

Synthesis and Polymerase-Mediated Transcription of Base-Modified 2'-Fluoroarabinose Nucleic Acid in Preparation for Particle Display Selection with Modified Aptamers:

Author: Christopher J. A. Skrodzki

Persistent link: <http://hdl.handle.net/2345/bc-ir:108611>

This work is posted on [eScholarship@BC](#),
Boston College University Libraries.

Boston College Electronic Thesis or Dissertation, 2019

Copyright is held by the author. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0>).

Synthesis and
Polymerase-Mediated Transcription
of Base-Modified
2'-Fluoroarabinose Nucleic Acid
in Preparation for Particle Display
Selection with Modified Aptamers

Christopher J. A. Skrodzki

A Thesis

Submitted to the Graduate Faculty of

the Department of Chemistry

in Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

Boston College
Morrissey College of Arts and Sciences
Graduate School

August 2019

Copyright © 2019 Christopher J. A. Skrodzki

All rights reserved

SYNTHESIS AND POLYMERASE-MEDIATED TRANSCRIPTION OF
BASE-MODIFIED 2'-FLUOROARABINOSE NUCLEIC ACID
IN PREPARATION FOR PARTICLE DISPLAY
SELECTION WITH MODIFIED APTAMERS

Christopher J. A. Skrodzki, M.S.

Advisor: Jia Niu, Ph.D.

Nucleic acid aptamers are promising alternatives to antibodies for a wide array of diagnostic and therapeutic applications. However, state-of-the-art aptamers suffer from poor pharmacokinetics and diversity, limiting their affinity and specificity for many therapeutically relevant targets. The emerging field of glycoscience provides opportunities to improve the utility of aptamers over antibodies. Combining synthetic chemistry with modern molecular biology and polymer science, the synthesis of Xeno Nucleic Acid monomers and chemoenzymatic polymerization via engineered polymerase enzymes allows the production of nucleic acid drugs with superior resistance to endogenous nucleases. The modular structure of nucleic acids provides for the design of sequence-defined polymers capable of post-synthetically appending complex synthetic glycans, extending the catalytic geometry of aptamers. Our SELEX inspired FACS based particle display approach allows for high-throughput screening. Additionally, we expect this method has the capability of screening aptamers in human serum.

Our synthetic approach utilizes a Sonogashira cross-coupling reaction to install a flexible alkyne to the major groove of 2'-deoxy-2'-fluoro-arabinose uracil base. By incorporating recent advances in nucleic acid synthesis, one-pot nucleobase activation and sugar glycosylation is achieved and bis-oxybenzyl phosphoramidite synthesis can afford gram scale HPLC-free purification of the triphosphates. The FANA C8-alkyl-uridine triphosphate will be incorporated by an engineered Tgo DNA polymerase to allow systematic introduction of alkynyl conjugation handles into a DNA-templated FANA polymer. Subsequent conjugation with azido-modified glycans via the Huisgen copper-catalyzed alkyne-azide cycloaddition (CuAAC) click reaction will generate sequence controlled nucleic acid-carbohydrate hybrid molecules amendable for directed evolution.

TABLE OF CONTENTS

List of Figures	vii
List of Tables	viii
List of Abbreviations	ix
Acknowledgements	xi
Dedication	xii
Chapter One: Introduction	1
Advantages of XNA Technology over Traditional Nucleic Acids	2
Benefits of Glyco-mimetics and Polysaccharide Hybrid Materials	5
Promise of Aptamer Science over Traditional Affinity Reagents	7
Chapter Two: Literature Review	10
Historical Development of XNA Technology	10
Historical Overview of Glyco-mimetics and Polysaccharide Hybrid Materials	12
Chapter Three: Materials and Methods	19
Experimental Design	19
DNA transcription into 2'-Deoxy-2'-Fluoro-Arabinose Nucleic Acids (FANA)	20
Recombinant Tgo plasmid verification	21
Recombinant Tgo expression and purification	21
Recombinant Tgo purification QC	23
Polymerase screen for FANA Transcription Activity	24

Polymerase Dilution Series.....	27
Tgo Timecourse Optimization	28
FANA Library Isolation Using RNA PAGE Methods	29
DNA Polymerase Mediated FANA ‘Reverse’ Transcription into cDNA	33
Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)	34
Coupling of FP to Magnetic Microparticles	35
Surface Pacification of FP-Magnetic Microparticles.....	36
Testing Conjugation of FP-Magnetic Microparticles	37
Emulsion PCR Generation of Magnetic Microparticles Library	38
Testing Extension Efficiency of Emulsion PCR.....	39
Proposed – FANA ePCR & Click Conjugation of Aptamer Microparticles	40
Proposed – Screening of Modified Aptamer Microparticles	41
Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA.....	42
5-Iodouracil.....	43
2’-Deoxy-2’-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose	43
2’-Deoxy-2’-Fluoro-3,5-di-O-benzoyl- β -D-Arabinofuranosyl-5-Iodouridine	44
Trimethyl(octa-1,7-diyn-1-yl)silane	51
2’-Deoxy-2’-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl- β -C8-alkyl-Uridine....	52
Chapter Four: Results	54
DNA transcription into 2’-Deoxy-2’-Fluoro-Arabinose Nucleic Acids (FANA)	54
Polymerase screen for FANA Transcription Activity	54
Polymerase Dilution Series.....	56

Tgo Timecourse Optimization	57
DNA Polymerase Mediated FANA ‘Reverse’ Transcription into cDNA	59
Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)	60
Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA.....	61
5-Iodouracil.....	62
2’-Deoxy-2’-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose	63
2’-Deoxy-2’-Fluoro-3,5-di-O-benzoyl- β -D-Arabinofuranosyl-5-Iodouridine	63
Trimethyl(octa-1,7-diyn-1-yl)silane	67
2’-Deoxy-2’-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl- β -C8-alkyl-Uridine....	69
Chapter Five: Discussion	71
DNA transcription into 2’-Deoxy-2’-Fluoro-Arabinose Nucleic Acids (FANA)	71
Polymerase screen for FANA Transcription Activity	71
DNA Polymerase Mediated FANA ‘Reverse’ Transcription into cDNA	72
Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)	72
Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA.....	73
Final Thoughts	73
References Cited	75
Appendices.....	A1

LIST OF FIGURES

Figure 1. – Purification of Tgo polymerase expressed in XL1-blue cells.	24
Figure 2. – Tgo-like polymerase displaying known gof RNA transcription mutations....	26
Figure 3. – 12.5 % 29:1 PAGE extraction of 100X FANA XPE & 10 % analytical gel.	32
Figure 4. – DNA polymerase screen for DNA templated FANA transcription.....	55
Figure 5. – Tgo dilution series indicating laddering due to polymerase concentration...	57
Figure 6. – Timecourse optimization of Tgo Cy5-primer extension into FANA library	58
Figure 7. – Preliminary ‘reverse’ transcription of FANA library with Cy3-primer	59
Figure 8. – Conjugation of Cy5-DNA FP to Dynabeads MyOne CA magnetic beads. ...	60
Figure 9. – Proposed scheme for synthesis of FANA C8-alklyl-uridine triphosphates....	62
Figure 10. – Nucleobase uracil iodine halogenation mediated by CAN oxidation	63
Figure 11. – Bz protected 2-F-arabinose bromine halogenation in glacial acetic acid.....	65
Figure 12. – X-ray diffractometric structure of FANA glycoside precursor 2''	66
Figure 13. – Fraction distillate GC-MS trace containing target product 13 TMS-Linker	68
Figure 14. – Variable FANA-Iodouracil Pd cross-coupling to a variable alkyne linker. .	70

LIST OF TABLES

Table 1 – Oligonucleotides used within this study, and their respective sequence.	20
Table 2. – Summary of different FANA glycoside precursor glycosylation conditions. .	64
Table 3. – TMS Linker distillation with regards to different data sources.	68
Table 4. – Summary of different FANA -Iodouracil Pd cross-coupling conditions.....	70

LIST OF ABBREVIATIONS

Å – Ångström, 0.1 nm

Ab – Antibody

bs-mAbs – bispecific monoclonal Antibodies

sdAb – single-domain Antibodies

Nb-mAbs – monoclonal Nanobodies

Fab – antigen-binding antibody Fragment

Fc – crystallizable antibody Fragment

scFv – single-chain variable antibody Fragment

V_{HH} – ‘camelid’ Variable Heavy-chain Heterodimer domain

V_L – Variable Light-chain domain

ACN – ACetoNitrile

BSA – n,o-Bis(trimethylSilyl)-Acetamide

Bz – Benzoyl

CAN – Ceric Ammonium Nitrate

CST-SCA – Compartmentalized Self-Tagging-Statistical Correlation Analysis

cryo-EM – cryogenic Electron

Da – Dalton, g·mol⁻¹

DAST – DiethylAminoSulfur Trifluoride

ddH₂O – double distilled/ultra-pure water

DDS – Double bond Diethylenetriamine Succinic acid

DIPEA – n,n-DiIsoPropylEthylAmine

DMF – DiMethylFormamide

DMSO – DiMethylSulfOxide

DMPA – 2,2-DiMethoxy-2-PhenylAcetophenone

DNA – 2’-Deoxyribose Nucleic Acid

EDC – Ethyl-3-(3-Dimethylaminopropyl)Carbodiimide

EDS – EthyleneDioxy Succinic acid

EDTA – EthyleneDiamineTetraAcetic acid

Fmoc – Fluorenylmethyloxycarbonyl

Alloc – Allyloxycarbonyl

Boc – tert-Butoxycarbonyl

gof – gain-of-function
HMDS – HexaMethylDiSilazane
IPTG – IsoPropyl β -D-1-ThioGalactopyranoside
Kd – kinetic Konstant of dissociation
KOD – *thermococcus KODakaraensis* polymerase
MicroED – micro Electron Diffraction
NCL – Native Chemical Ligation
nt(s) – nucleotide(s)
xNTPs – xeno Nucleotide TriPhosphates
PCR – Polymerase Chain Reaction
PEG – PolyEthylene Glycol
PG – Protecting Group
PMO – Phosphoroamidate Morpholino Oligomer
PNA – Peptide Nucleic Acid
RNA – Ribose Nucleic Acid
 tRNA – transfer RNA
SNP – Single Nucleotide Polymorphism.
SPA – Single-Particle Analysis
SPPS – Solid-Phase Peptide Synthesis
STA – SubTomogram Averaging
TEA – TriEthylAmine
TBTA – Tris((1-Benzyl-4-Triazolyl)methyl)Amine
TCEP – Tris(2-CarboxyEthyl)-Phosphine hydrochloride
TDS – Triple bond functionalized diethyleneTriamine Succinic acid
TEC – Thiol-Enol Click
TFA – TriFluoroAcetic acid
TFO – Triplex Forming Oligonucleotides
TLC – Thin Layer Chromatography
TMSOTf – TriMethylSilyl Trifluoromethanesulfonate
UAA – Unnatural Amino Acid
UTR – UnTranslated Region
XNA – Xeno Nucleic Acid
XPE – XNA Primer Extension
XRD – X-Ray Diffraction crystallography
2'MOE – 2'-o-MethOxyEthyl
2'OMe – 2'-O-Methyl

ACKNOWLEDGMENTS

I would like to first and foremost sincerely thank Dr. Niu for helping me maintain a clear sense of focus over the year encompassing this experimental work, willingness to include me in correspondences with collaborators, excitement and trust to share my preliminary results with visiting speakers, and for inexorable patience in assisting me through this entire process.

This material is based upon work supported by Boston College Dept. of Chemistry start-up funds for new investigators.

Additionally, I cordially thank Chao Liu, Dr. Cangji Yang, and Dr. Hanchu Haung for experimental support *and* the occasional good bit of lab banter.

To my wonderful wife and son,
Although I did the mental work, you sacrificed the most of us all.
~Thank you for being there, even when I wasn't.

CHAPTER ONE

Introduction

In the wake of the “protein-drug revolution” a new frontier of drug amiable chemistries has begun gaining traction, stemming from the foundations of traditional small molecule organic synthesis and biochemistry- nucleic acid and glycomimetic therapeutic agents. These new models of study are very well suited to the emerging field of *Chemical Biology*, where a deep understanding of organic chemical synthesis and Biochemistry does more than just empower biological research, as per biological chemistry, but highly suggests even *demand*s not just the analysis but also development of novel biological polymers. Although there are significantly developments worth noting in the development of synthetic peptides, aptly named polyamides, as per the amide bond linking joining each monomer, or more exotically as *xenoproteins*¹; however, at the core of these developments is a sense of ‘similarity’ to natural proteins or the ability to augment natural systems to include unnatural amino acids (UAAs), either pre-/post-synthetically or in vivo systems through directed evolution. A light etymological analysis suggests a reason as to perhaps why there is a subsequent distance between the adoption of the term *xeno* in the context of proteins- stemming from the Ancient Greek *ξένος* (*xénos*) for stranger or foreigner, this affix contextually describes any species that exhibits unusual characteristics from within their class; surely, since proteins are arguably the state-of-the-field in biotechnology and modification of amino acid into varying flavors of β or γ linked side chains does little to obfuscate the similarity of protein-like

polymers. This is perhaps why the term *xeno nucleic acid* (XNA) has found moderately more acceptance in literature.

Advantages of XNA Technology over Traditional Nucleic Acids

While it is attractive to believe that the foundations of XNA research stem from the discovery of ribozymes, RNA molecules capable of acting similarly to proteins as the ‘functional’ elements of life and not necessarily as purely ‘informational,’ the true origins are more-likely to be in the creativity of early 20th century chemists and partially rooted in the discovery of the structure of DNA^{2,3}. While certainly there is something to be said concerning ribozymes’ ability to have the same chemical species, act up- & down-stream in molecular biological central dogma, XNA polymers have technological value due to three key factors: 1) multi-stranded character of nucleic acids allows redundant self-encoding; 2) modular design composed of generally three main motifs- nucleobase, sugar ring, and backbone linkage; & 3) single duplex stabilization, through dihedral angles or π - π stacking.

The multi-stranded character of nucleic acids is best known by Watson, Crick, and Franklin double helix model^{4,5}; yet, the amphiprotic characteristics of nucleic acid bases also allows the formation of triplex helices through a combination of Watson-Crick and Hoogsteen base pairing of multiple base-pairs⁶, given the correct sequence or modified chemical backbone such as aminoethyl glycine Peptide Nucleic Acid (aegPNA)⁷⁻⁹. Biologically, each strand of the double helix allows nucleic acids to have the unique property to self-template their partner strand, allowing sequence-specific replication of the genetic information of a *living* system through a sort of insured “self-redundancy.” This also provides for flexibility to polymer by defining an ability to specifically identify

a target, biologically its own complement strand, but this need not necessarily be the case. Indeed, this is the structural basis of aptamer technology. Additionally, the ability for nucleic acids to form complex secondary structure triplexes with similar sequences of nucleotides is a common feature within transfer RNA (tRNA) nucleotides and forms the basis of the development of minor-groove Triplex Forming Oligonucleotides (TFOs) for the purpose of specific DNA duplex interaction and subsequent gene silencing ^{10,11}, similar to function of natural triplex and G-quadruplexes within genetic Untranslated Regions (UTRs) notable telomeres ^{12,13}. Perhaps then it is no surprise that even Linus Pauling and his contemporaries considered a triplex helix structure ^{14,15}. Hence, modification of natural nucleic acid bases has also allowed the generation of even higher unnatural multiplexes ¹⁶.

This is possible due to the modularity of nucleic acids as a polymer. While the term XNA seemed to have stemmed originally from an \underline{x} NA denotation the different \underline{x} sugar ring incorporated within a specific species of nucleic acid, such as D for deoxyribose or R for Ribose, this term has grown in scale to essentially detonate any unnatural/synthetic nucleic acid structure- including modified bases and backbones, regardless of their authors' original intentions ^{17,18}. This "XNA Zoo," so-to-speak, exists because nucleic acid structures are generally tolerant of modification to each of the three elements in a mutually exclusive manner; some even going so far as to classify any possible XNA as existing along a three-dimensional modified base, sugar, backbone continuum ¹⁹. This arguable classifies XNA as a side-chain polymer with respect to their sugar moieties, which are mainly responsible for determining the steric effects and thus thermal stability of XNAs ²⁰

The sugar ring of XNAs are a likely contributor to the ability for XNAs to form stable helices, as it is a well-known phenomenon that polysaccharides are capable of forming complex helices in solution ²¹⁻²³. The effect of these sugar rings to affect the resulting secondary structure has been suggested to be mainly due to the average sugar pucker orientation of the furanose envelope or pseudorotational phase angle, which is defined by the atom which is displaced furthest from the plane of the other atoms either on the same/opposite side of the plane as C5 (endo/exo) ^{24,25}. Additionally, while there has been significant interest in developing nucleobases lacking hydrogen bonding and more reliant on shape complementarity for base-pairing complementarity ^{17,26,27}, there is growing evidence which suggests that a major factor in determining helix stability is due to the ability for aromatic nucleobases to base-stack or have quadrupole moment π - π interactions ²⁸⁻³¹. Presumably, these structural characteristics help produce the most important emergent property of XNA- nuclease resistance, by limiting the accessibility of the polymer backbone to degradative enzymes ³², yet maintaining the ability for many sequence complimentary XNAs to hybridize with natural DNA/RNA and form helices ^{33,24}.

From this research, applications of XNA have already begun emerging, mainly in therapeutics. The ability for XNAs to be increasingly resistant to endogenous human nucleases has been extensively studied and has laid the basis for XNA as effective Anti-Sense Oligo (ASO) reagents ^{34,35}. It seems the current ultimate measure of an XNA's stability is it's resistance to 3' exonuclease snake venom phosphodiesterase, while endonucleases ribonuclease H1 and nuclease S1 have also been used throughout literature ^{36,34,37-39}. Additionally, since many XNA modifications have been found to increase

duplex melting temperatures, a measure of helix stability, XNA have also recently found applications in enhancing gene-editing platforms, yielding higher target sequence specificity as well as lower off-target editing events ⁴⁰.

Benefits of Glyco-mimetics and Polysaccharide Hybrid Materials

Inherently, glyco-mimetics can be loosely defined as a class of synthetic polymers which seek to replicate the activity of natural polysaccharides, glycolipids, and glycosylaminoglycans. These natural polymers are instrumental in regulating biological systems through intercellular signaling in a structure known as the glycocalyx or pericellular matrix. Many integral proteins within the cellular membrane are capable of interacting with this glycocalyx covalently by direct glycosylation as glycoproteins or non-covalently by as lectin proteins ⁴¹. A growing body of evidence suggests that the majority of these non-covalent interactions are highly dependent on the binding of several glycan epitopes, which are not only specific to the configuration of sugars, but have enhanced avidity through multivalency or a precisely evolved length and unique spacing of epitopes.

This allows cells to communicate through chemical non-hormonal means for self-recognition, cell fate determination, and immune response ^{41,42}. Hence, glycol-mimetics have been a tool to develop various drugs treating ailments such as organ rejections following transplantation, paralysis and stem cell research, as well as vaccine research ⁴³⁻⁴⁶. Indeed, many viruses are known for their relative structural simplicity and innate high mutation rate; however, it has been discovered that displayed glycan epitopes of viruses are highly conserved, despite rapidly changing primary amino acid sequence, since these structures are highly specific for maintaining virulence ⁴⁷. Hence, research in

glycomimetics has led to the development of glycan-epitope specific affinity reagents, such as broadly neutralizing antibodies and aptamers. Consequently, by understanding how pathogens are capable of selectivity through glyco-science this has also allowed the development of cell-specific targeting technologies ⁴⁸⁻⁵⁰.

The basis of these phenomenon is based in what is known as the *cluster glycoside effect* ⁵¹. Several aspects constitute the cluster glycoside effect: it has long been known that carbohydrate/ligand-protein interactions are capable of being drastically enhanced by increasing saccharide lengths, allowing cell receptors to discriminate between mono- or multi-valent saccharide interactions due to the many possible modes available for multi-receptor binding, multi-ligand binding, or receptor spacing along the membrane ⁵²; this clearly suggests, and has been substantiated by experiment, that a successful glycol-mimetic reagent is capable of optimal binding with increasing similarity to the natural receptor epitope in terms of spacing, as well as a unique multivalent ligand ⁵³⁻⁵⁵; additionally, theoretical experiments suggesting that the physics necessary for binding and reception of carbohydrates is similar to the CH- π interactions well known between aromatic amino acids and proline ^{56,57}, similarly sharing electron density as in the π - π interactions in aromatic previously described in nucleobases, yet capable of doing so through hydrogen-bond acceptor/donor interactions ⁵⁸; the consequence of this understanding suggests that not only length or spacing but also the electrostatic interactions of each specific carbohydrate within a glycomimetic have unique electronic and steric components ⁵⁹, which synthetic experiments have substantiated as capable of influencing the total avidity of epitope recognition interactions, demonstrating that some heterologous glycomimetics have enhanced affinities over homopolymers ⁶⁰. Considering

these holistically, perhaps it is not surprising that another ‘fine-grain’ determinate of specific carbohydrate bindings are configurationally specific modifications, specifically sulfations or the *sulfation code*^{61,62}.

Current state-of-the science biological hybrid polymers are capable targeting specific organ tissues for gene therapy via glyco-conjugation, while preventing rapid renal clearance or face liver or kidney accumulation and the resulting toxicity with polyethylene glycol (PEG)-conjugation and optimized backbone chemistries⁶³⁻⁶⁹.

However, while there is a significant ability to augment the specificity and stability of such polymers with end-modifications, the overall efficacy and ease of synthesis of the core drug can drastically differ. Hence, in addition to main-chain glycol-mimetics which seek to emulate naturally occurring polysaccharides and glycosylaminoglycans, several side-chain chemistries have been explored to take advantage of the specificity of glycol-mimetics such as amino acid, XNA, dendrimers, and more traditional polymers^{70,44,71,72}.

A complete analysis comparing each of these chemistries in-depth is outside the scope of this overview; however it is worth noting some major accomplishments and differences in approaches. Particularly, while applications are currently limited in scope, this technology could generate a renaissance in drug design by limiting off-target drug effects and consequently side-effects, as glycol-mimetics are better understood.

Promise of Aptamer Science over Traditional Affinity Reagents

There have already been a number of insightful reviews over the years specifically detailing the benefits of aptamers over antibodies and the emerging benefits of dendrimers and cationic lipids^{68,73}. In a general sense, all major monoclonal affinity reagents are side-chain polymers which take advantage of the variability of their side-

chains in order to non-covalently interact with the expected target. For meaningful therapeutic applications, multivalency, monodispersity, and cell permeability are also highly desired; hence water-soluble sequenced polymers are the standard from a practical and analytical standpoint. Antibodies are the human immune systems' answer to the need for affinity reagents in order to fight infectious disease and are well suited to these applications; however, the major limitation to the medical applications of antibodies as ideal therapeutic agents is the synthetic availability. While selection for antibodies is reasonably high-throughput with display techniques, sequence controlled antibodies must be generated recombinantly, which is complicated by the necessity to use specialized cell-lines expressing non-natural amino acids if the activity of a specific antibody is discovered in vivo and requires post-translational modifications, or require solid-phase synthesis^{74,75}. Significant advances have been made in solid-phase peptide synthesis (SPPS), allowing for the reduction of bulk reagents and time from liters and hours into milliliters and minutes⁷⁶. Additionally, variants on Dawson's native chemical ligation (NCL) process for generating synthetic polypeptides in pieces has allowed SPPS to generate small proteins larger than 110 residues or ~12.1 kDa roughly double the traditional limit of about 60 residues or ~ 6.6 kDa⁷⁷⁻⁸⁰, which can additionally be coupled with photo-flow-reactor systems for automation⁸¹; however, while nanobodies (Nb-mAbs) also known as single-domain antibodies (sdAb), and bispecific antibodies (bs-mAbs) have recently been developed with the aims of significant reducing the complexity of traditional antibodies. Nanobodies require only single variable domains (V_{HH}), removing the need for immunogenic light variable (V_L) domains in the antigen-binding fragment (Fab) region as single-chain variable fragment (scFv), and even

allowing removal of the crystallizable fragment (Fc) region through the linkage of two or more Nb-mAbs together. Overall, these advancements are capable of generating Nb-mAbs at the scales of 15 kDa, increase cell-permeability and structural stability, decrease immune responses and aggregation effects, significant “camelization”/residue-specific-modification, replacing hydrophobic residues with hydrophilic ones, is necessary to allow the water solubility of these constructs ⁸². Ultimately, the generation of polypeptides via this method requires substantial project specific optimizations, involving resin selection and protective group selection; these limit the synthetic accessibility, and thus ease of mutant analysis for understanding the mechanisms of action, of evolved antibodies ⁸³. Similarly, many dendrimers are in part based on SPPS generated polypeptides but assembled using coupling chemistries ⁸⁴⁻⁸⁶.

Nucleic acid bases affinity reagents are capable of overcoming many of these issues, generally phosphoramidite chemistries regularly allow for coupling efficiencies exceeding 99 % and generation of up to 100 bases ^{87,88}, as previously described are amiable specific chemical modifications of drugs in order to increase cell permeability, and are capable of selection of targets with sub 20 nM Kd values ⁸⁹ and epitopes with multivalences approaching or equivalent to natural glycan ligands ^{55,60}. It is thus, unsurprising that aptamer-hybrid materials are novelly being considered alongside -dendrimer and -peptide modifications ^{73,85,90}.

CHAPTER TWO

Literature Review

Historical Development of XNA Technology

Significant advances in XNA research came in the late 70s; particularly in therapeutic applications. While the study of glycomimetics has experienced a resurgence in research due to the influx of polymer chemists to this field of study^{52,72}. Many current XNA chemistries were developed to suit similar requirements of genetic therapy. The XNA chemistry of 2'-deoxy-2'-Fluoro-Arabinose Nucleic Acid (FANA) has previously been utilized as so-called 3rd generation Anti-Sense Oligo (ASO) RNA-based therapies³⁴; similar chemistries have found successful FDA drug approval, such as the 2nd generation 2'-*O*-methoxyethyl (2'MOE) *Nusinersen*, and 2-*O*-Methyl (2'OMe) *Patisiran*, and 3rd generation phosphoroamidate morpholino oligomer (PMO) *Eteplirsen*.

XNAs open up a new lens in understanding the limitations of traditional molecular biology and how scientists should go about understanding the intersections of the various disciplines of: systems chemistry, exobiology, synthetic biology, and research into the origins of life, in the quest of understanding the molecular world; Xenobiology opens the door to the development of new biotechnologies, the ability to synthesize novel “never born proteins,” artificially expand genetic information systems, and provides a deep understanding of the origin of all Earthly life, in addition to allowing safe and ethical scientific progress into ultimately synthesis of wholly man-made organisms. With new generations of replicative multiplexed polymerase chain reaction (PCR) and single

nucleotide polymorphism (SNP) sensing available through XNA technology, the largest hurdle science has yet to overcome is understanding the advantages and limitations of zoo of different XNAs both historically and currently being researched: from humble Homo-DNA, isoGs:isoC, dZ:dP, y:x, z:s, xDNA:yDNA, ImO^N:NaN^O & ImN^O:NaO^N, to more exotic Q:F, P:H, Ds:Pa, PICS:PICS leading to NaM/MMO2:5SICS, N^d:M^d, HNA, tc-DNA, CeNA, 2'OMe/SeMe/F/N₃, and the almost absurd class of ANA, FANA, TNA, GNA, LNA/BNA, PNA, MNA, PMO, and αS/B XNAs. Regardless, it seems imperative to no longer think of genetic information being stored in the archetypical A, T, C, G, U breakthrough scheme that has formed the foundation of modern molecular, especially as XNA technology advances to no longer limit itself to single changes in nucleic acids either in bases, sugar rings, or polymer linkages. Via the advents compartmentalized self-tagging-statistical correlation analysis (CST-SCA), the waves of research are already being made in attempting to provide the proper cellular machinery in order to allow XNAs to not only replicate and be synthesized for use in genetic assays but also to reversibly transcribe *in vivo*, allowing genetic information to be stored and be communicated across both DNA and XNA organismal platforms. While full XNA organisms may seem to be a reality only the distant future, the pace of genetic technology may be compared Moore's Law and the development of computer electronics, and thus the question of when is ultimately an inevitability which may even be tackled with this generation of scientists. This highlights many of these historical developments and explicitly describes current technology in the development of promiscuous, or "unfussy," polymerases which allow not only the synthesis of a variety XNAs but their reverse transcription. This is the beginning of an era where proteins will not be modified to suit

purposes but designed *de novo* and synthesized in orthogonal organisms which safely live alongside Earthly-life- anecdotal living aliens from our imaginations.

Historical Overview of Glyco-mimetics and Polysaccharide Hybrid Materials

Polysaccharides have traditionally been in the realm of organic chemists; however as research has uncovered the increasing role of not just ligand length but also spacing in defining functional multivalency. For an early review of advances in synthetic polymerization the review by Grubbs and Grubbs is highly recommended, detailing the development of the various methods discovered for controlling and initiating polymerization and its origins in from the early 19th century ⁹¹. More recently, the necessity of sulfation patterns to regulate axon growth, and not necessarily the length of polysaccharides was elucidated by a retro-synthetic approach, which accomplished the synthesis of a tetrasaccharide from the coupling of two disaccharides, each originating from two highly protected pyrano-sugars ⁶¹. The length sugar synthesis of this report clearly demonstrated that the activity of chondroitin sulfate glycosylaminoglycans could be replicated in part through synthetically available specific sulfated polysacchirdes. This research was further expanded to by generating similarly sulfated side-chain mimetic glycopolymer molecules of varying lengths via a ring opening polymerization strategy and controlled by the catalyst concentration ⁶². This marked a transition of the study of glycomimetics into the world of polymer chemistry, as the report demonstrated that sulfated disaccharide glycomimetics had an increasing level of axon inhibition with increase length, yet overall had a lower activity compared to a sulfated tetrasaccharide polymer, synthetically derived from two disaccharide monomers, which additionally had activity resembling the natural isolate despite the tetrasaccharide having only a moderate

length in comparison to the disaccharide. The functional necessity of sulfation/functionalized glycan-ligand multivalency was figured to not be enough to define glycosylaminoglycan activity, similar length disaccharide polymers with varying numbers and location of sulfations were synthesized via previous methods in order and found only to be active when a specific combination or *sulfation code* was present ⁴³. Additionally, this synthetic approach was able to produce a clear application by allowing the selection of a specific monoclonal antibody, which novelly was able to selectively target a unique class of sulfated natural glycosylaminoglycan and able to additively enhance axon regrowth in combination with axon stimulating drugs.

While the mechanisms of the natural structure-function relationship of glycosylaminoglycans were being teased out by organically synthesized or variable length polymer glycomimetics, other research was focused on generating glycomimetics in a sequence-specific and monodisperse methodology. This was accomplished by the generation of glycomimetics from the Huisgen copper-catalyzed alkyne-azide cycloaddition (CuAAC) click amiable triple bond diethylenetriamine succinic acid (TDS) ‘side-chain’ monomers in combination with a ethylenedioxy succinic acid (EDS) ‘spacer’ monomers assembled via fluorenylmethyloxycarbonyl (Fmoc) protecting group (PG) solid-phase synthesis methods ⁵⁴. The products of this synthesis were capable of demonstrating that the spacing of the glycan residues are significant with trivalent side-chain glycopolymer have ~250 times the activity compared to that of the natural α -methyl-D-mannose ligand and 8 times the activity as a monovalent polymer of similar length, producing a polymer displaying more ligands per total polymer length (7 nm) and better approximating the natural lectin binding site distance (6.5 nm). Additionally, this

method was superior to previous synthesis as minimal sugar functionalization was necessary to generate acetyl protected 2-azidoethyl-O- α -D-mannopyranoside, and the resulting polymer was capable of being functionalized prior to resin cleavage, and generated with purities ≥ 90 % requiring minimal purification; additionally the use of Fmoc methods allowed a mild 25 % piperidine in dimethylformamide (DMF) automated deprotection method ⁸⁰, easy tracking of this deprotection and final resin cleavage procedure, as well as allow a asymmetric synthesis using trifluoroacetic acid (TFA) as a protecting group starting with a diamine molecule. Variations on this methodology utilizing Fmoc, tert-Butoxycarbonyl (Boc), and Allyloxycarbonyl (Alloc) amine protection for orthogonal deprotection and functionalization also allow the generation of branching monomers ⁹². Furthermore, by functionalizing TDS monomers with sugars immediately following coupling, polymers displaying different glycan species are capable of being generated, in addition to having precise polymer length and multivalency control ⁶⁰. Similar work was attempted with the thiol-ene click (TEC) chemistry and olefin diethylenetriamine succinic acid (DDS) ‘side-chain’ monomers via photo-flow-reactor setup, yet these investigations only seem to have attempted to generate hetero-glycomimetics via already glycan functionalized monomers; however, it has been previously demonstrated that the CuAAC, TEC, and several click chemistries are orthogonal to one another and could be potentially used to functionalize different ‘side-chain’ DS-like monomers following solid-phase synthesis of the backbone ^{93,94}. This generation of hetero-glycomimetics demonstrated that high affinity lectin binding is possible to be enhanced by the presence of different species of glycans up to 6 times for binding ligands per polymer at surface interfaces, but in solution homo-glycopolymers

are capable of generating aggregate complex due to multi-target interactions, highlighting that inclusion of non-binding glycans of varying steric diameter are capable of controlling in solution binding modes, especially on flexible backbones.

Recently, these techniques have been combined in order to synthesis multiblock gycopolymers from a TEC step growth polymerization of DDS-TDS-EDS-TDS-EDS-TDS-DDS blocks generated from solid-phase synthesis with Cys-EDS₅-Cys ‘spacing’ blocks, mediated by 2,2-dimethoxy-2-phenylacetophenone (DMPA) radical initiator and tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) disulfide reducing agent ⁹⁵. While this method incorporated polymer chemistry techniques leading to the synthesis of polydispersed polymer with a mass dispersity ($\bar{M}_w/\bar{M}_n, D_M$) of 1.76 – 1.87, typical of a step polymerization ⁹⁶, this allows the efficient synthesis of higher MW glycomimetics, similar to the NCL methods of SPPS; this demonstrated that molecules with higher multivalency tend to have higher affinity, yet the efficiency per binding ligand efficiency tends to maximize at a specific multivalency, independent of the MW of the molecule and seemingly the ligand density and not ligand spacing. The emphasis on the total MW of glycomimetics is significant for applications of polymer chemistry to glycomimetics, as previous studies in the context of maintaining peptide complex stability suggest that while the attachment of polyethylene glycol (PEG) and certain glycoforms of non-natural glycosylations were beneficial in decreasing biopolymer immunogenicity in vitro and in vivo mice, the increased MW of similar glycosylation modifications was found to decrease the mean absorption in dogs, without effects on the biological activity of agent ⁹⁷. This demonstrates that glycosylation modification of biopolymers holds promise in improving the stability of complex biologically medicines without eliciting immune

response and perhaps even decreasing it, here specifically insulin; but, care needs to be taken as to the method of glycosylation, as there seems to be an interplay between the ability for glycosylated/increased MW drugs to adsorb through tissues and with decreasing immunogenicity. In response to these findings, studies utilizing a single natural serine/threonine glycosylation of an insulin biopolymer at non-structured regions were found to have a similar biological stability to in vitro proteases, 4.7 times higher chemical/complex stability, in addition to an increase in the activity of the drug in vitro adipocytes; while triglycosylated glycopolymers were found to have increased biological stability, 2.7 times higher chemical/complex stability, with no change in the activity of the drug in vitro⁹⁸. This seems to suggest that the adage 'less is more' holds true, at least with the applications of glycosylations to improve biopolymer stability, without decreasing biological activity.

Current methods to discover natural glycosylations are mostly reliant on glycan arrays⁹⁹ and mass spectrometry, each of which present unique challenges¹⁰⁰. With carbohydrate arrays, discovery of particularly lectin protein specificity to a specific class of carbohydrates is reasonably easy but is ultimately limited by the synthetic availability of carbohydrates, either from natural sources or current synthetic methods. This in particular can be troubling for discovering associations with highly branched glycans, as synthetic methods become limited due to the significant combinatorics involved in carbohydrate polymers. Respectively, while lectin arrays are becoming more common commercially, these suffer from the same combinatorics issues as glycan arrays. However, while mass spectrometry does not limit the possible scope of analysis, this approach is highly limited by the scope of current glycan databases, additionally this

analysis is significantly easier with *a priori* knowledge of the glycosylated agent, particularly proteins. Additionally, considering that some glycans are not available in significant quantities for traditional spectroscopic analysis, this is the current standard.

Recent advances in cryogenic electron microscopy (cryo-EM) allowed the specialized analysis of peptide polymers particles larger than 50 kDa and smaller than 500 nm to a resolution of 3.8 Å within at hydrated or membrane-bound conditions, possibly down to a theoretical limit of 38 kDa¹⁰¹; it is worth noting the key advantage that cryo-EM provides is the ability to provide reasonably high-resolution absolute structure determination for molecules traditionally difficult to crystalize due to their size, such as many proteins, otherwise X-ray diffraction/crystallography (XRD) techniques are standard use. Alternatives to XRD have been developed in order to overcome the destructive nature of X-rays, specifically using electron diffraction¹⁰², significantly requiring fewer and smaller crystals; iteration of this technique in combination with cryo-EM has developed yet another technique which is perhaps ideal for the analysis of polymers smaller than limits afforded from cryo-EM- micro Electron Diffraction (MicroED)¹⁰³. MicroED is similarly accessible as cry-EM, also utilizing transmission electron microscopy, the workflow is also quite similar to that of subtomogram averaging (STA) but for the analysis of single-‘nanocrystals,’ and is thus inherently not capable of single-particle analysis (SPA) of molecules in native-like conditions¹⁰⁴. Continuous rotation methods with microED have proven to be capable of high resolution imaging of enzymes (187 kDa) to a resolution of 1.8 Å¹⁰⁵, 9 residue peptides (816 Da) to a resolution of 1.85 Å¹⁰⁶ and 6 residue peptides (905 Da) to a resolution of 1.1 Å with the

addition of divalent heavy metal dopants ¹⁰⁷, and milligram scale heterogeneous powdered mixtures of cyclic small molecules (≥ 133 Da) to a resolution of 1.0 Å ¹⁰⁸.

CHAPTER THREE

Materials and Methods

Experimental Design

This study was focused on generating a FANA sequence controlled polymer backbone from DNA library templates using minimally modified DNA polymerase transcription. These transcripts are designed to additionally incorporate CuAAC clickable alkynyl modified uracil bases, replacing all corresponding thymine bases complementary to the library adenosines. The incorporation of this modified-DNA base has been previously found to be well tolerated in some exonuclease deficient (exo-) DNA polymerases; hence, the total synthesis of this FANA base was begun. The incorporation of this base into the FANA backbone was designed to allow easy appending of ethyl azide saccharides, and subsequently selection for sequence-evolvable and nuclease resistant FANA-lectin aptamers.

Table 1 – Oligonucleotides used within this study, and their respective sequence. All DNA oligonucleotides were purchased from Integrated DNA Technologies. Primers were ordered with standard desalting. PCR templates and modified primers were ordered with PAGE purification.

Use	ID	Sequence (5'-3')
Tgo Plasmid Sequencing	Terminator Seq -1	CTT CCT TAG CTC CTG AAA ATC TC
	Tgo-T5 +1	TCG TCT TCA CCT CGA GAA ATC ATA
	Terminator Seq+1	CAA TTG TGA GCG GAT AAC AAT TTC
	Tgo Seq +1	GTT ATT ACC TGG AAG AAT ATC
	Tgo Seq +2	GAG AGG GAG CTG GCA AGA AG
	Tgo Seq +3	CTT CTA CAA GCG CGG CTT C
FANA Transcription	DNA library: T-Lib	ATC CAG AGT GAC GCA GCA N ₄₀ (N=A:T:C:G=25:25:25:25)TG GAC ACG GTG GCT TAG T
	Heavy DNA lib: T-Lib XL	ATC CAG AGT GAC GCA GCA N ₄₀ (N=A:T:C:G=25:25:25:25)TG GAC ACG GTG GCT TAG T (AAC) ₈
	QC DNA Rev: T-CY5-RP	/5Cy5/ACT AAG CCA CCG TGT CCA
	QC DNA For: T-CY3-FP	/5Cy3/ATC CAG AGT GAC GCA GCA
	Forward 1A: T1-AM6-FP	/5AmMC6//iSp18/ATC CAG AGT GAC GCA GCA
Click-Particle Display	Forward 2: T1-TTB-RC	/5Hexynyl/TGG ACA CGG TGG CTT AGT
	Forward 1 QC: T1-CY5-FC	/5Cy5/TGC TGC GTC ACT CTG GAT
	Forward: T1-FP	ATC CAG AGT GAC GCA GCA
	Reverse: T-RP	ACT AAG CCA CCG TGT CCA
	QC DNA Reverse: T1-Cy5-RP	/5Cy5/ACT AAG CCA CCG TGT CCA
	DNA library: T1-Lib	ATC CAG AGT GAC GCA GCA CGG AAC GTC TTT GTA ACT TGA AAT ACC GTG GTA GGT TGG CTA GGT TGG ACA CGG TGG CTT AGT

DNA transcription into 2'-Deoxy-2'-Fluoro-Arabinose Nucleic Acids (FANA)

Working initially from protocols developed by a collaborating laboratory^{109,110}, a primer-extension based method was implemented for ‘translation’ of DNA 76-mer libraries, consisting of a random region of 40 nt flanked by 18 nt two primer binding sequences, into FANA via minimally-mutated polymerases. This work involved screening commercially-available polymerases and exonuclease deficient (exo-) *Thermococcus gorgonarius* (Tgo) recombinant polymerase¹¹¹- a B-family polymerase best known for previous research introducing translesion synthesis^{112,113} RNA synthesis activity¹¹⁴, and overall similarity to *Pyrococcus furiosus* (Pfu) polymerase¹¹⁵⁻¹¹⁷.

Recombinant Tgo plasmid verification

Purified pGDR11-Tgo plasmid (GenBank KP682507.1) was obtained as a generous gift from the Chaput laboratory in UC Irvine (carrying the D141A and E143A but not D215A exo- mutations¹¹⁸, responsible for active-site ligand anchoring of two divalent ions¹¹⁹⁻¹²¹) and transformed into electrocompetent TOP10 *E. coli* cells in SOC broth. These cells were incubated overnight at 37 °C and transferred to 100 µg·mL⁻¹ ampicillin LB-Lennox plates. Colony TempliPhi amplification was used to generate transcripts for sequence verification of the transforming plasmid. A 25 % glycerol stock and 5 mL liquid culture of verified colony was generated. From the overnight liquid culture a SpinSmart miniprep columns with Qiagen buffers were used to reisolate the pGDR11-Tgo plasmid. These plasmid were confirmed to be identical by digestion with NEB NheI-HF and AvrII or SacI-HF restriction enzymes and ran on a 10 % TAE agarose gel in 1X TBE buffer at 200 V for 40 minutes with 0.25 µL·mL⁻¹ ethidium bromide-1X TBE staining for 5 minutes (Appendix A). DNA libraries and primers were purchased from Integrated DNA Technologies (Coralville, IA).

Recombinant Tgo expression and purification

This verified plasmid was transformed into electrocompetent XL1-blue *E. coli* cells. Tgo polymerase expression was induced and was purified from cells using the aforementioned previous protocol¹⁰⁹. Briefly, a 100 mL starter culture inoculated a 1 L solution of LB-ampicillin to OD₆₀₀ = 0.6 at 37 °C with shaking at 225 rpm. Expression was induced by the addition of 500 µL 1M isopropyl β-D-1-thiogalactopyranoside (IPTG) to final concentration of 0.5 mM followed by incubation at 37 °C with shaking at 225 rpm for four hours. Cells were harvested with centrifugation at 7000 × g for 10 minutes at 4 °C, resuspended in 40 mL buffer (10 mM Tris, 500 mM NaCl, 10 %

glycerol, pH=8.00) for 5 cycles of sonication in a 4 °C cold room, and immediately placed into a 80 °C water bath for 60 minutes. The cell lysate was then cooled for 30 minutes on ice to allow non-heat stable proteins to aggregate, followed by pelleting by centrifugation at 40,000 × g for 30 minutes at 4 °C. Nucleic acids were removed from the clarified cell lysate by adding 10 % (v/v) polyethyleneimine to a final concentration of 0.5%, incubating for 15 minutes on ice, followed by pelleting by centrifugation at 40,000 × g for 20 minutes at 4 °C. The polymerases were precipitated from this supernatant with ammonium sulfate powder to a final concentration of 60% (w/v), incubation on ice for 30 minutes, and centrifugation at 40,000 × g for 30 minutes at 4 °C. This pellet was resuspended in a low salt buffer (10 mM Tris, 50 mM NaCl, 10 % glycerol, pH=8.00) and clarified by a final centrifugation at 40,000 × g for 8 minutes at 4 °C. Polymerases were purified in 10 mL aliquots via fast protein liquid chromatography (FPLC) through a GE Healthcare (Chicago, IL) HiTrap Heparin column with a 10x0.5 mL 50, 375, 500, 1000 mM NaCl step elution program. The fractions were monitored by UV spectroscopy at Abs₂₈₀ on a Thermo Scientific (Waltham, MA) NanoDrop One. The 500 mL fractions were reinjected into the column and repurified. Purified polymerase was concentrated using Millipore (Burlington, MA) Microcon YM-10 membrane 10k MWCO centrifugal microconcentrator and resuspended in a moderate salt, 50 % glycerol buffer (10 mM Tris, 250 mM NaCl, 50 % glycerol, pH=8.00). Alternatively, dialysis with Thermo Scientific Slide-A-Lyzer 20k MWCO MINI Dialysis Cups followed by microconcentrators has been used to resuspend the polymerase without glycerol; however, it was found that the activity of the polymerase seemingly increased with storage in 50 % glycerol (Appendix D).

Recombinant Tgo purification QC

Aliquots of proteins throughout the purification process were suspended in 1X Cold Harbor Spring sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer, denatured by heating for 10 min at 90 °C on the heating block, and loaded into Invitrogen (Waltham, MA) NuPage 4-12% Bis-Tris gels with 1X MOPS SDS running buffer. Gels were run for 50 minutes at 200 V alongside Thermo Scientific PageRuler Plus MW ladder and stained with Coomassie R-250 dye for visualization (Figure 1). The final polymerase concentration was determined via the amino acid sequence of Tgo (GenBank ALL53335.1) and the ExpASy ProtParam W,Y,C method for extinction coefficient estimation in water^{122,123}.

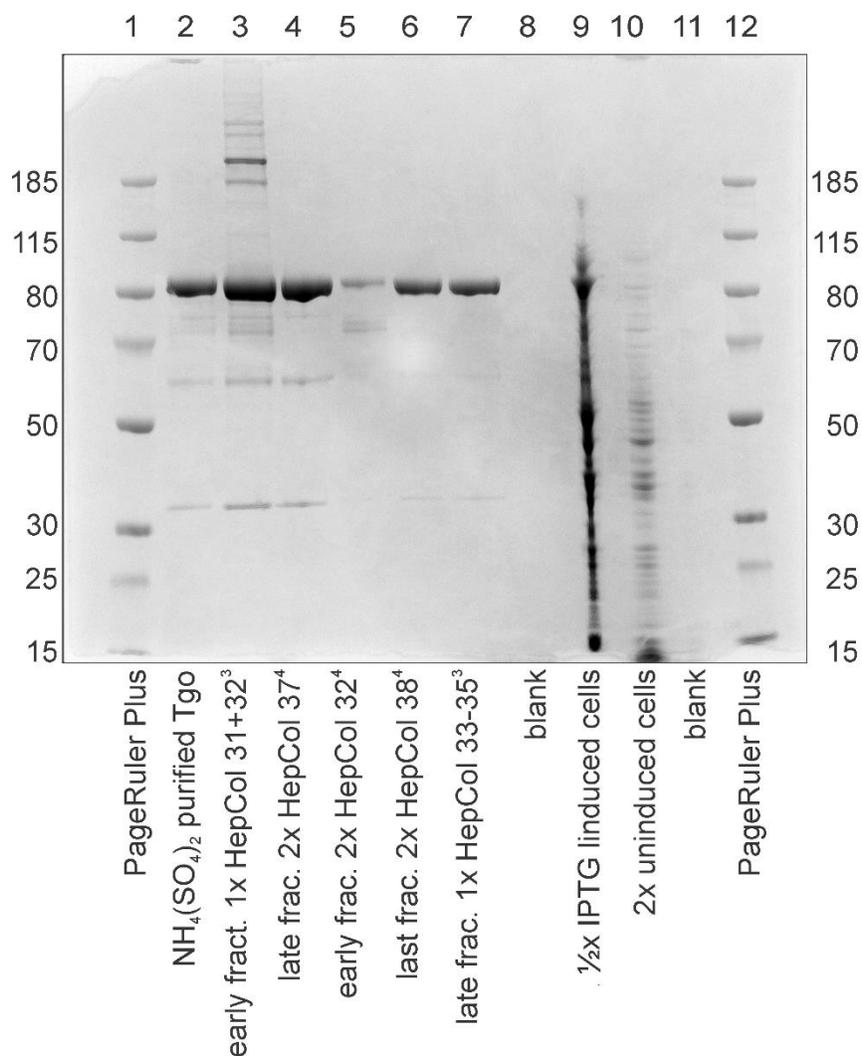


Figure 1. – Purification of Tgo polymerase expressed in XL1-blue cells. PageRuler Plus loaded on each side 2.5 μ L each, 10 μ L of each purification sample, 5 μ L of IPTG induced cells after 4 hrs at 37°C with 225 rpm shaking, and 20 μ L of cells prior to cell lysis.

Polymerase screen for FANA Transcription Activity

In order to identify the ideal DNA polymerase for transcribing a FANA oligonucleotide from a DNA template, New England Biolab's (Ipswich, MA) *Thermus aquaticus* (Taq), Deep Vent (exo-), ThermoPol (exo-), and Milipore' *Thermococcus kodakaraensis* (KOD) XL polymerase were screened alongside the previously described purified recombinant Tgo polymerase. Each polymerase was selected in order provide a different perspective the characteristics necessary for a B family polymerase to accept FANA monomers in ThermoPol buffer (20 mM Tris, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM KCl,

2.0 mM MgCl₂, 0.1 % Triton X-100, pH = 8.8),: Tgo (exo-), as previously mentioned lacking 3'→5' proof reading, and is reported as having x26 Taq fidelity as a wildtype polymerase¹²⁴; Taq polymerase, an A family polymerase and the standard for DNA transcription, capable of 5'→3' proof reading and nominal strand displacement, at a concentration of 5.0 U·μL⁻¹; Deep Vent (exo-), a B family polymerase with similar exo-mutations^{125,126}, thus with limited 3'→5' proof reading but moderate strand displacement, and is reported as having x5 Taq fidelity, at a concentration of 2.0 U·μL⁻¹; Terminator (exo-), a B family polymerase is a (D141A, E143A, & A485L are known gain-of-function (gof) mutations to enhance incorporation of unnatural substrates) mutant 9°N polymerase^{127,114}, thus lacking 3'→5' proof reading activity but nominal strand displacement (NEB), at a 2.0 U·μL⁻¹ concentration; KOD XL, also known as KOD-Dash is a 40:1 mixture of N210D (exo-) and wildtype KOD B family¹²⁸, thus with limited 3'→5' proof reading, has been previously reported to incorporate a wide variety of unnatural modified uracil DNA nucleobases^{129,130} and was specifically chosen as it has been used to generate monosaccharide-modified uracil nucleobase appended⁸⁹, at a 2.0 U·μL⁻¹ concentration in a proprietary reaction buffer.

The ideal polymerase was identified as having the capability of successfully generating full length FANA products, as well as consuming all of the supplied Cy5 tagged DNA template reverse primer. The fidelity was not examined in this screen, as this would be reserved for future studies, once reverse-transcription DNA→FANA→DNA capabilities were developed, including modified nucleobases. Additionally, this method could potentially be used to study the efficiency kinetics of these systems¹³¹.

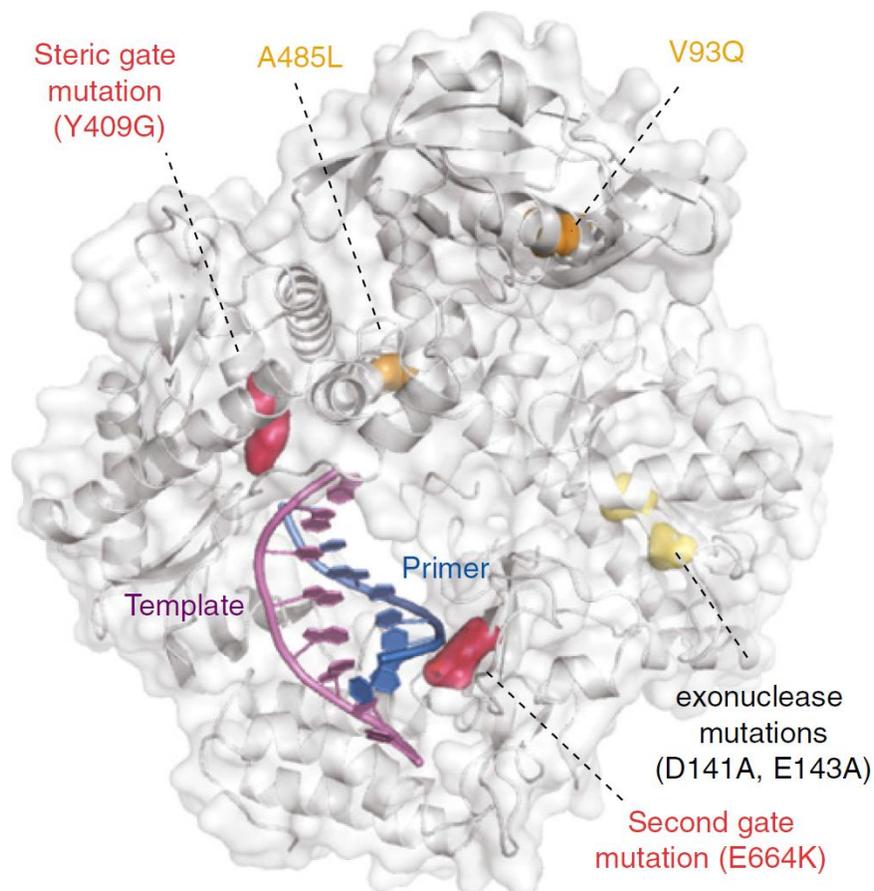


Figure 2. – Tgo-like polymerase displaying known gof RNA transcription mutations Pfu polymerase shown interfacing with DNA template via RNA primer, emphasizing key mutations: V93Q responsible for uracil stalling, reducing RNA processivity; A485L the ‘Terminator’ mutation also known to increase substrate promiscuity; Y409G and E664K steric gate mutations discovered alongside TGK, affecting the thumb subdomain, as previously demonstrated to be key in maintaining substrate discrimination^{27,132–135}; and aforementioned D141 & E143A (in black text), 3’→5’ exonuclease proof reading mutations, with similar D215A is not shown¹¹⁸. Reproduced without permission from original authors¹¹⁴.

Polymerase activity was assessed by primer (XNA) extension (XPE) reactions in 10 μ L reaction volumes containing 1X reaction buffer (ThermoPol or otherwise), 1 μ M T-Cy5-RP primer, 1 μ M T-Lib DNA library, 100 μ M of each FANA A/T/G/U xNTPs purchased from Metkinen Chemistry (Kuusisto, FIN), and 1 μ L polymerase (crude Tgo, purified Tgo, Taq, Deep Vent, Terminator, or KOD XL)¹⁰⁹. Primers and templates were denatured by heating at 90 °C on a thermocycler for 5 minutes for proper annealing on ice for 10 minutes, prior to the addition of polymerase. Upon addition of the polymerase,

the samples were incubated for 3 hours at 55 °C. Reactions were quenched with the addition of 100 µL (10 equivalents v/v) of 99 % dionized (DI) formamide and 25 mM EDTA (ethylenediaminetetraacetic acid) to a final volume concentration of 2.27 mM. DI formamide was stored at -20 °C, decanted, and thawed for aliquots in order to maintain consistent results. Due to the increased DNA/FANA duplex stability, 15 µL samples were denatured at 95 °C on a heating block for 10 minutes, prior to loading onto denaturing urea-PAGE gels. After denaturation, samples were loaded onto Invitrogen Novex 10% TBE-Urea gel, preheated at 8 W for at least 30 minutes. Gels were run for 60 minutes at 180 V alongside 1 µL Thermo Scientific GeneRuler Ultra Low Range DNA Ladder with 5 µL 2X bromophenol blue + xylene cyanol loading dye and a primer-template complex, imaged with a Cy5 sensitive channel, and stained with 0.25 µL/mL ethidium bromide-1X TBE buffer for composite visualization.

Polymerase Dilution Series

Tgo, Therminator, and Deep Vent polymerases were diluted in order analyzing the concentration sensitivity of these polymerases in producing full DNA-FANA extension, as per previously reported methods¹¹⁰. The procedure was conducted in reaction conditions similar to polymerase screen XPE reactions. Dilutions were made of Tgo in moderate salt buffer (10 mM Tris, 500 mM NaCl, 10 % glycerol, pH=8.00) and Therminator and Deep Vent were diluted in the respective NEB buffers E & D. Serial 1:1 dilutions were made to a final dilution of 1:512, providing a measure of the optimal polymerase concentration for DNA-FANA extension, as just higher than the concentration producing partial extension ‘laddering;’ additionally, these dilutions demonstrated that the polymerases did not contain any significant nuclease

contaminations by comparing the intensity of the starting material template bands to the non-polymerase containing negative control.

The results of this study found differences between the activity of Deep Vent polymerase in the polymerase screen and the dilution series. This suggested that perhaps the polymerase storage conditions of polymerases were consequential to the final activity, more so than % 50 glycerol prospectively allowing long term cold storage. A batch of Tgo polymerase known to have activity at a 1:4 dilution and a batch prepared with dialysis and microconcentrators known to have activity at a 1:8 dilution, were used to test the ability of NEB polymerase storage buffer compositions with variable KCl, Tween 20, IGEPAL CA-630, and Triton X-100 content, in addition to a constant 10 mM Tris, 1 mM DTT (dithiothreitol), 0.1 mM EDTA, and 50% glycerol, while our Tgo polymerase is currently storage in 500 mM NaCl, 10 mM Tris, & 10 % glycerol buffer. Buffer exchange was performed using a similar Microcon YM-10 membrane 10 MWCO centrifugal microconcentrator, instead replacing the previous volume buffer with a buffer augmented with a single additive of interest.

Tgo Timecourse Optimization

Tgo diluted 1:4 in a 500 mM NaCl, 10 mM Tris, & 50 % glycerol buffer was henceforth used within each procedure involving purified recombinant Tgo polymerase. FANA have 18 Da difference from DNA per monomer, allowing PAGE separation of FANA from DNA; however, in order to insure high quality FANA separation from DNA, a 'heavy' DNA library T-Lib XL was used with an extended 3' region, consisting of 6xAAC repeats downstream of the primer-binding site (Table 1). This sequence was chosen to increase synthetic yield, as purines tend to couple more efficiently than pyrimidines on solid-phase. Additionally, this experiment was found to be most

successful run at 3X volume conditions, with 3 μL sample aliquots taken at time points in 30 μL DI formamide with 25 mM EDTA, heated 10 minutes 95 $^{\circ}\text{C}$ and consequently cooled on ice for 10 minutes prior to loading on a 1:30 hr 8 W preheated TBE-Urea PAGE gel.

FANA Library Isolation Using RNA PAGE Methods

In order to generate an adequate amount of FANA library for testing reverse transcription conditions as well as anticipating purification losses, 100X scale XPE reactions were setup following a protocol similar to the previous XNA XPE reactions. Care was taken to split the master mix into PCR tubes with no more than $\frac{1}{2}$ the maximum operating thermocycler well volume, in order to maintain optimal heat transfer. However, extensions were conducted for 120 minutes, in order to maximize the generation of full length product but minimize over-extension by products, and using the ‘heavy’ T-Lib XL library as the DNA template. T-Lib XL was critical to this procedure as the additional 5.4 kDa of mass of the DNA library is approximately 3.9 kDa more massive than the produced FANA product; hence, the FANA product runs past the DNA library, minimizing the amount of DNA contamination possible, as opposed to the opposite case. The reactions were stopped by pooling into a 1 mL LoBind Eppendorf microcentrifuge tube (Hamburg, Germany) and 52.63 μL ($\frac{1}{19}$ equivalents v/v) 0.5 M EDTA was added to a final concentration of 25 mM to stop the reactions. A hole was punched in the cap using a 20G needle, frozen in liquid nitrogen for at least 3 minutes, and lyophilized for at least 6 hours. Afterwards, the sample was a pale blue liquid of approximately 50 μL , the volume of the polymerase glycerol. This sample was resuspended in 1 mL DI formamide. Samples were denatured for 20 minutes at 95 $^{\circ}\text{C}$ on a heating block, followed by 10 minutes cooling on ice.

The most efficient purification of the FANA products was investigated in a variety of different gel conditions. It was found that loading 25 μ L of denatured sample into commercial Invitrogen Novex 10% TBE-Urea gels produced reasonable separation between the FANA XPE and T-Lib XL library; however, even though preheating for 1 hr over 30 minutes was found to produce better band migration, this did not have a significant effect on separation (data not shown). These early separation attempts allowed isolation of FANA XPE product material or, as was known internally, X-Lib, although it did contain DNA contamination (Appendix E). This X-Lib was used to investigate the potential separation of 100X XPE reaction products and post-purification purity with analytical gel techniques- diluted Cy5-RP, T-Lib, T-Lib XL, purified X-Lib, and XPE reactions to concentrations similar to those of previous experiments, 1.36 pmol nucleic acid 10:1 in DI formamide (DIF). Self-cast gels were cast using Bio-Rad (Hercules, CA) Mini-Protean Tetra Cell 1.5 mm combs. Discontinuous 10 % 19:1/29:1 stacking/resolving 7.5 M urea TBE and continuous 10 % 29:1 7.5 M urea TBE PAGE gels were found to have minimal differences; however, were cast without a gel foot, hence there was a tendency for the control primer to run off the gel (Appendix F). The final preparative gel conditions allowed loading 500 μ L of DI formamide directly into an extended 680 μ L capacity well on a 12.5 % 29:1 7.5 M urea TBE PAGE gel, preheated for 15 minutes at 16 W and ran at 16 W for 30 minutes (Figure 3). These conditions allowed UV shadowing of the DNA template library as well as two lower FANA XPE product bands that could even be seen by the naked eye under normal library lighting. The purity of this extraction was monitored via a continuous 10 % 29:1 7.5 M urea TBE PAGE gel.

Products were extracted from gels through cutting bands directly from the gel and processing by placing one half of the gel slices each into a 500 μ L microcentrifuge tube with a hole punched via a 20G needle, for centrifugation at 21,000X g for 5 minutes into 1.5 mL LoBind tubes, followed by either a crush-and-soak or electroelution methods- soaking the crushed gel overnight in 1 mL DPBS overnight at 37 °C with shaking at 225 rpm; or overnight electroelution in dialysis tubing at 100 V, reversing the polarity for 2 minutes, and recovering the solution with rinsing in ddH₂O. The final slurry was loaded into two 800 μ L Thermo Scientific 30 μ m pore Pierce centrifuge columns or Corning (Corning, NY) Costar Spin-X 0.45 μ m pore centrifuge tubes and concentrated with Millipore Amicon Ultra 3 MWCO centrifugal microconcentrators into 10 mM Tris, pH = 8.00 five times at 10 minutes each into 500 μ L LoBind microcentrifuge tubes for a combined recovery of 24.0 % or 22.4 %, respectively. Samples were quantified by UV spectroscopy at Abs₂₆₀ on a Thermo Scientific NanoDrop One and the IDT supplied extinction coefficient of the T-Lib DNA library.

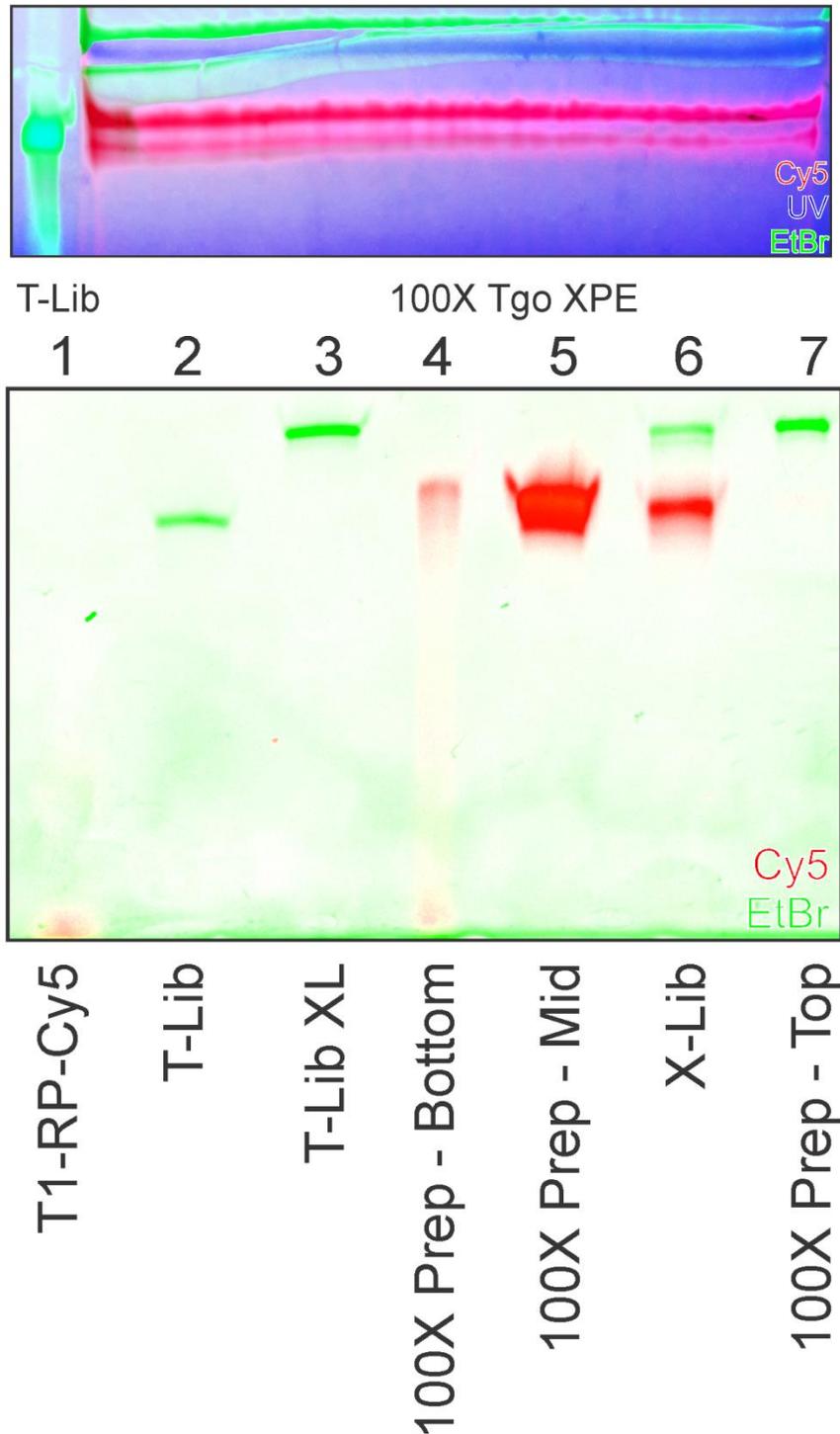


Figure 3. – 12.5 % 29:1 PAGE extraction of 100X FANA XPE & 10 % analytical gel
 The self-cast gel (top) was loaded with 150 pmol T-lib in the reference well to visually track the separation of 500 pmol FANA XPE reaction mixture; pre-heated at 16 W for 15 min and run at 16 W for 30 min. Composite channels allow FANA product (in red) to be clear separated from DNA, T-Lib XL (in green). Cut outs indicate the top (DNA) and bottom bands (FANA) were able to be excised under UV shadowing (in blue). 10 % 29:1 continuous TBE-Urea analytical gel (bottom): 15 μ L sample at 0.091 μ M in DI formamide, 1 hr at 180 V, 1 hr preheated 8 W.

DNA Polymerase Mediated FANA 'Reverse' Transcription into cDNA

A preliminary method for 'reverse' transcription of the generated FANA library back into DNA ¹¹⁰. The DNA polymerases chosen were a family DNA polymerases based on the polymerase found in *Bacillus stearothermophilus*; these polymerases are also known for being 3'→5' exonuclease activity deficient ¹³⁶, while the large fragment of this polymerase (Bst LF) is also 5'→3' exo-, similar to Klenow fragment. Bst 2.0 and Bst 3.0 are mutants based on this Bst LF polymerase, with increasing optimizations for isothermal DNA replication and DNA/RNA reverse transcription activity. The polymerases were originally screened at the stock concentrations of 5.0 U·μL⁻¹, 8.0 U·μL⁻¹, 8.0 U·μL⁻¹, and 5.0 U·μL⁻¹ respectively; but, were rescreened in concentrations of 1.6 U·μL⁻¹ and 0.8 U·μL⁻¹ after dilution in NEB buffers D & specific Bst 2.0 buffer.

The reaction conditions were similar to those of FANA XPE- reactions in were 5 μL reaction volumes, due to the scarcity of X-Lib material, containing 1X reaction buffer (ThermoPol or otherwise) reaction buffer supplemented with 0.3 mM MgCl₂, 1 μM T-Cy3-FP primer (relative to T-Lib), 1 μM X-Lib FANA library, 500 μM of standard dNTPs purchased from NEB, and 1 μL isothermal DNA polymerase (Bst, Bst LF, Bst 2.0, and Bst 3.0) ¹⁰⁹. Primers and templates were denatured by heating at 90 °C on a thermocycler for 5 minutes for proper annealing on ice for 10 minutes, prior to the addition of polymerase. Upon addition of the polymerase, the samples were incubated for 3 hours at 50 °C. Reactions were quenched with the addition of 100 μL (10 equivalents v/v) of 99 % DI formamide and 25 mM EDTA. These were analyzed in 15 μL volumes onto 30 minute 8 W preheated denaturing Invitrogen Novex 10% TBE-Urea gels, ran at 180 V for 1 hr.

Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)

The current state-of-the-art method for generating nucleic acid aptamers involves the use of Systematic Evolution of Ligands by EXponential enrichment (SELEX). This method allows for the systematic amplification of ‘hit’ sequences which are capable of binding to a stationary target; however, this method limits the total sequence-scope due to the labor intensive screening. A new method to overcome this limitation is Particle Display (PD), which works with this paradigm but allowing for a high throughput method of screening by functionalizing the aptamer library members to magnetic beads, capable of screening $\sim 10^7$ particles/hour, as well as multiparametric simultaneous positive/target and negative/non-target selection via a Fluorescence-Activated Cell Sorting (FACS) instrument^{137,138}. By anchoring the nucleic acid library to the relatively stationary magnetic particles, recombinant fluorescent targets are instead washed over- the inverse of SELEX.

In order to successfully perform PD it is important to generate magnetic particles which are monoclonal, only displaying one species of each member of the starting nucleic acid library. This is accomplished through emulsion PCR (ePCR). In this workflow amino-modified DNA forward primers (T1-AM6-FP, Forward 1) were attached to carboxylic acid functionalized Invitrogen Dynabeads MyOne beads via 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) carbodimide crosslinking chemistry. This requires a specialized forward primer with a 5' amino group modification for covalent coupling, anchoring nucleic acids to the beads during thermal cycling. In order to prospectively screen a FANA library, starting from a bead attached DNA library,

a DNA-strand reverse-complement/XNA-strand forward primer was designed that was capable of bead functionalizing with a second coupling chemistry.

It was found that the bead to phosphorus distance between Forward 1 was found to be 33.7 Å from molecular mechanic stimulations on structures built in Avogadro and optimized to the Universal Force Field (UFF) ¹³⁹; hence, utilizing that the functionalized FANA library forward primer terminal triple bond modification available through IDT (T1-TTB-RC, Forward 2), a comparable distance of 33.0 Å is possible with the additional functionalization of an amino-dPEG₅-azide linker molecular, available through BroadPharm (San Diego, CA), following a CuAAC click coupling.

Coupling of FP to Magnetic Microparticles

This method was conducted according to previous methods ^{137,138}. Beads were prepared in Eppendorf LoBind tubes for use with a DynaMag magnet; while handling beads, care was taken to not keep the magnetic beads longer than 1 min nor to dry, in order to prevent aggregation and changes in bead surface chemistry. Invitrogen Dynabeads MyOne CA (10⁷/μl) were vortexed and a 500 μL aliquot was washed once with 500 μL 0.01 M NaOH, three times with treated Thermo Scientific Ambion non-DEPC treated nuclease-free water, and resuspended in 150 μL of an EDC reaction solution containing 0.2 mM T1-AM6-FP, 200 mM NaCl, 1 mM imidazole chloride, and 250 mM EDC prepared in dimethyl sulfoxide (DMSO, 50 % v/v). Particles were mixed well with three 1 sec vortex pulses and dip sonication, and incubated overnight at room temperature on a 10 rpm rotator.

Prospectively, these beads could be washed with nuclease-free water and the above protocol could be repeated with an EDC reaction mixture containing amino-dPEG₅-azide in order to additionally couple and extend the secondary T1-TTP-RC XNA-

strand forward primer later with the CuAAC click reaction and DNA templating bridge amplification. This would allow each bead to essentially contain both a DNA ‘informational’ and FANA ‘functional’ strand without the need of a ‘reverse transcription’ process for aptamer ‘hit’ sequencing. However, the orthogonality of this method remains to be tested. The remainder of this work focuses on merely coupling a single library to these beads.

Surface Pacification of FP-Magnetic Microparticles

The inclusion of the additional PEG18 spacer between the terminal amino moiety and primer sugar was intentional. This minimizes potential interactions of aptamer candidates to the particle surface which could potentially lead to false positive selection and reduce the functional ‘hit’ library diversity for a target protein. In order to minimize undesired nucleic acid/bead-surface interactions, we incorporated a PEG18 molecule as a spacer between the particle surface and the forward primer at the 5’ end. However, in order to prevent further non-specific surface effects between the beads, the beads were pacified in excess with an amino-dPEG12-alcohol molecule purchased from BroadPharm.

The particles were washed with 500 μ L 0.1 M (*N*-morpholino)ethanesulfonic (MES), pH 4.7 buffer three times, followed by three 1 sec vortex pulses and dip sonication. Beads were then resuspended in a fresh pacification solution containing 100 mM *N*-Hydroxysuccinimide (NHS) and 200 mM EDC in cold 0.1 M MES, pH = 4.7 buffer. Particles were mixed well with three 1 sec vortex pulses and dip sonication, and incubated for 20 minutes at room temperature on a 10 rpm rotator. The use of NHS over imidazole is designed to preferentially convert any free bead carboxyls into highly amino-reactive NHS-esters; additionally, this step is conducted in MES buffer as the

EDC coupling reaction is ideal at low pH values, but the stabilized NHS-ester is capable of coupling in physiological pH~7 conditions.

The particles were then washed two times in 500 μ L cold DPBS and resuspended in 20 mM amino-dPEG12-alcohol in DPBS buffer. Particles were mixed well with one 1 sec vortex pulse and dip sonication, and incubated for 30 minute at room temperature or 2 hours at 4 $^{\circ}$ C on a 10 rpm rotator. The reaction was quenched by washing the particles once with 500 μ L TT buffer (250 mM Tris, 0.1 % v/v Tween 20, pH=8.00) and three times with 500 μ L STT buffer of (10 mM Tris·HCl, 0.1 % v/v Tween 20, pH=8.00) buffer for storage at in 500 μ L STE at 4 $^{\circ}$ C.

Testing Conjugation of FP-Magnetic Microparticles

In order to verify that the coupling of F1-AM6-FP was successful, a Cy5 fluorescent compliment is hybridized to the beads in order to indicate the presence of the primer after washing. Additionally, this is an important checkmark to track whether any further optimizations are necessary for the previous steps. To test the conjugation efficiency, 5 μ L 10^7 FP-particles were resuspended in 50 μ L PBST (1X DPBS + 0.05 % v/v Tween 20) in an aluminum foil wrapped microcentrifuge tube and washed three times with 50 μ L PBST. This 50 μ L of 10^6 particles was again resuspended with 20 μ M fluorescently-modified T1-CY5-FC in 500 μ L PBST room temperature for 30 minutes on a 10 rpm rotator. The beads were washed twice with 500 μ L PBST before resuspension in 500 μ L for storage. A dilution of 10^5 beads, 10^5 FP-particles, and 10^5 FP-particles + fluorescent tag was analyzed on a Bio-Rad S3 Cell Sorter-FACS at 648-668 nm (Appendix G).

Emulsion PCR Generation of Magnetic Microparticles Library

Generation of the T-Lib-Particle library was conducted according to previously reported methods, utilizing only natural dNTPs⁹⁴. A PCR reaction was conducted on the surface of each bead, by controlling the concentrations of the library to primer-attached-beads, such that probabilistically each water in oil (w/o) emulsion contains one of each, with the addition of a reverse primer in excess. The emulsion PCR mixture was setup such that the generated second strand would present a Cy5 fluorescent tag, allowing the second strand to be detected only prior to helix denaturation. A 12 μL aliquot of 10^7 T1-AM6-FP conjugated particles was resuspended in 1000 μL aqueous reaction mix containing 1X Mg-free Taq reaction buffer, 1.5 mM MgCl_2 , 10 nM T1-FP regenerative forward primer, 1 μM T1-Cy5-RP QC DNA reverse primer, 1.5 pM T-Lib DNA library, 200 μM of each dNTPs purchased from NEB, and $0.05 \text{ U} \cdot \mu\text{L}^{-1}$ Taq polymerase. The reaction mixture was transferred to an IKA DT-20 tube and placed on an IKA Ultra-Turrax device. Over 5 minutes at 620 rpm an emulsion was formed by adding a freshly prepared oil phase consisting of biological grade Sigma Aldrich (St. Louis, MO) 4.5% Span 80, 0.40% Tween 80 and 0.05% Triton X-100 in mineral oil. This emulsion was transferred in 100 μL aliquots to a 96 well plate, sealed with thermal resistant film, and run on thermocycler with the following program- 95 $^\circ\text{C}$, 120 sec + [95 $^\circ\text{C}$, 15 sec + 56 $^\circ\text{C}$, 30 sec + 68 $^\circ\text{C}$, 60 sec]*39 + 68 $^\circ\text{C}$, 300 sec + 4 $^\circ\text{C}$ hold. After PCR, Applied Biosystems (Foster City, CA) SOLiD emulsion collection trays were attached to the top of the trays with adhesive tape, inverted, and spun on a centrifuge at 300 x g for 2 minutes, in order to transfer the emulsions into the trays.

The emulsion was broken by adding 10 mL sec-butanol to the tray in 1 mL aliquots. Afterwards, the sample was transferred to a 50 mL centrifuge tube and mixed

with vortexing for 30 seconds, and spun on a centrifuge at 3,000 x g for 5 minutes, in order to pellet the particles. The oil phase was carefully removed and the particles were resuspended in 600 μ L of emulsion breaking buffer (10 mM Tris, 1 mM EDTA, 100 mM NaCl, and 1% v/v Triton-X) and transferred into a 1.5 mL microcentrifuge tube. The remaining supernatant was removed by vortexing for 30 seconds, followed by centrifugation at 15,000 x g for 90 seconds, and the use of the DynaMag magnet. The particles were washed three times with DPBS buffer and stored in 300 μ L PBST.

Testing Extension Efficiency of Emulsion PCR

In order to confirm whether the generated T1-particles had the appropriate PCR extension ratio of $\sim 20 \pm 15$ %, necessary for efficient emulsion PCR combinatorics, in which adding too little template limits particle-library diversity but adding too much template material leads to non-monoclonal beads¹³⁷. To test the extension efficiency, 25 μ L 1:25 10^7 T1-particles were washed once with non-DEPC treated nuclease-free water and resuspended in 250 μ L 0.1 M NaOH at room temperature for 5 minutes in an aluminum foil wrapped microcentrifuge tube on a 10 rpm rotator. The denatured T1-particles were washed three times with 250 μ L nuclease-free water and once with 250 μ L PBST. A 50 μ L aliquot of these 10^6 particles was again resuspended in 1 μ M fluorescently-modified T1-CY5-FC in 500 μ L PBST at room temperature for 30 minutes on a 10 rpm rotator. The beads were washed twice with 250 μ L PBST before resuspension in 250 μ L for storage.

A dilution of 10^5 beads, T1-particles, 10^5 ssT1-particles + fluorescent tag, and 10^5 dsT1-particles + fluorescent tag were analyzed on an Accuri C6+ BD Biosciences FACS at 648-668 nm (Figure 8). This procedure was repeated with 500 pM T1-Lib in response to preliminary data to reduce the extension efficiency (Figure X). Although this was not

conducted, prospectively this should be repeated once more with 250 pM library to further reduce the ePCR particle extension efficiency.

Proposed – FANA ePCR & Click Conjugation of Aptamer Microparticles

In order to utilize the emulsion PCR technique in order to generate nuclease resistant aptamers, it is key that a XNA library members are capable of being attached to the magnetic microparticles, as was accomplished with DNA. The ability for KOD XL in order to incorporate modified DNA C8-alkyl-uridine triphosphates in the generation of DNA aptamers has previously been described ^{140,89} ; additionally, we have demonstrated that Tgo (exo-) and KOD XL have DNA templated FANA transcription activity (Figure 4). The most ideal path to generating these aptamers may well depend on the future results from this project, specifically the activity and fidelity of saccharide clicked- C8-alkyl-uridine FANA ‘reverse’ transcription back into easily sequenced DNA. This is why we have designed a monoclonal two primer (T1-AM6-FP, amino-dPEG₅-azide + T1-TTB-RC) system, if this ‘reverse’ transcription is problematic for functionalized FANA aptamers.

Monoclonal, particle-displaying non-natural C8-alkyl-uridine DNA sequences have been previously clicked to monosaccharides using the CuAAC reaction ⁸⁹. This allows SELMA using FACS for lectin targets, screening for the optimum number of multivalency and GAG length necessary for positive selection and screening against a non-target lectin negative selection ¹³⁸. Prospectively, the T1-TTB-RC primer would be CuAAC clicked to a DNA library + PEG₅-azide functionalized magnetic microparticle, following emulsion PCR, following a similar procedure as is necessary for click-conjugation of DNA aptamer particles to saccharides.

The optimal conditions for the click-conjugation of azido-sugar monosaccharides was previously found to be 0.4 mM Cu(I)Br & 0.4 mM tris((1-benzyl-4-triazolyl)methyl)amine (TBTA) with acetyl protected 1' azido-ethyl monosaccharides suspended in 100 mM methanol⁸⁹. In full, T1-particles are resuspended in an aluminum foil wrapped tube with 10 μ L DPBS buffer pH 8.0 and degassed for at least 15 minutes with N₂ bubbling. Afterwards, 25 μ L of 20 mM Na₂HPO₄ and 5 μ L 10 % v/v Tween 20 are added and the click reaction is initiated by the addition of 5 μ L 100 mM (methanol) 2-azidoethyl 2,3,4,6-tetra-O-acetyl-pyranoside and 2.5 μ L of premixed 1:1 solution of 10 mM (3:1 DMSO:tBuOH) Cu:TBTA. The reaction tube is mixed well with three 1 sec vortex pules, placed into a 20 mL vial with a septum, purged with N₂ for 5 minutes. Following incubation for 2 hours at room temperature, the dsDNA particles were washed five times with 300 μ L TE buffer and resuspended in 250 μ L 0.1 M NaOH for 5 minutes at room temperature on a 10 rpm rotator. The sugars were deprotected from the ssDNA-particles by resuspension in 300 μ L 18 M ammonium hydroxide solution for 3 hours at room temperature on a 10 rpm rotator. The final particles are washed five times with 300 μ L TE buffer and stored in 300 μ L 10 mM Tris.

Proposed – Screening of Modified Aptamer Microparticles

In order to finally screen the generated microparticles for aptamers, the target and counter-selection target are necessary. It is proposed that the best counter-selection target is a target lectin protein which is within the same clade as the target lectin, but sensitive to different ligands. Working from previously reported procedures, denatured modified aptamer microparticles are resuspended with 1 nM fluorescently tagged target-lectin 250 nM differently fluorescently tagged counter-selection target in PBSMCT buffer (1X DPBS + 2.5 mM MgCl₂, 1 mM CaCl₂, 0.01% v/v Tween 20) for 1 hour at room

temperature on a 10 rpm rotator⁸⁹. It is possible to extend this protocol to alternatively screen in biotinylated human serum and label all serum proteins with 50 nM streptavidin-conjugated fluorescent dye¹³⁸. Additionally, SELMA can be enhanced by, decreasing the target lectin concentration every round after 95% of the initial microparticles are sorted out and with screening at 37 °C instead of room temperature for the final four rounds⁵⁵. Similarly, we propose screening could additionally be enhanced by increasing the selection entropy with increasing salt concentrations.

Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA

All synthesis targets were verified by NMR spectroscopy in conjunction with either Direct Analysis in Real Time (DART) or LC/MS TOF. All heteronuclear ¹⁹F-NMR spectra were recorded on a Varian 500 MHz H-19F/15N-31P OneNMR PFG probe instrument, while 2D NMR spectra were primarily recorded on a Varian 600 MHz 1H-19F/15N-31P PFG AutoX DB NB probe and a 600 MHz 1H{13C/15N} 5mm PFG automation-enabled triple-resonance probe instrument; all spectra were recorded by the primary author and cordially maintained by Boston College's Magnetic Resonance Staff Dr. Thusitha Jayasundera & Dr. Jing Jin. Mass spectrometry samples were submitted for analysis on JEOL AccuTOF DART or Agilent 6220 Time-of-Flight LC/MS systems, based on the expected stability of the generated molecules and overall pre-/post-silica gel chromatography purity, by Boston College's Mass Spectrometry Center's Dr. Marek Domin. The correct synthesis of the β -glycoside was confirmed relative to previous reported NMR spectra and by recrystallization followed by X-ray diffractometry. X-ray crystals were submitted for analysis on Bruker Kappa Apex Duo by the Boston College's Dept. of Chemistry's X-ray Crystallography Center's Dr. Bo Li.

5-Iodouracil

The synthesis was conducted according to previously reported conditions with minor modifications ¹⁴¹ and in good agreement with previously reported spectra ¹⁴². A mixture of uracil (300 mg, 2.68 mmol) and ceric ammonium nitrate (CAN) (733.04 mg, 1.34 mmol), in MeOH (16 mL) was stirred at 70 °C for 6.5 hours. Reaction progress was monitored by thin layer chromatography (TLC) in EtOAc/i-PrOH/H₂O, 4:1:2. The solvent was evaporated in vacuo and gave a colorless solid powder that was crystallized from 46 mL EtOH/H₂O (1:1) to yield two crops of crystals 312.4 mg (with a slight orange hue) and 203.7 mg of colorless fine needles. The flask was covered in foil to prevent light sensitive damage. Further material was purchased from TCI (Tokyo, JPN)

¹H NMR (500 MHz, DMSO-*d*₆) δ : 11.39 (brs, 1H, NH-3), 11.15 (brs, 1H, NH-1), 7.88 (s, 1 H, H-6). ¹³C NMR (126 MHz, DMSO-*d*₆) δ : 161.46 (C-4), 151.21 (C-2), 146.94 (C-6), 67.56 (C-5). DART-MS *m/z*: 238.93216 [M + H]⁺.

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose

The synthesis was conducted according to previously reported conditions with minor modifications ¹⁴³ and in good agreement with previously reported spectra ¹⁴⁴. The reaction was conducted in oven-dried glassware under a N₂ atmosphere in a multi-neck round bottom flask. To a stirred solution of 1,3,5-tri-O-benzoyl- α -D-ribofuranose (5.74 g, 12.41 mmol) in DCM (86.2 mL), diethylaminosulfur trifluoride (DAST) (6.00 g, 37.22 mmol, 4.92 mL) was added dropwise at 0 °C, and the resulting mixture allowed to stir at room temperature for 20.5 hr. The reaction progress was monitored by TLC, using DCM and developed using 10 % H₂SO₄ in EtOH staining with heating. The reaction was cooled and quenched with saturated aqueous NaHCO₃ (57.47 mL). The organic layer was first washed with water (2 \times 250 mL) and then with saturated aqueous brine (250 mL). This was then dried over Na₂SO₄, filtered and evaporated in vacuo. The crude product was purified by silica gel column chromatography using 5.5 cm x 25 cm column with a 10-20

g loading capacity; the column was run using 500 mL 8:1 hex:EtOAc in order to first remove multi-fluorinated byproducts, followed by a 1 L of 4:1 hex:EtOAc. This yielded 4.17 g (72.3 %) of sticky yellow oil.

¹H NMR (600 MHz, Chloroform-d) δ : 8.13 – 8.03 (m, 6H, m-Bz), 7.66 – 7.52 (m, 3H, p-Bz), 7.48 – 7.38 (m, 6H, o-Bz), 6.76 (d, J = 9.1 Hz, 1H, 1'H), 5.64 (dd, J = 19.5, 3.3 Hz, 1H, 3'H), 5.40 (d, J_{HF} = 48.2 Hz, 1H, 2'H), 4.82 – 4.68 (m, 4H, 4'H & 5'H). ¹⁹F NMR (470 MHz, Chloroform-d) δ : -190.84 (ddd, J = 48.1, 19.6, 9.2 Hz). ¹³C NMR (151 MHz, Chloroform-d) δ : 166.20, 165.39, 164.49, 133.93, 133.80, 133.20, 129.94, 129.89, 129.87, 129.77, 129.63, 129.62, 129.16, 129.14, 128.75, 128.73, 128.65, 128.55, 128.41, 99.48, 99.23, 97.59, 96.36, 84.03, 76.87, 76.66, 63.59. DART-MS m/z : 482.15982 [M + NH₄]⁺.

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- β -D-Arabinofuranosyl-5-Iodouridine

o *Entry 1*

The synthesis was conducted with modifications to similar syntheses^{145,146} utilizing an and in good agreement with previously reported spectra^{147,148}. A solution of (2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-arabinofuranose (3.89 g, 8.38 mmol) was azeotropically dried with 3x30 mL toluene (at 40 mbar, 40 °C) and 5-iodouracil (2.08 g, 8.76 mmol) in anhydrous acetonitrile (ACN) (13.37 mL) was added N,O-bis(trimethylsilyl)-acetamide (BSA) (3.66 g, 21.78 mmol, 4.40 mL) and the mixture was stirred for 25 minutes at 60 °C. Trimethylsilyl trifluoromethanesulfonate (TMSOTf) (5.58 g, 25.13 mmol, 4.86 mL) was added dropwise over 10 min, and stirring was continued for another 28:45 hours at 60 °C, after which time TLC analysis (3:2 hexanes–EtOAc) showed the reaction to be complete.

The solution became a light yellow-brown color 5-10 minutes after adding BSA, and the reaction mixture became a light chestnut oil color immediately after adding the TMSOTf Lewis acid. After overnight (O/N), at 14:15 hr after adding the TMSOTf, the reaction took on a dark-reddish brown color.

The mixture was cooled to room temperature, diluted with 27.9 mL (3/2 vol) of EtOAc, and poured into 18.6 mL (1 vol) of cold sat. aq. NaHCO₃ solution with stirring. This yielded unreacted base as a pure white powder, this was filtered off through a Hirsch

funnel. The organic layer was then evaporated off under pressure. The organic layer was separated and washed with H₂O (2x250 mL) and brine (2x250 mL), dried over Na₂SO₄, and concentrated in vacuo.

After purification, a significant amount of starting materials was recovered- 2.1549 g of sugar SM and 1.3253 g unreacted base SM, limiting the reaction to a theoretical yield of 36.3 %. This clearly indicates that the base is not being activated properly, since it is recovered in excess compared to the sugar; while it is possible that some sugar is being removed as a salt due to existing as an oxycarbenium ion^{149,150}, yet upon addition of BSA the solution becomes clear, suggesting that all of the base is being in fact silylated. Hence, there does not seem to be enough TMSOTf promoter being added to drive the reaction. However, 1.18 g of the alpha anomer and 577.10 mg of the beta anomer were recovered. This 2.05:1 alpha:beta ratio, in good agreement to the previous report for the use of this starting sugar and Friedel-Craft's catalyst¹⁵¹. The final products are silvery-white in color and include small amounts of dark yellow-orange contaminants. This is particularly obvious when running the column, as the beta product co-eluted with this contaminant, giving a lime green colored band. The majority of this contaminant was leached from the pure beta product by washing the beta product in EtOAc and decanting the colored solution off of the white crystal. Additionally, 90.1 mg of bis-beta nucleoside product was also recovered, exhibiting a unique NMR signature and elution order (data not shown), as expected.

○ *2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide*

In order to explore the potential application of a one-pot sialylation and glycosylation, halogenation of the 1' positioned for a SN₂ 'Walden inversion' enantioselective β glycosylation was conducted with bromine¹⁵². This approach is well documented for the

synthesis of inductively hindered 2'halo furanoses and especially bulky nucleobases^{153–155}. The synthesis was conducted without modifications to previous reports¹⁴³ and in good agreement with previously reported spectra¹⁴⁷.

The reaction was conducted in oven-dried glassware under a N₂ atmosphere in a 25 mL round bottom flask. The fluoro-arabinose sugar 2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-arabinofuranose (1.47 g, 3.16 mmol), DCM (7.02 mL), and 33% HBr (5.792 M, 1.64 mL) were stirred for 10.5 hours at 22 °C. Monitoring the reaction with DCM TLC. On addition of the HBr, the solution took on a light orange color.

After 10.5 hours the reaction had a fine ruby orange color. The solution was evaporated in vacuo, redissolved in DCM (7.02 mL), and then washed with water (2x10 mL) and brine (2x10 mL). The solution was dried with NaSO₄ and concentrated to a viscous dark red syrup. Following drying over 36 hr on a high pressure vacuum pump gave a final yield of 1.2167 g of product for 91.23% isolated yield. Afterwards this was azeotropically dried with toluene (3x15 mL), which was further dried under high vacuum 18 hr.

¹H NMR (500 MHz, Chloroform-*d*) δ : 8.23 – 8.02 (m, 4H, o-Bz), 7.74 – 7.52 (m, 2H, p-Bz), 7.52 – 7.40 (m, 4H, m-Bz), 6.64 (d, *J* = 12.2 Hz, 1H, 1'H), 5.60 (d, *J* = 50.0 Hz, 1H, 2'H), 5.55 (ddd, *J* = 22.1, 4.0, 1.3 Hz, 1H, 3'H), 4.86 – 4.77 (m, 2H, 5'H), 4.72 (dd, *J* = 11.6, 3.9 Hz, 1H, 4'H). ¹⁹F NMR (470 MHz, Chloroform-*d*) δ : -165.95 (ddd, *J* = 50.0, 22.1, 12.2 Hz). DART-MS *m/z*: 423.02267 [M + H]⁺.

o *Entry 2*

In a separate 10 mL round bottom, 5-iodouracil base (720.31 mg, 3.03 mmol) was dissolved in dry ACN (4.6 mL) BSA (1.52 g, 7.49 mmol, 1.83 mL), was added the mixture, and was stirred for 30 minutes at 60 °C. Once the solution was clear, it was allowed to cool and the excess silylation agent was removed by evaporation in vacuo to afford the silylated 2,4- bis-O-(trimethylsilyl)-5-iodouracil base. To this, NaI (345.67 mg, 2.31 mmol) was added and was redissolved in dry ACN (793.13 μ L).

The reaction was conducted in oven-dried glassware under a N₂ atmosphere in a 50 mL round bottom flask. Azeotropically dried fluoro sugar 2-deoxy-2-fluoro-3,5-di-O-benzoyl- α -D-arabinofuranosyl-bromide (1.22 g, 2.88 mmol) was dissolved in DCM (3.81 mL). The silylbase solution was injected and the reaction was monitored by TLC 1:1 hex:EtOAc, showing the reaction to be complete after 99 hrs, or 4 days.

After adding BSA the first solution became a light yellow-brown orange color, and immediately after adding the base the reaction mixture became red orange and slowly developed into a light chestnut oil color. After O/N, at 14:15 hours, the reaction took on a dark-reddish brown color similar suggesting that this reaction is proceeding correctly; however also forming colored byproducts.

The mixture was cooled to room temperature, diluted with 2.88 mL (3/2 vol) of DCM, and quenched with 1 mL of cold sat. aq. NaHCO₃ solution, filtered through a Buchner funnel, washing with CAN. This yielded unreacted base as a pure white powder. The organic layer was then evaporated in vacuo to dryness and rediluted in DCM. The organic layer was separated and washed with H₂O (2x125 mL) and brine (2x125 mL), dried over Na₂SO₄, and concentrated under in vacuum.

According to TLC after workup, the starting materials side-reacted, likely with the residual BSA in solution. ¹⁹F-NMR analysis found that the combined integral of these sugar byproducts was 2.42 (data not shown), while the alpha and beta were found to have a 1.00 and 4.06. This resulted in a 4.06⁻¹ α : β selectivity and a 67.65 % yield, with 54.28 % the target product.

o *Entry 3*

The reaction was conducted in oven-dried glassware under a N₂ atmosphere in a 25 mL round bottom flask. The day prior, 5-iodouracil was silylated to 2,4- bis-O-

(trimethylsilyl)-5-iodouracil with HMDS ¹⁴⁷. To a mixture of 5-iodo-1H-pyrimidine-2,4-dione (1.10 g, 4.61 mmol), dry ACN (9.45 mL), was added HMDS (1.68 g, 10.41 mmol, 2.17 mL) at 40 °C. The resulting mixture was heated at reflux (85 °C) for 24 hr, then cooled, and concentrated under reduced pressure to syrup, which was used in the subsequent step. On drying, this solution became cloudy. However, this solution was capable of being dried in vacuo, as HMDS was counter-intuitively found to be easy to evaporate over BSA.

The next reaction was also conducted in oven-dried glassware under a N₂ atmosphere in a 50 mL round bottom flask. The fluoro-arabinose sugar 2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-arabinofuranose (1.62 g, 4.19 mmol), in DCM (5.27 mL), was stirred with ACN (1.10 mL) containing the HMDS prepared 2,4- bis-O-(trimethylsilyl)-5-iodouracil base (1.62 g, 4.19 mmol) and added NaI (478.83 mg, 3.19 mmol) at 22 °C. The reaction was monitored by TLC with 1:1 hex:EtOAc, showing the reaction to be complete in approximately 114 hrs, or 4.75 days.

After adding the base the solution became a light yellow-brown color. After O/N, at 14:15 hours, the reaction took on a dark-reddish brown color suggesting that this reaction is proceeding correctly; however also forming colored byproducts.

The reaction mixture was worked-up similar to Entry 2. After purification, 129.9 mg of the alpha anomer and 687.3 mg of the beta anomer were recovered; however, it was found that the formation of a yellow, very EtOAc soluble, decomposition product was formed with increased temperature rotovaping. Despite this, we report 4.09⁻¹ α : β selectivity, which may well have been higher.

○ *Entry 4*

Similarly to Entry 2, 5-iodouracil base (785.26 mg, 3.30 mmol) was prepared in dry ACN (5.02 mL) with BSA (1.66 g, 8.17 mmol, 2.00 mL). To the dry flask, NaI (376.83 mg, 2.51 mmol, 94.19 μ L) was added and was dissolved in dry THF (5.02 mL).

The reaction was conducted in oven-dried glassware under a N_2 atmosphere in a 50 mL round bottom flask. Azeotropically dried fluoro sugar 2-deoxy-2-gluoro-3,5-di-O-benzoyl- α -D-arabinofuranosyl-bromide (1.33 g, 3.14 mmol) dissolved in THF (5.02 mL). The silylbase solution was injected and the reaction was monitored by TLC 1:1 hex:EtOAc, showing the reaction to be complete after 99 hrs, or 4 days.

Unlike the other reactions, after adding BSA the first solution had no change in color. After O/N, at 14:15 hours, the reaction took on a dark-reddish brown color similar suggesting that this reaction is proceeding correctly; however also forming colored byproducts.

The reaction mixture was worked-up similar to Entry 2. According to TLC after workup, the starting materials side-reacted, likely with the residual BSA in solution. ^{19}F -NMR analysis found that the combined integral of these sugar byproducts was 21.46 (data not shown), while the alpha and beta were found to have a 1.00 and 13.76. This resulted in a 13.76^{-1} α : β selectivity and a 40.75 % yield, with 37.99 % the target product.

○ *Purification*

We found that a silica gel hex:EtOAc gradient of 1.8 L 8:1, 5.4 L 3:1, & 600 mL 2:1, with a 100 mL 20% MeOH/EtOAc flush, it is possible to get complete separation of the SM and each anomer on a 4 cm x 45 cm column. TLC of these gradients suggests that only 8:1 and 3:1 are necessary to elute well, holding to a rule of thumb that gives the next highest polarity spot have a target Rf of 0.1 in the new solvent, respectively to the sugar

SM and the alpha anomer. The 2:1 could still be used to elute the beta anomer, if it is the only one left on the column.

We found the most successful crystals were generated by adding warm DCM until the solution was clear, cooling to room temperature, adding hexanes until the solution was cloudy, then heating the and adding hexanes dropwise until the solution stayed cloudy, followed by one last drop of DCM resulting in a clear solution, producing a light cream colored crystal. Afterwards, iPrOH was added with heating to 60 °C, cooling with evaporation overnight lightly covered flask in the back of a fumehood, while cooling overnight at 4 °C within a refrigerator was found to increase impure precipitates.

β : ^1H NMR (600 MHz, DMSO- d_6) δ : 11.94 (s, 1H, NH-3), 8.03 (ddd, $J = 27.4, 7.8, 1.4$ Hz, 4H, o-Bz), 7.99 (d, $J_{5'H} = 2.0$ Hz, 1H, H-6), 7.69 (dtt, $J = 24.7, 7.5, 1.4$ Hz, 2H, p-Bz), 7.55 (dt, $J = 21.3, 7.8$ Hz, 4H, m-Bz), 6.30 (dd, $J = 19.7, 3.8$ Hz, 1H, 1'H), 5.69 (ddd, $J = 19.7, 4.7, 1.6$ Hz, 1H, 3'H), 5.53 (ddd, $J_{HF} = 50.4, J_{HH} = 3.9, 1.7$ Hz, 1H, 2'H), 4.79 – 4.71 (m, 2H, 5'H), 4.62 (q, $J = 4.3$ Hz, 1H, 4'H)

α : ^1H NMR (600 MHz, DMSO- d_6) δ : 11.83 (s, 1H, NH-3), 8.00 (dtt, $J = 21.6, 6.9, 1.3$ Hz, 4H, o-Bz), 8.20 (s, 1H, H-6), 7.69 (dtt, $J = 11.7, 7.3, 1.3$ Hz, 2H, p-Bz), 7.54 (t, $J = 7.7$ Hz, 4H, m-Bz), 6.19 (dd, $J = 17.0, 1.1$ Hz, 1H, 1'H), 5.74 (dt, $J_{HF} = 49.3, J_{HH} = 1.7$ Hz, 1H, 2'H), 5.71 (dt, $J = 17.0, 3.3, 1.8$ Hz, 1H, 3'H), 5.26 (q, $J = 4.7$ Hz, 1H, 4'H), 4.62 – 4.46 (m, 2H, 5'H)

mix: gHSQC ^{13}C NMR (101 MHz, DMSO- d_6) δ : 129.36, 128.90, 144.68, 133.56, 128.51, 83.18, 83.11, 76.18, 75.86, 91.95, 93.63, 62.95, 78.45/ ^1H NMR (400 MHz, DMSO- d_6) δ : 8.05, 8.01, 7.99, 7.70, 7.55, 6.31, 6.27, 5.72, 5.66, 5.60, 5.47, 4.75, 4.62.

mix: ^{13}C NMR (101 MHz, DMSO- d_6) δ : 165.54 (C=O), 164.82 (C=O), 160.28 (C4), 149.71 (C2), 144.95 (C6), 133.97 (p-Bz), 133.64 (p-Bz), 129.66 (o-Bz x2), 129.24 (o-Bz x2), 129.14 (Bz), 128.87 (m-Bz x2), 128.80 (m-Bz x2), 128.63 (Bz), 94.04 ($\beta^2\text{H}$), 92.14 ($\alpha^2\text{H}$), 83.56($\beta^1\text{H}$), 83.40($\alpha^1\text{H}$), 78.76(4'H), 76.47 ($\alpha^3\text{H}$), 76.17($\beta^3\text{H}$), 69.59(C5), 63.26(5')

mix: ^{19}F NMR (470 MHz, DMSO- d_6) δ : -186.57 (dt, $J = 49.5, 16.8$ Hz, α), -190.90 (ddd, $J = 47.2, 19.5, 8.7$ Hz, SM), -199.17 (dt, $J = 50.6, 19.9$ Hz, β)

β : ^1H NMR (400 MHz, Acetone- d_6) δ : 10.85 (bs, 1H, NH-3), 8.13 (t, $J = 8.2, 7.2$ Hz, 4H, o-Bz), 8.10 (d, $J = 3.0$ Hz, 1H, H-6), 7.69 (dt, $J = 14.1, 7.5, 6.3$ Hz, 2H, p-Bz), 7.56 (td, $J = 7.5, 4.6$ Hz, 4H, m-Bz), 6.41 (dd, $J = 20.8, 3.2$ Hz, 1H, 1'H), 5.77 (dd, $J = 19.0, 3.7$ Hz, 1H, 3'H), 5.54 (dd, $J = 50.5, 3.2$ Hz, 1H, 2'H), 4.88 (dd + t, $J = 30.4, 4.0$ Hz + $J = 5.3, 4.9$ Hz, 2H, 5'H), 4.76 (q, $J = 3.9$ Hz, 1H, 4'H)

α : ^1H NMR (500 MHz, Acetone- d_6) δ : 8.24 (bs, 1H, H-6), 8.07 (ddd, $J = 35.4, 7.2, 1.4$ Hz), 7.70 (ddd, $J = 49.3, 5.7, 3.2$ Hz), 7.54 (td, $J = 7.3, 0.8$ Hz), 6.33 (d, $J = 15.8$ Hz, 1H, 1'H), 5.84 (d, $J = 13.9$ Hz, 1H, 3'H), 5.78 (d, $J = 48.3$ Hz, 1H, 2'H), 5.38 (td, $J = 5.6, 2.3$ Hz, 1H 4'H), 4.75 – 4.56 (m, 2H, 5'H)

mix: ^{19}F NMR (470 MHz, Acetone- d_6) δ : -187.44 (dt, $J = 48.4, 15.9$ Hz, α), -191.64 (ddd, $J = 48.0, 20.3, 9.2$ Hz, SM), -201.06 (dt, $J = 50.7, 20.2$ Hz, β)

β : ^{19}F NMR (470 MHz, CDCl_3) δ : -201.35 (dt, $J = 50.2, 19.6$ Hz, β)

β : DART-MS m/z : 581.02026 $[\text{M} + \text{H}]^+$. ESI-MS m/z : 581.020996 $[\text{M} + \text{H}]^+$.

α : DART-MS m/z : 581.02196 $[\text{M} + \text{H}]^+$. ESI-MS m/z : 581.021179 $[\text{M} + \text{H}]^+$.

Trimethyl(octa-1,7-diyn-1-yl)silane

The reaction was conducted in oven-dried glassware under a N_2 atmosphere in a 100 mL two neck flask. The synthesis was conducted according to previously reported conditions with minor modifications^{156,157} and in good agreement with previously reported mass fragmentation patterns¹⁵⁸. To a solution of 1,7-octadiyne (5 g, 47.10 mmol, 6.17 mL) in THF (50.00 mL) at -78 °C, waiting 15 minutes to cool, nBuLi (2.5 M, 18.84 mL) was added slowly over 10 minutes. The mixture was allowed to stir for 20 min and then for 1 h at 0 °C. The solution was recooled to -78 °C over 15 minutes, TMSCl (5.12 g, 47.10 mmol, 5.98 mL) was added. The mixture was allowed to warm to ambient temperature and then stirred for O/N for 19 hours. The reaction was quenched using 5 mL of water and the organic solvent was removed in vacuo. The layers were washed with the addition of 3x50 mL of diethylether and 2x100 mL water and 100 mL brine followed by drying of the organic layer over NaSO_4 . The solvent was removed in vacuo; however, a stronger 21 mbar vacuum source was used in combined with single collection flask, thermometer, and single distillation column, as per previous procedures.

Practically, the reported 0.3 mbar with a Vigreux column is equivalent to my conditions, since a simple distillation was used but with higher pressure. ^1H -NMR analysis gave very disperse peaks with little resolution. Additionally, attempting to solve for the fractions using the terminal proton and TMS integral gave contradicting numbers. Hence, GC-MS was conducted, using the integral under each curve to solve for the

composition of each fraction. Overall, 186.4 mg of still material was recovered, suggesting 3.73 % of byproducts of heavy alkanes.

The MS spectra matches the expected mass for the SM–H fragment, product–Me fragment, and byproduct–Me fragment, indicating that TMS is in fact being incorporated. Using two major fractions were collected and analyzed for further synthesis efforts- the first fraction consisted of 2.50 g of material which was composed of 0.02 g SM, 1.82 g the target single-TMS product, and 0.65 g double-TMS non-reactive byproducts; and the second fraction consisted of 1.63 g of material which was composed of 0.00 g SM, 0.40 g target product, and 1.23 g byproducts. Thus, the overall yield of the reaction was 2.23 g or 26.49 % and 15.97 % byproducts.

EI-MS *m/z*: 105.1 [M – H][–]; 163.2 [M – CH₄][–]; 235.2 [M – CH₄][–]

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-C8-alkyl-Uridine

This screen of synthetic conditions was conducted according to previously reported conditions with minor modifications^{159,140,160} and in good agreement with previously reported spectra¹⁶¹. The reaction was conducted in oven-dried glassware under a N₂ atmosphere in a 10 mL Schlenk round bottom, 2'-Deoxy-2'-Fluoro-3,5-di-o-benzoyl-β-D-arabinofuranosyl-5-iodouridine (50.2 mg, 86.51 μmol), bis(triphenylphosphine) palladium(II) dichloride catalyst (6.07 mg, 8.65 μmol), and iodocopper (3.30 mg, 17.30 μmol) solids were dissolved in dry DMF (500 uL), followed by a 3 cycles of freeze-pump-thaw, over room temperature acetone. Separately prepared N,N-Diisopropylethylamine (DIPEA) (22.36 mg, 173.01 μmol, 29.58 uL) was freeze-pump-thawed three times and the stirred reaction mixture. After 10 min at room temperature, three times freeze-pump-thawed trimethyl(octa-1,7-diynyl)silane (44.03 mg, 173.01 μmol, 55.04 uL) linker and the reaction was stirred at room temperature for 5:45 hrs.

Upon addition of DMF the reaction mixture took on a burnt orange color; upon the addition of DIPEA the solution took on a brown color, which developed into a lime green with freeze-pump-thaw, ; with the addition of the linker the reaction took on a light cream color to which developed into a black color with mixing. The reaction progress was monitored by TLC, using with 1:1 hex:EtOAc, and developed using 10 % H₂SO₄ in EtOH staining with heating. The reaction was cooled and quenched with saturated aqueous NaHCO₃ (57.47 mL). Reaction catalyst was removed by filtering twice through celite, with ~30 mL 1:1 DCM:MeOH, in order to avoid leaching the Pd metal again from the dark grey celite. The sample was evaporated to a mixture of DMF and MeOH, silica gel was added to dry load a 3 cm x 18 cm column, with approximately 250 mg loading capacity. We found that a silica gel hex:EtOAc:DCM gradient of 200 mL hex + 1 % triethyl amine (TEA), 400 mL 4:1:0, 500 mL 1:1:0, 300 mL DCM to get complete adequate separation of the SM and the expected product. All final vials had a dark brown coloration, likely due to catalyst impurities.

TMS: ¹H NMR (500 MHz, CDCl₃) δ: 8.94 (s, 1H, N3H), 8.07 (dd, J = 25.4, 7.8 Hz, 4H, o-Bz), 7.78 (s, 1H, C6H), 7.62 (dt, J = 23.8, 7.4 Hz, 2H, p-Bz), 7.48 (dt, J = 15.8, 7.7 Hz, 4H, m-Bz), 6.32 (d, J = 20.8 Hz, 1H, 1'H), 5.63 (d, J = 18.0, 1H, 3'H), 5.34 (d, J_{HF} = 45.9, 1H, 2'H), 4.80 (ddd, J = 25.4, 12.1, 4.2 Hz, 2H, 5'H), 4.54 (m, 1H, 4'H), 2.33 (m, 2H, Alk3H), 2.21 (m, 2H, Alk6H), 1.59 (m, 4H, Alk4H & Alk5H), 0.14 (s, 9H, TMS). ¹⁹F NMR (470 MHz, CDCl₃) δ: -201.01 (dtd, J = 50.4, 20.7, 19.9, 2.2 Hz). ¹³C NMR (151 MHz, CDCl₃) δ: -201.01 (dtd, J = 50.4, 20.7, 19.9, 2.2 Hz). ¹³C NMR (151 MHz, o-d₆) δ: 166.77 (C=O), 166.06 (C=O), 160.54 (C4), 150.77 (C2), 146.19 (C6), 134.81 (Bz), 134.39 (p-Bz), 134.37 (p-Bz), 130.82 (o-Bz x2), 130.59 (o-Bz), 130.53 (o-Bz), 130.10 (Bz), 129.73 (m-Bz), 129.68 (m-Bz x2), 129.66 (m-Bz), 122.56 (C5), 100.9 (Alk C1), 94.98 (β²H), 93.62 (α²H), 85.55 (β¹H), 85.45 (α¹H), 81.52 (4'), 78.00 (α³H), 77.70 (β³H), 73.09 (Alk C2), 70.22 (Alk C7), 68.72 (Alk C8), 64.36 (α⁵H), 64.21 (β⁵H), 28.53 (Alk C4), 28.49 (Alk C5), 19.49 (Alk C3), 18.42 (Alk C6), 0.14 (TMS). ESI-MS *m/z*: 631.227417 [M + H]⁺.

H: ESI-MS *m/z*: 581.021179 [M + H]⁺.

CHAPTER FOUR

Results

DNA transcription into 2'-Deoxy-2'-Fluoro-Arabinose Nucleic Acids (FANA)

Nucleic acid aptamers are promising alternatives to antibodies for a wide array of diagnostic and therapeutic applications. However, state-of-the-art aptamers suffer from poor pharmacokinetics and diversity, limiting their affinity and specificity for many therapeutically relevant targets. Combining synthetic chemistry with modern molecular biology and polymer science, the synthesis of xeno nucleic acid (XNA) monomers and chemoenzymatic polymerization via engineered polymerase enzymes allow the production of nucleic acid drugs with superior resistance to endogenous nucleases.

Polymerase screen for FANA Transcription Activity

The screening primer extension assay demonstrated that Tgo polymerase, Terminator, and KOD XL polymerase produced full length FANA products; however, KOD XL did not seem to completely consume the starting DNA primer, suggesting limited activity or processivity. Deep Vent (exo-) polymerase seemed to also have limited activity as truncated DNA/FANA products were observed.

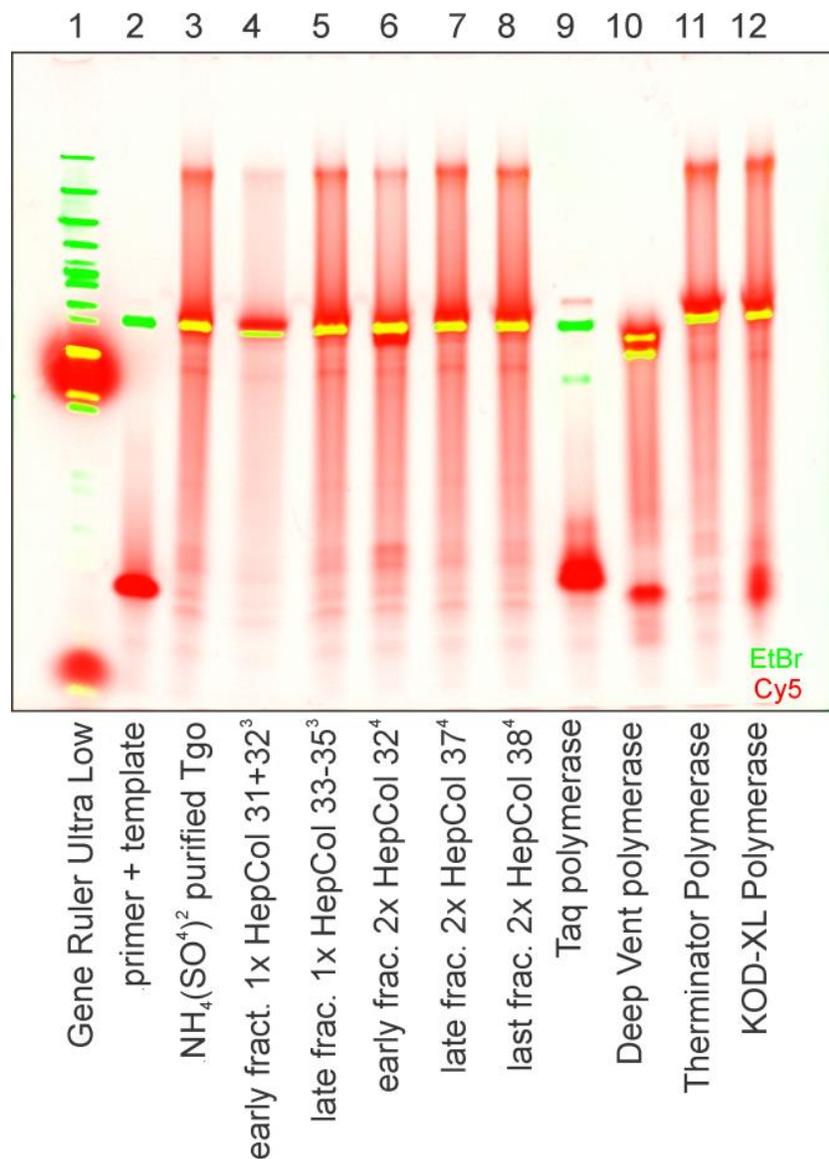


Figure 4. – DNA polymerase screen for DNA templated FANA transcription
Loaded with 15 μL samples at 0.091 μM in DI formamide, 1 hr at 180 V, 1:00 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Reaction conditions: 3:00 hr, non-diluted polymerases, 25 mM EDTA. Ethidium bromide (EtBr) selectively stains DNA T-Lib 78-mer, while staining FANA poorly. NT – no template, NP – no primer

The preferential staining of ethidium bromide stain for DNA over FANA was remarkably useful for being able to discriminate between starting material DNA and product FANA bands. However, the chosen dsDNA ladder was not informative, as in these denaturing conditions it seemed to decompose into several non-descript bands. Additionally, it was observed that the addition of the loading dye was deleterious to

obtaining an accurate signal, as Cy5 artefacts were observed within lane 1 designating the ladder, most likely caused by bromophenol blue (lower) and xylene cyanol (higher). However, this method alternatively indicated that lane 2, comprising of DNA library and the Cy5-RP without added polymerase, is a respectable relative size standard. This negative control consistently allows full-length FANA transcription products to be visible from the starting DNA library, as the 18 Da difference in FANA sugars across the length of the remaining 58 nts to be approximately 1.5 kDA (including the mass of 5' Cy5 tag) higher in mass, despite the resolution between these bands being inadequate for purification purposes.

Polymerase Dilution Series

The optimal purified polymerase concentration was determined for optimal XNA activity by performing a primer extension assay with FANA xNTPs similarly as to the procedure necessary for the polymerase screen. The optimal purified polymerase concentration necessary for DNA templated FANA transcription with Tgo was determined to 1:8 following recombinant expression and 2x Heparin column purification (Figure 4). Since Terminator and Deep Vent (exo-) were also observed to produce reasonable additionally determined 1:8 and at most 1:4 respectively (Appendix B & C). However, unlike the screening conditions, the Deep Vent polymerase seemingly was capable of FANA transcription even without dilute.

This informed further interrogation of the possibility for polymerase storage additives having an enhancing or debilitating effects on FANA transcription, since the screening was conducted on a gel additionally containing Tgo polymerase purification aliquots with variable salt and glycerol buffer compositions. This follow-up study suggests that not only does 50 % glycerol content within a polymerase storage buffer

allow -80 deg C storage, but that 50 % glycerol is capable of enhancing Tgo DNA polymerase FANA XPE activity (Appendix D). This data could explain why the Deep Vent (exo-) polymerase did not generate full length FANA product in the polymerase screen, yet in a gel containing only Deep Vent (exo-) polymerase consistently containing 50 % glycerol, full length transcription of FANA was observed.

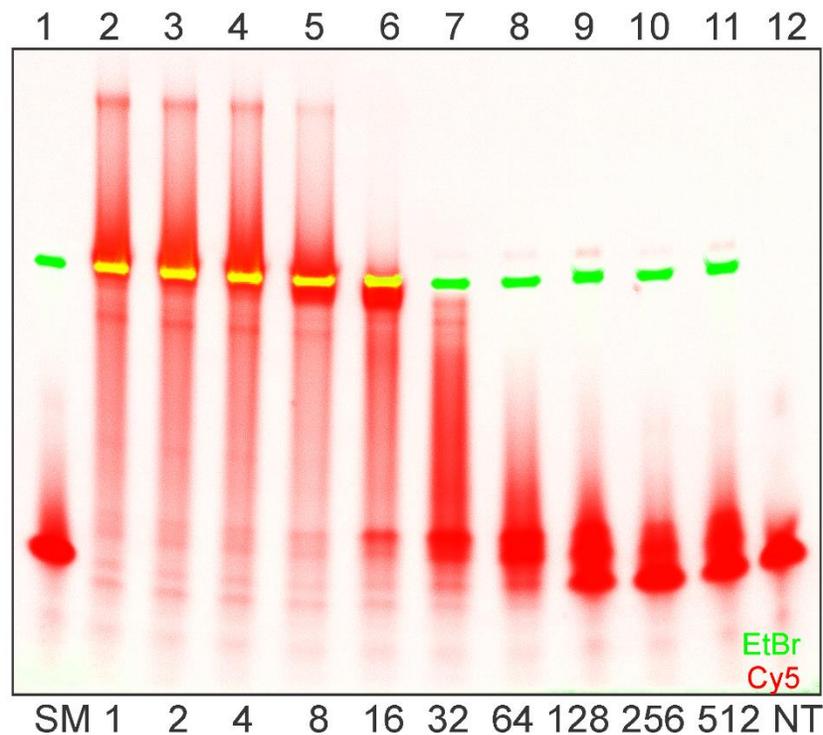


Figure 5. – Tgo dilution series indicating laddering due to polymerase concentration Loaded with 15 μ L samples at 0.091 μ M in DI formamide, 1 hr at 180 V, 1:00 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Reaction conditions: 3:00, 0.25 μ M Tgo (exo-) polymerase (optimum), 25 mM EDTA. EtBr selectively stains DNA T-Lib 78-mer, while staining FANA poorly. NT – no template, NP – no primer. Indicates that a 1:4 dilution is optimal, as 1:8 dilution seems to have increased incomplete products beneath overlapping DNA band (in yellow).

Tgo Timecourse Optimization

The optimal purified polymerase concentration The optimal purified polymerase concentration was determined for optimal XNA activity by performing a primer extension assay with FANA xNTPs similarly as to the procedure necessary for the polymerase screen. It was observed that the ideal FANA XPE reaction time for Tgo was

90 minutes; however, 120 minutes could likely lead to the generation of more product. A key utility of these findings were that the final band after 180 minutes was consistently aberrant (Figure 1). This band is thought to indicate the generation of overextension by-products. Additionally, the importance of EDTA concentration was found to be key in controlling the rate of FANA extension, as 0.25 mM EDTA was found complete extension after at least 150 minutes (data not reported).

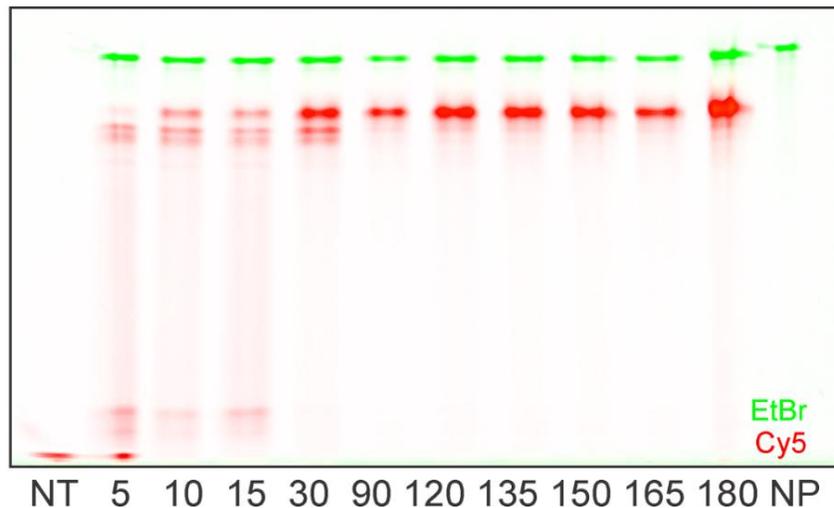


Figure 6. – Timecourse optimization of Tgo Cy5-primer extension into FANA library Loaded with 15 μ L samples at 0.091 μ M in DI formamide, 1 hr at 180 V, 1:30 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Optimum reaction conditions: 1:30 hr (2:00 high yield), 0.25 μ M Tgo (exo-) polymerase, 25 mM EDTA. EtBr selectively stains DNA T-Lib-XL 96-mer, while staining FANA poorly. NT – no template, NP – no primer.

Further work involves the recovery of ‘light’ DNA libraries from FANA libraries, following a similar screen for ‘reverse transcription-like’ activity. The goal of this work involved the implementation of a primer-based protocol for generating FANA and subsequently checking the fidelity of the ‘forward-’ and ‘reverse-translation’ process, through FANA.

DNA Polymerase Mediated FANA 'Reverse' Transcription into cDNA

It is key that less formamide is not used to resuspend extended FANA samples, as the downstream separation becomes more difficult with more concentrated samples. Particularly, it was found when XPE products were resuspended with 500 μL (1:2 μL :pmol) DI formamide the product was not found to UV shadow on an analytical gel equivalent to the previous polymerase studies; instead, it was observed that the majority of UV shadowing was present on a higher artefact band, due to lack of resuspension of ssFANA material and the consequent non-separation of the dsDNA/FANA hybrid helix.

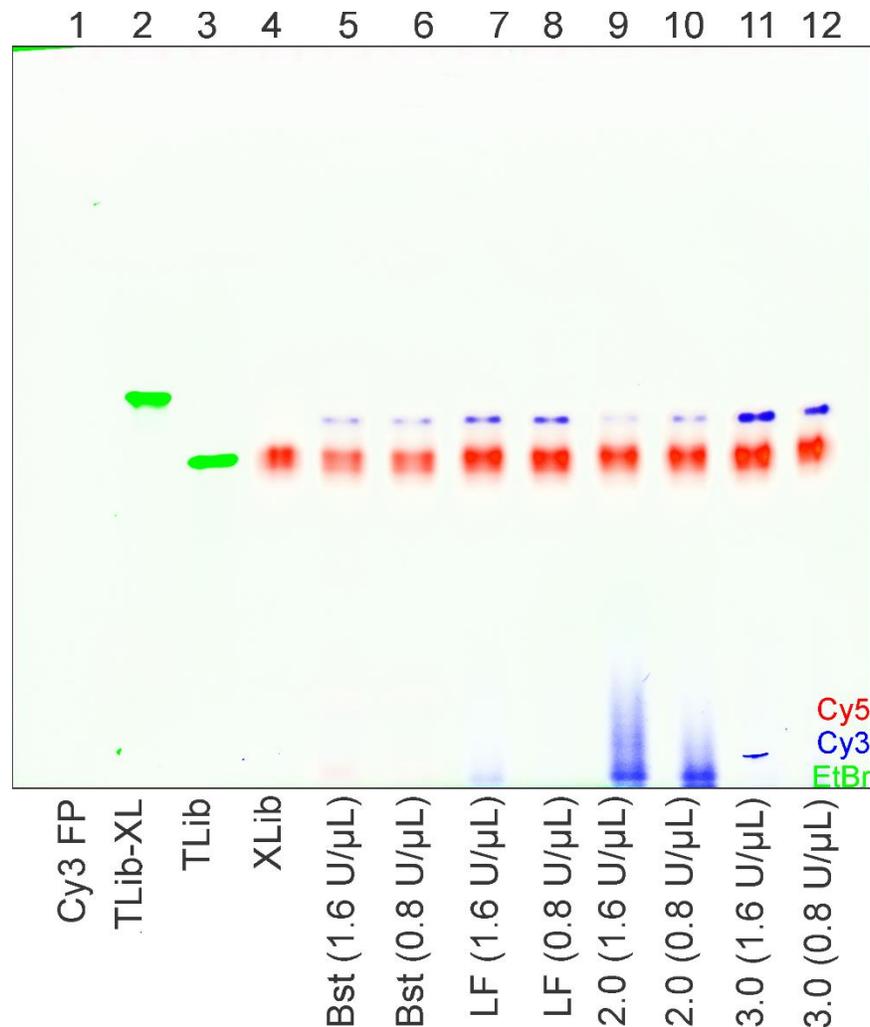


Figure 7. – Preliminary 'reverse' transcription of FANA library with Cy3-primer
Loaded with 15 μL samples at 0.091 μM in DI formamide, 1 hr at 180 V, 1 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Reaction conditions: 3:00 hr at 50 $^{\circ}\text{C}$, 2.3 mM MgCl_2 , with variable concentration polymerases. EtBr selectively stains DNA 100-mer, while staining FANA poorly.

Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)

The modular structure of nucleic acids allows the design of sequence-defined polymers capable of post-synthetically appending complex synthetic glycans, extending the potential catalytic geometry of aptamers. After breaking the emulsion, Click-PD is accomplished by incorporating modified NTPs displaying alkynyl handles facile for the Huisgen copper-catalyzed alkyne-azide cycloaddition (CuAAC) reaction to other azide-reagents⁸⁹. The optimum monosaccharide reagents for this click reaction have been previously determined to be acetyl-protected 1'-azidoethyl pyranoses. Overall, this selection of modified-aptamers (SELMA) technique produces glycan-sensitive affinity reagents as well as utilizing the library size of the nucleic acid sequence to evolve the optimum multivalent number and spacing for binding.

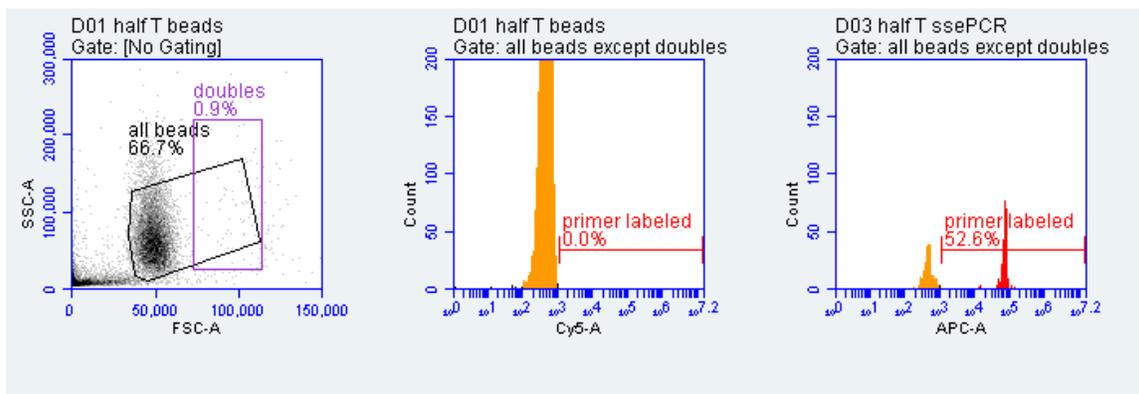


Figure 8. – Conjugation of Cy5-DNA FP to Dynabeads MyOne CA magnetic beads. As displayed on Accuri C6+ BD Biosciences. Size and shape of initial particles for flow cytometry (left); 640 nm laser excitation and detection of Cy5 fluorescence and bead autofluorescence on 670 nm LP filter, relative to count of beads before (middle) and after (right). These data indicates that roughly half of all beads were labeled with primers following ePCR, suggesting a even less than 0.75 pM DNA library must be used in order to reach the optimum value of $\sim 20\%$ ¹³⁷.

Thus, the preliminary work with DNA library functionalized microparticles allows access to the large library diversity and high-throughput ideal for these screening procedures. It was found that 99.96 % of beads had different fluorescence profile after functionalization, indicating the facile conjugation of amino-primers to these beads

(Appendix G). Further work was able identify the sensitivity of these beads to DNA, capable of interfacing with as little as 0.75 pmol of DNA in emulsion (Figure 8).

Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA

The extension of Click-PD technique should allow for individual XNA library members to be grown on the surface of the microparticle beads with our proposed two primer (T1-AM6-FP, amino-dPEG₅-azide + T1-TTB-RC) system . This could be accomplished by initially functionalizing the beads with both the amino-forward primer and an amino-dPEG₅-azide adapter, which is facile for the CuAAC yet utilizing orthogonal conjugation chemistry than necessary for functionalizing the bead. This could be used to generate an evolvable XNA biopolymer from the template DNA library member on each bead, via a bridge amplification primer extension reaction. However, this requires the generation of non-commercially available xNTPs.

Our approach utilizes a Sonogashira cross-coupling reaction to install a flexible alkyne to the major groove of 2'-deoxy-2'-fluoro-arabinose uracil base. Uracil was chosen as the modified nucleobase, since this base was previously demonstrated to be minimally stabilized by the elongated alkyne handle which not only has less steric effects than a direct alkyne but also is less likely to destabilize a duplex, overall allowing the highest efficiency for click-conjugation¹⁶²⁻¹⁶⁵. By incorporated recent advances in nucleic acid synthesis, one-pot nucleobase activation and sugar glycosylation^{145,154,166-169} or two-step high selectivity^{143,147,170} can be achieved, additionally we propose to utilize bis-oxybenzyl phosphoramidite synthesis in order to generate gram scale HPLC-free purification of 5' triphosphates¹⁷¹ (Figure 9).

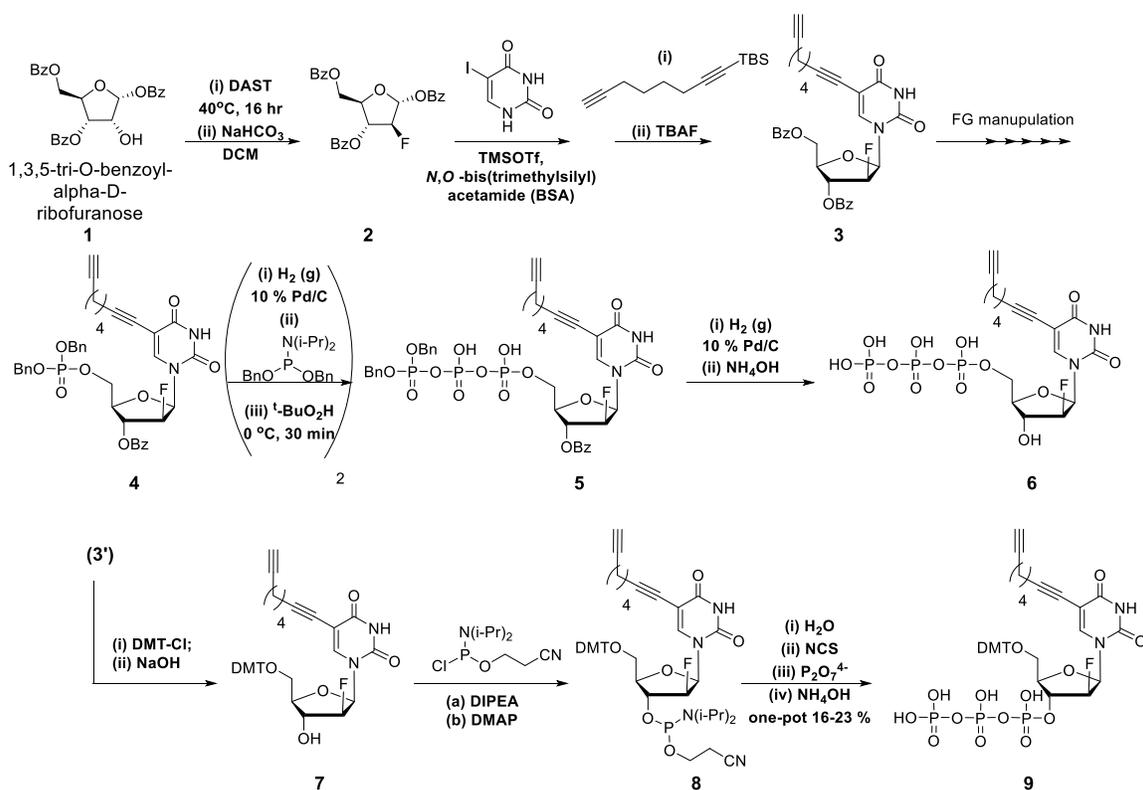


Figure 9. – Proposed scheme for synthesis of FANA C8-alkyl-uridine triphosphates. These monomers could be useful for either polymerase incorporation into any sequence chain (6) or solid-phase synthesis of a target sequence, from the deprotected complete nucleoside (9).

5-Iodouracil

This reaction was previously reported in 80 % isolated yield (Figure 10). We reported a similar yield of 81.03 % from two crops of crystals- the first of which had a slight orange hue, likely due to light cerium metal impurities. However, the presence of these minor impurities was not found to affect the next step of the reaction. While, this reaction step was found to be unnecessary as the product was found to be available commercially. The utility of this reaction lies within the mechanism, wherein the strong oxidative properties of cerium ammonium nitrate (CAN) in situ generate element iodine that causes nuclear iodination of electron rich aromatic compounds^{172,173}, allowing even protected uracil nucleosides to halogenated with this procedure¹⁷⁴.

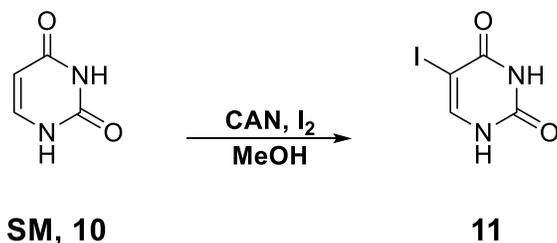


Figure 10. – Nucleobase uracil iodine halogenation mediated by CAN oxidation
 The ideal reaction conditions were with 0.6 mol equiv I₂, 0.5 mol equiv CAN are reported to functionalize protected and deprotected DNA & RNA nucleosides ¹⁴¹.

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose

Indeed, the majority of these reactions were found to finish the production of products after 16 hours with heating to 40 °C. Despite this, while instrumental failure was the main cause, it was found that the reaction progresses readily at room temperature, even potentially suppressing the generation of multi-fluorinated products. These multi-fluorinated products were prominently observed with ¹⁹F-NMR (data not shown) with a larger 10+ gram scale reaction; these byproducts are the result of DAST's well documented ability to insert germinal fluorines at the position of aldehydes, here the benzoyl (Bz) protecting groups (PGs). In an attempt to mediate the potentially hazardous properties of DAST reagent, an alternate deoxyfluorinating agent 1.5 mol equiv of XtalFluor-E with 1.5 mol equiv DBU was briefly tested ¹⁷⁵; however, the resulting reaction produced significantly more byproducts (data not shown). These byproducts are a non-issue considering the mechanism necessary for the originally planned S_N1 Friedel-Crafts Lewis acid catalyzed Vorbrüggen reaction mechanism (Appendix H) ^{168,167,152}; however, their identification and removal became more crucial to minimizing future byproducts with the final bromination intermediate mechanism.

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- β -D-Arabinofuranosyl-5-Iodouridine

We report successful synthesis of this molecule using the previously reportedly non-productive Vorbrüggen S_N1 reaction mechanism (Table 2, entry 1) ¹⁴⁷; however,

considering the target product is the β anomer, the effective isolated yield is only 11.9 %. Hence, a bromination 1'OBz PG SN₁ substitution allowed the conversion of **2** into **2'**, potentially due to the anomeric effect and bromine's relatively large atomic radius. This essentially sets up Br⁻ adducts as a good leaving group allowing amiable SN₂ substitutions to the 2'fluoro-arabino sugar 1' position, forcing β stereochemistry of the nucleobase. A small study was conducted in which the previously reported solvent in combination with the 1-pot Vorbrüggen sialylation agent BSA^{147,145} (Table 2, entry 2), the solvent and weaker sialylation agent as hexamethyldisilazane (HMDS)¹⁴⁷ (Table 2, entry 3), and the solvent used in a more recent report for a similar transformation alongside the strong BSA sialylation agent¹⁷⁰ (Table 2, entry 4).

Table 2. – Summary of different FANA glycoside precursor glycosylation conditions. The differing effects of glycosylation solvents and base sialylation agents were explored. The superior reaction (in italics) was selected based on increased β production, relieving purification efforts.

Donor	Mech	Scale	Glyco Temp	Solvent	Sialy Agent	Steps	Time	Yield	$\alpha:\beta$
Sugar cation	SN ₁	3.89 g	60 °C	ACN	BSA	1	15 hr	36.2 %	2.05
<i>α-Bromo sugar</i>	SN ₂	1.47 g	22 °C	1:4.8 ACN/ DCM	BSA	2	45 hr	67.7 % ^a	4.06 ⁻¹
<i>α-Bromo sugar</i>	SN ₂	1.46 g	22 °C	1:4.8 ACN/ DCM	HMDS	3	60 hr	43.2 % ^b	4.09 ⁻¹
<i>α-Bromo sugar</i>	SN ₂	1.86 g	22 °C	<i>THF</i> ²	BSA	2	45 hr	40.8 % ^a	13.76 ⁻¹
a – Calculated recovery yield									
b – Isolated yield, including decomposition									

These reactions were easily tracked by the generation of a dark reddish brown color. Previous failed reactions (data not shown), conducted with large excess TMSOTf to be indicative of the formation of N3-nucleoside byproducts, due to the sialylation of the formed products. Our hypothesis for this mechanism is that the nucleobase N3 is likely being the nitrogen with highest electron density (Appendix H).

Heteronuclear ^{19}F -NMR was found to be invaluable in the monitoring of this reaction, as it allowed accurate quantification of the α : β ratio of products, even in a crude reaction mixture. This was key for screening the selectivity of these reactions, as it was found that Entries 2 & 4 underwent side-reactions during the process of drying in vacuo, most likely due to the presence of residual BSA reagent. Future work could easily avoid this issue, with the addition of another quenching step, allowing the removal of BSA in solution and/or low temperature/high vacuum removal of solvent, instead of at 50 °C.

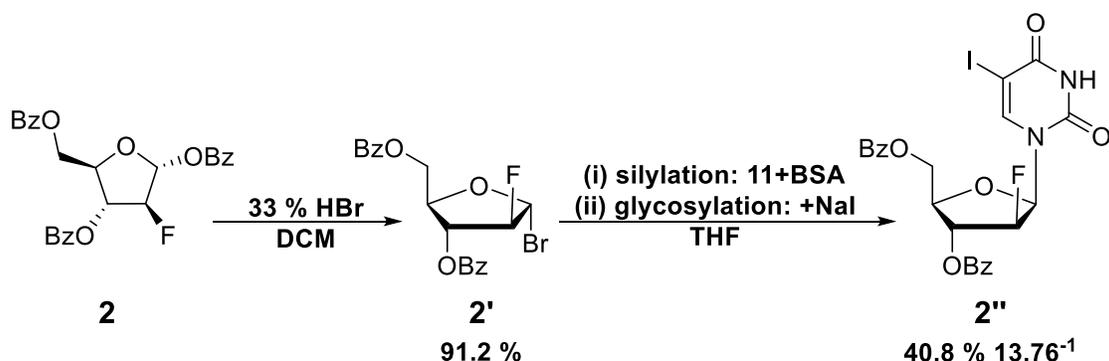


Figure 11. – Bz protected 2-F-arabinose bromine halogenation in glacial acetic acid
The ideal reaction conditions were with 0.6 mol equiv I₂, 0.5 mol equiv CAN are reported to functionalize protected and deprotected DNA & RNA nucleosides ¹⁴¹.

It was not surprising that Entry 4 was superior reaction since its solvent minimized stabilization of a potential $\text{S}_{\text{N}}1$ mechanism charged intermediates, producing an effective isolated yield of 38.0 % with 93.1 % purity (Figure 11). This reaction was selected despite the fact that Entry 2 generated a higher effective isolated yield of 54.3 % β anomer but with 80.2 % purity, requiring more precise chromatographic separation of diaestereomer.

This is non-trivial at the preparative scale required for this total synthesis; recrystallization was attempted on these products utilizing previously reported recrystallization conditions 1:4 DCM:hex followed by *i*PrOH ¹⁴⁷. This purification method was successful in producing a reasonably pure crystal, capable of verifying the

absolute orientation of the product via XRD (Figure 12, Appendix I); however, this methodology was found to be highly skill-dependent and with low reproducibility, recovering 2.91^{-1} to 21.60^{-1} α : β purities, as monitored by ^{19}F -NMR.

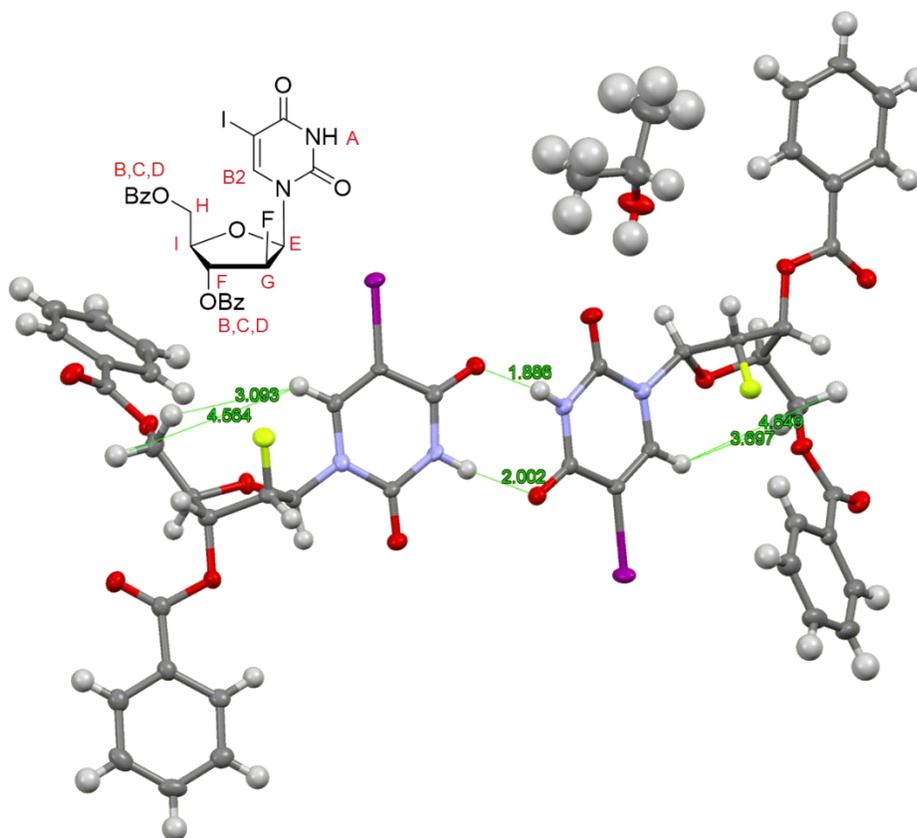


Figure 12. – X-ray diffractometric structure of FANA glycoside precursor **2''**

Reference structure indicates NMR unique hydrogens. This data was generated following iPrOH recrystallization. The full crystallographic data is available in Appendix H. Atomic distance measurements indicate that crystal self-paired during recrystallization process and NOE between C6H and 5'H is possible, as distance is $<5\text{\AA}$.

What was surprising was that the use of a weaker silylation agent did not drastically change the selectivity of the $\text{S}_{\text{N}}2$ reactions, as previous reports have found that the strength of silylation agent is capable of increasing β anomer for Vorbrüggen $\text{S}_{\text{N}}1$ reaction mechanism,¹⁷⁶; however, the use of HMDS was found to drastically reduce yield of this reaction. This was complicated by the fact that the BSA is reported to have very high boiling point, boiling at $71\text{-}73\text{ }^{\circ}\text{C}$ at 35 mmHg . However, the TMS resonance

forms were calculated to have a lower bp, an estimated 119.6 ± 23.0 °C at 760 mmHg, via ChemSpider. Experimentally, it was found that a BSA solution boils at approximately 35 °C at 50 mbar, and thus easily boil over a water bath of 50 °C. Hence, with the drying of HMDS the bump tramp displayed white crystals with evaporation. Additionally, it was found that this silylation procedure can actually proceed quite rapidly, within 15 minutes at 85 °C, while BSA works similarly within 2 minutes.

Trimethyl(octa-1,7-diyn-1-yl)silane

Conducting vacuum distillation using a simple distillation column, it was found that the **SM, 12** distills around 53 °C, placing the vacuum strength at an apparent strength of around 71 mbar (assuming a ACD/Labs Percepta Platform - PhysChem Module on ChemSpider.com predicted 35.8 kJ/mol), according to the Clausius-Capryon equations. The single-TMS protected product **13** was found to distill around 98 °C. This predicts an estimated bp of 188 °C at 1 atm for the product, using the 42.2 kJ/mol estimated by ChemSpider; or this estimates 37.1 kJ/mol if using 203.5 °C, as calculated using ChemSpider, a better match to the experimental data. Additionally, it was found that in an oil bath, the double-TMS protect byproduct **13'** distills around 135 °C; following the ChemSpider estimated bp of 264.3 °C at 1 atm, this estimates an enthalpy of vaporization of 37.5 kJ/mol- this is likely within error since it does not seem reasonable for the enthalpy of vaporization to increase so little with a second TMS group, while it increases with a first. Hence, using this logic, using 38.4 kJ/mol (assuming 1.7 increase in magnitude per TMS group), the estimated bp is 137 °C, suggesting a good match. If the ChemSpider predicted enthalpy of 48.2 kJ/mol is used, then the vacuum bp should be around 258 °C or the 1 atm bp should be around 213 °C.

Table 3. – TMS Linker distillation with regards to different data sources. Data collected from mixed product NMR and GC-MS analysis (Figure 13), alongside calculated physical properties and experimental elution temperatures.

	Starting Material	Single-TMS Product	Double-TMS Byproduct
NMR ratio	0 ^c	0.700	0.300
area ratio	0.0086	0.7298	0.2616
bp, 760 torr	135 °C ^a	203.5 °C ^b	264.5 °C ^b
bp, 18 torr	22 °C ^c	32 °C ^a	63 °C ^a
a – experimental data			
b – estimated using physical properties, ACD/Labs Percepta			
c - assumed			

In an attempt to clear up the inconsistency in these data, the vacuum line was tested in the lab. The house vac gave a reading of 140-160 torr (~200 mbar), while the vacuum pump gave 16 torr (21 mbar). This gave ~ 15 kJ/mol lower values, suggesting that the boiling points observed would be higher than that observed. This gives credence to the strength of the vacuum being in fact smaller than 71 mbar, a highly efficient system; since the temperature inside each flask is likely lower than the bath temperature.

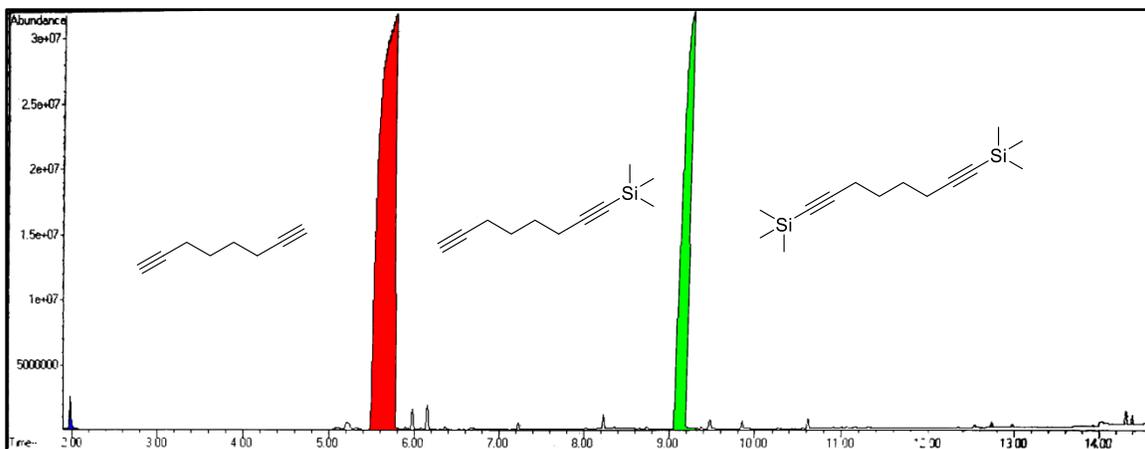


Figure 13. – Fraction distillate GC-MS trace containing target product **13** TMS-Linker. The integral of these traces was used to find the mass ratio of the **SM**, **12** (in blue), target single-TMS protected product **13** (in red), and non-reactive double-TMS protected byproduct **13'** (in green).

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-C8-alkyl-Uridine

Initial attempts to conduct this reaction resulted in complex product mixtures that required extensive purification and contained many reaction impurities. Although the possibility of metal scavengers or an additional post purification reverse-phase activated charcoal filtration step was considered, in order to remove the presumed Pd metal contamination of the final products, despite silica gel purification, the inability to adequately detect the Pd metal (Appendix O) and the proposed additional 5+ steps of the proposed reaction pathway (Figure 9) suggested the celite filtration alone was sufficient. Additionally, while freeze-pump-thaw techniques were utilized in order to generate 'utmost' oxygen free samples for the screening, in order to minimize Glaser linker-homocoupling byproducts¹⁷⁷, in our experience this step was likely excessive. It was found that the largest affect on the reaction mixture purity/control of byproducts was by the age of the Pd catalyst. Upon switching to a newly purchased catalyst, regardless of type (data not shown), the resulting TLC clearly indicated fewer impurities.

In response to the widely reported sugar and linker protections presented in the literature, a screen of the optimal reaction conditions was initiated in order to identify the affect on reactivity that these factors could have on FANA based XNA (Table 4). It was identified that, despite literature reports that 10 equivalents non-TMS-protected linker is enough to drive this reaction, increasing the equivalents of linker conserves one synthetic and purification step; however, with TMS protection only 2 equivalents liker is necessary in order to produce a similar result. Preliminary work (data not shown)has also suggested that the use of deprotected sugars in order to simply the purification and limit the amount of unintended byproducts, considering that the Sonogashira chemistry is insensitive to alcohols^{178,179}.

Table 4. – Summary of different FANA -Iodouracil Pd cross-coupling conditions. The differing effects of sugar and linker protection were explored. The superior reaction (in italics) was selected based high purity and overall step yield, suggesting scalability.

SM purity	$R_1 = R_2$	R_3	equiv linker	scale SM	Step Yield	Purity
<i>only β</i>	<i>Bz</i>	<i>H</i>	<i>100</i>	<i>50 mg 90 μmol</i>	<i>39.31 %^b</i>	unknown byproducts
only β	Bz	H	10	600 mg 1 mmol	6.83 % ^b	<i>pure</i>
only β	Bz	TMS	2	75 mg 130 μ mol	23.69 % ^b	deprotection byproducts
<i>only β</i>	<i>Bz</i>	<i>TMS</i>	<i>2</i>	<i>50 mg 100 μmol</i>	<i>40.3 %^b</i>	<i>pure</i>
27.62 % β	H	H	2	60 mg 165 μ mol	87.29 % ^b	SM byproducts
only β	H	TMS	4	200 mg 450 μ mol	–	–

a – Calculated recovery yield
b – Isolated yield, including decomposition

Likely, the additional step of adding acid in order to generate the TMS-linker was found to potentially deprotect the nucleoside during cross-coupling (Table 4, entry 3). The workflow trade-off in deprotecting subsequent nucleoside one-pot immediately following cross-coupling & filtration of the reaction mixture was not explored. However, further PG manipulation of a deprotected 5' would produce a stoichiometrically sensitive 1' OH.

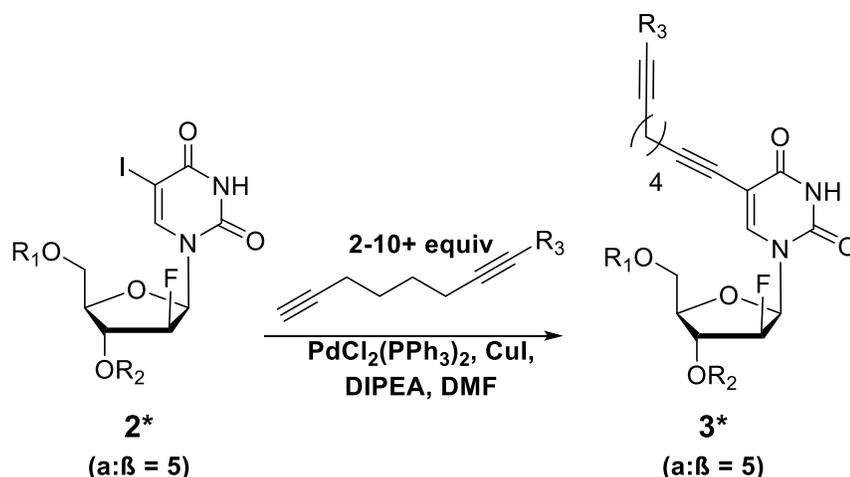


Figure 14. – Variable FANA-Iodouracil Pd cross-coupling to a variable alkyne linker.

Two ideal reaction conditions were identified (Table 4, entry 1 & 4). Any preference between the two is very likely one of scale and overall cost. For this reason, both have been identified as being ideal candidates, as use of the unprotected linker simplifies the total synthesis procedure.

CHAPTER FIVE

Discussion

DNA transcription into 2'-Deoxy-2'-Fluoro-Arabinose Nucleic Acids (FANA)

Nucleic acid aptamers are promising alternatives to antibodies for a wide array of diagnostic and therapeutic applications. However, state-of-the-art aptamers suffer from poor pharmacokinetics and diversity, limiting their affinity and specificity for many therapeutically relevant targets. Combining synthetic chemistry with modern molecular biology and polymer science, the synthesis of xeno nucleic acid (XNA) monomers and chemoenzymatic polymerization via engineered polymerase enzymes allow the production of nucleic acid drugs with superior resistance to endogenous nucleases.

Polymerase screen for FANA Transcription Activity

Indeed, while Tgo polymerase was identified to be a useful polymerase for the transcription of DNA into FANA, it should be emphasized that the focus of this work was in no manner to identify the optimal FANA polymerase. This is outside the scope of this work; however, the previous work to identify functional B family Tgo polymerase¹¹⁰ and the evolution of A family SFM4-3 from the, Klenow-like Stoffel fragment of Taq, DNA polymerase¹⁸⁰⁻¹⁸². With this in retrospect, the activity of Klenow (-exo) fragment polymerase compared to Klenow and Klentaq could have additionally been screened, in order to determine if it was in fact the 3'→5' proof reading activity of Taq polymerase that made FANA transcription inactive or if this was due to a larger structural morphology affecting substrate promiscuity.

DNA Polymerase Mediated FANA 'Reverse' Transcription into cDNA

There exist many polymerases which are Mn^{+2} dependent for unnatural promiscuity and activity, such as for incorporating unnatural nucleobases or sugars. While the current scope of this research project does not make any efforts to directly mutate polymerases on a grand scale, in order to overcome these limitations. It was found that the mechanism for Mn^{+2} to impart RNA template reverse transcription-like properties in A family DNA polymerases, could be related to the ability for DNA polymerase to accept unnatural substrate. This has been reported to be caused by the divalent metal ion catalyzed hydrolysis of RNA into shorter sequences¹⁸³, this seems a plausible explanation for the observations of reverse transcription activity in –exo B family DNA polymerases superior or similar to that of obligate AMV RNA polymerase with RNA templates up to 125 nts¹⁸⁴. Regardless, we have identified potential new mutations to induce in Tgo or Tgo-like Pfu B family polymerases, in order to better harness this mechanism.

Click-Particle Display (Click-PD) Selection with Modified Aptamers (SELMA)

The bead to terminal oxygen distance for the pacification PEG was found to be 45.0 Å; while it is unsure whether this value is critical for appropriate screening, as to the authors' knowledge this has not been screened in previous literature, and as it stands the pacification would be approximately equal in length to the other reagents with 9 PEG units at 34.4 Å. Ultimately, this presents further potential for process optimization and the screening of further conditions.

Additionally, it is not known whether each of the proposed screening methods is ideal for only glycoprotein or glyconucleic acids. It would be particularly surprising to discover that each method additionally is compatible with one another. For instance,

while our proposed increased salt concentration gradient theoretically would produce an increasing entropic penalty, similar to increased temperature conditions, it may be likely that this could cause serum protein aggregation. Thus, indirectly forcing selection for/against large protein aggregates, or even the target of interest. Regardless, the ability to screen for targets in human serum, as well as at physiological conditions of 37 °C, would certainly be a boon for selecting for interaction of ‘true’ binding partners, in addition to providing a better in vitro model for in vivo therapeutic conditions.

Synthesis of Saccharide Conjugatable Click Facile Base-Modified FANA

Although the synthesis target was not reached during the primary author’s duration of time spent on this project, the proposed synthesis route would allow the generation of the reagents necessary to allow the screen of novel modified FANA aptamers. The produced C8-alkyl-uridine FANA triphosphates could be incorporated by an engineered Tgo (exo-) DNA polymerase to allow systematic introduction of alkynyl conjugation handles into a DNA-templated FANA polymer. Subsequent conjugation with azido-modified glycans via the CuAAC click reaction would generate sequence controlled nucleic acid-carbohydrate hybrid molecules amendable for directed evolution. Additionally, this synthesis places the phosphoramidite within reach¹⁸⁵, allowing solid-phase synthesis of any generated ‘hit’ FANA polymer for validation.

Final Thoughts

The great variability in the ability for XNA to take and sustain a great variety of different structures could very well someday allow individual companies to develop their own strains of living cells which utilize a proprietary genetic polymer system to store or even, at very least, translate biological information- orthogonally to all other living or

currently known beings. While this may seem excess today, as this technology further develops along its experience curve and the barriers to entry is further reduced, this could be a conceivable measure in order to control in vivo intellectual property and/or provide ultimate biosafety, more so than metabolic limitations ¹⁸⁶. Already there are mounting concerns for the need of biodefense in this “post-genomic age” ¹⁸⁷; however, this only seems to highlight a major limitation in this technology- the inability for it to be rapidly sequenced de novo nor a priori without knowledge of the chemical makeup of an XNA and the appropriate enzyme biomachinery necessary for transcription into DNA for traditional sequencing.

REFERENCES CITED

1. Gates, Z. P. *et al.* Xenoprotein engineering via synthetic libraries. *Proc. Natl. Acad. Sci.* **115**, E5298–E5306 (2018).
2. Kruger, K. *et al.* Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31**, 147–157 (1982).
3. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.* **15**, (2017).
4. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).
5. Arnott, S. Historical article: DNA polymorphism and the early history of the double helix. *Trends Biochem. Sci.* **31**, 349–354 (2006).
6. Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **16**, 907–916 (1963).
7. Frank-Kamenetskii, M. D. & Mirkin, S. M. Triplex DNA Structures. *Annu. Rev. Biochem.* **64**, 65–95 (1995).
8. Bentin, T., Hansen, G. I. & Nielsen, P. E. Measurement of PNA Binding to Double-Stranded DNA. in *Peptide Nucleic Acids* **208**, 91–110 (Humana Press, 2002).
9. Dorn, S. *et al.* Side chain modified peptide nucleic acids (PNA) for knock-down of six3 in medaka embryos. *BMC Biotechnol.* **12**, 50 (2012).
10. Chen, D., Meena, Sharma, S. K. & McLaughlin, L. W. Formation and Stability of a Janus-Wedge Type of DNA Triplex. *J. Am. Chem. Soc.* **126**, 70–71 (2004).
11. Chen, H., Meena & McLaughlin, L. W. A Janus-Wedge DNA Triplex with A-W1-T and G-W2-C Base Triplets. *J. Am. Chem. Soc.* **130**, 13190–13191 (2008).
12. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13**, 770–780 (2012).
13. Bugaut, A. & Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* **40**, 4727–4741 (2012).
14. Pauling, L. & Corey, R. B. Structure of the Nucleic Acids. *Nature* **171**, 346–347 (1953).
15. Fraser, R. D. B. The Structure of Deoxyribose Nucleic Acid. (1953).
16. Chaput, J. C. & Switzer, C. A DNA pentaplex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci.* **96**, 10614–10619 (1999).
17. Henry, A. A. & Romesberg, F. E. Beyond A, C, G and T: augmenting nature's alphabet. *Curr. Opin. Chem. Biol.* **7**, 727–733 (2003).

18. Wojciechowski, F. & Leumann, C. J. Alternative DNA base-pairs: from efforts to expand the genetic code to potential material applications. *Chem. Soc. Rev.* **40**, 5669–5679 (2011).
19. Pinheiro, V. B. & Holliger, P. The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* **16**, 245–252 (2012).
20. Hunziker, J. Warum Pentose- und nicht Hexose-Nucleinsaeuren. *HELVETICA Chem. ACTA* **76**, 259–352 (1993).
21. Stevens, E. S. Vacuum UV Circular Dichroism of Polysaccharides. *Photochem. Photobiol.* **44**, 287–293 (1986).
22. Morris, E. R., Rees, D. A. & Thom, D. Characterization of polysaccharide structure and interactions by circular dichroism: order–disorder transition in the calcium alginate system. *J Chem Soc Chem Commun* 245–246 (1973). doi:10.1039/C39730000245
23. Bystrický, S., Szu, S. C., Gotoh, M. & Kováč, P. Circular dichroism of the O-specific polysaccharide of *Vibrio cholerae* O1 and some related derivatives. *Carbohydr. Res.* **270**, 115–122 (1995).
24. Anosova, I. *et al.* The structural diversity of artificial genetic polymers. *Nucleic Acids Res.* **44**, 1007–1021 (2016).
25. Arnott, S. & Hukins, D. W. L. The dimensions and shapes of the furanose rings in nucleic acids. *Biochem. J.* **130**, 453–465 (1972).
26. Hirao, I. *et al.* An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **3**, 729–735 (2006).
27. Betz, K. *et al.* Structural Insights into DNA Replication without Hydrogen Bonds. *J. Am. Chem. Soc.* **135**, 18637–18643 (2013).
28. Brotschi, C. & Leumann, C. J. DNA with Hydrophobic Base Substitutes: A Stable, Zipperlike Recognition Motif Based On Interstrand-Stacking Interactions. *Angew. Chem. Int. Ed.* **42**, 1655–1658 (2003).
29. Bommarito, S., Peyret, N. & Jr, J. S. Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* **28**, 1929–1934 (2000).
30. Guckian, K. M. *et al.* Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA. *J. Am. Chem. Soc.* **122**, 2213–2222 (2000).
31. Brotschi, C., Mathis, G. & Leumann, C. J. Bipyridyl- and Biphenyl-DNA: A Recognition Motif Based on Interstrand Aromatic Stacking. *Chem. – Eur. J.* **11**, 1911–1923 (2005).
32. Williams, A. A. *et al.* The Impact of Sugar Pucker on Base Pair and Mismatch Stability. *Biochemistry* **48**, 11994 (2009).
33. Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure. **284**, 7 (1999).

34. Chan, J. H. P., Lim, S. & Wong, W. S. F. Antisense oligonucleotides: from design to therapeutic application. *Clin. Exp. Pharmacol. Physiol.* **33**, 533–540 (2006).
35. Yamamoto, T. *et al.* Cholesterol-lowering Action of BNA-based Antisense Oligonucleotides Targeting PCSK9 in Atherogenic Diet-induced Hypercholesterolemic Mice. *Mol. Ther. - Nucleic Acids* **1**, e22 (2012).
36. Abdur Rahman, S. M. *et al.* Design, Synthesis, and Properties of 2',4'-BNA NC: A Bridged Nucleic Acid Analogue. *J. Am. Chem. Soc.* **130**, 4886–4896 (2008).
37. Culbertson, M. C. *et al.* Evaluating TNA stability under simulated physiological conditions. *Bioorg. Med. Chem. Lett.* **26**, 2418–2421 (2016).
38. Noronha, A. M. *et al.* Synthesis and Biophysical Properties of Arabinonucleic Acids (ANA): Circular Dichroic Spectra, Melting Temperatures, and Ribonuclease H Susceptibility of ANA·RNA Hybrid Duplexes. *Biochemistry* **39**, 7050–7062 (2000).
39. Cummins, L. L. *et al.* Characterization of fully 2'-modified oligoribonucleotide hetero- and homoduplex hybridization and nuclease sensitivity. *Nucleic Acids Res.* **23**, 2019–2024 (1995).
40. Cromwell, C. R. *et al.* Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. *Nat. Commun.* **9**, 1448 (2018).
41. Reitsma, S., Slaaf, D. W., Vink, H., van Zandvoort, M. A. M. J. & oude Egbrink, M. G. A. The endothelial glycocalyx: composition, functions, and visualization. *Pflüg. Arch. - Eur. J. Physiol.* **454**, 345–359 (2007).
42. Yi, W. *et al.* Phosphofructokinase 1 Glycosylation Regulates Cell Growth and Metabolism. *Science* **337**, 975–980 (2012).
43. Brown, J. M. *et al.* A sulfated carbohydrate epitope inhibits axon regeneration after injury. *Proc. Natl. Acad. Sci.* **109**, 4768–4773 (2012).
44. Huang, M. L., Smith, R. A. A., Trieger, G. W. & Godula, K. Glycocalyx Remodeling with Proteoglycan Mimetics Promotes Neural Specification in Embryonic Stem Cells. *J. Am. Chem. Soc.* **136**, 10565–10568 (2014).
45. Krauss, I. J. *et al.* Fully Synthetic Carbohydrate HIV Antigens Designed on the Logic of the 2G12 Antibody. *J. Am. Chem. Soc.* **129**, 11042–11044 (2007).
46. Wang, L.-X. Synthetic carbohydrate antigens for HIV vaccine design. *Curr. Opin. Chem. Biol.* **17**, 997–1005 (2013).
47. Leonard, K., Spellman, W., Harris, R. J., Thomas, N. & GregorySII, J. Assignment of Intrachain Disulfide Bonds and Characterization of Potential Glycosylation Sites of the Type 1 Recombinant Human Immunodeficiency Virus Envelope Glycoprotein (gp120) Expressed in Chinese Hamster Ovary Cells. **11** (1990).
48. Spiess, M. The asialoglycoprotein receptor: a model for endocytic transport receptors. *Biochemistry* **29**, 10009–10018 (1990).
49. Hamer, C. J. A. van den, Morell, A. G., Scheinberg, I. H., Hickman, J. & Ashwell, G. Physical and Chemical Studies on Ceruloplasmin IX. The role of galactosyl

- residues in the clearance of ceruloplasmin from the circulation. *J. Biol. Chem.* **245**, 4397–4402 (1970).
50. Grewal, P. K. Chapter Thirteen - The Ashwell–Morell Receptor. in *Methods in Enzymology* (ed. Fukuda, M.) **479**, 223–241 (Academic Press, 2010).
 51. Lundquist, J. J. & Toone, E. J. The Cluster Glycoside Effect. *Chem. Rev.* **102**, 555–578 (2002).
 52. Kiessling, L. L. & Pohl, N. L. Strength in numbers: non-natural polyvalent carbohydrate derivatives. *Chem. Biol.* **3**, 71–77 (1996).
 53. Schwefel, D. *et al.* Structural Basis of Multivalent Binding to Wheat Germ Agglutinin. *J. Am. Chem. Soc.* **132**, 8704–8719 (2010).
 54. Ponader, D., Wojcik, F., Beceren-Braun, F., Dervede, J. & Hartmann, L. Sequence-Defined Glycopolymer Segments Presenting Mannose: Synthesis and Lectin Binding Affinity. *Biomacromolecules* **13**, 1845–1852 (2012).
 55. Temme, J. S., MacPherson, I. S., DeCoursey, J. F. & Krauss, I. J. High Temperature SELMA: Evolution of DNA-Supported Oligomannose Clusters Which Are Tightly Recognized by HIV bnAb 2G12. *J. Am. Chem. Soc.* **136**, 1726–1729 (2014).
 56. Avci, F. G., Altinisik, F. E., Ulu, D. V., Olmez, E. O. & Akbulut, B. S. An evolutionarily conserved allosteric site modulates beta-lactamase activity. *J. Enzyme Inhib. Med. Chem.* **31**, 33–40 (2016).
 57. Zondlo, N. J. Aromatic-Proline Interactions: Electronically Tunable CH/ π Interactions. *Acc. Chem. Res.* **46**, 1039–1049 (2013).
 58. Fernández-Alonso, M. del C., Cañada, F. J., Jiménez-Barbero, J. & Cuevas, G. Molecular Recognition of Saccharides by Proteins. Insights on the Origin of the Carbohydrate–Aromatic Interactions. *J. Am. Chem. Soc.* **127**, 7379–7386 (2005).
 59. Espinosa, E., Molins, E. & Lecomte, C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **285**, 170–173 (1998).
 60. Ponader, D. *et al.* Carbohydrate-Lectin Recognition of Sequence-Defined Heteromultivalent Glycooligomers. *J. Am. Chem. Soc.* **136**, 2008–2016 (2014).
 61. Tully, S. E. *et al.* A Chondroitin Sulfate Small Molecule that Stimulates Neuronal Growth. *J. Am. Chem. Soc.* **126**, 7736–7737 (2004).
 62. Gama, C. I. *et al.* Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2**, 467–473 (2006).
 63. Nair, J. K. *et al.* Multivalent N-Acetylgalactosamine-Conjugated siRNA Localizes in Hepatocytes and Elicits Robust RNAi-Mediated Gene Silencing. *J. Am. Chem. Soc.* **136**, 16958–16961 (2014).
 64. Matsuda, S. *et al.* siRNA Conjugates Carrying Sequentially Assembled Trivalent N-Acetylgalactosamine Linked Through Nucleosides Elicit Robust Gene Silencing *In Vivo* in Hepatocytes. *ACS Chem. Biol.* **10**, 1181–1187 (2015).

65. Janas, M. M. *et al.* Selection of GalNAc-conjugated siRNAs with limited off-target-driven rat hepatotoxicity. *Nat. Commun.* **9**, 723 (2018).
66. Moulton, H. M. & Moulton, J. D. Morpholinos and their peptide conjugates: Therapeutic promise and challenge for Duchenne muscular dystrophy. *Biochim. Biophys. Acta BBA - Biomembr.* **1798**, 2296–2303 (2010).
67. Jirka, S. M. G. *et al.* Cyclic Peptides to Improve Delivery and Exon Skipping of Antisense Oligonucleotides in a Mouse Model for Duchenne Muscular Dystrophy. *Mol. Ther.* **26**, 132–147 (2018).
68. Keefe, A. D., Pai, S. & Ellington, A. Aptamers as therapeutics. *Nat. Rev. Drug Discov.* **9**, 537–550 (2010).
69. Ng, E. W. M. *et al.* Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat. Rev. Drug Discov.* **5**, 123–132 (2006).
70. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
71. Cohen, M. *et al.* Synthetic Mucus Nanobarriers for Identification of Glycan-Dependent Primary Influenza A Infection Inhibitors. *ACS Cent. Sci.* **2**, 710–714 (2016).
72. Kiessling, L. L. & Grim, J. C. Glycopolymer probes of signal transduction. *Chem. Soc. Rev.* **42**, 4476–4491 (2013).
73. Pednekar, P. P., Jadhav, K. R. & Kadam, V. J. Aptamer-dendrimer bioconjugate: a nanotool for therapeutics, diagnosis, and imaging. *Expert Opin. Drug Deliv.* **9**, 1273–1288 (2012).
74. Zheng, Y., Gilgenast, M. J., Hauc, S. & Chatterjee, A. Capturing Post-Translational Modification-Triggered Protein–Protein Interactions Using Dual Noncanonical Amino Acid Mutagenesis. *ACS Chem. Biol.* **13**, 1137–1141 (2018).
75. Li, X. & Liu, C. C. Biological Applications of Expanded Genetic Codes. *ChemBioChem* **15**, 2335–2341 (2014).
76. Collins, J. M., Porter, K. A., Singh, S. K. & Vanier, G. S. High-Efficiency Solid Phase Peptide Synthesis (HE -SPPS). *Org. Lett.* **16**, 940–943 (2014).
77. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
78. Fang, G.-M. *et al.* Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angew. Chem. Int. Ed.* **50**, 7645–7649 (2011).
79. Vinogradov, A. A., Evans, E. D. & Pentelute, B. L. Total synthesis and biochemical characterization of mirror image barnase. *Chem Sci* **6**, 2997–3002 (2015).
80. Palomo, J. M. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC Adv* **4**, 32658–32672 (2014).

81. Chisholm, T. S., Clayton, D., Dowman, L. J., Sayers, J. & Payne, R. J. Native Chemical Ligation–Photodesulfurization in Flow. *J. Am. Chem. Soc.* **140**, 9020–9024 (2018).
82. Bannas, P., Hambach, J. & Koch-Nolte, F. Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics. *Front. Immunol.* **8**, (2017).
83. Li, S. *et al.* Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7**, 301–311 (2005).
84. Wu, P. *et al.* Efficiency and Fidelity in a Click-Chemistry Route to Triazole Dendrimers by the Copper(I)-Catalyzed Ligation of Azides and Alkynes. *Angew. Chem. Int. Ed.* **43**, 3928–3932 (2004).
85. Kwok, A., Eggimann, G. A., Reymond, J.-L., Darbre, T. & Hollfelder, F. Peptide Dendrimer/Lipid Hybrid Systems Are Efficient DNA Transfection Reagents: Structure–Activity Relationships Highlight the Role of Charge Distribution Across Dendrimer Generations. *ACS Nano* **7**, 4668–4682 (2013).
86. Liu, J., Liu, M., Zheng, B., Yao, Z. & Xia, J. Affinity Enhancement by Ligand Clustering Effect Inspired by Peptide Dendrimers–Shank PDZ Proteins Interactions. *PLOS ONE* **11**, e0149580 (2016).
87. *Solid-phase synthesis: a practical guide.* (Marcel Dekker, 2000).
88. Brown, T. & Brown, T. Solid-phase oligonucleotide synthesis. *ATDBio - Nucleic Acids Book* Available at: <https://www.atdbio.com/content/17/Solid-phase-oligonucleotide-synthesis>. (Accessed: 4th May 2019)
89. Gordon, C. K. L. *et al.* Click-PD: A Quantitative Method for Base-Modified Aptamer Discovery. *bioRxiv* 626572 (2019). doi:10.1101/626572
90. Wang, H. *et al.* Aptamer-Dendrimer Bioconjugates for Targeted Delivery of miR-34a Expressing Plasmid and Antitumor Effects in Non-Small Cell Lung Cancer Cells. *PLOS ONE* **10**, e0139136 (2015).
91. Grubbs, R. B. & Grubbs, R. H. 50th Anniversary Perspective : Living Polymerization—Emphasizing the Molecule in Macromolecules. *Macromolecules* **50**, 6979–6997 (2017).
92. Wojcik, F., Mosca, S. & Hartmann, L. Solid-Phase Synthesis of Asymmetrically Branched Sequence-Defined Poly/Oligo(amidoamines). *J. Org. Chem.* **77**, 4226–4234 (2012).
93. Thomas, B., Fiore, M., Daskhan, G. C., Spinelli, N. & Renaudet, O. A multi-ligation strategy for the synthesis of heterofunctionalized glycosylated scaffolds. *Chem. Commun.* **51**, 5436–5439 (2015).
94. Yang, C., Flynn, J. P. & Niu, J. Facile Synthesis of Sequence-Regulated Synthetic Polymers Using Orthogonal SuFEx and CuAAC Click Reactions. *Angew. Chem. Int. Ed.* **57**, 16194–16199 (2018).

95. Gerke, C. *et al.* Sequence-Controlled Glycopolymers via Step-Growth Polymerization of Precision Glycomacromolecules for Lectin Receptor Clustering. *Biomacromolecules* **18**, 787–796 (2017).
96. Polymer Molecular Weight Distribution and Definitions of MW Averages: Technical Overview. (2015).
97. Uchio, T., Baudyš, M., Liu, F., Song, S. C. & Kim, S. W. Site-specific insulin conjugates with enhanced stability and extended action profile. *Adv. Drug Deliv. Rev.* **35**, 289–306 (1999).
98. Guan, X. *et al.* Chemically Precise Glycoengineering Improves Human Insulin. *ACS Chem. Biol.* **13**, 73–81 (2018).
99. Liang, R. *et al.* Parallel synthesis and screening of a solid phase carbohydrate library. *Science* **274**, 1520–1522 (1996).
100. Stadlmann, J. *et al.* Comparative glycoproteomics of stem cells identifies new players in ricin toxicity. *Nature* **549**, 538–542 (2017).
101. Murata, K. & Wolf, M. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1862**, 324–334 (2018).
102. Gruene, T. *et al.* Rapid Structure Determination of Microcrystalline Molecular Compounds Using Electron Diffraction. *Angew. Chem. Int. Ed.* **57**, 16313–16317 (2018).
103. Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife* **2**, e01345 (2013).
104. Nannenga, B. L. & Gonen, T. MicroED opens a new era for biological structure determination. *Curr. Opin. Struct. Biol.* **40**, 128–135 (2016).
105. Nannenga, B. L. & Gonen, T. The cryo-EM method microcrystal electron diffraction (MicroED). *Nat. Methods* **16**, 369 (2019).
106. Rodriguez, J. A. *et al.* Structure of the toxic core of α -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
107. Sawaya, M. R. *et al.* Ab initio structure determination from prion nanocrystals at atomic resolution by MicroED. *Proc. Natl. Acad. Sci.* **113**, 11232–11236 (2016).
108. Jones, C. G. *et al.* The CryoEM Method MicroED as a Powerful Tool for Small Molecule Structure Determination. *ACS Cent. Sci.* **4**, 1587–1592 (2018).
109. Nikooban, A., Dunn, M. R. & Chaput, J. C. Engineered Polymerases with Altered Substrate Specificity: Expression and Purification: Engineered Polymerases: Expression and Purification. in *Current Protocols in Nucleic Acid Chemistry* (eds. Egli, M., Herdewijn, P., Matusda, A. & Sanghvi, Y. S.) 4.75.1-4.75.20 (John Wiley & Sons, Inc., 2017). doi:10.1002/cpnc.33
110. Wang, Y., Ngor, A. K., Nikooban, A. & Chaput, J. C. Evolution of a General RNA-Cleaving FANA Enzyme. *Nat. Commun.* **9**, (2018).

111. Miroshnichenko, M. L. *et al.* *Thermococcus gorgonarius* sp. nov. and *Thermococcus pacificus* sp. nov.: heterotrophic extremely thermophilic archaea from New Zealand submarine hot vents. *Int. J. Syst. Evol. Microbiol.* **48**, 23–29 (1998).
112. Jozwiakowski, S. K. & Connolly, B. A. A Modified Family-B Archaeal DNA Polymerase with Reverse Transcriptase Activity. *ChemBioChem* **12**, 35–37 (2011).
113. Whitfield, C. J., Turley, A. T., Tuite, E. M., Connolly, B. A. & Pike, A. R. Enzymatic Method for the Synthesis of Long DNA Sequences with Multiple Repeat Units. *Angew. Chem. Int. Ed.* **54**, 8971–8974 (2015).
114. Cozens, C., Pinheiro, V. B., Vaisman, A., Woodgate, R. & Holliger, P. A short adaptive path from DNA to RNA polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8067–8072 (2012).
115. Evans, S. J. *et al.* Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* **28**, 1059–1066 (2000).
116. Hansen, C. J., Wu, L., Fox, J. D., Arezi, B. & Hogrefe, H. H. Engineered split in Pfu DNA polymerase fingers domain improves incorporation of nucleotide γ -phosphate derivative. *Nucleic Acids Res.* **39**, 1801–1810 (2011).
117. Jozwiakowski, S. K., Keith, B. J., Gilroy, L., Doherty, A. J. & Connolly, B. A. An archaeal family-B DNA polymerase variant able to replicate past DNA damage: occurrence of replicative and translesion synthesis polymerases within the B family. *Nucleic Acids Res.* **42**, 9949–9963 (2014).
118. Firbank, S. J., Wardle, J., Heslop, P., Lewis, R. J. & Connolly, B. A. Uracil Recognition in Archaeal DNA Polymerases Captured by X-ray Crystallography. *J. Mol. Biol.* **381**, 529–539 (2008).
119. C M Joyce & Steitz, and T. A. Function and Structure Relationships in DNA Polymerases. *Annu. Rev. Biochem.* **63**, 777–822 (1994).
120. Derbyshire, V., Pinsonneault, J. K. & Joyce, C. M. Structure-function analysis of 3' \rightarrow 5'-exonuclease of DNA polymerases. in *Methods in Enzymology* **262**, 363–385 (Academic Press, 1995).
121. Hopfner, K.-P. *et al.* Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl. Acad. Sci.* **96**, 3600–3605 (1999).
122. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005).
123. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423 (1995).
124. Tgo DNA Polymerase: From *Thermococcus gorgonarius* Deoxynucleoside-triphosphate: DNA deoxynucleotidyltransferase, EC 2.7.7.7.

125. Kong, H., Kucera, R. B. & Jack, W. E. Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J. Biol. Chem.* **268**, 1965–1975 (1993).
126. Southworth, M. W. *et al.* Cloning of thermostable DNA polymerases from hyperthermophilic marine Archaea with emphasis on *Thermococcus* sp. 9 degrees N-7 and mutations affecting 3'-5' exonuclease activity. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 5281–5285 (1996).
127. Gardner, A. F. & Jack, W. E. Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and Taq DNA polymerases. *Nucleic Acids Res.* **30**, 605–613 (2002).
128. Nishioka, M. *et al.* Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J. Biotechnol.* **88**, 141–149 (2001).
129. Sawai, H. *et al.* Expansion of structural and functional diversities of DNA using new 5-substituted deoxyuridine derivatives by PCR with superthermophilic KOD Dash DNA polymerase. *Chem. Commun.* **0**, 2604–2605 (2001).
130. Sawai, H., Ozaki-Nakamura, A., Mine, M. & Ozaki, H. Synthesis of New Modified DNAs by Hyperthermophilic DNA Polymerase: Substrate and Template Specificity of Functionalized Thymidine Analogues Bearing an sp³-Hybridized Carbon at the C5 α -Position for Several DNA Polymerases. *Bioconjug. Chem.* **13**, 309–316 (2002).
131. Hwang, G. T. & Romesberg, F. E. Unnatural Substrate Repertoire of A, B, and X Family DNA Polymerases. *J. Am. Chem. Soc.* **130**, 14872–14882 (2008).
132. Malyshev, D. A. *et al.* Solution Structure, Mechanism of Replication, and Optimization of an Unnatural Base Pair. *Chem. – Eur. J.* **16**, 12650–12659 (2010).
133. Brown, J. A. & Suo, Z. Unlocking the Sugar “Steric Gate” of DNA Polymerases. *Biochemistry* **50**, 1135–1142 (2011).
134. Chim, N., Jackson, L. N., Trinh, A. M. & Chaput, J. C. Crystal structures of DNA polymerase I capture novel intermediates in the DNA synthesis pathway. *eLife* **7**, e40444 (2018).
135. Jackson, L. N., Chim, N., Shi, C. & Chaput, J. C. Crystal structures of a natural DNA polymerase that functions as an XNA reverse transcriptase. **11** (2019).
136. Aliotta, J. M. *et al.* Thermostable Bst DNA polymerase I lacks a 3' \rightarrow 5' proofreading exonuclease activity. *Genet. Anal. Biomol. Eng.* **12**, 185–195 (1996).
137. Wang, J. *et al.* Particle Display: A Quantitative Screening Method for Generating High-Affinity Aptamers. *Angew. Chem.* **126**, 4896–4901 (2014).
138. Wang, J. *et al.* Multiparameter Particle Display (MPPD): A Quantitative Screening Method for the Discovery of Highly Specific Aptamers. *Angew. Chem. Int. Ed.* **56**, 744–747 (2017).

139. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).
140. Gramlich, P. M. E., Wirges, C. T., Gierlich, J. & Carell, T. Synthesis of Modified DNA by PCR with Alkyne-Bearing Purines Followed by a Click Reaction. *Org. Lett.* **10**, 249–251 (2008).
141. Asakura, J. & Robins, M. J. Cerium (IV)-mediated halogenation at C-5 of uracil derivatives. *J. Org. Chem.* **55**, 4928–4933 (1990).
142. Župančić, N. *et al.* Synthesis and Biological Activity of Reversed Pyrimidine Nucleosides. *Croat. Chem. Acta* **88**, 43–52 (2015).
143. Wilds, C. J. & Damha, M. J. 2'-Deoxy-2'-fluoro- β -D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies. *Nucleic Acids Res.* **28**, 3625–3635 (2000).
144. Alauddin, M. M. *et al.* Stereospecific fluorination of 1,3,5-tri-O-benzoyl- α -D-ribofuranose-2-sulfonate esters: preparation of a versatile intermediate for synthesis of 2'-[18F]-fluoro-arabinonucleosides. *J. Fluor. Chem.* **106**, 87–91 (2000).
145. Sau, S. P., Fahmi, N. E., Liao, J.-Yu., Bala, S. & Chaput, J. C. A Scalable Synthesis of α -L-Threose Nucleic Acid Monomers. *J. Org. Chem.* **81**, 2302–2307 (2016).
146. Mei, H. *et al.* Synthesis and polymerase activity of a fluorescent cytidine TNA triphosphate analogue. *Nucleic Acids Res.* **45**, 5629–5638 (2017).
147. Tann, C. H., Brodfuehrer, P. R., Brundidge, S. P., Sapino, C. & Howell, H. G. Fluorocarbohydrates in synthesis. An efficient synthesis of 1-(2-deoxy-2-fluoro- β -D-arabinofuranosyl)-5-iodouracil (β -FIAU) and 1-(2-deoxy-2-fluoro- β -D-arabinofuranosyl)thymine (β -FMAU). *J. Org. Chem.* **50**, 3644–3647 (1985).
148. Derudas, M. *et al.* Evaluation of Novel Phosphoramidate ProTides of the 2'-Fluoro Derivatives of a Potent Anti-Varicella Zoster Virus Bicyclic Nucleoside Analogue. *Antivir. Chem. Chemother.* **21**, 15–31 (2010).
149. Paulsen, H. Cyclic Acyloxonium Ions in Carbohydrate Chemistry. in *Advances in Carbohydrate Chemistry and Biochemistry* **26**, 127–195 (Elsevier, 1971).
150. Mohamed, S., He, Q. Q., Lepage, R. J., Krenske, E. H. & Ferro, V. Glycosylations of Simple Acceptors with 2- O -Acyl L -Idose or L -Iduronic Acid Donors Reveal Only a Minor Role for Neighbouring-Group Participation: Glycosylations of Simple Acceptors with 2- O -Acyl L -Idose or L -Iduronic Acid Donors Reveal Only a Minor Role for Neighbouring-Group Participation. *Eur. J. Org. Chem.* **2018**, 2214–2227 (2018).
151. Gangangari, K. K., Humm, J. L., Larson, S. M., Vara, N. & Pillarsetty, K. TMSOTf assisted synthesis of 2'-deoxy-2'-[18F] fluoro- β -D-arabinofuranosylcytosine ([18F] FAC). *PLOS ONE* **8** (2018).
152. Simons, C. *Nucleoside Mimetics: Their Chemistry and Biological Properties*. (CRC Press, 2000).

153. Wright, J. Arthur., Wilson, D. Patrick. & Fox, J. J. Nucleosides. LXIV. Fluoro sugar analogs of arabinosyl- and xylosylcytosines. *J. Med. Chem.* **13**, 269–272 (1970).
154. Ritzmann, G., Klein, R. S., Hollenberg, D. H. & Fox, J. J. Nucleosides LXXXIX. Synthesis of 1-(2-chloro-2-deoxy- α - and - β -D-arabinofuranosyl)cytosines. *Carbohydr. Res.* **39**, 227–236 (1975).
155. Reichman, U., Watanabe, K. A. & Fox, J. J. A practical synthesis of 2-deoxy-2-fluoro-d-arabinofuranose derivatives. *Carbohydr. Res.* **42**, 233–240 (1975).
156. Hari, D. P. & Waser, J. Copper-Catalyzed Oxy-Alkynylation of Diazo Compounds with Hypervalent Iodine Reagents. *J. Am. Chem. Soc.* **138**, 2190–2193 (2016).
157. Hellbach, B., Gleiter, R. & Rominger, F. Cyclic Dienes by Alkyne Metathesis. *Synthesis* **2003**, 2535–2541 (2003).
158. Rontani, J.-F. & Aubert, C. Hydrogen and trimethylsilyl transfers during EI mass spectral fragmentation of hydroxycarboxylic and oxocarboxylic acid trimethylsilyl derivatives. *J. Am. Soc. Mass Spectrom.* **19**, 66–75 (2008).
159. Gierlich, J., Burley, G. A., Gramlich, P. M. E., Hammond, D. M. & Carell, T. Click Chemistry as a Reliable Method for the High-Density Postsynthetic Functionalization of Alkyne-Modified DNA. *Org. Lett.* **8**, 3639–3642 (2006).
160. Brunner, K. *et al.* Cell-Penetrating and Neurotargeting Dendritic siRNA Nanostructures. *Angew. Chem. Int. Ed.* **54**, 1946–1949 (2015).
161. Garg, N. K., Woodroffe, C. C., Lacenere, C. J., Quake, S. R. & Stoltz, B. M. A ligand-free solid-supported system for Sonogashira couplings: applications in nucleoside chemistry. *Chem. Commun.* 4551 (2005). doi:10.1039/b505737j
162. El-Sagheer, A. H. & Brown, T. Click chemistry with DNA. *Chem. Soc. Rev.* **39**, 1388 (2010).
163. Seela, F. & Sirivolu, V. R. DNA Containing Side Chains with Terminal Triple Bonds: Base-Pair Stability and Functionalization of Alkynylated Pyrimidines and 7-Deazapurines. *Chem. Biodivers.* **3**, 509–514 (2006).
164. Seela, F. & Sirivolu, V. R. Nucleosides And Oligonucleotides With Diynyl Side Chains: The Huisgen-Sharpless Cycloaddition “Click Reaction” Performed On Dna And Their Constituents. *Nucleosides Nucleotides Nucleic Acids* **26**, 597–601 (2007).
165. Seela, F. & Sirivolu, V. R. Nucleosides and Oligonucleotides with Diynyl Side Chains: Base Pairing and Functionalization of 2'-Deoxyuridine Derivatives by the Copper(I)-Catalyzed Alkyne–Azide ‘Click’ Cycloaddition. *Helv. Chim. Acta* **90**, 535–552 (2007).
166. Vorbrüggen, H. & Bennua, B. New simplified nucleoside synthesis. *Tetrahedron Lett.* **19**, 1339–1342 (1978).
167. Vorbrüggen, H. & Höfle, G. Nucleoside Syntheses, XXIII) On the Mechanism of Nucleoside Synthesis. *Chem. Ber.* **114**, 1256–1268 (1981).

168. Vorbrüggen, H., Krolkiewicz, K. & Bennua, B. Nucleoside syntheses, XXII Nucleoside synthesis with trimethylsilyl triflate and perchlorate as catalysts. *Chem. Ber.* **114**, 1234–1255 (1981).
169. Mukaiyama, T., Hirano, N., Nishida, M. & Uchiro, H. Catalytic Stereoselective Synthesis of Pyrimidine 2-Deoxyribonucleosides. 2 (1995).
170. Sivets, G. G. Syntheses of 2'-deoxy-2'-fluoro- β -d-arabinofuranosyl purine nucleosides via selective glycosylation reactions of potassium salts of purine derivatives with the glycosyl bromide. *Tetrahedron Lett.* **57**, 268–271 (2016).
171. Sau, S. P. & Chaput, J. C. A Gram-Scale HPLC-Free Synthesis of TNA Triphosphates Using an Iterative Phosphorylation Strategy. *Org. Lett.* **19**, 4379–4382 (2017).
172. Das, B., Krishnaiah, M., Venkateswarlu, K. & Reddy, V. S. A mild and simple regioselective iodination of activated aromatics with iodine and catalytic ceric ammonium nitrate. *Tetrahedron Lett.* **48**, 81–83 (2007).
173. Galvão, T. L. P., Rocha, I. M., Ribeiro da Silva, M. D. M. C. & Ribeiro da Silva, M. A. V. Is Uracil Aromatic? The Enthalpies of Hydrogenation in the Gaseous and Crystalline Phases, and in Aqueous Solution, as Tools to Obtain an Answer. *J. Phys. Chem. A* **117**, 5826–5836 (2013).
174. Asakura, J. & Robins, M. J. Cerium(iv) catalyzed iodination at C5 of uracil nucleosides. *Tetrahedron Lett.* **29**, 2855–2858 (1988).
175. L'Heureux, A. *et al.* Aminodifluorosulfonium Salts: Selective Fluorination Reagents with Enhanced Thermal Stability and Ease of Handling [†], [‡]. *J. Org. Chem.* **75**, 3401–3411 (2010).
176. Henschke, J. P. *et al.* A Stereoselective Process for the Manufacture of a 2'-Deoxy- β -d-Ribonucleoside Using the Vorbrüggen Glycosylation. *Org. Process Res. Dev.* **17**, 1419–1429 (2013).
177. Elangovan, A., Wang, Y.-H. & Ho, T.-I. Sonogashira Coupling Reaction with Diminished Homocoupling. *Org Lett* **5**, 4 (2003).
178. Chinchilla, R. & Nájera, C. The Sonogashira Reaction: A Booming Methodology in Synthetic Organic Chemistry [†]. *Chem. Rev.* **107**, 874–922 (2007).
179. Johansson Seechurn, C. C. C., Kitching, M. O., Colacot, T. J. & Snieckus, V. Palladium-Catalyzed Cross-Coupling: A Historical Contextual Perspective to the 2010 Nobel Prize. *Angew. Chem. Int. Ed.* **51**, 5062–5085 (2012).
180. Thirunavukarasu, D., Chen, T., Liu, Z., Hongdilokkul, N. & Romesberg, F. E. Selection of 2'-Fluoro-Modified Aptamers with Optimized Properties. *J. Am. Chem. Soc.* **139**, 2892–2895 (2017).
181. Chen, T. *et al.* Evolution of thermophilic DNA polymerases for the recognition and amplification of C2'-modified DNA. *Nat. Chem.* **8**, 556–562 (2016).

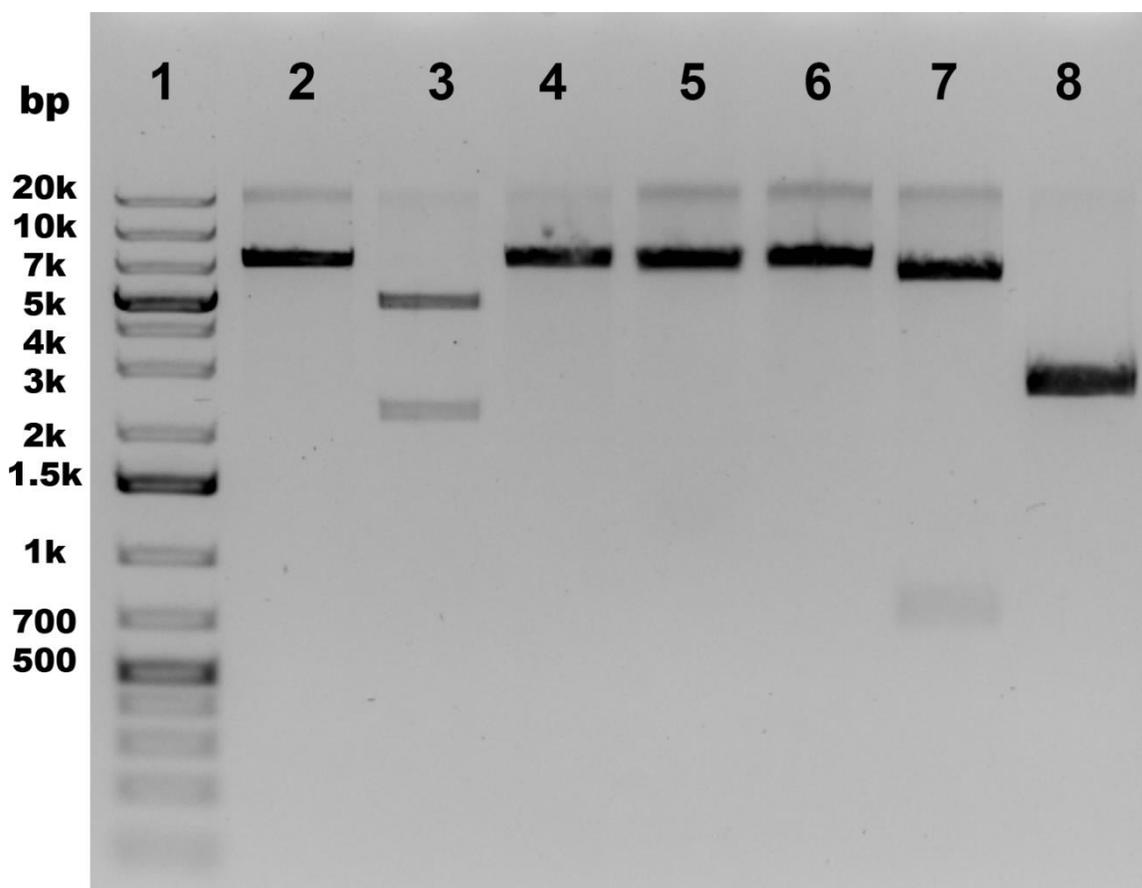
182. Fa, M., Radeghieri, A., Henry, A. A. & Romesberg, F. E. Expanding the Substrate Repertoire of a DNA Polymerase by Directed Evolution. *J. Am. Chem. Soc.* **126**, 1748–1754 (2004).
183. Myers, T. W. & Gelfand, D. H. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* **30**, 7661–7666 (1991).
184. Shi, C., Shen, X., Niu, S. & Ma, C. Innate Reverse Transcriptase Activity of DNA Polymerase for Isothermal RNA Direct Detection. *J. Am. Chem. Soc.* **137**, 13804–13806 (2015).
185. Sau, S. P. & Chaput, J. C. A one-pot synthesis of α -1 -threofuranosyl nucleoside triphosphates (tNTPs). *Bioorg. Med. Chem. Lett.* **26**, 3271–3273 (2016).
186. Schmidt, M. Xenobiology: A new form of life as the ultimate biosafety tool. *BioEssays* **32**, 322–331 (2010).
187. National Academies of Sciences, E. *Biodefense in the Age of Synthetic Biology*. (2018). doi:10.17226/24890

APPENDICES

APPENDIX A

Tgo Plasmid RD QC Agarose Gel

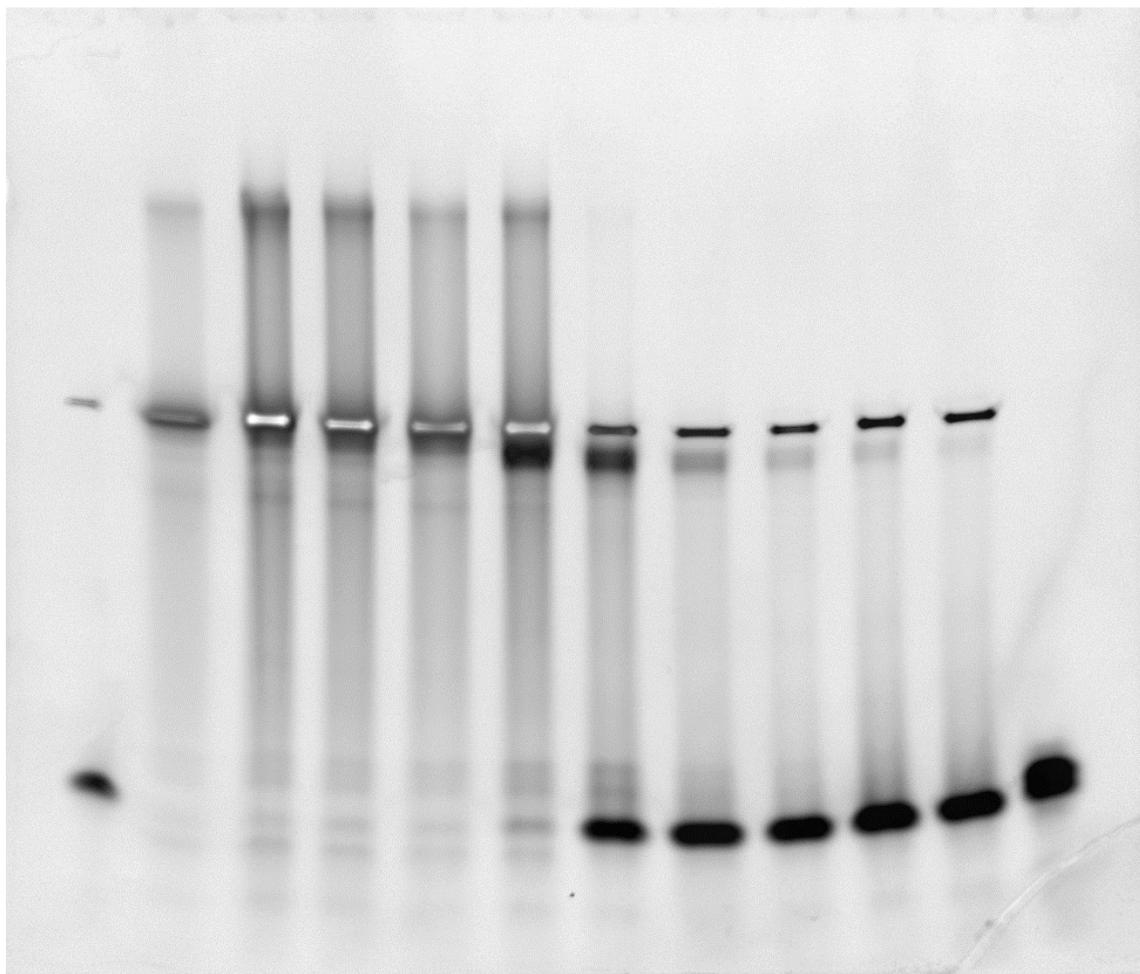
Restriction digests of TOP10 isolated pGDR11-Tgo & pGDR11-KOD RI plasmids treated with NEB NheI-HF (2,5) and AvrII (3,6) or SacI-HF (4,7), restriction enzymes, loaded alongside Puc19 with SacI-HF (8) with reference to 1 kb Plus GeneRuler DNA ladder (1). Ran on a 10 % TAE agarose gel in 1X TBE buffer at 200 V for 40 minutes with 0.25 $\mu\text{L}\cdot\text{mL}^{-1}$ ethidium bromide-1X TBE staining for 5 minutes.



APPENDIX B

Therminator XPE Dilution Series

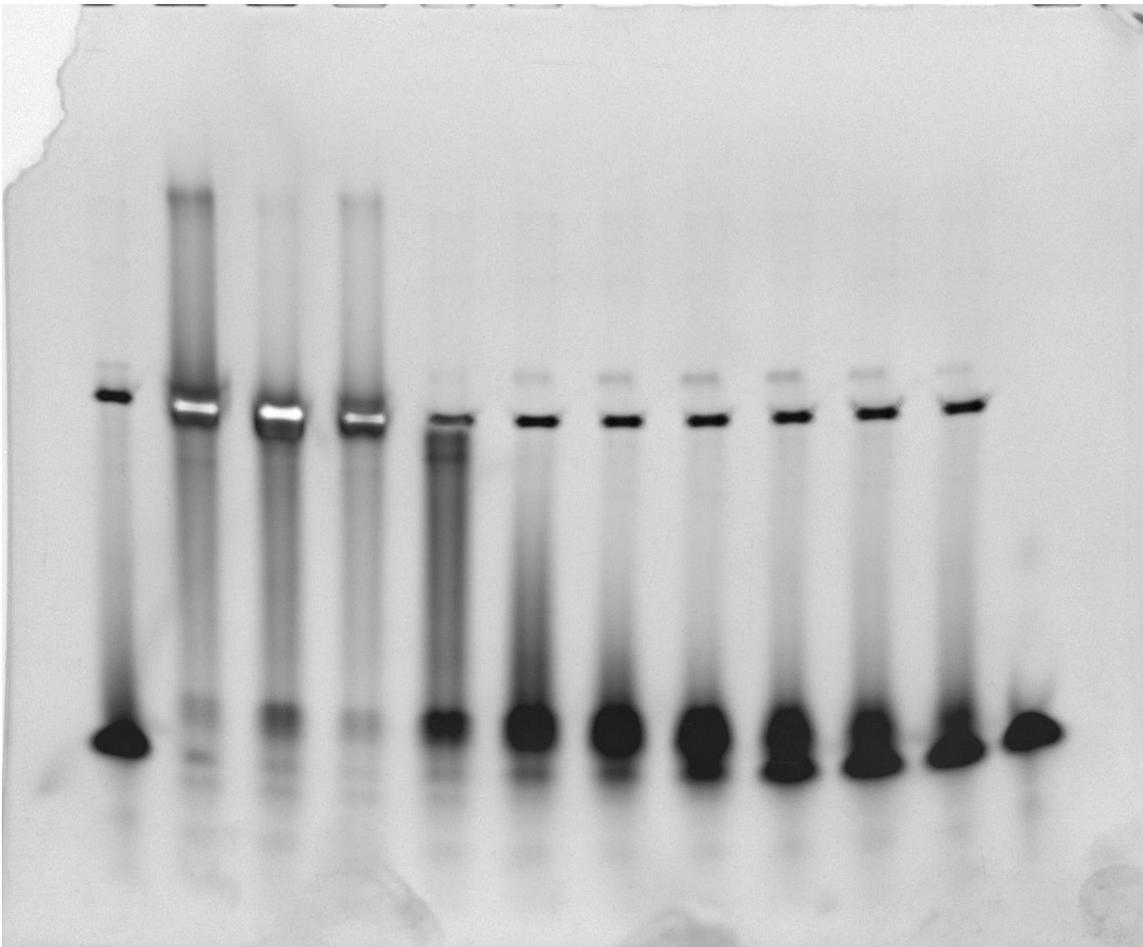
Therminator dilution series, conducted similarly to Tgo polymerase (Figure 5), displaying less laddering. Lanes increase 2X in dilution from left to right. Loaded with 20 μL samples at 0.091 μM in DI formamide, 1 hr at 180 V, 1:00 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Optimum reaction conditions: 1:30 hr (2:00 high yield), 0.25 μM Tgo (exo-) polymerase, 25 mM. EtBr selectively stains DNA T-Lib 78-mer, while staining FANA poorly. NP – no polymerase (leftmost most); NT – no template (right most). Black and white composite indicates overlapping bands with white. DNA T-Lib is clearly visible across the top of the gel. Indicates that optimum dilution is 1:8.



APPENDIX C

Deep Vent (exo-) XPE Dilution Series

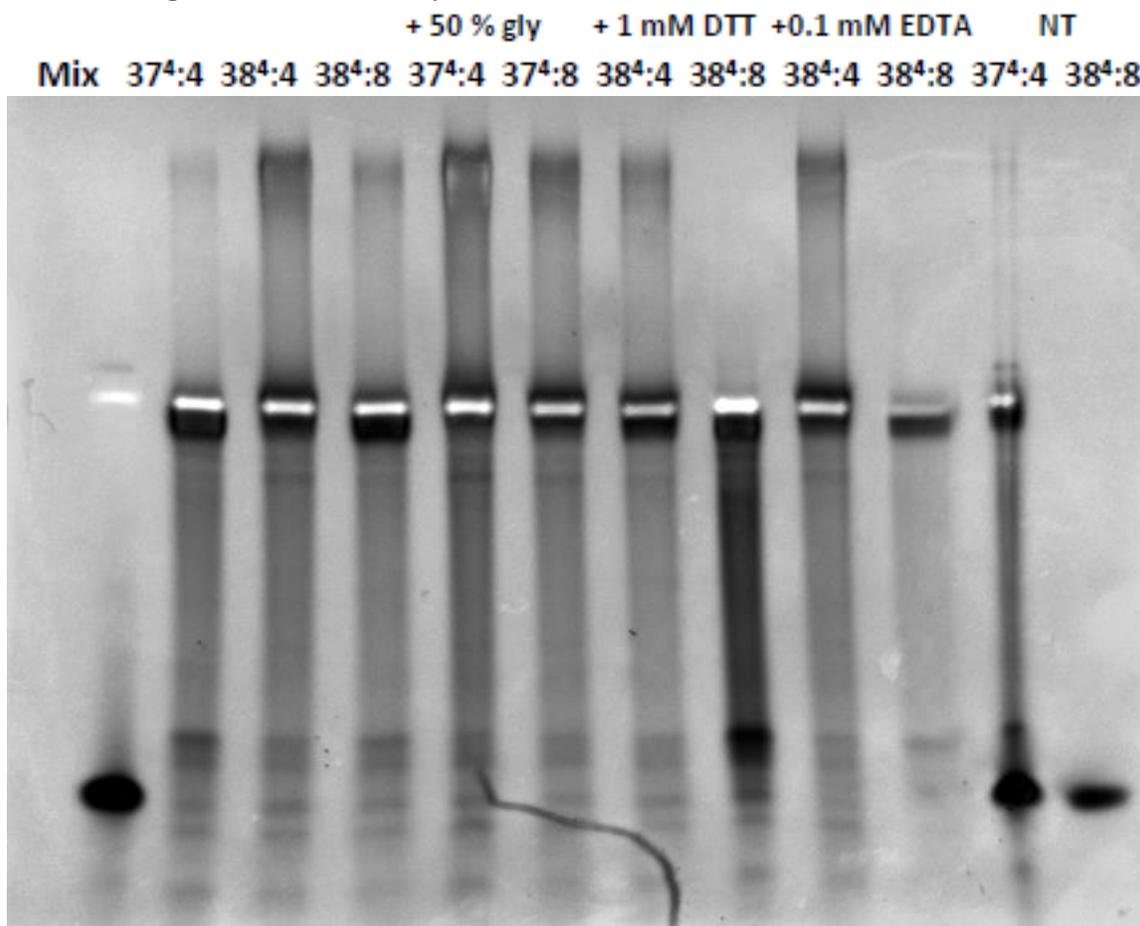
Deep Vent (exo-) dilution series, conducted similarly to Tgo polymerase (Figure 5), displaying less laddering. Lanes increase 2X in dilution from left to right. Loaded with 20 μL samples at 0.091 μM in DI formamide, 1 hr at 180 V, 1:00 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. Optimum reaction conditions: 1:30 hr (2:00 high yield), 0.25 μM Tgo (exo-) polymerase, 25 mM. EtBr selectively stains DNA T-Lib 78-mer, while staining FANA poorly. NP – no polymerase (left most); NT – no template, no DNA library (right most). Black and white composite indicates overlapping bands with white. DNA T-Lib is clearly visible across the top of the gel. Indicates that optimum dilution is 1:4.



APPENDIX D

Tgo Storage Additives Study

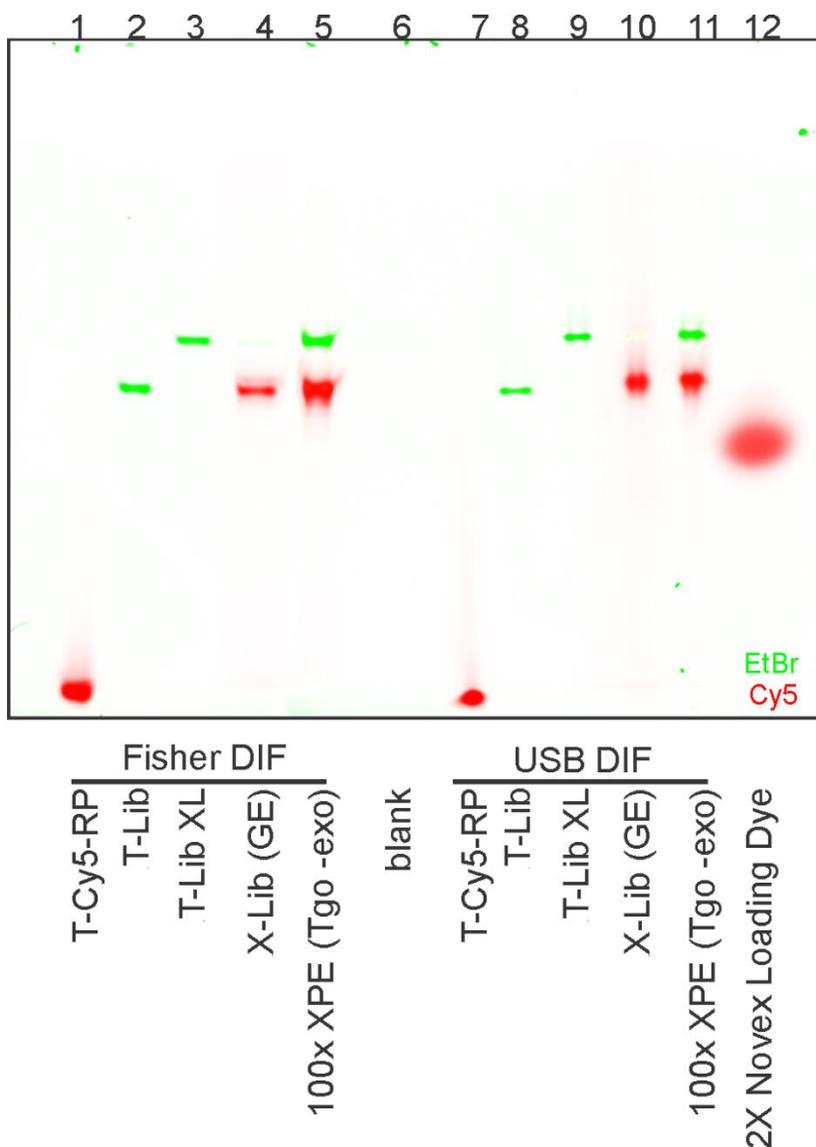
Loaded with 15 μL samples at 0.091 μM in DI formamide, 1 hr at 180 V, 1:00 hr preheated at 8 W, 4-12 % TBE-Urea PAGE. The KCl concentration over NaCl was not tested; however, Tween 20 IGEPAL CA-630, Triton X-100 was not found to be beneficial to enhancing FANA XPE activity (data not reported). However, excessive DDT was found to negatively affect FANA XPE activity, while 50 % glycerol was found to positively affect FANA XPE activity. The effect of EDTA was not found to be significant at the tested concentration. Mix – DNA library and primer, no polymerase; NT – no template, no DNA library.



APPENDIX E

DIF Purity and Early Gel Extractions of X-Lib

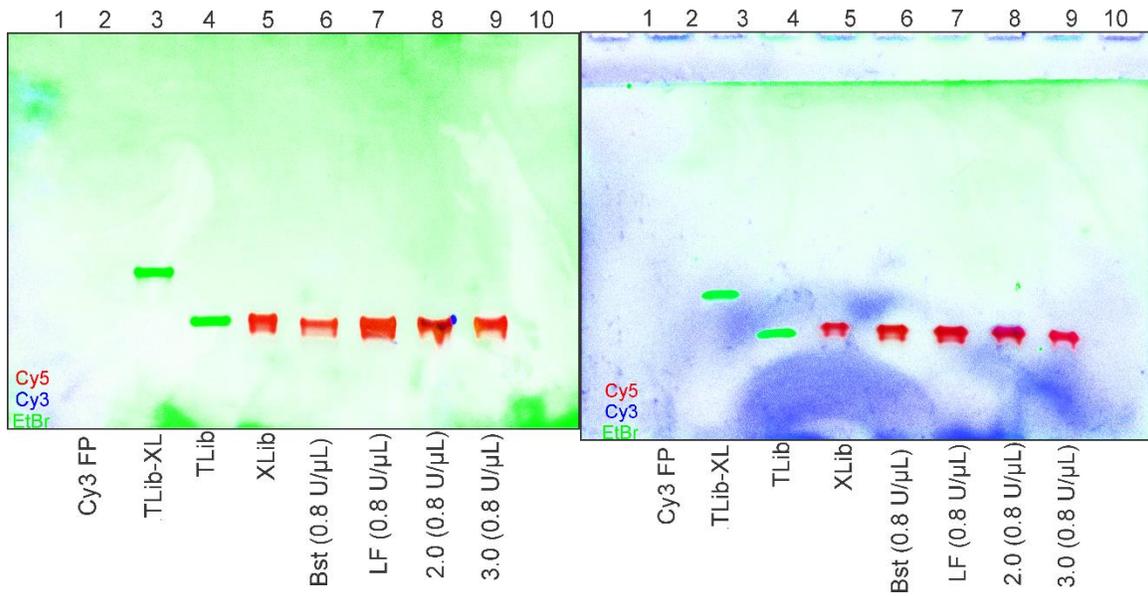
Effect of formamide purity was assayed using two different suppliers of formamide- synthesis grade Fisher/Thermo Scientific & Affymetrix/USB Biochemicals (Santa Clara, CA) UltraPure formamide. While results are minimal, it was observed that formamide purity, assayed via $^1\text{H-NMR}$ & $^{13}\text{C-NMR}$ (data not reported), correlated with tighter band width. Although X-Lib through gel extraction (GE) was primarily FANA material, some DNA contamination is still detectable. Samples were loaded in 15 μL aliquots, 10:1 formamide and 1X XPE concentrations (1.36 pmol) of nucleic acid.



APPENDIX F

DIF Purity and Early Gel Extractions of X-Lib

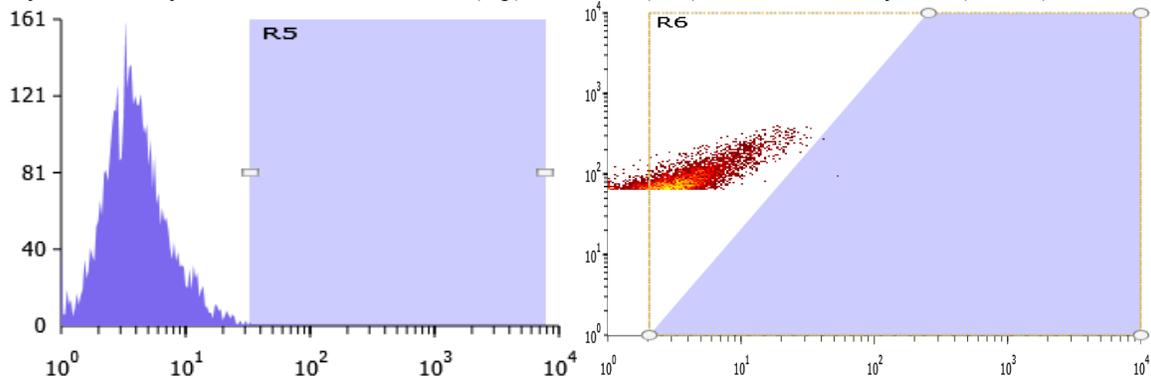
Samples were loaded in 15 μ L aliquots, except for X-Lib loaded at 25 μ L due to low concentration, 10:1 formamide and 1X XPE concentrations (1.36 pmol) of nucleic acid. The composite images detail unsuccessful cDNA transcription experiments.



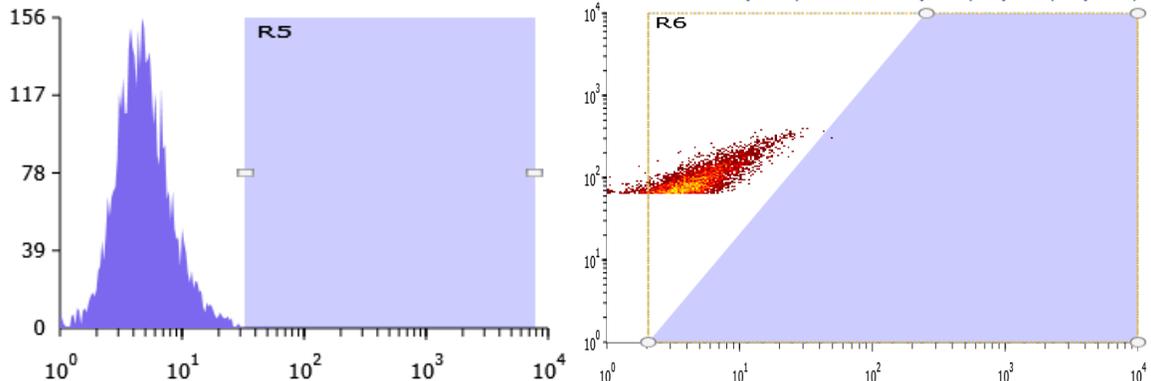
APPENDIX G

FP-Particle Encapsulation Affinity

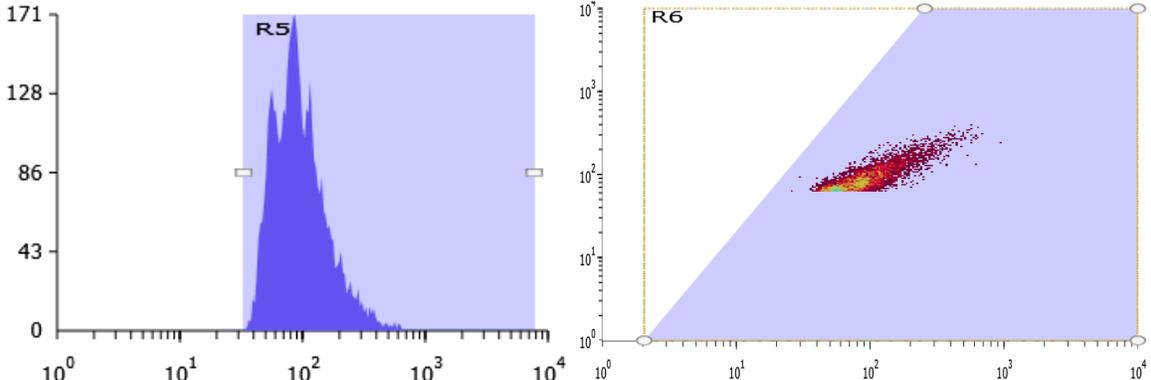
Bio-Rad S3 Cell Sorter-FACS captured 5000 events of 1:100 T1-AM6-FP functionalized 10^7 Dynabeads MyOne Carboxylic Acid beads. Bare beads (top), FP-beads (mid), FP-beads+T1-Cy5-RF (bottom).



Label	Count	Mean	% of Plot	CV	Std Dev	Median	Label	Count	Mean	% of Plot	CV	Std Dev	Median
Total	4537	4.86	100.00 %	79.65	3.87	4.00	Total	4537	(4.72, 9)	100.00 %	(81.95, 4)	(3.87, 4)	(4.00, 8)
R5	5	46.00	0.11 %	31.08	14.30	33.00	R6	3	(54.67, :	0.07 %	(22.33, 6)	(12.21, 8)	(41.00, 6)



Label	Count	Mean	% of Plot	CV	Std Dev	Median	Label	Count	Mean	% of Plot	CV	Std Dev	Median
Total	4699	5.60	100.00 %	63.77	3.57	5.00	Total	4699	(5.47, 10)	100.00 %	(65.30, 4)	(3.57, 4)	(5.00, 8)
R5	2	45.50	0.04 %	7.43	3.38	43.00	R6	1	(49.00, :	0.02 %	(0.91, 0)	(0.44, 0)	(1.00, 1)

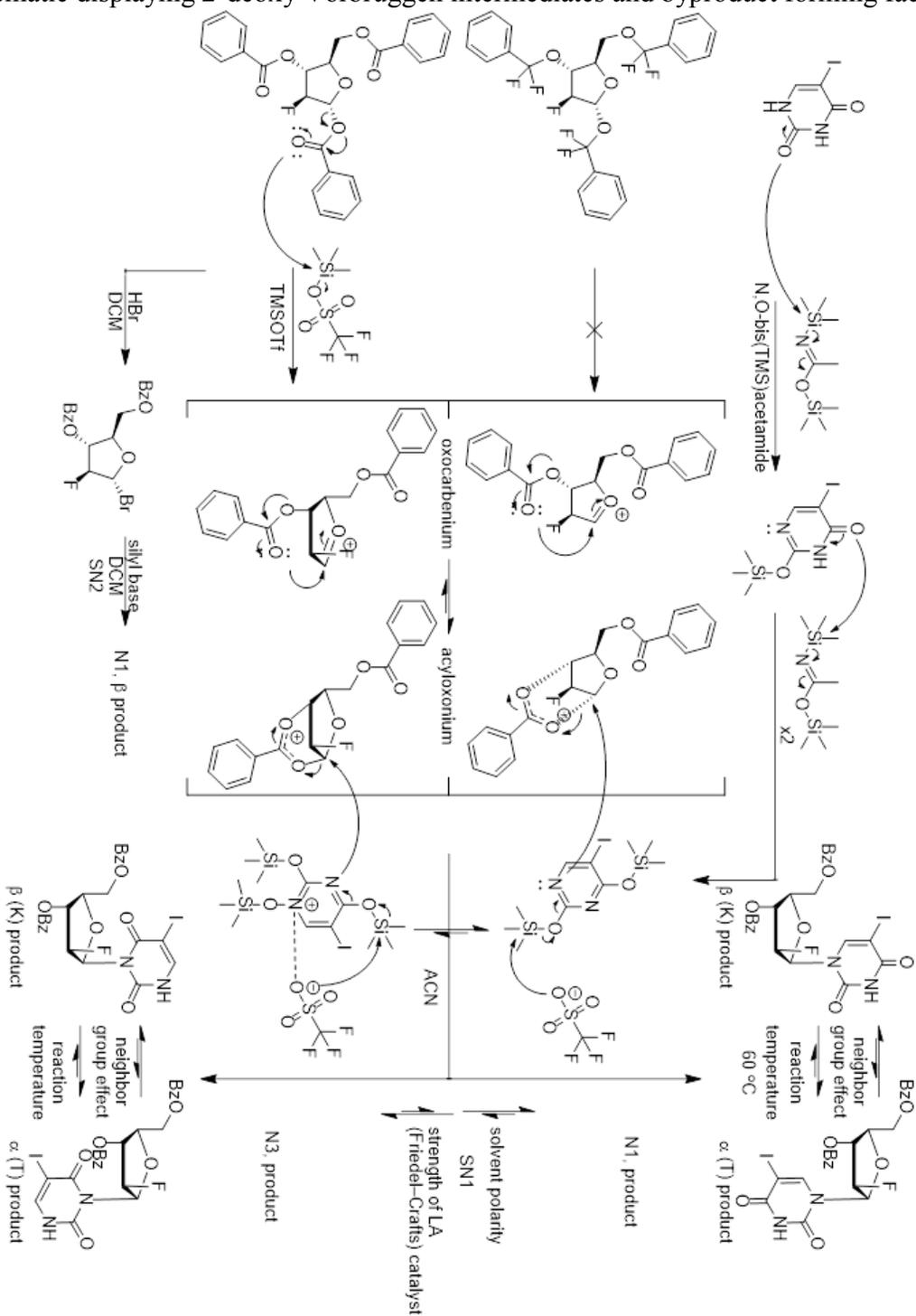


Label	Count	Mean	% of Plot	CV	Std Dev	Median	Label	Count	Mean	% of Plot	CV	Std Dev	Median
Total	4748	111.45	100.00 %	66.13	73.70	91.00	Total	4748	(111.33, 100)	100.00 %	(66.20, 4)	(73.70, 4)	(91.00, 8)
R5	4746	111.48	99.96 %	66.10	73.70	91.00	R6	4748	(111.33, 100)	100.00 %	(66.20, 4)	(73.70, 4)	(91.00, 8)

APPENDIX H

2'-Deoxy-2'-Fluoro-D-arabinofuranosyl Glycosylation Mechanisms & Conditions

Schematic displaying 2-deoxy Vorbrüggen intermediates and byproduct forming factors.



APPENDIX I

Crystal data and structure refinement for C₂₃H₁₈FIN₂O₇

Identification code	C23H18FIN2O7
Empirical formula	C49 H44 F2 I2 N4 O15
Formula weight	1220.68
Temperature	100(2) K
Wavelength	1.54178 Å
Crystal system	Orthorhombic
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell dimensions	a = 9.7073(3) Å a = 90°. b = 21.2291(7) Å b = 90°. c = 23.6686(9) Å g = 90°.
Volume	4877.6(3) Å ³
Z	4
Density (calculated)	1.662 Mg/m ³
Absorption coefficient	10.832 mm ⁻¹
F(000)	2440
Crystal size	0.480 x 0.120 x 0.090 mm ³
Theta range for data collection	2.796 to 66.649°.
Index ranges	-11<=h<=11, -15<=k<=24, -28<=l<=28
Reflections collected	32826
Independent reflections	8325 [R(int) = 0.0402]
Completeness to theta =	66.649° 99.6 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.7528 and 0.4461
Refinement method	Full-matrix least-squares on F ²
Data / restraints / parameters	8325 / 2 / 658
Goodness-of-fit on F ²	1.027
Final R indices [I>2sigma(I)]	R1 = 0.0246, wR2 = 0.0576
R indices (all data)	R1 = 0.0261, wR2 = 0.0581
Absolute structure parameter	0.039(3)
Extinction coefficient	n/a
Largest diff. peak and hole	0.715 and -0.460 e.Å ⁻³

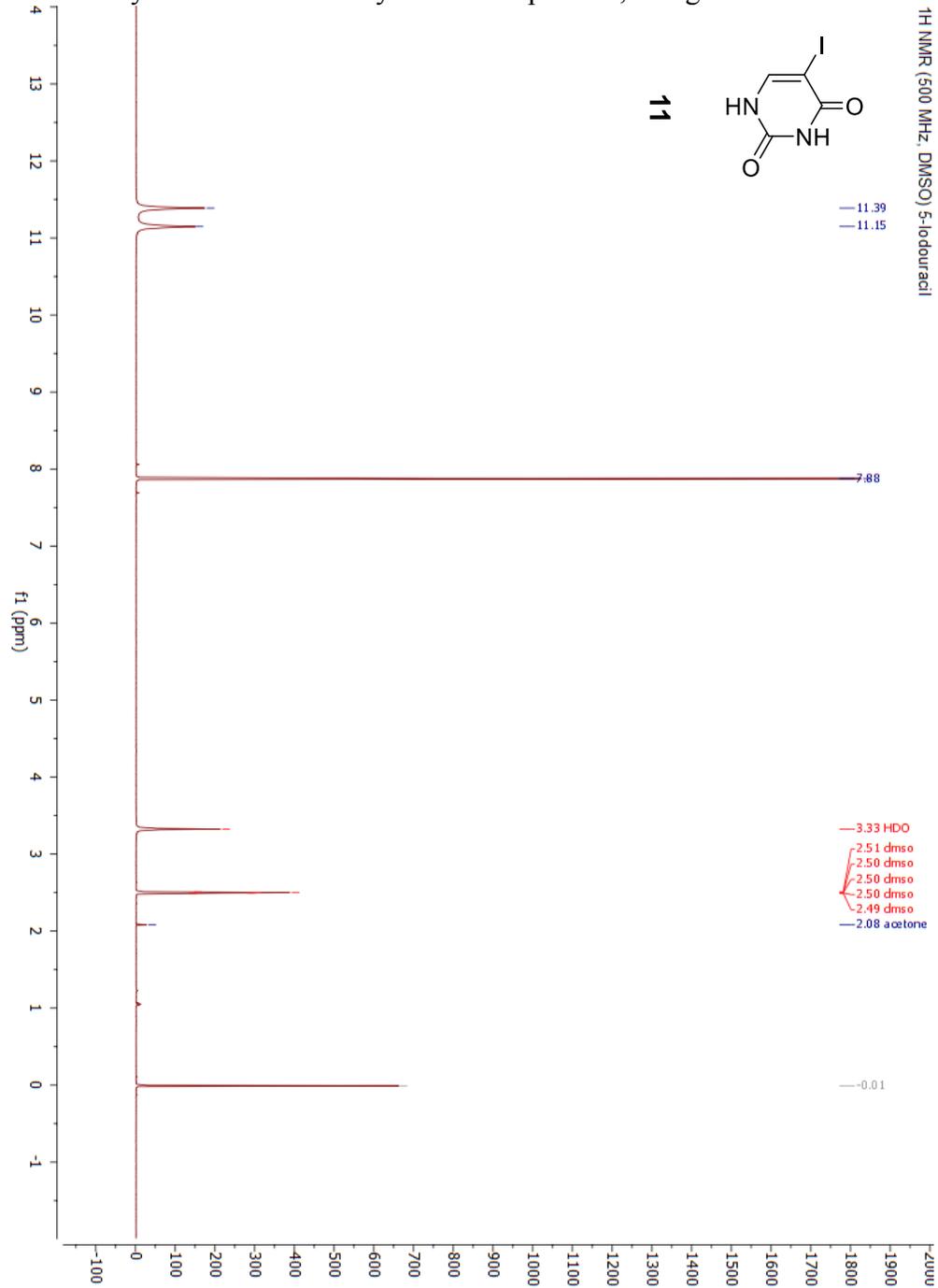
Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for C23H18FIN2O7. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
I(1)	6796(1)	9765(1)	3472(1)	22(1)
F(1)	12089(3)	9466(1)	3591(1)	24(1)
O(1)	11297(3)	8339(1)	4206(1)	21(1)
O(2)	10497(3)	7560(2)	2821(1)	23(1)
O(3)	6490(3)	8628(1)	2620(1)	24(1)
O(4)	14503(3)	8256(2)	3881(1)	21(1)
O(5)	15674(4)	8456(2)	4682(1)	28(1)
O(6)	11109(3)	8836(1)	5272(1)	21(1)
O(7)	11647(4)	9626(2)	5862(1)	28(1)
N(1)	10226(4)	8467(2)	3340(2)	19(1)
N(2)	8508(4)	8107(2)	2730(2)	18(1)
C(1)	11537(5)	8362(2)	3618(2)	19(1)
C(2)	12648(5)	8861(2)	3532(2)	21(1)
C(3)	13558(5)	8745(2)	4045(2)	20(1)
C(4)	12532(5)	8524(2)	4502(2)	19(1)
C(5)	9800(5)	8012(2)	2952(2)	16(1)
C(6)	7624(5)	8594(2)	2846(2)	19(1)
C(7)	8131(5)	9039(2)	3264(2)	17(1)
C(8)	9376(5)	8955(2)	3503(2)	19(1)
C(9)	15574(5)	8170(2)	4243(2)	21(1)
C(10)	16596(5)	7696(2)	4035(2)	22(1)
C(11)	16320(5)	7302(2)	3576(2)	22(1)
C(12)	17292(5)	6857(2)	3419(2)	27(1)
C(13)	18530(5)	6809(2)	3710(2)	30(1)
C(14)	18803(6)	7210(2)	4157(2)	31(1)
C(15)	17834(5)	7656(2)	4321(2)	26(1)
C(16)	12207(5)	9049(2)	4910(2)	25(1)
C(17)	10907(5)	9195(2)	5732(2)	20(1)
C(18)	9679(5)	8990(2)	6062(2)	20(1)
C(19)	9507(5)	9238(2)	6604(2)	25(1)
C(20)	8408(6)	9039(2)	6935(2)	29(1)
C(21)	7481(5)	8596(2)	6728(2)	31(1)
C(22)	7640(5)	8360(2)	6189(2)	30(1)
C(23)	8741(5)	8550(2)	5855(2)	24(1)
I(2)	6283(1)	5805(1)	1547(1)	23(1)
F(2)	2436(3)	7260(1)	678(1)	23(1)
O(8)	1074(3)	6639(1)	1646(1)	21(1)
O(9)	7188(3)	7164(1)	2056(1)	22(1)
O(10)	3104(3)	8191(2)	2156(1)	26(1)
O(11)	-997(3)	7650(1)	978(1)	21(1)
O(12)	-2508(3)	7022(2)	529(1)	25(1)
O(13)	1171(3)	5962(1)	649(1)	22(1)
O(14)	565(4)	5009(2)	327(2)	31(1)
N(3)	3101(4)	7196(2)	1778(2)	18(1)
N(4)	5110(4)	7647(2)	2121(2)	20(1)
C(24)	1640(4)	7249(2)	1631(2)	19(1)
C(25)	1398(5)	7490(2)	1031(2)	21(1)
C(26)	41(5)	7182(2)	872(2)	20(1)
C(27)	-93(5)	6613(2)	1272(2)	19(1)
C(28)	3831(5)	6671(2)	1629(2)	19(1)
C(29)	5194(5)	6623(2)	1721(2)	18(1)
C(30)	5936(5)	7147(2)	1972(2)	18(1)
C(31)	3724(5)	7715(2)	2028(2)	19(1)
C(32)	-2259(5)	7513(2)	771(2)	19(1)
C(33)	-3254(5)	8028(2)	871(2)	21(1)
C(34)	-2884(5)	8574(2)	1164(2)	27(1)
C(35)	-3851(6)	9050(2)	1233(2)	31(1)
C(36)	-5159(6)	8986(2)	1013(2)	31(1)
C(37)	-5533(5)	8443(2)	723(2)	28(1)
C(38)	-4585(5)	7958(2)	661(2)	23(1)
C(39)	-79(5)	5986(2)	975(2)	23(1)
C(40)	1390(6)	5430(2)	346(2)	24(1)
C(41)	2771(5)	5425(2)	84(2)	24(1)
C(42)	3194(6)	4884(2)	-199(2)	31(1)
C(43)	4514(6)	4849(3)	-414(2)	37(1)
C(44)	5422(6)	5345(3)	-348(2)	34(1)
C(45)	5009(6)	5894(3)	-72(2)	31(1)
C(46)	3691(6)	5932(2)	142(2)	26(1)
O(15)	648(5)	8905(2)	2060(2)	48(1)
C(47)	-452(8)	9876(3)	1951(4)	62(2)
C(48)	896(8)	9520(3)	1883(3)	52(2)
C(49)	2058(8)	9833(3)	2159(3)	59(2)

APPENDIX J.1

5-Iodouracil
¹H-NMR Spectrum

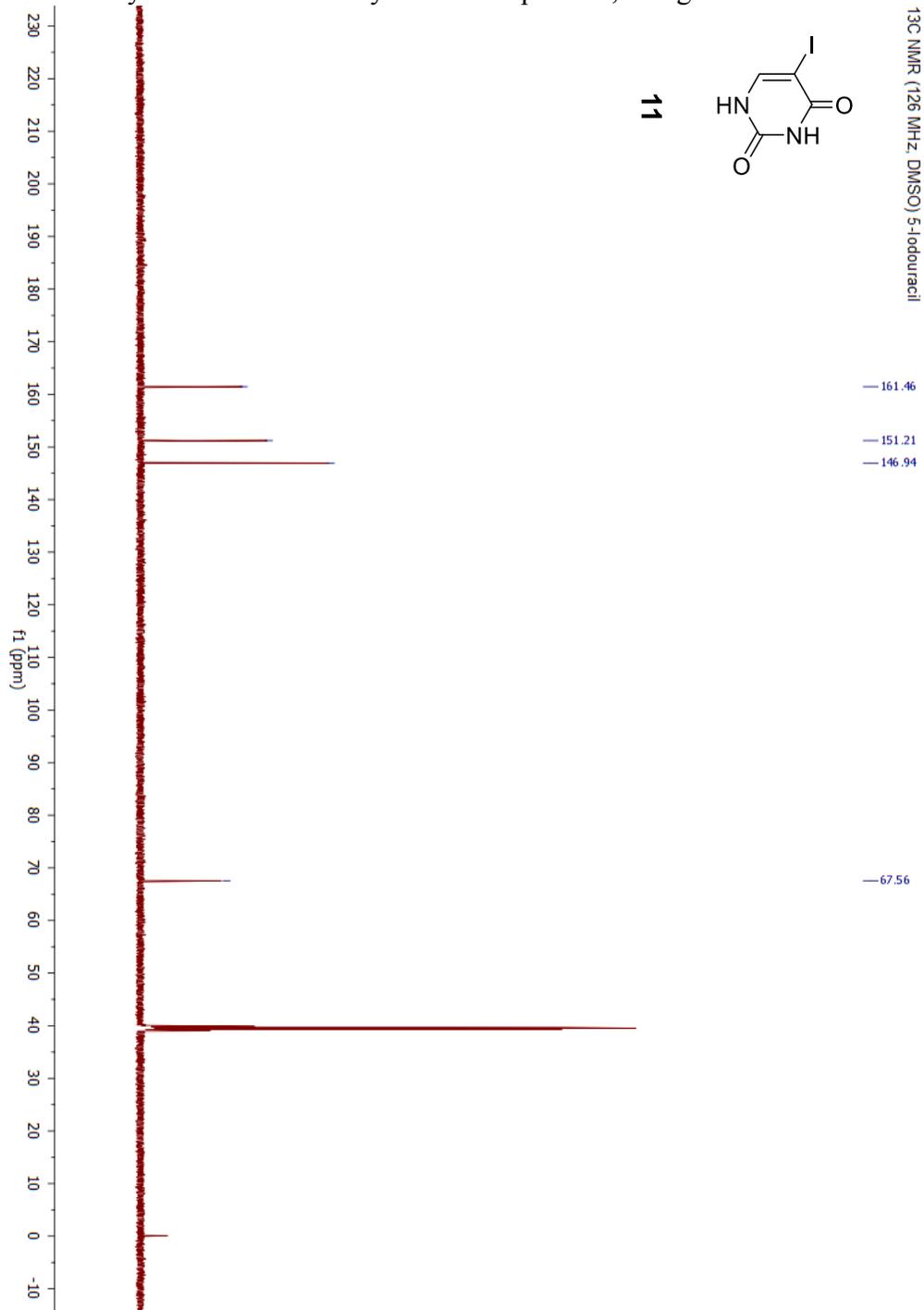
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX J.2

5-Iodouracil
¹³C-NMR Spectrum in DMSO

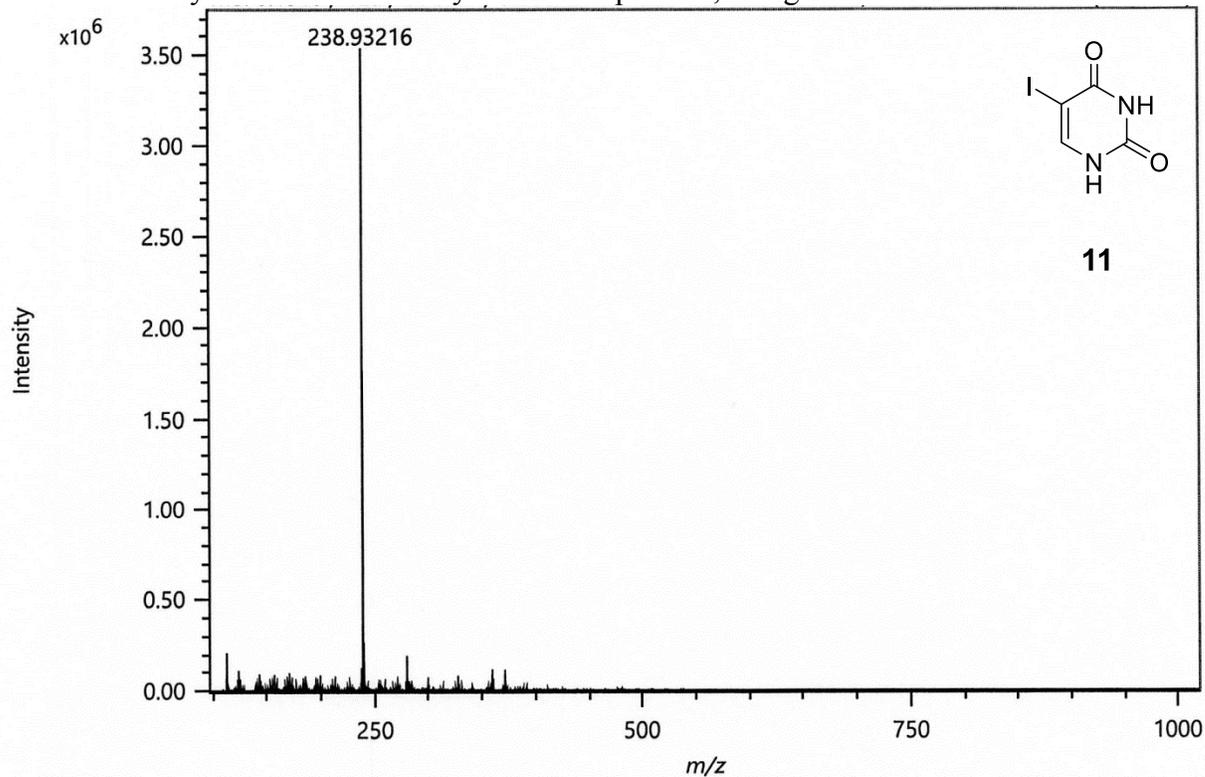
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX J.3

5-Iodouracil
DART Mass Spectrum

Chemical analysis of colorless recrystallization product, alongside reference structure.



Elemental Composition

Parameters

Tolerance: 10.00 ppm
Electron: Odd/Even
Charge: +1
DBE: -20.5 - 100.0

Elements Set 1:

Symbol	C	H	N	I	O
Min	0	0	1	1	1
Max	100	100	2	1	2

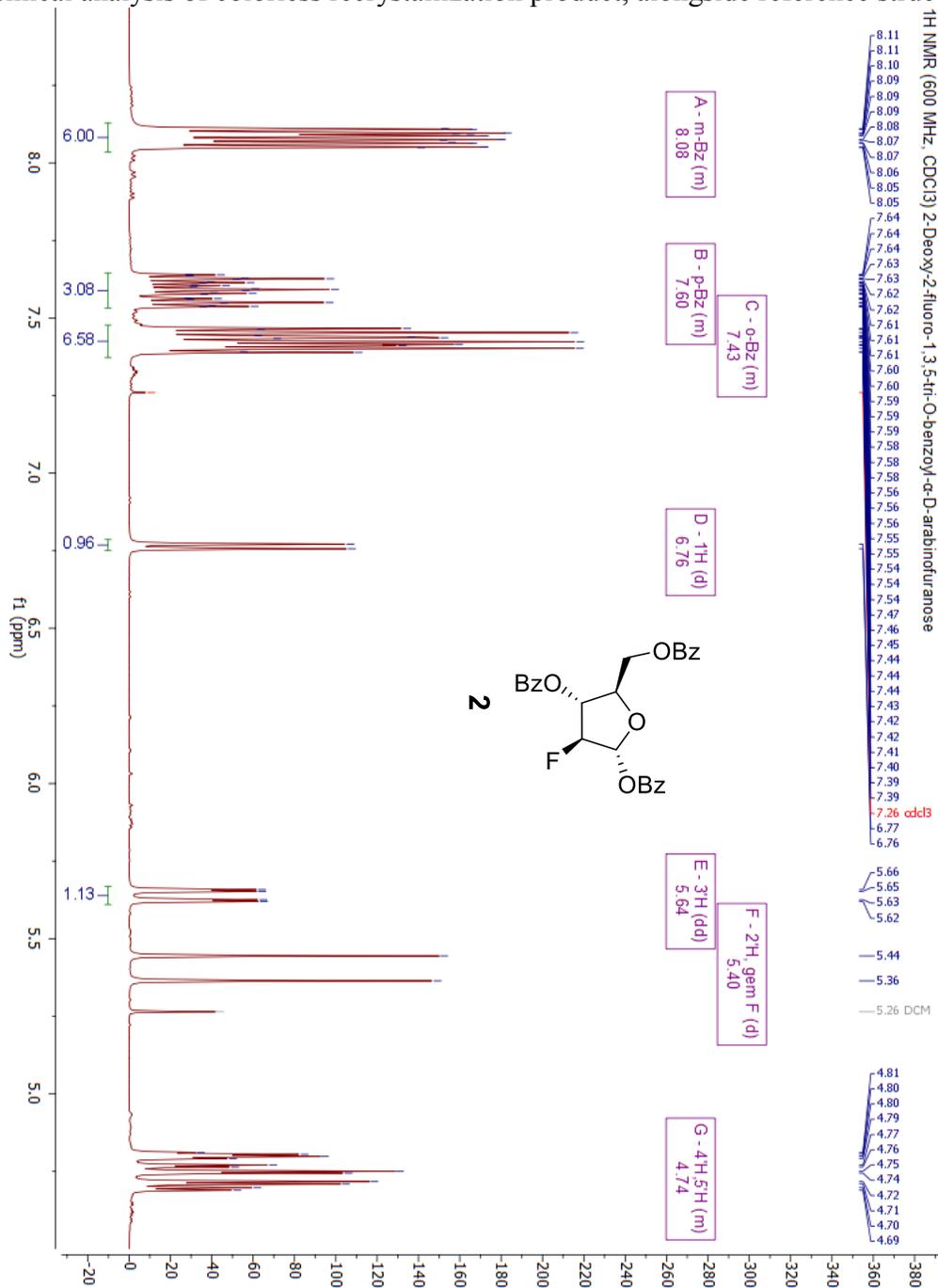
Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [ppm]	DBE
238.93216	3485166.00	C ₄ H ₄ N ₂ O ₂ I	238.93120	4.03	3.5

APPENDIX K.1b

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
 $^1\text{H-NMR}$ Spectrum in CDCl_3

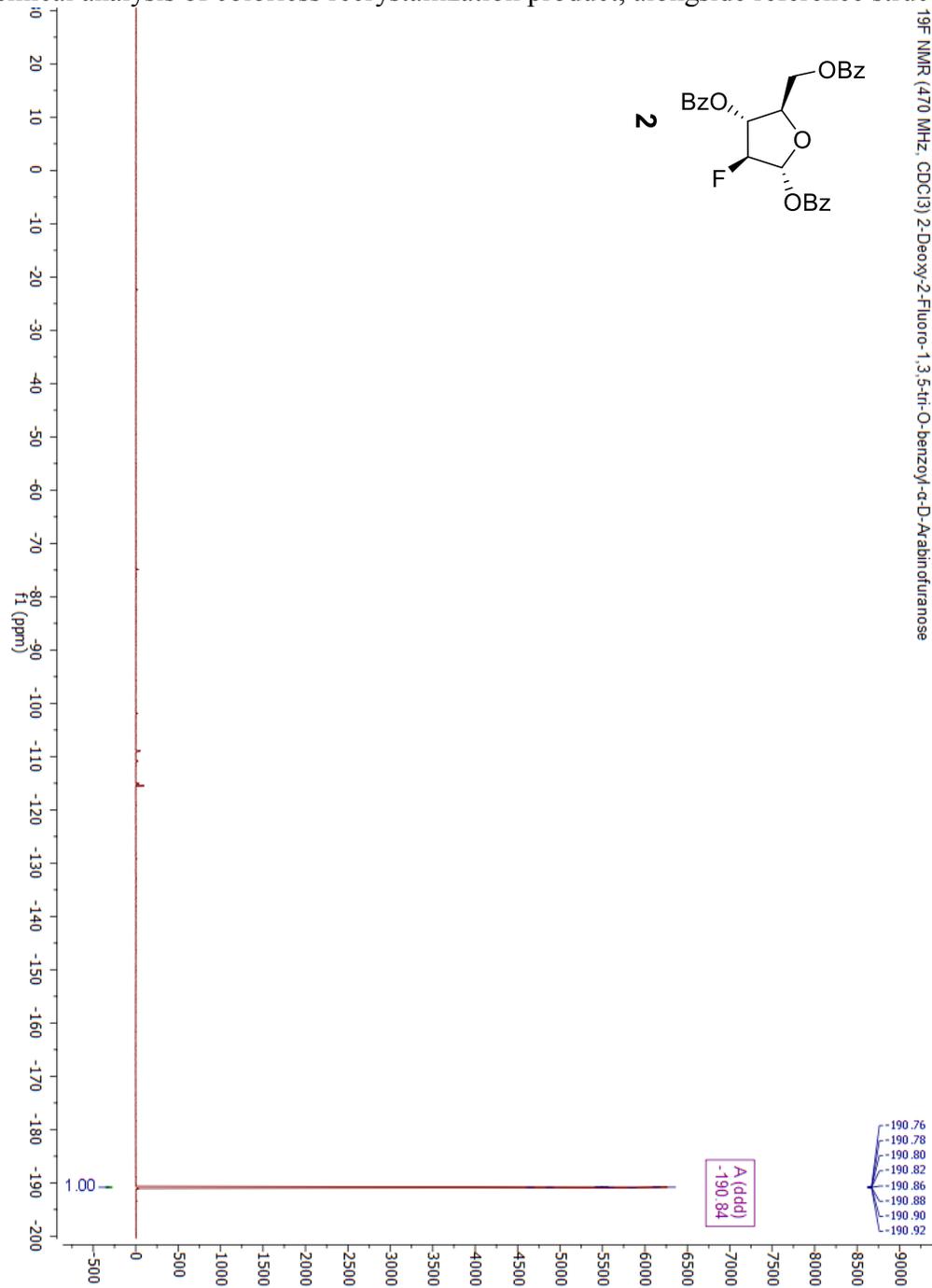
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.2a

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
 ^{19}F -NMR Spectrum in CDCl_3

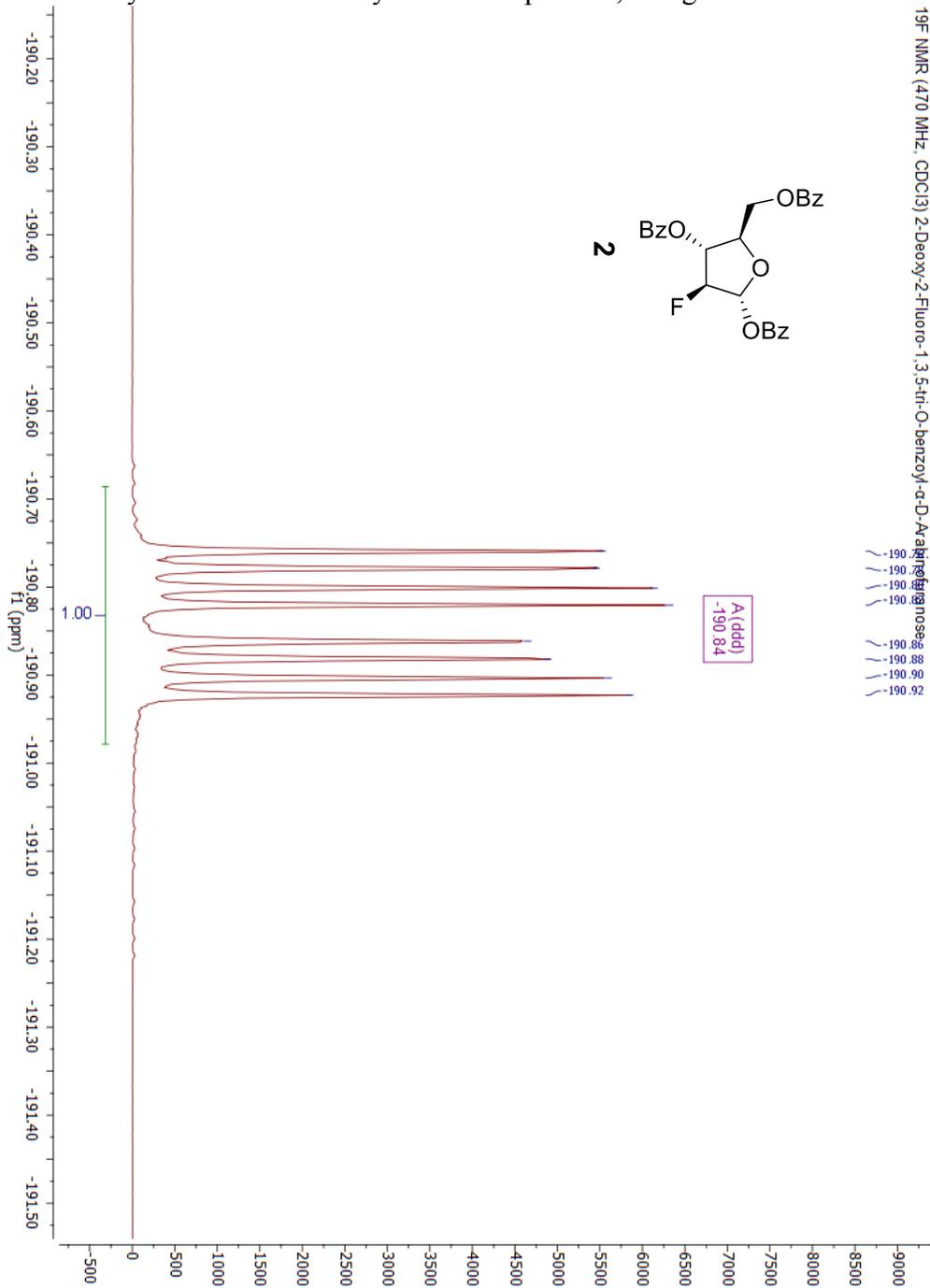
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.2b

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
 ^{19}F -NMR Spectrum in CDCl_3

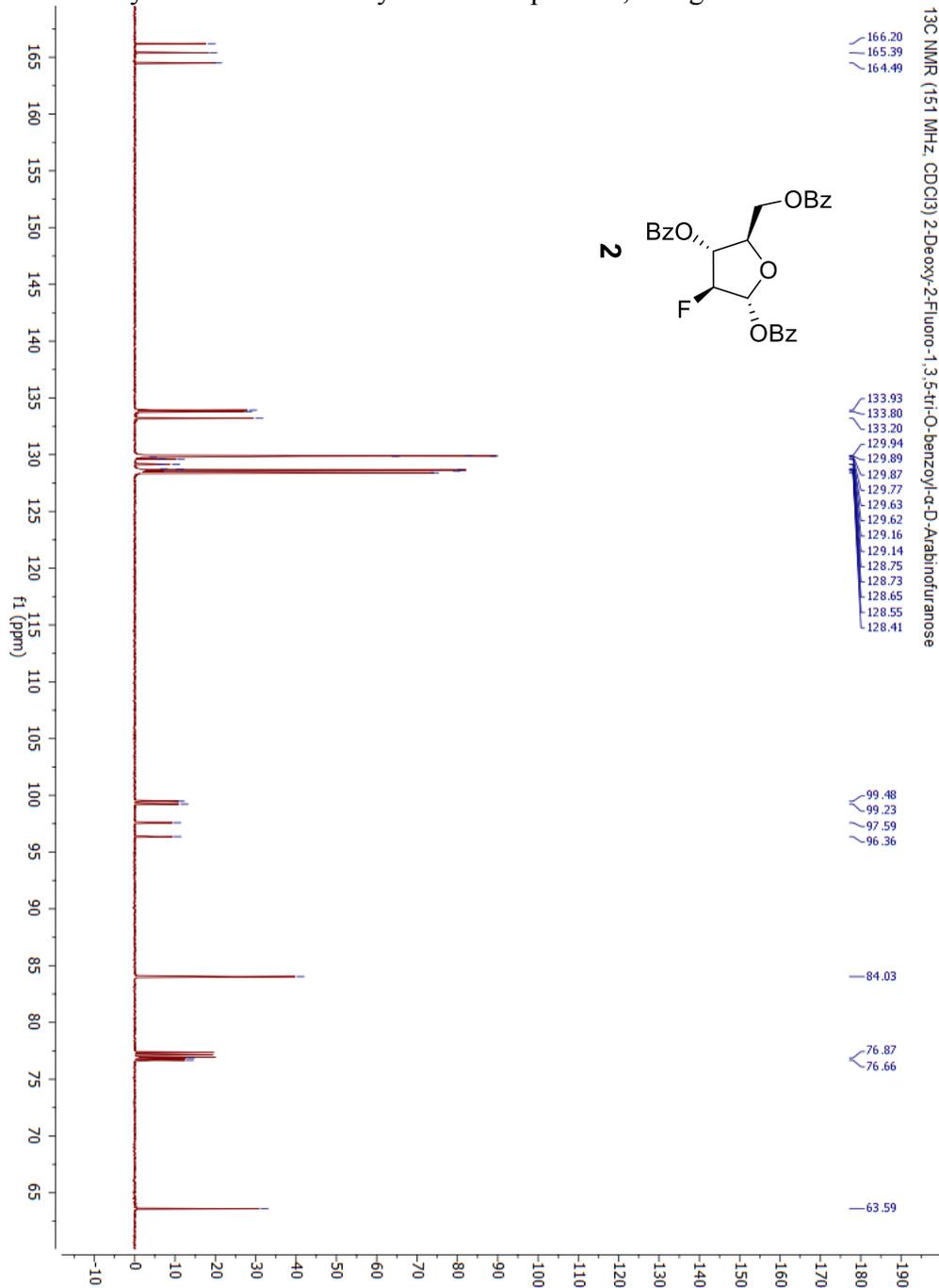
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.3a

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
 ^{13}C -NMR Spectrum in CDCl_3

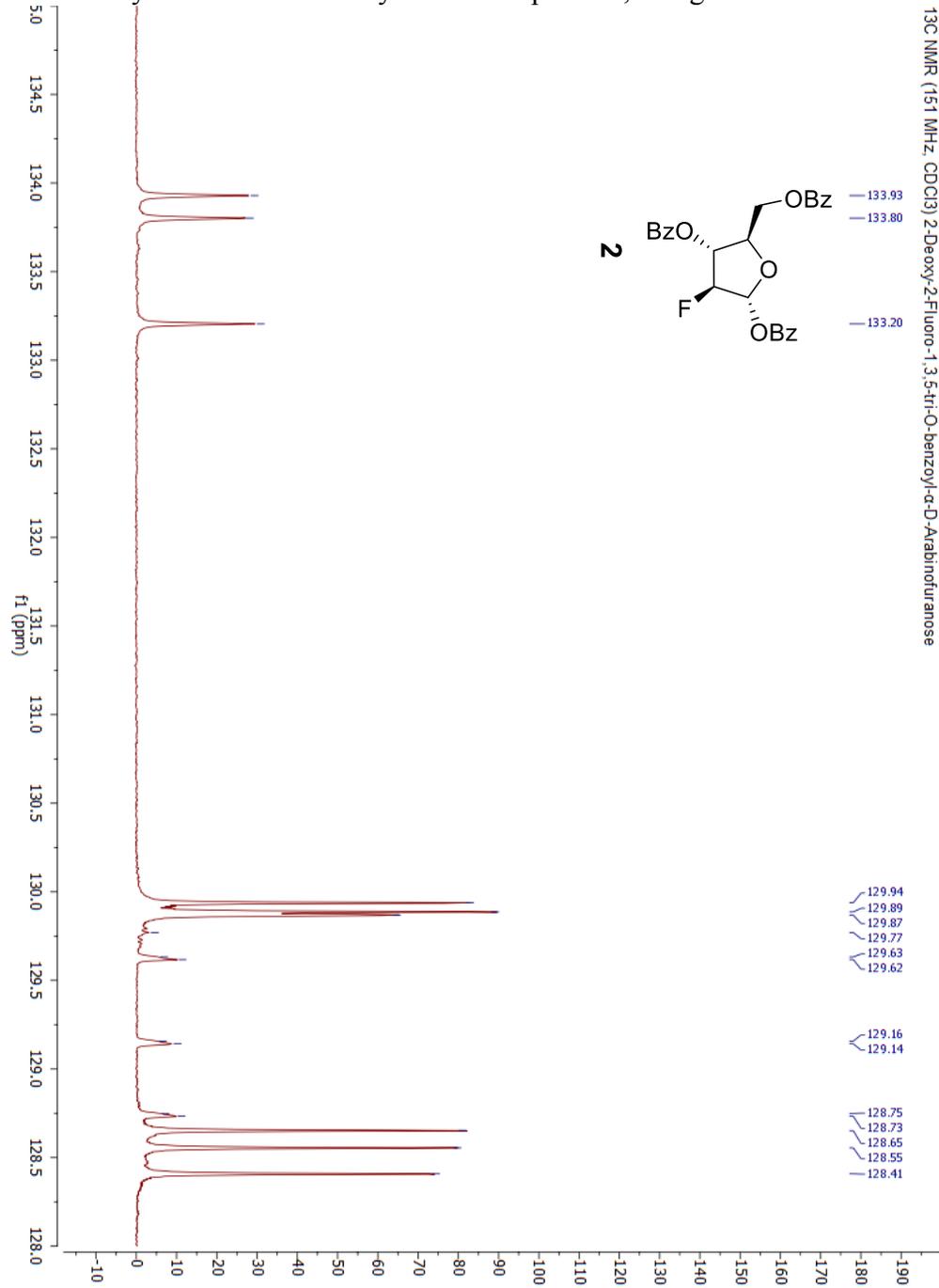
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.3b

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
 ^{13}C -NMR Spectrum in CDCl_3

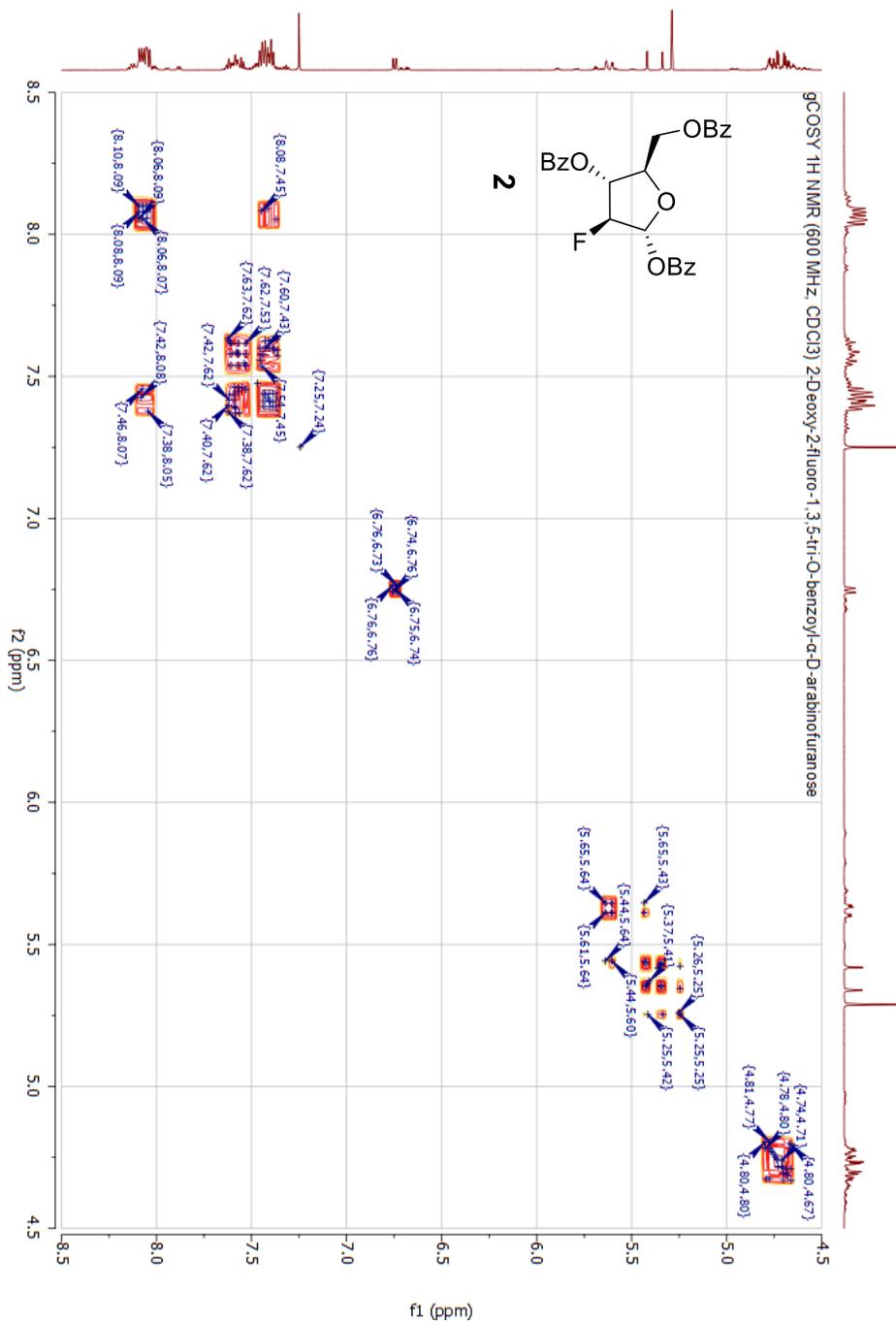
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.4a

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
gCOSY Spectrum in CDCl₃

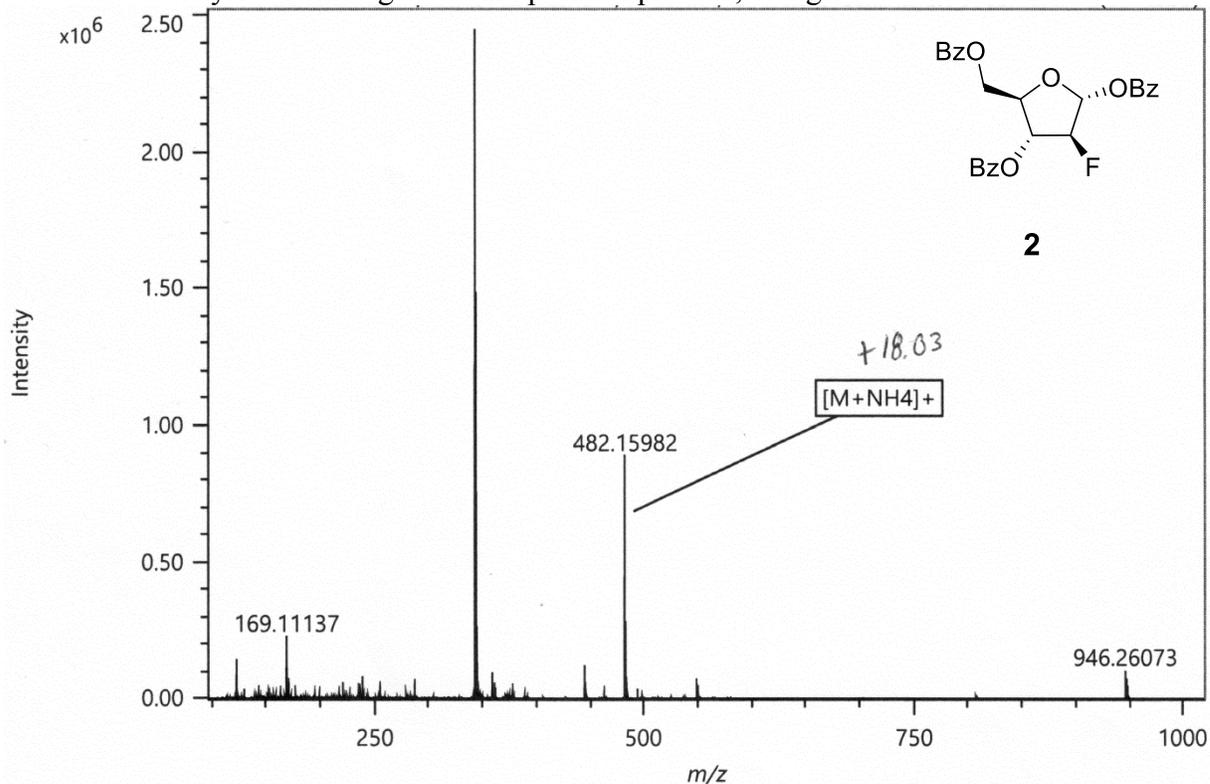
Chemical analysis of colorless recrystallization product, alongside reference structure.



APPENDIX K.5

2'-Deoxy-2'-Fluoro-1,3,5-tri-O-benzoyl- α -D-Arabinofuranose
DART Mass Spectrum in CDCl₃

Chemical analysis of silica gel column purified product, alongside reference structure.



Elemental Composition

Parameters

Tolerance: 10.00 ppm
Electron: Odd/Even
Charge: +1
DBE: -10.5 - 50.0

Elements Set 1:

Symbol	C	H	F	N	O
Min	0	0	1	1	1
Max	100	100	1	1	7

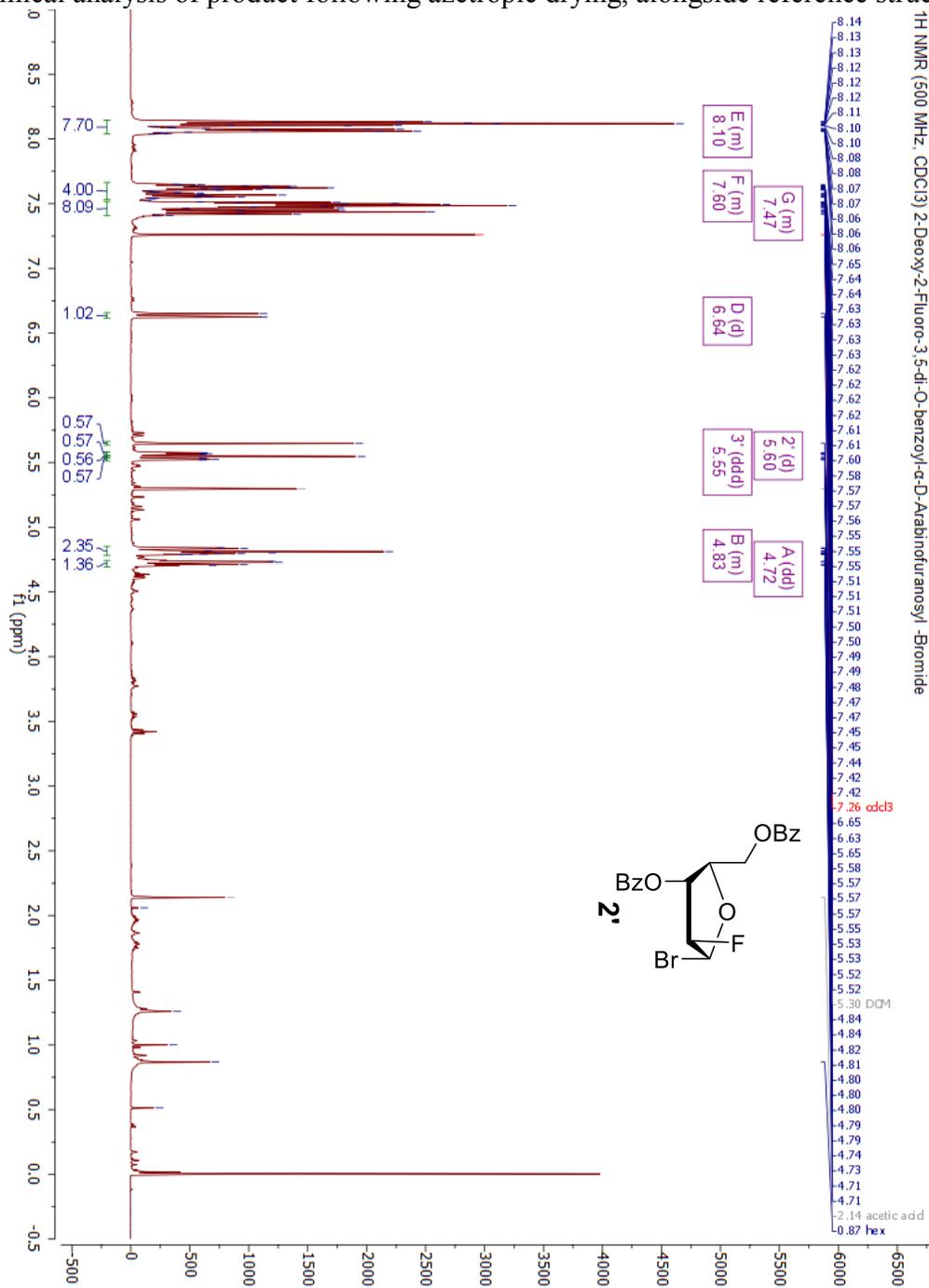
Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [ppm]	DBE
482.15982	857379.20	C ₂₆ H ₂₅ N O ₇ F	482.16096	-2.37	14.5
		C ₃₃ H ₂₁ N O ₂ F	482.15508	9.81	23.5

APPENDIX L.1

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide
 $^1\text{H-NMR}$ Spectrum in CDCl_3

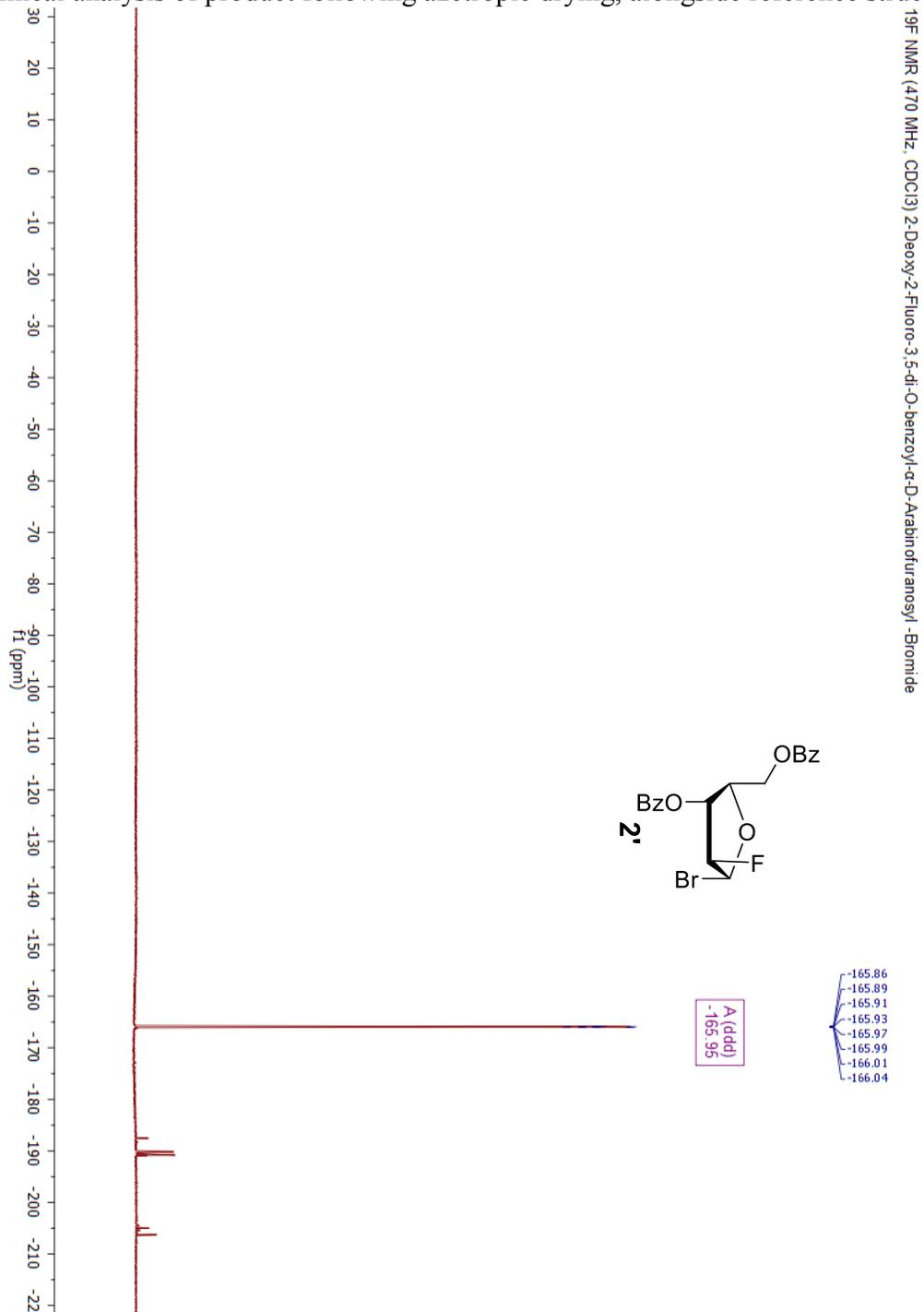
Chemical analysis of product following azeotropic drying, alongside reference structure.



APPENDIX L.2

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide
 ^{19}F -NMR Spectrum in CDCl_3

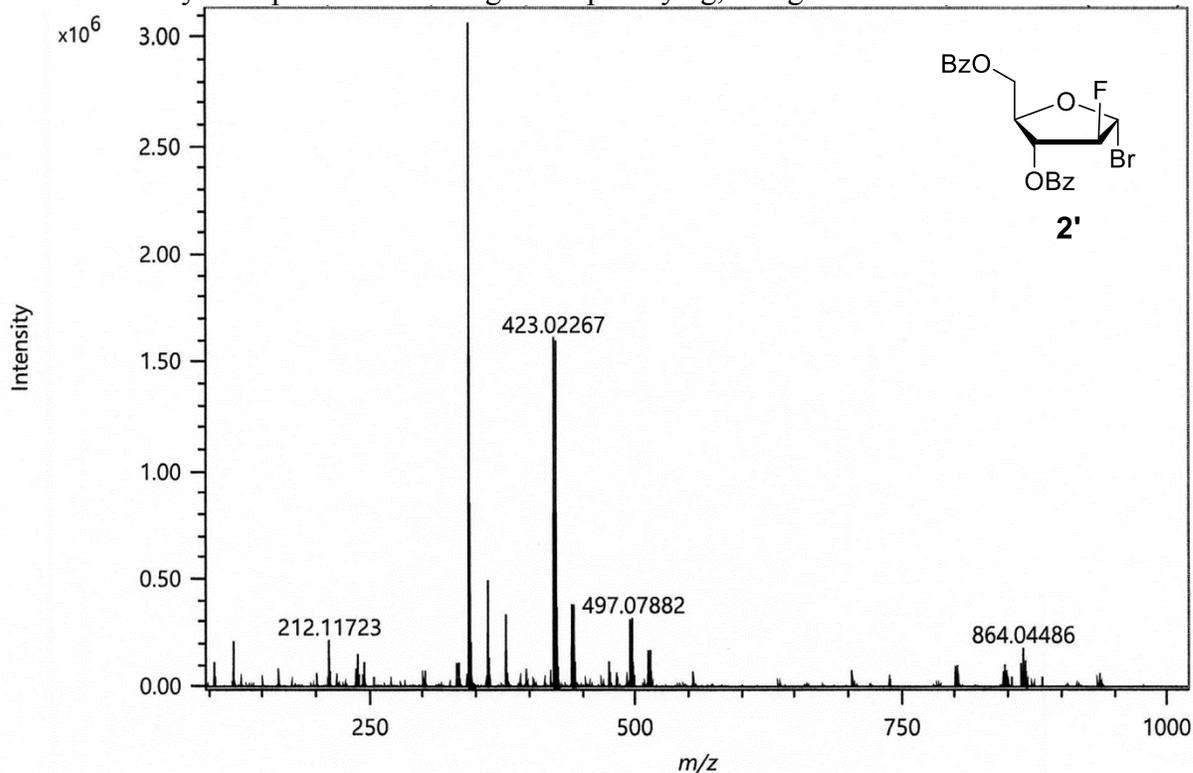
Chemical analysis of product following azeotropic drying, alongside reference structure.



APPENDIX L.3

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide
DART Mass Spectrum

Chemical analysis of product following azeotropic drying, alongside reference structure.



Elemental Composition

Parameters

Tolerance: 10.00 ppm
Electron: Odd/Even
Charge: +1
DBE: -20.5 - 100.0

Elements Set 1:

Symbol	C	H	Br	F	O
Min	0	0	1	1	1
Max	100	100	1	1	5

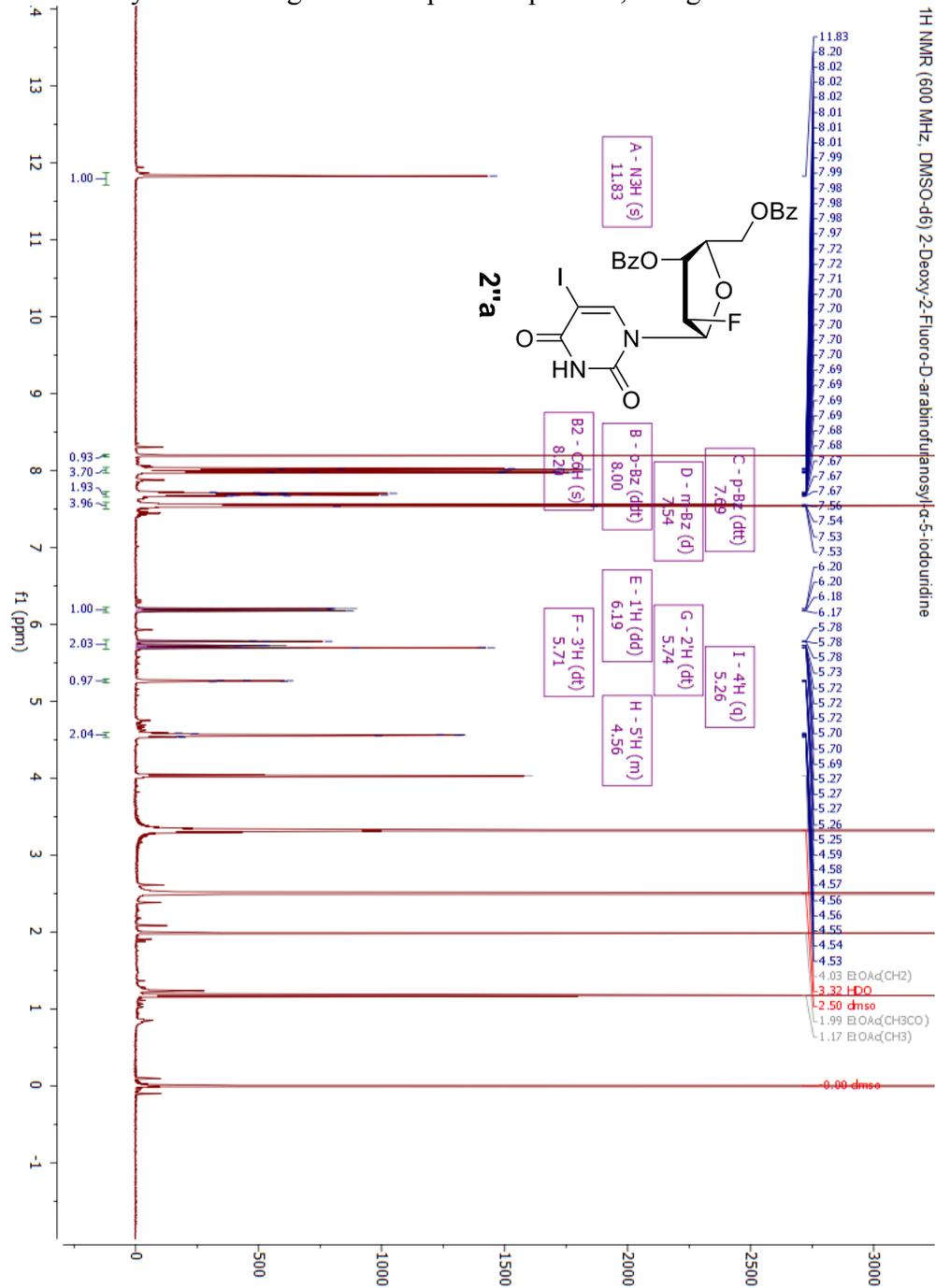
Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [ppm]	DBE
423.02267	1570386.00	C ₁₉ H ₁₇ O ₅ FBr	423.02379	-2.65	10.5

APPENDIX M.1a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum in DMSO

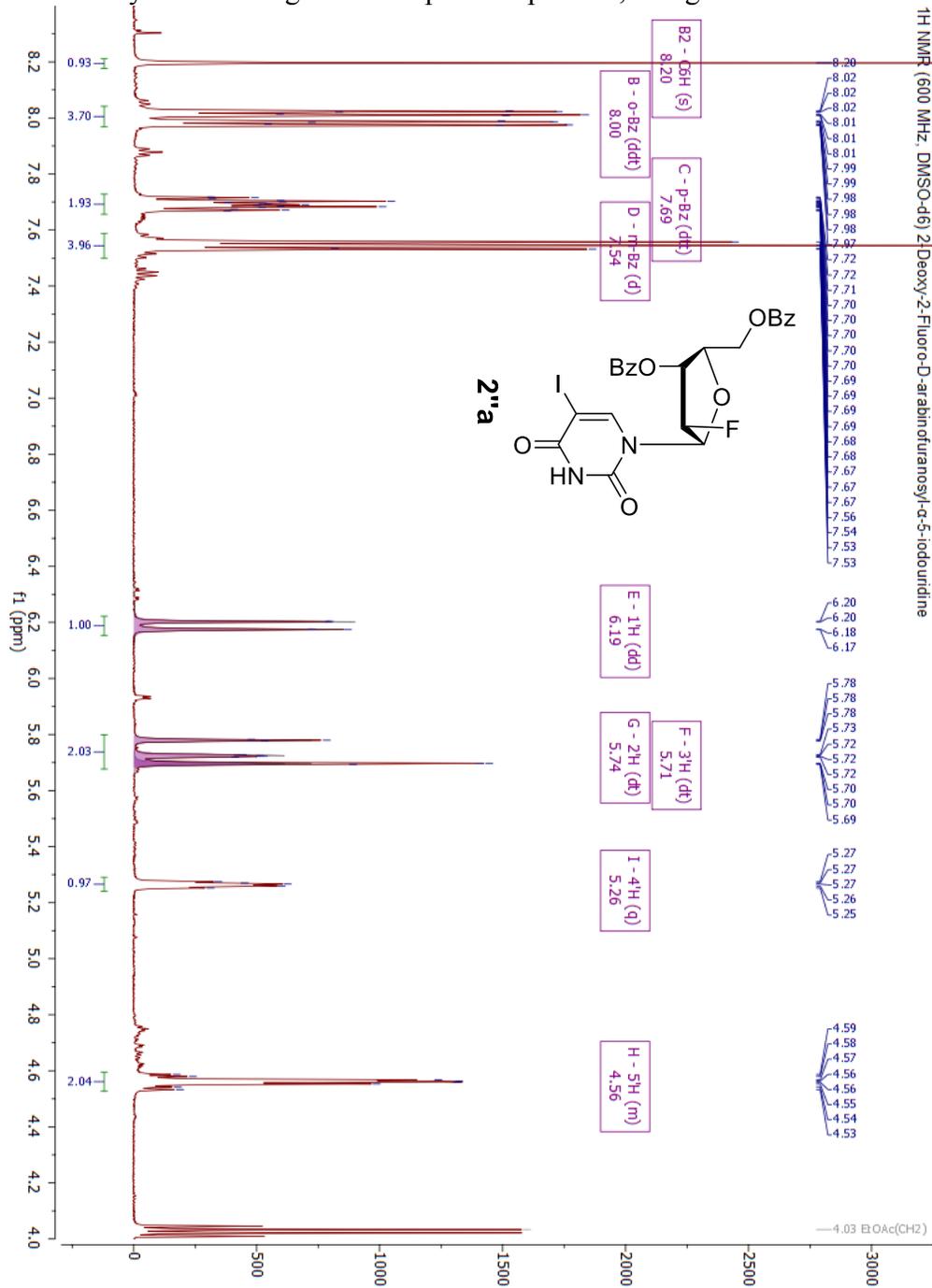
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.1b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-5-Iodouridine
 $^1\text{H-NMR}$ Spectrum in DMSO

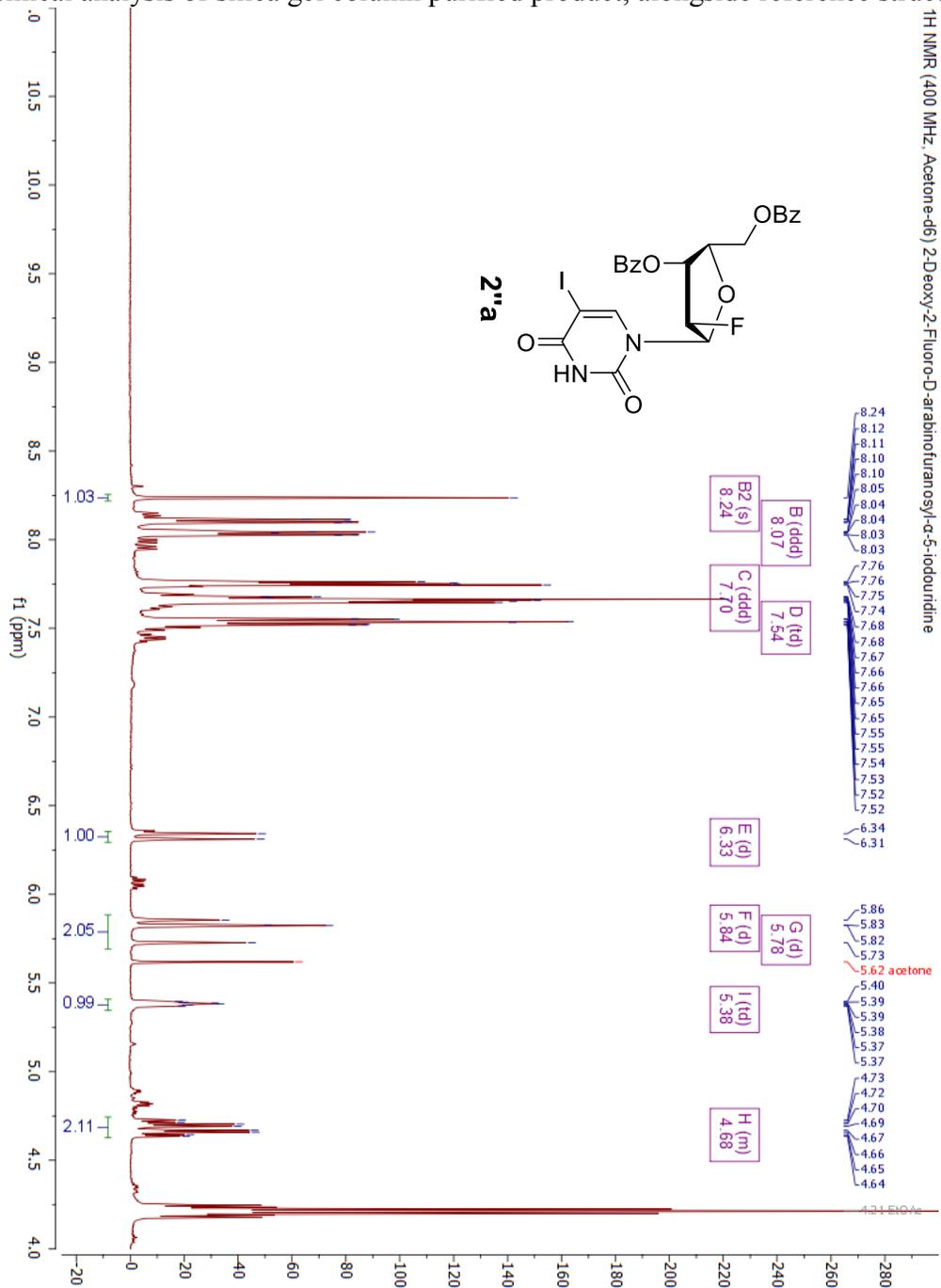
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.1c

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-5-Iodouridine
 $^1\text{H-NMR}$ Spectrum in Acetone

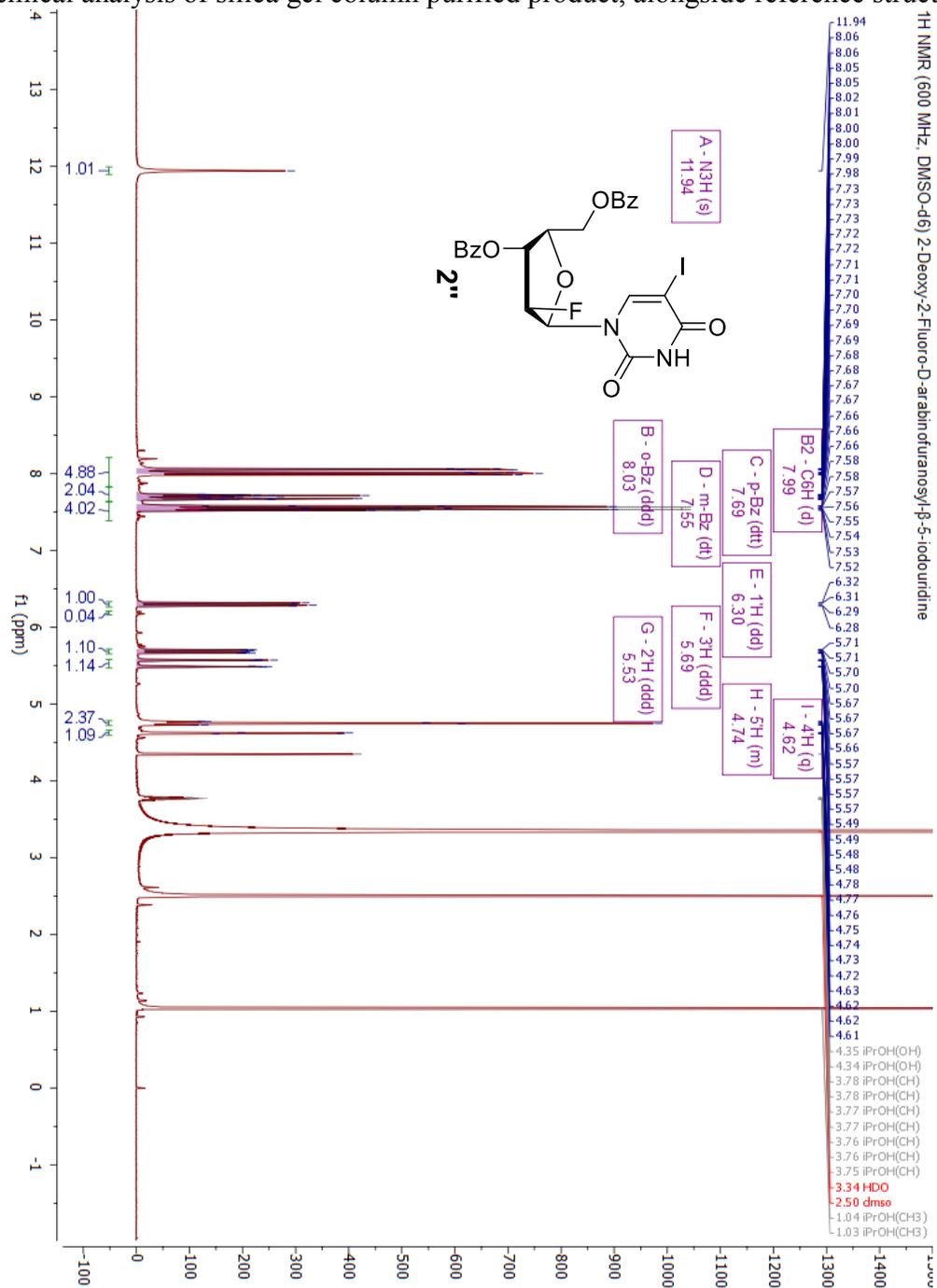
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.2a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum in DMSO

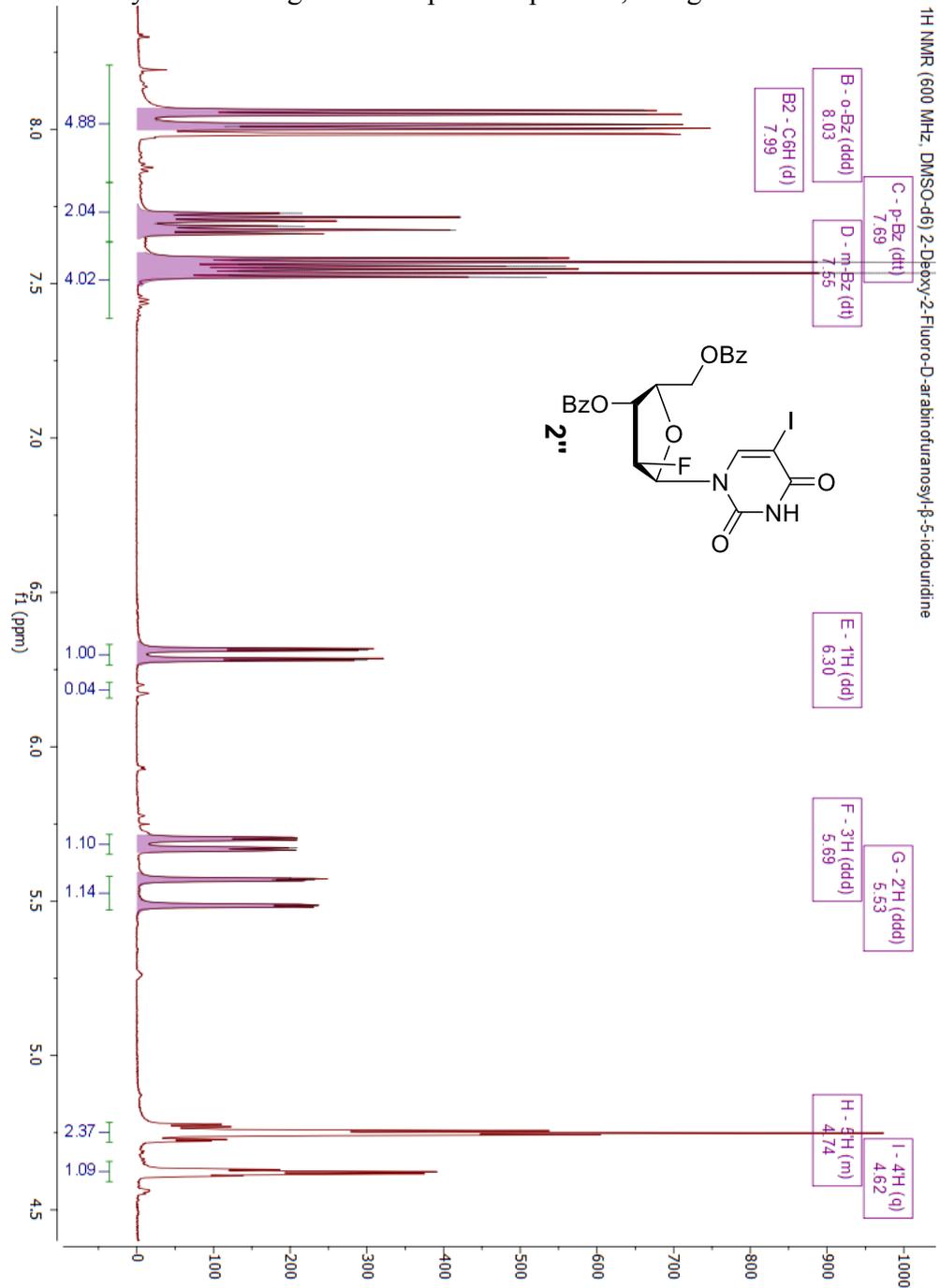
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.2b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum in DMSO

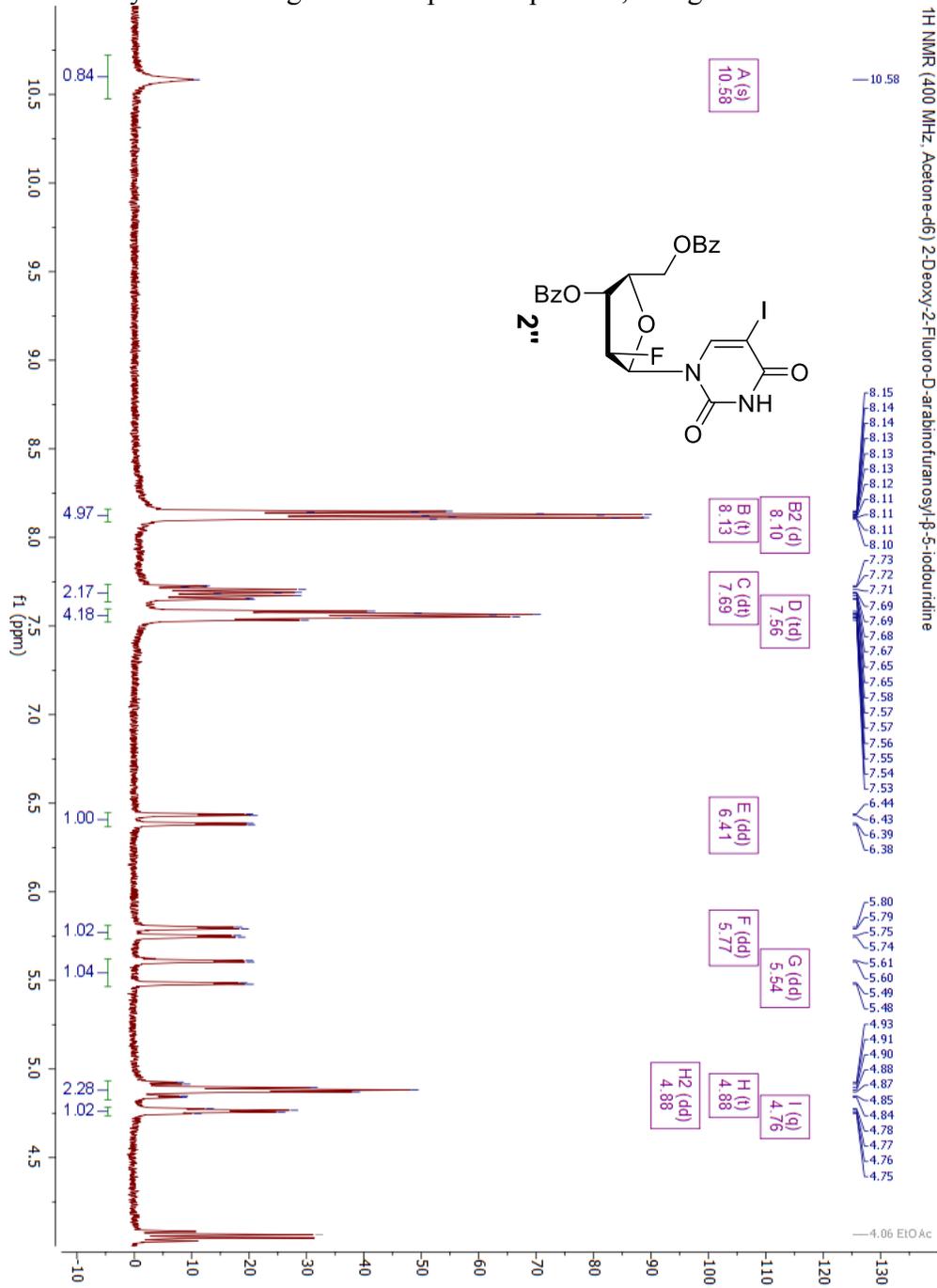
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.2c

-Deoxy-2-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum in Acetone

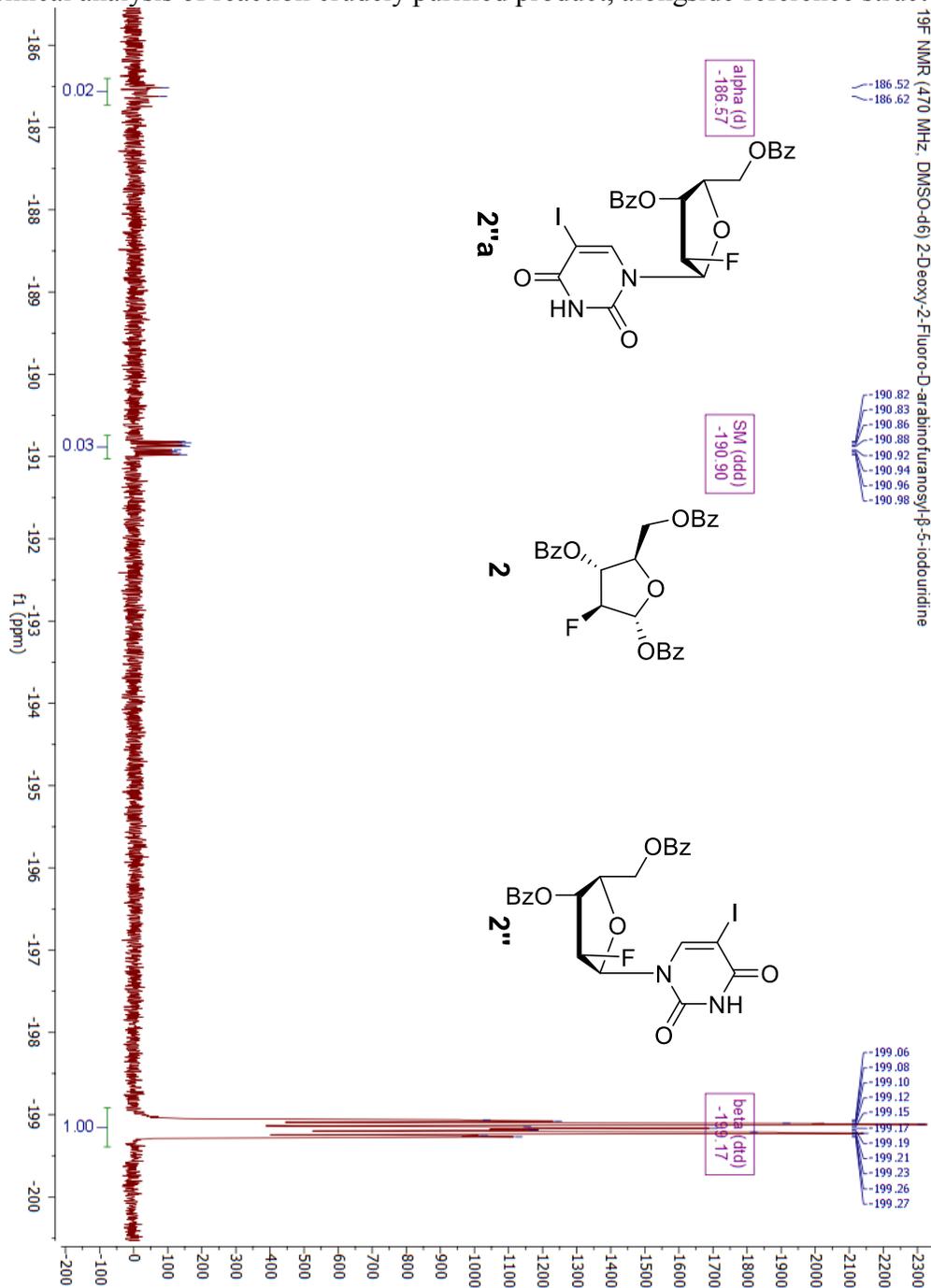
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.3a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-5-Iodouridine
¹⁹F-NMR Spectrum in DMSO

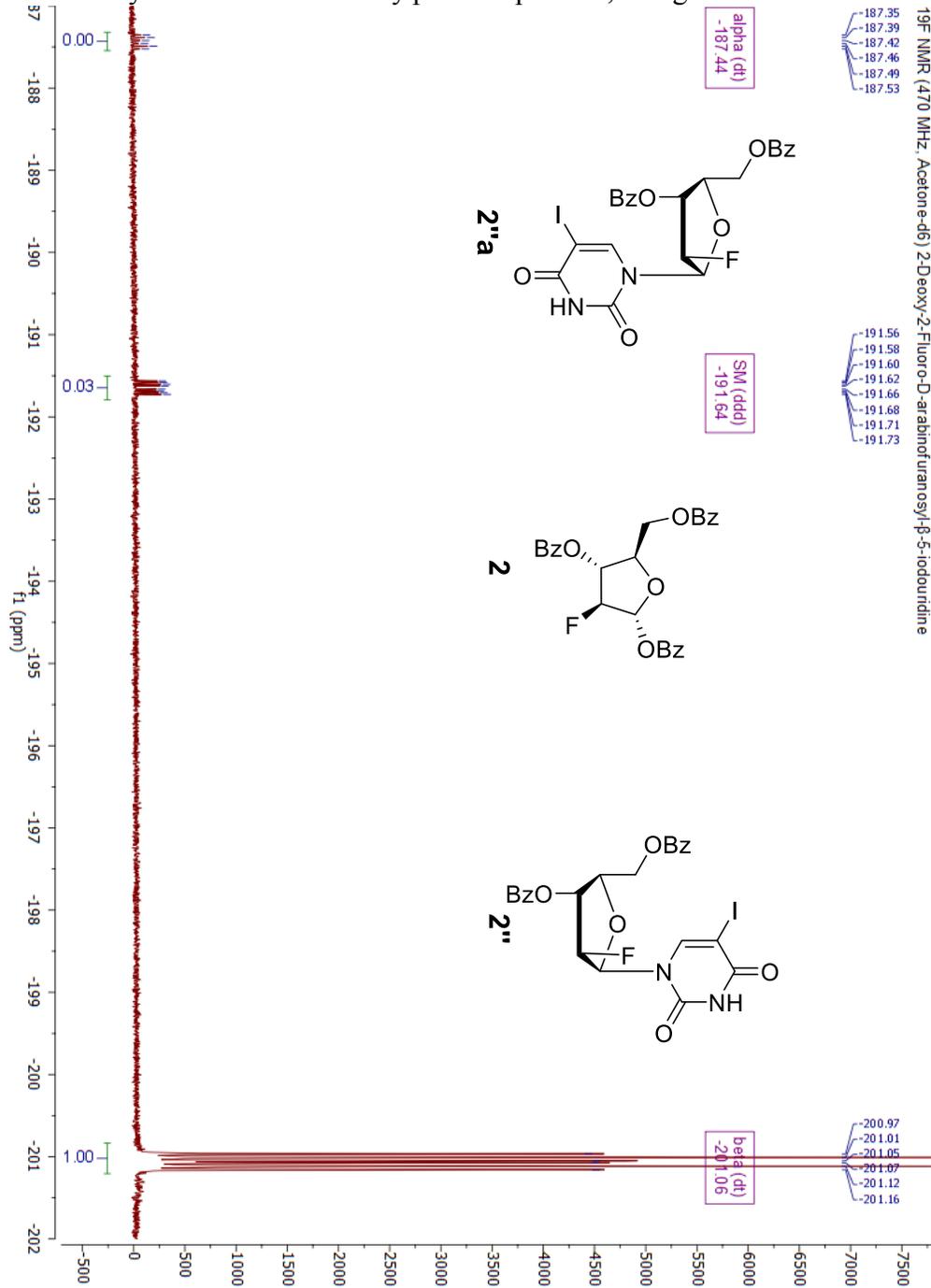
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX M.3b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-5-Iodouridine
¹⁹F-NMR Spectrum in Acetone

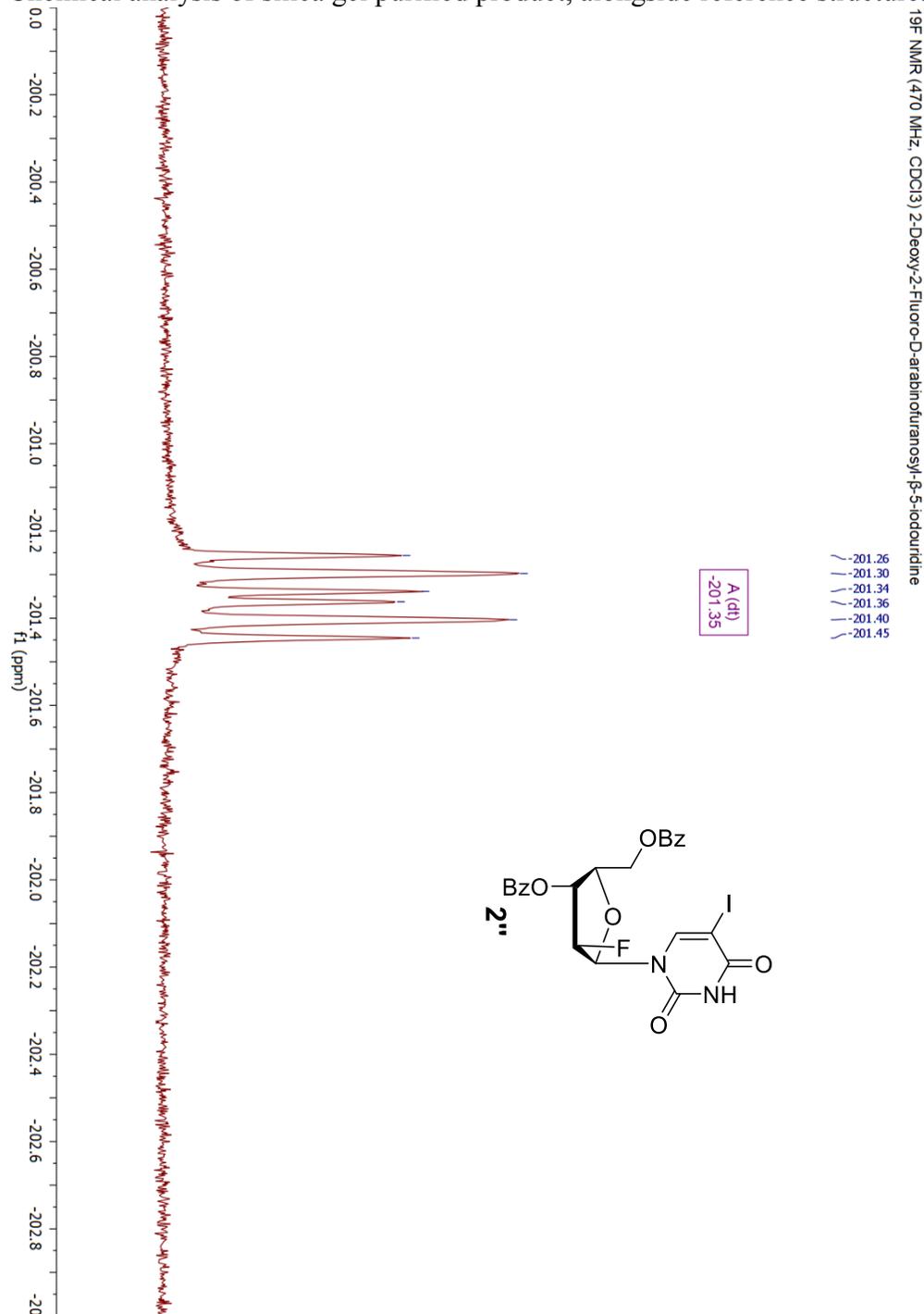
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX M.3c

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- β -D-Arabinofuranosyl-5-Iodouridine
 ^{19}F -NMR Spectrum in CDCl_3

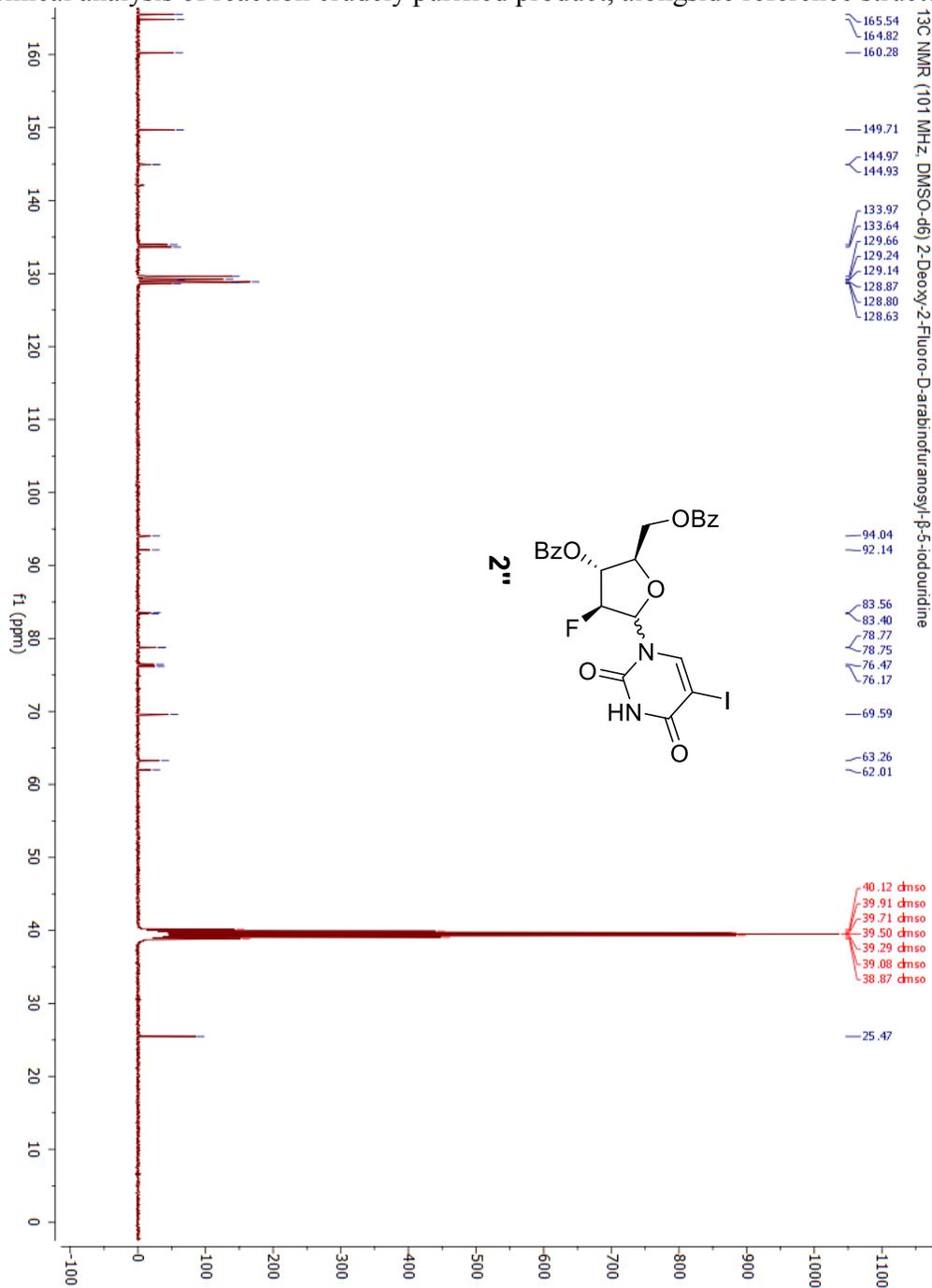
Chemical analysis of silica gel purified product, alongside reference structure.



APPENDIX M.4

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-5-Iodouridine
¹³C-NMR Spectrum in DMSO

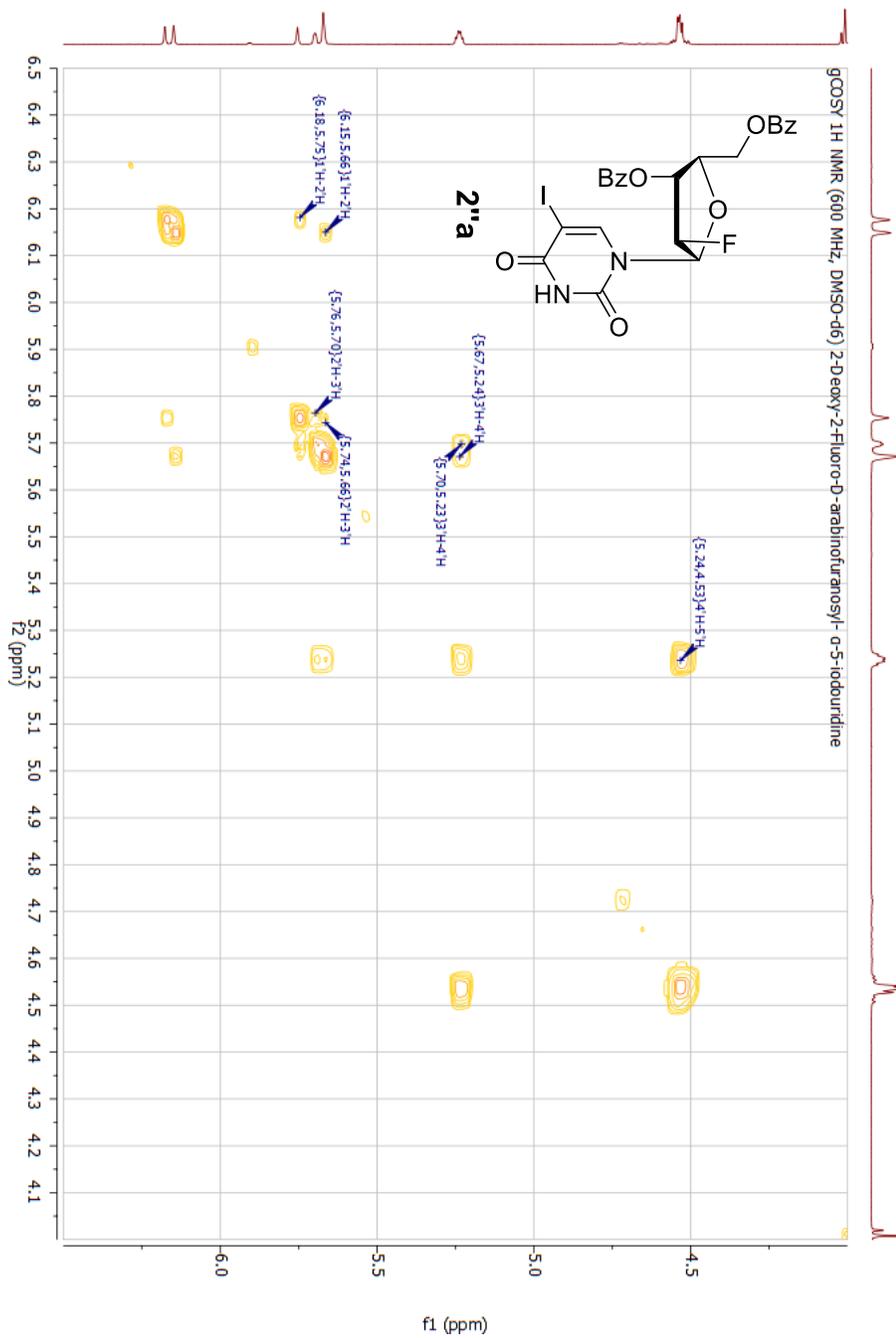
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX M.5a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-5-Iodouridine
gCOSY Spectrum in Acetone

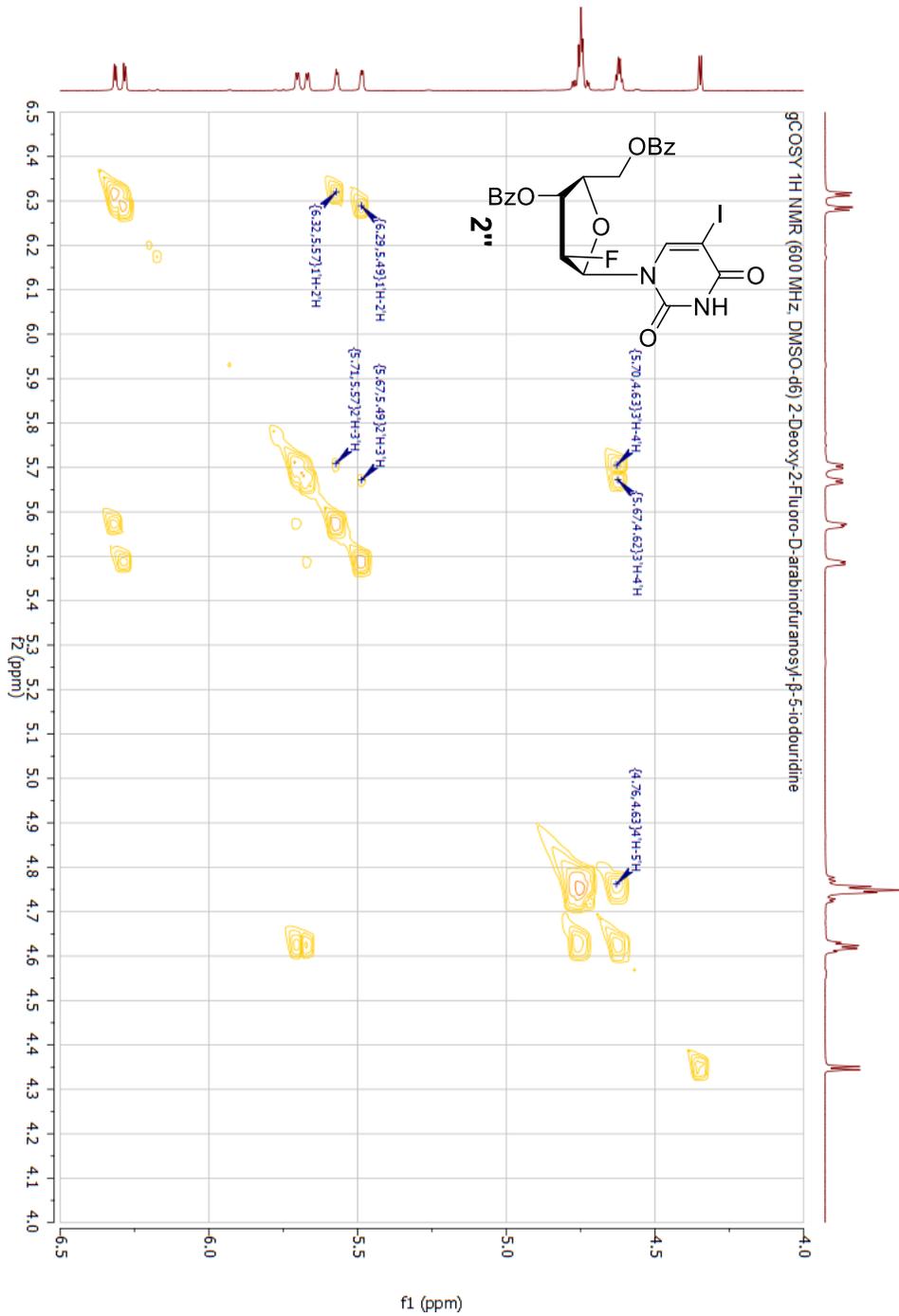
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.5b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
gCOSY Spectrum in DMSO

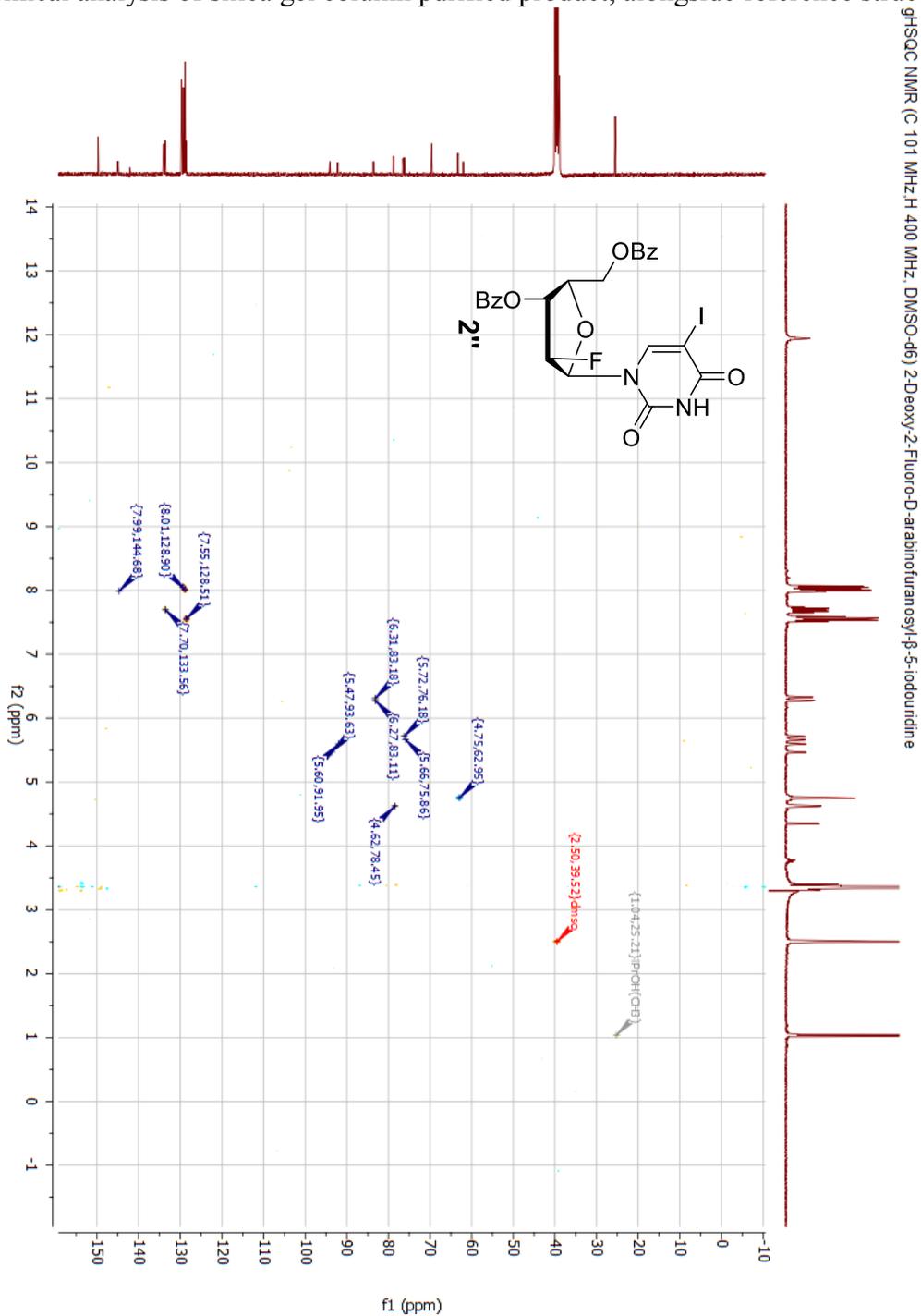
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.6

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
gHSQC Spectrum in DMSO

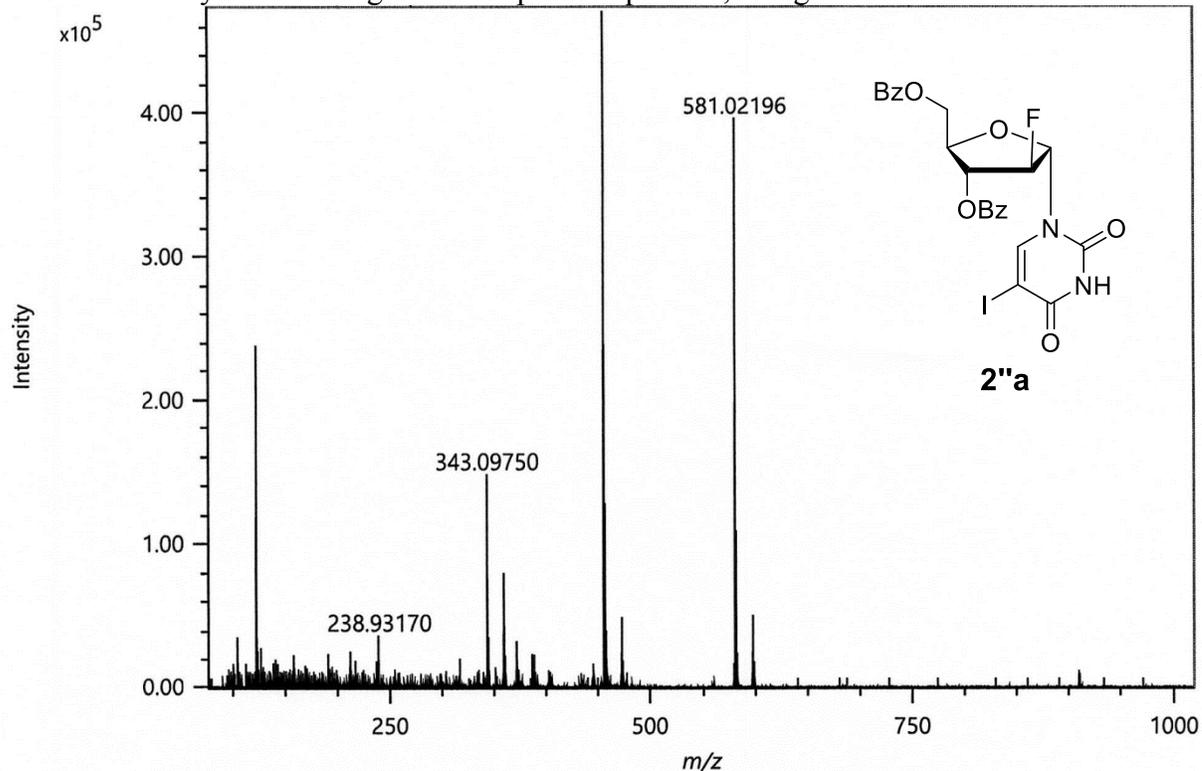
Chemical analysis of silica gel column purified product, alongside reference structure.



APPENDIX M.7a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide
DART Mass Spectrum

Chemical analysis of silica gel column purified product, alongside reference structure.



Elemental Composition

Parameters

Tolerance: 5.00 ppm
Electron: Odd/Even
Charge: +1
DBE: -20.5 - 100.0

Elements Set 1:

Symbol	C	H	F	I	N	O
Min	0	0	1	1	1	1
Max	100	100	1	1	2	7

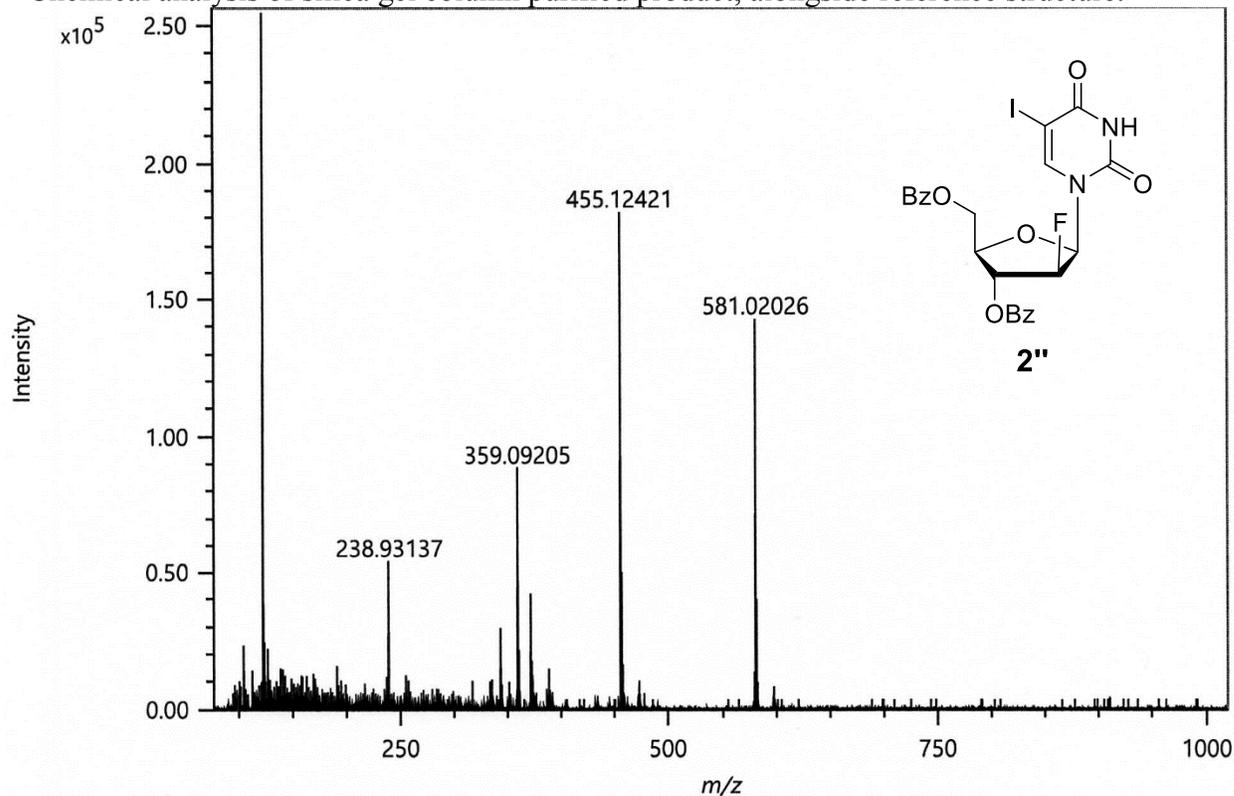
Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [ppm]	DBE
581.02196	389926.00	C ₂₃ H ₁₉ N ₂ O ₇ F I	581.02155	0.72	14.5

APPENDIX M.7b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-Bromide
DART Mass Spectrum

Chemical analysis of silica gel column purified product, alongside reference structure.



Elemental Composition

Parameters

Tolerance: 5.00 ppm
Electron: Odd/Even
Charge: +1
DBE: -20.5 - 100.0

Elements Set 1:

Symbol	C	H	F	I	N	O
Min	0	0	1	1	1	1
Max	100	100	1	1	2	7

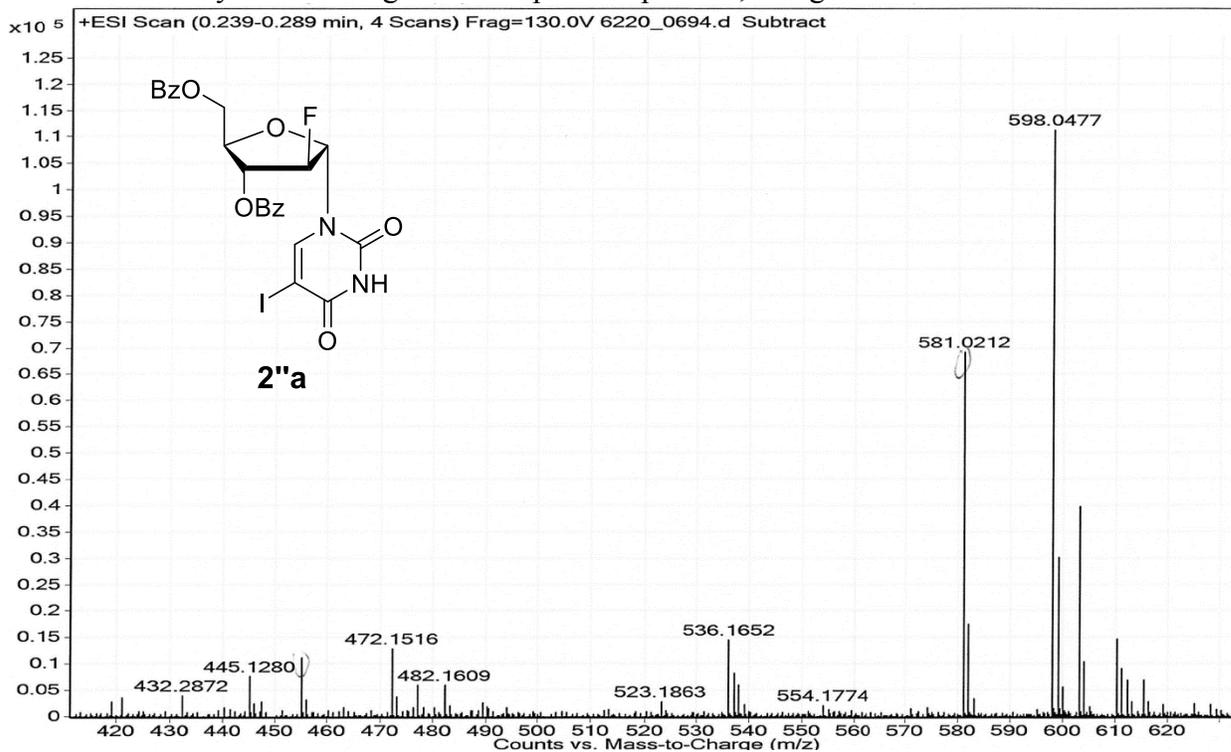
Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [ppm]	DBE
581.02026	139313.76	C ₂₃ H ₁₉ N ₂ O ₇ F I	581.02155	-2.22	14.5

APPENDIX M.8a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl- α -D-Arabinofuranosyl-Bromide
ESI Mass Spectrum

Chemical analysis of silica gel column purified product, alongside reference structure.



Element Limits:

C 1/100 H 1/100 F 1/1 I 1/1 N 1/2 O 1/7

Tolerance: 5.00 ppm

Low error bound (mmu): 5.0 (for masses < 1000)

High error bound (mmu): 20.0 (for masses > 4000)

Even or odd electron ion or both: BOTH

Minimum unsaturation: -5.0

Maximum unsaturation: 100.0

Elemental Composition Calculation

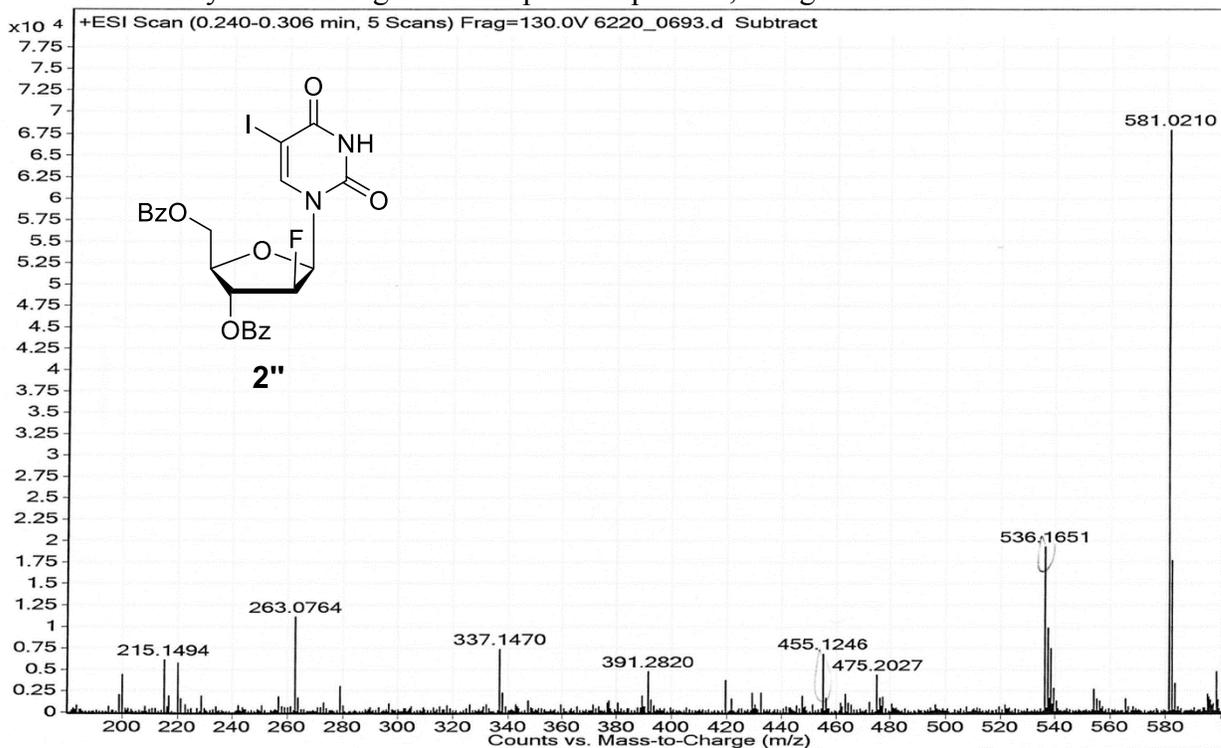
9/7/2018 9:11:29 AM

Meas. mass	Abund. %	Diff. ppm	Unsat.	Compositions
u				
581.021179	0.00	-1.61	14.5	C23 H19 F1 I1 N2 O7
		8.47	23.5	C30 H15 F1 I1 N2 O2

APPENDIX M.8b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-Bromide
ESI Mass Spectrum

Chemical analysis of silica gel column purified product, alongside reference structure.



Element Limits:
C 1/100 H 1/100 F 1/1 I 1/1 N 1/2 O 1/7
Tolerance: 5.00 ppm
Low error bound (mmu): 5.0 (for masses < 1000)
High error bound (mmu): 20.0 (for masses > 4000)
Even or odd electron ion or both: BOTH
Minimum unsaturation: -5.0
Maximum unsaturation: 100.0

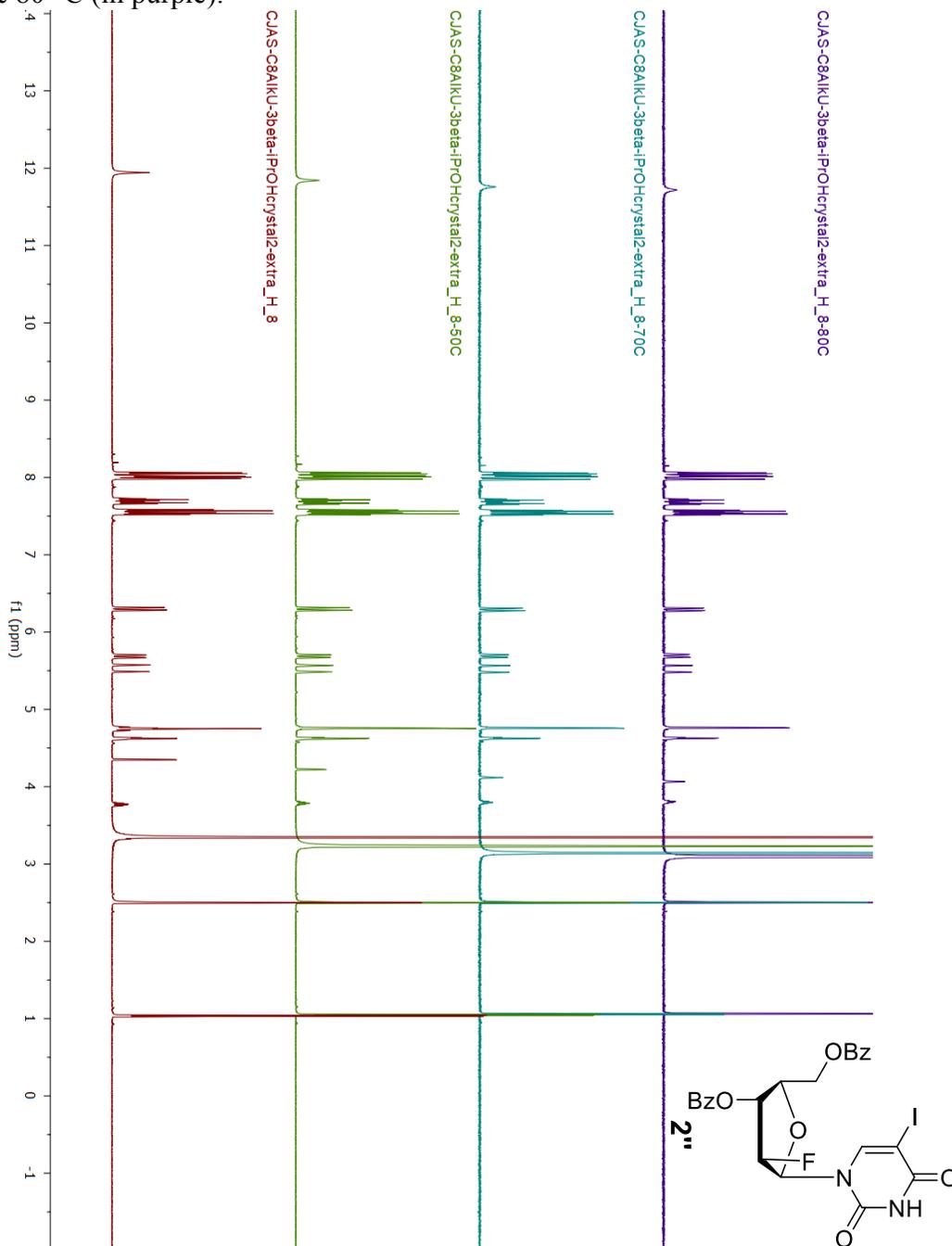
Elemental Composition Calculation
9/7/2018 9:09:10 AM

Meas. mass	Abund.	Diff.	Unsat.	Compositions
u	%	ppm		
581.020996	0.00	-1.92	14.5	C23 H19 F1 I1 N2 O7
		8.15	23.5	C30 H15 F1 I1 N2 O2

APPENDIX M.9a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum Temperature Study in DMSO

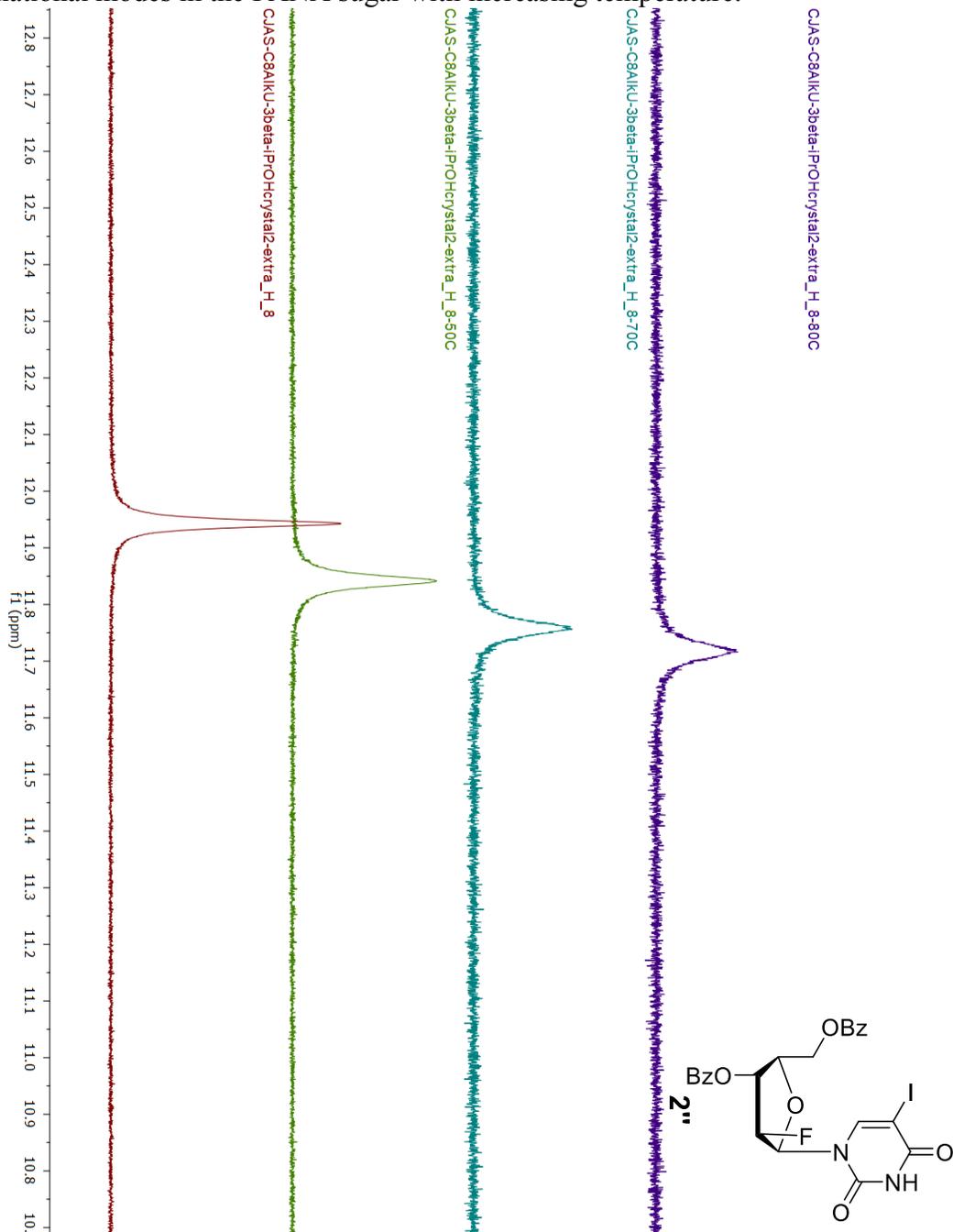
Investigation of C6H doublet coupling to 5;H in only β isomer, alongside reference structure. Temperature modulated (left to right) 20 °C (in red), 50 °C (in green), 70 °C (in blue), & 80 °C (in purple).



APPENDIX M.9b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum Temperature Study in DMSO

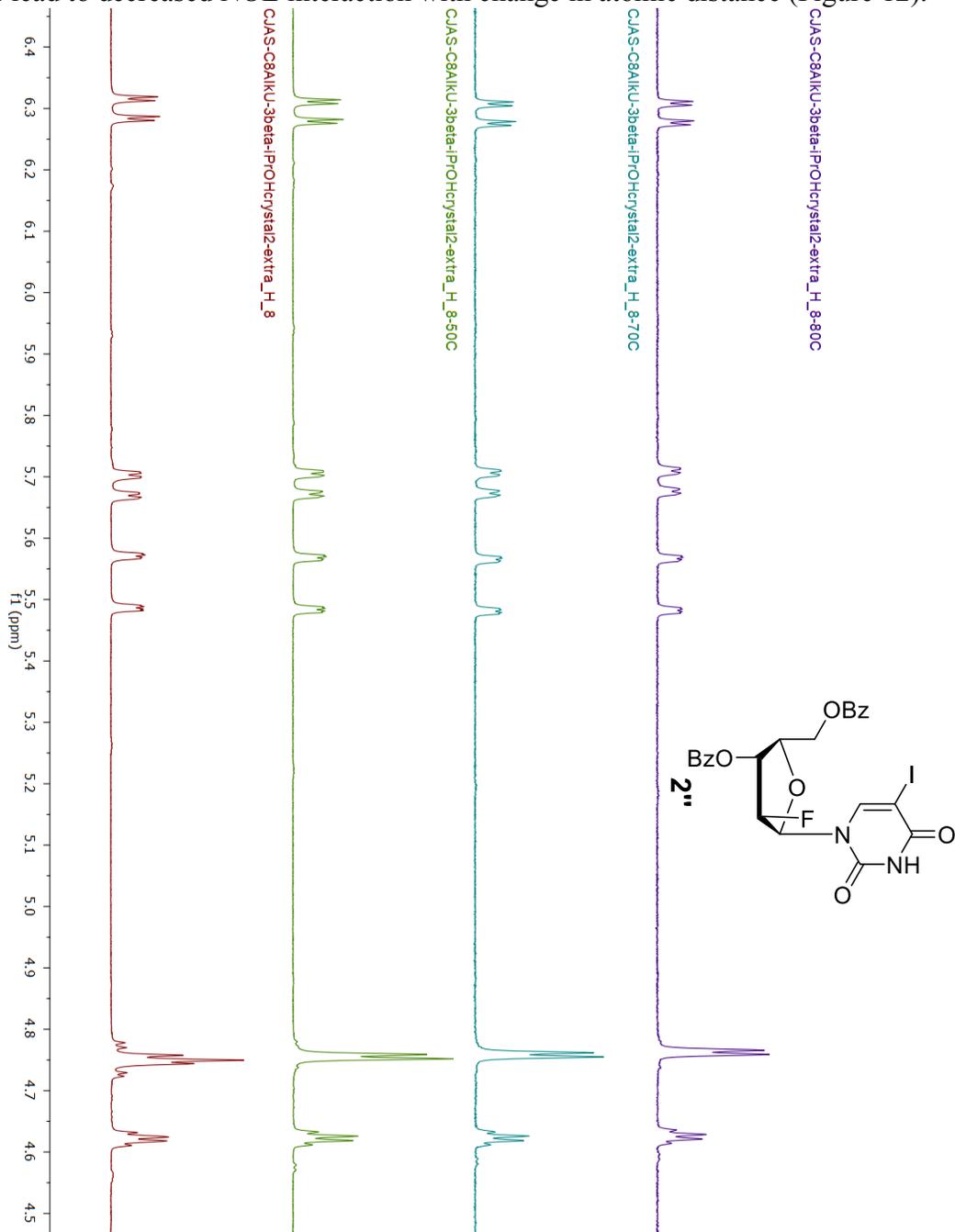
Investigation of C6H doublet coupling to 5;H in only β isomer, alongside reference structure. N3H upfield shifting indicating increased shielding likely due to change in conformational modes in the FANA sugar with increasing temperature.



APPENDIX M.9c

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum Temperature Study in DMSO

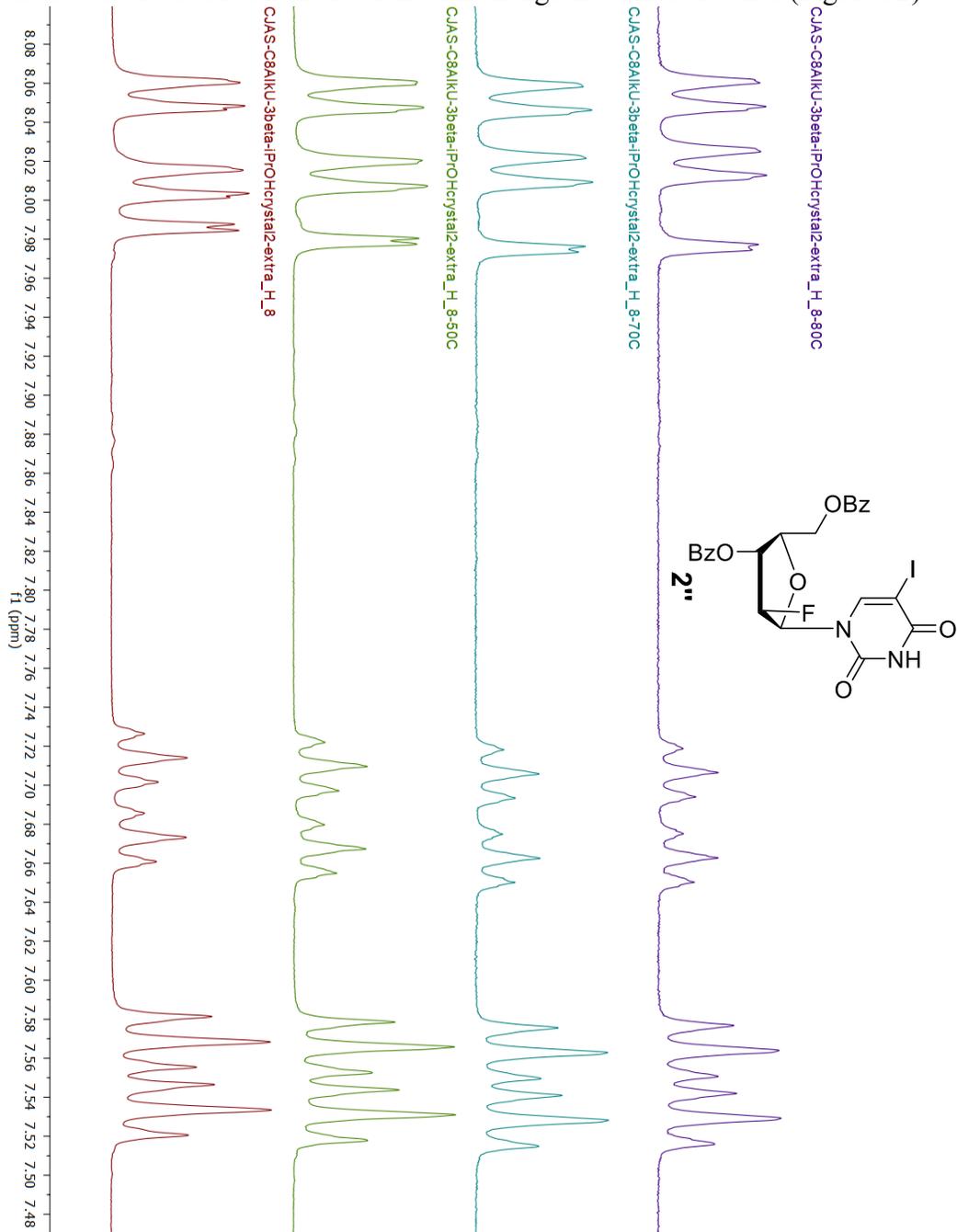
Investigation of C6H doublet coupling to 5_βH in only β isomer, alongside reference structure. 5' H coupling changes from multiplet into doublet as proposed conformational changes lead to decreased NOE interaction with change in atomic distance (Figure 12).



APPENDIX M.9d

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-β-D-Arabinofuranosyl-5-Iodouridine
¹H-NMR Spectrum Temperature Study in DMSO

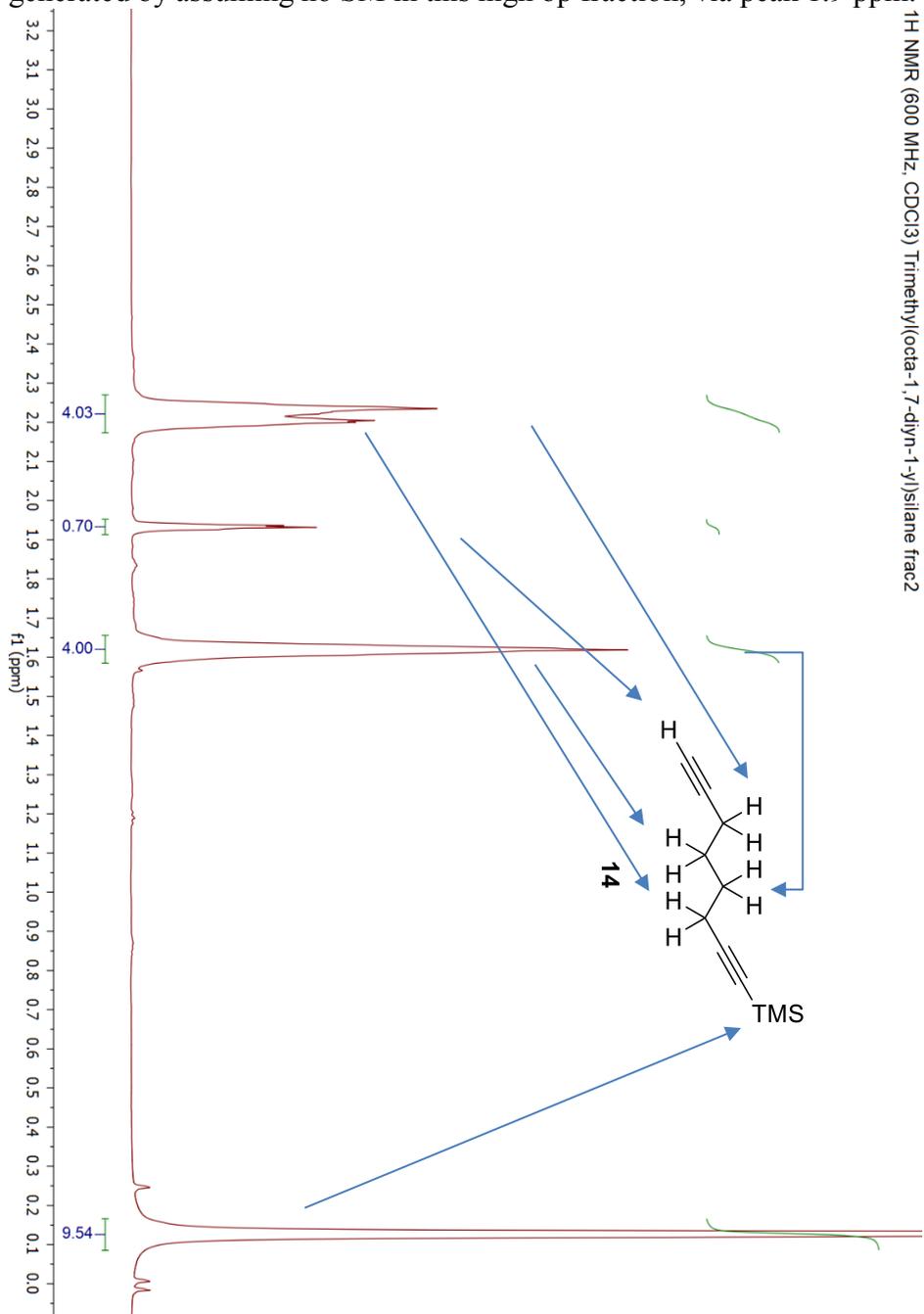
Investigation of C6H doublet coupling to 5;H in only β isomer, alongside reference structure. C6H coupling changes from doublet into singlet as proposed conformational changes lead to decreased NOE interaction with change in atomic distance (Figure 12).



APPENDIX N

Trimethyl(octa-1,7-diyn-1-yl)silane
¹H-NMR Spectrum in CDCl₃

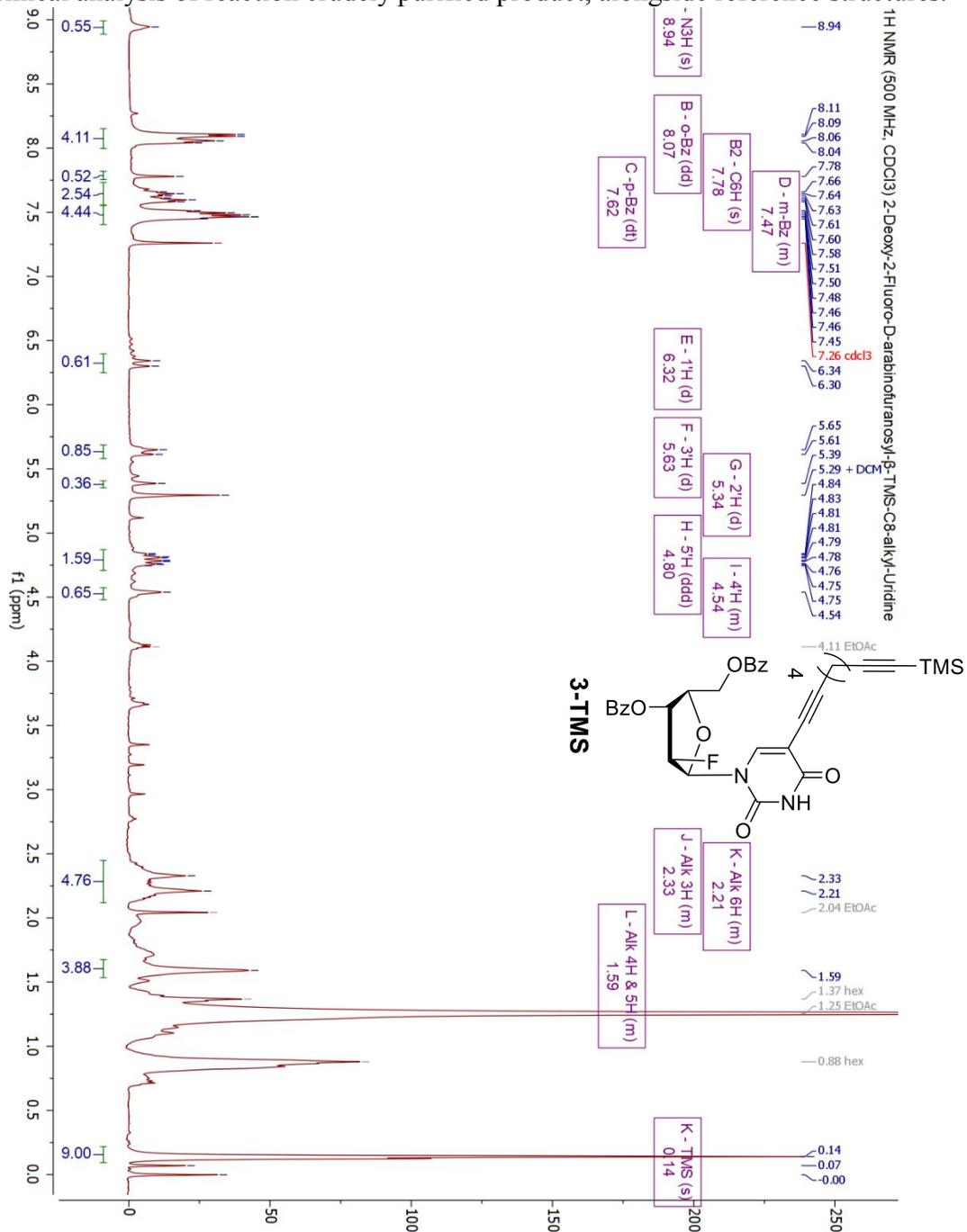
Spectrum of distillation fraction, indicating mixture of product and byproducts. Integrands below indicate molar ratio due to each proton (arrows) Data in Figure 13 generated by assuming no SM in this high bp fraction, via peak 1.9 ppm.



APPENDIX O.1a

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-TMS-C8-alkyl-Uridine
¹H-NMR Spectrum in CDCl₃

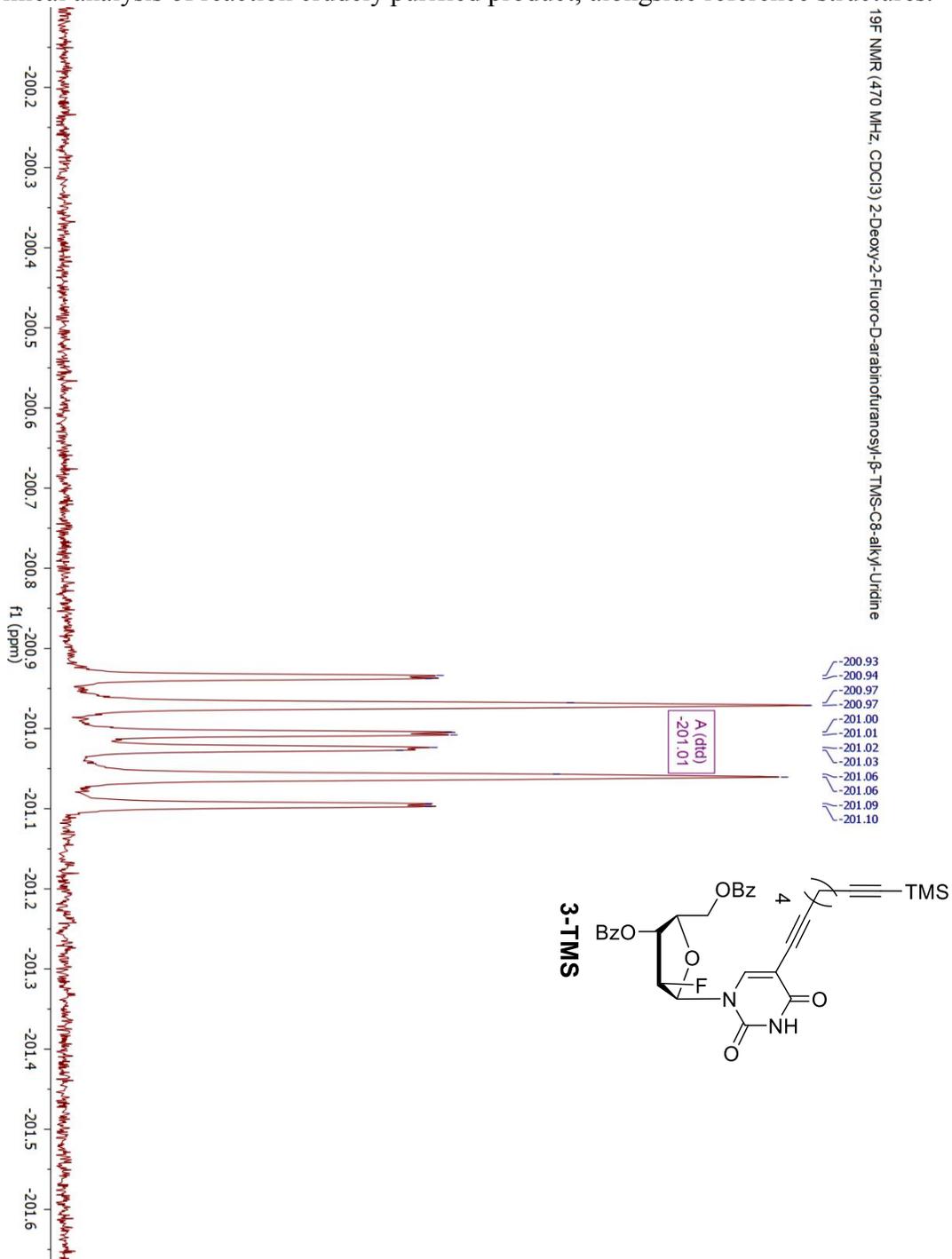
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX O.1b

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-TMS-C8-alkyl-Uridine
¹⁹F-NMR Spectrum in CDCl₃

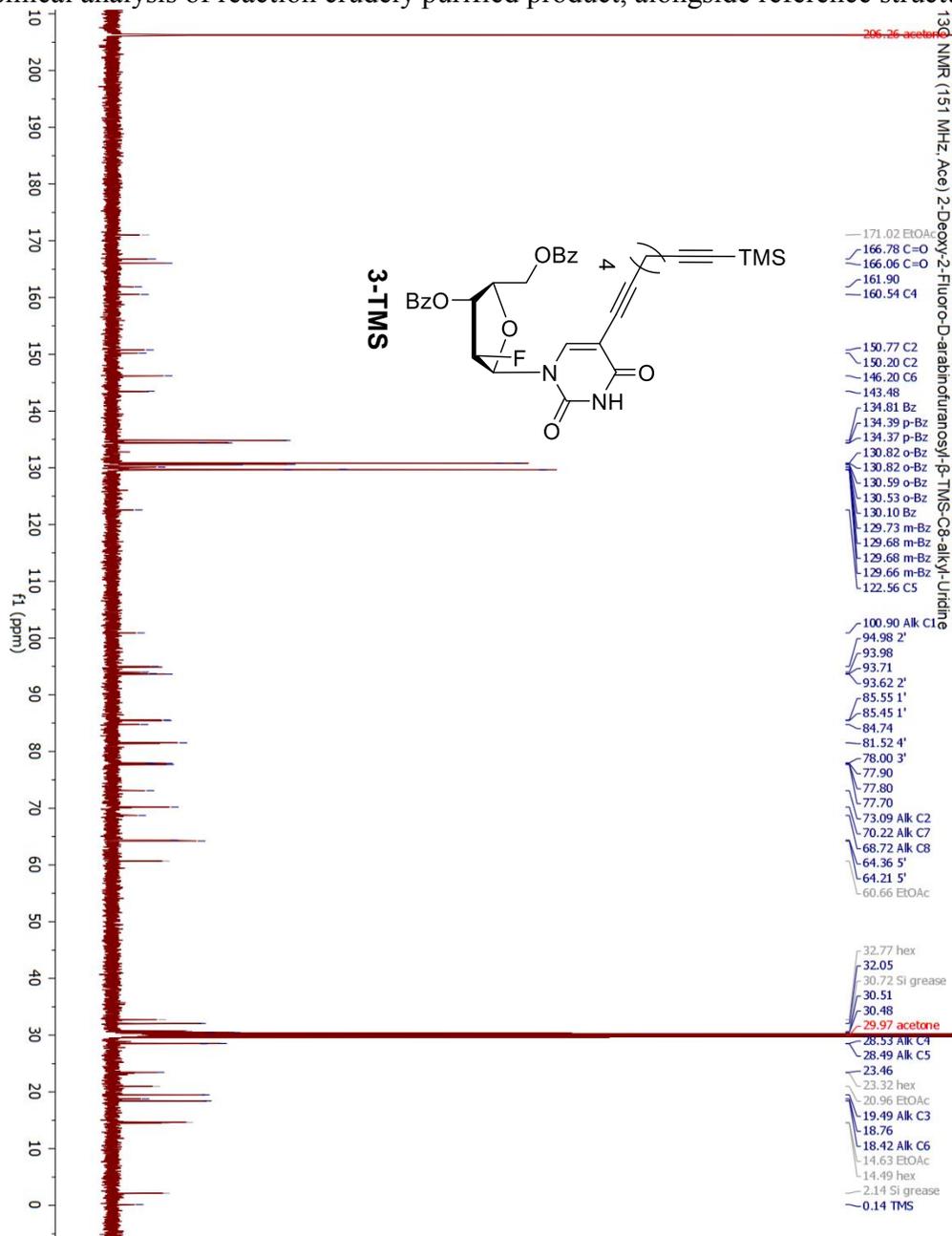
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX O.1c

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-TMS-C8-alkyl-Uridine
¹³C-NMR Spectrum in Acetone

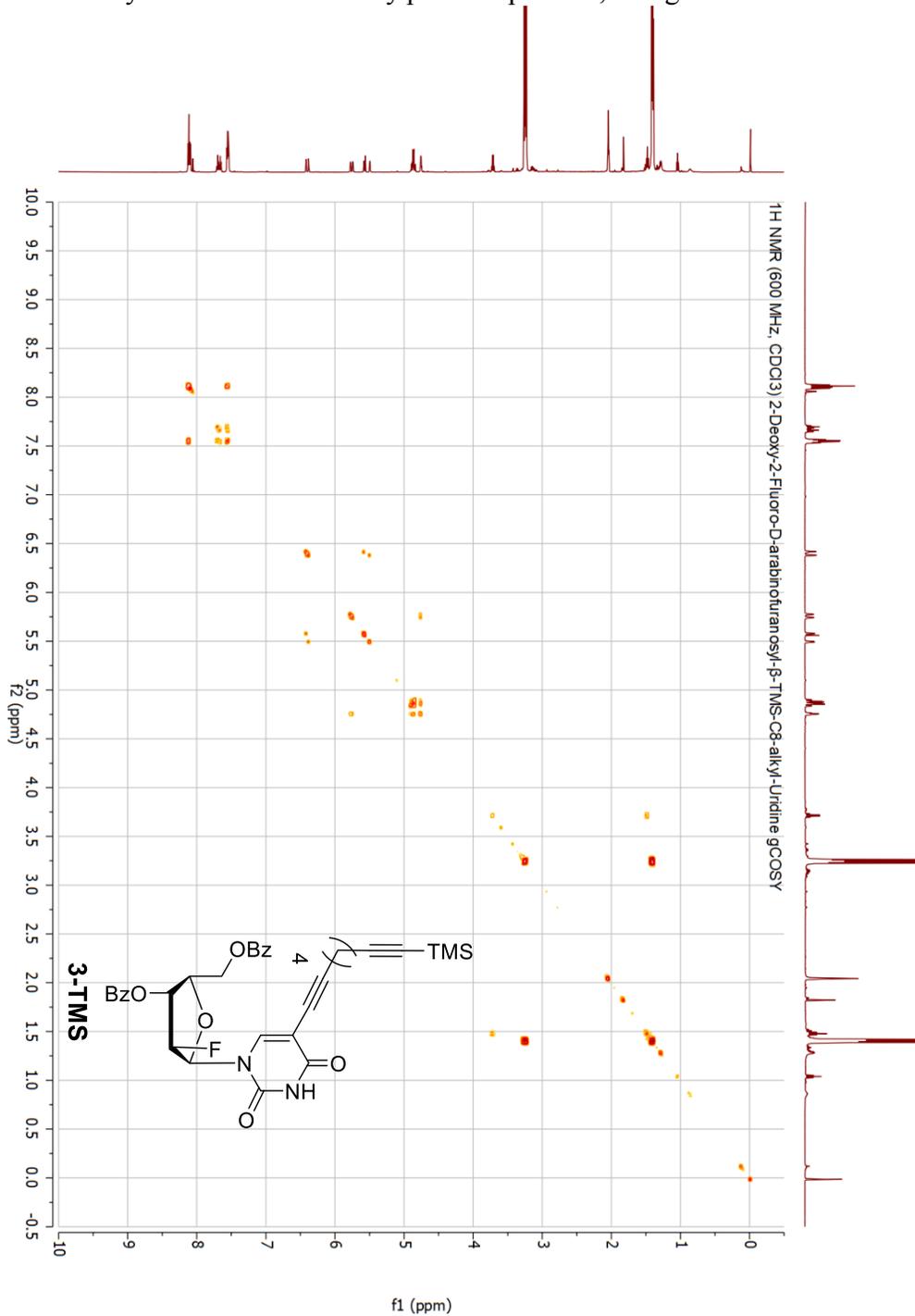
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX O.1d

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-TMS-C8-alkyl-Uridine
gCOSY Spectrum in CDCl₃

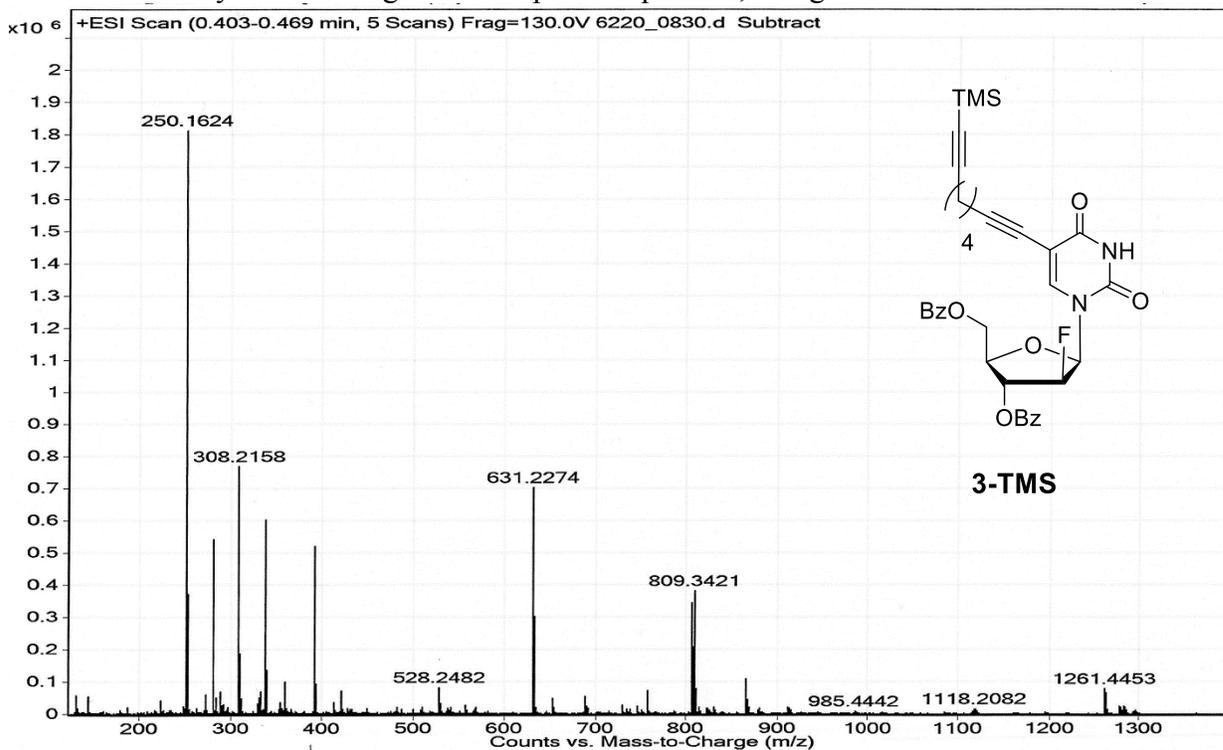
Chemical analysis of reaction crudely purified product, alongside reference structures.



APPENDIX O.1e

2'-Deoxy-2'-Fluoro-3,5-di-O-benzoyl-D-Arabinofuranosyl-β-TMS-C8-alkyl-Uridine
ESI Mass Spectrum

Chemical analysis of silica gel column purified product, alongside reference structure.



Element Limits:
C 1/120 H 1/120 F 1/1 N 1/2 O 1/7 Si 1/1
Tolerance: 5.00 ppm
Low error bound (mmu): 5.0 (for masses < 1000)
High error bound (mmu): 20.0 (for masses > 4000)
Even or odd electron ion or both: BOTH
Minimum unsaturation: -50.0
Maximum unsaturation: 100.0

Elemental Composition Calculation
11/30/2018 10:09:09 AM

Meas. mass u	Abund. %	Diff. ppm	Unsat.	Compositions
631.227417	0.00	-0.29	18.5	C34 H36 F1 N2 O7 Si1

