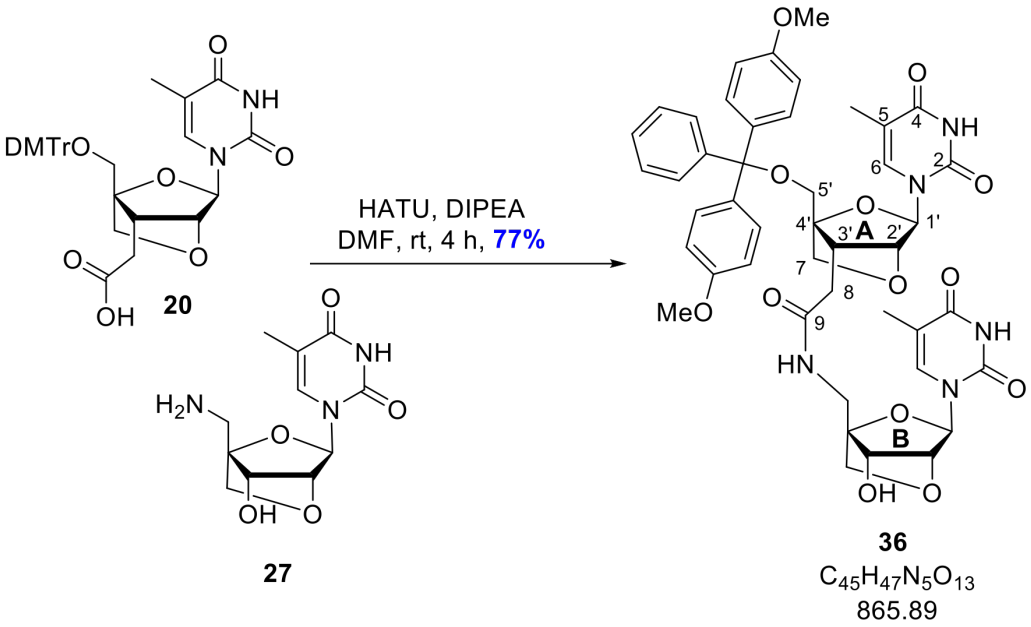
(1*S*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((((4-methoxyphenyl) diphenylmethyl)amino)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite (**35**)



Compound **34** (300 mg, 0.465 mmol) was dissolved in anhydrous, degassed CH_2Cl_2 (3 mL) and DIPEA (203 µL, 1.16 mmol) and chloro-phosphitylating reagent (155 µL, 0.698 mmol) were added. The reaction was stirred at rt under an inert atmosphere for 1 h. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was submitted to inert column chromatography (50% EtOAc/Pet. Ether with 1% pyridine) to give phosphoramidite **35** (351 mg, 0.414 mmol, 89%) as a white foam.

R_f: 0.85 (EtOAc:Pet. Ether:pyr, 50:49:1); ^{31}P NMR (162 MHz, $CDCl_3$) δ 148.87, 148.41; HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{47}H_{54}N_6O_7P$, 845.3786; found, 845.3770.

2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)-*N*-(((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl)acetamide (**36**)

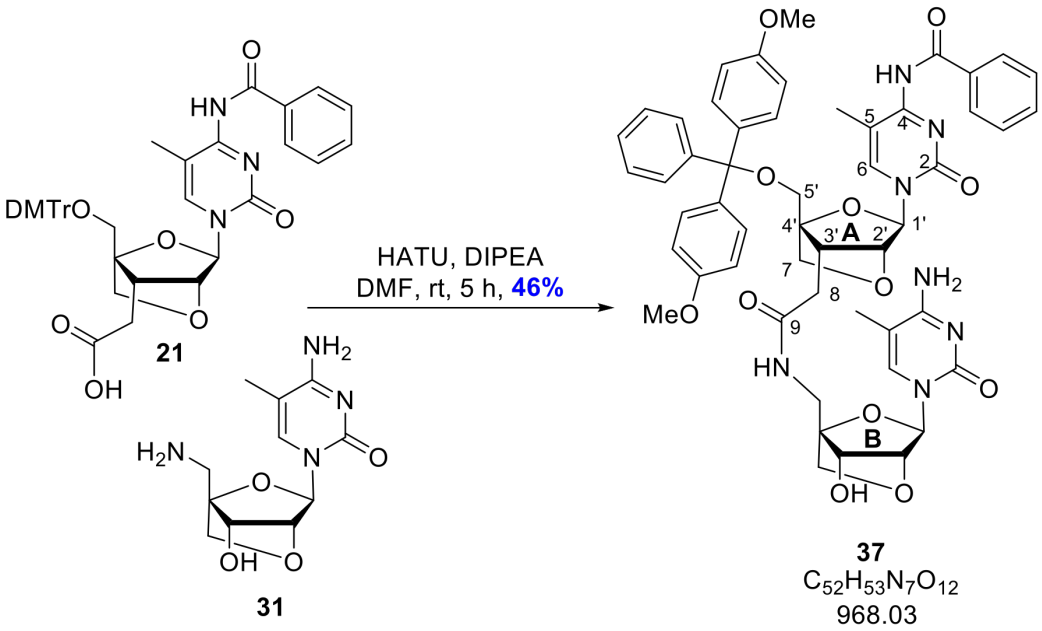


Compound **20** (450 mg, 0.732 mmol) was dissolved in anhydrous DMF (5 mL) and anhydrous diisopropylethylamine (DIPEA) (0.256 mL, 1.47 mmol) and 4Å molecular sieves were added. The reaction was stirred at rt for 15 min before hexafluorophosphate-azabenzotriazole-tetramethyl-uronium (HATU) (278 mg, 0.732 mmol) was added. The reaction was stirred for 30 min and then compound **27** (217 mg, 0.806 mmol) was added. The reaction was stirred at rt for 4 h and the solvent removed under reduced pressure. The residue was re-dissolved in EtOAc (10 mL), washed with sat. $NaHCO_3$ (2 × 10 mL), dried over Na_2SO_4, and the solvent removed under reduced pressure. The resulting residue was purified via column chromatography (0-20% MeOH/EtOAc with 1% pyridine) to give dimer **36** (488 mg, 0.564 mmol, 77%) as a white powder.

R_f: 0.45 (EtOAc:MeOH:pyridine, 79:20:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.43 (s, 1H, N*H*-thymine), 11.35 (s, 1H, N*H*-thymine), 8.28 (t, J = 6.0 Hz, 1H, N*H*-amide), 7.55 (d, J = 1.3 Hz, 1H, C(A6 or B6)*H*), 7.42 (dd, J = 8.5, 1.3 Hz, 1H, C(A6 or B6)*H*), 7.25-7.45 (m, 9H,

H-Ar), 6.89-6.92 (m, 4H, *H*-Ar), 5.53 (d, J = 3.9 Hz, 1H, C(B3')O*H*), 5.48 (s, 1H, C(A1' or B1')*H*), 5.38 (s, 1H, C(A1' or B1')*H*), 4.38 (s, 1H, C(A2' or B2')*H*), 4.12 (s, 1H, C(A2' or B2')*H*), 3.81 (d, J = 8.0 Hz, 1H, C(7A or 7B)H_A), 3.76 (d, J = 3.6 Hz, 1H, C(B3')*H*), 3.74 (d, J = 2.2 Hz, 6H, 2 × CH_3-OMe), 3.69 (d, J = 8.5 Hz, 1H, C(7A or 7B)H_A), 3.62 (d, J = 8.5 Hz, 1H, C(7A or 7B)H_B), 3.56 (d, J = 8.0 Hz, 1H, C(7A or 7B)H_B), 3.51 (t, J = 5.6 Hz, 2H, C(B5')H_2), 3.48 (d, J = 11.5 Hz, 1H, C(A5')H_A), 3.36 (d, J = 11.3 Hz, 1H, C(A5')H_B), 2.48 (s, 1H, C(A3')*H*), 2.20 (dd, J = 15.4, 8.6 Hz, 1H, C(8)H_A), 2.10 – 2.03 (m, 1H, C(8)H_B), 1.72 (d, J = 1.2 Hz, 3H, CH_3-thymine(A or B), 1.55 (d, J = 1.1 Hz, 3H, CH_3-thymine(A or B); ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.4 (*C*-9), 158.2 (*C*-A4 or B4), 144.6 (*C*-A2 or B2), 135.1 (*C*-A6 or B6), 135.1 (*C*-A7 or B6), 134.5 (*C*-Ar), 134.2 (*C*-Ar), 129.8 (*C*-Ar), 129.8 (*C*-Ar), 127.9 (*C*-Ar), 127.7 (*C*-Ar), 126.9 (*C*-Ar), 113.3 (*C*-Ar), 108.5 (*C*-A5 or B5), 108.5 (*C*-A5 or B5), 89.2 (*C*-DMTr or A4'), 87.3 (*C*-DMTr or A4'), 86.6 (*C*-A1' or B1'), 86.4 (*C*-A1' or B1'), 85.8 (*C*-B4'), 79.8 (*C*-A2' or B2'), 78.9 (*C*-A2' or B2'), 71.4 (*C*-7A or 7B), 71.4 (*C*-7A or 7B), 70.0 (*C*-B3'), 59.0 (*C*-A5'), 55.0 (*C*H_3-OMe), 40.1 (*C*-A3'), 35.3 (*C*-B5'), 29.5 (*C*-8), 12.3 (*C*H_3-thymine(A or B)), 12.1 (*C*H_3-thymine(A or B)); HRMS-ESI (*m/z*): $[M+Na]^+$: calcd. for $C_{45}H_{47}N_5O_{13}$, 888.3063; found, 888.3060.

N-(1-((1*R*,3*R*,4*R*,7*S*)-7-(2-((((1*S*,3*R*,4*R*,7*S*)-3-(4-amino-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl)amino)-2-oxoethyl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (**37**)

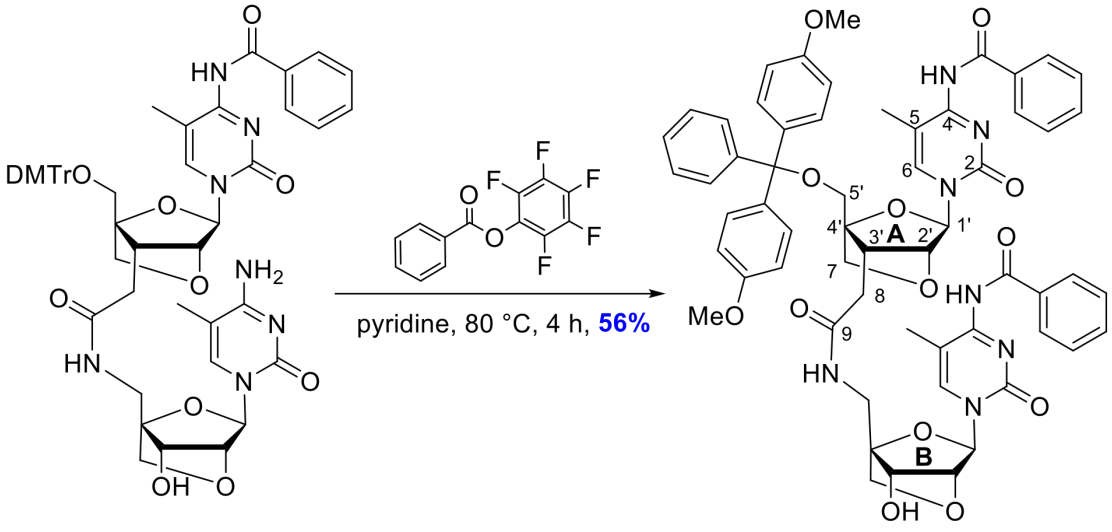


Compound **21** (541 mg, 0.754 mmol) was dissolved in anhydrous DMF (4 mL) in the presence of 3Å molecular sieves and DIPEA (26 µL, 1.51 mmol) and HATU (373 mg, 0.980 mmol) were added and stirred at rt for 30 min. Compound **31** (283 mg, 1.056 mmol) was added and the mixture was left to stir for 3 h. An additional portion of HATU (58 mg, 0.153 mmol) was added and the mixture was left to stir for another hour. The DMF was removed under reduced pressure and the residue was purified by column chromatography (0-10% MeOH/EtOAc with 1% pyridine) to give dimer **37** (340 mg, 0.351 mmol, 46%) as a white powder.

R_f: 0.23 (EtOAc:MeOH:pyr, 80:19:1); ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 7.54 – 7.47 (m, 3H, *H*-Ar and C(A6 and B6)*H*), 7.41 – 7.28 (m, 13H, *H*-Ar), 6.95 – 6.88 (m, 4H, *H*-Ar), 5.57 (s, 1H, C(A1' or B1')*H*), 5.45 (d, J = 3.9 Hz, 1H, C(3')-O*H*), 5.37 (s, 1H, C(A1' or B1')*H*), 4.49 (s, 1H, C(A2' or B2')*H*), 4.06 (s, 1H, C(A2' or B2')*H*), 3.80 (d, J = 7.9 Hz, 1H, C(A7 or B7)H_A), 3.74 (d, J = 2.3 Hz, 7H, 2 × CH_3-OMe and C(A7 or B7)H_B), 3.67 (dt, J =

4.9, 2.2 Hz, 2H, C(B3')H and C(A7 or B7)H_A), 3.58 – 3.50 (m, 3H, C(A7 or B7)H_B and C(A5')H_A and C(B5')H_A), 3.43 (d, J = 11.2 Hz, 2H, C(A5')H_B and C(B5')H_B), 2.54 (t, J = 6.9 Hz, 1H, C(A3')H), 2.24 (dd, J = 15.4, 8.1 Hz, 1H, C(8)H_A), 2.14 – 2.07 (m, 1H, C(8)H_B), 1.80 (d, J = 17.3 Hz, 6H, 2 × CH_3-^{Me}cytosine (A or B)); ^{13}C NMR (151 MHz, DMSO-d$_6$) δ 170.8 (C-9), 165.9 (C-B4), 158.7 (C-A4), 158.7 (C-Ar), 155.0 (C-B2), 145.0 (C-A2), 137.3 (C-A6 or B6), 137.1 (C-A6 or B6), 136.6 (C-Ar), 135.6 (C-Ar), 135.6 (C-Ar), 135.1 (C-Ar), 134.4 (C-Ar), 133.0 (C-Ar), 132.0 (C-Ar), 130.3 (C-Ar), 129.9 (C-Ar), 129.3 (C-Ar), 128.8 (C-Ar), 128.5 (C-Ar), 128.3 (C-Ar), 127.9 (C-Ar), 127.4 (C-Ar), 124.4 (C-Ar), 113.8 (C-Ar), 101.2 (C-A5 or B5), 101.1 (C-A5 or B5), 90.2 (C-A4'), 87.9 (C-DMTr), 87.5 (C-A1' or B1'), 87.5 (C-A1' or B1'), 86.4 (C-B4'), 80.2 (C-A2' or B2'), 79.5 (C-A2' or B2'), 72.1 (C-A7 or B7), 71.9 (C-A7 or B7), 70.3 (C-B3'), 59.4 (C-A5'), 55.5 (CH$_3$-OMe), 40.6 (C-A3'), 35.8 (C-B5'), 30.1 (C-8), 13.7 (CH$_3$-^{Me}cytosine (A or B)); LRMS-ESI (m/z): [M+H]$^+$ 968 (100%), [M+H]$^+$ 969 (82%), [M+Na]$^+$ 990 (16%); HRMS-ESI (m/z): [M+H]$^+$: calcd. for $C_{52}H_{54}N_7O_{12}$, 968.3825; found 968.3798.

N-(1-((1*S*,3*R*,4*R*,7*S*)-1-((2-((1*R*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl) acetamido)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (**38**)



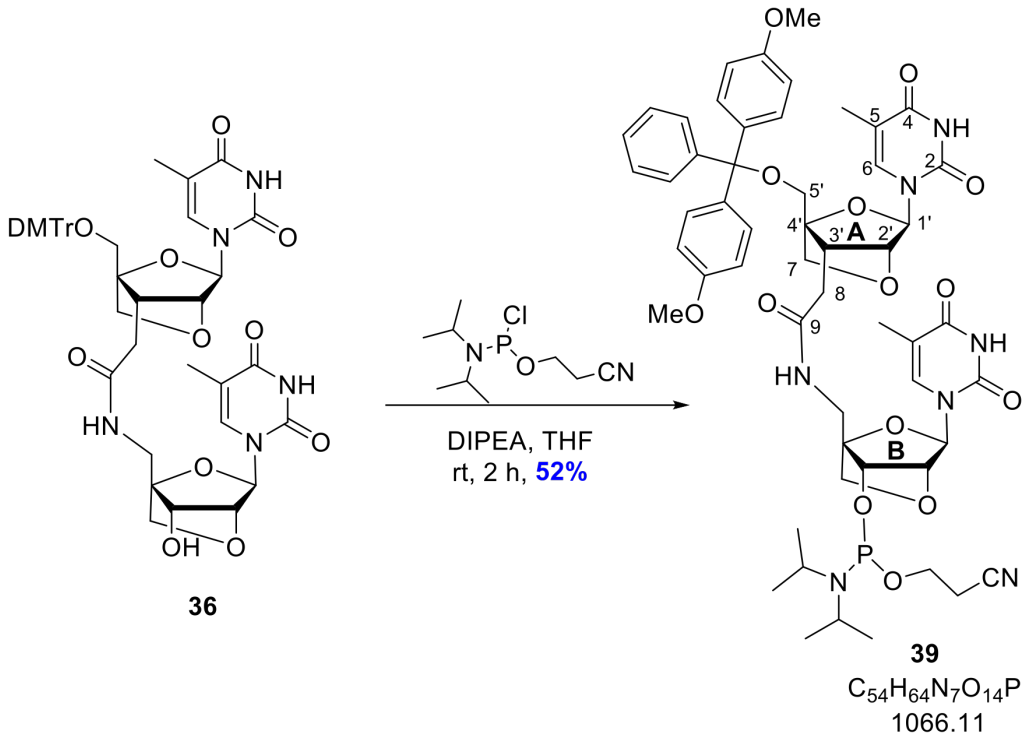
37

38
$C_{59}H_{57}N_7O_{13}$
1072.14

Compound **37** (340 mg, 0.035 mmol) was dissolved in anhydrous pyridine (5 mL) and pentafluorophenyl benzoate (141 mg, 0.489 mmol) was added and stirred at 80 °C for 4 h. After cooling to rt, the pyridine was removed under reduced pressure and the crude product was purified by column chromatography (0-5% MeOH/EtOAc with 1% pyridine) to give dimer **38** (211 mg, 0.197 mmol, 56%) as a white powder.

R_f: 0.62 (EtOAc:MeOH:pyridine, 95:4:1); ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.99 – 7.92 (m, 2H, *H*-Ar), 7.87 – 7.59 (m, 3H, *H*-Ar), 7.57 – 7.43 (m, 8H, *H*-Ar and C(A6 and B6)*H*), 7.40 – 7.24 (m, 8H, *H*-Ar), 6.92 (dd, J = 8.8, 4.6 Hz, 4H, *H*-Ar), 5.55 (d, J = 7.1 Hz, 2H, C(A1' or B1')*H*) and C(B3')O*H*), 5.45 (s, 1H, C(A' or B1')*H*), 4.52 (s, 1H, C(A2' or B2')*H*), 4.22 (s, 1H, C(A2' or B2')*H*), 3.86 (s, 1H, C(A7 or B7)H_A), 3.79 (d, J = 4.2 Hz, 1H, C(B3')*H*), 3.74 (d, J = 2.1 Hz, 7H, C(A7 or B7)H_B and 2 × CH_3-^{Me}cytosine), 3.68 (d, J = 8.5 Hz, 1H, C(A7 or B7)H_A), 3.61 (d, J = 8.1 Hz, 2H, C(A7 or B7)H_B and C(B5')H_A), 3.47 (s, 1H, C(B5')H_B), 3.54 (d, J = 11.3 Hz, 1H, C(A5')H_A), 3.43 (d, J = 11.3 Hz, 1H, C(A5')H_B), 2.55 (s, 1H, C(A3')*H*), 2.23 (d, J = 15.5 Hz, 1H, C(8)H_A), 2.12 (d, J = 4.9 Hz, 1H, C(8)H_B), 2.00 (s, 3H, CH_3-^{Me}cytosine), 1.80 (s, 3H, CH_3-^{Me}cytosine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 178.1 (*C*OBz), 170.4 (*C*-9), 159.3 (*C*-A4 or B4), 159.9 (*C*-A4 or B4), 158.2 (*C*-Ar), 146.9 (*C*-A2 or B2), 144.6 (*C*-A2 or B2), 136.9 (*C*-A6 or B6), 136.8 (*C*-A6 or B6), 136.6 (*C*-Ar), 136.1 (*C*-Ar), 135.1 (*C*-Ar), 133.8 (*C*-Ar), 133.3 (*C*-Ar), 132.5 (*C*-Ar), 129.8 (*C*-Ar), 129.6 (*C*-Ar), 129.3 (*C*-Ar), 129.1 (*C*-Ar), 128.8 (*C*-Ar), 128.4 (*C*-Ar), 128.3 (*C*-Ar), 127.9 (*C*-Ar), 127.7 (*C*-Ar), 127.4 (*C*-Ar), 127.4 (*C*-Ar), 126.9 (*C*-Ar), 123.9 (*C*-Ar), 113.3 (*C*-Ar), 109.4 (*C*-A5 or B5), 89.7 (*C*-A4'), 87.6 (*C*-A1' or B1'), 86.9 (*C*-A1' or B1'), 86.5 (*C*-DMTr), 85.9 (*C*-B4'), 79.6 (*C*-A2' or B2'), 78.6 (*C*-A2' or B2'), 71.6 (*C*-A7 or B7), 71.4 (*C*-A7 or B7), 70.1 (*C*-B3') 59.7 (*C*-A5'), 55.0 (*C*H$_3$-OMe), 40.1 (*C*-A3'), 35.2 (*C*-B5'), 30.7 (*C*-8), 13.2 (*C*H$_3$-^{Me}cytosine), 12.9 (*C*H$_3$-^{Me}cytosine); HRMS-ESI (m/z): $[M+H]^+$: calcd. for $C_{59}H_{58}N_7O_{13}$, 1072.4087; found, 1072.4077.

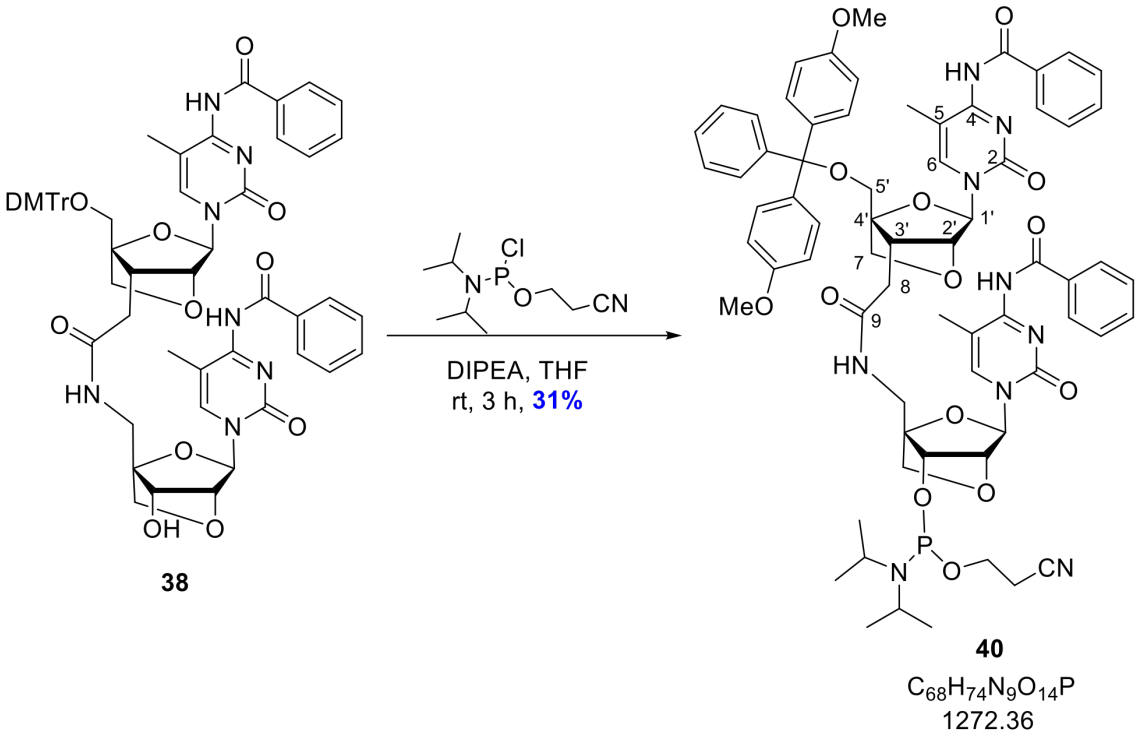
(1*S*,3*R*,4*R*,7*S*)-1-((2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl) acetamido)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1] heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite (**39**)



Compound **36** (302 mg, 0.349 mmol) was dissolved in anhydrous, degassed THF (6 mL) and anhydrous, degassed DIPEA (152 µL, 0.872 mmol) and chloro-phosphitylating reagent (117 µL, 0.0523 mmol) was added and the reaction was stirred under an inert atmosphere at rt for 2 h. The THF was removed under reduced pressure and the residue was dissolved in anhydrous, degassed CH_2Cl_2 (5 mL). An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was redissolved in the minimum volume of anhydrous degassed CH_2Cl_2 (0.2-0.5 mL) and precipitated from anhydrous, degassed hexanes (5 × 25 mL). The crude residue was submitted to inert column chromatography (99% EtOAc, 1% pyridine) to yield the phosphoramidite **39** (195 mg, 0.183 mmol, 52%) as a white foam.

R_f: 0.60 (EtOAc:pyridine, 99:1); ^{31}P NMR (162 MHz, $CDCl_3$); δ 149.2, 148.9; LRMS-ESI (m/z): [M-H-oxidised]$^-$ 926 (100%).

(1*S*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((2-((1*R*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)-methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetamido)methyl)-2,5-dioxabicyclo[2.2.1] heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite) (**40**)



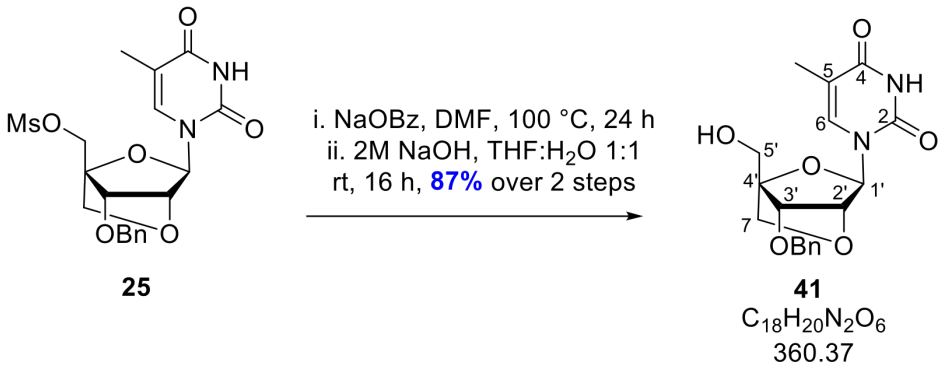
Compound **38** (178 mg, 0.166 mmol) was dissolved in anhydrous, degassed THF (3 mL) and anhydrous, degassed DIPEA (57 uL, 0.332 mmol) was added. Chloro-phosphitylating reagent (55 µL, 0.249 mmol) was added and the reaction was stirred under argon for 2 h. An additional portion of chloro-phosphitylating reagent (10 µL, 0.045 mmol) was added and the reaction was stirred for a further hour. The THF was removed under reduced pressure and the residue was dissolved in anhydrous, degassed CH_2Cl_2 (5 mL). An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was redissolved

in the minimum volume of anhydrous degassed CH_2Cl_2 (0.2-0.5 mL) and precipitated from anhydrous, degassed hexane (5 × 25 mL). The crude residue was submitted to inert column chromatography (99% EtOAc, 1% pyridine) to yield the phosphoramidite **40** (65 mg, 0.051 mmol, 31%) as a white foam.

R_f: 0.45 (EtOAc:pyridine, 99:1); ^{31}P NMR (243 MHz, $CDCl_3$) δ 149.43, 149.39, 149.04, 149.00. NB: ^{31}P impurity is dimer C(Bz)-C phosphoramidite.

7.1.3 Compounds in Chapter 3

1-((1*S*,3*R*,4*R*,7*S*)-7-(benzyloxy)-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**41**)



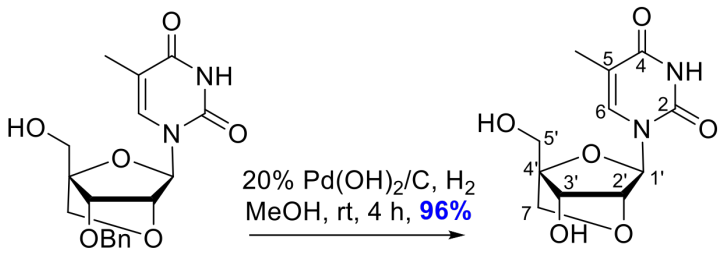
25

41
$C_{18}H_{20}N_2O_6$
360.37

Compound **25** (8.00 g, 18.2 mmol) was dissolved in DMF (75 mL) and sodium benzoate (5.26 g, 36.5 mmol) was added and the reaction was heated to 100 °C for 16 h. The reaction was cooled to rt, filtered, and the DMF was evaporated. The crude oil was redissolved in EtOAc, washed with $NaHCO_3$ (50 mL), H_2O (50 mL), brine (50 mL), dried over Na_2SO_4 and the solvent was removed under reduced pressure to give a clear oil. This was dissolved in THF:H_2O (1:1, v/v) (160 mL) and 2 M NaOH (40 mL) was added and the reaction was left to stir at rt for 16 h. The compound was extracted with CH_2Cl_2 (2 × 50 mL) and the organic layers were combined and washed with $NaHCO_3$ (2 × 50 mL), brine (50 mL) and dried over Na_2SO_4. The crude alcohol was purified by column chromatography (0-5% MeOH/CH_2Cl_2) to give **41** (5.73 g, 15.9 mmol, 87%) as a pale white foam.

R_f: 0.35 (CH_2Cl_2:MeOH, 95:5); ^{1}H NMR (600 MHz, MeOD) δ 7.69 (q, J = 1.2 Hz, 1H, C(6)H), 7.24 – 7.36 (m, 5H, H-Ar), 5.57 (d, J = 0.8 Hz, 1H, C(1')H), 4.59 – 4.66 (m, 2H, CH_2-OBn), 4.42 (d, J = 0.9 Hz, 1H, C(2')H), 3.98 (d, J = 7.5 Hz, 2H, C(3')H and C(5')H_A or C(7)H_A), 3.93 (d, J = 12.9 Hz, 1H, C(5')H_A or C(7)H_A), 3.90 (d, J = 12.9 Hz, 1H, C(5')H_B or C(7)H_B), 3.77 (d, J = 7.7 Hz, 1H, C(5')H_B or C(7)H_B), 1.87 (d, J = 1.2 Hz, 1H, CH_3-thymine); ^{13}C NMR (151 MHz, MeOD) δ 166.4 (C-2), 151.8 (C-4), 138.9 (C-Ar), 136.7 (C-6), 129.5 (C-Ar), 129.4 (C-Ar), 129.1 (C-Ar), 128.9 (C-Ar), 110.8 (C-5), 89.9 (C-4'), 88.5 (C-1'), 78.3 (C-3'), 76.9 (C-2'), 73.2 (C-7 or C-5'), 73.1, (C-7 or C-5'), 57.6 (CH$_2$-OBn), 12.6 (CH$_3$-thymine); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{18}H_{21}N_2O_6$, 361.1394; found, 361.1395.

1-((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**42**)



41

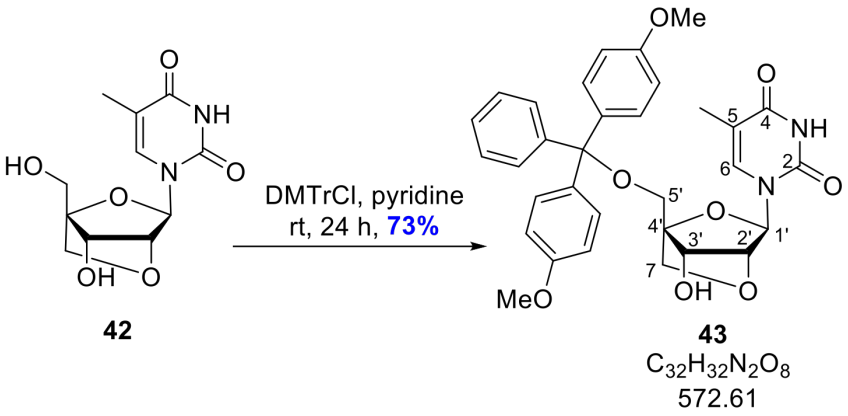
42
$C_{11}H_{14}N_2O_6$
270.24

Compound **41** (5.64 g, 15.6 mmol) was dissolved in MeOH (150 mL) and 20% Pd(OH)$_2$/C (2.60 g, 3.70 mmol) was added and the atmosphere was replaced by H_2 gas by bubbling H_2 gas through the solution for 15 min. The reaction was stirred for 4 h, then filtered over Celite and the solvent was removed under reduced pressure to give **42** (4.07 g, 15.1 mmol, 96%) as a pale white foam.

R_f: 0.30 (CH_2Cl_2:MeOH, 9:1); ^{1}H NMR (600 MHz, MeOD) δ 7.74 (q, J = 1.2 Hz, 1H, C(6)H), 5.54 (d, J = 0.7 Hz, 1H, C(1')H), 4.27 (s, 1H, C(2')H), 4.07 (s, 1H, C(3')H), 3.96 (d, J = 7.7 Hz, 1H, C(7)H_A), 3.91 (d, J = 1.0 Hz, 2H, C(5')H_2), 3.75 (d, J = 7.8 Hz, 1H, C(7)H_B), 1.89 (d, J = 1.3 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, MeOD) δ 166.5

(*C*-2), 151.9 (*C*-4), 136.8 (*C*-6), 110.7 (*C*-5), 90.4 (*C*-4'), 88.3 (*C*-1'), 80.9 (*C*-2'), 72.4 (*C*-7), 70.4 (*C*-3'), 57.6 (*C*-5'), 12.6 (CH_3-thymine); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{11}H_{15}N_2O_6$, 271.0925; found, 271.0919.

1-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**43**)

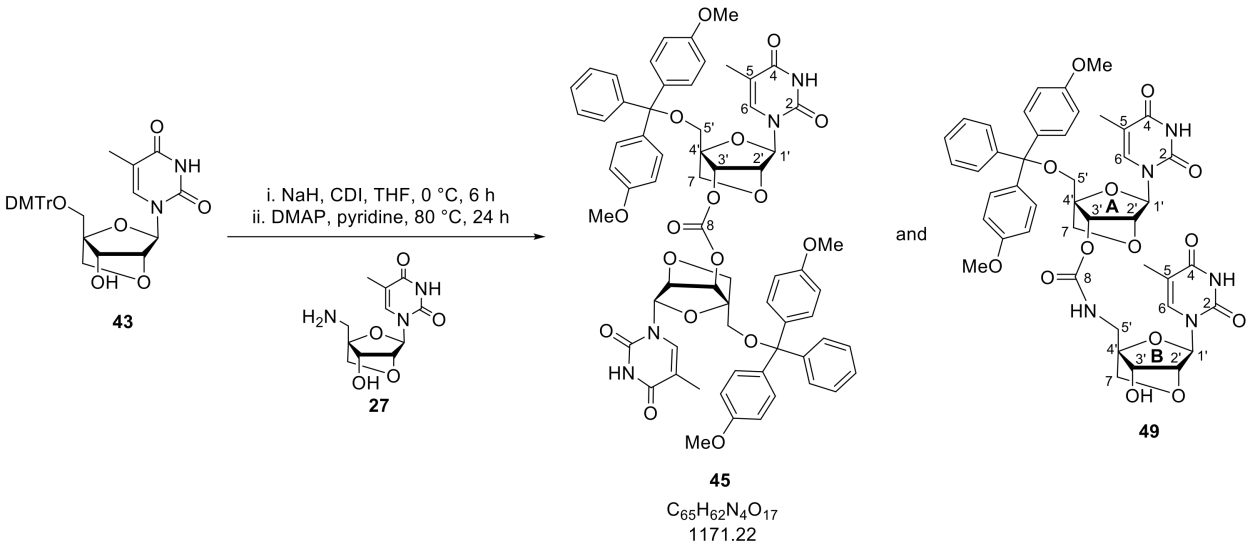


Compound **42** (4.00 g, 14.79 mmol) was dissolved in anhydrous pyridine (100 mL) and DMTrCl (6.26 g, 18.48 mmol) was added portion-wise over 1 h and stirred at rt overnight. Additional DMTrCl (1.80 g, 5.31 mmol) was added and the reaction was stirred for a further 2 h. The solvent was removed under reduced pressure and the residue purified by column chromatography (0-5% MeOH/CH_2Cl_2) to give **43** (6.17 g, 10.77 mmol, 73%) as a white foam.

R_f: 0.45 (CH_2Cl_2:MeOH, 19:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.82 (s, 1H, N*H*), 7.64 (q, J = 1.2 Hz, 1H, C(5)*H*), 7.42 – 7.49 (m, 2H, *H*-Ar), 7.29 – 7.39 (m, 7H, *H*-Ar), 6.80 – 6.91 (m, 4H, *H*-Ar), 5.62 (d, J = 0.7 Hz, 1H, C(1')*H*), 4.44 (s, 1H, C(2')*H*), 4.28 (d, J = 5.6 Hz, 1H, C(3')*H*), 3.88 (d, J = 8.2 Hz, 1H, C(7)H_A), 3.81 (d, J = 8.2 Hz, 1H, C(7)H_B), 3.79 (d, J = 1.0 Hz, 6H, 2 × C*H*$_3$-OMe), 3.57 (d, J = 11.1 Hz, 1H, C(5')H_A), 3.48 (d, J = 11.1 Hz, 1H, C(5')H_B), 2.58 (d, J = 5.9 Hz, 1H, O*H*), 1.69 (t, J = 1.1 Hz, 3H, C*H*$_3$-thymine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 164.0 (*C*-2), 158.9 (*C*-Ar), 149.9 (*C*-2), 144.6 (*C*-Ar), 135.5 (*C*-6), 134.7 (*C*-Ar), 130.2 (*C*-Ar), 129.3 (*C*-Ar), 128.2 (*C*-Ar), 127.3 (*C*-Ar), 113.5 (*C*-Ar), 110.6 (*C*-5), 88.3 (*C*-4'), 87.1 (*C*-1'), 86.9 (*C*-DMTr), 79.5 (*C*-2'), 71.8 (*C*-7), 70.6 (*C*-3'),

58.4 (*C*-5'), 55.4 (*C*H_3-OMe), 12.7 (*C*H_3-thymine); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{32}H_{32}N_2O_8Na$, 595.2051; found, 595.2052.

Bis((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl) carbonate (**45**)

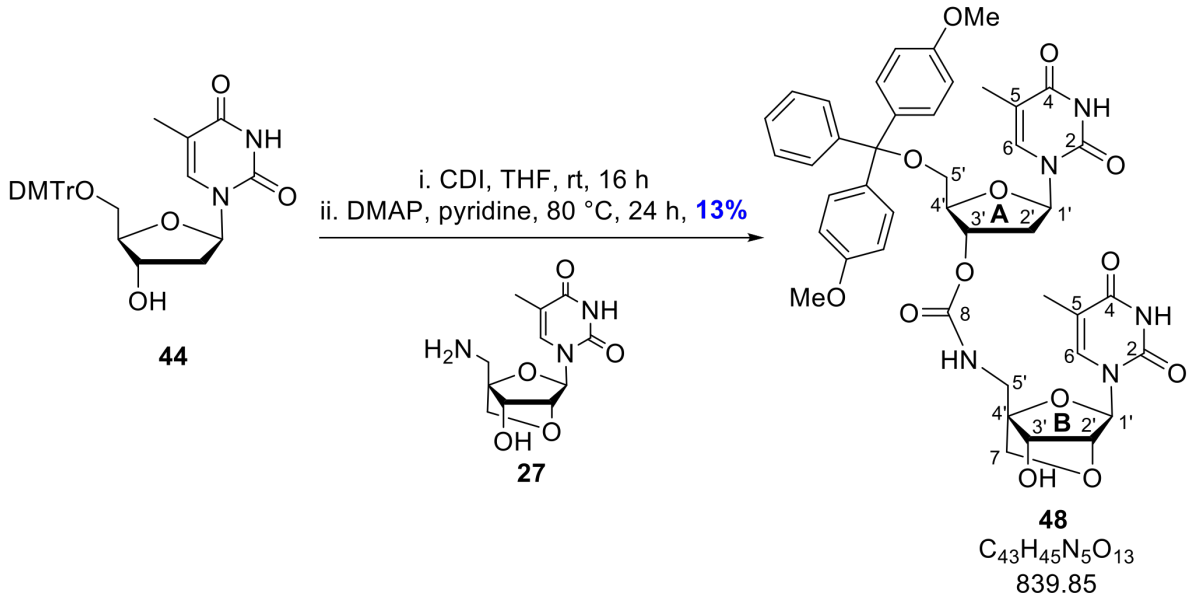


Compound **43** (366 mg, 0.639 mmol) was dissolved in dry THF (6 mL), cooled to 0 °C over ice, and NaH (60 mg, 1.28 mmol) was added and the reaction was stirred for 15 min. Carbonyl diimidazole (CDI) (155 mg, 0.956 mmol) was added. After 4 h, an additional portion of CDI (57 mg, 0.352 mmol) was added. After further 2 h, the THF was removed under reduced pressure and the crude was redissolved in CH_2Cl_2 (10 mL), washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over Na_2SO_4, evaporated under reduced pressure and used immediately without further purification. Compound **27** (200 mg, 0.767 mmol) and DMAP (8 mg, 0.064 mmol) were added and the reaction was heated to 80 °C overnight. A further portion of amine (60 mg, 0.230 mmol) was added, with further DMAP (15 mg, 0.120 mmol) after 1 hour and another portion of DMAP (20 mg, 0.160 mmol) 2.5 h later. After 2 h, the pyridine was removed under reduced pressure and crude residue was dissolved in CH_2Cl_2 (10 mL), washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over Na_2SO_4 and reduced under vacuum. The by-product **45** formed during the activation of compound **43** and was isolated and separated from the desired dimer **49** by column chromatography

(0-5% MeOH/EtOAc).

R_f: 0.90 (EtOAc, MeOH; 19:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.71 (s, 2H, 2 x NH), 7.62 (d, J = 1.4 Hz, 2H, 2 × C(6)H), 7.46 – 7.37 (m, 4H, H-Ar), 7.36 – 7.20 (m, 14H, H-Ar), 6.89 – 6.81 (m, 8H, H-Ar), 5.69 (d, J = 0.6 Hz, 2H, 2 × C(1')H), 4.93 (s, 2H, 2 × C(2' or 3')H), 4.75 (s, 1H, 2 × C(2' or 3')H), 3.95 – 3.82 (m, 4H, 2 × C(7)H_2), 3.79 (d, J = 1.3 Hz, 12H, 4 × CH_3-OMe), 3.53 (d, J = 11.2 Hz, 2H, 2 × C(5')H_A), 3.41 (d, J = 11.2 Hz, 2H, 2 × C(5')H_B), 1.73 (d, J = 1.2 Hz, 6H, 2 × CH_3-thymine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 164.3 (C-8), 158.9 (C-4), 152.8 (C-Ar), 149.9 (C-2), 144.2 (C-Ar), 135.2 (C-Ar), 135.1 (C-Ar), 134.0 (C-6), 130.2 (C-Ar), 130.1 (C-Ar), 128.3 (C-Ar), 128.2 (C-Ar), 127.4 (C-Ar), 113.6 (C-Ar), 111.1 (C-5), 87.4 (C-4' or C-DMTr), 87.1 (C-1'), 86.9 (C-4' or C-DMTr), 77.8 (C-2' or 3'), 74.9 (C-2' or 3'), 72.2 (C-7), 57.9 (C-5'), 55.4 (CH$_3$-OMe), 12.9 (CH$_3$-thymine); LRMS-ESI (m/z): $[M+H]^+$ 1171 (11%), $[M+Na]^+$ 1193 (84%), $[M+Na]^+$ 1194 (79%); HRMS-ESI+ (m/z): $[M+Na]^+$ calcd. for $C_{65}H_{62}N_4O_{17}Na$, 1193.4002; found, 1193.4019.

(2*R*,3*S*,5*R*)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)tetrahydrofuran-3-yl (((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl)carbamate (**48**)[a]

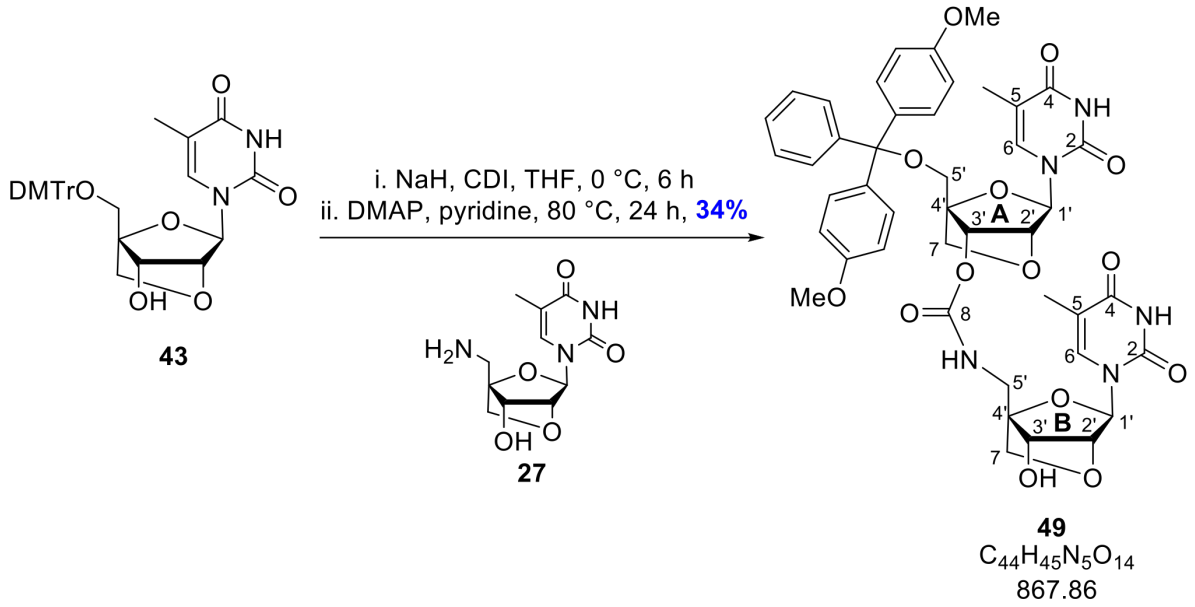


Compound **44** (500 mg, 0.918 mmol) was dissolved in anhydrous THF (10 mL) and CDI (0.223 g, 1.38 mmol) was added and the reaction was stirred at rt for 16 h. The solvent was removed under reduced pressure and the crude was re-dissolved in EtOAc (15 mL), washed with a KH_2PO_4 solution (5% w/v, 20 mL), sat. aqueous KCl (20 mL), dried over Na_2SO_4 and concentrated under reduced pressure to give an intermediate electrophile which was used without further purification. The electrophile was dissolved in anhydrous pyridine (12 mL) and a solution of amine **27** (0.310 g, 1.15 mmol) in anhydrous pyridine (12 mL) was added dropwise over the course of an hour. DMAP (14.6 mg, 0.12 mmol) was added, the reaction was heated to 80 °C, and stirred overnight. The solvent was removed under reduced pressure and the crude was redissolved in CH_2Cl_2 (20 mL), which washed with sat. aqueous $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by column chromatography (0-5% MeOH/EtOAc) to give dimer **48** (0.132 g, 0.157 mmol, 13%) as a white solid.

[a]This compound was synthesised by Autumn Usher, Part II.

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.36 (d, J = 9.1 Hz, 2H, 2 × NH), 7.77 (t, J = 6.2 Hz, 1H, NH-carbamate), 7.52 (d, J = 1.4 Hz, 1H, C(6A or 6B)H), 7.42 (d, J = 1.5 Hz, 1H, C(6A or 6B)H), 7.40 – 7.19 (m, 9H, H-Ar), 6.88 (dd, J = 9.0, 1.8 Hz, 4H, H-Ar), 6.24 (dd, J = 8.6, 5.9 Hz, 1H, C(A1')H), 5.55 (d, J = 4.1 Hz, 1H, C(B3')OH), 5.41 (s, 1H, C(B1')H), 5.27 (d, J = 6.2 Hz, 1H, C(A3')H), 4.12 (s, 1H, C(B2')H), 4.05 (d, J = 2.4 Hz, 1H, C(A4')H), 3.89 – 3.82 (m, 2H, C(7)H_{A}, C(B3')H), 3.73 (s, 6H, 2 × CH_3-OMe), 3.62 (d, J = 8.0 Hz, 1H, C(7)H_{B}), 3.54 (t, J = 6.2 Hz, 2H, C(B5')H_2), 3.31 (d, J = 3.5 Hz, 1H, C(A5')H_{A}), 3.21 (dd, J = 11.0, 7.6 Hz, 1H, C(A5')H_{B}), 2.44 (s, 1H, C(A2')H_{A}), 2.29 (dd, J = 13.3, 5.9 Hz, 1H, C(A2')H_{B}), 1.76 (d, J = 1.2 Hz, 3H, CH_3-thymine(A or B)), 1.41 (d, J = 1.1 Hz, 3H, CH_3-thymine(A or B)); ^{13}C NMR (151 MHz, DMSO-d_6) δ 163.8 (C-A4 or B4), 163.6 (C-A4 or B4), 158.2 (C-Ar), 158.2 (C-Ar), 155.8 (C-8), 150.4 (C-A2 or B2), 149.9 (C-A2 or B2), 144.6 (C-Ar), 135.4 (C-A6 or B6), 135.3 (C-Ar), 135.1 (C-Ar), 134.6 (C-A6 or B6), 129.7 (C-Ar),127.9 (C-Ar), 127.7 (C-Ar), 126.9 (C-Ar), 113.3 (C-Ar), 109.9 (C-A5 or B5), 108.5 (C-A5 or B5), 87.1 (C-B4'), 86.3 (C-B1'), 86.1 (C-DMTr), 83.6 (C-A1'), 83.3 (C-A4'), 79.0 (C-B2'), 74.6 (C-A3'), 71.4 (C-7), 69.9 (C-B3'), 63.7 (C-A5'), 55.0 (CH$_3$-OMe), 37.2 (C-B5'), 36.7 (C-A2'), 12.4 (CH$_3$-thymine(A or B)), 11.6 (CH$_3$-thymine(A or B)); HRMS-ESI (m/z): [M+Na]$^+$ calcd. for $C_{43}H_{45}N_5O_{13}Na$, 862.2912; found, 862.2934.

(1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl) carbamate (**49**)

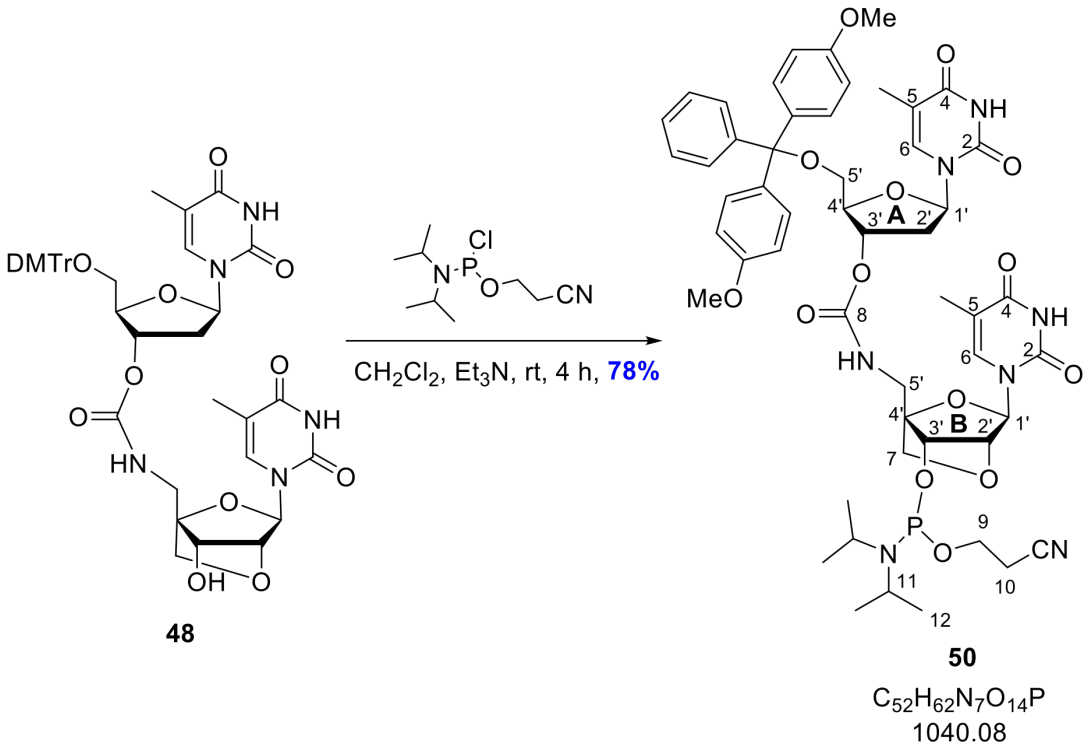


Compound **43** (366 mg, 0.639 mmol) was dissolved in dry THF (6 mL), cooled to 0 °C over ice, and NaH (60 mg, 1.28 mmol) was added and the reaction was stirred for 15 min. Carbonyl diimidazole (CDI) (155 mg, 0.956 mmol) was added. After 4 h, an additional portion of CDI (57 mg, 0.352 mmol) was added. After further 2 h, the THF was removed under reduced pressure and the crude was redissolved in CH_2Cl_2 (10 mL), washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over Na_2SO_4, evaporated under reduced pressure and used immediately without further purification. The electrophile was co-evaporated in anhydrous pyridine (2 × 5 mL) then dissolved in anhydrous pyridine (5 mL). Compound **27** (200 mg, 0.767 mmol) and DMAP (8 mg, 0.064 mmol) were added and the reaction was heated to 80 °C overnight. A further portion of amine (60 mg, 0.230 mmol) was added, with further DMAP (15 mg, 0.120 mmol) after 1 hour and another portion of DMAP (20 mg, 0.160 mmol) 2.5 h later. After 2 h, the pyridine was removed under reduced pressure and crude residue was dissolved in CH_2Cl_2 (10 mL), washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over Na_2SO_4 and reduced under vacuum. The crude residue was

purified by column chromatography (0-5% MeOH/EtOAc) to give dimer **49** (190 mg, 0.219 mmol, 34%) as a white powder.

R_f: 0.25 (MeOH:EtOAc; 1:19); ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.50 (s, 1H, N*H*), 11.37 (s, 1H, N*H*), 7.56 (d, J = 2.3 Hz, 1H, C(B6)*H*), 7.36 – 7.44 (m, 2H, C(A6)*H*), 7.19 – 7.32 (m, 7H, *H*-Ar), 6.90 (ddd, J = 8.9, 4.4, 2.1 Hz, 4H, *H*-Ar), 5.61 (d, J = 4.1 Hz, 1H, C(3'B)O*H*), 5.53 (s, 1H, C(A1')*H*), 5.39 (s, 1H, C(B1')*H*), 5.17 (s, 1H, C(A3')*H*), 4.46 (s, 1H, C(A2')*H*), 4.10 (s, 1H, C(B3')*H*), 3.81 (m, 3H, C(B2')*H* and C(A8)H_2), 3.71 – 3.74 (m, 6H, 2 × CH_3-OMe), 3.45 – 3.63 (m, 4H, C(B8)H_2 and C(B5')H_2), 3.36 – 3.44 (m, 2H, C(A5')H_2), 1.72 (s, 3H, CH_3-thymine(B)), 1.55 (s, 3H, CH_3-thymine(A)); ^{13}C NMR (151 MHz, DMSO-d_6) δ 163.8 (*C*-2), 163.7 (*C*-2), 158.2 (*C*-Ar), 155.0 (*C*-8), 149.9 (*C*-4), 149.9 (*C*-4), 144.5 (*C*-Ar), 135-134.7 (*C*-Ar and C-A6), 133.8 (*C*-B6), 129.7 (*C*-Ar), 128.0 (*C*-Ar), 127.6 (*C*-Ar), 126.9 (*C*-Ar), 113.3 (*C*-Ar), 108.9 (*C*-B5), 108.5 (*C*-A5), 87.0 (*C*-A1'), 86.4 (*C*-B1'), 85.9 (*C*-DMTr), 79.0 (*C*-B3'), 77.5 (*C*-A2'), 71.7 (*C*-A8), 71.4 (*C*-B8), 70.8 (*C*-A3'), 70.2 (*C*-B2'), 58.0 (*C*-A5'), 55.0 (*C*H$_3$-OMe), 55.0 (*C*H$_3$-OMe), 37.7 (*C*-B5'), 12.4 (*C*H$_3$-thymine(B)), 12.2 *C*H$_3$-thymine(A)); LRMS (ESI+) *m/z* (%): 303 (45%) [DMTr]$^+$, 890 (100%) [M+Na]$^+$, 891 (36%); HRMS-ESI (*m/z*): [M]$^+$ calcd. for $C_{44}H_{45}N_5O_{14}$, 867.2958, found 867.2962; [M+Na]$^+$ calcd. for $C_{44}H_{45}N_5O_{14}Na$, 890.2855, found 890.2838.

(2*R*,3*S*,5*R*)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)tetrahydrofuran-3-yl (((1*S*,3*R*,4*R*,7*S*)-7-(((2-cyanoethoxy)(diisopropylamino)phosphaneyl)oxy)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo-[2.2.1]heptan-1-yl)methyl)carbamate (**50**)

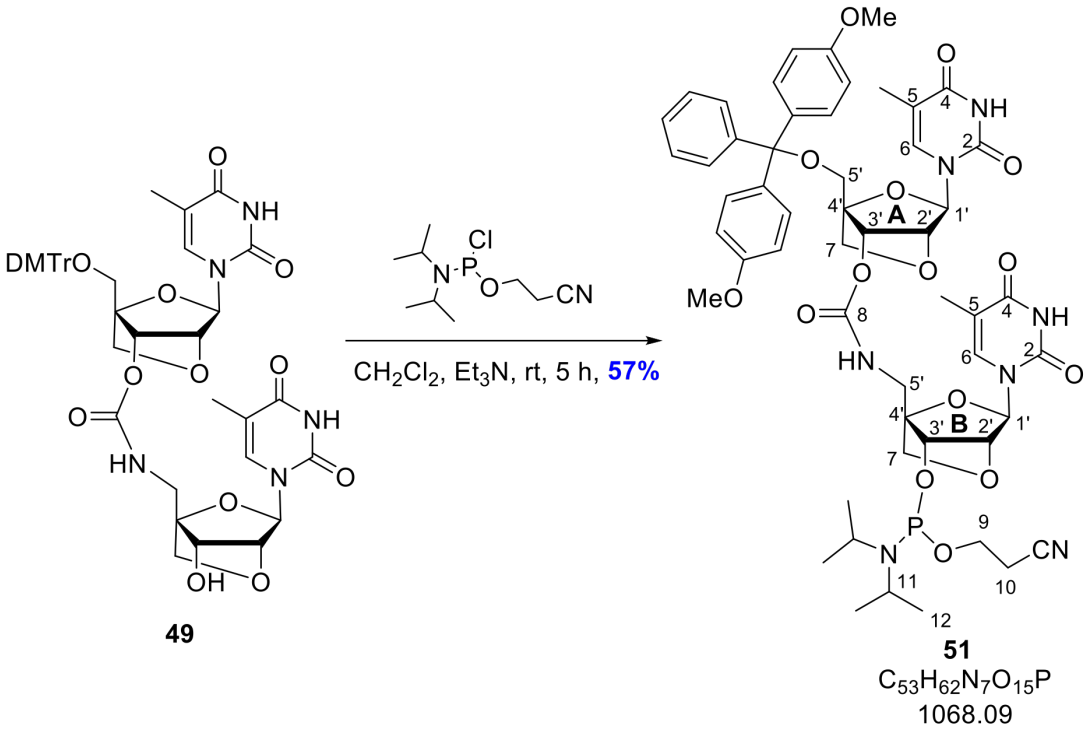


Compound **48** (178 mg, 0.212 mmol) was dissolved in anhydrous, degassed CH_2Cl_2 (3 mL) and anhydrous, degassed Et_3N (74 µL, 0.531 mmol) and chloro-phosphitylating reagent (71 µL, 0.318 mmol) were added and the reaction was stirred in an inert atmosphere at rt for 1.5 h. Additional portions of Et_3N (25 µL, 0.180 mmol) and chloro-phosphitylating reagent (25 µL, 0.112 mmol) were added once every 1.5 h for the next 3 h. After 1 hour, an inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was purified by inert column chromatography (100% EtOAc, 0.5% Et_3N) to give phosphoramidite **50** (170 mg, 0.172 mmol, 78%) as a white solid.

^{31}P NMR (162 MHz, DMSO-d_6) δ 148.26, 147.43; HRMS-ESI$^+$ (m/z): $[M+H]^+$ calcd. for

$C_{52}H_{63}N_7O_{14}P$, 1040.4165; found, 1040.4178.

(1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (((1*S*,3*R*,4*R*,7*S*)-7-(((2-cyanoethoxy)(diisopropylamino)phosphaneyl)oxy)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl)carbamate (**51**)

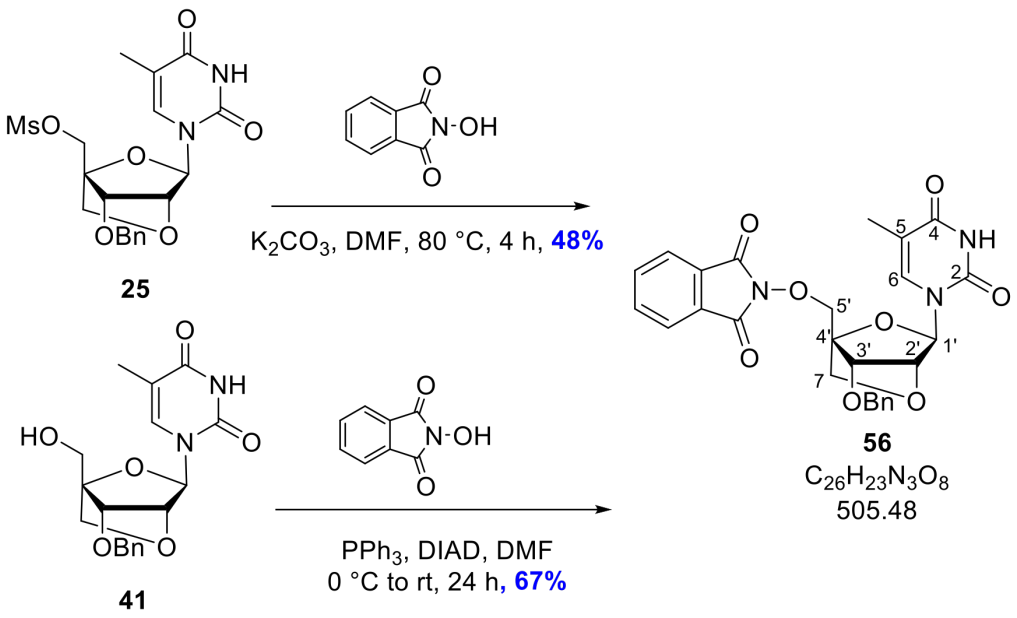


Compound **49** (160 mg, 0.191 mmol) was dissolved in anhydrous, degassed CH_2Cl_2 (2.5 mL) and anhydrous, degassed Et_3N (65 µL, 0.478 mmol) and chloro-phosphitylating reagent (64 µL, 0.287) were added and stirred at rt under an inert atmosphere. After 2 h, additional portions of Et_3N (20 µL) and chloro-phosphitylating reagent (20 µL) were added. After 3 h, an inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was purified by inert column chromatography (100% EtOAc, 0.5% Et_3N) to give phosphoramidite **51** (116 mg, 0.109 mmol, 57%) as a white foam.

R_f: 0.55 (EtOAc); ^{31}P NMR (162 MHz, DMSO-d_6) δ 148.08, 147.46; ^{1}H NMR (400 MHz,

DMSO-d_6) δ 11.48 (s, 1H), 11.37 (s, 1H), 7.98 (t, J = 6.2 Hz, 1H), 7.56 (dd, J = 2.7, 1.4 Hz, 1H), 7.46 (dd, J = 4.5, 1.4 Hz, 1H), 7.43 – 7.39 (m, 2H), 7.36 – 7.20 (m, 7H), 6.90 (ddd, J = 9.0, 3.9, 1.3 Hz, 4H), 5.53 (d, J = 2.1 Hz, 1H), 5.45 (d, J = 13.2 Hz, 1H), 5.17 (d, J = 19.2 Hz, 1H), 4.45 (d, J = 27.8 Hz, 1H), 4.34 (d, J = 14.2 Hz, 1H), 4.03 – 3.91 (m, 1H), 3.73 (p, J = 3.6 Hz, 9H), 3.67 – 3.47 (m, 4H), 3.41 (d, J = 3.7 Hz, 2H), 2.81 – 2.67 (m, 2H), 1.77 – 1.72 (m, 3H), 1.61 – 1.49 (m, 3H), 1.13 – 1.08 (m, 8H), 1.05 (d, J = 6.6 Hz, 7H); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{53}H_{63}N_7O_{15}P$, 1068.4114; found, 1068.4134; $[M+Na]^+$ calcd. for $C_{53}H_{62}N_7O_{15}PNa$, 1090.3934; found, 1090.4414. NB: Assignments for ^{1}H NMR were not possible due to compound instability.

2-(((1*R*,3*R*,4*R*,7*S*)-7-(benzyloxy)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy)isoindoline-1,3-dione (**56**)



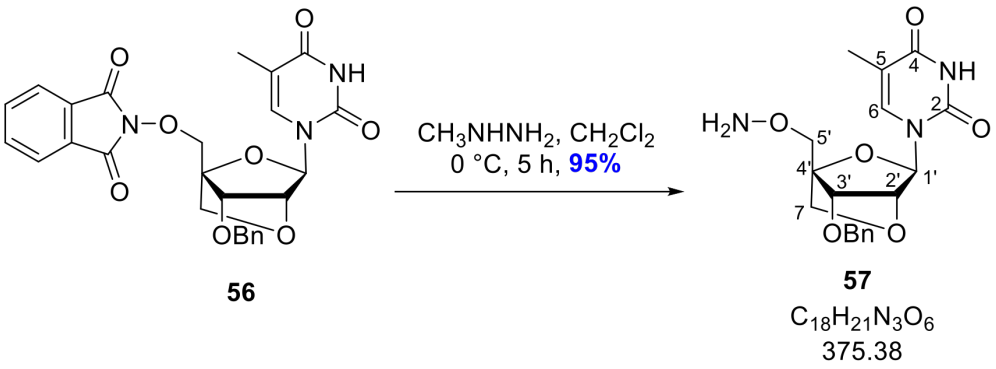
Method A (S_N2): Compound **25** (1.13 g, 2.57 mmol) was dissolved in DMF (25 mL) and K_2CO_3 (1.07 g, 7.71 mmol) and *N*-hydroxyphthalimide (545 mg, 3.34 mmol) were added. The reaction was heated to 80 °C for 4 h; the reaction was cooled to rt, partitioned between EtOAc (25 mL) and water (25 mL) and neutralised with 1 M HCl. The product was extracted from the aqueous layer with EtOAc (2 × 25 mL), the organic layers combined and washed with $NaHCO_3$ (25 mL), brine (25 mL), dried over Na_2SO_4, and the solvent removed

under reduced pressure. The residue was purified by column chromatography (50-100% EtOAc/Pet. Ether) to give **56** (621 mg, 1.23 mmol, 48%) as a white foam.

Method B (Mitsunobu): Compound **41** (590 mg, 1.64 mmol) was dissolved in anhydrous DMF (10 mL) and triphenylphosphine (558 mg, 2.13 mmol) and *N*-hydroxyphthalimide (347 mg, 2.13 mmol) were added and the reaction was cooled to 0 °C on ice. Diisopropyl azodicarboxylate (DIAD) (419 µL, 2.13 mmol) was added drop-wise over 1 h and the reaction was brought to rt and stirred overnight. The solvent was removed under reduced pressure and the compound was purified by column chromatography (50-100% EtOAc/Pet. Ether) to give **56** (551 mg, 1.09 mmol, 67%) as a white foam.

R_f 0.50 (EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (s, 1H, N*H*), 7.78 – 7.92 (m, 4H, *H*-Ar(Phth)), 7.75 (q, J = 1.2 Hz, 1H, C(6)*H*), 7.22 – 7.36 (m, 5H, *H*-Ar(OBn)), 5.65 (d, J = 0.9 Hz, 1H, C(1')*H*), 4.63 – 4.79 (m, 3H, CH_2-OBn and C(5')H_A), 4.48 – 4.61 (m, 2H, C(3')*H* and C(5')H_B), 4.23 (s, 1H, C(2')*H*), 4.09 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.91 (d, J = 7.8 Hz, 1H, C(7)H_B),), 1.93 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 163.6 (*C*-4), 163.1 (*C*O-Phth), 149.6 (*C*-2), 136.8 (*C*-Ar(OBn)), 135.1 (*C*-6), 134.9 (*C*-Ar(Phth)), 128.7 (*C*-Ar(Phth)), 128.5 (*C*-Ar(OBn)), 128.2 (*C*-Ar(OBn)), 127.8 (*C*-Ar(OBn)), 123.9 (*C*-Ar(Phth)), 110.6 (*C*-5), 87.6 (*C*-1'), 85.9 (*C*-4'), 77.2 (*C*-3'), 76.1 (*C*-2'), 73.0 (*C*-5'), 72.6 (*C*H_2-OBn), 72.0 (*C*-7), 12.6 (*C*H_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 506 (100%), $[M+Na]^+$ 538 (40%), $[2M+Na]^+$ 1035 (54%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{26}H_{23}N_3O_8$, 506.1558; found, 506.1560.

1-((1*R*,3*R*,4*R*,7*S*)-1-((aminooxy)methyl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**57**)



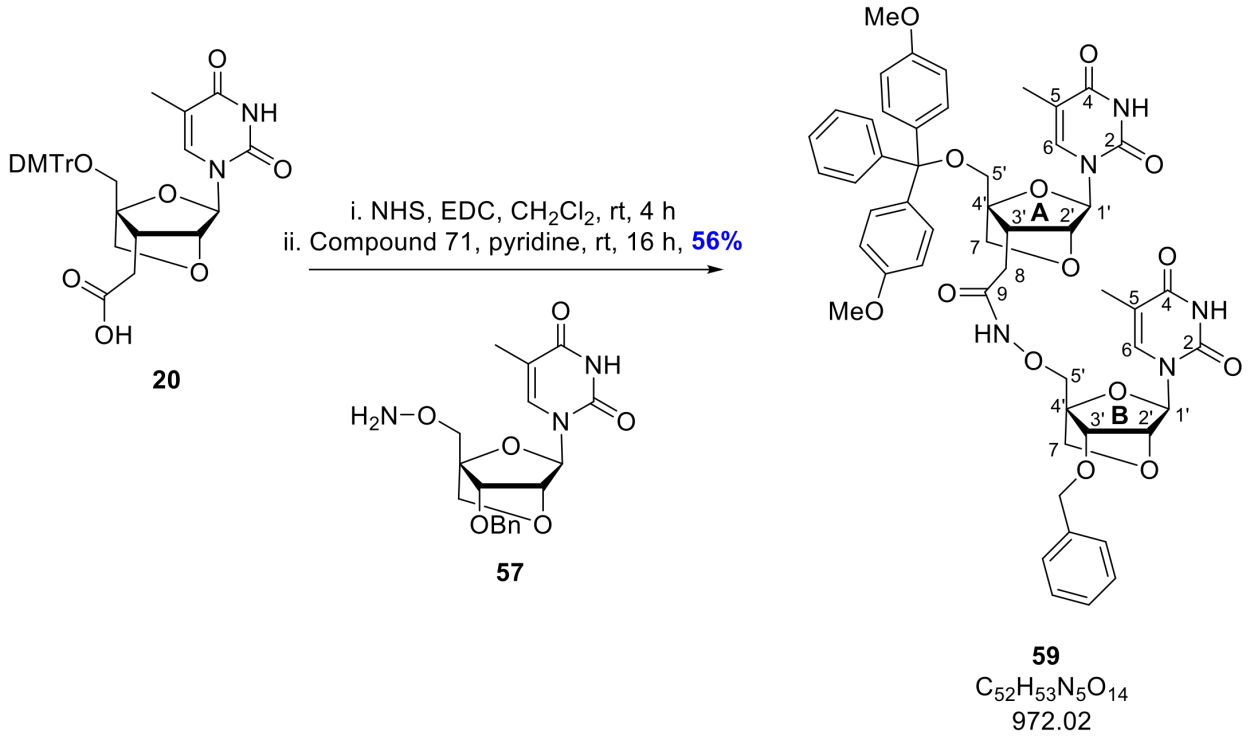
56

57
$C_{18}H_{21}N_3O_6$
375.38

Compound **56** (343 mg, 0.679 mmol) was dissolved in anhydrous CH_2Cl_2 (5 mL) and cooled to 0 °C on ice. Methylhydrazine (43 µL, 0.814 mmol) was added and the reaction was stirred for 2 h. Additional portions of methylhydrazine (21 µL, 0.407 mmol) were added every hour for the next 3 h. The reaction was brought to rt and the solvent removed under reduced pressure. The residue was purified by column chromatography (0-5% MeOH/CH_2Cl_2) to give **57** (242 mg, 0.644 mmol, 95%) as a white foam.

R_f: 0.35 (MeOH:CH_2Cl_2; 1:19); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.53 (s, 1H, N*H*), 7.37 (d, J = 1.4 Hz, 1H, C(6)*H*), 7.28 – 7.36 (m, 5H, *H*-Ar), 5.62 (s, 1H, C(1')*H*), 4.66 (d, J = 11.6 Hz, 1H, CH_A-OBn), 4.54 (t, J = 5.8 Hz, 2H, C(3')*H* and CH_B-OBn), 4.02 – 4.14 (m, 3H, C(5')H_2 and C(7)H_A), 3.89 (d, J = 8.0 Hz, 1H, C(7)H_B), 3.83 (s, 1H, C(2')*H*), 1.92 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 163.5 (*C*-4), 163.5 (*C*-2), 149.6 (*C*-Ar), 137.0 (*C*-Ar), 134.7 (*C*-6), 128.7 (*C*-Ar), 128.6 (*C*-Ar), 128.4 (*C*-Ar), 127.9 (*C*-Ar), 110.4 (*C*-5), 87.7 (*C*-4'), 86.9 (*C*-1'), 76.8 (*C*-3'), 76.6 (*C*-2'), 72.5 (CH_2-OBn or *C*-7), 72.4 (CH_2-OBn or *C*-7), 71.0 (*C*-5'), 12.9 (CH_3-thymine); LRMS-ESI m/z (%): $[M+H]^+$ 376 (90%), $[2M+H]^+$ 751 (100%), $[3M+H]^+$ 1126 (94%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{18}H_{22}N_3O_6$, 376.1503; found, 376.1504; $[M+Na]^+$ calcd. for $C_{18}H_{21}N_3O_6Na$, 398.1323, found, 398.1324.

N-(((1*R*,3*R*,4*R*,7*S*)-7-(benzyloxy)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy)-2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo-[2.2.1]heptan-7-yl)acetamide (**59**)

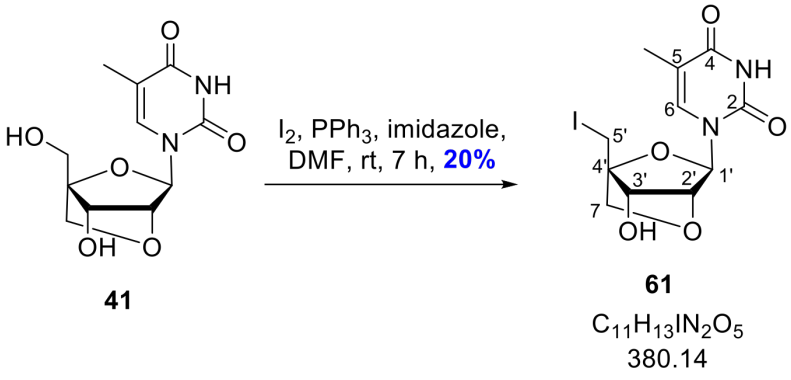


Compound **20** (150 mg, 0.244 mmol) was dissolved in anhydrous CH_2Cl_2 (3 mL) and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (70 mg, 0.366 mmol), *N*-hydroxysuccinimide (NHS) (31 mg, 0.268 mmol) were added. The reaction was stirred at rt for 2 h, after which the organic layer was washed with half sat. $NaHCO_3$ (2 × 10 mL), brine (10 mL), dried over Na_2SO_4 and the solvent was removed under reduced pressure. The intermediate activated ester **59** was used without further purification and dissolved in anhydrous pyridine (4 mL); compound **57** (100 mg, 0.268 mmol) was added and the reaction was stirred at rt overnight. The solvent was removed under reduced pressure and the residue was purified by column chromatography (20-100% EtOAc/Pet. Ether, then 0-2% MeOH/EtOAc) to give dimer **59** (133 mg, 0.137 mmol, 56%) as a white foam.

R_f: 0.25 (EtOAc); ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.46 (s, 1H, N*H*), 11.37 (s, 1H, N*H*),

7.67 – 7.53 (m, 5H, *H*-Ar(OBn)), 7.46 – 7.41 (m, 2H, C(6A)*H* and C(6B)*H*), 7.37 – 7.22 (m, 9H, *H*-Ar(DMTr)), 6.95 – 6.87 (m, 4H, *H*-Ar(DMTr)), 5.50 (d, J = 3.8 Hz, 1H, C(1'A)*H* and C(1'B)*H*), 4.60 (s, 2H, CH_2-OBn), 4.45 (d, J = 3.2 Hz, 2H, C(3'B)*H* and C(2'A)*H*), 4.35 (d, J = 11.6 Hz, 1H, C(7A)H_A), 4.12 (d, J = 11.5 Hz, 1H, C(7A)H_B), 4.04 (s, 1H, C(3'A)*H*), 3.92 (d, J = 8.0 Hz, 1H, C(5'A)H_A), 3.80 (d, J = 8.0 Hz, 1H, C(5'A)H_B), 3.74 (d, J = 2.3 Hz, 6H, 2 × CH_3-OMe), 3.66 (d, J = 8.5 Hz, 1H, C(5'B)H_A), 3.61 (d, J = 8.5 Hz, 1H, C(5'B)H_B), 3.51 (d, J = 11.3 Hz, 1H, C(7B)H_A), 3.35 (s, 1H, C(7B)H_B), 2.03 – 1.86 (m, 2H, C(8)H_2), 1.77 – 1.72 (m, 3H, CH_3-thymine(A or B)), 1.59 – 1.56 (m, 3H, CH_3-thymine(A or B)); ^{13}C NMR (151 MHz, DMSO-d$_6$) δ 167.4 (*C*-9), 163.9 (*C*-4A or 4B), 163.7 (*C*-4A or 4B), 158.2 (*C*-Ar), 149.9 (*C*-2A or 2B), 149.9 (*C*-2A or 2B), 144.6 (*C*-Ar(OBn)), 137.7 (*C*-Ar(OBn)), 135.2 (*C*-Ar(DMTr)), 135.1 (*C*-6A or 6B), 134.8 (*C*-6A or 6B), 134.2 (*C*-Ar), 133.1 (*C*-Ar), 132.4 (*C*-Ar), 132.0 (*C*-Ar), 132.0 (*C*-Ar), 131.5 (*C*-Ar), 131.4 (*C*-Ar), 129.7 (*C*-Ar), 129.7 (*C*-Ar), 128.8 (*C*-Ar), 128.7 (*C*-Ar), 128.2 (*C*-Ar), 128.2 (*C*-Ar), 127.9 (*C*-Ar), 127.6 (*C*-Ar), 127.5 (*C*-Ar), 126.9 (*C*-Ar), 113.3 (*C*-Ar(DMTr)), 108.6 (*C*-5A or 5B), 108.5 (*C*-5A or 5B), 89.3 (*C*-A4'), 86.6 (*C*-1'A or 1'B), 86.6 (*C*-1'A or 1'B), 85.8 (*C*-B4'), 85.4 (*C*-DMTr), 79.5 (*C*-3'B or 2'A), 76.5 (*C*-B3' or A2' or A3'), 76.1 (B3' or A2' or A3'), 71.5 (*C*-B5' or A5' or 7A), 71.3 (*C*-B5' or A5' or 7A), 71.2 (*C*-B5' or A5' or 7A), 58.6 (*C*-7B), 55.0 (*C*H$_3$-OMe), 26.6 (*C*-8), 12.3 (*C*H$_3$-thymine(A or B)), 12.1 (*C*H$_3$-thymine(A or B)); LRMS-ESI (*m/z*): [M+Na]$^+$ 994 (100%), [M-DMTr+Na]$^+$ 692 (50%); HRMS-ESI (*m/z*): [M+Na]$^+$calcd. for $C_{52}H_{53}N_5O_{14}Na$, 994.3481; found, 994.3494.

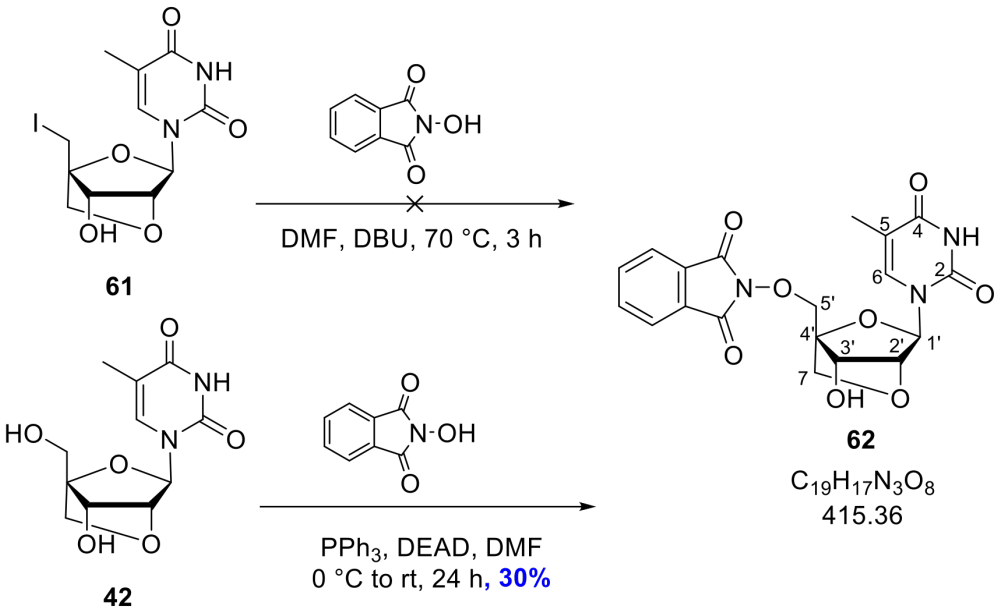
1-((1*S*,3*R*,4*R*,7*S*)-7-hydroxy-1-(iodomethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-pyrimidine-2,4(1*H*,3*H*)-dione (**61**)



Compound **41** (160 mg, 0.592 mmol) was dissolved in DMF (2.5 mL) and triphenylphosphine (233 mg, 0.888 mmol), iodine (225 mg, 0.888 mmol), and imidazole (121 mg, 1.78 mmol) were added and the reaction was stirred at rt. After 5 h, additional portions of triphenylphosphine (95 mg, 0.362 mmol), iodine (106 mg, 0.418 mmol), and imidazole (70 mg, 1.03 mmol) were added. After 1.5 h, the reaction was diluted in EtOAc (10 mL) and washed with 10% $Na_2S_2O_3$ (10 mL), brine (10 mL), dried over $MgSO_4$ and reduced under pressure. The crude residue was purified by column chromatography (0-20% MeOH/CH_2Cl_2) to give **61** (44 mg, 0.116 mmol, 20%) as a white foam.

R_f: 0.35 (MeOH:CH_2Cl_2; 1:9); ^{1}H NMR (400 MHz, MeOD) δ 7.98 (q, J = 1.2 Hz, 1H, C(6)H), 5.59 (s, 1H, C(1')H), 4.35 (s, 1H, C(2')H), 4.05 (d, J = 8.0 Hz, 1H, C(7)H_A), 3.96 (s, 1H, C(3')H), 3.87 (d, J = 7.9 Hz, 1H, C(7)H_B), 3.64 (q, J = 11.7 Hz, 2H, C(5')H_2), 1.92 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, MeOD) δ 166.5 (C-4), 151.8 (C-2), 137.0 (C-6), 110.8 (C-5), 88.7 (C-4'), 87.9 (C-1'), 81.1 (C-2'), 74.2 (C-3'), 73.4 (C-7), 12.6 (CH$_3$-thymine), −1.22 (C-5'); LRMS (ESI+) m/z (%): 413 (100%) $[M+CH_3OH+H]^+$, 887 (72%) $[2M+2ACN+Na]^+$; HRMS-ESI (m/z): $[M]^+$ calcd. for $C_{11}H_{13}IN_2O_5$, 379.9864; found, 380.2152; $[M+H]^+$ calcd. for $C_{11}H_{14}IN_2O_5$, 380.9942; found, 380.9940; $[M+Na]^+$ calcd. for $C_{11}H_{13}IN_2O_5Na$, 402.9762; found, 402.9761

2-(((1*R*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy)isoindoline-1,3-dione (**62**)

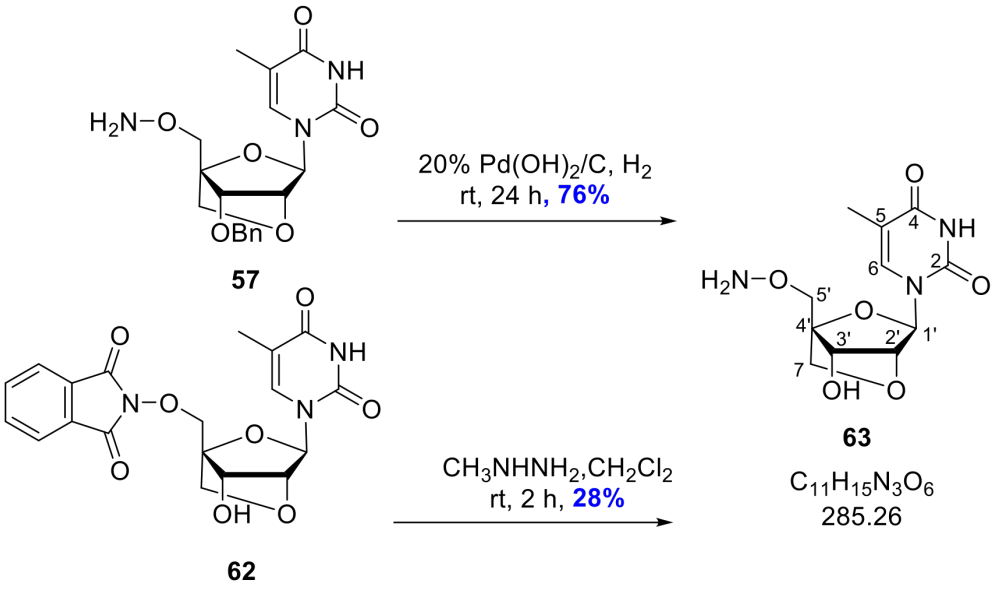


Method A (S_N2): Compound **61** (40 mg, 0.105 mmol) was dissolved in DMF (1 mL) and *N*-hydroxyphthalimide (34 mg, 0.210 mmol) and DBU (31 µL, 0.210 mmol) were added and the reaction was heated to 70 °C for 3 h. The reaction was cooled to rt and the solvent removed under reduced pressure. The residue was redissolved in EtOAc (2 mL), washed with sat. $NaHCO_3$ (3 mL), brine (3 mL), dried over $MgSO_4$ and reduced under reduced pressure. The crude residue was purified by column chromatography (0-5% MeOH/CH_2Cl_2), but unfortunately only starting material was recovered.

Method B (Mitsunobu): Compound **42** (184 mg, 0.68 mmol) was dissolved in DMF (7 mL), and N-hydroxyphthalimide (122 mg, 0.75 mmol), triphenylphosphine on resin (469 mg, 0.75 mmol), and DEAD (118 µL, 0.75 mmol) were added and the reaction was stirred at rt overnight. The resin was filtered off and the reaction was diluted with EtOAc (5 mL), washed with sat. $NaHCO_3$ (10 mL), brine (10 mL), dried over Na_2SO_4, and reduced under pressure. The crude residue was purified by column chromatography (0-5% MeOH/CH_2Cl_2) to give **62** (257 mg, 30% product by weight) as a white powder contaminated with triphenylphosphine oxide (TPPO).

R_f: 0.45 (MeOH: CH_2Cl_2; 1:19); ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.38 (s, 1H, NH), 7.93 – 7.80 (m, 4H, H-Ar), 7.73 (d, J = 1.3 Hz, 1H, C(6)H), 7.68 – 7.51 (m, 52H, H-Ar TPPO), 5.90 (d, J = 4.4 Hz, 1H, C(3')OH), 5.49 (s, 1H, C(1')H), 4.86 (d, J = 11.9 Hz, 1H, C(5')H_A), 4.54 (d, J = 12.0 Hz, 1H, C(5')H_B), 4.23 (s, 1H, C(2')H), 4.11 (d, J = 4.6 Hz, 1H, C(3')H), 4.03 (d, J = 7.6, 1H, C(7)H_A), 3.88 (d, J = 8.0 Hz, 1H, C(7)H_B), 1.81 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 163.9 (C-4), 162.9 (C-9), 156.6 (C-Ar TPPO), 156.2 (C-Ar TPPO), 155.3 (C-Ar TPPO), 151.8 (C-Ar TPPO), 149.9 (C-2), 134.9 (C-6), 134.8 (C-Ar), 128.6 (C-Ar), 123.3 (C-Ar), 108.7 (C-5), 86.6 (C-4'), 86.1 (C-1'), 79.0 (C-2'), 73.8 (C-5'), 70.9 (C-7), 69.5 (C-3'), 12.2 (CH$_3$-thymine); HRMS-ESI (m/z): $[M]^+$ calcd. for $C_{19}H_{17}N_3O_8$, 415.1010, found 415.2713; $[M+H]^+$ calcd. for $C_{19}H_{18}N_3O_8$, 416.1088, found 416.1097.

1-((1R,3R,4R,7S)-1-((aminooxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-pyrimidine-2,4(1H,3H)-dione (**63**)

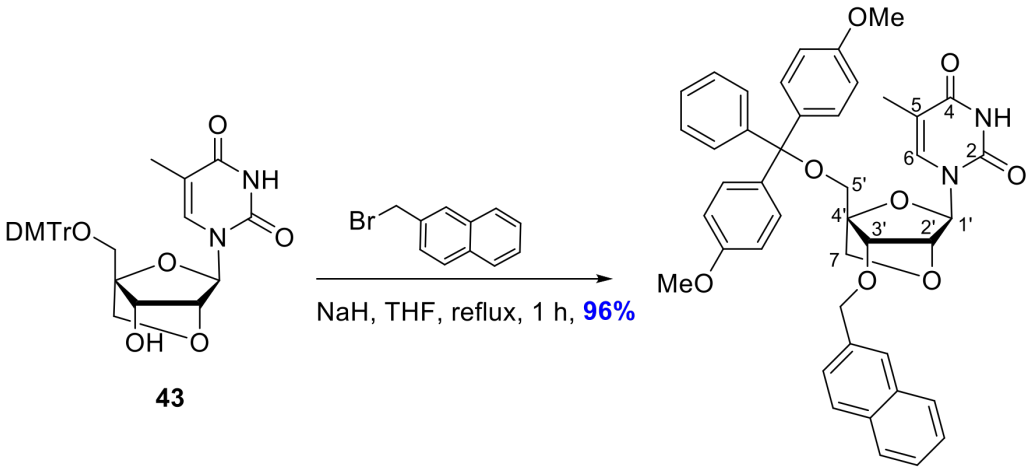


Method A: Compound **57** (76 mg, 0.208 mmol) was dissolved in MeOH (5 mL) and 20% $Pd(OH)_2/C$ (15 mg, 10 mol%) was added. H_2 gas was bubbled through the solution until the flask's atmosphere was completely replaced by H_2 3-5 times. The reaction was left to stir at rt overnight and filtered over Celite. The residue was purified by column chromatography (0-5% MeOH/EtOAc) to give **63** (45 mg, 0.158 mmol, 76%) as a white powder.

Method B: Compound **62** (208 mg, 0.501 mmol) was dissolved in CH_2Cl_2 (5 mL) and cooled to 0 °C over ice. Methylhydrazine (40 µL, 0.751 mmol) was added dropwise over an hour. A white precipitate formed and the solvent was removed under reduced pressure. The crude was purified by column chromatography (0-15% MeOH/CH_2Cl_2) to give **63** (41 mg, 0.144 mmol, 28%) as a white powder.

R_f: 0.20 (MeOH:CH_2Cl_2; 1:9); ^{1}H NMR (400 MHz, MeOD) δ 7.74 (q, J = 1.2 Hz, 1H, C(6)H), 5.54 (s, 1H, C(1')), 4.27 (s, 1H, C(3')H), 4.07 (s, 1H, C(2')H), 3.96 (dd, J = 7.9, 0.6 Hz, 1H, C(7)H_A), 3.91 (s, 2H, C(5')H_2), 3.75 (d, J = 7.9 Hz, 1H, C(7)H_B), 1.89 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, MeOD) δ 166.5 (C-4), 151.9 (C-2), 136.8 (C-6), 110.7 (C-5), 90.4 (C-4'), 88.3 (C-1'), 80.9 (C-3'), 72.4 (C-7), 70.2 (C-2'), 57.6 (C-5'), 12.6 (CH$_3$-thymine); LRMS (ESI+) m/z (%): $[M+Na]^+$ 293 (100%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{11}H_{15}N_3O_6Na$, 308.0853; found 308.0581.

1-((1R,3R,4R,7S)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-(naphthalen-2-yl-methoxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1H,3H)-dione (**67**)



43

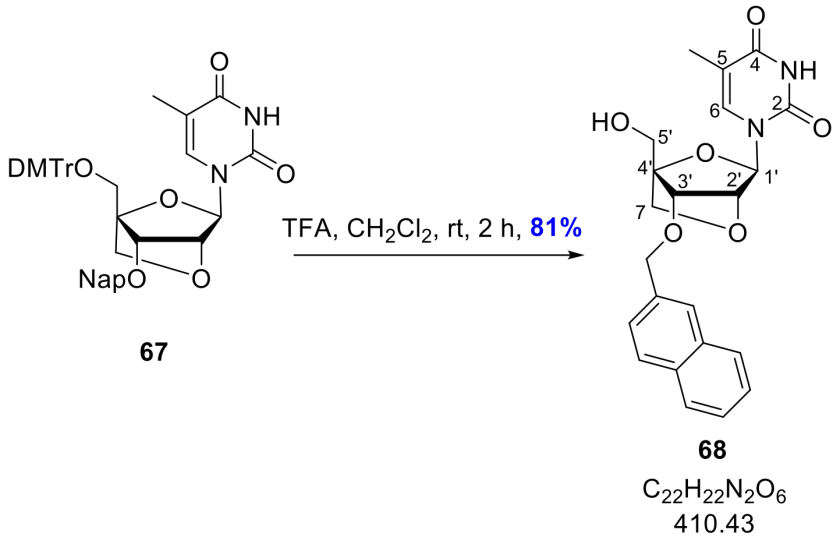
67
$C_{43}H_{40}N_2O_8$
712.80

Compound **43** (500 mg, 0.873 mmol) was dissolved in anhydrous THF (8 mL) and 60% NaH in mineral oil (140 mg, 3.49 mmol) was added in one portion. After 20 min, 2-bromomethylnapthalene (290 mg, 1.31 mmol) was added and refluxed for 1 h. The reaction was brought to rt, quenched with MeOH (10 mL), reduced under pressure, and the crude

residue was purified by column chromatography (0-5% MeOH/CH_2Cl_2) to give **67** (598 mg, 0.838 mmol, 96%) as a white foam.

R_f: 0.55 (MeOH:CH_2Cl_2; 1:19); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (s, 1H, N*H*), 7.87 – 7.70 (m, 3H, *H*-Ar(Nap)), 7.65 (d, J = 1.4 Hz, 1H, C(6)*H*), 7.54 – 7.45 (m, 2H, *H*-Ar), 7.44 – 7.37 (m, 2H, *H*-Ar), 7.35 (dd, J = 8.4, 1.7 Hz, 1H, *H*-Ar(Nap)), 7.32 – 7.19 (m, 7H, *H*-Ar(DMTr)), 6.81 – 6.70 (m, 4H, *H*-Ar(DMTr)), 5.66 (s, 1H, C(1')*H*), 4.85 (d, J = 11.6 Hz, 1H, CH_A-Nap), 4.68 – 4.63 (m, 2H, CH_B-Nap and C(2')*H*), 4.22 (s, 1H, C(3')*H*), 3.95 (d, J = 7.7 Hz, 1H, C(7)H_A), 3.77 (d, J = 7.7 Hz, 1H, C(7)H_B), 3.74 (s, 3H, CH_3-OMe), 3.71 (s, 3H, CH_3-OMe), 3.55 (d, J = 10.9 Hz, 1H, C(5')H_A), 3.42 (d, J = 11.0 Hz, 1H, C(5')H_B), 1.54 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.7 (*C*-4), 158.8 (*C*-Ar(DMTr)), 149.8 (*C*-2), 144.5 (*C*-Ar(Nap)), 135.4 (*C*-Ar(DMTr)), 135.3 (*C*-Ar(DMTr)), 134.6 (*C*-6), 134.3 (*C*-Ar(Nap)), 133.2 (*C*-Ar(DMT)), 133.2 (*C*-Ar(DMTr)), 130.2 (*C*-Ar), 130.1 (*C*-Ar), 128.6 (*C*-Ar), 128.2 (*C*-Ar), 128.1 (*C*-Ar), 128.0 (*C*-Ar), 127.9 (*C*-Ar), 127.3 (*C*-Ar), 127.1 (*C*-Ar), 126.6 (*C*-Ar), 126.4 (*C*-Ar(Nap)), 125.9 (*C*-Ar(Nap)), 113.5 (*C*-Ar(DMTr)), 113.4 (*C*-Ar(DMTr)), 110.6 (*C*-5), 87.9 (*C*-4'), 87.5 (*C*-1'), 86.8 (*C*-DMTr), 77.2 (*C*-2'), 75.8 (*C*-3'), 72.5 (CH_2-Nap or *C*-7), 72.5 (CH_2-Nap or *C*-7), 58.1 (*C*-5'), 55.3 (CH_3-OMe), 55.3 (CH_3-OMe) 12.7 (CH_3-thymine); LRMS-ESI (*m/z*): $[DMT]^+$ 303 (97%), $[M+Na]^+$ 735 (100%), $[M+Na]^+$ 736 (58%); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd. for $C_{43}H_{40}N_2O_8Na$: 735.2677, found 735.2691.

1-((1*S*,3*R*,4*R*,7*S*)-1-(hydroxymethyl)-7-(naphthalen-2-ylmethoxy)-2,5-dioxabicyclo[2.2.1]-heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**68**)



67

68

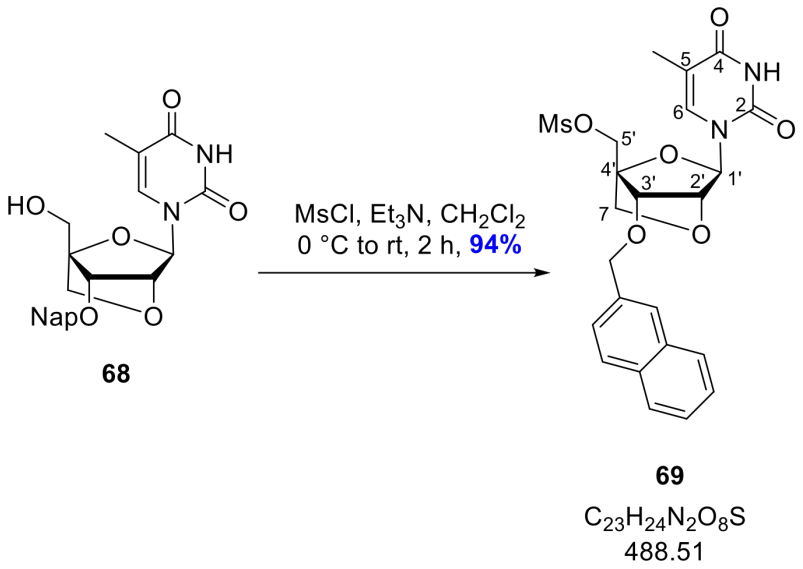
$C_{22}H_{22}N_2O_6$

410.43

Compound **67** (416 mg, 0.584 mmol) was dissolved in CH_2Cl_2 (6 mL) and trifluoroacetic acid (TFA) (1 mL) was added and the reaction was stirred at rt for 2 h. The reaction was quenched with MeOH (10 mL) and the solvent removed under reduced pressure. The crude residue was purified by column chromatography (0-6% MeOH/CH_2Cl_2) to give **68** (195 mg, 0.475 mmol, 81%) as a white foam.

R_f: 0.45 (MeOH:CH_2Cl_2; 1:9); ^{1}H NMR (400 MHz, MeOD) δ 7.89 – 7.77 (m, 4H, *H*-Ar), 7.65 (d, J = 1.3 Hz, 1H, C(6)*H*), 7.53 – 7.44 (m, 3H, *H*-Ar), 5.59 (s, 1H, C(1')*H*), 4.87 – 4.76 (m, 2H, CH_2-Nap), 4.47 (s, 1H, C(3')*H*), 4.08 – 4.01 (m, 2H, C(7)H_A and C(2')*H*), 3.95 (d, J = 2.0 Hz, 2H, C(5')H_2), 3.81 (d, J = 7.8 Hz, 1H, C(7)H_B), 1.78 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (126 MHz, MeOD) δ 166.4 (*C*-4), 151.8 (*C*-2), 136.5 (*C*-Ar), 136.2 (*C*-6), 134.6 (*C*-Ar), 134.6 (*C*-Ar), 129.2 (*C*-Ar), 128.9 (*C*-Ar), 128.7 (*C*-Ar), 128.1 (*C*-Ar), 127.3 (*C*-Ar), 127.2 (*C*-Ar), 127.0 (*C*-Ar), 110.8 (*C*-5), 89.9 (*C*-4'), 88.5 (*C*-1'), 78.2 (*C*-3'), 76.8 (*C*-2'), 73.2 (*C*H_2-Nap or *C*-7), 73.2 (*C*H_2-Nap or *C*-7), 57.6 (*C*-5'), 12.5 (*C*H_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 411 (22%), $[2M+H]^+$ 821 (100%), $[2M+Na]^+$ 843 (53%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{22}H_{23}N_2O_6$, 411.1551; found, 411.1583; $[M+Na]^+$ calcd. for $C_{22}H_{22}N_2O_6Na$, 433.1370; found, 433.1374.

((1*R*,3*R*,4*R*,7*S*)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-7-(naphthalen-2-ylmethoxy)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl methanesulfonate (**69**)



68 **69**

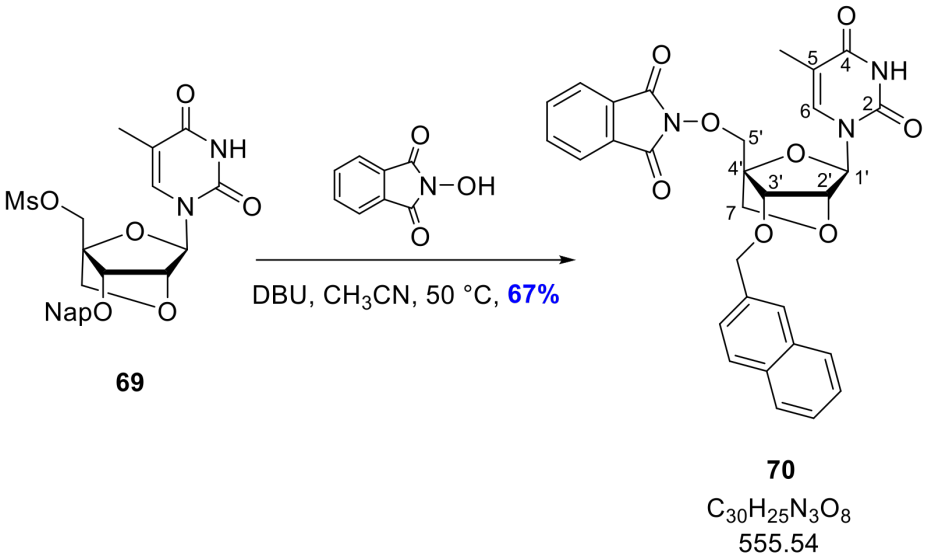
$C_{23}H_{24}N_2O_8S$
488.51

Compound **68** (690 mg 1.69 mmol) was dissolved in CH_2Cl_2 (10 mL) and Et_3N (354 µL, 2.54 mmol) and cooled to 0 °C on ice. Methanesulfonyl chloride (196 µL, 2.54 mmol) was added dropwise over 30 min and the reaction was allowed to warm to rt and stirred for 1.5 h. The reaction was diluted with additional CH_2Cl_2 (10 mL) and washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over $MgSO_4$ and reduced under pressure. The crude residue was purified by column chromatography (0-4% MeOH/CH_2Cl_2) to give **69** (773 mg, 1.58 mmol, 94%) as a white foam.

R_f: 0.45 (MeOH:CH_2Cl_2; 1:19); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.23 (s, 1H, N*H*), 7.87 – 7.77 (m, 3H, *H*-Ar), 7.77 – 7.70 (m, 1H, C(6)*H*), 7.55 – 7.34 (m, 4H, *H*-Ar), 5.66 (s, 1H, C(1')*H*), 4.85 (d, J = 11.5 Hz, 1H, CH_A-Nap), 4.72 (d, J = 11.5 Hz, 1H, CH_B-Nap), 4.64 – 4.49 (m, 3H, C(5')H_2 and C(2')*H*), 4.12 (d, J = 7.8 Hz, 1H, C(7)H_A), 3.96 (s, 1H, C(3')*H*), 3.90 (d, J = 7.8 Hz, 1H, C(7)H_B), 3.01 (s, 3H, CH_3-OMs), 1.84 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 163.3 (*C*-4), 149.5 (*C*-2), 134.2 (*C*-6), 133.9 (*C*-Ar), 133.3 (*C*-Ar), 133.2 (*C*-Ar), 128.7 (*C*-Ar), 128.0 (*C*-Ar), 127.9 (*C*-Ar), 127.4 (*C*-Ar), 126.7 (*C*-Ar), 126.7 (*C*-Ar), 125.9 (*C*-Ar), 110.9 (*C*-5), 87.8 (*C*-1'), 85.8 (*C*-4'), 76.9 (*C*-2'), 76.1 (*C*-3'), 72.9 (*C*H_2-Nap), 71.8 (*C*-7), 64.1 (*C*-5'), 38.0 (*C*H_3-OMs), 12.5

(*C*H$_3$-thymine); LRMS-ESI (*m/z*): [M+H]$^+$ 489 (74%), [M+Na]$^+$ 511 (100%), [2M+Na]$^+$ 999 (75%); HRMS-ESI (*m/z*): [M+H]$^+$ calcd. for $C_{23}H_{25}N_2O_8S$, 489.1326; found, 489.1333; [M+Na]$^+$ calcd. for $C_{23}H_{24}N_2O_8SNa$, 511.1146; found, 511.1151.

2-(((1*R*,3*R*,4*R*,7*S*)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-7-(naphthalen-2-ylmethoxy)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy)isoindoline-1,3-dione (**70**)



69

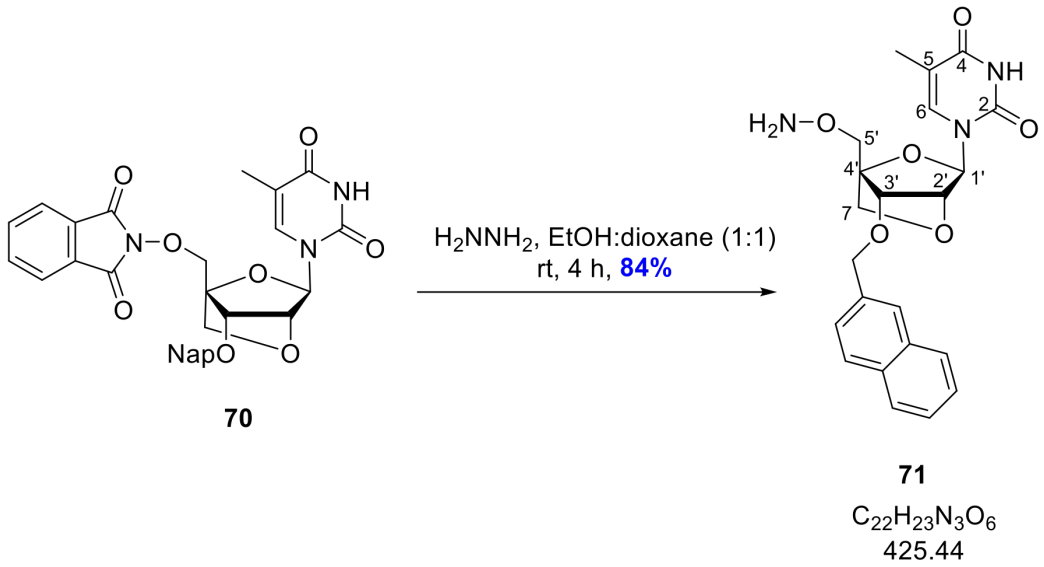
70
$C_{30}H_{25}N_3O_8$
555.54

Compound **69** (761 mg, 1.56 mmol) was dissolved in CH_3CN (15 mL) and *N*-hydroxyphthalimide (400 mg, 2.45 mmol) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (400 µL, 2.64 mmol) were added. The reaction was heated to 50 °C overnight. After being brought to rt, the reaction was neutralised with 1 M HCl, partitioned between EtOAc (20 mL) and washed with sat. $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over $MgSO_4$ before being reduced under pressure. The crude residue was purified by column chromatography (0-60% EtOAc/CH_2Cl_2) to give **70** (584 mg, 1.05 mmol, 67%) as white foam.

R_f: 0.65 (EtOAc:CH_2Cl_2; 8:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (s, 1H, N*H*), 7.89 – 7.70 (m, 9H, *H*-Ar and C(6)*H*), 7.51 – 7.39 (m, 3H, *H*-Ar), 5.66 (s, 1H, C(1')*H*), 4.84 (q, J = 11.6 Hz, 2H, CH_2-Nap), 4.75 (d, J = 11.4 Hz, 1H, C(5')H_A), 4.63 – 4.55 (m, 2H, C(2')*H* and C(5')H_B), 4.30 (s, 1H, C(3')*H*), 4.13 (d, J = 7.7 Hz, 1H, C(7)H_A), 3.93 (d, J = 7.7 Hz, 1H, C(7)H_B), 1.88 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.6 (*C*-4), 163.2 (*C*O-Phth), 149.7 (*C*-2), 135.2 (*C*-6), 134.9 (*C*-Ar), 134.4 (*C*-Ar), 133.3 (*C*-

Ar), 133.2 (*C*-Ar), 128.8 (*C*-Ar), 128.5 (*C*-Ar), 128.0 (*C*-Ar), 127.8 (*C*-Ar), 126.9 (*C*-Ar), 126.4 (*C*-Ar), 126.3 (*C*-Ar), 125.7 (*C*-Ar), 123.9 (*C*-Ar), 110.7 (*C*-5), 87.7 (*C*-1'), 86.1 (*C*-4'), 77.3 (*C*-2'), 76.3 (*C*-3'), 73.1 (*C*-5'), 72.9 (CH_2-Nap), 72.1 (*C*-7), 12.6 (CH_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 557 (94%), $[M+Na]^+$ 579 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{30}H_{26}N_3O_8$, 556.1714; found, 556.1724; $[M+Na]^+$ calcd. for $C_{30}H_{25}N_3O_8Na$, 578.1534; found, 578.1544.

1-((1*R*,3*R*,4*R*,7*S*)-1-((aminooxy)methyl)-7-(naphthalen-2-ylmethoxy)-2,5-dioxa-bicyclo-[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**71**)



70

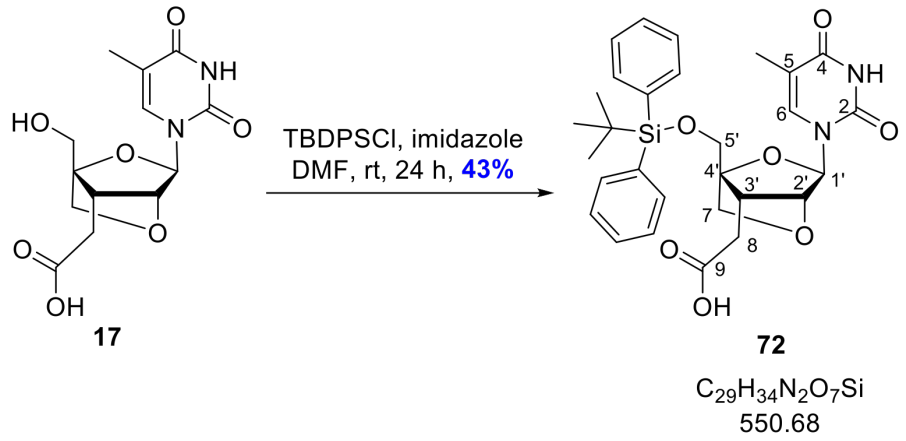
71
$C_{22}H_{23}N_3O_6$
425.44

Compound **70** (500 mg, 0.900 mmol) was dissolved in EtOH:dioxane (1:1, v/v) (10 mL) and hydrazine hydrate (85 µL, 1.36 mmol) was added and the reaction was stirred at rt for 4 h. The white precipitate was removed by filtration and the filtrate was reduced under pressure. The crude residue was purified by column chromatography (0-100% EtOAc/CH_2Cl_2) to give **71** (320 mg, 0.752 mmol, 84%) as a white powder.

R_f: 0.30 (EtOAc:CH_2Cl_2; 8:2); ^{1}H NMR (400 MHz, MeOD) δ 7.88 – 7.75 (m, 4H, *H*-Ar), 7.52 – 7.43 (m, 4H, *H*-Ar and C(6)*H*), 5.55 (s, 1H, C(1')*H*), 4.84 – 4.73 (m, 2H, CH_2-Nap), 4.46 (s, 1H, C(2')*H*), 4.14 – 4.03 (m, 3H, C(5')H_2 and C(7)H_A), 3.95 (s, 1H, C(3')*H*), 3.86 (d, J = 7.9 Hz, 1H, C(7)H_B), 1.75 (d, J = 1.2 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, MeOD) δ 166.3 (*C*-4), 151.7 (*C*-2), 136.3 (*C*-Ar), 136.1 (*C*-Ar), 134.6 (*C*-6), 134.6 (*C*-Ar), 129.3 (*C*-

Ar), 128.9 (*C*-Ar), 128.7 (*C*-Ar), 128.2 (*C*-Ar), 127.4 (*C*-Ar), 127.3 (*C*-Ar), 127.1 (*C*-Ar), 110.9 (*C*-5), 88.6 (*C*-4'), 88.0 (*C*-1'), 78.2 (*C*-2'), 77.2 (*C*-3') 73.5 (*C*-7) 73.3 (CH_2-Nap), 71.8 (*C*-5'), 12.5 (CH_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 426 (100%),$[2M+H]^+$ 851 (74%), $[3M+H]^+$ 1276 (57%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{22}H_{24}N_3O_6$, 426.1660; found, 426.1646; $[M+Na]^+$ calcd. for $C_{22}H_{23}N_3O_6Na$, 448.1479; found, 448.1478.

2-((1*R*,3*R*,4*R*,7*S*)-1-(((tert-butyldiphenylsilyl)oxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydro pyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)acetic acid (**72**)

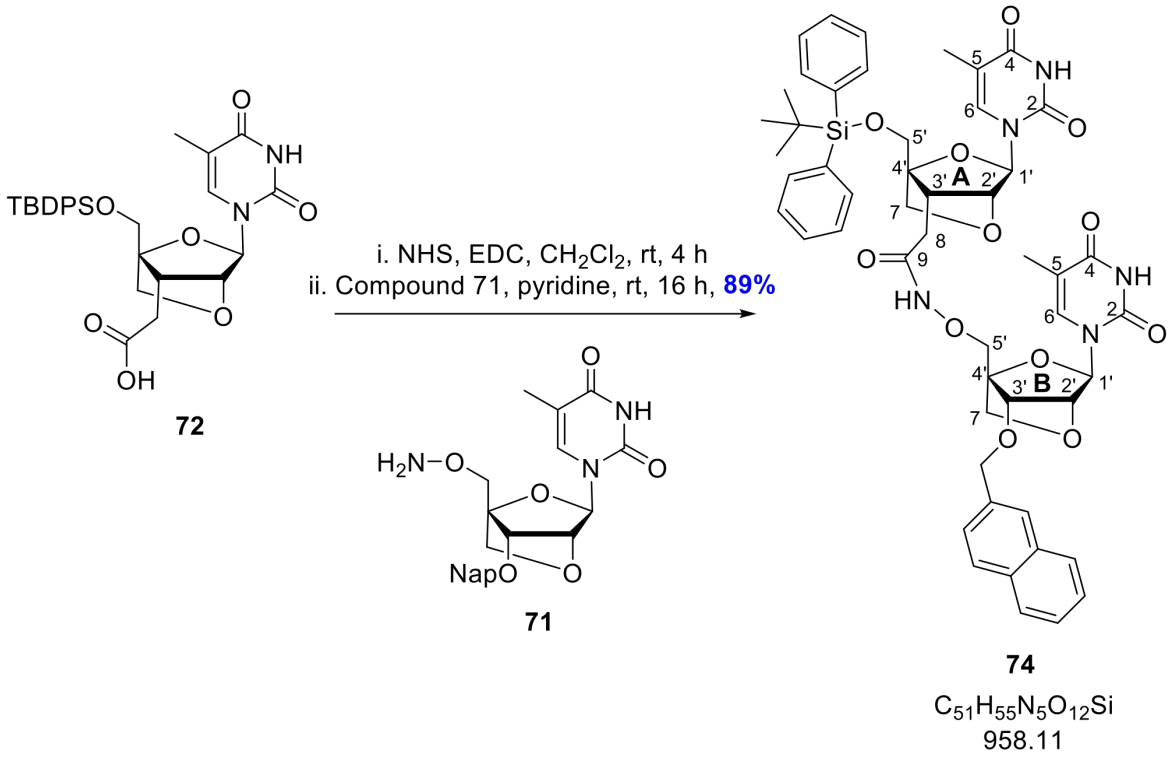


Compound **17** (426 mg, 1.36 mmol) and Et_3N (238 µL, 1.71 mmol) were dissolved in DMF (2 mL) and cooled to 0 °C over ice. TBDPSCl (443 uL, 1.71 mmol) was added dropwise and the reaction was stirred at rt for 16 h. The DMF was removed under reduced pressure and the crude was partitioned between CH_2Cl_2 (20 mL) and water (20 mL). The organic layer was washed with sat. $NaHCO_3$ (20 mL), brine (20 mL), and dried over $MgSO_4$. The crude was submitted to column chromatography (0-10% MeOH/CH_2Cl_2) to yield compound **72** (322 mg, 0.585 mmol, 43%) as a white powder.

R_f: 0.60 (MeOH:CH_2Cl_2; 1:9); ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.65 (s, 1H, N*H*), 7.72 – 7.65 (m, 4H, *H*-Ar), 7.52 (d, J = 1.4 Hz, 1H, C(6)*H*), 7.49 – 7.37 (m, 6H, *H*-Ar), 5.65 (s, 1H, C(1')*H*), 4.70 (s, 1H, C(2')*H*), 4.11 (d, J = 12.2 Hz, 1H, C(5')H_A), 3.81 (d, J = 12.1 Hz, 1H, C(5')H_B), 3.74 (d, J = 8.6 Hz, 1H, C(7)H_A), 3.64 (d, J = 8.6 Hz, 1H, C(7)H_B), 2.42 – 2.30 (m, 3H, C(8)H_2 and C(3')*H*), 1.74 (d, J = 1.2 Hz, 3H, CH_3-thymine), 1.11 (s, 9H, 3 × CH_3-tBu); ^{13}C NMR (151 MHz, $CDCl_3$) δ 175.1 (*C*-9), 164.7 (*C*-4), 150.2 (*C*-2), 135.8

(*C*-Ar), 135.5 (*C*-Ar), 134.9 (*C*-6), 132.7 (*C*-Ar), 132.4 (*C*-Ar), 130.5 (*C*-Ar), 130.3 (*C*-Ar), 128.2 (*C*-Ar), 128.2 (*C*-Ar), 128.1 (*C*-Ar), 128.1 (*C*-Ar), 110.5 (*C*-5), 90.8 (*C*-4'), 87.5 (*C*-1'), 80.4 (*C*-2'), 71.5 (*C*-7), 59.8 (*C*-5'), 40.2 (*C*-3'), 28.8 (*C*-8), 27.1 ($C(CH_3)_3$-tBu), 19.5 (CH_3-tBu), 12.5 (CH_3-thymine); LRMS-ESI (*m/z*): $[M+H]^+$ 551 (65%), $[M+Na]^+$ 573 (72%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{29}H_{35}N_2O_7Si$, 551.2208; found, 551.2207

2-((1*R*,3*R*,4*R*,7*S*)-1-(((tert-butyldiphenylsilyl)oxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)-*N*-(((1*R*,3*R*,4*R*,7*S*)-3-(5-methyl-2,4-dioxo-3,4- dihydropyrimidin-1(2*H*)-yl)-7-(naphthalen-2-ylmethoxy)-2,5-dioxabicyclo[2.2.1] heptan-1-yl)methoxy)acetamide (**74**)

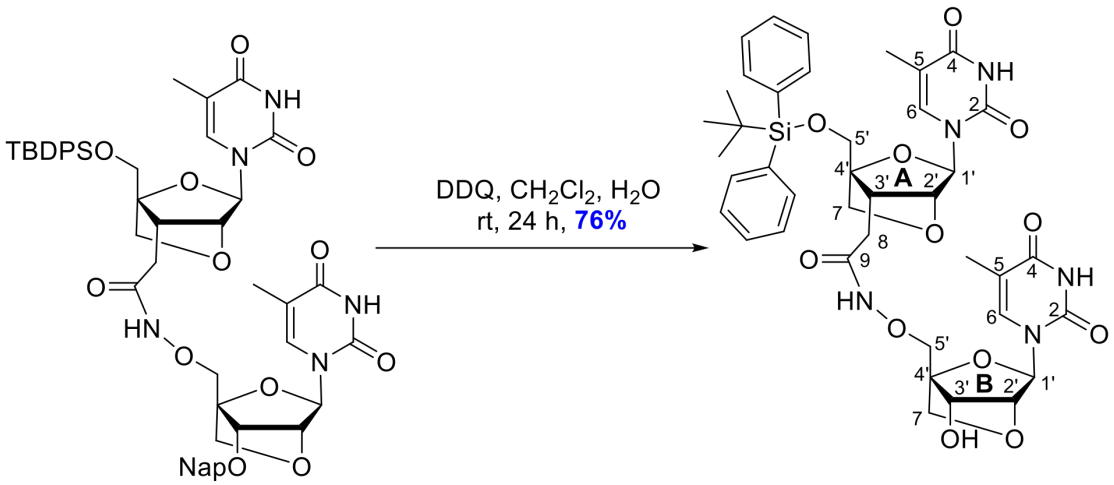


Compound **72** (188 mg, 0.342 mmol) was dissolved in CH_2Cl_2 (3 mL) and EDC (98 mg, 0.513 mmol) and NHS (43 mg, 0.376 mmol) were added and the reaction was stirred at rt for 4 h. The reaction mixture was diluted in CH_2Cl_2 (5 mL), washed with half sat. $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over $MgSO_4$. The solvent was removed under reduced pressure to give the intermediate electrophile, which was used without further purification. The electrophile was dissolved in pyridine (3 mL) and compound **71** (160 mg, 0.376 mmol)

was added. After stirring for 16 h at rt, the pyridine was removed under reduced pressure and the crude residue was purified by column chromatography (0-100% EtOAc/CH_2Cl_2) to give dimer **74** (290 mg, 0.303 mmol, 89%) as a white powder.

R_f: 0.75 (EtOAc:MeOH; 19:1); ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.41 (d, J = 11.4 Hz, 2H, NH-thymine(A or B) and NH-amino-oxyamide), 11.35 (s, 1H, NH-thymine(A or B)), 7.91 – 7.80 (m, 4H, H-Ar), 7.70 – 7.65 (m, 4H, H-Ar), 7.60 (s, 1H, C(A6 or B6)H), 7.53 – 7.39 (m, 10H, H-Ar and C(A6 or B6)H), 5.50 (d, J = 3.4 Hz, 2H, C(A1' and B1')H_2), 4.78 (s, 2H, CH_2-Nap), 4.51 (d, J = 4.7 Hz, 1H, C(B2')H), 4.41 (d, J = 16.3 Hz, 2H, C(A2')H and C(A5')H_A), 4.15 (d, J = 11.6 Hz, 1H, C(A5')H_B), 4.13 – 4.08 (m, 2H, C(B3')H and C(B5')H_A), 4.03 (d, J = 12.3 Hz, 1H, C(B5')H_B), 4.00 – 3.94 (m, 1H, C(B7)H_A), 3.82 (d, J = 8.1 Hz, 1H, C(B7)H_B), 3.74 – 3.66 (m, 2H, C(A7)H_2), 2.59 (s, 1H, C(NHS)H_4), 2.43 (t, J = 6.8 Hz, 1H, C(A3')H), 2.09 (d, J = 6.7 Hz, 2H, C(8)H_2), 1.72 – 1.65 (m, 3H, CH_3-thymine(A or B)), 1.48 (s, 3H, CH_3-thymine(A or B)), 1.01 (s, 9H, 3 × CH_3-tBu); ^{13}C NMR (151 MHz, DMSO-d_6) δ 172.7 (CO-NHS), 167.5 (C-9), 163.8 (C-A4 or B4), 163.7 (C-A4 or B4), 149.9 (C-A2 and B2), 135.3 (C-Ar), 135.1 (C-Ar), 134.9 (C-Ar), 134.8 (C-A6 or B6), 134.1 (C-A6 or B6), 132.7 (C-Ar), 132.5 (C-Ar), 132.3 (C-Ar), 1301 (C-Ar), 129.9 (C-Ar), 127.9 (C-Ar), 127.9 (C-Ar), 127.8 (C-Ar), 127.6 (C-Ar), 127.5 (C-Ar), 126.2 (C-Ar), 126.0 (C-Ar), 125.9 (C-Ar), 125.7 (C-Ar), 108.6 (C-A5 or B5), 108.4 (C-A5 or B5), 90.4 (C-A4'), 86.6 (C-A1' or B1'), 86.5 (C-A1' or B1'), 85.5 (C-B4'), 79.9 (C-A2'), 76.6 (C-B2' or B3'), 76.4 (C-B2' or C-3'), 71.6 (CH$_2$-Nap or C-A5' or C-B7), 71.1 (CH$_2$-ONap or C-A5' or C-B7), 71.3 (CH$_2$-Nap or C-A5' or C-B7), 70.9 (C-A7), 59.7 (C-B5'), 40.1 (C-A3'), 26.8 (C-8), 26.6 C(CH$_3$)$_3$-tBu), 25.2 (CH$_2$-NHS), (18.9 (CH$_3$-tBu), 12.1 (CH$_3$-thymine(A or B)), 11.9 (CH$_3$-thymine(A or B)); LRMS-ESI (m/z): [M+H]$^+$ 958 (24%), [M+Na]$^+$ 980 (100%), [M+Na]$^+$ 981 (56%); HRMS-ESI (m/z): [M+H]$^+$ calcd. for $C_{51}H_{56}N_5O_{12}Si$, 958.3689; found, 958.3690; [M+Na]$^+$ calcd. for $C_{51}H_{55}N_5O_{12}SiNa$, 980.3509; found, 980.3516.

2-((1*R*,3*R*,4*R*,7*S*)-1-(((tert-butyldiphenylsilyl)oxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydro pyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)-*N*-(((1*R*,3*R*,4*R*,7*S*)-7-hydroxy-3- (5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl) methoxy)acetamide (**75**)

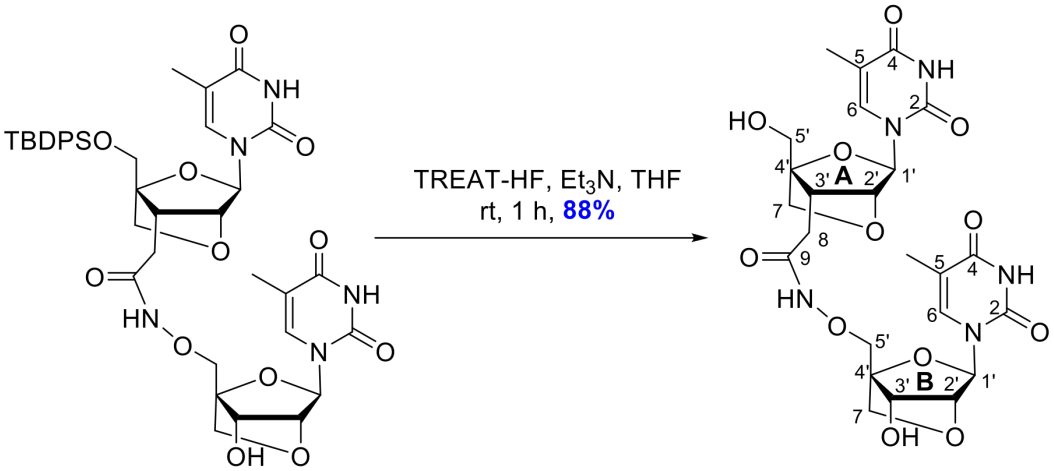


Compound **74** (285 mg, 0.297 mmol) was dissolved in CH_2Cl_2 (10 mL) and H_2O (50 µL) and 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) (101 mg, 0.446 mmol) was added. MeOH (500 µL) was added to solubilise the reaction and it was stirred at rt for 3 h. Additional portions of DDQ (35 mg, 0.155 mmol) and H_2O (25 µL) were added and the reaction was stirred at rt overnight. The solvent was removed under reduced pressure and the crude was re-dissolved in EtOAc (7 mL), washed with 10% (w/v) Na_2SO_3 (10 mL), sat. $NaHCO_3$ (10 mL), brine (10 mL), dried over $MgSO_4$, and reduced under pressure. The crude was purified by column chromatography (0-3% MeOH/EtOAc) to give the dimer **75** (165 mg, 0.226 mmol, 76%) as a white powder.

R_f: 0.45 (MeOH:EtOAc; 3:97); ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.38 (d, J = 52.2 Hz, 3H, 3 × N*H*), 7.72 – 7.64 (m, 4H, *H*-Ar), 7.62 (s, 1H, C(A6 or B6)*H*), 7.49 – 7.44 (m, 7H, *H*-Ar and C(A6 or B6)*H*), 5.77 (s, 1H, C(B3')O*H*), 5.50 (s, 1H, C(A1')*H*), 5.43 (s, 1H, C(B1')*H*), 4.41 (s, 1H, C(A2')*H*), 4.37 (d, J = 11.6 Hz, 1H, C(A5')H_A), 4.16 – 4.08 (m, 3H, C(B2')*H* and C(A5')H_B and C(B5')H_A), 4.03 (d, J = 12.3 Hz, 1H, C(B5')H_B), 3.95 (d, J = 2.5 Hz,

1H, C(B3')H), 3.91 (d, J = 8.0 Hz, 1H, C(B7)H_A), 3.77 (d, J = 8.0 Hz, 1H, C(B7)H_B), 3.74 – 3.66 (m, 2H, C(A7)H_2), 2.40 (td, J = 9.1, 3.4 Hz, 1H, C(A3')H), 2.14 – 2.03 (m, 2H, C(8)H_2), 1.78 – 1.72 (m, 3H, CH_3-thymine(A or B)), 1.50 (d, J = 1.2 Hz, 3H, CH_3-thymine(A or B)), 1.04 (s, 9H, 3 × CH_3-tBu); ^{13}C NMR (151 MHz, DMSO-d$_6$) δ 167.4 (C-9), 163.8 (C-A4 or B4), 163.7 (C-A4 or B4), 149.9 (C-A2 or B2), 149.9 (C-A2 or B2), 135.2 (C-Ar), 134.9 (C-A6 or B6), 134.1 (C-A6 or B6), 132.7 (C-Ar), 132.3 (C-Ar), 130.1 (C-Ar), 130.0 (C-Ar), 128.0 (C-Ar), 127.9 (C-Ar), 108.6 (C-A5 or B5), 108.4 (C-A5 or B5), 90.4 (C-A4'), 86.5 (C-A1' or B1'), 86.5 (C-A1' or B1'), 86.3 (C-B4'), 79.8 (C-A2'), 78.8 (C-B2'), 71.7 (C-A5'), 70.9 (C-A7 and B7), 69.5 (C-B3'), 59.8 (C-B5'), 40.1 (C-A3'), 26.8 (C-8), 26.7 $C(CH_3)_3$-tBu), (18.9 (CH_3-tBu), 12.0 (CH_3-thymine(A or B)), 11.9 (CH_3-thymine(A or B)); LRMS-ESI (m/z): $[M+H]^+$ 818 (37%), $[M+Na]^+$ 840 (100%), $[M+Na]^+$ 841 (62%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{40}H_{48}N_5O_{12}Si$, 818.3063; found, 818.3092.

2-((1*R*,3*R*,4*R*,7*S*)-1-(((tert-butyldiphenylsilyl)oxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)-*N*-(((1*R*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy) acetamide (**76**)



75

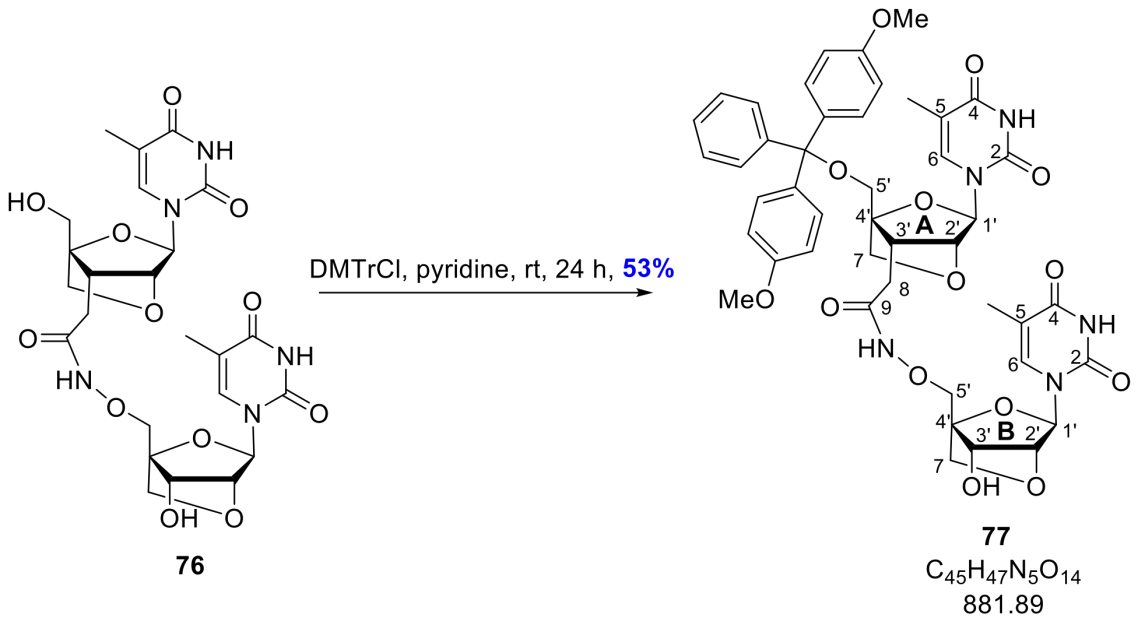
76
$C_{24}H_{29}N_5O_{12}$
579.52

Compound **75** (160 mg, 0.196 mmol) was dissolved in THF (3 mL) and *N,N*-diethylethanamine

trihydrofluoride (TREAT-HF) (160 µL, 0.980 mmol) and Et_3N (137 µL, 0.980 mmol) were added and the reaction was stirred at rt for 1 hour. The solvent was removed under reduced pressure and the crude residue was purified by reverse-phase column chromatography (0-100% MeCN/H_2O) and then again by silica column chromatography (0-10% MeOH/EtOAc) to give diol dimer **76** (100 mg, 0.173 mmol, 88%) as a white powder.

R_f: 0.30 (MeOH:EtOAc; 1:0); ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.40 (m, 3H, 3 × N*H*), 7.62 (dd, J = 3.1, 1.4 Hz, 2H, C(A6)*H* and C(B6)*H*), 5.77 (m, 1H, O*H*) 5.43 (d, J = 4.2 Hz, 2H, C(A1')*H* and C(B1')*H*), 5.23 (dd, J = 7.6, 4.2 Hz, 1H, O*H*)), 4.36 (d, J = 12.6 Hz, 2H, C(B3')*H* and C(A5')H_A), 4.16 – 4.08 (m, 2H, C(A5')H_B and C(A2')*H*), 3.98 – 3.87 (m, 2H, C(B2')*H* and C(B7)H_A), 3.78 (td, J = 12.1, 5.5 Hz, 3H, C(B5')H_2 and C(B7)H_B), 3.71 (d, J = 8.5 Hz, 1H, C(A7)H_A)), 3.62 (d, J = 8.5 Hz, 1H, C(A7)H_B)), 2.25 – 2.13 (m, 2H, C(8)H_A and C(A3')*H*), 2.04 (t, J = 7.0 Hz, 1H, C(8)H_B), 1.77 (dd, J = 11.6, 1.1 Hz, 6H, 2 × CH_3-thymine (A and B)); ^{13}C NMR (151 MHz, DMSO-d_6) δ 167.8 (*C*-9), 163.9 (*C*-A4 or B4), 163.8 (*C*-A4 or B4), 149.9 (*C*-A2 or B2), 149.9 (*C*-A2 or B2), 134.9 (*C*-A6 or B6), 134.8 (*C*-A6 or B6), 108.6 (*C*-A5 or B5), 108.3 (*C*-A5 or B5), 90.7 (*C*-A4'), 86.5 (*C*-A1' or B1'), 86.4 (*C*-A1' or B1'), 86.2 (*C*-B4'), 79.8 (*C*-B3'), 78.9 (*C*-A2'), 71.7 (*C*-A5'), 71.0 (*C*-A7 or B7) 70.9 (*C*-A7 or B7), 69.5 (*C*-B2'), 56.7 (*C*-B5'), 40.1 (*C*-A3'), 26.8 (*C*-8), 12.4 (*C*H_3-thymine(A or B)), 12.0 (*C*H_3-thymine(A or B); LRMS-ESI (*m/z*): $[M+H]^+$ 580 (18%), $[M+Na]^+$ 602 (89%), $[M+Na]^+$ 603 (21%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{24}H_{30}N_5O_{12}$, 580.1886; found, 580.1883; $[M+Na]^+$ calcd. for $C_{24}H_{29}N_5O_{12}Na$, 602.1705; found, 602.1730.

2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl)-*N*-(((1*R*,3*R*,4*R*,7*S*)-7-hydroxy-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methoxy)acetamide (**77**)

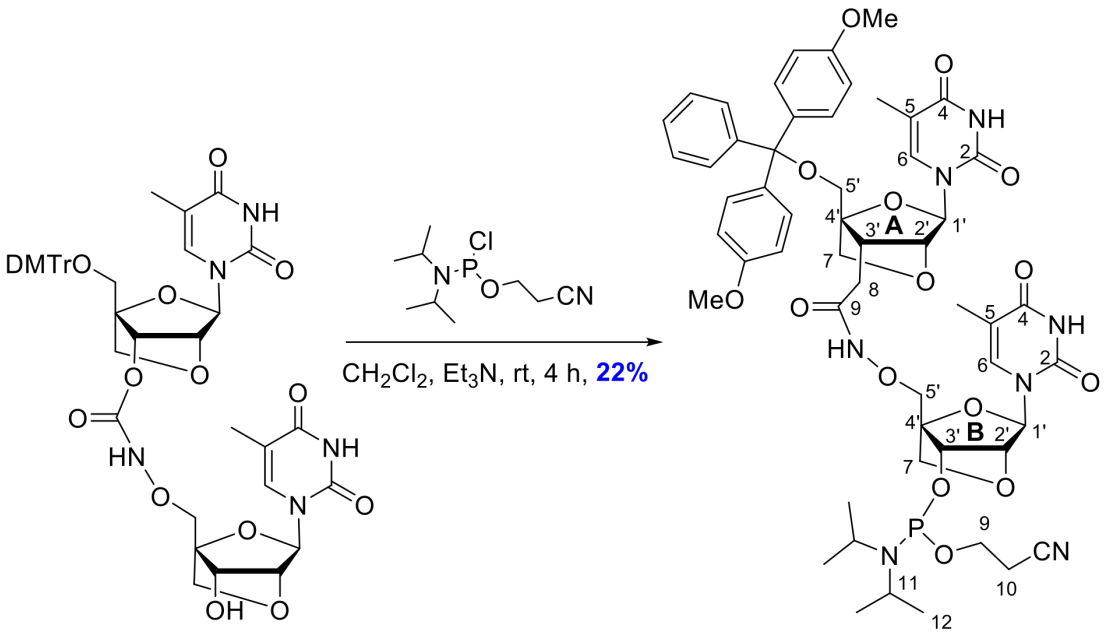


Compound **76** (98 mg, 0.169 mmol) was dissolved in anhydrous pyridine (3 mL) and activated 4Å molecular sieves were added. DMTrCl (75 mg, 0.220 mmol) dissolved in anhydrous pyridine (500 µL), was added drop-wise to the reaction over 30 min. After 6 h, an additional portion of DMTrCl (75 mg, 0.220 mmol), dissolved in anhydrous pyridine (500 µL), was added dropwise over the course of an hour and the reaction was left to stir overnight. The solvent was removed under reduced pressure and the crude was purified by column chromatography (0-5% MeOH/EtOAc) to give dimer **77** (79 mg, 0.090 mmol, 53%) as a white powder.

R_f: 0.40 (EtOAc); ^{1}H NMR (600 MHz, acetone-d$_6$) δ 10.57 (s, 1H, N*H*), 10.14 (s, 1H, N*H*), 10.03 (s, 1H, N*H*), 7.73 (s, 1H, C(A6 or B6)*H*), 7.67 (s, 1H, C(A6 or B6)*H*), 7.56 – 7.50 (m, 2H, *H*-Ar), 7.43 – 7.33 (m, 6H, *H*-Ar), 7.28 – 7.24 (m, 1H, *H*-Ar), 6.92 (dq, J = 8.1, 3.2 Hz, 4H, *H*-Ar), 5.61 (s, 1H, C(B3')*H*), 5.59 (s, 1H, C(A1')*H*), 5.50 (s, 1H, C(B1')*H*), 4.57 (s, 1H, C(A2')*H*), 4.35 – 4.22 (m, 3H, C(B2')*H*) and C(B7 or B5')*H*$_2$), 4.00 (d, J = 7.9 Hz,

1H, C(A7 or B7 or B5')H_{A}), 3.81 – 3.74 (m, 8H, 2 × CH_3-OMe and C(A7 or B7 or B5')H_{B} and C(A7 or B7 or B5')H_{A}), 3.69 (dd, J = 10.1, 3.6 Hz, 2H, C(A7 or B7 or B5')H_{B} and C(A5')H_{A}), 3.47 (d, J = 11.2 Hz, 1H, C(A5')H_{B}), 2.73 (dd, J = 9.2, 4.5 Hz, 1H, C(A3')H), 2.21 – 2.10 (m, 2H, C(8)H_2), 1.81 – 1.77 (m, 3H, CH_3-thymine(A or B)), 1.64 (s, 3H, CH_3-thymine(A or B)); ^{13}C NMR (151 MHz, acetone-d_6) δ 169.2 (C-9), 164.4 (C-A4 and B4), 159.8 (C-Ar), 150.9 (C-A2 or B2), 150.9 (C-A2 or B2), 145.8 (C-Ar), 136.5 (C-Ar), 136.3 (C-Ar), 135.7 (C-A6 or B6), 135.4 (C-A6 or B6), 131.0 (C-Ar), 129.0 (C-Ar), 128.8 (C-Ar), 127.8 (C-Ar), 114.1 (C-Ar), 110.0 (C-A5 and B5), 90.7 (C-A4') 88.2 (C-A1'), 88.1 (C-B1'), 87.7 (C-B4'), 87.4 (C-DMTr), 81.0 (C-A2'), 80.2 (C-B2'), 72.7 (C-A7 or B7 or B5'), 72.5 (C-A7 or B7 or B5'), 72.0 (C-A7 or B7 or B5'), 71.0 (C-B3'), 59.9 (C-A5'), 54.9 (CH$_3$-OMe), 40.9 (C-A3'), 27.9 (C-8), 12.9 (CH$_3$-thymine(A or B)), 12.7 (CH$_3$-thymine(A or B)); LRMS-ESI (m/z): [M+Na]$^+$ 904 (100%) 615 (31%, T-acid); HRMS-ESI (m/z): [M+Na]$^+$ calcd. for $C_{45}H_{47}N_5O_{14}Na$, 904.3012, found 904.3057.

(1*R*,3*R*,4*R*,7*S*)-1-(((2-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl) acetamido)oxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxa bicyclo[2.2.1]heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite (**78**)



77

78

$C_{54}H_{64}N_7O_{15}P$

1082.11

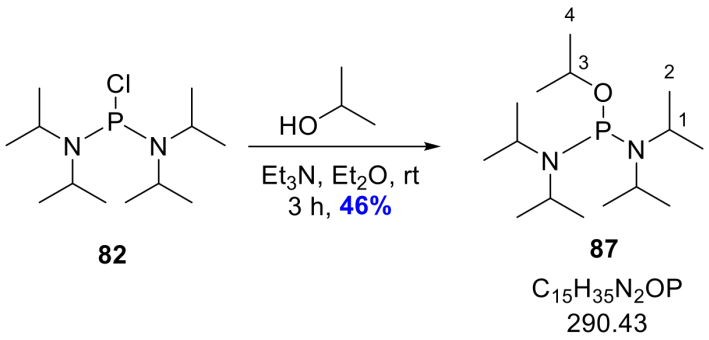
Compound **77** (79 mg, 0.089 mmol) dissolved in anhydrous, degassed CH_2Cl_2 (1.5 mL) and anhydrous, degassed Et_3N (12.5 µL, 0.178 mmol) and chloro-phosphitylating reagent (20 µL, 0.134 mmol) were added and the reaction was stirred in an inert atmosphere at rt for 1.5 h. Additional portions of Et_3N (20 µL, 0.285 mmol) and chloro-phosphitylating reagent (10 µL, 0.067 mmol) were added once every 1.5 h for the next 3 h. After 1 h, an inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. A reverse precipitation from anhydrous, degassed hexanes (100 mL) was attempted. A portion of the hexanes (80 mL) were siphoned off and the remaining solvents were removed under reduced pressure. The residue was purified by inert column chromatography (100% EtOAc, 0.5% Et_3N) to give

phosphoramidite **78** (21 mg, 0.019 mmol, 22%) as a white solid. NB: Only a portion was siphoned in the interest of yield.

R_f: 0.50 (EtOAc).

7.1.4 Compounds in Chapter 4

1-isopropoxy-*N,N,N',N'*-tetraisopropylphosphanediamine (**87**)



82 **87**

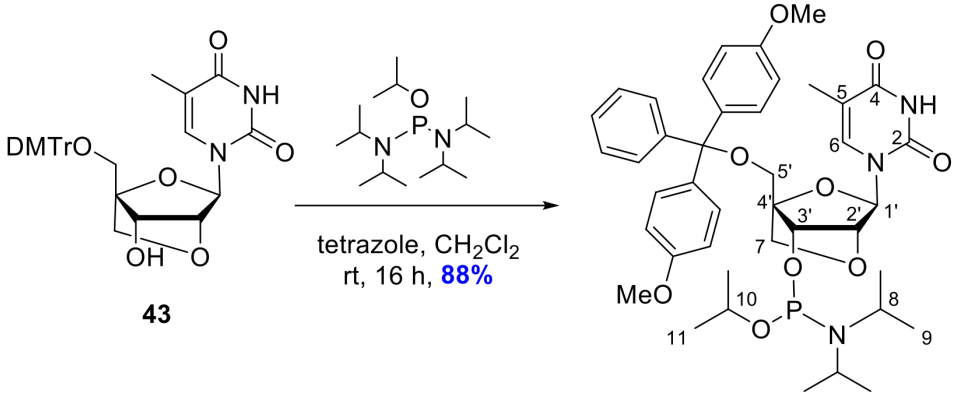
$C_{15}H_{35}N_2OP$

290.43

Chloro-bis(diisopropyl)phosphine (**82**) (2.00 g, 7.50 mmol) was dissolved in anhydrous, degassed Et_2O (25 mL) and anhydrous, degassed Et_3N (3.14 mL, 22.5 mmol) was added. Anhydrous, degassed isopropanol (1.15 mL, 15 mmol) was added and the reaction was stirred at rt for 3 h and monitored by ^{31}P NMR. Upon completion, a white precipitate had formed. The solution was filtered from the salt by-product by cannula and purified by distillation (84 °C at >1 mbar) to give **87** (835 µL, 3.45 mmol, 46%) as a colourless oil.

d = 1.2 g/mol; ^{31}P NMR (162 MHz, $CDCl_3$) δ 114.99; ^1H NMR (400 MHz, $CDCl_3$) δ 3.91 (dp, J = 10.3, 6.1 Hz, 1H, C(3)H), 3.51 (dp, J = 10.7, 6.7 Hz, 4H, 4 × C(1)H), 1.20 (d, J = 6.1 Hz, 6H, 2 x C(4)H_3), 1.16 (dd, J = 6.9, 5.0 Hz, 24H, 8 × C(2)H_3); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{15}H_{36}H_2OP$, 291.2560; found, 291.2556.

(1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl isopropyl diisopropyl phosphoramidite (**91**)



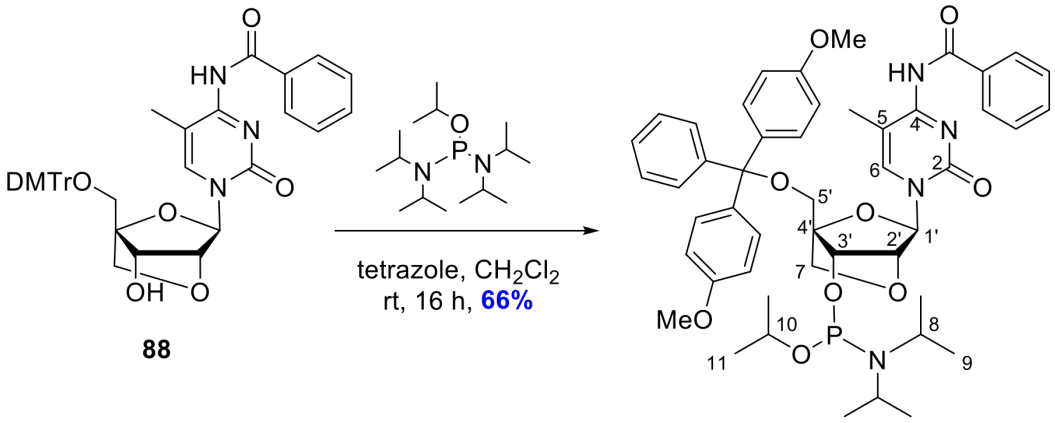
91
$C_{41}H_{52}N_3O_9P$
761.85

Compound **43** (200 mg, 0.349 mmol) was dissolved in dry, degassed CH_2Cl_2 (3 mL) and tetrazole (0.45 M in MeCN) (776 µL, 0.349 mmol) and compound **87** (169 µL, 0.700 mmol) were added and the reaction was stirred in an inert atmosphere at rt overnight. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude residue was purified by inert column chromatography (40% EtOAc/Pet. Ether, 0.5% Et_3N) to give the phosphoramidite **91** (235 mg, 0.308 mmol, 88%) as a white foam.

R_f: 0.45 (EtOAc:Pet Ether, 2:3); ^{31}P NMR (162 MHz, $CDCl_3$) δ 147.44, 146.25; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.36 – 8.89 (m, 1H, N*H*), 7.70 (dt, J = 5.9, 1.3 Hz, 1H, C(6)*H*), 7.47 (ddd, J = 8.1, 4.1, 1.4 Hz, 2H, *H*-Ar), 7.19 – 7.40 (m, 7H, *H*-Ar), 6.84 (dp, J = 8.5, 3.2 Hz, 4H, *H*-Ar), 5.66 (d, J = 5.2 Hz, 1H, C(1')*H*), 4.55 (d, J = 15.3 Hz, 1H, C(2')*H*), 4.30 (d, J = 9.3 Hz, 0.5H, C(3')*H*), 4.20 (d, J = 7.1 Hz, 0.5H, C(3')*H*), 4.11 (m, 0.5H, C(10)*H*), 4.02 – 3.90 (m, 0.5H, C(10)*H*), 3.92 – 3.82 (m, 2H, C(7)H2 or C(5')H_2), 3.79 (t, J = 1.2 Hz, 6H, 2 × CH_3-OMe), 3.51 (m, 4H, C(7)H_2 or C(5')H_2 and 2 × C(8)*H*), 1.60 (dd, J = 20.3, 1.2 Hz, 3H, CH_3-thymine), 1.15 – 0.99 (m, 18H, 2 × C(11)H_3 and 4 × C(9)H_3); LRMS-ESI

(m/z): $[M+H]^+$ 762 (100%), $[M+Na]^+$ 784 (46%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{41}H_{53}N_3O_9P$, 762.3514; found, 762.3504.

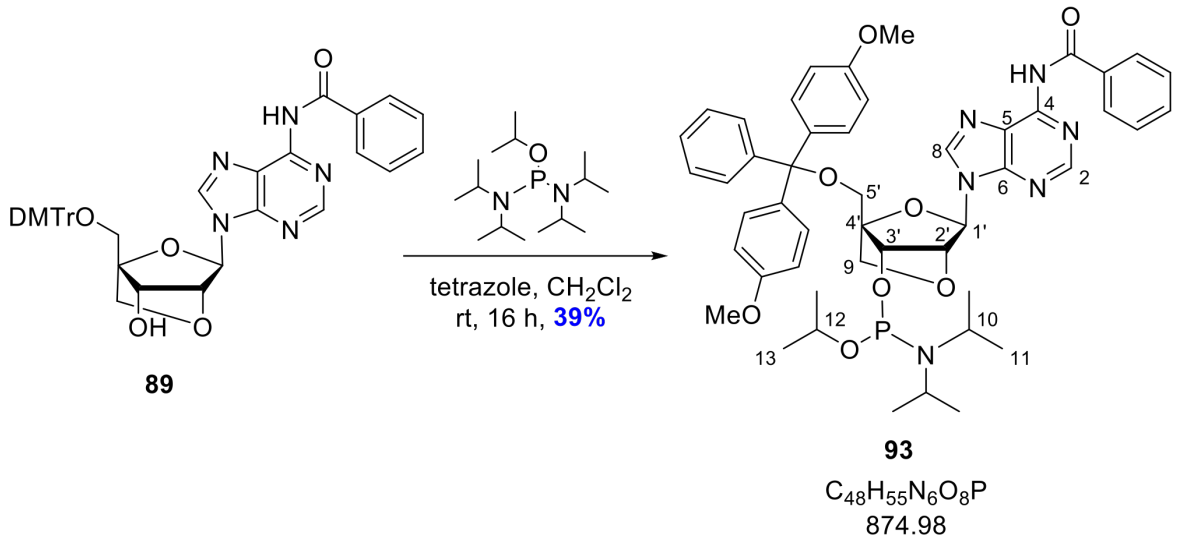
(1*R*,3*R*,4*R*,7*S*)-3-(4-benzamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((bis(4-methoxy phenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl isopropyl diisopropyl phosphoramidite (**92**)



Compound **88** (200 mg, 0.300 mmol) was dissolved in dry, degassed CH_2Cl_2 (3 mL) and tetrazole (0.45 M in MeCN) (667 µL, 0.300 mmol) and compound **87** (reagent) (109 µL, 0.450 mmol) was added and the reaction was stirred in an inert atmosphere at rt overnight. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude residue was purified by inert column chromatography (20% EtOAc/Pet. Ether, 0.5% Et_3N) to give the phosphoramidite **92** (169 mg, 0.195 mmol, 66%) as a white foam.

R_f: 0.45 (EtOAc:Pet. Ether, 1:4); ^{31}P NMR (162 MHz, $CDCl_3$) δ 147.56, 146.4; HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{48}H_{58}N_4O_9P$, 865.3936; found, 865.3921.

(1*R*,3*R*,4*R*,7*S*)-3-(6-benzamido-9*H*-purin-9-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy) methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl isopropyl diisopropylphosphoramidite (**93**)



89

93

$C_{48}H_{55}N_6O_8P$

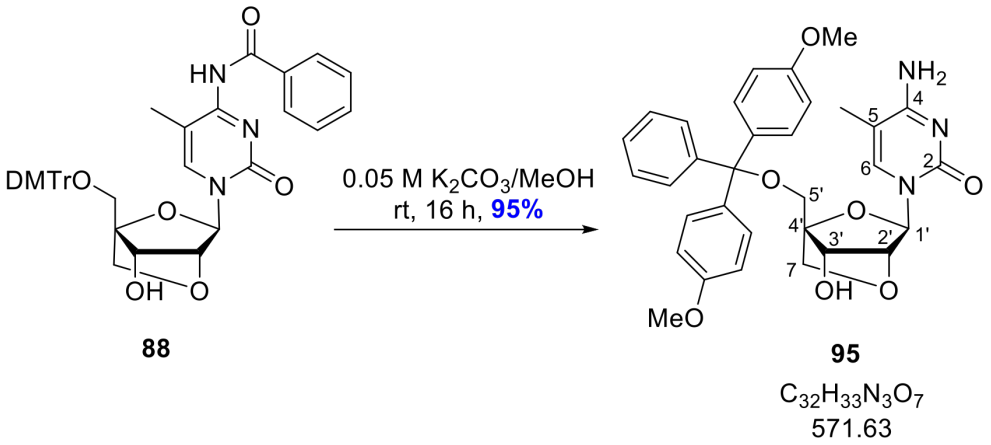
874.98

Compound **89** (200 mg, 0.292 mmol) was dissolved in dry, degassed CH_2Cl_2 (3 mL) and tetrazole (0.45 M in MeCN) (649 µL, 0.292 mmol) and compound **87** (reagent) (140 µL, 0.580 mmol) was added and the reaction was stirred in an inert atmosphere at rt overnight. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude residue was purified by inert column chromatography (40% EtOAc/Pet. Ether, 0.5% Et_3N), precipitated from degassed hexanes (3 × 100 mL), and submitted to column chromatography a second time (10-30% EtOAc/Pet. Ether, 0.5% Et_3N) to give the phosphoramidite **93** (99 mg, 0.113 mmol, 39%) as a white foam.

R_f: 0.75 (EtOAc:Pet. Ether, 2:3); ^{31}P NMR (203 MHz, $CDCl_3$) δ 147.29, 146.28; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.05 (s, 1H, N*H*), 8.80 (d, J = 4.8 Hz, 1H, C(2)*H*), 8.31 (d, J = 6.1 Hz, 1H, C(8)*H*), 8.06 – 7.99 (m, 2H, *H*-Ar), 7.66 – 7.44 (m, 5H, *H*-Ar), 7.38 – 7.18 (m, 7H, *H*-Ar), 6.88 – 6.79 (m, 4H, *H*-Ar), 6.16 (d, J = 1.5 Hz, 1H, C(1')*H*), 4.91 (d, J = 5.1 Hz, 1H, C(2')*H*), 4.36 (d, J = 8.9 Hz, 0.5H, C(3')*H*), 4.28 (d, J = 6.9 Hz, 0.5H, C(3')*H*), 4.15 – 3.99 (m, 2H, C(5')H2 or C(9)H_2), 3.98 – 3.89 (m, 1H, C(12)*H*), 3.79 (d, J = 0.7 Hz, 6H,

2 × CH_3-OMe), 3.60 – 3.36 (m, 4H, C(5')H_2 or C(9)H_2 and 2 × C(10)H), 1.14 – 0.76 (m, 18H, 2 × C(13)CH_3 and 4 × C(11)H_3); LRMS-ESI (m/z): $[M]^+$ 875 (100%), $[M+H]^+$ 876 (81%), [M(oxidised)]$^+$ 891 (36%); HRMS-ESI (m/z): $[M+Na]^+$ calcd. for $C_{48}H_{55}N_6O_8PNa$, 897.3711; found, 897.3757.

4-amino-1-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidin-2(1H)-one (**95**)

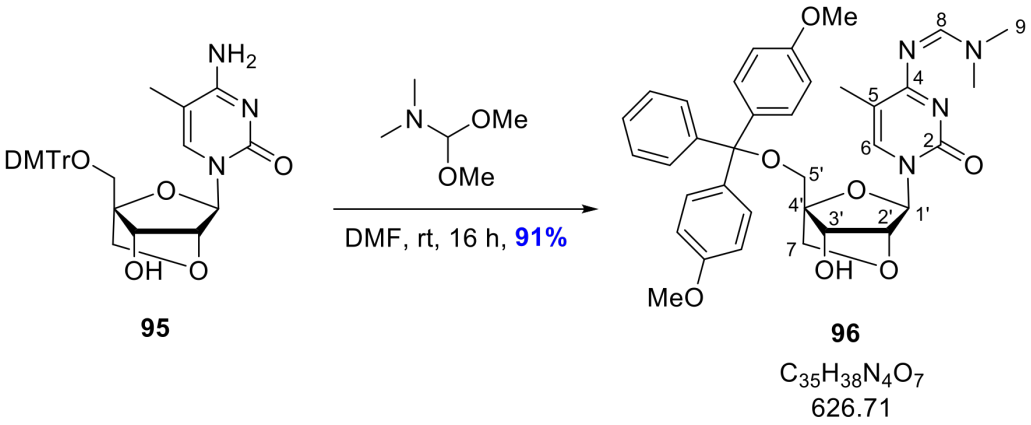


Compound **88** (500 mg, 0.740 mmol) was dissolved in a solution of 0.05 M K_2CO_3in MeOH (10 mL) and stirred at rt overnight. The solvent was removed under reduced pressure and the crude was purified by column chromatography (0-10% MeOH/CH_2Cl_2) to give **95** (402 mg, 0.703 mmol, 95%) as a white powder.

R_f: 0.38 (CH_2Cl_2:MeOH, 9:1); ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 7.62 – 7.57 (m, 1H, C(6)H), 7.31 – 7.16 (m, 5H, H-Ar), 7.11 – 7.05 (m, 4H, H-Ar), 6.87 – 6.81 (m, 4H, H-Ar), 5.39 (d, J = 0.7 Hz, 1H, C(1')H), 4.07 (s, 1H, C(2')H), 3.85 (s, 1H, C(3')H), 3.81 (d, J = 7.7 Hz, 1H, C(7)H_A), 3.76 (d, J = 2.1 Hz, 2H, C(5')H_2), 3.73 (s, 6H, 2 × CH_3-OMe), 3.63 (d, J = 7.6 Hz, 1H, C(7)H_B), 1.85 (d, J = 1.0 Hz, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, DMSO-d$_6$) δ 174.9 (C-4), 167.3 (C-Ar), 164.1 (C-2), 157.8 (C-Ar), 149.7 (C-Ar), 146.6 (C-6), 138.7 (C-Ar), 138.4 (C-Ar), 137.9 (C-Ar), 137.1 (C-Ar), 136.9 (C-Ar), 135.9 (C-Ar), 122.2 (C-Ar), 110.0 (C-5), 98.0 (C-4'), 96.2 (C-1'), 89.4 (C-DMTr), 88.5 (C-2'), 80.4 (C-7), 77.9 (C-3'), 65.6 (C-5'), 64.5 (CH$_3$-OMe), 22.8 (CH$_3$-thymine); LRMS-ESI (m/z):

$[DMTr]^+$ 303 (100%), $[2(M\text{-}DMTr)+H]^+$ 539 (97%); HRMS-ESI (m/z): $[M+Na]^+$ calcd for $C_{32}H_{33}N_3O_7Na$, 594.2211; found, 594.2212.

(*Z*)-*N*'-(1-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)-*N*,*N*-dimethyl formimidamide (**96**)

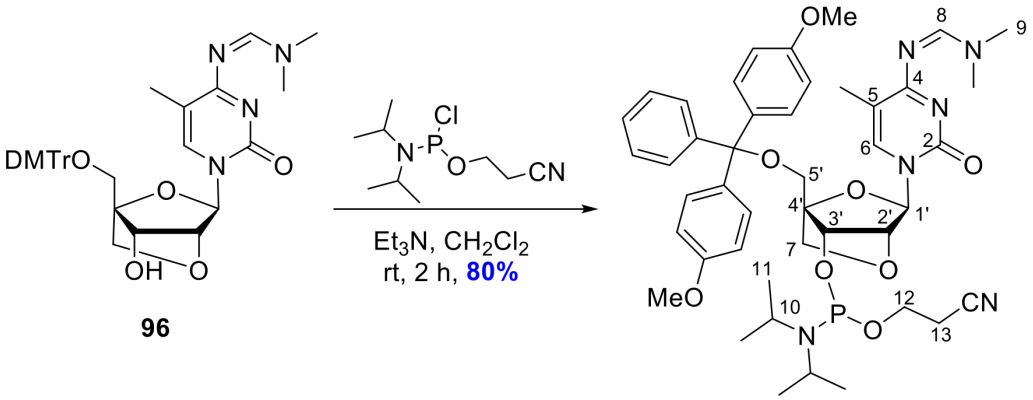


Compound **95** (200 mg, 0.350 mmol) was dissolved in anhydrous DMF (4 mL) and *N*,*N*-dimethylformamide dimethylacetal (196 µL, 1.40 mmol) was added and left to stir at rt overnight. The solvent was removed under reduced pressure and the residue co-evaporated in toluene (3 × 5 mL). The residue was dissolved in CH_2Cl_2, washed with sat. $NaHCO_3$ (2 × 10 mL), dried over Na_2SO_4, and reduced under pressure. The crude was purified by column chromatography in (0-10% MeOH/CH_2Cl_2) to give **96** (200 mg, 0.319 mmol, 91%) as a white powder.

R_f: 0.75 (MeOH:CH_2Cl_2, 1:9); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.75 (s, 1H, C(8)*H*), 7.77 (d, J = 1.1 Hz, 1H, C(6)*H*), 7.45 – 7.53 (m, 2H, *H*-Ar), 7.37 (dd, J = 9.0, 1.0 Hz, 5H, *H*-Ar), 7.20 – 7.35 (m, 2H, *H*-Ar), 6.81 – 6.89 (m, 4H, *H*-Ar), 5.76 (d, J = 0.7 Hz, 1H, C(1')*H*), 4.55 (s, 1H, C(2')*H*), 4.25 (s, 1H, C(3')*H*), 3.80 – 3.90 (m, 2H, C(7)H_2), 3.79 (d, J = 0.7 Hz, 6H, 2 × CH_3-OMe), 3.54 (d, J = 10.9 Hz, 1H, C(5')H_A), 3.46 (d, J = 10.9 Hz, 1H, C(5')H_B), 3.10 – 3.16 (m, 6H, 2 × C(9)H_3), 1.88 (d, J = 1.0 Hz, 3H, CH_3-thymine); ^{13}C NMR (101 MHz, $CDCl_3$) δ 158.6 (*C*-4), 156.0 (*C*-8), 144.1 (*C*-2), 137.7 (*C*-6), 135.6 (*C*-Ar), 130.1 (*C*-Ar), 128.1 (*C*-Ar), 128.0 (*C*-Ar), 127.0 (*C*-Ar), 113.3 (*C*-Ar), 110.3 (*C*-5), 87.8

(*C*-1'), 86.6 (*C*-4'), 79.4 (*C*-2'), 77.4 (*C*-DMTr), 71.7 (*C*-7), 70.4 (*C*-3'), 58.7 (*C*-5'), 55.3 (CH_3-OMe), 41.3 (*C*-9), 35.1 (*C*-9), 14.2 (CH_3-thymine); LRMS-ESI (*m/z*): $[DMTr]^+$ 303 (4%), $[M+H]^+$ 627 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{35}H_{39}N_4O_7$, 627.2813; found, 627.3588.

(1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(4-(((*Z*)-(dimethylamino) methylene)amino)-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite (**98**)

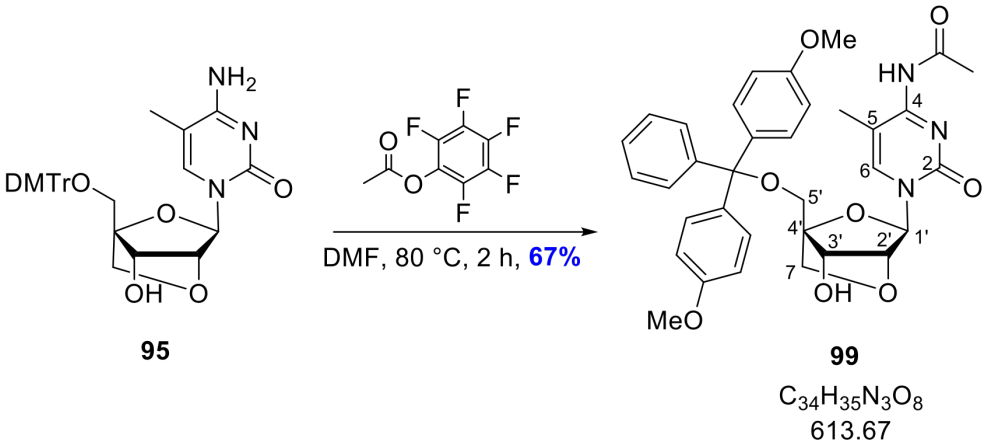


Compound **96** (185 mg, 0.295 mmol) was dissolved in anhydrous degassed CH_2Cl_2 (3 mL) and anhydrous, degassed Et_3N (103 µL, 0.738 mmol) and chloro-phosphitylating agent (105 µL, 0.443 mmol) were added and the reaction was stirred under an inert atmosphere at rt for 2 h. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was redissolved in the minimum volume of anhydrous degassed CH_2Cl_2 (0.2-0.5 mL) and precipitated from anhydrous, degassed hexane (5 × 50 mL) and the solvent was removed under reduced pressure to give **98** (195 mg, 0.234 mmol, 80%) as a white foam.

R_f: 0.60 (MeOH:CH_2Cl_2, 1:19); ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.22, 149.05; LRMS-ESI (*m/z*): $[M(oxidised)\text{-dmf}+Na]^+$ 758 (51%), $[M]^+$ 827 (100%), $[M+Na]^+$ 849 (27%); HRMS-

ESI (m/z): $[M+Na]^+$ calcd. for $C_{44}H_{55}N_6O_8PNa$, 849.3711; found, 849.3717.

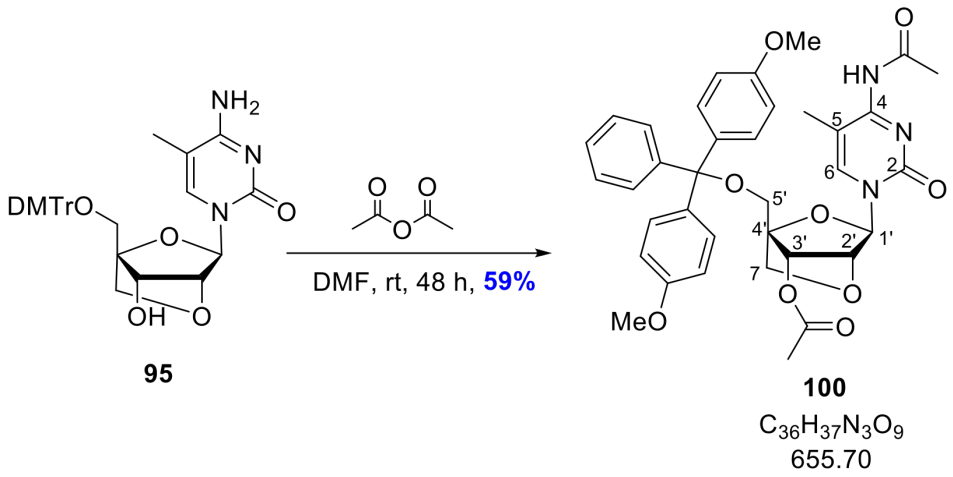
N-(1-((1*R*,3*R*,4*R*,7*S*)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)acetamide (**99**)



Compound **95** (108 mg, 0.189 mmol) was dissolved in DMF (4 mL) and pentafluorophenyl acetate (64 mg, 0.284 mmol) was added and the reaction was stirred at rt for 5 h and heated to 80 °C for 2 h. The reaction was cooled to rt, the solvent removed under reduced pressure, and the crude was purified by column chromatography (0-10% MeOH/CH_2Cl_2) to give **99** (78 mg, 0.127 mmol, 67%) as a white powder.

R_f: 0.55 (MeOH:CH_2Cl_2, 1:9); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (s, 1H, N*H*), 7.42 – 7.35 (m, 2H, C(5)*H* and *H*-Ar), 7.32 – 7.14 (m, 8H, *H*-Ar), 6.83 – 6.74 (m, 4H, *H*-Ar), 5.62 (s, 1H, C(1')*H*), 4.48 (s, 1H, C(2')*H*), 4.17 (s, 1H, C(3')*H*), 3.83 – 3.75 (m, 2H, C(7)H_2), 3.73 (d, J = 1.0 Hz, 6H, 2 × CH_3-OMe), 3.52 (d, J = 11.1 Hz, 1H, C(5')H_A), 3.41 (d, J = 11.1 Hz, 1H, C(5')H_B), 1.68 (d, J = 1.0 Hz, 3H, CH_3-thymine), 1.18 (s, 3H, CH_3-OAc); ^{13}C NMR (151 MHz, $CDCl_3$) δ 162.6 (*C*-4), 158.9 (*C*-Ar), 144.6 (*C*-Ar), 135.5 (*C*-Ar), 130.3 (*C*-Ar), 130.2 (*C*-Ar), 130.2 (*C*-Ar), 128.2 (*C*-Ar or C-6), 128.2 (*C*-Ar or C-6), 127.3 (*C*-Ar), 113.5 (*C*-Ar), 88.4 (*C*-4'), 87.8 (*C*-1'), 86.9 (*C*-DMTr), 79.1 (*C*-2'), 71.7 (*C*-7), 70.4 (*C*-3'), 58.3 (*C*-5'), 55.4 (*C*H_3-OMe), 29.8 (*C*H_3-OAc), 13.8 (*C*H_3-thymine); LRMS-ESI (m/z): $[DMTr]^+$ 303 (100%), $[M+H]^+$ 614 (94%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{34}H_{36}N_3O_8$, 614.2497; found, 614.2489; $[M+Na]^+$ calcd. for $C_{34}H_{35}N_3O_8Na$, 636.2316; found, 636.2312.

(1*R*,3*R*,4*R*,7*S*)-3-(4-acetamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (**100**)



95

100

$C_{36}H_{37}N_3O_9$

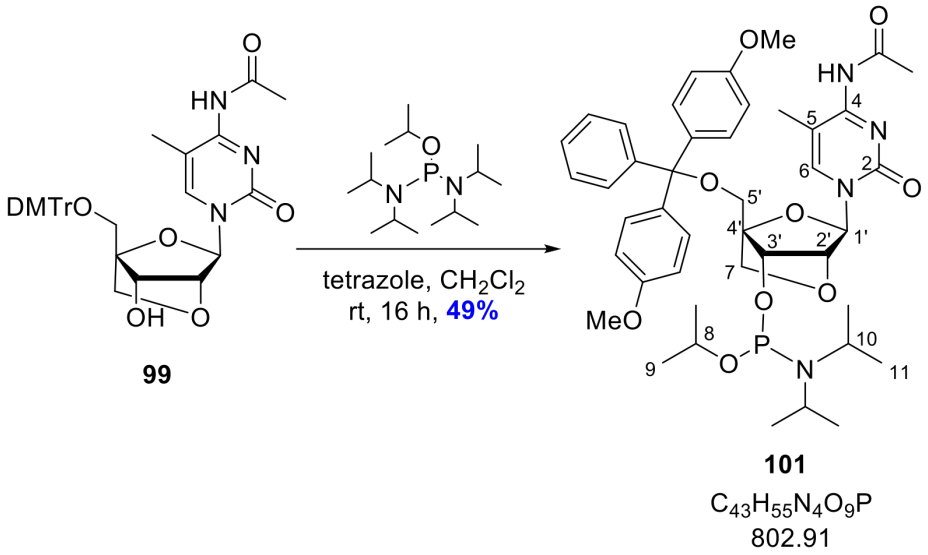
655.70

Compound **95** (450 mg, 0.787 mmol) was dissolved in anhydrous DMF (3 mL) and Ac_2O was added (93 µL, 0.980 mmol) and stirred at rt for 24 h. Over the course of the next 24 h, additional portions of Ac2O were added (50 µL, 0.527 mmol, then 100 µL, 1.05 mmol). The solvent was removed under reduced pressure and the crude residue was purified by column chromatography (0-10% MeOH/CH_2Cl_2) to give the doubly-acetylated by-product **100** (303 mg, 0.462 mmol, 59%) as well as the single-acetylated product **99** (110 mg, 0.179 mmol, 23%).

R_f: 0.60 (MeOH:CH_2Cl_2, 1:9); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.95 (s, 1H, N*H*), 7.87 (s, 1H, C(6)*H*), 7.43 – 7.15 (m, 9H, *H*-Ar), 6.96 – 6.88 (m, 4H, *H*-Ar), 5.59 (s, 1H, C(1')*H*), 5.07 (s, 1H, C(3')*H*), 4.57 (s, 1H, C(2')*H*), 3.84 (d, J = 8.4 Hz, 1H, C(7)H_A), 3.74 (s, 7H, C(7)H_B and 2 × CH_3-OMe), 3.43 (d, J = 12.4 Hz, 2H, C(5')H_2), 2.27 (s, 3H, C(4)NH-CH_3-OAc), 2.01 (s, 3H, C(3')-CH_3-OAc), 1.82 – 1.78 (m, 3H, CH_3-thymine); ^{13}C NMR (151 MHz, DMSO-d_6) δ 169.3 (*C*O-OAc), 162.3 (*C*-4), 158.2 (*C*-Ar), 153.3 (*C*-2), 144.4 (*C*-Ar), 140.4 (*C*-6), 135.1 (*C*-Ar), 134.9 (*C*-Ar), 129.7 (*C*-Ar), 129.6 (*C*-Ar), 128.0 (*C*-Ar), 127.6 (*C*-Ar), 126.9 (*C*-Ar), 113.4 (*C*-Ar), 105.7 (*C*-5), 87.4 (*C*-1'), 86.7 (*C*-4'), 85.9 (*C*-DMTr), 76.8 (*C*-2'), 71.7 (*C*-7), 70.3 (*C*-3'), 57.7 (*C*-5'), 55.1 (*C*H_3-OMe), 24.9 (*C*(4)NH-CH_3-OAc), 20.5 (C(3')-*C*H_3-OAc), 14.1 (*C*H_3-thymine); LRMS-ESI (m/z): $[DMTr]^+$ 303 (28%), $[M+H]^+$

656 (100%), $[M+Na]^+$ 698 (31%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{36}H_{38}N_3O_9$, 656.2603; found, 656.2612; $[M+Na]^+$ calcd. for $C_{36}H_{37}N_3O_9Na$, 678.2422; found, 678.2413.

(1*R*,3*R*,4*R*,7*S*)-3-(4-acetamido-5-methyl-2-oxopyrimidin-1(2*H*)-yl)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl isopropyl diisopropyl phosphoramidite (**101**)



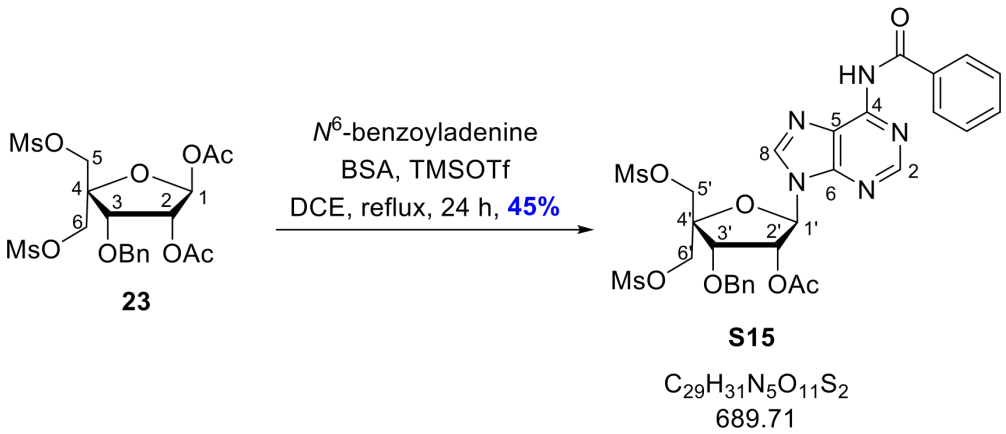
Compound **99** (173 mg, 0.282 mmol) was dissolved in dry degassed CH_2Cl_2 (3 mL) and tetrazole (0.45M in MeCN) (626 µL, 0.282 mmol) and compound **87** (reagent) (103 µL, 0.423 mmol) were added, and the reaction was stirred under an inert atmosphere at rt overnight. An inert aqueous work-up was conducted with degassed sat. KCl (10 mL), which was washed with anhydrous, degassed CH_2Cl_2 (2 × 5 mL); the organic layers were combined and dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude residue was purified by inert column chromatography (100% EtOAc, 0.5% Et_3N) to give the phosphoramidite **101** (116 mg, 0.144 mmol, 49%) as a white foam.

R_f: 0.85 (EtOAc); ^{31}P NMR (203 MHz, CD2Cl2) δ 147.59 (dd, J = 10.4, 7.4 Hz), 146.39 (p, J = 9.9 Hz); ^1H NMR (400 MHz, $CDCl_3$) δ 7.97 (s, 1H, NH), 7.47 (ddt, J = 6.0, 4.8, 1.4 Hz, 2H, C(6)H and H-Ar), 7.40 – 7.19 (m, 8H, H-Ar), 6.89 – 6.78 (m, 4H, H-Ar), 5.72 (s, 1H, C(1')H), 4.67 (s, 1H, C(2')H), 4.29 (d, J = 9.1 Hz, 0.5H, C(3')H), 4.18 (d, J = 7.0 Hz, 0.5H, C(3')H), 4.09 – 3.96 (m, 0.5H, C(10)H), 3.87 (m, 2.5H, C(8)H and C(7)H_2

or C(5')H_2), 3.79 (t, J = 1.4 Hz, 6H, 2 × CH_3-OMe), 3.62 – 3.38 (m, 4H, C(10)H_2 and C(7)H_2 or C(5')H_2), 2.65 (s, 3H, CH_3-OAc), 1.62 (dd, J = 23.9, 1.0 Hz, 3H, CH_3-thymine), 1.18 – 0.94 (m, 18H, 2 × C(9)CH_3 and 4 × C(11)H_3); LRMS-ESI (m/z): $[M]^+$ 803 (100%), $[M+H]^+$ 804 (52%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{43}H_{56}N_4O_9P$, 803.3779; found, 803.3765.

7.1.5 Compounds in Appendix

(2*R*,3*R*,4*S*)-2-(6-benzamido-9*H*-purin-9-yl)-4-(benzyloxy)-5,5-bis(((methylsulfonyl)oxy) methyl)tetrahydrofuran-3-yl acetate (**S15**)

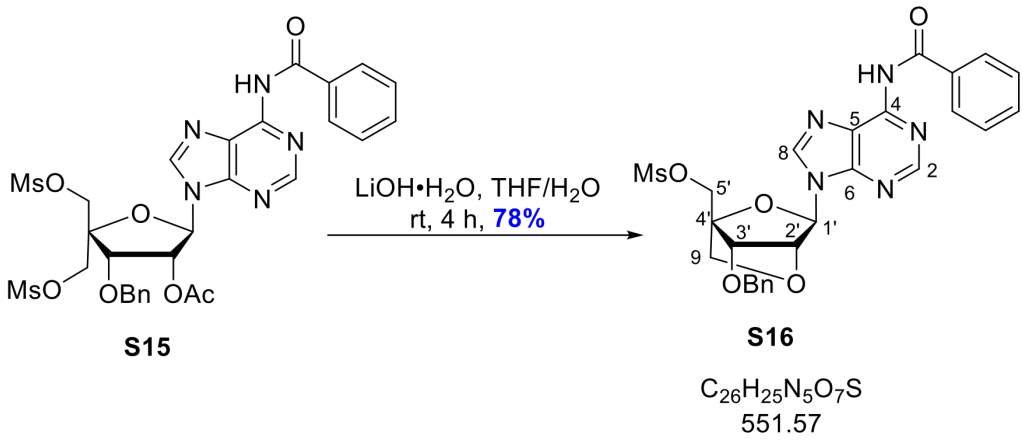


Compound **23** (5.5 g, 10.8 mmol), N^6-benzoyladenine (3.22 g, 13.5 mmol), and BSA (7.26 mL, 29.7 mmol) were suspended in anhydrous dichloroethane (75 mL). The reaction mixture was refluxed under argon for 1 h. After cooling to rt, TMSOTf was added (3.90 mL, 21.5 mmol) and refluxed for a further 24 h. Once cooled to rt, the solution was poured into ice cold sat. $NaHCO_3$ (100 mL) and stirred vigorously for 30 min. The solution was filtered and the phases separated. The organic phase was washed with sat. $NaHCO_3$ (3 × 50 mL). The organic layers were combined, dried over Na_2SO_4, and the solvent was removed under reduced pressure. The residue was purified by column chromatography (0-5% MeOH/CH_2Cl_2) and again (0-80% EtOAc/Pet. Ether) to give **S15** as a light yellow foam (3.35 g, 4.86 mmol, 45%).

R_f: 0.75 (CH_2Cl_2:MeOH, 95:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.01 (s, 1H, NH), 8.78 (s, 1H,

C(2)*H*), 8.08 (s, 1H, C(8)*H*), 8.03 (m, 2H, *H*-Ar(Bz)), 7.60-7.62 (m, 1H, *H*-Ar(Bz)), 7.51-7.55 (m, 2H, *H*-Ar(Bz)), 7.35-7.40 (m, 5H, *H*-Ar(OBn)), 6.21 (d, J = 3.3 Hz, 1H, C(1')*H*), 6.09 (dd, J = 5.9, 3.3 Hz, 1H, C(2')*H*), 5.12 (d, J = 5.9 Hz, 1H, C(3')*H*), 4.63-4.71 (m, 3H, C(5' or 6')H2 and CH_A-OBn), 4.35-4.45 (m, 3H, C(5' or 6')H2 and CH_B-OBn), 3.04 (s, 3H, CH_3-OMs), 2.87 (s, 3H, CH_3-OMs), 2.14 (s, 3H, CH_3-OAc); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.8 (*C*OCH$_3$-OAc), 153.0 (*C*-2), 142.6 (*C*-8), 136.6 (*C*-Ar(OBn)), 133.1 (*C*-Ar(Bz)), 128.9 (*C*-Ar), 128.7 (*C*-Ar), 128.0 (*C*-Ar), 88.2 (*C*-1'), 84.4 (*C*-4'), 77.6 (*C*-3'), 74.9 (*C*H$_2$-OBn), 73.7 (*C*-2'), 67.5 (*C*-5' or 6'), 67.5 (*C*-5' or 6'), 38.0 (*C*H$_3$-OMs), 37.7 (*C*H$_3$-OMs), 20.8 (*C*H$_3$-OAc); LRMS-ESI (*m/z*): $[M]^+$ 689 (100%); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{29}H_{32}N_5O_{11}S_2$, 690.1534; found, 690.1574.

((1*R*,3*R*,4*R*,7*S*)-3-(6-benzamido-9*H*-purin-9-yl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl methanesulfonate (**S16**)

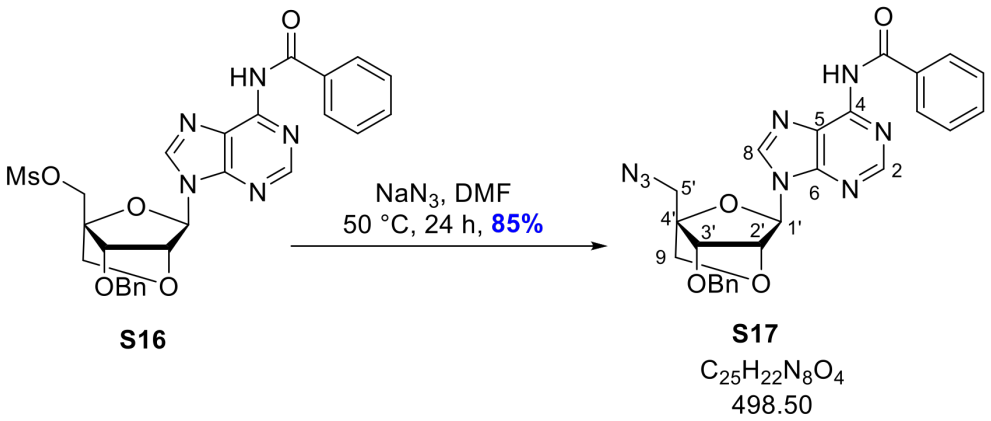


Compound **S15** (453 mg, 0.657 mmol) was dissolved in a solution of THF/H_2O (3:2, v/v) (10 mL). LiOH•H_2O (138 mg, 3.29 mmol) was added and the mixture was stirred at rt for 4 h. The reaction mixture was neutralised with acetic acid (150 µL) and water (10 mL) was added to produce a precipitate. The mixture was filtered and the precipitate washed with water and dried under vacuum to yield **S16** (281 mg, 0.510 mmol, 78%) as an off-white powder.

R_f: 0.67 (EtOAc:Pet. Ether, 8:2); ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.24 (bs, 1H, N*H*), 8.74 (s, 1H, C(2)*H*), 8.53 (s, 1H, C(8)*H*), 8.05 (d, 2H, *H*-Ar(Bz)), 7.66 (m, 1H, *H*-Ar(Bz)),

7.56 (m, 2H, *H*-Ar(Bz)), 7.29-7.36 (m, 5H, *H*-Ar(OBn)), 6.19 (s, 1H, C(1')*H*), 4.97 (s, 1H, C(2')*H*), 4.83 (d, J = 12.0 Hz, 1H, CH_A-OBn), 4.70 (s, 2H, C(5')H_2), 4.64 (d, J = 12.0 Hz, 1H, CH_B-OBn), 4.58 (s, 1H, C(3')*H*), 4.10 (d, J = 8.1 Hz, 1H, C(9)H_A), 3.98 (d, J = 8.1 Hz, 1H, C(9)H_B), 3.25 (s, 3H, CH_3-OMs); ^{13}C NMR (101 MHz, DMSO-d$_6$) δ 165.6 (*C*O-Bz), 151.7 (*C*-2), 151.4 (*C*-6), 150.4 (*C*-4), 141.7 (*C*-8), 137.6 (*C*-Ar(OBn)), 133.3 (*C*-Ar(Bz)), 132.4 (*C*-Ar(Bz)), 127.6-128.5 (*C*-Ar(OBn)), 125.6 (*C*-5), 85.6 (*C*-1'), 84.7 (*C*-4'), 77.8 (*C*-3'), 76.9 (*C*-2'), 71.7 (*C*-9), 71.3 (*C*-5'), 65.9 (*C*H$_2$-OBn), 36.9 (*C*H$_3$-OMs); LRMS-ESI (m/z): [M]$^+$ 551 (100%); HRMS-ESI (m/z): [M+H]$^+$ calcd. for $C_{26}H_{26}N_5O_7S$, 552.1548; found, 552.1596; [M+Na]$^+$ calcd. for $C_{26}H_{25}N_5O_7SNa$, 574.1367; found, 574.1367.

N-(9-((1*S*,3*R*,4*R*,7*S*)-1-(azidomethyl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-9*H*-purin-6-yl)benzamide (**S17**)



S16

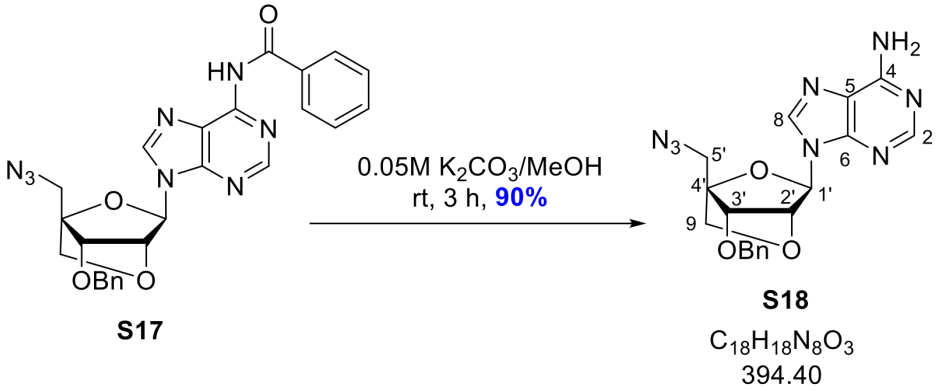
S17
$C_{25}H_{22}N_8O_4$
498.50

Compound **S16** (2.70 g, 4.89 mmol) was dissolved in DMF (80 mL) and sodium azide (954 mg, 14.67 mmol) was added and the reaction mixture was stirred at 50 °C for 24 h. The solvent was removed under vacuum and the residue dissolved in EtOAc (50 mL). The organic layer was washed with sat. $NaHCO_3$ (2 × 50 mL), dried over Na_2SO_4, and the solvent was removed under vacuum to give **S17** (2.08 g, 4.17 mmol, 85%) as a crystalline off-white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.05 (s, 1H, N*H*), 8.75 (s, 1H, C(2)*H*), 8.18 (s, 1H, C(8)*H*), 8.02-8.05 (m, 2H, *H*-Ar(Bz)), 7.59-7.64 (m, 1H, *H*-Ar(Bz)), 7.51-7.55 (m, 2H, *H*-Ar(Bz)), 7.28-7.30 (m, 5H, *H*-Ar(OBn)), 6.09 (s, 1H, C(1')*H*), 4.94 (s, 1H, C(2')*H*), 4.65 (d, J = 11.7 Hz, 1H, CH_A-OBn), 4.58 (d, J = 11.7 Hz, 1H, CH_B-OBn), 4.24 (s, 1H, C(3')*H*), 4.14

(d, J = 8.0 Hz, 1H, C(9)H_A), 3.98 (d, J = 7.8 Hz, 1H, C(9)H_B), 3.74 (d, J = 13.8 Hz, 1H, C(5')H_A), 3.65 (d, J = 13.7 Hz, 1H, C(5')H_B); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.7 (CO-Bz), 152.9 (C-2), 150.9 (C-6), 149.8 (C-4), 140.7 (C-8), 136.8 (C-Ar(OBn)) 133.6 (C-Ar(Bz)), 133.0 (C-Ar(Bz)), 129.1 (C-Ar(Bz)), 128.5 (C-Ar(OBn)), 127.9 (C-Ar(Bz)), 123.7 (C-5), 87.1 (C-1'), 87.0 (C-4'), 77.8 (C-3'), 77.0 (C-2'), 73.0 (C-9), 72.7 (CH$_2$-OBn), 47.7 (C-5'); LRMS-ESI (m/z): $[M]^+$ 498 (100%); HRMS-ESI (m/z): $[M+H]^+$ calcd. for $C_{25}H_{22}N_8O_4$, 499.1837; found, 499.1829.

9-((1*S*,3*R*,4*R*,7*S*)-1-(azidomethyl)-7-(benzyloxy)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-9*H*-purin-6-amine (**S18**)



S17

S18

$C_{18}H_{18}N_8O_3$

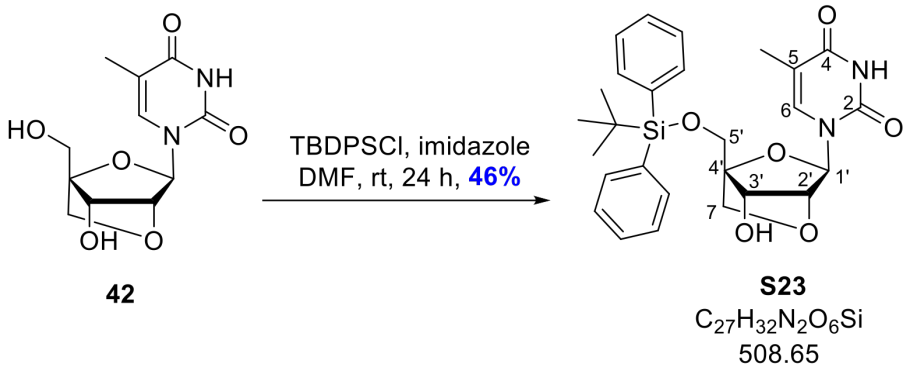
394.40

Compound **S17** (223 mg, 0.447 mmol) was dissolved in 0.05 M K_2CO_3/MeOH (20 mL) and the reaction mixture was stirred at rt for 3 h. The solvent was removed under vacuum and purified by column chromatography (0-10% MeOH/CH_2Cl_2) to give **S18** (159 mg, 0.403 mmol, 90%) as a white powder.

R_f: 0.67 (CH_2Cl_2:MeOH, 9:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.32 (s, 1H, C(2)H), 7.95 (s, 1H, C(8)H), 7.27-7.32 (m, 5H, H-Ar), 6.03 (s, 1H, C(1')H), 4.92 (s, 1H, C(2')H), 4.66 (d, J = 11.7 Hz, 1H, H, CH_A-OBn), 4.57 (d, J = 11.6, 1H, CH_B-OBn), 4.23 (s, 1H, C(3')H), 4.12 (d, J = 7.8 Hz, 1H, C(9)H_A), 3.97 (d, J = 7.8 Hz, 1H, C(9)H_B), 3.73 (d, J = 13.7 Hz, 1H, C(5')H_A), 3.64 (d, J = 13.7 Hz, 1H, C(5')H_B); ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.4 (C-6), 152.7 (C-2), 148.9 (C-4), 138.4 (C-8), 136.9 (C-Ar), 128.0 (C-Ar), 128.7 (C-Ar), 120.1 (C-5), 87.0 (C-1'), 86.8 (C-4'), 77.9 (C-3'), 77.0 (C-2'), 73.0 (C-9), 72.6 (CH$_2$-OBn), 47.7

(*C*-5'); HRMS-ESI (*m/z*): $[M+H]^+$ calcd. for $C_{18}H_{18}N_8O_3$, 395.1575; found, 395.1574. NB: Used in the monomer approach on route to 5'-NH-MMTr adenine LNA phosphoramidite.

1-((1*R*,3*R*,4*R*,7*S*)-1-(((tert-butyldiphenylsilyl)oxy)methyl)-7-hydroxy-2,5-dioxa bicyclo[2.2.1]heptan-3-yl)-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (**S23**)



42

S23
$C_{27}H_{32}N_2O_6Si$
508.65

Compound **42** (300 mg, 1.11 mmol) was dissolved in anhydrous DMF (5 mL) and tert-butyldiphenylsilyl chloride (TBDSCl) (375 µL, 1.44 mmol) and imidazole (98 mg, 1.44 mmol) were added. After stirring at rt overnight, the DMF was removed under reduced pressure and the crude residue was dissolved in CH_2Cl_2 (20 mL) and washed with H_2O (20 mL), sat. $NaHCO_3$ (20 mL), brine (20 mL), dried over $MgSO_4$ and reduced under pressure. The residue was purified by column chromatography (0-3% MeOH/CH_2Cl_2) to give **S23** (258 mg, 0.51 mmol, 46%) as a white foam.

R_f: 0.33 (MeOH:CH_2Cl_2; 1:19); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.11 (s, 1H, N*H*), 7.76 – 7.66 (m, 4H, *H*-Ar), 7.52 (q, J = 1.2 Hz, 1H, C(6)*H*), 7.50 – 7.35 (m, 6H, *H*-Ar), 5.65 (s, 1H, C(1')*H*), 4.42 (s, 1H, C(2')*H*), 4.20 (d, J = 6.4 Hz, 1H, C(3')*H*), 4.08 (d, J = 12.0 Hz, 1H, C(5')H_A), 3.97 (d, J = 12.1 Hz, 1H, C(5')H_B), 3.91 (d, J = 8.4 Hz, 1H, C(7)H_A), 3.82 (d, J = 8.4 Hz, 1H, C(7)H_B), 1.96 (d, J = 6.5 Hz, 1H, C(3')-O*H*), 1.73 (d, J = 1.3 Hz, 3H, CH_3-thymine), 1.12 (s, 9H, 3 × CH_3-tBu); ^{13}C NMR (151 MHz, $CDCl_3$) δ 163.5 (*C*-4), 149.7 (*C*-2), 135.8 (*C*-Ar), 135.6 (*C*-Ar), 134.4 (*C*-6), 132.8 (*C*-Ar), 132.5 (*C*-Ar), 130.4 (*C*-Ar), 130.3 (*C*-Ar), 128.2 (*C*-Ar), 128.2 (*C*-Ar), 110.6 (*C*-5), 89.3 (*C*-4'), 87.0 (*C*-1'), 79.7 (*C*-2'), 71.4 (*C*-7), 70.5 (*C*-3'), 59.1 (*C*-5'), 27.1 (C(*C*H$_3$)$_3$-tBu), 19.6 (*C*H$_3$-tBu), 12.4 (*C*H$_3$-thymine); LRMS-ESI (*m/z*): $[M+Na]^+$ 531 (100%), $[2M+Na]^+$ 1039 (88%); HRMS-

ESI (m/z): $[M+H]^+$ calcd. for $C_{27}H_{33}N_2O_6Si$, 509.2102 ; found, 509.2097; $[M+Na]^+$ calcd. for $C_{27}H_{32}N_2O_6SiNa$, 531.1922; found, 531.1917.

7.2 Oligonucleotide synthesis and characterisation

7.2.1 Solid-phase oligonucleotide synthesis

Automated solid-phase oligonucleotide synthesis was carried out on an Applied Biosystems 394 synthesiser. Synthesis was performed on a 1.0 µmol scale using 2'-OMe CPG solid supports (A(Bz)) and U), 2'OMe-phosphoramidites (U, A(Bz), C(Ac), G(dmf)), LNA-phosphoramidites (T, A(Bz), ^{Me}C(Bz), G(dmf)), and 5'-biotin phosphoramidite, all purchased from *Biosearch Technologies* (acquired *LINK*). Standard reagents such as Cap A, Cap B, BTT, oxidiser, and ACN were purchased from *Sigma-Aldrich*. EDITH was purchased from *Biosearch Technologies*. CH_2Cl_2 (alcohol-free) was purchased from *Rathburn Chemicals*.

Solid-phase oligonucleotide synthesis was carried out in the cycles of acid-catalysed detritylation, coupling, capping, and oxidation or sulfurisation outlined in Chapter 1, Section 1.1.3. 2'OMe phosphoramidites were coupled for 5 min and custom phosphoramidites (monomers and dimers), as well as LNA-phosphoramidites, were coupled for an additional 5 min. Sulfurisation was carried out for 4 min using 3-ethoxy-1,2,4-dithiazoline-5-one (EDITH) (*Biosearch Technologies*). For oligonucleotides requiring on-resin amide coupling, the 15 second capping step using acetic anhydride was omitted. Coupling efficiency was estimated using the trityl cation conductivity monitoring system built into the synthesiser.

DNA and RNA complement oligonucleotides (shown in Table B.1) were purchased from *Integrated DNA Technologies*.

7.2.2 Optimised on-resin amide coupling conditions

On-resin amide coupling via the monomer approach was optimised and discussed as in Chapter 2, Figure 2.7. The custom 5'-MMTr-amino-phosphoramidite was coupled using the conditions outlined above. Detritylation was performed on the synthesiser for 3 min using deblock (3% TCA/CH_2Cl_2) (*Biosearch Technologies*). The column was removed from the synthesiser and a solution containing the 5'-DMTr-LNA-3'-COOH monomer (compound **20-21**) (10 eq.), HATU (10 eq.) and *N*-methylmorpholine (30 eq.) in 400 µL anhydrous DMF, reacted previously for 30 min, was pushed through the column by attaching a 1 mL plastic syringe to each side. The solution was agitated every 10 min for 1 hour by pushing the solution back and forth through the column. The reaction mixture was discarded and the resin was washed with anhydrous ACN (5 mL), anhydrous THF (3 mL), and dried under argon before being transferred back to the synthesiser for solid phase synthesis to continue as described in Chapter 1.

7.2.3 Deprotection conditions

For standard ASOs, the oligonucleotides were deprotected and cleaved from the solid support by heating the resin in concentrated aqueous ammonia in a sealed glass vial at 55 °C for 5 h.

For ASOs containing triester backbones (Chapter 4), the cyanoethyl groups were deprotected first by incubating the resin in 20% diethylamine in toluene for 30 min at rt (using 2 × 1 mL syringes as described previously). The resin was washed with ACN (3 mL) and then treated with a mixture of anhydrous THF and anhydrous ethylene diamine (1:1, v/v) (1 mL) and heated at 55 °C for 15 min. The resin was washed in anhydrous THF (3 mL), anhydrous ACN (5 mL), and the oligos were eluted using EtOH/H_2O (1:1, v/v) (2 mL).

7.2.4 Oligonucleotide purification

Oligonucleotides were purified by reverse phase high-performance liquid chromatography (RP-HPLC) on one of two systems:

1. Gilson system using a Luna 10 µm C8(2) 100 Å pore Phenomenex column (250 × 10 mm)
2. Agilent 1260 Infinity system using a Clarity 5 µm Oligo-RP^{TM} C18 Phenomenex column (250 × 10 mm)

The oligonucleotides were purified using a gradient of ACN in 0.1 M triethylammonium acetate buffer (TEAA) (Buffer A). For oligonucleotides purified with the 5'-DMTr protecting group remaining (DMTr-on), the gradient used was 20-80% Buffer B (60% ACN, 0.1 M TEAA) over 20 min at a flow rate of 4 mL min^{-1} (Gilson) or 30 min at a flow rate of 5 mL min^{-1} (Agilent). For oligonucleotides purified without the 5'-DMTr protecting group (DMTr-off), the gradient used was 20-50% Buffer B (50% ACN, 0.1 M TEAA) over the same flow rates as mentioned previously. Elution was monitored by UV absorption between 260-295 nm.

If the purified oligonucleotide remained DMTr-on post-purification, the 5'-DMTr protecting group was removed by treating the lyophilised oligonucleotide in 80% acetic acid (500 µL) for 30 min, then neutralised with 1M TEAA stock (500 µL) before the entire 1 mL solution was submitted to salt exchange.

7.2.5 Salt exchange

Pure, fully deprotected oligonucleotides were exchanged from the Et_3N^+ salt to the Na^+ salt using the RP-HPLC Gilson system. The column mentioned above was equilibrated with 0.1 M NaOAc (Buffer A). The oligonucleotide was injected (in volumes < 1 mL) and this method was run: 10 min of 0.1 M NaOAc (Buffer A), 10 min of H_2O (Buffer B), 10 min of 50% ACN (Buffer A/C). Elution was monitored by UV absorption between 260-295 nm.

7.2.6 Oligonucleotide characterisation by mass spectrometry

Oligonucleotide LC-MS was recorded on a Waters Xevo G2-QTOF which is coupled to an Acquity UHPLC system using a C18 column, 130Å (1.7 µm, 2.1 mm × 50 mm). Buffer

A is composed of 8.15 mM Et_3N and 200 mM hexafluoroisopropanol (HFIP) made in 5% MeOH/H_2O. Buffer B is composed of 20% Buffer A in MeOH. A gradient of 0-70% Buffer B at flow rate of 0.2 mL min^{-1} was used over 8 min to produce the HPLC chromatograms. Raw data were processed and de-convoluted using the Waters MassLynx v4.1 software package.

Oligonucleotide concentration was determined by three independent replicates using the UV-Vis spectrophotometer Nanodrop^{TM} 2000 (*Thermo Fisher Scientific*).

7.3 Cell culture

7.3.1 HeLa pLuc/705 cell culture

HeLa pLuc/705 cells were obtained from our collaborators Prof. Matthew Wood (Institute of Developmental and Regenerative Medicine, University of Oxford). They were cultured in growth media (DMEM with GlutaMAX-1 (*Gibco*, cat. no. 10566016) supplemented with 10% (v/v) FBS (*Gibco*, cat. no. A5256701), 1% (v/v) Antibiotic-Antimycotic (*Gibco*, cat. no. 15240062) and incubated at 37 °C, 5% CO_2.

7.3.2 Treatment of HeLa pLuc/705 cells with ASO

For transfected activity, the HeLa pLuc/705 cells were seeded on a 96-well plate at 10,000 cells/well and incubated for 24 h.

Transfection with Lipofectamine 2000: Lipofectamine 2000 (*Invitrogen*, cat no. 11668019) was diluted in OptiMEM (*Gibco*, cat. no. 31985062) (1 µL/100 µL). ASO (20 µM stock) was diluted in OptiMEM (1 µL/100 µL). The lipofectamine and ASO were added together (1:1, v:v) and the mixture was incubated at rt for 10 min (final ASO concentration = 1000 nM). Each well was treated with ASO:lipofectamine complex (10 µL) in OptiMEM (90 µL) (final ASO concentration = 100 nM) for 4 h. The media was removed and replaced with growth media and incubated for a further 20 h.

Transfection with Lipofectamine 3000: Lipofectamine 3000 (*Invitrogen*, cat. No. L3000001)

was diluted in OptiMEM (*Gibco*, cat. no. 31985062) (1 µL/100 µL) and incubated for 5 min at rt. ASO (20 µM stock) was diluted in OptiMEM to give concentrations ranging from 125-2000 nM. The lipofectamine and ASO stocks were added together (1:1, v:v) and the mixture was incubated at rt for 10 min (final ASO concentrations = 62.5-1000 nM). Each well was treated with the ASO:lipofectamine complex (10 µL) diluted in 100 µL growth media (final concentration = 6.25-100 nM) and the cells were incubated for a further 24 h.

For gymnotic activity, the HeLa pLuc/705 cells were seeded on a 96-well plate at 1,000 cells/well and incubated for 24 h. Lyophilised ASOs (lypohilised at 20 µM/mL) were reconstituted in 1 mL of OptiMEM (final ASO concentration = 20 µM) which was further diluted in OptiMEM to concentrations required (2.5-20 µM). The cells were treated with the OptiMEM media containing ASOs (100 µL) and incubated for 96 h.

7.3.3 H2k *mdx* cell culture

H2k *mdx* cells were obtained from our collaborators Prof. Matthew Wood (Institute of Developmental and Regenerative Medicine, University of Oxford). They were cultured in growth media (DMEM with GlutaMAX-1 supplemented with 20% (v/v) FBS, 1.5% (v/v) chicken embryo extract (*Sera Laboratories*, cat. no. CE-650-J), and 1:10,000 mouse interferon-γ (*Millipore*, cat. no. IF005)) in incubators at 33 °C, 10% CO_2 on flasks coated with Matrigel (*Scientific Laboratory Supplies*, cat. no. 354234).

7.3.4 Treatment of H2k *mdx* cells with ASO

Cells were seeded on a 24-well plate coated with Matrigel at 20,000 cells/well. After 48 h (or at 60-75% confluency), the media was changed to differentiation media (DMEM with GlutaMAX-1 supplemented with 5% (v/v) horse serum inactivated (*Thermo Fisher Scientific*, cat. no. 26050–088), 1% (v/v) Antibiotic-Antimycotic (*Gibco*)) and incubated at 37 °C, 5% CO_2 for 24-48 h, or until the myoblasts have formed striated myotubes.

RNAiMax (*Thermo Fisher Scientific*, cat. no. 13778075) transfection reagent was diluted in OptiMem (6 µL/100 µL OptiMEM). ASO (20 µL stock) was diluted in OptiMEM to give

final concentration of 2000 nM. The RNAiMax and ASO were added together (1:1, v:v) and the mixture was incubated for 5 min at rt (final concentration = 1000 nM). Each well was treated with 50 µL of ASO-lipofectamine complex diluted in 500 µL of differentiation media (final concentration = 100 nM) and incubated for 24 h. The media was changed to fresh differentiation media and incubated for a further 72 h.

For gymnotic activity, lyophilised ASOs (lyophilised at 20 µM/mL) were reconstituted in 2 mL differentiation media (final ASO concentration = 10 µM). The cells were treated with the differentiation media containing ASOs (500 µL per well) and incubated for 96 h.

7.3.5 Luciferase activity assay

HeLa pLuc/705 cells were seeded and cultured according to the treatment with ASO as described above. Once the treatment course was complete, the media was removed and the cells were washed with PBS (*Gibco*, cat. no. 10010023) (200 µL) and treated with GloLysis^{TM} buffer (*Promega*, cat. no. E2661) (100 µL). The cells were shaken on an orbital shaker for 10 min at rt to ensure full lysis. To quantify protein, lysate (25 µL) was used in a BCA protein quantification kit (*Pierce*, cat. no. 23227) using Glo-Lysis^{TM} buffer as a standard and following the manufacturer's instructions. To quantify luciferase production, lysate (50 µL) was added to Bright-Glo^{TM} luciferase reagent (*Promega*, cat. no. E2610) (50 µL) in white 96-well plates and luminescence was immediately measured on a CLARIOstar microplate reader (*BMG Labtech*) and analysed with CLARIOstar software version 5.21.R2.

The luminescence of each well was divided by the total protein quantity of that well and this value was normalised to the average of the untreated wells to give "fold increase over untreated" activity. The final activity was given as the average and propogated error of each biological replicates' mean and standard deviation. A 2-way mixed ANOVA test was used to determine statistical significance within a concentration (using the 2'OMe/PS ASO as the control).

7.3.6 WST-1 cytotoxicity assay

HeLa pLuc/705 cells were seeded and cultured according to the treatment with ASO as described above. Once the treatment course was complete, WST-1 reagent (*Roche*, cat. no. 5015944001) was added to each well and incubated for 30 min. The cells were shaken on an orbital shaker for 1 min and the absorbance at 440 nm was measured using a CLARIOstar microplate reader (*BMG Labtech*) and analysed with CLARIOstar software version 5.21.R2. Untreated cells were used as a 100% viability reference.

7.3.7 Exon skipping and (q)PCR

H2k *mdx* cells were seeded and cultured according to the treatment with ASO as described above. The cells were washed with PBS (500 µL) and RNA was extracted using an RNease Mini Kit (*Qiagen*, cat. no. 74106) following the manufacturer's instructions. RNA quantity was measured using a Nanodrop™ 2000/2000c spectrophotometer (Thermo Fisher Scientific). The RNA was reverse-transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (*Applied Biosystems*™, cat. no. 4368814) following the manufacturer's instructions. cDNA quantity was measured using a Nanodrop™ 2000/2000c spectrophotometer and each sample was diluted to ∼50-60 ng/µL. 400 ng cDNA was amplified using AmpliTaq Gold™ 360 Master Mix (*Applied Biosystems*™, cat. no. 4398881) in a 50 µL reaction following the cycle conditions: 95 ° for 10 min, followed by 30 cycles of 95 °C for 30 sec, 58 °C for 1 min, 72 °C for 2 min, and a final extension of 10 min at 72 °C. The forward outer primer (covering exon20) (5'-CAGAATTCTGCCAATTGCTGAG-3') and reverse outer primer (covering exon 26) (5'-TTCTTCAGCTTGTGTCATCC-3') were purchased from *Integrated DNA Technologies*. 2 µL of this reaction was amplified again using AmpliTaq Gold™ 360 Master Mix in a 50 µL reaction following the cycle conditions: 95 ° for 10 min, followed by 22 cycles of 95 °C for 30 sec, 58 °C for 1 min, 72 °C for 2 min, and a final extension of 10 min at 72 °C. The forward inner primer (covering exon20) (5'-CCCAGTCTACCACCCTATCAGAGC-3') and reverse inner primer (covering exon 26) (5'-CCTGCCTTTAAGGCTTCCTT-3') were purchased from *Integrated DNA Technologies*.

1 µL of this reaction was separated by 2% agarose gel (using SYBR™ Safe DNA Gel Stain (*Invitrogen*, cat. no. S33102)) (100 V for 1.25 h).

7.4 Protein-binding studies and proteomics

7.4.1 Biotin pull-down assay

1. **Cell lysate preparation:** Trypsinised HeLa pLuc/705 cells in growth media were centrifuged at 1,000 g for 5 min at rt to yield a cell pellet. The pellet was washed with ice cold PBS (*Gibco*, cat. no. 10010023) and centrifuged at 1,000 g for 5 min at rt. The cell pellet was lysed using IP lysis buffer (*Pierce*, cat. no. 87787) according to the manufacturer's protocol with the addition of Halt™ protease inhibitor cocktail *(Thermo Fisher Scientific*, cat. no. 78425). The supernatant protein quantity was quantified using a BCA protein quantification kit (*Pierce*, cat. no. 23227) and Nanodrop™ 2000 spectrophotometer (*Thermo Fisher Scientific*).

2. **Streptavidin bead pull-down:** To a low protein binding microcentrifuge tube (*Thermo Fisher Scientific*, cat. no. 90410) was added 25 µL of PureProteome streptavidin magnetic beads (*Millipore*, cat. no. LSKMAGT). The beads were washed (3 × 200 µL) in Binding Buffer (PBS with 0.2% Tween-20 (*Sigma Aldrich*, cat. no. P1379). The beads were incubated with biotinylated-ASO (25 µM) (20 µL of 250 µM stock diluted in 180 µL Binding Buffer) for 30 min at rt (in a shaker at 1000 rpm). The beads were washed (3 × 200 µL) in Binding Buffer.

3. **Incubation and washing:** The beads were incubated with 500 µg cell lysate (prepared as above) for 2 h at 4 °C (in a shaker at 1000 rpm). The beads were washed (10 × 200 µL) in Washing Buffer (PBS with 0.2% Tween-20 with additional 100 mM NaCl). This washing was done by pipetting twice up and down (rather than vortexing). On the 10th wash, the beads, diluted in 200 µL Washing Buffer, were split into 3 low protein binding microcentrifuge tubes (50 µL each) and pulled down.

4. **Competitive elution:** The proteins were eluted by incubating the beads with 50 µL of capture ASO diluted in Binding Buffer (0.33X, 1X, and 3X concentrations were made using 1.67 µL, 5 µL, and 15 µL of 250 µM ASO stock, respectively) for 5 min at rt (in a shaker at 1000 rpm). The beads were pulled down and the supernatant was transferred to a clean low protein binding microcentrifuge tube.

7.4.2 SDS-PAGE gel electrophoresis and staining

The proteins were separated and visualised using the contents of a NuPAGE Bis-Tris Welcome Pack 4-12%, (*Invitrogen*, cat. no. NP032A), which contains Mini Gel Tank, NuPAGE Bis-Tris mini gels (10-well), NuPAGE MES SDS Running Buffer, 20X (cat. no. NP0002), NuPAGE LDS Sample Buffer, 4X (cat. no. NP0007), NuPAGE Sample Reducing Agent, 10X (cat. no. NP0009), and PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa (cat. no. 26619).

Samples were made in a total volume of 20 µL, containing 10 µL of protein supernatant (eluted as described above); samples were reduced and loaded, and the gel was run according to the manufacturer's protocol (35 min at 200 V constant). The gels were stained using the ProteoSilver™ Silver Stain Kit (*Sigma-Aldrich*, cat. no. PROTSIL1) following the manufacturer's instructions.

7.4.3 Label-free quantification

The biotin pull-down assay was conducted as listed previously, with the following changes:

1. **Washing:** The beads were washed (7 × 200 µL) in washing buffer (PBS with 0.2% Tween-20 with additional 100 nM NaCl). The beads were then washed with (3 × 200 µL) in clean PBS. On the 10th wash, the beads, diluted in 200 µL clean PBS, were split into 3 low protein-binding microcentrifuge tubes (50 µL each) and pulled down.

2. **Competitive elution:** The proteins were eluted by incubating the beads with 50 µL of capture ASO diluted in PBS-only (1X concentration made using 5 µL of 250 µM

ASO stock) for 5 min at rt (in a shaker at 1000 rpm). The beads were pulled down and the supernatant was transferred to a clean low protein-binding microcentrifuge tube.

These samples were transferred on ice to the Advanced Proteomics Facility in the Department of Biochemistry where they were digested with benzonase, trypsin, and samples were run on an Orbitrap Elite.[b]

[b]I thank Dr. Marjorie Fournier and all members of the Advanced Proteomics Facility for conducting LFQ on these samples.

Bibliography

1. Crick, F. H. On protein synthesis. *Symp Soc Exp Biol* **12,** 138–63 (1958).
2. Koonin, E. V. Does the central dogma still stand? *Biol Direct* **7,** 27 (2012).
3. Brenner, S., Jacob, F. & Meselson, M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190,** 576–581 (1961).
4. Alberts, B. *et al.* in *Molecular Biology of the Cell. From DNA to RNA* 4th Edition (Garland Science, New York, 2002).
5. Berget, S. M., Moore, C. & Sharp, P. A. Spliced segments at the 5 terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences* **74,** 3171–3175 (1977).
6. Wang, Y. *et al.* Mechanism of alternative splicing and its regulation. *Biomed Rep* **3,** 152–158 (2015).
7. Chen, L., Bush, S. J., Tovar-Corona, J. M., Castillo-Morales, A. & Urrutia, A. O. Correcting for Differential Transcript Coverage Reveals a Strong Relationship between Alternative Splicing and Organism Complexity. *Molecular Biology and Evolution* **31,** 1402–1413 (2014).
8. Gilbert, W. Why genes in pieces? *Nature* **271,** 501.
9. Montes, M., Sanford, B. L., Comiskey, D. F. & Chandler, D. S. RNA Splicing and Disease: Animal Models to Therapies. *Trends Genet* **35,** 68–87 (2019).

10. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PLOS Biology* **15,** e2003243 (2017).
11. Cobb, M. & Comfort, N. What Rosalind Franklin truly contributed to the discovery of DNA's structure. *Nature* **616,** 657–660 (2023).
12. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171,** 737–738 (1953).
13. Alberts, B *et al. The Structure and Function of DNA. Molecular Biology of the Cell* Book Section. 2002.
14. Alberts, B. *et al.* in *Molecular Biology of the Cell* 4th Edition (Garland Science, New York, 2002).
15. Chargaff, E. Structure and function of nucleic acids as cell constituents. *Fed Proc* **10,** 654–9 (1951).
16. Ussery, D. W. DNA Structure: A-, B-and Z-DNA Helix Families. *Encyclopedia of Life Sciences* **1,** e003122 (2002).
17. Ho, P. & Carter, M. in *DNA Replication - Current Advances. DNA Structure: Alphabet Soup for the Cellular Soul* (IntechOpen, Aug. 2011).
18. Leslie, E. & Grover, N. in *Fundamentals of RNA Structure and Function. RNA: Composition and Base Pairing* 1–19 (Springer International Publishing, 2022).
19. Wing, R. *et al.* Crystal structure analysis of a complete turn of B-DNA. *Nature* **287,** 755–8 (1980).
20. Deleavey, G. & Damha, M. J. Designing Chemically Modified Oligonucleotides for Targeted Gene Silencing. *Chemistry Biology* **19,** 937–54 (2012).
21. Wang, A. H. *et al.* Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **282,** 680–6 (1979).
22. Rohs, R. *et al.* Origins of specificity in protein-DNA recognition. *Annu Rev Biochem* **79,** 233–69 (2010).
23. Timsit, Y. DNA structure and polymerase fidelity. *Journal of Molecular Biology* **293,** 835–853 (1999).

24. Michelson, A. M. & Todd, A. R. Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3':5'-internucleotidic linkage. *Journal of the Chemical Society (Resumed),* 2632–2638 (1955).
25. Smith, M, Rammler, D., Goldberg, I. & Khorana, H. Studies on Polynucleotides. XIV.[1] Specific Synthesis of the $C_{3"}$-$C_{5"}$ Interribonucleotide Linkage. Syntheses of Uridylyl-(3"→5")-Uridine and Uridylyl-(3"→ 5")-Adenosine[2]. *Journal of the American Chemical Society* **84,** 430–440 (1962).
26. Merrifield, R. B. Solid phase peptide synthesis: The synthesis of a tetrapeptide. *Journal of the American Chemical Society* **85,** 2149–2154 (1963).
27. Froehler, B. C., Ng, P. G. & Matteucci, M. D. Synthesis of DNA via deoxynucleoside H-phosphonate intermediates. *Nucleic Acids Res* **14,** 5399–407 (1986).
28. Letsinger, R. L. & Mahadevan, V. Stepwise Synthesis of Oligodeoxyribonucleotides on an Insoluble Polymer Support[1,2]. *Journal of the American Chemical Society* **88,** 5319–5324 (1966).
29. Caruthers, M. H. *et al.* in *Methods in Enzymology* 287–313 (Academic Press, 1987).
30. Matteucci, M. D. & Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *Journal of the American Chemical Society* **103,** 3185–3191 (1981).
31. Caruthers, M. H. The chemical synthesis of DNA/RNA: our gift to science. *Journal of Biological Chemistry* **288,** 1420–1427 (2013).
32. Hoose, A., Vellacott, R., Storch, M., Freemont, P. S. & Ryadnov, M. G. DNA synthesis technologies to close the gene writing gap. *Nature Reviews Chemistry* **7,** 144–161 (2023).
33. in. *Nucleic Acids Book. Solid-phase Oligonucleotide Synthesis* (ATD Bio).
34. Link Technologies. *Guidebook for the Synthesis of Oligonucleotides* (2015).
35. LGC Biosearch Technologies. *Oligonucleotide synthesis basics: Key takeaways for post-synthesis processing* https://blog.biosearchtech.com/oligonucleotide-synthesis-basics-key-takeaways-for-post-synthesis-processing.

36. Fornstedt, T. & Enmark, M. Separation of therapeutic oligonucleotides using ion-pair reversed-phase chromatography based on fundamental separation science. *Journal of Chromatography Open* **3,** 100079 (2023).
37. Glen Research. *Preparing oligonucleotides for antisense experiments* https://www.glenresearch.com/reports/gr10-12.
38. Liu, A., Cheng, M., Zhou, Y. & Deng, P. Bioanalysis of Oligonucleotide by LC-MS: Effects of Ion Pairing Regents and Recent Advances in Ion-Pairing-Free Analytical Strategies. *Int J Mol Sci* **23** (2022).
39. Basiri, B., van Hattum, H., van Dongen, W. D., Murph, M. M. & Bartlett, M. G. The Role of Fluorinated Alcohols as Mobile Phase Modifiers for LC-MS Analysis of Oligonucleotides. *J Am Soc Mass Spectrom* **28,** 190–199 (2017).
40. Apffel, A., Chakel, J. A., Fischer, S., Lichtenwalter, K. & Hancock, W. S. Analysis of Oligonucleotides by HPLC Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **69,** 1320–1325 (1997).
41. Polo, L. M. & Limbach, P. A. Analysis of oligonucleotides by electrospray ionization mass spectrometry. *Curr Protoc Nucleic Acid Chem* **10** (2001).
42. Pomerantz, S. C., McCloskey, J. A., Tarasow, T. M. & Eaton, B. E. Deconvolution of Combinatorial Oligonucleotide Libraries by Electrospray Ionization Tandem Mass Spectrometry. *Journal of the American Chemical Society* **119,** 3861–3867 (1997).
43. Millipore Sigma Aldrich. *Oligonucleotide Quality Control by Mass Spectrometry* https://www.sigmaaldrich.com/US/en/technical-documents/technical-article/genomics/qpcr/custom-dna-oligos-qc-analysis-by-mass-spectrometry.
44. Alwan, W. & Rault, M. *Fast Determination of Thermal Melt Temperature of Double-Stranded Nucleic Acids by UV-Vis Spectroscopy* https://www.agilent.com/cs/library/applications/Application-tm-ramps-cary-3500-uv-5994-0384en-us-agilent.pdf.
45. Schreiber-Gosche, S. & Edwards, R. A. Thermodynamics of Oligonucleotide Duplex Melting. *Journal of Chemical Education* **86,** 644 (2009).
46. Kypr, J., Kejnovská, I., Renciuk, D. & Vorlícková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res* **37,** 1713–25 (2009).

47. Bishop, G. R. & Chaires, J. B. Characterization of DNA structures by circular dichroism. *Curr Protoc Nucleic Acid Chem* **7** (2003).

48. Pohl, F. M. & Jovin, T. M. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). *Journal of Molecular Biology* **67,** 375–396 (1972).

49. Baker, E. S. & Bowers, M. T. B-DNA Helix Stability in a Solvent-Free Environment. *Journal of the American Society for Mass Spectrometry* **18,** 1188–1195 (2007).

50. Studdert, D. S., Patroni, M. & Davis, R. C. Circular dichroism of DNA: Temperature and salt dependence. *Biopolymers* **11,** 761–779 (1972).

51. Stephenson, M. L. & Zamecnik, P. C. Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proceedings of the National Academy of Sciences* **75,** 285–288 (1978).

52. De Mesmaeker, A., Waldner, A., Fritsch, V. & Wolf, R. M. Improvement of Antisense Oligonucleotides by Synthetic Modifications. *European Journal of Medicinal Chemistry* **30,** 479–484 (1995).

53. Li, Y., Chen, S., Rahimizadeh, K., Zhang, Z. & Veedu, R. N. Inhibition of survivin by 2-O-methyl phosphorothioate-modified steric-blocking antisense oligonucleotides. *RSC Advances* **14,** 13336–13341 (2024).

54. Havens, M. A. & Hastings, M. L. Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res* **44,** 6549–63 (2016).

55. Khorkova, O., Stahl, J., Joji, A., Volmar, C.-H. & Wahlestedt, C. Amplifying gene expression with RNA-targeted therapeutics. *Nature Reviews Drug Discovery* **22,** 539–561 (2023).

56. Egli, M. & Manoharan, M. Chemistry, structure and function of approved oligonucleotide therapeutics. *Nucleic Acids Res* **51,** 2529–2573 (2023).

57. Zhang, L., Vickers, T. A., Sun, H., Liang, X. H. & Crooke, S. T. Binding of phosphorothioate oligonucleotides with RNase H1 can cause conformational changes in the protein and alter the interactions of RNase H1 with other proteins. *Nucleic Acids Res* **49,** 2721–2739 (2021).

58. Wan, W. B. & Seth, P. P. The Medicinal Chemistry of Therapeutic Oligonucleotides. *Journal of Medicinal Chemistry* **59,** 9645–9667 (2016).

59. Dana, H. *et al.* Molecular Mechanisms and Biological Functions of siRNA. *Int J Biomed Sci* **13,** 48–57 (2017).

60. Watts, J. K. & Corey, D. R. Silencing disease genes in the laboratory and the clinic. *J Pathol* **226,** 365–79 (2012).

61. Kang, S.-H., Cho, M.-J. & Kole, R. Up-Regulation of Luciferase Gene Expression with Antisense Oligonucleotides: Implications and Applications in Functional Assay Development. *Biochemistry* **37.** doi: 10.1021/bi980300h, 6235–6239 (1998).

62. Baker, Y. R. *et al.* An LNA-amide modification that enhances the cell uptake and activity of phosphorothioate exon-skipping oligonucleotides. *Nature Communications* **13,** 4036 (2022).

63. Eckstein, F. & Gindl, H. Polyribonucleotides Containing a Phosphorothioate Backbone. *European Journal of Biochemistry* **13,** 558–564 (1970).

64. Eckstein, F. Phosphorothioate Oligodeoxynucleotides: What Is Their Origin and What Is Unique About Them? *Antisense and Nucleic Acid Drug Development* **10,** 117–121 (2000).

65. Crooke, S. T., Vickers, T. A. & Liang, X. H. Phosphorothioate modified oligonucleotide-protein interactions. *Nucleic Acids Res* **48,** 5235–5253 (2020).

66. Liang, X.-h., Sun, H., Shen, W. & Crooke, S. T. Identification and characterization of intracellular proteins that bind oligonucleotides with phosphorothioate linkages. *Nucleic Acids Research* **43,** 2927–2945 (2015).

67. Freier, S. M. & Altmann, K.-H. The ups and downs of nucleic acid duplex stability: Structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Research* **25,** 4429–4443 (1997).

68. Iwamoto, N. *et al.* Control of phosphorothioate stereochemistry substantially increases the efficacy of antisense oligonucleotides. *Nature Biotechnology* **35,** 845–851 (2017).

69. Knouse, K. W. *et al.* Unlocking P(V): Reagents for chiral phosphorothioate synthesis. *Science* **361,** 1234–1238 (2018).

70. Fedoroff, O. Y., Salazar, M. & Reid, B. R. Structure of a DNA : RNA Hybrid Duplex: Why RNase H Does Not Cleave Pure RNA. *Journal of Molecular Biology* **233,** 509–523 (1993).

71. Kurreck, J., Wyszko, E., Gillen, C. & Erdmann, V. A. Design of antisense oligonucleotides stabilized by locked nucleic acids. *Nucleic Acids Research* **30,** 1911–1918 (2002).

72. Hagedorn, P. H. *et al.* Locked nucleic acid: modality, diversity, and drug discovery. *Drug Discovery Today* **23,** 101–114 (2018).

73. Yildirim, I., Kierzek, E., Kierzek, R. & Schatz, G. C. Interplay of LNA and 2'-O-Methyl RNA in the Structure and Thermodynamics of RNA Hybrid Systems: A Molecular Dynamics Study Using the Revised AMBER Force Field and Comparison with Experimental Results. *The Journal of Physical Chemistry B* **118,** 14177–14187 (2014).

74. Hagedorn, P. H. *et al.* Locked nucleic acid: modality, diversity, and drug discovery. *Drug Discovery Today* **23,** 101–114 (2018).

75. Eichert, A. *et al.* The crystal structure of an 'All Locked' nucleic acid duplex. *Nucleic Acids Research* **38,** 6729–6736 (2010).

76. Xu, Y., Villa, A. & Nilsson, L. The free energy of locking a ring: Changing a deoxyribonucleoside to a locked nucleic acid. *Journal of Computational Chemistry* **38,** 1147–1157 (2017).

77. Kaur, H., Babu, B. R. & Maiti, S. Perspectives on Chemistry and Therapeutic Applications of Locked Nucleic Acid (LNA). *Chemical Reviews* **107,** 4672–4697 (2007).

78. Straarup, E. M. *et al.* Short locked nucleic acid antisense oligonucleotides potently reduce apolipoprotein B mRNA and serum cholesterol in mice and non-human primates. *Nucleic Acids Res* **38,** 7100–11 (2010).

79. You, Y., Moreira, B. G., Behlke, M. A. & Owczarzy, R. Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Res* **34,** e60 (2006).

80. Swayze, E. E. *et al.* Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals. *Nucleic Acids Research* **35,** 687–700 (Dec. 2006).

81. Pfeiffer, N. *et al.* First-in-human phase I study of ISTH0036, an antisense oligonucleotide selectively targeting transforming growth factor beta 2 (TGF-2), in subjects with open-angle glaucoma undergoing glaucoma filtration surgery. *PLOS ONE* **12,** e0188899 (2017).

82. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38,** 23–38 (2013).

83. Kumar, P. *et al.* High-Affinity RNA Targeting by Oligonucleotides Displaying Aromatic Stacking and Amino Groups in the Major Groove. Comparison of Triazoles and Phenyl Substituents. *The Journal of Organic Chemistry* **79,** 2854–2863 (2014).

84. Burel, S. A. *et al.* Early-Stage Identification and Avoidance of Antisense Oligonucleotides Causing Species-Specific Inflammatory Responses in Human Volunteer Peripheral Blood Mononuclear Cells. *Nucleic Acid Therapeutics* **32,** 457–472 (2022).

85. Kundu, J. *et al.* Synthesis of Phosphorodiamidate Morpholino Oligonucleotides Using Trityl and Fmoc Chemistry in an Automated Oligo Synthesizer. *The Journal of Organic Chemistry* **87,** 9466–9478 (2022).

86. Pradeep, S. P., Malik, S., Slack, F. J. & Bahal, R. Unlocking the potential of chemically modified peptide nucleic acids for RNA-based therapeutics. *RNA* **29,** 434–445 (2023).

87. Brazil, R. Peptide Nucleic Acids Promise New Therapeutics and Gene Editing Tools. *ACS Central Science* **9,** 3–6 (2023).

88. Migawa, M. T. *et al.* Site-specific replacement of phosphorothioate with alkyl phosphonate linkages enhances the therapeutic profile of gapmer ASOs by modulating interactions with cellular proteins. *Nucleic acids research* **47,** 5465–5479 (2019).

89. Rozners, E. & Strömberg, R. Synthesis and Properties of Oligoribonucleotide Analogs Having Formacetal Internucleoside Linkages. *The Journal of Organic Chemistry* **62,** 1846–1850 (1997).

90. Vasquez, G. *et al.* Evaluation of Phosphorus and Non-Phosphorus Neutral Oligonucleotide Backbones for Enhancing Therapeutic Index of Gapmer Antisense Oligonucleotides. *Nucleic Acid Therapeutics* **32,** 40–50 (2022).

91. Patutina, O. A. *et al.* Mesyl phosphoramidate backbone modified antisense oligonucleotides targeting miR-21 with enhanced in vivo therapeutic potency. *Proceedings of the National Academy of Sciences* **117,** 32370–32379 (2020).
92. Michel, T. *et al.* Cationic phosphoramidate α-oligonucleotides efficiently target single-stranded DNA and RNA and inhibit hepatitis C virus IRES-mediated translation. *Nucleic Acids Research* **31,** 5282–5290 (2003).
93. Rajasekaran, T. *et al.* Backbone Hydrocarbon-Constrained Nucleic Acids Modulate Hybridization Kinetics for RNA. *Journal of the American Chemical Society* **144,** 1941–1950 (2022).
94. Collotta, D., Bertocchi, I., Chiapello, E. & Collino, M. Antisense oligonucleotides: a novel Frontier in pharmacological strategy. *Front Pharmacol* **14,** 1304342 (2023).
95. Khvorova, A. siRNAs—A New Class of Medicines. *JAMA* **329,** 2185–2186 (2023).
96. Nie, T. Eplontersen: First Approval. *Drugs* **84,** 473–478 (2024).
97. Roberts, T. C., Langer, R. & Wood, M. J. A. Advances in oligonucleotide drug delivery. *Nature Reviews Drug Discovery* **19,** 673–694 (2020).
98. Migliorati, J. M. *et al.* Absorption, Distribution, Metabolism, and Excretion of US Food and Drug Administration–Approved Antisense Oligonucleotide Drugs. *Drug Metabolism and Disposition* **50,** 888 (2022).
99. Haque, U. S. & Yokota, T. Enhancing Antisense Oligonucleotide-Based Therapeutic Delivery with DG9, a Versatile Cell-Penetrating Peptide. *Cells* **12** (2023).
100. Alhamadani, F. *et al.* Adverse Drug Reactions and Toxicity of the Food and Drug Administration–Approved Antisense Oligonucleotide Drugs. *Drug Metabolism and Disposition* **50,** 879 (2022).
101. Shen, W. *et al.* Chemical modification of PS-ASO therapeutics reduces cellular protein-binding and improves the therapeutic index. *Nature Biotechnology* **37,** 640–650 (2019).
102. Wu, H. *et al.* Nephrotoxicity of marketed antisense oligonucleotide drugs. *Curr Opin Toxicol* **32** (2022).
103. Cui, H., Zhu, X., Li, S., Wang, P. & Fang, J. Liver-Targeted Delivery of Oligonucleotides with N-Acetylgalactosamine Conjugation. *ACS Omega* **6,** 16259–16265 (2021).

104. Tran, P., Weldemichael, T., Liu, Z. & Li, H. Y. Delivery of Oligonucleotides: Efficiency with Lipid Conjugation and Clinical Outcome. *Pharmaceutics* **14** (2022).

105. Hou, X., Zaks, T., Langer, R. & Dong, Y. Lipid nanoparticles for mRNA delivery. *Nature Reviews Materials* **6,** 1078–1094 (2021).

106. Southey, M. W. Y. & Brunavs, M. Introduction to small molecule drug discovery and preclinical development. *Frontiers in Drug Discovery* **3** (2023).

107. Lombardino, J. G. & Lowe, J. A. The role of the medicinal chemist in drug discovery — then and now. *Nature Reviews Drug Discovery* **3,** 853–862 (2004).

108. Idziak, I., Just, G., Damha, M. J. & Giannaris, P. A. Synthesis and hybridization properties of amide-linked thymidine dimers incorporated into oligodeoxynucleotides. *Tetrahedron Letters* **34,** 5417–5420 (1993).

109. De Mesmaeker, A. *et al.* Amides as a New Type of Backbone Modification in Oligonucleotides. *Angewandte Chemie International Edition in English* **33,** 226–229 (1994).

110. De Mesmaeker, A. *et al.* Amide Backbones with Conformationally Restricted Furanose Rings: Highly Improved Affinity of the Modified Oligonucleotides for Their RNA Complements. *Angewandte Chemie International Edition in English* **35,** 2790–2794 (1996).

111. Pallan, P. S., von Matt, P., Wilds, C. J., Altmann, K.-H. & Egli, M. RNA-Binding Affinities and Crystal Structure of Oligonucleotides Containing Five-Atom Amide-Based Backbone Structures. *Biochemistry* **45,** 8048–8057 (2006).

112. Lebreton, J., Waldner, A., Lesueur, C. & De Mesmaeker, A. Antisense Oligonucleotides with Alternating Phosphodiester. *Synlett* **1994,** 137–140 (1994).

113. Lebreton, J., Waldner, A., Fritsch, V., Wolf, R. M. & De Mesmaeker, A. Comparison of two amides as backbone replacement of the phosphodiester linkage in oligodeoxynucleotides. *Tetrahedron Letters* **35,** 5225–5228 (1994).

114. Mesmaeker, A. D. *et al.* Amides as Substitute for the Phosphodiester Linkage in Antisense Oligonucleotides. *Synlett,* 733–736 (1993).

115. Lebreton, J. *et al.* Synthesis of thymidine dimer derivatives containing an amide linkage and their incorporation into oligodeoxyribonucleotides. *Tetrahedron Letters* **34,** 6383–6386 (1993).

116. Mahesh, S., Tang, K. C. & Raj, M. Amide Bond Activation of Biological Molecules. *Molecules* **23** (2018).

117. Mutisya, D. *et al.* Amides are excellent mimics of phosphate internucleoside linkages and are well tolerated in short interfering RNAs. *Nucleic Acids Res* **42,** 6542–51 (2014).

118. Epple, S., Thorpe, C., Baker, Y. R., El-Sagheer, A. H. & Brown, T. Consecutive 5- and 3-amide linkages stabilise antisense oligonucleotides and elicit an efficient RNase H response. *Chemical Communications* **56,** 5496–5499 (2020).

119. Richter, M. *et al.* Amide Modifications in the Seed Region of the Guide Strand Improve the On-Target Specificity of Short Interfering RNA. *ACS Chem Biol* **18,** 7–11 (2023).

120. Tanui, P. *et al.* Synthesis, biophysical studies and RNA interference activity of RNA having three consecutive amide linkages. *Org. Biomol. Chem.* **12,** 1207–1210 (2014).

121. Kotikam, V., Viel, A. J. & Rozners, E. Synthesis and Biological Activity of Short Interfering RNAs Having Several Consecutive Amide Internucleoside Linkages. *Chemistry – A European Journal* **26,** 685–690 (2020).

122. Mutisya, D. *et al.* Amide linkages mimic phosphates in RNA interactions with proteins and are well tolerated in the guide strand of short interfering RNAs. *Nucleic Acids Research* **45,** 8142–8155 (2017).

123. Kotikam, V., Gajula, P. K., Coyle, L. & Rozners, E. Amide Internucleoside Linkages Are Well Tolerated in Protospacer Adjacent Motif-Distal Region of CRISPR RNAs. *ACS Chemical Biology* **17,** 509–512 (2022).

124. Lauritsen, A. & Wengel, J. Oligodeoxynucleotides containing amide-linked LNA-type dinucleotides: synthesis and high-affinity nucleic acid hybridization. *Chemical Communications,* 530–531 (2002).

125. Viel, J. A., Pal, C. & Rozners, E. Optimization of Automated Synthesis of Amide-Linked RNA. *ACS Omega* **7,** 20420–20427 (2022).

126. Iwase, R., Toyama, T. & Nishimori, K. Solid-Phase Synthesis of Modified RNAs Containing Amide-Linked Oligoribonucleosides at Their 3-End and Their Application to siRNA. *Nucleosides, Nucleotides Nucleic Acids* **26,** 1451–1454 (2007).

127. Devor, E. J. & Behlke, M. A. *Oligonucleotide Yield, Resuspension, and Storage* https://www.crchudequebec.ulaval.ca/wp-content/uploads/2015/10/Oligonucleotide_Yield_Resuspension_and_Storage.pdf.

128. Kochetkova, V. S., Varizhuk, M. A., Kolganova, A. N., Timofeev, N. E. & Florent'Ev, L. V. Synthesis of 3-deoxy-3-carboxymethylnucleosides, precursors of oligonucleotides with an amide internucleoside bond. *Russian Journal of Bioorganic Chemistry* **35,** 68–74 (2009).

129. Arzel, L. *et al.* Synthesis of ribonucleosidic dimers with an amide linkage from d-xylose. *The Journal of Organic Chemistry* **81,** 10742–10758 (2016).

130. Vorbrüggen, H. & Bennua, B. New simplified nucleoside synthesis. *Tetrahedron Letters* **19,** 1339–1342 (1978).

131. Choi, W. J. *et al.* Fluorocyclopentenyl-cytosine with Broad Spectrum and Potent Antitumor Activity. *Journal of Medicinal Chemistry* **55,** 4521–4525 (2012).

132. Kotikam, V. & Rozners, E. Concurrent Hydrogenation of Three Functional Groups Enables Synthesis of C3-Homologated Nucleoside Amino Acids. *Organic Letters* **19,** 4122–4125 (2017).

133. Glen Research. *Chemical Modifications of the 5'-Terminus of Oligonucleotides* https://www.glenresearch.com/reports/gr1-21.

134. Uhlmann, E., Peyman, A., Breipohl, G. & Will, W. D. PNA: Synthetic Polyamide Nucleic Acids with Unusual Binding Properties. *Angewandte Chemie International Edition* **37,** 2796–2823 (1998).

135. Carpino, L. A. *et al.* Comparison of the Effects of 5- and 6-HOAt on Model Peptide Coupling Reactions Relative to the Cases for the 4- and 7-Isomers. *Organic Letters* **2,** 2253–2256 (2000).

136. Schildkraut, C. & Lifson, S. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* **3,** 195–208 (1965).

137. K. Singh, S., A. Koshkin, A., Wengel, J. & Nielsen, P. LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition. *Chemical Communications,* 455–456 (1998).

138. Evich, M., Spring-Connell, A. M. & Germann, M. W. Impact of modified ribose sugars on nucleic acid conformation and function. *Heterocyclic Communications* **23,** 155–165 (2017).

139. Bondensgaard, K. *et al.* Structural Studies of LNA:RNA Duplexes by NMR: Conformations and Implications for RNase H Activity. *Chemistry – A European Journal* **6,** 2687–2695 (2000).

140. Wang, T., Larcher, L., Ma, L. & Veedu, R. Systematic Screening of Commonly Used Commercial Transfection Reagents towards Efficient Transfection of Single-Stranded Oligonucleotides. *Molecules* **23,** 2564 (Oct. 2018).

141. Kita, R., Osawa, T. & Obika, S. Conjugation of oligonucleotides with activated carbamate reagents prepared by the Ugi reaction for oligonucleotide library synthesis. *RSC Chem Biol* **3,** 728–738 (2022).

142. Prhavc, M., Lesnik, E. A., Mohan, V. & Manoharan, M. 2'-O-Carbamate-containing oligonucleotides: synthesis and properties. *Tetrahedron Letters* **42,** 8777–8780 (2001).

143. Waldner, A., De Mesmaeker, A. & Lebreton, J. Synthesis of Oligodeoxyribonucleotides containing dimers with carbamate moieties as replacement of the natural phosphodiester linkage. *Bioorganic Medicinal Chemistry Letters* **4,** 405–408 (1994).

144. Ghosh, A. K. & Brindisi, M. Organic carbamates in drug design and medicinal chemistry. *J Med Chem* **58,** 2895–940 (2015).

145. Thorpe, C., Epple, S., Woods, B., El-Sagheer, A. H. & Brown, T. Synthesis and biophysical properties of carbamate-locked nucleic acid (LNA) oligonucleotides with potential antisense applications. *Organic Biomolecular Chemistry* **17,** 5341–5348 (2019).

146. Datta, D. *et al.* Aminooxy Click Chemistry as a Tool for Bis-homo and Bis-hetero Ligand Conjugation to Nucleic Acids. *Organic Letters* **24,** 4496–4501 (2022).

147. Pifferi, C. *et al.* Aminooxylated Carbohydrates: Synthesis and Applications. *Chemical Reviews* **117,** 9839–9873 (2017).

148. Winans, K. A. & Bertozzi, C. R. An inhibitor of the human UDP-GlcNAc 4-epimerase identified from a uridine-based library: a strategy to inhibit O-linked glycosylation. *Chem Biol* **9,** 113–29 (2002).

149. Song, Z., He, X.-P., Chen, G.-R. & Xie, J. 6-O-Amino-2-O-carboxymethyl Glucopyranoside as Novel Glycoaminoxy Acid Building Block for the Construction of Oligosaccharide Mimetics. *Synthesis* **2011,** 2761–2766 (2011).

150. Li, X., Wu, Y.-D. & Yang, D. α-Aminoxy Acids: New Possibilities from Foldamers to Anion Receptors and Channels. *Accounts of Chemical Research* **41,** 1428–1438 (2008).

151. Gong, Y., Peyrat, S., Sun, H. & Xie, J. Synthesis of nucleoside aminooxy acids. *Tetrahedron* **67,** 7114–7120 (2011).

152. Peyrat, S. & Xie, J. Synthesis of Thymidine Dimers from 5'-O-Aminothymidine. *Synthesis* **44,** 1718–1724 (2012).

153. Vasseur, J. J., Debart, F., Sanghvi, Y. S. & Cook, P. D. Oligonucleosides: synthesis of a novel methylhydroxylamine-linked nucleoside dimer and its incorporation into antisense sequences. *Journal of the American Chemical Society* **114,** 4006–4007 (1992).

154. Morvan, F., Sanghvi, Y. S., Perbost, M., Vasseur, J.-J. & Bellon, L. Oligonucleotide Mimics for Antisense Therapeutics: Solution Phase and Automated Solid-Support Synthesis of MMI Linked Oligomers. *Journal of the American Chemical Society* **118,** 255–256 (1996).

155. Prakash, T. P. *et al.* Antisense Oligonucleotides Containing Conformationally Constrained 2,4-(N-Methoxy)aminomethylene and 2,4-Aminooxymethylene and 2-O,4-C-Aminomethylene Bridged Nucleoside Analogues Show Improved Potency in Animal Models. *Journal of Medicinal Chemistry* **53,** 1636–1650 (2010).

156. Koshkin, A. A., Fensholdt, J., Pfundheller, H. M. & Lomholt, C. A Simplified and Efficient Route to 2'-O, 4'-C-Methylene-Linked Bicyclic Ribonucleosides (Locked Nucleic Acid). *The Journal of Organic Chemistry* **66,** 8504–8512 (2001).

157. Shin, I. *et al.* Synthesis of Optically Active Phthaloyl d-Aminooxy Acids from l-Amino Acids or l-Hydroxy Acids as Building Blocks for the Preparation of Aminooxy Peptides. *The Journal of Organic Chemistry* **65,** 7667–7675 (2000).

158. Filira, F. *et al.* Opioid peptides: synthesis and biological properties of [(N-glucosyl,N-methoxy)-,-diamino-(S)-butanoyl]4-deltorphin-1-neoglycopeptide and related analogues. *Organic Biomolecular Chemistry* **1,** 3059–3063 (2003).

159. Fernando, R., Shirley, J. M., Torres, E., Jacobs, H. K. & Gopalan, A. S. Preparation of bifunctional isocyanate hydroxamate linkers: Synthesis of carbamate and urea tethered polyhydroxamic acid chelators. *Tetrahedron Lett* **53,** 6367–6371 (2012).

160. Chen, N. & Xie, J. Synthesis of O-Amino Sugars and Nucleosides. *Molecules* **23** (2018).

161. Noel, O. & Xie, J. Synthesis of Nucleo Aminooxy Acid Derivatives. *Synthesis* **45,** 134–140 (2013).

162. Lloyd, D., Bylsma, M., Bright, D. K., Chen, X. & Bennett, C. S. Mild Method for 2-Naphthylmethyl Ether Protecting Group Removal Using a Combination of 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) and -Pinene. *The Journal of Organic Chemistry* **82,** 3926–3934 (2017).

163. Seth, P. P. & Swayze, E. E. *Oligomeric compounds comprising neutrally linked terminal bicyclic nucleosides* Patent number WO2009124238A1. Oct. 2009.

164. Sharma, V. K. *et al.* Synthesis and biological properties of triazole-linked locked nucleic acid. *Chemical Communications* **53,** 8906–8909 (2017).

165. Miller, P. S., Fang, K. N., Kondo, N. S. & Ts'o, P. O. P. Conformation and interaction of dinucleoside mono- and diphosphates. V. Syntheses and properties of adenine and thymine nucleoside alkyl phosphotriesters, the neutral analogs of dinucleoside monophosphates. *Journal of the American Chemical Society* **93,** 6657–6665 (1971).

166. Moody, H. M. *et al.* Regiospecific inhibition of DNA duplication by antisense phosphate-methylated oligodeoxynucleotides. *Nucleic Acids Research* **17,** 4769–4782 (1989).

167. Miller, P. S., Barrett, J. C. & Ts'o, P. O. P. Alkyl phosphotriesters of dinucleotides and oligonucleotides. 4. Synthesis of oligodeoxyribonucleotide ethyl phosphotriesters and their specific complex formation with transfer ribonucleic acid. *Biochemistry* **13,** 4887–4896 (1974).

168. Buck, H. M. *et al.* Phosphate-Methylated DNA Aimed at HIV-1 RNA Loops and Integrated DNA Inhibits Viral Infectivity. *Science* **248,** 208–212 (1990).

169. Buck, H. The Biochemical Impact by Covalent Shielding of the Anionic Oxygen of the Phosphate Group in DNA and RNA as Methylated Phosphotriester Linkage on the Inhibition of DNA Duplication and on HIV-1 RNA Viral Infectivity Has Been Seriously Overlooked. *Journal of Biophysical Chemistry* **14,** 59–66 (2023).

170. Monfregola, L. & Caruthers, M. H. Solid-Phase Synthesis, Hybridizing Ability, Uptake, and Nuclease Resistant Profiles of Position-Selective Cationic and Hydrophobic Phosphotriester Oligonucleotides. *The Journal of Organic Chemistry* **80,** 9147–9158 (2015).

171. Iyer, R. P., Yu, D., Jiang, Z. & Agrawal, S. Synthesis, biophysical properties, and stability studies of mixed backbone oligonucleotides containing segments of methylphosphotriester internucleotidic linkages. *Tetrahedron* **52,** 14419–14436 (1996).

172. Meade, B. R. *et al.* Efficient delivery of RNAi prodrugs containing reversible charge-neutralizing phosphotriester backbone modifications. *Nat Biotechnol* **32,** 1256–61 (2014).

173. Tsubaki, K. *et al.* Synthesis and Evaluation of Neutral Phosphate Triester Backbone-Modified siRNAs. *ACS Medicinal Chemistry Letters* **11,** 1457–1462 (2020).

174. Hammill, M. L. *et al.* SiRNAs with Neutral Phosphate Triester Hydrocarbon Tails Exhibit Carrier-Free Gene-Silencing Activity. *ACS Medicinal Chemistry Letters* **13,** 695–700 (2022).

175. Lartia, R., Vallée, C. & Defrancq, E. Post-synthetic transamination at position N4 of cytosine in oligonucleotides assembled with routinely used phosphoramidites. *Organic Biomolecular Chemistry* **18,** 9632–9638 (2020).

176. Hogrefe, R. I., Vaghefi, M. M., Reynolds, M. A., Young, K. M. & Arnold L. J., J. Deprotection of methylphosphonate oligonucleotides using a novel one-pot procedure. *Nucleic Acids Res* **21,** 2031–8 (1993).

177. Johnsson, R. A., Bogojeski, J. J. & Damha, M. J. An evaluation of selective deprotection conditions for the synthesis of RNA on a light labile solid support. *Bioorganic Medicinal Chemistry Letters* **24,** 2146–2149 (2014).

178. Song, Q. *et al.* High yield protection of purine ribonucleosides for phosphoramidite RNA synthesis. *Tetrahedron Letters* **40,** 4153–4156 (1999).

179. Hu, T. *et al.* "One-Pot" Synthesis of Molnupiravir from Cytidine. *Organic Process Research Development* **26,** 358–364 (2022).

180. Iyer, R. P. in *Current Protocols in Nucleic Acid Chemistry. Nucleobase Protection of Deoxyribo- and Ribonucleosides* chap. 2 (2000).

181. Ucal, S. *et al.* Selective acetylation of primary amino groups with phenyl acetate: Simple conversion of polyamines into N,N'-diacetyl derivatives. *Archive for Organic Chemistry* **2015,** 42–49 (2015).

182. Dowdy, S. F. Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnology* **35,** 222–229 (2017).

183. Bailey, J. K., Shen, W., Liang, X. H. & Crooke, S. T. Nucleic acid binding proteins affect the subcellular distribution of phosphorothioate antisense oligonucleotides. *Nucleic Acids Res* **45,** 10649–10671 (2017).

184. Mendell, J. R. & Lloyd-Puryear, M. Report of MDA muscle disease symposium on newborn screening for Duchenne muscular dystrophy. *Muscle Nerve* **48,** 21–6 (2013).

185. Duan, D., Goemans, N., Takeda, S., Mercuri, E. & Aartsma-Rus, A. Duchenne muscular dystrophy. *Nature Reviews Disease Primers* **7,** 13 (2021).

186. Bellayou, H. *et al.* Duchenne and Becker muscular dystrophy: contribution of a molecular and immunohistochemical analysis in diagnosis in Morocco. *J Biomed Biotechnol* **2009,** 325210 (2009).

187. Gao, Q. Q. & McNally, E. M. The Dystrophin Complex: Structure, Function, and Implications for Therapy. *Compr Physiol* **5,** 1223–39 (2015).

188. Nowak, K. J. & Davies, K. E. Duchenne muscular dystrophy and dystrophin: pathogenesis and opportunities for treatment. *EMBO Rep* **5,** 872–6 (2004).

189. Manini, A., Abati, E., Nuredini, A., Corti, S. & Comi, G. P. Adeno-Associated Virus (AAV)-Mediated Gene Therapy for Duchenne Muscular Dystrophy: The Issue of Transgene Persistence. *Front Neurol* **12,** 814174 (2021).

190. Hoy, S. M. Delandistrogene Moxeparvovec: First Approval. *Drugs* **83,** 1323–1329 (2023).

191. Lim, K. R., Maruyama, R. & Yokota, T. Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des Devel Ther* **11,** 533–545 (2017).

192. Matsuo, M. Antisense Oligonucleotide-Mediated Exon-skipping Therapies: Precision Medicine Spreading from Duchenne Muscular Dystrophy. *JMA Journal* **4,** 232–240 (2021).

193. McGreevy, J. W., Hakim, C. H., McIntosh, M. A. & Duan, D. Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Dis Model Mech* **8,** 195–213 (2015).

194. Bulfield, G., Siller, W. G., Wight, P. A. & Moore, K. J. X chromosome-linked muscular dystrophy (mdx) in the mouse. *Proc Natl Acad Sci U S A* **81,** 1189–92 (1984).

195. Jat, P. S. *et al.* Direct derivation of conditionally immortal cell lines from an H-2Kb-tsA58 transgenic mouse. *Proceedings of the National Academy of Sciences* **88,** 5096–5100 (1991).

196. Morgan, J. E. *et al.* Myogenic Cell Lines Derived from Transgenic Mice Carrying a Thermolabile T Antigen: A Model System for the Derivation of Tissue-Specific and Mutation-Specific Cell Lines. *Developmental Biology* **162,** 486–498 (1994).

197. Lehto, T. *et al.* Cellular trafficking determines the exon skipping activity of Pip6a-PMO in mdx skeletal and cardiac muscle cells. *Nucleic Acids Research* **42,** 3207–3217 (2013).

198. Crooke, S. T., Vickers, T. A. & Liang, X.-h. Phosphorothioate modified oligonucleotide–protein interactions. *Nucleic Acids Research* **48,** 5235–5253 (2020).

199. Geary, R. S., Norris, D., Yu, R. & Bennett, C. F. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv Drug Deliv Rev* **87,** 46–51 (2015).

200. Bijsterbosch, M. K. *et al.* Modulation of plasma protein binding and in vivo liver cell uptake of phosphorothioate oligodeoxynucleotides by cholesterol conjugation. *Nucleic Acids Research* **28,** 2717–2725 (2000).

201. Gaus, H. J. *et al.* Characterization of the interactions of chemically-modified therapeutic nucleic acids with plasma proteins using a fluorescence polarization assay. *Nucleic Acids Research* **47,** 1110–1122 (2019).

202. Crooke, S. T., Wang, S., Vickers, T. A., Shen, W. & Liang, X. H. Cellular uptake and trafficking of antisense oligonucleotides. *Nat Biotechnol* **35,** 230–237 (2017).

203. Liang, X.-h. *et al.* Golgi-endosome transport mediated by M6PR facilitates release of antisense oligonucleotides from endosomes. *Nucleic Acids Research* **48,** 1372–1391 (2020).

204. Lorenz, P., Baker, B. F., Bennett, C. F. & Spector, D. L. Phosphorothioate Antisense Oligonucleotides Induce the Formation of Nuclear Bodies. *Molecular Biology of the Cell* **9,** 1007–1023 (1998).

205. Liang, X.-h., Shen, W., Sun, H., Prakash, T. P. & Crooke, S. T. TCP1 complex proteins interact with phosphorothioate oligonucleotides and can co-localize in oligonucleotide-induced nuclear bodies in mammalian cells. *Nucleic Acids Research* **42,** 7819–7832 (2014).

206. Shen, W., Liang, X.-h. & Crooke, S. T. Phosphorothioate oligonucleotides can displace NEAT1 RNA and form nuclear paraspeckle-like structures. *Nucleic Acids Research* **42,** 8648–8662 (2014).

207. Liang, X.-H. *et al.* Hsp90 protein interacts with phosphorothioate oligonucleotides containing hydrophobic 2-modifications and enhances antisense activity. *Nucleic Acids Research* **44,** 3892–3907 (2016).

208. Dean, J. M. & DeLong, R. K. A high-throughput screening assay for the functional delivery of splice-switching oligonucleotides in human melanoma cells. *Methods Mol Biol* **1297,** 187–96 (2015).

209. Novak, J. S. *et al.* Author Correction: Myoblasts and macrophages are required for therapeutic morpholino antisense oligonucleotide delivery to dystrophic muscle. *Nat Commun* **9,** 1256 (2018).

210. Akpulat, U. *et al.* Shorter Phosphorodiamidate Morpholino Splice-Switching Oligonucleotides May Increase Exon-Skipping Efficacy in DMD. *Molecular Therapy - Nucleic Acids* **13,** 534–542 (2018).

211. Aoki, Y. *et al.* Fine Tuning of Phosphorothioate Inclusion in 2'-O-Methyl Oligonucleotides Contributes to Specific Cell Targeting for Splice-Switching Modulation. *Front Physiol* **12,** 689179 (2021).

212. Tri Bao, L., Hornum, M., Sharma, P., Nielsen, P. & Veedu, R. Nucleobase-modified antisense oligonucleotides containing 5-(phenyltriazol)-2-deoxyuridine nucleotides induce exon-skipping: In vitro. *RSC Advances* **7,** 54542–54545 (Nov. 2017).

213. Chelly, J., Kaplan, J.-C., Maire, P., Gautron, S. & Kahn, A. Transcription of the dystrophin gene in human muscle and non-muscle tissues. *Nature* **333,** 858–860 (1988).

214. Sano, T., Vajda, S. & Cantor, C. R. Genetic engineering of streptavidin, a versatile affinity tag. *Journal of Chromatography B: Biomedical Sciences and Applications* **715,** 85–91 (1998).

215. Kakiuchi-Kiyota, S., Whiteley, L. O., Ryan, A. M. & Mathialagan, N. Development of a Method for Profiling Protein Interactions with LNA-Modified Antisense Oligonucleotides Using Protein Microarrays. *Nucleic Acid Ther* **26,** 93–101 (2016).

216. Deng, N., Li, Z., Pan, C. & Duan, H. freeQuant: A Mass Spectrometry Label-Free Quantification Software Tool for Complex Proteome Analysis. *ScientificWorldJournal* **2015,** 137076 (2015).

217. Laskay, A., Lobas, A. A., Srzentić, K., Gorshkov, M. V. & Tsybin, Y. O. Proteome Digestion Specificity Analysis for Rational Design of Extended Bottom-up and Middle-down Proteomics Experiments. *Journal of Proteome Research* **12,** 5558–5569 (2013).

218. Woessmann, J. *et al.* Assessing the Role of Trypsin in Quantitative Plasma and Single-Cell Proteomics toward Clinical Application. *Analytical Chemistry* **95,** 13649–13658 (2023).

219. Frank, A. M. A ranking-based scoring function for peptide-spectrum matches. *J Proteome Res* **8,** 2241–52 (2009).

220. Scott, A. M. *et al.* Generalized precursor prediction boosts identification rates and accuracy in mass spectrometry based proteomics. *Communications Biology* **6,** 628 (2023).

221. Danko, K., Lukasheva, E., Zhukov, V. A., Zgoda, V. & Frolov, A. Detergent-Assisted Protein Digestion—On the Way to Avoid the Key Bottleneck of Shotgun Bottom-Up Proteomics. **23** (2022).

222. Weidner, D. A., Valdez, B. C., Henning, D., Greenberg, S. & Busch, H. Phosphorothioate oligonucleotides bind in a non sequence-specific manner to the nucleolar protein C23/nucleolin. *FEBS Lett* **366,** 146–50 (1995).

223. Nicol, S. M., Causevic, M., Prescott, A. R. & Fuller-Pace, F. V. The nuclear DEAD box RNA helicase p68 interacts with the nucleolar protein fibrillarin and colocalizes specifically in nascent nucleoli during telophase. *Exp Cell Res* **257,** 272–80 (2000).

224. Cheng, X. *et al.* Interactions between single-stranded DNA binding protein and oligonucleotide analogs with different backbone chemistries. *J Mol Recognit* **10,** 101–7 (1997).

225. Piñol-Roma, S. & Dreyfuss, G. hnRNP proteins:Localization and transport between the nucleus and the cytoplasm. *Trends in Cell Biology* **3,** 151–155 (1993).

226. Saarikettu, J. *et al.* The RNA-binding protein Snd1/Tudor-SN regulates hypoxia-responsive gene expression. *FASEB Bioadvances* **5,** 183–198 (2023).

227. Vickers, T. A. & Crooke, S. T. Development of a Quantitative BRET Affinity Assay for Nucleic Acid-Protein Interactions. *PLOS ONE* **11,** e0161930 (2016).

228. Liang, X. H. *et al.* Antisense oligonucleotides targeting translation inhibitory elements in 5' UTRs can selectively increase protein levels. *Nucleic Acids Res* **45,** 9528–9546 (2017).

229. Jeong, S. SR Proteins: Binders, Regulators, and Connectors of RNA. *Mol Cells* **40,** 1–9 (2017).

230. Hock, R., Furusawa, T., Ueda, T. & Bustin, M. HMG chromosomal proteins in development and disease. *Trends Cell Biol* **17,** 72–9 (2007).

231. Xie, Y. *et al.* Structural basis for high-order complex of SARNP and DDX39B to facilitate mRNP assembly. *Cell Rep* **42,** 112988 (2023).

232. Guo, S., Nguyen, L. & Ranum, L. P. W. RAN proteins in neurodegenerative disease: Repeating themes and unifying therapeutic strategies. *Curr Opin Neurobiol* **72,** 160–170 (2022).

233. Alam, M. R. *et al.* Intracellular delivery of an anionic antisense oligonucleotide via receptor-mediated endocytosis. *Nucleic Acids Research* **36,** 2764–2776 (2008).

234. Cooper, G. M. in *The Cell: A Molecular Approach. Translation of mRNA* 2nd Edition (Sinauer Associates, Sunderland, MA, 2000).

235. Weinger, J. S., Parnell, K. M., Dorner, S., Green, R. & Strobel, S. A. Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nature Structural Molecular Biology* **11,** 1101–1106 (2004).

A

Chapter 1 Supplementary Figures

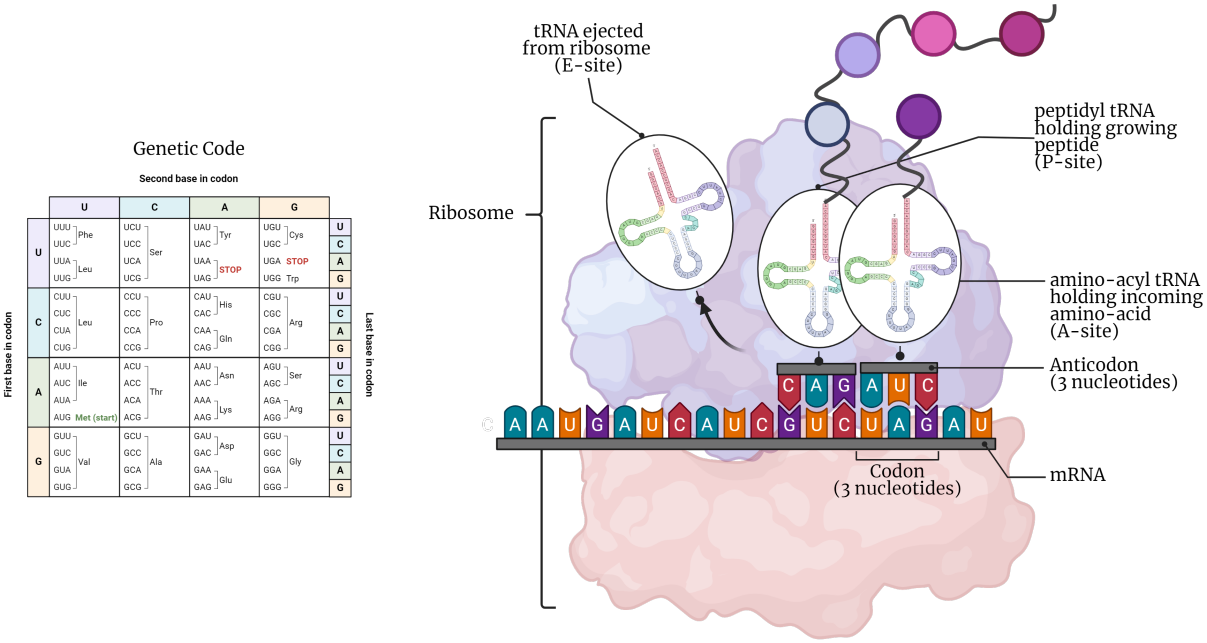


Figure A.1: Chart containing the genetic code (right) and schematic representation of translation conducted by the ribosome (left). Nucleic acids are "read" in groups of three termed "codons"; each codon specifies a particular amino acid (e.g. AUG codes for methionine), recruited to the ribosome by a tRNA molecule bearing the anti-codon at one loop and the correct amino-acid covalently bound at the 3'-end.[234] The ribosome, a structure formed by a collection of proteins and rRNA, holds both the incoming amino-acyl tRNA moiety in place (A-site) as well as the peptidyl-tRNA (P-site) moiety; the peptidyl tRNA is made up of the tRNA from the penultimate coupling and the growing peptide chain. Between these two sites within the ribosome, the formation of the amide bond between the α-amino group on the amino-acyl tRNA and the COO-group of the peptidyl tRNA is catalysed, linking the next amino acid. The ribosome moves to the next codon and the process begins again with the next incoming amino-acyl tRNA moiety.[235]

Gibbs free energy equation:

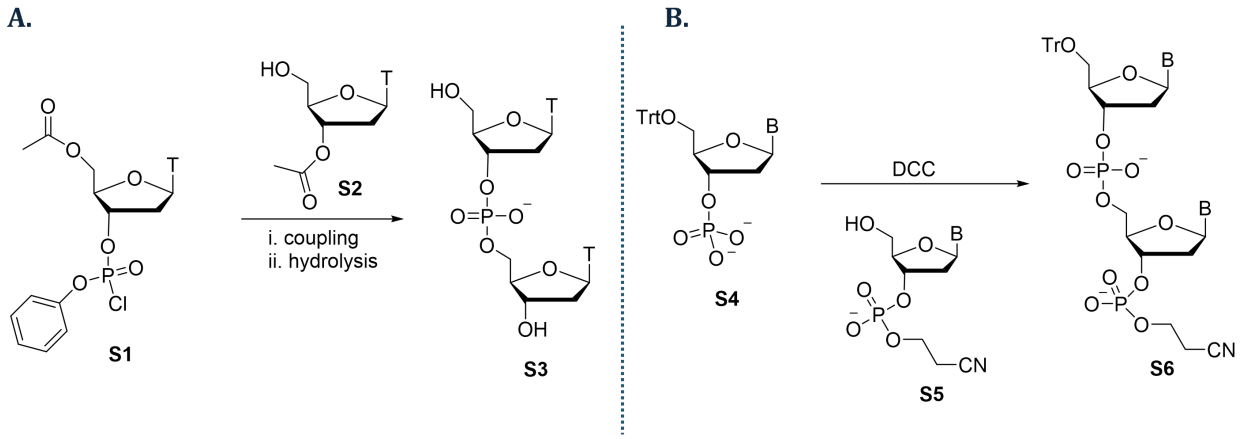
$$\Delta G = \Delta H - T\Delta S \tag{A.1}$$

Gibbs free energy for dissociation equation:

$$\Delta G^{\circ} = -RT\ln K_{\text{eq}} \tag{A.2}$$

Beer-Lambert Law:

$$A = \varepsilon c l \tag{A.3}$$



Scheme A.1: A. The first dinucleotide **S3**, synthesised by Michelson and Todd in 1955.[24] **B.** Synthesis of the dinucleotide **S6** using the phosphodiester approach, introduced by Khorana et al.[25]

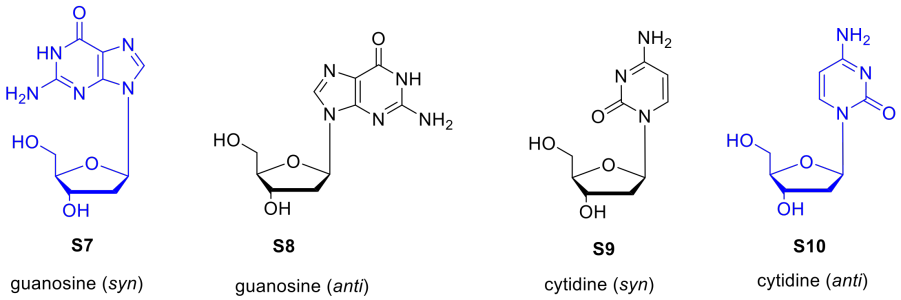


Figure A.2: Guanosine (**S7** and **S8**) and cytidine (**S9** and **S10**) nucleosides in the *syn* and *anti* conformations, respectively. Z-form DNA often forms through alternating *syn*-G and *anti*-C sequences (blue).

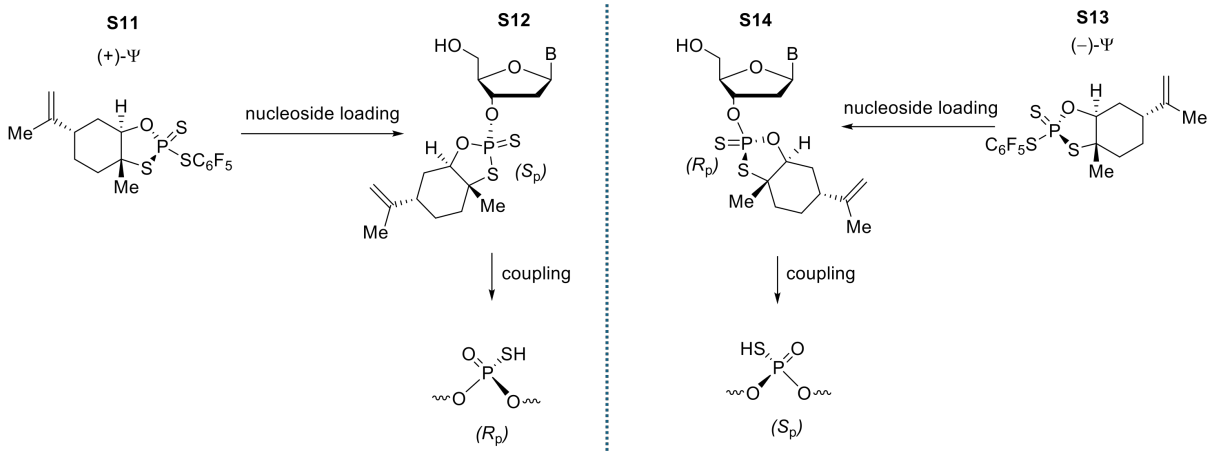


Figure A.3: Chiral P(V) reagents (**S11** and **S13**), developed by members of the Phil Baran group in order to install stereopure phosphorothiaote linkages. Coupling of an S_p monomer **S12** results in an R_p phosphorothioate linkage (left), while coupling of an R_p monomer **S14** results in a S_p linkage (right).[69]

B

Chapter 2 Supplementary Figures

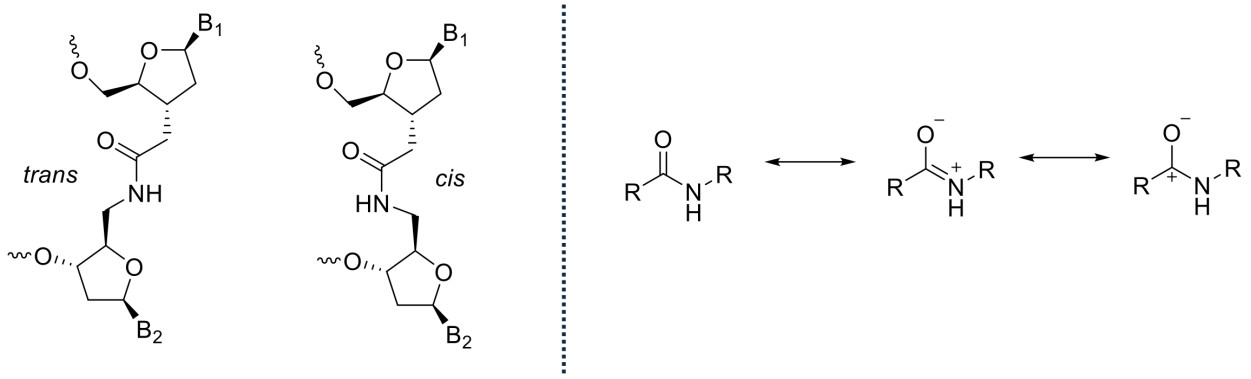
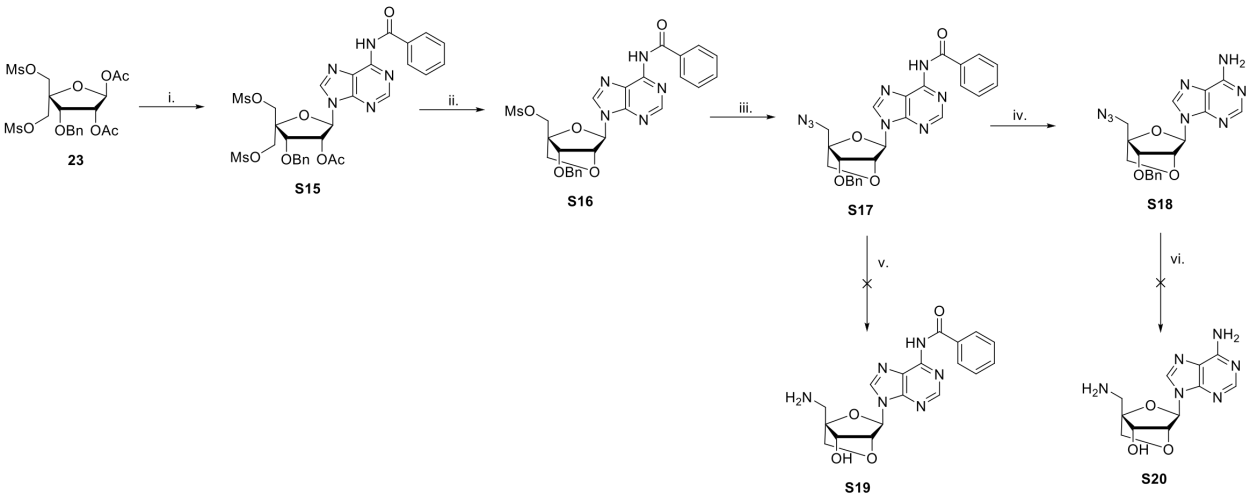


Figure B.1: *trans* and *cis* geometries of amide bonds (left). The AM1 linkage is shown by crystal structure to adopt the more thermodynamically favourable *trans* geometry. The amide linkage is planar due to the resonance of the CO-N bond (right).



Scheme B.1: Scheme of 5'-NH_2 adenine(Bz) monomer syntheses. i. N^6-benzoyladenine, BSA, TMSOTf, DCE, reflux, 24 h, 45%; ii. $LiOH \bullet H_2O$, THF/H_2O (3:2, v:v), rt, 4 h, 78%; iii. NaN_3, DMF, 50 °C, 24 h, 85%; iv. 0.05 M K_2CO_3/MeOH, rt, 3 h, 90%.

ASO#	Sequence (5' → 3')	Expected Mass (Da)	Found Mass (Da)	Notes
S-ASO-1	TGTAACTGAGGTAAGAGG	5627.6	5628.0	Complementary DNA
S-ASO-2	UGUAACUGAGGUAAGAGG	5859.6	5861.0	Complementary full RNA
S-ASO-3	UGAGGUAAGA	3247.9	3248.0	Complementary fragment RNA

Table B.1: Table of oligonucleotides which were used as reverse complements in the formation of ASO:DNA and ASO:RNA heteroduplexes in Chapters 2-4. All oligonucleotides were purchased from *Integrated DNA Technologies*.

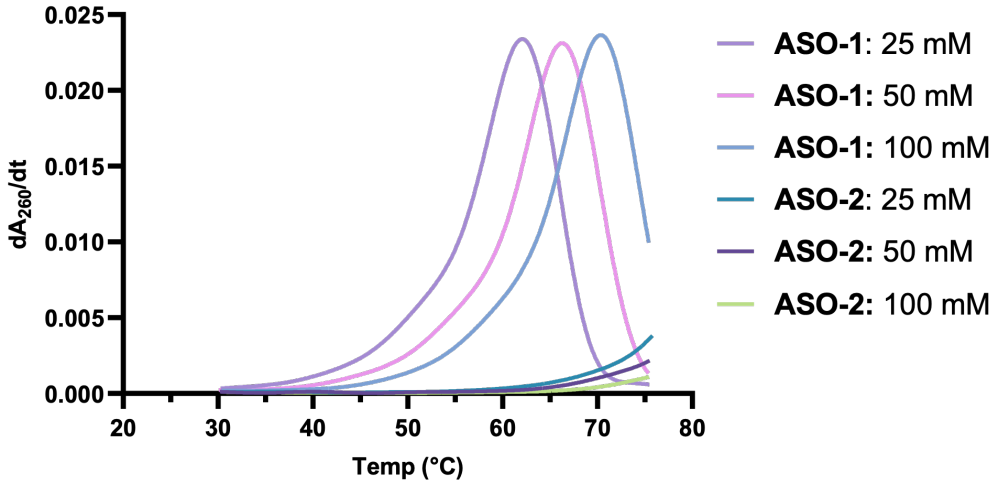


Figure B.2: First derivative curves of melting temperatures for the 2'OMe/PS control (ASO-1) and LNA control ASO (ASO-2) against the full complementary RNA target at varying concentrations of NaCl (10 mM phosphate, pH = 7.0.

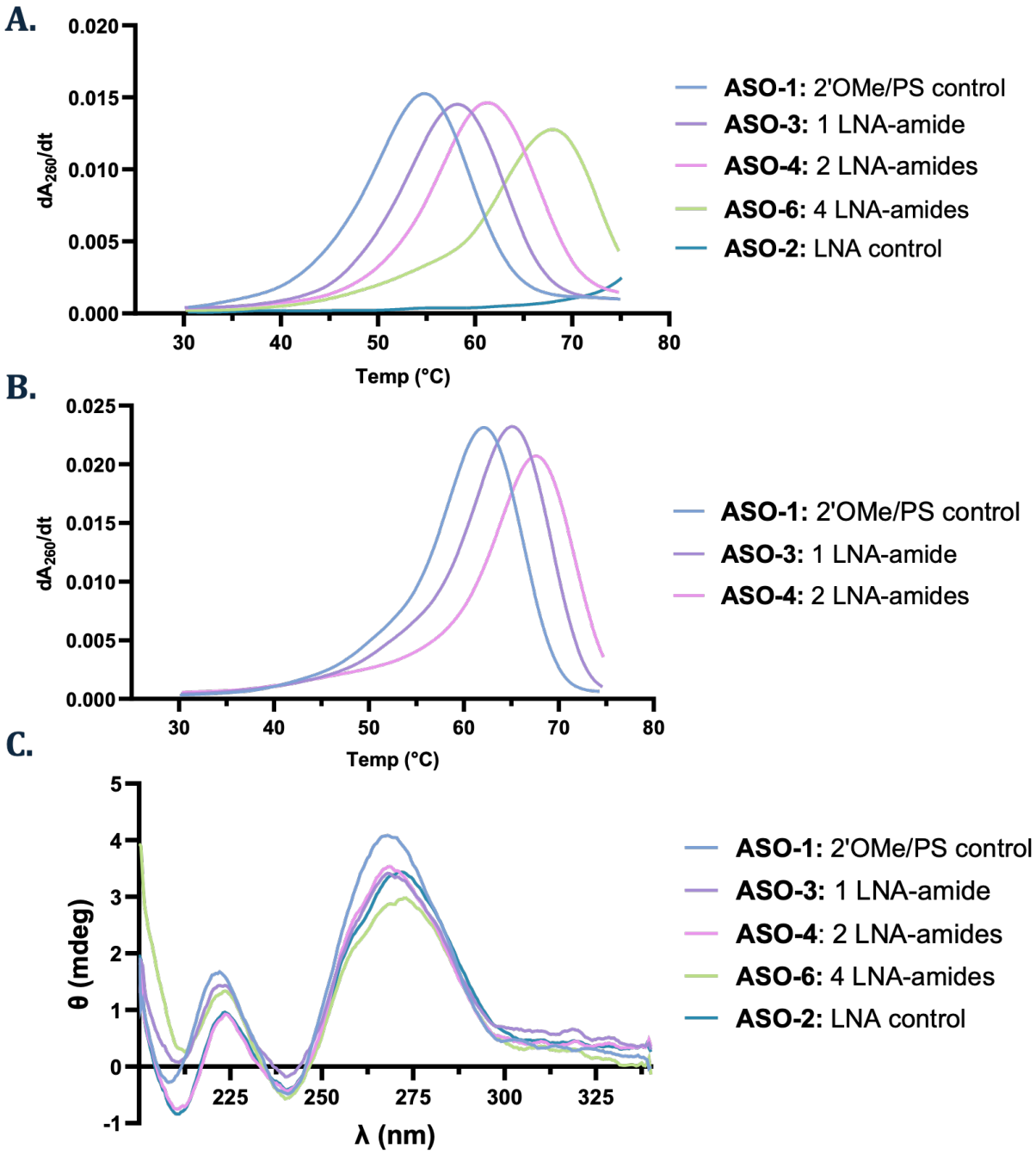


Figure B.3: First derivative curves of melting temperatures for ASOs containing LNA-amide linkages against complementary DNA (200 mM NaCl, 10 mM phosphate, pH = 7.0) (**A.**) and against the full complementary RNA (25 mM NaCl, 10 mM phosphate, pH = 7.0) (**B.**). **C.** Circular dicroism spectra of ASOs containing LNA-amide linkages in an ASO:DNA heteroduplex (200 mM NaCl, 10 mM phosphate, pH = 7.0).

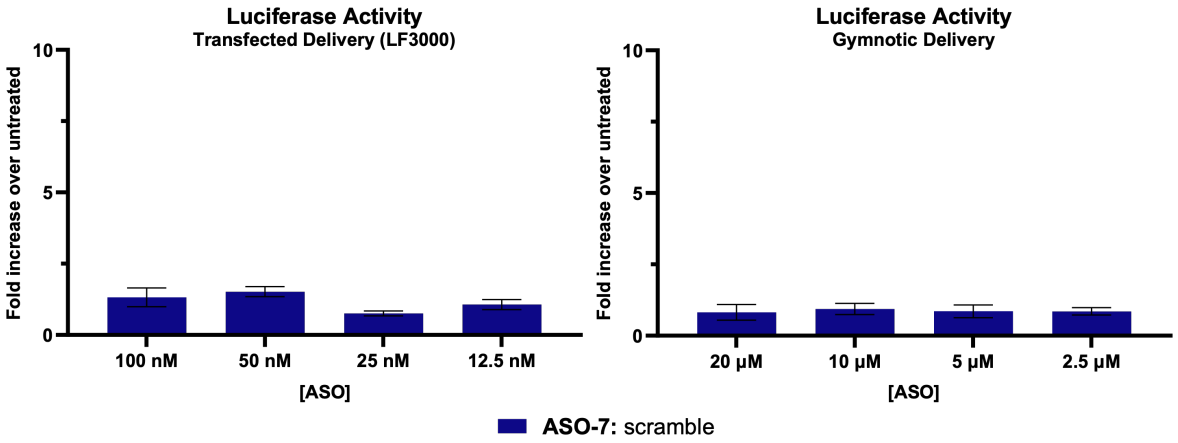


Figure B.4: Dose-response splice-switching activity of the scramble (ASO-7) using the luciferase reporter assay (HeLa pLuc/705 cell line), delivered via transfection with Lipofectamine 3000 (left) and via gymnosis (right). Activity was measured as luminescence normalised first to protein quantity then to untreated cells. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

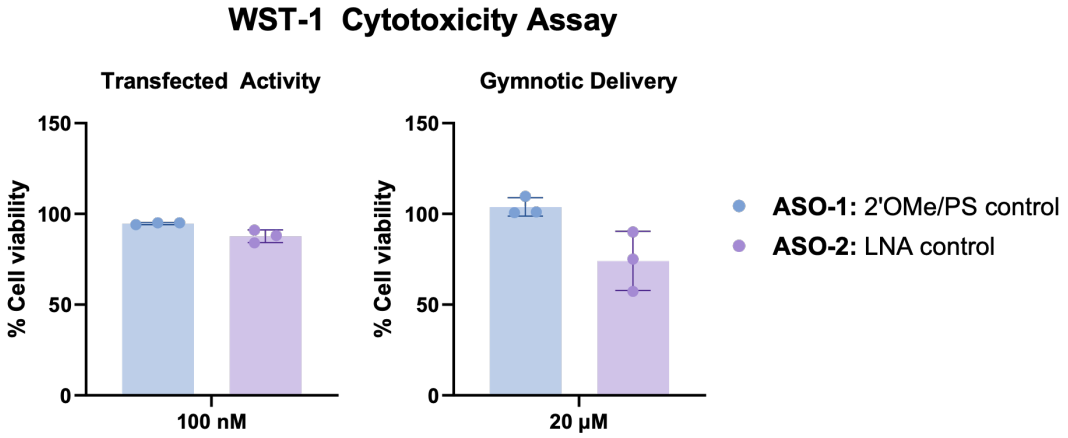
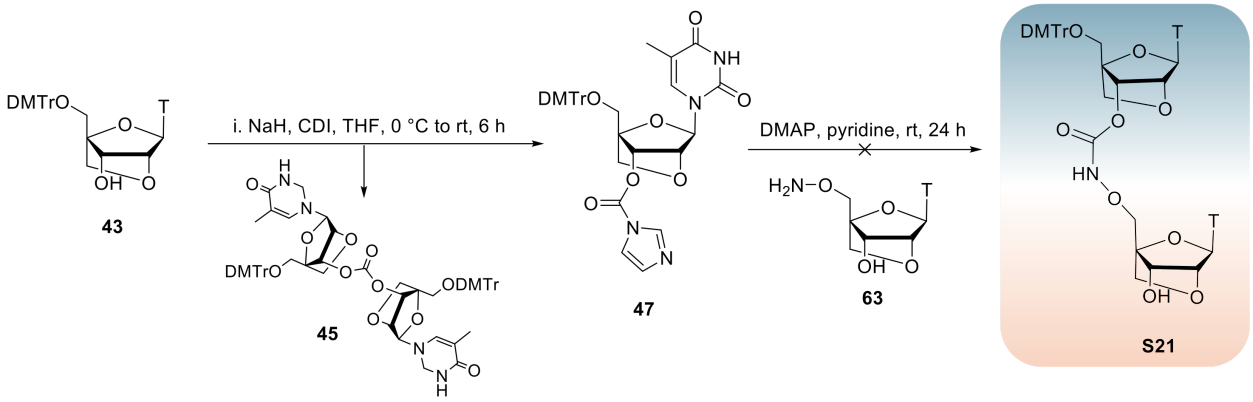


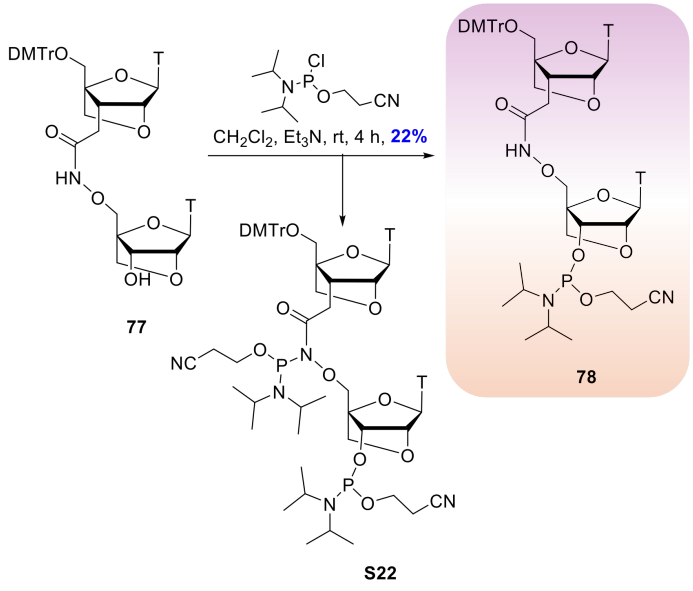
Figure B.5: Cytotoxicity (as % cell viability using untreated cells as 100% viability) of positive control ASOs (ASOs 1-2) using the highest transfected dose (100 nM, left) and gymnotic dose (20 µM, right). Each dot is representative of a single technical replicate.

C

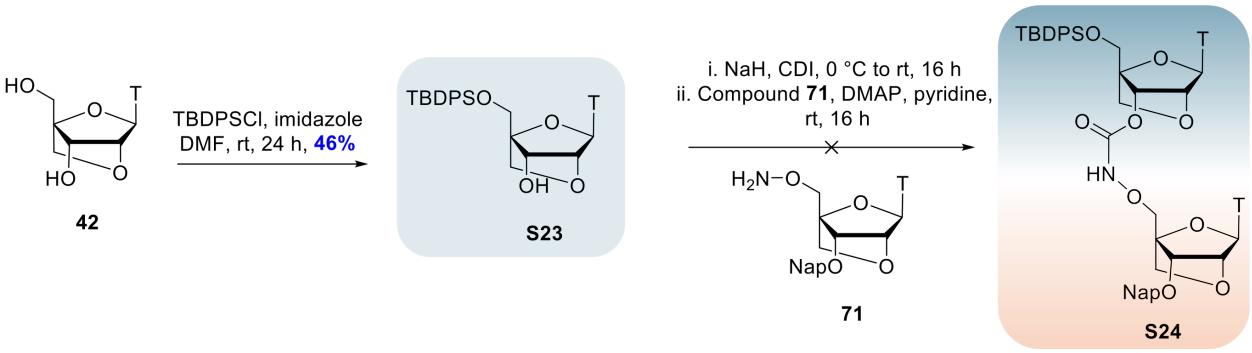
Chapter 3 Supplementary Figures



Scheme C.1: Attempted synthesis of amino-oxycarbamate dimer **S21** via the 5'-amino-oxy monomer **63** and the activated electrophile **47**. The carbonate dimer **45** by-product was formed upon deprotonation of the 3'-alcohol **43**.



Scheme C.2: Phosphitylation of LNA-amino-oxyamide dimer **77** to form phosphoramidite dimer **78**. Second phosphitylation at the reactive NH position in the linkage was suspected during the reaction to form the by-product **S22**.



Scheme C.3: Attempted synthetic route to achieve amino-oxy-carbamate dimer **S24**. 5'-TBPDS-protected alcohol **S23** was not activated in sufficient quantities (via NaH and CDI) to achieve dimer **S24**.

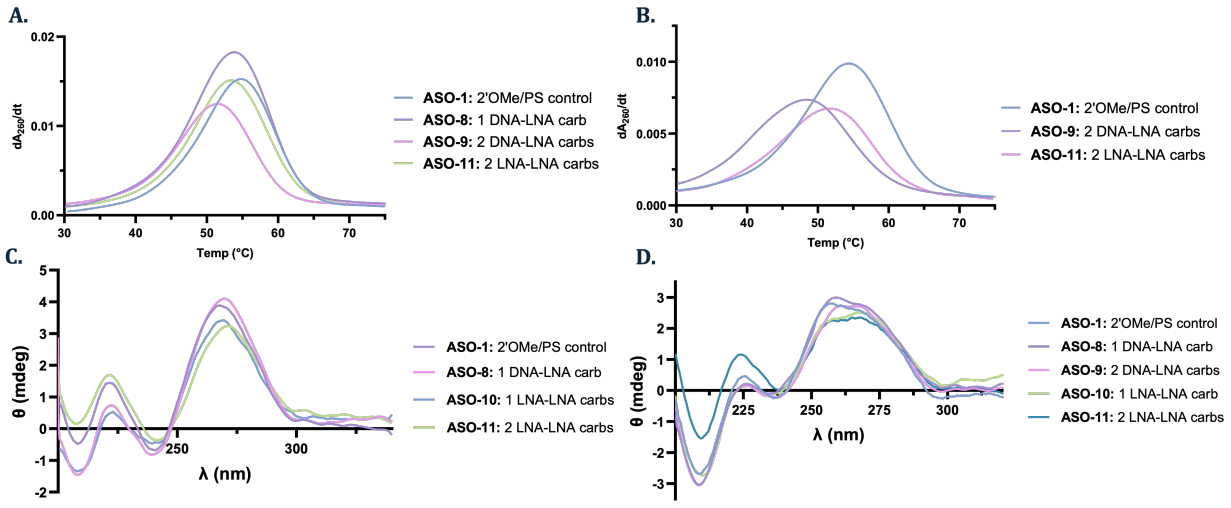


Figure C.1: Additional biophysical properties of ASOs containing carbamate linkages. First derivative curves of melting temperatures of ASOs containing carbamate linkages against the DNA complement (200 mM NaCl, 10 mM phosphate, pH = 7.0) (**A.**) and the RNA fragment complement (25 mM NaCl, 10 mM phosphate, pH = 7.0) (**B.**). Circular dichroism spectra of ASOs containing carbamate linkages in ASO:DNA heteroduplexes (**C.**) (200 mM NaCl, 10 mM phosphate, pH = 7.0) and ASO:RNA heteroduplexes (**D.**) (25 mM NaCl, 10 mM phosphate, pH = 7.0).

D

Chapter 4 Supplementary Figures

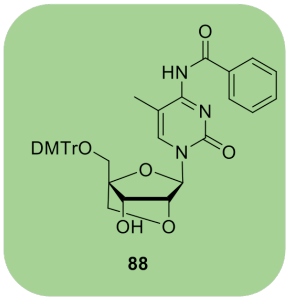
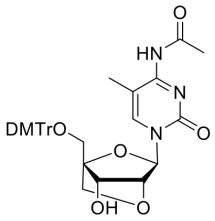
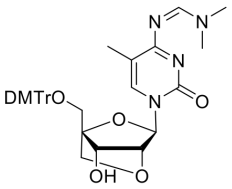
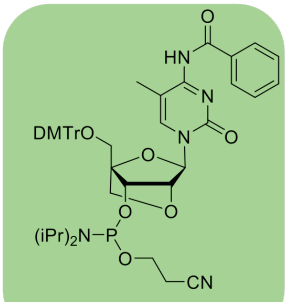
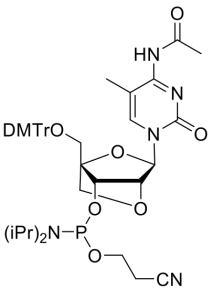
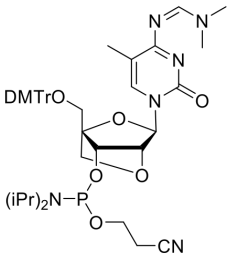
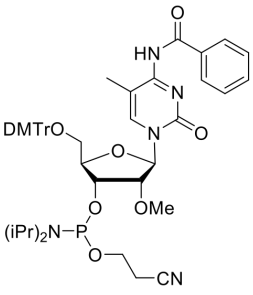
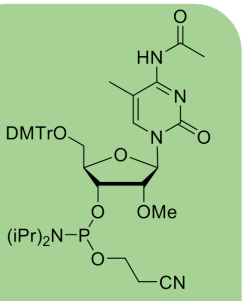
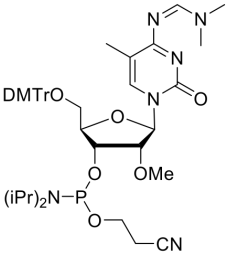
Commercially available	^{Me}C(Bz)	^{Me}C(Ac)	^{Me}C(dmf)
locked 3'-OH monomer	 **88**	 **99**	 **96**
locked phosphor-amidite	 **S25**	 **S26**	 **98**
2'OMe phosphor-amidite	 **S27**	 **S28**	 **S29**

Table D.1: Summary of commercially available modified ^{Me}cytosine monomers, organised by protecting group and sugar modifications. The commercially available LNA monomer **88** and corresponding LNA phosphoramidite **S25** are benzoyl protected. Due to the deprotection conditions discussed in Chapter 4, the oligonucleotides containing LNA-LNA isopropyl triesters required acetyl or dimethylformamide protecting groups for ^{Me}C incorporations, leading to the synthesis of monomers **99**, **96**, and **98**. The remainder of the oligonucleotide was synthesised using 2'OMe sugar modifications; the ^{Me}C (and C) version of this is commercially available as acetyl-protected (**S28**).

E

Chapter 5 Supplementary Figures

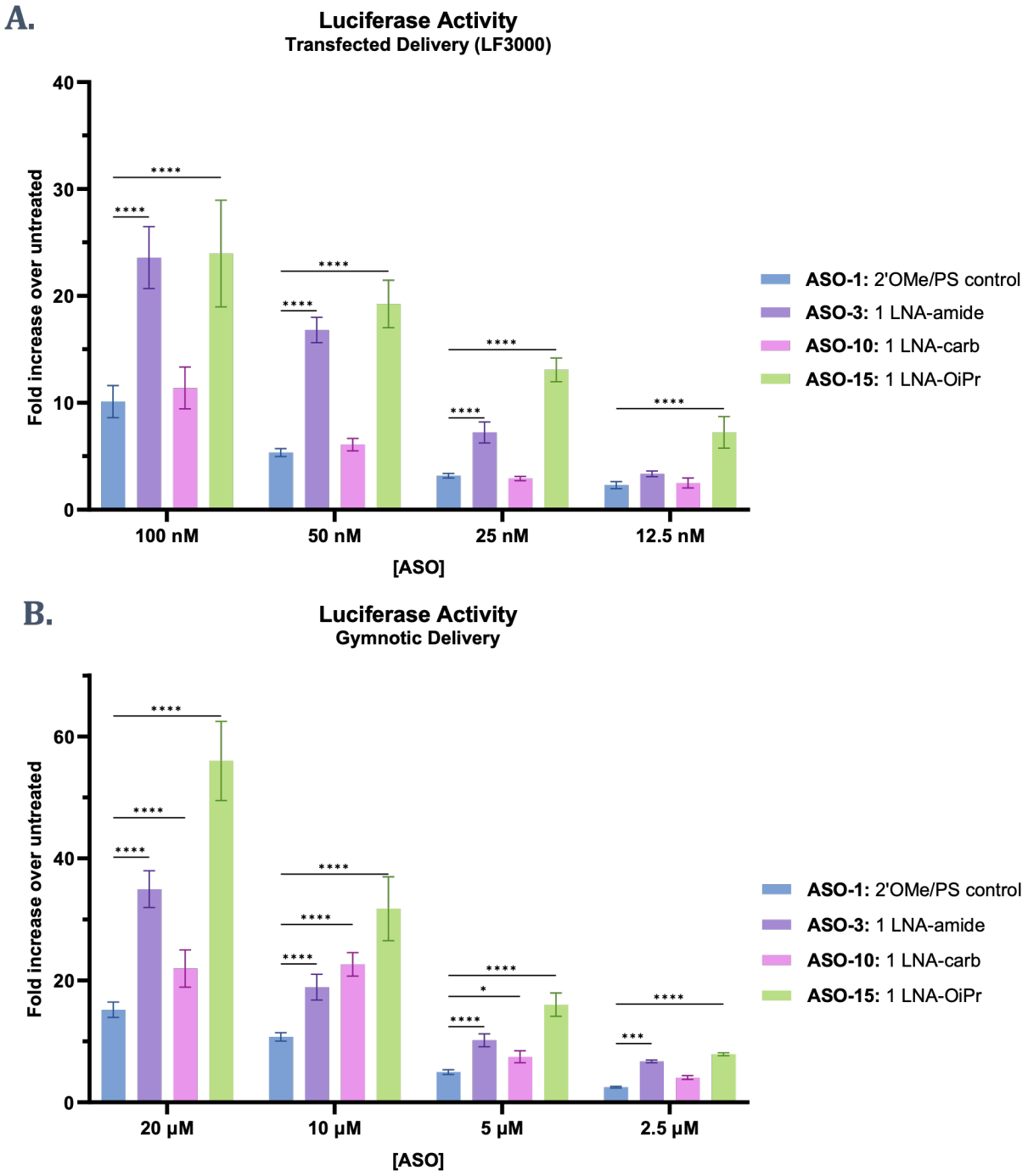


Figure E.1: Dose-response splice-switching activity of ASOs containing 1 LNA-neutral linkage delivered via Lipofectamine 3000 (**A.**) or via gymnosis (**B.**). Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS ASO as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

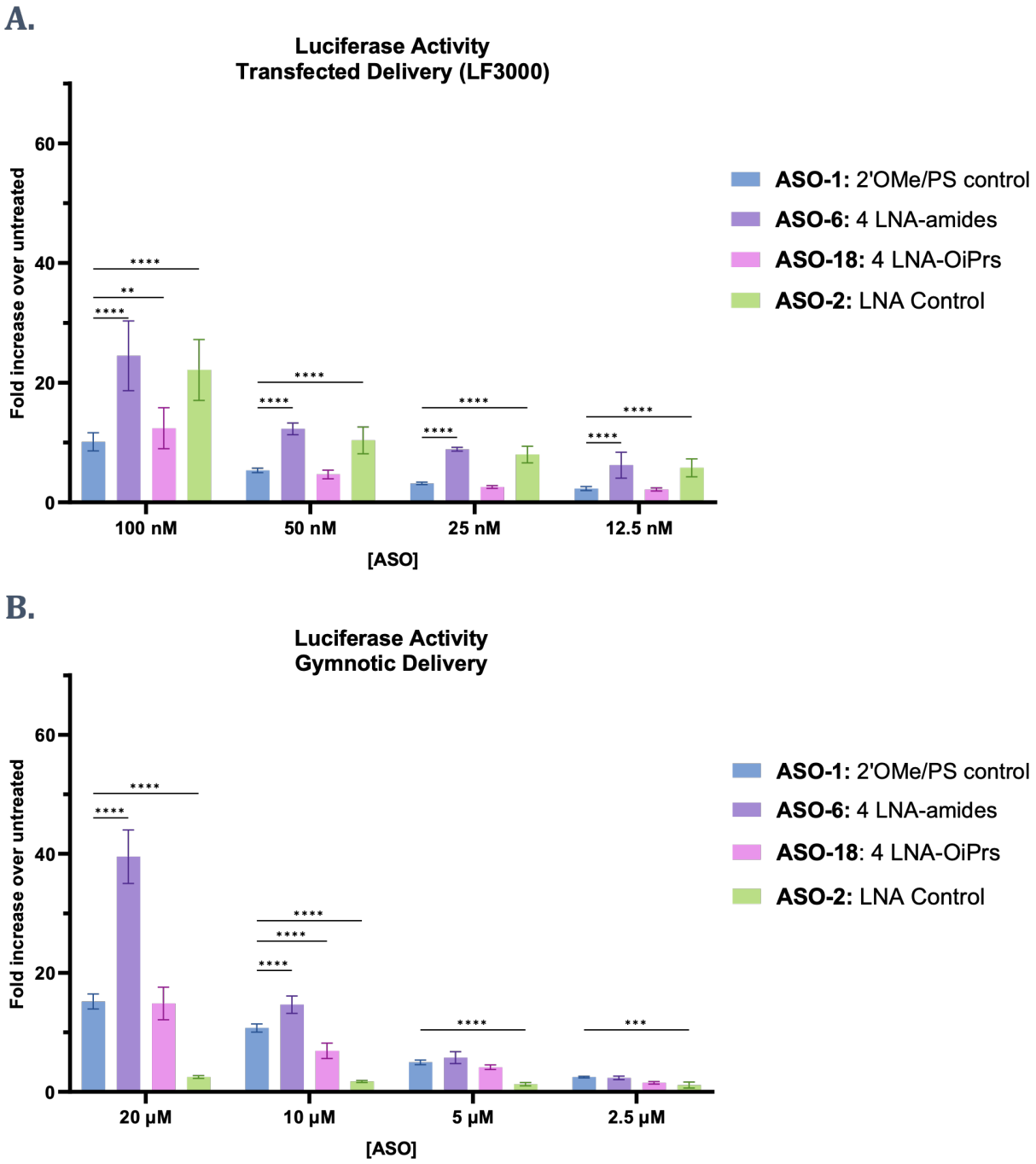


Figure E.2: Dose-response splice-switching activity of ASOs containing 4 LNA-neutral linkages delivered via Lipofectamine 3000 (**A.**) or via gymnosis (**B.**). Activity was measured as luminescence normalised first to protein quantity then to untreated cells. Statistical significance was determined using a 2-way ANOVA test using 2'OMe/PS ASO as the control within each concentration. *represents $P < 0.05$, **represents $P < 0.01$, ***represents $P < 0.001$, ****represents $P < 0.0001$. All data are given as the means of distinct biological replicates ($n = 3$); each biological replicate is the mean ±SD of technical replicates ($n = 3$).

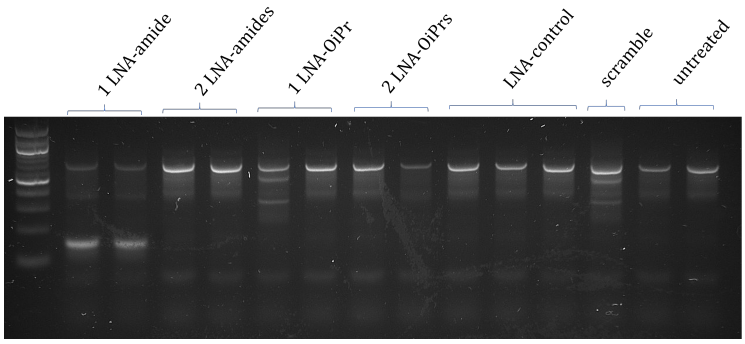


Figure E.3: Agarose gel produced by entry 4 in Figure 5.6. Amplification conditions: Luna One-Step polymerase enzyme, 100 ng cDNA, nested PCR (primers: 20-26 outer, 20-26 inner).

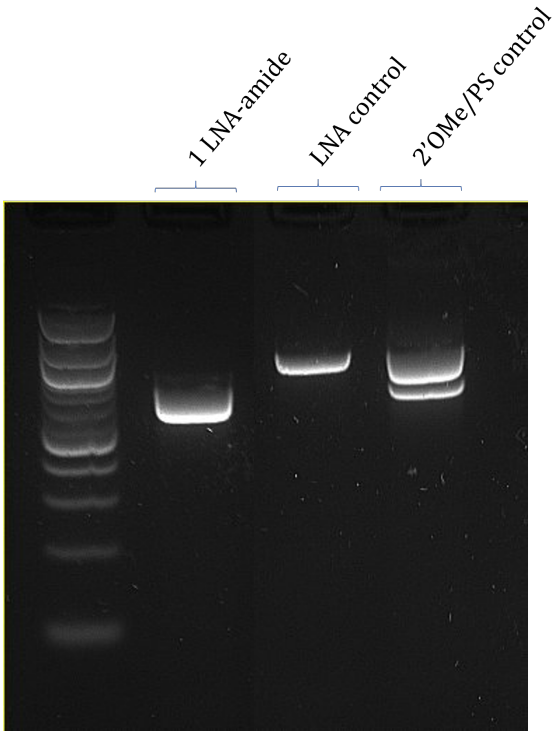
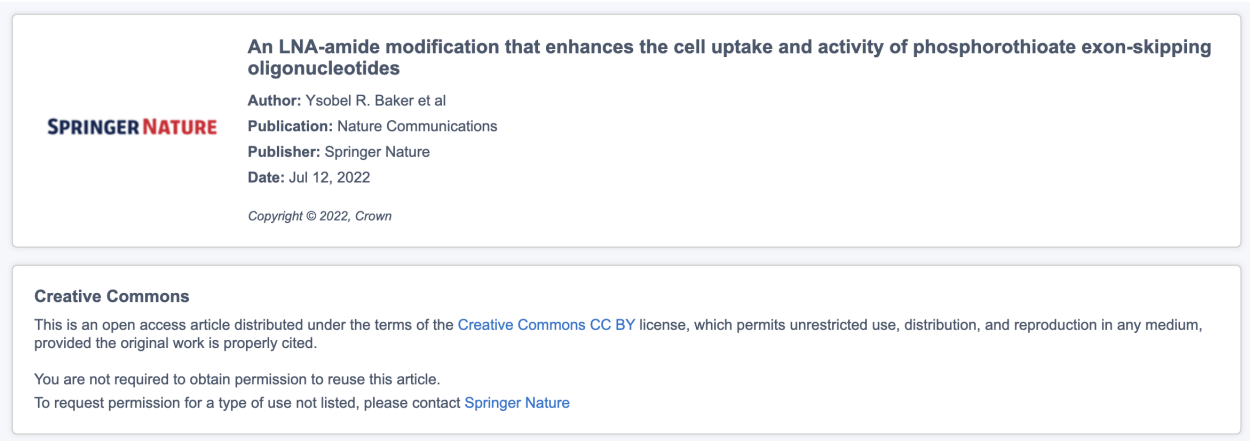


Figure E.4: Agarose gel produced by entry 8 in Figure 5.6. Amplification conditions: Amplitaq Gold polymerase enzyme, 400 ng cDNA, nested PCR (primers: 20-26 outer, 20-26 inner).

Copyright

This thesis includes data reproduced with permission according to Creative Commons Licences:

- Baker et al.[62]



SPRINGER NATURE

An LNA-amide modification that enhances the cell uptake and activity of phosphorothioate exon-skipping oligonucleotides

Author: Ysobel R. Baker et al
Publication: Nature Communications
Publisher: Springer Nature
Date: Jul 12, 2022

Copyright © 2022, Crown

Creative Commons

This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

You are not required to obtain permission to reuse this article.
To request permission for a type of use not listed, please contact Springer Nature

- Thorpe et al.[145]

Synthesis and biophysical properties of carbamate-locked nucleic acid (LNA) oligonucleotides with potential antisense applications

C. Thorpe, S. Epple, B. Woods, A. H. El-Sagheer and T. Brown, *Org. Biomol. Chem.*, 2019, **17**, 5341 **DOI:** 10.1039/C9OB00691E

This article is licensed under a Creative Commons Attribution 3.0 Unported Licence. **You can use material from this article in other publications without requesting further permissions** from the RSC, provided that the correct acknowledgement is given.

Read more about how to correctly acknowledge RSC content.

"I got power, poison, pain, and joy inside my DNA"

– Kendrick Lamar

⁂