

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Kyle E. Giesler

Date

The Design, Synthesis, and Biological Evaluation of Novel Prodrugs and Antiviral Agents for HIV
and Other Chronic Viral Infections

By

Kyle E. Giesler

Doctor of Philosophy

Chemistry

Dennis Liotta, PhD

Advisor

Khalid Salaita, PhD

Committee Member

Frank McDonald, PhD

Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

The Design, Synthesis, and Biological Evaluation of Novel Prodrugs and Antiviral Agents for HIV
and Other Chronic Viral Infections

By

Kyle E. Giesler

B.S., University of Central Florida

Advisor: Dennis C. Liotta, PhD

An abstract of

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry

2017

Abstract

The Design, Synthesis, and Biological Evaluation of Novel Prodrugs and Antiviral Agents for HIV and Other Chronic Viral Infections

By Kyle E. Giesler

Drug delivery is a critical challenge that must be overcome to achieve therapeutic efficacy *in vivo*. Many drug discovery programs fail to identify targets with desirable absorption, distribution, metabolic, and excretion (ADME) profiles which severely limits their clinical prospect. The conjugation of a prodrug to a pharmacologically active agent provides an opportunity to overcome these limitations and significantly alter the pharmacological profile of the parent. With respect to nucleoside analogues, a large repertoire of prodrugs exists to improve their biological availability, however, many of these prodrugs generate toxic byproducts, experience inefficient cleavage, or are subject to premature degradation prior to delivery of the target. Herein, we have designed and synthesized novel lipid prodrugs for nucleoside analogues that significantly improve the antiviral activity of the parent, are stable in human plasma for 24 h and have the potential to achieve enhanced tissue distribution *in vivo*. Further, we prepared a series of promising acyclic nucleoside analogues for the treatment of chronic HIV-1 and are actively in the process of targeting other HIV proteins, particularly negative factor (Nef), to combat resistance and introduce novel therapies to market.

The Design, Synthesis, and Biological Evaluation of Novel Prodrugs and Antiviral Agents for HIV
and Other Chronic Viral Infections

By

By Kyle E. Giesler
B.S., University of Central Florida

Advisor: Dennis C. Liotta, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Chemistry

2017

Acknowledgements

First and foremost, I would like to thank Dr. Liotta for providing me the intellectual freedom to pursue problems I found interesting and supporting my (many) ambitious ideas. Despite his busy schedule, he was always willing to sit down with me to discuss new ideas, listen to my concerns, and hear me out when I was experiencing problems in the lab. His hands off approach allowed me to discover things for myself and facilitated significant professional and personal growth for which I am grateful. I would also like to thank the late Dr. Snyder for our engaging conversations that always left me with a sense of peace and clarity once I left his office. He was truly a source of infinite wisdom with a calm and collective demeanor and a flare of tenacity that I deeply admired. Dr. Snyder and the entire computational laboratory really helped me during my first year at Emory as a new graduate student finding my path and put me on the road to success that ultimately culminated with this dissertation. My committee members Khalid Salaita and Frank McDonald are true academics at heart and always ask profound and detail-oriented questions that have made me a better critical thinker and a better scientist in general. My peers Brooke Katzman and Kristen Stoltz have become dear friends of mine that were there for me in my lowest of lows when I didn't think I was going to make it through the program. To all my friends I've made here in Atlanta-thank you. You have enriched my life in ways that I cannot begin to describe. We've spent countless nights at Ten and Blake's on the Park, shared thousands of laughs (often at each other's expense), bonded at summer pool parties, escaped to the cool Appalachian Mountains, danced all night at the club, played video games over boxed wine, and unified together against adversity during Atlanta Pride. These experiences and many others shaped me into the person I am today and helped stave off depression and frustration throughout the years. I am truly thankful for your friendship and will never forget the memories we have shared. Finally, I would like to thank my family for their unconditional support; Mom, Dad, David, Paige, and Wilson. I hope I have made you proud.

Table of Contents

List of Illustrations

Figures

Tables

Schemes

List of Abbreviations

Chapter 1: Novel Lipid Prodrugs of Tenofovir: Design, Synthesis, and Biological Evaluation

1.1	Introduction	2
1.1.1	Nucleoside Analogues as Antiviral Drugs	4
1.1.2	Phosphorylation is Required for Metabolic Activation of NAs	5
1.1.3	Kinase Bypass Strategies: Phosphonates	8
1.1.4	Tenofovir and the Use of Prodrugs	9
1.1.5	Tenofovir Alafenamide (TAF)	11
1.1.6	CMX157	12
1.2	The Design of Novel Lipid TFV Conjugates	13
1.3	Synthesis of 1 st Gen TFV Disulfide Conjugates	15
1.4	Biological Evaluation of 1 st Gen TFV Disulfide Prodrug Conjugates	17
1.5	Stability Studies for our 1 st Generation Conjugates	20
1.6	Probing the Mechanism of Cleavage	21
1.7	Second Generation Disulfide Conjugates	23
1.8	Revisiting Our 1 st Generation Conjugates	26
1.9	What is the Disulfide Advantage?	28
1.10	Conclusions	30
1.11	Experimental Details	31
1.11.1	Anti-HIV Assay	31
1.11.2	Anti-HBV Assay	32
1.11.3	Cytotoxicity Studies	33
1.11.4	Stability Studies	33
1.11.5	Kinetic Studies	34
1.11.6	Chemical Synthesis and General Procedures	35
1.11.7	Synthesis and Characterization of Lipids	37
1.11.8	Synthesis and Characterization of TFV Conjugates	46
1.12	References	71

Chapter 2: Synthesis and Evaluation of Other Disulfide Prodrug Conjugates

2.1	Introduction	83
2.2	Synthesis of Phosphate Disulfide Conjugates	84
2.2.1	Preparation of Phosphoramidite Coupling Reagents	85
2.2.2	Lipid Conjugates of Nucleoside Analogue 2.1	86

2.2.3	Lipid Conjugates of Emtricitabine (FTC)	88
2.2.4	Lipid Conjugates of Nucleoside Analogue 2.2	89
2.2.5	Lipid Conjugates of AFV	90
2.3	Biological Evaluation of Phosphate Lipid Prodrugs	90
2.4	Experimental Details	92
2.4.1	Anti-HIV Assay	92
2.4.2	Anti-HBV Assay	93
2.4.3	Anti-HCV Assay	94
2.4.4	Cytotoxicity Studies	95
2.4.5	Chemical Synthesis and General Procedures	95
2.4.6	Crystallographic Data	108
2.5	References	110

Chapter 3: A Green and Expedient Synthesis of Acyclic Thioaminal Nucleoside Analogues from Purine and Pyrimidine Hemiaminals

3.1	Introduction	112
3.2	Synthesis and Substrate Scope	112
3.3	Biological Evaluation	117
3.4	The Pursuit for More Interesting Analogues	118
3.5	Conclusions	120
3.6	Experimental Details	121
3.6.1	Anti-HIV Assay	121
3.6.2	Anti-HBV Assay	122
3.6.3	Anti-HSV Assay	123
3.6.4	Cytotoxicity Studies	123
3.6.5	Chemical Synthesis and General Procedures	123
3.7	Crystallographic Data	140
3.8	References	148

Chapter 4: Towards the Design and Synthesis of Novel HIV-1 Nef Inhibitors

4.1	Introduction	150
4.1.1	Structure of HIV-1 Nef	152
4.1.2	Nef and HIV-1 Pathogenesis	152
4.1.3	Mechanistic Details of Nef-Mediated MHC-I Down-regulation	153
4.2	HTVS Efforts to Identify Novel Nef Inhibitors	156
4.3	<i>In Silico</i> Design of Novel α -Helical Mimetics	159
4.4	Retrosynthetic Analysis of Compound 4.4	161
4.4.1	Synthesis of the R ¹ , R ² , and R ³ Fragments	161
4.4.2	Towards the Synthesis of Compound 4.4	163
4.5	Conclusions	164
4.6	Experimental Details	165
4.6.1	Computational Analysis	165
4.6.2	Chemical Synthesis	168

4.6.3	Crystallographic Data	181
4.7	References	186

List of Illustrations

List of Figures

Figure 1.1.	Canonical nucleosides ubiquitously found in Nature	2
Figure 1.2.	Nucleoside analogues used in the treatment of viral disease	4
Figure 1.3.	The phosphorylation of emtricitabine and other NAs requires three independent enzymes before reaching the viral polymerase to arrest viral replication	5
Figure 1.4.	Acyclic phosphonate NAs to bypass the initial phosphorylation by dNKs	9
Figure 1.5.	Prodrugs used to improve the bioavailability of NAs	10
Figure 1.6.	Evolution of alkoxyalkyl prodrugs and the structure of CMX157	13
Figure 1.7.	Concept design for our disulfide prodrug strategy	14
Figure 1.8.	Decomposition profile of 1.6c and 1.6d at 37 °C in human plasma, sodium carbonate/bicarbonate buffer (pH 9.0), Dulbecco's modified eagle medium (DMEM) without fetal bovine serum, and phosphate-buffered saline (PBS, pH 7.4)	20
Figure 1.9.	The decomposition profile of compound 1.8 at various pH values obeys first order kinetics with respect to [1.8] in aqueous media	22
Figure 1.10.	Decomposition profile of 1.10 in various buffer solutions containing 0.1 M DTT at 37.4 °C	25
Figure 1.11.	Modified cleavage mechanism for 1.9 and related analogues	26
Figure 2.1.	Target nucleosides for disulfide prodrug conjugation	83
Figure 2.2.	Cleavage sites for phosphonate and phosphate lipid prodrugs by PLC and PLD	84
Figure 3.1.	Crystal structure of a) Purine 3.8 and b) Pyrimidine 3.22.	117
Figure 3.2.	Structural features from TFV and AFV were adapted in the design of novel thioaminal phosphonates 3.29 and 3.30	119
Figure 4.1.	Incomplete NMR overlay structure of HIV-1 Nef (PDB code: 2Nef)	152
Figure 4.2.	A canonical SYSQV signaling peptide bound to the surface of μ 1	154
Figure 4.3.	The crystal structure of the Nef-MHC- μ 1 ternary complex (PDB code: 4EN2)	155
Figure 4.4.	A close up view of the φ recognition socket occupied by A-323 of the MHC-I CD	155
Figure 4.5.	Structures of the three hit compounds from the Aldrich collection	159
Figure 4.6.	The pseudo- α -helical nature of the crystallized MHC-I CD	160
Figure 4.7.	A representative pose of a 5-6-5 imidazole-phenyl-thiazole mimetic bound to the surface of μ 1	161

List of Schemes

Scheme 1.1. Decomposition of TDF <i>in vivo</i>	10
Scheme 1.2. Decomposition of Tenofovir Alafenamide <i>in vivo</i> .	11
Scheme 1.3. Comparison of Relevant Prodrugs and Their Proposed Mechanisms of Cleavage	14
Scheme 1.4. Initial Synthesis of TFV Disulfide Conjugates	15
Scheme 1.5. Successful Synthesis of 1 st Generation TFV Disulfide Conjugates	16
Scheme 1.6. Kinetic Model Used to Study Intramolecular Cyclization for 1 st Gen Prodrugs	22
Scheme 1.7. Kinetic Model Used to Study the Potential Intramolecular Cyclization of 1.9	25
Scheme 1.8. Using Disulfide Prodrugs to Probe Enzymatic Hydrolysis	27
Scheme 2.1. General Synthesis of Phosphate NA Lipid Conjugates.	85
Scheme 2.2. Intramolecular Reduction of 2.3	86
Scheme 2.3. Attempted Lipid Conjugation of 2.1 with an acetonide PG	87
Scheme 2.4. Lipid Conjugation of 2.1 with a Phenylboronate PG	88
Scheme 2.6. HDP and Disulfide Lipid Conjugates of NA 2.2 .	89
Scheme 2.5. HDP and Disulfide Lipid Conjugates of FTC	89
Scheme 2.7. Synthesis of Compound 2.20	90
Scheme 3.1. Synthesis and Reactivity of Hemiaminals 3.1 and 3.2	113
Scheme 3.2. Synthesis of Phosphonates 3.29 and 3.30	119
Scheme 4.1. Retrosynthetic analysis of compound 4.4	161
Scheme 4.2. Synthesis of the R ¹ , R ² , and R ³ Fragments	162
Scheme 4.3. Attempted Installation of R ¹ and R ²	163
Scheme 4.4. Alternative Strategy Towards the Synthesis of 4.4 and Related Analogues	164

List of Tables

Table 1.1. Anti-HIV-1 and HBV Activity of Lysogenic Phospholipids 1.6a-e	18
Table 1.2. Anti-HIV-1 and HBV Activity of Bis-disulfide Conjugates 1.7a-e	19
Table 1.3. Anti-HIV-1 and HBV Activity of 2 nd Generation TFV Disulfide Prodrug Conjugates	24
Table 1.4. Anti-HIV-1 Activity of Conjugates 1.14-1.22 in PBMCs	28
Table 1.5. Anti-HIV-1 Activity of Non-Disulfide TFV Lipid Conjugates	29
Table 2.1. Antiviral Activity Various NAs and Prodrug Conjugates	91
Table 2.2. Crystal Data and Structure Refinement for 1-((2R,3R,4R,5R)-4-((<i>tert</i> -Butyldimethylsilyloxy)-3-fluoro-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (2.17)	109
Table 3.2. Prepared Acyclic Purine Thioaminal Nucleoside Analogues	115
Table 3.3. Prepared Acyclic Pyrimidine Thioaminal Nucleoside Analogues	116
Table 3.4. Antiviral Activity of Select Thioaminals against HIV-1, HBV, and HSV-1	118
Table 3.5. Crystal Data and Structure Refinement for 2-(((6-amino-9H-purin-9-yl)methyl)thio)ethanol	141
Table 3.6. Crystal Data and Structure Refinement for 2-(1-((((9H-purin-6-yl)amino)methyl)thio)methyl)cyclopropyl)acetic acid (3.15)	143
Table 3.7. Crystal Data and Structure Refinement for 5-Fluoro-4-(((1-hydroxypropan-2-yl)thio)methyl)amino)pyrimidin-2(1H)-one (3.22)	145
Table 3.8. Crystal Data and Structure Refinement for diethyl (2-(((2,4-dichloropyrimidin-5-yl)amino)methyl)thio)ethyl)phosphonate (3.28)	147
Table 4.1. Predicted binding affinity of various truncated MHC-I peptides to the surface of μ 1	157
Table 4.2. Crystal Data for (1-methyl-1 <i>H</i> -imidazol-2-yl)methanol (4.13)	182
Table 4.3. Crystal Data and Structure Refinement for 4-Bromo-2-((1-trityl-1 <i>H</i> -imidazol-2-yl)methoxy)benzotrile (4.16)	183
Table 4.4. Crystal Data and Structure Refinement for 4-Bromo-2-((1-methyl-1 <i>H</i> -imidazol-2-yl)methoxy)benzotrile (4.17)	184
Table 4.5. Crystal Data and Structure Refinement for 4-(1 <i>H</i> -imidazol-1-yl)-2-((1-trityl-1 <i>H</i> -imidazol-2-yl)methoxy)benzotrile	185

List of Abbreviations

A: adenosine

Ac: acetate

ADP: adenosine diphosphate

AFV: adefovir

AMP: adenosine monophosphate

AP-1: adapter protein 1

AZT: azidothymidine

BnOH: benzyl alcohol

Bp: base pair

BuOH: butanol

C: cytidine

CC₅₀: effective concentration required to reduce cell viability by 50%

CD: cytoplasmic domain

CMV: cytomegalovirus

CTL: cytotoxic T-lymphocyte

dCK: deoxycytosine kinase

DCM: dichloromethane

dCMP: deoxycytidine monophosphate

dGK: deoxyguanosine kinase

DMEM: Dulbecco's Modified Eagle Medium

DMF: dimethylformamide

DMSO: dimethylsulfoxide

DNA: deoxyribonucleic acid

dNK: deoxyribonucleoside kinase

DTE: dithioethanol

DTT: dithiothreitol

EBOV: ebola virus

EC₅₀: effective concentration required to inhibit viral replication by 50%

EC₉₀: effective concentration required to inhibit viral replication by 90%

ESI: electrospray ionization

EtOAc: ethyl acetate

EtOH: ethanol

FDA: Food and Drug Administration

FTC: emtricitabine

G: guanosine

GALT: gut-associated lymphatic tissue

GDP: guanosine diphosphate

GSH: glutathione

GTP: guanosine triphosphate

HAART: highly active antiretroviral therapy

hAMPK1: human adenylate kinase 1

hAMPK2: human adenylate kinase 2

HBV: hepatitis B virus

HCV: hepatitis C virus

HDE: hexadecyldisulfanylethanol

HDP: hexadecyloxypropyl

hGMPK: human guanylate kinase

HINT-1: histidine triad nucleotide binding protein 1

HIV-1: human immunodeficiency virus type-1

hNDK: human nucleoside diphosphate kinase

HRMS: high resolution mass-spectrometry

HSV: herpes simplex virus

HSV-1 TK: herpes simplex virus type-1 thymidine kinase

HSV-2 TK: herpes simplex virus type-2 thymidine kinase

hTMPK: human thymidylate kinase

HTVS: high-throughput virtual screening

hUMP-CMPK: human uridylate-cytidylate kinase

iPrOH: isopropanol
LC-MS: liquid chromatography mass-spectrometry
MeOH: methanol
MHC-I: major histocompatibility complex type-1
NA: nucleoside analogue
NDPK A: nucleoside diphosphate kinase A
NDPK: nucleoside diphosphate kinase
NMPK: nucleotide monophosphate kinase
NMR: nuclear magnetic resonance
NNRTI: non-nucleoside reverse transcriptase inhibitor
NRTI: nucleoside reverse transcriptase inhibitors
NuH: general nucleophile
PBMC: peripheral blood mononuclear cell
PBS: phosphate-buffered saline
PDB: protein databank
Ph: phenyl
PLC: phospholipase C
PLD: phospholipase D
PreP: pre-exposure prophylaxis
PrOH: propanol
PTSA: paratoluene sulfonic acid
qPCR: quantitative polymerase chain reaction
RNA: ribonucleic acid
ROH: general alcohol
RSH: general thiol
RT: reverse transcriptase
SAR: structure activity relationships
SATE: S-acyl thioethyl
s-BuOH: *sec*-butanol

T: thymidine

TAF: tenofovir alafenamide

TBSCl: *t*-butyldimethylsilyl chloride

t-BuOH: *tert*-butanol

TCA: trichloroacetic acid

TDF: tenofovir disoproxil fumarate

TFA: trifluoroacetic acid

TFV: tenofovir

TFVdpp: tenofovir diphosphate

THF: tetrahydrofuran

TI: therapeutic index

TK1: thymidine kinase 1

TK2: thymidine kinase 2

TLC: thin layer chromatography

TMS: tetramethylsilane

Ts: tosylate

U: uracil

UHV: ultra-high vacuum

UV: ultraviolet

Chapter 1

Novel Lipid Prodrugs of Tenofovir: Design, Synthesis, and Biological Evaluation

Abstract: The therapeutic value of numerous small molecules hinges on their ability to permeate the plasma membrane. This is particularly true for tenofovir (TFV), adefovir, and other antiviral nucleosides that demonstrate potent antiviral activity but poor bioavailability. Using TFV as a model substrate, we hybridized two disparate prodrug strategies to afford novel reduction-sensitive lipid conjugates of TFV that exhibit sub-nanomolar activity towards HIV-1 and are stable in human plasma for more than 24 h with a therapeutic index > 100,000. These compounds significantly rival the clinically approved formulation of TFV and revitalize the potential of disulfide-bearing prodrugs which have seen limited *in vitro* and *in vivo* success since their debut over 20 years ago. We provide evidence for enzymatic hydrolysis and contend that the disulfide linkage is not reduced prior to TFV release.

1.1 Introduction

Nucleosides are the building blocks of life. Without exception, the canonical nucleosides adenosine (A), guanosine (G), cytidine (C), thymidine (T), uracil (U), and their 2'-deoxy counterparts are the foundation of the genetic code and the RNA/DNA of every lifeform on the planet. The DNA nucleoside template features a 2'-deoxy ribose ring flanked with two hydroxyl moieties at the 5' and 3' position and a hydrophobic nucleobase at the 1'-position as a β anomer. RNA nucleosides also share these design features and bear an additional 2'-OH moiety that is thought to reduce the hydrolytic stability of RNA *in vivo*. Figure 1.1 shows the structures of the 2'-deoxy nucleosides and uridine. The incorporation of a nucleoside into a growing strand of DNA requires thrice phosphorylation of the 5'-OH by deoxyribonucleoside kinases (dNKs), nucleotide monophosphate kinases (NMPKs), and nucleoside diphosphate kinases (NDPKs) to yield a nucleoside triphosphate.¹ The nucleoside triphosphate is then recognized by a polymerase and is subsequently hydrolyzed to the monophosphate and covalently linked to the 3'-OH of a terminal nucleotide in the nascent strand. Complementary base pairing ensures that the newly synthesized strand complements the template sequence to yield two binary polymers of nucleic acids that wind together into a double helix.

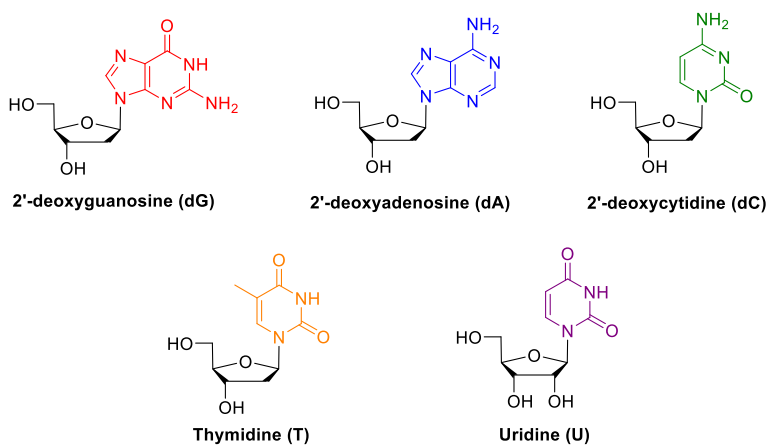


Figure 1.1. Canonical nucleosides ubiquitously found in Nature.

The fundamental role that DNA plays in the central dogma of modern biology requires DNA replication to be a precise and exacting process. A high fidelity polymerase is essential to maintaining the integrity of the human genome (3×10^9 bp) and is dependent on a healthy and well-balanced supply of the four deoxyribonucleotides. Base pair substitutions introduced during DNA synthesis often result in cellular aberrations that are detrimental to survival or come with the cost of reduced fitness.² Consequently, RNA and DNA polymerases are equipped with a proof-reading 3'-5'-exonuclease to eject and correct a given nucleotide during DNA or RNA synthesis so that it exactly complements the template sequence (A pairs with T, G pairs with C).³ This level of scrutiny is one of many mechanisms⁴ in place to preserve the integrity of the genome and is a common feature to most if not all DNA polymerases including the herpes simplex virus (HSV) DNA polymerase.⁵

Unlike most organisms, retroviruses have RNA genomes that require reverse transcription from RNA to DNA to replicate and produce viral progeny. Reverse transcription is performed by the viral enzyme reverse transcriptase (RT) which lacks a discretionary 3'-5'-exonuclease.⁶ The omission of this piece of machinery results in a relatively high error rate and subjects retroviruses to substantial genotypic variation.⁷ Human immunodeficiency virus (HIV) is a retrovirus and subsequently relies on its own viral RT for replication. Several studies have demonstrated that HIV approaches (but never reaches) error catastrophe, a term used to describe a phenomenon where a genome becomes completely compromised due to replication errors.⁸⁻¹⁰ The average mutation rate for HIV is approximately 5×10^{-5} mutations per nucleotide which is about one million times greater than the mutation rate observed in eukaryotic cells.¹¹⁻¹²

1.1.1 Nucleoside Analogues as Antiviral Drugs

Rapid and unfastidious replication can be exploited by nucleoside analogues (NAs) to target viruses that pose a threat to human health. Nucleoside analogues have achieved renown clinical success for the treatment of numerous viral infections including HIV,¹³ hepatitis B virus (HBV),¹⁴ hepatitis C virus (HCV),¹⁵ HSV,¹⁶ and cytomegalovirus (CMV).¹⁷⁻¹⁸ A handful of FDA-approved and clinically promising NAs are shown in Figure 1.2. These agents mimic the canonical nucleoside template in Figure 1.1 and are recognized by a viral polymerase to induce obligate or non-obligate chain termination which in turn compromises the production of viral progeny and arrests disease progression. Obligate chain terminator NAs are incorporated into viral DNA/RNA, but lack the critical 3'-OH necessary for DNA/RNA elongation. Non-obligate terminators are endowed with a 3'-OH, but adopt a conformation that prohibits the polymerase enzyme from functionalizing this position.

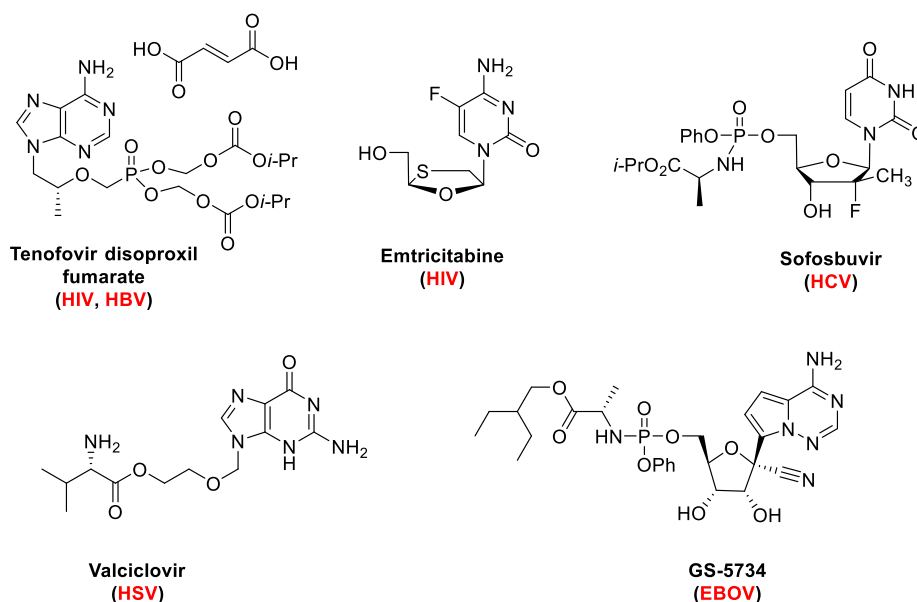


Figure 1.2. Nucleoside analogues used in the treatment of viral disease.

Emtricitabine and tenofovir are obligate chain terminators and are a key component of most once-a-day antiviral regimens for the treatment of chronic HIV. In 2012, the FDA approved the use of emtricitabine and tenofovir disoproxil fumarate as a chemical pre-exposure prophylaxis (PreP) to prevent the spread and transmission of HIV in men who have sex with men¹⁹ -a feat that was thought to be impossible fifteen years ago. Sofosbuvir received approval in 2013 for the treatment of chronic HCV and represents the first curative therapy for HCV with a cure rate exceeding 95% in most patients.²⁰⁻²¹ Valaciclovir has been on the market since 1995 and remains a prominent antiviral against genital HSV²² and herpes zoster virus.²³ GS-5734 was recently identified by Gilead Sciences in a high-throughput screening assay and demonstrates nanomolar potency against Ebola virus (EBOV) and has emerged as a promising clinical trial candidate to treat *Filoviridae* viral infections.²⁴⁻²⁵

1.1.2 Phosphorylation is Required for Metabolic Activation of NAs

All NAs require conversion to the active triphosphate in order to exert their antiviral activity on the viral polymerase (Figure 1.3). This is performed by several different kinases. For natural deoxynucleosides, the first phosphorylation step to the monophosphate is irreversibly catalyzed by deoxynucleoside kinases (dNKs); two cytosolic enzymes, thymidine kinase 1 (TK1) and deoxycytosine kinase (dCK), and two mitochondrial enzymes, deoxyguanosine kinase (dGK) and thymidine kinase

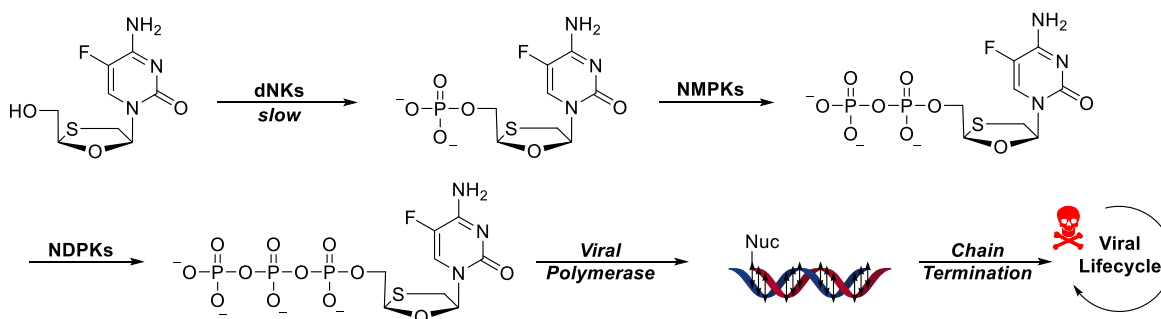


Figure 1.3. The phosphorylation of emtricitabine and other NAs requires three independent enzymes before reaching the viral polymerase to arrest viral replication.

2 (TK2).²⁶ Together, these four families collectively phosphorylate all natural deoxynucleosides to their corresponding nucleotides with distinct but overlapping specificity.

The catalytic efficiency of the dNK family is relatively poor compared to nucleoside monophosphate kinases (NMPKs) and nucleoside diphosphate kinases (NDPKs) that are responsible for the latter two phosphorylation events to yield the nucleoside triphosphate. Phosphorylation of a nucleoside to the monophosphate is therefore believed to be the rate-limiting step in the conversion of a NA to its active triphosphate.

In general, TK2 phosphorylates nucleosides bearing a thymidine or uracil base, dCK phosphorylates nucleosides with cytosine or adenine, and dGK phosphorylates nucleosides attached to guanine or hypoxanthine.²⁷ The specificity and efficiency of dNK enzymes towards various NAs has been examined. TK1 is able to phosphorylate NAs with structural variations at the 3'-position such as AZT and also tolerates modifications at the C5 position on the pyrimidine ring (5-fluoro, 5-chloro, 5-bromo, and 5-ethyl dU) as long as the structural perturbation doesn't introduce too much steric bulk.²⁸ NAs that lack a 3' substituent (such as d4T) are poor substrates for this enzyme and are not readily phosphorylated. With respect to the stereochemistry of the arabinose ring, TK1 is enantioselective and selectively phosphorylates the D-enantiomer of thymidine and its analogues.²⁹ In contrast, the mitochondrial enzyme dCK exhibits low enantioselectivity and readily phosphorylates L-dC and lamivudine with near equal catalytic efficiency to their D-enantiomers.

Several DNA viruses (herpes viruses and poxyviruses) synthesize their own viral thymidine kinases and thymidylate kinases that sensitize these viruses to NAs. Herpes simplex virus type-1 thymidine kinase (HSV-1 TK) exhibits a much broader range of activity than human TK1 and has relatively loose enantioselective substrate demands.³⁰ The enzyme also has the capacity to phosphorylate thymidine analogues with relatively large substituents at the C5 position, such as bromoethynyl in the case of brivudin.³¹ Acyclovir and ganciclovir are also selectively phosphorylated

to the monophosphate by HSV-1 TK and HSV-2 TK over human TK enzymes. Resistance to acyclovir and ganciclovir develops from mutations in the viral TK that diminish or silence its catalytic activity.³² Note that these viral kinases do not continue the activation of NAs to the diphosphate and subsequent triphosphate. This process is reserved for host kinases.

The second phosphorylation step toward a nucleoside triphosphate is catalyzed by the NMPK family to furnish the corresponding nucleoside diphosphate. In humans, the major NMPKs are human adenylate kinase 1 and 2 (hAMPK1 and hAMPK2), human guanylate kinase (hGMPK), human uridylate-cytidylate kinase (hUMP-CMPK), human thymidylate kinase (hTMPK). The enzymes are named according to their preferred substrate (hAMPK1 preferentially binds to adenosine monophosphate, AMP). The NMPK family generally exhibits higher catalytic activity in the conversion of their natural substrate to the corresponding diphosphate when compared to dNKs, however, many nucleoside monophosphates struggle to achieve diphosphate conversion at this step. AZT is rapidly phosphorylated to the monophosphate by TK1 with an efficacy comparable to that of thymidine,³³ however, AZT is a poor substrate for hTMPK which causes accumulation of the monophosphate in the cell.³⁴⁻³⁵ Other cytidine and uridine analogues have also been reported to accumulate at the monophosphate level which have the potential to disrupt cell signaling pathways and result in toxicity.³⁶⁻³⁷

The third and final reversible phosphorylation to an active nucleoside triphosphate is catalyzed by a variety of different enzymes at the basal level of metabolism that convert ADP to ATP or GDP to GTP. With respect to triphosphate formation, six different families of kinases have been identified, one of which is the human nucleoside diphosphate kinase (hNDK) family. hNDPKs were identified more than half a century ago and are abundant in most living organisms, have high catalytic activity, high turnover rates, and are relatively promiscuous.²⁶ Phosphorylated ribovarin and its derivatives are

processed by NDPK A at a rate comparable to cytosine-containing substrates.³⁸ In fact, most diphosphate NAs are efficiently converted to the triphosphate without significant kinetic barriers.

1.1.3 Kinase Bypass Strategies: Phosphonates

From the above discussion, it is clear that phosphorylation of a NA is required to unlock its activity, however, NAs that are poor substrates for host and viral kinases are not necessarily poor substrates for a viral polymerase. For instance, 2'-deoxy-5-(2-thienyl)uridine is nontoxic and inactive against HIV-1 in cell culture models, but the corresponding triphosphate is a potent inhibitor of HIV-1 RT and DNA polymerase α .³⁹ One way to completely bypass the need for kinase activation is to administer the active triphosphate directly to the patient. This is not currently feasible as triphosphates are extremely sensitive to chemical and enzymatic hydrolysis and typically exist in cells as transient species. For these reasons, no triphosphate NA has been developed for clinical use, but the idea has permeated the literature.⁴⁰

Nucleoside monophosphates (nucleotides) are markedly more stable toward chemical hydrolysis than nucleoside triphosphates and the installation of a phosphate generally improves the antiviral activity of the parent nucleoside. Unfortunately, the phosphate is readily clipped by 5'-nucleotides to yield the unphosphorylated nucleoside.⁴¹ This hydrolysis reaction can be overcome by replacing the frail P-O phosphate bond with a hydrolytically stable P-C bond to afford a phosphonate. This strategy hasn't been particularly yielding in the discovery of active NAs that feature a cyclic ribose or arabinose ring but has been successful for acyclic NAs like tenofovir, adefovir, and cidofovir (Figure 1.4).

Adefovir and tenofovir are slowly phosphorylated by hAMPK1 and hAMPK2 at rates significantly slower than AMP.⁴² Cidofovir is phosphorylated by hUMP-CMPK with low catalytic efficiency at a rate that is less than 0.1% that of dCMP phosphorylation.²⁶ Despite relatively poor phosphorylation to the diphosphate, the permanent nature of the P-C bond de-sensitizes acyclic

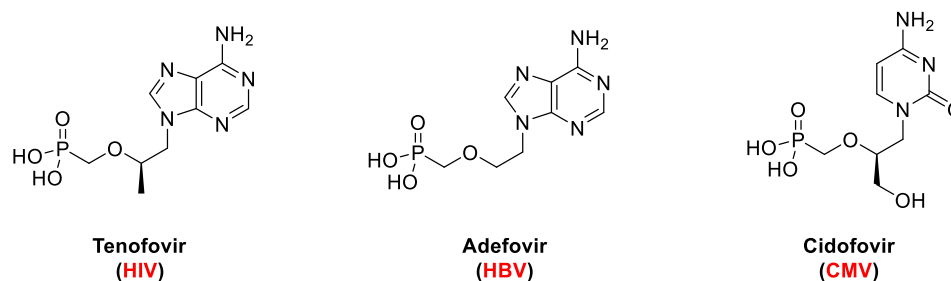


Figure 1.4. Acyclic phosphonate NAs to bypass the initial phosphorylation by dNKs.

phosphonate NAs to 5'-nucleotidases and the dianionic charge of the phosphonate moiety traps these compounds inside cells resulting in long intracellular half-life. These factors contribute to the relatively high accumulation of phosphorylated phosphonate acyclic NAs and their potent antiviral activities.

1.1.4 Tenofovir and the Use of Prodrugs

Tenofovir (TFV) demonstrates broad spectrum anti-viral activity against HIV,⁴³⁻⁴⁵ HBV,⁴⁶⁻⁴⁷ and HSV-2.⁴⁸ TFV structurally resembles 2',3'-dideoxyadenosine which lacks the requisite 3' hydroxyl moiety necessary for DNA polymerization and triggers obligate chain termination upon incorporation of tenofovir diphosphate (TFVdpp) into a growing viral DNA strand. The persistent nature of the phosphonate moiety successfully bypasses the initial phosphorylation event by dNKs, but plagues TFV and other acyclic phosphonate NAs with significant dianionic character at physiological pH that restricts diffusion across the plasma membrane resulting in rapid renal clearance, depreciated bioavailability, and reduced antiviral activity *in vivo*. When orally administered to mice,⁴⁹ the bioavailability of TFV is approximately 2% and that of adefovir has been reported to be <1% in monkeys⁵⁰ and 8-11% in rats.⁵¹ These undesirable properties can be ameliorated by masking the anionic phosphonic acid with various prodrugs to alter the pharmacokinetic profile of the parent nucleoside, enhance cellular permeability, and improve bioavailability. Several eclectic prodrug

strategies have been developed for this purpose and have been extensively reviewed.⁵²⁻⁵³ Figure 1.5 shows a handful of common prodrugs used to improve the lipophilicity of nucleotides and phosphonate NAs.

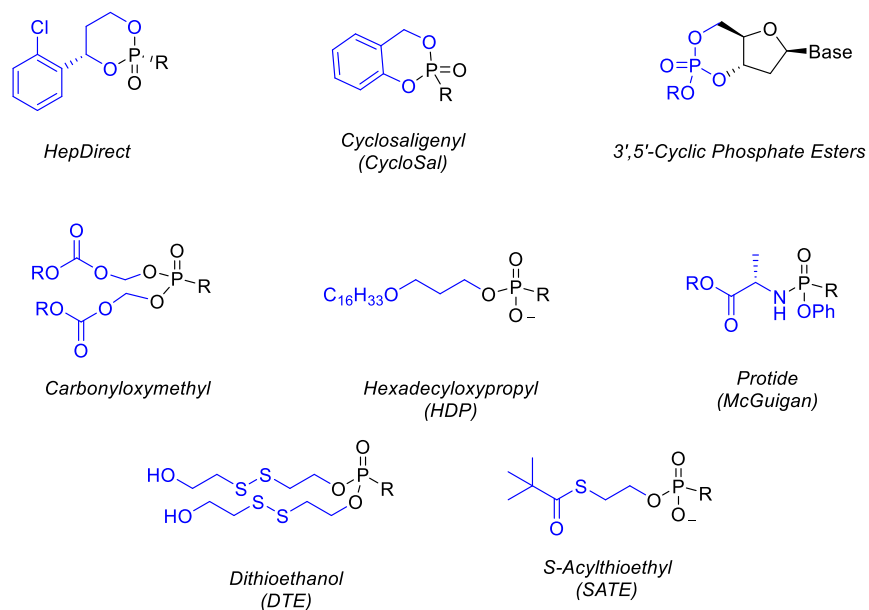
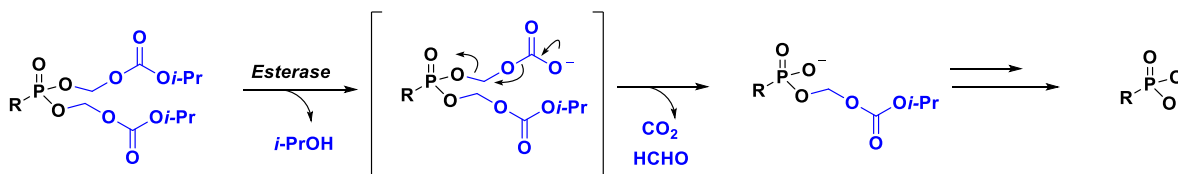


Figure 1.5. Prodrugs used to improve the bioavailability of NAs.

The clinically-approved prodrug formulation of TFV is tenofovir disoproxil fumarate (TDF) which is manufactured by Gilead Sciences. TDF is a carbonyloxymethyl prodrug that features two isopropoxycarbonyl groups esterified to the phosphonate. This prodrug relies on an esterase-activated cleavage mechanism to liberate TFV shown in Scheme 1.1.⁵⁴

Scheme 1.1. Decomposition of TDF *in vivo*.

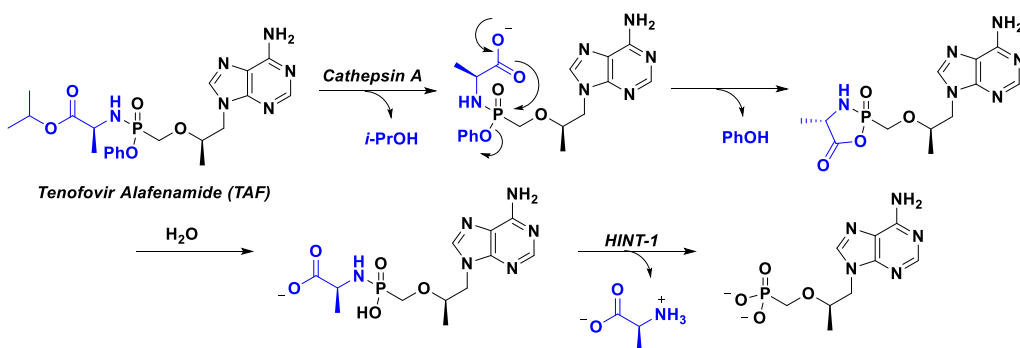


The installation of two isopropyl carbonate esters increases the oral bioavailability of TFV to 25%, enhances tissue distribution, and improves biological stability.⁴⁹ However, the relatively high concentration of esterases in the liver and plasma renders a significant fraction of TDF susceptible to premature hydrolysis resulting in systemic exposure to TFV, a known nephrotoxin.⁵⁵⁻⁵⁶ Continuous administration of TDF has been reported to induce lactic acidosis, Fanconi syndrome, acute renal failure, and bone which are a direct consequence of TFV's nephrotoxicity.⁵⁷⁻⁵⁸ It is therefore desirable to exploit alternative prodrug strategies to enhance intracellular delivery and reduce systemic exposure of tenofovir. To date, two lead candidates have surfaced: tenofovir alafenamide (TAF)⁵⁹⁻⁶⁰ and hexadecyloxypropyl tenofovir (CMX157).⁶¹

1.1.5 Tenofovir Alafenamide (TAF)

TAF is an isopropylalaninyl phenyl ester (protide) of TFV that requires two disparate enzymes for TFV release; cathepsin A and histidine triad nucleotide binding protein 1 (HINT-1).⁶² Scheme 1.2 shows the detailed mechanism of TAF degradation *in vivo*. Cathepsin A is a serine protease localized almost exclusively to lysosomal endosomes and initiates the hydrolysis of TAF via cleavage of the isopropyl ester. The resulting carboxylate anion cyclizes to eject phenoxide and subsequent hydrolysis of this intermediate furnishes an acyclic alinanylphosphoramidate that is removed from the

Scheme 1.2. Decomposition of Tenofovir Alafenamide *in vivo*.



phosphonate head group by HINT-1. TFV is then phosphorylated twice to the diphosphate to arrest viral replication.

In preliminary trials, TAF demonstrated little to no nephrotoxicity in thirty-eight subjects and was more potent than TDF against HIV at 1/10th the dose (25 mg/day) as a ten day monotherapy in HIV-1 positive adults.⁵⁹ Furthermore, a > 90% reduction in plasma TFV was observed in patients dosed with 25 mg/day of TAF compared to those given 300 mg TDF. These findings and subsequent clinical evaluation led to the FDA-approval of TAF in 2015 for the treatment of HIV. Gilead has since marketed new single tablet combination therapies for HIV-1 positive adults including Genvoya[®], Odefsy[®], and Descovy[®] to phase out TDF-containing regimens.

TAF has clear advantages over TDF, however, approximately 70% of the absorbed dose is extracted by the liver.⁶³ This means that only 7.5 mg of a single 25 mg dose reaches the target tissues as intact TAF and is sufficient to suppress HIV replication *in vivo*. Extensive hepatic extraction limits absorption into viral sanctuaries including the brain, gut-associated lymphatic tissues, and genitalia. Sub-optimal drug concentrations and low-levels of viral replication in these compartments can give rise to drug resistance over a prolonged period of time. Lipid prodrugs have the ability to penetrate viral sanctuaries to provide a sufficient concentration of antiviral drug to cells. One such lipid prodrug of TFV is CMX157.

1.1.6 CMX157

CMX157 is a hexadecyloxypropyl (HDP) prodrug of TFV that relies on the catalytic activity of an intracellular hydrolase, phospholipase C (PLC) and/or sphingomyelinase, to clip the entire HDP chain and liberate TFV exclusively within the cytosol in a single enzymatic step.^{61, 64} The HDP strategy evolved from a series of judicious modifications to the glycerophospholipid backbone to furnish a robust ethereal prodrug with improved metabolic stability (Figure 1.6). CMX157 is currently in Phase II clinical trials for the treatment of chronic HBV and has yielded promising results for the treatment

of HIV.⁶⁵ Preliminary data disclosed on Chimerix's website⁶⁶ reveal CMX157 is well-tolerated and achieves significant concentrations of TFVdpp up to one week after a single 400 mg dose, suggesting the potential for a convenient, once-a-week dosing regimen for HIV.

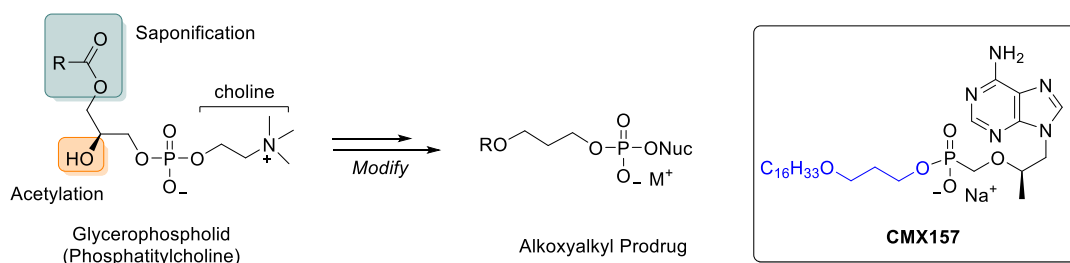


Figure 1.6. Evolution of alkoxyalkyl prodrugs and the structure of CMX157.

CMX157 is amphipathic and experiences a different metabolic fate than TAF and TDF that explains its long systemic half-life. Long chain amphipathic lysophospholipids are neither absorbed in the stomach nor transported to the liver by the portal vein. Rather, these substances are absorbed by intestinal enterocytes via passive diffusion to circumvent first-pass metabolism.⁶⁷ Once inside the intestinal cell, CMX157 is likely incorporated into milky fat particles called chylomicrons that contain triglycerides, fat-soluble vitamins, cholesterol, and other phospholipids. Chylomicrons exit the enterocyte by exocytosis where they go into lymphatic circulation and deliver lipids to distal tissues throughout the body.⁶⁸

1.2 The Design of Novel Lipid TFV Conjugates

CMX157's potential for reduced dosing frequency inspired us to design novel lipid conjugates of TFV. This lipid conjugate, shown in Figure 1.7, was designed to readily permeate the plasma membrane and rapidly release the nucleoside on contact with glutathione (GSH) or other intracellular reductases within the target cell *without* the need for enzymatic activation as shown in Scheme 1.3.

GSH is typically stockpiled in millimolar concentrations (2-10 mM) within the cytosol and participates in a multitude of biological functions to promote the reduction of disulfides, detoxification of xenobiotics⁶⁹⁻⁷⁰, mediate immunoregulation⁷¹, and curb oxidative stress⁷² amongst others. GSH is seldom found in the extracellular space and thus creates a gradient that fosters intracellular reduction of disulfides and simultaneously promotes an oxidative extracellular environment to support S-S formation, thus preserving the integrity of the disulfide until delivery.⁷³⁻⁷⁴

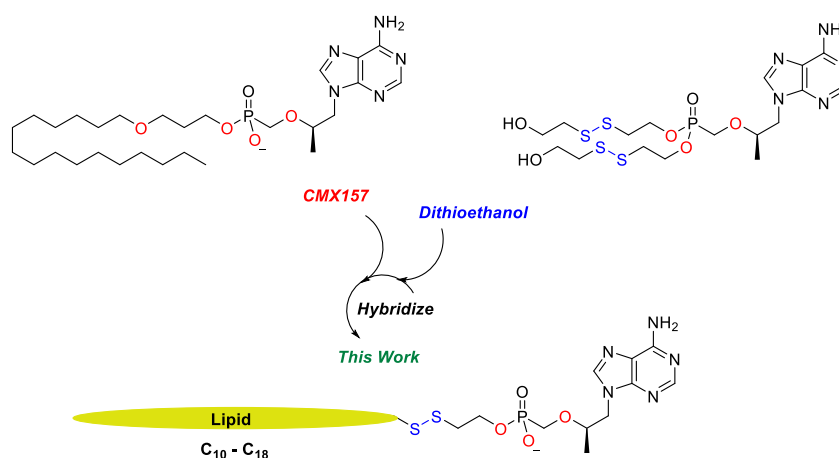
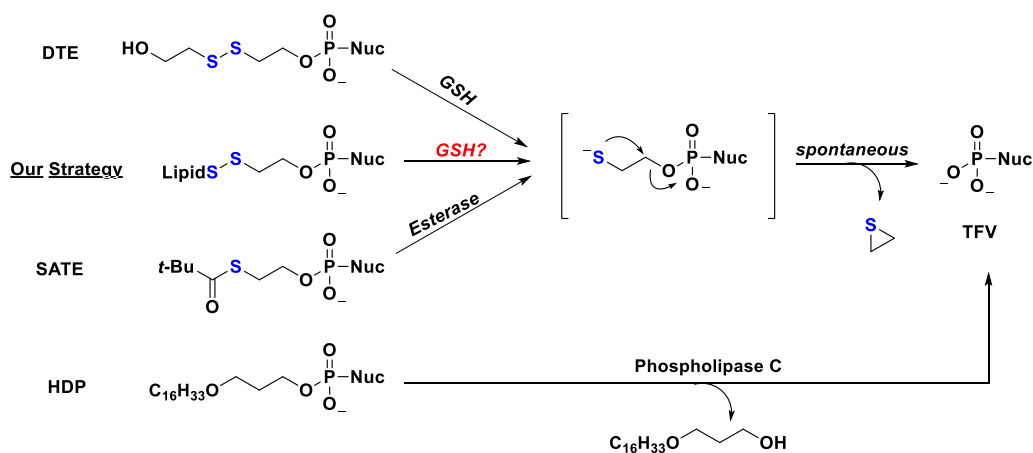


Figure 1.7. Concept design for our disulfide prodrug strategy.

Scheme 1.3. Comparison of Relevant Prodrugs and Their Proposed Mechanisms of Cleavage.



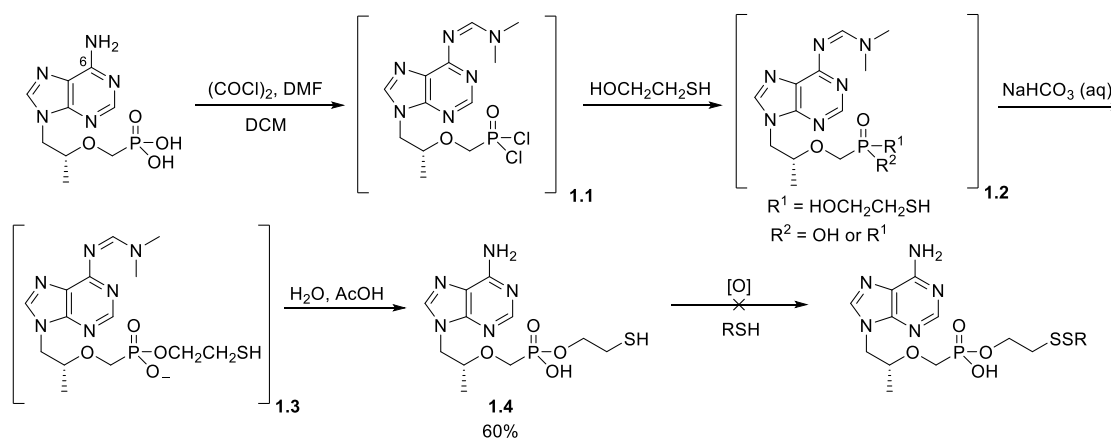
The genesis of our idea stemmed from the work of Gosselin and collaborators who developed

dithioethanol (DTE) prodrugs in the mid-1990s (Figure 1.2). The DTE prodrug strategy was exploited to mediate the delivery of adefovir, AZT, and 3'-deoxyuridine to HIV-1 infected MT-4, CEM-SS, and CEM-TK⁻ cell lines *in vitro*.⁷⁵⁻⁷⁶ Gosselin reported that conjugation of bis(DTE) to adefovir increased the HIV-1 activity of the parent nucleoside by ten-fold in CEM-TK⁻ cells and noted that the conjugates were exceptionally stable ($t_{1/2} > 24$ h) at pH 2, 7.4, culture medium, and human gastric juice. However, DTE prodrugs were found to rapidly degrade in human serum ($t_{1/2} < 5$ min) which has stalled clinical efforts to advance this technology forward. Herein, we hybridized the DTE approach with the hexadecyloxypropyl strategy parlayed for CMX157 to afford novel reduction-sensitive lipid prodrugs using tenofovir as a model substrate. The synthesis and biological activity of these compounds will now be discussed.

1.3 Synthesis of 1st Gen TFV Disulfide Conjugates

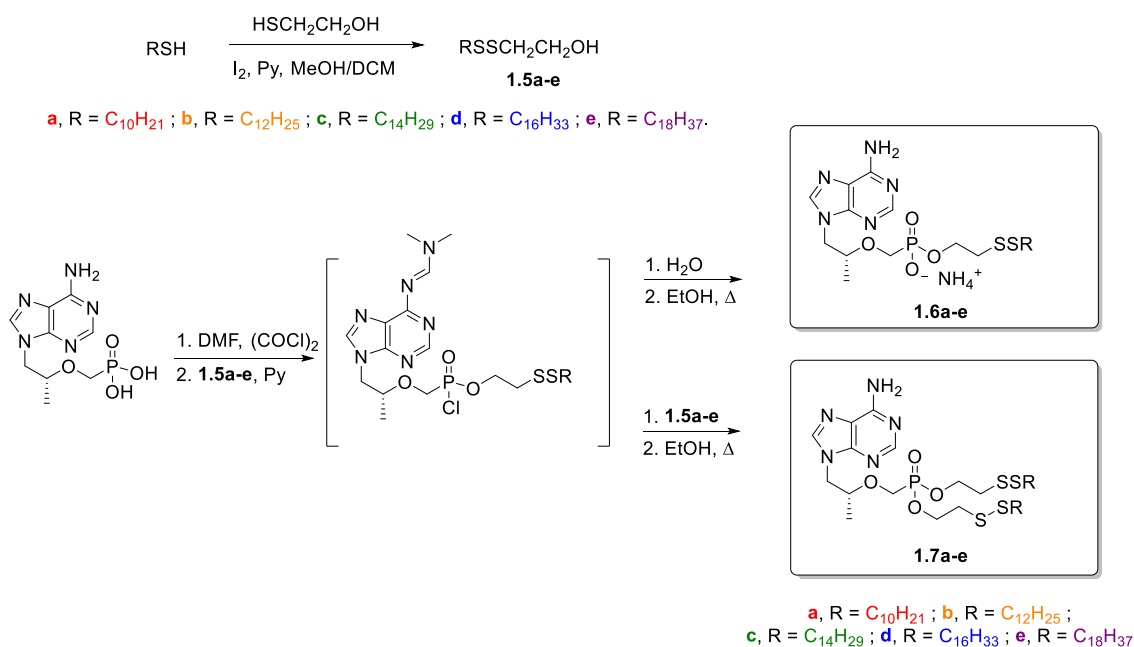
Our original synthesis toward the desired disulfide conjugates is shown in Scheme 1.4. The synthesis began with commercially available TFV and was converted to bischloridate **6** with excess oxalyl chloride and DMF in DCM. When catalytic DMF was used, the reaction failed to reach completion due to the presence of the C-6 amino moiety on the purine ring system that depletes the

Scheme 1.4. Initial Synthesis of TFV Disulfide Conjugates.



catalytic Vilsmeier chlorinating agent and becomes concomitantly protected as *N*-formimidine. This was resolved by adding 1.2 equivalents of DMF to rapidly afford the formimidine protected bischloridate **1.1**. Compound **1.1** was subsequently treated with excess β -mercaptoethanol to generate a mixture of bis- and mono- adducts of **1.2** that were not isolated. When **1.2** stirred with aqueous saturated sodium bicarbonate for 30 minutes, the complexity of the mixture reduced to a single species whose m/z ratio was consistent with **1.3**. Acid-mediated hydrolysis of **1.3** afforded **1.4** in moderate yield whose structural assignment was unambiguously characterized by HRMS, ^1H , ^{13}C , and ^{31}P NMR. The key step in Scheme 1.4 is the construction of a disulfide between **1.4** and an aliphatic thiol such as hexadecanethiol. Unfortunately, this endeavor proved unsuccessful with a variety of oxidizing agents including iodine, oxone, O_2 , and H_2O_2 . This reaction sequence was thus abandoned and an alternative synthesis of was sought. Our new synthesis (Scheme 1.5) features the oxidation of β -mercaptoethanol with a handful of aliphatic thiols (C_{10} - C_{18}) to produce the corresponding disulfide lipids **1.5a-e** in moderate yield (34-45%). These compounds were crystalline solids with sharp melting

Scheme 1.5. Successful Synthesis of 1st Generation TFV Disulfide Conjugates.



points (**1.5a** is an oil) that could be cleanly purified by silica gel column chromatography. Alcohols **1.5a-e** were subsequently coupled to TFV using the same DMF/oxalyl chloride methodology introduced in Scheme 1.3 to yield conjugates **1.6a-e** and **1.7a-e**.

The I₂-mediated oxidation of β -mercaptoethanol with aliphatic thiols consistently generated mixed disulfides in ~40% yield with the symmetrical disulfide byproducts producing the remaining 60%. This procedure generated a sufficient quantity of desired mixed disulfide for our purposes, however, there are established protocols for the selective production of unsymmetrical disulfides that may be useful on large scale. Two popular strategies from the synthetic literature exploit thioimide⁷⁷⁻⁷⁸ and thiocarbonate⁷⁹ methodologies which exploit reactive sulfenyl intermediates that react with a second thiol, RSH, to furnish unsymmetrical disulfides. In 2006, Hunter and collaborators described a convenient, one-pot procedure utilizing 1-chlorobenzotriazole (BtCl) to produce mixed disulfides in high yield.⁸⁰ During the discovery synthesis, this method was initially attempted to synthesize **1.5a-e**, however, the low reaction temperature (-20 °C) caused rapid precipitation of the long aliphatic thiols from the solution resulting in no appreciable reaction. This necessitated us to switch to a more conventional coupling strategy using the I₂/Py system shown in Scheme 1.5.

1.4 Biological Evaluation of 1st Gen TFV Disulfide Prodrug Conjugates

Compounds **1.6a-e** and **1.7a-e** were evaluated against HIV-infected PBMCs and antiviral activity was assessed by measuring reverse transcriptase in cell supernatants. HBV activity was also measured and qPCR was used to quantify viral DNA following incubation in HepG2.2.15 cells after 6 days. Table 1.1 details the antiviral activity of conjugates **1.6a-e** with modifications to the lipid chain length. These compounds resemble lysogenic phospholipids characterized by an anionic phosphate head group and a single aliphatic tail and presumably have detergent-like properties that disrupt lipid bilayers.⁸¹⁻⁸²

Table 1.1. Anti-HIV-1 and HBV Activity of Lysogenic Phospholipids **1.6a-e** Compared to TDF, TFV, TAF, and CMX157.^a

Compound	HIV-1			HBV		
	EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	TI (CC ₅₀ /EC ₅₀)	EC ₅₀ ^b (HepG2)	CC ₅₀ ^c (HepG2)	TI (CC ₅₀ /EC ₅₀)
TAF	0.0007	21	30,000	-	-	-
TDF	0.0046	44	9,500	0.34	65	190
TFV	0.319	> 100	> 300	-	-	-
CMX157	0.02	> 100	> 5,000	-	-	-
1.6a	0.085	> 50	> 590	> 50	> 50	> 1
1.6b	0.0030	25	8,300	1.67	> 50	> 30
1.6c	0.00050	14	28,000	0.444	> 25	> 56
1.6d	0.00065	14	22,000	0.020	> 25	>1,200
1.6e	0.00060	6.4	11,000	0.505	> 50	> 100

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1 or HBV replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce cell viability by 50%.

As shown in Table 1.1, compound **1.6a** possesses the shortest lipid (C₁₀) and is 18-fold less potent than TDF (85 nM vs. 4.6 nM) with a relatively poor therapeutic index (TI) of 590 at the EC₅₀. Extending the length of the alkyl chain by two carbon atoms results in **1.6b** whose HIV-1 activity (3.0 nM) is comparable to TDF and seven-fold more active than CMX157 (20 nM). Maximum antiviral activity was obtained for conjugates **1.6c-e** with alkyl chain lengths ranging from 14-18 carbon atoms, respectively. **1.6c-e** exhibit sub-nanomolar EC₅₀ values of ~0.5 nM that outrivals TDF and compare well with TAF. Notably, **1.6d** boasts a TI₅₀ that exceeds 20,000 whereas that of **1.6c** approaches 30,000. This broad therapeutic window is sustained even at the EC₉₀ for both **1.6c** and **1.6d** and is nearly an order of magnitude wider than the TI₉₀ of TDF (11,000 vs. 1,300 –data not shown). The observed correlation between chain length and antiviral activity corresponds to Hostetler’s work who found that ~20 atoms is the ideal linker length necessary to achieve maximum potency for alkyloxypropyl

conjugates of cidofovir.⁶⁴ Note that compounds **1.6c-e** possess linker lengths of 19, 20, and 21 atoms, respectively, when the mercaptoethyl bridge is taken into account. Table 1.1 also reveals that increasing chain length is associated with a concomitant increase in cytotoxicity –a phenomenon that has been documented for a variety of surfactants in numerous aquatic organisms.⁸³

With respect to HBV activity, all conjugates in Table 1.1 demonstrated rather unremarkable activity with the exception of **1.6d** whose potency (20 nM) and TI₅₀ (1,250) bested that of TDF which further supports the advantage of our lipids over carbonate prodrugs.

Table 1.2. Anti-HIV-1 and HBV Activity of Bis-disulfide Conjugates **1.7a-e**^a

Compound	HIV-1			HBV		
	EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	TI (CC ₅₀ /EC ₅₀)	EC ₅₀ ^b (HepG2)	CC ₅₀ ^c (HepG2)	TI (CC ₅₀ /EC ₅₀)
1.7a	0.349	> 100	> 287	> 100	> 100	> 1
1.7b	1.18	> 100	> 85	> 100	> 100	> 1
1.7c	0.189	> 100	> 529	> 100	> 100	> 1
1.7d	6.32	> 25	4.0	> 25	> 25	> 1
1.7e	0.331	> 25	> 76	> 25	> 25	> 1

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1 or HBV replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce cell viability by 50%.

In contrast to **1.6a-e**, **1.7a-e** resemble conventional phospholipids with two aliphatic tails that assume a cylindrical shape and do not readily traverse the plasma membrane on their own accord.⁸⁴ Previous works have demonstrated that phospholipid translocation depends on several factors such as lipid chain length, composition of the polar head group, and the presence of various ATP-dependent enzymes (flippase, floppase, scramblase, aminophospholipid translocase, etc) to affect this locomotion.⁸⁴⁻⁸⁶ We therefore speculated that **1.7a-e** would exhibit compromised antiviral activity when compared to their lysogenic brethren. Indeed, all phosphonodiester in Table 1.2 proved to be

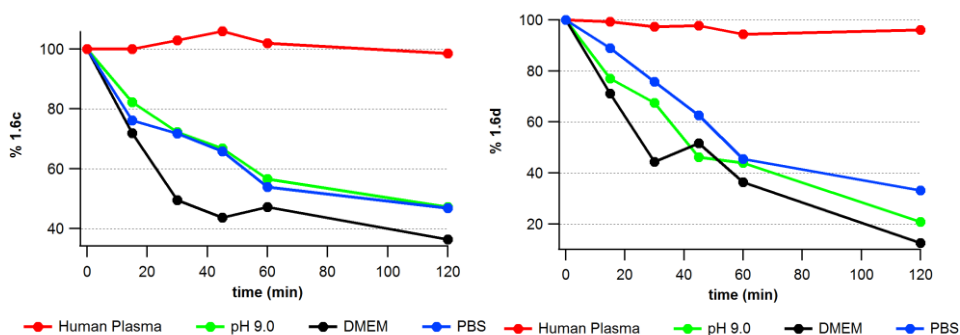


Figure 1.8. Decomposition profile of **1.6c** and **1.6d** at 37 °C in human plasma, sodium carbonate/bicarbonate buffer (pH 9.0), Dulbecco's modified eagle medium (DMEM) without fetal bovine serum, and phosphate-buffered saline (PBS, pH 7.4).

significantly less potent than **1.6a-e**. Conjugates **1.7a** and **1.7e** achieved potencies comparable to TFV against HIV-1 whereas **1.7c** was the only compound to demonstrate a two-fold increase in activity when compared to the parent nucleoside. All compounds in Table 1.2 were relatively inactive against HBV.

1.5 Stability Studies for our 1st Generation Conjugates

In order to be considered clinical candidates for oral delivery, conjugates **1.6a-e** must demonstrate sufficient hydrolytic, nucleophilic, and plasma stability. Despite the dismal reported plasma stability profile of bis(DTE)-conjugates ($t_{1/2} < 5$ min)⁷⁵, we reasoned **1.6a-e** could form micelles in solution and protect the delicate S-S linkage from reductases circulating in human serum.⁸⁷ Compounds **1.6c** and **1.6d** were selected as model compounds and subjected to various media including human serum, PBS buffer (pH 7.4), Dulbecco's Modified Eagle Medium (DMEM), and carbonate/bicarbonate buffer (pH 9). Aliquots of each sample were analyzed by LC-MS at varying times points over the course of 2 h with a final time point at 24 h (Figure 1.8).

Compounds **1.6c** and **1.6d** demonstrated robust stability in human plasma with a half-life > 24 h which provides some evidence to support our micelle theory. We then wondered how **1.6c** and **1.6d** would fare in the presence of base (pH 9.0), PBS (pH 7.4), and upon exposure to nucleophilic

media (DMEM). This task proved to be more demanding than expected. Figure 1.8 reveals that **1.6c** and **1.6d** extensively “decompose” at pH 9, PBS, and DMEM after 3 h, yet apparently remain intact when incubated with human serum. We suspected that the lipid component of **1.6c** and **1.6d** gradually adsorb to the glass surface of the vessel in these highly aqueous environments to cause an artificial decline in analyte concentration. To test this hypothesis, we assessed the stability of hexadecyloxypropyl 2'-deoxy-2'-fluorouridine, an alkoxyalkyl lipid conjugate of 2'-deoxy-2'-fluorouridine in PBS buffer using Pyrex, Kimax, and silanized glassware. These experiments revealed that both Pyrex and silanized glass encourage substantial lipid adsorption and gave rise to an artificial decomposition profile similar to that in Figure 1.8 (data not shown). Interestingly, Kimax glassware did not produce this effect despite its near identical composition to Pyrex. All stability experiments were repeated with Kimax glassware using hexadecyloxypropyl 2'-deoxy-2'-fluorouridine as an internal standard. With these conditions, both **1.6c** and **1.6d** exhibited similar stabilities at pH 9, PBS, and DMEM, however, neither compound exhibited a half-life of > 2 h in these media. We believe that precipitation is responsible for the poor half-life of these compounds in PBS, carbonate buffer, and DMEM. This is consistent with the observation that **1.6c** (C₁₄ alkyl chain) procured a near two-fold stability increase over **1.6d** (C₁₆ alkyl chain) in all examined media (with the exception of human serum). Taken together, the stability data shown in Figure 1.8 for compounds **1.6c** and **1.6d** in PBS, DMEM, and at pH 9 is of little value, however, the long plasma stability profile of **1.6c** and **1.6d** reveals a clear advantage of our lipid prodrugs and suggests that these compounds resist premature degradation in serum.

1.6 Probing the Mechanism of Cleavage

To this point we have assumed that our disulfide conjugates cleave in a reductive manner illustrated in Scheme 1.3 to reveal the antiviral activity of TFV. This mechanism invokes the liberation of mutagenic thiirane which has also been implicated in the decomposition of S-acyl-2-thioethyl

Scheme 1.6. Kinetic Model Used to Study Intramolecular Cyclization for 1st Gen Prodrugs.

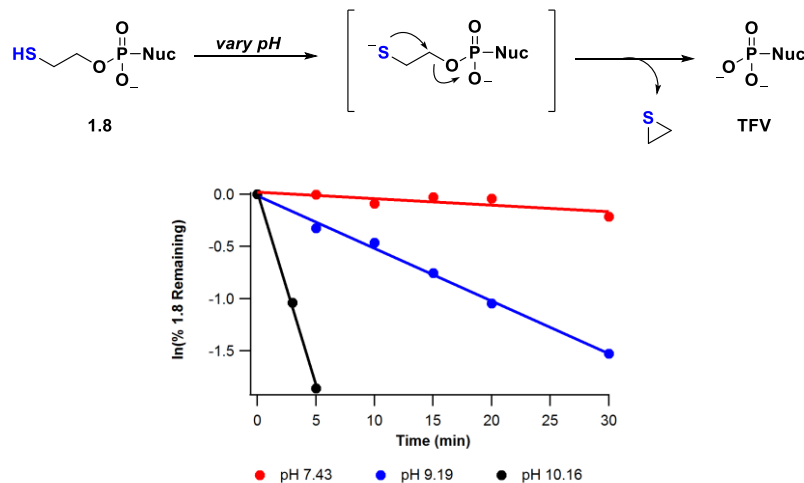


Figure 1.9. The decomposition profile of compound **1.8** at various pH values obeys first order kinetics with respect to **[1.8]** in aqueous media.

(SATE) and dithioethanol (DTE) prodrugs. SATE prodrugs are popular for the *in vitro* delivery of nucleoside analogues, however, they are rarely considered for clinical development due to toxicity concerns. This prompted us to probe the cleavage mechanism of our disulfide conjugates using **1.6d** as a model compound. To validate the mechanism shown in Scheme 1.3, we initially treated **1.6d** with GSH in PBS at 37.4 ± 0.2 °C and monitored decomposition by LC-MS. However, as previously mentioned, the poor aqueous solubility of **1.6d** precluded accurate spectral measurements and necessitated an alternative model system. We therefore chose to access the posited thiolate intermediate via the deprotonation of **1.8** rather than subjecting **1.6d** to GSH with the underlying assumption that intramolecular cyclization governs TFV release (Scheme 1.6). Thiol **1.8** was synthesized and incubated in various aqueous solutions buffered at pH 2.0, 9.19, and 11.56 spiked with 0.1 M DTT at 37.4 °C. Compound **1.8** demonstrated high stability in glycine/HCl buffer (pH = 2.0) with no observed decomposition. Moderate stability was noted in PBS (pH = 7.43) with the appearance of a new UV signal that faintly presented itself at a significantly polar retention time ($t_r =$

0.66 min). This species was identified as TFV when compared to an authentic sample. When **1.8** was incubated in aqueous carbonate/bicarbonate buffer (pH = 9.19) significant decomposition occurred and complete consumption of starting material was noted within 8 minutes at pH 10.16. Using the method of initial rates, the differential rate law was found to be first order with respect to $[\text{OH}^-]$ resulting in the overall second-order rate expression for the decomposition of **1.8**.

$$\frac{d[\mathbf{6}]}{dt} = k[\mathbf{1.8}][\text{OH}^-] \quad \text{Equation 1}$$

The $[\text{OH}^-]$ dependence in Equation 1 suggests that the thiolate performs the *3-exo-tet* cyclization to release TFV rather than thiol **1.8** which is consistent with the mechanism presented in Schemes 1.3 and 1.5. Figure 1.9 graphically illustrates how the pH influences the reaction rate. These results are also in agreement with previous work⁸⁸⁻⁸⁹ and confirm that thiolates of SATE, DTE, and **1.6d** have the capacity to cyclize and release the parent nucleoside and generate thiirane. When assayed against HIV-infected PBMCs, **1.8** is 58-fold less active than TFV ($\text{EC}_{50} = 18.7 \mu\text{M}$) indicating that the mercaptoethyl moiety is not readily cleaved by enzymatic means and reinforces the notion that a single proton governs TFV release in this system.

1.7 Second Generation Disulfide Conjugates

The kinetic experiments outlined in Scheme 1.5 and Figure 1.9 suggest that the toxicity associated with **1.6d** and its congeners may be due to thiirane production. This led us to prepare second generation disulfide conjugates that would be better suited for clinical development (Table 1.3).⁹⁰ Compound **1.9** was of primary interest and was designed to undergo intramolecular *5-exo-tet* cyclization to yield non-electrophilic tetrahydrothiophene and TFV after disulfide reduction (Scheme 1.6). To our delight, **1.9** exhibited sub-nanomolar anti-HIV activity comparable to **1.6d** with reduced toxicity and a $\text{TI}_{50} > 100,000$. This therapeutic index is likely wider as the poor aqueous solubility of **1.9** prevented accurate CC_{50} measurements beyond $50 \mu\text{M}$.

Table 1.3. Anti-HIV-1 and HBV Activity of 2nd Generation TFV Disulfide Prodrug Conjugates^a

Cmpd	Structure	HIV-1			HBV		
		EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	TI (CC ₅₀ /EC ₅₀)	EC ₅₀ ^b (HepG2)	CC ₅₀ ^c (HepG2)	TI (CC ₅₀ /EC ₅₀)
TFV		0.320	> 100	> 300	-	-	-
TDF		0.0045	44.0	9500	0.34	64.5	190
1.6d		0.00065	14.3	22000	0.020	>25	> 1200
1.8		18.6	> 100	> 5.38	41.9	>100	> 2
1.9		< 0.0005	> 50	> 100000	0.248	> 50	> 200
1.10		5.13	> 100	> 18	> 100	> 100	> 1
1.11		0.0229	17.2	751	4.48	17.5	3.9
1.12		< 0.0005	15.9	> 31800	0.152	32.5	214
1.13		0.007	> 50	> 7000	1.05	> 50	> 47

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1 or HBV replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce cell viability by 50%. R = TFV

To determine if **1.9** has the ability to release TFV via a 5-*exo-tet* intramolecular cyclization reaction, thiol **1.10** was synthesized and subjected to various aqueous buffers in a similar manner done for **1.8** (Scheme 1.7). Surprisingly, **1.10** was stable at all tested pH values, even at pH 11.56 with no detectable decomposition after several hours at 37 °C (Figure 1.10). This finding gave us pause; if **1.10** and its corresponding thiolate do not cyclize to liberate TFV, then reduction is not expected to facilitate TFV release from **1.9**. Furthermore, **1.10** is weakly active against HIV and HBV, revealing

Scheme 1.7. Kinetic Model Used to Study the Potential Intramolecular Cyclization of **1.9**.

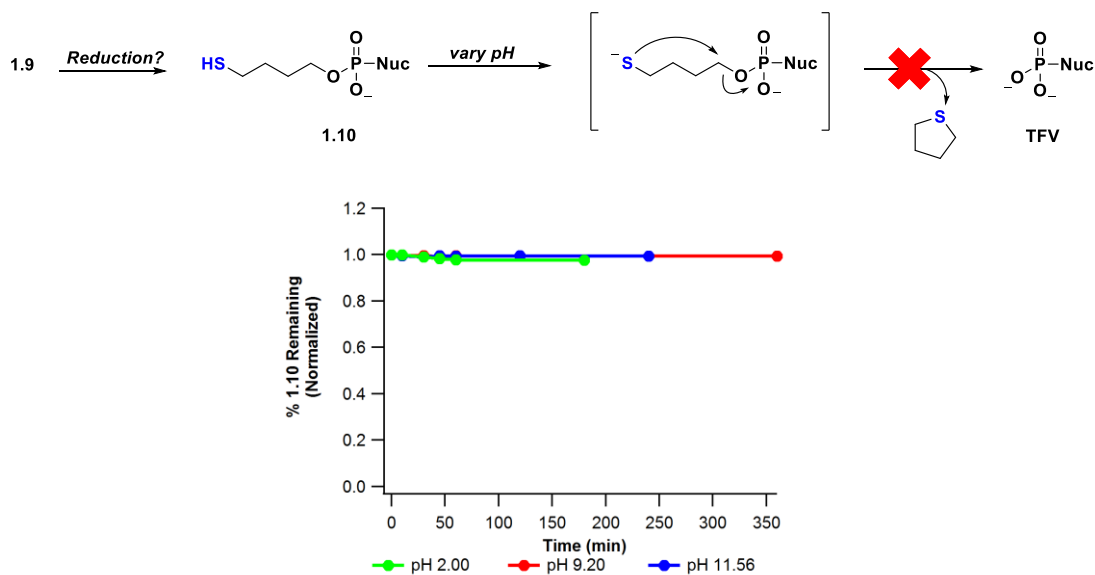


Figure 1.10. Decomposition profile of **1.10** in various buffer solutions containing 0.1 M DTT at 37.4 °C.

that this compound is a relatively poor substrate for enzymatic hydrolysis. When taken together, the evidence suggests that disulfide reduction is not necessary for TFV release. Additional support for this theory comes from the synthesis and subsequent antiviral evaluation of compound **1.13**, which features the disulfide eight atoms away from the phosphonate head group to prohibit the potential for kinetic and thermodynamic cyclization following reduction. Compound **1.13** potently inhibited HIV-1 replication ($EC_{50} = 7.0$ nM) indicating facile removal of the disulfide lipid *in vitro*. This leads us to believe that **1.9** and **1.13** are substrates for cellular enzymes, specifically phospholipase C and/or sphingomyelinase,^{64,91} to sever the P–O bond before disulfide reduction (if reduction occurs at all) as shown in Figure 1.11. It is likely that all of the conjugates in Tables 1.1, 1.2, and 1.3 are subject to this mechanism of cleavage despite the fact that our first generation conjugates have the ability to release TFV via intramolecular cyclization. It remains to be seen if thiol **1.8** is realized *in vitro*, however, it's

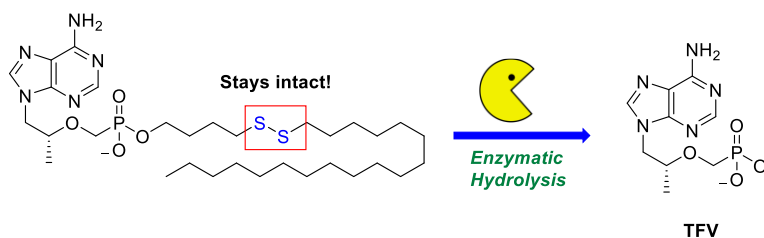


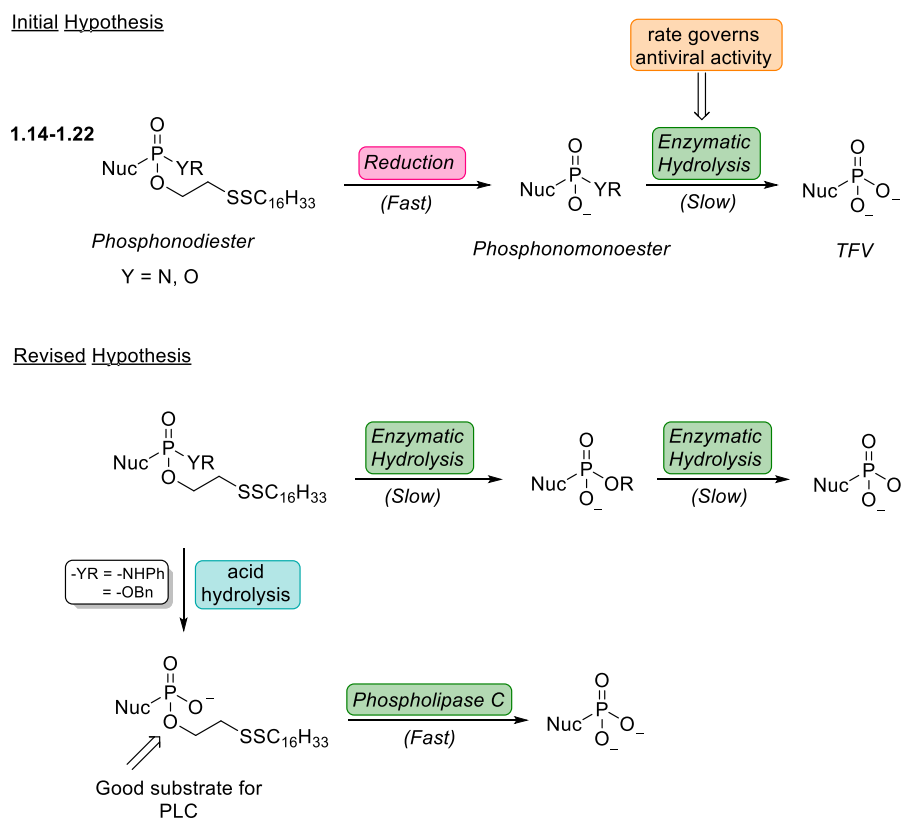
Figure 1.11. Modified cleavage mechanism for **1.9** and related analogues.

possible that micelle formation hinders GSH and other hydrophilic reductants from approaching the S—S bond in our 1st gen series so that enzymatic hydrolysis may be the exclusive mechanism of release for these compounds.

With the cleavage mechanism partially elucidated, we now turn to evaluate some of the structure activity relationships (SAR) in Table 1.3. Various structural perturbations of the aliphatic chain (i.e translocation of the disulfide, the introduction of aromatic rings) tend to result in relatively potent TFV conjugates. One exception is **1.11** whose anti-HIV activity is compromised when compared to the other conjugates in Table 1.3. It's interesting to note that **1.9** is substantially less toxic than **1.6d** for different reasons than previously thought. After prodrug cleavage, one can envision that the resulting disulfide lipid eventually undergoes reduction at a later time to release mercaptoethanol and 4-mercaptobutanol, respectively, to produce different toxicity profiles. The 4-mercaptobutanol spacer exploited for **1.9** appears to achieve maximum antiviral activity and minimum cytotoxicity and was therefore chosen as the optimum thiol going forward.

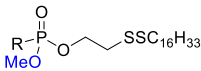
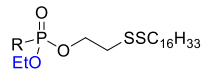
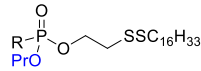
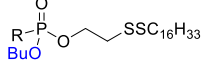
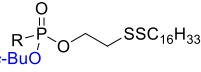
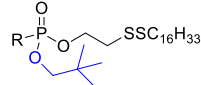
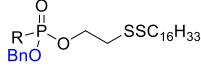
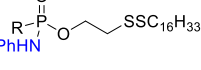
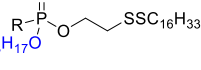
1.8 Revisiting Our 1st Generation Conjugates

Prior to our kinetic studies, we prepared a third series of phosphonodiester 1st generation conjugates (**1.14-1.22**) featuring hexadecyldisulfanylethanol (HDE) and another smaller alcohol esterified to the phosphonate head group. Our initial intent for these compounds was to utilize the self-immolative properties of the HDE lipid to indirectly probe the enzymatic hydrolysis of the

Scheme 1.8. Using Disulfide Prodrugs to Probe Enzymatic Hydrolysis.


resulting phosphonomonoester shown in Scheme 1.8. However, we cannot rule out the possibility that both esters are enzymatically hydrolyzed in a two-step fashion. Table 1.4 shows the anti-HIV of compounds **1.14-1.22**. There is no apparent trend in increasing or decreasing anti-HIV activity with respect to the identity of the alcohol shown in blue. Notably, all of the compounds in Table 1.4 are less potent than our lysogenic disulfide prodrugs shown in Table 1.1. Compounds **1.14**, **1.20**, and **1.21** are the most potent analogues with **1.21** nearly equipotent to TDF. The potency of **1.20** and **1.21** may stem from the acid-labile nature of the benzyl and phosphoramidate moieties that can come off at

Table 1.4. Anti-HIV-1 Activity of Conjugates **1.14-1.22** in PBMCs^a

Cmpd	Structure	EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	TI (CC ₅₀ /EC ₅₀)
1.14		0.026	38	> 1,500
1.15		0.35	36	102
1.16		0.216	10.6	49.0
1.17		0.19	27.9	> 150
1.18		0.049	23.7	480
1.19		0.14	> 100	> 700
1.20		0.018	> 41.1	> 2,300
1.21		0.005	34.2	7,000
1.22		0.911	> 50	> 50

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1 replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce the viability of uninfected cells by 50%.

R = tenofovir

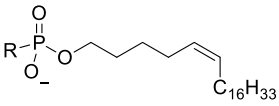
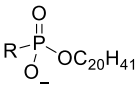
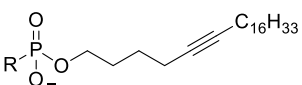
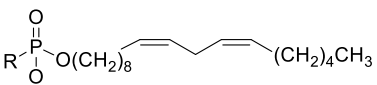
mildly acidic pH values (~5) present in lysosomes. Facile chemical removal of these groups generates a lysogenic phospholipid prodrug akin to **1.6a-e** that is a good substrate for PLC to rapidly release the disulfide lipid and liberate TFV.

1.9 What is the Disulfide Advantage?

Out of all the disulfide conjugates prepared, it is significant to note that **1.6d**, **1.9**, and **1.13** are more potent than CMX157 suggesting an inherent advantage of our disulfide prodrug over HDP conjugates. These differences are not subtle; **1.9** is 40-fold more potent than CMX157 against HIV-1. This raises the question: what is the disulfide advantage? Is it the “kinked” S–S bond (i.e., its 90°

dihedral angle)? If so, how does its presence improve antiviral activity of TFV over HDP lipids? One possible explanation is that the disulfide increases membrane fluidity to accelerate translocation of the conjugate from the outer leaflet of the plasma membrane to the inner leaflet where it is then enzymatically cleaved to release TFV within the cytosol. Alternatively, the added “kink” may enhance the rate of PLC cleavage. To gain insight, a fourth series of TFV-lipid conjugates **1.23-1.26** were prepared that feature various geometries, double bonds, and triple bonds in lieu of a disulfide to see if our technology is truly unique. These compounds were prepared in a slightly different manner from our disulfide conjugates using 2,4,6-trisopropylbenzenesulfonylchloride as the coupling reagent (see Experimental). The anti-HIV activity of these compounds is shown in Table 1.5.

Table 1.5. Anti-HIV-1 Activity of Non-Disulfide TFV Lipid Conjugates^a

Cmpd	Structure	EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	TI (CC ₅₀ /EC ₅₀)
1.23		0.0003*	> 18.0	60,000
1.24		< 0.0003*	> 12.0	> 40,000
1.25		0.0004*	> 25.0	63,000
1.26		0.038	> 50.0	13,000

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1 replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce the viability of uninfected cells by 50%.

*Compound did not fully solubilize in DMSO stock solution for biological analysis. R = TFV

Compounds **1.23-1.25** were relatively nontoxic toward PBMCs and exhibited potent anti-HIV activity with sub-nanomolar EC₅₀ values. The activity profile of these compounds is similar to those observed for our disulfide prodrugs and suggests that the disulfide bond does not provide a potency

gain over other simple aliphatic prodrugs like **1.24**. It's apparent that the enzymes responsible for the cleavage of our TFV-lipid prodrug conjugates have a broad substrate scope and that our disulfide technology is not unique in this regard. Interestingly, linoleic acid conjugate **1.26** was 95-fold less potent than **1.23-1.25** indicating that multiple units of unsaturation may negatively influence PLC hydrolysis.

Despite the fact that our disulfide prodrugs fail to confer a unique potency advantage, CMX157 and other lipid prodrugs experience significant ω -oxidative catabolism in the intestinal lumen that shortens the lipid chain and reduces their lipophilicity and affinity for PLC (personal communication). ω -oxidation occurs at the termini of the aliphatic chain and serves as a rescue pathway for long chain (>12 C) lipids that cannot undergo α or β cleavage. With this new knowledge, we wondered if the sharp dihedral angle of our disulfide lipids would deter enzymatic oxidative catabolism which is predominantly mediated by the CYP450 superfamily. These studies are ongoing in collaboration with the Emory Institute for Drug Development (EIDD). Unfortunately, no data are available at the time of this writing.

In addition to examining the metabolic stability of our disulfide prodrugs, we are also interested in synthesizing non-disulfide lipid prodrugs to deter ω -oxidation in the intestinal lumen. Terminal trifluoromethylation or alkynylation of the ω carbon on CMX157 are our immediate targets. It is hoped that the installation of these moieties preserves the structural integrity of the lipid chain until PLC cleavage releases TFV. Then, oxidative decomposition can occur at the hydroxy terminus of the released lipid.

1.10 Conclusions

Our results demonstrate that disulfide-bearing lipid conjugates of tenofovir significantly improve the potency of the parent nucleoside by over 600-fold and that these conjugates readily permeate the cell membrane, are stable in human plasma for > 24 h, and have significantly higher

therapeutic indices than TDF against both HIV and HBV at the EC₅₀ and EC₉₀. This strategy has the potential for selective intracellular delivery and enhanced permeability of tenofovir to viral sanctuaries such as the gut-associated lymphatic tissue, central nervous system, and reproductive organs. Although a disulfide is not required to confer subnanomolar potency, it may offer unique advantages to deter oxidative lipid catabolism which is known to occur for CMX157 in the intestinal lumen. We have ongoing studies that seek to evaluate the stability of our disulfide conjugates in intestinal microsomes to better characterize their metabolic fate and additional analogues are being designed to suppress undesired ω -oxidation.

1.11 Experimental Details

1.11.1 Anti-HIV Assay

All HIV assays were performed by ImQuest BioSciences (Frederick, MD). All compounds were solubilized at 40 mM in DMSO and stored at -20 °C. Test materials were evaluated up to 100 μ M, and five serial logarithmic dilutions. AZT and 3TC were obtained from the NIH AIDS Research and Reference Reagent Program and used as controls in the anti-HIV and anti-HBV assay, respectively. Fresh human PBMCs were obtained from a commercial source and determined to be HIV and HBV negative. The leukophoresed blood cells were washed repeatedly with PBS, then diluted 1:1 with Dulbecco's phosphate buffered saline (PBS) and layered over 15 ml of Ficoll-Hypaque density gradient in a 50 ml conical centrifuge tube. The tubes were centrifuged for 30 min at 600 g. Banded PBMCs were gently aspirated from the resulting interface and washed three times with PBS. After the final wash, cell number was determined by Trypan Blue dye exclusion and cells re-suspended at 1×10^6 cells/mL in RPMI 1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 μ g/mL PHA-P, 100 U/ml penicillin and 100 μ g/mL streptomycin and allowed to incubate for 48-72 hr at 37°C. After incubation, PBMCs were centrifuged and re-suspended in tissue culture medium (RPMI

1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 $\mu\text{g}/\text{mL}$ PHA-P, 100 U/ml penicillin and 100 $\mu\text{g}/\text{mL}$ streptomycin, 3.6 ng/mL recombinant human IL-2). The cultures were maintained until use by half-volume culture changes with fresh IL-2 containing tissue culture medium every 3 days. Assays were initiated with PBMCs at 72 hr post PHA-P stimulation. Immediately prior to use, target cells were re-suspended in fresh tissue culture medium at 1×10^6 cells/ml and plated in the interior wells of a 96-well round bottom microliter plate at 50 $\mu\text{L}/\text{well}$. Then, 100 μL of 2X concentrations of compound-containing medium was transferred to the 96-well plate containing the cells in 50 μL of the medium. Following addition of test compound to the wells, 50 μL of a predetermined dilution of HIV virus (prepared at 4x of final desired in-well concentration) was added, and mixed well. For infection, 50-150 TCID_{50} of each virus was added per well (final MOI approximately 0.002). PBMCs were exposed in triplicate to virus and cultured in the presence or absence of the test compound at varying concentrations as described for the 96-well microliter plates. After 7 days, HIV-1 replication was quantified in the tissue culture supernatant by measurement of reverse transcriptase activity. Wells with cells and virus only served as virus controls. Separate plates were identically prepared without virus for cytotoxicity studies. Reverse transcriptase activity was measured in cell-free supernatants using a standard radioactive incorporation polymerization assay.

1.11.2 Anti-HBV Assay

One hundred microliters (100 μL) of wells of a 96-well flat-bottom plate at a density of 1×10^4 cells per well were incubated at 37 °C in 5% CO_2 for 24 h. Following incubation, six ten-fold serial dilutions of test compound prepared in RPMI1640 medium with 10% fetal bovine serum were added to individual wells of the plate in triplicate. Six wells in the plate received medium alone as a virus control only. The plate was incubated for 6 days at 37 °C at 5% CO_2 . The culture medium was changed on day 3 with medium containing the indicated concentration of each compound. One hundred microliters (100 μL)

of supernatant was collected from each well for analysis of viral DNA by qPCR and cytotoxicity was evaluated by XTT staining of the cell culture monolayer on the sixth day (see below). Ten microliters (10 μ L) of cell culture supernatant collected on the sixth day was diluted in qPCR dilution buffer (40 μ g/ml sheared salmon sperm DNA) and boiled for 15 minutes. Quantitative real time PCR was performed in 386 well plates using an Applied Biosystems 7900HT Sequence Detection System and the supporting SDS 2.4 software. Five microliters (5 μ L) of boiled DNA for each sample and serial 10-fold dilutions of a quantitative DNA standard were subjected to real time Q-PCR using Platinum Quantitative PCR SuperMix-UDG (Invitrogen) and specific DNA oligonucleotide primers (IDT, Coralville, ID) HBV-AD38-qF1 (5'-CCG TCT GTG CCT TCT CAT CTG-3'), HBV-AD38-qR1 (5'-AGT CCA AGA GTY CTC TTA TRY AAG ACC TT-3') and HBV-AD38-qP1 (5'-FAM-CCG TGT GCA /ZEN/CTT GCG TTC ACC TCT GC-3'BHQ1) at a final concentration of 0.2 μ M for each primer in a total reaction volume of 15 μ L. The final HBV DNA copy number in each sample was interpolated from the standard curve by the SDS2.4 software and the data were analyzed by Excel.

1.11.3 Cytotoxicity Studies

Cytotoxicity was evaluated by staining uninfected cells with tetrazolium dye XTT with spectrophotometric readings at 450/650 nm with a Molecular Devices Vmax plate reader. These experiments were run in triplicate for each compound tested.

1.11.4 Stability Studies

Stock solutions of the target compound (2.0 mg) were prepared in methanol at a concentration 100-fold that of the desired concentration. 10 μ L of this solution was added to either 990 μ L of human plasma, PBS, DMEM, or sodium carbonate buffer at pH = 9 in Kimax glass tubes that were pre-warmed in a reciprocal water bath shaker at 37 $^{\circ}$ C prior to addition of compound. The mixture was then vortexed and compound (**12c**, **12d**, or DTE-TFV) was added to the solution. Vortexing

continued at 150 rpm at 37 °C. At each time point (0, 15, 30, 45, 120, 1440 min) a 100 μ L aliquot was taken from the solution and mixed with 400 μ L of an internal standard (0.5 μ L hexadecyloxypropyl 2'-deoxy-2'-fluorodine in methanol with 0.1% formic acid) in 1.5 mL polypropylene micro-centrifuge tubes. The aliquots were vortexed for 10 seconds at maximum speed, centrifuged at 13,000 rpm, and decanted into LC vials. The supernatant was analyzed by LC-MS by noting the relative UV intensity of the title compound at 254 nm.

1.11.5 Kinetic Studies

All buffer solutions (pH = 2, 7.4, 9.19, 10.16) were freshly prepared and the pH of each solution was measured with an Accumet Basic AB15 Plus pH meter at 23 °C. Quantitative LC-MS (QLC-MS) was performed with an Agilent Technologies 6100 quadrupole instrument equipped with UV detection at 254 and 210 nm and a Varian C8 analytical column using a H₂O/MeOH gradient (35-75% MeOH) over three minutes with intermittent washing (35% MeOH isocratic, 1 min) between runs. For analysis, a four dram vial equipped with a magnetic stir bar was charged with the desired buffer (3.0 ml) and incubated in a water bath at 37.4 ± 0.2 °C prior to the introduction of analyte (compound **6** or **7**). Dithiothreitol (DTT) was also added so that the concentration of DTT was ~ 0.1 M. After stirring for 10 minutes, 60 μ L of a stock solution of **6** or **7** (dissolved in water and MeOH) was added to the buffer solution. Aliquots of this solution (25 μ L) were removed at time points $t = 0$ min, 5 min, 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, 12 h, and 24 h. Shorter time points were required for some analyses. Note that the samples were not removed from the water bath during aliquot acquisition. The aliquots were diluted with an aqueous solution containing H₂O/MeOH/HCOOH (75:25:0.1, respectively) to a fixed volume of 1025 μ L and analyzed by LC-MS. Integration of the 254 nm absorption signal was the preferred method of quantitation. Linear and exponential regressions were performed with IgorPro v.6.36 and Microsoft Excel, where convenient.

1.11.6 Chemical Synthesis and General Procedures

All reagents were obtained from commercial suppliers and used without further purification unless otherwise specified. Reaction progress was monitored by either thin layer chromatography (TLC) using pre-coated aluminum-backed silica gel plates (60 F₂₅₄ Merk, article 5554), by NMR, or liquid chromatography-mass spectrometry (LC-MS) on an Agilent Technologies 6100 quadrupole instrument equipped with UV detection at 254 and 210 nm and a Varian C8 analytical column. LC-MS analyses were performed using a step-wise H₂O/MeOH gradient with the % MeOH increasing from 75-95% over the course of 3 min unless otherwise specified. Flash column chromatography was conducted using CombiFlash Rf 200 (Telendyne-Isco) automated flash chromatography system with hand-packed RediSep columns. Evaporation of solvents was carried out on a rotary evaporator under reduced pressure and under ultra-high vacuum (UHV) where appropriate. ¹H NMR and ¹³C NMR spectra were recorded at ambient temperature on a Varian 400 spectrometer. ³¹P spectra were recorded at ambient temperature on either a Mercury 300 or Varian 400 spectrometer. Unless otherwise specified, all NMR spectra were obtained in deuterated chloroform (CDCl₃) and referenced to the residual solvent peak. Chemical shifts are given in δ values and coupling constants are reported in hertz (Hz). Melting points were determined on a MelTemp melting apparatus and are uncorrected. High resolution mass-spectra (HRMS) were acquired on a VG 70-S Nier Johnson or JEOL mass spectrometer. Elemental analyses were performed by Atlantic Microlabs (Norcross, GA) for C, H, N analysis and are in agreement with the proposed structures with purity \geq 95%.

General Procedures

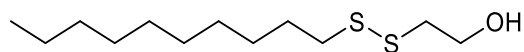
Procedure A: To a stirring solution of thiol A (28.4 mmol) and thiol B (28.4 mmol) in MeOH/DCM (25:75, 200. mL) was added pyridine (56.8 mmol) followed by the gradual added iodine (7.21 g, 28.4 mmol). The solution stirred for 4 h at room temperature with progress monitored by TLC. Then, the mixture was then diluted with MeOH (100 mL) and the suspension was filtered and the supernatant

collected. The solvents were evaporated under reduced pressure and the resulting solid was washed with dilute acid and extracted into DCM. The organic layer was concentrated and the residue purified on a silica column using a hexanes/EtOAc gradient to afford the titled disulfide.

Procedure B: To a stirring solution of tenofovir (0.174 mmol) in anhydrous DCM (6.00 mL) and *N,N*-dimethylformamide (0.209 mmol) was gradually added excess oxalyl chloride (0.870 mmol). The mixture stirred open to atmosphere for 15 min at room temperature until a colorless, transparent solution was observed and no starting material coated the walls of the vessel. The solvent and excess oxalyl chloride were evaporated under reduced pressure to produce a pale yellow foam which was redissolved in anhydrous DCM (5 mL) and placed under argon. The vessel was equipped with a magnetic stir bar and chilled to 0 °C. Then, a solution of X (0.209 mmol) and pyridine (1.05 mmol) in anhydrous DCM was slowly added, drop-wise. After stirring for 10 min at this temperature, the mixture was naturally warmed to room temperature and stirred for an hour. Then, Y was added (if X = Y, 2 equiv. of X are added in one portion) and stirring continued for 30 minutes. Aqueous 1.2 M HCl or methanolic ammonia were used to cleave the undesired formimidine and progress could be monitored by LC-MS (isocratic 95% MeOH, 5% H₂O, 7 min). Upon completion, the solvents were evaporated and the resulting residue was purified via silica column chromatography using a DCM/DCM:MeOH:NH₄OH (90:10:0.3) gradient (0-26%) to furnish the titled tenofovir conjugate.

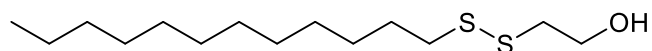
Procedure C: Tenofovir (0.348 mmol) and lipid L (0.522 mmol) in anhydrous pyridine (3 mL) were mixed with 2,4,6-triisopropylbenzene-1-sulfonyl chloride (1.05 mmol) at room temperature and stirred for 24 h. Then, aqueous dilute sodium hydroxide was added and stirring continued for 2 h. The mixture was concentrated under reduced pressure and partitioned between DCM and dilute aqueous acid. The organic layer was collected and purified on silica using a DCM/DCM:MeOH:NH₄OH (80:20:3) gradient to furnish the title compound.

1.11.7 Synthesis and Characterization of Lipids



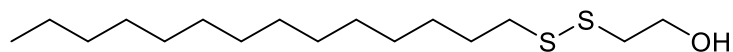
2-(Decyldisulfanyl)ethanol (1.5a)

Procedure A, thiol A = decane-1-thiol, thiol B = 2-mercaptoethanol. Purification gradient: hexanes/EtOAc (0-12% EtOAc), 43.2 % yield. Colorless oil. ^1H NMR (400 MHz, CDCl_3) δ 3.90 (dd, $J = 11.9, 5.9$ Hz, 2H), 2.85 (dd, $J = 7.0, 4.6$ Hz, 2H), 2.75 - 2.66 (m, 2H), 2.11 (t, $J = 6.3$ Hz, 1H), 1.75 - 1.62 (m, 2H), 1.58 - 1.19 (m, 14H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 60.3, 41.1, 39.0, 31.9, 29.5, 29.47, 29.3, 29.2, 29.1, 28.5, 22.3, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{12}\text{H}_{26}\text{OS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 273.1317, found 273.1312.



2-(Dodecyldisulfanyl)ethanol (1.5b)

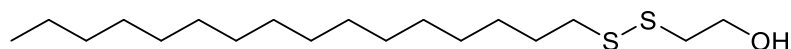
Procedure A, thiol A = dodecane-1-thiol, thiol B = 2-mercaptoethanol. Purification gradient: hexanes/EtOAc (0-15% EtOAc), 44.6 % yield. Waxy solid. ^1H NMR (400 MHz, CDCl_3) δ 3.90 (dd, $J = 11.8, 6.0$ Hz, 2H), 2.85 (t, $J = 5.8$ Hz, 2H), 2.75 - 2.67 (m, 2H), 2.04 (t, $J = 6.3$ Hz, 1H), 1.77 - 1.58 (m, 2H), 1.47 - 1.16 (m, 18H), 0.89 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 60.2, 41.1, 39.02, 31.9, 29.62, 29.61, 29.56, 29.47, 29.3, 29.2, 29.1, 28.5, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{14}\text{H}_{30}\text{OS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 301.1630, found 301.1625. Melting Point: 30-32°C



2-(Tetradecyldisulfanyl)ethanol (1.5c)

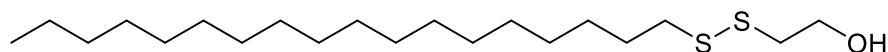
Procedure A, thiol A = tetradecane-1-thiol, thiol B = 2-mercaptoethanol. Purification gradient: hexanes/EtOAc (0-12% EtOAc), 44.2 % yield. White powder. ^1H NMR (400 MHz, CDCl_3) δ 3.90 (dd, $J = 11.9, 6.0$ Hz, 2H), 2.85 (t, $J = 5.8$ Hz, 2H), 2.75 - 2.66 (m, 2H), 2.10 (t, $J = 6.3$ Hz, 1H), 1.74

- 1.62 (m, 2H), 1.47 - 1.06 (m, 22H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 60.3, 41.1, 39.0, 31.9, 29.67, 29.65, 29.63, 29.62, 29.57, 29.48, 29.3, 29.2, 29.1, 28.5, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{16}\text{H}_{34}\text{OS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 329.1943, found 329.1936. Melting Point: 41-42 °C.



2-(Hexadecyldisulfanyl)ethanol (1.5d)

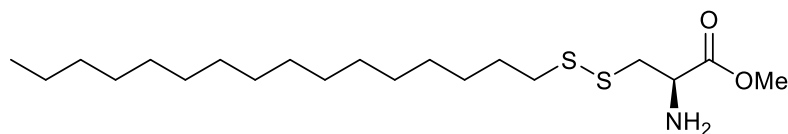
To a stirring solution of hexadecane-1-thiol (8.74 mL, 28.4 mmol) and 2-mercaptoethanol (2.00 mL, 28.4 mmol) in MeOH/DCM (50:50, 200 mL) was added pyridine (4.94 mL, 56.8 mmol) followed by the gradual added iodine (7.21 g, 28.4 mmol) until the color of the solution remained brown. The solution stirred for 2 h at room temperature then the resulting suspension was filtered and the supernatant collected. The solvents were evaporated under reduced pressure and the resulting solid was washed with water and extracted into DCM. The organic layer was concentrated and the residue purified on a silica column using hexanes/EtOAc (0-8%) gradient to afford the title compound 2-(hexadecyldisulfanyl)ethanol (3.98 g, 11.89 mmol, 41.9 % yield) as a fluffy white powder. ^1H NMR (400 MHz, CDCl_3) δ 3.90 (q, $J = 5.6$ Hz, 2H), 2.95 – 2.80 (m, 2H), 2.77 – 2.63 (m, 2H), 2.01 (t, $J = 6.0$ Hz, 1H), 1.69 (dt, $J = 14.9, 7.3$ Hz, 2H), 1.46 – 1.17 (m, 26H), 0.89 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 60.3, 41.2, 39.1, 31.9, 29.7 (4), 29.66, 29.63, 29.58, 29.49, 29.36, 29.2, 29.1, 28.5, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{18}\text{H}_{38}\text{OS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 357.2256, found 357.2253. Melting Point: 50-51°C.



2-(Octadecyldisulfanyl)ethanol (1.5e)

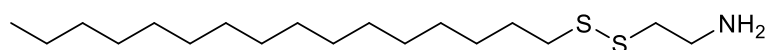
Procedure A, thiol A = octadecane-1-thiol, thiol B = 2-mercaptoethanol. Purification gradient: hexanes/EtOAc (0-15% EtOAc), 34.3 % yield. White powder. ^1H NMR (400 MHz, CDCl_3) δ 3.90 (dd, $J = 11.8, 5.9$ Hz, 2H), 2.85 (t, $J = 5.8$ Hz, 2H), 2.75 - 2.68 (m, 2H), 2.03 (t, $J = 6.2$ Hz, 1H), 1.74

- 1.63 (m, 2H), 1.45 - 1.17 (m, 30H), 0.89 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 60.3, 41.1, 39.0, 31.9, 29.68 (3), 29.66 (2), 29.64 (2), 29.63, 29.57, 29.48, 29.35, 29.2, 29.1, 28.5, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{20}\text{H}_{42}\text{OS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 385.2569, found 385.2563. Melting Point: 57-58 °C.



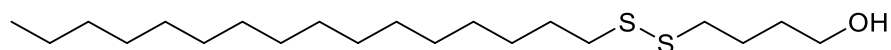
(R)-Methyl 2-amino-3-(hexadecyldisulfanyl)propanoate (SI-1)

Procedure A, thiol A = hexadecane-1-thiol, thiol B = and L-cysteine methyl ester hydrochloride. Isolated upon filtration as a third layer that presented itself during extraction and required no further purification, 32.0 % yield. White solid. All spectra are referenced to CD_3OD . ^1H NMR (400 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$) δ 4.66 (s, 2H), 4.32 (dd, $J = 7.4, 4.6$ Hz, 1H), 3.85 (s, 3H), 3.25 (dd, $J = 14.8, 4.6$ Hz, 1H), 3.15 (dd, $J = 14.8, 7.4$ Hz, 1H), 2.81 - 2.65 (m, 2H), 1.75 - 1.60 (m, 2H), 1.46 - 1.14 (m, 26H), 0.85 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$) δ 168.2, 53.1, 51.3, 37.9, 36.8, 31.7, 29.5 (4), 29.47 (2), 29.42, 29.3, 29.2, 29.1, 28.8, 28.3, 22.5, 13.6. HRMS (ESI) m/z calculated for $\text{C}_{20}\text{H}_{42}\text{O}_2\text{N}_2$ $[\text{M} + \text{H}]^+$: 392.2651, found 392.2651. Melting Point: 114 - 116 °C.



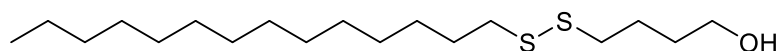
2-(Hexadecyldisulfanyl)ethanamine (SI-2)

Procedure A, thiol A = hexadecane-1-thiol, thiol B = cystamine hydrochloride. Isolated upon filtration as a third layer that presented itself during extraction and required no further purification, 47.0 % yield. White powder. ^1H NMR (400 MHz, $\text{MeOD}/\text{CDCl}_3$) δ 2.94 (t, $J = 6.7$ Hz, 2H), 2.58 (t, $J = 6.7$ Hz, 2H), 2.41 - 2.34 (m, 2H), 1.34 (dt, $J = 14.9, 7.3$ Hz, 2H), 1.10 - 1.00 (m, 3H), 1.00 - 0.87 (m, 25H), 0.53 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{MeOD}/\text{CDCl}_3$) δ 37.7, 37.6, 33.4, 31.4, 29.1 (3), 29.10 (3), 29.04, 28.97, 28.82, 28.69, 28.50, 27.9, 22.1, 13.2. HRMS (ESI) m/z calculated for $\text{C}_{18}\text{H}_{40}\text{NS}_2$ $[\text{M} + \text{H}]^+$: 334.2596, found 334.2596. Melting Point: 99 - 102 °C.



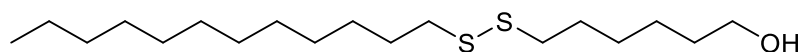
4-(Hexadecyldisulfanyl)butan-1-ol (SI-3)

Procedure A, thiol A = hexadecane-1-thiol, thiol B = 4-mercaptobutan-1-ol. Purification gradient: hexanes/EtOAc (0-16% EtOAc), 38.6 % yield. Pearlescent white powder. ^1H NMR (400 MHz, CDCl_3) δ 3.68 (t, $J = 6.3$ Hz, 2H), 2.78 - 2.65 (m, 4H), 1.85 - 1.73 (m, 2H), 1.73 - 1.62 (m, 4H), 1.43 - 1.22 (m, 26H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 62.4, 39.1, 38.7, 31.9, 31.4, 29.67 (2), 29.66 (2), 29.64 (2), 29.58, 29.50, 29.4, 29.23, 29.20, 28.51, 25.4, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{20}\text{H}_{42}\text{OS}_2$ $[\text{M} + \text{Na}]^+$: 385.2569, found 385.2562. Melting Point: 51-52 °C.



4-(Tetradecyldisulfanyl)butan-1-ol (SI-4)

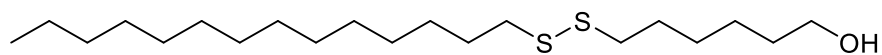
Procedure A, thiol A = tetradecane-1-thiol, thiol B = 4-mercaptobutan-1-ol. Purification gradient: hexanes/EtOAc (0-15% EtOAc), 38.5 % yield. White solid. ^1H NMR (400 MHz, CDCl_3) δ 3.68 (t, $J = 6.3$ Hz, 2H), 2.78 - 2.62 (m, 4H), 1.84 - 1.73 (m, 2H), 1.73 - 1.61 (m, 4H), 1.42 - 1.21 (m, 22H), 0.88 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 62.4, 39.1, 38.7, 31.9, 31.4, 29.67, 29.66, 29.64 (2), 29.58, 29.50, 29.4, 29.23, 29.20, 28.5, 25.4, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{18}\text{H}_{38}\text{OS}_2$ $[\text{M} + \text{H}]^+$: 335.2436, found 335.2436. Melting Point: 41-42°C.



6-(Dodecyldisulfanyl)hexan-1-ol (SI-5)

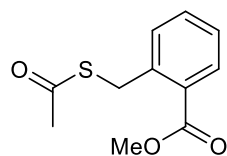
Procedure A, thiol A = dodecane-1-thiol, thiol B = 6-mercaptohexan-1-ol. Purification gradient: hexanes/EtOAc (0-10% EtOAc), 46 % yield. White solid. ^1H NMR (400 MHz, CDCl_3) δ 3.65 (t, $J = 6.6$ Hz, 2H), 2.71 - 2.66 (m, 4H), 1.75 - 1.63 (m, 4H), 1.63 - 1.54 (m, 2H), 1.48 - 1.35 (m, 6H), 1.34 - 1.21 (m, 16H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 62.9, 39.1, 38.9, 32.6, 31.9,

29.6, 29.6, 29.58, 29.50, 29.33, 29.23, 29.20, 29.1, 28.5, 28.2, 25.3, 22.7, 14.1. HRMS (ESI) m/z calculated for $C_{18}H_{38}OS_2$ $[M + H]^+$: 335.2436, found 335.2439. Melting Point: 38-39°C.



6-(Tetradecylsulfanyl)hexan-1-ol (SI-6)

Procedure A, thiol A = tetradecane-1-thiol, thiol B = 6-mercaptohexan-1-ol. Purification gradient: hexanes/EtOAc (0-10% EtOAc), 16 % yield. White solid. 1H NMR (400 MHz, $CDCl_3$) δ 3.65 (t, $J = 5.3$ Hz, 2H), 2.68 (ddd, $J = 14.4, 7.3, 3.1$ Hz, 4H), 1.74 - 1.55 (m, 6H), 1.48 - 1.21 (m, 26H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 62.9, 39.1, 38.9, 32.6, 31.9, 29.68, 29.66, 29.64 (2), 29.59, 29.50, 29.4, 29.23, 29.20, 29.1, 28.5, 28.2, 25.4, 22.7, 14.1. HRMS (ESI) m/z calculated for $C_{20}H_{43}OS_2$ $[M + H]^+$: 363.2749, found 363.2749. Melting Point: 42-43°C.

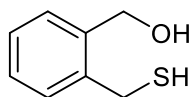


Methyl 2-((acetylthio)methyl)benzoate(SI-7)

To a stirring solution of methyl 2-(bromomethyl)benzoate (7.00 g, 30.6 mmol) in THF (100 mL) and DMF (5 mL) at 0°C was added potassium ethanethioate (3.84 g, 33.6 mmol) followed by catalytic tetrabutylammonium iodide (2.26 g, 6.11 mmol). The reaction mixture stirred at this temperature for 15 min, then naturally warmed to room temperature and stirred for 12 h. Progress was monitored by TLC (hexanes/EtOAc, 4:1). Then, the THF was evaporated under reduced pressure and the resulting oil was partitioned between EtOAc and brine (3x). The organic layer was collected, dried over anhydrous sodium sulfate, and purified on a silica column using a hexanes/EtOAc gradient (0-6% EtOAc) to afford the title compound methyl 2-((acetylthio)methyl)benzoate (6.53 g, 29.1 mmol, 95 % yield) as a foul-smelling yellow oil. 1H NMR (400 MHz, $CDCl_3$) δ 7.95 (d, $J = 7.8$ Hz, 1H), 7.51 (d, $J = 7.7$ Hz, 1H), 7.47 - 7.39 (m, 1H), 7.35 - 7.26 (m, 1H), 4.47 (s, 2H), 3.90 (s, 3H), 2.29 (s, 3H). ^{13}C

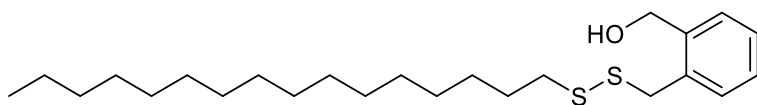
NMR (101 MHz, CDCl₃) δ 195.6, 167.3, 140.0, 132.5, 131.6, 131.0, 128.5, 127.4, 52.1, 32.2, 30.2.

HRMS (ESI) m/z calculated for C₁₁H₁₃O₃S [M + H]⁺ : 225.0579, found 225.0580.



(2-(Mercaptomethyl)phenyl)methanol (SI-8)

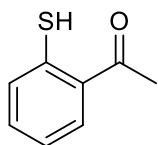
To a stirring solution of lithium aluminum hydride (82.0 mL, 163 mmol) in THF under inert atmosphere at 0°C was added methyl 2-((acetylthio)methyl)benzoate (9.15 g, 40.8 mmol), drop-wise. The solution stirred at this temperature for 15 min then naturally warmed to room temperature. Progress was monitored by LC-MS (H₂O/MeOH gradient, 75-95% MeOH, 3 min). The reaction reached completion after stirring for one hour and was subsequently chilled to 0°C. Then, acetone (9.00 mL, 122 mmol) was added slowly, drop-wise followed by the slow addition of 15% aqueous NaOH (37 mL) with vigorous stirring under inert atmosphere. The reaction mixture was then diluted with a saturated solution of sodium potassium tartrate and vigorously stirred for 2 h. After stirring, the pH was adjusted to 9 with solid ammonium chloride and the reaction mixture was allowed to settle. The supernatant was collected and concentrated under reduced pressure and the resulting aqueous mixture was partitioned with DCM. The solids were also partitioned with DCM and brine. The organic layers were collected, dried over anhydrous sodium sulfate, and purified via flash chromatography on a silica column using a hexanes/EtOAc gradient (0-41% EtOAc) to afford the title compound (2-(mercaptomethyl)phenyl)methanol (3.76 g, 24.38 mmol, 59.8 % yield) as a yellow oil. ¹H NMR (400 MHz, CDCl₃) δ 7.41 - 7.33 (m, 1H), 7.30 - 7.23 (m, 3H), 4.76 (s, 2H), 3.83 (d, J = 7.2 Hz, 2H), 1.87 (t, J = 7.2 Hz, 1H). ¹³C NMR (101 MHz, CDCl₃) δ 139.4, 138.2, 129.4, 129.3, 128.5, 127.8, 63.1, 26.1. HRMS (ESI) m/z calculated for C₈H₁₀SOCl [M + Cl]⁻ : 189.0146, found 189.0146.



(2-((Hexadecyldisulfanyl)methyl)phenyl)methanol (SI-9)

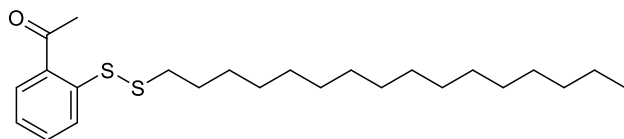
Procedure A, thiol A = hexadecane-1-thiol, thiol B = (2-(mercaptomethyl)phenyl)methanol.

Purification gradient: hexanes/EtOAc (0-7% EtOAc), 44.8 % yield. Pale yellow solid. ^1H NMR (400 MHz, CDCl_3) δ 7.44 - 7.37 (m, 1H), 7.34 - 7.21 (m, 3H), 4.79 (d, J = 5.6 Hz, 2H), 4.01 (s, 2H), 2.42 - 2.33 (m, 2H), 2.08 (t, J = 5.8 Hz, 1H), 1.53 (dt, J = 14.8, 7.3 Hz, 2H), 1.43 - 1.02 (m, 26H), 0.88 (t, J = 6.9 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 139.0, 135.2, 131.0, 129.1, 128.1, 127.9, 63.0, 40.4, 38.7, 31.9, 29.69 (2), 29.68 (2), 29.66, 29.65, 29.59, 29.48, 29.36, 29.16, 28.98, 28.46, 22.7, 14.1. HRMS (ESI) m/z calculated for $\text{C}_{24}\text{H}_{42}\text{OS}_2$ $[\text{M} + \text{Na}]^+$: 433.2569, found 433.2561. Melting Point: 38-40°C.

**1-(2-Mercaptophenyl)ethanone (SI-10)**

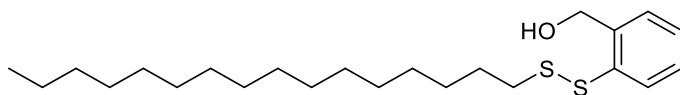
To a stirring solution of 2-mercaptobenzoic acid (4.00 g, 25.9 mmol) in anhydrous THF (130 mL) under nitrogen at 0 °C was added methyllithium (56.7 mL, 91 mmol) slowly, drop-wise. The reaction stirred at this temperature for 15 min then naturally warmed to room temperature. Progress was monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 75-95% MeOH, 3 min). After stirring for 5 h at room temperature, complete conversion was observed by monitoring the disappearance of the negative ionization peak of the starting material by LC-MS. The mixture was chilled to 0°C and quenched with water (10 mL). Then, the THF was evaporated under reduced pressure and the resulting aqueous layer was acidified with 1.2 M aqueous HCl to pH 5 and partitioned between DCM (100 mL, 3x). The organic layer was collected, dried under anhydrous sodium sulfate and purified on a silica column using a hexanes/EtOAc gradient (0-8%) to afford the title compound 1-(2-mercaptophenyl)ethanone (3.75 g, 24.64 mmol, 95 % yield) as an orange oil. ^1H NMR (400 MHz, CDCl_3) δ 7.88 (dd, J = 4.5, 3.9 Hz, 1H), 7.32 - 7.29 (m, 2H), 7.20 (ddd, J = 7.9, 5.3, 3.2 Hz, 1H), 4.46 - 4.42 (m, 1H), 2.62 (s, 3H). ^{13}C

NMR (101 MHz, CDCl₃) δ 198.8, 137.5, 132.6, 132.3, 131.73, 131.66, 124.7, 27.7. HRMS (ESI) m/z calculated for C₈H₈OSNa [M + Na]⁺ : 175.0188, found 175.0188.



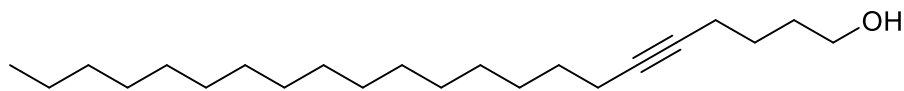
1-(2-(Hexadecyldisulfanyl)phenyl)ethanone (SI-11)

Procedure A, thiol A = hexadecane-1-thiol, thiol B = 1-(2-mercaptophenyl)ethanone. Purification gradient: hexanes/EtOAc (0-4% EtOAc), 22.6 % yield. Pale yellow solid. ¹H NMR (400 MHz, CDCl₃) δ 8.28 (dd, J = 8.2, 1.1 Hz, 1H), 7.90 (dd, J = 7.8, 1.4 Hz, 1H), 7.57 (ddd, J = 8.3, 7.3, 1.4 Hz, 1H), 7.28 (td, J = 7.7, 1.1 Hz, 1H), 2.73 - 2.68 (m, 2H), 2.64 (s, 3H), 1.66 (dt, J = 15.0, 7.3 Hz, 2H), 1.42 - 1.17 (m, 26H), 0.88 (t, J = 6.9 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 198.9, 142.4, 134.2, 132.6, 131.4, 126.1, 124.8, 38.2, 31.9, 29.69, 29.68 (2), 29.65 (2), 29.62, 29.57, 29.48, 29.36, 29.2, 29.1, 28.6, 27.5, 22.7, 14.1. HRMS (ESI) m/z calculated for C₂₄H₄₁OS₂ [M + H]⁺ : 409.2593, found 409.2588. Melting Point: 51-52 °C.



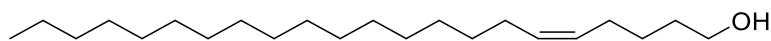
(2-(Hexadecyldisulfanyl)phenyl)methanol (SI-12)

Procedure A, thiol A = hexadecane-1-thiol, thiol B = (2-mercaptophenyl)methanol. Purification gradient: hexanes/EtOAc (0-6% EtOAc), 49.3 % yield. Off-white solid. ¹H NMR (400 MHz, CDCl₃) δ 7.78 - 7.72 (m, 1H), 7.46 - 7.40 (m, 1H), 7.33 - 7.25 (m, 2H), 4.84 (s, 2H), 2.75 - 2.69 (m, 2H), 1.65 (dt, J = 14.9, 7.3 Hz, 2H), 1.39 - 1.17 (m, 26H), 0.88 (t, J = 6.9 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 140.2, 135.7, 129.78, 128.5, 128.2, 127.7, 63.3, 38.7, 31.9, 29.69 (3), 29.67 (2), 29.63, 29.57, 29.46, 29.37, 29.2, 28.7, 28.4, 22.7, 14.1. HRMS (ESI) m/z calculated for C₂₃H₄₀S₂O [M + Na]⁺ : 419.2412, found 419.2409. Melting Point: 40-42 °C.



Docos-5-yn-1-ol (SI-13)

To a stirring solution of hex-5-yn-1-ol (4.50 g, 45.9 mmol) in anhydrous THF (153 mL) was added butyllithium (38.5 ml, 96 mmol) at 0 °C. The yellow mixture gelatinized at this temperature and required warming to free the stir bar. After stirring for 15 minutes at room temperature, 1-bromohexadecane (9.59 ml, 45.9 mmol) was added in one portion. The reaction exothermed and was quickly placed in an ice bath. After 15 minutes, the ice bath was removed and reaction progress was monitored by TLC (4:1 hexanes/EtOAc, KMnO₄). After stirring overnight at room temperature, no further progress occurred by TLC and a significant amount of starting material remained. The mixture was quenched with dilute acid and the organic solvent was evaporated under reduced pressure. The resulting orange gum was partitioned between brine and EtOAc. Then, the organic layer was collected, dried over anhydrous sodium sulfate, concentrated, then purified on silica gel using a hexanes/EtOAc gradient (0-8% EtOAc) to furnish the title compound docos-5-yn-1-ol (2.97 g, 9.21 mmol, 20 % yield) as a white solid. ¹H NMR (400 MHz, CDCl₃) δ 3.68 (t, J = 6.4 Hz, 2H), 2.41 - 2.29 (m, 1H), 2.06 (d, J = 3.1 Hz, 1H), 1.88 - 1.41 (m, 10H), 1.35 - 1.20 (m, 24H), 0.88 (t, J = 6.7 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 87.8, 69.4, 62.7, 35.0, 31.9, 31.3, 31.1, 30.5, 29.67 (4), 29.64 (3), 29.61, 29.55, 29.46, 29.34, 27.2, 22.7, 14.1. HRMS (ESI) m/z calculated for C₂₂H₄₃O [M + H]⁺ : 323.3308, found 323.3307. Melting Point: 59-61 °C.

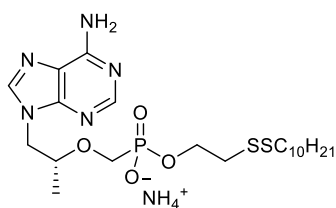


(Z)-Docos-5-en-1-ol (SI-14)

To a stirring solution of docos-5-yn-1-ol (1.00 g, 3.10 mmol) in THF (35 mL) was added quinoline (0.073 ml, 0.620 mmol) and 5% palladium on carbon (200 mg, 20% by weight). The resulting mixture

was placed under an atmosphere of hydrogen gas and stirred at room temperature with progress monitored by TLC (hexanes/EtOAc 4:1, KMnO₄ stain). Complete conversion was observed by TLC after 2 h. Note that the product is slightly more polar than the starting alkyne. The mixture was filtered over celite and the supernatant was collected and concentrated under reduced pressure. Then, the residue was taken up with DCM and washed with dilute aqueous acid to remove the quinoline. The organic layer was again collected, dried over anhydrous sodium sulfate, and concentrated to afford the title compound (Z)-docos-5-en-1-ol (1.00 g, 3.08 mmol, 99 % yield) as an off-white solid that required no additional purification. ¹H NMR (400 MHz, CDCl₃) δ 5.49 (ddd, *J* = 16.9, 10.4, 8.8 Hz, 1H), 4.94 (ddd, *J* = 16.9, 11.2, 2.1 Hz, 1H), 3.60 (t, *J* = 6.5 Hz, 2H), 1.92 (dtd, *J* = 13.1, 8.6, 4.2 Hz, 1H), 1.64 - 1.37 (m, 4H), 1.37 - 1.14 (m, 31H), 0.86 (t, *J* = 6.8 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 143.1, 114.3, 62.9, 44.0, 35.1, 31.9, 31.0, 30.4, 29.8, 29.70 (5), 29.69 (2), 29.66 (2), 29.37, 27.2, 22.3, 14.1. HRMS (ESI) *m/z* calculated for C₂₂H₄₅O [M + H]⁺ : 325.3464, found 325.3460. Melting Point: 33-34 °C.

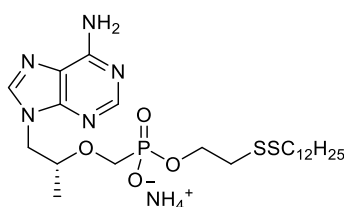
1.11.8 Synthesis and Characterization of TFV Conjugates



2-(Decyldisulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.6a)

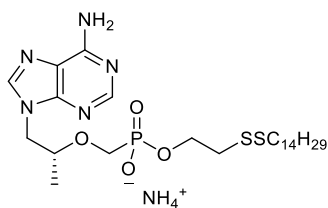
Procedure B, X = 2-(decyldisulfanyl)ethanol, Y = H₂O, 49 % yield. Amorphous solid. ¹H NMR (400 MHz, CD₃OD) δ 8.36 (s, 1H), 8.24 (s, 1H), 4.42 (dd, *J* = 14.4, 3.1 Hz, 1H), 4.27 (dd, *J* = 14.4, 6.7 Hz, 1H), 4.08 – 4.00 (m, 2H), 3.95 (ddd, *J* = 9.5, 6.3, 3.1 Hz, 1H), 3.76 (dd, *J* = 12.8, 9.4 Hz, 1H), 3.53 (dd,

$J = 12.8, 10.0$ Hz, 1H), 2.89 – 2.76 (m, 2H), 2.69 (dd, $J = 15.7, 8.5$ Hz, 2H), 1.65 (dt, $J = 14.7, 7.2$ Hz, 2H), 1.44 – 1.24 (m, 14H), 1.21 (d, $J = 6.2$ Hz, 3H), 0.92 (t, $J = 6.9$ Hz, 3H). Carbon referenced to CD_3OD (49.0 ppm). ^{13}C NMR (100 MHz, CD_3OD) δ 157.4, 153.3, 151.7, 145.2, 120.4, 77.86 (d, $J = 12.9$ Hz), 66.4 (d, $J = 160.4$ Hz), 65.2 (d, $J = 5.6$ Hz), 50.0, 41.2 (d, $J = 5.9$ Hz), 40.8, 33.9, 31.5, 31.3, 31.2, 31.0, 30.3, 24.6, 17.8, 15.3. ^{31}P NMR (162 MHz, CD_3OD) δ 16.9. HRMS (ESI) m/z calculated for $\text{C}_{21}\text{H}_{39}\text{N}_5\text{O}_4\text{PS}_2$ $[\text{M} + \text{H}]^+$: 520.2175, found 520.2178. Anal. calculated for $\text{C}_{21}\text{H}_{43}\text{N}_6\text{O}_5\text{PS}_2$ (as a monohydrate ammonium salt): C, 45.47; H, 7.81; N, 15.15. Found: C, 45.56; H, 7.53; N, 15.10.



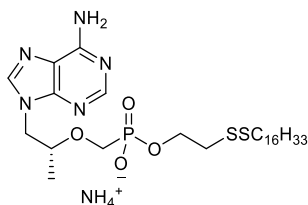
2-(Dodecylsulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.6b)

Procedure B, X = 2-(dodecylsulfanyl)ethanol, Y = H_2O , 21 % yield. Amorphous solid. ^1H NMR (399 MHz, CD_3OD) δ 8.32 (s, 1H), 8.21 (s, 1H), 4.39 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.23 (dd, $J = 14.4, 6.8$ Hz, 1H), 4.06 - 3.96 (m, 2H), 3.91 (td, $J = 6.5, 3.1$ Hz, 1H), 3.73 (dd, $J = 12.8, 9.4$ Hz, 1H), 3.50 (dd, $J = 12.8, 10.1$ Hz, 1H), 2.88 - 2.72 (m, 2H), 2.65 (dd, $J = 14.8, 7.6$ Hz, 2H), 1.67 - 1.55 (m, 2H), 1.41 - 1.22 (m, 18H), 1.18 (d, $J = 6.2$ Hz, 3H), 0.89 (t, $J = 6.9$ Hz, 3H). Carbon referenced to CD_3OD signal (49.00 ppm). ^{13}C NMR (100 MHz, CD_3OD) δ 156.8, 152.8, 150.9, 144.3, 119.6, 77.0 (d, $J = 12.9$ Hz), 65.5 (d, $J = 160.7$ Hz), 64.4 (d, $J = 5.6$ Hz), 49.1, 40.4 (d, $J = 5.9$ Hz), 40.0, 33.0, 30.74, 30.72, 30.68, 30.62, 30.4, 30.3, 30.2, 29.5, 23.7, 17.0, 14.4. ^{31}P NMR (162 MHz, CD_3OD) δ 17.0. HRMS (ESI) m/z calculated for $\text{C}_{23}\text{H}_{43}\text{N}_5\text{O}_4\text{PS}_2$ $[\text{M} + \text{H}]^+$: 548.2488, found 548.2490. Anal. calculated for $\text{C}_{23}\text{H}_{45}\text{N}_6\text{O}_4\text{PS}_2$ (as an ammonium salt): C, 48.92; H, 8.03; N, 14.88. Found: C, 48.51; H, 7.96; N, 14.06.



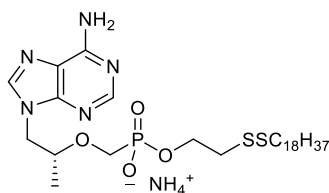
2-(Tetradecyldisulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.6c)

Procedure B, X = 2-(tetradecyldisulfanyl)ethanol, Y = H₂O, 14 % yield. Amorphous solid. Proton spectrum referenced to CD₃OH (3.31 ppm). ¹H NMR (399 MHz, CD₃OD/CDCl₃) δ 8.33 (s, 1H), 8.21 (s, 1H), 4.39 (dd, *J* = 14.4, 3.0 Hz, 1H), 4.22 (dd, *J* = 14.4, 6.8 Hz, 1H), 4.03 (q, *J* = 6.9 Hz, 2H), 3.90 (td, *J* = 6.5, 3.0 Hz, 1H), 3.75 (dd, *J* = 12.8, 9.4 Hz, 1H), 3.50 (dd, *J* = 12.7, 10.0 Hz, 1H), 2.88 - 2.73 (m, 2H), 2.65 (dd, *J* = 14.0, 6.7 Hz, 2H), 1.68 - 1.56 (m, 2H), 1.41 - 1.21 (m, 24H), 1.18 (d, *J* = 6.2 Hz, 3H), 0.88 (t, *J* = 6.9 Hz, 3H). Carbon spectrum referenced to CD₃OD (49.0 ppm). ¹³C NMR (100 MHz, CD₃OD/CDCl₃) δ 155.8, 151.6, 150.5, 144.4, 119.3, 76.7 (d, *J* = 12.7 Hz), 65.3 (d, *J* = 160.0 Hz), 64.1 (d, *J* = 5.7 Hz), 40.2 (d, *J* = 6.0 Hz), 39.81, 32.82, 30.53, 30.51, 30.46, 30.40, 30.2, 30.1, 30.0, 29.3, 23.5, 16.8, 14.4. ³¹P NMR (162 MHz, CD₃OD/CDCl₃) δ 16.8. HRMS (ESI) *m/z* calculated for C₂₅H₄₇N₅O₄PS₂ [M + H]⁺ : 576.2801, found 576.2803. Anal. calculated for C₂₅H₅₀N₆O₅PS₂ (as a monohydrate ammonium salt): C, 49.24; H, 8.26; N, 13.78. Found: C, 49.33; H, 8.19; N, 13.55. Melting Point: 149-154 °C.



2-(Hexadecyldisulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.6d)

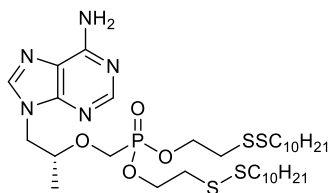
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = H₂O, 28 % yield. White solid. ¹H NMR (400 MHz, CD₃OD) δ 8.32 (s, 1H), 8.20 (s, 1H), 4.43 - 4.33 (m, 1H), 4.23 (dd, J = 14.4, 6.7 Hz, 1H), 4.00 (qd, J = 7.0, 1.7 Hz, 2H), 3.90 (pd, J = 6.3, 3.1 Hz, 1H), 3.73 (dd, J = 12.8, 9.4 Hz, 1H), 3.49 (dd, J = 12.7, 10.1 Hz, 1H), 2.86 - 2.73 (m, 2H), 2.64 (dd, J = 7.7, 6.8 Hz, 2H), 1.68 - 1.55 (m, 2H), 1.39 - 1.25 (m, 26H), 1.17 (d, J = 6.2 Hz, 3H), 0.94 - 0.85 (m, 3H). ¹³C NMR (101 MHz, CD₃OD) δ 155.4, 151.5, 149.41, 142.9, 139.9, 118.1, 75.5 (d, J = 13.1 Hz), 64.1 (d, J = 160.2 Hz), 63.0 (d, J = 5.7 Hz), 38.9 (d, J = 6.2 Hz), 38.5, 31.7, 29.38, 29.35, 29.34, 29.28, 29.23, 29.1, 28.9, 28.7, 28.1, 22.3, 15.4, 13.0. ³¹P NMR (162 MHz, CD₃OD) δ 16.9. HRMS (ESI) *m/z* calculated for C₂₇H₅₁O₄N₅PS₂ [M + H]⁺: 604.3114, found 604.3114. Anal. calculated for C₂₇H₅₃O₅N₆PS₂ (as an ammonium salt monohydrate): C, 50.92; H, 8.39; N, 13.20. Found: C, 51.45; H, 8.42; N, 12.79. Melting Point: 138-142°C.



2-((Octadecyldisulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.6e)

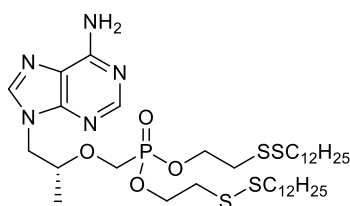
Procedure B, X = 2-(octadecyldisulfanyl)ethanol, Y = H₂O, 25 % yield. Amorphous solid. Proton referenced to residual TMS. ¹H NMR (399 MHz, CD₃OD/CDCl₃) δ 8.23 (s, 1H), 8.12 (s, 1H), 4.30 (dd, J = 14.4, 2.9 Hz, 1H), 4.13 (dd, J = 14.4, 6.7 Hz, 1H), 3.93 (q, J = 6.9 Hz, 2H), 3.81 (td, J = 6.3, 3.0 Hz, 1H), 3.65 (dd, J = 12.7, 9.4 Hz, 1H), 3.40 (dd, J = 12.6, 10.1 Hz, 1H), 2.72 (td, J = 6.8, 2.9 Hz, 2H), 2.56 (dd, J = 14.3, 6.9 Hz, 2H), 1.59 - 1.45 (m, 2H), 1.28 - 1.13 (m, 30H), 1.09 (d, J = 6.2 Hz, 3H), 0.79 (t, J = 6.6 Hz, 3H). Carbon referenced to CD₃OD signal (49.0 ppm). ¹³C NMR (100 MHz, CD₃OD/CDCl₃) δ 156.4, 152.5, 150.6, 144.1, 119.4, 76.7 (d, J = 12.7 Hz), 65.3 (d, J = 159.9 Hz), 64.1 (d, J = 5.6 Hz), 49.0, 40.2 (d, J = 6.8 Hz), 39.8, 32.8, 30.5, 30.46, 30.40, 30.2, 30.1, 30.0, 29.3, 23.5,

16.8, 14.4. ^{31}P NMR (162 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$) δ 16.8. HRMS (ESI) m/z calculated for $\text{C}_{29}\text{H}_{55}\text{N}_5\text{O}_4\text{PS}_2$ $[\text{M} + \text{H}]^+$: 632.3427, found 632.3431. Anal. calculated for $\text{C}_{29}\text{H}_{59}\text{N}_6\text{O}_5\text{PS}_2$ (as a monohydrate ammonium salt): C, 52.23; H, 8.92; N, 12.60. Found: C, 52.48; H, 8.88; N, 12.66



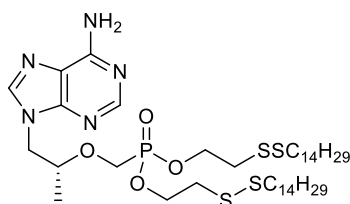
2-(Decyldisulfanyl)ethyl (2-(decylsulfinothioyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.7a)

Procedure B, X = Y = 2-(decyldisulfanyl)ethanol, 37.3 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.33 (s, 1H), 7.96 (s, 1H), 6.10 (s, 2H), 4.35 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.26 (ddt, $J = 14.7, 11.5, 5.2$ Hz, 4H), 4.13 (dd, $J = 14.4, 7.6$ Hz, 1H), 3.98 - 3.92 (m, 1H), 3.89 (dd, $J = 13.7, 8.9$ Hz, 1H), 3.64 (dd, $J = 13.7, 9.5$ Hz, 1H), 2.86 (dt, $J = 16.6, 6.7$ Hz, 4H), 2.70 - 2.63 (m, 4H), 1.69 - 1.57 (m, 4H), 1.41 - 1.10 (m, 31H), 0.86 (t, $J = 6.8$ Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.6, 152.9, 150.1, 141.7, 119.2, 76.5 (d, $J = 11.9$ Hz), 64.2 (apparent t, $J = 6.6$ Hz), 62.7 (d, $J = 168.5$ Hz), 48.1, 39.07, 39.06, 38.38 (d, $J = 5.7$ Hz), 38.34 (d, $J = 5.8$ Hz), 31.9, 29.51, 29.48, 29.3, 29.2, 29.1, 28.5, 22.6, 16.5, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 21.4. HRMS (ESI) m/z calculated for $\text{C}_{33}\text{H}_{63}\text{N}_5\text{O}_4\text{PS}_4$ $[\text{M} + \text{H}]^+$: 752.3495, found 752.3509. Anal. calculated for $\text{C}_{33}\text{H}_{62}\text{N}_5\text{O}_4\text{PS}_4$: C, 52.70; H, 8.31; N, 9.31. Found: C, 52.80; H, 8.46; N, 9.29.



2-(Dodecyldisulfanyl)ethyl (2-(dodecylsulfinothioyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.7b)

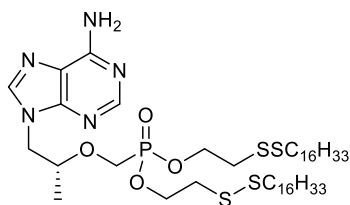
Procedure B, X = Y = 2-(dodecyldisulfanyl)ethanol, 30.8 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.35 (s, 1H), 7.96 (s, 1H), 5.90 (s, 2H), 4.36 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.33 – 4.21 (m, 4H), 4.14 (dd, $J = 14.5, 7.6$ Hz, 1H), 4.00 – 3.93 (m, 1H), 3.90 (dd, $J = 13.7, 8.9$ Hz, 1H), 3.64 (dd, $J = 13.7, 9.5$ Hz, 1H), 2.87 (dt, $J = 17.8, 6.7$ Hz, 4H), 2.72 – 2.62 (m, 4H), 1.70 – 1.59 (m, 4H), 1.44 – 1.11 (m, 39H), 0.87 (t, $J = 6.9$ Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.7, 152.9, 150.0, 141.7, 119.2, 76.5 (d, $J = 12.0$ Hz), 64.3 (d, $J = 5.4$ Hz), 64.2 (d, $J = 6.4$ Hz), 62.7 (d, $J = 168.5$ Hz), 48.1, 39.04, 39.03, 38.37 (d, $J = 3.4$ Hz), 38.31 (d, $J = 3.4$ Hz), 31.86, 29.6, 29.58, 29.55, 29.47, 29.3, 29.2, 29.1, 28.5, 22.6, 16.5, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 21.4. HRMS (ESI) m/z calculated for $\text{C}_{37}\text{H}_{71}\text{N}_5\text{O}_4\text{PS}_4$ $[\text{M} + \text{H}]^+$: 808.4121, found 808.4133. Anal. calculated for $\text{C}_{37}\text{H}_{70}\text{N}_5\text{O}_4\text{PS}_4$: C, 54.98; H, 8.73; N, 8.67. Found: C, 54.31; H, 8.97; N, 8.43. Melting Point: 41 - 42 °C.



2-(Tetradecyldisulfanyl)ethyl (2-(tetradecylsulfinothioyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl phosphonate (1.7c)

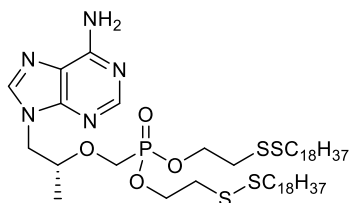
Procedure B, X = Y = (tetradecyldisulfanyl)ethanol, 13 % yield. White foam. ^1H NMR (400 MHz, CDCl_3) δ 8.31 (s, 1H), 7.98 (s, 1H), 6.40 (s, 2H), 4.41 - 4.16 (m, 5H), 4.12 (dd, $J = 14.4, 7.6$ Hz, 1H), 4.00 - 3.81 (m, 2H), 3.62 (dd, $J = 13.7, 9.5$ Hz, 1H), 2.85 (dt, $J = 16.7, 6.7$ Hz, 4H), 2.70 - 2.60 (m, 4H), 1.70 - 1.54 (m, 4H), 1.42 - 1.13 (m, 47H), 0.91 - 0.77 (m, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.0, 151.7, 149.9, 142.1, 119.1, 76.4 (d, $J = 11.7$ Hz), 64.3 (d, $J = 12.8$ Hz), 64.2, 62.6 (d, $J = 168.7$ Hz), 48.2, 39.06, 39.05, 38.37 (d, $J = 5.9$ Hz), 38.34 (d, $J = 5.8$ Hz), 31.9, 29.67, 29.66, 29.63, 29.59, 29.51, 29.3, 29.2, 29.1, 28.5, 22.7, 16.5, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 21.2. HRMS (ESI) m/z calculated for $\text{C}_{41}\text{H}_{79}\text{N}_5\text{O}_4\text{PS}_4$ $[\text{M} + \text{H}]^+$: 864.4747, found 864.4751. Anal. calculated for

$C_{41}H_{78}N_5O_4PS_4$: C, 57.00; H, 9.10; N, 8.10. Found: C, 56.00; H, 9.23; N, 7.46. Melting Point: 29-31 °C.



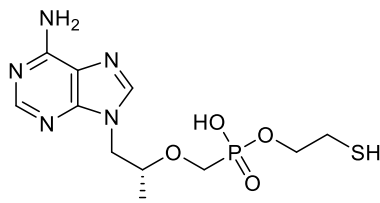
2-(Hexadecyldisulfanyl)ethyl (2-(hexadecylsulfinothioyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.7d)

Procedure B, X = Y = 2-(hexadecyldisulfanyl)ethanol, 30.8 % yield. Amorphous solid. 1H NMR (400 MHz, $CDCl_3$) δ 8.36 (s, 1H), 8.06 (s, 1H), 6.07 (s, 2H), 4.37 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.33 - 4.22 (m, 4H), 4.16 (dd, $J = 14.4, 7.6$ Hz, 1H), 3.99 - 3.93 (m, 1H), 3.90 (dd, $J = 13.7, 9.0$ Hz, 1H), 3.65 (dd, $J = 13.7, 9.5$ Hz, 1H), 2.88 (dt, $J = 17.6, 6.7$ Hz, 4H), 2.73 - 2.63 (m, 4H), 1.70 - 1.59 (m, 4H), 1.46 - 1.14 (m, 55H), 0.87 (t, $J = 6.9$ Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.5, 153.0, 150.0, 142.1, 118.7, 76.3 (d, $J = 12.0$ Hz), 64.3 (d, $J = 4.6$ Hz), 64.2 (d, $J = 4.6$ Hz), 62.6 (d, $J = 168.7$ Hz), 48.2, 39.03, 39.02, 38.38 (d, $J = 2.4$ Hz), 38.32 (d, $J = 2.5$ Hz), 31.9, 29.67, 29.63, 29.59, 29.50, 29.3, 29.2, 29.1, 28.5, 22.7, 16.5, 14.1. ^{31}P NMR (162 MHz, $CDCl_3$) δ 21.4. HRMS (ESI) m/z calculated for $C_{45}H_{87}N_5O_4PS_4$ $[M + H]^+$: 920.5373, found 920.5398. Anal. calculated for $C_{45}H_{86}N_5O_4PS_4$: C, 58.72; H, 9.42; N, 7.61. Found: C, 58.05; H, 9.50; N, 7.44.



2-(Octadecyldisulfanyl)ethyl (2-(octadecylsulfinothioyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.7e)

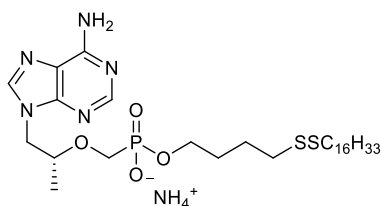
Procedure B, X = Y = 2-(octadecyldisulfanyl)ethanol, 33.6 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.35 (s, 1H), 7.97 (s, 1H), 5.86 (s, 2H), 4.36 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.27 (ddt, $J = 16.3, 14.7, 5.2$ Hz, 5H), 4.14 (dd, $J = 14.4, 7.7$ Hz, 1H), 3.99 - 3.93 (m, 1H), 3.90 (dd, $J = 13.7, 8.9$ Hz, 1H), 3.65 (dd, $J = 13.7, 9.5$ Hz, 1H), 2.87 (dt, $J = 18.1, 6.7$ Hz, 4H), 2.72 - 2.63 (m, 4H), 1.70 - 1.59 (m, 4H), 1.45 - 1.16 (m, 63H), 0.87 (t, $J = 6.9$ Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.6, 152.8, 150.0, 141.7, 119.2, 76.5 (d, $J = 11.9$ Hz), 64.3 (d, $J = 5.8$ Hz), 64.2 (d, $J = 5.7$ Hz), 62.7 (d, $J = 168.4$ Hz), 48.1, 39.04, 39.03, 38.36 (d, $J = 5.8$ Hz), 38.32 (d, $J = 5.7$ Hz), 31.9, 29.67, 29.62, 29.58, 29.50, 29.3, 29.2, 29.1, 28.5, 22.7, 16.5, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 21.4. HRMS (ESI) m/z calculated for $\text{C}_{49}\text{H}_{95}\text{N}_5\text{O}_4\text{PS}_4$ $[\text{M} + \text{H}]^+$: 976.5999, found 976.6033. Anal. calculated for $\text{C}_{49}\text{H}_{94}\text{N}_5\text{O}_4\text{PS}_4$: C, 60.27; H, 9.70; N, 7.17. Found: C, 59.99; H, 9.82; N, 7.14. Melting Point: 56-57°C.



2-Mercaptoethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.8)

To a stirring solution of dry tenofovir (1.00 g, 3.48 mmol) in anhydrous DCM (100 mL) and *N,N*-dimethylformamide (0.537 mL, 6.96 mmol) was gradually added excess oxalyl chloride (1.19 mL, 13.93 mmol) at room temperature and stirred for 1 h. The mixture was then cooled to 0°C, and quenched with excess 2-mercaptoethanol (2.451 mL, 34.8 mmol). Progress was monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 50-95% MeOH, 3 min). The mixture stirred for an additional hour at this temperature. Aqueous HCl was then added (3 mL) followed by 15 mL of methanol (pH = 1). The mixture stirred at room temperature overnight, then the pH was gradually raised to 5 using saturated aqueous sodium bicarbonate. The organic solvents were evaporated under reduced pressure at 30°C and the resulting residue was purified on a C18 reverse phase column using a $\text{H}_2\text{O}/\text{MeOH}$ gradient

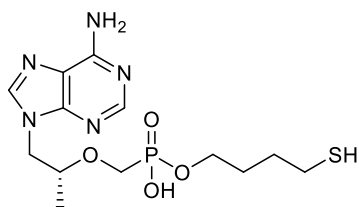
(isocratic 10% MeOH) to afford the title compound 2-mercaptoethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (0.600 g, 1.74 mmol, 50 % yield) as a white solid. ^1H NMR (400 MHz, D_2O) δ 8.38 (s, 1H), 8.37 (s, 1H), 4.44 (dd, $J = 14.7, 3.1$ Hz, 1H), 4.26 (dd, $J = 14.7, 7.7$ Hz, 1H), 4.04 – 3.92 (m, 1H), 3.77 – 3.66 (m, 3H), 3.48 (dd, $J = 13.4, 9.6$ Hz, 1H), 2.53 (td, $J = 6.4, 4.2$ Hz, 2H), 1.19 (d, $J = 6.3$ Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 149.7, 148.5, 145.5, 144.4, 117.5, 75.6 (d, $J = 12.5$ Hz), 66.1 (d, $J = 5.6$ Hz), 62.9 (d, $J = 159.5$ Hz), 48.4, 24.4 (d, $J = 6.5$ Hz), 15.6. ^{31}P NMR (162 MHz, D_2O) δ 20.8. HRMS (ESI) m/z calculated for $\text{C}_{11}\text{H}_{19}\text{O}_4\text{N}_5\text{PS}$ $[\text{M} + \text{H}]^+$: 348.0889, found 348.0892. Anal. calculated for $\text{C}_{11}\text{H}_{20}\text{N}_5\text{O}_5\text{PS}$ (as a monohydrate): C, 36.16; H, 5.52; N, 19.17. Found: C, 35.73; H, 5.53; N, 19.07. Melting point: Decomposes at 100°C.



(*R*)-4-(Hexadecyldisulfanyl)butyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.9)

Procedure B, X = 4-(hexadecyldisulfanyl)butan-1-ol, Y = H_2O , 63.7 % yield. Amorphous solid. ^1H and ^{13}C spectra referenced to CD_3OD (3.31 δ and 49.15 δ , respectively). ^1H NMR (400 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 8.30 (s, 1H), 8.20 (s, 1H), 4.37 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.22 (dd, $J = 14.4, 6.8$ Hz, 1H), 3.94 - 3.82 (m, 1H), 3.82 - 3.65 (m, 3H), 3.45 (dd, $J = 12.7, 10.2$ Hz, 1H), 2.68 - 2.58 (m, 4H), 1.74 - 1.54 (m, 6H), 1.42 - 1.20 (m, 25H), 1.18 (d, $J = 6.2$ Hz, 3H), 0.88 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 156.9, 153.2, 150.8, 144.1, 119.5, 77.0 (d, $J = 13.2$ Hz), 65.4 (d, $J = 160.3$ Hz), 65.4 (d, $J = 5.9$ Hz), 66.2, 65.4, 65.3, 64.6, 39.7, 39.3, 33.1, 30.81, 30.80, 30.79, 30.78, 30.74, 30.69 (d, $J = 5.9$ Hz), 30.5, 30.4, 30.2, 29.5, 26.6, 23.8, 17.0, 14.6. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 16.5. HRMS (ESI) m/z calculated for $\text{C}_{29}\text{H}_{55}\text{N}_5\text{O}_4\text{PS}_2$ $[\text{M} + \text{H}]^+$: 632.3427, found 632.3451. Anal.

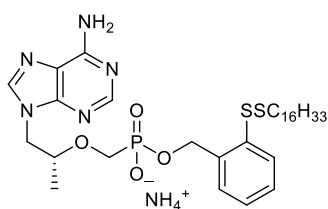
calculated for $C_{29}H_{58}N_6O_5PS_2$ (NH^{4+} monohydrate): C, 52.31; H, 8.78; N, 12.62. Found: C, 52.49; H, 8.78; N, 12.64.



4-Mercaptobutyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.10)

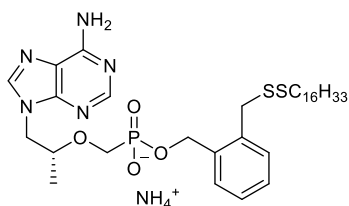
To a stirring solution of tenofovir (1.0 g, 3.5 mmol) in anhydrous DCM (34.8 mL) and *N,N*-dimethylformamide (0.32 mL, 4.2 mmol) was added excess oxalyl chloride (1.50 mL, 17.41 mmol) at room temperature. The mixture stirred for 15 min with progress monitored with LC-MS ($H_2O/MeOH$ gradient, 35-75% MeOH, 3 min) by quenching an aliquot of the reaction mixture with MeOH. Upon completion, the solvents and excess oxalyl chloride were evaporated under reduced pressure and the resulting yellow foam was further dried under UHV. The foam was re-dissolved in anhydrous DCM (34 mL) and the mixture was chilled to 0°C. Then, a solution of 4-mercaptobutan-1-ol (0.40 mL, 3.83 mmol) and pyridine (1.7 mL, 21 mmol) in DCM (2 mL) was added drop-wise and the reaction stirred at this temperature for 15 min then naturally warmed to room temperature with progress again monitored by LC-MS. After stirring for 2 h, the mixture was quenched with water (0.25 mL) and stirring continued at room temperature for 30 min. The solution was then acidified with HCl (1.2 M) and homogenized with MeOH until the aqueous/organic interface disappeared and stirred overnight to facilitate the cleavage of formimidine. The solvents were then evaporated under reduced pressure and the resulting yellow oil was purified on a C18 reverse phase column using a $H_2O/MeOH$ gradient with 0.1 % formic acid (0-25% MeOH) to afford the title compound 4-mercaptobutyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (0.71 g, 1.9 mmol, 54 % yield) as a white foam. 1H NMR (400 MHz, CD_3OD) δ 8.40 (s, 1H), 8.27 (s, 1H), 4.46 (dd, $J =$

14.4, 3.0 Hz, 1H), 4.27 (dd, $J = 14.5, 7.1$ Hz, 1H), 4.08 - 3.98 (m, 1H), 3.83 (ddd, $J = 22.4, 12.6, 7.7$ Hz, 3H), 3.63 (dd, $J = 13.0, 9.5$ Hz, 1H), 2.99 (t, $J = 6.9$ Hz, 1H), 2.49 (t, $J = 6.7$ Hz, 2H), 1.72 - 1.58 (m, 4H), 1.19 (d, $J = 6.3$ Hz, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 152.2, 150.4, 146.3, 146.2, 119.2, 76.6 (d, $J = 12.2$ Hz), 65.6 (d, $J = 6.0$ Hz), 64.7 (d, $J = 162.9$ Hz), 31.5, 30.6 (d, $J = 6.1$ Hz), 27.4, 24.7, 17.0. ^{31}P NMR (162 MHz, CD_3OD) δ 17.7. HRMS (ESI) m/z calculated for $\text{C}_{13}\text{H}_{23}\text{N}_5\text{O}_4\text{PS}$ $[\text{M} + \text{H}]^+$: 376.1202, found 376.1198. Anal. calculated for $\text{C}_{13}\text{H}_{24}\text{O}_5\text{N}_5\text{PS}$ (as a monohydrate): C, 39.69; H, 6.15; N, 17.80. Found: C, 39.88; H, 5.73; N, 17.76. Melting Point: 151-152°C.



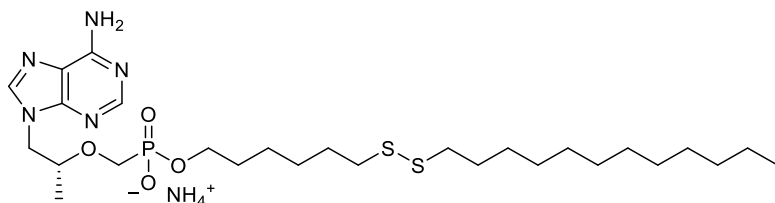
(*R*)-2-(Hexadecyldisulfanyl)benzyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.11)

Procedure B, X = (2-(hexadecyldisulfanyl)phenyl)methanol, Y = H_2O , 40.8 % yield. Off-white solid. ^1H NMR (400 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$, referenced to CD_3OD) δ 8.26 (s, 1H), 8.18 (s, 1H), 7.67 (dd, $J = 7.7, 1.3$ Hz, 1H), 7.46 (dd, $J = 7.5, 1.3$ Hz, 1H), 7.22 (dtd, $J = 21.2, 7.4, 1.5$ Hz, 2H), 5.13 - 5.03 (m, 2H), 4.32 (dd, $J = 14.4, 3.1$ Hz, 1H), 4.17 (dd, $J = 14.4, 6.5$ Hz, 1H), 3.88 - 3.81 (m, 1H), 3.77 (dd, $J = 12.7, 9.4$ Hz, 1H), 3.53 (dd, $J = 12.7, 10.0$ Hz, 1H), 2.65 (t, $J = 7.2$ Hz, 2H), 1.57 (dt, $J = 14.8, 7.2$ Hz, 2H), 1.34 - 1.14 (m, 26H), 1.11 (d, $J = 6.3$ Hz, 3H), 0.86 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$, referenced to CD_3OD) δ 156.2, 152.3, 150.4, 143.9, 138.8 (d, $J = 6.9$ Hz), 135.9, 129.88, 128.86, 128.77, 128.1, 119.2, 76.8 (d, $J = 13.2$ Hz), 65.5 (d, $J = 159.8$ Hz), 65.2 (d, $J = 5.0$ Hz), 48.9, 39.6, 32.85, 30.60, 30.59, 30.56, 30.54, 30.48, 30.4, 30.28, 30.1, 29.6, 29.2, 23.6, 16.8, 14.5. ^{31}P NMR (121 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 16.5. HRMS (ESI) m/z calculated for $\text{C}_{32}\text{H}_{51}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} - \text{H}]^-$: 664.3114, found 664.3130. Anal. calculated for $\text{C}_{32}\text{H}_{57}\text{N}_6\text{O}_5\text{PS}_2$ (NH_4^+ monohydrate): C, 54.83; H, 8.20; N, 11.99. Found: C, 54.23; H, 8.16; N, 12.38. Melting Point: 177-180°C.



(R)-2-((Hexadecyldisulfanyl)methyl)benzyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.12)

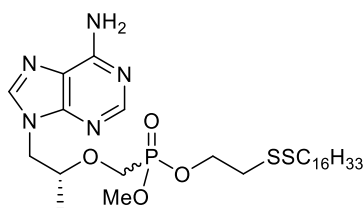
Procedure B, X = (2-((hexadecyldisulfanyl)methyl)phenyl)methanol, Y = H₂O, 48 % yield. Amorphous solid. Proton spectrum referenced to CD₃OD (3.31 ppm). ¹H NMR (400 MHz, CDCl₃/CD₃OD) δ 8.23 (s, 1H), 8.19 - 8.15 (m, 1H), 7.40 - 7.33 (m, 1H), 7.26 - 7.15 (m, 3H), 5.07 (d, *J* = 6.9 Hz, 2H), 4.29 (dd, *J* = 14.4, 3.0 Hz, 1H), 4.11 (dd, *J* = 14.4, 6.7 Hz, 1H), 3.92 (s, 2H), 3.74 (ddd, *J* = 17.0, 10.3, 6.2 Hz, 2H), 3.42 (dd, *J* = 12.7, 10.3 Hz, 1H), 2.31 - 2.22 (m, 2H), 1.44 (dd, *J* = 14.2, 7.1 Hz, 2H), 1.31 - 1.13 (m, 26H), 1.09 (d, *J* = 6.3 Hz, 3H), 0.84 (t, *J* = 6.9 Hz, 3H). Carbon referenced to CD₃OD (49.00 ppm). ¹³C NMR (101 MHz, CD₃OD/CDCl₃) δ 155.5, 151.9, 150.0, 143.4, 136.9 (d, *J* = 6.5 Hz), 135.8, 131.2, 129.2, 128.2, 128.1, 118.7, 76.3 (d, *J* = 13.3 Hz), 64.96 (d, *J* = 5.3 Hz), 64.90 (d, *J* = 159.4 Hz), 48.4, 40.5, 39.0, 32.4, 30.13 (3), 30.09 (3), 30.03, 29.9, 29.8, 29.6, 29.4, 28.9, 23.1, 16.6, 14.4. ³¹P NMR (162 MHz, CDCl₃/CD₃OD) δ 16.3. HRMS (ESI) *m/z* calculated for C₃₃H₅₃O₄N₅PS₂ [M - H]⁻: 678.3282, found 678.3274. Anal. calculated for C₃₃H₅₉N₆O₅PS₂ (NH⁴⁺ monohydrate): C, 55.44; H, 8.32; N, 11.75. Found: C, 56.09; H, 8.10; N, 11.42.



6-(Dodecyldisulfanyl)hexyl hydrogen (((R)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.13)

Procedure B, X = 6-(dodecyldisulfanyl)hexan-1-ol, Y = H₂O, 19 % yield. Amorphous solid.

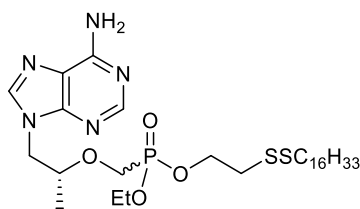
^1H NMR (400 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$ referenced to TMS) δ 8.31 (s, 1H), 8.24 (s, 1H), 4.39 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.20 (dd, $J = 14.4, 6.9$ Hz, 1H), 3.85 (ddd, $J = 9.9, 6.4, 3.1$ Hz, 1H), 3.83 - 3.71 (m, 3H), 3.45 (dd, $J = 12.6, 10.2$ Hz, 1H), 2.67 (dt, $J = 10.8, 3.8$ Hz, 4H), 1.72 - 1.61 (m, 4H), 1.56 (dt, $J = 13.5, 6.8$ Hz, 2H), 1.44 - 1.23 (m, 22H), 1.20 (d, $J = 6.2$ Hz, 3H), 0.89 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$ referenced to CD_3OD) δ 155.8, 152.5, 150.0, 143.3, 118.8, 76.3 (d, $J = 13.2$ Hz), 65.2 (d, $J = 6.0$ Hz), 64.6 (d, $J = 159.7$ Hz), 48.5, 39.3, 39.2, 34.9, 32.3, 31.3 (d, $J = 6.3$ Hz), 30.05, 30.03, 29.99, 29.92, 29.75, 29.65, 29.58, 29.47, 28.9, 28.6, 25.8, 23.1, 16.6, 14.3. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 16.4. HRMS (ESI) m/z calculated for $\text{C}_{27}\text{H}_{49}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} - \text{H}]^-$: 602.2969, found 602.2970. Anal. calculated for $\text{C}_{27}\text{H}_{53}\text{N}_6\text{O}_4\text{PS}_2$ (as an ammonium salt): C, 52.23; H, 8.60; N, 13.54. Found: C, 52.21; H, 8.65; N, 12.42.



2-(Hexadecyldisulfanyl)ethyl methyl (((1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.14)

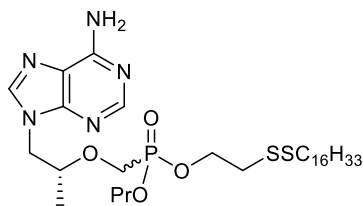
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = MeOH, 35.3 % yield. White solid. ^1H NMR (400 MHz, CDCl_3) δ 8.34 (s, 1H), 7.96 (s, 1H), 6.08 (s, 2H), 4.36 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.30 - 4.20 (m, 2H), 4.13 (dd, $J = 14.5, 7.8$ Hz, 1H), 3.96 - 3.84 (m, 2H), 3.71 (dd, $J = 23.5, 10.8$ Hz, 3H), 3.62 (dd, $J = 13.7, 9.0$ Hz, 1H), 2.86 (ddd, $J = 13.3, 12.4, 6.7$ Hz, 2H), 2.72 - 2.63 (m, 2H), 1.69 - 1.59 (m, 2H), 1.33 - 1.22 (m, 29H), 0.87 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.4, 152.7, 150.1, 141.8, 119.2, 76.6, 76.5, 76.4, 64.3, 64.24, 64.19, 64.13, 63.2, 61.5, 60.2, 53.0, 52.89, 52.85, 52.78, 48.2, 39.07, 39.05, 38.46, 38.42, 38.4, 38.37, 31.9, 29.67, 29.64, 29.58, 29.50, 29.34, 29.22, 29.1, 28.5, 28.48, 22.68, 16.56, 16.50, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 22.2, 22.1. HRMS (ESI) m/z calculated for $\text{C}_{28}\text{H}_{53}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 618.3271, found 618.3259. Anal. calculated for $\text{C}_{28}\text{H}_{52}\text{N}_5\text{O}_4\text{PS}_2$: C,

54.43; H, 8.48; N, 11.34. Found: C, 54.20; H, 8.56; N, 11.08. Melting Point: 48-50°C. d.r = 1:1 by ^{31}P NMR.



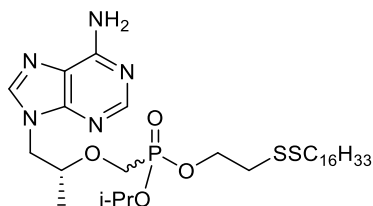
(2-(Hexadecyldisulfanyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.15)

Procedure B, X= 2-(hexadecyldisulfanyl)ethan-1-ol, Y = EtOH, 13 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.35 (s, 1H), 7.98 (s, 1H), 5.86 (s, 2H), 4.37 (dd, $J = 14.4, 2.9$ Hz, 1H), 4.26 (ddd, $J = 21.3, 10.7, 4.3$ Hz, 2H), 4.19 – 3.97 (m, 4H), 3.94 (dd, $J = 7.8, 3.2$ Hz, 1H), 3.88 (dd, $J = 13.6, 8.9$ Hz, 1H), 3.67 – 3.57 (m, 1H), 2.87 (dt, $J = 19.2, 6.7$ Hz, 2H), 2.68 (td, $J = 7.2, 5.7$ Hz, 2H), 1.65 (td, $J = 14.5, 7.2$ Hz, 2H), 1.49 – 1.14 (m, 31H), 0.88 (t, $J = 6.7$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.3, 152.8, 141.8, 76.4, 76.3, 64.14, 64.08, 64.04, 63.97, 63.5, 62.69, 62.62, 62.56, 61.8, 48.2, 39.07, 39.05, 38.46, 38.43, 38.40, 38.37, 31.9, 29.67, 29.65, 29.63, 29.58, 29.49, 29.34, 29.22, 29.10, 29.10, 28.48, 22.7, 16.53, 16.48, 16.41, 16.36, 14.1. ^{31}P NMR (162 MHz, CDCl_3) δ 21.91, 21.86. HRMS (ESI) m/z calculated for $\text{C}_{29}\text{H}_{55}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 632.3427, found 632.3417. d.r = 1:1 by ^{31}P NMR. Anal. calculated for $\text{C}_{29}\text{H}_{54}\text{N}_5\text{O}_4\text{PS}_2$: C, 55.12; H, 8.61; N, 11.08. Found: C, 55.48; H, 8.37; N, 10.44.



2-(Hexadecyldisulfanyl)ethyl propyl (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.16)

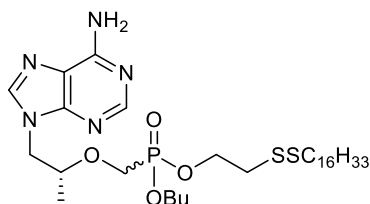
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = propanol, 41 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.33 (s, 1H), 7.96 (d, $J = 1.2$ Hz, 1H), 6.08 (s, 3H), 4.35 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.30 – 4.19 (m, 2H), 4.13 (ddd, $J = 14.4, 7.6, 2.2$ Hz, 1H), 4.03 – 3.83 (m, 4H), 3.62 (ddd, $J = 13.6, 9.6, 2.5$ Hz, 1H), 2.86 (dt, $J = 16.2, 6.7$ Hz, 2H), 2.71 – 2.59 (m, 2H), 1.73 – 1.56 (m, 4H), 1.44 – 1.04 (m, 28H), 0.89 (ddd, $J = 13.7, 11.7, 7.2$ Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.52, 152.87, 150.08, 141.73, 119.16, 76.44, 76.42, 76.32, 76.31, 68.09, 68.03, 67.97, 64.15, 64.09, 64.06, 64.00, 63.46, 61.79, 48.17, 48.14, 39.07, 39.05, 38.48, 38.45, 38.42, 38.39, 31.88, 29.66, 29.64, 29.62, 29.57, 29.48, 29.33, 29.21, 29.09, 28.48, 23.87, 23.83, 23.81, 23.77, 22.66, 16.52, 16.50, 14.11, 9.96, 9.92. ^{31}P NMR (121 MHz, CDCl_3) δ 21.92, 21.87. HRMS (ESI) m/z calculated for $\text{C}_{30}\text{H}_{57}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 646.3584, found 646.3575. Anal. calculated for $\text{C}_{30}\text{H}_{56}\text{N}_5\text{O}_4\text{PS}_2$: C, 55.79; H, 8.74; N, 10.84. Found: C, 54.97; H, 8.78; N, 10.07.



2-(Hexadecyldisulfanyl)ethyl isopropyl (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (SI-15)

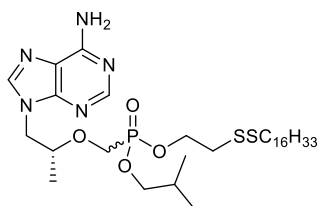
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = *i*PrOH, 33.9 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.29 (s, 1H), 7.94 (s, 1H), 6.17 (d, $J = 22.9$ Hz, 2H), 4.74 – 4.61 (m, 1H), 4.31 (dd, $J = 14.4, 2.3$ Hz, 1H), 4.25 – 4.02 (m, 3H), 3.94 – 3.73 (m, 2H), 3.55 (ddd, $J = 13.6, 9.6, 7.8$ Hz, 1H), 2.82 (dt, $J = 13.5, 6.8$ Hz, 2H), 2.69 – 2.54 (m, 2H), 1.59 (dt, $J = 11.5, 9.1$ Hz, 2H), 1.49 – 0.97 (m, 35H), 0.82 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 171.89, 155.54, 155.52, 152.78, 150.02, 141.72, 119.08, 76.36, 76.24, 71.66, 71.59, 63.98, 63.92, 63.86, 62.23, 62.16, 48.12, 48.08, 39.06, 39.04, 38.45, 38.38, 31.87, 29.64, 29.60, 29.55, 29.46, 29.31, 29.19, 29.07, 28.46, 25.34, 24.06, 24.02,

24.01, 23.97, 23.92, 23.88, 23.83, 22.64, 16.50, 16.48, 14.75, 14.09. ^{31}P NMR (162 MHz, CDCl_3) δ 20.08, 20.03. HRMS (ESI) m/z calculated for $\text{C}_{30}\text{H}_{57}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 646.3584, found 646.3583. d.r = 1.2:1 by ^{31}P NMR. Anal. calculated for $\text{C}_{30}\text{H}_{56}\text{N}_5\text{O}_4\text{PS}_2$: C, 55.79; H, 8.74; N, 10.84. Found: C, 53.79; H, 8.12; N, 10.38.



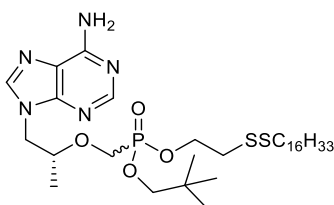
Butyl (2-(hexadecyldisulfanyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methylphosphonate (1.17)

Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = *n*-BuOH, 54 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.34 (s, 1H), 7.96 (d, $J = 1.4$ Hz, 1H), 5.99 (s, 2H), 4.35 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.29 – 4.19 (m, 2H), 4.13 (ddd, $J = 14.5, 7.6, 2.8$ Hz, 1H), 4.10 – 3.91 (m, 4H), 3.87 (dd, $J = 13.6, 9.0$ Hz, 1H), 3.62 (ddd, $J = 13.6, 9.5, 3.0$ Hz, 1H), 2.87 (dt, $J = 15.2, 6.7$ Hz, 2H), 2.71 – 2.61 (m, 2H), 1.68 – 1.54 (m, 4H), 1.36 – 1.17 (m, 30H), 0.96 – 0.81 (m, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.47, 152.85, 150.07, 150.05, 141.74, 119.12, 77.33, 77.02, 76.70, 76.44, 76.42, 76.31, 66.34, 64.14, 64.08, 64.01, 63.46, 61.78, 48.16, 48.13, 39.06, 39.05, 38.48, 38.45, 38.42, 38.39, 32.47, 32.43, 32.41, 32.38, 31.88, 29.65, 29.64, 29.62, 29.56, 29.48, 29.32, 29.21, 29.09, 28.47, 22.66, 18.63, 18.59, 16.52, 16.49, 14.10, 13.57, 13.56. ^{31}P NMR (121 MHz, CDCl_3) δ 21.14, 21.11. HRMS (ESI) m/z calculated for $\text{C}_{31}\text{H}_{59}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 660.3740, found 660.3748. d.r: 1:1 by ^{31}P NMR. Anal. calculated for $\text{C}_{31}\text{H}_{59}\text{N}_5\text{O}_4\text{PS}_2$: C, 56.42; H, 8.86; N, 10.61. Found: C, 55.63; H, 8.87; N, 10.06.



2-(Hexadecyldisulfanyl)ethyl isobutyl (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.18)

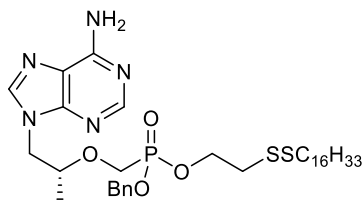
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = isobutanol, 7.0 % yield. Amorphous solid. ¹H NMR (400 MHz, CDCl₃) δ 8.35 (s, 1H), 7.97 (s, 1H), 5.74 (s, 2H), 4.37 (dd, *J* = 14.4, 3.0 Hz, 1H), 4.31 – 4.20 (m, 2H), 4.15 (dd, *J* = 14.4, 7.6 Hz, 1H), 3.98 – 3.74 (m, 4H), 3.64 (ddd, *J* = 13.6, 9.5, 1.4 Hz, 1H), 2.94 – 2.82 (m, 2H), 2.73 – 2.64 (m, 2H), 1.90 (dtd, *J* = 20.1, 13.4, 6.7 Hz, 1H), 1.69 – 1.58 (m, 2H), 1.39 – 1.21 (m, 28H), 0.98 – 0.81 (m, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 155.30, 152.87, 141.80, 119.14, 76.39, 76.37, 76.27, 76.25, 72.44, 72.38, 72.31, 64.19, 64.13, 64.09, 64.02, 63.38, 61.70, 48.21, 48.17, 39.09, 39.07, 38.49, 38.46, 38.44, 38.40, 31.90, 29.67, 29.63, 29.58, 29.50, 29.34, 29.22, 29.17, 29.15, 29.11, 28.49, 22.67, 18.62, 18.58, 16.51, 14.12. ³¹P NMR (121 MHz, CDCl₃) δ 21.06, 21.03. HRMS (ESI) *m/z* calculated for C₃₁H₅₈O₄N₅PS₂ [M + H]⁺ : 660.3740, found 660.3744. d.r: 1:1 by ³¹P NMR. Anal. calculated for C₃₁H₅₇N₅O₄PS₂: C, 56.42; H, 8.86; N, 10.61. Found: C, 56.65; H, 8.95; N, 10.25.



2-(Hexadecyldisulfanyl)ethyl neopentyl (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.19)

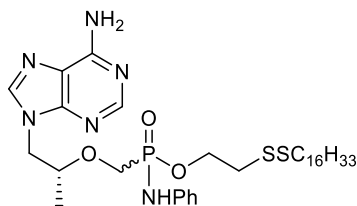
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = neopentanol, 44 % yield. Amorphous solid. ¹H NMR (400 MHz, CDCl₃) δ 8.31 (s, 1H), 7.98 (s, 1H), 6.33 (s, 2H), 4.36 (dd, *J* = 14.4, 2.8 Hz, 1H), 4.32 – 4.20 (m, 2H), 4.15 (dd, *J* = 14.5, 7.4 Hz, 1H), 3.95 (dd, *J* = 6.0, 4.6 Hz, 1H), 3.88 (dd, *J* = 13.6, 9.2 Hz, 1H), 3.75 – 3.55 (m, 3H), 2.87 (dt, *J* = 13.5, 6.7 Hz, 2H), 2.73 – 2.58 (m, 2H), 1.62 (dd, *J* = 10.7, 6.8 Hz, 2H), 1.36 – 1.11 (m, 29H), 0.94 – 0.75 (m, 12H). ¹³C NMR (101 MHz, CDCl₃) δ 155.43, 152.45, 149.94, 149.93, 141.74, 118.84, 118.82, 76.30, 76.24, 76.18, 76.13, 75.73, 75.71, 75.66, 75.64,

64.22, 64.15, 64.09, 63.29, 61.61, 48.20, 48.16, 39.05, 39.04, 38.48, 38.45, 38.42, 38.39, 32.08, 32.03, 32.02, 31.97, 31.88, 29.65, 29.64, 29.61, 29.56, 29.47, 29.32, 29.20, 29.08, 28.47, 25.94, 25.90, 22.65, 16.48, 16.45, 14.10. ^{31}P NMR (162 MHz, CDCl_3) δ 20.95, 20.93. HRMS (ESI) m/z calculated for $\text{C}_{32}\text{H}_{61}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 674.3897, found 674.3892. d.r: 1.2:1 by ^{31}P NMR. Anal. calculated for $\text{C}_{32}\text{H}_{60}\text{N}_5\text{O}_4\text{PS}_2$: C, 57.03; H, 8.97; N, 10.39. Found: C, 57.11; H, 8.92; N, 10.11.



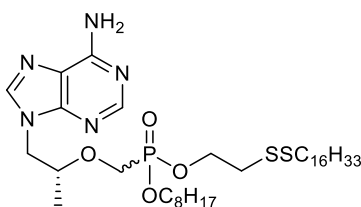
Benzyl (2-(hexadecyldisulfanyl)ethyl) (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.20)

Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = BnOH, 54 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.34 (s, 1H), 7.94 (s, 1H), 7.34 (ddd, $J = 16.8, 7.7, 6.2$ Hz, 5H), 6.31 (s, 2H), 5.15 – 5.01 (m, 2H), 4.33 – 4.07 (m, 4H), 3.95 – 3.78 (m, 2H), 3.65 – 3.50 (m, 1H), 2.82 (dt, $J = 13.4, 6.5$ Hz, 2H), 2.72 – 2.60 (m, 2H), 1.68 – 1.58 (m, 2H), 1.45 – 1.09 (m, 28H), 0.88 (t, $J = 6.5$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.64, 155.63, 152.87, 150.00, 141.68, 135.90, 135.84, 135.78, 128.58, 128.57, 128.07, 119.07, 76.45, 76.42, 76.33, 76.30, 68.07, 68.01, 64.13, 64.09, 64.07, 64.03, 63.69, 62.01, 48.07, 39.04, 38.34, 38.31, 38.28, 38.25, 31.88, 29.65, 29.61, 29.56, 29.48, 29.32, 29.20, 29.09, 29.06, 28.48, 28.46, 22.65, 16.53, 16.43, 14.11. ^{31}P NMR (121 MHz, CDCl_3) δ 21.65 (2). HRMS (ESI) m/z calculated for $\text{C}_{34}\text{H}_{56}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 694.3584, found 694.3578. Anal. calculated for $\text{C}_{34}\text{H}_{55}\text{N}_5\text{O}_4\text{PS}_2$: C, 58.85; H, 8.13; N, 10.09. Found: C, 58.50; H, 8.08; N, 9.69.



2-(Hexadecyldisulfanyl)ethyl P-(((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)-*N*-phenylphosphonamidate (1.21)

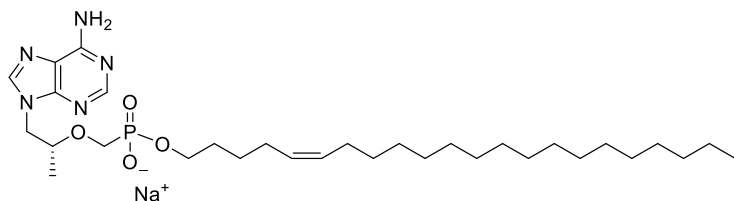
Procedure B, X = 2-(hexadecyldisulfanyl)ethan-1-ol, Y = aniline, 34 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.29 (s, 1H), 8.27 (s, 1H), 7.95 (s, 1H), 7.91 (s, 1H), 7.18 (dt, $J = 13.5, 7.8$ Hz, 2H), 6.93 (ddd, $J = 18.7, 15.8, 7.9$ Hz, 3H), 6.69 (s, 2H), 4.44 – 4.15 (m, 4H), 4.12 – 3.89 (m, 3H), 3.83 – 3.67 (m, 2H), 2.93 – 2.84 (m, 2H), 2.62 (t, $J = 7.4$ Hz, 2H), 1.64 – 1.56 (m, 2H), 1.34 – 1.21 (m, 22H), 1.19 (d, $J = 6.2$ Hz, 2H), 1.12 (d, $J = 6.2$ Hz, 2H), 0.88 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.57, 152.48, 152.45, 149.96, 149.81, 141.85, 139.44, 139.25, 129.45, 129.38, 122.23, 118.86, 118.80, 118.12, 118.06, 118.00, 117.94, 76.48, 76.36, 76.11, 76.00, 63.23, 63.21, 62.56, 62.50, 61.68, 61.66, 48.33, 48.22, 38.95, 38.28, 38.22, 31.89, 29.67, 29.65, 29.63, 29.57, 29.48, 29.33, 29.20, 29.07, 28.46, 22.67, 16.40, 16.22, 14.12. ^{31}P NMR (121 MHz, CDCl_3) δ 21.83, 21.76. HRMS (ESI) m/z calculated for $\text{C}_{33}\text{H}_{56}\text{O}_3\text{N}_6\text{PS}_2$ $[\text{M} + \text{H}]^+$: 679.3587, found 679.3591. d.r: 1.5:1 by ^{31}P NMR. Anal. calculated for $\text{C}_{33}\text{H}_{56}\text{N}_5\text{O}_4\text{PS}_2$: C, 58.38; H, 8.17; N, 12.38. Found: C, 57.69; H, 7.81; N, 11.58.



2-(Hexadecyldisulfanyl)ethyl octyl (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.22)

Procedure B, X = 2-(hexadecyldisulfanyl)ethanol, Y = 1-octanol, 60.3 % yield. Amorphous solid. ^1H NMR (400 MHz, CDCl_3) δ 8.35 (s, 1H), 7.97 (s, 1H), 5.85 (s, 2H), 4.36 (dd, $J = 14.4, 3.0$ Hz, 1H), 4.30 – 4.19 (m, 2H), 4.14 (ddd, $J = 14.4, 7.6, 2.4$ Hz, 1H), 4.08 – 3.83 (m, 4H), 3.62 (ddd, $J = 13.5, 9.5, 3.6$ Hz, 1H), 2.94 – 2.81 (m, 2H), 2.73 – 2.63 (m, 2H), 1.72 – 1.51 (m, 4H), 1.41 – 1.10 (m, 39H), 0.93 – 0.71 (m, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 155.40, 155.39, 152.88, 150.13, 150.11, 141.79, 119.21, 119.20, 76.44, 76.33, 66.71, 66.65, 66.58, 64.15, 64.09, 64.06, 63.99, 63.50, 61.82, 48.17, 48.13, 39.08,

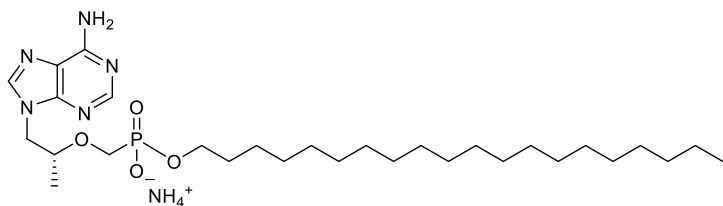
39.07, 38.49, 38.45, 38.43, 38.40, 31.89, 31.74, 30.53, 30.49, 30.47, 30.43, 29.67, 29.63, 29.58, 29.50, 29.34, 29.23, 29.16, 29.15, 29.10, 29.09, 29.07, 28.49, 25.41, 25.37, 22.67, 22.61, 16.54, 16.52, 14.11, 14.08. ^{31}P NMR (162 MHz, CDCl_3) δ 21.09, 21.07.d.r: 1.2:1. HRMS (ESI) m/z calculated for $\text{C}_{35}\text{H}_{67}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 716.4366, found 716.4355. Anal. calculated for $\text{C}_{35}\text{H}_{66}\text{O}_4\text{N}_5\text{PS}_2$: C, 58.71; H, 9.29; N, 9.78. Found: C, 58.53; H, 9.35; N, 9.52.



(Z)-docos-5-en-1-yl hydrogen (((R)-1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.23)

Procedure C: L = (Z)-docos-5-en-1-ol. Purification gradient: 0-50% DCM:MeOH: NH_4OH , 5% yield.

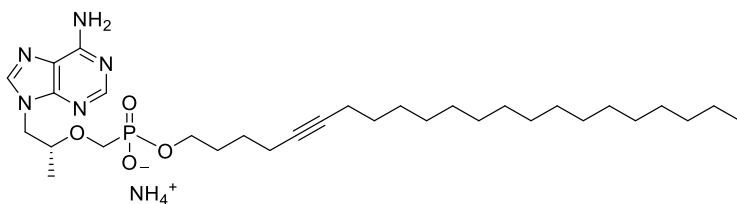
^1H NMR (399 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 8.23 (s, 1H), 8.13 (s, 1H), 6.99 (s, 2H), 5.43 - 5.31 (m, 1H), 4.87 - 4.78 (m, 2H), 4.29 (dd, $J = 14.2, 2.6$ Hz, 1H), 4.02 (dd, $J = 14.4, 7.6$ Hz, 1H), 3.81 - 3.63 (m, 3H), 1.85 - 1.75 (m, 1H), 1.55 - 1.36 (m, 1H), 1.36 - 1.01 (m, 38H), 0.79 (t, $J = 6.2$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 149.7, 149.1, 147.7, 143.1, 142.7, 122.3, 114.1, 75.4 (d, $J = 12.1$ Hz), 64.9 (d, $J = 6.6$ Hz), 63.8 (d, $J = 157.2$ Hz), 43.8, 34.8, 33.8, 31.7, 30.8, 29.5, 29.4, 29.38, 29.13, 29.09, 28.5 (d, $J = 6.3$ Hz), 26.9, 24.4, 23.4, 22.4, 16.1, 13.8. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 16.2. HRMS (ESI) m/z calculated for $\text{C}_{31}\text{H}_{55}\text{O}_4\text{N}_5\text{P}$ $[\text{M} - \text{H}]^-$: 592.3997, found 592.3998. Anal. calculated for $\text{C}_{32}\text{H}_{61}\text{N}_5\text{NaO}_6\text{P}$ and $\text{C}_{32}\text{H}_{64}\text{N}_5\text{O}_7\text{P}$ (a 1:1 mixture of the sodium salt and free acid solvated with $2\text{H}_2\text{O}$ and MeOH): C, 57.90; H, 9.49; N, 10.55. Found: C, 57.99; H, 9.55; N, 10.26.



Icosyl hydrogen ((((*R*)-1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate

(1.24)

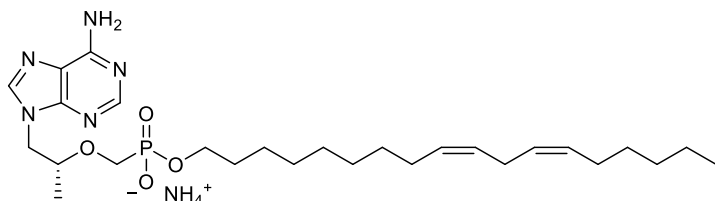
Procedure C: L = icosan-1-ol. Purification gradient: 0-58% DCM:MeOH:NH₄OH, 20% yield. Waxy solid. ¹H NMR (400 MHz, CDCl₃/CD₃OD) δ 8.25 (s, 1H), 8.17 (s, 1H), 4.32 (dd, *J* = 14.4, 2.7 Hz, 1H), 4.06 (dd, *J* = 14.4, 7.5 Hz, 1H), 3.81 - 3.65 (m, 4H), 3.33 (dd, *J* = 12.9, 10.3 Hz, 2H), 1.52 - 1.43 (m, 2H), 1.24 - 1.11 (m, 39H), 0.82 (t, *J* = 6.8 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃/CD₃OD) δ 154.4, 151.0, 149.3, 143.1, 118.0, 75.7 (d, *J* = 12.7 Hz), 65.1 (d, *J* = 5.5 Hz), 64.0 (d, *J* = 158.8 Hz), 48.0, 30.88 (d, *J* = 6.1 Hz), 30.9, 29.67, 29.60, 29.36, 29.30, 25.7, 22.6, 16.4, 14.0. ³¹P NMR (162 MHz, CDCl₃/CD₃OD) δ 16.1. HRMS (ESI) *m/z* calculated for C₂₉H₅₄N₅O₄P [M - H]⁻: 566.3840, found 566.3833. Anal. calculated for C₂₉H₅₇N₆O₄P and C₂₉H₅₆N₅O₅P (as a 1:1 mixture of hydrated free acid and the ammonium salt): C, 59.51; H, 9.73; N, 13.16. Found: C, 59.51; H, 9.70; N, 12.86.



Docos-5-yn-1-yl hydrogen ((((*R*)-1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.25)

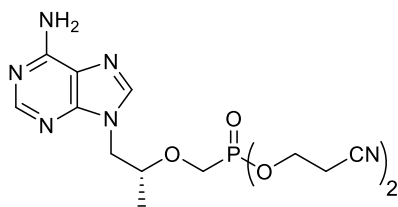
Procedure C: L = docos-5-yn-1-ol. Purification gradient: 0-43% DCM:MeOH:NH₄OH, 27% yield. ¹H NMR (399 MHz, CDCl₃/CD₃OD) δ 8.28 (s, 1H), 8.08 (s, 1H), 6.99 (s, 2H), 4.30 (dt, *J* = 13.6, 3.9 Hz, 1H), 4.01 (dd, *J* = 14.4, 8.0 Hz, 1H), 3.91 - 3.82 (m, 1H), 3.78 (q, *J* = 6.5 Hz, 1H), 3.72 (dd, *J* = 12.9, 8.8 Hz, 2H), 3.38 (dd, *J* = 12.9, 9.3 Hz, 1H), 2.25 - 2.16 (m, 1H), 1.97 (d, *J* = 2.3 Hz, 1H), 1.77 - 1.66 (m, 1H), 1.64 - 1.50 (m, 1H), 1.40 - 1.12 (m, 35H), 0.80 (t, *J* = 6.9 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃/CD₃OD) δ 153.2, 149.1, 147.9, 143.8, 118.0, 87.6, 75.4 (d, *J* = 11.1 Hz), 73.5, 69.6, 64.8 (d, *J* = 6.1 Hz), 63.9 (d, *J* = 159.9 Hz), 35.0, 31.9, 31.3, 31.1, 29.7, 29.7, 29.6, 29.5, 29.4, 28.8 (d, *J* = 6.2 Hz), 27.3, 24.7, 23.8, 22.7, 16.5, 14.1. ³¹P NMR (162 MHz, CDCl₃/CD₃OD) δ 16.6. HRMS (ESI) *m/z*

calculated for $C_{31}H_{53}O_4N_5P$ $[M - H]^-$: 590.3840, found 590.3834. Anal. calculated for $C_{31}H_{55}N_5O_5PNa$ and $C_{31}H_{57}N_6O_4P$ (as a 1:1 mixture of ammonium and sodium salts): C, 60.03; H, 9.10; N, 12.42. Found: C, 60.15; H, 9.20; N, 12.32.



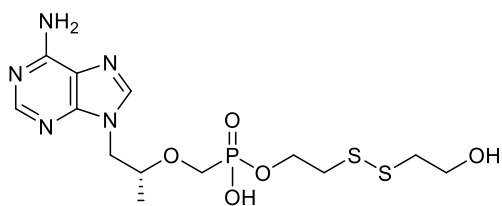
(9Z,12Z)-Octadeca-9,12-dien-1-yl hydrogen (((R)-1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (1.26)

Procedure C: L = (9Z,12Z)-octadeca-9,12-dien-1-ol. Purification gradient: 0-47% DCM:MeOH:NH₄OH, 50% yield. Waxy solid. ¹H NMR (400 MHz, CDCl₃/CD₃OD) δ 8.22 (s, 1H), 8.15 (s, 1H), 5.34 – 5.19 (m, 4H), 4.29 (dd, *J* = 14.3, 2.7 Hz, 1H), 4.04 (dd, *J* = 14.4, 7.5 Hz, 2H), 3.93 (bs, 4H), 3.77 – 3.64 (m, 4H), 3.34 – 3.26 (m, 1H), 2.68 (t, *J* = 6.6 Hz, 2H), 1.99 – 1.90 (m, 4H), 1.44 (dd, *J* = 13.6, 6.7 Hz, 2H), 1.31 – 1.14 (m, 18H), 1.12 (d, *J* = 6.2 Hz, 3H), 0.80 (t, *J* = 6.9 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃/CD₃OD) δ 154.6, 151.1, 149.3, 142.9, 130.1, 129.9, 127.9, 127.7, 118.1, 75.7 (d, *J* = 13.0 Hz), 65.0 (d, *J* = 5.8 Hz), 64.0 (d, *J* = 160.2 Hz), 31.4, 30.8 (d, *J* = 6.3 Hz), 29.6, 29.42, 29.24, 29.21, 29.18, 27.09, 27.04, 25.6, 25.5, 22.4, 16.3, 13.9. ³¹P NMR (162 MHz, CDCl₃/CD₃OD) δ 16.2. HRMS (ESI) *m/z* calculated for $C_{27}H_{47}O_4N_5P$ $[M + H]^+$: 536.3360, found 536.3356. Anal. calculated for $C_{27}H_{49}N_6O_6P$ (as an ammonium salt dihydrate): C, 55.46; H, 8.45; N, 14.37. Found: C, 55.61; H, 8.40; N, 14.06.

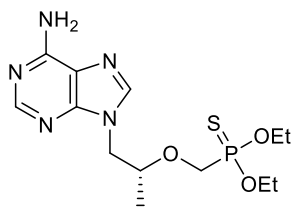


(*R*)-bis(2-Cyanoethyl) (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate**(SI-16)**

Procedure B, X = Y = 2-cyanoethanol, 56 % yield. Hygroscopic white solid. ¹H NMR (400 MHz, CDCl₃) δ 8.21 (d, *J* = 0.7 Hz, 1H), 7.89 (s, 1H), 6.83 (s, 2H), 4.31 – 4.00 (m, 5H), 3.93 – 3.80 (m, 2H), 3.68 – 3.53 (m, 2H), 2.74 – 2.57 (m, 4H), 1.16 (d, *J* = 6.2 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 155.8, 152.8, 149.8, 141.6, 118.8, 116.83, 116.81, 76.6 (d, *J* = 12.9 Hz), 62.3 (d, *J* = 168.8 Hz), 60.95 (d, *J* = 6.3 Hz), 60.94 (d, *J* = 5.9 Hz), 48.1, 19.8 (d, *J* = 4.8 Hz), 19.8 (d, *J* = 4.5 Hz), 16.4. ³¹P NMR (162 MHz, CDCl₃) δ 22.1. HRMS (ESI) *m/z* calculated for C₁₅H₂₁O₄N₇P [M + H]⁺: 394.138, found 394.138.

**2-((2-Hydroxyethyl)disulfanyl)ethyl hydrogen (((*R*)-1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (SI-17)**

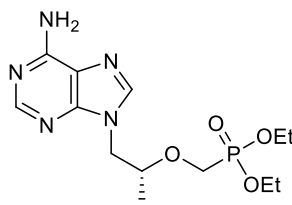
Procedure B, X = 2,2'-disulfanediybis(ethan-1-ol), Y = H₂O, 15 % yield. Amorphous solid. Purified on a C18 reverse phase column using a H₂O/MeOH gradient (10-20% MeOH). ¹H NMR (400 MHz, D₂O) δ 8.39 (s, 1H), 8.38 (s, 1H), 8.19 (s, 2H), 4.44 (dd, *J* = 14.7, 2.9 Hz, 1H), 4.26 (dd, *J* = 14.7, 7.8 Hz, 1H), 4.04 – 3.90 (m, 1H), 3.89 – 3.78 (m, 2H), 3.73 (dt, *J* = 11.7, 7.5 Hz, 3H), 3.47 (dd, *J* = 13.3, 9.8 Hz, 1H), 2.80 – 2.65 (m, 4H), 1.20 (d, *J* = 6.3 Hz, 3H). ¹³C NMR (101 MHz, D₂O) δ 165.7, 149.7, 148.5, 145.5, 144.4, 117.6, 75.5 (d, *J* = 12.7 Hz), 62.8 (d, *J* = 159.4 Hz), 62.5 (d, *J* = 5.5 Hz), 59.0, 48.5, 39.8, 38.5 (d, *J* = 6.2 Hz), 15.5. ³¹P NMR (D₂O) δ: 17.9. HRMS (ESI) *m/z* calculated for C₁₃H₂₃O₅N₅PS₂ [M]⁺: 424.087, found 424.0877. Anal. calculated for C₁₃H₂₂N₅O₅PS₂: C, 36.87; H, 5.24; N, 16.54. Found: C, 36.58; H, 5.24; N, 16.43.



(R)-O,O-diethyl (((1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonothioate

(SI-18)

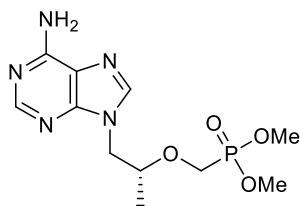
To a stirring solution of (*R*)-diethyl (((1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (0.100 g, 0.291 mmol) in anhydrous dioxane was added Davy Reagent (0.108 g, 0.379 mmol) at room temperature. The solution was then heated to 80°C and stirred for 2 h. Additional Davy reagent was added to encourage completion. Then, the solvent was evaporated under reduced pressure and the resulting residue was suspended in DCM and purified on a silica column using a DCM/MeOH (0-7%) gradient to afford the title compound (*R*)-O,O-diethyl (((1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonothioate (47.6 mg, 0.132 mmol, 45.5 % yield) as a white solid. ¹H NMR (400 MHz, CDCl₃) δ 8.41 (s, 1H), 8.35 (s, 1H), 4.40 (d, *J* = 13.9 Hz, 1H), 4.18 (d, *J* = 15.3 Hz, 1H), 4.07 – 3.83 (m, 6H), 3.71 – 3.65 (m, 1H), 1.23 – 1.12 (m, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 149.7, 149.3, 145.3, 144.3, 118.0, 75.8 (d, *J* = 9.4 Hz), 69.0 (d, *J* = 136.0 Hz), 63.0 (d, *J* = 6.9 Hz), 62.8 (d, *J* = 6.9 Hz), 16.7 (d, *J* = 3.1 Hz), 16.3 (d, *J* = 3.0 Hz). ³¹P NMR (CDCl₃) δ: 87.1. HRMS (ESI) *m/z* calculated for C₁₃H₂₃O₃N₅PS [M + H]⁺ : 360.1253, found 360.1247.



(R)-Diethyl (((1-(6-amino-9H-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (SI-19)

To a stirring solution of dry tenofovir (1.00 g, 3.48 mmol) in anhydrous DCM (50 mL) and *N,N*-dimethylformamide (0.268 mL, 3.48 mmol) C was gradually added excess oxalyl dichloride (0.896 mL,

10.45 mmol). The reaction mixture stirred for 30 min at this temperature and then poured into a stirring solution of sodium ethoxide/ethanol at 0°C. The cloudy white solution stirred at room temperature for 15 min, then neutralized with aqueous acid to pH 5. The mixture was heated to 50°C and stirred until complete cleavage of the formamide was observed by LC-MS (H₂O/MeOH gradient, 50-95% MeOH, 3 min). The solvents were evaporated under reduced pressure and the residue was purified on a silica column using a DCM/DCM:MeOH:NH₄OH (80:20:3) gradient to afford the title compound (*R*)-diethyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (0.850 g, 2.47 mmol, 71 % yield) as an off-white solid. ¹H NMR (400 MHz, CDCl₃) δ 8.26 (s, 1H), 7.91 (s, 1H), 6.57 (s, 2H), 4.29 (dd, *J* = 14.4, 3.0 Hz, 1H), 4.14 - 3.67 (m, 7H), 3.60 - 3.46 (m, 1H), 1.28 - 1.10 (m, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 155.8, 152.8, 149.9, 141.6, 119.0, 76.3 (d, *J* = 12.0 Hz), 62.7 (d, *J* = 168.5 Hz), 62.3, 16.5, 16.4 (d, *J* = 4.4 Hz), 16.3 (d, *J* = 4.3 Hz). ³¹P NMR (CDCl₃) δ 20.8. HRMS (ESI) *m/z* calculated for C₁₃H₂₃O₄N₅P [M + H]⁺ : 344.1482, found 344.1483.



(*R*)-Dimethyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (SI-20)

To a stirring solution of dry tenofovir (1.00 g, 3.48 mmol) in anhydrous DCM (50 mL) and *N,N*-dimethylformamide (0.268 mL, 3.48 mmol) was gradually added excess oxalyl dichloride (0.896 mL, 10.45 mmol). After stirring for one hour at room temperature, the reaction mixture was poured into a stirring solution of sodium methoxide/Methanol at 0°C. The cloudy white solution stirred at room temperature for 15 min, then neutralized with aqueous HCl to pH 5. The solution was then heated to 50°C for 8 h. Progress monitored by LC-MS. Upon completion, the solvents were evaporated under reduced pressure and the residue was purified on a silica column using a DCM/DCM:MeOH:NH₄OH

(80:20:3) gradient (0-33%) to afford the title compound (*R*)-dimethyl (((1-(6-amino-9*H*-purin-9-yl)propan-2-yl)oxy)methyl)phosphonate (0.990 g, 3.14 mmol, 90 % yield) as a white powder. ¹H NMR (400 MHz, CDCl₃) δ 8.29 (d, *J* = 0.5 Hz, 1H), 7.92 (s, 1H), 6.48 - 6.43 (s, 2H), 4.32 (dd, *J* = 14.4, 2.9 Hz, 1H), 4.08 (dd, *J* = 14.5, 7.9 Hz, 1H), 3.94 - 3.78 (m, 2H), 3.73 - 3.49 (m, 7H), 1.21 (dd, *J* = 6.2, 0.5 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 155.7, 152.8, 150.0, 141.6, 119.0, 76.5 (d, *J* = 11.8 Hz), 62.0 (d, *J* = 168.4 Hz), 53.0 (d, *J* = 6.6 Hz), 52.8 (d, *J* = 6.6 Hz), 48.2, 16.5. ³¹P NMR (CDCl₃) δ 20.7. HRMS (ESI) *m/z* calculated for C₁₁H₁₉O₄N₅P [M + H]⁺: 316.1169, found 316.1160.

1.12 References

- (1) Van Rompay, A. R.; Johansson, M.; Karlsson, A., Phosphorylation of Nucleosides and Nucleoside Analogs by Mammalian Nucleoside Monophosphate Kinases. *Pharmacol. Ther.* 2000, 87 (2), 189-198.

- (2) Loeb, L. A.; Springgate, C. F.; Battula, N., Errors in DNA Replication as a Basis of Malignant Changes. *Cancer Res.* 1974, 34 (9), 2311-2321.
- (3) Kunkel, T. A., DNA Replication Fidelity. *J. Biol. Chem.* 2004, 279 (17), 16895-16898.
- (4) Echols, H.; Goodman, M. F., Fidelity Mechanisms in DNA Replication. *Annu. Rev. Biochem.* 1991, 60 (1), 477-511.
- (5) Zhu, Y.; Trego, K. S.; Song, L., et al., 3' to 5' Exonuclease Activity of Herpes Simplex Virus Type 1 DNA Polymerase Modulates Its Strand Displacement Activity. *J. Virol.* 2003, 77 (18), 10147-10153.
- (6) Preston, B. D.; Poiesz, B. J.; Loeb, L. A., Fidelity of HIV-1 Reverse Transcriptase. *Science* 1988, 242 (4882), 1168-1171.
- (7) Lauring, A. S.; Andino, R., Quasispecies Theory and the Behavior of Rna Viruses. *PLoS Pathog.* 2010, 6 (7), e1001005.
- (8) Bull, J. J.; Sanjuán, R.; Wilke, C. O., Theory of Lethal Mutagenesis for Viruses. *J. Virol.* 2007, 81 (6), 2930-2939.
- (9) Smith, R. A.; Loeb, L. A.; Preston, B. D., Lethal Mutagenesis of HIV. *Virus Res.* 2005, 107 (2), 215-28.
- (10) Svarovskaia, E. S.; Cheslock, S. R.; Zhang, W. H., et al., Retroviral Mutation Rates and Reverse Transcriptase Fidelity. *Front. Biosci.* 2003, 8, d117-34.
- (11) Drake, J. W., The Distribution of Rates of Spontaneous Mutation over Viruses, Prokaryotes, and Eukaryotes. *Ann. N. Y. Acad. Sci.* 1999, 870, 100-7.
- (12) Drake, J. W.; Charlesworth, B.; Charlesworth, D., et al., Rates of Spontaneous Mutation. *Genetics* 1998, 148 (4), 1667-86.

- (13) Gallant , J. E.; DeJesus , E.; Arribas , J. R., et al., Tenofovir Df, Emtricitabine, and Efavirenz Vs. Zidovudine, Lamivudine, and Efavirenz for HIV. *N. Engl. J. Med.* 2006, 354 (3), 251-260.
- (14) Marcellin , P.; Chang , T.-T.; Lim , S. G., et al., Adefovir Dipivoxil for the Treatment of Hepatitis B E Antigen–Positive Chronic Hepatitis B. *N. Engl. J. Med.* 2003, 348 (9), 808-816.
- (15) Lawitz , E.; Mangia , A.; Wyles , D., et al., Sofosbuvir for Previously Untreated Chronic Hepatitis C Infection. *N. Engl. J. Med.* 2013, 368 (20), 1878-1887.
- (16) Gupta, R.; Wald, A.; Krantz, E., et al., Valacyclovir and Acyclovir for Suppression of Shedding of Herpes Simplex Virus in the Genital Tract. *J. Infect. Dis.* 2004, 190 (8), 1374-1381.
- (17) Paya, C.; Humar, A.; Dominguez, E., et al., Efficacy and Safety of Valganciclovir Vs. Oral Ganciclovir for Prevention of Cytomegalovirus Disease in Solid Organ Transplant Recipients. *Am. J. Transplantation* 2004, 4 (4), 611-620.
- (18) Beadle, J. R.; Hartline, C.; Aldern, K. A., et al., Alkoxyalkyl Esters of Cidofovir and Cyclic Cidofovir Exhibit Multiple-Log Enhancement of Antiviral Activity against Cytomegalovirus and Herpesvirus Replication in Vitro. *Antimicrob. Agents Chemother.* 2002, 46 (8), 2381-2386.
- (19) Grant, R. M.; Lama, J. R.; Anderson, P. L., et al., Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *N. Engl. J. Med.* 2010, 363 (27), 2587-2599.
- (20) Jacobson , I. M.; Gordon , S. C.; Kowdley , K. V., et al., Sofosbuvir for Hepatitis C Genotype 2 or 3 in Patients without Treatment Options. *N. Engl. J. Med.* 2013, 368 (20), 1867-1877.
- (21) Lawitz, E.; Mangia, A.; Wyles, D., et al., Sofosbuvir for Previously Untreated Chronic Hepatitis C Infection. *N. Engl. J. Med.* 2013, 368 (20), 1878-1887.
- (22) Corey, L.; Wald, A.; Patel, R., et al., Once-Daily Valacyclovir to Reduce the Risk of Transmission of Genital Herpes. *N. Engl. J. Med.* 2004, 350 (1), 11-20.

- (23) Beutner, K. R.; Friedman, D. J.; Forszpaniak, C., et al., Valaciclovir Compared with Acyclovir for Improved Therapy for Herpes Zoster in Immunocompetent Adults. *Antimicrob. Agents Chemother.* 1995, 39 (7), 1546-1553.
- (24) Siegel, D.; Hui, H. C.; Doerffler, E., et al., Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1-F][Triazin-4-Amino] Adenine C-Nucleoside (Gs-5734) for the Treatment of Ebola and Emerging Viruses. *J. Med. Chem.* 2017, 60 (5), 1648-1661.
- (25) Warren, T. K.; Jordan, R.; Lo, M. K., et al., Therapeutic Efficacy of the Small Molecule Gs-5734 against Ebola Virus in Rhesus Monkeys. *Nature* 2016, 531 (7594), 381-385.
- (26) Deville-Bonne, D.; El Amri, C.; Meyer, P., et al., Human and Viral Nucleoside/Nucleotide Kinases Involved in Antiviral Drug Activation: Structural and Catalytic Properties. *Antiviral Res.* 2010, 86 (1), 101-120.
- (27) Johansson, N. G.; Eriksson, S., Structure-Activity Relationships for Phosphorylation of Nucleoside Analogs to Monophosphates by Nucleoside Kinases. *Acta Biochim. Pol.* 1996, 43 (1), 143-60.
- (28) Munch-Petersen, B.; Cloos, L.; Tyrsted, G., et al., Diverging Substrate Specificity of Pure Human Thymidine Kinases 1 and 2 against Antiviral Dideoxynucleosides. *J. Biol. Chem.* 1991, 266 (14), 9032-8.
- (29) Wang, J.; Choudhury, D.; Chattopadhyaya, J., et al., Stereoisomeric Selectivity of Human Deoxyribonucleoside Kinases. *Biochemistry* 1999, 38 (51), 16993-16999.
- (30) Spadari, S.; Maga, G.; Focher, F., et al., L-Thymidine Is Phosphorylated by Herpes Simplex Virus Type 1 Thymidine Kinase and Inhibits Viral Growth. *J. Med. Chem.* 1992, 35 (22), 4214-20.
- (31) Fyfe, J. A., Differential Phosphorylation of (E)-5-(2-Bromovinyl)-2'-Deoxyuridine Monophosphate by Thymidylate Kinases from Herpes Simplex Viruses Types 1 and 2 and Varicella Zoster Virus. *Mol. Pharmacol.* 1982, 21 (2), 432-7.

- (32) Burrel, S.; Bonnafeous, P.; Hubacek, P., et al., Impact of Novel Mutations of Herpes Simplex Virus 1 and 2 Thymidine Kinases on Acyclovir Phosphorylation Activity. *Antiviral Res.* 2012, 96 (3), 386-390.
- (33) Furman, P. A.; Fyfe, J. A.; St Clair, M. H., et al., Phosphorylation of 3'-Azido-3'-Deoxythymidine and Selective Interaction of the 5'-Triphosphate with Human Immunodeficiency Virus Reverse Transcriptase. *Proc. Natl. Acad. Sci. U. S. A.* 1986, 83 (21), 8333-8337.
- (34) Balzarini, J.; Herdewijn, P.; De Clercq, E., Differential Patterns of Intracellular Metabolism of 2',3'-Didehydro-2',3'-Dideoxythymidine and 3'-Azido-2',3'-Dideoxythymidine, Two Potent Anti-Human Immunodeficiency Virus Compounds. *J. Biol. Chem.* 1989, 264 (11), 6127-33.
- (35) Lavie, A.; Schlichting, I.; Vetter, I. R., et al., The Bottleneck in AZT Activation. *Nat. Med.* 1997, 3 (8), 922-4.
- (36) Schinazi, R. F.; Goudgaon, N. M.; Fulcrand, G., et al., Cellular Pharmacology and Biological Activity of 5-Carboranyl-2'-Deoxyuridine. *Int. J. Radiation Oncology Biol. Phys.* 28 (5), 1113-1120.
- (37) Vilpo, J. A.; Vilpo, L. M., Nucleoside Monophosphate Kinase May Be the Key Enzyme Preventing Salvage of DNA 5-Methylcytosine. *Mutat. Res.* 1993, 286 (2), 217-20.
- (38) Gallois-Montbrun, S.; Chen, Y.; Dutartre, H., et al., Structural Analysis of the Activation of Ribavirin Analogs by NDP Kinase: Comparison with Other Ribavirin Targets. *Mol. Pharmacol.* 2003, 63 (3), 538-46.
- (39) Persson, T.; Hörnfeldt, A.-B.; Gronowitz, S., et al., Thienyl-Substituted Nucleosides and Their Triphosphates. *Antiviral Chem. Chemother.* 1994, 5 (6), 395-402.
- (40) Gollnest, T.; Dinis de Oliveira, T.; Rath, A., et al., Membrane-Permeable Triphosphate Prodrugs of Nucleoside Analogues. *Angew. Chem. Int. Ed. Engl.* 2016, 55 (17), 5255-8.
- (41) Rinaldo-Matthis, A.; Rampazzo, C.; Reichard, P., et al., Crystal Structure of a Human Mitochondrial Deoxyribonucleotidase. *Nat. Struct. Mol. Biol.* 2002, 9 (10), 779-787.

- (42) Robbins, B. L.; Greenhaw, J.; Connelly, M. C., et al., Metabolic Pathways for Activation of the Antiviral Agent 9-(2-Phosphonylmethoxyethyl)Adenine in Human Lymphoid Cells. *Antimicrob. Agents Chemother.* 1995, 39 (10), 2304-8.
- (43) Gallant, J. E.; DeJesus, E.; Arribas, J. R., et al., Tenofovir Df, Emtricitabine, and Efavirenz Vs. Zidovudine, Lamivudine, and Efavirenz for HIV. *N. Engl. J. Med.* 2006, 354 (3), 251-60.
- (44) Fung, H. B.; Stone, E. A.; Piacenti, F. J., Tenofovir Disoproxil Fumarate: A Nucleotide Reverse Transcriptase Inhibitor for the Treatment of HIV Infection. *Clin. Ther.* 2002, 24 (10), 1515-48.
- (45) De Clercq, E., Clinical Potential of the Acyclic Nucleoside Phosphonates Cidofovir, Adefovir, and Tenofovir in Treatment of DNA Virus and Retrovirus Infections. *Clin. Microbiol. Rev.* 2003, 16 (4), 569-96.
- (46) Ying, C.; De Clercq, E.; Nicholson, W., et al., Inhibition of the Replication of the DNA Polymerase M550v Mutation Variant of Human Hepatitis B Virus by Adefovir, Tenofovir, L-Fmau, Dapd, Penciclovir and Lobucavir. *J. Viral Hepat.* 2000, 7 (2), 161-5.
- (47) Delaney, W. E. t.; Ray, A. S.; Yang, H., et al., Intracellular Metabolism and in Vitro Activity of Tenofovir against Hepatitis B Virus. *Antimicrob. Agents Chemother.* 2006, 50 (7), 2471-7.
- (48) Andrei, G.; Lisco, A.; Vanpouille, C., et al., Topical Tenofovir, a Microbicide Effective against HIV, Inhibits Herpes Simplex Virus-2 Replication. *Cell Host Microbe* 2011, 10 (4), 379-89.
- (49) Kearney, B. P.; Flaherty, J. F.; Shah, J., Tenofovir Disoproxil Fumarate - Clinical Pharmacology and Pharmacokinetics. *Clin. Pharmacokinet.* 2004, 43 (9), 595-612.
- (50) Balzarini, J.; Naesens, L.; Slachmuylders, J., et al., 9-(2-Phosphonylmethoxyethyl)Adenine (Pmea) Effectively Inhibits Retrovirus Replication in Vitro and Simian Immunodeficiency Virus Infection in Rhesus Monkeys. *AIDS* 1991, 5 (1), 21-8.

- (51) Starrett, J. E., Jr.; Tortolani, D. R.; Russell, J., et al., Synthesis, Oral Bioavailability Determination, and in Vitro Evaluation of Prodrugs of the Antiviral Agent 9-[2-(Phosphonomethoxy)Ethyl]Adenine (Pmea). *J. Med. Chem.* 1994, 37 (12), 1857-1864.
- (52) Pradere, U.; Garnier-Amblard, E. C.; Coats, S. J., et al., Synthesis of Nucleoside Phosphate and Phosphonate Prodrugs. *Chem. Rev.* 2014, 114 (18), 9154-218.
- (53) Ray, A. S.; Hostetler, K. Y., Application of Kinase Bypass Strategies to Nucleoside Antivirals. *Antiviral Res.* 2011, 92 (2), 277-91.
- (54) Hecker, S. J.; Erion, M. D., Prodrugs of Phosphates and Phosphonates. *J. Med. Chem.* 2008, 51 (8), 2328-45.
- (55) Karras, A.; Lafaurie, M.; Furco, A., et al., Tenofovir-Related Nephrotoxicity in Human Immunodeficiency Virus-Infected Patients: Three Cases of Renal Failure, Fanconi Syndrome, and Nephrogenic Diabetes Insipidus. *Clin. Infect. Dis.* 2003, 36 (8), 1070-3.
- (56) Coca, S.; Perazella, M. A., Rapid Communication: Acute Renal Failure Associated with Tenofovir: Evidence of Drug-Induced Nephrotoxicity. *Am. J. Med. Sci.* 2002, 324 (6), 342-4.
- (57) Lewis, W.; Day, B. J.; Copeland, W. C., Mitochondrial Toxicity of Nrti Antiviral Drugs: An Integrated Cellular Perspective. *Nat. Rev. Drug Discovery* 2003, 2 (10), 812-22.
- (58) Herlitz, L. C.; Mohan, S.; Stokes, M. B., et al., Tenofovir Nephrotoxicity: Acute Tubular Necrosis with Distinctive Clinical, Pathological, and Mitochondrial Abnormalities. *Kidney Int.* 2010, 78 (11), 1171-7.
- (59) Ruane, P. J.; DeJesus, E.; Berger, D., et al., Antiviral Activity, Safety, and Pharmacokinetics/Pharmacodynamics of Tenofovir Alafenamide as 10-Day Monotherapy in HIV-1-Positive Adults. *J. Acquir. Immune Defic. Syndr.* 2013, 63 (4), 449-455.
- (60) Markowitz, M.; Zolopa, A.; Squires, K., et al., Phase I/II Study of the Pharmacokinetics, Safety and Antiretroviral Activity of Tenofovir Alafenamide, a New Prodrug of the HIV Reverse Transcriptase Inhibitor Tenofovir, in HIV-Infected Adults. *J. Antimicrob. Chemother.* 2014, 69 (5), 1362-1369.

- (61) Painter, G. R.; Almond, M. R.; Trost, L. C., et al., Evaluation of Hexadecyloxypropyl-9-R-[2-(Phosphonomethoxy)Propyl]-Adenine, Cmx157, as a Potential Treatment for Human Immunodeficiency Virus Type 1 and Hepatitis B Virus Infections *Antimicrob. Agents Chemother.* 2007, 51 (12), 4538-4538.
- (62) Birkus, G.; Kutty, N.; He, G. X., et al., Activation of 9-[(R)-2-[[S]-[[S]-1-(Isopropoxycarbonyl)Ethyl]Amino] Phenoxyphosphinyl]-Methoxy]Propyl]Adenine (Gs-7340) and Other Tenofovir Phosphonoamidate Prodrugs by Human Proteases. *Mol. Pharmacol.* 2008, 74 (1), 92-100.
- (63) Babusis, D.; Phan, T. K.; Lee, W. A., et al., Mechanism for Effective Lymphoid Cell and Tissue Loading Following Oral Administration of Nucleotide Prodrug Gs-7340. *Mol. Pharm.* 2013, 10 (2), 459-466.
- (64) Hostetler, K. Y., Alkoxyalkyl Prodrugs of Acyclic Nucleoside Phosphonates Enhance Oral Antiviral Activity and Reduce Toxicity: Current State of the Art. *Antiviral Res.* 2009, 82 (2), A84-98.
- (65) Thornton, P. J.; Kadri, H.; Miccoli, A., et al., Nucleoside Phosphate and Phosphonate Prodrug Clinical Candidates. *J. Med. Chem.* 2016, 59 (23), 10400-10410.
- (66) Chimerix, I. Chimerix's Antiviral Cmx157 Demonstrates Positive Phase 1 Results with Favorable Pharmacokinetics, Safety & Tolerability 2010.
<http://ir.chimerix.com/releasedetail.cfm?releaseid=752310>.
- (67) Porter, C. J.; Trevaskis, N. L.; Charman, W. N., Lipids and Lipid-Based Formulations: Optimizing the Oral Delivery of Lipophilic Drugs. *Nat. Rev. Drug Discovery* 2007, 6 (3), 231-48.
- (68) Wahbeh, G. T.; Christie, D. L., Chapter 2 - Basic Aspects of Digestion and Absorption A2 - Wyllie, Robert. In *Pediatric Gastrointestinal and Liver Disease (Third Edition)*, Hyams, J. S.; Editor, A.; Kay, M., Eds. W.B. Saunders: London, 2006; pp 11-23.
- (69) Coles, B.; Ketterer, B., The Role of Glutathione and Glutathione Transferases in Chemical Carcinogenesis. *Crit. Rev. Biochem. Mol. Biol.* 1990, 25 (1), 47-70.

- (70) Pastore, A.; Federici, G.; Bertini, E., et al., Analysis of Glutathione: Implication in Redox and Detoxification. *Clin. Chim. Acta* 2003, 333 (1), 19-39.
- (71) Droge, W.; Schulze-Osthoff, K.; Mihm, S., et al., Functions of Glutathione and Glutathione Disulfide in Immunology and Immunopathology. *FASEB J.* 1994, 8 (14), 1131-8.
- (72) Schulz, J. B.; Lindenau, J.; Seyfried, J., et al., Glutathione, Oxidative Stress and Neurodegeneration. *Eur. J. Biochem.* 2000, 267 (16), 4904-11.
- (73) Schafer, F. Q.; Buettner, G. R., Redox Environment of the Cell as Viewed through the Redox State of the Glutathione Disulfide/Glutathione Couple. *Free Radic. Biol. Med.* 2001, 30 (11), 1191-1212.
- (74) Gilbert, H. F., [2] Thiol/Disulfide Exchange Equilibria and Disulfidebond Stability. *Methods Enzymol.* 1995, 251, 8-28.
- (75) Benzaria, S.; Pelicano, H.; Johnson, R., et al., Synthesis, in Vitro Antiviral Evaluation, and Stability Studies of Bis(S-Acyl-2-Thioethyl) Ester Derivatives of 9-[2-(Phosphonomethoxy)Ethyl] Adenine (PMEA) as Potential PMEA Prodrugs with Improved Oral Bioavailability. *J. Med. Chem.* 1996, 39 (25), 4958-4965.
- (76) Gosselin, G. M., (FR), Imbach, Jean-louis (Montpellier, FR), Perigaud, Christian (Montpellier, FR) Biologically Active Phosphotriester-Type Compounds. 2002.
- (77) Harpp, D. N.; Ash, D. K.; Back, T. G., et al., A New Synthesis of Unsymmetrical Disulfides. *Tetrahedron Lett.* 1970, 11 (41), 3551-3554.
- (78) Boustany, K. S.; Sullivan, A. B., Chemistry of Sulfur Compounds-VI. A Novel Method for the Preparation of Disulfides. *Tetrahedron Lett.* 1970, 11 (41), 3547-3549.
- (79) Brois, S. J.; Pilot, J. F.; Barnum, H. W., New Synthetic Concepts in Organosulfur Chemistry. I. New Pathway to Unsymmetrical Disulfides. The Thiol-Induced Fragmentation of Sulfenyl Thiocarbonates. *J. Am. Chem. Soc.* 1970, 92 (26), 7629-7631.
- (80) Hunter, R.; Cairra, M.; Stellenboom, N., Inexpensive, One-Pot Synthesis of Unsymmetrical Disulfides Using 1-Chlorobenzotriazole. *J. Org. Chem.* 2006, 71 (21), 8268-8271.

- (81) Riggers, R. J.; Pomorski, T.; Holthuis, J. C. M., et al., Lipid Traffic: The Abc of Transbilayer Movement. *Traffic* 2000, 1 (3), 226-234.
- (82) Contreras, F. X.; Sanchez-Magraner, L.; Alonso, A., et al., Transbilayer (Flip-Flop) Lipid Motion and Lipid Scrambling in Membranes. *FEBS Lett.* 2010, 584 (9), 1779-86.
- (83) Showell, M., *Handbook of Detergents: Formulation.* Taylor & Francis: 2005.
- (84) Boon, J. M.; Smith, B. D., Chemical Control of Phospholipid Distribution across Bilayer Membranes. *Med. Res. Rev.* 2002, 22 (3), 251-281.
- (85) Bevers, E. M.; Comfurius, P.; Dekkers, D. W., et al., Lipid Translocation across the Plasma Membrane of Mammalian Cells. *Biochim. Biophys. Acta* 1999, 1439 (3), 317-30.
- (86) Daleke, D. L., Regulation of Transbilayer Plasma Membrane Phospholipid Asymmetry. *J. Lipid Res.* 2003, 44 (2), 233-242.
- (87) Oumzil, K.; Khiati, S.; Grinstaff, M. W., et al., Reduction-Triggered Delivery Using Nucleoside-Lipid Based Carriers Possessing a Cleavable Peg Coating. *J. Control. Release* 2011, 151 (2), 123-30.
- (88) Iyer, S.; Hengge, A. C., The Effects of Sulfur Substitution for the Nucleophile and Bridging Oxygen Atoms in Reactions of Hydroxyalkyl Phosphate Esters. *J. Org. Chem.* 2008, 73 (13), 4819-4829.
- (89) Xia, F.; Zhu, H., Density Functional Calculations on the Effect of Sulfur Substitution for 2'-Hydroxypropyl-P-Nitrophenyl Phosphate: Co Vs. Po Bond Cleavage. *Bioorg. Chem.* 2012, 40, 99-107.
- (90) Giesler, K. E.; Liotta, D. C., Next-Generation Reduction Sensitive Lipid Conjugates of Tenofovir: Antiviral Activity and Mechanism of Release. *J. Med. Chem.* 2016, 59 (22), 10244-10252.

- (91) Pertusati, F.; Serpi, M.; McGuigan, C., Medicinal Chemistry of Nucleoside Phosphonate Prodrugs for Antiviral Therapy. *Antivir. Chem. Chemother.* 2012, 22 (5), 181-203.

Chapter 2

Synthesis and Evaluation of Other Disulfide Prodrug Conjugates

Abstract. In chapter 1, our disulfide prodrugs successfully improved the antiviral activity of tenofovir and demonstrated pronounced stability in human serum. The question remains: can this technology be exploited to improve the antiviral profile of other NAs, particularly nucleotides? In this chapter, we prepared disulfide prodrug conjugates of emtricitabine, adefovir, **2.1** and **2.2** and compared their antiviral activity to either the parent nucleoside, the hexadecyloxypropyl conjugate, or the clinically approved prodrug formulation to assess if our technology is comparable.

2.1 Introduction

The preliminary success of our TFV disulfide lipid conjugates prompted us to prepare disulfide prodrugs of other NAs. We chose to focus on NAs currently in the clinic or have demonstrated promising pre-clinical data. This led to the selection of a handful of compounds shown in Figure 2.1. Emtricitabine (FTC) and adefovir (AFV) are approved for the treatment of HIV and HBV, respectively. Compound **2.1** and **2.2** are potent anti-HCV NAs (**2.2** is the sofosbuvir nucleoside without the prodrug) and require phosphate prodrugs to achieve sufficient antiviral activity.

In contrast to phosphonate lipid prodrugs, phosphate lipid prodrugs have two potential

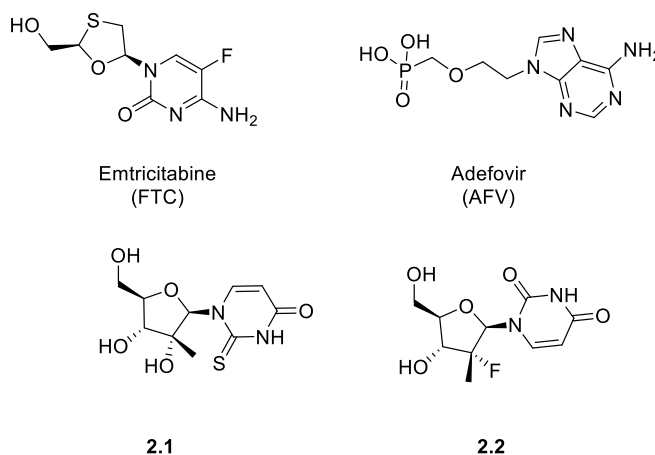


Figure 2.1. Target nucleosides for disulfide prodrug conjugation.

hydrolysis sites, **a** and **b**, shown in Figure 2.2. Site **a** is clipped by phospholipase C and site **b** is targeted by phospholipase D (PLD). Hydrolysis by PLD removes the phosphate head group entirely to generate the unphosphorylated nucleoside. The latter reaction is undesirable, however, it remains unclear what structural factors favor PLC hydrolysis over PLD hydrolysis, especially for unnatural lipids. We proceeded to synthesize lipid conjugates of the nucleoside analogues in Figure 2.1 with two questions in mind 1) does our technology improve the antiviral activity of these compounds and 2)

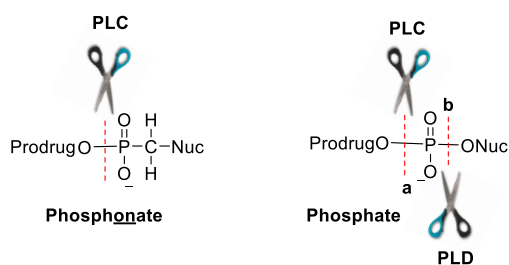


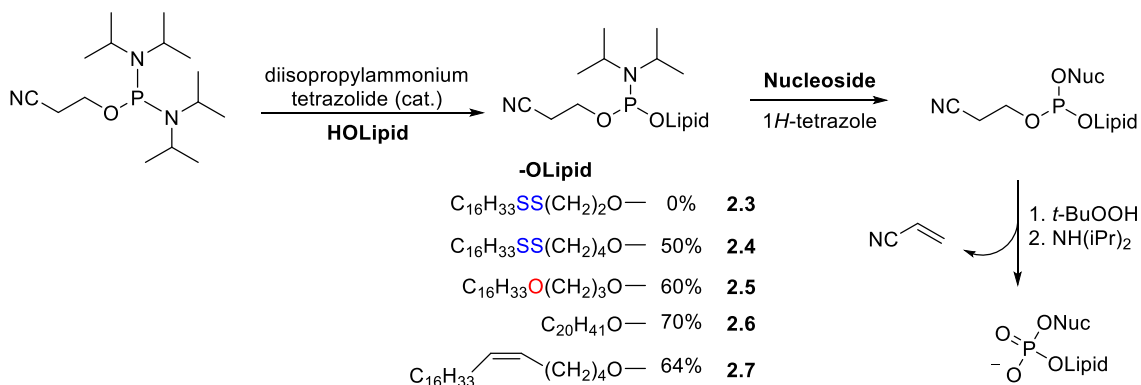
Figure 2.2. Cleavage sites for phosphonate and phosphate lipid prodrugs by PLC and PLD.

do our disulfide prodrugs confer an advantage over other phosphate prodrugs? These questions are addressed herein.

2.2 Synthesis of Phosphate Disulfide Conjugates

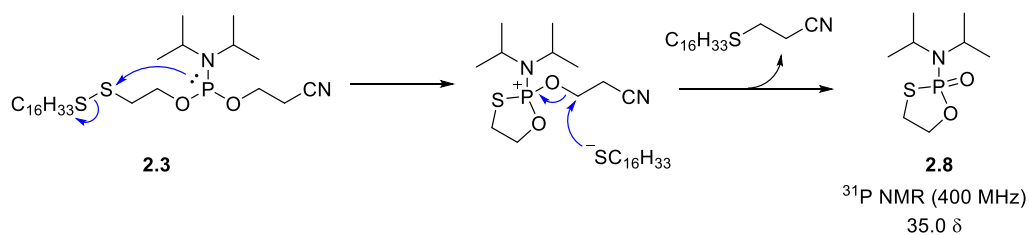
To install the requisite phosphate moiety on FTC, **2.1**, and **2.2**, we chose to use phosphoramidite chemistry that has seen extensive use in the synthesis of oligonucleotides, phosphorothioates, phosphoramidates, and other molecules of biological interest.⁹²⁻⁹³ A salient feature of phosphoramidites is their sensitivity towards weak acids, including silica gel, to promote nucleophilic attack at phosphorous and facilitate departure of the amino moiety. In the case of oligonucleotide synthesis, the nucleophile is often the 5'-hydroxyl to furnish an intermediate phosphite which is readily oxidized to a phosphate using *t*-BuOOH. The oxidized product is then treated with base to remove the cyanoethyl moiety in a final step to furnish the desired conjugates. This strategy is outlined in Scheme 2.1.

Scheme 2.1. General Synthesis of Phosphate NA Lipid Conjugates.



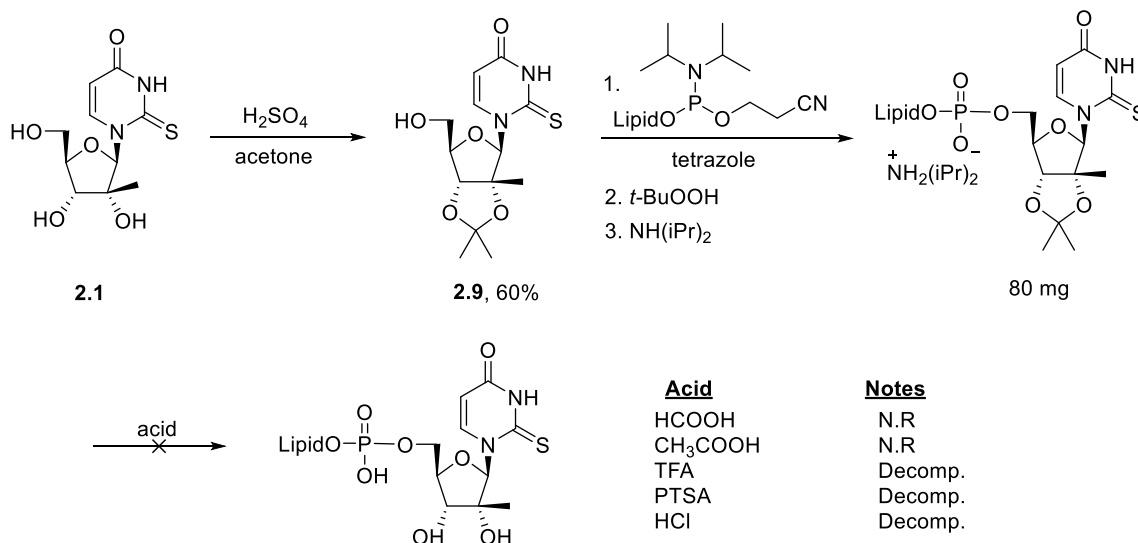
2.2.1 Preparation of Phosphoramidite Coupling Reagents

The first step in Scheme 2.1 involves the synthesis of the requisite phosphoramidite coupling reagent. 2-cyanoethyl *N,N,N',N'*-tetraisopropyl phosphoramidite (1.1 equiv.) was reacted with lipid **1.5d** (1.0 equiv.) in the presence of catalytic diisopropylammoniumtetrazolide (0.6 equiv.) in an attempt to furnish the corresponding phosphoramidite **2.3** where -OLipid is -O(CH₂)₂SSC₁₆H₃₃. Unfortunately, **2.3** could not be accessed. The reaction yielded a single product at 35.0 δ (³¹P NMR) consistent with phosphorous oxidation. It was later discovered that **2.3** experiences spontaneous intramolecular reduction shown in Scheme 2.2 to expel hexadecylthiolate which subsequently removes the cyanoethyl moiety to exclusively yield heterocycle **2.8**. Literature characterization data for compound **2.8** is limited, however, the observed ³¹P resonance matches that of similar compounds.⁹⁴ The remaining phosphoramidite intermediates **2.4-2.7** were synthesized without incident, however, **2.4** gradually decomposed at -20 °C and had to be prepared fresh prior to coupling.

Scheme 2.2. Intramolecular Reduction of **2.3**.

2.2.2 Lipid Conjugates of Nucleoside Analogue **2.1**

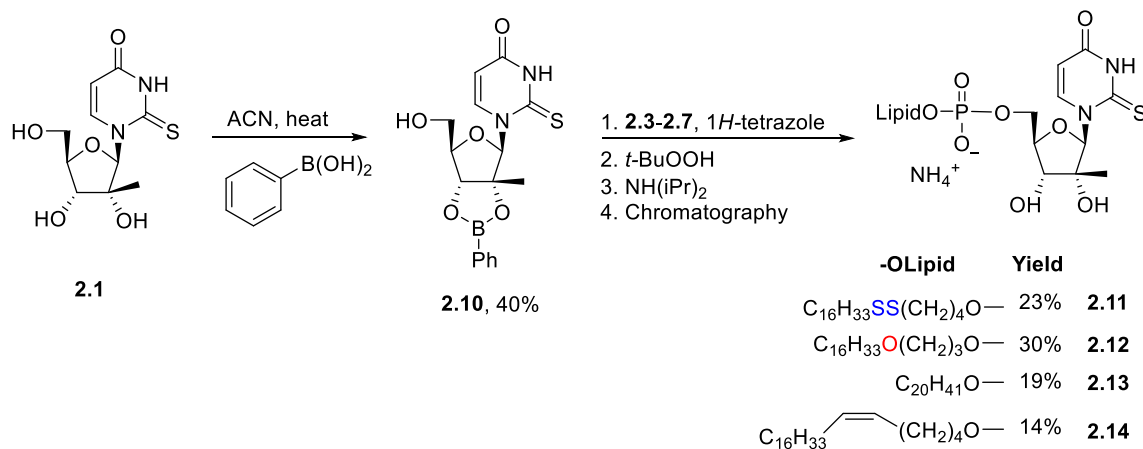
Compound **2.1** was identified by the Emory Institute for Drug Development (EIDD) and was found to exhibit anti-HCV and anti-RSV activity.⁹⁵ Preliminary studies revealed that **2.1** is poorly phosphorylated to the monophosphate and would therefore benefit from phosphate prodrug conjugation. Analogue **2.1** was synthesized on a multi-gram scale by Manohar Saindane and provided to us for further derivatization. Initial treatment of **2.1** with 1*H*-tetrazole and **2.4** resulted in six different phosphite signals in the ³¹P NMR, indicating no selectivity for the primary alcohol over the secondary and tertiary hydroxyls present on the ribose ring. We therefore sought to protect **2.1** as an acetonide before proceeding with the 1*H*-tetrazole coupling reaction. Protection of **2.1** with acetone and sulfuric acid proceeded relatively smoothly to furnish acetonide **2.9** in 60% yield (~1 g). Subsequent coupling of **2.9** to **2.4-2.7** was capricious and low-yielding (Scheme 2.3). The reaction was found to be extremely sensitive to adventitious water and a significant amount of the phosphoramidite was undesirably converted to the H-phosphonate (8.5 ppm). Drying the nucleoside with acetonitrile/toluene was insufficient to remove any residual hydrates and heating **2.1** at 100 °C overnight also did not improve the yield. Despite this setback, the reaction was sufficient to produce

Scheme 2.3. Attempted Lipid Conjugation of **2.1** with an Acetonide Protecting Group.

80 mg of conjugated material after several attempts. The conjugate was then subjected to acid hydrolysis to remove the acetonide in an attempt to furnish the final product shown in Scheme 2.3. This transformation proved to be very challenging. Formic and acetic acid failed to result in any reaction after 24 h and stronger acids such as TFA, PTSA, and HCl resulted in extensive decomposition. All attempts to optimize the reaction time, concentration, and solvent were unsuccessful. We therefore required an alternative protecting group that could 1) selectively mask the 1,2-diol and 2) could be gently removed without harsh acids.

In 2014, researchers at Idenix reported the use of a phenylboronate protecting group to mask the 1,2-diol of 2'-C-methylcytidine.⁹⁶ Boronates are readily oxidized/cleaved by peroxides and are also sensitive to water. Given the use of *t*-BuOOH in our coupling reaction, we thought that phenylboronate would be an ideal protecting group that could be readily removed without introducing another step to the reaction sequence. Compound **2.1** was suspended in acetonitrile and refluxed for 3 h in the presence of anhydrous sodium sulfate to furnish phenylboronate **2.10** in 40% yield (Scheme

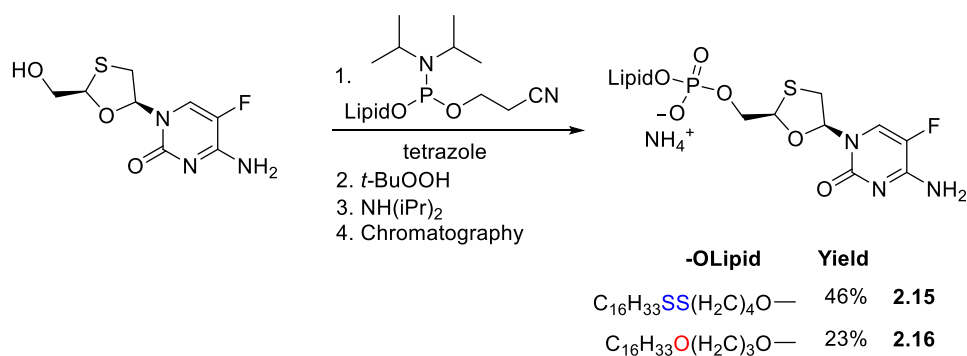
Scheme 2.4. Lipid Conjugation of **2.1** with a Phenylboronate Protecting Group.



2.4). This product was not stable to silica gel and could not be purified by column chromatography. Pure compound was obtained by recrystallizing **2.10** from DCM and acetonitrile. Compound **2.10** was then coupled to phosphoramidites **2.4-2.7** to produce the desired lipid conjugates **2.11-2.14**.

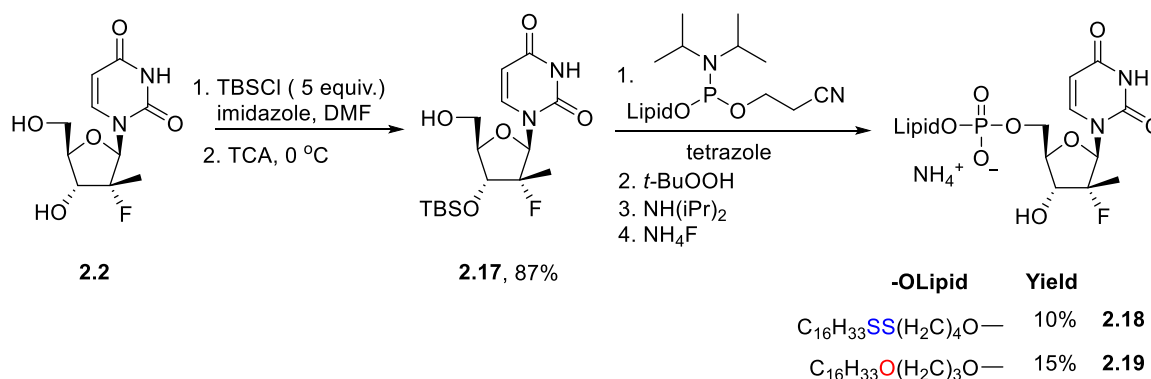
2.2.3 Lipid Conjugates of Emtricitabine (FTC)

FTC is readily anabolized to the monophosphate, diphosphate and eventually triphosphate by cellular kinases and is administered to HIV-positive patients as the free nucleoside.⁹⁷ The plasma elimination half-life for FTC is roughly 8-10 hours and the maximum concentration of 5'-FTC triphosphate peaks after 12 hours from a single 200 mg dose to permit a convenient once-a-day dosing regimen in combination with TAF and another anti-HIV drug from a different drug class. However, if once-a-week dosing regimens are to be the future of HIV therapy, the half-life of FTC must be significantly augmented. This may be possible with phosphate lipid prodrugs. In contrast to compound **2.1**, FTC bears only one hydroxyl unit and does not require protecting groups prior to the introduction of the coupling reagent. FTC was successfully coupled with **2.4** and **2.5** to yield conjugates **2.15** and **2.16** (Scheme 2.5).

Scheme 2.5. HDP and Disulfide Lipid Conjugates of FTC.

2.2.4 Lipid Conjugates of Nucleoside Analogue 2.2

Compound **2.2** features a 3'-OH that requires protection prior to prodrug coupling. Note that **2.2** is the sofosbuvir nucleoside without the phosphoramidite prodrug attached. Without the phosphoramidate headgroup, **2.2** is poorly phosphorylated to the triphosphate and demonstrates poor anti-HCV activity, but with the prodrug, this molecule successfully cures HCV. To generate the HDP and disulfide prodrugs, **2.2** was persilylated with TBSCl and imidazole then selectively deprotected with trichloroacetic acid (TCA) at 0 °C to furnish the 3'-OTBS protected nucleoside **2.17** (Scheme 2.6). **2.17** was then reacted with our standard coupling conditions which proceeded smoothly.

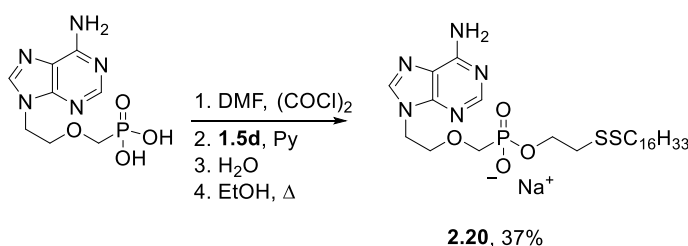
Scheme 2.6. HDP and Disulfide Lipid Conjugates of NA **2.2**.

Final TBS protection with ammonium fluoride was slow in aqueous methanol and took almost a week to furnish **2.18** and **2.19**.

2.2.5 Lipid Conjugates of AFV

AFV disulfide conjugates were readily prepared using the DMF/oxalyl chloride methodology exploited for TFV. Scheme 2.7 shows the preparation of compound **2.20**.

Scheme 2.7. Synthesis of Compound **2.20**.



2.3 Biological Evaluation of Phosphate Lipid Prodrugs

Compounds **2.11-2.16** and **2.18-2.20** were evaluated against HIV-1, HBV, or HCV. Nucleosides **2.1** and FTC were also assessed for comparison. These results are summarized in Table 2.1. With respect to anti-HIV activity, emtricitabine (FTC) exhibited an EC_{50} value of $0.01 \mu\text{M}$ without any prodrug attachment. This is the FDA-approved formulation of FTC that is administered to patients suffering from chronic HIV-1 infection. When FTC is conjugated to HDP or our disulfide prodrug, no dramatic change in activity was observed. Compound **2.15** is somewhat less potent ($EC_{50} = 0.07 \mu\text{M}$) than FTC whereas **2.16** is nearly equipotent to FTC ($EC_{50} = 0.008 \mu\text{M}$). Both of these EC_{50} values are nearly identical when the degree of experimental error is taken into account. The data suggest that PLD cleavage is rapid and significant in PBMCs to remove the phosphate prodrug entirely and release FTC rather than FTC monophosphate. Thus, there appears to be no potency advantage with lipid prodrug conjugates of emtricitabine.

Table 2.1. Antiviral Activity Various NAs and Prodrug Conjugates^a

Cmpd	Parent Nuc	Prodrug	HIV-1		HBV		HCV	
			EC ₅₀ ^b (PBMCs)	CC ₅₀ ^c (PBMCs)	EC ₅₀ ^b (HepG2)	CC ₅₀ ^c (HepG2)	EC ₅₀ ^b (Huh7)	CC ₅₀ ^c (Huh7)
2.1	2.1	none	-	-	-	-	27.8	> 100
2.11	2.1	C ₁₆ H ₃₃ SS(CH ₂) ₄ O-	-	-	-	-	12.9	37.8
2.12	2.1	C ₁₆ H ₃₃ O(CH ₂) ₃ O-	-	-	-	-	1.38	> 100
2.13	2.1	C ₂₀ H ₄₁ O-	-	-	-	-	2.37	50.5
2.14	2.1	C ₁₆ H ₃₃ CH=CH(CH ₂) ₄ O-	-	-	-	-	2.64	38.2
02072-1*	2.1	Protide	-	-	-	-	0.187*	> 400*
FTC	FTC	none	0.01	> 100	-	-	-	-
2.15	FTC	C ₁₆ H ₃₃ SS(CH ₂) ₄ O-	0.07	74	-	-	-	-
2.16	FTC	C ₁₆ H ₃₃ O(CH ₂) ₃ O-	0.008	67	-	-	-	-
2.18	2.2	C ₁₆ H ₃₃ SS(CH ₂) ₄ O-	-	-	-	-	N/A	N/A
2.19	2.2	C ₁₆ H ₃₃ O(CH ₂) ₃ O-	-	-	-	-	N/A	N/A
Sofosbuvir	2.2	Protide	-	-	-	-	0.05	> 1.0
2.20	AFV	C ₁₆ H ₃₃ SS(CH ₂) ₂ O-	0.002	0.09	8.67 [†]	0.54 [†]	-	-

^aAll data represent an average of triplicate experiments. ^bEC₅₀, effective concentration (in μM) required to inhibit HIV-1, HBV, or HCV replication by 50%. ^cCC₅₀, effective concentration (in μM) required to reduce the viability of uninfected cells by 50%.

[†]Assayed in AD38 cells. *Reported in reference 95.

Conjugate **2.20** was also assayed against HIV-1 and features the parent nucleoside, adefovir. **2.20** exhibited potent anti-HIV activity (EC₅₀ = 0.002 μM), but a relatively poor therapeutic index of 45 in PBMCs. Furthermore, **2.20** demonstrated unremarkable activity against HBV in HepG2 cells. This is interesting given that AFV is exceptionally active against HBV and suggests that PLC is not highly expressed in this cell line to remove the prodrug. Relatively poor antiviral activity was also noted for compounds **2.1** and **2.11-2.14** which were assayed in a HCV replicon assay using Huh7 hepatocytes. In the absence of a prodrug, compound **2.1** poorly inhibited HCV replication with an EC₅₀ = 27.8 μM. Attachment of our disulfide prodrug (**2.11**) only improved activity by 2-fold (EC₅₀ = 12.9 μM), however, this improvement is not consistent with the antiviral activity profile observed for other lipid prodrugs **2.12-2.14**. This anomaly raises questions about the accuracy of the results for compound **2.11**. Nonetheless, none of the lipid conjugates of **2.1** achieved the antiviral activity of 02072-1, the protide prodrug of compound **2.1** disclosed in reference 90. Our data seems to indicate that **2.12-2.14**

cleave slowly in Huh7 cells and therefore fail to completely “unlock” the antiviral potential of **2.1** once released into the cytosol. Whether this is due to poor PLC expression in Huh7 cells or some other mechanism is unknown at this time and leaves us with inconclusive evidence about the advantage of lipid prodrugs for the treatment of HCV.

2.4 Experimental Details

2.4.1 Anti-HIV Assay

All HIV assays were performed by ImQuest BioSciences (Frederick, MD). All compounds were solubilized at 40 mM in DMSO and stored at -20°C. Test materials were evaluated up to 100 μ M, and five serial logarithmic dilutions. AZT and 3TC were obtained from the NIH AIDS Research and Reference Reagent Program and used as controls in the anti-HIV and anti-HBV assay, respectively. Fresh human PBMCs were obtained from a commercial source and determined to be HIV and HBV negative. The leukophoresed blood cells were washed repeatedly with PBS, then diluted 1:1 with Dulbecco’s phosphate buffered saline (PBS) and layered over 15 ml of Ficoll-Hypaque density gradient in a 50 ml conical centrifuge tube. The tubes were centrifuged for 30 min at 600 g. Banded PBMCs were gently aspirated from the resulting interface and washed three times with PBS. After the final wash, cell number was determined by Trypan Blue dye exclusion and cells re-suspended at 1×10^6 cells/mL in RPMI 1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 μ g/mL PHA-P, 100 U/ml penicillin and 100 μ g/mL streptomycin and allowed to incubate for 48-72 hr at 37°C. After incubation, PBMCs were centrifuged and re-suspended in tissue culture medium (RPMI 1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 μ g/mL PHA-P, 100 U/ml penicillin and 100 μ g/mL streptomycin, 3.6 ng/mL recombinant human IL-2). The cultures were maintained until use by half-volume culture changes with fresh IL-2 containing tissue culture medium

every 3 days. Assays were initiated with PBMCs at 72 hr post PHA-P stimulation. Immediately prior to use, target cells were re-suspended in fresh tissue culture medium at 1×10^6 cells/ml and plated in the interior wells of a 96-well round bottom microliter plate at 50 μ L/well. Then, 100 μ L of 2X concentrations of compound-containing medium was transferred to the 96-well plate containing the cells in 50 μ L of the medium. Following addition of test compound to the wells, 50 μ L of a predetermined dilution of HIV virus (prepared at 4x of final desired in-well concentration) was added, and mixed well. For infection, 50-150 TCID₅₀ of each virus was added per well (final MOI approximately 0.002). PBMCs were exposed in triplicate to virus and cultured in the presence or absence of the test compound at varying concentrations as described for the 96-well microliter plates. After 7 days, HIV-1 replication was quantified in the tissue culture supernatant by measurement of reverse transcriptase activity. Wells with cells and virus only served as virus controls. Separate plates were identically prepared without virus for cytotoxicity studies. Reverse transcriptase activity was measured in cell-free supernatants using a standard radioactive incorporation polymerization assay.

2.4.2 Anti-HBV Assay

All HBV assays were performed by ImQuest BioSciences (Frederick, MD). One hundred microliters (100 μ L) of wells of a 96-well flat-bottom plate at a density of 1×10^4 cells per well were incubated at 37°C in 5% CO₂ for 24 h. Following incubation, six ten-fold serial dilutions of test compound prepared in RPMI1640 medium with 10% fetal bovine serum were added to individual wells of the plate in triplicate. Six wells in the plate received medium alone as a virus control only. The plate was incubated for 6 days at 37°C at 5% CO₂. The culture medium was changed on day 3 with medium containing the indicated concentration of each compound. One hundred microliters (100 μ L) of supernatant was collected from each well for analysis of viral DNA by qPCR and cytotoxicity was evaluated by XTT staining of the cell culture monolayer on the sixth day (see below). Ten microliters (10 μ L) of cell

culture supernatant collected on the sixth day was diluted in qPCR dilution buffer (40 µg/ml sheared salmon sperm DNA) and boiled for 15 minutes. Quantitative real time PCR was performed in 386 well plates using an Applied Biosystems 7900HT Sequence Detection System and the supporting SDS 2.4 software. Five microliters (5 µL) of boiled DNA for each sample and serial 10-fold dilutions of a quantitative DNA standard were subjected to real time Q-PCR using Platinum Quantitative PCR SuperMix-UDG (Invitrogen) and specific DNA oligonucleotide primers (IDT, Coralville, ID) HBV-AD38-qF1 (5'-CCG TCT GTG CCT TCT CAT CTG-3'), HBV-AD38-qR1 (5'-AGT CCA AGA GTY CTC TTA TRY AAG ACC TT-3') and HBV-AD38-qP1 (5'-FAM-CCG TGT GCA /ZEN/CTT GCG TTC ACC TCT GC-3'BHQ1) at a final concentration of 0.2 µM for each primer in a total reaction volume of 15 µL. The final HBV DNA copy number in each sample was interpolated from the standard curve by the SDS2.4 software and the data were analyzed by Excel.

2.4.3 Anti-HCV Assay

All HCV assays were performed by ImQuest BioSciences (Frederick, MD). Huh-7 luc/neo ET cells bearing a discistronic HCV genotype 1b luciferase reporter replicon were plated at 7.5×10^3 cells/mL in duplicate 96-well plates for the parallel determination of antiviral efficacy and cytotoxicity. The plates were cultured for 24 h prior to the addition of compounds. Six serial dilutions of the test compounds (highest concentration 100 µM) and Sofosbuvir (highest test 1.0 µM) were prepared in cell culture medium and added to the cultured cells in triplate wells for each dilution. Six wells in the test plates received medium alone as an untreated control. Following 72 hours of culture in the presence of compound, one of the plates was used for the determination of cytotoxicity by staining with XTT and the other for antiviral efficacy by determination of luciferase reporter activity.

2.4.4 Cytotoxicity Studies

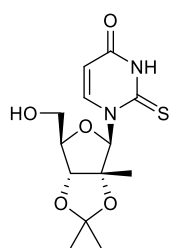
Cytotoxicity was evaluated by staining uninfected cells with tetrazolium dye XXT with spectrophotometric readings at 450/650 nm with a Molecular Devices Vmax plate reader. These experiments were run in triplicate for each compound tested.

2.4.5 Chemical Synthesis and General Procedures

All reagents were obtained from commercial suppliers and used without further purification unless otherwise specified. Reaction progress was monitored by either thin layer chromatography (TLC) using pre-coated aluminum-backed silica gel plates (60 F₂₅₄ Merk, article 5554), by NMR, or liquid chromatography-mass spectrometry (LC-MS) on an Agilent Technologies 6100 quadrupole instrument equipped with UV detection at 254 and 210 nm and a Varian C8 analytical column. LC-MS analyses were performed using a step-wise H₂O/MeOH gradient with the % MeOH increasing from 75-95% over the course of 3 min unless otherwise specified. Flash column chromatography was conducted using CombiFlash Rf 200 (Telendyne-Isco) automated flash chromatography system with hand-packed RediSep columns. Evaporation of solvents was carried out on a rotary evaporator under reduced pressure and under ultra-high vacuum (UHV) where appropriate. ¹H NMR and ¹³C NMR spectra were recorded at ambient temperature on a Varian 400 spectrometer. ³¹P spectra were recorded at ambient temperature on either a Mercury 300 or Varian 400 spectrometer. Unless otherwise specified, all NMR spectra were obtained in deuterated chloroform (CDCl₃) and referenced to the residual solvent peak. Chemical shifts are given in δ values and coupling constants are reported in hertz (Hz). Melting points were determined on a MelTemp melting apparatus and are uncorrected. High resolution mass-spectra (HRMS) were acquired on a VG 70-S Nier Johnson or JEOL mass spectrometer. Elemental analyses were performed by Atlantic Microlabs (Norcross, GA) for C, H, N analysis and are in agreement with the proposed structures with purity \geq 95%.

General Procedures

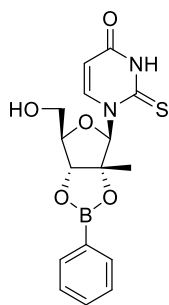
Procedure 3A: To a stirring solution of diisopropylammonium tetrazolide (0.126 mmol) and lipid L (0.336 mmol) in CDCl_3 (2 mL) was added 3-((bis(diisopropylamino)phosphino)oxy)propanenitrile (0.315 mmol). The mixture stirred at room temperature for 1 h with progress monitored by ^{31}P NMR. Upon completion (~2 h) this mixture was added directly to a stirring solution of nucleoside N (0.210 mmol) and excess 1*H*-tetrazole (0.840 mmol) in anhydrous ACN under argon. Stirring continued for an additional hour with progress monitored by ^{31}P NMR. Then, excess *tert*-butylhydroperoxide (1.050 mmol) was added and the mixture stirred for an additional 30 minutes. The solvents were then evaporated under reduced pressure and the resulting residue was purified via silica gel chromatography using a DCM/MeOH gradient (0-5% MeOH). The desired intermediate was isolated and dissolved in $\text{CDCl}_3/\text{MeOH}$ (3 mL) and reacted with diisopropylamine (1 mL). The resulting mixture stirred for 5 h at room temperature. Then, the solvents were evaporated and the vessel was placed under UHV for 1 h to remove any residual diisopropylamine. The residue was purified via silica gel column chromatography using a DCM/DCM:MeOH: NH_4OH (80:20:3) gradient to furnish the title compound as an ammonium salt.



1-((3aR,4R,6R,6aR)-6-(Hydroxymethyl)-2,2,3a-trimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (2.9)

To a stirring mixture of 1-((2R,3R,4R,5R)-3,4-dihydroxy-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (0.50 g, 1.8 mmol) in acetone (4.0 mL) was added catalytic sulfuric acid. The mixture stirred at room temperature overnight then concentrated under

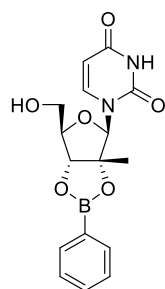
reduced pressure. Water was added followed by solid sodium bicarbonate. The solution was chilled at 0 °C to precipitate the title compound 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-2,2,3a-trimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (0.25 g, 0.80 mmol, 44 % yield) as an off-white solid that required no further purification. The product was dried under UHV and co-evaporated with DCM and acetonitrile (10:1) three times. ¹H NMR (400 MHz, *d*₆-DMSO) δ 8.05 (d, *J* = 8.2 Hz, 1H), 6.96 (s, 1H), 6.01 (d, *J* = 8.2 Hz, 1H), 5.41 (s, 1H), 4.50 (d, *J* = 3.3 Hz, 1H), 4.19 (q, *J* = 3.0 Hz, 1H), 3.69 (dd, *J* = 25.8, 12.0 Hz, 2H), 1.50 (s, 3H), 1.36 (s, 3H), 1.23 (s, 3H). ¹³C NMR (101 MHz, *d*₆-DMSO) δ 176.6, 159.8, 141.7, 113.4, 106.6, 95.8, 90.9, 85.6, 84.9, 61.2, 28.6, 27.8, 20.0. HRMS (ESI) *m/z* calculated for C₁₃H₁₉N₂O₅S [M + H]⁺ : 315.1009, found 315.1011.



1-((3aR,4R,6R,6aR)-6-(Hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (2.10)

A suspension of 1-((2R,3R,4R,5R)-3,4-dihydroxy-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (250. mg, 0.911 mmol) in acetonitrile (6 mL) was mixed with phenylboronic acid (122 mg, 1.00 mmol) at room temperature. Powdered 4 Å molecular sieves were added and the mixture was heated to 80 °C under inert atmosphere for 3 h. Note that reaction progress could not be monitored by TLC. The mixture was cooled to room temperature, diluted with DCM filtered over Whatman filter paper. The supernatant was collected and concentrated under reduced pressure then the resulting solid was taken up in chloroform and chilled in an ice bath to

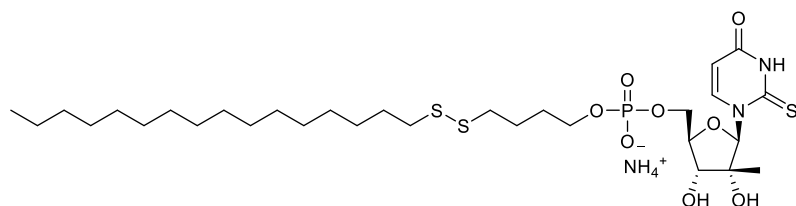
crystallize the title compound 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one (300 mg, 0.911 mmol, 91 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 12.84 (s, 1H), 8.07 (d, $J = 8.2$ Hz, 1H), 7.76 (d, $J = 6.7$ Hz, 2H), 7.60 - 7.53 (m, 1H), 7.45 (t, $J = 7.4$ Hz, 2H), 7.08 (s, 1H), 6.09 (dd, $J = 8.1, 2.0$ Hz, 1H), 5.44 (t, $J = 5.1$ Hz, 1H), 4.73 (d, $J = 5.2$ Hz, 1H), 4.19 (dt, $J = 5.2, 2.7$ Hz, 1H), 3.85 (ddd, $J = 12.2, 4.7, 2.7$ Hz, 1H), 3.79 (ddd, $J = 12.3, 5.3, 3.2$ Hz, 1H), 1.31 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 176.6, 159.8, 141.7, 135.1, 132.3, 128.4, 110.0, 106.89 (2), 95.3, 92.1, 86.3, 84.8, 60.6, 55.4, 21.0. HRMS (ESI) m/z calculated for $\text{C}_{16}\text{H}_{17}\text{BN}_2\text{O}_5\text{S}$ $[\text{M} + \text{H}]^+$: 361.1024, found 361.1022.



1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)pyrimidine-2,4(1H,3H)-dione (SI-21)

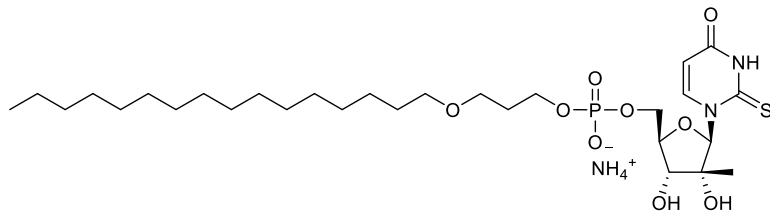
A suspension of 1-((2R,3R,4R,5R)-3,4-dihydroxy-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (1.00 g, 3.87 mmol) was mixed with phenylboronic acid (0.519 g, 4.26 mmol) at room temperature. Anhydrous magnesium sulfate was added and the mixture stirred at 80 °C for 3 h. Note that reaction progress could not be monitored by TLC. After 3 h, the mixture was then diluted with toluene, cooled to room temperature, then filtered over filter paper. The supernatant was collected and the solvent evaporated under reduced pressure to furnish the titled compound 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)pyrimidine-2,4(1H,3H)-dione (1.20 g, 3.49 mmol, 90 % yield) as an off-white solid. The material

was re-crystallized from DCM and extensively dried via the co-evaporation of toluene before use. ^1H NMR (400 MHz, d_6 -DMSO) δ 11.51 (s, 1H), 7.89 (d, $J = 8.1$ Hz, 1H), 7.75 (d, $J = 6.9$ Hz, 2H), 7.57 (t, $J = 7.4$ Hz, 1H), 7.45 (t, $J = 7.5$ Hz, 2H), 6.14 (s, 1H), 5.72 (dd, $J = 8.1, 1.7$ Hz, 1H), 5.30 (t, $J = 5.4$ Hz, 1H), 4.68 (d, $J = 4.9$ Hz, 1H), 4.13 (dd, $J = 7.8, 3.3$ Hz, 1H), 3.86 - 3.71 (m, 2H), 1.30 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 163.4, 150.7, 135.0, 132.3, 128.4, 101.9, 91.4, 85.3, 60.9, 20.6. HRMS (ESI) m/z calculated for $\text{C}_{16}\text{H}_{18}\text{BN}_2\text{O}_6$ $[\text{M} + \text{H}]^+$: 345.1252, found 345.1253.



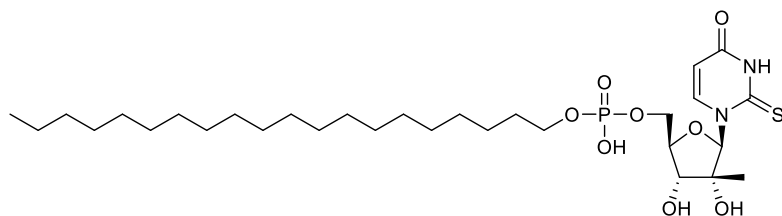
((2R,3R,4R,5R)-3,4-Dihydroxy-4-methyl-5-(4-oxo-2-thioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)methyl (4-(hexadecylsulfanyl)butyl) hydrogen phosphate (2.11)

Procedure 3A: L = (hexadecylsulfanyl)butan-1-ol, N = 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one. Purification gradient: 0-55% DCM:MeOH:NH₄OH, 23% yield. ^1H NMR (400 MHz, CDCl₃/CD₃OD) δ 7.94 (d, $J = 8.2$ Hz, 1H), 6.85 (s, 1H), 6.00 (d, $J = 8.1$ Hz, 1H), 4.16 (ddd, $J = 12.2, 4.6, 2.2$ Hz, 1H), 4.04 (ddd, $J = 12.1, 6.4, 2.0$ Hz, 1H), 3.97 (dd, $J = 8.9, 2.2$ Hz, 1H), 3.81 (d, $J = 8.9$ Hz, 1H), 3.84 - 3.76 (m, 2H), 2.65 - 2.56 (m, 4H), 1.74 - 1.63 (m, 4H), 1.63 - 1.53 (m, 2H), 1.34 - 1.10 (m, 30H), 0.80 (t, $J = 6.5$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl₃/CD₃OD) δ 176.3, 160.5, 141.33, 106.6, 95.0, 81.1 (d, $J = 7.7$ Hz), 79.2, 73.2, 72.0, 65.38 (d, $J = 5.9$ Hz), 62.6 (d, $J = 5.6$ Hz), 38.9, 38.3, 31.9, 29.7, 29.64, 29.60, 29.5, 29.4, 29.3, 29.3, 29.2, 28.5, 25.4, 22.7, 20.5, 14.1. ^{31}P NMR (162 MHz, CDCl₃/CD₃OD) δ 0.5. HRMS (ESI) m/z calculated for $\text{C}_{30}\text{H}_{55}\text{O}_8\text{N}_2\text{PS}_3$ $[\text{M} - \text{H}]^-$: 697.2785, found 697.2775. Anal. calculated for $\text{C}_{30}\text{H}_{61}\text{N}_4\text{O}_8\text{PS}_3$ (as a bisammonium salt): C, 49.16; H, 8.39; N, 7.64. Found: C, 48.85; H, 8.12; N, 7.66.



((2R,3R,4R,5R)-3,4-Dihydroxy-4-methyl-5-(4-oxo-2-thioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)methyl (3-(hexadecyloxy)propyl) hydrogen phosphate (2.12)

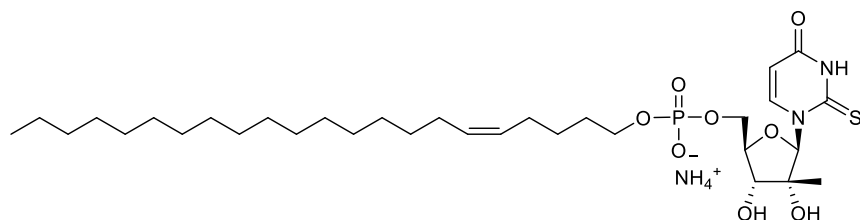
Procedure 3A: L = 3-(hexadecyloxy)propan-1-ol, N = 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one. Purification gradient: 0-55% DCM:MeOH:NH₄OH, 30% yield. ¹H NMR (399 MHz, CDCl₃/CD₃OD) δ 8.61 (s, 1H), 7.89 (d, J = 8.1 Hz, 1H), 6.79 (s, 1H), 5.96 (d, J = 7.8 Hz, 1H), 4.23 - 4.13 (m, 1H), 4.10 - 4.02 (m, 1H), 3.99 (d, J = 9.2 Hz, 1H), 3.92 - 3.84 (m, 2H), 3.82 (d, J = 8.8 Hz, 1H), 3.44 (t, J = 6.6 Hz, 2H), 3.33 (t, J = 6.8 Hz, 2H), 1.86 - 1.77 (m, 2H), 1.52 - 1.43 (m, 2H), 1.29 - 1.15 (m, 26H), 1.10 (s, 3H), 0.83 (t, J = 6.6 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃/CD₃OD) δ 176.1, 160.6, 141.3, 106.5, 94.9, 80.9 (d, J = 7.5 Hz), 79.2, 73.0, 71.9, 71.1, 67.0, 63.0 (d, J = 5.9 Hz), 62.5 (d, J = 6.7 Hz), 31.8, 30.6 (d, J = 7.3 Hz), 29.61, 29.56, 29.48, 29.44, 29.3, 26.0, 22.6, 20.4, 14.0. ³¹P NMR (162 MHz, CD₃OD/CDCl₃) δ -0.5. HRMS (APCI) *m/z* calculated for C₂₉H₅₂O₉N₂PS [M - H]⁻: 635.3136, found 635.3140. Anal. calculated for C₂₉H₅₈N₃O₁₀PS (as a monohydrate): C, 51.85; H, 8.70; N, 6.25. Found: C, 51.83; H, 8.80; N, 6.60.



((2R,3R,4R,5R)-3,4-dihydroxy-4-methyl-5-(4-oxo-2-thioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)methyl icosyl hydrogen phosphate (2.13)

Procedure 3A: L = eiconasol, N = 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one.

Purification gradient: 0-55% DCM:MeOH:NH₄OH, 19% yield. ¹H NMR (399 MHz, CDCl₃/CD₃OD) δ 8.63 (s, 1H), 7.94 (d, *J* = 8.2 Hz, 1H), 6.82 (s, 1H), 5.95 (d, *J* = 8.1 Hz, 1H), 4.07 (ddd, *J* = 45.4, 11.2, 5.3 Hz, 2H), 3.94 (dd, *J* = 9.1, 1.9 Hz, 1H), 3.81 (d, *J* = 9.2 Hz, 1H), 3.78 - 3.69 (m, 2H), 1.56 - 1.45 (m, 2H), 1.27 - 1.07 (m, 38H), 0.77 (t, *J* = 6.8 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃/CD₃OD) δ 176.2, 160.6, 144.7, 141.3, 106.4, 95.0, 81.0 (d, *J* = 8.3 Hz), 79.1, 73.64, 71.7, 65.9 (d, *J* = 5.9 Hz), 31.8, 30.5 (d, *J* = 7.8 Hz), 29.53, 29.48, 29.2 (d, *J* = 5.5 Hz), 25.6, 22.5, 20.2, 13.9. ³¹P NMR (162 MHz, CDCl₃/CD₃OD) δ 0.5. HRMS (ESI) *m/z* calculated for C₃₀H₅₄O₈N₂PS [M - H]⁻: 633.3342, found 633.3344. Anal. calculated for C₃₀H₆₀N₃O₉PS (as a monohydrate): C, 53.79; H, 9.03; N, 6.27. Found: C, 53.79; H, 9.07; N, 6.33.

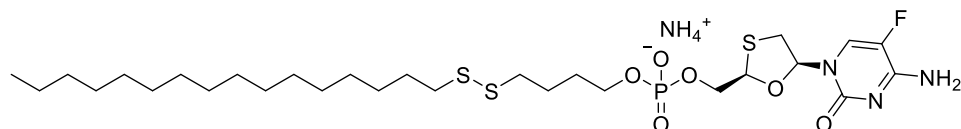


((2R,3R,4R,5R)-3,4-dihydroxy-4-methyl-5-(4-oxo-2-thioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)methyl (Z)-docos-5-en-1-yl hydrogen phosphate (2.14)

Procedure 3A: L = (Z)-docos-5-en-1-ol, N = 1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-3a-methyl-2-phenyltetrahydrofuro[3,4-d][1,3,2]dioxaborol-4-yl)-2-thioxo-2,3-dihydropyrimidin-4(1H)-one.

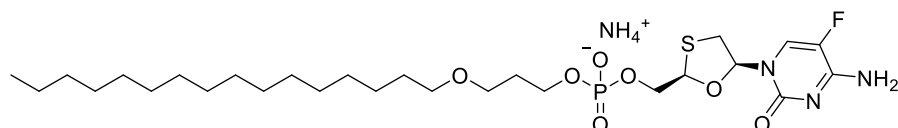
Purification gradient: 0-55% DCM:MeOH:NH₄OH, 14% yield. ¹H NMR (399 MHz, CDCl₃/CD₃OD) δ 8.73 (s, 1H), 7.97 (d, *J* = 8.1 Hz, 1H), 6.85 (s, 1H), 6.01 (d, *J* = 8.1 Hz, 1H), 5.48 - 5.37 (m, 1H), 4.94 - 4.83 (m, 2H), 4.12 (ddd, *J* = 18.1, 11.4, 5.5 Hz, 2H), 4.00 (dd, *J* = 9.2, 2.3 Hz, 1H), 3.84 (d, *J* = 9.1 Hz, 1H), 3.81 - 3.72 (m, 2H), 1.91 - 1.80 (m, 1H), 1.65 - 1.32 (m, 2H), 1.30 - 1.10 (m, 36H), 0.83 (t, *J* = 6.5 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃/CD₃OD) δ 176.2, 160.5, 142.8, 141.3, 114.4, 106.5, 95.0,

81.0 (d, $J = 8.1$ Hz), 79.2, 73.80, 73.79, 71.9, 66.2 (d, $J = 5.7$ Hz), 62.5, 44.0, 35.0, 31.8, 31.0, 29.7, 29.6, 29.6, 29.3, 28.4 (d, $J = 7.9$ Hz), 27.1, 22.6, 20.4, 14.1. ^{31}P NMR (162 MHz, $\text{CD}_3\text{OD}/\text{CDCl}_3$) δ -0.1. HRMS (ESI) m/z calculated for $\text{C}_{32}\text{H}_{56}\text{O}_8\text{N}_2\text{PS}$ $[\text{M} - \text{H}]^-$: 659.3500, found 659.3505. Anal. calculated for $\text{C}_{33}\text{H}_{64}\text{N}_5\text{O}_9\text{PS}$ (methanol solvate with ammonium salts): C, 53.71; H, 8.74; N, 9.49. Found: C, 53.66; H, 8.77; N, 9.42.



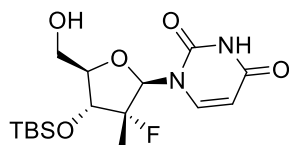
[(2R,5S)-5-(4-Amino-5-fluoro-2-oxo-pyrimidin-1-yl)-1,3-oxathiolan-2-yl]methyl 4-(hexadecylsulfanyl)butyl hydrogen phosphate (2.15)

Procedure 3A: L = 4-(hexadecylsulfaneyl)butan-1-ol, N = emtricitabine. Purification gradient: 0-46% DCM:MeOH:NH₄OH, 46% yield. ^1H NMR (399 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 8.05 (d, $J = 6.6$ Hz, 1H), 6.15 (t, $J = 4.5$ Hz, 1H), 5.28 (t, $J = 3.4$ Hz, 1H), 4.22 (dd, $J = 10.3, 5.6$ Hz, 1H), 4.08 – 4.00 (m, 1H), 3.83 (q, $J = 6.0$ Hz, 2H), 3.43 (dd, $J = 12.4, 5.4$ Hz, 1H), 3.09 (dd, $J = 12.3, 3.7$ Hz, 1H), 2.66 – 2.57 (m, 4H), 1.75 – 1.63 (m, 4H), 1.59 (p, $J = 7.4$ Hz, 2H), 1.37 – 1.12 (m, 26H), 0.81 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 157.8, 157.7, 154.4, 125.7 (d, $J = 31.4$ Hz), 87.5, 85.4 (d, $J = 8.5$ Hz), 65.4 (d, $J = 4.1$ Hz), 65.2 (d, $J = 6.1$ Hz), 38.8, 38.2, 37.9, 31.8, 29.59, 29.57, 29.55, 29.52, 29.46, 29.33, 29.25, 29.19, 29.11, 28.5, 25.3, 22.6, 14.0. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ -0.5. ^{19}F NMR (376 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ -167.0. HRMS (ESI) m/z calculated for $\text{C}_{28}\text{H}_{50}\text{FN}_3\text{O}_6\text{PS}_3$ $[\text{M} - \text{H}]^-$: 670.2593, found 670.2588. Anal. calculated for $\text{C}_{28}\text{H}_{54}\text{N}_4\text{O}_6\text{PS}_3\text{F}$: C, 48.82; H, 7.90; N, 8.13. Found: C, 48.13; H, 7.97; N, 7.67.



[(2R,5S)-5-(4-Amino-5-fluoro-2-oxo-pyrimidin-1-yl)-1,3-oxathiolan-2-yl]methyl 3-hexadecyloxypropyl hydrogen phosphate (2.16)

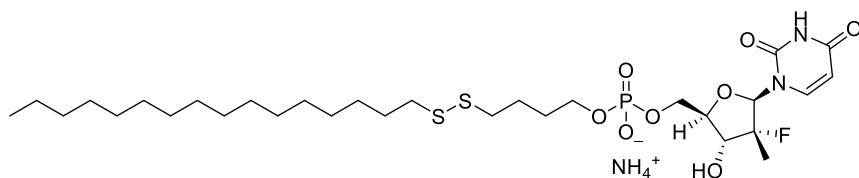
Procedure 3A: L = 3-(hexadecyloxy)propan-1-ol, N = emtricitabine. Purification gradient: 0-46% DCM:MeOH:NH₄OH, 23% yield. ¹H NMR (399 MHz, CDCl₃/CD₃OD) δ 8.78 (s, 2H), 8.07 (d, *J* = 6.6 Hz, 1H), 6.16 (t, *J* = 4.5 Hz, 1H), 5.30 (t, *J* = 3.3 Hz, 1H), 4.25 (dd, *J* = 10.4, 5.8 Hz, 1H), 4.10 – 4.00 (m, 1H), 3.91 (q, *J* = 6.4 Hz, 2H), 3.47 (t, *J* = 6.3 Hz, 2H), 3.45 – 3.41 (m, 1H), 3.35 (t, *J* = 6.9 Hz, 2H), 3.10 (dd, *J* = 12.5, 3.5 Hz, 1H), 1.85 (p, *J* = 6.4 Hz, 2H), 1.50 (p, *J* = 6.8 Hz, 2H), 1.29 – 1.18 (m, 26H), 0.84 (t, *J* = 6.8 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃/CD₃OD) δ 157.7, 157.6, 154.4, 125.7 (d, *J* = 32.7 Hz), 87.4, 85.4 (d, *J* = 8.1 Hz), 74.0, 71.1, 67.1, 65.4, 62.9 (d, *J* = 5.3 Hz), 37.9, 31.9, 30.7 (d, *J* = 7.7 Hz), 29.65, 29.62, 29.59, 29.50, 29.3, 26.0, 22.6, 14.0. ¹⁹F NMR (376 MHz, CDCl₃/CD₃OD) δ -167.0. ³¹P NMR (162 MHz, CDCl₃) δ -0.7. HRMS (ESI) *m/z* calculated for C₂₇H₄₈O₇N₃FPS [M - H]⁻: 608.2940, found 608.2939. Anal. calculated for C₂₇H₅₂N₄O₇PSF (as a hydrate with inorganic salts): C, 50.41; H, 8.38; N, 9.80. Found: C, 50.20; H, 8.32; N, 9.78.



1-((2R,3R,4R,5R)-4-((*tert*-Butyldimethylsilyloxy)-3-fluoro-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (2.17)

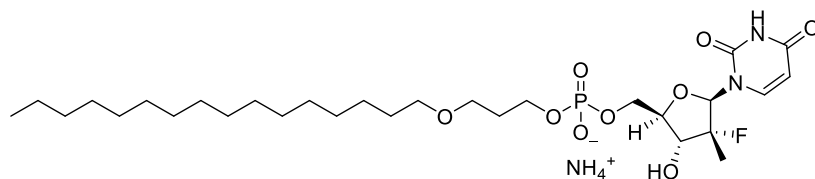
To a stirring solution of 1-((2R,3R,4R,5R)-3-fluoro-4-hydroxy-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (1.00 g, 3.84 mmol) in DMF (5 mL) was added 1H-imidazole (1.36 g, 20.0 mmol) and *tert*-butylchlorodimethylsilane (2.90 g, 19.2 mmol). The mixture stirred at room temperature overnight and was treated with 25 mL of saturated aqueous sodium bicarbonate. After stirring for 10 min, the mixture was extracted with DCM. The organic layer was collected and washed with water and brine, dried over anhydrous Na₂SO₄ and filtered. The

volatiles were removed under reduced pressure and the residue was purified via silica gel column chromatography (EtOAc:hexane = 30:70 v/v) to give 1-((2R,3R,4R,5R)-4-((*tert*-butyldimethylsilyl)oxy)-5-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-fluoro-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (1.87 g, 3.84 mmol). Subsequently, to a solution of 1-((2R,3R,4R,5R)-4-((*tert*-butyldimethylsilyl)oxy)-5-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-fluoro-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (1.87 g, 3.84 mmol) in 13 mL of THF was added, at 0 °C, a solution of 2,2,2-trichloroacetic acid (6.16 ml, 61.5 mmol) in 4 mL of water. After stirring for 7 h at 0 °C, the solution was neutralized with solid NaHCO₃ till gas evolution ceased. 40 mL of water were then added and the solution was extracted with CH₂Cl₂ (3 x 20 mL). The combined organic layers were dried over Na₂SO₄, filtered and the volatiles were concentrated under reduced pressure. The residue was purified by silica gel column chromatography (EtOAc:hexane, 0-70% EtOAc) to give the title compound 1-((2R,3R,4R,5R)-4-((*tert*-butyldimethylsilyl)oxy)-3-fluoro-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (1.25 g, 3.34 mmol, 87 % yield). Spectral data is in accord with literature precedent. ¹H NMR (400 MHz, CDCl₃) δ 8.38 (s, 1H), 7.85 (d, *J* = 7.8 Hz, 1H), 6.11 (d, *J* = 18.0 Hz, 1H), 5.76 (dd, *J* = 8.2, 2.3 Hz, 1H), 4.13 (dd, *J* = 24.7, 11.3 Hz, 1H), 4.05 (dd, *J* = 9.1, 1.6 Hz, 1H), 3.84 (d, *J* = 12.3 Hz, 1H), 1.36 (d, *J* = 21.9 Hz, 3H), 0.94 (s, 9H), 0.15 (s, 3H), 0.15 (s, 3H). HRMS (ESI) *m/z* calculated for C₁₆H₂₈O₅N₂FSi [M + H]⁺ : 375.1746, found 375.1742.



((2R,3R,4R,5R)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-4-fluoro-3-hydroxy-4-methyltetrahydrofuran-2-yl)methyl (4-(hexadecyldisulfaneyl)butyl) hydrogen phosphate (ammonium salt) (2.18)

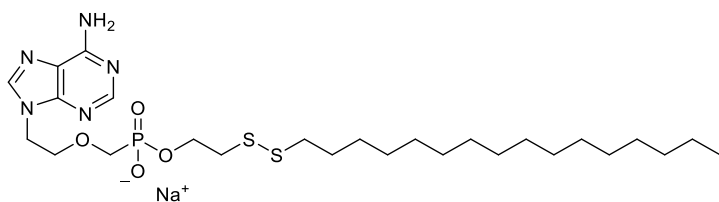
Procedure 3A: L = 4-(hexadecyldisulfaneyl)butan-1-ol, N = compound **2.2**: ^1H NMR (399 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 7.86 (d, $J = 8.2$ Hz, 1H), 6.13 (d, $J = 18.4$ Hz, 1H), 5.75 (d, $J = 8.1$ Hz, 1H), 4.22 – 4.19 (m, 1H), 4.07 (dd, $J = 12.3, 6.2$ Hz, 1H), 4.03 – 3.98 (m, 1H), 3.92 (dd, $J = 24.1, 9.4$ Hz, 1H), 3.86 – 3.78 (m, 2H), 2.66 – 2.58 (m, 4H), 1.77 – 1.64 (m, 4H), 1.63 – 1.54 (m, 2H), 1.32 (d, $J = 22.2$ Hz, 3H), 1.25 – 1.16 (m, 28H), 0.81 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 164.2, 150.8, 140.1, 102.5, 100.6 (d, $J = 182.0$ Hz), 80.5 (d, $J = 7.3$ Hz), 77.4, 70.5 (d, $J = 19.0$ Hz), 65.2 (d, $J = 5.7$ Hz), 61.9 (d, $J = 5.0$ Hz), 38.8, 38.2, 31.8, 29.56, 29.53, 29.49, 29.42, 29.34, 29.27, 29.24, 29.1 (d, $J = 8.0$ Hz), 28.4, 25.3, 22.6, 16.1 (d, $J = 25.5$ Hz), 13.9. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 0.9. ^{19}F NMR (376 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ -163.3. HRMS (ESI) m/z calculated for $\text{C}_{30}\text{H}_{55}\text{O}_8\text{N}_2\text{FPS}_2$ [M - H] $^-$: 683.2970, found 683.2970.



((2R,3R,4R,5R)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-4-fluoro-3-hydroxy-4-methyltetrahydrofuran-2-yl)methyl dihydrogen phosphate (ammonium salt) (2.19)

Procedure 3A: L = hexadecyloxypropanol, N = compound **2.2**: ^1H NMR (399 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 7.85 (d, $J = 8.2$ Hz, 1H), 6.12 (d, $J = 18.5$ Hz, 1H), 5.74 (d, $J = 8.1$ Hz, 1H), 4.19 (ddd, $J = 12.1, 4.5, 1.9$ Hz, 1H), 4.10 – 4.03 (m, 2H), 4.00 (dt, $J = 9.5, 1.9$ Hz, 1H), 3.92 (dd, $J = 24.2, 9.5$ Hz, 1H), 3.89 – 3.84 (m, 2H), 3.47 (t, $J = 6.4$ Hz, 2H), 3.34 (t, $J = 6.9$ Hz, 2H), 1.83 (p, $J = 6.4$ Hz, 2H), 1.48 (p, $J = 6.7$ Hz, 2H), 1.32 (d, $J = 22.2$ Hz, 3H), 1.26 – 1.16 (m, 26H), 0.81 (t, $J = 6.7$ Hz, 3H). ^{13}C NMR (101 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 164.2, 150.7, 140.1, 102.4, 100.6 (d, $J = 182.2$ Hz), 88.9 (d, $J = 42.8$ Hz), 80.4 (d, $J = 7.5$ Hz), 71.1, 70.4 (d, $J = 17.5$ Hz), 67.1, 62.8 (d, $J = 4.7$ Hz), 61.9, 31.8, 30.6 (d, $J = 7.2$ Hz), 29.6, 29.4, 29.3, 26.0, 22.6, 16.1 (d, $J = 25.4$ Hz), 13.9. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ

0.4. ^{19}F NMR (376 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ -159.4. HRMS (ESI) m/z calculated for $\text{C}_{29}\text{H}_{51}\text{O}_9\text{N}_2\text{FP}$ $[\text{M} - \text{H}]^-$: 621.3321, found 621.3321.

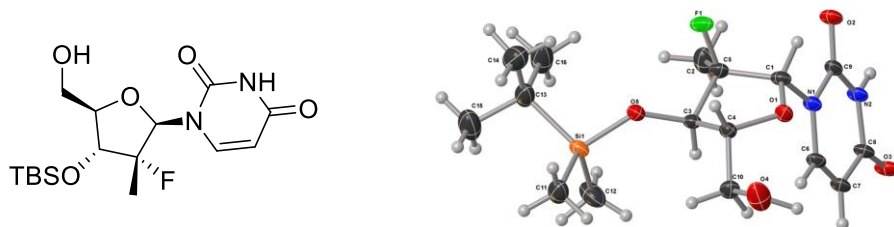


2-(Hexadecylsulfinothioyl)ethyl ((2-(6-amino-9*H*-purin-9-yl)ethoxy)methyl)phosphonic acid (2.20)

To a stirring solution of dry adefovir (0.10 g, 0.37 mmol) in anhydrous DCM (6 mL) and *N,N*-dimethylformamide (0.034 mL, 0.439 mmol) was gradually added excess oxalyl chloride (0.157 mL, 1.830 mmol). The mixture stirred at r.t (open to air) for 15 min until complete dissolution of starting material was observed. The solvent and excess oxalyl chloride were evaporated under reduced pressure to afford a pale yellow foam which was re-dissolved in anhydrous DCM (6 mL) to afford a clear colorless solution that was chilled to 0 °C. Then a mixture of 2-(hexadecylsulfanyl)ethanol (0.122 g, 0.366 mmol) and pyridine (0.177 mL, 2.196 mmol) in anhydrous DCM (2 mL) was slowly added, dropwise. The mixture stirred at this temperature for 15 min, then naturally warmed to room temperature and stirred for 3 h. Then, water (0.1 mL, 5.5 mmol) was added and the mixture continued stirring for an additional 30 min. The solvent was evaporated under reduced pressure and the resulting residue was dried under UHV. The residue was re-dissolved in 190 proof EtOH (5 mL) and stirred at 40°C overnight. The resulting milky reaction mixture was diluted with water (3 mL) and filtered. The solid was washed with cold water (2 mL), MeOH (1 mL), and DCM (2 mL) in that order and subsequently dried under UHV to afford the title compound 2-(hexadecylsulfinothioyl)ethyl ((2-(6-amino-9*H*-purin-9-yl)ethoxy)methyl)phosphonic acid (84 mg, 0.137 mmol, 37 % yield) as a white solid that required no further purification. ^1H NMR (400 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$, referenced to CDCl_3) δ 8.17

(s, 1H), 7.89 (s, 1H), 4.33 - 4.24 (m, 2H), 4.02 (q, $J = 6.7$ Hz, 2H), 3.96 - 3.85 (m, 2H), 3.66 (d, $J = 7.7$ Hz, 2H), 2.73 (t, $J = 6.5$ Hz, 2H), 2.57 - 2.48 (m, 2H), 1.49 (dt, $J = 14.8, 7.2$ Hz, 2H), 1.25 - 1.04 (m, 26H), 0.71 (t, $J = 6.9$ Hz, 3H). ^{13}C NMR (100 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$, referenced to CD_3OD) δ 150.0, 149.0, 145.1, 144.2, 117.4, 70.8 (d, $J = 11.8$ Hz), 65.4 (d, $J = 152.1$ Hz), 64.1 (d, $J = 6.9$ Hz), 44.6, 39.5, 39.1 (d, $J = 5.7$ Hz), 32.1, 30.0 (3), 29.9 (2), 29.8, 29.74, 29.71, 29.6, 29.44, 29.35, 28.7, 22.9, 14.2. ^{31}P NMR (162 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$) δ 16.63. HRMS (ESI) m/z calculated for $\text{C}_{26}\text{H}_{49}\text{O}_4\text{N}_5\text{PS}_2$ $[\text{M} + \text{H}]^+$: 590.2958, found 590.2965. Anal. calculated for $\text{C}_{26}\text{H}_{47}\text{N}_5\text{O}_4\text{PS}_2\text{Na}$ (sodium salt): C, 51.04; H, 7.74; N, 11.45. Found: C, 50.55; H, 7.74; N, 11.79. Melting Point: Decomposes at 165°C .

2.4.6 Crystallographic Data



1-((2R,3R,4R,5R)-4-((*tert*-Butyldimethylsilyl)oxy)-3-fluoro-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (2.17)

Single crystals of compound **2.17** were recrystallized from DCM. A suitable crystal was selected and mounted on a loop with paratone oil on a XtaLAB Synergy, Dualflex, HyPix diffractometer. The crystal was kept at 100(2) K during data collection. Using Olex2,⁹⁸ the structure was solved with the ShelXS⁹⁹ structure solution program using Direct Methods and refined with the XL⁹⁹ refinement package using Least Squares minimisation. Crystal Data for C₁₆H₂₇FN₂O₅Si (*M* = 374.48 g/mol): orthorhombic, space group P2₁2₁2₁ (no. 19), *a* = 6.1767(3) Å, *b* = 6.9222(2) Å, *c* = 44.0327(17) Å, *V* = 1882.67(13) Å³, *Z* = 4, *T* = 100(2) K, $\mu(\text{CuK}\alpha) = 1.443 \text{ mm}^{-1}$, *D*_{calc} = 1.321 g/cm³, 5890 reflections measured (8.032° ≤ 2 Θ ≤ 130.174°), 2851 unique (*R*_{int} = 0.0506, *R*_{sigma} = 0.0524) which were used in all calculations. The final *R*₁ was 0.0573 (*I* > 2 σ (*I*)) and *wR*₂ was 0.1595 (all data). Additional crystallographic data is shown in Table 2.2.

Table 2.2. Crystal Data and Structure Refinement for 1-((2R,3R,4R,5R)-4-((*tert*-Butyldimethylsilyl)oxy)-3-fluoro-5-(hydroxymethyl)-3-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (**2.17**)

Empirical formula	C ₁₆ H ₂₇ FN ₂ O ₅ Si
Formula weight	374.48
Temperature/K	100(2)
Crystal system	orthorhombic
Space group	P2 ₁ 2 ₁ 2 ₁
a/Å	6.1767(3)
b/Å	6.9222(2)
c/Å	44.0327(17)
α/°	90
β/°	90
γ/°	90
Volume/Å ³	1882.67(13)
Z	4
ρ _{calc} /cm ³	1.321
μ/mm ⁻¹	1.443
F(000)	800.0
Crystal size/mm ³	0.326 × 0.171 × 0.072
Radiation	CuKα (λ = 1.54184)
2θ range for data collection/°	8.032 to 130.174
Index ranges	-2 ≤ h ≤ 7, -8 ≤ k ≤ 7, -51 ≤ l ≤ 40
Reflections collected	5890
Independent reflections	2851 [R _{int} = 0.0506, R _{sigma} = 0.0524]
Data/restraints/parameters	2851/204/233
Goodness-of-fit on F ²	1.136
Final R indexes [I ≥ 2σ (I)]	R ₁ = 0.0573, wR ₂ = 0.1466
Final R indexes [all data]	R ₁ = 0.0649, wR ₂ = 0.1595
Largest diff. peak/hole / e Å ⁻³	0.41/-0.62
Flack parameter	0.00(3)

2.5 References

- (92) Sinha, N. D.; Biernat, J.; McManus, J., *et al.*, Polymer Support Oligonucleotide Synthesis Xviii.1.2): Use of B-Cyanoethyl-N,N-Dialkylamino-/N-Morpholino Phosphoramidite of Deoxynucleosides for the Synthesis of DNA Fragments Simplifying Deprotection and Isolation of the Final Product. *Nucleic Acids Res.* **1984**, *12* (11), 4539-4557.
- (93) Beaucage, S. L.; Iyer, R. P., The Synthesis of Modified Oligonucleotides by the Phosphoramidite Approach and Their Applications. *Tetrahedron* **1993**, *49* (28), 6123-6194.
- (94) Tebby, J. C., *Crc Handbook of Phosphorus-31 Nuclear Magnetic Resonance Data*. CRC Press: Boca Raton, 1991; p 573 p.
- (95) Liotta, D. C.; Painter, G. R.; Bluemling, G. R., *et al.* Preparation of Amino Acid-Containing Nucleotide and Nucleoside Triphosphates as Antiviral Agents. WO2015038596A1, 2015.
- (96) Mayes, B. A.; Arumugasamy, J.; Baloglu, E., *et al.*, Synthesis of a Nucleoside Phosphoramidate Prodrug Inhibitor of Hcv Ns5b Polymerase: Phenylboronate as a Transient Protecting Group. *Organic Process Research & Development* **2014**, *18* (6), 717-724.
- (97) Wang, L. H.; Begley, J.; St Claire, R. L., 3rd, *et al.*, Pharmacokinetic and Pharmacodynamic Characteristics of Emtricitabine Support Its Once Daily Dosing for the Treatment of HIV Infection. *AIDS Res. Hum. Retroviruses* **2004**, *20* (11), 1173-82.
- (98) Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J., *et al.*, Olex2: A Complete Structure Solution, Refinement and Analysis Program. *J. Appl. Crystallogr.* **2009**, *42* (2), 339-341.
- (99) Sheldrick, G. M., Crystal Structure Refinement with Shelxl. *Acta Crystallographica. Section C, Structural Chemistry* **2015**, *71* (Pt 1), 3-8.

Chapter 3

A Green and Expedient Synthesis of Acyclic Thioaminal Nucleoside Analogues from Purine and Pyrimidine Hemiaminals

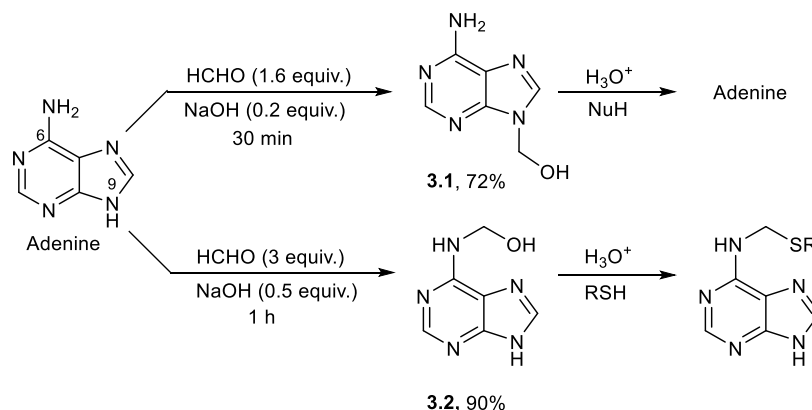
Abstract: Nucleoside analogues are an important class of drugs for the treatment of acute and chronic viral infections. However, their synthesis is generally plagued by the use of multiple protecting groups, toxic solvents, and low yields which collectively hinder the discovery of novel therapeutics. Herein, we couple purine and pyrimidine hemiaminals with thiols in the presence of other functional groups to furnish a series of hydrolytically stable thioaminal nucleoside analogues in good yield without the use of organic solvents or protecting groups. The desired nucleosides cleanly precipitate from the reaction mixture after 2 h to generate a library of lead compounds that may be further optimized and developed into potent antiviral agents.

3.1 Introduction

Acyclic nucleoside analogues feature a flexible linker in place of a rigid ribofuranose and demonstrate potent broad spectrum antiviral activity.¹⁰⁰ Acyclovir, cidofovir, tenofovir, and adefovir are notable examples that have achieved renowned clinical success for the treatment of HSV,¹⁶ CMV,¹⁸ HIV,¹⁰¹ and HBV,¹⁴ respectively. Such compounds are generally endowed with limited stereochemical information which obviates the need for successive stereoselective transformations. These features inspired us to pursue novel acyclic nucleoside scaffolds that may demonstrate promising anti-HIV, anti-HBV, and anti-HSV activity. Herein, we disclose a convenient two step protocol that exploits the selective reactivity of purine and pyrimidine hemiaminals towards thiols to afford acyclic thioaminal nucleoside analogues in good yield without the use of organic solvents and protecting groups. The biological activity of certain analogues against HIV-1 and HBV is also described along with crystallographic data that provides direct evidence of an anomeric effect at the N–C–S linkage.

3.2 Synthesis and Substrate Scope

Our initial work began with adenine which undergoes hydroxymethylation with formaldehyde to yield hemiaminals (Scheme 3.1).¹⁰²⁻¹⁰⁵ Exposure of adenine to a slight excess of formaldehyde (1.6 equiv.) and catalytic NaOH in water over the course of 30 min furnished **3.1** which required no further purification beyond filtration after neutralization with ammonium chloride. Hydroxymethylation at N-9 was confirmed by noting the presence of both amino protons and the absence of the characteristic N-9 proton at 13.0 δ in the ¹H NMR spectrum. Unfortunately, **3.1** was exceptionally labile under a variety of conditions and not amenable to functionalization which thwarted our efforts to advance this material forward.

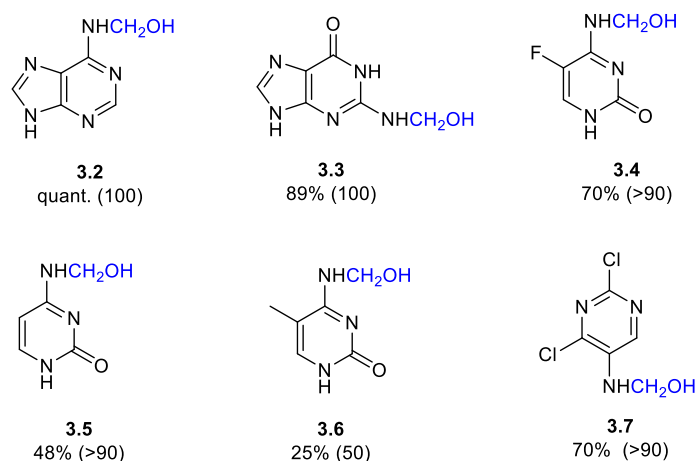
Scheme 3.1. Synthesis and Reactivity of Hemiaminals **3.1** and **3.2**.

When additional NaOH (0.5 equiv.) and formaldehyde (3.0 equiv.) stirred with aqueous adenine for 1 h, a second distinct hemiaminal precipitated from the mixture following NH_4Cl neutralization. This species was identified as compound **3.2** by ^1H NMR and LC-MS. In contrast to **3.1**, compound **3.2** was exceptionally robust and reacted with thiols in dilute acid ($\text{pH} = 0$) at room temperature to furnish thioaminal ethers that were resistant to protonolysis. Interestingly, these same set of conditions failed to produce stable hemiaminal ethers when alcohols (MeOH, EtOH, iPrOH, $\text{HOCH}_2\text{CH}_2\text{OH}$) were used in lieu of a thiol. The distinct reactivity difference between **3.1** and **3.2** can be traced to the nucleofugacity of their respective anions. In the presence of acid, hemiaminal **3.1** experiences protonation at the imidazole nitrogen which poises the entire heterocycle to leave as a neutral species when the methylene carbon is attacked by a nucleophile, Nu. This pathway is not accessible for compound **3.2** and is therefore resistant to displacement of the heterocycle.

From these preliminary findings, it became clear that acyclic thioaminal scaffolds could be rapidly and selectively accessed by simply mixing the requisite purine hemiaminal with a desired thiol in dilute sulfuric acid at room temperature. We therefore chose to expand our repertoire of hemiaminal precursors beyond adenine with the intent to prepare a library of thioaminal ethers from a handful of purine and pyrimidine nucleobases. These adducts are shown in Table 3.1. Hydroxymethylation

proceeded smoothly for all entries in Table 3.1 except 5-methylcytosine whose conversion to **3.6** was slow and failed to reach completion after several hours suggesting that the reaction is sensitive to the steric environment of the amino group. Importantly, no detectable conversion was observed for hypoxanthine or other heterocycles that lacked an exocyclic amine. Isolated yields were largely governed by the precipitation of product from the reaction mixture and for this reason purines consistently produced higher yields than pyrimidines.

Table 3.1. Synthesis of Hemiaminals from Purines and Pyrimidines^{a,b,c}



^aYields shown are isolated yields.

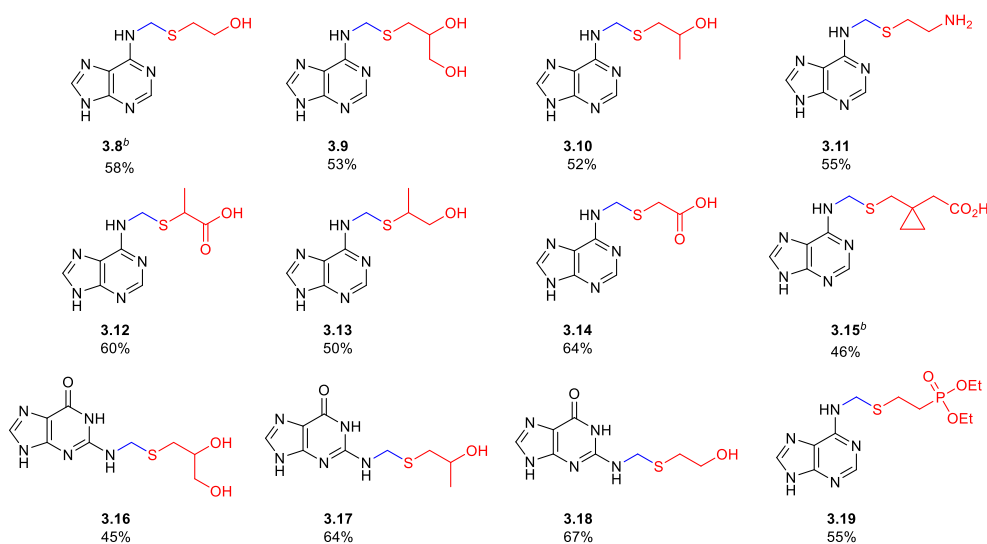
^bValues in parenthesis are yields determined by LC-MS.

^cConditions: The corresponding purine or pyrimidine (1.0 equiv.) was reacted with HCHO (3.0 equiv.) and aqueous NaOH (0.5 equiv.) for 1 h at room temperature.

With a variety of hemiaminals in hand, we then sought to react these species with a series of thiols in aqueous acid to build a novel library of acyclic thioaminal nucleoside analogues. The presence of a distal –OH was a critical design consideration for the coupling thiol as this would be the designated site for cellular phosphorylation and conversion to the active triphosphate. In this respect, it is fortuitous that alcohols failed to produce stable hemiaminal ethers in aqueous acid which allowed us to circumvent the use of protecting groups.

After some minor optimization, we arrived at the ideal coupling protocol for the preparation of purine thioaminals. Best results were obtained when hemiaminal **3.2** or **3.3** (1.0 equiv.) stirred with the target thiol (1.2 equiv.) in aqueous acid (pH = 0) at room temperature for 2 hours. The reaction was insensitive to the identity of the acid (HCl, H₂SO₄), exposure to atmospheric oxygen, concentration, and the presence of alcoholic media (MeOH, EtOH), however, no reaction occurred when the pH exceeded 5. Table 3.2 illustrates the substrate scope with various thiols. It is apparent that secondary thiols are equally reactive as primary thiols and the presence of other functional groups (amines, carboxylic acids, alcohols) do not suppress or retard thioaminal formation. Efforts to consolidate hydroxymethylation and thioaminal formation into a one-pot procedure were not successful and resulted in the formation of byproducts. In addition, it was far more convenient to prepare the bench stable hemiaminal in bulk and introduce diversity in a second step. Many of the products shown in Table 3.2 were isolated in > 85% purity after precipitation from the reaction

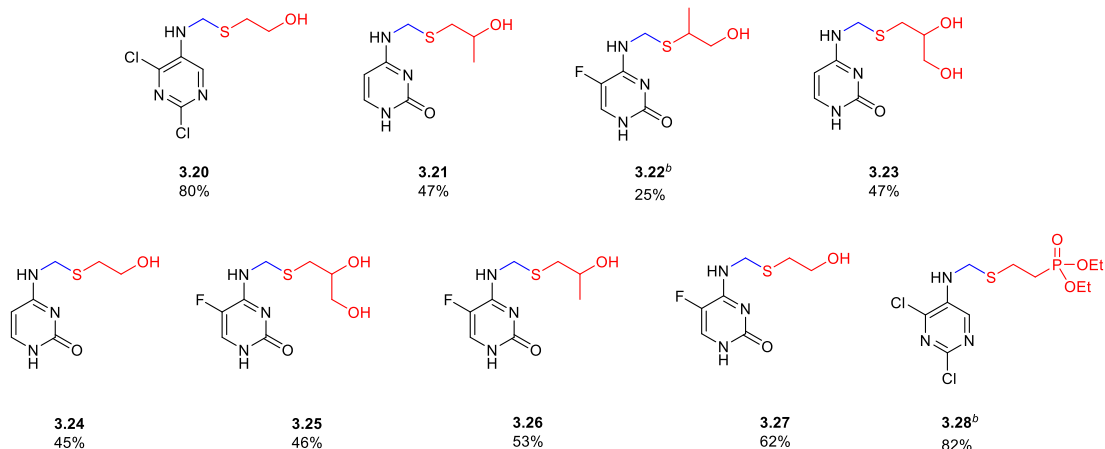
Table 3.2. Prepared Acyclic Purine Thioaminal Nucleoside Analogues^{a,c}



^aYields shown are isolated yields.

^bCrystal structure available.

^cConditions: Purine hemiaminal (1.0 equiv.) was reacted with the desired thiol (1.2 equiv.) in dilute sulfuric acid (pH = 0) for 2 h at room temperature.

Table 3.3. Prepared Acyclic Pyrimidine Thioaminal Nucleoside Analogues^{a,c}^aYields shown are isolated yields.^bCrystal structure available^cConditions: Pyrimidine hemiaminal (1.0 equiv.) was reacted with the desired thiol (1.2 equiv.) in concentrated phosphoric acid at room temperature for 7 h.

mixture, however, several compounds were further purified via reverse phase column chromatography to assess their biological activity.

In contrast to **3.2** and **3.3**, pyrimidine hemiaminals **3.4-3.7** failed to produce thioaminal ethers in aqueous acid. It was later discovered that this transformation could be effected by concentrated phosphoric acid, however, longer reaction times (7 h) were required. These conditions are reminiscent of the dehydrative conditions necessary for conventional acetal formation¹⁰⁶ and the judicious choice of H₃PO₄ circumvents the need for organic solvents and acids. The thiols used to generate purine thioaminals **3.8-3.10**, **3.13**, **3.16**, and **3.17** were also exploited to produce pyrimidine thioaminals **3.19-3.26** in Table 3.3. Despite the use of a relatively weak acid and a longer reaction time, the yields presented in Table 3.3 are comparable to those shown in Table 3.2.

Unambiguous assignment of the N-CH₂-S linkage was provided by suitable X-ray quality crystals of **3.8**, **3.15**, **3.22**, and **3.28**. Figures 3.1a and 3.1b show the X-ray structure of **3.8** and **3.22**, respectively. Both structures feature a prominent anomeric effect with the antibonding σ^* C-N

orbital aligned antiperiplanar to one of the donor lone pairs on sulfur. In addition, **3.22** also features a weak electrostatic interaction between the polar O—H hydrogen and the pyrimidine π cloud to produce the observed conformation in the solid state. Such stereoelectronic and electrostatic effects are ubiquitous throughout nature in sugars, nucleosides, and natural products and may be successfully exploited to control the conformational flexibility of more complex thioaminal derivatives. To the best of our knowledge, this is the first crystal structure deposited in the literature that features an acyclic thioaminal.

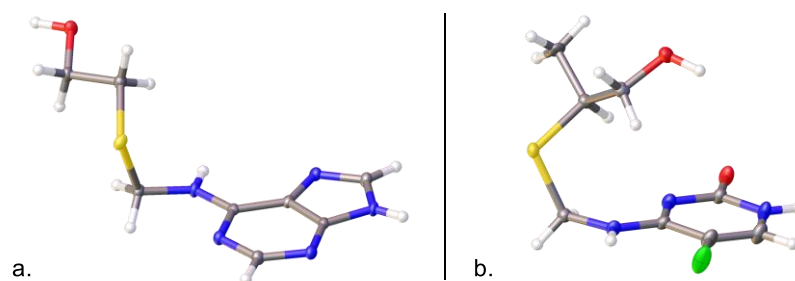


Figure 3.1. Crystal structure of a) Purine **3.8** and b) Pyrimidine **3.22**.

3.3 Biological Evaluation

To characterize the biological activity of our nucleoside analogues, compounds **3.8-3.10**, **3.17-3.18**, **3.19** and **3.25-3.27** were evaluated against HIV-1, HBV, and HSV-1 for potential antiviral activity. These results are shown in Table 3.4. Purines **3.8**, **3.9**, **3.17**, and **3.18** exhibited micromolar activity against HIV-1 whereas all tested compounds were inactive up to 100 μM against HBV and HSV-1. Compound **3.9** emerged as the most potent anti-HIV agent in this series with an $\text{EC}_{50} = 22 \mu\text{M}$, however, the unremarkable activity of these compounds and the closeness of their EC_{50} values to the 100 μM cutoff makes it difficult to identify convincing structure-activity relationships. It is not unusual for biological data to show two or four-fold differences in activity. Nonetheless, it is clear that these compounds are poor anti-HIV, anti-HBV, and anti-HSV agents.

Table 3.4. Antiviral Activity of Select Thioaminals against HIV-1, HBV, and HSV-1^a

Cmpd	HIV-1	HBV	HSV-1
	EC ₅₀ ^b (PBMCs)	EC ₅₀ ^b (HepG2)	EC ₅₀ ^b (Vero)
3.8	44	> 100	-
3.9	22	> 100	-
3.10	> 100	> 100	
3.17	46	> 100	> 100
3.18	52	> 100	> 100
3.19	> 100	> 100	-
3.25	> 100	> 100	> 100
3.26	> 100	> 100	> 100
3.27	> 100	> 100	-

^aAll data represent an average of triplicate experiments.

^bEC₅₀, effective concentration (in μ M) required to inhibit HIV-1, HBV, or HSV-1 by 50%.

Weak antiviral activity may be due to poor kinase phosphorylation, poor incorporation of the triphosphate by the viral polymerase, both, or may be due to another mechanism entirely. In an attempt to significantly improve activity, diethyl phosphonates **3.19** and **3.28** were prepared to overcome dNK phosphorylation. Unfortunately, **3.19** was inactive against HIV-1 at 100 μ M and discouraged us from pursuing these compounds further. As a result, **3.28** was not submitted for analysis.

3.4 The Pursuit for More Interesting Analogues

Commercial availability of the purine and pyrimidine precursors was the main driver in the preparation of the compounds shown in Tables 3.2 and 3.3. The relative inactivity of these compounds prompted us to exploit our thioaminal methodology toward the synthesis of potentially interesting thioaminal acyclic NAs. The broad spectrum antiviral activity of adefovir (AFV) and tenofovir (TFV) inspired us to design analogues **3.29** and **3.30** shown in Figure 3.2. Key structural differences between our scaffold and AFV and TFV are shaded in color.

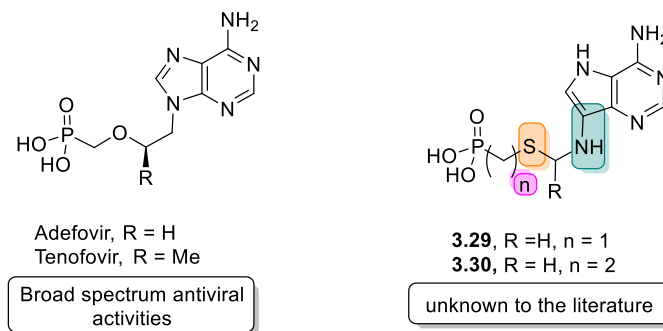
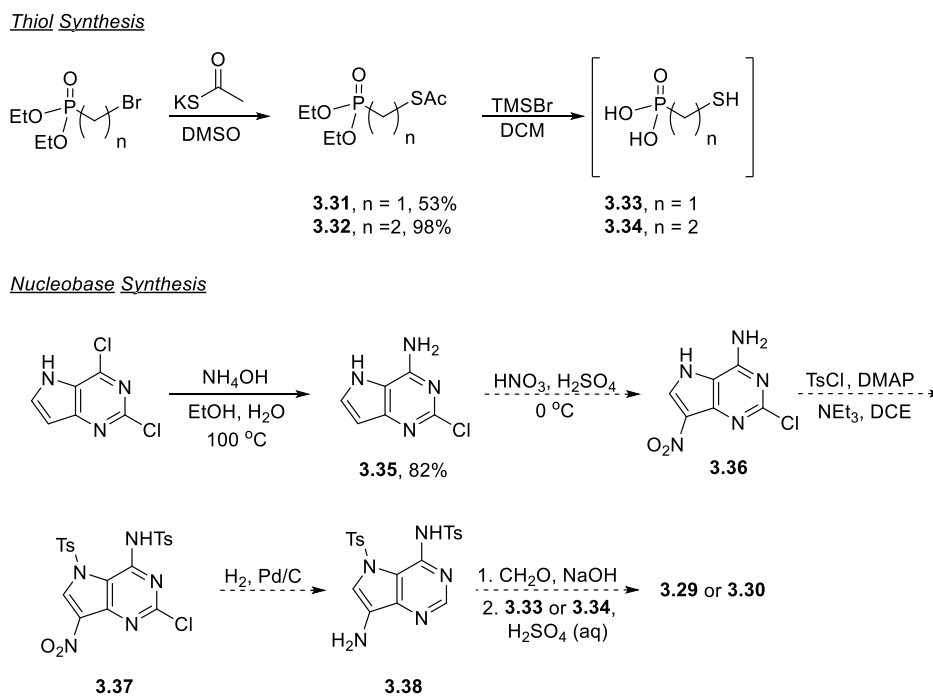


Figure 3.2. Structural features from TFV and AFV were adapted in the design of novel thioaminal phosphonates **3.29** and **3.30**.

To determine if these modifications are sufficient to preserve antiviral activity, compounds **3.29** and **3.30** were pursued as initial targets. Our synthesis towards **3.29** and **3.30** is presented in Scheme 3.2.

Scheme 3.2. Synthesis of Phosphonates **3.29** and **3.30**.



We began our synthesis with the preparation of thiols **3.33** and **3.34**. Diethyl (bromomethyl)phosphonate and diethyl (bromoethyl)phosphonate were mixed with potassium thioacetate in DMSO to yield thioacetates **3.31** and **3.32**, respectively. Subsequent treatment with excess TMSBr removed both ethyl protecting groups and concomitantly cleaved the acetate to furnish **3.33** and **3.34** which were used crude after the removal of residual TMSBr and volatile byproducts under ultra-high vacuum.

With respect to the nucleobase synthesis and subsequent coupling reaction, 2,4-dichloro-5H-pyrrolo[3,2-d]pyrimidine was aminated with concentrated ammonium hydroxide at elevated temperatures to generate **3.35**. The brown solid is currently being nitrated with HNO₃/H₂SO₄ to afford **3.36**. Tosylation of **3.36** with TsCl and DMAP furnishes **3.37** and subsequent reductive dehalogenation is expected to yield **3.38**. Hydroxymethylation, acid-catalyzed deprotection, and coupling to thiols **3.33** and **3.34** will hopefully produce desired targets **3.29** and **3.30**. This synthesis is currently in progress and the biological activity of **3.29** and **3.30** will be evaluated in due course.

3.5 Conclusions

We have successfully exploited the reactivity of purine and pyrimidine hemiaminals with thiols to access hydrolytically stable thioaminal nucleoside analogues without the use of organic solvents or protecting groups. The N–C–S linkage is readily forged in either aqueous acid or concentrated H₃PO₄ depending on the nature of the hemiaminal. An anomeric effect is clearly evident in the crystal structures of **3.8**, **3.15**, **3.22**, and **3.28** which, to the best of our knowledge, are the first known acyclic thioaminals deposited in the crystallographic literature. Few of the prepared compounds demonstrate micromolar antiviral activity against HIV-1, however, additional targets are currently underway that closely resemble adefovir and tenofovir. It is hoped that these compounds will boast broad spectrum antiviral activity and demonstrate sub-micromolar potency against HIV-1 and HBV.

3.6 Experimental Details

3.6.1 Anti-HIV Assay

All HIV assays were performed by ImQuest BioSciences (Frederick, MD). All compounds were solubilized at 40 mM in DMSO and stored at -20°C. Test materials were evaluated up to 100 μ M, and five serial logarithmic dilutions. AZT and 3TC were obtained from the NIH AIDS Research and Reference Reagent Program and used as controls in the anti-HIV and anti-HBV assay, respectively. Fresh human PBMCs were obtained from a commercial source and determined to be HIV and HBV negative. The leukophoresed blood cells were washed repeatedly with PBS, then diluted 1:1 with Dulbecco's phosphate buffered saline (PBS) and layered over 15 ml of Ficoll-Hypaque density gradient in a 50 ml conical centrifuge tube. The tubes were centrifuged for 30 min at 600 g. Banded PBMCs were gently aspirated from the resulting interface and washed three times with PBS. After the final wash, cell number was determined by Trypan Blue dye exclusion and cells re-suspended at 1×10^6 cells/mL in RPMI 1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 μ g/mL PHA-P, 100 U/ml penicillin and 100 μ g/mL streptomycin and allowed to incubate for 48-72 hr at 37°C. After incubation, PBMCs were centrifuged and re-suspended in tissue culture medium (RPMI 1640 with 15% Fetal Bovine Serum (FBS), 2 mmol/L L-glutamine, 2 μ g/mL PHA-P, 100 U/ml penicillin and 100 μ g/mL streptomycin, 3.6 ng/mL recombinant human IL-2). The cultures were maintained until use by half-volume culture changes with fresh IL-2 containing tissue culture medium every 3 days. Assays were initiated with PBMCs at 72 hr post PHA-P stimulation. Immediately prior to use, target cells were re-suspended in fresh tissue culture medium at 1×10^6 cells/ml and plated in the interior wells of a 96-well round bottom microliter plate at 50 μ L/well. Then, 100 μ L of 2X concentrations of compound-containing medium was transferred to the 96-well plate containing the cells in 50 μ L of the medium. Following addition of test compound to the wells, 50 μ L of a predetermined dilution of HIV virus (prepared at 4x of final desired in-well concentration) was added,

and mixed well. For infection, 50-150 TCID₅₀ of each virus was added per well (final MOI approximately 0.002). PBMCs were exposed in triplicate to virus and cultured in the presence or absence of the test compound at varying concentrations as described for the 96-well microliter plates. After 7 days, HIV-1 replication was quantified in the tissue culture supernatant by measurement of reverse transcriptase activity. Wells with cells and virus only served as virus controls. Separate plates were identically prepared without virus for cytotoxicity studies. Reverse transcriptase activity was measured in cell-free supernatants using a standard radioactive incorporation polymerization assay.

3.6.2 Anti-HBV Assay

All HBV assays were performed by ImQuest BioSciences (Frederick, MD). One hundred microliters (100 μ L) of wells of a 96-well flat-bottom plate at a density of 1×10^4 cells per well were incubated at 37°C in 5% CO₂ for 24 h. Following incubation, six ten-fold serial dilutions of test compound prepared in RPMI1640 medium with 10% fetal bovine serum were added to individual wells of the plate in triplicate. Six wells in the plate received medium alone as a virus control only. The plate was incubated for 6 days at 37°C at 5% CO₂. The culture medium was changed on day 3 with medium containing the indicated concentration of each compound. One hundred microliters (100 μ L) of supernatant was collected from each well for analysis of viral DNA by qPCR and cytotoxicity was evaluated by XTT staining of the cell culture monolayer on the sixth day (see below). Ten microliters (10 μ L) of cell culture supernatant collected on the sixth day was diluted in qPCR dilution buffer (40 μ g/ml sheared salmon sperm DNA) and boiled for 15 minutes. Quantitative real time PCR was performed in 386 well plates using an Applied Biosystems 7900HT Sequence Detection System and the supporting SDS 2.4 software. Five microliters (5 μ L) of boiled DNA for each sample and serial 10-fold dilutions of a quantitative DNA standard were subjected to real time Q-PCR using Platinum Quantitative PCR SuperMix-UDG (Invitrogen) and specific DNA oligonucleotide primers (IDT, Coralville, ID) HBV-AD38-qF1 (5'-CCG TCT GTG CCT TCT CAT CTG-3'), HBV-AD38-qR1 (5'-AGT CCA AGA

GTY CTC TTA TRY AAG ACC TT-3') and HBV-AD38-qP1 (5'-FAM-CCG TGT GCA /ZEN/CTT GCG TTC ACC TCT GC-3'BHQ1) at a final concentration of 0.2 μ M for each primer in a total reaction volume of 15 μ L. The final HBV DNA copy number in each sample was interpolated from the standard curve by the SDS2.4 software and the data were analyzed by Excel.

3.6.3 Anti-HSV Assay

All HSV-1 assays were performed by ImQuest BioSciences (Frederick, MD). No additional details were available at the time of this writing.

3.6.4 Cytotoxicity Studies

Cytotoxicity was evaluated by staining uninfected cells with tetrazolium dye XXT with spectrophotometric readings at 450/650 nm with a Molecular Devices Vmax plate reader. These experiments were run in triplicate for each compound tested.

3.6.5 Chemical Synthesis and General Procedures

All reagents were obtained from commercial suppliers and used without further purification unless otherwise specified. Reaction progress was monitored by either thin layer chromatography (TLC) using pre-coated aluminum-backed silica gel plates (60 F₂₅₄ Merk, article 5554), by NMR, or liquid chromatography-mass spectrometry (LC-MS) on an Agilent Technologies 6100 quadrupole instrument equipped with UV detection at 254 and 210 nm and a Varian C8 analytical column. LC-MS analyses were performed using a step-wise H₂O/MeOH gradient with the % MeOH increasing from 75-95% over the course of 3 min unless otherwise specified. Flash column chromatography was conducted using CombiFlash Rf 200 (Telendyne-Isco) automated flash chromatography system with hand-packed RediSep columns. Evaporation of solvents was carried out on a rotary evaporator under reduced pressure and under ultra-high vacuum (UHV) where appropriate. ¹H NMR and ¹³C NMR spectra were recorded at ambient temperature on a Varian 400 spectrometer. ³¹P spectra were recorded

at ambient temperature on either a Mercury 300 or Varian 400 spectrometer. Unless otherwise specified, all NMR spectra were obtained in deuterated chloroform (CDCl_3) and referenced to the residual solvent peak. Chemical shifts are given in δ values and coupling constants are reported in hertz (Hz). Melting points were determined on a MelTemp melting apparatus and are uncorrected. High resolution mass-spectra (HRMS) were acquired on a VG 70-S Nier Johnson or JEOL mass spectrometer. Elemental analyses were performed by Atlantic Microlabs (Norcross, GA) for C, H, N analysis and are in agreement with the proposed structures with purity $\geq 95\%$. Crystallographic data was collected using a Bruker APEX-II CCD diffractometer and the structures were solved with the SIR2004 structure solution program in Olex2⁹⁸ using Direct Methods and the XH refinement package using CGLS minimization.

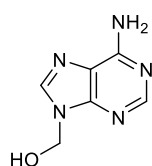
General Procedures

Procedure 2A: To a stirring suspension of paraformaldehyde (3.0 equiv.) in water was added sodium hydroxide (0.5 equiv.). After solubilization of paraformaldehyde, the desired purine (1 equiv.) was added. The concentration of purine was 0.5 M. The mixture stirred at room temperature with progress monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 35-75% MeOH, 3 min). After 45 min of stirring, the solution was chilled to 0 °C and the pH was neutralized with saturated NH_4Cl . The precipitate was filtered over a glass frit and the supernatant continued stirring to induce further precipitation which was subsequently filtered. The solid was dried over solid P_2O_5 and further dried under UHV to furnish the desired hemiaminal.

Procedure 2B: To a stirring suspension of hemiaminal H (0.25 g, 1.0 equiv.) in water (0.5 M) was added dilute sulfuric acid until the mixture became homogenous (pH = 0). Then, desired thiol T (1.2 equiv) was added and the solution stirred at room temperature for 2 h with progress monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 35-75% MeOH, 3 min). Upon completion, the solution was

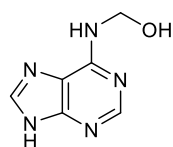
neutralized with NaHCO_3 and chilled to $0\text{ }^\circ\text{C}$ to precipitate the title compound in $> 85\%$ purity that was further purified via C18 reverse phase column chromatography using a $\text{H}_2\text{O}/\text{MeOH}$ gradient.

Procedure 2C: To a suspension of pyrimidine hemiaminal H (0.150 g, 1.0 equiv.) in concentrated phosphoric acid (1.0 mL) was added desired thiol T (1.2 equiv.). The resulting solution stirred at room temperature for 7 h with progress monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 35-75% MeOH, 3 min). Upon completion, the solution was directly purified via C18 reverse phase column chromatography using a $\text{H}_2\text{O}/\text{MeOH}$ gradient to afford the title compound.



((6-Amino-9H-purin-9-yl)methanol (3.1))

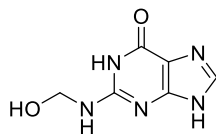
To a stirring suspension of adenine (0.25 g, 1.9 mmol) in water (7.40 mL) was added aqueous sodium hydroxide (0.37 mL, 0.37 mmol) followed by paraformaldehyde (0.089 g, 3.0 mmol). The reaction mixture stirred at room temperature for 2 h, then chilled to $0\text{ }^\circ\text{C}$ and filtered cold. The collected solid was dried over solid P_2O_5 overnight and further dried under UHV to afford the title compound ((6-amino-9H-purin-9-yl)methanol (0.22 g, 1.3 mmol, 72 % yield) as a white powder. ^1H NMR (400 MHz, d_6 -DMSO) δ 8.17 (s, 1H), 8.16 (s, 1H), 7.24 (s, 2H), 5.49 (s, 2H), 3.36 (s, 1H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 156.0, 152.6, 149.1, 140.7, 118.6, 65.6. HRMS could not be obtained for this compound.



((9H-Purin-6-yl)amino)methanol (3.2)

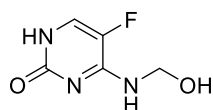
General Procedure 2A: (8.56 g, 100 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 13.05 (s, 1H), 8.25 (s, 1H), 8.18 (s, 1H), 5.64 (s, 1H), 4.99 (s, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 154.3,

152.7, 150.4, 140.0, 119.2, 63.9. HRMS (ESI) m/z calculated for $C_6H_8ON_5$ $[M + H]^+$: 166.0723, found 166.0723. Melting Point: > 260 °C.



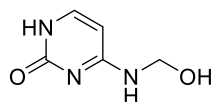
2-((Hydroxymethyl)amino)-1H-purin-6(9H)-one (3.3)

General Procedure 2A: (10.7 g, 89% yield) as an off-white solid. Characterized by LC-MS. m/z $[M+H]^+ = 182$



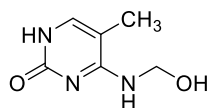
5-Fluoro-4-((hydroxymethyl)amino)pyrimidin-2(1H)-one (3.4)

General Procedure 2A: (0.86 g, 70% yield) as a white solid. Characterized by LC-MS. m/z $[M+H]^+ = 160$



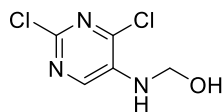
4-((Hydroxymethyl)amino)pyrimidin-2(1H)-one (3.5)

General Procedure 2A: (3.1 g, 48% yield) as a white solid. Characterized by LC-MS. m/z $[M+H]^+ = 142$



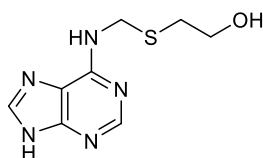
4-((Hydroxymethyl)amino)-5-methylpyrimidin-2(1H)-one (3.6)

General Procedure 2A: (0.62 g, 25% yield) as a white solid. Characterized by LC-MS. m/z $[M+H]^+ = 156$



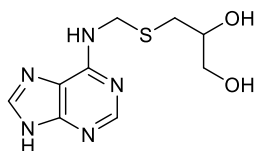
((2,4-Dichloropyrimidin-5-yl)amino)methanol (3.7)

General Procedure 2A: (0.21 g, 71% yield) as a light brown solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 8.29 (s, 1H), 7.08 (t, $J = 6.9$ Hz, 1H), 4.81 (d, $J = 6.9$ Hz, 2H), 3.34 (s, 1H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 153.2, 146.1, 145.9, 137.4, 70.0. HRMS (APCI) m/z calculated for $\text{C}_5\text{H}_6\text{ON}_3\text{Cl}_2$ $[M + H]^+$: 193.9882, found 193.9884. Melting Point: 144-146°C.



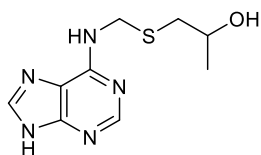
2-(((6-Amino-9H-purin-9-yl)methyl)thio)ethanol (3.8)

General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 2-mercaptoethan-1-ol, (0.241 g, 57.8 % yield) as a white powder. ^1H NMR (400 MHz, D_2O) δ 8.52 (s, 1H), 8.37 (s, 1H), 4.82 (s, 2H), 3.74 (t, $J = 6.1$ Hz, 2H), 2.83 (t, $J = 6.1$ Hz, 2H). ^{13}C NMR (101 MHz, D_2O) δ 150.4, 146.6, 145.7, 143.7, 114.1, 60.2, 42.8, 32.6. HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{12}\text{ON}_5\text{S}$ $[M + H]^+$: 226.0757, found 226.0753. Anal. calculated for $\text{C}_8\text{H}_{11}\text{N}_5\text{OS}$: C, 42.65; H, 4.92; N, 31.09. Found: C, 42.23; H, 4.75; N, 30.70. Melting Point: 149-151°C. X-ray quality crystals were produced from vapor diffusion of a solution of **3.8** in water and ammonium hydroxide.



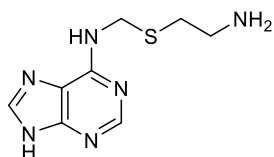
3-(((9H-Purin-6-yl)amino)methyl)thio)propane-1,2-diol (3.9)

General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 3-mercaptopropane-1,2-diol, (0.207 g, 53 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 8.64 (s, 1H), 8.46 (s, 1H), 8.25 (s, 1H), 8.16 (s, 1H), 4.72 (s, 2H), 3.62 (dt, J = 10.9, 5.4 Hz, 1H), 3.34 (d, J = 5.5 Hz, 2H), 2.79 (dd, J = 13.6, 4.6 Hz, 1H), 2.65 (dd, J = 13.6, 6.8 Hz, 1H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 166.3, 152.3, 152.0, 140.4, 71.88, 64.6, 42.9, 34.6. Unobserved resonance at approximately 120 ppm. HRMS (ESI) m/z calculated for $\text{C}_9\text{H}_{14}\text{O}_2\text{N}_5\text{S}$ $[\text{M} + \text{H}]^+$: 256.0862, found 256.0860. Anal. calculated for $\text{C}_9\text{H}_{17}\text{N}_5\text{O}_4\text{S}$ (as a dihydrate): C, 37.10; H, 5.88; N, 24.04. Found: C, 36.25; H, 5.66; N, 24.55. Melting Point: Decomposes at 160 °C.



1-(((9H-Purin-6-yl)amino)methylthio)propan-2-ol (3.10)

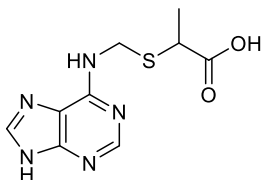
General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 1-mercaptopropan-2-ol, (0.188 g, 51.8 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 13.01 (s, 1H), 8.33 (s, 1H), 8.26 (s, 1H), 8.16 (s, 1H), 4.83 (s, 1H), 4.71 (s, 2H), 3.77 (m, 1H), 2.67 (m, 2H), 1.08 (d, J = 6.1 Hz, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 153.4, 152.2, 149.9, 139.5, 119.4, 66.5, 42.9, 39.6, 22.6. HRMS (ESI) m/z calculated for $\text{C}_9\text{H}_{14}\text{ON}_5\text{S}$ $[\text{M} + \text{H}]^+$: 240.0913, found 240.0909. Anal. calculated for $\text{C}_9\text{H}_{13}\text{N}_5\text{OS}$: C, 45.17; H, 5.48; N, 29.27. Found: C, 43.92; H, 5.19; N, 28.77. Melting Point: Decomposes at 225°C.



N-(((2-Aminoethyl)thio)methyl)-9H-purin-6-amine (3.11)

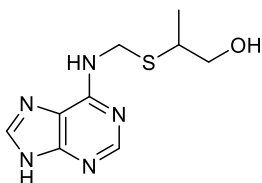
General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = cystamine, (0.185 g, 54.5 % yield) as a white powder. ^1H NMR (400 MHz, D_2O) δ 7.77 (s, 1H), 7.75 (s, 1H), 4.86 (s, 2H), 4.47 (s, 2H),

3.27 (t, $J = 6.6$ Hz, 2H), 2.92 (t, $J = 6.6$ Hz, 2H). ^{13}C NMR (101 MHz, D_2O) δ 151.4, 151.1, 148.7, 140.1, 116.2, 41.9, 38.4, 27.2. HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{13}\text{N}_6\text{S}$ $[\text{M} + \text{H}]^+$: 225.0916, found 225.0915. Melting Point: Decomposes at 136 °C.



2-(((9H-Purin-6-yl)amino)methylthio)propanoic acid (3.12)

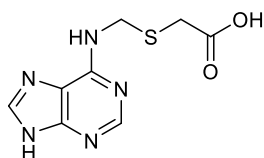
General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = thiolactic acid, (0.23 g, 60 % yield) as a white powder. ^1H NMR (400 MHz, d_6 -DMSO) δ 12.89 (s, 1H), 8.26 (s, 1H), 8.17 (s, 1H), 4.81 (s, 2H), 3.74 (q, $J = 13.5, 6.5$ Hz, 1H), 1.33 (d, $J = 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 174.5, 152.1, 139.7, 42.3, 40.5, 18.1. Unobserved resonances at 154 ppm, 151 ppm, and 120 ppm. HRMS (ESI) m/z calculated for $\text{C}_9\text{H}_{12}\text{O}_2\text{N}_5\text{S}$ $[\text{M} + \text{H}]^+$: 254.0706, found 254.0705. Melting Point: 179-180 °C.



2-(((9H-Purin-6-yl)amino)methylthio)propan-1-ol (3.13)

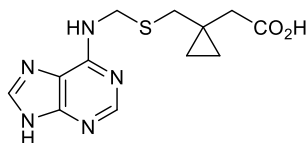
General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 2-mercaptopropan-1-ol, (0.177 g, 49 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 12.98 (s, 1H), 8.33 - 8.20 (m, 2H), 8.15 (s, 1H), 4.77 (s, 1H), 4.69 (s, 1H), 3.54 (dd, $J = 10.7, 5.5$ Hz, 1H), 3.33 (dd, $J = 10.7, 7.3$ Hz, 1H), 3.04 (sex, $J = 13.6, 6.8$ Hz, 1H), 1.18 (d, $J = 6.9$ Hz, 3H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 153.1, 152.10, 152.08, 139.5, 118.6, 66.4, 41.3, 18.3. Masked ^{13}C NMR resonance behind solvent peak. HRMS (ESI)

m/z calculated for $C_9H_{14}N_5OS$ $[M + H]^+$: 240.0913, found 240.0914. Melting Point: Decomposes at 180 °C.



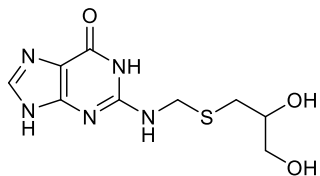
2-(((9H-Purin-6-yl)amino)methyl)thio)acetic acid (3.14)

General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 2-mercaptoacetic acid (0.23 g, 64% yield) as a white powder. 1H NMR (400 MHz, d_6 -DMSO) δ 12.69 (s, 2H), 8.39 (s, 1H), 8.26 (s, 1H), 8.17 (s, 1H), 4.77 (s, 2H), 3.46 (s, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 171.8, 153.0, 152.1, 139.8, 118.9, 42.7, 39.5, 32.7. HRMS (ESI) m/z calculated for $C_8H_{10}N_5O_2S$ $[M + H]^+$: 240.0549, found 240.0548. Melting Point: Decomposes at 210 °C.



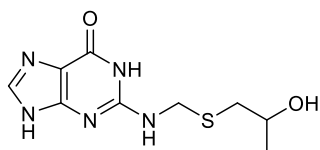
2-(1-(((9H-Purin-6-yl)amino)methyl)thio)methyl)cyclopropyl)acetic acid (3.15)

General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = 2-(1-(mercaptomethyl)cyclopropyl)acetic acid, (0.171 g, 46 % yield) as a white powder. 1H NMR (400 MHz, d_6 -DMSO) δ 12.89 (s, 1H), 12.24 (s, 1H), 8.32 (s, 1H), 8.25 (s, 1H), 8.15 (s, 1H), 4.69 (s, 2H), 2.81 (s, 2H), 2.30 (s, 2H), 0.51 (s, 2H), 0.45 (t, $J = 5.1$ Hz, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 173.2, 153.4, 152.1, 149.9, 139.4, 119.2, 42.4, 38.6, 17.2, 11.8. Unobserved resonance. HRMS (ESI) m/z calculated for $C_{12}H_{16}O_2N_5S$ $[M + H]^+$: 294.1019, found 294.1010. Melting Point: > 260 °C. X-ray quality crystals were produced from vapor diffusion of a solution of **3.15** in water and ammonium hydroxide.



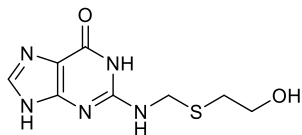
2-(((2,3-Dihydroxypropyl)thio)methylamino)-1,9-dihydro-6H-purin-6-one (3.16)

General Procedure 2B: H = 2-((hydroxymethyl)amino)-1,9-dihydro-6H-purin-6-one, T = 3-mercaptopropane-1,2-diol (0.20 g, 45% yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 12.57 (s, 1H), 10.71 (s, 1H), 7.66 (s, 1H), 6.93 (t, $J = 6.2$ Hz, 1H), 4.97 - 4.81 (m, 1H), 4.69 - 4.57 (m, 1H), 4.56 - 4.46 (m, 2H), 3.66 - 3.57 (m, 1H), 2.76 (dd, $J = 13.6, 4.5$ Hz, 1H), 2.60 (dd, $J = 13.7, 6.9$ Hz, 1H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 156.9, 151.5, 151.1, 135.7, 116.9, 71.8, 64.5, 43.9, 34.6. HRMS (ESI) m/z calculated for $\text{C}_9\text{H}_{14}\text{O}_3\text{N}_5\text{S}$ [$\text{M} + \text{H}$] $^+$: 272.0811, found 272.0811. Anal. calculated for $\text{C}_9\text{H}_{13}\text{N}_5\text{O}_3\text{S}$: C, 39.84; H, 4.83; N, 25.81. Found: C, 38.83; H, 4.54; N, 25.08. Melting Point: Decomposes at 210 °C.



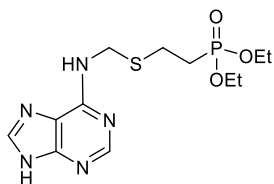
2-(((2-Hydroxypropyl)thio)methylamino)-1,9-dihydro-6H-purin-6-one (3.17)

General Procedure 2B: H = 2-((hydroxymethyl)amino)-1,9-dihydro-6H-purin-6-one, T = 1-mercaptopropan-2-ol, (225 mg, 64 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 12.57 (s, 1H), 10.76 (s, 1H), 7.66 (s, 1H), 6.96 (s, 1H), 4.80 (s, 1H), 4.50 (d, $J = 6.1$ Hz, 2H), 3.78 (m, 1H), 2.68 - 2.57 (m, 2H), 1.10 (d, $J = 8.0$ Hz, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 157.3, 151.9, 151.5, 136.1, 117.3, 66.9, 44.2, 39.8, 23.1. HRMS (ESI) m/z calculated for $\text{C}_9\text{H}_{14}\text{O}_2\text{N}_5\text{S}$ [$\text{M} + \text{H}$] $^+$: 256.0862, found 256.0862. Anal. calculated for $\text{C}_9\text{H}_{13}\text{N}_5\text{O}_2\text{S}$: C, 42.34; H, 5.13; N, 27.43. Found: C, 41.35; H, 4.98; N, 27.24. Melting Point: > 260 °C.



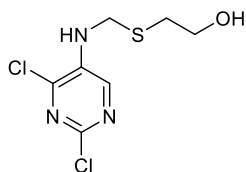
2-(((2-Hydroxyethyl)thio)methyl)amino)-1,9-dihydro-6H-purin-6-one (3.18)

General Procedure 2B: H = 2-((hydroxymethyl)amino)-1,9-dihydro-6H-purin-6-one, T = 2-mercaptoethan-1-ol, (0.223 g, 67% yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 7.67 (s, 1H), 7.00 (s, 1H), 4.87 (s, 1H), 4.50 (d, $J = 6.4$ Hz, 2H), 3.57 (t, $J = 6.7$ Hz, 2H), 2.69 (t, $J = 6.7$ Hz, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 156.9, 151.5, 151.0, 135.7, 117.0, 61.4, 43.4, 33.4. HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{12}\text{O}_2\text{N}_5\text{S}$ $[\text{M} + \text{H}]^+$: 242.0706, found 242.0705. Anal. calculated for $\text{C}_8\text{H}_{11}\text{N}_5\text{O}_2\text{S}$: C, 39.82; H, 4.60; N, 29.03. Found: C, 38.88; H, 4.47; N, 29.13. Melting Point: > 260 °C.



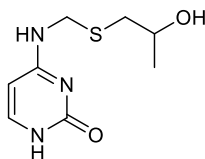
Diethyl (2-(((9H-purin-6-yl)amino)methyl)thio)ethyl)phosphonate (3.19)

General Procedure 2B: H = ((9H-purin-6-yl)amino)methanol, T = diethyl (2-mercaptoethyl)phosphonate, 55 % yield. Amorphous solid. ^1H NMR (399 MHz, CDCl_3) δ 8.48 (s, 1H), 8.06 (s, 1H), 7.74 (s, 1H), 4.96 (s, 2H), 4.07 (p, $J = 7.2$ Hz, 4H), 2.93 (td, $J = 9.6, 5.8$ Hz, 2H), 2.33 - 2.19 (m, 2H), 1.27 (t, $J = 7.1$ Hz, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ 154.2, 152.4, 150.0, 138.8, 119.4, 61.8 (d, $J = 6.5$ Hz), 42.1, 27.0 (d, $J = 135.5$ Hz), 23.8 (d, $J = 2.6$ Hz), 16.4 (d, $J = 6.0$ Hz). ^{31}P NMR (162 MHz, CDCl_3) δ 28.8. HRMS (ESI) m/z calculated for $\text{C}_{12}\text{H}_{21}\text{N}_5\text{O}_3\text{PS}$ $[\text{M} + \text{H}]^+$: 346.1097, found 346.1094.



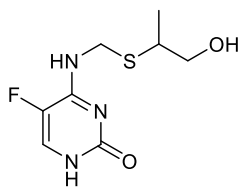
2-((((2,4-Dichloropyrimidin-5-yl)amino)methyl)thio)ethan-1-ol (3.20)

General Procedure 2C: H = ((2,4-dichloropyrimidin-5-yl)amino)methanol, T = 2-mercaptoethan-1-ol, (0.105 g, 80 % yield) as an off-white solid. ¹H NMR (400 MHz, CDCl₃) δ 8.06 (s, 1H), 4.54 (s, 2H), 3.89 (t, *J* = 6.2 Hz, 2H), 2.77 (t, *J* = 5.5 Hz, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 147.6, 147.1, 141.5, 136.1, 62.8, 46.4, 33.7. HRMS (ESI) *m/z* calculated for C₇H₁₀ON₃Cl₂S [M + H]⁺ : 253.9916, found 253.9918.



4-((((2-Hydroxypropyl)thio)methyl)amino)pyrimidin-2(1H)-one (3.21)

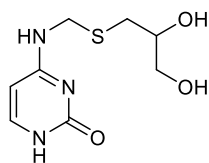
General Procedure 2C: H = 4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 1-mercaptopropan-2-ol, (0.107 g, 47 % yield) as a white powder. ¹H NMR (400 MHz, *d*₆-DMSO) δ 10.43 (s, 1H), 8.14 (t, *J* = 5.8 Hz, 1H), 7.32 (d, *J* = 6.9 Hz, 1H), 5.62 (d, *J* = 7.0 Hz, 1H), 4.81 (s, 1H), 4.47 (d, *J* = 6.1 Hz, 2H), 3.77 (sex, *J* = 12.0, 6.4 Hz, 1H), 2.63 (dd, *J* = 13.3, 6.1 Hz, 1H), 2.56 (dd, *J* = 13.2, 5.9 Hz, 1H), 1.10 (d, *J* = 6.2 Hz, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO) δ 164.2, 156.4, 142.2, 93.3, 66.3, 41.8, 39.3, 22.6. HRMS (ESI) *m/z* calculated for C₈H₁₄O₂N₃S [M + H]⁺ : 216.0801, found 216.0801. Anal. calculated for C₈H₁₃N₃O₂S: C, 44.63; H, 6.09; N, 19.52. Found: C, 43.79; H, 6.16; N, 18.87. Melting Point: 156-157°C.



4-Amino-5-fluoro-1-(((1-hydroxypropan-2-yl)thio)methyl)pyrimidin-2(1H)-one (3.22)

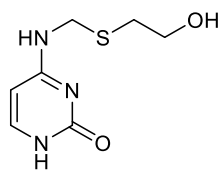
General Procedure 2C: H = 5-fluoro-4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 2-mercaptopropan-1-ol, (56 mg, 25 % yield) as a white solid. ¹H NMR (400 MHz, *d*₆-DMSO) δ 10.32 (s,

1H), 8.46 (t, $J = 5.6$ Hz, 1H), 7.65 (d, $J = 6.2$ Hz, 1H), 4.98 (s, 1H), 4.56 (dd, $J = 13.6, 6.1$ Hz, 1H), 4.44 (dd, $J = 13.6, 5.8$ Hz, 1H), 3.52 (dd, $J = 10.8, 5.5$ Hz, 1H), 3.35 (dd, $J = 10.7, 7.1$ Hz, 1H), 3.02 (sex, $J = 13.5, 6.7$ Hz, 1H), 1.18 (d, $J = 6.9$ Hz, 3H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 155.2 (d, $J = 12.9$ Hz), 154.7, 136.2 (d, $J = 237.9$ Hz), 126.7 (d, $J = 29.3$ Hz), 66.2, 41.7, 40.5, 18.2. ^{19}F NMR (376 MHz, d_6 -DMSO) δ -172.7 (d, $J = 6.0$ Hz). HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{13}\text{O}_2\text{N}_3\text{FS}$ $[\text{M} + \text{H}]^+$: 234.0707, found 234.0708. Anal. calculated for $\text{C}_8\text{H}_{12}\text{N}_3\text{O}_2\text{SF}$: C, 41.19; H, 5.19; N, 18.01. Found: C, 40.71; H, 5.18; N, 17.63. Melting Point: 172-174°C. X-ray quality crystals were produced from vapor diffusion of a solution of **3.22** in water and methanol.



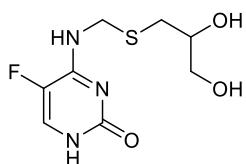
4-(((2,3-Dihydroxypropyl)thio)methyl)amino)pyrimidin-2(1H)-one (**3.23**)

General Procedure 2C: H = 4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 3-mercapto propane-1,2-diol, (116 mg, 47.2 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 10.44 (s, 1H), 8.14 (t, $J = 6.2$ Hz, 1H), 7.32 (d, $J = 7.0$ Hz, 1H), 5.63 (d, $J = 7.0$ Hz, 1H), 4.92 (s, 1H), 4.61 (s, 1H), 4.53 - 4.42 (m, 2H), 3.66 - 3.58 (m, 1H), 3.35 (d, $J = 5.2$ Hz, 2H), 2.73 (dd, $J = 13.5, 4.9$ Hz, 1H), 2.58 (dd, $J = 13.5, 6.9$ Hz, 1H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 164.2, 156.5, 142.1, 93.4, 71.6, 64.5, 42.0, 34.5. HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{14}\text{O}_3\text{N}_3\text{S}$ $[\text{M} + \text{H}]^+$: 232.0750, found 232.0752. Anal. calculated for $\text{C}_8\text{H}_{13}\text{N}_3\text{O}_3\text{S}$: C, 41.55; H, 5.67; N, 18.17. Found: C, 41.14; H, 5.65; N, 17.87. Melting Point: 155-156°C.



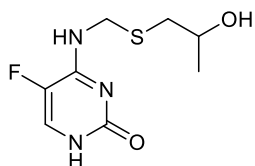
4-Amino-1-(((2-hydroxyethyl)thio)methyl)pyrimidin-2(1H)-one (**3.24**)

General Procedure 2C: H = 4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 2-mercaptoethanol, (97 mg, 45 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 10.43 (s, 1H), 8.14 (t, J = 6.2 Hz, 1H), 7.32 (d, J = 7.0 Hz, 1H), 5.63 (d, J = 7.0 Hz, 1H), 4.86 (s, 1H), 4.46 (d, J = 6.3 Hz, 2H), 3.56 (t, J = 6.8 Hz, 2H), 2.67 (t, J = 6.8 Hz, 2H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 164.2, 156.4, 142.2, 93.3, 61.1, 41.4, 33.2. HRMS (ESI) m/z calculated for $\text{C}_7\text{H}_{12}\text{O}_2\text{N}_3\text{S}$ $[\text{M} + \text{H}]^+$: 202.0644, found 202.0644. Anal. calculated for $\text{C}_7\text{H}_{11}\text{N}_3\text{O}_2\text{S}$: C, 41.78; H, 5.51; N, 20.88. Found: C, 41.50; H, 5.44; N, 20.72. Melting Point: 143-145°C.



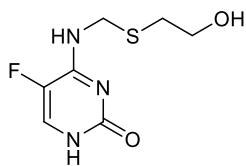
4-(((2,3-Dihydroxypropyl)thio)methyl)amino)-5-fluoropyrimidin-2(1H)-one (3.25)

General Procedure 2C: H = 5-fluoro-4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 3-mercaptopropane-1,2-diol, (0.107 g, 46.1 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 10.44 (s, 1H), 8.45 (t, J = 6.1 Hz, 1H), 7.65 (d, J = 6.2 Hz, 1H), 4.96 (s, 1H), 4.66 (s, 1H), 4.51 (dd, J = 12.3, 5.1 Hz, 1H), 4.46 (dd, J = 12.3, 5.1 Hz, 1H), 3.61 (dt, J = 11.4, 5.5 Hz, 1H), 3.36 (dd, J = 9.1, 3.6 Hz, 1H), 3.32 (dd, J = 9.1, 4.1 Hz, 1H), 2.80 (dd, J = 13.6, 4.7 Hz, 1H), 2.64 (dd, J = 13.6, 6.9 Hz, 1H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ , 155.3 (d, J = 12.8 Hz), 154.8, 136.2 (d, J = 238.1 Hz), 126.7 (d, J = 30.0 Hz), 71.5, 64.5, 42.3, 35.2. ^{19}F NMR (376 MHz, d_6 -DMSO) δ -172.7 (d, J = 6.1 Hz). HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{13}\text{O}_3\text{N}_3\text{S}$ $[\text{M} + \text{H}]^+$: 250.0656, found 250.0658. Anal. calculated for $\text{C}_8\text{H}_{12}\text{N}_3\text{O}_3\text{SF}$: C, 38.55; H, 4.85; N, 16.86. Found: C, 38.26; H, 4.69; N, 16.76. Melting Point: 163-164°C.



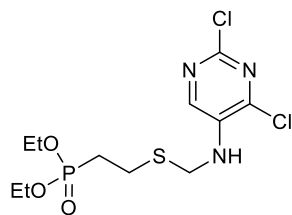
5-Fluoro-4-(((2-hydroxypropyl)thio)methyl)amino)pyrimidin-2(1H)-one (3.26)

General Procedure 2C: H = 5-fluoro-4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 1-mercaptopropan-2-ol (0.115 g, 52.5 % yield) as a white solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 8.47 (t, $J = 5.9$ Hz, 1H), 7.65 (d, $J = 6.2$ Hz, 1H), 4.47 (d, $J = 6.1$ Hz, 2H), 3.78 (sep, 1H), 2.66 (d, $J = 5.5$ Hz, 2H), 1.10 (d, $J = 6.1$ Hz, 3H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 155.2 (d, $J = 13.1$ Hz), 154.7, 136.2 (d, $J = 238.0$ Hz), 126.7 (d, $J = 29.3$ Hz), 66.2, 42.1, 39.9, 22.5. ^{19}F NMR (376 MHz, d_6 -DMSO) δ -172.7 (d, $J = 6.0$ Hz). HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{13}\text{O}_2\text{N}_3\text{SF}$ $[\text{M} + \text{H}]^+$: 234.0707, found 234.0707. Anal. calculated for $\text{C}_8\text{H}_{12}\text{N}_3\text{O}_2\text{SF}$: C, 41.19; H, 5.19; N, 18.01. Found: C, 40.43; H, 5.11; N, 17.79. Melting Point: 168-170 °C.



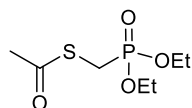
5-Fluoro-4-(((2-hydroxyethyl)thio)methyl)amino)pyrimidin-2(1H)-one (3.27)

General Procedure 2C: H = 5-fluoro-4-((hydroxymethyl)amino)pyrimidin-2(1H)-one, T = 2-mercaptoethan-1-ol, (0.128 g, 61.9 % yield) as a white powder. ^1H NMR (399 MHz, d_6 -DMSO) δ 8.49 (t, $J = 6.0$ Hz, 1H), 7.65 (d, $J = 6.2$ Hz, 1H), 4.47 (d, $J = 6.2$ Hz, 2H), 3.56 (t, $J = 6.7$ Hz, 2H), 2.72 (t, $J = 6.7$ Hz, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 155.27 (d, $J = 13.0$ Hz), 154.8, 136.3 (d, $J = 238.0$ Hz), 126.8 (d, $J = 28.8$ Hz), 61.1, 41.7, 33.8. ^{19}F NMR (376 MHz, d_6 -DMSO) δ -172.6 (d, $J = 6.0$ Hz). HRMS (ESI) m/z calculated for $\text{C}_7\text{H}_{11}\text{O}_2\text{N}_3\text{SF}$ $[\text{M} + \text{H}]^+$: 220.0550, found 220.0551. Anal. calculated for $\text{C}_7\text{H}_{10}\text{N}_3\text{O}_2\text{SF}$: C, 38.35; H, 4.60; N, 19.17. Found: C, 37.96; H, 4.47; N, 18.93. Melting Point: 168-172°C.



Diethyl (2-(((2,4-dichloropyrimidin-5-yl)amino)methyl)thio)ethylphosphonate (3.28)

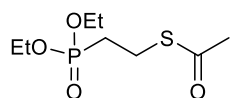
General Procedure 2C: H = ((2,4-dichloropyrimidin-5-yl)amino)methanol, T = diethyl (2-mercaptoethyl)phosphonate, 82 % yield. Light brown solid. ^1H NMR (400 MHz, d_6 -DMSO) δ 8.29 (s, 1H), 7.08 (t, $J = 6.6$ Hz, 1H), 4.58 (d, $J = 6.6$ Hz, 2H), 4.03 - 3.93 (m, 4H), 2.68 - 2.60 (m, 2H), 2.11 - 2.00 (m, 2H), 1.21 (t, $J = 7.1$ Hz, 6H). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 146.1, 145.0, 143.4, 136.7, 61.5 (d, $J = 6.2$ Hz), 44.8, 26.2 (d, $J = 133.6$ Hz), 22.9 (d, $J = 3.0$ Hz), 16.7 (d, $J = 5.8$ Hz). ^{31}P NMR (162 MHz, d_6 -DMSO) δ 29.1. HRMS (ESI) m/z calculated for $\text{C}_{11}\text{H}_{19}\text{O}_3\text{N}_3\text{Cl}_2\text{PS}$ $[\text{M} + \text{H}]^+$: 374.0256, found 374.0256. Melting Point: 118-119°C.



S-((Diethoxyphosphoryl)methyl) ethanethioate (3.31)

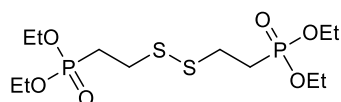
To a round bottom flask equipped with a magnetic stir bar was added ethanol (25 mL) and pyridine (3.35 mL, 39.4 mmol). The mixture was chilled to 0°C then (chloromethyl)phosphonic dichloride (1.83 mL, 17.9 mmol) was added in one portion. After 30 min, the solvent was evaporated under reduced pressure and the undesired pyridine salt was partitioned between acidic water and DCM. The organic layer was collected, concentrated, and the yellow oil was dissolved in DMSO (50 mL). Then, potassium ethanethioate (2.252 g, 19.72 mmol) was added and the mixture stirred at room temperature overnight. Progress was monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 75-95% MeOH, 3 min). The red solution was partitioned between ethyl acetate and brine. The organic layer was collected, dried under

anhydrous sodium sulfate and concentrated to furnish a red oil that was purified on a silica column using a hexanes/EtOAc gradient (0-65% EtOAc) to afford the title compound S-((diethoxyphosphoryl)methyl) ethanethioate (2.15 g, 9.53 mmol, 53.2 % yield). ^1H NMR (399 MHz, CDCl_3) δ 4.11 (dq, $J = 8.1, 7.1$ Hz, 4H), 3.20 (d, $J = 14.0$ Hz, 2H), 2.37 (s, 3H), 1.30 (t, $J = 7.1$ Hz, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ 193.0, 62.7 (d, $J = 6.4$ Hz), 30.0, 22.2 (d, $J = 151.4$ Hz), 16.3 (d, $J = 6.0$ Hz). ^{31}P NMR (162 MHz, CDCl_3) δ 22.6. HRMS (ESI) m/z calculated for $\text{C}_7\text{H}_{16}\text{O}_4\text{PS}$ $[\text{M} + \text{H}]^+$: 227.0501, found 227.0500.



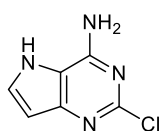
S-(2-(Diethoxyphosphoryl)ethyl) ethanethioate (3.32)

To a stirring solution of diethyl (2-bromoethyl)phosphonate (4.48 mL, 24.48 mmol) in DMSO (82.0 mL) was added potassium ethanethioate (3.08 g, 26.9 mmol). The mixture stirred at room temperature under inert atmosphere for 2 h with progress monitored by LC-MS ($\text{H}_2\text{O}/\text{MeOH}$ gradient, 75-75% MeOH, 3 min). Upon completion, the solution was diluted with water (200 mL) and the crude was extracted into EtOAc (100 mL, 3x). The organic layer was collected, concentrated under reduced pressure, and purified on a silica column using a hexanes/EtOAc gradient to afford the title compound S-(2-(diethoxyphosphoryl)ethyl) ethanethioate (5.77 g, 24.0 mmol, 98 % yield) as a pale yellow oil. ^1H NMR (400 MHz, CDCl_3) δ 4.18 - 4.03 (m, 4H), 3.09 - 3.00 (m, 2H), 2.32 (s, 3H), 2.09 - 1.98 (m, 2H), 1.33 (t, $J = 7.1$ Hz, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ 195.1, 61.8 (d, $J = 6.4$ Hz), 30.5, 26.4 (d, $J = 136.9$ Hz), 22.7 (d, $J = 3.2$ Hz), 16.4 (d, $J = 5.9$ Hz). HRMS (ESI) m/z calculated for $\text{C}_8\text{H}_{17}\text{O}_4\text{PS}$ $[\text{M} + \text{H}]^+$: 241.0657, found 241.0656.



Tetraethyl(disulfanediylbis(ethane-2,1-diyl))bis(phosphonate) (SI-22)

To a stirring solution of diethyl (2-bromoethyl)phosphonate (3.00 mL, 16.3 mmol) in DMSO (54.0 mL) was added potassium ethanethioate (2.05 g, 18.0 mmol). The mixture stirred at room temperature for 2 h with progress monitored by LC-MS (H₂O/MeOH gradient, 75-75% MeOH, 3 min). Upon completion, the solution was diluted with water (200 mL) and the crude was extracted into EtOAc (100 mL, 3x). The organic layer was collected and concentrated under reduced pressure to afford a pale yellow oil that was mixed with 100 mL EtOH and 0.1 M aqueous NaOH (100 mL) at room temperature. Upon completion by TLC, the pH was neutralized with concentrated HCl and the organic solvents evaporated under reduced pressure. The product was extracted into DCM (100 mL, 3x), concentrated under reduced pressure, and purified on a silica column using a DCM/MeOH gradient (0-5%) to afford the title compound tetraethyl (disulfanediylbis(ethane-2,1-diyl))bis(phosphonate) (2.78 g, 7.05 mmol, 43.2 % yield) as a pale yellow oil. ¹H NMR (400 MHz, CDCl₃) δ 4.18 - 4.03 (m, 8H), 2.89 - 2.81 (m, 4H), 2.20 - 2.10 (m, 4H), 1.32 (t, *J* = 7.1 Hz, 12H). ¹³C NMR (101 MHz, CDCl₃) δ 61.9 (d, *J* = 6.6 Hz), 30.9 (d, *J* = 3.7 Hz), 26.2 (d, *J* = 137.0 Hz), 16.4 (d, *J* = 6.0 Hz). HRMS (ESI) *m/z* calculated for C₁₂H₂₉O₆P₂S₂ [M + H]⁺: 395.0875, found 395.0868.

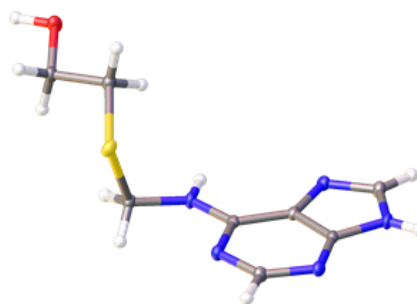
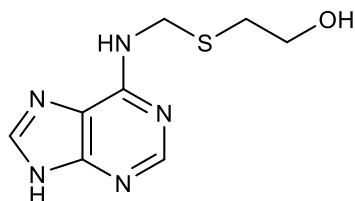


2-Chloro-5H-pyrrolo[3,2-d]pyrimidin-4-amine (3.35)

Concentrated ammonium hydroxide (38.5 mL, 558 mmol) was added to a stirring mixture of 2,4-dichloro-5H-pyrrolo[3,2-d]pyrimidine (7.00 g, 37.2 mmol) in EtOH (74.5 mL). The vessel was sealed and heated to 100 °C overnight. Progress was monitored by LC-MS (H₂O/MeOH gradient, 35-75% MeOH, 3 min). Upon completion, the mixture was cooled to room temperature, chilled to 0 °C and filtered cold. The solid was washed with cold water and dried under UHV to afford the title compound

2-chloro-5*H*-pyrrolo[3,2-*d*]pyrimidin-4-amine (5.15 g, 30.5 mmol, 82 % yield) as a brown solid which required no additional purification and was taken directly to the next step.

3.7 Crystallographic Data

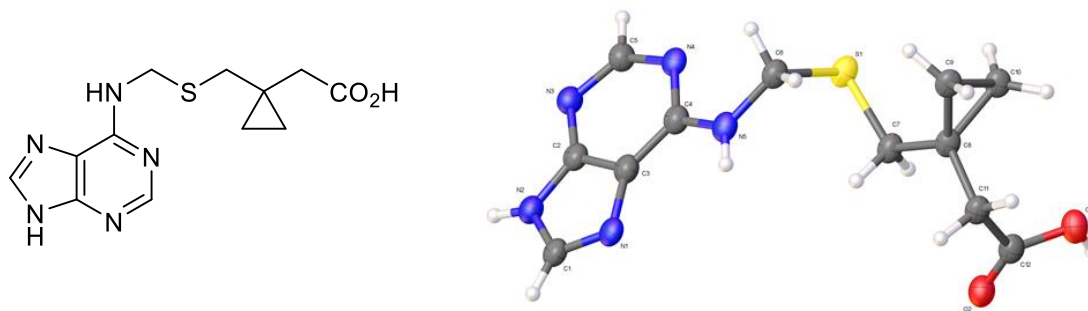


2-(((6-Amino-9H-purin-9-yl)methyl)thio)ethanol (**3.8**)

Single crystals of 2-(((6-Amino-9H-purin-9-yl)methyl)thio)ethanol (**3.8**) were recrystallized from a mixture of water and ammonium hydroxide by diffusion. A suitable crystal was selected and mounted on a loop with paratone oil on a Bruker APEX-II CCD diffractometer. The crystal was kept at 100(2) K during data collection. Using Olex2, the structure was solved with the SIR2004 structure solution program using Direct Methods and refined with the XH refinement package using CGLS minimisation. Crystal Data for $C_8H_{11}N_5OS$ ($M = 225.28$ g/mol): triclinic, space group P-1 (no. 2), $a = 4.9970(12)$ Å, $b = 7.8306(19)$ Å, $c = 12.915(3)$ Å, $\alpha = 73.104(4)^\circ$, $\beta = 86.336(3)^\circ$, $\gamma = 85.176(3)^\circ$, $V = 481.4(2)$ Å³, $Z = 2$, $T = 100(2)$ K, $\mu(\text{MoK}\alpha) = 0.316$ mm⁻¹, $D_{\text{calc}} = 1.554$ g/cm³, 4420 reflections measured ($3.298^\circ \leq 2\Theta \leq 54.968^\circ$), 2215 unique ($R_{\text{int}} = 0.0576$, $R_{\text{sigma}} = 0.0711$) which were used in all calculations. The final R_1 was 0.0548 ($I > 2\sigma(I)$) and wR_2 was 0.1236 (all data). Data was collected by John Bacsá, PhD at the Emory X-ray crystallography core facility. Additional crystallographic data is summarized in Table 3.5.

Table 3.5. Crystal Data and Structure Refinement for 2-(((6-amino-9H-purin-9-yl)methyl)thio)ethanol

Empirical formula	C ₈ H ₁₁ N ₅ OS
Formula weight	225.28
Temperature/K	100(2)
Crystal system	triclinic
Space group	P-1
a/Å	4.9970(12)
b/Å	7.8306(19)
c/Å	12.915(3)
α/°	73.104(4)
β/°	86.336(3)
γ/°	85.176(3)
Volume/Å ³	481.4(2)
Z	2
ρ _{calc} /cm ³	1.554
μ/mm ⁻¹	0.316
F(000)	236.0
Crystal size/mm ³	0.756 × 0.298 × 0.178
Radiation	MoKα (λ = 0.71073)
2θ range for data collection/°	3.298 to 54.968
Reflections collected	4420
Data/restraints/parameters	2215/34/155
Goodness-of-fit on F ²	1.082
Final R indexes [I ≥ 2σ (I)]	R ₁ = 0.0548, wR ₂ = 0.1168
Final R indexes [all data]	R ₁ = 0.0712, wR ₂ = 0.1236
Largest diff. peak/hole / e Å ⁻³	0.57/-0.33

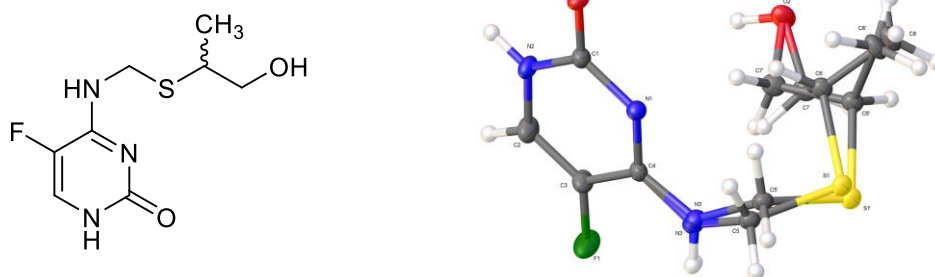


2-(1-((((9*H*-Purin-6-yl)amino)methyl)thio)methyl)cyclopropyl)acetic acid (3.15)

Single colorless needle-shaped crystals of 2-(1-((((9*H*-Purin-6-yl)amino)methyl)thio)methyl)cyclopropyl)acetic acid (3.15) were recrystallised from a mixture of water and ammonium hydroxide by diffusion. A suitable crystal (0.70×0.02×0.02 mm) was selected and mounted on a loop with paratone oil on a Bruker APEX-II CCD diffractometer. The crystal was cooled to $T = 100(2)$ K during the data collection. The structure was solved with the XT structure solution program⁹⁹, using combined Patterson and dual-space recycling methods and by using Olex2 as the graphical interface.⁹⁸ The structure was refined by Least Squares using version 2014/7 of ShelXL. Crystal Data. $C_{12}H_{15}N_5O_2S$, $M_r = 293.35$, monoclinic, $P2_1/c$ (No. 14), $a = 10.375(14)$ Å, $b = 5.066(7)$ Å, $c = 25.75(3)$ Å, $\beta = 96.79(3)^\circ$, $\alpha = \gamma = 90^\circ$, $V = 1344(3)$ Å³, $T = 100(2)$ K, $Z = 4$, $Z' = 1$, $\mu(\text{MoK}\alpha) = 0.251$, 8245 reflections measured, 1754 unique ($R_{int} = 0.3561$) which were used in all calculations. The final wR_2 was 0.2615 (all data) and R_1 was 0.1342 ($I > 2(I)$). Data was collected by John Bacsa, PhD at the Emory X-ray crystallography core facility. Additional crystallographic data is summarized in Table 3.6.

Table 3.6. Crystal Data and Structure Refinement for 2-(1-((((9H-purin-6-yl)amino)methyl)thio)methyl)cyclopropyl)acetic acid (3.15)

Formula	C ₁₂ H ₁₅ N ₅ O ₂ S
$D_{calc.}/\text{g cm}^{-3}$	1.450
ρ/mm^{-1}	0.251
Formula Weight	293.35
Colour	colourless
Shape	needle
Max Size/mm	0.70
Mid Size/mm	0.02
Min Size/mm	0.02
T/K	100(2)
Crystal System	monoclinic
Space Group	P2 ₁ /c
$a/\text{\AA}$	10.375(14)
$b/\text{\AA}$	5.066(7)
$c/\text{\AA}$	25.75(3)
$\alpha/^\circ$	90
$\beta/^\circ$	96.79(3)
$\gamma/^\circ$	90
$V/\text{\AA}^3$	1344(3)
Z	4
Z'	1
$\rho_{min}/^\circ$	1.593
$\rho_{max}/^\circ$	22.667
Measured Refl.	8245
Independent Refl.	1754
Reflections $I >$	701
2sigma(I)	
R_{int}	0.3561
Parameters	182
Restraints	186
Largest Peak	0.456
Deepest Hole	-0.555
GooF	1.046
wR_2 (all data)	0.2615
wR_2	0.1934
R_1 (all data)	0.2978
R_1	0.1342

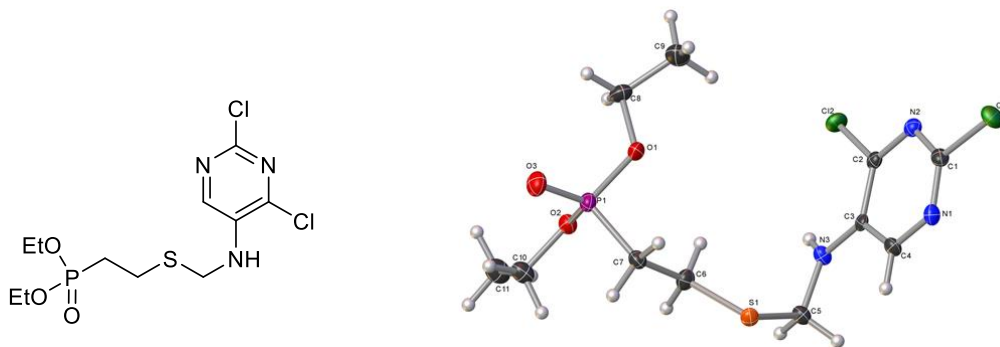


5-Fluoro-4-(((1-hydroxypropan-2-yl)thio)methyl)amino)pyrimidin-2(1H)-one (3.22)

Single colourless prism-shaped crystals of (5-Fluoro-4-(((1-hydroxypropan-2-yl)thio)methyl)amino)pyrimidin-2(1H)-one (**3.22**) were recrystallized from a mixture of methanol and water by vapor diffusion. A suitable crystal (0.57×0.56×0.20 mm) was selected and mounted on a loop with paratone oil on a Bruker APEX-II CCD diffractometer. The crystal was cooled to $T = 100(2)$ K during the data collection. The structure was solved with XT structure solution program⁹⁹, using combined Patterson and dual-space recycling methods. and by using Olex2 as the graphical interface.⁹⁸ The model was refined with version of ShelXL-97 using Least Squares minimisation. Crystal Data: $C_8H_{12}FN_3O_2S$, $M_r = 233.27$, monoclinic, $P2_1/n$ (No. 14), $a = 4.9842(7)$ Å, $b = 12.2199(16)$ Å, $c = 16.942(2)$ Å, $\beta = 92.150(2)^\circ$, $\alpha = \gamma = 90^\circ$, $V = 1031.2(2)$ Å³, $T = 100(2)$ K, $Z = 4$, $Z' = 1$, μ (MoK α) = 0.313, 12695 reflections measured, 3084 unique ($R_{int} = 0.0303$) which were used in all calculations. The final wR_2 was 0.1327 (all data) and R_1 was 0.0541 ($I > 2(I)$). Data was collected by John Bacsá, PhD at the Emory X-ray crystallography core facility. Additional crystallographic data is summarized in Table 3.7.

Table 3.7. Crystal Data and Structure Refinement for 5-Fluoro-4-(((1-hydroxypropan-2-yl)thio)methyl)amino)pyrimidin-2(1H)-one (**3.22**)

Formula	C ₈ H ₁₂ FN ₃ O ₂ S
<i>D</i> _{calc.} / g cm ⁻³	1.503
□ /mm ⁻¹	0.313
Formula Weight	233.27
Colour	colourless
Shape	prism
Max Size/mm	0.57
Mid Size/mm	0.56
Min Size/mm	0.20
<i>T</i> /K	100(2)
Crystal System	monoclinic
Space Group	P2 ₁ /n
<i>a</i> /Å	4.9842(7)
<i>b</i> /Å	12.2199(16)
<i>c</i> /Å	16.942(2)
□ /°	90
□ /°	92.150(2)
□ /°	90
<i>V</i> /Å ³	1031.2(2)
<i>Z</i>	4
<i>Z'</i>	1
□ _{min} /°	2.406
□ _{max} /°	30.471
Measured Refl.	12695
Independent Refl.	3084
Reflections I >	2775
2σ(I)	
<i>R</i> _{int}	0.0303
Parameters	169
Restraints	70
Largest Peak	0.480
Deepest Hole	-0.339
GooF	1.194
<i>wR</i> ₂ (all data)	0.1327
<i>wR</i> ₂	0.1296
<i>R</i> ₁ (all data)	0.0603
<i>R</i> ₁	0.0541



Diethyl (2-(((2,4-dichloropyrimidin-5-yl)amino)methyl)thio)ethyl)phosphonate (3.28)

Single colourless needle-shaped crystals of (Diethyl (2-(((2,4-dichloropyrimidin-5-yl)amino)methyl)thio)ethyl)phosphonate (**3.28**) were recrystallised from a mixture of methanol and water by slow evaporation. A suitable crystal (0.42×0.14×0.13 mm) was selected and mounted on a loop with paratone oil on a Bruker APEX-II CCD diffractometer. The crystal was cooled to $T = 100(2)$ K during data collection. The structure was solved with the ShelXT-2014/4 structure solution program⁹⁹ using combined Patterson and dual-space recycling methods and by using Olex2 as the graphical interface.⁹⁸ The crystal structure was refined with version 2014/7 of ShelXL-2014/7 using Least Squares minimisation. Crystal Data: $C_{11}H_{18}Cl_2N_3O_3PS$, $M_r = 374.21$, orthorhombic, $Pbca$ (No. 61), $a = 18.299(3)$ Å, $b = 9.6099(14)$ Å, $c = 18.744(3)$ Å, $\alpha = \beta = \gamma = 90^\circ$, $V = 3296.2(8)$ Å³, $T = 100(2)$ K, $Z = 8$, $Z' = 1$, $\mu(\text{MoK}\alpha) = 0.629$ mm⁻¹, 17250 reflections measured, 3012 unique ($R_{int} = 0.0854$) which were used in all calculations. The final wR_2 was 0.1393 (all data) and R_1 was 0.0657 ($I > 2\sigma(I)$). Data was collected by Thomas C. Pickel at the Emory X-ray crystallography core facility. Additional crystallographic data is summarized in Table 3.8.

Table 3.8. Crystal Data and Structure Refinement for diethyl (2-(((2,4-dichloropyrimidin-5-yl)amino)methyl)thio)ethyl)phosphonate (**3.28**)

Formula	C ₁₁ H ₁₈ Cl ₂ N ₃ O ₃ PS
$D_{calc.}/\text{g cm}^{-3}$	1.508
μ/mm^{-1}	0.629
Formula Weight	374.21
Colour	colourless
Shape	needle
Max Size/mm	0.42
Mid Size/mm	0.14
Min Size/mm	0.13
T/K	100(2)
Crystal System	orthorhombic
Space Group	Pbca
$a/\text{\AA}$	18.299(3)
$b/\text{\AA}$	9.6099(14)
$c/\text{\AA}$	18.744(3)
$\alpha/^\circ$	90
$\beta/^\circ$	90
$\gamma/^\circ$	90
$V/\text{\AA}^3$	3296.2(8)
Z	8
Z'	1
$\mu_{min}/^\circ$	2.442
$\mu_{max}/^\circ$	25.347
Measured Refl.	17250
Independent Refl.	3012
Reflections $I > 2\sigma(I)$	2236
R_{int}	0.0854
Parameters	192
Restraints	72
Largest Peak	0.525
Deepest Hole	-0.453
GooF	1.059
wR_2 (all data)	0.1393
wR_2	0.1190
R_1 (all data)	0.0992
R_1	0.0657

3.8 References

- (100) De Clercq, E.; Holy, A., Acyclic Nucleoside Phosphonates: A Key Class of Antiviral Drugs. *Nat. Rev. Drug Discovery* **2005**, *4* (11), 928-40.
- (101) Nelson, M. R.; Katlama, C.; Montaner, J. S., *et al.*, The Safety of Tenofovir Disoproxil Fumarate for the Treatment of HIV Infection in Adults: The First 4 Years. *AIDS* **2007**, *21* (10), 1273-1281.
- (102) Le Botlan, D. J., ¹h and ¹³c Nmr Study of the Interaction of Formaldehyde on Adenine and Its Derivatives. *Magn. Reson. Chem.* **1989**, *27* (3), 295-298.
- (103) Eyring, E. J.; Ofengand, J., Reaction of Formaldehyde with Heterocyclic Imino Nitrogen of Purine and Pyrimidine Nucleosides. *Biochemistry* **1967**, *6* (8), 2500-&.
- (104) Alexander, J., Reaction of Formaldehyde with Nucleosides - Addition to 2',3',5'-Triacetyl 9-Beta-D-Arabinofuranosyladenine. *J. Org. Chem.* **1984**, *49* (8), 1453-1454.
- (105) Lewin, S., 155. Reactions of Nucleic Acids and Their Components. Part Iii. The Interaction of Adenine and Uracil with Formaldehyde. *Journal of the Chemical Society (Resumed)* **1964**, (0), 792-809.
- (106) Smith, M., *March's Advanced Organic Chemistry : Reactions, Mechanisms, and Structure*. 7th Edition / ed.; Wiley: Hoboken, New Jersey, 2013; p xxv, 2047 pages.

Chapter 4

Towards the Design and Synthesis of Novel HIV-1 Nef Inhibitors

Abstract: Highly aggressive anti-retroviral therapy has transformed HIV from a death sentence into a chronic, manageable condition for millions of patients world-wide. However, a clinically-viable cure for HIV remains to be found due to the presence and persistence of long-lived viral reservoirs. Histone D-acetylase inhibitors have been shown to reactivate HIV from latency, but T-cell mediated clearance of these newly revealed virus-harboring cells is hampered by a combination of host and viral factors. One such factor is the viral protein, Nef. Nef plays a prominent role in HIV pathogenesis and immune evasion and may provide HIV with a distinct advantage to counter and evade the immune response to thwart viral clearance when HIV is “awakened” from latency. Herein, we describe the *in silico* design and synthesis of novel Nef inhibitors with the aim of targeting Nef-mediated MHC-I downregulation to improve T-cell recognition and facilitate viral clearance.

4.1 Introduction

The human immunodeficiency virus type-1 (HIV-1) is a pervasive pathogen that continues to elude eradication efforts despite the advent of highly active antiretroviral therapy (HAART). At present, there are more than thirty clinically approved drugs for the treatment of HIV. These drugs antagonize critical events in the viral lifecycle and may be divided into six classes: protease inhibitors, integrase inhibitors, nucleoside reverse transcriptase inhibitors (NRTI), non-nucleoside reverse transcriptase inhibitors (NNRTI), fusion inhibitors and entry inhibitors. A “cocktail” of three inhibitors either from the same or different classes constitutes a personalized HAART regimen which must be taken for life. HAART effectively arrests ongoing viral replication resulting in undetectable (< 50 HIV RNA copies/mL) viral loads and delayed disease progression, however, numerous studies have demonstrated that upon interruption or cessation of HAART, virus leaches from a pool of long-lived reservoirs capable of reestablishing the infection.¹⁰⁷⁻¹¹⁰ It is thought that latently infected cells transcribe fragments of HIV at low levels, but these transcripts are retained in the nucleus and are not translated to viral proteins so they evade immune recognition.¹¹¹⁻¹¹³ The exact location and frequency of these latent cells remains ill-defined, however, a large body of evidence indicates a majority of replication competent provirus resides in resting CD4⁺ T cells and gut-associated lymphatic tissue (GALT).¹¹⁴⁻¹¹⁸ The longevity and transcriptional silence of these provirus-containing cells precludes viral clearance and presents a major challenge for the eradication of HIV.

In order to effectively purge the latent reservoir, cells harboring competent HIV-1 provirus require exogenous stimulation to trigger the synthesis of viral proteins. To this end, several small molecules have been identified to reawaken the virus *in vitro* including vorinostat, bryostatins, and disulfiram amongst others.¹¹⁹⁻¹²⁰ The underlying assumption currently adopted by the HIV/AIDS community is that the production of viral proteins from a latent host will stimulate the effector functions of cytotoxic T-lymphocytes (CTLs) to promptly recognize and eliminate the infected cell. It

is also thought that the administration of HAART will prevent new rounds of replication and trigger apoptosis via the accumulation of cytopathic viral proteins. Despite the popularity of this hypothesis, the literature is relatively silent in addressing the fate of latent cells following reactivation. Recent data obtained from clinical trials suggest co-administration of anti-latency drugs with HAART does not significantly reduce the size of the viral reservoir *in vivo*.¹²¹⁻¹²² Further, Siliciano and coworkers revealed that CTLs derived from HAART-compliant patients fail to lyse reactivated CD4⁺ cells integrated with a NL4-3- Δ vpr- Δ env-drEGFP reporter virus *in vitro* unless the CTLs are pre-stimulated with Gag-derived antigenic peptides and IL-2.¹²³ These findings suggest that anti-latency drugs may display limited efficacy in purging the viral reservoirs *in vivo* and reveal yet another unprecedented barrier towards the eradication of HIV.

Given that a majority of HIV-infected cells are eliminated by CTLs¹²⁴, it is critical to understand why these cells fail to eradicate reactivated virus. One well-documented explanation is that prolonged administration of HAART decreases the surface presentation of viral epitopes, diminishing the capacity of CTLs to recognize and eliminate infected cells.¹²⁵ However, there may be other forces that conspire to prevent a robust CTL response during the early stages of viral replication. Of particular interest is the viral protein Nef. Nef is charged with several roles to subvert and counter the adaptive and innate immune response and is specifically tailored to antagonize CTL function, prolong viral persistence, and may promote reseeding of the reservoir during reactivation *in vivo*.¹²⁶

With revived curative efforts within the HIV/AIDS community, the ability of CTLs to purge the viral reservoirs is of substantial interest. Herein, the design and synthesis of novel Nef inhibitors is discussed in an attempt to increase the recognition of infected cells and promote viral clearance.

4.1.1 Structure of HIV-1 Nef

Nef is a flexible 27-kDa *N*-myristoylated viral adapter protein that relies on protein-protein interactions to optimize conditions for viral replication. The protein is characterized by a small structured core composed of two layered α -helicies and a small β -sheet (residues 62-147 and 179-200) and three unstructured loops (residues 2-61, 201-206 and 148-178) that are responsible for Nef's considerable adaptability and versatility (Figure 4.1, standard NL4-3 numbering). Due to the extreme conformational flexibility of the unstructured loops, a complete NMR or crystallographic structure of Nef remains to be realized.

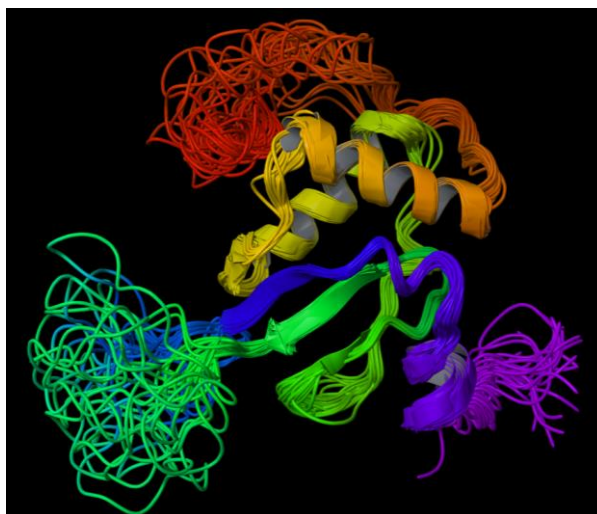


Figure 4.1. Incomplete NMR overlay structure of HIV-1 Nef (PDB code: 2Nef). The structured core is delimited by the presence of a small β sheet and two overlapping α -helicies. A majority of the protein is dominated by several unstructured loops that play several roles in the progression to AIDS.

4.1.2 Nef and HIV-1 Pathogenesis

The link between Nef expression and HIV-1 pathogenesis has been known for some time and several reports have demonstrated that Nef alone is responsible for terminal CD4 decline and the onset of a characteristic AIDS-like disease in transgenic mice.¹²⁷⁻¹²⁸ Clinical evidence in human subjects also support Nef's requirement for disease progression. A cohort of HIV-positive patients infected

with Nef-deleted virus exhibit undetectable viral loads, high CD4 counts, and show few symptoms of clinical disease in the absence of suppressive HAART.¹²⁹⁻¹³¹

There are several mechanisms by which Nef protects HIV from cellular detection and eradication that make Nef an essential player in AIDS progression.¹³²⁻¹³⁵ First, Nef is one of the first and most abundantly expressed genes in the HIV genome accounting for approximately 80% of HIV-1 mRNA during the initial stages of viral replication.¹³⁶ Early expression of Nef is advantageous in that Nef is able to modulate the activity of numerous surface receptors including CD8, CD28, CD1, CD80, CD86, CXCR4, CCR5, DC-SIGN, CCR3, MHC-I, MHC-II and CD4, all of which are involved in the coordination of multiple immune functions.¹³² With respect to the innate immune response, Nef inhibits the expression of NKp44L on the surface of CD4⁺ T-cells to avoid lysis of infected cells by natural killer cells.¹³⁷ With regards to CTL recognition, major histocompatibility complex type-1 (MHC-I) presentation is of critical importance. Engulfed antigens are processed by the cellular machinery and are presented on the cell surface by MHC-I molecules. The presence of foreign antigenic fragments alerts CTLs which subsequently lyse the offending cell. Nef antagonizes this process by restricting the presentation of MHC-I molecules and redirects their fate to lysosomes for degradation. This highly conserved viral function shrouds infected cells from CTL surveillance and may be one of the operative mechanisms that hinders the elimination of reactivated latent cells following treatment with anti-latency drugs.

4.1.3 Mechanistic Details of Nef-Mediated MHC-I Down-regulation

An assembly of sorting proteins are required to recruit MHC-I to the lysosomal compartment. One such protein is adapter protein 1 (AP-1) that is directly targeted by HIV-1 Nef. AP-1 belongs to a small family of heterotetrameric adapter proteins that recruit and sort transmembrane proteins to specific subcellular locations within the endosomal network. The final destination of the cargo protein is determined by a sorting signal encrypted in the amino acid sequence of the target's cytoplasmic

domain (CD). This sequence is deciphered by AP-1 which subsequently transports the protein to the necessary compartment. The $\mu 1$ subunit of AP-1 recognizes the tyrosine sorting motif $Yxx\phi$ (where ϕ is any bulky hydrophobic residue) which connects transmembrane proteins to lysosomes and lysosomal organelles. Crystallographic analysis of $\mu 1$ in complex with a $Yxx\phi$ -bearing peptide reveals a two-pronged recognition site where Y and ϕ fit snugly into two hydrophobic pockets on the surface of $\mu 1$ as shown in Figure 4.2. The identity of ϕ and the adjacent residues tailor the binding properties of the signal and adjusts the sub-cellular localization of the target protein.

In the presence of Nef, $\mu 1$ targets MHC-I for lysosomal degradation. Interestingly, the only sorting signal present on the MHC-I CD is YSQA which lacks the bulky hydrophobic residue at the Y+3 position necessary for $\mu 1$ recognition. This suggests that Nef somehow forces $\mu 1$ to interpret the

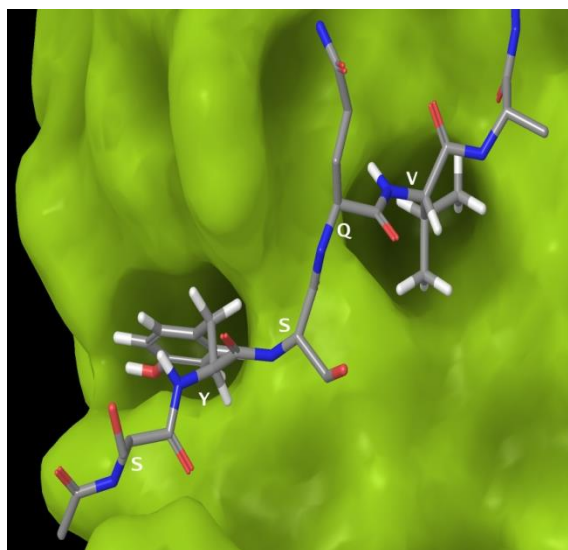


Figure 4.2. A canonical SYSQV signaling peptide bound to the surface of $\mu 1$. Note that valine and tyrosine on the signaling peptide serve as the “prongs” for the two hydrophobic “sockets” on the surface of $\mu 1$. This is the mechanism by which AP-1 sorts $Yxx\phi$ flanked cargo throughout the endosomal network.

cryptic $YxxA$ sorting signal as $Yxx\phi$ which triggers the transport of MHC-I to the lysosome. Several biochemical studies have shown this to be the case and recent crystallographic evidence reveals the

exclusive molecular interactions between $\mu 1$, Nef, and the MHC-I CD. The crystal structure of the ternary Nef- $\mu 1$ -MHC-I complex is shown in Figure 4.3 and reveals the formation of a non-canonical binding groove between Nef and $\mu 1$ that readily accommodates the MHC-I CD. It can be seen that Nef constitutes one of the side walls of the binding groove and locks A-323 into the φ recognition socket of $\mu 1$ which dupes the sorting machinery to interpret YxxA as Yxx φ . A close up of A-323 embedded in the φ pocket is shown in Figure 4.4.

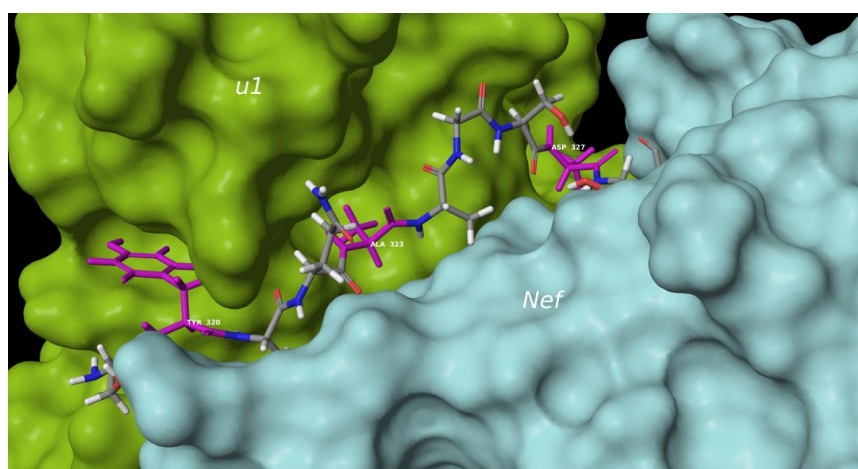


Figure 4.3. The crystal structure of the Nef-MHC- $\mu 1$ ternary complex (PDB code: 4EN2) shows the exquisite co-evolution between virus and host. Both Nef and $\mu 1$ form a clamp-like groove that binds the MHC-I CD and effectively pulls the attached protein from the plasma and endosomal membranes. Hot spot residues Y-320, A-323, and D-327 on the MHC-I CD are colored magenta.

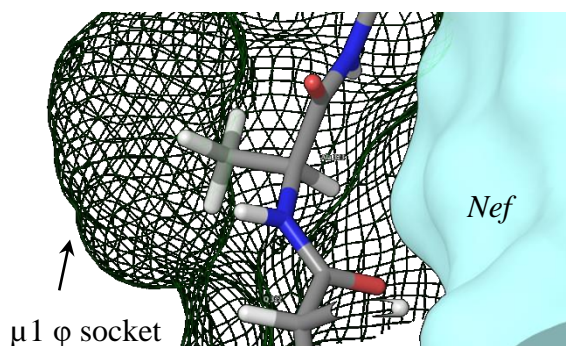


Figure 4.4. A close up view of the φ recognition socket occupied by A-323 of the MHC-I CD. The pocket is largely unoccupied by the small methyl side chain of alanine which precludes binding under normal cellular conditions. Nef's presence prevents dissociation of this residue and thus recodes the YxxA sorting signal to read Yxx φ .

This fascinating display of viral evolution underpins the importance of virus-host interactions in HIV immune evasion and illuminates new potential antiviral targets that may complement existing HAART regimens. It is thought that pharmacological inhibition of the Nef- μ 1 binding groove may restrict HIV's access to the secretory pathway and attenuate Nef's ability to down modulate MHC-I and other immunoregulatory surface receptors that support a robust antiviral response. It is believed that such inhibition, in combination with HAART and anti-latency drugs, may bolster the localized immune response against reawakened HIV and promote the clearance of reactivated virus infected cells *in vivo*. This is the motivation for the current work that will now be presented.

4.2 HTVS Efforts to Identify Novel Nef Inhibitors

In an attempt to develop a small molecule that binds to the Nef-AP-1 binding groove, 4EN2 was retrieved from the PDB and the MHC-I chain was subsequently removed to prepare the resulting Nef- μ 1 binding groove for high-throughput virtual screening (HTVS) using grid-based ligand docking with energetics (Glide).¹³⁸⁻¹⁴⁰ Crystallized water molecules were removed from the protein before subjecting the site of interest to more than 50,000 compounds from the Maybridge collection. MMGBSA calculations were performed on the top 1000 hits to predict relative $\Delta G_{\text{binding}}$ and ligand strain energy for each hit using Schrodinger's Prime software package.

The results of this initial screen proved unsatisfactory and necessitated a deeper analysis of the crystal structure. It was recognized that all of the hot-spot residues on the MHC-I CD were embedded into the μ 1 surface, not Nef, suggesting that a majority of the binding energy of the ternary complex stems from the MHC-I- μ 1 interaction rather than the interaction between μ 1 and Nef or Nef and MHC-I. This hypothesis is supported by several literature studies that indicate residues Y-321, A-324, and D-327 on the MHC-I CD are absolutely required for Nef-mediated MHC-I downregulation¹⁴¹⁻¹⁴³ and their role in maintaining the structural integrity of the ternary complex is clearly evident in the crystal structure in Figure 4.3 (shown in magenta). To this end, Nef was removed from all subsequent

Table 4.1. Predicted binding affinity of various truncated MHC-I peptides to the surface of $\mu 1$

Entry	MHC-1 Tail Sequence	Predicted Relative Ligand Binding (kcal/mol)	H-bonds to Receptor	Ligand Strain Energy (kcal/mol)
1	SYSQAAGSDSAQ	-141.906	8	16.064
2	SYSQAAGSDSA	-122.479	8	26.471
3	SYSQAAGSDS	-129.521	11	22.268
4	SYSQAAGSD	-134.926	8	9.571
5	SYSQAAGS	-113.780	6	7.553
6	SYSQAA	-103.665	7	8.979
7	YSQAA	-97.618	7	8.042
8	SYSQA	-96.505	5	4.550
9	YSQA	-90.042	6	3.802

docking studies and the investigation turned toward targeting solely the $\mu 1$ face of AP-1 instead of the Nef- $\mu 1$ heterodimer itself.

Before additional docking studies were pursued, it was necessary to evaluate the relative binding energy of the native MHC-I peptide in complex with $\mu 1$ to develop a molecule with a similar predicted binding affinity. As such, the 12-mer crystallized MHC-I CD was docked into the $\mu 1$ face of AP-1 “as is” followed by MMGBSA calculations. The native ligand was predicted to bind to the $\mu 1$ surface with a binding affinity of -142 kcal/mol. It should be noted that the numerical result of this calculation is completely arbitrary and invokes little thermodynamic insight when left unaccompanied. As such, the calculation was repeated for a series of peptides derived from the crystallized MHC-1 CD to evaluate the contributions of various residues to the total binding energy. These results are summarized in Table 4.1.

The data presented in Table 4.1 support the notion that shorter peptides have fewer contact points with the protein surface and therefore exhibit a reduction in binding energy when compared to their elongated congeners. However, this trend is skewed by entries 3 and 4 which warrant explanation. For entry 3, minor conformational changes in D-327 and S-328 during the post-docking minimization process resulted in the formation of three additional hydrogen bonds (11 total) between MHC-I and the $\mu 1$ surface that are not present in entries 1 and 2. This accounts for the larger than expected

binding energy. With regards to entry 4, the removal of S-328 results in an unexpected increase in predicted ΔG_{bind} similar to the value in entry 1, but does not result in the formation of any new hydrogen bonds, significant conformation changes, or the strengthening of pre-existing hydrogen bonds as determined by bond length and orientation. This apparent anomaly is readily explained when ligand strain energies are considered. Deletion of S-328 alleviates inherent conformational stresses of the peptide in entry 3 which negatively influences binding affinity. Smaller ligand strain energies thus correspond to an enhanced binding affinity within a series if no additional contacts are made. Conversely, the relatively large ligand strain energy observed in entry 2 is likely responsible for such a dramatic decrease in binding affinity upon deletion of glutamine from entry 1. Taken together, this series serves as a reference for identifying competitive binders that may block MHC-I recruitment to the surface of $\mu 1$.

With this fiducial series in hand, additional screenings were performed using the Aldrich and Princeton collections against $\mu 1$. A majority of the sampled compounds exhibited undesirable MMGBSA scores of higher than -70 kcal/mol, however, three targets scored exceptionally well relative to the series presented in Table 4.1 with an avg. MMGBSA score of -100 kcal/mol. These targets, **4.1**, **4.2**, and **4.3**, closely resemble the peptides in Table 1 in that all bear a tyrosine moiety and a hydrophobic appendage that accommodate the tyrosine and φ pockets on the surface of $\mu 1$, respectively. Compounds **4.1**, **4.2**, and **4.3** were purchased from Aldrich and submitted for testing in collaboration with Dr. Eric Hunter's laboratory at the Emory HIV Vaccine Center. Two assays are currently being developed in parallel to assess their biological activity against HIV-1 Nef.

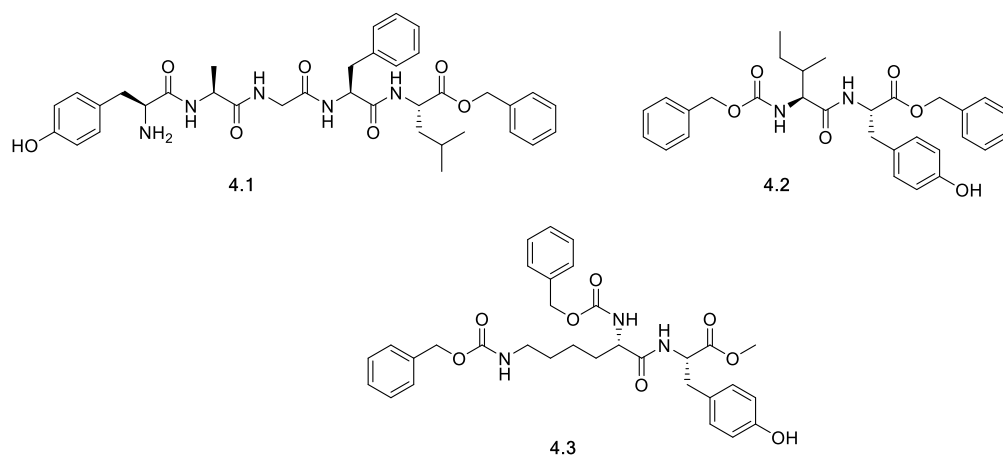


Figure 4.5. Structures of the three hit compounds from the Aldrich collection that have been submitted for testing.

4.3 *In Silico* Design of Novel α -Helical Mimetics

While 4.1-4.3 compounds await testing, it was quickly realized that the peptidic nature of 4.1-4.3 may hamper their development beyond academic interest. This prompted the design of non-peptidic small molecules that resist proteosomal degradation *in vivo*. Given that the MHC-I CD is the native ligand for the Nef- μ 1 complex, a ligand-based approach was pursued. Inspection of the shape of the crystallized MHC-I CD revealed a pseudo- α -helical structure with residues D-237, A-324, and Y-321 located at the i , $i+4$, and $i+7$ positions on the helix, respectively, shown in Figure 4.6a and resulted in the pursuit of non-peptidic α -helical mimetics.

In 2012, Cummins and Hamilton reported the synthesis of water soluble 5-6-5 imidazole-phenyl-thiazole based α -helical mimetics (Figure 4.6).¹⁴⁴ The relatively facile synthesis of these compounds –combined with their superior aqueous solubility profile over their terphenyl-based predecessors –made these particular α -helical mimetics an attractive scaffold for further development.

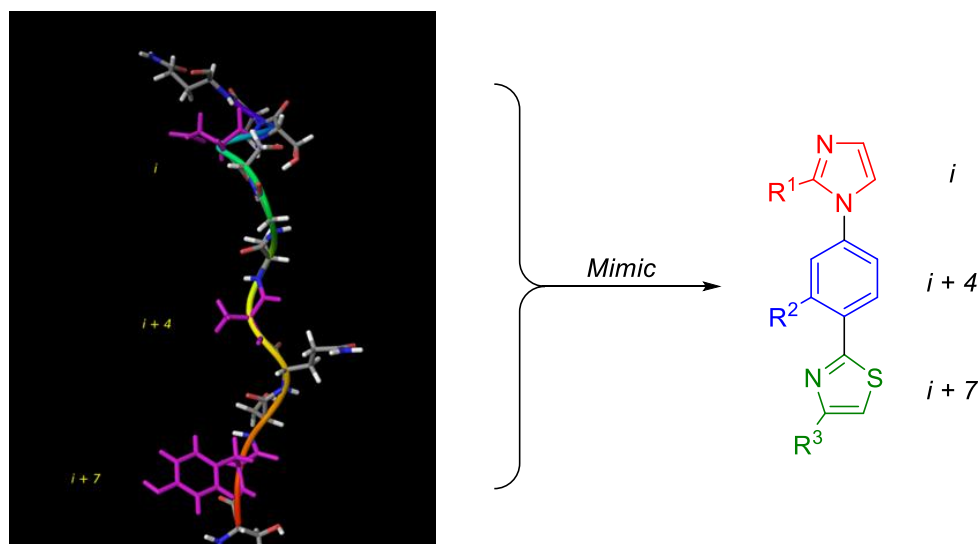


Figure 4.6. The pseudo- α -helical nature of the crystallized MHC-I CD (Nef and μ 1 have been removed for clarity). Key residues D-237, A-324, and Y-321 are colored magenta. The design strategy to capture these critical elements in a small molecule is shown on the right.

To this end, several 5-6-5 imidazole-phenyl-thiazole analogues were rendered *in silico* with a variety of appendages at R^1 , R^2 , and R^3 . After some additional computational preparation these molecules were conformationally sampled and docked onto the surface of μ 1 and ranked according to MMGBSA score. It was initially thought that R^3 and R^2 would occupy tyrosine- and φ binding pockets, respectively as these appendages correspond to the $i + 7$ and $i + 4$ positions on the ligand. However, due to the pseudo α -helical nature of the MHC-I CD, the R^3 and R^1 appendages, not R^2 and R^1 , occupy the tyrosine and φ binding pockets of μ 1, respectively as shown in Figure 4.7. Only a handful of 5-6-5 imidazole-phenyl-thiazole targets bound to the surface of μ 1 with comparable binding affinity to the truncated MHC-I CD peptides in Table 1 (entries 6-9). Target **4.4** (MMGBSA score = -95.6 kcal/mol) was initially chosen for synthesis as a proof-of-concept to determine if we could effectively capture the pseudo- α -helical properties of the MHC-I CD in a small molecule with suitable binding energy to target the Nef- μ 1 binding groove.

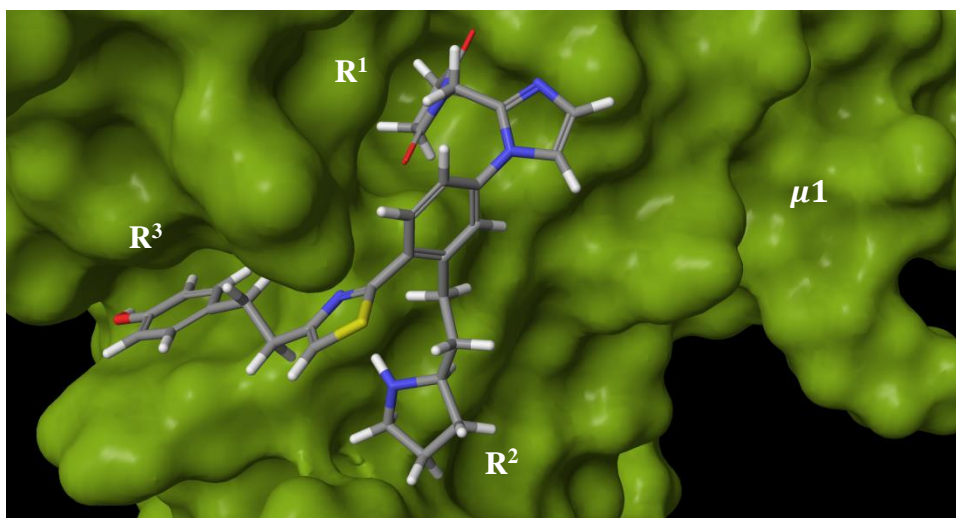
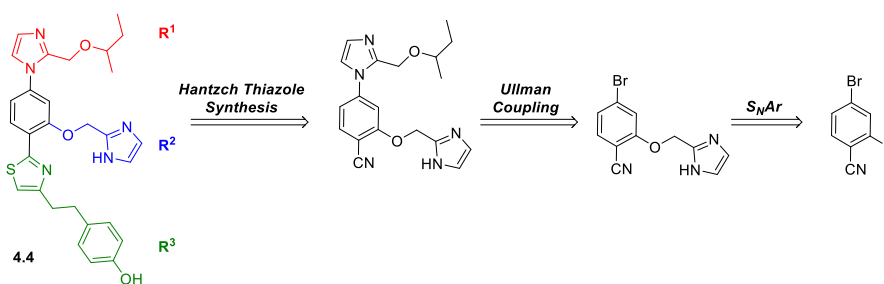


Figure 4.7. A representative pose of a 5-6-5 imidazole-phenyl-thiazole mimetic bound to the surface of $\mu 1$. Note that R^1 and R^3 occupy the two binding pockets while R^2 provides additional contact to the protein surface.

4.4 Retrosynthetic Analysis of Compound 4.4

A retrosynthetic analysis of **4.4** is presented in Scheme 4.1 with three key steps. S_NAr displacement of fluoride installs the R^2 substituent while an Ullman coupling and a Hantzsch thiazole reaction builds the northern and southern regions of **4.4**, respectively.

Scheme 4.1. Retrosynthetic analysis of compound **4.4**.



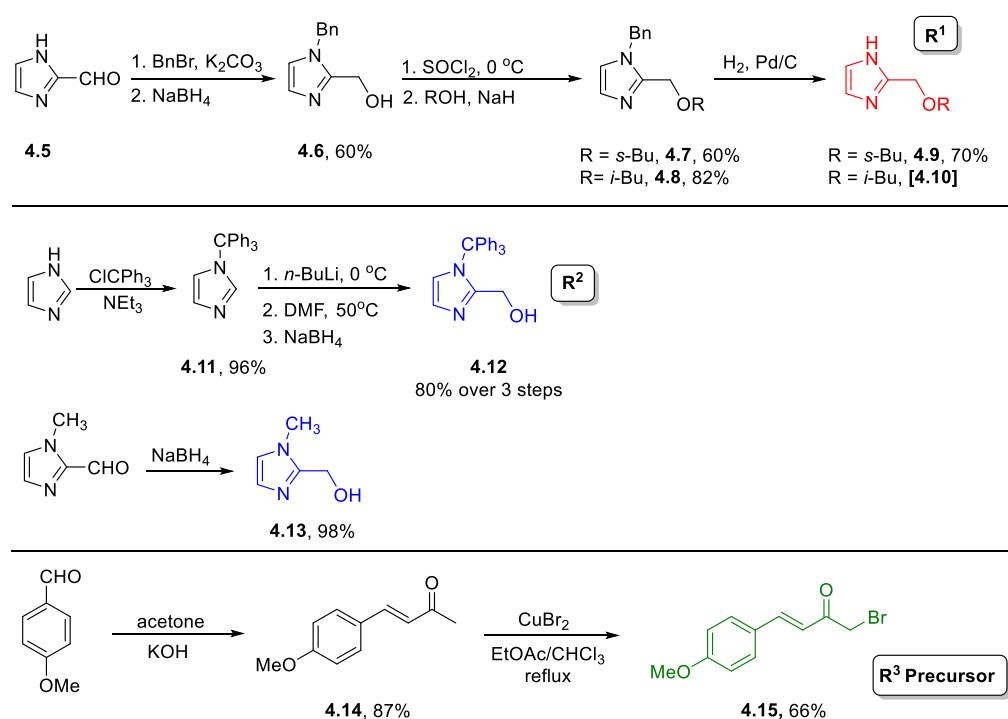
4.4.1 Synthesis of the R^1 , R^2 , and R^3 Fragments

The functionalized imidazole core containing R^1 was prepared according to the route outlined in Scheme 4.2 (top line). Protection of 1*H*-imidazole-2-carbaldehyde with benzyl bromide followed

by NaBH₄ reduction furnished compound **4.6** which was then converted to the alkyl chloride upon reaction with thionyl chloride at 0 °C. This intermediate was subjected to a Williamson ether synthesis protocol with *sec*-butanol or isobutanol and sodium hydride to afford **4.7** and **4.8**, respectively. Final reductive de-benzylation of the protecting group yielded the desired fragments **4.9** and **4.10**.

The R² fragment **4.12** was readily prepared from imidazole in a four step reaction sequence shown in the middle line of Scheme 4.2. Briefly, imidazole was converted to trityl imidazole and subsequently formylated with DMF at elevated temperature. Reduction of the intermediate aldehyde with NaBH₄ yielded **4.12**. Compound **4.13** was also prepared upon reduction. The desired R³ Hatzch thiazole precursor was prepared over a two-step sequence from 4-methoxybenzaldehyde (Scheme 4.2, last line). Intermediate **4.14** was easily obtained from the condensation of 4-methoxybenzaldehyde with acetone and subsequent α -halogenation was achieved using CuBr₂ in refluxing chloroform/ethyl

Scheme 4.2. Synthesis of the R¹, R², and R³ Fragments.

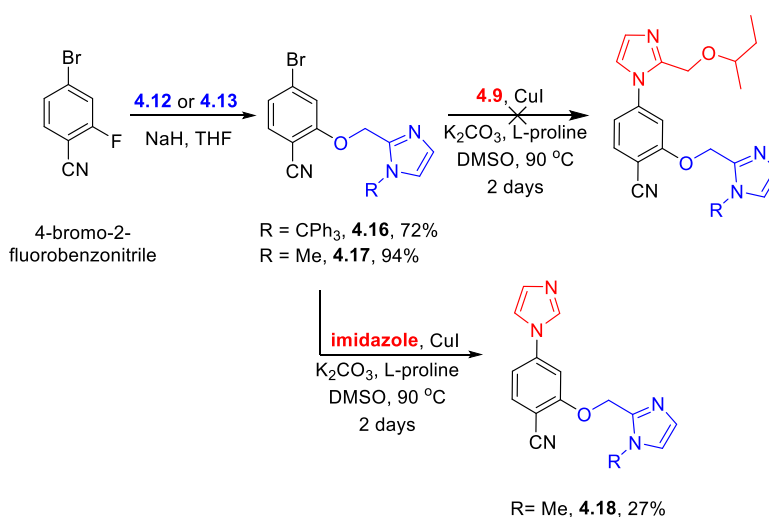


acetate to furnish **4.15**. With the desired fragments in hand, the three coupling reactions presented in Scheme 4.1 were then attempted.

4.4.2 Towards the Synthesis of Compound 4.4

The assembly of fragments R^1 , R^2 , and R^3 to the 4-bromo-2-fluorobenzonitrile core began with S_NAr fluoride displacement shown in Scheme 4.3 to afford **4.14** in good yield. The observed chemoselectivity for fluorine over bromine displacement is a kinetic phenomenon that stems from the inherent polarization of the C-F bond and the stabilization incurred by the electronegative fluorine atom in the Meisenheimer complex.¹⁰⁶ Unfortunately, Ullman coupling of **4.16** with fragment R^1 (**4.9**) failed to forge the desired N-C linkage. Extended reaction times resulted in degradation and additional CuI had no appreciable effect. To determine if steric encumbrance around the reactive nitrogen in **4.9** hinders the reaction, we chose to couple imidazole with **4.17** directly. These efforts were successful, but only barely as the yield was exceptionally poor. In light of these findings, we chose to change our approach and install R^1 using the S_NAr protocol shown in Scheme 4.4. 2-bromo-4-fluorobenzonitrile was reacted with **4.10** and heated with Na_2S in *t*-BuOH at 90 °C to produce thioamide **4.20** in 56% over two steps. Subsequent heating of **4.20** in the presence of **4.14** successfully furnished thiazole **4.21**.

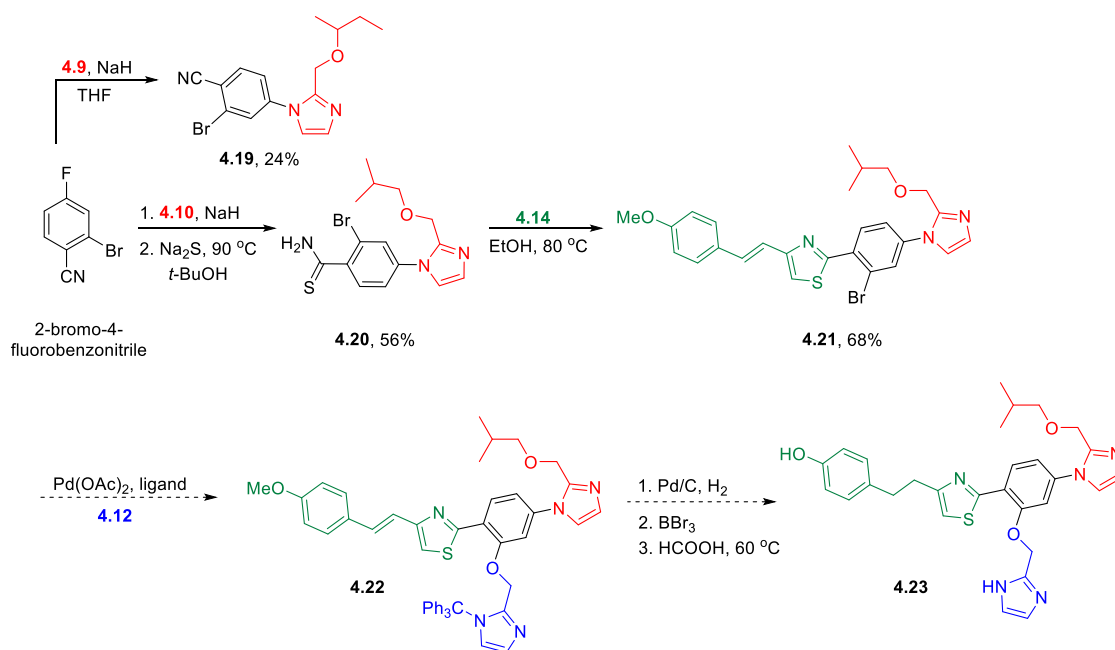
Scheme 4.3. Attempted Installation of R^1 and R^2 .



It was envisioned that aryl ether **4.22** could be constructed from a Pd-catalyzed C-O bond forming reaction¹⁴⁵ between **4.12** and **4.21**. Reduction of the olefin moiety, followed by demethylation and de-tritylation would provide access to **4.23** and related analogues, including compound **4.4**.

The synthesis described in Scheme 4.4 is currently underway and data for **4.1-4.4** and **4.23** will be collected and reported in due course with our collaboration with the Emory Vaccine Center. It is hoped that these compounds will display micromolar activity against Nef-mediated MHC-1 downregulation in a cell-based assay and that future analogues can be accessed using the synthetic protocol outlined in Scheme 4.4 to improve activity towards CTL recognition of virus-harboring cells.

Scheme 4.4. Alternative Strategy Towards the Synthesis of **4.4** and Related Analogues.



4.5 Conclusions

A sterilizing cure for HIV-1 remains to be found due to the presence of long-lived latent viral reservoirs. CTLs play a critical role in the elimination of latent HIV-1 provirus and this response may be bolstered by both vaccination and chemical therapies. The inhibition of HIV Nef may facilitate

CTL recognition and subsequent removal of HIV-infected cells. This work aims to design and synthesize a series of small molecules to disrupt Nef-mediated MHC-I downregulation by targeting the AP1-Nef interface. These efforts are currently in motion and the performance of these compounds will be reported in the next several months. We believe that vaccination, in combination with HAART, HDAC inhibitors, and a potent Nef inhibitor, may bring us one step closer to a silver bullet that has the capacity to hobble HIV's edge and cure AIDS.

4.6 Experimental Details

4.6.1 Computational Analysis

Maybridge Library Screen (4EN2 entire complex)

Protein Preparation & Docking: 4EN2 was retrieved from the PDB and the entire protein was optimized using the Protein Wizard in Schrodinger's Maestro package. Briefly, all hydrogen bonds were optimized for all three proteins (μ 1, Nef, and MHC-I CD) using the interactive optimizer and all water molecules were removed from the structure. The MHC-I CD was selected as the ligand and all docking parameters were set to defaults using GLIDE. The entire Maybridge Collection was then prepared using LigPrep and docked into the target area. Subsequent MMGBSA calculations also set to default parameters.

Maybridge Library Screen (μ 1 only)

Protein Preparation & Docking: 4EN2 was retrieved from the PDB and Nef was removed from the complex. Both μ 1 and MHC-I CD were optimized using the Protein Wizard in Schrodinger's Maestro package. Briefly, all hydrogen bonds were optimized using the interactive optimizer and all water molecules were removed from the structure. This grid is referred to as u1_grid2. The MHC-I CD was selected as the ligand and all docking parameters were set to defaults using GLIDE. The

entire Maybridge Collection was then prepared using LigPrep and docked into the target area on $\mu 1$. Subsequent MMGBSA calculations were performed in duplicate: once with default parameters and another calculation that grants residues within 3 Å of the ligand flexibility to optimize MMGBSA score. Both calculations proved unsatisfactory and no compounds from the Maybridge Library were selected for testing. For rigid MMGBSA calculations, the tightest binder exhibited an MMGBSA score of -66 kcal/mol whereas flexible MMGBSA scoring resulted in a tightest binder with -74 kcal/mol.

Princeton and Aldrich Library Screen ($\mu 1$ only)

Protein Preparation & Docking: grid u1_grid2 was used as the prepared protein for docking. The Princeton and Aldrich collections were previously filtered using several criteria in PipeLine Pilot, prepared using LigPrep, and docked to the surface of u1_grid2 using the following constraints: ligand must bind to the electrostatic network in the Y-pocket on $\mu 1$ and the φ -binding pocket required occupation by a hydrophobic moiety. Post-docking MMGBSA scores performed which were set to their default parameters. Three hits came out of this screen and were purchased for testing (4.1-4.3)

Native Docking to the Surface of $\mu 1$ (Table 4.1)

Protein Preparation & Docking: u1_grid2 was used as the prepared protein for these sets of experiments and is prepared as previously described. The native MHC-I CD was docked in place on the crystallized surface of $\mu 1$ with ligand sampling set to “none” and all other docking settings set to their defaults. This calculation was repeated for the series of peptides presented in Table I. Subsequent MMGBSA calculations were performed in duplicate: one with default parameters and another calculation that grants residues within 3 Å of the ligand flexibility to optimize MMGBSA score. The results presented in Table 1 show the rigid MMGBSA scores from these experiments.

Docking Peptidomimetics to the Surface of μ 1 (Experiment 1)

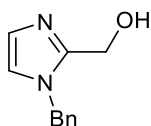
Protein Preparation & Docking: u1_grid2 was used as the prepared μ 1 protein and is prepared as previously described. Docking of several in silico generated 5-6-5 imidazole-phenyl-thiazole α -helical mimetics was performed with different appendages at R1, R2, and R3. Two constraints were imposed during docking: ligand must bind to the electrostatic network in the Y-pocket on μ 1 and the φ -binding pocket required occupation by a hydrophobic moiety. The α -helical mimetics were prepared using LigPrep and imported into Glide. Exhaustive conformational searches not performed prior to docking. Post-minimization poses set to 10 with 5 poses per ligand. Rigid MMGBSA calculations performed after docking. Results indicate that these structures do indeed fit into the two-pronged binding pockets of μ 1 and prompted further experimentation with the identities of R¹, R², and R³ to maximize binding affinity.

Docking Peptidomimetics to the Surface of μ 1 (Experiment 2)

Protein Preparation & Docking: u1_grid2 was used as the prepared μ 1 protein and is prepared as previously described. Additional 5-6-5 imidazole-phenyl-thiazole α -helical mimetics were prepared in silico with different appendages for R¹, R², and R³ than those explored in Experiment 1. Docking details identical to those found in Experiment 1. However, MMGBSA scores were performed both rigidly and flexibly as previously described. These targets bound the surface of μ 1 better than the original targets in Experiment 1. A large library of 5-6-5 imidazole-phenyl-thiazole α -helical mimetics was desired to further explore other unique side chains. Instead of pursuing this endeavor manually, other programs such as Combiglide, Pipeline Pilot, and others were sought to produce a large enumerated library of peptidomimetics. Targets from these studies ultimately proved fruitful and were selected for synthesis.

4.6.2 Chemical Synthesis

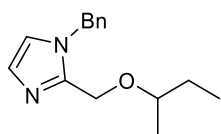
All reagents were obtained from commercial suppliers and used without further purification unless otherwise specified. Reaction progress was monitored by either thin layer chromatography (TLC) using pre-coated aluminum-backed silica gel plates (60 F₂₅₄ Merk, article 5554), by NMR, or liquid chromatography-mass spectrometry (LC-MS) on an Agilent Technologies 6100 quadrupole instrument equipped with UV detection at 254 and 210 nm and a Varian C8 analytical column. LC-MS analyses were performed using a step-wise H₂O/MeOH gradient with the % MeOH increasing from 75-95% over the course of 3 min unless otherwise specified. Flash column chromatography was conducted using CombiFlash Rf 200 (Telendyne-Isco) automated flash chromatography system with hand-packed RediSep columns. Evaporation of solvents was carried out on a rotary evaporator under reduced pressure and under ultra-high vacuum (UHV) where appropriate. ¹H NMR and ¹³C NMR spectra were recorded at ambient temperature on a Varian 400 spectrometer. ³¹P spectra were recorded at ambient temperature on either a Mercury 300 or Varian 400 spectrometer. Unless otherwise specified, all NMR spectra were obtained in deuterated chloroform (CDCl₃) and referenced to the residual solvent peak. Chemical shifts are given in δ values and coupling constants are reported in hertz (Hz). Melting points were determined on a MelTemp melting apparatus and are uncorrected. High resolution mass-spectra (HRMS) were acquired on a VG 70-S Nier Johnson or JEOL mass spectrometer.



(1-Benzyl-1H-imidazol-2-yl)methanol (4.6)

To a stirring solution of 1H-imidazole-2-carbaldehyde (0.500 g, 5.20 mmol) and benzyl bromide (0.80 mL, 6.76 mmol) in acetonitrile was added potassium carbonate (1.44 g, 10.41 mmol). The reaction stirred at 40 °C under nitrogen for 5 h. Significant conversion was noted by LC-MS and no starting

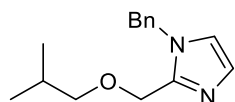
material remained. The mixture was removed from heat, diluted with brine, and extracted into DCM. The organic layer was collected, dried over sodium sulfate, and filtered. After being concentrated under reduced pressure, the resulting oil was dissolved in 50 mL DCM/MeOH (1:1) and chilled to 0°C. Then, sodium borohydride (0.197 g, 5.20 mmol) was added and the solution naturally warmed to room temperature over the course of an hour. Upon completion, the reaction was again chilled and quenched with water (15 mL). The organic solvents were evaporated under reduced pressure and the crude was extracted into DCM after being washed with brine. The organic layer was collected, dried over sodium sulfate, concentrated, and purified via flash chromatography using a silica column and a DCM/DCM:MeOH:NH₄OH (90:10:0.1) gradient (0-60%) to afford the title compound (1-benzyl-1*H*-imidazol-2-yl)methanol (0.583 g, 3.10 mmol, 59.5 % yield) as an off-white solid. ¹H NMR (400 MHz, CDCl₃) δ 7.35 – 7.26 (m, 2H), 7.21 (s, 1H), 7.18 – 7.10 (m, 2H), 6.84 (d, *J* = 1.3 Hz, 1H), 6.77 (d, *J* = 1.3 Hz, 1H), 5.22 (s, 2H), 4.62 (s, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 148.2, 136.4, 128.9, 128.0, 127.2, 126.7, 120.5, 55.6. HRMS (ESI) *m/z* calculated for C₁₁H₁₃N₂O [M + H]⁺ : 189.1022, found 189.1023.



1-Benzyl-2-(sec-butoxymethyl)-1*H*-imidazole (4.7)

To stirring thionyl chloride (2.31 mL, 31.9 mmol) in DCM at 0°C was added (1-benzyl-1*H*-imidazol-2-yl)methanol (2.00 g, 10.6 mmol). The reaction stirred at 0°C for 30 min then naturally warmed to room temperature. Progress was monitored by LC-MS. Reaction complete within an hour. The solution was quenched with aqueous sodium bicarbonate until the pH reached ~5. Then, the product was extracted into DCM and dried over sodium sulfate. The organic layer was concentrated to approximately 50 mL under reduced pressure and slowly added to a stirring solution of butan-2-ol (10

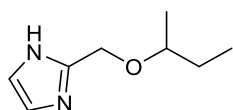
mL, 109 mmol) and the corresponding sodium alkoxide in THF (prepared fresh from 2-butanol and sodium hydride (2.55 g, 63.8 mmol). After addition, the mixture stirred for 10 min, then sodium iodide (0.796 g, 5.31 mmol) was added and the reaction was heated to 40°C. Progress was monitored by LC-MS. Complete conversion was observed after 30 min. The mixture was cooled to room temperature, quenched with aqueous ammonium chloride, and the crude extracted into DCM. The organic layer was collected, dried over sodium sulfate and purified on a silica column using DCM/DCM:MeOH:NH₄OH (90:10:0.1) (0-20%) with UV monitoring at 254 and 210 nm to afford the title compound 1-benzyl-2-(sec-butoxymethyl)-1*H*-imidazole (1.56 g, 6.40 mmol, 60.2 % yield) as a yellow oil. ¹H NMR (400 MHz, CDCl₃) δ 7.34 – 7.21 (m, 3H), 7.13 – 7.05 (m, 2H), 6.95 (d, *J* = 1.0 Hz, 1H), 6.83 (s, 1H), 5.19 (s, 2H), 4.51 (qd, *J* = 12.4, 0.8 Hz, 2H), 3.46 – 3.33 (m, 1H), 1.58 – 1.31 (m, 2H), 1.08 (dd, *J* = 6.1, 0.8 Hz, 3H), 0.81 (td, *J* = 7.3, 0.7 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 145.2, 136.5, 128.8, 127.9, 127.7, 127.0, 121.1, 76.1, 62.6, 49.6, 28.9, 18.8, 9.6. HRMS (ESI) *m/z* calculated for C₁₅H₂₁N₂O [M + H]⁺: 245.1648, found 245.1648.



1-Benzyl-2-(isobutoxymethyl)-1*H*-imidazole (4.8)

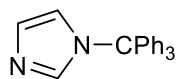
To stirring thionyl chloride (9.25 mL, 128 mmol) in DCM (100 mL) at 0°C was slowly added (1-benzyl-1*H*-imidazol-2-yl)methanol (8.00 g, 42.5 mmol). The reaction stirred at 0°C for 30 min, then naturally warmed to room temperature. Progress was monitored by LC-MS. The solution was quenched with saturated aqueous sodium bicarbonate until the pH reached ~5. Then, the product was extracted into DCM and dried over sodium sulfate. The organic layer was concentrated to approximately 50 mL under reduced pressure and slowly added to a stirring solution of 2-methylpropan-1-ol (63.0 g, 850 mmol) and the corresponding sodium alkoxide in THF (prepared fresh from 2-butanol and sodium hydride (10.2 g, 255 mmol)). After addition, the mixture stirred for 10

min at room temperature, then sodium iodide (3.19 g, 21.25 mmol) was added and the reaction was heated to 40°C. Progress was monitored by LC-MS. Upon completion, the mixture was cooled to room temperature, quenched with aqueous ammonium chloride, and the crude extracted into DCM. The organic layer was collected, dried over sodium sulfate and purified on a silica column using DCM/DCM:MeOH:NH₄OH (90:10:0.1) (0-20%) with UV monitoring at 254 and 210 nm to afford the title compound 1-benzyl-2-(isobutoxymethyl)-1*H*-imidazole (8.50 g, 34.8 mmol, 82 % yield) as a yellow oil which was taken directly to the next step.



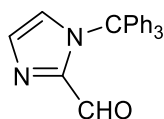
2-(*sec*-Butoxymethyl)-1*H*-imidazole (4.9)

To a stirring solution of 1-benzyl-2-(*sec*-butoxymethyl)-1*H*-imidazole (2.00 g, 8.19 mmol) in EtOH (50 mL) was added an equal weight of 10 wt. % palladium on carbon (2.00 g, 1.879 mmol) followed by the addition of ammonium formate (2.58 g, 40.9 mmol) in a single portion and the reaction was heated to 80°C. After one hour, an additional 5 equiv. of ammonium formate was added. Progress was monitored by LC-MS. Upon completion, the mixture was cooled to room temperature, filtered over celite, and washed with DCM. The supernatant was then collected and filtered over celite a second time to reduce contamination with Pd. The organic layer was collected and the solvents were evaporated under reduced pressure. The crude product was purified on a silica column using a DCM/DCM:MeOH:NH₄OH (90:10:0.1) gradient to afford the title compound 2-(*sec*-butoxymethyl)-1*H*-imidazole (0.884 g, 5.73 mmol, 70.1 % yield) as a yellow oil. ¹H NMR (400 MHz, CDCl₃) δ 7.00 (s, 2H), 4.70 - 4.50 (m, 2H), 3.43 (q, *J* = 6.1 Hz, 1H), 1.61 - 1.34 (m, 2H), 1.10 (d, *J* = 6.1 Hz, 3H), 0.83 (t, *J* = 7.4 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 145.9, 121.9, 63.6, 28.9, 18.9, 9.6, 1.0. HRMS (ESI) *m/z* calculated for C₈H₁₅N₂O [M+ H]⁺ : 155.1170, found 155.1171.



1-Trityl-1*H*-imidazole (4.11)

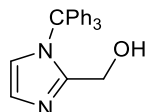
A round bottom flask was charged with 1*H*-imidazole (4.95 g, 72.8 mmol) and (chloromethanetriyl)tribenzene (20.2 g, 72.8 mmol) and dissolved in DCM (200 mL) at 0 °C. Upon dissolution, triethylamine (10.68 mL, 72.8 mmol) was added drop-wise over a thirty min interval. The solution was then removed from the ice bath and stirred for approximately 4 h at room temperature. Evaporation of the solvent under reduced pressure and re-crystallization in cold ethanol afforded the product 1-trityl-1*H*-imidazole (21.0 g, 67.7 mmol, 93% yield) which required no further purification. ¹H NMR (600 MHz, CDCl₃) δ 7.47 (s, 1H), 7.36 – 7.31 (m, 9H), 7.14 (m, 7H), 7.07 (s, 1H), 6.83 (s, 1H). ¹³C NMR (151 MHz, CDCl₃) δ 142.7, 139.2, 130.0, 128.5, 128.2, 121.9, 75.4. HRMS (ESI) *m/z* calculated for C₂₂H₁₉N₂ [M + H]⁺ :311.1542, found 311.11538.



1-Trityl-1*H*-imidazole-2-carbaldehyde (SI-23)

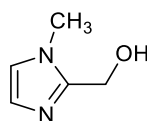
A 250 mL flame dried round bottom flask was charged with 1-trityl-1*H*-imidazole (6.87 g, 22.1 mmol) and placed under argon atmosphere followed by the addition of THF (100 mL). The reaction contents were stirred and subsequently cooled to 0 °C whereupon butyllithium (10.6 mL, 26.6 mmol) was added drop-wise to produce a red solution. The solution was naturally warmed to room temperature and stirred for an hour and then reintroduced to the ice bath prior to the drop-wise addition of *N,N*-dimethylformamide (5.14 mL, 66.4 mmol). After an additional 15 min at 0 °C, the solution was heated to 50 °C and stirred for an additional 30 min, quenched with saturated NH₄Cl, and washed with brine. The yellow organic layer was collected and dried over magnesium sulfate. Filtration and subsequent evaporation of the solvent afforded the pale yellow solid 1-trityl-1*H*-imidazole-2-carbaldehyde (6.25

g, 18.5 mmol, 83 % yield) which was used directly without purification. Spectral data is consistent with the literature. ^1H NMR (400 MHz, CDCl_3) δ 9.24 (s, 1H), 7.36 – 7.31 (m, 9H), 7.30 (s, 1H), 7.16 – 7.10 (m, 6H), 7.03 (s, 1H). ^{13}C NMR (151 MHz, CDCl_3) δ 178.9, 142.7, 142.3, 130.0, 129.9, 129.7, 128.4, 128.3, 128.2, 127.1, 121.9, 77.0. HRMS (ESI) m/z calculated for $\text{C}_{23}\text{H}_{21}\text{N}_2\text{O}$, 339.1491 $[\text{M} + \text{H}]^+$, found 341.1491.



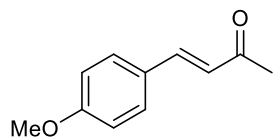
(1-Trityl-1*H*-imidazol-2-yl)methanol (4.12)

A round bottom flask was charged with 1-trityl-1*H*-imidazole-2-carbaldehyde (9.80 g, 29.0 mmol) and dissolved in DCM/MeOH (50:50) to produce an orange/brown solution. The solution was cooled to 0 °C and sodium borohydride (1.09 g, 29.0 mmol) was added portion-wise. After stirring for 15 min the solution was naturally warmed to room temperature and continued stirring for 1 hour. Upon completion, the reaction was quenched with water and washed with brine. The organic layer was collected and the solvent was evaporated under reduced pressure to afford the crude product which was triturated with EtOAc and filtered. The pale yellow solid was recovered and dried under UHV to afford the pure product (1-trityl-1*H*-imidazol-2-yl)methanol (7.95 g, 23.3 mmol, 81 % yield). ^1H NMR (400 MHz, CDCl_3) δ 7.41 - 7.32 (m, 9H), 7.18 - 7.08 (m, 6H), 7.01 (dd, $J = 1.5, 0.6$ Hz, 1H), 6.81 (dd, $J = 1.7, 0.8$ Hz, 1H), 3.65 (d, $J = 5.4$ Hz, 2H), 2.97 (t, $J = 5.5$ Hz, 1H). ^{13}C NMR (101 MHz, CDCl_3) δ 148.1, 141.3, 129.6, 128.9, 128.8, 126.1, 120.5, 76.9, 56.2. HRMS (ESI) m/z calculated for $\text{C}_{23}\text{H}_{21}\text{N}_2\text{O}$, 341.1648 $[\text{M} + \text{H}]^+$, found 341.1647. Melting Point: Decomposes at 210 °C.



(1-Methyl-1*H*-imidazol-2-yl)methanol (4.13)

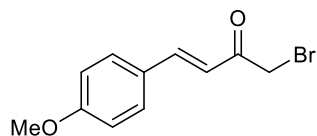
To a solution of 1-methyl-1*H*-imidazole-2-carbaldehyde (5.00 g, 45.4 mmol) in ethanol/THF (160 mL, 1:1) cooled to 0 °C was added sodium borohydride (2.00 g, 54.5 mmol). The reaction was stirred for two h at room temperature and the solvents were subsequently evaporated under reduced pressure. The resulting viscous residue was quenched with water (15 mL) and dissolved in 100 mL CHCl₃. To force the water-soluble product into the organic layer, an excess of anhydrous MgSO₄ was added to the mixture to produce a residue containing aggregates of sodium borates/boronic acid and MgSO₄. Filtration of the mixture and evaporation of the solvent afforded pure (1-methyl-1*H*-imidazol-2-yl)methanol (5.00 g, 44.3 mmol, 98 % yield). Re-crystallization in DCM afforded the title compound as a white solid. ¹H NMR (600 MHz, CDCl₃) δ 6.87 (d, 1H), 6.82 (d, *J* = 1.2 Hz, 1H), 4.63 (s, 2H), 3.73 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 148.2, 126.8, 121.7, 55.9, 33.0. HRMS (ESI) *m/z* calculated for C₅H₉N₂O [M + H]⁺ : 113.0709, found 113.0708. Melting Point: 110-113 °C.



(*E*)-4-(4-Methoxyphenyl)but-3-en-2-one (4.14)

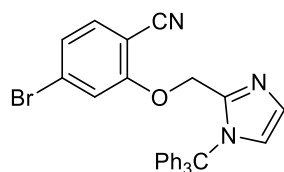
To a stirring solution of acetone (3.50 mL, 47.7 mmol) and 4-methoxybenzaldehyde (4.47 mL, 36.7 mmol) in a solution of water and acetonitrile (50:50, 25 mL) was added sodium hydroxide (1.47 g, 36.7 mmol). The reaction mixture was heated to 60 °C and stirred for 5 h after which TLC indicated completion (hexanes/EtOAc 4:1). The reaction was quenched with saturated NH₄Cl, washed with water, and extracted into DCM. Evaporation of the solvent afforded a bright yellow solid which was purified via flash chromatography using a gentle hexanes/EtOAc gradient (0-22%) to afford the title compound white solid (*E*)-4-(4-methoxyphenyl)but-3-en-2-one (16.9 g, 96.0 mmol, 87 % yield). ¹H NMR (400 MHz, CDCl₃) δ 7.52 - 7.47 (m, 2H), 7.46 (s, 1H), 6.92 (d, *J* = 8.7 Hz, 1H), 6.61 (d, *J* = 16.3 Hz, 1H), 3.85 (s, 3H), 2.36 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 198.4, 161.6, 143.3, 129.9, 127.0,

125.0, 114.4, 55.4, 27.4. HRMS (ESI) m/z calculated for $C_{11}H_{13}O_2$, $[M + H]^+$:117.0910, found 117.0911. Melting Point: 72-74°C (lit. 72-74°C).



(E)-1-Bromo-4-(4-methoxyphenyl)but-3-en-2-one (4.15)

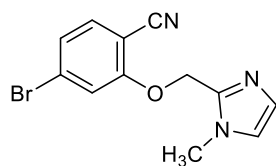
To a stirring solution of (E)-4-(4-methoxyphenyl)but-3-en-2-one (0.500 g, 2.84 mmol) in $CHCl_3/EtOAc$ (50 mL, 3:2) was added copper (II) bromide (1.15 g, 5.16 mmol). The vessel was attached to a reflux condenser and heated to 75°C for one hour. Progress was monitored by TLC (4:1 hexanes:EtOAc). Upon completion, the mixture was cooled to room temperature, filtered to remove the resulting brown cuprous residue, and concentrated under reduced pressure. The crude was purified on a silica column using a hexanes/EtOAc gradient (0-17%) to afford the title compound (E)-1-bromo-4-(4-methoxyphenyl)but-3-en-2-one (0.430 g, 1.67 mmol, 64.7 % yield) as a tan solid. 1H NMR (400 MHz, $CDCl_3$) δ 7.67 (d, $J = 16.0$ Hz, 1H), 7.55 (dd, $J = 8.7, 0.9$ Hz, 1H), 6.93 (dd, $J = 8.8, 0.8$ Hz, 2H), 6.88 - 6.79 (m, 1H), 4.08 (s, 2H), 3.86 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 191.2, 162.3, 145.4, 130.7, 126.8, 120.1, 114.7, 55.7, 33.4. HRMS (ESI) m/z calculated for $C_{11}H_{11}O_2BrNa$ $[M + Na]^+$: 276.9834, found 276.9842.



4-Bromo-2-((1-trityl-1H-imidazol-2-yl)methoxy)benzonitrile (4.16)

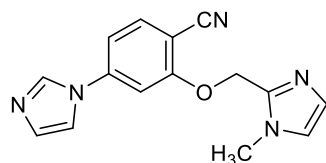
To a stirring solution of (1-trityl-1H-imidazol-2-yl)methanol (3.50 g, 10.3 mmol) and 4-bromo-2-fluorobenzonitrile (2.47 g, 12.3 mmol) in anhydrous THF at 0 °C was added sodium hydride (1.65 g, 41.1 mmol) portionwise. The reaction was stirred at that temperature for one hour then naturally

warmed to room temperature, stirred for 3 h and monitored by TLC using DCM:MeOH:NH₄OH (90:10:0.1) to afford a yellow solution. Upon completion, the solvent was evaporated under reduced pressure and the resulting residue was washed with water and extracted into chloroform. The organic layer was further treated with Et₂O and MeOH and the product recrystallized from this solution to afford 4-bromo-2-((1-trityl-1*H*-imidazol-2-yl)methoxy)benzonitrile (3.88 g, 7.46 mmol, 72.5 % yield) and was used directly in the next reaction.



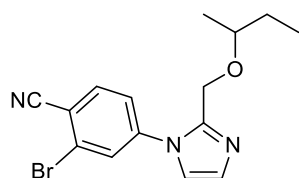
4-Bromo-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.17)

To stirring solution of (1-methyl-1*H*-imidazol-2-yl)methanol (4.00 g, 35.7 mmol) and sodium hydride (5.71 g, 143 mmol) in anhydrous THF (250 mL) at 0 °C was added 4-bromo-2-fluorobenzonitrile (7.85 g, 39.2 mmol). TLC indicated complete conversion within thirty min at that temperature (DCM:MeOH:0.1 NH₄OH). The solvent was evaporated under reduced pressure and the resulting residue was washed with water and extracted into chloroform. The organic layer was recovered, dried over MgSO₄, filtered, and diluted with Et₂O to triturate the product from the supernatant. The solid was collected and dried under vacuum to afford 4-bromo-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (9.83 g, 33.6 mmol, 94 % yield) which required no further purification. A suitable x-ray quality crystal was obtained by re-crystallizing the product from hexanes/EtOAc (60:40). ¹H NMR (600 MHz, CDCl₃) δ 7.57 (d, *J* = 1.1 Hz, 1H), 7.40 (d, *J* = 8.5 Hz, 1H), 7.20 (dd, *J* = 8.2, 1.0 Hz, 1H), 7.03 (s, 1H), 6.94 (s, 1H), 5.31 (s, 2H), 3.80 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 160.0, 141.7, 134.5, 128.3, 125.3, 123.2, 120.6, 117.5, 115.7, 101.5, 63.9, 33.5. HRMS (ESI) *m/z* calculated for C₁₂H₁₁N₃OBr, [M + H]⁺ : 292.0080, found 292.0075. Melting Point: Decomposes at 177 °C.



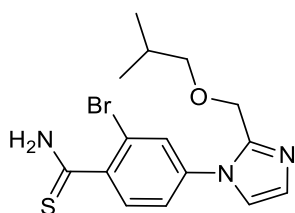
4-(1*H*-imidazol-1-yl)-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.18)

A microwave vessel equipped with a magnetic stir-bar was charged with 4-bromo-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.00 g, 13.7 mmol), 1*H*-imidazole (0.84 g, 12 mmol), pyrrolidine-2-carboxylic acid (0.28 g, 2.5 mmol), and copper(I) iodide (0.23 g, 1.2 mmol). The vessel was sealed, evacuated, and purged with nitrogen (x2). Dry DMSO (5 mL) was then added to the reaction mixture. The mixture was heated to 95 °C and stirred for 48 h. Upon completion, the vessel was cooled to room temperature and the resulting suspension was poured into water (5 mL) and CHCl₃. The bilayer was treated with 30% NH₄OH (aq) and the yellow organic layer was recovered and further washed with water to remove any residual DMSO. The organic layer was again collected, concentrated under reduced pressure, and purified via flash chromatography using a silica column and a DCM/DCM:MeOH:NH₄OH (90:10:0.1) gradient to afford 4-(1*H*-imidazol-1-yl)-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (1.04 g, 3.74 mmol, 27.3 % yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.48 (d, *J* = 1.2 Hz, 1H), 7.94 (d, *J* = 1.3 Hz, 1H), 7.90 (dd, *J* = 8.5, 1.8 Hz, 1H), 7.82 (d, *J* = 1.9 Hz, 1H), 7.47 (d, *J* = 8.5 Hz, 1H), 7.25 (d, *J* = 1.2 Hz, 1H), 7.18 (d, *J* = 1.2 Hz, 1H), 6.92 (d, *J* = 1.2 Hz, 1H), 5.49 (s, 2H), 3.74 (s, 3H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 161.0, 142.1, 142.0, 136.4, 135.7, 131.0, 127.6, 123.8, 118.4, 116.3, 113.2, 106.1, 99.0, 63.8, 33.1.



2-Bromo-4-(2-(sec-butoxymethyl)-1*H*-imidazol-1-yl)benzonitrile (4.19)

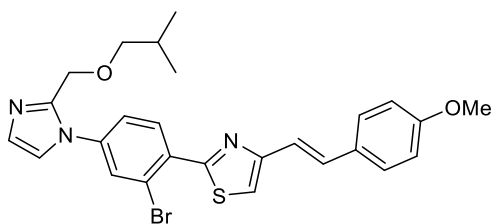
To a stirring solution of 1-benzyl-2-(sec-butoxymethyl)-1*H*-imidazole (0.650 g, 2.66 mmol) in EtOH (4 mL) was added an equal weight of 10 wt. % palladium on carbon (0.650 g, 0.611 mmol) followed by the addition of ammonium formate (0.168 g, 2.66 mmol) in a single portion and the reaction was heated to 80°C. After one hour, and additional 5 equivalents of ammonium formate was added. Complete conversion was observed by LC-MS after an hour and a half. The mixture was cooled to room temperature, filtered over celite, and washed with DCM. The organic layer was collected and the mixture was concentrated and added to a solution containing potassium carbonate (0.478 g, 3.46 mmol) and 2-bromo-4-fluorobenzonitrile (0.692 g, 3.46 mmol). The mixture was heated to 40°C under nitrogen and progress was monitored by LC-MS. After several hours, quantitative conversion was observed. The mixture was cooled to room temperature, washed with brine, and extracted into DCM. The organic layer was collected, dried over sodium sulfate and concentrated. The crude was purified on a silica column using a hexanes/EtOAc gradient (0-50%) to afford the title compound 2-bromo-4-(2-(sec-butoxymethyl)-1*H*-imidazol-1-yl)benzothioamide (0.211 g, 0.631 mmol, 23.7 % yield over two steps) as a white solid. ¹H NMR (400 MHz, CDCl₃) δ 7.97 (d, *J* = 2.1 Hz, 1H), 7.73 (d, *J* = 8.3 Hz, 1H), 7.60 (ddd, *J* = 8.4, 2.1, 0.6 Hz, 1H), 7.12 (d, *J* = 1.4 Hz, 1H), 7.07 (d, *J* = 1.4 Hz, 1H), 4.48 – 4.35 (m, 2H), 3.47 (q, *J* = 6.1 Hz, 1H), 1.60 – 1.33 (m, 2H), 1.09 (dd, *J* = 6.1, 0.6 Hz, 3H), 0.79 (dd, *J* = 7.7, 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 144.7, 141.8, 135.1, 129.3, 129.1, 126.1, 123.8, 120.9, 116.4, 114.9, 76.6, 61.5, 28.8, 18.8, 9.6. HRMS (ESI) *m/z* calculated for C₁₅H₁₇BrN₃O [M + H]⁺ 334.0549, found 334.0555.



2-Bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)benzothioamide (4.20)

To a stirring solution of 1-benzyl-2-(isobutoxymethyl)-1*H*-imidazole (2.00 g, 8.19 mmol) in EtOH (50 mL) was added an equal weight of 10 wt. % palladium on carbon (2.00 g, 1.88 mmol) followed by the addition of ammonium formate (2.58 g, 40.9 mmol) in a single portion and the reaction was heated to 80°C. After one hour, and additional 5 equivalents of ammonium formate was added. Progress was monitored by LC-MS. Upon completion, the mixture was cooled to room temperature, filtered over celite, and washed with DCM. The supernatant was then collected and filtered over celite a second time to reduce contamination with Pd. The organic layer was collected and the solvents were evaporated under reduced pressure. The resulting residue of 2-(isobutoxymethyl)-1*H*-imidazole (0.88 g, 5.7 mmol) was dissolved in THF and treated with sodium hydride (0.29 g, 7.4 mmol). The mixture stirred for 15 min, then 2-bromo-4-fluorobenzonitrile (1.71 g, 8.56 mmol) was added. The mixture stirred at room temperature. After several hours, the reaction stalled at ~50% completion. The mixture was then heated to 55°C for several h. No effect observed. Finally, DMF and potassium carbonate (excess) was added and the mixture was heated to 100°C for an hour. Complete conversion observed by LC-MS. The yellow reaction mixture cooled to room temperature and was washed with brine. The crude was extracted into DCM, dried over sodium sulfate, filtered, and concentrated under reduced pressure. The resulting oil was purified on a silica column using a hexanes/EtOAc gradient (with monitoring at 254 and 210 nm) to afford the title compound 2-bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)benzonitrile (0.495 g, 1.48 mmol, 26.0 % yield) as a white solid and pressed forward to the next step. Then, to a stirring solution of 2-bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)benzonitrile (0.500 g, 1.46 mmol) in *tert*-butanol (10 mL) was added excess aqueous ammonium sulfide (2.55 mL, 7.48 mmol). The vessel was sealed, purged with nitrogen and heated to 90°C. After several hours, complete conversion observed. The mixture was cooled to room temperature, washed with brine and the crude extracted into DCM. The organic layer was collected, dried over sodium sulfate, filtered, and purified on a silica column using a DCM/DCM:MeOH:NH₄OH (90:10:0.1)

gradient to afford the title compound 2-bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)benzothioamide (0.311 g, 0.846 mmol, 56.5 % yield) as a yellow solid. ¹H NMR (400 MHz, CDCl₃) δ 7.92 (s, 1H), 7.79 (d, *J* = 2.1 Hz, 1H), 7.70 (d, *J* = 8.3 Hz, 1H), 7.48 (ddd, *J* = 8.3, 2.1, 0.5 Hz, 1H), 7.26 (s, 1H), 7.13 – 7.06 (m, 2H), 4.41 (s, 2H), 3.27 (dd, *J* = 6.6, 0.5 Hz, 2H), 1.87 (dp, *J* = 13.3, 6.7 Hz, 1H), 0.88 (dd, *J* = 6.7, 0.5 Hz, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 201.5, 144.4, 142.5, 138.9, 130.4, 129.3, 128.8, 123.9, 121.2, 117.7, 77.3, 63.9, 28.4, 19.4. HRMS (ESI) *m/z* calculated for C₁₅H₁₉N₃OBrS [M + H]⁺: 368.0426, found 368.0431.

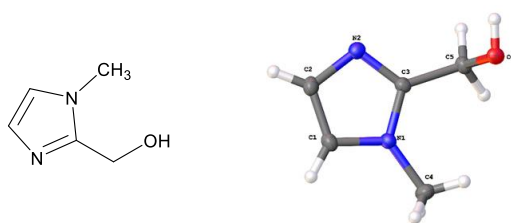


(*E*)-2-(2-Bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)phenyl)-4-(4-methoxystyryl)thiazole (4.21)

To a stirring solution of 2-bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)benzothioamide (0.250 g, 0.679 mmol) in ethanol was added (*E*)-1-bromo-4-(4-methoxyphenyl)but-3-en-2-one (0.173 g, 0.679 mmol). The mixture was purged with nitrogen and stirred at 80°C. Progress was monitored by LC-MS. After 3 h of stirring, the reaction was cooled to room temperature, extracted into DCM and washed with brine. The organic layer was collected, dried over anhydrous sodium sulfate, and purified on a silica column using a DCM/DCM:MeOH:NH₄OH gradient (90:10:0.1) to afford the title compound (*E*)-2-(2-bromo-4-(2-(isobutoxymethyl)-1*H*-imidazol-1-yl)phenyl)-4-(4-methoxystyryl)thiazole (0.243 g, 0.464 mmol, 68.3 % yield) as an orange foam. ¹H NMR (400 MHz, CDCl₃) δ 8.29 (dd, *J* = 8.4, 1.0 Hz, 1H), 7.94 (dd, *J* = 2.1, 1.0 Hz, 1H), 7.63 – 7.43 (m, 4H), 7.28 (d, *J* = 0.9 Hz, 1H), 7.16 (dt, *J* = 5.7, 1.3 Hz, 2H), 7.07 – 6.99 (m, 1H), 6.93 – 6.84 (m, 2H), 4.47 (s, 2H), 3.81 (s, 3H), 3.31 (dd, *J* = 6.5, 1.0 Hz, 2H), 1.98 – 1.82 (m, 1H), 0.92 (dd, *J* = 6.7, 1.0 Hz, 6H). ¹³C

NMR (101 MHz, CDCl₃) δ 163.7, 159.5, 154.4, 144.5, 138.7, 133.7, 132.4, 131.5, 130.0, 129.5, 128.9, 128.0, 123.8, 121.9, 121.2, 119.0, 116.3, 114.1, 77.3, 64.1, 55.3, 28.5, 19.5. HRMS (ESI) m/z calculated for C₂₆H₂₇O₂N₃BrS [M + H]⁺ : 524.1001, found 524.1004.

4.6.3 Crystallographic Data

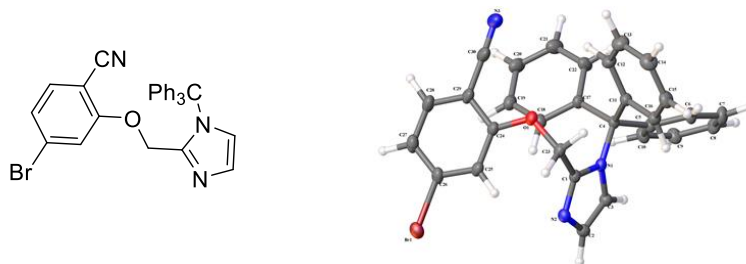


(1-Methyl-1*H*-imidazol-2-yl)methanol (4.13)

Single crystals of (1-Methyl-1*H*-imidazol-2-yl)methanol (**4.13**) were obtained by evaporation from an water and methanol. A suitable crystal (0.906×0.665×0.660 ~ mm³) was selected and mounted on a Bruker APEX-II CCD diffractometer. The crystal was kept at 109(2) K during data collection. Using Olex2⁹⁸ the structure was solved with the Superflip structure solution program, using the Charge Flipping solution method. The model was refined with the ShelXL⁹⁹ refinement package using Least Squares minimisation. Crystal Data. C₅H₈N₂O, M = 112.13, orthorhombic, Pbc_a (No. 61, a = 12.3509 Å, b = 6.9244 Å, c = 13.076 Å, $\alpha = \beta = \gamma = 90^\circ$, V = 1118.29(14) Å³, T = 109(2) K, Z = 8, μ (Mo K α) = 0.096, 4124 reflections measured, 1688 unique ($R_{\text{int}} = 0.0227$) which were used in all calculations. The final wR_2 was 0.1133 (all data) and R_1 was 0.0409 ($I > 2(I)$). Data was collected by John Bacsá, PhD at the Emory X-ray crystallography core facility. Additional crystallographic data is summarized in Table 4.2.

Table 4.2. Crystal Data for (1-methyl-1*H*-imidazol-2-yl)methanol (**4.13**)

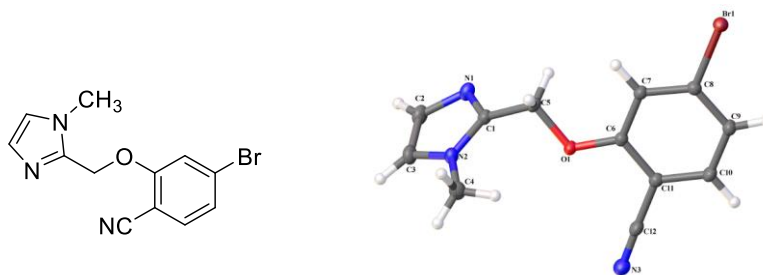
Empirical Formula	C ₅ H ₈ N ₂ O
Formula Weight	112.13
T/K	109.01
Crystal System	orthorhombic
Space Group	Pbca
a/Å	12.3509(9)
b/Å	6.9244(5)
c/Å	13.0760(10)
α/°	90
β/°	90
γ/°	90
V/Å ³	1118.29(14)
Z	8
D calc./g cm ⁻³	1.332
μ/mm ⁻¹	0.096
Theta max/°	30.480
Measured Refl.	4124
Independent Refl.	1688
Reflections Used	1562
R(int)	0.0227
Parameters	103
Restraints	0
GooF	1.063
Max Size/mm	0.906
Mid Size/mm	0.665
Min Size/mm	0.660



4-Bromo-2-((1-trityl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.16)

Table 4.3. Crystal Data and Structure Refinement for 4-Bromo-2-((1-trityl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.16)

Empirical formula	C ₃₀ H ₂₂ BrN ₃ O	
Formula weight	520.41	
Temperature	1102 K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	P -1	
Unit cell dimensions	a = 9.5754(17) Å	∠ = 89.244(3)°.
	b = 10.5422(18) Å	∠ = 88.217(3)°.
	c = 11.817(2) Å	∠ = 83.079(3)°.
Volume	1183.5(4) Å ³	
Z	2	
Density (calculated)	1.460 Mg/m ³	
Absorption coefficient	1.766 mm ⁻¹	
F(000)	532	
Crystal size	1.054 x 0.492 x 0.2 mm ³	
Theta range for data collection	1.724 to 30.831°.	
Reflections collected	19307	
Independent reflections	7324 [R(int) = 0.0317]	
Completeness to theta = 30.000°	99.6 %	
Absorption correction	None	
Max. and min. transmission	0.7461 and 0.5372	
Refinement method	Full-matrix least-squares on F ²	
Data / restraints / parameters	6534 / 0 / 448	
Goodness-of-fit on F ²	1.069	
Final R indices [I > 2σ(I)]	R1 = 0.038, wR2 = 0.1115	



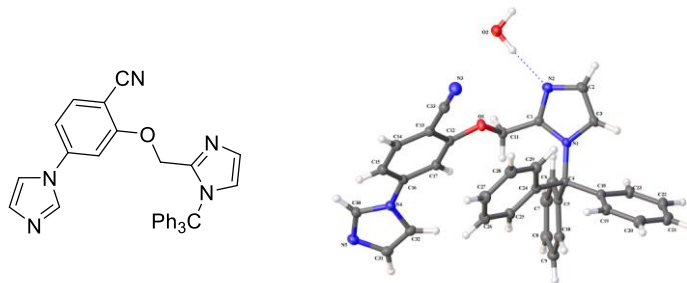
4-Bromo-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.17)

Table 4.4. Crystal Data and Structure Refinement for 4-Bromo-2-((1-methyl-1*H*-imidazol-2-yl)methoxy)benzonitrile (4.17)

CCDC	n/a
Formula	C ₁₂ H ₁₀ BrN ₃ O
$D_{calc.}/\text{g cm}^{-3}$	1.691
μ/mm^{-1}	3.568
Formula Weight	292.14
Colour	colourless
Shape	prism
Size/ mm^3	0.520 × 0.280 × 0.208
T/K	110(2)
Crystal System	triclinic
Space Group	$P\bar{1}$
a/Å	7.2328(6)
b/Å	7.6735(6)
c/Å	11.3261(9)
$\alpha/^\circ$	100.4605(9)
$\beta/^\circ$	96.0499(10)
$\gamma/^\circ$	109.3793(9)
V/Å ³	573.77(8)
Z	2
$\Theta_{min}/^\circ$	2.892
$\Theta_{max}/^\circ$	31.358
Measured Refl.	10626
Independent Refl.	3744
Reflections Used	3454
R_{int}	0.0215
Parameters	190
Restraints	10
Largest Peak Largest Peak	0.743
Deepest Hole	-0.399
GooF	1.049
wR_2 (all data)	0.0652
wR_2	0.0637
R_1 (all data)	0.0294
R_1	0.0262

Crystal Data. C₁₂H₁₀BrN₃O, M = 292.14, triclinic, $P\bar{1}$ (No. 2, $a = 7.2328(6)$ Å, $b = 7.6735(6)$ Å, $c = 11.3261(9)$ Å, $\alpha = 100.4605^\circ$, $\beta = 96.0499^\circ$, $\gamma = 109.3793^\circ$, $V = 573.77(8)$ Å³, $T = 110(2)$ K, $Z = 2$, $\mu(\text{Mo K}\alpha) = 3.568$, 10626 reflections measured, 3744 unique ($R_{int} = 0.0215$) which were used in all calculations. The final wR_2 was 0.0652 (all $I > 2\sigma(I)$), and R_1 was 0.0262 ($I > 2\sigma(I)$).

Experimental. Single crystals of C₁₂H₁₀BrN₃O (**KEG-1-30**) were recrystallised from DCM by slow evaporation. A suitable crystal (0.520 × 0.280 × 0.208 mm³) was selected and mounted on a loop in perfluorotether oil on a Bruker APEX-II CCD diffractometer. The crystal was kept at 110(2) K during data collection. Using Olex2 [1], the structure was solved with the Superflip [2] structure solution program, using the Charge Flipping solution method. The model was refined with the ShelXL [3] refinement package using Least Squares minimisation.



4-(1*H*-imidazol-1-yl)-2-((1-trityl-1*H*-imidazol-2-yl)methoxy)benzonitrile

Table 4.5. Crystal Data and Structure Refinement for 4-(1*H*-imidazol-1-yl)-2-((1-trityl-1*H*-imidazol-2-yl)methoxy)benzonitrile

CCDC	n/a
Formula	$C_{33}H_{27}N_5O_2$
$D_{calc.}/g\ cm^{-3}$	1.309
μ/mm^{-1}	0.084
Formula Weight	525.59
Colour	colourless
Shape	prism
Size/ mm^3	$0.744 \times 0.512 \times 0.407$
T/K	110(2)
Crystal System	triclinic
Space Group	$P\bar{1}$
a/Å	9.7072(6)
b/Å	10.0253(6)
c/Å	14.6984(9)
$\alpha/^\circ$	98.4976(8)
$\beta/^\circ$	104.4835(8)
$\gamma/^\circ$	100.5470(8)
$V/\text{Å}^3$	1333.11(14)
Z	2
$\Theta_{min}/^\circ$	2.110
$\Theta_{max}/^\circ$	31.134
Measured Refl.	22116
Independent Refl.	8499
Reflections Used	7505
R_{int}	0.0203
Parameters	449
Restraints	115
Largest Peak Largest Peak	0.513
Deepest Hole	-0.350
Goof	1.056
wR_2 (all data)	0.1319
wR_2	0.1239
R_1 (all data)	0.0523
R_1	0.0471

Crystal Data. $C_{33}H_{27}N_5O_2$, $M = 525.59$, triclinic, $P\bar{1}$ (No. 2, $a = 9.7072(6)\text{Å}$, $b = 10.0253(6)\text{Å}$, $c = 14.6984(9)\text{Å}$, $\alpha = 98.4976^\circ$, $\beta = 104.4835^\circ$, $\gamma = 100.547^\circ$, $V = 1333.11(14)\text{Å}^3$, $T = 110(2)\text{K}$, $Z = 2$, $\mu(\text{Mo K}\alpha) = 0.084$, 22116 reflections measured, 8499 unique ($R_{int} = 0.0203$) which were used in all calculations. The final wR_2 was 0.1319 (all data) and R_1 was 0.0471 ($I > 2\sigma(I)$).

Experimental. Single crystals of $C_{33}H_{27}N_5O_2$ (**KEG-1-37**) were recrystallised from a mixture of DCM and chloroform by vapor diffusion. A suitable crystal ($0.744 \times 0.512 \times 0.407\text{mm}^3$) was selected and mounted on a loop in perfluorotether oil on a Bruker APEX-II CCD diffractometer. The crystal was kept at 110(2) K during data collection. Using Olex2 [1], the structure was solved with the Superflip [2] structure solution program, using the Charge Flipping solution method. The model was refined with the ShelXL [3] refinement package using Least Squares minimisation.

4.7 References

- (107) Durand, C. M.; Blankson, J. N.; Siliciano, R. F., Developing Strategies for HIV-1 Eradication. *Trends Immunol.* **2012**, *33* (11), 554-62.
- (108) Harrigan, P. R.; Whaley, M.; Montaner, J. S., Rate of HIV-1 Rna Rebound Upon Stopping Antiretroviral Therapy. *AIDS* **1999**, *13* (8), F59-62.
- (109) Le, T.; Farrar, J.; Shikuma, C., Rebound of Plasma Viremia Following Cessation of Antiretroviral Therapy Despite Profoundly Low Levels of HIV Reservoir: Implications for Eradication. *AIDS* **2011**, *25* (6), 871-2; author reply 872-3.
- (110) Zhang, L.; Chung, C.; Hu, B. S., *et al.*, Genetic Characterization of Rebounding HIV-1 after Cessation of Highly Active Antiretroviral Therapy. *J. Clin. Invest.* **2000**, *106* (7), 839-45.
- (111) Iglesias-Ussel, M. D.; Romerio, F., HIV Reservoirs: The New Frontier. *AIDS Rev.* **2011**, *13* (1), 13-29.
- (112) Lassen, K. G.; Bailey, J. R.; Siliciano, R. F., Analysis of Human Immunodeficiency Virus Type 1 Transcriptional Elongation in Resting Cd4+ T Cells in Vivo. *J. Virol.* **2004**, *78* (17), 9105-14.
- (113) Adams, M.; Sharmeen, L.; Kimpton, J., *et al.*, Cellular Latency in Human Immunodeficiency Virus-Infected Individuals with High Cd4 Levels Can Be Detected by the Presence of Promoter-Proximal Transcripts. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (9), 3862-6.
- (114) Dahl, V.; Josefsson, L.; Palmer, S., HIV Reservoirs, Latency, and Reactivation: Prospects for Eradication. *Antiviral Res.* **2010**, *85* (1), 286-94.
- (115) Chun, T. W.; Stuyver, L.; Mizell, S. B., *et al.*, Presence of an Inducible HIV-1 Latent Reservoir During Highly Active Antiretroviral Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (24), 13193-7.
- (116) Siliciano, J. D.; Kajdas, J.; Finzi, D., *et al.*, Long-Term Follow-up Studies Confirm the Stability of the Latent Reservoir for HIV-1 in Resting Cd4+ T Cells. *Nat. Med.* **2003**, *9* (6), 727-8.
- (117) Poles, M. A.; Boscardin, W. J.; Elliott, J., *et al.*, Lack of Decay of HIV-1 in Gut-Associated Lymphoid Tissue Reservoirs in Maximally Suppressed Individuals. *J. Acquir. Immune Defic. Syndr.* **2006**, *43* (1), 65-8.

- (118) Chun, T. W.; Nickle, D. C.; Justement, J. S., *et al.*, Persistence of HIV in Gut-Associated Lymphoid Tissue Despite Long-Term Antiretroviral Therapy. *J. Infect. Dis.* **2008**, *197* (5), 714-20.
- (119) Mehla, R.; Bivalkar-Mehla, S.; Zhang, R., *et al.*, Bryostatins Modulates Latent HIV-1 Infection Via Pkc and Ampk Signaling but Inhibits Acute Infection in a Receptor Independent Manner. *PLoS One* **2010**, *5* (6), e11160.
- (120) Wightman, F.; Ellenberg, P.; Churchill, M., *et al.*, Hdac Inhibitors in HIV. *Immunol. Cell Biol.* **2012**, *90* (1), 47-54.
- (121) Sagot-Lerolle, N.; Pallier, C., Prolonged Valproic Acid Treatment Does Not Reduce the Size of Latent HIV Reservoir (Vol 22, Pg 1125, 2008). *AIDS* **2008**, *22* (13), li-li.
- (122) Siliciano, J. D.; Lai, J.; Callender, M., *et al.*, Stability of the Latent Reservoir for HIV-1 in Patients Receiving Valproic Acid. *J. Infect. Dis.* **2007**, *195* (6), 833-6.
- (123) Shan, L.; Deng, K.; Shroff, N. S., *et al.*, Stimulation of HIV-1-Specific Cytolytic T Lymphocytes Facilitates Elimination of Latent Viral Reservoir after Virus Reactivation. *Immunity* **2012**, *36* (3), 491-501.
- (124) Deng, K.; Perteua, M.; Rongvaux, A., *et al.*, Broad Ctl Response Is Required to Clear Latent HIV-1 Due to Dominance of Escape Mutations. *Nature* **2015**, *517* (7534), 381-385.
- (125) McIlroy, D., Do HIV-Specific Ctl Continue to Have an Antiviral Function During Antiretroviral Therapy? If Not, Why Not, and What Can Be Done About It? *Front. Immunol.* **2013**, *4*, 52.
- (126) Foster, J. L.; Garcia, J. V., HIV-1 Nef: At the Crossroads. *Retrovirology* **2008**, *5*, 84.
- (127) Hanna, Z.; Priceputu, E.; Chrobak, P., *et al.*, Selective Expression of Human Immunodeficiency Virus Nef in Specific Immune Cell Populations of Transgenic Mice Is Associated with Distinct Aids-Like Phenotypes. *J. Virol.* **2009**, *83* (19), 9743-58.
- (128) Gorry, P. R.; McPhee, D. A.; Verity, E., *et al.*, Pathogenicity and Immunogenicity of Attenuated, Nef-Deleted HIV-1 Strains in Vivo. *Retrovirology* **2007**, *4*, 66.
- (129) Kirchhoff, F.; Greenough, T. C.; Brettler, D. B., *et al.*, Brief Report: Absence of Intact Nef Sequences in a Long-Term Survivor with Nonprogressive HIV-1 Infection. *N. Engl. J. Med.* **1995**, *332* (4), 228-32.

- (130) Gulizia, R. J.; Collman, R. G.; Levy, J. A., *et al.*, Deletion of Nef Slows but Does Not Prevent Cd4-Positive T-Cell Depletion in Human Immunodeficiency Virus Type 1-Infected Human-Pbl-Scid Mice. *J. Virol.* **1997**, *71* (5), 4161-4164.
- (131) Mwimanzi, P.; Markle, T. J.; Martin, E., *et al.*, Attenuation of Multiple Nef Functions in HIV-1 Elite Controllers. *Retrovirology* **2013**, *10*, 1.
- (132) Landi, A.; Iannucci, V.; Van Nuffel, A., *et al.*, One Protein to Rule Them All: Modulation of Cell Surface Receptors and Molecules by HIV Nef. *Curr. HIV Res.* **2011**, *9* (7), 496-504.
- (133) Das, S. R.; Jameel, S., Biology of the HIV Nef Protein. *Indian J. Med. Res.* **2005**, *121* (4), 315-32.
- (134) Quaranta, M. G.; Mattioli, B.; Giordani, L., *et al.*, Immunoregulatory Effects of HIV-1 Nef Protein. *Biofactors* **2009**, *35* (2), 169-74.
- (135) Tokarev, A.; Guatelli, J., Misdirection of Membrane Trafficking by HIV-1 Vpu and Nef: Keys to Viral Virulence and Persistence. *Cell Logist* **2011**, *1* (3), 90-102.
- (136) Shugars, D. C.; Smith, M. S.; Glueck, D. H., *et al.*, Analysis of Human Immunodeficiency Virus Type 1 Nef Gene Sequences Present in Vivo. *J. Virol.* **1993**, *67* (8), 4639-50.
- (137) Fausther-Bovendo, H.; Sol-Foulon, N.; Candotti, D., *et al.*, HIV Escape from Natural Killer Cytotoxicity: Nef Inhibits Nkp441 Expression on Cd4+ T Cells. *AIDS* **2009**, *23* (9), 1077-87.
- (138) Friesner, R. A.; Murphy, R. B.; Repasky, M. P., *et al.*, Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177-6196.
- (139) Halgren, T. A.; Murphy, R. B.; Friesner, R. A., *et al.*, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750-1759.
- (140) Friesner, R. A.; Banks, J. L.; Murphy, R. B., *et al.*, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739-1749.
- (141) Wonderlich, E. R.; Williams, M.; Collins, K. L., The Tyrosine Binding Pocket in the Adaptor Protein 1 (Ap-1) Mu1 Subunit Is Necessary for Nef to Recruit Ap-1 to the Major Histocompatibility Complex Class I Cytoplasmic Tail. *J. Biol. Chem.* **2008**, *283* (6), 3011-22.

- (142) Cohen, G. B.; Gandhi, R. T.; Davis, D. M., *et al.*, The Selective Downregulation of Class I Major Histocompatibility Complex Proteins by HIV-1 Protects HIV-Infected Cells from Nk Cells. *Immunity* **1999**, *10* (6), 661-71.
- (143) Jia, X. F.; Singh, R.; Homann, S., *et al.*, Structural Basis of Evasion of Cellular Adaptive Immunity by HIV-1 Nef. *Nat. Struct. Mol. Biol.* **2012**, *19* (7), 701-706.
- (144) Cummings, C. G.; Hamilton, A. D., Expedient Route to Functionalized and Water Soluble 5-6-5 Imidazole-Phenyl-Thiazole Based α -Helix Mimetics. *Tetrahedron* **2013**, *69* (5), 1663-1668.
- (145) Gowrisankar, S.; Sergeev, A. G.; Anbarasan, P., *et al.*, A General and Efficient Catalyst for Palladium-Catalyzed C–O Coupling Reactions of Aryl Halides with Primary Alcohols. *J. Am. Chem. Soc.* **2010**, *132* (33), 11592-11598.