# Backbone-Modified Oligonucleotides for Biology and Medicine




Sven Epple

Keble College, University of Oxford

*A thesis submitted for the degree of Doctor of Philosophy*

November 2021

© Sven Epple

# Declaration

I, Sven Epple, am the author of the thesis titled "*Backbone-Modified Oligonucleotides for Biology and Medicine*" and declare that the relating research presented in this thesis was performed and written by the author under the supervision of Prof. Tom Brown at the University of Oxford. The research is original and has not been submitted in full or in part for any other degree at this or at any other university. Appropriate citations/acknowledgements are included where the presented content was adapted from literature or resulted from someone else's contributions. The copyright of this thesis rests with the author.

# Acknowledgements

I would like to thank Prof. Tom Brown for this fantastic opportunity to carry out my research in his lab, and for all his guidance and support, especially during such a difficult time when the coronavirus pandemic caused many uncertainties. I would like to thank Afaf for her encouragement during my DPhil studies and for many good discussions—science and non-science related. I am extremely grateful to Cam for all the time that we have spent together in and outside from the lab. Working with Cam and Ben during the early days of my time in the lab was a great joy, and we were truly a dream team trio. I would like to thank Pui for being an inspiration in any possible way. I was always striving for her level of kindness, joy, creativity, tenacity, and strength. I would like to thank Yssy for her guidance and dedication in proofreading my transfer report and the paper manuscripts, and for her contributions to my research. I am incredibly lucky to being part of a fantastic research group, so I want to express my gratitude to all the Brownies: Cam, thanks for being such a great labmate, you are amazing! Nicolò, you are amazing, too, and thanks for introducing the Wednesday lunchtime trips to the Gloucester market—everybody knows how I was planning my whole week around this. Sarah, our CRISPR discussions were great, and you were truly the social butterfly of the group keeping everybody connected and dragging us lazy people out on a Friday night "after school". Thank you Thanishta, Chunsen, Jerry and Phoebe for an amazing time and for going through the ups and downs of a DPhil with me. I am very grateful to Nicolò and Giorgio for proofreading this thesis and for their helpful comments. I would like to thank Batool for inviting me for a site visit to GSK and for many insightful virtual meeting that guided me on my scientific and professional journey.

Outside from the lab, I want to thank Tanya, Ben, Edward, Mariel, Danny and Julien for being great friends and for many amazing moments. I am incredibly grateful to Giorgio who supported me in any possible way that I could think of: Ti amo con tutto il mio cuore! Meinen Eltern danke ich für ihre bedingungslose Unterstützung, die mir alles auf meinem bisherigen Lebensweg erst ermöglicht hat.

# Publications

- Baker, R. Y.; Thorpe, C.; Cheng, J.; Poller, M. L.; Cox, L.; Kumar, P.; Lim, F. W.; Lie, L.; McClorey, G.; **Epple, S.**; Singleton, D.; McDonough, M.; Hardwick, J.; Christensen, E. K.; Wood, M.; Hall, P. J.; El-Sagheer, H. A.; Brown, T. Oligonucleotide analogues with locked-amide linkages have enhanced therapeutic potential. (*submitted*).

- **Epple, S**.; Modi, A.; Baker, Y. R.; Węgrzyn, E.; Traoré, D.; Wanat, P.; Tyburn, A. E. S.; Shivalingam, A.; Taemaitree, L.; El-Sagheer, A. H.; Brown, T. A New 1,5-Disubstituted Triazole DNA Backbone Mimic with Enhanced Polymerase Compatibility. *J. Am. Chem. Soc.* 2021, *143* (39), 16293-16301.

- **Epple, S**.; El-Sagheer, A. H.; Brown, T. Artificial Nucleic Acid Backbones and Their Applications in Therapeutics, Synthetic Biology and Biotechnology. *Emerg. Top. Life Sci.* 2021, *5* (5), 691–697.

- **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. *Chem. Commun.* 2020, *56* (41), 5496–5499.

- 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. *Org. Biomol. Chem.* 2019, *17* (21), 5341–5348.

The right to reproduce the articles, or adaption of the articles, is retained by the authors as per the publishers' copyright policies[a,b] or has been granted upon request.[c] Appropriate acknowledgements are included where the above-listed articles are reproduced in whole or in part.

[a] https://www.rsc.org/journals-books-databases/author-and-reviewer-hub/authors-information/licences-copyright-permissions/#acknowledgements
[b] https://portlandpress.com/pages/copyright_and_permissions
[c] see Copyright, page 382.

# Table of Contents

**Declaration** ..... **III**

**Acknowledgements** ..... **V**

**Publications** ..... **VII**

**Table of Contents** ..... **IX**

**Abbreviations** ..... **XIII**

**Abstract** ..... **17**

**CHAPTER I INTRODUCTION** **19**

**1. The Structure of Nucleic Acids** ..... **23**
1.1. Primary Structure of Nucleic Acids 23
1.2. Secondary Structure of Nucleic Acids 28

**2. Oligonucleotide Synthesis** ..... **35**

**3. Modified Oligonucleotides** ..... **43**
3.1. Sites of Modifications 44
3.2. Synthesis of Modified Oligonucleotides 45
3.3. Synthesis of Backbone-Modified Oligonucleotides 47
3.3.1. Backbone Formation During Solid Phase Oligonucleotide Synthesis 47
3.3.1.1. The Monomer Approach 48
3.3.1.2. The Dimer Approach 51
3.3.2. Post-Synthetic Artificial Backbone Formation 52

**4. Antisense Oligonucleotides** ..... **55**
4.1. Introduction to Antisense Oligonucleotides 55
4.2. Mechanisms of Antisense Oligonucleotides 56
4.3. Chemical Modifications of Antisense Oligonucleotides 60

**5. Backbone Modifications in Synthetic Biology** ..... **65**
5.1. Construction of Active Biomolecules 65
5.2. Synthetic Genes 67

**6. Overall Structure of the Thesis** ..... **71**

**CHAPTER II CARBAMATE–LOCKED NUCLEIC ACID CHIMERAS FOR POTENTIAL ANTISENSE APPLICATIONS** **73**

**1. Background** ..... **75**
1.1. The Carbamate Internucleoside Linkage 75
1.2. The Locked Nucleic Acid Modification 78
1.2.1. Biophysical Properties 78

**2. Aim of the Chapter ..... 81**

**3. Results and Discussion ..... 83**
3.1. Synthesis of Carbamate-Linked Dinucleoside Phosphoramidites 83
3.1.1. Synthesis of Nucleophiles 83
3.1.2. Synthesis of Electrophiles 85
3.1.3. Synthesis of Dimers 86
3.2. Biophysical Studies of Carbamate-Linked LNA Oligonucleotides 88
3.2.1. Duplex Stabilities 88
3.2.2. Structural Integrity and Nucleolytic Resistance 92

**4. Conclusion ..... 95**

**5. Future Work ..... 97**

**CHAPTER III BUILDING BLOCKS FOR THE SYNTHESIS OF AMIDE-LINKED OLIGONUCLEOTIDES 101**

**1. Background ..... 103**
1.1. The Amide Internucleoside Linkage 103
1.1.1. Biophysical Properties 103
1.1.2. Synthesis 106

**2. Aim of the Chapter ..... 111**

**3. Results and Discussion ..... 113**
3.1. Part I: Synthesis of an Amide-Linked Dinucleoside Phosphoramidite Dimer for *En Bloc* Amide Addition. 113
3.2. Part II: Synthesis of an Amino–Carboxylic Acid Nucleoside for Consecutive On-Resin Amide Bond Formations 114
3.2.1. Determination of a Suitable 5'-Protecting Group 116
3.2.2. Synthesis of a Suitably Protected Amino–Carboxylic Acid Nucleoside 119

**4. Conclusion ..... 121**

**CHAPTER IV AMIDE-LINKED OLIGONUCLEOTIDES FOR RNASE H ACTIVATION 123**

**1. Background ..... 125**
1.1. RNase H in Antisense Applications 125
1.1.1. RNase H Mechanism 125
1.2. RNase H-Compatible Modifications 130
1.3. The Gapmer Strategy 131
1.4. Charge-Neutral Backbone Modifications for Enhanced Cellular Uptake 133

**2. Aim of the Chapter ..... 137**

**3. Results and Discussion ..... 141**
3.1. Oligonucleotides with Isolated Amide Linkages 141
3.1.1. Synthesis of Oligonucleotides with Isolated Amides 141
3.1.2. Oligonucleotides with Isolated Amides for RNase H Activation 142
3.2. Oligonucleotides with Consecutive Amide Linkages for Antisense Application 144
3.2.1. Synthesis of Oligonucleotides with Consecutive Amides 144
3.2.2. Target Engagement of Oligonucleotides with Consecutive Amides 150
3.2.2.1. Duplex Stability 151

3.2.2.2. Duplex Structure 152
3.2.3. Oligonucleotides with Consecutive Amides for RNase H Activation 154
3.2.3.1. Determination of the Cutting Site 156
3.2.3.2. Cutting Rate 158
3.2.3.3. Cutting Rate: Further Results 163
3.2.3.4. Mismatch Discrimination 165
3.2.4. Serum Stability of Oligonucleotides with Consecutive Amides 169
3.2.5. Cellular Uptake of Oligonucleotides with Consecutive Amides 171

**4. Conclusion ..... 177**

**5. Future Work ..... 179**
5.1. New Chemical Approaches for Monomer Synthesis 179

**CHAPTER V TRIAZOLE-LINKED OLIGONUCLEOTIDES FOR SYNTHETIC BIOLOGY 183**

**1. Background ..... 185**
1.1. Formation of Triazoles 185
1.1.1. The Copper(I)-Catalysed Azide–Alkyne Cycloaddition 186
1.1.2. The Ruthenium(II)-Catalysed Azide–Alkyne Cycloaddition 187
1.1.3. The Structural Diversity of the Triazole Internucleoside Linkage 189
1.2. Triazoles in Genetic Templates 191
1.2.1. Preparation of Libraries 191
1.2.2. Triazole-Modified Primers in PCR 193
1.2.3. Synthesis of Triazole-Linked Genes 194
1.2.4. Circular Genetic Templates 195

**2. Aim of the Chapter ..... 197**

**3. Results and Discussion ..... 199**
3.1. Synthesis of Triazole-Linked Dinucleoside Phosphoramidites 199
3.1.1. Synthesis 199
3.2. Biocompatibility of the Triazole Linkages 203
3.2.1. Target Engagement of Oligonucleotides with Triazole Linkages 204
3.2.1.1. Duplex Stability 204
3.2.1.2. Duplex Structure 206
3.2.2. Replication Competency of Oligonucleotides with Triazole Linkages 207
3.2.2.1. Primer Extension Assay 208
3.2.2.2. Quantitative PCR 214
3.2.2.3. Sequencing 221
3.2.3. Structure–Biocompatibility Relationship 222

**4. Conclusion ..... 227**

**5. Future Work ..... 229**

**CHAPTER VI OVERALL CONCLUSIONS AND FUTURE DIRECTIONS 231**

**CHAPTER VII MATERIALS AND METHODS 239**

**1. Experimental Procedures ..... 241**
1.1. UV Melting Studies 241

1.2. Circular Dichroism Studies 241
1.3. Foetal Bovine Serum Assay 242
1.4. RNase H Assay 242
1.5. Gel Extraction 243
1.6. Cell Culture 244
1.7. Cell Preparation for Microscopy 244
1.8. Confocal Microscopy 245
1.9. Linear Copying Assays 246
1.10. Linear Copying for Mass Spectrometry 247
1.11. qPCR Kinetics 247
1.12. Sequencing of PCR Amplicons 248

**2. Oligonucleotide Synthesis ..... 251**
2.1. Synthesis of DNA Oligonucleotides 251
2.2. Synthesis of RNA Oligonucleotides 252
2.3. Synthesis of PO–AM Chimeras 253

**3. Purification and Characterisation of Oligonucleotides ..... 255**
3.1. Purification by HPLC 255
3.2. Purification by Gel Electrophoresis 255
3.3. Characterisation by Mass Spectrometry 256

**4. Small Molecule Synthesis ..... 257**
4.1. General Procedures 257
4.2. Compounds of Chapter II 259
4.3. Compounds of Chapter III 286
4.4. Compounds of Chapter IV 293
4.5. Compounds of Chapter V 297

**REFERENCES 315**

**APPENDIX 331**

**1. List of Oligonucleotides ..... 333**

**2. Additional Experimental Data ..... 337**
2.1. Data of Chapter II 337
2.2. Data of Chapter IV 339
2.3. Data of Chapter V 341

**3. NMR Spectra ..... 361**

**4. Copyright ..... 382**

# Abbreviations

| | |
|---|---|
| **α** | stereochemical descriptor |
| **Å** | angstrom |
| **A** | adenosine |
| ***A*** | absorbance |
| **Ac** | acetyl |
| **Ade** | adenine |
| **AF488** | Alexa Fluor™ 488 |
| **AF647** | Alexa Fluor™ 647 |
| **AIBN** | azaisobutyronitrile |
| **AM** | amide internucleoside linkage |
| **ANA** | arabinonucleic acid |
| **anhyd** | anhydrous |
| **aq** | aqueous |
| **Ar** | aryl |
| **ASO(s)** | antisense oligonucleotide(s) |
| **a.u.** | arbitrary units |
| **β** | stereochemical descriptor |
| **b** | broad |
| **Bn** | benzyl |
| **bp** | base pair |
| **BSA** | *N*,*O*-bis(trimethylsilyl)acetamide |
| **BTT** | 5-(benzylthio)-1H-tetrazole |
| **Bu** | butyl |
| **Bz** | benzoyl |
| **C** | Celsius<br>cytidine |
| **c** | centi ($10^{-2}$)<br>concentration |
| **CA** | carbamate internucleoside linkage |
| **calcd** | calculated |
| **Cas9** | CRISPR-associated protein 9 |
| **CD** | circular dichroism |
| **CE** | 2-cyanoethyl |
| **CeNA** | cyclohexene nucleic acid |
| **concd** | concentrated |
| **COSY** | correlation spectroscopy |
| **CPG** | controlled pore glass |
| **CRISPR** | clustered regularly interspaced short palindromic repeats |
| **CSLM** | confocal scanning laser miscroscopy |
| $C_t$ | threshold cycle |
| **CuAAC** | copper(I)-catalysed azide–alkyne cycloaddition |
| **Cyt** | cytosine |
| **δ** | NMR chemical shift in ppm downfield from a standard |
| **Δ** | difference, change |
| **d** | doublet (spectra) |
| **D** | absolute configuration |
| **Da** | dalton(s) |
| **dA** | deoxyadenosine |
| **DAPI** | 4′,6-diamidino-2-phenylindole |
| **dC** | deoxycytidine |
| **DDAB** | didodecyldimethylammonium bromide |
| **DEA** | *N*,*N*-diethylamine |
| **DFT** | density functional theory |
| **dG** | deoxyguanidine |
| **DIPEA** | *N*,*N*-diisopropylethylamine |
| **DMAP** | 4-dimethylaminopyridine |
| **DMD** | Duchenne Muscular Dystrophy |
| **DMEM** | Dulbecco's modified eagle medium |
| **DMF** | *N*,*N*-dimethylformamide |
| **DMP** | Dess–Martin periodinane |
| **DMSO** | dimethylsulfoxide |
| **DMSO-$d_6$** | deuterated dimethylsulfoxide |
| **DMTr** | 4,4′-dimethoxytrityl |
| **dNTP** | deoxynucleoside triphosphate |
| **DTT** | dithiothreitol |
| **ε** | molar absorptivity |
| **E** | envelope sugar pucker |
| ***E*** | entgegen (configurational) |
| **EDC** | 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride |
| **EDTA** | ethylenediaminetetraacetic acid |
| **EGF** | epidermal growth factor |
| **EGFP** | enhanced green fluorescent protein |
| **ESI** | electrospray ionisation |
| **Et** | ethyl |
| **EU** | enzyme unit |
| **FAM** | fluorescein amidite |
| **FANA** | 2′-deoxy-2′-fluoro-arabinonucleic acid |
| **FBS** | foetal bovine serum |
| **FDA** | Food and Drug Administration |
| **FGI** | functional group interconversion |
| **FL** | fluorescein |
| **Fmoc** | fluorenylmethyloxycarbonyl |
| **g** | gram(s) |
| **G** | guanosine |
| ***G*** | Gibbs free energy |
| **Gua** | guanine |
| **h** | hour(s) |
| ***H*** | enthalpy |
| **HATU** | 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate |

| | |
|---|---|
| **HIV** | human immunodeficiency virus |
| **HMBC** | heteronuclear multiple bond correlation |
| **HPLC** | high performance liquid chromatography |
| **HRMS** | high resolution mass spectrometry |
| **HSQC** | heteronuclear single quantum coherence |
| **HTT** | huntingtin |
| **Hz** | Hertz |
| **iBu** | isobutyryl |
| **$^{i}$Pr** | isopropyl |
| **IR** | infrared |
| ***J*** | coupling constant |
| **λ** | wavelength |
| **L** | litre(s) |
| **LA** | Lewis acid |
| **LB** | Luria–Bertani |
| **LNA** | locked nucleic acid |
| **µ** | micro ($10^{-6}$) |
| **m** | metre(s)<br>milli ($10^{-3}$)<br>multiplet (spectra) |
| **M** | molar |
| **M** | mega ($10^{6}$)<br>monoisotopic mass |
| **Me** | methyl |
| **$^{Me}$C** | 5-methyl cytosine |
| **MEM** | minimal essential medium |
| **mesyl** | methanesulfonyl (also Ms) |
| **mHTT** | mutated huntingtin |
| **min** | minute(s) |
| **MMTr** | 4-monomethoxytrityl |
| **mod** | modification |
| **MOE** | methoxyethyl |
| **mol** | mole(s) |
| **MP** | methylphosphonate |
| **mRNA** | messenger RNA |
| **Ms** | methanesulfonyl (also mesyl) |
| **MS** | mass spectrometry<br>molecular sieves |
| **MW** | molecular weight |
| ***m/z*** | mass-to-charge ratio |
| **n** | nano ($10^{-9}$)<br>unspecific ribonucleoside |
| **N** | normal |
| **N** | unspecific deoxynucleoside |
| **NGS** | next-generation sequencing |
| **NiAAC** | nickel-catalysed azide–alkyne cycloaddition |
| **NMI** | *N*-methylimidazole |
| **NMM** | *N*-methylmorpholine |
| **NMR** | nuclear magnetic resonance |
| **ON(s)** | oligonucleotide(s) |
| **org** | organic |
| **p** | pico ($10^{-12}$) |
| ***p*** | para |
| ***P*** | pseudorotation phase angle |
| **PA** | phosphoramidate |
| **PACE** | phosphonoacetate |
| **PAGE** | polyacrylamide gel electrophoresis |
| **PBS** | phosphate-buffered saline |
| **PCR** | polymerase chain reaction |
| **PDB** | protein data bank |
| **PG** | protecting group |
| **pH** | negative logarithm of hydrogen ion concentration |
| **Ph** | phenyl |
| **PMO** | phosphorodiamidate morpholino oligonucleotide |
| **PNA** | peptide nucleic acid |
| **PO** | phosphodiester |
| **PO-BP** | phosphodiester bond-binding pocket |
| **ppm** | parts per million |
| **Pr** | propyl |
| **PS** | phosphorothioate |
| **Py** | pyridine |
| **PyBOP** | (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate |
| **q** | quartet (spectra) |
| **qPCR** | quantitative polymerase chain reaction |
| **quant** | quantitative |
| ***R*** | stereochemical descriptor |
| **$R_f$** | retention factor |
| **RNAi** | RNA interference |
| **RNase H** | ribonuclease H |
| **RCA** | rolling circle amplification |
| **rpm** | revolutions per minute |
| **rt** | room temperature |
| **RuAAC** | ruthenium(II)-catalysed azide–alkyne cycloaddition |
| **s** | second(s)<br>singlet (spectra) |
| ***S*** | entropy<br>stereochemical descriptor |
| **S-PACE** | thiophosphonoacetate |
| **sat.** | saturated |
| **SATE** | *S*-acyl-2-thioethyl |
| **SEM** | standard error of the mean |
| **sgRNA** | single guide RNA |
| **SSO** | splice-switching oligonucleotide |
| ***syn*** | stereochemical descriptor |
| **t** | triplet (spectra) |
| **T** | thymidine<br>twist sugar pucker |
| ***T*** | temperature |
| **TBAF** | tetra-*n*-butylammonium fluoride |
| **TBDPS** | *tert*-butyldiphenylsilyl |
| **TBE** | tris base, boric acid, EDTA buffer |
| **TBS** | *tert*-butyldimethylsilyl |
| **TCA** | trichloroacetic acid |
| **TdT** | terminal deoxynucleotidyl transferase |

| | |
|---|---|
| **TEAB** | triethylammonium bicarbonate buffer |
| ***tert*** | tertiary |
| **TFA** | trifluoroacetic acid |
| **THF** | tetrahydrofuran |
| **Thy** | thymine |
| **TL** | triazole internucleoside linkage |
| **TLC** | thin layer chromatography |
| $T_m$ | melting temperature |
| **TMS** | trimethylsilyl |
| **TMSOTf** | trimethylsilyl trifluoromethanesulfonate |
| **tosyl** | 4-toluenesulfonyl (also Ts) |
| **trityl** | triphenylmethyl |
| **Tris** | tris (hydroxymethyl) aminomethane |
| **Ts** | 4-toluenesulfonyl (also tosyl) |
| **TTR** | transthyretin |
| **Tz** | triazole |
| **U** | uridine |
| **Ura** | uracil |
| **UTR** | untranslated region |
| **wtHTT** | wild-type huntingtin |
| ***Z*** | zusammen (configurational) |

# Abstract

Chemically modified oligonucleotides find myriad applications in therapeutics and synthetic biology because they can be utilised to modulate biophysical or therapeutic properties, install fluorophores for detection, or to allow orthogonal reactivity through additional functional groups. Whilst oligonucleotides can be principally modified anywhere at their nucleotide subunits, the presented work investigates modifications of the oligonucleotide backbone for therapeutic applications and synthetic biology.

Antisense oligonucleotides (ASOs) have become an emerging therapeutic approach to modulate protein expression at the mRNA level to combat hard-to-treat diseases. Stability against nucleolytic degradation, good target hybridisation properties, and efficient delivery into the cell are essential. The chemical modification of the hydrolysable, anionic phosphodiester backbone holds great potential for protection against phosphodiesterases, and to reduce the polyanionic charges for enhanced cellular uptake. In the presented work, two charge-neutral backbone surrogates—the carbamate and amide internucleoside linkages—were investigated for their ability to confer drug-like properties by replacing the natural phosphodiester backbone. Combinations with additional bicyclic sugar modifications and different designs of the ASO segments were explored. The antisense assessment criteria comprised their general hybridisation properties with

an RNA target, their stability against nucleolytic degradation, and their potential for enhanced cellular uptake.

Backbone-modified oligonucleotides are also versatile constructs for synthetic biology as exemplified by the construction of synthetic genetic polymers, ribozymes, aptamers, or gene editing systems. Amongst those, artificial genetic polymers can be applied for gene synthesis, sequencing, or the production of nanomaterials. A crucial element for backbone modifications in genetic templates is their compatibility to be read by polymerases. In this context, triazole-based internucleoside linkages have been extensively studied as promising replication compatible phosphodiester mimics. This thesis presents a novel triazole backbone structure and its assessment for replication compatibility in terms of continuation, speed, and correctness of the read-through by DNA polymerases.

# CHAPTER I

# INTRODUCTION

"*We are dealing with an entity sui generis not comparable to any hitherto known group.*" (Johannes Friedrich Miescher, 1871)



When Miescher reported a one-of-its-kind entity back in 1871,[1] the world did not know what this enigmatic substance might be.[2] 150 years and uncountable breakthroughs later, we are still dealing with a one-of-its-kind substance, but we found a name for it—nucleic acids. Many ground-breaking discoveries along the way have shed light on nucleic acids and shaped our modern understanding of their structure and function. We know that nucleic acids are polymeric biomacromolecules that hold a cell's genetic information in the form of a molecular code. Moreover, this code can be replicated, deciphered, and executed by the cell's biomolecular machinery, thereby providing the blueprint of life on our planet. All this genetic information is the sum of the underlying molecular structure of nucleic acids, which determines their cellular functions.

# 1. The Structure of Nucleic Acids

## 1.1. Primary Structure of Nucleic Acids

Nucleic acids are polymeric biomacromolecules that form long, unbranched strings of molecular subunits called nucleotides. Many of these subunits are repeatedly attached to each other to form an oligonucleotide (ON) (Figure I-1A). The nucleotide itself is the sum of a heterocyclic nucleobase, a sugar, and a phosphate; a nucleoside only comprises the sugar and the nucleobase. Nucleosides are attached to each other *via* phosphodiester (PO) linkages between the substituents of adjacent sugar subunits. The sugar part can be a 2′-deoxyribose or a ribose

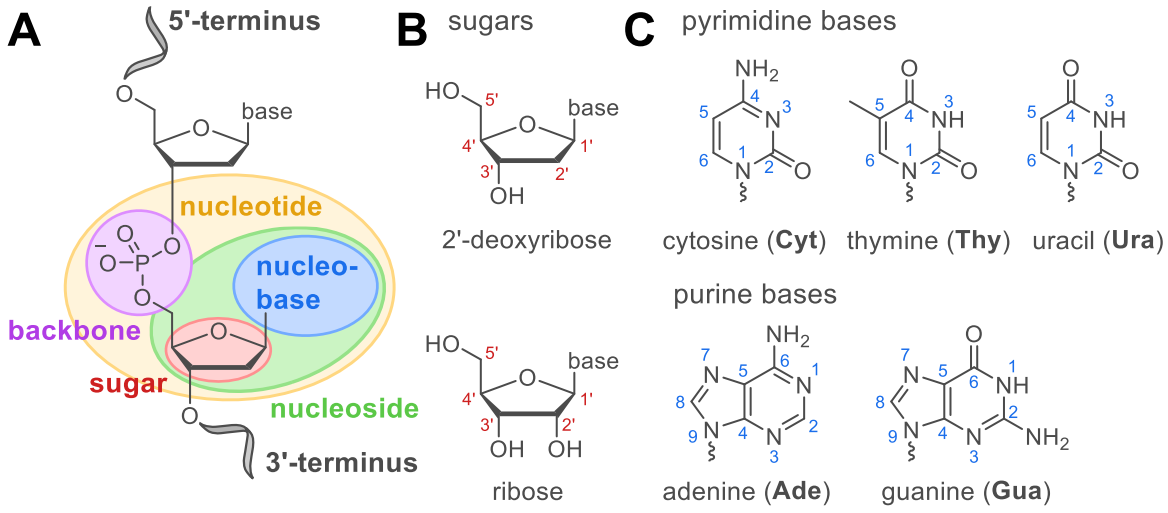



**Figure I-1: Overview of the molecular structure of nucleic acids.**
(**A**) The terminology of different chemical moieties within an oligonucleotide. The nucleobase, sugar and phosphate backbone are highlighted. Together, sugar and nucleobase make a nucleoside whilst sugar, nucleobase and phosphate make a nucleotide. The different ends of the oligonucleotide are labelled as 3′- and 5′-terminus. (**B**) The two sugars found in DNA (top) and RNA (bottom). The numbering of the sugar atoms is indicated. (**C**) Examples of pyrimidine and purine bases and their molecular structures. The numbering of the atoms is indicated.

(Figure I-1B), and determines whether the oligonucleotide is a 2′-deoxyonucleic acid (DNA) or a ribonucleic acid (RNA), respectively. In both cases, the phosphodiester linkages are formed between the 3′- and 5′-substituents of adjacent nucleosides. This asymmetry induces a directionality of the oligonucleotide resulting in a 3′- and a 5′-terminus (Figure I-1A, B). Ribo- and deoxyonucleosides have different nucleobases attached to the sugars (Figure I-1C). In nature, five different nucleobases are mainly found: adenine (Ade), cytosine (Cyt), guanine (Gua), thymine (Thy) and uracil (Ura). Deoxynucleosides of adenine, cytosine, guanine, or thymine are called deoxyadenosine (dA), deoxycytidine (dC), deoxyguanosine (dG), or thymidine (T)—the suffix ‘deoxy’ or ‘d’ can be omitted as thymine naturally only occurs in deoxynucleosides. Ribonucleosides of adenine, cytosine, guanine, or uracil are called adenosine (A), cytidine (C), guanosine (G), or uridine (U). Naturally, RNA does not contain any thymine nucleobases, which are replaced by uracil.

The order in which nucleosides are connected to each other composes an oligonucleotide’s code, also known as the oligonucleotide sequence. Here, the nucleobases form the distinctive chemical moieties, and the phosphate and sugar parts remain constant (Figure I-2). With four different nucleobases in DNA (Ade, Cyt, Gua and Thy) and RNA (Ade, Cyt, Gua and Ura), an oligonucleotide can be read as a four-letter code. Known as the central dogma of molecular biology, a DNA sequence can be transcribed into a messenger RNA (mRNA) sequence, which itself can be translated into a protein sequence.[3] Thus, DNA holds the genetic information—i.e., the blueprint—for cellular protein biosynthesis.

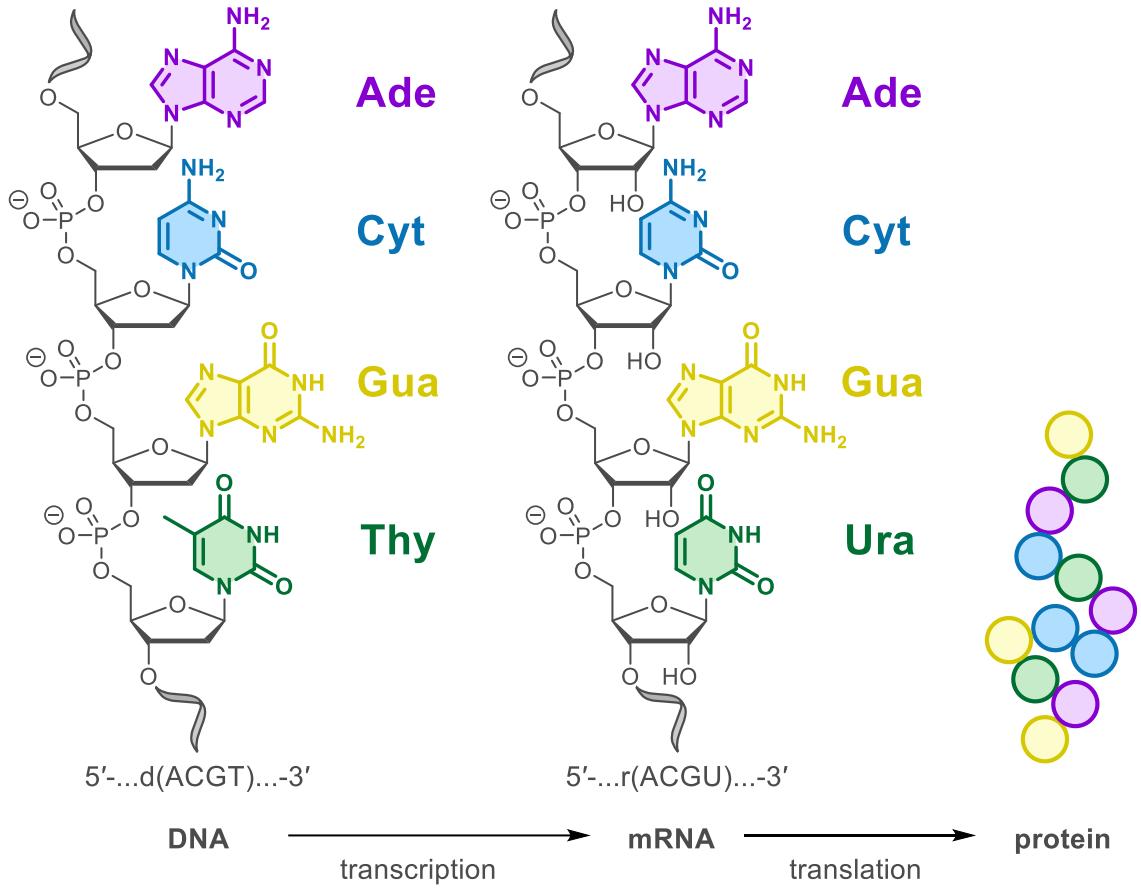



**Figure I-2: The central dogma of molecular biology.**
The molecular details of a DNA and an RNA oligonucleotide section are shown. The nucleobases Ade, Cyt, Gua and Thy/Ura are highlighted. The conventional four-letter string code is shown for the highlighted oligonucleotide sections. The protein is schematically represented as a string of pearls to represent its modular structure. Note that the colours of the protein subunits do not correspond to the individual nucleobases in the oligonucleotide structures.

The individual nucleotide subunits carry certain stereochemical and conformational characteristics, which determine an oligonucleotide's biophysical properties.[4] The nucleobases are attached to the sugar at the chiral, anomeric centre C1′. Thus, nucleosides can exist as stereoisomers at this anomeric centre, i.e., anomers (Figure I-3). Naturally, oligonucleotides only contain β-nucleosides; α-oligonucleosides are extremely rare.[5] It is therefore important to control the

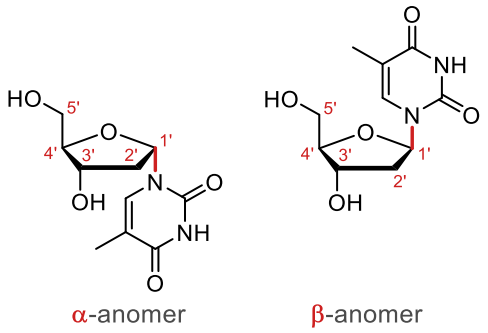



**Figure I-3: Sugar configurations in DNA and RNA.**
Configurational isomers at C1′ (anomeric centre) of the 2′-deoxyribose are exemplified for β- and α-T and the corresponding glycosidic bonds are highlighted. (**B**) Sugar puckers typically found in DNA (top row) and RNA (bottom row). The numbering of the sugar atoms is indicated.

formation of the glycosidic bond when chemically synthesising nucleosides, which will be discussed in Chapter III of this thesis.

Depending on the electronic properties of the sugar substituents, the pentose ring favours different conformations. In general, a five-membered ring does not adopt a flat conformation because of the resulting high torsional strain.[4] Instead, the torsional strain of the ribose is relieved by adopting conformations in which three or four atoms are coplanar, and one or two atoms deviate from the plane (Figure I-4). The sugar pucker can adopt an envelope (E) conformation in which only one atom is deviating from the plane that is formed by the remaining four atoms. Another conformation is the twist (T) sugar pucker in which two adjacent atoms deviate to opposite sides from the plane that is formed by the remaining three atoms. In (deoxy)ribose, atoms deviating from the plane to the same and opposite side of C5′ are called *endo* and *exo*, respectively. Different sugar puckers can be represented along a pseudorotation phase angle (*P*) cycle, which also shows that the puckers are

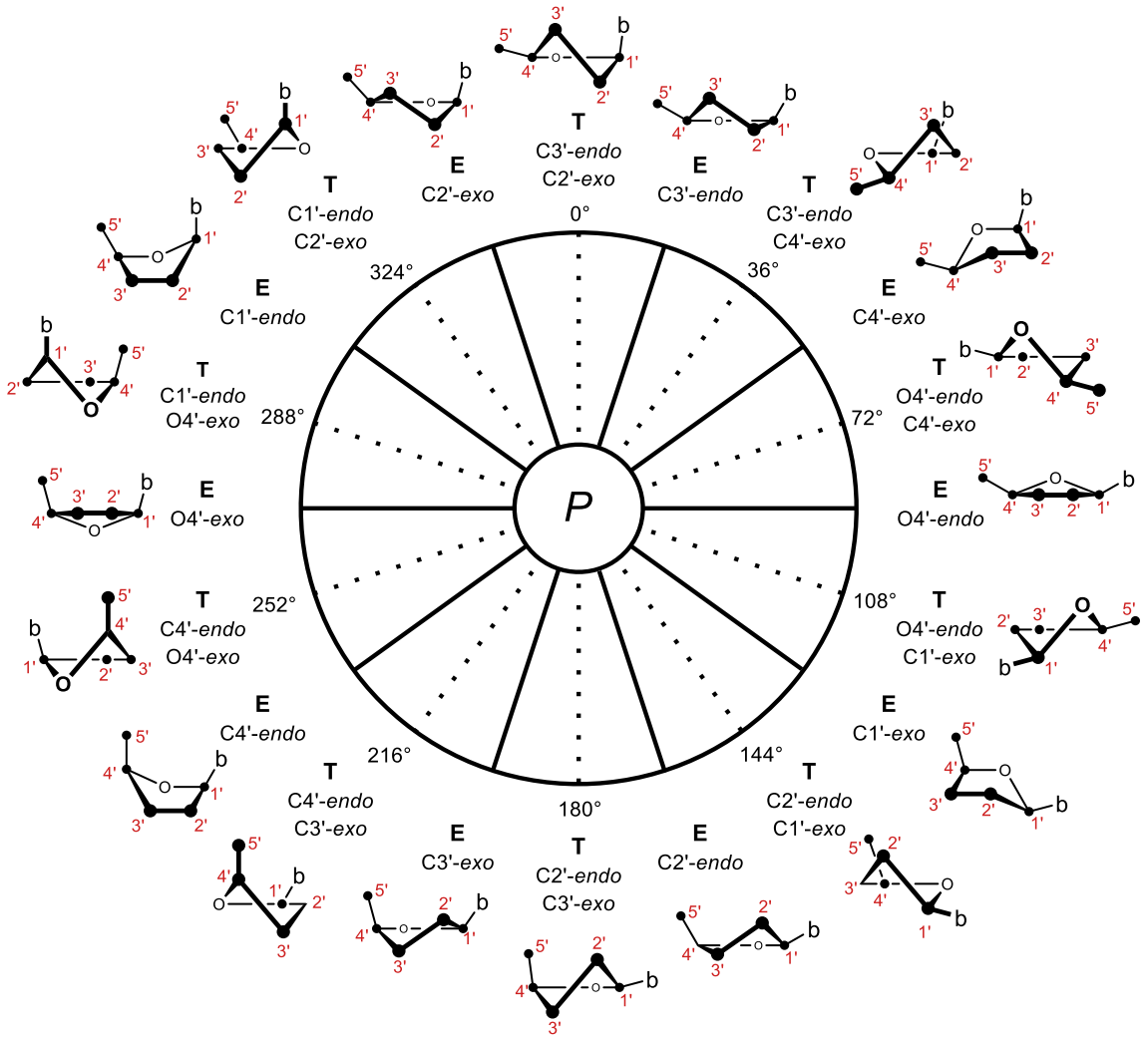



**Figure I-4: Sugar conformations in DNA and RNA.**
Sugar puckers represented as ball-and-stick structures along the pseudorotation phase angle (*P*) cycle. Twist and envelope puckers are marked with T and E, respectively. The numbering of the sugar atoms is indicated. Abbreviations: b, nucleobase.

not fixed but could continuously transition through the cycle giving rise to a theoretically infinite number of possible conformations.[4] However, due to the unsymmetrical substitution of the sugar ring in DNA and RNA, some puckers are favoured over others: On the nucleoside monomer level, pyrimidine-substituted ribonucleosides favour the C3′-*endo*, and purine-substituted ribonucleosides the C2′-*endo* conformation;[4] deoxynucleosides generally favour the C2′-*endo*

conformation.[4] However, those states should be considered as a preferential adoption of the respective sugar puckers (60–80%) in the equilibrium of the rapid transition between C2′- and C3′-*endo* conformations.[4] In fact, the energy barrier between the C2′- and C3′-*endo* conformations is low, and transition is fast—so fast that distinct puckers cannot be observed by NMR under standard conditions.[6] On the polymer level, the sugars in DNA preferentially adopt a C3′-*endo* or a C3′-*exo* conformation—a close variant of the C2′-*endo* conformation (see Figure I-4).[4] Whether the deoxyriboses of a DNA polymer adopt a C3′-*endo* or a C3′-*exo* conformation depends on the solvent and salt concentrations.[4] In RNA, the sugars prefer to adopt a C3′-*endo* conformation.[4] The preferences of adopting the C2′-*endo*/C3′-*exo* or the C3′-*endo* conformation are energetically driven. In detail, an axial orientation of electronegative substituents is favoured to minimise the overall dipole moment.[4] This can be utilised to favour a certain ribose conformation by introducing electronegative substituents.[7] One example is the introduction of a 2′-F substituent at a deoxyribose to favour the C3′-*endo* conformation.[8] Other strategies, such as introducing bicyclic systems to constrain the sugar pucker,[9] are relevant to the presented work and will be discussed in the corresponding paragraphs.

## 1.2. Secondary Structure of Nucleic Acids

DNA does generally not occur as a single strand but is hybridised to a second strand of DNA.[4] Together, the two strands form a double helical structure (left-hand side; Figure I-5). The orientation of the two strands is antiparallel with opposite termini of the individual oligonucleotides at each end of the duplex. Naturally, DNA

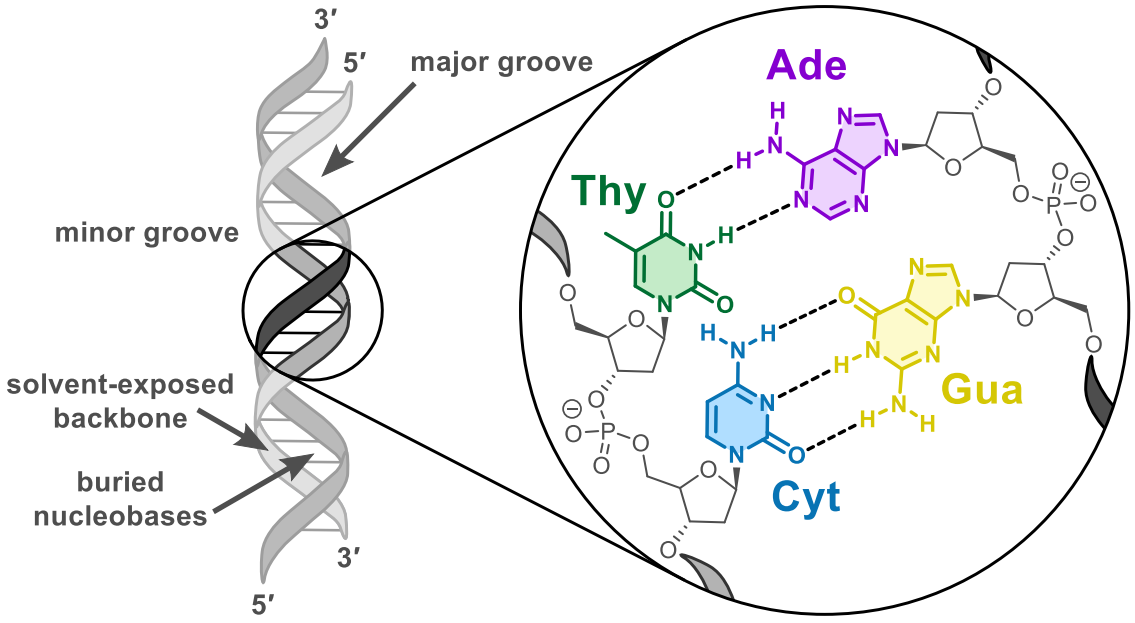



**Figure I-5: The Watson–Crick base pairing model.**
A DNA double helix is schematically represented at the left-hand side. The ribose/phosphate backbone is shown as a ribbon structure and the bases are represented by horizontal lines. Major and minor grooves are indicated. Hydrogen bonds between two complementary, antiparallel DNA strands are represented as dotted lines. Adenine, cytosine, guanine and thymine nucleobases are highlighted.

forms a right-handed duplex structure that contains a major and a minor groove. Several factors contribute to the hybridization of two oligonucleotide strands:

*The hydrophobic effect.* In water, hybridisation is partly driven by the hydrophobic effect.[4,10] Different to the negatively charged backbone, the nucleobases are hydrophobic and more difficult to dissolve in water. The water molecules occupying the hydrophobic surfaces of the nucleobases have limited possibilities of forming hydrogen bonds with neighbouring molecules. Thus, these water molecules become more organised and lose entropy. Upon hybridisation of two oligonucleotide strands, their solvent-exposed hydrophobic surface area decreases because the nucleobases are buried inside the double helix and the

polyanionic backbones are exposed to the solvent (left-hand side; Figure I-5). Along with a reduced hydrophobic surface area, more organised water molecules are liberated to gain a higher level of disorder (entropy gain). Although the entropy gain of the liberated water molecules partially counteracts the entropy loss of the two hybridising oligonucleotide strands, the overall hybridisation event remains entropically disfavoured, i.e., the change in entropy is negative. The change in Gibbs free energy is defined as $\Delta G$ in Equation 1, wherein $\Delta H$ is the change in enthalpy, $T$ is the temperature in Kelvin, and $\Delta S$ is the change in entropy.

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

Hybridisation is favourable when $\Delta G$ is negative. Thus, based on the overall negative $\Delta S$ for duplex formation, Equation 1 shows that hybridisation becomes less favourable with increasing temperature. Indeed, heating up a duplex favours its de-hybridisation, which is also referred to as denaturation or melting. The concept of heating up an oligonucleotide duplex to favour its denaturation will be important later in this section when introducing UV melting experiments.

*Nucleobase stacking.* Another factor contributing to the hybridisation of two oligonucleotide strands is nucleobase stacking—an exothermic process in which the nucleobases stack vertically.[4] Nucleobase stacking is driven by dipoles, dipole-induced dipole moments, and interaction of π-systems. As such, the nucleobases are permanent dipoles leading to electronic attractions of neighbouring nucleobases. Moreover, dipole-induced parallel dipole moments contribute to attraction of the polarisable nucleobases, which is an electronic effect called London

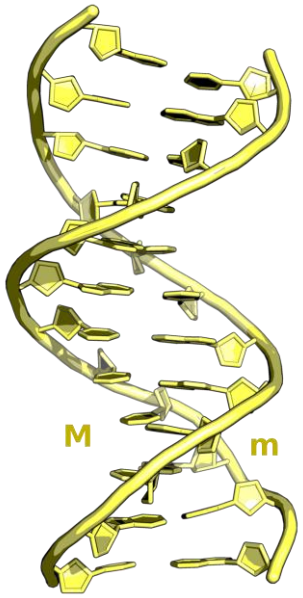
dispersion.[4,10] Lastly, nucleobases are aromatic and interactions of their π-systems of adjacent nucleobases are energetically favourable.[4] Because nucleobase stacking is more prevalent in double than in single stranded oligonucleotides, duplex formation is partly driven by the exothermic stacking of nucleobases.

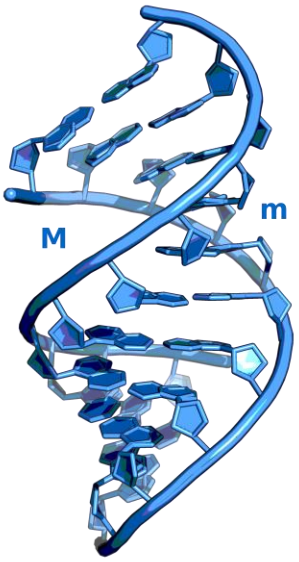
*Hydrogen bonding*. The nucleobases contain hydrogen bond donors and acceptors allowing the formation of interstrand hydrogen bonds. In water, hydrogen bonding between the nucleobases of two hybridised strands is relatively weak due to competing interactions with the solvent.[4] Nevertheless, interstrand hydrogen bonding is an important factor when describing duplex formation—especially with respect to the observed sequence specificity. Based on observations made by Erwin Chargaff,[11,12] Rosalind Franklin,[13,14] and Maurice Wilkins,[15] this hybridisation model was first proposed by Nobel laureates James Watson and Francis Crick and describes the hydrogen bonding interaction between adenine–thymine and cytosine–guanine base pairs. This model was subsequently coined 'Watson–Crick base pairing' (right-hand side; Figure I-5).[16] In detail, the adenine–thymine pair forms two, and the cytosine–guanine pair forms three hydrogen bonding interactions. The specificity of the Watson–Crick base pairing has profound consequences: (**i**) the hybridisation of oligonucleotide strands displays profound sequence specificity, and two oligonucleotides with matching sequences allowing Watson–Crick base pairing are called 'complementary'. Moreover, this specificity increases with the lengths of the oligonucleotides.[17] (**ii**) The design of a complementary oligonucleotide can be exceptionally rational when knowing the sequence of the target oligonucleotide. Indeed, the rationality of the Watson–Crick base pairing model is at the core of a

rapidly expanding therapeutic area, i.e., antisense oligonucleotides,[17] which will be introduced later in this chapter.

Whilst DNA mainly exists as a double strand, RNA usually occurs as a single-stranded biomolecule.[4] Despite its single-stranded nature, RNA can hybridise to a complementary oligonucleotide based on the same Watson–Crick base pairing principles as mentioned for DNA.[4] However, duplexes containing RNA (i.e., DNA:RNA or RNA:RNA duplexes) differ from DNA:DNA duplexes in their shape and biophysical properties (Table I-1).[4,18] A DNA:DNA duplex forms a right-handed double helix with a wide major and a narrow minor groove. One helical turn consists of 10.4 base pairs and has a pitch of 33.2 Å. In comparison, the right-handed RNA:RNA duplex is more compressed with 11 base pairs per helical turn and a helical pitch of only 24.6 Å. This results in a very deep and narrow major groove and a wide and shallow minor groove. The two different helical structures are often referred to as B-form (B-DNA; DNA:DNA duplex) or A-form (A-RNA; RNA:RNA duplex). Furthermore, the different duplex structures are characterised by having different sugar conformations: the C3′-*exo* conformation for the B-form DNA duplex, and the C3′-*endo* conformation for the A-form RNA duplex. As mentioned above (see Section 1.1), the preferred sugar pucker in DNA can be shifted towards the C3′-*endo* conformation with varying solvents and salt concentration, leading to the existence of A-form DNA.[4] Such helical structures can be evaluated by spectroscopic methods, e.g., circular dichroism (CD) spectroscopy,[19] which will be used throughout this thesis to assess the global structures of oligonucleotide duplexes. Lastly, the different duplex structures are correlated with different duplex

**Table I-1: Parameters describing an oligonucleotide double helix.**
The double helical structures of a DNA:DNA duplex (B-DNA; PDB, 1BNA) and an RNA:RNA duplex (A-RNA; PDB, 1QCU) are shown as yellow and blue cartoons, respectively. Abbreviations: bp, base pair; M, major groove; m, minor groove. *Table adapted from literature*.[4,18]




| | B-DNA | A-RNA |
|---|---|---|
| Orientation | right-handed | right-handed |
| Pitch/turn of helix | 33.2 Å | 24.6 Å |
| Base pairs per helix turn | 10.4 | 11 |
| Sugar pucker | C2′-*endo* | C3′-*endo* |
| Major groove (M) | Wide, deep | Narrow, very deep |
| Minor groove (m) | Narrow, deep | Very broad, shallow |

stabilities. As such, RNA:RNA duplexes are generally more stable than DNA:DNA duplexes.[20] Heteroduplexes (DNA:RNA) are less stable than RNA:RNA but can be more or less stable than DNA:DNA duplexes, depending on their sequence.[20,21] Different duplex stabilities can be evaluated by thermal melting experiments in which an oligonucleotide duplex is gradually heated up to determine the melting temperature ($T_m$)—the temperature at which half of the duplexes are denatured.[4] Such melting experiments will be used throughout this thesis to measure duplex stabilities.

# 2. Oligonucleotide Synthesis

*"From a Dream to Machine."*[d]

Nowadays, oligonucleotides are mainly produced by chemical synthesis through chemical bond formation between nucleosides. It is crucial to control the order in which the nucleotides are linked to each other because this will ultimately determine the oligonucleotide's sequence. Hence, single nucleosides must be added iteratively resulting in a theoretical minimum of n–1 reactions for an n-mer oligonucleotide. Traditional organic synthesis routines—i.e., purifying a product from reagents, starting material and side products after each synthetic step—would become unfeasible in this scenario. An answer to this problem is the concept of solid phase synthesis, which was introduced by Robert Bruce Merrifield and honoured with a Nobel Prize in 1984.[22,23] The growing oligonucleotide chain is immobilised to a solid support, typically a polystyrene or controlled pore glass (CPG) resin, and the reagents are delivered in solution. This results in several advantages: (**i**) reagents can be added in excess accelerating the reaction rates and resulting in near-quantitative yields; (**ii**) the product can be purified from left-over reagents by filtration; and (**iii**) the process can be automated. Several chemistries to form the phosphodiester bond have been explored[24–27] but the gold standard is the phosphoramidite chemistry,[28] for which oligonucleotide synthesis progresses in the

[d] Quotation: Prof. Andreas Marx, Universität Konstanz, 2015.

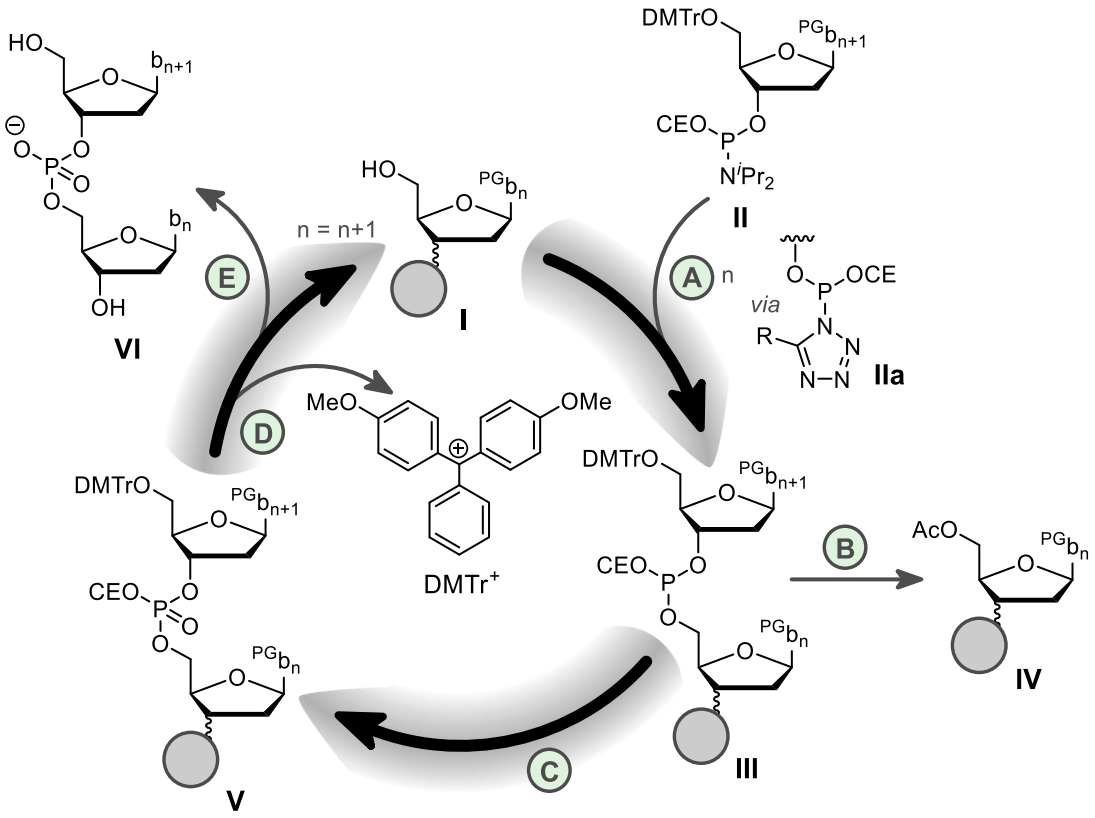



**Figure I-6: Schematic overview of the solid phase oligonucleotide synthesis cycle using CE phosphoramidites.**
Standard reagents and conditions: (**A**) **Activation and Coupling:** CE phosphoramidite **II** (0.1 M in MeCN), for R = SBn: 5-benzylthio-1H-tetrazole (BTT, 0.25 M in MeCN), rt, 30 s. (**B**) **Capping:** $Ac_2O$/pyridine/THF [1/1/8 (v/v/v)], *N*-methyl imidazole (NMI) [17.6% (w/v) in MeCN], rt, 30 s. (**C**) **Oxidation:** 0.015 M $I_2$ in $H_2O$/pyridine/THF [2/20/78 (v/v/v)], rt, 45 s. (**D**) **Detritylation:** trichloroacetic acid (TCA) [3% (w/v) in $CH_2Cl_2$], rt, 50 s. (**E**) a) **Cleavage:** concd aq $NH_3$, rt, 20 min; b) **Deprotection:** concd aq $NH_3$, 55 °C, 5 h. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; $^i$Pr, isopropyl; PG, protecting group. *Figure adapted from literature.*[30]

3′→5′ direction. A typical solid phase oligonucleotide synthesis cycle is depicted in Figure I-6 and involves: (**A**) activation and coupling of a protected nucleoside 2-cyanoethyl (CE) phosphoramidite to a resin-bound oligonucleotide; (**B**) capping of unreacted oligonucleotides; (**C**) oxidation of phosphorus(III) to phosphorus(V); (**D**) deprotection of the 5′-OH; and (**E**) cleavage of the oligonucleotide from the solid support and global deprotection. Each synthesis step will be discussed in the following paragraphs.

*(A) Activation and Coupling.* A standard nucleoside CE phosphoramidite **II** is rather unreactive and can be stable in acetonitrile for weeks.[29] In order to react with the free 5′-OH of the nascent oligonucleotide **I**, the phosphoramidite must be activated, which is typically achieved with tetrazole or a derivative thereof (Figure I-7). Protonation of the diisopropylamino group in phosphoramidite **II-H**$^+$ ($pK_a$ of 4.8 for tetrazole[29]) and successive nucleophilic displacement by the resulting tetrazolide results in the formation of a reactive phosphorotetrazolidite intermediate **IIa**, which readily reacts with the free 5′-OH of **I** to form phosphite triester **III**. This catalytic process greatly enhances the rate of the reaction and couplings can be easily performed within 30 seconds at room temperature.[30]

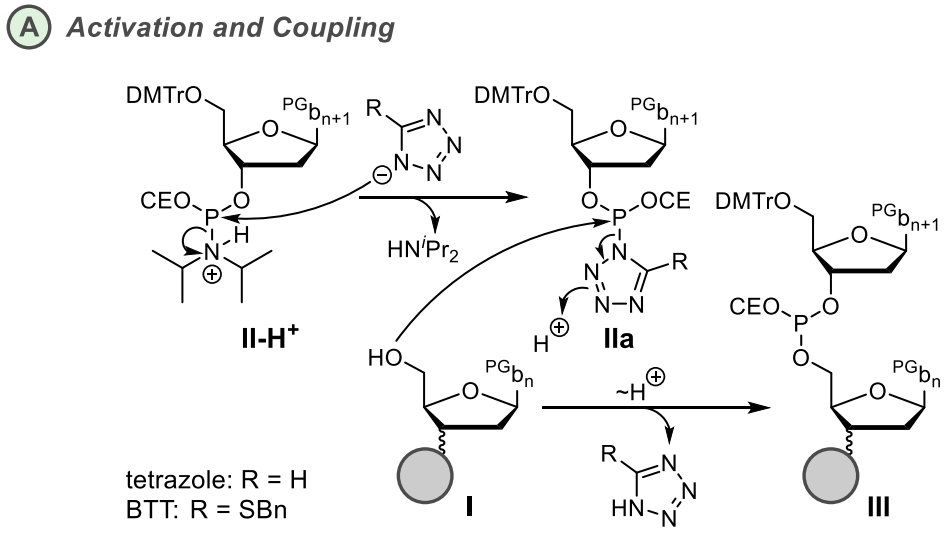



**Figure I-7: The mechanisms of the activation and coupling steps.**
Standard activators include tetrazole (R = H) or BTT (R = SBn). The solid support is represented as a grey circle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; $^i$Pr, isopropyl; PG, protecting group. *Figure adapted from literature.*[30]

***(B) Capping.*** Even though the coupling reaction is near to completion (up to ca. 99.5%),[30] the remaining 0.5% of unreacted **I** can become a substrate in the next synthesis cycle. If carried forward, this would lead to single deletion mutations, which are challenging to separate—if not practically impossible. Thus, a capping step is usually performed (Figure I-8). Acetic anhydride is used to acetylate the free 5′-OH of unreacted **I** to capped **IV** in the presence of *N*-methylimidazole as a catalyst. This will prevent addition of phosphoramidites in successive synthesis cycles. Capped oligonucleotide strands result in truncated sequences, which are easier to separate from the desired product. Note that the capping step can be performed before and after the oxidation step.[30]

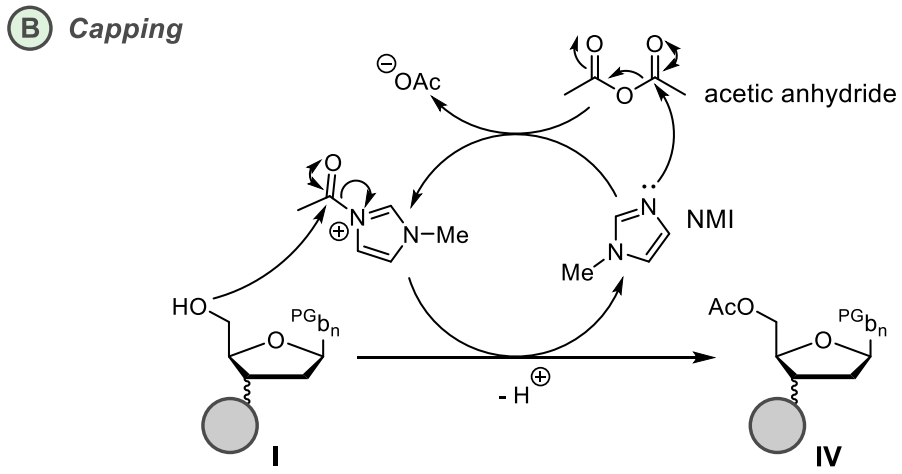



**Figure I-8: The mechanism of the capping step.**
The standard capping protocol uses NMI to catalyse the acetylation of the 5′-OH with acetic anhydride. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; PG, protecting group. *Figure adapted from literature.*[30]

*(C) Oxidation*. Phosphite triester **III** has a phosphorus(III) centre, which is unstable under the acidic conditions used in the deprotection step.[30] Thus, the phosphorus(III) is oxidised to phosphorus(V) using iodine and water in the presence of pyridine as a weak base (Figure I-9). The resulting phosphotriester **V** has the same oxidation state as the natural phosphodiester bond in oligonucleotides and can be easily deprotected to the phosphodiester form in the final cleavage and deprotection step.

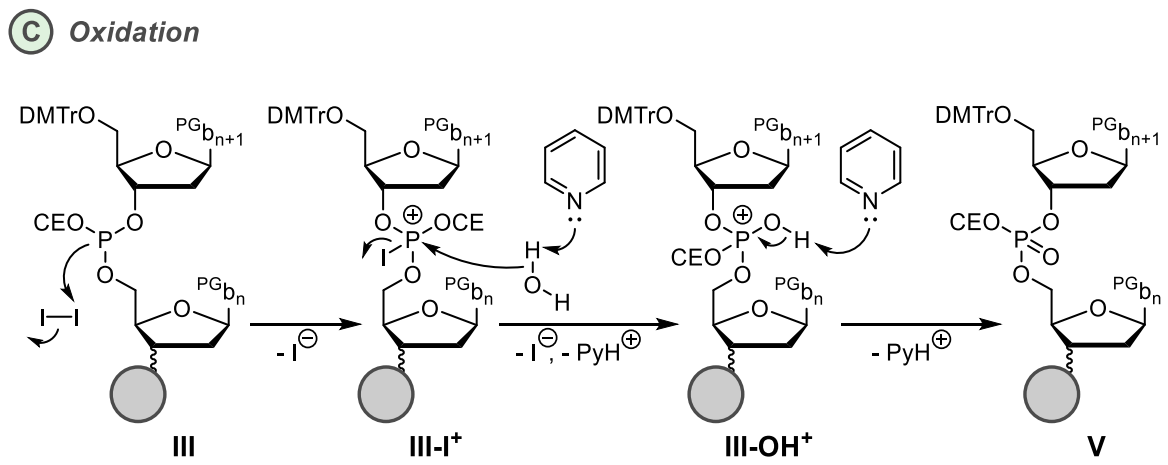



**Figure I-9: The mechanism of the oxidation step.**
The oxidation step is performed under basic aq conditions. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; PG, protecting group; Py, pyridine*. Figure adapted from literature.*[30]

***(D) Detritylation.*** The 5′-end of the most recently coupled nucleoside CE phosphoramidite is protected with a trityl group—e.g., 4,4′-dimethoxytrityl (DMTr)—to prevent the uncontrolled addition of multiple monomers. This trityl protecting group must be removed to release a free 5′-OH before coupling of any further nucleoside CE phosphoramidites. Detritylation is achieved by treatment with an acid—e.g., trichloroacetic acid—which protonates the 5′-oxygen of **V** whereby turning it into a leaving group and cleaving off a trityl carbocation (Figure I-10).[30] The trityl carbocation is stabilised by electron delocalisation from the three adjacent aromatic rings and by electron donation from the two methoxy substituents in the case of $DMTr^+$. Moreover, the $DMTr^+$ cation absorbs at 495 nm, allowing the detritylation step to be monitored. The detritylated, resin-bound oligonucleotide can either enter the next coupling cycle or be cleaved from the resin and deprotected.

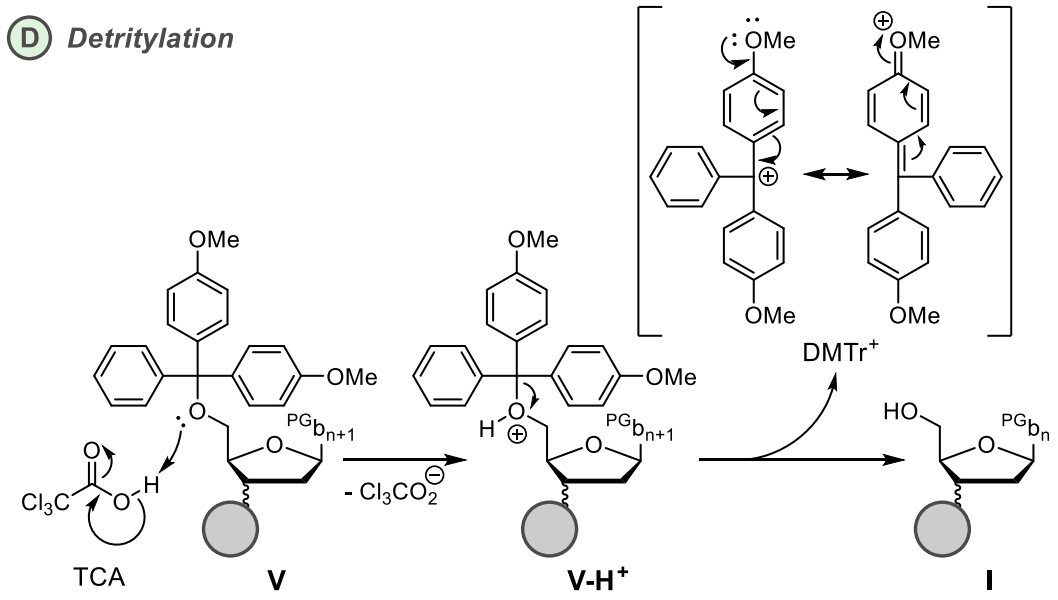



**Figure I-10: The mechanism of the detritylation step.**
Acid-catalysed detritylation is usually achieved with TCA. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; DMTr, 4,4′-dimethoxytrityl; PG, protecting group. *Figure adapted from literature.*[30]

Note that the final nucleoside of a sequence can be left DMTr-protected for purification purposes and is then detritylated after purification of the oligonucleotide.

***(E) Cleavage and Deprotection.*** The synthesised oligonucleotide is usually bound to the resin *via* a succinyl linker forming an ester bond with the 3′-oxygen of the oligonucleotide (Figure I-11).[30] Esters can be hydrolysed under basic conditions and treatment with a solution of concentrated aqueous ammonia releases the synthesised oligonucleotide from the resin within 20 minutes at room temperature. The CE protecting group of the phosphotriester backbone is removed under the same conditions in an $E1_{cb}$ mechanism releasing acrylonitrile. The nucleobases A, C and G are protected at their exocyclic amines, typically with an

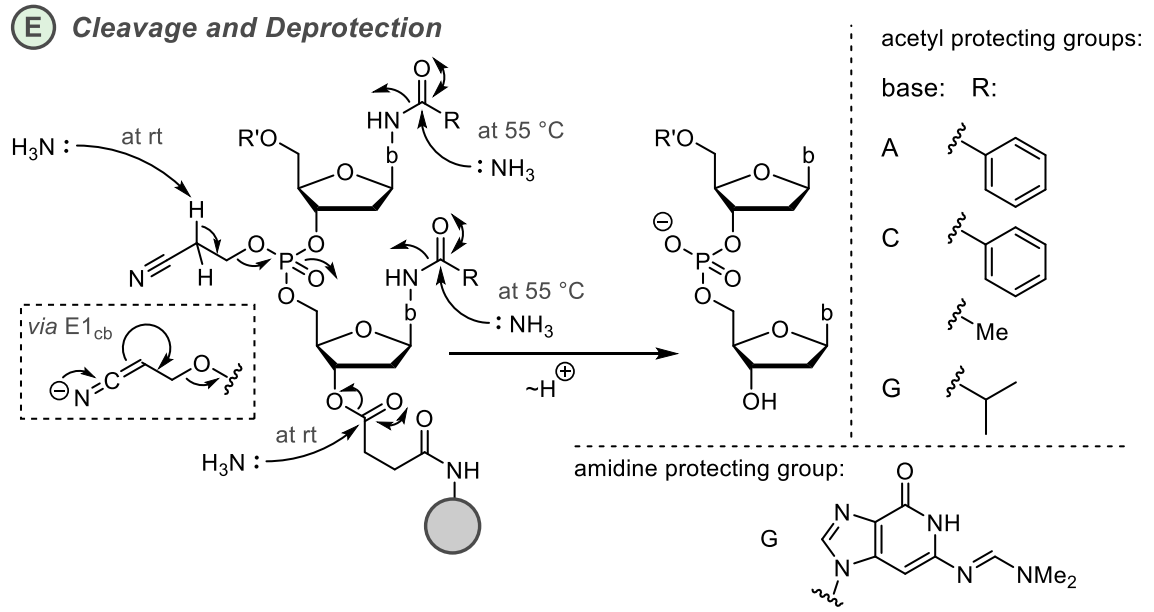



**Figure I-11: Mechanism of the cleavage and deprotection steps.**
Cleavage and deprotection are both carried out in concd aq $NH_3$ at rt or 55 °C, respectively. The standard nucleobase protecting groups are shown. Note that T is unprotected during solid phase synthesis. If a detritylation step was performed prior to the cleavage and deprotection: R′ = H. If no detritylation was performed before: R′ = DMTr. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; DMTr, 4,4′-dimethoxytrityl. *Figure adapted from literature.*[30]

acetyl (Ac, $dC^{Ac}$), benzoyl (Bz; $dA^{Bz}$, $dC^{Bz}$), isobutyryl (iBu; $dG^{iBu}$), or *N*,*N*-dimethylformamide (dmf; $dG^{dmf}$) group, to prevent any side reactions during the synthesis. Acetyl, benzoyl and isobutyryl protecting groups form amide bonds with the exocyclic amines and require harsher conditions to be removed. Incubation of the unbound oligonucleotide in concentrated aqueous ammonia at 55 °C will deprotect $dA^{Bz}$, $dC^{Ac}$ or $dC^{Bz}$ within 5 hours, whereas deprotection of $dG^{iBu}$ takes 16 hours.[30,31] Thus, $dG^{dmf}$ is often preferred which can be deprotected with concentrated aqueous ammonia at 55 °C within 5 hours.[31]

# 3. Modified Oligonucleotides

DNA is the blueprint of life on our planet—a blueprint with the capacity to store all the world's data in one room[32]—and yet, the chemical space of unmodified oligonucleotides can often be too narrow for certain applications in biotechnology and synthetic biology.[33] Moreover, oligonucleotides are products and substrates of naturally occurring processes which makes them susceptible to enzymatic degradation.[34] Consequently, there are many reasons to introduce a chemical modification into an oligonucleotide: (**i**) expansion of the genetic alphabet beyond A, C, G, and T;[35] (**ii**) introduction of chemical handles for ligation;[36,37] (**iii**) enhancement of hybridisation properties;[33] (**iv**) protection against degradation by nucleases,[33,34] etc. Indeed, the possibilities of utilising chemical modifications of oligonucleotides are so vast that a comprehensive discussion is beyond the scope of this introduction. Therefore, the following sections will introduce concepts and applications of oligonucleotide modifications that are relevant to the research chapters of this thesis. General modification sites will be illustrated and synthetic strategies for their incorporation into oligonucleotides will be described with a focus on backbone modifications. Moreover, the concepts of chemically modified therapeutic oligonucleotides and unnatural genetic polymers will be discussed to provide examples for possible applications in the antisense technology and synthetic biology.

## 3.1. Sites of Modifications

Chemical modifications have been reported for every part of a nucleotide and can be roughly categorised as nucleobase, sugar and backbone modifications (Figure I-12).[33] Furthermore, the two ends of an oligonucleotide (3′- and 5′-terminus) are potential modification sites. In nature, nucleobases are subject to modifications such as methylations (5-methyl cytosine ($^{Me}$C) or 6-methyl A), that serve, for example, as epigenetic markers.[38] More unnatural base modifications commonly aim to influence Watson–Crick base pairing properties,[39] or to introduce novel base pairs expanding the genetic code.[35] Sugar modifications often influence the sugar pucker, which is known to affect an oligonucleotide's duplex hybridisation properties.[40] The phosphate backbone attributes a polyanionic character to an oligonucleotide. Thus, modifications of the backbones eliminating or reverting the negative charges can substantially alter an oligonucleotide's physicochemical properties. Because the phosphodiester bond can be hydrolysed by nucleases, the

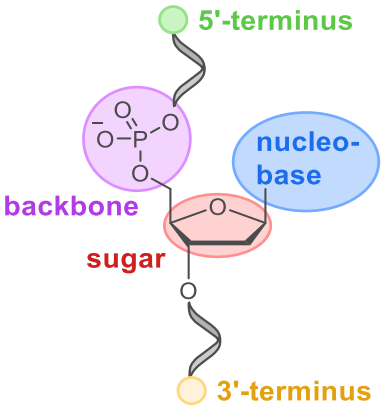



**Figure I-12: Modification sites.**
Oligonucleotides can be modified internally at the nucleobase, sugar or phosphodiester backbone as highlighted in the nucleotide structure. Moreover, oligonucleotides can be modified at either end of the oligonucleotide (termini) as indicated by the coloured circles.

backbone also resembles an attractive site for modifications to confer higher stability against nucleolytic degradation.[34] The termini are attractive sites to install chemical handles that can further undergo bioorthogonal ligation reactions,[36,37] or to incorporate fluorescent dyes.[30] The following two sections will discuss selected examples to illustrate general synthetic strategies to introduce nucleobase, sugar, backbone, and terminal modifications into an oligonucleotide.

## 3.2. Synthesis of Modified Oligonucleotides

The number of possible oligonucleotide modifications is immense. Although there are several examples utilising enzymatic processes to modify oligonucleotides,[41,42] the precise incorporation of a desired modification mainly relies on solid phase oligonucleotide synthesis. Its robustness and adaptability allow the incorporation of modified nucleosides during the synthesis cycles. The most common strategy makes use of modified CE phosphoramidites that already contain the desired modification. Many of them are commercially available with online catalogues listing more than 50 frequently requested '*common*' modifications.[43–46] Commercially available examples of nucleoside CE phosphoramidites that are used to introduce 2′-OMe, locked nucleic acid (LNA) or $^{Me}$C modifications are shown in Figure I-13. Both **1** and **2** are relevant to this thesis and have been used to synthesise 2′-OMe- and LNA-modified oligonucleotides. Structures **1**–**3** contain a CE phosphoramidite moiety and a DMTr-protected oxygen, facilitating internal strand modification by continuation of the oligonucleotide synthesis after coupling. In contrast, fluorescein CE phosphoramidite (FAM) **4** is an example for labelling the

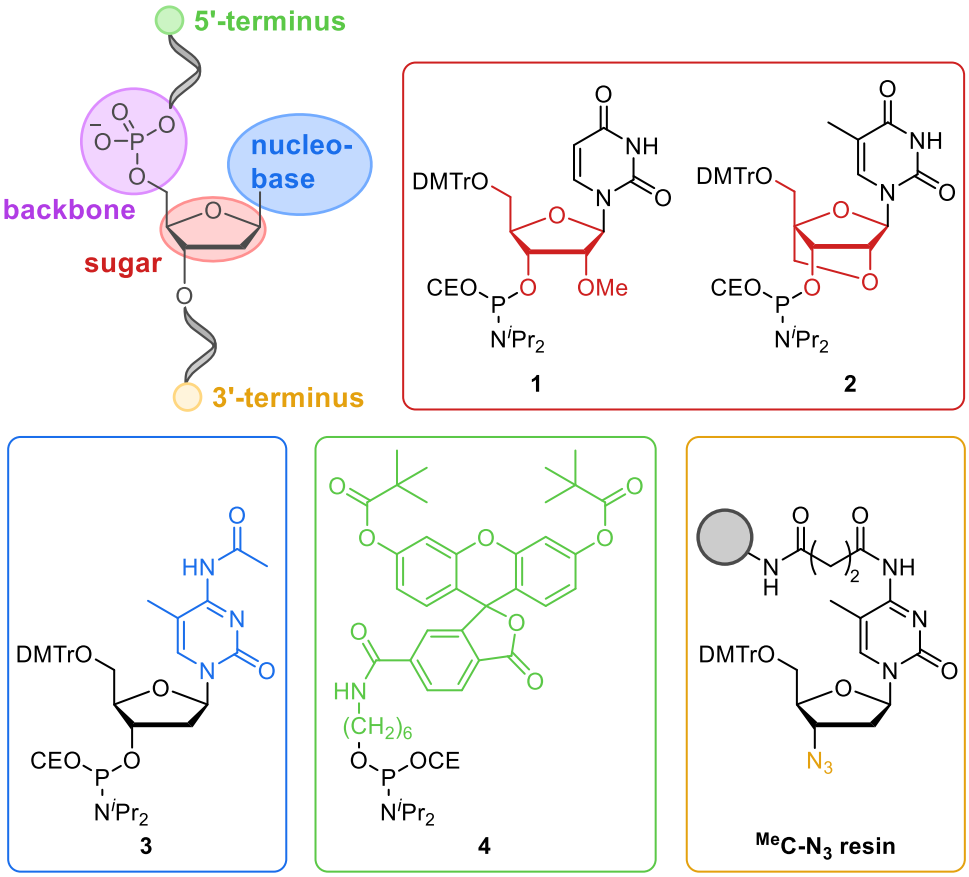



**Figure I-13: Selected structures for the introduction of chemical modifications during solid phase oligonucleotide synthesis.**
The moieties of the chemical structures corresponding to modifications of the sugar, nucleobase, 3′-terminus, and 5′-terminus are highlighted. The solid support is represented as a grey cycle. Abbreviations: CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; $^i$Pr, isopropyl.

5′-terminus of an oligonucleotide with fluorescein (FL). The absence of a DMTr-protected oxygen precludes continuation of the synthesis. The 3′→5′ direction of the solid phase oligonucleotide synthesis makes introduction of 3′-terminal modifications more complex. Although some solid supports with 3′-modification are commercially available,[47] modification of the 3′-terminus often requires the preparation of custom-made resins. One example is the immobilisation of 3′-azido $^{Me}$C to a CPG resin *via* the exocyclic amine of the base ($^{Me}$C-N$_3$ resin; Figure I-13).[48]

Solid phase synthesis using $^{Me}C\text{-}N_3$ resin will produce an oligonucleotide which is labelled with a 3′-azide.

Apart from direct incorporation during solid phase oligonucleotide synthesis, modifications can be introduced post-synthetically. Reasons for this approach can be the unavailability of a modified CE phosphoramidite or instability of the modification under conditions used during solid phase oligonucleotide synthesis. However, such post-synthetic modifications often require a chemical handle with orthogonal reactivity to other functional groups within the oligomer to ensure chemoselectivity of the post-synthetic labelling reaction. A typical example is the labelling of a 5′-$NH_2$-modified oligonucleotide with the active ester of a florescent dye.[30] Another example, the click ligation of azide- and alkyne-modified oligonucleotides, will be subject of Chapter V and discussed in the relevant section.

## 3.3. Synthesis of Backbone-Modified Oligonucleotides

### 3.3.1. Backbone Formation During Solid Phase Oligonucleotide Synthesis

The previous section explained how nucleobase, sugar and terminal modifications can be introduced into an oligonucleotide *via* modified CE phosphoramidites, resins, or post-synthetic labelling. In all cases, the general steps of the synthesis cycle essentially remain unchanged, and the modification is a result of the type of CE phosphoramidite or resin used. Introduction of artificial backbones is inherently more complicated. This is because solid phase oligonucleotide

synthesis is optimised for the reaction of a CE phosphoramidite with an alcohol, which will eventually form to the natural phosphodiester bond (see Section 2; Chapter I). Therefore, introduction of an artificial backbone requires different strategies, which can be roughly divided into the monomer and the dimer approach.

#### 3.3.1.1. The Monomer Approach

During the monomer approach, the artificial backbone is formed on-resin (Figure I-14A). One possibility to generate an artificial backbone is to alter the steps of the solid phase synthesis cycle itself whilst retaining alcohol **I** and CE phosphoramidite **II** as the reaction partners (Figure I-14B). For example, performing the oxidation step of the phosphite triester **III** with a sulphurising reagent[49] will give rise to a thiophosphotriester **IV**, which will eventually give the unnatural phosphorothioate (PS) backbone.[50] In this artificial backbone, one of the non-bridging oxygens of the phosphodiester bond is replaced by a sulphur. The straightforward integration of sulphurisation into the standard oligonucleotide synthesis protocol arguably advanced the phosphorothioate linkage to the most commonly used artificial backbone. However, replacement of a non-bridging oxygen by sulphur renders the phosphorothioate bond *P*-chiral. Although considerable progress has been made to form phosphorothioates stereospecifically,[51] these developments have not yet fully progressed to being a standard procedure in solid phase oligonucleotide synthesis.

Further to alteration of the synthesis cycle, alternative reaction partners can be employed to install an unnatural backbone. A commercially available example is the

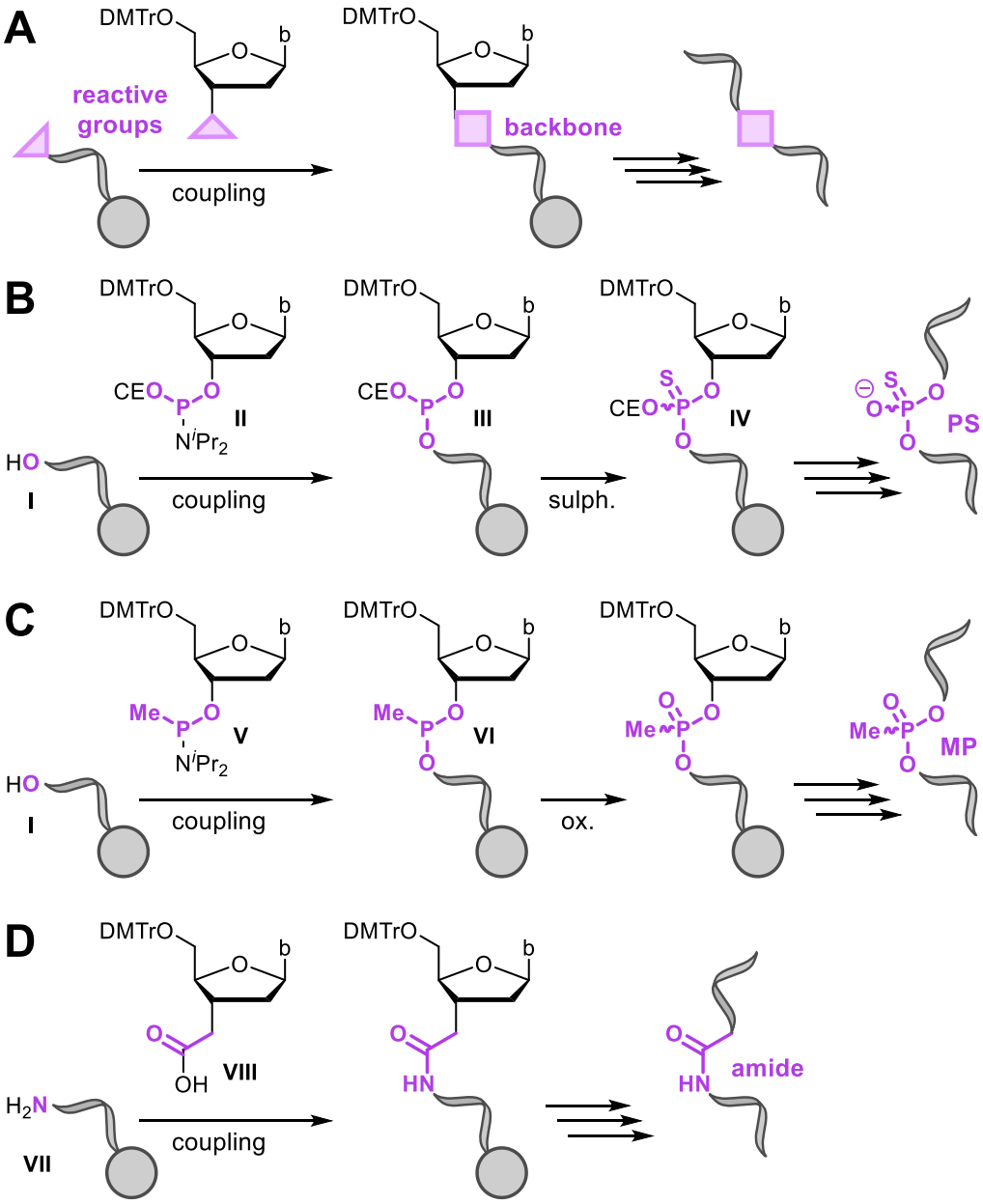



**Figure I-14: The on-resin monomer approach.**
(**A**) Schematic representation of the on-resin formation of the backbone linkage (square) by two reactive groups (triangles). (**B**) Formation of a PS backbone: Coupling is performed using standard CE phosphoramidite coupling protocols. Oxidative sulphurisation (sulph.) is water-free and makes use of a sulphurising reagent instead of the standard $I_2$-mediated oxidation under aq conditions.[49,50] (**C**) Formation of an MP backbone: Coupling of a methyl phosphonamidite **V** as an alternative reactive moiety. Oxidation (ox.) follows standard protocols. Note that conditions for deprotection and purification must be adapted.[52] (**D**) Formation of an amide backbone: The amide bond is formed between the resin-bound 5′-$NH_2$ in **VII** and the carboxylic acid in **VIII** using peptide coupling conditions. No oxidation is needed for the amide backbone.[55]
The solid support is represented as a grey cycle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethyxytrityl; $^i$Pr, isopropyl.

reaction of alcohol **I** with a methyl phosphonamidite **V**[52] (Figure I-14C)—as

opposed to the conventional CE phosphoramidite **II**. Coupling of a methyl phosphonamidite **V** results in methyl phosphonite diester **VI**, which can be oxidised to the methyl phosphonate (MP) backbone.[53,54] Note that due to certain chemical requirements of this modification, different deprotection and purification protocols must be followed.[52]

Both the phosphorothioate and methyl phosphonate linkages are phosphorus-based backbone modifications and closely resemble the natural phosphodiester linkage. However, the structural diversity of modified backbones can be much higher. One example is the amide bond which has been reported as an artificial backbone (Figure I-14D).[55] Because of its structural difference, on-resin formation of the amide backbone requires different reaction partners (i.e., carboxylic acid and amine) and different coupling reaction conditions (i.e., peptide coupling). Indeed, on-resin amide bond formation has been reported for the reaction between a resin-bound 5′-$NH_2$ oligonucleotide **VII** and a 3′-carboxymethyl nucleoside **VIII**.[55] When combined with standard solid phase oligonucleotide synthesis, such 'exotic' backbone formations must be compatible with the standard reaction conditions and synthesis equipment.

Note that on-resin formation during solid phase oligonucleotide synthesis has been reported for many more artificial backbones,[56–61] but those fall outside the scope of this introduction.

### 3.3.1.2. The Dimer Approach

Whenever the monomer approach is deemed futile, the dimer approach can offer an alternative solution to incorporate a desired backbone modification. During the dimer approach, the artificial backbone is introduced *en bloc* as part of a dinucleoside CE phosphoramidite (Figure I-15A). This allows the formation of the artificial backbone under conditions that might not be compatible with the standard oligonucleotide synthesis protocol or equipment (e.g., solvents, temperature, pH,

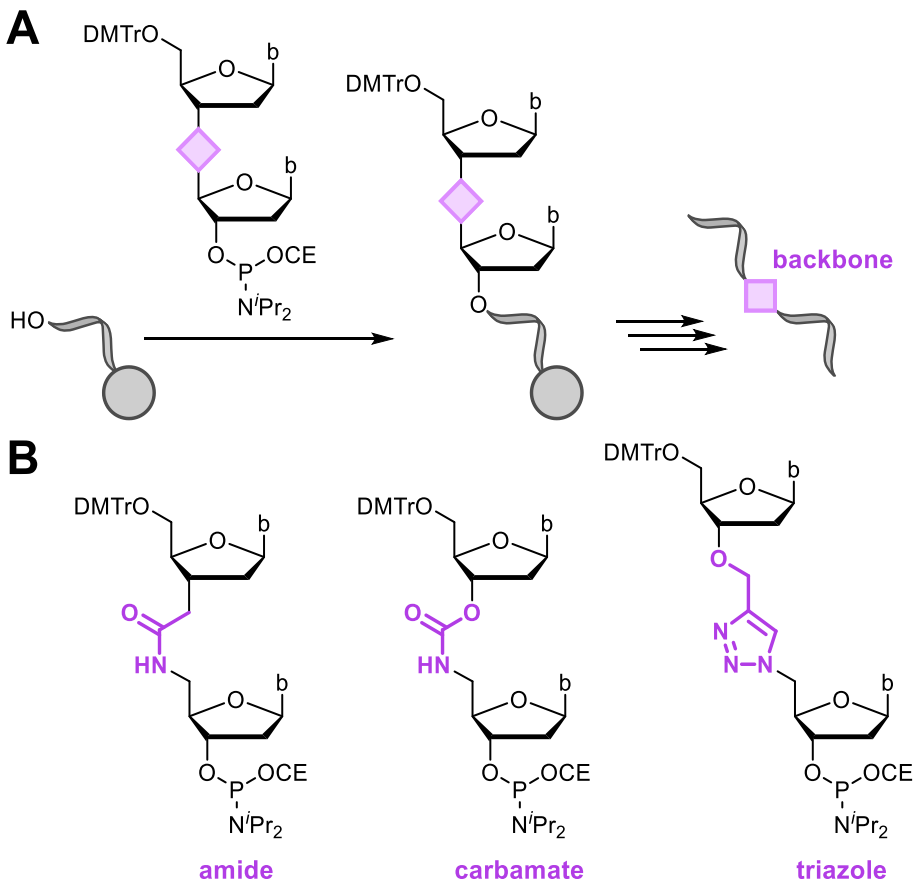



**Figure I-15: The *en bloc* dimer approach.**
(**A**) Schematic representation of the *en bloc* introduction of an artificial backbone (square) by coupling of a dinucleoside CE phosphoramidite containing the modified backbone linkage. (**B**) Examples of dinucleoside CE phosphoramidites containing an amide,[62,67] carbamate,[69] or triazole[73] as a modified internucleoside linkage. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethyxytrityl; $^i$Pr, isopropyl.

etc). Literature-reported unnatural backbones in dinucleoside CE phosphoramidites include amides,[62–68] carbamates,[69] triazoles[48,70–73] and many more[56,74–76] (Figure I-15B).

### 3.3.2. Post-Synthetic Artificial Backbone Formation

One major limitation of solid phase oligonucleotide synthesis is the accessible length of the oligomers[77] (up to 200 bases for DNA[78] and up to 120 bases for RNA[79] oligonucleotides). Many applications in biotechnology and synthetic biology require the preparation of significantly longer oligonucleotides, which deems solid phase oligonucleotide synthesis unsuitable for these purposes. Thus, the introduction of modified backbones for long (>200 base pairs) oligonucleotides must rely on a different approach. Indeed, long oligomers can be assembled by chemical ligation of shorter, chemically modified oligonucleotides with reactive terminal groups (Figure I-16A). Whilst the chemically modified short oligonucleotides can still be produced by solid phase synthesis, the modified backbone is generated post-synthetically by chemical ligation resulting in a longer, backbone-modified oligonucleotide. A splint oligonucleotide which is complementary to the reactive ends of the modified oligonucleotides is often used to bring the reactive groups together and facilitate templated ligation. However, one limitation of this approach is the required compatibility/orthogonality with other functional groups in the oligonucleotides. Moreover, the ligation must proceed under reaction conditions in which the oligonucleotides are soluble and stable. Literature-reported ligation reactions that fulfil these requirements include the formation of a

phosphorothioester,[80] phosphoramidate,[81–84] amide,[83] squaramide,[85] urea[85] or triazoles[83,86] as artificial backbones (Figure I-16B).

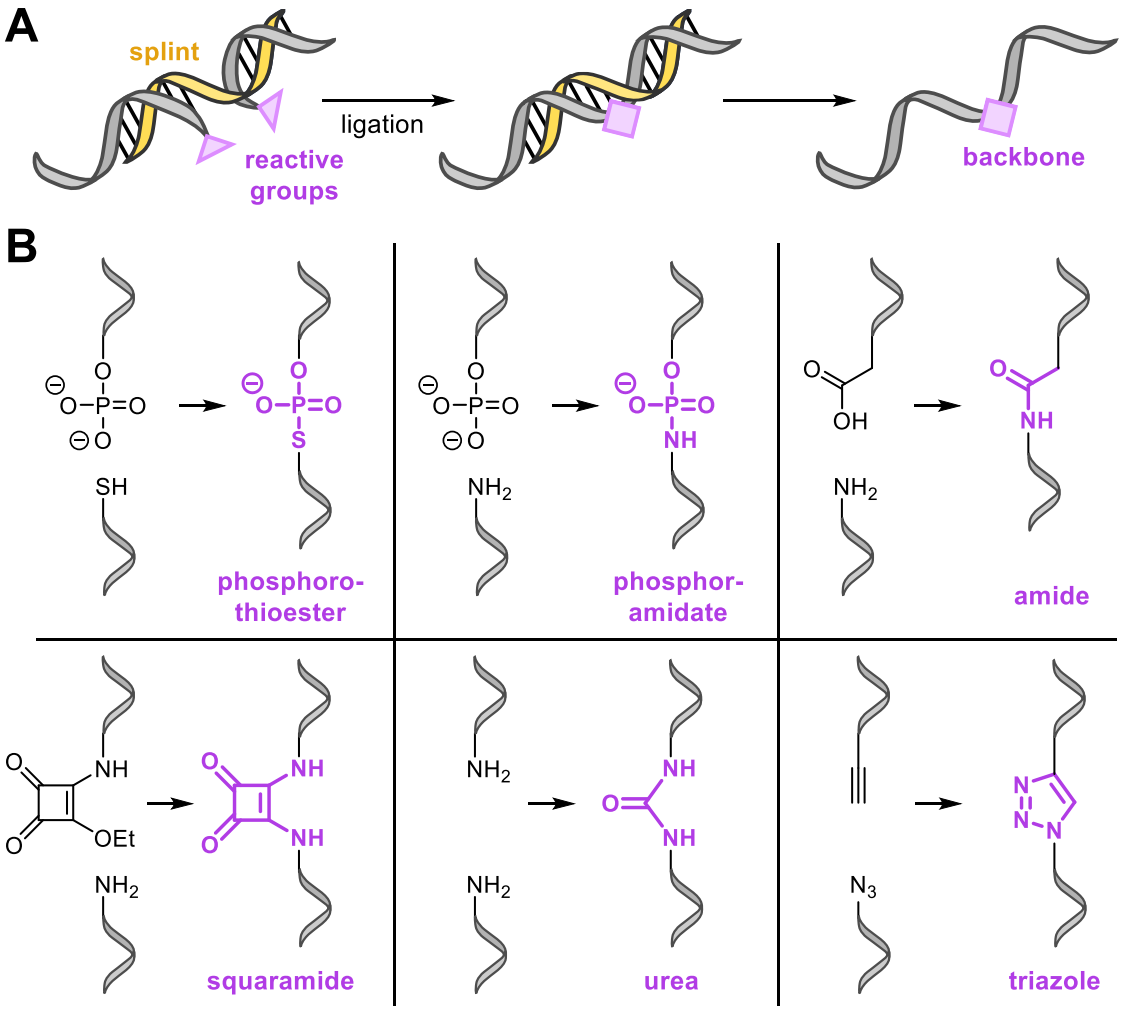



**Figure I-16: Post-synthetic formation of modified backbones.**
(**A**) A schematic representation of the general strategy to form backbone-modified oligonucleotides by ligation. Reactive ends of shorter oligonucleotides (triangles) can be ligated to form the artificial backbone (square). The splint oligonucleotide is shown in yellow. (**B**) Literature-reported[80–86] examples of artificial backbones that can be produced by (splint-mediated) ligation.

# 4. Antisense Oligonucleotides

## 4.1. Introduction to Antisense Oligonucleotides

All **anti-sense** oligonucleotides (ASOs) have in common that they are complementary (i.e., **anti**) to an RNA strand (i.e., **sense** strand) to modulate protein expression. Thus, ASOs ensure target specificity through Watson–Crick base pairing which makes ASO drug design exceptionally rational compared to more traditional medicinal chemistry approaches.[87,88] In theory, any protein can be targeted by designing an ASO sequence which is complementary to the mRNA of the selected target protein (see the central dogma of molecular biology; Figure I-2). ASOs are typically short (~20-mers) but long enough to ensure complementarity to a unique sequence that is not found in other mRNAs (off-targets).[17] Moreover, protein modulation at the mRNA level opens the possibility to target proteins that are considered undruggable through conventional approaches. Thus, ASOs hold the potential to combat hard-to-treat diseases such as cancers or genetic disorders. Unfortunately, the past has shown that drug development is not as straightforward as depicted here. Even though the idea of therapeutic ASOs intrigued researchers for decades, only a few ASO-based drugs have made it to the clinic.[89] The low success rate for ASO drugs can be pinned down to several unexpected obstacles that must be overcome: (**i**) RNA is a 3D-structured macromolecule and some sections of the chosen mRNA target might be buried inside the 3D structure, and not accessible for the ASO. Thus, determination of a suitable target sequence to ensure high

antisense activity is not always straightforward.[90] (**ii**) Unmodified oligonucleotides are normally rapidly degraded by nucleases in the serum. Typically, ASOs must be chemically modified to protect them from degradation and to ensure optimal serum lifetime.[91,92] However, chemical modifications of ASOs come with their own challenges and convolute ASO drug design. (**iii**) Chemical modifications can add toxicity or immunogenicity to ASOs.[93] (**iv**) Some chemical modifications are chiral but their introduction is not stereocontrolled. Iterative introduction without stereocontrol of such chiral modifications generates millions of different ASO stereoisomers, all having potentially different biophysical or therapeutic properties.[94] (**v**) Some modifications prevent the activation of mechanisms that regulate protein expression.[93] This is, for example, the case for many modifications at the 2′-position which prevent the activation of certain ribonucleases to degrade the mRNA target. A more detailed explanation for this phenomenon will be provided in Chapter IV. (**vi**) The typical size and polyanionic character of oligonucleotides impedes their cellular uptake and tissue distribution.[93] Advanced formulation, conjugation and molecular design strategies are usually applied to facilitate cellular uptake. Indeed, the delivery of macromolecules such as ASOs remains a major challenge for modern medicine.

## 4.2. Mechanisms of Antisense Oligonucleotides

ASOs can modulate protein expression *via* different mechanisms, and it is important to understand these for effective ASO drug design. Indeed, the molecular

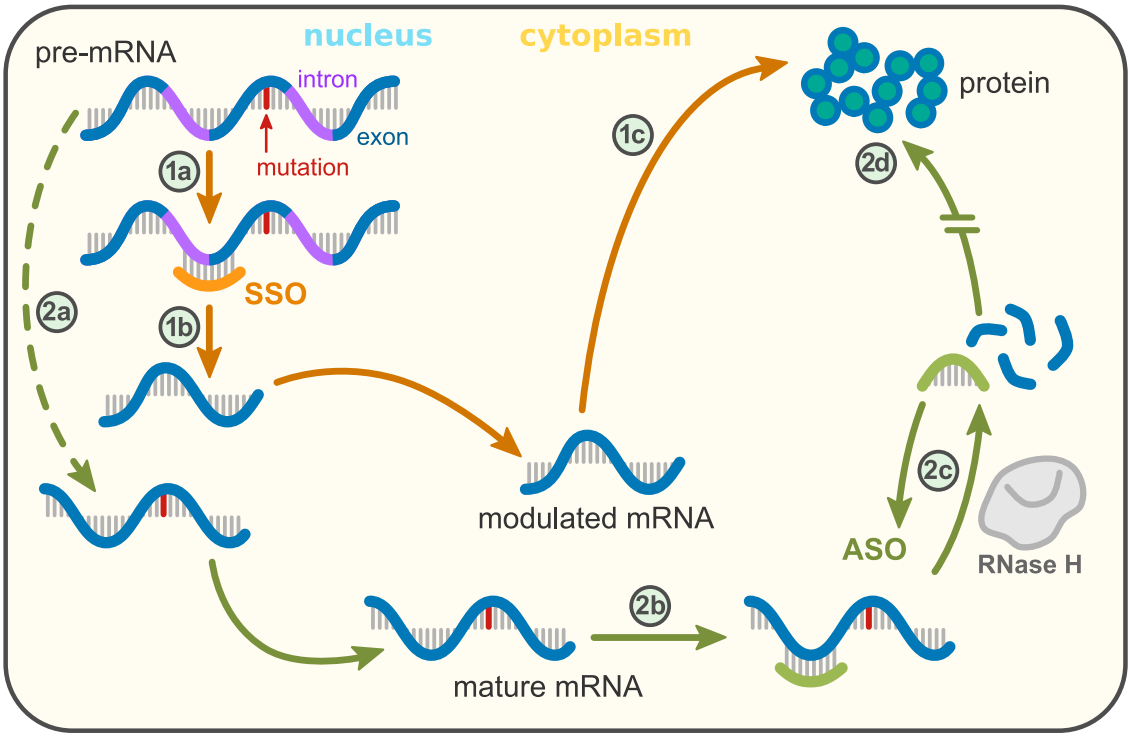



**Figure I-17: Overview of antisense mechanisms.**
**Splice-switching**: (**1a**) Binding of a splice-switching oligonucleotide (SSO) to its RNA target. (**1b**) Modulated splicing leads to the exclusion of a targeted exon (exon skipping). (**1c**) The modulated mRNA is translated into an alternative splice isoform of the target protein.
**Activation of RNase H**: (**2a**) Unmodulated maturation of mutated mRNA and export into the cytoplasm. (**2b**) Binding of ASO to its mRNA target. (**2c**) Recruitment of RNase H which degrades the mRNA and releases the ASO for the next catalytic cycle. (**2d**) Degradation of the mRNA prevents protein expression (knockdown). Note that RNase H can degrade the RNA target inside the nucleus or in the cytoplasm.[99]

design of an ASO determines its mode of action. A summary of the two main antisense mechanisms that are currently found in the clinic is given in Figure I-17.

The first example (1a–1c; Figure I-17) describes the modulation of mRNA maturation, which leads to the expression of a different protein splicing isoform.[95] This approach takes advantage of the fact that transcribed RNA undergoes a maturation process before being translated into a protein. Part of this maturation process is the splicing, which involves the exclusion of untranslated RNA sequences

(introns) and ligation of RNA sequences to be translated (exons). If an exon carried a mutation causing the expression of a malfunctional protein, the entire exon could be excluded (i.e., skipped) from the mature mRNA.[96] This can produce a splice variant of the protein that might be different/shorter but still functional. ASOs that interfere with the splicing process are called splice-switching oligonucleotides (SSOs). For example, splice-switching is a therapeutic approach for the treatment of Duchenne muscular dystrophy (DMD).[95] Around 14% of DMD patients have a mutation in exon 51 which leads to the expression of an unfunctional dystrophin protein.[97] Eteplirsen (or Exondys 51$^{TM}$) is a clinically approved SSO targeting exon 51 to produce a shorter but functional splice variant of the dystrophin protein by exclusion of mutated exon 51.[98]

The second antisense approach is activation of ribonuclease H (RNase H), which is an intracellular enzyme that can recognise a DNA:RNA heteroduplex, and selectively degrade the RNA strand.[99–101] This endogenous mechanism can be utilised by antisense oligonucleotides that function as the DNA part in the DNA:RNA heteroduplex to trigger RNase H-mediated degradation of an RNA, e.g., the mRNA which codes for a therapeutically relevant target protein (2a–2d; Figure I-17).[102,103] The activation of RNase H does not generate an alternative protein isoform, but leads to the degradation of its mRNA and thus to the downregulation of protein expression (i.e., knockdown). Degradation of an mRNA can be useful if it carries a mutation leading to the expression of an altered but unessential protein with undesired effects (e.g., oncogenic, prone to aggregation, etc). For instance, transthyretin (TTR) is an alternative, non-essential carrier protein

for thyroid hormones and retinol.[104–106] Mutations in the *TTR* gene can lead to the expression of a mutated TTR form which aggregates due to misfolding of the protein. The resulting aggregates are cytotoxic and cause familial transthyretin amyloidosis.[107] Inotersen (or Tagsedi™) is a clinically approved, RNase H-activating ASO that targets the *TTR* mRNA to prevent its expression and aggregation, thereby treating familial transthyretin amyloidosis.[108] Note that the RNase H-activating ASO does not necessarily have to hybridise to a mutated site of the mRNA target (2b; Figure I-17). In fact, Inotersen hybridises to the 3′-untranslated region (3′UTR) of the *TTR* mRNA coding for TTR, thereby inhibiting the expression of wild-type and mutated TTR.[109,110] However, simultaneous knockdown of wild-type and mutated protein can be problematic if the wild-type form performs important cellular functions whilst the mutated version—originating from one mutated allele—causes a disease phenotype. For example, Huntington's disease leads to neurodegeneration caused by a mutated *huntingtin* gene (*mHTT*) with an increased number of CAG trinucleotide repeats.[111] Therefore, knockdown of *mHTT* represents a promising therapeutic approach, but the simultaneous knockdown of wild-type *huntingtin* (*wtHTT*) raises concerns surrounding the long-term impact of reduced *wtHTT* levels.[112,113] Several strategies have been explored to selectively or disproportionally target *mHTT*, e.g., by targeting the CAG tract or single nucleotide polymorphisms that are often found in the mutated gene.[111,114] However, such allele-selective approaches are harder to deploy, and two such candidates have just recently failed clinical trials.[115]

Next to splice-switching and RNase H activation, RNA interference (RNAi) is a rapidly evolving antisense technology. Indeed, RNAi is now a dominant antisense approach in clinical studies and some RNAi-based drugs have already been approved.[116] Of those, the recent approval of Inclisiran (or Leqvio™)[117] is widely perceived as a game changer in the therapeutic oligonucleotides field treating a more common disease with a large number of patients that could benefit from this treatment.[118] Such RNAi therapeutics are administered as double-stranded RNA-type ASOs to trigger an alternative antisense mechanism inside the cell.[119,120] Its discussion, however, is beyond the scope of this introduction.

## 4.3. Chemical Modifications of Antisense Oligonucleotides

As mentioned above, therapeutic oligonucleotides must be chemically modified to enhance their enzymatic stability, target engagement, and delivery properties.[34,91,92] From the plethora of possible oligonucleotide modifications,[40] only a limited number is found in clinically approved therapeutics (Figure I-18). The phosphorothioate backbone (1st generation) is the earliest and most commonly adopted backbone modification, with the first approval of a PS-modified drug in 1998 (Table I-2).[121] The phosphorothioate backbone is compatible with RNase H[94] and stabilises PS-modified oligonucleotides against nucleolytic degradation.[34,122] Moreover, PS–ASOs show improved cellular uptake and tissue distribution properties.[123] The protein binding properties of the phosphorothioate linkage protect PS–ASOs from being cleared too rapidly from the blood, and aid cellular uptake

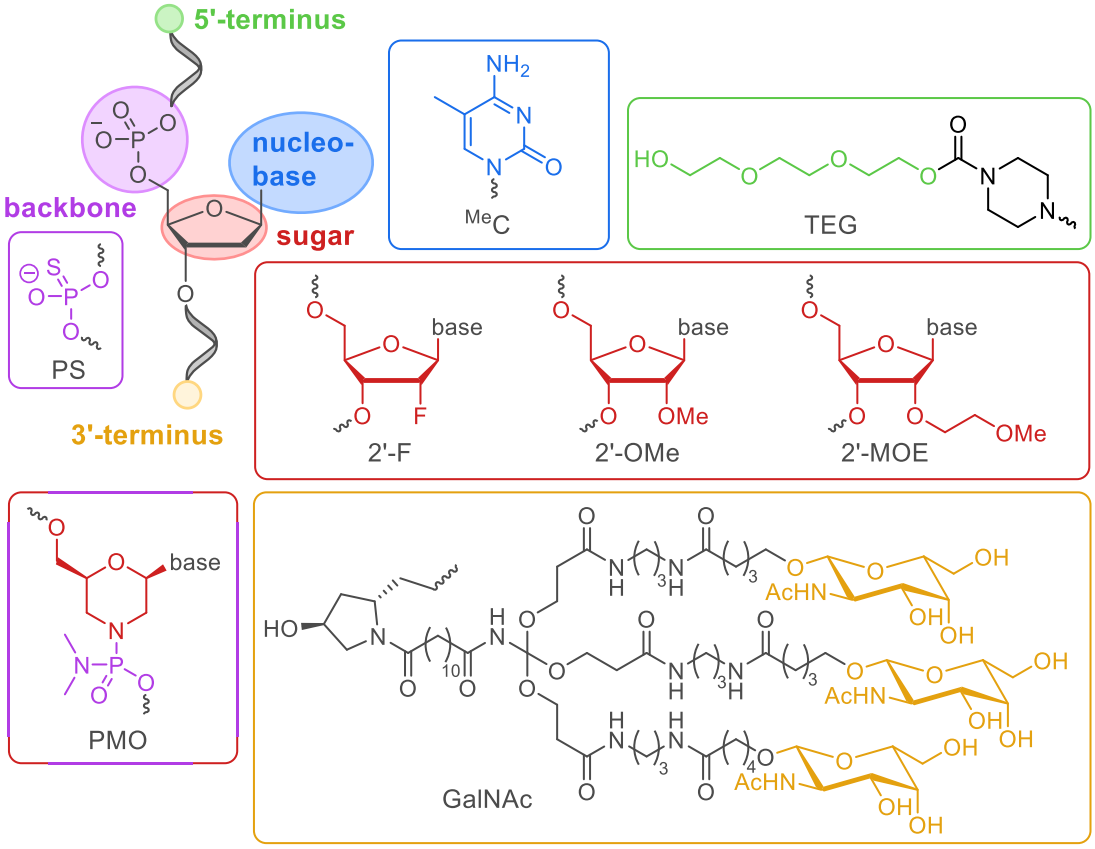



**Figure I-18: Chemical modifications found in clinically approved oligonucleotide drugs.** Modifications of the nucleobase, backbone, sugar, 3′-, or 5′-terminus are highlighted in blue, purple, red, green, or yellow, respectively. Abbreviations: GalNAc, *N*-acetylgalactosamine; $^{Me}$C, 5-methylcytosine; MOE, *O*-methoxyethyl; PMO, phosphorodiamidate morpholino oligonucleotide; PS, phosphorothioate; TEG, triethylene glycol.

through interactions with cell surface proteins.[124] However, the phosphorothioate linkage is a double-edged sword as its *P*-chirality makes PS–ASO drug discovery convoluted.[94] As such, the different stereoisomers of the phosphorothioate linkage have different RNase H compatibilities,[94] and only one of the two stereoisomers is mainly protected from nucleolytic degradation.[122] Moreover, the all-phosphorothioate backbone is often associated with severe toxicity and side effects.[125–127] Whilst protein binding can be beneficial for serum half-life, tissue distribution, and cellular uptake, non-specific protein binding can also contribute to

the toxicity of PS–ASOs.[126,127] Interestingly, there is increasing evidence that the toxicity of PS–ASOs with an all-phosphorothioate backbone can be considerably reduced when introducing other backbone modifications: amongst others, precisely positioned charge-neutral alkyl phosphonate,[128,129] phosphotriester,[129] (thio)formacetal,[129] or amide[129] backbones have been shown to improve the therapeutic indexes of their toxic all-phosphorothioate parent PS–ASOs; the presence of negatively-charged mesylphosphoramidates[130] can reduce non-specific protein binding of PS–ASOs, reducing their pro-inflammatory profiles; and reducing the phosphorothioate content was shown to improve tolerability of PS/PO-mixed backbone ASOs compared to their all-phosphorothioate equivalents whilst retaining efficacy.[131] Together, those studies suggest that mixed backbones with a reduced phosphorothioate content, or ASOs in which some of the phosphorothioates are replaced by alternative backbone linkages, are a promising way forward towards next-generation antisense therapeutics.

Modification of the nucleobases are less adopted, but the $^{Me}$C is a nucleobase modification that is commonly found in ASOs[108,132–134] to increase target affinity and reduce immunogenicity.[40,93] For three decades, chemical modifications of ASOs have mainly focused on improving serum stability and target affinity. To this end, modifications are mostly centred around the 2′-position of the ribose moiety (2$^{nd}$ generation). As mentioned above, introduction of electronegative substituent at the 2′-position favours the RNA-typical C3′-*endo* conformation which results in a

**Table I-2: List of clinically approved oligonucleotide therapeutics.**

| Year | Drug | Indication | Mechanism | Modifications |
|---|---|---|---|---|
| 1998 | Fomivirsen[121] (Vitravene™) | Cytomegalovirus retinitis | RNase H | PS |
| 2013 | Mipomersen[132] (Kynamro™) | Familial hypercholesterolemia | RNase H | PS, 2′-MOE, 5-methyl C |
| 2016 | Nusinersen[133] (Spinraza™) | Spinal muscular atrophy | Splice switching | PS, 2′-MOE, 5-methyl C |
| 2016 | Eteplirsen[98] (Exondy 51™) | Duchenne muscular dystrophy | Splice switching | PMO |
| 2018 | Inotersen[108] (Tagsedi™) | Hereditary transthyretin-mediated amyloidosis | RNase H | PS, 2′-MOE, 5-methyl C |
| 2018 | Patisiran[119] (Onpattro™) | Hereditary transthyretin-mediated amyloidosis | RNAi | 2′-OMe |
| 2019 | Volanesorsen[134] (Waylivra™) | Familial chylomicronemia syndrome | RNase H | PS, 2′-MOE, 5-methyl C |
| 2019 | Golodirsen[137] (Vyondys 53™) | Duchenne muscular dystrophy | Splice switching | PMO, TEG |
| 2019 | Givosiran[120] (Givlaari™) | Acute hepatic porphyria | RNAi | PS, 2′-OMe, 2′-F, GalNAc |
| 2020 | Viltolarsen[138] (Viltepso™) | Duchenne muscular dystrophy | Splice switching | PMO |

Abbreviations: GalNAc, *N*-acetylgalactosamine; $^{Me}$C, 5-methylcytosine; MOE, *O*-methoxyethyl; PMO, phosphorodiamidate morpholino oligonucleotide; PS, phosphorothioate; TEG, triethylene glycol.

higher affinity to RNA targets.[7,20,21] The 2′-modifications commonly found in oligonucleotide therapeutics are 2′-F, 2′-OMe and 2′-MOE (Table I-2). A structurally more different motif is found in phosphorodiamidate morpholino oligonucleotides (PMO), which combine a charge-neutral backbone with a 6-membered morpholine ring as a sugar mimic.[135] Interaction with scavenger receptors enhances cellular delivery[136] and several PMO-based therapeutics have been clinically approved to date (Table I-2).[98,133,137,138] Moreover, aqueous solubility of the charge-neutral PMOs can be enhanced by conjugation with a 5′-terminal triethylene glycol (TEG).[139] Despite said advantages, PMOs are associated with concerns over off-target effects[140,141] and solid phase synthesis of PMO–DNA

chimeras has not yet been established due to the opposite directionalities of DNA and traditional PMO synthesis.[135] More recently, the focus has shifted towards modifications that enhance delivery and enable tissue targeting. For instance, a tridendrically conjugated *N*-acetylgalactosamine (GalNAc) at the 3′-terminus can be used to specifically target hepatocytes.[89,142]

# 5. Backbone Modifications in Synthetic Biology

The modification of an oligonucleotide can have myriad applications in synthetic biology and other biotechnological applications.[56] Like ASOs, oligonucleotides for synthetic biology can be modified at the nucleobase, sugar, or backbone. However, a comprehensive introduction of all modification sites and their possible applications in synthetic biology is beyond the scope of this chapter. Instead, the following section will describe selected examples of backbone modifications for the construction of catalytically active biomolecules and gene synthesis.

## 5.1. Construction of Active Biomolecules

Many applications in synthetic biology involve long DNA or RNA with several hundreds of base pairs. However, the precise introduction of chemical modifications *via* solid phase synthesis is not feasible for such long oligomers. Thus, ligations are often pursued to assemble chemically modified, long oligonucleotides.[56] Amongst those, the copper(I)-catalysed azide–alkyne cycloaddition (CuAAC)[143,144] is a versatile reaction to ligate short, azide- and alkyne-modified oligonucleotides, forming triazoles as artificial backbone mimics (Figure I-19A).[56]

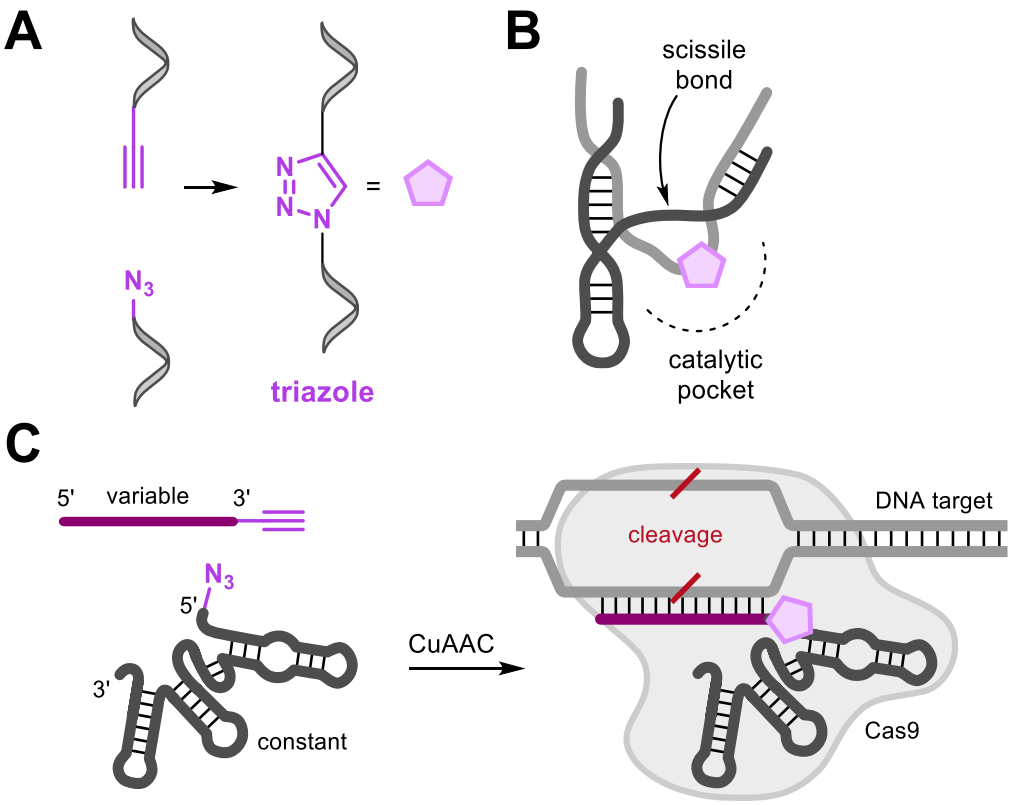



**Figure I-19: Construction of bioactive, longer oligonucleotides through chemical ligation of modified, shorter oligonucleotides.**
(**A**) CuAAC of alkyne- and azide-modified ONs to generate a triazole as an internucleoside linkage. (**B**) Schematic representation of a triazole backbone-modified hammerhead ribozyme. The triazole backbone is in the catalytic pocked, which does not alter the catalytic activity of the hammerhead ribozyme to hydrolyse the scissile phosphate bond. (**C**) CuAAC ligation of a 3′-alkyne variable and a 3′-azide constant RNA to assemble functional sgRNAs that can induce the sequence-specific cleavage of a target DNA by the Cas9 enzyme.

Although not producing a 'long' oligonucleotide with several hundred bases, one report exemplified the templated CuAAC ligation of azide- and alkyne-labelled short RNAs as a proof-of-concept to assemble a triazole-modified, functional hammerhead ribozyme (Figure I-19B).[145] The hammerhead ribozyme is a short RNA motif with self-cleaving activity[146,147] and the authors showed that a triazole backbone within the catalytic pocket does not alter the self-cleavage rate or pattern compared to its unmodified analogue.[145]

Another powerful application of the CuAAC ligation was demonstrated by the construction of a triazole-modified, bioactive single guide RNA (sgRNA) for gene editing applications (Figure I-19C).[148] Importantly, this approach allows the cost-effective assembly of sgRNA libraries *via* the ligation of an invariant, azide-modified RNA (79-mer) with a variable, DNA-targeting, alkyne-modified RNA (20-mer). The resulting, triazole-modified sgRNAs were shown to be functional for the sequence-specific cleavage of a DNA target by the Cas9 nuclease.[148]

## 5.2. Synthetic Genes

The introduction of unnatural internucleoside linkages into DNA or RNA generates backbone-modified oligonucleotides that can retain their genetic information to a remarkable extent.[56] Such artificial genetic polymers can then serve as templates for transcription or translation by polymerases (Figure I-20A). It is important that the polymerase can continue replication/transcription through the backbone modification to produce a replicon/transcript with the correct, complementary sequence.[48,56,81,85]

One literature-reported gene synthesis example uses the phosphoramidate internucleoside linkage which is formed by the condensation of amine- and phosphate-modified oligonucleotides (Figure I-20B). Furthermore, several of such condensations allow the assembly of whole genes *via* self-templated ligation of chemically modified, shorter oligonucleotides (Figure I-20C). In this study, a total of eight chemical ligations have been performed between amine- and phosphate-

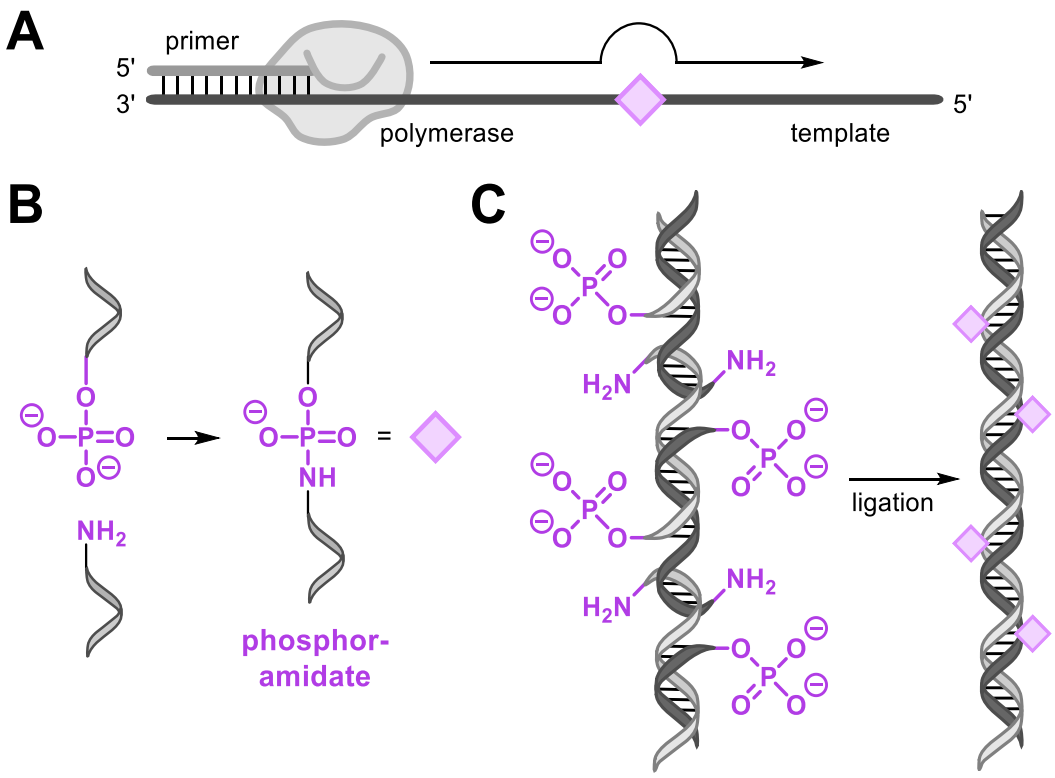



**Figure I-20: Formation of phosphoramidate internucleoside linkages to assemble a whole gene that can be replicated by a polymerase.**
(**A**) Polymerase-mediated extension of a primer along a backbone-modified template. It is crucial that the polymerase can continue replication through the artificial backbone linkage (square) to produce a replicon with the correct, complementary sequence. (**B**) Condensation of amine- and phosphate-modified ONs to generate a phosphoramidate as an internucleoside linkage. (**C**) Self-templated condensation of amine- and phosphate-modified, short ONs to assemble a 762-base pair *egfp* gene. Note that not all eight phosphoramidates (squares) are shown for simplicity.

labelled shorter oligonucleotides to assemble a 762-base pair enhanced green fluorescent protein (*egfp*) gene that could be correctly replicated by DNA polymerases *in vitro*.[81] Moreover, the authors showed that the phosphoramidate linkage can be transcribed by DNA and RNA polymerases.[81]

Other artificial backbone motifs such as triazoles have been reported as part of genetic polymers[48,83,149–155] and will be introduced in more detail in Chapter V,

which will focus on the replication compatibility of triazole backbone-modified genetic templates.

# 6. Overall Structure of the Thesis

The presented work describes the modification of the oligonucleotide backbone with applications in two areas: Chapter II–IV focus on antisense oligonucleotides with charge-neutral, achiral backbone modifications for therapeutic applications; and Chapter V reports backbone-modified oligonucleotides as artificial genetic templates.

*Chapter II* focuses on the carbamate bond (see Figure I-15) as a charge-neutral, achiral mimic of the natural phosphodiester backbone. Moreover, the chapter reports the combination of the carbamate backbone with the locked nucleic acid sugar modification and discusses their combined antisense properties.

*Chapter III* introduces the amide bond (see Figure I-15) as another charge-neutral, achiral backbone surrogate and describes synthetic strategies towards the solid phase synthesis of backbone chimeras.

*Chapter IV* builds on the findings of its preceding chapter to synthesise amide backbone chimeras as antisense oligonucleotides. This chapter provides a systematic evaluation of the amide backbone and its compatibility with inducing an *in vitro* RNase H response. Moreover, the chapter reports target hybridisation, serum stability and cellular uptake studies of amide-modified oligonucleotides discussing their therapeutic potential.

*Chapter V* shows the synthesis and incorporation of a new triazole backbone structure into oligonucleotides. This chapter focuses on the read-through compatibility of various triazole backbone structures by DNA polymerases as part of backbone modified genetic templates. Moreover, a triazole structure–biocompatibility discusses the structural differences of the triazoles in the context of their observed read-through compatibilities.

# CHAPTER II

# CARBAMATE–LOCKED NUCLEIC ACID CHIMERAS FOR POTENTIAL ANTISENSE APPLICATIONS

# 1. Background

## 1.1. The Carbamate Internucleoside Linkage

In general, the modification of backbones is a promising strategy to enhance the serum stability of oligonucleotides,[34] which is a key requirement for therapeutic approaches. Moreover, good hybridisation properties with RNA are important for potential antisense applications.[156,157] This chapter focuses on the charge-neutral, 4-atom carbamate linkage (CA), which has two constitutional isomers ($CA_X$ and $CA_Y$; Figure II-1). Because of their polyanionic character, unmodified oligonucleotides do not readily enter the cell, limiting their therapeutic potential.[93] Consequently, the charge-neutrality of the carbamate linkage is a promising characteristic to reduce the overall charge and enhance cellular uptake.[158,159] Both carbamate linkages—$CA_X$

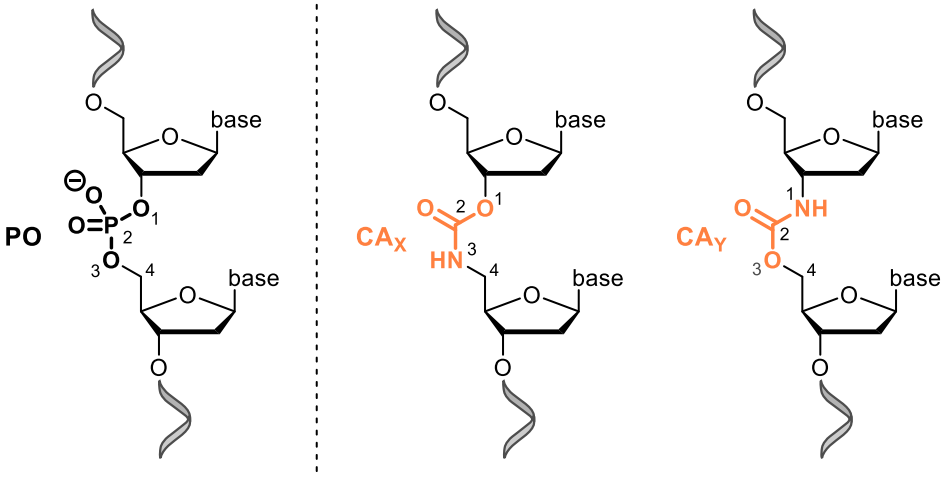



**Figure II-1: The constitutional isomers of the carbamate linkage.**
The carbamate moiety is highlighted. The natural PO is shown for reference. The linker length is defined as the minimum number of atoms connecting the sugar rings as indicated. $CA_X$, $CA_Y$, and the natural PO are all 4-atom internucleoside linkages.

and $CA_Y$—have been previously described as phosphodiester bond mimics. As such, the $CA_X$ linkage is stable under basic and acidic conditions, and against degradation by spleen or snake venom phosphodiesterases.[160] Moreover, oligonucleotides in which the phosphodiester bonds have been fully substituted by the $CA_X$ linkage have been evaluated for their hybridisation properties with RNA and DNA targets (ON1 and ON2; Table II-1).[161,162] Whilst the $CA_X$ linkages in ON1 increased duplex stabilities with RNA and DNA targets,[161] the same internucleoside modification led to non-detectable interactions of ON2 with its targets.[162] Despite this discrepancy, most experimental data suggest that the $CA_X$ linkage destabilises duplexes with RNA (ON3–ON6; Table II-1).[69] The constitutional isomer $CA_Y$ has been evaluated in duplex stability experiments, too. As such, $CA_Y$ decreased the affinity to RNA and DNA targets substantially more than $CA_X$ (ON7 and ON8; Table II-1), but no stability data in acidic or basic conditions, or against nucleolytic degradation, has

**Table II-1: Duplex stability properties of oligonucleotides with a carbamate backbone.**
Data is taken from literature[69,161–163] and the reported $\Delta T_m$'s are relative to the isosequential, unmodified DNA controls.

| | | | RNA target | DNA target |
|---|---|---|---|---|
| CA | ON | Sequence (5′→3′) | $\Delta T_m$[a] (°C)/mod | $\Delta T_m$[a] (°C)/mod |
| **X** | ON1 | $H_2N$-C$^X$C$^X$C$^X$C$^X$C$^X$C$^X$C$^X$C | +5.6[161] | +5.4[161] |
| **X** | ON2 | T$^X$T$^X$T$^X$T$^X$T$^X$T$^X$T | no interaction[162] | no interaction[162] |
| **X** | ON3 | CTCGTACCT$^X$TTCCGGTCC | −3.4[69] | n. a. |
| **X** | ON4 | CTCGTACT$^X$TT$^X$TCCGGTCC | −3.2[69] | n. a. |
| **X** | ON5 | GCGT$^X$TT$^X$TT$^X$TT$^X$TT$^X$TGCG | −4.8[69] | n. a. |
| **X** | ON6 | TTTT$^X$TCTCTCTCTCT | −2.9[69] | −1.8[69] |
| **Y** | ON7 | GCGTTT$^Y$TTTGCT | −7.3[163] | −8.0[163] |
| **Y** | ON8 | GCGT$^Y$TT$^Y$TT$^Y$TGCT | −12.0[163] | −9.4[163] |

$^X$ and $^Y$ denote **$CA_X$** and **$CA_Y$** linkages (see Figure II-1).
[a] the $\Delta T_m$ is the difference of the observed $T_m$ to the $T_m$ of the isosquential unmodified control with the same target. The presented $\Delta T_m$'s are normalised to the number of modifications.

been reported.[163] Thus, this chapter focuses on the $CA_X$ linkage for potential antisense applications. Hereafter, the simplified term ‘carbamate’ will refer to the $CA_X$ linkage if not otherwise stated.

Apart from its nuclease resistance and hybridisation properties, the rigid carbamate linkage can lead to poor solubility of all-carbamate-modified oligonucleotides.[164] However, a single ionisable group (see amine in ON1, which is protonated under physiological pH; Table II-1) or natural phosphodiester linkages (ON3–ON8; Table II-1) seem sufficient to confer aqueous solubility. It is noteworthy that the mentioned oligonucleotides with consecutive carbamate linkages (ON1 and ON2; Table II-1) have been prepared in solution.[161,162] Indeed, a solid phase synthesis protocol for the on-resin carbamate formation has not yet been reported.

The carbamate modification alone is less than ideal to confer desired drug-like properties because it decreases affinity to an RNA target. A possible solution could be the combination of the carbamate backbone with other modifications that are known to enhance RNA affinity. For example, a modified nucleobase has been previously described to restore RNA target affinity of oligonucleotides with duplex-destabilising triazole backbones.[39] Moreover, the introduction of locked nucleic acids has been proven as an excellent strategy to increase RNA target affinities.[165–168]

## 1.2. The Locked Nucleic Acid Modification

### 1.2.1. Biophysical Properties

The bicyclic locked nucleic acid ribose (LNA; Figure II-2)[169,170] is a widely adopted modification to confer several drug-like properties. Although LNAs do not change the phosphodiester backbone, the bicyclic sugar protects its surrounding phosphodiester bonds to some extent against nucleolytic degradation.[171] As described in the introduction, RNA nucleosides preferentially adopt a C3′-*endo* sugar pucker. The perspective formula of the locked nucleic acid structure in Figure II-2 shows that the oxymethylene bridge between C2′ and C4′ fixes the sugar pucker in the C3′-*endo* conformation. This pre-organises the sugar pucker into the preferred RNA conformation and thus, reduces the entropy penalty for binding to an RNA target. As a result, locked nucleic acids largely increase duplex stabilities with RNA targets.[172–174] However, this stabilisation can cause LNA-modified

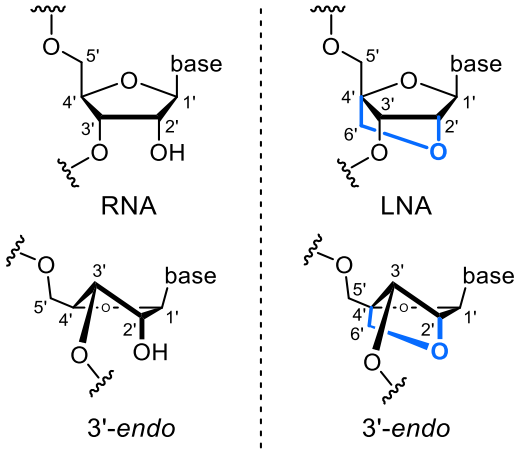
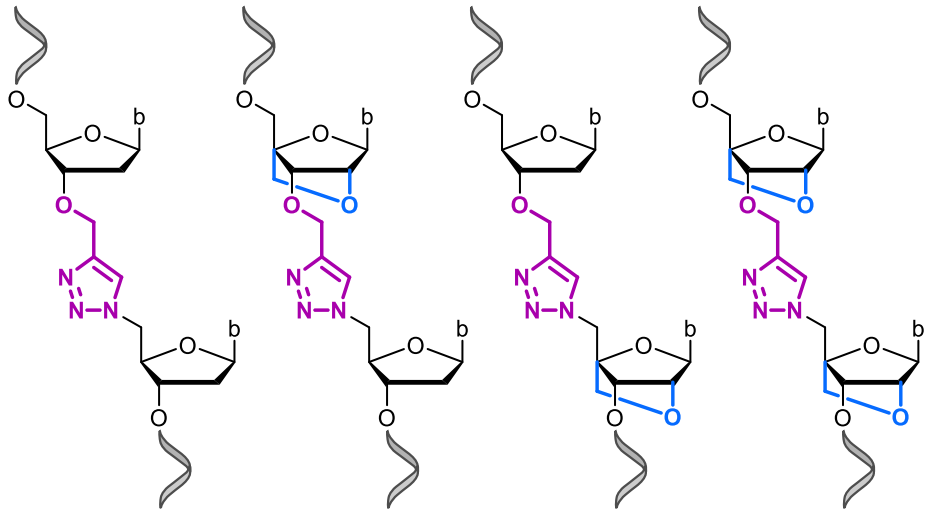



**Figure II-2: The rigid sugar pucker of the locked nucleic acid modification.**
Comparison of ribose and LNA sugars. Top row: structural formula of the ribose (left) and LNA (right) nucleosides. Bottom row: perspective structural formula of the twisted C3′-*endo* conformations of the ribose (left) and LNA (right) sugar puckers. The oxymethylene bridge is highlighted.

oligonucleotides binding to imperfectly matched RNAs (off-targets), which can lead to toxicity.[175]

The duplex-stabilising properties of locked nucleic acids have been previously utilised in combination with other, duplex-destabilising modifications to mitigate their otherwise reduced binding affinities.[165–168] One example is the combination of locked nucleic acids with a duplex-destabilising triazole backbone (Table II-2).[165] The duplex destabilisation by the triazole linkage with an RNA target could be mitigated by an LNA modification on the 3′-side, or on both sides of the triazole. Interestingly, LNA on the 5′-side of the triazole alone did not restore the duplex stability with an RNA target. In fact, the presence of an LNA on the 5′-side of a

**Table II-2: Duplex stability properties of LNA-modified oligonucleotides with a triazole backbone.**
Data is taken from literature[165] and the reported $\Delta T_m$'s[a] are relative to the isosequential, unmodified DNA control.




| target strand | $\Delta T_m$ (°C) | | | |
|---|---|---|---|---|
| RNA | −6.2[165] | −7.2[165] | −0.8[165] | +0.1[165] |
| DNA | −8.9[165] | −11.5[165] | −6.0[165] | −5.8[165] |

[a] the $\Delta T_m$ is the difference of the observed $T_m$ to the $T_m$ of the isosquential unmodified control with the same target.
The triazole backbone and the oxymethylene bridge of the locked nucleic acid are highlighted. Abbreviations: b, nucleobase.

triazole[165–167] or amide[168] modification mostly failed to stabilise duplexes with RNA targets. Moreover, combining duplex-destabilising backbones with LNAs does not guarantee rescued RNA target affinity. In fact, previous reports showed that, depending on their position, LNAs can even further reduce duplex stabilities of backbone-modified oligonucleotides with their RNA targets.[166] As expected, the extent of duplex stabilisation by LNAs is generally less for DNA targets, because of the discrepancy between the fixated C3′-*endo* LNA pucker and the preferred C2′-*endo* conformation in DNA.

# 2. Aim of the Chapter

Both the carbamate backbone and the LNA sugar confer some desirable drug-like properties but are insufficient as individual modifications to ensure adequate drug-like properties. The aim of this chapter is to investigate combinations of the carbamate backbone with the LNA to overcome the limitation of reduced target affinity from carbamate-only modified oligonucleotides. Because the $CA_X$ backbone is known to be less destabilising than $CA_Y$, this study focuses on the combination of LNAs with the former. Several structural designs are required to test the interrelation between the carbamate backbone and the LNA sugar (Figure II-3A). Because of the lack of a synthesis protocol for the on-resin formation of the carbamate backbone, a dimer strategy seems most feasible to access the desired $CA_X$–LNA combinations. For this purpose, four different dinucleoside CE phosphoramidites **5**–**8** are required when in the context of a T–T sequence motif (Figure II-3B). Those phosphoramidites will allow the solid phase synthesis of carbamate- and LNA-modified oligonucleotides for duplex stability studies with DNA and RNA targets. The results obtained from these studies will give valuable insights into the effects of combining the carbamate backbone with LNAs for potential antisense applications.

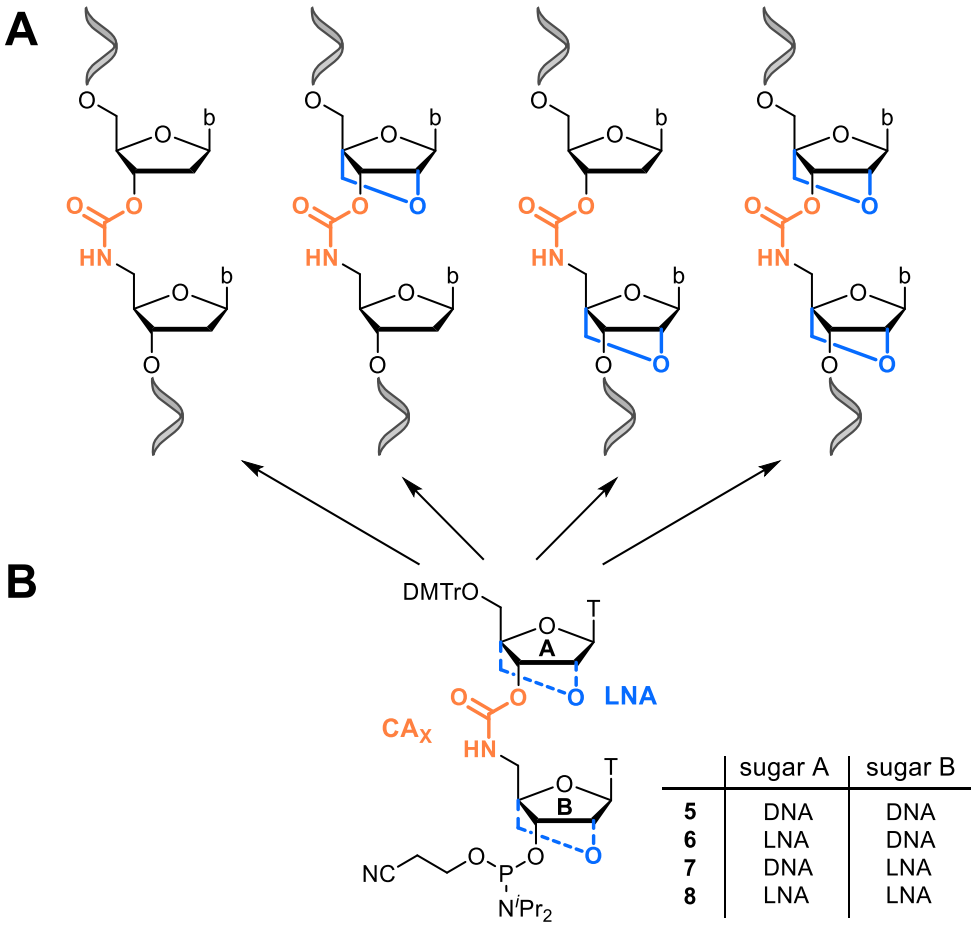



**Figure II-3: Structural designs of CA$_X$–LNA chimeras to test the interrelation between the carbamate backbone and the LNA modification.**
(**A**) The four different combinations of the carbamate backbone with LNAs. The CA$_X$ backbone and the oxymethylene bridge of the LNA sugar are highlighted. (**B**) Dinucleoside CE phosphoramidites **5**–**8** required for this study. Abbreviations: b, nucleobase; DMTr, 4,4′-dimethoxytrityl; T, thymine.

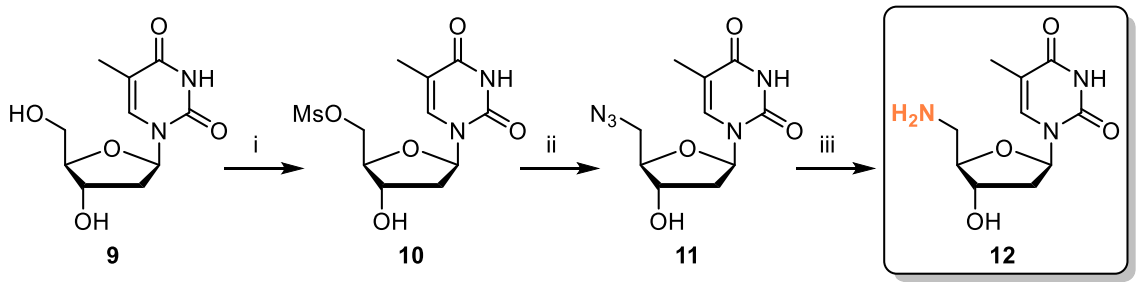
# 3. Results and Discussion

## 3.1. Synthesis of Carbamate-Linked Dinucleoside Phosphoramidites

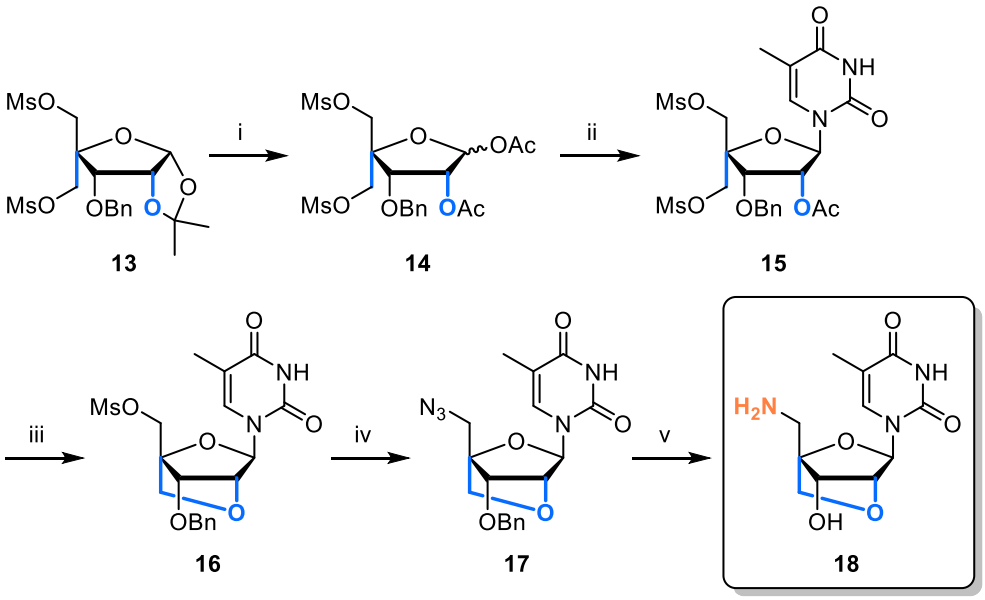
In line with the pursued dimer strategy to synthesise carbamate-modified oligonucleotides with LNAs, dinucleoside CE phosphoramidites **5**–**8** were synthesised *via* solution phase carbamate bond formation between suitable nucleophilic or electrophilic nucleoside monomers.

### 3.1.1. Synthesis of Nucleophiles

DNA nucleophile **12** was synthesised from thymidine (**9**) applying a chiral pool strategy according to literature procedures (Scheme II-1).[176] This involved the regioselective mesylation of the 5′-OH in thymidine (**9**) to **10**[176] and successive substitution by sodium azide to **11**.[176] In a last step, azide **11** was hydrogenated to




**Scheme II-1: Synthesis of DNA nucleophile 12.**
Reagents and conditions: (i) methanesulfonyl chloride, pyridine, $CH_2Cl_2$, 0 °C to rt, 2.5 h, 66%; (ii) $NaN_3$, DMF, 60 °C, 16 h, 94%; (iii) $H_2$, Pd/C (10%), MeOH, rt, 5 h, 94%. Abbreviations: Ms, methanesulfonyl.




**Scheme II-2: Synthesis of LNA nucleophile 18.**
Reagents and conditions: (i) a) TFA/$H_2O$, 8/2 (v/v), rt, 2.5 h; b) $Ac_2O$, pyridine, rt, 16 h, 75% (2 steps); (ii) thymine, *N*,*O*-bis(trimethylsilyl)acetamide (BSA), trimethylsilyl trifluoromethanesulfonate (TMSOTf), MeCN, reflux, 4 h, 81%; (iii) NaOH, 1,4-dioxane/$H_2O$, 1/1 (v/v), rt, 1 h, 75%; (iv) $NaN_3$, DMF, 65 °C, 20 h, 90%; (v) a) $H_2$, $Pd(OH)_2$ (20%), MeOH, rt, 3 h; b) $HCO_2NH_4$, $Pd(OH)_2$, reflux, 4 h, 87% (2 steps). Abbreviations: Ac, acetyl; Bn, benzyl; Ms, methanesulfonyl.

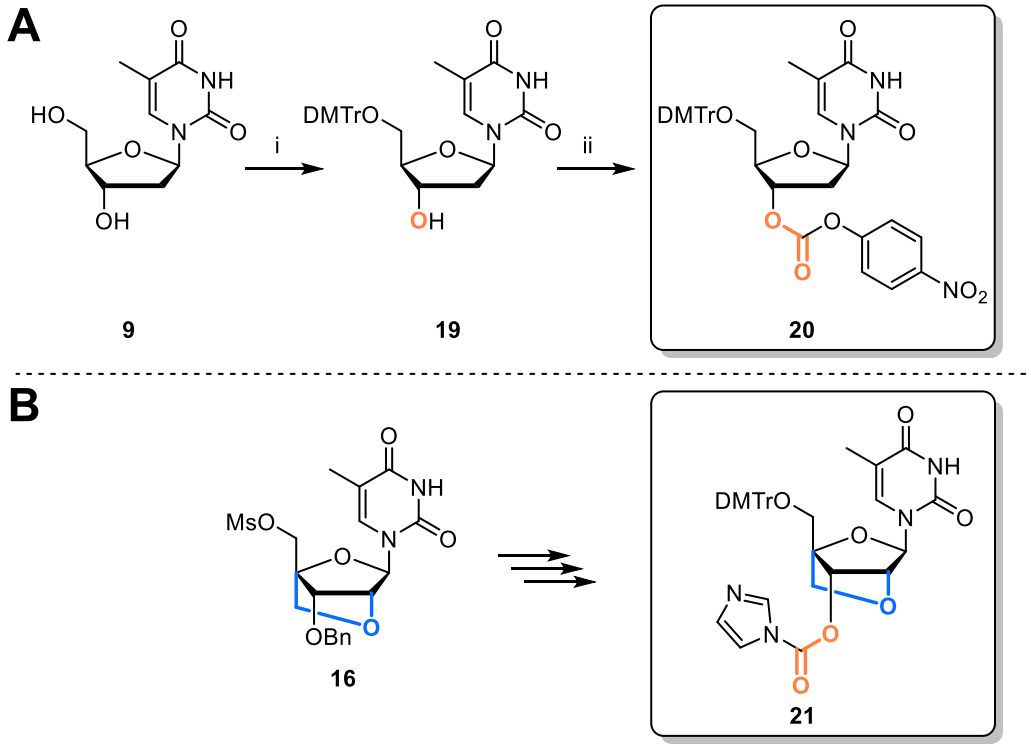
afford amine **12**,[176] which is the first nucleophilic building block for synthesising the proposed dinucleoside CE phosphoramidites **5**–**8**.

The synthesis of LNA nucleophile **18** followed a modified procedure of the chiral pool-based LNA synthesis developed by Koshkin et al. (Scheme II-2).[170,177] In a first step, commercially available acetonide **13** was cleaved under acidic conditions to give **14**,[177] which was glycosylated using Vorbrüggen conditions to thymidine **15**.[177] Successive, treatment under basic conditions hydrolysed the 2′-OAc group to release a 2′-OH, which displaced the α-faced mesyl group *via* an intramolecular $S_N2$ reaction to afford LNA **16**[170] in a one-pot procedure. The β-faced mesyl group was further substituted with sodium azide to obtain **17**, which was hydrogenated in a two-step procedure to LNA nucleophile **18**. The two-step

hydrogenation was necessary because initial attempts using a palladium-catalysed reduction under a hydrogen atmosphere only reduced the azide, but addition of ammonium formate at an elevated temperature reduced the remaining 3′-OBn protecting group.

### 3.1.2. Synthesis of Electrophiles

The synthesis of DNA electrophile **20**[76,178,179] followed established literature procedures using a chiral pool strategy (Scheme II-3A).[76,178–181] In a first step, commercially available thymidine (**9**) was regioselectively protected at the 5′-




**Scheme II-3: Synthesis of DNA electrophile 20 and LNA electrophile 21.**
Reagents and conditions: (i) 4,4′-dimethoxytrityl chloride, pyridine, rt, 16 h, 85%; (ii) 4-nitrophenyl chloroformate, $CH_2Cl_2$/pyridine, 2/1 (v/v), 4 °C to rt, 16 h, 46%. Abbreviations: Bn, benzyl; DMTr, 4,4′-dimethoxytrityl; Ms, methanesulfonyl. The synthesis of electrophile **21** is described elsewhere.[182]

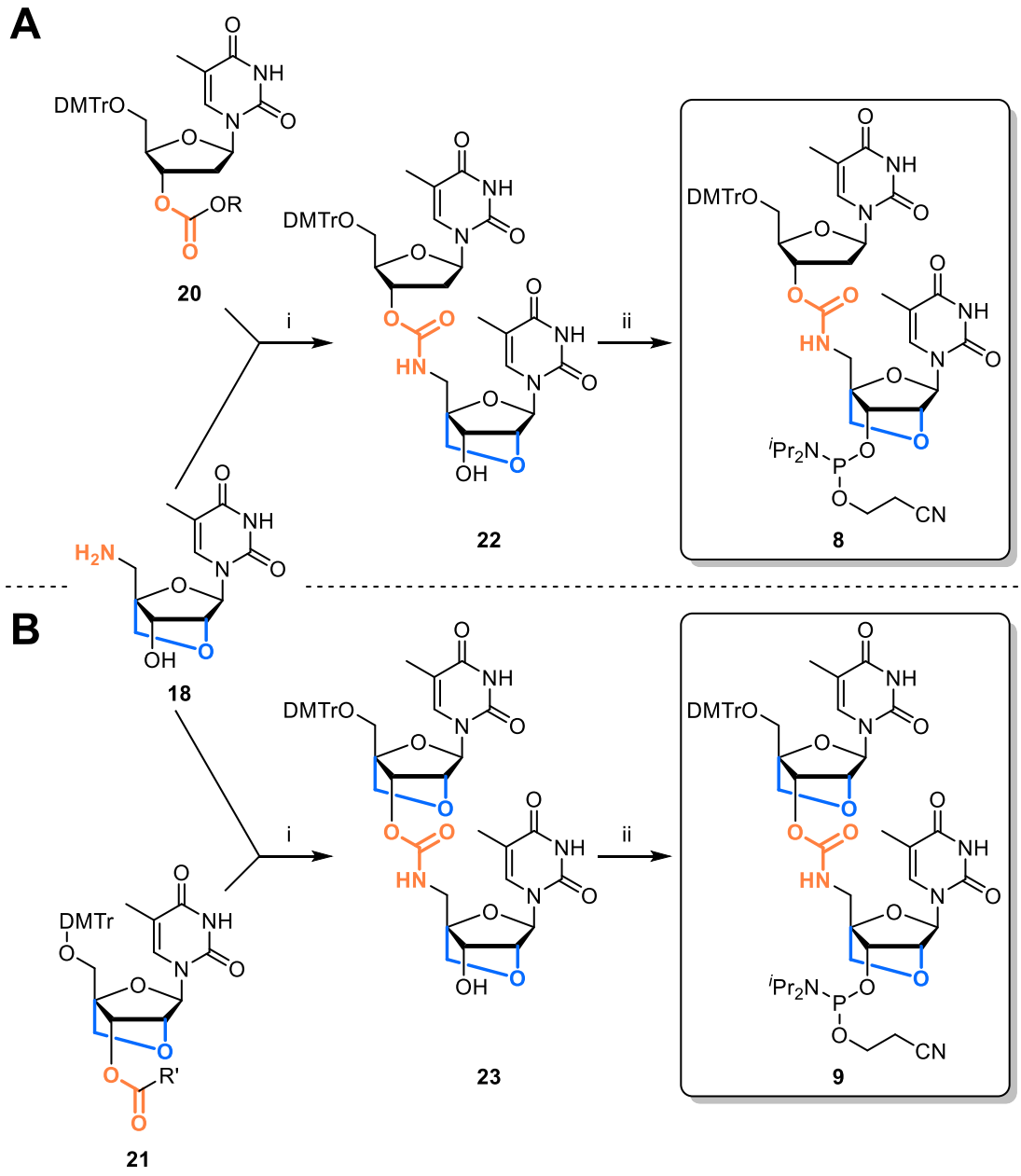
position to afford **19**.[180,181] The 3′-position was then activated as a *para*-nitrophenyl carbonate to introduce the electrophilic carbonyl in **20**.[76,178,179]

The LNA-type electrophile **21** was prepared from **16** in four steps (Scheme II-3B).[e,182] In this case, activation of the 3′-position was achieved as the carbonylimidazolide **21** because activation *via* a *para*-nitrophenyl carbonate similar to **20** was futile.[162]

### 3.1.3. Synthesis of Dimers

After successful preparation of nucleophiles **12** and **18**, and electrophiles **20** and **21**, the monomers were reacted in different combinations to form the desired $CA_X$-linked dinucleosides with various LNA modification sites. As such, carbonate **20** was reacted with LNA nucleophile **18** to form DNA–$CA_X$–LNA dimer **22**, which was phosphitylated to DNA–$CA_X$–LNA phosphoramidite **7** for solid phase oligonucleotide synthesis (Scheme II-4A). Condensation of LNA electrophile **21** with LNA nucleophile **18** produced LNA–$CA_X$–LNA dimer **23**, and successive phosphitylation afforded LNA–$CA_X$–LNA phosphoramidite **8** (Scheme II-4B). The remaining two target dimers (DNA–$CA_X$–DNA **5** and LNA–$CA_X$–DNA **6**; see Figure II-3) were synthesised in a similar way.[f,182]

[e] Prepared by Benjamin Woods (Part II), University of Oxford.
[f] dimer **5** was synthesised by Cameron Thorpe (DPhil), University of Oxford; and dimer **6** was synthesised by Benjamin Woods (Part II), University of Oxford.




**Scheme II-4: Convergent synthesis of dinucleosides 7 and 8 for solid phase oligonucleotide synthesis.**

(**A**,**B**) Reagents and conditions: (i) pyridine, 80 °C, 16 h, 51% **22**, 70% **23**; (ii) 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, $CH_2Cl_2$, rt, 2 h, 57%, **7**, 88%, **8**. R = *para*-nitrophenyl, R′ = imidazol-1-yl. Abbreviations: DMTr, 4,4′-dimethoxytrityl; $^{i}$Pr, isopropyl.

## 3.2. Biophysical Studies of Carbamate-Linked LNA Oligonucleotides

Following successful preparation of dimer CE phosphoramidites **5**–**8**, oligonucleotides with various combinations of carbamate backbones and LNA sugars were synthesised for evaluation in biophysical and nucleolytic degradation studies. Importantly, no solubility issues in aqueous buffers were observed for any of the synthesised oligonucleotides. The studies included thermal melting experiments of the duplexes that are formed by modified oligonucleotides upon hybridisation with a target strand to evaluate their duplex stabilities. Moreover, the overall structural integrities of duplexes formed by modified oligonucleotides and their respective targets were assessed by circular dichroism experiments. The stability against nucleolytic degradation was evaluated by incubation of modified oligonucleotides with nucleases.

### 3.2.1. Duplex Stabilities

Two sets of modified oligonucleotides were prepared for thermal melting experiments[g] and compared to an unmodified DNA control (ON9; Table II-3). The two sets included oligonucleotides containing either a single or a triple incorporation of dimers **5**–**8**.

The duplex stability effects resulting from a single carbamate linkage in different combinations with adjacent LNAs are summarised in Table II-3. As

[g] in collaboration with Cameron Thorpe (DPhil), University of Oxford.

**Table II-3: $T_m$'s of ON9–ON13 against their complementary RNA and DNA targets.**
ON9–ON13 were incubated with their respective RNA or DNA targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0) and absorbance was measured at λ = 260 nm in relation to *T*. $T_m$'s were obtained from maxima of the 1st derivatives ($dA_{260}/dT$) (see Figure A-1; Appendix) and $\Delta T_m$'s are reported as the difference to $T_m$(ON9). *Table and caption reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

| | | RNA target | DNA target |
|---|---|---|---|
| ON | Sequence (5′→3′) | $T_m$ ($\Delta T_m$)/°C | $T_m$ ($\Delta T_m$)/°C |
| ON9 | GCTTGCTTCGTTCC | 63.6 | 60.2 |
| ON10 | GCTTGCTTCGTTCC | 57.2 (−6.4) | 57.5 (−2.7) |
| ON11 | GCTTGCtTCGTTCC | 56.1 (−7.5) | 50.5 (−9.7) |
| ON12 | GCTTGCTtCGTTCC | 61.3 (−2.3) | 60.3 (+0.1) |
| ON13 | GCTTGCttCGTTCC | 62.1 (−1.5) | 53.4 (−6.8) |

N DNA  n LNA  PO  CBM

expected, the presence of a carbamate backbone inside a DNA strand decreased duplex stabilities with RNA and DNA targets by −6.4 and −2.7 °C, respectively (ON10; Table II-3). Similar to previous reports,[69] the carbamate seemed better tolerated in a DNA:DNA than in a DNA:RNA duplex. Introduction of an LNA adjacent to the 5′-side of the carbamate led to further destabilisation of the duplexes with both targets (ON11; Table II-3). This phenomenon has been previously described for other backbone modifications[165–168] and can be explained by the effect of the LNA sugar on the backbone linkage to its 3′-side (Figure II-4). Whilst the sugars preferably adopt a C3′-*endo* or C2′-*endo* conformation in RNA or DNA, respectively, those states are in an equilibrium and there is some flexibility.[4] However, the bicyclic system of the LNA sugar fixates the pucker into the C3′-*endo* conformation and restricts flexibility. This affects the position of the internucleoside linkage on the 3′-side of the LNA (Figure II-4).[183] In the case of a phosphodiester

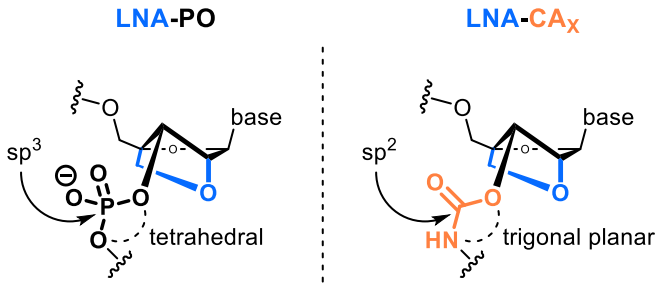



**Figure II-4: The influence of the LNA sugar on the 3′-adjacent backbone linkage.**
Perspective structural formula of the LNA sugar with a 3′-PO (left) or 3′-$CA_X$ (right). The carbamate backbone and oxymethylene bridge are highlighted.

bond, the tetrahedral linkage may be able to better accommodate its fixed position because of its more flexible $sp^3$-hybridised phosphorus. On the contrary, the trigonal planar carbamate with its $sp^2$-hybridised carbon might not have enough flexibility to fully accommodate its fixed position, thereby destabilising the duplex. This positional dependency of the LNA to affect the carbamate linkage is further exemplified for ON12. For this oligonucleotide, the carbamate is attached to the 5′-side of the LNA sugar, whilst the phosphodiester linkage is at its 3′-side. In this design, the influence of the LNA mitigated destabilisation for duplexes with RNA and DNA by +4.1 and +2.8 °C when compared to the carbamate without any LNA (ON12 vs. ON10; Table II-3). As expected, the stabilisation effect is higher for the RNA target. Surrounding the carbamate with an LNA on both sides of the linkage resulted in an even larger differentiation between RNA and DNA affinities (ON13; Table II-3). As such, the duplex of ON13 with the RNA target was destabilised by −1.5 °C whilst the affinity to DNA was reduced by −6.8 °C. Because the target of an antisense oligonucleotide is an RNA strand, such discrimination between DNA and RNA targets is highly desirable for potential antisense applications.

Expanding on the previous results, oligonucleotides with three carbamates in combination with LNAs were tested in thermal melting experiments (Table II-4). Incorporation of three carbamates without any LNA destabilised duplexes with RNA and DNA by −18.8 and −8.1 °C, respectively (ON14; Table II-4). Compared to the single carbamate in ON10, the observed destabilisations are cumulative, i.e., the observed destabilisation was approximately 3× that of a single carbamate. Addition of an LNA on the 5′-side of the carbamate led to further destabilisation (ON15; Table II-4). Again, the destabilisation of the RNA duplex was approximately cumulative. Moreover, the duplex with DNA was so much destabilised that a melting temperature could not be accurately determined. The presence of an LNA on the 3′-side of the carbamate linkages mitigated the destabilising effects when compared to the carbamate-modified oligonucleotide without any LNAs (ON16 vs. ON14; Table II-4). Indeed, the ON16:DNA duplex showed almost identical stability properties to the completely unmodified DNA

**Table II-4: $T_m$'s of ON14–ON17 against their complementary RNA and DNA targets.** ON14–ON17 were incubated with their respective RNA or DNA targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0) and absorbance was measured at λ = 260 nm in relation to *T*. $T_m$'s were obtained from maxima of the 1st derivatives ($dA_{260}/dT$) (see Figure A-2) and $\Delta T_m$'s are reported as the difference to the $T_m$ of the unmodified control (ON9; Table II-3). *Table and caption reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

| | | RNA target | DNA target |
|---|---|---|---|
| ON | Sequence (5′→3′) | $T_m$ ($\Delta T_m$)/°C | $T_m$ ($\Delta T_m$)/°C |
| ON14 | GCTTGCTTCGTTCC | 44.8 (−18.8) | 52.1 (−8.1) |
| ON15 | GCtTGCtTCGtTCC | 40.6 (−23.0) | <30.0 |
| ON16 | GCTtGCTtCGTtCC | 59.6 (−4.0) | 60.1 (−0.1) |
| ON17 | GCttGCttCGttCC | 61.6 (−2.0) | 41.2 (−19.0) |

N DNA  n LNA  PO  CBM

duplex (see ON9:DNA; Table II-3). Incorporation of three carbamates surrounded by LNAs reversed the duplex stabilising effect of the 3′-LNA with a DNA target (ON17; Table II-4). Indeed, the LNA–$CA_X$–LNA modification destabilised the ON17:DNA duplex by −19.0 °C. On the contrary, the ON17:RNA duplex was only destabilised by −2.0 °C and showed the highest absolute melting temperature (61.6 °C) of all modified oligonucleotides tested. These results demonstrate excellent selectivity properties of the LNA–$CA_X$–LNA modification for an RNA target whilst only slightly decreasing target affinity (0.63 °C per mod. for ON17:RNA; Table II-4).

### 3.2.2. Structural Integrity and Nucleolytic Resistance

The abilities of modified oligonucleotides ON14–ON17 to form A- and B-form duplexes with RNA and DNA targets were evaluated in circular dichroism experiments (data not shown).[h,182] In essence, the respective A- and B-form helical structures of duplexes with RNA and DNA were retained for ON14–ON16. Only the ON17:DNA showed a hypsochromic shift towards a duplex structure with more A-form character. This is not surprising because of the increasing influence of the six LNAs with their locked C3′-*endo* conformation. Indeed, similar observations have been reported for other backbones in combination with LNAs.[165]

The resistance of modified oligonucleotides ON14–ON17 against nucleolytic degradation was tested by incubation in foetal bovine serum (FBS) and with snake

[h] Measured by Cameron Thorpe (DPhil), University of Oxford.

venom phosphodiesterase (data not shown).[i,182] In both cases, the nucleolytic resistance of LNA–$CA_X$–LNA-modified ON17 was greatly enhanced compared to unmodified ON9. Moreover, the LNA–$CA_X$–LNA motif in ON17 conferred higher resistance against hydrolysis by snake venom phosphodiesterase than its LNA–PO–LNA-modified analogue.

[i] Tested by Cameron Thorpe (DPhil), University of Oxford.

# 4. Conclusion

Chemical modifications of oligonucleotides are essential to enhance their drug-like properties for antisense applications. However, the introduction of a single type of modification is often not sufficient to fulfil all therapeutic requirements. Thus, combining different types of modifications can be a promising strategy to overcome their individual limitations. Whilst many modifications have already been investigated, the interrelation of combined modifications remains poorly understood. This chapter reports the combination of a carbamate backbone with locked nucleic acids to evaluate their joined effects on duplex formation properties, target discrimination, and nucleolytic resistance. In total, four different carbamate-linked dinucleoside CE phosphoramidites (**5**–**8**) were prepared and used to synthesise oligonucleotides with carbamate backbones surrounded by LNAs in various combinations. Evaluation in melting experiments revealed that the LNA modification can mitigate reduced affinity of carbamate-modified oligonucleotides to RNA targets. Importantly, the stabilising effects were dependent on the location of the LNA with respect to the carbamate backbone. Moreover, three incorporations of the LNA–$CA_X$–LNA motif resulted in the most stable duplex formed by a carbamate-modified oligonucleotide with an RNA target (ON17:RNA; Table II-4). This motif also showed excellent discrimination against a DNA target. Combined with the largely enhanced stability against nucleases, the LNA–$CA_X$–LNA motif

represents a promising combination of backbone and sugar modifications for potential antisense applications.

# 5. Future Work

Although being a valuable combination for antisense applications, the solid phase synthesis of carbamate-modified oligonucleotides currently depends on the preparation of suitable dimer CE phosphoramidites. Considering the expansion of all four T–T phosphoramidites **5**–**8** to any base combination, a total of 64 different dimer phosphoramidites would be required to cover all sequence possibilities and $CA_X$–LNA combinations. Moreover, this approach would still not permit fully carbamate-modified oligonucleotides because the coupling of dimers would inevitably introduce a phosphodiester bond; the highest extent of modified backbones would be 50% with alternating carbamate and phosphodiester/phosphorothioate linkages. Whilst recently reported data suggests that even sporadic incorporation of charge-neutral linkages can greatly enhance the therapeutic properties of PS–ASOs,[128,129] a dimer approach would still limit the flexibility and throughput of screening campaigns aiming to identify the ideal number and position of such charge-neutral internucleoside linkages. Thus, a solid phase synthesis-compatible protocol for the on-resin formation of the carbamate bond is highly desirable to expedite discoveries around $CA_X$-modified oligonucleotides. A proposed integration of the on-resin carbamate bond formation into the solid phase synthesis cycle of oligonucleotides is illustrated in Figure II-5.

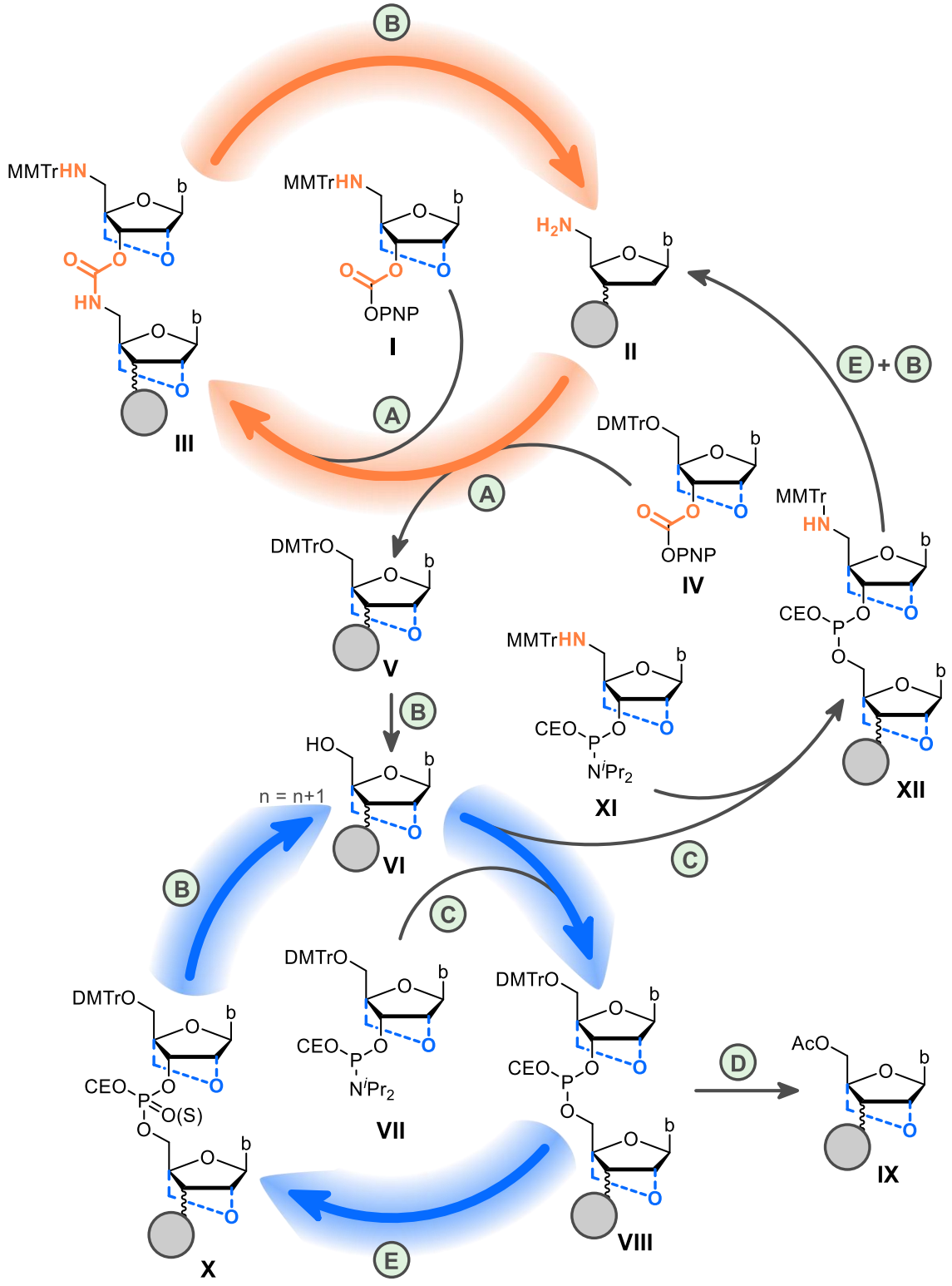



**Figure II-5: Schematic overview of transitions between carbamate couplings (top cycle) and standard CE phosphoramidite oligonucleotide synthesis (bottom cycle).**
(**A**) Carbamate coupling. (**B**) Detritylation. (**C**) Coupling of CE phosphoramidites. (**D**) Capping. (**E**) Oxidation or sulphurisation. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; $^{i}$Pr, isopropyl; MMTr, 4-methoxytrityl; PNP, *para*-nitrophenyl.

The carbamate bond could be formed *via* coupling of a 5′-protected (LNA) 3′-carbonate **I** with a free 5′-amine of the growing oligonucleotide **II**. High coupling efficiencies could be achieved by supplying an excess of carbonate **I**, and/or by adding a catalyst to accelerate the coupling rate, e.g., 4-dimethylaminopyridine (DMAP). The acid stability of the formed carbamate internucleoside linkage would allow acid-mediated deprotection of the 5′-MMTr in **III** to release a free 5′-amine that can enter the next coupling cycle. The standard solid phase oligonucleotide synthesis cycle could be initiated by coupling of (LNA) carbonate **IV** to introduce the 5′-oxygen functionality on the growing end of resin-bound oligonucleotide **V**. Detritylation would result in a free 5′-OH in **VI**, which is a substrate for the steps of the standard solid phase synthesis protocol: coupling of (LNA) CE phosphoramidite 5′-amine **VII** to generate phosphite triester **VIII**; capping of unreacted **VI** to acetylated **IX**; oxidation of phosphite triester **VIII** to phosphotriester **X**; and detritylation to re-generate 5′-OH in **VI** for the next cycle. Alternatively, the standard oxidation could be replaced by a sulphurisation method as explained in Section 3.3.1.1 of Chapter I to produce the phosphorothioate linkage, thereby further expanding the chemical space of possible backbone modifications and combinations thereof. Transition to the carbamate cycle could then be achieved by coupling of MMTr-protected (LNA) CE phosphoramidite **XI**. Oxidation/sulphurisation and detritylation of the resulting phosphite triester **XII** would re-introduce **II** with a free 5′-amine that can enter the carbamate cycle.

As indicated in Figure II-5, coupling could be performed with LNA or DNA nucleosides, largely increasing the flexibility for the design of

PO/PS/$CA_X$/DNA/LNA chimeras. The proposed on-resin carbamate formation cycle could give rise to all-carbamate-modified oligonucleotides and facilitate a flexible approach to screen different mixed backbone designs with variations of phosphodiester, phosphorothioate and carbamate content and positions. In light of recent reports demonstrating that the precise incorporation of charge-neutral linkages[128,129] or 2′-modifications[127] can greatly enhance the therapeutic indexes of PS–ASOs, the establishment of an on-resin carbamate formation protocol would be highly desirable to enable the synthesis of large sets of mixed-backbone and 2′-modified chimeras to facilitate a deeper understanding for potential antisense applications.

# CHAPTER III

# BUILDING BLOCKS FOR THE SYNTHESIS OF AMIDE-LINKED OLIGONUCLEOTIDES

# 1. Background

## 1.1. The Amide Internucleoside Linkage

### 1.1.1. Biophysical Properties

The 4-atom amide linkages **I** and **II** were first described in 1993 as an internucleoside backbone surrogate to replace the natural 4-atom phosphodiester bond (Figure III-1).[63,64] Similar to the carbamate linkage of the previous chapter, the amide modification can be introduced into oligonucleotides *via* amide-linked

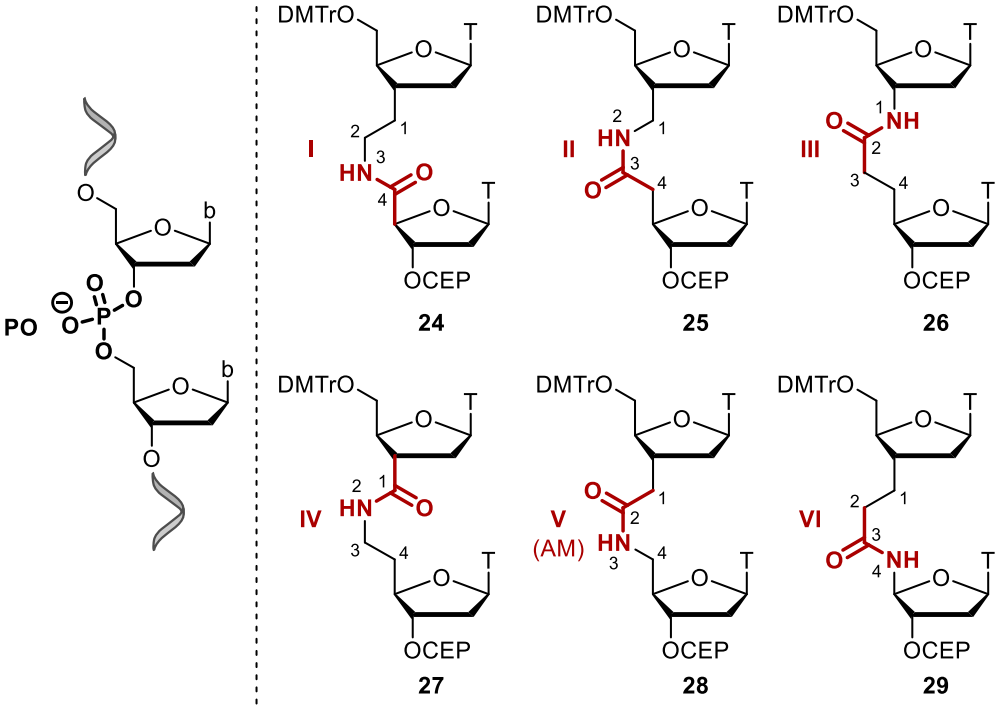



**Figure III-1: Summary of possible 4-atom amide internucleoside linkages.**
The natural PO and amides (**I**–**VI**) are 4-atom internucleoside linkages. Dimer CE phosphoramidites **24**–**28** have been reported in the literature (**24**;[63] **25**;[65] **26**;[64]; **27**;[66] and **28**[62,67]). Dimer phosphoramidite **29** is theoretical and has not yet been reported. The linker length is defined as the minimum number of atoms connecting the sugar rings as indicated. Abbreviations: AM, amide internucleoside linkage **V**; b, nucleobase; CEP, 2-cyanoethyl *N*, *N*-diisopropyl phosphoramidite; DMTr, 4,4′-dimethoxytrityl; T, thymine.

dinucleoside CE phosphoramidites. Indeed, a diverse set of dinucleoside CE phosphoramidites (**24**,[63] **25**,[65] **26**,[64] **27**,[66] and **28**;[62,67] Figure III-1) has been reported to introduce amides of type **I**, **II**, **III**, **IV** and **V** during the solid phase synthesis of amide-modified oligonucleotides. Dimer CE phosphoramidite **29** has not yet been reported in the literature and is a hypothetical consideration to introduce amide **VI** to complete the set of theoretically possible 4-atom amide internucleoside linkages.

Some of the reported amide-modified oligonucleotides were further evaluated for their binding affinity towards an RNA target (Table III-1). With exception of type **II** amides, incorporation of the unnatural phosphodiester mimics (**I**–**IV**)

**Table III-1: Summary of reported 4-atom amide backbones and their effects on duplex stabilities with RNA and DNA targets.**

| | | | RNA target | DNA target |
|---|---|---|---|---|
| Type | ON | Sequence (5'→3') | $\Delta T_m$[a] (°C)/mod | $\Delta T_m$[a] (°C)/mod |
| **I** | ON18 | CTCGTACCT$^{I}$TTCCGGTCC | −1.3[63] | n. a. |
| | ON19 | GCGT$^{I}$TT$^{I}$TT$^{I}$TT$^{I}$TT$^{I}$TGCC | −2.8[63] | n. a. |
| **II** | ON20 | CTCGTACCT$^{II}$TTCCGGTCC | −0.5[65] | n. a. |
| | ON21 | GCGT$^{II}$TT$^{II}$TT$^{II}$TT$^{II}$TT$^{II}$TGCC | +0.4[65] | −1.4[65] |
| **III** | ON22 | CTCGTACCT$^{III}$TTCCGGTCC | −2.2[64] | n. a. |
| | ON23 | GCGT$^{III}$TT$^{III}$TT$^{III}$TT$^{III}$TT$^{III}$TGCC | −3.9[64] | n. a. |
| **IV** | ON24 | CTCGTACCT$^{IV}$TTCCGGTCC | −3.8[66] | n. a. |
| | ON25 | CTCGTACT$^{IV}$TT$^{IV}$TCCGGTCC | −3.3[66] | n. a. |
| **V** | ON26 | CTCGTACCT$^{V}$TTCCGGTCC | +0.9[62] | n. a. |
| | ON27 | GCGT$^{V}$TT$^{V}$TT$^{V}$TT$^{V}$TT$^{V}$TGCC | −0.3[62] | −1.5[62] |
| | ON28 | T$^{V}$TT$^{V}$TT$^{V}$CT$^{V}$CT$^{V}$CT$^{V}$CT$^{V}$CT | +0.6[67] | −0.3[67] |
| | ON29 | T$^{V}$T$^{V}$T$^{V}$T$^{V}$T$^{V}$C̲TCTCTCTCT | +0.3[55] | −0.6[55] |
| | ON30 | TTTTTCTCTC̲$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$T | −0.7[55] | −0.5[55] |
| | ON31 | TTTTT$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$TCTCT | −0.8[55] | −0.5[55] |
| | ON32 | pT$^{V}$T$^{V}$T$^{V}$T$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$T$^{V}$C̲$^{V}$Tp | −0.4[55] | −0.2[55] |

$^{I–V}$ denote amide internucleoside linkages (see Figure III-1); C̲ = 5-methylcytosine; p = phosphate.
[a] the $\Delta T_m$ is the difference of the observed $T_m$ to the $T_m$ of the isosquential unmodified control with the same target. The presented $\Delta T_m$'s are normalised to the number of modifications.

generally destabilise the duplexes formed by the modified oligonucleotides with an RNA target.[55,62–67] Intriguingly, a single type **V** amide can stabilise hybridisation to an RNA target by +0.9 °C (ON26; Table III-1).[62] Multiple incorporations of type **V** amides into an oligonucleotide can either increase or slightly decrease RNA target affinity, depending on the sequence (ON27 and ON28; Table III-1).[62,67] The type **V** amide internucleoside linkage is the most explored among the types represented in Figure III-1. Indeed, *N*-alkylated analogues,[62] exploration towards higher sequence variation (T–T,[62] T–C[67] and C–T[184] dimers), combinations with 2′-substituted sugars,[68,185–190] and consecutive incorporations[55,189,190] (see ON29–ON32; Table III-1) have been reported. Moreover, incorporation of type **V** amides was shown to increase resistance against nucleolytic degradation of the amide-modified oligonucleotides.[62] Molecular dynamic simulations[62,191,192] and X-ray crystal structures[193] of amide **V**-modified oligonucleotides in duplex with their RNA targets revealed a high structural integrity of the backbone surrogate to mimic the natural phosphodiester bond. Although excellent biocompatibility of amide **V** was reported for transcription[194] and RNAi,[193,195] possible antisense applications in the context of RNase H activation have not been reported and is focus of Chapter IV.

This chapter (Chapter III) focuses on the preparation of 3′-homologated 2′-deoxy mono- or dinucleosides for the synthesis of amide-modified oligonucleotides bearing the type **V** amide backbone, which is hereafter referred to as AM or simply amide.

### 1.1.2. Synthesis

Nucleoside building blocks are complex and chiral molecules that can have various nucleobases. Naturally occurring nucleosides are readily available which makes them a desirable starting material for the synthesis of unnatural nucleosides using a chiral pool strategy. This strategy was demonstrated for the LNA monomers in Chapter II to synthesise carbamate-linked LNA oligonucleotides. The nucleoside building blocks for the formation of the amide bond are shown in Figure III-2. The amide-linked dimer can be disconnected at the amide bond giving carboxylic acid **I** and amine **II** as monomeric fragments. Whilst the natural 3′- and 5′-heteroatom functionality is retained in amine **V**, carboxylic acid **I** contains a 3′-methylene. Hence, amine **II** can be obtained by functional group interconversion (FGI) from a chiral pool, but disconnection of the 3′-carboxymethyl in **I** requires the formation of a C−C bond and loss of the natural 3′-heteroatom functionality. This step is the major synthetic challenge for the preparation of this building block. The first synthesis to

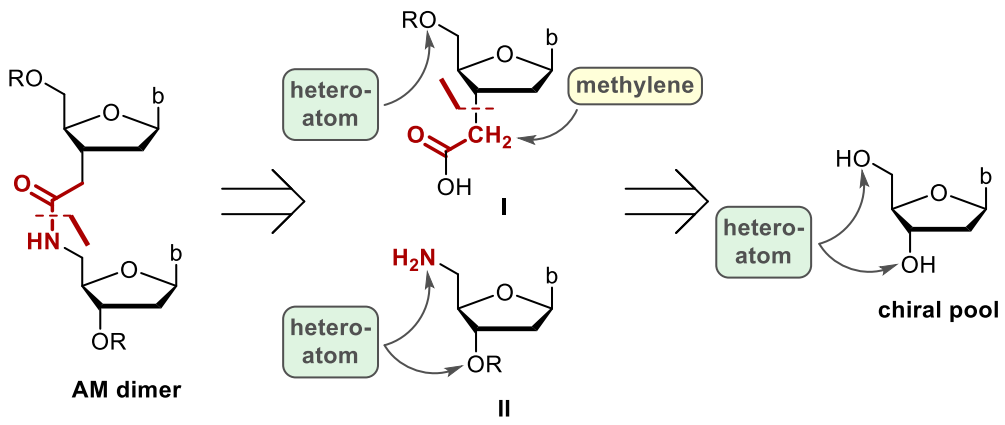
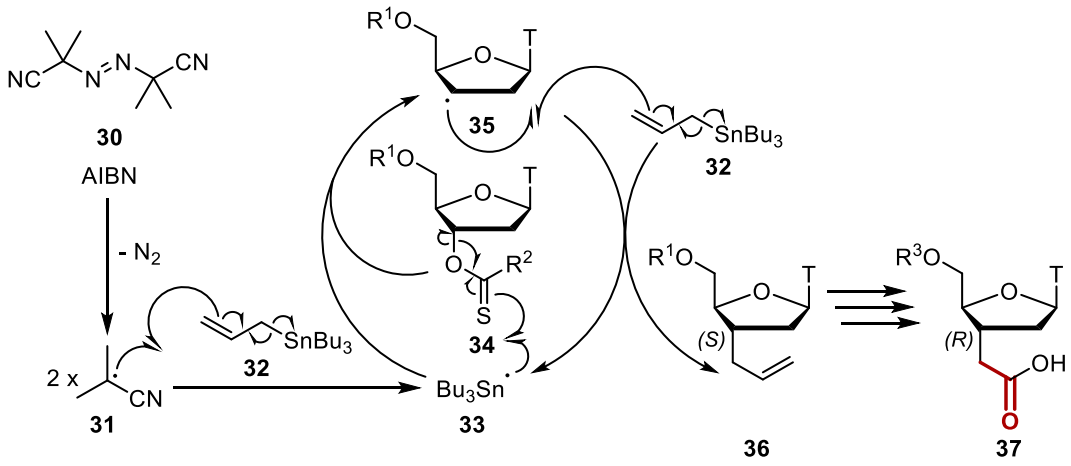



**Figure III-2: Retrosynthetic analysis of the amide internucleoside linkage.**
The amide bond is highlighted in red. The 3′- and 5′-functionalities with respect to their chiral pool counterparts are indicated with arrows. Abbreviations: b, nucleobase.




**Scheme III-1: Radical chemistry for C3′-homologisation.**
$R^1$ = *tert*-butyldimethylsilyl (TBS), $R^2$ = OPh, $R^3$ = *tert*-butyldiphenylsilyl (TBDPS) Abbreviations: AIBN, azobisisobutyronitrile; T, thymine.

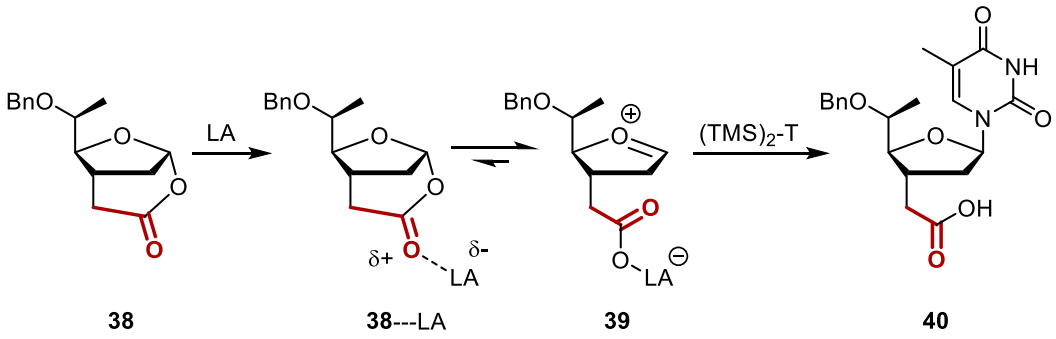
establish this 3′-C−C bond was reported in 1989 where the authors used free radical chemistry to stereoselectively install a 3′-allyl group (Scheme III-1).[196] The reaction follows a Barton–McCombie deoxygenation mechanism and starts with the generation of a tributylstannyl radical **33** upon decomposition of AIBN as a radical initiator. This leads to homolytic cleavage of the 3′-C–O bond of xanthate **34** to form nucleoside radical **35**. The unpaired electron in **35** pairs with an electron of the allyl substituent of allyltributyltin (**32**) to form the 3′-C–C bond regenerating a tributylstannyl radical **33** that enters the next cycle. The Barton–McCombie-type homologisation was later adapted by De Mesmaeker et al. to convert allyl nucleoside **36** into acid **37**.[62,63] This reaction sequence (not shown) includes osmium tetroxide-mediated cleavage of the C=C bond followed by oxidation of a resulting aldehyde to the carboxylic acid in **37**. Interestingly, the stereoselective outcome of the 3′-C–C bond formation exclusively gives *S*-isomer **36**[196,197] (note: the orientation of the 3′-substituent in *R*-isomer **37** is retained but named differently due to changes

in substituent priorities). The observed stereoselectivity could be explained by steric factors of the bulky TBS-group hindering addition of radical **34** to allyltributyltin (**35**) from its β-face.[196] Indeed, addition of riboside radicals having a 2′-substituent adjacent to the radical centre result in inverse stereoselectivity giving the undesired *R*-epimer of **36**.[68] It is important to emphasise that C3′-homologated nucleoside synthesis by the free radical method (Scheme III-1) is solely reported for pyrimidine bases (C, $^{Me}$C, U, T). Among those, $^{Me}$C- and C-nucleosides were obtained after the radical-mediated 3′-C–C bond formation *via* conversion of C3′-homologated T- and U-bases, respectively.[55,67,184] This limitation can be explained by the general incompatibility of protection groups during the radical reaction. In an attempt to expand the base variation of C3′-homologated nucleosides to purines, Balkrishen et al. observed a significant loss of standard purine protection groups (benzoyl, isobutyryl) during the radical-mediated 3′-C–C bond formation, leaving a mixture of products presumably from unpredictable reactions by fragments formed from protecting group cleavage.[198] Moreover, amidine protecting groups (acetamidine, formamidine) resulted in similar side reactions leading to low yields and tedious purifications.

Other strategies to introduce the 3′-C–C bond include addition of a Grignard reagent to a preformed C3′-ketone followed by a Barton–McCombie deoxygenation.[199] However, C3′-ketone deoxyribosides are unstable as they undergo elimination of the bases even under mild conditions,[200–202] and a 2′-functionality of ribosides directs the deoxygenation towards formation of the undesired epimer.[202] A widely adapted strategy is the Wittig reaction of a 3′-ketone

followed by hydrogenation to introduce the carboxymethyl group on ribosides.[203–206] However, this approach also relies on a 3′-ketone riboside and is thus not compatible for deoxyribosides due to elimination of the base as outlined above.

A rather different approach is to modify an abasic sugar moiety which can then serve as a glycosyl donor in a later step to introduce the nucleobase. This strategy is exemplified in Chapter II where modified ribose analogue **14** (see Scheme II-2) was used as the glycosyl donor in the Vorbrüggen glycosylation reaction. Similar strategies have been reported for C3′-homologated ribosides. However, this strategy only works particularly well with a 2′-substituent that can direct the glycosylation step to form the β-anomer through neighbouring group participation. Hence, C3′-homologated deoxyribosides require a different approach to ensure the correct stereochemical outcome from the glycosylation reaction. To this end, Wendeborn et al. reported bicyclic glycosyl donor **38** for the stereoselective installation of a thymine base (Scheme III-2).[207] Bicyclic lactone **38** can be activated as a glycosyl donor by addition of a Lewis acid to form oxonium ion **39**. The authors found that




**Scheme III-2: Late-stage glycosylation of a bicyclic lactone.**
After activation of lactone **38** by a Lewis acid (LA), the equilibrium is shifted towards oxonium ion **39**. The bulkiness of the LA determines the stereoselectivity of the glycosylation reaction. Abbreviations: Bn, benzyl; T, thymine; TMS, trimethylsilyl. *Scheme adapted from literature.*[207]

the use of a bulky Lewis acid can direct the glycosylation towards exclusive formation of the desired β-anomer **40**. This approach has the potential to allow the stereoselective glycosylation of different bases, thereby facilitating sequence variation of amide-modified oligonucleotides for antisense studies. This late-stage divergence of the synthesis also avoids the use of toxic stannanes or osmium tetroxide.

# 2. Aim of the Chapter

The aim of this chapter is the preparation of chemically modified di- or mononucleosides to facilitate the incorporation of the amide backbone as an unnatural internucleoside linkage in oligonucleotides for further antisense studies. Considering the synthetic challenges that come with 3′-homologated 2′-deoxynucleosides, the synthesis is initially focused on thymidines (Figure III-3). This aim comprises two parts: the synthesis of an amide-linked dinucleoside CE phosphoramidite **I** to synthesise amide-modified oligonucleotides *via* the monomer approach (Part I); and the synthesis of an amino–carboxylic acid monomer **II** for consecutive on-resin amide bond formations (Part II).

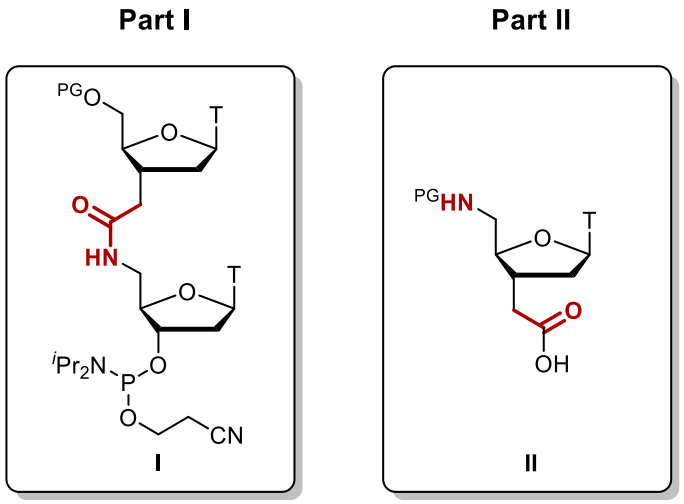



**Figure III-3: Target structures for Part I and II of this chapter.**
The two target structures of this chapter are shown, and the amide bond is highlighted. Abbreviations: $^{i}$Pr, isopropyl; PG, protecting group; T, thymine.

**Part I: *En bloc* amide addition.** The first part of the chapter is aiming at the synthesis of an amide-linked 2′-deoxydinucleoside CE phosphoramidite **I** for standard solid phase oligonucleotide synthesis. This dinucleoside contains the amide bond and will allow the straightforward incorporation of isolated amides at various sites within an oligonucleotide for further tests in antisense studies. Therefore, the 5′-protecting group in **I** should be acid-labile (e.g., 4,4′-dimethoxytrityl, or 4-methoxytrityl).

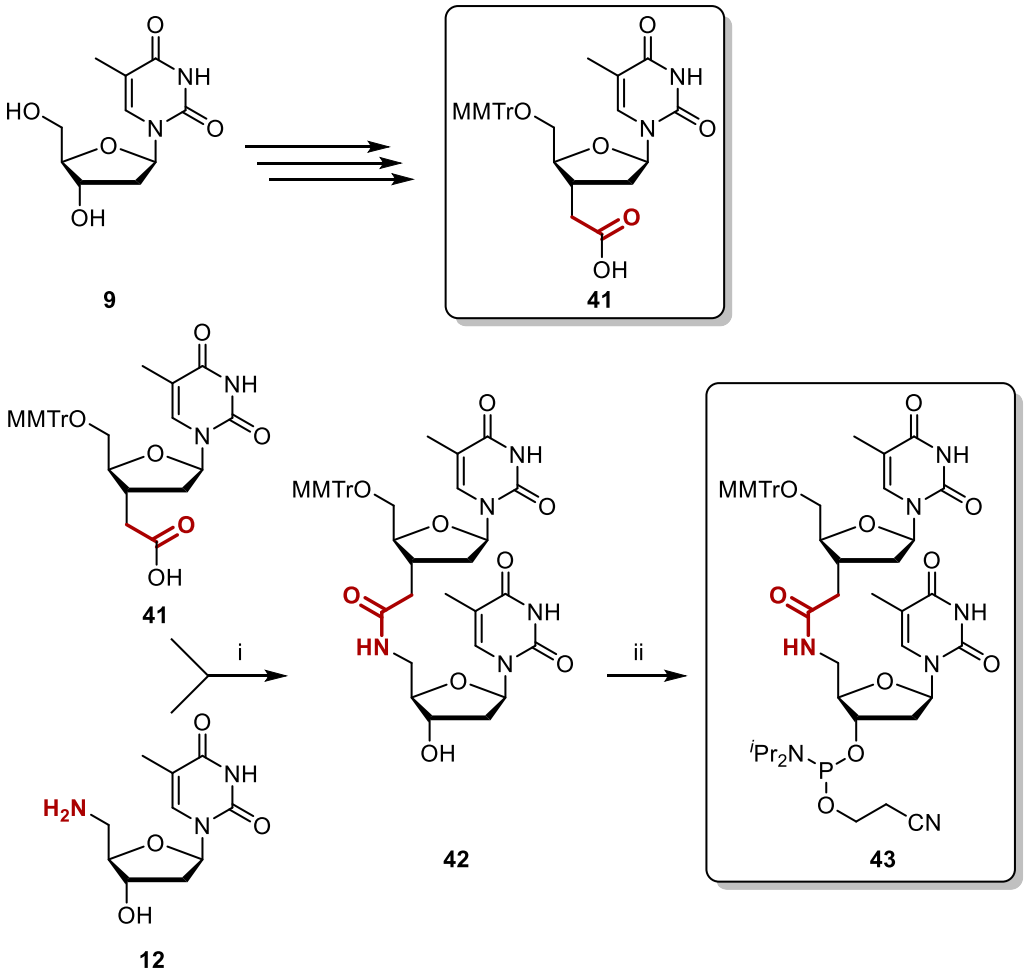
**Part II: On-resin amide formation.** The second part of this chapter is the synthesis of a 2′-deoxynucleoside **II** bearing a 3′-carboxymethyl and a 5′-nitrogen functionality. This monomer will allow the controlled on-resin amide bond formation for the synthesis of oligonucleotides with consecutive amide linkages. Firstly, a suitable 5′-nitrogen protecting group must be determined to ensure amide bond formation in the 3′→5′ direction. Determining a suitable 5′-nitrogen protecting group requires consideration of its deprotection conditions and its compatibility with the solid phase oligonucleotide synthesis protocol. The successful integration of on-resin amide couplings into the solid phase oligonucleotide synthesis protocol would allow a flexible approach to synthesise AM–PO chimeras for further antisense studies.

# 3. Results and Discussion

## 3.1. Part I: Synthesis of an Amide-Linked Dinucleoside Phosphoramidite Dimer for *En Bloc* Amide Addition.

Chemically modified internucleoside linkages can be introduced *via* nucleoside dimers containing the desired chemical modification for standard oligonucleotide synthesis (see Section 3.3.1.2; Chapter I). Introduction of the amide internucleoside linkage *via* the dimer strategy has been previously reported for oligonucleotides containing isolated amides.[62,67,68,184] Dinucleoside phosphoramidite **43** was prepared *via* a convergent synthesis route starting from thymidine (**9**) (Scheme III-3): The first building block—acid **41**—was synthesised[j] from thymidine (**9**) according to published procedures.[62,67,68,184] Because of stability concerns regarding the acid-labile 5′-protection in the presence of a free carboxylic acid, the more stable 4-methoxytrityl (MMTr) was chosen over the conventional 5′-*O*-DMTr protecting strategy. Indeed, each *para*-OMe substituent at the trityl protecting group increases its cleavage rate approximately 10 times.[208] The second building block—amine **12**—was synthesised as previously described in Scheme II-1; Chapter II. Amide coupling of amine **12** and acid **41** gave dimer **42** which was then phosphitylated to phosphoramidite **43** for solid phase oligonucleotide synthesis.

[j] Monomer **41** synthesised by Cameron Thorpe (DPhil), University of Oxford.




**Scheme III-3: Convergent synthesis of dinucleoside 43 for solid phase oligonucleotide synthesis.**
Reagents and conditions: (i) 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU), DIPEA, DMF, rt, 1 h, 69%, (ii) 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, $CH_2Cl_2$, rt, 1 h, 56%. Abbreviations: $^{i}$Pr, isopropyl; MMTr; 4-methoxytrityl.

## 3.2. Part II: Synthesis of an Amino–Carboxylic Acid Nucleoside for Consecutive On-Resin Amide Bond Formations

The highest percentage of backbone modifications using the dimer strategy is 50% in the case of alternating phosphodiester and amide linkages (see ON28;

Table III-1). This example would require oligonucleotide synthesis from nucleoside CE phosphoramidite dimers only. Moreover, 16 dimers would be required to cover all possible sequences with the standard nucleobases A, C, G and T (Figure III-4). The number of required monomers could be reduced by on-resin amide bond formation: A carboxylic acid nucleoside could be coupled to a 5′-amine of the growing oligonucleotide, which itself can be introduced *via* the coupling of a 5′-amino-modified CE phosphoramidite monomer. Hence, the number of required monomers to cover all sequence possibilities for oligonucleotides with alternating

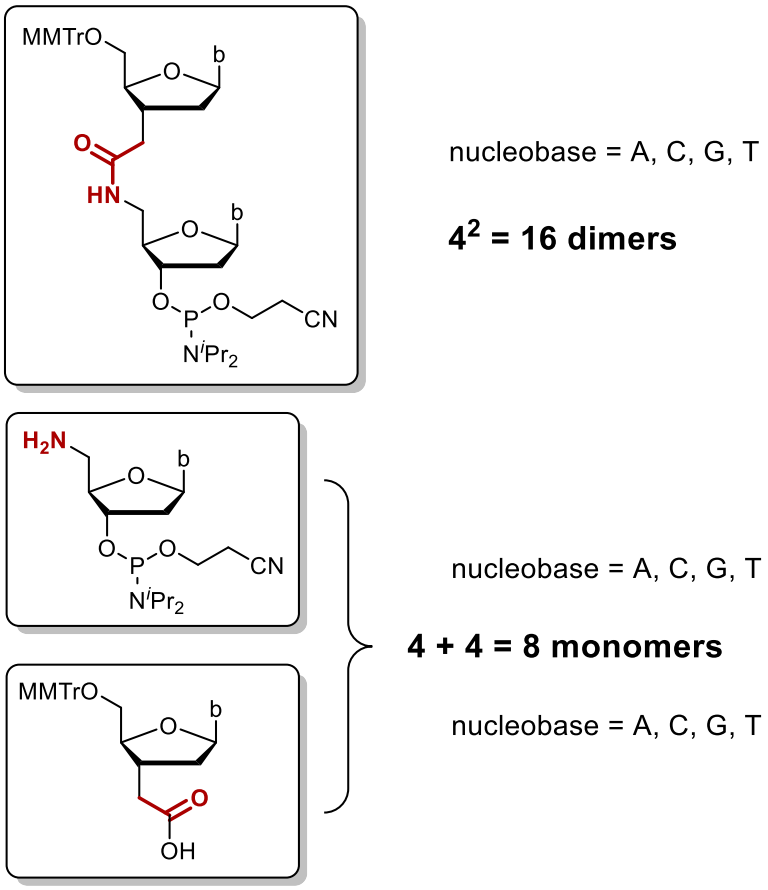



**Figure III-4: Number of required mono- and dinucleosides to cover all possible sequences for the synthesis of amide-linked oligonucleotides.**
Assuming four different nucleoside nucleobases (excluding chemical or epigenetic modifications), the number of required CE phosphoramidites is 16 or 8 for the dimer strategy or on-resin amide bond formation, respectively. Abbreviations: b, nucleobase; $^{i}$Pr, isopropyl; MMTr, 4-methoxytrityl.

phosphodiester and amide linkages would be reduced to 8 monomers (Figure III-4). However, both strategies would still not allow for the synthesis of oligonucleotides with consecutive amides.

### 3.2.1. Determination of a Suitable 5′-Protecting Group

The synthesis of oligonucleotides with consecutive amides requires a different nucleoside monomer having both an amino and a carboxylic acid functional group—i.e., amino–carboxylic acid. It is crucial that the chosen protecting group for the amino–carboxylic acid nucleoside is compatible with amide bond formations and standard phosphoramidite oligonucleotide synthesis. Moreover, orthogonal protection is required at the 5′-end to comply with the directionality of solid phase oligonucleotide synthesis (3′→5′) allowing the synthesis of backbone chimeras by combining phosphoramidite and amide chemistry. Possible 5′-amino protecting groups are MMTr and Fmoc as they might be compatible with both oligonucleotide and peptide synthesis. The acid-stable Fmoc protecting group might be promising to circumvent possible self-deprotection in the presence of a free 3′-carboxylic acid. Thus, the compatibility of the Fmoc protecting group with solid phase oligonucleotide synthesis was initially evaluated in a simplified model system (Figure III-5). The general strategy involved coupling and deprotection cycles of Fmoc-protected amino acids to a 5′-amino-modified oligonucleotide on a solid support. (**A**) The CE protecting groups of the immobilised oligonucleotide strand were removed by an initial wash with *N*, *N*-diethylamine (DEA) to prevent Michael addition of any amino groups to acrylonitrile. (**B**) The carboxylic acid of Fmoc-

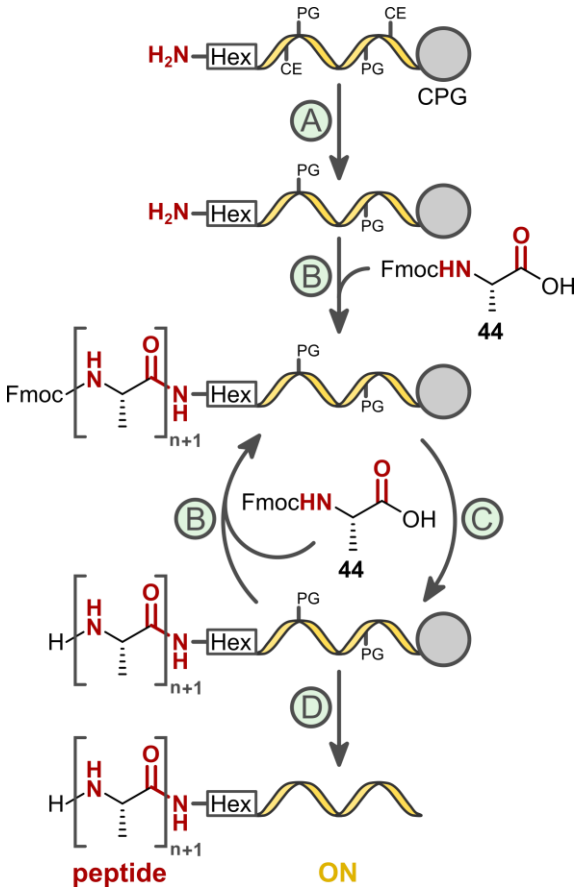



**Figure III-5: Fmoc strategy for the synthesis of oligonucleotide–peptide chimeras.** Reagents and conditions reported for 0.5 µmol scale: (**A**) DEA wash: [DEA/MeCN, 1/4 (v/v), 1 mL], rt, 20 min; (**B**) coupling: (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP) (10 eq in 300 µL anhyd DMF), Fmoc-L-alanine **44** (10 eq, stock: 100 mM in anhyd DMF), *N*-methylmorpholine (NMM) (3 eq), rt, 1.5 h; (**C**) deprotection: [piperidine/DMF, 1/4 (v/v), 2 mL], rt, 20 min; (**D**) cleavage: concd aq $NH_3$ (4 mL), 55 °C, 5 h. Abbreviations: CE, 2-cyanoethyl; CPG, controlled pore glass; Hex, *n*-hexyl linker; PG, protecting group.

protected alanine **44** was activated with PyBOP for amide bond formation with the terminal amine of the growing chain. (**C**) Successive deprotection with piperidine should result in the formation of a terminal amino group that can serve as a substrate for further coupling cycles (**B**). In total, four coupling–deprotection cycles were performed to evaluate the feasibility of the Fmoc strategy to form peptide–oligonucleotide chimeras. (**D**) The resulting chimera was then cleaved from the solid support and analysed by UPLC-MS (Figure III-6).

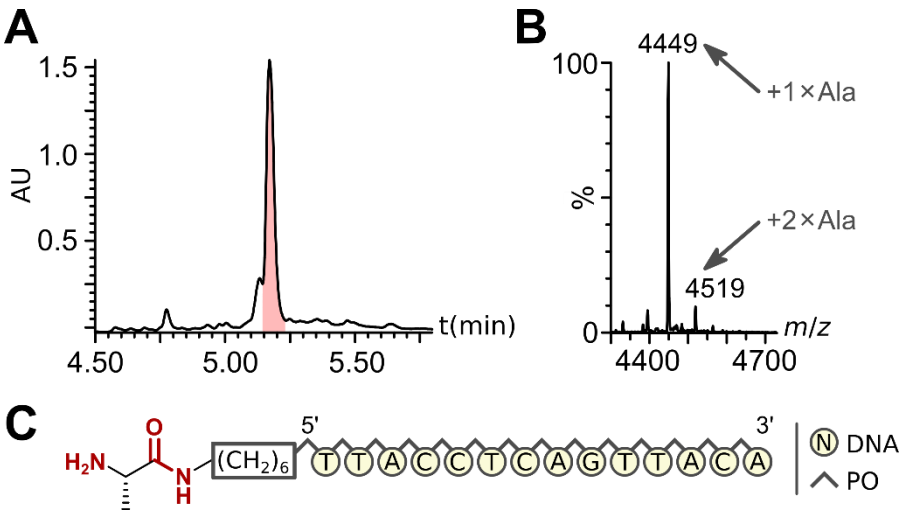


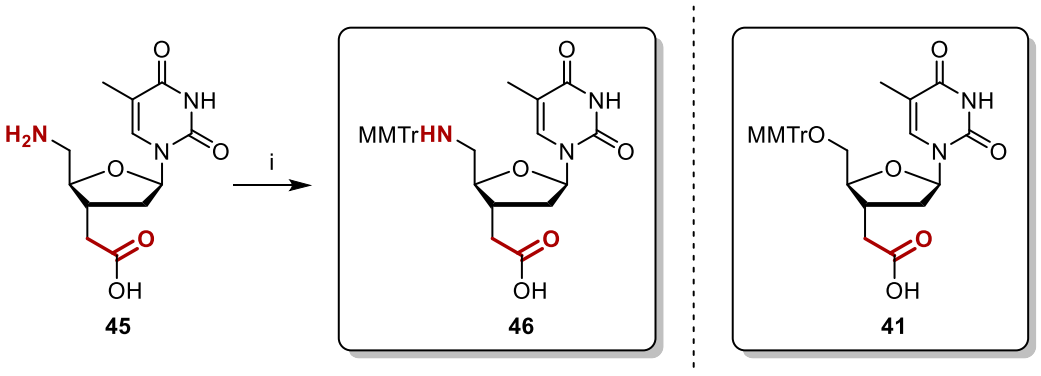

**Figure III-6: Analysis of oligonucleotide–peptide chimeras.**
(**A**) HPLC analysis of oligonucleotide–peptide chimeras that were cleaved off the resin and deprotected following four coupling cycles with Fmoc-L-alanine **44**. The main peak is highlighted. (**B**) MS analysis of the main peak from (**A**). The masses corresponding to the coupling of one and two L-alanines are indicated. (**C**) Structure of the major oligonucleotide–peptide chimera obtained after four couplings.

The mass of the main UPLC peak (4449 Da; Figure III-6A, B) confirmed coupling of a single alanine (Figure III-6C) instead of the expected addition of four alanines. Hence, the Fmoc strategy was considered unfeasible for this purpose, and further exploration of a suitable protecting group strategy was required. It is important to have an orthogonally protected monomer that can be activated for amide bond formation and successively deprotected for further couplings under conditions that are compatible with standard solid phase oligonucleotide synthesis. The MMTr group is an excellent candidate for said requirements because it is stable during activation of the carboxylic acid and during the oligonucleotide synthesis cycles as previously reported for the synthesis of AM–PO backbone chimeras.[55] Indeed, further experiments in our group showed that a 5′-terminal MMTr protecting

group for the amino functionality resulted in excellent coupling of MMTr-protected amino acids and formation of peptide–oligonucleotide chimeras.[k]

### 3.2.2. Synthesis of a Suitably Protected Amino–Carboxylic Acid Nucleoside

Based on the results of the peptide coupling experiments described above, MMTr-protected amino–carboxylic acid **46** (Scheme III-4) was chosen as target monomer for the synthesis of oligonucleotides with consecutive amides. To this end, tritylation of previously prepared amine **45**[l] gave **46** with the desired MMTr-protecting group for integration of on-resin amide bond formations into the solid phase oligonucleotide synthesis protocol. Importantly, no self-cleavage of the 5′-MMTr in the presence of a carboxylic acid was observed for compound **46**. This agrees with similar reported molecules having an MMTr-protected primary amine and a free carboxylic acid.[209,210] In addition to amino–carboxylic acid **46**, 5′-*O*-




**Scheme III-4: Protection of amino carboxylic acid 46.**
Reagents and conditions: (i) 4-methoxytriphenyltrityl chloride, $Et_3N$, 4-dimethylaminopyridine, pyridine, 16 h, rt, 30%. Abbreviations: MMTr; 4-methoxytrityl.

[k] Conditions optimised by Dr Ysobel R. Baker, University of Oxford.
[l] Monomer **45** synthesised by Cameron Thorpe (DPhil), University of Oxford.

MMTr protected carboxylic acid **41** is a valuable analogue for non-consecutive on-resin amide couplings, or to transition from on-resin amide couplings to standard solid phase oligonucleotide synthesis. This will be discussed in detail in Chapter IV.

# 4. Conclusion

The preparation of chemically modified nucleosides for the introduction of amide backbones during oligonucleotide synthesis requires a challenging C–C bond formation. Therefore, the synthetic endeavours were focused on thymidine analogues according to previously established synthesis strategies. This chapter reports the synthesis of chemically modified thymine nucleosides that will be valuable for the synthesis of amide-modified oligonucleotides. These include dimer CE phosphoramidite **43** and MMTr-protected amino–carboxylic acid **46**, that can be used to introduce amide bonds *via* the dimer and monomer strategies, respectively. Importantly, the 5′-MMTr protecting group was determined as being suitable for integration of amide couplings into the solid phase oligonucleotide synthesis protocol—as opposed to the 5′-Fmoc protecting group. As outlined in the background section of this chapter, the synthetic preparation of 3′-homologated DNA-type nucleosides with all four bases is challenging and indeed, the nucleosides prepared in this chapter were limited to thymidine analogues. The next chapter will describe the application of successfully prepared mono- and dinucleosides for the synthesis of chemically modified oligonucleotides with an amide backbone. Moreover, a discussion about a future direction for the synthetic accessibility of amino–carboxylic acid nucleosides with different bases will be provided at the end of Chapter IV, which will provide the context for the underlying structural considerations.

# CHAPTER IV

# AMIDE-LINKED OLIGONUCLEOTIDES FOR RNASE H ACTIVATION

# 1. Background

## 1.1. RNase H in Antisense Applications

Antisense oligonucleotides are emerging as versatile and powerful tools to combat hard-to-treat diseases. Activation of RNase H to degrade an mRNA target is among the most clinically relevant antisense mechanisms (see Table I-2; Chapter I).[102,103,108,121,132,134]

### 1.1.1. RNase H Mechanism

The activation of RNase H leads to the sequence specific degradation of an mRNA target whilst specificity is conferred by the ASO sequence.[17] The overall mechanism of the catalytic RNase H cycle is shown in Figure IV-1. The ASO must enter the cell where RNase H-mediated degradation of the mRNA happens (cytosol and nucleus).[99] As mentioned in Chapter I, cellular uptake of oligonucleotides remains a major therapeutic challenge.[93] Once inside the cell, the ASO must finds its mRNA target to form an ASO:mRNA heteroduplex through Watson–Crick base pairing. The ASO:mRNA duplex can be recognised and bound by the ubiquitously present RNase H, which selectively degrades the mRNA strand. RNase H and the ASO are then released and can enter the next catalytic cycle with a new, intact mRNA.[99–101] The degraded mRNA can no longer be translated into a functional protein and thus, RNase H-activating ASOs lead to the knockdown of a target

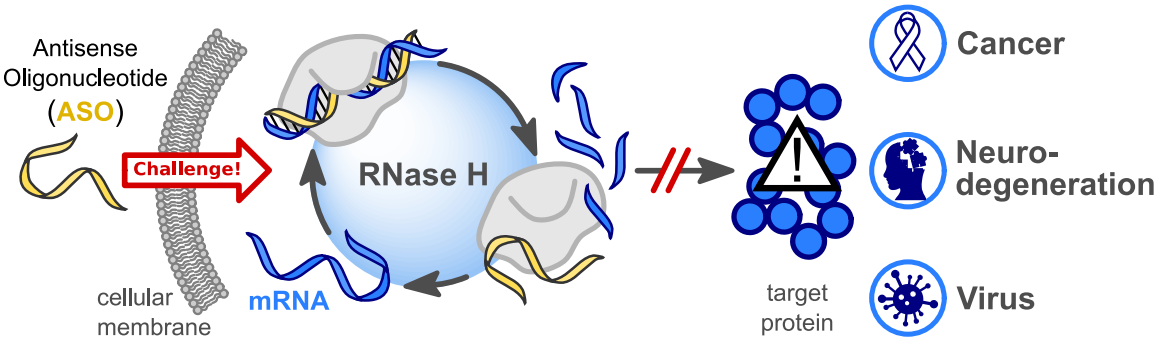



**Figure IV-1: RNase H-mediated mRNA degradation for antisense applications.**
Cellular uptake remains a major challenge for the efficacy of ASOs. Once inside the cell, ASOs can bind to their mRNA target through Watson–Crick base pairing to trigger an RNase H response. RNase H recognises and binds an ASO:mRNA duplex and degrades the mRNA by hydrolysis of PO bonds. Degradation of the mRNA coding for a therapeutically relevant target protein prevents translation into the protein and has implications for several diseases such as cancers, Huntington's disease, viral infections and others.[89]

protein. The modulation of protein expression at the mRNA level has implications for the treatment of various diseases such as cancers, neurodegeneration, or viral infections.[89]

The RNase H approach relies on the fact that the ASO takes the part of the DNA in a DNA:RNA heteroduplex to trigger the RNase H-mediated degradation of a complementary RNA. Thus, the characteristics of the ASO and the resulting ASO:RNA duplex are essential to elicit an RNase H response. Although RNase H can bind to all hybrids, i.e., DNA:DNA, RNA:RNA and DNA:RNA duplexes, only the latter is a substrate for the subsequent hydrolysis of the RNA strand.[211] Previous reports showed that various structural requirements of the double helix determine RNase H activation: The O4′-*endo* sugar pucker lies between the C2′- and C3′-*endo* conformations (Figure IV-2) and was determined as a commonly adapted conformation of the ASO strand.[212,213] Importantly, the O4′-*endo* sugar pucker

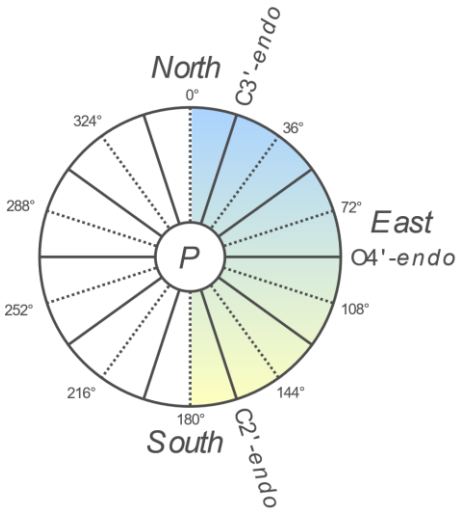



**Figure IV-2: The pseudorotation cycle of the different sugar puckers and their nomenclature.** Because of their interdependency, the dihedral angles determining the sugar pucker can be described with a single pseudorotation angle *P*. The relevant descriptions of different sugar puckers are indicated according to their pseudorotation angles. *Figure adapted from literature.*[213]

results in a duplex conformation with a large minor groove width—closer to the A-form RNA:RNA than the B-form DNA:DNA duplex—which is important for substrate recognition.[214,215] However, bicyclic sugar analogues that are locked into the O4′-*endo* conformation do not elicit an RNase H response.[215] Thus, a flexible sugar pucker is essential for RNase H activation.[212] Indeed, later studies confirmed that RNase H activation occurs with a transition from the typical A-form DNA:RNA heteroduplex to a more B-like conformation—i.e., A→B transition.[216]

Discrimination between the ASO and the RNA target results from specific interactions with the phosphodiester bond-binding pocket (PO-BP) and the catalytic centre of RNase H.[94,217] The latter involves hydrogen bonds between amino acid residues and the 2′-OH groups of four consecutive ribosides inside the catalytic centre of RNase H. The interactions of human RNase H with the DNA strand of a

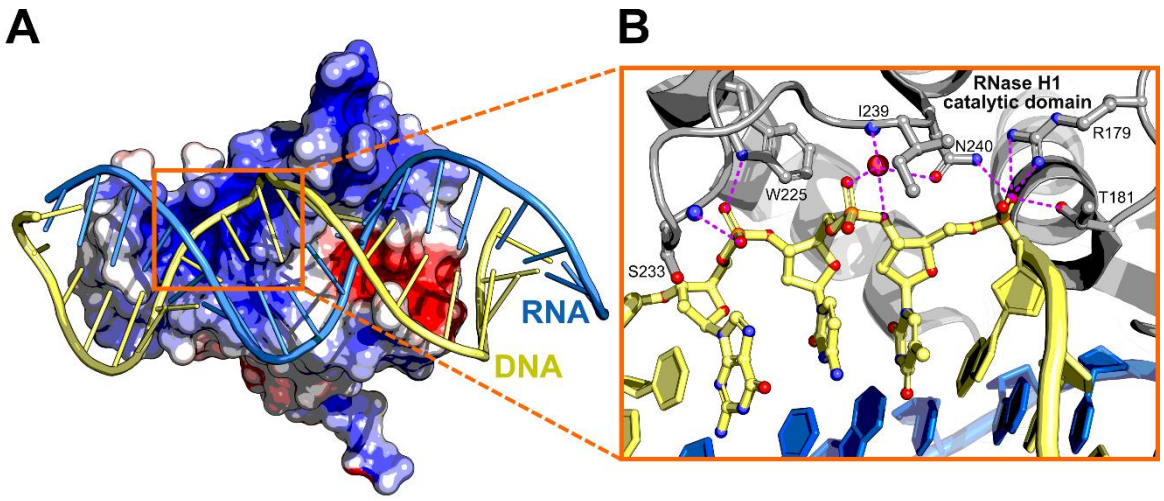



**Figure IV-3: Substrate recognition by the catalytic domain of human RNase H1 (h-RNase HC).** (**A**) Representation of the surface potential of h-RNase HC (PDB, 2QK9) in complex with a 20-mer DNA:RNA duplex (aligned from PDB, 2QKB). Negative, neutral and positive surface potentials are coloured in red, grey and blue, respectively. The DNA and RNA strands are represented as cartoon structures. (**B**) Close-up view of the PO-BP of h-RNase HC (PDB, 2QKB). The protein is represented as a grey cartoon structure and the PO-binding residues are highlighted in a ball-and-stick representation. The three consecutive DNA nucleotides interacting with the PO-BP are shown as a stick and ball structure and interactions with h-RNase HC are represented by dotted lines. *Figure adapted from literature.*[94,217]

DNA:RNA duplex are detailed in Figure IV-3 and Figure IV-4. The DNA strand is buried deep inside the PO-binding pocket of RNase H (box; Figure IV-3A). The surface potential indicated in Figure IV-3A shows that the PO-binding pocket is positively charged, facilitating ionic interactions with the negatively charged phosphodiester backbone of the DNA strand. The detailed interactions of the DNA strand with the PO-binding pocket of human RNase H are shown in Figure IV-3B. In detail, three consecutive phosphates of the DNA strand are interacting through hydrogen bonds with the peptide backbone and amino acid residues of RNase H. These interactions form the basis for the ASO designs later in this chapter. It is worth mentioning that the PO-binding pocket creates a chiral environment and the diastereotopic oxygen atoms of the phosphodiester backbone form different

interactions with RNase H. This affects the design of chemically modified ASOs—e.g., PS-modified ASOs.[94]

A different angle of the DNA strand inside the PO-binding pocket is shown in Figure IV-4. Expansion of the middle DNA nucleotide (box; Figure IV-4A) and rotation by roughly 90° gives Figure IV-4B. The enlarged and rotated view shows how perfectly the DNA strand fits into the PO-binding pocket. Indeed, any 2′-substituent of a ribose analogue would clash with tryptophan 221, which explains why RNase H only tolerates DNA-type ASOs.[217] This steric clash of (modified) riboses limits the range of RNase H-compatible ASO designs.

In summary, the activation of RNase H is highly susceptible to chemical modifications of the ASO: (**i**) a flexible structure is required to enable the A→B transition of the ASO:RNA heteroduplex.[212] (**ii**) The possibility to form hydrogen

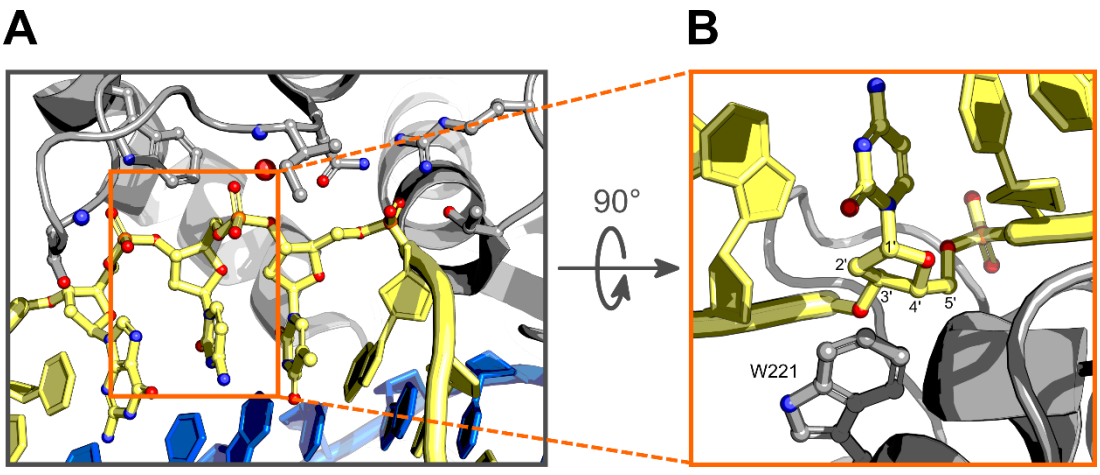



**Figure IV-4: Sugar specificity of RNase H for the ASO strand.**
(**A**) PO-BP of h-RNase HC from Figure IV-3B. The protein, DNA and RNA are represented as cartoon structures, and the relevant amino acid residues and DNA nucleotides are highlighted as ball-and-stick structures. (**B**) Close-up and rotated view of the cytosine inside the PO-BP from (**A**). The carbons of the deoxyribose ring are numbered and W221 is highlighted in a ball-and-stick structure. *Figure adapted from literature.*[94,217]

bonds between the ASO backbone and the PO-binding pocket is likely to be necessary for accommodation of the ASO.[94,217] (**iii**) Riboses with their 2′-substituents (and analogues thereof) are not tolerated inside the PO-binding pocket and prevent RNase H activity.[217]

## 1.2. RNase H-Compatible Modifications

Despite the susceptibility of RNase H activation to chemical modifications of the ASO, a limited set of tolerated backbone and sugar modifications have been reported (Figure IV-5).[94,218–225] Notably, all backbone modifications listed in

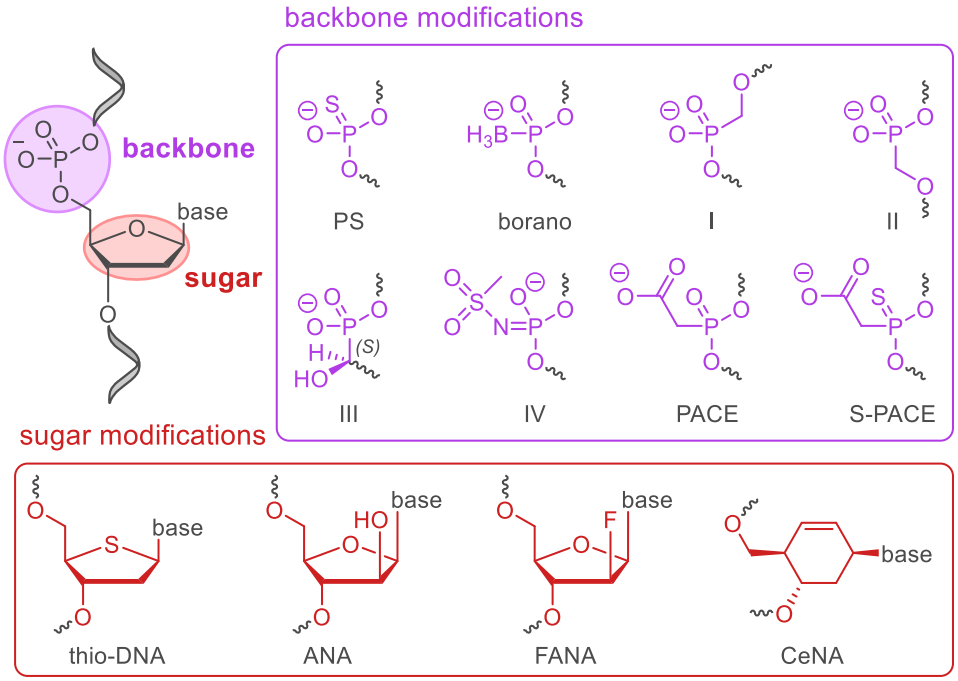



**Figure IV-5: RNase H-compatible backbone and sugar modifications.**
**Backbone modifications**: phosphorothioate (PS),[94] boranophosphate (borano),[218] methylenephosphonate (I and II),[219] 5′-hydroxyphosphonate (III),[219] mesyl phosphoramidate (IV),[220,221] phosphonoacetate (PACE)[222] and thiophosphonoacetate (S-PACE).[222]
**Sugar modifications**: 2′,4′-dideoxy-4′-thio ribose (thio-DNA),[225] arabinofuranose (ANA),[223] 2′-deoxy-2′-fluoro-arabinonucleic acid (FANA)[223] and cyclohexene nucleic acid (CeNA).[224]

Figure IV-5 are negatively charged, which allows ionic interactions with the positively charged PO-binding pocket (see Figure IV-3A). The phosphorothioate,[94] boranophosphate,[218] mesyl phosphoramidate,[220] phosphonoacetate[222] and thiophosphonoacetate[222] modifications are all *P*-chiral, adding another layer of complexity to the design of ASOs when considering the chiral environment of the PO-binding pocket (see Figure IV-3B). Indeed, Iwamoto et al. reported differences in the ability to elicit an RNase H response between different PS–ASO diastereoisomers.[94] The 19 *P*-chiral phosphorothioate linkages in the therapeutic oligonucleotide Mipomersen give rise to more than half a million individual isomers, all potentially having different biophysical and therapeutic properties. Even though there have been considerable advances in the stereo-controlled synthesis[51] of PS–ASOs, this has not yet progressed to the clinic.

Sugar modifications exclude riboses due to the steric clash of the 2′-substituent with a tryptophan inside the PO-binding pocket (see Figure IV-4B). Thus, modifications are mainly limited to deoxyriboses or arabinofuranoses—the 2′-epimer of ribose (Figure IV-5). The only exception that does not fall into one of these categories is the cyclohexene nucleic acid (CeNA) modification, which was reported to be tolerated by RNase H.[224]

## 1.3. The Gapmer Strategy

The limited number of backbone and ribose modifications that are tolerated inside the PO-binding pocket of RNase H prevents the application of a plethora of

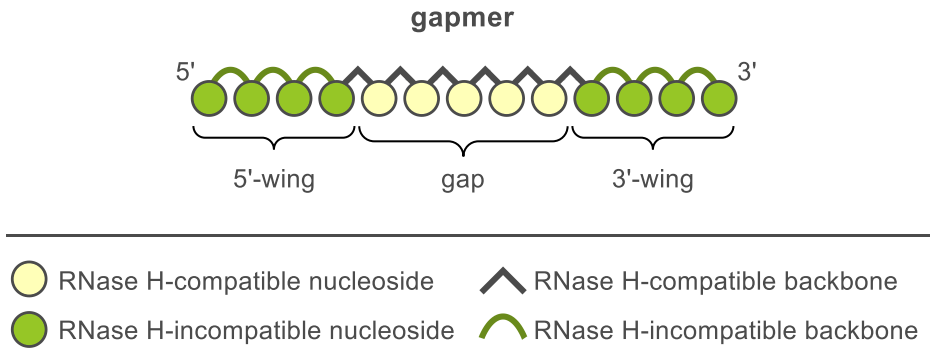



**Figure IV-6: The general design of an RNase H-activating gapmer.**
The gapmer comprises a central gap, which is composed of RNase H-compatible nucleosides and backbone linkages, and two terminal wings, which are composed of RNase H-incompatible nucleosides and backbone linkages.

ASO modifications that are desirable to enhance drug-like properties. A solution is to restrict these modifications to the 3′- and 5′-terminal sections (i.e., wings) of the ASO. Consequently, the ASO retains a central section (i.e., gap) that is fully tolerated by RNase H whilst the modified wings confer desirable drug-like properties such as enzymatic stability or enhanced target affinity (Figure IV-6). Indeed, a central gap of five RNase H-compatible nucleotides is enough to elicit an efficient RNase H response.[226] This molecular design of ASOs was first proposed in a patent[227] in 1987 and was subsequently coined the gapmer design.

Since this initial patent, several sugar modifications have been reported as part of the gapmer wings.[128,228,229] The benefits of introducing sugar modifications were further demonstrated in a study by Shen et al., where the authors reported that a single 2′-OMe modification in the gap of a gapmer can substantially improve its therapeutic index.[127] Although modified backbones are less explored than modified sugars, non-ionic methylphosphonate gapmer wings can increase specificity of

RNase H-mediated target cleavage.[230] Moreover, replacement of a single phosphorothioate by a charge-neutral methylphosphonate in the 5′-wing can eliminate toxicity of a hepatotoxic PS–ASO gapmer.[128] However, methylphosphonates are *P*-chiral and multiple incorporations will exponentially increase the complexity of the resulting ASOs. Nevertheless, the incorporation of multiple charge-neutral phosphodiester surrogates—ideally achiral—might facilitate better cellular uptake by reducing the overall negative charge of the ASO.

## 1.4. Charge-Neutral Backbone Modifications for Enhanced Cellular Uptake

Decreasing the overall negative charge of an ASO provides a rationale for enhanced cellular uptake across the negatively charged cell membrane.[159] Early proposals rationalised that charge-neutral ASOs would be able to diffuse across the cellular membrane.[158,159] This early hypothesis led to the evaluation of several charge-neutral phosphodiester analogues for enhanced cellular uptake (Figure IV-7). Oligonucleotides with a methylphosphonate backbone (MP; Figure IV-7) showed significantly altered uptake characteristics compared to unmodified oligonucleotides.[231] However, the authors reported that MP-modified oligonucleotides do not freely diffuse across the cellular membrane but are internalised *via* an endocytic pathway which leads to their entrapment inside subcellular compartments. In fact, the early view that neutral compounds the size of oligonucleotides could readily diffuse across the cell membrane is outdated, and more recent models explain oligonucleotide internalisation through interaction with

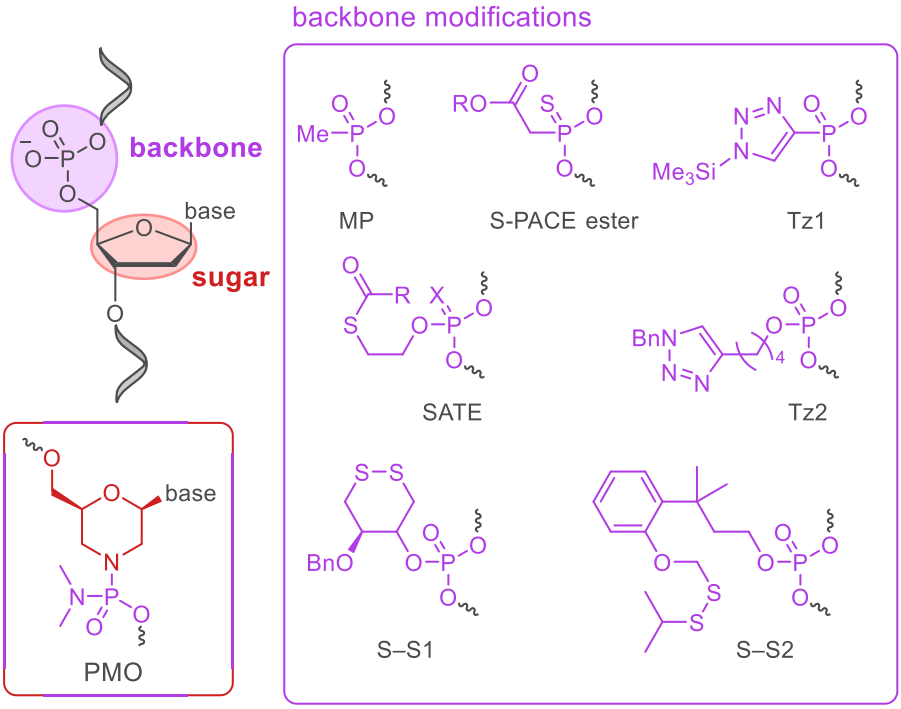



**Figure IV-7: Charge-neutral backbone modifications enhancing cellular uptake.**
The backbone modifications are highlighted in purple and comprise: MP,[231] S-PACE ester (R = Me, *n*-Bu),[222] Tz1,[234] *S*-acyl-2-thioethyl (SATE) (R = *tert*-Bu; X = O, or S),[235,236] Tz2,[237] S–S1,[238] and S–S2.[239] PMOs[136] combine sugar (red) and backbone modifications.

cell surface receptors and endocytic pathways.[124] Moreover, such receptor-mediated uptake mechanisms are not limited to charge-neutral oligonucleotides, and PS–ASOs are a prime example of polyanionic oligonucleotides that show enhanced cellular uptake through interaction with cell surface proteins.[124] *Vice versa*, reduction of the anionic charges does not necessarily enhance internalisation as exemplified by charge-neutral peptide nucleic acids (PNAs), which show very limited membrane permeability and cellular uptake.[232,233] Nevertheless, various other studies demonstrated the potential of charge reduction for enhanced cellular uptake—although most certainly not *via* passive diffusion across the membrane. As such, reduction of the overall charge in thiophosphonoacetate ester chimeras was reported to substantially increase cellular uptake (S-PACE; Figure IV-7).[222] The

authors found that a charge-to-linkage ratio of 0.5 (anionic PO/total number of internucleoside linkages) was sufficient for enhanced uptake. Moreover, the ester functions as a prodrug moiety that can be hydrolysed by subcellular esterases, resulting in the RNase H-compatible thiophosphonoacetate modification (see S-PACE; Figure IV-5). The triazole-modified backbone (Tz1; Figure IV-7) is formed by click chemistry between an alkyne phosphonate and an azide.[234] Gapmers with such charge-neutral triazoles in the wings are taken up by cells and do not result in endosomal entrapment, but subcellular localisation was dependent on the cell line.[234] Moreover, the triazole can be formed from different azides which broadens the possibilities of exploring alternative substituents to facilitate cellular uptake. The phosphotriester modification with a thioester in the sidechain (SATE; Figure IV-7) significantly enhances cellular uptake and can be administered as a prodrug.[235,236] Hydrolysis of the thioester by subcellular thioesterases results in complete deprotection and formation of the natural phosphodiester (or phosphorothioate) bonds. Thus, the stereochemistry of the *P*-chiral phosphotriester might be less important as it will be eliminated after internalisation. Oligonucleotides bearing triazole-modified phosphotriesters (Tz2; Figure IV-7) can aggregate and thereby deliver a cargo (e.g., fluorescent dye) inside cells.[237] However, it seems unlikely that the charge-neutrality of the backbone mediates cellular uptake in this case, which might be more likely due to unspecific internalisation of the formed aggregates. Phosphotriesters with disulfides in their side chain (S–S1 and S–S2; Figure IV-7) represent an interesting class of backbone modifications that increase cellular uptake through formation of disulphide bonds with a cysteines of cell surface proteins

followed by internalisation.[238–240] However, charge-neutrality of the backbone might not be a specific requirement for potential cellular uptake *via* disulphide bond formation, which was recently proposed to play a role in the uptake of negatively charged PS–ASOs, too[241] Both disulphides in Figure IV-7 are prodrug structures that can be reduced by cytosolic glutathione and release an unmodified oligonucleotide. Phosphoramidate morpholino oligonucleotides (PMO; Figure IV-7) combine a sugar modification with a charge-neutral backbone and are taken up by the cell through interactions with scavenger receptors.[136]

It is important to emphasise that all backbone modifications in Figure IV-7 are *P*-chiral, which can significantly complicate the drug discovery process especially when the ASOs are not designed as prodrugs that release a non-chiral internucleoside linkage.

Even though the effects on serum stability, toxicity and cellular uptake of charge-neutral backbone modifications are striking, antisense drug development is mainly focused on the phosphorothioate backbone modification. Therefore, charge-neutral backbone alternatives are underexplored but urgently needed to improve drug-like properties of ASOs.

# 2. Aim of the Chapter

Based on the duplex stability, nuclease resistance, and solid phase oligonucleotide synthesis properties of the two charge-neutral backbones presented in Chapter II and III, i.e., the carbamate and amide linkages, the latter was chosen for further studies in the context of RNase H activation (AM; Figure IV-8). The amide backbone is achiral and charge-neutral, thereby providing promising

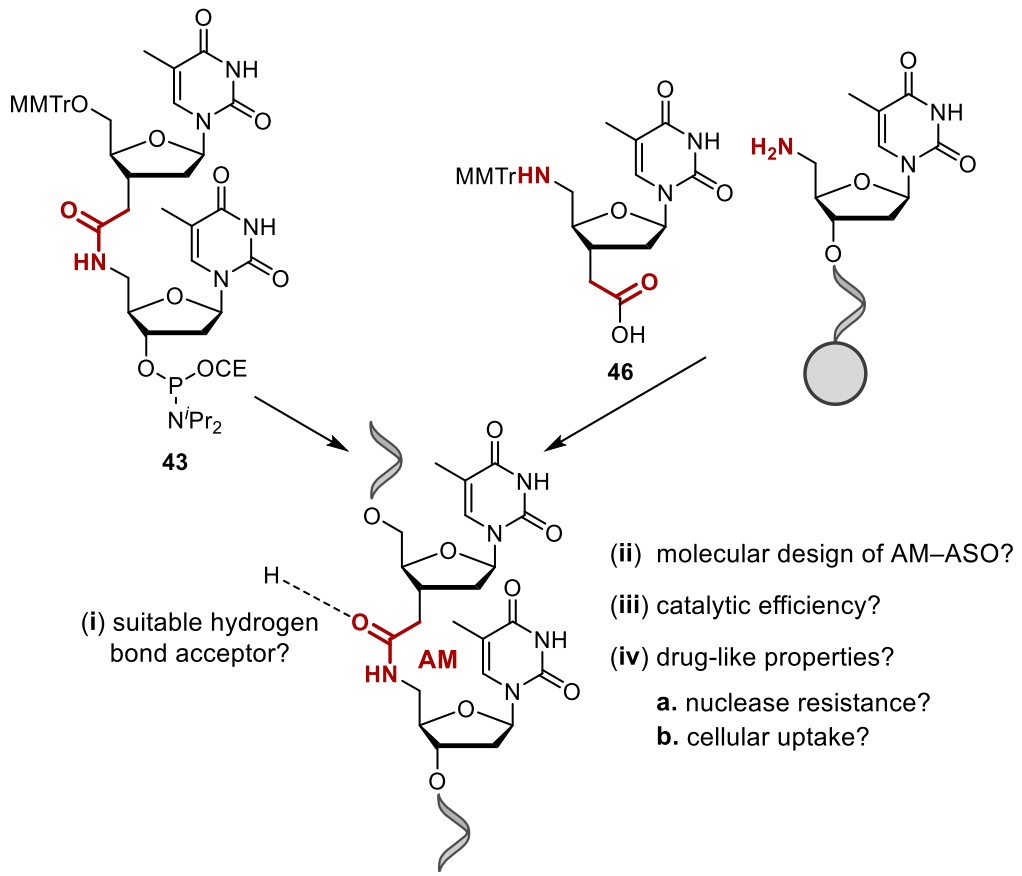



**Figure IV-8: The proposed approach to investigate the compatibility of the amide modification in RNase H-activating ASOs.**
(**i**–**iv**) The questions defining RNase H compatibility. Hydrogen bonds inside the PO-BP of RNase H is indicated with dotted lines. The solid support is represented by a grey cycle. Abbreviations: CE, 2-cyanoethyl; $^i$Pr, isopropyl; MMTr, 4-methoxytrityl.

characteristics to potentially enhance cellular uptake without introducing an additional sterocentre, which deconvolutes ASO drug design. Moreover, amide-modified oligonucleotides have been previously shown to retain good hybridisation properties and the amide is an excellent mimic of the phosphodiester bond with reported applications in RNAi.[190195193] However, its suitability in RNase H-activating ASOs has not yet been elucidated. Thus, the aim of this chapter is the evaluation of amide-modified ASOs for RNase H activation, and their drug-like properties.

The proposed synthetic approach to access amide-modified oligonucleotides is based on both the dimer and the monomer approach (see Section 3.3.1; Chapter I). This combines the ease of amide-modified oligonucleotide synthesis using CE phosphoramidite **43** and the possibility to introduce consecutive amides using amino–carboxylic acid **46**—the synthesis of both has been described in Chapter III. The latter requires integration of a robust peptide coupling protocol with solid phase oligonucleotide synthesis. Together, the monomer and dimer approaches give access to a variety of AM–PO chimera designs for subsequent evaluation of their RNase H compatibility, which will be addressed with the following questions: (**i**) is the amide modification tolerated inside the phosphodiester bond-binding pocket of RNase H? In detail, can the amide modification replace one of the three consecutive phosphodiester bonds that form hydrogen bond interactions with RNase H? (**ii**) What are the possible sites for amide modifications inside an AM–ASO? (**iii**) How efficient is RNase H activation by a given AM–ASO design? (**iv**) Does the amide

modification confer other desirable drug-like properties such as enhanced nuclease stability or cellular uptake?

The answers to those questions will provide invaluable insights into the applicability of the amide modification in ASOs to confer desirable drug-like properties whilst eliciting an efficient RNase H response.

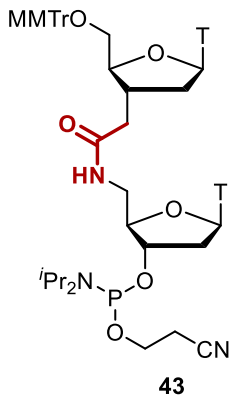
# 3. Results and Discussion

## 3.1. Oligonucleotides with Isolated Amide Linkages

### 3.1.1. Synthesis of Oligonucleotides with Isolated Amides

To investigate the compatibility with RNase H activation, oligonucleotides with isolated amides were synthesised *via* the dimer strategy. Because ribose modifications at the 2′-position generally eradicate RNase H activity, the amide modification was examined as part of deoxynucleosides. For this purpose, previously synthesised dinucleoside **43** was used during solid phase oligonucleotide synthesis to introduce isolated amides at various sites inside the AM–PO chimeras (Table IV-1). ASOs ON33–ON37 are all 12-mers complementary to a fluorescein-labelled polyA RNA (FL-RNA1). ON33 is an unmodified DNA polyT 12-mer, which was used as a positive control to induce RNase H-mediated degradation of FL-RNA1. Three isolated amide modifications were introduced in ON34 where each amide is interspaced by two consecutive phosphodiester bonds. This design is based on the hydrogen bonds formed by the three consecutive phosphodiesters of the DNA strand within the PO-binding pocket of human RNase H (see Figure IV-3B). The possibility of replacing at least one of these three consecutive phosphodiester bonds by an amide was being tested by this design. ON35 is a gapmer control, which narrows the window for RNase H activation. The central part of ON35 is composed of six thymidines, which are flanked by three 2′-OMe uridines

**Table IV-1: List of synthesised ONs for RNase H studies.**
Abbreviations: $^{i}$Pr, isopropyl; MMTr, 4-methoxytrityl; T, thymine.




43

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| ON33 | TTTTTTTTTTTTTT | 3588 | 3589 |
| ON34 | TTTTTTTTTTTTTT | 3472 | 3473 |
| ON35 | uuuTTTTTTTTuuu | 3683 | 3685 |
| ON36 | uuuTTTTTTTTuuu | 3645 | 3646 |
| ON37 | uuuTTTTTTTTuuu | 3606 | 3607 |
| FL-RNA1 | FL aaaaaaaaaaaaaa | 4425 | 4426 |

N DNA  n RNA  n 2′-OMe  PO  AM  FL fluorescein

at each terminus. The $T_6$-window leaves enough space for efficient RNase H activation when hybridised to an RNA target.[226] ON36 and ON37 are 2′-OMe gapmers with isolated amides inside the $T_6$-window. ON36 has only one central amide whilst ON37 has two amides which are interspaced by two consecutive phosphodiester bonds.

### 3.1.2. Oligonucleotides with Isolated Amides for RNase H Activation

ASOs ON33–ON37 were evaluated for their ability to induce an RNase H response upon binding to their target FL-RNA1. *E. coli* RNase H was used for all enzymatic studies as its sequence preferences for target RNAs are nearly identical to human RNase H.[242] The amino acid sequences of the catalytic domains between *E. coli* and human RNase H display strong homology and comparison of different RNase H classes shows a conserved core structure of the catalytic domains (*E. coli*, *Bacillus halodurans* and HIV).[243] The general approach to test an RNase H response

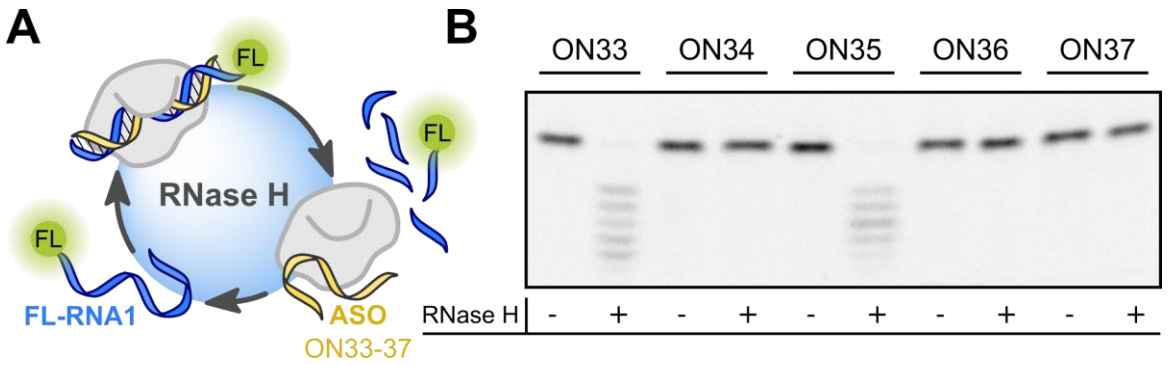



**Figure IV-9: RNase H-mediated degradation of FL-RNA1.**
(**A**) Schematic representation of the RNase H-mediated degradation of FL-RNA1 which is induced by formation of an ASO:RNA heteroduplex. (**B**) RNase H-mediated degradation of FL-RNA1 after incubation with ASOs at rt for 20 h. ASOs are ON33–ON37. ASO/FL-RNA1, 1.1/1 (mole/mole). Degradation fragments were separated by 20% denaturing PAGE and the FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

and the obtained results are shown in Figure IV-9. Each ASO (ON33–ON37) was incubated with FL-RNA1 and RNase H at room temperature for 20 h. Upon formation of an ASO:FL-RNA1 duplex, RNase H should cleave the FL-RNA1 target, a process that could be monitored *via* the fluorescent label (Figure IV-9A). The oligomers were then analysed by denaturing gel electrophoresis and emission of the fluorescein was detected using a fluorescence imager (Figure IV-9B). No cleavage of FL-RNA1 occurred in the absence of RNase H, which excludes unspecific target degradation. Positive controls ON33 and ON35 induced complete degradation of the target after 20 h, confirming RNase H activity. However, all amide-modified ASOs (ON34, ON36 and ON37) did not induce any RNase H response and FL-RNA1 remained un-degraded. Thus, the amide is not tolerated as a replacement for any of the three phosphodiester bonds within the PO-binding pocket of RNase H. The reported RNase H-compatible modifications of Figure IV-5 suggest that steric factors do not seem to be too limiting (see the relatively bulky

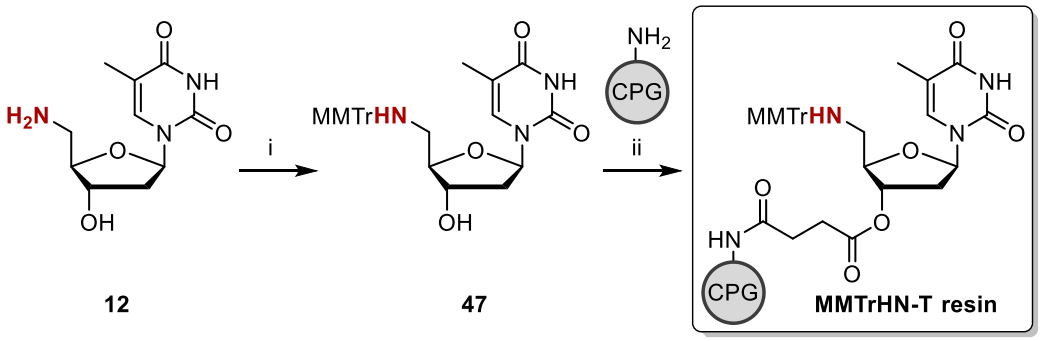
mesyl phosphoramidates; Figure IV-5). However, all RNase H-compatible backbone mimics reported to date are negatively charged. This suggests that the negative charge is crucial for accommodation within the basic PO-binding pocket of RNase H, which also agrees with the surface potential shown in Figure IV-3A.

## 3.2. Oligonucleotides with Consecutive Amide Linkages for Antisense Application

The observed incompatibility of the amide modification inside the PO-binding pocket of RNase H demanded a redesign of the AM–ASOs. To this end, an efficient on-resin peptide coupling protocol was necessary to facilitate the synthesis of an AM–PO gapmer with consecutive amide modifications in the wings. The consecutiveness of the amide modification was desired to obtain a higher overall degree of amide modifications whilst leaving a window wide enough for RNase H activation. The following sections describe the on-resin peptide coupling protocol that was used to synthesise AM–PO gapmers with consecutive amides. Those AM–PO gapmers were then further evaluated for their ability to elicit an efficient RNase H response, their resistance against degradation by nucleases, and their cellular uptake potential.

### 3.2.1. Synthesis of Oligonucleotides with Consecutive Amides

The introduction of an amide as the first internucleoside linkage at the 3′-wing requires resin functionalisation as a starting point for the oligonucleotide synthesis (Scheme IV-1). Amine **12** was first protected with MMTr to obtain **47**[244] under




**Scheme IV-1: Resin functionalisation with monomer 47.**
Reagents and conditions: (i) 4-methoxytriphenyltrityl chloride, pyridine, 4 °C to rt, 24 h, 72% (ii) succinic anhydride, 4-dimethylaminopyridine, pyridine, 20 h, rt, 29%. Abbreviations: CPG, controlled pore glass; MMTr, 4-methoxytrityl.

standard protection conditions. The MMTr protection group was preferred over the Fmoc protection group for reasons explained in Chapter III. Monomer **47** was then coupled to amine-modified controlled pore glass *via* a succinyl ester linker. This serves as a starting point for the synthesis of oligonucleotides containing an amide at their 3′-ends. AM–PO chimeras were synthesised by combining solid phase oligonucleotide synthesis with on-resin amide bond formation, which was optimised[m] from protocols reported by von Matt et al.,[55] Brandsen et al.,[245] Tanui et al.,[189] and Kotikam et al.[246] A general overview of the transition between amide couplings and standard CE phosphoramidite oligonucleotide synthesis is described in Figure IV-10. An MMTr-protected amino–carboxylic acid (**I**) can be coupled to the 5′-amino group of the growing oligonucleotide chain (**II**) to form an amide bond (**III**). The combination of PyBOP as the coupling agent and 4-methyl morpholine as the base, gave the best results. Standard detritylation deprotects the 5′-amino group (**II**), which serves as a substrate for further coupling/deprotection cycles of amino–

[m] Conditions optimised by Dr. Ysobel R. Baker (University of Oxford)

carboxylic acids (**I**) to form oligonucleotide sections with consecutive amides. Alternatively, a carboxylic acid monomer (**IV**) with a 5′-oxygen functionality can be coupled *via* amide bond formation. Successive detritylation of (**V**) produces a 5′-OH group (**VI**) which serves as a substrate for standard CE phosphoramidite oligonucleotide synthesis. Importantly, all steps—coupling of phosphoramidites (**VII**), capping and oxidation of phosphite triester (**VIII**), and detritylation of phosphate triester (**X**)—are compatible in the presence of the amide backbones. Coupling of CE phosphoramidites with a 5′-nitrogen functionality (**XI**) forms a phosphite triester (**XII**), which can be oxidised and deprotected to install a 5′-$NH_2$ (**II**). These transitions between amide- and phosphodiester-forming oligonucleotide synthesis allow the flexible synthesis of AM–PO chimeras. Even though not explored as part of this thesis, the presented synthesis of AM–PO chimeras might be applicable for the synthesis of AM–PS chimeras by changing the standard oxidation to a sulphurisation procedure (step E; Figure IV-10). Such chimeras are particularly relevant in light of the growing evidence suggesting that combinations of phosphorothioate linkages and alternative backbones have the potential to produce mixed-backbone ASOs with reduced toxicities.[93,117,128,129]

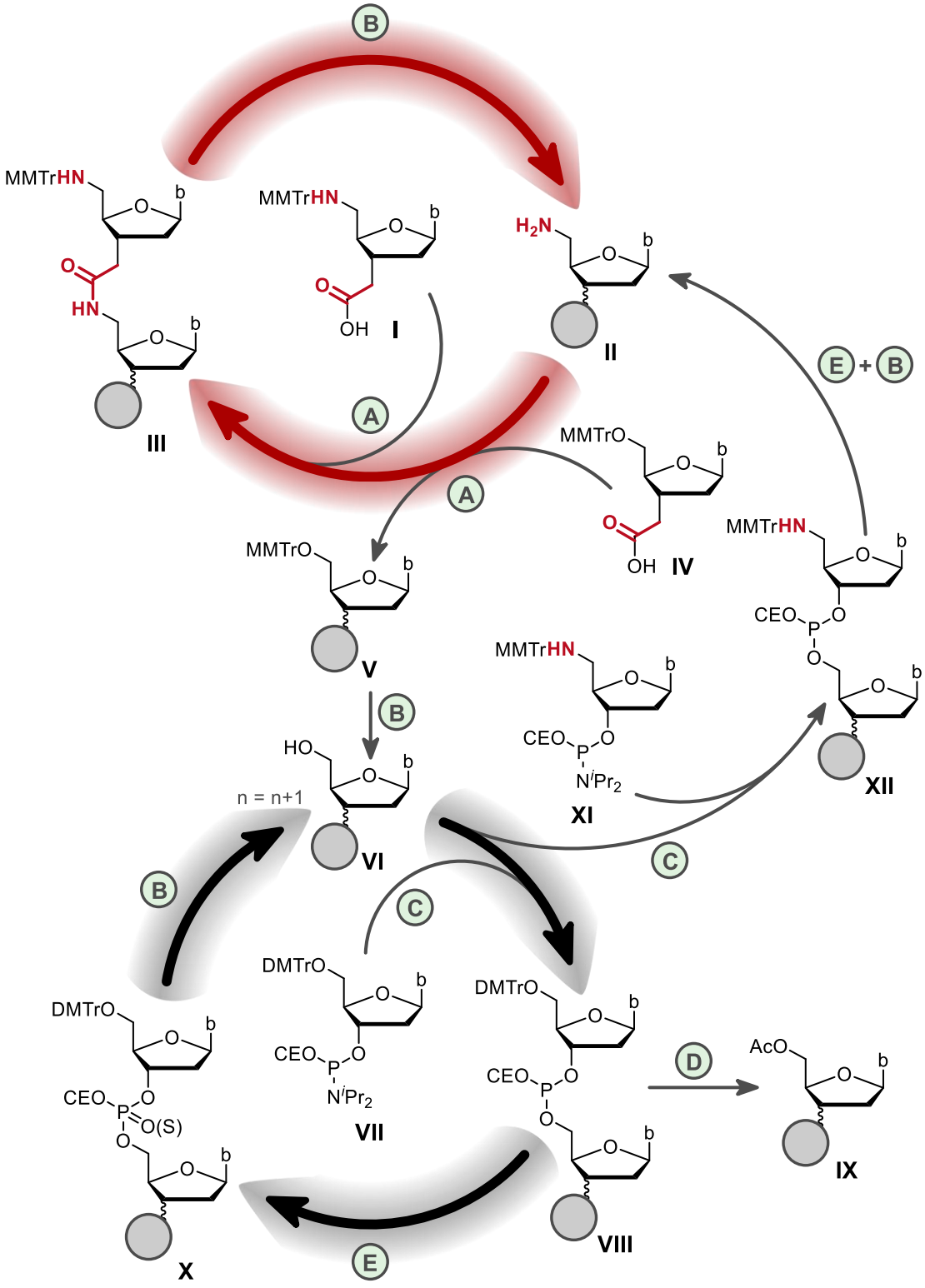



**Figure IV-10: Schematic overview of transitions between amide couplings (red cycle) and standard CE phosphoramidite oligonucleotide synthesis (grey cycle).**
(**A**) Amide coupling. (**B**) Detritylation. (**C**) Coupling of CE phosphoramidites. (**D**) Capping. (**E**) Oxidation or sulphurisation. The solid support is represented as a grey circle. Abbreviations: b, nucleobase; CE, 2-cyanoethyl; DMTr, 4,4′-dimethoxytrityl; MMTr, 4-methoxytrityl. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

Using the strategy outlined in Figure IV-10, AM–PO gapmer chimeras with consecutive amide backbones were synthesised (Table IV-2). AM–PO gapmer chimera ON38 contains eight consecutive amides (four at each wing) and a central unmodified window with nine consecutive phosphodiester bonds for RNase H activation. With the retained nine negatively charged phosphodiester bonds and a total number of 17 internucleoside linkages, ON38 has a charge-to-linkage ratio of 0.53. Importantly, no water solubility problems were observed for ON38. This aligns with a previous report showing that a single phosphate group at either terminus of an all-amide 15-mer oligonucleotide is sufficient for adequate water solubility.[55] As mentioned in Section 3.2 of Chapter I, FAM **4** can be coupled using CE phosphoramidite chemistry to install a fluorescent label at the 5′-terminus (FL-ON38, Table IV-2). FL-ON38 is the fluorescently labelled analogue of ON38 and has a charge-to-linkage ratio of 0.56. No water solubility issues were found for FL-ON38, regardless of the flat and rather hydrophobic 5′-fluorescein label.

The purities of ON38 and FL-ON38 were evaluated by UPLC analysis of the oligonucleotides after global deprotection and cleavage from the resin. The traces of the crude and purified oligonucleotides and their corresponding mass spectra are

**Table IV-2: List of synthesised ONs with consecutive amide internucleoside linkages.**

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| ON38 | TTTTTCCTGATAGTTTTT | 5140 | 5141 |
| FL-ON38 | FL-TTTTTCCTGATAGTTTTT | 5677 | 5678 |

N DNA  PO  AM  FL fluorescein

shown in Figure IV-11. The UPLC traces of crude ON38 (Figure IV-11A) and crude FL-ON38 (Figure IV-11D) show relatively clean chromatograms and each major peak corresponds to the desired oligonucleotide. Purification of ON38 and FL-ON38 by HPLC resulted in pure oligonucleotide samples (Figure IV-11B, E) and the identities of the oligonucleotides were confirmed by mass spectrometry (Figure IV-11C, F). The high purity of the crudes suggests that the conditions for the on-resin amide bond formation resulted in excellent coupling efficiencies. Moreover, the overall isolated yield for ON38 was 16%. Assuming a hypothetical average coupling efficiency of 98.5%[247] to establish the nine phosphodiester bonds in ON38, the estimated efficiency of each of the eight amide couplings is 81%. The

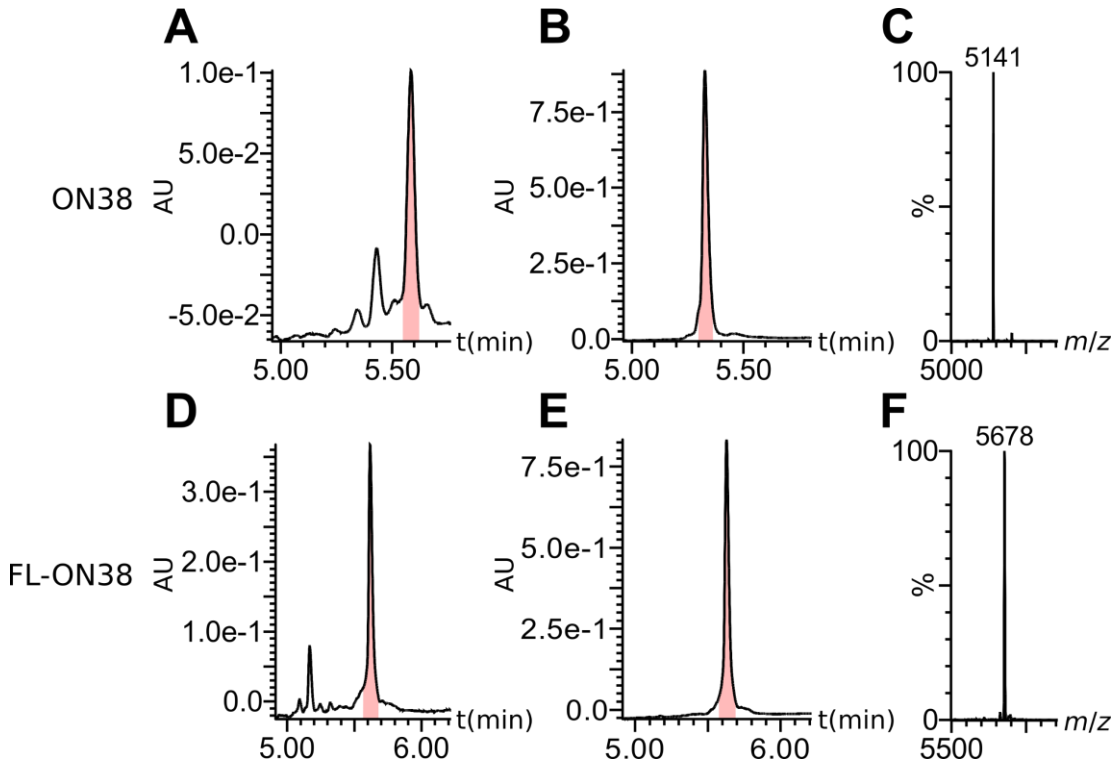



**Figure IV-11: Purity of synthesised ONs with consecutive amide internucleoside linkages.** (**A**) UPLC trace of crude ON38. (**B**) UPLC trace of purified ON38. (**C**) MS analysis of purified ON38. (**D**) UPLC trace of crude FL-ON38. (**E**) UPLC trace of purified FL-ON38. (**F**) MS analysis of purified FL-ON38. UPLC peaks corresponding to the desired ONs are highlighted. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

estimation results from Equation 2 where *Y*, *ce* and *n* describe overall yield, coupling efficiency and number of couplings, respectively.

$$Y = ce^n \quad (2)$$

However, this estimation can still be considered as an underestimation because the yield was determined as the isolated yield after HPLC purification of ON38. This is striking as previously reported couplings of 2′-*O*-TBS-protected amino–carboxylic acids for amide bond formation were only performed for up to six consecutive amides with low overall yields (0.14% overall yield for n = 6).[190] Even though the authors reported individual coupling efficiencies of 80–85%, the bulky 2′-*O*-TBS group seemed to hamper the synthesis of oligonucleotides with more than six consecutive amides. Hence, the absence of a 2′-functionality in 2′-deoxy amino–carboxylic acid **46** facilitates high coupling efficiencies and the synthesis of oligonucleotides with consecutive amides greater than six. In fact, eight consecutive amides were incorporated in ON38 and FL-ON38.

### 3.2.2. Target Engagement of Oligonucleotides with Consecutive Amides

Target engagement of ASOs with RNA is essential for high antisense activity.[157,229] This includes the sequence-specific formation of stable ASO:RNA duplexes through Watson–Crick base pairing.[17] Moreover, the duplex structure (A-type or B-type) is important for antisense applications involving RNase H[215] or RNAi.[248,249] In this context, duplex stabilities and helical structures were evaluated by UV melting and circular dichroism experiments, respectively.

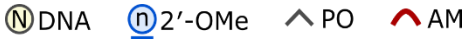
### 3.2.2.1. Duplex Stability

Previous reports of isolated amides showed that the amide modification can slightly increase duplex stabilities with target RNAs (see ON26; Table III-1).[62] The authors also reported a destabilising effect against a DNA target. Discrimination between RNA and DNA is desirable as the target for antisense applications is an RNA strand. A later study evaluated the effect of consecutive amides on duplex stabilities with RNA and DNA targets (see ON29–ON32; Table III-1).[55] However, none of those tested amide-modified oligonucleotides included a gapmer design. Thus, affinities of AM–PO gapmer ON38 to complementary RNA and DNA targets were evaluated by melting temperature measurements (Table IV-3). Because 2′-modifications can improve affinities to RNA targets, a 2′-OMe gapmer (ON39, Table IV-3) was tested as a typical example of a commonly adapted gapmer design for therapeutic approaches.[250] ON38 and ON39 were both compared to unmodified ON40. All measured melting temperatures and their differences to

**Table IV-3: Melting temperatures of ASOs ON38–40 against their complementary RNA and DNA targets.**

ON38–40 were incubated with their respective RNA or DNA targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0) and absorbance was measured at λ = 260 nm in relation to *T*. $T_m$'s were obtained from maxima of the 1st derivatives ($dA_{260}/dT$) (see Figure A-3; Appendix) and $\Delta T_m$'s are reported as the difference to $T_m$(ON40). *Table and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

| | | RNA target | DNA target |
|---|---|---|---|
| ON | Sequence (5′→3′) | $T_m$ ($\Delta T_m$)/°C | $T_m$ ($\Delta T_m$)/°C |
| ON38 |  | 54.6 (−1.5) | 52.3 (−3.2) |
| ON39 |  | 58.6 (+2.5) | 47.8 (−7.7) |
| ON40 |  | 56.1 | 55.5 |



ON40 ($\Delta T_m$'s) are summarised in Table IV-3. The consecutive amides in AM–PO gapmer ON38 decreased duplex stability with RNA by −1.5 °C (−0.19 °C per mod.). Conversely, the 2′-OMe modification in gapmer ON39 increased target affinity to RNA by +2.5 °C (+0.25 °C per mod.). Both observations are consistent with the literature.[55,251] Hybridisation of ON39 and ON38 against a DNA target resulted in decreased melting temperatures: −7.7 °C (−0.77 °C per mod.) and −3.2 °C (−0.40 °C per mod.), respectively. Thus, both oligonucleotides show selectivity for RNA over DNA, although this effect is more prevalent for 2′-OMe gapmer ON39.

#### 3.2.2.2. Duplex Structure

In addition to the duplex stabilities, the helical duplex structures formed by ASOs and their respective targets are relevant for target engagement. Whilst DNA:DNA duplexes adapt a B-form structure, RNA:RNA and DNA:RNA duplexes generally form A-form helices.[4] The effects of the modified wings of gapmers ON38 and ON39 on the global duplex structure were evaluated by circular dichroism experiments (Figure IV-12). The ON40:RNA duplex displayed all the characteristic peaks for a global A-type RNA structure: a maximum near 260 nm; a minimum near 210 nm; a small negative θ-values between 290 and 300 nm.[19] The duplex of AM–PO gapmer ON38 with RNA resulted in an almost identical CD spectrum to the ON40:RNA duplex. Hence, the global structural integrity of ON38 to form an A-type helix with RNA is retained. The duplex formed by 2′-OMe gapmer ON39 and RNA retained the characteristic minimum at 210 nm and the small negative θ-values between 290 and 300 nm but the maximum showed a bathochromic shift from

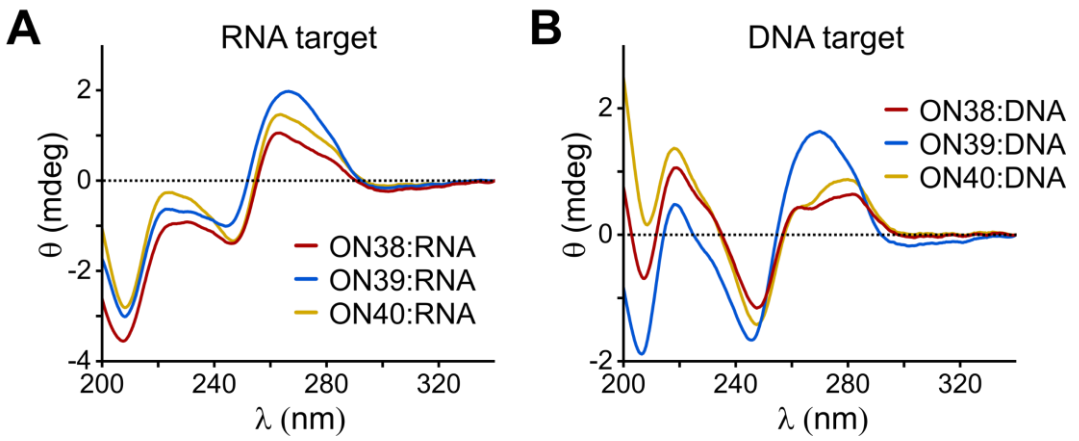



**Figure IV-12: Circular dichroism spectra of ON38–40 in duplex with their RNA or DNA targets.** CD spectra of duplexes formed by ON38–40 with their respective targets (2 μM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). Data points were taken as the average of four scans. (**A**) RNA target: 5′-r(AAA AAC UAU CAG GAA AAA)-3′; (**B**) DNA target: 5′-d(AAA AAC TAT CAG GAA AAA)-3′. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

around 260 to around 270 nm. The maximum around 270 nm is more characteristic for an A-type DNA structure but the global distortion compared to ON40:RNA was minor.

The circular dichroism spectra of ON38–40 in duplex with a DNA target are given in Figure IV-12B. The reference duplex ON40:DNA showed the characteristics for a global B-type DNA: approximately equal negative and positive peaks at around 245 nm and 275 nm, respectively, and centred around 260 nm.[19] Again, ON38:DNA was almost identical to ON40:DNA. This demonstrates a striking ability of ON38 to adapt either an A-type or a B-type global structure depending on the respective target. ON39:DNA was mostly like ON40:DNA but had a distinct hypsochromic shift of the maximum from around 275 nm to around 270 nm. This peak has strong similarities to the ON40:RNA maximum around 270 nm in Figure IV-12A and indicates a substantial amount of A-type character for

the ON39:DNA duplex. This is not surprising as it is well-known that the 2′-OMe modification favours a C3′-*endo* sugar pucker and therefore an A-type helical structure.[7]

### 3.2.3. Oligonucleotides with Consecutive Amides for RNase H Activation

Previous results showed that isolated amides are not tolerated inside the PO-binding pocket of RNase H to induce RNA degradation (see Section 3.1.2; Chapter IV). Thus, the amide modification was only introduced as charge-neutral wings of AM–PO gapmer ON38, which was incubated with a complementary, fluorescently labelled RNA target (FL-RNA2) in the presence of *E. coli* RNase H. 2′-OMe gapmer ON39 and unmodified DNA ON40 were used as positive controls (Figure IV-13A). Aliquots of the reaction mixtures containing ASO/FL-RNA2, 1/4 (mole/mole) and *E. coli* RNase H were quenched after set time points and analysed by gel electrophoresis (Figure IV-13B–D). All ASOs (ON38–40) induced RNase H-mediated degradation of FL-RNA2. Individual degradation rates will be discussed later in this chapter. Interestingly, all tested ASOs resulted in the formation of a major degradation product (triangles; Figure IV-13B–D). The cutting pattern induced by AM–PO gapmer ON38 mostly resulted in a single degradation product (triangle; Figure IV-13B) with only few additional faint bands. The identity of the major degradation will be discussed later in this chapter. The 2′-OMe gapmer ON39 induced hydrolysis of FL-RNA2 to form a major degradation product along with another distinct degradation product (filled and unfilled triangles; Figure IV-13C).

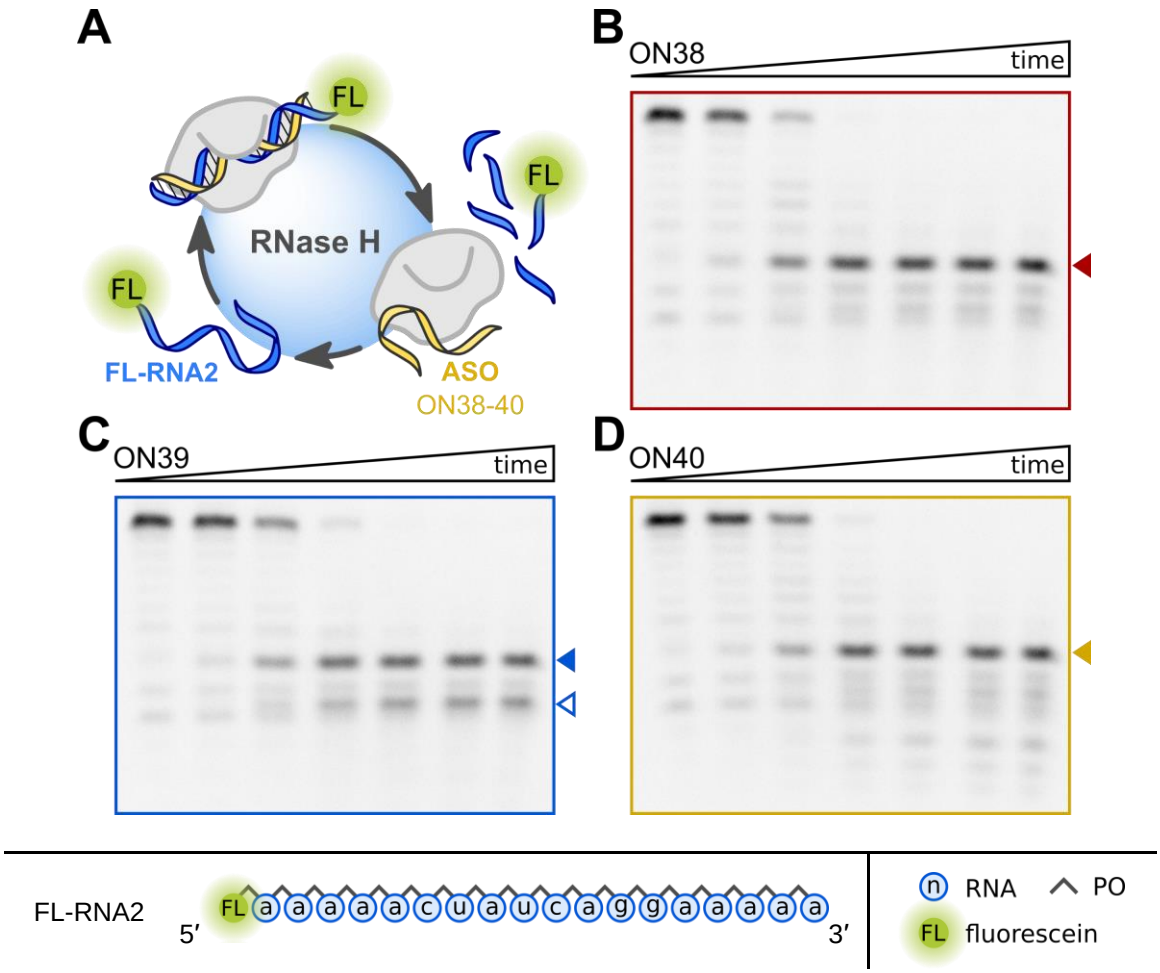



**Figure IV-13: RNase H-mediated degradation of FL-RNA2.**
(**A**) Schematic representation of the RNase H-mediated degradation of FL-RNA2 which is induced by formation of an ASO:FL-RNA2 heteroduplex. (**B**–**D**) RNase H-mediated degradation of FL-RNA2 after incubation with ASOs ON38–40 over 2 h at 37 °C. ASO/FL-RNA2, 1/4 (mole/mole). Aliquots of the reactions were quenched at time points: 0 min, 1 min, 5 min, 15 min, 30 min, 1 h and 2 h. Degradation fragments of the aliquots were separated by 20% denaturing PAGE and the FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Filled triangles represent bands resulting from major cutting site which were extracted from the gel for further analysis. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

Unmodified DNA ON40 produced a major degradation product (triangle; Figure IV-13D), but also led to other faint degradation bands. Indeed, ON40 induced more degradation fragments compared to AM–PO gapmer ON38 and 2′-OMe gapmer ON39. Both the 2′-OMe[226] and the amide modification prevent an RNase H response in their proximity (see Section 3.1.2; Chapter IV). Hence, it is not surprising that

both ON38 and ON39 produced a more defined cutting pattern. Furthermore, comparison of the cutting patterns induced by ON38 and ON39 (Figure IV-13B, C) suggests that the specificity and cutting precision of the AM–PO gapmer design could be superior to the 2′-OMe gapmer design. The possibility of target specificity and mismatch discrimination will be evaluated later in this chapter.

#### 3.2.3.1. Determination of the Cutting Site

As discussed in the previous section, ASO-induced degradation of FL-RNA2 by RNase H led to the formation of a major degradation product. In order to identify this product, the bands indicated with filled triangles in Figure IV-13B–D were extracted from the gel and analysed by mass spectrometry. The found masses (4002 Da) correspond to an 11-mer fragment of FL-RNA2, containing the fluorescent label (calcd mass: 4001 Da). The schematic representations of duplexes formed by ON38–40 with FL-RNA2 are shown in Figure IV-14. The triangles in Figure IV-14 indicate the cutting site of FL-RNA2 by RNase H when in duplex with the respective ASOs. Previous empirical evidence suggests that ASOs with a 5′-d(TCCC)-3′ motif usually perform better than other ASOs in gene knockdown experiments—the corresponding RNA target complement would be 5′-r(GGGA)-3′.[252] A later study determined that the preferred cutting site of human and *E. coli* RNase H is indeed downstream of a G-rich sequence in the RNA strand.[242] The observed major degradation product resulting from a cut in the sequence context 5′-

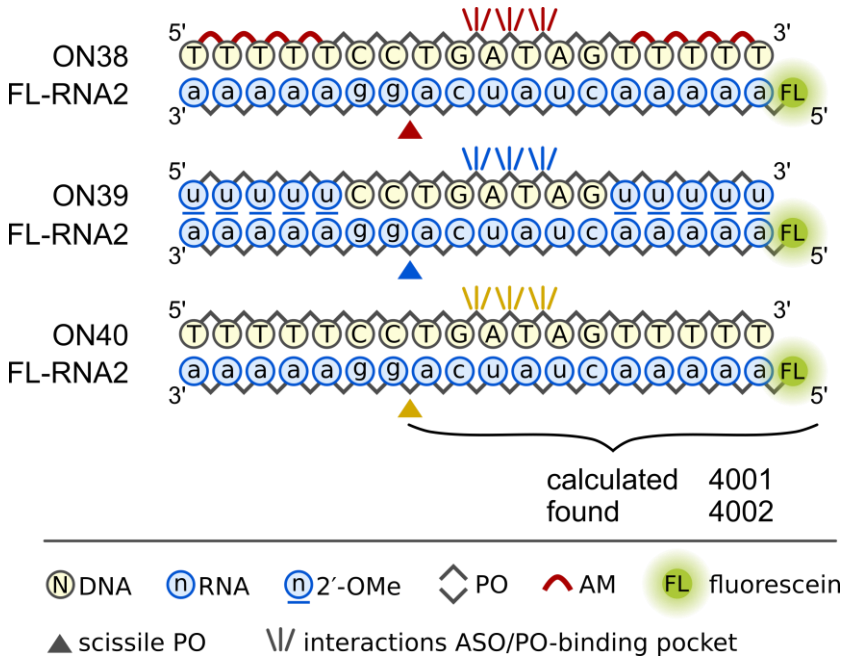



**Figure IV-14: ASOs ON38–40 in duplex with FL-RNA2.**
The sequences of ON38–40 are illustrated in duplex with FL-RNA2. Triangles indicate the cutting sites based on the major fragments found for gel extracts (see triangles in Figure IV-13). Lines indicate the major interactions of the ASO with the PO-BP of human RNase H based on the major cutting sites (scissile phosphate). *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

r(CA↓GGA)-3′ (triangles; Figure IV-14) agrees with these previously determined sequence preferences of RNase H.

The determined major cutting site of FL-RNA2 can be used to discuss the most prevalent interactions of the respective ASOs with the PO-binding pocket of RNase H (Figure IV-14). Previous studies showed that the interactions of the PO-binding pocket of human RNase H with the three consecutive phosphodiesters of the ASO occur 2–4 base pairs from the scissile phosphodiester bond (lines; Figure IV-14).[217,243] Note that the described interactions are taken from a crystal structure of the catalytic domain of human RNase H, and although human and *E. coli* RNase H share a remarkably similar structure, deviations mainly occur in the

basic protrusion, which harbours the interactions with the phosphate backbone of the ASO.[217] Because the cleavage studies in this thesis are performed with *E. coli* RNase H, the interpretation of the observed cutting sites in relation to the potential interactions inside the PO-binding pocket of RNase H cannot be directly deducted from the crystal structure with human RNase H. Moreover, duplex binding by human RNase H is not only mediated by the catalytic domain, but also driven by a RNA:DNA hybrid-binding domain, which is absent in *E. coli* RNase H.[253] Nevertheless, assuming reasonably similar interactions inside the PO-binding pocket of human and *E.coli* RNase H, and based on the observed cutting sites form Figure IV-14, both ASO gapmers (ON38 and ON39) would be bound within their unmodified windows. The possibility of RNase H to 'slide' along the duplex towards the 3′-end of the ASO whilst forming interactions with the phosphodiester bonds of the unmodified window is very limited for ON38 and ON39. On the contrary, the ON40:FL-RNA2 duplex provides more phosphodiester bonds of unmodified nucleotides towards the 3′-end of the ASO. The possible interaction of RNase H with the ON40:FL-RNA2 duplex closer towards the 3′-end of the ASO would lead to a shorter fluorescent fragment of FL-RNA2. This would explain the fluorescently labelled RNA fragments with low molecular weights observed for ON40 (Figure IV-13D) which are absent for ON38 or ON39 (Figure IV-13B, C).

#### 3.2.3.2. Cutting Rate

The efficiencies of *in vitro* activation of *E. coli* RNase H by ON38–40 were compared after gel electrophoresis. To this end, the band intensities for full-length

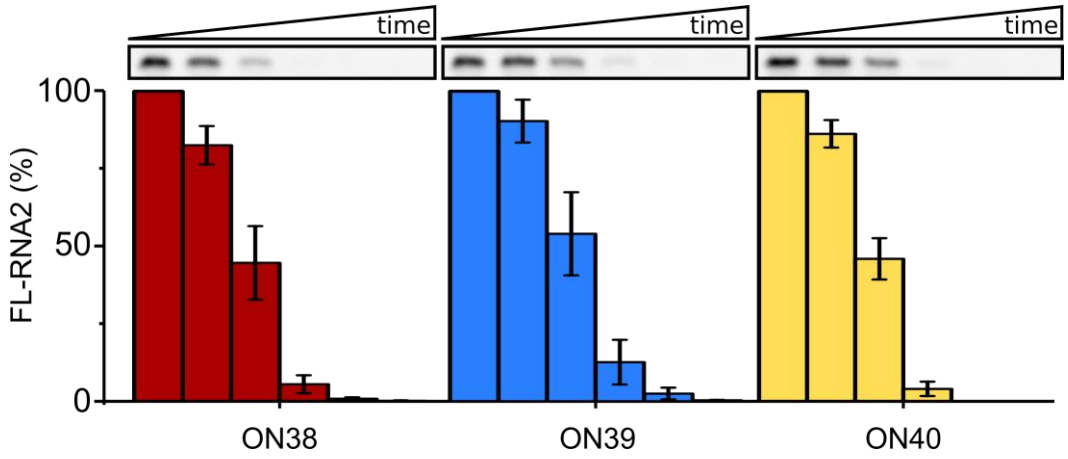



**Figure IV-15: ASO-induced degradation of FL-RNA2 by RNase H at 37 °C.**
Quantification of full-length FL-RNA2 from gels shown in Figure IV-13 using ImageJ. Time points are 0 min, 1 min, 5 min, 15 min, 30 min and 1 h. Quantification of Figure IV-15 is normalised to t = 0 min for each set respectively. Data is represented as the average of triplicates ± standard error of the mean. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

FL-RNA2 (i.e., fluorescent bands with highest molecular weight; Figure IV-13B–D) were integrated using ImageJ from triplicates for each set of ASO, and plotted as bar charts (Figure IV-15). The top of Figure IV-15 shows representative gel sections of full-length FL-RNA2 with visualisation of the fluorescein label from aliquots taken over time. The measured intensities for each time point (bars) indicate the amount of remaining full-length FL-RNA2. Each bar describes the average from triplicates with the respective standard errors of the mean. The first bar of each set of ASO (t = 0 min) was set to 100% and the remaining bars for each set are given as a percentage thereof. All tested ASOs (ON38–40) induced complete degradation of FL-RNA2 by RNase H after 30 min. As described above, the ratio of ASOs (ON38–40) to the RNA target was ASO/FL-RNA2, 1/4 (mole/mole). This means that the ASO-induced degradation of FL-RNA2 by RNase H is catalytic. The catalytic activity of AM–PO gapmer ON38 to induce target degradation is remarkable.

Indeed, previously reported gapmer chimeras with charge-neutral PNA wings only induced non-catalytic degradation of an RNA target.[254] A later study by Debacker et al. identified an improved PNA-chimera design which led to catalytic activities.[255] This PNA-chimera relied on an optimal linker length between the PNA and phosphodiester sections for high catalytic activity. Moreover, the charge-neutral PNA section was restricted to the 3′-wing of the gapmer ASO whilst the 5′-wing was composed of LNA modifications. In comparison, the simple AM–PO gapmer design of ON38 with charge-neutral wings at both termini elicits a catalytic RNase H response without any need for linker optimisation. It is worth mentioning that the RNase H assays reported by Debacker et al.[255] were performed at room temperature whilst the experiments described in this section were incubated at 37 °C. Further RNase H cutting rates with ASOs ON38–40 at 20 °C are described later in this chapter.

When comparing the three sets in Figure IV-15, no notable differences of catalytic rates were observed for any of the tested ASOs at any time point. Hence, AM–PO gapmer ON38 induced RNase H-mediated degradation of FL-RNA2 as efficiently as positive controls ON39 and ON40. No obvious correlation can be drawn between the determined RNA binding affinities (see Table IV-3) and the induced RNA cleavage rates by RNase H. Although ASOs ON38, ON39 and ON40 all had different affinities to their RNA target, no differences were found in activating RNase H to cleave FL-RNA2. Despite the general view that increased target affinity correlates with higher antisense activity,[156] there is increasing evidence that the stabilities of ASO:RNA duplexes do not directly correlate with the

effectiveness of an RNase H response. In fact, several studies observed that shorter oligonucleotides with lower target affinities are more potent than their longer analogues targeting the same sites with increased affinities.[228,256–258] A possible explanation for such observations was previously provided on the basis of a kinetic model describing the events of the ASO-induced RNA degradation by RNase H involving the following steps:[259] (**i**) ASO:RNA duplex formation, (**ii**) recruitment of RNase H, (**iii**) RNA cleavage by RNase H, and (**iv**) dissociation of RNase H and the cleaved RNA fragments from the ASO. The dissociation of the cleaved RNA target affects the rate at which the ASO can be available for the next catalytic cycle (see Figure IV-13A). Whilst dissociation of the ASO from the intact RNA target is slower than from the cleaved products, increased target affinity—e.g., through increased ASO lengths—will reduce the dissociation rate in both cases. In other words, an ASO with higher affinity to the intact RNA target will also have higher affinity to the cleaved RNA products. Moreover, high affinity of the ASO to the cleaved RNA products will slow down the release of RNase H from the cleaved complex and dissociation of the ASO from the RNA fragments. Thus, there is an optimal affinity of the ASO to the RNA target to ensure high catalytic efficiency. As such, sufficiently high affinity is needed to form an ASO:RNA heteroduplex and recruit RNase H, but increased affinity beyond the optimum is counterproductive to the catalytic efficiency. In summary, this model explains why longer ASOs (higher target affinity) can be less potent than their shorter analogues (lower target affinity) even when targeting the same sites of an RNA.

Apart from increasing the length of an ASO, several duplex-stabilising chemical modifications have previously failed to increase RNase H cleavage rates. For instance, the ASO:RNA duplex-stabilising FANA modification ($\Delta T_m$ = +5 °C, or +0.27 °C per mod.) is able to induce RNase H-mediated degradation of the RNA but with no enhanced efficiency over unmodified DNA.[223] However, the FANA modification is less flexible and less likely to adopt a B-form typical C2′-*endo* sugar pucker.[260] As discussed in Section 1.1.1 of this chapter, flexibility and an A→B transition are important for RNase activity, explaining why FANA-modified ASOs fail to induce increased RNase H activity despite their higher RNA binding affinity. Moreover, the duplex-stabilising 5′-hydroxyphosphonate internucleoside linkage is compatible with RNase H activation but does also not increase cleavage rates.[219] On the contrary, various duplex-destabilising modifications were shown to elicit an RNase H response as efficiently as—or even more efficiently than—unmodified DNA.[218,222] The retained antisense activity of AM–PO gapmer ON38 agrees with these reports and a high target affinity does not seem to be the determining factor for efficient RNase H activation. Apart from an optimal binding affinity for catalytic efficiency, toxic effects have been previously described for ASOs with very high target affinities. As such, ASOs that bind their RNA target too tightly are also more likely to bind other RNAs that may not be perfectly matched but still share enough sequence similarities.[261–263] Such off-target binding, and successive degradation of the unintended targets, can lead to toxic effects. Therefore, higher target affinity does not necessarily mean better RNA targeting in a therapeutic context, which must be considered when determining an ASO's optimal target affinity.

As discussed in the circular dichroism experiments above, the AM–PO gapmer chimera ON38 can either adopt an A- or a B-form, depending on its target. This suggests that the transition barrier from the A-form duplex to an 'RNase H-cleavable B-form' is low. Hence, the ability of ON38 to elicit an efficient RNase H response agrees with the flexibility of adopting A- and B-form helical structures. In fact, this structural flexibility of ON38 might be more important than a higher RNA target-binding affinity to elicit an efficient RNase H response.

#### 3.2.3.3. Cutting Rate: Further Results

The ability of ASOs ON38–40 to trigger an RNase H response was evaluated at 20 °C for better comparison to previously reported RNase H assays with PNA gapmers.[255] The ASOs were incubated with FL-RNA2 in the presence of RNase H at 20 °C and aliquots were quenched at different time points and analysed by gel electrophoresis as described in the previous section. All ASOs induced the same cutting pattern of FL-RNA2 upon RNase H degradation as observed in Figure IV-13B–D (see Figure A-4; Appendix). Quantification of remaining full-length FL-RNA2 over time is plotted in Figure IV-16A for each set of ASO. AM–PO gapmer ON38 induced catalytic degradation of FL-RNA2, but at a slower rate than in the previous assay at 37 °C. Whilst FL-RNA2 was completely degraded after 30 min at 37 °C (see ON38; Figure IV-15), ca. 17% was still present after 30 min at 20 °C (ON38; Figure IV-16A). Almost identical results were obtained for degradation of

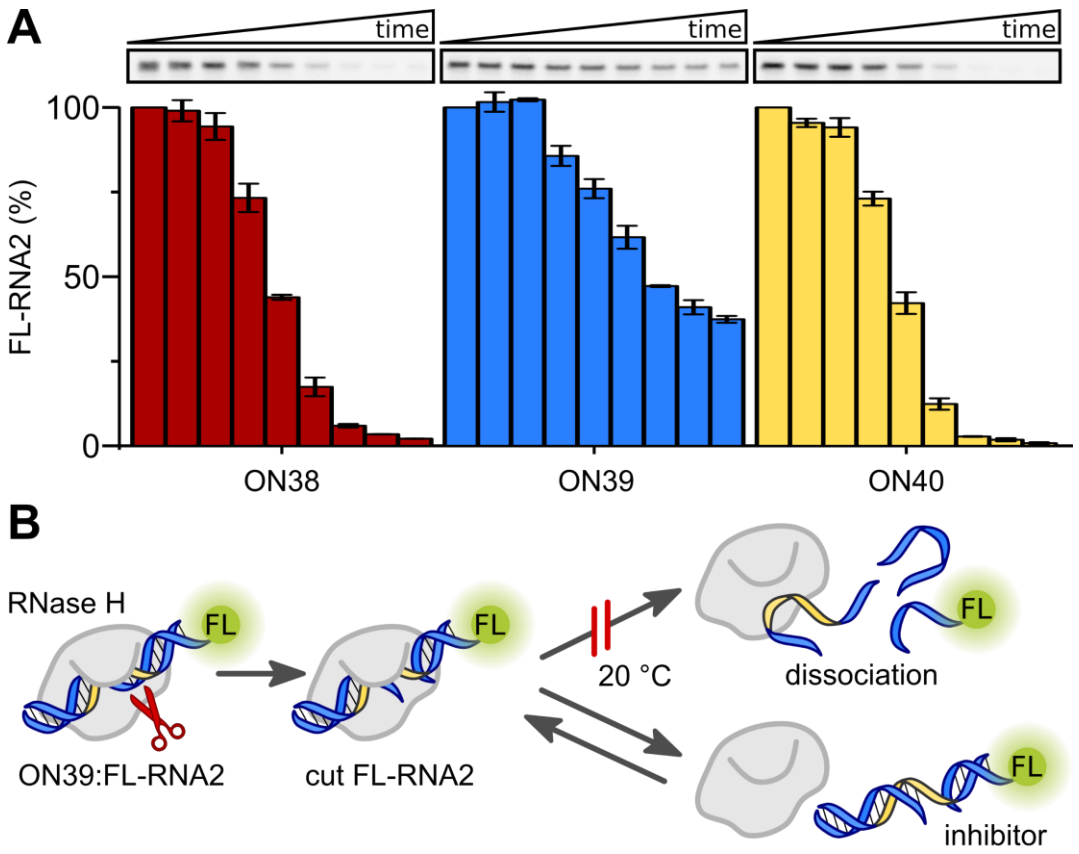



**Figure IV-16: ASO-induced degradation of FL-RNA2 by RNase H at 20 °C.**
(**A**) RNase H-mediated degradation of FL-RNA2 after incubation with ASOs ON38–40 over 3 h at 20 °C. ASO/FL-RNA2, 1/4 (mole/mole). Aliquots of the reactions were quenched at time points: 0 s, 30 s, 1 min, 5 min, 15 min, 30 min, 1 h, 2 h and 3 h. Degradation fragments of the aliquots were separated by 20% denaturing PAGE and the FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Representative bands for full-length FL-RNA2 are shown at the top of the graph. Bars represent quantification of full-length FL-RNA2 as the average of duplicates ± standard error of the mean; those were calculated from gels using ImageJ. Quantification of FL-RNA2 is normalised to t = 0 min for each set respectively. (**B**) Proposed rational for the observed inhibition of degradation for ASO = ON39. Upon hybridisation, RNase H recognises the ON39:FL-RNA2 heteroduplex and induces cutting of FL-RNA2. However, the duplex-stabilising 2′-OMe modifications prevent dissociation of the fragments at 20 °C which then serve as a competitive inhibitor. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

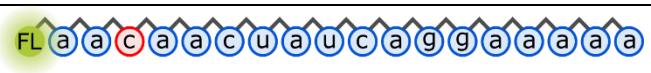
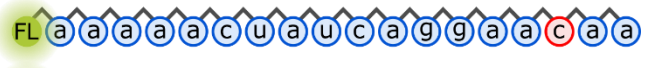
FL-RNA2 induced by unmodified DNA ON40. Surprisingly, target degradation induced by 2′-OMe gapmer ON39 was slower than for AM–PO gapmer ON38 and unmodified ON40. Around 37% of full-length FL-RNA2 remained undegraded even after 3 h of incubation at 20 °C. This means that the previously observed duplex

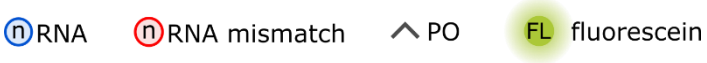
stabilisation for 2′-OMe gapmer ON39 is in an inverse relationship with its catalytic activity at 20 °C. Such an inverse relationship has been previously reported for boranphosphates—note that the RNase H assays in this report are performed at ambient temperature, too.[218] The observed temperature effect on the catalytic performance of 2′-OMe gapmer ON39 compared to ON38 and ON40 is surprising. A possible explanation can be based on the kinetic model[259] described in Section 3.2.3.2 of this chapter, and is summarised in Figure IV-16B. The ON39:FL-RNA2 heteroduplex can be recognised and bound by RNase H to induce hydrolysis of the scissile phosphodiester bond in the FL-RNA2 strand. However, the combination of duplex stabilisation by 2′-OMe modifications together with a lowered incubation temperature (20 °C) could allow ON39 and the cleaved FL-RNA2 fragments to stay hybridised as a duplex with a single strand break. This outcome would prevent the 2′-OMe gapmer from hybridising with a new strand of full-length FL-RNA2 to induce the next catalytic cycle. Moreover, the stabilised duplex with a single strand break would no longer be substrate for RNase H cutting but could still be a substrate for RNase H binding due to the promiscuity of RNase H.[211] Thus, the cut but hybridised duplex would serve as a competitive inhibitor against an uncut ON39:FL-RNA2 heteroduplex for RNase H binding. Both scenarios would eventually lead to an overall inhibition of the catalytic rate.

#### 3.2.3.4. Mismatch Discrimination

Mismatch discrimination is important to ensure minimal off-target effects caused by unintended degradation of RNAs having a high sequence similarity to the

**Table IV-4: List of FL-labelled RNA targets bearing mismatches to ASO sequences FL-RNA3–FL-RNA5.**

| ON | Sequence (5′ → 3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| FL-RNA3 |  | 6314 | 6315 |
| FL-RNA4 |  | 6314 | 6315 |
| FL-RNA5 |  | 6314 | 6316 |




on-target RNA. Previous studies reported enhanced target specificity of chimeric gapmers with charge-neutral methylphosphonates in the wings.[230] As discussed above, the precise cutting pattern resulting from RNase H activation by AM–PO gapmer ON38 (see Figure IV-13B) raises the question of potentially enhanced target specificity. To address this question, ASOs ON38–40 were evaluated to discriminate FL-RNA targets bearing single nucleotide mismatches at various positions. A list of the fluorescently labelled target RNAs is given in Table IV-4. The complementary adenine was exchanged to a cytosine towards the 5′-terminus (FL-RNA3), in the central region (FL-RNA4), or towards the 3′-terminus (FL-RNA5) of the newly synthesised target RNAs to introduce a C–T mismatch at these positions. Note that the central mismatch in FL-RNA4 is within the preferred binding site of the ASO within the PO-binding pocket of RNase H (see Figure IV-14). In addition, all target RNAs (FL-RNA3–5) contain a 5′-terminal fluorescein for monitoring. Each ASO (ON38–40) was separately incubated with each mismatched RNA target (FL-RNA3–5) in the presence of RNase H at 37 °C. In total, this results in nine possible ASO/RNA combinations (Figure IV-17). Thus, aliquots were only taken after 10

and 15 min for each ASO/RNA combination to ensure a reasonable throughput of the experiments. The aliquots were analysed by gel electrophoresis as described in Section 3.2.3.2 of this chapter. Quantification of remaining full-length FL-RNA3–5 from the gels is plotted in Figure IV-17 from triplicates for each set of ASO. All tested ASOs induced retarded degradation of the mismatched RNA targets. This becomes obvious by comparison to the degradation rates of the perfectly matched FL-RNA2 (see Figure IV-15). Around 60% of all mismatched RNA targets were degraded after 15 min when induced by the AM–PO gapmer (ON38 + FL-RNA3, 4, or 5; Figure IV-17). In comparison, the same AM–PO gapmer previously induced

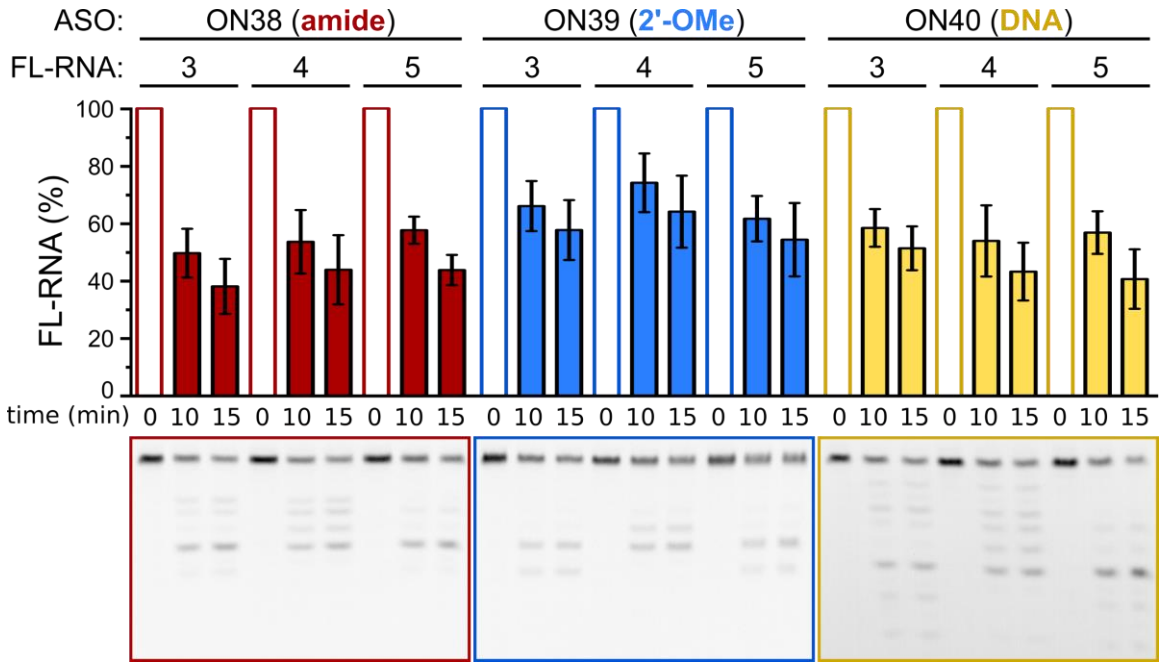



**Figure IV-17: RNase H-mediated degradation of mismatched FL-RNA3–5 after hybridisation with ASOs ON38–40.**
RNA targets FL-RNA3–5 were incubated with ASOs ON38–40 at 37 °C. ASO/RNA target, 1/4 (mole/mole). Aliquots of the reactions were quenched at time points: 0 min, 10 min and 15 min. Degradation fragments of the aliquots were separated by 20% denaturing PAGE and the FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Bars represent quantification of full-length FL-RNA3–5 as the average of triplicates ± standard error of the mean; those were calculated from gels using ImageJ. Quantification of FL-RNA3–5 is normalised to t = 0 min for each set respectively. The reference (t = 0 min) is represented as a no-fill bar for easier comparison between the degradation bars (filled).

target degradation of ca. 95% when hybridised to perfectly matched FL-RNA2 after 15 min under the same conditions (see ON38; Figure IV-15). A similar decrease was observed for the degradation of the mismatched RNAs induced by the 2′-OMe gapmer, and by the unmodified DNA (ON39 and ON40, respectively; Figure IV-17). Whilst all ASOs show mismatch discrimination when compared to degradation of a perfectly matched RNA target, degradation of the mismatched RNA targets tends to be slowest when induced by the 2′-OMe gapmer (ON39 + FL-RNA3, 4, or 5; Figure IV-17).

Interestingly, introduction of a mismatch caused different cutting patterns for all ASOs depending on the location of the mismatch (gels; Figure IV-17). All ASOs resulted in the formation of a main degradation product regardless of the mismatch location, but the formed side products varied in size and amount. In general, the cutting patterns induced by the 2′-OMe gapmer ON39 produced the fewest fragmentation bands visible on the gel. This suggests that the 2′-OMe modification at the wings of a gapmer could enhance cutting precision of an RNA target. Given that the amide modification also has the potential to prevent RNase H activity in its proximity (see results of Section 3.1.2; Chapter IV), further experiments with gapmers combining 2′-OMe and amide modifications might be interesting to evaluate their combined effects on mismatch discrimination and cutting precision.

### 3.2.4. Serum Stability of Oligonucleotides with Consecutive Amides

Serum stability of ASOs is fundamental for therapeutic applications. Resistance against enzymatic degradation by serum phosphodiesterases ensures optimal biological lifetime and therapeutic efficacy of the ASO. To this end, ASOs (ON38–40) were incubated in FBS at 37 °C. Aliquots of the mixtures were quenched over a course of 12 h and analysed by gel electrophoresis. The ASOs, or degradation products thereof, were visualised from the gels by UV irradiation (Figure IV-18). The amide modifications in AM–PO gapmer ON38 substantially increased serum lifetime of the oligomer compared to 2′-OMe gapmer ON39 or unmodified ON40. Whilst ON40 and ON39 were completely degraded after 3 h and 6 h, respectively, full-length ON38 was still present after 12 h in FBS. Both the amide and the 2′-OMe modification have been previously reported to enhance serum stability of oligonucleotides.[62,264] However, a head-to-head comparison of both modifications has not yet been reported. The results in Figure IV-18 show that the serum stability conferred by the amide modification is substantially higher than the enhanced stability resulting from the 2′-OMe wings. Note that AM–PO gapmer ON38 migrated with an apparent molecular weight greater than 28 bases whilst 2′-OMe gapmer ON39 and unmodified ON40 both migrated with molecular weights lower than 28 bases. This difference can be explained by the influence of the non-ionic amide modification on the overall charge of ON38. Indeed, the charge-neutral character of the amide backbone reduces the charge-to-linkage ratio of ON38 to 0.53 (9 negative charges divided by 17 internucleoside linkages). During gel electrophoresis, macromolecules are separated based on their electrophoretic

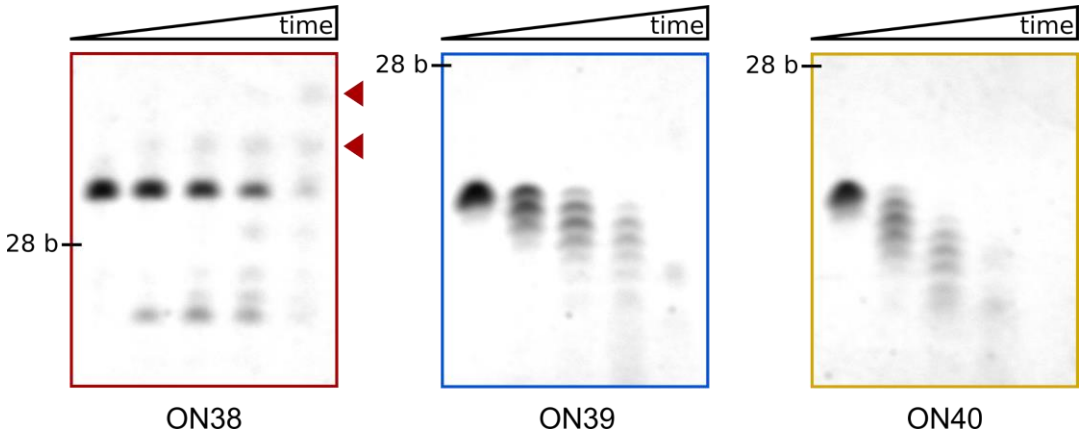



**Figure IV-18: Serum stability of ASOs ON38–40.**
ASOs ON38–40 were incubated at 37 °C in phosphate-buffered saline (PBS)/FBS, 1/1 (v/v). Aliquots were quenched at time points: 0 min, 1 h, 3 h, 6 h and 12 h and separated by 20% denaturing PAGE. Oligonucleotide fragments were viewed under short wavelength UV light. Approximate oligonucleotide lengths of 28 bases (28 b) are indicated for each set. Triangles indicate fragments with a higher apparent molecular weight emerging over time. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

mobility which itself depends on the charge, conformation and length of the oligomer. The length is the same for ON38–40, and the denaturing conditions of the PAGE eliminates possible conformational effects. Because of the reduced charge-to-linkage ratio of ON38, its electrophoretic mobility is reduced leading to slower migration through the gel and a higher apparent molecular weight. Furthermore, degradation of AM–PO gapmer ON38 led to the formation of degradation products that migrated with an even higher apparent molecular weight than full-length ON38 (red triangles; Figure IV-18). This can be explained by further reduction of the charge-to-linkage ratio upon degradation of AM–PO gapmer ON38. For example, a degradation fragment containing only the amide backbone and no phosphodiester bonds would have a charge-to-linkage ratio of 0.00 (0 negative charges). This example will not migrate through the gel. The differences in the electrophoretic

mobilities between ON38–40 is striking and emphasise the profound effect of the amide modification on the biophysical properties of AM–PO gapmer ON38.

### 3.2.5. Cellular Uptake of Oligonucleotides with Consecutive Amides

As discussed in the introduction of this chapter, cellular uptake of macromolecules remains a major challenge for their therapeutic application. Fluorescently labelled analogues of ON38–40 were synthesised to monitor their cellular uptake properties *via* the fluorescent tag (FL-ON38–40, respectively; Table IV-5). Introduction of the fluorescent label was achieved by coupling of FAM **4**, leading to the formation of an additional phosphodiester bond. The four consecutive amides at each wing of FL-ON38 reduce its charge-to-linkage ratio to 0.56. In contrast, all 17 internucleoside linkages of FL-ON39, and FL-ON40 consist of anionic phosphodiester bonds plus one additional 5′-terminal phosphodiester bond for the fluorescein label. Cellular uptake of FL-ON38–40 was evaluated without addition of any transfection agent. HeLa cells were grown for 16 h in serum-free medium containing 5 µM of each fluorescently labelled oligonucleotide. The medium was then removed, and the cells were washed with PBS, fixed and mounted

**Table IV-5: Set of FL-labelled ONs for cellular uptake studies.**

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| FL-ON38 |  | 5677 | 5678 |
| FL-ON39 |  | 6147 | 6149 |
| FL-ON40 |  | 5988 | 5989 |



on a glass slide—conditions: pH 8.4 at 20 °C containing 4′,6-diamidino-2-phenylindole (DAPI) as a nucleus stain. Cellular uptake of FL-ON38–40 was evaluated by visualisation of the fluorescein label using confocal scanning laser microscopy (CSLM) (Figure IV-19). Amide-modified FL-ON38 showed enhanced cellular internalisation compared to FL-ON39 and FL-ON39, for which a clear signal in the FL-channel was mainly absent. In some cases, weak fluorescent signals could be observed but were mostly localised outside from the cell (FL-ON39, FL/DAPI/brightfield; Figure IV-19). In contrast, FL-ON38 showed a clear accumulation of punctuated fluorescent signals and a fluorescent haze inside the cell. The localisation of fluorescence in the form of 'fluorescent dots' suggests entrapment of FL-ON38 inside subcellular compartments. Despite previous suggestions that charge-neutral ASOs can enter the cell through passive diffusion across the cell membrane,[235] FL-ON38 does not seem to freely pass membrane barriers—which is in fact not surprising. Indeed, it has been reported that fully charge-neutral methyl phosphonate-modified ASOs seem to enter the cell *via* fluid phase endocytosis.[231] In this case, uptake was slow and plateaued between 10–12 hours after incubation. Likewise, cellular uptake of amide-modified FL-ON38 was only observed after incubation at a concentration of 5 μM for 16 h but was undetectable for any shorter incubation times (data not shown). Fluid phase endocytosis results in unspecific uptake from the extracellular fluid into vesicles whilst their fate upon maturation can be highly variable.[265] This maturation process can range from the formation of early to late endosomes, lysosomes or even to fusion with the cell membrane for cargo recycling. It is important to mention that the pH

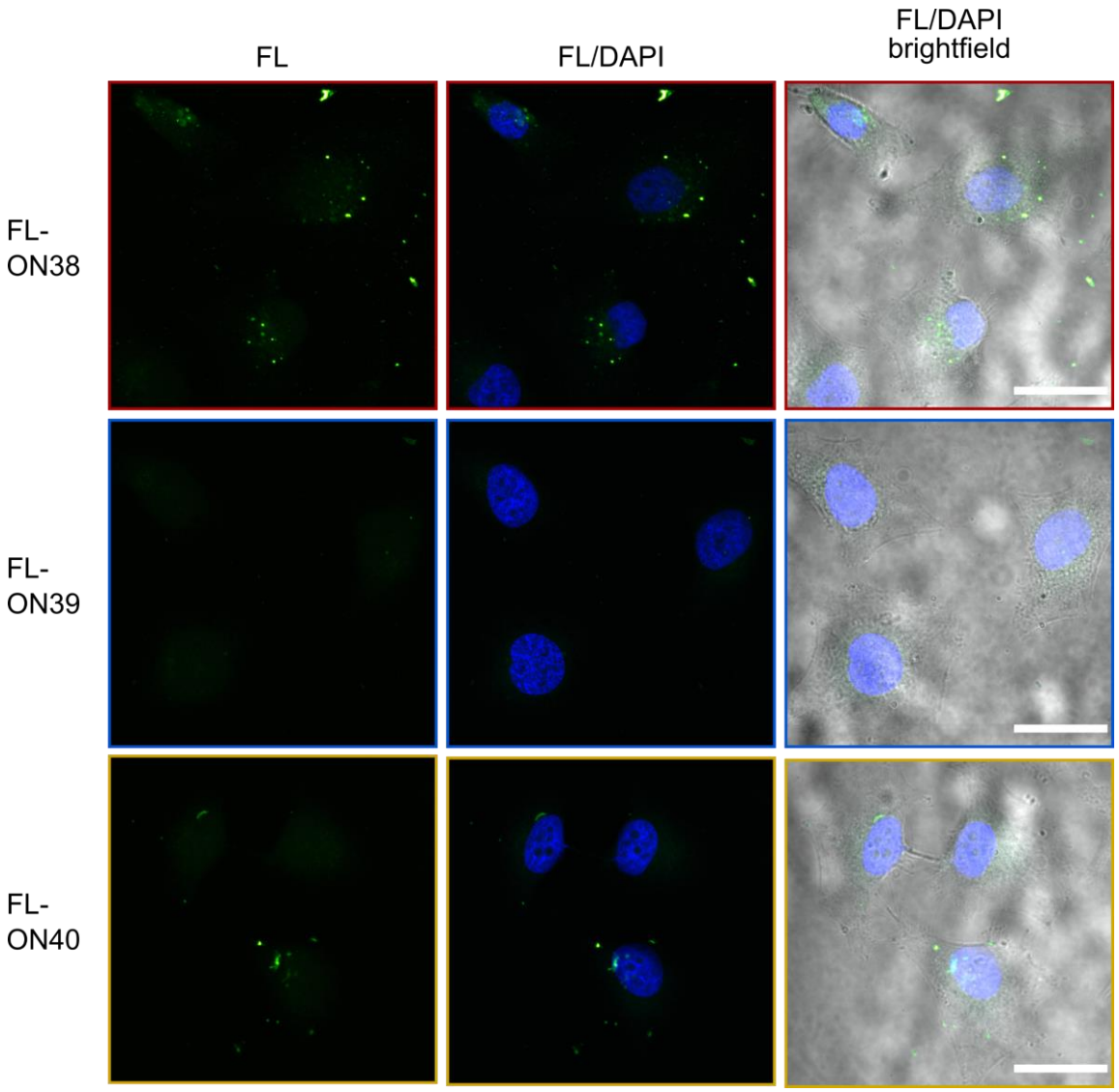



**Figure IV-19: Gymnotic cellular uptake of FL-ON38–40.**
HeLa cells were treated with 5 µM FL-ON38–40 in serum-free growth medium for 16 h at 37 °C without any transfection reagent. CSLM was performed on fixed cells and nuclei were stained with DAPI. Scale bar: 30 µm. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

inside the vacuoles decreases from pH 6.5 (early endosome) to pH 5.5 (late endosome) to pH 4.5 (lysosome).[266] Note that fluorescein is highly sensitive to pH and its fluorescence intensity decreases drastically upon acidification. A change from pH 7.7 to pH 6.7 results in the loss of fluorescence intensity of more than 40% and around 90% of fluorescein is quenched at pH 5.9.[267] This means that detection

of FL-ON38 would be impossible under physiological conditions in the case of its entrapment in acidified vesicles. Indeed, punctuated fluorescence could not be observed by live cell imaging (data not shown) but only after aldehyde cell fixation, which is known to influence endosomal integrity.[268] Successive mounting and storage at pH 8.4 might be the reason why a fluorescent signal could be detected in fixed cells. Even though this does not prove the entrapment of FL-ON38 in acidified subcellular compartments, it emphasises the complexity of cellular uptake and underlines the caution that must be exercised when interpreting the CSLM results.

In order to shed more light on the cellular uptake mechanism, HeLa cells were incubated with 5 μM FL-ON38 in serum-free medium for 16 h and endosomes were stained by incubation with AF647-labelled epidermal growth factor (EGF–AF647) for 30 min prior to cell fixation. Cells were then mounted (pH 8.4, DAPI) and analysed by CSLM (Figure IV-20). Incubation with EGF–AF647 stimulates receptor-mediated endocytosis leading to the formation of subcellular vesicles such as endosomes or lysosomes.[265] Co-localisation of a fluorescent signal in the FL- and the AF647-channels confirms entrapment of FL-ON38 inside subcellular compartments. Figure IV-20 shows bright punctuated fluorescent signals inside the cell for the AF647-channels confirming efficient internalisation of EGF–AF647. Moreover, the fluorescein signals are comparable in intensity and size as seen in the previous experiment (see FL-ON38; Figure IV-19). However, co-localisation of AF647 and fluorescein was only found for a low number of punctuated fluorescent signals as indicated by the arrows in Figure IV-20. This is not surprising because

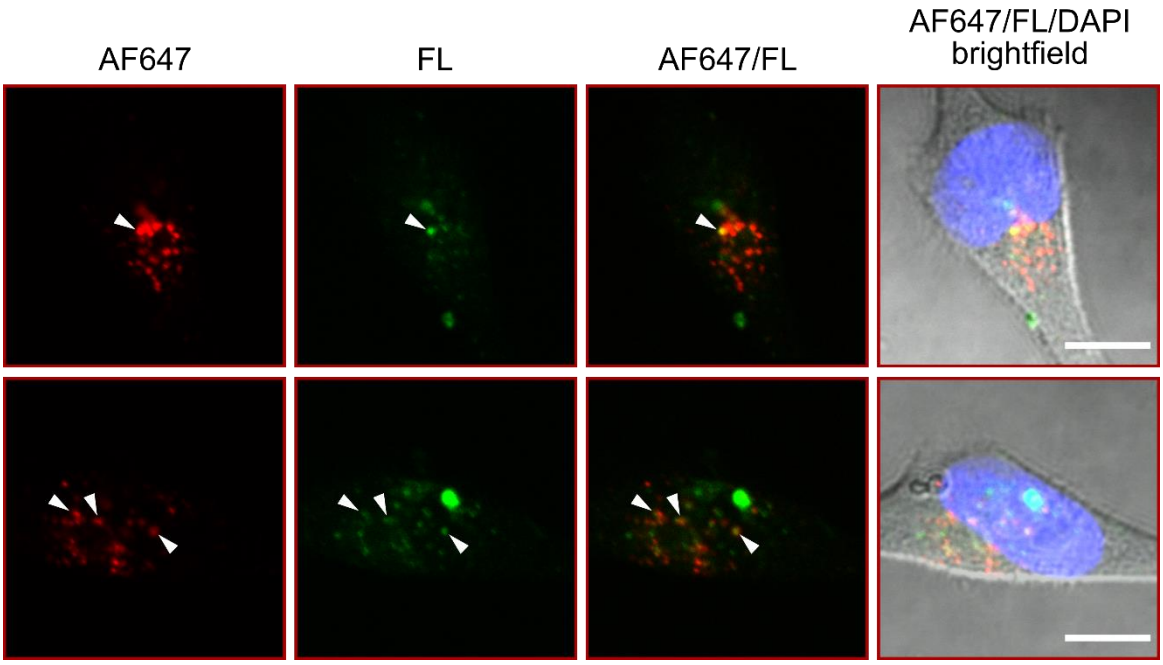



**Figure IV-20: Gymnotic cellular uptake of FL-ON38 and receptor-mediated endocytosis of EGF–AF647.**
HeLa cells were treated with 5 µM FL-ON38 in serum-free growth medium for 16 h at 37 °C without any transfection reagent. Endosomes were stained by incubation with an EGF–AF647 complex (2 µg·mL$^{-1}$) for 30 min at 37 °C before cell fixation. Nuclei were stained with DAPI and cells analysed by CSLM. Scale bar: 10 µm. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

different uptake mechanisms (e.g., receptor-mediated vs. non-receptor-mediated) can lead to a different fate upon vesicle maturation.[265] Simultaneous uptake or fusion of vesicles containing FL-ON38 and EGF–AF647 would lead to co-localisation of the respective signals. In fact, the sporadic co-localisation of AF647 and FL signals supports the hypothesis that FL-ON38 is indeed entrapped inside subcellular compartments and part of the intracellular vesicle trafficking system, even though the exact type of vesicle remains elusive.

# 4. Conclusion

The amide bond is an attractive chemical modification to mimic the natural phosphodiester bond in antisense oligonucleotides. However, its applicability in the context of RNase H activation has not yet been investigated. This chapter reports the synthesis of amide-modified oligonucleotides for RNase H, serum stability, and cellular uptake studies. To this end, amide-modified oligonucleotides were synthesised using the dimer and monomer strategies during solid phase oligonucleotide synthesis. Importantly, on-resin amide bond formation was achieved in high purity and good yield, allowing the introduction of consecutive amide backbones. Considering recent reports suggesting that mixed-backbone PS–ASOs are promising for next-generation antisense therapeutics,[93,117,128,129] the on-resin amide bond formation represents a valuable platform to develop AM–PS mixed-backbone ASOs. In fact, the authors of a study highlighting the beneficial effects of charge-neutral backbones—such as the presented amide linkage—specifically state that introduction of artificial backbones through the dimer approach *"severely limits their versatility and broader application."*[129]

Oligonucleotides with isolated amide backbones were rationally designed to test whether the amide bond can replace one of the three natural phosphodiester bonds interacting with RNase H. However, the inability of these amide-modified oligonucleotides to elicit an RNase H response demonstrated that the amide

modification is not tolerated inside the phosphodiester bond-binding pocket of RNase H.

A revised gapmer design was applied for the synthesis of ON38, which contained four consecutive amide linkages at either wing, and an unmodified gap of nine nucleotides. AM–PO gapmer ON38 activated RNase H to degrade a fully complementary RNA target as efficiently as its 2′-OMe gapmer and its unmodified DNA analogues. It is likely that the efficient activation of RNase H originates from the flexibility of the amide linkage to adopt both an A- and a B-form helical structure, depending on the target.

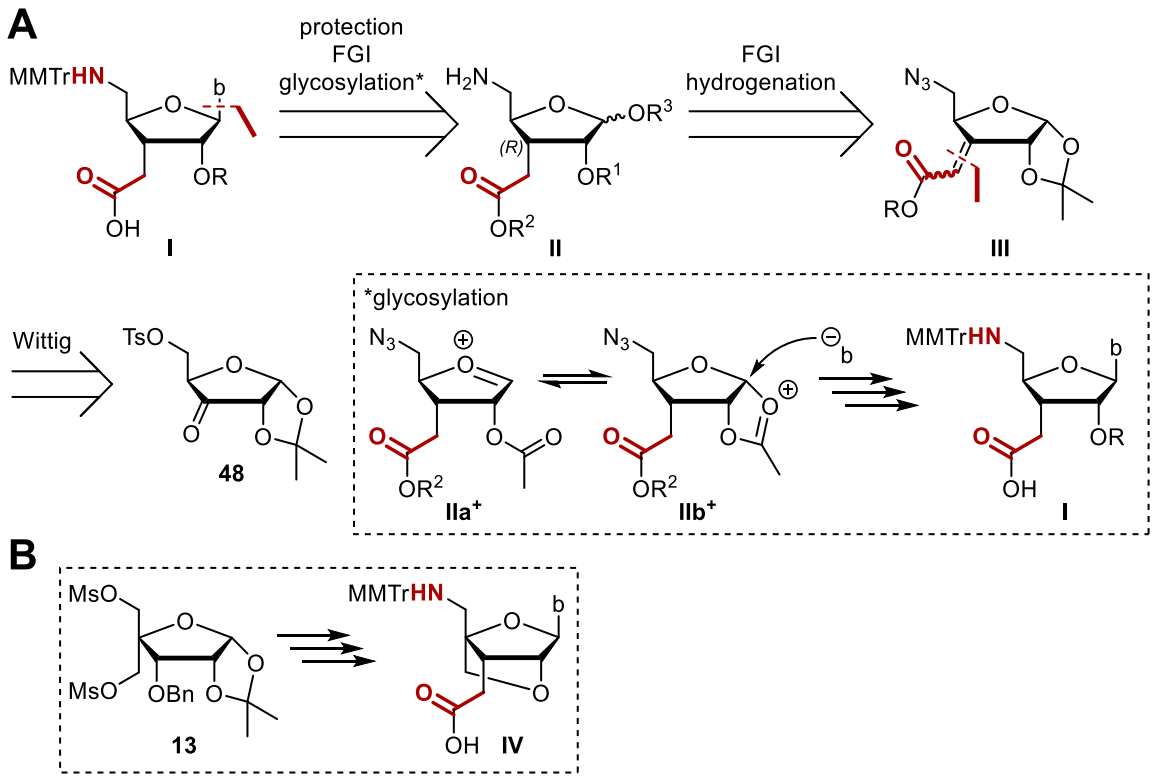
The amide backbone in AM–PO gapmer ON38 and its fluorescent analogue FL-ON38 enhanced desirable drug-like properties. As such, AM–PO gapmer ON38 was significantly more stable in serum than the corresponding 2′-OMe gapmer or unmodified DNA. Moreover, the fluorescently labelled AM–PO gapmer FL-ON38 showed potentially enhanced cellular uptake, which remains a major challenge for therapeutic oligonucleotides. However, the observed punctuated fluorescence indicated the entrapment of AM–PO gapmer FL-ON38 in subcellular compartments such as endosomes. Further assays with a functional read-out are necessary to determine whether AM–PO gapmer FL-ON38 can exit the subcellular compartments to produce enhanced biological activity.

# 5. Future Work

Whilst the amide backbone represents a promising modification for antisense oligonucleotides, its further development is largely hampered by the synthetic inaccessibility of the amino–carboxylic acids nucleosides with different bases (see Chapter III). The synthetic endeavours of Chapter III focused on the preparation of DNA-type amino–carboxylic acids nucleosides to test whether the amide bond is tolerated by RNase H. However, the results of Chapter IV show that the amide modification was not tolerated within the phosphodiester bond-binding pocked of RNase H but was well tolerated in the wings of an AM–PO gapmer. Moreover, 2′-modifications of the ribose are often used in gapmer designs (see Table I-2, Chapter I). Together, this means that 2′-substituted amino–carboxylic acids nucleosides might be a promising way forward, which opens the possibility for new synthetic approaches.

## 5.1. New Chemical Approaches for Monomer Synthesis

The presence of a 2′-substitutent would allow the stereoselective late-stage glycosylation of a nucleoside precursor through neighbouring group participation (Scheme IV-2A). Target amino–carboxylic acid nucleoside **I** could be accessible through functional group interconversions from **II** to introduce the required MMTr protection group at the 5′-nitrogen, and by hydrolysing an ester to the free 3′-carboxylic acid. Moreover, the base in **I** could be stereoselectively glycosylated *via*




**Scheme IV-2: Introduction of sugar modifications.**
(**A**) Retrosynthetic analysis to introduce further sugar modifications at the ACA nucleoside. The neighbouring group participation to direct the β-stereochemistry of the glycosidic bond in the glycosylation reaction is shown in the box. (**B**) A proposed LNA–amino–carboxylic acid nucleoside that could be accessible from commercially available **13** through chemical transformations based on (**A**).
Abbreviations: b, nucleobase; FGI, functional group interconversion; MMTr, 4-methoxytrityl; Ms, methanesulfonyl.

the oxonium ion intermediate **IIa**$^+$ (box; Scheme IV-2A). This oxonium ion can be stabilised by neighbouring group participation of the 2′-substituent (e.g., *O*-acetyl) to form intermediate **IIb**$^+$, which directs the incoming nucleobase to the β-side. Glycosyl donor **II** could be accessible from a global hydrogenation of **III**, similar to reported procedures for the synthesis of C3′-homologated nucleosides.[206] Importantly, the stereochemical outcome of the reaction is likely to exclusively produce the *R*-isomer because of hydrogenation from the convex face of the bicyclic

system.[269] The C–C bond formation at C3′ could arise from a Wittig reaction with ketone **48**.[270]

Expanding on the possibility to introduce a 2′-functionality for the synthesis of 2′- and amide-modified gapmers, the locked nucleic acid modification might be another promising candidate to combine sugar and backbone modification as seen in Chapter II for LNA–carbamate chimeras. A suitable bicyclic amino–carboxylic acid nucleoside (**IV**; Scheme IV-2B) could be accessible form acetonide **13** based on the same neighbouring group participation principles as outlined in Scheme IV-2A.

A synthetic procedure to access sugar-modified amino–carboxylic acid nucleosides would further allow a revised gapmer design to investigate cellular antisense activities. This could comprise the synthesis of 2′-functionalised AM–PO gapmers to target a reporter gene (e.g., *luciferase*) in readily available reporter cell lines, or to replicate the sequence of a clinically approved gapmer for direct comparison studies. Alternatively, cell lines could be genetically engineered to produce an mRNA with a suitable ASO target sequence within its 3′ or 5′UTR. RNase H-mediated degradation of either the 3′ or 5′UTR would lead to destabilisation and rapid degradation of the corresponding mRNA due to its loss of the stabilising 5′ cap or 3′ poly(A) tail, respectively. Because changes in the 3′ or 5′UTR do not affect the protein's sequence, such alterations could be applied to reporter genes encoding for a protein with a direct expression level readout, e.g., different light emission intensities correlating with different levels of luciferase

expression. Moreover, targetable sites could be engineered into endogenous genes, either within the untranslated regions or within the protein coding sequence if not lethal for the cell. Several RNA quantification kits are readily available for a variety of endogenous genes.[271] Such assays rely on labelled reporter oligonucleotide constructs that hybridise to the mRNA target for direct quantification from cell lysates. Thus, suitable ASO target sequences could be introduced within an endogenous gene and knockdown efficiencies could be evaluated by quantification of the corresponding mRNA levels.

# CHAPTER V

# TRIAZOLE-LINKED OLIGONUCLEOTIDES FOR SYNTHETIC BIOLOGY

# 1. Background

## 1.1. Formation of Triazoles

The formation of triazoles has myriad applications in the life sciences. Its power lies in the modular approach of reacting an alkyne with an azide to form a 1,2,3-triazole *via* a 1,3-dipolar cycloaddition (Figure V-1).[272] The regiochemistry of the cycloaddition determines whether the resulting triazole is 1,4- or 1,5-disubstituted. Moreover, the reaction has full atom economy and the involved reactive groups (azide/alkyne) are absent from biological systems, rendering this reaction bioorthogonal to other functional groups of biomolecules.[273] Unsurprisingly, the ligation of alkynes and azides to form triazoles has found many applications for the chemoselective labelling of biomolecules.[274–277] However, the traditional, non-catalysed azide–alkyne Huisgen cycloaddition requires high temperatures, long reaction times, and results in poor regioselectivity between 1,4- and 1,5-disubstituted 1,2,3-triazoles.[272] The development of metal ion-catalysed azide–alkyne cycloadditions has led to a significant enhancement of reaction rates and

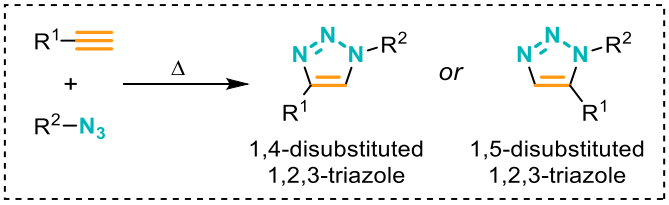



**Figure V-1: The azide–alkyne Huisgen cycloaddition.**
The heat-activated 1,3-dipolar cycloaddition of an alkyne and an azide leads to the formation of a 1,2,3-triazole. Depending on the regiochemistry, the resulting 1,2,3-triazole can be 1,4- or 1,5-disubstituted.

facilitated the control over the regiochemistry.[143,144,278–283] The following paragraphs will introduce the copper(I)-catalysed azide–alkyne cycloaddition[143,144] and the ruthenium(II)-catalysed azide–alkyne cycloaddition (RuAAC).[279,280]

### 1.1.1. The Copper(I)-Catalysed Azide–Alkyne Cycloaddition

The CuAAC is arguably the most widely used metal ion-catalysed version of the azide–alkyne Huisgen cycloaddition. Its broad scope, high yield, and the low amount of byproducts define the CuAAC as a 'click' reaction.[284] Even though the term 'click chemistry' comprises many other types of reactions, the CuAAC is so predominant that the term 'click' is often used synonymously. The copper(I)-catalysed cycloaddition of an azide and an alkyne always leads to a 1,4-disubstituted 1,2,3-triazole.[36] The underlying mechanism has been intensively studied and involves two copper nuclei of oxidation state +1 (Figure V-2).[36,285–287] Whilst the reaction outcome still resembles the result of the concerted, pericyclic 1,3-dipolar cycloaddition (box; Figure V-2), copper(I) catalysis follows a stepwise mechanism in which the copper(I) catalyst has two functions: (**i**) addition to the alkyne and complexation with the azide to accelerate the reaction (ca. $10^7$–$10^8$-fold);[284] and (**ii**) organisation of the reaction partners to regioselectively form the 1,4-disubstitited 1,2,3-triazole (Figure V-2).[36] The formation of the Cu-acetylide intermediate has significant consequences as it only allows terminal alkynes to enter the catalytic cycle. The short reaction times under physiological conditions (room temperature, aqueous), and the bioorthogonality make the CuAAC a powerful ligation reaction for research in chemical biology. Moreover, the stereochemical control of the

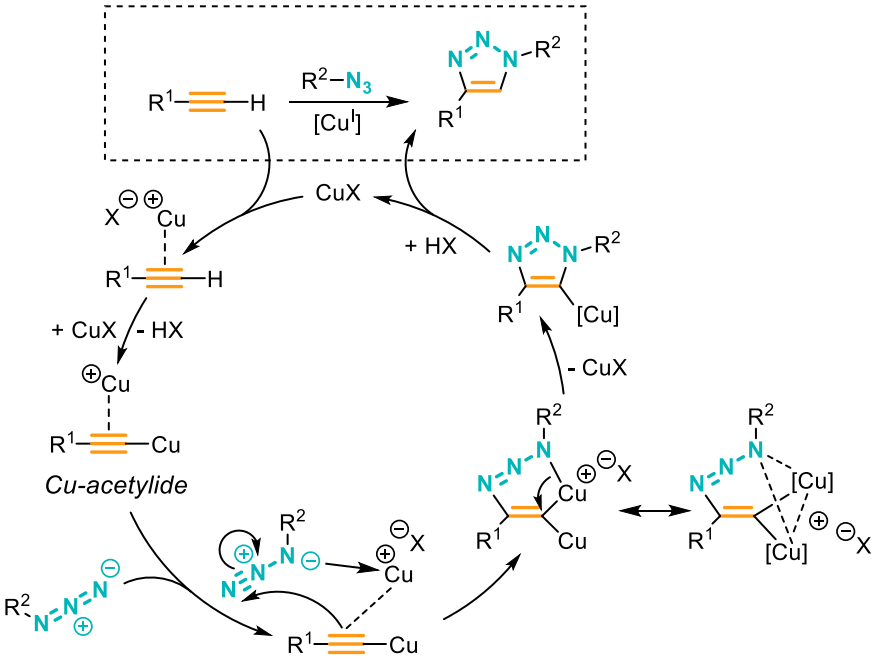



**Figure V-2: The mechanism of the dinuclear copper(I) catalysis involved in the CuAAC.** The cycloaddition of an alkyne and an azide to form a 1,4-disubstituted 1,2,3-triazole is mostly catalysed by two copper nuclei. Note that the $Cu^{I}$ species is normally stabilised by the addition of ligands, which are omitted in this representation for simplicity. *Figure adapted from literature.*[36,285–287]

CuAAC enables the construction of defined (biomimetic) macromolecules. As such, 1,2,3-triazoles with a defined substitution pattern mimicking internucleoside backbones will be the focus of this chapter.[36]

### 1.1.2. The Ruthenium(II)-Catalysed Azide–Alkyne Cycloaddition

Several alternative synthetic approaches have been developed to obtain 1,5-disubstituted 1,2,3-triazoles.[279,280,283,288–292] Of those, the RuAAC[292] has emerged as the most widely adopted metal ion-catalysed azide–alkyne cycloaddition.[279,280] Many ruthenium-based catalysts have been developed and the stereochemical outcome of the RuAAC depends on the ligands of the catalyst. As such, ruthenium catalysts with little steric hindrance can afford the 1,4-disubstitution, whilst the

presence of a sterically more hindered pentamethylcyclopentadienyl (Cp*) ligand results in 1,5-disubstituted 1,2,3-triazoles.[36,279] Although the mechanism of the RuAAC is not yet fully elucidated, a computational study led to the proposed intermediates of Figure V-3.[36,280] The reaction is catalysed by the alkyne's increased nucleophilicity through π-interactions with the ruthenium core. Moreover, no Ru-acetylide is formed in comparison to the Cu-acetylide in the CuAAC (see Figure V-2), which allows internal alkynes to be a substrate in the RuAAC. Despite the metal ion catalysis, RuAAC requires elevated temperatures and only proceeds in aprotic solvents under an inert gas atmosphere.[279] Nevertheless, RuAAC can be performed for ligation of biomolecules such as peptides[293] or nucleosides.[72] The exploration of

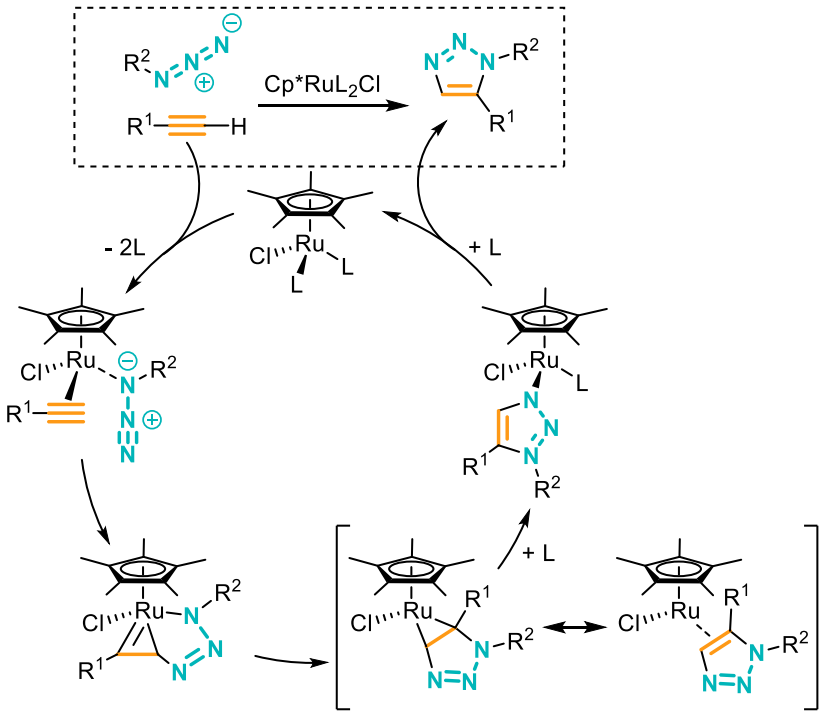



**Figure V-3: Proposed intermediates of the Ru-catalysis involved in RuAAC.**
The proposed intermediates are derived from experimental observations and DFT calculations. Note that the proposed absence of a Ru-acetylide rationalises the fact that internal alkynes can be a substrate for RuAAC. *Figure adapted from literature.*[36,280]

the RuAAC to generate a 1,5-disubstituted triazole as an internucleoside backbone mimic will be part of this chapter.

### 1.1.3. The Structural Diversity of the Triazole Internucleoside Linkage

The different 1,2,3-triazole regioisomers that originate from CuAAC or RuAAC enable the investigation of structurally diverse triazoles as internucleoside backbone mimics (Figure V-4). Combined with different azide and alkyne nucleoside substrates, ligation *via* CuAAC or RuAAC can generate various triazole linkages (TLs). The molecular details of a diverse set of TLs that are relevant for this chapter are shown in Figure V-4 (TL**1**,[294] novel TL**2**, TL**3**,[72] TL**4**,[149] TL**5**,[70] TL**6**,[153] TL**7**[48]). The TLs are arranged by increasing linkage length, which is defined as the minimum number of atoms connecting the sugar rings. The linker length of the 4-atom TL**1** closely resembles the natural phosphodiester bond. The 5-atom TLs **2**–**4** are slightly longer, and the 6-atom TLs **5**–**7** are the longest internucleoside linkages in this set. The effect of CuAAC *versus* RuAAC results in a different number of $sp^2$-hybridised centres along the bonds connecting the sugar rings, and thus results in a different number of atoms that are constrained in one plane along the linkage (starred atoms; Figure V-4). As a result of the 1,5-disubstitution, TL**2** and TL**3** only have four atoms in one plane, whilst 1,4-disubstituted TL**1** and TLs **4**–**7** have five atoms in one plane along the internucleoside linkage. The effect on the overall flexibility of the triazole linkages will be subject of the discussion later in this chapter. Except for TL**2**, all other triazole linkages of Figure V-4 have been previously reported in the literature.[48,70,72,149,153,294] For some of them, additional applications in synthetic

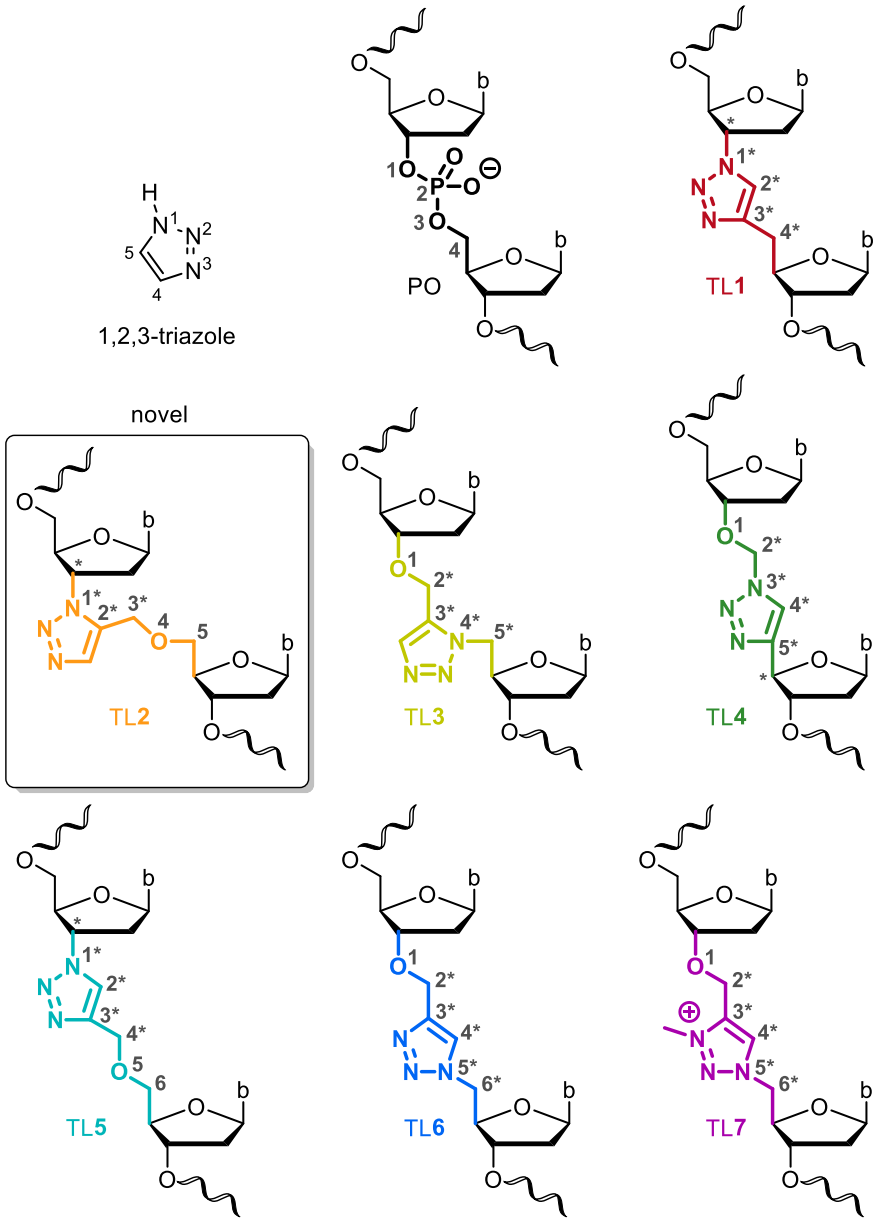



**Figure V-4: Overview of the molecular structures of selected triazole internucleoside linkages investigated in this chapter.**
The general structure of a 1,2,3-triazole and the natural PO are shown for reference. The number of atoms connecting the sugar rings (= linker length) are indicated. The atoms along the linkage that are constrained in one plane are marked with (*). Novel TL**2** is highlighted in the box. Abbreviations: b, nucleobase.

biology have been reported. The following paragraphs will introduce the literature-reported TLs and their application in synthetic biology with a focus on the read-through compatibility by polymerases.

## 1.2. Triazoles in Genetic Templates

The assembly of backbone-modified oligonucleotides can be achieved by a variety of ligation reactions to form various artificial backbones (see Section 3.3.2; Chapter I). As discussed in Chapter I, the ability to be read by polymerases is fundamental for any application of a backbone modification in genetic templates. As such, some of the above-mentioned triazoles (Figure V-4) have been studied as part of genetic templates for various applications in synthetic biology and biotechnology.[48,83,149–155]

### 1.2.1. Preparation of Libraries

Replication-compatible triazole linkages have the potential to facilitate the preparation of Next-Generation Sequencing (NGS) libraries. NGS—also known as sequencing by synthesis—uses a strand with an unknown sequence as a template to elongate a primer. Sequencing is performed by iterative cycles in which fluorescently labelled dNTPs compete for addition to the growing primer.[295] This process can be monitored and used to deduct the sequence of the template. However, initial hybridisation of the primer with the template is necessary, which requires a sequence-defined region of the template. This is usually achieved by enzymatic ligation of adaptor sequences to the unknown template.[295] The CuAAC can be a powerful tool for chemical ligation of oligonucleotides. In this context, a study by Miura et al. reported the concept of using CuAAC to ligate adaptor oligonucleotides to templates for the preparation of NGS libraries (Figure V-5).[86] In this study, the authors used CuAAC to form TL**1** as the artificial internucleoside linkage (Figure V-

5A). The 3′-azide labelling of the template with the unknown sequence relied on enzymatic azide tailing with a 3′-azide dNTP using terminal deoxynucleotidyl transferase (TdT, Figure V-5B). The resulting 3′-azide-labelled template could then be ligated to a synthetic, 5′-alkyne-modified adaptor oligonucleotide to form TL**1** *via* CuAAC. The ligation of the adaptor oligonucleotide introduces a section with a

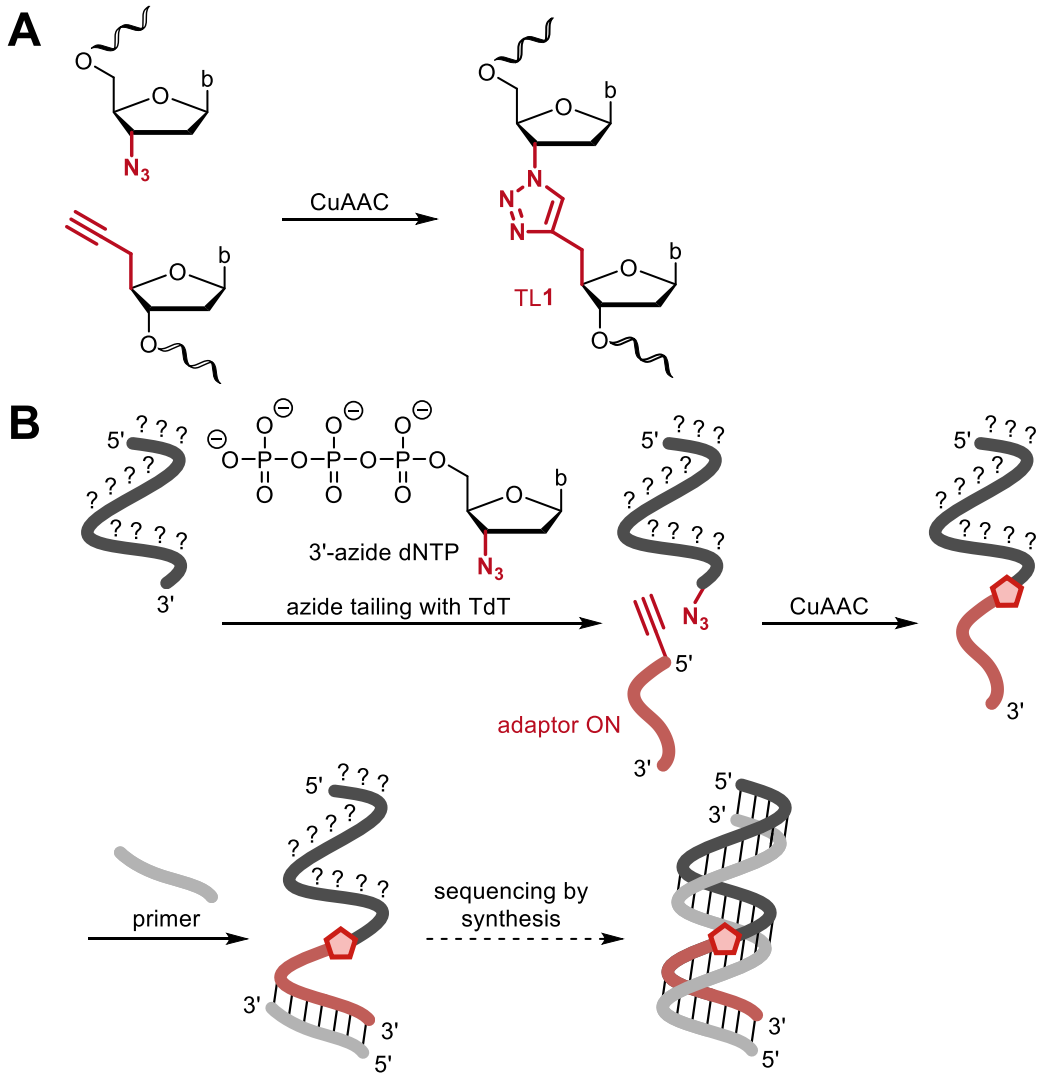



**Figure V-5: Synthesis and application of TL1 for the preparation of NGS libraries.**
(**A**) The principle of generating TL**1**-linked longer ONs *via* CuAAC of alkyne- and azide-modified shorter ONs. (**B**) Application of TL**1** in NGS. The addition of a 3′-azide group onto an unknown ON can be achieved by azide tailing with TdT and a 3′-azide dNTP as a substrate. The template-free ligation of a 5′-alkyne adaptor ON attaches a known sequence (light red) to the 3′-azide-labelled template *via* TL**1** (red pentagon). Hybridisation of a primer (light grey) to the adaptor region can facilitate sequencing of the template by synthesis.

known sequence that can be used for primer hybridisation. In theory, the primer could then be extended along the unknown template using NGS cycles to determine the sequence of the template. Unfortunately, the authors reported inefficient read-through of TL**1** by 16 different polymerases, which prevents the sequencing by synthesis approach.[86] Moreover, TL**1** has been reported to be mutagenic when read by several polymerases,[48] which further limits its applications when correct read-through is required for sequencing.

### 1.2.2. Triazole-Modified Primers in PCR

Short, triazole-linked oligonucleotides can be used as primers in PCR reactions to produce longer oligonucleotides with triazole backbones.[149,153] Varizhuk et al. investigated the compatibility of TL**4**-modified primers in PCR amplifications (Figure V-6).[149] In this study, primers with single incorporations of TL**4** at different sites were elongated by Taq or Pfu polymerase. Note that this experiment will lead

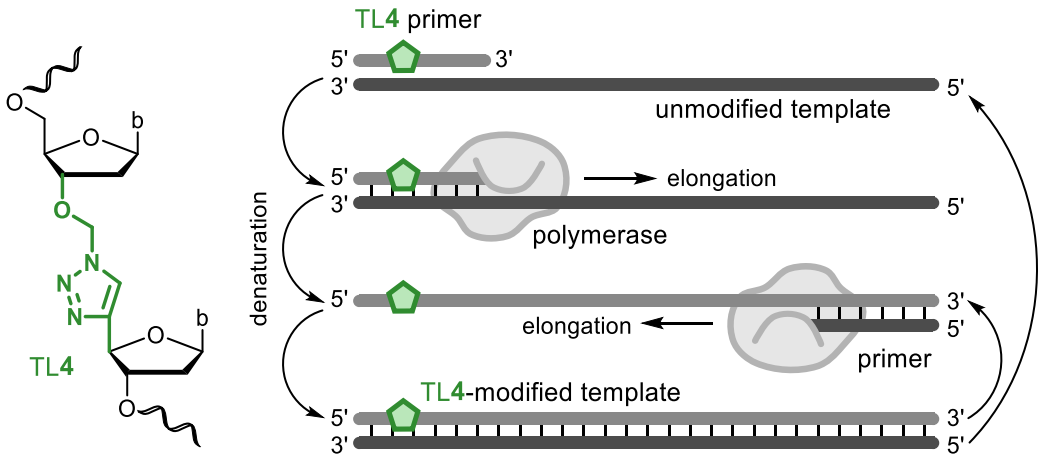



**Figure V-6: Applications of TL4-modified primers.**
The TL**4** modification can be part of a primer that is extended by a polymerase during PCR. After denaturation of the formed hybrid, the extended TL**4**-modified primer can serve as a TL**4**-modified template in the next PCR cycle.

to the formation of a TL**4**-modified template, meaning that the compatibility of TL**4** as part of a primer and as part of a template is simultaneously assessed. The authors reported efficient amplification with Taq polymerase whereas Pfu polymerase did generally not tolerate the TL**4**-modified primers. However, incorporation of TL**4** close to the 3′-end of the primer also resulted in a substantial decrease of product formation in quantitative PCR (qPCR) experiments with Taq polymerase.[149]

### 1.2.3. Synthesis of Triazole-Linked Genes

The formation of replication-compatible artificial backbones *via* ligation reactions of chemically modified shorter oligonucleotides can be utilised to assemble long oligonucleotides (>200 base pairs), which can be the size of small genes (Figure V-7A). Importantly, this allows the precise introduction of additional modifications (e.g., $^{Me}$C as epigenetic markers) at defined sites of the gene. To this end, TL**6** has been intensively studied for the synthesis of genes from azide- and alkyne-modified gene fragments. The first study of synthesising genetic material by CuAAC of alkyne- and azide-modified shorter oligonucleotides to form TL**6** reported the assembly of a 300-mer from three chemically modified 100-mers.[153] Moreover, the authors showed that the assembled 300-mer was copied during PCR with high fidelity.[153] A later study reported the self-templated CuAAC ligation of alkyne- and azide-labelled shorter oligonucleotides to assemble a 335-base pair *ilov* gene *via* formation of eight triazoles (Figure V-7B). This gene was further inserted into a plasmid and introduced into *E. coli* which then expressed the green-fluorescent iLov protein (Figure V-7C).[154] In a related study, a plasmid with a TL**6**-

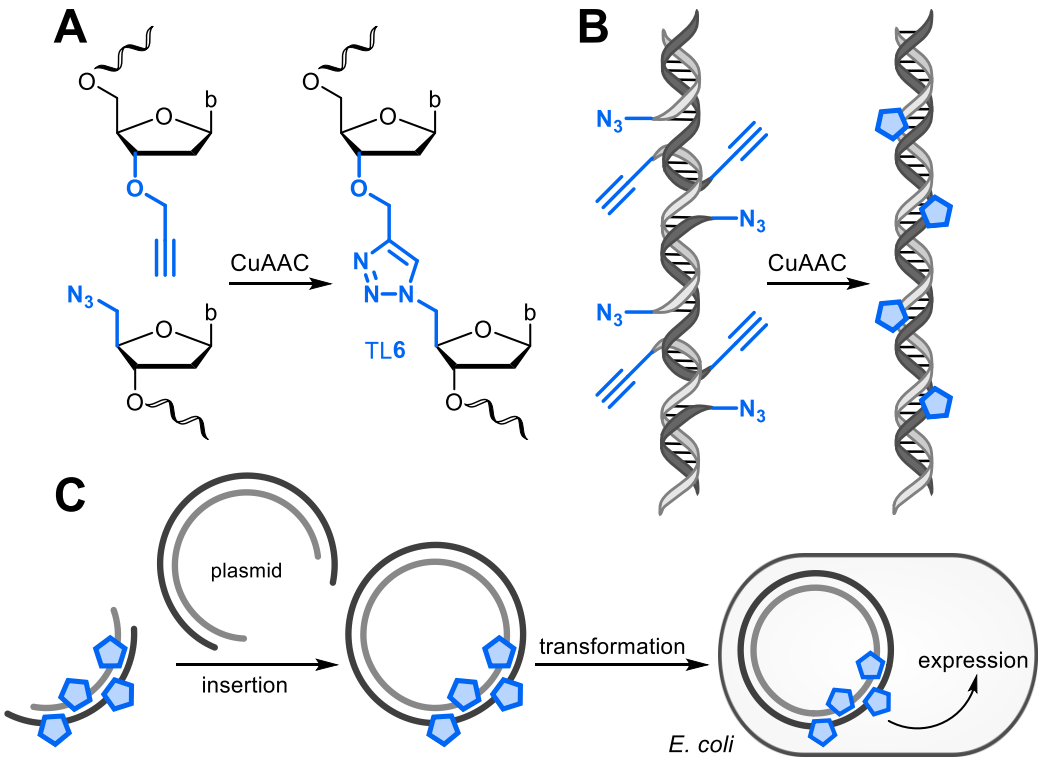



**Figure V-7: Formation of TL6 to assemble TL6-modified genetic polymers.**
(**A**) The principle of generating TL**6**-linked longer ONs *via* CuAAC of alkyne- and azide-modified shorter ONs. (**B**) Self-templated CuAAC of alkyne- and azide-modified, shorter ONs to assemble a 335-base pair, TL**6**-modified *ilov* gene. Note that not all eight triazoles (blue pentagons) are shown for simplicity. (**C**) Ligation of a TL**6**-modified gene into a vector and successive transformation of *E. coli* can lead to expression of the protein that is encoded by the TL**6**-modified region of the plasmid.

modified *mcherry* gene has been introduced into mammalian[150] and bacterial[155] cells, which then expressed the red-fluorescent mCherry protein. However, the presence of TL**6** destabilises the duplex with a complementary DNA or RNA strand[72] and induces a TL**6**-dependent slowdown in PCR experiments.[154]

### 1.2.4. Circular Genetic Templates

Besides gene synthesis, the formation of TL**6** by CuAAC has been reported as a method of self-ligation to produce circular genetic templates (Figure V-8). Circular templates can be a substrate for rolling circle amplification (RCA), which

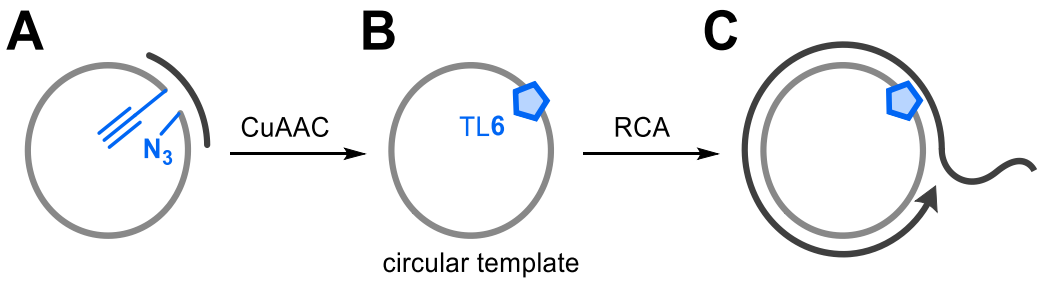



**Figure V-8: Synthesis and RCA application of TL6-modified circular templates.**
(**A**) The modified ON has terminal azide and alkyne groups that are highlighted in blue. The splint template (dark grey) brings the reactive ends in close proximity. (**B**) CuAAC forms TL**6** (blue pentagon) to produce a circular template. (**C**) Extension of a primer by a strand-displacing polymerase (not shown) leads to replication around the circular template to produce concatemers as indicated by the dark grey arrow.

is a versatile tool for chemical biology, materials science, and medicine.[296] Extension of a primer by a strand-displacing polymerase generates a replicon with multiple copies of the circular template sequence—these are called concatemers. Such RCA products can give rise to fascinating nanostructures like DNA nanoflowers.[83] It is essential that the strand-displacing polymerase can continue replication through the artificial linkage to produce the RCA concatemers. As such, the presence of TL**6** in a circular template is compatible with RCA[153] and can be utilised to produce DNA nanoflowers.[83]

# 2. Aim of the Chapter

Whilst a diverse set of triazole linkages has been reported to date (Figure V-9), there are certain limitations for their applications in genetic templates. For example, TL**1** is incompatible with efficient polymerase read-through[48,86] and TL**4** and TL**6** slow down PCR amplifications.[149,154] Moreover, the read-through compatibilities of TL**3** and TL**5** have not yet been reported. The aim of this chapter is to introduce a novel triazole—TL**2**—as an internucleoside linkage into oligonucleotides and test

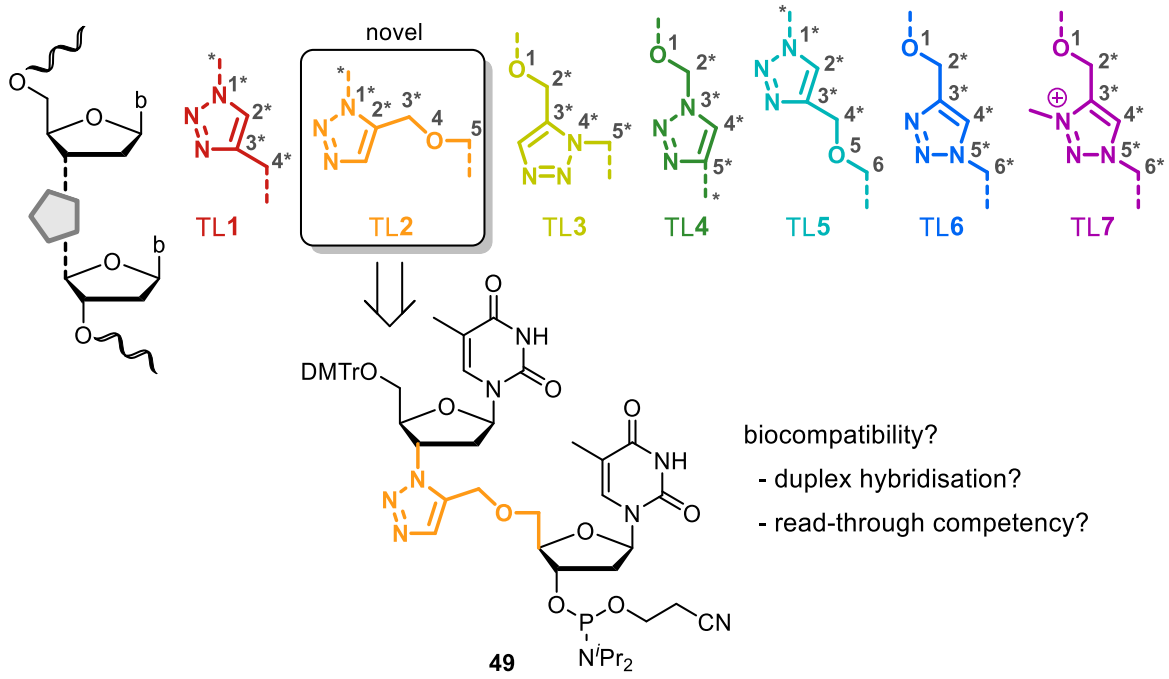



**Figure V-9: The proposed synthetic accessibility of TL2-modified ONs through T$^{TL2}$T-dimer CE phosphoramidite 49.**
The set of TLs which are proposed for the biocompatibility study are shown at the top and novel TL**2** is highlighted in the box. The number of atoms connecting the sugar rings (= linker length) are indicated in grey. The atoms along the linkage that are constrained in one plane are marked with (*). Compound **49** shows the proposed structure through which TL**2**-modified ONs are accessible and the TL**2** moiety is highlighted. Abbreviations: b, nucleobase; DMTr, 4,4′-dimethoxytrityl; $^{i}$Pr, isopropyl.

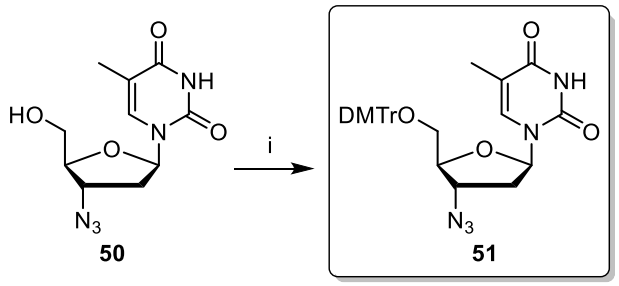
its biocompatibility. The proposed linker contains a 1,5-disubstituted 1,2,3-triazole and connects the sugar ring through five atoms. These properties could confer favourable biocompatibility properties to TL**2**. Its introduction requires $T^{TL2}$T-dimer CE phosphoramidite **49** for solid phase oligonucleotide synthesis. To this end, RuAAC will be explored to generate the desired 1,5-disubstituted 1,2,3-triazole from suitable alkyne and azide nucleoside precursors. Upon its successful synthesis, $T^{TL2}$T-dimer CE phosphoramidite **49** will be used to synthesise TL**2**-modified oligonucleotides for biocompatibility assessment. This will comprise the evaluation of their ability to form duplexes with complementary DNA and RNA targets and their read-through compatibility by polymerases.

# 3. Results and Discussion

## 3.1. Synthesis of Triazole-Linked Dinucleoside Phosphoramidites

### 3.1.1. Synthesis

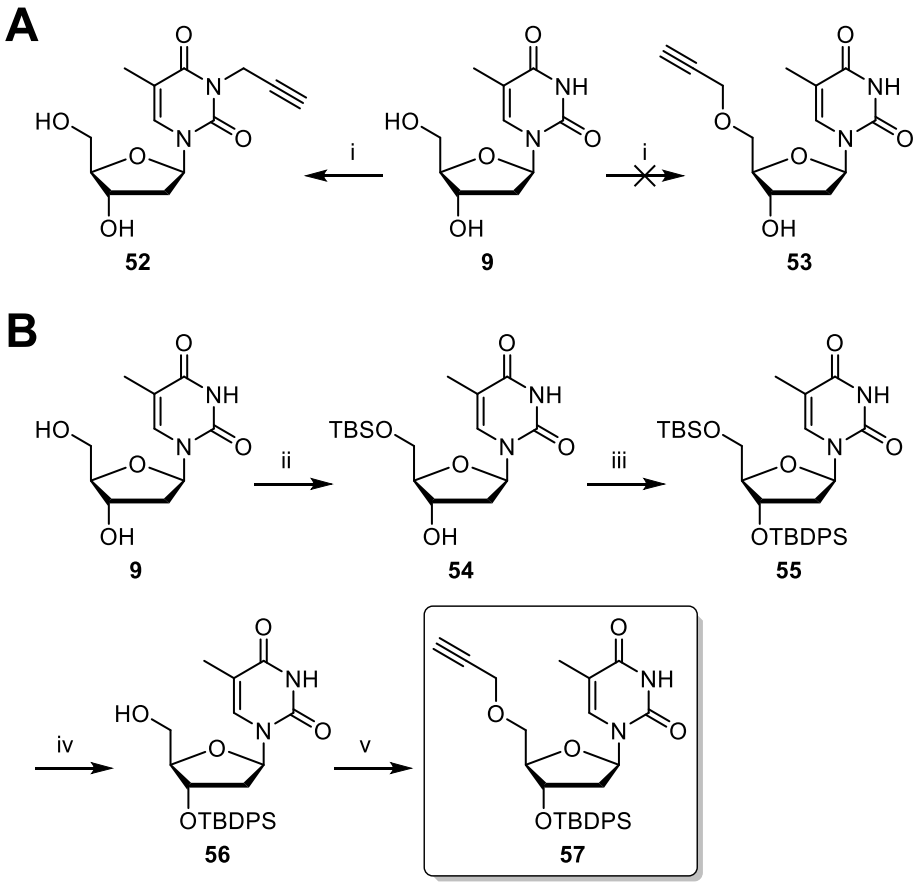
The preparation of a dinucleoside CE phosphoramidite containing the novel 1,5-disubstituted TL**2** required the preparation of 3′-azide- and 5′-alkyne-functionalised nucleosides for successive RuAAC. The synthesis of azide **51**[297] was achieved in a single step to protect the 5′-OH of commercially available azidothymidine (**50**) with a DMTr-group (Scheme V-1).




**Scheme V-1: Synthesis of DMTr-protected azidothymidine 51.**
Reagents and conditions: (i) 4,4′-dimethoxytrityl chloride, pyridine, 16 h, rt, 89%. Abbreviations: DMTr, 4,4′-dimethoxytrityl.

The synthesis of the 5′-alkyne-modified nucleoside required an orthogonal protecting group strategy to ensure chemo- and regioselectivity of the propargylation reaction. Indeed, the attempted propargylation of unprotected thymidine (**9**) did not lead to the desired 5′-porpargylated thymidine **53**, but only

yielded *N*-3-propargylated thymidine **52** (Scheme V-2A).[n] Despite a contradicting literature report for this reaction,[298] chemoselective propargylation at the *N*-3 of the nucleobase is not surprising, because the added base (1,8-diazabicyclo[5.4.0]undec-7-ene, DBU; pKa ca. 12.5) is not strong enough to deprotonate the primary 5′-OH (pKa ca. 15). Instead, deprotonation of the nucleobase (pKa of N-H, 9.8) converts the *N*-3 position into a nucleophile for propargylation. Thus, an alternative synthesis




**Scheme V-2: Synthesis of alkyne 57.**
Reagents and conditions: (**A**) (i) DBU, propargyl bromide, DMF, 60 °C, 16 h, 86%. (**B**) (ii) imidazole, *tert*-butyldimethylsilyl chloride, DMF, 4 °C to rt, 3 h, 94%; (iii) imidazole, *tert*-butyldiphenylsilyl chloride, rt, 16 h, 84%; (iv) AcOH/$H_2O$, 3/1 (v/v), THF, rt, 72 h, 83%; (v) NaH, propargyl bromide, THF, 4 °C to rt, 30%. Abbreviations: DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; DMTr, 4,4′-dimethoxytrityl; TBDPS, *tert*-butyldiphenylsilyl; TBS, *tert*-butyldimethylsilyl.

[n] Synthesised by Ewa Węgrzyn (Part II), University of Oxford.

route using orthogonal protecting groups was chosen to control the propargylation at the 5′-OH (Scheme V-2B).

In a first step, thymidine (**9**) was regioselectively protected at the 5′-OH to afford **54**.[299] Regioselectivity was achieved by slow addition of one equivalent of the silylating reagent at low temperatures and can be explained by the lower steric hindrance of the primary 5′-OH compared to the secondary 3′-OH. Compound **54** was then silylated at the 3′-position to obtain fully protected **55**.[299] Treatment of **55** with diluted acetic acid selectively deprotected the *tert*-butyldi**methyl**silyl (TBS) group whilst leaving the *tert*-butyldi**phenyl**silyl (TBDPS) group in place to yield 3′-silylated **56**.[300] Propargylation of **56** was achieved with NaH as the base (pKa 35) to give 5′-propargylated **57**. Note that it is important to add at least two equivalents of the base because one equivalent is first consumed by deprotonation of the nucleobase. The correct propargylation site was determined by HMBC spectroscopy (Figure V-10). In contrast to the coupling between the C2/C6 and the propargyl H6′ in **52** (Figure V-10A),[o] the H6′ of the propargyl group couples with the C5′ of the sugar in compound **57** (Figure V-10B). This confirms the correct propargylation site.

[o] Recorded by Ewa Węgrzyn (Part II), University of Oxford.

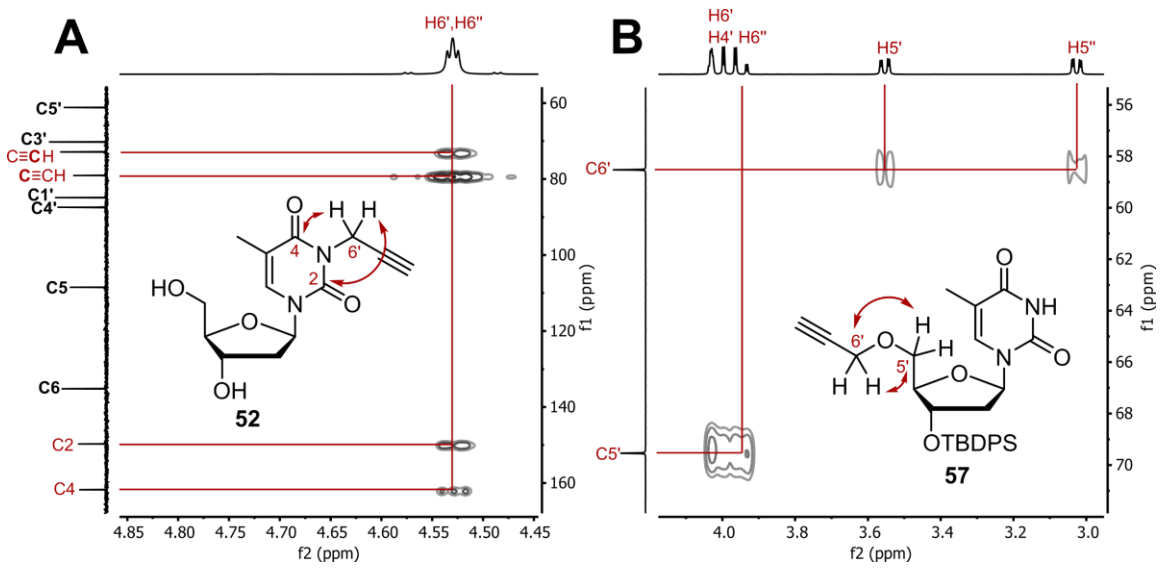
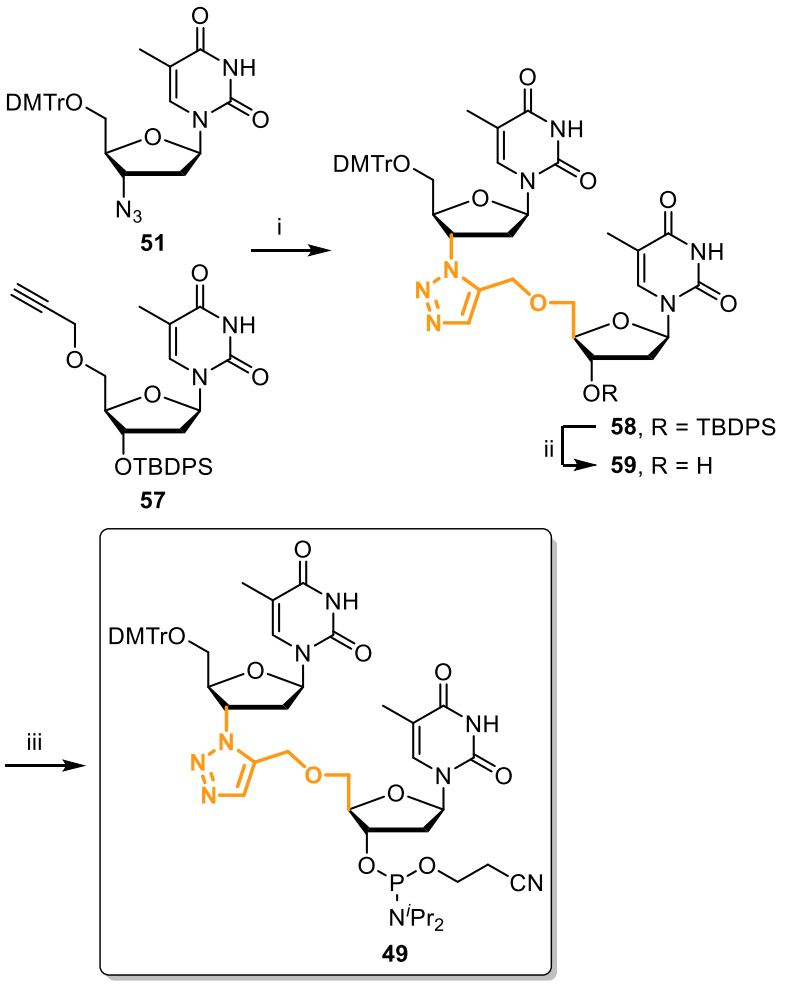



**Figure V-10: Determination of the propargylation site.**
(**A**) HMBC spectrum of *N*-3-propargylated **52**. The $^{1}H–^{13}C$ correlation between the C2/C4 and the propargyl H6′ is indicated as red lines. The corresponding correlations are also indicated with red arrows in the molecular structure. (**B**) HMBC spectrum of 5′-propargylated **57**. The $^{1}H–^{13}C$ correlations between the 5′-position of the sugar and the H6′-position of the propargyl group are indicated with red lines in the spectrum and with red arrows in the molecular structure. Abbreviations: TBDPS, *tert*-butyldiphenylsilyl. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

Following their successful preparation, the azide and alkyne ligation partners **51** and **57** were coupled by RuAAC to form dimer **58** (Scheme V-3). The desired 1,5-disubstitution was confirmed by HMBC spectroscopy (Figure A-5; Appendix). The 3′-position of dimer **57** was then deprotected to give alcohol **59**, which was phosphitylated to $T^{TL2}T$-dimer CE phosphoramidite **49** for solid phase oligonucleotide synthesis.




**Scheme V-3: Synthesis of T$^{TL2}$T-dimer CE phosphoramidite 49.**
Reagents and conditions: (i) Cp*Ru(PPh$_3$)$_2$Cl, 1,4-dioxane, 70 °C, 1 h, 65%; (ii) TBAF, THF, rt, 3.5 h, 82%; (iii) 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, $CH_2Cl_2$, rt, 2 h, 90%. Abbreviations: Cp*, pentamethylcyclopentadienyl; DIPEA, *N,N*-diisopropylethylamine; DMTr, 4,4′-dimethoxytrityl; TBAF, tetrabutylammonium fluoride; TBDPS, *tert*-butyldiphenylsilyl.

## 3.2. Biocompatibility of the Triazole Linkages

The following paragraphs describe and discuss the biocompatibility assessment of different triazole linkages. The biocompatibility criteria include the ability to form stable duplexes with unmodified target oligonucleotides without perturbing the overall duplex structure. Moreover, the ‘readability’ of the triazole linkage by DNA

polymerases will be assessed, which includes the evaluation of continuation, speed, and correctness of the read-through process.

### 3.2.1. Target Engagement of Oligonucleotides with Triazole Linkages

Different triazole-modified oligonucleotides are required for the evaluation of their duplex hybridisation properties and the effects conferred by each triazole linkage. Some of the triazole linkages from Figure V-4 have already been evaluated in short isosequential oligonucleotides and include TL**1**, TL**3**, TL**4**, and TL**6**.[72] Moreover, limited hybridisation properties of TL**5** have been reported, but in a different sequence context.[70] Thus, TL**2**- and TL**5**-modified[p] oligonucleotides were synthesised with the same sequence as the triazole-modified oligonucleotides in the reported study[72] investigating TL**1**, TL**3**, TL**4**, and TL**6**. This allows the accurate comparison of the duplex hybridisation properties among TLs **1**–**6**.

#### 3.2.1.1. Duplex Stability

The thermal stabilities of duplexes formed by triazole-modified oligonucleotides with an unmodified complementary DNA or RNA target were assessed through thermal melting experiments.[q] All sequences of the modified oligonucleotides are summarised in Table V-1. The melting temperatures presented for ON41, ON43, ON44, and ON46 were taken from a literature report for the shown sequences.[72] To complement the thermal melting data, duplex stabilities of ON42 or

[p] $T^{TL5}T$-dimer CE phosphoramidite and TL**5**-modified ONs were prepared by Aman Modi (Part II), University of Oxford.
[q] $T_m$'s were determined in collaboration with Aman Modi (Part II), University of Oxford.

**Table V-1: $T_m$'s of triazole modified oligonucleotides with an RNA or DNA target.**
Values were obtained from the maxima $dA_{260}/dT$ vs. $T$ for 3 µM of each ON in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0. The reported $T_m$'s are the averages of two independent experiments. $\Delta T_m$'s are relative to the unmodified control ON47. *Table and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

| | | | | RNA target[b] | DNA target[c] |
|---|---|---|---|---|---|
| ON | TL | length[a] | Sequence (5′→3′) | $T_m$ ($\Delta T_m$)/°C | $T_m$ ($\Delta T_m$)/°C |
| ON41 | TL**1** | 4 |  | (−13.6)[72] | (−11.2)[72] |
| ON42 | TL**2** | 5 |  | 55.5 (−3.4) | 57.0 (−5.1) |
| ON43 | TL**3** | 5 |  | (−3.3)[72] | (−2.9)[72] |
| ON44 | TL**4** | 5 |  | (−0.8)[72] | (−3.2)[72] |
| ON45 | TL**5** | 6 |  | 50.3 (−8.6) | 53.5 (−8.6) |
| ON46 | TL**6** | 6 |  | (−5.3)[72] | (−7.4)[72] |
| ON47 | PO | 4 |  | 58.9 | 62.1 |




[a] Defined as the minimum number of atoms connecting the sugar rings (see Figure V-4).
[b] Sequence of DNA target = 5′-d(GCT GCA AAC GTC G)-3′.
[c] Sequence of RNA target = 5′-r(GCU GCA AAC GUC G)-3′.

ON45 with their complementary RNA or DNA targets were measured using the same conditions as reported for the other oligonucleotides in Table V-1.[72] The obtained melting temperatures of modified ON42 and ON45, and their differences to unmodified ON47 ($\Delta T_m$'s) are summarised in Table V-1. In comparison to unmodified ON47, all triazole-modified oligonucleotides decreased the duplex stabilities with an RNA and a DNA target, albeit to different degrees. In general, the 5-atom TLs **2**–**4** are the least destabilising, ranging between −5.1 and −0.8 °C. It is noteworthy that a single incorporation of TL**4** was reported to destabilise the duplex with a complementary RNA target by only −0.8 °C. Although being closest to the phosphodiester linker length, 4-atom TL**1** induces the highest degree of

destabilisation for duplexes with RNA and DNA (−13.6 and −11.2 °C, respectively). The longer 6-atom TL**5** and TL**6** are more destabilising than the 6-bond TLs **2**–**4**, but not to the extent of TL**1**. In summary, the 5-atom triazole linkages tend to form the most stable duplexes—or the least destabilised duplexes—with DNA or RNA. However, no other obvious correlation can be drawn from the molecular structures of TLs **1**–**6** and their associated duplex stability properties.

### 3.2.1.2. Duplex Structure

The reported studies of TL**1**, TL**3**, TL**4**, and TL**6** also investigated the duplex structures of the corresponding oligonucleotides from Table V-1 with RNA and DNA.[72] As such, all reported triazole linkages do not significantly disturb the overall duplex structures compared to the unmodified control ON47 as shown by circular dichroism experiments.[72] In this context, duplexes of ON42 (TL**2**) and ON45 (TL**5**)

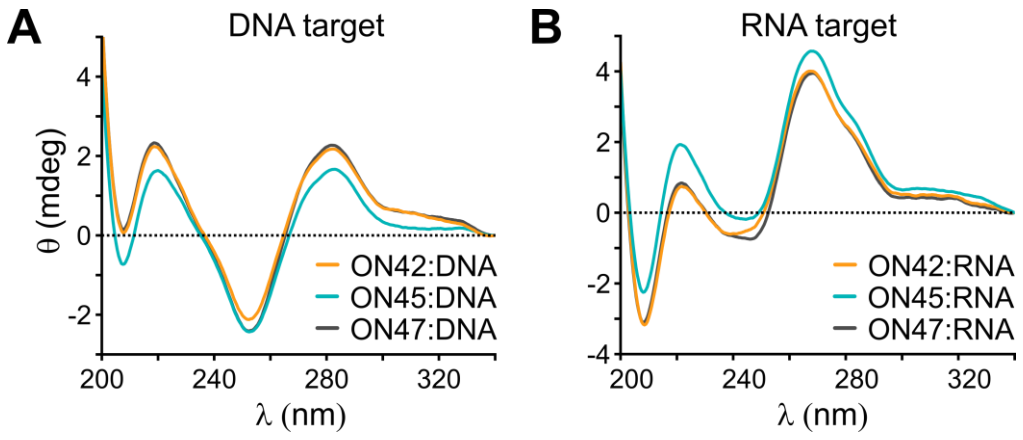



**Figure V-11: Circular dichroism spectra of ON42 (TL2), ON45 (TL5) and unmodified ON47 in duplex with their RNA or DNA targets.**
CD spectra of duplexes formed by ON42 (TL**2**), ON45 (TL**5**) and unmodified ON47 with their respective targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). Data points were taken as the average of four scans. (**A**) RNA target: 5′-r(GCU GCA AAC GUC G)-3′; (**B**) DNA target: 5′-d(GCT GCA AAC GTC G)-3′. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

with RNA and DNA were examined by CD (Figure V-11).[r] The duplexes of ON42 (TL**2**) with DNA and RNA show almost identical spectra to the equivalent duplexes formed by unmodified ON47. Moreover, duplexes of ON45 (TL**5**) with DNA or RNA retain global B- or A-form helical structures, respectively. In detail, both TL**2**- and TL**5**-modified ON42 and ON45 retain the negative and positive peaks at around 245 nm and 275 nm with a DNA target, indicating a B-type DNA helical structure (Figure V-11A). Likewise, the A-RNA indicative minimum at around 210 nm and maximum at around 260 nm are retained for the duplexes with RNA (Figure V-11B).

### 3.2.2. Replication Competency of Oligonucleotides with Triazole Linkages

The replication competencies of TLs **1**–**7** were assessed by polymerase read-through experiments (Figure V-12). For this assay, templates with a single triazole linkage (TLs **1**–**7**) were synthesised using either $T^{TL\mathbf{2}}$T-dimer CE phosphoramidite **49**, or analogous T–T-dimer CE phosphoramidites containing one of the other triazole linkages.[s] The sequences of the templates can be found in Table A-3; Appendix. The assessment of the read-through competencies comprised the ability of polymerases to continue replication through the triazole linkages. Moreover, the speed and correctness with which a triazole linkage can be 'read' by polymerases add to the definition of its polymerase competency. As such, 'efficient

[r] CD experiments were performed in collaboration with Aman Modi (Part II), University of Oxford.
[s] $T^{TL\mathbf{1}}$T-dimer CE phosphoramidite, $T^{TL\mathbf{6}}$T-dimer CE phosphoramidite, $T^{TL\mathbf{7}}$T-dimer CE phosphoramidite, and corresponding templates synthesised by Agnes E. S. Tyburn (DPhil), University of Oxford. $T^{TL\mathbf{3}}$T-dimer CE phosphoramidite and template synthesised by Diallo Traoré (visiting student), University of Oxford. $T^{TL\mathbf{4}}$T-dimer CE phosphoramidite synthesised by Przemyslaw Wanat (visiting student), University of Oxford. $T^{TL\mathbf{5}}$T-dimer CE phosphoramidite and template synthesised by Aman Modi (Part II), University of Oxford.

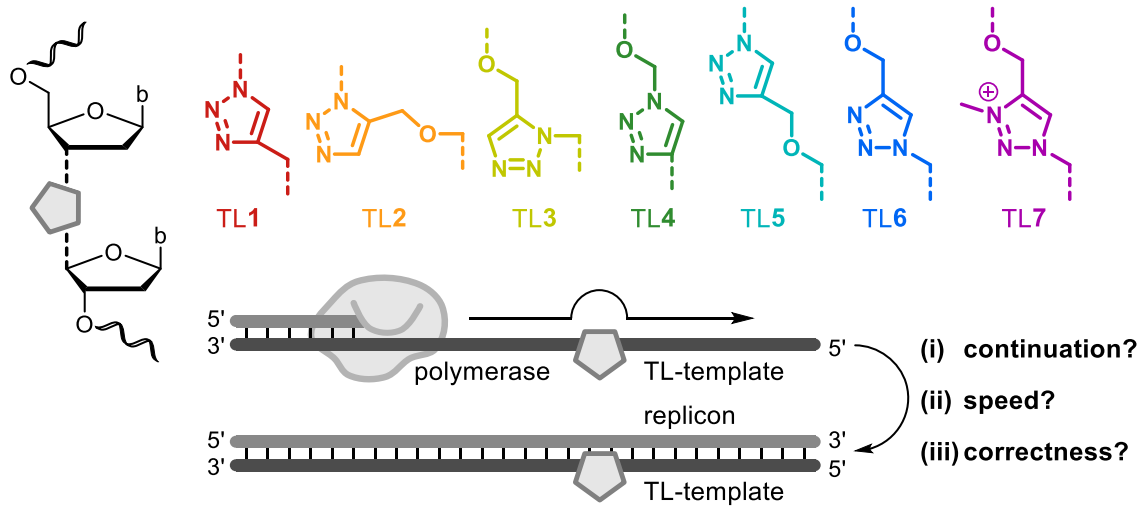



**Figure V-12: The categories to assess the read-through compatibilities of TLs 1–7.**
The different triazole structures tested in this chapter are summarised at the top for reference. The read-through compatibility assesses the extension of a primer (light grey) through a TL (grey pentagon) along the modified template (dark grey) by a polymerase. The three categories that constitute the read-through compatibility are (i) Continuation: does the polymerase generally continue primer extension through the TL? (ii) Speed: does the TL impose a kinetic barrier on the read-through process? (iii) Correctness: does the TL induce mutations in the replicon?

read-through' is defined as the rapid formation of a replicon with the correct sequence by a polymerase using the triazole-modified oligonucleotide as a genetic template.

### 3.2.2.1. Primer Extension Assay

*Klenow Fragment.* The question whether a polymerase can continue replication when encountering an artificial triazole backbone was addressed by primer extension assays through TLs **1**–**7** (Figure V-13A). Unmodified control Template 0 (PO) and triazole-modified Templates 1–7 (TLs **1**–**7**) were incubated with a fluorescently labelled primer and the large fragment of DNA polymerase I (Klenow fragment) at 37 °C. Extension of the fluorescent primer along the template will result in fluorescently labelled replicons that can be analysed by gel (Figure V-

13B). As expected, elongation of the fluorescent primer along unmodified Template 0 results in the formation of a single fluorescently labelled product (FL-prod. 0**; Figure V-13B). In parallel, a 4-fold scale with the same reaction conditions was analysed by mass spectrometry confirming the formation of the expected fluorescently labelled full-length product (FL-prod. 0**; Figure V-13C).

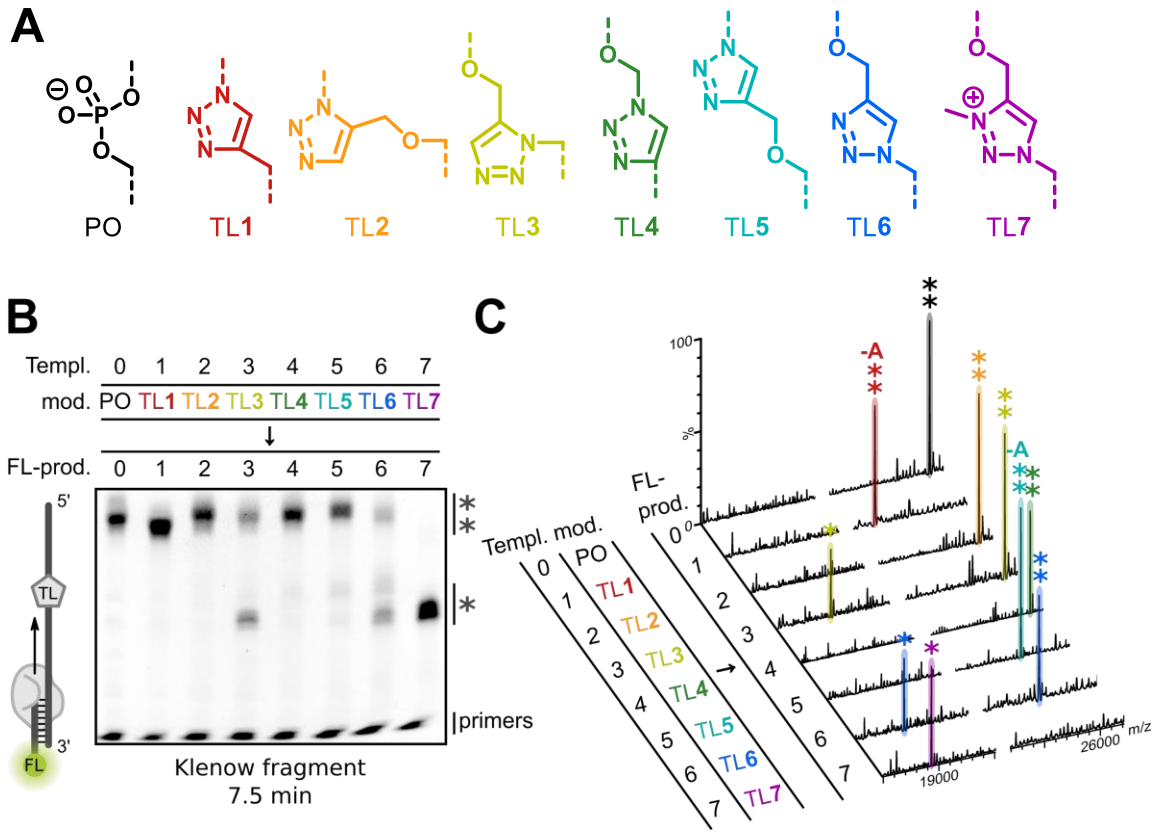



**Figure V-13: Primer extension assay with Klenow fragment at 37 °C for 7.5 min.**
(**A**) Summary of the internucleoside linkage structures. (**B**) PAGE analysis of fluorescent products (FL-prod.) after extension of fluorescently labelled primers along templates (Templ.) with the specified modifications (mod.) by Klenow fragment at 37 °C for 7.5 min. The FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Templates and FL-primers were: Template 0 (PO), FL-Primer 1; Template 1 (TL**1**), FL-Primer 2; Template 2 (TL**2**), FL-Primer 1; Template 3 (TL**3**), FL-Primer 1; Template 4 (TL**4**), FL-Primer 1; Template 5 (TL**5**), FL-Primer 1; Template 6 (TL**6**), FL-Primer 3; or Template 7 (TL**7**), FL-Primer 1. All TLs were surrounded by the sequence motif: d(CT$^{TL}$TA). The sequences of the templates and FL-primers can be found in Table A-3; Appendix. (**C**) MS analysis of 4-fold scale reactions using the same conditions as the corresponding reactions described in (**B**). Identified read-through and truncated products are labelled with (**) or (*), respectively. –A denotes a missing dA. The full spectra and the calculated and found masses are shown in Figure A-7–A-9 and Table A-4; Appendix. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

Replication of the TL**1** template afforded a product that roughly migrated with a similar rate as the full-length product from the PO template (FL-prod. 1**; Figure V-13B). However, mass analysis revealed an A-deletion for this replicon (FL-prod. 1**$^{-A}$; Figure V-13C). Indeed, a reported sequencing study showed that replication of the TL**1** template by Klenow fragment leads to a single nucleotide deletion opposite from the triazole in 97% of the replicons.[48] The TL**2**-modified template retains the ability to be read by Klenow fragment to produce a single product of the expected mass (FL-prod. 2**; Figure V-13B, C). No additional products were detected on the gel or by mass analysis. Although the TL**3** template produces the expected full-length product (FL-prod. 3**; Figure V-13B, C), an additional shorter fragment was observed on the gel. This fragment was identified by mass analysis to be the result of the incomplete replication that stopped directly before the triazole modification (FL-prod. 3*; Figure V-13B, C). Like TL**2**, the TL**4**-template resulted in the exclusive formation of the expected full-length replicon (FL-prod. 4**; Figure V-13B, C). As observed for TL**1**, the presence of TL**5** induces an A-deletion in the replicon of the corresponding template (FL-prod. 5**$^{(-A)}$; Figure V-13B, C). Moreover, faint bands of truncated products were visible on the gel, which were, however, not resolved by mass spectrometry. Replication of the TL**6** template led to the formation of multiple products as observed by gel electrophoresis. Of those, the expected full-length product and a truncated product for which replication stopped directly before the triazole were identified by mass analysis (FL-prod 6* and FL-prod. 6**; Figure V-13B, C). The methylated analogue—TL**7**—induced the most drastic effect: replication did not

proceed across the TL**7** modification and a full-length product was absent. Instead, an intense band for a truncated product was visible on the gel, and mass analysis identified a major product for which replication stopped directly before the triazole (FL-prod. 7*; Figure V-13B, C).

*Hot-Start Taq Polymerase.* The primer extension experiment from the previous paragraph was conducted with Klenow fragment at 37 °C. However, possible applications of the triazole linkages in PCR experiments require them to be biocompatible with thermostable polymerases such as hot-start Taq polymerase. Thus, replication competencies of TLs **1**–**7** were further tested using hot-start Taq polymerase at 60 °C for 160 min (Figure V-14A). The long incubation time was based on previous primer extension assays using Taq polymerases reporting a significant slow-down of replication because of the high template concentrations.[48,301] It is well-known that replication by hot-start Taq polymerase adds an additional A-nucleotide to the replicon,[302] and indeed, the unmodified control Template 0 (PO) afforded a single replicon including the expected A-adduct (FL-prod. 0**; Figure V-14B, C). The presence of TL**1** in Template 1 resulted in the production of several replicons: the product resulting from continuation of replication through TL**1** was only formed in low amounts and included an A-deletion mutation (FL-prod. 1**$^{(-A)}$; Figure V-14B, C). This is not surprising as a previous study showed that replication through TL**1** by a Taq polymerase induces a single nucleotide deletion around the corresponding position of the template's triazole in 87% of the replicons.[48] The main products of primer extension along the TL**1** template were truncated products resulting from termination before and after

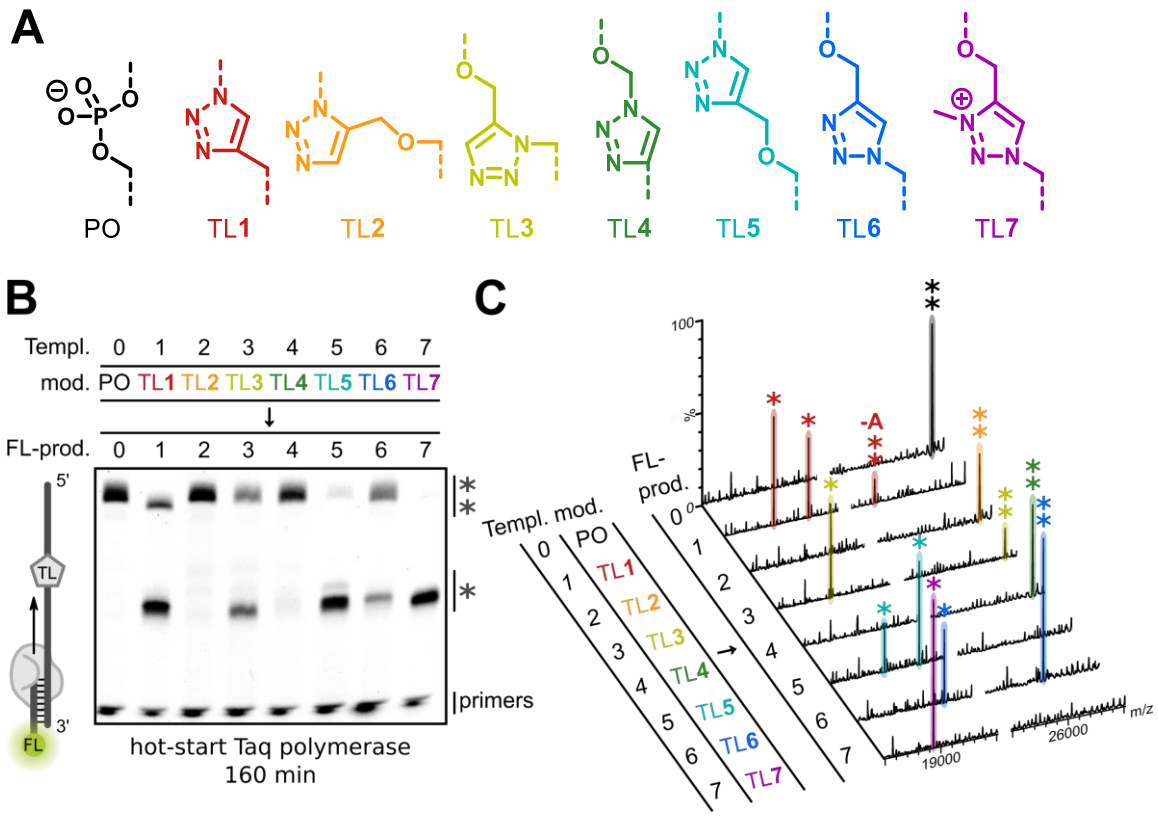



**Figure V-14: Primer extension assay with hot-start Taq polymerase at 60 °C for 160 min.** (**A**) Summary of the internucleoside linkage structures. (**B**) PAGE analysis of fluorescent products (FL-prod.) after extension of fluorescently labelled primers along templates (Templ.) with the specified modifications (mod.) by hot-start Taq polymerase at 60 °C for 160 min. The FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Templates and FL-primers were: Template 0 (PO), FL-Primer 1; Template 1 (TL**1**), FL-Primer 2; Template 2 (TL**2**), FL-Primer 1; Template 3 (TL**3**), FL-Primer 1; Template 4 (TL**4**), FL-Primer 1; Template 5 (TL**5**), FL-Primer 1; Template 6 (TL**6**), FL-Primer 3; or Template 7 (TL**7**), FL-Primer 1. All TLs were surrounded by the sequence motif: d(CT$^{TL}$TA). The sequences of the templates and FL-primers can be found in Table A-3; Appendix. (**C**) MS analysis of 4-fold scale reactions using the same conditions as the corresponding reactions described in (**B**). Identified read-through and truncated products are labelled with (**) or (*), respectively. –A denotes a missing dA. The full spectra and the calculated and found masses are shown in Figure A-10–A-12 and Table A-5; Appendix. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

the triazole linkage (FL-prod. 1*; Figure V-14B, C). Remarkably, the TL**2** template was replicated without any truncated products detectable by gel or MS and exclusively produced the expected full-length product including the hot-start Taq typical A-adduct (FL-prod. 2**; Figure V-14B, C). The TL**3** modification was partially read by the polymerase to produce the expected full-length product, but

also stopped replication directly before the triazole (FL-prod. 3* and FL-prod. 3**; Figure V-14B, C). Replication through TL**4** was again comparable to TL**2** and exclusively afforded the full-length product (FL-prod. 4**; Figure V-14B, C). When replicated by hot-start Taq polymerase, TL**5** mostly stopped replication resulting in truncated products. Only a very faint band with a similar migration rate to the full-length products was observed on the gel but could not be detected by mass spectrometry. Moreover, mass analysis showed that replication stopped directly before and after TL**5** (FL-prod. 5*; Figure V-14B, C). Read-through of TL**6** produced the correct full-length replicon but also a truncated fragment (FL-prod. 6* and FL-prod. 6**; Figure V-14B, C). This confirms previous results showing that continuation of replication through TL**6** by Taq polymerase generates the correct replicon sequence with high fidelity.[48] Like in the previous primer extension experiment with Klenow fragment, TL**7** blocks replication through the artificial backbone directly before the triazole, confirming reported results for this linkage (FL-prod. 7*; Figure V-14B, C).[48]

*Alternative Template Sequence*. The excellent read-through compatibility of TL**2** was further confirmed with an alternative TL**2** template of a different sequence. Whilst incorporation of $T^{TL\mathbf{2}}T$-dimer CE phosphoramidite **49** inevitably leads to a $T^{TL\mathbf{2}}T$ sequence around the artificial linkage, the bases around the dimer were changed in the alternative template—from d($CT^{TL\mathbf{2}}TA$) to d($GT^{TL\mathbf{2}}TT$). Primer extensions by Klenow fragment and hot-start Taq polymerase were performed under the same conditions as described in the previous sections. Figure V-15 shows that both primer extension experiments of the TL**2**-modified

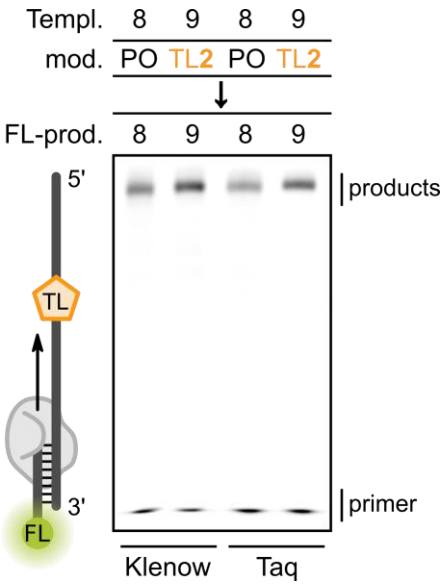



**Figure V-15: Primer extension assay with an alternative sequence**
PAGE analysis of fluorescent products (FL-prod.) after extension of fluorescently labelled primers along templates (Templ.) with the specified modifications (mod.) by Klenow fragment at 37 °C for 7.5 min, or by hot-start Taq polymerase at 60 °C for 160 min. The FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm). Templates and FL-primers were: Template 8 (PO), FL-Primer 4; Template 9 (TL**2**), FL-Primer 4. TL**2** was surrounded by the sequence: d(GT$^{TL}$**$^{2}$**TT). The sequences of the templates and FL-primers can be found in Table A-3; Appendix. MS analysis of 4-fold scale reactions using the same conditions and the calculated and found masses are shown in Figure A-13 and Table A-6; Appendix. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

alternative templates resulted in the exclusive formation of a single replicon that migrated with the same rate as the replicon from the isosequential unmodified control Template 8 (PO). The identities of the FL-products 9 were confirmed by mass spectrometry and showed the same masses as their corresponding FL-products 8 (Figure A-13 and Table A-6; Appendix).

### 3.2.2.2. Quantitative PCR

Apart from the ability of polymerases to continue replication through a triazole linkage, the speed at which the polymerase overcomes the artificial backbone constitutes an important element to characterise its read-through compatibility.

Thus, the read through kinetics of TLs **1**–**7** by hot-start Taq polymerase were evaluated in qPCR experiments. In general, PCR passes through consecutive amplification cycles and each cycle consists of three steps: (**i**) annealing of primers—typically at around 55 °C; (**ii**) extension of the primers—at 68 °C for hot-start Taq polymerase;[303] and (**iii**) denaturation of the replicon from the template—typically at 95 °C. The annealing and extension steps in the performed qPCR were combined to a single annealing/extension step at 60 °C (Figure V-16A). The duration of this step is referred to as the 'extension time' and will be relevant for evaluating the replication kinetics of the triazole-modified templates. During a qPCR experiment, the original template is being replicated exponentially, which can be detected as an increasing fluorescent signal from the added intercalating dye (Figure V-16B). A commonly used descriptor for qPCRs is the threshold cycle ($C_t$)—the cycle at which the fluorescent signal overcomes a set threshold (t). This $C_t$ value will be used to describe and compare the reaction kinetics of qPCRs with different triazole backbones. By the nature of PCR amplification with unmodified dNTPs, a modified template will be replicated into an unmodified replicon, which itself will serve as an unmodified template in the next cycle. Thus, the impact of the template modification on the replication rate is particularly relevant during the first cycles of qPCR, when the modified template species still constitutes a significant percentage. If the modification impairs the replication kinetics, the onset of the qPCR will be delayed, and the threshold cycle will be increased. The qPCR kinetics experiment is based on the following hypothesis: impediment of the replication rate by an unnatural backbone can be overcome by prolonging the extension time,

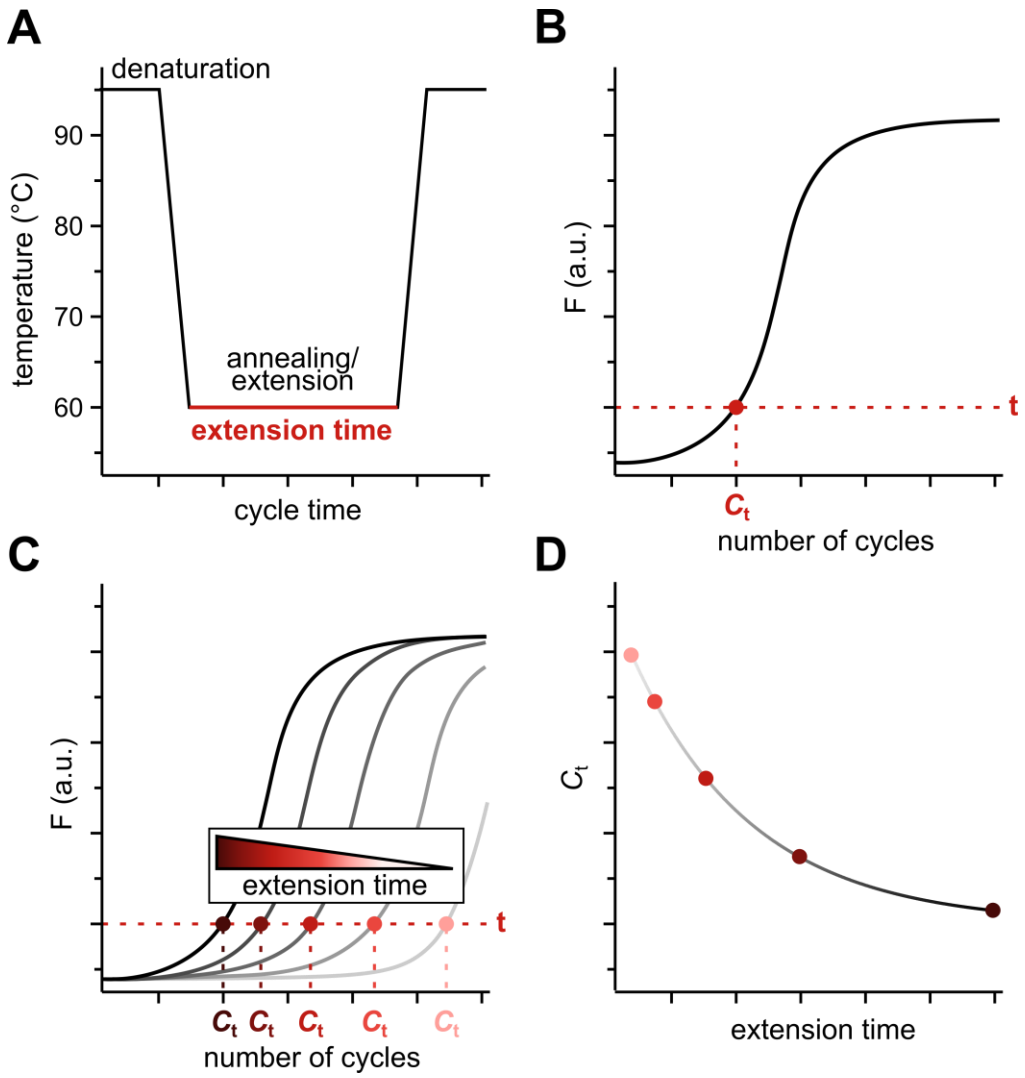



**Figure V-16: The concept of evaluating the impediment of an artificial backbone on the speed of replication.**

(**A**) Representation of one qPCR cycle. The denaturation temperature is set to 95 °C and the annealing/extension temperature is set to 60 °C. The extension time is defined as the period during which the reaction mixture is kept at 60 °C for each cycle (red line). (**B**) Schematic representation of the recorded fluorescence (F) as the qPCR passes through consecutive cycles. The set threshold (t) and the corresponding $C_t$ are indicated in red. (**C**) Schematic representation of the hypothesis that prolonged extension times will lead to decreased $C_t$'s for impeding template modifications. The duration of the extension time is indicated with a coloured wedge. The gradient from light to dark red represents short to long extension times. The grey curves represent different fluorescence readouts of qPCRs with short (light grey) and long (dark grey) extension times. The $C_t$'s are indicated for each qPCR experiment with different extension times using the light to dark red colour scheme. (**D**) Schematic $C_t$–extension time correlation for qPCR amplifications of a template with an impeding backbone modification. The data points use the same colour code as the $C_t$'s in (**C**). The fitted line shows the extension time dependency of the $C_t$'s and uses a gradient from light grey (short extension time) to dark grey (long extension time).

thereby granting more time to the polymerase to read through the artificial linkage.

A long extension time will render the impediment of the artificial backbone less impactful during the early cycles of qPCR and the $C_t$ value will decrease (Figure V-16C, D). In other words: the threshold cycle for a template with an extension time-dependent modification will decrease with a prolonged extension time when within the range of time dependency. Based on this hypothesis, the triazole-modified templates (TLs **1**–**7**; Figure V-17A) were assessed in qPCR experiments with varying extension times from 15–240 s (Figure V-17B). Isosequential unmodified controls (Templates 0, and 10–12) were used as a reference. The replication rates of the unmodified templates are considered to be extension time-independent because the phosphodiester bonds are quickly read by hot-start Taq polymerase—the vendor's recommended extension time is 1 min per kb[303] and the templates used in this experiment are ~80-mers (see Table A-3; Appendix). Indeed, the obtained threshold cycles of the unmodified control templates did not show any dependency on the extension time (Templates 0, and 10–12; Figure V-17B). Moreover, the unmodified templates generally resulted in lower $C_t$'s than the triazole-modified templates (Templates 1–7). Whilst the obtained $C_t$'s appeared to be time-independent, the overall shapes of the qPCR curves showed some time-dependency and resulted in an almost linear amplification for short extension times of <120 s (see Figure A-14–16; Appendix). This observation disagrees with the above-mentioned assumption that the amplification of unmodified templates is generally extension time-independent, which should result in sigmoidal curves for all extension times tested. Instead, it appears that the qPCR becomes less efficient at shorter extension times which might result from altered primer efficiencies caused

by the changed qPCR parameters. Another cause could be inefficient amplification by the hot-start Taq polymerase. An extension time of 1 min per kb is recommended for hot-start Taq polymerase[303] and the used extension times of 15–240 s lie within these calculations for an ~80-mer template. However, the 1 min per kb rule is most probably not scalable to such short templates, and hot-start Taq might be too slow for highly efficient qPCR amplification at very short extension times. Other polymerases are much faster, e.g., Phusion polymerase with a recommended extension time of 15–30 s per kb.[304] Previous experiments have in fact shown that similar apparent linear qPCR amplifications with hot-start Taq polymerase at short extension times give the expected sigmoidal curves when Phusion polymerase was used instead.[48] It is important to consider that the short extension times chosen for the presented qPCR experiments might be close to or beyond the limits for hot-start Taq polymerase, thereby causing a potential problem in this system, and care should be taken when interpreting the following qPCR results for the TL-modified templates. Thus, all TL-modified templates were compared to their respective unmodified control templates having the same sequences and same primers.

The differences in the threshold cycles of the triazole-modified templates to their respective isosequential controls ($\Delta C_t$) are given in Figure V-17C. Almost all tested modified templates show extension time dependencies but to different extents. TL**5** and TL**6** showed the highest extension time dependency (Templates 5 and 6; Figure V-17C): the $\Delta C_t$'s start high (ca. 8–9 at 15 s) and drop rapidly with longer extension times. Indeed, the $\Delta C_t$'s of the TL**5**- and TL**6**-modified templates keep decreasing until an extension time of 240 s, which is the highest in this

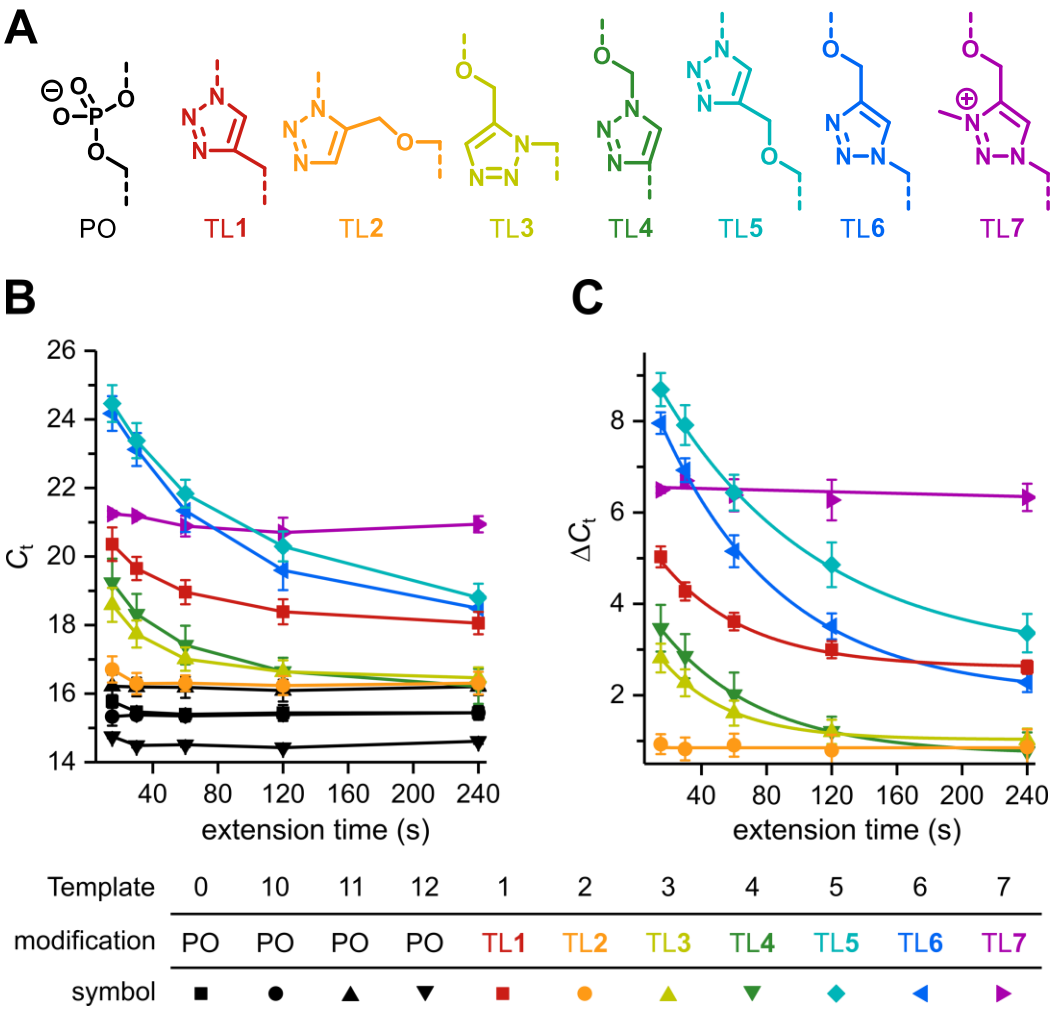



**Figure V-17: Results from qPCR experiments with different extension times.**
(**A**) Summary of the triazole internucleoside linkage structures. (**B**) $C_t$'s from qPCR experiments with hot-start Taq polymerase using different extension times from 15–240 s. $C_t$'s were determined from qPCRs with a threshold fluorescence set to F = 300 (a.u.). Data is presented as the average from triplicates ± standard error of the mean. Templates and primers were: Template 0 (PO), FP 1, RP1; Template 1 (TL**1**), FP 2, RP 2; Template 2 (TL**2**), FP 1, RP 1; Template 3 (TL**3**), FP 1, RP 1; Template 4 (TL**4**), FP 1, RP 1; Template 5 (TL**5**), FP 1, RP 1; Template 6 (TL**6**), FP 3, RP 3; Template 7 (TL**7**), FP 1, RP 1; Template 10 (PO), FP 2, RP 2; Template 11 (PO), FP 3, RP 3; Template 12 (PO), FP 1, RP 1. The sequences and representative qPCRs can be found in Table A-3 and Figure A-14–A-19; Appendix. (**C**) $\Delta C_t$'s for the TL-modified templates (TLs **1**–**7**) to their respective isosequential unmodified control (POs) calculated from (**B**). $\Delta C_t = C_t$(TL-modified) − $C_t$(unmodified) for each extension time. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

experimental setting. The observed extension time dependency of TL**6** is not surprising and confirms a previously reported TL**6**-induced slowdown in PCR experiments.[48,154] The $\Delta C_t$'s of the TL**1** template are extension time-dependent

between 15–120 s and the fitted curve flattens towards an extension time of 240 s (Template 1; Figure V-17C). The extension time dependency profiles of the TL**3** and TL**4** templates did not significantly differ from each other (Templates 3 and 4; Figure V-17C). Both are extension time-dependent from 15–120 s and become extension time-independent towards 240 s. Moreover, their $\Delta C_t$'s were generally lower than the ones obtained for the templates modified with TL**1** and TLs **5**–**7**. The only triazoles that did not show any extension time dependency were TL**2** and TL**7**. However, the $\Delta C_t$'s of the TL**7**-modified template were significantly higher than those of the TL**2**-modified template (~6–7 vs. ~1). The apparent extension time independency of TL**7** has been previously explained on the basis of the steps in the qPCR cycle (see Figure V-16).[48] It is reasonable that TL**7** imposes a kinetic barrier that is too high to be overcome by prolonging the extension time—thus, extension time-independent. Instead, the TL**7** barrier can only be overcome with an increase in temperature that happens every cycle when transitioning from annealing/extension (60 °C) to denaturation (95 °C). Moreover, previous studies[48] showed that TL**7** is not correctly amplified, which will be discussed later in this chapter. In contrast, the low $\Delta C_t$'s of the TL**2** template resemble a genuine extension time independency of replication through this triazole. This is remarkable and means that TL**2** is the only triazole of the tested set that does not impose a kinetic barrier on the replication rates by hot-start Taq polymerase.

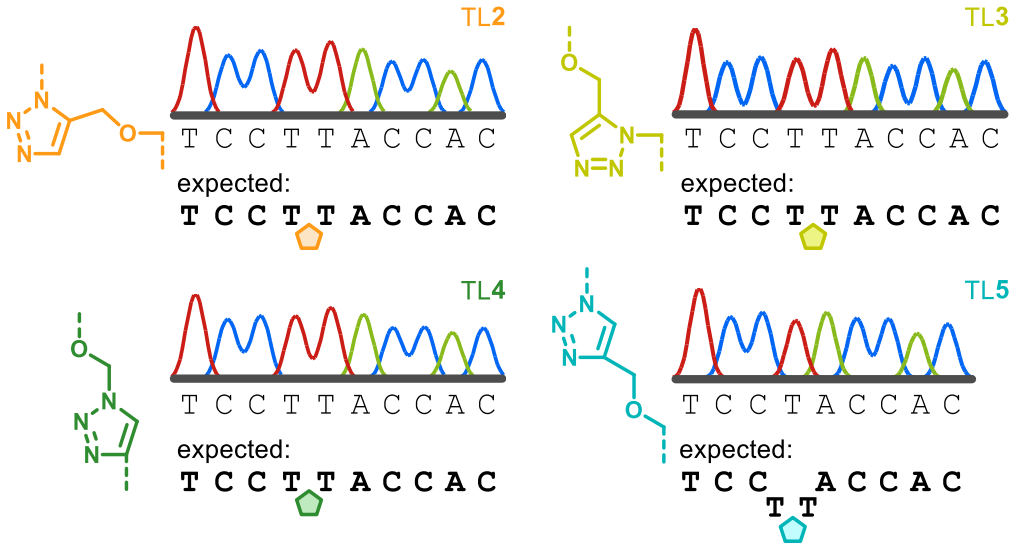



**Figure V-18: Sanger sequencing results of PCR amplicons.**
The relevant sections of representative Sanger sequencing chromatograms are shown for PCR amplicons of the TL-modified Templates 2–5 (TLs **2**–**5**). The corresponding sequences are indicated below the chromatograms. The expected sequences of the original TL-modified templates are shown in bold letters and the position of the TL is indicated with a pentagon. The full Sanger sequencing results can be found in Figure A-20–A-21; Appendix.

### 3.2.2.3. Sequencing

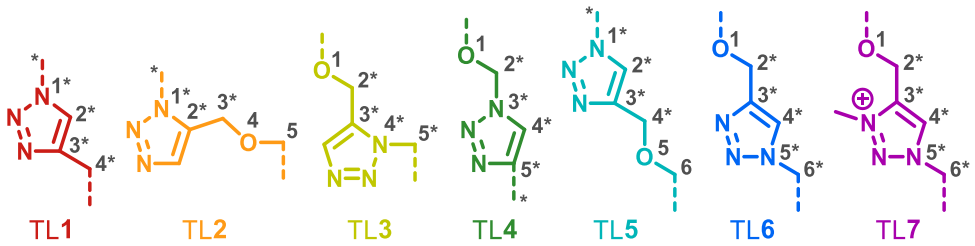
The final category to assess the replication competency of TLs **1**–**7** is the 'correctness' of the read-through event by polymerases. As such, a triazole-modified template is being read correctly when producing a replicon with the expected complementary sequence. The read-through fidelities of TL**1**, TL**6** and TL**7** by Taq polymerase have been previously reported.[48] Thus, replicons from templates with TLs **2**–**5** were sequenced to complement the read-through fidelity data for the set of triazole linkages. For this experiment, each triazole-modified template (Templates 2–5; TLs **2**–**5**) was amplified by hot-start Taq polymerase in PCR experiments using tailed primers. The exact sequences of the tailed primers can be

found in the Appendix (LIC-FP, LIC-RP; Table A-3). Next, the amplicons were annealed to a vector with an ampicillin resistance cassette, transformed into *E. coli* where ligation happened, and selected from agar plates supplemented with ampicillin. Four to five colonies were picked for each modified template, grown in media, and the vectors extracted for Sanger sequencing (Figure V-18). In all colonies picked (n = 4), read-through of TLs **2**–**4** by hot-start Taq polymerase resulted in the expected replicon sequence around the triazoles. In contrast, TL**5** is mostly mutagenic and induced a single deletion mutation around the triazole in four out of five sequenced colonies.

### 3.2.3. Structure–Biocompatibility Relationship

The duplex hybridisation properties and read-through competencies provide insight into the overall biocompatibilities of TLs **1**–**7**. Combining these findings with the molecular structures of the individual triazoles allows the discussion of a structure–biocompatibility relationship. The molecular structures, duplex hybridisation properties and read-through competencies of TLs **1**–**7** are summarised in Table V-2. The molecular characteristics can be divided into linker length and number of atoms in one plane along the internucleoside linkage. The set of tested triazoles linkages includes 4-atom (TL**1**), 5-atom (TLs **2**–**4**), and 6-atom (TLs **5**–**7**). Moreover, there are two sets of triazole regioisomers originating from either CuAAC or RuAAC, resulting in different numbers of rigid atoms in one plane along the internucleoside linkage: 1,4-disubstituted triazoles with five atoms (TL**1** and TLs **4**–**7**), and 1,5-disubstituted triazoles with only four atoms in one plane along

**Table V-2: Summary of biocompatibilities.**
The duplex stabilities are summarised from the melting temperature experiments with complementary DNA and RNA targets. The read-through competencies are summarised from the primer extension assays with Klenow fragment and hot-start Taq polymerase, the qPCR kinetics, and the sequencing experiments. The data is complemented with literature-known characteristics for some of the TLs as indicated.




| TL | linker length[a] | atoms in plane[b] | duplex stability[c] RNA target | duplex stability[c] DNA target | read-through (Klenow) | read-through (Taq) | kinetic barrier (Taq) | correctness[d] (Taq) |
|---|---|---|---|---|---|---|---|---|
| TL**1** | 4 | 5 | −13.6[72] | −11.2[72] | yes, mutagenic | partially | medium/high | no, single deletion[48] |
| TL**2** | **5** | **4** | **−3.4** | **−5.1** | **yes** | **yes** | **not significant** | **yes** |
| TL**3** | 5 | 4 | −3.3[72] | −2.9[72] | partially | partially | low | yes |
| TL**4** | 5 | 5 | −0.8[72] | −3.2[72] | yes | yes | low | yes |
| TL**5** | 6 | 5 | −8.6 | −8.6 | yes, mutagenic | partially | high | no, single deletion |
| TL**6** | 6 | 5 | −5.3[72] | −7.4[72] | partially | partially | high | yes[48] |
| TL**7** | 6 | 5 | n/a | n/a | no | no | very high | no, multiple deletions[48] |

[a] Defined as the minimum number of atoms connecting the sugar rings (grey numbers).
[b] Along the internucleoside linkage (starred numbers).
[c] $\Delta T_m$'s to unmodified ON47.
[d] Defined as having the correct sequence after read-through has occurred.

the linkages (TL**2** and TL**3**). The bonds of the triazole rings cannot freely rotate which results in a rigid character of the $sp^2$-hybridised triazole and its directly attached atoms, forming one plane. Thus, the combination of linker length and number of atoms in one plane along the internucleoside linkage has an influence on the overall flexibility of the triazole backbone. For example, TL**1** has a linker length of four atoms which are all in one plane. In addition, the C3′-atom of the 5′-sugar is also part of this plane. Consequently, the overall rigidity of the linker can be

considered as very high. In contrast, TL**2** has only four atoms in one plane along the linkage, but its linker length is five atoms, which make it generally more flexible. Thus, the 1,5-dusubstituted 1,2,3-triazole motif from RuAAC is an effective strategy to increase the overall flexibility compared to the 1,4-disubstituted 1,2,3-triazoles from CuAAC.

The 5-atom TLs **2**–**4** showed the most favourable duplex hybridisation properties with a DNA and RNA target. TL**1** has the same number of atoms directly connecting the sugar rings as the natural phosphodiester bond, and yet induces the highest degree of destabilisation with DNA and RNA. This might be explained by the very high rigidity of TL**1** in contrast to the flexible phosphodiester. The 6-atom TL**5** and TL**6** are more destabilising than the 5-atom TLs **2**–**4**. This suggests that the increased linker length is unfavourable for duplex hybridisation.

In general, the read-through compatibilities are most favourable for the 5-atom TLs **2**–**4**. Indeed, all 5-atom triazole linkages are, at least partially, read by Klenow fragment and hot-start Taq polymerase, and result in replicons with the correct sequences. Moreover, the kinetic barriers in the qPCR experiments were lowest for TLs **2**–**4**. The shorter 4-atom TL**1** and the longer 6-atom TL**5** are both read by Klenow fragment and hot-start Taq polymerase but result in mutations and the kinetic barriers to overcome TL**1** and TL**5** in qPCR experiments were elevated. TL**6** is an exception amongst the 6-atom linkers as it can be read by Klenow fragment and Tag polymerases with high fidelity.[48] However, TL**6** imposes a high kinetic barrier leading to a slowdown in PCR experiments.[48,154] The replication

incompatibility of the cationic, methylated TL**7** provides insights into another molecular requirement for read-through competency. Next to introducing a cationic group, the methyl substituent in TL**7** removes the capability of the nitrogen to form a hydrogen bond. It is known that polymerases interact with the template strand through multiple hydrogen bonds[305] and the deprivation of the nitrogen lone pair as a potential hydrogen bond acceptor in TL**7** might be detrimental to its read-through competency. However, introduction of the cationic character is likely to have an even larger effect on the read-through compatibility of this triazole. The hydrogen bond acceptor model is further supported by the different nitrogen lone pair accessibilities between TL**2** and TL**3**—those might explain why TL**2** has a better overall read-through competency despite similar linker length and flexibility. Indeed, TL**2** stands out as the artificial backbone with the highest read-through compatibility; it is the only triazole that is fully read by Klenow fragment and hot-start Taq polymerase, imposing no kinetic barrier, and leading to the correct replicon sequence.

# 4. Conclusion

Triazoles are powerful structural elements that can mimic the natural phosphodiester bond in oligonucleotides to generate remarkable biomimetic biomolecules. Most triazoles are formed by CuAAC to generate 1,4-disubstituted 1,2,3-triazoles,[48,70,149,153,294] but recent advances in adopting RuAAC facilitate the formation of 1,5-disubstituted 1,2,3-triazole linkages between nucleosides.[72] Altogether, a structurally diverse set of different triazole linkages has been reported and some of them have found versatile applications in synthetic biology and biotechnology.[86,145,149,150,152–155] However, each reported triazole linkage comes with its own limitations: e.g., reduced duplex stability of triazole-modified oligonucleotides,[72] or impaired read-through efficiency of triazole-modified templates by polymerases.[149,154] This chapter reports the synthesis of a novel 1,5-disubstituted 1,2,3-triazole (TL**2**) as part of a dimer CE phosphoramidite for incorporation into oligonucleotides. Evaluation of the biocompatibilities of seven structurally diverse TLs provides insight into a structure–biocompatibility relationship. Comparison with published duplex stability data[72] shows that 5-atom triazole linkages generally formed more stable duplexes with DNA and RNA compared to their shorter or longer isoforms. Moreover, the read-through compatibility of 5-atom TL**2** was superior to all other tested triazole backbones. Indeed, TL**2**-modified templates were efficiently and correctly copied by hot-start Taq polymerase without imposing a kinetic barrier on the replication.

# 5. Future Work

Although TL**2** showed superior biocompatibility over all other tested triazoles, it comes with its limitations, too. Formation of 1,5-disubstituted 1,2,3-triazoles by RuAAC requires elevated temperature and only proceeds in aprotic solvents.[279] This limits the current scope of applications for TL**2**, as the dimer CE phosphoramidite approach is impractical to introduce backbone modifications in long oligonucleotides (> 200 base pairs). Ideally, a 'click' protocol would be required to ligate azide- and alkyne-modified shorter oligonucleotides to form TL**2** (Figure V-19A). One possibility might be a recently reported nickel-catalysed azide–alkyne cycloaddition (NiAAC) that can form 1,5-disubstituted 1,2,3-triazoles at room temperature and under aqueous conditions.[283] Indeed, the authors reported ligation with protected azido carbohydrates and azido amino acids. However, whether this reaction can be applied to unprotected oligonucleotides is unclear.

An alternative approach might be the RuAAC-mediated ligation of azide- and alkyne modified oligonucleotides in an aprotic organic solvent. The polyanionic character of oligonucleotides prevents them from being dissolved in aprotic organic solvents. One possibility is to exchange the counterion from typically sodium to a long-alkyl-chain ammonium ion (Figure V-19B). This can be done, for example, with didodecyldimethylammonium bromide (DDAB),[306] and could allow RuAAC of azide- and alkyne-labelled oligonucleotides in aprotic organic solvents.

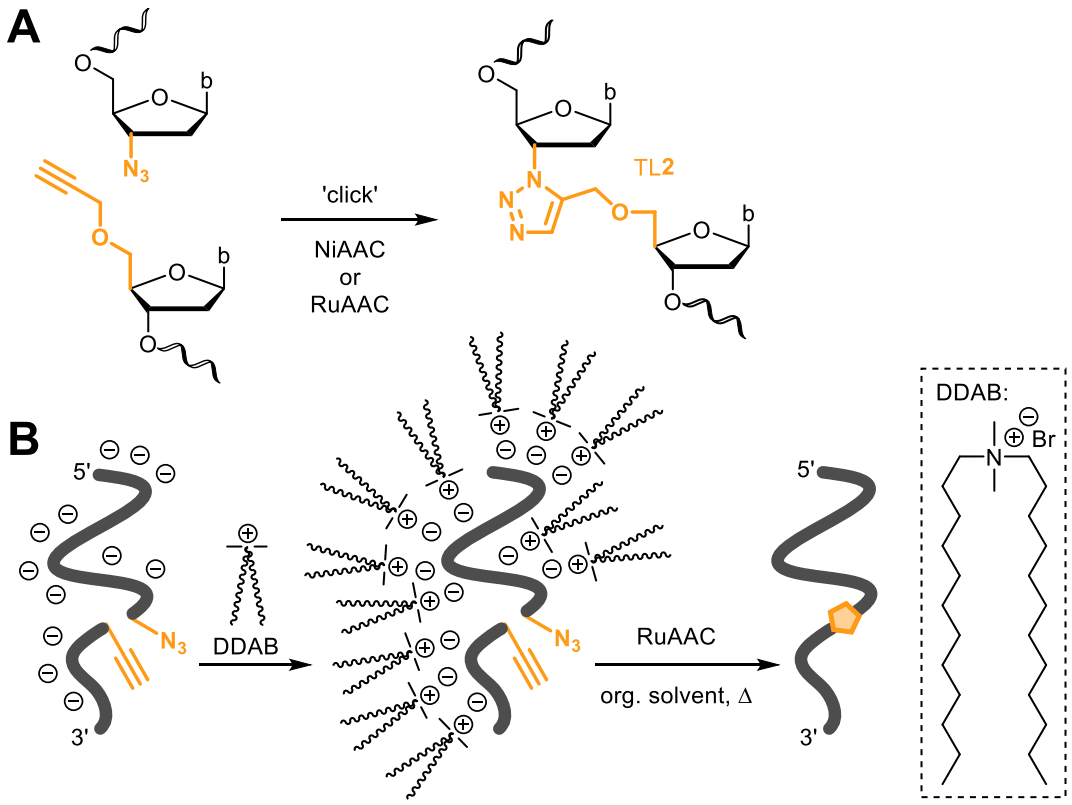



**Figure V-19: Proposed formation of TL2 with alkyne- and azide-labelled oligonucleotides.** (**A**) Schematic representation of the general formation of TL**2** with a 3′-azide- and a 5′-alkyne-modified ONs. The 1,5-disubstituted 1,2,3-triazole could be formed by RuAAC or NiAAC. (**B**) Schematic representation of RuAAC to form TL**2** with solubilised 3′-azide- and a 5′-alkyne-modified ONs in organic solvent. The ONs are shown in dark grey, the polyanionic character is illustrated with (−)-symbols, and the azide and alkyne functional groups are highlighted in orange. The formed TL**2** is symbolised by an orange pentagon. The sodium and bromonium counter-ions are not shown for simplicity.

With the potential of developing a RuAAC protocol to ligate azide- and alkyne-labelled oligonucleotides in organic solvent, and considering the recent progress towards NiAAC in aqueous solutions, TL**2** constitutes a promising backbone mimic for future applications in synthetic biology and biotechnology.

# CHAPTER VI

# OVERALL CONCLUSIONS AND FUTURE DIRECTIONS

"*We are dealing with an entity sui generis not comparable to any hitherto known group.*" (Johannes Friedrich Miescher, 1871)



150 years after Friedrich Miescher's discovery we are still dealing with a one-of-its-kind biomolecule, but we are continuously finding new ways of utilising and manipulating nucleic acids to our advantage. The controlled synthesis of oligonucleotides is no longer a dream, and chemically modified oligonucleotides have become readily accessible tools for research across the life sciences. The modification of oligonucleotide backbones has emerged as a challenging, yet promising approach to generate artificial oligomers with applications in therapeutics, diagnostics, synthetic biology, and biotechnology. The presented work investigated backbone-modified oligonucleotides as potential antisense oligonucleotides, or as artificial genetic templates.

*Chapter II* focused on the carbamate ($CA_X$) bond as a charge-neutral, achiral mimic of the natural phosphodiester backbone. Although the carbamate backbone is known to lower the affinity of carbamate-modified oligonucleotides with their complementary DNA and RNA targets, its nucleolytic resistance, charge-neutrality, and achiral character are promising for therapeutic applications. Thus, the carbamate backbone was combined with a bicyclic sugar modification—the locked nucleic acid—which is known to increase RNA binding affinity. Thermal melting experiments of $CA_X$–LNA-modified oligonucleotides showed that the additional

bicyclic sugar can indeed restore binding affinities to RNA targets, albeit it being dependent on the position of the LNA. As such, the LNA modification only mitigated reduced RNA affinity when positioned on the 3′-side of the carbamate linkage, or as an LNA–$CA_X$–LNA sandwich motif. The LNA–$CA_X$–LNA modification also showed a high selectivity in restoring RNA- over DNA-binding affinity. Such discrimination between DNA and RNA targets is highly desirable for potential antisense applications.

As a future direction, the integration of the on-resin carbamate bond formation into the solid phase oligonucleotide synthesis protocol could lower the required number of dimers necessary to cover all sequence and LNA–$CA_X$ combinations. This would expedite the research of LNA–$CA_X$-modified oligonucleotides towards *in vitro* or cellular antisense studies.

*Chapter III* introduced the amide bond (AM) as an alternative charge-neutral, achiral backbone mimic. Although requiring a synthetically challenging 3′-C–C bond formation at the monomer level, the amide linkage has some desirable attributes: (**i**) retained RNA target affinity; (**ii**) achiral; (**iii**) charge-neutral; and (**iv**) robust and high yielding solid phase amide bond formation protocols established for peptide synthesis. Regarding the latter, the integration of an Fmoc-based peptide coupling protocol into the solid phase oligonucleotide synthesis was explored but did not give the desired coupling results. Thus, a trityl protecting group strategy was further pursued and valuable trityl-protected nucleosides were prepared for the solid phase synthesis of amide-modified oligonucleotides.

*Chapter IV* reports the synthesis of AM–PO oligonucleotide chimeras for RNase H activation in antisense studies. The reported AM–PO oligonucleotide syntheses were based on the findings and monomers from its preceding chapter. In detail, the dimer approach was used to synthesise AM–PO chimeras with isolated amide linkages, and on-resin amide bond formations were used for the synthesis of AM–PO chimeras with consecutive amide linkages. The question whether the amide bond can be accommodated inside the phosphodiester bond-binding pocket of RNase H to trigger the degradation of an RNA target was initially addressed using modified oligonucleotides with isolated amides. The observed incompatibility of the amide bond to be accommodated inside the PO-binding pocket to activate RNase H prompted a revised gapmer design with consecutive amide linkages in the wings and an unmodified central gap. This AM–PO gapmer induced catalytic, RNase H-mediated degradation of an RNA target as efficiently as unmodified DNA or a 2′-OMe gapmer control. Moreover, the amide backbone greatly enhanced the nucleolytic resistance of the AM–PO gapmer compared to its unmodified or 2′-OMe-modified analogues. Cell studies indicated potentially enhanced cellular uptake of a fluorescently labelled AM–PO gapmer, but punctuated fluorescence suggested its entrapment inside endosomes.

A future objective could be the synthesis of 2′-homologated amino–carboxylic acid nucleosides with all four canonical nucleobases. Paradoxically, the necessity to restrict the amide backbones to the wings of a gapmer to elicit an RNase H response expands the chemical space of possible nucleosides. As such, amino–carboxylic acid nucleosides could be combined with additional 2′-modifications that are tolerated in

the wings but not in the central gap—e.g., LNA, 2′-OMe, etc. The presence of a 2′-functionality can be utilised to stereocontrol the glycosylation reaction during the synthesis of the amino–carboxylic acid nucleosides. This increases the number of possible synthetic strategies, thereby increasing the likelihood to establish a successful procedure for the synthesis of amino–carboxylic acid nucleosides with all four nucleobases. The accessibility of such nucleosides would ultimately allow cellular studies with a functional read out of antisense activities (e.g., luciferase knock down) or the mimicking of a clinically approved ASO therapeutic.

*Chapter V* investigated different triazole internucleoside linkages as part of genetic templates. Replication-compatible triazole linkages are versatile backbone mimics for gene synthesis, sequencing, nucleic acids detection, or the production of nanomaterials. Although many different triazole linkages have been studied for their read-through compatibilities by polymerases, all of them have their limitations such as PCR slow-down, mutagenicity, or read-through incompatibility. Chapter V reported the RuAAC-mediated synthesis of a dinucleoside CE phosphoramidite with a novel triazole linkage structure (TL**2**) for the solid phase synthesis of triazole-modified oligonucleotides. The novel triazole linkage induced no global structural distortion of duplexes with DNA and RNA targets, but destabilised them. The degree of destabilisation was in the average range when compared to the other tested triazoles. Importantly, TL**2**-modified genetic templates had the best polymerase read-through compatibility amongst the tested triazole-modified templates. This included continuation, speed, and correctness of the polymerase read-through. However, the formation of TL**2** requires RuAAC, which proceeds at elevated

temperatures and only in aprotic solvents, thereby limiting the application of TL**2** for gene synthesis.

A future direction could be the exploration of TL**2** formation *via* water-compatible NiAAC, or *via* RuAAC in organic solvent. The latter requires solubility of oligonucleotides in organic solvents, which could be achieved by exchanging sodium through long-alkyl-chain ammonium as the counter cation. Such a ligation protocol for the formation of a TL**2** backbone would drastically expand its possible applications for gene synthesis and across the life sciences.

# CHAPTER VII

# MATERIALS AND METHODS

# 1. Experimental Procedures

## 1.1. UV Melting Studies

UV DNA melting curves were recorded on a Cary 4000 Scan UV-Visible Spectrophotometer using 3 μM of each ON in a 10 mM phosphate buffer containing 200 mM NaCl at pH 7.0. Samples were annealed by heating to 85 °C (10 °C/min) and then slowly cooling to 20 °C (1 °C/min). Six successive cycles (heating and cooling) were then performed at a gradient of 0.5 °C/min and the change in UV absorbance at 260 nm was recorded. The melting temperatures for each experiment were calculated from the average of the 1st derivatives of the melting curves using in built software. The melting temperatures are reported as the average from two independent experiments. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.2. Circular Dichroism Studies

Circular dichroism in Chapter IV was recorded for duplexes formed from 2 μM ON38, ON39, or ON40 and 2 μM of a complementary DNA, or RNA target in 10 mM phosphate buffer with 200 mM NaCl (pH 7.0). Circular dichroism in Chapter V was recorded for duplexes formed from 3 μM ON42, ON45, or ON47 and 3 μM of a complementary DNA, or RNA target in 10 mM phosphate buffer with

200 mM NaCl (pH 7.0). In all cases, circular dichroism was performed on a Chirascan Plus spectrometer using a quartz cuvette with a pathlength of 1 mm. Scans were taken at rt from 200–340 nm with 0.5 a step and 1.0 s time point intervals. The average of four scans was taken and smoothed to 20 points using a third order polynomial (Savitzky-Golay, Origin 2017). The spectra were then baseline corrected to the θ-value at 340 nm. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.3. Foetal Bovine Serum Assay

5 nmol of ON38, ON39, or ON40 were dissolved in 50 μL of Dulbecco's PBS and 50 μL of FBS (Gibco, standard sterile filtered). This was then vortexed and 20 μL were removed (t = 0 min; control), mixed with formamide 1/1 (v/v) and kept at −20 °C. The remaining solution was incubated at 37 °C and 20 μL aliquots were removed at specified time points, mixed with formamide 1/1 (v/v) and kept at −20 °C. Samples were analysed by 20% PAGE (400 V) and viewed under short wavelength UV light using a G:Box (*Syngene*). *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.4. RNase H Assay

*Escherichia coli* RNase H (*New England Biolabs*, cat. no. M0523) was used for all RNase H studies. Oligonucleotides (1.4 μM for ON33–ON37; 0.33 μM for ON38,

ON39, or ON40), corresponding targets (1.3 μM for FL-RNA1, FL-RNA2, FL-RNA3, FL-RNA4, or FL-RNA5) and *E. coli* RNase H (20 U/mL) were incubated at 37 °C or 20 °C in 1 × RNase H reaction buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, 10 mM dithiothreitol (DTT), pH 8.3). At each time point, 15 μL aliquots were taken and quenched with aq EDTA-$Na_2$ (5 μL, 160 mM) at a concentration of 40 mM, denatured with formamide 1/1 (v/v) and kept at –20 °C in the dark. The aliquots were submitted to a 20% PAGE (400 V) and analysed using a G:Box (*Syngene*) with an excitation wavelength of λ = 460 nm and an emission filter of λ = 516–600 nm for detection of the FL label. Quantification of full-length FL-RNA2, FL-RNA3, FL-RNA4, or FL-RNA5 from gel images was performed using ImageJ. For quantification, rectangular sections of the gel lanes were selected from 8-bit inverted grayscale images and plotted using the inbuilt ImageJ analysis tool. Peaks corresponding to full-length FL-RNA2, FL-RNA3, FL-RNA4, or FL-RNA5 were enclosed using the straight-line tool and quantified using the wand (tracing) tool. All obtained values were normalised to t = 0 for the individual data set and plotted in Origin 2017. *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.5. Gel Extraction

After 20% PAGE as described above, the gel was visualised under UV and bands of interest were excised. The polyacrylamide bands were then crushed and incubated in a ThermoMixer (*Eppendorf*) at 37 °C, 950 rpm in RNA extraction buffer (50 mM Tris-HCl, 25 mM NaCl, pH 7.5, 7 mL) for 16 h. The polyacrylamide

was filtered off and the filtrate concentrated under reduced pressure and desalted using illustra™ NAP-10 and NAP-25 columns (*GE Healtcare*) consecutively. Extracted ONs were analysed by MS as described in the corresponding section. *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.6. Cell Culture

HeLa cells (regularly tested for mycoplasma contamination) were cultured in Dulbecco's modified eagle media (DMEM) [(high glucose, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-buffered, no phenol red, with glutamine (*Thermo Fisher Scientific*, cat. no. 21063029)], supplemented with 10% (v/v) FBS (*Thermo Fisher Scientific*, cat. no. 10270098) in a humidified incubator at 37 °C with 5% $CO_2$. *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.7. Cell Preparation for Microscopy

8-Well chambered cover glass slides (Nunc™ Lab-Tek™ from *VWR Int.*, cat. no. 734-2062) were coated with a 0.01% poly-L-lysine solution (*Sigma Aldrich*, cat. no. P8920) for 1 h at 37 °C. The poly-L-lysine solution was removed, and wells were washed two times with PBS. Then, HeLa cells (passage between 11–35) were seeded (5000 cells/well) and allowed to attach in DMEM supplemented with 10% (v/v) FBS for 24 h at 37°C. The medium was removed, and the wells were washed

two times with PBS. Fluorescently labelled oligonucleotides (FL-ON38, FL-ON39, FL-ON40; 5 μM in serum-free, colourless Opti-MEM™) were added and the cells were incubated at 37 °C for 16 h. Optionally EGF-AF647 (*Thermo Fischer Scientific*, cat. no. E35351) was added at a final concentration of 2 $\mu g \cdot mL^{-1}$ 30 min before removal of the medium. The medium was removed, the wells were washed with PBS and the cells fixed with 4% paraformaldehyde in PBS for 20 min at rt. After removal of the fixation buffer, the wells were washed with PBS, disassembled and a coverslip was mounted using ProLong™ Diamond Antifade Mountant with DAPI (*Thermo Fisher Scientific*, cat. no. P36966). *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.8. Confocal Microscopy

Cell imaging was performed on an Olympus FluoView FV1200 confocal microscope with a 40× UPlan FLN 1.30 oil immersion inverted objective. The samples were excited using internal lasers and emission detected using the inbuilt filter systems. Different channels were imaged using sequential excitation and detection of emission. Images were then processed using ImageJ software. DAPI, FL and AF647 were excited at 405, 488 and 635 nm respectively and emission detected using the corresponding inbuilt filter set ups (DAPI, AF488 and AF647 respectively). *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 1.9. Linear Copying Assays

Primer extension reactions were performed using the large fragment of DNA polymerase I (Klenow fragment) (*New England Biolabs*, cat. no. M0210S) or hot-start Taq polymerase (*New England Biolabs*, cat. no. M0495S) using a Bio-Rad T100 PCR thermal cycler. Used FL-primers and templates are listed in Table A-3. For Klenow fragment, a master-mix composed of NEB buffer 2 (10×, 1 μL), dNTPs (10 mM, 0.2 μL), polymerase (5 U/μL, 0.2 μL) and water (3.6 μL) was prepared. This mix (5 μL) was then added to a solution of FL-primer (10 μM, 1.5 μL), template (10 μM, 2.3 μL) and water (1.2 μL). Practically, oligonucleotide solutions were mixed with the master mix and immediately incubated on a pre-warmed thermocycler at 37 °C for 7.5 min. Reactions were quenched at 37 °C using a stop solution [1 mM EDTA-$Na_2$/formamide, 1/1 (v/v), 20 μL] before storage at –20 °C in the dark. For hot-start Taq polymerase, a master-mix composed of standard Taq buffer (10×, 1 μL), dNTPs (10 mM, 0.2 μL), polymerase (5 U/μL, 0.5 μL) and water (3.3 μL) were prepared. This mix (5 μL) was then added to a solution of FL-primer (10 μM, 1.5 μL), template (10 μM, 2.3 μL) and water (1.2 μL). Practically, oligonucleotide solutions were mixed with the master mix and immediately heat activated at 95 °C for 2 min before primer extension at 60 °C for 160 min. Reactions were quenched using a stop solution [1.0 mM EDTA-$Na_2$/formamide, 1/1 (v/v), 20 μL] before storage at –20 °C in the dark.

The samples were separated by 8% denaturing PAGE (400 V) and imaged using a G:Box (*Syngene*) with an excitation wavelength of 460 nm and an emission filter

of 516–600 nm for detection of the FL-labelled oligonucleotides. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

## 1.10. Linear Copying for Mass Spectrometry

Primer extension reactions were performed as described in Section 1.9; Chapter VII but scaled up 4×. After the specified extension times, Klenow fragment reactions were heated to 75 °C for 20 min to inactivate the polymerase. Hot-start Taq polymerase reactions were inactivated by addition of aq EDTA-$Na_2$ (160 mM, 3.33 μL). Each reaction was then mixed with PhOH/$CHCl_3$ (40 μL, *Thermo Fisher Scientific*, cat. no. 15593031), vortexed for 30 s and centrifuged (1000 rpm, 5 min). The aq layer was then transferred to a new tube, mixed with aq NaOAc (4 μL, 3 M, pH 5.2) and EtOH (132 μL), before incubating at –80 °C overnight. The samples were then centrifuged (13,000 rpm, 20 min), the supernatant removed and the pellet re-dissolved in water (10 μL). These reaction mixtures were analyzed by UPLC-MS as described in Section 3.3; Chapter VII. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

## 1.11. qPCR Kinetics

qPCR reactions were performed using hot-start Taq polymerase (*New England Biolabs*, cat. no. M0495S) on a Bio-Rad CFX96 qPCR detection system. Template and primers can be found in Table A-3. Master-mixes composed of standard Taq

buffer (10×, 2 μL), EvaGreen (20×, 1 μL), dNTPs (10 mM, 0.4 μL), polymerase (5 U/μL, 0.1 μL) and water (13.5 μL) were prepared. This mix (17 μL) was then added to a solution of FP (10 μM, 1 μL), RP (10 μM, 1 μL) and template (18.7 pM, 1 μL). Samples for all qPCR extension times were prepared together and stored at 4 °C before sequential use. qPCR thermal cycling conditions consisted of thermal activation (120 s, 95 °C), and 31 cycles of denaturation (15 s, 95 °C) and annealing/extension (15, 30, 60, 120 or 240 s, 60 °C), with emission recorded at the end of each extension step. Samples were excited at 450–490 nm and emission monitored at 510–530 nm. Threshold cycles were determined at a fluorescence set to F = 300 (a.u.). *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

## 1.12. Sequencing of PCR Amplicons

TL-containing templates [Template 2 (TL**2**), Template 3 (TL**3**), Template 4 (TL**4**) and Template 5 (TL**5**)] were amplified using hot-start Taq polymerase (*New England Biolabs*, cat. no. M0495S) and tailed primers (LIC-FP, LIC-RP; Table A-3) for cloning using an aLICator LIC cloning and expression kit (*Thermo Fisher*, cat. no. K1241). Master-mixes composed of standard Taq buffer (10×, 2 μL), dNTPs (10 mM, 0.4 μL), polymerase (5 U/μL, 0.1 μL) and water (14.5 μL) were prepared. This mix (17 μL) was then added to a solution of LIC-FP (10 μM, 1 μL), LIC-RP (10 μM, 1 μL) and template (18.7 pM, 1 μL). PCR thermal cycling conditions consisted of thermal activation (120 s, 95 °C), and 31 cycles of denaturation (15 s, 95 °C) and annealing/extension (45 s, 60 °C) and a final extension time of 7 min at

60 °C. Samples were then purified using a QIAquick PCR purification kit (*QIAGEN*, cat. no. 28104) and successful amplification was confirmed by 2% agarose gel electrophoresis with 5 µL of the purified samples. The purified PCR products were annealed to a pLATE11 vector according to the aLICator LIC cloning and expression protocol. The annealing mixture (2 µL) was used for transformation of 20 µL of NEB 5-alpha competent *E. coli* cells (C2987I) according to the vendor's High Efficiency Transformation Protocol. Cells were spread onto an LB agar selection plate supplemented with ampicillin (100 µg/mL). Following incubation at 37 °C for 16 h, colonies were picked and grown in a small-scale culture (10 mL) containing LB media and 100 µg/mL ampicillin at 37 °C for 16 h with 250 rpm. The plasmid DNA was extracted from the bacterial cells using the QIAprep Spin Miniprep kit according to the manufacturer's protocol (*QIAGEN*, cat. no. 27104). The plasmid was eluted with 50 µL water and the concentration was measured using a NanoDrop™ 2000/2000c spectrophotometer (*Thermo Fisher Scientific*). Plasmids were sent for Sanger sequencing (*Source BioScience*) with a suitable primer (Seq-FP; Table A-3) according to the sample preparation requirements. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

# 2. Oligonucleotide Synthesis

## 2.1. Synthesis of DNA Oligonucleotides

Standard DNA phosphoramidites, solid supports and reagents were purchased from *Link Technologies* and *Applied Biosystems*. 5′-Fluorescein CE phosphoramidite **4** and 5′-amino-dT CE phosphoramidite were purchased from *Link Technologies* (cat. no. 2134) and *Glen Research* (cat. no 10-1932-02), respectively. Automated solid phase synthesis of oligonucleotides (trityl off) was performed on an Applied Biosystems 394 synthesiser. Synthesis was performed on a 1.0 μmol scale involving cycles of acid-catalysed detritylation, coupling, capping, and iodine oxidation. Standard DNA phosphoramidites were coupled for 60 s, and chemically modified phosphoramidites for 10 min. Coupling efficiencies and overall synthesis yields were determined by the inbuilt automated trityl cation conductivity monitoring facility. The oligonucleotides were then cleaved from the solid support, and the nucleobase and backbone protecting groups were removed by exposure to concd aq for 60 min at rt followed by heating in a sealed glass vial at 55 °C for 5 h. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 2.2. Synthesis of RNA Oligonucleotides

RNA synthesis was performed using 2′-*O*-*tert*-butyldimethylsilyl (2′-*O*-TBS)-protected RNA CE phosphoramidites with tertbutylphenoxyacetyl-protected A, G and C nucleobases. During each cycle, RNA phosphoramidites were coupled at rt for 10 min using benzylthiotetrazole (BTT) as the coupling agent, obtaining coupling efficiencies of >97%. *tert*-Butylphenoxyacetic anhydride was used as the capping agent and 0.1 M iodine as the oxidizing agent (*Sigma-Aldrich*). Cleavage of oligonucleotides from the solid support and deprotection of the nucleobases and backbone were achieved by exposure to concd aq $NH_3$/EtOH [3/1 (v/v), 4 mL] for 2 h at rt, followed by heating in a sealed glass vial for 2 h at 55°C.

The oligonucleotides were then concentrated to a small volume under reduced pressure, transferred to 15 mL plastic tubes and freeze dried (lyophilised). The residue was dissolved in DMSO (300 μL) and $Et_3N{\cdot}3HF$ (300 μL) was added after which the reaction mixtures were kept at 65 °C for 2.5 h. Aqueous NaOAc (3 M, 50 μL) and *n*-BuOH (3 mL) were added, the mixture was vortexed, and the samples were kept at −80 °C for 30 min before centrifugation at 13,000 rpm at 4 °C for 10 min. The supernatant was decanted, and the precipitate was washed twice with 70% EtOH (0.75 mL), then dried *in vacuo*. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 2.3. Synthesis of PO–AM Chimeras

Integration of on-resin amide bond formation into the solid phase oligonucleotide synthesis cycle was achieved as outlined in Figure IV-10. The following describes the synthesis of ON38 and FL-ON38 in detail. The 3′-wings of ON38 and FL-ON38 were synthesised on MMTrHN-T resin (42 mg, loading 23.6 μmol·g$^{-1}$, 1.0 μmol; see Scheme IV-1). Detritylation was performed on an Applied Biosystems 394 synthesiser as for standard DNA synthesis for 80 s. For amide couplings, the synthesis column was removed from the synthesiser and equipped with a 1 mL plastic syringe on each side. In a separate vial, 4-methylmorpholine (3.5 μL, 32 μmol, 32 equiv), PyBOP (5.2 mg, 10 μmol, 10 equiv) and monomer **46** (6 mg, 10 μmol, 10 equiv) were mixed in anhyd DMF (400 μL) and incubated at rt for 5 min. The solution was added to the column containing the growing oligonucleotide with the 5′-$NH_2$ *via* one of the syringes and the solution was pushed back and forth three times through the column using the two attached syringes. The reaction was left at rt for 2 h agitating the reaction mixture using the syringes approx. every 20 min. The solvents were then removed, and the resin was washed with anhyd MeCN (3 × 1 mL) using the two syringes as described before. The column was transferred back to the synthesiser for detritylation and then washed with MeCN. This allows for repetitive amide couplings as described in Figure IV-10. To switch from amide couplings to standard oligonucleotide synthesis, the 5′-deprotected oligomer on the solid support was incubated with monomer **41** as described above for amide couplings and transferred back to the synthesiser for detritylation. Solid phase oligonucleotide synthesis was performed as described in

Section 2.1; Chapter IV. To switch from standard oligonucleotide synthesis to amide couplings, commercially available 5′-amino-dT CE phosphoramidite (*Glen Research*; cat. no 10-1932-02) was coupled with an extended coupling time of 10 min to introduce a 5′-$NH_2$ functionality (without the capping step) and amide couplings were performed as outlined above. The final amide coupling was performed using monomer **41** to introduce a 5′-OH after standard detritylation. After synthesis, the resin was divided (2 × 25 mg) and one half was cleaved and deprotected as described in Section 2.1; Chapter IV to give ON38. The other half was transferred back to the synthesiser a coupled with 5′-fluorescein CE phosphoramidite **4** for an extended coupling time of 10 min introducing the 5'-terminal fluorescein in FL-ON38. Cleavage from the resin and deprotection was carried out as described in Section 2.1; Chapter IV. *Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

# 3. Purification and Characterisation of Oligonucleotides

## 3.1. Purification by HPLC

The fully deprotected oligonucleotides were purified by RP-HPLC on a Gilson system using a Luna 10 μm C8(2) 100 Å pore Phenomenex column (250 × 10 mm) with a gradient of MeCN in aq triethylammonium bicarbonate (TEAB) (buffer A, 0.1 M TEAB, pH 7.5; buffer B, 50% buffer A in MeCN). Buffer B was increased from 20–50% over 20 min at a flow rate of 4 mL·min$^{-1}$, and elution was monitored by UV absorption between 260–295 nm. For ON38, buffer B was increased from 30–50%, and for FL-ON38, buffer B was increased from 25–50%. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 3.2. Purification by Gel Electrophoresis

The samples were separated by 8% denaturing PAGE (400 V, ca. 1 nmol of oligonucleotide per lane) and the gel was visualised under UV. Bands of interest were excised, crushed and incubated in a ThermoMixer (*Eppendorf*) at 37 °C, 950 rpm in water (ca. 1 mL per band) for 16 h. The polyacrylamide was then filtered off and the filtrate was concentrated under reduced pressure. Oligonucleotides were

desalted using illustra™ NAP-10 and NAP-25 columns (*GE Healthcare*), consecutively. Extracted oligonucleotides were then analysed by MS. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

## 3.3. Characterisation by Mass Spectrometry

Oligonucleotide MS was recorded on a UPLC-MS Waters XEVO G2-QTOF ($ESI^-$) using an ACQUITY UPLC ON BEH C18 column, 130 Å (1.7 μm, 2.1 mm × 50 mm). A gradient of MeOH in $Et_3N$ and hexafluoroisopropanol (HFIP) was used (buffer A, 8.60 mM $Et_3N$, 200 mM HFIP in 5% MeOH/$H_2O$ (v/v); buffer B, 20% buffer A in MeOH). Buffer B was increased from 0–70% over 8 min, at a flow rate of 0.2 $mL \cdot min^{-1}$. Raw data were processed and deconvoluted using the Waters MassLynx v4.1 software package. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

# 4. Small Molecule Synthesis

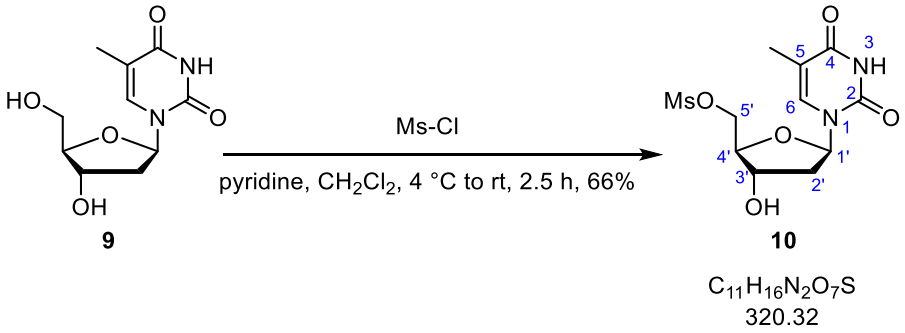
## 4.1. General Procedures

All reagents for synthesis were purchased from *Sigma-Aldrich, Acros Organics*, *Fluorochem*, *Carbosynth* and *Fisher Scientific*, and used without further purification. Anhydrous solvents ($CH_2Cl_2$, MeCN, THF, 1,4-dioxane, pyridine, DMF and $Et_3N$) were obtained using an MBraun Bench Top solvent purification system. All air/moisture sensitive reactions were carried out under inert atmosphere (argon) in oven-dried glassware. Solvents used for phosphitylation reactions and purification were also degassed by bubbling the solvents with argon for 20 min. Reactions were monitored by thin-layer chromatography using Merck Kieselgel 60 F24 silica gel plates (0.22 mm thickness, aluminium-backed), visualised by UV irradiation at 254/265 nm and by staining with *p*-anisaldehyde followed by gentle heating. Column chromatography was carried out under pressure (Biotage® SP4) using Biotage® Sfär columns. $^{1}$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) spectra were measured on a Bruker AVIIIHD 400, a Bruker AVII 500 (with a $^{13}$C cryoprobe) or a Bruker AVIII 600 (with a prodigy $N_2$ broadband cryoprobe) spectrometer. $^{1}$H and $^{13}$C NMR spectra were referenced to the appropriate deuterated solvent signal, chemical shifts are given in ppm, and all coupling constants (*J*) are quoted in Hertz (Hz). Assignment of the compounds was aided by correlation spectroscopy (COSY; $^{1}$H–$^{1}$H), heteronuclear single quantum correlation (HSQC; $^{1}$H–$^{13}$C) and

heteronuclear multiple bond correlation (HMBC; $^{1}H–^{13}C$) experiments when feasible. High-resolution mass spectra (HRMS) were recorded by the Departmental Mass Spectrometry Service, University of Oxford on a Thermo Scientific Exactive Mass Spectrometer (using a Waters Equity autosampler and pump) for electrospray ionisation (ESI) and an Agilent 7200 Accurate Mass QTOF GCMS (using a SIM Direct Insertion Probe) for electron ionization (EI). High-resolution values are calculated to 4 decimal places from the molecular formula, and all values are within a tolerance of 5 ppm. *Experimental procedure reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society, and from Ref.*[182] *and Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 4.2. Compounds of Chapter II

2′-Deoxy-5′-*O*-(methanesulfonyl)thymidine (**10**)[176]




Thymidine (**9**) (2.00 g, 8.26 mmol) was co-evaporated with anhyd pyridine (2 × 8 mL) and was then suspended in anhyd pyridine (50 mL). A solution of methanesulfonyl chloride (674 µL, 8.68 mmol) in anhyd $CH_2Cl_2$ (10 mL) was added dropwise at 4 °C over 1 h and the reaction mixture was then stirred at rt for 2.5 h. The resulting clear solution was quenched with MeOH (5 mL) and the solvents were removed *in vacuo* followed by co-evaporation with MePh (3 × 10 mL). The resulting solid was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–10%) to give mesylate **10** (1.74 g, 5.44 mmol, 66%) as a white solid.

***R*f** 0.49 ($CH_2Cl_2$/MeOH, 9/1);

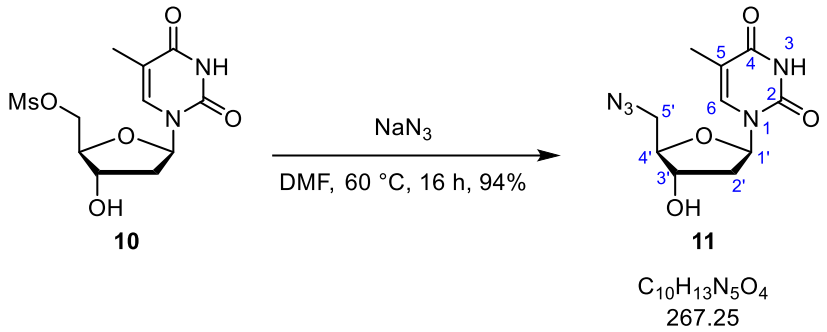
**IR** ($cm^{-1}$): 3366 (O–H), 3023, 2935 (O–H), 2843, 1714 (C=O), 1660, 1351;

**$^1$H NMR** (400 MHz, DMSO-$d_6$): δ 11.32 (s, 1H, NH), 7.48 (d, *J* = 1.3 Hz, 1H, H6), 6.22 (dd, *J* = 7.7, 6.3 Hz, 1H, H1′), 5.50 (d, *J* = 4.4 Hz, 1H, OH), 4.40 (dd, *J* = 11.0, 3.7 Hz, 1H, H5′), 4.35 (dd, *J* = 11.0, 5.5 Hz, 1H, H5″), 4.29 – 4.23 (m, 1H, H3′), 3.97 (dt, *J* = 5.5, 3.7 Hz, 1H, H4′), 3.22 (s, 3H, $CH_3$-*mesyl*), 2.19 (ddd, *J* = 13.8, 7.7, 6.3 Hz, 1H, H2′), 2.10 (ddd, *J* = 13.8, 6.3, 3.5 Hz, 1H, H2″), 1.78 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$): δ 163.7 (C4), 150.5 (C2), 135.8 (C6), 109.9 (C5), 84.0 (C1′), 83.4 (C4′), 70.2 (C3′), 69.6 (C5′), 38.4 (C2′), 36.7 ($CH_3$-*mesyl*), 12.1 ($CH_3$-*thymine*);

***m/z* (ESI$^+$)**: 321.098 ($[M + H]^+$, 8.97%), 343.078 ($[M + Na]^+$, 19.8%), 641.195 ($[2M + H]^+$, 31.8%), 663.130 ($[2M + Na]^+$, 100%), 983.213 ($[3M + Na]^+$, 65.7%), 1303.283 ($[4M + Na]^+$, 19.7%).

5′-Azido-2′,5′-dideoxythymidine (**11**)[176]




Sodium azide (1.83 g, 28.2 mmol) was added to a solution of mesylate **10** (1.70 g, 5.31 mmol) in DMF (20 mL) and the resulting suspension was stirred at 60 °C for 16 h. The solvents were then evaporated *in vacuo* and the remaining solid was re-suspended in MeOH and filtered through celite. The solvent was removed under reduced pressure and the crude was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–10%) to give azide **11** (1.33 g, 4.99 mmol, 94%) as a white solid.

***R*f** 0.65 ($CH_2Cl_2$/MeOH, 9/1);

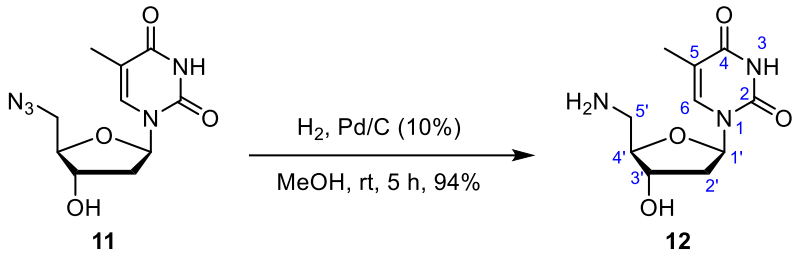
**IR** ($cm^{-1}$): 3381 (O–H), 3184, 3113, 3045, 2928 (O–H), 2092 ($N_3$), 1719 (C=O), 1652, 1217;

**$^1$H NMR** (400 MHz, DMSO-$d_6$): δ 11.33 (s, 1H, NH), 7.49 (d, *J* = 1.3 Hz, 1H, H6), 6.43 – 5.95 (m, 1H, H1′), 5.41 (d, *J* = 4.4 Hz, 1H, OH), 4.25 – 4.12 (m, 1H, H3′), 3.84 (td, *J* = 5.3, 3.6 Hz, 1H, H4′), 3.55 (d, *J* = 5.3 Hz, 2H, H5′ and H5″), 2.25 (ddd, *J* = 13.8, 7.6, 6.6 Hz, 1H, H2′), 2.08 (ddd, *J* = 13.8, 6.5, 3.7 Hz, 1H, H2″), 1.79 (d, *J* = 1.3 Hz, 3H, $CH_3$);

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$): δ 163.7 (C4), 150.5 (C2), 136.1 (C6), 109.8 (C5), 84.6 (C1′), 83.9 (C4′), 70.7 (C3′), 51.7 (C5′), 38.1 (C2′), 12.1 ($CH_3$);

***m/z* (ESI$^+$)**: 290.086 ($[M + Na]^+$, 9.20%), 535.235 ($[2M + H]^+$, 49.7%), 557.187 ($[2M + Na]^+$, 86.1%), 824.286 ($[3M + Na]^+$, 100%), 1091.389 ($[4M + Na]^+$, 44.8%), 1358.490 ($[5M + Na]^+$, 41.6%), 1626.628 ($[6M + Na]^+$, 10.3%);

5′-Amino-2′,5′-dideoxythymidine (**12**)[176]




$C_{10}H_{15}N_3O_4$
241.25

A solution of azide **11** (1.15 g, 4.31 mmol) in MeOH (100 mL) was stirred at rt for 5 h under a hydrogen gas atmosphere in the presence of 10% Pd/C (221 mg, 2.08 mmol). The reaction mixture was then filtered through celite and the solvent was removed under reduced pressure to give amine **12** (0.98 g, 4.05 mmol, 94%) as a white solid, which was used without any further purification.

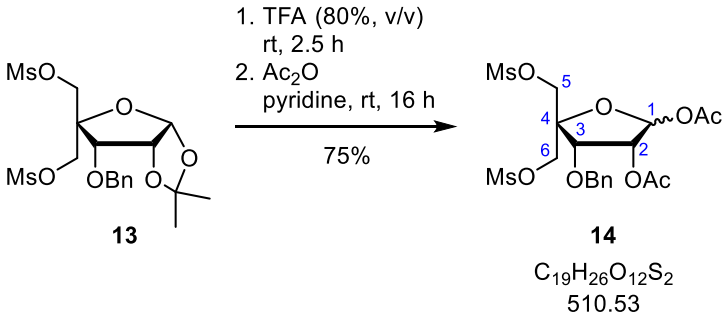
**IR** ($cm^{-1}$): 3346, 3023, 2923 (O–H), 2841, 1690 (C=O), 1645, 1447, 1274;

**$^1$H NMR** (400 MHz, $H_2O$-suppression, DMSO-$d_6$): δ 7.62 (d, *J* = 1.3 Hz, 1H, H6), 6.14 (dd, *J* = 7.7, 6.2 Hz, 1H, H1′), 4.19 (dt, *J* = 6.4, 3.4 Hz, 1H, H3′), 3.64 (td, *J* = 5.2, 3.4 Hz, 1H, H4′), 2.72 (dd, *J* = 5.2, 1.2 Hz, 2H, H5′ and H5″), 2.12 (ddd, *J* = 13.6, 7.7, 6.4 Hz, 1H, H2'), 2.03 (ddd, *J* = 13.6, 6.2, 3.4 Hz, 1H, H2″), 1.78 (d, *J* = 1.3 Hz, 3H, $CH_3$);

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$): δ 164.1 (C4), 150.7 (C2), 136.2 (C6), 109.6 (C5), 87.8 (C1′), 83.4 (C4′), 70.8 (C3′), 43.7 (C5′), 38.9 (C2′), 12.2 ($CH_3$);

***m/z* (ESI$^+$)**: 242.140 ($[M + H]^+$, 57.1%), 483.240 ($[2M + H]^+$, 100%), 505.221 ($[2M + Na]^+$, 31.4%), 724.320 ($[3M + H]^+$, 54.0%), 346.349 ($[3M + Na]^+$, 51.8%), 965.432 ($[4M + H]^+$, 10.9%), 987.546 ($[4M + Na]^+$, 11.5%).

1,2-Di-*O*-acetyl-3-*O*-benzyl-4-*C*-methanesulfonoxy-methyl-5-*O*-methanesulfonyl-D-*erythro*-pentofuranose (**14**)[177]




Commercially available compound **13** (1.03 g, 2.15 mmol) was dissolved in aq TFA [80% (v/v), 5 mL] and stirred at rt for 2.5 h. The solvents were evaporated, and the resulting yellow oil was diluted with $CH_2Cl_2$ (12 mL). The organic layer was washed with sat. aq $NaHCO_3$ (2 × 15 mL), dried over $Na_2SO_4$, and filtered. Evaporation of the solvent gave a yellow oil which was co-evaporated with anhyd pyridine (2 × 4 mL), and diluted in anhyd pyridine (3 mL). $Ac_2O$ (800 μL) was added and the reaction was stirred at rt for 16 h. The reaction was quenched with sat. aq $NaHCO_3$ (15 mL) and extracted with EtOAc (2 × 13 mL). The combined organic layers were successively washed with 1 N aq HCl (3 × 10 mL) and brine (10 mL), and dried over $Na_2SO_4$. Filtration and evaporation of the solvent gave acetylated glycosyl donor **14** (0.82 g, 1.60 mmol, 75%) which was used without any further purification.

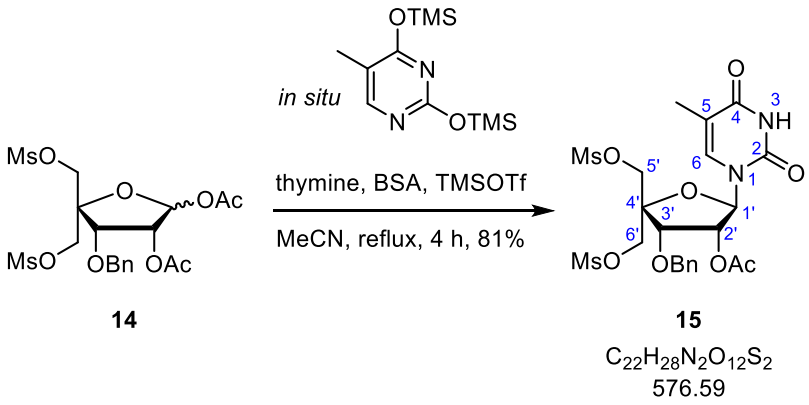
***R*f** 0.37 (petroleum ether (40–60)/EtOAc, 1/1);

**$^{1}$H NMR** (400 MHz, $CDCl_3$) δ 7.42 – 7.27 (m, 5H, H-Ar), 6.71 (s, 1H, H1′), 5.38 (d, *J* = 4.8 Hz, 1H, H2′), 4.78 – 4.05 (m, 7H, H3′ *and* 3 × $CH_2$), 3.02 (s, 3H, $CH_3$-*mesyl*), 3.01 (s, 3H, $CH_3$-*mesyl*), 2.15 (s, 3H, $CH_3$-*acetyl*), 2.11 (s, 3H, $CH_3$-*acetyl*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 169.4 (O**C**(O)$CH_3$), 169.0 (O**C**(O)$CH_3$), 136.6 (C-Ar), 128.8 (CH-Ar), 128.6 (CH-Ar), 128.3 (CH-Ar), 97.5 (C1′), 82.9 (C4′), 78.9 (C3′), 74.2 (O**C**$H_2$Ph), 73.6 (C2′), 69.0 ($CH_2$OMs), 68.6 ($CH_2$OMs), 37.9 ($CH_3$-*mesyl*), 37.7 ($CH_3$-*mesyl*), 21.1 ($CH_3$-*acetyl*), 20.9 ($CH_3$-*acetyl*);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{19}H_{26}O_{12}S_2Na$, 533.0758; found, 533.0753.

1-(2-*O*-Acetyl-3-*O*-benzyl-4-*C*-methanesulfonoxymethyl-5-*O*-methanesulfonyl-β-D-*erythro*-pentofuranosyl)thymine (**15**)[177]




**14** **15**

$C_{22}H_{28}N_2O_{12}S_2$
576.59

*N*,*O*-Bis(trimethylsilyl)acetamide (BSA) (1.09 mL, 4.43 mmol) was added to a solution of glycosyl donor **14** (820 mg, 1.60 mmol) and thymine (254 mg, 2.02 mmol) in anhyd MeCN (4 mL), and refluxed for 1 h. Trimethylsilyl triflate (TMSOTf) (380 mL, 2.10 mmol) was added, and the reaction mixture was refluxed for another 3 h. The reaction was cooled to rt, $CH_2Cl_2$ (10 mL) was added, and the solution was washed with half-saturated $NaHCO_3$ (2 × 10 mL) and brine (10 mL). The combined $NaHCO_3$ washes were extracted with $CH_2Cl_2$ (6 mL) and the combined organic layers were dried over $Na_2SO_4$, filtered, and the solvent was evaporated under reduced pressure. The residue was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–2%) to give product **15** (747 mg, 1.29 mmol, 81%).

***R*f** 0.26 ($CH_2Cl_2$/MeOH, 97/3);

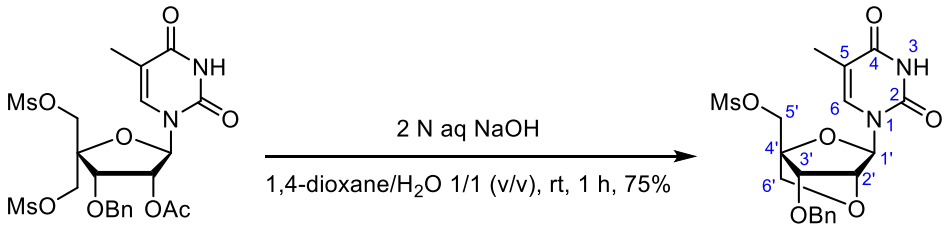
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.99 (s, 1H, NH), 7.44 – 7.27 (m, 4H, H-Ar), 7.07 (d, *J* = 1.3 Hz, 1H, H6), 5.71 (d, *J* = 3.3 Hz, 1H, H1′), 5.56 (dd, *J* = 6.5, 3.3 Hz, 1H,

H2′), 4.69 (d, *J* = 6.5 Hz, 1H, H3′), 4.63 – 4.25 (m, 6H, 3 × $CH_2$), 3.02 (s, 3H, $CH_3$-*mesyl*), 3.00 (s, 3H, $CH_3$-*mesyl*), 2.11 (s, 3H, $CH_3$-*acetyl*), 1.92 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 170.1 (O**C**(O)$CH_3$), 163.6 (C4), 150.2 (C2), 138.0 (C6), 136.7 (C-Ar), 128.8 (CH-Ar), 128.5 (CH-Ar), 128.3 (CH-Ar), 112.0 (C5), 92.5 (C1′), 84.1 (C4′), 77.9 (C3′), 74.9 (O**C**$H_2$Ph), 73.8 (C2′), 68.5 ($CH_2$OMs), 67.6 ($CH_2$OMs), 37.8 ($CH_3$-*mesyl*), 37.7 ($CH_3$-*mesyl*), 20.9 ($CH_3$-*acetyl*), 12.4 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{22}H_{28}O_{12}N_2S_2Na$, 599.0981; found, 599.0973.

5′-*O*-Methanesulfonyl-3′-*O*-benzyl-LNA thymidine (**16**)[177]




**15**

**16**

$C_{19}H_{22}N_2O_8S$

438.45

Compound **15** (676 mg, 1.17 mmol) was dissolved in 1,4-dioxane (1.5 mL), $H_2O$ (1.5 mL) and 2 N aq NaOH (3.0 mL) and stirred at rt for 1 h. The reaction mixture was diluted with sat. aq $NaHCO_3$ (3.0 mL) and neutralised with 1 N aq HCl. The neutralised solution was extracted with $CH_2Cl_2$ (2 × 7 mL) and the combined org layers were dried over $Na_2SO_4$, filtered, and the solvents were evaporated. The remaining solid was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–3%) to give mesylated LNA **16** (383 mg, 0.87 mmol, 75%).

***R*f** 0.43 ($CH_2Cl_2$/MeOH, 9/1);

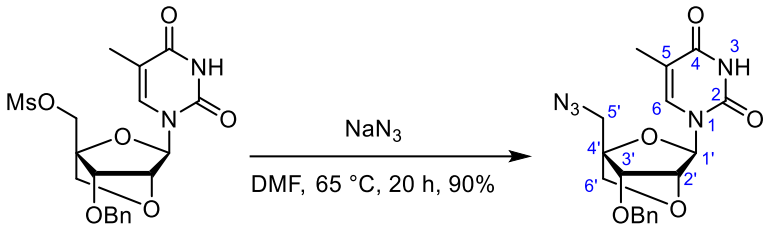
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.65 (bs, 1H, NH), 7.42 (d, *J* = 1.3 Hz, 1H, H6), 7.39 – 7.28 (m, 5H, H-Ar), 5.67 (d, *J* = 0.7 Hz, 1H, H1′), 4.67 (d, *J* = 11.4 Hz, 1H, 1 × $OCH_2Ph$), 4.63 – 4.50 (m, 4H, H2′, H5′, H5″ *and* 1 × $OCH_2Ph$ ), 4.09 (d, *J* = 7.8 Hz, 1H, H6′), 3.92 (s, 1H, H3′), 3.88 (d, *J* = 7.8 Hz, 1H, H6″), 3.08 (s, 3H, $CH_3$-*mesyl*), 1.93 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.6 (C4), 149.7 (C2), 136.6 (C-Ar), 134.3 (C6), 128.8 (CH-Ar), 128.6 (CH-Ar), 128.2 (CH-Ar), 111.0 (C5), 87.8 (C1′), 85.7 (C4′),

76.8 (C2′), 76.2 (C3′), 72.6 (O**C**$H_2$Ph), 71.8 (C6′), 64.2 (C5′), 38.1 ($CH_3$-*mesyl*), 12.6 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{19}H_{22}O_8N_2SNa$, 461.0989; found, 461.0986.

5′-Azido-5′-deoxy-3′-*O*-benzyl-LNA thymidine (**17**)[182]




**16**

**17**

$C_{18}H_{19}N_5O_5$

385.38

Sodium azide (1.01 g, 15.5 mmol) was added to a solution of mesylated LNA **16** (1.00 g, 2.28 mmol) in DMF (7.0 mL), and the reaction mixture was stirred at 65 °C for 20 h. The solvent was evaporated *in vacuo*, the crude re-suspended in EtOAc (70 mL), and filtered over celite. Evaporation of the solvent gave an oil which was submitted to column chromatography (petroleum ether(40–60)/EtOAc, 15–100%) to give azide **17** (790 mg, 2.05 mmol, 90%).

***R*f** 0.60 ($CH_2Cl_2$/MeOH, 9/1);

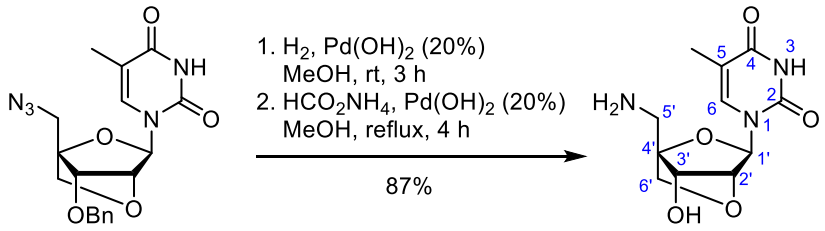
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.88 (bs, 1H, NH), 7.40 (d, *J* = 1.3 Hz, 1H, H6), 7.39 – 7.27 (m, 5H, H-Ar), 5.64 (s, 1H, H1′), 4.66 (d, *J* = 11.4 Hz, 1H, 1 × $OCH_2Ph$), 4.58 (s, 1H, H2′), 4.54 (d, *J* = 11.4 Hz, 1H, 1 × $OCH_2Ph$), 4.03 (d, *J* = 7.9 Hz, 1H, H6′), 3.87 (s, 1H, H3′), 3.84 (d, *J* = 7.9 Hz, 1H, H6′), 3.75 (d, *J* = 13.7 Hz, 1H, H5′), 3.67 (d, *J* = 13.7 Hz, 1H, H5′′), 1.94 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.7 (C4), 149.8 (C2), 136.7 (C-Ar), 134.4 (C6), 128.7 (CH-Ar), 128.5 (CH-Ar), 128.0 (CH-Ar), 110.8 (C5), 87.8 (C1′), 86.9 (C4′), 76.8 (C2′), 76.4 (C3′), 72.5 ($OCH_2Ph$), 72.5 (C6′), 47.5 (C5′), 12.9 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{18}H_{20}O_5N_5$, 386.1459; found, 386.1460.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

5′-Amino-5′-deoxy-LNA thymidine (**18**)[182]




**17** **18**

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

$Pd/(OH)_2$ (20%, 338 mg, 2.41 mmol) was added to a solution of azide **17** (1.50 g, 3.90 mmol) in MeOH (50 mL), and the reaction mixture was stirred at rt for 3 h under a hydrogen gas atmosphere. Ammonium formate (1.84 g, 29.2 mmol) was then added and the reaction was refluxed for 4 h under an argon gas atmosphere. The reaction mixture was cooled down to rt, diluted with $CH_2Cl_2$ (30 mL) and MeOH (30 mL), filtered over celite, and the solvents were evaporated. The residue was submitted to column chromatography ($CH_2Cl_2$/MeOH, 3–5%; +1% $Et_3N$) to give amine **18** (963 mg, 3.58 mmol, 87%).

***R*f** 0.25 ($CH_2Cl_2$/MeOH, 17/3; +1% $Et_3N$);

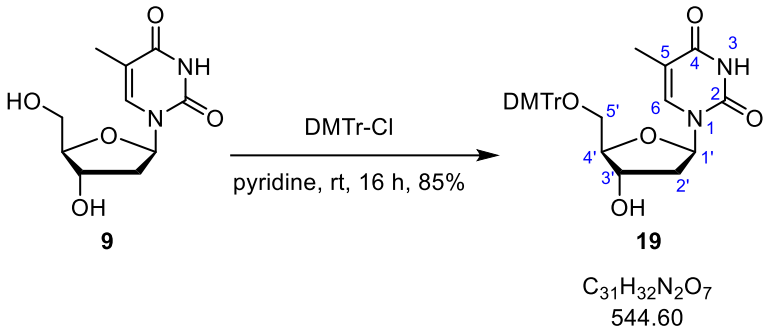
**$^1$H NMR** (400 MHz, $D_2O$) δ 7.53 (d, *J* = 1.3 Hz, 1H, H6), 5.66 (s, 1H, H1′), 4.50 (s, 1H, H2′), 4.20 (s, 1H, H3′), 4.02 (2 × d, *J* = 18.0 Hz, 2H, H6′ *and* H6″), 3.37 (d, *J* = 14.4 Hz, 1H, H5′), 3.29 (d, *J* = 14.4 Hz, 1H, H5″), 1.88 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $D_2O$) δ 166.7 (C4), 151.2 (C2), 136.0 (C6), 110.6 (C5), 86.9 (C4′), 86.9 (C1′), 79.4 (C2′), 71.6 (C6′), 70.4 (C3′), 37.0 (C5′), 11.6 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{11}H_{16}O_5N_3$, 270.1085; found, 270.1085.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

5′-*O*-(4,4′-Dimethoxytrityl)thymidine (**19**)[180,181]




**9** **19**

$C_{31}H_{32}N_2O_7$
544.60

Thymidine (**9**) (2.00 g, 8.28 mmol) was co-evaporated with anhyd pyridine (2 × 5 mL) and suspended in anhyd pyridine (15 mL). 4,4′-Dimethoxytrityl chloride (3.35 g, 9.91 mmol), dissolved in anhyd pyridine (15 mL), was added dropwise, and the mixture was stirred at rt for 16 h. The reaction was quenched with MeOH (7 mL) and the solvents were evaporated *in vacuo*. The remaining solid was co-evaporated with MePh (3 × 10 mL) and the residue was dissolved in $CH_2Cl_2$ (40 mL), washed with brine (2 × 20 mL), dried over $Na_2SO_4$, and evaporated under reduced pressure. The remaining solid was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–10%), and re-purified by column chromatography (petroleum ether(40–60)/EtOAc, 1/4; +0.5% $Et_3N$) to give DMTr-protected deoxythymidine **19** (3.82 g, 7.01 mmol, 85%).

***R*$_f$** 0.30 (petroleum ether (40–60)/EtOAc, 20/80; +1% $Et_3N$);

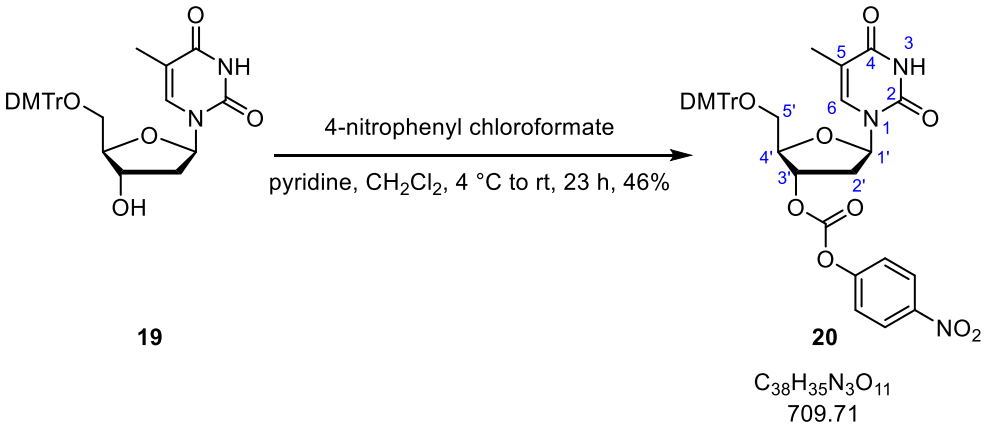
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.88 (bs, 1H, NH), 7.59 (d, *J* = 1.3 Hz, 1H, H6), 7.44 – 7.36 (m, 2H, H-Ar), 7.33 – 7.21 (m, 7H, H-Ar), 6.92 – 6.72 (m, 4H, H-Ar), 6.43 (dd, *J* = 7.9, 5.9 Hz, 1H, H1′), 4.58 (dt, *J* = 6.0, 3.0 Hz, 1H, H3′), 4.07 (q, *J* = 3.0 Hz, 1H, H4′), 3.79 (s, 6H, 2 × $OCH_3$), 3.47 (dd, *J* = 10.5, 3.0 Hz, 1H, H5′), 3.37 (dd,

*J* = 10.5, 3.0 Hz, 1H, H5″), 2.58 (bs, 1H, OH), 2.43 (ddd, *J* = 13.6, 6.0, 3.0 Hz, 1H, H2′), 2.32 (ddd, *J* = 13.6, 7.9, 6.1 Hz, 1H, H2″), 1.46 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.9 (C4), 158.8 (C-Ar), 150.5 (C2), 144.4 (C-Ar), 135.8 (C6), 135.5 (C-Ar), 130.2 (CH-Ar), 128.2 (CH-Ar), 128.1 (CH-Ar), 127.3 (CH-Ar), 113.4 (CH-Ar), 111.4 (C5), 87.1 ($C_{quat}$-*DMTr*), 86.3 (C4′), 84.8 (C1′), 72.7 (C3′), 63.7 (C5′), 55.4 ($OCH_3$), 41.1 (C2′), 12.0 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{31}H_{32}O_7N_2Na$, 567.2102; found, 567.2102.

5′-*O*-Dimethoxytrityl-3′-*O*-(4-nitrophenoxycarbonyl)thymidine (**20**)[76,178,179]




**19** **20**

$C_{38}H_{35}N_3O_{11}$
709.71

A suspension of 4-nitrophenyl chloroformate (582 mg, 2.88 mmol) in anhyd $CH_2Cl_2$ (20 mL) and anhyd pyridine (5 mL) was added at to a solution of DMTr-protected thymidine **19** (1.02 g, 1.87 mmol) in anhyd pyridine (5 mL) at 4 °C, and the reaction mixture was stirred at rt for 16 h. Another portion of 4-nitrophenyl chloroformate (265 mg, 1.32 mmol) was added and the reaction was stirred for another 7 h. The solvents were evaporated, and the crude submitted to column chromatography (petroleum ether (40–60)/EtOAc, 1/1; +1% $Et_3N$) to give activated carbonate **20** (0.61 g, 0.86 mmol, 46%).

***R*f** 0.29 (petroleum ether (40–60)/EtOAc, 1/1; +1% $Et_3N$);

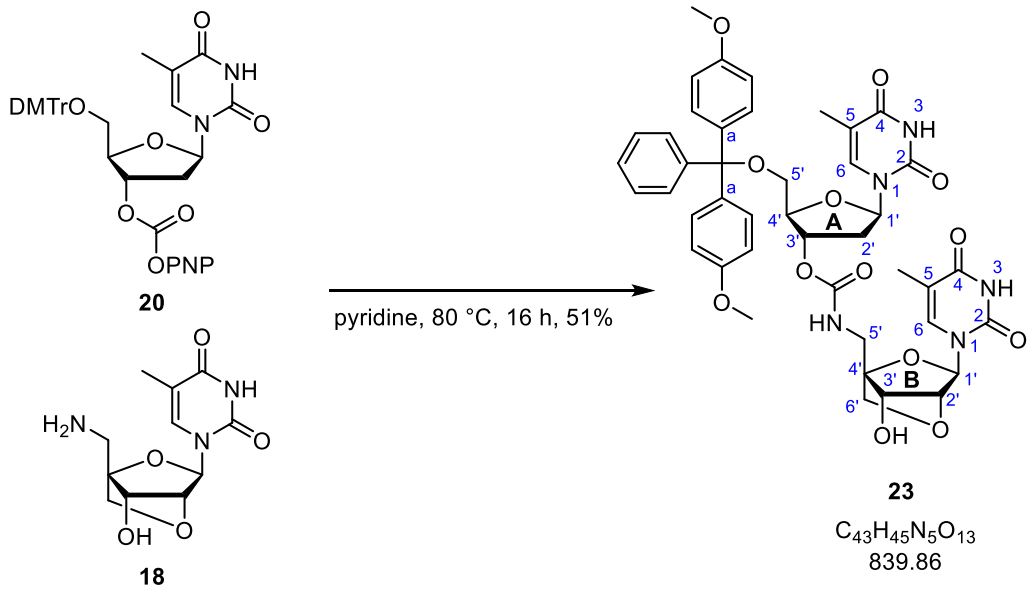
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.63 (bs, *J* = 4.4 Hz, 1H, HN), 8.36 – 8.25 (m, 2H, H-Ar), 7.62 (d, *J* = 1.3 Hz, 1H, H6), 7.42 – 7.35 (m, 4H, H-Ar), 7.35 – 7.22 (m, 7H, H-Ar and solvent), 6.85 (dd, *J* = 8.9, 1.2 Hz, 4H, H-Ar), 6.52 (dd, *J* = 9.2, 5.3 Hz, 1H, H1′), 5.46 – 5.42 (m, 1H, H3′), 4.33 – 4.30 (m, 1H, H4′), 3.79 (2 × s, 6H, 2 × $OCH_3$), 3.59 (dd, *J* = 10.7, 2.8 Hz, 1H, H5′), 3.48 (dd, *J* = 10.7, 2.4 Hz, 1H, H5″),

2.67 (ddd, *J* = 14.5, 5.3, 1.2 Hz, 1H, H2′), 2.52 (ddd, *J* = 14.5, 9.2, 6.0 Hz, 1H, H2″), 1.42 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.4 (C4), 158.9 (C-Ar), 155.1 (C-Ar *or* C=O-*carbonate*), 151.9 (C-Ar *or* C=O-*carbonate*), 150.3 (C2), 145.6 (C-Ar), 144.1 (C-Ar), 135.1 (C6), 135.0 (C-Ar), 130.1 (CH-Ar), 128.1 (CH-Ar), 128.1 (CH-Ar), 127.4 (CH-Ar), 125.4 (CH-Ar), 121.7 (CH-Ar), 113.4 (CH-Ar), 111.9 (C5), 87.4 ($C_{quat}$-*DMTr*), 84.3 (C1′), 83.5 (C4′), 80.5 (C3′), 63.7 (C5′), 55.3 ($OCH_3$), 37.9 (C2′), 11.7 ($CH_3$-*thymine*);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{38}H_{35}O_{11}N_3Na$, 732.2164; found, 732.2167.

5′-*O*-(Dimethoxytrityl)-DNA-thymidine–$CA_X$–LNA-thymidine dimer (**22**)[182]




Amine **18** (282 mg, 1.05 mmol) and carbonate **20** (750 mg, 1.06 mmol) were separately co-evaporated with anhyd pyridine (3 × 5 mL) and each dissolved in anhyd pyridine (5 mL). The solution with activated carbonate **20** was added dropwise to the solution with amine **18** at rt, and the reaction was stirred at 80 °C for 16 h. The solvent was evaporated *in vacuo* and the residue submitted to column chromatography (EtOAc/MeOH, 2–4%; +1% $Et_3N$). The obtained white solid was re-dissolved in $CHCl_3$/$^i$PrOH [3/2 (v/v), 50 mL] and washed with sat. aq $NaHCO_3$ until the aq layer stayed colourless. The org layer was dried over $Na_2SO_4$, and filtered over celite. Evaporation of the solvents gave DNA–$CA_X$–LNA dimer **23** as a white solid (452 mg, 0.54 mmol, 51%).

***R*f** 0.38 ($CH_2Cl_2$/MeOH, 9/1);

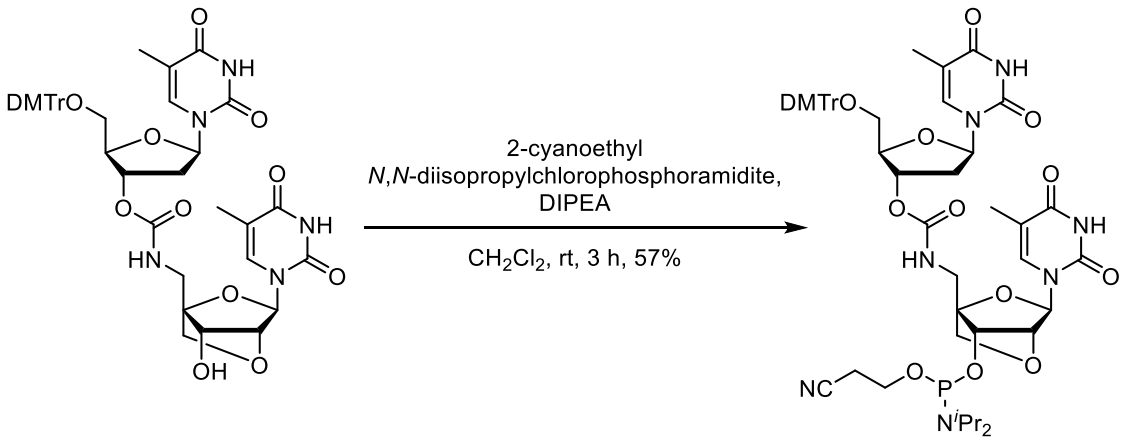
**$^{1}$H NMR** (400 MHz, DMSO-$d_6$) δ 11.42 (s, 1H, NH-*thymine*), 11.40 (s, 1H, NH-*thymine*), 7.82 (t, *J* = 6.1 Hz, 1H, NH-*carbamate*), 7.55 (d, *J* = 1.1 Hz, 1H, H6), 7.45 (d, *J* = 1.1 Hz, 1H, H6), 7.43 – 7.23 (m, 9H, H-Ar), 6.97 – 6.87 (m, 4H, H-Ar), 6.28 (dd, *J* = 8.6, 5.9 Hz, 1H, H1′A), 5.59 (d, *J* = 4.1 Hz, 1H, OH), 5.44 (s, 1H, H1′B), 5.36 – 5.24 (m, 1H, H3′A), 4.15 (s, 1H, H2′B), 4.12 – 4.03 (m, 1H, H4′A), 3.92 – 3.85 (m, 2H, H6′B *and* H3′B), 3.76 (s, 6H, 2 × $OCH_3$), 3.65 (d, *J* = 8.0 Hz, 1H, H6′′B), 3.61 – 3.53 (m, 2H, H5′B *and* H5′′B), 3.36 – 3.32 (m, 1H, H5′A), 3.27 – 3.21 (m, 1H, H5′′A), 2.51 – 2.43 (m, 1H, H2′A), 2.37 – 2.27 (m, 1H, H2′′A), 1.79 (d, *J* = 1.1 Hz, 3H, $CH_3$-*thymine*), 1.43 (d, *J* = 1.1 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$) δ 163.8 (C4), 163.6 (C4), 158.2 (C-Ar), 155.8 (C=O-*carbamate*), 150.4 (C2), 149.9 (C2), 144.6 (C-Ar), 135.4 (C6), 135.3([a]C-Ar), 135.1 ([a]C-Ar), 134.6 (C6), 129.7 (CH-Ar), 128.0 (CH-Ar), 127.7 (CH-Ar), 126.9 (CH-Ar), 113.3 (CH-Ar), 109.9 (C5), 108.5 (C5), 87.1 (C4′B *or* $C_{quat}$-*DMTr*), 86.3 (C4′B *or* C1′B *or* $C_{quat}$-*DMTr*), 86.1 (C4′B *or* C1′B *or* $C_{quat}$-*DMTr*), 83.6 (C1′A), 83.3 (C4′A), 79.0 (C2′B), 74.7 (C3′A), 71.4 (C6′B), 69.9 (C3′B), 63.7 (C5′A), 55.1 ($OCH_3$), 37.2 (C5′B), 36.7 (C2′A), 12.4 ($CH_3$-*thymine*), 11.6 ($CH_3$-*thymine*); [a]Note that the marked carbons can resolve in distinct peaks due to different environments.

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{43}H_{45}O_{13}N_3Na$, 862.2906; found, 862.2906.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

5′-*O*-(Dimethoxytrityl)-DNA-thymidine–$CA_X$–LNA-thymidine dimer CE phosphoramidite (**8**)[182]




**23** **8**

$C_{52}H_{62}N_7O_{14}P$
1040.08

2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.18 mL, 0.81 mmol) was added dropwise over 10 min to a solution of degassed anhyd DIPEA (0.25 mL, 1.44 mmol) and DNA–$CA_X$–LNA dimer **23** (480 mg, 0.57 mmol) in degassed anhyd $CH_2Cl_2$ (5.0 mL), and the reaction was stirred at rt for 1 h. More degassed anhyd DIPEA (0.12 mL, 0.69 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.12 mL, 0.54 mmol) were added, and the reaction was stirred at rt for another 2 h. The reaction mixture was then diluted with degassed anhyd $CH_2Cl_2$ (10 mL), and washed with degassed, sat. aq KCl (10 mL) under an argon gas atmosphere. The org layer was dried over $Na_2SO_4$, and the solvents were evaporated under reduced pressure. The remaining residue was dissolved in a minimum amount of degassed $CH_2Cl_2$ (2.5 mL), and the product was then precipitated by addition of hexanes (250 mL). The supernatant was cooled to 4 °C, decanted, and the solid dried *in vacuo*. Purification by column chromatography

using degassed solvents (hexanes/EtOAc, 1/10; +1% pyridine) under an argon gas atmosphere gave DNA–$CA_X$–LNA phosphoramidite **8** (338 mg, 57%).

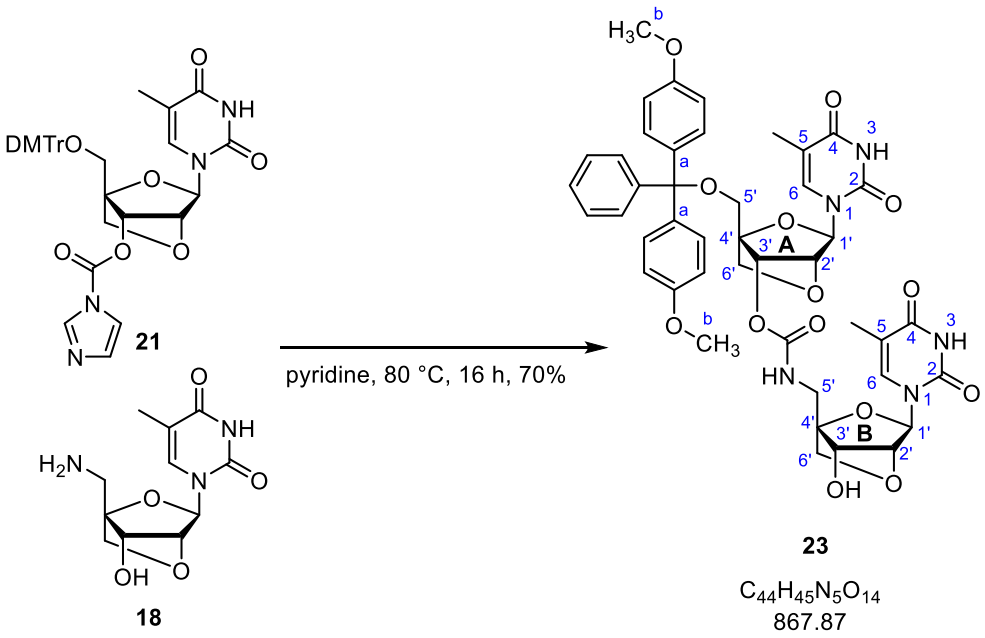
***R*f** 0.51 (hexanes/EtOAc, 1/10; +1% pyridine);

**$^{31}$P NMR** (162 MHz, $CDCl_3$) δ 149.5, 149.1;

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{52}H_{63}O_{14}N_7P$, 1040.4165; found, 1040.4183.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

## 5′-*O*-(Dimethoxytrityl)-LNA-thymidine–CA$_X$–LNA-thymidine dimer (**23**)[182]




Amine **18** (202 mg, 0.75 mmol) and carbonate **21**[t] (500 mg, 0.75 mmol) were separately co-evaporated with anhyd pyridine (2 × 4 mL) and each dissolved in anhyd pyridine (5 mL). The solution with activated carbonate **21** was added dropwise to the solution with amine **18** at rt, and the reaction was stirred at 80 °C for 16 h. A catalytic amount of DMAP (9.0 mg, 0.08 mmol) was added and the mixture was stirred at 80 °C for another 24 h. The solvent was evaporated *in vacuo* and the residue re-dissolved in $CH_2Cl_2$ (70 mL), washed with a solution of sat. aq $NaHCO_3$ (2 × 70 mL) and brine (1 × 50 mL). The separated org layer was dried over $MgSO_4$, filtered and the solvent was evaporated. The residue was submitted to column chromatography (EtOAc/MeOH 0–6%; +1% $Et_3N$) to give LNA–CA$_X$–LNA dimer **23** as a white solid (453 mg, 0.52 mmol, 70%).

[t] Prepared by Benjamin Woods (Part II), University of Oxford.

***R*f** 0.41 (EtOAc/MeOH, 24/1; +1% $Et_3N$);

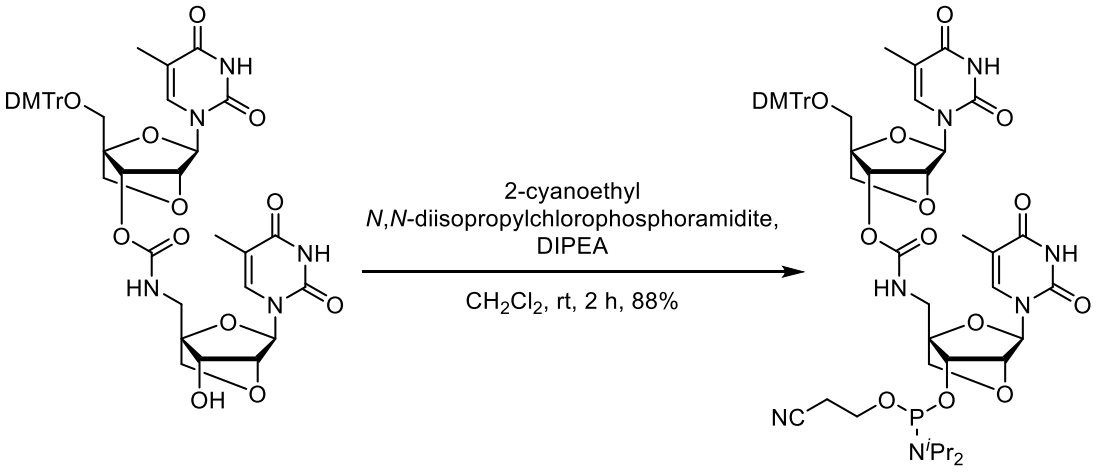
**$^1$H NMR** (400 MHz, DMSO-*d*6) δ 11.50 (s, 1H, NH-*thymine*), 11.39 (s, 1H, NH-*thymine*), 7.91 (t, *J* = 6.1 Hz, 1H, NH-*carbamate*), 7.63 – 7.48 (m, 1H, H6), 7.45 – 7.17 (m, 10H, 9 × H-Ar and H6), 6.96 – 6.81 (m, 4H, H-Ar), 5.62 (d, *J* = 4.1 Hz, 1H, OH), 5.53 (s, 1H, H1′A *or* H1′B), 5.39 (s, 1H, H1′A *or* H1′B), 5.18 (s, 1H, H3′A), 4.47 (s, 1H, H2′A), 4.11 (s, 1H, H2′B), 3.91 – 3.76 (m, 5H, H6′A, H6′′A, H6′B, H6′′B *and* H3′B), 3.72 (2 × s, 6H, 2 × $OCH_3$), 3.65 – 3.47 (m, 2H, H5′A *or* H5′B), 3.47 – 3.38 (m, 2H, H5′A *or* H5′B), 1.72 (d, J = 1.1 Hz, 3H, $CH_3$-*thymine*), 1.55 (d, J = 1.1, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$) δ 163.8 (C4), 163.8 (C4), 158.2 (C-Ar), 155.0 (C=O-*carbamate or* C2), 150.0 (C=O-*carbamate or* C2), 149.9 (C=O-*carbamate or* C2), 144.6 (C-Ar), 135.0 ($^a$C-Ar *or* C6), 134.9 ($^a$C-Ar *or* C6), 134.8 ($^a$C-Ar *or* C6), 133.8 (C6), 129.7 (CH-Ar), 128.0 (CH-Ar), 127.6 (CH-Ar), 126.9 (CH-Ar), 113.3 (CH-Ar), 109.0 (C5), 108.5 (C5), 87.1 (C4′A *or* C4′B *or* $C_{quat}$-*DMTr*), 86.4 (C1′A *or* C1′B *or* C4′A *or* C4′B *or* $C_{quat}$-*DMTr*), 86.3 (C1′A *or* C1′B *or* C4′A *or* C4′B *or* $C_{quat}$-*DMTr*), 85.9 (C4′A *or* C4′B *or* $C_{quat}$-*DMTr*), 79.1 (C2′B), 77.5 (C2′A), 71.7 (C6′A *or* C6′B), 71.4 (C6′A *or* C6′B), 70.9 (C3′A), 70.2 (C3′B), 58.0(C5′A), 55.1 ($O^bCH_3$), 55.0 ($O^bCH_3$), 37.8 (C5′B), 12.4 ($CH_3$-*thymine*), 12.2 ($CH_3$-*thymine*); $^{a,b}$Note that the marked carbons can resolve in distinct peaks due to different environments.

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{44}H_{45}O_{14}N_5Na$, 890.2855; found, 890.2857.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

5′-*O*-(Dimethoxytrityl)-LNA-thymidine–$CA_X$–LNA-thymidine dimer CE phosphoramidite (**9**)[182]




**23** **9**

$C_{53}H_{62}N_7O_{15}P$
1068.09

2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.12 mL, 0.54 mmol) was added dropwise over 10 min to a solution of degassed anhyd DIPEA (0.17 mL, 0.95 mmol) and LNA–$CA_X$–LNA dimer **23** (334 mg, 0.38 mmol) in degassed anhyd $CH_2Cl_2$ (3.5 mL), and the reaction was stirred at rt for 1 h. More 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.1 mL, 0.45 mmol) was added, and the reaction was stirred at rt for another 1 h. The reaction mixture was then diluted with degassed anhyd $CH_2Cl_2$ (10 mL), and washed with degassed, sat. aq KCl (10 mL) under an argon gas atmosphere. The org layer was dried over $Na_2SO_4$, and the solvents were evaporated under reduced pressure. The solid was re-dissolved in a minimum amount of degassed anhyd $CH_2Cl_2$ (2.5 mL) and precipitated by addition of degassed anhyd hexanes (150 mL). The supernatant was decanted, and the remaining slurry was dried *in vacuo* to give a white solid. Precipitation was repeated once more and successive purification by column chromatography using degassed

solvents (hexanes/EtOAc, 3/20; +1% pyridine) under an argon gas atmosphere gave LNA–$CA_X$–LNA phosphoramidite **9** (356 mg, 88%).

***R*f** 0.57 (hexanes/EtOAc, 1/10; +1% pyridine);

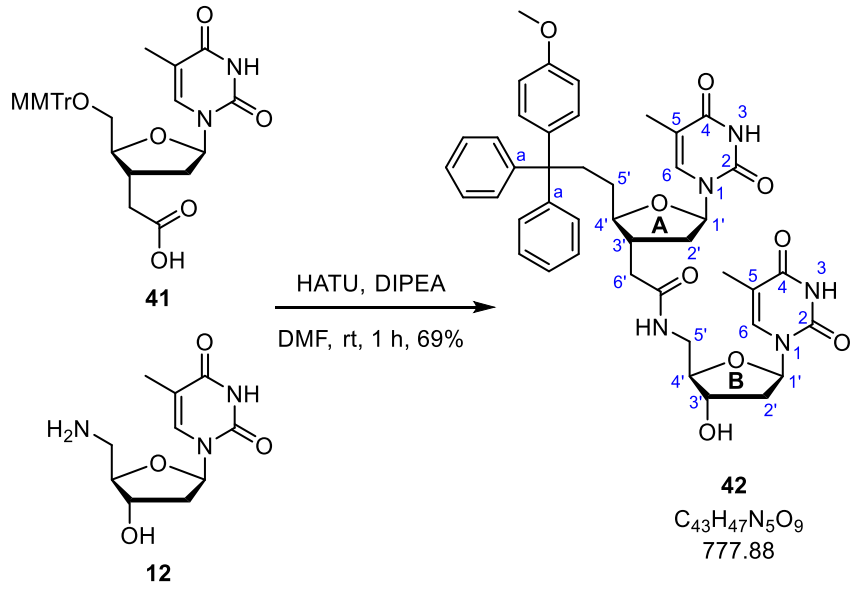
**$^{31}$P NMR** (162 MHz, $CDCl_3$) δ 148.1, 147.5;

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{53}H_{63}O_{15}N_7P$, 1068.4114; found, 1068.4109.

*Experimental procedure and characterisation reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

## 4.3. Compounds of Chapter III

5′-*O*-(Methoxytrityl)-dT–AM–dT dimer (**42**)[308]




Anhyd DIPEA (0.25 mL, 1.44 mmol, pre-dried over activated 4 Å molecular sieves) an HATU (271 mg, 0.72 mmol) were added to a mixture of carboxylic acid **41**[u] (389 mg, 0.72 mmol) and activated 4 Å molecular sieves (200 mg) in anhyd DMF (5 mL), and the mixture was stirred at rt for 20 min. A suspension of amine **12**[176] [189 mg, 0.78 mmol, pre-incubated with activated 4 Å molecular sieves (100 mg) in anhyd DMF (4 mL) for 10 min] was added and the reaction was stirred at rt for 1 h. The resulting yellow suspension was evaporated *in vacuo* and re-suspended in $CH_2Cl_2$/pyridine [99/1 (v/v), 75 mL] and filtered through celite. The filtrate was washed with sat. aq $NaHCO_3$ (50 mL), and the aq layer was re-extracted with $CH_2Cl_2$ (3 × 30 mL). The combined org phases were dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. The residue was submitted to column chromatography ($CH_2Cl_2$/MeOH, 0–10%; +0.5% $Et_3N$). The resulting white solid

[u] Prepared by Cameron Thorpe (DPhil), University of Oxford.

was re-dissolved in EtOAc (70 mL), and washed with sat. aq $NaHCO_3$ (50 mL) and brine (50 mL) to remove remaining $Et_3N$. The org phase was dried over $Na_2SO_4$, filtered, and evaporated to give dimer **42** (387 mg, 0.50 mmol, 69%) as a white solid.

***R*f** 0.25 ($CH_2Cl_2$/MeOH, 9/1);

**IR** ($cm^{-1}$): 3692, 3026, 2923 (O–H), 2840, 1684 (C=O);

**$^1$H NMR** (400 MHz, DMSO-$d_6$) δ 11.28 (bs, 2H, 2 × NH-*thymine*), 8.09 (t, *J* = 5.8 Hz, 1H, NH-*amide*), 7.49 (d, *J* = 1.3 Hz, 1H, H6), 7.47 (d, *J* = 1.3 Hz, 1H, H6), 7.43 – 7.35 (m, 4H, H-Ar), 7.35 – 7.28 (m, 4H, H-Ar), 7.28 – 7.20 (m, 4H, H-Ar), 6.92 – 6.85 (m, 2H, H-Ar), 6.11 (dd, *J* = 7.7, 6.3 Hz, 1H, H1′B), 6.02 (dd, *J* = 7.0, 4.2 Hz, 1H, H1′A), 5.29 (d, *J* = 4.2 Hz, 1H, OH), 4.15 – 4.06 (m, 1H, H3′B), 3.82 (dt, *J* = 7.7, 3.6 Hz, 1H, H4′A), 3.77 – 3.67 (m, 4H, $OCH_3$ *and* H4′B), 3.29 – 3.14 (m, 4H, H5′A, H5″A, H5′B *and* H5″B), 2.77 – 2.62 (m, 1H, H3′A), 2.26 – 2.13 (m, 3H, H6′A, H6″A *and* H2′A), 2.13 – 1.99 (m, 3H, H2″A, H2′B *and* H2″B), 1.77 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*), 1.45 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (126 MHz, DMSO-$d_6$) δ = 170.6 (C=O-*amide*), 163.7 (C4), 163.7 (C4), 158.3 (C-Ar), 150.4 (C2), 150.4 (C2), 144.2 ($^a$C-Ar), 144.1 ($^a$C-Ar), 136.1 (C6A *or* C6B), 135.6 (C6A *or* C6B), 134.8 (C-Ar), 130.1 (CH-Ar), 128.0 (CH-Ar), 127.9 (CH-Ar), 127.0 (CH-Ar), 113.3 (CH-Ar), 109.7 (C5), 109.2 (C5), 86.0 ($C_{quat}$-*MMTr*), 84.9 (C4′B), 83.9 (C1′A *and* C1′B), 83.4 (C4′A), 71.3 (C3′B), 63.7 (C5′A), 55.1 ($OCH_3$), 41.0 (C5′B), 38.5 (C2′B), 37.7 (C6′), 37.4 (C2′A), 35.2 (C3′A), 12.0

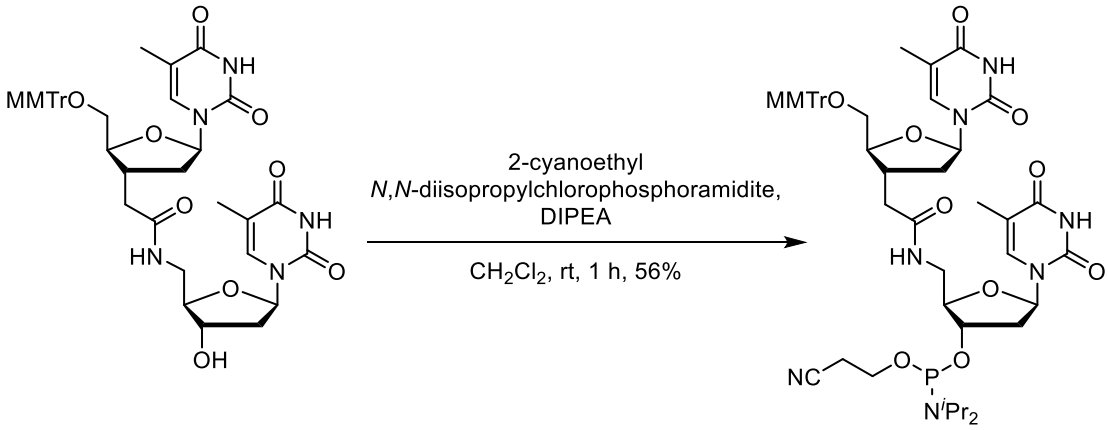
($CH_3$-*thymine*), 11.9 ($CH_3$-*thymine*); [a]Note that the marked carbons can resolve in distinct peaks due to different environments.

***m/z* (ESI$^+$)**: 273.145 ($[MMTr]^+$, 100%), 802.326 ($[M + Na]^+$, 83.3%), 1581.633 ($[2M + Na]^+$, 27.1%);

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{42}H_{45}N_5O_{10}Na$, 802.3059; found, 802.3062.

*Experimental procedure and characterisation reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 5′-*O*-(Methoxytrityl)-dT–AM–dT dimer CE phosphoramidite (**43**)[308]




**42**

**43**
$C_{51}H_{62}N_7O_{11}P$
980.07

2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.12 mL, 0.54 mmol) was added dropwise over 10 min to a solution of degassed anhyd DIPEA (0.20 mL, 1.15 mmol) and dimer **42** (359 mg, 0.46 mmol) in degassed anhyd $CH_2Cl_2$ (4 mL), and the reaction was stirred at rt for 1 h. The reaction mixture was then diluted with degassed anhyd $CH_2Cl_2$ (3 mL), and washed with degassed, sat. aq KCl (10 mL) under an argon gas atmosphere. The aq layer was extracted with degassed $CH_2Cl_2$ (2 × 5 mL) and the combined organic layers were dried over $Na_2SO_4$, filtered and the solvents evaporated under reduced pressure. The remaining residue was dissolved in a minimum amount of degassed $CH_2Cl_2$ (2.5 mL) and the product was then precipitated by addition of petroleum ether (40–60) (70 mL). The supernatant was decanted and the solid dried *in vacuo*. The precipitation was repeated as described above to give a white solid. Column chromatography using degassed anhyd solvents ($CH_2Cl_2$/MeCN, 50% with a solvent switch to EtOAc/MeCN, 30%)

under an argon gas atmosphere gave phosphoramidite **43** (254 mg, 0.26 mmol, 56%) as a white solid.

***R*f** 0.39 ($CH_2Cl_2$/MeOH, 19/1);

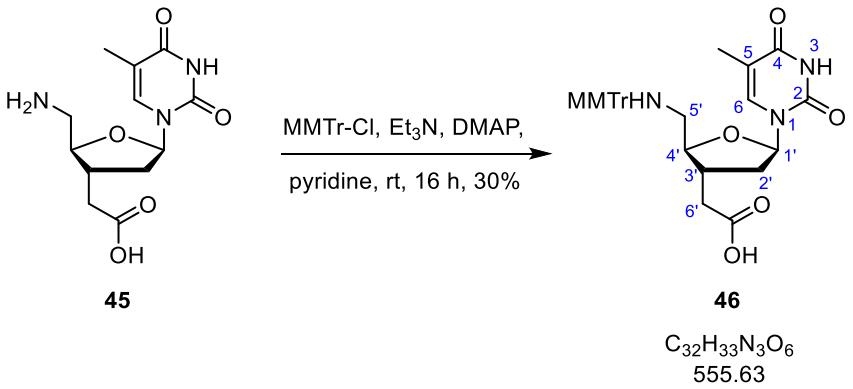
**$^{31}$P NMR** (162 MHz, $CDCl_3$) δ 149.4, 149.0;

***m/z* (ESI$^+$):** 273.133 ($[MMTr]^+$, 100%), 980.427 ($[M + H]^+$, 15.7%), 1002.457 ($[M + Na]^+$, 42.5%);

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{51}H_{63}N_7O_{11}P$, 980.4318; found, 980.4319.

*Experimental procedure and characterisation reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

### 3′-Carboxymethyl-2′,5′-dideoxy-5′-(monomethoxytritylamino)thymidine (**46**)[55]




**45**

**46**

$C_{32}H_{33}N_3O_6$
555.63

Amino acid **45**[v] (136 mg, 0.48 mmol) was co-evaporated with anhyd pyridine (5 mL) and dissolved in anhyd pyridine (4.8 mL). Anhyd $Et_3N$ (0.97 mL, 0.70 mmol), DMAP (6 mg, 0.05 mmol) and 4-methoxytriphenylmethyl chloride (331 mg, 1.07 mmol) were added and the reaction mixture was stirred at rt for 16 h. The solvents were then evaporated *in vacuo* and the residue was re-dissolved in $CH_2Cl_2$ (50 mL) and successively washed with sat. aq $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (30 mL). The organic phase was dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. Purification of the residue by column chromatograph ($CH_2Cl_2$/MeOH, 0–10%; +1% pyridine) gave monomer **46** (79 mg, 0.14 mmol, 30%) as a white solid.

***R*f** 0.29 ($CH_2Cl_2$/MeOH, 9/1; +1% pyridine)

**IR** ($cm^{-1}$): 2924 (O–H), 2854, 1688 (C=O), 1509, 1465, 1265 (C–O-*acid*), 1181;

**$^1$H NMR** (400 MHz, $CDCl_3$) δ 9.47 (bs, 1H, NH-*thymine*), 7.49 – 7.41 (m, 4H, H-Ar), 7.39 – 7.33 (m, 2H, H-Ar), 7.31 – 7.24 (m, 4H, H-Ar *and* solvent), 7.23 (d, *J* = 1.2 Hz, 1H, H6), 7.21 – 7.15 (m, 2H, H-Ar), 6.87 – 6.74 (m, 2H, H-Ar), 6.08 (dd, *J* = 7.2, 2.9 Hz, 1H, H1′), 3.85 – 3.80 (m, 1H, H4′), 3.77 (s, 3H, $OCH_3$), 2.63 (dd, *J* = 13.0, 2.9 Hz, 1H, H5′), 2.53 – 2.40 (m, 2H, H3′, N**H**–*MMTr*), 2.36 – 2.22 (m, 4H,

[v] Prepared by Cameron Thorpe (DPhil), University of Oxford.

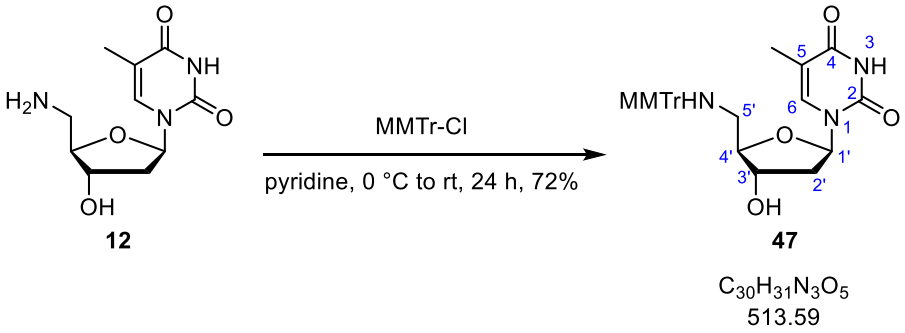
H2′, H5″, H6′, H6″), 2.22 – 2.10 (m, 1H, H2″), 1.82 (d, *J* = 1.2 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 175.7 (C=O-*acid*), 164.4 (C4), 158.2 (C-Ar), 150.5 (C2), 145.9 (C-Ar), 137.7 (C-Ar), 135.7 (C6), 129.9 (CH-Ar), 128.6 (CH-Ar), 128.1 (CH-Ar), 126.6 (CH-Ar), 113.4 (CH-Ar), 110.8 (C5), 85.2 (C4′), 85.1 (C1′), 70.4 ($C_{quat}$-*MMTr*), 55.3 ($OCH_3$), 46.0 (C5′), 39.3 (C2′), 36.2 (C6′), 36.1 (C3′), 12.7 ($CH_3$-*thymine*);

***m/z* (ESI$^+$)**: 273.464 ($[MMTr]^+$, 18.8%), 556.937 ($[M + H]^+$, 100%).

## 4.4. Compounds of Chapter IV

2′,5′-Dideoxy-5′-(monomethoxytritylamino)thymidine (**47**)[244]




A solution of 4-methoxytriphenylmethyl chloride (128 mg, 0.42 mmol) in anhyd pyridine (1.3 mL) was added dropwise to a stirred solution of amine **12**[176] (100 mg, 0.42 mmol) in anhyd pyridine (2.7 mL) at 4 °C. The reaction mixture then stirred at rt for 24 h. The solvent was evaporated *in vacuo*, the residue was re-dissolved in $CH_2Cl_2$ (50 mL), and washed with $H_2O$ (40 mL). The aq layer was re-extracted with $CH_2Cl_2$ (20 mL), and the combined org layers were dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. Column chromatography ($CH_2Cl_2$/MeOH, 0–5%) gave MMTr-protected amine **47** (154 mg, 0.30 mmol, 72%).

***R*f** 0.16 ($CH_2Cl_2$/MeOH, 97:3);

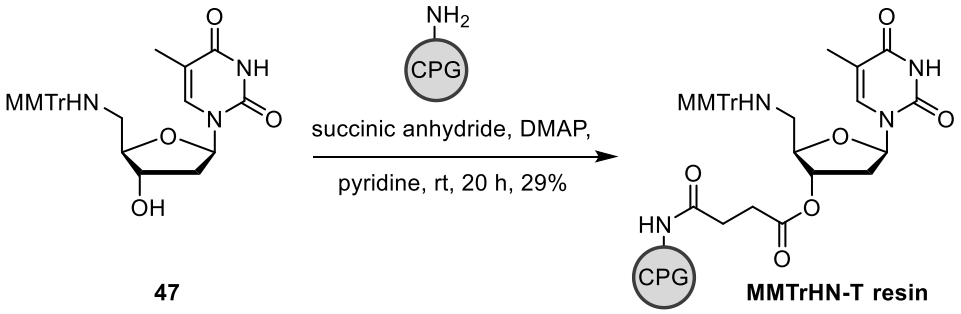
**IR** ($cm^{-1}$): 3426 (O–H), 3055 (CH-Ar), 2925, 2851, 2360, 2341, 1685 (C=O), 1508, 1466, 1250, 1181;

**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.46 (s, 1H, NH-*thymine*), 7.49 – 7.42 (m, 4H, H-Ar), 7.38 – 7.33 (m, 2H, H-Ar), 7.31 – 7.25 (m, 4H, H-Ar), 7.23 – 7.17 (m, 2H, H-Ar), 7.06 (d, *J* = 1.3 Hz, 1H, H6), 6.89 – 6.75 (m, 2H, H-Ar), 6.27 (t, *J* = 6.7 Hz, 1H, H1′), 5.30 (s, 1H, N**H**–*MMTr or* OH), 4.32 (dt, *J* = 7.0, 4.6 Hz, 1H, H3′), 3.98 (dt, *J* = 6.5, 4.6 Hz, 1H, H4′), 3.78 (s, 3H, $OCH_3$), 2.58 (dd, *J* = 12.4, 4.6 Hz, 1H, H5′), 2.42 – 2.22 (m, 3H, H5″, H2′, OH *or* N**H**–*MMTr*), 2.11 (dt, *J* = 13.7, 6.7 Hz, 1H, H2″), 1.84 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.4 (C4), 158.2 (C-Ar), 150.2 (C2), 145.9 (C-Ar), 137.7 (C-Ar), 135.2 (C6), 129.9 (CH-Ar), 128.6 (CH-Ar), 128.1 (CH-Ar), 126.7 (CH-Ar), 113.4 (CH-Ar), 111.5 (C5), 86.0 (C4′), 84.4 (C1′), 72.8 (C3′), 70.4 ($C_{quat}$-*MMTr*), 55.4 ($OCH_3$), 46.1 (C5′), 40.4 (C2′), 12.8 ($CH_3$-*thymine*);

***m/z* (ESI$^+$)**: 273.133 ($[MMTr]^+$, 100%), 536.2 ($[M + Na]^+$, 0.73%);

Functionalisation of CPG resin (MMTrHN-T resin)[244,309,310]



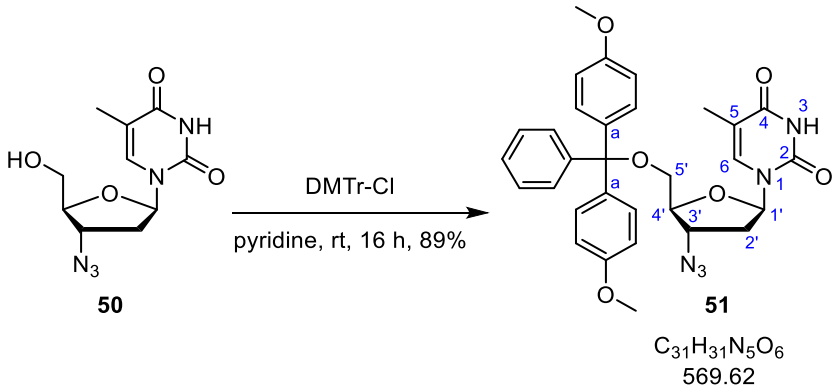

Long chain alkylamine controlled pore glass (LCAA CPG) 1000/110 (*LGC Genomics Ltd.*) resin (1.50 g, loading 81.1 μmol·g$^{-1}$) was rotated at rt for 1 h in a glass vessel containing a solution of 3% (w/v) trichloroacetic acid in anhyd $CH_2Cl_2$ (15 mL). The resin was filtered and consecutively washed with $Et_3N$/DIPEA [9/1 (v/v), 10 mL], anhyd $CH_2Cl_2$ (10 mL) and $Et_2O$ (10 mL) and dried *in vacuo*. The resin was soaked in freshly distilled anhyd pyridine (4 mL), a solution of succinic anhydride (750 mg, 7.50 mmol) and 4-dimethylaminopyridine (DMAP) (75 mg, 0.62 mmol) in freshly distilled pyridine (15 mL) was added, and the reaction vessel was rotated at rt for 20 h. The solvent was removed by filtration and the resin was consecutively washed with anhyd pyridine (10 mL), anhyd $CH_2Cl_2$ (10 mL) and $Et_2O$ (10 mL) and dried *in vacuo*. The resin was soaked with anhyd pyridine (4 mL) and a solution of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) (386 mg, 2.01 mmol), $Et_3N$ (78 μL, 0.56 mmol) and DMAP (15 mg, 0.12 mmol) in anhyd pyridine (5 mL) was added followed by addition of nucleoside **47**[244] (70 mg, 0.14 mmol) in anhyd pyridine (2 mL). The reaction mixture was gently agitated at rt for 24 h, pentachlorophenol (45 mg, 0.17 mmol) was then added, and the mixture was rotated at rt for another 1 h before filtration and successive washing with anhyd pyridine (3 × 5 mL), anhyd $CH_2Cl_2$ (3 × 5 mL),

and $Et_2O$ (3 × 5 mL). Piperidine/anhyd DMF [1/9 (v/v), 5 mL] was added to the dried resin and the vessel was rotated at rt for 10 min. Afterwards, the resin was filtered and consecutively washed with anhyd $CH_2Cl_2$ (3 × 5 mL) and $Et_2O$ (3 × 5 mL). Oligonucleotide synthesis grade capping reagent (*Applied Biosystems*) $Ac_2O$/pyridine/THF/*N*-methyl imidazole in THF [1/1 (v/v), 10 mL] was added and the vessel was rotated for 1 h. The resin was filtered and consecutively washed with anhyd $CH_2Cl_2$ (3 × 5 mL) and $Et_2O$ (3 × 5 mL) and dried *in vacuo* to give the functionalised resin (MMTrHN-T resin) as a white solid (1.46 g, loading 23.6 $\mu mol \cdot g^{-1}$, 29%). Loading efficiency was determined by measuring $A_{478}$ after MMTr-deprotection in a solution of 3% (w/v) trichloroacetic acid in anhyd $CH_2Cl_2$ applying an extinction coefficient of $\varepsilon = 56\ mL \cdot cm^{-1} \cdot \mu mol^{-1}$ for the MMTr-cation.

*Experimental procedure reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

## 4.5. Compounds of Chapter V

5′-*O*-(4′′-Dimethoxytriphenylmethyl)-3′-azido-3′-deoxythymidine (**51**)[297]




4,4'-Dimethoxytrityl chloride (9.52 g, 28.1 mmol) was added to a solution of zidovuidine (**50**) (5.00 g, 18.7 mmol) in anhyd pyridine (70 mL), and the reaction was stirred at rt for 16 h. The solvent was evaporated *in vacuo*, the residue dissolved in $CH_2Cl_2$ (200 mL), and washed with $H_2O$ (2 × 100 mL). The organic layer was dried over $Na_2SO_4$, filtered, and evaporated. Column chromatography (petroleum ether (40–60)/EtOAc, 0–80%) gave monomer **51** as a yellow foam (9.48 g, 16.6 mmol, 89%).

***R*f** 0.60 (petroleum ether (40–60)/EtOAc, 2/3);

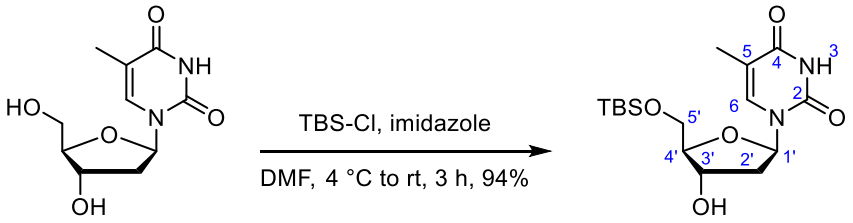
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.96 (s, 1H, NH), 7.54 (d, *J* = 1.2 Hz, 1H, H6), 7.39 – 7.31 (m, 2H, H-Ar), 7.30 – 7.16 (m, 7H, H-Ar), 6.84 – 6.76 (m, 4H, H-Ar), 6.22 (t, *J* = 6.4 Hz, 1H, H1′), 4.30 (dt, *J* = 7.0, 4.6 Hz, 1H, H3′), 3.93 (dt, *J* = 4.6, 2.8 Hz, 1H, H4′), 3.75 (s, 6H, 2 × $OCH_3$), 3.52 (dd, *J* = 10.9, 2.8 Hz, 1H, H5′), 3.28 (dd, *J*

= 10.9, 2.8 Hz, 1H, H5″), 2.48 – 2.31 (m, 2H, H2′ *and* H2″), 1.45 (d, *J* = 1.2 Hz, 3H, $CH_3$-*thymine*).

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.8 (C4), 158.9 (C-Ar), 150.4 (C2), 144.3 (C-Ar), 135.3 (C6), 135.3 ([a]C-Ar), 135.3 ([a]C-Ar), 130.2 (CH-Ar), 128.2 (CH-Ar), 127.4 (CH-Ar), 113.5 (CH-Ar), 113.5 (CH-Ar), 111.5 (C5), 87.2 ($C_{quat}$-*DMTr*), 84.6 (C1′), 83.6 (C4′), 63.0 (C5′), 60.7 (C3′), 55.4 ($OCH_3$), 38.1 (C2′), 12.0 ($CH_3$-*thymine*); [a]Note that the marked carbons can resolve in distinct peaks due to different environments.

***m/z* (ESI$^-$)**: 568.1 ([M – H]$^-$, 100%).

5′-*O*-(tert-Butyldimethylsilyl)-thymidine (**54**)[299]




**9**

**54**

$C_{16}H_{28}N_2O_5Si$

356.49

Imidazole (4.44 g, 63.7 mmol) was added to a stirred solution of thymidine (**9**) (7.00 g, 28.9 mmol) in anhyd DMF (100 mL) at 4 °C and stirred for 30 min. A solution of *tert*-butyldimethylsilyl chloride (4.79 g, 31.7 mmol) in anhyd DMF (40 mL) was slowly added over 1 h and the reaction mixture was then stirred at rt for another 3 h. The solvent was evaporated *in vacuo*, the residue dissolved in $CH_2Cl_2$ (200 mL), and washed with $H_2O$ (200 mL). The organic layer was dried over $Na_2SO_4$, filtered, and evaporated. Column chromatography (petroleum ether (40–60)/EtOAc, 0–100%) gave silylated thymidine **54** (9.68 g, 27.2 mmol, 94%) as a white foam.

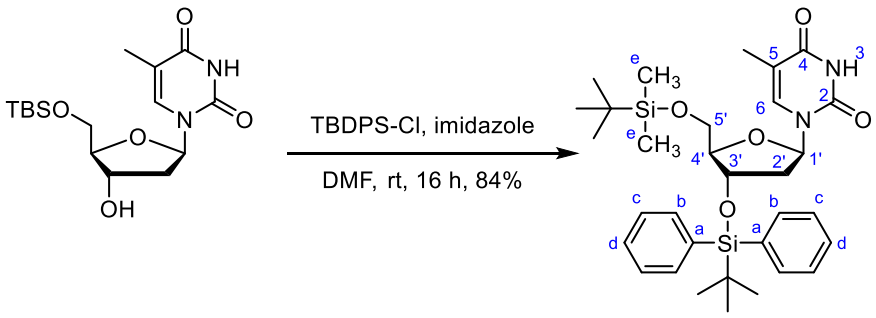
***R*$_f$** 0.40 ($CH_2Cl_2$/MeOH, 19/1);

**$^1$H NMR** (400 MHz, $CDCl_3$) δ 9.43 (s, 1H, NH), 7.53 (q, *J* = 1.2 Hz, 1H, H6), 6.39 (dd, *J* = 8.3, 5.6 Hz, 1H, H1′), 4.44 (ddt, *J* = 6.5, 4.5, 2.5 Hz, 1H, H3′), 4.06 (q, *J* = 2.5 Hz, 1H, H4′), 3.89 (dd, *J* = 11.3, 2.5 Hz, 1H, H5′), 3.82 (dd, *J* = 11.3, 2.5 Hz, 1H, H5″), 3.19 (d, *J* = 4.5 Hz, 1H, OH), 2.39 (ddd, *J* = 13.3, 5.6, 2.5 Hz, 1H, H2′), 2.17 – 1.99 (m, 1H, H2″), 1.90 (d, *J* = 1.2 Hz, 3H, $CH_3$-*thymine*), 0.91 (s, 9H, 3 × $CH_3$-*tert butyl*), 0.10 (2 × s, 6H, $Si(CH_3)_2$);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 164.1 (C4), 150.7 (C2), 135.6 (C6), 111.1 (C5), 87.5 (C4′), 85.2 (C1′), 72.7 (C3′), 63.8 (C5′), 41.3 (C2′), 26.0 (C(**C**$H_3$)$_3$), 18.5 (**C**($CH_3$)$_3$), 12.7 ($CH_3$-*thymine*), –5.2 (Si(**C**$H_3$)$_2$);

***m/z* (ESI$^+$)**: 379.2 ([M + Na]$^+$, 100%).

3′-*O*-(*tert*-Butyldiphenylsilyl)-5′-*O*-(*tert*-butyldimethylsilyl)-thymidine (**55**)[299]




**54** **55**

$C_{32}H_{46}N_2O_5Si_2$
594.90

Imidazole (4.95 g, 72.7 mmol) was added to a stirred solution of silylated thymidine **54** (9.60 g, 27.2 mmol) in anhyd DMF (100 mL) and the solution was stirred at rt for 30 min. *tert*-Butyl(chloro)diphenylsilane (10.0 g, 36.4 mmol) was added and the reaction mixture was stirred at rt for 16 h. The solvent was evaporated *in vacuo*, the residue dissolved in EtOAc (150 mL), and washed with $H_2O$ (150 mL). The organic layer was dried over $Na_2SO_4$, filtered, and evaporated. Column chromatography (petroleum ether (40–60)/EtOAc, 0–30% gave disilylated thymidine **55** as a white solid (13.6 g, 22.9 mmol, 84%).

***R*f** 0.84 (petroleum ether (40–60)/EtOAc, 3/2);

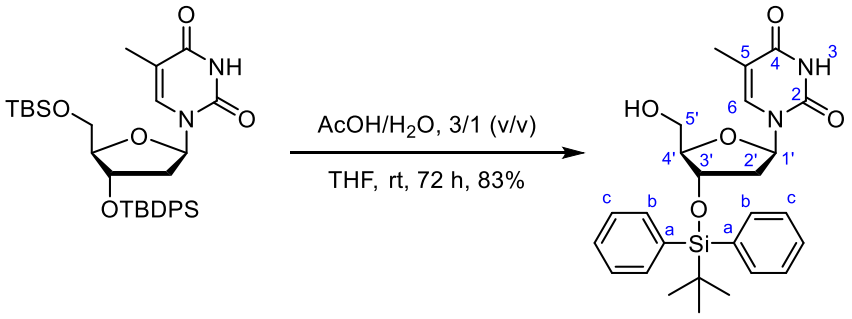
**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.92 (s, 1H, NH), 7.75 – 7.55 (m, 4H, H-Ar), 7.49 – 7.35 (m, 7H, H6 *and* 6 × H-Ar), 6.50 (dd, *J* = 9.0, 5.4 Hz, 1H, H1′), 4.37 – 4.30 (m, 1H, H3′), 3.99 (q, *J* = 2.0 Hz, 1H, H4′), 3.62 (dd, *J* = 11.4, 2.0 Hz, 1H, H5′), 3.13 (dd, *J* = 11.4, 2.0 Hz, 1H, H5″), 2.32 (ddd, *J* = 13.0, 5.4, 1.5 Hz, 1H, H2′), 1.87 (d, *J* = 1.2 Hz, 3H, $CH_3$-*thymine*), 1.85 – 1.78 (m, 1H, H2″), 1.09 (s, 9H, 3 × $CH_3$-*tert*

*butyl*), 0.78 (s, 9H, 3 × $CH_3$-*tert butyl*), –0.08 (s, 3H, $Si(CH_3)_2$), –0.12 (s, 3H, $Si(CH_3)_2$);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 164.0 (C4), 150.5 (C2), 135.9 ($^{b}$CH-Ar), 135.8 ($^{b}$CH-Ar), 135.6 (C6), 133.5 ($^{a}$C-Ar), 133.3 ($^{a}$C-Ar), 130.1 ($^{c}$CH-Ar), 130.1 ($^{c}$CH-Ar), 128.0 ($^{d}$CH-Ar), 128.0 ($^{d}$CH-Ar), 110.9 (C5), 88.2 (C4′), 85.2 (C1′), 74.4 (C3′), 63.4 (C5′), 41.5 (C2′), 27.0 (C(**C**$H_3)_3$), 25.9 (C(**C**$H_3)_3$), 19.2 (**C**$(CH_3)_3$), 18.4 (**C**$(CH_3)_3$), 12.6 ($CH_3$-*thymine*), –5.4 (Si($^{e}$$CH_3)_2$), –5.5 (Si($^{e}$$CH_3)_2$); $^{a,b,c,d,e}$Note that the marked carbons can resolve in distinct peaks due to different environments.

***m/z* (ESI$^+$)**: 617.2 ($[M + Na]^+$, 100%).

3′-*O*-(*tert*-Butyldiphenylsilyl)-thymidine (**56**)[300]




**55** **56**

$C_{26}H_{32}N_2O_5Si$
480.64

Disilylated thymidine **55** (13.2 g, 22.2 mmol) was dissolved in a solvent mixture of THF (44.4 mL) and AcOH/$H_2O$ [3/1 (v/v), 178 mL] and stirred at rt for 72 h. The solvents were evaporated under reduced pressure, the residue was dissolved in $CH_2Cl_2$ (100 mL), and washed with sat. aq $NaHCO_3$ (200 mL). The aq layer was extracted with $CH_2Cl_2$ (50 mL) and the combined org layers were dried over $Na_2SO_4$, filtered, and evaporated. The crude was submitted to column chromatography (petroleum ether (40–60)/EtOAc, 0–70%) to give alcohol **56** (8.80 g, 18.3 mmol, 83%).

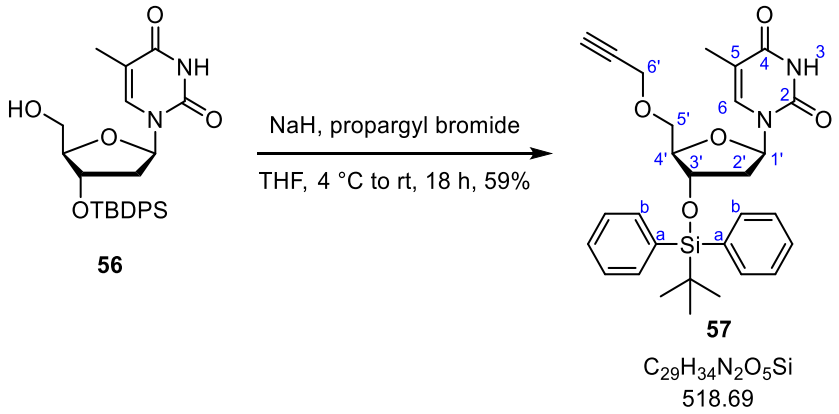
***R*$_f$** 0.21 (petroleum ether (40–60)/EtOAc, 1/1);

**$^1$H NMR** (400 MHz, $CDCl_3$) δ 8.68 (s, 1H, NH), 7.69 – 7.59 (m, 4H, H-Ar), 7.51 – 7.35 (m, 6H, H-Ar), 7.28 (d, *J* = 1.3 Hz, 1H, H6), 6.23 (dd, *J* = 7.8, 6.0 Hz, 1H, H1′), 4.45 (dt, *J* = 6.0, 3.0 Hz, 1H, H3′), 3.97 (q, *J* = 3.0 Hz, 1H, H4′), 3.63 (dt, *J* = 12.0, 3.0 Hz, 1H, H5′), 3.29 – 3.19 (m, 1H, H5″), 2.27 (ddd, *J* = 13.4, 6.0, 3.0 Hz, 1H, H2′), 2.23 – 2.10 (m, 2H, H2″ *and* OH), 1.84 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*), 1.09 (s, 9H, 3 × $CH_3$-*tert butyl*);

**$^{13}$C NMR** (101 MHz, $CDCl_3$) δ 163.7 (C4), 150.4 (C2), 137.0 (C6), 135.9 ($^{b}$CH-Ar), 135.8 ($^{b}$CH-Ar), 133.4 ($^{a}$C-Ar), 133.2 ($^{a}$C-Ar), 130.3 ($^{c}$CH-Ar), 130.2 ($^{c}$CH-Ar), 128.0 (CH-Ar), 111.1 (C5), 87.8 (C4′), 86.9 (C1′), 73.1 (C3′), 62.2 (C5′), 40.4 (C2′), 27.0 (C(**C**$H_3$)$_3$), 19.2 (**C**($CH_3$)$_3$), 12.6 ($CH_3$-*thymine*); $^{a,b,c}$Note that the marked carbons can resolve in distinct peaks due to different environments.

***m/z* (ESI$^+$)**: 503.2 ([M + Na]$^+$, 100%).

5′-*O*-Propargyl-3′-*O*-(*tert*-butyldiphenylsilyl)-thymidine (**57**)[307]




A 60% dispersion of NaH in mineral oil (1.33 g, 33.1 mmol) was added to solution of alcohol **56**[300] (5.31 g, 11.0 mmol) in anhyd THF (35 mL) at 4 °C, and the mixture was stirred at 4 °C for 20 min. Propargyl bromide (0.83 mL, 11.0 mmol) was added at 4 °C over 1 h, and the mixture was then stirred at rt for 18 h. The reaction was neutralised to pH 7 with 1 N aq AcOH, and extracted with $Et_2O$ (75 mL). The org layer was washed with brine (2 × 75 mL), dried over $Na_2SO_4$, filtered, and evaporated. The residue was submitted to column chromatography (petroleum ether (40–60)/EtOAc, 0–50%) to give alkyne **57** as a white solid (3.35 g, 6.46 mmol, 59%).

***R*f** 0.69 (petroleum ether (40–60)/EtOAc, 1/1);

**$^1$H NMR** (500 MHz, $CDCl_3$) δ 8.52 (s, 1H, NH), 7.68 – 7.60 (m, 4H, H-Ar), 7.52 (d, *J* = 1.3 Hz, 1H, H6), 7.49 – 7.43 (m, 2H, H-Ar), 7.43 – 7.36 (m, 4H, H-Ar), 6.48 (dd, *J* = 8.1, 5.8 Hz, 1H, H1′), 4.43 (dt, *J* = 5.8, 2.3 Hz, 1H, H3′), 4.06 – 3.98 (m, 2H, H4′ *and* H6′), 3.95 (dd, *J* = 15.7, 2.4 Hz, 1H, H6″), 3.55 (dd, *J* = 10.4, 2.4 Hz, 1H, H5′), 3.03 (dd, *J* = 10.4, 2.4 Hz, 1H, H5″), 2.36 (t, *J* = 2.4 Hz, 1H, C≡CH), 2.32

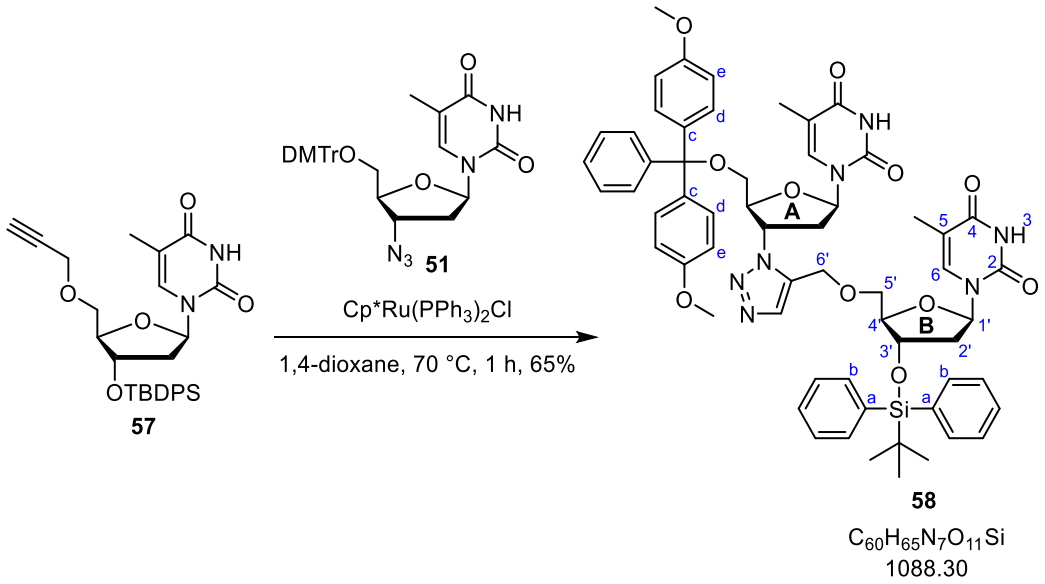
(ddd, *J* = 13.5, 5.8, 2.3 Hz, 1H, H2′), 1.94 (ddd, *J* = 13.5, 8.1, 5.8 Hz, 1H, H2″), 1.88 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*), 1.08 (s, 9H, 3 × $CH_3$-*tert butyl*);

**$^{13}$C NMR** (126 MHz, $CDCl_3$) δ 163.8 (C4), 150.4 (C2), 136.1 (C6), 135.9 ($^{b}$CH-Ar), 135.9 ($^{b}$CH-Ar), 133.4 ($^{a}$C-Ar), 133.2 ($^{a}$C-Ar), 130.2 (CH-Ar), 128.0 (CH-Ar), 111.0 (C5), 86.6 (C4′), 85.2 (C1′), 78.9 (**C**≡CH), 75.3 (C≡**C**H), 74.0 (C3′), 69.5 (C5′), 58.5 (C6′), 41.2 (C2′), 27.0 (C(**C**$H_3$)$_3$), 19.1 (**C**($CH_3$)$_3$), 12.7 ($CH_3$-*thymine*); $^{a,b}$Note that the marked carbons can resolve in distinct peaks due to different environments.

**HRMS–ESI** (*m/z*): $[M + H]^+$ calcd for $C_{29}H_{35}N_2O_5Si$, 519.2310; found, 519.2307.

*Experimental procedure and characterisation reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

5′-*O*-(4,4′-Dimethoxytrityl)-thymidine-TL**2**-3′-*O*-(*tert*-butyldiphenylsilyl)-thymidine dimer (**58**)[307]




Azide **51** (1.27 g, 2.23 mmol) and alkyne **57** (1.85 g, 3.57 mmol) were co-evaporated with anhyd 1,4-dioxane (2 × 10 mL) and dissolved in anhyd 1,4-dioxane (17 mL). Cp*RuCl(PPh$_3$)$_2$ (355 mg, 0.45 mmol) was added and the reaction was stirred at 70 °C for 1 h. The solvent was evaporated under reduced pressure and the residue re-dissolved in EtOAc (150 mL) and consecutively washed with sat. aq EDTA-Na$_2$ (150 mL), sat. aq NaHCO$_3$ (150 mL) and brine (150 mL). The separated organic layer was dried over Na$_2$SO$_4$, filtered and evaporated. Column chromatography of the crude (petroleum ether 40–60/EtOAc, 30–100%) gave dimer **58** (1.59 g, 1.46 mmol, 65%).

***R*$_f$** 0.45 (EtOAc);

**$^1$H NMR** (400 MHz, CDCl$_3$) δ 8.98 (2 × s, 2H, 2 × NH), 7.65 – 7.52 (m, 5H, H6 *and* 5 × H-Ar), 7.46 (s, 1H, H-*triazole*), 7.44 – 7.39 (m, 2H, H-Ar), 7.38 – 7.29 (m,

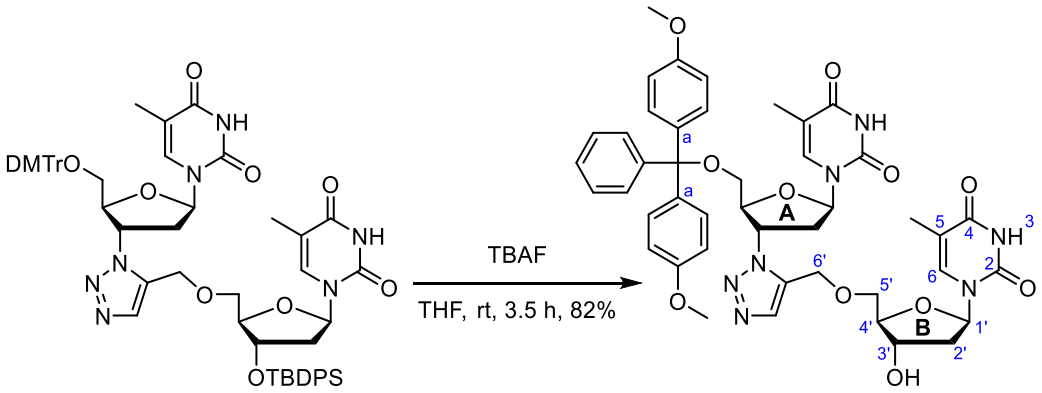
6H, H-Ar), 7.28 – 7.24 (m, 2H, H-Ar *and* $CDCl_3$), 7.24 – 7.17 (m, 5H, H-Ar), 6.88 (d, *J* = 1.3 Hz, 1H, H6), 6.85 – 6.74 (m, 4H, H-Ar), 6.49 (t, *J* = 6.5 Hz, 1H, H1′A), 6.24 (dd, *J* = 7.4, 6.3 Hz, 1H, H1′B), 5.16 (dt, *J* = 8.7, 4.8 Hz, 1H, H3′A), 4.38 (ddd, *J* = 4.8, 3.6, 3.0 Hz, 1H, H4′A), 4.34 – 4.23 (m, 2H, H6′ *and* H6″), 4.20 (dt, *J* = 6.8, 3.4 Hz, 1H, H3′B), 3.96 – 3.90 (m, 1H, H4′B), 3.76 (2 × s, 6H, 2 × $OCH_3$), 3.47 (dd, *J* = 10.8, 3.6 Hz, 1H, H5′A), 3.26 (dd, *J* = 10.8, 3.0 Hz, 1H, H5″A), 3.16 (dd, *J* = 10.5, 2.9 Hz, 1H, H5′B), 3.06 – 2.93 (m, 2H, H2′A *and* H5″B), 2.64 (ddd, *J* = 13.7, 8.7, 6.5 Hz, 1H, H2″A), 2.28 (ddd, *J* = 13.4, 6.3, 3.4 Hz, 1H, H2′B), 1.93 – 1.77 (m, 1H, H2″B), 1.66 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*), 1.60 (d, *J* = 1.3 Hz, 3H, $CH_3$-*thymine*), 1.04 (s, 9H, $CH_3$-*tert butyl*);

**$^{13}$C NMR** (126 MHz, $CDCl_3$) δ 163.9 (C4), 163.6 (C4), 158.9 (C-Ar), 150.3 (C2), 150.2 (C2), 144.3 (C-Ar), 136.2 (C6), 135.8 ($^{b}$CH-Ar), 135.7 ($^{b}$CH-Ar), 135.4 (C6), 135.2 ($^{c}$C-Ar), 135.1 ($^{c}$C-Ar), 134.1 (H**C**=C-*triazole*), 133.1 ($^{a}$C-Ar *or* HC=C-*triazole*), 133.0 ($^{a}$C-Ar *or* HC=**C**-*triazole*), 132.9 ($^{a}$C-Ar *or* HC=**C**-*triazole*), 130.3 (CH-Ar), 130.0 ($^{d}$CH-Ar), 130.0 ($^{d}$CH-Ar), 128.2 (CH-Ar), 128.1 (CH-Ar), 128.0 (CH-Ar), 127.4 (CH-Ar), 113.5 ($^{e}$CH-Ar), 113.4 ($^{e}$CH-Ar), 111.4 (C5), 111.3 (C5), 87.2 ($C_{quat}$-*DMTr*), 86.6 (C1′A), 85.5 (C4′B), 85.2 (C1′B), 83.8 (C4′A), 73.1 (C3′B), 70.1 (C5′B), 63.2 (C5′A), 61.0 (C6′), 58.4 (C3′A), 55.4 ($OCH_3$), 40.3 (C2′B), 38.6 (C2′A), 26.9 (C(**C**$H_3$)$_3$), 19.1 (**C**($CH_3$)$_3$), 12.6 ($CH_3$-*thymine*), 12.1 ($CH_3$-*thymine*); $^{a,b,c,d,e}$Note that the marked carbons can resolve in distinct peaks due to different environments.

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{60}H_{65}N_7O_{11}SiNa$, 1110.4404; found, 1110.4397.

*Experimental procedure and characterisation reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

5′-*O*-(4,4′-Dimethoxytrityl)-thymidine-TL**2**-thymidine dimer (**59**)[307]




**58** **59**

$C_{44}H_{47}N_7O_{11}$
849.90

Tetrabutylammonium fluoride (TBAF, 1 M in THF, 2.05 mL, 2.05 mmol) was added to a solution of dimer **58** (1.49 g, 1.37 mmol) in THF (12 mL) and stirred at rt for 3.5 h. The solvent was evaporated and the remaining solid submitted to column chromatography (EtOAc/propan-2-ol, 0–20%) to give dimer **59** as a white foam (0.95 g, 1.12 mmol, 82%).

***R*f** 0.51 (EtOAc/propan-2-ol, 4/1);

**$^1$H NMR** (500 MHz, $CDCl_3$) δ 10.54 (bs, 1H, NH), 9.96 (s, 1H, NH), 7.60 (s, 1H, H-*triazole*), 7.44 – 7.40 (m, 1H, H6), 7.37 – 7.32 (m, 2H, H-Ar), 7.29 – 7.16 (m, 8H, H6 *and* 7 × H-Ar), 6.84 – 6.77 (m, 4H, H-Ar), 6.15 (t, *J* = 5.9 Hz, 1H, H1′B), 6.08 (dd, *J* = 7.7, 4.4 Hz, 1H, H1′A), 5.57 (q, *J* = 7.8 Hz, 1H, H3′A), 4.58 – 4.51 (m, 2H, H3′B *and* H4′A), 4.48 (bs, 1H, OH), 4.44 (bs, 2H, H6′ *and* H6″), 4.03 – 3.96 (m, 1H, H4′B), 3.75 (2 × s, 6H, 2 × $OCH_3$), 3.73 – 3.63 (m, 2H, H5′B and H5″B), 3.54 (dd, *J* = 10.8, 3.8 Hz, 1H, H5′A), 3.23 – 3.11 (m, 2H, H2′A *and* H5″A), 2.97 –

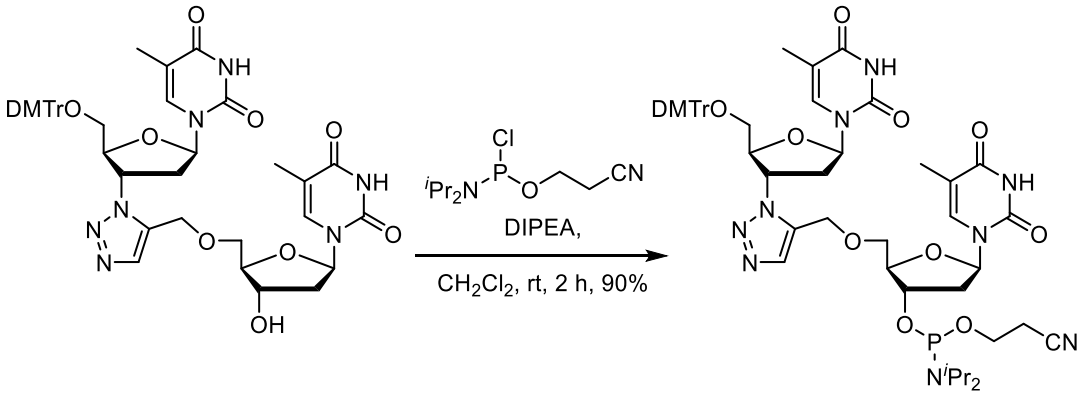
2.87 (m, 1H, H2″A), 2.35 (dt, *J* = 13.7, 5.9 Hz, 1H, H2′B), 2.11 – 2.03 (m, 1H, H2″B), 1.71 (d, *J* = 1.1 Hz, 3H, $CH_3$-*thymine*), 1.49 (d, *J* = 1.2 Hz, 3H, $CH_3$-*thymine*);

**$^{13}$C NMR** (126 MHz, $CDCl_3$) δ 165.0 (C4), 164.2 (C4), 158.8 (C-Ar), 150.9 (C2), 150.6 (C2), 144.2 (C-Ar), 138.5 (C6), 135.7 (C6), 135.4 ([a]C-Ar), 135.2 ([a]C-Ar), 133.8 (H**C**=C-*triazole or* HC=**C**-*triazole*), 133.7 (H**C**=C-*triazole or* HC=**C**-*triazole*), 130.0 (CH-Ar), 128.1 (CH-Ar), 128.1 (CH-Ar), 127.3 (CH-Ar), 113.4 (CH-Ar), 111.2 (C5), 111.0 (C5), 88.6 (C1′A), 86.9 ($C_{quat}$-*DMTr*), 85.2 (C1′B), 84.9 (C4′B), 83.6 (H4′A), 70.6 (H3′B *or* C5′B), 70.5 (H3′B *or* C5′B), 62.3 (C5′A), 60.8 (C6′), 57.7 (C3′A), 55.4 ($OCH_3$), 40.9 (C2′B), 38.3 (C2′A), 12.3 ($CH_3$-*thymine*), 12.2 ($CH_3$-*thymine*); [a]Note that the marked carbons can resolve in distinct peaks due to different environments.

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{44}H_{47}N_7O_{11}Na$, 872.3226; found, 872.3234.

*Experimental procedure and characterisation reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

5′-*O*-(4,4′-Dimethoxytrityl)-thymidine-TL**2**-thymidine dimer CE phosphoramidite (**49**)[307]




**59** **49**

$C_{53}H_{64}N_9O_{12}P$
1050.12

Degassed *N*,*N*-diisopropylethylamine (DIPEA, 370 μL, 2.13 mmol) and chloro(diisopropyl-amino)-β-cyanoethoxyphosphine (240 μL, 1.07 mmol) were added to a solution of dimer **59** (400 mg, 0.47 mmol) in degassed, anhyd $CH_2Cl_2$ (3 mL) and the reaction mixture was stirred at rt for 2 h. The solution was diluted with degassed $CH_2Cl_2$ (10 mL) and then washed with sat. aq KCl (10 mL). The aq layer was extracted with degassed $CH_2Cl_2$ (2 × 10 mL) and the combined organic layers were dried over $Na_2SO_4$, filtered and the solvents evaporated under reduced pressure. Purification by column chromatography (EtOAc) gave phosphoramidite **49** as a white foam (438 mg, 0.42 mmol, 90%).

***R*f** 0.40 (EtOAc);

**$^{31}$P NMR** (162 MHz, $CDCl_3$) δ 148.5, 148.1;

**HRMS–ESI** (*m/z*): $[M + Na]^+$ calcd for $C_{53}H_{64}N_9O_{12}PNa$, 1072.4304; found, 1072.4295.

*Experimental procedure and characterisation reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

# REFERENCES


(1) Miescher, F. J. Ueber Die Chemische Zusammensetzung Der Eiterzellen. *Med.-Chem. Unters.* **1871**, *4*, 441–460.
(2) Dahm, R. Friedrich Miescher and the Discovery of DNA. *Dev. Biol.* **2005**, *278* (2), 274–288.
(3) Crick, F. H. C. On Protein Synthesis. In *Symp Soc Exp Biol*; 1958; Vol. 12, p 8.
(4) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag New York Inc., 1984.
(5) Ni, G.; Du, Y.; Tang, F.; Liu, J.; Zhao, H.; Chen, Q. Review of α-Nucleosides: From Discovery, Synthesis to Properties and Potential Applications. *RSC Adv.* **2019**, *9* (25), 14302–14320.
(6) Röder, O.; Lüdemann, H.-D.; Von Goldammer, E. Determination of the Activation Energy for Pseudorotation of the Furanose Ring in Nucleosides by 13C Nuclear-Magnetic-Resonance Relaxation. *Eur. J. Biochem.* **1975**, *53* (2), 517–524.
(7) Evich, M.; Spring-Connell, A. M.; Germann, M. W. Impact of Modified Ribose Sugars on Nucleic Acid Conformation and Function. *Heterocycl. Commun.* **2017**, *23* (3), 155–165.
(8) Østergaard, M. E.; Dwight, T.; Berdeja, A.; Swayze, E. E.; Jung, M. E.; Seth, P. P. Comparison of Duplex Stabilizing Properties of 2′-Fluorinated Nucleic Acid Analogues with Furanose and Non-Furanose Sugar Rings. *J. Org. Chem.* **2014**, *79* (18), 8877–8881.
(9) Meldgaard, M.; Wengel, J. Bicyclic Nucleosides and Conformational Restriction of Oligonucleotides. *J. Chem. Soc. Perkin Trans. 1* **2000**, No. 21, 3539–3554.
(10) Gilbert, H. F. *Basic Concepts in Biochemistry: A Student's Survival Guide*; McGraw-Hill Education, 1992.
(11) Elson, D.; Chargaff, E. On the Desoxyribonucleic Acid Content of Sea Urchin Gametes. *Experientia* **1952**, *8* (4), 143–145.
(12) Chargaff, E.; Lipshitz, R.; Green, C. Composition of the Desoxypentose Nucleic Acids of Four Genera of Sea-Urchin. *J Biol Chem* **1952**, *195* (1), 155–160.
(13) Franklin, R. E.; Gosling, R. G. The Structure of Sodium Thymonucleate Fibres. I. The Influence of Water Content. *Acta Crystallogr.* **1953**, *6* (8–9), 673–677.
(14) Franklin, R. E.; Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171* (4356), 740–741.
(15) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids. *Nature* **1953**, *171* (4356), 738–740.
(16) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737–738.
(17) Crooke, S. T. *Antisense Drug Technology: Principles, Strategies, and Applications*; CRC press, 2007.
(18) Pal, A. K.; Bahadur, B. Chirality in Nature and Biomolecules: An Overview. *LS- Int. J. Life Sci.* **2020**, *9* (1), 72–84.
(19) Ranjbar, B.; Gill, P. Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses- A Review. *Chem. Biol. Drug Des.* **2009**, *74* (2), 101–120.
(20) Lesnik, E. A.; Freier, S. M. Relative Thermodynamic Stability of DNA, RNA, and DNA:RNA Hybrid Duplexes: Relationship with Base Composition and Structure. *Biochemistry* **1995**, *34* (34), 10807–10815.
(21) Suresh, G.; Priyakumar, U. D. DNA–RNA Hybrid Duplexes with Decreasing Pyrimidine Content in the DNA Strand Provide Structural Snapshots for the A- to B-Form Conformational Transition of Nucleic Acids. *Phys. Chem. Chem. Phys.* **2014**, *16* (34), 18148–18155.
(22) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.
(23) Nobel Prize Outreach AB. *The Nobel Prize in Chemistry 1984*. Retrieved from https://www.nobelprize.org/prizes/chemistry/1984/summary (accessed Jun 15, 2021).
(24) Garegg, P. J.; Lindh, I.; Regberg, T.; Stawinski, J.; Strömberg, R.; Henrichson, C. Nucleoside H-Phosphonates. III. Chemical Synthesis of Oligodeoxyribonucleotides by the

Hydrogenphosphonate Approach. *Tetrahedron Lett.* **1986**, *27* (34), 4051–4054.
(25) Froehler, B. C.; Ng, P. G.; Matteucci, M. D. Synthesis of DNA via Deoxynudeoside H-Phosphonate Intermediates. *Nucleic Acids Res.* **1986**, *14* (13), 5399–5407.
(26) Efimov, V. A.; Buryakova, A. A.; Reverdatto, S. V; Chakhmakhcheva, O. G.; Ovchinnikov, Y. A. Rapid Synthesis of Long-Chain Deoxyribooligonucleotides by the N-Methylimidazolide Phosphotriester Method. *Nucleic Acids Res.* **1983**, *11* (23), 8369–8387.
(27) Efimov, V. A.; Molchanova, N. S.; Chakhmakhcheva, O. G. Approach to the Synthesis of Natural and Modified Oligonucleotides by the Phosphotriester Method Using O-Nucleophilic Intramolecular Catalysis. *Nucleosides, Nucleotides Nucleic Acids* **2007**, *26* (8–9), 1087–1093.
(28) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites-A New Class of Key Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862.
(29) Glen Research. *Glen Report 10.11 - DCI - A LOGICAL ALTERNATIVE ACTIVATOR*. Retrieved from https://www.glenresearch.com/reports/gr10-11 (accessed Jun 16, 2021).
(30) Brown, T.; Brown, T. (Jnr). *Nucleic Acids Book*. Retrieved from https://www.atdbio.com/nucleic-acids-book (accessed Jun 4, 2021).
(31) Glen Research. *Glen Report 20.24 - Deprotection - Volume 1 - Deprotect to Completion*. Retrieved from https://www.glenresearch.com/reports/gr20-24 (accessed Feb 12, 2022).
(32) Service, R. F. *DNA could store all of the world's data in one room*. [Blog post]. Retrieved from https://www.sciencemag.org/news/2017/03/dna-could-store-all-worlds-data-one-room (accessed Jun 11, 2021).
(33) Ochoa, S.; Milam, V. T. Modified Nucleic Acids: Expanding the Capabilities of Functional Oligonucleotides. *Molecules* **2020**, *25* (20), 4659.
(34) Clavé, G.; Reverte, M.; Vasseur, J.-J.; Smietana, M. Modified Internucleoside Linkages for Nuclease-Resistant Oligonucleotides. *RSC Chem. Biol.* **2021**, *2* (1), 94–150.
(35) Henry, A. A.; Romesberg, F. E. Beyond A, C, G and T: Augmenting Nature's Alphabet. *Curr. Opin. Chem. Biol.* **2003**, *7* (6), 727–733.
(36) Fantoni, N. Z.; El-Sagheer, A. H.; Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem. Rev.* **2021**, *121* (12), 7122–7154.
(37) Ganz, D.; Harijan, D.; Wagenknecht, H.-A. Labelling of DNA and RNA in the Cellular Environment by Means of Bioorthogonal Cycloaddition Chemistry. *RSC Chem. Biol.* **2020**, *1* (3), 86–97.
(38) Kumar, S.; Chinnusamy, V.; Mohapatra, T. Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. *Front. Genet.* **2018**, *9*, 640.
(39) El-Sagheer, A. H.; Brown, T. Combined Nucleobase and Backbone Modifications Enhance DNA Duplex Stability and Preserve Biocompatibility. *Chem. Sci.* **2014**, *5* (1), 253–259.
(40) Brad Wan, W.; Seth, P. P. The Medicinal Chemistry of Therapeutic Oligonucleotides. *J. Med. Chem.* **2016**, *59* (21), 9645–9667.
(41) Hocek, M. Synthesis of Base-Modified 2′-Deoxyribonucleoside Triphosphates and Their Use in Enzymatic Synthesis of Modified DNA for Applications in Bioanalysis and Chemical Biology. *J. Org. Chem.* **2014**, *79* (21), 9914–9921.
(42) Sarac, I.; Hollenstein, M. Terminal Deoxynucleotidyl Transferase in the Synthesis and Modification of Nucleic Acids. *ChemBioChem* **2019**, *20* (7), 860–871.
(43) ATDBio Ltd. *Common modifications*. Retrieved from https://www.atdbio.com/content/9/Common-modifications (accessed Jun 11, 2021).
(44) Integrated DNA Technologies Inc. *Oligo modifications*. Retrieved from https://eu.idtdna.com/pages/products/custom-dna-rna/oligo-modifications (accessed Jun 11, 2021).
(45) TriLink BioTechnologies. *Oligonucleotide Modification Information | TriLink BioTechnologies*. Retrieved from https://www.trilinkbiotech.com/omi (accessed Jun 15, 2021).
(46) Glen Research. *Modification and Labeling - Products*. Retrieved from https://www.glenresearch.com/products/labels-and-modifiers-1.html (accessed Jun 17, 2021).
(47) LGC Genomics Ltd. *Conjugation Reagents - Link Technologies - Reagents for Oligonucleotide Synthesis & Modification*. Retrieved from http://www.linktech.co.uk/products/bioconjugation_and_immobilisation (accessed Jun 17, 2021).
(48) Shivalingam, A.; Tyburn, A. E. S.; El-Sagheer, A. H.; Brown, T. Molecular Requirements of

High-Fidelity Replication-Competent DNA Backbones for Orthogonal Chemical Ligation. *J. Am. Chem. Soc.* **2017**, *139* (4), 1575–1583.
(49) Glen Research. *Sulfurizing Reagent Products | Glen Research*. Retrieved from https://www.glenresearch.com/products/ancillary-reagents/sulfurizing-reagents.html (accessed Jul 16, 2021).
(50) Stec, W. J.; Zon, G.; Egan, W.; Stec, B. Automated Solid-Phase Synthesis, Separation, and Stereochemistry of Phosphorothioate Analogues of Oligodeoxyribonucleotides. *J. Am. Chem. Soc.* **1984**, *106* (20), 6077–6079.
(51) Knouse, K. W.; DeGruyter, J. N.; Schmidt, M. A.; Zheng, B.; Vantourout, J. C.; Kingston, C.; Mercer, S. E.; Mcdonald, I. M.; Olson, R. E.; Zhu, Y.; Hang, C.; Zhu, J.; Yuan, C.; Wang, Q.; Park, P.; Eastgate, M. D.; Baran, P. S. Unlocking P(V): Reagents for Chiral Phosphorothioate Synthesis. *Science* **2018**, *361* (6408), 1234–1238.
(52) Glen Research. *Methyl Phosphonamidites - Backbone Modification - Label Phosphoramidites and Supports - Modification and Labeling - Products*. Retrieved from https://www.glenresearch.com/products/labels-and-modifiers/label-phosphoramidites-and-supports/backbone-modification-phosphoramidites-and-supports/methyl-phosphonamidites.html (accessed Jun 21, 2021).
(53) Miller, P. S.; Yano, J.; Yano, E.; Carroll, C.; Jayaraman, K.; Ts'o, P. O. P. Nonionic Nucleic Acid Analogues. Synthesis and Characterization of Dideoxyribonucleoside Methylphosphonates. *Biochemistry* **1979**, *18* (23), 5134–5143.
(54) Hogrefe, R. I.; Vaghefi, M. M.; Reynolds, M. A.; Young, K. M.; Arnold, L. J. Deprotection of Methylphosphonate Oligonucleotides Using a Novel One-Pot Procedure. *Nucleic Acids Res.* **1993**, *21* (9), 2031–2038.
(55) von Matt, P.; De Mesmaeker, A.; Pieles, U.; Zürcher, W.; Altmann, K. H. 2'-Deoxyribo-PNAs: A Structurally Novel Class of Polyamide Nucleic Acids with Good RNA and DNA Binding Affinity. *Tetrahedron Lett.* **1999**, *40* (15), 2899–2902.
(56) Epple, S.; El-Sagheer, A. H.; Brown, T. Artificial Nucleic Acid Backbones and Their Applications in Therapeutics, Synthetic Biology and Biotechnology. *Emerg. Top. Life Sci.* **2021**, *143* (39), 16293–16301.
(57) Li, P.; Sergueeva, Z. A.; Dobrikov, M.; Shaw, B. R. Nucleoside and Oligonucleoside Boranophosphates: Chemistry and Properties. *Chem. Rev.* **2007**, *107* (11), 4746–4796.
(58) Brill, W. K. D.; Tang, J. Y.; Ma, Y. X.; Caruthers, M. H. Synthesis of Oligodeoxynucleoside Phosphorodithioates via Thioamidites. *J. Am. Chem. Soc.* **1989**, *111* (6), 2321–2322.
(59) Kers, I.; Stawiński, J.; Kraszewski, A. Aryl H-Phosphonates. 10. Synthesis of Nucleoside Phosphoramidate and Nucleoside Phosphoramidothioate Analogues via H-Phosphonamidate Intermediates. *Tetrahedron* **1999**, *55* (38), 11579–11588.
(60) Isobe, H.; Fujino, T. Triazole-Linked Analogues of DNA and RNA (TLDNA and TLRNA): Synthesis and Functions. *Chem. Rec.* **2014**, *14* (1), 41–51.
(61) Ghimire, C.; Park, S.; Iida, K.; Yangyuoru, P.; Otomo, H.; Yu, Z.; Nagasawa, K.; Sugiyama, H.; Mao, H. Direct Quantification of Loop Interaction and π–π Stacking for G-Quadruplex Stability at the Submolecular Level. *J. Am. Chem. Soc.* **2014**, *136* (44), 15537–15544.
(62) De Mesmaeker, A.; Waldner, A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M.; Freier, S. M. Amides as a New Type of Backbone Modification in Oligonucleotides. *Angew. Chem., Int. Ed.* **1994**, *33* (2), 226–229.
(63) De Mesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. Amides as Substitute for the Phosphodiester Linkage in Antisense Oligonucleotides. *Synlett* **1993**, *1993* (10), 733–736.
(64) Lebreton, J.; De Mesmaeker, A.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. Synthesis of Thymidine Dimer Derivatives Containing an Amide Linkage and Their Incorporation into Oligodeoxyribonucleotides. *Tetrahedron Lett.* **1993**, *34* (40), 6383–6386.
(65) 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 Lett.* **1994**, *35* (29), 5225–5228.
(66) De Mesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M. Replacement of the Phosphodiester Linkage in Oligonucleotides: Comparison of Two Structural Amide Isomers. *Bioorganic Med. Chem. Lett.* **1994**, *4* (7), 873–878.
(67) Lebreton, J.; Waldner, A.; Lesueur, C.; De Mesmaeker, A. Antisense Oligonucleotides with Alternating Phosphodiester-"Amide-3" Linkages. *Synlett* **1994**, *1994* (02), 137–140.
(68) De Mesmaeker, A.; Lesueur, C.; Fritsch, V.; Waldner, A.; Bévière, M.-O.; Wolf, R. M. Amide Backbones with Conformationally Restricted Furanose Rings: Highly Improved Affinity of

the Modified Oligonucleotides for Their RNA Complements. *Angew. Chem., Int. Ed.* **1996**, *35* (24), 2790–2794.
(69) Waldner, A.; De Mesmaeker, A.; Lebreton, J. Synthesis of Oligodeoxyribonucleotides Containing Dimers with Carbamate Moieties as Replacement of the Natural Phosphodiester Linkage. *Bioorganic Med. Chem. Lett.* **1994**, *4* (3), 405–408.
(70) Madhuri, V.; Kumar, V. A. Design and Synthesis of Dephosphono DNA Analogues Containing 1,2,3-Triazole Linker and Their UV-Melting Studies with DNA/RNA. *Nucleosides, Nucleotides Nucleic Acids* **2012**, *31* (2), 97–111.
(71) Varizhuk, A. M.; Kaluzhny, D. N.; Novikov, R. A.; Chizhov, A. O.; Smirnov, I. P.; Chuvilin, A. N.; Tatarinova, O. N.; Fisunov, G. Y.; Pozmogova, G. E.; Florentiev, V. L. Synthesis of Triazole-Linked Oligonucleotides with High Affinity to DNA Complements and an Analysis of Their Compatibility with Biosystems. *J. Org. Chem.* **2013**, *78* (12), 5964–5969.
(72) Baker, Y. R.; Traoré, D.; Wanat, P.; Tyburn, A.; El-Sagheer, A. H.; Brown, T. Searching for the Ideal Triazole: Investigating the 1,5-Triazole as a Charge Neutral DNA Backbone Mimic. *Tetrahedron* **2020**, *76* (7), 130914.
(73) Palframan, M. J.; Alharthy, R. D.; Powalowska, P. K.; Hayes, C. J. Synthesis of Triazole-Linked Morpholino Oligonucleotides via CuI Catalysed Cycloaddition. *Org. Biomol. Chem.* **2016**, *14* (11), 3112–3119.
(74) Ueno, Y.; Naito, T.; Kawada, K.; Shibata, A.; Kim, H. S.; Wataya, Y.; Kitade, Y. Synthesis of Novel SiRNAs Having Thymidine Dimers Consisting of a Carbamate or a Urea Linkage at Their 3′ Overhang Regions and Their Ability to Suppress Human RNase L Protein Expression. *Biochem. Biophys. Res. Commun.* **2005**, *330* (4), 1168–1175.
(75) Sato, K.; Seio, K.; Sekine, M. Squaryl Group as a New Mimic of Phosphate Group in Modified Oligodeoxynucleotides: Synthesis and Properties of New Oligodeoxynucleotide Analogues Containing an Internucleotidic Squaryldiamide Linkage. *J. Am. Chem. Soc.* **2002**, *124* (43), 12715–12724.
(76) Li, H.; Miller, M. J. Syntheses and Binding Studies of Oligonucleotides Containing N-Hydroxycarbamate Linkages: Potential DNA Cleaving Antisense Oligomers. *Tetrahedron Lett.* **2000**, *41* (22), 4323–4327.
(77) Integrated DNA Technologies Inc. *Oligo Synthesis: Coupling Efficiency & Quality Control | IDT*. Retrieved from https://eu.idtdna.com/pages/education/decoded/article/oligo-synthesis-why-idt-leads-the-oligo-industry (accessed Jul 7, 2021).
(78) Integrated DNA Technologies Inc. *Ultramer DNA Oligos*. Retrieved from https://eu.idtdna.com/pages/products/custom-dna-rna/dna-oligos/ultramer-dna-oligos (accessed Jul 7, 2021).
(79) Integrated DNA Technologies Inc. *Long, custom RNA Oligos-Ultramer RNA Oligonucleotides*. Retrieved from https://eu.idtdna.com/pages/education/decoded/article/long-custom-rna-oligos-ultramer-rna-oligonucleotides (accessed Jul 7, 2021).
(80) Xu, Y.; Kool, E. T. Chemical and Enzymatic Properties of Bridging 5'-S-Phosphorothioester Linkages in DNA. *Nucleic Acids Res.* **1998**, *26* (13), 3159–3164.
(81) El-Sagheer, A. H.; Brown, T. Single Tube Gene Synthesis by Phosphoramidate Chemical Ligation. *Chem. Commun.* **2017**, *53* (77), 10700–10702.
(82) Nakamoto, K.; Abe, N.; Tsuji, G.; Kimura, Y.; Tomoike, F.; Shimizu, Y.; Abe, H. Chemically Synthesized Circular RNAs with Phosphoramidate Linkages Enable Rolling Circle Translation. *Chem. Commun.* **2020**, *56* (46), 6217–6220.
(83) Chen, J.; Baker, Y. R.; Brown, A.; El-Sagheer, A. H.; Brown, T. Enzyme-Free Synthesis of Cyclic Single-Stranded DNA Constructs Containing a Single Triazole, Amide or Phosphoramidate Backbone Linkage and Their Use as Templates for Rolling Circle Amplification and Nanoflower Formation. *Chem. Sci.* **2018**, *9* (42), 8110–8120.
(84) Maruyama, H.; Oikawa, R.; Hayakawa, M.; Takamori, S.; Kimura, Y.; Abe, N.; Tsuji, G.; Matsuda, A.; Shuto, S.; Ito, Y.; Abe, H. Chemical Ligation of Oligonucleotides Using an Electrophilic Phosphorothioester. *Nucleic Acids Res.* **2017**, *45* (12), 7042–7048.
(85) Shivalingam, A.; Taemaitree, L.; El-Sagheer, A. H.; Brown, T. Squaramides and Ureas: A Flexible Approach to Polymerase-Compatible Nucleic Acid Assembly. *Angew. Chem., Int. Ed.* **2020**, *59* (28), 11416–11422.
(86) Miura, F.; Fujino, T.; Kogashi, K.; Shibata, Y.; Miura, M.; Isobe, H.; Ito, T. Triazole Linking for Preparation of a Next-Generation Sequencing Library from Single-Stranded DNA. *Nucleic Acids Res.* **2018**, *46* (16), e95.
(87) Ambadipudi, S.; Zweckstetter, M. Targeting Intrinsically Disordered Proteins in Rational

Drug Discovery. *Expert Opin. Drug Discov.* **2016**, *11* (1), 65–77.
(88) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat. Rev. Drug Discov.* **2018**, *17* (8), 547–558.
(89) Crooke, S. T.; Liang, X. H.; Baker, B. F.; Crooke, R. M. Antisense Technology: A Review. *J. Biol. Chem.* **2021**, *296*, 100416.
(90) Ho, S. P.; Bao, Y.; Lesher, T.; Malhotra, R.; Ma, L. Y.; Fluharty, S. J.; Sakai, R. R. Mapping of RNA Accessible Sites for Antisense Experiments with Oligonucleotide Libraries. *Nat. Biotechnol.* **1998**, *16* (1), 59–63.
(91) Rinaldi, C.; Wood, M. J. A. Antisense Oligonucleotides: The next Frontier for Treatment of Neurological Disorders. *Nat. Rev. Neurol.* **2017**, *14*, 9–21.
(92) Geary, R. S.; Watanabe, T. A.; Truong, L.; Freier, S.; Lesnik, E. A.; Sioufi, N. B.; Sasmor, H.; Manoharan, M.; Levin, A. A. Pharmacokinetic Properties of 2′-O-(2-Methoxyethyl)-Modified Oligonucleotide Analogs in Rats. *J. Pharmacol. Exp. Ther.* **2001**, *296* (3), 890–897.
(93) Kuijper, E. C.; Bergsma, A. J.; Pijnappel, W. W. M. P.; Aartsma-Rus, A. Opportunities and Challenges for Antisense Oligonucleotide Therapies. *J. Inherit. Metab. Dis.* **2021**, *44* (1), 72–87.
(94) Iwamoto, N.; Butler, D. C. D.; Svrzikapa, N.; Mohapatra, S.; Zlatev, I.; Sah, Di. W. Y.; Meena; Standley, S. M.; Lu, G.; Apponi, L. H.; Frank-Kamenetsky, M.; Zhang, J. J.; Vargeese, C.; Verdine, G. L. Control of Phosphorothioate Stereochemistry Substantially Increases the Efficacy of Antisense Oligonucleotides. *Nat. Biotechnol.* **2017**, *35* (9), 845–851.
(95) Havens, M. A.; Hastings, M. L. Splice-Switching Antisense Oligonucleotides as Therapeutic Drugs. *Nucleic Acids Res.* **2016**, *44* (14), 6549–6563.
(96) Aartsma-Rus, A. *Exon Skipping: Methods and Protocols*; Springer, 2012.
(97) Bladen, C. L.; Salgado, D.; Monges, S.; Foncuberta, M. E.; Kekou, K.; Kosma, K.; Dawkins, H.; Lamont, L.; Roy, A. J.; Chamova, T.; Guergueltcheva, V.; Chan, S.; Korngut, L.; Campbell, C.; Dai, Y.; Wang, J.; Barišić, N.; Brabec, P.; Lahdetie, J.; Walter, M. C.; Schreiber-Katz, O.; Karcagi, V.; Garami, M.; Viswanathan, V.; Bayat, F.; Buccella, F.; Kimura, E.; Koeks, Z.; van den Bergen, J. C.; Rodrigues, M.; Roxburgh, R.; Lusakowska, A.; Kostera-Pruszczyk, A.; Zimowski, J.; Santos, R.; Neagu, E.; Artemieva, S.; Rasic, V. M.; Vojinovic, D.; Posada, M.; Bloetzer, C.; Jeannet, P.-Y.; Joncourt, F.; Díaz-Manera, J.; Gallardo, E.; Karaduman, A. A.; Topaloğlu, H.; El Sherif, R.; Stringer, A.; Shatillo, A. V; Martin, A. S.; Peay, H. L.; Bellgard, M. I.; Kirschner, J.; Flanigan, K. M.; Straub, V.; Bushby, K.; Verschuuren, J.; Aartsma-Rus, A.; Béroud, C.; Lochmüller, H. The TREAT-NMD DMD Global Database: Analysis of More than 7,000 Duchenne Muscular Dystrophy Mutations. *Hum. Mutat.* **2015**, *36* (4), 395–402.
(98) Syed, Y. Y. Eteplirsen: First Global Approval. *Drugs* **2016**, *76* (17), 1699–1704.
(99) Liang, X. H.; Sun, H.; Nichols, J. G.; Crooke, S. T. RNase H1-Dependent Antisense Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the Nucleus. *Mol. Ther.* **2017**, *25* (9), 2075–2092.
(100) Hausen, P.; Stein, H. Ribonuclease H. *Eur. J. Biochem.* **1970**, *14* (2), 278–283.
(101) Stein, H.; Hausen, P. Enzyme from Calf Thymus Degrading the RNA Moiety of DNA-RNA Hybrids: Effect on DNA-Dependent RNA Polymerase. *Science* **1969**, *166* (3903), 393–395.
(102) Smith, C. I. E.; Zain, R. Therapeutic Oligonucleotides: State of the Art. *Annu. Rev. Pharmacol. Toxicol.* **2019**, *59* (1), 605–630.
(103) Bennett, C. F. Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* **2019**, *70* (1), 307–321.
(104) Peterson, P. A.; Rask, L.; Östberg, L.; Andersson, L.; Kamwendo, F.; Pertoft, H. Studies on the Transport and Cellular Distribution of Vitamin A in Normal and Vitamin A-Deficient Rats with Special Reference to the Vitamin A-Binding Plasma Protein. *J. Biol. Chem.* **1973**, *248* (11), 4009–4022.
(105) Power, D. M.; Elias, N. P.; Richardson, S. J.; Mendes, J.; Soares, C. M.; Santos, C. R. A. Evolution of the Thyroid Hormone-Binding Protein, Transthyretin. *Gen. Comp. Endocrinol.* **2000**, *119* (3), 241–255.
(106) Rosen, H. N.; Moses, A. C.; Murrell, J. R.; Liepnieks, J. J.; Benson, M. D. Thyroxine Interactions with Transthyretin: A Comparison of 10 Different Naturally Occurring Human Transthyretin Variants. *J. Clin. Endocrinol. Metab.* **1993**, *77* (2), 370–374.
(107) Hawkins, P. N.; Ando, Y.; Dispenzeri, A.; Gonzalez-Duarte, A.; Adams, D.; Suhr, O. B. Evolving Landscape in the Management of Transthyretin Amyloidosis. *Ann. Med.* **2015**, *47* (8), 625–638.
(108) Keam, S. J. Inotersen: First Global Approval. *Drugs* **2018**, *78* (13), 1371–1376.

(109) Gertz, M. A.; Scheinberg, M.; Waddington-Cruz, M.; Heitner, S. B.; Karam, C.; Drachman, B.; Khella, S.; Whelan, C.; Obici, L. Inotersen for the Treatment of Adults with Polyneuropathy Caused by Hereditary Transthyretin-Mediated Amyloidosis. *Expert Rev. Clin. Pharmacol.* **2019**, *12* (8), 701–711.
(110) Mahfouz, M.; Maruyama, R.; Yokota, T. Inotersen for the Treatment of Hereditary Transthyretin Amyloidosis. In *Gapmers*; Springer, 2020; pp 87–98.
(111) Aslesh, T.; Yokota, T. Development of Antisense Oligonucleotide Gapmers for the Treatment of Huntington's Disease. In *Gapmers*; Springer, 2020; pp 57–67.
(112) Chan, L.; Yokota, T. Development and Clinical Applications of Antisense Oligonucleotide Gapmers. In *Gapmers*; Springer, 2020; pp 21–47.
(113) Kaemmerer, W. F.; Grondin, R. C. The Effects of Huntingtin-Lowering: What Do We Know so Far? *Degener. Neurol. Neuromuscul. Dis.* **2019**, *9*, 3–17.
(114) Wild, E. J.; Tabrizi, S. J. Therapies Targeting DNA and RNA in Huntington's Disease. *Lancet Neurol.* **2017**, *16* (10), 837–847.
(115) Sheridan, C. Questions Swirl around Failures of Disease-Modifying Huntington's Drugs. *Nat. Biotechnol.* **2021**, *39* (6), 650–652.
(116) Roberts, T. C.; Langer, R.; Wood, M. J. A. Advances in Oligonucleotide Drug Delivery. *Nat. Rev. Drug Discov.* **2020**, *19* (10), 673–694.
(117) Lamb, Y. N. Inclisiran: First Approval. *Drugs* **2021**, *81* (3), 389–395.
(118) Vijay, N.; Ali, A. Inclisiran: A Game Changer in a Changing Game? *J. Am. Coll. Cardiol.* **2021**, *77* (9), 1194–1196.
(119) Hoy, S. M. Patisiran: First Global Approval. *Drugs* **2018**, *78* (15), 1625–1631.
(120) Scott, L. J. Givosiran: First Approval. *Drugs* **2020**, *80* (3), 335–339.
(121) Perry, C. M.; Balfour, J. A. B. Fomivirsen. *Drugs* **1999**, *57* (3), 375–380.
(122) Koziolkiewicz, M.; Wojcik, M.; Kobylanska, A.; Karkowski, B.; Rebowska, B.; Guga, P.; Stec, W. J. Stability of Stereoregular Oligo(Nucleoside Phosphorothioate)s in Human Plasma: Diastereoselectivity of Plasma 3`-Exonuclease. *Antisense Nucleic Acid Drug Dev.* **1997**, *7* (1), 43–48.
(123) Geary, R. S.; Norris, D.; Yu, R.; Bennett, C. F. Pharmacokinetics, Biodistribution and Cell Uptake of Antisense Oligonucleotides. *Adv. Drug Deliv. Rev.* **2015**, *87*, 46–51.
(124) Crooke, S. T.; Wang, S.; Vickers, T. A.; Shen, W.; Liang, X. H. Cellular Uptake and Trafficking of Antisense Oligonucleotides. *Nat. Biotechnol.* **2017**, *35* (3), 230–237.
(125) Liang, X. H.; Sun, H.; Shen, W.; Crooke, S. T. Identification and Characterization of Intracellular Proteins That Bind Oligonucleotides with Phosphorothioate Linkages. *Nucleic Acids Res.* **2015**, *43* (5), 2927–2945.
(126) Vickers, T. A.; Rahdar, M.; Prakash, T. P.; Crooke, S. T. Kinetic and Subcellular Analysis of PS-ASO/Protein Interactions with P54nrb and RNase H1. *Nucleic Acids Res.* **2019**, *47* (20), 10865–10880.
(127) Shen, W.; De Hoyos, C. L.; Migawa, M. T.; Vickers, T. A.; Sun, H.; Low, A.; Bell, T. A.; Rahdar, M.; Mukhopadhyay, S.; Hart, C. E.; Bell, M.; Riney, S.; Murray, S. F.; Greenlee, S.; Crooke, R. M.; Liang, X. hai; Seth, P. P.; Crooke, S. T. Chemical Modification of PS-ASO Therapeutics Reduces Cellular Protein-Binding and Improves the Therapeutic Index. *Nat. Biotechnol.* **2019**, *37* (6), 640–650.
(128) Migawa, M. T.; Shen, W.; Wan, W. B.; Vasquez, G.; Oestergaard, M. E.; Low, A.; De Hoyos, C. L.; Gupta, R.; Murray, S.; Tanowitz, M.; Bell, M.; Nichols, J. G.; Gaus, H.; Liang, X.; Swayze, E. E.; Crooke, S. T.; Seth, P. P. Site-Specific Replacement of Phosphorothioate with Alkyl Phosphonate Linkages Enhances the Therapeutic Profile of Gapmer ASOs by Modulating Interactions with Cellular Proteins. *Nucleic Acids Res.* **2019**, *47* (11), 5465–5479.
(129) Vasquez, G.; Migawa, M. T.; Wan, W. B.; Low, A.; Tanowitz, M.; Swayze, E. E.; Seth, P. P. Evaluation of Phosphorus and Non-Phosphorus Neutral Oligonucleotide Backbones for Enhancing Therapeutic Index of Gapmer Antisense Oligonucleotides. *Nucleic Acid Ther.* **2021**, *32* (1), 40–50.
(130) Anderson, B. A.; Freestone, G. C.; Low, A.; De-Hoyos, C. L.; III, W. J. D.; Østergaard, M. E.; Migawa, M. T.; Fazio, M.; Wan, W. B.; Berdeja, A.; Scandalis, E.; Burel, S. A.; Vickers, T. A.; Crooke, S. T.; Swayze, E. E.; Liang, X.; Seth, P. P. Towards next Generation Antisense Oligonucleotides: Mesylphosphoramidate Modification Improves Therapeutic Index and Duration of Effect of Gapmer Antisense Oligonucleotides. *Nucleic Acids Res.* **2021**, *49* (16), 9026–9041.
(131) Tran, H.; Moazami, M. P.; Yang, H.; McKenna-Yasek, D.; Douthwright, C. L.; Pinto, C.;

Metterville, J.; Shin, M.; Sanil, N.; Dooley, C.; Puri, A.; Weiss, A.; Wightman, N.; Gray-Edwards, H.; Marosfoi, M.; King, R. M.; Kenderdine, T.; Fabris, D.; Bowser, R.; Watts, J. K.; Brown, R. H. Suppression of Mutant C9orf72 Expression by a Potent Mixed Backbone Antisense Oligonucleotide. *Nat. Med.* **2022**, *28* (1), 117–124.
(132) Hair, P.; Cameron, F.; McKeage, K. Mipomersen Sodium: First Global Approval. *Drugs* **2013**, *73* (5), 487–493.
(133) Hoy, S. M. Nusinersen: First Global Approval. *Drugs* **2017**, *77* (4), 473–479.
(134) Paik, J.; Duggan, S. Volanesorsen: First Global Approval. *Drugs* **2019**, *79* (12), 1349–1354.
(135) Summerton, J.; Weller, D. Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense Nucleic Acid Drug Dev.* **1997**, *7* (3), 187–195.
(136) Miyatake, S.; Mizobe, Y.; Tsoumpra, M. K.; Lim, K. R. Q.; Hara, Y.; Shabanpoor, F.; Yokota, T.; Takeda, S.; Aoki, Y. Scavenger Receptor Class A1 Mediates Uptake of Morpholino Antisense Oligonucleotide into Dystrophic Skeletal Muscle. *Mol. Ther. - Nucleic Acids* **2019**, *14*, 520–535.
(137) Heo, Y.-A. Golodirsen: First Approval. *Drugs* **2020**, *80* (3), 329–333.
(138) Dhillon, S. Viltolarsen: First Approval. *Drugs* **2020**, *80* (10), 1027–1031.
(139) Food and Drug Administration. *Product Quality Review(s)*. Retrieved from https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/211970Orig1s000ChemR.pdf (accessed Jul 12, 2021).
(140) Gerety, S. S.; Wilkinson, D. G. Morpholino Artifacts Provide Pitfalls and Reveal a Novel Role for Pro-Apoptotic Genes in Hindbrain Boundary Development. *Dev. Biol.* **2011**, *350* (2), 279–289.
(141) Blum, M.; De Robertis, E. M.; Wallingford, J. B.; Niehrs, C. Morpholinos: Antisense and Sensibility. *Dev. Cell* **2015**, *35* (2), 145–149.
(142) Nair, J. K.; Attarwala, H.; Sehgal, A.; Wang, Q.; Aluri, K.; Zhang, X.; Gao, M.; Liu, J.; Indrakanti, R.; Schofield, S.; Kretschmer, P.; Brown, C. R.; Gupta, S.; Willoughby, J. L. S.; Boshar, J. A.; Jadhav, V.; Charisse, K.; Zimmermann, T.; Fitzgerald, K.; Manoharan, M.; Rajeev, K. G.; Akinc, A.; Hutabarat, R.; Maier, M. A. Impact of Enhanced Metabolic Stability on Pharmacokinetics and Pharmacodynamics of GalNAc–SiRNA Conjugates. *Nucleic Acids Res.* **2017**, *45* (19), 10969–10977.
(143) Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem., Int. Ed.* **2002**, *41* (14), 2596–2599.
(144) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase:[1, 2, 3]-Triazoles by Regiospecific Copper (I)-Catalyzed 1, 3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, *67* (9), 3057–3064.
(145) El-Sagheer, A. H.; Brown, T. New Strategy for the Synthesis of Chemically Modified RNA Constructs Exemplified by Hairpin and Hammerhead Ribozymes. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (35), 15329–15334.
(146) Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. Capturing the Structure of a Catalytic RNA Intermediate: The Hammerhead Ribozyme. *Science* **1996**, *274* (5295), 2065 LP – 2069.
(147) Scott, W. G.; Finch, J. T.; Klug, A. The Crystal Structure of an AII-RNAhammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. *Cell* **1995**, *81* (7), 991–1002.
(148) Taemaitree, L.; Shivalingam, A.; El-Sagheer, A. H.; Brown, T. An Artificial Triazole Backbone Linkage Provides a Split-and-Click Strategy to Bioactive Chemically Modified CRISPR SgRNA. *Nat. Commun.* **2019**, *10* (1), 1610.
(149) Varizhuk, A. M.; Kaluzhny, D. N.; Novikov, R. A.; Chizhov, A. O.; Smirnov, I. P.; Chuvilin, A. N.; Tatarinova, O. N.; Fisunov, G. Y.; Pozmogova, G. E.; Florentiev, V. L.; Zelinsky, N. D. Synthesis of Triazole-Linked Oligonucleotides with High Affinity to DNA Complements and an Analysis of Their Compatibility with Biosystems. *J. Org. Chem.* **2013**, *78* (12), 5964–5969.
(150) Birts, C. N.; Sanzone, A. P.; El-Sagheer, A. H.; Blaydes, J. P.; Brown, T.; Tavassoli, A. Transcription of Click-Linked DNA in Human Cells. *Angew. Chem., Int. Ed.* **2014**, *53* (9), 2362–2365.
(151) El-Sagheer, A. H.; Brown, T. Synthesis and Polymerase Chain Reaction Amplification of DNA Strands Containing an Unnatural Triazole Linkage. *J. Am. Chem. Soc.* **2009**, *131* (11), 3958–3964.
(152) El-Sagheer, A. H.; Brown, T. Efficient RNA Synthesis by in Vitro Transcription of a Triazole-

Modified DNA Template. *Chem. Commun.* **2011**, *47* (44), 12057–12058.
(153) El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T. Biocompatible Artificial DNA Linker That Is Read through by DNA Polymerases and Is Functional in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (28), 11338–11343.
(154) Kukwikila, M.; Gale, N.; El-Sagheer, A. H.; Brown, T.; Tavassoli, A. Assembly of a Biocompatible Triazole-Linked Gene by One-Pot Click-DNA Ligation. *Nat. Chem.* **2017**, *9* (11), 1089–1098.
(155) Sanzone, A. P.; El-Sagheer, A. H.; Brown, T.; Tavassoli, A. Assessing the Biocompatibility of Click-Linked DNA in Escherichia Coli. *Nucleic Acids Res.* **2012**, *40* (20), 10567–10575.
(156) Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. Evaluation of 2'-Modified Oligonucleotides Containing 2'-Deoxy Gaps as Antisense Inhibitors of Gene Expression. *J. Biol. Chem.* **1993**, *268* (19), 14514–14522.
(157) Wahlestedt, C.; Salmi, P.; Good, L.; Kela, J.; Johnsson, T.; Hokfelt, T.; Broberger, C.; Porreca, F.; Lai, J.; Ren, K.; Ossipov, M.; Koshkin, A.; Jakobsen, N.; Skouv, J.; Oerum, H.; Jacobsen, M. H.; Wengel, J. Potent and Nontoxic Antisense Oligonucleotides Containing Locked Nucleic Acids. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (10), 5633–5638.
(158) Miller, P. S. Non-Ionic Antisense Oligonucleotides. In *Oligodeoxynucleotides: Antisense Inhibitors of Gene Expression*; Cohen, J. S., Ed.; Macmillan Education UK: London, 1989; pp 79–95.
(159) Jaroszewski, J. W.; Cohen, J. S. Cellular Uptake of Antisense Oligodeoxynucleotides. *Adv. Drug Deliv. Rev.* **1991**, *6* (3), 235–250.
(160) Mungall, W. S.; Kaiser, J. K. Carbamate Analogs of Oligonucleotides. *J. Org. Chem.* **1977**, *42* (4), 703–706.
(161) Kang, H.; Chou, P.; Johnson, W. C.; Weller, D.; Huang, S.; Summerton, J. E. Stacking Interactions of ApA Analogues with Modified Backbones. *Biopolymers* **1992**, *32* (10), 1351–1363.
(162) Coull, J. M.; Carlson, D. V; Weith, H. L. Synthesis and Characterization of a Carbamate-Linked Oligonucleoside. *Tetrahedron Lett.* **1987**, *28* (7), 745–748.
(163) Kutterer, K. M. K.; Just, G. Synthesis and Hybridization Studies of Urea and Carbamate Linked Thymidine Dimers Incorporated into Oligodeoxynucleosides. *Bioorg. Med. Chem. Lett.* **1994**, *4* (3), 435–438.
(164) Stirchak, E. P.; Summerton, J. E.; Weller, D. D. Uncharged Stereoregular Nucleic Acid Analogs: 2. Morpholino Nucleoside Oligomers with Carbamate Internucleoside Linkages. *Nucleic Acids Res.* **1989**, *17* (15), 6129–6141.
(165) Kumar, P.; El-Sagheer, A. H.; Truong, L.; Brown, T. Locked Nucleic Acid (LNA) Enhances Binding Affinity of Triazole-Linked DNA towards RNA. *Chem. Commun.* **2017**, *53* (63), 8910–8913.
(166) Sharma, V. K.; Singh, S. K.; Krishnamurthy, P. M.; Alterman, J. F.; Haraszti, R. A.; Khvorova, A.; Prasad, A. K.; Watts, J. K. Synthesis and Biological Properties of Triazole-Linked Locked Nucleic Acid. *Chem. Commun.* **2017**, *53* (63), 8906–8909.
(167) Kumar, P.; Truong, L.; Baker, Y. R.; El-Sagheer, A. H.; Brown, T. Synthesis, Affinity for Complementary RNA and DNA, and Enzymatic Stability of Triazole-Linked Locked Nucleic Acids (t-LNAs). *ACS Omega* **2018**, *3* (6), 6976–6987.
(168) Lauritsen, A.; Wengel, J. Oligodeoxynucleotides Containing Amide-Linked LNA-Type Dinucleotides: Synthesis and High-Affinity Nucleic Acid Hybridization. *Chem. Commun.* **2002**, No. 5, 530–531.
(169) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.; In, Y.; Ishida, T.; Imanishi, T. Synthesis of 2′-O,4′-C-Methyleneuridine and -Cytidine. Novel Bicyclic Nucleosides Having a Fixed C3, -Endo Sugar Puckering. *Tetrahedron Lett.* **1997**, *38* (50), 8735–8738.
(170) Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. LNA (Locked Nucleic Acids): Synthesis of the Adenine, Cytosine, Guanine, 5-Methylcytosine, Thymine and Uracil Bicyclonucleoside Monomers, Oligomerisation, and Unprecedented Nucleic Acid Recognition. *Tetrahedron* **1998**, *54* (14), 3607–3630.
(171) Kurreck, J.; Wyszko, E.; Gillen, C.; Erdmann, V. A. Design of Antisense Oligonucleotides Stabilized by Locked Nucleic Acids. *Nucleic Acids Res.* **2002**, *30* (9), 1911–1918.
(172) Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. Structural Studies of LNA:RNA Duplexes by NMR: Conformations and Implications for RNase H Activity. *Chem. – A Eur. J.* **2000**, *6* (15), 2687–2695.
(173) K. Singh, S.; Wengel, J. Universality of LNA-Mediated High-Affinity Nucleic Acid

Recognition. *Chem. Commun.* **1998**, No. 12, 1247–1248.
(174) K. Singh, S.; A. Koshkin, A.; Wengel, J.; Nielsen, P. LNA (Locked Nucleic Acids): Synthesis and High-Affinity Nucleic Acid Recognition. *Chem. Commun.* **1998**, No. 4, 455–456.
(175) Hagedorn, P. H.; Persson, R.; Funder, E. D.; Albæk, N.; Diemer, S. L.; Hansen, D. J.; Møller, M. R.; Papargyri, N.; Christiansen, H.; Hansen, B. R.; Hansen, H. F.; Jensen, M. A.; Koch, T. Locked Nucleic Acid: Modality, Diversity, and Drug Discovery. *Drug Discov. Today* **2018**, *23* (1), 101–114.
(176) Van Daele, I.; Munier-Lehmann, H.; Froeyen, M.; Balzarini, J.; Van Calenbergh, S. Rational Design of 5′-Thiourea-Substituted α-Thymidine Analogues as Thymidine Monophosphate Kinase Inhibitors Capable of Inhibiting Mycobacterial Growth. *J. Med. Chem.* **2007**, *50* (22), 5281–5292.
(177) 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). *J. Org. Chem.* **2001**, *66* (25), 8504–8512.
(178) Leuck, M.; Vagle, K. E.; Shawn Roach, J.; Wolter, A. A Novel Reagent for the Chemical Phosphorylation of Oligonucleotides. *Tetrahedron Lett.* **2004**, *45* (2), 321–324.
(179) Prykota, T. I.; Pfleiderer, W. Nucleotides Part LXXX: Synthesis of 3′-O Fluorescence Labeled Thymidine Derivatives and Their 5′-O-Triphosphates. *Nucleosides, Nucleotides Nucleic Acids* **2011**, *30* (7–8), 544–551.
(180) Kulikov, A.; Arumugam, S.; Popik, V. V. Photolabile Protection of Alcohols, Phenols, and Carboxylic Acids with 3-Hydroxy-2-Naphthalenemethanol. *J. Org. Chem.* **2008**, *73* (19), 7611–7615.
(181) Moriguchi, T.; Sekine, M.; Shinozuka, K. Novel Method of the Synthesis and Hybridization Properties of an Oligonucleotide Containing Non-Ionic Diisopropylsilyl Internucleotide Linkage. *Bioorg. Med. Chem.* **2013**, *21* (24), 8013–8018.
(182) 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. *Org. Biomol. Chem.* **2019**, *17* (21), 5341–5348.
(183) Petersen, M.; Bondensgaard, K.; Wengel, J.; Jacobsen, J. P. Locked Nucleic Acid (LNA) Recognition of RNA: NMR Solution Structures of LNA:RNA Hybrids. *J. Am. Chem. Soc.* **2002**, *124* (21), 5974–5982.
(184) Grunder-Klotz, E.; Just, G. Synthesis of Deoxycytidine (DC) Building Blocks for Amide-Linked Dimers or Oligomers: Example of a Deoxycytidine-Amide-Thymidine Dimer. *Nucleosides Nucleotides* **1994**, *13* (9), 1829–1841.
(185) Selvam, C.; Thomas, S.; Abbott, J.; Kennedy, S. D.; Rozners, E. Amides as Excellent Mimics of Phosphate Linkages in RNA. *Angew. Chem., Int. Ed.* **2011**, *50* (9), 2068–2070.
(186) Hardcastle, T.; Novosjolova, I.; Kotikam, V.; Cheruiyot, S. K.; Mutisya, D.; Kennedy, S. D.; Egli, M.; Kelley, M. L.; Smith, A. V. B.; Rozners, E. A Single Amide Linkage in the Passenger Strand Suppresses Its Activity and Enhances Guide Strand Targeting of SiRNAs. *ACS Chem. Biol.* **2018**, *13* (3), 533–536.
(187) Mutisya, D.; Selvam, C.; Lunstad, B. D.; Pallan, P. S.; Haas, A.; Leake, D.; Egli, M.; Rozners, E. Amides Are Excellent Mimics of Phosphate Internucleoside Linkages and Are Well Tolerated in Short Interfering RNAs. *Nucleic Acids Res.* **2014**, *42* (10), 6542–6551.
(188) Rozners, E.; Katkevica, D.; Bizdena, E.; Strömberg, R. Synthesis and Properties of RNA Analogues Having Amides as Interuridine Linkages at Selected Positions. *J. Am. Chem. Soc.* **2003**, *125* (40), 12125–12136.
(189) Tanui, P.; Kennedy, S. D.; Lunstad, B. D.; Haas, A.; Leake, D.; Rozners, E. Synthesis, Biophysical Studies and RNA Interference Activity of RNA Having Three Consecutive Amide Linkages. *Org. Biomol. Chem.* **2014**, *12* (8), 1207–1210.
(190) Kotikam, V.; Viel, J. A.; Rozners, E. Synthesis and Biological Activity of Short Interfering RNAs Having Several Consecutive Amide Internucleoside Linkages. *Chem. - Eur. J.* **2020**, *26* (3), 685–690.
(191) De Mesmaeker, A.; Jouanno, C.; Wolf, R. M.; Wendeborn, S. Replacement of the Phosphodiester Linkage in Oligonncleotides by an Amide: Effect of Backbone Length on Duplex Stability with RNA Complement. *Bioorganic Med. Chem. Lett.* **1997**, *7* (4), 447–452.
(192) Fritsch, V.; De Mesmaeker, A.; Waldner, A.; Lebreton, J.; Blommers, M. J. J.; Wolf, R. M. Molecular Mechanics and Dynamics Studies on Two Structurally Related Amide-Modified DNA Backbones for Antisense Technology. *Bioorganic Med. Chem.* **1995**, *3* (3), 321–335.
(193) Mutisya, D.; Selvam, C.; Lunstad, B. D.; Pallan, P. S.; Haas, A.; Leake, D.; Egli, M.; Rozners, E. Amides Are Excellent Mimics of Phosphate Internucleoside Linkages and Are Well

Tolerated in Short Interfering RNAs. *Nucleic Acids Res.* **2014**, *42* (10), 6542–6551.
(194) Shivalingam, A.; Tyburn, A. E. S.; El-Sagheer, A. H.; Brown, T. Molecular Requirements of High-Fidelity Replication-Competent DNA Backbones for Orthogonal Chemical Ligation. *J. Am. Chem. Soc.* **2017**, *139* (4), 1575–1583.
(195) Mutisya, D.; Hardcastle, T.; Cheruiyot, S. K.; Pallan, P. S.; Kennedy, S. D.; Egli, M.; Kelley, M. L.; Smith, A. van B.; Rozners, E. Amide Linkages Mimic Phosphates in RNA Interactions with Proteins and Are Well Tolerated in the Guide Strand of Short Interfering RNAs. *Nucleic Acids Res.* **2017**, *45* (14), 8142–8155.
(196) Chu, C. K.; Doboszewski, B.; Schmidt, W.; Roey, P. Van; Ullas, G. V. Synthesis of Pyrimidine 3′-Allyl-2′,3′dideoxyribonucleosides by Free-Radical Coupling. *J. Org. Chem.* **1989**, *54* (11), 2767–2769.
(197) Fiandor, J.; Y. Tam, S. Synthesis of 3′-Deoxy-3′-(2-Propynyl)Thymidine and 3′-Cyanomethyl-3′-Deoxythymidine, Analogs of AZT. *Tetrahedron Lett.* **1990**, *31* (5), 597–600.
(198) Bhat, B.; Swayze, E. E.; Wheeler, P.; Dimock, S.; Perbost, M.; Sanghvi, Y. S. Synthesis of Novel Nucleic Acid Mimics via the Stereoselective Intermolecular Radical Coupling of 3′-Iodo Nucleosides and Formaldoximes. *J. Org. Chem.* **1996**, *61* (23), 8186–8199.
(199) Hattori, H.; Nozawa, E.; Iino, T.; Yoshimura, Y.; Shuto, S.; Shimamoto, Y.; Nomura, M.; Fukushima, M.; Tanaka, M.; Sasaki, T.; Matsuda, A. Nucleosides and Nucleotides, 175. Structural Requirements of the Sugar Moiety for the Antitumor Activities of New Nucleoside Antimetabolites, 1-(3- C-Ethynyl-β-d-Ribo-Pentofuranosyl)Cytosine and -Uracil. *J. Med. Chem.* **1998**, *41* (15), 2892–2902.
(200) Jones, A. S.; Williamson, A. R.; Winkley, M. The Chromium Trioxide-Pyridine Oxidation of Deoxyribonucleosides and Deoxyribonucleotides. *Carbohydr. Res.* **1965**, *1* (3), 187–195.
(201) Pfitzner, K. E.; Moffatt, J. G. Sulfoxide-Carbodiimide Reactions. I. A Facile Oxidation of Alcohols. *J. Am. Chem. Soc.* **1965**, *87* (24), 5661–5670.
(202) Calvo-Mateo, A.; Camarasa, M. J.; Díaz-Ortíz, A.; De las Heras, F. G.; Alemany, A. Synthesis of 3'-c-Cyano-3'-Deoxy- Pentofuranosylthymine Nucleosides. *Tetrahedron* **1988**, *44* (15), 4895–4903.
(203) Robins, M. J.; Sarker, S.; Xie, M.; Zhang, W.; Peterson, M. A. Synthesis of 2',3'-Fused (3.3.0) γ-Butyrolactone-Nucleosides and Coupling with Amino-Nucleosides to Give Amide-Linked Nucleotide-Dimer Analogues. *Tetrahedron Lett.* **1996**, *37* (23), 3921–3924.
(204) Robins, M. J.; Doboszewsk, B.; Timoshchuk, V. A.; Peterson, M. A. Glucose-Derived 3'-(Carboxymethyl)-3'-Deoxyribonucleosides and 2',3'- Lactones as Synthetic Precursors for Amide-Linked Oligonucleotide Analogues. *J. Org. Chem.* **2000**, *65* (10), 2939–2945.
(205) Peterson, M. A.; Nilsson, B. L.; Sarker, S.; Doboszewski, B.; Zhang, W.; Robins, M. J. Amide-Linked Ribonucleoside Dimers Derived from 5'-Amino-5'-Deoxy- and 3'-(Carboxymethyl)-3'-Deoxynucleoside Precursors. *J. Org. Chem.* **1999**, *64* (22), 8183–8192.
(206) Kotikam, V.; Rozners, E. Concurrent Hydrogenation of Three Functional Groups Enables Synthesis of C3′-Homologated Nucleoside Amino Acids. *Org. Lett.* **2017**, *19* (15), 4122–4125.
(207) Wendeborn, S.; Nussbaumer, H.; Robert, F.; Jörg, M.; Pachlatko, J. P. Towards Novel Amide-Modified Oligonucleotides: Asymmetric Synthesis of 5′-(S)-Methyl-3′-Carboxymethylene-3′-Deoxythymidine. *Tetrahedron Lett.* **2002**, *43* (31), 5461–5464.
(208) Smith, M.; Rammler, D. H.; Goldberg, I. H.; Khorana, H. G. Studies on Polynucleotides. XIV.1 Specific Synthesis of the C3″-C5″ Interribonucleotide Linkage. Syntheses of Uridylyl-(3″→5″)-Uridine and Uridylyl-(3″→5″)-Adenosine2. *J. Am. Chem. Soc.* **1962**, *84* (3), 430–440.
(209) Polushin, N. N.; Chen, B. C.; Anderson, L. W.; Cohen, J. S. Synthesis and Characterization of Imidazoyl-Linked Synthons and 3'-Conjugated Thymidine Derivatives. *J. Org. Chem.* **1993**, *58* (17), 4606–4613.
(210) Dubowchik, G. M.; Radia, S. Monomethoxytrityl (MMT) as a Versatile Amino Protecting Group for Complex Prodrugs of Anticancer Compounds Sensitive to Strong Acids, Bases and Nucleophiles. *Tetrahedron Lett.* **1997**, *38* (30), 5257–5260.
(211) Lima, W. F.; Crooke, S. T. Binding Affinity and Specificity of Escherichia Coli RNase H1: Impact on the Kinetics of Catalysis of Antisense Oligonucleotide-RNA Hybrids. *Biochemistry* **1997**, *36* (2), 390–398.
(212) Lima, W. F.; Nichols, J. G.; Wu, H.; Prakash, T. P.; Migawa, M. T.; Wyrzykiewicz, T. K.; Bhat, B.; Crooke, S. T. Structural Requirements at the Catalytic Site of the Heteroduplex Substrate for Human RNase H1 Catalysis. *J. Biol. Chem.* **2004**, *279* (35), 36317–36326.
(213) Minasov, G.; Teplova, M.; Nielsen, P.; Wengel, J.; Egli, M. Structural Basis of Cleavage by

RNase H of Hybrids of Arabinonucleic Acids and RNA. *Biochemistry* **2000**, *39* (13), 3525–3532.
(214) Fedoroff, O. Y.; Salazar, M.; Reid, B. R. Structure of a DNA : RNA Hybrid Duplex: Why RNase H Does Not Cleave Pure RNA. *J. Mol. Biol.* **1993**, *233* (3), 509–523.
(215) Zamaratski, E.; Pradeepkumar, P. I.; Chattopadhyaya, J. A Critical Survey of the Structure-Function of the Antisense Oligo/RNA Heteroduplex as Substrate for RNase H. *J. Biochem. Biophys. Methods* **2001**, *48* (3), 189–208.
(216) Davisy, R. R.; Shaban, N. M.; Perrino, F. W.; Hollis, T. Crystal Structure of RNA-DNA Duplex Provides Insight into Conformational Changes Induced by RNase H Binding. *Cell Cycle* **2015**, *14* (4), 668–673.
(217) Nowotny, M.; Gaidamakov, S. A.; Ghirlando, R.; Cerritelli, S. M.; Crouch, R. J.; Yang, W. Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Mol. Cell* **2007**, *28* (2), 264–276.
(218) Rait, V. K.; Shaw, B. R. Boranophosphates Support the RNase H Cleavage of Polyribonucleotides. *Antisense Nucleic Acid Drug Dev.* **1999**, *9* (1), 53–60.
(219) Šípová, H.; Špringer, T.; Rejman, D.; Šimák, O.; Petrová, M.; Novák, P.; Rosenbergová, Š.; Páv, O.; Liboska, R.; Barvík, I.; Štěpánek, J.; Rosenberg, I.; Homola, J.; Šípová, H.; Špringer, T.; Rejman, D.; Šimák, O.; Petrová, M.; Novák, P.; Rosenbergová, Š.; Páv, O.; Liboska, R.; Barvík, I.; Štěpánek, J.; Rosenberg, I.; Homola, J. 5′-O-Methylphosphonate Nucleic Acids—New Modified DNAs That Increase the Escherichia Coli RNase H Cleavage Rate of Hybrid Duplexes. *Nucleic Acids Res.* **2014**, *42* (8), 5378–5389.
(220) Miroshnichenko, S. K.; Patutina, O. A.; Burakova, E. A.; Chelobanov, B. P.; Fokina, A. A.; Vlassov, V. V.; Altman, S.; Zenkova, M. A.; Stetsenko, D. A. Mesyl Phosphoramidate Antisense Oligonucleotides as an Alternative to Phosphorothioates with Improved Biochemical and Biological Properties. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (4), 1229–1234.
(221) Patutina, O. A.; Gaponova (Miroshnichenko), S. K.; Sen'kova, A. V; Savin, I. A.; Gladkikh, D. V; Burakova, E. A.; Fokina, A. A.; Maslov, M. A.; Shmendel', E. V; Wood, M. J. A.; Vlassov, V. V; Altman, S.; Stetsenko, D. A.; Zenkova, M. A. Mesyl Phosphoramidate Backbone Modified Antisense Oligonucleotides Targeting MiR-21 with Enhanced in Vivo Therapeutic Potency. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (51), 32370–32379.
(222) Sheehan, D.; Lunstad, B.; Yamada, C. M.; Stell, B. G.; Caruthers, M. H.; Dellinger, D. J. Biochemical Properties of Phosphonoacetate and Thiophosphonoacetate Oligodeoxyribonucleotides. *Nucleic Acids Res.* **2003**, *31* (14), 4109–4118.
(223) Noronha, A.; Brukner, I.; Borkow, G.; Arion, D.; Parniak, M. A. Hybrids of RNA and Arabinonucleic Acids (ANA and 2′F-ANA) Are Substrates of Ribonuclease H. *J. Am. Chem. Soc.* **1998**, *120* (49), 12976–12977.
(224) Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P. Cyclohexene Nucleic Acids (CeNA): Serum Stable Oligonucleotides That Activate RNase H and Increase Duplex Stability with Complementary RNA. *J. Am. Chem. Soc.* **2000**, *122* (36), 8595–8602.
(225) Jones, G. D.; Lesnik, E. A.; Owens, S. R.; Risen, L. M.; Walker, R. T. Investigation of Some Properties of Oligodeoxynucleotides Containing 4'-Thio-2'-Deoxynucleotides: Duplex Hybridization and Nuclease Sensitivity. *Nucleic Acids Res.* **1996**, *24* (21), 4117–4122.
(226) Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. Evaluation of 2'-Modified Oligonucleotides Containing 2'-Deoxy Gaps as Antisense Inhibitors of Gene Expression. *J. Biol. Chem.* **1993**, *286* (19), 14514–14522.
(227) Nucleic Acid Hybridization and Amplification Method for Detection of Specific Sequences in Which a Complementary Labeled Nucleic Acid Probe Is Cleaved. US5403711A, July 6, 1993.
(228) Stanton, R.; Sciabola, S.; Salatto, C.; Weng, Y.; Moshinsky, D.; Little, J.; Walters, E.; Kreeger, J.; Dimattia, D.; Chen, T.; Clark, T.; Liu, M.; Qian, J.; Roy, M.; Dullea, R. Chemical Modification Study of Antisense Gapmers. *Nucleic Acid Ther.* **2012**, *22* (5), 344–359.
(229) Khvorova, A.; Watts, J. K. The Chemical Evolution of Oligonucleotide Therapies of Clinical Utility. *Nat. Biotechnol.* **2017**, *35* (3), 238–248.
(230) Giles, R. V.; Tidd, D. M. Increased Specificity for Antisense Oligodeoxynucleotide Targeting of RNA Cleavage bY RNase H Using Chimeric Methylphosphonodiester/Phosphodiester Structures. *Nucleic Acids Res.* **1992**, *20* (4), 763–770.
(231) Shoji, Y.; Akhtar, S.; Periasamy, A.; Herman, B.; Juliano, R. L. Mechansism of Cellular

Uptake of Modified Oligodeoxynucleotides Containing Methylphosphonate Linkages. *Nucleic Acids Res.* **1991**, *19* (20), 5543–5550.
(232) Wittung, P.; Kajanus, J.; Edwards, K.; Nielsen, P.; Nordén, B.; Malmström, B. G. Phospholipid Membrane Permeability of Peptide Nucleic Acid. *FEBS Lett.* **1995**, *365* (1), 27–29.
(233) Noble, S. A.; Bonham, M. A.; Bisi, J. E.; Bruckenstein, D. A.; Brown, P. H.; Brown, S. C.; Cadilla, R.; Gaul, M. D.; Hanvey, J. C.; Fred Hassman, C.; Josey, J. A.; Luzzio, M. J.; Myers, P. M.; Pipe, A. J.; Ricca, D. J.; Su, C. W.; Stevenson, C. L.; Thomson, S. A.; Wiethe, R. W.; Babiss, L. E. Impact of Biophysical Parameters on the Biological Assessment of Peptide Nucleic Acids, Antisense Inhibitors of Gene Expression. *Drug Dev. Res.* **1995**, *34* (2), 184–195.
(234) Krishna, H.; Caruthers, M. H. Alkynyl Phosphonate DNA: A Versatile “Click”Able Backbone for DNA-Based Biological Applications. *J. Am. Chem. Soc.* **2012**, *134* (28), 11618–11631.
(235) Vivès, E.; Dell’Aquila, C.; Bologna, J. C.; Morvan, F.; Rayner, B.; Imbach, J. L. Lipophilic Pro-Oligonucleotides Are Rapidly and Efficiently Internalized in HeLa Cells. *Nucleic Acids Res.* **1999**, *27* (20), 4071–4076.
(236) Meade, B. R.; Gogoi, K.; Hamil, A. S.; Palm-Apergi, C.; Van Den Berg, A.; Hagopian, J. C.; Springer, A. D.; Eguchi, A.; Kacsinta, A. D.; Dowdy, C. F.; Presente, A.; Lönn, P.; Kaulich, M.; Yoshioka, N.; Gros, E.; Cui, X. S.; Dowdy, S. F. Efficient Delivery of RNAi Prodrugs Containing Reversible Charge-Neutralizing Phosphotriester Backbone Modifications. *Nat. Biotechnol.* **2014**, *32* (12), 1256–1261.
(237) Tanabe, K.; Ando, Y.; Hara, D.; Ito, T.; Nishimoto, S. I. Oligonucleotides with Consecutive Alkylated Phosphate Units: Aggregation Characteristics and Drug Transport into Living Cells. *RSC Adv.* **2014**, *4* (26), 13367–13370.
(238) Hayashi, J.; Samezawa, Y.; Ochi, Y.; Wada, S. ichi; Urata, H. Syntheses of Prodrug-Type Phosphotriester Oligonucleotides Responsive to Intracellular Reducing Environment for Improvement of Cell Membrane Permeability and Nuclease Resistance. *Bioorganic Med. Chem. Lett.* **2017**, *27* (14), 3135–3138.
(239) Saneyoshi, H.; Ohta, T.; Hiyoshi, Y.; Saneyoshi, T.; Ono, A. Design, Synthesis, and Cellular Uptake of Oligonucleotides Bearing Glutathione-Labile Protecting Groups. *Org. Lett.* **2019**, *21* (4), 862–866.
(240) Shu, Z.; Tanaka, I.; Ota, A.; Fushihara, D.; Abe, N.; Kawaguchi, S.; Nakamoto, K.; Tomoike, F.; Tada, S.; Ito, Y.; Kimura, Y.; Abe, H. Disulfide-Unit Conjugation Enables Ultrafast Cytosolic Internalization of Antisense DNA and SiRNA. *Angew. Chem., Int. Ed.* **2019**, *58* (20), 6611–6615.
(241) Laurent, Q.; Martinent, R.; Moreau, D.; Winssinger, N.; Sakai, N.; Matile, S. Oligonucleotide Phosphorothioates Enter Cells by Thiol-Mediated Uptake. *Angew. Chem., Int. Ed.* **2021**, *60* (35), 19102–19106.
(242) Kiełpiński, Ł. J.; Hagedorn, P. H.; Lindow, M.; Vinther, J. RNase H Sequence Preferences Influence Antisense Oligonucleotide Efficiency. *Nucleic Acids Res.* **2017**, *45* (22), 12932–12944.
(243) Nowotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W. Crystal Structures of RNase H Bound to an RNA/DNA Hybrid: Substrate Specificity and Metal-Dependent Catalysis. *Cell* **2005**, *121* (7), 1005–1016.
(244) Reddy, P. M.; Bruice, T. C. Solid-Phase Synthesis of Positively Charged Deoxynucleic Guanidine (DNG) Tethering a Hoechst 33258 Analogue: Triplex and Duplex Stabilization by Simultaneous Minor Groove Binding. *J. Am. Chem. Soc.* **2004**, *126* (12), 3736–3747.
(245) Brandsen, B. M.; Hesser, A. R.; Castner, M. A.; Chandra, M.; Silverman, S. K. DNA-Catalyzed Hydrolysis of Esters and Aromatic Amides. *J. Am. Chem. Soc.* **2013**, *135* (43), 16014–16017.
(246) Kotikam, V.; Viel, J. A.; Rozners, E. Synthesis and Biological Activity of Short Interfering RNAs Having Several Consecutive Amide Internucleoside Linkages. *Chem. – A Eur. J.* **2019**, chem.201903754.
(247) Integrated DNA Technologies Inc. *coupling efficiency for longer, full-length oligos | IDT*. Retrieved from https://eu.idtdna.com/pages/education/decoded/article/getting-enough-full-length-oligo- (accessed Nov 6, 2021).
(248) Sipa, K.; Sochacka, E.; Kazmierczak-Baranska, J.; Maszewska, M.; Janicka, M.; Nowak, G.; Nawrot, B. Effect of Base Modifications on Structure, Thermodynamic Stability, and Gene Silencing Activity of Short Interfering RNA. *RNA* **2007**, *13* (8), 1301–1316.
(249) Chernikov, I. V; Vlassov, V. V; Chernolovskaya, E. L. Current Development of SiRNA

Bioconjugates: From Research to the Clinic. *Front. Pharmacol.* **2019**, *10*, 444.
(250) Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. Evaluation of 2'-Modified Oligonucleotides Containing 2'-Deoxy Gaps as Antisense Inhibitors of Gene Expression. *J. Biol. Chem.* **1993**, *268* (19), 14514–14522.
(251) Freier, S.; Altmann, K.-H. The Ups and Downs of Nucleic Acid Duplex Stability: Structure-Stability Studies on Chemically-Modified DNA:RNA Duplexes. *Nucleic Acids Res.* **1997**, *25* (22), 4429–4443.
(252) Tu, G. C.; Cao, Q. N.; Zhou, F.; Israel, Y. Tetranucleotide GGGA Motif in Primary RNA Transcripts: Novel Target Site for Antisense Design. *J. Biol. Chem.* **1998**, *273* (39), 25125–25131.
(253) Cerritelli, S. M.; Crouch, R. J. Ribonuclease H: The Enzymes in Eukaryotes. *FEBS J.* **2009**, *276* (6), 1494–1505.
(254) Malchére, C.; Verheijen, J.; Van der Laan, S.; Bastide, L.; Van Boom, J.; Lebleu, B.; Robbins, I. A Short Phosphodiester Window Is Sufficient to Direct RNase H-Dependent RNA Cleavage by Antisense Peptide Nucleic Acid. *Antisense Nucleic Acid Drug Dev.* **2011**, *10* (6), 463–468.
(255) Debacker, A. J.; Sharma, V. K.; Meda Krishnamurthy, P.; O'Reilly, D.; Greenhill, R.; Watts, J. K. Next-Generation Peptide Nucleic Acid Chimeras Exhibit High Affinity and Potent Gene Silencing. *Biochemistry* **2019**, *58* (6), 582–589.
(256) Frieden, M.; Christensen, S. M.; Mikkelsen, N. D.; Rosenbohm, C.; Thrue, C. A.; Westergaard, M.; Hansen, H. F.; Ørum, H.; Koch, T. Expanding the Design Horizon of Antisense Oligonucleotides with Alpha-l-LNA. *Nucleic Acids Res.* **2003**, *31* (21), 6365–6372.
(257) Straarup, E. M.; Fisker, N.; Hedtjärn, M.; Lindholm, M. W.; Rosenbohm, C.; Aarup, V.; Hansen, H. F.; Ørum, H.; Hansen, J. B. R.; Koch, T. Short Locked Nucleic Acid Antisense Oligonucleotides Potently Reduce Apolipoprotein B MRNA and Serum Cholesterol in Mice and Non-Human Primates. *Nucleic Acids Res.* **2010**, *38* (20), 7100–7111.
(258) Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Wancewicz, E. V; Witchell, D.; Swayze, E. E. Short Antisense Oligonucleotides with Novel 2′−4′ Conformationaly Restricted Nucleoside Analogues Show Improved Potency without Increased Toxicity in Animals. *J. Med. Chem.* **2009**, *52* (1), 10–13.
(259) Pedersen, L.; Hagedorn, P. H.; Lindholm, M. W.; Lindow, M. A Kinetic Model Explains Why Shorter and Less Affine Enzyme-Recruiting Oligonucleotides Can Be More Potent. *Mol. Ther. - Nucleic Acids* **2014**, *3*.
(260) Li, F.; Sarkhel, S.; Wilds, C. J.; Wawrzak, Z.; Prakash, T. P.; Manoharan, M.; Egli, M. 2′-Fluoroarabino- and Arabinonucleic Acid Show Different Conformations, Resulting in Deviating RNA Affinities and Processing of Their Heteroduplexes with RNA by RNase H. *Biochemistry* **2006**, *45* (13), 4141–4152.
(261) Hagedorn, P. H.; Hansen, B. R.; Koch, T.; Lindow, M. Managing the Sequence-Specificity of Antisense Oligonucleotides in Drug Discovery. *Nucleic Acids Res.* **2017**, *45* (5), 2262–2282.
(262) Hagedorn, P. H.; Pontoppidan, M.; Bisgaard, T. S.; Berrera, M.; Dieckmann, A.; Ebeling, M.; Møller, M. R.; Hudlebusch, H.; Jensen, M. L.; Hansen, H. F.; Koch, T.; Lindow, M. Identifying and Avoiding Off-Target Effects of RNase H-Dependent Antisense Oligonucleotides in Mice. *Nucleic Acids Res.* **2018**, *46* (11), 5366–5380.
(263) Papargyri, N.; Pontoppidan, M.; Andersen, M. R.; Koch, T.; Hagedorn, P. H. Chemical Diversity of Locked Nucleic Acid-Modified Antisense Oligonucleotides Allows Optimization of Pharmaceutical Properties. *Mol. Ther. - Nucleic Acids* **2020**, *19*, 706–717.
(264) Martin, P. Ein Neuer Zugang Zu 2′-O-Alkylribonucleosiden Und Eigenschaften Deren Oligonucleotide. *Helv. Chim. Acta* **1995**, *78* (2), 486–504.
(265) Lim, J. P.; Gleeson, P. A. Macropinocytosis: An Endocytic Pathway for Internalising Large Gulps. *Immunol. Cell Biol.* **2011**, *89* (8), 836–843.
(266) Hu, Y. B.; Dammer, E. B.; Ren, R. J.; Wang, G. The Endosomal-Lysosomal System: From Acidification and Cargo Sorting to Neurodegeneration. *Transl. Neurodegener.* **2015**, *4* (18), 1–20.
(267) Bennet, M. A. Multi-Parameter Quantitative Mapping of Microfluidic Devices, University of Edinburgh, 2011.
(268) Murk, J. L. A. N.; Posthuma, G.; Koster, A. J.; Geuze, H. J.; Verkleij, A. J.; Kleijmeer, M. J.; Humbel, B. M. Influence of Aldehyde Fixation on the Morphology of Endosomes and Lysosomes: Quantitative Analysis and Electron Tomography. *J. Microsc.* **2003**, *212* (1), 81–

90.
(269) Albrecht, H.; Jones, G.; Moffatt, J. 3'-Deoxy-3'-(Dihydroxyphosphinylmethyl)Nucleosides. Isophosphonate Analogs of Nucleoside 3'-Phosphates. *J. Am. Chem. Soc.* **1970**, *92* (18), 5511–5513.
(270) Xavier, N. M.; Madeira, P. J. A.; Florêncio, M. H.; Rauter, A. P. Synthetic Approaches to Novel Thiosugar Scaffolds Containing α,β-Unsaturated Carbonyl Groups. *Eur. J. Org. Chem.* **2009**, *2009* (29), 4983–4991.
(271) Thermo Fisher Scientific. *QuantiGene RNA Assays for Gene Expression Profiling | Thermo Fisher Scientific - UK*. Retrieved from https://www.thermofisher.com/uk/en/home/life-science/gene-expression-analysis-genotyping/quantigene-rna-assays.html?ef_id=Cj0KCQiAgP6PBhDmARIsAPWMq6kSeDgAmp0hMJ_bFfNOi-jKrCitX7_oZeJBLo6RegkAreDYMwmx8cAaAnfSEALw_wcB:G:s&s_kwcid=AL!3652!3!572685280554!e!!g! (accessed Feb 6, 2022).
(272) Huisgen, R. 1,3-Dipolar Cycloadditions. Past and Future. *Angew. Chem., Int. Ed.* **1963**, *2* (10), 565–598.
(273) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chem., Int. Ed.* **2009**, *48* (38), 6974–6998.
(274) Pickens, C. J.; Johnson, S. N.; Pressnall, M. M.; Leon, M. A.; Berkland, C. J. Practical Considerations, Challenges, and Limitations of Bioconjugation via Azide–Alkyne Cycloaddition. *Bioconjug. Chem.* **2018**, *29* (3), 686–701.
(275) Thirumurugan, P.; Matosiuk, D.; Jozwiak, K. Click Chemistry for Drug Development and Diverse Chemical–Biology Applications. *Chem. Rev.* **2013**, *113* (7), 4905–4979.
(276) Lahann, J. *Click Chemistry: A Universal Ligation Strategy for Biotechnology and Materials Science*; John Wiley & Sons, Ltd Chichester, UK, 2009.
(277) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. Increasing the Efficacy of Bioorthogonal Click Reactions for Bioconjugation: A Comparative Study. *Angew. Chem., Int. Ed.* **2011**, *50* (35), 8051–8056.
(278) Hein, J. E.; Fokin, V. V. Copper-Catalyzed Azide–Alkyne Cycloaddition (CuAAC) and beyond: New Reactivity of Copper(i) Acetylides. *Chem. Soc. Rev.* **2010**, *39* (4), 1302–1315.
(279) Johansson, J. R.; Beke-Somfai, T.; Said Stålsmeden, A.; Kann, N. Ruthenium-Catalyzed Azide Alkyne Cycloaddition Reaction: Scope, Mechanism, and Applications. *Chem. Rev.* **2016**, *116* (23), 14726–14768.
(280) Boren, B. C.; Narayan, S.; Rasmussen, L. K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V. V. Ruthenium-Catalyzed Azide−Alkyne Cycloaddition: Scope and Mechanism. *J. Am. Chem. Soc.* **2008**, *130* (28), 8923–8930.
(281) Sultana, J.; Sarma, D. Ag-Catalyzed Azide-Alkyne Cycloaddition: Copper Free Approaches for Synthesis of 1,4-Disubstituted 1,2,3-Triazoles. *Catal. Rev.* **2020**, *62* (1), 96–117.
(282) Smith, C. D.; Greaney, M. F. Zinc Mediated Azide–Alkyne Ligation to 1,5- and 1,4,5-Substituted 1,2,3-Triazoles. *Org. Lett.* **2013**, *15* (18), 4826–4829.
(283) Kim, W. G.; Kang, M. E.; Lee, J. Bin; Jeon, M. H.; Lee, S.; Lee, J.; Choi, B.; Cal, P. M. S. D.; Kang, S.; Kee, J. M.; Bernardes, G. J. L.; Rohde, J. U.; Choe, W.; Hong, S. Y. Nickel-Catalyzed Azide-Alkyne Cycloaddition to Access 1,5-Disubstituted 1,2,3-Triazoles in Air and Water. *J. Am. Chem. Soc.* **2017**, *139* (35), 12121–12124.
(284) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem., Int. Ed.* **2001**, *40* (11), 2004–2021.
(285) Marrocchi, A.; Facchetti, A.; Lanari, D.; Santoro, S.; Vaccaro, L. Click-Chemistry Approaches to π-Conjugated Polymers for Organic Electronics Applications. *Chem. Sci.* **2016**, *7* (10), 6298–6308.
(286) Meldal, M.; Diness, F. Recent Fascinating Aspects of the CuAAC Click Reaction. *Trends Chem.* **2020**, *2* (6), 569–584.
(287) Liqun, J.; R., T. D.; Mohand, M.; Guy, B. Isolation of Bis(Copper) Key Intermediates in Cu-Catalyzed Azide-Alkyne "Click Reaction." *Sci. Adv.* **2021**, *1* (5), e1500304.
(288) Krasiński, A.; Fokin, V. V; Sharpless, K. B. Direct Synthesis of 1,5-Disubstituted-4-Magnesio-1,2,3-Triazoles, Revisited. *Org. Lett.* **2004**, *6* (8), 1237–1240.
(289) Himbert, G.; Frank, D.; Regit, M. Untersuchungen an Diazoverbindungen Und Aziden, XXV1) Azid-Additionen an (Silyäthinyl)-, (Germyläthinyl)- Und (Stannyläthinyl)Amine. *Chem. Ber.* **1976**, *109* (1), 370–394.
(290) Boyer, J.; Mack, C.; Goebel, N.; Morgan L, J. Reactions of Sodium Phenylacetylide and Sodium Alkoxide with Tosyl and Mesyl Azides. *J. Org. Chem.* **1958**, *23* (7), 1051–1053.

(291) Kwok, S. W.; Fotsing, J. R.; Fraser, R. J.; Rodionov, V. O.; Fokin, V. V. Transition-Metal-Free Catalytic Synthesis of 1,5-Diaryl-1,2,3-Triazoles. *Org. Lett.* **2010**, *12* (19), 4217–4219.
(292) Zhang, L.; Chen, X.; Xue, P.; Sun, H. H. Y.; Williams, I. D.; Sharpless, K. B.; Fokin, V. V; Jia, G. Ruthenium-Catalyzed Cycloaddition of Alkynes and Organic Azides. *J. Am. Chem. Soc.* **2005**, *127* (46), 15998–15999.
(293) Tala, S. R.; Singh, A.; Lensing, C. J.; Schnell, S. M.; Freeman, K. T.; Rocca, J. R.; Haskell-Luevano, C. 1,2,3-Triazole Rings as a Disulfide Bond Mimetic in Chimeric AGRP-Melanocortin Peptides: Design, Synthesis, and Functional Characterization. *ACS Chem. Neurosci.* **2018**, *9* (5), 1001–1013.
(294) Isobe, H.; Fujino, T.; Yamazaki, N.; Guillot-Nieckowski, M.; Nakamura, E. Triazole-Linked Analogue of Deoxyribonucleic Acid (TLDNA): Design, Synthesis, and Double-Strand Formation with Natural DNA. *Org. Lett.* **2008**, *10* (17), 3729–3732.
(295) Illumina. *Illumina Sequencing by Synthesis - YouTube*. Retrieved from https://www.youtube.com/watch?v=fCd6B5HRaZ8 (accessed Aug 12, 2021).
(296) Ali, M. M.; Li, F.; Zhang, Z.; Zhang, K.; Kang, D.-K.; Ankrum, J. A.; Le, X. C.; Zhao, W. Rolling Circle Amplification: A Versatile Tool for Chemical Biology, Materials Science and Medicine. *Chem. Soc. Rev.* **2014**, *43* (10), 3324–3341.
(297) Eisenhuth, R.; Richert, C. Convenient Syntheses of 3′-Amino-2′,3′-Dideoxynucleosides, Their 5′-Monophosphates, and 3′-Aminoterminal Oligodeoxynucleotide Primers. *J. Org. Chem.* **2009**, *74* (1), 26–37.
(298) Kawagoe, N.; Kasori, Y.; Hasegawa, T. Highly C6-Selective and Quantitative Modification of Cellulose: Nucleoside-Appended Celluloses to Solubilize Single Walled Carbon Nanotubes. *Cellulose* **2011**, *18* (1), 83–93.
(299) Bürckstümmer, E.; Carell, T. Synthesis and Properties of DNA Containing a Spore Photoproduct Analog. *Chem. Commun.* **2008**, No. 34, 4037–4039.
(300) Szabó, T.; Stawiński, J. Synthesis and Some Conformational Features of the 5′-Deoxy-5′-Methylphosphonate Linked Dimer, 5′-Deoxy-5′-C-(Phosphonomethyl)Thymidin-3′-Yl (Thymidin-5′-Yl)Methylphosphonate [p(CH2)Tp(CH2)T]. *Tetrahedron* **1995**, *51* (14), 4145–4160.
(301) Kainz, P. The PCR Plateau Phase – towards an Understanding of Its Limitations. *Biochim. Biophys. Acta - Gene Struct. Expr.* **2000**, *1494* (1), 23–27.
(302) New England Biolabs. *A-Tailing with Taq Polymerase | NEB*. Retrieved from https://international.neb.com/protocols/2013/11/01/a-tailing-with-taq-polymerase (accessed Aug 5, 2021).
(303) New England Biolabs. *PCR Using Hot Start Taq DNA Polymerase (M9404) | NEB*. Retrieved from https://international.neb.com/protocols/2012/10/04/pcr-using-hot-start-taq-dna-polymerase-m0495 (accessed Aug 5, 2021).
(304) New England Biolabs. *PCR Protocol Phusion DNA Polymerase | NEB*. Retrieved from https://international.neb.com/protocols/0001/01/01/pcr-protocol-m0530 (accessed Feb 19, 2022).
(305) Li, Y.; Korolev, S.; Waksman, G. Crystal Structures of Open and Closed Forms of Binary and Ternary Complexes of the Large Fragment of Thermus Aquaticus DNA Polymerase I: Structural Basis for Nucleotide Incorporation. *EMBO J.* **1998**, *17* (24), 7514–7525.
(306) Liu, K.; Zheng, L.; Liu, Q.; de Vries, J. W.; Gerasimov, J. Y.; Herrmann, A. Nucleic Acid Chemistry in the Organic Phase: From Functionalized Oligonucleotides to DNA Side Chain Polymers. *J. Am. Chem. Soc.* **2014**, *136* (40), 14255–14262.
(307) Epple, S.; Modi, A.; Baker, Y. R.; Węgrzyn, E.; Traoré, D.; Wanat, P.; Tyburn, A. E. S.; Shivalingam, A.; Taemaitree, L.; El-Sagheer, A. H.; Brown, T. A New 1,5-Disubstituted Triazole DNA Backbone Mimic with Enhanced Polymerase Compatibility. *J. Am. Chem. Soc.* **2021**.
(308) 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. *Chem. Commun.* **2020**, *56* (41), 5496–5499.
(309) Damha, M. J.; Giannaris, P. A.; Zabarylo, S. V. An Improved Procedure for Derivatization of Controlled-Pore Glass Beads for Solid-Phase Oligonucleotide Synthesis. *Nucleic Acids Res.* **1990**, *18* (13), 3813–3821.
(310) Fonvielle, M.; Li De La Sierra-Gallay, I.; El-Sagheer, A. H.; Lecerf, M.; Patin, D.; Mellal, D.; Mayer, C.; Blanot, D.; Gale, N.; Brown, T.; Van Tilbeurgh, H.; Ethève-Quelquejeu, M.; Arthur, M. The Structure of FemX(Wv) in Complex with a Peptidyl-RNA Conjugate: Mechanism of Aminoacyl Transfer from Ala-tRNA(Ala) to Peptidoglycan Precursors. *Angew. Chem., Int.*

*Ed.* **2013**, *52* (28), 7278–7281.
(311) Baraniak, D.; Baranowski, D.; Ruszkowski, P.; Boryski, J. Nucleoside Dimers Analogues with a 1,2,3-Triazole Linkage: Conjugation of Floxuridine and Thymidine Provides Novel Tools for Cancer Treatment. Part II. *Nucleosides, Nucleotides Nucleic Acids* **2019**, *38* (11), 807–835.

# APPENDIX

# 1. List of Oligonucleotides

**Table A-1: List of oligonucleotides for Chapter II.**

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| ON9 | GCTTGCTTCGTTCC | 4197 | 4196 |
| ON10 | GCTTGCT$^{X}$TCGTTCC | 4160 | 4160 |
| ON11 | GCTTGC**T**$^{X}$TCGTTCC | 4188 | 4187 |
| ON12 | GCTTGCT$^{X}$**T**CGTTCC | 4188 | 4188 |
| ON13 | GCTTGC**T**$^{X}$**T**CGTTCC | 4216 | 4215 |
| ON14 | GCT$^{X}$TGCT$^{X}$TCGT$^{X}$TCC | 4086 | 4086 |
| ON15 | GC**T**$^{X}$TGC**T**$^{X}$TCG**T**$^{X}$TCC | 4170 | 4167 |
| ON16 | GCT$^{X}$**T**GCT$^{X}$**T**CGT$^{X}$**T**CC | 4170 | 4170 |
| ON17 | GC**T**$^{X}$**T**GC**T**$^{X}$**T**CG**T**$^{X}$**T**CC | 4254 | 4255 |
| DNA target | GGAACGAAGCAAGC | 4331 | 4330 |
| RNA target | ggaacgaagcaagc | 4555 | 4554 |

*Table reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

**Table A-2: List of oligonucleotides for Chapter IV.**

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| ON33 | TTTTTTTTTTT | 3588 | 3589 |
| ON34 | TTT$^{AM}$TTT$^{AM}$TTT$^{AM}$TTT | 3472 | 3473 |
| ON35 | **uuu**TTTTTT**uuu** | 3683 | 3685 |
| ON36 | **uuu**TTT$^{AM}$TTT**uuu** | 3645 | 3646 |
| ON37 | **uuu**TT$^{AM}$TTT$^{AM}$T**uuu** | 3606 | 3607 |
| FL-RNA1 | FL-aaaaaaaaaaa | 4425 | 4426 |
| ON38 | T$^{AM}$T$^{AM}$T$^{AM}$T$^{AM}$TCCTGATAGT$^{AM}$T$^{AM}$T$^{AM}$T$^{AM}$T | 5140 | 5141 |
| ON39 | **uuuuu**CCTGATAG**uuuuu** | 5610 | 5610 |
| ON40 | TTTTTCCTGATAGTTTTT | 5451 | 5451 |
| DNA target | AAAAACTATCAGGAAAAA | 5542 | 5542 |
| RNA target | aaaaacuaucaggaaaaa | 5802 | 5803 |
| FL-RNA2 | FL-aaaaacuaucaggaaaaa | 6338 | 6339 |
| FL-RNA3 | FL-aacaacuaucaggaaaaa | 6314 | 6315 |
| FL-RNA4 | FL-aaaaacucucaggaaaaa | 6314 | 6315 |
| FL-RNA5 | FL-aaaaacuaucaggaacaa | 6314 | 6316 |
| FL-ON38 | FL-T$^{AM}$T$^{AM}$T$^{AM}$T$^{AM}$TCCTGATAGT$^{AM}$T$^{AM}$T$^{AM}$T$^{AM}$T | 5677 | 5678 |
| FL-ON39 | FL-**uuuuu**CCTGATAG**uuuuu** | 6147 | 6149 |
| FL-ON40 | FL-TTTTTCCTGATAGTTTTT | 5988 | 5989 |

*Table reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

**Table A-3: List of oligonucleotides for Chapter V.**

| ON | Sequence (5′→3′) | Calcd (Da) | Found (Da) |
|---|---|---|---|
| ON42 | CGACGT$^{TL2}$TTGCAGC | 3952 | 3953 |
| ON45 | CGACGT$^{TL5}$TTGCAGC | 3952 | 3953 |
| ON47 | CGACGTTTGCAGC | 3951 | 3952 |
| DNA target | GCTGCAAACGTCG | 3960 | 3961 |
| RNA target | gcugcaaacgucg | 4140 | 4140 |
| Template 0 (PO) | GCATTCGAGCAACGTAAGATCCTTACCACACA ATCTCACACTCTGGAATTCACACTGACAATAC TGCCGACACACATAACC | 24675 | 24675 |
| Template 1 (TL**1**) | GCGTACGAGAACCGAATGATCT$^{TL1}$TAGCCACA ATCTCACACTCTGGAATTCACACTGACAATAC GTCGCACACAACAATCC | 24423 | 24424 |
| Template 2 (TL**2**) | GCATTCGAGCAACGTAAGATCCT$^{TL2}$TACCACA CAATCTCACACTCTGGAATTCACACTGACAAT ACTGCCGACACACATAACC | 24677 | 24677 |
| Template 3 (TL**3**) | GCATTCGAGCAACGTAAGATCCT$^{TL3}$TACCACA CAATCTCACACTCTGGAATTCACACTGACAAT ACTGCCGACACACATAACC | 24677 | 24677 |
| Template 4 (TL**4**) | GCATTCGAGCAACGTAAGATCCT$^{TL4}$TACCACA CAATCTCACACTCTGGAATTCACACTGACAAT ACTGCCGACACACATAACC | 24663 | 24663 |
| Template 5 (TL**5**) | GCATTCGAGCAACGTAAGATCCT$^{TL5}$TACCACA CAATCTCACACTCTGGAATTCACACTGACAAT ACTGCCGACACACATAACC | 24677 | 24677 |
| Template 6 (TL**6**) | GCTTACGACGAAGAACGGATCT$^{TL6}$TAGCACAC AATCTCACACTCTGGAATTCACACTGACAATA CCTCGGCCAATACACACA | 24453 | 24455 |
| Template 7 (TL**7**) | GCATTCGAGCAACGTAAGATCT$^{TL7}$TAGCACAC AATCTCACACTCTGGAATTCACACTGACAATA CTGCCGACACACATAACC | 24443 | 24443 |
| Template 8 (PO) | TCGTAGCACCTTGCAGCTGACGTTTACATCCA GTCTATTGACTATGCAGTAATAACGGCCACCC TCGCGAACACACCCAAT | 24736 | 24738 |
| Template 9 (TL**2**) | TCGTAGCACCTTGCAGCTGACGT$^{TL2}$TTACATC CAGTCTATTGACTATGCAGTAATAACGGCCAC CCTCGCGAACACACCCAAT | 24738 | 24739 |
| Template 10 (PO) | GCGTACGAGAACCGAATGATCTTAGCACACAA TCTCACACTCTGGAATTCACACTGACAATACG TCGCACACAACAATCC | 24451 | 24453 |
| Template 11 (PO) | GCTTACGACGAAGAACGGATCTTAGCACACAA TCTCACACTCTGGAATTCACACTGACAATACC TCGGCCAATACACACA | 24451 | 24453 |
| Template 12 (PO) | GCATTCGAGCAACGTAAGATCTTAGCACACAA TCTCACACTCTGGAATTCACACTGACAATACT GCCGACACACATAACC | 24426 | 24427 |
| FL-Primer 1 | FL-TGGTTATGTGTGTCGGCAG | 6442 | 6445 |
| FL-Primer 2 | FL-TGGATTGTTGTGTGCGACG | 6442 | 6444 |
| FL-Primer 3 | FL-TTGTGTGTATTGGCCGAGG | 6442 | 6444 |
| FL-Primer 4 | FL-TATTGGGTGTGTTCGCGAG | 6442 | 6444 |
| FP1 | GCATTCGAGCAACGTAAG | 5533 | 5534 |
| FP2 | GCGTACGAGAACCGAATG | 5558 | 5559 |
| FP3 | GCTTACGACGAAGAACGG | 5558 | 5559 |

| | | | |
|---|---|---|---|
| RP1 | GGTTATGTGTGTCGGCAG | 5602 | 5603 |
| RP2 | GGATTGTTGTGTGCGACG | 5602 | 5603 |
| RP3 | TGTGTGTATTGGCCGAGG | 5602 | 5603 |
| LIC-FP | AGAAGGAGATATAACTGCATTCGAGCAACGTAAG | 10557 | 10558 |
| LIC-RP | GGAGATGGGAAGTCAGGTTATGTGTGTCGGCAG | 10370 | 10371 |
| Seq-FP | TAATACGACTCACTATAGGG | 6125 | 6126 |

*Table reproduced with permission from Ref.[307] Copyright 2021 American Chemical Society.*

# 2. Additional Experimental Data

## 2.1. Data of Chapter II

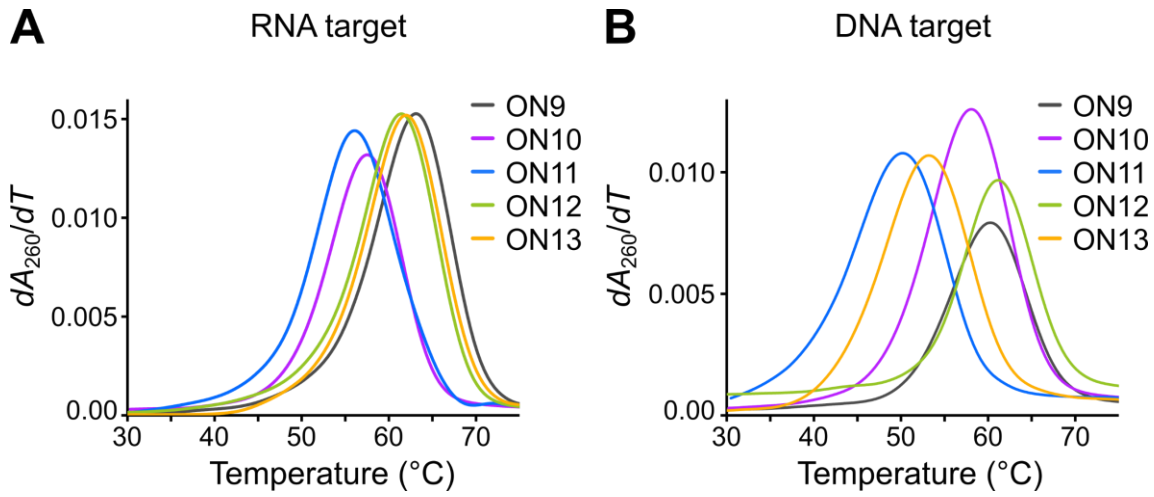



**Figure A-1: First derivatives of UV melting curves.**
Representative plots of first derivatives (d$A_{260}$/d$T$) of $T_m$ curves. $T_m$ curves were obtained from measuring the absorbance at λ = 260 nm in relation to $T$ for ON9–ON13 incubated with their respective targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). (**A**) target = complementary RNA: 5′-r(GGA ACG AAG CAA GC)-3′; (**B**) target = complementary DNA: 5′-d(GGA ACG AAG CAA GC)-3′. *Figure and caption reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

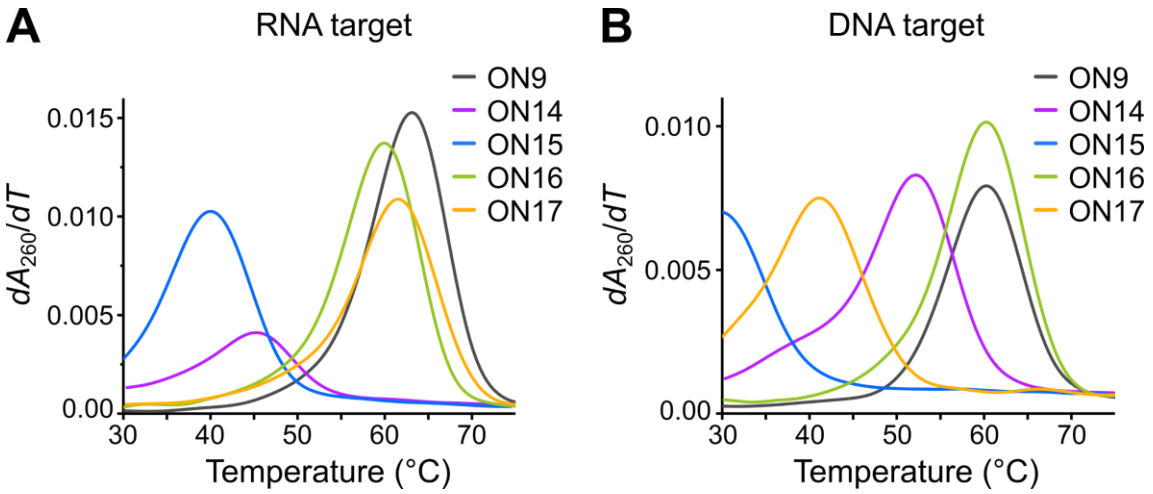



**Figure A-2: First derivatives of UV melting curves.**
Representative plots of first derivatives (d$A_{260}$/d$T$) of $T_m$ curves. $T_m$ curves were obtained from measuring the absorbance at λ = 260 nm in relation to $T$ for ON9 or ON14–ON17 incubated with their respective targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). (**A**) target = complementary RNA: 5′-r(GGA ACG AAG CAA GC)-3′; (**B**) target = complementary DNA: 5′-d(GGA ACG AAG CAA GC)-3′. *Figure and caption reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*

## 2.2. Data of Chapter IV

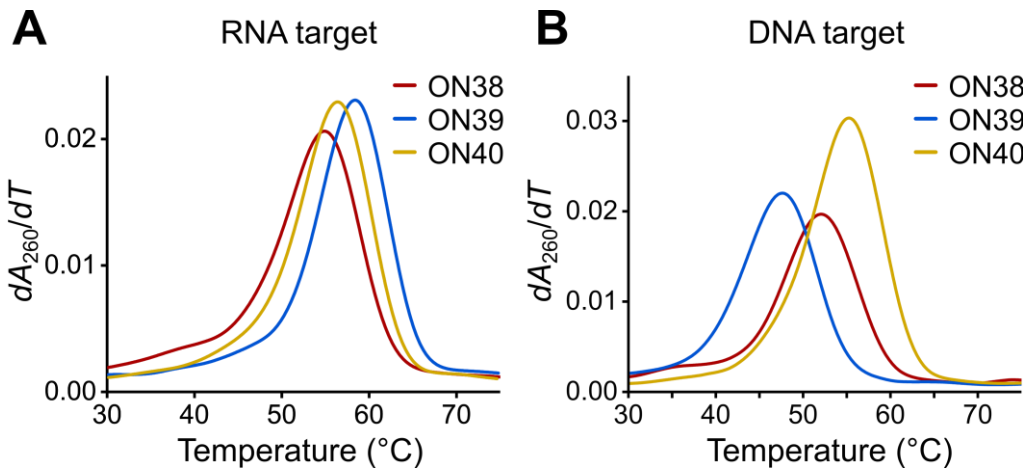



**Figure A-3: First derivatives of UV melting curves.**
Representative plots of first derivatives ($dA_{260}/dT$) of $T_m$ curves. $T_m$ curves were obtained from measuring the absorbance at λ = 260 nm in relation to $T$ for ON38, ON39 and ON40 incubated with their respective targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). (**A**) target = complementary RNA: 5′-r(AAA AAC UAU CAG GAA AAA)-3′; (**B**) target = complementary DNA: 5′-d(AAA AAC TAT CAG GAA AAA)-3′. *Figure and caption reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*

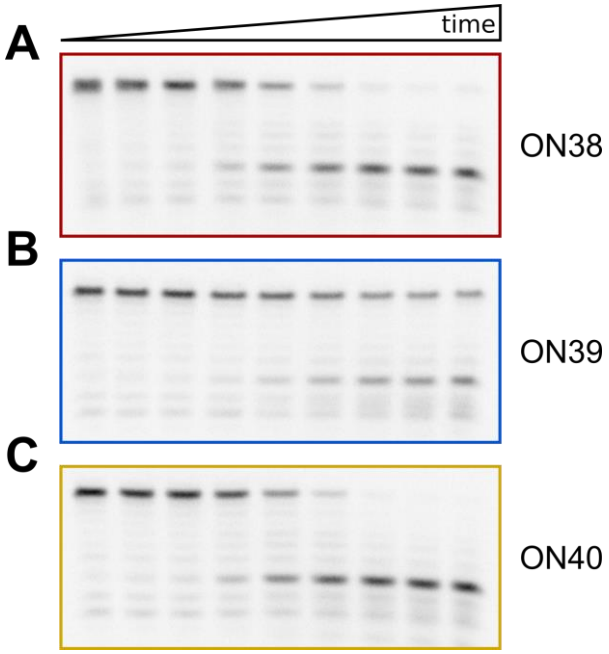



**Figure A-4: RNase H-mediated degradation of a FL-labelled RNA.**
(**A**–**C**) RNase H-mediated degradation of FL-labelled after incubation with ASOs ON38, ON39 and ON40 over 3 h at 20 °C. ASO/FL-RNA2, 1/4 (mole/mole). Aliquots of the reactions were quenched at time points: 0 s, 30 s, 1 min, 5 min, 15 min, 30 min, 1 h, 2 h and 3 h. Degradation fragments of the aliquots were separated by 20% denaturing PAGE and the FL label was visualised using a G:Box ($\lambda_{ex}$ = 460 nm, $\lambda_{em}$ = 516–600 nm).

## 2.3. Data of Chapter V

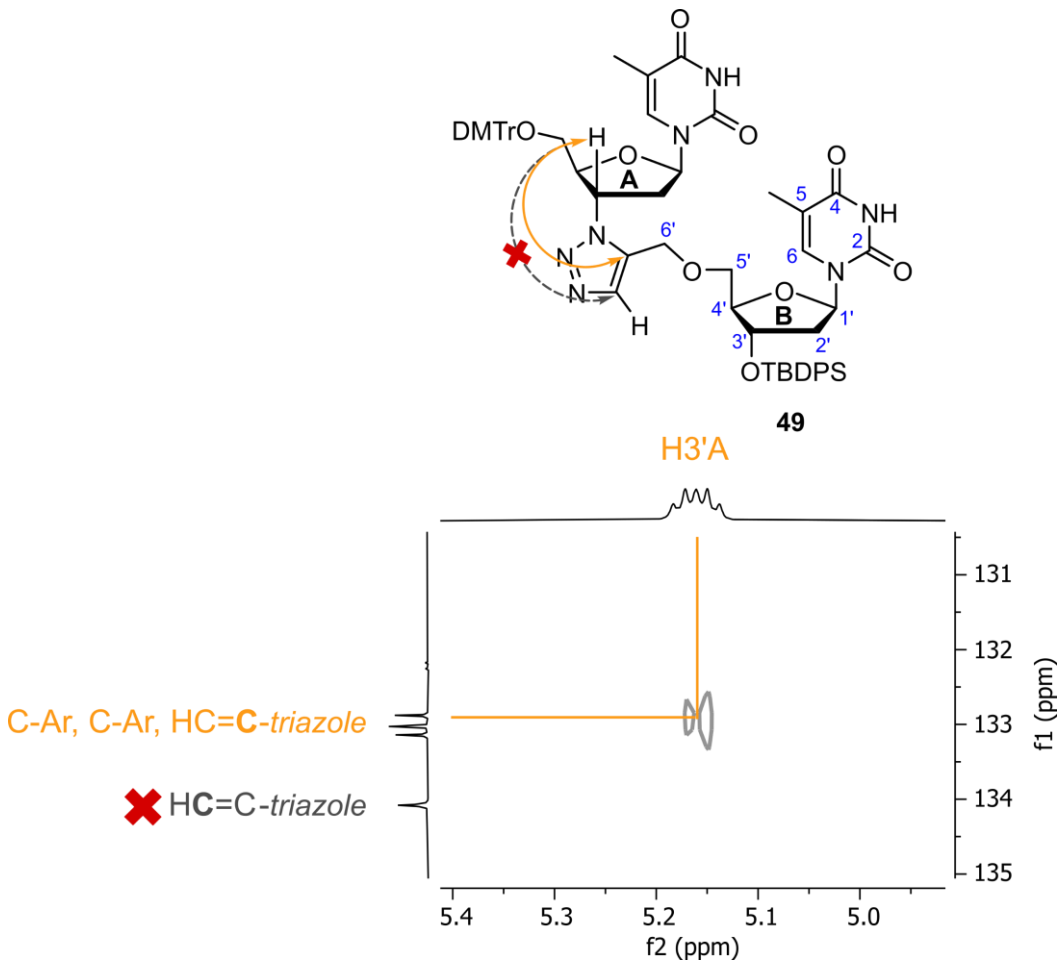



**Figure A-5: Determination of the 1,5-disubstitution of dimer 49 by HMBC spectroscopy.** $^{1}H–^{13}C$ correlation between H3′A and HC=**C**-triazole is indicated with orange lines. No $^{1}H–^{13}C$ correlation between H3′A and H**C**=C-triazole was observed. Note that $^{1}H–^{13}C$ correlation between H3′A and H**C**=C-triazole but not between H3′A and HC=**C**-triazole was previously reported for the 3′-*O*-tert-butyldimethylsilyl (TBS)-protected analogue of the 1,4-disubstituted 1,2,3-triazole T–T dimer.[311] *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

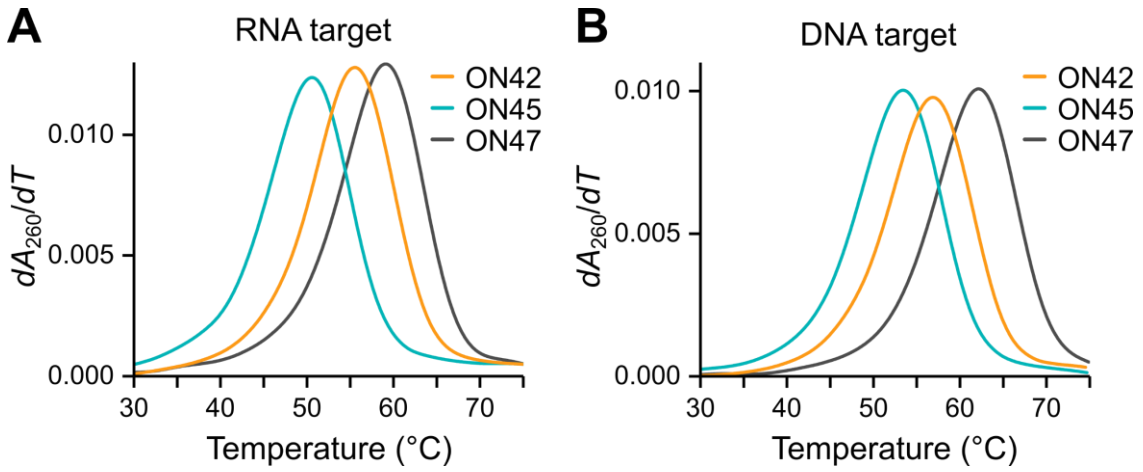



**Figure A-6: First derivatives of UV melting curves.**
Representative plots of first derivatives (d$A_{260}$/d$T$) of $T_m$ curves. $T_m$ curves were obtained from measuring the absorbance at λ = 260 nm in relation to $T$ for ON42, ON45 and ON47 incubated with their respective targets (3 µM, in 10 mM phosphate buffer, 200 mM NaCl, pH 7.0). (**A**) target = complementary RNA: 5′-r(GCU GCA AAC GUC G)-3′; (**B**) target = complementary DNA: 5′-d(GCT GCA AAC GTC G)-3′. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

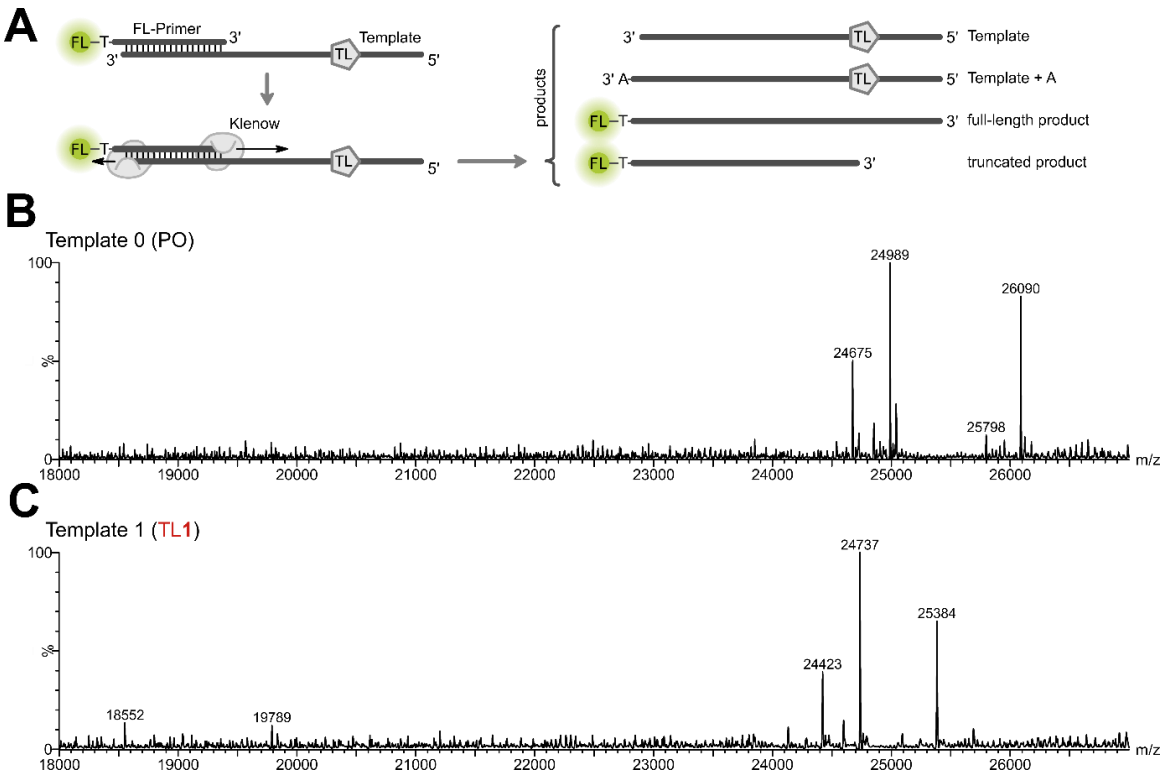



**Figure A-7: MS analysis of primer extension assay with Klenow fragment at 37 °C for 7.5 min.** (**A**) Schematic overview of 3′-extension with Klenow fragment and the resulting products. Note that the single nucleotide overhang (T) of FL-primers at the 5′-end produces an addition of an A-nucleotide to the templates during the assay. (**B**, **C**) MS analyses of ONs after extension of FL-primers with Klenow fragment using (**B**) Template 0 (PO), FL-Primer 1 or (**C**) Template 1 (TL**1**), FL-Primer 2. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-4. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

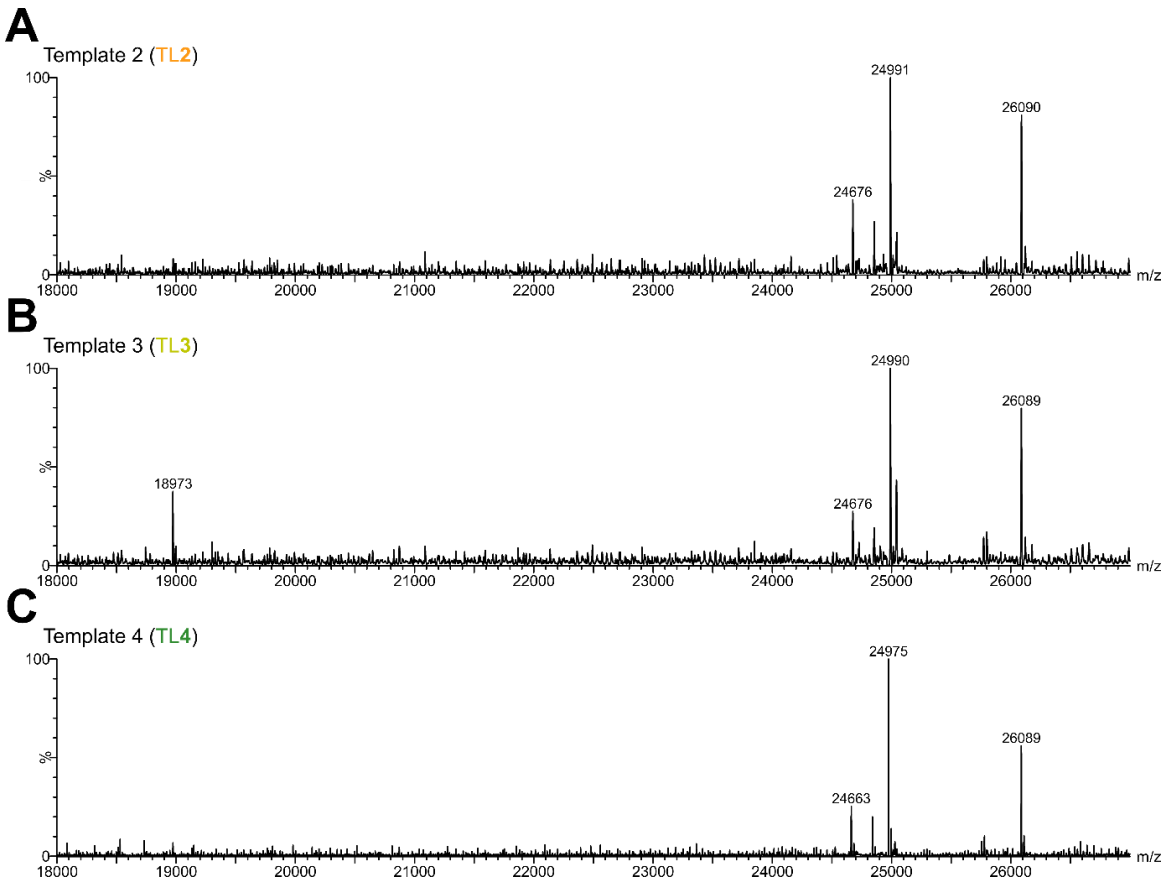



**Figure A-8: MS analysis of primer extension assay with Klenow fragment at 37 °C for 7.5 min.** (**A**–**C**) MS analyses of ONs after extension of FL-primers with Klenow fragment using (**A**) Template 2 (TL**2**), FL-Primer 1; (**B**) Template 3 (TL**3**), FL-Primer 1 or (**C**) Template 4 (TL**4**), FL-Primer 1. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-4. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

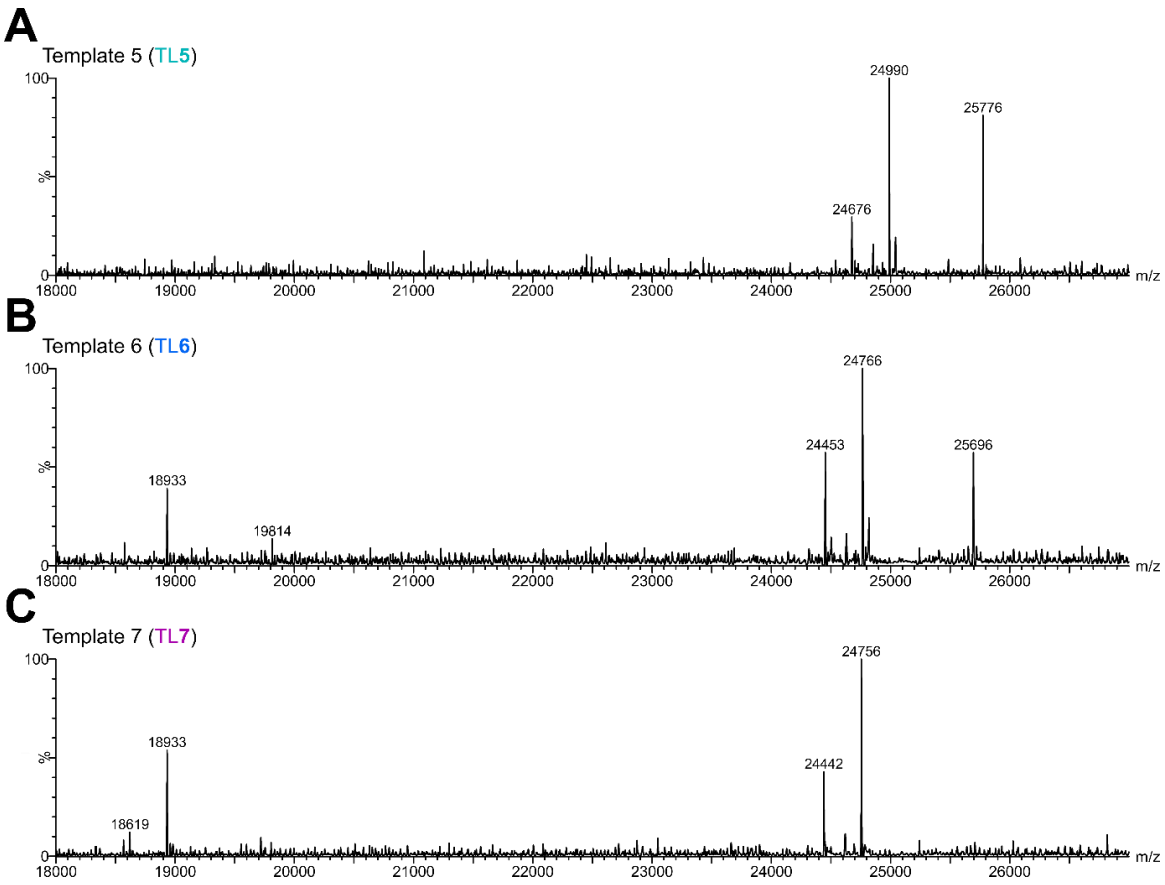



**Figure A-9: MS analysis of primer extension assay with Klenow fragment at 37 °C for 7.5 min.** (**A**–**C**) MS analyses of ONs after extension of FL-primers with Klenow fragment using (**A**) Template 5 (TL**5**), FL-Primer 1; (**B**) Template 6 (TL**6**), FL-Primer 3 or (**C**) Template 7 (TL**7**), FL-Primer 1. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-4. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

**Table A-4: List of calculated and found ON masses after primer extension with Klenow fragment at 37 °C for 7.5 min.**

| Template (modification) | Calcd (Da) | Found (Da) | Comment |
|---|---|---|---|
| Template 0 (PO) | 26089 | 26090 | Full-length product |
| | 24988 | 24989 | Template 0 + A |
| | 24675 | 24675 | Template 0 |
| Template 1 (TL**1**) | 25382 | 25384 | Full-length product with –A deletion |
| | 24736 | 24737 | Template 1 + A |
| | 24423 | 24423 | Template 1 |
| Template 2 (TL**2**) | 26089 | 26090 | Full-length product |
| | 24990 | 24991 | Template 2 + A |
| | 24677 | 24676 | Template 2 |
| Template 3 (TL**3**) | 26089 | 26089 | Full-length product |
| | 24990 | 24990 | Template 3 + A |
| | 24677 | 24676 | Template 3 |
| | 18972 | 18973 | Truncated product; stopped directly before TL**3** |
| Template 4 (TL**4**) | 26089 | 26089 | Full-length product |
| | 24976 | 24975 | Template 4 + A |
| | 24663 | 24663 | Template 4 |
| Template 5 (TL**5**) | 25776 | 25776 | Full-length product with –A deletion |
| | 24990 | 24990 | Template 5 + A |
| | 24677 | 24676 | Template 5 |
| Template 6 (TL**6**) | 25695 | 25696 | Full-length product |
| | 24766 | 24766 | Template 6 + A |
| | 24453 | 24453 | Template 6 |
| | 18932 | 18933 | Truncated product; stopped directly before TL**6** |
| Template 7 (TL**7**) | 24756 | 24756 | Template 7 + A |
| | 24443 | 24442 | Template 7 |
| | 18932 | 18933 | Truncated product; stopped directly before TL**7** |

*Table reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

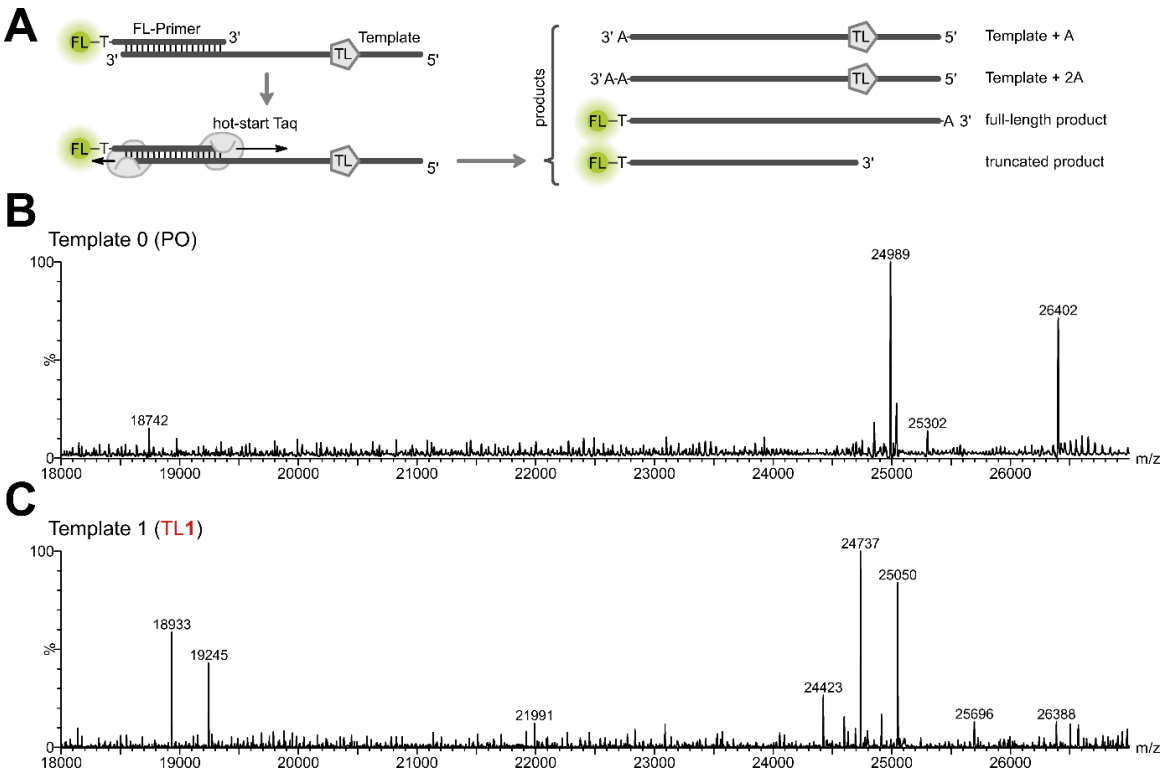



**Figure A-10: MS analysis of primer extension assay with hot-start Taq polymerase at 60 °C for 160 min.**

(**A**) Schematic overview of 3′-extension with hot-start Taq polymerase and the resulting products. Note that the single nucleotide overhang (T) of FL-primers at the 5′-end produces an addition of an A-nucleotide to the templates during the assay. The formation of a 3′-A overhang by hot-start Taq is included in the products. (**B**, **C**) MS analyses of ONs after extension of FL-primers with hot-start Taq polymerase using (**B**) Template 0 (PO), FL-Primer 1 or (**C**) Template 1 (TL**1**), FL-Primer 2. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-5. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

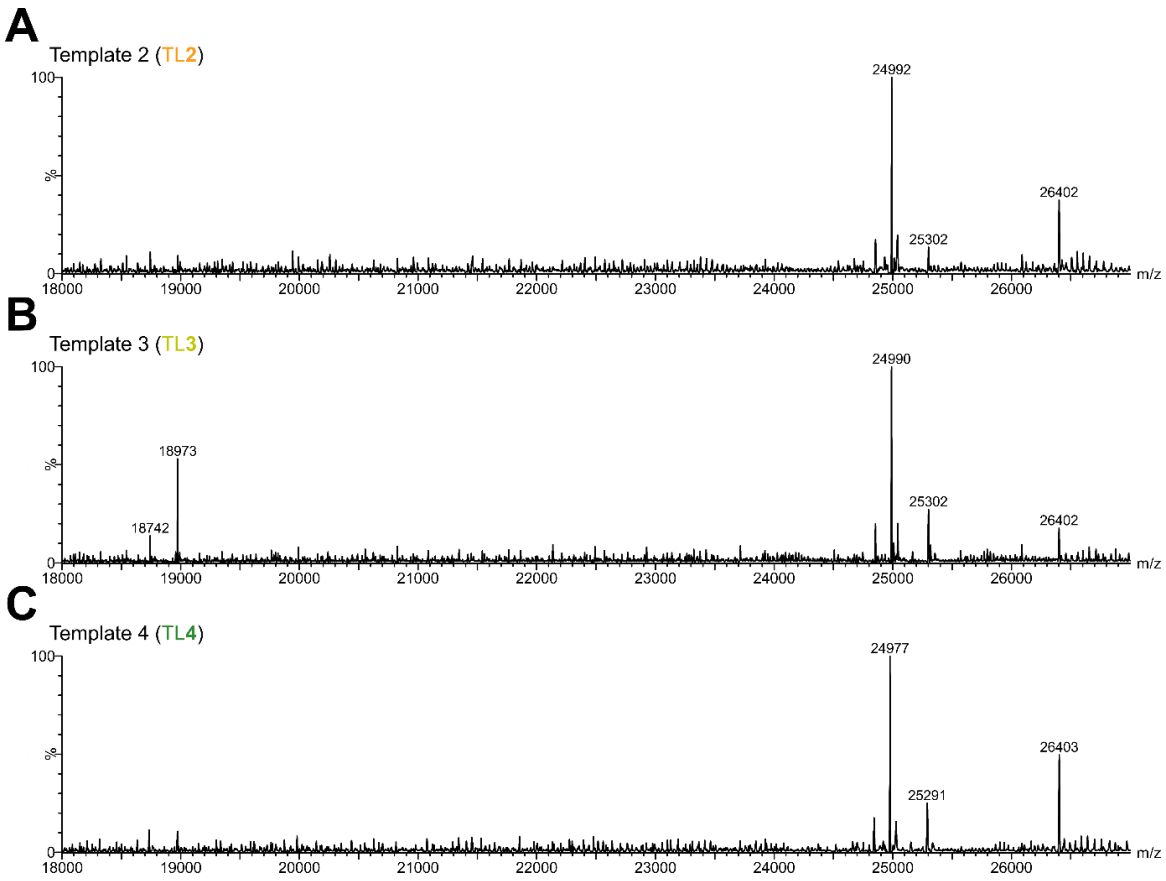



**Figure A-11: MS analysis of primer extension assay with hot-start Taq polymerase at 60 °C for 160 min.**

(**A**–**C**) MS analyses of ONs after extension of FL-primers with hot-start Taq polymerase using (**A**) Template 2 (TL**2**), FL-Primer 1; (**B**) Template 3 (TL**3**), FL-Primer 1 or (**C**) Template 4 (TL**4**), FL-Primer 1. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-5. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

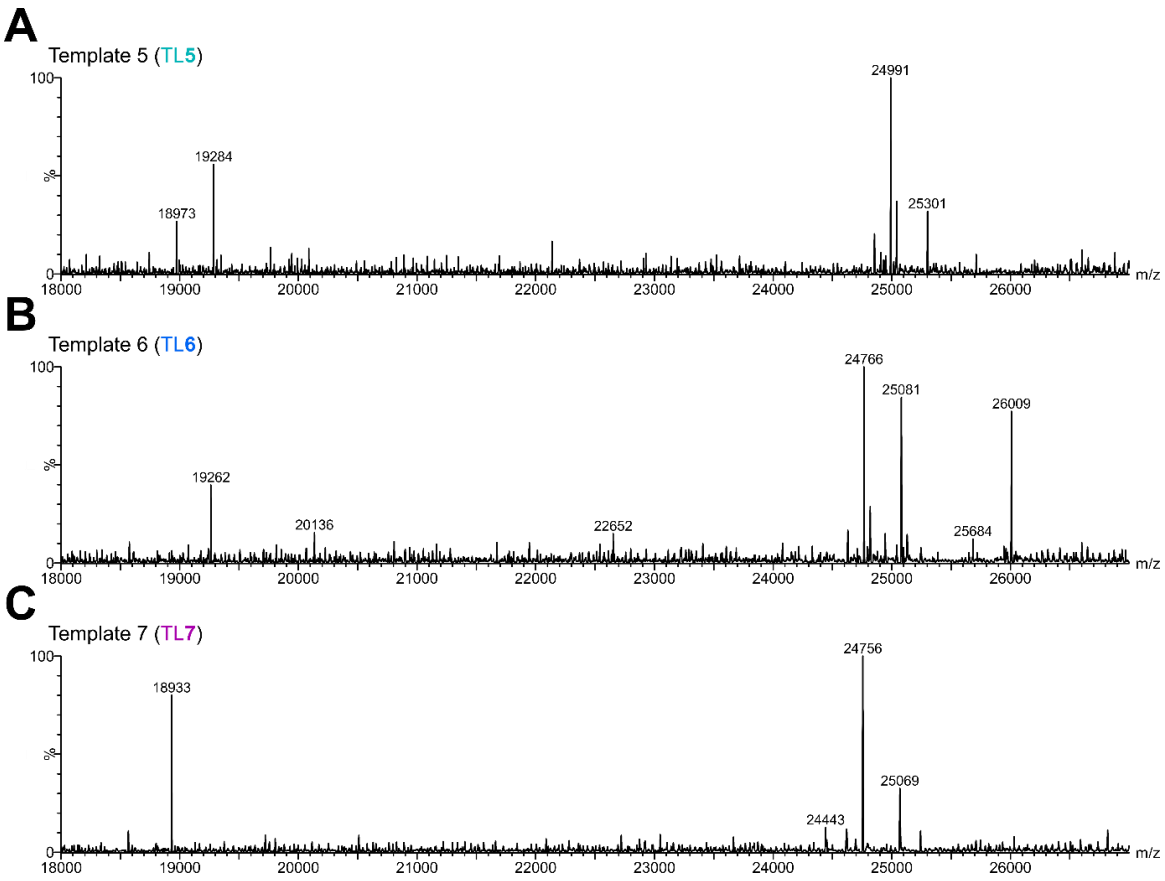



**Figure A-12: MS analysis of primer extension assay with hot-start Taq polymerase at 60 °C for 160 min.**

(**A**–**C**) MS analyses of ONs after extension of FL-primers with hot-start Taq polymerase using (**A**) Template 5 (TL**5**), FL-Primer 1; (**B**) Template 6 (TL**6**), FL-Primer 3 or (**C**) Template 7 (TL**7**), FL-Primer 1. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-5. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

**Table A-5: List of calculated and found ON masses after primer extension with hot-start Taq polymerase at 60 °C for 160 min.**

| Template (modification) | Calcd (Da) | Found (Da) | Comment |
|---|---|---|---|
| Template 0 (PO) | 26402 | 26402 | Full-length product (including + 3′-A overhang adduct) |
| | 25302 | 25302 | Template 0 + 2 × A |
| | 24988 | 24989 | Template 0 + A |
| Template 1 (TL**1**) | 25695 | 25696 | Full-length product (without 3′-A overhang adduct or with –A deletion) |
| | 25049 | 25050 | Template 1 + 2 × A |
| | 24736 | 24737 | Template 1 + A |
| | 24423 | 24423 | Template 1 |
| | 19245 | 19245 | Truncated product; stopped directly after TL**1** |
| | 18932 | 18933 | Truncated product; stopped directly before TL**1** |
| Template 2 (TL**2**) | 26402 | 26402 | Full-length product (including + 3′-A overhang adduct) |
| | 25304 | 25302 | Template 2 + 2 × A |
| | 24990 | 24992 | Template 2 + A |
| Template 3 (TL**3**) | 26402 | 26402 | Full-length product (including + 3′-A overhang adduct) |
| | 25304 | 25302 | Template 3 + 2 × A |
| | 24990 | 24990 | Template 3 + A |
| | 18972 | 18973 | Truncated product; stopped directly before TL**3** |
| Template 4 (TL**4**) | 26402 | 26403 | Full-length product (including + 3′-A overhang adduct) |
| | 25290 | 25291 | Template 4 + 2 × A |
| | 24976 | 24977 | Template 4 + A |
| Template 5 (TL**5**) | 25304 | 25301 | Template 5 + 2 × A |
| | 24990 | 24991 | Template 5 + A |
| | 19285 | 19284 | Truncated product; stopped directly after TL**5** |
| | 18972 | 18973 | Truncated product; stopped directly before TL**5** |
| Template 6 (TL**6**) | 26009 | 26009 | Full-length product (including + 3′-A overhang adduct) |
| | 25080 | 25081 | Template 6 + 2 × A |
| | 24766 | 24766 | Template 6 + A |
| | 19261 | 19262 | Truncated product; stopped directly after TL**6** but with A→G mutation |
| Template 7 (TL**7**) | 25070 | 25069 | Template 7 + 2 × A |
| | 24756 | 24756 | Template 7 + A |
| | 24443 | 24443 | Template 7 |
| | 18932 | 18933 | Truncated product; stopped directly before TL**7** |

*Table reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

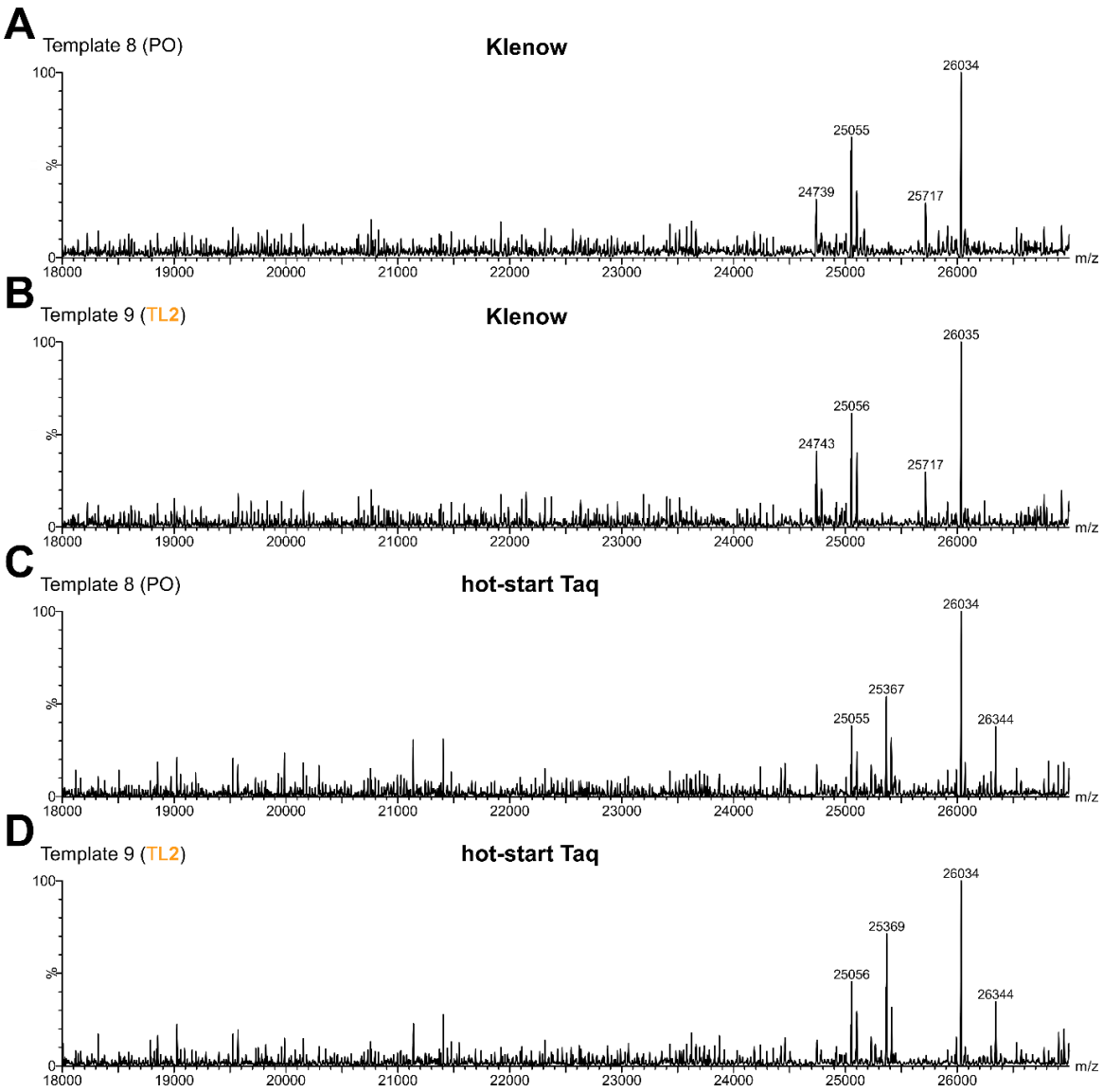



**Figure A-13: MS analyses of ONs after extension of FL-Primer 4 using alternative templates.** Alternative templates were: Template 8 (PO) or Template 9 (TL**2**). (**A**, **B**) Klenow fragment at 37 °C for 7.5 min. (**C**, **D**) hot-start Taq polymerase at 60 °C for 160 min. Sequences can be found in Table A-3. Calculated and found masses and their assignments are summarised in Table A-6. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

**Table A-6: List of calculated and found ON masses after primer extension with Klenow fragment at 37 °C for 7.5 min or hot-start Taq polymerase at 60 °C for 160 min.**

| | Template (modification) | Calcd (Da) | Found (Da) | Comment |
|---|---|---|---|---|
| Klenow | Template 8 (PO) | 26031 | 26034 | Full-length product |
| | | 25717 | 25717 | Full-length product – A[a] |
| | | 25049 | 25055 | Template 8 + A |
| | | 24736 | 24739 | Template 8 |
| | Template 9 (TL**2**) | 26031 | 26035 | Full-length product |
| | | 25717 | 25717 | Full-length product – A[a] |
| | | 25051 | 25056 | Template 9 + A |
| | | 24738 | 24743 | Template 9 |
| hot-start Taq | Template 8 (PO) | 26344 | 26344 | Full-length product (including + 3′-A overhang adduct) |
| | | 26031 | 26034 | Full-length product (without the 3′-A overhang adduct)[b] |
| | | 25363 | 25367 | Template 8 + 2 × A |
| | | 25049 | 25055 | Template 8 + A |
| | Template 9 (TL**2**) | 26344 | 26344 | Full-length product (including + 3′-A overhang adduct) |
| | | 26031 | 26034 | Full-length product (without the 3′-A overhang adduct)[b] |
| | | 25365 | 25369 | Template 9 + 2 × A |
| | | 25051 | 25056 | Template 9 + A |

[a] distinct peaks with a missing A-nucleotide were detected for the unmodified Template 8 and isosequential TL**2**-modified Template 9, both having a 5′-T. This suggests that the missing A-nucleotide in the replicon results in both cases from incomplete replication without addition of the 3′-terminal A-nucleotide (complementary to the 5′-T of the templates) and is thus determined by the template sequence and not by the TL**2** modification.

[b] major peaks without the Taq-characteristic 3′-A overhang were detected for the unmodified Template 8 and isosequential TL**2**-modified Template 9. This suggests that formation of a 3′-A overhang is suppressed in both cases and not specific to the TL**2** modification.

*Table reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

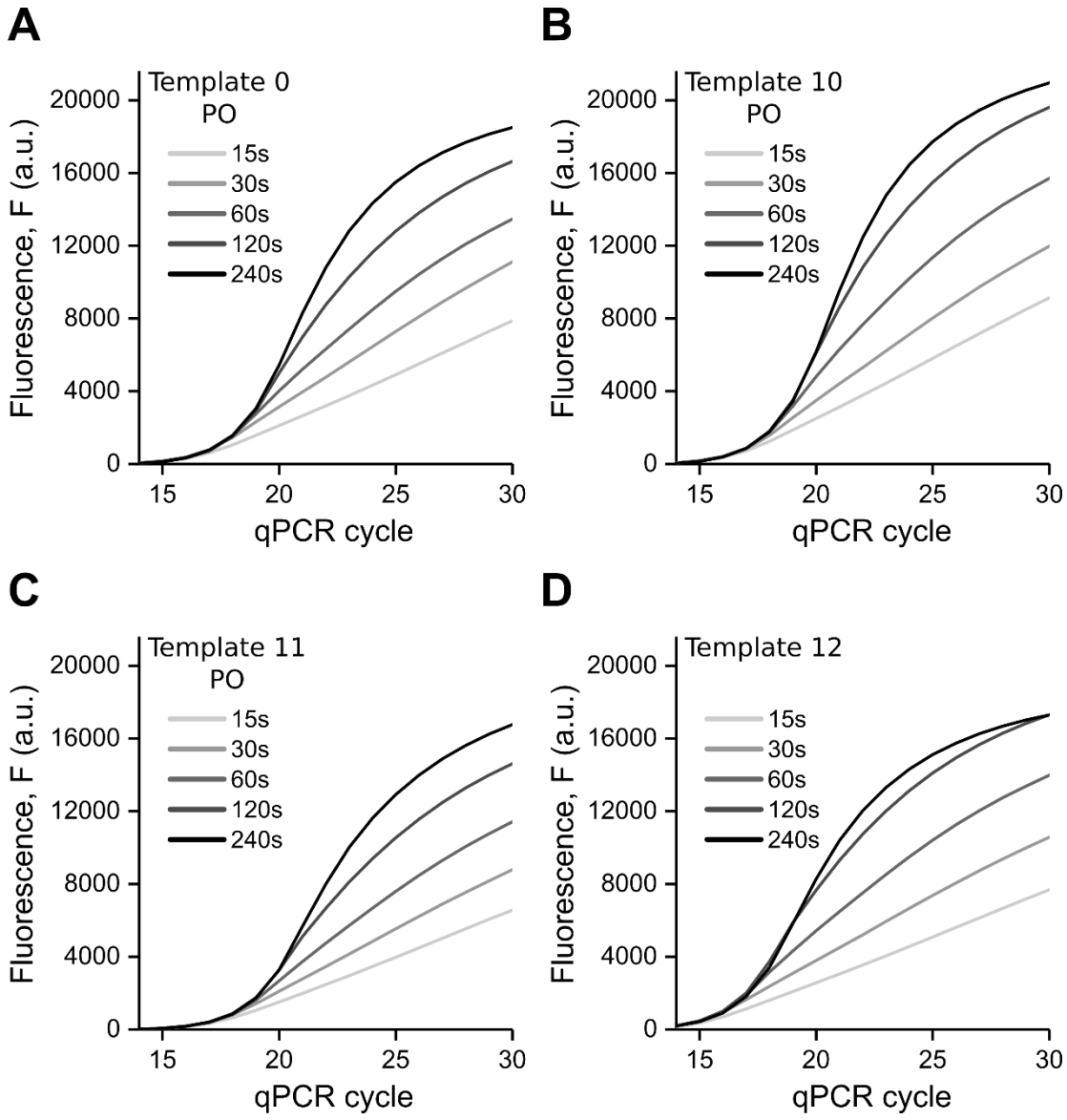



**Figure A-14: Representative qPCR amplifications of Template 0 and Templates 10–12 by hot-start Taq polymerase with different extension times.**
Fluorescence emission was monitored at the end of each amplification cycle. The extension time during the qPCR cycle was varied from 15–240 s. Templates, forward and reverse primers (FP, RP) were the following: (**A**) Template 0 (PO), FP1, RP1; (**B**) Template 10 (PO), FP2, RP2; (**C**) Template 11 (PO), FP3, RP3 and (**D**) Template 12 (PO), FP1, RP1. Sequences can be found in Table A-3. Enlarged sections for each panel showing the threshold used for the $C_t$ calculations are shown in Figure A-17. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

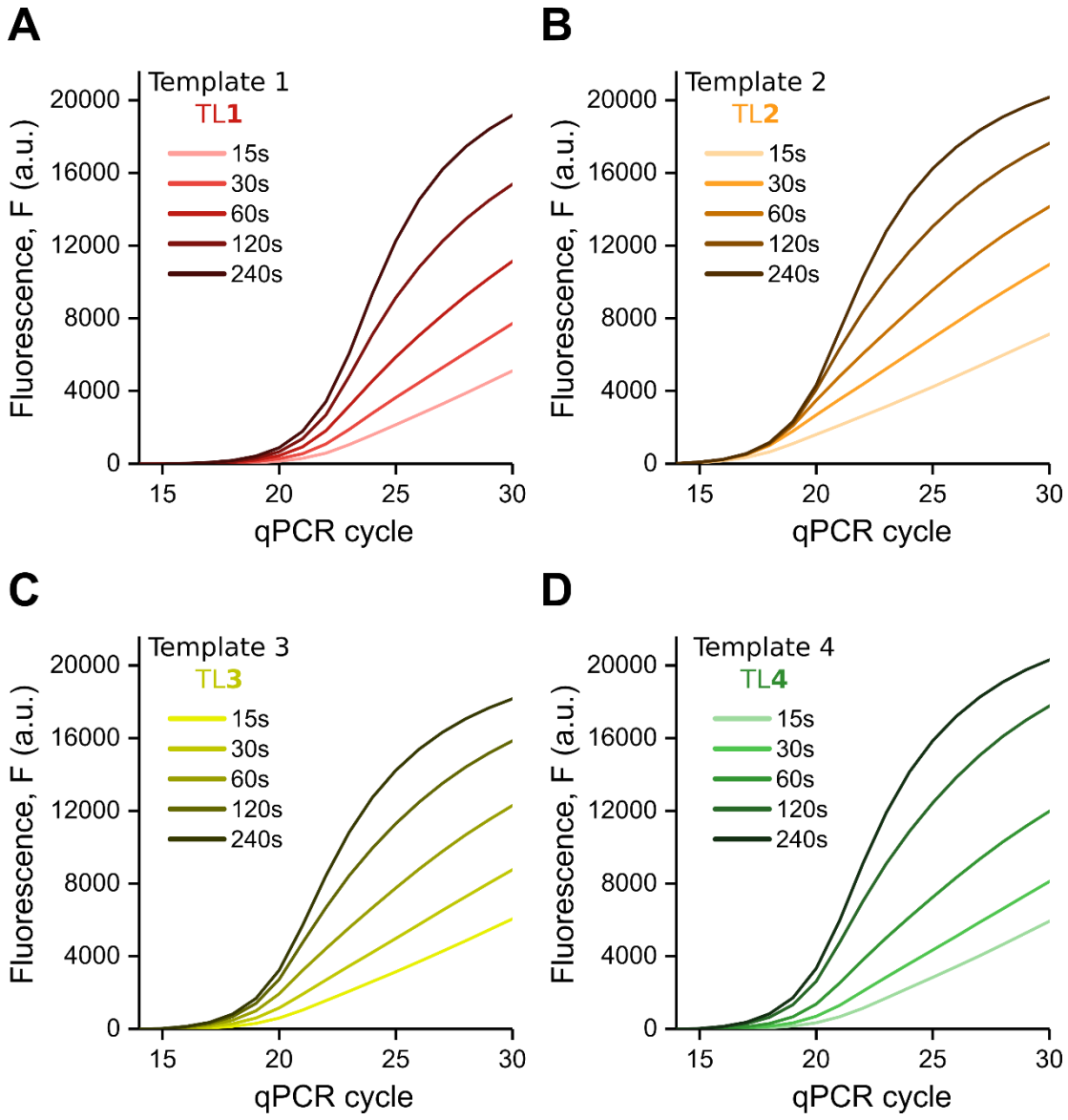



**Figure A-15: Representative qPCR amplifications of Templates 1–4 by hot-start Taq polymerase with different extension times.**
Fluorescence emission was monitored at the end of each amplification cycle. The extension time during the qPCR cycle was varied from 15–240 s. Templates, FP and RP were the following: (**A**) Template 1 (TL**1**), FP2, RP2; (**B**) Template 2 (TL**2**), FP1, RP1; (**C**) Template 3 (TL**3**), FP1, RP1 and (**D**) Template 4 (TL**4**), FP1, RP1. Sequences can be found in Table A-3. Enlarged sections for each panel showing the threshold used for the $C_t$ calculations are shown in Figure A-18. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

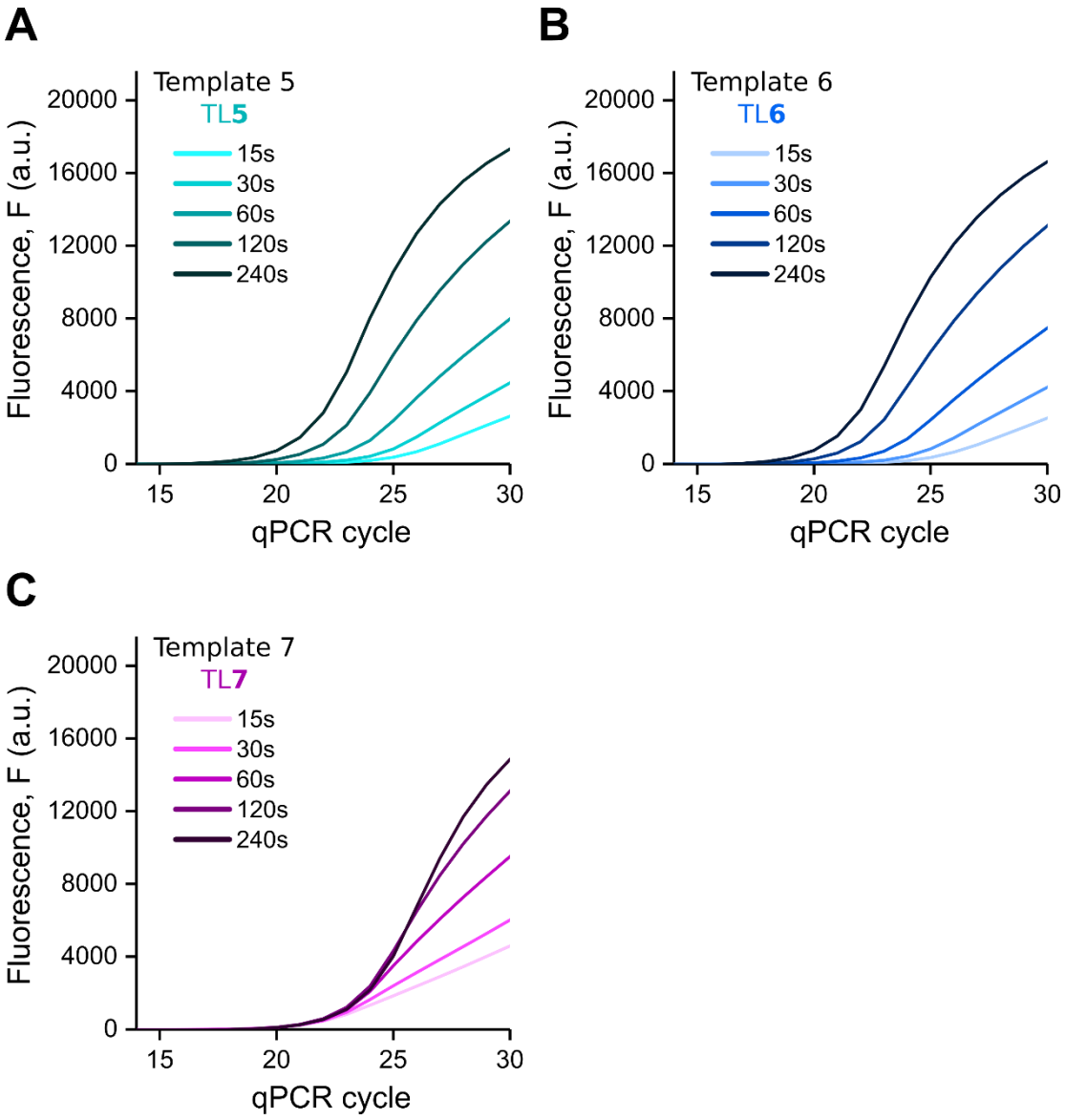



**Figure A-16: Representative qPCR amplifications of Templates 5–7 by hot-start Taq polymerase with different extension times.**
Fluorescence emission was monitored at the end of each amplification cycle. The extension time during the qPCR cycle was varied from 15–240 s. Templates, FP and RP were the following: (**A**) Template 5 (TL**5**), FP1, RP1; (**B**) Template 6 (TL**6**), FP3, RP3 and (**C**) Template 7 (TL**7**), FP1, RP1. Sequences can be found in Table A-3. Enlarged sections for each panel showing the threshold used for the $C_t$ calculations are shown in Figure A-19. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

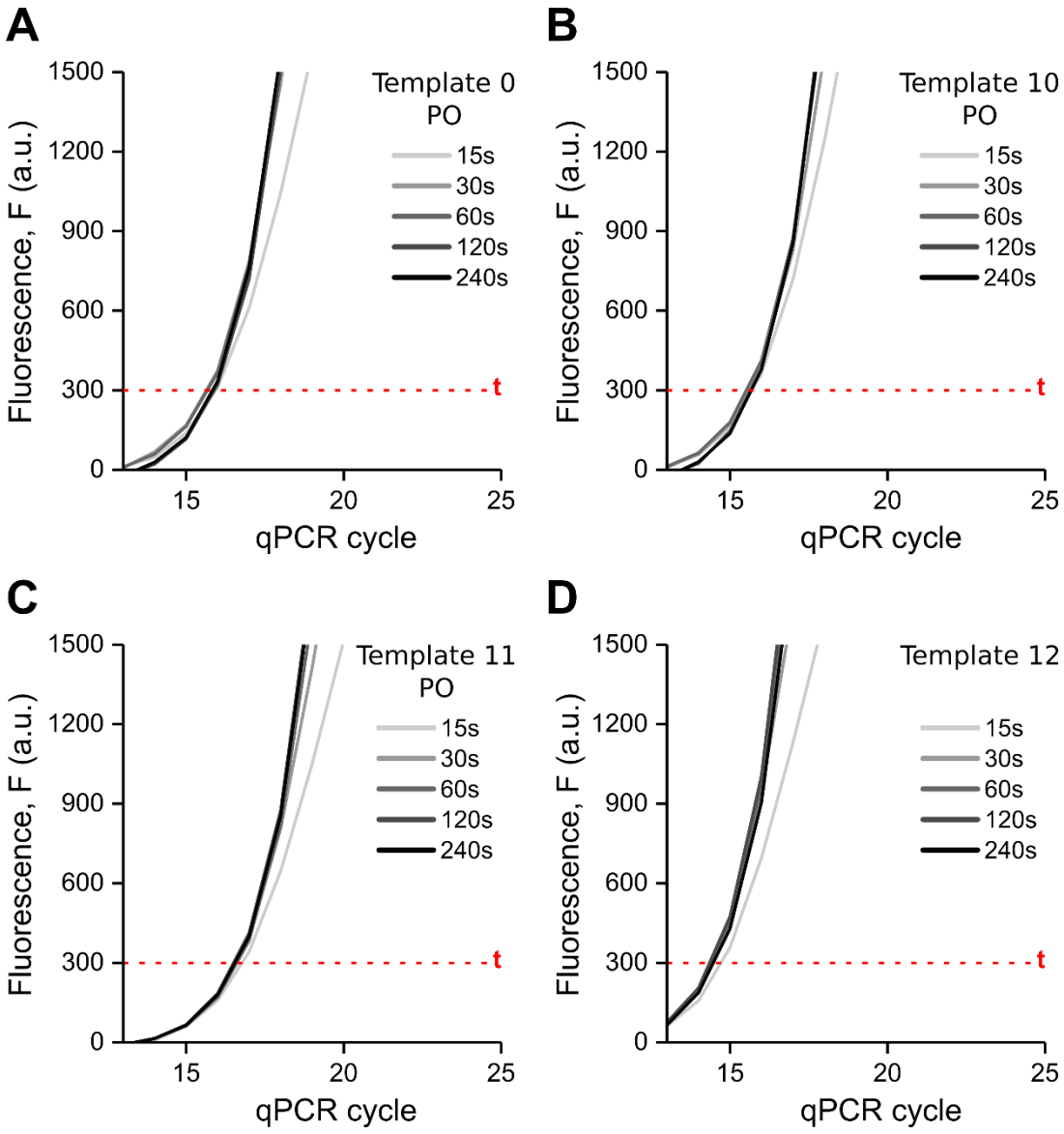



**Figure A-17: Enlarged, representative sections of qPCR amplifications of Template 0 and Templates 10–12 by hot-start Taq polymerase with different extension times.**
The threshold (t) at F = 300 (a.u.) is indicated as a red dotted line. (**A**–**D**) correspond to qPCR samples from (**A**–**D**) in Figure A-14. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

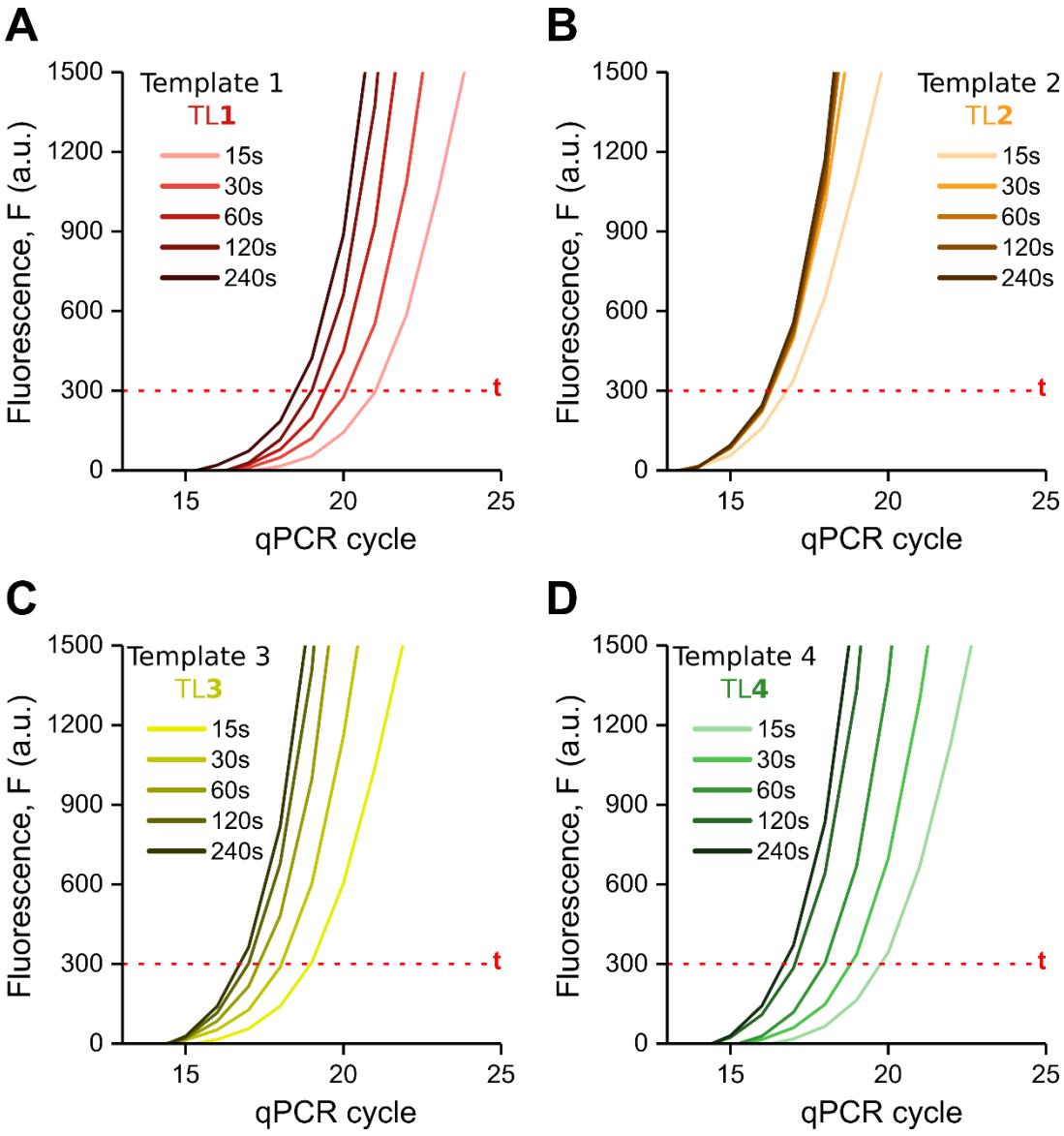



**Figure A-18: Enlarged, representative sections of qPCR amplifications of Templates 1–4 by hot-start Taq polymerase with different extension times.**
The threshold (t) at F = 300 (a.u.) is indicated as a red dotted line. (**A**–**D**) correspond to qPCR samples from (**A**–**D**) in Figure A-15. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

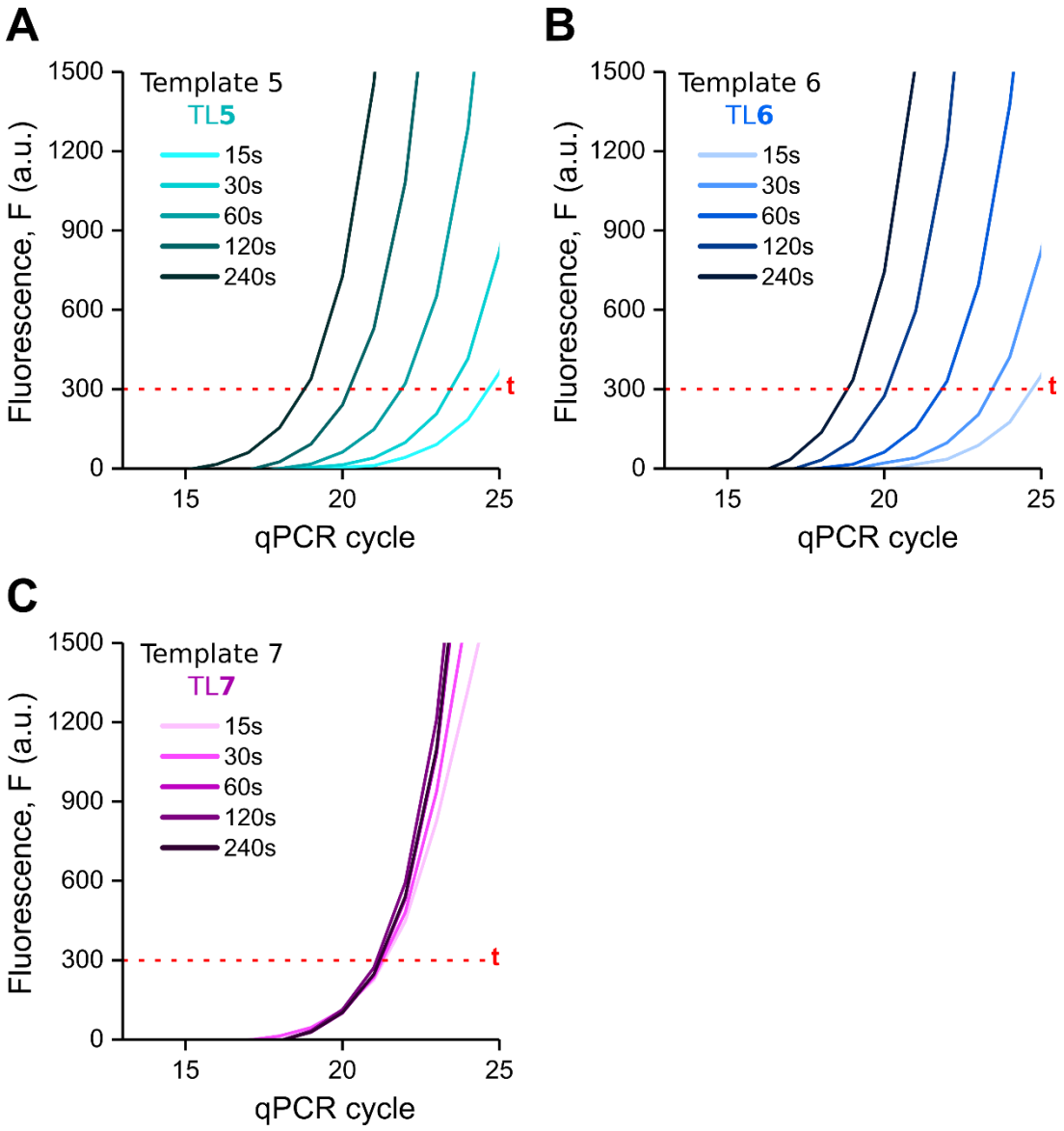



**Figure A-19: Enlarged, representative sections of qPCR amplifications of Templates 5–7 by hot-start Taq polymerase with different extension times.**
The threshold (t) at F = 300 (a.u.) is indicated as a red dotted line. (**A**–**C**) correspond to qPCR samples from (**A**–**C**) in Figure A-16. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

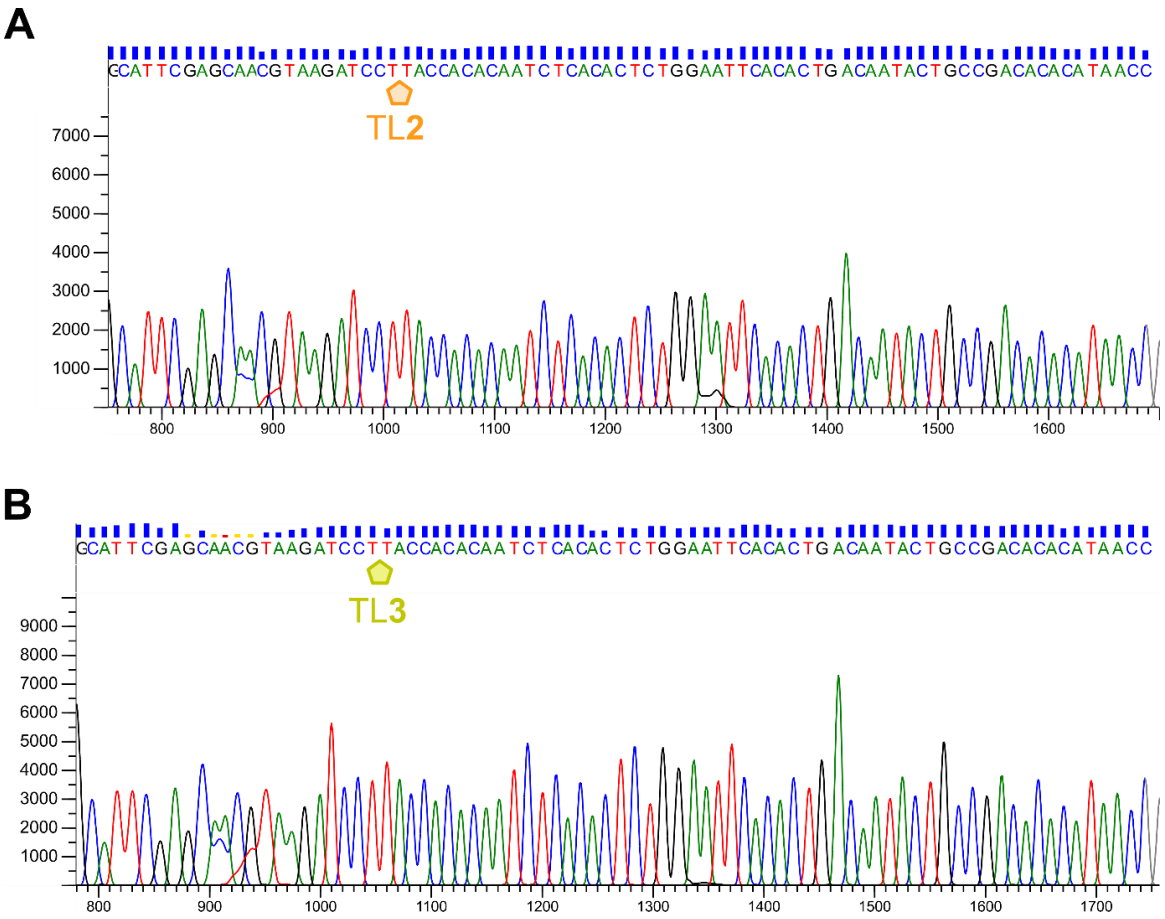



**Figure A-20: Representative Sanger sequencing results.**
Amplicons from PCR products of the TL-containing templates using tailed primers (LIC-FP and LIC-RP; Table A-3) and hot-start Taq polymerase were cloned into a vector and the insertion sites were sequenced by Sanger sequencing. The original position of the TL in the corresponding template is indicated as a pentagon. Templates used: (**A**) Template 2 (TL**2**), n = 4. (**B**) Template 3 (TL**3**), n = 4. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

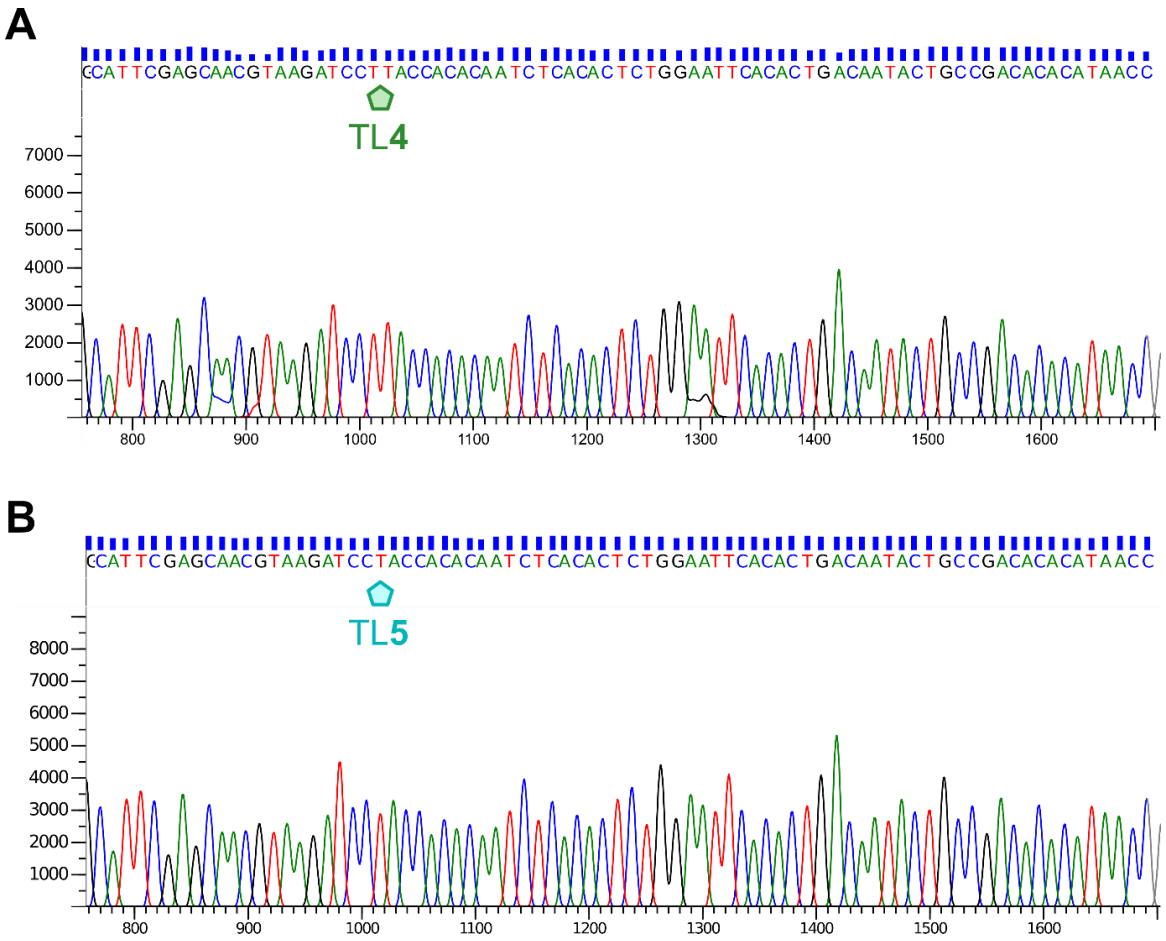
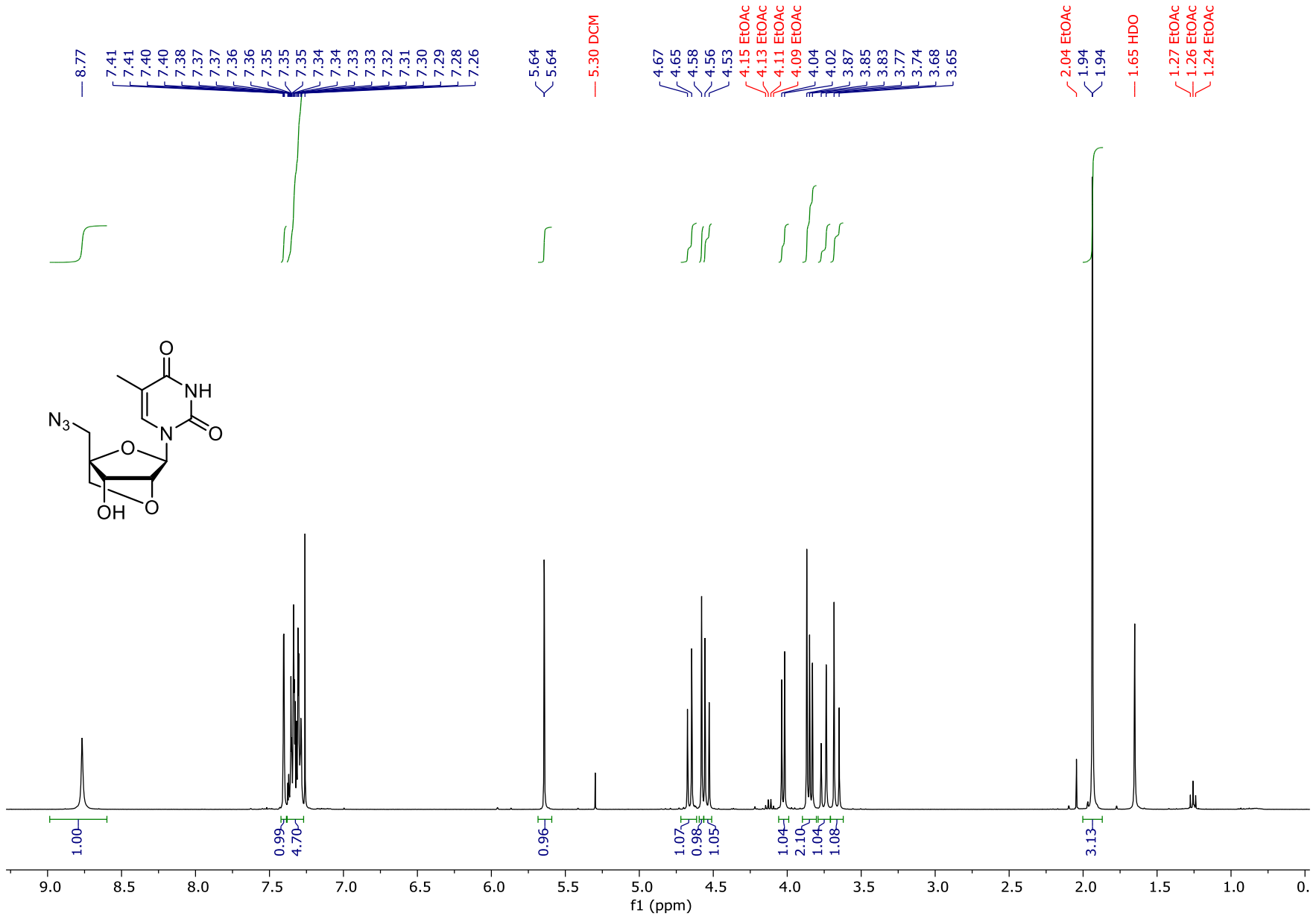


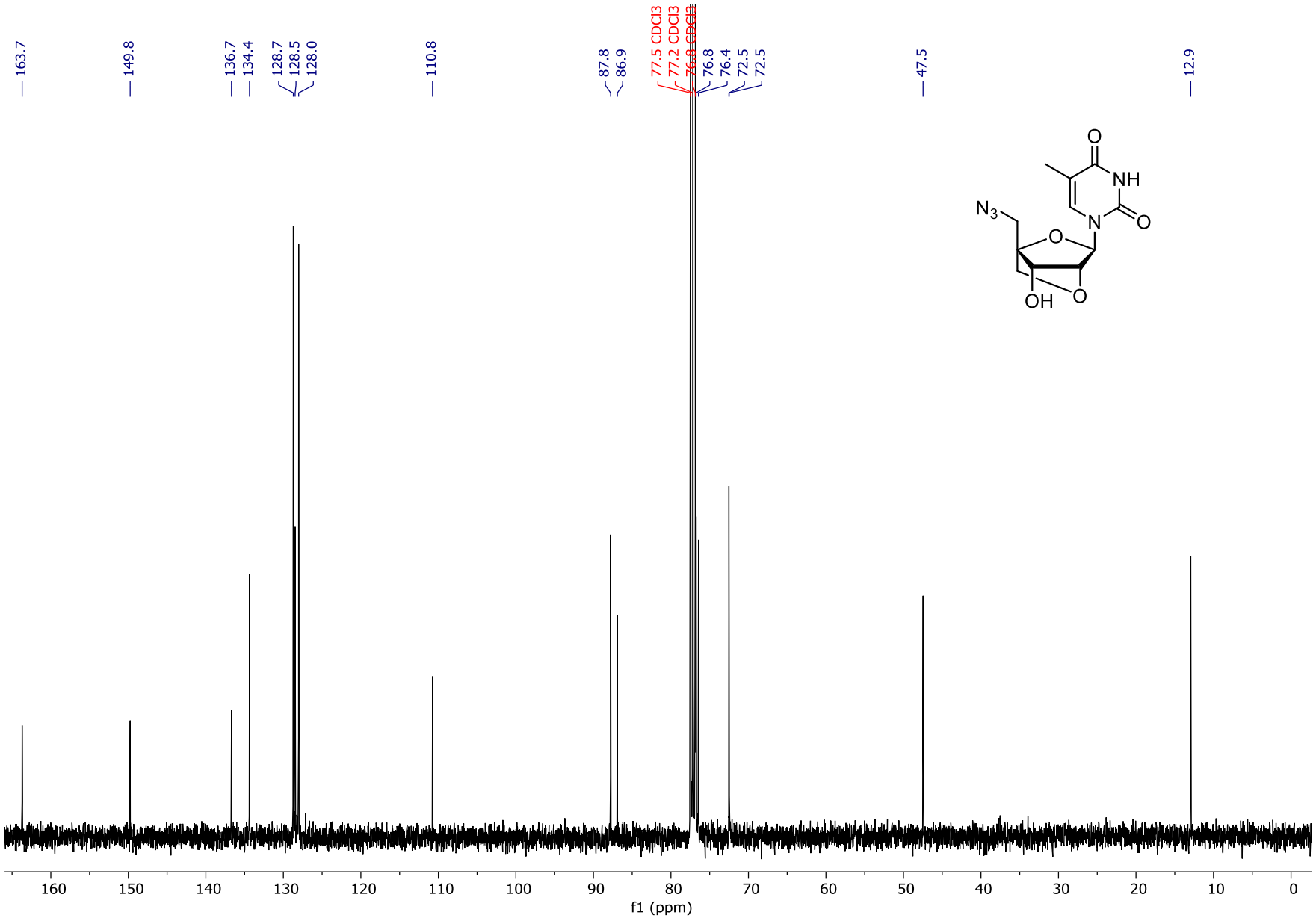

**Figure A-21: Representative Sanger sequencing results.**
Amplicons from PCR products of the TL-containing templates using tailed primers (LIC-FP and LIC-RP; Table A-3) and hot-start Taq polymerase were cloned into a vector and the insertion sites were sequenced by Sanger sequencing. The original position of the TL in the corresponding template is indicated as a pentagon. Templates used: (**A**) Template 4 (TL**4**), n = 4. (**B**) Template 5 (TL**5**), n = 5. Note that TL**5** causes a deletion mutation around the triazole for which the original $T^{TL5}T$ motif has been read as a single T. *Figure and caption reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

# 3. NMR Spectra

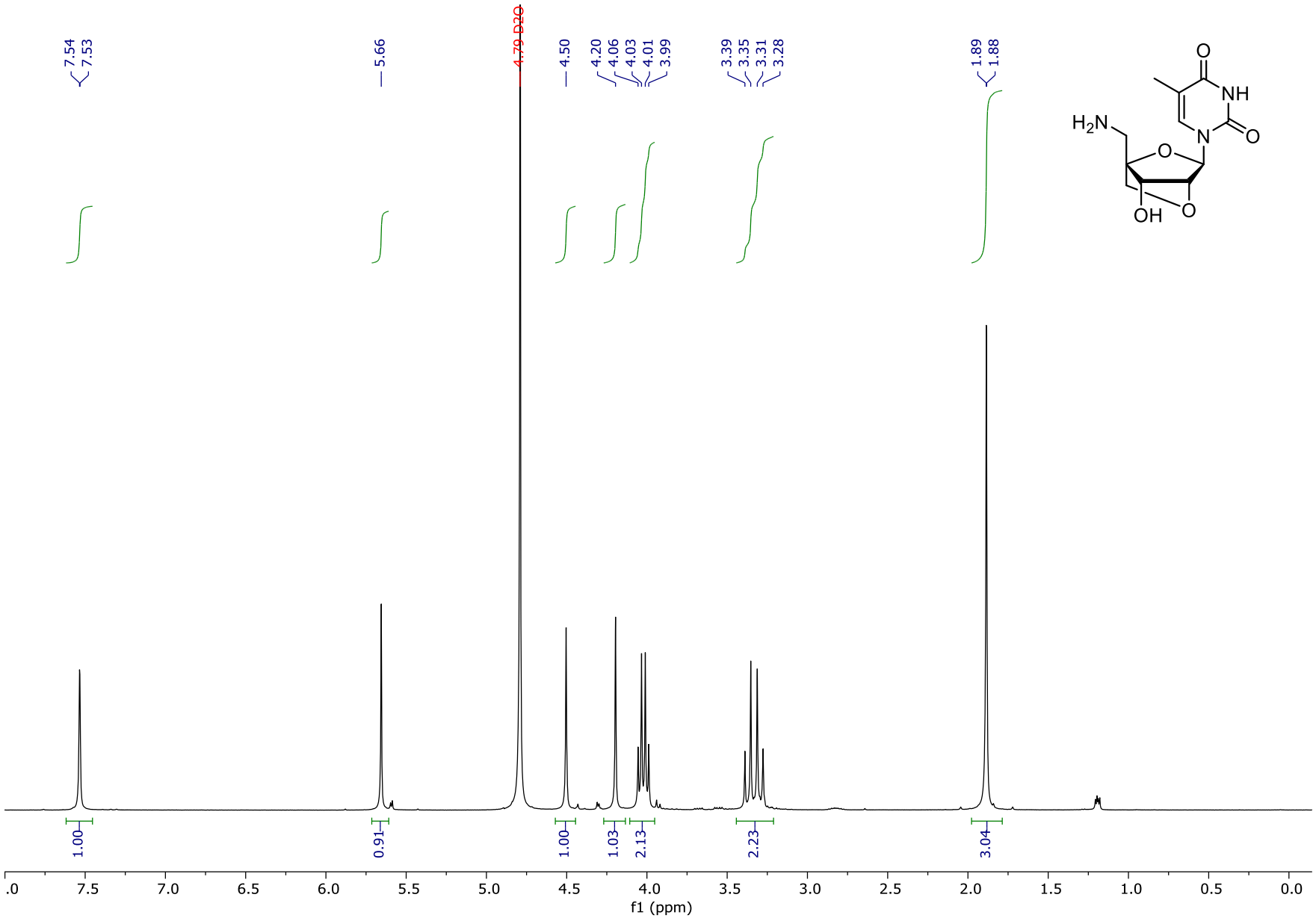
NMR spectra are shown for novel compounds. References and copyrights are indicated for compounds that have been published elsewhere.



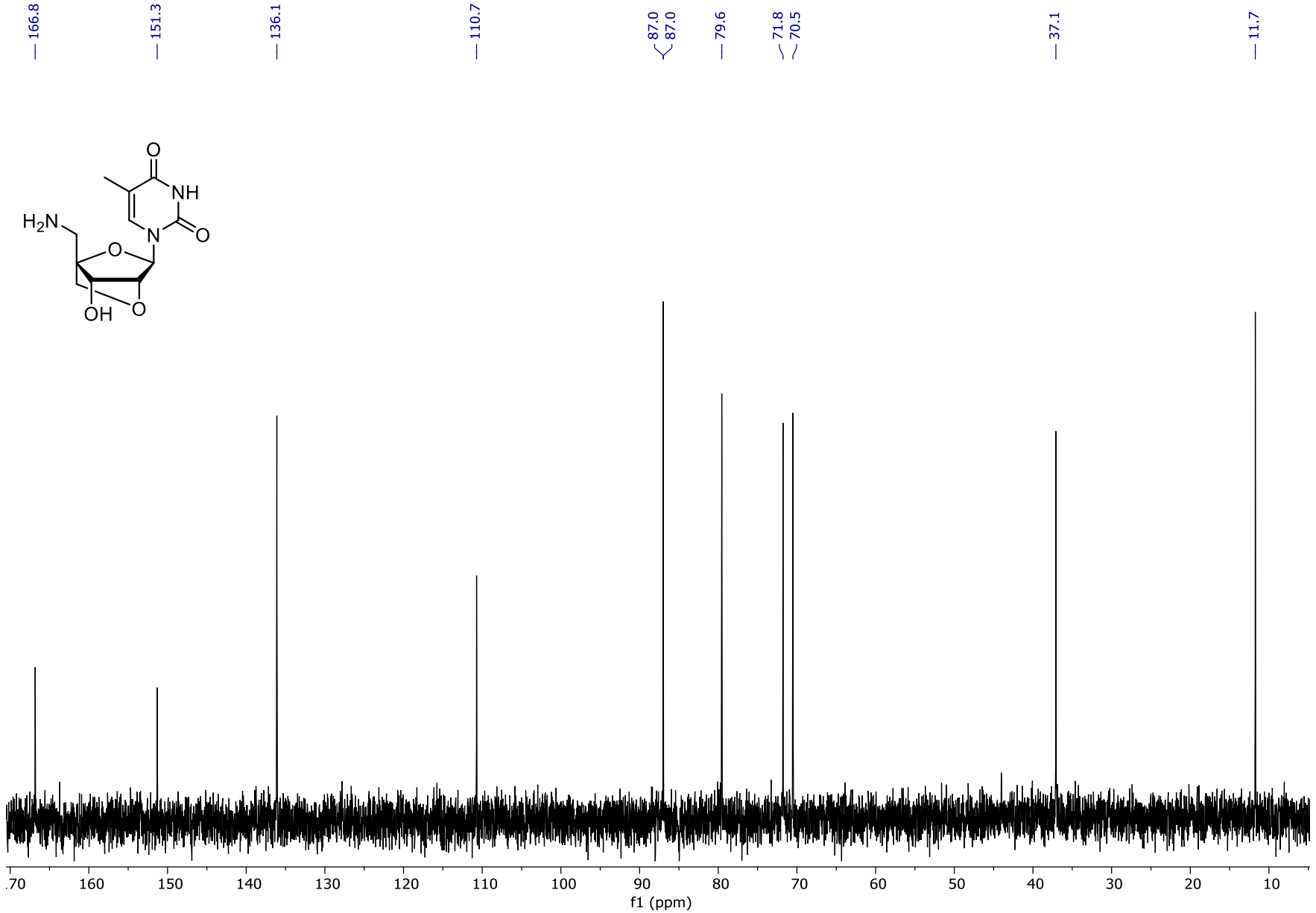

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*



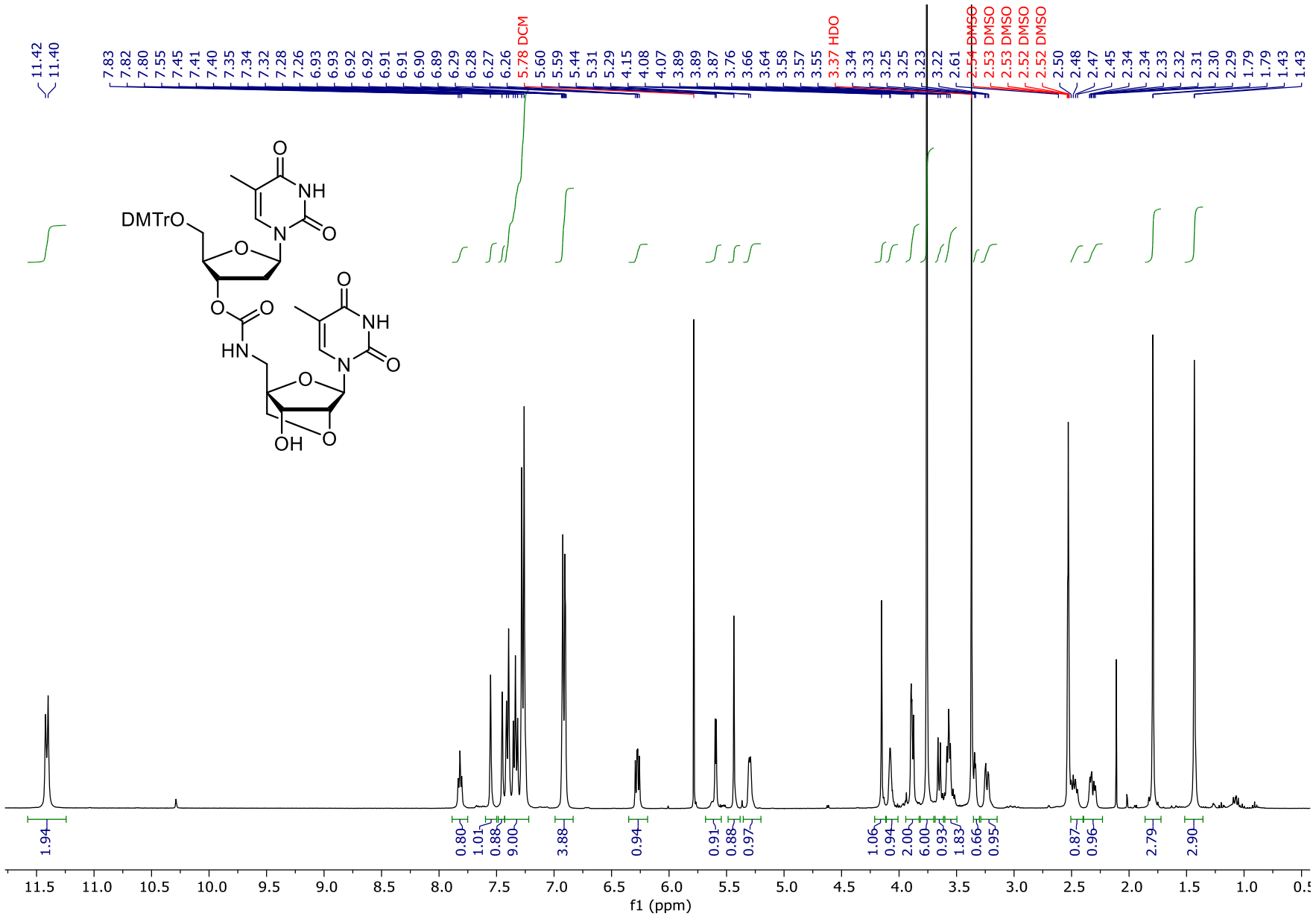

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry*.



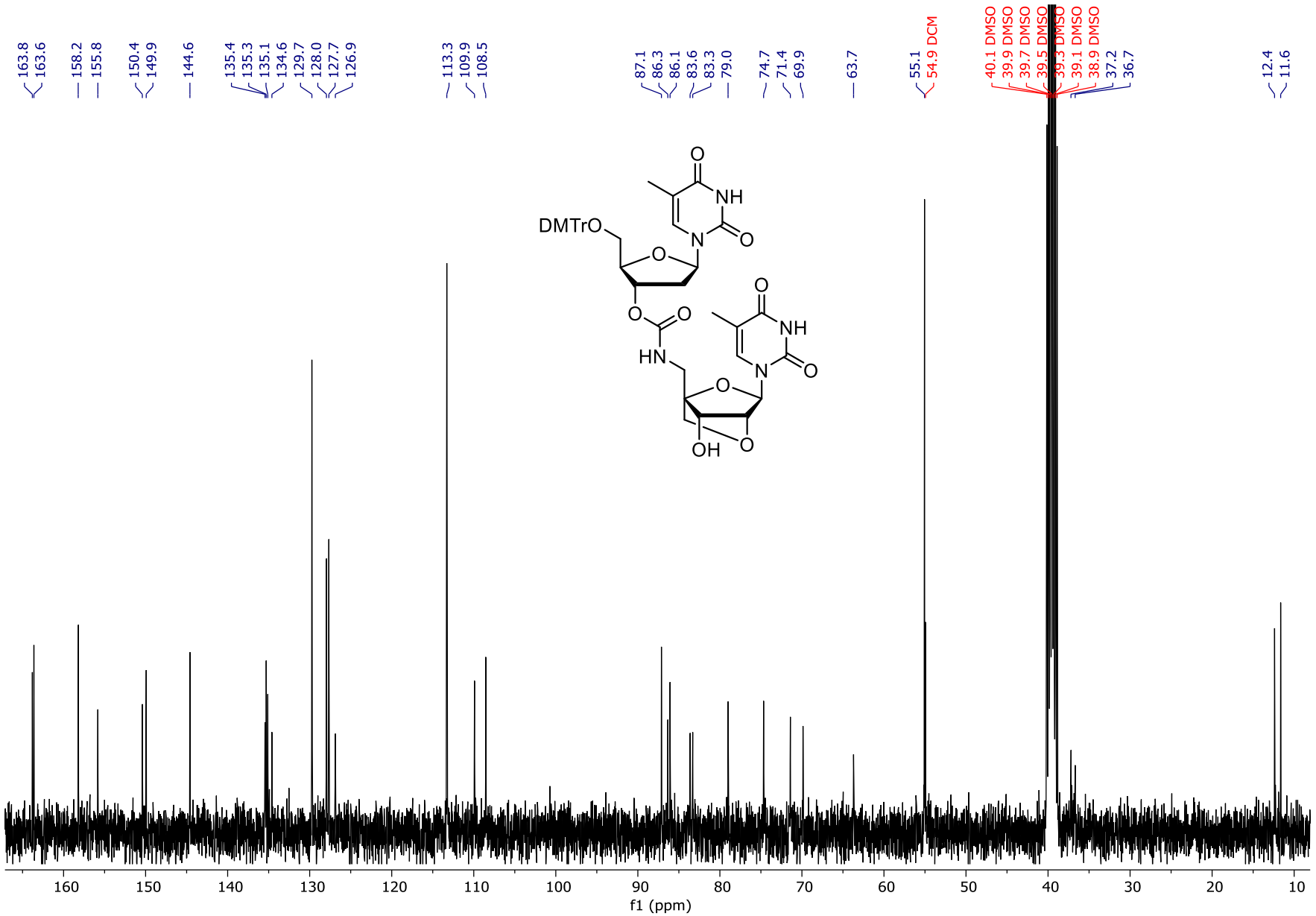

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry*.



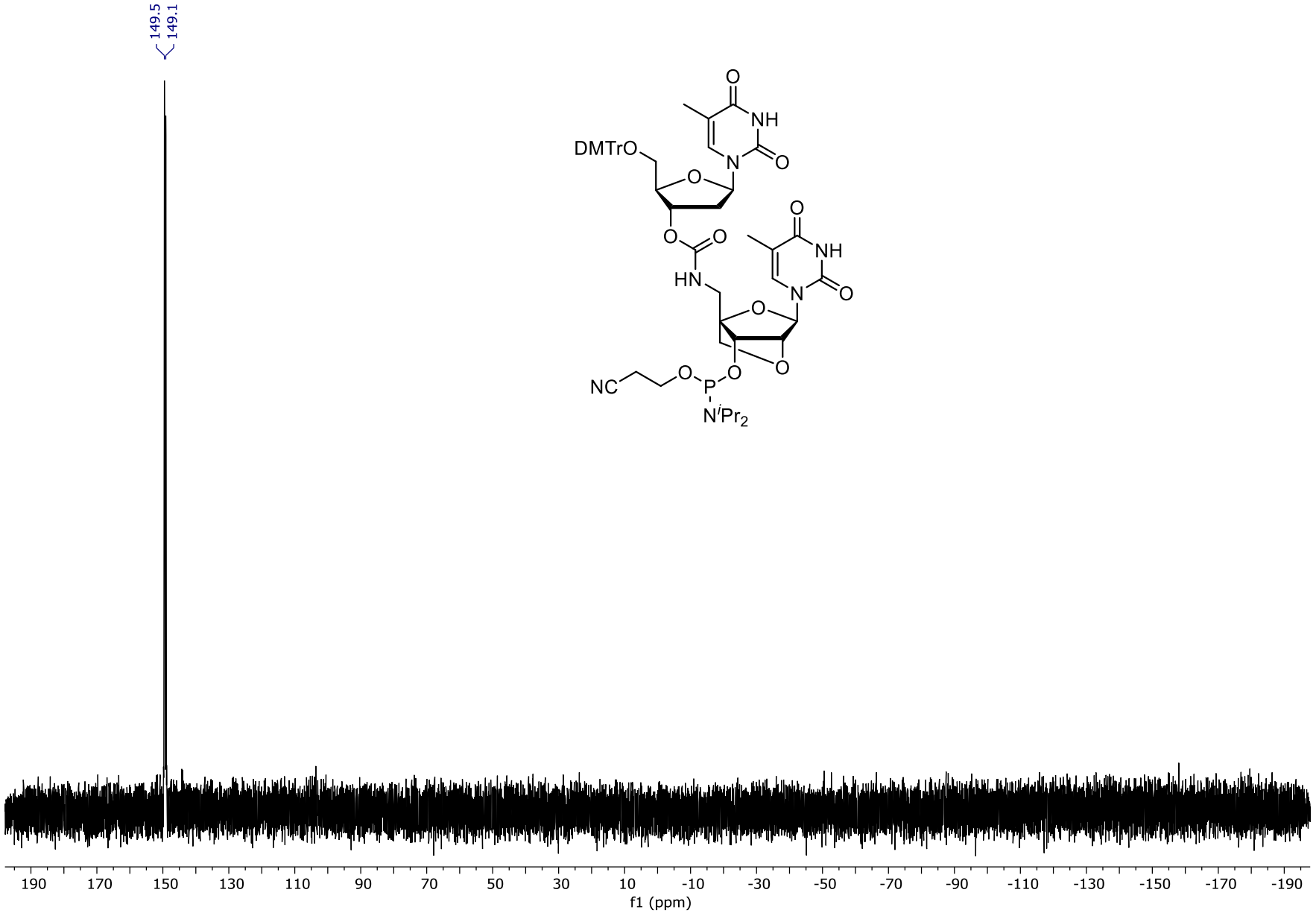

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry*.



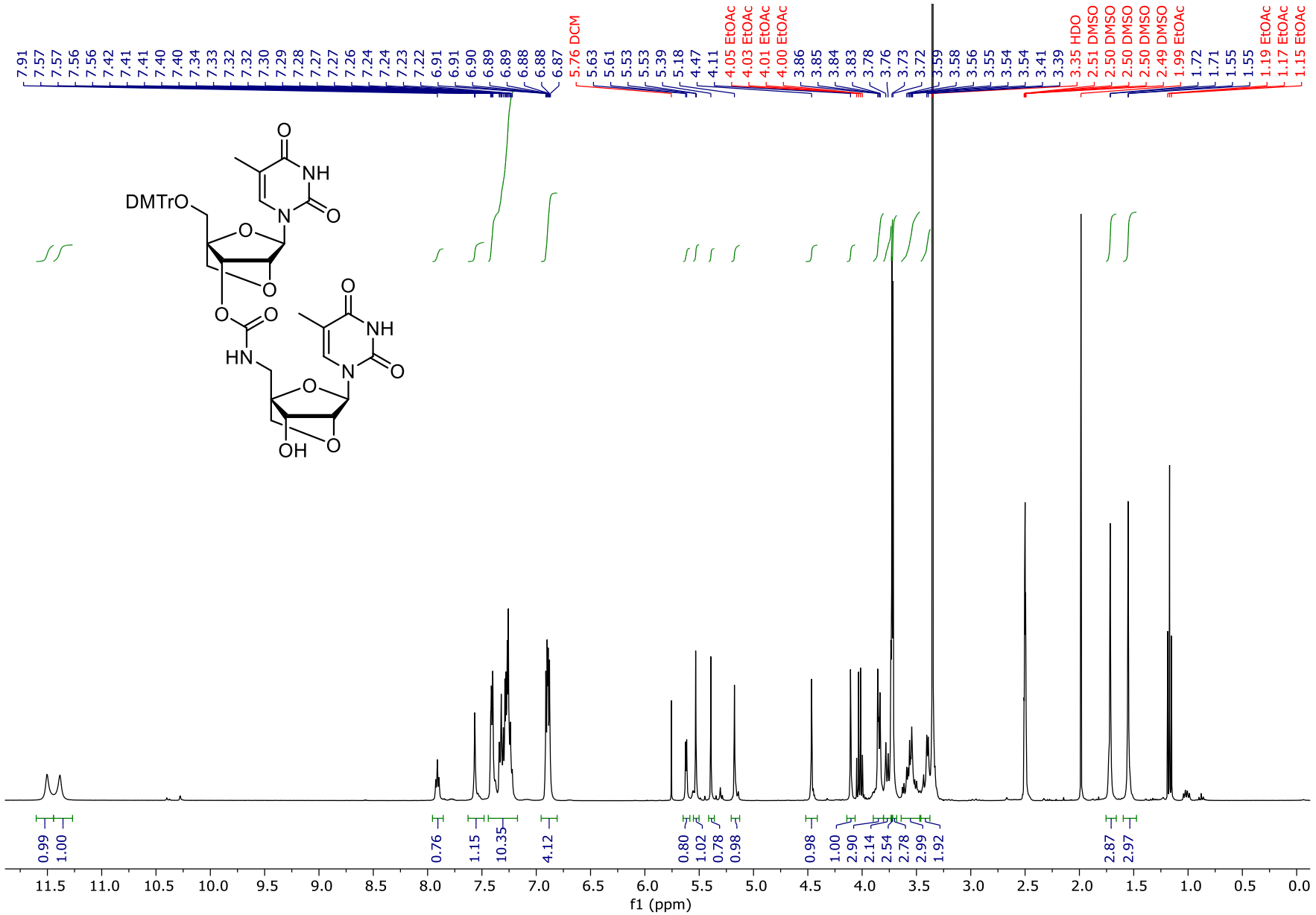

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*



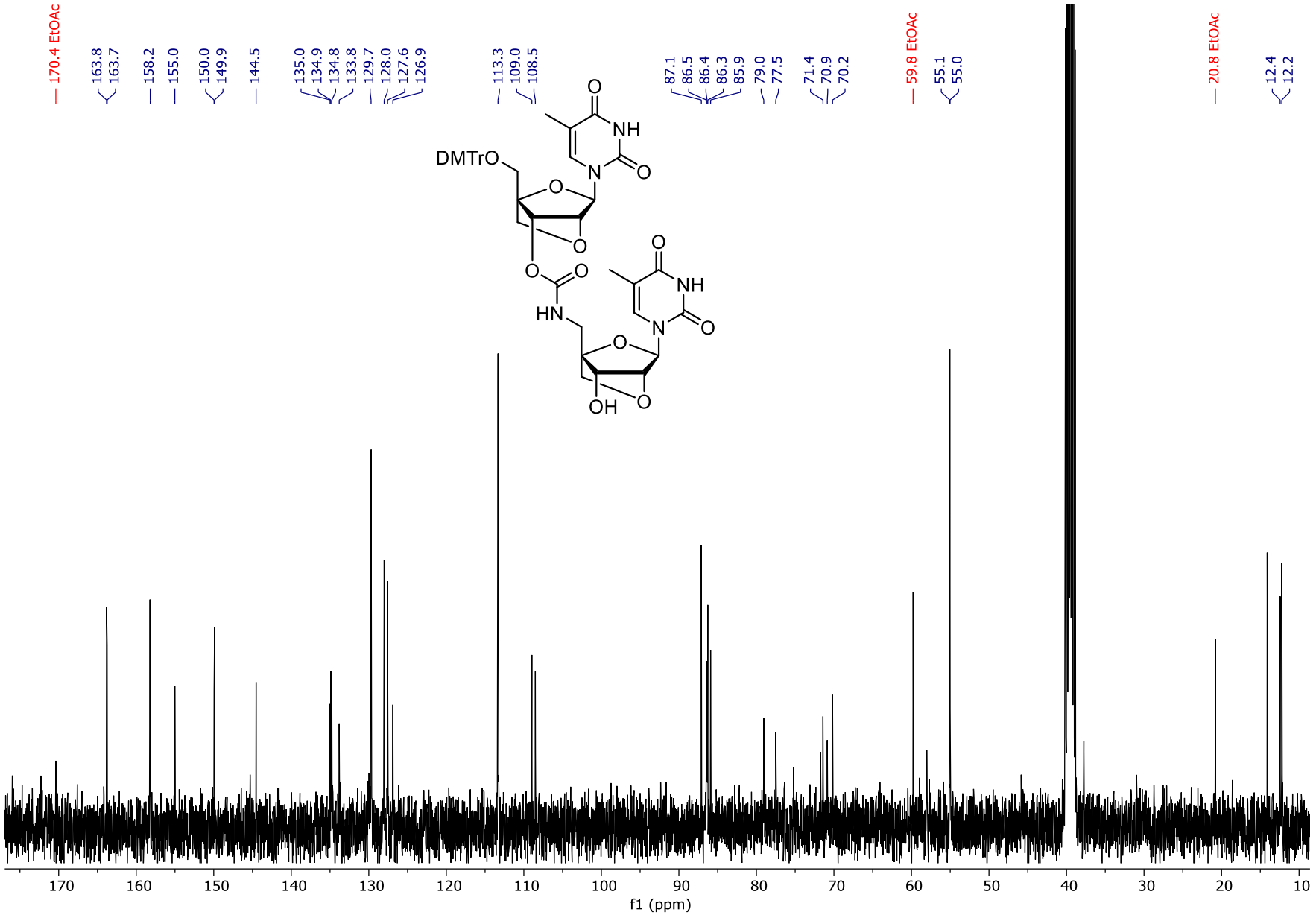

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*



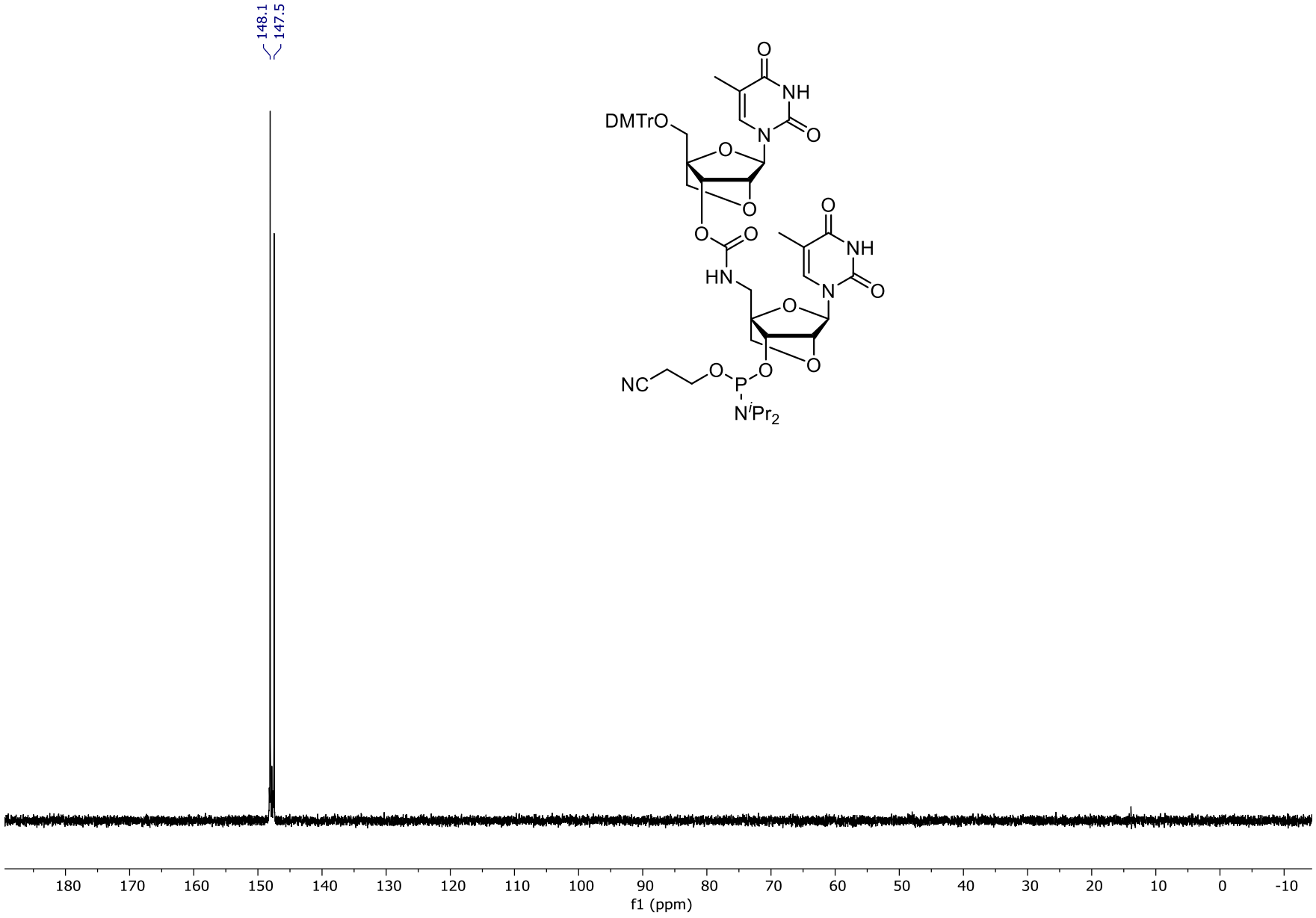

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry*.



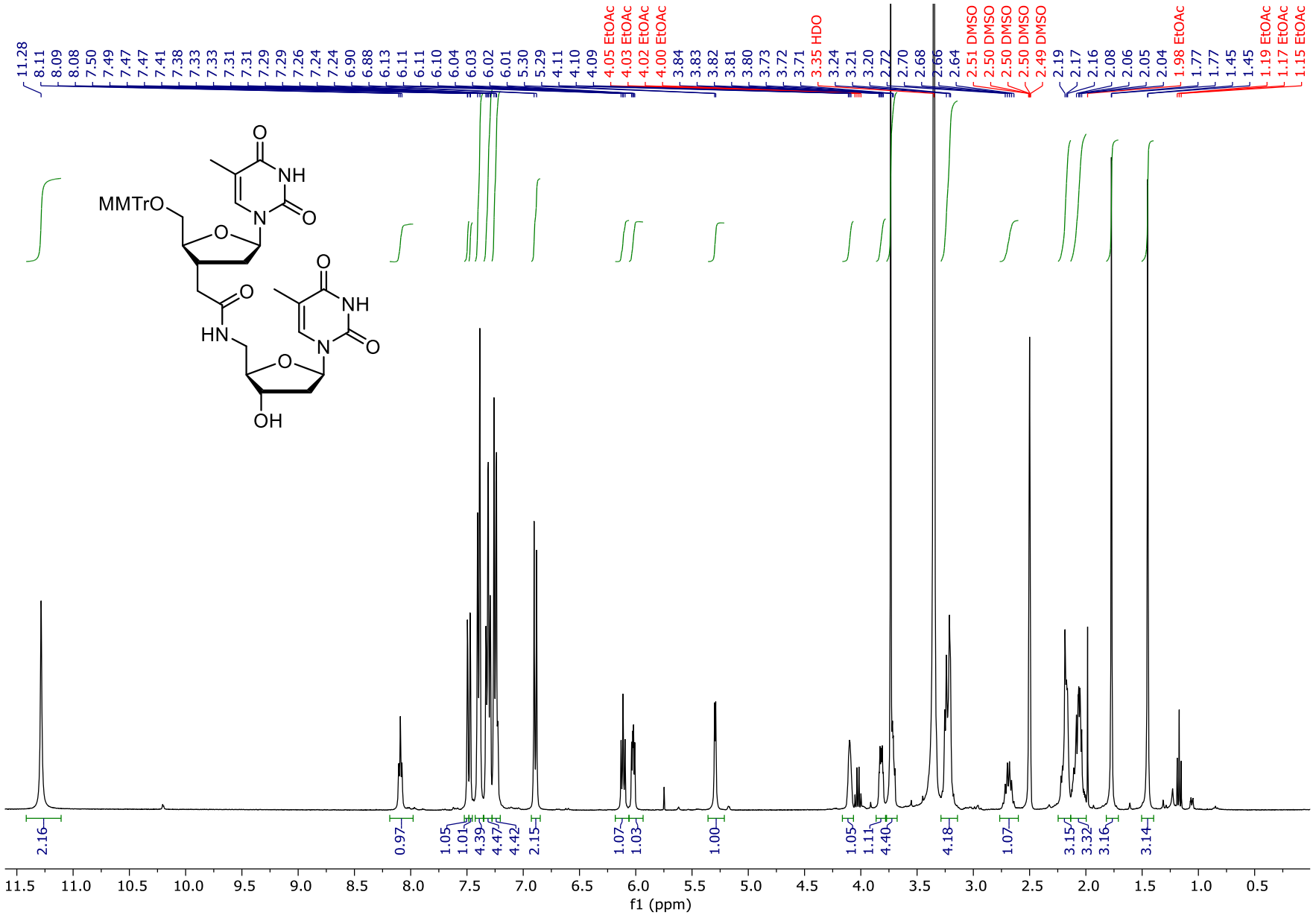

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*



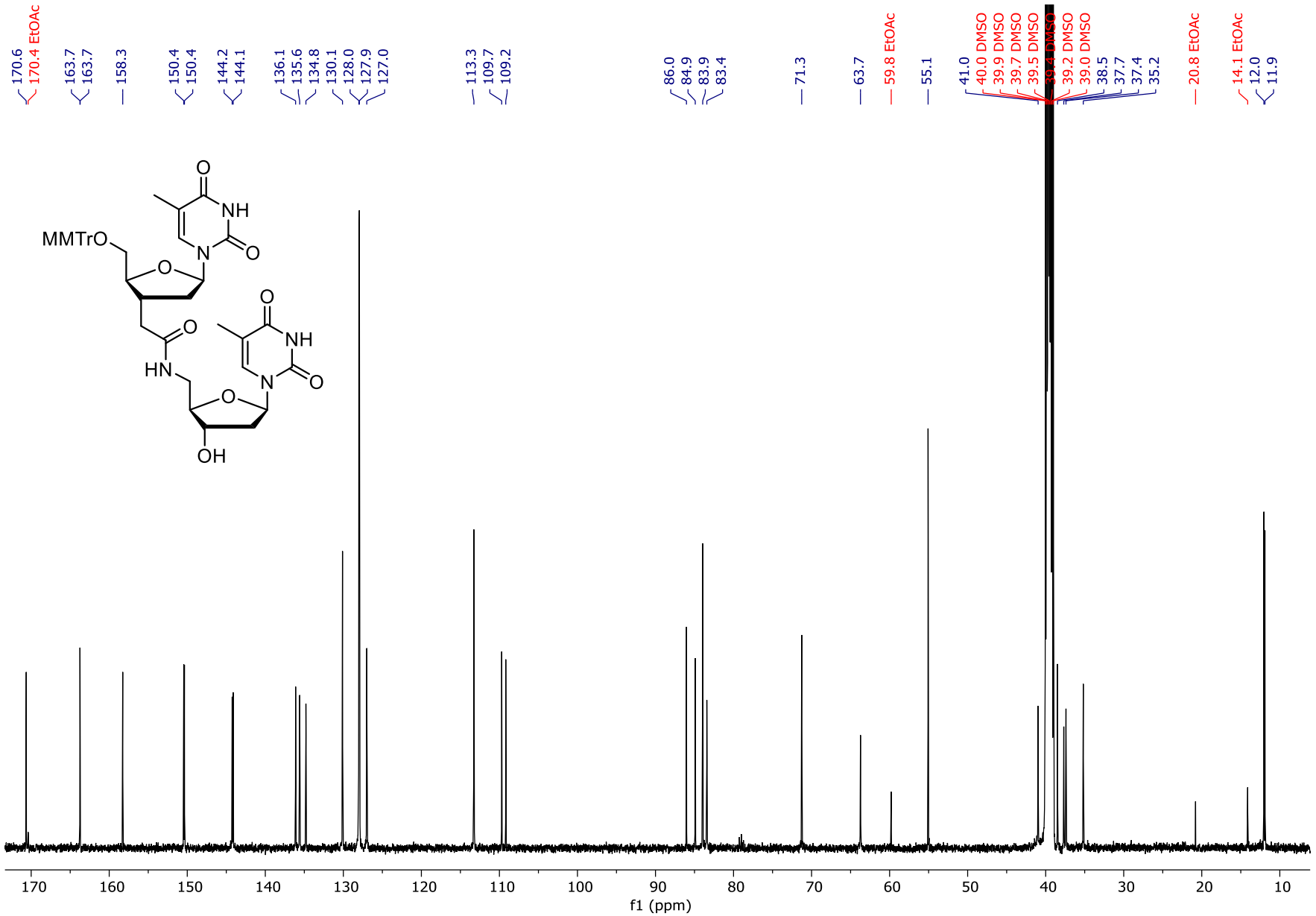

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry.*



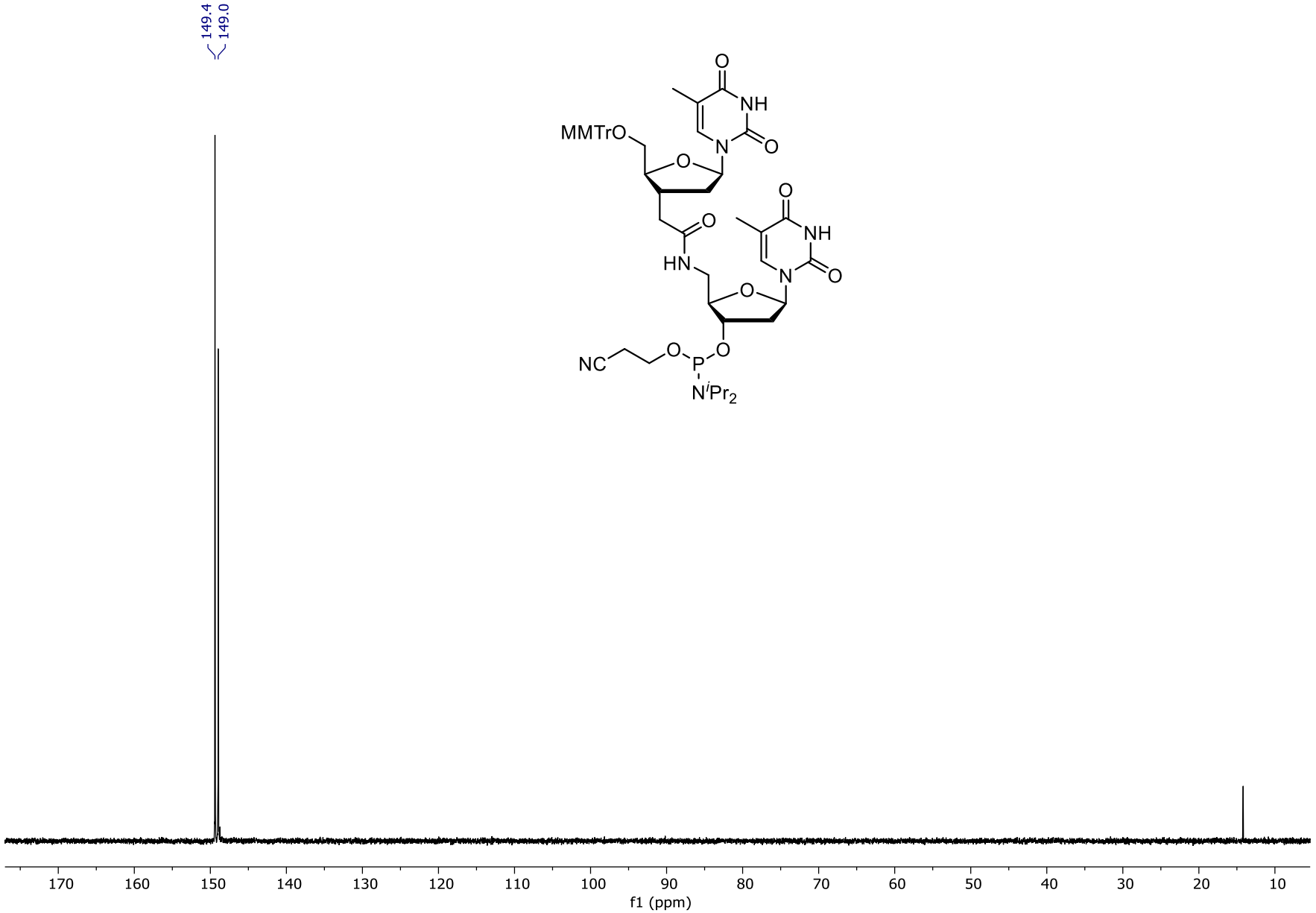

*Reproduced from Ref.*[182] *with permission from the Royal Society of Chemistry*.



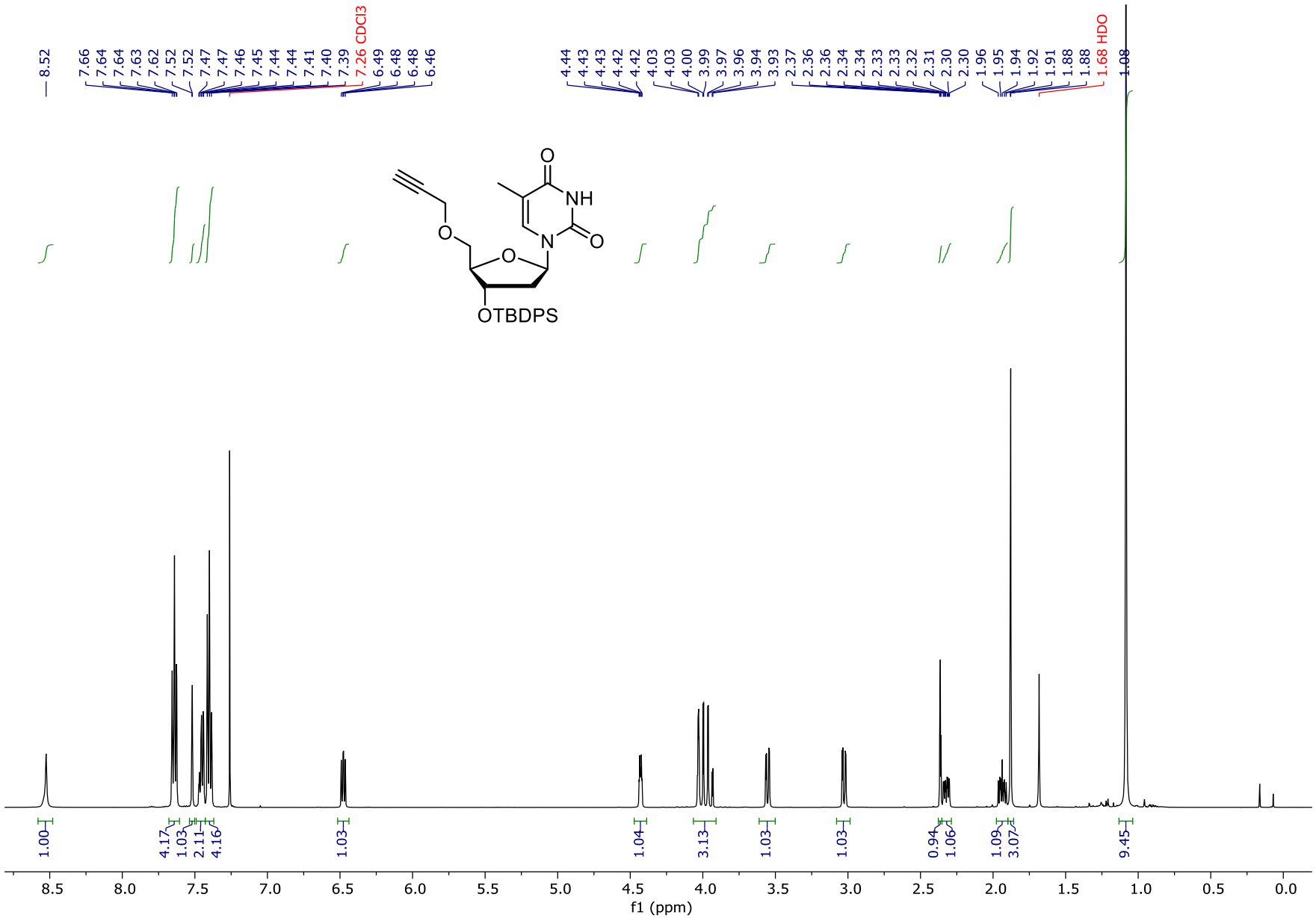

*Reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*



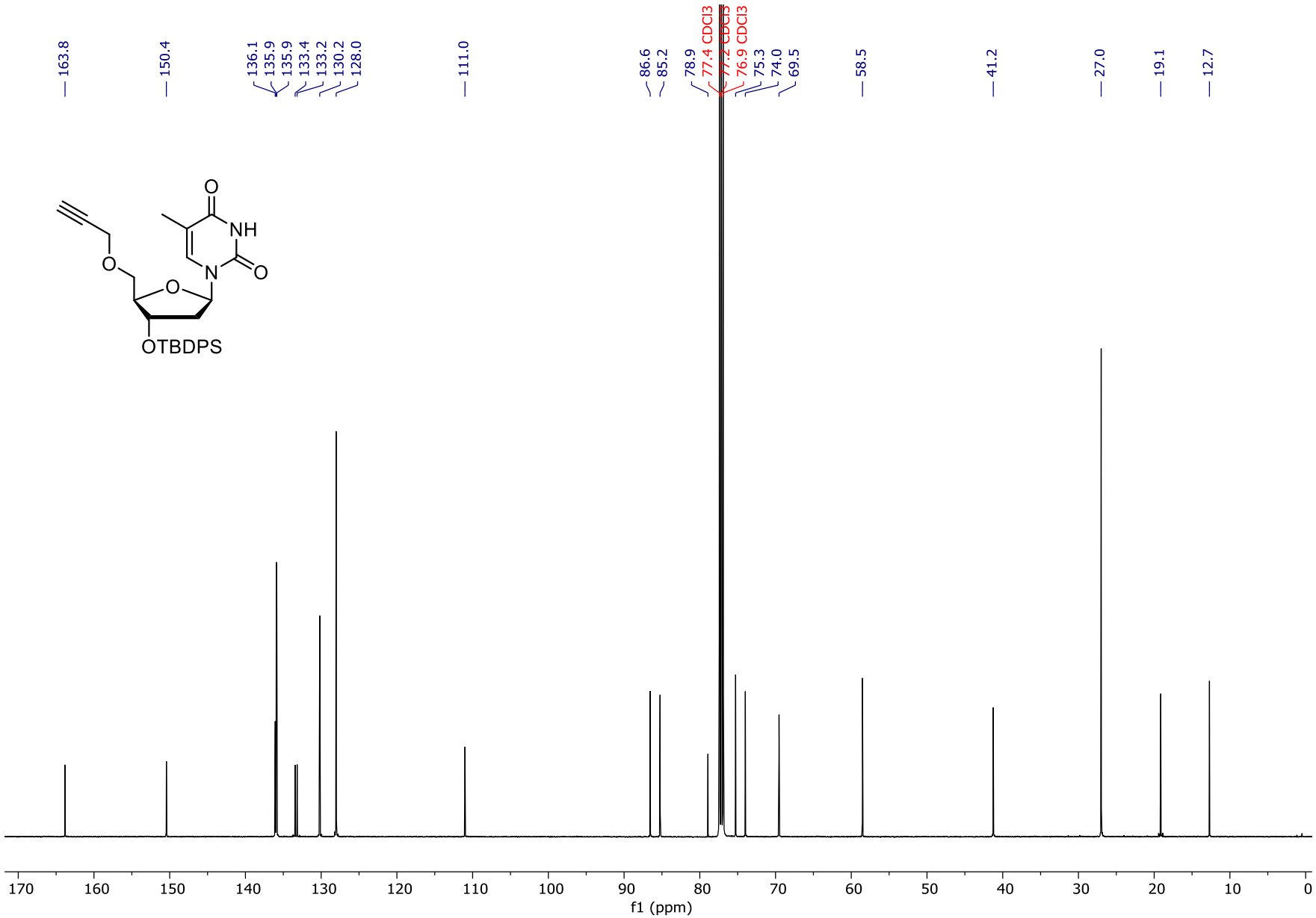

*Reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry.*



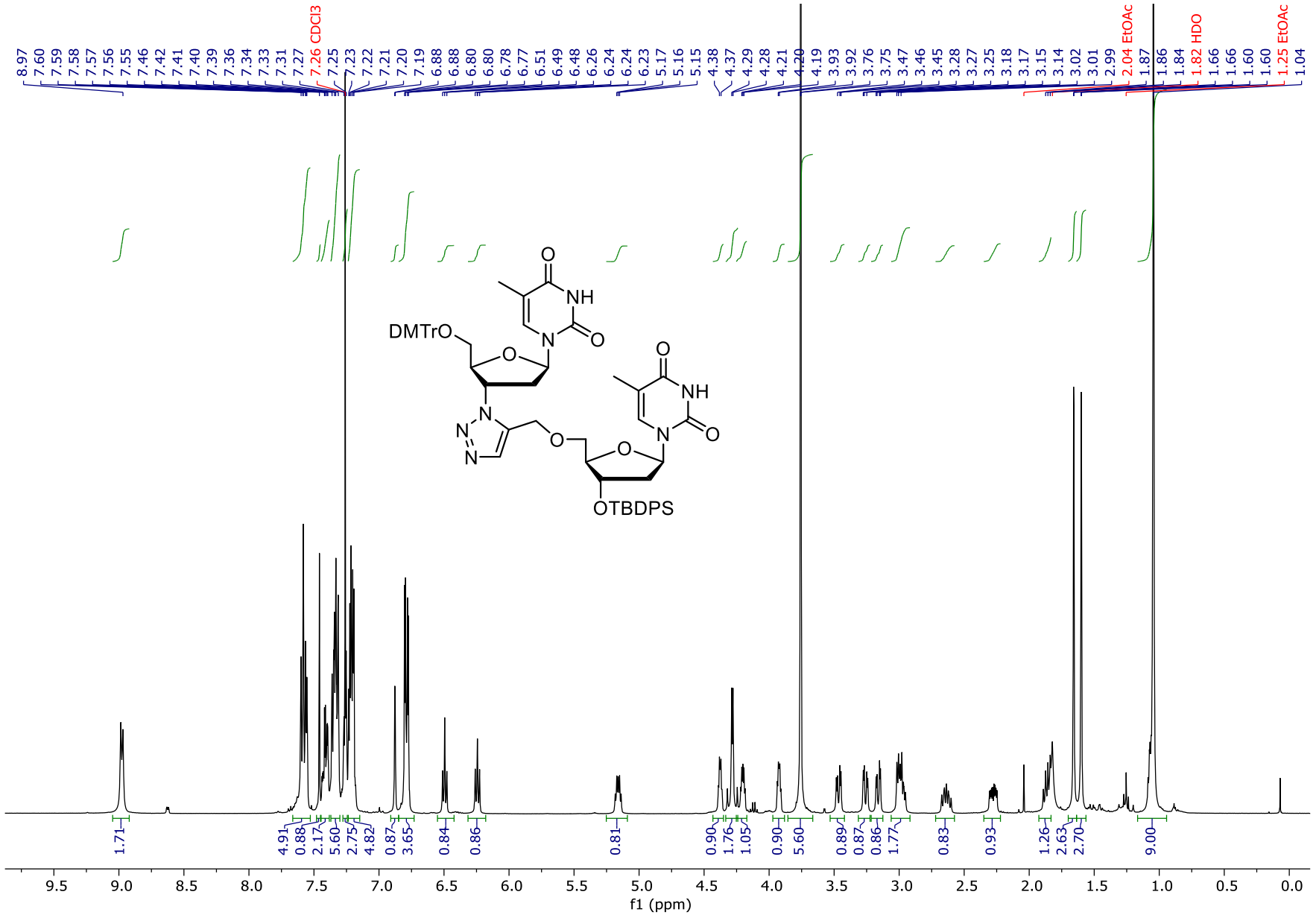

*Reproduced from Ref.*[308] *with permission from the Royal Society of Chemistry*.



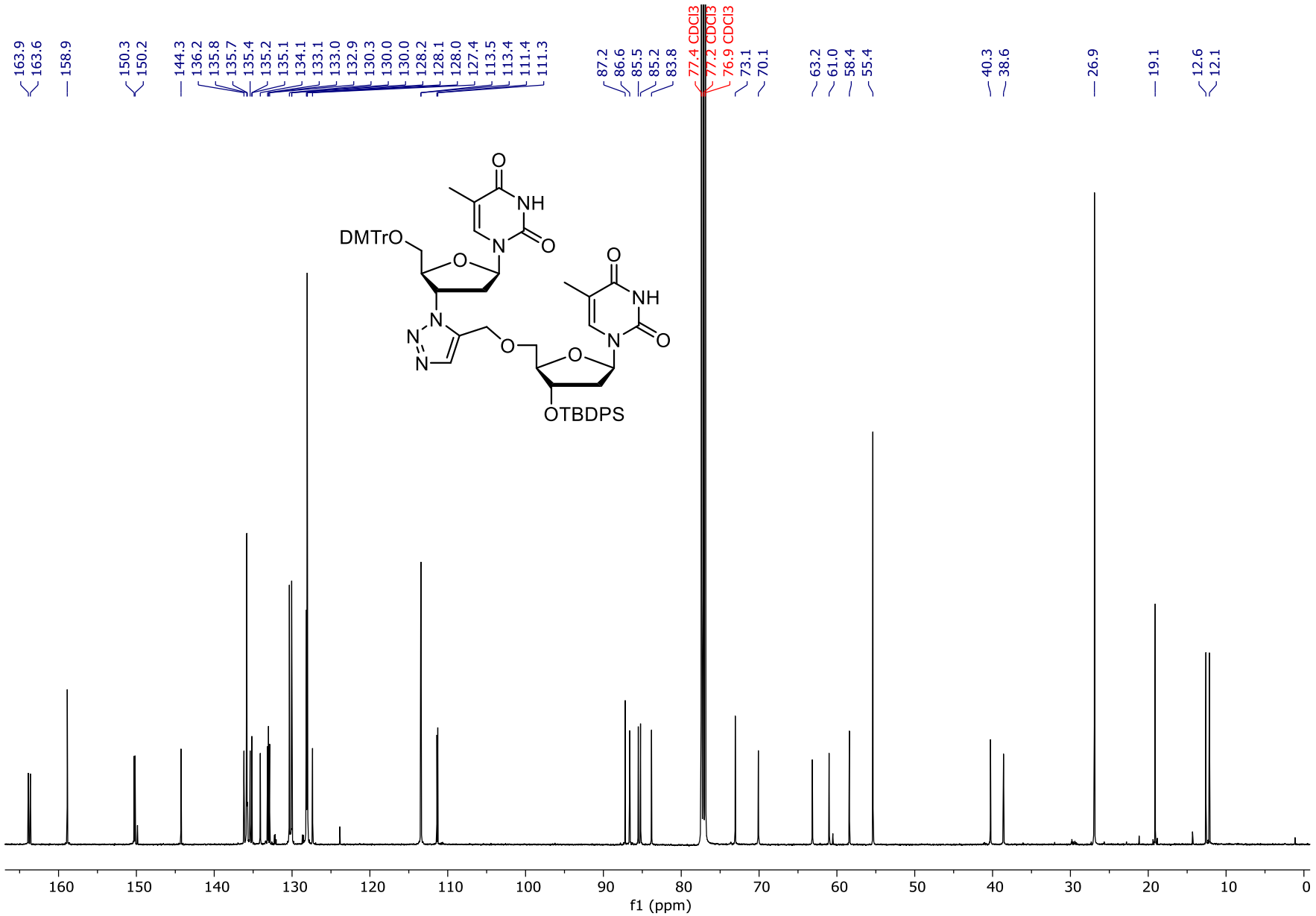

*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*



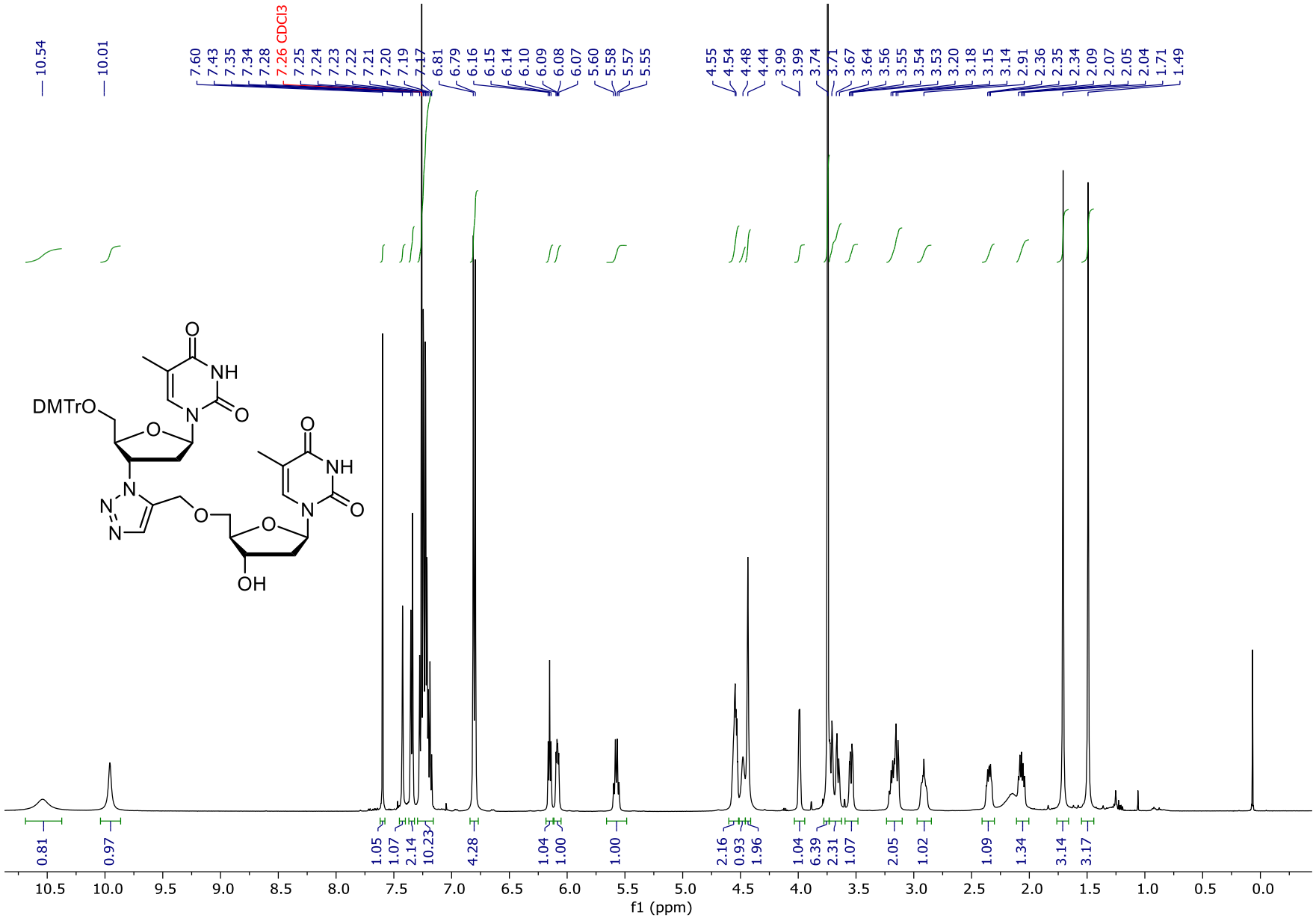

*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*



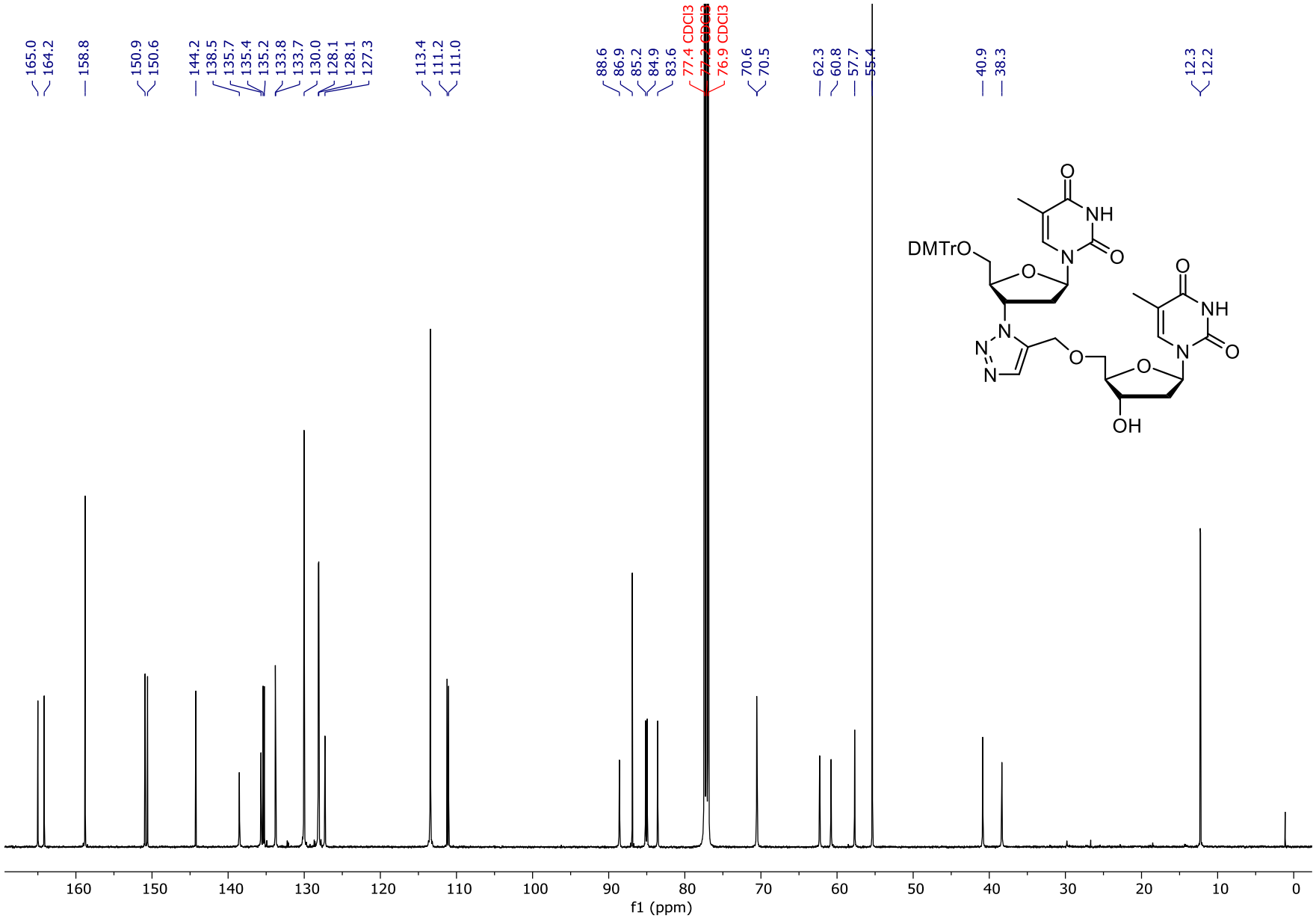

*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*



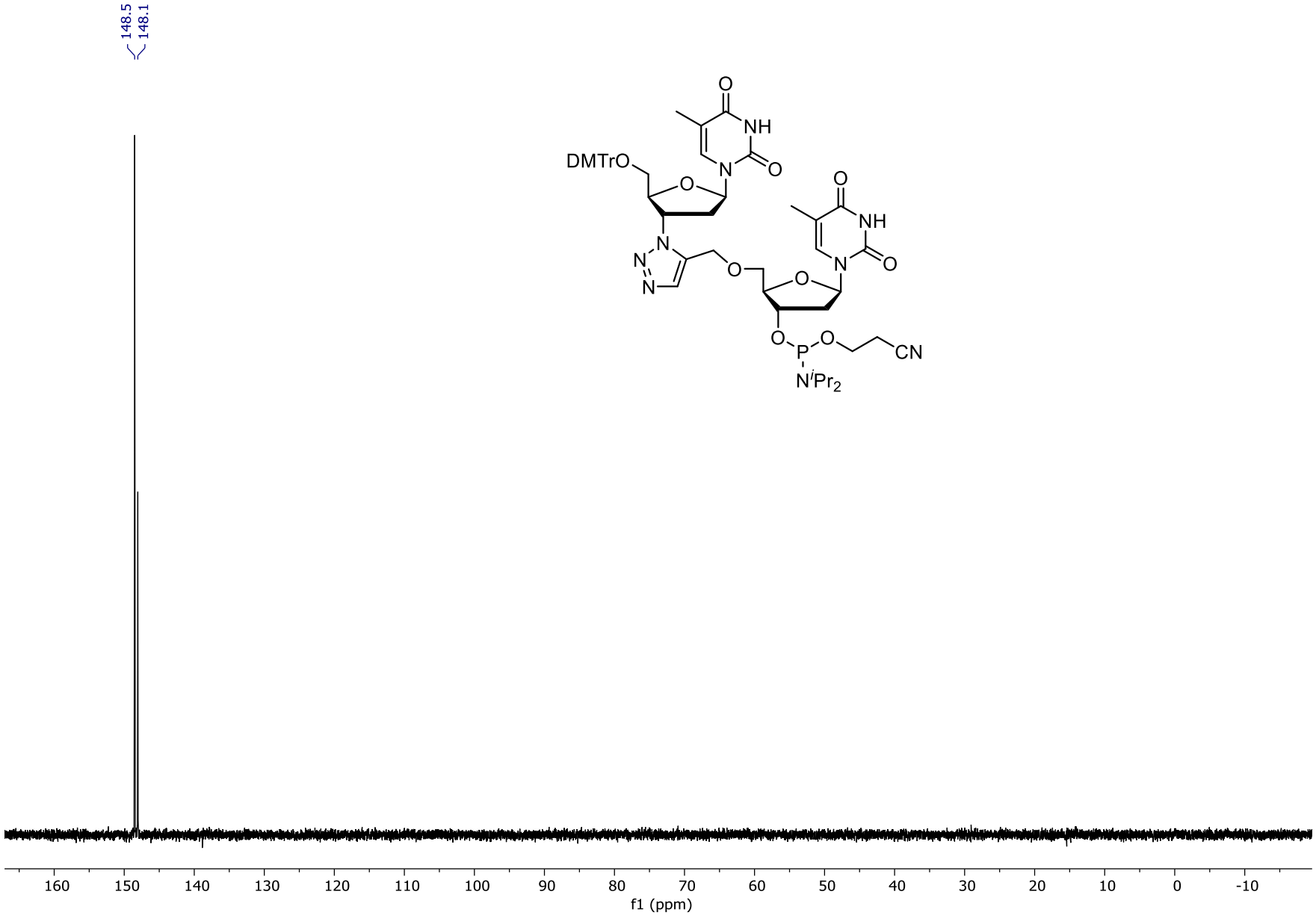

*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*




*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*

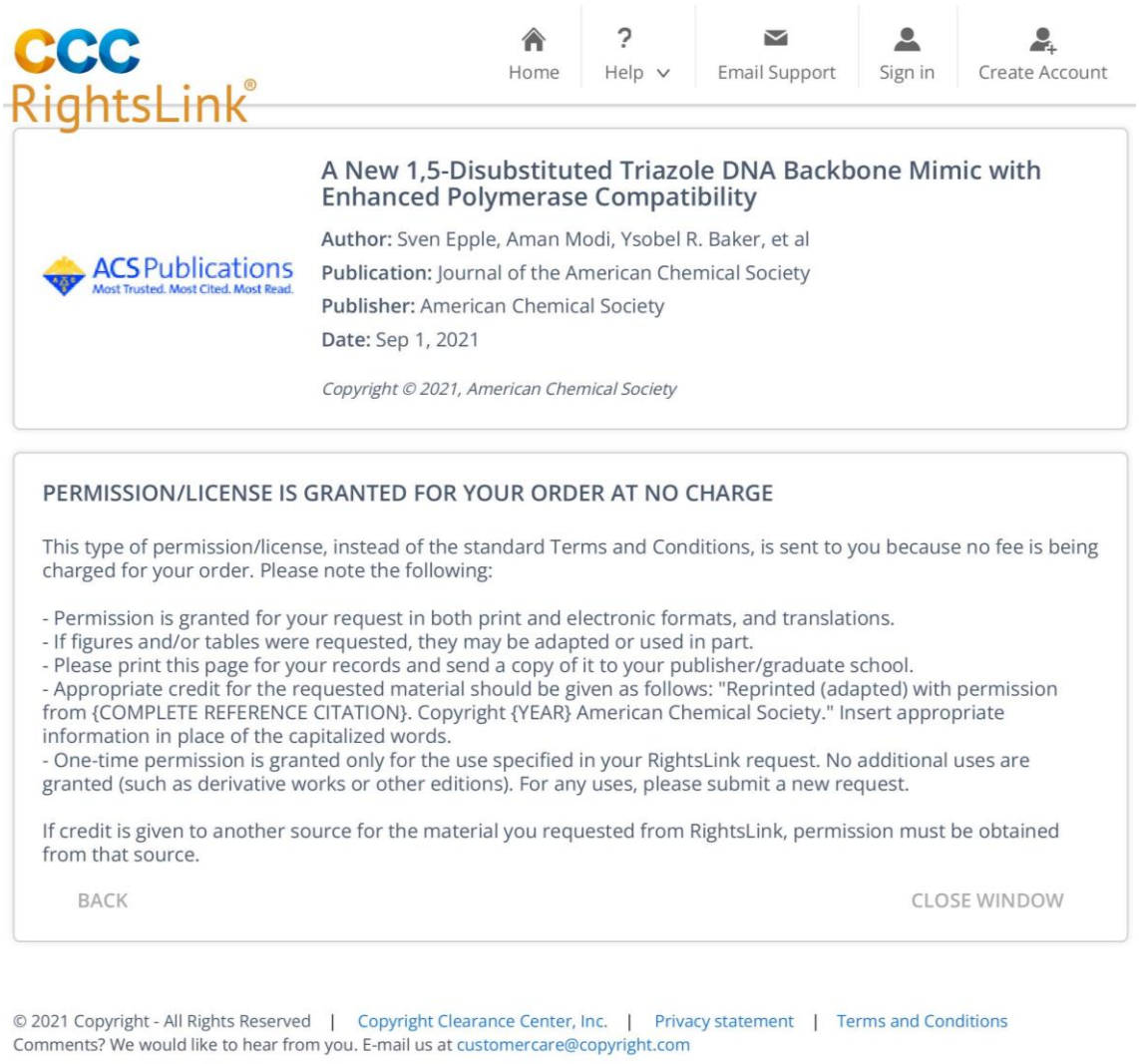



*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society.*




*Reproduced with permission from Ref.*[307] *Copyright 2021 American Chemical Society*.

# 4. Copyright



CCC RightsLink®

Home | Help | Email Support | Sign in | Create Account

**A New 1,5-Disubstituted Triazole DNA Backbone Mimic with Enhanced Polymerase Compatibility**

ACS Publications
Most Trusted. Most Cited. Most Read.

**Author:** Sven Epple, Aman Modi, Ysobel R. Baker, et al
**Publication:** Journal of the American Chemical Society
**Publisher:** American Chemical Society
**Date:** Sep 1, 2021

*Copyright © 2021, American Chemical Society*

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK CLOSE WINDOW

© 2021 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com